HAL open science

Contribution à l'amélioration de l'expertise en situation de crise par l'utilisation de l'informatique distribuée : application aux crues à cinétique rapide

Vincent Thiérion

To cite this version:

Vincent Thiérion. Contribution à l'amélioration de l'expertise en situation de crise par l'utilisation de l'informatique distribuée : application aux crues à cinétique rapide. Gestion et management. École Nationale Supérieure des Mines de Paris, 2010. Français. NNT : 2010ENMP0019 . pastel-00566227

HAL Id: pastel-00566227

https://pastel.hal.science/pastel-00566227

Submitted on 15 Feb 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ParisTech
INSTITUT DES SCIENCES ET TECHNOLOGIES
PARIS INSTITUTE OF TECHNOLOGY

MINES ParisTech

École doctorale n° 432 : Sciences des Métiers de l'Ingénieur

Doctorat ParisTech

T H È S E

pour obtenir le grade de docteur délivré par

l'École nationale supérieure des mines de Paris

Spécialité " Sciences et génie des activités à risques "

présentée et soutenue publiquement par

Vincent Thierion

le 09 juillet 2010

Contribution à l'amélioration de l'expertise en situation de crise par l'utilisation de l'informatique distribuée

Application aux crues à cinétique rapide

Directrice de thèse : **Sophie Sauvagnargues-Lesage**
Co-encadrement de la thèse : **Pierre-Alain Ayral**

Jury

M. Ernesto BONOMI, Professeur, CRS4, Parco Scientifico e Tecnologico della Sardegna	Rapporteur
M. Freddy VINET, Professeur, GESTER, Université Montpellier 3	Rapporteur
M. Franck GUARNIERI, Professeur, CRC, École des Mines PARISTECH	Examinateur
M. Didier GRAILLOT, Directeur de recherche, SITE, École des Mines de Saint-Etienne	Examinateur
M. Olivier PAYRASTRE, Docteur, LCPC, MEEDDM et MESR	Examinateur
Mme Monique PETITDIDIER, Docteur d'État, CETP, Institut Pierre Simon Laplace	Examinateur
Mme Sophie SAUVAGNARGUES-LESAGE, Maître-assistant, LGEI, École des Mines d'Alès	Directrice
M. Pierre-Alain AYRAL, Maître-assistant, LGEI, École des Mines d'Alès	Co-encadrant

MINES ParisTech
Centre de Recherche sur les Risques et les Crises
Rue Claude Daunesse, B.P. 207
06904 Sophia-Antipolis Cedex

T H E S E

ParisTech
INSTITUT DES SCIENCES ET TECHNOLOGIES
PARIS INSTITUTE OF TECHNOLOGY

MINES ParisTech

École doctorale n° 432 : Sciences des Métiers de l'Ingénieur

Doctorat ParisTech

T H È S E

pour obtenir le grade de docteur délivré par

l'École nationale supérieure des mines de Paris

Spécialité " Sciences et génie des activités à risques "

présentée et soutenue publiquement par

Vincent Thierion

le 09 juillet 2010

Contribution à l'amélioration de l'expertise en situation de crise par l'utilisation de l'informatique distribuée

Application aux crues à cinétique rapide

Directrice de thèse : **Sophie Sauvagnargues-Lesage**
Co-encadrement de la thèse : **Pierre-Alain Ayral**

Jury

M. Ernesto BONOMI, Professeur, CRS4, Parco Scientifico e Tecnologico della Sardegna	Rapporteur
M. Freddy VINET, Professeur, GESTER, Université Montpellier 3	Rapporteur
M. Franck GUARNIERI, Professeur, CRC, École des Mines PARISTECH	Examinateur
M. Didier GRAILLOT, Directeur de recherche, SITE, École des Mines de Saint-Etienne	Examinateur
M. Olivier PAYRASTRE, Docteur, LCPC, MEEDDM et MESR	Examinateur
Mme Monique PETITDIDIER, Docteur d'État, CETP, Institut Pierre Simon Laplace	Examinateur
Mme Sophie SAUVAGNARGUES-LESAGE, Maître-assistant, LGEI, École des Mines d'Alès	Directrice
M. Pierre-Alain AYRAL, Maître-assistant, LGEI, École des Mines d'Alès	Co-encadrant

MINES ParisTech
Centre de Recherche sur les Risques et les Crises
Rue Claude Daunesse, B.P. 207
06904 Sophia-Antipolis Cedex

T H E S E

REMERCIEMENTS

Je tiens à remercier en premier lieu Ernesto Bonomi et Freddy Vinet qui ont accepté d'être les rapporteurs de cette thèse. En analysant de manière critique ce mémoire, ils ont permis d'alimenter et de poursuivre la réflexion nécessaire à la validation scientifique de ces travaux. Dans ce sens, il convient de remercier Didier GRAILLOT, président du jury, dont l'implication et le dynamisme ont permis la pleine réussite de ma soutenance de thèse.

« Remercier » représente une part infime du travail de thèse, mais pourtant essentielle tant le « réseau » apparait indispensable à toute apprentissage de la recherche. L'instigateur … Jean-François Galtié, du laboratoire GEODE à l'université Toulouse Le Mirail, qui a été le premier à croire en mes motivations et m'aiguiller vers l'indispensable « financement ». La directrice … Sophie Sauvagnargues-Lesage qui, par son esprit de synthèse et sa décontraction, m'a toujours rassuré dans l'aventure sinueuse de l'élève-chercheur. L'encadrant … Pierre-Alain Ayral … une histoire commune qui a forgé un sentiment fort et permis des discussions franches et constructives. À trois, l'épreuve était moins incertaine !

Débuter une thèse multidisciplinaire abordant des thèmes aussi différents que « l'eau qui coule » et « le bit qui transite » exige de développer un certain nombre de compétences. L'échange social m'est apparu essentiel dans cette tâche. Ainsi, Jean-François Musso m'a ouvert les portes du monde de la technologie grille, je l'en remercie cordialement. J'exprime de plus ma gratitude à David Weissenbach et Monique Petitdidier dont l'expérience respective m'a permis de comprendre la dimension technique, politique et économique de cette technologie innovante. A ce titre, les nombreuses discussions avec Monique Petitdidier m'ont permis de réaliser cette recherche avec l'assurance de la légitimité et de la pertinence. Qui ne connait pas Monique Petitdidier dans le monde EGEE ? Cette personne par qui tout le monde se rencontre.

Les orientations scientifiques d'un travail de recherche sont alimentées par l'expérience acquise, la « confrontation avec ses pairs » et la prospection intellectuelle. Les heures passées en compagnie des partenaires du projet européen CYCLOPS sont représentatives des efforts que peuvent représenter une thèse, même si ces rencontres se déroulaient entre le Portugal, l'Italie, la Grèce. Dans ce sens, je suis extrêmement reconnaissant envers l'équipe du laboratoire italien ESSI-Lab composée de Stefano Nativi, Paolo Mazzetti, Lorenzo Bigagli, Enrico Boldrini, Fabrizio Vitale, Mattia Santoro et Valerio Angelini. En liant amitié et professionnalisme, nous avons progressé avec Valerio dans l'univers de l'anglophonie et échangé nos savoirs respectifs afin de mener à bien l'association entre science des risques et informatique distribuée. Je tiens à remercier en suivant Carlos Morais-Pires, coordinateur du projet Cyclops auprès de l'Union Européenne.

Définitivement orientée vers l'opérationnel, cette thèse, reflète dans un sens les projets de recherche des Mines. Ici, en l'occurrence, il s'agissait d'appréhender au mieux les forces en présence et les besoins du service de prévision des crues Grand Delta (… du Rhône). L'équipe actuellement en place dans ce service mérite toute ma gratitude. Olivier Payrastre a su me présenter scrupuleusement l'environnement qu'il connait bien, celui de l'hydrologie opérationnelle… de la pluie à l'inondation. Autour de lui, je remercie Fabrice Mannessiez, Boris Aymes, Yann Laborda et François Bressand qui chacun à leur manière ont su partager leur expérience respective. Je terminerais par remercier Arthur Marchandise, Dominique Bielli et Christophe Bouvier.

Ricardo Marques, de l'univeristé Do Minho au Portugal, Cyrille Bronner, du LGEI, Renaud Burck, Simon Galbode, Serge Lhomme, Olivier Michelon et Guillaume Olive, de l'ENSG, Laurent Jégou et Jacob Geisel, de

l'université Toulouse Le Mirail, ont participé de prêt ou de loin à cette recherche, mais tous avec la même motivation. Je ne peux que les remercier pour l'intérêt qu'ils ont porté à cette recherche.

Il me reste à remercier deux groupes de personne qui ont gravité autour de moi pendant ces 3 années, mes collègues de travail et mes « collègues » de vie.

Pour le premier, ils se reconnaîtront quand je parle de « l'équipe risque » et ses séminaires, des « partenaires de foot » et des « anciens » de la cafet'. Je remercie d'autre part Miguel Lopez-Ferber, directeur du laboratoire LGEI, et Gilles Dusserre d'avoir accepté de m'accueillir au sein de leur laboratoire. Et la grande famille des thésards, Boris, Virginie, Marie, Léa, Clémence, Nicolas, Serge, Aurélia, Carole, Cyrille, Marc, Florian, Julien, Vincent, Line, …

Et puis … il reste … les amis ! Par qui commençait … par qui finir ? Je ne sais pas … je vais essayer de n'oublier personne ! Je commencerais par mes collocs de « Chez Felix » à Anduze, PE, Nico, Léa, Fred, Vincent, Greg et Nideux… entre thésards et noctambules, on se comprend ; de Méjannes, Julie, Poussinet et Fred ; de Saint Privat de Vieux, Vincent, Pia, Gaëtane et Damien. Et puis n'oublions pas les racines toulousaines et pyrénéennes, Benou, Jordi, Peio, Matteous, Rockicki, Jacky, Anne, Lili, Julie, Flonathan, Guigui, Chloé, Yomeuh, Thib', Fred, Sandrine, et les autres ! Enfin… Merci aux grimpeurs ardéchois !

Denis, mon compagnon de recherche, je te remercie pour tout !

Chaque personne se construit au travers de son entourage proche, de sa famille. Mes parents, Alain et Danielle, et ma sœur, Marie, m'ont donné toute la force et les armes pour accomplir la recherche présentée ici. Il me semble que mon plus grand « Merci » doit leur revenir !

Il n'est pas conseillé pour autant d'oublier Lolo, pour son grand soutien dans les moments difficiles. Je me demande encore comment elle a fait pour me supporter et je ne la remercierais jamais assez d'avoir surmonté avec moi cette étape ô combien périlleuse.

À Papy et Jul' …

SOMMAIRE

Sommaire 1

Introduction Générale 5

Partie 1 : Cadre institutionnel et État de l'art 9

Chapitre 1 : Les crues à cinétique rapide : phénomène et gestion opérationnelle 11

I - 1. Les crues à cinétique rapide : un risque majeur 13

I - 1.1. La gestion du risque « inondation » 13

I - 1.2. Dimension hydrométéorologique des crues éclair 19

I - 1.3. La modélisation des phénomènes de crues à cinétique rapide 26

I - 1.4. Synthèse des crues à cinétique rapide 33

I - 2. Cadre institutionnel et opérationnel de la gestion de crise « inondation » 34

I - 2.1. La notion de vigilance 34

I - 2.2. La vigilance météorologique : l'instigateur de l'alerte 34

I - 2.3. La vigilance hydrologique : une émergence justifiée 37

Chapitre 2 : L'informatique distribuée : une composante nouvelle des sociétés modernes 45

II - 1. Propos introductifs 47

II - 1.1. les technologies de l'information et de la communication (TIC) 47

II - 1.2. Collaboration et organisation virtuelle 50

II - 1.3. Interopérabilité 50

II - 2. L'informatique distribuée et la technologie grille 52

II - 2.1. Généralités sur l'informatique distribuée 52

II - 2.2. Les grands modèles de communication dans un système distribué 55

II - 2.3. La technologie grille : définitions et fonctionnement 57

II - 2.4. Synthèse sur la technologie grille 72

II - 3. La cartographie en ligne entre géomatique et services web 73

II - 3.1. Gestion des données géoréférencées : SIG et géomatique 73

II - 3.2. Les SIG et les réseaux informatiques 76

Partie 2 : Expertise hydrométéorologique en situation de crise : existant et propositions d'amélioration 87

Chapitre 3 : L'expertise hydrométéorologique des crues à cinétique rapide 89

III - 1. L'existant technologique du Service de Prévision des Crues Grand Delta 91

III - 1.1. Les méthodes RM-ODP et UML dédiées à la spécification des systèmes d'information 91

III - 1.2. Le système d'information du SPC-GD : rôles, compétences et capacités 93

III - 2. Retours d'expérience : Besoins opérationnels, scientifiques et technologiques 107

III - 2.1. Les crues à cinématique rapide : les cas « Aude 1999 » et « Gard 2002 » 107

III - 2.2. De la Sécurité Civile à l'expertise hydrométéorologique : un perfectionnement perpétuel 115

Chapitre 4 : La technologie grille et la géomatique : vecteurs d'innovation 127

IV - 1. Les contraintes technologiques de la prévision hydrologique 129

IV - 1.1. De la prévision hydrologique… 129

IV - 1.2. … aux contraintes technologiques inhérentes 130

IV - 2. La grille EGEE : un support informatique performant et adaptable 132

IV - 2.1. La technologie grille : des caractéristiques technologiques et des champs d'application favorables à une adoption opérationnelle 132

IV - 2.2. La technologie grille et la science de l'information géographique : une synergie innovante 138

IV - 3. G-ALHTAÏR : un système spatial d'aide à la décision hydrologique 145

IV - 3.1. L'hypothèse de recherche centrale : le système spatial d'aide à la décision 145

IV - 3.2. Caractérisation globale de G-ALHTAÏR dédié à l'anticipation des crues à cinétique rapide 147

IV - 3.3. Spécification détaillée du système spatial d'aide à la décision au sein du SPC-GD 147

Partie 3 : Méthodologie expérimentale 155

Chapitre 5 : Grillification de l'application ALHTAÏR 157

V - 1. L'intergiciel gLite : un accès unique aux ressources d'EGEE 159

V - 1.1. Les services d'accès 161

V - 1.2. les services de sécurité 161

V - 1.3. les services d'information et de surveillance des ressources 162

V - 1.4. les services de gestion de données 163

V - 1.5. Les services de gestion des tâches 163

V - 2. Méthodologie fondamentale 167

V - 2.1. Principe de fonctionnement du scénario de prévision 167

V - 2.2. Stratégie de portage 169

Chapitre 6 : Interfaçage entre la grille EGEE et le SPC-GD 177

VI - 1. Notes introductives : la problématique d'interfaçage 179

VI - 2. G-ALHTAÏR : un outil garant de l'interopérabilité et de l'orchestration des traitements 182

VI - 2.1. Point de vue « Entreprise » : l'interface de cartographie en ligne garante de l'opérationnalité du SPC-GD 182

VI - 2.2. Points de vue « Information et Traitement » : les services web de l'OGC 185

VI - 2.3. Point de vue « Ingénierie » et « Technologie » de G-ALHTAÏR 199

VI - 3. RRM-Grid : un outil garant de l'ordonnancement et d'une qualité de service 206

VI - 3.1. Propos introductifs : limites de performance de l'intergiciel gLite et contraintes de qualité de service 206

VI - 3.2. Ordonnanceur et qualité de service 207

VI - 3.3. RRM-Grid : une couche applicative garante d'une efficacité opérationnelle 209

Partie 4 : Résultats et Perspectives : vers l'adoption potentielle de la technologie grille 219

Chapitre 7 : Expérimentations de l'ordonnanceur RRM-Grid 221

VII - 1. Plan d'expérience 223

VII - 1.1. Ressources de grille : la VO ESR 223

VII - 1.2. Méthodes de traitement expérimentées 224

VII - 1.4. Critères d'analyse 226

VII - 1.3. Étude de cas en situation opérationnelle 228

VII - 2. Résultats expérimentaux de RRM-Grid 232

VII - 2.1. Performances technologiques 232

VII - 2.2. Étude de cas de l'évènement du 22 octobre 2008 240

Chapitre 8 : Perspectives opérationnelles et scientifiques de G-ALHTAÏR 255

VIII - 1. G-ALHTAÏR : un système spatial d'aide à la décision en devenir 257

VIII - 1.1. L'optimisation des traitements de grille 257

VIII - 1.2. L'interopérabilité : entre sophistication et opérationnalité 261

VIII - 2. Enrichissement de la dimension hydrométéorologique 264

VIII - 2.1. L'approche multi-modèle en temps réel 264

VIII - 2.2. Vers l'optimisation du modèle 268

VIII - 2.3. Gestion des systèmes de mesure et des données collectées 269

Conclusion générale 273

Bibliographie 281

Liste des figures 305

Liste des tableaux 311

Liste des sigles 313

Table des matières 319

Annexe 329

INTRODUCTION GENERALE

L'expertise en situation de crise est une composante essentielle des missions de la Sécurité Civile. Elle représente généralement l'instigateur initial de la gestion de crise, en informant les services institutionnels et opérationnels de la survenance d'un évènement dangereux. La recherche présentée dans ce mémoire s'intéresse au risque naturel le plus fréquent à la surface de la Terre et l'un des plus dévastateurs : les inondations. Les occurrences d'inondation sont nombreuses et de formes différentes. Parmi ces différentes manifestations, un phénomène plus spécifique provoque régulièrement des dommages importants et un dérèglement des procédures opérationnelles de gestion de crise : la crue à cinétique rapide ou crue éclair.

Les crues à cinétique rapide

En France, la surveillance et l'expertise hydrologique des cours d'eau principaux sont assurées par les services de prévision des crues (SPC) au niveau régional[1] et par le service central d'hydrométéorologie et d'appui à la prévision des inondations (SCHAPI) au niveau national. Dans le cadre des crues à cinétique rapide, le service de prévision des crues « Grand Delta » (SPC-GD), de part sa localisation géographique à proximité du massif des Cévennes, représente un service opérationnel compétent, dans la mesure où une grande partie du territoire placé sous sa surveillance est soumis à ce type de phénomène. Il est, de part l'insuffisance des connaissances sur les processus hydrologiques relatifs aux crues à cinétique rapide, régulièrement confronté à des problématiques complexes d'expertise en situation de crise.

La dernière occurrence catastrophique provoquée par ce type de crues correspond au 8 et 9 septembre 2002 dans le département du Gard. Pendant près de 24h, le département et ses bassins versants amont sont touchés par des pluies diluviennes provoquant une montée soudaine des cours d'eau. La rapidité de cet événement et son étendu géographique ont provoqué des dommages importants et de fortes défaillances dans la chaîne d'alerte hydrométéorologique (HUET et *al*., 2003a).

Après cette situation de crise de grande ampleur, de nombreux retours d'expérience et recherches en hydrométéorologie se sont focalisés sur l'analyse du phénomène et l'évaluation des causes possibles des défaillances organisationnelles. Dès l'autonome 2002, une réforme des services d'annonce de crues (SAC), alors en place, a été annoncée et matérialisée par loi du 30 juillet 2003 relatif à la création des services de prévision des crues. Cette réforme avait ainsi pour objectifs principaux, une meilleure prévision de ces phénomènes potentiellement dangereux et une amélioration nécessaire des circuits de communication entre les partenaires de la gestion de crise « inondation » (VINET, 2007).

[1] Régional s'entends ici plutôt par grands bassins hydrologiques

Ainsi, depuis cet évènement, de nombreuses initiatives ont vu le jour pour répondre à ces nouveaux objectifs opérationnels. L'amélioration de la gestion de crise « inondation » nécessite l'implication de nombreux champs disciplinaires, tant technologiques que scientifiques. Il n'existe donc *a priori* pas une réponse univoque qui permettrait de mener à bien la protection des citoyens et de leurs infrastructures mais une complémentarité dans les approches menées. L'un des axes d'amélioration passe par une culture quotidienne de veille technologique permettant une évolution des outils et des méthodes existants dédiés à l'expertise hydrométéorologique. La présente recherche s'inscrit donc clairement dans ce courant de perfectionnement de la nouvelle mission des services de prévision des crues, en empruntant plus exactement la voie technologique.

L'informatique distribuée

Depuis le milieu des années 90, l'informatique connait une révolution dans la manière de traiter et de partager les données. A la vision originelle du poste informatique centralisant l'ensemble des logiciels et des données, est venue se greffer une conception nouvelle basée sur la distribution des ressources informatiques : l'informatique distribuée.

De nombreux champs de recherche se sont développés autour de ce nouveau concept. Globalement, il est possible de distinguer plusieurs caractéristiques principales :

- le calcul parallèle cherchant à optimiser et à accélérer les traitements informatiques entre des processeurs distincts,
- le calcul intensif qui, en s'appuyant sur un grand nombre de ressources informatiques, permet de prendre en charge un nombre important de données,
- la dimension collaborative offerte par la connexion de ressources informatiques géographiquement distantes et dont le porte-drapeau est certainement le réseau Internet.

L'ensemble de ces caractéristiques et les progrès relatifs au matériel (mémoire, processeurs, bande passante, etc.) ont progressivement amené Ian Foster du laboratoire national américain Argonne et Carl Kesselman de l'université de Californie du Sud (FOSTER et KESSELMAN, 2004) à concevoir une solution technologique capable d'intégrer et de virtualiser des ressources informatiques géographiquement distantes au sein d'un système d'information unique : la technologie grille.

À l'heure actuelle, et grâce aux efforts de recherche internationaux depuis plus de 10 ans, cette technologie tend à devenir une solution informatique générique de grande échelle offrant des capacités de calcul et de stockage importantes et un pouvoir de collaboration jamais atteints. Dans ce sens, elle ouvre de nouvelles opportunités dans le besoin croissant d'interdisciplinarité requis pour la résolution de problématiques complexes, telles que celles des sciences de la Terre.

En Europe, un effort similaire a été mené depuis le début des années 2000 et a donné naissance à l'architecture de grille EGEE (Enabling Grid for E-sciencE). Récemment passée en phase de production, cette architecture assure une offre de calcul et de stockage à la demande à plus de 140 institutions, dont la plus connue est le CERN (Organisation européenne pour la recherche nucléaire) et son

accélérateur de particules LHC (Large Hadron Collider). Sa structuration en sous-ensembles thématiques, appelés « organisation virtuelle » (VO), permet d'une part une gouvernance harmonisée et d'autre part l'intégration constante de nouveaux partenaires tant dans la fourniture de ressources informatiques que dans l'utilisation de ces dernières.

Ainsi, une telle situation permet d'entrevoir un recours à cette solution technologique, pour les problématiques relatives à l'expertise hydrométéorologique des crues à cinétique rapide, et plus largement, pour l'amélioration des systèmes opérants de la Sécurité Civile. À titre d'exemple, le projet européen CYCLOPS (Cyber Infrastructure for Civil protection Operative ProcedureS) tend à illustrer cette hypothèse, part le rapprochement des communautés « grille » et « Sécurité Civile » dans le contexte d'expertise des risques naturel (MAZZETTI, 2006).

La problématique générale

Ainsi, la problématique générale de cette recherche s'appuie sur les fondements théoriques de l'expertise en situation de crise et de la technologie grille. Elle cherche à évaluer les aptitudes technologiques de la grille, en tant que solution informatique innovante, à soutenir les activités opérationnelles relatives à l'expertise hydrométéorologique des crues à cinétique rapide en situation de crise.

Cette problématique a fait émerger deux objectifs sous-jacents :

- la nécessaire amélioration des capacités de prévision des crues à cinétique rapide par le recours à des ressources informatiques supplémentaires et à des méthodes propres à la géomatique,
- la consolidation des moyens de collaboration et de partage des services opérationnels au vu des défaillances des circuits de communication durant l'inondation de 2002.

Le rôle central et la longue expérience du service de prévision des crues « Grand Delta » dans la gestion de l'expertise hydrologique des crues à cinétique rapide en font un objet d'étude naturel. Ce choix est de plus renforcé par l'implication continue de ce service dans l'amélioration générale de l'hydrologie opérationnelle.

Le plan de thèse

Ce manuscrit est composé de 4 parties représentant la démarche globale développée dans cette recherche.

La partie 1 « Cadre institutionnel et états de l'art scientifiques » s'attache à définir et décrire le contexte général de la gestion des crues à cinétique rapide (chapitre 1) et à analyser les fondements théoriques des technologies sélectionnées dans cette recherche (chapitre 2). Plus précisément, le chapitre 1 « Les crues à cinétique rapide : phénomène et gestion opérationnelle » est dédié à la définition de l'aléa « crue à cinétique rapide » et au positionnement de son expertise dans le champ plus large de la gestion de crise. Dans le chapitre 2 « L'informatique distribuée : une composante nouvelle des sociétés modernes », l'informatique distribuée et ses concepts fondamentaux sont présentés dans l'objectif d'établir une vision précise des motivations scientifiques ayant amenés au développement de la tech-

nologie grille. En fin de chapitre, le concept de science de l'information géographique est décrit et corrélé à l'informatique distribuée.

La partie 2 « Technologie grille et géomatique pour l'expertise hydrométéorologique en situation de crise » correspond à la mise en problématique de cette recherche. Le chapitre 3 « l'expertise hydrométéorologique des crues à cinétique rapide » décrit l'existant technologique et organisationnel du service de prévision des crues « Grand Delta » et les besoins en matière d'expertise hydrologique. En parallèle, le chapitre 4 « La technologie grille et la géomatique : vecteurs d'innovation » développe une analyse précise des capacités informatiques de la grille et la prise en charge des problématiques liées aux sciences de la Terre et à la géomatique. En fin de partie, une proposition méthodologique est développée grâce à une analyse croisée entre besoins, contraintes et opportunités.

La partie 3 « Méthodologie expérimentale» décrit de manière précise les développements méthodologiques menés dans cette recherche. Le chapitre 5 « Grillification de l'application ALHTAÏR » présente la méthodologie mise en oeuvre pour exécuter un modèle hydrologique sur les ressources de l'architecture de grille EGEE. Le chapitre 6 « Interfaçage entre la grille EGEE et le SPC-GD » s'attache à décrire deux solutions pour accéder aux ressources de grille dans un contexte opérationnel de gestion de crise.

La partie 4 « Résultats et perspectives : vers l'adoption potentielle de la technologie grille » établit une évaluation des performances de la méthodologie proposée dans cette recherche, et par extension des capacités offertes par la technologie grille (chapitre 7 « Expérimentations de l'ordonnanceur RRM-Grid »). Enfin, le chapitre 8 « Perspectives opérationnelles et scientifiques de G-ALHTAÏR » s'appuie sur l'ensemble des besoins détaillés dans le chapitre 3 et des potentialités générales des technologies abordées dans cette recherche pour proposer plusieurs axes d'innovation à mener à la suite de cette thèse.

Finalement, cette thèse, en établissant des relations fortes entre la communauté opérationnelle de l'expertise hydrologique et certains membres français et européens de la communauté « grille » tend à consolider sa dimension opérationnelle.

PARTIE 1

Cadre institutionnel et états de l'art scientifiques

« … Dans la nuit, le vent est passé au sud, poussant vers les Cévennes un plafond de nuages lourds qui défilent devant la lune. Les arbres, les volets, les portes de la maison, tout a commencé à respirer sous l'influence du souffle tiède aux odeurs de mer et de garrigue. A l'intérieur des maisons, une agitation inquiète a pris possession des corps endormis. Abandonné par le sommeil, debout devant la fenêtre, je contemple la nuit noire… tout le vallon bruisse, et la somme des ces milliers de frottements de feuilles forment une rumeur ample et sourde. Poc ! Un impact unique mais puissant retentit sur le velux. Je lève les yeux. Une large auréole d'eau s'achève de s'étaler à l'endroit où est tombée la première goutte. Durant quelques secondes on pourrait croire qu'elle restera la seule. Mais une seconde arrive, et bientôt les autres suivent à une cadence qui accélère comme une locomotive à vapeur quittant le quai. Un crépitement continu envahit bientôt le silence de la pièce obscure. AU lever du jour, la pluie forme un rideau continu devant les fenêtres, transformant le versant opposé de la vallée en une simple ombre à peine perceptible. Pour un normand comme moi, la pluie peut prendre deux formes extrêmes : le fin crachin qui dure, ou la violente mais passagère pluie d'orage. L'épisode cévenol, c'est les deux à la fois. Une pluie lourde, et sui dure, dure, tant qu'on n'arrive pas à comprendre comment le ciel peut produire tant d'eau sans s'assécher à jamais. »

Rêve éveillé, Marc Lemonier

Chapitre 1

Les crues à cinétique rapide : phénomène et gestion opérationnelle

1.1. Les crues à cinétique rapide : un risque majeur

1.2. Cadre institutionnel et opérationnel de la gestion de crise « Inondation »

Ce chapitre définit les processus météorologiques et hydrologiques impliqués dans le phénomène de crue à cinétique rapide et la modélisation pluie-débit adaptée à la reconstitution de son débit. Sa faible occurrence et les dégâts qu'elle occasionne parfois la positionne en tant risque naturel majeur. Dans ce contexte, le cadre opérationnel qui permet son suivi et sa prévision en situation de crise est approfondi.

INTRODUCTION

L'élément central de cette recherche correspond à la « crue éclair ». Ce phénomène naturel a généralement lieu dans des bassins versants de petite taille, caractérisés par des pentes raides, et soumis à des précipitations orageuses de forte intensité et de durée relativement longue. La période la plus propice, selon différentes études, s'étale du mois de septembre au mois de novembre (Automne). De manière générale, les zones touchées par ces crues éclair, sont majoritairement caractérisées par des conditions climatiques de type méditerranéen, bien que certains auteurs les considèrent de la même manière que les crues torrentielles touchant des zones montagneuses. La définition précise de ce phénomène peut s'avérer variable suivant les critères choisis (hydrologiques, pluviométriques, comportementaux, etc.) pour le caractériser. Dans le cadre de cette étude, la crue éclair est considérée comme un phénomène naturel de faible extension géographique mais dont l'intensité, la rapidité de génération et la faible occurrence limite l'expertise hydrométéorologique des opérations de prévision de la gestion de crise. Ainsi, le terme crue à cinétique rapide est préféré à celui de crue éclair et est considéré ici comme un risque majeur menaçant les territoires anthropisés.

En France, la zone géographique concernée par ce type d'évènement se situe dans le quart sud-est du territoire métropolitain, et plus précisément dans les régions montagneuses proches du pourtour méditerranéen. Il s'agit globalement des contreforts orientaux du massif pyrénéen, et méridionaux du massif alpin et du Massif Central. La zone de référence de cette, de part la plus forte probabilité d'occurrence des crues éclair, correspond au territoire amont des Cévennes surveillé par le service de prévision « Grand Delta » ou SPC-GD.

L'objet de ce chapitre, est donc de préciser les processus météorologiques et hydrologiques qui caractérisent ce type de phénomène, ainsi que les moyens scientifiques et opérationnels mis en œuvre pour les étudier et les comprendre. Dans une seconde partie, étant donné les fortes incertitudes et la difficulté globale de les prévoir, les moyens institutionnels et opérationnels mis en place au niveau français pour organiser leur analyse et leur caractérisation en situation de crise sont présentés. Ce premier chapitre doit donc amener le lecteur à comprendre les enjeux actuels de la gestion opérationnelle des crues à cinétique rapide.

I - 1. Les crues à cinétique rapide : un risque majeur

Dans ce paragraphe, il s'agit de présenter les phénomènes hydrométéorologiques générateurs des crues éclair et les moyens informatiques utilisés pour les modéliser. Le terme de « crue à cinétique rapide » est préféré au terme « crue éclair », plus représentatif de la contrainte temporelle de cet aléa naturel.

I - 1.1. La gestion du risque « inondation »

Malgré ces particularités physiques, les crues à cinétique rapide peuvent être considérées dans le champ plus large du risque « inondation ». Ainsi, il semble nécessaire de définir précisément les termes fondamentaux relatifs à la science des risques, et en particulier, le positionnement de la prévision hydrométéorologique en situation de crise « inondation ».

I - 1.1.1. Fondements théoriques : aléa, vulnérabilité et risques majeurs

I - 1.1.1.1. La notion de risque

De manière générale, le risque résulte de la relation entre la société et son environnement. Comme l'aborde DAUTUN (2007), « les risques ont toujours été inhérents à nos sociétés » au point que DUPONT (2003) aborde le concept de « civilisation du risque ». Le concept de risque tel qu'on l'utilise de nos jours est né avec l'avènement socio-économique des sociétés industrielles. De part, les activités de plus en plus sophistiquées qu'elle génère, la société met potentiellement en danger sa population et ses infrastructures. De la même manière, l'urbanisation croissante et l'extension des villes confrontent de plus en plus les populations avec des éléments naturels potentiellement menaçant (feux de forêts, inondations, etc.).

Le risque peut être d'ordre naturel, technologique ou anthropique. Il existe de nombreuses définitions suivant le point de vue de l'observateur et les problématiques étudiées (AYRAL et GRIOT, 2001). Selon les Nations Unies, le risque correspond à l' « espérance mathématique de pertes en vies humaines, blessés, dommages aux biens et atteinte à l'activité économique au cours d'une période de référence et dans une région donnée, pour un aléa particulier ». Il s'agit d'une définition générale, qui permet cependant de relever plusieurs notions fondamentales du risque : l'espace, le temps et l'incertitude. Ce point de vue est confirmé par VEYRET (2004) qui considère le risque comme un « objet géographique » caractérisé par « sa dimension spatiale et ses aspects temporels ». Plus simplement, selon les Nations Unies, « Le risque est le produit de l'aléa par la vulnérabilité », alors que DAUPHINE (2001) propose de définir le risque plutôt comme une fonction de l'aléa et de la vulnérabilité, étant donné la subjectivité de la perception du risque par les sociétés :

$$Risque = F(aléa, vulnérabilité)$$

Ainsi, le risque naturel est le résultat d'un phénomène naturel, dont la gravité dépend de l'occurrence et de l'intensité d'un aléa, et de la vulnérabilité d'une cible (enjeux environnementaux, humains, et économiques) (Figure 1).

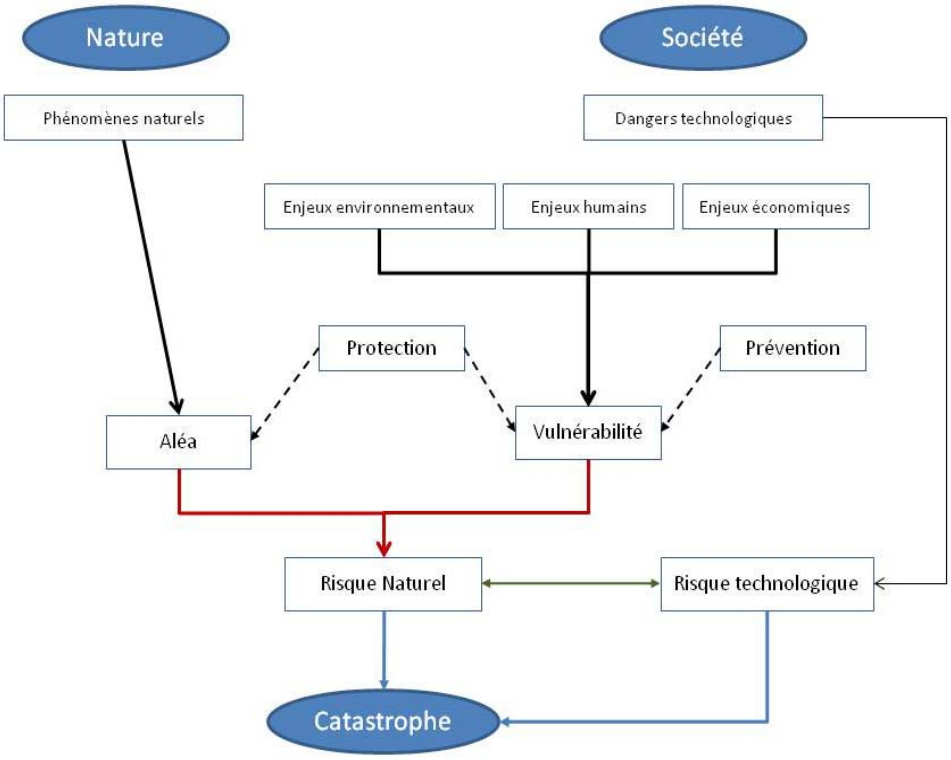


Figure 1 : Le risque naturel entre Nature et Société (tirée de AYRAL, 2001)

De plus, toujours d'après la Figure 1, la prise en compte et la réduction d'un risque naturel dépend des opérations de protection et de prévention développées par une société.

I - 1.1.1.2. L'aléa

Un aléa, principal d'objet d'étude de cette recherche, est caractérisé par une intensité, une fréquence, une extension spatiale, et une période temporelle d'occurrence (VEYRET, 2004). Il existe comme pour le risque un grand nombre de définitions. L'Institut des Risques Majeurs (IRMa) définit l'aléa comme une « possibilité de l'apparition d'un phénomène résultant de facteurs ou de processus qui échappent au moins en partie à l'homme » (AYRAL et GRIOT, 2001), ce qui correspond bien à l'aspect fortement imprévisible d'une crue à cinétique rapide ou crue éclair. L'évaluation des caractéristiques d'un aléa nécessite d'appréhender les processus physiques en jeu. Cette évaluation passe généralement par l'utilisation d'une instrumentation adaptée capable de collecter et transmettre les données relatives aux processus physiques en jeu. De manière générale, la caractérisation précise et stricte d'un aléa reste encore de nos jours un idéal à atteindre. Cette difficulté est d'autant plus exacerbée que la fréquence de l'aléa en jeu est faible (HAZIZA, 2007).

I - 1.1.1.3. La vulnérabilité

La vulnérabilité quant à elle est polysémique et peut avoir une dimension qualitative, quantitative ou semi-quantitative. Pour AYRAL (2001), « la vulnérabilité est l'étude des dommages potentiels sur les enjeux ». La notion d'enjeux est définie par le Ministère de l'Écologie, de l'Énergie, du Développement Durable et de la Mer (MEEDDM) comme les « personnes, biens, équipements, environnement menacés par l'aléa et susceptibles de subir des préjudices ou des dommages ». De manière générale, étudier la vulnérabilité consiste à évaluer la valeur économique des enjeux susceptibles d'être endomma-

gés par un aléa. VEYRET (2004) distingue une composante objective, la valeur socio-économique des territoires menacés, et une composante subjective, liée à la perception du danger. Finalement, la vulnérabilité permet d'intégrer une dimension sociale à l'étude des risques (HAZIZA, 2007).

I - 1.1.1.4. Le risque majeur

Si le terme « gravité » correspond à une valeur quantitative de dommages, elle augmente avec l'imprédictibilité et l'intensité de l'aléa, et l'importance des dommages (LEFEVRE et SCHNEIDER, 2003). Dans ce mémoire, l'aléa considéré est relatif aux crues à cinétique rapide qui sont par nature difficilement prévisible en temps et en lieu, et qui occasionnent régulièrement des dommages socio-économiques paralysant le fonctionnement des territoires concernés pendant quelques jours. Dans ce contexte, le terme de « risque majeur » est employé. D'un point de vue social, le risque majeur est rattaché à l'idée de *catastrophe*, souvent « révélateur d'une rupture socio-économique » (VEYRET, 2004)

Ainsi, cette recherche se focalise sur les risques naturels, et en particulier sur les inondations provoquées par des phénomènes de crues à cinétique rapide. Ce type d'aléa a de tout temps menacé les populations. La manière d'appréhender les risques naturels, suit globalement les conceptions du rapport Société – Nature (VEYRET, 2004) :

- Une séparation entre l'aléa naturel et la société, considérés indépendamment pour réduire le risque,
- Les éléments naturels sont fortement transformés par les sociétés, la nature est donc considérée comme anthropisée[2]. L'aléa devient donc un « révélateur de la vulnérabilité » (VEYRET, 2004).

Dans le massif des Cévennes, inclus dans la zone d'étude « Grand Delta » de cette recherche, des crues à cinétique rapide touchent régulièrement les populations. Les actions anthropiques passées et présentes jouent certainement un rôle non négligeable sur le déclenchement de ce phénomène naturel, cependant leur contribution dans la génération de ce type de crues n'est toujours pas avéré (MARTIN et *al*., 2008). Dans les zones à plus fort développement économique, touchées par le champ d'inondation, résultant de la crue amont, l'effet anthropique est certainement plus marqué. L'imperméabilisation des sols, résultant de l'urbanisation, et plus largement la modification du cycle de l'eau augmente significativement les volumes ruisselés et *a fortiori* la vulnérabilité des populations face à l'inondation (HAZIZA, 2007). Cette question de l'étude du risque naturel est complexe, la manière de l'aborder dépend fortement de la finalité des travaux engagés.

Dans le cadre de ces travaux, l'objectif sous-jacent est d'améliorer la prise en compte de l'aléa « crue éclair », par les services opérationnels en vue de diminuer la vulnérabilité des populations et de leurs biens soumises à cet aléa. La réduction de la vulnérabilité d'un territoire, et par extension les risques qui le concerne, passe par de nombreuses mesures tant structurelles qu'organisationnelles. VEYRET (2004) identifie trois actions principales d'amélioration :

[2] Terme concernant l'action de la société sur les espaces naturels

- La meilleure connaissance des processus physiques,
- La mise en place d'aménagements adaptés,
- Un usage des territoires intégrant le danger.

L'ensemble de ces actions peuvent être réalisées à différentes étapes de la gestion du risque, avant en phase de prévention, pendant (prévision et gestion de crise) ou après un évènement catastrophique (retours d'expérience). Dans le cadre de cette étude, une attention particulière est donnée à la composante « prévision » de la gestion de crise, de l'émergence de l'évènement naturel à risque au retour à la normale météorologique (Figure 2).

Figure 2 : Cycle du risque et positionnement scientifique de cette recherche

I - 1.1.2. La prévision hydrométéorologique en situation de crise

I - 1.1.2.1. La crise : définitions et caractéristiques fonctionnelles

L'intensité de l'aléa et la difficulté de le prévoir précisément sont clairement des facteurs prépondérants dans l'apparition d'une crise, cependant de nombreux autres facteurs peuvent créer des conditions propices à son établissement.

Ainsi, avant d'aborder la gestion de crise, à part entière, il convient de définir cette notion de crise. Une situation de crise trouve son origine dans l'émergence d'un évènement soudain de faible probabilité d'occurrence, assimilable à un risque majeur. Il est possible de caractériser une crise par des (DAUTUN, 2007 ; LAGADEC, 1991 *in* DAUTUN, 2004) :

- défaillances organisationnelles et techniques,
- capacités de réaction des organisations affectées,
- processus de décision faussés par de nombreuses incertitudes,
- temps d'action limités corrélés à une évolution rapide de la situation.

Il en résulte une situation où les services opérationnels et décisionnels ont, d'une part des difficultés à appréhender précisément l'aléa et, sont d'autre part potentiellement vulnérables à ce dernier.

Ainsi, la gestion de crise, avec la réflexion multi-institutionnelle et le développement de stratégies adaptées qu'elle engendre, concoure à réduire le nombre de « crises de grande ampleur » (DAUTUN, 2007). La gestion de crise, centrée ici sur la phase de réponse, regroupe ainsi un ensemble d'institutions, aux compétences souvent différentes, qui coopèrent et s'unissent pour endurer le phénomène menaçant et ses conséquences directes. La réussite de cet objectif impose au préalable le développement de cadres réglementaires permettant une formalisation des rôles et des missions de tous les partenaires impliqués dans la gestion de crise (phase de prévention). Les plans d'urgence et de secours sont le reflet de cette nécessité. Ils cherchent à définir précisément l'organisation des secours à adopter pour répondre aux dommages engendrés par une catastrophe. Par exemple, le Plan Communal de Sauvegarde (PCS) permet à l'échelle communale d'inventorier le ou les risques et de mettre en place des stratégies opérationnelles, des systèmes opérants et des politiques publiques pour les réduire. Aux échelles départementale, zonale et nationale, le plan ORSEC (Organisation de la Réponse de SEcurité Civile) répond aux mêmes objectifs par une organisation multi-institutionnelle compétente et effective pour gérer les secours, et plus largement la réponse opérationnelle, de manière unifiée et homogène (MIAT, 2006).

La gestion de crise permet, par les actions de prévision en situation de crise, d'une part de limiter la vulnérabilité de la société face à un évènement dommageable et d'autre part de tenir informé, tout au long de la phase de réponse, la population sur les risques la menaçant. Ce devoir de communication en situation de crise (LEFEVRE et SCHNEIDER, 2003), garant de la sécurité des citoyens et réduisant les risques de crise, passe généralement par les médias, tels que la télévision, les radios et Internet.

Finalement, ces situations catastrophiques, où la vulnérabilité « globale » de la société est temporairement accrue, ont un effet souvent dramatique sur les enjeux mais permettent potentiellement une amélioration de la gestion des catastrophes à venir, par l'établissement de retours d'expérience[3].

I - 1.1.2.2. La prévision hydrométéorologique comme élément fondamental de la gestion de crise « inondation »

La prévision cherche à spécifier voire anticiper les caractéristiques physiques de l'aléa, juste avant et tout le long de la gestion de crise. L'objectif principal de la prévision est d'informer en temps réel les services de la Sécurité Civile, en charge de la protection de la population, et la population elle-même, sur le développement potentiel d'un aléa. Dans le cas des inondations, cette information cruciale « fait appel à de multiples disciplines scientifiques », telles que l'hydrologie, la météorologie, la climatologie et l'hydraulique, (DAUPHINE, 2004).

Selon DAUPHINE (2004), il existe trois types de prévision :

- empirique, basée sur les faits historiques, elle ne permet pas de spécifier précisément les caractéristiques de l'aléa,
- déterministe, basée sur les causes réelles de l'aléa et qui doit permettre de spécifier en date et lieu l'occurrence de ce dernier grâce à la connaissance des processus physiques en jeu,

[3] Le retour d'expérience (REX) est « l'analyse détaillée d'un événement passé, dans le but d'améliorer la connaissance et de renforcer si nécessaire, la chaîne de sécurité » (HUET, 2005).

- probabiliste, basée sur la probabilité d'occurrence d'un phénomène de type décennale, centennale ou milléniale où une période de retour[4] est prévue ou prévisible pour l'évènement à venir. Dans le cas des crues, si l'intensité d'un évènement hydrologique peut être appréhendée au travers de cette approche, ses caractéristiques spatio-temporelles sont souvent difficiles à spécifier.

A ces trois types de prévision vient se rajouter l'approche par scénarios qui peut être rattachée aux autres approches et qui cherche à définir différentes situations à venir en se basant sur des conditions initiales et des cheminements différents (DAUPHINE, 2004). Ainsi, un jeu de scénarios peut être développé grâce à des prévisions basées sur des équations déterministes sur lesquelles les paramètres et les données d'entrée varient d'un scénario à l'autre.

L'ensemble de ces méthodes de prévision sont souvent basées sur la mise en place de bases de connaissances ainsi que sur le développement de systèmes d'information prenant en charge des modèles. Cependant, l'anticipation d'un aléa, qui se distingue de la prévision par sa plus grande précision spatio-temporelle, appelle à une expertise humaine capable de détecter les « signaux avertisseurs » en « phase d'incubation » de l'aléa (Dautun, 2007). Ainsi, Veyret (2004) distingue trois stades d'appréhension de l'aléa en phase de prévision :

- la vigilance,
- la pré-alerte,
- l'alerte ou l'alarme.

Pour appuyer cette expertise tout au long du processus de prévision, des systèmes de surveillance en temps-réel sont utilisés. Ces moyens d'observation sont nombreux et essentiels pour la caractérisation de l'aléa (images radar et satellitales, réseaux de mesures et de télémesures, etc.). Si les prévisions manquent généralement de précisions quantitatives sur l'aléa menaçant, elles indiquent quasi-systématiquement une situation dangereuse à venir, notamment grâce à ces systèmes d'observation. Cependant, ces faiblesses dans la prévision peuvent conditionner l'irruption d'une situation de crise. Dans le cas des crues éclair, les services institutionnels doivent avoir une véritable capacité d'anticipation, en améliorant d'une part la connaissance des processus hydrologiques en jeu et d'autre part en offrant une capacité de gestion de l'information hydrométéorologique en temps-réel, permettant d'endosser les délais courts spécifiques aux crues à cinétique rapide. Dans ce sens, il s'agit de mieux spécifier les caractéristiques hydrométéorologiques de ces crues.

[4] Moyenne à long terme du temps ou du nombre d'années séparant un événement de grandeur donnée d'un second événement d'une grandeur égale ou supérieure (Glossaire International d'Hydrologie : http://hydrologie.org/glu/aglo.htm)

I - 1.2. Dimension hydrométéorologique des crues éclair

I - 1.2.1. Fondements théoriques

Il existe plusieurs définitions du terme « inondation ». Par exemple, le dictionnaire français d'hydrologie[5] s'accorde sur une définition assez générale où l'inondation correspond à « la submersion temporaire, naturelle ou artificielle, d'un espace terrestre » par les eaux. D'un point de vue hydrologique, une inondation désigne « un recouvrement d'eau qui déborde du lit mineur » (LEFEVRE et SCHNEIDER, 2003). SALOMON (1997) insiste sur le caractère exceptionnel de ce phénomène et le relie à l'augmentation du débit d'un cours d'eau, généralement nommé « crue », lorsque la crue est assez puissante elle est ainsi susceptible d'entraîner le débordement du cours d'eau provoquant ainsi une inondation dans le lit majeur.

La caractérisation d'une crue se fait au moyen d'une hauteur d'eau (m), d'un débit (m^3/s), voire d'une vitesse (m/s) (LEFEVRE et SCHNEIDER, 2003). Le volume d'eau observable en un point donné (exutoire) résulte de l'accumulation directe ou indirecte des eaux provenant de zones géographiques amont. Il est communément admis que le bassin versant représente « l'unité fonctionnelle fondamentale » pour l'analyse des phénomènes hydrologiques (AMBROISE, 1991). Selon le dictionnaire français d'hydrologie, un bassin versant est une région délimitée, drainée par un cours d'eau et ses tributaires, dont elle constitue l'aire d'alimentation. Ainsi, la compréhension et l'analyse d'une crue à un exutoire donné ainsi que l'inondation potentielle qu'elle provoque, est basée sur les phénomènes géographiques (géologie, sol, occupation du sol, artefacts, etc.) agissant sur le bassin versant correspondant à cet exutoire.

D'une région à l'autre du globe, il existe de nombreux types d'inondations régies par des situations physiques et anthropiques différentes. La typologie exhaustive proposée par LEFEVRE et SCHNEIDER (2003) peut être revisitée en insistant sur la dimension temporelle de la prévision (anticipation) de ces évènements :

- Anticipation longue :
    - les inondations littorales où les eaux des cours d'eau en crue ne sont pas évacuées du fait d'un niveau de mer élevé (marée),
    - les inondations de plaine associées à des bassins versants de très grande taille et à des pluies peu intenses mais de longue durée ;
- Anticipation courte :
    - les ruptures de digues, de barrages ou d'embâcles,
    - les débâcles glaciaires résultant de la rupture d'embâcles glaciaires,
    - les inondations urbaines dont l'infiltration de l'eau est fortement limitée par l'imperméabilité des aménagements et la saturation du pluvial,

[5] http://webworld.unesco.org/water/ihp/db/glossary/glu/indexdic.htm

- les inondations torrentielles dont la durée est souvent courte et caractérisées par une charge solide considérable,
- les inondations rapides caractérisées par des temps de concentration courts (entre 6 et 12h), et des débits très importants (KOBIYAMA et GOERL, 2007).

Les phénomènes hydrologiques abordés dans ce travail concernent exclusivement les inondations rapides, c'est-à-dire causées par un évènement pluvieux intense et des conditions hydrologiques et morphologiques relativement propices. LEFEVRE et SCHNEIDER (2003) en distingue trois types (instantanée, subite et rapide) qui différent par leur extension géographique, leur durée et leur débit de pointe.

A ces considérations d'ordre hydrologique, le présent travail, adoptant une vision opérationnelle, utilise plutôt l'expression, « crue à cinétique rapide », représentative de délais opérationnels courts. L'analyse hydrométéorologique qui suit, s'appuie cependant sur le cas des crues éclair régulièrement observées dans le massif des Cévennes.

I - 1.2.2. Phénomènes météorologiques

Le facteur prépondérant de la génération d'une crue éclair est relatif aux précipitations. De manière générale, une augmentation de débit sur de courts délais est provoquée par des pluies orageuses caractérisées par des intensités horaires parfois très élevées (75 mm/h à 300 mm/h) (KELSCH et *al.*, 2001). Plusieurs types de perturbations atmosphériques sont susceptibles d'engendrer de tels phénomènes hydrologiques (FAURE et *al.*, 1994) :

- des « champs pluvieux frontaux […] de 5 à 15 km de large » (FAURE et *al.*, 1994) dus à la rencontre d'un front froid avec une masse d'air chaud et humide. La localisation des précipitations est directement corrélée à celle du front froid,
- des masses d'air, chargées en humidité, formées au dessus de la mer Méditerranée et remontant vers le Nord. Dans certains cas, l'ensemble de ces structures d'orages monocellulaires se répartissent en ligne pour former un système occupant une vaste zone (centaines de km) : un « système convectif de méso-échelle » ou SCM (Figure 3) (RIVRAIN, 1997).

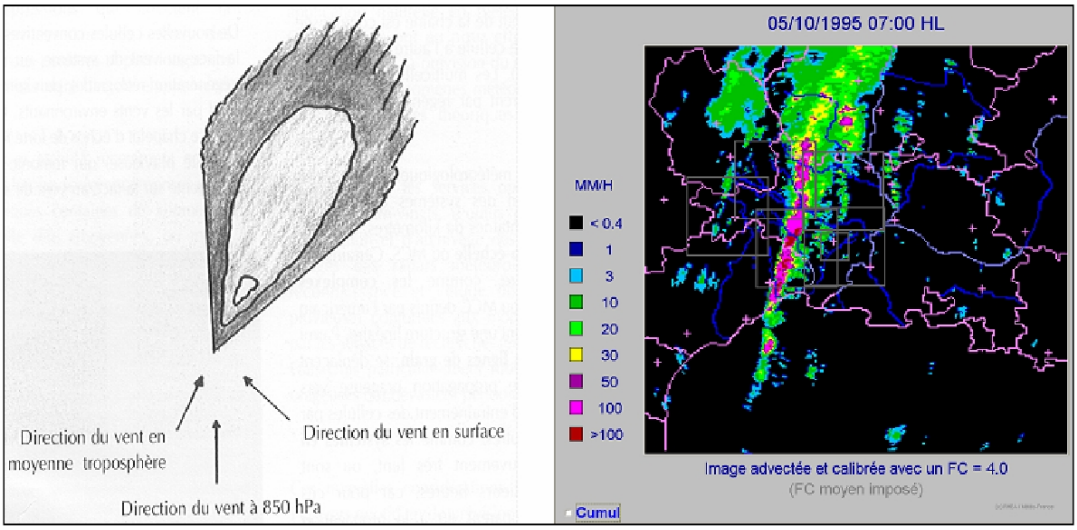


Figure 3 : Système convectif de méso-échelle en V (tirée de RIVRAIN, 1997 - Image radar RHEA)

De manière schématique, dans les Cévennes, les mois d'automne offrent des conditions climatiques propices au développement d'un SCM, caractérisées par la présence :

- d'une situation dépressionnaire au Nord de la France, relativement stationnaire, formée par des masses d'air froid provenant de l'Arctique, le terme de « goutte froide » est souvent employé pour désigner cette masse d'air froid polaire situé en altitude dans nos régions. Selon l'altitude de ces gouttes froides, on observe des pluies continues ou de forte intensité associées à des orages,
- d'un anticyclone centré sur le sud-est de l'Europe,
- d'un pouvoir convectif fort de la mer Méditerranée due à ses eaux chaudes.

Les masses d'air froid de basses pressions remontent du sud-ouest de l'Europe et rencontrent les masses de hautes pressions provoquant ainsi un flux de sud à sud-est. Au dessus de la Méditerranée, le fort gradient thermique entre les basses et les hautes altitudes provoquent une augmentation de la turbulence de l'air et de la condensation. Cette situation découle généralement sur la formation de systèmes orageux dont la dimension peut croître avec l'apport d'air humide. Une situation météorologique spécifique en V est souvent observée où les cellules orageuses se forment en Méditerranée à la pointe du V et se déplacent vers le Nord (Figure **3**). Cette situation météorologique est généralement caractérisée par une capacité régénérative forte et des cellules orageuses fortement pluvieuses. La difficulté de prévoir leur formation et leur évolution (répartition spatiales des champs pluvieux et durée de vie) rend par conséquent l'anticipation des réactions hydrologiques problématiques.

De plus, à une échelle plus grande, les systèmes convectifs perturbés arrosant régulièrement les Cévennes, correspondant contreforts méridionaux du Massif Central, peuvent engendrer un effet orographique brutal (VIERS, 1968), du fait de la proximité entre la mer méditerranée et le massif montagneux des Cévennes, qui favorise des précipitations régulières à forte intensité (Figure **4**). Dans ces conditions, souvent stationnaires, il est question d'épisode cévenol[6] qui peut durer plusieurs heures du fait du blocage orographique par le massif des Cévennes.

[6] Si ce phénomène, encore mal expliqué (NEPPEL et *al*., 1998), peut s'activer tout au long de l'année, les relevés statistiques de ces 50 dernières années dénotent leur prédominance à l'automne (65% des évènements ont eu lieu en septembre, octobre et novembre

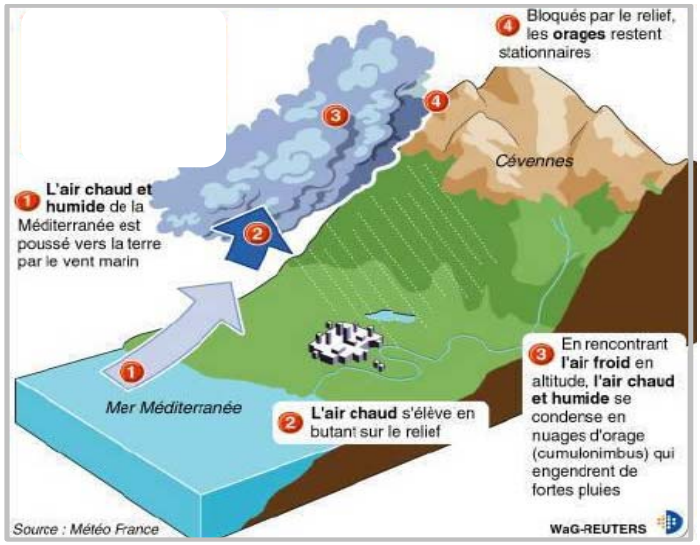


Figure 4 : L'épisode cévenol (source : Météo-France)

La forte localisation de ces phénomènes et leur rapidité de génération représentent, comme dans le cas d'un SCM, des facteurs limitant à leur prévision en situation opérationnelle.

Finalement, dans les deux cas, il apparait indéniable que les précipitations engendrées provoquent des réactions hydrologiques sur les bassins versants concernés. Cependant, à pluie supposée équivalente, cette réaction hydrologique et les débits observés aux exutoires sont régulièrement différents. Ainsi, au facteur pluviométrique déjà considéré viennent se greffer d'autres éléments d'ordre hydrologique, géologique, géomorphologique, pédologique, écologique et anthropique intéressants à analyser.

I - 1.2.3. Définition et processus hydrologiques d'une crue à cinétique rapide

D'une manière générale, la structure géologique en place et de manière plus précise la perméabilité de la roche mère sous-jacente conditionne d'une certaine façon les caractéristiques des écoulements en profondeur et en surface. De plus, les modelés du paysage, dépendants du substratum et du climat, ont des caractéristiques structurelles qui contraignent de manière prépondérante les modalités d'action des autres facteurs géographiques (GAUME, 2002). Suivant l'échelle considérée, la forme du paysage caractérisable par le système « pente – orientation – altitude » va aussi agir sur l'état et les processus hydrologiques. De ce point de vue, il est essentiel de considérer les phénomènes hydrologiques à plusieurs échelles. La forme et la longueur du bassin versant ainsi que la densité et l'organisation du réseau hydrographique sont à l'origine de réponses hydrologiques différentes. A une échelle plus grande (au sens géographique[7] du terme), c'est-à-dire celle du versant, ces variations géométriques déterminent les propriétés hydrodynamiques de la surface et du sol.

De manière schématique, les processus hydrologiques concourant à l'alimentation d'un cours d'eau sont liés aux échanges hydrologiques entre l'atmosphère, la biosphère et le sol et, aux échanges hydriques entre la biosphère et le sol. Il s'agit au travers de cette analyse d'identifier les processus mis en jeu pour déterminer la part des eaux précipitées qui contribue directement ou indirectement à l'écoulement des cours d'eau, le terme de « pluie nette » est utilisé. Les gouttes d'eau issues des pré-

[7] Une grande échelle couvrira un petit territoire (1/5 000ème par exemple) et une petite échelle un grand territoire (1/100 000ème par exemple)

cipitations arrivent dans la biosphère par gravité. Une proportion de cette eau rentre en contact avec le sol et une autre avec la végétation (biosphère). Pour la seconde, deux processus principaux sont observés :

- l'interception, variable d'une espèce végétale à l'autre, qui retient une part de la pluie brute. Une faible part de l'eau interceptée peut ruisseler le long du tronc et contribué aux processus hydrologiques au sol,
- l'évapotranspiration, qui à l'inverse retransmet à l'atmosphère l'eau transpirée par le couvert végétal.

Dans le cadre des crues à cinétique rapide, l'intensité et la durée des précipitations observées, particulièrement importantes, tendent à minimiser la contribution des processus « atmosphère – biosphère » et à exacerber ceux agissant au niveau du sol. Ainsi, le devenir des précipitations qui atteignent effectivement le sol se partage entre l'infiltration et le ruissellement direct.

Les processus hydrologiques, sont régis par la perméabilité et l'infiltrabilité du sol (COSANDEY et *al.*, 2003). Elles représentent respectivement la capacité d'un sol à « laisser passer l'eau » et la quantité d'eau susceptible de s'infiltrer dans le sol. Ainsi, le ruissellement de surface ainsi que les transferts souterrains, contribuant au débit des cours d'eau, sont appréhendés au travers de ces deux caractéristiques. La Figure **5**, proposée par AMBROISE (1998), illustre les principaux processus à l'origine des écoulements de crue.

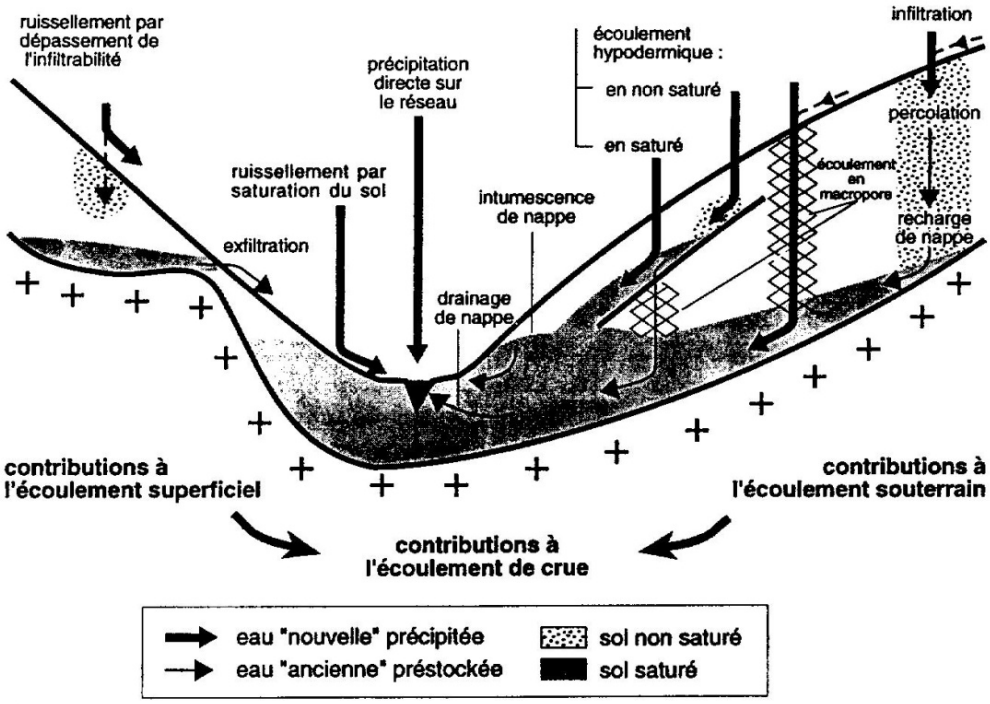


Figure 5 : Genèse des débits de crues : principaux processus superficiels et souterrains (tirée d'AMBROISE, 1998)

Plusieurs types d'écoulement sont générés lors d'un évènement de crue pour les petits bassins versants touchés par des crues à cinétique rapide (Tableau **1**). Leur localisation et la part de leur contribution à l'écoulement de crue sont relativement difficiles à évaluer.

	Types de ruissellement	Processus
Écoulements de surface	**Dépassement de l'infiltrabilité**	L'intensité des précipitations dépasse le seuil d'infiltrabilité du sol et provoque un ruissellement direct (HORTON, 1933).
	Saturation du sol	les précipitations tombent sur ces surfaces quasiment imperméables, il est question de ruissellement par le bas
	Exfiltration	Les précipitations sur des sols non saturés peuvent chasser des eaux souterraines préexistantes. Il est question de l'effet piston
Écoulements de subsurface	**Écoulement hypodermique**	Eau circulant dans la frange supérieure du sol dont l'infiltration est bloquée par un milieu sous-jacent fortement saturé
	Intumescence de nappe	Excédent hydrique d'une nappe peu profonde souvent perchée
	Écoulement par macro-pores	Transfert d'eau du à la présence de fissures et de tunnels formés par les animaux ou les racines.

Tableau 1 : Synthèse des types de ruissellement (d'après GAUME, 2002)

Qualitativement, la contribution à l'écoulement de crue provient d'une part du ruissellement de surface et d'autre part de processus hydrologiques dans le sol et les substratums. Ces processus concourent à ce ruissellement et alimentent directement le réseau hydrographique du bassin versant. Les moyens technologiques (disponibilité des données, moyens d'observations, etc.) ne permettent pas à l'heure actuelle d'identifier précisément la répartition de ces phénomènes sur un bassin versant. Il est possible d'assigner globalement (Figure **6**) :

- des phénomènes de saturation de sol aux zones aux pentes les plus faibles et les abords des cours d'eau (fond de vallée),
- le processus hortonien aux sols peu profonds de plus grandes pentes (haut de versant).

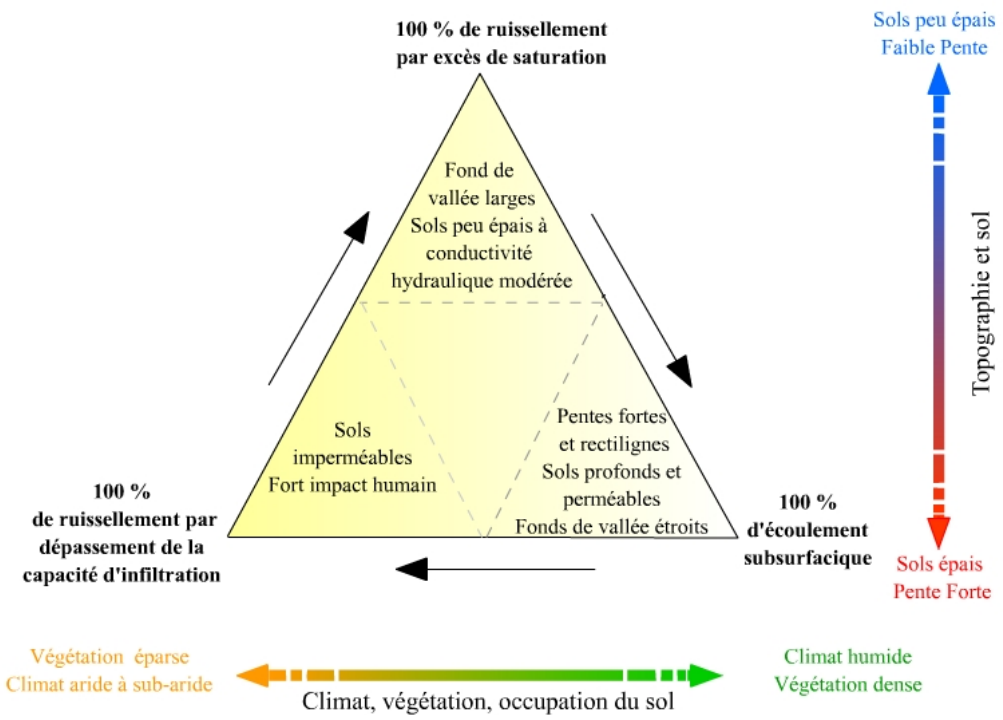


Figure 6 : Représentation des différents processus de génération de l'écoulement en fonction du milieu naturel (tirée d'ANDERSON et BURT, 1990)

Cependant, cette répartition des phénomènes reste très conceptuelle étant donné d'une part le manque de connaissance sur l'épaisseur et la structure du sol, et d'autre part, les processus de transfert, moteurs prépondérant de l'écoulement de crue, qui restent difficilement identifiables. Tous ces types de ruissellement, sans exception, ont lieu à chaque évènement pluvieux et contribuent au débit d'eau à l'exutoire du bassin versant. C'est donc bien dans la quantification et la localisation précise de ces processus que les efforts doivent être menés. Cette variabilité spatio-temporelle des processus météorologiques et hydrologiques et, des vitesses de transferts hydrauliques, conditionnés par des facteurs structuraux (morpho-pédologiques et topographiques) du bassin versant, est abordée par de nombreux auteurs dont AMBROISE (1998), AYRAL (2005), BEVEN (2001), BOUVIER et *al*. (2004), CAPORALI (2007), CHIANG et *al*. (2007), COLLIER (2007), ESTUPINA BORRELL (2004), GAUME (2002), HAZIZA (2007) et MARCHANDISE (2007).

A terme, l'augmentation de systèmes de mesures performants *in situ* et la prise en compte de chaque processus par des modèles hydrologiques et hydrauliques spécifiques pourraient permettre une meilleure compréhension du fonctionnement hydrologique et du cheminement de l'eau sur les bassins versants concernés par ces crues éclair. Cependant, à l'heure actuelle, les services opérationnels utilisent des modèles parcimonieux permettant uniquement de caractériser globalement le débit à l'exutoire et d'évaluer un niveau potentiel de risque.

Ce travail de recherche a pour objet d'étude, les crues à cinétique rapide**,** qui sont théoriquement concernés par les processus hydrologiques abordés précédemment. D'un point de vue hydrologique, cette appellation générale manque certainement de précision en englobant des évènements caractérisés par différents processus hydrologiques, débits de pointe et temps de réponse. Cependant, dans un contexte opérationnel, l'information critique concerne la prévisibilité de dépassement de seuils de débits,

plutôt que l'identification précise des processus hydrologiques et leur répartition à l'échelle du bassin versant. L'ensemble des éléments cités et une synthèse du recueil de définitions de KOBIYAMA et GOERL (2007), permettent de définir une crue à cinétique rapide comme :

Une augmentation soudaine du débit à l'exutoire d'un bassin versant encaissé et de petite taille (quelques centaines de km²) provoquée par des pluies orageuses localisés et stationnaires (de l'ordre de la centaine de mm/h) dans des délais courts (généralement de 6 à 12h) limitant leur prévisibilité et la réactivité des services opérationnels

Si la durée de la crue éclair est généralement de l'ordre de quelques heures, les évènements catastrophiques observés sont souvent caractérisés par plusieurs pointes de crue, rallongeant ainsi la durée du phénomène et *a fortiori* l'inondation des plaines avals.

I - 1.3. La modélisation des phénomènes de crues à cinétique rapide

Il s'agit dorénavant d'analyser les modèles et les plateformes hydrologiques existants destinés à la simulation et à la prévision des crues à cinétique rapide en situation opérationnelle.

I - 1.3.1. La modélisation pluie-débit

Dans ce paragraphe, il est question d'identifier les grandes approches pour la modélisation hydrologique utilisées pour simuler et/ou prévoir les crues à cinétique rapide. Il existe un grand nombre de modèles dédiés, mais aucune solution parfaite n'existe à l'heure actuelle, certainement du fait de l'hétérogénéité spatio-temporelle des phénomènes hydrométéorologiques et du manque de connaissance les concernant (cf. §.I - 1.2.3). L'objet de ce paragraphe n'est donc pas de développer une étude exhaustive des différents courants de pensée en hydrologie, mais d'aborder la question suivant un œil opérationnel, un accent sera ainsi donné aux approches qui ont été testées dans le cadre de la prévision opérationnelle des crues éclair.

De manière générale, les modèles s'étendent d'une approche physique, à une approche empirique selon leur niveau de conceptualisation (ESTUPINA BORRELL, 2004). Les caractéristiques des modèles répondent à des objectifs et des philosophies différentes :

- stochastique/déterministe suivant que le modèle soit basé sur des paramètres faisant appel au calcul de probabilités ou à des lois physiques connues,
- continu/évènementiel suivant que le modèle soit activé lors d'une occurrence exceptionnelle ou exécuté quelque soit les conditions hydrologiques,
- globaux/ distribués suivant que le bassin versant soit considéré comme homogène hydrologiquement ou constitué de sous-ensembles représentant son hétérogénéité fonctionnelle.

Différentes classifications existent donc, celle abordée ici s'inspire des classifications proposées par AMBROISE (1998) ; CONSANDEY et *al*. (2002) ; GAUME (2002) ; GNOUMA (2006).

Ainsi, il convient dorénavant de présenter les trois types de modèles principalement utilisés dans la modélisation des crues éclair, à savoir les modèles empiriques, physiques et conceptuels.

I - 1.3.1.1. Modèles empiriques

L'étude statistique d'échantillons d'évènements passés permet de définir une relation statistique pouvant exister entre la pluie et le débit. Ces études ne prennent finalement en compte que les entrées et sorties de l'hydrosystème[8] étudié. L'inconvénient de ces modèles se situe dans l'impossibilité d'identifier les lois physiques du phénomène qui s'y rattachent. Cependant, dans le cadre de l'ingénierie hydrologique et notamment la prévision opérationnelle des crues rapides ce type d'approche a montré son efficacité. A titre d'exemple, l'approche par réseaux de neurones, critiqué pour sa trop grande abstraction mathématique et son nombre « trop important de variables », a montré des résultats très encourageants dans la prévision à court terme des crues à cinétique rapide (TOUKOUROU et *al*., 2009).

I - 1.3.1.2. Modèles physiques

Ce type de modèle implique un réseau de mesures relativement dense et étendu. Ces modèles sont basés sur les lois physiques de la mécanique des fluides pour modéliser l'ensemble des processus hydrologiques, présentés précédemment. Les données d'entrée sont en principe des données collectées sur le terrain ou par télédétection. Cependant, la méconnaissance des phénomènes hydrologiques à l'échelle du bassin versant et la difficulté de les mesurer *in situ* apparaît comme un facteur limitant à l'utilisation de ces modèles en hydrologie opérationnelle (ESTUPINA BORRELL, 2004 ; MARCHANDISE, 2007). Dans ce sens, certains modèles opérationnels sont basés sur une approche physique où certains paramètres sont calés suivant des approches empiriques et/ou conceptuelles. Le modèle spatialisé[9] SHE (Système Hydrologique Européen) appartient à cette famille (ABBOTT et *al*., 1986). SHE est utilisé pour modéliser l'ensemble des processus hydrologiques présents dans le cycle de l'eau, c'est-à-dire de la pluie jusqu'au débit du cours d'eau. Il peut être utilisé pour modéliser la contribution neigeuse, l'évapotranspiration ainsi que tous les processus d'échanges en atmosphère et sol. Il utilise la loi des équations de Darcy pour la description des zones saturées et non saturées (BEVEN, 2001).

En conclusion, ces modèles physiques sont difficilement utilisables en situation opérationnelle du fait de leur complexité et de leur utilisation sur des bassins versants expérimentaux de petite taille. D'autre part, les besoins en calcul et la difficulté de disposer de l'ensemble des données d'entrée qu'ils requièrent apparaissent comme des facteurs limitant.

I - 1.3.1.3. Modèles conceptuels

Dans le cadre de ces modèles, les processus hydrologiques et hydrauliques sont simplifiés par l'utilisation de réservoirs interconnectés, conditionnés par les données d'entrée, et des paramètres préalablement calés sur des évènements passés. Par exemple, le modèle semi-empirique américain SCS (Soil Conservation Service) est basé sur le seul paramètre de coefficient d'aptitude au ruisselle-

[8] Cette approche permet d'analyser les processus hydrologiques longitudinaux, latéraux et verticaux composant le cycle hydrologique d'un bassin versant (Cosandey et *al*., 2003)
[9] Le terme « spatialisé » signifie que les processus hydrologiques sont spatialement caractérisés

ment (GAUME, 2002 ; BOUVIER et *al*., 2004). Dans d'autres cas, un réservoir peut être utilisé pour modéliser les processus « atmosphère – sol » et/ou pour décrire ceux agissant dans le sol. Généralement, on parle respectivement de modules de production, qui transforme la pluie brute en pluie nette et de transfert qui propage l'eau à l'exutoire ou aux mailles connexes (AYRAL, 2005). Le modèle TOPMODEL, basé sur des aires contributives variables, en est un exemple (BEVEN et KIRBY, 1979 ; ESTUPINA BORRELL, 2004). En mode spatialisé, ces modèles conceptuels permettent d'aborder les processus hydrologiques en se rapprochant d'une approche physique où différents processus hydrologiques peuvent être appréhendés au travers par exemple de calage de paramètres différents (AMBROISE, 1998). Il est aisé de comprendre que ces modèles dont les paramètres peuvent être nombreux souffrent d'une déconnexion entre le paramètre (calé le plus souvent) et la réalité physique qu'il propose de modéliser (perméabilité, capacité d'infiltration, etc.).

I - 1.3.1.4. Synthèse des capacités de modélisation des crues à cinétique rapide

Ces différents types de modèle ont montré leur efficacité et leurs limites dans différentes applications. La particularité hydrologique des crues éclair et la difficulté de les observer précisément rendent leur modélisation délicate et souvent imprécise. A l'heure actuelle, il ne semble pas exister de solutions opérationnelles idéales pour les modéliser et surtout les anticiper précisément. De manière générale, AMBROISE (1998) rappelle que le « *choix d'un modèle pose un dilemme :*

- *soit un modèle simple, facile à caler mais à domaine de validité étroit, ayant un faible pouvoir d'extrapolation ;*
- *soit un modèle détaillé, à large domaine de validité mais difficile à caler, fournissant des simulations correctes mais assorties d'une large incertitude ».*

Dans ce contexte, ESTUPINA BORRELL (2004) s'attache malgré tout à définir le modèle à envisager pour répondre au mieux à ces contraintes de modélisation des crues éclair :

« un modèle hydrologique a priori déterministe, distribué en surface et en temps, perceptuel et à base physique […] il sera préférable que les temps de calcul soient courts et que les paramètres du modèle soient porteurs d'un maximum de sens physique de façon à être mesurables, anticipables ou portables (il faut limiter les procédures de calage à partir de longues chroniques d'observations pluie - débit) »

Cette difficulté à anticiper les crues éclair par la modélisation, impose donc d'adapter les capacités de modélisation des services opérationnels, en développant des approches intégrant plusieurs capacités de décision et/ou en s'attachant à combler l'incertitude de la prévision par le recours à des indicateurs de niveaux de risque. Enfin, la variabilité des processus hydrologiques engagés durant les crues éclair limitent *a priori* l'utilisation d'un modèle unique. Ces dernières années, l'hydrologie opérationnelle s'est donc orientée vers des solutions :

- intégrant et conjuguant des capacités d'expertise pluri-thématique telle que la météorologie et hydrologie,
- diversifiant et multipliant les approches de modélisation de ces deux champs scientifiques,

- offrant des fonctionnalités de visualisation, d'interrogation, de comparaison et de collaborations.

Elles ont suggéré le recours à des plateformes de modélisation, qu'il convient maintenant d'approfondir.

I - 1.3.2. Les plateformes de modélisation

Les 4 plateformes de modélisation présentées ci-après ont en commun d'être associées à une approche opérationnelle. En fait, leur concepteur les considèrent comme des modèles à part entière, alors que dans cette étude, elles sont abordées au travers de la notion de plateforme du fait de leur aspect composite, intégrant fonctionnalités SIG (Systèmes d'information Géographique) et différents modèles hydrologiques. Ainsi, le terme de « modèle » est réservé aux fonctions « métier » d'une plateforme que l'on désignera donc comme le support technologique de la modélisation hydrométéorologique multi-modèles.

I - 1.3.2.1. ATHYS, ATelier HYdrologique Spatialisé

Il s'agit d'une plateforme développée par l'Institut de Recherche pour le Développement (IRD) pour gérer les problématiques liées à l'eau (inondation, études d'impact, etc.) (BOUVIER et DELCLAUX, 1996). Elle intègre de manière homogène plusieurs modèles hydrologiques avec en parallèle des fonctionnalités de gestion et d'analyse de données spatiales basées sur le logiciel libre GRASS (NETELER et MITASOVA, 2004). Elle est donc composée de trois modules principaux :

- MERCEDES (Maillage Élémentaire Régulier Carré pour l'Étude Des Écoulements Superficiels) fournissant plusieurs modules de production et de transfert (TOPMODEL, SCS, ondes cinématiques, etc.),
- VISHYR permettant la préparation et l'analyse statistique des données hydro-climatiques stationnelles,
- VICAIR regroupe un ensemble de fonctionnalités traitement des données géoréférencées et d'analyse spatiale.

Au vue de cette plateforme, l'information géographique et hydrologique peut être gérée de manière continue en adéquation avec les différentes étapes de la recherche en hydrologie (ANCTIL et *al.*, 2005). Des analyses statistiques peuvent être effectuées sur l'ensemble des données. De plus, le calage des modèles et les analyses post-évènementielles (BOUVIER et *al.*, 2004) sont facilités du fait de l'intégrité préalable des données.

I - 1.3.2.2. MARINE (Model of Anticipation of Runoff and INondations for Extreme events

La plateforme MARINE offre des capacités de modélisation hydrologique et hydraulique à base physique spatialement et temporelle distribuée (ESTUPINA-BORRELL et *al.*, 2006). L'originalité de cette approche réside dans l'intégration de données satellitales permettant d'intégrer les caractéristiques structurelles des bassins versants et des cours d'eau, et dans les 4 différents algorithmes de modélisation :

- de l'infiltration (hydrologie)
  - par le concept de Horton
  - par le modèle de Green & Ampt
- du ruissellement (hydraulique)
  - par le concept des isochrones variables
  - par une résolution locale de l'équation de l'onde cinématique

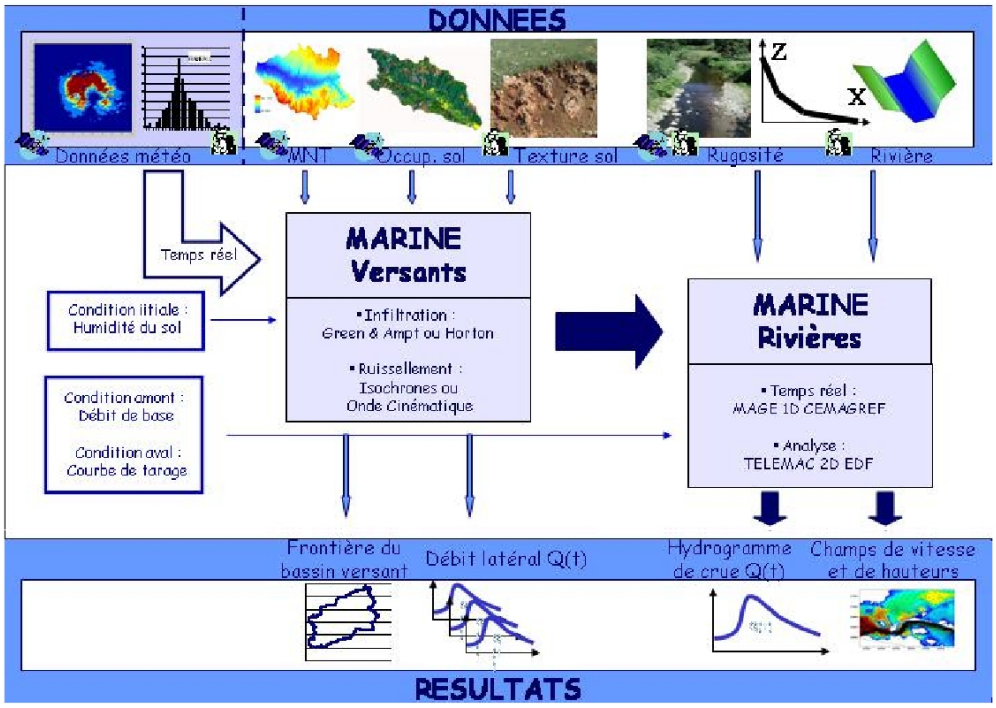


Figure 7 : Structure de la plateforme MARINE (tirée d'ESTUPINA-BORRELL, 2004)

L'objectif principal de cette plateforme est sa capacité à construire dynamiquement des chaînes de modélisation adaptées à la variabilité des processus hydrologiques des crues éclair (Figure **7**). Elle permet de modéliser les débits des bassins versants amont et leur propagation sur les plaines avals. L'efficacité de cette approche a été démontrée et justifiée par son intégration dans le projet opérationnel PACTES (Prévention et Anticipation des Crues au moyen des TEchniques Spatiales), qui « vise à mettre à disposition des utilisateurs des produits permettant de réaliser une prévision et une prévention efficaces des phénomènes de crues » (ESTUPINA-BORRELL et *al.*, 2006). Enfin, les fonctionnalités SIG sont implicitement intégrées à l'aspect modélisation hydrologique pour soutenir l'approche multi-scalaire de cet outil.

I - 1.3.2.3. ALHTAÏR, Alarme Hydrologique Territoriale Automatisée par Indicateur de Risque

Cette plateforme est aujourd'hui opérationnelle dans le service de prévision des crues « Grand Delta » (SPC-GD) (BRESSAND, 2002 ; AYRAL *et* al., 2007). Sa structure est caractérisée par un faible couplage entre les fonctionnalités SIG et la modélisation hydrologique. On peut distinguer trois composants principaux :

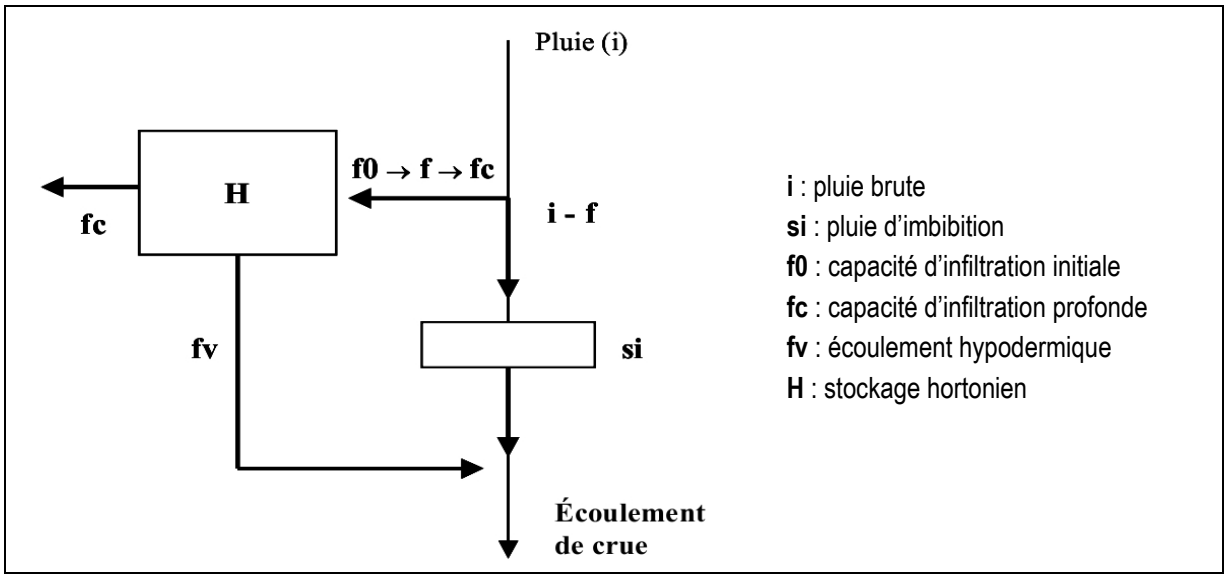
- HYDROKIT qui fournit un ensemble de fonctionnalités SIG pour la préparation des donnés topographiques et des paramètres du modèle principal. Cet outil a été développé par une compagnie privée, Strategis®,

- CALAMAR (CAlcul de LAMes d'eau à l'Aide du Radar) qui est développé et alimenté par le bureau d'études RHEA®. Il fournit des lames d'eau observées et prévues grâce à un réseau de radars au sol. De plus, une application permet leur visualisation, leur interrogation et leur traitement,

- La modélisation distribuée permettant le suivi et la prévision hydrologique est basée sur des données d'entrée fournies par les deux précédents composants. Elle est gérée grâce à une interface graphique permettant de comparer les données collectées avec les hydrogrammes modélisés, et d'établir les dépassements de seuils hydrologiques critiques (notion d'indicateurs). Son module de production est basé sur le principe d'Horton (Figure 8) et le module de transfert sur la loi de transfert géomorphologique donnée par :

$$V = (1 + \frac{p-1}{9}) \times L^{0,25}$$

où V est la vitesse de propagation exprimée en m/s,
p est la pente en %,
L est la distance hydraulique à l'exutoire en km



« *La pluie brute (i) est soustraite à la capacité d'infiltration, la pluie efficace (qui « ruisselle » au sens hortonien) doit remplir le volume que représente la pluie d'imbibition (si) avant de participer activement à l'écoulement de crue. La capacité d'infiltration initiale (f0) diminue au cours de l'événement pour tendre vers l'infiltration profonde (fc). L'écoulement hypodermique (fv) se génère par la vidange du stockage hortonien (H) et participe à « retardement » aux écoulements générateurs de la crue. A la fin de l'événement, le volume d'imbibition (si) se vide au travers du stockage hortonien* (AYRAL, 2005).

Figure 8 : Fonctionnement du module de production d'ALHTAÏR (d'après AYRAL, 2005)

Faiblement couplés, ces 3 modules offrent trois points de vue différents pour l'expertise hydrologique du SPC-GD. Les interfaces graphiques conviviales et adaptées permettent aux prévisionnistes de développer des simulations hydrologiques paramétriques, d'analyser des données hydrométéorologiques multi-sources afin de prévoir les situations de crues extrêmes.

I - 1.3.2.4. AIGA, Adaptation de l'Information Géographique pour l'Alerte en crue

Il s'agit d'une méthode développée par le CEMAGREF et Météo-France (Direction Interrégionale Sud-Est - DIRSE) basée sur le traitement des données pluviométriques (radar et réseaux de mesures) et une modélisation « pluie-débit » (LAVABRE et GREGORIS, 2005). Cette plateforme est connectée à une base de données hydrométéorologiques de référence générée par la méthode SHYPRE (Simulation d'HYDrogrammes pour la PREdétermination des crues) (ARNAUD et LAVABRE, 2000). Au travers de SHYPRE ou de sa version régionale (SHYREG) (LAVABRE et *al*., 2003), un générateur stochastique de pluie horaire, utilisant des chroniques hydrométéorologiques de référence, permet d'établir des cartographies de quantiles de pluie pour différents horizons de prévision et périodes de retour.

Ainsi, les champs de pluie observés ou prévus, fournis par le réseau ARAMIS (Application Radar A la Météorologie Infra Synoptique) de 24 radars géré par Météo-France, sont comparés avec les données de référence obtenues par la méthode SHYPRE, afin d'établir un risque pluviométrique. En second lieu, ces données de pluie sont injectées dans un modèle « pluie-débit », le débit modélisé ainsi obtenu est lui aussi comparé avec la référence SHYPRE en vue d'établir un risque hydrologique.

L'intérêt d'AIGA réside dans sa capacité à établir un risque hydrométéorologique spatialisé pour l'ensemble d'un territoire soumis à des crues de différents types. Cette plateforme cherche à répondre aux besoins opérationnels, en privilégiant l'identification de dépassement de seuils, plutôt qu'une évaluation quantitative plus incertaine. La capacité de prévision de cette plateforme est de l'ordre de quelques heures. D'un point de vue technologique, cette plateforme intègre de manière transparente des fonctionnalités géomatiques, statistiques et de modélisation hydrologique.

L'ensemble de ces plateformes répondent à des besoins opérationnels de gestion de crise hydrologique. ATHYS et surtout MARINE montrent certainement la meilleure voie à suivre en termes d'intégration et de modularité des fonctions « métier ». Cependant, ces deux approches ne sont actuellement pas adaptées à un contexte temps-réel, malgré les résultats prometteurs de MARINE en prévision opérationnelle dans le cadre du projet PACTES (Prévention et Anticipation des Crues au moyen des Techniques Spatiales). Même si AIGA fournit une lisibilité concise et opérationnelle du risque hydrométéorologique, en améliorant la logique d'indicateurs de risque initiée par ALHTAÏR, sa faible pro-

pension pour la prévision précise des débits limite son efficacité opérationnelle. La simplicité d'utilisation est certainement à mettre à l'actif d'ALHTAÏR qui a été entièrement conçue dans l'optique opérationnelle de modélisation et de prévision des crues éclair en temps-réel du service de prévision des crues Grand Delta (SPC-GD) particulièrement concerné par les phénomènes de crues à cinétique rapide.

I - 1.4. Synthèse des crues à cinétique rapide

Cette première partie du chapitre 1 a permis de positionner les crues à cinétique rapide en tant que risque majeur, de part leur faible occurrence et leur intensité généralement forte. Leur rapidité de génération et les processus hydrométéorologiques complexes et variés qui les caractérisent en font un phénomène naturel particulier, difficile à prévoir en situation de crise. Le recours à des modèles pluie-débit et des plateformes hydrologiques permettant d'appréhender les processus physiques dominants, représente une solution opérationnelle prometteuse, ayant motivé plusieurs projets de recherche. En parallèle, les efforts menés dans les dernières années sur la caractérisation d'ensemble de cet aléa et les « réflexes » opérationnels correspondants se sont avérés indispensables à sa gestion efficace en situation de crise. Ainsi, dans la deuxième partie de ce chapitre, il convient de présenter précisément le cadre institutionnel et opérationnel de la gestion de crise relative à cet aléa.

I - 2. Cadre institutionnel et opérationnel de la gestion de crise « inondation »

Dans les paragraphes précédents (cf. §.I - 1), le contexte général de la recherche a été présenté. Il a permis d'identifier de manière générale le cadre théorique dans lequel cette recherche s'intègre. Il convient de présenter dans cette partie le cadre institutionnel dans lequel la prévision hydrologique s'inscrit, et les éléments fondamentaux de la chaîne d'alerte relative aux inondations.

I - 2.1. La notion de vigilance

Le déclenchement d'une gestion de crise « inondation » correspond aux mises en vigilances hydrologique et météorologique d'une ou plusieurs zones du territoire français. Initialement placés en vigilance, la situation hydrométéorologique sur ces territoires peut évoluer vers une situation de pré-alerte et d'alerte des services de secours. Ces nuances syntaxiques prennent véritablement leur sens dans une optique de gestion de crise. Ce dispositif de mise en vigilance progressive permet une montée en puissance progressive des moyens d'expertise hydrométéorologique et de protection des populations. Cette notion de vigilance, par laquelle une situation de catastrophe peut naître, se définit comme :

« une mise en garde sur la potentialité de survenance d'un évènement dangereux, sans en assurer sa venue pour autant » (HAZIZA, 2007)

Elle permet à la population d'adopter certaines précautions et aux services de la sécurité civile de mobiliser progressivement les équipes d'intervention.

I - 2.2. La vigilance météorologique : l'instigateur de l'alerte

I - 2.2.1. L'expertise sur les phénomènes météorologiques

De part les conclusions du paragraphe I - 1.2, l'observation météorologique apparaît essentielle à la prévision des phénomènes d'inondation et contribue donc au déclenchement de la surveillance hydrologique et plus globalement à la mise en vigilance des services opérationnels de la Sécurité Civile. Il convient donc de présenter les moyens mis en œuvre par Météo-France et ses services déconcentrés pour développer et diffuser la prévision météorologique.

I - 2.2.1.1. Observations météorologiques

A la base de la prévision météorologique à l'échelle française, Météo-France utilise les données météorologiques de l'Organisation Météorologique Mondiale (OMM) basées sur le Système Mondial d'Observation (SMO) pour anticiper dans un délai de quelques heures la survenance de phénomènes météorologiques dangereux. Ce réseau de mesures est constitué de stations au sol, de ballons-sondes, de systèmes de télémesure tels que les radars ou les satellites météorologiques (RAINER, 2003).

A l'échelle nationale, Météo-France utilise un important réseau d'observation, constitué de stations au sol, de radars météorologiques, de radiosondes et d'observatoires répartis sur le territoire. En ce qui concerne les observations au sol, le réseau RADOME (Réseau d'Acquisition et de Données

d'Observation Météorologique Etendue) permet la surveillance et une prévision immédiate des conditions météorologiques à une finesse infra-départementale (TARDIEU et LEROY, 2003). Avec le réseau ARAMIS, Météo-France dispose d'un réseau de 23 radars météorologiques pour la surveillance des précipitations, nouvellement densifié au travers du projet PANTHERE (Projet ARAMIS Nouvelles Technologies en Hydrométéorologie Extension et REnouvellement). L'ensemble de ce réseau de radars doit permettre une meilleure estimation régionale des taux de pluie et du type de précipitations.

I - 2.2.1.2. Modélisation météorologique

En termes de modélisation numérique, Météo-France utilisent trois modèles imbriqués, ARPEGE (Action de Recherche Petite/Echelle Grande Echelle), ALADIN (Aire Limitée, Adaptation dynamique, Développement InterNational) et AROME (Application de la Recherche à l'Opérationnel à Méso-Echelle). Ces modèles sont intégrés dans une chaîne de simulation numérique de l'atmosphère. Cette imbrication de modèles permet à Météo-France de développer une approche multi-échelles (FISCHER et *al.*, 2006) où (Figure **9**):

- Le modèle ARPEGE fait partie partie des modèles de circulation générale (GCM en anglais) et couvre l'ensemble du globe avec une résolution variable ; celle-ci est de 15 km sur le France. Au-delà de trois jours, Météo-France utilise le modèle IFS (Integrated Forecasting System) du Centre Européen pour les Prévisions Météorologiques à Moyen Terme (CEPMMT). Cela permet de fournir des prévisions deux fois par jour avec un horizon de prévision de 72 h.

- Le modèle ALADIN est centré sur l'Europe de l'Ouest et fourni une prévision météorologique à 36 h avec une résolution de 10 km en se basant sur des conditions initiales fournies par les sorties de simulation d'ARPEGE et des cycles d'assimilation de données météorologiques.

- Le modèle AROME, nouvellement créé en 2008, couvre seulement le territoire métropolitain avec une maille de 2,5 km. Il utilise les sorties du modèle ALADIN pour son initialisation. Avec ce modèle, Météo-France est capable de prévoir des phénomènes météorologiques de petite échelle à des horizons de prévision de quelques heures, en utilisant les observations au sol pour réajuster ses prévisions (assimilation de données). Il est destiné en particulier aux services opérationnels de gestion de crise,

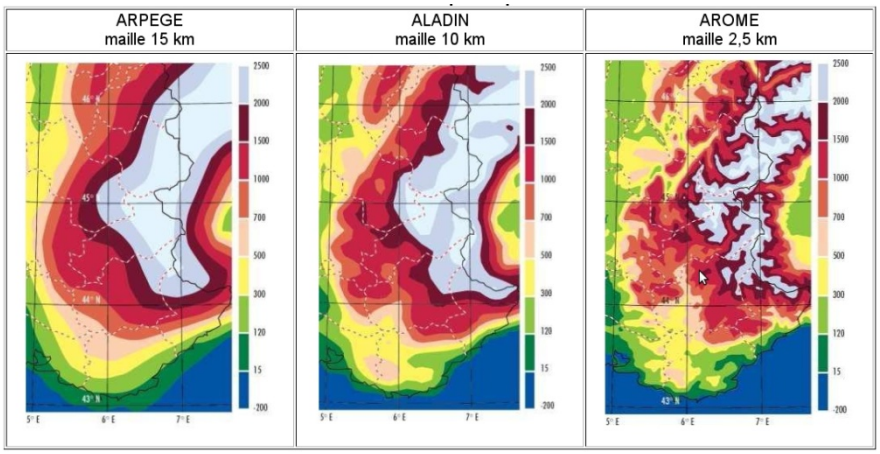


Figure 9 : L'approche multi-échelles de Météo-France : les trois principaux modèles (source : Météo-France)

Ainsi, avec l'ensemble du réseau d'observation et l'imbrication de ses modèles numériques, Météo-France est capable de fournir une prévision à différentes échéances avec une approche multi-échelles. D'un point de vue opérationnel, dans le contexte de gestion de crise des risques naturels d'origine climatique, Météo-France a donc un rôle central dans l'élaboration de l'expertise prévisionnelle.

I - 2.2.2. Les grandes caractéristiques de la vigilance météorologique

De manière générale, Météo-France dispose d'une organisation territoriale multi-échelle adaptée à l'ampleur des aléas météorologiques et aux besoins de communication avec l'ensemble des interlocuteurs de la gestion de crise (HAZIZA, 2007). Cette organisation permet de développer une mise en vigilance précise et adaptée à l'ampleur de l'aléa.

La mise en vigilance météorologique s'appuie sur la circulaire interministérielle du 28 septembre 2001[10], relative à la refonte de la procédure d'alerte météorologique, qui définit univoquement les procédures d'alerte à mettre en œuvre. Cette procédure formalise la transmission d'informations critiques sur l'aléa météorologique. Techniquement, l'outil de communication privilégié est la cartographie, en l'occurrence la carte de vigilance de Météo-France[11] développée par les centres météorologiques régionaux et actualisée 2 fois par jour à 6h et à 16h (Figure 10a). Cette carte a une précision spatiale départementale et précise les dangers météorologiques à un horizon de 24h. Le niveau de vigilance est basé sur une échelle à 4 couleurs :

- le niveau de vigilance « vert » correspond à une situation sans danger,
- le niveau de vigilance « jaune » correspond à des phénomènes potentiellement dangereux mais habituels,
- le niveau de vigilance « orange » justifie la mise en pré-alerte et l'éventuelle montée en puissance des moyens des services dédiés au risque météorologique identifié. La carte de vigilance est accompagnée de la description des phénomènes attendus et de recommandations de sécurité auprès de la population,
- le niveau de vigilance « rouge » s'apparente à une situation de catastrophe naturelle pouvant impliquée le développement d'une situation de crise pour l'ensemble des partenaires engagés.

[10] http://www.anena.org/jurisque/reglement3/systalert/c280901vigil.htm
[11] http://france.meteofrance.com/vigilance/Accueil

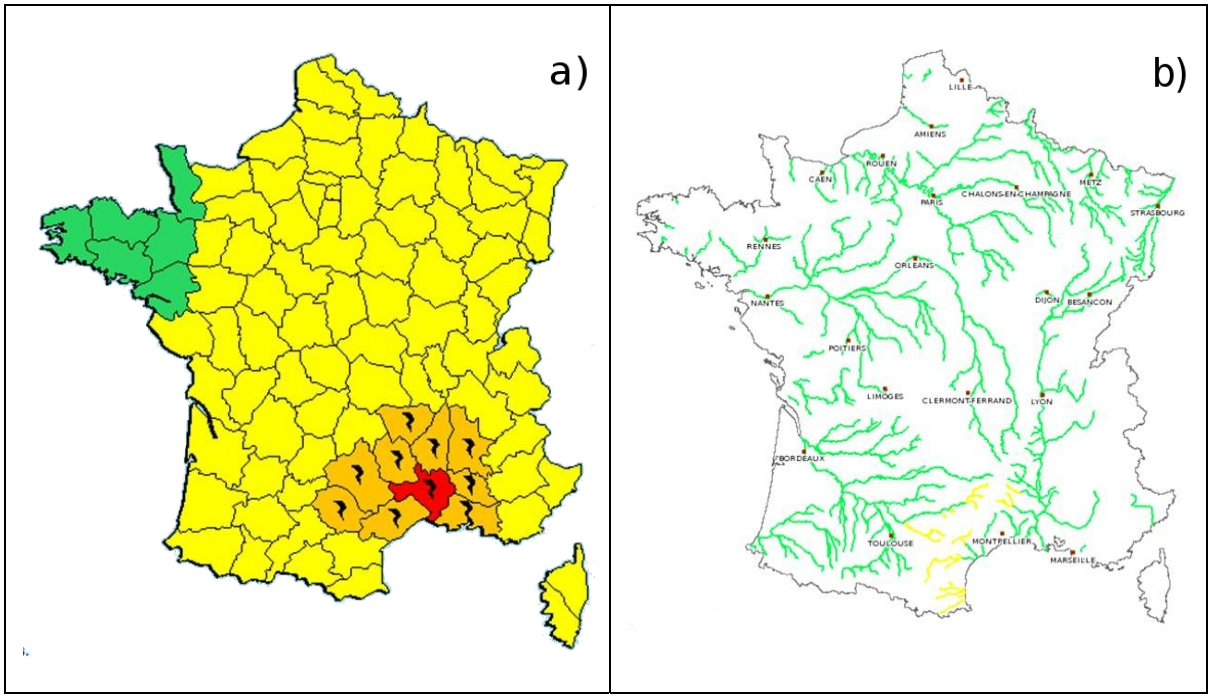


Figure 10 : Carte de vigilance météorologique du 09 septembre 2002 à 01h37 (a) et carte de vigilance hydrologique du 13 avril 2007 à 09h52 (b) (sources : Météo-France et Vigicrues : http://www.vigicrues.gouv.fr/)

En accompagnement des niveaux de vigilance, le centre météorologique régional concerné développe, quand la situation le requiert, un bulletin météorologique spécial (BMS) qui détaille le phénomène météorologique susceptible de toucher un territoire (localisation, période, durée, conséquences, conseils de comportement, …) et qui est susceptible d'appuyer le processus d'aide à la décision du dispositif de gestion de crise.

I - 2.3. La vigilance hydrologique : une émergence justifiée

I - 2.3.1. De l'annonce à la prévision des crues

L'annonce des crues a commencé à se structurer en France à partir du milieu du 19ème siècle grâce aux travaux du Conseil des Ponts et Chaussées (HOUDRE, 2001). Ce n'est véritablement qu'au début du 20ème siècle (1910), que l'aléa hydrologique est géré sur l'ensemble du territoire métropolitain par 4 services centraux des inondations (le bassin de la Garonne, le bassin de la Loire, le bassin de la Seine, du Nord et de l'Est, et le bassin du Rhône) assistés de services hydrométriques et d'annonce des crues départementaux qui surveillent le niveau des principaux cours d'eau français :

Progressivement le nombre de services d'annonce des crues a augmenté afin de gérer l'ensemble de la variabilité des réponses hydrologiques que connait le territoire français et de mieux couvrir le territoire national. Ainsi, cette évolution a permis la création de 52 services d'annonce de crues (SAC) répartis sur l'ensemble du territoire français, système qui a perduré jusqu'en 2002.

Les crues catastrophiques de ces 10 dernières années ont remis en cause ce système initial. Les crues de l'Aude en 1999 et du Gard en 2002 ont été révélatrices des difficultés organisationnelles et décisionnelles de la gestion de crise « inondation ». De nombreux dysfonctionnements, mettant souvent en

cause la responsabilité des services opérationnels, ont été identifiés lors des retours d'expérience qui ont suivi (SAUVAGNARGUES-LESAGE et *al*., 2007) :

- recouvrement irrégulier du territoire par les réseaux de surveillance,
- faible utilisation de l'imagerie radar et obsolescence d'une partie des équipements,
- mauvaise répartition des rôles et/ou redondance de missions,
- difficulté d'identification des interlocuteurs parmi les partenaires,
- incohérence du découpage administratif peu pertinent d'un point de vue hydrologique.

L'ensemble de ces difficultés contribue à « atténuer l'efficacité du dispositif global d'annonce des crues en allongeant le délai nécessaire à la transmission de l'alerte » (SAUVAGNARGUES-LESAGE et *al*., 2007). Dans ce contexte, la loi du 30 juillet 2003[12] a proposé la création de 22 services de prévision des crues (SPC) remplaçant les 52 services d'annonce des crues (MEDD, 2002). L'arrêté interministériel du 2 juin 2003[13] a imposé la création d'un service de coordination et d'appui aux services de prévision des crues, le SCHAPI (Service Central d'Hydrométéorologie et d'Appui à la Prévision des Inondations).

I - 2.3.2. Le SCHAPI : entre opérationnalité et innovation

Ce service de compétence nationale a été créé dans l'optique d'assurer une expertise hydrométéorologique et d'assurer une compétence hydrologique à l'échelle national en situation de crise (TANGUY *et al*., 2004). Son rôle est central dans la gestion du risque « inondation » et réaffirme le rôle central de l'Etat dans la prévention des risques technologiques et naturels. En situation de crise, il assure un soutien aux SPC, pour mener à bien leur mission de prévision de crues en les assistant dans leur diagnostic hydrologique tout en produisant une carte de vigilance « inondation » actualisée deux fois par jour. Hors crise, il « exerce une mission d'animation et d'assistance, de conseil et de formation auprès des services et des établissements intervenant dans le domaine de la prévision des crues » (SAUVAGNARGUES-LESAGE et *al*., 2007). Progressivement, en identifiant les différents besoins des SPC, le SCHAPI formalise et harmonise les moyens technologiques et logiciels à mettre à disposition des SPC, en développant « une architecture informatique visant à créer un véritable réseau informatique des SPC » (TANGUY et *al*., 2004).

I - 2.3.3. Les services de prévision des crues : la compétence hydrologique locale

I - 2.3.3.1. Rôle des services de prévision des crues

La création des services de prévision de crues en remplacement des services d'annonce des crues répond principalement à un besoin hydrologique. En effet, les territoires de compétence des SPC cadrent mieux avec les logiques hydrologiques en action. Ainsi, ce territoire « ne se limite plus à des linéaires de cours d'eau mais couvre autant que possible un ensemble de bassins versants hydrologiquement homogènes » (TANGUY et *al*., 2004). Ainsi, les SPC ont généralement à leur charge des terri-

[12] http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000604335
[13] http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000229756

toires englobant aussi bien les processus hydrologiques amont à l'origine de la génération des crues que ceux aval relatifs aux inondations. Ils sont en charge (DIREN, 2005) :

- de la formation de leurs agents,
- du développement des modèles hydrométéorologiques adaptés au territoire de compétence,
- de leur calage[14] avec le soutien du SCHAPI,
- de la diffusion des prévisions réalisées (vigilance) au public,
- du développement progressif d'une capacité de prévision des inondations.

Pour mener à bien leur mission, les SPC ont à leur disposition deux outils réglementaires de surveillance et de prévision visant à formaliser au mieux leurs missions et relations avec les acteurs de la gestion de crise « inondation » :

- le schéma directeur de prévision des crues (SDPC)
- le règlement de surveillance, de prévision et de transmission de l'information des crues (RIC)

I - 2.3.3.2. Le schéma directeur de prévision des crues (SDPC)

Il est développé par le préfet coordonnateur de bassin en s'appuyant sur la DIREN (Direction régionale de l'environnement) de bassin et en collaboration avec les préfets de zone de défense et les préfets de départements concernés. Il vise à définir l'organisation de la surveillance hydrologique pour les SPC de la région concernée. Ainsi, il définit (DIREN, 2005, TANGUY et *al*., 2004, SAUVAGNARGUES-LESAGE et *al*., 2007) :

- le territoire de compétence, les cours d'eau réglementaires et les fonctionnements hydrologiques des bassins versants pour chaque SPC,
- l'organisation mise en œuvre par l'Etat pour mener à bien cette mission de surveillance, tels que les réseaux d'observations (stations au sol et radars),
- les coopérations et la cohérence entre les dispositifs mis en place par les collectivités territoriales, les établissements publics de l'Etat (par exemple Météo-France) et « les maîtres d'ouvrage de grands aménagements hydrauliques ».

En s'appuyant sur ce document chaque service de prévision des crues de la région concernée développe un règlement de surveillance, de prévision et de transmission de l'information des crues (RIC).

[14] Définition du jeu de paramètre d'un modèle adapté au contexte hydrologique en présence

I - 2.3.3.3. Le règlement de surveillance, de prévision et de transmission de l'information des crues (RIC)

Ce règlement vise à définir les moyens mis en œuvre par chaque SPC pour mener à bien les objectifs définis par le SDPC (TANGUY *et al*., 2004). Chaque SPC identifie clairement :

- le fonctionnement et l'historique hydrologiques des bassins versants surveillés,
- la liste des cours d'eau et bassins versants surveillés avec les seuils hydrométéorologiques au-delà desquels des dommages peuvent avoir lieu,
- les données collectées par les réseaux de mesure qui lui sont alloués,
- les données fournies par des services annexes (autres SPC, SCHAPI, EDF, Météo-France, etc.),
- les modèles développés, calés, et utilisés pour mener à bien leur mission,
- les ouvrages hydrauliques susceptibles d'avoir un impact sur les crues (TANGUY et *al*., 2004),
- la liste des communes couvertes par l'information sur les crues élaborée par le SPC et la description de ce dispositif d'information (SAUVAGNARGUES-LESAGE et *al*., 2007).

