C2S@Exa Inria Project Lab
Computer and Computational Sciences at Exascale

Journées scientifiques Inria
June 25-26, 2013
Inria Rennes – Bretagne Atlantique

Stéphane Lanteri and several coauthors

Contact: Stephane.Lanteri@inria.fr

ANDRA Mastering radioactive waste

DE LA RECHERCHE À L'INDUSTRIE cea

Inria INVENTEURS DU MONDE NUMÉRIQUE

C2S@Exa Inria Project Lab
Computer and Computational Sciences at Exascale
General context

Computational science (also scientific computing or scientific computation) is concerned with constructing mathematical models and quantitative analysis techniques and using computers to analyze and solve scientific problems (Wikipedia)

In practical use, it is typically the application of computer simulation (numerical simulation) tools to study problems in various scientific disciplines

- Computational fluid dynamics
- Computational electromagnetics
- Computational geoseismics
- Computational chemistry
- Etc.

Typical scenario

Physical phenomena → Mathematical model → Numerical model → Simulation tool

C2S@Exa Inria Project Lab
Computer and Computational Sciences at Exascale
General context: driving applications

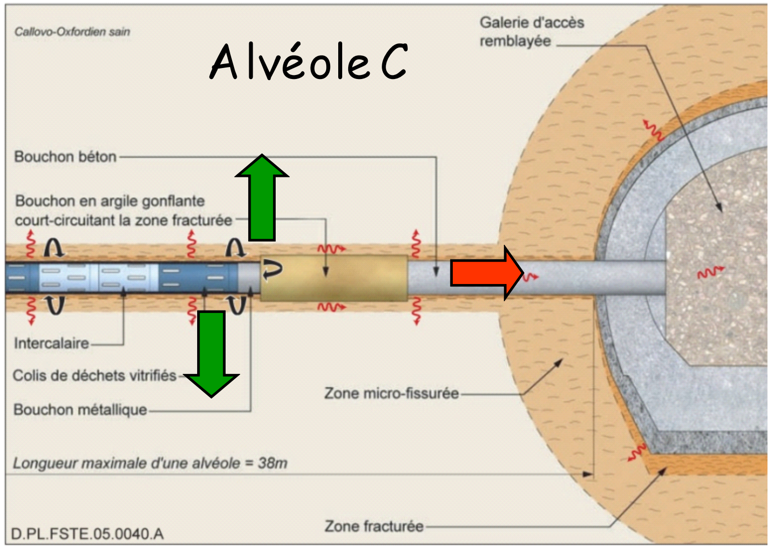
External partner: ANDRA
(Agence Nationale de Traitement des Déchets Radioactifs)
Reactive transport of radionuclide in porous media

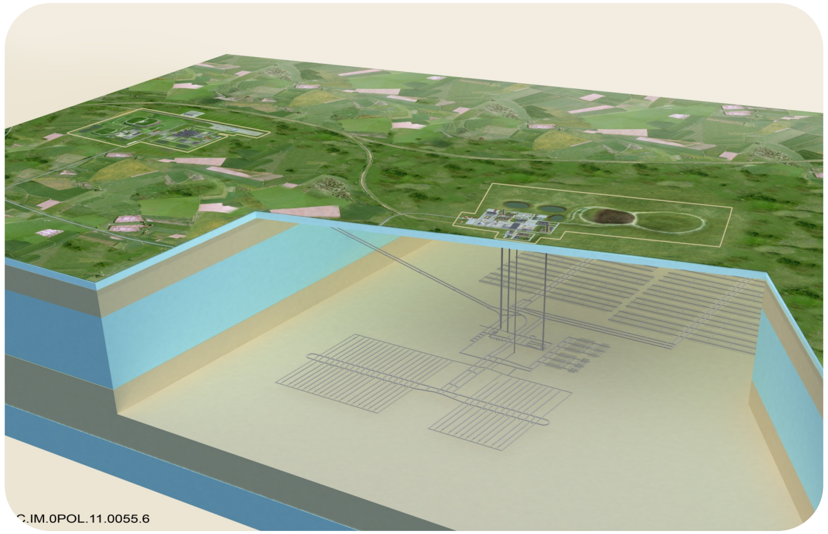
In the field of phenomenological description and performance/safety assessment, ANDRA has to perform many numerical simulations, in particular to quantify flow, from the waste package to the human being, and through the repository and geological environment

