

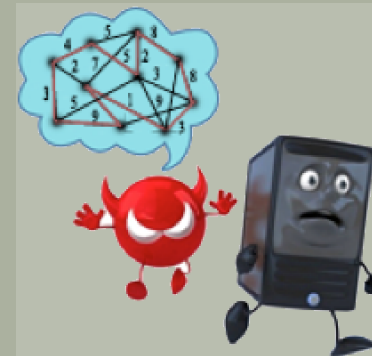
Dans la Saga du Calcul Haute-Performance



CLAUDE TADONKI

Enseignant-Chercheur en informatique

MINES ParisTech ♦ PSL Research University ♦ CRI (Centre de Recherche en Informatique)



Exposé de Pérennisation ♦ Commission de la Recherche ♦ 26 juin 2014



Mon chemin vers le CRI et mon positionnement thématique



Université de Rennes/IRISA → Université de Genève (HEC/CUI)

→ Laboratoire Européen de Biologie Moléculaire → Université Paris-Sud → MINES ParisTech - PSL

→ **Doctorat** en informatique

Université de Rennes I – mars 2001

→ **HDR** en informatique

Université Paris-Sud XI – mai 2013

Compétences dans le champ du CRI

- ✓ Analyse et transformation de codes
- ✓ Parallélisation automatique

Compétences complémentaires

- ✓ Algorithmique et complexité
- ✓ Calcul haute performance
- ✓ Recherche opérationnelle et prog math.



Responsabilités ponctuelles

- ✓ Responsable de la constitution du dossier AERES du CRI
- ✓ Coordinateur des réunions de réflexion sur le projet à cinq ans du CI

Mission

Animer la recherche en HPC au CRI ainsi que les collaborations avec les autres centres

Statistiques Générales

Rubriques	Avant 2011	2011-2014
Journaux	12	4
Conférences	35	10
Rapports de recherche	15	8
Projets (participation)	8 (3 européens)	3 (PetaQCD – BioCloud – TIMC)
Projets (soumission)	9 (4 ✓)	7 (2 ✓ et 1 en cours)
Doctorants (co-encadrements)	7	5 (1 soutenance et 2 en 09/04)
Stagiaires	15	3 (niveau master 2)
Logiciels importants	2	1
Chapitre d'ouvrage	1	1

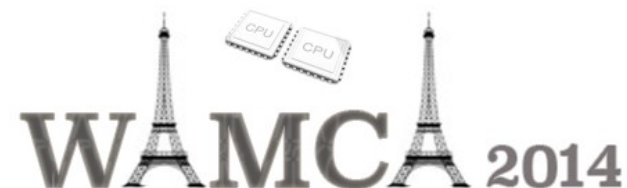
→ Rayonnement (trois dernières années)

Reviews (journal): **6**

Reviews (conférence): **20**

Jury (thèse et qualification de thèse): **3**

Comité de programme: **7**



Habilitation à Diriger les Recherches (HDR)

Fait marquant de ma période de tenure-track !!!



METHODE + PROGRAMMATION + MACHINE = CALCUL HAUTE PERFORMANCE

Habilitation à **D**iriger les **R**echerches
Soutenue à
Université Paris-Sud – Orsay
16 mai 2013

High Performance Computing as a Combination of Methods and Machines and Programming