L'information réglementaire développée par le RIC, sur la surveillance, la prévision et la transmission des informations relatives aux crues peut être mise à jour par chaque SPC au travers d'un rapport annuel approuvé par le préfet concerné.

I - 2.3.4. Les grandes caractéristiques de la vigilance hydrologique

Une vigilance « crue » a été mise en place à partir du 11 juillet 2006[15], en suivant le modèle de la vigilance météorologique. De la même manière, une échelle à 4 niveaux de vigilance (vert, jaune, orange et rouge) a été établie. Le SPC concerné par un évènement hydrologique spécifique et potentiellement dangereux, développe une expertise basée sur ses réseaux de mesures, les sorties des modèles hydrologiques, l'expérience du territoire de ses prévisionnistes et l'appui scientifique du SCHAPI. En utilisant la codification opérationnelle et la description hydrologique du phénomène, le SPC fournit son expertise au moins deux fois par jour (10h et 16h) au SCHAPI qui produit la carte nationale de vigilance crue (Figure 10b). En phase de gestion de crise, et si les conditions hydrologiques sont susceptibles d'évoluer rapidement, la fréquence de renouvellement de l'information de vigilance peut augmenter. Comme pour la vigilance météorologique, le passage en vigilance « crue » orange, implique une montée en puissance des services opérationnels dédiés à la gestion de crise « inondation ». Si le vert indique une situation hydrologique conforme à la normale, et le jaune des évènements hydrologiques ponctuels, à partir de la vigilance orange, le risque d'inondation, avec une menace non négligeable sur les enjeux, impose une attention particulière de la part de tous (y compris la population).

En complément de la publication d'une vigilance hydrologique et lorsque les conditions hydrologiques le requièrent, les SPC ont la possibilité de diffuser des bulletins de suivi aux partenaires opérationnels

[15] http://texteau.ecologie.gouv.fr/texteau/ServletUtilisateurAffichageTexte?idTexte=795

décrivant de manière plus précise le phénomène hydrologique attendu (débit prévu). Il est alors question de prévision des crues (au sens d'anticipation).

SYNTHÈSE : LA GESTION DE CRISE « INONDATION »

Les vigilances météorologiques et hydrologiques, telles qu'elles ont été abordées précédemment, et l'organisation interinstitutionnelle qui en découle, permettent une articulation et une synchronisation avec l'organisation multi-échelles de la Sécurité Civile. Ainsi, ces approches réglementaires qui permettent une structuration forte et univoque de la gestion de crise « inondation » apparaissent de prime abord applicable au cas des crues à cinétique rapide. D'autant que la gestion des crues à cinétique rapide, de part les inondations de plaines qu'elles provoquent généralement, ne revêt pas une organisation particulière de la gestion de crise.

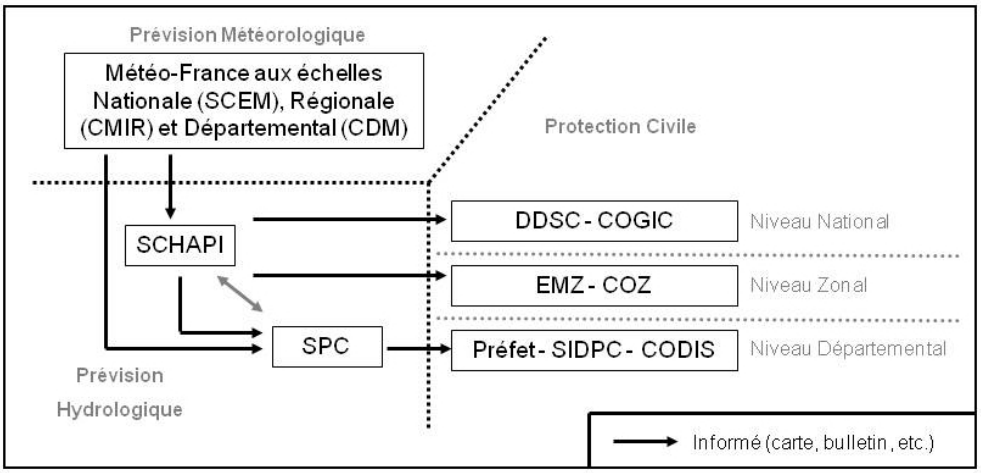


Figure 11 : Liens organisationnels entre les composantes météorologiques, hydrologiques et Sécurité Civile de la gestion de crise "inondation"

D'après la Figure **11**, il est possible d'identifier une organisation tripartite qui se compose tout d'abord de la double expertise météorologique (Météo-France) et hydrologique (SCHAPI et SPC), et d'autre part, les services de la sécurité civile développant une gestion adaptée des secours. Elle est, comme généralement dans la gestion opérationnelle des risques, multi-institutionnelles et multi-échelles.

Le déclenchement d'une situation de vigilance « crue » intervient systématiquement à la suite d'une mise en vigilance météorologique. Le déclenchement des opérations propres à la gestion de crise « inondation » dépend quant à elle de l'intensité et de l'extension géographique de l'aléa, mais surtout de la décision basée sur l'expertise multithématiques et multi-échelles développée par les décideurs. Si en amont, les services dédiés à l'analyse de l'aléa fournissent une information objective et critique sur l'aléa, seuls les responsables de la gestion de crise que sont le préfet (et le maire) sont aptes à activer les secours en adéquation avec les plans d'urgence préalablement définis et à donner l'alerte. Ces services de secours préparent alors leur montée en puissance en entrant en communication avec l'ensemble des experts de leur échelon et en activant les moyens techniques dont ils disposent (DDSC, 2001).

L'une des particularités principales d'un crue à cinétique rapide en comparaison avec une inondation de plaine est son aspect temporel (cf. §.I - 1.2.3). En effet, la rapidité des réponses hydrologiques limite l'efficacité des opérations de mise en sécurité des populations et relègue les services de la Sécurité Civile à des opérations de secours à la personne qui, de part les situations hydrologiques extrêmes en présence, accentuent *a priori* la vulnérabilité de ces services et de la population. Dans ce contexte, la nécessité d'anticiper de tels évènements hydrologiques et de transmettre une information critiquée et critique dans des délais courts apparaît primordiale.

Dans ce sens, et dans l'objectif d'améliorer la prévision hydrométéorologique, une nouvelle vigilance « pluie-inondation » a été mise en place à partir de décembre 2007. Elle cherche à améliorer « l'efficacité de la chaîne d'alerte » fondée sur une « démarche tripartite Intérieur[16]-Écologie-Météo-France » (CHASSAGNEUX et DOLMIERE, 2007). Cette nouvelle vigilance basée sur un nouveau pictogramme (« pluie-inondation » à la place de « précipitations fortes ») propose d'intégrer les deux vigilances, météorologique et hydrologique, en associant clairement des pluies fortes à un phénomène d'inondation et en établissant une concertation intégrée à tous les niveaux de la gestion de crise (Figure **12**). Une telle évolution apparait mieux adaptée à la prise en compte du risque « crue à cinétique rapide ».

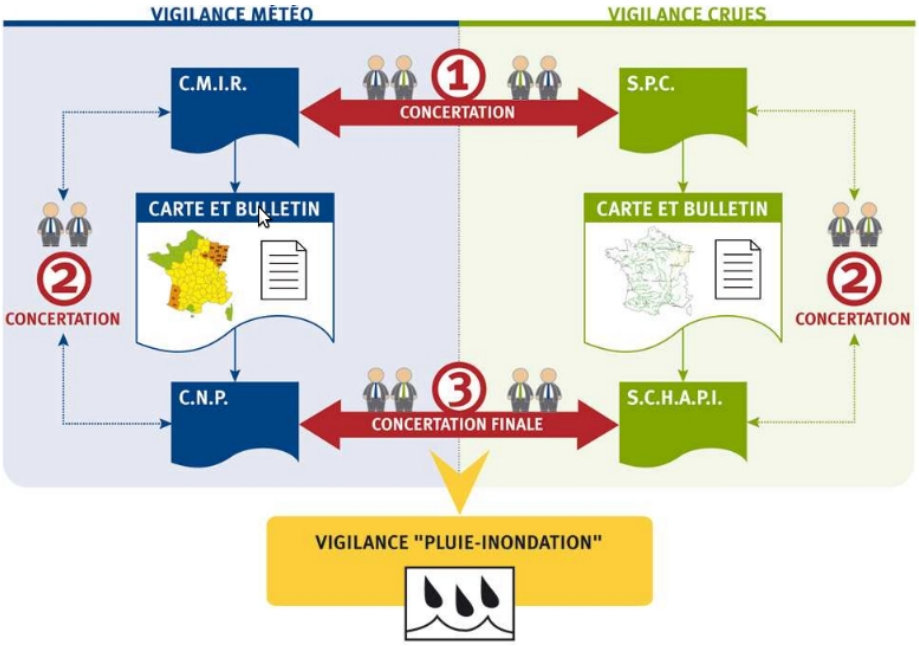


Figure 12 : Développement de l'expertise hydrométéorologique concertée : la vigilance "pluie-inondation" (d'après CHASSAGNEUX et DOLMIERE, 2007)

Cette nouvelle organisation permet d'adapter la réponse opérationnelle (ensemble des actions des 4 groupes de partenaires) à l'évolution de l'intensité de l'aléa et à son extension géographique. L'utilisation de cette nouvelle vigilance est donc pleinement justifiée, sa propension à informer progressivement les partenaires sur la survenance possible d'évènements hydrologiques permet aux différents acteurs de développer progressivement leur montée en puissance, afin de se placer, au moment où la situation le requiert, en gestion de crise. Elle peut permettre une amélioration importante de la gestion

[16] Ministère de l'Intérieur

de crise « inondation » sous réserve de moyens technologiques capables de soutenir ce besoin croissant de collaboration.

Ainsi, le présent travail de recherche tente de s'aligner avec les objectifs de cette nouvelle organisation et cherche à expérimenter des solutions technologiques basées sur l'informatique distribuée pour améliorer l'efficacité de cette organisation multi-partenariale et de la prise en compte de l'aléa « crue à cinétique rapide ». Ainsi, la prévision des inondations provoquées par des crues à cinétique rapide, de part sa faible extension géographique, requiert principalement une vigilance hydrométéorologique à l'échelle locale. Le service de prévision des crues « Grand Delta » (SPC-GD) apparait donc comme le service opérationnel hydrologique le plus pertinent pour opérer une amélioration technologique de la capacité de prévision hydrométéorologique des crues à cinétique rapide en situation de crise. En effet, le territoire « Grand Delta » situé au sud-est de la France entre le sud du Massif Central et le sud des Alpes est régulièrement touché par ce type de crues, en particulier sur les contreforts des Cévennes. Dans ce contexte, l'objet d'étude de cette recherche correspond aux objectifs de ce service et les relations organisationnelles qu'il entretient avec l'ensemble des partenaires de la crise.

Chapitre 2

L'informatique distribuée : une composante nouvelle des sociétés modernes

2.1. Propos introductifs sur les TIC

2.2. L'informatique distribuée et la technologie grille

2.3. La cartographie en ligne entre géomatique et services web

Les technologies de l'information et de la communication (TIC) font aujourd'hui partie intégrante des institutions publiques. L'ensemble des technologies qu'elles englobent offre un éventail de solutions informatiques avec lesquelles ses institutions peuvent composer pour résoudre leurs problématiques de fonctionnement. Dans le cadre de la gestion opérationnelle des crues à cinétique rapide, la technologie grille et les systèmes d'information géographique pourraient s'avérer adapter.

INTRODUCTION

L'informatique distribuée et plus largement les technologies de l'information et de la communication (TIC) s'est, durant les 20 dernières années, progressivement introduit dans le quotidien de la Société et de ses structures décisionnelles. Au niveau sociétal, l'Internet, le World Wide Web (WWW) et la cartographie en ligne se sont progressivement étendu à l'ensemble des organisations et des particuliers. Pour en témoigner, la transmission des vigilances météorologiques et hydrologiques sous forme cartographique (chapitre 1), par Météo-France et le SCHAPI, aux partenaires de la Sécurité Civile et aux populations, se font à l'heure actuelle majoritairement au travers du Web. Cette pérennisation croissante de ces nouveaux moyens de communication offre donc *a priori* une solution en adéquation avec les contraintes fonctionnelles des services en charge de la prévision des crues à cinétique rapide.

Avec l'avènement de ces technologies, des termes nouveaux sont apparus, tels qu'« interopérabilité », « ubiquité » ou bien encore « organisation virtuelle ». Ces termes renvoient à plusieurs technologies informatiques récemment apparues et appartenant au champ de l'informatique distribuée. Dans le cadre de cette étude, deux de ces nouvelles technologies de l'information et de la communication sont plus particulièrement abordées. D'une part, la technologie grille, née, à la fin des années 90, de la volonté croissante de regrouper des réseaux informatiques locaux et des organisations administrativement distinctes, semble offrir un moyen performant de collaboration interinstitutionnelle et de démultiplication des capacités informatiques de chaque partenaire. D'autre part, les services web, par leur capacité de mise en communication de systèmes d'information géographiquement distants, ont enrichi les potentialités de partage et de traitements de l'information géographique au travers des réseaux informatiques.

L'objectif de ce chapitre est donc d'établir un état de l'art de ces nouvelles technologies, en approfondissant dans un premier temps les concepts fondamentaux des nouvelles technologies de l'information et de la communication. Dans une seconde partie, la technologie grille est présentée, en s'attachant à la positionner dans le large champ de l'informatique distribuée et en s'appuyant sur l'historique de sa création. Enfin, les aspects fondamentaux du concept de science de l'information géographique sont présentés, du système d'information géographique au service d'information géographique, et corrélé à la problématique de l'informatique distribuée. Ce chapitre doit donc permettre d'orienter l'hypothèse méthodologique de cette recherche vers une articulation fonctionnelle de ces deux concepts technologiques, au service de l'expertise hydrométéorologique des crues à cinétique en situation de crise.

II - 1. Propos introductifs

II - 1.1. les technologies de l'information et de la communication (TIC)

Il n'existe pas une définition univoque des TIC et les définitions varient en fonction de leur utilisation. Globalement, elles peuvent être apparentées à la dénomination générale de la « Société de l'information » et englobent les concepts gravitant autour de l'informatique. L'OCDE (Organisation de coopération et de développement économiques) définit le secteur d'activité des technologies de l'information et de la communication comme « une combinaison des entreprises manufacturières et de services qui capturent, transmettent et affichent électroniquement des données et des informations » (OCDE, 2002). L'office québécois de la langue française définit les TIC comme :

« l'ensemble des technologies issues de la convergence de l'informatique et des techniques évoluées du multimédia et des télécommunications, qui ont permis l'émergence de moyens de communication plus efficaces, en améliorant le traitement, la mise en mémoire, la diffusion et l'échange de l'information. »

Il est donc possible de caractériser les TIC suivant deux caractéristiques principales, le pendant technologique correspondant à la numérisation de l'information et le pendant social de leur capacité à mettre en relation (CASTELLS, 1998 ; MUCCHIELLI, 1995 ; RALLET, 2006 ; SEGUY, 2008).

Les TIC sont donc composés de :

- l'équipement informatique classique tel que les ordinateurs et les périphériques d'entrée et de sortie connectés,
- réseaux informatiques et de télécommunications,
- logiciels et plus largement de l'ensemble des outils capables de gérer les données numériques,
- activités de « contenu » comme les multimédias et l'audiovisuel,

A partir du milieu des années 90, une nouvelle révolution technologique, basée sur les principes de l'informatique distribuée, a progressivement vu le jour, permettant la connexion croissante d'ordinateurs au travers de réseaux câblés. Pour en témoigner, la croissance exponentielle du nombre d'utilisateurs d'Internet illustrée par la Figure 13.

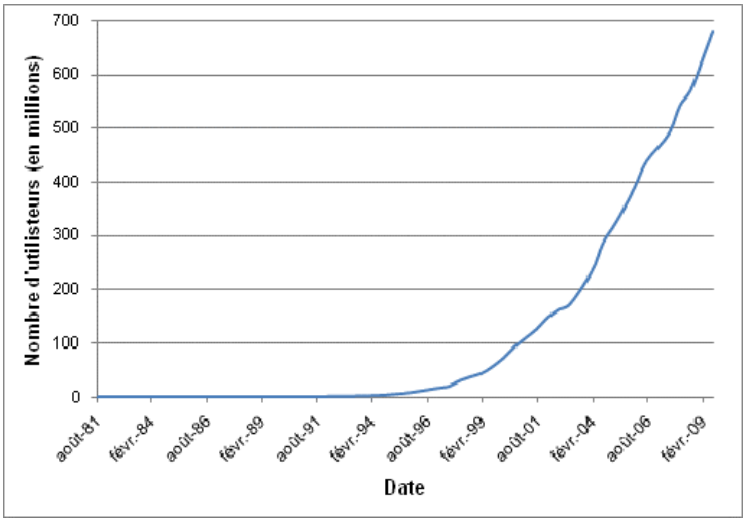


Figure 13 : Évolution du nombre d'utilisateurs d'Internet (source www.isc.org)

D'autre part, les systèmes informatiques se sont progressivement étendus et distribués (Rallet, 2006) (Figure 14). Si, dans un premier temps, la recherche en informatique a permis la connexion de postes informatiques au sein d'une même organisation (Intranet), elle permet aujourd'hui d'en connecter plusieurs géographiquement distantes (Extranet). Cette évolution matérielle s'accompagne d'autre part, d'un développement croissant de solutions logicielles adaptées, le World Wide Web (le terme Web est utilisé dans la suite de ce mémoire), les courriers électroniques et le transfert de fichiers par FTP (File Transfer Protocol) étant sans doute les applications les plus connues et les plus utilisées à l'heure actuelle.

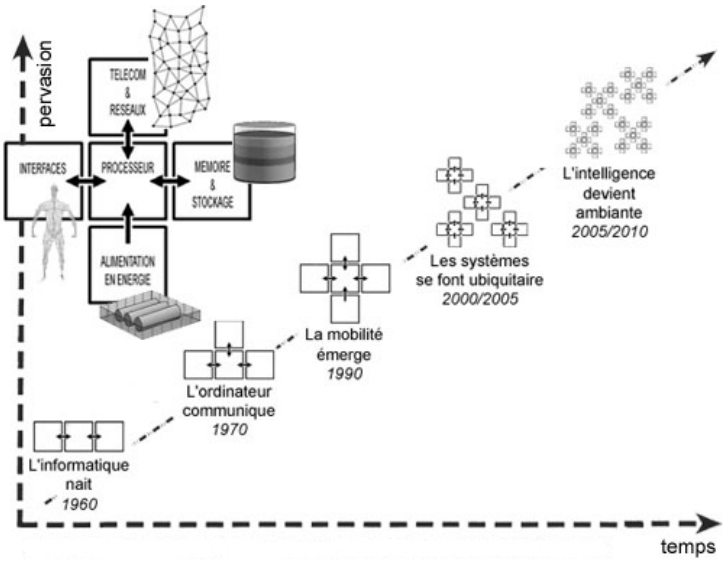


Figure 14 : Évolution de l'informatique de 1960 à nos jours
(adapté de WALDNER, 2007)

Dans ce contexte, l'utilisation de l'informatique et des réseaux à grande échelle se généralise et s'initie dans le quotidien des entreprises. Les TIC sont devenus une composante à part entière de l'organisation et permettent d'entretenir, de soutenir et de faire survivre les compétences acquises, par

leur pouvoir de « capitalisation des savoirs et de gestion des connaissances » (RALLET, 2006). Les fonctions principales des TIC sont (Figure 15) :

- le partage des ressources,
- le partage d'information,
- la communication,
- la coordination.

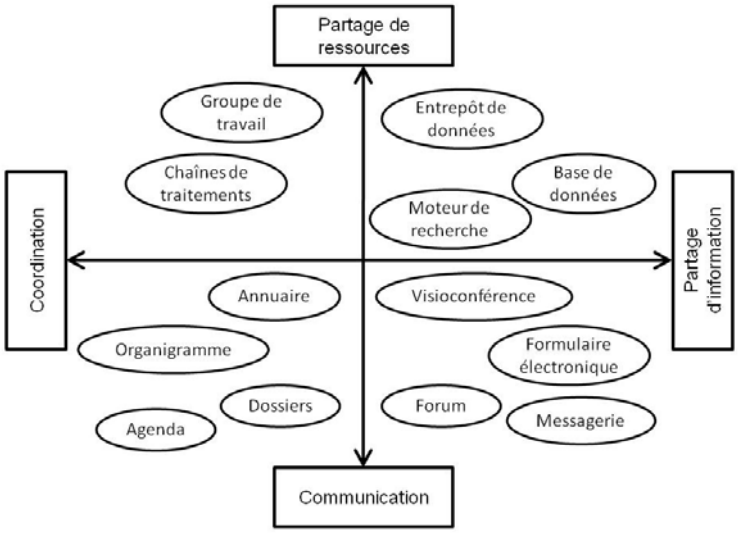


Figure 15 : Ensemble des fonctions de l'intranet, et par extension des TIC

A l'origine appliquée à l'Intranet, ce modèle est extensible à l'Internet et à l'ensemble des technologies « réseau » existantes. Cette « révolution culturelle » comme le titre Le Monde diplomatique « Manière de voir » de février-mars 2010, déterritorialise progressivement le fonctionnement des sociétés modernes (CHOLLET et RIVIERE, 2010). Physiquement distribués, ces réseaux informatiques englobent de plus en plus d'entités matérielles géographiquement distantes et invisibles pour l'utilisateur, au point d'utiliser le terme « ubiquitaire » ou « en nuages » : « environnement d'intelligence artificielle dans lequel les ordinateurs et réseaux sont enfouis et intégrés dans le monde réel. L'utilisateur a accès à un ensemble de services au travers d'interfaces distribuées intelligentes. Ces interfaces s'appuient sur des technologies intégrées dans les objets familiers » (PUZENAT, 2008).

À cette vision fortement virtualisée des ressources informatiques vient se greffer des pratiques sociétales nouvelles. D'après AYACHE (2008), les TIC représentent « le pivot de la société hyperinformationnelle » aboutissant au développement d'un « intelligence collective » que LEVY (1997) définit comme « une intelligence partout distribuée, sans cesse valorisée, coordonnée en temps réel, qui aboutit à une mobilisation effective des compétences ». Plus pragmatiquement, cette vision « philosophique » se vérifie quotidiennement sur le Web, au travers d'outils du Web 2.0 défini comme participatif et social (O'REILLY, 2007). Il est depuis peu question de Web 3.0 ou web sémantique ou « Graphe Global Géant ». Au travers de cette approche, les pages internet sont capables de communiquer entre elles

sans l'intervention humaine grâce à un langage formalisé (BERNERS-LEE, 2007 ; PISANI et PIOTET, 2008 ; ZELDMAN, 2006).

II - 1.2. Collaboration et organisation virtuelle

Globalement, l'avènement des TIC a progressivement ouvert la voie de la collaboration entre des partenaires scientifiques, académiques et industriels géographiquement distants. Regroupés autour d'un ensemble de technologies informatiques, ces utilisateurs poursuivent des objectifs communs (coopération) ou tirent profit de cette organisation pour résoudre leurs propres problématiques (collaboration). Il est question d' « e-collaboration » ou « organisation virtuelle » (BOURAS et *al*., 2008). Ce concept apparait comme un des éléments fondateurs des recherches engagées, durant la fin des années 90 et poursuivies à l'heure actuelle, sur la technologie grille et son utilisation par des communautés scientifiques. Il constitue donc un axe de réflexion privilégié de cette recherche.

II - 1.3. Interopérabilité

L'interopérabilité, terme utilisé dans de nombreux domaines, est fortement liée à cette notion de partage de l'information et donc *a fortiori* de collaboration. Elle correspond « à l'aptitude, pour plusieurs systèmes hétérogènes, à pouvoir communiquer/échanger une information mécanique, électronique ou logique dans un environnement prédéterminé » (DUFLOT, 2007). D'après la Figure 16b, un client (C1) est capable, si les systèmes d'information sont interopérables, d'accéder à différents serveurs distants. A *contrario*, si aucun effort d'interopérabilité n'est engagé, il est souvent nécessaire d'utiliser des clients différents (C1, C2 et C3) pour accéder à des systèmes d'informations distants (Figure 16a).

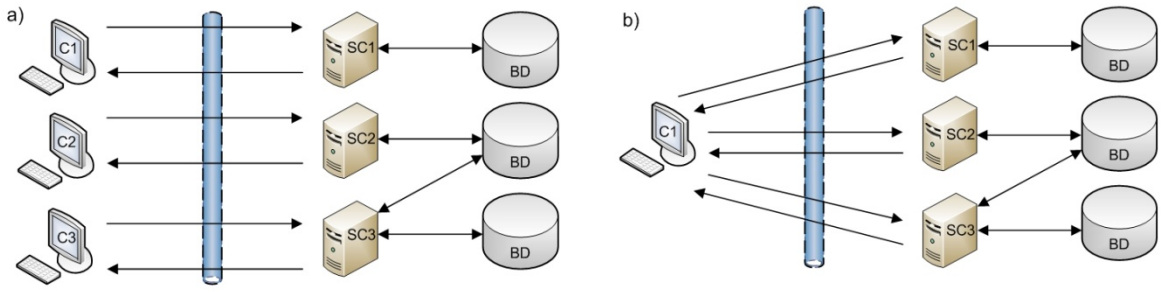


Figure 16 : L'interopérabilité selon l'OpenGIS® (adapté de KOLODZIEJ, 2004)

Ces concepts épistémologiques permettent de mieux situer les avancées technologiques actuelles du monde des TIC et d'intégrer cette recherche dans le vaste champ disciplinaire les englobant. En effet, la technologie grille, et les intergiciels permettant son fonctionnement, peuvent être inclus dans cette vision de l'informatique collaborative et interopérable. Ils sont largement décrits et discutés dans les paragraphes suivants.

Dans ce sens, cette deuxième partie de chapitre permet d'approfondir les concepts fondamentaux de l'informatique distribuée, tels qu'elle est a été présentée précédemment. En se basant sur les caractéristiques principales de ce champ d'application, une attention particulière est développée sur la technologie grille. Dans une seconde partie, la notion de système d'information géographique est présentée pour mieux appréhender la branche plus spécifique des SIG basés sur une architecture réseau, pro-

gressivement apparue dans le courant des années 90 en parallèle de la pérennisation de l'informatique distribuée. L'aboutissement de cette réflexion amène, en fin de chapitre, à considérer la gestion de l'information géographique au travers d'un réseau informatique distribué par l'utilisation de services web dédiés, permettant d'entrevoir la nécessaire interopérabilité des systèmes d'information temps-réel. Finalement, l'objectif principal de cette partie est de mettre en exergue les nouvelles potentialités technologiques existantes, dans la continuité des efforts développés autour de l'amélioration progressive de la chaîne informationnelle de la gestion de crise.

II - 2. L'informatique distribuée et la technologie grille

II - 2.1. Généralités sur l'informatique distribuée

II - 2.1.1. Définition de l'informatique distribuée

L'enjeu technologique, scientifique et politique auquel correspond cette mise en connexion de processeurs ou de systèmes d'information distants a donné naissance au champ disciplinaire de l'informatique distribuée. D'après TANENBAUM et VAN STEEN (2002), un système informatique distribué est une « collection d'ordinateurs indépendants qui apparaissent à l'utilisateur comme un système unique et cohérent ». Une autre définition plus proche du concept de grille est celle de KSHEMKALYANI et SINGHAL (2008), pour qui un système distribué est « une collection d'entités indépendantes qui coopèrent pour résoudre un problème qui ne peut être résolu individuellement ». Du point de vue utilisateur, fondamental dans cette recherche, un système distribué fournit des services (au sens de capacité) à des entités en dissimulant les efforts technologiques mis en œuvre pour y parvenir. Ainsi, au travers d'un système distribué, des utilisateurs peuvent accéder à des ressources[1] que leur seul poste informatique ne leur fournit pas.

Finalement, au quotidien, le développement croissant des systèmes informatiques distribués des entreprises correspond clairement à une logique économique. Le ratio coût - performance tend ainsi à augmenter avec l'utilisation d'un système distribué (KSHEMKALYANI et SINGHAL, 2008).

II - 2.1.2. Une typologie des architectures distribuées

Globalement, un système distribué est caractérisé par des indicateurs structuraux tels que :

- le nombre de nœuds de calcul,
- leur distribution géographique,
- leur homogénéité ou leur hétérogénéité matérielle et logicielle.

A une échelle plus fine, d'autres considérations d'ordre matériel doivent être identifiées, suivant que le système soit constitué de plusieurs processeurs avec une mémoire commune partagée ou de plusieurs ordinateurs ou processeurs disposant chacun de leur mémoire.

En fonction de l'ensemble de ces critères, il est possible de distinguer trois principales architectures distribuées :

- le superordinateur
- le cluster
- la grappe de serveurs ou ferme de calcul

[1] Le terme « ressource » est polysémique, employé seul il fait référence à l'ensemble des composants d'un système d'information tant matériels, logiciels et informationnels.

Dans le cas des superordinateurs et des clusters, l'ensemble des composants informatiques est regroupé géographiquement. D'autre part, les processeurs (unité centrale de traitement ou CPU) et les ordinateurs qui respectivement composent ces architectures, ont des caractéristiques matérielles et logicielles proches, il est question de systèmes fortement couplés. A l'inverse, les grappes de serveurs ou fermes de calcul correspondent à une architecture faiblement couplée. Enfin, une connexion très haut débit leur permet d'échanger rapidement des instructions et des données, stockées sur des mémoires partagées ou distribuées, par l'envoi de messages, dont le standard de communication le plus répandu est le Message Passing Interface (MPI) (WALKER et DONGARRA, 1996). Suivant la taxonomie de Flynn (FLYNN, 1972), de telles architectures s'appuient sur le principe « instructions multiples, plusieurs mémoires » (MIMD) (Foster, 1995).

L'objectif principal de ces architectures est de permettre l'exécution d'un grand nombre d'instructions, le plus souvent interdépendantes, de manière simultanée. Ces architectures permettent ainsi du calcul haute performance (HPC). Les mesures de performance de ces systèmes se font généralement au travers du nombre d'opérations à virgule flottante par seconde (ou flops). Ainsi, en juin 2008, le superordinateur Roadrunner a passé la barre de la puissance « pétaflopique » (10^{15} flops), c'est-à-dire qu'il est capable de réaliser un million de milliards d'opérations par seconde (KOMORNICKI et *al*., 2009). Une telle capacité permet par exemple de modéliser des systèmes complexes comprenant un grand nombre de paramètres (climat, électronique, réactions chimiques, etc.) (FOSTER, 1995).

Le terme de système distribué englobe des architectures informatiques très différentes rendant leur classification délicate, aux systèmes fortement couplés déjà présentés s'opposent des architectures aux éléments souvent très hétérogènes (GOSCINSKI, 1991). Il s'agit de réseaux d'entreprises, d'universités ou de firmes internationales :

- les réseaux locaux ou LAN (Local Area Network)
- les réseaux étendus ou WAN (Wide Area Network)

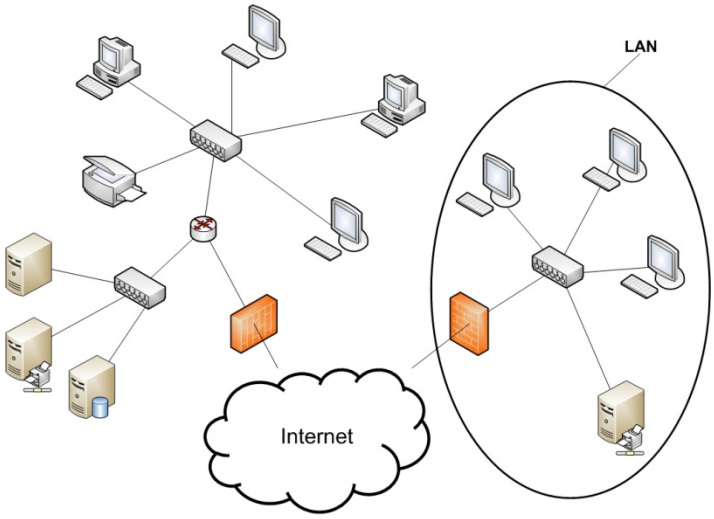


Figure 17 : Exemple d'un réseau informatique étendu englobant plusieurs réseaux locaux (LAN)

A une échelle locale, les réseaux locaux (LAN) permettent d'interconnecter des utilisateurs, des serveurs d'applications (impression, base de données, messagerie, …) et d'accéder à des données distantes, grâce notamment aux protocoles TCP/IP et à l'approche de communication « client - serveur ». A plus grande échelle, des réseaux étendus (WAN) relient plusieurs réseaux locaux géographiquement dispersés et administrativement indépendants (Figure 17).

A ce jour, Internet représente le réseau le plus important en termes d'utilisateurs, géographiquement distants, connectés (Figure 13). Il permet d'accéder à des ressources sans pour autant connaître leur localisation exacte et les systèmes informatiques permettant de les gérer (BERNERS-LEE et *al*., 1994). Cependant, majoritairement informationnel, le réseau Internet n'offre pas de véritables solutions pour le traitement des données.

L'appréhension de ces concepts généraux apparait primordiale pour comprendre précisément le fonctionnement et les enjeux scientifiques gravitant autour d'une architecture de grille qui, dans un sens, cherche à enrichir la capacité calculatoire des réseaux informatiques.

II - 2.1.3. les nouveaux enjeux de l'informatique distribuée

TANENBAUM et VAN STEEN (2002) définissent plusieurs critères fondamentaux de performance d'un système distribué :

- la transparence, dans le sens où la plupart des opérations sur les ressources connectées sont invisibles pour l'utilisateur. La localisation, la réplication, la concurrence d'accès et les erreurs ne sont pas perceptibles par l'utilisateur final.
- l'ouverture qui dénote la capacité du système à accueillir de nouveaux composants aux caractéristiques informatiques différentes sans altérer le fonctionnement du système déjà existant. Au travers de cette notion, Il est aussi question de portabilité des programmes lorsqu'ils sont capables d'être exécutés sur des environnements différents, et d'interopérabilité des systèmes quand ces derniers sont capables de communiquer ensemble malgré des caractéristiques technologiques différentes.
- l'extensibilité ou mise à l'échelle, dans la mesure où le système peut fonctionner à des échelles différentes, en d'autres termes quand il est apte à maîtriser une charge de calcul et/ou de données croissante.

Comme le souligne TANENBAUM et VAN STEEN (2002), un réseau étendu (WAN) offre une certaine ouverture et une extensibilité par rapport à une architecture fortement couplée, en contre partie la transparence tend à diminuer avec l'accroissement du nombre de participants. Ainsi, l'un des défis de l'informatique distribuée est d'intégrer des composants hétérogènes et appartenant à des domaines administratifs différents (tels que des ordinateurs monoprocesseurs, des fermes de calcul ou des clusters) dans un système global, appréhendé comme homogène par l'utilisateur final. La technologie grille, approfondie dans les paragraphes suivants, tend précisément à atteindre cet objectif.

Au travers de cet exposé, il s'agit de souligner les grandes lignes d'un système distribué de manière à appréhender au mieux le fonctionnement d'une grille informatique. Ainsi, loin d'être exhaustive, cette

étude cherche à relever les composantes tant techniques qu'organisationnelles d'un système distribué de manière à les confronter, plus loin dans cet exposé, aux caractéristiques fonctionnelles et aux besoins engendrés par une situation de gestion de crise.

II - 2.2. Les grands modèles de communication dans un système distribué

Le fonctionnement d'un système distribué s'appuie sur des modèles de communication standards qu'il convient de synthétiser afin de mieux appréhender les subtilités fonctionnelles d'une architecture de grille.

II - 2.2.1. Client-serveur

Le modèle client-serveur représente le modèle de communication le plus utilisé dans le monde des systèmes distribués. Il s'agit de permettre la communication entre un client et un serveur. Ce dernier offre une panoplie de service, comme l'accès à des fichiers de données, à des bases de données ou à des opérations informatiques. La mise en place d'une telle communication s'effectue généralement à l'aide des protocoles standards TCP/IP de l'Internet, permettant une connexion directe et univoque entre deux postes informatiques. Le développement de ce modèle implique une organisation en tiers. Il s'agit de clairement séparer les développements et *a fortiori* les opérations de chaque acteur. Dans ce cas, deux modèles principaux existent :

- l'architecture 2-tiers, qui considère le client et le serveur comme deux entités fonctionnelles distinctes
- l'architecture 3-tiers, qui considère trois niveaux de communication (Figure 18) :
    - le client et son interface utilisateur
    - le serveur d'application qui contient les fonctionnalités de communication avec les deux autres tiers et les programmes spécifiques à l'application
    - le serveur de données qui interagit avec le serveur d'application en interprétant ses requêtes et en lui transmettant les données correspondantes

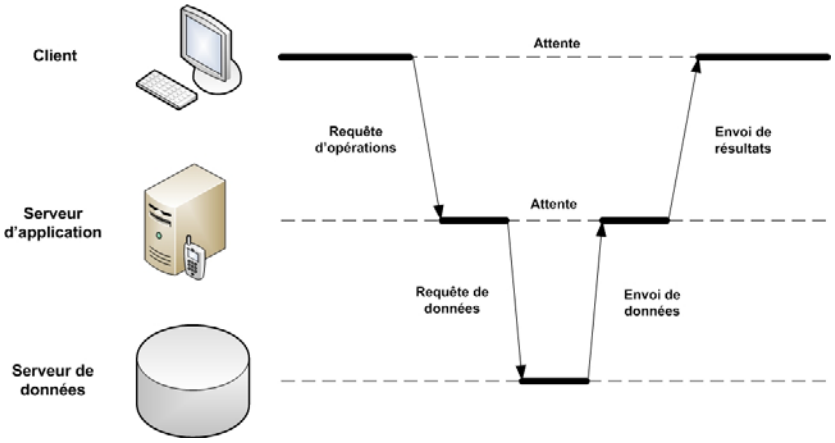


Figure 18 : Architecture de communication client-serveur 3 tiers
(d'après TANENBAUM et VAN STEEN, 2002)

Ce type de communication est largement utilisé à l'heure actuelle dans des réseaux locaux ou pour le fonctionnement d'un site web classique.

II - 2.2.2. Pair-à-pair

Ce modèle de communication est une évolution du modèle client-serveur. Les postes informatiques qui participent à la communication sont à tour de rôle client et serveur (SCHULER, 2005). Les deux (voire les plus) participants exécutent le même programme permettant de créer une session de communication et offrant la double fonctionnalité de client et de serveur. Chaque participant peut faire une requête à un autre client (serveur) et vice et versa. Ce type de communication se démarque de la précédente par son aspect distribué. En effet, ce modèle de communication ne nécessite pas de serveur central en charge de l'ordonnancement des opérations ou de l'accès à des données distantes. Les postes informatiques sont directement interconnectés entre eux suivant un schéma deux à deux, grâce aux mêmes standards TCP/IP que la communication client-serveur.

Le succès de ce mode de communication se retrouve dans les systèmes d'échange de fichiers multimédias, tels que Napster ou eMule. L'utilisateur télécharge un fichier sur plusieurs machines tout en fournissant la partie du fichier qu'il a déjà en sa possession. De cette manière, les systèmes pair-à-pair permettent de distribuer les données sur plusieurs nœuds indépendants du réseau. De la même manière, comme cela sera abordé dans les paragraphes suivants, les nœuds présents dans ce type de réseau peuvent mettre à disposition leur puissance de calcul.

A grande échelle, un tel modèle de communication n'est cependant efficace que par l'ajout de serveurs centraux permettant le recensement et l'indexation des ressources et des utilisateurs (MINAR et HEDLUND, 2001), rendant le modèle quelque peu hybride entre le modèle client-serveur et le pair-à-pair. Ainsi, si à l'origine ce modèle était principalement conçu pour l'échange de fichiers distants, ce modèle s'est progressivement complexifié pour parvenir à « une fourniture de services multiutilisateurs relativement proches de ceux fournis par les GRID[2] » (SOBERMAN, 2003).

II - 2.2.3. L'intergiciel

Un intergiciel utilise schématiquement les modèles de communication client-serveur et pair-à-pair. Une des réponses au développement d'un système distribué grande échelle, tel que la grille, transparent, ouvert et extensible passe par l'utilisation d'un intergiciel. L'intergiciel ou logiciel médiateur correspond à :

« un logiciel qui permet le fonctionnement de plusieurs ordinateurs en coordination, en attribuant à chacun une tâche spécifique, comme les échanges avec les utilisateurs, l'accès aux données ou aux réseaux »[3]

Les définitions sont multiples, souvent reliées à la finalité fonctionnelle de l'intergiciel. D'un point de vue informatique, KSHEMKALYANI et SINGHAL (2008) explique qu' « un système distribué utilise une architec-

[2] L'Action Concertée Incitative (ACI) du ministère de la recherche a créé l'acronyme GRID pour « Globalisation des Ressources Informatiques et des Données » (http://www-sop.inria.fr/aci/grid/public/)
[3] Définition de la délégation générale à la langue française et aux langues (http://www.culture.gouv.fr/culture/dglf/)

ture par couches pour décomposer la complexité de la conception du système. L'intergiciel est un logiciel distribué qui pilote le système distribué tout en fournissant une transparence de l'hétérogénéité au niveau de la plateforme ». TANENBAUM et VAN STEEN (2002) considèrent l'intergiciel simplement comme une couche de logiciels qui permet de « plus ou moins cacher l'hétérogénéité de l'ensemble des plateformes sous-jacentes mais aussi pour améliorer la transparence de la distribution ». Structurellement, il s'agit du « niveau intermédiaire installé au-dessus des systèmes d'exploitation et des protocoles de communication » (KRAKOWIAK, 2008). Enfin, JEANVOINE (2007) en considérant l'intergiciel comme « une couche logicielle qui est positionnée entre le système d'exploitation des ressources et les applications des utilisateurs » complète l'ensemble de ces points de vue en intégrant la communication de l'intergiciel avec les postes utilisateurs.

D'un point de vue pratique, souvent de finalités différentes (JEANVOINE, 2007), l'intergiciel se comporte comme un système de médiation, comme mandataire ou comme interface avec les composants logiciels (KRAKOWIAK, 2008). Il cache la répartition géographique des entités physique du système et l'hétérogénéité de ces dernières ainsi que celle de leurs systèmes d'exploitation.

Le développement d'un intergiciel implique avant tout une standardisation des protocoles de communication de manière à interfacer efficacement les services réseau de chaque nœud et à permettre l'extensibilité de l'architecture. Par exemple, le Remote Procedure Call (RPC) est un standard permettant d'échanger des messages entre machines distantes. Enfin, en cachant les interactions avec le matériel, un intergiciel a un niveau d'abstraction élevé en assurant une certaine « indépendance vis-à-vis des langages et des plates-formes » (KRAKOWIAK, 2008) et assure une vue globale du système distribué (TANENBAUM et VAN STEEN, 2002).

II - 2.3. La technologie grille : définitions et fonctionnement

II - 2.3.1. Historique et définitions

Les avancées informatiques en termes de gestion et d'intégration de ressources distantes, illustrées par le développement d'intergiciels autonomes et standardisés permet progressivement de répondre aux nouveaux enjeux de l'informatique distribuée (cf. §.II - 2.1.3). On parle communément d'infrastructure de grille ou de technologie grille.

II - 2.3.1.1. Grille de 1ère génération

D'après BERMAN (2003), le premier projet de grande ampleur destiné au développement et à la mise en production d'une infrastructure de grille est I-WAY. Il correspond à la première génération de grilles, nommé metacomputing, dont l'objectif était de fournir « des ressources de calcul pour un ensemble d'applications haute-performance » (DE ROURE et *al.*, 2003). Au travers de cette approche, il s'agissait avant tout de fournir une architecture adaptée au calcul parallèle en connectant « des superordinateurs, des systèmes de stockage de masse et des dispositifs de visualisation évolués » (FOSTER et *al.*, 1997, VICAT-BLANC PRIMET et *al.*, 2003). Il s'agissait donc de grille de calcul définit comme une « infrastructure matérielle et logicielle qui fournit un accès fiable, cohérent, omniprésent et bon marché à des capacités de calcul sophistiquées » (FOSTER et KESSELMAN, 1999).

Cette première génération a donc permis de développer des solutions informatiques offrant des capacités de calcul parallèle pour la prise en charge de problèmes complexes. Vers la fin des années 90, d'autres besoins concernant la gestion de données massives ont émergé. Il est question de calcul intensif où un même algorithme peut être répliqué à « l'infini » pour analyser des jeux de données différents.

II - 2.3.1.2. Grille de 2ème génération

Ainsi, la technologie grille s'oriente vers une architecture globale intégrant des systèmes informatiques géographiquement distants et accessibles au travers d'un système distribué unique. Les auteurs abordent le terme d'omniprésence de la grille (DE ROURE et *al*., 2003). Cette nouvelle génération a permis l'intégration de systèmes d'information hétérogènes tout en garantissant progressivement l'intégration de nouvelles ressources, telles que des systèmes d'observation, et la gestion dynamique de ces dernières (Figure 19).

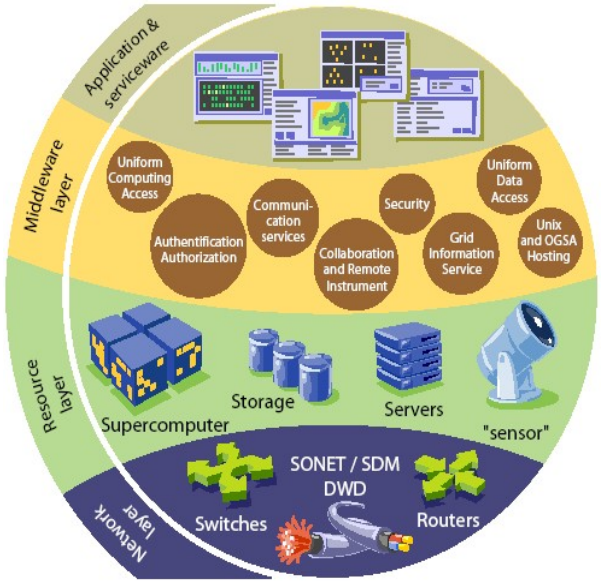


Figure 19 : Architecture globale de la technologie grille, les couches correspondent de bas en haut aux ressources réseau, matérielles, de l'intergiciel et applicatives (source : Grid Café : http://www.gridcafe.org/)

L'interopérabilité est le maître mot de cette génération. Elle a donné naissance à un intergiciel de haut niveau, la boîte à outil Globus, offrant une panoplie de services standardisés, comme l'authentification des participants, la découverte de ressources et l'accès à ces dernières (FOSTER et KESSELMAN, 1997).

II - 2.3.1.3. Grille de 3ème génération

Finalement, la dernière génération de grille, celle utilisée dans ce travail de recherche, s'appuie sur les avancées des deux premières générations et correspond à une Architecture Orientée Service (SOA). L'approche par services apporte une forte flexibilité dans l'utilisation des ressources distantes, et d'autre part permet l'automatisation des tâches de gestion et d'administration (DE ROURE et *al*., 2003). Une architecture de grille est alors appréhendée comme un système à haut niveau d'abstraction regroupant plusieurs communautés d'utilisateurs au travers d'une infrastructure informatique globale et fortement distribuée. Dans ce type d'architecture, l'utilisation de métadonnées, « ensembles de don-

nées structurées décrivant des ressources physiques ou numériques […] essentiel pour l'interopérabilité de l'information et sa gestion » (MOREL-PAIR, 2005), est à la base de l'harmonisation et du partage des ressources. Techniquement, cette dernière génération se base sur la norme OGSA[4] (Open Grid Services Architecture) qui définit des mécanismes standards de fonctionnement de la grille (FOSTER et *al*., 2002), où chaque composant de l'infrastructure est appréhendée comme un service de grille (Grid service), facilitant ainsi l'ouverture, l'extensibilité et la transparence du système.

La grille de « 3ème génération » s'appuie sur le concept d'organisation virtuelle (VO), déjà définie précédemment (cf. §.II - 1.2). Si les deux premières générations se sont attachées à construire une architecture informatique grande échelle, FOSTER et *al*. (2001) introduit la 3ème génération comme étant relative au « grid problem », c'est-à-dire parvenir à « un partage de ressources coordonné et la résolution de problèmes dans des organisations virtuelles dynamiques et multi-institutionnelles » (Figure 20). Ainsi, cette génération soulève la problématique du regroupement de multiples utilisateurs et/ou organisations administrativement indépendants autour d'une architecture de grille intégrant un nombre important de ressources informatiques géographiquement distantes.

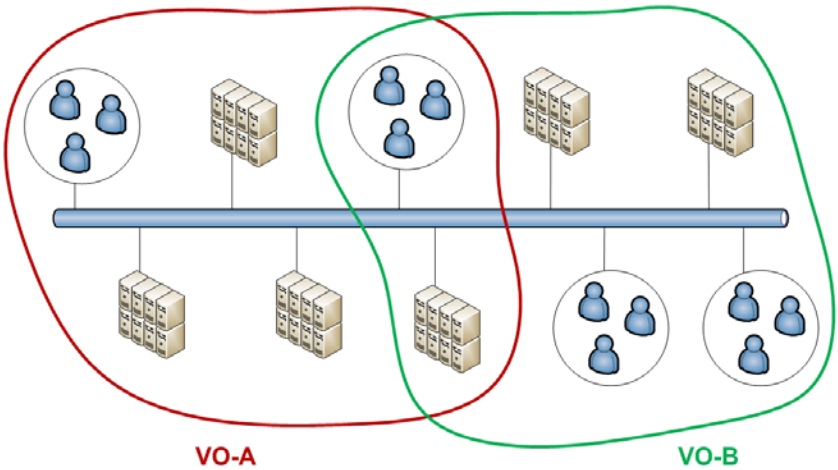


Figure 20 : Relation entre organisation virtuelle, ressources matérielles et utilisateurs : le « grid problem »

II - 2.3.1.4. L'organisation virtuelle garante de l'approche collaborative

Les questions gravitant autour du « grid problem » concernent la résolution de problèmes tels que :

- La virtualisation des ressources pour permettre leur accès transparent par les utilisateurs autorisés (MAGOULES et *al*., 2009),
- La flexibilité et la coordination dans l'utilisation des ressources en facilitant leur accès à la demande (FOSTER et *al*., 2001),
- La sécurisation des fermes des calculs intégrées dans l'organisation virtuelle ainsi que des opérations effectuées par l'ensemble des utilisateurs (intrusion).

[4] http://www.globus.org/ogsa/

Cette approche repose donc sur le regroupement de communautés d'utilisateurs et de fournisseurs de ressources informatiques (fermes de calcul) au sein d'un environnement collaboratif régi par l'intergiciel de grille permettant la virtualisation de ces ressources. Cette virtualisation, au sens de système unique, doit être basé sur une notion de confiance mutuelle entre l'ensemble des partenaires entre utilisateurs et fournisseurs, et entre fournisseurs eux-mêmes), qui s'engage à respecter une politique d'utilisation préalablement définie (BROOKE et PARKIN, 2009 ; NAQVI, 2005).

La problématique de sécurité est donc d'une importance capitale du fait du nombre important de systèmes informatiques gravitant autour d'une architecture de grille. Ainsi, l'organisation virtuelle permet « d'assurer un certain niveau de confiance à tous ses membres » (CONTES, 2005). Grâce à des mécanismes d'authentification tels que la cryptographie et les courtiers de sécurité (proxy) et d'autorisation comme la reconnaissance mutuelle d'un certain niveau de confiance défini par le « contrat » d'adhésion à l'organisation virtuelle, une relative sécurité est garantie dans une architecture de grille (KNOOPS, 2007).

Finalement, la définition de la technologie grille retenue dans le cadre de cette recherche est (BERMAN et *al*., 2003 ; CHERVENAK et *al*., 2000 ; FOSTER et *al*., 2002 ; FOSTER et KESSELMAN, 2004 ; SOBERMAN, 2003) :

Une architecture informatique distribuée qui consiste à la mise en connexion sécurisée de fermes de calcul et/ou de clusters géographiquement distants offrant, par leur association et leur fonctionnement collaboratif, des capacités de calcul et de stockage à la demande pour un utilisateur inscrit dans une organisation virtuelle l'autorisant à accéder à ces ressources.

L'analyse de ce glissement technologique et sémantique que le concept de grille a connu permet de mieux entrevoir les multiples facettes que cette technologie peut fournir à des communautés scientifiques, à la recherche informatique ou à des entreprises (RICHARD et *al*., 2008). La grille évolue quotidiennement, cependant, de nombreux obstacles technologiques subsistent, en effet l'extension à grande échelle d'une architecture distribuée ne rime pas pleinement avec l'intégration de ressources fortement hétérogènes. Dans ce sens, des contraintes fortes telles que l'harmonisation des systèmes d'exploitation des nœuds de calcul subsistent.

II - 2.3.2. Les 5 couches fonctionnelles de la technologie grille

Le modèle OSI (Open Systems Interconnection) est le modèle de communication entre ordinateurs défini par la norme ISO[5]. Il permet de décrire et de mettre en relation l'ensemble des protocoles informatiques, de la couche matérielle (Physique) au navigateur web (Application), utilisés lors d'une communication réseau (Figure 21). Par exemple, Internet est basé sur les protocoles TCP/IP qui se situent au niveau des couches Réseau (IPv4) et Transport (TCP et UDP). Au niveau supérieur (Application), le web s'appuie sur l'ensemble des couches inférieures pour prendre en charge les requêtes des utilisateurs basées sur le protocole HTTP. Une architecture de grille s'appuie sur ces 7 couches pour intégrer et mettre en communication l'ensemble des ressources matérielles (Site). Le modèle théorique de FOS-

[5] ISO/IEC standard 7498-1:1994 « Information technology - Open Systems Interconnection - Basic Reference Model: The basic model »

TER et *al*. (2001) fait office de référence dans la communauté scientifique. Il est constitué de 5 couches et est largement repris dans les paragraphes qui suivent (Figure 21).

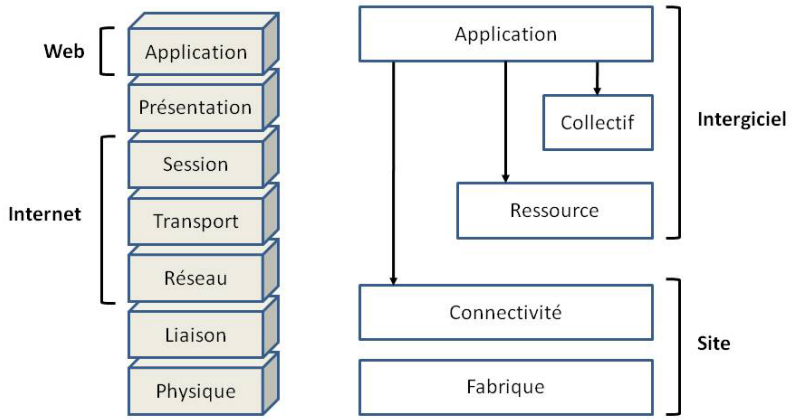


Figure 21 : Relations entre le modèle OSI et les couches d'une architecture de grille
(d'après FOSTER et *al*., 2001)

Il s'agit de présenter succinctement ces 5 couches de protocoles.

II - 2.3.2.1. la couche « Fabrique »

Elle regroupe l'ensemble des ressources matérielles fournis par chaque site de calcul :

- les ressources de calcul sur lesquelles les algorithmes sont exécutés,
- les ressources de stockage enrichies de protocoles de transfert spécifiques,
- les ressources « réseau » qui permettent de relier les deux précédentes au sein d'un site et les sites entre eux,
- les gestionnaires de versions de ressources,
- les catalogues des ressources matérielles, logicielles et de données.

II - 2.3.2.2. la couche « Connectivité »

Cette couche regroupe l'ensemble des protocoles de communication et d'authentification nécessaires à un fonctionnement transparent et sécurisé caractérisant une architecture de grille. Généralement, elle utilise les protocoles TCP/IP de l'Internet. Une des solutions originale de la technologie grille concerne le principe « single sign on » qui permet à un utilisateur de s'identifier une seule fois à l'ouverture de sa session alors qu'un système de délégation permet d'assurer l'authentification auprès des ressources distantes utilisées (couche « Fabrique ») pour l'ensemble des traitements requis.

II - 2.3.2.3. la couche « Ressource »

Cette couche s'appuie sur les deux couches précédentes pour implémenter des protocoles permettant « la négociation, l'initialisation, la gestion, le contrôle, la comptabilité et le paiement » des opérations sur les ressources matérielles partagées de la couche « Fabrique ». Il existe deux grands types de protocoles :

- les protocoles d'information qui décrivent en temps-réel l'état des ressources de la grille,
- les protocoles de gestion qui sont utilisés pour « négocier l'accès aux ressources partagés » en fonction des requêtes utilisateurs.

II - 2.3.2.4. la couche « Collectif »

Elle permet d'intégrer et de virtualiser l'ensemble des ressources matérielles dans un système d'information unique, en fournissant des fonctionnalités de haut niveau permettant de gérer l'ensemble des interactions entre l'utilisateur et les ressources de calcul et de stockage. Il est possible d'identifier la gestion :

- des annuaires des ressources
- de l'ordonnancement des tâches à exécuter
- des diagnostiques d'utilisation des ressources
- des réplications de données et d'algorithmes
- de fonctionnalités avancées comme des environnements collaboratifs, des services de paiements, etc.

II - 2.3.2.5. la couche « Application »

Elle représente la couche de plus haut niveau (équivalente à celle du modèle OSI) et s'appuie sur les protocoles des couches « Connectivité », « Ressource » et « Collectif » sous-jacentes. Elle permet à l'utilisateur d'une organisation virtuelle d'interagir de manière simplifiée avec l'ensemble des services de la grille implémentés dans les couches inférieures.

De manière schématique, les services de l'intergiciel de grille sont inclus dans les 3 couches supérieures, alors que les protocoles de plus bas niveau, permettant l'interaction des ressources matérielles sont gérés à l'échelle de chaque site de calcul (Figure 21). Au travers de cette description, il est important de relever l'aspect prépondérant de l'intergiciel de grille qui offre à l'utilisateur des facilités fonctionnelles pour les tâches informatiques qu'il exécute sur la grille, et qui gère les interactions de l'ensemble de ses services de manière transparente et automatique.

Dans cette étude, l'ensemble de ces concepts sont abordés et expérimentés au travers d'un point de vue opérationnel. Les infrastructures et les projets de grille sont nombreux, leurs architectures varient selon les attentes et les besoins technologiques des utilisateurs finaux. Le paragraphe suivant s'attache à présenter les types de grille les plus courants et quelques initiatives récentes afin d'entrevoir les solutions technologiques retenues pour répondre à la problématique de cette recherche.

II - 2.3.3. Une typologie des infrastructures de grille

Suivant les classifications, il existe plusieurs types de grille :

- les grilles de PC ou « Public-Resource Computing » partent du principe que les ordinateurs individuels sont de manière générale sous-utilisés et disposent ainsi des ressources informa-

tiques disponibles. Ce type de grille s'appuie par exemple sur l'intergiciel BOINC[6] (Berkeley Open Infrastructure for Network Computing) (ANDERSON, 2004). Des projets comme SETI@home[7] (ANDERSON et *al*., 2002), Folding@home[8] (SHIRTS et PANDE, 2001) et climate*prediction*.net[9] (STAINFORTH et *al*., 2004) regroupent plusieurs milliers de postes individuels en vue de traiter des données massives. Par exemple, SETI@home a totalisé plus de 400.000 ans de temps de calcul en 3 ans en fournissant à des ordinateurs individuels distant, un programme et des données à traiter, grâce à la technologie peer-to-peer (ANDERSON et *al*., 2002). Les programmes Folding@home et climate*prediction*.net se basent sur le même principe technologique pour respectivement modéliser un grand nombre de scénarios de repliement de protéine et d'évolution du climat jusqu'en 2100. Plus récemment, le projet EDGeS (Enabling Desktop Grids for e-Science) cherche à connecter une grille de PC, composée de 100.000 unités, avec l'architecture de grille grand échelle EGEE (Enabling Grids for E-sciencE), sur laquelle la méthodologie de cette recherche est basée (URBAH et *al*., 2009).

- les grilles de sites mutualisent des ressources de calcul, de stockage et de mesure appartenant à des sites administrativement différents. Ces sites sont généralement des clusters ou des fermes de calcul avec une forte stabilité de leurs ressources (CAPPELLO, 2006). En France, elles ont vu le jour en 2001 sous l'acronyme GRID (SOBERMAN, 2003) (cf. §.II - 2.2.2). La majorité de ces grilles ont un accès contrôlé basé sur des autorisations et des relations de confiance entre les fournisseurs et les utilisateurs. C'est dans ce contexte que l'organisation virtuelle prend tout son sens. Ce type de grilles peut être scindé en plusieurs sous-types suivant la finalité de leur usage :

    - les grilles de calcul dont l'objectif est d'intégrer des fermes de serveurs de manière à obtenir des calculs distribués à grand échelle. L'objectif est de disposer de capacités calculatoires très importantes afin de résoudre des problèmes que même les superordinateurs ne parviennent pas à prendre en charge. (SOBERMAN, 2003)

    - les grilles de données si les capacités de calcul restent nécessaires à la gestion et à l'analyse de données, ce type de grille cherche à intégrer des bases de données ou des fichiers volumineux dans un système unique accessible par les utilisateurs de manière virtuelle. Elles s'appuient sur des fonctionnalités de découverte d'information et d'établissement de schémas de connaissance permettant une gestion « intelligente » de données. Le projet européen *Earth Science Grid*[10], en cours de réalisation, est une illustration de ce type de grille.

La tendance actuelle des grilles de sites, comme cela est démontrée dans les paragraphes qui suivent, est de tendre vers des grilles collaboratives fondées sur « l'établissement d'un environnement virtuel qui permet à des individus ou groupes de personnes géographiquement dispersés de coopérer dans la

[6] http://boinc.berkeley.edu/
[7] http://setiathome.berkeley.edu/
[8] http://folding.stanford.edu/
[9] http://climateprediction.net/
[10] www.euearthsciencegrid.org

poursuite d'un objectif commun » (BRETON et *al*., 2005). L'infrastructure matérielle, composée de fermes de calcul, de stockages et d'instruments de mesure, est assez hétérogène mais permet de regrouper par exemple des laboratoires autour d'un projet de recherche commun.

A terme, l'objectif ultime est le regroupement de l'ensemble des architectures de grille mondiales sous l'égide d'une architecture unique et globale de plus grande ampleur. Il est souvent question de *World Wide Grid* (SOBERMAN, 2003 ; VICAT-BLANC PRIMET et *al*., 2003).

II - 2.3.4. Des grilles d'ampleur nationale

La distinction entre ces différents types de grille varie suivant les auteurs. Il existe par exemple des grilles d'entreprise (RICHARD et *al*., 2008) ou des portails scientifiques permettant « l'accès à des capacités de simulation » (BAXEVANIDIS et *al*., 2002). Il est à retenir que les grilles permettent de fédérer des utilisateurs et leurs ressources. Ces dernières pouvant être des processeurs (ou des ordinateurs), des systèmes de stockage, et des systèmes d'observation et de mesure, l'ensemble étant connecté par des réseaux aux débits importants (de quelques MB/s à plusieurs GB/s).

II - 2.3.4.1. GRIDA3

Il s'agit d'une grille installée en Sardaigne et gérée par le laboratoire CRS4. Son objectif est de maintenir une organisation virtuelle orientée vers la surveillance et la gestion de problématiques environnementales et basée sur un système d'aide à la décision collaboratif *(*MURGIA et *al., 2009)*.

Cette grille regroupe plusieurs domaines administratifs publics et privés sardes. D'un point de vue matériel, elle compte 32 nœuds de calcul permettant de disposer d'une puissance de calcul cumulée de 1,1 Tflops et connectés entre eux par un réseau de 2Gbps. Les services de l'intergiciel commercial EnginFrame[11] permettent aux utilisateurs d'accéder aux ressources matérielles virtualisées au travers d'un portail web. Cet intergiciel s'appuie sur une architecture 3 tiers composée du tiers client (portail web), d'un tiers serveur basé sur l'approche client-serveur et contrôlant l'ensemble des services d'utilisation de la grille (authentification, gestion des données, gestion des jobs, interactions de serveurs, etc.) et d'un 3ème tiers de gestion des ressources matérielles par des agents informatiques autonomes.

Le développement de cette grille répond avant tout à une logique applicative. L'intergiciel et les services de gestion des ressources matérielles et logicielles sont principalement basés sur les spécifications de ces applications. L'objectif est de mutualiser des données environnementales afin les intégrer dans différents modèles de simulation et établir un environnement collaboratif d'aide à la décision environnementale.

II - 2.3.4.2. GRID'5000

Cette grille française étendue sur l'ensemble du territoire métropolitain regroupe 9 centres de calcul, soit approximativement 5.000 processeurs (CAPPELLO, 2006). Cette grille pair-à-pair est dédiée à la recherche en informatique et plus particulièrement sur l'informatique distribuée. « Cette plateforme peut être définie comme un système distribué expérimental grande échelle hautement reconfigurable, con-

[11] http://nice-software.com/web/nice/products/enginframe

trôlable et maîtrisable » (ORGERIE et LEFEVRE, 2009). L'aspect « recherche » de cette plateforme offre des services spécifiques aux utilisateurs comme la réservation en avance de ressources de calcul, ainsi que le choix du système d'exploitation et l'installation de « piles logicielles » spécifiques pour mettre en place leurs expérimentations informatiques.

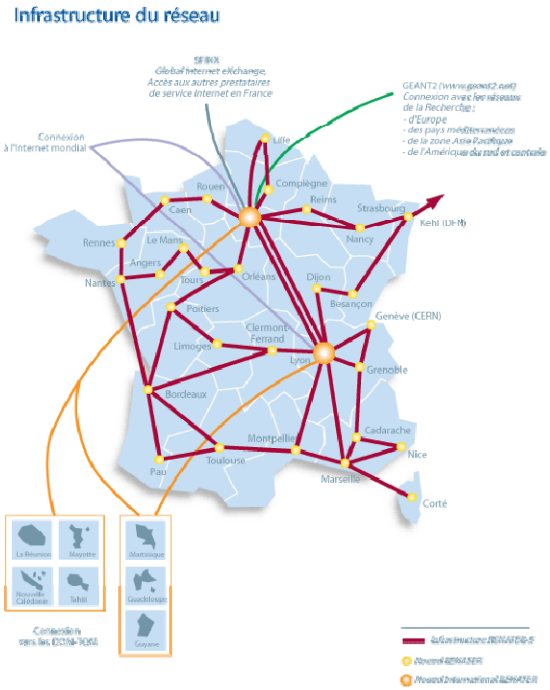


Figure 22 : Infrastructure du réseau RENATER
(source : http://www.renater.fr/)

D'un point de vue matériel, l'ensemble des nœuds de cette grille sont relativement homogènes de manière à faciliter son administration et la reproductibilité des expérimentations (D'ANFRAY et JEANNOT, 2007). Ils sont interconnectés à 10 Gbps au moyen du réseau national français RENATER (Réseau National de télécommunications pour la Technologie l'Enseignement et la Recherche) (LE GUIGNER, 1994) (Figure 22), séparé du réseau Internet classique afin de faciliter la gestion de la sécurité des ressources. D'un point de vue typologique, cette grille peut être rattachée au champ du métacomputing et cherche à développer des algorithmes spécifiques pour la gestion de ressources informatiques distantes.

II - 2.3.5. Le projet EGEE : une révolution technologique, scientifique et politique

II - 2.3.5.1. La vision E-Science

Les deux précédentes grilles GRIDA3 et GRID'5000 ont en commun une relative uniformité de leurs champs d'applications. Cependant, au regard de la définition de Foster, elles n'offrent a priori pas un niveau d'abstraction suffisamment élevé pour permettre la résolution de problématiques indépendantes, nécessitant chacune l'intégration de ressources (au sens le plus large) pluri-thématiques et administrativement indépendantes.

Au début des années 2000, John Taylor aborde le terme de « e-science » relatif à « la collaboration globale dans les secteurs clés de la science et de la prochaine génération d'infrastructures qui l'a rendront possible » (HEY et TREFETHEN, 2002). Par cette phrase, Taylor, alors directeur général des conseils de recherche au Royaume-Uni, soulignait la nécessité d'établir des collaborations multidisciplinaires fortes basées sur des solutions technologiques communes dans les domaines de la recherche. MINETER et *al*. (2003) reprend le même terme pour décrire une infrastructure scientifique entreprise « par des organisations virtuelles de groupes étendus de personnes, avec en général des jeux de données énormes et des ressources de calcul haute-performance distribués ». D'après GURNEY et *al*. (2009), elle correspond aux challenges actuels de la société moderne, tels que la modélisation météorologique et du climat, les problématiques de la biosphère ou encore la chimie moléculaire. JONES (2009) préfère le terme de « e-infrastructure » capable d'intégrer l'ensemble des recherches actuelles autour d'une architecture informatique commune et d'harmoniser les solutions technologiques offertes aux chercheurs. Cette nouvelle philosophie, rendue possible par les récents progrès technologiques, fait implicitement référence à la technologie grille. D'après la Figure 23, une telle révolution impose le glissement d'infrastructures distinctes et cloisonnées à une infrastructure globale et interopérable permettant de fédérer des communautés aux problématiques communes par le biais d'organisations virtuelles thématiquement dédiées.

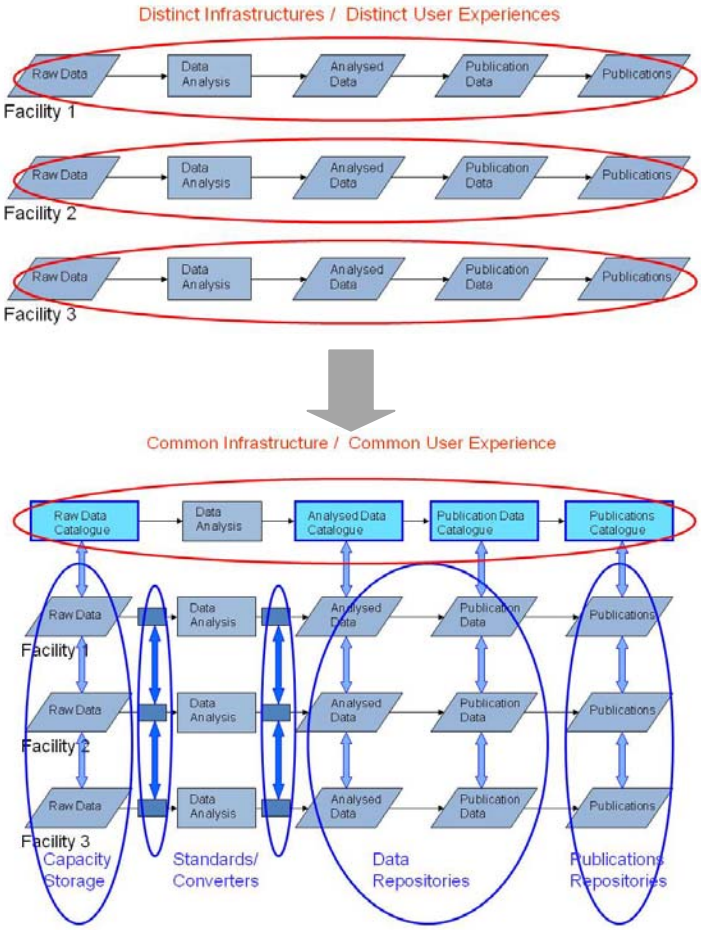


Figure 23 : Évolutions technologiques pour le développement d'une e-Infrastructure (d'après JONES, 2009)

Cette philosophie peut être apparentée au « grid problem » de Foster et comme le confirme HEY et TREFETHEN (2002) l'infrastructure de Taylor tend à correspondre implicitement à la technologie grille. Ainsi, des projets tels qu'e-Science au Royaume-Uni (HEY et TREFETHEN, 2002) et EGEE (Enabling Grids for E-sciencE) (GAGLIARDI et *al*., 2005) en Europe cherchent encore à l'heure actuelle à répondre à ce nouveau postulat.

II - 2.3.5.2. La grille EGEE

La grille EGEE, sur laquelle les développements méthodologiques de cette recherche sont fondés, résulte à l'origine de l'intégration internationale de différentes équipes scientifiques gravitant autour de la physique des particules, et en particulier rassemblées autour de données massives devant être produites par l'accélérateur de particules Large Hadron Collider (LHC) à Genève (BRITTON et *al*., 2009). Le projet initial European DataGrid (EDG) (KUNSZT, 2003), initié en 2000 par le CERN (Organisation européenne pour la recherche nucléaire) et financé par l'Union Européenne, était une étude de faisabilité et a permis de développer une architecture de grille capable de répondre aux fortes contraintes de réplication de ces données afin de les rendre accessible à l'ensemble des partenaires géographiquement dispersés (SEGAL et *al*., 2000). La pérennité progressive de ce projet pilote a donné le jour à la grille EGEE, un des systèmes distribués les plus importants au monde apte à englober des communautés scientifiques de plus en plus diversifiées (IDG, 2009; GAGLIARDI et *al*., 2005).

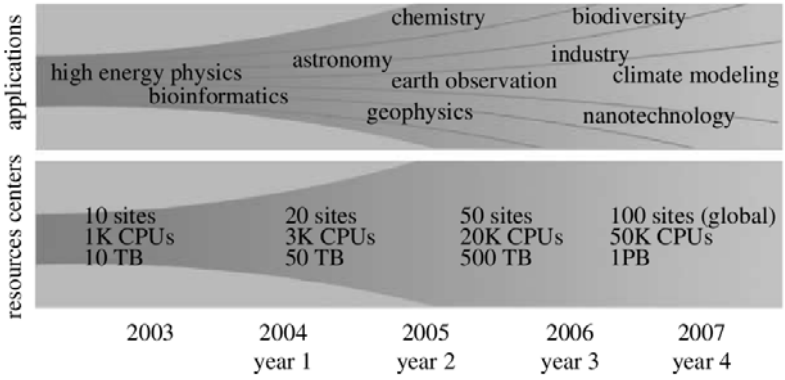


Figure 24 : Évolution de l'infrastructure d'EGEE (d'après GAGLIARDI et *al*., 2005)

Après trois phases successives, l'implication de ces communautés et leur collaboration avec les partenaires informatiques impliqués dans le développement de cette grille ont progressivement porté en production cette architecture devenue mature (Figure 24). Ainsi, la grille EGEE s'articule entre la fourniture intensive de ressources informatiques aux différentes communautés scientifiques et les recherches informatiques dédiées à son perfectionnement et à sa mise en production. Par les multiples domaines scientifiques qu'elle intègre, cette architecture offre un socle d'innovations technologiques et thématiques (EGEE, 2009). En octobre 2009, le site internet officiel du projet[12] recensait plus de 260 sites de calcul répartis sur une cinquantaine de pays (majoritairement européens). Au total, cette infrastructure regroupe plus de 150.000 unités de calcul et 28 petabytes d'espace de stockage, offerts à environ 140 organisations virtuelles regroupant pas moins de 10.000 utilisateurs dans plusieurs domaines scientifiques.

[12] http://project.eu-egee.org

D'un point de vue technologique, le cœur de l'infrastructure EGEE s'appuie sur un réseau de centres de calcul européens interconnectés par le réseau haut débit GEANT2 (Figure 25) (PATIL et *al*., 2006). Ce réseau transnational est relié à des réseaux haut débit nationaux, comme RENATER pour la France. Ces réseaux, qui permettent à des communautés académiques et scientifiques européennes de communiquer et de connecter des systèmes informatiques disparates, offre un débit de l'ordre du Gbps au sein de la grille et de l'ordre du Mbps pour les utilisateurs externes[13].

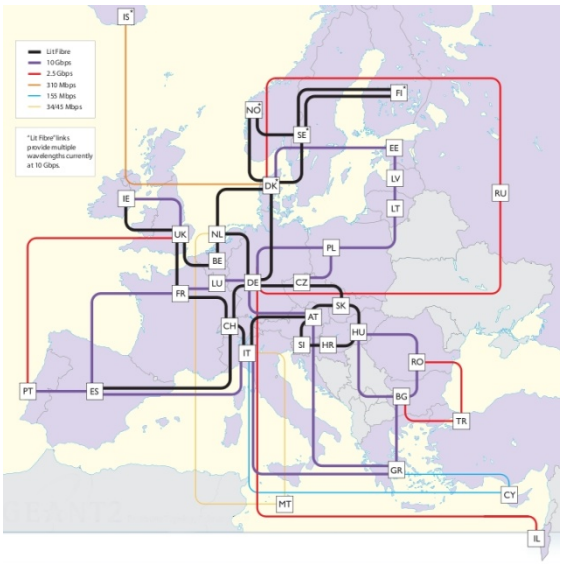


Figure 25 : Caractéristiques de l'infrastructure réseau de GEANT2 (d'après PATIL et *al*., 2006)

Dans ce sens l'évolution du projet EGEE envisage la création d'une « e-Infrastructure » : l'*European Grid Initiative* (EGI), intégrant l'ensemble des recherches actuelles autour d'une architecture informatique commune et permettant d'harmoniser les solutions technologiques offertes aux chercheurs (JONES, 2009). Cependant, le développement de telles infrastructures implique des enjeux forts de gouvernance entre les différents participants et les pays concernés, qui limitent encore sa pleine réalisation (JEFFERY, 2007).

Il convient maintenant de présenter les deux principales organisations virtuelles qui constituent l'architecture EGEE.

II - 2.3.5.2.1 *La physique des particules et la grille W-LCG*

Sans le développement du LHC à Genève, la grille EGEE n'aurait sans doute jamais vu le jour à cette échelle. L'accélérateur de particules situé 150 m sous terre dans un tunnel de 27 km de circonférence ouvre de nouvelles perspectives dans la compréhension de l'univers, de ses constituants et de sa formation. Les expériences de physique des hautes énergies qui s'y dérouleront ont pour objectif de produire des données afin « de réaliser des études de simulation massive des « événements » qui seront produits quand les faisceaux de haute énergie composés de protons ou d'ions lourds se heurtent »

[13] Ce débit est fortement variable suivant la localisation géographique de l'utilisateur

(lettre d'information sur EGEE[14]). Ainsi, 4 projets principaux ont été développés autour de la grille de calcul W-LCG[15] (World Wide LHC Computing and Grid) pour mener à bien les expérimentations :

- ATLAS (A Toroidal LHC ApparatuS)
- CMS (Compact Muon Solenoid)
- ALICE (A Large Ion Collider Experiment)
- LHCb (Large Hadron Collider beauty)

Ces 4 expériences menées sont génératrices de données réelles et simulées massives de l'ordre du pétabyte (ESPINAL et *al*., 2008) et nécessitent une infrastructure de grille conséquente permettant de regrouper une dizaine de milliers de chercheurs appartenant à des centaines de laboratoires dispersés sur 3 continents. Cette architecture est structurée et hiérarchisée en tiers (AVERY, 2009). La production des données brutes a lieu au niveau des 4 expériences déjà citées, il est question du Tier0. Une partie des données sont transmises au Tier1 constitué de 11 laboratoires nationaux (par exemple le CC-IN2P3[16] de Villeurbanne) qui « assurent la reconstruction et la pérennisation » de ces données (MALEK, 2009). Plus de 150 laboratoires constituent le Tier2 et « assurent la production d'événements simulés et participent aux tâches d'analyses centralisées et stochastiques » sur la base des données qui leur sont transmises par le Tier1. Enfin, le Tier3, sans engagement direct dans le projet W-LCG, peuvent suppléer les traitements effectués au niveau du Tier2.

La mise en production de ces expériences a représenté un défi informatique majeur. Ainsi, depuis plus de 2 ans, des opérations « d'épreuve par les données » ou « data challenge » sont effectuées par les 4 organisations enfin de valider son fonctionnement (CHOLLET et *al*., 2007; MALEK, 2009). Le 30 mars 2010, la plus grande expérience du LHC a été initiée avec succès, générant des résultats prometteurs pour l'ensemble des laboratoires mondiaux engagés dans le domaine et engageant l'infrastructure de grille sous-jacente dans un rôle opérationnel jamais atteint[17].

II - 2.3.5.2.2 *La recherche biomédicale et l'organisation virtuelle BIOMED*

La deuxième organisation virtuelle la plus consommatrice de ressources au sein d'EGEE est BIOMED, dédiée à la recherche biomédicale. En effet, plus de 20.000 nœuds de calcul sont utilisés par une dizaine d'application impliquées dans cette organisation virtuelle (BRETON, 2009). Cette organisation virtuelle, qui utilise environ 5% des ressources d'EGEE, est centrée sur trois domaines principaux :

- l'imagerie médicale,
- la bioinformatique qui s'intéresse à l'analyse de séquences génétiques,

[14] http://projects.um.edu.mt/eumedgrid/information/fr-apps-running-on-egee.pdf
[15] http://lcg.web.cern.ch/LCG/
[16] Centre de Calcul de l'Institut National de Physique Nucléaire et de Physique des Particules (http://cc.in2p3.fr/)
[17] http://press.web.cern.ch/press/PressReleases/Releases2010/PR07.10F.html

- la découverte d'antiviraux en développant des expérimentations *in sillico*[18].

Dans ces domaines de recherche, la technologie grille, par ses potentialités informatiques accrues, permet par exemple d'accroître le nombre de scénarios d'évolution de virus, d'améliorer la visualisation des structures moléculaires (3D), de tester des combinaisons de conformation virus / principes actifs et de fédérer des laboratoires de recherche autour de problématiques communes comme dans le cas des pandémies (JACQ et *al*., 2008).

En comparaison avec la physique des particules dont les architectures informatiques, telle que la technologie grille, étaient initialement adaptées aux problématiques de calcul, la recherche biomédicale était, et reste, confrontée à plusieurs défis majeurs. Ainsi, afin d'éviter le recodage complet de ces applications, l'utilisation de services web ont permis de les interfacer avec la grille EGEE. De plus, ces applications reposent sur des stockages externes qu'il a fallu connecter de manière transparente. En ce qui concerne l'imagerie médicale, le recours à du calcul parallèle pour perfectionner l'analyse d'images médicales se révèle à l'heure actuelle comme un des défis majeurs de cette communauté (BRETON, 2009). D'autre part, le regroupement de chercheurs spécialisés dans les sciences de la vie et leurs contributions respectives ont nécessité le développement de portails d'accès aux outils, qui soient conviviaux et faciles d'utilisation (BENOIT-CATTIN et *al*., 2006). Enfin, les retours d'expériences des différents projets ont confirmé la viabilité d'une grille comme EGEE face aux questions de sécurité. En effet, la confidentialité des données médicales a été préservée grâce à leur cryptage préalable.

Finalement, si ces deux communautés représentent les principaux consommateurs et fournisseurs de ressources de la grille, EGGE d'autres communautés ont progressivement rejoint ce consortium telles que :

- l'archéologie
- l'astronomie et l'astrophysique
- la chimie
- la finance
- la recherche en informatique
- les sciences de la planète

Ainsi, l'initiative EGEE et sa mise en production progressive démontre le bien fondé de la technologie grille pour prendre en charge les aspects tant organisationnels que technologiques de projets opérationnels impliquant des ressources informatiques à la demande et conséquentes, et des compétences multidisciplinaires (ESFRI, 2008).

[18] Néologisme correspondant à la simulation numérique dans les sciences de la vie

II - 2.3.5.3. La grille : une technologie adaptée aux sciences de la Planète

À l'échelle des sciences de la Planète, des efforts coordonnés, comme par exemple le projet EU FP6 DEGREE[19] (Dissemination and Exploitation of GRids in Earth sciencE), permettent progressivement de formaliser une infrastructure de grande ampleur, accessible par un grand nombre d'utilisateurs désireux de tirer profit de solutions technologiques viables et matures. En effet, les applications dédiées à l'expertise environnementale et plus largement aux sciences de la Planète ont en commun de nécessiter la gestion de données multidimensionnelles aux échelles spatiales et temporelles variables et dont la taille est souvent importante voire problématique. Ces données multi-sources, provenant de systèmes d'observations et de collecte, doivent être regroupées autour d'une même infrastructure informatique afin de servir l'aspect multithématique et multi-organisationnel des sciences de la Planète (DEGREE PROJECT CONSORTIUM, 2008). Durant ces dernières années, la communauté des sciences de la Terre et les recherches technologiques associées ont montré un grand intérêt pour le web et surtout la technologie grille pour répondre à leurs besoins (DI, 2006; PETITDIDIER et *al*., 2009; RENARD et *al*., 2009; WOOLF et NATIVI, 2008). D'autre part, le traitement de ce type de données nécessite d'une part des puissances de calcul et de la capacité de stockage, et d'autre part la construction dynamique de chaînes de traitement. Enfin, cette modularité doit être soutenue par des services de découverte de services et de métadonnées capables d'informer les utilisateurs sur la disponibilité des données et des algorithmes et leurs caractéristiques fonctionnelles de manière transparente (GAGLIARDI et *al*., 2005).

À titre d'exemple, HLUCHY et *al*. (2005) développent le concept d'organisation virtuelle autour de l'objectif de gestion de crise « inondation » en Slovaquie. L'infrastructure de grille permet le regroupement des acteurs responsables de la gestion de crise, comme les décisionnaires de la gestion des opérations de sécurité civile, les experts en charge de l'évaluation météorologique et hydrologique, les services institutionnels et privés, et dans une moindre mesure la population (Figure 26).

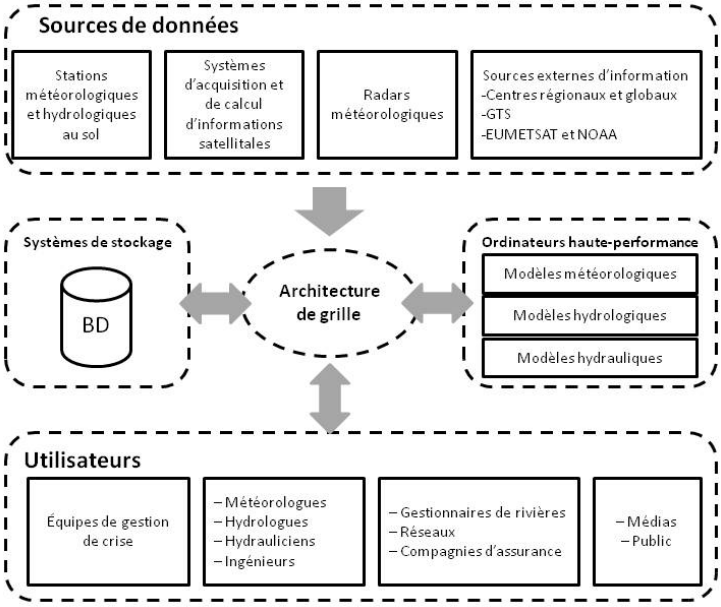


Figure 26 : Organisation virtuelle dédiée aux inondations de plaine
(tirée de HLUCHY et *al*., 2003)

[19] http://www.eu-degree.eu/

L'architecture développée permet d'accéder à des données pluri-thématiques et multi-sources permettant d'observer et d'interpréter précisément l'aléa hydrométéorologique. Cette architecture informatique, accessible par différents acteurs adhérant à l'organisation virtuelle permet de rechercher et de visualiser ces jeux de données, et de les modéliser dynamiquement sans se préoccuper de l'aspect technologique. En effet, il est souvent considéré comme un facteur limitant à l'élaboration d'une démarche scientifique performante et *a fortiori* au bon déroulement des opérations de gestion de crise qui nécessitent une assise technologique pérenne, opérationnelle et fiable.

II - 2.4. Synthèse sur la technologie grille

Dans les paragraphes précédents, la technologie grille a été présentée en détail au travers de son historique et de son appartenance au champ plus large de l'informatique distribuée. S'il existe plusieurs types de grille, la tendance actuelle s'oriente vers un élargissement et une généralisation des solutions, au travers de standards promulgués par le consortium OGF[20]. Cet organisme s'attache à harmoniser et motiver le développement de services de grille interopérables et standards. Dans ce sens, l'architecture de grille EGEE, qui s'attache à respecter les recommandations de l'OGF, représente la plus importante initiative de grille collaborative à l'échelle européenne. En intégrant des ressources informatiques géographiquement distantes et un grand nombre d'utilisateurs, cette architecture permet de répondre progressivement aux besoins de collaboration requis par la majorité des communautés scientifiques. Le succès de sa mise en production repose vraisemblablement sur le concept d'organisation virtuelle (VO), un élément organisationnel central permettant la fourniture, l'accès et le partage sécurisés de ressources informatiques (données, processeurs et stockage). Ainsi, l'architecture de grille EGEE offre *a priori* plusieurs avantages pour mener l'étude de faisabilité proposée dans cette recherche.

[20] Open Grid Forum : http://www.ogf.org/

II - 3. La cartographie en ligne entre géomatique et services web

L'avènement progressif des réseaux informatiques offre de nouvelles potentialités technologiques à de nombreux domaines scientifiques, tel que la géomatique. En effet, la gestion et le partage de l'information géographique, essentielle à la caractérisation des phénomènes naturels ayant lieu à la surface de la Terre, a pu être améliorée par le recours à des solutions tirant profit d'Internet et du Web. Ainsi, cette deuxième partie de chapitre s'attache à présenter les fondements conceptuels de la géomatique, science de l'information géographie, et son évolution croissante vers une distribution géographique de ses composants logiciels et de ses sources de données.

II - 3.1. Gestion des données géoréférencées : SIG et géomatique

II - 3.1.1. Définitions

La géomatique ou science de l'information géographique, définie comme « l'ensemble des techniques de traitement informatique de l'information géographique » (DENEGRE et SALGE, 2004), peut être assimilée à une science à part entière, la science de l'information géographique (GOODCHILD, 1992) et correspond à un des champs de recherche de cette étude. L'information géographique peut être représentée par un point, une ligne ou une aire, et par conséquent permet de représenter tout phénomène spatial[21]. Dans un contexte technologique de plus en plus propice, de nombreuses initiatives ont vu le jour afin de faciliter la manipulation et la réutilisation de l'information géographique pour répondre à différents besoins par des utilisateurs non-experts (BURROUGH, 1986 ; LONGLEY et *al*., 2005). D'un point de vue scientifique, l'apparition et le succès rapide des systèmes d'information géographique (SIG), dédiés à la prise en charge de l'information géographique, peuvent être corrélé à la conviction que les phénomènes à la surface de la Terre et leur compréhension nécessitent une approche multidisciplinaire. Ainsi, le SIG, en superposant des données de disciplines différentes mais complémentaires, permettent d'avoir une vision plus générale et synthétique d'un problème donné (BURROUGH, 1986 ; LAURINI et THOMPSON, 1992 ; ROCHE et CARON, 2004). Enfin, le « SIG est fondamental pour la résolution de problèmes du monde réel » (LONGLEY et *al*., 2005), ce qui correspond pleinement à la problématique d'expertise hydrométéorologique nécessaire à la gestion de crise « inondation ».

Il existe de nombreuses définitions et appellations faisant référence à un système d'information géographique (BORDIN, 2002). Certains auteurs accordent une importance particulière à sa dimension « base de données » (PORNON, 1992 ; SERVIGNE et LIBOUREL, 2006). BORDIN (2002) résume cette approche en considérant les SIG comme des « systèmes de gestion de base de données pour la saisie, le stockage, l'extraction, l'interrogation, l'analyse et l'affichage de données localisées ». L'accent peut être mis sur l'aspect « traitement » en assimilant un SIG à « un groupe de procédures qui permettent la saisie, le stockage, la récupération, la cartographie et l'analyse spatiale de données tant spatiales qu'attributaires pour supporter l'aide à la décision de l'organisation » (GRIMSHAW, 1994). Enfin, JOLIVEAU (1996) globalise le SIG comme :

[21] Il est possible de distinguer l'approche « vecteur » principalement utilisée pour représenter des informations structurelles et l'approche « raster » dédiée à la représentation de phénomènes spatialement continus

« l'ensemble des structures, des méthodes, des outils et des données constitué pour rendre compte des phénomènes localisés dans un espace spécifique et faciliter les décisions à prendre sur cet espace ».

C'est donc dans une optique d'aide à la décision qu'il est appréhendé dans cette recherche.

Cette multiplicité de points de vue ne doit pas occulter le fait que l'information géographique constitue le cœur même d'un SIG. La meilleure façon de l'appréhender est certainement d'énumérer les questions fondamentales auxquelles il se doit de répondre (BORDIN, 2002 ; BURROUGH, 1986) :

- les questions simples, telles que Où ? Quoi ? (Qui ? Quel ?) qui correspondent à des fonctions de gestion permettant de localiser un objet et de se renseigner sur ses caractéristiques, elles traduisent directement la démarche de lecture d'une carte.
- les requêtes complexes, comme Comment ? Et si ? qui ouvre la voie à des fonctionnalités évoluées basées sur les tâches basiques du premier niveau. Au travers de ces questions, le SIG s'apparente à de véritables outils d'aide à la décision ou de simulations, comme par exemple des études d'impacts environnementaux.
- le cas particulier du Quand ? qui n'apparait pas être un élément fondateur du SIG, et qui encore de nos jours n'a pas réellement de fonctionnalités dédiées et nécessitant souvent un effort technique non négligeable.

L'une des plus grandes caractéristiques du SIG est sa capacité à superposer des couches représentants différents types d'information géographique (AYRAL et SAUVAGNARGUES, 2009). Par conséquence, il existe un grand nombre de façons de représenter cartographiquement les données brutes et l'information géographique en résultant. Ainsi, la sémiologie graphique s'intéresse aux différentes manières de représenter les données géographiques sur une carte de manière à présenter et à mettre en valeur l'information souhaitée, au point de pouvoir « faire mentir les cartes » (MONMONIER, 1993). Ainsi, le résultat cartographique obtenu est souvent « plus facile à lire qu'un photographie, plus précise qu'un texte et plus expressive qu'un tableau […] apporte une information traitée et sélectionnée, mise à portée du lecteur […] bref la traduction cartographique d'une réalité » (Brunet, 1987). La cartographie et le SIG et plus largement l'information géographique apparaissent donc comme des éléments primordiaux dans un contexte d'aide à la décision, tels que la gestion de crise « inondation » (SAUVAGNARGUES-LESAGE et AYRAL, 2007).

II - 3.1.2. Structuration et finalité de l'information géographique dans un SIG

D'un point de vue structurel, un SIG est structuré autour de plusieurs composantes (AYRAL et SAUVAGNARGUES-LESAGE, 2009 ; BURROUGH, 1986 ; PORNON, 1992) (Figure 27).

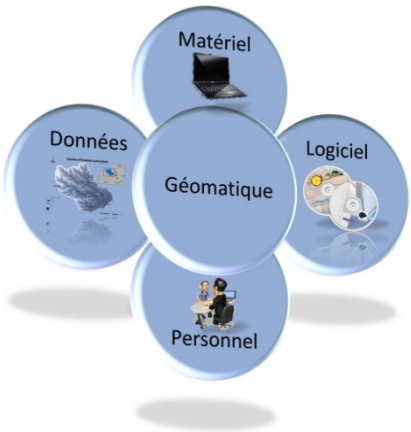


Figure 27 : Les 4 composantes du SIG et par extension de la géomatique

Ces composantes sont au nombre de 4 :

- le matériel, à savoir un ordinateur, un serveur ou un réseau local, du matériel d'acquisition (théodolite, GPS, scanner, etc.…), de stockage, de visualisation et d'impression. Certains auteurs complètent cette liste non exhaustive avec le réseau, tel que l'Internet (BURROUGH, 1986).
- les données qui sont de deux ordres, géographiques et attributaires. La base de données sous-jacente permet de les relier intrinsèquement, afin de pouvoir considérer un objet géographique tant par son aspect spatial que par ses caractéristiques attributaires. D'un point de vue économique, les données représentent entre 60 et 80% des coûts de mise en place d'un SIG (PORNON, 1992).
- les applications qui composent la dimension logicielle du SIG et qui opèrent en tant qu'interface entre les trois autres composantes. On retrouve 4 grandes familles de fonctionnalités :
    - l'acquisition en vue de manipuler les données,
    - l'archivage en vue de stocker et d'extraire les données sous la forme d'une base de données à composante géographique,
    - l'analyse en vue d'interroger, de transformer et d'étudier les données
    - l'affichage en vue de visualiser les données et de produire un résultat cartographique

  Au sein de ces 4 familles de nombreuses fonctionnalités existent et évoluent suivant la finalité même du SIG.

- les personnes que l'on retrouve tout au long du cycle de vie du SIG. La démocratisation de l'information géographique dans le quotidien de la population implique de plus en plus cette dernière en tant qu'utilisateur à part entière.

De nombreux domaines utilisent les SIG pour gérer et traiter l'information géographique. Ainsi, la classification des SIG varie d'un auteur à l'autre, BORDIN (2002) identifie 4 types d'usage principaux :

- inventaire-observatoire qui permet de collecter, d'intégrer et de mettre en cohérence un ensemble de données autour d'un thème donné,
- étude-aide à la décision dont l'objectif est d'analyser les données géographiques entre elles pour répondre à une problématique donnée,
- gestion-suivi qui correspond à la dimension « application » d'un SIG,
- communication qui facilite la diffusion d'information géographique entre plusieurs utilisateurs.

Dans ce contexte, l'utilisation croissante des SIG dans le monde industriel et institutionnel est tout à fait justifiée. Si la plupart des défis technologiques inhérents à l'utilisation des SIG par des utilisateurs non-spécialistes sont à l'heure actuelle maîtrisés aux travers de solutions commerciales de plus en plus performantes, il subsiste toujours des difficultés d'ordre technologique et politico-organisationnelle pour le partage de l'information géographique à grande échelle.

II - 3.2. Les SIG et les réseaux informatiques

II - 3.2.1. Cartographie en ligne : définitions et principe

Dans la suite de cet exposé, une attention particulière est donnée à l'utilisation et le partage de l'information géographique au travers des réseaux informatiques. En effet, le perfectionnement des réseaux en termes de largeur de bande passante et de débit a permis d'entrevoir une nouvelle façon d'appréhender et de manipuler l'information géographique (Figure 28).

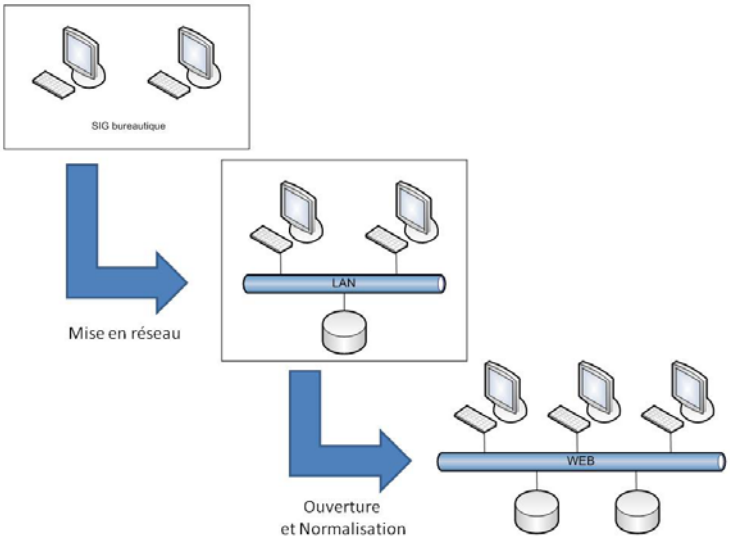


Figure 28 : Évolution architectural des SIG en ligne

Ainsi, au travers d'Internet et plus largement des réseaux informatiques, le domaine des SIG a vu se développer des capacités d'échanges de données à grande échelle mais aussi la possibilité « d'incorporer interaction et animation dans l'affichage » (PETERSON, 2005).

L'internet apparait comme un moyen de diffusion et de partage à l'échelle planétaire puisque quiconque le désire à l'heure actuelle peut se connecter à ce réseau mondial et communiquer avec d'autres utilisateurs géographiquement distants. La diffusion de l'information géographique ne représente pas une exception en soi, elle peut a priori être échangée à travers le réseau Internet. Cependant, l'aspect composite de l'information géographique, caractérisée par ses propriétés attributaires et spatiales, est un défi technologique majeur pour l'utilisation des SIG en ligne (CARTWRIGHT, 1999 ; KRAAK et BROWN, 2001 ; VAN ELZAKKER, 2001a).

Il est possible de définir la cartographie en ligne ou « Webmapping » comme un « processus de distribution de données géoréférencées via un réseau tel qu'Internet ou un intranet et de leur visualisation sur des applications cartographiques via une interface web » (MITCHELL, 2005). PENG et TSOU (2003) préfèrent le terme de « SIG Internet » (en anglais « Internet GIS ») qu'ils définissant comme « le domaine de recherche et d'application qui utilise Internet et les autres systèmes réseau (incluant les communications sans fil et les Intranets) pour faciliter l'accès, le traitement et la diffusion d'information géographique ». Finalement, l'omniprésence des technologies multimédias apparaît comme une caractéristique essentielle des SIG en ligne (LEPRINCE et *al*., 2003).

La définition suivante est ainsi proposée pour la cartographie en ligne :

Processus de distribution et de manipulation de données géoréférencées via un réseau informatique visant à faciliter leur partage et leur visualisation au moyen d'une interface web enrichie de technologies multimédias

II - 3.2.2. Une typologie de la cartographie en ligne

Une revue bibliographique (AVRIL et *al*., 2005 ; AYRAL et *al*., 2010 ; CARTWRIGHT, 2005 ; KÖBBEN, 2001 ; LEPRINCE et *al*., 2003 ; MACEACHREN, 1998 ; MITCHELL, 2005 ; PENG et TSOU, 2003 ; PETERSON, 2005 ; STANGER, 2008 ; VAN ELZAKKER, 2001a ; 2001b), a permis d'établir une typologie des solutions cartographiques au travers de réseaux locaux ou d'Internet. La précision de la typologie et les critères de distinction peuvent varier d'un auteur à l'autre, cependant ces classifications s'appuient sur les potentialités fonctionnelles des solutions et les architectures informatiques sous-jacentes.

De manière générale, la cartographie en ligne s'appuie historiquement sur le modèle informatique client-serveur couplé à une base de données distante ou intégrée au serveur d'application (respectivement une architecture deux tiers et trois tiers) (Figure **29**).

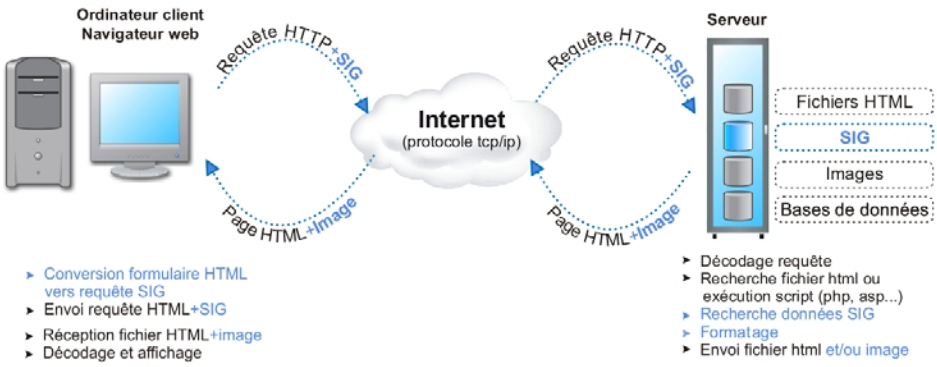


Figure 29 : Modèle client-serveur dédié à la cartographie en ligne (JEGOU, 2009)

D'autre part, trois critères principaux permettent de caractériser une application de cartographie du point de vue utilisateur (VAN ELZAKKER, 2001a) :

- l'audience, selon que l'application soit dédiée à des fins personnelles ou à un groupe d'utilisateurs,
- l'interaction (ou dynamicité) au sens d'interactivité relative à la possibilité « d'interroger n'importe quelle partie de la carte et ensuite demander toute information supplémentaire disponible » (CARTWRIGHT, 2005),
- la diversité et la richesse des données et des fonctionnalités, qui sous-entend la connaissance préalable des ressources mises à disposition ou leur découverte « à la volée ».

En se basant sur ces critères de performance, il est possible d'identifier deux grandes familles de solutions de SIG en ligne dans lesquelles l'interaction et la diversité fonctionnelle apparaissent croissantes :

- La diffusion de cartes statiques ou « cartes mortes » scannées ou dessinées sous forme d'images (format JPEG, GIF, etc.), ou de formats graphiques vectoriels (PDF, SVG, Flash, etc.) permettant d'améliorer la dynamicité[22] des solutions cartographiques,
- La mise à disposition de données géographiques par des serveurs cartographiques, dédiés qui offrent des moyens de cartographie à la demande et sont interactifs. Cette méthodologie inclut une dimension cliente importante

II - 3.2.2.1. La publication de cartes statiques

Il s'agit d'intégrer des cartes dans des sites web par l'intermédiaire de formats « image » (balises HTML). Cette approche offre peu d'interactivité à l'utilisateur, seule la définition préalable de zones cliquables peut permettre d'offrir des fonctionnalités de zoom. Une évolution vers des solutions plus dynamiques, a été initiée par l'utilisation de formats graphiques tels que le PDF (Portable Document Format), le SVG (Scalable Vector Graphics), format graphique à échelle variable, et le format propriétaire Flash (CARTWRIGHT, 2005 ; PENG et TSOU, 2003 ; PETERSON, 2005). Ces formats permettent « une écriture graphique vectorielle en 2D » (AVRIL et *al.*, 2005). Le serveur est capable de produire ces for-

[22] Ce terme correspond à la capacité d'actualisation graphique

mats en fonction des requêtes de l'utilisateur et de les intégrer dans le site web. Côté client, un plugin est souvent nécessaire pour visualiser et manipuler ces formats.

II - 3.2.2.2. La cartographie en ligne interactive et l'enrichissement des solutions clientes

L'idée fondamentale de cette approche repose sur l'enrichissement du serveur avec des fonctionnalités cartographiques, permettant à l'utilisateur des requêtes sur l'information géographique. Cette solution s'appuie sur deux technologies logicielles majeures :

- les formulaires qui correspondent à des pages Internet que l'utilisateur remplit et soumet au serveur (HTTP GET),
- la Common Gateway Interface (CGI) qui est une interface intégrée dans le serveur et qui permet l'interprétation de paramètres fournis par l'utilisateur et le contrôle de serveurs d'applications et de base de données.

Cette solution est basée sur une architecture trois tiers (client – serveur). Le premier tiers correspond au navigateur web côté client qui permet de faire des requêtes, le second tiers est le serveur web enrichi de fonctionnalités permettant la prise en charge de données géographiques et l'interprétation de requêtes paramétriques (CGI), enfin le dernier tiers représente les serveurs de données et d'applications avec lesquels le serveur central peut communiquer et produire le résultat publié sur le navigateur du client (Figure 30).

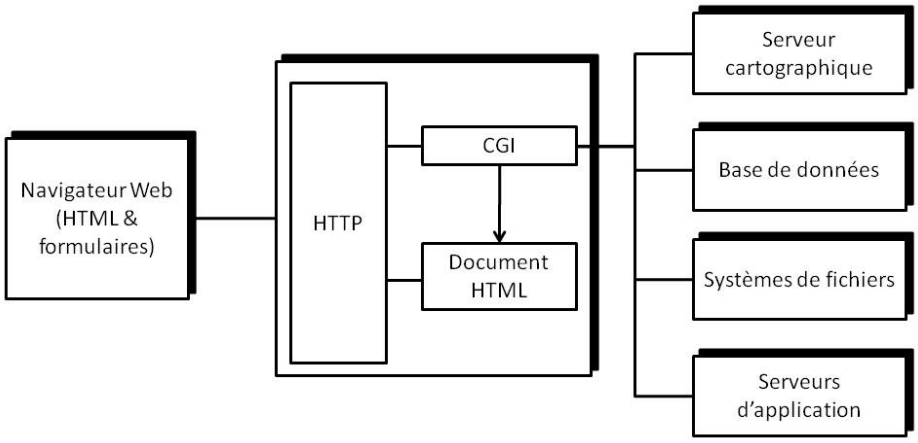


Figure 30 : Serveur cartographique enrichi d'une interface CGI (d'après PENG et TSOU, 2003)

Cette solution est à l'heure actuelle la plus répandue. L'une des premières initiatives correspond au serveur cartographique Mapserver développé durant les années 90 par l'université du Minnesota (MITCHELL, 2005). La solution commerciale ArcIMS® d'ESRI se base sur la même philosophie de fonctionnement pour laquelle des extensions fonctionnelles sont intégrées au serveur web enfin de supporter la gestion de l'information géographique (JIANG, 2005).

A l'origine, le navigateur web affichait la cartographie demandée par l'utilisateur sans que ce dernier puisse réellement naviguer et changer son apparence. En effet, chaque changement d'affichage (zoom, recadrage, graphisme, etc.) nécessitait le renvoie d'une requête au serveur afin de reconstruire

la sortie cartographique adéquate (PENG et TSOU, 2003 ; MITCHELL, 2005). Ainsi, pour des sites web à grande audience, une surcharge du réseau pouvait subvenir et limiter la dynamicité de l'interface (JIANG, 2005). L'une des solutions, pour améliorer ces problématiques de mise en charge du réseau, est d'avoir recours à des solutions cartographiques clientes.

Dans ce sens, le recours à des *plugins* côté client comme la plateforme JAVA et les contrôles ActiveX, ou aux langages JavaScript et au canevas AJAX permettent d'atteindre de tels objectifs (AVRIL et *al*., 2005 ; JIANG, 2005 ; KÖBBEN, 2001 ; PENG et TSOU, 2003). D'autre part, des clients cartographiques, tels que MapFish[23] et OpenLayers[24], permettent la prise en charge des données géographiques sur le navigateur client et fournissent un canevas graphique de base pour le développement de l'interface graphique. Dans ce sens, les fonctionnalités de base, telles que l'interrogation des données, le zoom ou la sémiologie sont effectuées sur le poste client de manière à limiter les interactions avec le serveur cartographique distant.

Les solutions opérationnelles de cartographie en ligne tirent souvent profit des différentes solutions abordées et sont de manière générale basées sur un serveur cartographique. Les paragraphes précédents se sont attachés à montrer les potentialités d'un SIG en ligne pour supporter les besoins de d'interrogation, de visualisation et de partage de l'information géographique qui apparaissent nécessaires dans un contexte d'aide à la décision relatif à une situation de crise.

II - 3.2.3. La science de l'information géographique

Les paragraphes suivants présentent l'évolution du SIG et de la cartographie en ligne, telle qu'elle a été envisagée par GOODCHILD (1992) au travers de l'appellation « science de l'information géographique »

II - 3.2.3.1. Nouveaux enjeux du SIG en ligne

L'origine des solutions de SIG en ligne est souvent reliée à l'initiative de plusieurs partenaires à des fins de communication et de partage de données géographiques. La phase de conception est basée sur les caractéristiques informatiques des systèmes d'informations auxquels cette solution sera connectée, ainsi que sur les besoins fonctionnels des futurs utilisateurs. Cette situation engendre de nombreuses solutions algorithmiques et de nombreux formats de données rendant la généralisation et la réutilisation souvent difficile et limitant clairement l'interopérabilité des systèmes opérants (TALADOIRE, 2003). D'un point de vue applicatif, le SIG en ligne permet de « collecter des données provenant de plusieurs sources hétérogènes pour constituer le capital de connaissance sur lequel fonder son fonctionnement » (CULLOT et *al*., 2003), peut engendrer des problèmes critiques de compatibilité (TALADOIRE, 2003) et une multiplication des interfaces de communication (Figure 31a).

À titre d'exemple, la directive européenne INSPIRE (Infrastructure for Spatial Information in Europe) vise à créer une infrastructure de données spatiales pour mieux formuler, mettre en œuvre et assurer le suivi des politiques communautaires environnementales (AFIGéO, 2007). Cette infrastructure a pour objectif de permettre la découverte de données spatiales et la construction de chaînes de services d'information géographique par des utilisateurs finaux.

