La Région Lorraine

Pôle de Partenariat Économique
Secteur Recherche et Enseignement Supérieur

Liberté • Égalité • Fraternité
RÉPUBLIQUE FRANÇAISE

MINISTÈRE
DE L'ENSEIGNEMENT SUPÉRIEUR
ET DE LA RECHERCHE

PÔLE DE RECHERCHE SCIENTIFIQUE ET TECHNOLOGIQUE

CONTRAT DE PROJETS ÉTAT-RÉGION 2007-2013

—ooOoo—

Bilan de l'année 2010
Demande de subvention 2011

Intitulé du Pôle de Recherche Scientifique et Technologique
MISN – Modélisations, informations et systèmes numériques

Projet
Expérimentations et calculs Distribués à Grande Échelle (EDGE)

Maître d'ouvrage
INRIA Nancy–Grand Est

Porteurs du projet
Martin Quinson et Lucas Nussbaum

Mots-clés (3 à 5)
Calcul à haute performance, industrie de la connaissance, gestion de ressources, évaluation de performances

Secteur Recherche et Enseignement Supérieur
Place Gabriel Hocquard - BP 81004 - 57036 METZ CEDEX 1

Tél. : 03 87 33 64 04 – Fax. : 03 87 33 64 08
e-mail : desr@lorraine.eu

SOMMAIRE

Volet administratif **3**

Liens avec les pôles de compétitivité **3**

Liens avec le programme « investissements d'avenir » du grand emprunt **4**

Liens avec les objectifs prioritaires de la région **4**

Bilan scientifique du projet pour l'année 2010 **5**

1 Contexte et problématiques du projet **5**

2 Réponses aux expertises 2010 **6**

3 Avancement du projet de recherche **7**

4 Production scientifique **10**

Description du projet pour les années 2011-2013 **12**

1 Objectifs du projet 2011-2013 **12**

2 Positionnement dans le contexte national et international **13**

3 Programme de recherche 2011-2013 **13**

3.1 Axe 1 – Méthodologies et outils de mise au point d'applications distribuées 14

3.1.1 Opération 1.1 – Méthodes, outils et services pour les plates-formes expérimentales 14

3.1.2 Opération 1.2 – Modéliser et simuler les systèmes informatiques de grande taille 20

3.2 Axe 2 – Usage maîtrisé des infrastructures de calcul scientifique 25

3.2.1 Opération 2.1 – Factorisation et logarithmes discrets, applications en cryptanalyse 26

3.2.2 Opération 2.2 – Prouver de grandes formules avec des symboles interprétés 29

3.2.3 Opération 2.3 – Communauté de calcul scientifique en Lorraine . . 30

4 Moyens utilisés et organisation du projet 32

4.1 Personnes 32

4.1.1 Participants actifs du projet EDGE 32

4.1.2 Utilisateurs des ressources expérimentales 32

4.2 Équipements et plates-formes 33

5 Organisation du projet 34

6 Retombées sur la région Lorraine 35

6.1 Avantage stratégique pour les scientifiques locaux 35

6.2 Rayonnement de la région 36

7 Valorisation de la recherche 36

8 Formation 37

Description synthétique actualisée pour 2011-2013 38

Indicateurs actualisés 40

1 Indicateurs de moyens 40

2 Indicateurs de production scientifique 40

3 Autres actions 41

I – VOLET ADMINISTRATIF

Projet : Expérimentations et calculs Distribués à Grande Échelle (EDGE)

Coordonnées des porteurs de projet

Nom : Quinson *Prénom :* Martin *Qualité :* Maître de conférence
Adresse du laboratoire :
Laboratoire Lorrain de Recherche en Informatique et ses Applications (LORIA)
Campus Scientifique
B.P. 239, 54506 Vandœuvre-lès-Nancy Cedex
Téléphone : 0383 59 2098
Adresse électronique : Martin.Quinson@loria.fr

Nom : Nussbaum *Prénom :* Lucas *Qualité :* Maître de conférence
Adresse du laboratoire :
Laboratoire Lorrain de Recherche en Informatique et ses Applications (LORIA)
Campus Scientifique
B.P. 239, 54506 Vandœuvre-lès-Nancy Cedex
Téléphone : 0383 59 8619
Adresse électronique : Lucas.Nussbaum@loria.fr

Établissement maître d'ouvrage : INRIA Nancy–Grand Est

Coordonnées du responsable administratif et financier du projet :

Nom : Moine *Prénom :* Elisabeth
Qualité : Responsable administrative et financière du centre INRIA Nancy - Grand Est
Adresse : Centre de Recherche INRIA-Nancy Grand-Est
Adresse électronique : elisabeth.moine@loria.fr

II – LIENS AVEC LES PÔLES DE COMPÉTITIVITÉ

➢ Ce projet (ou une partie de ce projet) s'inscrit-il dans les thématiques d'un pôle de compétitivité ?
☒Non
☐Oui – lequel ? ☐Materalia
☐Fibres Grand Est (FGE)
☐HYDREOS - Pôle de l'eau

➢ Ce projet (ou une partie de ce projet) sera-t-il soumis à labellisation d'un pôle de compétitivité ou est-il déjà labellisé ?
☒Non
☐Oui – lequel ? ☐Materalia
☐Fibres Grand Est (FGE)
☐HYDREOS - Pôle de l'eau

III – LIENS AVEC LE PROGRAMME « INVESTISSEMENTS D'AVENIR » DU GRAND EMPRUNT

➢ Ce projet (ou une partie de ce projet) est-il en lien avec une réponse à l'un des appels à projets du programme « Investissements d'avenir » du grand emprunt ?
□Non
⊠Oui – lequel ? □EquipEx
⊠LabEx
□IHU
□IRT
□IEED

Précisez les liens :
Ce projet s'inscrit pleinement dans le cadre du **LabEx Charles Hermite – Dynamique Sûreté Géométrie des Systèmes**. L'opération Opération 2.1 (*Factorisation et logarithmes discrets, applications en cryptanalyse*) est centrale dans le thème *Cryptography* du LabEx, tandis que l'opération 2.2 (*Prouver de grandes formules avec des symboles interprétés*) s'inscrit dans le cadre du thème *Formal methods for safety/security issues.*

IV – LIENS AVEC LES OBJECTIFS PRIORITAIRES DE LA RÉGION

➢ Impacts sur le développement durable ; Préciser :
Le calcul scientifique est un gros consommateur d'énergie, à la fois pour l'alimentation électrique et le refroidissement des calculateurs. Ce projet vise à permettre une utilisation plus raisonnée et plus efficace des ressources calcul, et donc à réduire l'impact sur l'environnement de ces calculs.

➢ Contacts avec des entreprises ; Préciser :

➢ Impacts sur les territoires ; Préciser :
L'université du Luxembourg est en train de mettre en place un nœud Grid'5000, qui sera opérationnel courant 2011. Son interconnection réseau au reste de Grid'5000 se fera d'ailleurs à travers Nancy. Grid'5000 est donc un important vecteur de collaborations au sein de la grande région. D'autre part, avec deux sites Grid'5000 (et un autre en cours de montage à Reims), la grande région joue un rôle central dans Grid'5000.

V – BILAN SCIENTIFIQUE DU PROJET POUR L'ANNÉE 2010

1 Contexte et problématiques du projet

Les évolutions technologiques permettent de construire des plates-formes matérielles de plus en plus grandes et complexes. La puissance potentielle des systèmes ainsi constitués offre de nouvelles possibilités en termes d'applications, qu'elles soient scientifiques comme les simulations multi-physiques, grand public comme les systèmes pair-à-pair ou commerciales comme le cloud computing.

Des systèmes informatiques de ces dimensions posent des problèmes méthodologiques et scientifiques spécifiques. Il s'agit par exemple d'assurer l'extensibilité, la tolérance aux fautes, la gestion des données et les performances en général des applications développées pour ces systèmes, grâce à des modèles de programmation, des algorithmes et des outils dédiés.

Pour répondre à ces défis, il est courant d'utiliser des middlewares de plus en plus complets et complexes, comme gLite sur la grille de production européenne EGI, J2EE dans le monde de l'entreprise ou les systèmes de cloud computing récemment proposés par Amazon ou Google. Malheureusement, ces environnements sont eux-mêmes complexes, et il est difficile d'évaluer si les programmes les utilisant offrent des performances satisfaisantes ou s'ils méritent d'être améliorés. De plus, l'évaluation et la mise au point de ces outils eux-mêmes posent des défis techniques et scientifiques d'importance.

Approche expérimentale en informatique. Ces difficultés imposent une approche expérimentale (au sens de la physique, de la chimie ou de la biologie), en complément ou à la place d'approches purement théoriques. Mais réaliser une expérience scientifiquement pertinente dans des environnements distribués à large échelle reste ardu, ne serait-ce qu'à cause de leur très grande dynamicité, qui rend difficile la reproduction d'une expérience donnée dans les mêmes conditions sur une plate-forme réelle donnée.

Grid'5000 [1] est un instrument scientifique permettant de réaliser des expériences à large échelle en informatique distribuée. Il a été spécifiquement conçu pour permettre des expérimentations à tous les niveaux du système (réseau, système d'exploitation, middleware, applications) tout en garantissant un contrôle maximal des conditions expérimentales. Il s'agit d'un projet d'envergure nationale, composé d'environ 2000 machines (5000 cœurs de processeurs) réparties sur neuf sites en France (Bordeaux, Grenoble, Lille, Lyon, Nancy, Orsay, Rennes, Sophia-Antipolis, Toulouse) et reliées par un réseau très haut débit fourni par Renater. Des extensions vers les Pays-Bas et le Japon ont été réalisées, et des extensions vers le Luxembourg et le Brésil sont en cours de réalisation. On peut d'ailleurs noter que l'interconnexion du site Grid'5000 du Luxembourg se fait par la Lorraine.

Simulation de systèmes informatiques. L'expérimentation directe est indispensable pour comprendre les performances d'une application existante sur une plate-forme donnée. Ce protocole expérimental n'est cependant pas adapté à toutes les études. Ainsi, il ne permet pas de déterminer quelle plate-forme construire pour maximiser l'efficacité d'une

1. `http://www.grid5000.fr/`

application donnée, et reste difficile à mettre en œuvre pour le prototypage rapide d'algorithmes. En pareille situation, le recours à la simulation permet de résoudre ces problèmes, comme l'ont démontré d'autres disciplines scientifiques comme la physique ou la biologie. Cette approche pose cependant des défis méthodologiques supplémentaires, tels que la modélisation des systèmes étudiés et le contrôle du biais expérimental induit.

SimGrid [2] est l'un des simulateurs les plus utilisés par les chercheurs en informatique distribuée pour tester leurs algorithmes avant de les mettre en œuvre dans des applications. Il dispose d'une communauté de plusieurs centaines d'utilisateurs, et rayonne largement au delà des frontières de la région lorraine où se trouvent la plupart de ses concepteurs.

Le calcul scientifique pour la recherche en informatique. Héberger un nœud de Grid'5000 est une chance pour la région Lorraine, puisque cela permet à la fois d'accroître la visibilité des chercheurs locaux du domaine et de faire bénéficier leurs collègues non-spécialistes d'un accès à un instrument unique au monde et de l'expertise indispensable pour utiliser des plates-formes de calcul modernes. De même, la proximité de l'équipe réalisant un outil comme SimGrid assure aux scientifiques lorrains la possibilité d'utiliser une expertise précieuse dans la compréhension des systèmes et dans la méthodologie expérimentale adaptée à notre discipline. La proximité géographique constitue un avantage indéniable, mais n'est pas suffisante. C'est pourquoi un travail d'accompagnement des scientifiques non-spécialistes est nécessaire pour leur permettre de maîtriser les outils mis à leur disposition et accroître ainsi leur potentiel de recherche.

Cette démarche n'est ni vraiment une action d'animation classique permettant à des spécialistes d'un domaine donné de se rencontrer, ni seulement une action de transfert à sens unique visant à diffuser un savoir faire des spécialistes vers les utilisateurs potentiels, mais à mi-chemin de ces deux approches. Nous souhaitons mettre en place un cercle vertueux consistant à voir les recherches sur les infrastructures distribuées à large échelle guidées pour répondre du mieux possible aux besoins des utilisateurs potentiels.

Description du projet Le projet « Expérimentations et calculs distribués à grande échelle » (EDGE) vise à développer les méthodologies expérimentales en informatique distribuée à large échelle afin de simplifier l'usage de plates-formes de calcul modernes. Concrètement, le projet s'articule autour de deux grands axes, eux même subdivisés en opérations de recherche spécifiques. Le premier axe porte sur l'amélioration des pratiques et méthodes dans notre domaine tandis que le second constitue des applications pratiques de ces avancées. Ce découpage est présenté brièvement au début de la section 3, page 7 tandis que son détail est rappelé dans le programme de recherche 2011-2013, page 13.

2 Réponses aux expertises 2010

La principale remarque des experts en 2010 était le relatif manque d'**organisation du projet et du document le présentant**. Le dit document a été complètement réécrit. Nous avons explicité le découpage du projet en deux axes principaux, eux mêmes découpés en opérations. Cette organisation est rappelée dans la section suivante. Nous avons également présenté avec plus de rigueur l'organisation de la gouvernance du projet (voir page 34).

Cette réorganisation du document nous a permis de mieux situer chaque opération

2. `http://simgrid.gforge.inria.fr/`

dans l'**état de l'art du domaine**, ce qui constituait une autre remarque des experts en 2010.

Enfin, nous avons présenté la **vue complète de nos objectifs scientifiques** dans EDGE par rapport au document 2010 afin de mieux expliciter le fait que nous ne visons pas seulement à constituer le nœud lorrain de Grid'5000, mais bien à utiliser les ressources expérimentales mises à disposition par ce biais dans un cadre plus général. Ce cadre est présenté dans l'introduction présentée dans la section précédente, ainsi que dans la description synthétique du projet page 38.

3 Avancement du projet de recherche

- Description détaillée des résultats obtenus en 2010
- Description de l'évolution des objectifs et du projet de recherche
- Point sur l'acquisition et la mise en œuvre des équipements

Le projet EDGE n'ayant été introduit dans le CPER MISN qu'en 2010, son organisation en axes et opérations est reconduit en l'état pour la période 2011-2013. Ce découpage est rappelé ci-après et les objectifs précis de chaque partie sont détaillés pages 13 à 31. Nous présentons ici les différents résultats obtenus au cours de l'année 2010. Les publications issues de ces recherches sont rappelées dans le texte, et leur liste complète est donnée page 10.

- Axe 1 « *Méthodologies et outils de mise au point d'applications distribuées* »
  - Opération 1.1 « *Méthodes, outils et services pour les plates-formes expérimentales* »
  - Opération 1.2 « *Modéliser et simuler les systèmes informatiques de grande taille* »
- Axe « *Usage maîtrisé des infrastructures de calcul scientifique* »
  - Opération 2.1 « *Factorisation et logarithmes discrets, applications en cryptanalyse* »
  - Opération 2.2 « *Prouver de grandes formules avec des symboles interprétés, applications à la preuve* »
  - Opération 2.3 « *Communauté de calcul scientifique en Lorraine* »

Opération 1.1 : *Méthodes, outils et services pour les plates-formes expérimentales.*
(voir page 14 pour la présentation du contexte et des objectifs de cette opération)

Les principaux résultats de l'année 2010 concernent l'émulateur Wrekavoc. À l'occasion du stage de master de Tomasz Buchert[1] dans l'équipe AlGorille du LORIA, l'émulation de performances CPU de Wrekavoc a été complètement retravaillée pour permettre d'émuler correctement des nœuds multicoeurs de performances différentes. Ce travail a déjà résulté en une publication [BNG10a], et un deuxième article est en cours de préparation [BNG10b]. Il reste à intégrer ce travail dans l'émulateur Wrekavoc, ce qui sera fait en 2011.

