Héméra
Scientific Challenges using Grid'5000

Christian Perez
GRAAL project-team, Lyon

April 19th, 2011
Reims, France

Outline of the talk

- Some data about Grid’5000
- Overview of Héméra
- Scientific Challenges of Héméra
- Working Groups of Héméra
- Conclusion

Outline of the talk

- **Some data about Grid'5000**
  - **2008-now**
- Overview of Héméra
- Scientific Challenges of Héméra
- Working Groups of Héméra
- Conclusion

ANR (1/2)

- **USS Simgrid** (ANR 08 SEGI 022) 2009-2011.
- **CHOC** (CHallenge en Optimisation Combinatoire) (CIGC) 2006-2009.
- **DOCK** (Conformation sampling and docking on grids) (CIGC) 2006-2009.
- **Decert** (Déduction Certifiée) 2009-2011
- **GWENDIA** (ANR-06-MDCA-009) 2007-2010
- **Spreads**, (ANR Télécom) 2008-2010
- **MyCloud** (ANR Arpège), 2010-2014
- **SafeScale** (ANR ARA SSIA), 2005-2008
- **FP3C** (ANR JST)
- **SelfXL** (Self-management of compleX and Large scale systems) (SEGI) 2009-2012
- **HIPCAL** (Virtual Cluster for Bioinformatics applications) (CIGC) 2007-2010
- **PETAFLOW** (Peta-scale data intensive computing and vizualisation with transactional high-speed networking: application to upper airway flow), (Blanc International), 2009-2012.
- **MapReduce** (ANR ARPEGE 2010), 2010-2014
- **FP3C** (ANR-JST 2010), 2010-2013

ANR (2/2)

- **NUMASIS** (ANR CI) 2006-2009
- **DISCOGRID** (ANR CI) 2006-2008
- **COOP** (Multi Level Cooperative Resource Management), 2009-2012
- **LEGO** (ANR CICG), 2005-2009
- **SPADES** (ANR SEGI), 2009-2012
- **FVNano**
- **Dalia**
- **GCPMF**, 2006-2009
- **Soceda**, 2010-2013
- **Submitted**
  - GG-DOCK (Hybrid Adaptive Grid and GPU-based Optimization for Flexible Molecular Docking) (appel "Investissements d'avenir - Bioinformatique") 2011-2014.
  - Green-Wave (GREEN energy aWAre Virtualization for reducing Electrical consumption in computing and networking in large scale distributed systems) (ANR MD) 2011-2014.
  - SOP (ANR INFRA 2011)
  - SONGS, (ANR INFRA 2011)
  - Modulab (COSINUS 2011), 3 ans

European Projects (1/2)

- **Grid4All**, "develop a grid infrastructure and middleware for the collaboration of dynamic, small virtual organizations such as communities, schools and families", STREP, 2004-2009.
- **XtreemOS**, "Building and Promoting a Linux-based Operating System to Support Virtual Organizations for Next Generation Grids", IP, FP6-033576, 2006-2010.
- **GOSSPLE**, "Gossple aims at providing a radically new approach to navigating the digital information", ERC Starting Grant, 2008-2013.
- **Autonomic Internet Project**, Future Internet Network Program, 2008-2011.
- **SOA4ALL** (STREP), 2008-2011
- **SCALUS**: Scaling by means of ubiquitous storage (MCITN EU), 2009-2013
- European COST Action **IC804** on "Energy efficiency for large scale distributed systems",2009-2013
- **BonFIRE**, "Testbeds for Internet of Services Experimentation", IP, 2010-2013.
- **Contrail**, "Open computing infrastructures for elastic services", IP, FP7-ICT-257438, 2010-2013.

European Projects (2/2)

- **GEYSERS** "Generalised Architecture for Dynamic Infrastructure ServiceS", 7th framework programme theme ICT-2009.1.1: Call 4 /The Network of the Future, 2010-2012.
- **SAIL** "Scalable & Adaptive Internet soLutions", 7th framework programme theme ICT-2009: Call 5 /Large-scale integrating project. 2010-2013.
- **PrimeEnergyIT** "Supporting the market for energy efficient central IT" - Intelligent Energy in Europe Programme 2010-2012.
- **PLAY** (STREP), 2010-2013
- **Submitted**
  - **P2PC2**: (Peer-to-Peer Computing on multi-Cloud Infrastructures) (Call IP - BonFire), 2011-2012.
  - **SOFEXA**: Software for exascale. submitted 01/2011.
  - **CoolEmAll**, FP7 STREP

Defended PhD (1/2)

- **Lakhdar Loukil**. « Méta-heuristiques hybrides parallèles pour le Q3AP sur environnements à grande échelle ». Univ. d'Oran le 13/12/2010. (Lille)
- **Ye Zhang**. « Méthodes itératives hybrides asynchrones sur plateformes de calcul hétérogènes pour la résolution accélérée de grands systèmes linéaires ». 2009. (Lille)
- **Alexandru Tantar**. « Méta-heuristiques parallèles coopératives pour le problème d'échantillonnage de protéines sur grilles de calcul ». 2009. (Lille)
- **Jean-Charles Boisson**. « Modélisation et résolution par metaheuristiques coopératives : de l'atome a la séquence protéique ». 2008. (Lille)
- **Diego Caminha Barbosa De Oliveira**, « Fragments de l'arithmétique dans une combinaison de procédures de décision », 14/03/2011, Nancy.
- **Julien Sopena**, « Algorithmes d'exclusion mutuelle : tolérance aux fautes et adaptation aux grilles ». (2008) (Orsay)
- **Olivier Delannoy**, « YML : un workflow scientifique pour le calcul haute performance » (2008) (Orsay)
- **Rémi Sharrock** « Gestion autonomique de performance, d'énergie et de qualité de service. Application aux réseaux filaires, aux réseaux de capteurs et aux grilles de calcul », 2010, (Toulouse)
- **Ludovic Hablot**. « Réseau longue distance et application distribuée dans les grilles de calcul : étude et propositions pour une interaction efficace », 2009 (Lyon)
- **Patrick Loiseau**. « Contributions to the analysis of scaling behavior and quality of service in networks: experimental and theoretical aspects ». 2009 (Lyon)
- **Bogdan Nicolae**, nov 2010 (Rennes)
- **Julien Bigot**, « Du support générique d'opérateurs de composition dans les modèles de composants logiciels, application au calcul scientifique », 2010 (Rennes/Lyon)

Defended PhD (2/2)

- **Boris Daix**, « Abstraction des systèmes informatiques à haute performance pour l'automatisation du déploiement d'applications dynamiques », 2009 (Rennes)
- **Hinde Lilia Bouziane**, « De l'abstraction des modèles de composants logiciels pour la programmation d'applications scientifiques distribuées », 2008 (Rennes)
- **Jean Arnaud**, 2010 (Grenoble)
- **Yiannis Georgiou**, 2010 (Grenoble)
- **Brice Videau**, 2009 (Grenoble)
- **Carlos Jaime Barrios Hernández**, 2010 (Grenoble)
- **Lucas Schnorr**, 2009 (Grenoble)
- **Xavier Besseron**, 2010 (Grenoble)
- **I. Filali**, « Improving Resource Discovery in P2P Systems », 2011 (Nice)
- **Tram Truong Huu**. « Workflow-based applications performance and execution cost optimization on cloud infrastructures », 2010 (Nice)
- **V. D. Doan**, « Grid computing for Monte Carlo based intensive calculations in financial derivative pricing applications », 2010 (Nice)
- **Elton Mathias**, « Hierarchical Multi-Domain computing based upon a component-oriented approach », 2010 (Nice)
- **M. Leyton**, « Advanced Features for Algorithmic Skeleton Programming », 2008 (Nice)
- **G. Charrier**, (Lyon)
- **B. Depardon,** (Lyon)