[23] http://mapfish.org/
[24] http://openlayers.org/

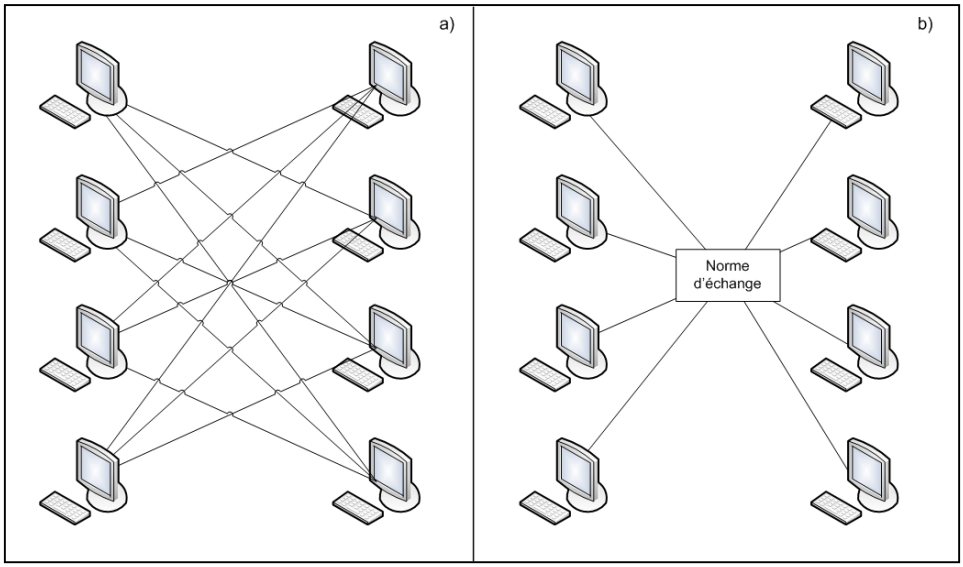


Figure 31 : Connexion de SIG distants sans norme d'échange (a)) et avec une norme d'échange (b)) (BORDIN, 2002)

D'autre part, l'orientation informationnelle du World Wild Web, a longtemps limité les potentialités fonctionnelles des sites web et donc de la cartographie en ligne (JIANG, 2005). Dans ce contexte, les nouvelles tendances technologiques et politiques de la cartographie en ligne, et plus largement des SIG s'attachent d'une part à normaliser la gestion de l'information géographique à l'échelle internationale (Figure 31b) et d'autre part à promouvoir *de facto* le développement de services d'information géographique offrant une véritable capacité de géotraitement aux SIG en ligne.

II - 3.2.3.2. Principe fonctionnel des services d'information géographique

Les technologies relatives à l'information géographique ont de tout temps suivi de près les évolutions des technologies de l'information et de la communication (PENG et TSOU, 2003). La maîtrise croissante d'architectures informatiques grande échelle administrées par des intergiciels de plus en plus performants (cf. §.II – 2.2.3) garantit progressivement une base technologique fiable pour envisager le SIG de manière distribuée.

L'article de TSOU et BUTTENFIELD (2002) « A dynamic architecture for distributing geographic information services » pose les bases fondamentales de ce que doivent être les services d'information géographique. Cette philosophie promeut le développement de solutions multi-tiers (CUTHBERT, 1999), dans un mode de communication proche du pair-à-pair (cf. §.II - 2.2.2). Il est question de « nœud SIG » mis en relation afin de partager un ensemble de services, tels que des données et des géotraitements (Figure 32).

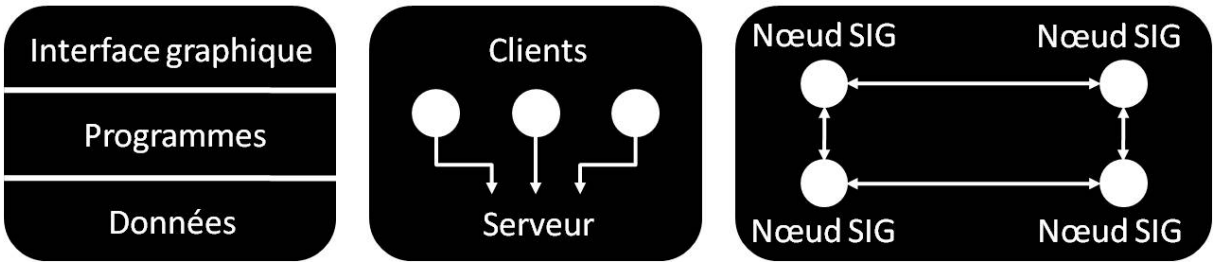


Figure 32 : Les trois grandes types d'architecture SIG (PENG et TSOU, 2003 ; TSOU et BUTTENFIELD (2002)

L'*internet GIS* (PENG et TSOU, 2003), le web geospatial (LAKE et FARLEY, 2007), les services d'information géographique (TSOU et BUTTENFIELD, 2002) ou bien encore le GeoWeb 2.0[25] sont autant d'appellations faisant référence à cette nouvelle approche de la gestion de l'information géographique de manière ubiquitaire. Ces différents néologismes constituent le socle technologique et philosophique d'une « néogéographie[26] » dans laquelle les composants fondamentaux des SIG sont géographiquement distribués et permettent de résoudre le « manque d'interopérabilité, de modularité et de flexibilité » des SIG en ligne actuels (TSOU et BUTTENFIELD, 2002).

Au travers de cette nouvelle approche, l'architecture, au sens large, incluant les données et les outils, est construite « temporairement en connectant des objets (données) et des composants logiciels à travers les réseaux informatiques » (TSOU et BUTTENFIELD, 2002). LAKE et FARLEY (2007) élargissent cette vision en positionnant cette architecture à travers de « multiples juridictions et régions géographiques ».

Si la mise en place de tels systèmes requiert une certaine révolution des mentalités, par rapport à la propriété intellectuelle et la nécessité de reconfiguration des systèmes monolithiques existants (TSOU et BUTTENFIELD, 2002), cette nouvelle approche offre plusieurs avantages indéniables (PENG et TSOU, 2003) :

- la globalisation de l'accès à l'information géographique et parallèlement sa décentralisation qui semble faciliter le traitement de données géographiques de plus en plus conséquentes. Il en résulte une révolution dans la manière de traiter ces données, à la méthode classique de rapatriement des données s'oppose un nouvelle vision où les algorithmes vont là où la donnée se trouve.

- la construction d'une chaîne de traitements « à la demande » peut être sérieusement envisagée, en opposition à l'observation classique que « 90 % des utilisateurs utilisent moins de 10 % des logiciels » (PENG et TSOU, 2003).

- cette approche basée sur les récentes révolutions informatiques rend l'information géographique et ses outils de traitement accessible par n'importe qui à n'importe quel endroit (Pornon et *al*., 2008).

Comme le souligne TSOU et BUTTENFIELD (2002), une architecture informatique distribuée représente un support adapté pour réaliser une telle infrastructure. Il paraît dès lors envisageable d'utiliser la technologie grille comme support de services d'information géographiques dédiés au traitement de fonctionnalités SIG plus complexes (Comment ? et Si ? du §.II - 3.1.1).

[25] D'après le GéoSéminaire 2008 - Géoweb 2.0 : vers des SIG collaboratifs ? (cf. http://geoseminaire2008.teledetection.fr/)
[26] http://fr.wikipedia.org/wiki/Néogéographie

II - 3.2.3.3. L'Open Geospatial Consortium : une formalisation opérationnelle des services d'information géographique

Ces services géographiques résultent clairement de la politique mondiale de normalisation de l'information géographique principalement guidée par le l'Open Geospatial Consortium (OGC). Associé au comité technique ISO/TC211, l'effort de normalisation de l'OGC « illustre la détermination de résoudre les problèmes associés à l'intégration d'applications SIG en ligne » (TSOU et BUTTENFIELD, 2002).

L'OGC, anciennement OpenGIS Consortium, représente actuellement la plus importante initiative de normalisation et de standardisation de l'information géographique. Ce consortium à but non lucratif, international et volontaire a été fondé en 1994 et regroupe 388 industriels, agences gouvernementales et universités (OGC, 2009). Il est rattaché à l'Organisation Internationale de Normalisation (ISO) (LAKE et FARLEY, 2007). Son objectif principal est d'établir un ensemble de normes dans le but de faciliter l'échange de données géographiques et plus largement de services dédiés à leur manipulation. La première motivation de l'OGC concerne donc l'interopérabilité (cf. §.II – 1.3) des systèmes d'information géographique. Au niveau logiciel, elle signifie que deux composants (A et B) peuvent interopérer si A peut envoyer des requêtes pour un service à B, basé sur une compréhension mutuelle de cette requête par A et B, et si B peut de la même manière retourner une réponse à A mutuellement compréhensible (PERCIVALL, 2002) (Figure 33).

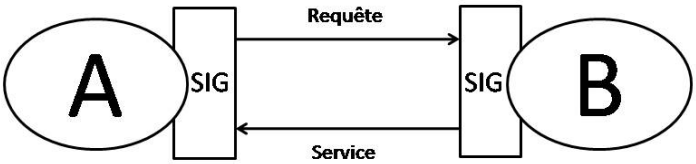


Figure 33 : Interopérabilité à l'échelle des composants
(D'après PERCIVALL, 2002)

Dans ce sens, l'OGC définit des spécifications techniques standardisées, à l'attention des développeurs, décrivant les protocoles, l'encodage des données et l'interface des logiciels à respecter pour garantir l'interopérabilité de leur système d'information géographique (OGC, 2009). Ce consortium n'a pas la vocation de développer des solutions opérationnelles mais plutôt des documents d'ingénierie. Concrètement, ces spécifications définissent le vocabulaire et la syntaxe des opérations de communication et de manipulation de l'information géographique[27]. Les membres industriels et académiques développent ensuite des solutions en accord avec ces recommandations (certifié *OGC Compliant*) et participent activement à leur généralisation.

II - 3.2.3.4. Interfaces, opérations et services

Les concepts fondamentaux de l'approche de l'OGC sont le service, l'interface et l'opération. Le service est fourni par une entité logicielle ou physique au travers d'interfaces. Ces dernières définissent l'ensemble des opérations dont dispose le service et la manière de les utiliser. Ainsi, une interface standardisée permet à un composant d'envoyer un message complexe (données et instructions) à un

[27] Il est souvent question d'ontologie (http://www.techno-science.net/?onglet=glossaire&definition=324)

autre composant qui est lui-même apte à interpréter cette requête et à retourner une réponse adéquate grâce à sa propre interface standardisée. D'autre part, elle détaille à la demande l'ensemble des opérations qui caractérise le service invoqué (KLOPFER, 2005).

En définissant les spécifications d'interfaces que les fournisseurs de données ou de géotraitements doivent respecter, l'OGC rend possible la mise en communication de systèmes d'information technologiquement indépendants et l'utilisation transparente de services décentralisés (Cuthbert, 1999). Il est question de web services de l'Open Geospatial Consortium (OWS). Ces services web standardisés encapsulent des données et/ou des géotraitements, et sont considérés comme des composants indépendants capables de communiquer et d'interagir afin d'automatiser des chaînes de traitements (OGC, 2009). Ainsi, à terme l'OGC sera capable de fournir les mêmes services, au sens d'opérations, que les SIG mais au travers du Web

Selon, le consortium W3C[28] relatif à la standardisation du World Wide Web, un service web est en substance (W3C, 2005):

un système logiciel construit pour supporter l'interaction de machine à machine de manière interopérable au travers d'un réseau. Il dispose d'une interface décrite dans un format interprétable par une machine […]. Les autres systèmes interagissent avec le service web, d'une manière prévue par sa description, grâce à des messages SOAP, utilisant le HTTP avec une sérialisation XML et en conjonction avec les autres standards relatifs au Web

Les services web « permettent aux implémentations d'être réutilisées sans impacter les clients […] et correspondent à un système faiblement couplé », indépendants des langages de programmation, ils sont développés au dessus du système d'information afin que « les processus internes ne soient jamais affectés » (KRALIDIS, 2007). Il en résulte pour les institutions une « réduction de la gestion des données […], l'intégration et le développement rapide d'applications et […] un engagement réduit ». Il est question d'architecture orientée services, déjà abordée dans le cadre de la 3ème génération de grille (cf. §.II - 2.3.1.3).

Ainsi, l'analyse bibliographique développée dans les paragraphes précédents permet d'établir une définition d'un service d'information géographique :

Un composant logiciel auto-descriptif supportant l'accès à une donnée géoréférencée ou la prise en charge d'un géotraitement à travers un réseau informatique distribué, et capable d'interopérer dynamiquement avec un service web, un système d'information ou un utilisateur distants.

La première hypothèse de recherche, d'ordre technologique, peut être dès lors envisagée. Le prévisionniste peut s'appuyer sur une interface graphique interactive (cartographie en ligne) afin d'interroger et visualiser l'ensemble des données hydrométéorologiques (géoréférencées) collectées et/ou résultantes de géotraitements pris en charge par des services d'information géographique et exécutés sur

[28] http://www.w3.org/

les nœuds de calcul d'une architecture de grille. Cependant, même si les tendances actuelles de l'informatique distribuée tendent à permettre un enrichissement technologique de la cartographie en ligne, le fossé reste encore important entre un SIG bureautique (tels qu'ArcGis de la société ESRI ou MapInfo chez Pitney Bowes) et un SIG en ligne enrichi de services d'information géographique. En illustration, le témoignage de Mike Hickey, le patron de Pitney Bowes (MapInfo) pour qui "l'explosion de la *néogéographie* suscite de l'intérêt et pousse à la convergence des données collaboratives, mais ce n'est pas du SIG" (FRANCICA, 2007).

SYNTHESE

Ce chapitre s'est attaché à développer une analyse bibliographique sur les technologies informatiques, et leurs courants de pensées, expérimentées dans cette recherche. Deux éléments centraux de réflexion peuvent d'ores et déjà être soulignés :

- Les caractéristiques informatiques de grande ampleur offertes par l'architecture EGEE et la dimension collaborative qu'elle fournit apparaissent comme des garanties fortes à l'accès à des capacités de calcul et de partage dans un contexte sécurisé et fiable,
- Les initiatives récentes d'intégration de nouvelles potentialités fonctionnelles, offertes par les réseaux informatiques, dans les méthodes et outils propres à la géomatique, en font une solution pérenne pour intégrer, analyser, et publier les données relatives à l'expertise hydrométéorologique,

Ainsi, la suite de ce mémoire de recherche, et la proposition méthodologique qu'il développe, doit être appréhendée au travers de cette articulation technologique entre l'architecture de grille EGEE et les méthodes modernes de manipulation de l'information géographique, dans la même philosophie que la collaboration récente entre les consortiums de l'Open Grid Forum (OGF) et de l'Open Geospatial Consortium (OGC).

PARTIE 2

Technologie grille et géomatique pour l'expertise hydrométéorologique en situation de crise

« La machine elle-même, plus elle se perfectionne, plus elle s'efface derrière son rôle. Il semble que tout l'effort industriel de l'homme, tous ses calculs, toutes ses nuits de veille sur les épures, n'aboutissent, comme signes visibles, qu'à la seule simplicité, comme s'il fallait l'expérience de plusieurs générations pour dégager peu à peu la courbe d'une colonne, d'une carène, ou d'un fuselage d'avion, jusqu'à leur rendre la pureté élémentaire de la courbe d'un sein ou d'une épaule »

Antoine de Saint-Exupéry, Terre des hommes, 1931

Chapitre 3

L'expertise hydrométéorologique des crues à cinétique rapide

3.1. L'existant technologique du service de prévision des crues « Grand Delta »

3.2. Retours d'expérience : besoins opérationnels, scientifiques et technologiques

Le Service de Prévision des Crues « Grand Delta » (SPC-GD) correspond au cas d'utilisation sélectionné pour expérimenter la technologie grille en tant que solution technologique susceptible de répondre à la nouvelle mission de prévision de ce service. La confrontation entre le fonctionnement existant du SPC-GD et les retours d'expériences élaborés à la suite de crues à cinétique rapide exceptionnelles amène à la définition précise des besoins récurrents de ce service.

INTRODUCTION

La partie 1 de ce mémoire de recherche s'est attaché à établir un état de l'art sur les crues à cinétique rapide et les moyens mis en œuvre pour les prendre en charge en situation opérationnelle, et sur les nouvelles technologies de l'information et de la communication offrant des solutions de collaboration, de partage et d'aide à la décision efficaces et opérationnelles. L'intégration, au sens d'adoption, de ces technologies dans les systèmes d'information des services opérationnels de la Sécurité Civile, requiert une appréhension précise des moyens informatiques et organisationnels existants à l'heure actuelle dans ces services. Étant donné la répartition géographique de ces phénomènes de crues à cinétique rapide sur le quart sud-est du territoire métropolitain français, le service de prévision des crues « Grand Delta » apparaît comme un acteur opérationnel approprié aux objectifs de recherche de cette étude.

L'exigence d'amélioration continue et récurrente de ce service, dans la philosophie générale de la Sécurité Civile, impose une définition précise de ses faiblesses et de ses besoins en matière de prise en charge de l'information hydrométéorologique, relative aux crues à cinétique rapide, en situation de crise. La connaissance, la plus exhaustive possible, de ces besoins opérationnels et de l'existant fonctionnel du SPC-GD doivent permettre de spécifier précisément les orientations méthodologiques à mener pour répondre à son nécessaire perfectionnement.

De plus, l'objectif sous-jacent de généralisation des méthodes et outils proposés dans cette recherche impose une méthodologie d'analyse de l'existant rigoureuse et formalisée que des méthodes comme le *Reference Model of Open Distributed Processing* (RM-ODP) et le langage UML (*Unified Modelling Language*) peuvent fournir. Ainsi, la première partie de chapitre s'attache à décrire l'existant technologique du SPC-GD au moyen des méthodes d'analyse proposées ci-dessus. Dans une seconde partie, l'analyse de deux évènements hydrométéorologiques extrêmes, à savoir les inondations dans le département de l'Aude en 1999 et celles dans le département du Gard en 2002, doit permettre de déceler les besoins opérationnel de ce service en matière de prévision des crues à cinétique rapide. Ce chapitre, en définissant une base de connaissance solide sur la gestion opérationnelle des crues à cinétique rapide par le SPC-GD, doit permettre de définir précisément les orientations méthodologiques proposées dans cette recherche.

III - 1. L'existant technologique du Service de Prévision des Crues « Grand Delta »

III - 1.1. Les méthodes « Reference Model of Open Distributed Processing » et «Unified Modelling Language » dédiées à la spécification des systèmes d'information

En vue de présenter les objectifs d'améliorations technologiques du SPC-GD basée sur la technologie grille, il convient au préalable de décrire son système actuel d'information en précisant toutes ses caractéristiques. Il s'agit d'une étape nécessaire et indispensable pour la description complète d'un service opérationnel, tel que le SPC-GD, et l'adoption efficace et adaptée de la technologie grille au sein de ce dernier (VICAT-BLANC PRIMET et *al*., 2003). Pour décrire le système d'information du SPC-GD, Il est important d'avoir recours à des méthodologies formalisées et reconnues par les communautés scientifiques et d'ingénierie. Le cadre conceptuel pour l'ingénierie des systèmes d'information RM-ODP a été choisi (Reference Model of Open Distributed Processing) pour fournir des réponses au « Quoi décrire », et le langage UML (Unified Modelling Language) pour répondre au « Comment décrire ».

III - 1.1.1. La méthode RM-ODP ou modèle de référence pour le « traitement réparti ouvert »

L'intérêt pour RM-ODP se situe dans sa capacité à modéliser des systèmes d'information distribués, c'est-à-dire composés de « ressources informatiques hétérogènes appartenant à des domaines administratifs différents » (ISO/IEC 10746-1). Il est décrit par les normes ISO suivantes : ISO/IEC 10746-1[1], ISO/IEC 10746-2[2], ISO/IEC 10746-3[3], ISO/IEC 10746-4[4]. Ce cadre conceptuel définit des moyens de spécification de systèmes d'information au moyen de plusieurs points de vue (Tab. 2). De plus, son utilisation par l'Open Geospatial Consortium pour décrire ses spécifications techniques a motivé ce choix (PERCIVALL, 2002).

Point de vue	Description
Entreprise	Objectif, compétences et politiques du système d'information dans son environnement
Information	Les types d'informations prises en charge par le système et les contraintes sur leur utilisation et leur interprétation
Traitement	Décomposition fonctionnelle du système sous la forme d'objets interagissant
Ingénierie	L'infrastructure nécessaire pour supporter la distribution du système d'information
Technologie	Les technologies pour supporter la distribution du système

Tableau 2 : Points de vue du cadre conceptuel RM-ODP

Ainsi, la méthode RM-ODP, en fournissant des critères de description précis, permet une présentation hiérarchisée et objective du système d'information du SPC-GD et des améliorations technologiques

[1] ITU-T Rec. X.901 | ISO/IEC 10746-1 « Overview »
[2] ITU-T Rec. X.902 | ISO/IEC 10746-2 « Foundations »
[3] ITU-T Rec. X.903 | ISO/IEC 10746-3 « Architecture »
[4] ITU-T Rec. X.904 | ISO/IEC 10746-4 « Architectural Semantics »

proposées. Dans le cadre de cette spécification, elle a seulement été utilisée pour structurer la description UML du système existant et du système proposé. Les recommandations strictes définies dans la norme ISO/IEC 19793[5] ont été utilisées uniquement dans le choix des diagrammes UML pour décrire chaque point de vue RM-ODP.

III - 1.1.2. Le langage UML

Le langage UML apparaît bien adapté pour réaliser la spécification du système existant. Il s'agit d'un langage visuel standardisé qui permet de représenter l'état et le fonctionnement d'un système d'information (RATIONAL PARTNERS, 1997). Comme le montre la figure 34, ce langage est basé sur des diagrammes dédiés aux descriptions structurelles, faisant abstraction des aspects temporels et dynamiques du système, et d'autres comportementaux décrivant l'enchaînement des opérations pour la réalisation des tâches principales.

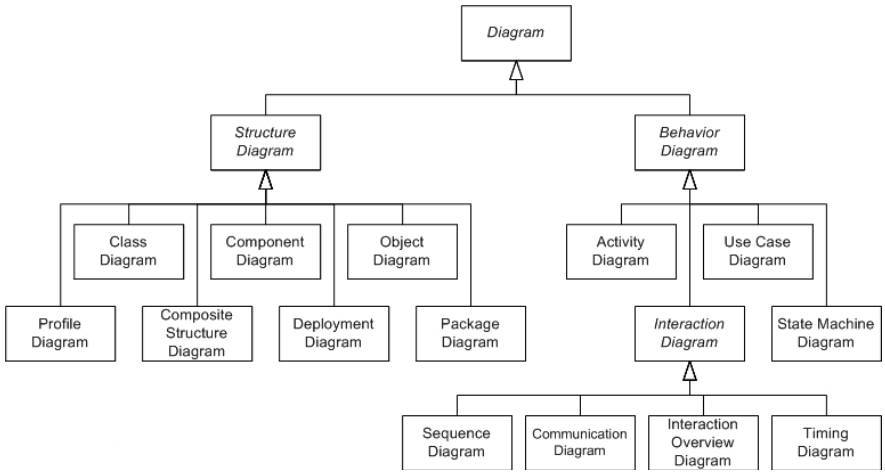


Figure 34 : Hiérarchie des diagrammes UML (d'après RATIONAL PARTNERS, 1997)

Pour ce second groupe, il existe une sous-classe de diagrammes intitulée les diagrammes d'interaction permettant de décrire précisément les interactions entre les composants du système d'information et celles avec les acteurs de ce système.

Dans le cadre de cette recherche, il s'agit de mieux comprendre dans un premier temps le comportement global et donc les fonctions « métier » du système d'information, et dans un second temps son fonctionnement proprement dit et sa structure technologique. Les diagrammes UML permettent d'appréhender cet existant technologique avec des niveaux d'abstraction différents.

L'analyse qui suit se base sur 5 de ces diagrammes UML :

- Le diagramme des cas d'utilisation qui montre les relations entre les acteurs et les fonctions principales du système
- Le diagramme d'activité dont l'objectif est de se focaliser sur l'enchaînement des traitements au sein de chaque cas d'utilisation

[5] ITU-T Rec X.906 | ISO/IEC 19793 « Use of UML for ODP systems specifications »

- Le diagramme de séquence permettant de représenter les interactions entre les différents acteurs et les composants du système d'information
- Le diagramme de classes qui représente une vue statique des objets composés de caractéristiques et de fonctions, ainsi que les relations entre ces objets.
- Le diagramme de déploiement qui décrit la structure statique de l'infrastructure physique et logicielle en montrant les relations de dépendance de l'ensemble des composants.

Finalement, l'intérêt d'utiliser des diagrammes UML réside dans la spécification précise et univoque des évolutions organisationnelles et technologiques apportées par l'adoption de la technologie grille et engagées dans cette recherche.

III - 1.2. Le système d'information du SPC-GD : rôles, compétences et capacités

III - 1.2.1. Présentation générale

Le service de prévision des crues « Grand Delta » est l'un des 22 SPC français. Il est rattaché à la Direction Départementale des Territoires et de la Mer (DDTM, anciennement DDE). Comme son nom l'indique et cela a été abordé dans le chapitre 1 de ce mémoire, il est en charge de la prévision des phénomènes hydrologiques extrêmes.

Son territoire de compétence s'étend sur 3 régions administratives (Provence-Alpes-Côte d'Azur, Languedoc-Roussillon, Rhône-Alpes), 7 départements (Ardèche, Drôme, Gard, Hérault, Vaucluse, Alpes de Haute Provence, Bouches du Rhône) et se sont donc 397 communes qui bénéficient de son dispositif de surveillance (DDE, 2006).

D'un point de vue hydrologique, 20 tronçons[6] réglementaires, c'est-à-dire pour lesquels une information sur les crues et leur prévision doit être obligatoirement développée (SPC-GD, 2006), sont effectivement surveillés par le SPC. Ces tronçons sont situés sur 12 cours d'eau principaux dont certains disposent de barrages EDF, de la Compagnie Nationale du Rhône (CNR) ou de barrages écrêteurs de crue pour le département du Gard (Fig. 35).

[6] Un tronçon est une section de cours d'eau

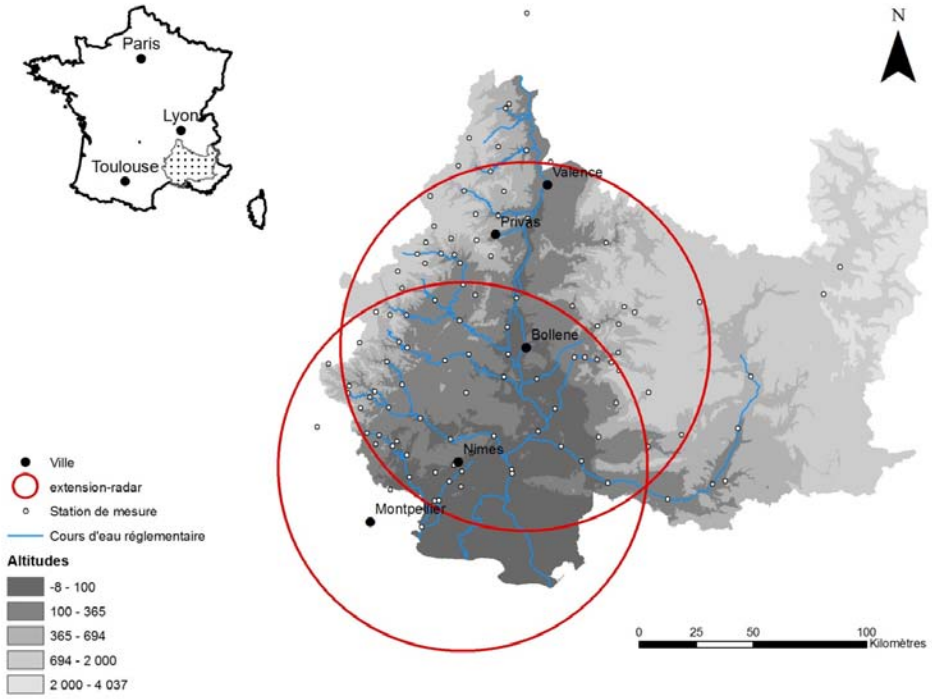


Figure 35 : Situation générale du territoire de compétence du SPC-GD
(sources : SPC-GD et ASTER GEDM 30 m[7])

Deux types principaux d'inondations affectent ce territoire, les inondations de plaine pour le Rhône, les parties aval de la Durance et de l'Ardèche ; les inondations relatives aux crues rapides telles que le bassin Vistre-Vidourle, les Gardons, la Cèze, la partie amont de l'Ardèche, l'Aygues, l'Ouvèze vauclusien et dans une moindre mesure l'Eyrieux et l'Ouvèze.

Afin de développer une proposition méthodologique dans la partie suivante, il convient de préciser les caractéristiques technologiques du système d'information du SPC-GD.

III - 1.2.2. Point de vue « Entreprise »

L'analyse du cadre institutionnel et opérationnel de la gestion de crise « inondation » (cf. §.I – 2.) a souligné l'importance des interrelations informationnelles entre les acteurs de la gestion de crise. Il convient donc de préciser la place du SPC-GD, et par extension de son système d'information, dans l'organisation opérationnelle institutionnelle, en mettant en valeur les échanges d'informations et de données entre le système et les acteurs externes (Fig. 36).

[7] http://www.gdem.aster.ersdac.or.jp/

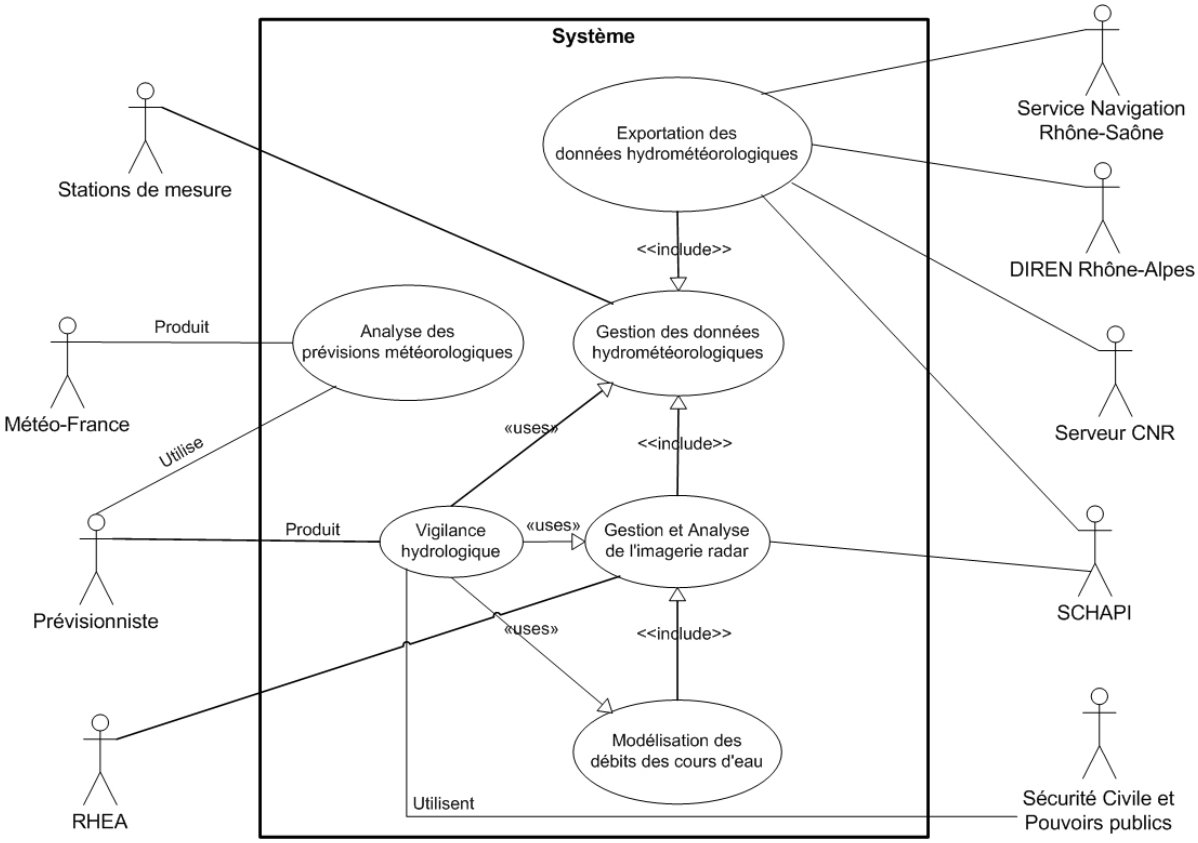


Figure 36 : Cas d'utilisation du système d'information du SPC-GD

L'acteur principal du système est le prévisionniste qui utilise et gère les différentes unités du système pour élaborer un bulletin de prévision le plus précis possible. Ce système d'information est basé sur des données hydrométéorologiques collectées depuis des stations de mesures réparties sur l'ensemble du territoire de compétence « Grand Delta » et des radars météorologiques situés à Bollène et Nîmes. En parallèle, le CMIR (Centre Météorologique InterRégional) d'Aix en Provence assiste régulièrement le prévisionniste par l'envoi de bulletins météorologiques complétés par des conversations téléphoniques régulières. L'expertise hydrologique est fondée sur 5 cas d'utilisation[8] :

- La gestion des données hydrométéorologiques,
- L'analyse, l'interprétation et la diffusion de l'imagerie radar,
- La modélisation du débit des cours d'eau,
- L'exportation des données hydrométéorologiques,
- L'analyse des prévisions météorologiques Météo-France, faiblement informatisé.

A noter que la modélisation hydrologique implique la bonne réception de l'imagerie radar, dépendante elle-même de la viabilité des données hydrométéorologiques collectées. L'ensemble des données radar et des outils permettant leur gestion, sont fournies par la société privée RHEA© et sont décrits dans les paragraphes qui suivent (cf. §.III - 1.2.3.2).

[8] Terme emprunté à la théorie de l'UML, correspondant aux fonctions principales d'un systèmes d'information et/ou d'une organisation

En ce qui concerne les données et informations sortantes du système, les données collectées sont répliquées sur des serveurs institutionnels tels que le Service Navigation Rhône-Saône, la DIREN Rhône-Alpes et la compagnie nationale du Rhône (CNR). Ces données sont également redirigées vers le système informatique du SCHAPI qui centralise l'ensemble des données des SPC et a de plus accès aux images radar du SPC-GD intégrées en temps réel à un site web.

Le développement de l'expertise hydrologique par les prévisionnistes du SPC s'établit en relation avec ceux du SCHAPI qui, en fonction de l'intensité de l'aléa, communique cette expertise aux services de la sécurité civile et aux pouvoirs publics de leurs échelons respectifs.

III - 1.2.3. Points de vue « Information » et « Traitement »

Dans ce paragraphe, il s'agit de présenter en détail les cas d'utilisation du SPC-GD, afin de comprendre le fonctionnement interne de son système d'information. Trois des cinq cas d'utilisation sont plus particulièrement analysés grâce un diagramme d'activité UML permettant de mettre en valeur leur dynamique temporelle.

III - 1.2.3.1. Gestion des données hydrométéorologiques

Ce cas d'utilisation est basé sur un réseau de mesures de 174 stations au sol appartenant en majorité à la DDTM mais aussi à EDF et au CNR (Fig. 35). Ces stations sont constituées d'un pluviomètre et/ou d'un limnimètre[9], L'enchaînement des opérations de cette activité s'effectue toutes les 5 minutes à chaque nouvelle réception de données collectées depuis les stations de mesures au sol (pluviomètres et limnimètres).

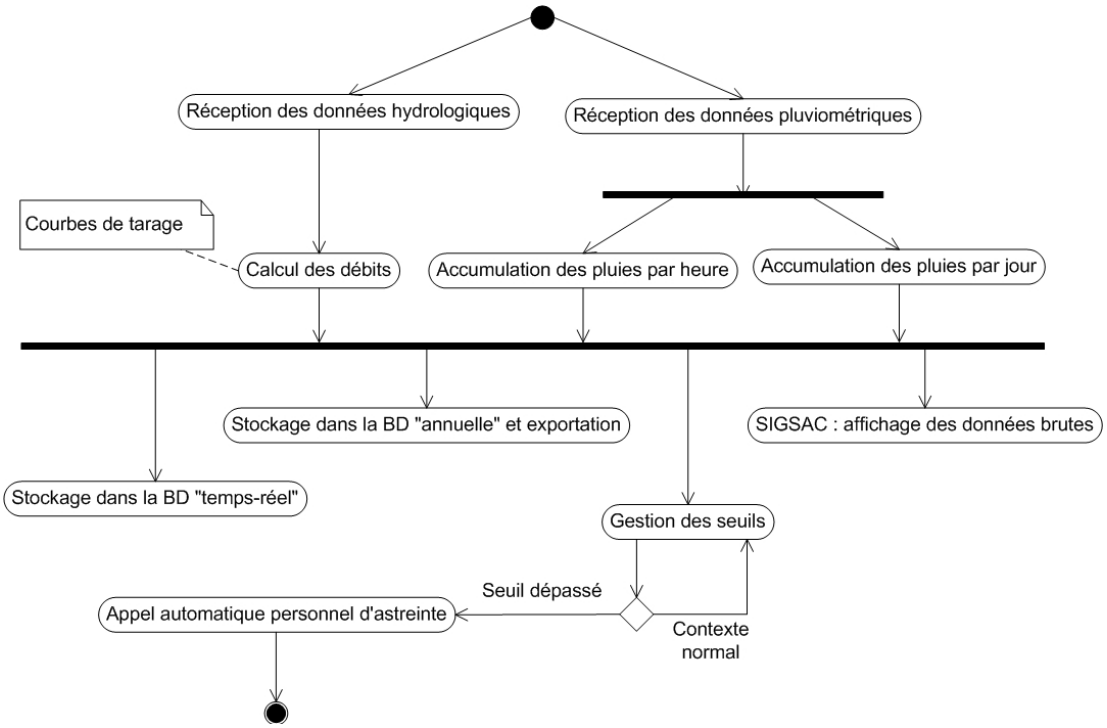


Figure 37 : Diagramme d'activité « gestion des données hydrométéorologiques »

[9] Instrument qui enregistre les variations du niveau de l'eau en fonction du temps (d'après le glossaire international d'hydrologie : http://webworld.unesco.org/water/ihp/db/glossary/glu/aglu.htm)

D'après la figure 37, après chaque réception, les fonctions principales de la gestion des données hydrométéorologiques brutes sont :

- Le décodage des informations radio et leur transformation en fichiers XML aux autres composants du système.
- La transformation des mesures limnimétriques en débits grâce aux courbes de tarage préalablement définies (jaugeages, retours d'expérience, etc.).
- Le calcul de cumuls pluviométriques horaires et journaliers.
- L'ensemble de ces données sont enregistrées dans deux bases de données, une pour la partie « temps réel » basée sur un cycle de 7 jours, et l'autre pour le stockage permanent avec un archivage annuel. Ces données permettent de développer en temps différé des opérations d'analyse et de calibration des modèles.
- Un système de surveillance automatique contrôle en temps-réel les seuils pluviométriques et limnimétriques et alerte visuellement et par téléphone le prévisionniste d'astreinte.
- L'interface graphique SIGSAC permet d'intégrer, de visualiser et d'interroger l'ensemble de ces données (BLEUSE, 2001).

III - 1.2.3.2. Gestion et analyse de l'imagerie radar

Ce cas d'utilisation offre aux prévisionnistes des moyens d'évaluer les caractéristiques et l'évolution des champs de pluie au travers de l'application CALAMAR 2 intégrant des images radar provenant de deux radars météorologiques situés à Nîmes et Bollène (Fig. 35 : cercles). Les observations radar sont découpés en 9 zones, dont la surface varie de 2500 km^2 à 4096 km^2 avec une résolution est de 1 km (Fig. 38). Cette application fournit des images radar « temps réel » (1 pas de temps) et une prévision à 1h (12 pas de temps) sur l'ensemble des 9 zones toutes les 5 minutes.

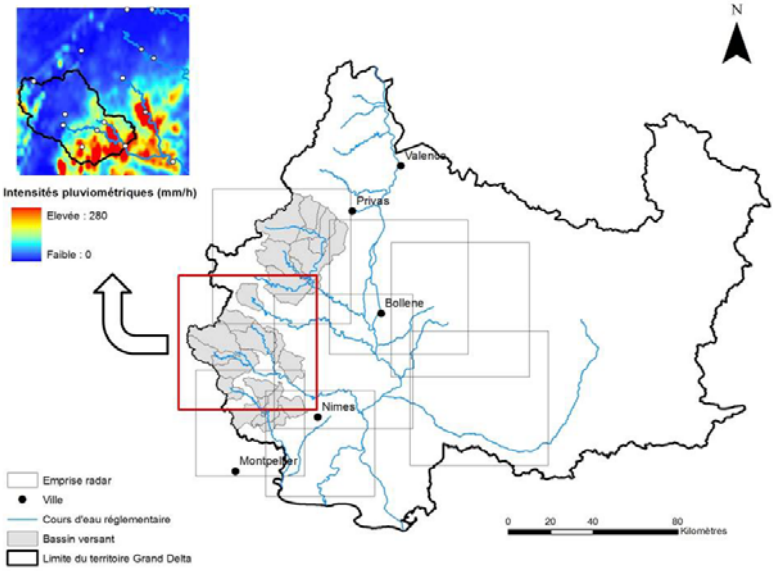


Figure 38 : Positions des zones radar et recouvrement d'une image radar sur le gardon d'Anduze (source : SPC-GD)

Les opérations les plus importantes concernent l'étalonnage et l'advection des images radar brutes. Après le géoréférencement de l'image, il est possible de distinguer deux phases de traitement (Fig. 39).

Il est possible de distinguer :

- les prétraitements durant lesquels les échos de sol et les masques causés par les reliefs sont retirés grâce à une méthodologie de voisinage et une carte de référence de ces échos. La réflectivité radar[10] est convertie en intensité de pluie, alors que le déplacement de la pluie est reconstitué entre deux images "instantanées" successives (advection).
- l'étalonnage permet d'ajuster l'intensité du signal radar en étalonnant l'image prétraitée avec les mesures pluviométriques au sol. Cet étalonnage s'appuie sur le traitement des pixels situés dans l'entourage des pluviomètres considérés (FAURE et *al*., 1994).

Des opérations de calcul de cumuls pluviométriques sont aussi effectuées sur ces données radar (horaire et quotidien).

L'ensemble des données relatives à l'imagerie radar est stocké pour une période de 7 jours et archivé pour une utilisation ultérieure. En temps-réel, ces données sont exportées sur un serveur distant et accessible pour visualisation par les services externes autorisés (SCHAPI, etc.).

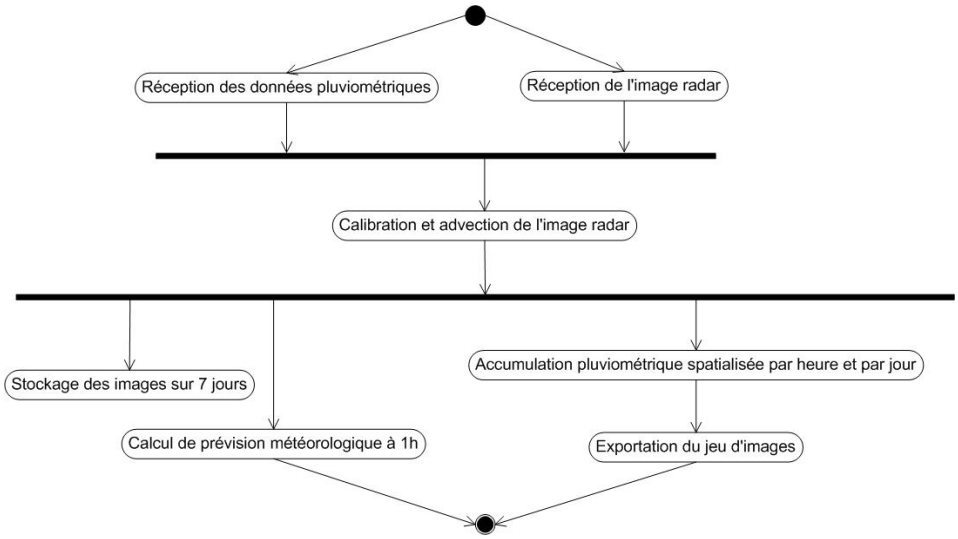


Figure 39 : Diagramme d'activité « gestion et analyse de l'imagerie radar »

Ce cas d'utilisation permet enfin de générer une prévision à une heure des champs de pluie. Afin d'établir cette prévision, la méthode utilisée se base sur les caractéristiques des échos des cellules, leur appariement avec des échos « imaginaires » contrôlé grâce à un arbre de décision et l'extrapolation des caractéristiques dans l'avenir (RHEA, 1998). Elle permet d'obtenir une estimation probabiliste de la pluie à un horizon d'une heure.

III - 1.2.3.3. Analyse des prévisions météorologiques et exportations de données hydrométéorologiques

[10] La relation réflectivité/intensité de pluie a été fixée empiriquement

Ces deux cas d'utilisation demandent peu de ressources informatiques, puisqu'ils correspondent à la réception des prévisions météorologiques de Météo France sous forme textuelle et le transfert de données produites au sein du système d'information du SPC-GD à des services externes. Ainsi, leur aspect technologique n'est pas approfondi dans cette étude de l'existant.

En situation de vigilance, le SPC-GD reçoit des bulletins Météo-France sur des zones AP (Avertissement Précipitation) (Fig. 40) qui détaillent pour chaque zone la moyenne des hauteurs de précipitations observées dans les 24 dernières heures et un intervalle de valeurs pour les hauteurs prévues à 24h et 48h d'échéance.

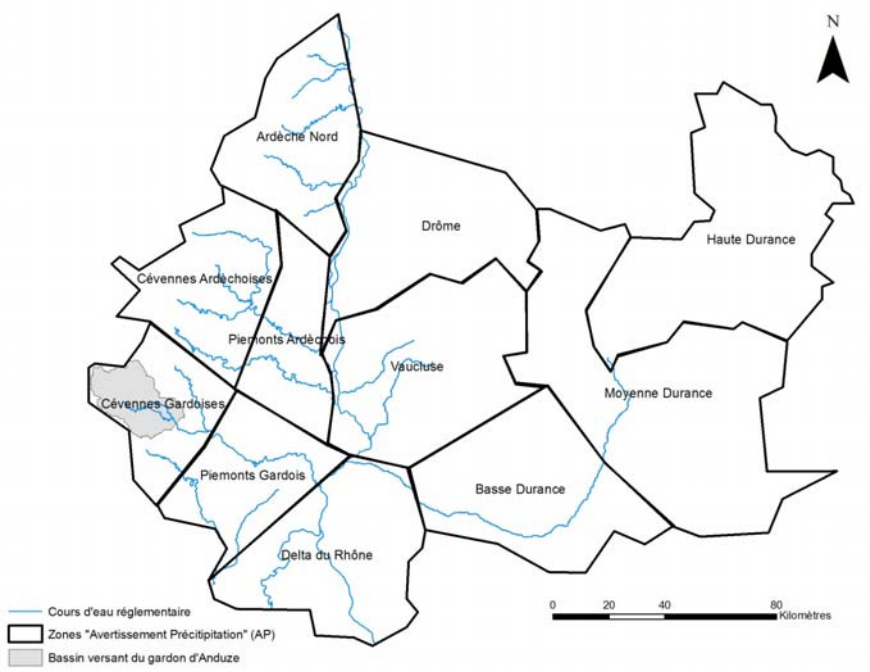


Figure 40 : Zones "Avertissement Précipitation" sur le territoire de compétence Grand Delta (source : SPC-GD)

Cette information de prévision, corrélée aux vigilances météorologiques, diffusée par Météo France, constitue, dans le fonctionnement existant, la seule véritable capacité d'anticipation de réponses hydrologiques, du fait de l'expérience des prévisionnistes sur les réponses hydrologiques des cours d'eau réglementaires.

Quant à l'exportation des données hydrométéorologiques vers les services institutionnels, elle constitue le seul moyen de partage informatique offert par le système, au moyen de transfert FTP[11].

III - 1.2.3.4. Modélisation du débit des cours d'eau

L'aspect modélisation représente la dernière fonction du système d'information (Fig. 41). L'application ALHTAÏR (ALarme Hydrologique Territoriale Automatisée par Indicateur de Risque) permet de compléter l'information relative à l'hydrométéorologie des bassins versants.

[11] File Transfert Protocole

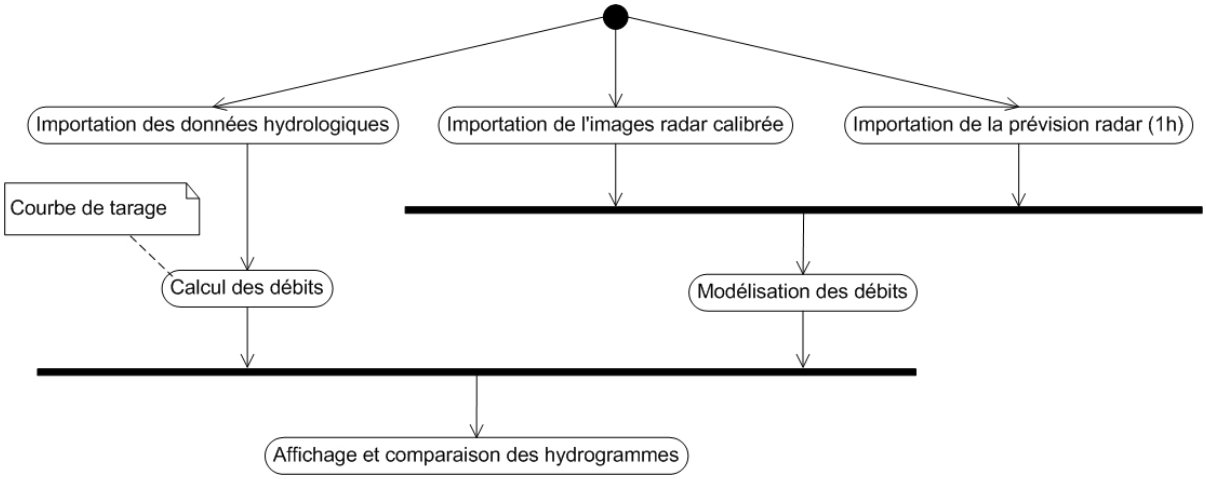


Figure 41 : Diagramme d'activité « modalisation du débit des cours d'eau »

Comme pour les autres cas d'utilisation, les données collectées sont nécessaires pour la phase de modélisation. D'une part, les courbes de tarage permettent d'approcher les débits potentiels grâce aux hauteurs d'eau, et d'autre part, les données radar temps-réel et en prévision constituent, avec les informations topographiques des bassins versants, les données d'entrée du modèle pluie-débit spatialisé ALHTAÏR.

La comparaison des hydrogrammes modélisés et ceux dérivés des observations, dont l'incertitude peut être malgré tout importante, permet de mieux préciser les débits en cours et envisager l'évolution de la situation dans un délai de quelques heures (décharge du bassin versant et prévision radar). Cet aspect « modélisation » est largement approfondi dans la suite de l'exposé puisqu'il constitue le cœur de la méthodologie proposée.

III - 1.2.4. Points de vue « Ingénierie » et « Technologique »

La dimension technologique de cette recherche requiert une attention particulière sur les moyens et les capacités techniques du système d'information du SPC-GD. Dans un premier temps, sa structure physique et ses relations fonctionnelles sont identifiées et mises en relation avec les cas d'utilisation concernés. Dans un second temps, la compétence « modélisation » est appréhendée à l'échelle de l'application ALHTAÏR.

III - 1.2.4.1. Description technique du système d'information

Afin de finaliser la description du système d'information, il s'agit de présenter sa structure physique et ses relations fonctionnelles en identifiant à chaque fois les cas d'utilisation concernés (Fig. 42).

Les données hydrométéorologiques en entrée proviennent :

- du réseau de mesures au sol collectées grâce à 3 concentrateurs (Privas, Avignon et Nîmes) par une connexion radio 1,4 GHz maillée et transmise à la partie « collecte » du serveur central SIGMA2000.
- des radars météorologiques au sol et transmises par satellite au serveur RETIM (Météo-France).

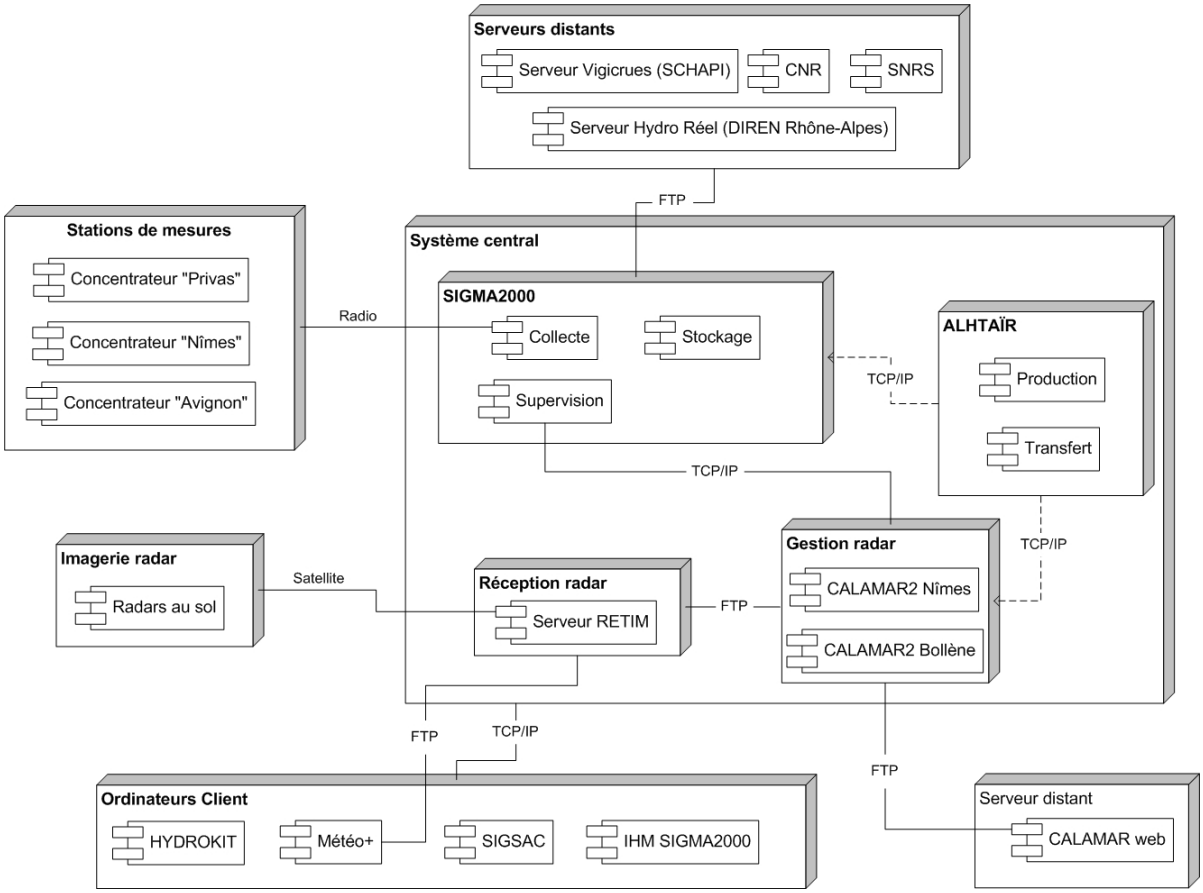


Figure 42 : Diagramme de déploiement du SPC-GD

Le serveur de concentration SIGMA2000, en particulier son composant « supervision », effectue la « gestion des données hydrométéorologiques » et gère les échanges de données entre les différents modules du service, et la transmission de ces données aux services externes (les serveurs Hydro Réel[12] et Vigicrues[13]). Il alimente les serveurs CALAMAR 2 (Nîmes et Bollène), développé par RHEA© (RHEA, 1998), en charge de la « gestion et d'analyse de l'imagerie radar ». De plus, ces serveurs accèdent aux données radars, stockées sur le serveur RETIM, par FTP et transmettent en temps-réel les images produites sur le serveur distant CALAMAR.

Enfin, le composant ALHTAÏR (nœud en UML) est en charge de la « modélisation du débit des cours d'eau ». Il est constitué de deux composants, la production et le transfert, propres à la modélisation hydrologique. Ce nœud utilise les résultats des opérations des nœuds SIGMA2000 et CALAMAR2.

Le prévisionniste s'appuie sur ces nœuds et les applications clientes disponibles sur une demi-douzaine de postes clients pour mener à bien son expertise :

- Météo+, logiciel développé par Météo France afin de visionner ses images radars et satellites, et les observations météorologiques (températures, vent, impacts de foudre, etc.),
- SIGSAC, développé en interne et basé sur le logiciel ARCGIS© (BLEUSE, 2001),

[12] http://www.rdbrmc.com/hydroreel2/
[13] http://www.vigicrues.ecologie.gouv.fr/

- L'IHM SIGMA2000, présente sur tous les postes clients, qui permet de suivre et de contrôler les opérations administratives du système,
- L'outil HYDROKIT® qui permet l'exploitation du Modèle Numérique de Terrain de la BD ALTI de l'IGN® afin de générer les informations utiles des bassins versants pour la modélisation (temps de concentration, temps de transfert, etc.). La résolution des bassins versants est de 50 m.

Cette architecture informatique est implantée dans la salle de crise de la DDTM. Il s'agit d'un réseau local (LAN) d'une bande passante de 1 GB/s. Les 3 serveurs d'application (SIGMA2000, CALAMAR et ALHTAÏR) sont des biprocesseurs de 1,5 GHz. La capacité totale de stockage du système est inférieure à 1 TB (820 GB). Ce système est relié à un réseau virtuel privé (VPN) à un débit de 4 Mb/s.

III - 1.2.4.2. L'application ALHTAÏR : « modélisation des cours d'eau »

Dans cette section, il s'agit de décrire le fonctionnement existant du composant « ALHTAÏR » en charge de la modélisation pluie-débit. L'application ALHTAÏR a été développée grâce au langage WINDEV©. Cette application est enrichie d'une interface graphique permettant au prévisionniste de paramétrer ces instances de modélisation (paramètres du modèle, liste des bassins versants, niveaux d'alerte et de pré-alerte, affichage de l'hydrogramme, etc.). Outre le cœur de l'application dédiée à la modélisation hydrologique, une série d'algorithmes concernent (Fig. 43) :

- les prétraitements :
    - la gestion des données d'entrée (bassin versant et image radar)
    - le calcul des correspondances géographiques entre zones radar et bassins versants
    - L'initialisation des conditions hydrologiques initiales
    - la lecture de la pluie mesurée ou estimée (prévision)
- Post-traitements :
    - La construction des hydrogrammes
    - Le contrôle des niveaux de pré-alerte et d'alerte

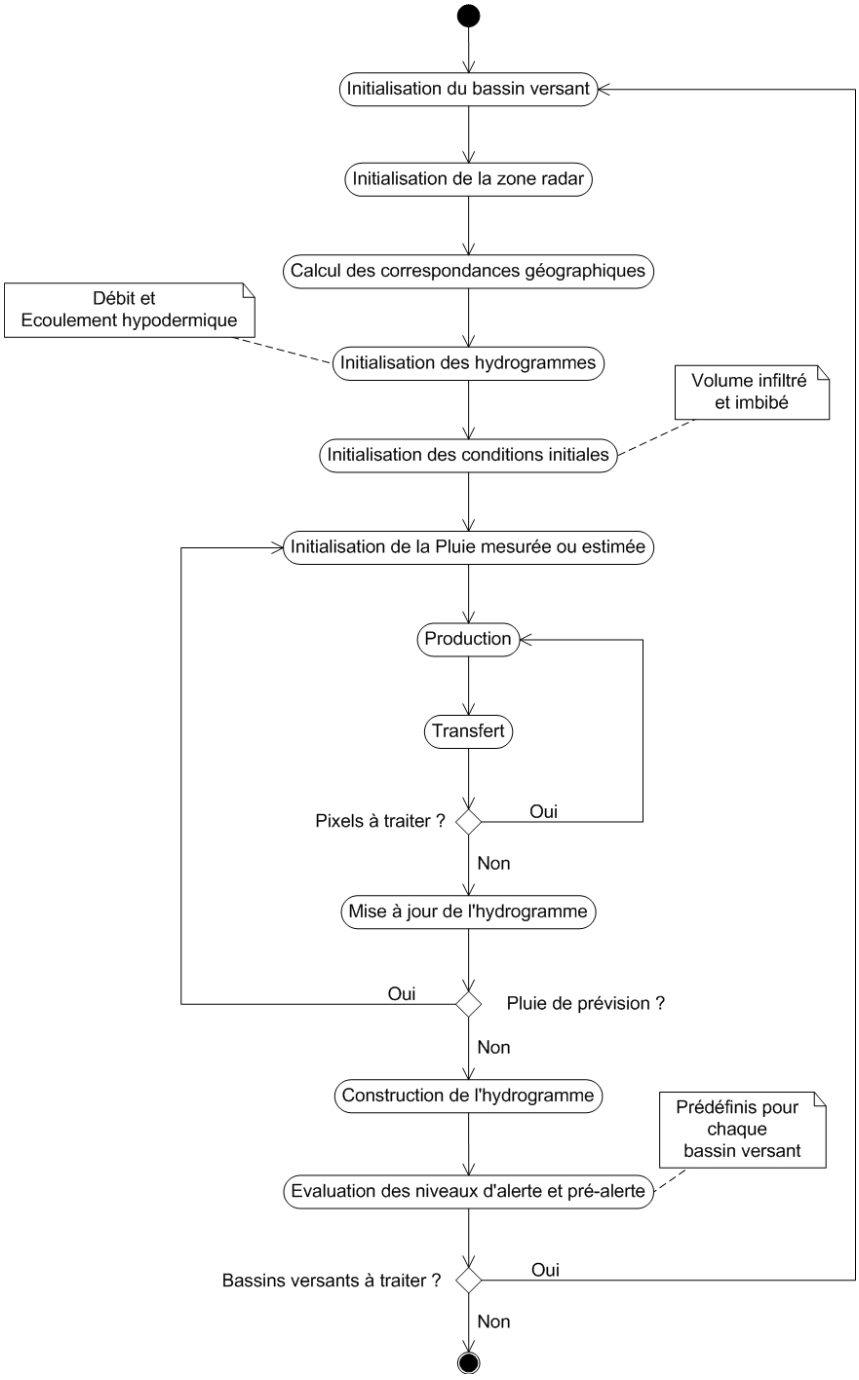


Figure 43 : Diagramme d'activité simplifié de l'application ALHTAÏR

L'application ALHTAÏR s'exécute en continu sur une trentaine de bassins versants (préalablement générés par HYDROKIT) situés dans la partie Est du territoire « Grand Delta ». Ces bassins versants sont traités séquentiellement en prenant en compte la pluie mesurée, et éventuellement celle prévue sur requête du prévisionniste, fournie par l'application CALAMAR 2. Les besoins en calcul pour cette tâche de modélisation étant relativement importants, le prévisionniste a la possibilité de diminuer la résolution géographique du bassin versant, dans ce cas la modélisation hydrologique ne traite qu'une cellule sur n cellules du bassin versant.

Dans le chapitre 1 (cf. §.I – 1.3.2.3), les principes hydrologiques des fonctions de production et de transfert ont été abordés, il convient de détailler le fonctionnement de ces modules d'un point de vue informatique (Fig. 44).

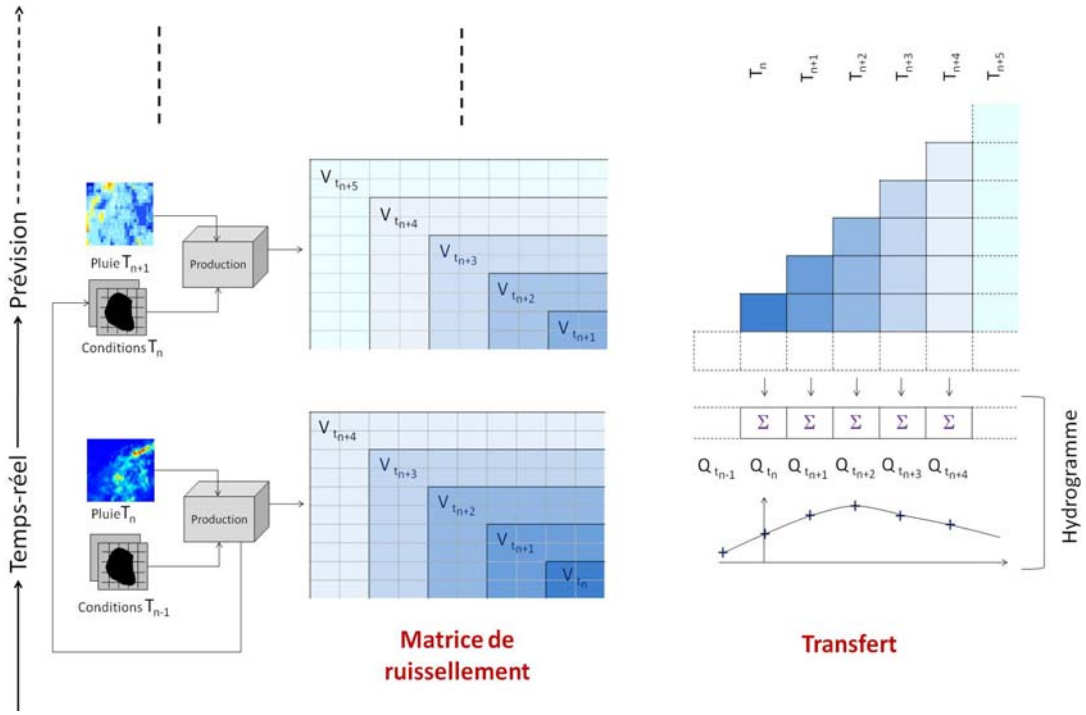


Figure 44 : Fonctionnement des modules de production et de transfert d'ALHTAÏR

Les deux données d'entrée évènementielles du module de production sont la pluie, et les conditions hydrologiques initiales, correspondant aux conditions hydrologiques modélisées au pas de temps précédent. D'un point de vue algorithmique, il n'existe aucune distinction algorithmique entre les modes « temps-réel » et « prévision », la seule différence intervient au niveau de la donnée de pluie en entrée. En effet, en mode « prévision », l'application ALHTAÏR prend en charge 12 matrices de pluie, correspondant à une prévision d'une heure fournie par RHEA (pas de temps 5 de minutes).

- La production permet donc de traiter chaque cellule du bassin versant de manière itérative, afin d'obtenir une matrice de ruissellement (V_{tn}).

- Le module de transfert traite cette matrice de ruissellement en sommant respectivement les volumes ruisselés des cellules ayant les même temps de transfert (intervalles de 5 minutes), de manière à construire l'hydrogramme. En pratique, les cellules, dont le temps de transfert est compris entre 0 et 5 minutes, correspondent au débit temps-réel (Q_{tn}), les cellules comprises entre 5 et 10 minutes au débit Q_{tn+1}, etc. Il en résulte une prévision à pluie nulle dont l'horizon est égal au temps de concentration du bassin versant.

En présence d'une pluie estimée, 13 itérations de modélisation (production et transfert) sont effectuées (1 pour le temps-réel et 12 pour la prévision). Les volumes ruisselés prévus grâce à la pluie estimée s'additionnent aux débits de la prévision à pluie nulle calculée durant le mode « temps-réel ».

Finalement, les grandes caractéristiques de ce cas d'utilisation sont :

- l'implémentation des ces modules au moyen du langage propriétaire WINDEV©,
- la rapidité d'exécution de ces modules (inférieur à 5 minutes),
- l'enchaînement presque séquentiel des tâches élémentaires,
- l'indépendance géographique des cellules du bassin versant dans leur traitement par les modules de production et de transfert.

De telles remarques permettent de mieux situer les contraintes informatiques de la modélisation pluie-débit du SPC-GD dans la perspective de l'utiliser au sein d'une architecture de grille, telle que celle d'EGEE.

III - 1.2.5. Bilan de l'analyse UML : premières hypothèses de recherche

L'analyse UML a permis de clairement spécifier le système d'information existant du SPC-GD à plusieurs niveaux d'abstraction. Elle a permis de positionner le système d'information dans son environnement opérationnel et de définir les fonctions principales (cas d'utilisation) dont elle est en charge en situation de crise afin de développer en continu une expertise hydrométéorologique à l'attention des services opérationnels de la Sécurité Civile et du pouvoir public (Fig. 36). L'objectif principal de cette analyse UML est d'envisager la grillification[14] de certaines tâches élémentaires du système d'information du SPC-GD. À la vue du diagramme de déploiement (Fig. 42), la structuration du système d'information en blocs indépendants permet *a priori* un déport de certains cas d'utilisation sur une architecture distante. Dans ce sens, les points de vue « Traitement » et « Ingénierie » (cf. §.III - 1.2.3) ont permis de détailler chacune de ces tâches afin de relever leur dynamique et leur dépendance fonctionnelle, et d'envisager le traitement de certaines d'entre elles sur une architecture de grille sans entraver le bon enchaînement du fonctionnement existant.

Les délais d'actualisation des données collectées (5 minutes) semblent limiter la prise en charge des traitements de ces données brutes sur une architecture distante, et par extension des fonctions relatives au fonctionnement temps réel du système d'information. D'après l'analyse UML et les discussions avec les prévisionnistes du SPC-GD, l'expertise hydrométéorologique est à l'heure actuelle principalement développée grâce aux cas d'utilisation « Gestion des données hydrométéorologiques » et « Gestion et analyse de l'imagerie radar ». En effet, l'expérience des prévisionnistes sur le fonctionnement hydrométéorologique du territoire de compétence Grand Delta permet, à la vue des réponses hydrologiques observées en temps réel et des caractéristiques quantitatives et qualitatives des cellules pluvieuses et de leur comportement, d'identifier l'intensité de l'aléa et *a fortiori* d'établir un niveau de vigilance « crue » conforme à la situation observée. La dimension modélisation (« Modélisation du débit des cours d'eau »), encore en cours de pérennisation (AYRAL, 2005), nécessite une validation à long terme de l'efficacité des modules hydrologiques et le développement d'une véritable capacité d'anticipation des pointes de crue. Dans ce sens, l'analyse UML a permis de mieux préciser les fonctions informatiques dédiées à la transformation de la pluie observée (radar) en débit à l'exutoire des bassins versants (modélisation) (cf. §.III - 1.2.4.2). Les caractéristiques fonctionnelles ainsi identifiées

[14] Terme utilisé pour la procédure d'implémentation d'algorithmes sur une architecture de grille (Soberman, 2003)

permettent dorénavant de confirmer la propension de ce cas d'utilisation à être pris en charge, au moins en partie, par des nœuds de calcul délocalisés.

Afin de spécifier précisément les caractéristiques organisationnelles et fonctionnelles de l'amélioration technologique attendue du système d'information du SPC-GD, il convient de compléter cette analyse UML avec l'identification de besoins récurrents identifiés dans le cadre de retours d'expérience sur des situations catastrophiques passées dues à des crues à cinétique rapide.

III - 2. Retours d'expérience : Besoins opérationnels, scientifiques et technologiques

III - 2.1. Les crues à cinétique rapide : les cas « Aude 1999 » et « Gard 2002 »

De nombreuses crues à cinétique rapide ont régulièrement lieu de part le monde. Certaines ne provoquent qu'une montée des eaux en ne mettant pas en danger les enjeux socio-économiques concernés. Dans des situations plus extrêmes, quand il est question de risque majeur, les dégâts occasionnés peuvent paralyser une région entière pendant plusieurs jours. ANTOINE et *al*. (2001) recensent une soixantaine de crues meurtrières, entre le XIVème siècle et la fin du XXème siècle dans le Languedoc-Roussillon, avec des caractères hydrométéorologiques variant d'une crue à l'autre, de même que les dégâts occasionnés. Durant la dernière décennie, deux évènements majeurs ont affectés le Sud-Est de la France, au point de provoquer la mort de dizaines de personnes et d'endommager les infrastructures vitales des départements de l'Aude et du Gard.

Les retours d'expériences de ces évènements sont variés et nombreux en termes de champs thématiques abordés. Les leçons tirées de ces analyses apportent des « leviers d'optimisation » qui permettent d'améliorer la gestion politique, organisationnelle et scientifique de ce type d'évènements.

En s'intéressant aux évènements du 12 et 13 novembre 1999 dans le département de l'Aude et du 8 et 9 septembre 2002 dans le département du Gard, de nombreuses clefs d'interprétation peuvent être identifiées et ouvrir la voie à des propositions d'améliorations et de perfectionnement des systèmes opérants de la Sécurité Civile et des administrations publiques rattachées.

III - 2.1.1. Évènement du 12 et 13 novembre 1999 dans le département de l'Aude

III - 2.1.1.1. Caractéristiques hydrométéorologiques

L'extension géographique de cet évènement dépasse les simples frontières du département de l'Aude, puisque les Pyrénées-Orientales, le Tarn et l'Hérault ont eux aussi été touchés par la violence exceptionnelle des intempéries de novembre 1999. Habituées à ce type d'inondations ravageuses, ces régions ont pourtant été confrontées durant ces deux jours à un évènement « rare et même exceptionnel dans son ensemble par l'intensité totale du phénomène et la surface touchée » (LEFROU et *al*., 2000). La caractéristique principale de cet évènement est l'accumulation importante de pluie dans un délai court, de l'ordre de 600 mm en 24h (Fig. 45) (GAUME et *al*., 2004) avec des intensités de pluie atteignant les 150 mm/h (VINET, 2003). Cet épisode se caractérise par deux pics pluviométriques successifs espacés de quelques heures dans la journée du 12 novembre provoqués par une forte activité orageuse. « Ces fortes précipitations ont été engendrées par un système convectif de méso-échelle typique de régions méditerranéennes » (VINET, 2003).

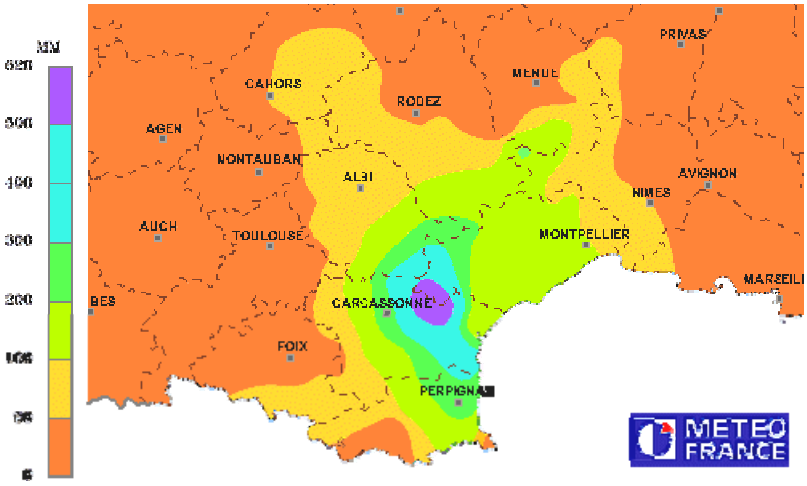


Figure 45 : Cumul des précipitations entre le 12 et le 13 novembre 1999
(source : Météo-France)

Évaluer la période de retour d'un tel évènement n'est pas chose aisée du fait en partie du manque de données climatiques historiques. Dans un certain sens les cumuls pluviométriques de cet évènement, bien qu'exceptionnels, sont régulièrement observés dans les zones touchées par ces évènements météorologiques. Le rapport de l'Inspection Générale de l'Environnement (IGE) sur cet évènement évalue approximativement cette date de retour entre 50 et 100 ans en moyenne sur l'ensemble de la région, ce qui permet de rendre compte de son exceptionnalité (LEFROU et *al.*, 2000). En outre, comme le rappelle Vinet (2003), l'extension géographique de cet évènement apparait comme le principal facteur explicatif de la situation de crise dans laquelle s'est retrouvée cette région géographique.

Les réponses hydrologiques de l'Orbieu et de l'Aude, et d'une grande partie des cours d'eau de la région, qui en résultent, sont tout aussi exceptionnelles par leur puissance et leur rapidité. Comme souvent lorsque les pluies se répartissent sur un large territoire, les types de crues sont nombreux (Vinet, 2003). Dans ce contexte, la fourchette des débits spécifiques (i.e. rapporté à la superficie du bassin versant) observés avoisinent la dizaine de $m^3.s^{-1}.km^2$ sur les bassins versants amont. Il en résulte un débit de crue aux abords des basses plaines de l'Aude de l'ordre de 4.500 $m^3.s^{-1}$. Ainsi, en moins de 12h, la plaine de l'Aude, elle-même touchée par de violentes précipitations, est envahie par les eaux des bassins versants amont. La présence de digues, correspondant au canal de jonction entre le canal du Midi et de la Robine qui barre le lit de l'Aude, a eu un double effet. D'une part, elles ont joué leur rôle de retenue d'eau protégeant d'un côté les villages situés en aval mais *a contrario* ont provoqué la montée des eaux sur les territoires voisins amonts. Dans une seconde phase, la rupture soudaine de ces digues, et les montées brutales des eaux qu'elles sont susceptibles de provoquer, a entrainé l'inondation des zones avals initialement protégées. Le dernier facteur aggravant concerne la présence d'un fort vent d'Est sur la Méditerranée qui a provoqué des surcotes marines empêchant l'évacuation normale de l'eau dans la mer (HAZIZA, 2007 ; VINET, 2003).

Ainsi, il est important de rappeler l'ambivalence des facteurs explicatifs de cette catastrophe, où l'exceptionnalité de l'aléa s'ajoute à la présence de territoires potentiellement vulnérables.

III - 2.1.1.2. Gestion de crise et dommages

Météo-France, comme généralement lors de catastrophes naturelles, émet le premier Bulletin Régional d'Alerte Météorologique (BRAM) à l'attention des services de Sécurité Civile et des Bulletins d'Alerte Précipitations (BAP) aux services d'annonce des crues concernés. Lors de cet évènement plusieurs BRAM ont été édités par Météo-France pour les 3 départements concernés. En suivant, et à la vue des premières réactions hydrométéorologiques et des dégâts occasionnés, dans la nuit du 12 au 13 novembre un plan ORSEC[15] a été mis en place pour chaque département touché par l'évènement (Aude, Pyrénées-Orientales et Tarn) (LEFROU et *al*., 2000 ; VINET, 2003).

Malgré la réussite apparente des opérations d'alerte et de secours, la soudaineté et l'intensité de ces phénomènes hydrologiques ont coûté la vie à 35 personnes alors que plus de 300 communes ont été sinistrées et ont mis pour certaines plus de trois ans pour effectuer leur reconstruction (Vinet, 2003). Le coût global de ces inondations a été estimé à environ 600 millions d'euros (LEFROU et *al*., 2000 ; VINET, 2003). D'une manière générale, une grande partie des réseaux de communications a été dégradée par submersion ou à cause de glissements de terrain, paralysant ainsi pendant quelques heures les zones sinistrées. Au total, plus de 30.000 foyers ont été privés d'électricité et de téléphone pendant toute la durée de l'évènement. Enfin, les réseaux d'alimentation en eau potable et d'assainissement ont eux aussi été fortement perturbés.

III - 2.1.2. Évènement du 8 et 9 septembre 2002 dans le département du Gard

III - 2.1.2.1. Caractéristiques hydrométéorologiques

Cet évènement est caractérisé par des conditions hydrométéorologiques assez proches de l'évènement de novembre 1999 puisque un système convectif de méso-échelle est aussi à l'origine de cet évènement (DELRIEU et *al*., 2005). Situés à proximité de l'arc méditerranéen, les départements du Gard, de l'Hérault, du Vaucluse, des Bouches du Rhône, de l'Ardèche et de la Drôme ont connu d'intenses précipitations alimentées par « de multiples super-cellules, qui se sont régénérées continuellement en restant quasiment stationnaires […] leurs mouvements convectifs ont refroidi un air chaud et humide provenant de la mer » (HAZIZA, 2007) (Fig. 46). Dans ce contexte, l'évènement a duré 28 heures entre le 8 septembre à 8h et le 9 septembre à 12h. Les cumuls de pluie observés, bien que très variables, restent exceptionnels pour ce type d'évènement, avec des valeurs extrêmes autour d'Anduze et Alès de 600 mm en 24h alors qu'un isohyète moyen de 200 mm sur une surface supérieure à 4.000 km^2 a été estimé (BOUVIER et *al*., 2004). Il est possible de distinguer trois phases météorologiques (DELRIEU et *al*., 2005):

- le système SCM a une orientation sud-ouest/nord-est et stagne sur le département du Gard, où d'importants cumuls ont lieu au nord d'une ligne Nîmes-Avignon entre 12h et 22h le 08 septembre,
- l'orientation devient Nord/Sud à partir de 22h et se positionne aux contreforts des Cévennes pendant 6h,

[15] Ce dispositif réglementaire, qui définit au préalable les procédures opérationnelles et les échanges interservices à effectuer en cas de catastrophe, permet une mise en place rapide des opérations de protection et de secours à la personne.

- au petit matin, le SCM commence à perdre en activité à cause de la présence d'un front froid et les précipitations encore conséquentes se décalent vers l'Est. A partir de 12h, les pluies ne concernent plus le département du Gard.

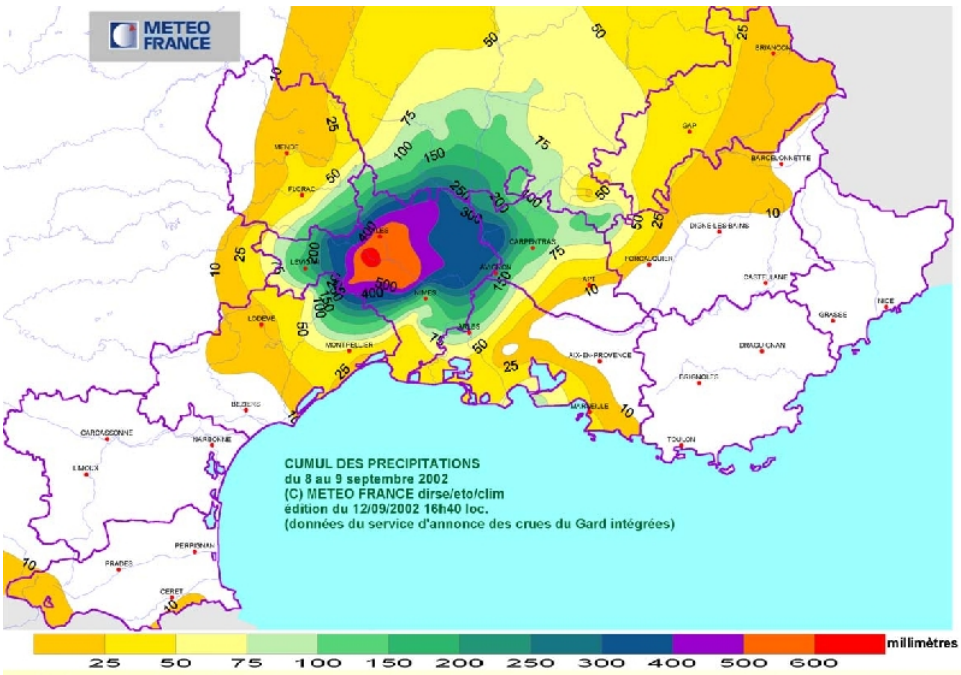


Figure 46 : Cumul des précipitations du 8 au 9 septembre 2002 (tirée de HUET et *al*., 2003a)

D'un point de vue hydrologique, la Cèze, le Gardon et le Vidourle sont les rivières dont la réaction hydrologique a été la plus forte (DELRIEU et *al*., 2005). Les réponses hydrologiques observées sont clairement en adéquation avec la répartition des pluies sur le département du Gard (GAUME, 2004). Ce sont les contreforts des Cévennes, et les plaines du Gard qui ont connu les plus fortes réactions hydrologiques. De la même manière qu'en 1999 dans l'Aude, des inondations de plaines ont fait suite à des crues torrentielles des bassins versants amont. Les plus fortes réactions ont été observées à l'aval du bassin versant d'Anduze, avec par exemple des débits spécifiques de l'ordre de 30 $m^3.s^{-1}.km^2$ sur le bassin versant de l'Alzon à Saint-Jean-du-Pin (HUET et *al*., 2003a). Il en résulte un débit de pointe de crue remarquable à la sortie des gorges du Gard à Remoulins estimé entre 5.000 et 7.000 $m^3.s^{-1}$. Concernant le bassin versant du Vidourle, le débit à l'exutoire à Sommières a été estimé entre 2.000 et 3.000 $m^3.s^{-1}$ (GAUME, 2004). Ces débits importants, dépassent ceux de l'évènement historique de référence de 1958, excepté sur les bassins versants cévenols amont. En effet, au pont Saint-Nicolas la hauteur d'eau relevée en 2002 dépasse de plus de 1,5 mètre la côte jusque là historique de 1958 (HAZIZA, 2007 ; HUET et *al*., 2003a). Le caractère exceptionnel de l'évènement du 8 et 9 septembre 2002 se confirme par la difficulté d'évaluation précise des débits en fonction des hauteurs d'eau, suivant les courbes de tarage établies préalablement (HUET et *al*., 2003a).

Ainsi, cet évènement peut être caractérisé par :

- des précipitations très intenses et des débits d'eau aux périodes de retour souvent supérieures à 50 ans,

- une grande extension géographique du phénomène

- une forte variabilité spatio-temporelle des précipitations.

III - 2.1.2.2. Gestion de crise et dommages

Ainsi, à la différence de l'évènement de 1999 dans l'Aude et comme cela a été décrit dans le chapitre 1, le cadre institutionnel de prévision météorologique a changé en 2001 (cf. §.I – 2.2.2.). Ainsi, en 2002, Météo-France diffuse des niveaux de vigilance accompagnés de bulletins descriptifs du phénomène météorologique. La première vigilance météorologique, de couleur orange, a été émise le 8 septembre à 05h53, annonçant des précipitations « de 30 à 50 mm en plaine et 70 à 150 mm sous orage fort » (HUET et *al*., 2003a). A compter de cette heure, des bulletins sont diffusés toutes les 6 heures à l'attention des services de Sécurité Civile, des services d'annonce des crues et du public. Sur le terrain, les premières demandes d'intervention ont lieu à milieu de journée du dimanche 8 septembre à Saint-Géniès-de-Malgoirès. A ce moment, la situation est déjà critique sur le bassin versant du Vidourle. Le 9 septembre à 1h27 du matin, le niveau de vigilance météo est élevé au niveau 4 (rouge) et des cumuls de précipitations de 300 mm sont annoncés. Après une accalmie, les précipitations redoublent au petit matin, au point que la situation oblige le préfet du Gard à déclencher le plan ORSEC à 10h16. Les difficultés opérationnelles sont nombreuses, « les gestes quotidiens les plus banals […] deviennent très difficiles » (SAUVAGNARGUES-LESAGE et SIMONET, 2004), et l'ensemble des réseaux de communications sont endommagés, au point qu'Alès et son centre de secours se retrouvent totalement isolés. Le maximum des interventions se situe dans l'après-midi du lundi 9 septembre (Fig. 47). Enfin, le mardi 10 septembre à 5h30, les digues d'Aramon (Rhône) et de Gallargues (Vidourle) cèdent sous la pression de l'eau provoquant de nouvelles victimes et d'importants dégâts.

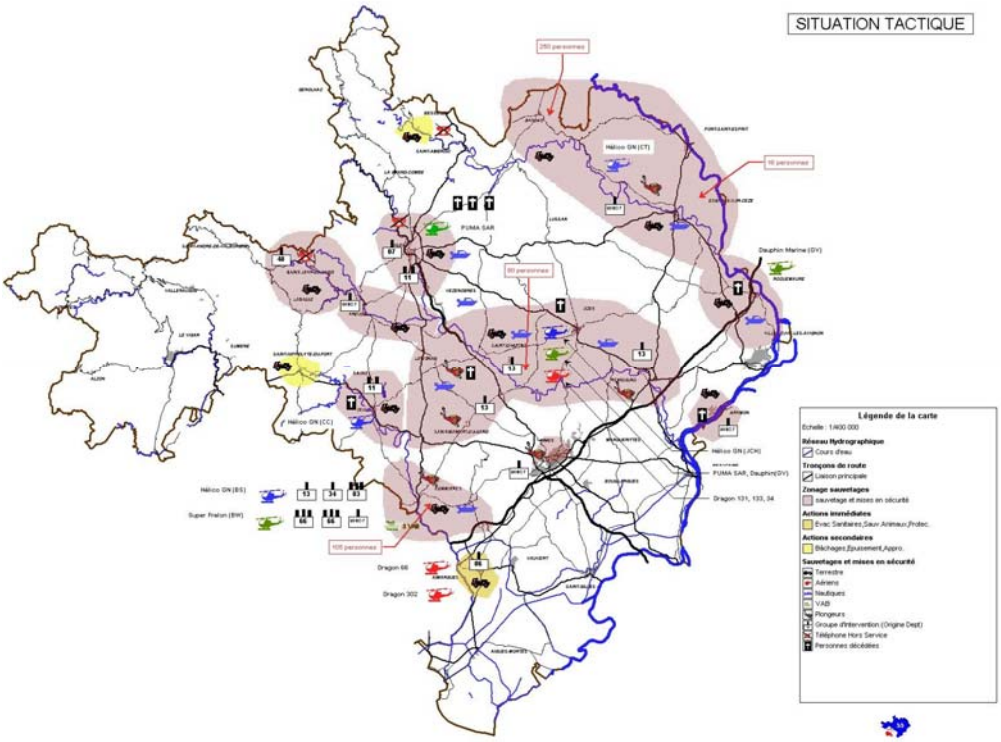


Figure 47 : Activité opérationnelle dans le Gard le lundi 09 septembre entre 12h et 18h (SDIS 30, 2002)

Le caractère exceptionnel de cet évènement se retrouve dans ses conditions hydrométéorologiques extrêmes et les dommages occasionnés qui ont paralysé l'ensemble du département du Gard pendant

plusieurs semaines après la catastrophe. De plus, 23 personnes sont décédées et 4.200 personnes ont été évacuées durant cet évènement. Le coût global de cette inondation a été évalué à 1,2 milliard d'euros sur les 6 départements touchés, plus de 85% des coûts sont imputés au simple département du Gard (Huet et *al*., 2003a). Dans ce même département, 295 communes ont été sinistrées soit 83% d'entre elles (SAUVAGNARGUES-LESAGE et SIMONET, 2004). « Il s'agit de la catastrophe naturelle la plus coûteuse parmi celles qui relèvent strictement du système d'indemnisation des dommages dus aux catastrophes naturelles (Catnat) » (HUET et *al*., 2003b).

III - 2.1.3. Synthèse sur les évènements gardois et audois

Ces deux catastrophes naturelles, en provoquant des situations de crise dans les services opérationnels et institutionnels de l'Etat et de la Sécurité Civile, ont fortement impacté les cadres institutionnels et opérationnels dédiés à la gestion de crise « inondation ». Les caractéristiques les plus importantes de ces deux évènements ont été synthétisées dans le tableau 3. La durée de chacun des évènements est basée sur la première mise en alerte (ou vigilance) météorologique et la levée de ces alertes.

Évènements / Caract.	12 et 13 Novembre 1999 dans l'Aude	8 et 9 Septembre 2002 dans le Gard
Durée de l'aléa	66 h (T0 : 11/11/1999 à 15h31)	40 h (T0 : 08/09/02 à 05h53)
Phénomènes météorologiques	système convectif de méso-échelle	système convectif de méso-échelle
Phénomènes hydrologiques	Crue torrentielle, inondation de plaine, rupture de digues, surcotes marines	Crue torrentielle, inondation de plaine, résurgences karstiques, rupture de digues
Débit spécifique maximal	20 $m^3.s^{-1}.km^2$	30 $m^3.s^{-1}.km^2$
Débit maximal instantané	4500 $m^3.s^{-1}$	5000-7000 $m^3.s^{-1}$
Périodes de retour (météo)	50-100 ans	> 100 ans
Heure approximative du débit de pointe[16]	**T33** (13/11/1999 à 01h00) Orbieu	**T26** (09/09/02 à 08h00) Gardons et Vidourle[17]
Alertes Météorologiques (Météo-France)	**T0** 1ers BRAM[18] et BAP[19] Aude, P.-O., Hérault 1er BRAM Tarn **T16** 1ère ALARME[20] Aude, P.-O., Hérault **T23**	**T0** Vigilance orange Gard (08/09/02 à 05h53) **T19** Vigilance rouge Gard (09/09/02 à 01h37) Bulletins (BRS[21]) émis toutes

[16] Les débits de point varient d'un cours d'eau à l'autre, l'heure retenue correspond aux débits maximaux observés
[17] Les évènements hydrologiques du Rhône du mardi 10 septembre n'ont pas été pris en compte
[18] Bulletin régional d'alerte météorologique émis par Météo-France
[19] Bulletins d'Alerte Précipitations émis par Météo-France
[20] Alerte aux risques météorologiques exceptionnels émis par Météo-France

		Au total 10 BRAM, 8 BAP, 3 ALARME et 1 communiqué de presse		les 6 h
Alertes hydrologiques	**T2** **T2** **T16** **T21**	Mise en vigilance du SAC de l'Aude (3 acquittements[22]) Mise en vigilance du SAC des P.-O. (3 acquittements et une mise en alerte à T25) Mise en pré-alerte du SAC du Tarn (2 échecs d'acquittement) Mise en vigilance du SAC de l'Hérault (3 échecs d'acquittement + contacts horaires avec la préfecture) Difficultés dans la remontée et la transmission de l'information	**T12** **T29**	Mise en vigilance du SAC du Gard (08/09/02 à 17h50) mises en alerte successives des bassins versants à partir de 19h (135 messages émis) Saturation du système d'observation (09/09/02 à 11h)
Dispositifs de Sécurité Civile	**T24** **T28** **T35** **T35** **T37** **T38**	Cellule de crise P.-O. Cellule de crise Aude Plan ORSEC Aude (du 13/11/99 à 02h30 au 15/12/99) Plan ORSEC P.-O. (13/11/99 à 02h30 au 15/11/99) PC de crise Hérault et Tarn Plan ORSEC Tarn (du 13/11/99 à 6h au 15/11/99)	**T14** **T28**	Ouverture du COD[23] (08/09/02 à 19h30) Plan ORSEC Gard (du 09/09/02 à 10h16 au 27/09/02 à 20h)
Dommages humains		35 morts, 6000 personnes évacuées et 435 communes sinistrées (26,3%)		23 morts, 4200 personnes évacuées et 295 communes sinistrées (83,5%)
Coût global (millions d'euros)		600		1200

Tableau 3 : Synthèse des caractéristiques hydrométéorologiques et opérationnelles des évènements dans l'Aude (1999) et dans le Gard (2002)

Plusieurs observations peuvent être extraites de cette synthèse et des paragraphes précédents :

- Ces deux évènements correspondent à deux situations hydrométéorologiques exceptionnelles de part les dégâts occasionnés et les débits maximaux instantanés mesurés[24].
- La mise en alerte des services opérationnels (Sécurité Civile et Préfecture) dépend fortement de la réactivité de Météo-France pour fournir des prévisions météorologiques.

[21] Bulletin régional de suivi émis par Météo France
[22] L'acquittement correspond à la validation de la réception d'un bulletin de prévision
[23] Centre Opérationnel Départemental
[24] Les valeurs indiqués sont issues des retours d'expérience et ont été réévalués et critiqués par les experts en hydrologie

- Les débits de pointe, souvent représentatifs de la période la plus menaçante pour les populations et les infrastructures, interviennent rapidement (approximativement 24h) après la mise en alerte météorologique, d'autant que ces observations occultent les réponses encore plus précoces des bassins versants amont souvent largement anthropisés.

- L'extension géographique de l'évènement dans l'Aude est très importante (plusieurs départements) mais ne doit pas occulter le maximum de dégâts dans la plaine de l'Aude.

- Le rôle des SAC semble s'être amélioré entre les deux évènements, avec une meilleure surveillance des cours d'eau à risque pour l'évènement de 2002. Cependant, les difficultés technologiques ont dans les deux cas limité l'efficacité de ces services, en particulier du fait d'une mauvaise remontée d'information depuis le terrain et un mauvais acquittement des bulletins météorologiques pour l'évènement de 1999.

- La difficulté d'anticipation de ces évènements, en particulier pour la dimension hydrologique, retarde la mise en place des réponses opérationnelles et limite ainsi l'efficacité des opérations de mise en sécurité et de secours aux personnes, du fait de l'exposition directe de ces services aux montées d'eau.

L'analyse de ces évènements et des réponses opérationnelles qui en résultent éclairent de plus le lecteur sur le fonctionnement et l'imbrication des réponses des services opérationnels tant au niveau de l'expertise hydrométéorologique que de l'organisation des secours et de la protection des populations. D'autre part, l'analyse des différentes défaillances permettent d'identifier les besoins relatifs à l'organisation de la gestion de crise dans son ensemble et donc les recommandations des experts pour parvenir à son amélioration. Dans ce sens, l'ensemble des ces besoins est appréhendé dans la partie suivante.

Il est important de signaler que l'ensemble des facteurs relatifs à la prévention des inondations peuvent prendre une importance non négligeable dans l'accentuation d'une situation de crise déjà engagée (HUET et *al*., 2003b) (changements d'usage, réseau routier croissant, destruction de bandes enherbées, mauvais entretien des cours d'eau, déviation des lits mineurs, etc.). Cependant, ils ne seront pas directement considérés dans l'analyse des besoins qui suit. Ce travail de recherche s'appuie avant tout sur l'hypothèse que l'éclosion d'une situation de crise « inondation » est principalement due aux conditions hydrométéorologiques en présence. Dans ce contexte, il semble primordial de relever, au travers des différents retours d'expérience et des recherches scientifiques sur la thématique, les leviers opérationnels et scientifiques susceptibles de perfectionner le suivi, l'anticipation et la caractérisation des crues à cinétique rapide par les experts et les « profanes ».

III - 2.2. De la Sécurité Civile à l'expertise hydrométéorologique : un perfectionnement perpétuel

III - 2.2.1. Définition de la typologie des besoins

Les principales sources bibliographiques utilisées dans cette revue sont des rapports parlementaires et des missions d'expertises produits par les administrations publiques et des documents scientifiques. Il est possible de relever deux grandes orientations des problématiques :

- la première concerne les besoins, d'ordre informationnel, des services de secours, inspirant les aspects stratégiques et tactiques de la crise,
- la seconde s'intéresse aux problématiques relatives à l'hydrométeorologie, de la mesure *in situ* à la modélisation numérique

Ces deux points de vue, bien que fondamentalement différents, argumentent conjointement pour un perfectionnement des méthodes et des outils de la gestion de crise. Ils suggèrent entre autre la mise en place de travaux de recherche multidisciplinaires conduisant au perfectionnement voire à la restructuration des systèmes opérants.

Afin d'identifier les besoins auxquels ces travaux cherche à répondre, il est possible de les hiérarchiser suivant les deux grandes missions opérationnelles, correspondantes aux missions des services opérationnels en charge de l'expertise hydrologique (SPC et SCHAPI) :

- le développement de l'expertise hydrologique composée de l'anticipation de l'aléa et sa caractérisation tout au long de l'évènement,

- la transmission de cette expertise aux services opérationnels de la Sécurité Civile, aux services de l'État et à la population

Afin de développer cette expertise, les hydrologues des services concernés s'appuient, d'une part, sur les prévisions météorologiques, fournies tout au long de l'évènement par les services de Météo-France, potentiellement responsables de l'augmentation du débit des cours d'eau, et les mesures *in situ* caractérisant les processus hydrologiques en cours (débit, hauteur d'eau, pluviométrie, etc.), et d'autre part, sur des prévisions hydrologiques à court terme simulées grâce à des modèles hydrologiques préalablement calibrés (Fig. 48).

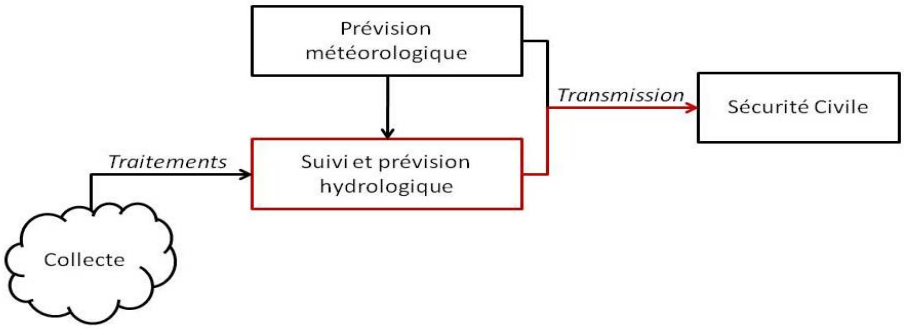


Figure 48 : Synopsis de l'expertise hydrologique en situation de crise (l'objet d'étude est représenté en rouge)

Comme le souligne MONTZ et GRUNTFEST (2002) « mieux nous comprendrons ces facteurs, et plus particulièrement les relations entre eux sur les petits bassins versants, plus grande sera l'amélioration de la prévision et de l'alarme ». Enfin, cette expertise discutée, critiquée et validée par le prévisionniste doit être régulièrement transmise aux services de la protection civile et plus largement à la population. A noter, que ces informations, originellement spécialisées, doivent être adaptées afin de fournir une vision synthétique et vulgarisée de l'évènement aux non-spécialistes.

En ce qui concerne la dimension « métier », qui engage cette thèse dans une orientation clairement opérationnelle, il est possible de fonder cette typologie des besoins sur deux fondements du système d'information que peut représenter le système d'aide à la décision permettant de développer l'expertise hydrométéorologique :

- celle relative à la collecte et à l'observation des données factuelles permettant d'appréhender le phénomène menaçant
- celle relative au traitement de ces données dans un objectif de modélisation et d'anticipation du phénomène.

Fortement liées dans le processus global d'aide à la décision, ces deux approches sont souvent différenciées dans un processus de conceptualisation d'un système d'information, comme par exemple dans le cas de la méthode de génie logiciel MERISE (TARDIEU et *al*., 1983). La première concerne la circulation des données et des informations dans le système d'information alors que la seconde implique des traitements informatiques visant à « faire parler les données ».

III - 2.2.2. Les missions de la Sécurité Civile : objectif ultime de la gestion de crise

En considérant que l'objectif ultime d'une gestion de crise est d'éviter des dommages à la société et son organisation (VINET, 2007), objectif utopique s'il en est, il apparaît primordial de considérer les besoins informationnels de la Sécurité Civile et de la population sur l'aléa, dans une situation de gestion de crise « inondation ». CLAUDET et BOUVIER (2004) identifient trois niveaux d'interlocuteurs en situation de crise :

- les institutionnels, acteurs prépondérants de la gestion de crise,

- les établissements et la population des zones inondables,
- le grand public dont la vulnérabilité est variable suivant ses activités.

L'information sur l'aléa doit donc être adaptée à chacun de ces protagonistes afin que chacun puisse accommoder son comportement et ses actions à la situation catastrophique. Ces considérations reviennent donc à identifier les informations (et leurs formes) que les services opérationnels de l'expertise hydrologique se doivent de fournir aux décisionnaires et au public. Il a été relevé, dans l'analyse de la littérature, plusieurs points critiques qu'il convient maintenant de détailler.

III - 2.2.2.1. Gestion des opérations de secours et prévision hydrométéorologique

III - 2.2.2.1.1 *L'information de prévision*

Un effort dans l'adaptation de l'information relative à la prévision hydrométéorologique, destinée aux décisionnaires de la gestion de crise, doit être mené, elle doit être en « cohérence avec les plans d'interventions » de gestion de crise (CLAUDET et BOUVIER, 2004). Cette information doit être régulièrement actualisée, au travers de bulletins concis et conformes aux compétences des services opérationnels et aux connaissances générales de la population (CLAUDET et BOUVIER, 2004 ; DROBOT et PARKER 2007 ; GRUNTFEST, 2007 ; HUET et *al*., 2003a). Cette amélioration de la pertinence des bulletins hydrologiques et météorologiques permet d'éviter des « confusions sémantiques » altérant la relation entre les données factuelles et le « niveau de crise à venir » (VINET, 2007).

III - 2.2.2.1.2 *Le partage et la diffusion des données*

Il parait essentiel de mieux intégrer et partager les observations, collectées depuis les réseaux de mesure, et les sorties de modèles, nécessaires à la gestion de la crue, en les rendant accessible à chaque acteur de la crise avec des niveaux d'interprétation adaptés (HUET et *al*., 2003a ; SAUVAGNARGUES-LESAGE et SIMONET, 2004 ; SAUVAGNARGUES-LESAGE et *al*., 2007a). En effet, il 'agit « d'améliorer la mise à disposition de l'information vers les responsables en charge de l'organisation des secours » (SAUVAGNARGUES-LESAGE et *al*., 2007a) comme le prévoit la réforme de la loi du 30 juillet 2003 destinée à moderniser le système d'annonce des crues français. Le développement de réseaux de communications qui permettent la création de « circuits d'alerte et d'information courts et redondants » apparait primordial (HUET et *al*., 2003a). Au travers de cette remarque, il s'agit plus largement d'affermir la politique de partage et de diffusion de l'information critique entre les acteurs de la crise. Clairement multidisciplinaire, la gestion de crise et notamment celle associée aux crues éclair, requiert le recoupement, de manière intégrée, de données, d'informations et de compétences fortement hétérogènes (DROBOT et PARKER, 2007 ; DUBAND, 2000 ; GRUNTFEST et HANDMER, 2007 ; HANDMER et *al*., 2007 ; HAZIZA, 2007 ; HUET et *al*., 2003a ; MONTZ et GRUNTFEST, 2002)

III - 2.2.2.1.3 *Le rôle prépondérant de l'information géographique*

La pérennisation de l'utilisation de l'information géographique et plus largement d'outils spatiaux d'aide à la décision est une composante essentielle de la gestion de crise. La littérature argumentant dans ce sens est conséquente. Cette synthèse permet d'apprécier la part de plus en plus prépondérante de

l'utilisation de supports visuels et plus spécifiquement de systèmes d'information géographique, dans les processus d'aide à la décision de la gestion de crise. De prime abord, l'intérêt de la représentation graphique est d'ordre cognitif, comme l'affirmait Confucius « une image vaut mille mots ». L'utilisation de l'information géographique en gestion de crise est multiple. Il est possible de dégager trois finalités principales de l'utilisation de SIG par la Sécurité Civile (SAUVAGNARGUES-LESAGE, 2001) :

- la gestion de la prévention où le SIG permet de rassembler les informations réglementaires
- la prise en charge de la prévision de l'aléa en analysant « les conditions d'éclosion et de développement du sinistre »
- la gouvernance des opérations de secours

Ainsi, du point de vue de la Sécurité Civile ou des décisionnaires, un outil de « cartographie évènementielle » en temps-réel permet a priori une meilleure visibilité de la situation tactique et stratégique du terrain en augmentant les capacités de surveillance et de contrôle du système opérant (SAUVAGNARGUES-LESAGE, 2001 ; ISTED, 2005 ; PORNON, 1992). Il offre en outre une harmonisation dans la remontée d'information (liaisons cartographiques entre les salles opérationnelles et les postes de commandement sur site) en admettant que les services de secours disposent de systèmes embarqués, il est alors question de télégéomatique (Fig. 49).

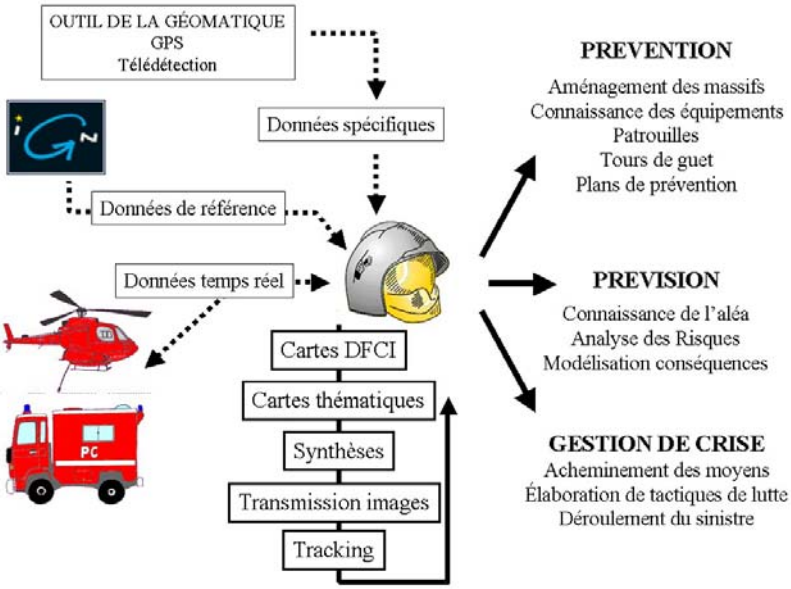


Figure 49 : La télégéomatique chez les pompiers (SAUVAGNARGUES-LESAGE, 2001)

Concernant l'appréhension de l'aléa, le recours à un système d'information géographique offre la capacité à intégrer l'ensemble des données collectées sur la zone menacée (HANDMER et *al*., 2007 ; SAUVAGNARGUES-LESAGE et AYRAL, 2007), de modéliser des scénarios d'évolution de l'aléa hydrométéorologique (FEIDAS et *al*., 2007 ; RABUFFETTI et BARBERO, 2005 ; SUI et MAGGIO, 1999) et de confronter le champ d'action de l'aléa avec les enjeux socio-économiques du territoire concerné (CUTTER, 2003). Sur ce dernier point, il s'agit, dans le cas des inondations, d'évaluer au mieux le niveau de risque en temps-réel en croisant l'extension du champ d'inondation (aléa) avec l'occupation du sol et a fortiori les enjeux matériels de la société (vulnérabilité). Finalement, couplé aux nouvelles technologies que représentent

Internet ou les architectures informatiques distribuées, le SIG facilite la diffusion et la mise à disposition de l'information à un réseau de partenaires (AL-SABHAN et *al*., 2003). Comme le confirme COVA (1999) en substance, « l'espace géographique est un cadre cohérent pour la réflexion sur les problèmes soulevés dans le contexte de la gestion d'urgence » alors que BRUGNOT (2001) rappelle que « les systèmes d'information géographiques, et plus particulièrement les cartes qui en sont extraites, constituent un support de transmission de l'information indispensable à la décision ».

III - 2.2.3. L'expertise hydrométéorologique : vers la modélisation numérique performante en situation opérationnelle

Il convient dorénavant d'axer l'analyse des besoins sur le fonctionnement des services opérationnels de l'expertise hydrologique, objet de recherche principal de cette recherche. Ces besoins concernent d'une part la gestion des données brutes collectées par les réseaux de mesure et d'autre part le traitement de ces données afin d'enrichir le corpus des connaissances relatif à l'aléa en cours. Il existe une certaine interdépendance avec les besoins de la Sécurité Civile abordés précédemment.

III - 2.2.3.1. Besoins relatifs à la gestion des données

Ces besoins sont relatifs à la gestion des données collectées durant la situation de crise. Ils concernent leur intégration parmi l'ensemble des partenaires institutionnels, la pérennisation des réseaux de mesure et la validité des informations extraites. D'autre part, un effort d'analyse de ces données semble nécessaire dans la phase post-événementielle d'une situation de crise.

III - 2.2.3.1.1 *L'intégration de l'expertise*

Dans la même philosophie que les besoins de la Sécurité Civile, il apparaît primordial de développer des outils et des plateformes de gestion de crise communs aux différents services en charge de l'expertise hydrométéorologique (HANDMER et *al*., 2007 ; SAUVAGNARGUES-LESAGE et *al*., 2007a). L'intégration des données collectées, des sorties de modèles hydrologiques (et hydrauliques) et des expertises s'impose comme une nécessité accrue de la gestion de crise « crue éclair » (DUBAND, 2000 ; GRUNTFEST, 2007 ; HUET et *al*. (2003a) rappellent l'indispensabilité que les services hydrologiques et météorologiques disposent « en temps réel des informations issu de tous les réseaux d'observation ».

III - 2.2.3.1.2 *Les réseaux de mesures*

Il convient de développer « un réseau fiable et pérenne de stations de mesure » (DUBAND, 2000) afin de disposer de données hydrométéorologiques en temps-réel. Les systèmes de mesure hydrologique sont particulièrement vulnérables lors de crues torrentielles à forte capacité de charriage. La perte, l'arrêt d'émission ou le dérèglement de ces appareils est fréquent. Cependant, il n'existe pas de solutions véritablement efficaces, la généralisation de l'imagerie radar, moins vulnérable aux aléas, apparait être une voie fiable (FAURE et *al*., 1994 ; KELSCH et *al*., 2001), alors que la densification des stations de mesure au sol implique de forts investissements financiers et humains.

III - 2.2.3.1.3 *Les courbes de tarage*

Les crues extrêmes, comme celles dans l'Aude en 1999 et dans le Gard en 2002, révèlent une difficulté latente dans l'établissement d'une relation univoque entre le débit et la hauteur d'eau (CLAUDET et BOU-

VIER, 2004). En situation extrême, « les côtes se trouvent dans la zone extrapolée des courbes de tarage » (HUET et *al*., 2003a), ce qui rend l'évaluation des débits difficile (avec une marge d'erreur de 10 à 20% (VINET, 2003) mais fournit paradoxalement, après réévaluation, de nouvelles valeurs de référence. Dans ce contexte, le retour d'expérience de l'IGE suggère de s'appuyer dorénavant sur les PHEC (Plus Hautes Eaux Connues) pour évaluer les périodes de retour hydrologiques (HUET et *al*., 2003a).

III - 2.2.3.1.4 *Les données historiques*

La mise en place d'une telle plateforme doit inciter l'implémentation de bases de données historiques sur les données gravitant autour de la problématique des inondations. Il convient ainsi « d'assurer la critique, l'archivage, la mise à jour, la consultation, la représentation et l'extraction de cette masse d'information » (Ambroise, 1999). Au travers de cette base de connaissances rémanente et fortement diversifiée, en termes de types de conditions hydrométéorologiques, la spécification objective et univoque de l'aléa en situation de crise semble envisageable. L'intérêt d'une telle base de données se situe dans la capacité à développer des méthodes d'analogie afin de générer des prévisions météorologiques probabilistes d'une part et mieux spécifier les périodes de retour des évènements (DELRIEU et *al*., 2005).