Des travaux préparatoires ont également été réalisés sur les autres tâches, notamment à travers un stage sur la diffusion efficace de données à grande échelle [2].

[1] Tomasz Buchert. Methods for Emulation of Multi-Core CPU Performance. Master's thesis, Poznań University of Technology, Poznań, Poland, 2010.

[2] Guillaume Gallani. Diffusion P2P de données sur grille de calcul. Master's thesis, Ecole des Mines de Nancy, Nancy, France, 2010.

Opération 1.2 : *Modéliser et simuler les systèmes informatiques de grande taille.*
(voir page 20 pour la présentation du contexte et des objectifs de cette opération)

En 2010, une nouvelle API, SMPI, a été ajoutée au simulateur SimGrid afin de permettre la simulation d'applications MPI sur clusters. Les spécificités de ces applications ainsi que les caractéristiques des topologies réseau visées font que les modèles de communication pré-existants ne sont que partiellement applicables (pour les communications de moyenne et grande taille).

Un nouveau modèle a donc été développé, puis testé intensivement à différentes échelles : communications point-à-point, communications collectives et benchmarks.

L'instantiation du modèle a été faite par une régression linéaire par morceaux entre le meilleur modèle linéaire pré-existant et les données expérimentales mesurées à l'aide de l'outil SKaMPI [3] sur une communication point-à-point. Cette opération donne une instance du modèle plus précise que toutes les instances de modèles linéaires tout en restant acceptablement portable sur des plates-formes différentes avec un réseau d'interconnexion similaire (éventuellement hiérarchique).

Les comparaisons de prédiction par simulation effectuées sur des opérations collectives ont permis de mettre clairement en évidence l'impact de la prise en compte de la contention du réseau sur la précision de la prédiction. SimGrid est le seul simulateur à ce jour à revendiquer un modèle de contention précis, ce qui fait de SMPI un outil de choix pour la simulation précise des applications MPI.

Afin d'améliorer la capacité de SMPI à simuler des grosses applications (en terme de quantité de processus impliqués), deux techniques de réduction de l'usage des ressources ont été ajoutées. La première permet de partager l'accès aux données globales de l'application entre les processus, diminuant la quantité de mémoire requise. Cette technique a permis de rendre possible la simulation de problème autrement inaccessibles, sur une machine considérée. La deuxième technique permet de substituer l'exécution réelle de certaines itérations de calcul par une valeur moyenne calculée à partir des mesures des itérations précédentes, sous la contrainte d'une erreur relative contrôlée. Cette technique permet de réduire linéairement le temps de simulation, en fonction du nombre d'itérations simulées.

Les travaux réalisés sur SMPI ont donné lieu à une publication [CSG+10], un deuxième article dans une revue internationale étant en cours de préparation.

Opération 2.1 : *Factorisation et logarithmes discrets, applications en cryptanalyse.*
(voir page 26 pour la présentation du contexte et des objectifs de cette opération)

Début 2010, le projet Caramel et ses partenaires ont achevé la factorisation du nombre RSA-768, établissant un nouveau record mondial de factorisation d'entiers. Le concours de la grille de calcul Grid'5000 a été déterminant pour ce calcul.

L'originalité essentiel du calcul RSA-768 réside dans la grande variété des environnements de calcul utilisés. L'algorithme employé, dit algorithme du crible algébrique, se décompose en deux phases. La première phase est déjà connue, de longue date, comme s'adaptant assez bien à un contexte de calcul largement distribué. Toutefois, les volumes de

[3] Ralf Reussner, Peter Sanders, and Jesper Larsson Träff. SKaMPI : a Comprehensive Benchmark for Public Benchmarking of MPI. *Scientific Programming*, 10(1) :55–65, 2002.

données manipulées (plusieurs téraoctets), et la longue durée des calculs (un peu plus d'un an pour cette première phase) ont conduit à développer des programmes de distribution des calculs automatisant au maximum les tâches. Cela est particulièrement important dans un cadre où les calculs qui ont été menés, de par leur longue durée, sont naturellement placés comme sous-prioritaires par rapport aux travaux d'autres utilisateurs. La plate-forme Grid'5000 dispose de la possibilité d'étiqueter certaines tâches comme « best-effort » signifiant qu'en l'absence de ressources libres, ces tâches sont évincées immédiatement pour laisser la place à d'autres tâches. Ce mode d'utilisation est bien adapté aux calculs de la première phase du crible algébrique. Toutefois, un désordre important est provoqué dans l'organisation des calculs par la fréquence des calculs interrompus. Les programmes qui ont été développés pour la gestion de cet effort de calcul permettent de traiter ces interruptions de calcul comme des événements normaux, et d'automatiser leur traitement : récupération des fichiers de résultats tronqués, définition et replanification des plages de travail non explorées, supervision des tâches en cours. Le développement de tels outils de supervision de calcul a permis de faire de la plate-forme Grid'5000 l'outil principal de l'effort de calcul pour la première phase du calcul RSA-768, tout en conservant le souci de maintenir l'utilisabilité de la plate-forme Grid'5000 en tant que ressource partagée avec d'autres utilisateurs.

La seconde phase de l'algorithme du crible algébrique est la résolution d'un système linéaire. De tels calculs d'algèbre linéaire sont omniprésents dans le domaine du calcul à hautes performances. Toutefois, le système considéré dans le cadre du crible algébrique est défini sur le corps fini GF(2). Dès lors, les calculs à mener sont exacts, et les notions de convergence, ou de rayon spectral, centrales dans le domaine de l'algèbre linéaire numérique, n'ont plus cours.

La résolution des systèmes linéaires intervenant dans le crible algébrique nécessitait encore, il y a dix ans, l'emploi d'un supercalculateur. Pour RSA-768, grâce à l'emploi de l'algorithme de Wiedemann par blocs, il a été possible de mener à bout (en plusieurs mois) la résolution de ce système de manière partiellement distribuée. Le niveau de distribution des calculs menés est sans précédent, avec des ressources de calcul exploitées à la fois en France, en Suisse et au Japon, la contribution des clusters de la plate-forme Grid'5000 (rendus accessibles par l'existance du nœud lorrain) s'avérant là encore déterminante. En particulier, il a été possible de montrer comment les calculs ont pu être menés sur Grid'5000 de manière à la fois efficace, et compatible avec le fait que cette ressource de calcul est partagée avec d'autres utilisateurs.

Opération 2.2 : *Prouver de grandes formules avec des symboles interprétés.*
(voir page 29 pour la présentation du contexte et des objectifs de cette opération)

L'utilisation de Grid'5000 dans notre contexte a nécessité la mise au point de l'infrastructure logicielle de test, qui consiste essentiellement à distribuer les cas de test individuels sur les nombreux noeuds réservés, à faire du « load balancing » et enfin à regrouper et présenter dans un rapport clair les résultats obtenus. Ce logiciel lui-même, GridTPT, est disponible pour la communauté de la preuve automatique, et a fait l'objet d'une publication à un Workshop spécialisé (Pragmatic Aspects of Automatic Reasoning) adossé à la fédération de conférence de référence du domaine (FLoC)[BOD+10b].

Depuis début 2009, l'infrastructure Grid'5000 et GridTPT sont utilisés pour le développement et l'évaluation du logiciel veriT [BOD+10a]. Pour ce, régulièrement, un

grand nombre de noeuds sont réservés pour une courte période, de façon à avoir un retour rapide sur l'impact des techniques mises en oeuvre au sein du solveur veriT, et essentiellement, le nombre de formules inspectées avec succès, et le temps nécessaire à ces inspections. Les ressources sont réservées sur Grid'5000, en priorité basse, c'est-à-dire en mode « best-effort » .

L'utilisation de Grid'5000 a permis de stabiliser le solveur. Il est maintenant distribué, et plusieurs plates-formes de vérification (notamment Coq, Rodin, TLAPS) ont entamé le développement d'un plug-in permettant d'utiliser notre solveur. Nous avons aussi participé à la compétition internationale de solveurs SMT, en 2009 et 2010 avec des résultats encourageants. À la compétition SMT-COMP 2010, veriT s'est classé deuxième dans plusieurs catégories.

Opération 2.3 : *Calcul scientifique en Lorraine.*
(voir page 30 pour la présentation du contexte et des objectifs de cette opération)

Bien que moins formalisés au sein de notre projet, les travaux des équipes lorraines bénéficiant des ressources expérimentales de EDGE sont remarquables. À titre d'exemple, l'équipe Score du LORIA utilise Grid'5000 depuis 2009 afin de valider les performances des architectures distribuées des nouveaux systèmes collaboratifs qu'elle propose.

Ainsi en 2010, l'utilisation de Grid'5000 a permis de valider l'architecture du système UniWiki (`http://sourceforge.net/projects/uniwiki/`) en mesurant ses performances à large échelle. Uniwiki permet de stocker, distribuer et éditer des contenus de type wiki tout en garantissant une haute disponibilité et une capacité de stockage quasi-infini. L'architecture d'Uniwiki repose sur la combinaison innovante d'une table de hachage distribuée et d'un mécanisme de réplication optimiste. L'implémentation met en œuvre un canevas d'interception distribuée orienté aspect (*distributed aspect oriented framework*).

L'expérimentation s'est faite en utilisant des traces d'exécution issues de scénarios réels d'utilisation de Wikipédia. Cette expérimentation, qui a donné lieu à une publication [OMM+10], a été réalisée dans le cadre d'une thèse CIFRE avec la société XWiki SAS et d'une collaboration avec l'Universitat Rovira i Virgili (Catalogne, Espagne). Cette dernière collaboration n'aurait jamais vu le jour sans la présence de Grid'5000 sur le site lorrain.

4 Production scientifique

Publications résultant du projet

[AMQ10] Sabina Akhtar, Stephan Merz and Martin Quinson. *A High-Level Language for Modeling Algorithms and their Properties.* 13th Brazilian Symposium on Formal Methods, Natal, Rio Grande do Norte, Brazil, Nov 8-12, 2010.

[BOD+10a] T. Bouton, D. C. B. de Oliveira, D. Déharbe, and P. Fontaine. veriT : an open, trustable and efficient SMT-solver. In R. Schmidt, editor, *Proc. Conference on Automated Deduction (CADE)*, volume 5663 of *Lecture Notes in Computer Science*, pages 151–156, Montreal, Canada, 2009. Springer.

[BOD+10b] T. Bouton, D. C. B. de Oliveira, D. Déharbe, and P. Fontaine. GridTPT : a distributed platform for Theorem Prover Testing. In B. Konev and R. Schmidt, editors, *Workshop on Practical Aspects of Automated Reasoning (PAAR)*, Edinburgh, UK, 2010.

[BNG10a] Tomasz Buchert, Lucas Nussbaum, and Jens Gustedt. Accurate emulation of CPU performance. In *8th International Workshop on Algorithms, Models and Tools for Parallel Computing on Heterogeneous Platforms (HeteroPar'2010)*, Ischia, Italy, 2010.

[BNG10b] Tomasz Buchert, Lucas Nussbaum, and Jens Gustedt. Methods for Emulation of Multi-Core CPU Performance. Research Report RR-7450, INRIA, 11 2010.

[BSQ10] Laurent Bobelin, Martin Quinson and Frédéric Suter. *Synthesizing Generic Experimental Environments for Simulation.* 5th International Conference on P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC'10), Fukuoka, Japan, Nov 4-6 2010.

[CSG+10] Pierre-Nicolas Clauss, Mark Stillwell, Stéphane Genaud, Frédéric Suter, Henri Casanova, Martin Quinson. *Single Node On-Line Simulation of MPI Applications with SMPI.* 25th IEEE International Parallel & Distributed Processing Symposium (IPDPS'11), May 16-20, 2011, Anchorage (Alaska) USA.

[KAF+10] T. Kleinjung, K. Aoki, J. Franke, A. K. Lenstra, E. Thomé, J. W. Bos, P. Gaudry, A. Kruppa, P. L. Montgomery, D. A. Osvik, H. Te Riele, A. Timofeev, and P. Zimmermann. Factorization of a 768-bit RSA modulus. In T. Rabin, ed., *CRYPTO 2010*, vol. 6223 of *Lecture Notes in Computer Science*, pp. 333–350, États-Unis Santa Barbara, 2010. Springer Verlag. The original publication is available at www.springerlink.com.

[KBL+10] T. Kleinjung, J. Bos, A. Lenstra, D. Arne Osvik, K. Aoki, S. Contini, J. Franke, E. Thomé, P. Jermini, M. Thiémard, P. Leyland, P. Montgomery, A. Timofeev, and H. Stockinger. A heterogeneous computing environment to solve the 768-bit RSA challenge. *Cluster Computing*, 2010.

[KNT10] T. Kleinjung, L. Nussbaum, and E. Thomé. Using a grid platform for solving large sparse linear systems over GF(2). In *11th ACM/IEEE International Conference on Grid Computing (Grid 2010)*, Belgique Brussels, Oct 2010.

[OMM+10] G. Oster, R. Mondéjar, P. Molli and S. Dumitriu. Building a collaborative peer-to-peer wiki system on a structured overlay. *Computer Networks*, 54 :1939–1952, 2010.

[RMQ10] Cristian Rosa, Stephan Merz and Martin Quinson. *A Simple Model of Communication APIs – Application to Dynamic Partial-order Reduction.* 10th International Workshop on Automated Verification of Critical Systems (AVOCS'10), Düsseldorf, Germany, Sept 20-23, 2010.

VI – DESCRIPTION DU PROJET POUR LES ANNÉES 2011-2013

1 Objectifs du projet 2011-2013

L'objectif du projet EDGE est de renforcer l'expertise présente en Lorraine autour du calcul scientifique informatique. Il s'agit à la fois d'établir et de renforcer un groupe de scientifiques spécialistes des problématiques associées et de faire profiter des utilisateurs potentiels de l'expertise ainsi constituée. Les travaux actuels et projetés se découpent donc en deux axes. Le premier vise à permettre l'industrialisation du processus expérimental en calcul scientifique tandis que le second vise à tirer parti des progrès méthodologiques ainsi réalisés pour donner un avantage compétitif aux scientifiques de la région.

L'élément fédérateur du projet est la plate-forme Grid'5000, qui constitue la plus grande partie des moyens demandés dans le projet, mais les travaux dépassent largement le cadre des plates-formes expérimentales avec des recherches sur d'autres méthodes expérimentales (simulation, émulation), des recherches transverses sur la mise au point et l'optimisation de plans expérimentaux, et des applications de ces recherches à d'autres domaines nécessitant la mise en œuvre de calcul scientifique.

Le projet EDGE est relativement neuf dans le CPER MISN puisqu'il n'existe en tant que tel que depuis 2010. Auparavant, ces problématiques étaient étudiées au sein d'une opération spécifique du projet AOC (Analyse, Optimisation et Contrôle). Durant la première phase du CPER (2008-2010), nous nous sommes attachés à renforcer les bases de notre communauté de recherche en Lorraine. Les efforts ont porté plus particulièrement sur :

- Acquisition de clusters reliés à la plate-forme Grid'5000 pour servir de catalyseurs à la fois aux recherches *sur* l'expérimentation distribuée en informatique, et aux recherches *utilisant* cette approche (février 2009 puis septembre 2010).
- Renforcement du pôle de recherche portant sur ces problématiques, par exemple par le recrutement de Lucas Nussbaum (co-porteur du projet) en septembre 2009.
- Formation de chercheurs intéressés par le calcul scientifique, dans des domaines tels que la cryptanalyse ou la preuve automatique.

La seconde phase (2011-2013) vise à continuer et approfondir les travaux et collaborations initiés lors de la première phase. Notre objectif à long terme est de permettre de faire plus avec moins, d'améliorer la rentabilité des plates-formes de calcul scientifique par un usage raisonné des moyens disponibles. Plus particulièrement, les travaux projetés sur la période consisteront à :

- Prolonger les travaux sur les outils nécessaires à l'expérimentation de systèmes informatiques de grande taille pour permettre une **industrialisation du processus**.
- Continuer les efforts sur la modélisation des systèmes informatiques de grande taille pour améliorer la **compréhension de ces systèmes** afin d'optimiser leur usage potentiel.
- Élargir la communauté des **utilisateurs avancés** de ces outils au travers de collaborations afin de leur donner un avantage compétitif dans leurs propres travaux tout en validant les recherches des spécialistes.