HDR

- **Stéphane Genaud**, « Exécutions de programmes parallèles à passage de messages sur grille de calcul », Université Henri Poincaré - Nancy I, 2009, Strasbourg
- **Gabriel Antoniu**, « Contribution à la conception de services de partage de données pour les grilles de calcul », ENS Cachan, 2009, Rennes
- **Thierry Monteil,** « Du cluster à la grille sous l'angle de la performance », 2010, Toulouse.
- **Eddy Caron**, « Contribution to the management of large scale platforms: the DIET experience », ENS Lyon, 2010, Lyon

Software Validated on Grid'5000 (1/2)

- **CONFIIT,** Computation Over Network with Finite number of Independent and Irregular Tasks (Reims)
- **ParadisEO-G,** Parallel and Distributed Evolving Objects on top of Globus (Lille)
- **DeployWhere/FDF,** framework open source orienté composant pour le déploiement de logiciels distribués et hétérogènes (Lille)
- **Wrekavoc** (Nancy)
- **GridTPT**, plateforme de test distribuée pour prouveurs de formules (Nancy)
- **veriT**, solveur de formules SMT (Nancy)
- **GSOC**, Grid Security Operation Center (Besançon)
- **dPerf**, prédiction de performances des applications distribué en pair-à-pair (Besançon)
- **XtreemOS** (Rennes)
- **BlobSeer** (Rennes)
- Bibliothèque de mesures de la consommation électrique. Placement de tâches Energy-aware (Toulouse)

Software Validated on Grid'5000 (2/2)

- **VMdeploy** / Saline (Nantes)
- **KEntropy** (Nantes)
- **Kargo** (Nantes)
- **KaStore** (Nantes)
- **kDFS** (Nantes)
- **Metroflux** (Lyon)
- **ANPI** (Lyon)
- **OVNI5000** (Lyon)
- **SHOWATTS** (Lyon)
- **MPI5000** (Lyon)
- **Green Grid5000** (Lyon)
- **ULCMi** (Lyon)
- **HLCMi** (Lyon)
- **DHICO** (Lyon)
- **DIET** (Lyon)
- **Grudu** (Lyon)
- **P2P-MPI** (Strasbourg)
- **MOTEUR** workflow manager (Nice)

Industrial Relations

- **Alcatel-Lucent Bell Labs**
  - Traffic aware routers
- **Orange Labs**
  - Data placement algorithmes on P2P architectures, thèse Cifre (Orange Labs, Issy les Moulineaux, UPMC (Regal))
- Projet **MSR**-INRIA
  - Microsoft Azure: A-Brain (AzureBrain), « cloud » testbed for experimenting storage technologies (Kerdata)
- **EDF R&D** (Myriads, GRAAL)
- **BULL** (GRAAL, Runtime)
  - Placement of applications
- **IBM**
  - BlueWaters, Clouds

Startup companies

- Three startups companies started by Grid'5000 researchers
  - LYaTiss (LIP, ENS Lyon) around virtualization et network QoS
  - SysFera (LIP, ENS Lyon) around large scale computing over Grids and Clouds
  - Activeon (INRIA Sophia) around distributed computing

Examples of Ongoing Research Impact

- SuperComputer
  - IBM BlueWaters
    - IO, *cf WG "Data"*
  - PRACE-2P
    - Component model, *cf WG "Programming Models"*
- Clouds
  - Lots of works
    - IO: Nimbus, MS Azure, *cf WG "Data" / SC "MapReduce"*
    - Energy: Entropy, *cf WGs "Energy" & "Virtualization" / SC "Energy"*
- Production Grids
  - "EGEE" (gLite)
    - Behavior, *cf WG "Methodology"/SC "Orchestration"*
- Desktop Grids
  - EDGES, EDGI, DEGISCO (GRAAL, MESCAL, IN2P3-LAL)
    - G5K used to test, validate middleware, to obtain traces, ...
- …

Outline of the talk

- Some data about Grid'5000
- **Overview of Héméra**
- Scientific Challenges of Héméra
- Working Groups of Héméra
- Conclusion

Overview of Hemera

- Goals
  - Demonstrate ambitious up-scaling techniques for large scale distributed computing by carrying out several dimensioning experiments on the Grid'5000 infrastructure
  - Animate the scientific community around Grid'5000
  - Enlarge the Grid'5000 community by helping newcomers to make use of Grid'5000.
- Open to everyone (not only INRIA)
- Structure
  - A direction committee
    - The same as Aladdin-G5K
  - Scientific challenges
  - Working groups
    - In cooperation with GDR ASR

Hemera: Current Participant List

1. *ACADIE* - Assistance à la Certification d'Applications DIstribuées et Embarquées
2. ALGORILLE - Algorithms for the Grid
3. *APO* - Algorithmes Parallèles et Optimisation
4. ASAP - As Scalable As Possible: foundations of large scale dynamic distributed systems
5. ASCOLA - Aspect and composition languages
6. *ASTRE* - Architecture, Systèmes, Temps-Réel, Embarqués
7. *CC-IN2P3* - Equipe de recherche du Centre de Calcul de l'IN2P3
8. CEPAGE - Chercher et Essaimer dans les Plates-formes A Grande Echelle
9. DOLPHIN - Parallel Cooperative Multi-criteria Optimization
10. GRAAL - Algorithms and Scheduling for Distributed Heterogeneous Platforms.
11. GRAND-LARGE - Global parallel and distributed computing
12. *ICPS* - Scientific Parallel Computing and Imaging
13. KERDATA - Cloud and Grid Storage for Very Large Distributed Data
14. OASIS - Active objects, semantics, Internet and security
15. MAESTRO - Models for the performance analysis and the control of networks
16. MESCAL - Middleware efficiently scalable
17. *MINC* - MIcro et Nanosystèmes pour les Communications sans fils
18. *MRS* – Modélisation et contrôle des Réseaux et Signaux
19. MYRIADS - Design and Implementation of Autonomous Distributed Systems
20. REGAL - Large-Scale Distributed Systems and Applications
21. RESO - Protocols and Software for Very High-Performance Network
22. RUNTIME - Efficient runtime systems for parallel architectures
23. SAGE - Simulations and Algorithms on Grids for Environment

Hemera: Current Participant List

Outline of the talk

- Some data about Grid’5000
- Overview of Héméra
- **Scientific Challenges of Héméra**
- Working Groups of Héméra
- Conclusion

List of Open Challenges

- Network
  - Traffic Awareness
- System
  - Energy Profiling of Large Scale Applications
  - Robustness of Large Systems in Presence of High Churn
  - Orchestrating Experiments on the gLite Production Grid Middleware
    - Since 2011, April 1st
- Programming Paradigm
  - Large Scale Computing for Combinatorial Optimization Problems
  - Scalable Distributed Processing Using the MapReduce Paradigm
- Domain Specific
  - Multi-parametric Intensive Stochastic Simulations for Hydrogeology
  - Thinking GRID for Electromagnetic Simulation of Oversized Structures

