Héméra
INRIA Large Scale Initiative
(INRIA Project Lab)

https://www.grid5000.fr/Hemera

Christian Perez
AVALON
INRIA, France

Agenda

10:00-10:15. *Welcome*. C. Perez (Avalon)

10:15-10:45. *La plate-forme Grid'5000*. F. Desprez (Avalon)

10:45-11:15. *Supporting Experimental Science in Distributed Systems Research*. L. Nussbaum (Algorille)

11:15-12:00. *Presentation et Bilan des 2 ans d'Héméra*. C. Perez (Avalon)

12h00-13h30. Déjeuner

13:30 – 13:55. *Large Scale Management of Virtual Machine*. A. Lèbre (Ascola)

13:55 – 14:20. *Scalable Data Management*. G. Antoniu (Kerdata)

14:20 – 14:45. *Solving Large Scale Combinatorial Optimization Problems*. B. Derbel (Dolphin)

14:45 – 15:10. *Modeling Large Scale Systems and Validating their Simulators*. M. Quinson (Algorille)

15:10 – 15:35. *Efficient large electromagnetic problem solving on Grid'5000*. T. Monteil (LAAS)

15:35 – 16:00. Pause

16:00 – 17:00. *Discussion*

Outline of the talk

- Overview of Héméra
- Scientific Challenges of Héméra
- Working groups of Héméra
- Héméra Resources
- Conclusion

Motivations

- Scientific issues
  - Large scale, volatile, complex systems
    - Performance, fault tolerance, scalability, data storage, programming models, algorithms, resource management, etc.
    - Methodological challenges
- Positioning
  - Mathematics,
  - Simulation
  - Emulation
  - ***Experimental testbed (Grid'5000)***
  - Production environment

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)

Organizing the (French) Community

- No structured community to exchange around Grid5000
- GDR ASR
  - Broader scope than Grid5000
  - C. Perez invited to be a member of the scientific committee
- Production Grid
  - Join call research grid/production grid
  - Handled by Aladdin

Hemera: 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. AVALON - Algorithms and Software Architectures for Service Oriented Platforms
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. GRAND-LARGE - Global parallel and distributed computing
11. ICPS - Scientific Parallel Computing and Imaging
12. KERDATA - Cloud and Grid Storage for Very Large Distributed Data
13. OASIS - Active objects, semantics, Internet and security
14. *MAESTRO - Models for the performance analysis and the control of networks*
15. MESCAL - Middleware efficiently scalable
16. MINC - MIcro et Nanosystèmes pour les Communications sans fils
17. MYRIADS - Design and Implementation of Autonomous Distributed Systems
18. REGAL - Large-Scale Distributed Systems and Applications
19. *RESO - Protocols and softwares for very high-performance network*
20. ROMA - Resource Optimization: Models, Algorithms, and scheduling
21. RUNTIME - Efficient runtime systems for parallel architectures
22. SAGE - Simulations and Algorithms on Grids for Environment
23. SARA - Services and Architectures for Advanced Networks
24. SEPIA - Système d'exploitation, systèmes répartis, de l'intergiciel à l'architecture
25. ZENITH - Scientific Data Management

Hemera: 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. **AVALON - Algorithms and Software Architectures for Service Oriented Platforms**
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. GRAND-LARGE - Global parallel and distributed computing
11. ICPS - Scientific Parallel Computing and Imaging
12. **KERDATA - Cloud and Grid Storage for Very Large Distributed Data**
13. OASIS - Active objects, semantics, Internet and security
14. MAESTRO - Models for the performance analysis and the control of networks
15. **MESCAL - Middleware efficiently scalable**
16. **MINC - MIcro et Nanosystèmes pour les Communications sans fils**
17. **MYRIADS - Design and Implementation of Autonomous Distributed Systems**
18. REGAL - Large-Scale Distributed Systems and Applications
19. RESO - Protocols and softwares for very high-performance network
20. ROMA - Resource Optimization: Models, Algorithms, and scheduling
21. RUNTIME - Efficient runtime systems for parallel architectures
22. **SAGE - Simulations and Algorithms on Grids for Environment**
23. **SARA - Services and Architectures for Advanced Networks**
24. **SEPIA - Système d'exploitation, systèmes répartis, de l'intergiciel à l'architecture**
25. **ZENITH - Scientific Data Management**

Hemera: Current Participant List

Hemera: Organization

- A direction committee
  - Aladdin comdir + C. Perez
  - Defines research directions around the Grid5000 testbed
  - Select & evaluate scientific challenges
  - Select & evaluate the working groups
- Scientific challenges
  - A large-scale “experiment” on Grid5000
- Working groups
  - Identified set of teams dealing with scientific challenges