2 Positionnement dans le contexte national et international

Bien qu'étant relativement récent, le projet EDGE bénéficie d'un très bon ancrage au niveau national et international qui se caractérise par :

- Un bassin important de recherche en informatique en Lorraine, concentrant des spécialistes de nombreux domaines de cette discipline, et une forte tradition de calcul scientifique partagée également par les informaticiens non-spécialistes du calcul haute performance. La Lorraine fut ainsi pionnière du domaine au travers du centre de calcul Charles Hermite dès les années 90.
- Une bonne visibilité scientifique nationale et internationale résultant d'une participation continue aux grandes conférences du domaine du calcul distribué à large échelle et de coordination et collaboration au travers de projets nationaux.
- Une expertise technique reconnue au niveau international dans le domaine des méthodologies d'expérimentation pour les systèmes informatiques de grande taille, notamment au travers d'outils développés localement :
  - SimGrid est le simulateur majeur pour le calcul distribué en informatique, utilisé à la fois pour étudier des grilles de calcul, des systèmes pair-à-pair, des plates-formes de *Volunteer computing* et des *Data Grids*.
  - Wrekavoc est l'un des seuls émulateurs diffusés permettant une étude conjointe des communications et des calculs dans un système informatique distribué.
  - La Lorraine est par ailleurs un site moteur de Grid'5000, et les compétences locales ont permis améliorer des outils tels que Kadeploy (déploiement de systèmes distribués).

Comme expliqué en section 6, page 35, l'expertise acquise au sein du projet constitue de plus à la fois un élément de rayonnement scientifique de la Lorraine reconnu par les régions voisines, et un avantage stratégique pour les scientifiques locaux. Nous visons à établir des passerelles pratiques et des mutualisations entre les plates-formes expérimentales comme celles proposées par EDGE et les ressource de calcul de production utilisées par les autres disciplines scientifiques.

3 Programme de recherche 2011-2013

Notre programme de recherche s'organise en deux axes, composés respectivement de deux et trois opérations :

- Axe 1 « *Méthodologies et outils de mise au point d'applications distribuées* »
  - Opération 1.1 « *Méthodes, outils et services pour les plates-formes expérimentales* »
  - Opération 1.2 « *Modéliser et simuler les systèmes informatiques de grande taille* »
- Axe « *Usage maîtrisé des infrastructures de calcul scientifique* »
  - Opération 2.1 « *Factorisation et logarithmes discrets, applications en cryptanalyse* »
  - Opération 2.2 « *Prouver de grandes formules avec des symboles interprétés, applications à la preuve* »
  - Opération 2.3 « *Communauté de calcul scientifique en Lorraine* »

3.1 Axe 1 – Méthodologies et outils de mise au point d'applications distribuées

Cet axe de travail vise à articuler les différentes approches de tests (simulation, émulation et test sur plate-forme réelle d'expérimentation telle que Grid'5000) afin de garantir aux solutions développées les meilleures performances possibles lors de leur utilisation en production. L'objectif à plus long terme est de permettre une industrialisation du processus expérimental en informatique distribuée afin d'optimiser les ressources et de conférer un avantage stratégique aux chercheurs dépendants du calcul scientifique dans leurs travaux.

Cet axe se découpe en deux opérations portant respectivement sur l'optimisation de l'usage des plates-formes expérimentales et sur l'usage de la simulation pour aider la compréhension et la prédiction des performances des systèmes informatiques distribués.

3.1.1 Opération 1.1 – Méthodes, outils et services pour les plates-formes expérimentales

Participants

- Équipe AlGorille/LORIA : Sébastien Badia (ingénieur), Sylvain Contassot (PR), Jens Gustedt (DR INRIA), Lucas Nussbaum (MCF, responsable de l'opération), Martin Quinson (MCF), Tina Rakotoarivelo (ingénieur).
- Équipe Score/LORIA : François Charoy (MCF).

Contexte et objectifs

Les plates-formes d'expérimentation comme EDGE, ou plus généralement Grid'5000, sont souvent vues comme des environnements parfaits pour réaliser des expériences complexes à grande échelle sur les systèmes distribués (calcul à hautes performances, systèmes peer-to-peer, Grilles, Cloud Computing). Grid'5000 fournit aux chercheurs un grand nombre de nœuds (2000 nœuds répartis dans 9 sites), et plusieurs fonctionnalités avancées pour réaliser des expériences : la possibilité de déployer son propre système d'exploitation sur les nœuds, un réseau dédié et performant, le support de la virtualisation, et une grande variété de technologies matérielles.

Toutefois, si ces fonctionnalités fournissent tout ce qui peut être attendu d'une plate-forme expérimentale, les utilisateurs de Grid'5000 rencontrent de nombreux problèmes quand ils essaient de réaliser des expériences complexes ou à grande échelle. En octobre 2010, nous avons réalisé un sondage auprès des utilisateurs de Grid'5000 pour les interroger sur les facteurs limitants lors d'expériences sur Grid'5000. Les réponses indiquent que certains services de l'infrastructure posent des problèmes de fiabilité (ce qui n'est pas surprenant puisque ces services fournissent des fonctionnalités uniques au monde, qui ont été développées spécifiquement pour Grid'5000). Ce manque de fiabilité est difficile à gérer pour les utilisateurs. Un utilisateur écrit : « *Nous avons utilisé 4 sites. Nous aurions aimé en utiliser 8, mais il est déjà compliqué de réserver et faire fonctionner 4 clusters en même temps.* » La grande échelle multiplie les problèmes et limite également les expériences. Un autre utilisateur écrit : « *Des expériences à plus grande échelle étaient prévues (plusieurs centaines de machines) mais les problèmes pratiques rencontrés ont limité nos ambitions* ».

Il apparaît que la plupart des expériences à grande échelle réalisées sur Grid'5000 ont

nécessité soit d'avoir un expérimentateur possédant une grande expertise technique, capable de régler les problèmes les plus ardus, soit de limiter la complexité de l'expérience en évitant l'utilisation de certaines fonctionnalités avancées. Dans le cas de l'expérience RSA-768[4] par exemple, le principal expérimentateur a été confronté de nombreux problèmes techniques et a remonté plus de 30 bugs à l'équipe technique de Grid'5000.

Même dans le cadre d'expériences peu complexes, il est aujourd'hui difficile d'automatiser le déroulement d'une expérience, ce qui affecte la qualité méthodologique des expériences réalisées sur la plate-forme. Lors d'une expérience dont les différentes étapes sont exécutées manuellement par l'expérimentateur, il est beaucoup plus difficile d'assurer que les conditions expérimentales correspondaient à celles attendues ou que l'expérience est reproductible. De plus, reproduire l'expérience plusieurs fois devient rapidement fastidieux, alors que c'est nécessaire pour garantir la validité statistique des résultats.

L'objectif de cet axe du thème EDGE du CPER MISN est d'augmenter la qualité de la recherche sur les systèmes distribués en industrialisant le processus expérimental : nous allons concevoir et développer des méthodes et des logiciels qui permettront aux utilisateurs de réaliser des expériences plus facilement, ou de dépasser les limites de ce qui est actuellement atteignable sur EDGE et Grid'5000. Notre travail s'effectuera à différents niveaux (figure 1) :

D'abord, nous concevrons et développerons *des services de base qui sont requis par la plupart des expériences* : contrôle d'un grand nombre de nœuds, gestion des données, émulation de conditions expérimentales, injection de charge et de fautes, instrumentation et monitoring, ... Dans certains cas, nous nous baserons sur des solutions existantes, mais dans la plupart des cas, aucune solution acceptable n'existe, et les expérimentateurs utilisent des solutions ad-hoc qu'ils ont développé eux-mêmes. Ce travail bénéficiera à toute la communauté en apportant des outils efficaces et fiables pour leurs expériences.

Ensuite, nous concevrons et développerons *un intergiciel pour l'orchestration d'expériences complexes et/ou à grande échelle*. Les expériences sont typiquement composées de nombreuses étapes qui peuvent échouer partiellement (seulement sur certains nœuds) ou complètement. Actuellement, la plupart des experimenteurs conduisent leurs expériences manuellement, car les *programmer* est trop difficile à cause de la nécessité de gérer différents types de pannes. Dans le meilleur des cas, les expérimentateurs écrivent des scripts ad-hoc pour contrôler leurs expériences, mais ignorent souvent la gestion des pannes.

Finalement, nous validerons nos travaux en les utilisant pour conduire des expériences complexes à grande échelle, largement au-delà de ce qui a été possible jusqu'à maintenant sur Grid'5000. En particulier, nous travaillerons avec des spécialistes des grilles de production pour déployer le middleware gLite sur Grid'5000 et permettre ainsi d'en évaluer des évolutions.

Positionnement scientifique

Le problème de l'industrialisation des processus expérimentaux se pose dans de nombreuses communautés. Des travaux similaires sont ainsi réalisés sur d'autres plates-formes

[4] Communiqué de presse INRIA : L'INRIA et ses partenaires battent un nouveau record de calcul et démontrent la vulnérabilité d'une clé RSA de 768 bits. `http://www.inria.fr/actualites/espace-presse/cp/pre210.fr.html`.

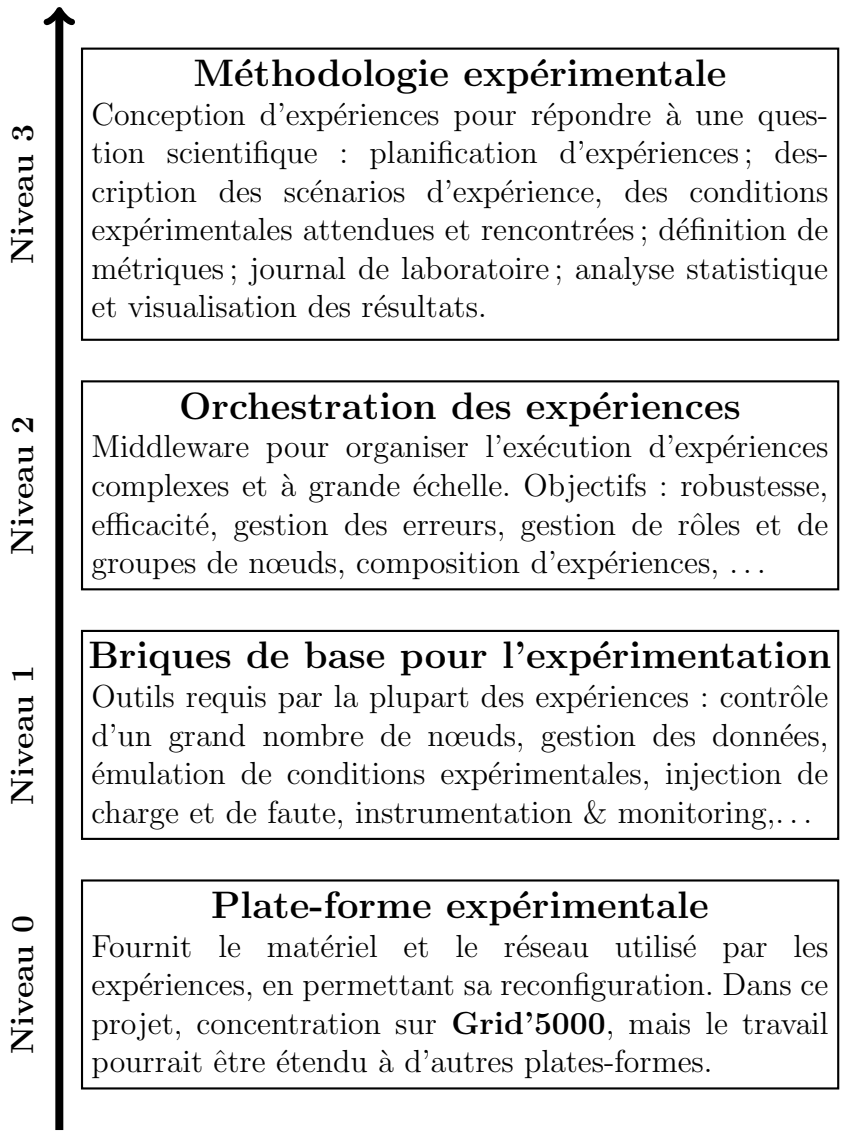

FIGURE 1 – Pile logicielle et méthodologique pour l'expérimentation.

pour l'expérimentation sur les systèmes distribués.

PlanetLab (`http://www.planet-lab.org/`, [5]) est une plate-forme d'expérimentation distribuée pour l'évaluation des applications et services de l'Internet du futur. La plate-forme contient actuellement plus de 500 nœuds actifs distribués sur Internet. Son objectif est légèrement différent de celui des plates-formes EDGE et Grid'5000, avec un focus sur les applications d'Internet : PlanetLab ne contient pas de *cluster* de machines connectées de manière proche. L'outil recommandé pour réaliser des expériences sur PlanetLab est Plush[6].

[5] Brent Chun, David Culler, Timothy Roscoe, Andy Bavier, Larry Peterson, Mike Wawrzoniak, and Mic Bowman. PlanetLab : an overlay testbed for broad-coverage services. *SIGCOMM Comput. Commun. Rev.*, 33(3) :3–12, 2003.

[6] Jeannie Albrecht, Christopher Tuttle, Alex C. Snoeren, and Amin Vahdat. PlanetLab application management using plush. *SIGOPS Oper. Syst. Rev.*, 40(1) :33–40, 2006.

Emulab[7] est une plate-forme initiée par l'université d'Utah. Bien que plus petite en taille que Grid'5000, elle fournit des fonctionnalités plus avancées sur la reconfiguration réseau. La gestion des expériences est intégrée à l'infrastructure, avec des outils pour gérer le cycle de vie d'une expérience dans Emulab, en incluant sa description, sa configuration, la séquence des différentes étapes, et l'enregistrement des résultats[8,9].

Le projet GENI (`http://www.geni.net/`) vise à construire une plate-forme expérimentale de grande taille qui capitalisera sur d'autres plates-formes comme PlanetLab et Emulab. Il est intéressant de noter que GENI inclut un groupe de travail actif dédié à la gestion des expériences. (`http://groups.geni.net/geni/wiki/GeniServices`), dont les objectifs sont proches de ceux de ce projet. Gush[10] est un outil de contrôle d'expériences basé sur Plush, et adapté aux besoins de GENI.

D'autres plates-formes sont également développées aux Etats-Unis. Le projet FutureGrid (`http://futuregrid.org/`) vise à construire une plateforme dont les buts sont similaires à Grid'5000. Le projet français est d'ailleurs régulièrement cité comme source d'inspiration par FutureGrid.

La communauté des réseaux de capteurs sans fil française construit également sa propre plate-forme expérimentale, financée par le projet ANR SensLAB (`http://www.senslab.info/`). Bien que ce domaine ait des spécificités qui rendent difficile une collaboration sur des outils logiciels, l'expertise et les méthodes de travail pourra être partagée entre SensLAB et les travaux entrepris au sein du projet EDGE.

Enfin, en Europe, les projets néerlandais *DAS-3* et *DAS-4* (2010) forment une grille expérimentale homogène constituée de clusters situés dans 4 universités participantes, reliés par un réseau d'interconnexion optique reconfigurable. Les réseaux de Grid'5000 et *DAS-3* sont interconnectés, permettant de réaliser des expériences distribuées sur les deux plates-formes.