HDR: Méthode + Programmation + Machine = HPC

MACHINE

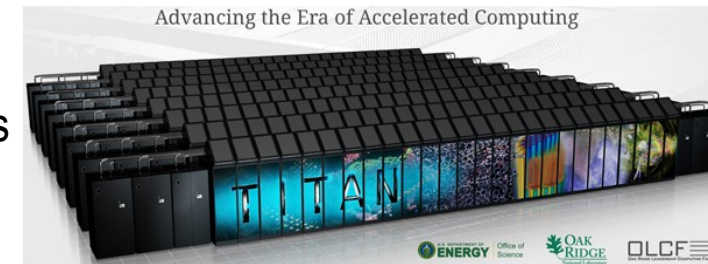
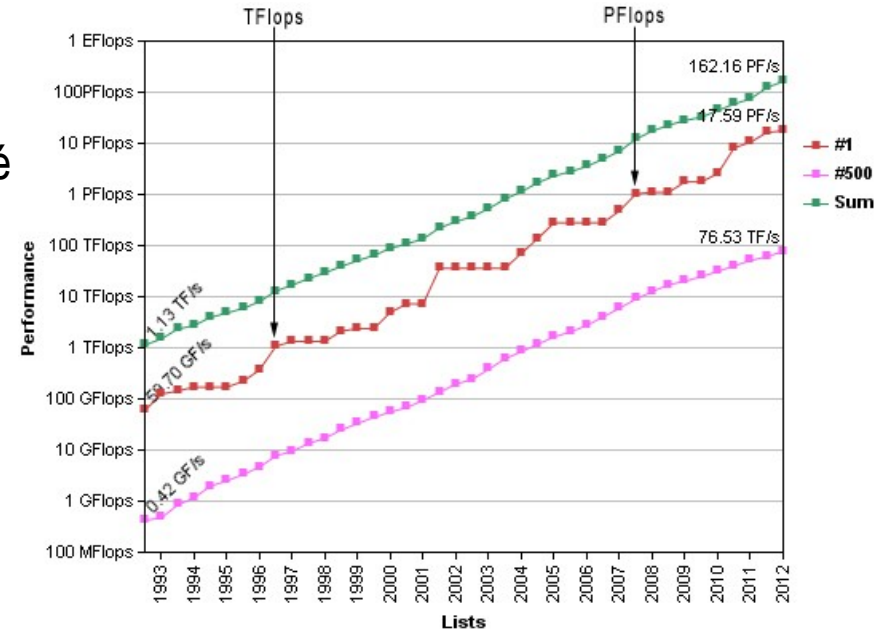
- La puissance des superordinateurs a beaucoup évolué
 $\approx 10^5$ depuis 1993
- Puissance = multi-cœur + vecteur + accélérateur
Complexité de la programmation
- Puissance = multi-cœur + vecteur + accélérateur
Complexité de la programmation

PROGRAMMATION

- Le Quicksort (N^2) reste préféré au Heapsort ($N \log(N)$)
- Le simplexe (exponentiel) supplante les solutions polynomiales
Les accès mémoire et le flot de contrôles justifient ces faits

METHODE

- Voyageur de Commerce → 25 nœuds = 25 ans sur TITAN
- Pourtant des tailles $> 80\,000$ ont été résolues en quelques jours
Combinaison des méthodes continues et discrètes



TITAN CRAY-XK7 the (2012) world fastest supercomputer

- 299 008 CPU cores (16-cores AMD Opteron 6274)
- 18 688 NVIDIA Tesla K20 GPUs
- Peak: 27.11 PFlop/s.
- Sustained: 17.59 PFlop/s (Linpack)

Activités d'enseignement

Principales activités d'enseignement

❖ Ecole des Mines de Paris

Cours: architecture matérielles et logicielles des ordinateurs

Type: enseignement spécialisé

Niveau: 3

Nombre d'heures: 30

❖ Université Paris-Sud Orsay

Cours: Modèle polyédrique et parallélisation automatique

Type: enseignement fondamental

Niveau: master 2 recherche

Nombre d'heures: 12

❖ Universités étrangères (Brésil, Egypte, Chili, ...)

Cours: programmation multithreadée et calcul vectoriel

Type: enseignement spécialisé

Niveau: master 2 et plus

Nombre d'heures: 20 par session



Activités d'encadrement

Activités d'encadrement (trois dernières années)

❖ Stages (niveau master 2)

2012: Wiktor Olko – ERASMUS Pologne – 6 mois – actuellement en CDI

2013: Pierre-Louis Caruana – Master 2 Orsay – 6 mois – actuellement en CDI

2014: Wajdi Farhani – Ecole Polytechnique de Tunisie – 6 mois – en cours

❖ Doctorats (principalement des co-encadrements officiels ou de fait)

✓ Alessandro Leite – Univ. Brasilia / Univ. Paris-Sud

Cloud Computing (modélisation, interface, évaluation) – 2 confs – soutenance prévue 09/2014