Challenges

List of 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
  - OpenStack on Grid'5000
  - *Large Scale Virtual Machine Deployment & Management*
  - Virtual Machine Live Migration
- Programming Paradigm
  - *Large Scale Computing for Combinatorial Optimization Problems*
  - *Scalable Distributed Processing Using the MapReduce Paradigm*
  - Low Level Component Model Enabling Performance Portability of HPC Application
- Application 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)
- Issues
  - Metrology based decision for Future Internet
    - Admission control, bandwidth sharing, differentiated flow treatment, congestion avoidance, etc
- Objective
  - Improve traffic analysis and characterization thanks to semantic networking and performance prediction
- Roadmap
  - Identify relevant features of traffic data to differentiate them
  - Validate the metrology testbed under real lived traffics
    - Need a full control experimental testbed

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

OpenStack on Grid'5000

Large Scale Virtual Machine Deployment & Management

Virtual Machine Live Migration

Energy Profiling of Large Scale Applications (Energy)

- Leaders
  - Laurent Lefèvre (AVALON), 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

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)

Energy: Some Results

- *Energy profiling and green leverages for services and applications in large scale distributed systems*
  - Joint PhD between INRIA AVALON (Lyon) – IRIT (Toulouse)
  - Started January 2011
- Some results
  - G. Tsafack, L. Lefevre, J.-M Pierson, P. Stolf, and G. Da Costa. "**A runtime framework for energy efficient HPC systems without a priori knowledge of applications**", ICPADS 2012 : 18th International Conference on Parallel and Distributed Systems, Singapore, December 2012
  - G.Tsafack, L. Lefevre, J.-. Pierson, P. Stolf, and G. Da Costa. **"Beyond CPU Frequency Scaling for a Fine-grained Energy Control of HPC Systems**", SBAC-PAD 2012 : 24th International Symposium on Computer Architecture and High Performance Computing, New York City, USA, October 2012
  - G.Tsafack, L. Lefevre, J.-M. Pierson, P. Stolf, and G. Da Costa. "**DNA-inspired Scheme for Building the Energy Profile of HPC Systems**", 1st Intl Workshop on Energy-Efficient Data Centres, Madrid, Spain, May 2012.

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)
    - Impact infrastructure integrity
- Objective
  - Maintain the platform connectivity in presence of high churn
- Roadmap
  - Develop a formal model to characterize the dynamics
    - Validate the model with traces (Skype, eDonkey, PlatenetLab, G5K)
  - Design algorithms for basic blocks of distributed systems on a churn-resilient overlay
    - Leader election, consensus, resource allocation, data storage
  - Experiments these algorithms on G5K
    - Tools for injecting dynamicity

Orchestrating Experiments on the gLite Production Grid Middleware (Orchestration)

- Leaders
  - Lucas Nussbaum (ALGORILLE), Frédéric Suter (AVALON/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

Orchestrating Experiments on the gLite Production Grid Middleware (Orchestration)

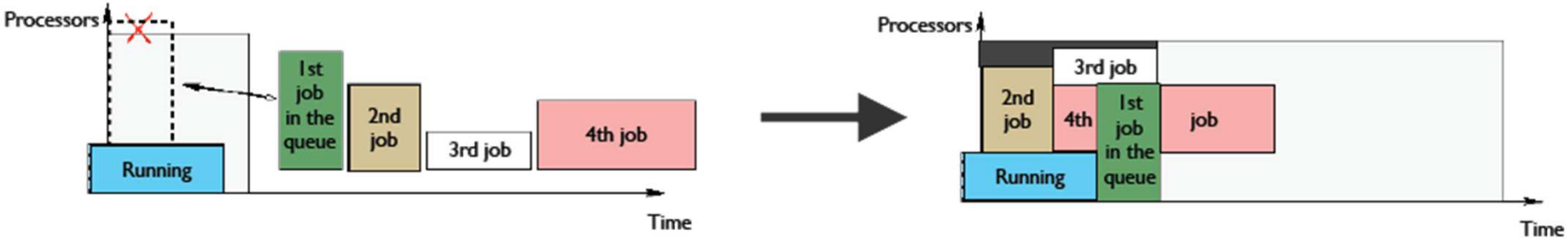
- Work done
  - Designed a set of tools to instantiate a gLite Grid on Grid'5000
  - Best poster award for at Rencontres France Grilles 2011
- Ongoing work to further automate this deployment using an experiment orchestration framework (with Tomasz Buchert)

OpenStack on Grid'5000

- Leader
  - Lucas Nussbaum (Algorille)
- Issue
  - Automatize the deployment of OpenStack on Grid’5000
- Results (S. Badia)
  - Designed a set of tools to instantiate an OpenStack cloud on Grid'5000
  - Included in g5k-campaign (standard set of deployment scripts)
  - Featured in a tutorial at Grid'5000 School 2012
  - Already used by an Inria startup (Harmonic Pharma)
    - Evaluate opportunities regarding data processing in the Cloud

Dynamic Management of Virtual Machines (DynVM)