Bien que les outils développés sur d'autres plates-formes soient évidemment d'intérêt pour les objectifs de ce projet, il y a plusieurs raisons pour lesquelles ces logiciels ne peuvent pas facilement être transférés vers notre contexte. D'abord, ces applications ont été développées pour une plate-forme cible précise, et sont fortement couplées avec l'infrastructure de cette plate-forme. Ensuite, le champ des expériences possibles sur Grid'5000 est plus large que sur les autres plates-formes, qui sont plus spécialisées. Ces applications n'ont pas forcément été conçues pour aller au-delà de ce qui est possible sur leur plate-forme d'origine. Pour s'adapter à la variété des expériences possibles sur Grid'5000, il est nécessaire de concevoir les logiciels avec une approche multi-couches et orientée services,

[7] Brian White, Jay Lepreau, Leigh Stoller, Robert Ricci, Shashi Guruprasad, Mac Newbold, Mike Hibler, Chad Barb, and Abhijeet Joglekar. An Integrated Experimental Environment for Distributed Systems and Networks. In *OSDI'02*, Boston, MA, 2002.

[8] Eric Eide, Leigh Stoller, Tim Stack, Juliana Freire, and Jay Lepreau. Integrated scientific workflow management for the Emulab network testbed. In *ATEC '06 : Proceedings of the annual conference on USENIX '06 Annual Technical Conference*, pp 33–33, Berkeley, CA, USA, 2006. USENIX Association.

[9] Eric Eide, Leigh Stoller, and Jay Lepreau. An Experimentation Workbench for Replayable Networking Research. In *4th Symposium on Networked Systems Design and Implementation (NSDI 2007)*, 2007.

[10] Jeannie Albrecht and Danny Yuxing Huang. Managing Distributed Applications using Gush. In *Sixth International Conference on Testbeds and Research Infrastructures for the Development of Networks and Communities, Testbeds Practices Session (TridentCom)*, 2010.

pour arriver dans une situation où les expériences seront capables de choisir les briques de base correspondant aux besoins particuliers d'une expérience.

Au fil des ans, la communauté E-Science a augmenté son utilisation de ressources computationnelles pour s'attaquer à des problèmes scientifiques. Les infrastructures de production comme EGI (précédemment EGEE) ont été créées pour permettre de grandes campagnes de calcul, et plusieurs outils comme Taverna (`http://www.taverna.org.uk/`), Kepler (`https://kepler-project.org/`) et MOTEUR[11] ont été développés pour contrôler ces campagnes d'expériences.

Finalement, le réseau social `http://myexperiment.org/` [12] est utilisé par plus de 1200 utilisateurs pour partager des workflows scientifiques dans de nombreux domaines (principalement en bio-informatique). Cela montre l'importance de la méthodologie expérimentale dans certains domaines scientifiques. La structuration de ces expériences est aussi utilisée pour partager et réutiliser des morceaux de workflows expérimentaux entre des groupes de recherche, même si ce processus est loin d'être trivial[13].

Programme de travail

L'objectif de ce projet est d'industrialiser le processus expérimental en réalisant des travaux aux trois niveaux de la pile logicielle et méthodologique pour l'expérimentation (figure 1).

Tâche 1 : Briques et services de base pour l'expérimentation. Grid'5000 propose une API de type service web REST permettant d'effectuer les opérations de recherche de ressources correspondant à des critères, de réservation de ressources, et de déploiement. Toutefois, une fois les ressources réservées, et l'environnement de travail déployé, Grid'5000 ne propose pas d'outil standard pour réaliser des expériences.

De nombreux outils ont été développés au fil des ans pour réaliser les fonctions de base nécessaires à la plupart des expériences, soit dans le cadre de Grid'5000, soit dans d'autres contextes proches. Ils ont atteint différents niveaux de qualité et de finition : certains sont restés au stade de prototypes, d'autres ont été abandonnés après avoir été utilisés. Pour la plupart, ils n'ont pas été conçus pour être utilisés par un outil de plus haut niveau (typiquement, un outil de conduite d'expérience), mais plutôt directement par l'utilisateur, avec une interface *shell*.

Nous évaluerons les briques de base existantes pour identifier celles qui sont réutilisables et recommandables dans le cadre de l'expérimentation sur Grid'5000, en mettant l'accent sur :

- Leur robustesse ;
- Leur capacité à passer à l'échelle ;
- La possibilité de les transformer en *service* utilisable depuis une application de plus

[11] T. Glatard, J. Montagnat, D. Lingrand, and X. Pennec. Flexible and efficient workflow deployement of data-intensive applications on grids with MOTEUR. *Journal of High Performance Computing Applications*, 22(3) :347–360, 2008.

[12] David De Roure, Carole Goble, and Robert Stevens. The design and realisation of the Virtual Research Environment for social sharing of workflows. *Future Generation Computer Systems*, 25(5) :561 – 567, 2009.

[13] Wei Tan, Jia Zhang, and I. Foster. Network Analysis of Scientific Workflows : A Gateway to Reuse. *Computer*, 43(9) :54 –61, September 2010.

haut niveau.

Dans un premier temps, nous nous concentrerons sur les outils permettant de contrôler efficacement un grand nombre de machines (lanceurs parallèles, par exemple), sur la problématique de la gestion des données, et sur les outils de qualification de la plate-forme (permettant de vérifier les conditions expérimentales avant de commencer une expérience). Dans un deuxième temps, d'autres catégories d'outils pourront être abordés, tels que des émulateurs, des outils d'instrumentation et de monitoring ou encore des injecteurs de charge et de fautes.

Après avoir sélectionné des briques de base répondant aux critères de qualité requis, nous développerons des couches d'abstraction permettant de les utiliser comme des *services* utilisables à partir d'outils de plus haut niveau comme le moteur de workflow décrit au paragraphe suivant.

Tâche 2 : Conception d'un moteur de workflow pour l'expérimentation. En première approximation, une expérience menée sur Grid'5000 peut facilement s'apparenter à un flot de travail, modélisable, exécutable et contrôlable par un moteur de workflow. Les scripts lancent l'exécution de tâches ou de services sur les différents sites ou nœuds du système qui sont autant d'acteurs d'un processus. On peut cependant vite identifier des points particuliers au contexte qui nous intéresse qui ne sont pas classiques dans les métiers cibles des systèmes de gestion de workflow.

- le faible nombre d'exécution des processus
- le nombre potentiellement très important des activités dans un processus
- des synchronisations potentiellement partielles
- un fort usage de la multi-instanciation dans les processus
- la nécessité de corréler des exécutions successives de processus
- la nécessité de gérer des évolutions de processus et leur historique

Ces différents points nous font penser qu'il sera nécessaire de réaliser dans un premier temps une étude sur plusieurs moteurs de workflow pour les qualifier pour la conduite d'expérience Grid'5000. Nous pensons par exemple à ruote (`http://ruote.rubyforge.org/`, un DSL en ruby adapté au scripting), à Bonita (`http://www.bonitasoft.com/`) pour sa grande capacité d'intégration et sa souplesse, et à un moteur BPEL pour l'interopérabilité et le fait que ce soit un standard industriel.

Cette première étape nous permettra de comprendre les limites de ces systèmes par rapport au métier particulier de l'expérimentation. Elle devra nous permettre ensuite de choisir une des solutions expérimentées puis de proposer des adaptations à son moteur. Nous espérons ainsi fournir un réel support à l'industrialisation de l'expérimentation sur Grid'5000. Du point de vue de l'orchestration, nous espérons apprendre du comportement particulier d'un moteur de workflow confronté à l'exécution d'un processus composé d'un très grand nombre d'activités. Ceci pourrait éventuellement s'appliquer à d'autres contextes comme des processus de *crowdsourcing* par exemple.

Tâche 3 : Validation à l'aide d'expériences complexes à grande échelle. Afin de valider les développements, nous procéderons à une série d'expériences visant à démontrer l'apport des solutions proposées pour la conduite d'expériences complexes à grande échelle. En particulier, nous recréerons l'infrastructure de la grille de production EGEE sur Grid'5000 en y déployant le middleware gLite dans une architecture multi-VO (*Virtual Or-*

ganization), multi-CE (*Computing Element*, ≈ cluster) et multi-SE (*Storage Element*), et exécuterons une ou des applications typiques des grilles de production sur l'infrastructure déployée. Cela pourra servir de base à des expériences sur des évolutions de l'infrastructure gLite ou sur des aspects plus applicatifs (par exemple, l'évaluation des stratégies de soumissions de tâches sur les grilles de production).

Cette expérience pose de plus gros problèmes de configuration et de synchronisation pendant les étapes de configuration. Une implémentation efficace devra être capable de configurer simultanément certains services de gLite.

3.1.2 Opération 1.2 – Modéliser et simuler les systèmes informatiques de grande taille

Participants :

- Équipe AlGorille/LORIA : Pierre-Nicolas Clauss (post-doctorant), El Mehdi Fekari (ingénieur), Lucas Nussbaum (MCF), Martin Quinson (MCF, responsable de l'opération), Cristian Rosa (doctorant), Christophe Thiéry (ingénieur).

Contexte et motivation

L'expérimentation directe telle qu'envisagée dans la première opération est indispensable pour comprendre les performances d'une application existante sur une plate-forme donnée. Ce protocole expérimental est en effet particulièrement efficace pour étudier et optimiser une solution existante, mais il n'est cependant pas adapté à toutes les études. Ainsi, il ne permet pas de déterminer quelle plate-forme construire pour maximiser l'efficacité d'une application donnée, et reste difficile à mettre en œuvre pour le prototypage rapide d'algorithmes. La simulation permet de résoudre ces problèmes, mais au prix de défis méthodologiques supplémentaires, tels que la modélisation des systèmes informatiques ou le contrôle du biais expérimental induit, auxquels cette opération vise à répondre.

Les avantages de la simulation des phénomènes complexes ne sont plus à démontrer en science. Cette approche est par exemple tellement utilisée dans d'autres disciplines telles que la physique, la biologie ou l'ingénierie qu'il est d'usage de la considérer comme une troisième voie scientifique, aux côtés de la théorie et de l'expérimentation directe.

Paradoxalement, alors que l'outil informatique est à la base de la simulation dans les autres disciplines scientifiques, force est de constater que cette approche souffre d'une certaine défiance dans notre discipline. En témoigne l'absence de standards établis en la matière et de modèles communément acceptés de tous. Ce phénomène est particulièrement criant dans le domaine de l'informatique distribuée et du calcul à haute performance, où les spécialistes s'attachent à construire des solutions de très grande taille. Citons par exemple la machine Jaguar Cray XT5[14] aux États-Unis qui regroupe 18 688 nœuds de calcul, dotés chacun de deux processeurs hexa-cores et 16GB de mémoire, pour une consommation électrique totale de 7 MW[15,16]. Bien que largement inférieure à la consommation des

[14] The Jaguar XT5 system. `http://www.nccs.gov/jaguar`.

[15] The Green500 List : Environmentally Responsible Supercomputing. `http://www.green500.org`.

[16] W. Feng and T. Scogland. The Green500 List : Year one. In *IPDPS '09 : Proceedings of the 2009 IEEE International Symposium on Parallel&Distributed Processing*, pp 1–7, Washington, DC, USA, 2009. IEEE Computer Society.

Data Center de l'Internet commercial[17] (comme le Data Center Microsoft de Chicago dont la consommation dépasse 190 MW[18]), cette consommation électrique correspond déjà à celle d'une petite ville.

Ces systèmes ne sont certes pas construits uniquement à des fins de validation en informatique, mais des communications privées avec des utilisateurs des ressources de calcul offertes par le projet EGI en Europe nous laissent penser que plus d'un tiers des calculs sont effectués à seules fins de test, debug et optimisation.

S'il est clair que l'approche expérimentale a un rôle majeur à jouer en informatique (ne serait-ce que pour *valider* les modèles utilisés en simulation), il semble cependant difficile voire dangereux de se limiter à cette seule méthodologie. De notre point de vue, il est nécessaire d'apporter une réponse méthodologique globale reposant à la fois sur *la simulation*, en première approche afin comprendre les phénomènes étudiés, et sur *l'expérimentation* pour confirmer et affiner les conclusions de l'étude. C'est pourquoi nous nous proposons de mettre ces deux méthodologies à l'œuvre de manière conjointe dans le projet EDGE.

Positionnement scientifique

Bien que moins développée que dans d'autres disciplines, la simulation n'est pas complètement absente des études en informatique. Des outils et modèles standards existent par exemple en conception de circuits[19] ou pour l'étude des réseaux[20,21,22], même si les limites de l'approche par simulation uniquement ont parfois été mises à l'index[23].

Dans le domaine de l'informatique distribuée, nous assistons à la fois à une pauvreté d'outils, modèles et méthodologies standards, et une abondance d'outils spécifiques. Dans [24], les auteurs étudient 141 articles du domaine pair-à-pair dont les expérimentations ont été réalisées par simulation. Ils rapportent que 30% utilisent un outil développé aux seules fins de l'étude présentée tandis que 50% ne précisent même pas le simulateur utilisé !

Cette situation s'explique sans doute par l'apparente simplicité d'un simulateur de système informatique distribué. Jusque récemment, les systèmes restaient suffisamment simples pour pouvoir être « simulés » à la main : le matériel était homogène et sans hiérarchie profonde et les programmes alternaient sans recouvrement les étapes de cal-

[17] A. Qureshi, R. Weber, H. Balakrishnan, J. Guttag, and B. Maggs. Cutting the electric bill for internet-scale systems. *SIGCOMM Comput. Commun. Rev.*, 39(4) :123–134, 2009.

[18] Rich Miller. Microsoft's 198 Megawatts of Motivation. `http://www.datacenterknowledge.com/archives/2008/04/04/microsofts-198-megawatts-of-motivation/`.

[19] Derek Hower, Luke Yen, Min Xu, Milo Martin, Doug Burger, and Mark Hill. WWW Computer Architecture Page. `http://pages.cs.wisc.edu/~arch/www/tools.html`.

[20] The Network Simulator (ns2). `http://nsnam.isi.edu/nsnam/`.

[21] James H. Cowie, David M. Nicol, and Andy T. Ogielski. Modeling the Global Internet. *Computing in Science and Engineering*, 1(1) :42–50, 1999.

[22] George F. Riley. The Georgia Tech Network Simulator. In *ACM SIGCOMM workshop on Models, Methods and Tools for Reproducible Network Research*, pp 5–12, 2003.

[23] Sam Jansen and Anthony McGregor. Performance, Validation and Testing with the Network Simulation Cradle. In *Proceedings of the 14th IEEE International Symposium on Modeling, Analysis, and Simulation*, pp 355–362, Washington, DC, USA, 2006. IEEE Computer Society.

[24] S. Naicken, A. Basu, B. Livingston, S. Rodhetbhai, and I. Wakeman. Towards Yet Another Peer-to-Peer Simulator. In *Proceedings of The Fourth International Working Conference on Performance Modelling and Evaluation of Heterogeneous Networks (HET-NETs), Ilkley, UK*, 2006.

cul et communication. Dans ces conditions, nul besoin d'un simulateur sophistiqué pour établir un diagramme de Gantt ou dénombrer les messages échangés, et chacun pouvait créer un petit outil adapté à ses besoins. La situation est malheureusement bien différente maintenant que le matériel présente des hiérarchies importantes (multi-core, multi-processeurs, clusters de machines, grilles de clusters) et des différences tant quantitatives que qualitatives (CPU contre GPGPU et réseau infiniband contre ethernet, ou bien connexion ADSL en bordure de réseau contre connexion optique au cœur du réseau). En réaction, les algorithmes et logiciels deviennent toujours plus complexes, et leur étude ne devrait plus se faire au travers d'outils et modèles qui ne soient pas rigoureusement validés.