Challenge on Network

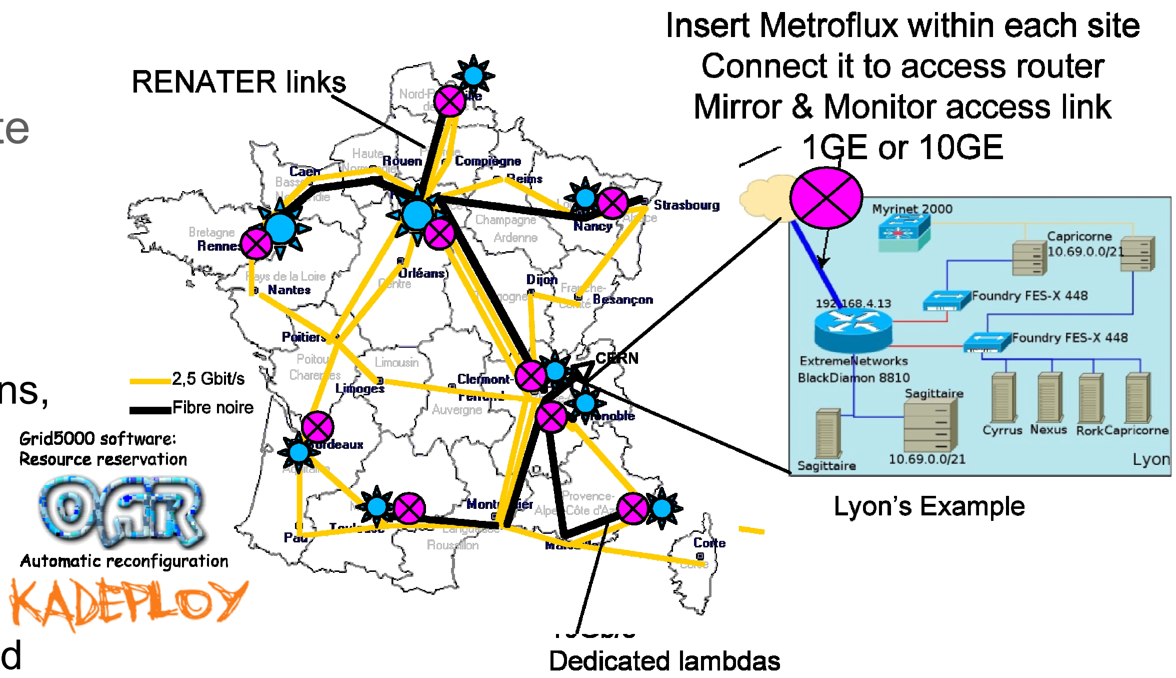
Traffic Awareness

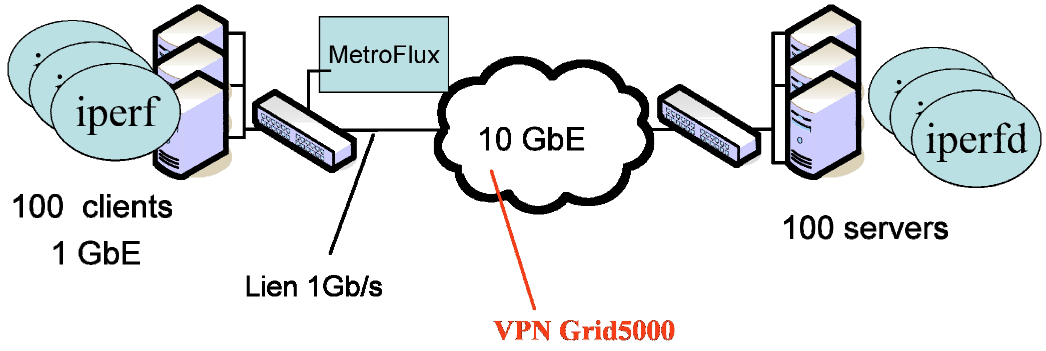
Traffic Awareness (Traffic)

- Leader
  - Paulo Gonçalves (RESO), K. Avrachenkov (MAESTRO)
- Context
  - Common Labs INRIA & Alcatel Bell Labs
  - Design of traffic aware routers for high-speed networks
- Objective
  - Identify application classes from the behavioral (semantic) analysis of corresponding traffic
    - How does traffic behavior relate to flows semantic?
    - Which traffic characteristics are capturable on high speed networks?
    - Which constraints to get meaningful characteristics on-line?
- Difficulties / Pitfalls
  - Initial program hampered by
    - Difficulty to obtain (download or simulate) traffic traces characteristic of different applications
    - Semi-supervised learning (as primarily thought) does not seem to over-perform traditional decision tree algorithms

Traffic Awareness & Grid5000

- How do we use Grid'5000?
  - As a controllable testbed to emulate large-scale, high speed networks
- Why do we use Grid'5000?
  - To reproduce the conditions of realistic environments …
    - Congestion, multi-scale aggregations, large size, heterogeneity.
  - that can alter the flows' semantic
- Technological advances
  - MetroFlux
    - Packet capture probe on high speed links and under controlled situations
  - Virtualization
    - Deployment of a physical infrastructure (open flow routers, switches) to emulate a virtual sub-network
  - Trans-national link
    - Construction, through Grid'5000, of a 1Gbps dedicated link between France and Japan





Challenges on Distributed Systems

Energy Profiling of Large Scale Applications

Robustness of Large Systems in Presence of High Churn

Orchestrating Experiments on the gLite Production Grid Middleware

Energy Profiling of Large Scale Applications (Energy)

- Leaders
  - Laurent Lefèvre (RESO), Jean-Marc Pierson (IRIT), Jean-Marc Menaud (ASCOLA)
- Issues
  - Reduce energy consumption of large-scale infrastructure
  - Management of physical resources & virtualized resources
- Objective
  - Handle energy efficiency aspects of large scale applications deployed on multiple sites
- Roadmap
  - Model (complex) energy consumptions of systems and applications
    - Need to profile applications
  - Develop software to log, store and expose energy usage
    - Make use of the G5K energy sensing infrastructure
  - Experiments on large scale and heterogeneous infrastructure

How to Decrease Energy Consumption without Impacting Performance?

- How to monitor and to analyze the usage and energy consumption of large scale platforms?
- How to apply energy leverages (large scale coordinated shutdown/slowdown)?
- How to design energy aware software frameworks?
- How to help users to express theirs Green concerns and to express tradeoffs between performance and energy efficiency?

Energy: Challenges

- Exploring energy aspects at large scale
- Two focus
  - Applications deployed on real physical resources
  - Applications and services deployed on virtualized resources
- Providing feedback on large scale applications
- Extending the Green Grid5000 infrastructure
- Analyzing energy usage of large scale applications per components
- Designing energy proportional frameworks (computing, memory or network usage)

Robustness of Large Systems in Presence of High Churn (P2P-Ch)

- Leaders
  - Pierre Sens (REGAL), Jean-Marc Vincent (MESCAL)
- Issues
  - Large scale distributed, heterogeneous platforms
    - 10K-100K nodes
  - Frequency of connections/disconnections (churn)
- Objective
  - Maintain the platform connectivity in presence of high churn
- Roadmap
  - Develop a formal model to characterize the dynamics
    - Failure Trace Archive – http://fta.inria.fr
  - Design algorithms for basic blocks of distributed systems on a churn-resilient overlay
  - Experiments these algorithms on G5K

Robustness of Large Systems in Presence of High Churn (P2P-Ch)

- Distributed algorithms for dynamic systems
  - Variable number of peer, dynamic topology, mobility
- Two approaches
  - Determinist
    - Consensus, mutual exclusion (1 internship Regal)
  - Probabilistic
    - High volatility, partitioning management
- Integrate models / traces in fault injection tools
  - FCI-FAIL – (Orsay)
- Large scale experiments on Grid’5000
  - 1 internship “Polytechnique” in REGAL (April – June 2011)

Orchestrating Experiments on the gLite Production Grid Middleware (Orchestration)

- Leaders
  - Lucas Nussbaum (ALGORILLE), Frédéric Suter (CC IN2P3)
  - Created 2011, April 1st (was research grid/production grid project)
- Issues
  - Production Grid Middleware
- Objective
  - Explore the use of the Grid'5000 testbed as a test environment for production grid software such as gLite and other related services
- Roadmap
  - Define a detailed procedure to deploy the gLite middleware on Grid'5000
  - Define reusable services
    - Control of a large number of nodes, data management, experimental condition emulations, load and fault injection, instrumentation and monitoring, etc.
  - Develop experiment orchestration middleware
  - Perform large-scale experiments involving the gLite middleware and applications from production grids