- Leader
  - Adrien Lebre (ASCOLA)
- Issue
  - Dynamic allocation of resources on a large amount of virtual machines
- Objective
  - Optimization of resource exploitation
  - Nation-wide management of virtual machines over Grid'5000
  - Use of advanced mechanisms to satisfy scheduling criterion (load balancing, consolidation, ...)
- Roadmap
  - Perform a series of large scale experiments on Grid'5000



Dynamic Management of Virtual Machines (DynVM)

- Experiment
  - Set of a nation-wide virtual network among 5 sites on 11 clusters (KaVLAN)
  - Simulation of charge by injecting load in the VM
  - Implementation of a global scenario
- Results
  - Management of **10000+** Virtual Machines on 512 Physical Machines
    - VM can be moved on another site using live migration capabilities
  - Implementation of portable and re-usable scripts
    - Resource reservation
    - Management of a large virtual network
    - Communication between the nodes

FLAUNCHER

Virtual Machine Live Migration

- Leader
  - J.Rouzaud-Cornabas (AVALON)
- Issues
  - Time for live migration depends on several factors
    - VM stress, host hardware, network performance
  - Performance degradation of the migrated VM and the collocated VMs (source and destination)
  - Erroneous VM placement and reconfiguration plan
- Objectives
  - Provide a performance model for Live Migration for the SimGrid Cloud Simulator (ANR INFRA SONGS)
  - Integrate the performance model in VM placement and migration algorithms to increase avoid costly migration plans
  - Validate the performance model thanks to large scale experiments involving thousands of VMs

Virtual Machine Live Migration

- Work done
  - Automatic live migration measurements on different clusters
  - Several migration strategies: parallel, sequential, crossed, one-to-many, many-to-one
  - Stress the VM (network, memory, CPU) during migration
  - Analysis of preliminary results to formalize a first model

- Future works
  - Run large scale experiments spanning multiple sites
  - Use large scale experiments to validate the simulator and the performance model
  - Evaluate the impact of micro-architectural components on live migration

Challenges on Programming Paradigms

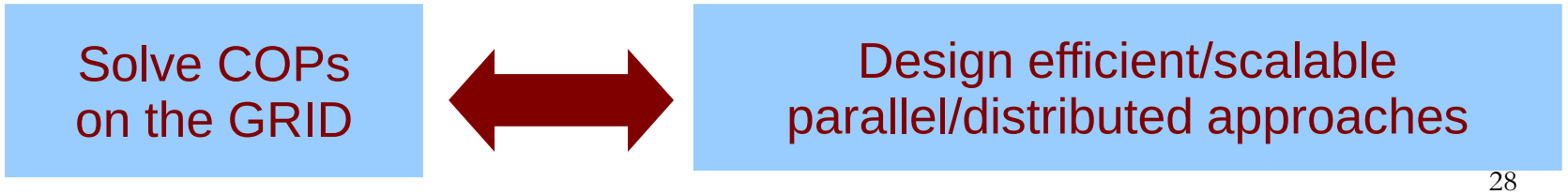
Large Scale Computing for Combinatorial Optimization Problems

Scalable Distributed Processing Using the MapReduce Paradigm

Low Level Component Model Enabling Performance Portability of HPC Applications

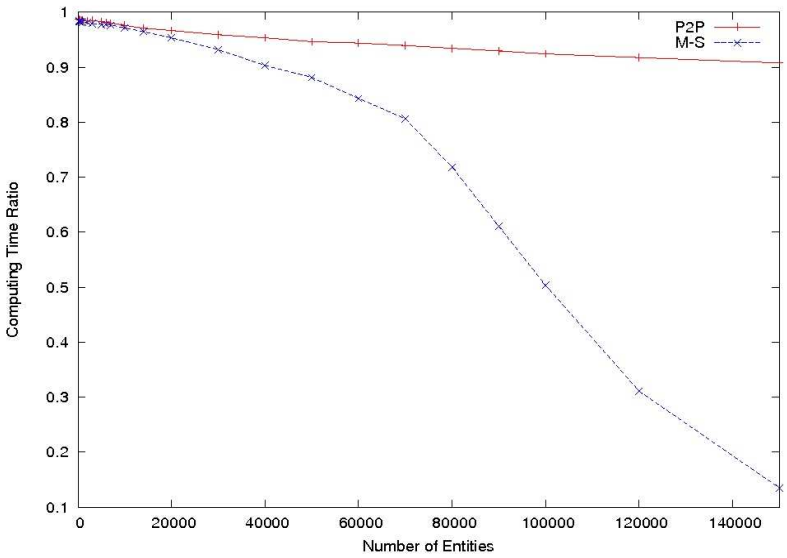
Large Scale Computing for Combinatorial Optimization Problems (COPs)