Avant de présenter les divers simulateurs existant dans la littérature, il convient de revenir sur les qualités attendues d'un tel outil pour qu'il constitue une réelle aide dans le travail des scientifiques. De notre avis, il faut qu'il possède les quatre caractéristiques suivantes pour cela :

- **Validité :** Il faut tout d'abord que les modèles utilisés soient réalistes, c'est à dire que les résultats obtenus par le biais du simulateur soient représentatifs des résultats qui auraient été obtenus par expérimentation directe.
- **Extensibilité :** Il faut également que le simulateur soit suffisamment rapide et économe en mémoire pour pouvoir être utilisé en pratique. Si chaque simulation d'intérêt demande un mois de calcul (ou pire, s'il est impossible de réaliser l'étude car elle demanderait trop de mémoire), il devient difficile de l'utiliser dans le cadre d'une étude donnée.
- **Utilisabilité :** La simulation n'est que l'une des étapes du processus expérimental, et des outils associés (visualisation, gestion de campagne) sont indispensables pour mener une approche cohérente.
- **Applicabilité :** Il faut que l'outil propose des objets représentant les concepts d'intérêt de la communauté de recherche visée : batch scheduler et serveurs de stockage pour les grilles, machine virtuelle pour les clouds, ainsi que peer et overlay pour les systèmes P2P.

Aucun des simulateurs de systèmes distribués existants ne nous semble satisfaire tous ces critères à la fois. La validité des simulateurs dédiés aux réseaux[20,21,22] est bonne, mais ils ne sont pas applicables aux études mettant en œuvre les calculs puisque seul le réseau est modélisé. Les simulateurs issus de la communauté du P2P[25,26] sont extrêmement extensibles, mais au prix d'un réalisme moindre. Dans PeerSim par exemple, le comportement de l'application étudiéee est représentée par un automate et non programmé

[25] Pedro García, Carles Pairot, Rubén Mondéjar, Jordi Pujol, Helio Tejedor, and Robert Rallo. PlanetSim : A New Overlay Network Simulation Framework. In *Software Engineering and Middleware, SEM 2004*, volume 3437 of *LNCS*, pp 123–137, Linz, Austria, March 2005.

[26] M. Jelasity, A. Montresor, G. P. Jesi, and S. Voulgaris. PeerSim. `http://peersim.sf.net/`.

directement. Les simulateurs issus des grilles[27,28] et des clouds[29,30] privilégient l'applicabilité. Ces outils présentent ainsi tous les concepts nécessaires aux études visées, mais n'offrent pas une extensibilité poussée, et la validation des outils reste cependant à montrer. Dans le cas de OptorSim et GroudSim, les auteurs ont même privilégié le génie logiciel de leurs outils à la validité des résultats, et le fait que les résultats sont faux quand la plate-forme n'est pas parfaitement homogène est même indiqué explicitement dans la documentation...

SimGrid est peut-être le plus générique des outils existants : Issu de la communauté des grilles, son champ d'applicabilité a depuis été accru au P2P et au volunteer computing[31]. Sa validation a fait l'objet d'études poussées de la part de nos collaborateurs[32], et son extensibilité a été prouvée par des experts extérieurs[33,34]. En ce qui concerne l'utilisabilité, des outils existent pour la visualisation et même le model-checking des algorithmes écrits dans ce cadre.

La figure 2 récapitule l'état de l'art du domaine.

Programme de travail

Les objectifs des travaux projetés dans cette opération visent tout d'abord une convergence méthodologique de l'expérimentation entre SimGrid et les plates-formes expérimentales comme Grid'5000. Par ailleurs, nous comptons utiliser les ressources expérimentales mises à disposition pour améliorer le simulateur lui-même, en particulier en ce qui concerne la validité de ses modèles.

Tâche 1 : Plan d'expériences et simulation. Un plan d'expérience est une suite ordonnée d'expériences à réaliser pour confirmer ou infirmer une hypothèse. Dans ce contexte, l'expérimentateur est intéressé par les effets d'un processus ou d'une intervention (le « traitement ») sur certains objets (les « sujets »), qui peuvent être des per-

[27] William H. Bell, David G. Cameron, Luigi Capozza, A. Paul Millar, Kurt Stockinger, and Floriano Zini. OptorSim - A Grid Simulator for Studying Dynamic Data Replication Strategies. *International J. of High Performance Computing Applications*, 17(4), 2003.

[28] Rajkumar Buyya and Manzur Murshed. GridSim : A Toolkit for the Modeling and Simulation of Distributed Resource Management and Scheduling for Grid Computing. *J. of Concurrency and Computation : Practice and Experience (CCPE)*, 14(13-15), Decembre 2002.

[29] R. Calheiros, R. Ranjan, A. Beloglazov, C. De Rose, and R. Buyya. CloudSim : A Toolkit for Modeling and Simulation of Cloud Computing Environments and Evaluation of Resource Provisioning Algorithms. *Software : Practice and Experience (SPE)*, 41(1) :23–50, January 2011.

[30] S. Ostermann, R. Prodan, and T. Fahringer. Dynamic Cloud Provisioning for Scientific Grid Workflows. In *The 11th ACM/IEEE International Conference on Grid Computing (Grid 2010)*, Brussels, Belgium, 2010.

[31] Bruno Donassolo, Henri Casanova, Arnaud Legrand, and Pedro Velho. Fast and Scalable Simulation of Volunteer Computing Systems Using Sim Grid. In *Proceedings of the Workshop on Large-Scale System and Application Performance (LSAP)*, 2010.

[32] Pedro Velho and Arnaud Legrand. Accuracy Study and Improvement of Network Simulation in the Simgrid Framework. In *Proceedings of the 2nd International Conference on Simulation Tools and Techniques for Communications, Networks and Systems, (SimuTools 2009)*, Rome, Italy, March 2009.

[33] Wim Depoorter, Nils Moor, Kurt Vanmechelen, and Jan Broeckhove. Scalability of Grid Simulators : An Evaluation. In *Proc. of the 14th Intl. Euro-Par Conf. on Parallel Processing*, 2008.

[34] Martin Barisits and Will Boyd. MartinWilSim Grid Simulator. Summer internship, Vienna UT and Georgia Tech, CERN, Switzerland, 2009. Available at `http://www.slideshare.net/wbinventor/slides-1884876`.

	CPU	Disk	Network	Application	Requirement	Scale
ns-2	-	-	fine d.e.	coarse d.e.	C++/tcl	<1,000
SSFNet	-	-	fine d.e.	coarse d.e.	Java	<100,000
GTNetS	-	-	fine d.e.	coarse d.e.	C++	<177,000
PlanetSim	-	-	cste time	coarse d.e.	Java	100,000
PeerSim	-	-	-	state machine	Java	1,000,000
OptorSim	coarse d.e.	amount	coarse d.e.	coarse d.e.	Java	few 100
GridSim	coarse d.e.	math	coarse d.e.	coarse d.e.	Java	few 1,000
CloudSim	coarse d.e.	coarse d.e.	coarse d.e.	coarse d.e.	Java	few 1,000
GroudSim	coarse d.e.	math	coarse d.e.	coarse d.e.	Java	few 1,000
SimGrid	math/d.e.	(some day)	math/d.e.	d.e./emul	C or Java	few 100,000

FIGURE 2 – Résumé des simulateurs de systèmes informatiques existants.

sonnes, des organes, des groupes de personnes, des plantes, des animaux, etc. Les plans d'expériences sont donc mis en œuvre dans de nombreux domaines en sciences humaines comme en sciences naturelles. Ils s'appliquent également naturellement en informatique, où l'expérience joue un rôle important. Il n'est malheureusement pas toujours possible de transposer dans un autre contexte les méthodes d'une discipline. Tester un médicament sur des individus pose des problèmes éthiques forçant à limiter au maximum les expériences tandis qu'il est assez aisé de tester un programme informatique. De même, les grandes échelles de temps des expériences en biologie imposent des protocoles expérimentaux très particuliers qui ne se justifient pas non plus dans notre contexte. De plus, quand les expériences sont menées sur simulateur, elles sont parfaitement reproductibles, au contraire d'expériences directes où il est impossible de faire abstraction du bruit. Cela rend possible d'autres techniques spécifiques.

Cette tâche vise à la mise en œuvre des techniques de mise au point de plan d'expériences adaptés à l'étude des systèmes informatiques. Nous allons mener notre étude en utilisant SimGrid comme outil expérimental, mais l'objectif à à plus long terme est de converger vers des outils et méthodes utilisables à la fois en simulation et en expérimentation directe. En effet, les plans expérimentaux sont naturellement également nécessaires dans ce contexte, comme en atteste la figure 1 page 16. La simplicité d'usage du simulateur nous permettra de faire abstraction des difficultés techniques afin de résoudre les problèmes théoriques qui se posent à nous. Combinées aux travaux de la première opération pour simplifier l'automatisation des expérimentations directes, ces avancées devraient être réutilisable sur les plates-formes expérimentales.

Tâche 2 : Émulation par interposition du simulateur. L'un des défauts majeurs de la simulation est qu'elle force le plus souvent à écrire un prototype de l'application étudiée dans un formalisme particulier. SimGrid ne déroge malheureusement pas à la règle, et il est actuellement impossible d'étudier une application existante par ce biais sans la réécrire au moins partiellement pour utiliser les interfaces du simulateur.

Le projet Simterpose vise à résoudre ce problème en permettant l'usage de SimGrid comme un émulateur. Ainsi, des applications réelles pourront être exécutées sur une plateforme virtuelle émulée par SimGrid : l'application s'exécutera, mais Simterpose interceptera toutes les communications et les calculs seront interceptés pour médiation, et retardés en fonction des paramètres simulés par SimGrid.

Une preuve de faisabilité était donnée par le projet MicroGrid[35], malheureusement abandonné depuis, ainsi que par un stage étudiant cet été dans notre équipe, mais de nombreux verrous techniques et scientifiques restent à lever pour rendre cette approche facilement utilisable.

Tâche 3 : Validation expérimentale des modèles du simulateur. Les modèles actuellement proposés par SimGrid offrent un niveau de réalisme acceptable (moins de 10% d'erreur) pour les réseaux de type métropolitain et pour des applications échangeant des messages de taille supérieure à quelques dizaines de kO. Cette validation a été faite par comparaison expérimentale au simulateur GTNetS, issu de la communauté réseau. Pour aller plus loin et proposer des modèles valides pour les réseaux rapides de type cluster et les petits échanges de messages, nous allons devoir mener notre étude par expérimentation directe. De part le contrôle expérimental offert, Grid'5000 constitue certainement l'outil méthodologique dont nous avons besoin pour cela.

Un autre axe d'amélioration de SimGrid porte sur les réseaux rapides de type infiniband ou myrinet, qui ne sont actuellement pas modélisés dans notre outil. Grid'5000 proposant de tels réseaux, nous comptons étudier dans un premier temps les limites des modèles analytiques existants dans la littérature. Nous intégrerons ensuite les meilleurs d'entre eux à SimGrid, possiblement en les raffinant pour les adapter à notre cas.

3.2 Axe 2 – Usage maîtrisé des infrastructures de calcul scientifique

Cet axe de travail vise à accompagner les scientifiques de la région Lorraine (tant académiques que privés) souhaitant utiliser les outils de calcul intensif dans le cadre de leurs recherches pour leur permettre de s'approprier les avancées mises en place dans le premier axe. Cela nous semble constituer une situation gagnant-gagnant, puisque cela donne aux spécialistes la garantie que les solutions qu'ils développent répondent de façon adéquate aux problèmes rencontrés lors de l'utilisation de ces infrastructures. Les utilisateurs bénéficieront quant à eux d'une expertise précieuse pour s'approprier l'usage de ces systèmes. Il est escompté que cela leur confère un avantage stratégique important pour leurs propres recherches, par une approche de science computationnelle tirant parti de la puissance des ordinateurs comme une troisième méthode scientifique aux côtés de la théorie et de l'expérimentation.

Plus généralement, nous travaillons à constituer et renforcer une communauté du calcul scientifique en Lorraine, dans l'objectif d'établir des synergies directes entre utilisateurs avancés ne nécessitant pas forcément de médiation des spécialistes du domaine.

Cet axe se décompose actuellement en trois opérations spécifiques. Les deux première portent respectivement sur l'application du calcul scientifique en cryptographie et en preuve automatique. La troisième opération est une sorte d'incubateur regroupant tous les utilisateurs des plates-formes expérimentales en Lorraine. L'un des objectifs attendus de la période 2011-2013 est d'étoffer encore cet axe en finalisant les collaborations initiées avec d'autres utilisateurs potentiels du calcul scientifique.

---

[35] H. Xia, H. Dail, H. Casanova, and A. Chien. The MicroGrid : Using Emulation to Predict Application Performance in Diverse Grid Network Environments. In *Workshop on Challenges of Large Applications in Distributed Environments*, Honolulu, June 2004.

3.2.1 Opération 2.1 – Factorisation et logarithmes discrets, applications en cryptanalyse

Participants

- Équipe AlGorille/LORIA : Lucas Nussbaum (MCF).
- Équipe Caramel/LORIA : Jérémie Detrey (CR CNRS), Pierrick Gaudry (CR INRIA), Emmanuel Thomé (CR INRIA, porteur de l'opération), Paul Zimmermann (DR INRIA).

Contexte et motivation

Cette opération, portée par le projet CARAMEL s'inscrit dans le cadre scientifique de la cryptanalyse des systèmes de chiffrements s'appuyant sur le problème de la factorisation (cryptosystème RSA), et du logarithme discret dans les corps finis. Ces travaux sont notamment menés en coopération avec le laboratoire LACAL de l'ÉPFL (Lausanne, Suisse).

Le cryptosystème RSA est aujourd'hui omniprésent (commerce électronique, chiffrement de courriers électroniques, cartes de crédit). La difficulté pour un attaquant de mettre en défaut la sécurité de ce système repose en particulier sur la difficulté du problème de la *factorisation d'entiers*. Il convient, afin de se prémunir contre des possibles attaques, de choisir la taille de clé de manière appropriée. Un tel choix se nourrit d'une évaluation précise de l'état de l'art en matière de cryptanalyse. Le projet CARAMEL et ses partenaires contribuent à cette évaluation en montrant comment un outil comme les grilles de calcul peut être mis à profit pour factoriser des nombres de taille record. L'intérêt est porté tout particulièrement à l'apport des grilles de calcul dans cette approche.

Positionnement scientifique

Les travaux de cryptanalyse sur le problème de la factorisation et du logarithme discret ont, de longue date, mobilisé des ressources de calcul importantes. L'algorithme phare utilisé dans ce contexte est l'algorithme du crible algébrique[36]. La structure générale de cet algorithme est semblable à celle de l'algorithme du crible quadratique qui lui a précédé[37]. Deux phases essentielles se distinguent dans ces algorithmes : une phase de crible, et une phase d'algèbre linéaire.

La phase de crible est très tôt apparue comme étant aisément distribuable. Des travaux anciens ont démontré par exemple comment cette phase pouvait se faire « par courrier électronique » [38]. Remis au goût du jour, ce mode de calcul se rapprocherait de l'utilisation d'une plate-forme comme BOINC (`http://boinc.berkeley.edu/`), à ceci près que les données produites par les calculs sont de volume significatif, et qu'il est nécessaire d'employer des ressources de calcul disposant de suffisamment de mémoire. Peu ou pas de calculs récents ont fait appel à un *grand* nombre de contributeurs distincts (dépassant 100

[36] A. K. Lenstra, H. W. Lenstra, Jr., M. S. Manasse, and J. M. Pollard. The number field sieve. In A. K. Lenstra and H. W. Lenstra, Jr., editors, *The development of the number field sieve*, volume 1554 of *Lecture Notes in Math.*, pp 11–42. Springer–Verlag, 1993.

[37] C. Pomerance. The quadratic sieve algorithm. In T. Beth, N. Cot, and I. Ingemarsson, editors, *Advances in Cryptology – EUROCRYPT '84*, volume 209 of *Lecture Notes in Comput. Sci.*, pp 169–182. Springer–Verlag, 1985. Proc. Eurocrypt '84, Paris (France), April 9–11, 1984.