Simulations have to take into account many physical processes, applied to different components (from the waste packages to the geological media) and material (clay, concrete, iron, glass, etc.) on very large time (up to one million years) and scale (from centimeter to tens of kilometers)

C2S@Exa Inria Project Lab
Computer and Computational Sciences at Exascale
General context: driving applications

Reactive transport of radionuclide in porous media
Modeling context of radioactive wave management





Hydraulic and migration of solute transfer from waste packages to the geological medium

C2S@Exa Inria Project Lab
Computer and Computational Sciences at Exascale
General context: driving applications

Reactive transport of radionuclide in porous media
Modeling context of radioactive wave management

Ex : deep geological repository

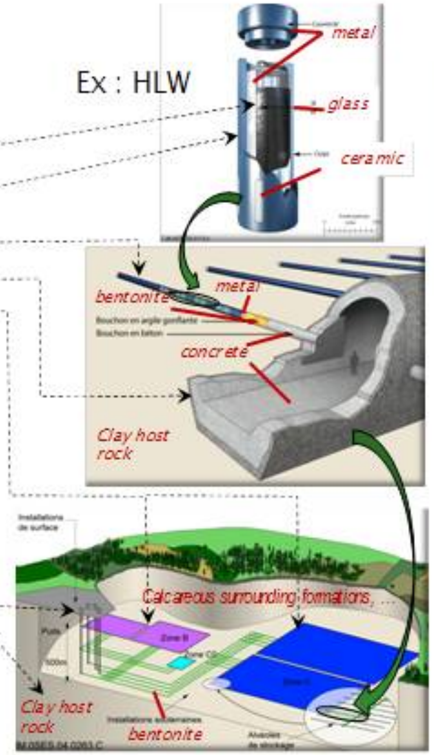
Multi components

Waste — Waste package
Galleries — Disposal cell
Access (shaft…) — Modules and zones
Geological medium

Multi materials

Glass Ceramic Metal
Concrete
Clay based (bentonite)
Clay host formation
calcareous, …

Ex : HLW



Multi processes

Thermal (T), Hydraulical-Gas (H) ,
Mechanical (M), Chemical (C), Radiological (R)
Solute Transfer in porous media (Tr).

with or without couplings (strong to low)

Multi spatial scale

-*centimetric to metric*:	waste
-*metric to decametric*	disposal cell
-*hectometric to kilometric*	disposal
-*pluri-kilometric*	geological medium

→ Management > 7 orders of magnitude

Multi temporal scale

Up to 1 million years

C2S@Exa Inria Project Lab
Computer and Computational Sciences at Exascale
General context: driving applications

Reactive transport of radionuclide in porous media
Classification of radioactive waste

<table>
<tr><th>Half-life
Activity</th><th>Very short life
< 100 days</th><th>Short life (SL)
< 31 years</th><th>Long life (LL)
> 31 years</th></tr>
<tr><td>Very low level (VLL)</td><td rowspan="4">Management through in situ radioactive decay</td><td colspan="2">Surface disposal at CSTFA (Aube District)
Recycling systems</td></tr>
<tr><td>Low level (LL)</td><td rowspan="2">Surface disposal at CSFMA (Aube District) , except for certain tritiated waste and sealed sources</td><td>Under study pursuant to Art. 4 of Planning Act of 28 June 2006.
Project for shallow disposal.</td></tr>
<tr><td>Intermediate level (IL)</td><td>Under study pursuant to Art. 3 of Planning Act of 28 June 2006. Project for a deep reversible repository.</td></tr>
<tr><td>High level (HL)</td><td colspan="2">Under study pursuant to Art. 3 of Planning Act of 28 June 2006. Project for a deep reversible repository.</td></tr>
</table>