✓ Mikael Kruse – Univ. Brasilia / Univ. Paris-Sud

Data Layout for Memory Optimization – 2 confs et 1 journal – soutenance prévue 09/2014

✓ Dadi Elwardani – Univ. Mohammed 1^{er} (Maroc)

Indexation efficace d'images 3D à grande échelle – 3 confs et 1 journal – soutenu en 03/2014

✓ Rafaelli Coutinho – Univ. De Niteroi (Brésil)

Optimization of Cloud Computing Management – 1 papier en cours – soutenance en 2015

✓ Jose Afonso Sanches – Univ. De Niteroi (Brésil)

Traffic Aware Scheduling on Distributed Systems – 1 papier en cours – soutenance en 2016

❖ Postdoctorat

2013 (septembre-Octobre): Taina Raiol – Univ. Brasilia – Bioinformatic on the Cloud

Collaborations internes: CMM — Projet TIMC

→ Cahier des charges

TIMC (Traitement d'Images **M**ulti-**C**ible) est un projet collaboratif impliquant le **CRI**, le **CMM** et le **CAOR**.

Dans le cadre de **TIMC**, la collaboration bilatérale CRI-CMM devait porter sur

- ◆ L'accélération des programmes de détection des zones de texte dans une image
- ◆ L'étude et la mise en œuvre efficace d'un algorithme d'inondation de graphes sous contraintes

→ Réalisations au sujet de l'inondation de graphes (C. TADONKI, F. MEYER, et F. IRIGOIN)

- ◆ Amélioration et validation théorique de l'algorithme basé sur les dendrogrammes
- ◆ Conception d'un simulateur en ligne pour les graphes linéaires
- ◆ Mise en œuvre de l'algorithme en C avec structuration adéquate des données
- ◆ Parallélisation de l'étape finale de l'algorithme et mise en œuvre avec la librairie Pthread
- ◆ Ecriture d'un article accepté et présenté à la conférence ICCS à Cairns (Australie) le 12/06/2014



Neighborhood graph
generated by morph-m



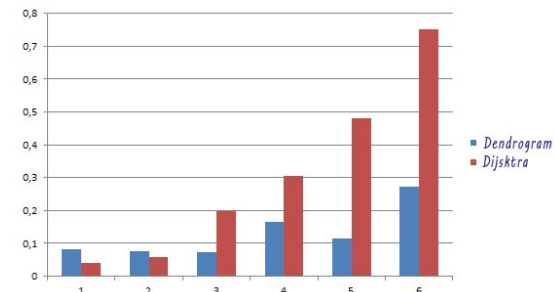
Number of nodes: 24 532
Number of edges: 96 138



Flooding values computed by the
dendrogram-based algorithm



Constructing the dendrogram: 1.737 s
Flooding process: 0.002 s
Whole algorithm: 1.739 s
Basic Dijkstra algorithm: 59.042 s



Collaborations internes: Géosciences

Premier contact avec Michel Tijani

Assistance relative aux travaux de thèse d'Irina Sin et collaboration autour du logiciel HYTEC

Forte collaboration avec Olivier Stab

- ✓ Aide ponctuelle au suivi du stage de Guillaume Caron (parallélisation avec OpenMP)
- ✓ Montage d'une pré-proposition ANR dénommée 3DCUT → **Acceptée**
- ✓ Montage de la proposition ANR complète avec VAREL Europe
- ✓ Thèse conjointe planifiée

3DCUT

Développement, optimisation et validation d'un simulateur de coupe 3D
d'aide à la conception des outils de forage de type PDC en conditions perturbées

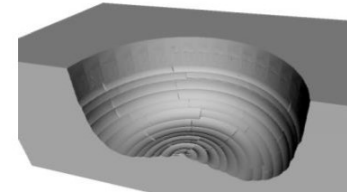
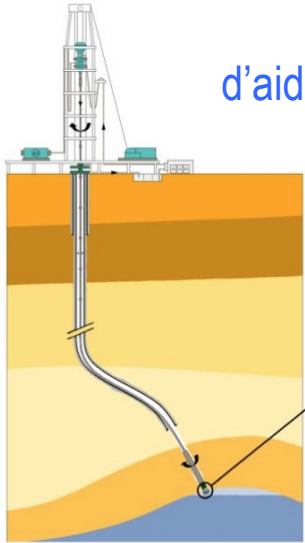