[38] A. K. Lenstra and M. S. Manasse. Factoring by electronic mail. In J.-J. Quisquater and J. Vandewalle, editors, *Advances in Cryptology – EUROCRYPT '89*, volume 434 of *Lecture Notes in Comput. Sci.*, pp 355–371. Springer–Verlag, 1990. Proc. Eurocrypt '89, Houthalen, April 10–13, 1989.

par exemple), notamment en raison des obstacles précités. L'approche privilégiée par les travaux récents est la coordination d'efforts d'un petit nombre de contributeurs, chacun à même de mobiliser des ressources de calcul importantes. Cette approche est notamment celle qui fut utilisée pour les derniers records de factorisation[39,4]. L'organisation des calculs pour le record RSA768[4] a introduit une étape d'« industrialisation » du processus, en automatisation une très grande majorité des étapes du calcul. Cette piste vise à être poursuivie dans le contexte du projet EDGE.

La phase d'algèbre linéaire du crible algébrique est par essence plus difficile à paralléliser. Les calculs d'algèbre linéaire n'ont rien d'exceptionnel dans le contexte du calcul à haute performance. Les matrices considérées ici sont *creuses*, ce qui est un cas fréquent. Toutefois les matrices considérées dans le cadre de l'algorithme du crible algébrique, que ce soit pour la factorisation ou pour le logarithme discret, sont très différentes des matrices rencontrées couramment dans les problèmes de nature « numérique ». Le matrices que nous rencontrons sont définies sur un corps fini, ce qui invalide immédiatement toute considération de convergence, d'algorithme de point fixe, ou encore des notions comme celle de valeur propre dominante.

Le travail logiciel et algorithmique réalisé dans le contexte de la résolution de systèmes linéaires numériques ne peut donc être que très partiellement réutilisé pour les matrices rencontrées dans le contexte de la cryptanalyse. Les algorithmes pertinents pour la résolution de systèmes linéaires sur des corps finis, particulièrement sur le corps GF(2), sont décrits dans la littérature, et dans les cas les plus importants il s'agit d'algorithmes spécifiques au cas des corps finis[40,41,42,43]. Ceci étant, les besoins matériels restent semblables. En particulier, la sensibilité de tels calculs à la disponibilité de réseaux rapides est importante. Les travaux passés sur cette étape d'algèbre linéaire ont longtemps requis l'emploi de « supercalculateurs » [44,45]. Lors de la décennie passée, ces étapes plus diffici-

[39] K. Aoki, J. Franke, T. Kleinjung, A. K. Lenstra, and D. A. Osvik. A kilobit special number field sieve factorization. In *Advances in Cryptology – ASIACRYPT 2007*, volume 4833 of *Lecture Notes in Comput. Sci.*, pp 1–12. Springer–Verlag, 2008. Proc. 13th International Conference on the Theory and Application of Cryptology and Information Security, Kuching, Malaysia, December 2-6, 2007.

[40] D. H. Wiedemann. Solving sparse linear equations over finite fields. *IEEE Trans. Inform. Theory*, IT–32(1) :54–62, January 1986.

[41] B. A. LaMacchia and A. M. Odlyzko. Solving large sparse linear systems over finite fields. In A. J. Menezes and S. A. Vanstone, editors, *Advances in Cryptology – CRYPTO '90*, volume 537 of *Lecture Notes in Comput. Sci.*, pp 109–133. Springer–Verlag, 1990. Proc. 10th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 11–15, 1990.

[42] P. L. Montgomery. A block Lanczos algorithm for finding dependencies over GF(2). In L. C. Guillou and J.-J. Quisquater, editors, *Advances in Cryptology – EUROCRYPT '95*, volume 921 of *Lecture Notes in Comput. Sci.*, pp 106–120, 1995. Proc. International Conference on the Theory and Application of Cryptographic Techniques, Saint-Malo, France, May 1995.

[43] D. Coppersmith. Solving linear equations over GF(2) via block Wiedemann algorithm. *Math. Comp.*, 62(205) :333–350, January 1994.

[44] D. M. Gordon and K. S. McCurley. Massively parallel computation of discrete logarithms. In E. F. Brickell, editor, *Advances in Cryptology – CRYPTO '92*, volume 740 of *Lecture Notes in Comput. Sci.*, pp 312–323. Springer–Verlag, 1993. Proc. 12th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 16–20, 1992.

[45] S. Cavallar, B. Dodson, A. K. Lenstra, W. Lioen, P. L. Montgomery, B. Murphy, H. J. J. te Riele, K. Aardal, J. Gilchrist, G. Guillerm, P. Leyland, J. Marchand, F. Morain, A. Muffett, C. Putnam, C. Putnam, and P. Zimmermann. Factorization of a 512-bit RSA modulus. In B. Preneel, editor, *Advances in Cryptology – EUROCRYPT 2000*, volume 1807 of *Lecture Notes in Comput. Sci.*, pp 1–18. Springer–Verlag, 2000. Proc. International Conference on the Theory and Application of

lement distribuables ont été progressivement adaptées pour fonctionner sur des grappes de machines, par exemple lors de la factorisation de RSA-200. L'originalité apportée par le calcul RSA-768 a été d'utiliser l'algorithme dit de Wiedemann par blocs[43,46], permettant de séparer le calcul en un petit nombre de sous-calculs indépendants. Ceci a permis d'utiliser la grille de calcul Grid'5000 pour une grande partie du calcul. L'un des objectifs du projet EDGE dans ce domaine est l'automatisation de cette mécanique.

L'instrument logiciel des recherches du projet Caramel sur le problème de la factorisation est le Cado-nfs (`http://cado-nfs.gforge.inria.fr/`), développé en collaboration avec plusieurs partenaires depuis 2007. Ce logiciel est une implantation complète et à jour du crible algébrique, qui incorpore l'ensemble de l'état de l'art.

Programme de travail

La ressource de calcul Grid'5000 a permis au projet Caramel de mener des recherches internationales de premier plan en 2009-2010. Plusieurs pistes de poursuite de ces travaux conduiront à nouveau à exploiter la ressource de calcul Grid'5000 dans les travaux du projet Caramel.

Tout d'abord, la poursuite de la collaboration entre le projet Caramel et l'Épfl amènera naturellement de nouvelles expériences semblables au calcul RSA-768. Les travaux de développement réalisés pour le calcul RSA-768 pourront être réinvestis, et de nouvelles améliorations sont attendues, notamment dans le but de rendre les solutions déployées aussi génériques que possible. C'est notamment dans cette perspective que seront explorées les continuations naturelles des efforts évoqués plus haut pour « automatiser » autant que possible les calculs.

Le développement du logiciel Cado-NFS par le projet Caramel est aussi un cadre naturel d'association avec les ressources de calcul fournies par la plate-forme Grid'5000. La validité des algorithmes déployés dans Cado-NFS, leur compétitivité, ne peuvent être illustrées qu'au moyen d'expériences de taille significative.

Enfin, un problème voisin du problème de la factorisation d'entiers est le problème du logarithme discret dans les corps finis. Ce problème est important en cryptologie notamment par le lien qu'il entretient avec les algorithmes de couplages, utiles dans certains protocoles. Attaquer le problème du logarithme discret dans les corps finis peut se faire par une variante de l'algorithme du crible algébrique. Toutefois, il existe une différence importante pour l'étape d'algèbre linéaire intervenant dans le calcul. Le ratio entre calculs et communication a vocation à être un peu modifié par rapport aux calculs de factorisation, les communications devenant moins importantes dans le cas du logarithme discret. Cette situation laisse envisager des travaux d'adaptation qui s'inscrivent parfaitement dans les thématiques de recherche du projet Caramel, et qui seront rendues possibles par la disponibilité d'un outil expérimental fort.

Cryptographic Techniques, Brugge, Belgium, May 2000.

[46] E. Thomé. Subquadratic computation of vector generating polynomials and improvement of the block Wiedemann algorithm. *J. Symbolic Comput.*, 33(5) :757–775, July 2002.

3.2.2 Opération 2.2 – Prouver de grandes formules avec des symboles interprétés

Participants :

- Équipe AlGorille/LORIA : Sébastien Badia (ingénieur), Lucas Nussbaum (MCF).
- Équipe VeriDis/LORIA : Diego Caminha Barbosa de Oliveira (Doctorant), Pascal Fontaine (MCF, porteur de l'opération).

Contexte et motivation

Cette opération, portée par le projet VERIDIS, se concentre sur la vérification d'algorithmes distribués, en utilisant des techniques basées sur la preuve de formules logiques. Il est en effet possible de traduire les propriétés de sûreté et de vivacité du logiciel en formules logiques (voir par exemple [47,48,49,50,51]). Ces formules peuvent être, pour un algorithme même de taille moyenne, en grand nombre et très longues. De plus, les algorithmes utilisant généralement des structures de données variées, des opérateurs sur ces structures se retrouvent aussi dans les formules générées au cours de la vérification. Les solveurs SMT (Satisfiability Modulo Theories, voir[52] pour une référence assez complète sur le domaine) décident de la satisfaisabilité de formules de logiques qui contiennent des symboles interprétés (par exemple de l'arithmétique, des tableaux, des listes, des pointeurs, des symboles non interprétés...). Par rapport aux solveurs SAT (non exhaustivement, MiniSAT[53], zChaff[54], picoSAT[55],...) maintenant utilisés pour de nombreuses applications industrielles, les solveurs SMT (par exemple CVC3[56], MathSAT[57], Z3[58]) proposent une expressivité beaucoup plus grande, tout en maintenant l'efficacité des solveurs

[47] Jean-Raymond Abrial. *The B-Book : Assigning Programs to Meanings*. Cambridge University Press, 1996.

[48] Jean-Christophe Filliâtre. Why : a multi-language multi-prover verification tool. Research Report 1366, LRI, Université Paris Sud, March 2003.

[49] Robert DeLine and K. Rustan M. Leino. BoogiePL : A typed procedural language for checking object-oriented programs, March 2005.

[50] Leslie Lamport. Specifying Concurrent Systems with TLA, March 1999.

[51] K. Rustan M. Leino and Rosemary Monahan. Automatic verification of textbook programs that use comprehensions. In *Formal Techniques for Java-like Programs (FTfJP)*, 2007.

[52] Clark Barrett, Roberto Sebastiani, Sanjit A. Seshia, and Cesare Tinelli. *Satisfiability Modulo Theories*, volume 185 of *Frontiers in Artificial Intelligence and Applications*, chapter 26, pp 825–885. IOS Press, February 2009.

[53] Niklas Eén, Alan Mishchenko, and Niklas Sörensson. Applying Logic Synthesis for Speeding Up SAT. In João Marques-Silva and Karem A. Sakallah, editors, *SAT*, volume 4501 of *Lecture Notes in Computer Science*, pp 272–286. Springer, 2007.

[54] Matthew W. Moskewicz, Conor F. Madigan, Ying Zhao, Lintao Zhang, and Sharad Malik. Chaff : Engineering an Efficient SAT Solver. In *Design Automation Conference (DAC)*, pp 530–535. ACM press, June 2001.

[55] Armin Biere. PicoSAT Essentials. *JSAT*, 4(2-4) :75–97, 2008.

[56] Clark Barrett and Cesare Tinelli. CVC3. In W. Damm and H. Hermanns, editors, *Computer Aided Verification (CAV)*, volume 4590 of *Lecture Notes in Computer Science*, pp 298–302. Springer, 2007.

[57] Marco Bozzano, Roberto Bruttomesso, Alessandro Cimatti, Tommi Junttila, Peter Rossum, Stephan Schulz, and Roberto Sebastiani. MathSAT : Tight Integration of SAT and Mathematical Decision Procedures. *Journal of Automated Reasoning*, 35(1-3) :265–293, 2005.

[58] Leonardo Mendonça de Moura and Nikolaj Bjørner. Z3 : An Efficient SMT Solver. In C. R. Ramakrishnan and Jakob Rehof, editors, *Tools and Algorithms for Construction and Analysis of Systems (TACAS)*, volume 4963 of *Lecture Notes in Computer Science*, pp 337–340. Springer, 2008.

SAT. Comme les solveurs SAT, les solveurs SMT ont un domaine potentiel d'application très large, qui ne se restreint pas seulement à la vérification.

Le solveur SMT veriT (`http://www.veriT-solver.org`) est développé au sein du projet Veridis depuis 2004. L'usage de la grille de calcul est un élément fondamental du développement et de l'évaluation des techniques utilisées dans veriT.

Programme de travail

Plusieurs améliorations à la plateforme GridTPT sont prévues : nous imaginons rendre possible le profiling, et des tests de couverture de code. Des techniques de recherche de minimum, appliquées aux temps d'exécution sur les nombreux cas de tests, permettront d'affiner la valeur des nombreux paramètres présents dans le solveur. Il se peut que les comportements observés lors de ces recherches de minimum suggèrent de nouvelles méthodes, permettant d'augmenter significativement l'efficacité de l'outil. Toujours dans GridTPT, le scheduler doit être remanié, et il faut rendre la plateforme tolérante aux erreurs. Plusieurs améliorations pour la présentation des résultats sont en cours. Nous espérons ainsi que GridTPT sera utilisé par d'autres développeurs de prouveurs. C'est le seul logiciel en son genre.

Les travaux sur veriT sont sur plusieurs axes : extension de l'expressivité, amélioration des performances, enrichissement de l'interface. Pour l'expressivité, les améliorations prévues sont une amélioration de la gestion des quantificateurs, et une extension de la couverture arithmétique de l'outil. Plusieurs techniques en cours d'élaboration permettront d'augmenter l'efficacité de veriT, sur les quantificateurs et l'arithmétique, mais aussi sur les formules sans quantificateurs. Pour ces points, le Grid'5000 sera, comme par le passé, un élément essentiel pour obtenir un retour rapide et exhaustif sur l'ensemble de la librairie de problèmes collectés à ce jour.

Plusieurs autres tâches liées à veriT, prévues ou en cours, mais plus éloignées de la plateforme Grid'5000, — comme l'amélioration de l'interface, la production de preuve, la compression de preuve, la production de noyau insatisfaisable — permettront d'accroître la visibilité de veriT dans le milieu de la recherche internationale. Des partenariats industriels se forment ; leur nombre devrait augmenter, et ces partenariats devraient s'inscrire dans la durée.

3.2.3 Opération 2.3 – Communauté de calcul scientifique en Lorraine

Partenaires

Nous listons ici les équipes comptant des utilisateurs réguliers ou occasionnels des ressources d'expérimentation mises à disposition par EDGE, mais avec qui les collaborations n'ont pas encore abouti à la mise en place d'une opération spécifique au sein du projet EDGE.

- Équipes du LORIA : Madynes, Cassis, Talaris, Calvi, Score.
- Équipes lorraines : Supélec (Metz), LITA (Metz)
- Équipes du Grand Est au delà de la Lorraine : Université du Luxembourg, CReSTIC (Reims), LSIIT (Strasbourg), OAS (Strasbourg), LIFC (Besançon, Belfort et Montbéliard), LMIA (Mulhouse).

Contexte et motivation

L'objectif de cette opération est de constituer une communauté de calcul scientifique en Lorraine. Nous pensons axer l'animation de cette communauté autour des applications en informatique, mais notre objectif à terme est de l'ouvrir plus largement à d'autres disciplines scientifiques. La plate-forme expérimentale mise à disposition par EDGE et Grid'5000 sera évidement amenée à jouer un rôle fédérateur primordial dans ce processus, mais l'objectif reste d'établir des interactions de recherche entre les membres.

Programme de travail

Il est très difficile d'établir le programme de travail de recherches aussi disparates que celles envisagées par les utilisateurs au sein de cette opération. Nous préférons lister ici les actions transverses pour l'élargissement de notre communauté.