C2S@Exa Inria Project Lab
Computer and Computational Sciences at Exascale
General context: driving applications

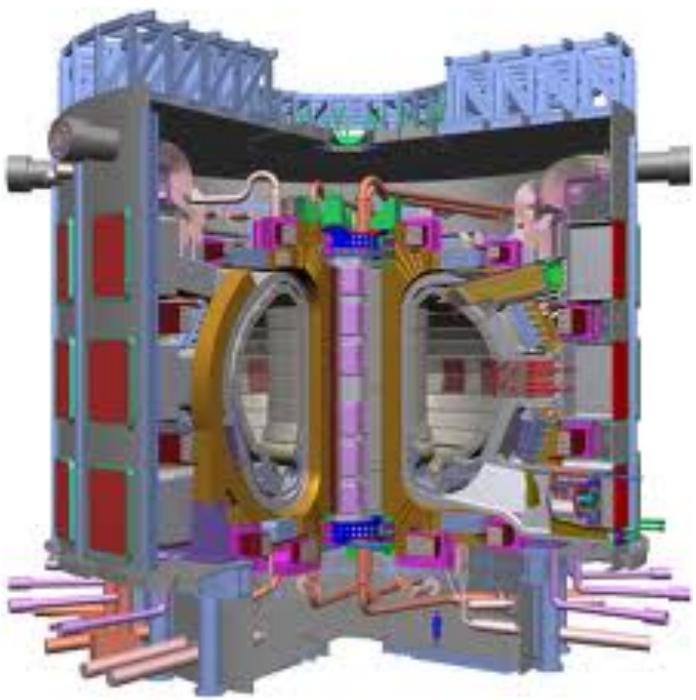
TRACES is a computer program for the simulation of flow and reactive transport in saturated porous media

Two levels of simulations are considered

- ✓ Complex physicochemical processes, involving strong couplings (such as chemical-transport, 2 phase-flow with radionuclide transfer, thermo-hydro-mechanics problems) but solved on small systems with grids up to many thousands of elements
- ✓ Simplified physicochemical processes (leak coupling) consisting of modifying in space and time hydraulic and solute transfer parameters, but solved on big systems with grids of many millions of elements. Global system is split into embedded compartments whose scales are bigger and bigger

C2S@Exa Inria Project Lab
Computer and Computational Sciences at Exascale
General context: driving applications

IRFM (Research Institute on Magnetic Fusion) institute at CEA (French Alternative Energies and Atomic Energy Commission - Cadarache center) is conducting research activities on nuclear fusion in the context of the ITER project with studies that are concerned with Magneto HydroDynamic (MHD) stability, turbulent transport, plasma-wall interaction, and RF heating



C2S@Exa Inria Project Lab
Computer and Computational Sciences at Exascale
General context: driving applications

External partner: CEA/IRFM
(Research Institute on Magnetic Fusion)

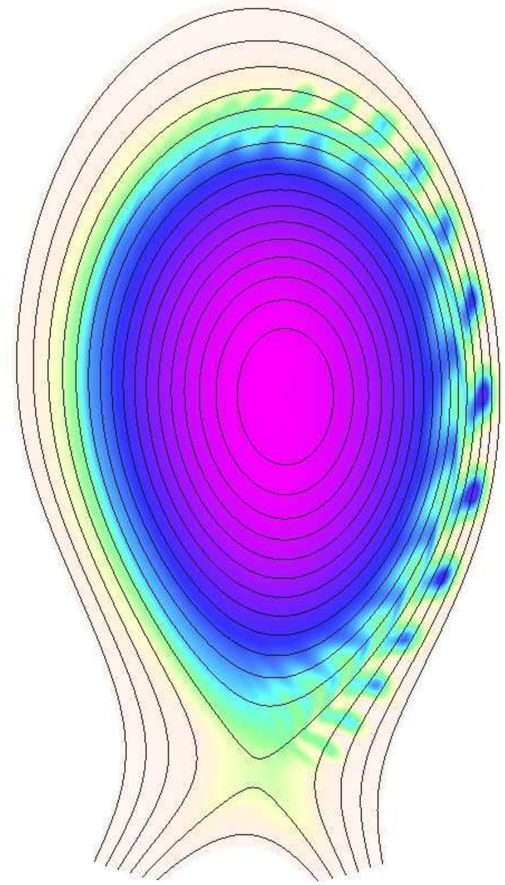
Significant recent progress in simulations of fine-scale turbulence and in large-scale dynamics of magnetically confined plasmas has been enabled by access to terascale supercomputers. These progress would have been unreachable without innovative analytic and computational methods for developing reduced descriptions of physics phenomena