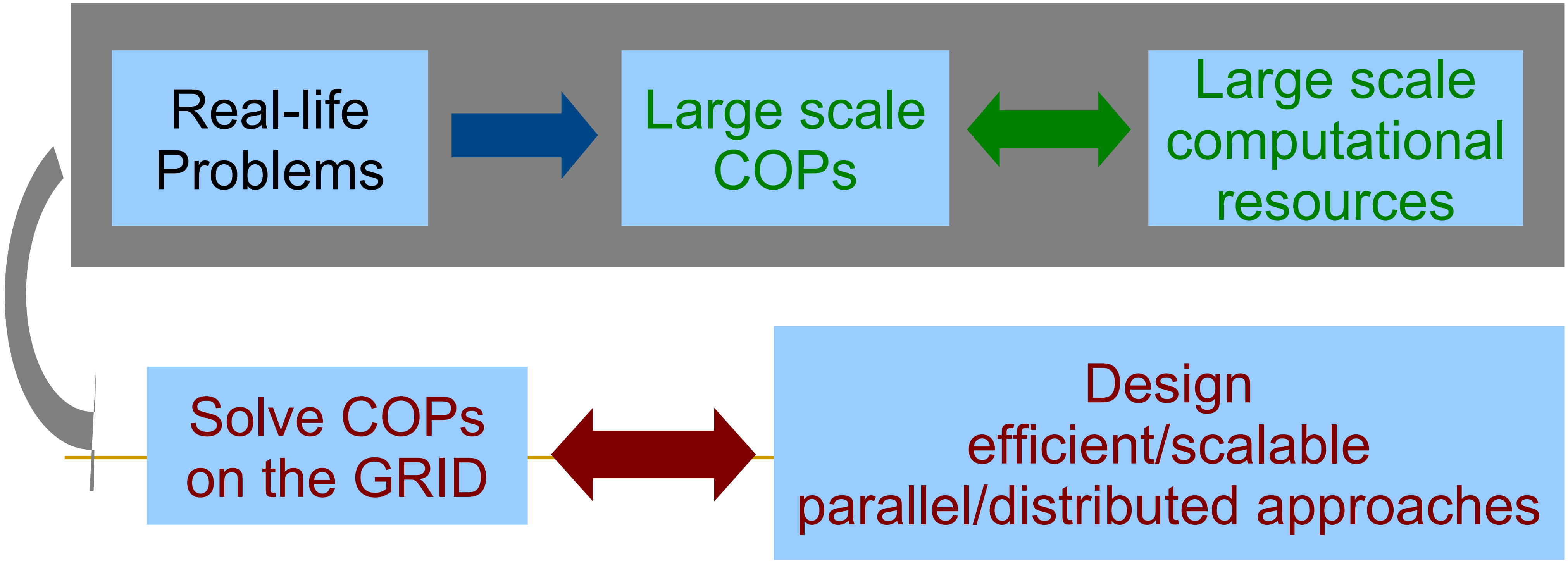
Challenges on Programming Paradigms

Large Scale Computing for Combinatorial Optimization Problems

Scalable Distributed Processing Using the MapReduce Paradigm

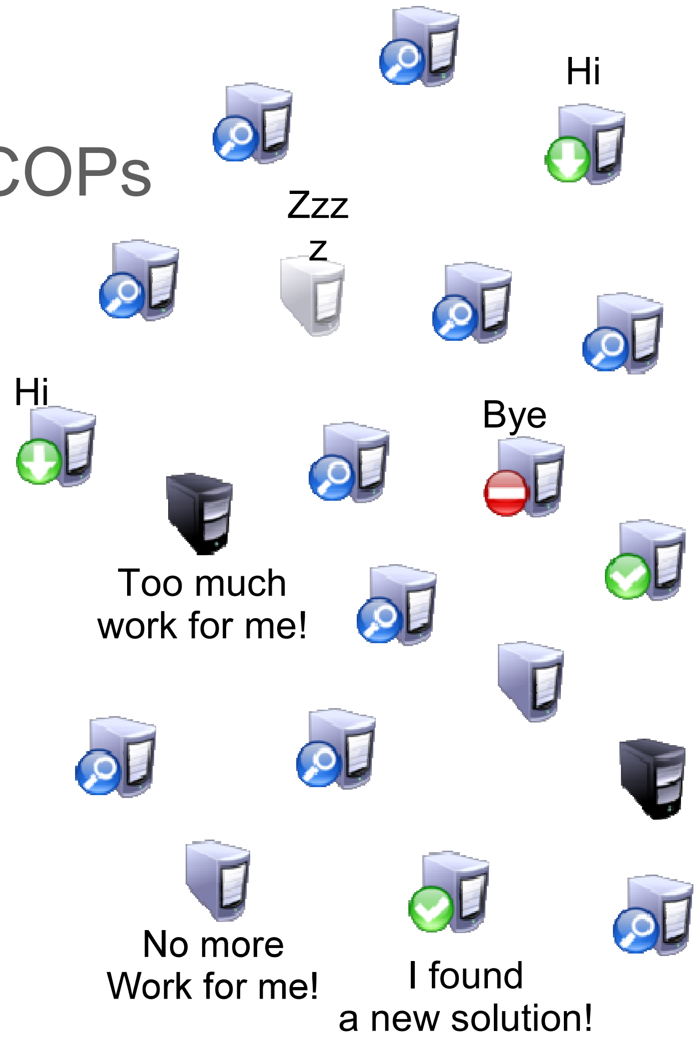
Large Scale Computing for Combinatorial Optimization Problems (COPs)

- Leaders
  - Bilel Derbel (Dolphin), Nouredine Melab (Dolphin)
- Objectives
  - Solve optimally large scale Combinatorial Optimization Problems (COPs) using huge amount of computational resources



Large Scale Computing for Combinatorial Optimization Problems (COPs)

- Goals at the application level
  - Solve optimally previously unsolved COPs
  - New specific COPs approaches
- Goals at the algorithmic level
  - How to gain in scalability?
    - Pure peer-to-peer approaches
    - Fully distributed algorithms
  - How to address latencies/resources volatility?
    - Fault-tolerant/dynamic algorithms
    - Redundancy vs efficiency

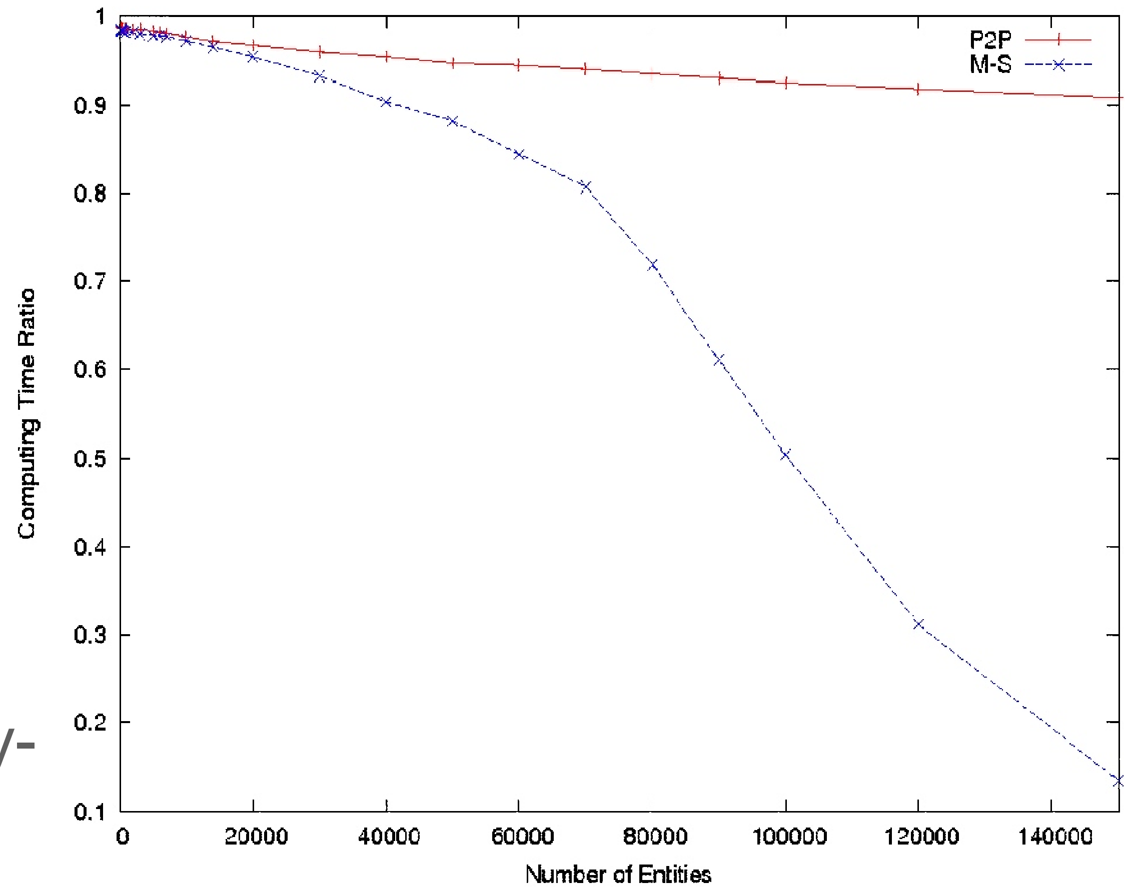


Large Scale Computing for Combinatorial Optimization Problems (COPs)