Tâche 1 : Formations à destination des utilisateurs potentiels. Malgré nos efforts, Grid'5000 reste un outil difficile à maîtriser techniquement. Dans l'opération 1.1, nous travaillons à simplifier cet usage en améliorant les outils de base disponibles pour les expérimentateurs (cf. page 18). Mais nous pensons que ces efforts ne sont pas suffisants et nous visons à organiser des formations régulières à destination des utilisateurs potentiels afin de réduire leur temps de prise en main et ainsi améliorer leur efficacité. Ces formations poursuivent des objectifs similaires à celles organisées pour aider les scientifiques souhaitant utiliser les ressources de calcul des grilles de productions, même si les outils à maîtriser restent légèrement différents à l'heure actuelle. Les formations porteront ainsi sur la réservation et le déploiement de ressources sur Grid'5000 avec les logiciels OAR et Kadeploy, l'utilisation de MPI, et l'utilisation de fonctionalités avancées de la plate-forme Grid'5000 pour faciliter l'automatisation des expériences.

Des supports de formations à notre instrument existent déjà et sont utilisées lors des écoles de printemps Grid'5000 annuelles. Notre objectif ici est d'en rafraîchir les supports si nécessaire, et surtout de donner ces formations à un public aussi large que possible lors de sessions régulières.

Afin d'accroître l'impact de ces formations, nous pensons donner des séminaires plus courts dans les réseaux scientifiques régionaux existants. Ainsi, le Réseau Grand Est regroupe des scientifiques autour des thématiques du calcul distribué dans l'est de la France. Il s'agit d'une action géographique du groupe de recherche ASR du CNRS jouant un rôle fédérateur important pour les acteurs concernés. Ce réseau constitue un public de choix pour des séminaires d'initiation à l'usage des plates-formes expérimentales, même si d'autres actions à destination d'un public plus large sont à prévoir.

Nous envisageons également de nous rapprocher du réseau de formation de France Grille qui organise régulièrement des journées de formation en région afin de démocratiser les grilles de production. Notre objectif serait de faire découvrir les grilles expérimentales à ce public et ainsi renforcer les liens entre les deux communautés. Si elles ne sont pas adaptées à du calcul intensif à but de production, les ressources expérimentales de EDGE peuvent s'avérer précieuses pour les scientifiques d'autres disciplines cherchant à optimiser leurs applications pour les rendre plus efficaces et moins gourmandes en ressources avant leur mise en production sur les grilles de production.

4 Moyens utilisés et organisation du projet

4.1 Personnes

Le projet EDGE regroupe de nombreux chercheurs et enseignants-chercheurs du LORIA à Nancy (de par la proximité des ressources expérimentales), mais également de l'IECN à Nancy et d'autres équipes de la région comme Supélec et le LITA à Metz, voire au delà de la Lorraine.

4.1.1 Participants actifs du projet EDGE

- **Opération 1.1 *Plates-formes expérimentales :*** Sébastien Badia (ingénieur AlGorille/LORIA), François Charoy (MCF Score/LORIA), Sylvain Contassot (PR AlGorille/LORIA), Jens Gustedt (DR INRIA AlGorille/LORIA), Lucas Nussbaum (MCF AlGorille/LORIA), Martin Quinson (MCF AlGorille/LORIA), Tina Rakotoarivelo (ingénieur AlGorille/LORIA).
- **Opération 1.2 *Simulation de systèmes informatiques :*** Pierre-Nicolas Clauss (post-doctorant AlGorille/LORIA), El Mehdi Fekari (ingénieur AlGorille/LORIA), Lucas Nussbaum (MCF AlGorille/LORIA), Martin Quinson (MCF AlGorille/LORIA), Cristian Rosa (doctorant AlGorille/LORIA), Christophe Thiéry (ingénieur AlGorille/LORIA).
- **Opération 2.1 *Applications à la cryptanalyse :*** Jérémie Detrey (CR INRIA – Caramel/LORIA), Pierrick Gaudry (DR CNRS – Caramel/LORIA), Emmanuel Thomé (CR INRIA – Caramel/LORIA – porteur de l'opération), Lucas Nussbaum (MCF AlGorille/LORIA), Paul Zimmermann (DR INRIA – Caramel/LORIA).
- **Opération 2.2 *Applications à la preuve automatique*** Sébastien Badia (ingénieur AlGorille/LORIA), Diego Caminha Barbosa de Oliveira (Doctorant VeriDis/LORIA), Pascal Fontaine (MCF VeriDis/LORIA), Lucas Nussbaum (MCF AlGorille/LORIA).

4.1.2 Utilisateurs des ressources expérimentales

D'autres chercheurs et enseignants-chercheurs lorrains sont associés au projet de façon plus informelle, par exemple par leur usage de la plate-forme Grid'5000. Il s'agit en particulier des équipes listées ci-après. L'un des objectifs de la période 2011-2013 est de les associer plus formellement au projet EDGE du CPER MISN afin de s'assurer que les outils mis en place dans ce cadre correspondent au mieux à leurs besoins scientifiques.

- L'équipe Madynes (LORIA) s'intéresse à la gestion des réseaux informatiques. Dans ce cadre, ses membres sont souvent amenés à expérimenter des systèmes informatiques distribués, par exemple pour en évaluer la sécurité ou la robustesse. Ces travaux s'insèrent naturellement dans le projet EDGE.
- Les travaux de recherche de l'équipe Score (LORIA) portent sur la science du web et des services. L'utilisation de la plate-forme Grid'5000 permet à l'équipe d'expérimenter et de valider *in-situ* les nouvelles architectures proposées pour permettre aux services collaboratifs distribuées de demain de supporter la charge générée par un très grand nombre d'utilisateurs. Ces travaux s'inscrivent logiquement dans le projet EDGE.

- L'équipe Carte (LORIA) étudie les systèmes adverses, comme les virus informatiques. Pour cela, les membres de cette équipe travaillent à la construction d'un laboratoire de haute sécurité dans le cadre du projet SSS du CPER MISN. Les virus constituent des systèmes informatiques très spécifiques, et leur étude implique des contraintes supplémentaires pour l'expérimentation. Dans le cadre d'une opération commune aux deux projets SSS et EDGE, nous avons engagé une réflexion pour la définition d'outils logiciels permettant de mener des expérimentations avec les virus informatiques. L'amont de cette réflexion trouve une place naturelle dans le projet EDGE tandis que la mise en production et l'étude proprement dite des virus trouvent leur place dans le projet SSS.
- Les recherches en bio-informatique de l'équipe Orpailleur (LORIA) nécessitent également de grosses infrastructures de calcul. Ces recherches ont naturellement lieu dans le projet MBI du CPER, mais la mise au point et l'amélioration des logiciels utilisés pour cela s'insère dans le projet EDGE, comme en atteste leur usage occasionnel des ressources expérimentales du projet.
- Les recherches de l'équipe IMS de Supélec couvrent un large spectre, depuis le traitement du signal jusqu'à l'intelligence artificielle en passant par les infrastructures distribuées de calcul intensif. Cet environnement fécond a donné naissance au projet Intercell, financé dans la première moitié du CPER MISN par le projet MIS. Comme dans le cadre des actions communes avec les projets SSS et MBI, les travaux amonts visant à établir cette plate-forme distribuée trouve une place naturelle dans EDGE tandis que la mise en production aval est menée dans un autre cadre par les scientifiques concernés.
- Claus Fieker s'installe en 2011 comme professeur à l'université de Kaiserslautern. Ses recherches portent sur la théorie algorithmique des nombres, et sont fortement liées aux travaux de l'équipe Caramel sur le crible algébrique. Les calculs menés par C. Fieker ont une structure semblable à celle du crible algébrique utilisé dans les travaux de cryptanalyse. Au même titre, l'emploi de ressources de calcul dans ces travaux est pertinent.

De nombreuses autres équipes recherche de Lorraine et du grand est utilisent les ressources expérimentales mises à disposition au sein du projet. Voici une liste non-exhaustive : LITA (Metz), CReSTIC (Reims), LSIIT (Strasbourg), OAS (Strasbourg), LIFC (Besançon, Belfort et Montbéliard), LMIA (Mulhouse).

4.2 Équipements et plates-formes

La majeure partie des demandes budgétaires dans le cadre du CPER porte sur l'instrument scientifique Grid'5000, puisque les autres besoins des recherches envisagées (en particulier dans l'opération *simulation de systèmes informatiques*) sont financés par ailleurs. En revanche, il est de la plus grande importance d'assurer la continuité de service pour faire vivre le site lorrain de Grid'5000, en conservant des machines fonctionnelles à tout moment ainsi que du personnel chargé de leur bon fonctionnement et utilisation.

Nous rappelons ici les ressources expérimentales mises à disposition afin d'établir l'instrument lorrain d'expérimentation sur les applications distribuées à large échelle.

Équipements composant le nœud lorrain de Grid'5000

- **cluster Griffon** : Installation : 02/2009
  92 nœuds (2 processeurs quadri-cœurs Intel Xeon L5420, 16 Go de RAM par nœud).

- **cluster Graphene** : Installation : 11/2010
  144 nœuds (1 processeur quadri-cœur Intel Xeon X3440, 16 Go de RAM par nœud).

Moyens humains
Il est impensable qu'un instrument scientifique de cette dimension puisse fonctionner sans des moyens humains suffisants. Depuis la création du site lorrain de Grid'5000, nous avons veillé à disposer des ressources humaines nécessaires.

La maintenance du site a été assurée par des ingénieurs débutants en CDD. Sur les périodes 2005-2007 et 2007-2009, ce rôle a été tenu par des ingénieurs associés financés par l'INRIA. Il est à remarquer que ce poste a permis au premier de nos ingénieurs (Xavier Delaruelle) d'obtenir depuis un poste d'ingénieur-chercheur au CEA. Depuis octobre 2009 et pour deux ans, ce poste est financé par l'INRIA sous forme d'IJD au travers de l'Action de Développement Technologique (ADT) Aladdin.

Le site lorrain de Grid'5000 est également moteur dans le développement d'outils permettant l'usage de l'instrument. Ceci est entre autres dû au fait que nous hébergeons deux ingénieurs permanents (Benjamin Dexheimer et Emmanuel Jeanvoine, respectivement ingénieur d'étude et ingénieur de recherche), qui travaillent sur le développement de l'infrastructure logicielle de Grid'5000.

5 Organisation du projet

Porteurs du projet. Martin Quinson et Lucas Nussbaum (LORIA/AlGorille). Les deux porteurs travaillent conjointement et font le point régulièrement sur l'avancement du projet et sa gestion :

- Pilotage des actions de recherche transverses au projet et pilotage de l'instrument scientifique
- Liens entre le projet et le reste du CPER MISN, ainsi qu'entre les recherches menées en Lorraine et celles menées en national et à l'international
- Encadrement des ingénieurs impliqués

Porteurs d'opérations. Chaque porteur a une fonction d'animation de l'opération et veille à l'avancement des travaux afférents. Il a également une fonction de lien entre les membres de l'opération et le reste du projet. Il assure la production du bilan et du projet annuel. Tous les porteurs se rencontrent régulièrement pour discuter de l'avancement global du projet et des efforts dans chaque opération et pour détecter d'éventuelles possibilités de collaboration.

- *Plates-formes expérimentales :* Lucas Nussbaum
- *Simulation de systèmes distribués :* Martin Quinson
- *Applications à la cryptanalyse :* Emmanuel Thomé
- *Applications à la preuve :* Pascal Fontaine

Assemblées générales : Elles servent à faire le point et à communiquer sur les différentes étapes du projet (demandes de subvention, bilans, appels à opération). Il est prévu d'avoir une ou deux AG par an.

6 Retombées sur la région Lorraine

6.1 Avantage stratégique pour les scientifiques locaux

Les scientifiques engagés dans ce projet disposent d'une bonne expertise dans le domaine de l'expérimentation d'applications distribuées à large échelle. Le site Grid'5000 de Lorraine jouit d'une grande visibilité, même au sein du réseau national. Il nous semble par exemple révélateur que la seconde école de printemps de l'instrument scientifique ait été organisée à Nancy en 2009. La participation de notre région à cet ensemble augmente également l'attractivité scientifique de la région, comme en atteste le recrutement récent de candidats de qualité dans nos équipes.

La présence de cet équipement et des équipes scientifiques associées constitue également un avantage stratégique d'importance pour les scientifiques n'étant pas spécialistes de ces systèmes. L'expérience de l'équipe CARAMEL est emblématique par sa participation à un record de factorisation visant à étudier la solution d'une clé de 768 bits. La présence de l'instrument scientifique en région, et surtout la proximité géographique des scientifiques spécialistes ont grandement facilité la prise en main du système, avec le succès et les retombées scientifiques et médiatiques que l'on sait.

La plupart des opérations scientifiques passées sur l'instrument constituent des *success story* comparables, où des scientifiques ont pu mettre au point par ce biais des infrastructures de calcul leur donnant un avantage stratégique important dans leur communauté propre. De tels cercles vertueux sont aussi présents au sein du CPER MISN, où des projets comme MBI et SSS profitent des expertises acquises au sein du projet EDGE pour leur permettre de maximiser le bénéfice de leurs ressources de calcul en leur permettant de faire mieux avec moins.

Étant donnée l'importance toujours croissante du calcul scientifique dans la recherche moderne, nous pensons qu'un tel avantage compétitif constitue un élément primordial pour la performance des chercheurs de notre région. Nous espérons pouvoir transmettre ainsi les connaissances en expérimentation informatisée à d'autres communautés de recherche lorraines, qu'elles soient issues de l'informatique ou d'autres disciplines telles que la physique (au travers de l'institut Jean Lamour) ou la biologie. Il est cependant important de noter que nous ne visons pas à devenir les opérateurs d'un centre de calcul traditionnel ou un centre de compétences en calcul haute performance comme l'était le centre Charles Hermite. Nous visons à étudier et améliorer sur notre infrastructure des applications scientifiques dont le devenir est d'être exécutées sur des plates-formes de productions. Il peut s'agir de ressources de calculs propres aux utilisateurs, des mésocentres de calcul fédérés par le projet EquipEx Equip@Meso porté par GENCI en France, ou de la grille de production européenne EGI par le biais de son incarnation nationale, France Grilles.

À terme, la suite logique de nos efforts est d'établir un mésocentre de calcul Lorrain hébergeant à la fois les applications de calcul scientifique nécessaires aux avancées de la physique ou de la biologie et les expérimentations spécifiques nécessaires aux avancées de l'informatique. Le défi reste de taille car les besoins opérationnels diffèrent grandement : les expérimentations informatiques nécessitent de pouvoir modifier toutes les couches de la pile logicielle mise à disposition (comme c'est le cas sur Grid'5000) tandis que les calculs scientifiques des autres disciplines nécessitent une couche logicielle mettant l'accent sur la stabilité, la compatibilité entre les solutions et des possibilités d'*accounting* permettant

un partage équitable des ressources.

Nous pensons que les expertises gagnées au travers du projet EDGE, les outils mis au point dans notre région ainsi que nos collaborations avec des acteurs majeurs des grilles de production (comme F. Suter, chargé de mission par l'institut des grilles pour les interactions entre grilles de production et grilles expérimentales, qui compte parmi nos collaborateurs proches) nous donnent les moyens de relever ce défi à moyen terme. Les avantages potentiels sont naturellement la mutualisation des moyens et l'accroissement des interactions entre les spécialistes du domaine et les utilisateurs prioritaires des techniques en découlant.

6.2 Rayonnement de la région

Le projet EDGE constitue par ailleurs un élément de rayonnement fort pour la région Lorraine. Il constitue en effet une expérience enviée par les régions du grand est de la France et de la Grande Région. Il est significatif de constater que le Luxembourg et la Champagne-Ardennes sont en train d'établir des projets comparables afin de doter leurs scientifiques de ressources expérimentales similaires au nœud Lorrain de Grid'5000. Notre expertise, mise à profit pour aider au lancement de ces opérations, assure donc un rayonnement du savoir-faire scientifique et technique de notre région.