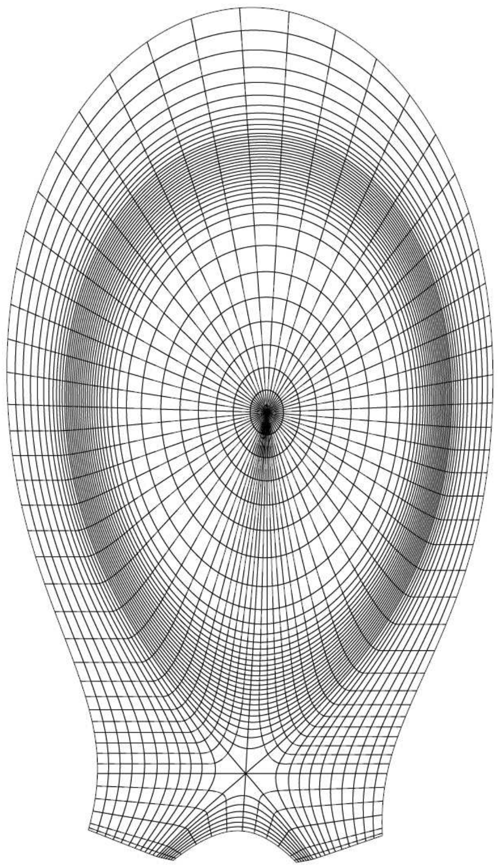
Accelerated progress on this critical issue is especially important for ITER, because the size and cost of a fusion reactor are determined by the balance between 1) loss processes and 2) self-heating rates of the actual fusion reactions. Realistic models, simulations and highly parallel algorithms are essential in dealing with such challenges because of the huge range of temporal and spatial scales involved

C2S@Exa Inria Project Lab
Computer and Computational Sciences at Exascale
General context: driving applications

Two simulations software are considered in C2S@Exa:

- JOREK is dedicated to the numerical study of Edge Localized Modes (ELMs) and disruptions
- GYSELA is a global non-linear electrostatic code which solves the gyrokinetic equations (Vlasov) in a five dimensional phase space with a semi-Lagrangian method