- How GRID5000 can help?
  - At the application level (make it a success story)
    - Effectively find unknown and optimal COPs solutions
  - At the algorithmic level (make it smart)
    - Experiments/simulations are mandatory to validate our algorithms
    - Measure the scalability / efficiency / congestion / fault-tolerance robustness of our approach

COPS: First Results

- P2P Branch&Bound
  - Fully distributed
    - Work sharing / Load balancing
    - Termination detection
    - Network congestion (messages)
  - Topology independent
  - Validated using a Pastry-like overlay and up to 150,000 processes



4th SCALE Challenge Finalist (CCGRID 2011)

COPs: Next Challenging Issues

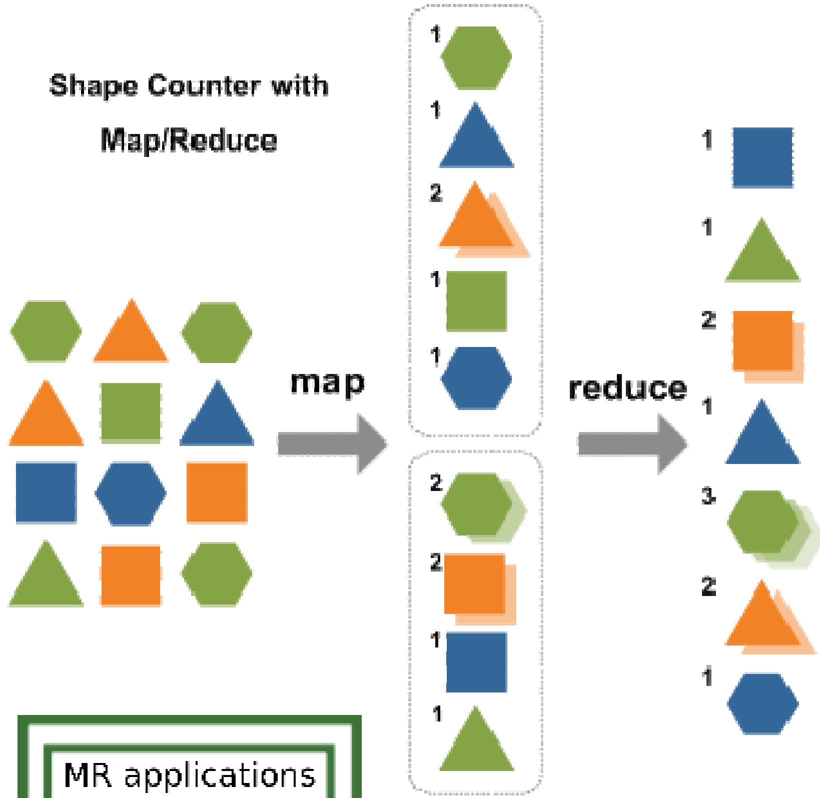
- Extensions to a dynamic, volatile and fully distributed environment
  - Post-doc started 2011, January
  - Maintain overlay connectivity distributely
  - Efficient fault-tolerant distributed algorithms
- Study the impact of network heterogeneity
- Study the proposed distributed protocol under some formal model capturing the dynamicity of the network
  - Related to high churn challenge
- Study the scalability of the proposed dynamic approach
  - Large scale experimentations, simulations, emulation

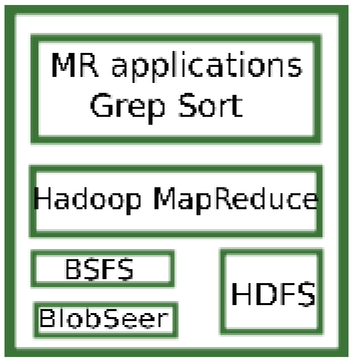
Scalable Distributed Processing Using the MapReduce Paradigm (Map-Red)

- Leaders
  - Gabriel Antoniu (KERDATA), Gilles Fedak (GRAAL)
- Issues
  - Distributed data-intensive applications (Peta-bytes)
  - Data storage layer
    - Efficient, fine-grain, high throughput accesses to huge files
    - Heavy concurrent access to the same file (R/W)
    - Data location awareness
    - Volatility
- Objective
  - Ultra-scalable MapReduce-based data processing on various physical platform (clouds, grids & desktop computing)
- Roadmap
  - Advanced data & meta-data management techniques
  - MapReduce on desktop grid platforms
  - Scheduling issues
    - Data & computation, heterogeneity, replication, etc.

Scalable Map-Reduce Processing

- ANR Project Map-Reduce (2010-2014) associated to the MapReduce HEMERA Challenge
  - Partners
    - INRIA (KerData, GRAAL), Argonne National Lab, UIUC, JLPC, IBM, IBCP, MEDIT
  - Goal
    - High-performance map-reduce processing through concurrency-optimized data processing
  - *http://mapreduce.inria.fr*
  - An objective of the project
    - Use BlobSeer as back-end storage for VM images and cloud application data
- Experiments done on Grid'5000
  - Up to 300 nodes
  - Plans: joint deployment G5K+FutureGrid (USA)
- Results to be transferred on real clouds
  - Nimbus (ANL): ANR MapReduce project
  - Microsoft Azure: A-Brain project (MSR-INRIA)
- First results : HPDC 2011 (12.9%)