7 Valorisation de la recherche

Les retombées socio-économiques attendues au plan local et régional sont multiples. Tout d'abord, nous travaillons à nous rapprocher des acteurs du tissu socio-économique local afin de les faire bénéficier de notre expertise et leur permettre d'utiliser les ressources de calcul pour expérimenter leurs logiciels. Nous avons établi des contacts préliminaires avec le département recherche et développement de l'entreprise Arcelor-Mittal, et nous espérons pouvoir concrétiser ce projet prochainement. Par ailleurs, la start-up Scalable Graphics a déjà utilisé nos infrastructures de manière ponctuelle afin d'étudier l'extensibilité d'un de ses produits logiciels. Ces collaborations entre la recherche publique et privée en Lorraine nous semblent très intéressantes, et nous souhaitons multiplier ces initiatives à l'avenir.

Les ressources expérimentales et l'expertise associée au projet EDGE sont également sources de collaborations avec des organismes de recherche semi-privés tels que le CEA (sur l'administration de grands centres de calcul) ou le CERN (sur la gestion des données dans les très grandes infrastructures de calcul).

De par leur nature même, les retombées des travaux de cryptanalyse ne sont quant à elles pas attendues à l'échelle d'une collaboration avec un partenaire industriel. En effet, les premiers acteurs sur le thème de l'analyse des cryptosystèmes sont gouvernementaux (ministère de la défense), tandis que les activités de préconisations de solution cryptographiques sont assurées par des acteurs comme l'ANSSI ou bien, dans des cadres spécifiques par des GIE, tels le GIE cartes bancaires. Ces derniers organismes sont des utilisateurs directs des travaux de cryptanalyse, puisque leurs préconisations sont dimensionnées directement en fonction des résultats obtenus. Aussi l'impact social et économique de ces travaux de recherche est important.

Enfin, EDGE constitue un vecteur de collaboration scientifique de la première importance pour les équipes lorraines. La proximité de la plate-forme et des compétences associées est un argument très souvent avancé lors du montage de projets en national (à

l'agence nationale de la recherche – ANR) ou au niveau de la communauté européenne. Les financements de la recherche obtenu par ces biais peuvent de plus se concentrer sur les projets propres de chaque équipe en s'appuyant sur les ressources expérimentales proposées par EDGE. Cette mutualisation des moyens matériels constitue donc une sorte de retour sur investissement pour la région en lui permettant d'augmenter les fonds nationaux et européens arrivant dans la région.

8 Formation

Les problématiques de la méthodologie expérimentale occupent une place importante dans la formation des futurs ingénieurs et scientifiques. Les outils que nous rendons disponibles constituent une base appréciée pour la pédagogie sur ces thèmes.

La plate-forme EDGE, et à travers elle la **plate-forme Grid'5000**, continuera à être utilisée dans plusieurs formations des universités lorraines. Les étudiants pourront ainsi acquérir des compétences sur le calcul à hautes performances ou l'administration d'infrastructures distribuées dans le cadre des formations suivantes :

- *Licence professionnelle Administration de systèmes, réseaux et applications à base de logiciels libres* (ASRALL) (IUT Nancy-Charlemagne, Univ. Nancy 2).
  Deux groupes de projets tutorés (8 étudiants) ont utilisé Grid'5000 en 2009/2010, et deux autres en 2010/2011. Il faut noter qu'un des étudiants ayant fait son projet tutoré en 2009/2010 a été recruté comme administrateur système sur la plate-forme EDGE.
- *Module « Algorithmique Répartie et Systèmes Distribués » du Master 1 Informatique* de l'université Henri Poincaré Nancy.
  Des travaux pratiques seront réalisés sur la plate-forme dans le cadre de ce module, afin de permettre aux étudiants de se former à la programmation OpenMP et MPI dans le contexte de clusters de taille importante.
- *Module « Algorithmique des Systèmes Parallèles et Distribués » de la deuxième année de l'École Supérieure d'Informatique et de ses Applications de Lorraine (ÉSIAL).*
  Comme pour le module de Master 1 à l'UHP, des travaux pratiques seront réalisés sur la plate-forme afin de former les élèves à la programmation OpenMP et MPI.

De son coté, le simulateur **SimGrid** est utilisé par exemple dans le cadre des formations suivantes :

- *Module « Algorithmique Distribuée » de la troisième année de l'École Nationale Supérieure d'Électricité et de Mécanique (ENSEM) de Nancy.*
- *Module « Algorithmique Distribuée Avancée et Grilles de Calcul » du master recherche d'informatique de l'université de Henri Poincaré de Nancy I*

Les algorithmes classiques de l'algorithme distribuée (tels que l'horloge de Lamport ou l'exclusion mutuelle), ne sont pas seulement présentés théoriquement, mais l'occasion est donnée aux élèves d'expérimenter en pratique ces algorithmes et leurs différentes variantes.

VII – DESCRIPTION SYNTHÉTIQUE ACTUALISÉE POUR 2011-2013

Contextes et enjeux

Les évolutions technologiques permettent de construire des plates-formes matérielles de plus en plus grandes et complexes. La puissance potentielle des systèmes ainsi constitués offre de nouvelles possibilités en termes d'applications, qu'elles soient scientifiques comme les simulations multi-physiques, grand public comme les systèmes pair-à-pair ou commerciales comme le cloud computing. En particulier, les bénéfices du calcul scientifique pour des disciplines comme la physique, la biologie ou l'ingénierie sont tels qu'il est d'usage de considérer que la simulation constitue une troisième approche scientifique, au même titre que la théorie et l'expérimentation. La maîtrise de ces technologies constitue un objectif majeur de nos sociétés de l'information.

Objectifs scientifiques

Des systèmes informatiques de ces dimensions posent cependant des problèmes méthodologiques et scientifiques spécifiques. Leurs caractéristiques complexes, hiérarchiques et dynamiques rendent leur usage très difficile. Les objectifs du projet EDGE sont les suivants :

i. Étudier ces systèmes en utilisant les différentes méthodologies scientifiques classiques (théorie, mais surtout expérimentation directe et simulation) ;

ii. Simplifier l'usage de ces systèmes par le développement d'outils au service d'une approche méthodologique unifiée ;

iii. Tirer partie de l'expertise ainsi acquise pour permettre des avancées scientifiques dans d'autres branches de l'informatique et dans d'autres disciplines scientifiques.

Principaux partenaires

Le projet « Expérimentations et calculs Distribués à Grande Échelle » (EDGE) vise à constituer en Lorraine d'une communauté regroupant les spécialistes des plates-formes distribuées à large échelle et les scientifiques utilisateurs de ces systèmes pour leurs recherches propres. Les chercheurs participants sont des membres de laboratoires nancéens (le LORIA et l'IECN), d'autres laboratoires de la région (Supélec et le LITA à Metz), voire au delà de la Lorraine (le CReSTIC à Reims, le LSIIT au Luxembourg, OAS à Strasbourg, le LIFC en Franche Comté, et le LMIA à Mulhouse).

Partenariats et soutiens européens ou internationaux

Les activités menées dans le cadre du projet EDGE sont des thématiques scientifiques mises en avant au niveau national européen et international. Les soutiens enregistrés sont de type ANR (programme Arpège : projet STREAM, projet USS-SimGrid ; programme blanc : projet ConcoRDant, projet CHIC ; programme Domaines Émergents : projet Decert), projets nationaux (programme ADT de l'INRIA : projet SimGrid Usability, projet Aladdin), projets de collaboration entre grilles expérimentales et grilles de production (programme joint Grid'5000/Institut des grilles : projet SimgLite, projet SimData) et projets bilatéraux internationaux (INRIA-CNPq : projet SMT-SAVeS avec le brésil ; PHC Germaine de Staël avec la suisse ; PHC Tournesol FL avec la belgique).

Principaux résultats visibles ou visés

Les résultats du projet feront l'objet de publications dans les revues et conférences nationales et internationales pertinentes à la fois dans les domaines de l'expérimentation sur plate-forme distribuée à large échelle et dans les domaines spécifiques des scientifiques tirant partie de ces ressources expérimentales pour leurs recherches propres.

Au niveau des outils permettant l'usage de ressources expérimentales, nous proposons une approche originale visant à constituer un **écosystème d'outils** de base, interconnectés par des **technologies ouvertes et standardisées** de type RESTful. Les avantages attendus de cette approche sont de pouvoir mutualiser et partager ces outils avec d'autres ressources expérimentales plus simplement qu'actuellement où chaque plate-forme propose ses propres outils, interconnectés par des solutions non standard.

Un autre avantage de cet usage d'une interface RESTful est que cela permet d'envisager des outils de **guidage automatisé d'expériences**. Actuellement, les expérimentateurs doivent réaliser ce processus de manière manuelle, ce qui complique encore leur tâche. De plus, cette automatisation permettra de réduire la quantité d'expériences à faire pour arriver à une conclusion donnée (et ainsi de maximiser la rentabilité des instruments scientifiques utilisés) tout en permettant le partage des scénarios expérimentaux entre scientifiques du domaine (et donc de maximiser l'impact des recherches réalisées). Une méthodologie expérimentale globale comme nous l'envisageons permettra également de réaliser une partie des expériences par simulation, afin de pré-filtrer les tests pertinents. L'objectif est de ne mobiliser des ressources de calcul lourdes que pour les expériences réellement importantes pour l'étude envisagée.

Les **applications de ces recherches** dans d'autres branches de l'informatique déboucheront à n'en pas douter sur des avancées majeures pour les domaines en question, même s'il est difficile d'établir un programme précis pour des recherches aussi vastes. Il reste cependant certain que d'autres événements marquants, comme le record mondial de factorisation du nombre RSA-768 établi en 2010 par l'équipe Caramel en crypto-analyse, feront partie des retombées du projet.

La suite logique de nos efforts serait d'établir **un mésocentre de calcul Lorrain** hébergeant à la fois les applications de calcul scientifique nécessaires aux avancées de la physique ou de la biologie et les expérimentations spécifiques nécessaires aux avancées de l'informatique. Le défi reste de taille car les besoins opérationnels diffèrent grandement : les expérimentations informatiques nécessitent de pouvoir modifier toutes les couches de la pile logicielle mise à disposition (comme c'est le cas sur Grid'5000) tandis que les calculs scientifiques des autres disciplines nécessitent une pile logicielle mettant l'accent sur la stabilité, la compatibilité entre les solutions et des possibilités d'*accounting* permettant un partage équitable des ressources. Il est cependant peu probable que nos travaux débouchent sur la création d'un tel centre dans les trois années à venir, mais ceci constitue clairement la direction vers laquelle nous tendons à plus long terme.

VIII – INDICATEURS ACTUALISÉS

1 Indicateurs de moyens

Équipes de recherche impliquées dans le projet

Trois équipes de recherche du LORIA sont fortement impliquées dans le projet (AlGorille, CARAMEL, VERIDIS), auxquelles s'ajoutent de nombreuses équipes utilisant les ressources expérimentales et l'expertise mises à disposition par le projet : au LORIA, au moins cinq équipes (Madynes, Cassis, Talaris, Calvi, Score), dans deux autres laboratoires de la région (Supélec et le LITA à Metz), ainsi que de nombreux autres partenaires au delà de la Lorraine (le CReSTIC à Reims, le LSIIT au Luxembourg, OAS à Strasbourg, le LIFC en Franche Comté, et le LMIA à Mulhouse).

Nombre de chercheurs ETP impliqués dans le projet : les équipes fortement impliquées représentent une dizaine de chercheurs et enseignants-chercheurs permanents tandis que les équipes associées représentent une trentaine de chercheurs et enseignants-chercheurs permanents.

Nombre de doctorants impliqués dans le projet : une demi-douzaine de doctorants sont impliqués dans le projet.

2 Indicateurs de production scientifique

Publications faisant apparaître des effets de synergie entre équipes

[KNT10] T. Kleinjung, L. Nussbaum, and E. Thomé. Using a grid platform for solving large sparse linear systems over GF(2). In *11th ACM/IEEE International Conference on Grid Computing (Grid 2010)*, Belgique Brussels, Oct 2010.

Autres publications résultant du projet

[AMQ10] Sabina Akhtar, Stephan Merz and Martin Quinson. *A High-Level Language for Modeling Algorithms and their Properties.* 13th Brazilian Symposium on Formal Methods, Natal, Rio Grande do Norte, Brazil, Nov 8-12, 2010.

[BOD+10a] T. Bouton, D. C. B. de Oliveira, D. Déharbe, and P. Fontaine. veriT : an open, trustable and efficient SMT-solver. In R. Schmidt, editor, *Proc. Conference on Automated Deduction (CADE)*, volume 5663 of *Lecture Notes in Computer Science*, pages 151–156, Montreal, Canada, 2009. Springer.

[BOD+10b] T. Bouton, D. C. B. de Oliveira, D. Déharbe, and P. Fontaine. GridTPT : a distributed platform for Theorem Prover Testing. In B. Konev and R. Schmidt, editors, *Workshop on Practical Aspects of Automated Reasoning (PAAR)*, Edinburgh, UK, 2010.

[BNG10a] Tomasz Buchert, Lucas Nussbaum, and Jens Gustedt. Accurate emulation of CPU performance. In *8th International Workshop on Algorithms, Models and Tools for Parallel Computing on Heterogeneous Platforms (HeteroPar'2010)*, Ischia, Italy, 2010.

[BNG10b] Tomasz Buchert, Lucas Nussbaum, and Jens Gustedt. Methods for Emulation of Multi-Core CPU Performance. Research Report RR-7450, INRIA, 11 2010.

[BSQ10] Laurent Bobelin, Martin Quinson and Frédéric Suter. *Synthesizing Generic Experimental Environments for Simulation.* 5th International Conference on P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC'10), Fukuoka, Japan, Nov 4-6 2010.

[CSG+10] Pierre-Nicolas Clauss, Mark Stillwell, Stéphane Genaud, Frédéric Suter, Henri Casanova, Martin Quinson. *Single Node On-Line Simulation of MPI Applications with SMPI.* 25th IEEE International Parallel & Distributed Processing Symposium (IPDPS'11), May 16-20, 2011, Anchorage (Alaska) USA.

[KAF+10] T. Kleinjung, K. Aoki, J. Franke, A. K. Lenstra, E. Thomé, J. W. Bos, P. Gaudry, A. Kruppa, P. L. Montgomery, D. A. Osvik, H. Te Riele, A. Timofeev, and P. Zimmermann. Factorization of a 768-bit RSA modulus. In T. Rabin, ed., *CRYPTO 2010*, vol. 6223 of *Lecture Notes in Computer Science*, pp. 333–350, États-Unis Santa Barbara, 2010. Springer Verlag. The original publication is available at www.springerlink.com.

[KBL+10] T. Kleinjung, J. Bos, A. Lenstra, D. Arne Osvik, K. Aoki, S. Contini, J. Franke, E. Thomé, P. Jermini, M. Thiémard, P. Leyland, P. Montgomery, A. Timofeev, and H. Stockinger. A heterogeneous computing environment to solve the 768-bit RSA challenge. *Cluster Computing*, 2010.

[OMM+10] G. Oster, R. Mondéjar, P. Molli and S. Dumitriu. Building a collaborative peer-to-peer wiki system on a structured overlay. *Computer Networks*, 54 :1939–1952, 2010.

[RMQ10] Cristian Rosa, Stephan Merz and Martin Quinson. *A Simple Model of Communication APIs – Application to Dynamic Partial-order Reduction.* 10th International Workshop on Automated Verification of Critical Systems (AVOCS'10), Düsseldorf, Germany, Sept 20-23, 2010.

3 Autres actions

L'un des fondements du projet EDGE est de mettre à disposition d'une communauté dédiée des ressources expérimentales permettant de mutualiser les équipements entre les partenaires. Constituer le nœud lorrain de Grid'5000 n'est certainement pas une fin en soi, mais la présence de cette équipement constitue un avantage stratégique pour les chercheurs impliqués, comme détaillé page 35 à 37.