III - 2.2.3.1.5 *Les retours d'expérience*

De la même manière, il convient de systématiser les retours d'expériences hydrologiques et/ou météorologiques afin d'alimenter en continu les bases de données et notamment établir « un contrôle et une analyse systématique des écarts prévision-réalisation » (DUBAND, 1990). Ces études *a posteriori* doivent établir des contre-validations et des inter-comparaisons parmi tous les bassins versants concernés, afin d'évaluer et d'améliorer les modèles opérationnels (GAUME et *al*., 2004). En complément, de l'analyse des données factuelles, de tels retours d'expériences doivent enfin permettre la mutualisation des interviews, des résultats de prospections « terrain » d'après crise ainsi que de tous les documents multimédias disponibles (GAUME, 2004). Dans le cadre des crues cévenoles, ce type de documents s'est avéré indispensable pour la caractérisation précise de l'aléa, étant donné l'endommagement fréquent des systèmes de mesure en phase de crue.

III - 2.2.3.2. Besoins relatifs aux traitements et à la modélisation des données

Comme cela a déjà été abordé, les opérations relatives à l'expertise hydrométéorologique impliquent le traitement et l'utilisation de ces données factuelles à des fins de modélisation et de caractérisation des réponses hydrologiques (cf. §.I – 1.3.). Dans ce sens, les retours d'expériences ont permis de relever plusieurs difficultés opérationnelles, telles la difficulté de prévision et la validité des modèles, qu'il convient de présenter.

III - 2.2.3.2.1 *La prévision d'ensembles*

La prédiction des crues éclair revêt de nombreuses incertitudes du fait en partie de l'inconsistance de la prévision météorologique à très court terme (quelques heures) et sur des zones géographiques réduites (quelques km²) (BEVEN, 2001) et de la complexité des processus hydrologiques en jeu, l'impréci-

sion des conditions hydriques initiales et les imperfections intrinsèques des modèles hydrologiques (AMBROISE, 1998 et 1999 ; BEVEN, 2001 ; DOSWELL et *al.*, 1996 ; RABUFFETTI et BARBERO, 2005). La figure 50 illustre les conséquences de l'incertitude sur la pluie prévue sur la performance de la modélisation hydrologique au sein du SPC-GD. En effet, si le débit simulé (c'est-à-dire basé sur la pluie évaluée à partir de données radar « temps réel ») se rapproche du débit mesuré (sur lequel une incertitude existe), la précision des hydrogrammes prévus à trois horizons différents tend à se dégrader avec l'augmentation de l'horizon de prévision.

Dans ces conditions, il devient difficile pour un prévisionniste du SPC-GD de diffuser une prévision hydrologique précise au-delà d'un délai d'une heure.

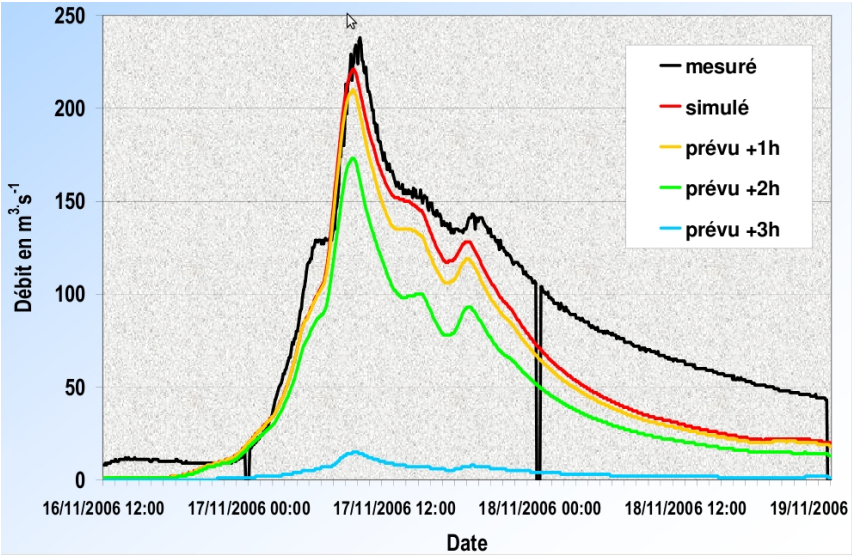


Figure 50 : Hydrogrammes mesuré, simulé et prévus par modélisation de la prévision de la pluie (1h, 2h et 3h) et une hypothèse de pluie nulle de la Beaume à Rosières avec le modèle ALHTAÏR (Tiré de PAYRASTRE, 2007)

Évaluer et diffuser en temps réel l'incertitude ou un « intervalle de confiance » de l'expertise hydrométéorologique aux décisionnaires apparaît comme une solution opérationnelle pertinente (AMBROISE, 1999 ; BEVEN, 2001 ; DUBAND, 2000 ; MONTZ et GRUNTFEST, 2002). Une autre issue opérationnelle concerne la simulation de différents scénarios météorologiques à différents délais de prévision basés sur des conditions initiales pouvant être définies par les experts pour simuler les débits correspondants (CLAUDET et BOUVIER, 2004 ; DUBAND, 2000). Cette approche s'apparente peu ou prou à la prévision d'ensembles utilisée dans la prévision météorologique.

III - 2.2.3.2.2 *Le calage en temps différé et l'assimilation en temps-réel*

L'une des étapes indispensables de l'utilisation d'un modèle en situation opérationnelle concerne le calage (étalonnage) de ses paramètres et correspond globalement à « vérifier la qualité d'un modèle en simulation » (AMBROISE, 1999). En fonction du modèle utilisé, la part entre les paramètres ayant une signification physique et les paramètres conceptuels, dépourvus d'une signification physique, varie. Ainsi, la phase de calibrage permet de fixer en laboratoire les valeurs de ces paramètres conceptuels. L'une des difficultés de la modélisation hydrologique est d'établir l'ensemble des paramètres d'un modèle permettant de simuler des situations hydrologiques diverses (MARCHANDISE, 2007). Cette situation

a pour conséquence une dérive du modèle aboutissant parfois à un fort décalage avec la situation réelle. L'une des solutions envisagées est d'assimiler en temps réel les mesures collectées in situ (BEVEN, 2001). Ainsi, le SCHAPI souhaite « améliorer la fiabilité des modèles météorologiques et hydrologiques » en soutenant « le développement de procédures d'assimilation d'un nombre de plus en plus grand de données d'observation » (TANGUY et *al*., 2004).

III - 2.2.3.2.3 *La valorisation de la modélisation opérationnelle*

Il apparaît indispensable d'intégrer des modèles hydrauliques et hydrologiques parcimonieux et robustes en vue de gérer l'hétérogénéité hydrologique des bassins versants (AMBROISE, 1999 ; DUBAND, 2000) (cf. §.). En mettant à disposition des services hydrologiques institutionnels une gamme de modèles hydrologiques pour simuler le débit aux exutoires des bassins versants amonts et de modèles hydrauliques pour calculer les transferts d'eau dans les plaines aval et *a fortiori* l'extension géographique des champs d'inondation ; ces derniers modèles pourraient accroître la capacité à anticiper les préjudices causés par les crues à cinétique rapide (BEVEN, 2001). Il en résulte le besoin de développer des capacités de discrétisation spatio-temporelle pour permettre une meilleure considération de l'hétérogénéité des processus hydrologiques locaux. Par exemple, une résolution spatiale fine pourrait être nécessaire sur les aires contributives alors qu'une résolution temporelle fine est nécessaire durant les évènements hydrologiques extrêmes au contraire des situations hydrologiques normales (Ambroise, 1999). Cette approche implique la complémentarité et la concertation entre ingénierie et recherche hydrologique en intégrant les « observation et mesures, expérimentation sur le terrain et au laboratoire fort de formalisation et de modélisation » (AMBROISE, 1999 ; DELRIEU et *al*., 2005). Finalement, comme cela est démontré dans l'article de VINET (2007), entre 1996 et 2006 dans la partie française de la Méditerranée, les victimes des inondations furent plus souvent localisées sur des bassins versants non jaugés. Ainsi, étendre la modélisation à ces bassins versant apparaît essentiel, comme le propose et l'expérimente Ayral sur les bassins versants amont des Cévennes dans ses travaux de recherche (AYRAL, 2005 ; AYRAL et *al*., 2007).

III - 2.2.3.3. Bilan des besoins

Qu'ils soient opérationnels, scientifiques ou politiques, ces besoins ont progressivement incité de nombreuses initiatives et perfectionnements du système opérationnel en place. Les réformes successives des services institutionnels, en particulier celle de la loi du 30 juillet 2003 relative à la modernisation des SAC ont concouru à une meilleure prise en compte de l'aléa « inondation », grâce notamment aux cartes de vigilance et aux bulletins de prévision mieux adaptés. Il subsiste cependant des difficultés opérationnelles nécessitant un perfectionnement continu des systèmes opérants.

L'évaluation du risque « inondation » en temps réel, et en particuliers celui lié aux crues à cinétique rapide, par les services institutionnels passe par le partage de l'information préventive et de prévision au travers de circuits de communication intégrés et redondants. Dans ce sens le recours à des méthodologies et des outils harmonisés et appropriés aux compétences de chaque partenaire apparaît indispensable. Ainsi, aux regards des retours d'expérience et des expertises précédents, l'information géographique et les systèmes d'information géographique, permettant sa prise en charge, semblent pouvoir offrir une solution efficace pour l'intégration des différents points de vue opérationnels de la gestion

de crise. En outre, les supports cartographiques qui en résultent, grâce à des approches sémiologiques différentes, permettent de diffuser une information adaptée à chaque compétence des acteurs opérationnels.

D'autre part, la meilleure prise en compte des évènements hydrologiques extrêmes à long terme impose une analyse hydrométéorologique intégrant différentes sources d'information relatives aux phénomènes tant du point de vue scientifique qu'opérationnel. Ainsi, la pérennisation de ces données, de ces observations et des outils qui facilitent leur interprétation, au sein de la communauté « hydrologie » des crues à cinétique rapide semble pouvoir garantir une meilleure compréhension de cet aléa. Cette appréhension précise des processus hydrologiques correspond *a priori* à la mission principale de ces services. Cependant, les efforts scientifiques et technologiques qu'elle implique, limitent une résolution complète de cette problématique à court terme. Néanmoins, en les dotant de solutions technologiques capables d'intégrer à grand échelle d'une part les données hydrométéorologiques relatives à des évènements passés et d'autre part une capacité de modélisation flexible, adaptable et facile d'utilisation ; cette mission critique pourrait progressivement être mieux appréhendée. Dans ce sens, une capacité de modélisation hydrométéorologique multiforme doit a priori être généralisée à tous les services de cette communauté.

En conséquence, l'émergence et la pérennisation croissante de technologies informatiques distribuées, telles qu'Internet et la technologie grille, motivent leur expérimentation dans le cadre de ces opérations d'expertise hydrométéorologique. En effet, le pouvoir de collaboration et de partage ainsi que les capacités calculatoires et de stockage qu'elles procurent s'alignent avec les exigences redondantes de la gestion de crise. Finalement, en connectant des communautés administrativement séparées, ces technologies semblent pouvoir assurer un support cohérent aux besoins d'intégration des retours d'expériences, des données historiques et surtout des données « temps réel ». Le gain de capacité de calcul offert par la technologie grille pourrait d'autre part offrir une solution viable au besoin croissant de modélisation hydrologique.

Cet éventail de besoins, bien que relativement exhaustif, montre les interdépendances existantes entre eux. En s'attachant à perfectionner certains besoins critiques, ce travail de recherche participe activement à l'évolution nécessaire et permanente des méthodes et des outils dédiés à l'expertise de l'aléa et plus largement à la gestion de crise. De plus, il cherche à vérifier ou infirmer des voies technologiques relativement nouvelles, à savoir la technologie grille, en évaluant la faisabilité d'un transfert technologique vers le monde de l'opérationnel.

Ainsi, la synthèse de ces besoins et l'objectif principal de perfectionnement de l'expertise hydrologique des crues à cinétique rapide ont orienté la problématique centrale de cette recherche vers l'enrichissement des capacités d'anticipation des crues éclair du SPC-GD afin de répondre aux objectifs énoncés par la loi du 30 juillet 2003. D'un point de vue opérationnel, cette amélioration semble devoir intégrer en définitive :

- une capacité d'anticipation :

- une prévision hydrologique multi-échelles et multi-modèles en vue de prendre en compte l'hétérogénéité des processus hydrologiques, ce qui implique l'augmentation du nombre de bassins versants à modéliser (DUBAND, 2000 ; RABUFFETTI et BARBERO, 2005)
- une prévision météorologique multi-horizons et multi-sources en vue de compenser l'incertitude des prévisions météorologiques à petite échelle (BEVEN, 2001; RABUFFETTI et BARBERO, 2005)

- une capacité de partage et de transmission de l'expertise hydrométéorologique en temps-réel

SYNTHESE

En début de chapitre, l'étude de l'existant a permis de relever des composants informatiques, et les compétences qu'ils permettent, susceptibles d'être impactés par l'adoption de la technologie grille. Le cas d'utilisation dédié à la modélisation hydrologique apparait comme le plus apte à subir des modifications dans l'enchaînement de ses activités (traitements). En effet, l'extraction des modules hydrologiques de production et de transfert de l'application ALHTAÏR, et leur exécution paramétrée sur des ressources distantes semblent une orientation technologique préservant l'intégrité du système existant.

Dans la même philosophie, l'analyse des besoins de la Sécurité Civile, et plus particulièrement ceux de l'expertise hydrologique ont fait émerger des besoins relatifs à la prévision hydrologique, et au partage des données et des expertises. Ainsi, à la vue des résultats de cette analyse, la prévision d'ensemble, basée sur des différents scénarios hydrométéorologiques, apparait comme une orientation de développement à approfondir pour enrichir les capacités de prévision du SPC-GD.

Ces deux nouvelles compétences, relatives à la modélisation et au partage en temps réel, qui ont émergé de l'analyse des besoins, constituent le postulat scientifique principal de cette recherche. De plus, elles permettent indirectement, de part les dépendances opérationnelles qu'elles ont avec l'ensemble des autres besoins, de légitimer leur sélection pour le perfectionnement global de l'expertise hydrologique en situation de crise et la résolution de la problématique générale, relative à l'adoption de la technologie grille.

Il s'agit dorénavant d'identifier les contraintes technologiques qu'elles engendrent et les potentialités techniques de la technologie grille et des services web afin de concevoir un système spatial d'aide à la décision hydrologique apte à répondre à l'ensemble de ces besoins.

Chapitre 4

La technologie grille et la géomatique : vecteurs d'innovation

4.1. Les contraintes technologiques de la prévision hydrologique

4.2. La grille EGEE : un support informatique performant et adaptable

4.3. G-ALHTAÏR : un système spatial d'aide à la décision hydrologique

Les besoins actuels de l'hydrologie opérationnelle, centrées sur la nécessité d'anticiper précisément l'occurrence d'une pointe de crue en temps et en lieu, permettent de définir les contraintes technologiques à respecter pour mener à bien le développement de solutions informatiques adaptées. Le socle informatique générique et robuste offert par la technologie grille permet d'entrevoir un respect de ces contraintes et un perfectionnement des missions du SPC-GD.

INTRODUCTION

Le chapitre 3 a permis de définir les opérations informatiques et les besoins du service de prévision des crues « Grand Delta » (SPC-GD), en matière de prévision hydrologique des crues à cinétique rapide en situation de crise. Cet état des lieux doit donc amener à l'appréhension précise des contraintes fonctionnelles de ce service et plus largement de ceux de la Sécurité Civile, pour la prise en charge de ce phénomène, susceptible de menacer les populations et leurs infrastructures. Les conclusions technologiques de la partie 1 ont permis de mieux préciser l'intérêt que revêt les technologies de l'informatique distribuée, et en particulier la technologie grille et les services web de l'Open Geospatial Consortium, et leur possible articulation fonctionnelle. Ainsi, il convient dorénavant d'approfondir les méthodologies conventionnelles propres à la technologie grille et le rapprochement pouvant exister avec les problématiques modernes de l'information géographique. Il en résulte une proposition méthodologique détaillée dédiée pour l'amélioration souhaitée du SPC-GD, dans sa capacité à prendre en charge des méthodes de prévision hydrologique et à transmettre son expertise hydrométéorologique en temps réel.

La première partie de ce chapitre, au vue des conclusions du chapitre 3, s'attache donc à repréciser les fondements et les besoins relatifs à la prévision hydrologique en situation de crise, et les contraintes fonctionnelles qui en découlent. Dans une seconde partie, les caractéristiques principales de la technologie grille sont présentées en comparaison avec une architecture informatique de petite échelle dédiée au calcul parallèle. Une présentation des rapprochements politico-scientifiques entre les consortiums de l'OGF, dédié à l'harmonisation des méthodes de grille, et l'OGC, dédié à la standardisation de la manipulation de l'information géographique, est ensuite développée. Dans la dernière partie de ce chapitre, la proposition méthodologique de cette recherche est présentée en s'appuyant sur l'ensemble des acquis développés dans les parties 1 et 2 de ce mémoire.

IV - 1. Les contraintes technologiques de la prévision hydrologique

IV - 1.1. De la prévision hydrologique…

Comme le suggèrent CLAUDET et BOUVIER (2004) :

> **« la qualité d'une prévision est caractérisée par la précision de la valeur annoncée et par l'anticipation dans le temps avec laquelle cette valeur est annoncée. Améliorer la prévision, c'est donc améliorer la précision de la valeur annoncée et/ou augmenter le délai de la prévision ».**

Ainsi, l'un des besoins les plus critiques concerne la rapidité d'exécution dont doit faire preuve les services hydrologiques, pour collecter, simuler, prévoir et informer les responsables de la gestion de crise. La connaissance rapide et précise des caractéristiques qualitatives et quantitatives de la pluie à venir peut permettre d'évaluer les risques hydrologiques en jeu (BEVEN, 2001). Différentes approches scientifiques viennent régulièrement alimenter le perfectionnement de cette prévision opérationnelle des crues éclair, qu'elle soit basée directement sur l'analyse des phénomènes pluvieux (CHIANG et *al*., 2007 ; DOSWELL et *al*., 1996; GUPTA et *al*., 2002) ou sur la prévision par modélisation hydrologique (AYRAL, 2005 ; BOUVIER et *al*., 2004 ; CAPORALI, 2007 ; DELRIEU et *al*., 2005 ; ESTUPINA BORRELL, 2004 ; GAUME et *al*., 2004 ; LAVABRE et GREGORIS, 2005 ; MARCHANDISE, 2007 ; NORBIATO et *al*., 2008 ; TOUKOUROU et *al*., 2009), elle connait encore de nos jours des faiblesses qui empêchent son réel succès en situation de crise. En effet, la prévision météorologique actuelle ne répond pas à la résolution spatio-temporelle que la prévision des crues éclair requiert (BEVEN, 2001). Enfin, la rapidité de génération d'une crue éclair, de 6 à 12 heures (KOBIYAMA et GOERL, 2007), impose à l'ensemble des acteurs de la gestion de crise une communication structurée, formelle et constante, afin de parvenir à « l'intégration de la chaîne d'alerte » (VINET, 2007) depuis l'expertise jusqu'à la gestion des opérations de secours (CLAUDET et BOUVIER, 2004 ; DRODOT et Parker, 2007 ; MONTZ et GRUNTFEST, 2002).

IV - 1.2. … aux contraintes technologiques inhérentes

Dans l'optique de perfectionnement de la chaîne d'alerte hydrométéorologique, dans laquelle ce travail de recherche s'engage, il s'agit de souligner les contraintes technologiques sous-jacentes à ces perspectives d'amélioration opérationnelle. Elles découlent directement des perspectives d'amélioration édictées par les retours d'expérience et/ou les conclusions des recherches hydrométéorologiques des crues éclair et permettent de conceptualiser les caractéristiques techniques du système expérimental développé dans cette recherche (THIERION et *al*., 2008).

C'est notamment au travers du projet européen CYCLOPS[1] (CYCLOPS, 2006), que la définition de ces contraintes technologiques a été menée (THIERION et *al*., 2007). Au travers de l'identification des contraintes de la gestion de l'expertise hydrologique des crues éclair et des feux de forêts, les résultats du projet cherchent à guider les recherches informatiques appropriées, telle que l'utilisation de technolo-

[1] http://www.cyclops-project.eu/

gies innovantes comme la technologie grille et les services web (MAZZETTI et *al.*, 2008). Le fond technologique de ce travail de recherche, déjà présenté dans le chapitre 2, est approfondi dans les paragraphes qui suivent. Il est possible de distinguer les besoins fonctionnels, relatifs à « ce que le système devrait faire » aux besoins non-fonctionnels correspondant à « comment le système devrait être », en d'autres termes aux critères qui permettent de juger de l'efficacité du système d'information (CYCLOPS, 2006 ; KOTONYA et SOMMERVILLE, 1998).

IV - 1.2.1. Besoins fonctionnels

Le système à prototyper doit donc s'appuyer sur les caractéristiques fonctionnelles suivantes :

1. l'accès et publication de l'information géospatiale en vue d'améliorer le partage de l'information géographique entre les acteurs, incluant la gestion des capteurs et des systèmes de collecte de données et la diffusion des informations modélisées et/ou expertisées,
2. la définition d'une politique d'utilisation des données afin d'harmoniser la gestion des données entre les différents utilisateurs,
3. la gestion d'autorisations et d'authentification pour limiter le risque d'intrusion étrangère sur les réseaux informatiques à risque des services de l'État,
4. l'accès à des capacités de stockage et de calcul à la demande, étant donné la variabilité de la production de données et de traitements informatiques nécessaires au cours de différentes crises,
5. la gestion de réservation et de priorisation des ressources informatiques qui permettent de garantir leur accès au moment de l'identification d'un phénomène dangereux proche,
6. la gestion dynamique de chaînes de traitements informatiques modulables afin de garantir la réutilisation des solutions logicielles existantes pour des tâches opérationnelles différentes.

IV - 1.2.2. Besoins non-fonctionnels

D'autre part, les méthodes implémentées dans ce travail de recherche se doivent de respecter un ensemble de contraintes fonctionnelles relatives à la gestion de crise :

1. l'adoption d'interfaces et de formats standards pour la gestion de l'information géospatiale pour garantir l'interopérabilité des systèmes d'information interconnectés,
2. doter les systèmes informatiques d'une bande passante conséquente pour éviter l'engorgement des réseaux et préserver les temps de réponse des applications,
3. assurer un fonctionnement en temps réel des applications au détriment de la qualité ou de la précision (mode dégradé), une étude préalable au cas par cas est indispensable pour identifier les informations critiques et primordiales,
4. l'augmentation de la résolution spatio-temporelle des résultats sans affecter le temps de réponse des applications pour répondre précisément aux requêtes des utilisateurs

5. l'utilisation d'interfaces graphiques conviviales et adaptées aux besoins des utilisateurs, selon une approche « métier » garantissant l'utilisation efficace des outils et la lisibilité des informations par des utilisateurs aux compétences différentes.

L'ensemble de ces contraintes technologiques constitue les postulats scientifiques et technologiques de cette recherche. A l'image d'un projet informatique, l'analyse des contraintes du système existant permet de conceptualiser le plus objectivement possible les solutions fonctionnelles à envisager et à expérimenter. Cette recherche propose donc d'identifier des solutions technologiques opérationnelles capables de résoudre les problèmes et les difficultés relevés précédemment (cf. §.III – 2.2.). Les progrès informatiques de ces dernières années, en particulier ceux ayant engendré le développement des architectures de grille, permettent d'entrevoir, au vue des besoins et des contraintes identifiés, une relative efficacité des technologies « réseau ». La distribution géographique des ressources informatiques, et les capacités calculatoires qu'elles engendrent offrent *a priori* certaines garanties dans les objectifs de partage informationnel et d'augmentation des capacités de modélisation du SPC-GD.

IV - 2. La grille EGEE : un support informatique performant et adaptable

IV - 2.1. La technologie grille : des caractéristiques technologiques et des champs d'application favorables à une adoption opérationnelle

Le chapitre 2 a permis de définir précisément le champ de l'informatique distribuée et plus précisément les fondements scientifiques et technologiques de la grille. Il apparait important dans ce chapitre méthodologique de caractériser les spécificités technologiques de cette architecture afin d'établir une proposition méthodologique en accord avec, d'une part les contraintes et les besoins du système existant et d'autre part les potentialités directes et indirectes de la technologie grille.

IV - 2.1.1. Le point central de la parallélisation informatique

En connectant des ressources de calcul et de stockage géographiquement distantes, la technologie grille offre la possibilité d'exécuter des tâches informatiques de manière simultanée capables d'accéder à des données partagées. La technologie grille s'appuie sur l'idée « que l'on dispose dans la plupart des cas de plus de puissance en mettant en parallèle des processeurs qu'à partir d'un seul, même plus rapide » (SOBERMAN, 2003). Elle permet donc de développer des méthodologies basées sur la parallélisation. Cette notion de parallélisation est complexe et multiforme, et sa mise en œuvre nécessite des efforts, en particulier pour la bonne synchronisation des tâches, souvent basés sur une approche empirique. GENGLER et *al*. (1996) et FOSTER (1995) distinguent deux types principaux de parallélisme :

- le parallélisme de contrôle,
- le parallélisme de données.

IV - 2.1.1.1. Le parallélisme de contrôle

Ce type de parallélisme correspond au découpage d'un programme séquentiel en sous-tâches indépendantes et communicantes au travers de messages (Figure **51**). Le développeur doit donc identifier les tâches exécutables simultanément et celles à exécuter en séquence, en respectant des recommandations telles que le nombre de tâches et leur durée en fonction de l'exécution globale, leur homogénéité temporelle, et la répartition des charges de calcul (SILVA et BUYYA, 1999).

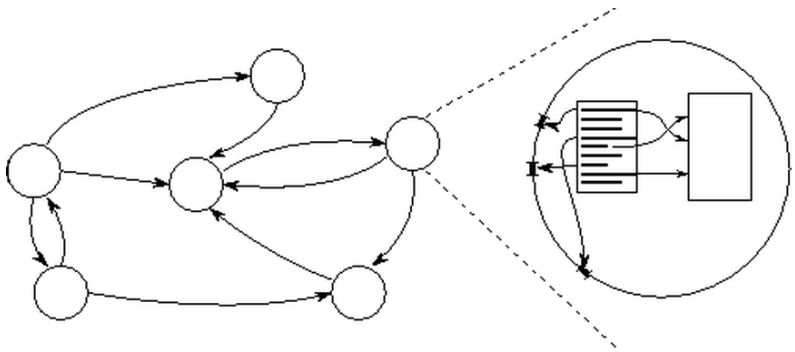


Figure 51 : Principe de parallélisation de contrôle, les cercles représentent des tâches, et les flèches des canaux de communication (tirée de FOSTER, 1995)

Le modèle informatique le plus répandu pour gérer ce type de parallélisme est le Message Passing Interface (MPI) (WALKER et DONGARRA, 1996). Il consiste à établir un ordonnancement de tâches s'exécutant en parallèle et/ou séquentiellement et communiquant par le biais de messages asynchrones. Un message est asynchrone lorsque l'émetteur ne voit pas son exécution bloquée dans l'attente d'une réponse du récepteur invoqué.

IV - 2.1.1.2. Le parallélisme de données

Ce parallélisme est implémenté sur le modèle SIMD (Single Instruction Multiple Data). Il consiste à la manipulation de structures de données régulières par un même algorithme. Les processeurs sont propriétaire de sous-ensembles d'une donnée et est « responsable de la réalisation de toutes les opérations les concernant ». Cette méthode de répartition des calculs est appelée le "propriétaire calcule" (Owner-computes rule, OCR) et (GENGLER et *al*., 1996). L'enjeu majeur de cette méthode est de déterminer la décomposition de la donnée initiale et la répartition du traitement de ces sous-ensembles par des tâches indépendantes afin de respecter les critères de performance préalablement cités (FOSTER, 1995). Trois schémas de répartition des données existent (Figure **52**) :

- par blocs
- cyclique
- par blocs cycliques

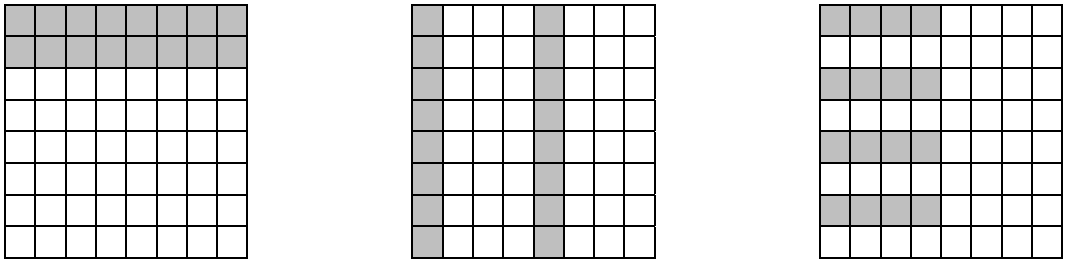


Figure 52 : Distribution des sous-ensembles dans la parallélisation des données avec 4 processeurs (d'après FOSTER, 1995)

Ces méthodes permettent de garantir une répartition des charges homogène entre les processeurs et le degré de parallélisme peut être très important et par conséquence garantir un gain considérable de temps d'exécution. De telles méthodes sont de manière générale bien adaptées pour les traitements de l'information géographique d'un SIG (ARMSTRONG, 1994 ; 1995). En effet, la représentation « raster » structurée sous forme matricielle est facilement décomposable.

IV - 2.1.1.3. Parallélisation et technologie grille

Comme le souligne GENGLER et *al*. (1996), les temps de communications et l'hétérogénéité des ressources peuvent amener à de grosses différences d'exécution des tâches élémentaires » et donc limiter l'efficacité de la parallélisation. La technologie grille de manière générale, et la grille EGEE en particulier, par sa structure faiblement couplée et l'hétérogénéité de ses ressources, ne correspond pas à une architecture totalement adaptée à la parallélisation de contrôle (BAUDE, 2002 ; PRIOL, 2005). Comme cela a été démontré dans le chapitre 2, les architectures de grille sont généralement hétéro-

gènes, et la dispersion géographique de ses ressources cause des temps de communication (temps de latence) plus longs que les superordinateurs et les clusters (cf. §.II – 2.1.1.).

Ainsi, même si des applications parallèles, basées sur le standard MPI (CHILDS, 2007), existent sur l'architecture EGEE, l'utilisation la plus courante de la grille concerne l'exécution d'un même algorithme sur différents jeu de données (ESPINAL et *al*., 2008) ou des études paramétriques (ASTALOS et *al*., 2008). Dans la même philosophie que la parallélisation classique des données, les études paramétriques correspondent au traitement de jeux de données indépendants par le même algorithme élémentaire. La seule dépendance concerne le jeu de paramètres fournis en entrée et le fichier de sortie contenant l'ensemble des résultats de chaque tâche (Figure **53**).

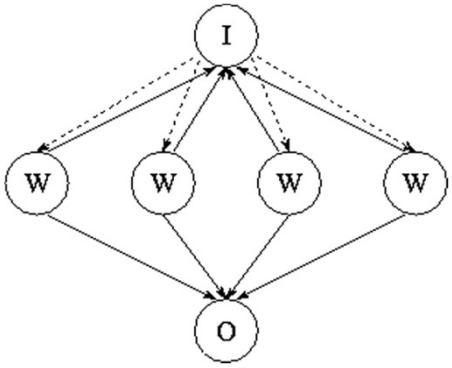


Figure 53 : Structure des tâches pour une étude paramétrique (tirée de FOSTER, 1995)

Dans le cadre de cette recherche, la technologie grille est donc abordée par sa capacité à exécuter un grand nombre de tâches indépendantes de manière simultanée. La seule dépendance pouvant exister entre ces différentes tâches concerne les données d'entrée qui peuvent être, dans certains cas, identique. Mais le grand nombre de ressources disponibles dans l'architecture EGEE permet la réplication intensive des données dans la grille et *a fortiori* de limiter la concurrence d'accès.

IV - 2.1.2. Les caractéristiques informatiques de la grille EGEE

Il s'agit d'une architecture grande échelle intégrant de manière dynamique un nombre quasi-infini de ressources. De cette manière, elle répond aux besoins d'utilisateurs en termes de calcul intensif à la demande et de flexibilité dans l'implémentation de leurs fonctions « métier ». Le calcul haute-performance (HPC), auquel cette architecture est associée, est clairement adaptée à supporter des charges de calcul importantes, en temps et en nombre de ressources requises (FOSTER et *al*., 2001). De nombreuses différences technologiques persistent entre un superordinateur, une ferme de calcul et une architecture de grille. Afin de souligner l'intérêt d'avoir recours à la technologie grille dans le cadre du traitement de la prévision hydrologique, plusieurs caractéristiques fondamentales des systèmes distribués ont été sélectionnées[2] et synthétisés dans le Tableau **4**. Il s'agit de comparer ses caractéristiques entre un système individuel (cluster, superordinateur, etc.) et la grille EGEE qui intègre plusieurs sites.

[2] Ces indicateurs sont principalement issus de la partie « 6.1 Objectives and motivation » de la norme ITU-T Rec. X.901 | ISO/IEC 10746-1.

Type / Caract.	EGEE	Cluster
Image unique du système	variable	oui
Hétérogénéité	faiblement couplée	fortement couplée
Inter-organisationnelle	oui	non
Nombre de ressources	illimité	<100
Extensibilité / Obsolescence	forte	limitée
Gestion des ressources	décentralisée	centralisée
Autonomie	plusieurs points de commande	1 seul point de commande
Point individuel de défaillance (SPOF)	non	fréquent
Temps de latence	de l'ordre de la minute	de l'ordre de la seconde
Concurrence	disjointe et compétitive	coopérative

Tableau 4 : Comparaison entre la grille EGEE et un cluster suivant les caractéristiques principales des systèmes distribués

Alors que les clusters et des superordinateurs intègrent des ressources comme s'il s'agissait d'un système informatique unique, la grille est intrinsèquement disparate de part les ressources hétérogènes et des multiples points de commande qu'elle intègre. Dans ce sens, une grille de l'envergure d'EGEE, en termes de nombre de ressources, n'offrent pas la visibilité globale que peut fournir un cluster, ou une grille de plus petite échelle telle que GRID'5000 (VICAT-BLANC PRIMET et *al*., 2003). Cependant, sa structuration en organisation virtuelle permet d'offrir à l'utilisateur un aperçu des ressources dont il dispose. Le nombre de processeurs disponibles peut cependant varier dans le temps, du fait de l'indépendance de chaque site vis-à-vis de l'organisation virtuelle. A l'inverse, une architecture de type cluster, fortement adaptée à la parallélisation, connait des difficultés dans sa capacité à partager dynamiquement des ressources.

En transcendant les frontières institutionnelles et géographiques, elle permet le regroupement de réseaux hétérogènes, facilite leur interaction et « abolit » les distances géographiques (FOSTER et KESSELMAN, 2004 ; MCCLATCHEY et *al*., 2007 ; VON LASZEWSKI et WAGSTROM, P., 2004). Le nombre presque illimité de ressources facilite de plus l'extensibilité de l'architecture de manière transparente et surtout les montées en charge soudaines[3], telles qu'elles peuvent avoir lieu lors de l'irruption de situations de crise. Dans ce sens, elle répond à l'un des fondements organisationnels de la gestion de crise, à savoir sa propension à établir un contexte collaboratif fiable et sécurisé.

Une architecture de grille comme EGEE fournit de manière continu (24h/24 et 7j/7) des ressources de calcul et de stockage avec un degré de fiabilité globalement important (GAGLIARDI et *al*., 2005), tout en permettant l'exécution simultanée et indépendante d'un grand nombre de tâches informatiques. Les motivations pour utiliser la technologie grille résident précisément dans la grande distribution de ses ressources et la décentralisation de sa gestion. Elles permettent d'éviter un point individuel de défaillance (en anglais *Single Point Of Failure*, SPOF) et d'offrir une redondance dans les circuits de com-

[3] Certains points de blocage persistent à ce niveau

munication (données et exécutions), favorisant la prise en charge des pannes et des erreurs d'exécution et garantissant ainsi une certaine réussite des traitements requis (au sens de qualité de service) (FOSTER et KESSELMAN, 2004 ; SOBERMAN, 2005). C'est au travers de ces fortes capacités de réplication, de communication asynchrone et donc de partage, que la technologie grille peut répondre aux besoins de l'expertise hydrologique en situation de crise.

En contre partie, la relative jeunesse de cette technologie et sa dimension toujours expérimentale la rend difficilement généralisable à l'ensemble des contraintes des utilisateurs finaux. En effet, initialement conçue dans l'optique de supporter la production massive de données du LHC et l'exécution de tâches de longue durée (CONDOR TEAM, 2009), elle ne s'adapte pas de prime abord à toutes les problématiques, en particulier quand elles ont des contraintes temporelles courtes comme celle établie dans cette recherche. Il est donc important de considérer dans la proposition méthodologique qui suit deux points critiques :

- l'allocation des tâches de calcul qui est confrontée à l'accès concurrent aux ressources de calcul par des utilisateurs multiples, à la différence d'un cluster où la répartition est généralement paramétrée et optimisée à l'avance. Cette difficulté est d'autant plus exacerbée que l'état des ressources en temps réel est plus difficilement identifiable sur EGEE.

- l'hétérogénéité des ressources implique des différences importantes au niveau des bandes passantes et des temps de latence, qui pour les seconds sont relativement longs en comparaison avec un cluster. Ainsi, il convient d'optimiser le ratio entre les temps de transfert (allocation, échanges de données et messages) et d'exécution des tâches[4].

Ces points particuliers sont approfondis dans le chapitre suivant dédié à l'implémentation du système d'aide à la décision.

IV - 2.1.3. L'utilisation de la grille pour la gestion de problématiques environnementales

L'entrée technologique est fondamentale dans cette recherche, cependant il semble important de s'appuyer sur les réussites scientifiques qu'ont connue différentes communautés scientifiques dans la prise en charge de leur thématique. Ainsi, un état de l'art concis a été développé de manière à justifier les choix scientifiques développés dans cette proposition méthodologique. Deux points sont particulièrement approfondis à la vue des besoins définis (cf. §.III – 2.2.) :

- la problématique d'aide à la décision liée au contexte de l'expertise hydrométéorologique

- la problématique des études paramétriques liée à l'incertitude de la modélisation hydrologique

IV - 2.1.3.1. Aide à la décision

Cette technologie, de part ses caractéristiques, apparait comme une solution technologique efficace pour intégrer des données évènementielles et des outils dédiés, parmi de multiples décisionnaires tout au long du processus d'expertise, tels que le SPC-GD et le SCHAPI dans le cas des crues à cinétique

[4] Il est question de granularité de la tâche (Vicat-Blanc Primet *et al*., 2003)

rapide (BOVOLO et *al*., 2008 ; CHERVENAK et *al*., 2000 ; FOLINO et *al*., 2005 ; GADGIL et *al*., 2004 ; HLUCHY et *al*., 2003 ; KUNSZT, 2003 ; MURGIA et *al*., 2009). Les recherches développées dans le cadre du projet MEDIGRID[5] suggèrent l'intégration de modèles de propagation de feux, hydrologiques, d'érosion des sols et de glissement de terrains dans un système unique d'aide à la décision basé sur la technologie grille. Au travers de cet outil, et de l'aspect modulaire de la grille, les chercheurs ont la possibilité de créer des *workflows*[6] en connectant différents modèles et de partager leurs données afin de déterminer les impacts potentiels d'un feu de forêt (BOVOLO et *al*., 2008). Dans la même philosophie, le projet AQUAGRID basé sur la plateforme GRIDA3 a implémenté une chaîne de modèles pour la gestion de la contamination des sols et des eaux souterraines (LECCA et *al*., 2009). Enfin, le système MOSÈ (Spatio-Temporal MOdelling of Environmental Evolutionary Processes by means of GeoSErvices) permet « d'analyser et de gérer l'identification et la mitigation de désastres naturels comme les inondations, les feux de forêts, les glissements de terrains, etc » (FOLINO et *al*., 2005). Le système développé permet de gérer à la volée des chaînes de traitements en fonctions de l'aléa à étudier.

IV - 2.1.3.2. Études paramétriques

La propension de la grille à soutenir des études paramétriques offre l'opportunité de développer des modélisations multi-scénarios adaptés aux besoins de chaque utilisateur (COVENEY, 2005 ; MINETER et *al*., 2002). En s'appuyant sur la capacité de la grille à prendre en charge de nombreuses exécutions en simultanée, HLUCHY et *al*. (2003) ont développé une plateforme permettant de simuler plusieurs scénarios hydrologiques basés sur des prévisions météorologiques déterministes et probabilistes, de la même manière que les objectifs précités d'amélioration de la prévision des crues à cinétique rapide (Synthèse du chapitre 3). MURGIA et *al*. (2009) suggèrent la même méthodologie de manière à :

- garantir un minimum de simulations valides (au sens informatique du terme), élément essentiel de la gestion de crise,
- tester des jeux de paramètres alternatifs, ce qui apparait adapté à l'optique d'aide à la décision hydrologique caractérisée par une forte incertitude de l'aléa.

Finalement, l'utilisation de la technologie grille, pour exécuter un grand nombre de simulations indépendantes, a été expérimentée grâce à la plateforme OpenSHA pour générer des cartes de risque sismique sur une centaine de zones géographiques (FIELD et *al*., 2005).

IV - 2.1.4. Synthèse sur l'intérêt d'utiliser la grille comme support technologique

En résumé et au vue de l'état de l'art préalablement développé, la grille peut être caractérisée par :

- des traitements informatiques sur des machines distantes
- des traitements parallèles et indépendants
- l'obtention de ressources informatiques à la demande

[5] http://www.eu-medin.org/
[6] Ce terme a été préféré à « chaîne de traitements » du fait de sa grande utilisation dans le monde de l'informatique

- le traitement de données volumineuses à distance
- un niveau élevé de sécurité et de confidentialité avec la possibilité de crypter les données
- la capacité de partager des données, des résultats et des algorithmes (déjà grillifiés)

En accédant à une architecture de grille, telle que celle fournie par EGEE, l'utilisateur final a les moyens d'augmenter les capacités informatiques dont il dispose à l'origine en déportant des traitements informatiques et des données sur des machines distantes et en laissant ainsi libre ses propres ressources. De plus, cette capacité de réplication améliore la fiabilité des traitements et la préservation des données, dans le cas d'un arrêt du système local. Enfin, la parallélisation des traitements combinée à l'accès à des données multi-sources permet d'implémenter des fonctionnalités plus complexes, telles que des *workflows* ou des études paramétriques enrichissant *a priori* les capacités offertes au SPC-GD pour soutenir leur mission de prévision des crues.

Finalement, l'une des questions les plus critiques concernant la Sécurité Civile, est la sécurité face aux intrusions et à la confidentialité des données et plus largement des systèmes opérants. Cette question, sous-jacente dans l'analyse technologique présentée, s'appuie sur l'état de l'art relatif à la technologie grille et en particulier sur les garanties offertes par le concept d'organisation virtuelle (cf. §.II – 2.3.1.4.). Cette dernière en basant ses relations informatiques, organisationnelles et politiques sur un contrat de confiance institutionnel semble offrir un niveau de sécurité convenable permettant l'utilisation simultanée et concurrente des ressources de la grille EGEE par des milliers d'utilisateurs.

Les caractéristiques technologiques de la grille semblent donc pouvoir répondre aux besoins informatiques de l'expertise hydrologique en situation de crise. Son utilisation est à l'heure actuelle encore prédestinée à la communauté informatique, le développement de telles applications nécessite des efforts financiers et humains que les services opérationnels ne peuvent certainement pas engager. Ainsi, la tendance actuelle au niveau de la communauté « grille » et des communautés scientifiques engagées dans sa démocratisation, au travers par exemple du projet EGEE, concerne d'une part la valorisation de leurs succès scientifiques et d'autre part, le développement de solutions applicatives adaptées aux compétences technologiques de services opérationnels. C'est dans ce contexte, ouvert vers une « adoption globale » de cette technologie que cette recherche s'intègre.

IV - 2.2. La technologie grille et la science de l'information géographique : une synergie innovante

IV - 2.2.1. L'intérêt croissant des services web

Un intérêt particulier doit être porté sur les initiatives de grille portant sur la gestion de l'information géographique, composante essentielle de l'expertise hydrométéorologique. D'après CRAGLIA et *al.* (2005), et WOOLF et NATIVI (2008), les infrastructures de données spatiales, telles qu'INSPIRE[7] (Infrastructure for Spatial Information in Europe), qui vont permettre à terme de mettre à disposition des données géospatiales à une large gamme d'utilisateurs, peuvent tirer profit de la technologie grille. Cepen-

[7] http://inspire.jrc.ec.europa.eu/

dant, l'utilisation opérationnelle de ces ressources distantes implique l'implémentation de couches logicielles intermédiaires (services évolués et applications) permettant l'interaction avec les couches de bas niveau de la grille (fabrique et intergiciel) et les portails web auxquels les utilisateurs ont accès (Figure **54**) (VON LASZEWSKI et WAGSTROM, 2004).

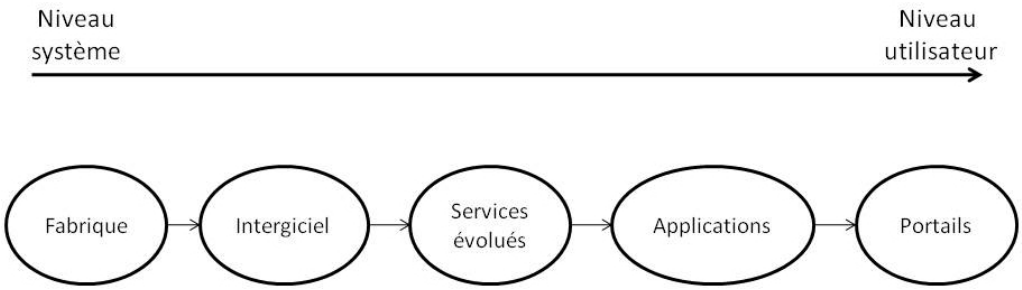


Figure 54 : Architecture de grille multi-tiers (tirée de Von LASZEWSKI et WAGSTROM, 2004)

Ces services évolués ou services web, déjà abordés dans le chapitre 2 dans le cadre des efforts d'interopérabilité menés par l'OGC, peuvent faciliter, grâce à leur rôle d'interface entre deux composants logiciels distribués :

- la découverte de données et de services de traitement,
- le rapatriement de ces données et l'exécution à distance de ces traitements,
- la gestion dynamique et modulaire des traitements informatiques qu'ils permettent de connecter (FOX, 2004 ; GADGIL et *al*., 2004 ; GHANEM et *al*., 2004 ; PRIOL, 2005)
- la composition à la volée et de manière automatique de *workflows* (SHEN et *al*., 2004 ; SHI et *al*., 2002).

Dans le cadre de l'information géographique, ils peuvent d'autre part enrichir les capacités fonctionnelles de la cartographie en ligne en effectuant des géotraitements sur une architecture de grille, et en diffusant les résultats sur l'interface de cartographie en ligne (CLARKE et *al*., 2006 ; HAWICK et *al*., 2003; SHEN et *al*., 2004 ; SHI et *al*., 2002 ; WANG et *al*., 2004). Les services web de l'OGC sont originellement destinés à la transmission et au traitement de l'information géographique au travers d'Internet (cf. §.II – 3.2.3.3.). Cependant, la valeur ajoutée de la technologie grille en comparaison avec Internet, en fait un nouveau support technologique performant.

Ainsi, au travers d'Internet (et du web), moyen privilégié pour accéder aux ressources de grille (COVENEY, 2005 ; HU et *al*., 2008 ; LIEN et *al*., 2004 ; YOUN et *al*., 2005; 2008), les services web permettent d'accéder aux services de l'intergiciel. De même, à l'échelle de l'intergiciel, l'utilisation de services web tend à se généraliser pour formaliser la communication entre les services de l'intergiciel.

L'enjeu principal de cette généralisation reste d'ordre ontologique[8], c'est pourquoi l'OGF[9] (Open Grid Forum) œuvre quotidiennement pour la standardisation du fonctionnement de la grille en se basant par exemple sur la norme OGSA (FOSTER et *al*., 2002) (cf. §.II – 2.3.1.3). Comme pour l'OGC, cet effort

[8] Qui est relatif à la sémantique et la syntaxe d'une communauté ou d'un champ scientifique
[9] http://www.gridforum.org/

nécessite l'engagement d'un grand nombre de partenaires afin de définir des spécifications les plus interopérables possibles. Dans le cadre de l'utilisation de l'information géographique sur une architecture de grille, l'intégration transparente de géotraitements et d'accès à des données implique notamment un rapprochement entre l'OGC et l'OGF pour formaliser des interfaces entre les services de l'intergiciel et les services géographiques.

IV - 2.2.2. Le Geospatial Grid

Cette approche ouvre la voie de la généralisation de services d'information géographique sur la grille, que Di (2006) formalise par l'utilisation de la technologie grille sous le terme de *Geospatial Grid*. L'approche développée par Di se base sur l'idée, partagée par plusieurs auteurs (AYDIN et *al*., 2006 ; 2008 ; CLARKE et *al*., 2006 ; DI et *al*., 2008 ; FOX, 2004 ; WOOLF et NATIVI, 2008 ; XIAO et FU, 2003), que l'accès à la grille par des applications à composante géographique doit être facilitée par le recours à des services web. D'après la Figure 55, l'interaction entre des utilisateurs et les ressources de calcul offertes par une architecture de grille pour accéder à des géotraitements ou des données géospatiales est rendue possible par l'implémentation de services web intermédiaires permettant d'interroger (*catalogues services*) de manière transparente les services d'information de la grille et d'utiliser (*geospatial web services*) les ressources (au sens large du terme) (DI, 2006). Des auteurs suggèrent l'utilisation des services web de l'Open Geospatial Consortium pour mener à bien ces implémentations (DI, 2006 ; WOOLF et NATIVI, 2008).

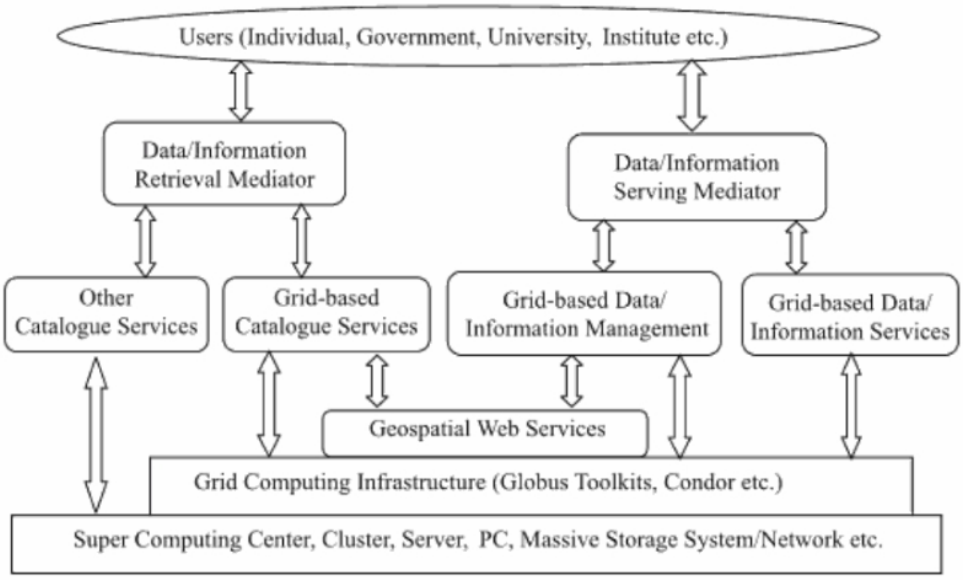


Figure 55 : Structure du Geospatial grid (tirée de DI, 2006)

En garantissant l'interopérabilité des systèmes d'information, ces standards semblent offrir au SPC-GD un moyen de consolider et d'étendre le champ des fournisseurs de données pendant la crise et d'accélérer l'accès à des traitements complexes (modèles) et la transmission d''informations critiques aux services de la Sécurité Civile.

Cependant, l'originalité de la donnée géographique en comparaison avec les données classiques impose encore le développement de nouveaux concepts technologico-scientifiques (définition d'une onto-

logie) pour permettre cette intégration dans une architecture de grille. C'est à cet objectif général que les développements présentés en suivant s'apparentent.

IV - 2.2.3. Des initiatives récentes dans l'utilisation de services web de l'OGC pour l'accès aux ressources de grille

Depuis quelques années, un effort a été initié pour permettre le rapprochement entre les communautés « grille » et « géospatiale », et plus exactement entre l'OGF et l'OGC afin de déterminer des standards communs (LEE et PERCIVALL, 2008). À terme, il s'agit de rendre possible l'interaction transparente entre les services de grille et ceux spécifiques à l'information géographique (FOX, 2004) afin de réaliser la véritable intégration des SIG et de la grille suggérée par DI (2006).

L'évolution des SIG bureautiques vers des SIG aux fonctionnalités distribuées sur un réseau informatique représente un défi majeur que le rapprochement entre l'OGC et l'OGF cherche à résoudre. Plusieurs initiatives ont cherché à développer de telles solutions, en intégrant des fonctions SIG au sein d'un intergiciel (CAI et *al*., 2009 ; FOLINO et *al., 2005* ; SHEN et *al*., 2004 ; WANG et SU, 2007), malgré un succès technologique certain, ces solutions sont difficilement généralisables à des infrastructures de grille grande échelle comme celle abordée dans cette étude.

Plusieurs initiatives ont cherché à implémenter des services web de l'OGC pour accéder aux ressources de la technologie grille, c'est-à-dire des données stockées ou des algorithmes déjà grillifiés. En présentant trois initiatives développées durant ces dernières années, et encore en cours à l'heure actuelle, il s'agit de souligner la généricité offerte par les services web de l'OGC, au même titre que les web services. L'enjeu apparaît donc technologique, scientifique et politique.

IV - 2.2.3.1. L'institut de recherches spatiales ukrainien

Les nœuds de l'architecture de grille utilisés par cet institut sont répartis sur 3 sites :

- Le centre RSGS-CAS (Remote Sensing Satellite Ground Station Chinese Academy of Sciences) en Chine relié à un système satellitaire (données géospatiales)
- L'institut of cybernétique du NASU (National Academy of Sciences of Ukraine)
- L'institut de recherches spatiales NASU- NSAU, l'équivalent du CNES français

Les recherches développées autour de cette architecture s'intègrent dans l'initiative GEOSS, au travers du projet CEOS / WGISS[10], en développant des applications et des méthodologies de traitement de données satellitales (KUSSUL et *al*., 2009). Dans ce contexte, une grillification des *Sensor Observation Service* (SOS) (NA et PRIEST, 2007), Sensor Alert Service (SIMONIS, 2006) et Sensor Planning Service (SPS) (SIMONIS, 2007) est proposée. Ces trois services de l'OGC sont spécifiques à la gestion de systèmes de mesure. Un tel développement peut permettre l'accès direct aux images satellites volumineuses par les nœuds de calcul, et un contrôle à distance des capteurs, afin d'adapter en temps-réel leur comportement aux différentes situations. Au travers de cette solution technologique, la modélisation météorologique numérique et celle de l'extension de champs d'inondation est accélérée et permet

[10] http://www.ceos.org/

de transmettre les informations aux services de secours dans des meilleurs délais opérationnels. Finalement, l'évaluation de la biodiversité nécessite un nombre important de données volumineuses, ainsi grâce à cette architecture leur traitement à distance facilite leur utilisation.

IV - 2.2.3.2. Le projet CYCLOPS et G-RISICO

Entre 2006 et 2008, le projet européen CYCLOPS[11] (CYber-Infrastructure for CiviL protection Operative ProcedureS) regroupant des spécialistes de la technologie grille, en environnement et en science du risque a concouru à « combler le fossé » existant entre les communautés de grille et de protection civile. Fondé sur deux expérimentations opérationnelles sur les crues éclair et les feux de forêts (approche « bottom-up »), ses résultats ont contribué à définir les besoins technologiques spécifiques aux services de protection civile en vue d'une adoption potentielle d'une infrastructure de grille.

En ce qui concerne les feux de forêts, le modèle RISICO, utilisé au sein de la sécurité civile italienne pour l'évaluation du risque « feu de forêts » sur l'ensemble du territoire italien, a été grillifié et nommé G-RISICO (MAZZETTI et *al*., 2009). Le portage de ce modèle sur la grille a permis une amélioration importante de ses performances de calcul. En effet, il n'existe aucune dépendance fonctionnelle entre les cellules des matrices traitées par G-RISICO ce qui a permis d'utiliser une méthode de parallélisation des données (cf. §.IV - 2.1.1.2) offrant des potentialités de changement de résolution et de montée en charge. Enfin de répondre aux objectifs de la Sécurité Civile en termes de facilité d'utilisation et d'interopérabilité, l'ensemble des interactions entre le portail utilisateur et l'intergiciel de grille ont été formalisées grâce aux services web de l'OGC, le *Web Coverage Service* (WCS) et le *Web Processing Service* (WPS) pour accéder respectivement aux données et aux algorithmes grillifiés sur les ressources de l'organisation virtuelle CYCLOPS de l'architecture EGEE. Ces services de l'OGC sont approfondis dans le chapitre 6 décrivant la méthodologie implémentée dans cette recherche (cf. §.IV - 1)

L'un des résultats principal du projet CYCLOPS est une proposition d'architecture informatique composée de plusieurs couches, des ressources de grille (*GRID Platform EGEE*) à l'interface graphique (*Presentation and Fruition Services*), en passant par la couche de services web dédiée à la dimension « logique » de l'architecture (*Grid Services for Earth Sciences*) (Figure **56**). Elle justifie le recours à des services web standardisé (en l'occurrence ceux de l'OGC) pour utiliser la grille en situation opérationnelle. De plus, cette architecture offre une forte capacité de reproductibilité des processus informatiques et cherche à soutenir les besoins relatifs à la Sécurité Civile.

[11] http://www.cyclops-project.eu/

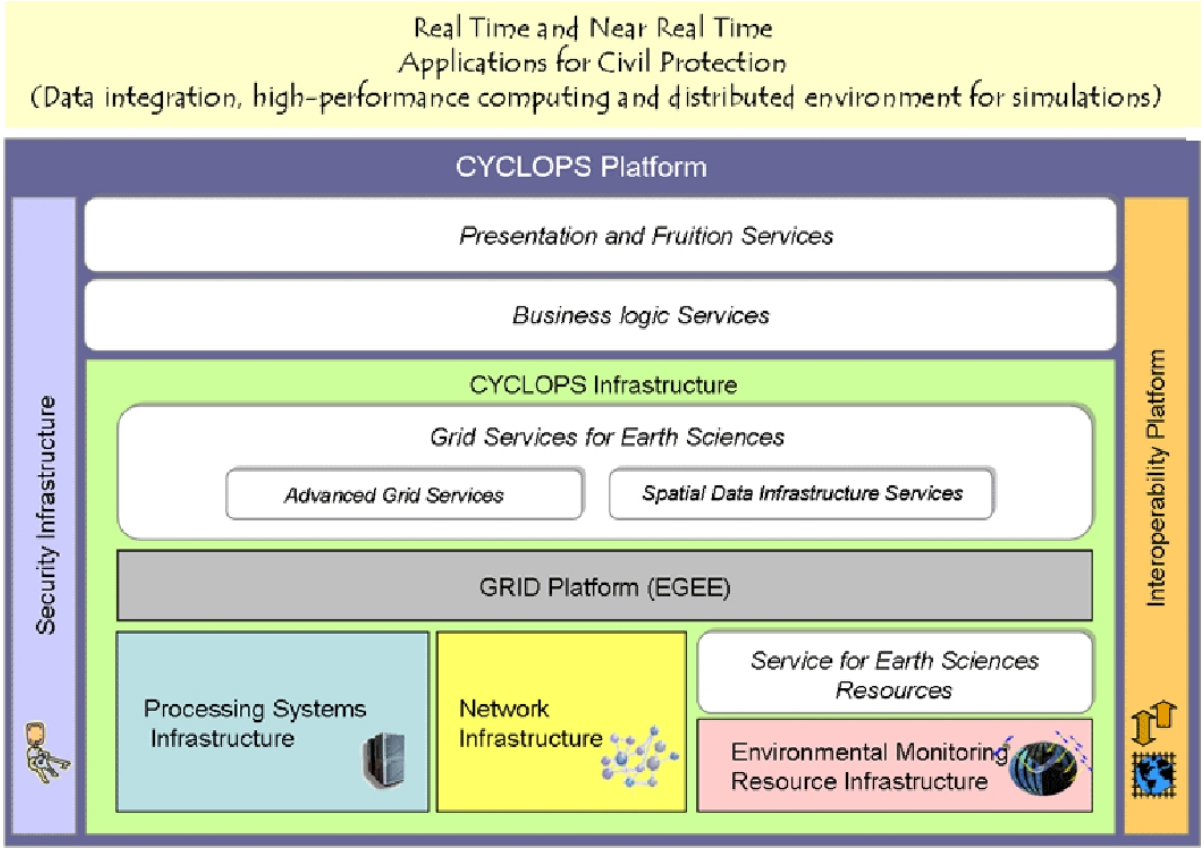


Figure 56 : Architecture proposée par le projet européen CYCLOPS (tirée de Mazzetti, 2006)

Le deuxième cas d'utilisation, concernant la gestion des crues éclair, est abordé en détail dans les chapitres suivants. Dans la même philosophie d'interopérabilité que G-RISICO, l'une des composantes méthodologiques de cette recherche s'efforce de proposer une implémentation de services web de l'OGC pour des besoins opérationnels spécifiques aux crues éclair.

IV - 2.2.3.3. Le projet GDI-GRID

L'objectif général de ce projet est d'intégrer les infrastructures de données spatiales allemandes au sein de l'architecture de grille D-GRID (initiative de grille allemande) (KURZBACH et *al*., 2009). La réalisation de cet objectif passe par le développement de services génériques permettant le traitement de données géospatiales sur l'infrastructure de grille (KUNZ et *al*., 2009). Plusieurs cas d'utilisation ont été imaginés pour démontrer l'utilité d'une telle implémentation :

- la simulation d'inondation
- la simulation de la propagation du son
- l'optimisation des itinéraires pour les services de secours

Comme pour les projets précédents, la gestion de ces opérations s'opère au travers des services web de l'OGC et de la technologie grille. La simulation d'inondation est basée sur des données géométriques dérivées du LIDAR et sur des modèles d'écoulement et de crues permettant de générer différents scénarios d'inondation suivant des jeux de paramètres et des résolutions spatiales différentes. Le traitement du LIDAR, ainsi que l'ensemble des scénarios nécessitent d'importantes capacités de calcul

que la grille peut fournir. L'ensemble de ces traitements est pris en charge par différents WPS. L'originalité de cette approche réside d'une part dans la capacité d'orchestration des WPS afin de générer dynamiquement des chaînes de traitements et d'autre part dans l'intégration complète des WPS comme services de grille à part entière.

Ces trois projets reflètent les possibilités technologiques et scientifiques qu'offre l'utilisation des services web de l'OGC pour accéder aux données géographiques et aux géotraitements présents dans une infrastructure de grille. Le nouvel intérêt porté par la société 52° North[12], mondialement reconnue dans le monde de la cartographie en ligne, pour les infrastructures de grille consolide finalement la viabilité de cette approche (BARANSKI, 2008). En d'autres termes, les services web de l'OGC ont été imaginés pour répondre aux besoins d'interopérabilité des organisations, que l'utilisation de la grille, en partie standardisée, permet de globalement préserver. Ainsi, la filiation originelle existante entre ces services web géographiques et la cartographie en ligne semble particulièrement adaptée pour le développement d'un système d'aide à la décision intégrant des modèles hydrologiques s'exécutant sur les ressources de calcul de la grille EGEE.

[12] http://52north.org/

IV - 3. G-ALHTAÏR : un système spatial d'aide à la décision hydrologique

Les deux parties précédentes ont permis d'identifier, au travers d'une analyse bibliographique centrée sur les questions technologiques de la grille, et en particuliers celle d'EGEE, des éléments structurels et fonctionnels pertinents *a priori* capables de répondre aux problématiques de l'expertise hydrométéorologique en situation de crise (cf. §.III.2.2.). Il s'agit dorénavant d'établir la proposition méthodologique de cette recherche, en adéquation avec les caractéristiques d'EGEE, les contraintes fonctionnelles du SPC-GD et les besoins inhérents à prévision hydrologique en situation de crise.

IV - 3.1. L'hypothèse de recherche centrale : le système spatial d'aide à la décision

L'objectif principal de la méthodologie de cette thèse est de contribuer à l'amélioration des capacités de prévision du service de prévision des crues « Grand Delta » (SPC-GD), alors que l'hypothèse technologique tacite correspond à la plus value que peut apporter l'informatique distribuée, et en particulier la technologie grille. Principalement méthodologique, cette démarche scientifique s'apparente à une expertise sur les potentialités de la technologie grille pour favoriser la mission d'anticipation de l'aléa « crue à cinétique rapide » en situation de crise.

Pour mener à bien cette démarche, la première hypothèse méthodologique correspond au recours à un système d'information expérimental capable de répondre aux contraintes et aux besoins fonctionnels du SPC-GD tout en permettant une étude de faisabilité sur l'architecture de grille EGEE. La dénomination choisie pour ce système d'information est « système spatial d'aide à la décision ».

Ce concept de système spatial d'aide à la décision (SSAD) a été formalisé à l'origine par DENSHAM (1991) pour qui le SSAD est « conçu pour fournir à l'utilisateur un environnement d'aide à la décision qui permet l'analyse de l'information géographique de manière souple ». Il est possible de compléter cette définition par « un système informatique interactif conçu pour supporter l'aide à la décision efficace d'un utilisateur ou d'un groupe d'utilisateurs tout en résolvant un problème spatial semi-structuré » (Malczewski**,** 1997).

L'architecture générale d'un tel système est basée sur (Figure **57**) :

- une ou plusieurs bases de données à composante géographique
- un système de modélisation relatif au problème à résoudre
- des outils de visualisation et d'analyse graphiques

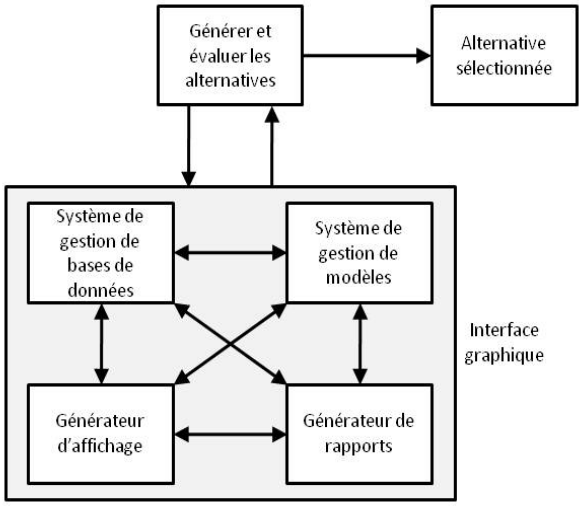


Figure 57 : Composants d'un système spatial d'aide à la décision
(d'après DENSHAM, 1991)

Il permet donc un accès facile aux données factuelles relatives au phénomène à étudier, à des outils de formalisme mathématique et à des informations perceptuelles (JANKOWSKI et *al*., 2006). Les capacités de visualisation correspondent à des fonctionnalités cartographiques et/ou à des sorties graphiques (courbes, statistiques, etc.) qui résultent en partie de l'interrogation (requêtes attributaires et spatiales) des données produites par le système (SOH et *al*., 2006). Des fonctionnalités SIG sont donc nécessaires pour prendre en charge la gestion de ces bases de données spatiales et l'affichage composite qui en résulte. Cependant, ZERGER et SMITH (2003) soulignent les limites d'utilisation d'un SIG monolithique dans les systèmes temps-réel de la gestion de crise pour :

- fournir une résolution spatiale fine des données géographiques dans des délais opérationnels (quelques minutes)
- partager et transmettre des données et des informations à des entités opérationnelles distantes
- prendre en charge des opérations de modélisation

Pour les deux premières limitations, la principale raison est le manque de ressources informatiques. En ce qui concerne la modélisation, les auteurs invoquent les capacités même d'un SIG comme facteur limitant de leur utilisation en temps-réel, et concluent qu'elle doit être effectuée séparément et suggèrent une « prochaine génération de systèmes spatiaux d'aide à la décision qui intègrent SIG, modèles dynamiques et interface utilisateur ».

La définition d'un SSAD retenue dans le cadre de cette recherche est donc :

Un système d'information permettant de visualiser, d'intégrer, de traiter et de diffuser des données et des informations géographiques distribuées, et dédiées à une thématique spécifique qui nécessite la collaboration de plusieurs systèmes opérants pour le développement de décisions opérationnelles.

En prenant en compte ces recommandations, la spécification rigoureuse de ce SSAD, nommé G-ALHTAÏR, doit d'autre part s'appuyer sur les conclusions de l'étude menée sur l'existant technologique, organisationnel et scientifique de l'expertise hydrométéorologique des crues à cinétique rapide dans le chapitre 3.

IV - 3.2. Caractérisation globale de G-ALHTAÏR dédié à l'anticipation des crues à cinétique rapide

Étant donné la thématique principale relative à l'hydrologie, un rapprochement entre « sciences hydro-météorologiques » et systèmes d'information géographique apparaît essentiel pour le développement d'un système d'aide à la décision performant (COLOSIMO et MENDICINO, 1996). Les SIG monolithiques existants, s'ils permettent de prendre en charge la gestion de données multithématiques, ne partagent pas entièrement les mêmes fondements scientifiques que l'hydrologie. Sui et Maggio (1999) s'appuient sur l'idée fondatrice de la science de l'information géographique de GOODCHILD (1992) (cf. §.III – 3.2.3) pour promouvoir l'intégration de concepts d'hydrologie dans des fonctionnalités SIG en développant une approche modulaire, versatile et interopérable, se démarquent des SIG monolithique. Cette approche se veut :

- distribuée de manière géographique,
- désagrégée en composants logiciels indépendants mais interopérables,
- découplée

Cette philosophie correspond à celle promulguée par l'Open Geospatial Consortium dont l'utilisation est justifiée par Zhang et Li (2005) pour des applications au temps d'exécution critique et nécessitant une grande accessibilité et maniabilité par des utilisateurs géographiques distants (DIAZ et *al*., 2007). Ce SSAD doit donc s'appuyer sur les technologies émergentes de l'informatique distribuée (CASTROGIOVANNI et *al*., 2005 ; FEDRA, 1998) comme les services web, l'Internet et la technologie grille pour soutenir la prise en charge de traitements hydrologiques intensifs dans des délais opérationnels courts et assurer l'aide à la décision interactive en temps-réel entre plusieurs partenaires (RADKE et *al*., 2000).

Il s'agit dorénavant de définir l'ensemble des hypothèses de recherche au travers de la spécification précise de ce SSAD.

IV - 3.3. Spécification détaillée du système spatial d'aide à la décision au sein du SPC-GD

IV - 3.3.1. Vers une nouvelle dimension organisationnelle

Les capacités du web et de la technologie grille à faire collaborer des utilisateurs géographiquement distants permet d'envisager un système de prévision hydrologique plus intégré entre le SPC-GD et le SCHAPI, et certainement extensible à l'ensemble des SPC. Dans ce sens, le diagramme de cas d'utilisation (Figure 58) permet de formaliser ces glissements fonctionnels.

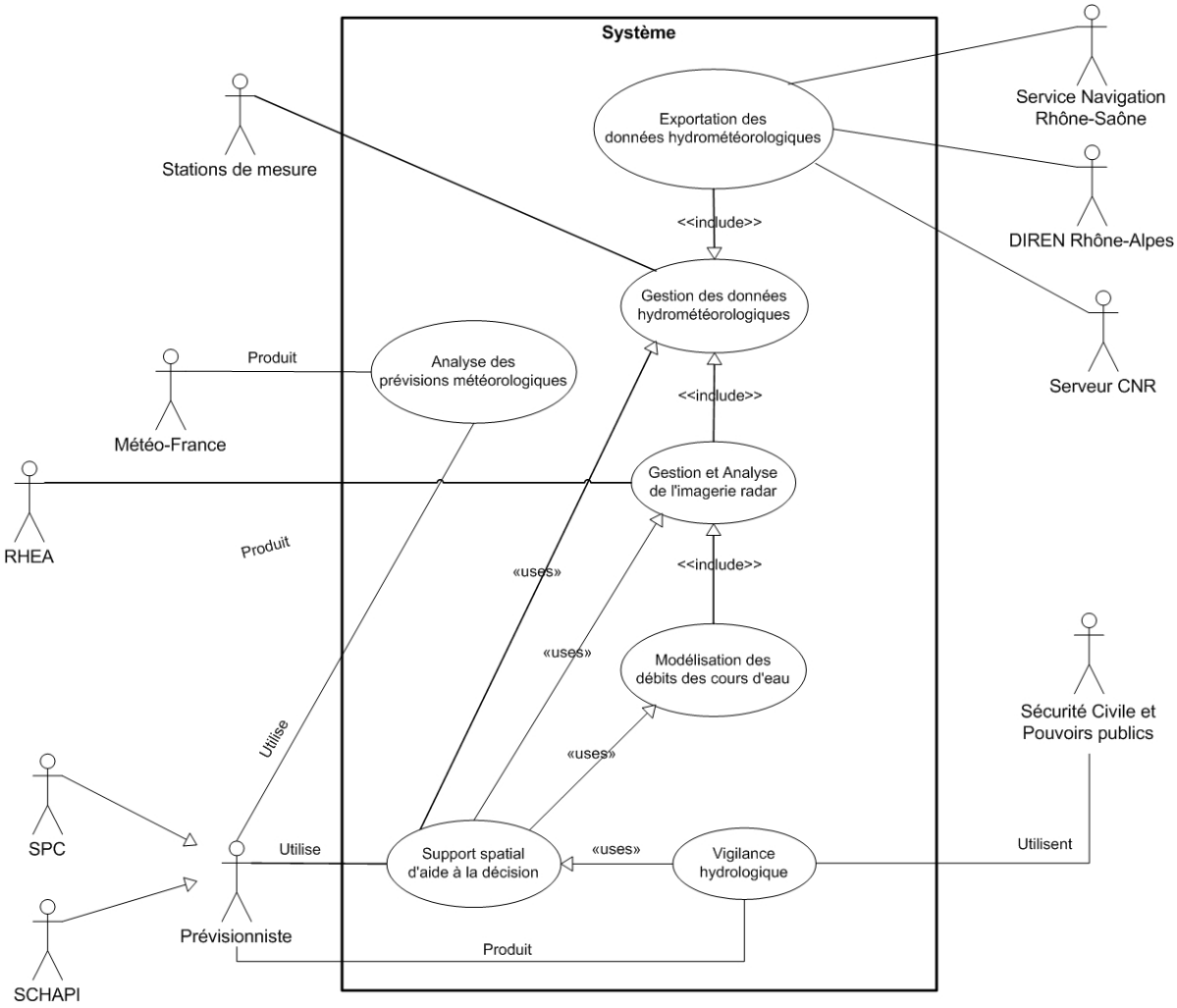


Figure 58 : Positionnement du cas d'utilisation « support spatial d'aide à la décision » dans l'organisation existante du SPC-GD

En comparaison avec le diagramme de cas d'utilisation du système existant (cf. §.III – 1.2.2. : Fig. 36), un nouveau cas d'utilisation a été ajouté, le « support spatial d'aide à la décision » qui correspond au rôle principal du SSAD. Il tend tout d'abord à concrétiser les efforts d'intégration de l'ensemble des données transitant dans le système d'information du SPC-GD, illustrés dans le passé par le développement de l'application SIGSAC (BLEUSE, 2001). Il doit permettre d'étendre les capacités technologiques de SIGSAC en intégrant les données collectées mais aussi les images radar traitées par CALAMAR 2 et les hydrogrammes modélisés par l'application ALHTAÏR.

Ce SSAD, basée sur une approche de cartographie en ligne, doit être accessible par Internet par les services autorisés en charge de la prévision hydrologique (SCHAPI et SPC). Cette amélioration permet de garantir le besoin de partage et de collaboration nécessaire au développement de l'expertise hydrologique et plus largement de la gestion de crise (cf. §.III – 2.2.2.1.2). D'autre part, la cartographie dynamique qu'offre ce type d'outils peut permettre d'adapter rapidement le rendu cartographique en fonction de l'utilisateur concerné (experts, sapeurs pompiers, maires, préfets, etc.) (HREBICEK et KONECNÝ, 2007). Enfin, les capacités d'interopérabilité et de communication du web assurent une transmission efficace de l'information. En effet, en accédant à ce portail, depuis leur propre système d'information, les utilisateurs distants peuvent lire, interroger et analyser les données disponibles sans risque d'incompatibilité logicielle.

L'assise informatique commune entre Internet (web) et la grille, et les nouvelles capacités informatiques offerte par cette dernière (cf. §.IV - 2.1.2), permet d'envisager une intégration et une complexification des opérations de modélisation hydrologique depuis ce portail. Dès lors, il apparaît envisageable de fournir un système capable de supporter plusieurs scénarios de prévision, c'est-à-dire multi-échelles, multi-modèles et basés sur des sources de données météorologiques variées, afin de mieux assister le prévisionniste dans son processus d'expertise opérationnelle.

IV - 3.3.2. Vers une capacité d'anticipation hydrologique

L'analyse du système d'information existant du SPC-GD (cf. §.III – 1.2.) a permis d'identifier ses fonctions « métier » susceptibles de tirer profit des ressources informatiques supplémentaires fournies par la grille. L'analyse des données collectées et simulées en temps réel, de part leur criticité temporelle, n'apparaît pas adaptée à une prise en charge distante par les ressources de grille, du fait en particuliers des temps de latence qui la caractérise intrinsèquement (cf. §.IV - 2.1.2). A contrario, la dimension « prévision » du cas d'utilisation « modélisation des débits des cours d'eau » offre d'une part des délais opérationnels plus longs (supérieur à 5 minutes) et d'autre part correspond à la motivation de perfectionnement principale du SPC-GD.

Ainsi, en se basant sur le diagramme UML du cas d'utilisation « Modélisation du débit des cours d'eau » (cf. §.III – 1.2.3.4. : Fig. 41), la composante modélisation du SSAD peut être formalisée par le diagramme d'activité de la Figure **59**.

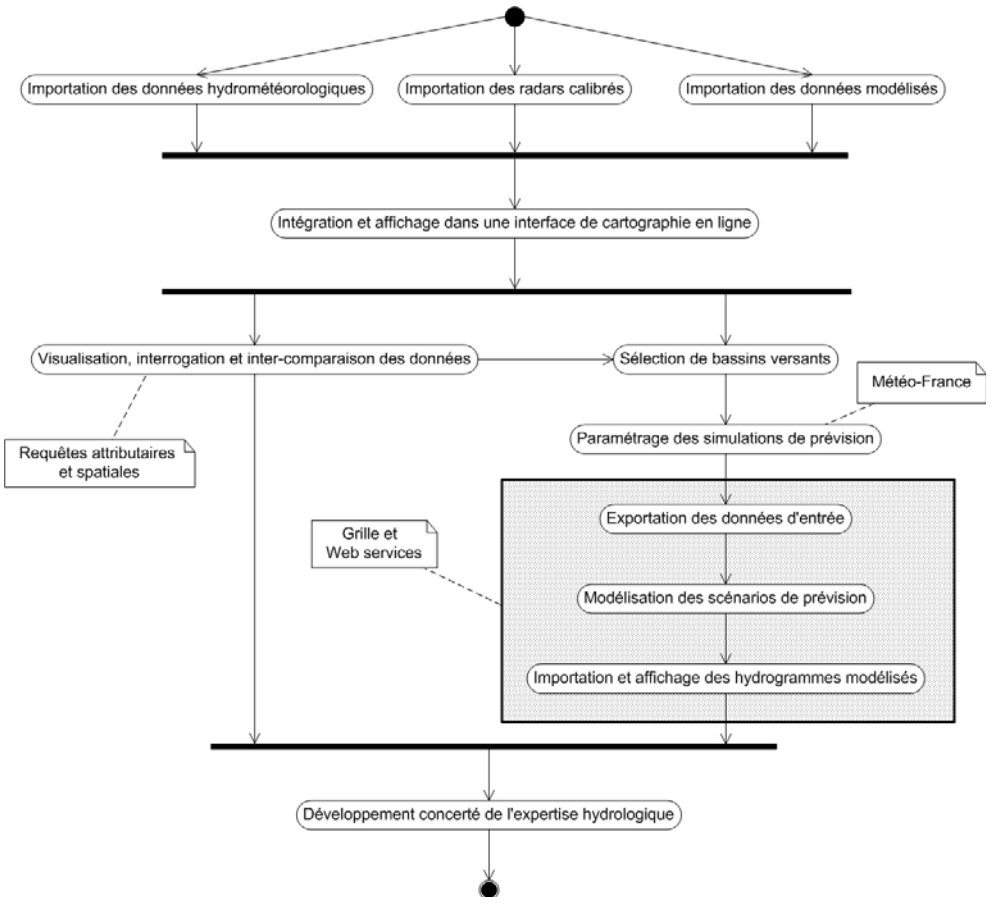


Figure 59 : Diagramme d'activité du cas d'utilisation "support spatial d'aide à la décision"

Le haut du diagramme correspond aux données transitant dans le système spatial d'aide à la décision. Une fois que l'intégration de ces données est réalisée, c'est-à-dire à chaque réactualisation des données collectées, des images radars calibrées et des hydrogrammes temps-réel modélisés par l'application ALHTAÏR, le prévisionniste analyse l'ensemble de ces données et développe une stratégie de prévision hydrologique. Pour cela, il s'appuie sur la situation hydrologique spatialisée et sur les différentes prévisions météorologiques dont il dispose, il définit alors les paramètres et données d'entrée des scénarios de prévisions qu'il souhaite exécuter.

A ce niveau, la technologie grille, de part sa capacité à effectuer des traitements en parallèle (cf. §.IV - 2.1.1.3), est utilisée pour prendre en charge les opérations de prévision hydrologique (Figure **59** : cadre gris). Le diagramme d'activité de l'application ALHTAÏR (cf. §.III – 1.2.4.2. : Fig. 43) montre un enchaînement séquentiel des différentes opérations informatiques pour transformer la pluie précipitée en débit pour chaque bassin versant. Ainsi, l'analyse de cet algorigramme permet de rejeter l'usage d'un parallélisme de contrôle qui, de plus, ne correspond pas à une utilisation classique de la grille. Ainsi, le gain de performance peut être envisagé par la prise en charge d'un nombre important de scénarios de prévision dans des délais opérationnels courts sans provoquer une surcharge de calcul sur le système local du SPC-GD (THIERION et AYRAL, 2008).

Le concept d'études paramétriques, fortement relié à l'utilisation principale d'une infrastructure de grille (cf. §.IV - 2.1.3.2), apparaît comme la meilleure plus value calculatoire que la grille puisse apporter au système d'information existant du SPC-GD.

Ainsi, afin de mieux anticiper les évènements de crues à cinétique rapide et, de réduire ou de spécifier l'incertitude sur la prévision hydrologique. Il a été démontré que l'une des sources principales de la difficulté de prévision hydrologique provient de l'incertitude existant sur les prévisions météorologiques à grand échelle et à court terme (cf. §.III – 2.2.3.2.). Le prévisionniste doit donc avoir la capacité de modéliser différents scénarios de prévision hydrologique caractérisés par :

- le(s) bassin(s) versant(s) à modéliser
- les conditions hydrologiques et les hydrogrammes « temps réel »
- le type de données de prévisions météorologiques,
- le(s) horizon(s) de prévision et des critères de variation de ces données de pluie

Ce critère de variation correspond à une pondération positive ou négative sur la donnée de prévision de pluie CALAMAR originale et constitue la méthode choisie pour injecter différents scénarios de prévision de pluie dans le modèle ALHTAÏR. Ce prototype G-ALHTAÏR a été conçu dans l'optique d'une fiabilité future des prévisions météorologiques à court terme et à petite échelle.

L'excédent d'instances de modélisation que représente cette procédure de prévisions d'ensembles et que le système d'information du SPC-GD ne semble pas apte à prendre en charge au vue de ses capacités informatiques (cf. §.III – 1.2.4.1), implique donc le recours aux ressources distantes offertes par l'architecture de grille EGEE capable de soutenir les fortes variations de charge que représente une

telle utilisation. Cependant, cette capacité nouvelle impose le recodage de l'application ALHTAÏR, initialement développée grâce au langage privé WINDEV©, en un langage interprétable par des nœuds de calcul de la grille EGEE[13].

Ainsi, le traitement d'algorithmes de manière simultanée, la forte capacité de réplication des données et de ces traitements et la politique d'utilisation stricte rendus possible par la technologie grille (cf. §.IV - 2.1.2), offre un intérêt indéniable pour faciliter leur partage sécurisé parmi des partenaires engagés dans la crise. Malgré tout, les compétences spécifiques que représentent ces opérations et la contrainte forte concernant l'approche « métier » de ce SSAD (cf. §.IV - 1.2.2) impose leur prise en charge par des « services évolués » capable d'interagir de manière autonome avec les ressources de l'intergiciel de grille (cf. §.IV - 2.2.2).

IV - 3.3.3. Vers l'orchestration des traitements de grille

L'intérêt croissant de l'utilisation des services web pour, d'une part permettre la réutilisation et la connexion de composants logiciels par des développeurs, et d'autre part faciliter l'accès transparent à ces composants et à des données distants par les utilisateurs, a été démontré (cf. §.II – 3.2.3.). L'utilisation des services web de l'OGC, déjà validé pour soutenir l'encapsulation de fonctions hydrologiques complexes (cf. §.IV - 3.2), prend tout son sens pour faciliter l'accès aux traitements de prévision par ensemble définis dans le paragraphe précédent et la transmission de ces résultats aux services en charge de la prévision hydrométéorologique (que l'on peut assimiler à la contrainte d'accès et de publication de l'information géospatiales défini dans le paragraphe IV - 1.2.1). Leur standardisation et leur reconnaissance par une large gamme de fournisseurs de logiciels et de données SIG garantissent de plus une politique commune de traitement de l'information à composante géographique et une plus grande fiabilité dans ces opérations de partage et de communication.

La Figure 60 présente l'architecture permettant la prise en charge des traitements informatiques sur la grille EGEE depuis le système d'information existant du SPC-GD.

[13] Les nœuds de calcul de la grille EGEE sont majoritairement basés sur le système d'exploitation Scientific Linux

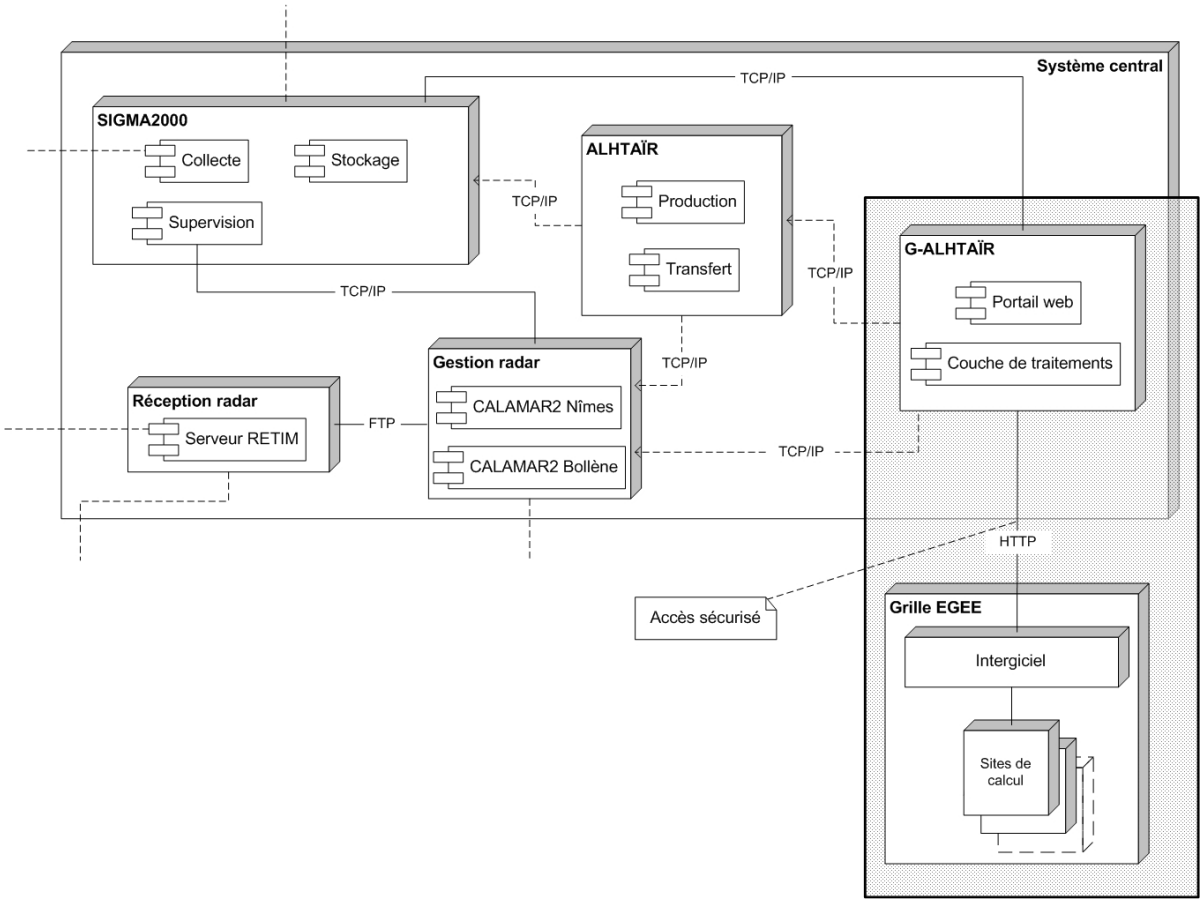


Figure 60 : Diagramme de déploiement synthétique du nouveau système

Le système spatial d'aide à la décision est intégré dans le réseau local du SPC-GD et connecté aux composants déjà existants lui fournissant l'ensemble des données et des informations nécessaires au développement de scénarios de prévision et donc *a fortiori* à l'expertise hydrologique (Figure **60** : cadre gris). Les simulations sont automatiquement envoyées et prises en charge par des instances de services web de l'OGC capables d'interagir avec les services de l'intergiciel de grille, dans la philosophie des initiatives du Geospatial Grid (cf. §.IV - 2.2.2).

Ainsi, l'instanciation et l'envoi de requêtes à ces services de l'OGC par le prévisionniste s'effectuent au travers de l'interface de cartographie en ligne présentée au IV - 3.2, connectée à Internet de manière à préserver la dimension ubiquitaire offerte par les services web et la technologie grille et intégrer de nouveaux acteurs tels que le SCHAPI.

SYNTHESE DES HYPOTHESES

Le chapitre 4 a permis de confronter les avancées scientifiques dans le domaine des grilles informatiques, tant au niveau fonctionnel qu'applicatif, avec l'existant technologique et organisationnel du SPC-GD décrit dans le chapitre 3. Ce service a ainsi été défini comme l'objet d'étude de cette recherche, de part sa longue implication dans l'innovation technico-scientifique appliquée à l'hydrologie opérationnelle et son activité opérationnelle fortement tournée vers la gestion des crues à cinétique rapide.

D'autre part, le rapprochement entre les communautés « grille » et « géomatique » offrent une nouvelle voie de recherche qu'il convient de mettre en exergue dans l'ensemble de ce manuscrit, et plus particulièrement dans la méthodologie adoptée. En effet, la majorité des données et des traitements du système d'information existant du SPC-GD relèvent de la manipulation de l'information géographique, véritable fil conducteur du développement de l'expertise hydrologique en situation de crise, et besoin avéré de la gestion de crise (chapitre 3). Ainsi, l'un des fondements principaux de cette recherche correspond au traitement de l'information géographique sur une architecture informatique distribuée, telle que la grille.

Cette situation technologique, les potentialités génériques de calcul et de partage de la technologie grille, et le rapprochement des deux communautés ont donc permis de définir les hypothèses scientifiques et méthodologiques de cette recherche permettant d'envisager les objectifs relatifs à :

- l'amélioration technologique du SPC-GD pour une meilleure prise en charge de sa mission de prévision hydrologique des crues à cinétique rapide,
- l'expérimentation de la technologie grille dans un cadre opérationnel tel que la gestion de crise.

Ainsi, ces hypothèse regroupées dans la proposition méthodologique de la fin de ce chapitre, et illustrées par le système spatial d'aide à la décision hydrologique G-ALHTAÏR, semblent susceptibles de répondre à certains besoins technologiques du SPC-GD. La structuration en composants de ce système expérimental permet d'envisager trois niveaux d'appréhension :

- l'exécution de modules hydrologiques paramétrés sur des nœuds de calcul de l'architecture de grille EGEE permettant de simuler de nombreux scénarios de prévision en vue d'appuyer la mission d'aide à la décision hydrologique du SPC-GD,
- un accès formalisé et ubiquitaire à ces fonctions de modélisation et à l'ensemble des données relatives à l'expertise hydrologique,
- une dimension collaborative forte entre les partenaires de l'expertise hydrologique.

La problématique générale d'évaluation de la technologie grille pour les besoins opérationnels de la Sécurité Civile peut donc être envisagée au travers ce système, représentatif des potentialités de la technologie grille et spécifié en fonction des besoins de l'expertise hydrologique des crues éclair.

La partie 3 de ce manuscrit s'attache donc à présenter en détail ces composants de la plateforme G-ALHTAÏR. Le chapitre 5 décrit la grillification des modules de production et de transfert présents dans l'application ALHTAÏR. Le chapitre 6 s'attache à décrire les couches d'interfaçage entre la grille EGEE et le système d'information du SPC-GD, incluses dans le système d'aide à la décision G-ALHTAÏR. Ce système d'information est susceptible d'offrir au prévisionniste un confort d'utilisation nécessaire à une situation de crise et permettent d'envisager la reproductibilité des développements informatiques garants d'une adoption progressive des technologies innovantes étudiées et évaluées dans cette recherche.

Finalement, la réalisation de cette méthodologie doit donc amener, dans la dernière partie de ce mémoire (partie 4) à une évaluation précise, et conforme aux contraintes fonctionnelles du SPC-GD, du système G-ALHTAÏR et par extension de la technologie grille.

PARTIE 3

Méthodologie expérimentale

« Le défi de la communication est moins de partager quelque chose avec ceux dont je suis proche que d'arriver à cohabiter avec ceux, beaucoup plus nombreux, dont je ne partage ni les valeurs ni les intérêts. Il ne suffit pas que les messages et les informations circulent vite pour que les Hommes se comprennent mieux. Transmission et interaction ne sont pas synonymes de communication. »

Dominique Wolton

Chapitre 5

Grillification de l'application ALHTAÏR

5.1. L'intergiciel gLite : un accès unique aux ressources d'EGEE

5.2. Méthodologie fondamentale : modélisation hydrologique sur la grille

L'intergiciel de grille gLite virtualise les ressources matérielles de calcul et de stockage offertes par la grille EGEE et d'interagit à distance avec les systèmes d'information d'utilisateurs. En grillifiant le modèle pluie-débit ALHTAÏR, utilisé en situation de crise par le SPC-GD, il devient envisageable de perfectionner la capacité de prévision hydrologique en situation de crise.

INTRODUCTION

Les conclusions du chapitre 4 ont permis d'identifier les principales composantes technologiques de G-ALHTAÏR en fonction des besoins et des contraintes fonctionnelles du SPC-GD. Trois niveaux de réflexion ont été identifiés :

- le niveau « Entreprise » correspondant à la collaboration entre le SPC-GD et le SCHAPI,
- le niveau « Traitement » permettant l'interaction entre les systèmes d'information,
- le niveau « Technologie » assurant la modélisation hydrologique sur la grille.

Ainsi, dans le cadre de ce chapitre, il est question de détailler les développements informatiques fondamentaux qui ont été menés pour permettre la modélisation hydrologique sur l'architecture de grille EGEE (niveau « Technologique »). Cette étape apparait essentielle à l'implémentation de la méthodologie proposée dans le chapitre précédent et détaillée dans cette partie. En effet, le portage ou la grillification des algorithmes requiert une connaissance précise des services de l'intergiciel de grille, et en particulier gLite l'intergiciel de la grille EGEE. En analysant les services de haut niveau de gLite, la première partie de ce chapitre doit permettre d'établir une stratégie de portage en adéquation avec les contraintes de calcul des modules de production et de transfert d'ALHTAÏR.

Dans la seconde partie du chapitre, la stratégie d'implémentation pour assurer et augmenter la capacité de modélisation allouée à la prévision hydrologique des crues à cinétique rapide, est détaillée. Les choix techniques de cette grillification sont principalement guidés par les contraintes informatiques des services web de l'OGC, pour faciliter dans un second temps l'encapsulation et la gestion standardisée de ces modules hydrologiques.

L'ensemble de ce chapitre, pivot de la partie méthodologique, doit permettre de mener les développements et les orientations du niveau supérieur, dédié à l'interaction (ou interfaçage) entre les systèmes d'information.

V - 1. L'intergiciel gLite : un accès unique aux ressources d'EGEE

Dans ce paragraphe, il est question de détailler la stratégie développée pour permettre l'exécution des modules de production et de transfert de l'application ALHTAÏR sur les ressources de calcul de l'infrastructure de grille EGEE. Ainsi, dans un premier temps, l'intergiciel de grille gLite, utilisé et développé dans le cadre du projet EGEE, est décrit. Une attention particulière est portée sur les fonctions dédiées à l'utilisation des ressources de calcul et de stockage par un utilisateur externe. Dans un second temps, la stratégie de recodage et de portage de l'application ALHTAÏR sur la grille EGEE sont développés. A ce niveau, étant donné le positionnement scientifique de cette recherche, les développements et les expérimentations qui suivent s'appuient sur les ressources informatiques de l'organisation virtuelle *Earth Science Research*[1] dédiée aux problématiques des sciences de la Terre au sein du projet EGEE.

Un intergiciel et l'ensemble de ses services permet donc de mettre en relation une communauté d'utilisateurs avec des ressources matérielles géographiquement distantes. D'autre part, « les services de l'intergiciel ont besoin de travailler ensemble de manière coordonnée, bien que ces services peuvent toujours être déployés et utilisés indépendamment » (LAURE et *al*., 2004).

En terme de services, il est possible de distinguer deux niveaux distincts, les services fondamentaux de l'intergiciel en charge des ressources matérielles (gestion des processus distants, co-allocation des ressources, accès aux stockages, découverte des ressources, sécurité et qualité de services) et ceux dédiés aux interactions avec l'utilisateur (environnements de développements, courtiers de ressources, planification des tâches) qui utilisent les interfaces fournies par les services de plus bas niveau (ASADZADEH et *al*., 2005 ; MAGOULES et *al*., 2009).

Le programme EGEE a développé son propre intergiciel de grille en s'appuyant sur les intergiciels Globus (SOTAMAYOR, 2005). Les fortes relations entre utilisateurs finaux et chercheurs en informatique, chargés du développement de cet intergiciel, ont permis de perfectionner progressivement ce dernier pour en faire à l'heure actuelle un outil puissant et adaptable aux besoins des différentes communautés impliquées dans EGEE. Comme tout intergiciel, gLite (Lightweight Middleware for Grid Computing) offre de nombreuses interfaces entre les protocoles de bas niveau des différentes ressources matérielles et les ressources externes des utilisateurs, en se basant sur la technologie des services web.

Etant donné le caractère clairement utilisateur de cette présente recherche, une attention particulière est donnée aux services de haut niveau permettant la gestion des données et des exécutions, de la sécurité et des services d'information et de découverte des ressources. Sous EGEE, Il est important de considérer deux niveaux de structure, la composition d'un site (aspect matériel) et la composition de la grille (aspect intergiciel). Concernant un site de calcul, il est possible de distinguer :

[1] https://cic.gridops.org/index.php?section=vo&vo=esr

- une interface utilisateur (UI) qui constitue le point d'accès au site et permet l'authentification des entrées externes.
- L'élément de calcul (CE) qui gère la file d'attente des tâches à exécuter sur le site
- Les nœuds de travail (WN) qui exécutent proprement dit les algorithmes
- Les éléments de stockage (SE) qui donnent accès aux serveurs d'accès aux données

Ce vocabulaire est spécifique à EGEE, il peut différer d'une grille à l'autre bien que les structurations logicielles et matérielles soient souvent les mêmes et basées sur le standard OGSA (cf. §.II – 2.3.1.3).

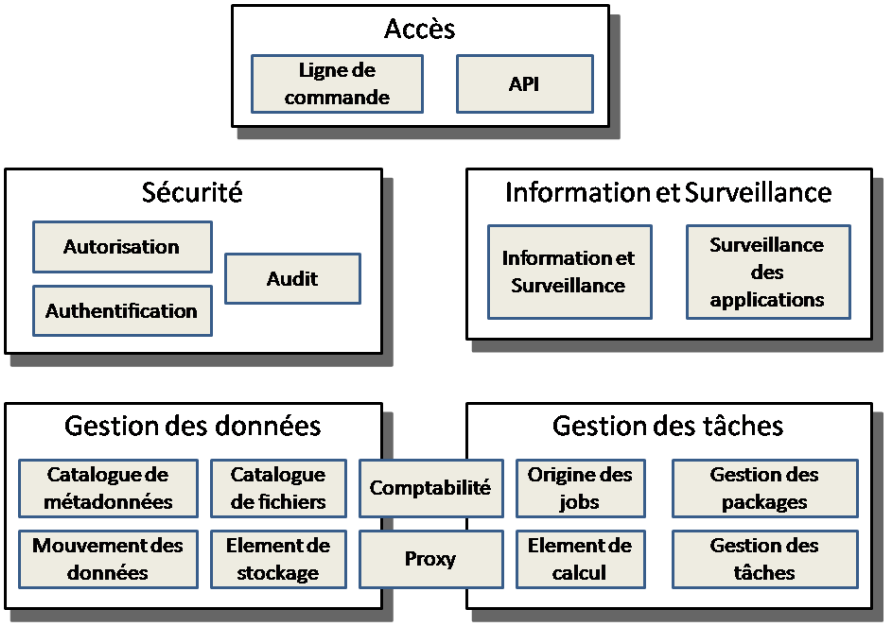


Figure 61 : Services de l'intergiciel gLite (d'après LAURE et *al.*, 2004)

Il est important de retenir que l'accès principal aux ressources de grille se fait par le biais d'une interface utilisateur, que l'organisation virtuelle, à laquelle l'utilisateur est inscrit et reconnu de confiance, met à disposition.

Au niveau de la grille dans son ensemble, il est possible d'identifier plusieurs services de haut niveau qui constituent gLite (Figure **61**). Ils sont au nombre de 5 :

- Les services d'accès
- les services de sécurité
- les services d'information et de surveillance des ressources
- les services de gestion des données
- les services de gestion des tâches

V - 1.1. Les services d'accès

L'accès à la grille se fait au travers d'une interface utilisateur (UI) généralement installée sur un site de la grille mais pouvant être aussi installé sur le système local de l'utilisateur. Il s'agit d'une interface en ligne de commande accessible au travers d'une connexion sécurisée SSH[2] qui est un protocole de communication sécurisé. De cette interface, l'utilisateur prépare, soumet, surveille et récupère les tâches de calcul (job) envoyées à distance. Il existe des interfaces graphiques permettant de faciliter cette interaction, tels que ceux développés dans les projets Ganga (BROCHU et *al*., 2009) ou g-Eclipse (GJERMUNDRØD et *al*., 2008).

V - 1.2. les services de sécurité

La sécurité est un point primordial dans une architecture de grille. En effet, le nombre d'utilisateurs d'institutions différentes et de ressources matérielles accessibles apparaissent *a priori* comme des facteurs critiques à la sureté d'une telle infrastructure. Il existe donc plusieurs mécanismes pour améliorer cette sécurité. Ainsi, tout utilisateur référencé doit posséder :

- un certificat électronique personnel
- une inscription dans une organisation virtuelle référencée dans EGEE
- un compte sur une interface utilisateur

Un certificat électronique permet d'identifier univoquement une personne, une machine ou un programme sur la grille (KNOOPS, 2007). Il correspond à une véritable carte d'identité numérique. EGEE s'appuie sur la norme X509 qui est un standard de cryptographie de l'union international des télécommunications. Ce certificat, à durée de validité limitée, est composé d'une clé publique et d'une clé privée. La première est signée par une autorité de certification (CA) et publiée sur le réseau de la grille. En France, l'autorité de certification officielle pour l'utilisation de la grille est GRID-FR[3], une sous-autorité de certification de l'autorité CNRS. La clé privée est déposée sur l'interface utilisateur de l'utilisateur (de la ressource en général) et est protégée par un mot de passe. Le chiffrement des communications repose sur l'utilisation d'algorithmes asymétriques permettant une authentification mutuelle : l'émetteur signe avec sa clé privée, alors que le destinataire vérifie la signature avec la clé publique de l'émetteur (CONTES, 2005).

Dans l'utilisation pratique de la grille, l'utilisateur accède aux ressources EGEE avec un courtier de sécurité (CONTES, 2005), appelé *proxy* dans gLite, valable 12h par défaut. Ce *proxy* accompagnera les jobs tout au long de leur processus d'exécution de manière à certifier sa légitimité ainsi que celles des sous-tâches dépendantes. Cette délégation de certification permet ainsi à l'utilisateur de ne s'identifier qu'une seule fois au moment de la création du *proxy* (single sign-on).

[2] http://tools.ietf.org/html/rfc4252
[3] http://igc.services.cnrs.fr/GRID-FR

Quant à l'inscription dans une organisation virtuelle (comme d'ailleurs l'obtention d'un compte sur une interface utilisateur), elle résulte avant tout d'un accord multipartite autorisant de nouveaux utilisateurs à accéder aux ressources allouées à une VO. D'un point de vue technique, chaque utilisateur de l'organisation virtuelle est identifié dans un serveur *VOMS* interrogé au moment de la création du proxy.

V - 1.3. les services d'information et de surveillance des ressources

La gestion de l'information dans un système grand échelle telle que la grille EGEE est cruciale pour garantir son fonctionnement efficace. S'il est encore complexe d'actualiser ces informations en temps-réel, la fréquence de mise à jour doit toutefois être la plus faible possible. Ainsi, comme tout système d'information, l'intergiciel gLite est doté de fonctionnalités permettant à chacun des éléments de la grille de connaître l'existence et l'état des autres ressources. Par exemple, un utilisateur ou un service de l'intergiciel peut à tout moment :

- vérifier l'état d'une ressource
- découvrir les ressources de la grille et leurs caractéristiques
- permettre le choix d'un nœud de calcul efficacement
- spécifier des critères d'exécution de ces tâches

Il est donc obligatoire pour chaque administrateur de site de publier et d'actualiser l'ensemble de ces informations, c'est à dire les informations générales relatives au site, la description et l'état des systèmes d'exploitation, des applications, et des données ainsi que les statistiques d'utilisation de l'ensemble des ressources.

Techniquement, deux mécanismes existent sur gLite pour gérer et accéder à ces informations de gestion :

- *Globus Monitoring and Discovery Service* (MDS) principalement utilisé pour la découverte et la publication de l'état des ressources,
- *Relational Grid Monitoring Architecture* (R-GMA) dont le fonctionnement, proche de celui du SQL (Structured Query Language), permet de surveiller et publier les informations des ressources au niveau utilisateur

Avec ces services, l'utilisateur accède aux informations relatives aux ressources matérielles dont il dispose dans sa VO :

- les éléments de calcul (*CE*)
- les éléments de stockage (*SE*)
- les courtiers de ressources (*Workload Management Server - WMS*)
-

V - 1.4. les services de gestion de données

Comme pour l'ensemble des ressources, les données utilisateur sont virtualisées sur la grille, c'est-à-dire que l'utilisateur ne sait pas exactement où ses données sont stockées. Ainsi, chaque fichier est identifié par un LFN (*Logical File Name*) et un GUID (*Grid Unique IDentifier*) qui identifient un fichier indépendamment de sa location physique. Alors que le SURL (*Storage URL*) (ou *Physical File Name* (PFN)) et le TURL (*Transport URL*) font référence aux informations physiques du fichier.

L'élément central de la gestion des données est le LFC (*LCG File Catalog*). Ce catalogue fait le lien entre un fichier, ses répliquas et ses différents types de nommage. Finalement, un ensemble de fonctions de base permettent, comme sur un système d'exploitation classique, de manipuler ces données de manière transparente. Elles sont regroupées sous le terme de *LCG utils*.

V - 1.5. Les services de gestion des tâches

V - 1.5.1. Un job de grille

Dans la suite de cet exposé, les termes « tâche » et « job » seront utilisés indifféremment pour faire référence aux traitements informatiques envoyés par l'utilisateur sur les ressources de la grille. Depuis l'interface utilisateur, les services de gestion des jobs permettent globalement à l'utilisateur :

- d'éditer les scripts des jobs grâce au langage JDL (*Job Description Language*) qui est un langage de haut niveau qui consiste à assigner des valeurs à des attributs (attribut = valeur) (PACINI et MARASCHINI, 2007)
- de soumettre ces scripts à la grille en le transmettant au courtier de ressources (WMS)
- de surveiller l'état de ces scripts tout au long de leur exécution (*Waiting*, *Submitted*, *Ready*, *Scheduled*, *Running*, *Cancelled*, *Aborted*, *Done*, *Cleared*) (Figure 62)
- de retourner les résultats (sorties) de ces scripts

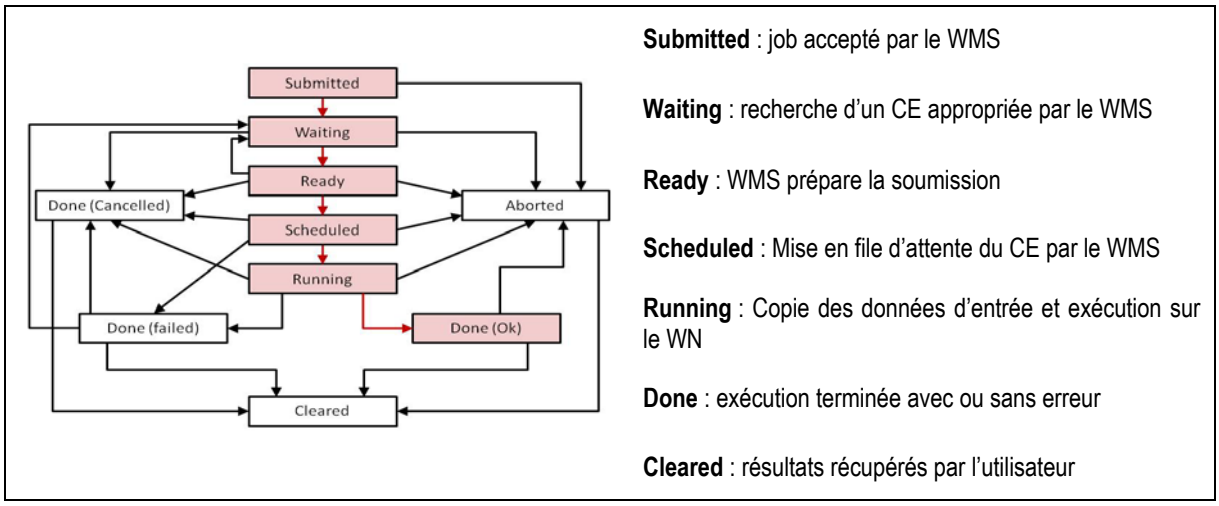


Figure 62 : États d'un job lors de son traitement par l'intergiciel (d'après BURKE et *al*., 2009 ; CHRISTODOULOPOULOS et *al*., 2008)

Schématiquement, la soumission de job aux nœuds de calcul de l'architecture EGEE correspond à l'envoi d'un paquet contenant la description du job en JDL, l'exécutable de l'algorithme, et éventuellement les données à traiter et les librairies nécessaires à l'exécution (InputSandbox). Si la

définition d'une grille s'apparente à une architecture distribuée grand échelle composée d'éléments hétérogènes, la grille EGEE impose aux sites fournisseurs de ressources une certaine homogénéité des sytèmes d'exploitation. En effet, il est fortement conseillé aux gestionnaires de sites de fournir des nœuds avec comme système d'exploitation la distribution Scientific Linux[4]. Par conséquence, un utilisateur désirant soumettre ses algorithmes sur la grille doit s'assurer qu'ils sont compatibles avec un tel environnement.

```
Type            =   "Job" ;
JobType         =   "Normal" ;
Executable      =   "/bin/sh" ;
Arguments       =   "test.sh" ;
InputSandbox    =   {"test.sh","input.txt"} ;
StdOutput       =   "std.out" ;
StdError        =   "std.err" ;
OutputSandbox   =   {"std.out","std.err","output.txt"} ;
```

Figure 63: Exemple d'un job de grille avec gLite

La description du job de la Figure 63 permet d'exécuter le script *test.sh* qui prend en entrée le fichier *input.txt* et génère la sortie *output.txt*. L'application qui permet d'exécuter ce script est le bash[5] (langage de base d'un environnement linux). La présence des attributs *InputSandbox* et *OutputSandbox* indique au courtier que l'utilisateur fournit un algorithme et une donnée d'entrée et désire recevoir le fichier de sortie et les sorties standards, une fois le job exécuté. Quant aux attributs *StdOutput* et *StdError*, ils permettent de recevoir les informations standards relatives à l'exécution du job. Il existe de nombreux autres attributs pour spécifier l'environnement d'exécution d'un job dont l'attribut *Requirements* qui permet de spécifier les contraintes d'exécution du job en se basant sur les attributs du schéma GLUE[6] (PACINI et MARASCHINI, 2007). A ce niveau, il est important de spécifier le type de jobs existants (PACINI et MARASCHINI, 2007) :

- Job
    - Normal
    - Interactif, entre le job et l'utilisateur
    - MPICH, approche de parallélisme MPI
    - Paramétrique, variation de paramètres et de données d'entrée
- DAG (Direct Acyclic Graph), gestion de chaînes de traitements
- Collection, ensemble de jobs indépendants

[4] https://www.scientificlinux.org/
[5] http://www.gnu.org/software/bash/bash.html
[6] Le schéma GLUE (Grid Laboratory for a Uniform Environment) est une description standardisée des ressources de calcul. Grâce à l'attribut *Requirements* l'utilisateur peut spécifier le type de système d'exploitation, un élément de calcul, une vitesse de processeur, etc.