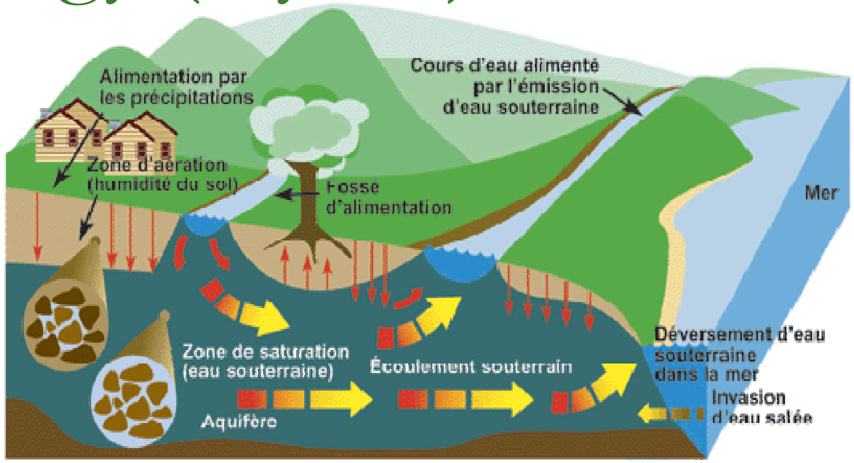


Challenges on Applications

Multi-parametric Intensive Stochastic Simulations for Hydrogeology

Thinking GRID for Electromagnetic Simulation of Oversized Structures

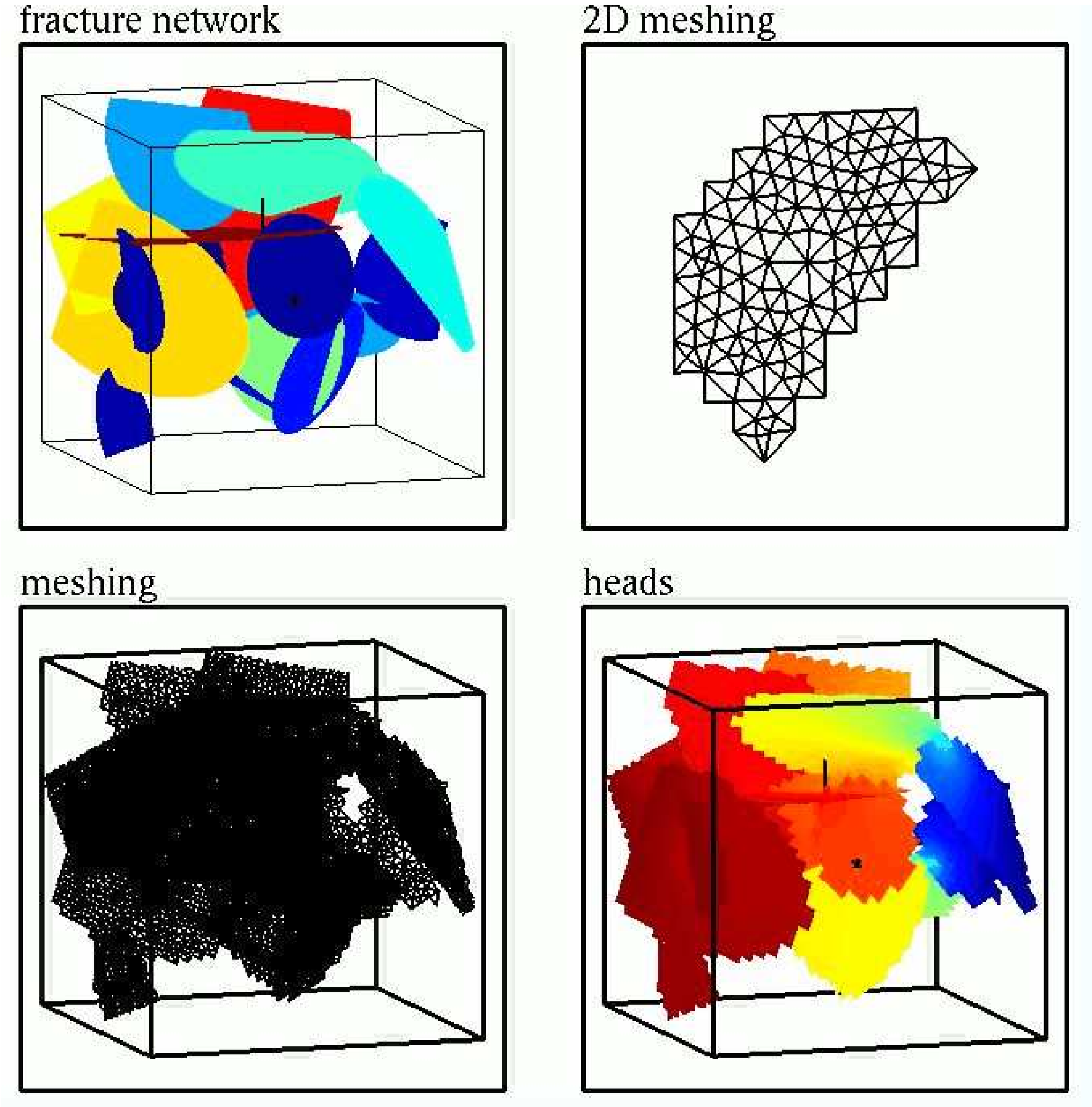
Multi-Parametric Intensive Stochastic Simulations for Hydrogeology (Hydro)



© http://www.ec.gc.ca/eau-water/

- Leader
  - Jocelyne Erhel (SAGE)
- Issues
  - Groundwater resource management & remediation
  - Limited knowledge
    - Highly heterogeneous and fractured geological formations
  - Numerical models
    - Probabilistic data + uncertainty quantification methods
      - Stochastic framework (multiple simulations)
    - Various physical parameters
  - Large size geological domain to discretize
- Objective
  - Efficient execution of multi-parametric heavy computation simulations
- Roadmap
  - Study how to program, deploy & schedule the application
  - Validate the approach for increasing level of parallelism for 2D problems then 3D problems
  - First challenge for the Hemera IJD

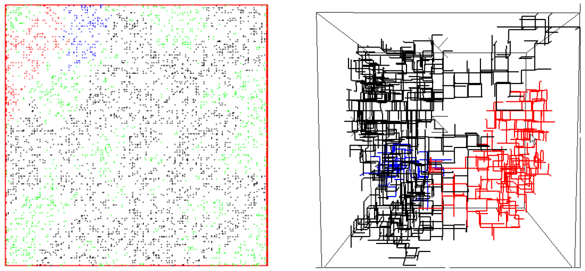
Example of Flow in Fractured Ground



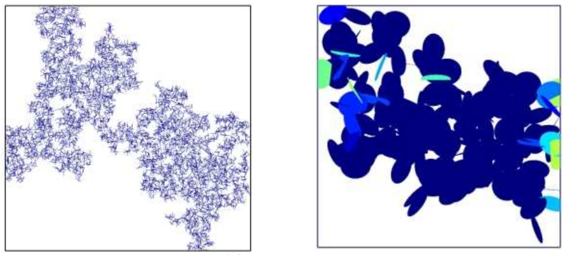
Influence of Fracture Network Complexity on Upscaling Hydrodynamic Laws

Objective: establish references results for more realistic fracture networks

STRUCTURES
Percolation fractals



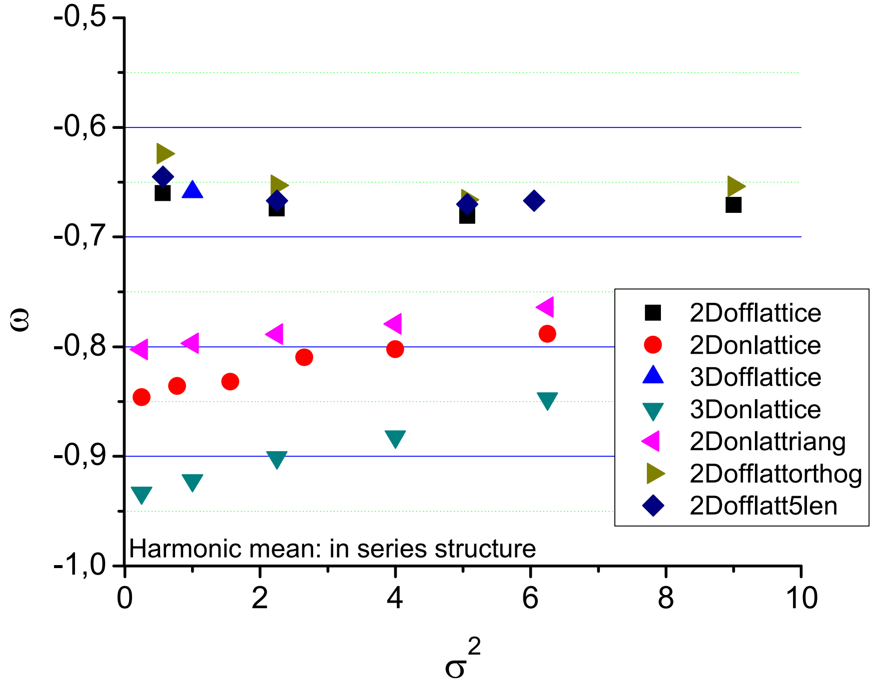
on lattice



off lattice

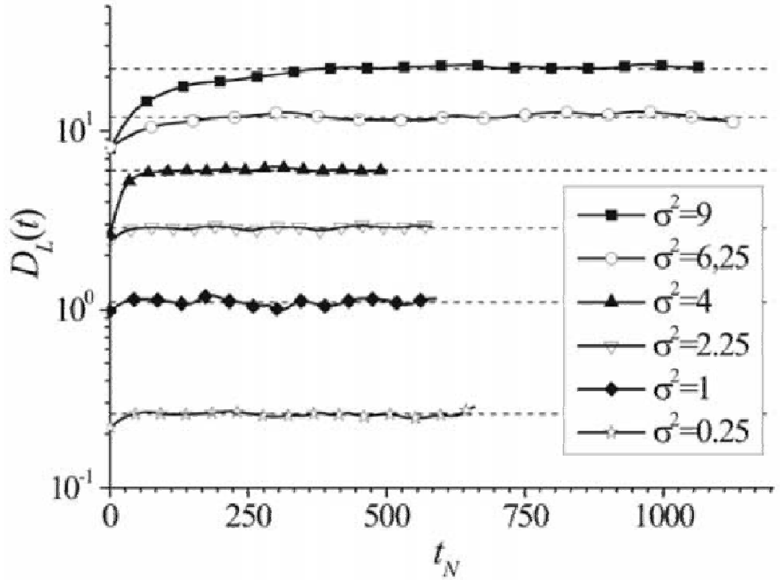
Local permeability distribution:
$LN(\mu,\sigma^2)$