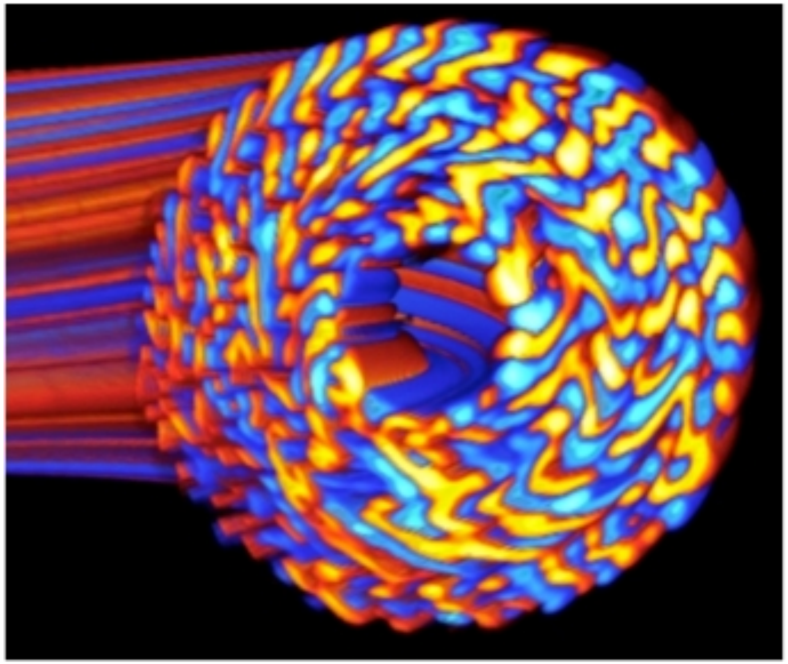




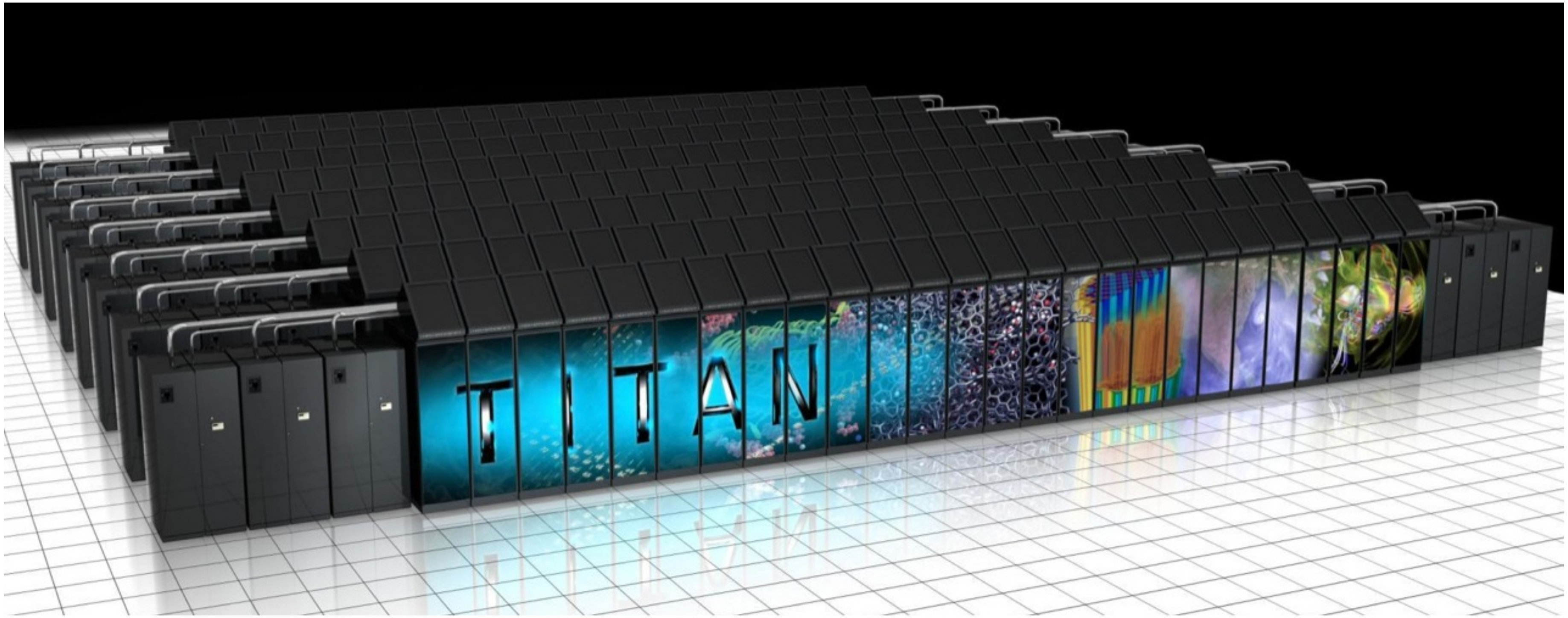
C2S@Exa Inria Project Lab
Computer and Computational Sciences at Exascale
General context: driving applications

Two simulations software are considered in C2S@Exa:

- JOREK is dedicated to the numerical study of Edge Localized Modes (ELMs) and disruptions
- GYSELA is a global non-linear electrostatic code which solves the gyrokinetic equations (Vlasov) in a five dimensional phase space with a semi-Lagrangian method

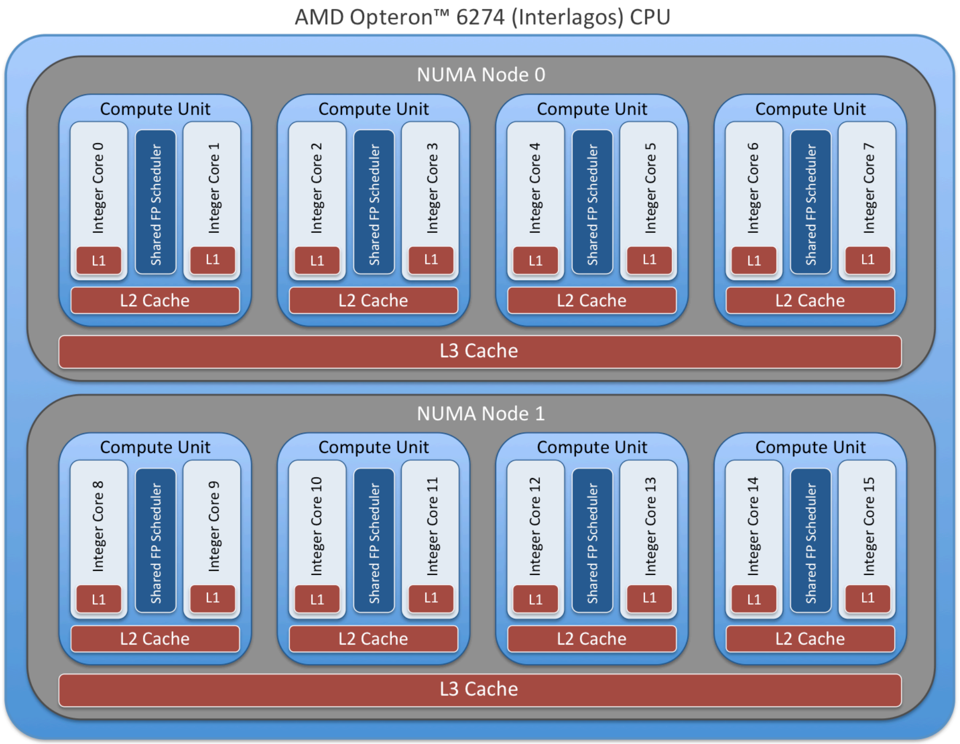


C2S@Exa Inria Project Lab
Computer and Computational Sciences at Exascale
General context: towards exascale computing



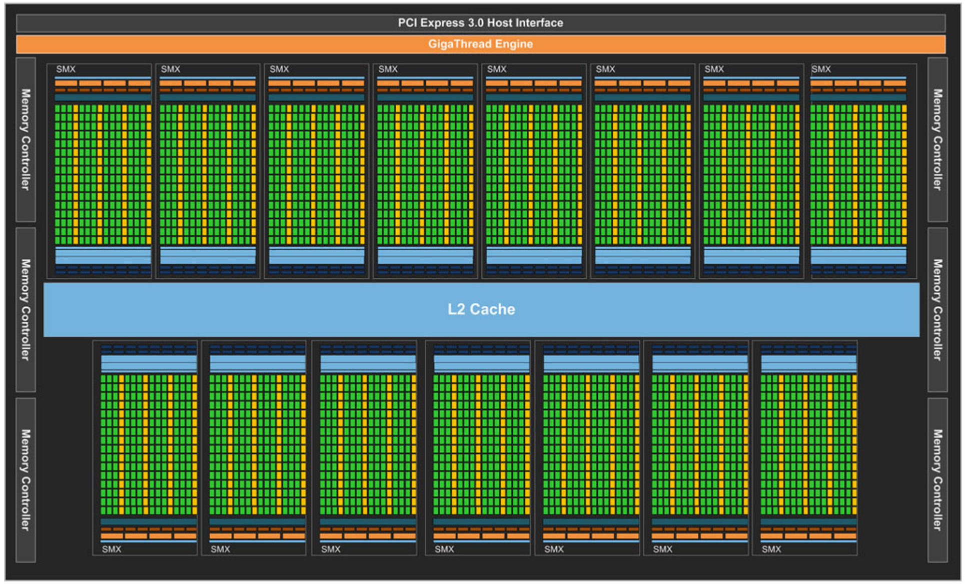
Titan system, Oak Ridge National Laboratory
Cray XK7 , AMD Opteron 6274 16C 2.2 GHz, Cray Gemini interconnect, NVIDIA K20x
560640 cores, 27112.5 TFlop/s peak