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



Large Scale B&B

- A new fully distributed B&B
  - Validated using a Pastry-like overlay and up to 150.000 processes
  - In : [IPDPS-LSPP'11, CCGRID-SCALE'11, BookChapter Wiley]
- Towards a Fault-tolerant Peer-to-peer B&B
  - A hybrid two level approach
    - Distributed work sharing and overlay maintenance
    - Centralized Checkpointing
  - Validated under several fault models
  - In : [Paper in preparation]
- People
  - Bilel Derbel, Nouredine Melab
  - Mathieu Djamaï (PhD MENRT)
  - Asim Ali (postdoc Hemera
  - Mohand Mezmaz (Researcher, Mons)



Scalable Dynamic Load Balancing

- Overlay-centric dynamic load balancing
  - Competitive against (B&B and UTS) :
    - Random Work Stealing
    - Master-Slave
    - Hierachical Master-Slave
    - Up to 1.000 Cores
  - In : [CLUSTER'12]
- Adaptive dynamic load balancing in B&B tree search
  - Near optimal with large scale heterogenous entities (3 clusters with up to 128 CPUs and 20 GPUs)
  - In : [LION'13]
- People
  - Bilel Derbel, Nouredine Melab
  - Trong Tuan Vu (PhD Hemera)
  - Asim Ali (postdoc Hemera)

COPS: Next Step

- Solving large scale COPs instances
  - Distributed and hierarchical checkpointing
  - Hybridizations with other COPs specific approaches to speed up the search (Hybrid B&B)
- HPC solutions for heterogeneous distributed networks and virtualized environment
  - Large scale multi-* distributed nodes
  - Efficient overlay mapping and topology awareness
  - In progress: Emulation of network latency on Grid'5000 using Distem

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

- Leaders
  - Gabriel Antoniu (KERDATA), Gilles Fedak (AVALON)
- 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.

ANR MapReduce

Low Level Component Model Enabling Performance Portability of HPC Application

- Leader
  - Christian Perez (AVALON)
- Objective
  - Evaluate the benefice of using L2C
  - Benchmarks of native and component based versions of a 2D stencil kernel code
- Work done in cooperation with D. Balouek
  - Many experiments on a lot of different kinds of hardware
  - Paper at the Workshop on Multi-Core Computing Systems 2012
  - PRACE 2IP deliverable

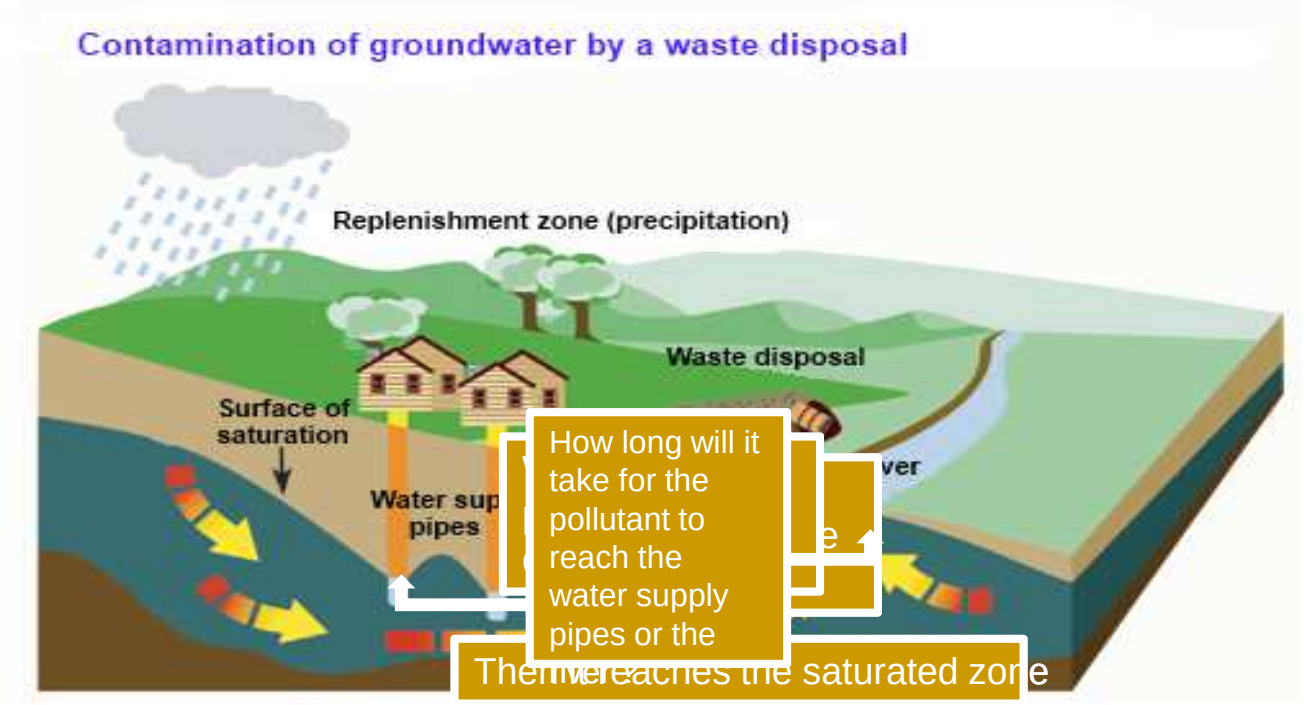
Challenges on Applications

Multi-parametric Intensive Stochastic Simulations for Hydrogeology

Thinking GRID for Electromagnetic Simulation of Oversized Structures

HYDRO challenge: groundwater numerical models

Issues: understand physical and chemical phenomena
Manage water resources, prevent risk of pollution, remedy polluted sites