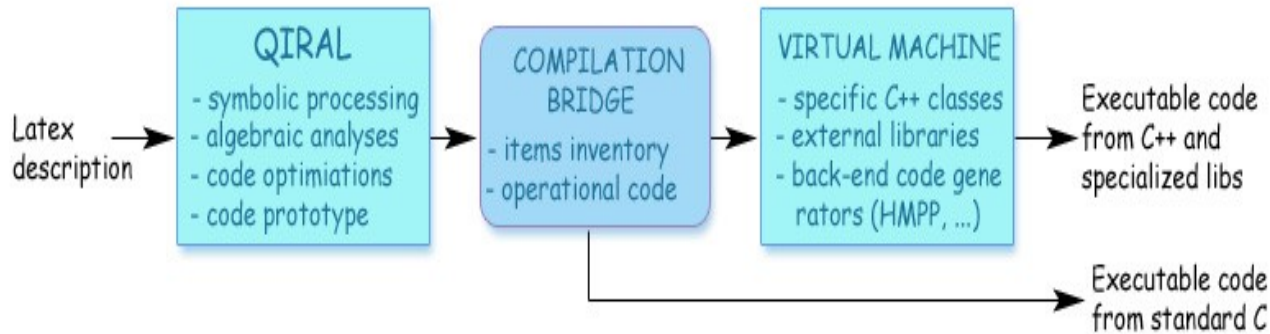
Fig. 2 : forage en conditions perturbées avec un outil PDC

→ **Collaborations potentielles:** CMA(RO & MARKAL) ♦ CBIO(Accélération) ♦ GEOSCIENCES(//)



Collaborations externes: LRI & INRIA

- ✓ INRIA-Scalay & INRIA Bordeaux & LRI & LAL & LPT: HPC et Generation de code pour la QCD



- ✓ Coopération spécifique avec INRIA-Scalay & LRI
 - Enseignement de master 2 (modèle polyédrique et parallélisation automatique)
 - Thèse en co-tutelle (Alessandro Leite) et co-encadrement (Mikael Kruse)
 - Coopérations internationales (STIC-AmSud)
 - Jury de thèse (Mouad Bahi)
 - Discussions scientifiques régulières

- ✓ Autres (potentielles) collaborations

Académiques: Rennes (IRISA) – Bordeaux (INRIA/LabRI)

Toulouse (CERFACS) - Avignon (LIA), Maison de la Simulation

Industriels: Amadeus – Air France – Thales – Intel



BioCloud Partners at LRI - Orsay



Collaborations internationales

✓ Collaboration avec le Brésil: **Université de Niteroi, Rio, Brasilia**

- Enseignement de master 2 (programmation parallèle - calcul vectoriel)
- Thèse en co-tutelle (Alessandro Leite) et co-encadrement (Jose Afonso et Rafaelli Coutinho)
- Projet de coopération (STIC-AmSud) **France – Brésil - Chili**
- Jury de qualification de thèse
- Visites réciproques (*chercheurs – doctorants – post-doctorants*) et publications conjointes
- Organisation d'événements scientifiques (workshops, tutoriels)



✓ Collaboration avec le Maroc: **Université Mohammed 1^{er} de Oujda**

- Séminaires (Calcul haute performance)
- Thèse en co-encadrement (Dadi Elwardani)
- Visites réciproques (*chercheurs*) et publications conjointes
- Comité de programme des conférences (ex: **ICMS 2014 à Marrakech**)

✓ Autres (potentielles) collaborations

USA, Suisse, Espagne (Matteo, BSC), Allemagne/Italie (ETMC), Tunisie



With Matteo Valero
Head of the BS Center



C. TADONKI

Collaborations internationales

BioCloud-EEAmSud



Federated Cloud Computing for Bioinformatics: Infrastructure, Algorithms and Applications