Global permeability
$K(macro)=\mu \exp(\omega\sigma^2/2)$

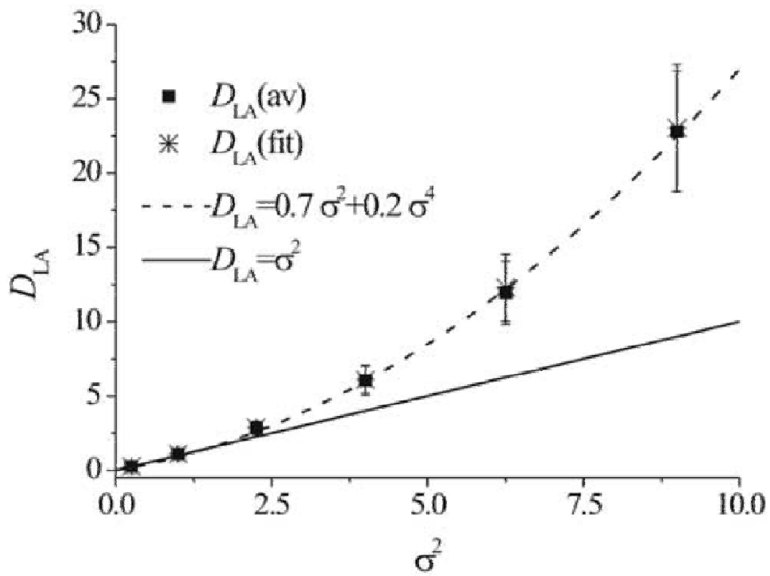


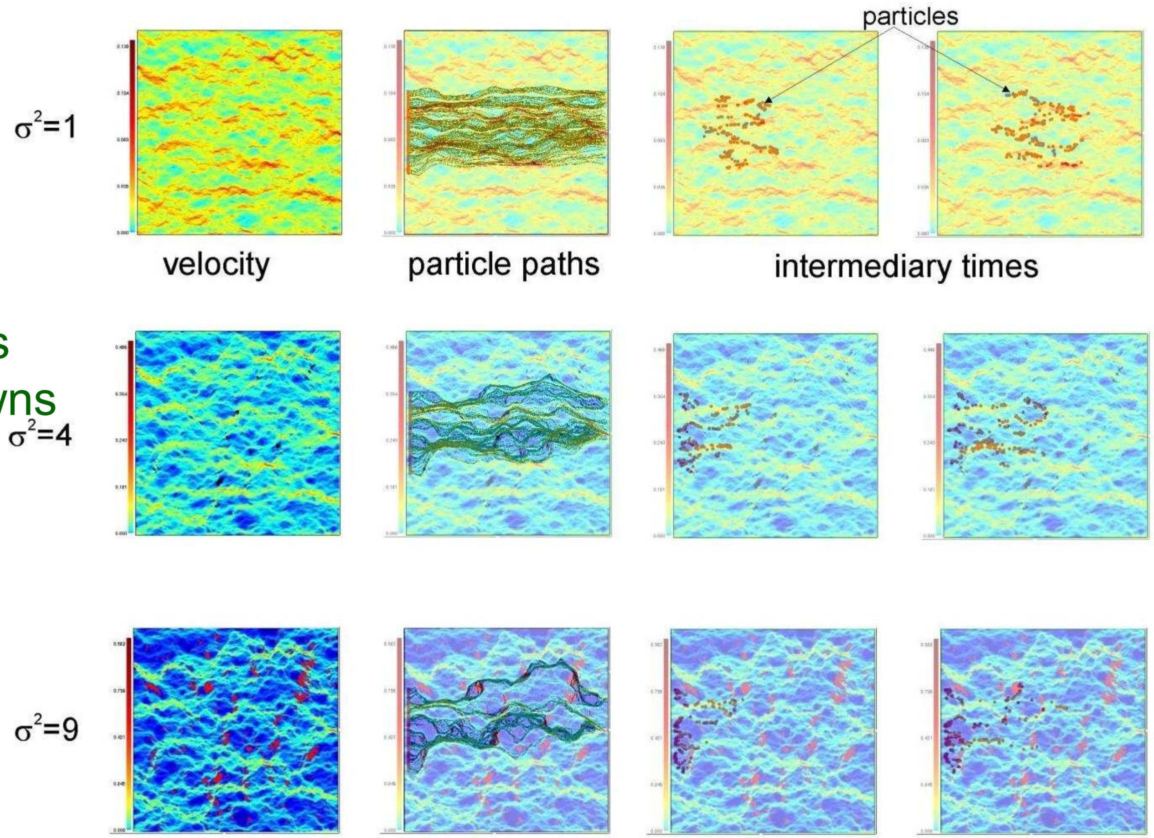
10^5 simulations of ~10^5-10^7 nodes
strong limitations for 3D off lattice networks

Longitudinal macro dispersion with pure convection



Each curve represents 100 simulations on domains with 67.1 millions of unknowns

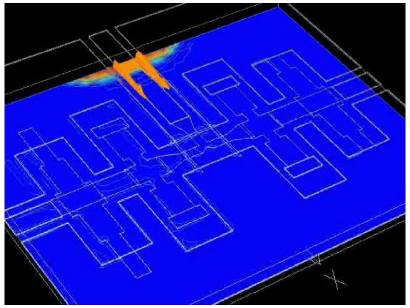




First numerical results obtained for such a large sigma!
From 2D to 3D

Thinking GRID for Electromagnetic Simulation of Oversized Structures (Electro)

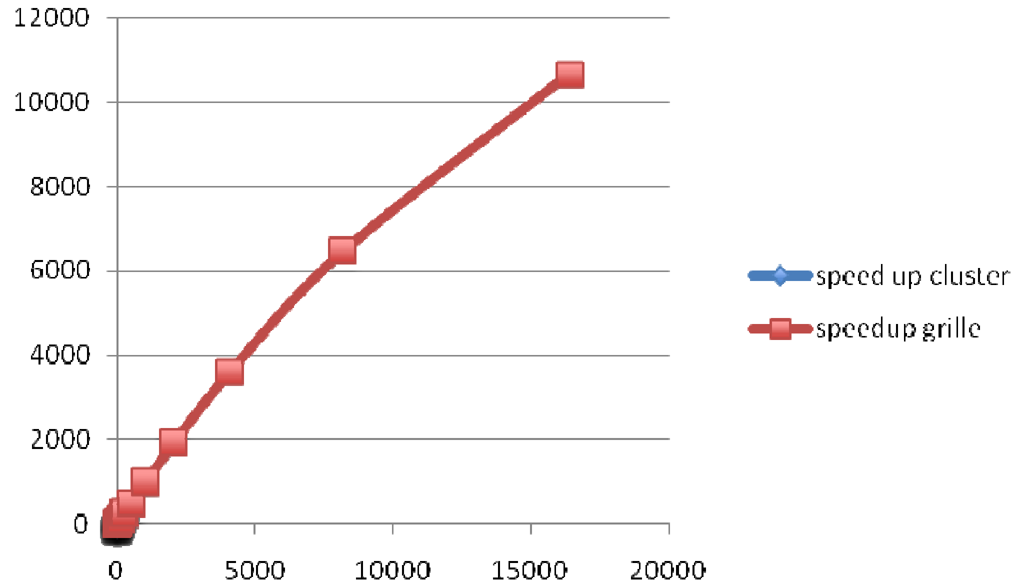
- Leaders
  - Hervé Aubert (MINC-LAAS), Thierry Monteil (MRS-LAAS), Patricia Stolf (ASTRE-IRIT)
- Design of sophisticated communication infrastructures
  - Transmission of signals from airborne sensors
  - Wheel antennas emitting data
    - E.g. tire pressure to a collector located inside a vehicle.