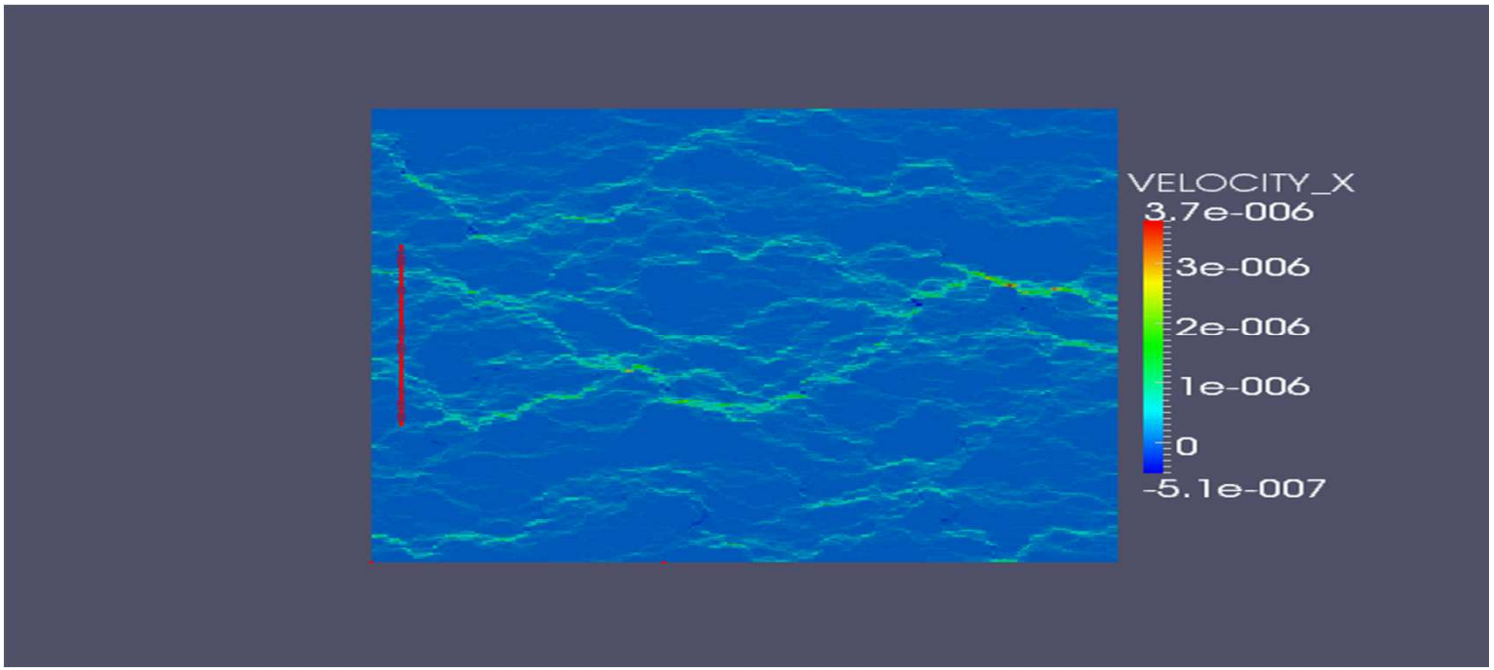


© http://www.ec.gc.ca/eau-water/

A COMPLEX DYNAMIC
Evolution of several phenomena with time: flow, transport, dispersion, chemistry/bio-chemistry,
multiphasic (non saturated zone: air/water, petroleum, gas)

HYDRO challenge: background and objectives

- ANR-projects MICAS (2008-2011), H2MNO4 (2012-2016)
  - H2Olab platform: software PARADIS, MPFRAC, etc
  - Heavy simulations on large domains
  - Multiparametric stochastic simulations in 2D and 3D domains
- HYDRO: efficient computation using multilevel parallelism
  - Leader: team Sage, Inria Rennes (J. Erhel)