- Project collaboratif France – Brésil – Chili
- Financé respectivement par le MAE - CAPES - CONYCIT
- Thème du Cloud pour la Bioinformatique
- Missions – séjours (post)doctoraux
- Première année validée et financement de la deuxième année accordée
- Bourses de thèse
- Publications conjointes et workshops



BiCloud Group
Porto de Galinhas (2013)

- Christine Eisenbeis
- Claude Tadonki
- Maria Emilia Machado
- Lucas Walter
- Lucia Drummond
- Bertero Coiti Togawa
- Antonio Inostrosa-Ponta
- Maria Emilia Machado
- Lucas Walter

Fort accent sur les visites et échanges !!!



Exposé de Taina Raiol au CRI (2013)



Département d'informatique de l'université de Niteroi (Rafaelli, Ubiratam, Lucia, Claude)



Alba, Claude, Maria Emilia au CRI (2012)

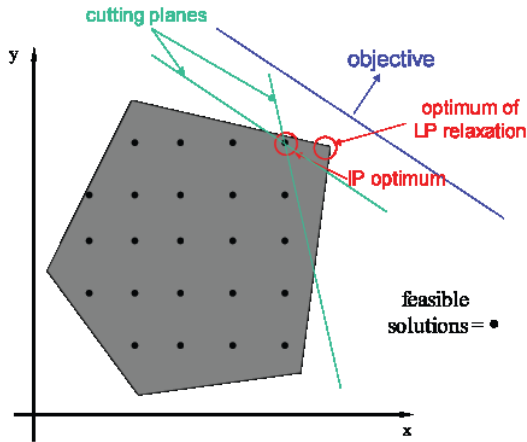




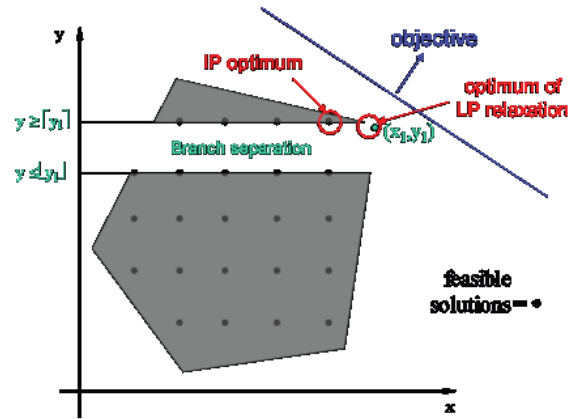
Principaux axes de recherche

➔ Recherche opérationnelle à grande échelle (méthode et calcul haute performance)

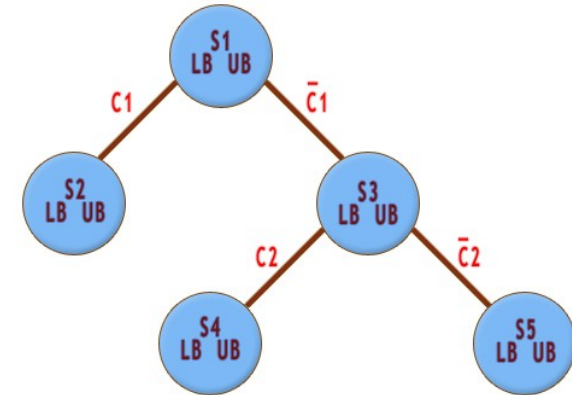
➔ De préférence dans le cadre de collaborations contractuelles



Integer Program: Cutting Planes



Integer Program: Branch and Bound
(or Divide and Conquer)



➔ Ordonnancement efficace dans les systèmes distribués ouverts (Cloud & Grilles)

➔ Modélisation et calcul scientifique (support pour la science et l'industrie)

➔ Programmation parallèle (mémoire partagée, mémoire distribuée, hybride, vectoriel)
➔ Vers l'exascale

➔ Algorithmique, complexité, et programmation avancée

MERCI BEAUCOUP DE VOTRE ATTENTION



QUESTIONS