C2S@Exa Inria Project Lab
Computer and Computational Sciences at Exascale
General context: towards exascale computing



Titan system, Oak Ridge National Laboratory
Cray XK7 , AMD Opteron 6274 16C 2.2 GHz, Cray Gemini interconnect, NVIDIA K20x
560640 cores, 27112.5 TFlop/s peak

C2S@Exa Inria Project Lab
Computer and Computational Sciences at Exascale
General context: towards exascale computing



	TESLA K10[a]	TESLA K20	TESLA K20X
Peak double precision floating point performance (board)	0.19 teraflops	1.17 teraflops	1.31 teraflops
Peak single precision floating point performance (board)	4.58 teraflops	3.52 teraflops	3.95 teraflops
Number of GPUs	2 x GK104s	1 x GK110	
Number of CUDA cores	2 x 1536	2496	2688
Memory size per board (GDDR5)	8 GB	5 GB	6 GB
Memory bandwidth for board (ECC off)[b]	320 GBytes/sec	208 GBytes/sec	250 GBytes/sec
GPU computing applications	Seismic, image, signal processing, video analytics	CFD, CAE, financial computing, computational chemistry and physics, data analytics, satellite imaging, weather modeling	
Architecture features	SMX	SMX, Dynamic Parallelism, Hyper-Q	
System	Servers only	Servers and Workstations	Servers only

Titan system, Oak Ridge National Laboratory
Cray XK7 , AMD Opteron 6274 16C 2.2 GHz, Cray Gemini interconnect, NVIDIA K20x
560640 cores, 27112.5 TFlop/s peak

C2S@Exa Inria Project Lab
Computer and Computational Sciences at Exascale
General context: towards exascale platforms

In recent computer systems, parallelism spreads over many architecture levels including nodes, processors, cores, threads, registers, SIMD-like and vector units

Several different levels of parallelism (from coarse to fine or very fine grain parallelism) need to be harnessed in order to maximize computational efficiency and scalability

Moreover, heterogeneity of the memory is growing at the node as well as at the chip level

The NUMA penalty in data accesses is one of the main critical issues for parallel performances and one must take care of locality access and memory affinity when distributing data on cores

All these heterogeneous characteristics of hardware resources keep effective performance far from theoretical peak

C2S@Exa Inria Project Lab
Computer and Computational Sciences at Exascale
General context: rationale

Most applications and algorithms are not yet ready to utilize these available architecture capabilities

Developing large-scale scientific computing tools that efficiently exploit this processing power, of the order of petaflops with current generation systems, is a very complicated task and will be an even more challenging one with future exascale systems

Heterogeneity characteristic and hierarchical organization of modern massively parallel computing systems are recognized as central features that impact at all the layers from the hardware to the software with issues related to computer science and numerical mathematics as well

At the current state of the art in technologies and methodologies, a multi-disciplinary approach is required to tackle the obstacles in manycore computing, with contributions from computer science, applied mathematics, high performance computing, and engineering disciplines

C2S@Exa Inria Project Lab
Computer and Computational Sciences at Exascale
General objective and foreseen contributions

Establishment of a continuum of skills in the applied mathematics and computer science fields for a multidisciplinary approach to the development of numerical simulation tools that will take full benefits of the processing capabilities of emerging high performance massively parallel architectures