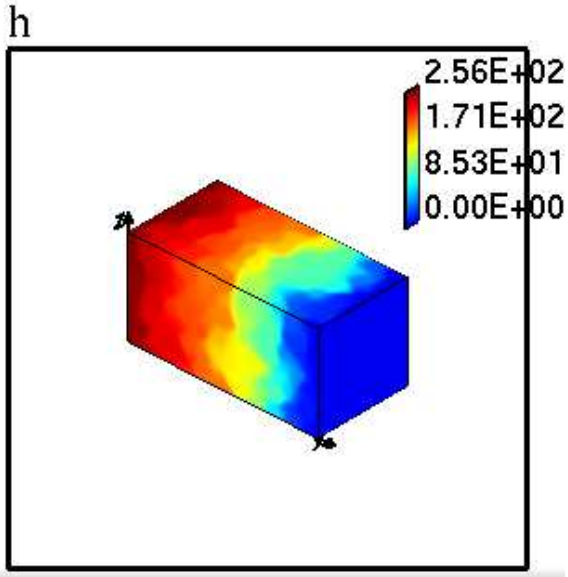


HYDRO challenge: Experiments

- Challenge done by N. Soualem (Sage engineer) and D. Balouek (Hemera engineer)
- Methodology
  - Automatic deployment
  - 200+ runs of multiparametric 2D simulations on Grid'5000
  - From a few minutes up to 60 hours
  - Total experiment time > 390 000 core-hours
  - Analysis of results on relevant metrics
- Results
  - Too much time in generation phase: corrected
  - Efficient parallel multigrid solver and parallel tracker
  - MPI relevant and efficient for parallel Monte Carlo simulations

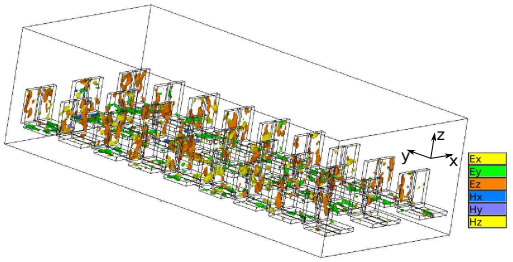
HYDRO challenge: outputs and perspectives

- 3D simulations
  - Production on GENCI machines and experiments on Grid'5000
  - Plenary talk at CANUM'2012
  - Two papers submitted to ESAIM-Proceedings and WRR
  - IHM H2OGuilde (ADT 2012-2013)
  - New project H2MNO4 (ANR MN 2012-2016)
- Numerical parallel algorithms
  - Improving generation phase
  - Improving particle tracker
  - Improving domain partition



Thinking GRID for Electromagnetic Simulation of Oversized Structures (Electro)

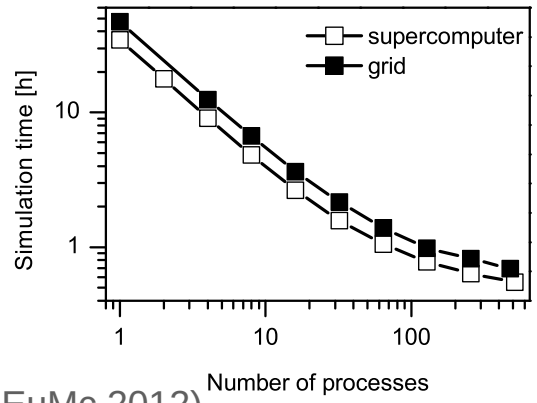
- Leaders
  - Hervé Aubert (MINC-LAAS), Thierry Monteil (MRS-LAAS), Patricia Stolf (SEPIA-IRIT)
- Design of sophisticated communication infrastructures
  - Transmission of signals from airborne sensors
  - Wheel antennas emitting data
    - E.g. tire pressure to a collector located inside a vehicle.



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

Thinking GRID for Electromagnetic Simulation of Oversized Structures (Electro)

- Results
  - Utilization of multithreading and MPI over grid (First measures: D. Balouek, C. Ruiz)
  - Collaboration between application, middleware and platform: experimentation
  - management (Under Work: C. Ruiz Thesis)
  - Uses of autonomic policies (To be done):
    - 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
  - Theoretical estimation of speedup for oversized problem (Done)
- Challenge objectives
  - Autonomic deployment and reconfiguration on grid5000
  - Parallel algorithm for electromagnetic simulation
- Héméra objectives
  - Large scale experiment
  - Experiment and support for electromagnetic researchers
- Production
  - Publications in the 2 best conferences on microwaves (IMS2012, EuMc 2012)
  - Creation of a new Software
  - Change of scale: Million to Billion of cells



Working Groups

Hemera: Working Groups

- What
  - A group of people
- Organization
  - Manage by two leaders
- Responsible of
  - Leading the working group and its community
  - Organizing workshops
  - Potentially proposing the organization of schools

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 (AVALON), 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), ric Vivien (ROMA)
- Efficient management of very large volumes of information for data-intensive applications
  - Gabriel Antoniu (KERDATA), Jean-Marc Pierson (SEPIA)
- 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

Stop

- 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

Continue

- Leaders
  - Laurent Lefèvre (AVALON), 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

Energy Efficient Large Scale Experimental Distributed Systems (Green)