V - 1.5.2. Soumission d'un job

L'envoi d'un job sur la grille déclenche une série d'opérations automatiques sur les différents services de la grille (Figure 64). Le WMS accepte (Submitted) et interprète le script du job et se renseigne auprès du système d'information (BDII, Berkeley Database Information Index) de la disponibilité de CE pour exécuter l'algorithme fourni par l'utilisateur (Waiting). Dès qu'un élément de calcul répond aux critères de sélection de l'utilisateur (Ready), le job est mis en file d'attente sur ce CE avant que le LRMS (Local Resource Management System) le place sur un WN pour être exécuté (Scheduled). Concernant les données d'entrées de l'algorithme, l'utilisateur a le choix de les transmettre avec le job (Input Sandbox) ou de les déposer au préalable sur un SE. Dans ce cas, au moment de la réception du job, le WMS vérifie la présence de ces données en interrogeant le catalogue LFC. Enfin, tout au long du processus de traitement du job, l'utilisateur peut se renseigner sur son état en interrogeant le service d'enregistrement et de comptabilité (LB, Logging and Bookkeeping) (Figure **64**).

Un tel fonctionnement est considéré comme asynchrone, dans le sens où l'opération de soumission se termine sans être contrainte d'attendre une réponse immédiate du courtier de ressource. Comme cela a déjà été abordé précédemment, le script du job, l'algorithme et les éventuelles données à traiter sont accompagnés du proxy temporaire permettant à ce paquet de transiter d'un élément de la grille à un autre en toute sécurité.

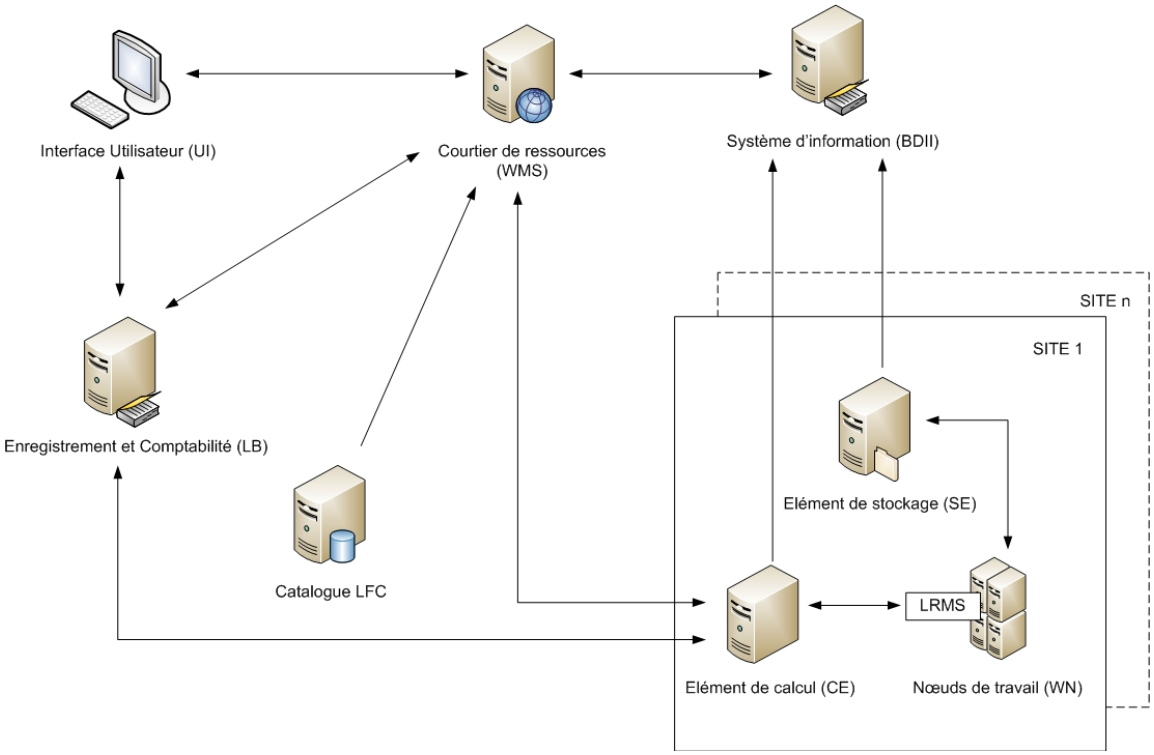


Figure 64 : Services de gLite impliqués dans la soumission d'un job (d'après BURKE et *al.*, 2009)

En conclusion, l'intergiciel gLite et les services de haut niveau qu'il fournit permet à l'utilisateur de disposer de nombreuses ressources de calcul et de stockage sans se préoccuper, s'il le souhaite, du choix de ces dernières (Attribut *Requirement* du JDL). En contrepartie, leur éloignement géographique et surtout les nombreuses interactions entre les services de l'intergiciel pour réaliser l'exécution demandée impliquent des temps de latence relativement long (GERMAIN-RENAUD et *al.*, 2008) qui sont

susceptibles de détériorer la performance des applications temps réel. Finalement, en ce qui concerne la grillification des applications, il est important de rappeler la nécessité d'implémenter des exécutables Linux et de s'assurer de la disponibilité des librairies nécessaires à leur exécution sur les éléments de calcul[7].

[7] Seules les librairies par défaut des distributions Scientific Linux sont présentes sur les éléments de calcul

V - 2. Méthodologie fondamentale

V - 2.1. Principe de fonctionnement du scénario de prévision

Comme cela a été décrit dans le chapitres 2 et 4, l'intérêt d'utiliser la grille dans cette étude réside initialement dans sa capacité à déporter des calculs susceptibles de surcharger les capacités informatiques du système d'information du SPC-GD (cf. §.III – 1.2.41). La méthodologie qui a émergé de l'analyse du chapitre 4 consiste donc à l'exécution d'un ensemble[8] de scénarios de prévisions de manière simultanée sur les nœuds de calcul de la grille. Par conséquence, un scénario de prévision correspondra à un job de grille.

D'un point de vue fonctionnel, l'analyse d'ALHTAÏR a permis d'identifier une indépendance entre les cellules du bassin versant mais une dépendance temporelle entre les instances de simulation successives (cf. §.III – 1.2.4.2). Si, en début d'évènement, les conditions hydrologiques du bassin versant sont considérées comme nulles (modèle hydrologique évènementiel), ses conditions assurent ensuite le « lien hydrologique » entre deux pas de temps successifs. Ainsi, pour permettre ce lien, entre la fonction « temps-réel » assurée par l'application ALHTAÏR existante et la version « prévision » portée sur la grille, il a été nécessaire d'enregistrer les conditions hydrologiques et l'hydrogramme à chaque pas de temps, de manière à permettre leur transfert et leur prise en charge ultérieure et à distance par le scénario de prévision. Ainsi, à chaque fois que le prévisionniste désire exécuter un scénario de prévision, le système d'aide à la décision doit récupérer les dernières conditions hydrologiques et l'hydrogramme du bassin versant, modélisées grâce à la version « temps réel » pour l'utiliser lors de l'exécution du scénario de prévision sur la grille (Figure **65**).

Afin de pouvoir modéliser un scénario de prévision sur la grille EGGE, les fichiers transférés au nœud de calcul (*InputSandbox*) correspondent aux :

- bassin versant à modéliser
- conditions hydrologiques « temps réel »
- l'hydrogramme « temps réel »
- la prévision de pluie CALAMAR
- le proxy créé grâce au certificat personnel
- les algorithmes d'ALHTAÏR et les librairies spécifiques

[8] Le terme de « simulation » est réservé dans la suite de ce mémoire pour évoquer cet ensemble de scénarios

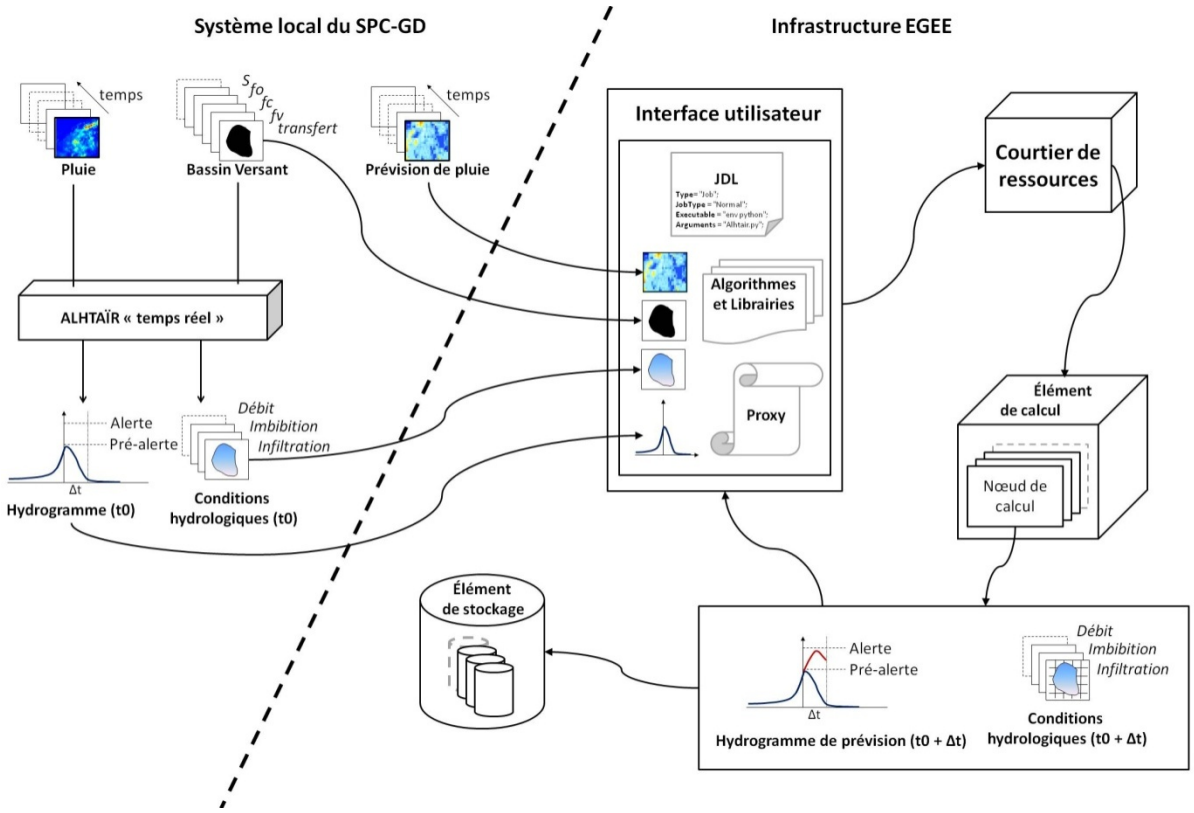


Figure 65 : Synopsis du traitement d'un scénario de prévision

Une fois que le prévisionniste soumet ce scénario les services de gestion de tâches de gLite se charge de soumettre le job correspondant à un élément de calcul disponible, qui, une fois l'exécution terminée retourne les nouvelles conditions hydrologiques et l'hydrogramme de prévision sur l'interface utilisateur et/ou les dépose sur un élément de stockage.

Il convient de détailler la méthodologie qui a permis d'une part l'exécution des algorithmes des modules de production et de transfert du modèle ALHTAÏR et d'autre part la gestion des données entre ces modules, et leur échange entre le système local et les ressources de la grille. Ces développements ont été menés depuis une interface utilisateur[9] installée dans les locaux de l'IPGP[10] à Paris, appartenant à l'organisation virtuelle ESR[11], constituant une des VO de la grille EGEE.

[9] Moyen par défaut d'accès aux ressources de grille (détaillée dans ce chapitre)
[10] Institut de physique du globe de Paris
[11] Earth Science Research.

V - 2.2. Stratégie de portage

V - 2.2.1. Recodage des algorithmes

La première étape de développement concerne la grillification des modules hydrologiques de l'application ALHTAÏR. Initialement développée sous WINDEV©, langage exclusivement réservé au système d'exploitation WINDOWS©, ces modules ont du être recodés avec un langage exécutable sur un environnement Linux.

Sur les nœuds de calcul de la grille EGEE, il existe majoritairement 4 langages de programmation[12] :

- JAVA[13]
- C[14]
- C++
- Python[15]

Il ne s'agit pas dans ce développement de discuter l'intérêt de chacun de ces langages. Dans le cadre de cette méthodologie, le langage de script Python a été choisi pour sa facilité d'utilisation, sa légèreté et sa modularité. En tant que langage interprété, il ne nécessite pas de compilation préalable du programme. D'autre part, il s'agit d'un langage de haut niveau, ne nécessitant pas la déclaration des variables. Malgré une certaine lenteur d'exécution, en comparaison avec les langages C et C++, sa rapidité de développement et la possibilité de développer une approche objet en fait un langage adapté à la problématique de cette recherche (PRECHELT, 2000).

Après l'analyse du code de l'application WINDEV©, un diagramme de classes UML a été créé pour faciliter le recodage de l'application (Figure **66**). Les éléments essentiels de la modélisation sont les caractéristiques du bassin versant et de l'image radar qui permettent de calculer le ruissellement de surface pour chaque cellule du bassin versant (production) et de construire l'hydrogramme à l'exutoire de ce dernier (transfert). Au préalable, les correspondances géographiques sont établies entre les cellules (au sens de pixel) du bassin versant et celles de l'image radar, c'est-à-dire qu'à une cellule de bassin versant correspond une cellule radar.

[12] Ces langages font partie des paquets officiels de la distribution Scientific Linux
[13] http://www.java.com/fr/
[14] http://gcc.gnu.org/
[15] http://www.python.org/

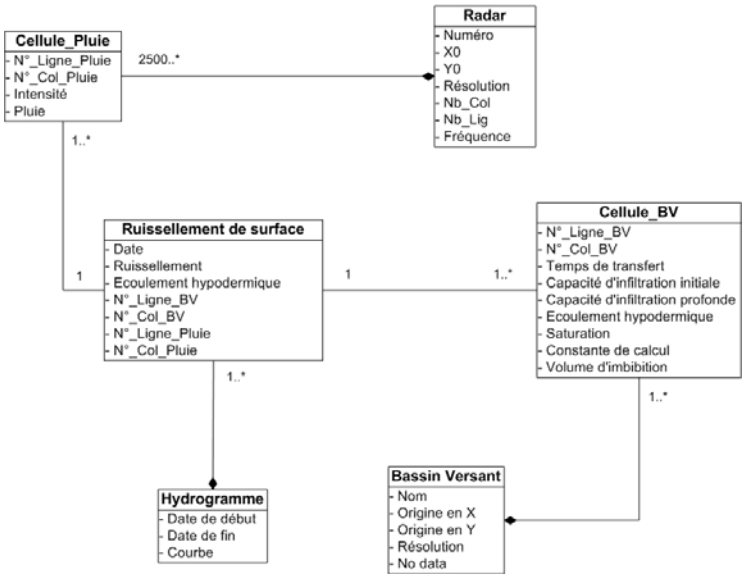


Figure 66 : Diagramme de classes des modules hydrologiques d'ALHTAÏR

La gestion de ces matrices est permise par le module externe NumPy[16] capable d'interagir avec les fonctions standards de Python. Finalement, la séquence de traitement du cœur de l'application ALHTAÏR décrit dans le chapitre 2 (cf. §.III – 1.2.4.2 : Fig. 43), correspondant aux modules de production et de transfert, n'a pas été modifiée, seule la structure, le format et l'accès aux données ont été adaptés notamment dans l'optique de l'utilisation des services web de l'OGC.

V - 2.2.2. Formatage des données

La seconde étape de la reconstruction des algorithmes fonctionnels d'ALHTAÏR concerne le formatage des données. Le format sélectionné est le NetCDF[17] (Network Common Data Form). Ce format de données permet la gestion efficace de données spatiotemporelles et multidimensionnelles. Il est basé sur des :

- dimensions,
- variables basées sur une ou plusieurs dimensions,
- attributs.

Un tel format permet de prendre en charge un grand nombre de données, dont celles issues du monitoring de l'environnement. L'intérêt principal est de pouvoir stocker un nombre important de matrices dans un seul fichier afin de réduire la multiplication des fichiers de stockage. Un ensemble de métadonnées, décrites grâce au langage CDL (Common Data Language) et basées sur la convention CF (Climate Forecast) (EATON et *al.*, 2009) (Figure **67**), est établi pour décrire et documenter les données stockées et recoder les données utilisées dans ALTHAÏR de manière standard.

Le recodage des données relatives à ALHTAÏR concerne :

- le bassin versant composé de 8 variables principales :
    - les limites du bassin versant

[16] http://numpy.scipy.org/
[17] http://www.unidata.ucar.edu/software/netcdf/

  - les temps de transfert de chaque cellule du bassin versant
  - les 6 paramètres spatialisés d'ALHTAÏR (AYRAL, 2005)
  - et les variables relatives au positionnement géographique des cellules
- la pluie issue des images radar composée d'une variable principale
  - la pluie brute spatialisée
  - et les variables relatives au positionnement géographique des cellules
- les conditions hydrologiques du bassin versant composées des situations :
  - d'infiltration
  - d'imbibition
  - de ruissellement
  - d'écoulement hypodermique
- le ou les hydrogrammes de chaque bassin versant

```
{
 dimensions:
  time = UNLIMITED ;    // (496 currently)
  x = 50 ;
  y = 50 ;
 variables:
  int Lambert_Conformal ;
    :grid_mapping_name = "lambert_conformal_conic" ;
    :…
  double rainfall(time=496, y=50, x=50) ;
    :units = "mm/h" ;
    :long_name = "Rainfall intensity" ;
    :coordinates = "lat lon" ;
    :grid_mapping = "Lambert_Conformal" ;
    :_CoordinateSystems = "ProjectionCoordinateSystem LatLonCoordinateSystem" ;
…
  double time(time=496) ;
    :long_name = "rainfall time" ;
    :standard_name = "time" ;
    :units = "seconds since 1970-01-01 02:00:00" ;
    :axis = "T" ;
    :_CoordinateAxisType = "Time" ;
    :bounds = "time_bnds" ;
  …
  float lat(x=50, y=50) ;
    :units = "degrees_north" ;
    :long_name = "latitude coordinate" ;
    :standard_name = "latitude" ;
    :_CoordinateAxisType = "Lat" ;
  float lon(x=50, y=50) ;
    :…
}
```

Figure 67 : Extrait de la description CDL d'une donnée de pluie CALAMAR, basée sur la convention CF

Grâce au format NetCDF, il est possible de mieux prendre en charge la dimension temporelle des données, essentielle dans le cas de la modélisation hydrologique. Il est donc possible de représenter l'ensemble d'un évènement hydrométéorologique au travers d'un seul fichier NetCDF. A noter que ce format est accessible par le module externe Scientific Python[18] permettant de construire, lire et modifier

[18] http://dirac.cnrs-orleans.fr/plone/software/scientificpython

ce type de données. L'intérêt du NetCDF réside dans sa capacité à intégrer des données multidimensionnelles de sources différentes permettant ainsi un accès facilité à des sous-ensembles par l'interrogation du fichier. Dans ce contexte, la motivation principale dans l'utilisation du NetCDF concerne sa prise en compte par l'OGC, et son rapprochement avec le service dédié à la gestion des couvertures géographiques (WCS) (DOMENICO et NATIVI, 2009). En effet, par sa capacité à s'auto-décrire, un service web peut l'interroger et accéder au sous-ensemble requis par un utilisateur ou un autre service. De plus, les efforts de développement sur le NetCDF et les différentes librairies qui le supportent, s'orientent vers une meilleure capacité à paralléliser leurs traitements (HARTNETT, 2009).

Finalement, le format NcML, qui est une représentation XML du NetCDF, a été utilisé pour coder les hydrogrammes des scénarios de prévision. Par sa capacité à décrire formellement les données et les métadonnées contenues dans un fichier NetCDF, il tend à améliorer l'interopérabilité dans les échanges de données. Son extension NcML-GML permet par exemple son intégration dans des logiciels SIG et sa meilleure prise en compte pour l'approche SOA (NATIVI et *al*., 2005).

L'implémentation de cette nouvelle version des modules de production et de transfert de l'application ALHTAÏR, grâce au langage python et au format de données NetCDF, peut être caractérisée par les indicateurs du Tableau **5**.

Données / algorithmes	Caractéristiques informatiques
Bassin versant	de 2 MB à 20 MB (19 MB pour 545 km^2)
Radar	454,3 KB (1h) et 3 MB (8h)
Conditions hydrologiques	14,6 MB (5 mn) et 13,2 GB (100h)
Hydrogramme	4,7 KB (NcML) et 3,8 KB (NetCDF)
Algorithmes et librairies	3,7 MB

Tableau 5 : Caractéristiques techniques des données et algorithmes après grillification

À la vue de ces caractéristiques, et étant donné les capacités technologiques de la grille EGEE l'ensemble de ces contraintes de traitement apparait relativement faible. Cependant, chaque instance de modélisation impose la prise en charge de l'ensemble de ces données et l'actualisation des conditions hydrologiques, de l'image radar et de l'hydrogramme à chaque opération de prévision. Cette mise à jour, peut représenter une limite à la rapidité de traitement, en particulier dans le cas où le réseau informatique externe utilisé pour accéder aux ressources de grille a une bande passante de faible capacité.

V - 2.2.3. Simulation hydrologique sur la grille EGEE

L'utilisation la plus classique de la grille EGEE se fait au travers du protocole SSH (cf. §.V - 1.1). Un utilisateur reconnu par une organisation virtuelle d'EGEE, dans le cas de cette étude *Earth Science Research*, c'est-à-dire disposant d'un certificat électronique se connecte à une des interfaces utilisateur fournie par certains sites de la VO. L'interface utilisateur peut permettre de stocker des algorithmes, des données et des librairies, après accord avec l'administrateur de cette UI. A partir de cette UI, l'utilisateur peut avoir accès aux services de haut niveau de gLite pour effectuer des traitements sur les ressources de la grille.

Dans le cadre de cette étude, une instance de prévision est basée sur le job de grille décrit dans la Figure **68**. Elle implique la présence, sur l'interface utilisateur, des fichiers contenus dans *InputSandbox*, qui, avec ce script et le proxy de sécurité, composera le paquet du job.

```
Type            =   "Job" ;
JobType         =   "Normal" ;
Executable      =   "env python" ;
Arguments       =   "Alhtair.py" ;
InputSandbox    =   {"Alhtair.py","alhtair.tar.gz","bassin-versant.nc","conditions-
                    initiales.nc","prévision-pluie.nc","hydrogramme-tps-reel.nc"} ;
StdOutput       =   "std.out" ;
StdError        =   "std.err" ;
OutputSandbox   =   {"std.out","std.err","hydrogramme.xml","post-conditions.nc", "logfile.log"} ;
  Requirements  =   "other.GlueHostArchitecturePlatformType == "i686"" ;
```

Figure 68 : Job de grille (JDL) pour exécuter les modules hydrologiques d'ALHTAÏR

La chaîne de traitements classique lors de l'exécution d'un job est la suivante :

- l'utilisateur créé un proxy de sécurité grâce à ses certificats (*voms-proxy-init*) et son mot de passe,
- il soumet ce job à l'intergiciel (glite-wms-job-submit) qui sélectionne un WMS disponible pour traiter la requête,
- un WMS accepte le job et prépare le paquet pour sa soumission au CE sélectionné grâce à la contrainte exprimée par l'attribut *Requirement* (en l'occurrence une machine 32 bits),
- le CE réceptionne le job, le transmet à un des ses WN, et transfère les données d'entrée (*InputSandbox*),
- l'exécution débute sur le WN par la préparation des données[19] et de l'environnement d'exécution (librairies *Scientific Python*, *Numpy* et *netcdf*) puis lance les modules hydrologiques (*Alhtair.py*),
- l'exécution se termine en préservant uniquement les données de sortie qui sont soit récupérées par l'utilisateur à la fin du job (*OutputSandbox*) soit/et enregistrées sur un élément de stockage avec un enregistrement dans le LFC,
- l'utilisateur tenu informé tout au long du traitement du job (glite-wms-job-status), récupère les données de sortie sur son UI (glite-wms-job-output).

L'envoi de manière asynchrone de ce job au courtier de ressource provoque une série d'opérations au niveau de l'intergiciel transparents à l'utilisateur qui peut à tout moment se renseigner sur son état d'avancement. Une fois terminée, il est en charge de la récupération des résultats de la simulation.

[19] À noter que des données déjà stockées sur des SEs peuvent être transférées sur le WN grâce aux fonctionnalités LCG

De prime abord, une telle architecture asynchrone permet de soumettre un nombre important de jobs en simultanée afin de répondre à un besoin ponctuel ou constant de ressources de calcul et/ou de stockage dans le cas d'une gestion intensive de données (réplication, échange, etc.), comme cette recherche le nécessite à travers la gestion d'ensembles de scénarios de prévision hydrologique. Dans une optique scientifique, visant à simuler différents scénarios, à traiter un jeu de données important, ou paralléliser des traitements intensifs[20], l'intergiciel gLite et ses services offrent un ensemble de commandes bash permettant de préparer les jobs, gérer des données sur SE, soumettre des jobs, suivre l'évolution de leur exécution, etc. Cependant, dans une optique opérationnelle, cette méthode de communication ne semble pas réellement approprier à :

- la compétence des prévisionnistes du SPC-GD
- la contrainte temporelle caractéristique d'une situation de crise

Ainsi, l'encapsulation de ces opérations de bas niveau par un composant logiciel capable d'interagir avec les services de gLite et le système d'information du SPC-GD se révèle nécessaire. Dans ce sens, la méthodologie de recours à une application web, tel qu'une interface de cartographie en ligne, connectée à des services web médiateurs doit être présentée.

[20] Il s'agit d'*embarrassingly parallel* (aucun équivalent français trouvé) où l'effort de parallélisation (au sens de parallélisation de contrôle) est limité voire inexistant

SYNTHESE

Le premier niveau de cette phase d'implémentation, détaillé dans ce chapitre, consiste donc à adapter le code existant à un fonctionnement de grille. Il est question de la grillification ou du portage de l'application. Une fois autorisé par l'organisation virtuelle à laquelle il appartient, l'utilisateur (ou son système d'information) se connecte à cette interface utilisateur et envoi des requêtes asynchrones pour exécuter l'application grillifiée aux services de l'intergiciel de grille. La relative autonomie de ces services, détaillée dans la première partie de ce chapitre, facilite la prise en charge des opérations soumises par l'utilisateur.

L'application ALHTAÏR, et en particulier ses modules de modélisation hydrologique, a donc été grillifiée sur les ressources de l'organisation virtuelle ESR. Un recodage de ces algorithmes a été nécessaire et une adaptation de la gestion des données d'entrée et de sortie a été menée pour faciliter la reproductibilité des traitements. Finalement, les capacités informatiques offertes par les nombreuses ressources de calcul de la VO ESR ont permis d'envisager la prise en charge de nombreux scénarios de prévision de manière simultanée.

Cependant, l'inconvénient principal de cette approche se situe dans l'excédent de charge de travail et les compétences spécifiques qu'elle implique. Dans une dimension opérationnelle temps réel, les aspects technologiques doivent être idéalement minimisés afin de libérer l'utilisateur de toutes contraintes fonctionnelles et lui permettre une réalisation optimale de la mission pour laquelle il est astreint. De plus, les premiers tests de modélisation hydrologique grâce à ce portage ont permis d'identifier une forte variabilité des performances de calcul et un taux d'échec incompatible avec les besoins de l'expertise hydrologique en situation de crise.

Dans ce sens, le recours à des services web de l'OGC interfacés avec les services de l'intergiciel gLite et une interface graphique par laquelle le prévisionniste établit ces scénarios de prévision apparait indispensable à la mission d'expertise hydrologique allouée au SPC-GD. D'autre part, une meilleure prise en charge des jobs de grille doit être envisagée pour améliorer les performances de modélisation.

Chapitre 6

Interfaçage entre la grille EGEE et le SPC-GD

6.1. G-ALHTAÏR : un outil garant de l'interopérabilité et de l'orchestration des traitements

6.2. RRM-Grid : un outil garant de l'ordonnancement et d'une qualité de service

Les moyens d'accès aux ressources informatiques de l'architecture de grille EGEE par le biais de l'intergiciel gLite ne coïncide pas fondamentalement avec les contraintes de fonctionnement du SPC-GD. En effet, l'antagonisme des compétences impose en particulier le recours à des services web géographiques et à une interface de cartographie permettant d'assurer l'opérationnalité requise du SPC-GD.

INTRODUCTION

Dans le cadre de ce chapitre, il est question de détailler les développements informatiques qui ont été menés pour implémenter l'ensemble des composants permettant la réalisation du niveau « traitement » de G-ALHTAÏR. Les conclusions du chapitre 5 ont soulevé la nécessité d'interfacer l'implémentation de bas niveau des modules hydrologiques (grillification) pour permettre :

- l'accès rapide et convivial à ces traitements par un prévisionniste durant une situation de crise,
- une performance de ces traitements en adéquation avec les contraintes fonctionnelles et non-fonctionnelles de la gestion de crise.

La prise en charge de ces contraintes nécessite donc des composants intermédiaires aptes à assurer une synchronisation et une qualité de service des traitements effectués à distance sur la grille EGEE. Deux axes méthodologiques ont donc été développés pour assurer cet interfaçage :

- des services web de l'OGC doivent assurer la prise en charge des scénarios construits, dynamiquement et en fonction de plusieurs paramètres hydrométéorologiques, par le prévisionniste.
- un ordonnanceur de haut-niveau doit assurer le traitement efficace de ces scénarios, en assurant leur prise en charge effective et ordonnancée par les nœuds de calcul de la grille EGEE

Ainsi, dans la première partie de ce chapitre, une réflexion méthodologique est développée de manière à identifier et justifier clairement les orientations d'implémentation choisies pour ces méthodes d'interfaçage entre grille et système local. Dans une seconde partie, le système spatial d'aide à la décision hydrologique G-ALHTAÏR est présenté, au travers des technologies des services web de l'OGC et d'une solution de cartographie en ligne. Enfin, la dernière partie s'attache à considérer la question fondamentale de la qualité de services dans les traitements informatiques inclus dans la gestion de crise des crues à cinétique rapide, par le biais de l'outil RRM-Grid.

Cette implémentation doit permettre une évaluation précise des performances de l'architecture de grille EGEE, afin de discuter la validité de cette technologie pour soutenir la nouvelle mission d'anticipation de l'aléa « crue à cinétique rapide » du SPC-GD.

VI - 1. Notes introductives : la problématique d'interfaçage

L'un des enjeux principal dans l'utilisation des ressources de grille correspond à la communication et à la synchronisation des systèmes informatiques géographiquement distribués. En l'occurrence, le système d'information du SPC-GD, dont le déploiement existant doit être faiblement impacté, doit échanger des messages et des données avec les ressources de calcul et de stockage d'EGEE. Il convient donc de détailler la structure logicielle des couches « logique » et « métier » de G-ALHTAÏR permettant cette interaction et l'utilisation de manière transparente des ressources de grille EGEE.

Le moyen par défaut offert par les organisations virtuelles pour accéder à ces ressources, concerne l'interface utilisateur (UI) qui constitue un des composants de l'intergiciel de grille, en l'occurrence gLite (Burke *et al*., 2009). Une fois autorisé par l'organisation virtuelle à laquelle il appartient, l'utilisateur (ou son système d'information) se connecte à cette interface utilisateur et envoi des requêtes asynchrones pour exécuter l'application grillifiée aux services de l'intergiciel de grille. La relative autonomie de ces services, détaillés dans le chapitre 5, facilite la prise en charge des opérations soumises par l'utilisateur. Cependant, , à la vue des conclusions du chapitre 5, cette méthode classique a plusieurs inconvénients en regard des contraintes opérationnelles.

Dans ce sens, l'une des contraintes fortes de l'utilisation d'opérations informatiques en situation de crise concerne l'obligation de s'appuyer sur des « interfaces graphiques conviviales » (cf. §.IV – 1.2.2.). Ainsi, la réussite opérationnelle de l'expertise hydrologique doit passer par l'implémentation d'un portail graphique enrichie de fonctionnalités cartographiques et graphiques afin d'améliorer le confort d'utilisation de l'expert et soutenir au mieux le processus d'aide à la décision sous-jacent. Cette interface a l'obligation d'être connectée à Internet ou à un réseau privé, de manière à être connectée à l'organisation virtuelle à laquelle elle est rattachée et indirectement aux services de l'intergiciel de grille. La deuxième partie de ce chapitre est donc dédiée à la description fonctionnelle et technologique de ce composant de haut niveau[1].

À ce niveau, une couche « logique » doit être implémentée pour enrichir les potentialités fonctionnelles de cette application de cartographie en ligne. Cette couche logique doit formaliser et favoriser l'interaction entre le composant de haut niveau et les services de l'intergiciel. Deux stratégies d'implémentation ont été adoptées. En effet, l'expérience acquise et les partenaires « ressources »[2] de ce travail de recherche ont permis de développer plusieurs orientations technologiques en fonction des différents besoins énoncés, qu'il convient de présenter dans ce chapitre et de discuter dans le chapitres 7 et 8.

À la vue des conclusions du chapitre 4 qui a démontré l'intérêt des services web pour soutenir l'interopérabilité des systèmes d'information, il semble primordial d'évaluer l'adaptabilité et l'opérationnalité de ces services web dans leur rôle de médiateur. De par leur formalisation, ils sont

[1] Le terme « haut niveau » est utilisé pour faire référence aux composants logiciels permettant la prise en charge des fonctions « métier » d'une application

[2] Il s'agit du laboratoire Earth and Space Science Informatics - Laboratory (ESSI-Lab) de l'IMAA (Institute of Methodologies for Environmental Analysis) en Italie, partenaire du projet européen CYCLOPS (cf. §.IV – 2.2.3.1.2) et des administrateurs de la VO Earth Science Research de la grille EGEE

interprétables et interrogeables aussi bien par l'homme que par un ordinateur. De plus, le constat positif du rapprochement entre l'OGC et l'OGF, garantissant progressivement la viabilité de l'approche SOA pour le fonctionnement et l'utilisation de la grille, a appuyé ce choix (cf. §.IV – 2.2.2.). Enfin, la spécificité des services web de l'OGC, relative à la prise en charge de géotraitements et de données géospatiales offrent un socle technologique solide pour enrichir les services classiques offerts par l'intergiciel de grille. Les détails de leur implémentation sont décrits dans la troisième partie de ce chapitre.

En parallèle de cette approche SOA, orientée « interopérabilité », des questions plus technologiques ont émergé dans cette phase d'implémentation. En effet, deux des contraintes définies dans le chapitre 4 (cf. §.IV – 1.2.) sont respectivement relatives à d'une part la « réservation et de priorisation des ressources informatiques » et d'autre part au « fonctionnement en temps réel des applications ». Or, le bilan des caractéristiques informatiques de la grille EGEE (cf. §.IV – 2.1.2.), bien que relativement positif, a permis de souligner deux manquements dans le fonctionnement intrinsèque de l'intergiciel d'EGEE :

- les temps de latence importants pour le traitement des tâches soumises
- l'absence de fonctionnalités pour prioritiser les tâches soumis et préempter à l'avance des ressources de calcul

Ainsi, le développement d'un ordonnanceur évolué apparaît nécessaire pour enrichir les capacités initiales de l'intergiciel gLite. Cet ordonnaceur-lanceur, dénommé RRM-Grid (*Rainfall-runoff modeling on Grid par la suite*), cherche à contrôler d'une part l'ordre de soumission des tâches de modélisation définies par le prévisionniste et d'autre part à diminuer les temps de latence de l'exécution de ces tâches en optimisant en particulier les transferts de données. Finalement, ce composant doit permettre de maximiser les taux de réussite des tâches soumises dans l'optique forte de la sécurité (au sens de fiabilité) des services opérants de la Sécurité Civile. Les paragraphes de cette dernière partie sont largement basés sur la publication soumise et acceptée à *Journal of Grid Computing* (THIERION et *al.*, 2010).

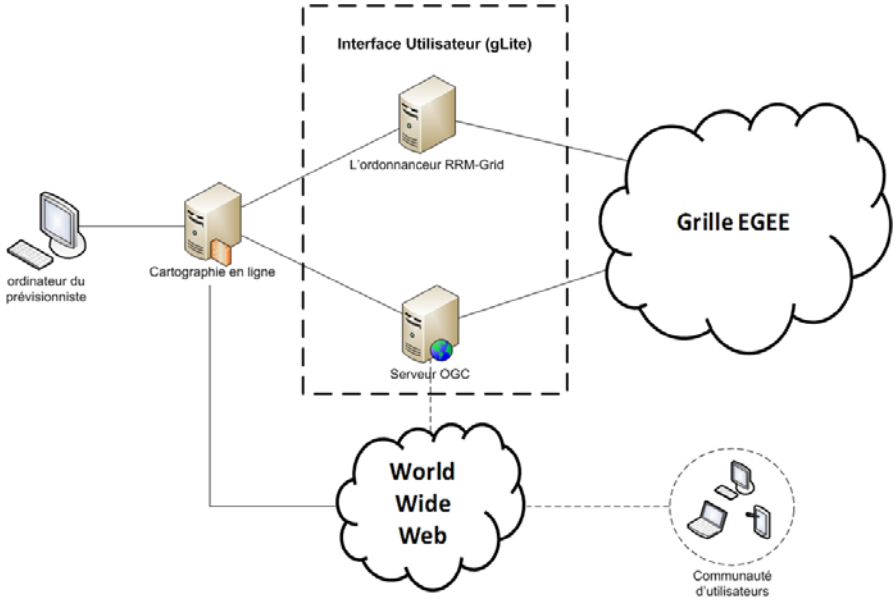


Figure 70 : Composants logiciels de l'application distribuée G-ALHTAÏR

La figure 70 synthétise cette note introductive, en schématisant les trois composants principaux du système spatial d'aide à la décision G-ALHTAÏR, c'est à l'interface de cartographie en ligne, le serveur OGC et l'ordonnanceur RRM-Grid. Ils doivent permettre au prévisionniste :

- de préparer à distance les scénarios de prévision hydrologique, en analysant la situation hydrométéorologique en cours et les différentes prévisions de pluie fournies par les partenaires délocalisés (Météo-France et RHEA), à partie de l'interface de cartographie en ligne,
- de soumettre ces scénarios à l'intergiciel de grille gLite grâce à 2 méthodes différentes :
    - un serveur OGC capable d'interpréter les requêtes du prévisionniste et de gérer la soumission, l'exécution et la publication des résultats de manière transparente et ubiquitaire
    - un ordonnanceur aux mêmes capacités que le serveur OGC enrichi de fonctionnalités permettant d'assurer une qualité de service suffisante en accord avec les besoins sécuritaires de la Sécurité Civile.

Enfin, l'interface de cartographie en ligne doit être accessible depuis Internet pour qu'une communauté d'utilisateurs institutionnels ait la possibilité de s'y connecter afin de participer aux opérations d'aide à la décision hydrologique.

VI - 2. G-ALHTAÏR : un outil garant de l'interopérabilité et de l'orchestration des traitements

Il s'agit de décrire le système spatial d'aide à la décision G-ALHTAÏR basé sur une application de cartographie en ligne collaborative et sur une couche logique composée de services web de l'OGC permettant l'accès aux ressources distantes de l'architecture de grille EGEE. Cette implémentation vient compléter la méthodologie de grillification présentée dans le chapitre 5 en argumentant en faveur d'une approche web garante de l'interopérabilité de ce nouveau système d'information. En effet, il s'agit de mettre en valeur l'intérêt d'utiliser des services web spécialisés favorisant une approche « métier » opérationnelle.

Comme pour la description du système existant, cette spécification s'appuie sur la méthodologie RM-ODP et le langage UML. Il s'agit d'approfondir le cas d'utilisation « support spatial d'aide à la décision » abordée dans la proposition méthodologique du chapitre 4 (cf. §.IV – 3.3.1. : Fig 58).

Ainsi, dans une première partie, les composants de l'interface graphique du système spatial d'aide à la décision implémentés, en adéquation avec les recommandations opérationnelles, sont décrits. Cette interface de cartographie en ligne doit assurer à terme l'intégration des données et des composants logiciels et matériels nécessaires au fonctionnement de cette plateforme. De plus, son fonctionnement sur le web offre et soutient une accessibilité continue et ubiquitaire aux fonctions de l'expertise hydrologique par l'ensemble des services opérationnels comme le SPC-GD et le SCHAPI. Dans un second temps, les points de vue « Traitement », « Ingénierie » et « Technologie » seront abordés de manière à présenter les orientations technologiques développées pour l'utilisation de services web de l'OGC afin d'accéder aux données et aux traitements grillifiés.

VI - 2.1. Point de vue « Entreprise » : l'interface de cartographie en ligne garante de l'opérationnalité du SPC-GD

Les chapitres précédents ont permis de souligner l'intérêt d'utiliser Internet et plus particulièrement les outils de cartographie en ligne pour assurer la gestion collaborative de l'expertise hydrométéorologique. Dans le cadre de cette étude, les deux acteurs institutionnels de cette mission sont le SPC-GD et le SCHAPI, cependant l'objectif de ce développement est de démontrer l'intérêt d'un tel outil pour supporter l'intégration rapide de nouveaux partenaires de gestion de crise, telles qu'ils ont été présentés dans le chapitre 3 (cf. §.III – 1.2.2.). Il est à noter que cet outil se démarque de la plateforme web Vigicrues[3] (LAPEYRE, 2006) principalement dédiée à la transmission d'information sur les situations de vigilance hydrologique aux populations et aux pouvoirs publics. Dans le cadre de cette réflexion, 5 cas d'utilisation ont été sélectionnés pour assister les prévisionnistes dans le développement de l'expertise hydrométéorologique des crues éclair (Fig. 71) :

- La gestion de la prévision météorologique,
- La gestion des cours d'eau,

[3] http://www.vigicrues.ecologie.gouv.fr/

- La gestion des stations mesures,
- La modélisation pour la prévision hydrologique,
- La gestion des commentaires.

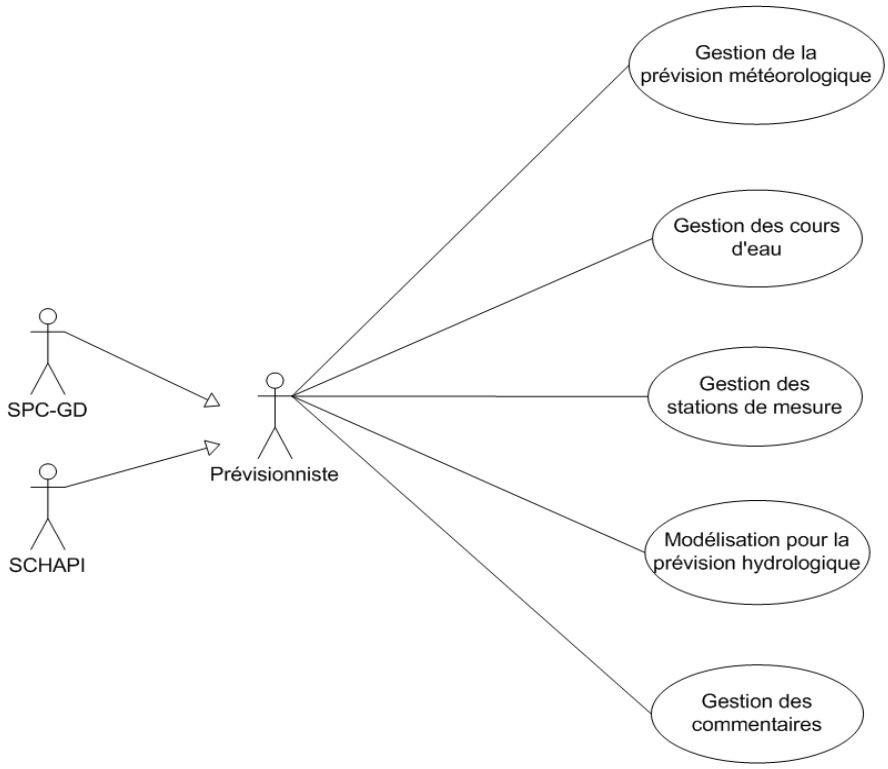


Figure 71 : Cas d'utilisation du système spatial d'aide à la décision

Ils permettent de visualiser et d'interroger les données collectées sur le territoire de compétence « Grand Delta » déjà décrites dans le chapitre 3 (cf. §.III – 1.2.3.1.). L'analyse de ces données et la connaissance du fonctionnement hydrologique des bassins versants de ce territoire permettent aux prévisionnistes d'établir des scénarios de prévision cohérents. De plus, un système de publication de messages permet de soutenir la dimension collaborative de ces opérations de crise. Enfin, cette interface cartographique offre la possibilité de partager l'ensemble des données stockées en local ou sur les éléments de stockage de la grille, entre le SPC-GD et le SCHAPI.

VI - 2.1.1. Gestion de l'imagerie radar

Cette activité consiste à intégrer les images radar fournies par RHEA© dans l'interface de cartographie en ligne, afin de les interpréter. Cette tâche apparaît essentielle pour le développement de l'expertise hydrologique. En effet, la connaissance des situations météorologiques « classiques », potentiellement propices au développement de précipitations orageuses intenses et de longues durées, est essentielle à l'anticipation de situations hydrologiques extrêmes. D'un point de vue fonctionnel, ce cas d'utilisation reprend les fonctionnalités offertes par l'application CALAMAR2 en permettant :

- le suivi des valeurs d'intensité de pluie issues de l'image radar,
- le paramétrage d'animations radar (plage et résolution temporelles),

- le calcul de cumuls pluviométriques sur une zone CALAMAR ou pour un bassin versant donné.

L'intérêt de l'animation radar est de pouvoir comprendre l'évolution des cellules de pluie afin d'anticiper leurs comportements à court terme. La connaissance des intensités maximales observées et des valeurs moyennes de précipitation peut permettre au prévisionniste de spécifier quantitativement les situations météorologiques à venir. Finalement, les prévisions Météo-France sur les zones AP (Avertissement Précipitation) (cf. §.III.1.2.3.3. : Fig. 40) fournissent aux prévisionnistes une information qualitative qui permet d'affiner l'évolution probable des champs de pluie, *a fortiori* de définir des scénarios de pluie spatialisés et de soutenir la modélisation de scénarios de prévision hydrologique cohérents et rigoureux.

VI - 2.1.2. Gestion des cours d'eau et des stations de mesure

Les linéaires des cours d'eau réglementaires et les stations de mesure sont aussi représentés sur cette interface. Comme pour les observations radar, les stations de mesure permettent de localiser les réactions hydrologiques en temps réel. En utilisant la codification des niveaux de vigilance et les côtes d'alerte établies pour la majorité des stations dans le règlement de surveillance (SPC-GD, 2006), le système d'aide à la décision adapte la sémiologie graphique de ces éléments en temps réel. L'automatisation de cette fonction permet de décharger du temps au prévisionniste pour se consacrer à la prévision proprement dite des débits et à l'évaluation du risque hydrologique (niveau de vigilance). Il est important de signaler que les niveaux de vigilance affichés sur cette interface ne sont pas diffusés au grand public, puisqu'ils nécessitent une critique par les experts avant leur diffusion. Couplés à la connaissance des champs de pluies, ces niveaux permettent d'établir une vision d'ensemble guidant le prévisionniste dans la sélection des bassins versants critiques qui requièrent une attention particulière quant à l'évolution de leurs débits.

VI - 2.1.3. Gestion des commentaires

Les retours d'expériences des deux évènements étudiés (cf. III.2.1.) ainsi que les recommandations des experts (cf. §.III – 2.2.) ont soulevé le besoin récurrent de collaboration entre les partenaires de la gestion de l'expertise. Les médias classiques tels que le téléphone ou le fax sont encore aujourd'hui largement utilisés en salle de crise, cependant le recours au web et en particulier les technologies du Web 2.0, peuvent enrichir cette capacité de collaboration et de dynamicité des interactions.

Ainsi, pour affermir l'opération de développement de l'expertise, cette interface de cartographie en ligne est enrichie d'un composant permettant au prévisionniste de poster des commentaires relatifs à son analyse de l'aléa. Cette coopération peut accompagner le prévisionniste du SPC-GD dans la définition et l'analyse des scénarios de prévision qu'il souhaite expérimenter. Ainsi, tout au long de l'évènement, elle favorise une caractérisation précise de l'aléa et le développement d'une expertise la plus univoque possible.

VI - 2.1.4. Modélisation pour la prévision hydrologique

Le dernier cas d'utilisation de ce système d'aide à la décision concerne le volet « prévision » des objectifs opérationnels du SPC-GD, plus largement des services institutionnels dédiés à l'expertise hydrologique. Elle représente l'élément central des développements méthodologiques de cette recherche et

est abordée dans le paragraphe suivant au travers des services web de l'OGC. Les trois premiers cas d'utilisation permettent d'analyser la situation hydrométéorologique en temps réel mais concourent surtout à fournir des éléments factuels susceptibles d'alimenter la définition et la mise en œuvre de scénarios de prévision. Il dépend de plusieurs paramètres que le prévisionniste peut faire varier :

- le bassin versant à modéliser,
- les conditions hydrologiques initiales,
- l'hydrogramme initial,
- la source de prévision de pluie,
- le délai de prévision,
- une pondération appliquée sur les intensités de pluie prévues.

L'application d'ALHTAÏR « temps réel » fournit, toutes les 5 minutes au prévisionniste, un débit modélisé à partir des pluies CALAMAR sur l'ensemble des bassins versants surveillés. Le prévisionniste en analysant les prévisions de pluie qu'il reçoit de Météo-France (zones AP) et de RHEA, et la situation hydrologique en temps réel (observations), a la possibilité de définir des scénarios de prévision de pluie à modéliser. Ces opérations de modélisation, par leur nombre et leur durée[4], nécessitent des ressources de calcul plus importantes que celles actuelles du SPC-GD. Elles sont donc traitées à distance sur l'architecture EGEE de manière à préserver une charge de calcul normale des ressources informatiques du SPC-GD.

L'envoi d'un nombre important de scénarios de prévision sur la grille représente un effort d'administration de la part du prévisionniste incompatible avec une situation de crise. En effet, la méthodologie classique de gestion d'un job de grille (cf. §.V - 1) doit être reproduite autant de fois que le nombre de scénarios de prévision requis. Dans ces conditions, le recours à des services web de l'OGC capables d'encapsuler les opérations de bas niveau pour le traitement et la gestion des données sur la grille, semble justifié. Il convient donc de décrire la méthodologie développée pour favoriser l'utilisation des services web pour supporter les traitements de grille correspondant aux exigences opérationnelles des prévisionnistes du SPC-GD.

VI - 2.2. Points de vue « Information et Traitement » : les services web de l'OGC

Le point de vue « Entreprise » de la méthode RM-ODP a permis de relever les cas d'utilisation qui composent le système spatial d'aide à la décision G-ALHTAÏR. L'analyse qui suit focalise sur le cas d'utilisation « Modélisation pour la prévision hydrologique » qui représente une augmentation de la charge de calcul susceptible d'affecter le bon fonctionnement du SPC-GD. Les autres cas d'utilisation semblent pouvoir être assurés par le système informatique existant puisqu'ils correspondent majoritairement à des opérations de visualisation et d'interrogation de données « temps réel ». Seule la gestion

[4] Un scénario de prévision modélise plusieurs matrices de pluie, 1h de prévision représentant 12 matrices (cf. §.III – 1.2.4.2)

de l'imagerie radar implique des opérations de géotraitements pour les calculs de cumuls et le paramétrage des animations. Cependant, la charge de calcul relativement faible qu'elle représente, permet d'envisager ces traitements sur le système local. Ainsi, en préambule de l'analyse UML dédiée à la formalisation de la méthodologie appliquée pour ce cas d'utilisation, il apparaît important de décrire les bases technologiques d'un service web de l'OGC.

VI - 2.2.1. Fondements technologiques des services web de l'OGC

VI - 2.2.1.1. Généralités sur les services de l'OGC

VI - 2.2.1.1.1 *Typologie générale*

L'OGC, dans son objectif de généralisation et de spécification de solutions interopérables, a un rôle de promotion auprès des développeurs et des utilisateurs du monde de l'information géographique (cf. §.I – 3.2.3.3.). Dans le cadre de cette méthodologie, un accent est mis sur l'utilisation de services ayant la capacité d'encapsuler des données et des géotraitements dans l'objectif de mettre en communication des systèmes d'information distribués.

Ces services web peuvent être divisés en 6 types/tiers différents, dont certains mettent en œuvre les spécifications d'implémentation relatives à l'information géographique et d'autres sont plus génériques aux technologies de l'information (PERCIVALL, 2002) :

- services d'interaction humaine ou d'application qui gèrent les interfaces utilisateurs et l'ensemble des éléments graphiques les composant,
- services de traitement qui nécessitent l'utilisation proprement dite de ressources de calcul pour interroger, modifier et construire de l'information géographique. Il est commun de les classer en 4 types : spatial, thématique, temporel, et relatif aux métadonnées,
- services de gestion de l'information qui permettent l'accès aux différents types de données, métadonnées ou plus généralement à des catalogues,
- services de gestion des systèmes, génériques aux technologies de l'information et prenant en charge la gestion des comptes utilisateurs, la sécurité, la comptabilité, ainsi que les opérations relatives à la gestion des systèmes d'instrumentation,
- services de *workflow* ou gestion de tâches, dédicacés à la construction de chaînes de traitements géographiques ou informatiques, et à leur gestion,
- services de communication relatifs aux mécanismes de messagerie, ou encore à la conversion de formats.

Les 3 types principaux des services de l'OGC sont les services d'applications, de traitements et de gestion de l'information (Fig. 72).

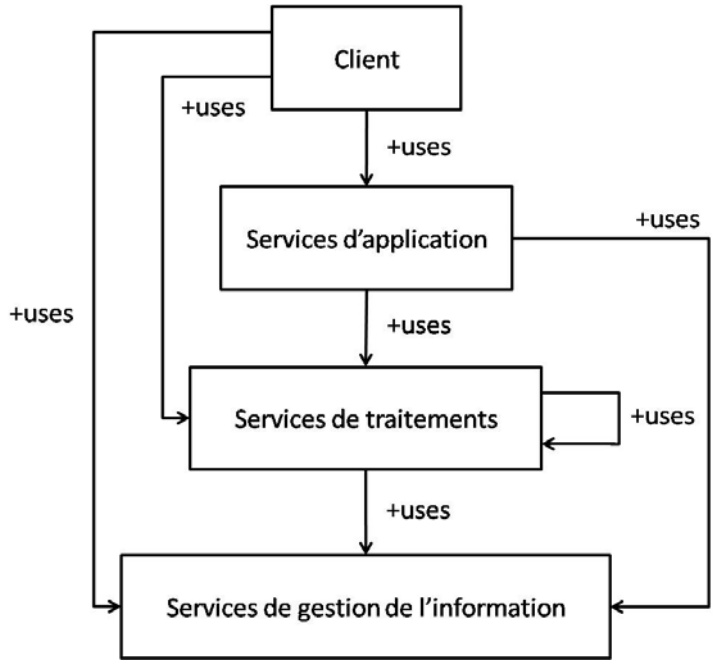


Figure 72 : Architecture des services web spécifiques de l'OGC
(d'après WHITESIDE, 2005)

Chaque service à la capacité de s'auto-décrire par des métadonnées, aussi bien interprétables par l'homme que par un autre service. Les opérations de transfert des données sont majoritairement basées sur un encodage XML (Extensible Markup Language). La standardisation de ces messages s'opère au travers des définitions de schémas XML[5] (XSD) définis par l'OGC. La transmission de ces données se fait au travers des protocoles standards du Web comme le HTTP[6] (Hypertext Transfer Protocol) ou le SOAP[7] (Simple Object Access Protocol) qui permettent d'interroger une ressource (service) identifiée grâce à une URL[8] (Uniform Resource Locator) (DANIEL, 2003 ; WHITESITE, 2007). Enfin, le format des données transférées doit répondre à un des types MIME[9] (Multipurpose Internet Mail Extensions). Le respect de ces recommandations garantit à l'utilisateur et au développeur l'accessibilité de leurs ressources par un composant (au sens de service) ou plus largement une entité physique situé à l'extérieur de son domaine administratif et juridique (WHITESITE, 2007).

VI - 2.2.1.1.2 *Exemples de services géographiques de l'OGC*

Les spécifications de l'OGC sont nombreuses et s'inscrivent dans une évolutivité constante de manière à répondre à un nombre de problématiques fonctionnelles toujours croissantes. Dans le cadre de cette étude, une attention particulière est portée sur les services de gestion de l'information géographique et de ses traitements particulièrement adaptés pour la récupération de données distantes et l'exécution de traitements à distance (originellement sur Internet). Il convient dorénavant d'approfondir les objectifs et le fonctionnement de certains services web de l'OGC. Cette analyse se focalise sur les services de l'OGC dédiés à la cartographie et aux géotraitements.

[5] http://schemas.opengis.net/
[6] http://www.w3.org/Protocols/
[7] http://www.w3.org/TR/soap/
[8] http://www.w3.org/Addressing/
[9] http://www.w3.org/Protocols/rfc1341/4_Content-Type.html

- le serveur de cartes web (*Web Map Service* - WMS) (DE LA BEAUJARDIERE, 2006) a pour rôle la préparation et/ou la mise à disposition sur un serveur de rendus cartographiques standardisés sous la forme d'images compressées (JPG, PNG, etc.) et dont les sources de données, les composant, peuvent être stockées sur plusieurs serveurs distants. De cette manière, des utilisateurs peuvent accéder à la volée, par exemple depuis leur SIG, à des fonds cartographiques par l'intermédiaire d'une requête au WMS.

- le serveur d'entités web *(Web Feature Service* - WFS) (VRETANOS, 2005) est spécifique à la gestion et la diffusion de données « vecteur ». Il permet la description et la modification[10] géométrique ce type de données grâce au langage de balisage géographique GML (*Geography Markup Language*), reconnu par l'OGC. Ce langage permet de spécifier les propriétés géométriques, géographiques et topologiques d'un objet géographique au point d'en être une copie. Le WFS peut aussi s'interfacer avec une base de données géographique (Fig. 73).

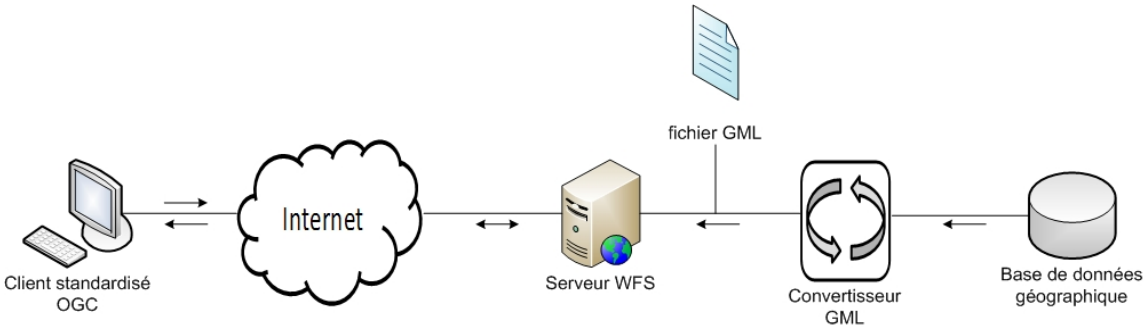


Figure 73 : Fonctionnement classique d'un WFS applicable aux autres services de l'OGC (d'après GEOCONNEXIONS, (2004)

- le serveur de couvertures web (*Web Coverage Service* - WCS) (WHITESITE et EVANS, 2006) permet d'accéder à des données de couvertures géographiques de la même manière qu'un WMS. La couverture représente un type de donnée capable de renseigner des valeurs, comprises dans un intervalle, en tout point du domaine spatio-temporel. Moins restrictif que le format raster, il est capable d'intégrer un plus grand nombre de type de données composites (images satellitales, observations radars, TIN, données ponctuelles, etc.). De plus, sa spécification inclus la possibilité d'implémenter des fonctionnalités basiques telles que le rééchantillonnage, le découpage et l'interpolation.

- le serveur de traitements web (*Web Processing Service* - WPS) (SCHUT, 2007) permet de servir et d'encapsuler des traitements ou des modèles utilisant l'information géographique. L'un des intérêts d'utilisation de ce service est sa capacité à accéder à des données directement fournies par l'utilisateur ou référencées sur le web (grâce en particuliers aux WCS et aux WFS) au point de le considérer comme un « service d'intergiciel pour données ». Enfin, il représente un service adapté à la construction de *workflows*, en gérant de manière autonome l'accès aux données, les géotraitements, la gestion des erreurs et le stockage des résultats.

[10] Il est question de mode transactionnel représenté sous la forme *-T, par exemple le WFS-T

L'ensemble de ces services dispose de plusieurs opérations[11] leur permettant de communiquer et d'interagir avec un composant ou un utilisateur distant. Ainsi, le WCS et le WPS (et dans la même mesure le WMS[12]) disposent de trois opérations, pour qu'un utilisateur ou un service :

- obtienne les métadonnées du service (*GetCapabilities*),
- obtienne la description des données contenues dans le WCS (*DescribeCoverage*) ou les traitements fournis par un WPS (*DescribeProcess*),
- extrait et récupère l'ensemble ou un sous-ensemble du WCS (GetCoverage) ou exécutent le(s) traitement(s) du WPS (*Execute*).

Les deux premières opérations sont très utiles pour connaître le contenu et les caractéristiques des éléments fournis par ces services web. Ainsi, un utilisateur ou un service web peut au préalable interroger un service pour connaître ses capacités en termes de données ou de traitements disponibles. La troisième opération permet d'exécuter un traitement ou d'obtenir une donnée fournie par le service web.

VI - 2.2.1.2. Invocation d'un service web de l'OGC

Un service ou un utilisateur invoque un service web de l'OGC grâce généralement à une requête HTTP avec les méthodes :

- GET[13] qui est structurée suivant le modèle « paire clé/valeur » ou KVP qui, à chaque paramètre, identifié dans la spécification standard du service web interrogé, associe une valeur,
- POST[14] qui est en charge de la transmission d'un fichier XML permettant, de la même manière que la méthode GET, de spécifier la requête de l'utilisateur au service web distant.

La manière la plus répandue d'invoquer l'exécution d'un service est d'envoyer un message HTTP GET à une URL donnée correspondant au service invoquée. L'URL est composée de l'adresse réseau du service et de paramètres permettant de spécifier la manière dont une des opérations du service est exécutée (Fig. 74). L'encodage KVP (Key Value Pair) est utilisée pour structurer le passage de paramètres, et correspond à des paires « nom / valeur » (WHITESITE, 2005). Certains paramètres sont obligatoires et d'autres optionnels, pour plus de détails, il faut se référer aux spécifications de chaque service web[15].

http://gacrue.ema.fr/geoserver/wms?bbox=3.6415473103523275,43.22836265563967,6.135699403285983,45.46761360168453&styles=&Format=application/openlayers&request=GetMap&version =1.1.1&layers=gacrue:troncons&width=800&height=673&srs=EPSG:4326

Figure 74 : Exemple de requête URL à un service WMS fourni par le serveur cartographique GeoServer

[11] Terme spécifique aux fonctionnalités offertes par chaque service
[12] Les 3 opérations d'un WMS sont le *GetCapabilities*, *GetFeatureInfo*, *GetMap*. Pour le WFS, il existe le *GetCapabilities*, *DescribeFeatureType*, *GetFEature* et pour la modification *InsertElement*, *DeleteElement*, *UpdateElement*, *LockFeature*
[13] http://www.w3.org/Protocols/rfc2616/rfc2616-sec9.html#sec9.3
[14] http://www.w3.org/Protocols/rfc2616/rfc2616-sec9.html#sec9.5
[15] http://www.opengeospatial.org/standards/

La requête URL de la figure 74 utilisée pour invoquer un service web distant est composée des éléments suivants :

- *http://gacrue.ema.fr/geoserver* est l'adresse URL du serveur contenant le service web invoqué
- *wms* est le type de service web invoqué
- *bbox=3.6415473103523275,43.22836265563967,6.135699403285983,45.46761360168453* représente l'extension de la carte c'est-à-dire longitude minimale, latitude minimale, longitude maximale et latitude maximale
- *styles=* est le style appliqué à la carte (nul dans ce cas)
- *Format=application/openlayers* est le format de sortie de la carte, de la couverture ou du vecteur
- *request=GetMap* est l'opération du service web invoquée par la requête
- *version=1.1.1* est la version de la spécification du web service
- *layers=gacrue:troncons* est la (les) couche(s) qui constituent la carte requise par le client
- *width=800&height=673* correspond aux dimensions en pixel de l'image de sortie
- *srs=EPSG:4326* est le code EPSG[16] (European Petroleum Survey Group) de la référence spatiale de la carte (ici WGS 1984)

La description de cet URL permet de montrer la grande flexibilité dans l'utilisation de ce type de services web. En effet, il est possible de demander une ressource (donnée ou traitement) en spécifiant ses paramètres de sortie ou d'exécution. Il est aussi possible d'invoquer un service web grâce à la méthode HTTP POST qui offre, en comparaison avec la méthode GET, la possibilité de transférer un fichier (généralement XML) pour spécifier les paramètres de la requête avec les mêmes attributs qu'une requête GET. Dans le cadre de la méthodologie présentée ci-après, les requêtes depuis l'interface de cartographie en ligne sont effectuées grâce aux deux méthodologies.

En retour, le service web publie une réponse sous forme XML à l'émetteur de la requête, basé sur les schémas standardisés de l'OGC, de manière à informer l'utilisateur ou un service de l'état d'avancement de la tâche ou d'une erreur relative à la requête envoyée par l'utilisateur ou d'une erreur d'exécution interne au service. Cette capacité de réponse permet l'enchaînement dynamique d'invocations de services (*workflows*).

[16] http://www.spatialreference.org/ref/epsg/

VI - 2.2.1.3. Positionnement technologique

Comme cela a été abordé dans le chapitre 4, l'utilisation des services web de l'OGC pour prendre en charge les traitements de l'information géographique sur la grille (Geospatial Grid) est en cours de développement et de formalisation (cf. §.IV – 2.2.2). La particularité de l'information géographique, voire géospatiale, dédiée à représenter des données multidimensionnelles et spatio-temporelles, rend cette adoption longue et implique la résolution de plusieurs challenges techno-ontologiques.

D'après BARANSKI (2008), l'implémentation de services web de l'OGC, pour accéder aux ressources de la grille, peut se faire de deux manières (Fig. 75) :

- à bas niveau, en installant les services de l'OGC sur un serveur extérieur à la grille, en tant qu'intermédiaire entre le portail utilisateur et les services de l'intergiciel (Fig. 75a),
- à haut niveau, en implémentant des services de grille ayant les spécifications de l'OGC (Fig. 75b).

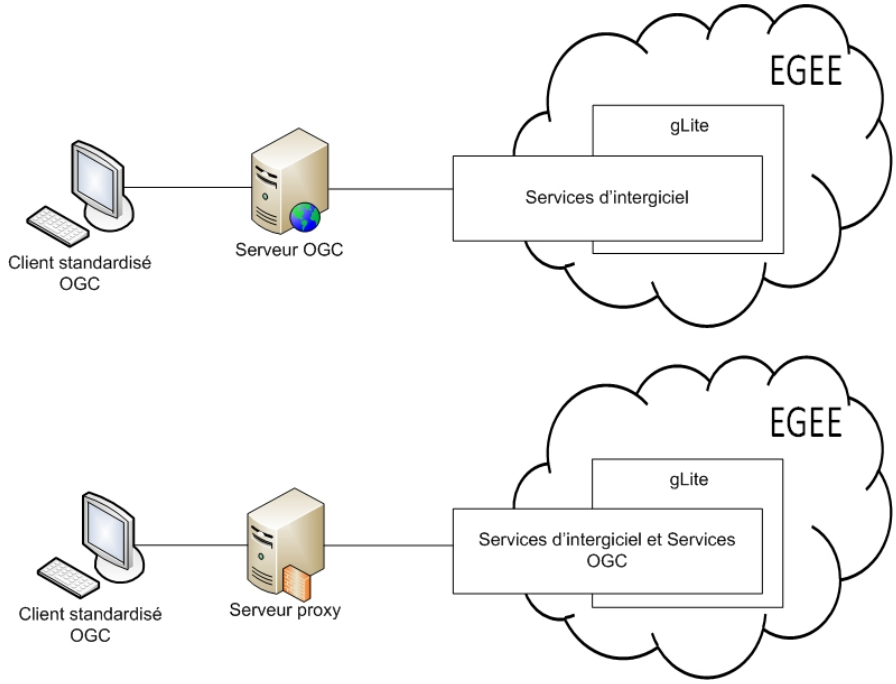


Figure 75 : Méthodes de grillification des services web de l'OGC. a) de bas niveau b) de haut niveau (d'après Baranski, 2008)

La seconde méthode apparaît comme étant la plus accomplie d'un point de vue technologique et aboutit à la prise en compte intrinsèque de l'information géospatiale par l'intergiciel de grille. De plus, l'intergiciel gLite ne permet pas l'installation de services web par la communauté des utilisateurs. Elle permettra à terme une véritable généralisation des traitements de l'information géographique et la construction dynamique de *workflows,* adaptés aux besoins spécifiques de chaque utilisateur, dans la même philosophie que les services d'intergiciel. Les défis scientifiques qu'elle implique, ne semblent pas permettre une opérationnalité de cette méthode avant plusieurs mois. Ainsi, la première méthode a été choisie dans le cadre de cette étude. Elle apparaît mieux adaptée aux problématiques de cette

recherche, à savoir son orientation « utilisateur » ne permettant pas une réelle appropriation des services de l'intergiciel gLite.

L'implémentation de ces services web pour permettre l'interaction entre l'intergiciel de grille gLite et les systèmes d'information distants a été développée dans le cadre du projet européen CYCLOPS, autour de deux cas d'utilisation :

- les feux de forêts en Italie (MAZZETTI et *al*., 2009)
- les crues éclair dans les Cévennes (THIERION et *al*., 2009a; 2009b)

La description technologique qui suit, s'appuie clairement sur les résultats de ces deux cas d'utilisation et les recherches menées dans le cadre du projet G-OWS (NATIVI et *al*., 2009). Ce projet vise à définir des méthodes et des outils pour la conception d'infrastructures de données spatiales basées sur la technologie grille.

Deux services web de l'OGC ont été sélectionnés pour permettre l'interaction entre le système d'information du SPC-GD et les ressources informatiques de la grille EGEE :

- le Web Coverage Service qui au travers de jobs de grille permet de (Fig. 76) :
    - récupérer une donnée qu'il référence et stocke sur un l'élément de calcul,
    - de traiter éventuellement cette donnée pour l'interpoler, la découper et/ou la rééchantillonner sur un nœud de calcul,
    - de la replacer sur un élément de stockage géographiquement proche de l'élément de calcul sur lequel le traitement principal est effectué. Le nouveau LFN (Logical File Name) de la donnée est enregistré pour sa réutilisation ultérieure.

  Dans le cas, de plusieurs requêtes au WCS, plusieurs jobs sont exécutés en parallèle et simultanément, ce qui permet une montée en charge de cette gestion des données.

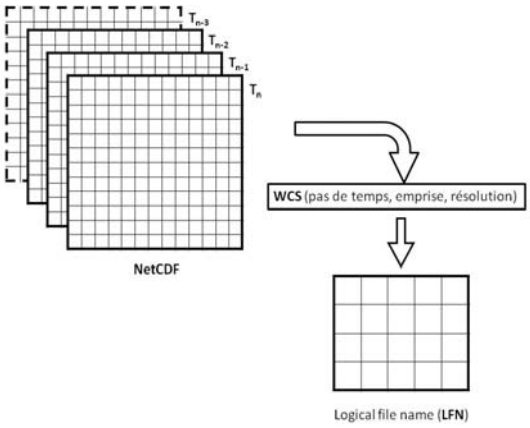


Figure 76 : Comportement classique du WCS dans le cas du traitement d'un fichier NetCDF

- le Web Processing Service, qui se substitue au rôle de l'utilisateur dans l'accès à la grille, a la capacité de :
    - générer des jobs de grille en JDL,
    - générer des requêtes vers un WCS pour récupérer les données nécessaires au traitement dont il est en charge,
    - soumettre ces jobs au courtier de ressources (WMS) qui exécutent des algorithmes qu'il référence sur un nœud de calcul,
    - suivre l'avancement du job et d'informer à la demande l'émetteur de la requête grâce à une réponse standard (XML).

  Le WPS prend ainsi en charge une chaîne de traitements de manière automatique et autonome, tout en ayant la capacité d'informer l'émetteur de la requête.

La compréhension précise du fonctionnement global des services web de l'OGC et celui plus spécifique pour l'accès aux services de l'intergiciel de grille, permet maintenant d'analyser l'utilisation et le fonctionnement de ces services pour la gestion de scénarios de prévision hydrologique sur la grille EGEE.

VI - 2.2.2. Les services web de l'OGC : le cas d'utilisation « modélisation pour la prévision hydrologique »

VI - 2.2.2.1. Point de vue « Information »

Dans la suite du développement, une convention de couleurs a été utilisée pour mieux comprendre l'organisation matérielle et géographique de G-ALHTAÏR, au travers des différents diagrammes UML présentés :

- le vert pour le système d'information du SPC-GD,
- le rouge pour le serveur OGC,
- le bleu pour les ressources de la grille EGEE.

Dans un premier temps, il est important de comprendre l'enchaînement des opérations pour développer et produire un scénario de prévision, défini précédemment (cf. §.VI - 2.1.4), et basé sur les paramètres suivants :

- l'horizon de prévision,
- le facteur multiplicateur des intensités de pluie prévues,

et les fichiers suivants :

- le bassin versant

- la pluie prévue,
- l'hydrogramme temps réel,
- les conditions hydrologiques initiales.

D'après la figure 77, le prévisionniste au moment où il le souhaite soumet une requête de scénarios de prévision depuis l'interface de cartographie en ligne. Les capacités de calcul de la grille permettent de traiter indépendamment et de manière simultanée ces différents scénarios. Les paragraphes qui suivent s'attachent à décrire les traitements effectués par le serveur OGC pour la prise en charge d'un scénario de prévision, reproductibles pour chaque scénario requis par le prévisionniste.

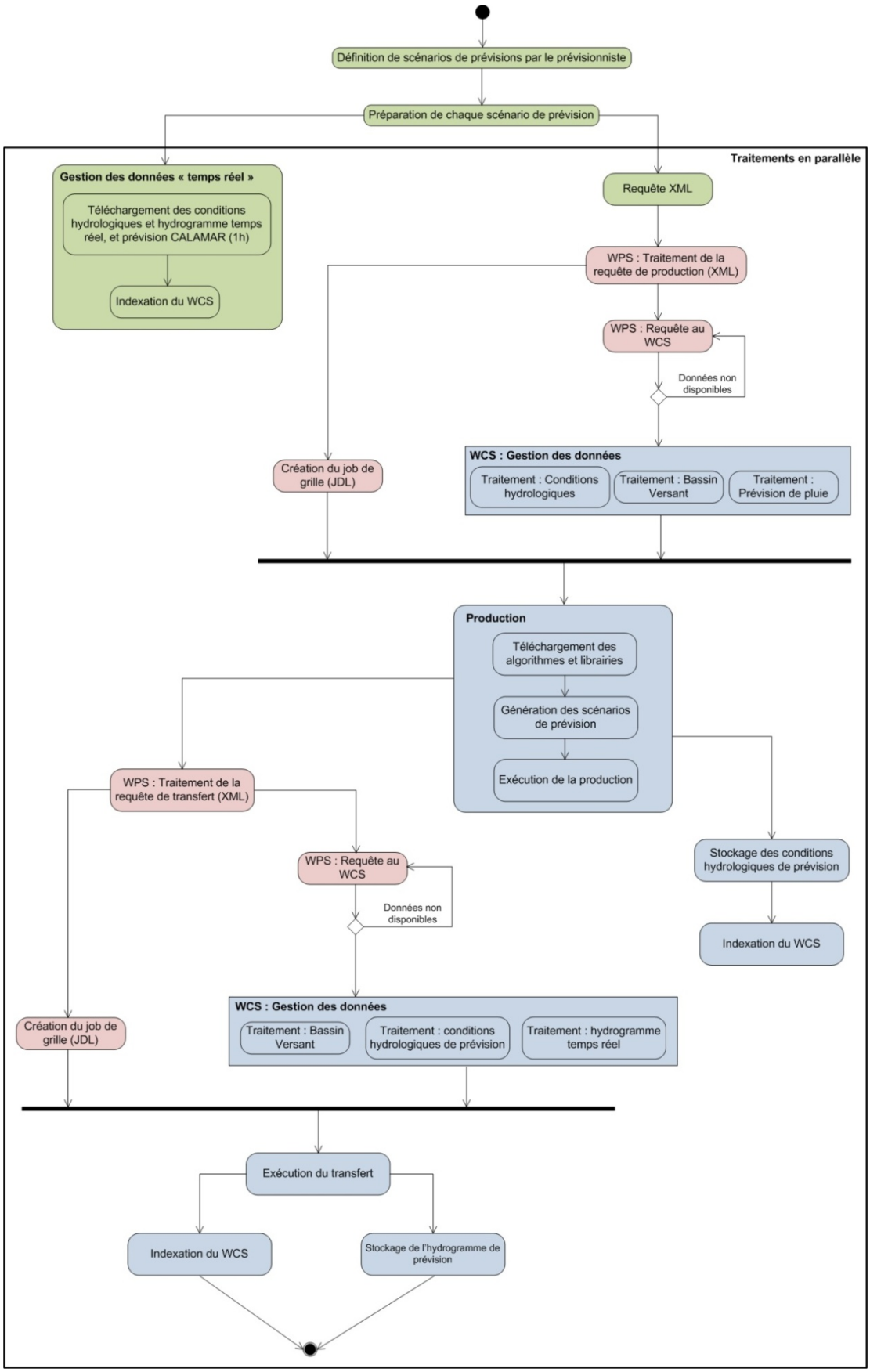


Figure 77 : Diagramme d'activité UML de G-ALHTAÏR, le vert représente le système d'information du SPC-GD, le rouge, le serveur OGC, et le bleu, les ressources de la grille EGEE

Les données « temps réel », telles que l'hydrogramme et les conditions hydrologiques, simulées à chaque pas de temps sur le système d'information du SPC-GD au moyen de l'application ALHTAÏR, correspondent aux données d'entrée potentielles des simulations de prévision effectuées sur la grille. Au lancement de la requête, ces données et la dernière prévision de pluie CALAMAR reçue sont donc automatiquement copiées sur un élément de calcul de la grille et référencées sur le serveur WCS.

En parallèle, le WPS reçoit une requête de production codée en XML qui lui fournit :

- le module de production à utiliser, en l'occurrence le module de production d'ALHTAÏR déjà grillifié (chapitre 5),
- les données d'entrée du module avec leur extension, leur projection et leur résolution géographique dans le cas du bassin versant, et en plus le pas de temps des données évènementielles (pluie, conditions hydrologiques et hydrogramme),
- les paramètres globaux de la modélisation que sont la pondération à effectuer sur la pluie de prévision, l'horizon de prévision, le pas de temps à modéliser et la résolution temporelle,
- la donnée de sortie correspondant aux conditions hydrologiques de prévision (contenant les matrices de ruissellement sur lesquelles le transfert doit être effectué), et son format.

L'interprétation de cette requête permet au WPS de préparer l'exécution du module de production en contactant le WCS charger de préparer les données d'entrée correspondant au bassin versant, aux conditions hydrologiques et à la prévision de pluie déjà stockés sur les éléments de stockage. Cette préparation ne peut être réalisée que dans le cas où les données d'entrée sont déjà téléchargées par le système d'information du SPC-GD sur un élément de stockage (Gestion des données « temps réel »).

En parallèle de cette préparation des données d'entrée, le WPS prépare et soumet un job de grille au courtier de ressource. Sur l'élément de calcul, le job de grille se charge successivement :

- de récupérer les algorithmes nécessaires à la production,
- de préparer le scénario de pluie grâce au délai de prévision et la pondération à appliquer sur la donnée prévision de pluie,
- d'exécuter la production,
- de stocker les données résultantes sur un élément de stockage et de les référencer sur le WCS (conditions hydrologiques de prévision).

A la fin de la production, une nouvelle instance du WPS réédite la même séquence pour exécuter le module de transfert grâce à une requête XML lui fournissant :

- le module de transfert à utiliser, en l'occurrence le module de transfert d'ALHTAÏR déjà grillifié (chapitre 5),

- les données d'entrée avec leurs paramètres spatio-temporels comme le bassin versant, l'hydrogramme « temps réel » et les conditions hydrologiques de prévision résultantes de la phase de production,
- la donnée de sortie correspondant à l'hydrogramme de prévision avec son format.

Finalement, le WPS gère, de la même manière que pour la production, un job de grille en charge de la récupération des algorithmes et des librairies spécifiques, de l'exécution du module de transfert, et du référencement de l'hydrogramme de prévision sur le WCS après un stockage préalable sur un élément de stockage.

VI - 2.2.2.2. Point de vue « Traitement »

La synchronisation et la communication entre l'interface de cartographie en ligne, le serveur OGC et l'architecture EGEE a nécessité le développement de composants, basés sur des patrons de conception[17], qui permettent la préparation et l'échange de données de manière automatique :

- l'échangeur : il prend en charge les prévisions de pluie fournies par RHEA©, les encode dans le format NetCDF et les place sur un élément de stockage de la grille (LCG utils). Il permet de plus le transfert des données issues de la modélisation temps réel d'ALHTAÏR (conditions hydrologiques et hydrogramme) utiles à la modélisation des différents scénarios de prévision. Enfin, il est en charge de la mise à jour de l'index de ces données sur le WCS (Fig. 78).

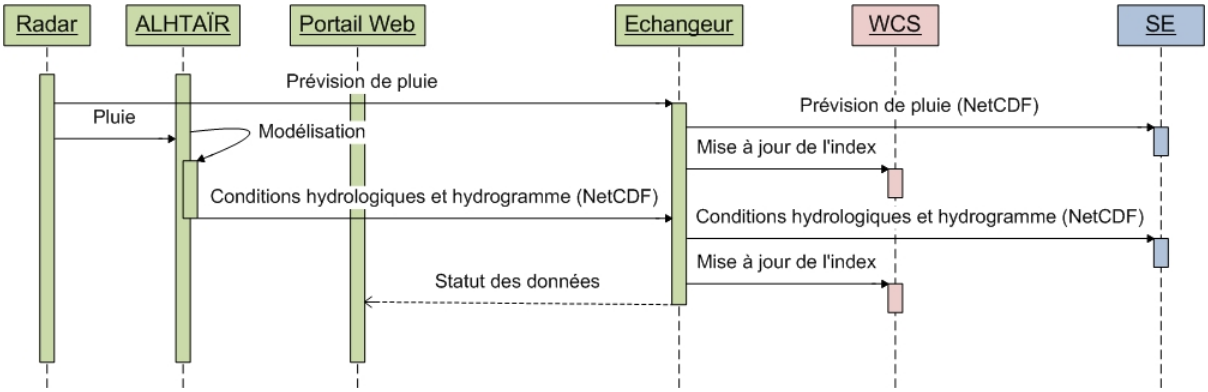


Figure 78 : Diagramme de séquence UML du rôle de l'échangeur, le vert représente le système d'information du SPC-GD, le rouge, le serveur OGC, et le bleu, les ressources de la grille EGEE

- le contrôleur : il facilite l'interaction entre le système local et le WPS. En effet, il formate chaque scénario de prévision requis par le prévisionniste (à partir de l'interface de cartographie en ligne) en fichier XML. Il se charge ensuite de l'envoi de ces fichiers au WPS, qu'il contacte régulièrement pour garder le prévisionniste informé de l'avancée de l'exécution des prévisions soumises.

La séquence complète du traitement d'un scénario peut être appréhendée grâce la figure 79.

[17] Un patron de conception décrit une solution standard pour répondre à des problèmes d'architecture et de conception des logiciels (Wikipedia)

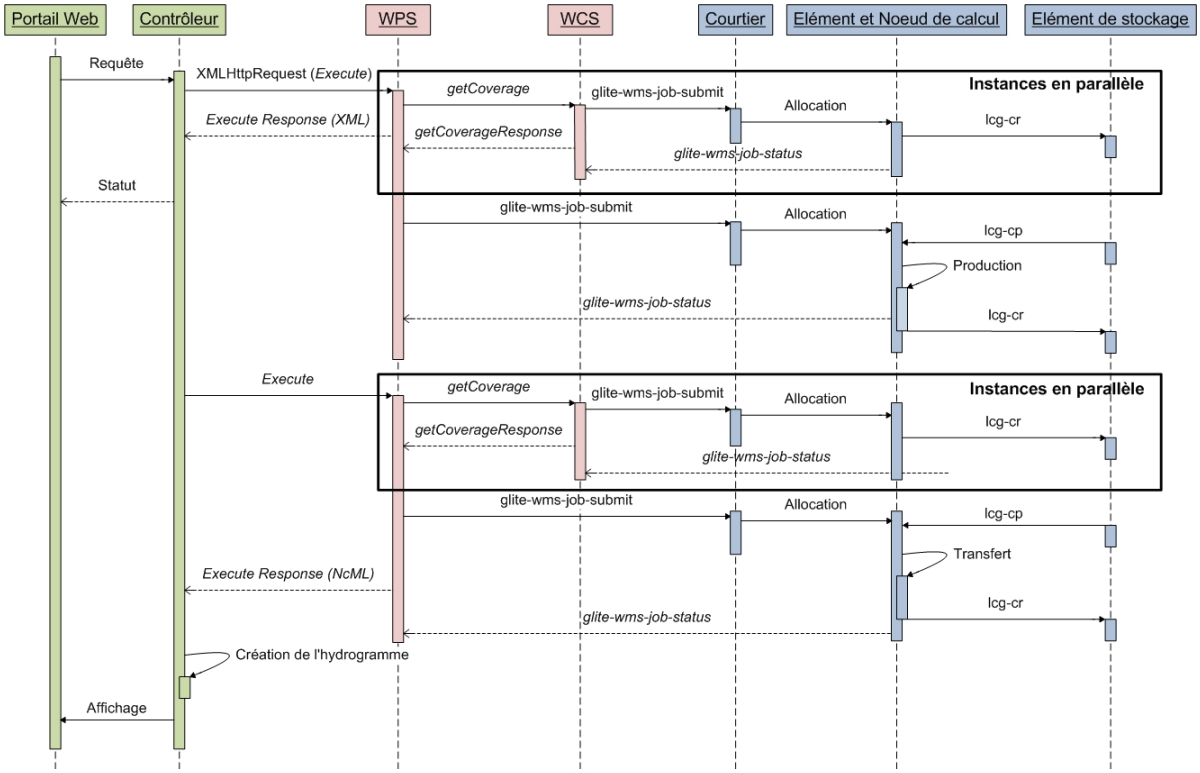


Figure 79 : Diagramme de séquence UML de GALHTAÏR, le vert représente le système d'information du SPC-GD, le rouge, le serveur OGC, et le bleu, les ressources de la grille EGEE

Après validation du scénario de prévision par le prévisionniste, le contrôleur envoie un message asynchrone codé en XML à l'opération *Execute* du WPS. Cet envoi est assuré par la fonction *XMLHttpRequest* du langage AJAX[18]. Le WPS réceptionne et interprète la requête XML de production. Ensuite, il prépare autant de requêtes XML que de données requises pour la modélisation (*Instances en parallèle*) et invoque le WCS pour la préparation de ces données grâce à son opération *GetCoverage*. Le WCS soumet un job de grille au courtier de ressource pour chaque donnée requise par le WPS (*glite-wms-jobs-submit*). Grâce à ces jobs, il prépare les données NetCDF à modéliser, mises à jour par l'échangeur sur un élément de stockage, en extrayant les conditions hydrologiques et la prévision de pluies correspondantes au pas de temps[19] indiqué dans la requête XML que lui a transmis le WPS, et le bassin versant. Il place ces données sur un élément de calcul (*lcg-cr*) et mémorise leurs LFN respectifs pour leur prise en charge par le WPS (Fig. 76).

Le WCS informe le WPS de l'avancée dans la préparation des données (*getCoverageResponse*), et une fois stockées, le WPS lance l'exécution du module de production grâce à un job de grille qu'il soumet au courtier de ressources pour exécution sur un nœud de calcul (*glite-wms-jobs-submit)*. Le job de grille accède aux données référencées sur le WCS (*lcg-cp*) et exécute les algorithmes relatifs à la production (scénario de pluie et module pseudo-hortonien). Lorsque la production est terminée, la matrice de ruissellement résultante est stockée sur un élément de stockage (*lcg-cr*) et le LFN de ce fichier est transmis au contrôleur pour permettre sa réutilisation dans le module de transfert exécuté en suivant

[18] Le langage AJAX, comme son acronyme l'indique au travers du terme « asynchronous », s'inscrit complètement dans la mouvance Web 2.0 et permet d'envoyer des requêtes asynchrones à un serveur distant

[19] Dans le cas de cette étude, ce pas de temps correspond à celui de la dernière simulation « temps réel » sur le système local au moment du lancement du scénario de prévision par le prévisionniste

(*Execute Response*). Une nouvelle requête XML est envoyée au WPS pour exécuter le module de transfert. La même séquence de traitements est effectuée par le WCS et le WPS pour respectivement récupérer, et modéliser le bassin versant, la matrice de ruissellement (conditions hydrologiques de prévision) et l'hydrogramme « temps réel » correspondant. Au total, la chaîne de traitements complète nécessite 8 jobs de grille pour respectivement traiter les données d'entrée (2 x 3 jobs) et exécuter les modules de production et de transfert (2 x 1 job).

Tout au long du cycle de vie du WPS, le contrôleur est informé de son état d'avancement grâce à la réponse *Execute Response*, qui est interfacée avec la réponse *getCoverageResponse* du WCS et la réponse de la fonction *glite-wms-job-status* relative au statut du job de grille. La dernière réponse *Execute Response* du WPS fournit au contrôleur l'hydrogramme de prévision codé grâce au format NcML (cf. §.V – 2.2.2.) Grâce à sa structure proche du XML, il peut être directement intégré dans cette réponse. Le contrôleur peut alors formater et afficher l'hydrogramme de prévision requis par le prévisionniste.

L'utilisation du WPS permet donc une gestion transparente de la modélisation d'un scénario de prévision sur la grille. Elle implique cependant le développement de composants logiciels, tels que l'échangeur et le contrôleur pour assurer la synchronisation et la communication entre le système d'information du SPC-GD et ce service. En effet, le serveur WPS est capable de gérer en simultanée de nombreuses instances de prévision, cependant le contrôleur doit se charger de préparer au préalable chaque scénario afin de les soumettre individuellement au WPS. Finalement, l'interface de cartographie en ligne couplée au serveur OGC assure une opérationnalité des fonctions de bas niveau de la grille susceptibles d'enrichir la capacité de calcul initiale du SPC-GD. Ainsi, dans une dernière partie, il s'agit de décrire les solutions logicielles retenues pour implémenter l'interface de cartographie en ligne et l'agencement de l'ensemble des composants logiciels inclus dans G-ALHTAÏR.

VI - 2.3. Point de vue « Ingénierie » et « Technologie » de G-ALHTAÏR

Ce paragraphe s'attache à décrire les solutions logicielles et matérielles adoptées pour soutenir le fonctionnement de G-ALHTAÏR, en particuliers celles dédiées au fonctionnement de l'interface de cartographie en ligne sur le système local du SPC-GD. Dans un second temps, l'architecture globale de G-ALHTAÏR est présentée en insistant sur l'organisation des composants décrits dans les paragraphes précédents.

VI - 2.3.1. L'interface de cartographie en ligne

Cette analyse est scindée entre les logiciels structurant le serveur web à composante géographique et les applications clientes enrichissant le fonctionnement global de l'interface de cartographie en ligne.

VI - 2.3.1.1. Le serveur web à composante géographique

D'après la figure 80, l'interface de cartographie en ligne permettant l'intégration et l'orchestration des fonctions « métier » du prévisionniste est permise grâce à 3 serveurs et un système de gestion de base de données (SGBD) (AYRAL et *al*., 2010 ; BURCK et *al*., 2008b).

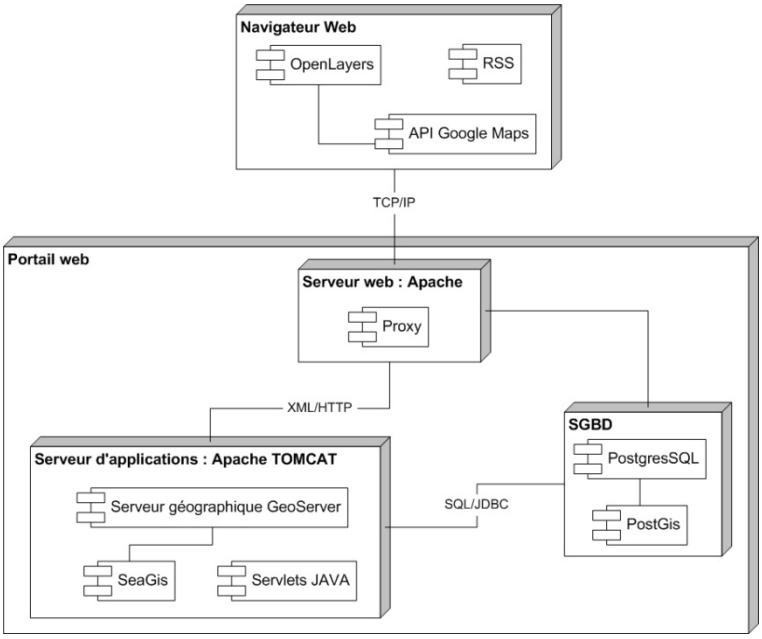


Figure 80 : Diagramme de déploiement UML de l'architecture de l'interface de cartographie en ligne

Il est donc possible de distinguer :

- le serveur Apache[20] supporte la prise en charge des scripts PHP, HMTL et JavaScript, dédiés à la conception du squelette du site web affiché chez le client et il interconnecte l'ensemble des composants logiciels de la plateforme.

- le serveur cartographique Geoserver[21], retenu pour sa compatibilité avec les normes de l'OGC (BURCK et *al*., 2008a), centralise l'information géographique afin de permettre sa diffusion et sa représentation graphique. Il est connecté à la base de données PostgreSQL[22]/PostGis[23] pour permettre la prise en charge des objets géographiques contextuels, tels que les bassins versants, les cours d'eau et les stations de mesure.

- le serveur Apache Tomcat[24] permettant l'implémentation :

    - de servlets Java, conçus pour étendre les fonctionnalités d'un serveur. Il a été utilisé pour servir des fonctionnalités de gestion des données de pluies (lecture, animations radar, calculs matriciels pour les cumuls de pluie, etc.). De plus, un de ces servlets permet la synchronisation temporelle des différents composants de l'application de cartographie en ligne.

    - l'application SeaGis[25], spécialisé dans l'affichage et la diffusion de données codées sous format multidimensionnel. Dans le cadre de cette étude, elle permet la prise en charge et l'affichage des données de pluie enregistrées dans des fichiers NetCDF.

[20] http://www.apache.org/
[21] http://geoserver.org/
[22] http://www.postgresql.org/
[23] http://postgis.refractions.net/
[24] http://tomcat.apache.org/

- Le SGBD PostreSQL associé à son extension PostGIS qui permet la centralisation des données géographiques diffusées et produites vers et depuis la plateforme. Il permet la gestion des commentaires ainsi que les opérations d'administration du site web.

Finalement, le serveur cartographique GeoServer supporte les services web cartographiques de l'OGC, tels que le *Web Feature Service* (WFS), le *Web Coverage Service* (WCS) et le *Web Map Service* (WMS). Ainsi, cette capacité revêt un intérêt tout particulier dans le contexte de gestion de crise. En effet, les rendus cartographiques et les données produits au travers de l'interface cartographique dédiée à l'expertise hydrologique, peuvent être mis à disposition à l'ensemble des partenaires par l'intermédiaire des services WMS, WFS et WCS. À ce niveau, il est important de nuancer l'utilisation des services web au travers du serveur cartographique GeoServer et ceux utilisés dans le cadre de l'accès aux ressources de grille. En effet, cette capacité fait partie intégrante de ce serveur alors que dans le cadre de l'interaction avec la grille, elle correspond à un défi technologique encore en cours.

VI - 2.3.1.2. L'application cartographique cliente

L'affichage des données géographiques, mises à disposition par le serveur cartographique, sur l'interface se fait au travers de la librairie orientée objet OpenLayers[26] basée sur JavaScript et AJAX. Elle offre un canevas graphique simplifiant l'implémentation de l'interface graphique et la sémiologie graphique des données SIG (Fig. 81).

Figure 81 : Interface graphique de l'application G-ALHTAÏR

Le fond cartographique de cette interface est fourni par Google© et son API Google Maps[27]. L'un des intérêts d'une API cartographique (*Application Programming Interface* ou Interface de programmation)

[25] http://www.geotoolkit.org/
[26] http://openlayers.org/
[27] http://code.google.com/apis/maps/

est d'accéder à des applications ou des fonctions web mises disposition par des fournisseurs au travers d'une autre application. Ainsi, dans le cadre de cette interface, au moment où le prévisionniste se connecte à l'interface de cartographie en ligne, une requête est envoyée à l'API Google Maps qui fournit le fond cartographique[28] avec l'extension et le niveau de zoom requis. Cette approche permet d'autre part de limiter la surcharge du serveur cartographique prioritairement dédié aux « tâches hydrologiques ».

L'affichage des hydrogrammes à chaque station de mesure se fait grâce à une interface de visualisation développée en Flash®, qui permet de visualiser plusieurs hydrogrammes simultanément. Etant donné l'incertitude importante des valeurs de débits, cette fonctionnalité apparaît primordiale pour établir une inter-comparaison entre le débit observé, et les différents débits modélisés (temps réel et scénarios de prévision). Enfin, la publication automatique des mesures collectées vers l'ensemble des partenaires s'appuie sur leur syndication grâce aux flux GeoRSS. Quant à la gestion des commentaires des utilisateurs, un service d'abonnement à un flux RSS permet à chaque partenaire d'être constamment informé sur les stratégies et les décisions des autres partenaires.

VI - 2.3.2. Synthèse : l'architecture globale de G-ALHTAÏR

Les moyens d'interfacer le système d'information existant du SPC-GD avec la grille EGEE ont été décrits précédemment. Il convient de synthétiser cette méthodologie au travers de l'analyse du diagramme de déploiement UML de la figure 82.

Le prévisionniste, ou toutes personnes autorisées accèdent à G-ALHTAÏR grâce à leur navigateur web. Au démarrage du système, les applications clientes se chargent et affichent l'ensemble des composants graphiques qui constituent l'interface. D'un point de vue hydrologique, les éléments géographiques spécifiques (cours d'eau, bassin versant et stations de mesure) sont présentés avec la sémiologie graphique correspondante à leur situation hydrologique (niveaux de vigilance).

Le fonctionnement web de G-ALHTAÏR est assuré par le serveur web et le serveur d'applications qui ont la capacité de répondre aux requêtes exécutées par le prévisionniste depuis l'interface graphique. Au niveau du serveur d'application, l'échangeur assure le lien entre la version « temps réel » d'ALHTAÏR et le scénario de prévision, en actualisant régulièrement sur les éléments de stockage de la grille EGEE les données issues de cette modélisation « temps réel » et en les indexant sur le serveur WCS. Quant au contrôleur, basé sur le fonctionnement dynamique et asynchrone du langage AJAX, il traduit, formate et transporte les messages et les réponses transitant entre le WPS et l'interface de cartographie en ligne. Ainsi, l'échangeur et le contrôleur garantissent l'interaction efficace entre le portail web (indirectement l'utilisateur) et les services web de l'OGC, rendant leur rôle essentiel dans la synchronisation du système local avec les composants distribués, que sont ces services et par extension les services de l'intergiciel gLite.

[28] Il est également envisageable d'utiliser le fond cartographie de l'IGC au travers de l'API Géoportail (https://api.ign.fr/geoportail/index.do)

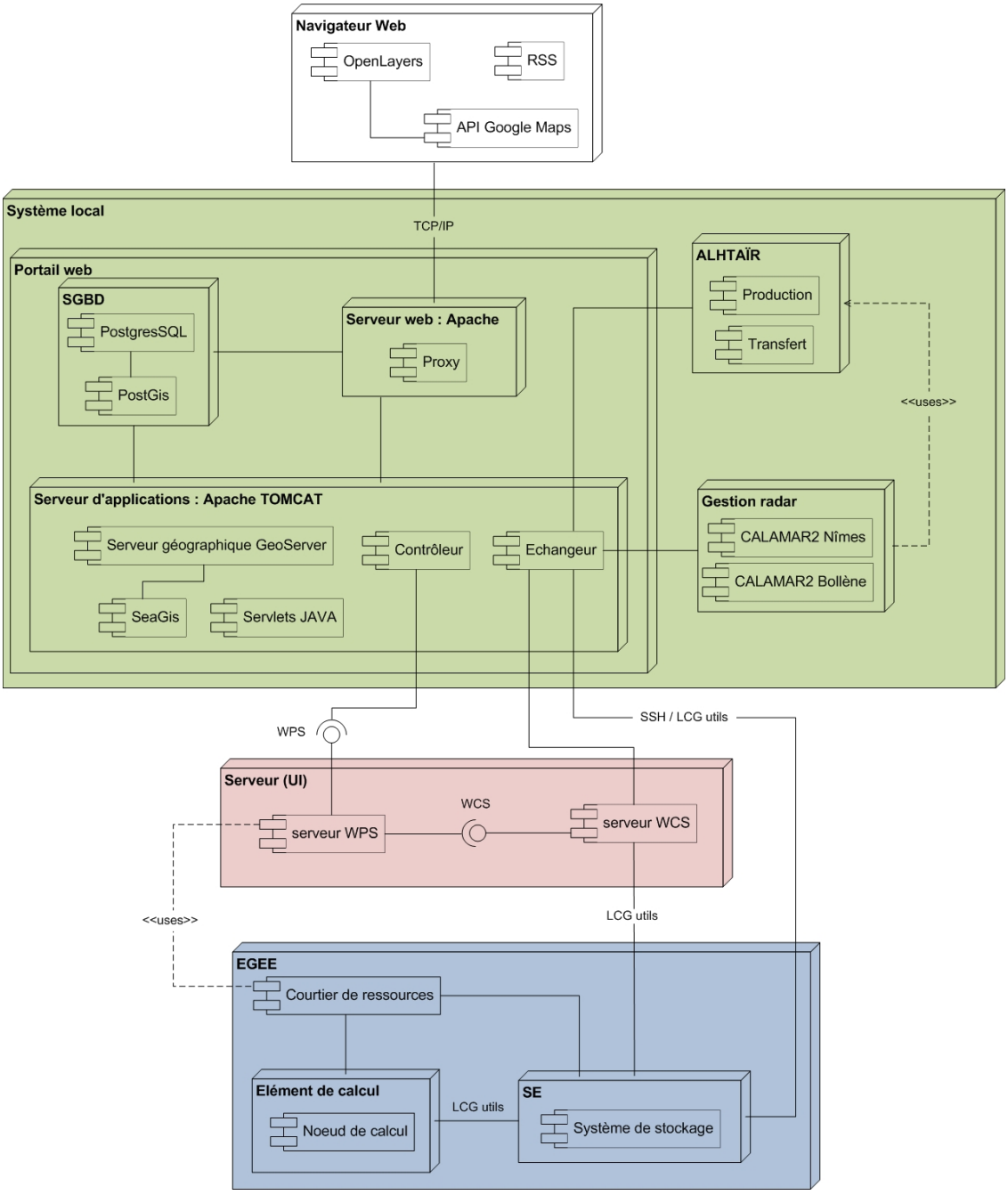


Figure 82 : Diagramme de déploiement UML de GALHTAÏR, le vert représente le système d'information du SPC-GD, le rouge, le serveur OGC, et le bleu, les ressources de la grille EGEE

Les serveurs WPS et WCS ont implémentés sur une interface utilisateur installée dans les locaux du laboratoire Earth and Space Science Informatics - Laboratory (ESSI-Lab) de l'IMAA (Institute of Methodologies for Environmental Analysis) en Italie, partenaire du projet européen CYCLOPS (cf. §.IV – 2.2.3.1.2.). Depuis cette interface utilisateur, ils ont donc la capacité de recevoir les requêtes HTTP envoyées par le contrôleur et les informations de mises à jour des données effectuées par l'échangeur. L'élément central de ce fonctionnement déporté est le WPS. À la réception d'une requête du contrôleur, de part ses capacités d'orchestration de traitements, il gère automatiquement la préparation et l'exécution du scénario de prévision sur les ressources de calcul de la grille et la transmission directe de l'hydrogramme prévu. De son côté, le WCS augmente la flexibilité de fonctionnement du WPS en

référençant[29] en temps réel et univoquement les données (et leurs métadonnées) fournies par l'échangeur, et en les prétraitant suivant les paramètres spatio-temporels requis par le scénario de prévision.

Finalement, les services de l'intergiciel et les ressources de la grille interagissent de manière autonome pour mener à bien l'exécution des opérations relatives à la modélisation hydrologique et informer l'émetteur de la demande de traitements (dans ce cas le WPS) sur l'état d'avancement de ces opérations informatiques.

En conclusion, l'imbrication et la mise en communication de l'ensemble de ces composants logiciels permet une gestion transparente de scénarios de prévision sur des ressources distantes, telles que celles fournies par l'architecture EGEE. La dynamicité et l'interactivité de l'interface de cartographie en ligne, permises par les technologies du Web 2.0, offrent au prévisionniste un moyen efficace d'analyser la situation hydrométéorologique « temps réel », de construire et traiter à distance des prévisions pertinentes dans un contexte collaboratif fort.

D'autre part, l'intérêt des services web, comme ceux de l'OGC, réside dans leur capacité de s'interfacer avec des systèmes d'information distants ne pouvant pas subir une refonte de leur structure et de leur fonctionnement. De plus, leur capacité à s'auto-décrire pour informer un système distant sur les ressources (traitements ou données) dont il dispose, offre à cette méthode une garantie dans le partage rapide et efficace de l'ensemble des éléments qui constitue l'aide à la décision hydrologique. Finalement, le cadre standardisé dans lequel ils s'inscrivent assure une reproductibilité et un élargissement des fonctions « métier » nécessaires et envisageables dans le cadre du perfectionnement de la modélisation hydrologique des crues à cinétique rapide. Il en résulte une forte capacité d'interopérabilité essentielle au fonctionnement efficace de la gestion de crise, permettant dans le cadre de l'expertise hydrologique de mieux intégrer les prévisionnistes du SPC-GD et du SCHAPI autour d'un système d'aide à la décision commun. Cette la méthodologie est donc susceptible de soutenir des capacités de calcul à la demande pour éviter la surcharge informatique du système d'information du SPC-GD. Dans ce contexte, plusieurs scénarios de prévision peuvent être modélisés sans altérer le fonctionnement du SPC-GD. Cette capacité technologique répond *a priori* à l'usage qu'un service opérationnel de gestion de crise est susceptible de développer pour répondre à des besoins ponctuels et imprévus.

En contrepartie, l'implémentation de ces services web implique une disponibilité importante des ressources de calcul, puisqu'un seul scénario de prévision nécessite pas moins de 8 jobs de grille. Dans une situation opérationnelle, le prévisionniste peut requérir plusieurs scénarios de prévision pour chaque bassin versant du territoire « Grand Delta », ce qui peut engendrer une forte surcharge dans la demande d'accès aux ressources de la grille (par ex. 4 scénarios de prévision pour 30 bassins versants nécessitent l'utilisation de 960 jobs). Ainsi, il semble important de s'intéresser à la dimension proprement calculatoire de la grille et en particulier à l'ordonnancement des jobs par le courtier de ressources sur les éléments de calcul de la grille. En effet, les premières expérimentations de G-ALHTAÏR, tel qu'il a été décrit précédemment, a permis de relever une faiblesse dans sa capacité à soutenir un

[29] Les opérations *GetCapabilities* et *DescribeCoverage* du WCS (cf. §.VI - 2.2.1.1.2), qui n'ont pas été présentées dans cette analyse, offrent au WPS la capacité de s'informer régulièrement sur les données servis par le WCS.

délai opérationnel court de ses réponses, contrainte pourtant essentielle à la gestion des crues à cinétique rapide. Les développements qui suivent ont donc cherché à perfectionner une latence, bien identifiée dans la littérature, des services de gestion des traitements de l'intergiciel de grille gLite (CHRISTODOULOPOULOS et *al*., 2008 ; GERMAIN-RENAUD et *al*., 2008 ; HUEDO et *al*., 2006 ; ROOD et LEWIS, 2009).

VI - 3. RRM-Grid : un outil garant de l'ordonnancement et d'une qualité de service

VI - 3.1. Propos introductifs : limites de performance de l'intergiciel gLite et contraintes de qualité de service

Une expérimentation préliminaire a été développée avec plusieurs challenges durant l'été 2009. Des séries de jobs (*Normal job*), exécutant le modèle ALHTAÏR, ont été soumis séquentiellement et sans spécifier de contraintes d'exécution (*Requirements*) aux courtiers de ressources de l'organisation virtuelle ESR. Les résultats obtenus (Fig. 83) reflètent les limites de performances de gLite pour la prise en charge de multiples jobs en simultanée.

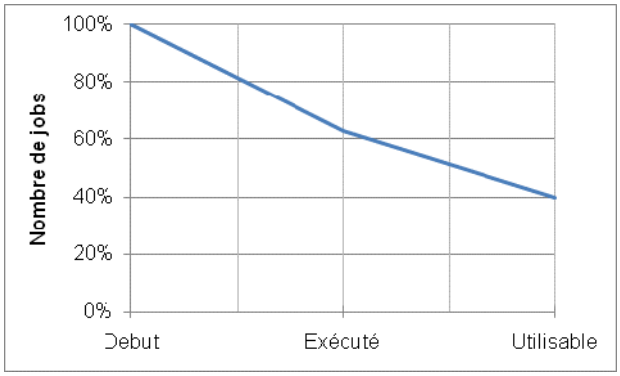


Figure 83 : Test préliminaire de l'exécution d'instances d'ALHTAÏR sur EGEE

En effet, sur l'ensemble des jobs soumis, environ 60% sont réellement exécutés (statut DONE(Ok)) mais seulement 40% sont considérés utilisables en regard des contraintes temporelles de la prévision des crues éclair. Deux observations principales ressortent de cette expérience :

- les erreurs d'exécution ont deux causes principales, le mauvais fonctionnement des services de l'intergiciel d'une part et d'autre part des problèmes de compatibilité logicielle (librairies),
- les temps de latence importants, pour atteindre l'exécution réelle du job et pour informer l'utilisateur de la fin de l'exécution (plusieurs minutes), limitent la rapidité informatique requise.

Au vue de ces premiers tests et des exigences de qualité informatique, baptisée « qualité de service » (QoS), de la prévision opérationnelle, il semble nécessaire de développer un gestionnaire de jobs permettant le contrôle autonome et efficace des jobs. Les web services à ce niveau ne permettent pas encore de palier à ce problème pourtant crucial dans les conditions opérationnelles de la gestion de crise associée aux crues à cinétiques rapides.

Le fonctionnement intrinsèque des services opérationnels, tel que ceux du SPC-GD, la particularité de la gestion des crues éclair et la nécessité d'accroitre les capacités de calcul imposent donc le respect d'une qualité de service de la part de la grille.

En premier lieu, la méthodologie développée dans la première partie de chapitre (cf. §.VI – 1), permet *a priori* de gérer un accès convivial et épisodique aux ressources de la grille EGEE, justifiant ainsi la pro-

pension de la grille EGEE à fournir à la demande un surplus de capacités informatiques. D'autre part, les délais opérationnels (c'est-à-dire la durée nécessaire à l'obtention des hydrogrammes) doivent être les plus courts possibles (COLLIER, 2007). Cette promptitude est fortement reliée à l'horizon de prévision requis. Par exemple, une prévision d'une heure est exploitable si l'hydrogramme simulé est disponible dans la demi-heure qui suit le lancement de la simulation. Ainsi, les jobs considérés dans ces travaux, permettant de prendre en charge différents scénarios de prévision, correspondent à des jobs de courte durée.