- Objectives
  - Increasing number of unknown parameters
  - Integrated in environments filled with various metallic and dielectric structures of different sizes
- Roadmap
  - Needs to develop a new 2D and 3D approach to simulate the electromagnetic behavior of large structures (planes, cars, buildings, etc).
  - Need of parallel execution for this oversized structure
  - Need to explore different configurations with multi-parametric executions

Thinking GRID for Electromagnetic Simulation of Oversized Structures (Electro)

- Utilization of multithreading and MPI over grid
- Collaboration between application, middleware and platform
- Uses of autonomic policies:
  - Breakdown or performance loss of a set of machines
  - Automatic execution of new simulations in self adapting network set-ups
  - Autonomic exploration of new solutions in multi-parametric mode
- First theoretical estimation of speedup for oversized problem
- Challenge objectives
  - Autonomic deployment and reconfiguration on grid5000
  - Parallel algorithm for electromagnetic simulation
- Héméra objectives
  - Large scale experiment
  - Experiment and support for electromagnetic researchers



Outline of the talk

- Some data about Grid'5000
- Overview of Héméra
- Scientific Challenges of Héméra
- **Working Groups of Héméra**
- Conclusion

List of Working Groups

- Transparent, Safe and Efficient Large Scale Computing
  - Stéphane Genaud (ICPS), Fabrice Huet (OASIS)
- Energy Efficient Large Scale Experimental Distributed Systems
  - Laurent Lefèvre (RESO), Jean-Marc Menaud (ASCOLA)
- Bring Grids Power to Internet-Users thanks to Virtualization Technologies
  - Adrien Lèbre (ASCOLA), Yvon Jégou (MYRIADS)
- Efficient exploitation of highly heterogeneous and hierarchical large-scale systems
  - Olivier Beaumont (CEPAGE), Frédéric Vivien (GRAAL)
- Efficient management of very large volumes of information for data-intensive applications
  - Gabriel Antoniu (KERDATA), Jean-Marc Pierson (ASTRE)
- Completing challenging experiments on Grid'5000
  - Lucas Nussbaum (ALGORILLE), Olivier Richard (MESCAL)
- Modeling Large Scale Systems and Validating their Simulators
  - Martin Quinson (ALGORILLE), Arnaud Legrand (MESCAL)
- Network metrology and traffic characterization
  - Paulo Gonçalves (RESO), Konstantin Avrachenkov (MAESTRO)

Transparent, Safe and Efficient Large Scale Computing

- Leaders
  - Stéphane Genaud (ICPS), Fabrice Huet (OASIS)
- Scientific challenges
  - Demonstrate which software architectural designs and programming models best match modern **large-scale** distributed systems
- Grid'5000 allows to experimentally reproduce characteristics of such systems
  - Network heterogeneity
    - High-latency WAN network links mixed with low-latency LAN
  - Hierarchical architecture
  - Virtualization of resources
- Grid'5000 allows to test
  - Programming Models
    - Combination of models ? New paradigms?
  - Middleware
    - Which abstractions for runtime libraries or users?
  - Complex Deployment
    - Workflows, code coupling, web services

Energy Efficient Large Scale Experimental Distributed Systems

- Leaders
  - Laurent Lefèvre (RESO), Jean-Marc Menaud (ASCOLA)
- Objective
  - Energy aware software approaches able to reduce the energy consumption needed for high performance computing and networking operations in large scale distributed systems (datacenters, Grids and Clouds)
- Working on three levels
  - Hardware
  - Infrastructure
  - Application
- Roadmap
  - **JTE «Aspects énergétiques du calcul» : 13/01/2011**
    - **Supported by Héméra**
  - **JTE «Energie dans les centres de données» : Juin/2011**

  - **SLA Energy / Cloud**
  - **Algorithms**
  - **Infrastructure/probe**

Bring Grids Power to Internet-Users thanks to Virtualization Technologies

- Leaders
  - Adrien Lèbre (ASCOLA), Yvon Jégou (MYRIADS)
- Context
  - Job schedulers
  - Exploit all VM capabilities
- Objectives
  - Cluster/Grid-Wide Context Switch
    - Manipulate vJobs (a job in VMs) instead of jobs
  - From the Grid to the Desktop
- Animation
  - Wiki page (2009), mailing list, JTE, …

Efficient exploitation of highly heterogeneous and hierarchical large-scale systems

- Leaders
  - Olivier Beaumont (CEPAGE), Frédéric Vivien (GRAAL)
- Potential research themes
  - Mapping of data and computations
  - (potentially with replication)
  - Resource management
  - Load-balancing
  - Scheduling in probabilistic contexts
  - (uncertainties, failures, etc.)
  - Distributed scheduling
  - Communication- and memory-aware scheduling
  - Platform modeling (mainly, use of)

Efficient management of very large volumes of information for data-intensive applications

- Leaders
  - Gabriel Antoniu (KERDATA), Jean-Marc Pierson (ASTRE)
- Objectives
  - Explore research issues related to high-level services for information management
    - Search, mining, visualization, processing)
  - For large volumes of distributed data
  - Taking into account
    - Security, efficiency and heterogeneity
    - Applications requirements
    - Execution infrastructure (grids, clouds)
- Issues
  - Fault-tolerance, caching, transport, security (encryption, confidentiality), consistency, location transparency
  - Interoperability among storage systems; Data indexing
  - Data mining, data classification, data assimilation, knowledge extraction, data visualization; Metadata management
- Communities involved
  - Distributed applications; Distributed systems; Clusters, grids, P2P, clouds; Fault-tolerant systems; Databases, data mining; Security; Numerical algorithms

Completing Challenging Experiments on Grid'5000

- Leaders
  - Lucas Nussbaum (ALGORILLE), Olivier Richard (MESCAL)
- Spin off the 'Orchestration' scientific challenge
- Axis of work
  - Methodology of the experimentation
    - Scenarios, experimental conditions, metrics, "cahier de laboratoire"
  - Tools for the experimentation
    - Increasing the confidence in experimental results
      - DSL?
- In conjunction with SimGrid

Modeling Large Scale Systems and Validating their Simulators

- Leaders
  - Martin Quinson (ALGORILLE), Arnaud Legrand (MESCAL)
- Context
  - Many studies rely on simulations
    - Easy to set up   Reproducible   Controlled   Enable exploration
    - Fast   Cheap   Not disruptive
  - Unfortunately models in most simulators are either simplistic, not assessed, or even plainly wrong.
- Challenges
  - Models need to be realistic, instantiable, and computationally tractable.
- Outcome
  - Better simulators with standard benchmark platforms
  - Better understanding of resources, applications, and platform
  - Interactions with other working groups regarding methodology (design of experiments, visualization, workload modeling, . . .)

Outline of the talk

- Some data about Grid’5000
- Overview of Héméra
- Scientific Challenges of Héméra
- Working Groups of Héméra
- **Conclusion**

Conclusion

- **Experimental platforms (and observation instruments) are essential** in the computer science methodology - like in other sciences!
- Many research kinds may need G5K
  - HPC, Grids (Classical/Desktop), Clouds, Distributed, Green, etc
    - Network, System, Middleware, Applications
  - A validation tool for applications/middleware before going to production
- Still a difficult tool to use
  - Interface: A lot of know-how
  - Methodology: Not yet recognized benefits
- Hemera
  - Target to solve challenges & to structure the French community
    - 23 teams in 11 cities
    - 8 scientific challenges
    - 8 working groups