Continue

- GreenDays events
  - GreenDays@Paris (May 31- June 1st, 2011)
    - Energy efficiency : how to apply our models and solutions in production infrastructures ?
  - GreenDays@Lyon (January 19-20, 2012)
    - Energy efficiency : how to monitor and impact on applications ?
  - GreenDays@Luxembourg (January 28-29, 2013)
    - Energy efficiency : "what else / what next ?
  - GreenDays@Lille scheduled in 2014
- COST Action IC804 on the energy efficiency topic: participation & animation
- Instrumentation of Grid'5000 with energy sensors
- Submitted ANR INFRA project in 2013
  - Energy Proportional Data Center for Clouds and HPC

Bring Grids Power to Internet-Users thanks to Virtualization Technologies **Continue**

- 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

Bring Grids Power to Internet-Users thanks to Virtualization Technologies **Continue**

- Meetings
  - Two working meetings (half-day) in 2011 and 2012
- "Tooling"
  - Several scripts to deploy cloudkit upon Grid'5000 (OpenNebula/Nimbus, CloudStack, OpenStack in progress).
    - From April 2012 to Dec 2012, with the support of EIT ICT Lab
    - Alexandra Carpen Amarie (PhD, Engineer MYRIADS)
    - Practical session during the Grid'5000 winter school
- Dissemintation
  - Flauncher and IaaS deployment tool have been presented in the latest publication of the Grid'5000 architecture (INRIA Research Report ''Adding virtualization capabilities to Grid'5000'')
- Two challenges + one new in 2013 (Myriads/Y. Jegou)
  - Large-scale deployment and exploitation of VEP & Snooze

Large Scale Management of Virtual Machines

- `Flauncher`
  - Deploy/configure/launch as many VMs as needed upon Grid'5000
- From Dec 2011 to May 2012
  - Involved people
    - Daniel Balouek, Engineer (HEMERA)
    - Flavien Quesnel, PhD student (ASCOLA Project Team)
    - Adrien Lèbre, Ass. Professor (ASCOLA Project Team)
  - A complete tutorial available on the Grid'5000
  - Used to evaluate the DVMS proposal [CPE 2012], on going submission at HPDC 2013
  - Live demonstration during the Grid'5000 winter school (1000 VMs / 135 PMs in 10 minutes)
- Active projects
  - Extend `Flauncher` to execute/control advanced workflows (including stress slots, migrations operations, ...)
    - Laurent Pouilloux, Engineer (HEMERA)
    - Jonathan Rouzaud Cornabas, Postdoc (AVALON Team)

Efficient exploitation of highly heterogeneous and hierarchical large-scale systems

Reformatted

- Leaders
  - Olivier Beaumont (CEPAGE), Frédéric Vivien (ROMA)
- 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

Continue

- Leaders
  - Gabriel Antoniu (KERDATA), Jean-Marc Pierson (SEPIA)
- 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

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

- Organization of workshops around MapReduce/Data management
  - MapReduce@HPDC (2011, 2012)
  - ScienceClouds@HPDC (2012,2013)
  - VTDC@HPDC (2012, 2013)
  - BDMC (Big Data Management in the Cloud) @Euro-Par 2012
- Post-doc of S. Ibrahim (KerData)
- ANR MapReduce on the MapReduce challenge
- Grid'5000 in European project MCITN SCALUS
  - Bring PhD studend to the Grid'5000 School.
- Need to improve collaborations with Zenith

Completing Challenging Experiments on Grid'5000

Continue

- 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

Completing Challenging Experiments on Grid'5000 Continue

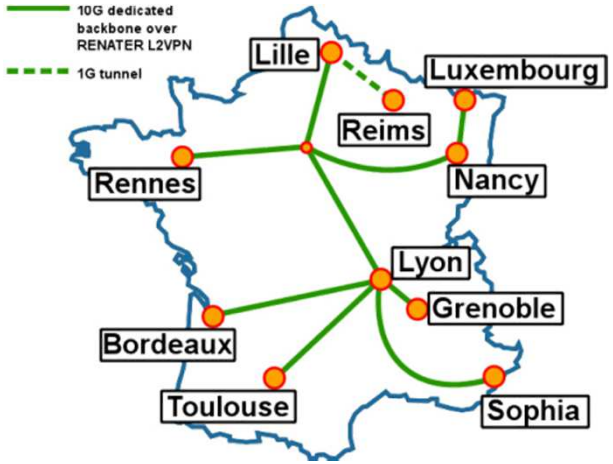
- Animation
  - BOF during Grid'5000 school 2011: best practices for experiments on Grid'5000
  - "Grid'5000 Gotchas" wiki page to list surprising testbed features
- INRIA ADT SOLFEGE
  - Development of Distem (emulator)
  - Tutorial during G5K School 2012
- Two ongoing PhDs
  - Cristian Ruiz
    - Olivier Richard / MESCAL, Thierry Monteil / LAAS
  - Tomasz Buchert
    - Lucas Nussbaum, Jean Gustedt / ALGORILLE
  - Several one-week visits to reinforce collaborations

Modeling Large Scale Systems and Validating their Simulators

Continue

- 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, . . .)
- ANR SimGrid, ANR SONGS

SimGrid Topology for Grid5000



- People
  - L.Nussbaum (ALGORILLE)
  - L. Pouilloux (IC Hemera)
- Work done
  - Use of Grid5000 API for global topology (network, sites, clusters)
  - Latency measurements of network links
  - Flops measurements for all clusters (M. Imbert)
- Future work
  - Add measurements results to API for links and nodes
  - Perform SimGrid simulations and compare with experiments
  - Generate other topology format ?