Plus globalement, la gestion de crise nécessite un fonctionnement en temps-réel des systèmes d'information permettant entre autres la tolérance aux pannes et la disponibilité continue des ressources informatiques (MAZZETTI et *al*., 2008). Pour la gestion des erreurs, le grand nombre de ressources disponibles permet de développer des solutions basées sur l'identification en temps-réel et la réallocation automatique des tâches dont l'exécution a été interrompue. Quant à la disponibilité des ressources de grille, elle apparaît plus critique et nécessite, en plus de solutions technologiques, le développement de politiques d'utilisation spécifiques. Cette garantie de disponibilité est limitée par la méconnaissance des quantités de ressources nécessaires, directement reliées à l'incertitude sur l'intensité et l'extension du phénomène à surveiller. De plus, une architecture de grille, telle que celle proposée par EGEE, n'offre pas de possibilité de prioritisation de tâches ni même de réservation préalable de ressources.

VI - 3.2. Ordonnanceur et qualité de service

Le respect de cette qualité de services implique *a priori* le développement d'une couche applicative permettant de coordonner de manière autonome la gestion automatique des jobs de grille (GERMAIN-RENAUD et *al*., 2008). Cet outil doit être capable de créer et soumettre automatiquement des jobs, tout en contrôlant régulièrement leur état. En effet, Christodoulopoulos et *al*. (2008) a établi une analyse statistique sur plus de 2 millions de jobs soumis sur la grille EGEE par l'intermédiaire de l'intergiciel gLite pendant un mois. Plusieurs indicateurs critiques ressortent de cette étude :

- 30% des jobs ont connu une erreur d'exécution,
- le temps alloué à l'enregistrement et la mise en attente d'un job peut représenter une part importante du cycle de vie de ce job,
- le transfert de données dont la taille est importante est significatif.

D'autre part, GERMAIN-RENAUD et *al*. (2008) focalisent leur étude sur la gestion des jobs de courte durée sur EGEE et observent que la soumission et l'ordonnancement de ce type de jobs prend plus de temps que leur exécution proprement dite.

Ainsi, étant donné la volatilité des ressources de grille (HUEDO et *al*., 2006; ROOD et LEWIS, 2009), une utilisation classique des services de l'intergiciel gLite ne semble pas permettre de répondre au niveau de qualité requis par la gestion de crise. Trois types principaux de problèmes sont identifiables :

- l'arrêt fréquent et inopiné du fonctionnement de ressources de grille (HUEDO et *al*., 2006),

- l'hétérogénéité des ressources peut être la source d'erreur d'exécution à cause de problèmes de compatibilité logicielle (librairies),
- les temps de latence importants des jobs de courte durée provoquent des temps d'exécution globaux trop importants (GERMAIN-RENAUD et *al*., 2008).

GERMAIN-RENAUD et *al*. (2008) plaident pour le développement d'une couche d'ordonnancement pour éliminer les temps de latence de l'intergiciel (en anglais *scheduler* ou ordonnanceur en français, les deux sont utilisés indifféremment dans la suite), et fournir des taux viables de jobs réussis et optimiser la récupération d'erreurs d'exécution. Au travers de cette recherche, les auteurs suggèrent l'utilisation d'outil tel que DIANE (Distributed Analysis Environment)[30], adapté à la grille EGEE, qui permet une meilleure prise en charge de la répartition des jobs et des erreurs de l'intergiciel au niveau utilisateur (KORKHOV et *al*., 2009). Il existe plusieurs *meta-schedulers* du même type, avec des fonctions spécifiques suivant les attentes des projets de recherche qui les ont initiés, comme DIANA (MCCLATCHEY et *al*., 2007), EMPEROR (ADZIGOGOV et *al*., 2005) ou GridWay[31] (HUEDO et *al*., 2004). DIANA évalue en temps-réel un indicateur global synthétisant les coûts pour l'exécution des jobs, le transfert des données et les caractéristiques du réseau de manière à sélectionner la meilleure ressource à utiliser. EMPEROR base sa sélection de ressources de calcul sur un algorithme d'estimation du temps d'exécution. Enfin, GridWay analyse régulièrement les performances et les disponibilités des ressources de calcul et offre des capacités de re-soumission des traitements en fonction de ces indicateurs.

En vue de répondre aux besoins de qualités de services de la prévision des crues éclair en situation de crise, il apparaît donc primordial de développer une couche d'ordonnancement automatique capable d'interagir avec l'intergiciel gLite et les requêtes utilisateurs, ayant les capacités suivantes :

- la création automatique de jobs avec le langage JDL,
- la soumission automatique de ces jobs en minimisant l'allocation de ces jobs par le WMS,
- le contrôle autonome de l'état d'exécution de ces jobs,
- la récupération des jobs ayant connu une erreur d'exécution et leur réallocation automatique,
- la récupération des résultats des modélisations aussitôt l'exécution terminée.

Une revue bibliographique a permis d'orienter les choix technologiques pour l'implémentation de ce *scheduler*. Si DIANA et EMPEROR représentent un bénéfice réel pour l'optimisation de l'allocation des jobs, ces *meta-schedulers* ont été développés au niveau de l'intergiciel alors que la présente recherche s'inscrit dans une dimension utilisateur, dépourvu de privilège administrateur sur les ressources et est donc dépendante des librairies et des logiciels fournis par ces ressources. Les ordonnanceurs de plus haut niveau, tels que DIANE et PilotJobs (ASTALOS et *al*., 2008), semblent mieux adapter aux exigences de cette étude. Ils sont basés sur une approche *pull* (le client « tire » les données vers lui à

[30] http://it-proj-diane.web.cern.ch/it-proj-diane/index.php
[31] http://www.gridway.org/

chaque besoin, à l'opposé de l'approche *push*, plus classique, où le serveur « pousse » les données vers le client) qui permet d'obtenir une meilleure répartition de charge (GERMAIN-RENAUD et *al*., 2008). Ces deux outils sont génériques et utilisables pour différentes thématiques. PilotJobs, utilisé dans différentes applications de l'organisation virtuelle *Earth Science Research* (VO ESR), a été retenu pour sa capacité :

- à accélérer la phase de soumission, en utilisant les collections de jobs (*job collections*) (BURKE et *al*., 2009),
- à utiliser le protocole RFIO (*Remote File Input/Output*) pour accéder aux fichiers stockés sur les éléments de stockage.

Le paragraphe suivant s'attache à décrire les composants logiciels de RRM-Grid basé sur les caractéristiques techniques de l'ordonnanceur PilotJobs et permettant de supporter les calculs intensifs la simulation hydrologique en situation de prévision opérationnelle.

VI - 3.3. RRM-Grid : une couche applicative garante d'une efficacité opérationnelle

L'application RRM-Grid (Rainfall-runoff modeling on Grid) est donc dédiée à l'interaction entre le système d'information client, en l'occurrence celui du SPC-GD, et les services de l'intergiciel gLite. La gestion des données est prise en charge par la fonctionnalité, baptisée RRM-Wrapper, et la gestion des tâches s'effectue au travers de l'ordonnanceur RRM-Parametric afin d'obtenir l'application RRM-Grid permettant le paramétrage dynamique de scénarios de prévision hydrologique, tels qu'ils ont été définis dans la partie précédente (cf. §.VI - 2.1.4).

VI - 3.3.1. La gestion des données avec RRM-Wrapper

Comme cela a été abordé, la gestion des données d'entrée dans l'exécution des jobs peut représenter une charge informatique importante ralentissant l'exécution globale du job. Dans le cadre du projet CYCLOPS, cette contrainte a été prise en charge grâce à un patron de conception Adaptateur (*Wrapper/Adapter Design Pattern*) utilisé pour le fonctionnement des services web. Il permet d'envoyer des messages d'état et de gérer les données d'entrée et les résultats de sortie (VERLATO et *al*., 2008). Cette capacité permet de contrôler la récupération préalable des données d'entrée, le démarrage de l'algorithme de modélisation et le stockage des hydrogrammes. L'utilisateur règle les paramètres d'exécution de l'adaptateur au travers du job de grille (JDL) (Fig. 84).

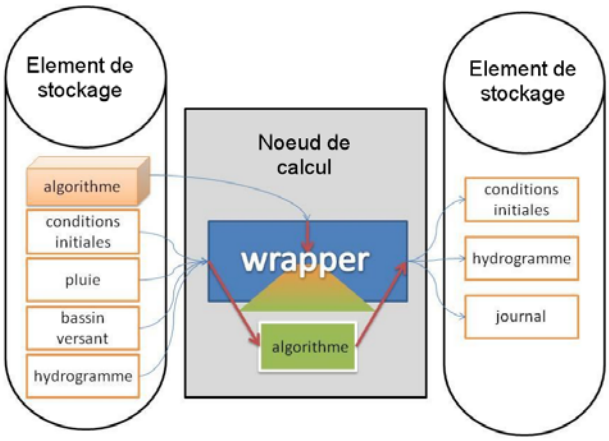


Figure 84 : Fonctionnement de RRM-Wrapper au niveau de la grille de calcul (GEISEL, 2009)

Ces paramètres de configuration sont :

- les fichiers à rapatrier avant exécution de l'algorithme :
    - l'emplacement à distance
    - les pré-conditions à remplir (décompresser, ...)
- le ou les exécutable(s) avec ses paramètres
- les fichiers de sortie
    - le nom local
    - l'emplacement de stockage à distance (SE)
    - les post-conditions à remplir (compresser, ...)
- les pré-conditions à respecter pour l'algorithme (optionnel)
    - type de CPU (32/64bit)
    - Version d'un logiciel (python 2.3 et non python 2.4...)
- la gestion des fichiers de journalisation (optionnelle)
- le choix du SE de gestion (optionnel)

D'après la figure 84, RRM-Wrapper télécharge les fichiers du bassin versant, des conditions hydrologiques initiales, de l'hydrogramme modélisé, de la prévision de pluie et le paquetage de l'algorithme ALHTAÏR avec toutes les bibliothèques nécessaires sur le nœud de calcul sélectionné (WN) par le courtier de ressources (WMS). Dès que l'ensemble de ces fichiers sont disponibles, RRM-Wrapper applique les pré-conditions de traitements, décompresse l'archive des algorithmes de production et de transfert d'ALHTAÏR et les exécute successivement. L'hydrogramme mis à jour et les nouvelles conditions hydrologiques sont stockés sur un SE prédéfini, avec les fichiers de journalisation de l'exécution.

En résumé, cet outil permet de diminuer les temps de latence au moment de la soumission et de l'allocation des jobs à l'élément de calcul en diminuant la quantité de données à transférer entre les services de l'intergiciel. D'autre part, en définissant des recommandations d'interface pour le développement d'autres algorithmes de modélisation, RRM-Wrapper assure une meilleure généricité de fonctionnement à RRM-Grid.

VI - 3.3.2. La gestion des jobs avec RRM-Parametric

L'amélioration des performances du courtier de ressources pour l'allocation des jobs est permise par RRM-Parametric fortement inspiré du canevas *Pilot jobs*, développé par l'institut d'Informatique de l'académie des sciences de Slovaquie II-SAS (Astalos et *al.*, 2008). Il s'agit d'un canevas open-source, facile à utiliser, légère et dédiée au pilotage de collections de jobs sur la grille, initialement utilisé dans le cadre de simulations astrophysiques. Cet outil est capable de gérer des simulations indépendantes. D'un point de vue informatique, il facilite et automatise les interactions avec les services de gLite.

D'après la figure 85, l'élément central de cette application est le *manager* qui supervise le traitement des différentes simulations paramétrées par l'utilisateur.

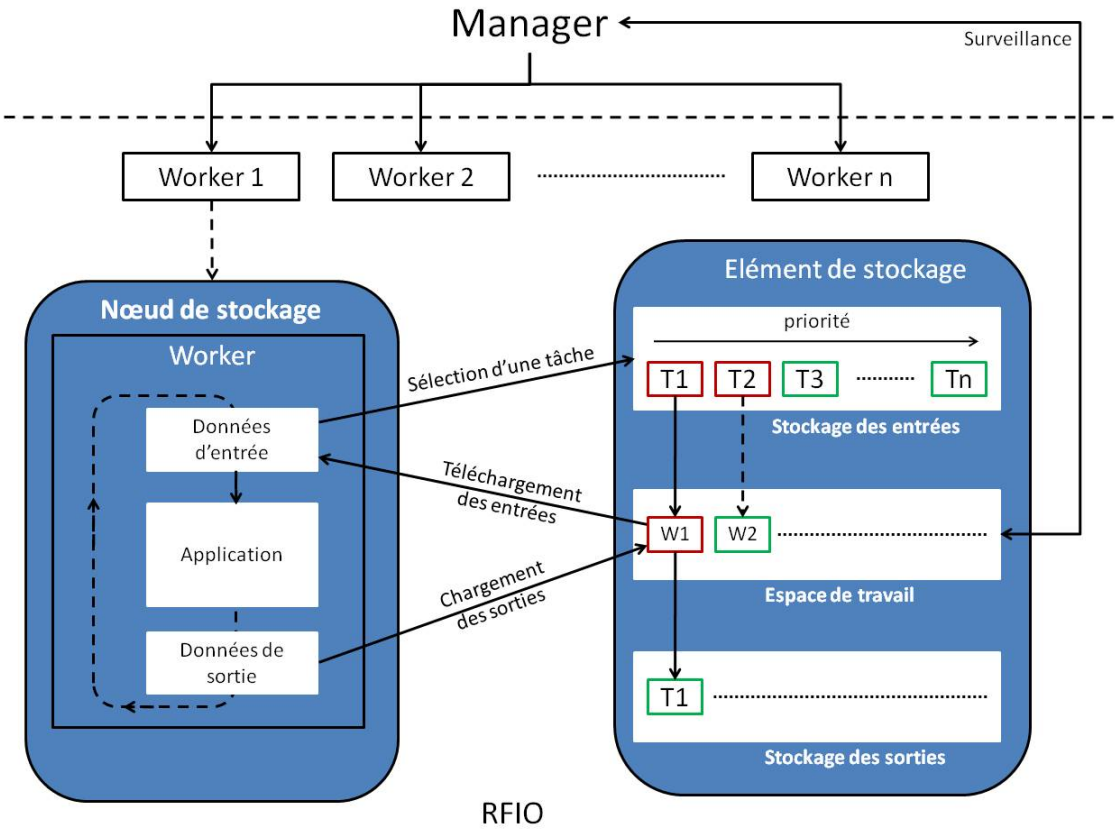


Figure 85 : Fonctionnement de RRM-Parametric pour l'allocation des jobs (GEISEL, 2009)

Depuis l'interface utilisateur, la *manager* communique continuellement avec des *workers* s'exécutant sur les noeuds de calcul et dont le nombre, souvent proportionnel au nombre de simulations à traiter, est défini par l'utilisateur. *Pilot jobs* utilise la collection de jobs de gLite pour soumettre presque simultanément tous les *workers (jobs)*. Cette fonctionnalité du courtier de ressource accélère la soumission des jobs en réutilisant la même authentification pour tous les jobs de la collection (BURKE *et* al., 2009).

Chaque *worker* prend en charge les simulations enregistrées dans une liste d'attente sur un élément de stockage prédéfini, en suivant la politique d'ordonnancement classique FIFO (*First In, First Out*), suivent une approche *pull*. Finalement, cette application est basée sur une approche *pull* où chaque *worker* prend en charge de manière autonome une nouvelle simulation aussitôt qu'il en a terminé une. Une fois que toutes les simulations sont exécutées, les *workers* libèrent progressivement les nœuds de calcul sur lesquels ils étaient exécutés.

Ce canevas générique peut être adapté pour différentes problématiques et est basé sur les services standards d'EGEE. Le *manager* s'exécute en arrière plan et automatise les tâches d'administration en vue de minimiser les interactions entre les utilisateurs et les services de l'intergiciel. Il prépare automatiquement les jobs, en générant les scripts JDL, soumet le nombre prédéfini de jobs à un courtier de ressources et les surveille grâce au service de journalisation (LB), de manière à garder informer l'utilisateur final. Le gain de performance est assuré par l'utilisation des collections de jobs, la limitation du nombre de jobs soumis, puisque un *worker* gère plusieurs simulations, et le fait qu'un *worker* reste actif sur le nœud de calcul tout au long de la procédure de prévision, limitant ainsi l'effort d'allocation du ou des courtiers de ressources.

Les contraintes « tems réel » de la prévision des crues éclair différent de celles associées habituellement à *Pilot Jobs*, du point de vue des délais de réponse et des durées d'exécution des jobs. Dans ce sens, le canevas *Pilot jobs* a été adapté pour prendre en charge :

- la priorité d'exécution des simulations,
- la réallocation automatique des jobs défectueux,
- la mise à jour d'indicateurs de performance identifiant les ressources de calcul performantes (liste blanche) et défaillantes (liste noire).

Dans la perspective d'un évènement hydrologique et avant le lancement effectif de la procédure de prévision, *RRM-Parametric* soumet automatiquement un job simple, c'est-à-dire un job fictif avec une commande Shell classique (SANCHIS, 2007) à tous les éléments de calcul alloués à l'organisation virtuelle et disponibles. Les éléments de calcul inscrits dans la liste blanche sont ceux où les jobs ont été exécutés avec succès en moins de 30 minutes.

Au début de la procédure de prévision, le *manager* prépare la liste d'attente sur un SE enrichi du protocole RFIO[32]. Basé sur la requête d'un utilisateur, il génère des fichiers de configuration pour chaque scénario de prévision (langage JSON[33]), utilisés pour créer les scripts JDL identifiant les éléments de calcul des listes blanche et noire (attribut GLUE *GlueCEUniqueID)*. La liste d'attente (contenant ces fichiers de configuration) est triée suivant l'ordre de priorité des scénarios défini par l'utilisateur. En phase de traitement, un *worker* traite toujours la simulation de plus haute priorité. Une fois que la simulation est terminée, le *worker* libère la mémoire allouée sur l'espace de travail (élément de stockage) et prend la suivante (Fig. 85). En ce qui concerne, la gestion des jobs défectueux, les simulations dont

[32] Ce protocole permet une meilleure gestion de l'accès concurrent aux fichiers, en n'autorisant qu'un seul accès simultanée à un même fichier
[33] http://www.json.org/

l'exécution a échoué, sont réallouées à la file d'attente (espace de travail) et la liste noire est automatiquement mise à jour.

VI - 3.3.3. RRM-Grid : un ordonnanceur adapté à la modélisation de scénarios de prévision hydrologique

RRM-Grid est basé sur l'encapsulation de RRM-Wrapper par RRM-Parametric. En effet, les jobs gérés par RRM-Parametric exécutent chacun une instance de RRM-Wrapper qui utilise des fichiers de configuration générés par RRM-Parametric, en fonction d'un fichier de configuration unique créé par l'utilisateur final. Ce fichier de configuration (langage JSON) définit les paramètres à utiliser pour créer l'ensemble des scénarios de prévision hydrologique :

- un élément de stockage enrichi du protocole RFIO,
- des éléments de stockage utilisés pour stocker les fichiers d'entrée et de sortie,
- les données locales à utiliser pour la modélisation (version « temps réel » du SPC-GD),
- les paramètres des scénarios de prévision,
- les priorités des simulations,
- Le nombre de *workers.*

Le diagramme de séquence UML de la figure 86 décrit l'enchaînement des opérations élémentaires de RRM-Grid.

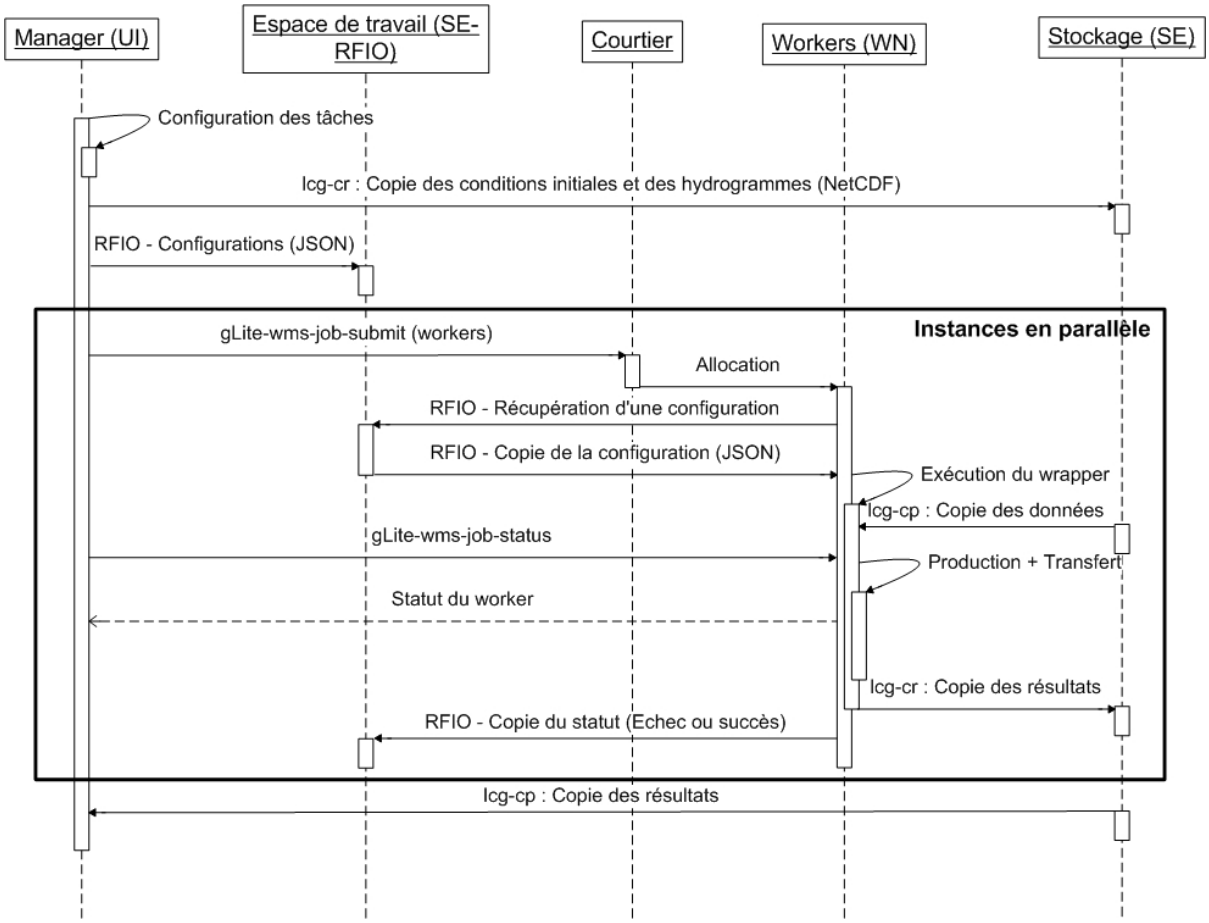


Figure 86 : Diagramme de séquence UML de RRM-Grid

À la réception de cette requête, le *manager* interprète ce fichier de configuration et génère automatiquement un fichier de configuration pour chaque scénario de prévision. Il transfère ces fichiers sur l'espace de travail grâce au protocole RFIO et les données d'entrée NetCDF (bassin versant, prévision de pluie, conditions hydrologiques initiales et hydrogramme temps-réel) sur les éléments de stockage prédéfinis, grâce aux fonctionnalités de gestion de données de gLite (LCG utils). Le manager prépare ensuite les *workers* (JDL) et les soumet au courtier de ressources en respectant les listes blanche et noire (*Requirements*). Dès que le *worker* est alloué à un nœud de calcul, le RRM-Wrapper copie les données relatives à ce scénario et lance l'exécution des fonctions de production et de transfert. Une fois cette exécution terminée, il place les hydrogrammes de prévision (NcML) sur les éléments de stockage prédéfinis. Le *worker* place le fichier de configuration du scénario terminé sur le l'espace de travail dans un répertoire destiné aux tâches réussies afin de signaler au *manager* la réussite de la tâche. En cas d'échec de la tâche, il place ce fichier dans un autre répertoire spécifique, pour permettre son traitement ultérieur. Finalement, le manager se charge de télécharger régulièrement les hydrogrammes de prévision modélisés sur le système local.

La figure 87 permet de résumer la structure logicielle et matérielle de l'ordonnanceur RRM-Grid.

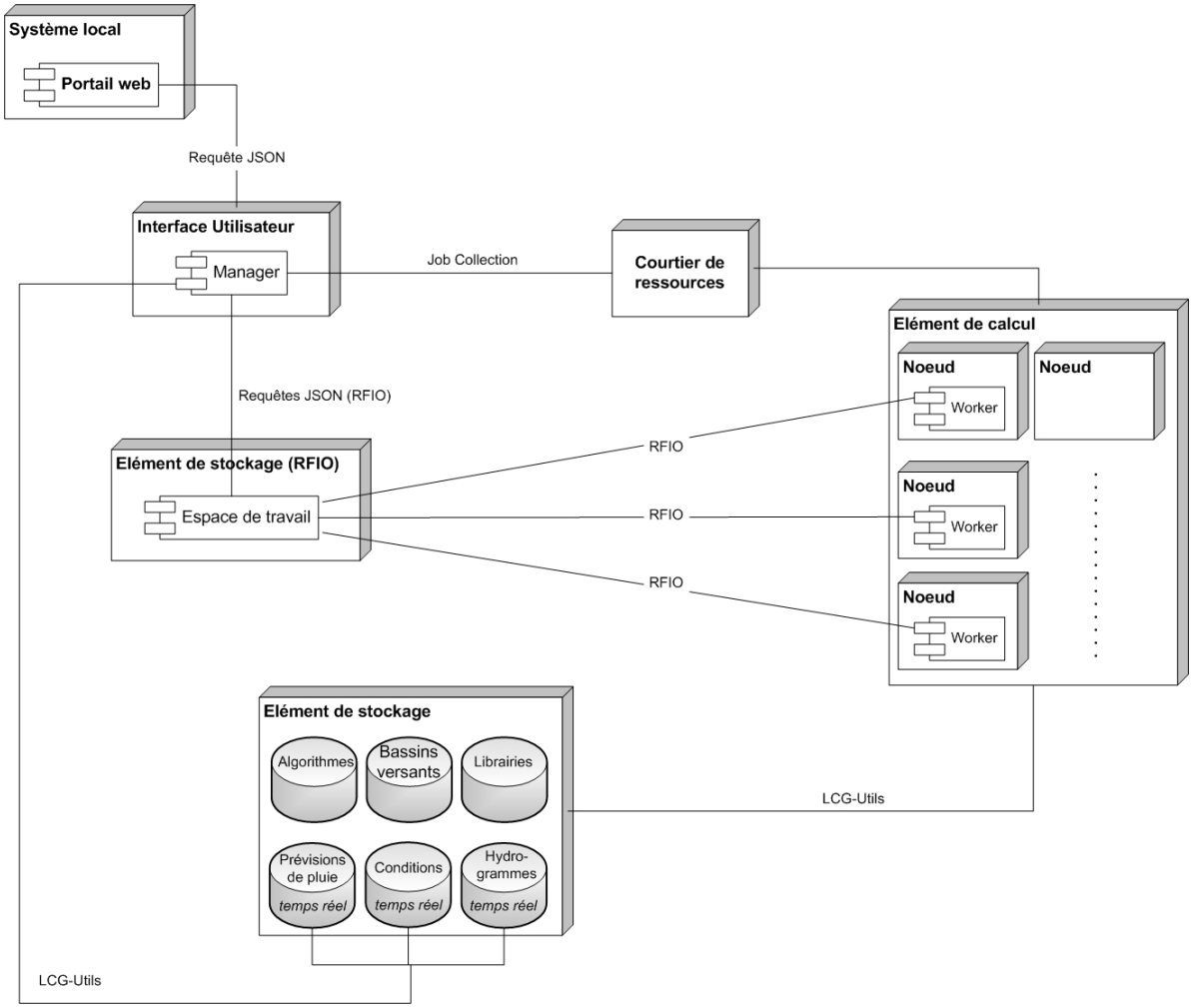


Figure 87 : Diagramme de déploiement de RRM-Grid

Le *manager* et les *workers* composants fonctionnels de l'application viennent assister le courtier de ressources dans la gestion d'un grand nombre de tâches. En effet, le schéma fonctionnel imaginé permet de limiter les demandes d'allocation au courtier de ressources, considéré comme le facteur limitant principal de la difficulté de l'intergiciel gLite à assurer une allocation rapide des jobs de grille. La méthode de soumission par collection, offre en premier lieu une solution permettant d'accélérer cette allocation (BURKE et *al*., 2009). En outre, l'exécution des *workers* sur des éléments de calcul tout au long du processus de simulation, accédant de manière autonome aux scénarios à modéliser stockés sur l'espace de travail, évite une surcharge potentielle du courtier de ressources. Dans ce sens, l'espace de travail correspond au composant principal de RRM-Grid en assurant la communication et les échanges entre le *manager* et le *worker*. Quant à la gestion de données, elle s'appuie sur le schéma classique de l'intergiciel gLite en permettant la réplication des données d'entrée afin de limiter la concurrence des *workers* dans leur accès. En effet, plusieurs scénarios de prévision, basés sur des paramètres de pondération et d'horizon de prévision différents, peuvent requérir le même jeu de données d'entrée.

Finalement, le formalisme intrinsèque du langage JSON, proche de celui du XML, et son affinité avec Internet permet d'envisager une prise en charge directe par le contrôleur, relatif à la méthode des services web de l'OGC (VI - 2.2.2.2), afin de permettre l'intégration transparente de RRM-Grid au sein du système spatial d'aide à la décision G-ALHTAÏR au travers du portail web du système local.

SYNTHESE SUR l'INTERFAÇAGE

Ce chapitre méthodologique a permis de présenter les moyens technologiques existants pour accéder aux ressources de la grille EGEE. L'interface utilisateur, est le point d'accès unique aux services de l'intergiciel gLite. Cependant, les opérations de bas niveau qu'elle offre ne calquent *a priori* pas avec les compétences techniques d'un utilisateur final, tel qu'un prévisionniste du SPC-GD.

Dans ce sens, le recours à des couches « métier » et « logique » semble indispensable pour assurer une opérationnalité suffisante au système spatial d'aide à la décision G-ALHTAÏR. Ces couches informatiques ont donc la capacité de traduire les besoins technologiques et scientifiques du prévisionniste et de mettre en communication des composants logiciels distribués. Au plus haut niveau, l'interface de cartographie en ligne intègre les données factuelles spécifiques à l'expertise hydrologique et fournit au prévisionniste une solution opérationnelle pour accéder à des services de plus bas niveau, tels que les services web de l'OGC. En tant qu'interface principale entre les systèmes d'information distribués, ces services web, et en particulier le WPS, orchestrent de manière autonome les traitements nécessaires à la réalisation des opérations de prévision hydrologique sur les ressources de la grille.

Ces traitements hydrologiques de bas niveau, exécutés sur les ressources de grille, ont nécessité une reconsidération de l'application ALHTAÏR existante. Le cœur de cette application opérationnelle, dédiée à la modélisation hydrologique des crues éclair, a été analysé et recodé en fonction des contraintes fonctionnelles et matérielles de la grille EGEE. En effet, la dimension hautement distribuée du système final, composé de l'existant technologique du SPC-GD et de G-ALHTAÏR, a impliqué une désagrégation de l'application en composants indépendants (modules hydrologiques, gestionnaire des données, etc.). Enfin, le recours à l'utilisation des services web de l'OGC a nécessité le remaniement et l'intégration des données dans un format multidimensionnel adapté, tel que le NetCDF.

Au niveau logique, l'axe principal de cette méthodologie concerne l'allocation optimale de ces instances de modélisation hydrologique aux éléments de calcul de la grille. RRM-Grid et les services de l'OGC permettent chacun à leur manière de planifier la séquence de traitements nécessaires à l'exécution de scénarios de prévision sur les ressources d'EGEE. Les services web de l'OGC, et la capacité d'interopérabilité qu'ils soutiennent, permettent d'assurer le référencement des ressources logicielles et de données, et *a fortiori* la facilité d'accès et la reproductibilité des opérations de modélisation. Quant à RRM-Grid, il permet d'aborder la dimension calculatoire et opérationnelle de l'architecture de grille EGEE et les contraintes nécessaires de qualité de service de G-ALHTAÏR. Son implémentation se base sur l'hypothèse, tirée de la littérature, que le gain de performance d'une application, nécessitant des délais de réponse courts, passe par la réduction du nombre de soumissions de jobs aux courtiers de ressources. Implémenté grâce à des langages et des services formalisés, RRM-Grid permet une certaine reproductibilité de ces méthodes de calcul. Ainsi, leur prise en charge par les services web de l'OGC semble envisageable, de manière à allier la capacité calculatoire de RRM-Grid et la propension

des services web à orchestrer les traitements sur la grille de calcul. Cette approche est approfondie dans les perspectives scientifiques de cette recherche (chapitre 8).

Finalement, au regard des caractéristiques technologiques de ces deux composants, RRM-Grid se base sur la configuration « un job de grille pour plusieurs scénarios » au contraire des services web de l'OGC qui correspondent à un schéma « plusieurs jobs de grille pour un scénario ». Ainsi, cette recherche prend le parti d'analyser et d'évaluer les performances de la technologie grille pour soutenir les opérations d'expertise hydrologique en situation de crise, principalement au travers des performances quantitatives de RRM-Grid. Cependant, les perspectives de cette recherche s'attacheront à dresser une évaluation, essentiellement qualitative, des services web de l'OGC, toujours en phase de développement sur le volet grille et non encore complètement opérationnels, et à définir des orientations scientifiques et techniques basées sur les caractéristiques et les potentialités technologiques de G-ALHTAÏR.

PARTIE 4

Résultats et Perspectives : vers l’adoption potentielle de la technologie grille

« Inutile de tirer votre épée pour couper de l'eau ; l'eau continuera à couler. » Li Po

Chapitre 7

Expérimentations de l'ordonnanceur RRM-Grid

8.1. Plan d'expérience

8.2. Résultats expérimentaux de RRM-Grid

Ce chapitre expérimente l'ordonnanceur RRM-Grid en tant que composant « logique » du système G-ALHTAÏR, afin d'évaluer sa capacité à supporter de multiples scénarios de prévision hydrologique dans des délais opérationnels courts. Dans ce sens, le faisceau de prévisions hydrologiques que RRM-Grid fournit au prévisionniste doit permettre d'enrichir sa capacité d'aide à la décision en situation de crise.

INTRODUCTION

Le chapitre 6 a permis de définir la structure et le fonctionnement de la plateforme G-ALHTAÏR dédiée à la prévision hydrologique des crues à cinétiques rapides. Principalement fondée sur les technologies du Web 2.0, et organisée en composants distribués, elle s'appuie sur les ressources informatiques de la grille EGEE, accessibles par un serveur WPS. Enfin, l'interface de cartographie en ligne, de part ses composants graphiques et son accès simplifié aux différents composants de G-ALHTAÏR, assure un niveau correct d'opérationnalité. Il a cependant été relevé, dans l'analyse des caractéristiques technologiques de la grille EGEE (cf. §.IV – 2.1.2.), des obstacles technologiques à son utilisation opérationnelle pour des traitements de courte durée, spécifiques à la prévision des crues eclar. Dans ce sens, la dernière partie du chapitre méthodologique (chapitre 6) s'est intéressée à l'aspect purement calculatoire de la grille EGEE, au travers de l'ordonnanceur RRM-Grid. L'objectif principal de cet outil est d'assister le courtier de ressources WMS de l'intergiciel gLite et de soutenir l'exécution simultanée d'un nombre important de scénarios de prévision hydrologique.

Il convient dorénavant de s'intéresser aux performances de ces outils. Comme cela a été abordé dans la synthèse du chapitre 6 les principaux résultats quantitatifs sont extraits d'expérimentations menées sur l'outil RRM-Grid. Ainsi, une attention particulière est portée sur les potentialités calculatoires de la grille EGEE pour supporter les contraintes de modélisation des crues à cinétique rapide. Une analyse comparative permet de confronter les potentialités intrinsèques de la grille EGEE avec celles *a priori* augmentées de RRM-Grid. Deux axes d'études sont privilégiés dans cette expérimentation :

- La dimension calculatoire consacrée à l'évaluation précise des capacités de prise en charge des requêtes de prévision du prévisionniste par RRM-Grid,
- La dimension hydrologique centrée sur la faculté d'anticipation des crues à cinétique rapide par un prévisionniste du SPC-GD, illustrée par un cas d'utilisation de RRM-Grid sur un évènement passé.

Le premier axe consiste à une montée en charge progressive des expérimentations de grille, de manière à mettre à l'épreuve les ressources de calcul et de stockage offertes par l'organisation virtuelle ESR pour traiter un nombre conséquent de scénarios de prévision. Le second axe cherche à démontrer l'utilité de RRM-Grid dans un contexte opérationnel pour suivre et prévoir dans des délais courts les réponses hydrologiques (débit) d'un cours d'eau affecté par un risque de crues éclair. Ainsi, l'évènement du 22 novembre 2008, qui s'est déroulé sur les bassins versants amont des Cévennes a été choisi pour présenter les capacités opérationnelles de cet outil.

VII - 1. Plan d'expérience

VII - 1.1. Ressources de grille : la VO ESR

Pour mener à bien cette phase expérimentale, un plan d'expérience a été développé en accord avec les prévisionnistes du Service de Prévision de Crues Grand Delta (SPC-GD), objet d'étude principal de cette recherche. En effet, au travers de ces expériences, il s'agit de répondre aux besoins opérationnels du SPC-GD en termes de gestion de calculs sur la grille EGEE. Dans ce sens, le modèle pluie-débit ALHTAÏR déjà porté sur la grille et l'ordonnanceur RRM-Grid ont été utilisés d'après les développements méthodologiques du chapitre précédent. Cette campagne expérimentale a eu lieu, entre 2009 et 2010, sur les ressources informatiques de l'organisation virtuelle Earth Science Research (VO ESR) (Fig. 87).

D'après les informations officielles de cette organisation virtuelle[1], les ressources de la VO ESR permettent de prendre en charge environ 1.200 jobs de grille en simultanée. D'autre part, en avril 2010, le nombre d'utilisateurs de cette VO avoisinait les 80[2]. Ainsi, lors de l'utilisation des ressources d'ESR, et dans l'interprétation des résultats de cette phase expérimentale, il est important de prendre en compte cette concurrence possible entre les différents utilisateurs de la VO. Les machines (processeurs) mises à disposition aux utilisateurs ont des puissances moyennes (c'est-à-dire a peu près équivalente à un ordinateur personnel).

Figure 87 : Répartition de la majorité des sites de calcul de la VO ESR

Les résultats expérimentaux qui suivent peuvent être scindés en deux catégories principales :

[1] https://cic.gridops.org/index.php?section=vo&vo=esr
[2] Il s'agit d'une VO de petite envergure en comparaison avec la VO BIOMED, deuxième plus grande VO, qui comptabilise presque 300 utilisateurs

- Les performances informatiques de RRM-Grid :
    - les résultats structuraux qui concernent l'état informatique de la VO ESR,
    - les résultats fonctionnels relatifs aux performances de RRM-Grid en fonction de différentes phases de simulations,
- Les résultats de prévision sur l'évènement sélectionné du 22 novembre 2008 sur le bassin versant du Gardon d'Anduze.

VII - 1.2. Méthodes de traitement expérimentées

Dans ce chapitre, il convient donc de mettre en exergue les potentialités de calcul de RRM-Grid et par extension celles de la VO ESR. À ce niveau, il est important de noter l'aspect expérimental de RRM-Grid, qui au moment de la rédaction de ce manuscrit nécessite certainement des améliorations technologiques permettant de profiter pleinement des ressources offertes par la VO ESR. Ainsi, l'ensemble des résultats technologiques doit être appréhendé en fonction des potentialités de l'ordonnanceur RRM-Grid plutôt que de celles de la VO ESR, qui *a priori* sont supérieures.

Pour rappel, un <u>scénario de prévision</u> est :

une instance de modélisation hydrologique intégrant les modules de production et de transfert de l'application ALHTAÏR, et prenant en charge une prévision de pluie à horizon variable sur laquelle une pondération positive ou négative est appliquée sur ses intensités de pluie.

L'objectif de cette analyse technologique est de comparer les performances de RRM-Grid avec une utilisation classique de la grille. Trois méthodes de traitement sur la grille sont comparées et discutées :

- Méthode 1 : une utilisation classique de la grille qui consiste à envoyer les jobs (scénarios de prévision) séparément au courtier de ressources (WMS). Ainsi, une routine informatique permet de créer automatiquement et d'envoyer les scénarios (JDL) au courtier. L'ensemble des données d'entrée (algorithmes et données d'entrée hydrologiques) est chargé dans l'*InputSandbox* de chaque job et transite entre les services de l'intergiciel tout au long du traitement du job.
- Méthode 2 (RRM-Wrapper) : une utilisation intermédiaire représentée par RRM-Wrapper (cf. §. VI – 2.3.1.) qui se base sur la même approche que la précédente, mais dans laquelle les données d'entrée sont gérées une fois que le job (scénario) est pris en charge par l'élément de calcul. Cette méthode permet d'éviter le transfert des données d'entrée au moment de la soumission.
- Méthode 3 (RRM-Grid) : l'utilisation de RRM-Grid qui encapsule RRM-Wrapper et qui assure l'ordonnancement des scénarios de prévision en limitant l'utilisation du coutier de ressources de l'intergiciel gLite. A la différence des deux premières méthodes, l'expérimentation de RRM-Grid se base sur la préparation du fichier de configuration globale (JSON) et le démarrage de l'outil depuis l'interface utilisateur.

La différence majeure de la méthode 3 en comparaison avec les deux premières, concerne la relation job / scénario. En effet, dans le cadre de la 3ème méthodologie, un job (*worker*) prend en charge successivement plusieurs scénarios, alors que les méthodes 1 et 2 ont un schéma « un job – un scénario ».

Ces 3 méthodes expérimentales prennent successivement en charge 4 simulations différentes dans lesquelles le nombre de scénarios, les horizons de prévisions et le pas de pondération varient (Tab. 6) :

- 33 scénarios de modélisation avec des horizons de prévision de 1, 2 et 3h et un intervalle de variation de -50 à + 50% avec un pas de 10%
- 66 scénarios de modélisation avec des horizons de prévision de 1, 2 et 3h et un intervalle de variation de -50 à + 50% avec un pas de 5%
- 105 scénarios de modélisation avec des horizons de prévision de 1, 2, 3, 4 et 5h et un intervalle de variation de -50 à + 50% avec un pas de 10%
- une simulation de plus grande ampleur, composée de 210 scénarios de modélisation avec des horizons de prévision de 1 à 10h et un intervalle de variation de -50 à + 50% avec un pas de 5%, est expérimentée. Cette simulation doit permettre d'affiner l'analyse de la capacité de RRM-Grid et *a fortiori* de la grille EGEE pour supporter une montée en charge, ainsi que l'effet du nombre de *workers* sur la performance générale de cette simulation. Ainsi, une analyse comparative est menée entre plusieurs simulations de 210 scénarios dans lesquelles le nombre de *workers* soumis varie de 40 à 200.

Scénarios / Méthodes	33	66	105	210
WMS	x	x	x	-
RRM-Wrapper	x	x	x	-
RRM-Grid	x	x	x	x

Tableau 6 : Synthèse de la méthodologie expérimentale

En choisissant trois horizons de prévision, il est ainsi possible de tester les capacités de RRM-Grid à assurer les traitements de manière ordonnée. En effet, l'ordre de traitement des scénarios de prévision est contraint par des niveaux de priorité directement reliés à l'horizon de prévision requis. Plus cet horizon est court, plus sa priorité est haute.

Le choix de quatre séquences de simulation progressives (de 33 à 210 scénarios) peut être justifié par deux arguments principaux :

- il semble intéressant d'évaluer l'effet du nombre de jobs à gérer durant la phase de soumission, qui apparait comme une phase critique de l'exécution des jobs de courte durée,
- en soumettant 210 jobs, cette expérimentation apparait représentative d'un processus complet de prévision en situation opérationnelle.

Pour le second argument, la prévision des débits sur les 30 bassins versants[3] surveillés par le SPC-GD, en appliquant des délais de prévision de 1, 2 et 3h et des variations de -50 à + 50%, implique la prise en charge de 990 scénarios en simultanée. Or, un nombre aussi important de scénarios apparait difficilement interprétable en situation de crise par les prévisionnistes du SPC-GD. Ainsi, en considérant ces 30 bassins, une simulation à 210 scénarios représente une situation proche de la réalité, puisqu'elle correspond à 7 scénarios de prévisions par bassin versant.

Plus généralement, le choix de ces 4 simulations est représentatif de la charge de calcul nécessaire pour analyser les évolutions potentielles des réponses hydrologiques des cours d'eau réglementaires. En effet, une simulation de prévision hydrologique durant une situation hydrométéorologique extrême, c'est-à-dire généralisée à la majorité des cours d'eau réglementaires du SPC-GD, nécessite la surveillance régulière et précise de tous les bassins versants.

VII - 1.3. Critères d'analyse

Il s'agit de définir les indicateurs de performance permettant de comparer l'efficacité des trois méthodologies de prise en charge des scénarios de prévision.

VII - 1.3.1. Phases principales de l'exécution d'un job

Tout d'abord, il est important de définir les phases temporelles permettant de mesurer l'efficacité de l'exécution des jobs de grille. Le tableau 7 présente les trois intervalles de temps pour lesquels un suivi spécifique a été mené.

Phase	Intervalle de temps
Soumission	Durée pour l'obtention du statut « Running »
Exécution	Durée pour l'obtention de la fin de l'exécution des algorithmes de modélisation hydrologique
Validation	Durée pour l'obtention du statut « Done »

Tableau 7 : Phases temporelles observées dans la phase expérimentale

La phase de « Soumission » est composée des statuts *Submitted*, *Waiting*, *Ready*, *Scheduled* d'un job. Dans certains cas, cette phase peut s'avérer la plus longue en comparaison avec l'exécution proprement dite de l'algorithme pour lequel le job est dédié.

La phase d' « Exécution » est incluse dans le statut *Running* et correspond à la récupération des données d'entrée, quand cela est nécessaire, et à l'exécution de l'algorithme de modélisation sur le nœud de calcul.

La phase de « Validation » est équivalente à la période de temps requise pour que le courtier de ressources informe l'utilisateur de la fin d'exécution du job et de la disponibilité des résultats. Cette phase

[3] En 2009, le modèle ALHTAÏR installé dans les locaux du SPC-GD comportait 30 bassins versants à modéliser.

a été sélectionnée pour illustrer sa faible réactivité dans le cas d'une utilisation classique de l'intergiciel gLite (Méthode 1).

Ainsi, ces trois phases ont pour objectif de synthétiser le cycle de vie d'un job, et par extension d'un scénario de prévision. En effet, il semblait nécessaire de regrouper les statuts initiaux de ce cycle de vie (*Submitted*, *Waiting*, *Ready*, *Scheduled*), dans un la phase de « Soumission », qui correspondent globalement à l'ensemble des opérations mises en œuvre pour allouer le job de grille au nœud de calcul.

VII - 1.3.2. Indicateurs de performance

Tout d'abord, dans le cadre de cette recherche et en accord avec les prévisionnistes du SPC-GD, le délai opérationnel du traitement des scénarios de prévision a été fixé à la moitié de l'horizon de prévision requis (c'est-à-dire 30 minutes pour une prévision d'une heure, 1h pour une prévision de 2h, etc.). Cet indicateur s'avère important pour distinguer les scénarios réellement utiles au prévisionniste aux scénarios exécutés avec succès, mais dont le temps trop long ne permet pas au prévisionniste d'effectuer une expertise hydrologique dans les délais qui lui sont alloués.

Ainsi, les principaux indicateurs de performance qui ont été sélectionnés dans le cadre de cette expérimentation correspondent :

- au temps de réponse pour l'exécution complète d'un scénario (ce temps correspond à la phase de soumission, d'exécution et de validation du job de grille). Cet indicateur est en partie dépendant de l'horizon de prévision du scénario. L'enjeu concernant cet indicateur se situe dans la capacité de RRM-Grid à traiter en priorité les scénarios dont l'horizon de prévision est le plus faible, en comparaison avec les deux autres méthodes dans lesquelles cette capacité n'existe pas,

- à la fiabilité des exécutions sur la grille étant donné le niveau de sécurité que la gestion de crise requiert. Il s'agit de relever le taux d'échec des jobs de grille ainsi que l'origine de ces erreurs (algorithmique, compatibilité, intergiciel, …).

- au nombre d'occurrences de simulation qu'il est possible de soumettre. Cet indicateur est *a priori* dépendant des ressources fournies par la VO ESR. Il semble difficile de spécifier une valeur précise. Dans ce sens, la simulation comprenant 210 scénarios est censée représenter une phase de prévision intense durant laquelle un grand nombre de cours d'eau réglementaires (bassins versants) sont concernés par l'évènement hydrométéorologique.

Étant donné la volatilité relative d'une architecture de grille telle qu'EGEE, et le positionnement « utilisateur » de cette recherche, ces indicateurs doivent être appréciés en tenant compte de nombreux facteurs contingents, correspondant à l'utilisation des ressources au moment de l'expérimentation ou encore la défaillance de sites de calcul (ou à une échelle plus petite de ressources informatiques), qu'il est difficile d'évaluer en temps réel.

Finalement, pour mener à bien cette expérimentation, une interface utilisateur a été installée dans les locaux du LGEI[4] à l'école des Mines d'Alès[5]. En effet, ce point matériel et logiciel permet une interaction rapide et efficace entre le système local, en charge de la prévision hydrologique « temps réel » et l'architecture de grille dédiée à la phase de prévision. Il évite d'une part une manipulation contraignante des données par session SSH, et d'autre part, l'augmentation des temps de réponse des jobs de grille.

VII - 1.4. Étude de cas en situation opérationnelle

VII - 1.4.1. Stratégie de construction des scénarios

Afin d'analyser la plus value opérationnelle que RRM-Grid est susceptible d'amener au SPC-GD, une démonstration de son utilisation est présentée. L'un des objectifs sous-jacents est de démontrer la viabilité éventuelle de cette méthodologie, et par extension de ces technologies émergentes, pour soutenir les opérations du SPC-GD liées à l'expertise hydrologique. Ce second volet d'analyse cherche donc à discuter la propension de RRM-Grid à soutenir une capacité de prévision hydrologique permettant aux prévisionnistes du SPC-GD d'anticiper à temps une pointe de crue et son délai de formation.

Cette fiabilité est de prime abord discutable, puisque comme cela a été démontré, l'une des principales difficultés dans l'anticipation des crues à cinétique rapide concerne l'impossibilité d'obtenir des prévisions météorologiques fiables à court terme et sur une extension géographique de faible taille. Étant donné cette incertitude sur la prévision de pluie CALAMAR, et la forte sensibilité du modèle ALHTAÏR, utilisé dans ce cadre, à la donnée de pluie (AYRAL, 2005 ; MARCHANDISE, 2007), les scénarios de pluie générés et injectés dans G-ALHTAÏR correspondent uniquement aux « prévisions de pluie » actuellement accessible aux prévisionnistes à savoir :

- une prévision CALAMAR 1h « extrapolée » à un délai de prévision supérieur[6]. La méthodologie pour construire ces scénarios de pluie fictifs consiste en une simple duplication de la prévision CALAMAR 1h pour les horizons suivants (Fig. 88),
- une prévision météorologique sur les zones AP à pluie constante et homogène sur un délai de prévision plus long terme (de 0 à 12h), délai de prévision généralement spécifié dans la prévision AP (par ex. 200 mm en 24h).

[4] Laboratoire de Génie de l'Environnement Industriel et des Risques Industriels et Naturels
[5] http://egee-uig.web.cern.ch/egee-uig/production_pages/InstallUI.html
[6] Pour préserver une certaine validité hydrométéorologique, ce délai de prévision météorologique ne dépasse pas 5h dans les expérimentations qui suivent

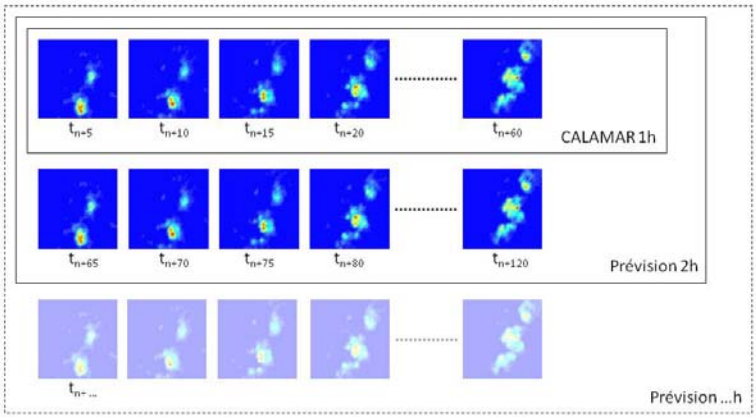


Figure 88 : Méthode de construction des scénarios de pluie basés sur CALAMAR

De plus, une pondération est effectuée sur les intensités de pluie construites par duplication et sur le scénario de pluie constante obtenu avec les prévisions AP (ces pondérations varient ici de -50% à +50%). Dans ce sens, en situation opérationnelle, la connaissance de l'incertitude sur la prévision météorologique doit permettre d'établir un intervalle de variations à appliquer sur cette donnée, utilisée comme entrée des modèles hydrologiques.

VII - 1.4.2. Matériel expérimental : Bassin versant et évènement hydrologique

La zone test choisie pour cette étude expérimentale est le bassin versant du Gardon d'Anduze (545 km^2). Ce bassin versant est situé en amont de la ville d'Anduze dans le Gard, qui correspond à l'exutoire de ce bassin versant. Le choix a été guidé par deux critères principaux (AYRAL, 2005) :

- le modèle hydrologique de l'application ALHTAÏR est principalement adapté aux bassins versants de petite taille où les processus hydrologiques prédominent sur les phénomènes hydrauliques, ce qui correspond au bassin versant du Gardon d'Anduze,
- de nombreuses recherches se sont attachées à décrire, comprendre et modéliser ces processus hydrologiques sur ce bassin versant que l'on peut qualifier de pilote.

Comme cela a été abordé dans le chapitre 1 (cf. §.I – 2.3.), la vigilance hydrologique correspond à l'outil réglementaire principal du SPC-GD pour informer les partenaires opérationnels et la population sur l'état hydrologique des cours d'eau surveillés. A ce titre, le tableau 8 présente les seuils des hauteurs d'eau et des débits pour le bassin versant d'Anduze correspondant aux différents niveaux de vigilance.

Niveau de Vigilance	Hauteur (m)	Débit (m^3/s)
Jaune	2	150
Orange	3,9	540
Rouge	5,5	1030

Tableau 8 : Hauteurs et débits correspondant aux niveaux de vigilance (d'après SPC-GD, 2006 et la courbe de tarage à Anduze du BVNE[7])

[7] Bassins Versants Numériques Expérimentaux : http://www.imft.fr/recherche/hydroeco/theme1-5.html

Cette information est primordiale pour l'analyse de la capacité de prévision hydrologique de RRM-Grid et par extension du SPC-GD. En effet, il s'agit d'examiner en premier lieu la propension de RRM-Grid à prévoir le dépassement de ces seuils dans une situation opérationnelle. A un second niveau, la capacité d'anticipation de la pointe de crue en temps et en intensité sera envisagée.

À la vue de la description de l'approche développée et des limites intrinsèques de la prévision de pluie, cette phase expérimentale doit être appréhendée comme une démonstration des potentialités opérationnelles offertes par RRM-Grid. L'objectif ne se situe donc pas dans l'analyse de la prévision en tant que telle qui nécessiterait une approche scientifique différente et une analyse rigoureuse de l'ensemble des incertitudes de la modélisation pluie-débit (MARCHANDISE, 2007), et une réflexion de fond sur la prévision météorologique et sa génération expérimentale. Cette analyse montre toutefois les capacités de prévision offertes par le modèle ALHTAÏR en intégrant dans ce dernier les « prévisions » de pluies actuellement à disposition du SPC-GD.

Ainsi, la démonstration de cette capacité de prise en charge de scénarios de prévision hydrologique de manière intensive s'appuie sur un évènement de crues à cinétique rapide d'intensité moyenne représentatif d'une activité routinière du SPC-GD pour prévoir les évènements de crues éclair.

L'évènement du 22 octobre 2008 correspond à une réaction hydrologique « classique[8] », de l'ordre de la crue décennale avec un débit de pointe de l'ordre de 1.000 m³/s à l'exutoire du bassin versant (Fig. 90) et des cumuls pluviométriques avoisinant les 400 mm sur la durée de l'évènement (Fig. 89).

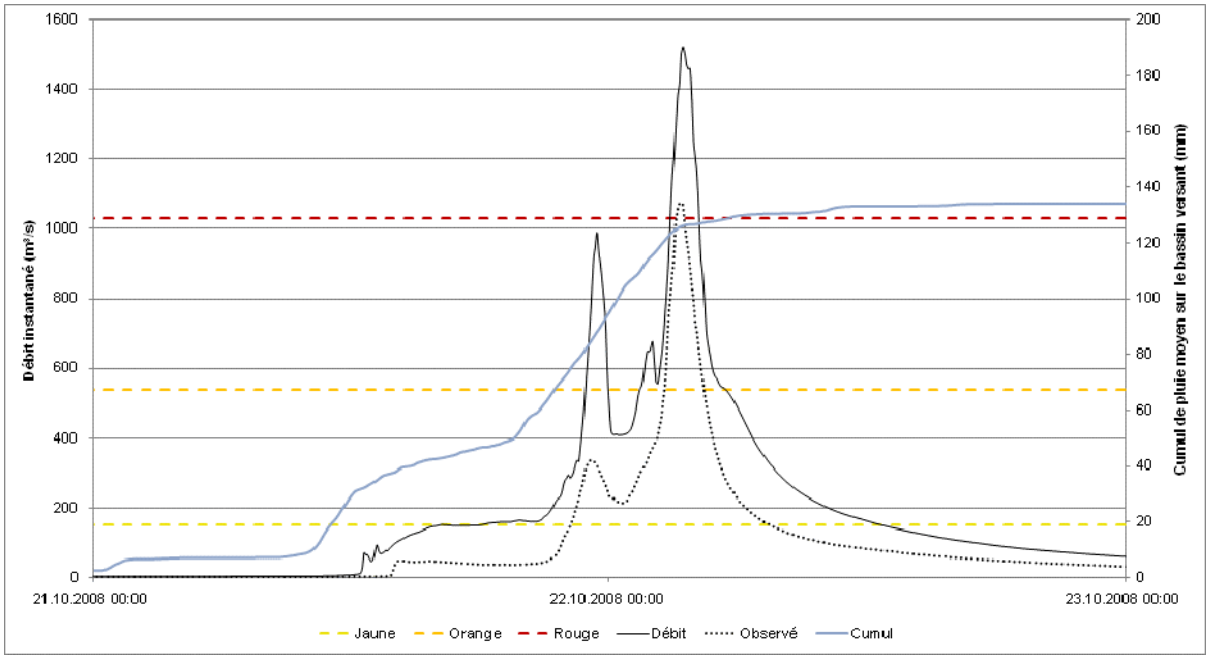


Figure 89 : Hydrogrammes de crue observé et modélisé de l'évènement du 22 novembre 2008 à l'exutoire du bassin versant d'Anduze

Ce débit correspond à une période de retour de 10 ans entrainant le dépassement du seuil de vigilance rouge et sur le terrain à la submersion du pont submersible d'Anduze. En modélisant cet évènement

[8] En opposition d'extrême, comme cela est le cas pour les deux évènements présentés (cf. §.III – 2.1.)

avec ALHTAÏR en temps différé, une faible performance peut être observée (Fig. 89), avec un critère de Nash[9] de 0,41. Cet indice correspond à la fonction critère suivante :

$$\text{Nash} = 1 - \frac{\sum_{i=0}^{n} (xi - yi)^2}{\sum_{i=0}^{n} \left(yi - \overline{yi}\right)^2}$$

où les xi désignent les n valeurs calculées et yi les n valeurs observées.

L'hydrogramme simulé par ALHTAÏR est fortement surestimé. Toutefois, la cinétique de l'événement est bien reproduite par le modèle, les dépassements des seuils des niveaux de vigilance étant particulièrement synchrones. Si cette simulation renvoie aux difficultés de modélisation de ce type d'événement (AYRAL, 2005 et MARCHANDISE, 2007), il montre toutefois la potentialité opérationnelle du dit modèle notamment pour rendre compte de la cinétique de l'événement.

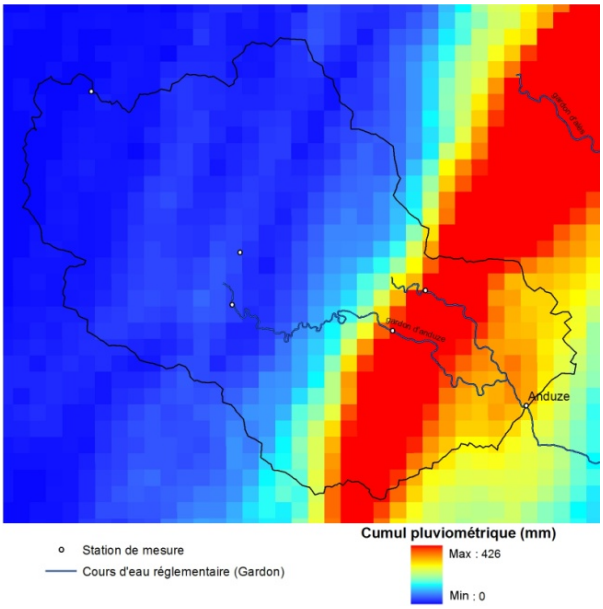


Figure 90 : Cumuls pluviométriques sur le bassin versant d'Anduze entre le 19/10/2008 et le 23/10/2008 (source : RHEA)

La stratégie développée pour cette phase expérimentale se base sur l'analyse des données de prévision de pluie CALAMAR et sur les cumuls pluviométriques annoncés par la prévision AP. En effet, en situation opérationnelle, le prévisionniste est susceptible d'utiliser en temps réel ces données et l'hydrogramme observé pour construire les scénarios de prévision qu'il souhaite modéliser.

[9] Le critère de Nash est le critère de comparaison des hydrogrammes de crues le plus répandu dans le domaine de l'hydrologie (NASH et SUTCLIFFE, 1970).

VII - 2. Résultats expérimentaux de RRM-Grid

La réalisation du plan d'expérience présenté précédemment permet d'établir un ensemble de résultats qu'il convient d'analyse dans ce paragraphe.

VII - 2.1. Performances technologiques

Il s'agit dans ce paragraphe de démontrer les capacités technologiques de RRM-Grid, en comparaison avec une utilisation de plus bas niveau telle qu'elle est offerte par défaut par les services de l'intergiciel gLite.

D'une manière générale, l'ensemble des résultats permettent de relever une tendance forte de la capacité technologique de RRM-Grid. En effet, le tableau 9 illustre la forte différence de performance de la méthode 3 (RRM-Grid) pour les trois premières simulations (33, 66 et 100 scénarios) en comparaison avec les deux premières méthodes (utilisation classique et RRM-Wrapper).

Méthodes	Taux d'exécution		Taux d'utilisation	
	Réussite	Échec	Réussite	Échec
1	63 %	37 %	40 %	60 %
2	57 %	43 %	46 %	54 %
3	97 %	3 %	96 %	4 %

Tableau 9 : Performances globales des trois méthodes dans le traitement des simulations prenant en charge 33, 66 et 105 scénarios de prévision

Deux indicateurs permettent de démontrer cette capacité. D'une part, le taux d'exécution, c'est-à-dire la prise en charge et l'exécution complète d'un scénario, approche les 100% alors que pour les deux premières méthodes cette valeur oscille autour de 50%. De la même manière, parmi l'ensemble de ces scénarios réellement exécutés seuls respectivement 40% et 46% d'entre eux sont considérés comme exploitables pour développer une prévision hydrologique dans le cas des méthodes 1 et 2. Cette exploitabilité est en relation avec le temps nécessaire pour l'obtention d'un scénario en fonction de son horizon de prévision. Ainsi, à la vue de ces résultats (96%), RRM-Grid semble apte à assurer une relative fiabilité dans les traitements requis par le prévisionniste.

VII - 2.1.1. Gestion des données d'entrée : RRM-Wrapper

Il s'agit dans un premier temps de s'intéresser à la capacité de prise en charge des données offerte par RRM-Wrapper. Les méthodes 1 et 2 qui correspondent respectivement à l'utilisation classique de la grille et à RRM-Wrapper, ont été testées sur les trois premières simulations (33, 66 et 105 scénarios). Les résultats sont présentés dans les deux graphiques de la figure 91.

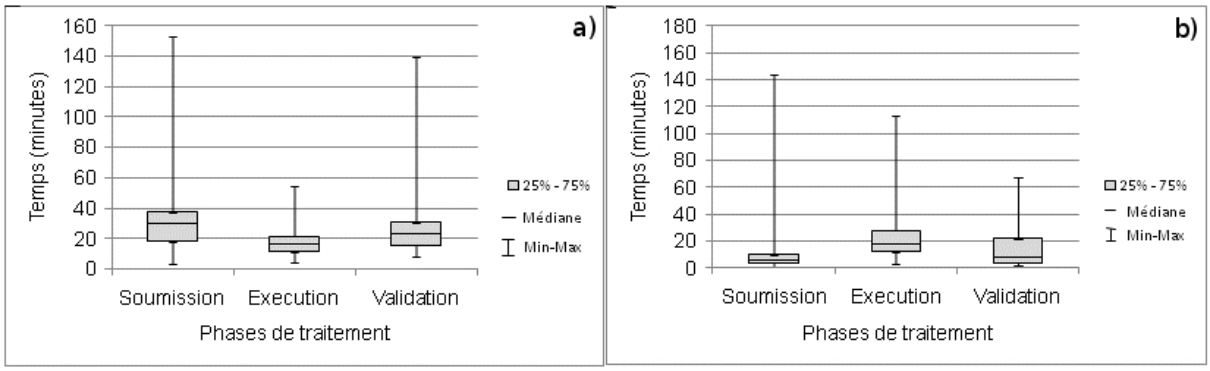


Figure 91 : Boîtes à moustaches des temps d'exécution des trois phases de traitement avec la méthode 1 (a) et la méthode 2 (b)

Ces deux diagrammes permettent de représenter la distribution des différentes durées des trois phases d'exécution d'un job sélectionnées pour cette expérience. Chaque « moustache » est constituée d'une ligne basse (du 0 au 25ème centile), d'une boîte (du 25ème au 75ème centile) et d'une ligne haute (du 75ème centile au 100ème centile). Les deux extrémités représentent enfin les valeurs maximale et minimale du délai.

L'observation des ces deux diagrammes permet de relever plusieurs tendances :

- La phase de soumission des scenarios de prévision est plus rapide avec la méthode 2 (Fig. 91b). En effet, 75% des jobs sont acceptés par le nœud de calcul pour exécution en moins de 10 minutes, alors que pour la méthode 1, 75% des jobs soumis sont acceptés après 40 minutes.

- La durée de l'exécution des modules hydrologiques (comprenant la gestion des données d'entrée et de sortie) est plus étalée pour RRM-Wrapper. En effet, l'écart interquartile se situe entre 12 et 28 minutes, alors que pour la figure 91a, il se situe entre 12 et 21.

- Pour la phase de validation, une différence de 10 minutes est observée pour 75% des jobs entre les deux méthodes. La méthode 2 permet d'obtenir un statut « Done » pour 75% des jobs après approximativement 20 minutes (Fig. 91b) à la différence de la méthode 1 qui nécessite une trentaine de minutes (Fig. 91a).

- Les maximums sont importants pour les deux méthodes, en particuliers pour la phase de soumission et de validation dans le cas de la méthode 1 et pour l'exécution dans le cas de la méthode 2.

En analysant ces observations, il est possible de relever l'intérêt d'utilisation de RRM-Wrapper pour accélérer la soumission et la validation des jobs. Ce gain de performance est relié :

- à l'allégement de l'*InputSandbox* de chaque job,

- au stockage des données de sortie sur un élément de stockage plutôt que leur transfert par l'*OutputSandbox*,

- à la notification de fin d'exécution du job par le service d'enregistrement et de comptabilité (LB, *Logging and Bookkeeping*), dès que la modélisation est terminée.

Cependant, la phase d'exécution de RRM-Wrapper est en moyenne plus longue qu'avec la méthode 1. Cette augmentation est due à la gestion des données d'entrée et de sortie au niveau de l'exécution des traitements sur le nœud de calcul. L'impact de cet accroissement doit toutefois être nuancé, du fait de temps d'exécution globaux plus bas que dans le cas d'une utilisation classique. En effet, la durée médiane des jobs est égale à 38 minutes avec RRM-Wrapper contre 66 minutes avec la méthode 1.

Il est à noter à ce niveau que RRM-Wrapper permet une généralisation de la méthode d'accès aux données[10] (cf. §.VI – 2.3.1.). Cependant, les performances relativement basses dans le traitement des scénarios, synthétisées dans le tableau 9, imposent une meilleure gestion du traitement des jobs et de leurs erreurs d'exécution éventuelles.

Finalement, le transfert des données entre le système local et les ressources de stockage de la grille est un point crucial de cette méthodologie. Il semble difficile d'évaluer précisément ces capacités étant donné leurs dépendances à la largeur de bande passante et au débit offert par la connexion du système local au réseau Internet. Au sein de la grille, cette gestion n'apparait plus à l'heure actuelle comme un facteur limitant étant donné la qualité de service requise par la majorité des applications de grille. De plus, la gestion de données de grande taille est permise grâce à la mobilité des algorithmes, principe fondamental de la grille, qui permet d'envoyer ces algorithmes à proximité géographique des données, évitant ainsi le transfert de ces données.

VII - 2.1.2. Gestion du traitement des scénarios : RRM-Grid

Le tableau 9 a permis de relever le gain de performance qu'offre l'outil RRM-Grid avec la prise en charge de la quasi-totalité des scénarios de prévision dans des délais opérationnels permettant au prévisionniste de tirer profit de l'information apportée par l'analyse hydrologique de ces scénarios.

Il s'agit, au travers de ce paragraphe de s'intéresser aux performances individuelles de la phase de soumission, du traitement des scénarios dans leur ensemble, et la réactivité de RRM-Grid à la montée en charge.

VII - 2.1.2.1. Phase de Soumission

Il s'agit de s'intéresser à l'une des phases les plus critiques de la gestion des tâches de l'intergiciel gLite, à savoir la soumission. La figure 92 montre les performances de soumission pour les méthodes 1 et 3 (respectivement l'utilisation classique de gLite et RRM-Grid) en fonction des 3 premières simulations, c'est-à-dire composées de 33, 66 et 105 scénarios.

[10] Ces données correspondent aux données d'entrée du modèle, aux librairies utiles au fonctionnement des algorithmes et aux algorithmes eux-mêmes

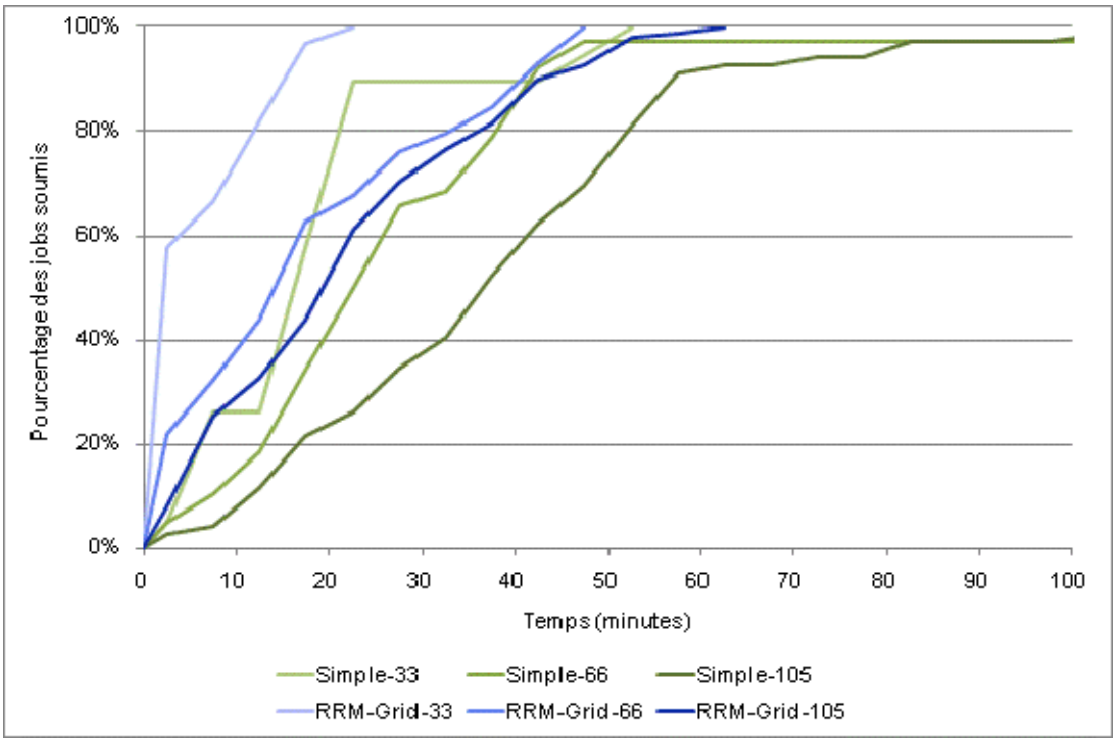


Figure 92 : Performance de la phase de soumission pour les méthodes 1 et 3 (pour plus de clarté, les mesures dont la durée dépassées les 100 minutes ont été retirées du graphique)

Il est tout d'abord important de signaler que l'ensemble de ces statistiques de soumission sont basées sur les scénarios réussis. Les scénarios ayant connus une erreur de soumission ou de traitement par les services de l'intergiciel sont donc exclus. Plusieurs remarques émergent de cette figure :

- un assez bon regroupement des trois simulations de chaque méthode. En effet, les courbes de la méthode RRM-Grid (bleues) sont systématiquement plus redressées que pour la méthode classique (courbes vertes). Cette observation plaide en faveur d'une spécificité fonctionnelle de chaque méthode.

- les courbes bleues (RRM-Grid) montrent une relative régularité dans la soumission des scénarios (forme rectiligne), au contraire des courbes de la méthode 1 qui montrent une plus grande intermittence dans la soumission des scénarios.

- les 2 méthodes sont influencées par le nombre de scénarios à traiter, en effet les 3 courbes les plus foncées sont systématiquement les plus inclinées des trois simulations (33, 63, 105 scénarios).

L'analyse de cette phase de soumission permet de plaider pour une utilisation de la méthode RRM-Grid. En effet, l'intérêt, déjà énoncé dans le paragraphe précédent de RRM-Wrapper, se combine à la plus grande capacité de soumission des collections de jobs de gLite, implémentées dans RRM-Grid. De plus, cette figure justifie l'utilisation de priorité de traitement dans RRM-Grid, du fait de la forte influence du nombre de jobs soumis en simultanée. Il convient donc d'assurer la soumission en priorité des jobs contraints par des délais opérationnels courts, c'est-à-dire ceux dont l'horizon de prévision est le plus faible.

VII - 2.1.2.2. Performance globale du traitement des scénarios

Il convient dorénavant de spécifier le gain de performance offert par RRM-Grid en comparant les performances d'exécution par délai de prévision (Fig. 93).

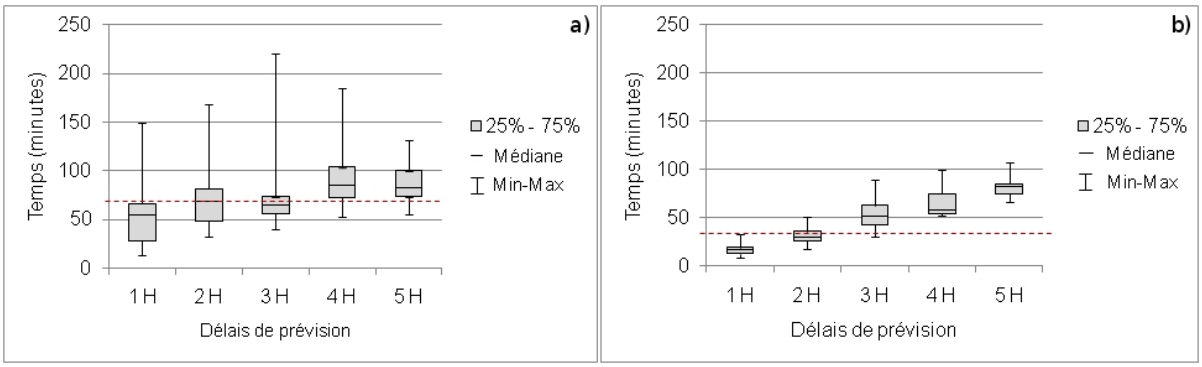


Figure 93 : Boîtes à moustaches des temps d'exécution pour chaque délai de prévision avec la méthode 1 (a) et la méthode 3 (b), la ligne horizontale en tiret rouge représente la médiane globale des temps d'exécution

Ces deux boîtes à moustaches montrent les percentiles de la distribution des durées d'exécution des scénarios en fonction de leur horizon de prévision. Les principales observations concernent :

- La forme regroupée du diagramme de RRM-Grid (Fig. 93b) où les extremums et les boîtes sont regroupés autour de la médiane (ligne en tirets) à la différence du diagramme de la méthode 1 (Fig. 93a) dans laquelle les boîtes sont plus larges et les extrémums étalés.
- Les écarts interquartiles (boîte) de la méthode 3 ne se superposent pas suivant l'axe de la durée alors que pour la méthode classique le traitement de la moitié des scénarios s'effectuent de manière quasi-simultanée pour les horizons inférieurs à 3h.
- Avec RRM-Grid (Fig. 93b), la plupart des prévisions d'horizon 1h (plus de 75%) sont traités et renvoyés à l'utilisateur en moins de 30 minutes et tous les autres horizons sont traités à temps.

Au travers de ces observations, la réallocation des jobs déficients implémentée dans RRM-Grid montre son efficacité. Elles confirment la rapidité plus grande de soumission des scénarios avec RRM-Grid grâce à l'utilisation des collections de jobs, corrélée à celle, déjà démontrée, de RRM-Wrapper. De plus, la séparation claire parmi les différents écarts interquartiles de chaque horizon confirme la capacité de RRM-Grid pour prendre en charge les traitements de manière ordonnée (priorité de traitements) et *a fortiori* sa faculté d'ordonnancement. En effet, il permet d'exécuter en priorité les scénarios d'horizon de prévision court dont le temps d'exécution est plus critique en situation opérationnelle. La petite taille des boites et la faible valeur des maximums illustrent la relative simultanéité de traitement des scénarios de même horizon.

L'illustration de cette efficacité est également visible sur la figure 94. Cette figure représente le détail temporel de chaque scénario requis par le prévisionniste.

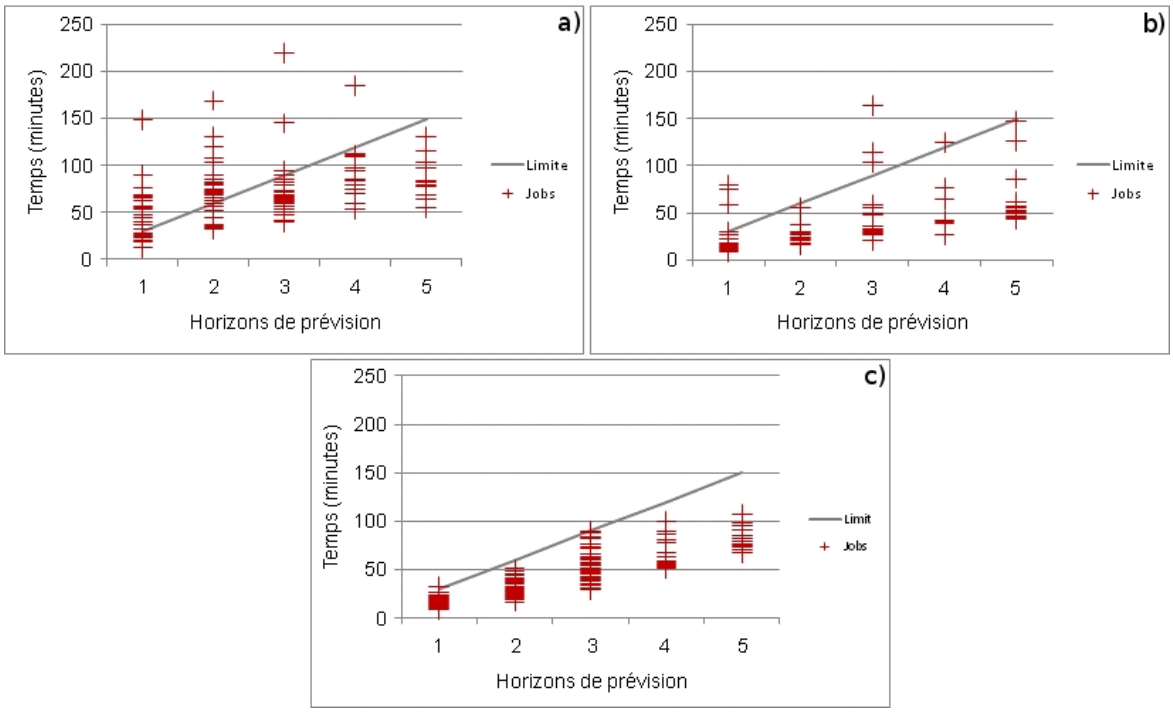


Figure 94 : Durée de l'exécution de tous les scénarios pour les 3 méthodes, usage classique (a), RRM-Wrapper (b) et RRM-Grid (c), la ligne diagonale représente le délai opérationnel pour chaque horizon

La ligne diagonale représente les seuils limites au-delà desquels les simulations sont considérées comme inutiles pour le développement de la prévision hydrologique. Il semble indéniable que RRM-Grid améliore la gestion des tâches offertes par défaut par l'intergiciel gLite et le fonctionnement de RRM-Wrapper dédiée à la gestion des données. En effet, avec RRM-Grid, tous les scénarios réussis ont été traités dans le délai opérationnel (Fig. 94c : diagonale), alors que dans le cas de la méthode 1 (Fig. 94a) et de RRM-Wrapper (Fig. 94b) des scénarios ont dépassés ces délais (respectivement 35% et 7%). Une deuxième interprétation est la confirmation du regroupement des traitements des scénarios de même horizon de prévision par RRM-Grid. Avec une utilisation simple et RRM-Wrapper, la durée de traitement des scénarios est beaucoup plus étalée. Finalement, la méthode de traitement des données d'entrée, offerte par RRM-Wrapper, confirme le poids de la phase de soumission des jobs dans la durée totale de ces derniers, en particulier quand ces données d'entrée ont une taille importante.

VII - 2.1.2.3. Montée en charge

Pour terminer, il semble intéressant d'expérimenter une simulation de plus grande ampleur composée de 210 scénarios. Cette simulation correspond à la prise en charge de plusieurs scénarios de prévision sur chaque bassin versant surveillé par le SPC-GD. De plus, cette simulation offre l'opportunité d'expérimenter l'impact du nombre de *workers* sur la performance globale de la simulation. Ainsi, les 210 scénarios sont successivement pris en charge par 40, 80, 120, 160 et 200 *workers* (Fig. 95).

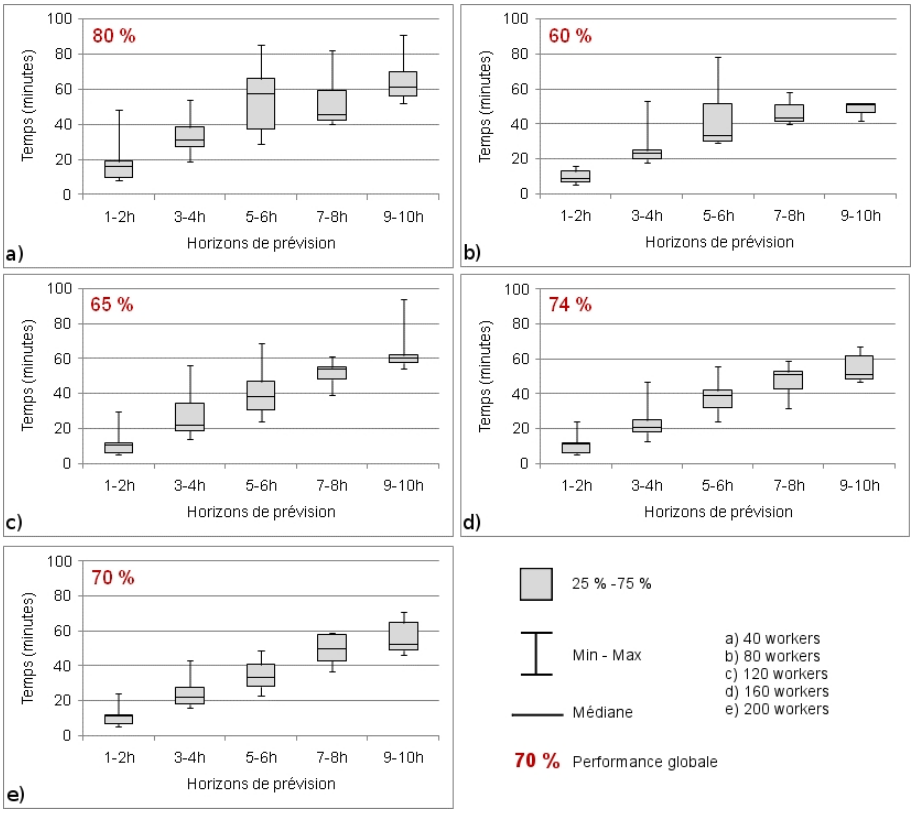


Figure 95 : Diagrammes à moustaches des temps d'exécution de 210 scénarios suivant le nombre de *workers*

Ces différents diagrammes à moustaches synthétisent les durées totales des scénarios réussis pour chaque simulation. Tout d'abord, la performance globale de l'ensemble de ces simulations avoisine 70%. Cette performance, bien qu'inférieure aux performances des simulations précédentes de RRM-Grid (Tab. 9), apparait acceptable pour assurer une prévision opérationnelle des réactions hydrologiques. D'autre part, une différence des écarts interquartiles peut être observée entre la simulation à 40 *workers* et celles à 160 et 200 *workers*. Cette différence est moins visible pour les simulations à 80 et 120 *workers.* En effet, pour ces simulations, les scénarios de délais de prévision intermédiaires (de 3 à 6h) semblent souffrir d'une plus grande irrégularité de traitement. L'augmentation du nombre de *workers* semble donc permettre une stabilisation des temps de traitement pour chaque intervalle d'horizons de prévision, du fait d'un plus grand nombre d'instances de scénarios traitées en simultané.

A partir de 80 *workers*, 75% des scénarios de prévision 1 et 2h sont traités en moins de 20 minutes. Un seul scénario de prévision 1h de la simulation à 120 *workers* a été traité au-delà des 30 minutes opérationnelles. En outre, malgré des performances moindres, la simulation à 40 *workers* contient seulement deux scénarios ayant dépassés les délais opérationnels. Cette dernière remarque, corrélée à la performance globale de cette simulation (80%), tend à opter pour un allégement du nombre de *workers*, afin de minimiser le nombre de nœuds de calcul requis et préserver un quota de ressources pour de nouvelles simulations.

VII - 2.1.3. Synthèse des capacités technologiques de RRM-Grid

Comme cela a déjà été abordé, la gestion de crise des crues à cinétique rapide est caractérisée par de fortes contraintes temporelles. D'autre part, la nécessité d'assurer un niveau élevé de fiabilité des systèmes informatiques apparait primordiale. L'ordonnanceur RRM-Grid implémenté pour accéder aux ressources de l'architecture de grille EGEE semble être adapté à l'ensemble de ces contraintes, en permettant l'exécution de nombreux scénarios de prévision sur des ressources de calculs délocalisées. De cette manière, les ressources informatiques du SPC-GD pourraient être préservées pour le fonctionnement « temps réel » et les opérations déjà existantes (cf. §.III – 1.) et primordiales à la mission d'expertise hydrologique du SPC-GD resteraient efficientes.

Comme cela a été présenté, les délais opérationnels sont mieux respectés en utilisant un tel système puisque les prévisions à très court terme (1 et 2h) ont été pour la plupart générées à temps. Dans ce sens, la gestion des priorités permet au prévisionniste d'adapter le fonctionnement de RRM-Grid dans les conditions correspondantes à la situation hydrologique en cours et de compenser la relative faiblesse de l'intergiciel gLite pour soumettre un grand nombre de simulations en simultanée. Enfin, la gestion des erreurs des services de l'intergiciel et les problèmes de compatibilité algorithmique permet d'assurer une plus grande fiabilité des traitements sur la grille EGEE.

L'ensemble de ces résultats, bien que probants, doit cependant être nuancé. En effet, l'utilisation de RRM-Grid pour cette phase expérimentale a permis de régulièrement relever des difficultés de fonctionnement limitant la capacité affichée de cet outil. Des points critiques, tels que l'allocation des *workers* par le courtier de ressources, l'exécution de *workers* qualifiables de « fantômes » bloqués au statut « running », la gestion de l'ensemble des erreurs et des faiblesses intrinsèques de la grille, due à sa volatilité. Une expérimentation à grande échelle, impliquant plusieurs phases expérimentales à différentes périodes et caractérisées par des besoins de ressources croissants, devra être mise en place pour améliorer progressivement RRM-Grid.

Enfin, les performances favorables de RRM-Grid ne doivent pas occulter la différence de représentativité existante entre ces résultats et la capacité technologique intrinsèque offerte par les ressources de calcul de l'organisation virtuelle ESR. En effet, le nombre de ces ressources (cf. §.VII - 1) offertes par la VO ESR semblent de prime abord adapté aux besoins de ressources de la prévision des crues éclair. Cependant l'effort d'interfaçage doit être poursuivi pour accéder à ces ressources avec une plus grande garantie dans le cas d'une montée en charge des traitements due un situation d'urgence émergeante.

VII - 2.2. Étude de cas de l'évènement du 22 octobre 2008

Il convient de s'intéresser dorénavant au point de vue du prévisionniste dans la construction de scénarios de prévision et la réception des hydrogrammes correspondants au travers de l'application RRM-Grid. En effet, il semble primordial de montrer l'intérêt de l'utilisation de cet outil, et par extension des ressources de calcul de la grille EGEE, à des fins opérationnelles. Ce paragraphe reflète toutefois les difficultés de prévision (des pluies et des débits) rencontrées actuellement par les services de prévision des crues.

Cette démonstration s'appuie donc sur le déroulement en temps réel d'un évènement passé. Cet évènement, ne correspond pas à une situation de crise, au sens où elle a été présentée dans le chapitre 1. Cependant, les valeurs de débit atteintes durant ces crues à cinétique rapide, représentent des seuils relativement élevés puisque le seuil de vigilance rouge a été dépassé dans la nuit du 22 octobre 2008 (Fig. 96). Cette démonstration propose donc de décrire le comportement d'un prévisionniste durant le déroulement de cet évènement. Il s'agit d'une présentation chronologique du début de l'évènement jusqu'à l'arrêt des précipitations et la diminution des débits.

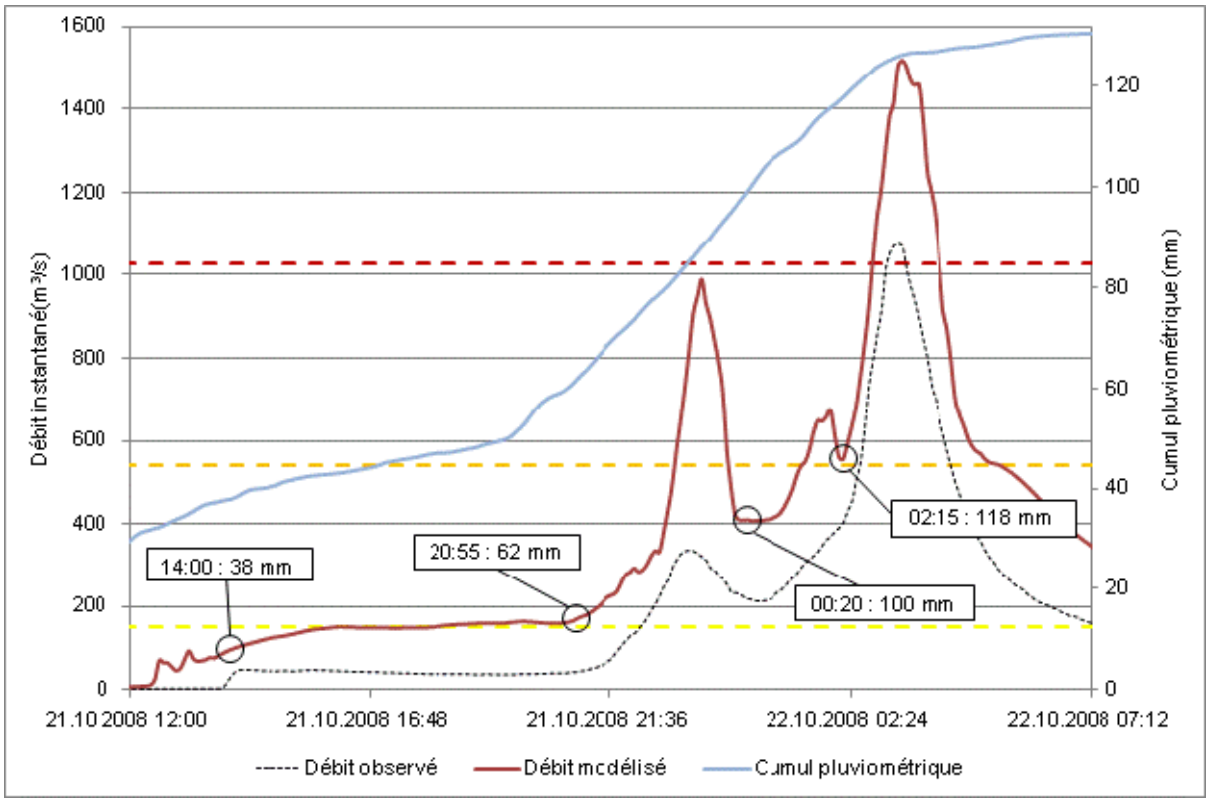


Figure 96 : Hydrogrammes observé et modélisé de l'évènement du 22 octobre 2008, et stratégie de prévision hydrologique pour la phase expérimentale (les encadrés correspondent aux cumuls pluviométriques)

Au vue de ces hydrogrammes, et des données de pluie prévue, 4 pas de temps ont été sélectionnés pour expérimenter RRM-Grid et développer une panoplie de scénarios hydrologiques. Ces simulations, au sens de groupe de scénarios de prévision, correspondent aux 4 sous-parties qui suivent. A chacun de ces pas de temps, les scénarios de prévision soumis à RRM-Grid ont l'objectif de soutenir le prévisionniste dans l'anticipation de la pointe de crue. Cette anticipation correspond à l'identification de l'heure et de l'intensité de la pointe de crue à venir. Elle doit être établie le plus longtemps à l'avance

avec le moins d'incertitude possible et transmise aux services de la Sécurité Civile et aux pouvoirs publics des zones géographiques concernées afin de garantir la sécurité des populations et de limiter les dommages (Tab. 10).

Horizon (heure)	Délai opérationnel (minutes)
1	30
2	60
3	90
n	$\frac{n \times 60}{2}$

Tableau 10 : Correspondances entre l'horizon de prévision du scénario et le délai opérationnel le concernant

Les prévisions de pluie utilisées sont basées sur la prévision CALAMAR originale et sur des prévisions à pluie constante à plus long terme, construites grâce à la différence approximative de cumuls pluviométriques entre le moment du lancement de la prévision et le pic de crue (125 mm).

Dans la suite de ce paragraphe, une simulation correspond à un ensemble de scénarios basés sur des prévisions météorologiques CALAMAR 1, 2 et 3h (CAL 1h, 2h, 3h) sur lesquelles des pondérations de -50 à +50% (avec un pas de 5%) sont appliquées (soit 11 scénarios), et sur des prévisions à pluie constante (AP) d'horizons variables[11] avec les mêmes pourcentages de pondération. Le prévisionniste demande toujours l'ordre de traitement suivant : 1h, 2h, 3h et les prévisions AP (priorité).

VII - 2.2.1. Phase n°1 : 21/10/2008 à 14h00

VII - 2.2.1.1. Situation hydrométéorologique

L'évènement météorologique du 22 novembre 2008 débute dans la matinée du 21 octobre par des pluies ponctuelles et s'intensifie en début d'après midi. Les premières réactions hydrologiques, observées à la station de mesure à l'exutoire du bassin versant d'Anduze, sont un peu avant 14h (Fig. 97). Les prévisions de Météo-France commencent à se concrétiser et des prévisions météorologiques sont transmises au SPC-GD. Dans ces conditions, le prévisionniste du SPC-GD va soumettre une première simulation lui permettant d'envisager plusieurs scénarios d'évolution de la situation hydrologique.

[11] Les horizons sont présentés dans la description de chaque phase

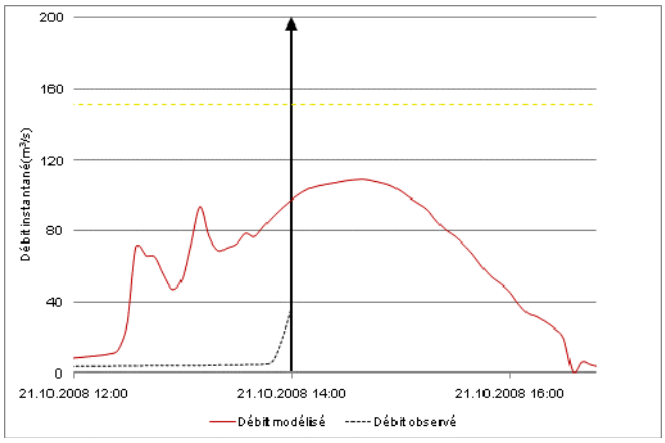


Figure 97 : Situation hydrologique du 21/10/2008 à 14h

VII - 2.2.1.2. Scénarios de prévision simulés

Les prévisions CALAMAR à 1h reçues le 21/10/2008 à 14h00 sont utilisées pour construire les prévisions météorologiques à 1h, 2h et 3h. Pour la prévision météorologique à pluie constante, un cumul de 80 mm a été injecté, équivalent à une intensité horaire de 10 mm/h au pas de temps 5 minutes pendant 12h. La simulation est soumise à RRM-Grid qui se charge de la construction du fichier global de configuration (JSON) et du lancement du *manager*.

VII - 2.2.1.3. Analyse des performances et des prévisions

À 14h10, le premier scénario basé sur la prévision CALAMAR 1h est retourné. Les réponses de RRM-Grid s'étalent jusqu'à 15h30 (Fig. 98).

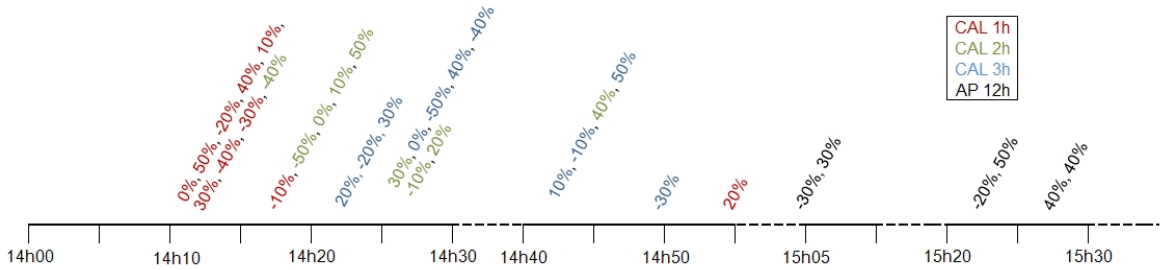


Figure 98 : Axe chronologique de la réception des scénarios de prévision de la phase n°1

Le taux de réussite de cette simulation atteint 80% (90% pour les prévisions 1h, 80% pour les 2h, 100% pour les 3h et 55% pour les prévisions AP). Un seul scénario (CAL 1h 20%) arrive après 55 minutes, soit trop tard pour être interprété par le prévisionniste. Il en résulte les hydrogrammes de la figure 99[12].

[12] Pour les prévisions CALAMAR, seuls les hydrogrammes de pondération maximale, minimale et nulle ont été affichés

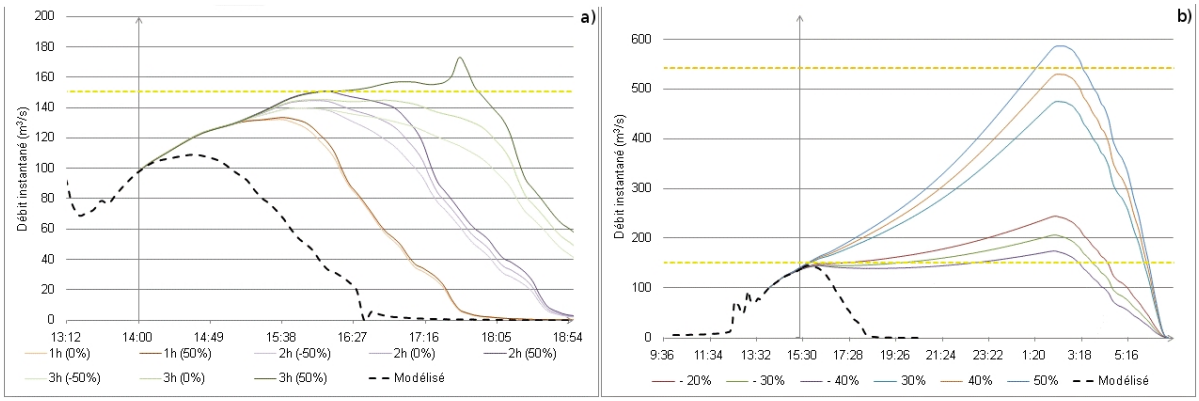


Figure 99 : Faisceaux de prévisions hydrologiques modélisées, basées sur les prévisions CALAMAR (a) et les prévisions AP à pluie constante (b) (T0 : 21/10/2008 à 14h00)

La prévision CALAMAR (Fig. 99a), en particuliers celle à 2h, permet d'entrevoir un dépassement potentiel de la vigilance jaune vers 16h. Dans la même philosophie, une pluie augmentée de 50% sur un délai de prévision de 3h tend à confirmer ce dépassement de vigilance jaune. La prévision AP à 12h indique des situations très différentes suivant le niveau de pondération appliqué. Les pondérations négatives (-20%, -30%, -40%) tendent à indiquer une relative stabilisation de la situation hydrologique. Alors que les pondérations positives indiquent un passage de la vigilance orange dans un délai de 12 à 15h.

VII - 2.2.2. Phase n°2 : 21/10/2008 à 20h55

VII - 2.2.2.1. Situation hydrométéorologique

Les précipitations continuent durant toute l'après midi du 21 octobre avec des intensités assez faibles (de l'ordre de quelques mm/h) provoquant une relative stabilisation du débit à l'exutoire d'Anduze. Peu après 18h, le modèle pluie-débit ALHTAÏR indique un dépassement de la vigilance jaune (150 m³/s) non confirmé par la station de mesure correspondante (38 m³/s) (Fig. 100). A partir de 19h30, les intensités augmentent de nouveau (de l'ordre de la dizaine de mm/h). Les prévisions de Météo-France indiquent une stabilisation de ces conditions dépressionnaires. La situation devient donc potentiellement à risque et nécessite une anticipation des réponses hydrologiques à court terme.

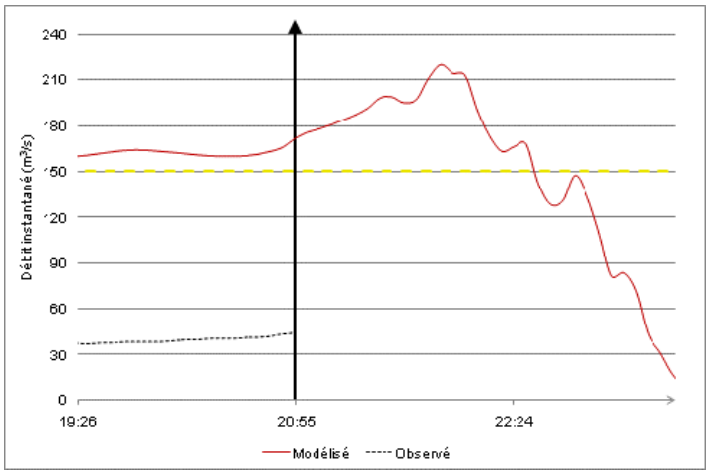


Figure 100 : Situation hydrologique du 21/10/2008 à 20h55

VII - 2.2.2.2. Scénarios de prévision simulés

Les prévisions CALAMAR à 1h reçues le 21/10/2008 à 20h55 sont utilisées pour construire les prévisions météorologiques à 1h, 2h et 3h. Pour la prévision météorologique à pluie constante, un cumul de 60 mm a été injecté, équivalent à une intensité horaire de 10 mm/h au pas de temps 5 minutes pendant 6h.

VII - 2.2.2.3. Analyse des performances et des prévisions

De la même manière qu'à 14h00, RRM-Grid est utilisé pour déporter ces scénarios de prévision sur l'architecture de grille EGEE. La seconde simulation commence à 20h55 pour s'étendre jusqu'à 22h00 (Fig. 101).

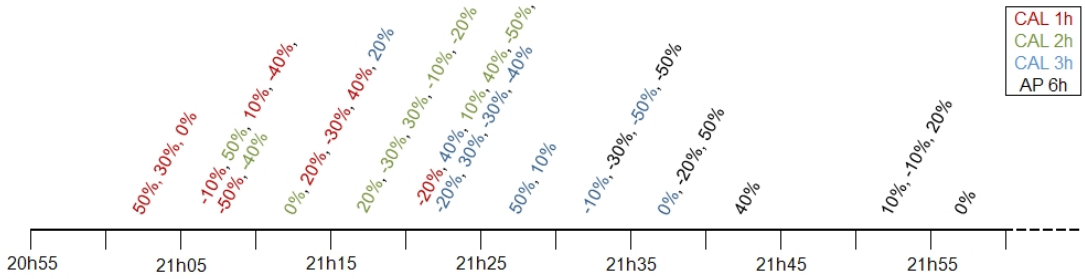


Figure 101 : Axe chronologique de la réception des scénarios de prévision de la phase n°2

Le taux de réussite de cette phase est important puisqu'il est égal à 95% (100% pour les 1h, 2h et 3h et 80% pour les prévisions AP). Cette phase est marquée par une rapidité des traitements sur la grille. En effet, après seulement 30 minutes, 70% des scénarios ont déjà été traités et tous les scénarios exécutés avec succès sont arrivés à temps pour être interprété. D'autre part, l'ordre de priorité défini au lancement de RRM-Grid semble avoir été respecté.

Il en résulte les hydrogrammes de la figure 102[13].

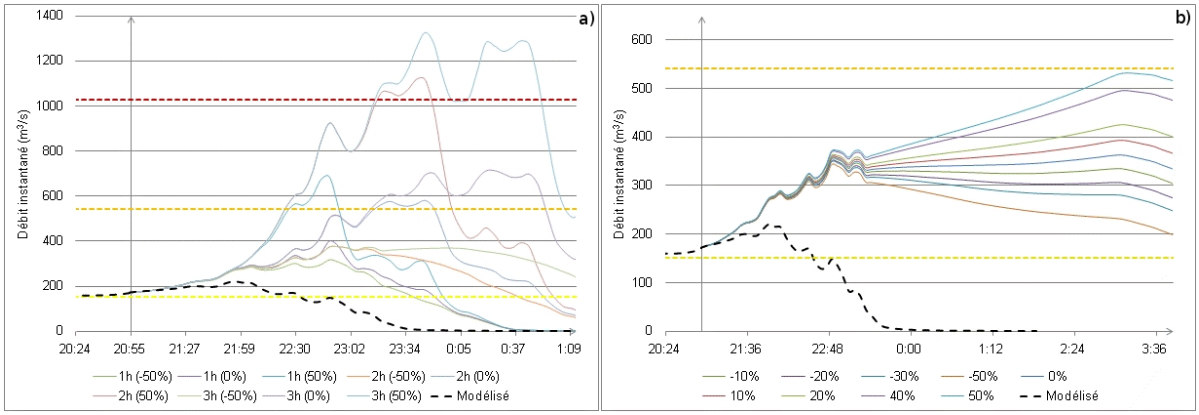


Figure 102 : Faisceaux de prévisions hydrologiques modélisées, basées sur les prévisions CALAMAR (a) et les prévisions AP à pluie constante (b) (T0 : 21/10/2008 à 20h55)

Les prévisions CALAMAR (Fig. 102a) indiquent des situations très diverses suivant la pondération et l'horizon concernés. Les horizons les plus courts avec les pondérations négatives tendent à indiquer

[13] Pour les prévisions CALAMAR, seuls les hydrogrammes de pondération maximale, minimale et nulle ont été affichés

une situation hydrologique stationnaire située entre les niveaux de vigilance jaune et orange. Les prévisions CAL 1h 50% et CAL 2h 0% et CAL 3h 0% indiquent en revanche un dépassement de la vigilance orange dans un délai de 2 à 3h. Enfin, les prévisions CAL 2h 50% et CAL 3h 50% montrent un dépassement de la vigilance rouge dans un délai inférieur à 3h. La prévision AP à plus long terme (Fig. 102b), tend à confirmer une augmentation de débit dans les 2h à venir quelque soit la pondération, sans toutefois un dépassement de la vigilance orange. Dans ces conditions, la connaissance de la pluie réellement tombée dans les minutes suivantes cette simulation peut permettre au prévisionniste d'affiner le choix du scénario le plus plausible.

VII - 2.2.3. Phase n°3 : 22/10/2008 à 00h20

VII - 2.2.3.1. Situation hydrométéorologique

Depuis la dernière simulation de prévision, les précipitations de pluies sont relativement constantes (de l'ordre de la dizaine de mm/h). Une augmentation de débit a été observée à 23h10 (338 m^3/s) confirmant le dépassement du niveau de la vigilance orange prévu lors de la première phase (prévision AP à 12h). Cependant, ce débit tend à diminuer alors que les prévisions météorologiques annoncent une consolidation des précipitations à venir (Fig. 103). Les niveaux d'infiltration du bassin versant d'Anduze, bien qu'en phase de vidange, sont vraisemblablement propices à une reprise des réponses hydrologiques. Il convient donc d'évaluer ce risque à court terme.

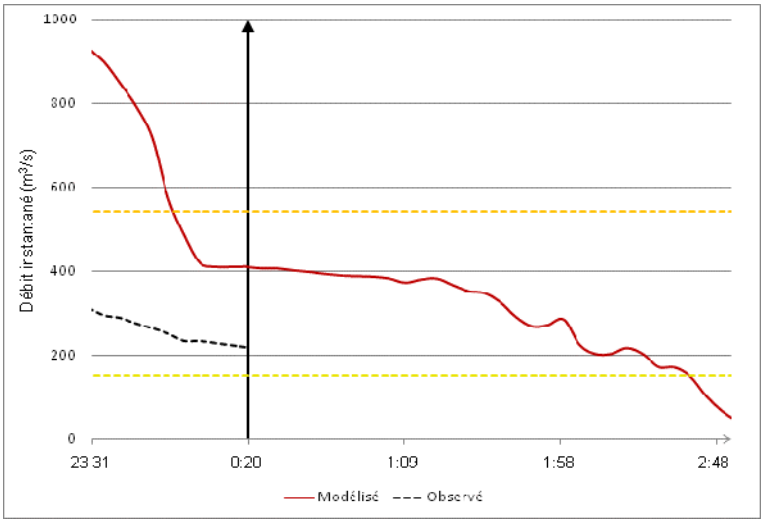


Figure 103 : Situation hydrologique du 22/10/2008 à 00h20

VII - 2.2.3.2. Scénarios de prévision simulés

Les prévisions CALAMAR à 1h reçues le 22/10/2008 à 00h20 sont utilisées pour construire les prévisions météorologiques à 1h, 2h et 3h. Pour la prévision météorologique à pluie constante, un cumul de 25 mm a été injecté, équivalent à une intensité horaire de 8 mm/h au pas de temps 5 minutes pendant 3h.

VII - 2.2.3.3. Analyse des performances et des prévisions

Ces scénarios sont injectés dans RRM-Grid à 00h20. Les réponses de l'outil s'étalent de 00h19 à 1h06 (Fig. 104). Cependant, ses performances sont en très net recul par rapport aux deux phases précédentes, puisque seulement 50% des scénarios ont été traités. Cette observation illustre la variabilité de la disponibilité des ressources de la grille ainsi que des performances de RRM-Grid.