Héméra Resources

Resources

- Post-doc/Internships
  - for short term contributions to challenges.
- PhD Students
  - for challenges requiring deeper studies.
- Engineers
  - for helping existing working groups to make experiments on Grid'5000.
    - Fill the gap between the Aladdin ADT and end users
    - Delegated to partners for a short period of time on a well-defined goal
- Missions
  - For organizing technical meetings, workshops

Resources - PhD

- 2010–2013: Energy – Astre/Avalon
  - Energy profiling and green leverages for services and applications in large scale distributed systems
- 2011–2014: COPS – Dolphin
  - Robust Peer-to-Peer Algorithms for Large Scale COPs
- 2011–2014: Experiments – Mescal/Laas
  - Methods and tools for challenging experiments on Grid'5000 : a use case on electromagnetic hybrid simulation
- 2012–2015: Data – Zenith
  - Optimizing a Cloud for Data Mining primitives

Resources: Post-doc

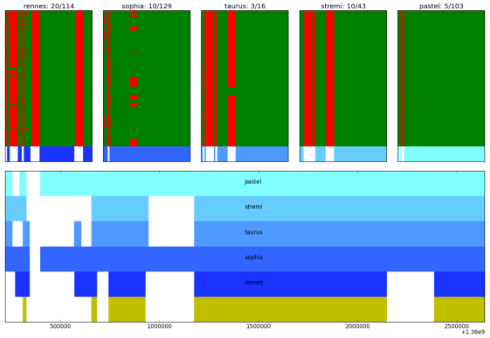
- 2010–2011: COPs – Dolphin
  - Fault-Tolerant Distributed Branch-and-Bound on the Grid
- 2011–2012: Data – Kerdata
  - Consistency, availability, scalability : Building the infrastructure for geo-distributed and geo-replicated cloud storage
- 02/2013–-01/2014*: Energy – Ascola/IRIT
  (*: on 2012 budget)
  - Energy monitoring, from the VM to the room

Hemera Engineers

- People
  - IJD – Daniel Balouek (2/2011 – 1/2013)
  - IC – Laurent Poulloux (10/2012 – 9/2013)
- Link to Grid'5000 Technical Committees
- Mission
  - D. Balouek: managed several challenges
    - Hydro, Electro, DynVM, L2C, Diet
  - L. Pouilloux
    - Help setting up challenges
    - Improve G5K tooling

Developing tools for large scale experiments on Grid5000 with M.Imbert (AVALON Research Engineer)

- Work done: find free slots for your experiments
  - Give some resources combinations (platform, sites, clusters)
  - Compute and draw the Gantt diagram from the API
  - Propose slots for the experiment



- Future work
  - Give a walltime and a minimal number of nodes
  - Compute the number of free nodes over time
  - Let the user choose when to perform the reservation
- Contribution to execo, http://execo.gforge.inria.fr/doc/

Grid'5000 School

- Editions
  - 2006, 2009, 2010, 2011, 2012
  - 100, 72, 80, 67, 69 registered participants
- Content
  - Tutorials
  - Scientific presentations
    - Invited talk
    - Research "papers" (~utilization of G5K)
      - Best paper supported by Hemera in 2012
    - Large Scale Deployment Challenge
      - In 2012, people deploys "easily" thousands of VM

Resources

	2010	2011	2012	2013
Total	**-**	**24000**	**28500**	**32600**
Mission	-	11000	24400	23600
Invest.	-	5000	1096	4000
Fonct.	-	825	1238	5000
Reste	-	~8000	~2000	-

- 2012: ~24 400 €
  - ~9 k€ pour les doctorants/post-doctorants Héméra
  - ~1 k€ pour les ingénieurs Héméra
  - ~5.6 k€ pour des participations à l'École Grid'5000
  - ~5 k€ pour le working group sur l'ordonnancement

Conclusion

- **Experimental platforms (and observation instruments) are essential** in the CS methodology - like in other sciences!
- Many research kinds are using Grid'5000
  - HPC, Grids (Classical/Desktop), Clouds, Distributed, Green, etc
  - A validation tool for applications/middleware before going to production
- Hemera
  - Target to solve challenges & to structure the French community
  - 12 challenges, rather dynamic, some failures, a lot of successes
    - Evolution of theme like cloud, virtualization, big data
  - 8 working groups, different organizations, low support of Hemera
    - Need to rethink them
  - IJD => IC: Evolution of user needs