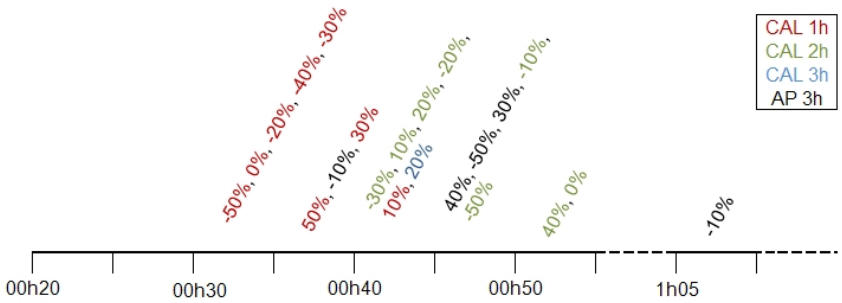


Figure 104 : Axe chronologique de la réception des scénarios de prévision de la phase n°3

Malgré tout, les scénarios et les hydrogrammes correspondants exécutés avec succès (Fig. 105) sont tous retournés dans les délais opérationnels permettant une bonne appréhension de la situation hydrologique à venir.

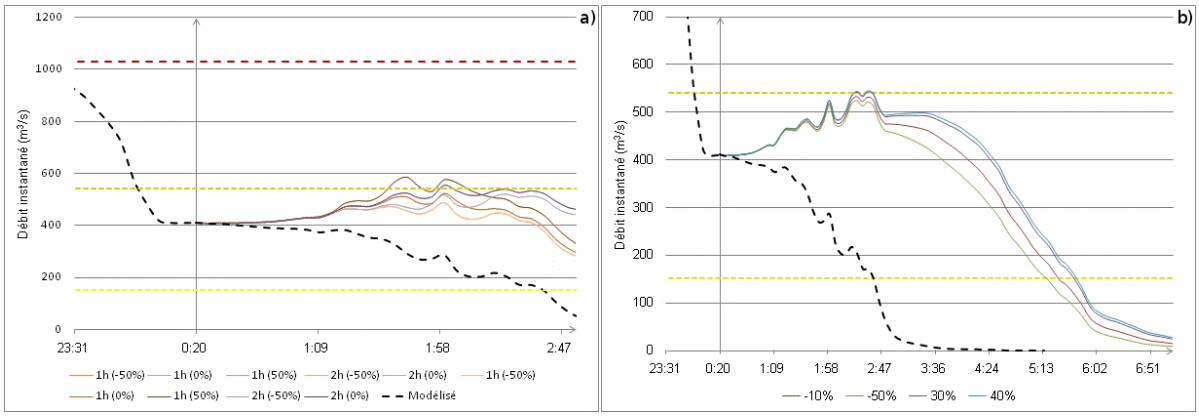


Figure 105 : Faisceaux de prévisions hydrologiques modélisées, basées sur les prévisions CALAMAR (a) et les prévisions AP à pluie constante (b) (T0 : 22/10/2008 à 00h20)

La différence entre les différents scénarios est peu marquée et la plupart des hydrogrammes résultants de cette simulation indiquent un retour à la vigilance orange dans un délai de 2 à 3h. L'horizon de la prévision AP (Fig. 105b), égal à 3h, ne permet pas d'envisager l'évolution de la situation sur un délai plus long alors que la prévision CALAMAR (Fig. 105a) indique une stabilisation de la situation hydrologique proche de la vigilance orange. Cette situation hydrométéorologique à risque astreint le SPC-GD à une attention continue sur ses cours d'eau réglementaire.

VII - 2.2.4. Phase n°4 : 22/10/2008 à 02h15

VII - 2.2.4.1. Situation hydrométéorologique

Le dépassement des seuils de vigilance orange, pressenti durant la phase précédente, ont été confirmé par le modèle ALHTAÏR « temps-réel », alors que la situation réellement observée est toujours si-

tuée à des valeurs inférieures mais tendant à atteindre ces seuils dans des délais très courts (Fig. 106). La pluie se maintient à des intensités proches de 10 mm/h et cette augmentation de débit laisse craindre un dépassement des seuils de vigilance rouge. D'autant plus, que l'hydrogramme modélisé par ALHTAÏR « temps réel », indique ce dépassement dans un délai d'une heure par vidange du bassin versant (pluie nulle). De plus, la dernière prévision CALAMAR réceptionnée indique des intensités pluviométriques maximales pour cet évènement (10 mm/h).

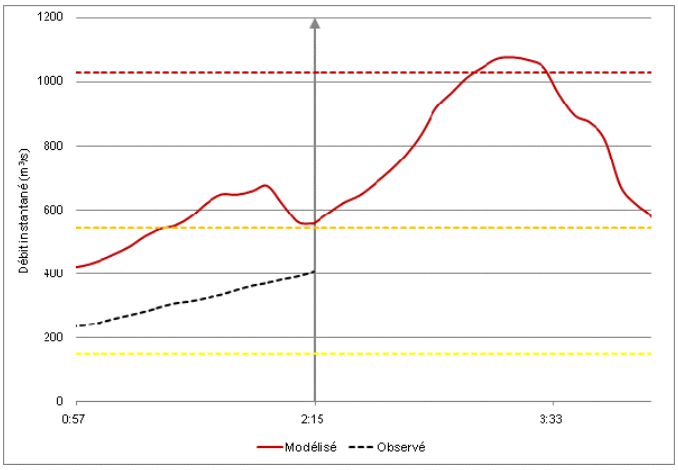


Figure 106 : Situation hydrologique du 21/10/2008 à 02h15

VII - 2.2.4.2. Scénarios de prévision simulés

Les prévisions CALAMAR à 1h reçues le 22/10/2008 à 02h15 sont utilisées pour construire les prévisions météorologiques à 1h, 2h et 3h. Pour la prévision météorologique à pluie constante, un cumul de 8 mm a été injecté, équivalent à une intensité horaire de 8 mm/h au pas de temps 5 minutes pendant 1h.

VII - 2.2.4.3. Analyse des performances et des prévisions

Dans la même philosophie que les 3 précédentes phases, 44 scénarios de prévision ont été soumis à RRM-Grid à partir de 2h15. Leur retour au sein du système d'information du SPC-GD commence à partir de 2h25 et se termine un peu avant 3h du matin (Fig. 107). Une observation intéressante concerne les scénarios de prévision basés sur la prévision AP 1h (noirs) dont les temps d'exécution sont proches de ceux de la prévision CAL 1h (rouge) et précédent les scénarios CAL 2h et CAL 3h dont la priorité était pourtant plus forte. Ainsi, le temps d'exécution de l'algorithme de modélisation (dont la durée dépend de l'horizon) influence le temps global de traitement du scénario dans le cas d'horizons de prévision courts.

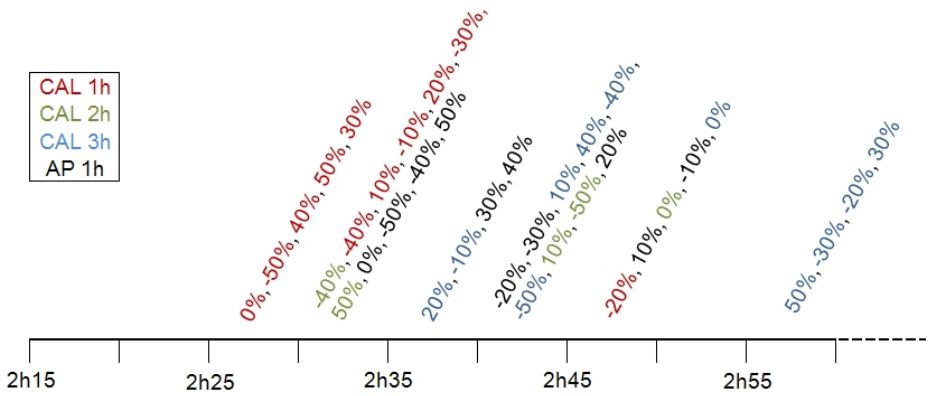


Figure 107 : Axe chronologique de la réception des scénarios de prévision de la phase n°4

Le taux de réussite de cette simulation est égal à 86% (100% pour les prévisions CAL 1h, 45% pour les CAL 2h, 100% pour les 3h et 100% pour les prévisions AP). Comme pour les simulations précédentes, les scénarios réussis sont tous arrivés à temps pour être interprété par le prévisionniste du SPC-GD.

Les prévisions hydrologiques correspondantes confirment le dépassement du seuil de vigilance rouge dans un délai de 1 à 2h (Fig. 108). En effet, que ce soit les prévisions avec des pondérations négatives ou celles avec des pondérations positives, le niveau de vigilance rouge est dépassé.

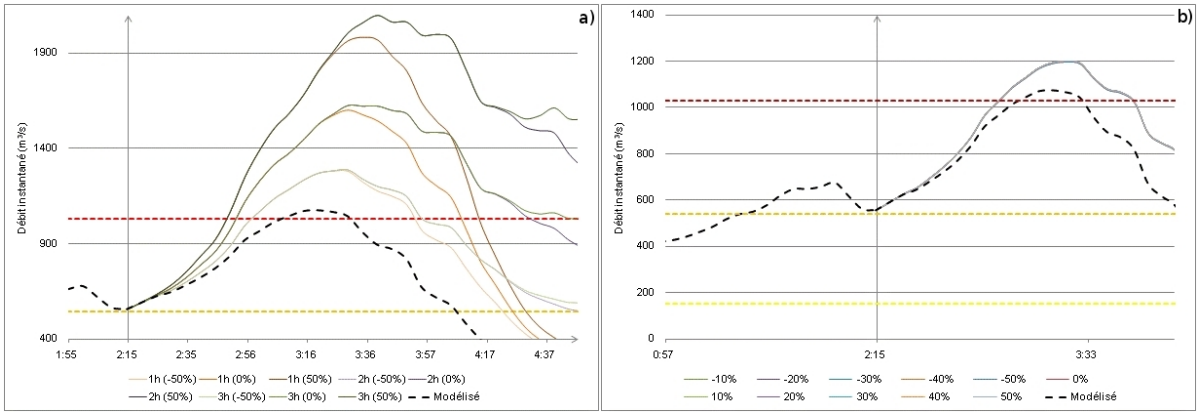


Figure 108 : Faisceaux de prévisions hydrologiques modélisées, basées sur les prévisions CALAMAR (a) et les prévisions AP à pluie constante (b) (T0 : 22/10/2008 à 02h15)

Les prévisions avec une pluie prévue augmentée de 50% (Fig. 108a) tendent à prévoir un doublement de la valeur de débit dans les 2h à venir. Cette situation correspond sur le terrain à l'inondation des habitations de la commune d'Anduze situées à proximité du cours d'eau du Gardon d'Anduze. Finalement, les prévisions basées sur la donnée de pluie AP ne permettent pas de discerner plusieurs situations hydrologiques potentielles.

VII - 2.2.5. Synthèse sur la capacité de prévision

L'ensemble de ce paragraphe a permis de démontrer l'intérêt d'utilisation de l'outil RRM-Grid et par extension du système d'aide à la décision G-ALHTAÏR dans une situation hydrologique menaçante. La forte réactivité de cet outil permet d'obtenir un faisceau de prévisions hydrologiques dans des délais courts en adéquation avec les contraintes temporelles spécifiques à l'expertise hydrométéorologique des crues à cinétique rapide.

Comme cela a été abordé précédemment, la validation des débits prévus semble incertaine dans le cadre de cette étude, du fait du peu d'intérêt que représentent les « prévisions » de pluie, qui peuvent, dans un contexte opérationnel, être injectées dans le modèle ALHTAÏR. Cependant, l'analyse au cas par cas des hydrogrammes résultants (Fig. 99, Fig. 102, Fig. 105, Fig. 108) permet de montrer toutefois l'intérêt de certaines de ces prévisions hydrologiques (Fig. 109).

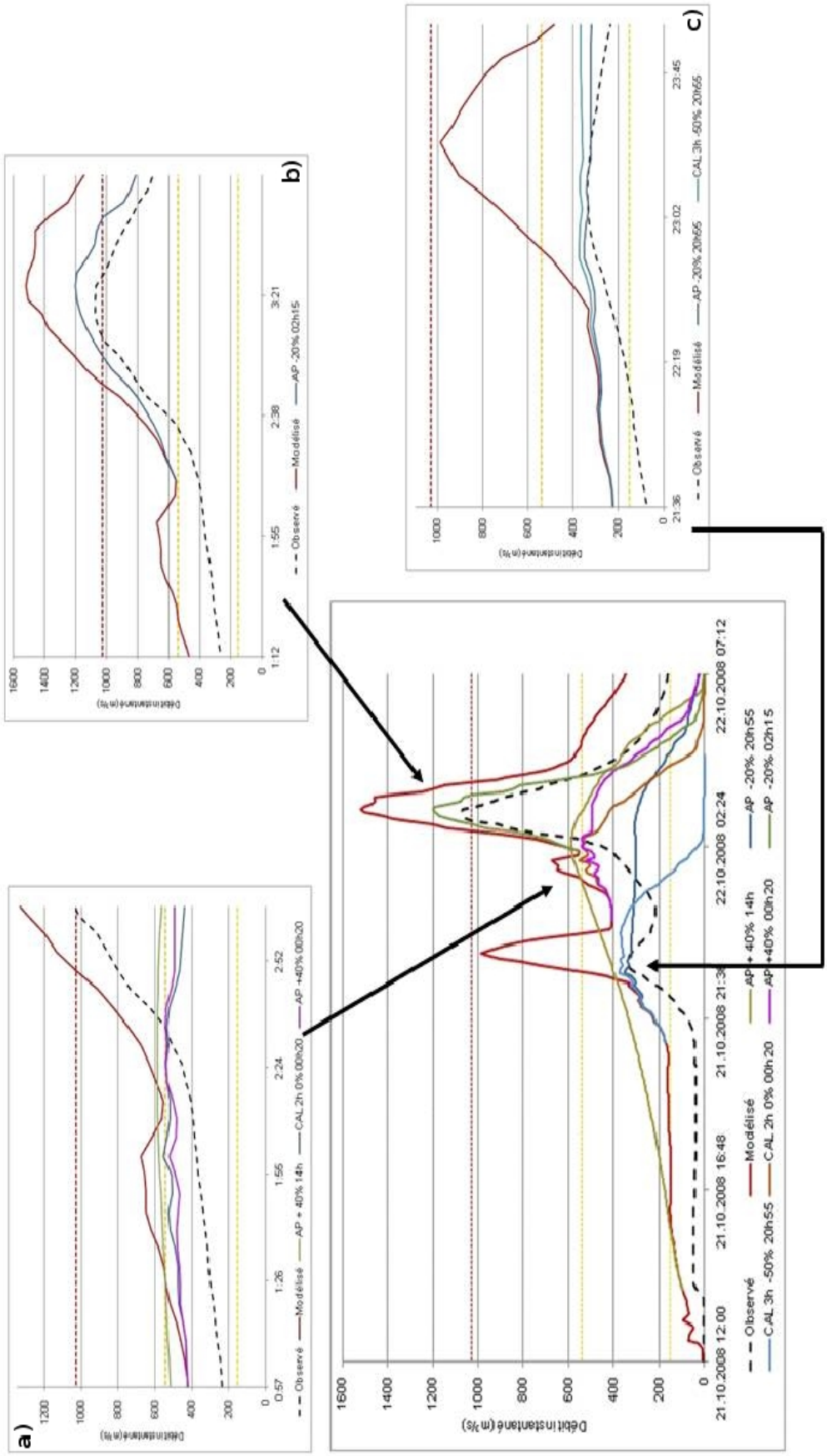


Figure 109 : Bilan des prévisions hydrologiques les plus performantes de l'évènement du 22 octobre 2008

Si l'on considère que la prévision des crues à cinétique consiste en partie à établir l'heure et l'intensité de la pointe de crues et dans une moindre mesure le dépassement de seuils de vigilance, trois situations hydrologiques auraient pu être prévues durant la phase d'expertise hydrologique présentée précédemment :

- Le premier pic du 21/10/2008 de 23h10 est bien anticipé par les prévisions *CAL 3h -50%* et surtout *AP 6h -20%* lancées à 20h55, respectivement réceptionnées à 21h30 et 21h35. En effet, le pic hydrologique observé a une valeur de 338 m^3/s, alors que les prévisions développées grâce à RRM-Grid prévoyaient un débit égal à 340 m^3/s pour la prévision AP et 370 m^3/s pour la prévision CALAMAR (Fig. 109c) au même moment.

- Le dépassement du seuil de la vigilance orange (540 m^3/s) à 02h35 par les prévisions *AP 3h 40%*, *CAL 2h 0%* lancées à 00h20 et respectivement réceptionnées à 00h45 et 00h50, et surtout par la prévision *AP 12h 40%*, lancée à 14h00 et réceptionnée à 15h30 dans l'après-midi du 21 octobre. Ces prévisions permettent d'anticiper à plus ou moins une demi-heure ce dépassement (Fig. 109a).

- Le dépassement du seuil de la vigilance rouge (1029 m^3/s) à 02h50 et le pic principal de cet évènement (1518 m^3/s) à 03h25, par la prévision *AP 1h -20%* soumise à 02h15 et réceptionnée à 02h40. Pour le pic, la différence entre la valeur de débit prévue et observée est égale à environ 300 m^3/s.

Il est clair que les délais offerts au prévisionniste pour diffuser cette information critique sont très courts dans la plupart des cas. En effet, il équivaut à environ 1h30 pour le premier pic, et le dépassement du seuil de vigilance orange, exceptée pour la prévision *AP 12h 40%* qui aurait permis d'anticiper la situation 11h à l'avance, et à 45 minutes pour la pointe de crue principale dans le cas de la prévision *AP 1h -20%*.

Le cas particuliers de dépassement du seuil de vigilance rouge, n'aurait sans doute pas pu être anticipé par l'outil. La criticité temporelle des prévisions sur les crues à cinétique rapide justifie d'autant plus le recours à des technologies innovantes telles que celles présentées dans cette recherche, au travers de G-ALHTAÏR (en considérant que G-ALHTAÏR encapsule l'ordonnanceur RRM-Grid), pour faciliter et accélérer la transmission de ces informations aux services de la Sécurité Civile et aux pouvoirs publics en charge de la sécurité des citoyens.

Finalement, d'un point de vue hydrométéorologique, la flexibilité d'utilisation de RRM-Grid permet d'offrir au prévisionniste des moyens technologiques propices au développement de multiples scénarios d'évolution hydrologique, permettant de compenser la forte incertitude existante sur la prévision météorologique à court terme et à grande échelle, dans des délais cours relatifs à la gestion des crues à cinétique rapide. Cette étude de cas montre toute l'importance de développer des prévisions de pluie efficaces et probabilisées, prévisions qui peuvent être exploitées de manière opérationnelle grâce à RRM-Grid et par extension G-ALHTAÏR. Finalement, Il est important de souligner le rôle prépondérant du prévisionniste dans l'anticipation des réponses hydrologiques des cours dont il est en charge. En

effet, sa longue expérience sur ces cours d'eau et les différents évènements qu'il a suivi, lui permettent :

- de développer les scénarios de prévision les mieux adaptés à la situation hydrométéorologique en cours et prenant en compte les incertitudes bien identifiées de la pluie et du modèle pluie-débit qu'il utilise,
- d'interpréter au mieux les faisceaux des hydrogrammes retournés par RRM-Grid susceptibles de lui fournir l'intervalle d'incertitude sur la réponse hydrologique à venir.

Finalement, les débits provenant des stations de mesures peuvent lui permettre de réajuster les scénarios de prévision et l'interprétation qu'il en fera durant sa démarche d'expertise hydrologique.

SYNTHESE

L'outil RRM-Grid a montré son efficacité en comparaison avec une utilisation classique de la grille, au travers des fonctionnalités de bas niveau de l'intergiciel gLite. En effet, en situation opérationnelle, il n'apparait pas pertinent d'accéder aux ressources de grille par la méthodologie offerte par défaut.

D'autre part, d'un point de vue calculatoire, RRM-Grid accélère drastiquement la phase de soumission des scénarios de prévision. Il a été démontré que l'effet du nombre de jobs, s'il influence cette phase, n'apparait pas véritablement comme un facteur limitant. Dans ce sens, la simulation de 210 scénarios à montrer une efficacité correspondante aux contraintes opérationnelle. RRM-Grid accélère enfin le traitement de ces scénarios en les exécutant de manière ordonnancée (horizon de prévision) et garantie un niveau de fiabilité en adéquation avec les contraintes de sécurité de la gestion de crise.

La disponibilité des ressources de calcul apparait cependant comme un des points clefs de l'opérationnalité requise dans cette étude. L'utilisation régulière des ressources offertes par la VO ESR a montré une forte variabilité de cette disponibilité d'une période à l'autre et au sein même d'une phase de simulation. Ainsi, une expérimentation de plus longue durée, et à des niveaux de disponibilité différents doit être menée pour parfaire le fonctionnement de cet ordonnanceur, et plus largement l'utilisation de la grille en situation de crise.

Finalement, l'intégration de cet outil dans la plateforme G-ALHTAÏR doit permettre au prévisionniste d'effectuer sa mission dans un environnement intégré, lui facilitant l'accès à l'ensemble des données hydrométéorologiques dont il dispose, et lui permettant :

- de définir les scénarios en adéquation avec la situation hydrométéorologique
- de transmettre rapidement cette information aux partenaires de la gestion de crise

Dans ce sens, l'interprétation des différents scénarios de prévision et son expérience doivent lui permettre de construire des cartes thématiques adaptés aux compétences des services de la Sécurité Civile et des pouvoirs publics, alors que les capacités offertes par l'interface de cartographie en ligne, et en particulier le WMS (Web Map Service), lui offre un moyen efficace et interopérable de les diffuser dans les plus brefs délais.

Du point de vue de l'hydrologie opérationnelle, RRM-Grid permet au prévisionniste d'évaluer, dans des délais opérationnels, la réponse hydrologique susceptible de menacer les territoires drainés par les cours d'eau concernés, sous réserve d'amélioration et de mise à disposition de prévisions de pluie et de modèles hydrologiques associés. Il est cependant important de noter, le rôle prépondérant du prévisionniste dans cette démarche. En effet, l'expertise hydrologique qu'il transmet aux services de Sécurité Civile doit estomper l'incertitude pouvant exister sur ces prévisions, afin d'éviter une surévaluation du niveau de vigilance et une réactivité disproportionnée de ces services.

Chapitre 8

Perspectives opérationnelles et scientifiques de G-ALHTAÏR

9.1. G-ALHTAÏR : un système spatial d'aide à la décision en devenir

9.2. Enrichissement de la dimension hydrologique

Le socle technologique cohérent que représente G-ALHTAÏR offre une opportunité nouvelle de répondre aux besoins récurrents de l'hydrologie opérationnelle. Les perspectives présentées dans ce chapitre doivent inciter les communautés scientifiques et opérationnels correspondantes à exploité cette solution pour soutenir leurs missions respectives.

INTRODUCTION

Le chapitre 7 s'est attaché à démontrer les capacités calculatoires de RRM-Grid. Ainsi, cette recherche a initié des développements susceptibles d'influencer à terme les méthodes et les outils de services opérationnels dédiés à l'expertise hydrométéorologique, tels que le SPC-GD. L'origine de cette recherche s'apparente à une mise en commun et une confrontation de fondements scientifiques foncièrement différents. Dans ce sens, l'éventail des besoins relatifs à l'expertise hydrométéorologique (cf. §.), se veut volontairement plus vaste que les besoins considérés dans la méthodologie de la partie 3. Cependant, en considérant ceux plus spécifiques à la prévision hydrologique, cette recherche a établi un premier socle expérimental sur lequel de nouvelles recherches et innovations peuvent dorénavant s'appuyer.

En effet, le système G-ALHTAÏR doit être appréhendé comme un prototype expérimental ouvert et adaptable à de nouvelles compétences opérationnelles et scientifiques. Ses fondements technologiques, basés sur une ontologie standardisée, lui permet *a priori* d'intégrer rapidement de nouvelles méthodes et composants logiciels. En effet, la voie de la standardisation ouverte par les consortiums de l'OGC et de l'OGF et leurs efforts pour susciter l'intérêt croissant de la part de communautés scientifiques et opérationnelles utilisatrices, tendent à confirmer la viabilité de cette plateforme. Elle encourage ainsi l'innovation nécessaire au perfectionnement de la chaîne d'alerte hydrométéorologique de la gestion de crise « inondation ».

Deux axes de recherche principaux composent la présentation des perspectives qui suit. D'une part, les développements technologiques, qu'il est possible de qualifier de « prototypaires », nécessitent vraisemblablement plusieurs améliorations au vue des résultats présentés dans le chapitre précédent. D'autre part, l'approche multi-modèles, l'étalonnage des modèles, l'assimilation des données, ainsi que la gestion des données collectées, identifiés dans le chapitre 3 en tant que besoins de l'expertise hydrologique en situation de crise, peuvent tirer profit des capacités nouvelles offertes par la technologie grille et le système spatial d'aide à la décision hydrologique G-ALHTAÏR.

VIII - 1. G-ALHTAÏR : un système spatial d'aide à la décision en devenir

Les résultats présentés dans le chapitre 7 sont centrés sur le composant RRM-Grid qui offre à G-ALHTAÏR sa capacité calculatoire. A l'inverse, G-ALHTAÏR et sa forte orientation « web » semble pouvoir fournir à RRM-Grid un niveau de généralisation en adéquation avec les contraintes de partage de la gestion de crise. Il s'agit donc, dans un premier temps, de recenser les points d'amélioration intrinsèque à RRM-Grid et dans un second de développer une proposition de rapprochement entre cet outil et les services web de G-ALHTAÏR.

VIII - 1.1. L'optimisation des traitements de grille

VIII - 1.1.1. Ordonnancement, accès aux ressources et rapidité d'exécution

Le composant RRM-Grid a confirmé sa capacité à communiquer avec les services de l'intergiciel gLite et à allouer des tâches informatiques de manière intensive, fiable et sécurisé. Les différentes phases expérimentales ont fait cependant émergées des baisses de performance due à des instabilités fonctionnelles des services de gLite difficilement prévisibles et à des limites technologiques intrinsèques.

À l'origine *Pilot jobs,* canevas fondamental de l'ordonnanceur RRM-Grid, a été conçu dans l'optique de soutenir la modélisation paramétrique sur des temps d'exécution longs et nécessitant moins de tâches élémentaires que celles requises dans le cadre de cette étude. Ainsi, la multiplication du nombre de scénarios de prévision et les délais de réponse nécessairement courts ont fait émerger plusieurs difficultés liées à la gestion des données et à la synchronisation des traitements. En effet, la concurrence d'accès (gestion des données et des traitements) est une question essentielle de l'informatique distribuée, au même titre que les problématiques relatives à la parallélisation.

VIII - 1.1.1.1. La gestion des données

À la vue des expérimentations du chapitre précédent et de manière à assurer une fiabilité plus régulière de son fonctionnement, il semble important d'apporter plusieurs modifications dans la gestion des données d'entrée et des fichiers de configurations des scénarios :

- une meilleure gestion des réplicas de données d'entrée en fonction du nombre de scénarios à simuler. En effet, ces données d'entrée identiques à tous les scénarios sont requises en simultanée par plusieurs de ces scénarios (notion de concurrence),

- l'installation d'un élément de stockage (SE) référencé dans l'organisation virtuelle de l'utilisateur, et donc accessible par l'ensemble des ressources de cette VO, doit permettre une gestion plus efficace des fichiers de configuration de chaque scénario, en limitant les opérations de gestion de données RFIO. Elle implique cependant une définition claire des droits d'accès et d'utilisation de ce SE pour les autres membres de la VO ESR, *a priori* exclus de la problématique de gestion de crise « inondation ».

VIII - 1.1.1.2. La gestion des traitements

La forte volatilité des ressources de grille, caractérisée par des changements d'état de fonctionnement de ces éléments imposent le développement de stratégies fonctionnelles spécifiques :

- L'augmentation du nombre d'instances de modélisation, allouées à un même scénario, peut permettre d'accroitre les chances de réussite de la prévision correspondante. En effet, le même scénario de prévision est exécuté plusieurs fois en simultanée.

- La concurrence intra-simulation peut être limitée en décomposant ces simulations en sous-ensembles contrôlés par différents *managers*. À titre d'exemple, chaque simulation (une instance de *manager*) peut être affectée au traitement de chaque bassin versant nécessitant une surveillance de ses réponses hydrologiques.

- La notion de listes blanche et noire (cf. §.VI – 2.3.2.) permet à RRM-Grid d'orienter ses opération d'ordonnancement, en ayant une connaissance régulière de l'état des éléments de calcul. Cette capacité manque cependant de capacité d'actualisation et demande un laps de temps qu'une situation de crise ne peut pas toujours permettre. Ainsi, l'une des solutions envisagées est d'utiliser les *workers*, en cours d'exécution sur un nœud de calcul, pour informer en continu le manager des performances de ce nœud.

Finalement, la problématique de communication entre les services de bas niveau de l'intergiciel de grille et les systèmes d'information utilisateurs est un axe de recherche récurrent abordé par un grand nombre d'utilisateurs. À terme, des solutions opérationnelles et adaptées à différentes contraintes fonctionnelles devraient émerger et faciliter la résolution des problèmes identifiés dans l'expérimentation de RRM-Grid (cf. §.VII – 2.).

VIII - 1.1.1.3. La question de la parallélisation

Dans le cadre de la méthodologie développée autour du système spatial d'aide à la décision G-ALHTAÏR, l'approche calculatoire s'est principalement orientée vers l'utilisation d'études paramétriques, pour lesquelles la technologie grille est fortement adaptée. La multiplication d'instances de modélisation fonctionnellement indépendantes est rendue possible à des niveaux de performance pouvant envisager son adoption dans un contexte opérationnel. D'autre part, le modèle ALHTAÏR, relativement simple, permet d'envisager des temps d'exécution courts et de garantir à RRM-Grid des réponses rapides. Cependant, la modélisation hydrologique et plus largement le traitement sur des données géospatiales requiert généralement des capacités de parallélisation de contrôle ou de données permettant de réduire fortement les temps de réponse pour être envisagée en situation opérationnelle

Ainsi, RRM-Grid, dans la manière dont il a été conçu, semble pouvoir offrir un socle technologique pour développer une véritable capacité de parallélisation, en particulier de données (cf. §.IV – 2.1.1.). Ainsi, le gain de performance du modèle ALHTAÏR[1] peut être envisagé par un parallélisme de données en

[1] Son fonctionnement en mailles indépendantes s'adapte clairement à la parallélisation de données. Certains modèles ou fonctions sont caractérisés par une dépendance géographique qui complexifie ou interdit cette approche.

décomposant le bassin versant en sous-ensembles sur lesquels un module de production et de transfert sont exécutés (MARQUES et THIERION, 2008) (cf. ANNEXE). Ainsi, sur l'interface utilisateur, un programme « maître » capable d'interagir avec le *manager* de RRM-Grid est en charge de la préparation des sous-ensembles de données d'entrée (bassin versant, conditions initiales et pluie) et des fichiers de configuration alloués à chaque sous-tâche. À la réception des hydrogrammes élémentaires modélisés grâce à chacune des sous-tâches, le programme « maître » les somme de manière à reconstruire l'hydrogramme global.

Cette approche impose donc une grande fiabilité des traitements de RRM-Grid afin d'assurer l'obtention de tous les hydrogrammes élémentaires nécessaires. D'autre part, une expérimentation doit être menée pour ajuster le traitement en parallèle susceptible peut permettre un gain de temps. En effet, ce gain de performance est relatif à la question de granularité, c'est-à-dire au ratio entre le temps de communication (par ex. la soumission des jobs) et la durée effective de chaque tâche élémentaire. Il dépend donc de :

- l'horizon de prévision,
- de la taille du bassin versant et de la donné de pluie,
- de la résolution géographique des données.

VIII - 1.1.2. Une meilleure visibilité des ressources de grille

L'une des difficultés initiales dans l'utilisation de la grille est relative à l'appréhension précise des ressources mises à disposition par les différents sites. Des fonctionnalités propres aux services gLite, le système d'information BDII et des sites Internet dédiés permettent d'avoir une vision générale de l'organisation virtuelle à la quelle l'utilisateur est rattachée. Cependant, la véracité des ces informations n'est pas complètement en accord avec les besoins de transparence propre à la gestion de crise.

Plusieurs solutions sont envisageables pour améliorer la visibilité des ressources informatiques de la grille. D'une part, un système de cartographie des sites de calcul peut simplifier l'analyse de ces ressources, d'autre part des méthodologies plus conséquentes, en termes de développement informatique, peuvent permettre d'effectuer des prévisions sur la disponibilité de ces ressources. Enfin, en amont de ces approches technologiques, une politique de fidélisation et de réservation des ressources devrait être menée par les services opérationnels.

VIII - 1.1.2.1. Cartographie des ressources en temps réel

Des efforts sont actuellement menés par les acteurs nationaux et internationaux de la communauté « grille » pour parvenir à une plus grande transparence des ressources offertes par chaque site de calcul qui compose l'architecture de grille. A titre d'exemple, le projet GStat 2.0[2] représente une initiative susceptible de fournir en temps réel la disponibilité et l'état des ressources de grille (FIELD et *al*., 2009). Dans ce sens et dans le contexte de gestion de crise, l'appréhension précise de la disponibilité de ces ressources apparait indispensable. En effet, un prévisionniste qui souhaite exécuter

[2] http://gstat-prod.cern.ch/

ces expérimentions hydrologiques, dans les meilleurs conditions, peut tirer profit d'une telle information, par exemple au travers d'une interface de cartographie en ligne, pour sélectionner les éléments de calcul susceptibles d'accepter ses traitements dans les meilleurs délais.

VIII - 1.1.2.2. Prévision et indicateurs de performance

En parallèle de cette approche, plusieurs initiatives se sont intéressées à la prévision des performances d'exécution et de disponibilité des ressources de calcul. En effet, un système de prédiction, basé sur une analyse multicritère prenant en compte tous les paramètres agissant sur la performance d'exécution d'un job de grille (machines, réseaux, systèmes de stockage, etc.), peut permettre de sélectionner les éléments de calcul capables d'exécuter les traitements informatiques suivant les contraintes de fonctionnement de l'utilisateur final (ROOD et LEWIS, 2009). Le suivi en continu des performances de chaque ressource de la grille peut permettre d'établir une base statistique capable de renseigner en temps réel l'outil RRM-Grid et de guider sa démarche d'ordonnancement. Il est possible de citer à ce titre la méthodologie développée au travers du « Network Weather Service » offrant des capacités de prédiction de la disponibilité des ressources dans une infrastructure de grille (WOLSKI, 2003).

VIII - 1.1.2.3. Réservation et/ou préemption des ressources

Le fonctionnement en situation de crise de services opérationnels, tels que le SPC-GD, peut être qualifié d'intermittent voire occasionnel. Les évènements nécessitant le recours à des ressources de calcul supplémentaires, de la manière dont il a été décrit dans l'ensemble de ce manuscrit sont rares et de courtes durées. La fréquence des évènements catastrophiques liés aux inondations et aux crues à cinétique rapide est relativement faible. A titre d'exemple, depuis 1999, il est possible de référencer moins de 10 évènements de ce type (ANTOINE et *al*., 2001). Dans ce contexte, et en considérant que l'utilité de la grille est avant tout dédié à l'expertise hydrométéorologique en temps réel, des solutions de réservation peuvent être envisagées. *A contrario*, cette utilisation modérée et ponctuelle des ressources de grille freine certainement la fidélisation de ces services opérationnels auprès des fournisseurs de ressources. Cette problématique apparait donc essentielle à l'adoption concrète de la technologie grille par les services opérationnels de la Sécurité Civile.

Dans le contexte expérimental de cette thèse, centré sur l'architecture de grille EGEE, la réservation préalable de ressources n'est pas explicitement offerte. Elle a n'a donc pu être véritablement approfondie et s'est appuyé sur le postulat qu'une architecture de grille fournie des ressources à la demande. Cependant, la variabilité des performances (cf. §.VII – 2.1.3.), implique une nouvelle considération de cette garantie de disponibilité. AL-ALI et *al*. (2004) synthétise cette problématique suivant trois besoins principaux :

- des ressources aux capacités spécifiques,
- la disponibilité des ressources pour une période donnée
- un accord (contrat) indiquant le niveau d'implication de ces ressources

Des efforts de gouvernance sont donc à mener pour clairement répondre à ces besoins. La présente recherche et les résultats prometteurs qu'elle suscite, doit permettre de motiver la négociation et la mise en place de contrats d'utilisation avec certains fournisseurs de ressources de l'organisation virtuelle à laquelle les services opérationnels de la Sécurité Civile sont rattachés. Dans ce sens, l'innovation continue initiée autour de l'architecture de grille EGEE, et la mise en production progressive des sites de calcul au travers de l'initiative européenne EGI (European Grid Initiative), devrait assurer progressivement une véritable capacité d'offre de ressources informatiques à la demande et clairement dédiées à des fonctions opérationnelles prédéfinies.

Cependant, l'architecture « grande échelle » que représente EGEE, en intégrant des communautés scientifiques et opérationnelles disparates n'offrent peut être pas toutes les garanties espérées par les acteurs opérationnels de la gestion de crise. Ainsi, dans la conclusion générale de cette recherche, dédiée à la définition d'une organisation virtuelle centrée sur la gestion du risque inondation, la question de la disponibilité des ressources (réservation/préemption) sera abordée et considérée comme un des enjeux majeurs de la potentielle adoption de la technologie grille par la Sécurité Civile au sens large (des pompiers aux prévisionnistes).

VIII - 1.2. L'interopérabilité : entre sophistication et opérationnalité

Le chapitre 6, dédié à la méthodologie expérimentale a permis d'analyser et de présenter les potentialités fonctionnelles et opérationnelles des services web de l'OGC pour permettre l'interaction transparente entre le système d'information du SPC-GD (portail web) et l'architecture de grille EGEE. Cependant, les limites bien définies de l'intergiciel gLite en matière d'ordonnancement, limite l'efficacité calculatoire du Web Processing Service (WPS). *A contrario*, RRM-Grid a été présenté comme une solution efficace pour enrichir la capacité initiale du courtier de ressources gLite, sur lequel le serveur WPS se base, afin d'obtenir de meilleurs délais de réponses. Dans ce contexte, l'une des premières évolutions nécessaire à l'opérationnalité complète de G-ALHTAÏR concerne le rapprochement technologique entre ces deux solutions.

En préalable, l'implémentation de RRM-Grid en accord avec le fonctionnement imposé par l'adoption des services web (NetCDF et interfaces des modules hydrologiques) apparait comme une première impulsion concrète pour engager cette interaction et l'augmentation des capacités transactionnelles du Web Coverage Service (WCS).

Tout d'abord, l'un des enjeux technologiques de l'utilisation de la grille par des applications temps réel caractérisées par une actualisation rapide des données, comme dans le cas de la problématique de cette recherche, concerne l'interaction entre les systèmes d'information. Or, le serveur WCS, décrit dans la méthodologie (cf. §.VI – 1.2.2.), s'il permet le référencement de données géospatiales, ne permet pas une mise à jour des données. Depuis 2007, une proposition d'amélioration a été initiée pour offrir à la spécification du WCS une véritable capacité transactionnelle : le WCS-T (WHITESITE, 2009). Dans ce sens, son implémentation semble pouvoir consolider le fonctionnement de G-ALHTAÏR et offrir une méthodologie standardisée.

Ainsi, l'amélioration continuelle des spécifications du WCS, permet donc d'envisager de nouvelles capacités dans G-ALHTAÏR :

- Le découpage automatique des données d'entrée pour supporter les propositions d'améliorations calculatoires offertes par la parallélisation de données (cf. §.IV – 2.1.1.2.).
- Le changement de la résolution spatio-temporelle des données afin par exemple d'accélérer leur traitement dans le cas d'une situation d'urgence exceptionnelle
- La création de bassins versants à la volée, depuis l'interface de cartographie en ligne, sur lesquels des instances de modélisation peuvent être appliquées (WPS)

Au niveau du WPS et dans la même philosophie que les algorithmes de production et de transfert de G-ALHTAÏR, les capacités du serveur WPS peuvent être enrichies en développant une interface (au sens de connexion) avec le *manager* de RRM-Grid (**Figure 110**). De cette manière, le WPS :

- interagit avec le WCS-T pour la gestion des données d'entrée et de sorties, comme cela est le cas dans la méthodologie existante,
- instancie le *manager* de RRM-Grid en lui spécifiant les scénarios de prévision à modéliser,
- prend en charge la gestion des messages entre le manager, les *workers* et l'interface de cartographie en ligne (prévisionniste).

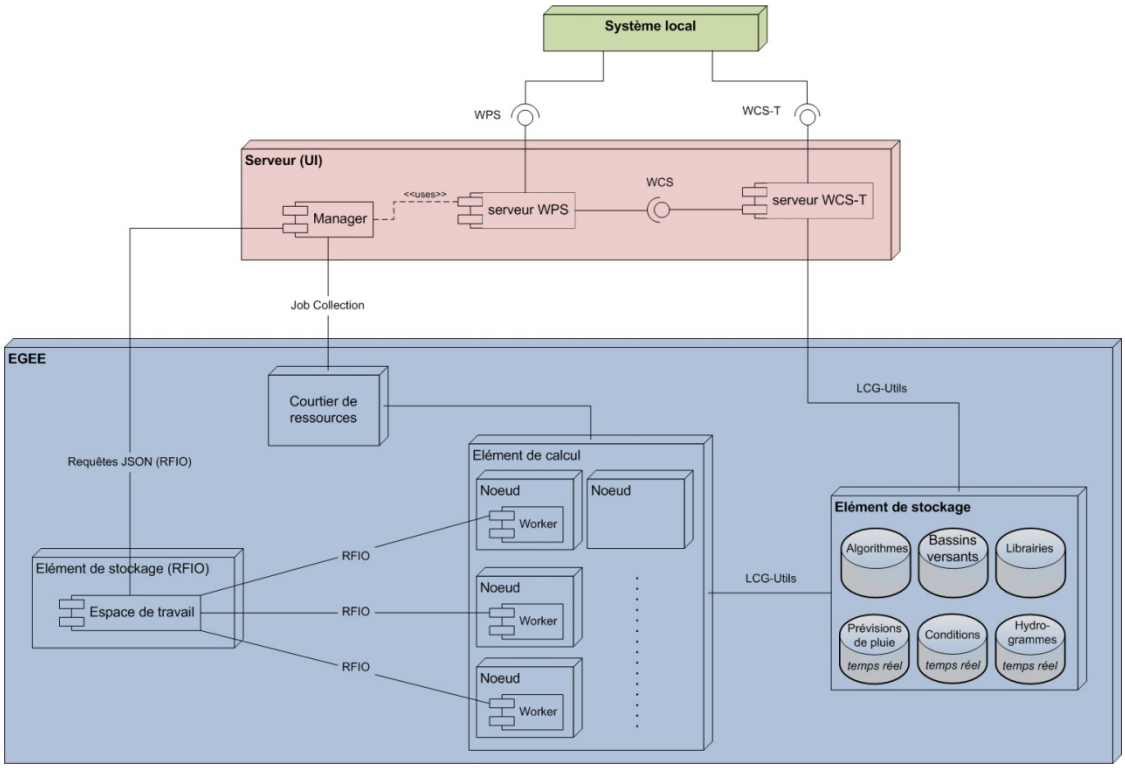


Figure 110 : Structure informatique de G-ALHTAÏR après intégration de RRM-Grid

D'autre part, le WCS-T permet de stocker et de référencer les données d'entrée des scénarios de prévision sur les éléments de stockage de l'architecture de grille au moyen de son opération « Transaction » (WHITESITE, 2009).

Enfin, comme cela est abordé dans la partie suivante, la dimension standardisée des services web de l'OGC peut faciliter l'intégration rapide d'autres modèles hydrologiques pluie-débit et de fonctions hydrologiques variées qu'il convient de présenter.

VIII - 2. Enrichissement de la dimension hydrométéorologique

L'ensemble des évolutions technologiques présentées dans le paragraphe précédent a pour objectif principal d'affermir l'opérationnalité du système spatial d'aide à la décision G-ALHTAÏR. Les présentes perspectives doivent cependant s'attacher à orienter les recherches à venir vers un enrichissement de nouveaux savoir-faire en hydrologie, en accord avec les besoins présentés dans le chapitre 3 (cf. §.III – 2.2.3.). En préalable, il est important de noter que la véracité hydrologique de ces solutions dépend en amont de celle de la prévision météorologique à une résolution spatio-temporelle fine (AYRAL, 2005 ; BEN DAOUD et *al*., 2009 ; CAUMONT et *al*., 2006 ; OBLED et BONTRON, 2005). Dans ce sens, OBLED et BONTRON (2005) proposent le recours aux méthodes par analogie, complétées récemment par les travaux de BEN DAOUD et *al*. (2009). En se basant sur une base de connaissance constituée d'évènements passés, dont le stockage et l'interrogation peuvent être effectués sur les ressources de grille, la méthode développée établit une comparaison statistique avec les pluies observées en temps réel et les prévisions originales, et génère des cellules pluvieuses potentielles en prévision. L'intégration du volet « prévision de pluies » reste toutefois une données d'entrée pour la plateforme G-ALHTAÏR, ce qui n'exclut pas d'utiliser, à plus long terme, les potentialités de la grille sur ce volet et pourquoi pas, si le prévisionniste en entrevoit le besoin (et dispose des compétences) de les intégrer à cette plateforme.

VIII - 2.1. L'approche multi-modèle en temps réel

La définition des besoins spécifiques à l'expertise hydrologique des crues à cinétique rapide (cf. §.III – 2.2.3.) corrélée à la caractérisation de ce phénomène (cf. §.I – 1.) a souligné la nécessité de prendre en charge les différents processus hydrologiques en jeu sur l'ensemble de la zone concernée. En effet, la modélisation en hydrologie impose clairement une approche multi-échelle capable de formaliser la génération des écoulements à l'échelle d'une placette, d'un versant, d'un bassin versant et d'un cours d'eau. Cette grande disparité de fonctionnement et les connaissances inégales les concernant imposent donc de prendre en charge chacun de ces phénomènes au cas par cas et de les intégrer dynamiquement du fait de leur interdépendance fonctionnelle.

VIII - 2.1.1. Une panoplie de modules de production et de transfert

Étant donné le complément de ressources informatiques offert par le recours à la technologie grille et la démonstration précédente de la capacité de RRM-Grid (par extension G-ALHTAÏR suivant les perspectives proposées dans le chapitre 7) à supporter différents scénarios hydrologiques, il apparaît envisageable de rajouter, comme paramètre d'entrée des requêtes du prévisionniste, les modules de production et de transfert à utiliser pour chaque scénario.

Comme cela a été montré dans le chapitre 1 (cf. §.I - 1.3), il existe un grand nombre de fonctions de production et de transfert utilisés pour respectivement calculer le ruissellement de surface et les débits correspondants dans le cas des phénomènes de crues éclair. L'objet de cette discussion n'est pas de sélectionner les mieux appropriés mais d'insister sur la possibilité offerte par ces nouvelles capacités de calcul. À titre d'exemple, il est donc possible de citer les fonctions de production des modèles suivants :

- SCS (Soil Conservation Service) (Bouvier et *al*., 2004 ; MARCHANDISE, 2007)
- TopModel (BEVEN et KIRBY, 1979)
- MARINE (Modélisation de l'Anticipation du Ruissellement et des Inondations pour des Évènements Extrêmes) (ESTUPINA-BORRELL et *al*., 2006)
- Le modèle développé par TOUKOUROU et *al*. (2009) basé les réseaux de neurones

En ce qui concerne les modules de transfert, au module de transfert géomorphologique implémenté dans cette recherche, il convient de rajouter un module basé sur le principe de l'onde cinématique. Cette fonction est basée sur le principe des mailles interactives[3]. Les nombreuses recherches qui se sont intéressées à l'ensemble de ces fonctions hydrologiques, en font des solutions de modélisation hydrologique performantes et adaptées au contexte opérationnel.

À titre d'exemple, pour le module de production de SCS, les modifications algorithmiques peuvent être minimisées. En effet, comme pour ALHTAÏR, ce module prend en charge une donnée de pluie et des conditions hydrologiques initiales. La principale différence concerne le calcul de la capacité d'infiltration qui, dans le cas de SCS, est basé sur le cumul pluviométrique, ce qui ne constitue pas un facteur limitant à la gestion de son exécution par le WPS et donc son intégration dans la plateforme G-ALHTAÏR.

Finalement, DELRIEU et *al*. (2009) argumentent dans le sens de ces propositions en rappelant les nouvelles opportunités offertes par les nouvelles approches informatiques et les ressources informatiques accrues pour la caractérisation de la variabilité des processus du cycle de l'eau à la surface de la terre et notamment le développement de modèles à base physique, modèles qui justifient l'utilisation d'un système permettant une capacité calculatoire importante.

VIII - 2.1.2. La modélisation hydraulique

La modélisation des débits en temps réel et en prévision est donc considérée comme une étape indispensable à la réussite de l'expertise hydrologique en situation de crise. Cependant, cette information implique une connaissance précise du comportement de chaque cours d'eau réglementaire, que le gestionnaire de crise (décideurs de la Sécurité Civile ou des pouvoirs publics) ne maitrise pas forcément. Dans ces conditions, l'objectif idéal des opérations d'expertise des SPC concerne l'évaluation du risque « inondation » en temps réel au niveau des plaines aval, où les enjeux sont *a priori* les plus importants. Cette capacité opérationnelle est tributaire de la performance de la modélisation hydraulique, permettant de prendre en compte les débits et les hauteurs des cours d'eau des plaines aval (BONNIFAIT et *al*., 2009). La connaissance de plusieurs valeurs de débit sur ces plaines aval permet alors d'évaluer l'extension potentielle du champ d'inondation[4], et *a fortiori* par croisement avec les enjeux socio-économiques en jeu, la cartographie du risque « inondation » en temps réel. La production de cette carte fait appel à des fonctionnalités, y compris de géotraitement, consommatrices

[3] Le ruissellement de surface d'une maille est transféré à une ou plusieurs mailles aval, une explication détaillée de ce principe peut être trouvée sur le site de la plateforme ATHYS (http://www.athys-soft.org/v1/aide/mercedes/interactives.html)

[4] Extension géographique de l'eau lorsque « le débit du cours d'eau excède la capacité de débit du lit principal » (Glossaire international d'hydrologie : http://hydrologie.org/glu/aglo.htm).

de puissance de calcul, limitant sa réalisation en situation opérationnelle avec les moyens informatiques du SPC-GD. Ainsi, cette opération peut elle aussi tirer profit des ressources supplémentaires offertes par l'architecture de grille et les méthodologies informatiques propres aux traitements parallèles.

À titre d'exemple, au niveau hydraulique, il pourrait être possible d'utiliser le modèle Dassflow basé sur les équation de Saint-Venant (HONNORAT et *al*., 2005), pour assurer la modélisation hydraulique de la plaine de la Gardonnenque.

VIII - 2.1.3. Prototype d'implémentation potentiel de l'approche multi-modèle

Une des implémentations possibles de ce couplage hydrologique-hydraulique est présenté dans la **Figure 111**.

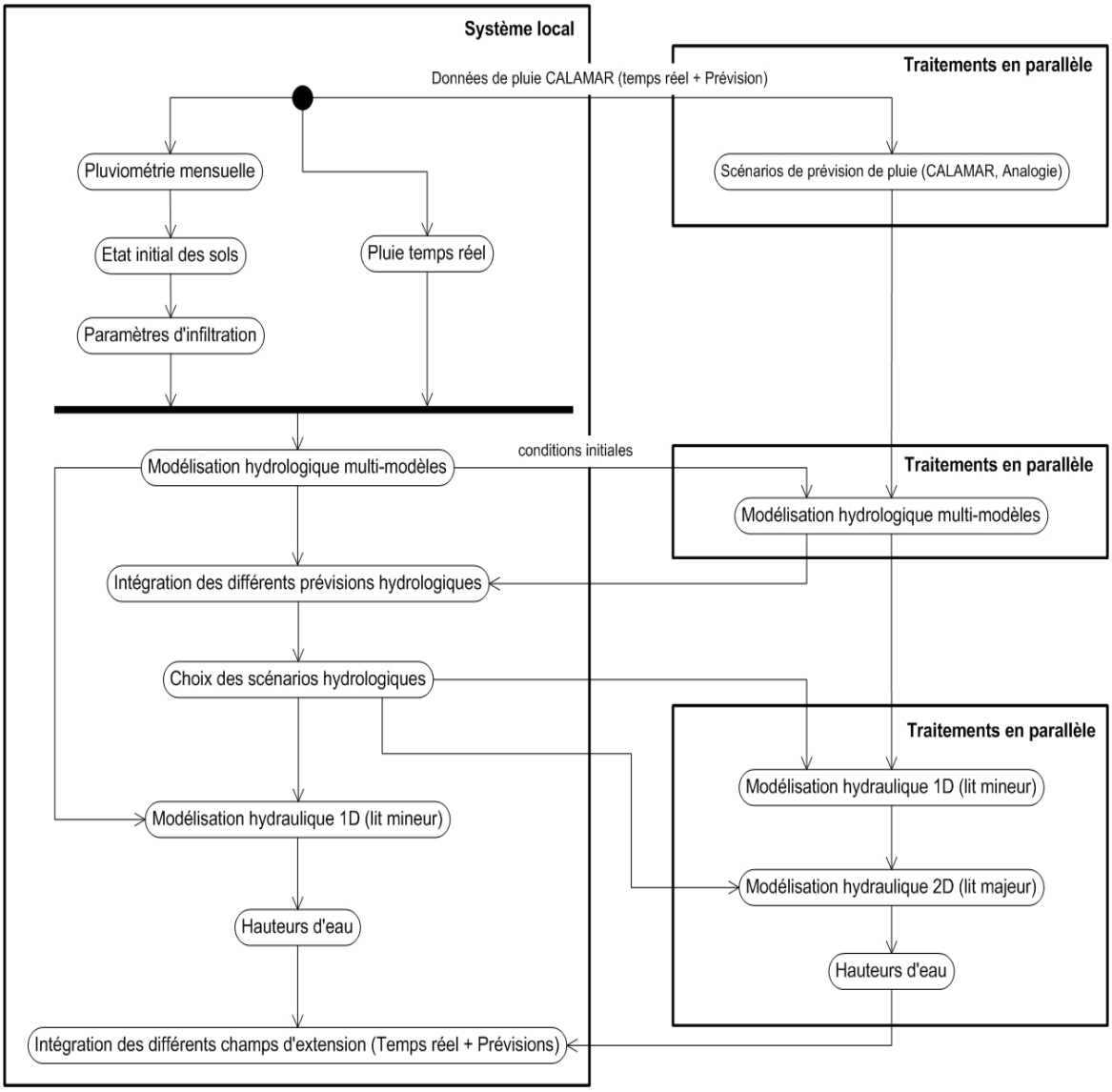


Figure 111 : Stratégie de modélisation hydrologique et hydraulique multi-modèles

Sur le système local, les conditions hydriques du sol, calculées au moyen de la pluviométrie mensuelle, permettent d'établir le jeu de paramètres d'infiltration. À partir de ce point, plusieurs modèles

hydrologiques simulent le débit temps réel à l'exutoire des bassins versants. Dans un second temps, le modèle hydraulique 1D propage les hauteurs d'eau, correspondantes à ces débits, à plusieurs points de référence aval du lit mineur en se basant sur les données topographiques de ce lit.

Dans le même temps, le prévisionniste soumet des prévisions hydrologiques sur l'architecture de grille avec les mêmes paramètres que celles de la méthodologie de cette recherche (cf. §.) à laquelle a été ajouté le modèle hydrologique à utiliser. À la réception des hydrogrammes correspondants, le prévisionniste établit son expertise hydrologique et sélectionne les débits à venir les plus probables. Il soumet en suivant, les prévisions hydrauliques qu'ils désirent tester grâce à une approche 1D (propagation du débit) et/ou une approche 2D (champ d'inondation).

Le recours à la grille peut donc permettre, comme dans le cadre de la méthodologie de cette recherche, de prendre en charge différents scénarios de prévision hydrologique et/ou hydraulique. En effet, la partie de droite de la **Figure 111** est composée de trois blocs de traitement dont les instances peuvent être exécutées en simultanée :

- Le premier bloc se base sur les données de pluie temps réel et en prévision CALAMAR et construit, comme pour la méthodologie initiale développée dans G-ALHTAÏR, différents scénarios de prévision météorologique. Il est possible d'imaginer l'utilisation de la méthode par analogie pour construire des prévisions supplémentaires, ou toutes autres techniques,

- Le second bloc correspond à la version initiale de G-ALHTAÏR qui, en fonction des conditions hydrologiques initiales modélisées sur le système local, des différents scénarios de prévision météorologique, du bassin versant et du modèle hydrologique, simule les débits potentiels à l'exutoire du bassin versant,

- Le troisième bloc prend lui en charge la modélisation hydraulique 1D et 2D en fonction des différents scénarios de débit que le second bloc lui a fourni depuis le système local.

Le développement de cette approche doit s'appuyer sur les développements de cette recherche. Étant donné, l'augmentation des communications entre le système local et les blocs de traitement eux-mêmes, et la nécessaire synchronisation entre les scénarios, cette approche justifie l'utilisation d'un serveur OGC constitué d'un WCS et d'un WPS dont les différentes capacités correspondent à :

- différents modules de production et de transfert hydrologique,

- les modèles hydrauliques 1D et 2D,

- les méthodes permettant de définir la pluie de prévision.

Quant au WCS, il doit faciliter l'échange de données entre le système local et les éléments de stockage de la grille.

Dépendant de la complexité algorithmique de ces modèles, et en particuliers les données d'entrée nécessaires à l'exécution de ces modèles, cette phase d'intégration peut s'avérer plus ou moins complexe.

VIII - 2.2. Vers l'optimisation du modèle

L'incertitude en modélisation hydrologique est relativement importante, en particulier du fait du caractère « conceptuel[5] » des modèles utilisés en opérationnel, qui cherchent à synthétiser des processus hydrologiques foncièrement différents. En effet, la plupart de ces modèles conceptuels ne sont pas transposable spatialement et doivent être « réglés » en fonction des conditions géographiques de chaque bassin versant à modéliser (MARCHANDISE, 2007). Ainsi, leur utilisation en situation opérationnelle implique donc au préalable l'étalonnage (calage) de leurs paramètres en temps différé ou l'assimilation des débits mesurés en temps réel.

VIII - 2.2.1. Étalonnage des modèles

En se basant sur un échantillon d'évènements passés, la phase de calage permet de déterminer les meilleurs paramètres du modèle à utiliser pour simuler au mieux le débit. De nombreux algorithmes existent pour soutenir cette opération. Pour approcher au mieux le jeu de paramètres « idéal[6] », il est donc nécessaire de tester l'ensemble des combinaisons de paramètres compris dans l'espace de paramètres défini au préalable, de généraliser cette démarche à de très nombreux bassins versants et de l'effectuer régulièrement après chaque évènement important. Cette opération consiste donc à un grand nombre d'itérations basées sur les différents jeux de paramètres et les évènements à tester. L'ensemble de ces hydrogrammes doit ensuite être comparé à une ou plusieurs références calculées sur la base d'évènements différents.

Il devient alors délicat, dans ces conditions, de calibrer le modèle à tous les types d'occurrence hydrométéorologique que le bassin versant peut connaître. L'un des inconvénients de ces méthodes se situe dans la quantité de ressources informatiques qu'elles requièrent, en particuliers quand les paramètres cherche à être spatialisés (AYRAL, 2005). Il devient alors intéressant de déporter ces traitements sur des ressources de calcul distantes d'une architecture de grille. En effet, les capacités de stockage offertes permettent, comme pour la méthode par analogie de pluie, de disposer de l'ensemble des évènements, alors qu'un nombre important d'instances de calage (jeux de paramètres) peuvent être exécutées en simultané sur différents nœuds offerts. Ainsi, le fonctionnement de RRM-Grid s'aligne parfaitement avec les besoins technologiques de cette approche, d'autant que la durée des simulations correspond à une utilisation plus classique de la grille (ASTALOS, 2008).

VIII - 2.2.2. Assimilation des débits

La prévision météorologique opérationnelle (Météo-France) a, depuis quelques années, recours à des méthodes d'assimilation des données stationnelles temps réel (température, humidité, etc.) pour améliorer la précision des prévisions à grande échelle qu'elle diffuse, en particuliers grâce au modèle ALADIN/France (FISCHER et *al*., 2006). Dans la même philosophie, des méthodes similaires sont en cours de développement dans le domaine de l'hydrologie opérationnelle (BROERSEN, 2007 ; COLLIER, 2007 ; RABUFFETTI, 2006) et de l'hydraulique (CASTAINGS et *al*., 2006 ; LAI et MONNIER, 2009).

---

[5] Dans le cas des modèles conceptuels, ces paramètres ont rarement une signification physique, et sont dédiés à la correction des incertitudes sur les données d'entrée et sur la conceptualisation du modèle proprement dit

[6] Il n'existe que très rarement qu'un seul jeu de paramètres, étant donné l'interdépendance des paramètres

Une des solutions envisagée concerne l'assimilation en temps réel des observations *in situ*, et en particuliers des débits ou hauteur d'eau. Cette approche permet de recalculer à la volée les paramètres de modélisation (dans le cas d'ALHTAÏR : fo, fc, i, si, fv et H) tout au long du processus de simulation. En situation opérationnelle, cette méthodologie offre un (des) résultat(s) de modélisation supplémentaires que le prévisionniste peut utiliser pour les confronter avec l'ensemble des hydrogrammes que le système G-ALHTAÏR lui fournit. Dans ce contexte, la solution opérationnelle offerte par les services web de l'OGC (ou par RRM-Grid) peut être adaptée en insérant ces paramètres hydrologiques directement dans la requête XML. Ainsi, en plus des paramètres de variation sur la donnée de pluie prévue, les scénarios de prévision soumis par le prévisionniste peuvent s'appuyer sur différents jeu de paramètres hydrologiques issus de la phase d'assimilation.

Cette méthode doit s'appuyer sur une stratégie opérationnelle en accord avec les contraintes de gestion de crise déjà définies (cf. §.IV – 1.). En effet, il parait difficilement concevable d'effectuer une assimilation des données à tous les pas de temps. Une solution à envisager pourrait être le lancement de cette procédure d'assimilation quand le prévisionniste identifie une divergence entre les débits modélisés et observés.

Enfin, cette méthode impose sans doute l'intégration automatique des données collectées sur les systèmes de stockage de la grille, afin de limiter le nombre d'interactions entre le système d'information local et les ressources de grille.

VIII - 2.3. Gestion des systèmes de mesure et des données collectées

De nombreuses recherches s'intéressent à l'heure actuelle à l'intégration transparente des capteurs et des systèmes de mesure au sein d'architectures de grille. Cette approche apparaît d'autant plus prometteuse dans le cas de données satellitales à la taille souvent contraignante pour des systèmes informatiques classiques (ordinateur personnel). Le projet CEOS / WGISS (cf. §.IV – 2.2.3.1.1.) consacre une partie de ses recherches à cette problématique. À une autre échelle, le projet e-CORCE qui projette de développer une constellation de satellites capables de fournir une information à résolution spatio-temporelle fine (1 m et 24h), implique l'utilisation d'une architecture de grille grand échelle capable de synchroniser, collecter et diffuser la production de données équivalent à un débit de 400 Gbit/s (ANTIKIDIS et *al*., 2008).

Dans le cadre de la gestion de l'expertise hydrologique et des perspectives déjà définies, cette approche pourrait offrir un certain intérêt. En effet, l'accès direct aux données collectées par les stations de mesure et à l'imagerie radar sur le territoire « Grand Delta » depuis les ressources de calcul et de stockage, pourrait simplifier leur manipulation et accélérer leur prise en compte par les composants de la méthodologie des présents travaux. La description du système d'information existant du SPC-GD (cf. §.III. 1.2.5.) a permis de souligner la structuration par composants particulièrement adaptée à leur reconsidération dans une optique d'architecture informatique distribuée. En se basant sur cette description, il est possible d'appréhender plusieurs niveaux d'intégration :

- la spécification du service web WCS-T constitue la première solution susceptible d'offrir un moyen de synchronisation entre les données collectées et stockées sur le système d'information du SPC-GD et les ressources de grille.

- Il semble envisageable de référencer les serveurs CALAMAR 2 et SIGMA2000 comme des éléments de stockage (SE) d'une organisation virtuelle dédiée aux inondations ou de ESR, de la même manière que la perspective énoncée dans le paragraphe VIII – 1.1.1.1.,

- HUGHES et *al*. (2006) ont développé un système d'information dédié à la gestion des mesures et des observations d'un cours d'eau, dans lequel chaque capteur est considéré comme un nœud de grille à part entière. De plus, chaque station est dotée de fonctionnalités de bas niveau lui permettant de contrôler en temps réel les informations dont elle est en charge et d'adapter son comportement en fonction de l'activité hydrologique ou des requêtes à distance du prévisionniste. Il est aussi possible de citer le travail de ZHOU et DE ROURE (2007) dans lequel les capteurs ont la capacité de gérer leur énergie électrique en fonction des conditions locales. Dans tous les cas, il est souvent question de « sensor grid ».

Cette dernière solution représente cependant une reconsidération forte de l'équipement existant et doit donc être envisagée à plus long terme. Au niveau de la recherche en hydrologie, elle peut cependant offrir un moyen technologique avantageux pour contrôler le cycle de vie et le rapatriement des données collectées depuis des stations de mesure, utilisées dans le suivi et le retour d'expérience d'évènements.

Quoiqu'il en soit, un transfert automatique de ces données sur l'infrastructure de grille, à chaque pas de temps mesuré, représente une solution initiale, limitant la refonte du système d'information existant du SPC-GD et permettant d'une part la sécurisation de ces données et d'autre part la rapidité de traitement *a posteriori*.

SYNTHESE

L'ensemble de ces nouvelles capacités technologiques et de ces fonctionnalités « métier », développées dans ces perspectives, sont dédiés à l'enrichissement progressif du système d'aide à la décision G-ALHTAÏR. Elles ont chacune un impact différent sur le système d'information existant et doivent être appréhendée selon les besoins des services opérationnels considérés et l'intégration des fonctions « métier » dans l'architecture de grille. Elles impliquent une phase de grillification préalable des algorithmes correspondants qui peut s'avérer dans certains cas relativement contraignante, et nécessiter une reconsidération de l'algorithme initial afin de l'adapter aux contraintes fonctionnelles de l'architecture de grille.

Ainsi, ces opérations de portage ne sont a priori pas généralisables et nécessitent une considération précise des objectifs opérationnels, suivant qu'ils soient centrés sur le gain de performance (parallélisation du code ou instances d'exécution en parallèle) ou le déport de ces opérations sur une architecture informatique distante. Dans tous les cas, un respect des recommandations développées dans cette recherche peut permettre à terme l'encapsulation de l'ensemble de ces fonctionnalités par un WPS et la gestion des données par un serveur WCS-T, afin de faciliter leur utilisation en situation de crise ou dans un contexte de recherche.

D'autre part, la capacité de synchronisation de ces services web de l'OGC permet une approche par composants favorisant la construction dynamique de chaînes de traitements adaptables aux différentes situations hydrologiques et objectifs de recherche. Dans ce sens, les fonctions opérationnelles sont clairement séparées du niveau algorithmique, suivant l'approche promue par Argent (2004). Ainsi, le système proposé doit permettre au prévisionniste d'accéder à des solutions logicielles faciles d'utilisation, qu'il peut sélectionner et exécuter sans se préoccuper des aspects technologiques sous-jacents ou de la contrainte récurrente de la gestion d'un grand nombre de jeux de données.

Finalement, l'ensemble de ces fonctionnalités centrées « hydrologie » doit donc permettre de modéliser le débit des cours d'eau et les zones inondées correspondantes, qu'ils soient situées sur les bassins versants amont répondant à des processus majoritairement hydrologiques ou sur les plaines avals contrôlées par un fonctionnement hydraulique. En situation de crise, l'expérience du prévisionniste sur son territoire de compétence doit lui permettre de caractériser spatialement les processus en jeu, afin de développer une stratégie de modélisation grâce à l'imbrication dynamique des fonctionnalités précédemment citées.

CONCLUSION

GÉNÉRALE

CONCLUSION GÉNÉRALE

Bilan scientifique

Cette recherche interdisciplinaire, qui s'articule autour de deux thématiques principales, l'informatique distribuée et l'hydrologie opérationnelle, a donc cherché à répondre à la problématique générale de la potentialité de la technologie grille pour la prise en charge de l'expertise hydrologique des crues à cinétique rapide en situation de crise, et plus particulièrement à la question délicate de leur prévision à court terme. En effet, le type d'aléa, de par ses caractéristiques physiques complexes et les multiples processus qui le composent, contraint les services opérationnels, et les laboratoires en charge de leur compréhension au développement de moyens technologiques, organisationnels et politiques spécifiques.

En définissant clairement les objets de cette étude à savoir le phénomène naturel de crue à cinétique rapide et le service de prévision des crues « Grand Delta » (SPC-GD) en charge de son expertise, cette recherche s'est appuyée sur un cas d'utilisation concret et représentatif du fonctionnement et des besoins plus généraux de la Sécurité Civile. Ainsi, l'expertise hydrologique des crues à cinétique rapide représente, durant une situation de crise, un élément essentiel à la protection des populations et des infrastructures. Dans ce sens, ce phénomène de crue à cinétique rapide, caractérisé par des temps de concentration courts, limite le délai opérationnel alloué à l'élaboration de l'expertise hydrométéorologique et à sa transmission aux services de secours et aux pouvoirs publics. Une telle situation a été illustrée au travers de l'étude des évènements de 1999 (Aude) et de 2002 (Gard) durant lesquels la chaîne d'alerte et de secours a été fortement impactée et a montré ses limites fonctionnelles.

Ainsi, le perfectionnement de ces opérations d'expertise doit s'appuyer sur un ensemble de leviers organisationnels, scientifiques et politiques que l'analyse des besoins du chapitre 3 a permis de préciser. La voie d'amélioration technologique ne peut *a priori* pas assurer à elle seule la mission d'amélioration édictée par ses recommandations. Cependant, elle semble pouvoir constituer une assise solide susceptible d'envisager et d'encourager le développement d'initiatives collectives, interdisciplinaires et intégrées entre « chercheurs et opérationnels » pour la résolution progressive des problématiques liées aux inondations et aux crues à cinétique rapide.

Dans ce sens, l'intérêt de la réflexion menée dans ce mémoire se situe dans une approche exhaustive dans l'analyse des besoins inhérents à la gestion de crise « inondation », incluant les crues à cinétique rapide. La légitimité scientifique de cette démarche semble consolidée par le large éventail de besoins

et de contraintes fonctionnelles fondamentales de la Sécurité Civile. L'analyse de ces besoins a permis de développer une proposition méthodologique basée sur quelques points d'amélioration préférentiels.

Dans ce sens, le choix de la technologie grille, pour soutenir cette nécessaire amélioration, en comparaison avec d'autres technologies de l'informatique distribuée, permet de fournir un socle délocalisé constitué d'un grand nombre de ressources informatiques et correspondant à un environnement collaboratif sécurisé et dynamique au travers du concept d'organisation virtuelle. L'analyse détaillée des caractéristiques technologiques de la grille a permis d'envisager un enrichissement fonctionnel et calculatoire des systèmes d'information de petite envergure tels qu'ils existent dans les services opérationnels régionaux de la Sécurité Civile. La généricité de fonctionnement offerte par la virtualisation des ressources informatiques qu'elle fournit et la formalisation progressive de ses opérations informatiques, grâce aux efforts de l'OGF (Open Grid Forum) et en Europe du projet EGEE (Enabling Grids for E-sciencE), offre un socle de calcul et de stockage adaptable à un grand nombre de problématiques. D'autre part, le rapprochement entre les sciences de l'information géographique (géomatique) et la communauté « grille »[1], durant ces dernières années, permet d'envisager une généralisation des services d'accès et de traitements de l'information géographique adaptables aux différentes problématiques des sciences de la Terre. La communauté de l'hydrologie opérationnelle, qu'il est possible d'inclure dans ce domaine, du fait de l'imbrication de problématiques géographiques et hydrologiques, peut donc *a priori* tirer profit de cette révolution technologique.

En intégrant plusieurs dimensions technologiques fortes, elle permet de décrire plusieurs orientations sur lesquelles les recherches relatives à l'expertise hydrologique peuvent s'appuyer. Il est donc important à ce niveau de rappeler ces trois dimensions expérimentales :

- La dimension « Technologie » qui correspond aux traitements de bas niveau sur les ressources de grille. Dans cette recherche, la capacité de prise en charge d'études paramétriques de la technologie grille a été plus précisément approfondie pour permettre le traitement d'un grand nombre de scénarios de modélisation hydrologique dans des délais opérationnels courts et une qualité de service proche des contraintes du fonctionnement de la Sécurité Civile
- La dimension « Traitement » qui est susceptible de supporter l'interaction entre des systèmes d'information distants et hétérogènes par le recours à des méthodes « ergonomiques » standardisées capables de s'auto-décrire et d'interagir de manière autonome, les services web de l'OGC (Open Geospatial Consortium). Le cadre générique qui caractérise cette dimension permet d'entrevoir une certaine efficacité dans la conception de nouvelles capacités et *a fortiori* de nouvelles compétences.
- La dimension « Entreprise » qui apparait indispensable aux opérations relatives à la gestion de crise. Elle s'appuie sur une interface graphique connectée aux composants du niveau « Traitement » et facilite d'une part l'appréhension de la situation « temps réel » et, d'autre part

[1] Il est question du rapprochement entre l'Open Geospatial Consortium (OGC) et l'Open Grid Forum (OGF)

la construction de scénarios de prévision hydrologique par le prévisionniste du SPC-GD. Son intégration dans un environnement « Web » prolonge l'interopérabilité et l'ubiquité offertes par la technologie grille et les services web de l'OGC.

Ainsi, cette méthodologie fondée sur des solutions logicielles innovantes, telles que les services web géographiques de l'OGC, confirme et s'aligne avec la nécessité de développer des interfaces entre les systèmes d'information géographiquement distants, afin d'assurer l'opérationnalité globale du système distribué. Le recours à des interfaces standardisées doit de plus faciliter et accélérer l'intégration rapide :

- de nouveaux composants logiciels et de nouvelles sources de données,
- de nouveaux partenaires inclus dans le processus d'aide à la décision hydrologique.

Il apparait dès lors envisageable de poursuivre les axes de développement initiés par cette recherche, en élargissant le champ d'application de cette méthodologie et en développant une véritable approche « métier » dans laquelle chaque acteur manipule et traite les données relatives à son champ d'expertise, sans contraintes informatiques fortes et dans un contexte collaboratif. G-ALHTAÏR peut donc s'orienter à terme vers un « système spatial d'aide à la décision hydrologique collaboratif ».

Finalement, la distribution géographique des composants du système G-ALHTAÏR, et plus largement la structuration de la grille en ressources géographiquement distantes semblent pouvoir permettre une meilleure sécurisation des traitements et des données, dans le sens où ils peuvent être gérés à l'extérieur de la zone à risque. L'aspect ubiquitaire de cette technologie facilite de plus la récupération de ces opérations en temps réel par des partenaires de la gestion de crise autorisés et épargnés par le phénomène dangereux, afin d'assurer la continuité de l'expertise. La relation opérationnelle SCHAPI[2] / SPC s'adapte bien à cette nouvelle capacité.

Finalement, il est important de noter que de nombreux travaux scientifiques abordent cette problématique de l'expertise hydrométéorologique sous l'angle technologique, mais comme le rappelle Montz et Gruntfest (2002) réduire la problématique de gestion des crues éclair à une simple question technologique ne suffit pas. Fort de cette considération, les orientations méthodologiques et les résultats de cette recherche doivent être appréhendés avec une certaine précaution et ne doivent pas occulter :

- la nécessité d'étudier cette problématique sous l'angle des sciences sociales et de l'effort de sensibilisation toujours nécessaire
- la nécessaire poursuite des recherches sur la connaissance des processus de genèse de ce type de phénomène qui permettra le développement des modèles de prévision plus pertinents et plus efficaces.

[2] Service Central d'Hydrométéorologie et d'Appui à la Prévision des Inondations

Une organisation virtuelle pour le risque « inondation »

La problématique initiale de cette recherche a orienté les développements méthodologiques vers une considération forte de la dimension « temps réel ». Elle s'est attachée à considérer l'expertise hydrologique en situation de crise. Cependant, certains arguments forts argumentent en faveur d'un élargissement de cette méthodologie vers une dimension plus globale de prise en compte des problématiques fondamentales du risque « inondation ». Cette approche s'appuie sur des méthodologies d'analyse et d'interprétation des données hydrologiques, stratégiques et opérationnelles en « temps différé[3] ». En effet, la dimension collaborative forte offerte par les technologies appréhendées dans ce mémoire et l'accès à des ressources de calcul mais aussi de stockage à la demande semble offrir un socle technologique et politique solide pour intégrer les opérations relatives à la gestion du risque « inondation ». Il peut permettre une continuité temporelle dans l'accès et l'analyse des données hydrométéorologiques parmi l'ensemble des acteurs institutionnels et académiques, de la phase de prévention aux retours d'expérience pratiqués après chaque occurrence de crue, en passant par l'expertise hydrométéorologique décrite dans ce mémoire (Fig. 112).

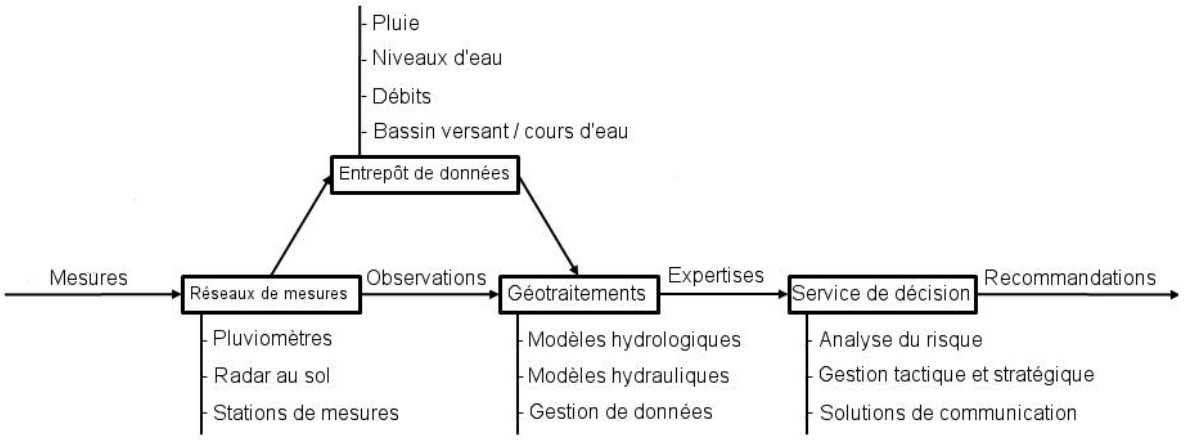


Figure 112 : Schéma synthétique d'un système d'information dédié à la gestion du risque "inondation" basé sur la technologie grille (adapté de Lee et Percivall, 2008)

Ainsi, en se basant sur les acquis technologiques offerts par G-ALHTAÏR, l'ensemble des acteurs pourraient être regroupés au sein d'une organisation virtuelle permettant de connecter dynamiquement et de manière interopérable l'ensemble de leurs systèmes d'information, dans la philosophie du concept d'e-Science (GAGLIARDI et *al.*, 2005 ; GURNEY et *al.*, 2009 ; HEY et TREFETHEN, 2002 ; JONES, 2009). Les données et les outils respectivement gérés et développés par chaque partenaire pourraient être partagés et intégrés (au sens de réplication) au sein des ressources informatiques offertes d'une organisation virtuelle à définir. Cette organisation virtuelle pourrait avoir la vocation d'une plateforme d'accès aux données historiques et « temps réel », constituant la base de connaissance

[3] Hors des périodes de crise

hydrométéorologique, et des capacités de traitement adaptées à ces données, le tout sur un environnement collaboratif pérenne.

Une telle vision, si elle s'aligne avec la volonté d'interdisciplinarité partagée par l'ensemble des acteurs, nécessite au préalable une expérimentation approfondie du fond méthodologique de cette recherche et plus particulièrement une réflexion globale sur des aspects relatifs à la gouvernance de cette organisation virtuelle, discutés plus globalement dans le dernier volet de ce mémoire. À titre d'exemple le projet BVNE (Bassins Versants Numériques Expérimentaux), qui regroupe des acteurs du monde de l'hydrologie pour la modélisation et la caractérisation des crues à cinétique rapide, représente une base solide sur laquelle il convient sans doute de s'appuyer.

Une adoption multiforme

En clôture de ce mémoire, il convient de s'intéresser à la problématique déjà abordée de disponibilité des ressources informatiques à la demande dans un contexte concurrentiel que représente la technologie grille. Le concept d'organisation virtuelle de la grille EGEE assure un niveau d'accessibilité élevé mais les résultats présentés dans le chapitre 7, tendent à relativiser la pleine disponibilité de ces ressources, d'autant que la gestion de crise revêt un fonctionnement imprévisible en temps et en charge de calcul.

Ainsi, les orientations politiques à développer pour poursuivre l'engagement technologique énoncé dans ce mémoire peuvent se baser sur une des deux dimensions de gouvernance suivantes :

1. Dans la philosophie des travaux présentés, les acteurs opérationnels doivent poursuivre les efforts de développement, initiés par le système G-ALHTAÏR, de manière intégrée avec les fournisseurs de ressources d'une organisation virtuelle associée au projet EGEE. En l'occurrence, la VO ESR semble s'aligner avec les problématiques de la gestion de crise. Ainsi, des campagnes de « fidélisation » de sites de calcul régionaux (par ex. Marseille et Montpellier) peuvent être envisagées afin d'assurer la disponibilité des ressources par la prioritisation des opérations d'ordonnancement provenant du SPC-GD lors de la survenance d'épisode d'inondation. Cette solution implique certainement la contribution de la part du SPC-GD à l'acquisition de ressources de calcul et de stockage dédiées à ces sites, dans la philosophie de partage initiée par la grille.

2. Une seconde approche concerne la réforme des services institutionnels dédiés à l'expertise hydrologique, dans laquelle cette recherche est rattachée. L'harmonisation technologique, explicitement décrite dans cette réforme, semble prédisposée à une réflexion collective envisageant la construction de sites de calculs localisés dans les services de prévision des crues, au SCHAPI ou chez les partenaires académiques et opérationnels de la gestion du risque « inondation ». Cette dimension

s'aligne sur l'organisation virtuelle dédiée au risque « inondation » proposée précédemment. Finalement, ces sites, principalement dédiés aux opérations d'expertise hydrométéorologique en situation de crise, pourraient être intégrés à une architecture de plus grande échelle, telle qu'EGEE, afin de leur assurer une viabilité technologique et une légitimité politique, leur permettant de disposer d'un surplus potentiel de ressources dans le cadre de recherches en hydrologie.

Dans tous les cas, l'orientation stratégique à poursuivre à la suite de ces travaux doit s'appuyer sur une réflexion à long terme, articulée entre la poursuite des expérimentations développées dans ces travaux, en enrichissant la méthodologie suivant les perspectives présentées dans le chapitre 8, et des « challenges » sur une architecture de grille, telle qu'EGEE, en reproduisant le déroulement d'évènements passés extrêmes, comme celui de 2002 dans le Gard, de manière à développer les stratégies d'adoption les mieux adaptées.

Finalement, les nouvelles orientations politico-organisationnelles du projet européen EGEE, en accord avec sa mise en production, doit permettre dans les mois à venir une structuration politique géographique, imbriquant des dimensions locales (nœuds de grille régionaux, Montpellier, Marseille, etc.), nationales (NGI[4]) et européenne (EGI[5]) de la gouvernance de cette grille. Cette orientation semble offrir un socle solide pour l'adhésion de communautés de plus en plus diverses. Au niveau français, une démarche prospective pour la définition des besoins, des rôles et des orientations en matière de grille de calcul a récemment été menée par l'Institut Des Grilles (IDG[6]). Au travers de son livre blanc (IDG, 2009), cet institut a défini les bases politiques et scientifiques de la future « NGI France », vision consolidée par « le protocole d'accord national signé à l'automne 2008 entre le Ministère de la recherche, le CNRS, le CEA, l'INRIA, l'INRA, RENATER et la conférences des Présidents d'Universités ». Ainsi, ces nouvelles garanties offrent certainement une opportunité nouvelle d'envisager une définition et un positionnement précis de la mission vitale dont les services de Sécurité Civile sont en charge, afin de garantir une mise à disposition de ressources informatiques suivant le principe d'intérêt général.

[4] National Grid Initiative
[5] European Grid Initiative
[6] http://idgrilles.fr/

BIBLIOGRAPHIE

Abbott, M, Bathurst, J., Cunge, J., O'connell, P., Rasmussen, J., (1986), An introduction to the European Hydrological System-Système Hydrologique Européen, SHE. II: Structure of a physically-based, distributed modelling system, Journal of hydrology (Amsterdam), 87, pp. 61-77

Adzigogov, L., Soldatos, J., Polymenakos, L., (2005), EMPEROR: An OGSA Grid meta-scheduler based on dynamic resource predictions, Journal of Grid computing, 3, 1, pp. 19-37

AFIGéO, (2007), La Directive INSPIRE - vue d'ensemble, Association française pour l'information Géographique, 12 p.

Al-Sabhan, W., Mulligan, M., Blackburn, G.A., (2003), A real-time hydrological model for flood prediction using GIS and the WWW, Computers, Environment and Urban Systems, 27, 1, pp. 9–32

Ambroise, B., (1991), « Hydrologie des petits bassins versants ruraux en milieu tempéré - processus et modèles », Séminaire du Conseil scientifique du Département Science du Sol de l'INRA, 26-27 mars 1991, Dijon, France, 34 p.

Ambroise, B., (1998), Genèse des débits dans les petits bassins versants ruraux en milieu tempéré : 1- processus et facteurs, Revue des Sciences de l'Eau, 4, pp. 471-495

Ambroise, B. (1999), Genèse des débits dans les petits bassins versants ruraux en milieu tempéré : 2- Modélisation systémique et dynamique, Revue des Sciences de l'Eau, 12, pp. 123-153

Anctil, F., Rousselle, J., Lauzon, N., (2005), *Hydrologie : Cheminements de l'eau*, Presses Internationales Polytechnique, Montréal, Canada, 317 p.

Anderson, M.G., Burt, T.P., (1990), Process studies in hillslope hydrology, Wiley, Chichester, Angleterre, 539 p

Anderson, D., Cobb, J., Korpela, E., Lebofsky, M., Werthimer, D., (2002), SETI@home: An Experiment in Public-Resource Computing, Communications of the ACM, 45, 11, pp. 56-61

Anderson, D., (2004), BOINC : A system for public-resource computing and storage, The 5th IEEE/ACM International Workshop on Grid Computing, Pittsburgh, Etats-Unis, 8 novembre, pp. 4-10

Antikidis, J.-P., Faviera, J.-J., Jeanjeana, H., Hussona, A., (2008), The one meter shop concept: , Acta Astronautica, 63, 1-4, pp. 156-164

Antoine, J.-M., Desailly, B., Gazelle, F., (2001), Les crues meurtrières, du Roussillon aux Cévennes, Annales de Géographie, 110, 622, pp. 597-623

Argent, R.M., (2004), An overview of model integration for environmental applications-components, frameworks and semantics, Environmental Modelling & Software, 19, 3, pp. 219–234

Armstrong, M.P., (1994), GIS and High Performance Computing, GIS/LIS '94 Proceedings, Phoenix, Arizona, 25-27 Octobre 1994, pp. 4-13.

Armstrong, M.P., (1995), Is There a Role for High-Performance Computing in GIS? Urban and Regional Information Systems Association (URISA) journal, 7, 2, pp. 7-10

Arnaud, P., Lavabre, J., (2000), La modélisation stochastique des pluies horaires et leur transformation en débits pour la prédétermination des crues, Revue des sciences de l'eau, 13, pp. 441-462

Asadzadeh, P., Buyya, R., Kei, C., Nayar, D., Venugopal, S., (2005), Global grids and software toolkits: A study of four grid middleware technologies. In : Yang, L., Guo, M., High Performance Computing : Paradigm and Infrastructure, Wiley, New Jersey, Etats Unis, pp. 431-458

Astalos, J., Huchly, L., Dobrucky, M., (2008), Experiences from porting the astrophysical simulation "the unified theory of Kuiper-belt and Oort-cloud formation" to EGEE Grid, 3rd EGEE User Forum, 11-14 Février, Clermont-Ferrand , France, 14 p.

Avery, P., (2009), Data Grids: a new computational infrastructure for data-intensive science, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 360, pp. 1191-1209

Avril, S., Bray, F., Denis, F., Ndiaye, L., Rabbe, L., (2005), « *SIG en ligne : panorama des fonctionnalités et modalités de mise en œuvre* », 17éme rencontre européenne de la géomatique, 30 mars - 1er avril 2005, Porte de Versailles, Paris, 38 p.

Aydin, G., Fox, G.C., Gadgil, H., Pierce, M.E., (2006), Streaming Data Services to Support Archival and Real-Time Geographical Information System Grids, the 6th Annual NASA Earth Science Technology, 7 p.

Aydin, G., Sayar, A., Gadgil, H., Aktas, M.S., Fox, G.C., Ko, S., Bulut, H., Pierce, M.E., (2008), Building and Applying Geographical Information System Grids Concurrency and Computation: Practice and experience, 20, 14, pp. 1653-1695

Ayache, G., (2008), Homo sapiens 2.0 - Introduction à une histoire naturelle de l'hyperinformation, Max Milo Editions, Paris, 284 p.

Ayral P-A., (2001), Contribution à l'élaboration d'une méthodologie d'évaluation de la vulnérabilité face à quatre risques « naturels » particuliers liés à la ressource en eau, Mémoire de DEA, Université Toulouse Le Mirail II, 134 p.

Ayral, P.-A., Griot, C., (2001), *Terminologie en science du risque*, Recueil de définitions, Document préparatoire au Colloque international : Dire le Risque : Le risque en examen, Sous l'égide du Conseil de l'Europe, Mèze, 18-20 mai 2001, 83p.

Ayral, P.-A., (2005), Contribution à la spatialisation du modèle opérationnel de prévision des crues éclairs ALHTAIR - Approche spatiale et expérimentale - Application au bassin du Gardon d'Anduze, Thèse de Doctorat, Université de Provence, Aix-Marseille 1, 300 p.

Ayral, P.-A., Sauvagnargues-Lesage, S., Gay, S., Bressand, F., (2007), *Forecasting Flash Floods with an Operational Model, Application in the South-East of France (Gard).* In : Begum, S., Stive M. J. F., Hall J. W., Flood Risk Management in Europe - Innovation in Policy and Practice, Advances in natural and technological hazards research, Springer Netherlands, Dordrecht, Les Pays-Bas, pp. 335-35

Ayral, P.-A., Sauvagnargues-Lesage, S., (2009), Systèmes d'Information Géographique : outil d'aide à la gestion territoriale, Techniques de l'ingénieur, 16 p.

Ayral, P.-A., Sauvagnargues-Lesage, S., Tena-Chollet, F., Thierion, V., (2010), Les Systèmes d'Information Géographique : mise en œuvre, Techniques de l'ingénieur, *in press.*

Baranski, B., (2008), Grid Computing enabled Web Processing Service. GI Days 2008, 16 - 18 Juin, Münster, Allemagne, 12 p.

Baude, F., (2002), Calcul réparti à grande échelle – Meta-computing, Hermès Science - Lavoisier, Paris, 190 p.

Baxevanidis, K., Davies, H., Foster, I., Gagliardi, F., (2002), Grids and research networks as drivers and enablers of future Internet architectures, Computer Networks, 40, pp. 5-17

Ben Daoud, A., Sauquet, E., Obled, C., Lang, M., Bontron, G., (2009), La prévision des précipitations par recherche d'analogues : état de l'art et perspectives, La Houille Blanche, 6, pp. 60-65

Benoit-Cattin, H., Pera, C., Bellet, F., (2006), Medical image processing web portal: Requirements analysis. An almost end user point of view ..., NA4 meeting: medical imaging sector, Genève, Suisse, 29 Novembre, 7 p.

Berman, F., Fox, G., Hey, T., (2003), The Grid: past, present, future. In : Berman, F., Fox, G., Hey, T., Grid computing: Making the global infrastructure a reality, John Wiley & Sons Ltd, Chichester, Angleterre, pp. 9-50

Berners-Lee, T., Cailliau, R., Luotonen, A., Nielsen, H., Secret, A., (1994), "The World-Wide Web", Communications of the ACM, 37, 8, pp. 76-82

Berners-Lee, T. (2007), « Giant Global Graph ». In *Decentralized Information Group*. En ligne. http://dig.csail.mit.edu/breadcrumbs/node/215. Consulté en 2010

Beven K.J., Kirby M.J., (1979), A physically based, variable contributing area model of *basin hydrology*, Hydrological Sciences, 24, 1, pp. 43-69

Beven, K., (2001), Rainfall-runoff modelling: the primer. Wiley, Chichester, 360 p.

Bleuse, N. (2001), Étude de définition et maquette d'un SIG opérationnel pour le Service d'Annonce des crues du Gard, Mémoire de Mastère, SILAT, École nationale supérieure d'agronomie de Montpellier 28 p.

Bonnifait, L., Delrieu, G., Le Lay, M., Boudevillain, B., Masson, A., Belleudy, P., Gaume, E., Saulnier, G.-M., (2009), Distributed hydrologic and hydraulic modelling with radar rainfall input: Reconstruction of the 8–9 September 2002 catastrophic flood event in the Gard region, France, Advances in Water Resources, 32, 7, pp. 1077-1089

Bordin, P., (2002), SIG - concepts, outils et données, Lavoisier, Paris, 259 p.

Bouras, C., Giannaka, E., Tsiatsos, T., (2008), e-Collaboration Concepts, Systems, and Applications. In : Freire, M.M., Periera, M., Encyclopedia of Internet Technologies and Applications, Information Science Reference, pp. 8-16

Bouvier, C., Delclaux, F., (1996), *ATHYS: a hydrological environment for spatial modelling and coupling with GIS***.** In : Kovar, K., Nachtnebel, H. P. (eds.), Application of geographic information systems in hydrology and water resources management**,** IAHS Publication**,** Wallingford, Royaume-Uni **,**pp. 19-28

Bouvier, C., Marchandise, A., Lequien, A., Brunet, P., Crespy, A., (2004), « *Distributed rainfall/runoff modelling of september 2002 flood in 2 southern France river basins* », BALWOIS 2004, 25-29 Mai 2004, Ohrid, République de Macédoine, 12 p.

Bovolo I., C., Abele, S., Bathurst, J., Caballero, D., Ciglan, M., Eftichidis, G., Simo, B., (2008), A distributed framework for multi-risk assessment of natural hazards used to model the effects of forest fire on hydrology and sediment yield, Computers and Geosciences, 35, pp. 924-945

Bressand, F., (2002), Le projet ALHTAÏR du service d'annonce des crues du Gard, La Houille Blanche, 2, pp. 64 - 68

Breton, V., Dean, K., Solomonides, T., (2005), The Healthgrid white paper, Studies in health technology and informatics, 112, pp. 249-321

Breton, V., (2009), « WISDOM: In silico docking against neglected and emerging diseases », IX International Conference on Science, Arts and Culture, Grid Computing : a new tool for Science and Innovation, Veli Lošinj, Croatie, 25-29 Août, 51 p.

Britton, D., Cass, A., Clarke, P.,Coles, J., Colling, D., Doyle, A., Geddes, N., Gordon, J., Jones, R., Kelsey, D., Lloyd, S., Middleton, R., Patrick, G., Sansum, R., Pearce, S., (2009), GridPP : the UK grid for particle physics, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 367, pp. 2447-2457

Brochu, F., Egede, U., Elmsheuser, J., Harrison, K., Jones, RWL, Lee, HC, Liko, D., Maier, A., Moscicki, J., Muraru, A., Patrick, G., Pajchel, K., Reece W., Samset B., Slater M., Soroko A., Tan C., Vanderster, D., Williams, M., (2009), Ganga : a tool for computational-task management and easy access to Grid resources Computer Physics Communications (submitted)

Broersen, P.M.T, (2007), Error Correction of Rainfall-Runoff Models With the ARMAsel Program, IEEE Transactions on instrumentation and measurement, 56, 6, pp. 2212-2219

Brooke, J.M., Parkin, M.S., (2009), Enabling scientific collaboration on the Grid, Future Generation Computer Systems, 26, 3, pp. 521–530

Brugnot, G., (2001), Gestion spatiale des risques, Lavoisier, Paris, 287 p.

Brunet, R., (1987), La carte, mode d'emploi, Fayard/Reclus, Paris, 270 p.

Burck, R., Gabolde, S., Lhomme, S., Michelon, O., Olive, G., (2008a), Cahier des charges technique : Cartographie Hydro-Météorologique Temps réel, Cycle des Ingénieurs diplômés de l'ENSG 2ème année (IT2), 50 p.

Burck, R., Gabolde, S., Lhomme, S., Michelon, O., Olive, G., (2008b), Guide programmeur : Cartographie Hydro-Météorologique Temps réel, Cycle des Ingénieurs diplômés de l'ENSG 2ème année (IT2), 69 p.

Burke, S., Campana, S., Lanciotti, E., Méndez Lorenzo, P., Miccio, V., Nater, C., Santinelli, R., Sciabà, A., (2009), gLite 3.1 User Guide, Manuals Series, LHC Computing Grid, 168 p.

Burrough, P.A., (1986), Principles of geographical information systems for land resources assessment, Oxford University Press, New York, Etats-Unis, 194 p.

Cai, Z., Weng, M., Li, Y., Du, Q., (2009), A system architecture of GIS middleware support for context-sensitivity, International Symposium on Spatial Analysis, Spatial-Temporal Data Modeling, and Data Mining, 13-15 Octobre, Wuhan, Chine

Caporali, E., (2007), A distributed hydrological model of flash-floods. In: Gruntfest, E. and Handmer, J. (eds.) Coping with flash floods, Kluwer Academic Publishers, Dordrecht, pp. 203-218

Cappello, F., (2006), Projet Grid'5000, Techniques de l'Ingénieur, IN48, 11p.

Cartwright, W., (1999), Development of multimedia. In : Cartwright, W., Peterson, M.P., Gartner, G.F., Multimedia cartography, Springer, Berlin, Allemagne, pp.11-26

Cartwright, W., (2005), Maps on the web. In : Peterson, M., Maps and the Internet, Elsevier Science Ltd, Oxford, Royaume-Uni, pp. 35-56

Castaings, W., Dartus, D., Honnorat, M., Dimet, F.X.L., Loukili, Y., Monnier, J., (2006), Automatic differentiation: a tool for variational data assimilation and adjoint sensitivity analysis for flood modeling. In : Bücker, M., Corliss, G., Naumann, U., Hovland, P., Norris, B., Automatic Differentiation: Applications, Theory, and Implementations, Lecture Notes in Computational Science and Engineering, 50, Springer, Berlin, Heidelberg, pp. 249-262

Castells, M., (1998), La société en réseaux - l'ère de l'information, Fayard, Paris, 671 p.

Castrogiovanni, E.M., La Loggia, G., Noto, L.V., (2005), Design storm prediction and hydrologic modeling using a web-GIS approach on a free-software platform, Atmospheric research, 77, 1-4, pp. 367-377

Caumont, O., Ducrocq, V., Delrieu, G., Gosset, M., Pinty, J.-P., Parent du Châtelet, J., Andrieu, H., Lemaître, Y., Scialom, G., (2006), Journal of Atmospheric and Oceanic Technology, 23, pp. 1049-1067

Chassagneux, P., Dolmière, J.-M., (2007), La vigilance « pluie – inondation », Congrès AFPCN, journées « Prévision, Vigilance et Alerte, 14-15 Novembre 2007, Toulouse, France

Chervenak, A., Foster, I., Kesselman, C., Salisbury, C., Tuecke, S., (2000), The data grid: Towards an architecture for the distributed management and analysis of large scientific datasets, Journal of Network and Computer Applications, 23, 3, pp. 187-200

Chiang, Y.-M., Hsu K.-L., Chang F.J, Hong Y., Sorooshian S., (2007), Merging multiple precipitation sources for flash flood forecasting, Journal of Hydrology, 340, pp. 183-196

Childs, S., (2007), Supporting MPI applications on the EGEE Grid, EGEE User Forum, 7-11 Mai, Manchester, Royaume-Uni, 15 p.

Christodoulopoulos, K., Gkamas, V., Varvarigos, E.A., (2008), Statistical Analysis and Modeling of Jobs in a Grid Environment, Journal of Grid Computing, 6, 1, pp. 77-101

Chollet, F., Hernandez, F., Malek, F., Shifrin, G., (2007), Actes du 2ème Colloque LCG-France, Clermont-Ferrand, 14 et 15 mars, 22 p.

Chollet, M., Rivière, P., (2010), Internet, révolution culturelle, Le Monde diplomatique – Manière de voir, février-mars 2010, 109, 98 p.

Clarke, K.C., Guan, Q., Zhang, T., (2006), The Anatomy and Physiology of the GeoGrid, AutoCarto 2006, 26-28 Juin, Vancouver, Etats-Unis, 13 p.

Claudet, R., Bouvier, C., (2004), Outils de prévision des crues rapides : les besoins de l'alerte et du suivi en temps réel, Colloque d'hydrotechnique « Crues méditerranéennes – Vulnérabilité, Aléa, Prévisions et gestion de crise », 176ème session du comité scientifique et technique, Société hydrotechnique de France, 15-16 juin 2004, Nîmes, France, pp. 105-112

Collier, C.G., (2007), Flash flood forecasting : What are the limits of predictability ? Quarterly Journal of the Royal Meteorological Society, 133, pp. 3-23

Colosimo, C., Mendicino, G., (1996), *GIS for distributed rainfall-runoff modeling*. In : Singh, V. P., Fiorentino, M., Geographical information systems in hydrology, Kluwer Academic Publishers, Dordrecht, Les Pays-Bas, pp. 195-236

CONDOR team, (2009), Condor® Version 7.2.4 Manual, University of Wisconsin–Madison, 962 p.

Contes, A., (2005), Une architecture de sécurité hiérarchique, adaptable et dynamique pour la grille, Thèse de Doctorat, Université de Nice – Sophia Antipolis, Sciences et Technologies de l'Information et de la Communication, 160 p.

Cosandey, C., Bigot, S., Dacharry, M., Gille, E., Laganier, R., Salvador, P.-G., (2003), *Les eaux courantes : Géographie et environnement* Éditions Belin, Paris, 240 p.

Cova, T., (1999), GIS in emergency management. In: Longley, P., Goodchild, M., Maguire, D., Rhind, D., Geographical Information Systems: Principles, Techniques, Applications and Management, 2nd Edition, John Wiley & Sons Inc., New York, Etats-Unis, pp. 845-858

Coveney, P.V., (2005), Scientific grid computing, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, The Royal Society, 363, pp. 1707-1713

Craglia, M., Grifiths, M., Costello, A., (2005), From SDIs towards the Spatial Grid, GIS Planet 2005 Conference, 30 Mai - 2 Juin, Estoril, Portugal, 14 p.

Cullot, N., Parent, C., Spaccapietra, S., Vangenot, C., (2003), Des ontologies pour données géographiques. Revue Internationale de Géomatique, 13, 3, pp. 285-306

Cuthbert, A., (1999), OpenGIS : Tales from a small Market Town, Lecture notes in computer science, Actes de la conférence INTEROP'99 : interoperating geographic information systems, 10-12 Mars 1999, Zurich, Suisse, pp. 17-28

Cutter, S.L., (2003), GI Science, Disasters, and Emergency Management, Transactions in GIS, 7, 4, pp. 439-445

CYber-Infrastructure for CiviL protection Operative ProcedureS (CYCLOPS), (2006), Annex I - "Description of Work", SIXTH FRAMEWORK PROGRAMME, Research Infrastructure – Communication Network Development, 112 p.

D'Anfray, P., Jeannot, E., (2007), « GRID'5000 une plate-forme d'expérimentation pour les systèmes distribués à large échelle », Actes de la conférence JRES, 6 p.

Daniel, J., (2003), Services Web - concepts, techniques et outils, Vuibert, Paris, 341 p.

Dauphiné, A., (2001), Risques et catastrophes : observer, spatialiser, comprendre, gérer, Armand Colin, Paris, 288 p.

Dautun, C., (2007), *Contribution à l'étude des crises de grande ampleur : Connaissance et aide à la décision pour la Sécurité Civile*, Thèse de Doctorat, Science et Génie de l'environnement, Ecole Nationale Supérieure des Mines Saint-Etienne, 408 p.

DDSC – Direction de la défense et de la sécurité civiles, (2001), Circulaire NORINTE0100268C du 28 septembre 2001 relative à la refonte de la procédure d'alerte météorologique - procédure Vigilance, récupéré en septembre 2009 de
http://www.anena.org/jurisque/reglement3/systalert/c280901vigil.htm

DEGREE Project Consortium, 2008, Earth science white paper on grids, ES Grid Roadmap Deliverable D6.1.2, DEGREE-WP6-D6.1, 63 p.

De la Beaujardiere, J., (2006), OpenGIS Web Map Server Implementation Specification, Version 1.3.0, OpenGIS Consortium Inc., Etats-Unis, 85 p.

Delrieu, G., Ducrocq, V., Gaume, E., Nicol, J., Payrastre, O., Yates, E., Kirstetter, P., Andrieu, H., Ayral, P.-A., Bouvier, C., (2005), The catastrophic flash-flood event of 8--9 September 2002 in the Gard region, France: a first case study for the Cévennes--Vivarais Mediterranean Hydrometeorological Observatory, Journal of Hydrometeorology, 6, pp. 34-52

Delrieu, G., Braud, I., Berne, A., Borga, M., Boudevillain, B., Fabry, F., Freer, J., Gaume, E., Nakakita, E., Seed, A., Tabary, P., Uijlenhoet, R., (2009), Weather radar and hydrology, Advances in Water Resources, 32, 7, pp. 969-974

Denègre, J., Salgé, F., (2004), Les Systèmes d'information géographique, Que sais-je ?, Presses Universitaires de France, Paris, 128 p.

Densham, P.J., (1991), Spatial decision support systems. In : Maguire, D.J., Goodchild, M.F., Rhind, D.W., Geographical information systems : principles and applications, John Wiley and Sons, New York, Etats-Unis, pp. 403 - 412

De Roure, D., Baker, M., Jennings, N., Shadbolt, N., (2003), The evolution of the Grid. In : Berman, F., Fox, G., Hey, T., Grid computing: Making the global infrastructure a reality, John Wiley & Sons Ltd, Chichester, Angleterre, pp. 65-100

Di, L. (2006), The geospatial grid. In : Rana, S., Sharma, J., Frontiers of geographic information technology, Springer-Verlag, Berlin, Allemagne, pp. 121-137

Di, L., Chen, A., Yang, W., Liu, Y., Wei, Y., Mehrotra, P., Hu, C., Williams, D. (2008), The development of a geospatial data Grid by integrating OGC Web services with Globus-based Grid technology. Concurrency and Computation: Practice and Experience, 20, 14, pp. 1617-1635

Díaz, L., Costa, S., Granell, C., Gould, M., (2007), Migrating geoprocessing routines to web services for water resource management applications, AGILE 2007 Conference, 13-17 Août, Washington DC, Etats-Unis, 8 p.

DIREN – Direction Régionale de l'Environnement, (2005), Schéma directeur de prévision des crues – bassin Rhône Méditerranée, 71 p.

Domenico, B., Nativi, S., (2009), Web Coverage Service (WCS) 1.1 extension for CF-netCDF 3.0 encoding, OpenGIS Discussion Paper, Version 0.2.2, OpenGIS Consortium Inc., Etats-Unis, 74 p.

Doswell, C.A., Brooks, H.E., Maddox, R.A. (1996), Flash flood forecasting: an ingredient-based methodology, weather and forecasting, 11, pp. 560-581

Dransch, D., (1999), Theoretical issues in multimedia cartography. In : Cartwright, W., Peterson, M.P., Gartner, G.F., Multimedia cartography, Springer, Berlin, Allemagne, pp. 41-48

Drobot, S., Parker, D.J., (2007), Advances and challenges in flash flood warnings, Environmental hazards, 7, 3, pp. 173-178

Duband, D., (2000), Réflexion actuelle sur la prévision hydrométéorologique, Houille blanche, 1, pp. 42-50

Duflot, F., (2007), L'interopérabilité dans tous ses états, Revue Lamy « Droit de l'immatériel », supplément n° 23, pp. 9-12

Dupont, Y., (2003), *Dictionnaire des risques*, Armand Colin, Paris, 421 p.

Eaton, B., Gregory, J., Drach, B., Taylor, K., Hankin, S., (2009), « NetCDF Climate and Forecast (CF) Metadata Conventions ». In : *CF Metadata - Climate and Forecast Metadata Convention*. En ligne. http://cf-pcmdi.llnl.gov/documents/cf-conventions/1.4/cf-conventions-multi.html. Consulté en 2010.

EGEE - Enabling Grids for E-sciencE, (2009), Success and Sustainability: EGEE Collaborating Projects' Achievements in 2009 and Future Plans, 36 p.

ESFRI - European Strategy Forum on Research Infrastructures, (2008), European roadmap for research infrastructures, European Communities, 104 p.

Espinal, X., Campana, S., Walker, R., (2008), Large-scale ATLAS production on EGEE, Journal of Physics : Conference Series, 119, 10 p.

Estupina Borrell, V., (2004), Vers une modélisation hydrologique adaptée à la prévision opérationnelle des crues éclair : Application à de petits bassins versants du sud de la France, Thèse de Doctorat, Science de l'Univers, de l'Environnement et de l'Espace, Institut National Polytechnique, Toulouse, 253 p.

Estupina-Borrell, V., Dartus, D., Ababou, R., (2006), Flash flood modeling with the MARINE hydrological distributed model, Hydrology and Earth System Sciences Discussions, 3, p. 3397–3438

Faure, D., Andrieu, H., Creutin, J., (1994), Application à l'hydrologie du radar météorologique (comparaison d'estimations radar et pluviométriques pour des lames d'eau horaires sur de petits bassins versants cévenols), Etudes et recherches des LPPC, 292 p.

Fedra, K., (1998), Integrated risk assessment and management: an overview and state of the art, Journal of hazardous materials, 61, 1-3, pp. 5-22

Feidas, H., Kontos, T., Soulakellis, N., Lagouvardos, K., (2007), A GIS tool for the evaluation of the precipitation forecasts of a numerical weather prediction model using satellite data, Computers and Geosciences, 33, 8, pp. 989—1007

Field, E. H., Gupta, V., Gupta, N., Maechling, P., Jordan, T. H., (2005), Hazard map calculations using grid computing, Seismological Research Letters, 76, 2, pp. 565-573

Field, L., Huang, J., Tsai, M., (2009), GStat 2.0: Grid Information System Status Monitoring, 17th International Conference on Computing in High Energy and Nuclear Physics, 21 - 27 Mars, Prague, République Tchèque, 7 p.

Fischer, C., Montmerle, T., Auger, L., Lacroix, B., (2006), *L'assimilation opérationnelle de données régionales à Météo-France*, Météorologie, 54, pp. 43-48

Flynn, M., (1972), Some computer organizations and their effectiveness, IEEE Transactions on Computers, 21, 9, pp. 948-960

Folino, G., Forestiero, A., Papuzzo, G., Spezzano, G., (2005), MOSE: A grid-enabled software platform to solve geoprocessing problems, Nuovo Cimento della Societa Italiana di Fisica sect. C, 28, pp. 225-235

Foster, I., (1995), Designing and building parallel programs: concepts and tools for parallel software engineering, Addison Wesley, Boston, Etats-Unis, 430 p.

Foster, I., Kesselman, C., (1997), Globus : A metacomputing infrastructure toolkit. International Journal of Supercomputer Applications, 11, 2, pp. 115-128

Foster, I., Geisler, J., Nickless, W., Smith, W., Tuecke, S. (1997), Software infrastructure for the I-WAY high-performance distributed computing experiment, Proc. 5th IEEE Symposium on High Performance Distributed Computing, pp. 562-571

Foster, I., Kesselman, C., (1999), Computational Grids. In : Foster, I., Kesselman, C., The Grid: Blueprint for a New Computing Infrastructure, Morgan-Kaufman, San Francisco, Etats-Unis, pp. 15-51

Foster, I., Kesselman, C., Tuecke, S., (2001), The Anatomy of the Grid : Enabling Scalable Virtual Organizations, International Journal Supercomputer Applications, 15, 3, pp. 200-222

Foster, I., Kesselman, C., Nick, J., Tuecke, S., (2002), « The Physiology of the Grid: An Open Grid Services Architecture for Distributed Systems Integration », Open Grid Service Infrastructure WG, Global Grid Forum, 22 juin 2002, 31 p.

Foster, I., Kesselman, C., (2004), Concepts and architecture. In : Foster, I., Kesselman, C., The Grid 2 : Blueprint for a New Computing Infrastructure, Morgan Kaufmann, San Francisco, Etats-Unis, pp. 37-63

Fox, G.C., (2004), Grids of Grids of Simple Services, Computing in Science & Engineering, 6, 4, pp. 84-87

Francica, J., (2007), « Neogeography is not GIS; not LI, journal en ligne ». In *All Points Blog*. En ligne. http://apb.directionsmag.com/archives/3703-Neogeography-is-not-GIS;-not-LI.html. Consulté en décembre 2009.

Gadgil, H., Choi, J., Engel, B., Fox, G., Ko, S., Pallickara, S., Pierce, M., (2004), Management of data streams for a real time flood simulation, Community Grids Lab Technical Report, 8 p.

Gagliardi, F., Jones, B., Grey, F., Begin, M., Heikkurinen, M., (2005), Building an infrastructure for scientific Grid computing: status and goals of the EGEE project, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, The Royal Society, 363, 1729-1742

Gaume, E., (2002), *Eléments d'analyse sur les crues éclair*, Thèse de Doctorat, Ecole Nationale du Génie Rural, des Eaux et des Forêts, 359 p.

Gaume, E., (2004), Analyse hydro-pluviométrique de la formation des crues éclair dans les bassins versants supérieurs du Gard et du Vidourle, Colloque d'hydrotechnique « Crues méditerranéennes – Vulnérabilité, Aléa, Prévisions et gestion de crise », 176ème session du comité scientifique et technique, Société hydrotechnique de France, 15-16 juin 2004, Nîmes, France, pp. 55-62

Gaume, E., Livet, M., Desbordes, M., Villeneuve, J., (2004), Hydrological analysis of the river Aude, France, flash flood on 12 and 13 November 1999, Journal of Hydrology, 286, pp. 135-154

Geisel, J., (2009), Intégration et développement d'une plateforme de modélisation hydrologique par la technologie grille, L3 NTIE, Université de Toulouse II - Le Mirail, 42 p.

GéoConnexions, (2004), Manuel pour les développeurs de l'ICDG : produire et publier l'information, les données et les services géographiques. En ligne. http://www.geoconnections.org/publica- tions/Technical_Manual/html_f/. Consulté en 2010

Gengler, M., Ubeda, F., Desprez, F., (1996), Initiation au parallélisme - concepts, architectures et algorithmes, Masson, Paris, 223 p.

Germain-Renaud, C., Loomis, C., Mościcki, J.T., Texier, R., (2008), Scheduling for Responsive Grids, Journal of Grid Computing, 6, 1, pp. 15-27

Ghanem, M., Guo, Y., Hassard, J., Osmond, M., Richards, M., (2004), Sensor Grids for Air Pollution Monitoring, Proceedings of 3rd UK e-Science All Hands Meeting, 31 août - 3 Septembre 2004, Nottingham, Angleterre, 8 p.

Gjermundrød, H., Dikaiakos, M., Stümpert, M., Wolniewicz, P., Kornmayer, H., (2008), « g-Eclipse - An Integrated Framework to Access and Maintain Grid Resources », The 9th IEEE/ACM International Conference on Grid Computing, Tsukuba, Japan, 29 septembre – 1er octobre, pp. 57-64

Gnouma, R., (2006), Aide à la calibration d'un modèle hydrologique distribué au moyen d'une analyse des processus hydrologiques : application au bassin versant de l'Yzeron, Mécanique, Energétique, Génie Civil et Acoustique, Institut National des Sciences Appliquées de Lyon, 412 p.

Goodchild, M.F., (1992), Geographical information science, International journal of geographical information systems, 6, 1, pp. 31-45

Goscinski, A., (1991), Distributed Operating Systems: The Logical Design, Addison-Wesley, Boston, Etats-Unis, 912 p.

Grimshaw, D., (1994), Bringing geographical information systems into business, John Wiley & Sons, New York, Etats-Unis, 273 p.

Gruntfest, E., (2007), Beyond flood detection: alternative applications of real-time data. In : Gruntfest, E. and Handmer, J. (eds.) Coping with flash floods, Kluwer Academic Publishers, Dordrecht, pp. 167-179

Gruntfest, E., Handmer, J., (2007), Dealing with flash floods: contemporary issues and future possibilities. In: Gruntfest, E. and Handmer, J. (eds.) Coping with flash floods, Kluwer Academic Publishers, Dordrecht, pp. 3-10

Gupta, H., Sorooshian, S., Gao, X., Imam, B., Hsu, K.-L., Bastidas, L., Li, J., Mahami, S., (2002), The challenge of predicting flash floods from thunderstorm rainfall, Philosophical Transactions of the royal society A: Mathematical, Physical and Engineering Sciences, 360, 1796, pp. 1363-1371

Gurney, R.J., Badcock, N.S., Garnett, E.A., Godfray, H.C.J., (2009), The environmental eScience revolution, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 367, pp. 803-813

Handmer, J., Henson, R., Sneeringer, P., Konieczny, R., Madej, P., (2007), Warning systems for flash floods: research needs, opportunities and trends. In: Gruntfest, E. and Handmer, J. (eds.) Coping with flash floods, Kluwer Academic Publishers, Dordrecht, pp. 77-89

Hartnett, E., (2009), NETCDF-4 and high performance computing, the 10th LCI international conference on high-performance clustered computing, 9-12 Mars, Boulder, Colorado, Etats-Unis, 91 p.

Hawick, K.A., Coddington, P.D., James, H.A., (2003), Distributed frameworks and parallel algorithms for processing large-scale geographic data, Parallel Computing, 29, 10, pp. 1297-1333

Haziza, E., (2007), *Contribution au développement de stratégies préventives face au risque de crue éclair*, Thèse de Doctorat, Sciences et métiers de l'Ingénieur, Ecole National Supérieure des Mines de Paris, Sophia-Antipolis, 336 p.

Hey, T., Trefethen, A., (2002), The UK e-science core programme and the grid, Future Generation Computer Systems, 18, pp. 1017-1031

Horton R.E., (1933), *The role of infiltration in the hydrological cycle*, Transactions of the American Geophysical Union, 14, pp. 446-460

Houdré, F., (2001), *L'annonce des crues. Histoires et évolution des services de 1847 à nos jours*, Ministère de l'Aménagement du Territoire et de l'Environnement (MATE), Rapport de stage, Ecole nationale supérieure d'hydraulique et de mécanique de Grenoble, 47 p.

Hluchy, L., Astalos, J., Dobrucky, M., Habala, O., Simo, B., Tran, V. D., (2003), Flood forecasting in a grid computing environment, Parallel Processing and Applied Mathematics, 3019, pp. 831-839

Hluchý, L., Habala, O., Tran, V., Gatial, E., Mališka, M., Šimo, B., Slížik, P., (2005), Collaborative environment for grid-based flood prediction, Computing and Informatics, 24, 1, pp. 1001–1022

Honnorat, M., Le Dimet, F.-X., Loukili, Y., Monnier, J., (2005), Dassflow : Un modè le direct et inverse d'écoulements Saint-Venant 2D, Rapport de recherche, INRIA, 24 p.

Hrebícek, J., **Konecný, M**., (2007), Introduction to ubiquitous cartography and dynamic geovisualization with implications for disaster and crisis management. In : Scharl, A., Tochtermann, K., The Geospatial Web. How Geobrowsers, Social Software and the Web 2.0 are Shaping the Network Society, Advanced Information and Knowledge Processing Series, Springer, Londres, Angleterre, pp. 209-214

Hu, W., Qing, Y., Ming-hui, Y., Qi, F., (2008), Grid-based platform for disaster response plan simulation over the Internet, Simulation modeling practice and theory, 16, 3, pp. 379-386

Huedo, E., Montero, R.S., Llorente, I.M., (2004), A framework for adaptive execution in grids, Software: Practice and Experience, 34, 7, pp. 631 - 651

Huedo, E., Montero, R., Llorente, I. M., (2006), Evaluating the reliability of computational grids from the end user's point of view, journal of Systems architecture, 52, 12, pp. 727-736

Huet, P., Martin, X., Prime, J.-L., Foin, P., Laurain, C., Cannard, P., (2003a), Retour d'expérience des crues de septembre 2002 dans les départements du Gard, de l'Hérault, du Vaucluse, des Bouches-du-Rhône, de l'Ardèche et de la Drôme. Rapport de l'Inspection Générale de l' Environnement (IGE), MEDD, 129 p.

Huet, P., Foin, P., Laurain, C., Martin, X., Prime, J.-L., (2003b), Retour d'expérience sur les inondations du Sud-Est de septembre 2002, Annales des Mines, Mai 2003, pp. 63-72

Huet, P. (2005), *La méthodologie des retours d'expériences après les accidents naturels,* Ministère de l'Environnement et du Développement Durable, 85 p.

Hughes, D., Greenwood, P., Coulson, G., Blair, G., (2006), GridStix: supporting flood prediction using embedded hardware and next generation grid middleware, International symposium on a world of wireless mobile and multimedia networks, 26-29 Juin 2003, Buffalo, Etats-Unis, pp. 621-626

IDG - Institut Des Grilles, (2009), Livre blanc sur les grilles de production, CNRS, 71 p.

ISTED - Institut des Sciences et des Techniques de l'Equipement et de l'Environnement pour le Développement, (2005), Systèmes d'information géographique et gestion des risques, ISTED, Paris, 101 p.

Jacq, N., Salzemann, J., Legré, Y., Reichstadt, M., Jacq, F., Medernach, E., Zimmermann, M., Maaß, A., Sridhar, M., Vinod-Kusam, K., Montagnat, J., Schwichtenberg, H., Hofmann, M., Breton, V., (2008), Grid enabled virtual screening against malaria, Journal of Grid Computing, 6, 1, pp. 29-43

Jankowski, P., Robischon, S., Tuthill, D., Nyerges, T., Ramsey, K., (2006), Design Considerations and Evaluation of a Collaborative, Spatio-Temporal Decision Support System, Transactions in GIS, 10, 3, pp. 335-354

Jeffery, K.J., (2007), Next generation GRIDs for environmental science, Environmental modelling and software, 22, 3, pp. 281-287

Jeanvoine, E., (2007), Intergiciel pour l'exécution efficace et fiable d'applications distribuées dans des grilles dynamiques de très grande taille, Thèse de Doctorat, Matisse, Université de Rennes 1, 182 p.

Jiang, B., (2005), Beyond serving maps : serving GIS functionality over the Internet. In : Peterson, M., Maps and the Internet, Elsevier Science Ltd, Oxford, Royaume-Uni, pp. 147-158

Joliveau, T., (1996), Gérer l'environnement avec des SIG. Mais qu'est-ce qu'un SIG ?, Revue de géographie de Lyon, 71, 2, pp. 101-110

Jones, B., (2009), EGI et les infrastructures de recherches ESFRI, Journées « Grilles France », 14 16 Octobre, Lyon, France, 16 p.

Kelsch, M., Caporali, E., Lanza, L., (2001), *Hydrometeorology of flash floods*. In : Gruntfest, E. and Handmer, J. (eds.) Coping with flash floods, Kluwer Academic Publishers, Dordrecht, pp. 19-35

Klopfer, M., (2005), Interoperability & open architectures: An analysis of existing standardization - processes & procedures, Open Geospatial Consortium Inc., OGC White Paper, 26 p.

Knoops, E., (2007), Sécurité sur le GRID, Tutorial utilisateur EGEE, 20 Mars, Lyon, 38 p.

Köbben, B., (2001), Publishing maps on the web. In : Kraak, M.-J., Brown, A., Web Cartography developments and prospects, Taylor & Francis, Londres, Angleterre, pp. 73-86

Kobiyama, M., Goerl, R., (2007), Quantitative method to distinguish flood and flash flood as disasters, SUISUI Hydrological Research Letters, 1, pp. 11-14

Kolodziej, K., (2004), OpenGIS® Web Map Server Cookbook, Version 1.0.2, OpenGIS Consortium Inc., Etats-Unis, 42 p.

Komornicki, A., Mullen-Schulz, G., Landon, D., (2009), Roadrunner: Hardware and Software Overview, IBM redbooks, 50 p.

Korkhov, V.V., Moscicki, J.T., Krzhizhanovskaya, V.V., (2009),Dynamic workload balancing of parallel applications with user-level scheduling on the Grid, Future Generation Computer Systems, 25, 1, pp. 28–34

Kotonya, G., Sommerville, I., (1998), Requirements Engineering: Processes and Techniques, John Wiley & Sons Ltd, New York, Etats-Unis, 294 p.

Kraak, M.-J., Brown, A., (2001), Web Cartography developments and prospects, Taylor & Francis, Londre, Angleterre, 213 p.

Krakowiak, S., (2008), Intergiciel et Construction d'Applications Réparties, Chap. 1 : Introduction à l'intergiciel, Rapport technique ICAR, 393 p.

Kralidis, A.T., (2007), Geospatial Web Services: The Evolution of Geospatial Data Infrastructure. In : Scharl, A., Tochtermann, K., The Geospatial Web. How Geobrowsers, Social Software and the Web 2.0 are Shaping the Network Society, Advanced Information and Knowledge Processing Series, Springer, Londres, Angleterre, pp. 223-228

Kshemkalyani, A., Singhal, M., (2008) Distributed computing: principles, algorithms, and systems, Cambridge University Press, Cambridge, Royaume-Uni, 752 p.

Kunszt, P., (2003), European DataGrid project: status and plans, Nuclear Inst. and Methods in Physics Research section A, 502, 2-3, pp. 376-381

Kunz, C., Gröper, R., Fleuren, T., Grimm, C., (2009), Geospatial workflows on the D-Grid infrastructure, Open Grid Forum 25, 2-6 Mars, Catane, Italie, 31 p.

Kurzbach, S., Pasche, E., Lanig, S., Zipf, A., (2009), Benefits of grid computing for flood modeling in service-oriented spatial data infrastructures, GIS.Science, 3, pp. 89-97

Kussul, N., Shelestov, A., Skakun, S., (2009), Grid and sensor web technologies for environmental monitoring, Earth Science Informatics, 2, 1-2, pp. 37-51

Lake, R., Farley, J., (2007), Infrastructure for the geospatial web. In : Scharl, A., Tochtermann, K., The Geospatial Web. How Geobrowsers, Social Software and the Web 2.0 are Shaping the Network Society, Advanced Information and Knowledge Processing Series, Springer, Londres, Angleterre, pp. 15-26

Lapeyre, G., (2006), *Production de la carte nationale de vigilance « crues » avec Mapserver*, Mémoire de Master professionnel SIGMA, Université Toulouse Le Mirail II et Ecole Nationale Supérieure Agronomique de Toulouse (ENSAT), 46 p.

Lavabre, J., Fouchier, C., Folton, N., Gregoris, Y., (2003), SHYREG : une méthode pour l'estimation régionale des débits de crue. Application aux régions méditerranéennes françaises, Ingénieries, spécial Risques naturels et aménagement du territoire, pp. 97-111

Lavabre, J., Grégoris, Y., (2005), *AIGA: un dispositif d'alerte des crues sur l'ensemble du réseau hydrographique*, Ingénieries, Centre National du Machinisme Agricole, du Génie Rural, des Eaux et Forêts, pp. 3-12

Laure, E., Hemmer, F., Prelz, F., Beco, S., Fisher, S., Livny, M., Guy, L., Barroso, M., Buncic, P., Kunszt, P., Di Meglio, A., Aimar, A., Edlund, A., Groep, D., Pacini, F., Sgaravatto, M., Mulmo, O., (2004), « Middleware for the next generation Grid infrastructure », Computing in High Energy Physics and Nuclear Physics 2004, Interlaken, Suisse, 27 Sep - 1 Oct 2004, 4p.

Laurini, R., Thompson, D., (1992), Fundamentals of spatial information systems, Academic press, Londre, Angleterre, 680 p.

Lecca, G., Lai, C., Murgia, F., Biddau, R., Fanfani, L., Maggi, P., (2009), AQUAGRID: an extensible platform for collaborative problem solving in groundwater protection, Earth Science Informatics, 2, 1-2, pp. 83-95

Lee, G., Percivall, G., (2008), Standards-Based Computing Capabilities for Distributed Geospatial Applications, Computer, 41, 11, pp. 50-57

Le Guigner, J., (1994), RENATER : réseau national de la technologie, de l'enseignement et de la recherche Bulletin des bibliothèques de France, Bulletin des bibliothèques de France, 39, 1, 39-44

Lefèvre, C., Schneider, J. L., (2003), *Les risques naturels majeurs*, GB Science, Paris, 306 p.

Lefrou, C., Martin, X., Labarthe, J., Varret, J., Maziere, B., Tordjeman, R., Feunteun, R., (2000), Les crues des 12, 13 et 14 novembre 1999 dans les départements de l'Aude, de l'Hérault, des Pyrénées-Orientales et du Tarn, Rapport au Ministre de l'Aménagement du Territoire et de l'Environnement, 140 p.

Leprince, F., Rabu, D., Quesseveur, E., (2003), Les usages de l'information géographique « en ligne » Essai de typologie et exemples d'application, Revue Internationale de Géomatique, 13, 3, pp. 307-322

Lévy, P., (1997), L'intelligence collective - Pour une anthropologie du cyberspace, La Découverte, Paris, 246 p.

Lien, H.-C., Shiau, Y.-H., Huang, C.-P., Wu, J.-H., Tsai, W.-F., (2004), The integration and application of the computational grid structure in flood forecast system, 7th International conference on high performance computing and grid in Asia pacific region, 20-22 Juillet, Tokyo, Japon, pp. 328-331

Longley, P., Goodchild, M., Maguire, D., Rhind, D., (2005), Geographic information systems and science, John Wiley & sons, Ltd, Chichester, Angleterre, 454 p.

Malczewski, J., (1997), Spatial Decision Support Systems, NCGIA Core curriculum in GIScience, 9 p.

Marques, R., Thierion, V., (2008), Parallelization and grid issues analysis, Rapport d'activité, Université de Minho (UMinho), 14 p.

McClatchey, R., Anjum, A., Stockinger, H., Ali, A., Willers, I., Thomas, M., (2007), Data intensive and network aware (DIANA) grid scheduling, Journal of Grid computing, 5, 1, pp. 43-64

MacEachren, A.M., (1998) Cartography, GIS and the World Wide Web, Progress in Human Geography, 22, 4, pp.575–585

Magoulès, F., Pan, J., Tan, K., Kumar, A., (2009), Introduction to Grid Computing, Taylor & Francis Group, CRC Press, 334 p.

Malek, F., (2009), Le calcul scientifique des expériences LHC, la grille de calcul LCG, Laboratoire de Physique Subatomique et de Cosmologie (LPSC), Rapport commandé par le ministère de la recherche: "Recueil d'études de cas permettant d'aborder les principaux enjeux soulevés par les données numériques de la recherche et leurs infrastructures", 18 p.

Marchandise, A. (2007), Modélisation hydrologique distribuée sur le Gardon d'Anduze ; étude comparative de différents modèles pluie-débit, extrapolation de la normale à l'extrême et tests d'hypothèses sur les processus hydrologiques, Thèse de Doctorat, Terre, Eau, Espace, Université Montpellier II, 215 p.

Martin, C., Didon-Lescot, J., Jolivet, J., (2008), *Réflexion sur le rôle hydrologique des aménagements anciens des versants et des talwegs : l'exemple des Cévennes*, Travaux du LGPA, n° 26 (Les eaux : problèmes de gestion et de qualité), pp. 31-40.

Mazzetti, P., (2006), Deliverable D1 – « Cyclops Project Open Conference », Report on CYCLOPS (CYber-Infrastructure for CiviL protectionOperative ProcedureS) Project. CYCLOPSWP01-D01-DPC: 8 p.

Mazzetti, P., Nativi, S., Esteves, A., Sa, V., Lebreiro, N., Thierion, V., Verlato, M., (2008), Deliverable D16 – « Research strategies for the development of a Civil Protection E-infrastructure », Report on CYCLOPS (CYber-Infrastructure for CiviL protection Operative ProcedureS) project, CYCLOPSWP04-D16-IMAA: 67 p.

Mazzetti, P., Nativi, S., Angelini, A., Verlato, M., Fiorucci, P., (2009), A Grid platform for the European Civil Protection e-Infrastructure: the Forest Fires use scenario, Earth Science Informatics, 2, 1-2, pp. 53-62

MEDD – Ministère de l'Ecologie et du Développement Durable, (2002), Circulaire du 1er octobre 2002 relative à la création de services de prévision des crues.

MIAT - Ministère de l'intérieur et l'aménagement du territoire, (2006), Guide ORSEC départemental, Méthode générale, Tome G.1, DDSC, 69 p.

Minar, N., Hedlund, M., (2001), A network of peers : peer-to-peer models through the history of the Internet. In : Oram, A., Peer-to-peer : Harnessing the power of disruptive technologies, O'Reilly & Associates, Inc., Sebastopol, Etats-Unis, pp. 9-20

Mineter, J.M., Skouloudis, A. N., Jarvis, C., Dowers, S., (2002), Towards utilisation of the grid in advanced environmental decision support, 8th International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes, 14-17 Octobre, Sofia, Bulgarie, pp. 434-438

Mineter, J.M., Dowers, S., Skouloudis, A.N., Jarvis, C., (2003), Towards use of grids in environmental research, management and policy, International journal of environment and pollution, 20, 1-6, pp. 297-308

Mitchell, T., (2005), Web Mapping Illustrated, O'Reilly Media, Sebastopol, Etats-Unis, 349 p.

Monmonier, M., (1993), Comment faire mentir les cartes. Du mauvais usage de la géographie, Flammarion, Paris, 232 p.

Lai, X., Monnier, J., (2009), Assimilation of spatially distributed water levels into a shallow-water flood model. Part I : Mathematical method and test case, Journal of Hydrology, 377, 1-2, pp. 1-11

Montz, B.E., Gruntfest, E., (2002), Flash flood mitigation: recommendations for research and applications, Environmental Hazards, 4, 1, pp. 15-22

Morel-Pair, C., (2005), « Panorama : des métadonnées pour les ressources électroniques », Atelier des réseaux de la documentation scientifique "Métadonnées et Archives Ouvertes", 11-13 octobre 2005, Arcachon, 33 p.

Mucchielli, A., (1995), Les sciences de l'information et de la communication, Hachette livre, Paris, 159 p.

Murgia, F., Biddau, R., Concas, A., Demontis, R., Fanfani, L., Heilmann, Z., Lai, C., Lecca, G., Lorrai, E., Marrocu, M., Marrone, V. A., Muscas, L., Peneva, E., Piras, A., Pisu, M., Pusceddu, G., Satta, G., Thies, D., Vacca, A., Valera, P., Ferrara, A., Bonomi, E. (2009), GRIDA3 - a shared resources manager for environmental data analysis and applications, Earth Science Informatics, 2, 1-2, pp. 5-21

Na, A., Priest, M., (2007), Sensor Observation Service, Version 1.0, OpenGIS Consortium Inc., Etats-Unis, 104 p.

Naqvi, S., (2005), Architecture de sécurité pour les grands systèmes, ouverts, répartis et hétérogènes, Thèse de Doctorat, Ecole Nationale Supérieure des Télécommunications, Informatique et réseaux, 160 p.

Nash J.E., Sutcliffe J.V., (1970), River flow forecasting through conceptual models. Part I. A discussion of principles, Journal of Hydrology, 27, 3, pp. 282-290

Nativi, S., Caron, J., Davis, E., Domenico, B., (2005), Design and implementation of netCDF markup language (NcML) and its GML-based extension (NcML-GML), Computers & Geosciences, 31, 9, pp. 1104-1118.

Nativi, S., Verlato, M., Pacini, F., (2009), G-OWS (gLite-OWS) Working Group Charter, Version 1.0, 19 p., En ligne. https://www.g-ows.org/. Consulté en 2010

Neppel, L., Desbordes, M., Masson, J., (1998), Caractérisation de l'aléa climatique pluvieux en régions méditerranéenne : analyse statistique des surfaces pluvieuses, Revue des sciences de l'eau (Paris), Lavoisier, 11, pp. 155-174

Neteler, M., Mitasova, H., (2004), Open source GIS : a GRASS GIS approach (second edition), Springer Science + Business Media, New York, Etats-Unis, 387 p.

Norbiato, D., Borga M., Esposti S.D., Gaume, E., Anquetin S., (2008), Flash flood warning based on rainfall thresholds and soil moisture conditions: An assessment for gauged and ungauged basins, Journal of Hydrology, 362, 3-4, pp. 274-290

Obled, C., Bontron, G., (2005), L'adaptation probabiliste des prévisions météorologiques pour la prévision hydrologique, La Houille Blanche, 1, pp. 23-28

OCDE (2002), Measuring the Information Economy, 94 p. En ligne. www.oecd.org/dataoecd/16/14/ 1835738.pdf. Consulté en 2010

OGC, (2009), Open Geospatial Consortium. En ligne. http://www.opengeospatial.org. Consulté en décembre 2009.

O'Reilly, T., (2007), What Is Web 2.0. Design Patterns and Business Models for the Next Generation of Software, International Journal of Digital Economics, 65, pp. 17-37

Orgerie, A.-C., Lefèvre, L., (2009), A year in the life of a large-scale experimental distributed system: usage of the Grid'5000 platform in 2007, Rapport technique de l'INRIA, 35 p.

Pacini, F., Maraschini, A., (2007), Job Description Language attributes specification, projet EGEE, 83 p.

Patil, A., Belter, B., Polyrakis, A., Rodwell, T., Przybylski, M., Grammatikou, M., (2006), GEANT2 advance multi-domain provisioning system, TERENA Networking Conference, 15-18 Mai 2006, Catane, Italie, 14 p.

Payrastre, O., (2007), Développement et utilisation des modèles au SPC Grand Delta, Congrès AFPCN, journées « Prévision, Vigilance et Alerte, 14-15 Novembre 2007, Toulouse, France

Peng, Z.-R., Tsou, M.-H., (2003), Internet GIS - Distributed Geographic Information Services for the Internet and Wireless Networks, Wiley, New Jersey, Etats-Unis, 679 p.

Percivall, C., (2002), The OpenGIS abstract specification. Topic 12: Open-GIS Service Architecture, (ISO 19119), Open GIS Consortium, Inc., 78 p.

Peterson, M., (2005), Maps and the Internet : an introduction. In : Peterson, M., Maps and the Internet, Elsevier Science Ltd, Oxford, Royaume-Uni, pp. 1-16

Petitdidier, M., Cossu, R., Mazzetti, P., Fox, P., Schwichtenberg, H., Som de Cerff, W., (2009), Grid in earth sciences, Earth Science Informatics, 2, 1-2, pp. 1-3

Pisani, F., Piotet, D., (2008), Comment le web change le monde - l'alchimie des multitudes, Pearson Education France, Paris, 266 p.

Pornon, H., (1992), Les SIG - mise en œuvre et applications, Hermès, Paris, 158 p.

Pornon, H., Yalamas, P., Pelegris, E. (2008), Services web géographiques, état de l'art et perspectives, Géomatique Expert, 65, pp. 44-50

Prechelt, L., (2000), An Empirical Comparison of Seven Programming Languages, IEEE Computer, 33, 10, pp. 23-29

Priol, T., (2005), Défis et perspectives scientifiques des grilles informatiques. In *Club des décideurs et professionnels en Informatique*. En ligne. http://priol.developpez.com/grid-computing/. Consulté en Février 2009.

Puzenat, D., (2008), WINDS@Antilles 2008 - atelier « recherche globale et locale en TIC », Séminaire GRIMAAG, 11-12 décembre 2008, Université Antilles Guyane, Pointe-à-Pitre, France, 25 p.

Rabuffetti, D., Barbero, S., (2005), Operational hydro-meteorological warning and real-time flood forecasting: the Piemonte Region case study, Hydrology and Earth System Sciences, 9, 4, pp. 457-466

Rabuffetti, D., (2006), Discharge assimilation in a distributed flood forecasting model, Advances in Geosciences, 7, pp. 355-360

Radke, J., Cova, T., Sheridan, M.F., Troy, A., Lan, M., Johnson, R., (2000), Application challenges for geographic information science: implications for research, education, and policy for emergency preparedness and response, URISA journal, 12, 2, pp. 15-30

Rainer, J., (2003), *Le Système mondial d'observation,* Météorologie, 24, pp. 28-32

Rallet, A., (2006), Technologies de l'information et de la communication – un enjeu économique de société. In : Olivesi, S., Sciences de l'information et de la communication – Objets, savoirs, discipline, Presses Universitaires de Grenoble, Grenoble, pp. 145-160

Renard, P., Badoux, V., Petitdidier, M., Cossu, R., (2009), Grid Computing for Earth Science, Eos, Transactions, American Geophysical Union, 90, 14, pp. 117-119

Richard, R., Ajay, A., Eswaran, C., (2008), Implementation of computational grid services in enterprise grid environments, American Journal of Applied Sciences, 5, 11, pp. 1442-1447

Rivrain, J.C., (1997), Les épisodes orageux à précipitations extrêmes dans les régions Méditerranéennes du sud de la France. Phénomènes Remarquables, No. 4, Météo France, SCEM, 93 p.

Roche, S., Caron, C., (2004) Aspects organisationnels des SIG, Lavoisier, Paris, 313 p.

Rood, B., Lewis, M. J., (2009), Grid Resource Availability Prediction-Based Scheduling and Task Replication, Journal of Grid Computing, 7, 4, pp. 479-500

Stainforth, D.A., Aina, T., Christensen, C., Collins, M., Faull, N., Frame, D.J., Kettleborough, J.A., Knight, S., Martin, A., Murphy, J.M., Piani, C., Sexton, D., Smith, L.A., Spicer, R.A., Thorpe, A.J., Allen, M.R., (2004), Uncertainty in predictions of the climate response to rising levels of greenhouse gases, Nature, 433, 7024, pp. 403-406

Salomon J-L., (1997), *L'homme face aux crues et aux inondations*, Presses Universitaires de Bordeaux, 136 p.

Sanchis, E., (2007), « Introduction à la programmation shell en Bash ». In *Club des décideurs et professionnels en Informatique*. En ligne. http://eric-sanchis.developpez.com/linux/shell/bash/. Consulté en 2009.

Sauvagnargues-Lesage, S., (2001), De la prévention à la gestion de crise : utilisation des Systèmes d'Information Géographiques. In Brugnot, G., Gestion spatiale des risques, Hermès Science Publications, Paris, 287 p.

Sauvagnargues-Lesage, S., Simonet, C., (2004), « Retour d'expérience sur la gestion de l'événement de Septembre 2002 par les services de Sécurité Civile », Colloque d'hydrotechnique « Crues méditerranéennes – Vulnérabilité, Aléa, Prévisions et gestion de crise », 176ème session du comité scientifique et technique, Société hydrotechnique de France, 15-16 juin 2004, Nîmes, France, pp. 63-69

Sauvagnargues-Lesage, S., Haziza, E., Piatyszek, E., Déchomets, R., Godfrin, V., Rasse, G., Kapfer, A., Roure, J.F., Benedettini, H., (2007), Contribution à la stratégie préventive des secours face à une alerte « inondation », Programme Risque-Décision-Territoire, Ministère de l'Ecologie et du Développement Durable, République Française, 251 p.

Sauvagnargues-Lesage, S., Ayral, P.-A., (2007), Using GIS for emergency management : a case study during the 2002 and 2003 flooding in south-east France, International Journal of Emergency Management, 4, 4, p. 682-703.

Service Départemental d'Incendie et de Secours du Gard (SDIS 30), (2002), Inondations du 8 au 9 septembre 2002 : Rapport général du Plan ORSEC du 9 Septembre au 25 Septembre, 95 p.

Segal, B., Robertson, L., Gagliardi, F., Carminati, F., (2000), « Grid computing : the European Data Grid Project », 47th IEEE Nuclear Science Symposium and Medical Imaging Conference, 15 - 20 Oct 2000, Lyon, France, 6 p.

Seguy A., (2008), Décision collaborative dans les systèmes distribués – application à la e-maintenance, Thèse de Doctorat, Institut National Polytechnique de Toulouse (INP), Systèmes Industriels, 193 p.

Silva, L.M., Buyya, R., (1999), Parallel Programming Models and Paradigms. In: Buyya, R., High Performance Cluster Computing: Architectures and Systems: Volume 2, Prentice Hall, New Jersey, 664 p.

Simonis, I. (2006), OGC Sensor Alert Service Candidate Implementation Specification, Version 0.9, OpenGIS Consortium Inc., Etats-Unis, 110 p.

Simonis, I. (2007), OpenGIS Sensor Planning Service Implementation Specification, Version 1.0, OpenGIS Consortium Inc., Etats-Unis, 186 p.

Servigne, S., Libourel, T., (2006), Fondements des bases de données spatiales, Lavoisier, Paris, 236 p.

Shen, Z., Luo, J., Zhou, C., Cai, S., Zheng, J., Chen, Q., Ming, D., Sun, Q., (2004), Architecture design of grid GIS and its applications on image processing based on LAN, Information Sciences, 166, 1-7, pp. 1-17

Schuler, F., (2005), « Etude et utilisation des technologies des P2P ». In *Club des décideurs et professionnels en Informatique*. En ligne. http://schuler.developpez.com/articles/p2p/. Consulté en 2009.

Schut, P., (2007), OpenGIS Web Processing Service, Version 1.0.0, OpenGIS Consortium Inc., Etats-Unis, 87 p.

Shi, Y., Shortridge, A., Bartholic, J., (2002), Grid Computing for Real Time Distributed Collaborative Geoprocessing, Symposium sur la théorie, les traitements et les applications des données Géospatiales, 9 – 12 juillet, Ottawa, Canada, 12 p.

Shirts, M.R., Pande, V.S., (2001), Mathematical Analysis of Coupled Parallel Simulations, Physivcal review letters, 86, 22, pp. 4983- 4987

Soberman, M., (2003), Les grilles informatiques, Hermès Science, Paris, 163 p.

Soberman, M., (2005), Les grilles informatiques - état de l'art et déploiement, JRES 2005 : Journées Réseaux 2005, 6 au 9 Décembre, Marseille, France, 10 p.

Soh, L. K., Zhang, J., Samal, A. A., (2006), Task-Based Approach to User Interface Design for a Web-Based Hydrologic Information Systems, Transactions in GIS, 10, pp. 417-449

SPC-GD – Service de Prévision des Crues Grand Delta, (2006), règlement de surveillance, de prévision et de transmission de l'information des crues (RIC), Direction Départementale de l'Equipement du Gard, 75 p.

Sui, D., Maggio, R., (1999), Integrating GIS with hydrological modelling: practices, problems, and prospects Computers, Environment and Urban Systems, 23, pp. 33-51

Taladoire, G., (2003), Les SIG sur le web, Lavoisier, Paris, 430 p.

Tanenbaum, A., Van Steen, M., (2002), Distributed systems - Principles and Paradigms, Prentice Hall, Upper Saddle River, New Jersey, 803 p.

Tanguy, J., Carriere, J., Le Trionnaire, Y., Schoen, R., (2004), *« Réorganisation de l'annonce des crues en France : Crues Méditerranéennes »,* Colloque d'hydrotechnique « Crues méditerranéennes – Vulnérabilité, Aléa, Prévisions et gestion de crise », 176ème session du comité scientifique et technique, Société hydrotechnique de France, 15-16 juin 2004, Nîmes, France, pp. 167-173

Tardieu, H., Rochfeld, A., Colletti, R., (1983), La méthode Merise - Principes et outils, Editions d'organisation, Paris, 328 p.

Tardieu, J., Leroy, M., (2003), Radome, le réseau temps réel d'observation au sol de Météo-France, Météorologie, 40, pp. 40-43

Toukourou, M., Johannet, A., Dreyfus, G., Ayral, P.-A., (2009), « *Rainfall-runoff modelling for flash floods : the case of Cévenol flash floods* », Journal of Applied Intelligence (soumis en 2009)

Thierion, V., Mazzetti, P., Oliveira, J., (2007), Deliverable D11 – « System Requirements Document », Report on CYCLOPS (CYber-Infrastructure for CiviL protection Operative ProcedureS) project, CYCLOPSWP03-D11-DDSC: 27 p.

Thierion, V., Ayral, P.-A., (2008), Flash floods applications, IBERGRID Conference, 12-14 Mai, Porto, Portugal, 18 p.

Thierion, V., Ayral, P.-A., Sauvagnargues-Lesage, S., Bressand, F., Nardin, P., Sorani, R., (2008), « Flood forecasting and warning operations by grid technology adoption in Civil Protection organization », 13th IWRA World Water Congress 2008, 1-4 septembre 2008, Montpellier, France

Thierion, V., Ayral, P.-A., Angelini, V., Bronner, C., Mazzetti, P., Sauvagnargues-Lesage, S., Payrastre, O., (2009a), Warning operations enhancement by Grid technology – GALHTAIR : a platform dedicated to the flash flood in the south of France, Demo session : End-user environments and portal technologies, 4th EGEE User Forum, 2-6 mars, Catane, Italie

Thierion, V., Ayral, P.-A., Angelini, V., Sauvagnargues-Lesage, S., Nativi, S., Payrastre, O., (2009b), Operational flash flood forecasting platform based on grid technology, European Geosciences Union, General Assembly, 19-24 Avril, Vienne, Austriche

Thierion, V., Ayral, P.-A., Geisel, J., Sauvagnargues-Lesage, S., Payrastre, O., (2010), Grid technology reliability for flash flood forecasting: end-user assessment, Journal of Grid Computing (Soumis)

Tsou, M.-H., Buttenfield, B. P., (2002), A dynamic architecture for distributing geographic information services, Transactions in GIS, 6, 4, pp. 355-381

Urbah, E., Kacsuk, P., Farkas, Z., Fedak, G., Kecskemeti, G., Lodygensky, O., Marosi, A., Balaton, Z., Caillat, G., Gombas, G., Kornafeld, A., Kovacs, J., He, H., Lovas, R., (2009), EDGeS: Bridging EGEE to BOINC and XtremWeb, Journal of Grid Computing, 7, 3, pp. 335 -354

Van Elzakker, C. (2001), Use of maps on the Web. In : Kraak, M.-J., Brown, A., Web Cartography developments and prospects, Taylor & Francis, Londres, Angleterre, pp. 21-36

Van Elzakker, C. (2001b), Users of maps on the Web. In : Kraak, M.-J., Brown, A., Web Cartography developments and prospects, Taylor & Francis, Londres, Angleterre, pp. 37-51

Verlato, M., Dal Pra, S., Angelini, V., D'Andrea, M., Frizziero, E., Squizzato, S., Pina, A., Esteves, A., Sa, V., Marques, R., Thierion, V., Ayral, P.-A., (2008), Deliverable D17 – « Dissemination and coordination activity », Report on CYCLOPS (CYber-Infrastructure for CiviL protection Operative ProcedureS) Project. CYCLOPS-WP02-D17-INFN : 51 p.

Veyret, Y., (2004), Géographie des risques naturels en France, de l'aléa à la gestion, Hatier, Paris, 251 p.

Vicat-Blanc Primet, P., D'Anfray, P., Blanchet, C., Bobelin, L., Boudeville, O., Chanussot, F., Cung, V., Denneulin, Y., Djerrah, A., Drabik, M., Jegou, Y., Huynh, P., Lacorne, N., Lacoste, D., Le cun, B., Lefevre, L., Mateo, D., Metery, R., Meurdesoif, Y., Mignot, J.-C., Pajot, A., Pham, C., Rémi, R., Romier, G., Touche, I., (2003), Réflexions et propositions pour une grille haute performance, Projet RNTL e-Toile, 81 p.

Viers, G., (1968), *Eléments de climatologie*, Fernand Nathan, Paris, 224 p.

Vinet, F., (2003), Crues et inondations dans la France Méditerranéenne. Les crues torrentielles des 12 et 13 novembre 1999, Éditions du Temps, Nantes, 224 p.

Vinet, F., (2007), Approches nationales de la prévention des risques et besoins locaux : le cas de la prévision et de l'alerte aux crues dans le Midi méditerranéen, Géocarrefour, 82,1-2, pp. 35-42

Von Laszewski, G., Wagstrom, P., (2004), Gestalt of the Grid. In: Hariri, S., Parashar, M., Tools and environments for parallel and distributed computing, Series on parallel and distributed computing, Wiley, Hoboken, New Jersey, pp. 149-187

Vretanos, P.A., (2005), Web Feature Service Implementation Specification, Version 1.1.0, OpenGIS Consortium Inc., Etats-Unis, 131 p.

W3C, (2005), « Web Services Architecture ». In *World WIde Web Consortium (W3C)*. En ligne. www.w3.org/TR/2004/NOTE-ws-arch-20040211/. Consulté en 2009.

Waldner, J.-B., (2007), Nano-informatique et Intelligence Ambiante – Inventer l'ordinateur du XXIème Siècle, Hermes-Lavoisier, 302 p.

Walker, D., Dongarra, J., (1996), MPI: a standard message passing interface Supercomputer, 12, 1, pp. 56-68

Wang, J., Su, G., (2007), Running architecture of Grid GIS based on the Quadruple Geo-spatial Grids model, MIPPR 2007 : Remote sensing and GIS data processing and applications, and Innovative multispectral technology and applications, 15-17 Novembre, Wuhan, Chine

Wang, Y., Ge, L., Rizos, C., Babu, R., (2004), Spatial data sharing on grid, Geomatics Research Australasia, 2004, 81, pp. 3-18

Whitesite, A., (2005), OpenGIS® web services architecture description, Version 0.1.0, OpenGIS Consortium Inc., Etats-Unis, 28 p.

Whitesite, A., **Evans, J.D**., (2006), Web Coverage Service (WCS) Implementation Specification, Version 1.1.0, OpenGIS Consortium Inc., Etats-Unis, 143 p.

Whitesite, A., (2007), OGC Web Services Common Specification, Version 1.1.0, OpenGIS Consortium Inc., Etats-Unis, 167 p.

Whitesite, A., (2009), Web Coverage Service (WCS) — Transaction operation extension, Version 1.1.4, OpenGIS Consortium Inc., Etats-Unis, 34 p.

Wolski , R., (2003), Experiences with Predicting Resource Performance On-line in Computational Grid Settings, ACM SIGMETRICS Performance Evaluation Review, 30, 4, pp. 41-49

Woolf, A., Nativi, S., (2008), How earth can contribute to and benefit from the spatial information infrastructure. In : Van Oosterom, P., Zlatanova, S., Creating spatial information infrastructures: towards the spatial Semantic Web, CRC Press, pp. 67-87

Xiao, N., Fu, W., (2003), SDPG : Spatial Data Processing Grid, Journal of Computer Science and Technology, 18, 4, pp. 523-531

Youn, C., Baru, C., Bhatia, K., Chandra, S., Lin, K., Memon, A., Memon, G., Seber, D., (2005), GEONGrid portal : Design and implementations, Concurrency and computation: practice and experience, 19, 12, pp. 1597-1607

Youn, C., Kaiser, T., Seber, D., Santini, C., (2008), Web-based simulating system for modeling earthquake seismic wavefields on the grid, Computers and Geosciences, 34, 12, pp. 1936-1946

Zeldman, J., (2006), « Web 3.0 ». In *A List Apart*. En ligne. www.alistapart.com/articles/web3point0. Consulté en 2010.

Zerger, A., Smith, D.I., (2003), Impediments to using GIS for real-time disaster decision support Computers, Environment and Urban Systems, 27, 1, pp. 123-141

Zhang, C., Li, W., (2005), The roles of web feature and web map services in real-time geospatial data sharing for time-critical applications, Cartography and geographic information science, 32, 4, pp. 269-283

Zhou, J., De Roure, D., (2007), FloodNet : Coupling adaptative sampling with energy aware routing in a flood warning system, Journal of computer science and technology, 22, 1, pp. 121-130

LISTE DES FIGURES

Partie 1 - Cadre institutionnel et états de l'art scientifiques

Figure 1 : Le risque naturel entre Nature et Société 14

Figure 2 : Cycle du risque et positionnement scientifique de cette recherche 16

Figure 3 : Système convectif de méso-échelle en V 20

Figure 4 : L'épisode cévenol 22

Figure 5 : Genèse des débits de crues : principaux processus superficiels et souterrains 23

Figure 6 : Représentation des différents processus de génération de l'écoulement en fonction du milieu naturel 25

Figure 7 : Structure de la plateforme MARINE 30

Figure 8 : Fonctionnement du module de production d'ALHTAÏR 32

Figure 9 : L'approche multi-échelles de Météo-France : les trois principaux modèles 35

Figure 10 : Carte de vigilance météorologique du 09 septembre 2002 à 01h37 (a) et carte de vigilance hydrologique du 13 avril 2007 à 09h52 (b) 37

Figure 11 : Liens organisationnels entre les composantes météorologiques, hydrologiques et Sécurité Civile de la gestion de crise "inondation" 41

Figure 12 : Développement de l'expertise hydrométéorologique concertée : la vigilance "pluie-inondation" 42

Figure 13 : Évolution du nombre d'utilisateurs d'Internet 48

Figure 14 : Évolution de l'informatique de 1960 à nos jours 48

Figure 15 : Ensemble des fonctions de l'intranet, et par extension des TIC 49

Figure 16 : L'interopérabilité selon l'OpenGIS® 50

Figure 17 : Exemple d'un réseau informatique étendu englobant plusieurs réseaux locaux (LAN) 53

Figure 18 : Architecture de communication client-serveur 3 tiers 55

Figure 19 : Architecture globale de la technologie grille, les couches correspondent de bas en haut aux ressources réseau, matérielles, de l'intergiciel et applicatives 58

Figure 20 : Relation entre organisation virtuelle, ressources matérielles et utilisateurs : le grid problem 59

Figure 21 : Relations entre le modèle OSI et les couches d'une architecture de grille 61

Figure 22 : Infrastructure du réseau RENATER 65

Figure 23 : Évolutions technologiques pour le développement d'une e-Infrastructure 66

Figure 24 : Évolution de l'infrastructure d'EGEE 67

Figure 25 : Caractéristiques de l'infrastructure réseau de GEANT2 68

Figure 26 : Organisation virtuelle dédiée aux inondations de plaine 71

Figure 27 : Les 4 composantes du SIG et par extension de la géomatique 75

Figure 28 : Évolution architectural des SIG en ligne 76

Figure 29 : Modèle client-serveur dédié à la cartographie en ligne 78

Figure 30 : Serveur cartographique enrichi d'une interface CGI 79

Figure 31 : Connexion de SIG distants sans norme d'échange (a)) et avec une norme d'échange (b)) 81

Figure 32 : Les trois grandes types d'architecture SIG 81

Figure 33 : Interopérabilité à l'échelle des composants 83

Partie 2 - Technologie grille et géomatique pour l'expertise hydrométéorologique en situation de crise

Figure 34 : Hiérarchie des diagrammes UML 92

Figure 35 : Situation générale du territoire de compétence du SPC-GD 94

Figure 36 : Cas d'utilisation du système d'information du SPC-GD 95

Figure 37 : Diagramme d'activité « gestion des données hydrométéorologiques » 96

Figure 38 : Positions des zones radar et recouvrement d'une image radar sur le gardon d'Anduze 97

Figure 39 : Diagramme d'activité « gestion et analyse de l'imagerie radar » 98

Figure 40 : Zones "Avertissement Précipitation" sur le territoire de compétence Grand Delta 99

Figure 41 : Diagramme d'activité « modalisation du débit des cours d'eau » 100

Figure 42 : Diagramme de déploiement du SPC-GD 101

Figure 43 : Diagramme d'activité simplifié de l'application ALHTAÏR 103

Figure 44 : Fonctionnement des modules de production et de transfert d'ALHTAÏR 104

Figure 45 : Cumul des précipitations entre le 12 et le 13 novembre 1999 108

Figure 46 : Cumul des précipitations du 8 au 9 septembre 2002 110

Figure 47 : Activité opérationnelle dans le Gard le lundi 09 septembre entre 12h et 18h 111

Figure 48 : Synopsis de l'expertise hydrologique en situation de crise 116

Figure 49 : La télégéomatique chez les pompiers 118

Figure 50 : Hydrogrammes mesuré, simulé et prévus par modélisation de la prévision de la pluie prévue (1h, 2h et 3h) de la Beaume à Rosières avec le modèle ALHTAÏR 121

Figure 51 : Principe de parallélisation de contrôle, les cercles représentent des tâches, et les flèches des canaux de communication 132

Figure 52 : Distribution des sous-ensembles dans la parallélisation des données avec 4 processeurs 133

Figure 53 : Structure des tâches pour une étude paramétrique 134

Figure 54 : Architecture de grille multi-tiers 139

Figure 55 : Structure du Geospatial grid 140

Figure 56 : Architecture proposée par le projet européen CYCLOPS 143

Figure 57 : Composants d'un système spatial d'aide à la décision 146

Figure 58 : Positionnement du cas d'utilisation « support spatial d'aide à la décision » dans l'organisation existante du SPC-GD 148

Figure 59 : Diagramme d'activité du cas d'utilisation "support spatial d'aide à la décision" 149

Figure 60 : Diagramme de déploiement synthétique du nouveau système 152

Partie 3 - Méthodologie expérimentale

Figure 61 : Services de l'intergiciel gLite 160

Figure 62 : États d'un job lors de son traitement par l'intergiciel 163

Figure 63 : Exemple d'un job de grille avec gLite 164

Figure 64 : Services de gLite impliqués dans la soumission d'un job 165

Figure 65 : Synopsis du traitement d'un scénario de prévision 168

Figure 66 : Diagramme de classes des modules hydrologiques d'ALHTAÏR 170

Figure 67 : Extrait de la description CDL d'une donnée de pluie CALAMAR, basée sur la convention CF 171

Figure 68 : Job de grille (JDL) pour exécuter les modules hydrologiques d'ALHTAÏR 173

Figure 70 : Composants logiciels de l'application distribuée G-ALHTAÏR 181

Figure 71 : Cas d'utilisation du système spatial d'aide à la décision 183

Figure 72 : Architecture des services web spécifiques de l'OGC 187

Figure 73 : Fonctionnement classique d'un WFS applicable aux autres services de l'OGC 188

Figure 74 : Exemple de requête URL à un service WMS fourni par le serveur cartographique GeoServer 189

Figure 75 : Méthodes de grillification des services web de l'OGC. a) de bas niveau b) de haut niveau 191

Figure 76 : Comportement classique du WCS dans le cas du traitement d'un fichier NetCDF 192

Figure 77 : Diagramme d'activité UML de G-ALHTAÏR 195

Figure 78 : Diagramme de séquence UML du rôle de l'échangeur 197

Figure 79 : Diagramme de séquence UML de GALHTAÏR 198

Figure 80 : Diagramme de déploiement UML de l'architecture de l'interface de cartographie en ligne 200

Figure 81 : Interface graphique de l'application G-ALHTAÏR 201

Figure 82 : Diagramme de déploiement UML de GALHTAÏR 203

Figure 83 : Test préliminaire de l'exécution d'instances d'ALHTAÏR sur EGEE 206

Figure 84 : Fonctionnement de RRM-Wrapper au niveau de la grille de calcul 210

Figure 85 : Fonctionnement de RRM-Parametric pour l'allocation des jobs 211

Figure 86 : Diagramme de séquence UML de RRM-Grid 213

Figure 87 : Diagramme de déploiement de RRM-Grid 214

Partie 4 - Résultats et Perspectives : vers l'adoption potentielle de la technologie grille

Figure 87 : Répartition de la majorité des sites de calcul de la VO ESR 223

Figure 88 : Méthode de construction des scénarios de pluie basés sur CALAMAR 229

Figure 89 : Hydrogrammes de crue observé et modélisé de l'évènement du 22 novembre 2008 à l'exutoire du bassin versant d'Anduze 230

Figure 90 : Cumuls pluviométriques sur le bassin versant d'Anduze entre le 19/10/2008 et le 23/10/2008 231

Figure 91 : Boîtes à moustaches des temps d'exécution des trois phases de traitement avec la méthode 1 (a) et la méthode 2 (b) 233

Figure 92 : Performance de la phase de soumission pour les méthodes 1 et 3 235

Figure 93 : Boîtes à moustaches des temps d'exécution pour chaque délai de prévision avec la méthode 1 (a) et la méthode 3 (b) 236

Figure 94 : Durée de l'exécution de tous les scénarios pour les 3 méthodes, usage classique (a), RRM-Wrapper (b) et RRM-Grid (c) 237

Figure 95 : Diagrammes à moustaches des temps d'exécution de 210 scénarios suivant le nombre de workers 238

Figure 96 : Hydrogrammes observé et modélisé de l'évènement du 22 octobre 2008, et stratégie de prévision hydrologique pour la phase expérimentale 240

Figure 97 : Situation hydrologique du 21/10/2008 à 14h 242

Figure 98 : Axe chronologique de la réception des scénarios de prévision de la phase n°1 242

Figure 99 : Faisceaux de prévisions hydrologiques modélisées, basées sur les prévisions CALAMAR (a) et les prévisions AP à pluie constante (b) (T0 : 21/10/2008 à 14h00) 243

Figure 100 : Situation hydrologique du 21/10/2008 à 20h55 243

Figure 101 : Axe chronologique de la réception des scénarios de prévision de la phase n°2 244

Figure 102 : Faisceaux de prévisions hydrologiques modélisées, basées sur les prévisions CALAMAR (a) et les prévisions AP à pluie constante (b) (T0 : 21/10/2008 à 20h55) 244

Figure 103 : Situation hydrologique du 22/10/2008 à 00h20 245

Figure 104 : Axe chronologique de la réception des scénarios de prévision de la phase n°3 246

Figure 105 : Faisceaux de prévisions hydrologiques modélisées, basées sur les prévisions CALAMAR (a) et les prévisions AP à pluie constante (b) (T0 : 22/10/2008 à 00h20) 246

Figure 106 : Situation hydrologique du 21/10/2008 à 02h15 247

Figure 107 : Axe chronologique de la réception des scénarios de prévision de la phase n°4 248

Figure 108 : Faisceaux de prévisions hydrologiques modélisées, basées sur les prévisions CALAMAR (a) et les prévisions AP à pluie constante (b) (T0 : 22/10/2008 à 02h15) 248

Figure 109 : Bilan des prévisions hydrologiques les plus performantes de l'évènement du 22 octobre 2008 250

Figure 110 : Structure informatique de G-ALHTAÏR après intégration de RRM-Grid 262

Figure 111 : Stratégie de modélisation hydrologique et hydraulique multi-modèles 266

Figure 112* : Schéma synthétique d'un système d'information dédié à la gestion du risque "inondation" basé sur la technologie grille 278

* Cette figure est située dans la conclusion générale

LISTE DES TABLEAUX

Partie 1 - Cadre institutionnel et états de l'art scientifiques

Tableau 1 : Synthèse des types de ruissellement 24

Partie 2 - Technologie grille et géomatique pour l'expertise hydrométéorologique en situation de crise

Tableau 2 : Points de vue du cadre conceptuel RM-ODP 91

Tableau 3 : Synthèse des caractéristiques hydrométéorologiques et opérationnelles des évènements dans l'Aude (1999) et dans le Gard (2002) 112

Tableau 4 : Comparaison entre la grille EGEE et un cluster suivant les caractéristiques principales des systèmes distribués 135

Partie 3 - Méthodologie expérimentale

Tableau 5 : Caractéristiques techniques des données et algorithmes après grillification 172

Partie 4 - Résultats et Perspectives : vers l'adoption potentielle de la technologie grille

Tableau 6 : Synthèse de la méthodologie expérimentale 225

Tableau 7 : Phases temporelles observées dans la phase expérimentale 226

Tableau 8 : Hauteurs et débits correspondant aux niveaux de vigilance 230

Tableau 9 : Performances globales des trois méthodes dans le traitement des simulations prenant en charge 33, 66 et 105 scénarios de prévision 232

Tableau 10 : Correspondances entre l'horizon de prévision du scénario et le délai opérationnel le concernant 241

LISTE DES SIGLES

ACI Action Concertée Incitative
AIGA Adaptation de l'Information Géographique pour l'Alerte en crue
AJAX Asynchronous JavaScript and XML
ALADIN Aire Limitée, Adaptation dynamique, Développement InterNational
ALHTAÏR Alarme Hydrologique Territoriale. Automatisée par Indicateur de Risque
ALICE A Large Ion Collider Experiment
API Application Programming Interface
ARAMIS Application Radar A la Météorologie Infra Synoptique
AROME Application de la Recherche à l'Opérationnel à Méso-Echelle
ARPEGE Action de Recherche Petite/Echelle Grande Echelle
ATHYS ATelier HYdrologique Spatialisé
ATLAS A Toroidal LHC ApparatuS

BAP Bulletins d'Alerte Précipitations
BMS Bulletin Météorologique Spécial
BRAM Bulletin Régional d'Alerte Météo
BRP Bulletin Régulier de Précipitations

CALAMAR CAlcul des LAMes d'eau à l'aide du radAR
CDL Common Data Language
CDM Centre Départemental de la Météorologie
CE Element de Calcul
CF Climat Convention
CGI Common Gateway Interface
CMIR Centres Météorologiques InterRégionaux
CMS Compact Muon Solenoid
CNRS Centre national de la recherche scientifique
CODIS Centre Opérationnel Départemental d'Incendie et de Secours
COGIC Centre Opérationnel de Gestion Interministérielle des Crises
COZ Centre Opérationnel de Zone
CPU Central Processing Unit
CYCLOPS CYber-Infrastructure for CiviL protection Operative ProcedureS

DAG Direct Acyclic Graph
DDSC Direction de la Défense et de la Sécurité Civile
DDTM Direction départementale des Territoires et de la Mer
DEGREE Dissemination and Exploitation of GRids in Earth sciencE
DIANA Data Intensive and Network Aware
DIANE Distributed Analysis Environment

EDG European Data Grid
EDGeS Enabling Desktop Grids for e-Science
EGEE Enabling Grids for E-sciencE
EMZ Etat-Major de Zone
EPSG European Petroleum Survey Group
ESR Earth Science Research

FIFO First In First Out
FTP File Tranfert Protocol

GIF Graphics Interchange Format
gLite Lightweight Middleware for Grid Computing
GLUE Grid Laboratory for a Uniform Environment
GML Geography Markup Language
GPS Global Positioning System
GRID Globalisation des Ressources Informatiques et des Données
GUID *Grid Unique IDentifier*

HPC High-Performance Computing
HTTP Hypertext Transfer Protocol

IFS Integrated Forecasting System
ISO International Organization for Standardization

JDL *Job Description Langage*
JPEG Joint Photographic Experts Group
JSON JavaScript Object Notation

LAN Local Area Network
LB Logging and Bookkeeping
LFC *LCG File Catalog*
LFN Logical File Name
LHC Large Hadron Collider
LHCb Large Hadron Collider beauty
LIDAR Light Detection and Ranging

MARINE Model of Anticipation of Runoff and INondations for Extreme events
MDS *Monitoring and Discovery Service*
MEEDDM Ministère de l'Écologie, de l'Énergie, du Développement Durable et de la Mer
MERCEDES Maillage Élémentaire Régulier Carré pour l'Étude Des Écoulements Superficiels
MPI Message Passing Interface
NcML NetCDF Markup Language
NetCDF Network Common Data Form

OCDE Organisation de coopération et de développement économiques
OGC Open Geospatial Consortium
OGF Open Grid Forum
OMM Organisation Météorologique Mondiale
ORSEC Organisation de la Réponse de SEcurité Civile
OSI Open Systems Interconnection

PANTHERE Projet ARAMIS Nouvelles Tech. en Hydrométéorologie Extension et REnouvellement
PCS Plan Communal de Sauvegarde
PDF Portable Document Format
PFN *Physical File Name*
PNG Portable Network Graphics
RADOME Réseau d'Acquisition et de Données d'Observation Météo Etendue
RENATER Réseau National de télécommunications pour la Tech. l'Enseignement et la Recherche
RFIO Remote File Input/Output
R-GMA *Relational Grid Monitoring Architecture*
RIC Règlement de surveillance, de prévision et de transmission de l'Information des Crues
RM-ODP Reference Model of Open Distributed Processing
RRM-Grid Rainfall-Runoff modeling on Grid
RSS Really Simple Syndication

SAC Service d'Annonce des Crues
SCEM Service Central d'Exploitation de la Météorologie
SCHAPI Service Central d'Hydrométéorologie et d'Appui à la Prévision des Inondations
SCM Système Convectif Méso-échelle
SCS Soil Conservation Service
SDPC Schéma Directeur de Prévision des Crues
SE Element de Stockage
SGBD Système de gestion de base de données
SHYPRE Simulation d'HYDrogrammes pour la PREdétermination des crues
SIDPC Service Interministériel de Défense et de Protection Civiles
SIG Système d'Information Géographique
SMO Système Mondial d'Observation
SOA Service-Oriented Approach
SOAP Simple Object Access Protocol
SPC Service de Prévision des crues
SPC-GD Service de Prévision des crues « Grand Delta »
SSAD Système Spatial d'Aide à la Décision
SURL *Storage URL*
SVG Scalable Vector Graphics

TIC Technologies de l'Information et de la Communication
TIN Triangulated irregular network
TURL *Transport URL*

UDP User Datagram Protocol
UI Interface Utilisateur
UML Unified Modelling Language
URL Uniform Resource Locator

VO Organisation Virtuelle
VOMS Virtual Organization Membership Service
VPN Réseau privé virtuel

WAN Wide Area Network
WCS Web Coveragde Service
WFS Web Feature Service

W-LCG The Worldwide LHC Computing Grid
WMS Web Map Service
WMS Workload Management System
WN Nœud de Calcul
WPS Web Processing Service
WWW World Wide Web

XML Extensible Markup Language
XSD XML Schema

TABLE DES MATIERES

Sommaire 1

Introduction Générale 5

Partie 1 : Cadre institutionnel et État de l'art 9

Chapitre 1 : Les crues à cinétique rapide : phénomène et gestion opérationnelle 11

I - 1. Les crues à cinétique rapide : un risque majeur 13

I - 1.1. La gestion du risque « inondation » 13

I - 1.1.1. Fondements théoriques : aléa, vulnérabilité et risques majeurs 13

I - 1.1.1.1. La notion de risque 13

I - 1.1.1.2. L'aléa 14

I - 1.1.1.3. La vulnérabilité 14

I - 1.1.1.4. Le risque majeur 15

I - 1.1.2. La prévision hydrométéorologique en situation de crise 16

I - 1.1.2.1. La crise : définitions et caractéristiques fonctionnelles 16

I - 1.1.2.2. La prévision hydrométéorologique comme élément fondamental de la gestion de crise « inondation » 17

I - 1.2. Dimension hydrométéorologique des crues éclair 19

I - 1.2.1. Fondements théoriques 19

I - 1.2.2. Phénomènes météorologiques 19

I - 1.2.3. Définition et processus hydrologiques d'une crue à cinétique rapide 22

I - 1.3. La modélisation des phénomènes de crues à cinétique rapide 26

I - 1.3.1. La modélisation pluie-débit 26

I - 1.3.1.1. Modèles empiriques 27

I - 1.3.1.2. Modèles physiques 27

I - 1.3.1.3. Modèles conceptuels 27

I - 1.3.1.4. Synthèse des capacités de modélisation des crues à cinétique rapide 28

I - 1.3.2. Les plateformes de modélisation 29

I - 1.3.2.1. ATHYS, ATelier HYdrologique Spatialisé 29
I - 1.3.2.2. MARINE (Model of Anticipation of Runoff and INondations for Extreme events 29
I - 1.3.2.3. ALHTAÏR, Alarme Hydrologique Territoriale Automatisée par Indicateur de Risque 31
I - 1.3.2.4. AIGA, Adaptation de l'Information Géographique pour l'Alerte en crue 32
I - 1.4. Synthèse des crues à cinétique rapide 33
I - 2. Cadre institutionnel et opérationnel de la gestion de crise « inondation » 34
I - 2.1. La notion de vigilance 34
I - 2.2. La vigilance météorologique : l'instigateur de l'alerte 34
I - 2.2.1. L'expertise sur les phénomènes météorologiques 34
I - 2.2.1.1. Observations météorologiques 34
I - 2.2.1.2. Modélisation météorologique 35
I - 2.2.2. Les grandes caractéristiques de la vigilance météorologique 36
I - 2.3. La vigilance hydrologique : une émergence justifiée 37
I - 2.3.1. De l'annonce à la prévision des crues 37
I - 2.3.2. Le SCHAPI : entre opérationnalité et innovation 38
I - 2.3.3. Les services de prévision des crues : la compétence hydrologique locale 38
I - 2.3.3.1. Rôle des services de prévision des crues 38
I - 2.3.3.2. Le schéma directeur de prévision des crues (SDPC) 39
I - 2.3.3.3. Le règlement de surveillance, de prévision et de transmission de l'information des crues (RIC) 40
I - 2.3.4. Les grandes caractéristiques de la vigilance hydrologique 40

Chapitre 2 : L'informatique distribuée : une composante nouvelle des sociétés modernes 45
II - 1. Propos introductifs 47
II - 1.1. les technologies de l'information et de la communication (TIC) 47
II - 1.2. Collaboration et organisation virtuelle 50
II - 1.3. Interopérabilité 50
II - 2. L'informatique distribuée et la technologie grille 52
II - 2.1. Généralités sur l'informatique distribuée 52
II - 2.1.1. Définition de l'informatique distribuée 52

II - 2.1.2. Une typologie des architectures distribuées 52
II - 2.1.3. les nouveaux enjeux de l'informatique distribuée 54
II - 2.2. Les grands modèles de communication dans un système distribué 55
II - 2.2.1. Client-serveur 55
II - 2.2.2. Pair-à-pair 56
II - 2.2.3. L'intergiciel 56
II - 2.3. La technologie grille : définitions et fonctionnement 57
II - 2.3.1. Historique et définitions 57
II - 2.3.1.1. Grille de 1ère génération 57
II - 2.3.1.2. Grille de 2ème génération 58
II - 2.3.1.3. Grille de 3ème génération 58
II - 2.3.1.4. L'organisation virtuelle garante de l'approche collaborative 59
II - 2.3.2. Les 5 couches fonctionnelles de la technologie grille 60
II - 2.3.2.1. la couche « Fabrique » 61
II - 2.3.2.2. la couche « Connectivité » 61
II - 2.3.2.3. la couche « Ressource » 61
II - 2.3.2.4. la couche « Collectif » 62
II - 2.3.2.5. la couche « Application » 62
II - 2.3.3. Une typologie des infrastructures de grille 62
II - 2.3.4. Des grilles d'ampleur nationale 64
II - 2.3.4.1. GRIDA3 64
II - 2.3.4.2. GRID'5000 64
II - 2.3.5. Le projet EGEE : une révolution technologique, scientifique et politique 65
II - 2.3.5.1. La vision E-Science 65
II - 2.3.5.2. La grille EGEE 67
II - 2.3.5.2.1 La physique des particules et la grille W-LCG 68
II - 2.3.5.2.2 La recherche biomédicale et l'organisation virtuelle BIOMED 69
II - 2.3.5.3. La grille : une technologie adaptée aux sciences de la Planète 71
II - 2.4. Synthèse sur la technologie grille 72
II - 3. La cartographie en ligne entre géomatique et services web 73
II - 3.1. Gestion des données géoréférencées : SIG et géomatique 73
II - 3.1.1. Définitions 73
II - 3.1.2. Structuration et finalité de l'information géographique dans un SIG 75
II - 3.2. Les SIG et les réseaux informatiques 76

II - 3.2.1. Cartographie en ligne : définitions et principe 76
II - 3.2.2. Une typologie de la cartographie en ligne 77
II - 3.2.2.1. La publication de cartes statiques 78
II - 3.2.2.2. La cartographie en ligne interactive et l'enrichissement des solutions clientes 79
II - 3.2.3. La science de l'information géographique 80
II - 3.2.3.1. Nouveaux enjeux du SIG en ligne 80
II - 3.2.3.2. Principe fonctionnel des services d'information géographique 81
II - 3.2.3.3. L'Open Geospatial Consortium : une formalisation opérationnelle des services d'information géographique 83
II - 3.2.3.4. Interfaces, opérations et services 83

Partie 2 : Technologie grille et géomatique pour l'expertise hydrométéorologique en situation de crise 87

Chapitre 3 : L'expertise hydrométéorologique des crues à cinétique rapide 89

III - 1. L'existant technologique du Service de Prévision des Crues « Grand Delta » 91

III - 1.1. Les méthodes RM-ODP et UML dédiées à la spécification des systèmes d'information 91

III - 1.1.1. La méthode RM-ODP ou modèle de référence pour le « traitement réparti ouvert » 91
III - 1.1.2. Le langage UML 92

III - 1.2. Le système d'information du SPC-GD : rôles, compétences et capacités 93

III - 1.2.1. Présentation générale 93
III - 1.2.2. Point de vue « Entreprise » 94
III - 1.2.3. Points de vue « Information » et « Traitement » 96
III - 1.2.3.1. Gestion des données hydrométéorologiques 96
III - 1.2.3.2. Gestion et analyse de l'imagerie radar 97
III - 1.2.3.3. Analyse des prévisions météorologiques et exportations de données hydrométéorologiques 98
III - 1.2.3.4. Modélisation du débit des cours d'eau 99
III - 1.2.4. Points de vue « Ingénierie » et « Technologique » 100
III - 1.2.4.1. Description technique du système d'information 100
III - 1.2.4.2. L'application ALHTAÏR : « modélisation des cours d'eau » 102
III - 1.2.5. Bilan de l'analyse UML : premières hypothèses de recherche 105

III - 2. Retours d'expérience : Besoins opérationnels, scientifiques et technologiques **107**

III - 2.1. Les crues à cinétique rapide : les cas « Aude 1999 » et « Gard 2002 » **107**

III - 2.1.1. Évènement du 12 et 13 novembre 1999 dans le département de l'Aude 107

III - 2.1.1.1. Caractéristiques hydrométéorologiques 107

III - 2.1.1.2. Gestion de crise et dommages 109

III - 2.1.2. Évènement du 8 et 9 septembre 2002 dans le département du Gard 109

III - 2.1.2.1. Caractéristiques hydrométéorologiques 109

III - 2.1.2.2. Gestion de crise et dommages 111

III - 2.1.3. Synthèse sur les évènements gardois et audois 112

III - 2.2. De la Sécurité Civile à l'expertise hydrométéorologique : un perfectionnement perpétuel **115**

III - 2.2.1. Définition de la typologie des besoins 115

III - 2.2.2. Les missions de la Sécurité Civile : objectif ultime de la gestion de crise 116

III - 2.2.2.1. Gestion des opérations de secours et prévision hydrométéorologique 117

III - 2.2.2.1.1 L'information de prévision 117

III - 2.2.2.1.2 Le partage et la diffusion des données 117

III - 2.2.2.1.3 Le rôle prépondérant de l'information géographique 117

III - 2.2.3. L'expertise hydrométéorologique : vers la modélisation numérique performante en situation opérationnelle 119

III - 2.2.3.1. Besoins relatifs à la gestion des données 119

III - 2.2.3.1.1 L'intégration de l'expertise 119

III - 2.2.3.1.2 Les réseaux de mesures 119

III - 2.2.3.1.3 Les courbes de tarage 119

III - 2.2.3.1.4 Les données historiques 120

III - 2.2.3.1.5 Les retours d'expérience 120

III - 2.2.3.2. Besoins relatifs aux traitements et à la modélisation des données 120

III - 2.2.3.2.1 La prévision d'ensembles 120

III - 2.2.3.2.2 Le calage en temps différé et l'assimilation en temps-réel 121

III - 2.2.3.2.3 La valorisation de la modélisation opérationnelle 122

III - 2.2.3.3. Bilan des besoins 122

Chapitre 4 : La technologie grille et la géomatique : vecteurs d'innovation **127**

IV - 1. Les contraintes technologiques de la prévision hydrologique **129**

IV - 1.1. De la prévision hydrologique… **129**

IV - 1.2. … aux contraintes technologiques inhérentes **129**

IV - 1.2.1. Besoins fonctionnels 130

IV - 1.2.2. Besoins non-fonctionnels 130

IV - 2. La grille EGEE : un support informatique performant et adaptable **132**

IV - 2.1. La technologie grille : des caractéristiques technologiques et des champs d'application favorables à une adoption opérationnelle **132**

IV - 2.1.1. Le point central de la parallélisation informatique 132

IV - 2.1.1.1. Le parallélisme de contrôle 132

IV - 2.1.1.2. Le parallélisme de données 133

IV - 2.1.1.3. Parallélisation et technologie grille 133

IV - 2.1.2. Les caractéristiques informatiques de la grille EGEE 134

IV - 2.1.3. L'utilisation de la grille pour la gestion de problématiques environnementales 136

IV - 2.1.3.1. Aide à la décision 136

IV - 2.1.3.2. Études paramétriques 137

IV - 2.1.4. Synthèse sur l'intérêt d'utiliser la grille comme support technologique 137

IV - 2.2. La technologie grille et la science de l'information géographique : une synergie innovante **138**

IV - 2.2.1. L'intérêt croissant des services web 138

IV - 2.2.2. Le Geospatial Grid 140

IV - 2.2.3. Des initiatives récentes dans l'utilisation de services web de l'OGC pour l'accès aux ressources de grille 141

IV - 2.2.3.1. L'institut de recherches spatiales ukrainien 141

IV - 2.2.3.2. Le projet CYCLOPS et G-RISICO 142

IV - 2.2.3.3. Le projet GDI-GRID 143

IV - 3. G-ALHTAÏR : un système spatial d'aide à la décision hydrologique **145**

IV - 3.1. L'hypothèse de recherche centrale : le système spatial d'aide à la décision **145**

IV - 3.2. Caractérisation globale de G-ALHTAÏR dédié à l'anticipation des crues à cinétique rapide **147**

IV - 3.3. Spécification détaillée du système spatial d'aide à la décision au sein du SPC-GD **147**

IV - 3.3.1. Vers une nouvelle dimension organisationnelle 147

IV - 3.3.2. Vers une capacité d'anticipation hydrologique 149
IV - 3.3.3. Vers l'orchestration des traitements de grille 151

Partie 3 : Méthodologie expérimentale 155

Chapitre 5 : Grillification de l'application ALHTAÏR 157

V - 1. L'intergiciel gLite : un accès unique aux ressources d'EGEE 159

V - 1.1. Les services d'accès 161

V - 1.2. les services de sécurité 161

V - 1.3. les services d'information et de surveillance des ressources 162

V - 1.4. les services de gestion de données 163

V - 1.5. Les services de gestion des tâches 163

V - 1.5.1. Un job de grille 163
V - 1.5.2. Soumission d'un job 165

V - 2. Méthodologie fondamentale 167

V - 2.1. Principe de fonctionnement du scénario de prévision 167

V - 2.2. Stratégie de portage 169

V - 2.2.1. Recodage des algorithmes 169
V - 2.2.2. Formatage des données 170
V - 2.2.3. Simulation hydrologique sur la grille EGEE 172

Chapitre 6 : Interfaçage entre la grille EGEE et le SPC-GD 177

VI - 1. Notes introductives : la problématique d'interfaçage 179

VI - 2. G-ALHTAÏR : un outil garant de l'interopérabilité et de l'orchestration des traitements 182

VI - 2.1. Point de vue « Entreprise » : l'interface de cartographie en ligne garante de l'opérationnalité du SPC-GD 182

VI - 2.1.1. Gestion de l'imagerie radar 183
VI - 2.1.2. Gestion des cours d'eau et des stations de mesure 184
VI - 2.1.3. Gestion des commentaires 184

VI - 2.1.4. Modélisation pour la prévision hydrologique 184

VI - 2.2. Points de vue « Information et Traitement » : les services web de l'OGC 185

VI - 2.2.1. Fondements technologiques des services web de l'OGC 186

VI - 2.2.1.1. Généralités sur les services de l'OGC 186

VI - 2.2.1.1.1 Typologie générale 186

VI - 2.2.1.1.2 Exemples de services géographiques de l'OGC 187

VI - 2.2.1.2. Invocation d'un service web de l'OGC 189

VI - 2.2.1.3. Positionnement technologique 191

VI - 2.2.2. Les services web de l'OGC : le cas d'utilisation « modélisation pour la prévision hydrologique » 193

VI - 2.2.2.1. Point de vue « Information » 193

VI - 2.2.2.2. Point de vue « Traitement » 197

VI - 2.3. Point de vue « Ingénierie » et « Technologie » de G-ALHTAÏR 199

VI - 2.3.1. L'interface de cartographie en ligne 199

VI - 2.3.1.1. Le serveur web à composante géographique 199

VI - 2.3.1.2. L'application cartographique cliente 201

VI - 2.3.2. Synthèse : l'architecture globale de G-ALHTAÏR 202

VI - 3. RRM-Grid : un outil garant de l'ordonnancement et d'une qualité de service 206

VI - 3.1. Propos introductifs : limites de performance de l'intergiciel gLite et contraintes de qualité de service 206

VI - 3.2. Ordonnanceur et qualité de service 207

VI - 3.3. RRM-Grid : une couche applicative garante d'une efficacité opérationnelle 209

VI - 3.3.1. La gestion des données avec RRM-Wrapper 209

VI - 3.3.2. La gestion des jobs avec RRM-Parametric 211

VI - 3.3.3. RRM-Grid : un ordonnanceur adapté à la modélisation de scénarios de prévision hydrologique 213

Partie 4 : Résultats et Perspectives : vers l'adoption potentielle de la technologie grille 219

Chapitre 7 : Expérimentations de l'ordonnanceur RRM-Grid 221

VII - 1. Plan d'expérience 223

VII - 1.1. Ressources de grille : la VO ESR 223

VII - 1.2. Méthodes de traitement expérimentées **224**

VII - 1.3. Critères d'analyse **226**

VII - 1.3.1. Phases principales de l'exécution d'un job 226

VII - 1.3.2. Indicateurs de performance 227

VII - 1.3. Étude de cas en situation opérationnelle **228**

VII - 1.3.1. Stratégie de construction des scénarios 228

VII - 1.3.2. Matériel expérimental : Bassin versant et évènement hydrologique 229

VII - 2. Résultats expérimentaux de RRM-Grid **232**

VII - 2.1. Performances technologiques **232**

VII - 2.1.1. Gestion des données d'entrée : RRM-Wrapper 232

VII - 2.1.2. Gestion du traitement des scénarios : RRM-Grid 234

VII - 2.1.2.1. Phase de Soumission 234

VII - 2.1.2.2. Performance globale du traitement des scénarios 236

VII - 2.1.2.3. Montée en charge 237

VII - 2.1.3. Synthèse des capacités technologiques de RRM-Grid 239

VII - 2.2. Étude de cas de l'évènement du 22 octobre 2008 **240**

VII - 2.2.1. Phase n°1 : 21/10/2008 à 14h00 241

VII - 2.2.1.1. Situation hydrométéorologique 241

VII - 2.2.1.2. Scénarios de prévision simulés 242

VII - 2.2.1.3. Analyse des performances et des prévisions 242

VII - 2.2.2. Phase n°2 : 21/10/2008 à 20h55 243

VII - 2.2.2.1. Situation hydrométéorologique 243

VII - 2.2.2.2. Scénarios de prévision simulés 244

VII - 2.2.2.3. Analyse des performances et des prévisions 244

VII - 2.2.3. Phase n°3 : 22/10/2008 à 00h20 245

VII - 2.2.3.1. Situation hydrométéorologique 245

VII - 2.2.3.2. Scénarios de prévision simulés 245

VII - 2.2.3.3. Analyse des performances et des prévisions 246

VII - 2.2.4. Phase n°4 : 22/10/2008 à 02h15 246

VII - 2.2.4.1. Situation hydrométéorologique 246

VII - 2.2.4.2. Scénarios de prévision simulés 247

VII - 2.2.4.3. Analyse des performances et des prévisions 247

VII - 2.2.5. Synthèse sur la capacité de prévision 249

Chapitre 8 : Perspectives opérationnelles et scientifiques de G-ALHTAÏR **255**

VIII - 1. G-ALHTAÏR : un système spatial d'aide à la décision en devenir **257**

VIII - 1.1. L'optimisation des traitements de grille **257**

VIII - 1.1.1. Ordonnancement, accès aux ressources et rapidité d'exécution 257

VIII - 1.1.1.1. La gestion des données 257

VIII - 1.1.1.2. La gestion des traitements 258

VIII - 1.1.1.3. La question de la parallélisation 258

VIII - 1.1.2. Une meilleure visibilité des ressources de grille 259

VIII - 1.1.2.1. Cartographie des ressources en temps réel 259

VIII - 1.1.2.2. Prévision et indicateurs de performance 260

VIII - 1.1.2.3. Réservation et/ou préemption des ressources 260

VIII - 1.2. L'interopérabilité : entre sophistication et opérationnalité **261**

VIII - 2. Enrichissement de la dimension hydrométéorologique **264**

VIII - 2.1. L'approche multi-modèle en temps réel **264**

VIII - 2.1.1. Une panoplie de modules de production et de transfert 264

VIII - 2.1.2. La modélisation hydraulique 265

VIII - 2.1.3. Prototype d'implémentation potentiel de l'approche multi-modèle 266

VIII - 2.2. Vers l'optimisation du modèle **268**

VIII - 2.2.1. Étalonnage des modèles 268

VIII - 2.2.2. Assimilation des débits 269

VIII - 2.3. Gestion des systèmes de mesure et des données collectées **270**

Conclusion générale **273**

Bibliographie **281**

Liste des figures **305**

Liste des tableaux **312**

Liste des sigles **313**

Annexe **329**

ANNEXE

Alhtair
-
Parallelization and Grid Issues Analysis

Ricardo Marques[1], Vincent Thierion[2]

[1]Universidade do Minho
ricardo@di.uminho.pt

[2]École des Mines D'Alés
vincent.thierion@ema.fr

July 10, 2008

Abstract

This document reports the analysis of *Althair*, a french civil protection flood modeling software, for future parallelization and porting to the Grid infrastructure. A parallel algorithm for the application is proposed, a a model for its execution in a Grid environment making use of Open Web Services is designed.

The proof of these two concepts is left for future work.

1 Introduction

In a world where time is crucial, and where everything must be done correctly, but at the same time as fast as possible, programmers are several times confronted with the need for speeding up their programs. This need is not only justified by the commodity of the users, but also for stronger reasons. There are situations where a faster program can make the difference between an efficient rescue plan and an uncontrolled situation. That is the case of applications used for floods and fire predictions. The sooner the authorities have the information about how will the situation evolve in the terrain, the sooner they will be able to react. That is why this kind of applications need to be able to make the best use of the computational power available, in order to model hazard situations efficiently and based in real-time data.

The aim of this report is to discuss how can *Alhtair* (a French civil protection flood modeling software) be parallelized and what could be the advantages executing it on the grid.

In section 2 the *Flash Flood* phenomena is presented, followed by the software *Alhtair* in the 3rd section. In the 4th, a suggestion for a possible parallelization of *Alhtair* is made. In section 5 is described a model for the integration of *Alhtair* with the Grid infrastructure. Finally, in 6, the main conclusions of this work are presented.

2 Flash Flood

Flash flood names a phenomena that occurs frequently in the South of France, in the *Gard Region* located between the Mediterranean and the first foothills of *Central Massif*, "*Les Cévènnes*" (Figure 1).

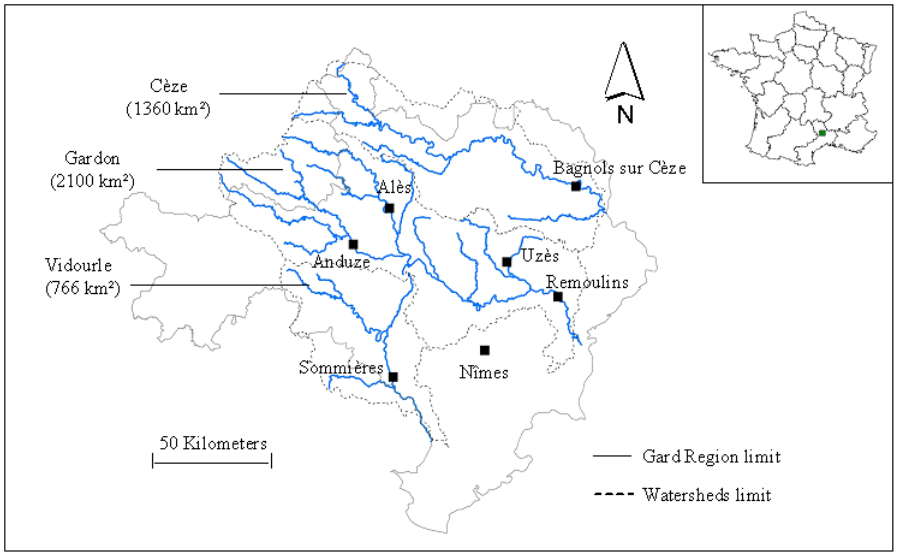


Figure 1: Gard Region and the principals rivers location, from [3].

When southern wind pushes the moist warm air from the Mediterranean to a much colder coast due to the presence of a mountain *massif*, the conditions for the disaster are set up [3]. The difference of the sea and the earth temperature makes the air cool down, and consequently the relative humidity increases. As a result we have high intensity rainfall event. This sequence of events is called "*épisode Cevenol*" (Figure 2), and generally occurs in the Autumn. Associated to this meteorological event, a watershed rapid response causes a potential flash flood.

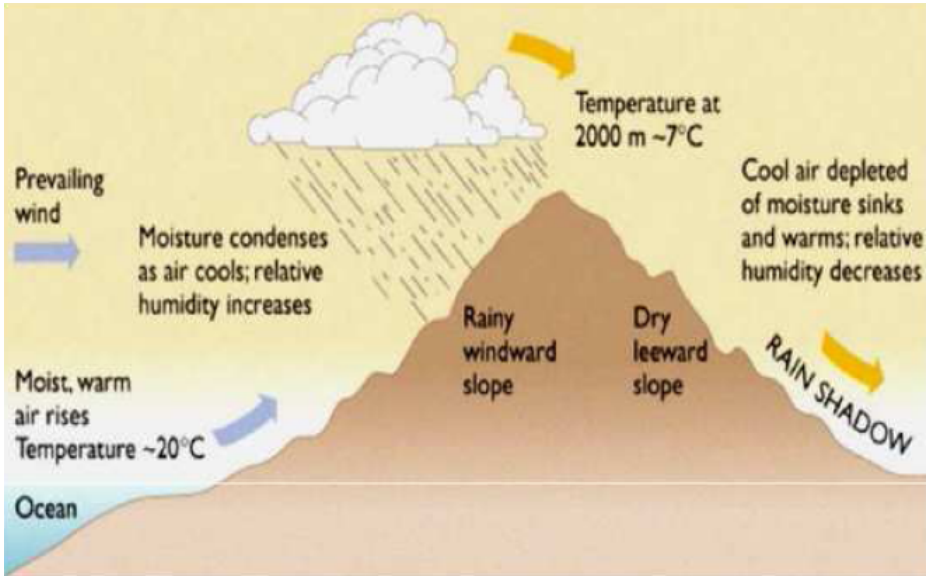


Figure 2: An "*episode Cevenol*", from [1].

Recently, in 2002, the damages caused by these floods made clear the need for improving real-time monitoring and flood forecast [2]. During the 8th and the 9th September, an extreme "*épisode Cèvenol*" occurred in the *Gard Region*. Its results were tragic, causing 22 victims and 1.2 billion euros damages [3].

3 *Alhtair*

Alhtair (***AL**arm **H**ydrologique **T**erritoriale **A**utomatisée par **I**ndicateur de **R**isque*) is an hydrological forecasting model (which corresponds to a distributed rainfall - runoff model) developed by the flood warning service of the *Gard Region*, in the South of France [3, 4].

The execution of *Althair* has two separate logic phases: the production phase and the propagation phase. The production phase consists of, based on the rainfall information and on the characteristics of the terrain, calculate the amount of water that infiltrates in the soil (flow), and the amount that stays in the surface (runoff). Then, in the propagation phase, the volume of watter per second that is expected to reach the outlet point, for all given time intervals that correspond to the time between two consecutive rainfall reports, is calculated.

Figure 3 shows an example of a transfer matrix, used both in the production and propagation phases.

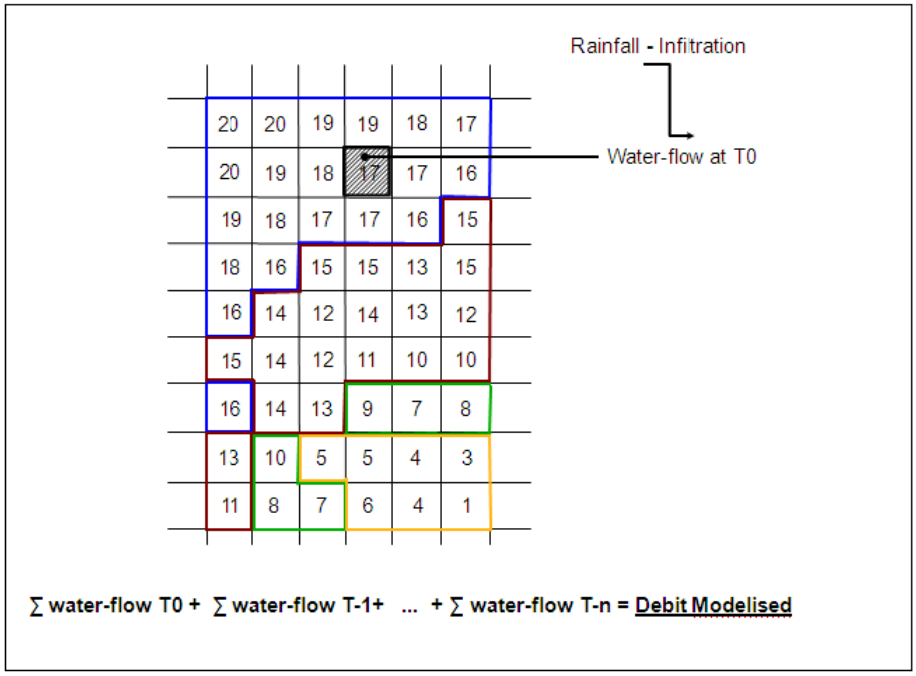


Figure 3: Transfer time matrix example.

The number inside each cell represents the time that the water-flow takes to reach the outlet point (transfer time). In this case, the outlet is located in the lower right corner, as we can see by the value of 1 present in this cell. The cells are grouped by transfer time. On the top right of the figure, we can see an example of how the water-flow is calculated, as described before. The final result of the simulation is achieved by summing the total runoff in the watershed for each time interval T_n.

3.1 Input Data

Following is the list of the data that *Alhtair* needs to execute:

- the number the of radar zone covering the watershed to model (source of rainfall reports)
- the list of rainfall reports
- the characteristics of the watershed

A set of radars is in charge of recollecting information on the rainfall. Each radar is responsible for a limited area, as seen in figure 4.

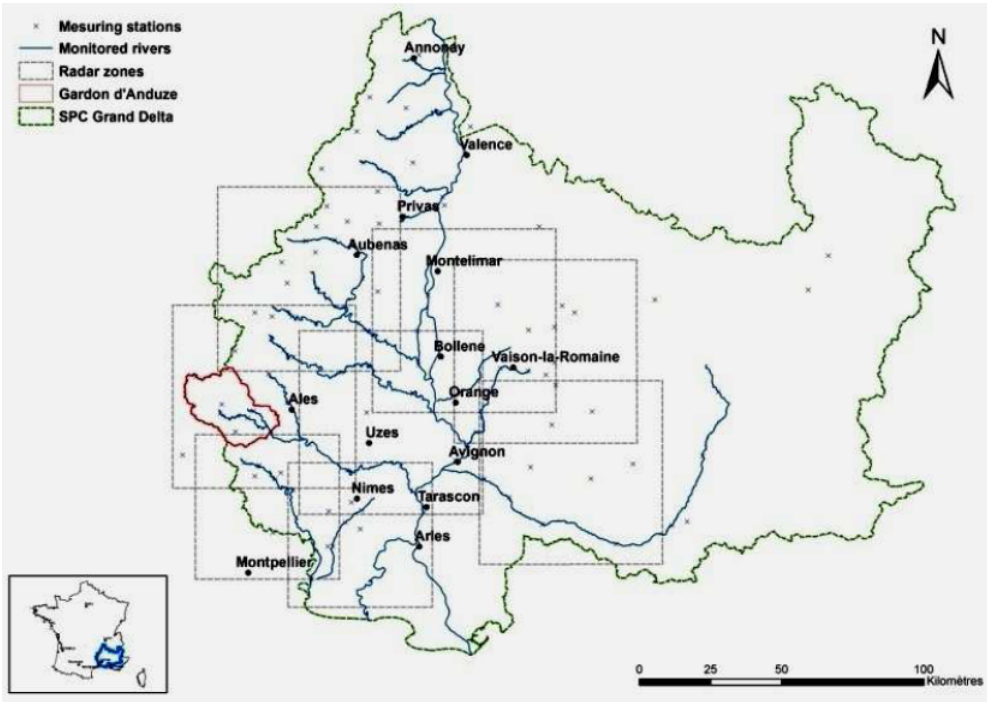


Figure 4: Grand Delta Flood Forecasting Service (SPC-CD), from [1].

Each square represents the range of a radar.

Each time we want to proceed to a simulation, a set of radar rainfall data and a file describing the watershed must be passed to the program. The information about the number of the radar used to produce the data and the coordinates of the lower left corner of the watershed must also be given to the application. This will be used to make the correspondence between the watershed map and the rainfall map. Obviously it makes no sense to use rainfall/watershed information that do not cover the same area.

The rainfall information consists on a list of reports. These reports have a given time interval between them, and are organized in a chronological order. The information contained is the spread and the intensity (mm/h) of the rain fall in the area covered by the radar, during that time interval. Currently, each radar produces a report every five minutes.

And finally, the characteristics of the watershed: a watershed file contains information about the flow time for each cell. This file is generated by a separate program, Hydrokit [4], which uses a *Digital Elevation Model* (DEM) of the area. Based on the elevation information, it calculates the transfer time, that is, the time that the water takes to reach a given outlet, for each cell of the watershed. The information about the capacity of the terrain to absorb the water is also relevant for this model, and is currently hard-coded in the application.

3.2 Output Data

The output produced by *Alhtair* is a text file that registers the volume of watter that is expected to reach the outlet point during the time covered by the rain fall reports, that is, an hydrograph (Figure 5).

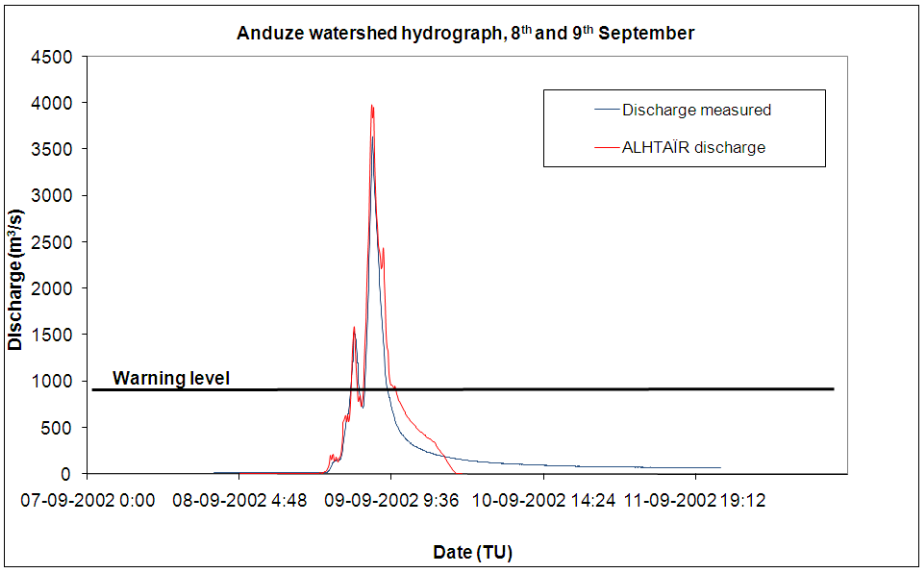


Figure 5: Hydrograph [THESIS_PIERRE].

In red is the graphic of the discharge calculated by *Alhtair* for a given event, as an example of a representation of its output.

3.3 Algorithm

The execution algorithm of *Althair* is described in Algorithm 3.3.

Algorithm 1 *Alhtair* Algorithm.

```
1. Initialize Data
2. Read the Watershed Information
3. Map the rainfall cells to the watershed cells
4. Start Simulation
    for each rainfall report:
        Read the rainfall information
        for each watershed cell:
            Retrieve raw rainfall
            Calculate the runoff
            Update water flow
5. Write the results to a file
```

The first step is the initialization of the data structures involved in the process (1). After this, *Alhtair* reads the information of the terrain over which the simulation will take place (2). The third step is to construct a correspondence vector between the radar cells and the watershed cells (3).

When everything is set up, the simulation starts. It consists on a cycle with three tasks:

- read the rain information, from the rainfall report;
- calculate the part of this rain that the terrain cannot absorb (e.g. the water flow);
- and finally, based on this information, update the watter flow;

This cycle is repeated for each rainfall report, in a chronological order. This process is illustrated in figure 6.

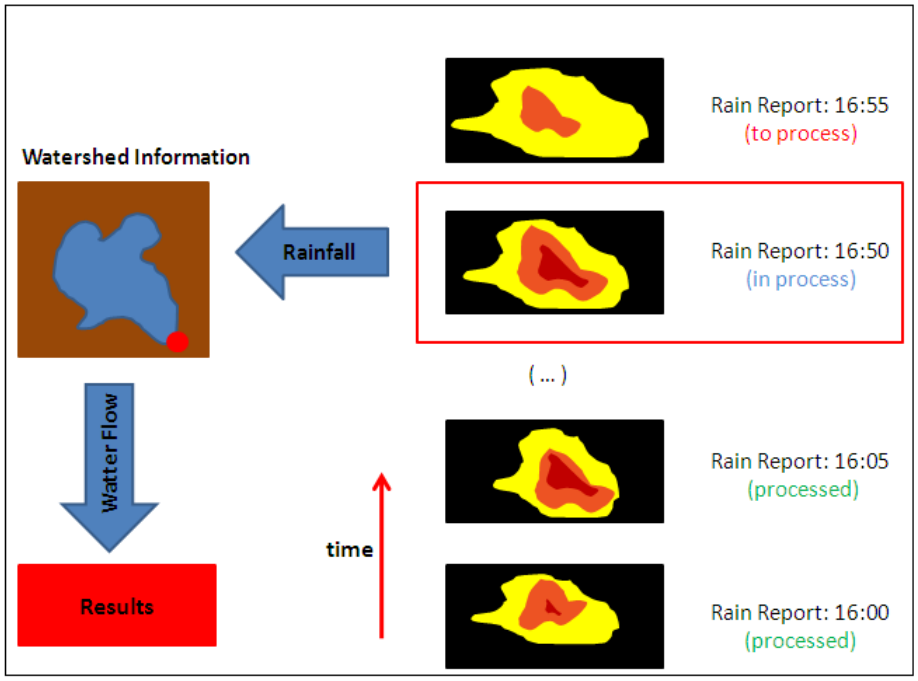


Figure 6: Alhtair Execution.

In the top left of the figure is represented the terrain information. When the simulation begins, the first report is read. Based on its information, the raw rainfall is calculated and applied to the terrain. The next step is to determine the runoff according to the infiltration capacity. The result (the water-flow is then stored, and the next report is read, repeating the process.

Note that the result is cumulative. That is, the final result is the accumulation of the volume of watter that each rainfall (report) will make reach the outlet point, for a fixed period of time (in this case, 5 minutes).

4 A Suggestion for the Parallelization of Alhtair

Currently, processing a rainfall report takes approximately 12 seconds, depending on the machine used to run it. If we add to this the (real) possibility of running the simulation with rainfall reports plus a set of forecast reports, then the simulation time can increase dramatically. As an example, if after receiving a rainfall report we want to add to the result a set of 60 forecast reports corresponding to a forecast of 5 hours, then the simulation would take an average of 12 minutes. This makes the need for reduction of the simulation time very clear. The idea is to make a simulation in the time interval between the arrival of two consecutive reports, that is, 5 minutes. This simulation should be as rich as possible, taking into account as much forecast reports as possible. The most

forecast reports are used, the more reliable the final result is. Figure 7 shows this characteristic of the model.

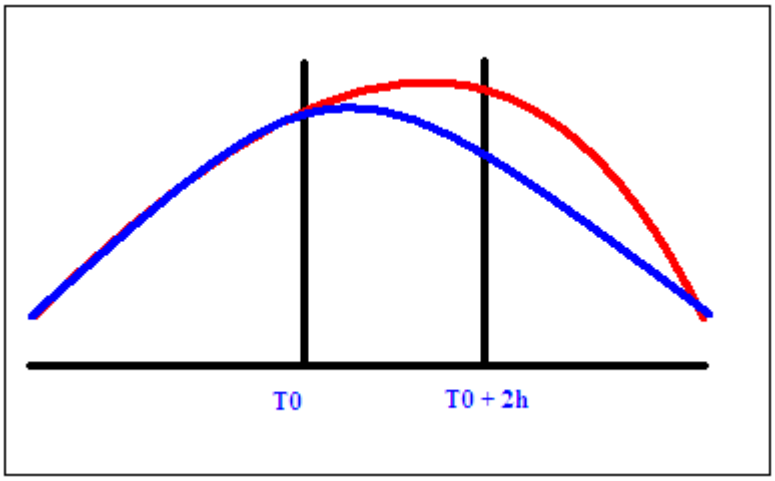


Figure 7: Reliability of the results of the simulation.

In blue we can see the expected hydrograph as a result of the execution without forecast reports, and in red the expected result with 2 hours of forecasts reports included. The results for both executions are the same until T_0. This is a consequence of using the same rainfall reports. From T_0 on, the blue hydrograph gives a constant decreasing of the flow, due to the absence of rainfall from that point on. On the other hand, the red graphic keeps increasing until T_{0+2h}, due to the rainfall prevision registered in the forecast reports for that period of time. From that point on, the hydrograph decreases due to the absence of rainfall forecast reports.

4.1 Domain Decomposition

The idea of paralellizing *Alhtair* is to distribute the calculus among all the processors available for the simulation. In this model, the calculus consists on determining, based on the rainfall reports, the infiltration for each watershed cell, and then traverse all the watershed cells to determine the total water-flow. Thus, the computational effort is a direct function of the number of watershed cells.

A simple approach to the decomposition is distributing blocks of lines form the watershed file to the worker nodes. But this approach has a weak point: the matrix contained in the watershed file is composed not only by cells that are a part of the watershed, but also by cells that are not and that are not involved in the simulation calculus. This is a consequence of the irregularity of the watershed shape. The border lines/columns have, in general, less cells than

the center lines/columns. This characteristic of the watershed file could cause a poor load balancing among the worker nodes, using this distribution technique.

To guarantee a good load balancing, the most safe technique is to divide the watershed cells among all the processes. In concrete, this can be done dividing resulting vector of step 3 in Algorithm 3.3. This vector contains the index of all the watershed cells, and, for each one of them, the index of the rainfall report cell that corresponds to it. This vector will later be traversed, and the rainfall will be calculated for each of the watershed cells present in the vector. Thus, dividing this vector equally among all the process guarantees the best load balance. In Algorithm 2 shows how to proceed to this division.

Algorithm 2 Watershed cells vector decomposition algorithm.

n_{procs}: number of processor involved in the simulation
n_{cells}: the number of watershed cells (size of the correspondence vector)

$remainder$: remainder of the division of n_{cells} by n_{procs}
$block$: integer division of n_{cells} by n_{procs}

if($myRank < remainder$):
    $myFirstCell = myRank * block + myRank$
    $myLastCell = myFirstCell + block$
else:
    $myFirstCell = myRank * block + remainder$
    $myLastCell = myFirstCell + block - 1$

Each processor is assigned, according to its rank, a contiguous block of cells. The size of this block is a function of the number of cells of the correspondence vector and the number of the processors involved in the simulation. The result is that each processor gets a part of the vector.

4.2 Algorithm

A way to parallelize *Athtair* is described in Algorithm 4.2.

Algorithm 3 Parallel *Alhtair*.

1. Each node reads the watershed file
2. Each node calculates the correspondence vector
3. Each node calculates the sub-vector that it is in charge of
4. Each node starts the simulation in it own subset of watershed cells
   for each rainfall report:
      Read the rain information relevant to process
         for each watershed cell that the process is in charge of:
            Calculate the rain that is not absorbed
            Calculate the watter flow
            Update water-flow
5. Send the results to the root node
6. The root node process the results (sum all the vectors)
7. The root node writes the results to file

In the beginning of program, each process reads the watershed file (1). The next step is to calculate the correspondence vector, that maps watershed cells to rainfall report cells (2). Then, according to the number of processes involved in the execution and to the number of watershed cells of the correspondence vector, each node calculates its sub-vector (3). This distribution is static during the execution and should be as balanced as possible, making each process be in charge of approximately the same number of cells. This way, the work balance is guaranteed, for all the cells require the same computation power. When each process knows which sub-vector it is responsible for, the simulation starts as in the sequential algorithm (4). After the simulation finishes, each process sends the results to the root node (5) which is in charge of receiving and processing all the results (6). Processing the results means, in this case, to sum all the result vectors received, to a final results vector. The last step is to write this final result to a file (7).

5 Thoughts on Porting *Althair* to the Grid

As an application used by civil protection to help decision makers, it is crucial that *Althair* is able to run fast and based on actual data. The first need can be achieved by running the parallel version of the application in a closed cluster environment. But this execution environment does not answer to the need for real time data, available through the *Open Web Services* (OWS) developed by the *Open Geospatial Consortium* (OGC).

The Grid infrastructure offers both computational power and real time data, by making powerful computational resources and Grid-enabled OWS [5] available for the users. Therefore, an integration of the *Althair* application with Grid and OWS is an important step towards its efficiency.

5.1 The Open Geospatial Web Services

The Open Geospatial Consortium is a non-profit, international, voluntary consensus standards organization. Its aim is to research and encourage the development and implementation of standards for geospacial contents and location based services. In particular, this organization has defined a specification for geospatial network web-based services: the OpenGeospatial Web Services. Following is the list specifications of services that comprise the OWS [11]:

- **Web Map Service (WMS):** provides a framework that supports the creation and display of registered and superimposed map-like views of information that come simultaneously from multiple remote and heterogeneous sources. Notice that a map is not the data itself. It is just the "interpretation" of the geospatial information requested, as a digital image suitable to be displayed in a computer screen [8].

- **Web Feature Service (WFS):** allows a client to retrieve and update geospatial data encoded in Geography Markup Language (GML) from multiple Web Services. It provides an interface that permits requests for geographical features across the web using platform independent calls. A feature can be seen as the description of a map, as its source code, whereas the WMS interface returns only an image that end-users cannot edit or spatially analyze. This image is a function of the source code that describes the map [9].

- **Web Coverage Service (WCS):** provides an interface allowing requests for interchange of geospatial data as coverages across the web using platform independent calls. The coverages are objects that contain digital geospatial information, that the user can edit and spacially analyze, representing a space-varying phenomena. A WCS provides access to potentially detailed and rich sets of geospatial information, in forms that are useful for client-side rendering, multi-valued coverages, and input into scientific models and other clients [6].

- **Web Processing Service (WPS):** this service is designed to standardize the way that geospatial calculations are made available to the Internet. WPS can describe any calculation (i.e. process) including all of its inputs and outputs, and trigger its execution as a Web Service. It facilitates the the publishing of geospatial processes and the discovery of and binding to those process by client applications. These process can consist on any algorithm, calculation or model that operates on spatially referenced data (although also process that operate on non-spatially referenced data are allowed). This process can then be published, that is, made available for other users [7].

- **Web Service Common (WSC):** specifies the aspects that are common to all or multiple OWSs. In these aspects are included the request and response operations, the parameters and data structures included in

operation requests and responses, and the XML and KPV encoding of operation requests and responses [10].

For the present work, two of these web services are particularly relevant: the WCS and the WPS.

5.2 A Model that Integrates Alhtair and OGS

Following is a description of what could be the behaviour of an *Alhtair* Web Processing Service.

When *Alhtair* WPS receives a request for a simulation describing the area and the period of the simulation, it looks for the data corresponding to that terrain and to the given period of time, that is, the watershed and the report files. This is done making use of the Grid-enabled WCG which, if the data requested is not available in Grid storage resources, sends a job to the Grid to get (and process if necessary) that data. Once the data is available for the execution of *Alhtair*, the WPS sends a job with the parallel version of *Alhtair* to the Grid, and waits for its completion. Once the results are available, they are fetched and place where the user specified.

6 Conclusions

The objectives of this work were achieved. The parallel version of *Alhtair* has the potential for decreasing the execution time of the application, and provide the decision makers with more reliable information in real time. This is a consequence of the use of a bigger number of reports, including forecast reports. Porting the application to the Grid Infrastructure, making use of Grid Enabled OWS brings to the application a set of real time information, crucial for this kind of models.

The proof of these two concepts is left for future works.

References

[1] Thierion Vincent, Ayral Pierre-Alain, "*Flash Flood Management: Thre French Case Study*", 2nd Iberian Grid Infrastructure Conference, 2008

[2] Claudet Roland, Bouvier Christophe, "*Flash Floods Forecasting Techniques and Operative Tools*", La Houille Blanche, 1-2005

[3] Ayral Pierre-Alain, Sauvagnargues-Lesage Sophie, Bressand François, Gay Stephan, "*Forecasting flash floods with an operational model: application in the South-East of France (Gard)*", Ecole des Mines d'Alès

[4] Aryal Pierre-Alain, Contribuion à la Spatialisation du Modèle Operationnel de Prévision des Crues Éclaires ALHTAIR - Approche Spatiale et Expérimentale - Application ao Bassin du Gardon d'Anduse. Ph. D. thesis, Université de Provence, Aix-Marseille 1, 300 p., 2005

[5] Paolo Mazzetti, Stefano Nativi, Valerio Angelini, Marco Verlato, Antonio Pina, Paolo Fiorucci, *"A Grid Platform for the European Civil Protection e-Infrastructure: the Forest Fires use Scenario"*, 2nd Iberian Grid Infrastructure Conference Proceedings (IBERGRID)

[6] OpenGIS, *"Web Coverage Service (WCS), Version 1.0.0"*, OGC 03-065r6

[7] OpenGIS, *"Web Processing Service (WPS), Version 1.0.0"*, OGC 05-007r7

[8] OpenGIS, *"Web Map Service, Version 1.3.0"*, OGC 03-109r1

[9] OpenGIS, *"Web Feature Service, Version 1.0"*, OGC 02-058

[10] OpenGIS, *"Web Services Common, Version 1.1.0"*, OGC 06-121r3

[11] Open GeoSpatial Consortium, "www.opengeospatial.org"

Contribution à l'amélioration de l'expertise en situation de crise par l'utilisation de l'informatique distribuée

Application aux crues à cinétique rapide

RESUME : Les crues à cinétique rapide, telles que l'évènement du 8 et 9 septembre 2002 dans le département du Gard, causent régulièrement d'importants dommages socio-économiques. A la suite de cette catastrophe, une réforme des services d'annonce des crues (SAC) a été initiée, et a engendré la transformation des 52 services existants en 22 services de prévision des crues (SPC) et la création d'un service central au niveau national, le SCHAPI. Plus cohérent hydrologiquement, ces services se sont vus assigner une nouvelle mission de prévision des crues. Dans le cadre des prévisions des crues à cinétique rapide, le service de prévision des crues « Grand Delta » (SPC-GD) utilise l'application ALHTAÏR capable d'intégrer des données hydrométéorologiques et la modélisation « temps réel » des débits des cours d'eau des bassins versants amont. Cependant, ce service ne dispose pas à l'heure actuelle d'une véritable capacité de prévision des crues à cinétique rapide. Cette thèse s'attache donc à perfectionner cette capacité de prévision en utilisant la vocation de la technologie grille, une technologie informatique innovante, à fournir de ressources informatiques à la demande et un environnement collaboratif. La méthodologie et les expérimentations ont permis de développer un système spatial d'aide à la décision hydrologique (G-ALHTAÏR) basé sur les ressources de l'architecture de grille européenne EGEE (Enabling Grids for E-sciencE). Des résultats prometteurs montrent une capacité à modéliser un nombre important de scénarios de prévision hydrologique dans un contexte opérationnel, et en adéquation avec les besoins du SPC-GD.

Mots clés : Prévision des crues à cinétique rapide, Modélisation pluie-débit, Technologie grille, Services web de l'OGC, Qualité de service.

Contribution to the enhancement of expertise in crisis situation through distributed computing

Application to flash flood

ABSTRACT: Flash floods, as it happened on September, the 8th and 9th 2002 in the Gard territory, can yield to important economic and human damages. After this catastrophic hydrological situation, a reform of flood warning services (SAC) was initiated. Thus, this political reform has transformed the 52 existing flood warning service into 22 flood forecasting service (SPC) which are more hydrologically consistent, and has created a national central service, the SCHAPI. A new effective hydrological forecasting mission has been assigned to these services. In the context of flash floods, the flood forecasting service Grand Delta (SPC-GD) uses the application ALHTAÏR integrating hydrometeorological data and real-time modeling capabilities of mountainous rivers' discharges. However, this service a limited hydrological forecast ability. Hence, this PhD attempts to enhance this forecasting competency using the purpose of the innovative Grid technology to provide on demand computing resources and collaborative environment. The methodology and experimentations have enabled to implement a hydrological spatial decision support system (G-ALHTAÏR) based on the resources of EGEE (Enabling Grids for E-sciencE) grid architecture. Promising results demonstrate a capacity to model an important number of hydrological forecasts scenarios in an operational context and suitable with SPC-GD requirements.

Keywords : Flash floods forecasting, rainfall-runoff modeling, Grid technology, OGC web services, Quality of service.

MINES ParisTech

ParisTech
INSTITUT DES SCIENCES ET TECHNOLOGIES
PARIS INSTITUTE OF TECHNOLOGY