Activities and contributions are organized along a three-level structure from generic building-blocks to large-scale applications

C2S@Exa Inria Project Lab
Computer and Computational Sciences at Exascale
Organization

Project structure and activities (1/3)
Level 1 – Towards generic and scalable algorithms

✓ **Computer science topics**

Upstream from the core topics which are centered on *the development of high performance numerical schemes and algorithms*

✓ **Algorithmic aspects**

Emphasis on the development of generic numerical libraries and solvers in order to benefit from all the parallelism levels with the main goal of optimal scaling on very large numbers of computing entities

✓ **Robustness, accuracy and scalability issues of numerical schemes**

Generic design issues of high performance numerical schemes for systems of partial differential equations

C2S@Exa Inria Project Lab
Computer and Computational Sciences at Exascale
Organization

Project structure and activities (2/3)
Level 2 – Towards robust, accurate and highly scalable numerical schemes for complex physical problems

Study of the systems of PDEs that model the scientific and engineering use cases considered in the project

Topics of interest include discretization in space of underlying systems of PDEs (high order approximation, adaptivity, etc.), solution algorithms base on continuous models (domain decomposition algorithms, physics based preconditioners, etc.) and numerical methods adapted to multi-scale and multi-physics problems

C2S@Exa Inria Project Lab
Computer and Computational Sciences at Exascale
Organization

Project structure and activities (3/3)
Level 3 – Towards exascale computing for the simulation of frontier problems

Large-scale simulations using high performance numerical computing methodologies resulting from the activities undertaken in the bottom and intermediate levels

With the involvement of external partners from research laboratories or industrial groups that will help in defining and dimensioning a number of frontier problems

C2S@Exa Inria Project Lab
Computer and Computational Sciences at Exascale
Organization: 5 thematic poles

Pole 1: Numerical linear algebra
Core numerical kernels, sparse direct solvers, preconditioned iterative solver and continuous solvers

Pole 2: Numerical schemes for PDE models
Optimize high order finite element and finite volume schemes for maximizing both the single core computational performances and the scalability in view of exploiting massive parallelism

Pole 3: Optimization of performances of numerical solvers
Optimization of the performances of numerical solvers by considering topics that often make a link or are at the interface between computer science techniques and tools for exploiting high performance computing systems and application software

Pole 4: Programming models
Component models and high level programming models

Pole 5: Resilience for exascale computing
Energy effective fault tolerant protocols for HPC applications, efficient fault tolerant protocols and algorithm-based fault tolerance, performance execution models for fault-tolerant applications, resilience for sparse linear algebra

C2S@Exa Inria Project Lab
Computer and Computational Sciences at Exascale

Core project-teams: numerical mathematicians

BACCHUS [INRIA Bordeaux – Sud-Ouest]
Parallel tools for numerical algorithms and resolution of essentially hyperbolic problems

HIEPACS [INRIA Bordeaux – Sud-Ouest]
High-end parallel algorithms for challenging numerical simulations

NACHOS [INRIA Sophia Antipolis – Méditerranée]
Numerical modeling and high performance computing for evolution problems in complex domains and heterogeneous media

SAGE [INRIA Rennes – Bretagne Atlantique]
Simulations and algorithms on Grids for environment

TONUS [INRIA Nancy – Grand-Est]
Tokamak numerical simulations

C2S@Exa Inria Project Lab
Computer and Computational Sciences at Exascale

Core project-teams: computer scientists

ALPINES [INRIA Paris - Rocquencourt]
Algorithms and parallel tools for integrated numerical simulations

AVALON [INRIA Grenoble – Rhône-Alpes]
Large algorithms and software architectures for service oriented platforms

MOAIS [INRIA Grenoble – Rhône-Alpes]
Programming and scheduling design for applications in interactive simulation

ROMA [INRIA Grenoble – Rhône-Alpes]
Resource optimization: models, algorithms, and scheduling

RUNTIME [INRIA Bordeaux – Sud-Ouest]
Efficient runtime systems for parallel architectures

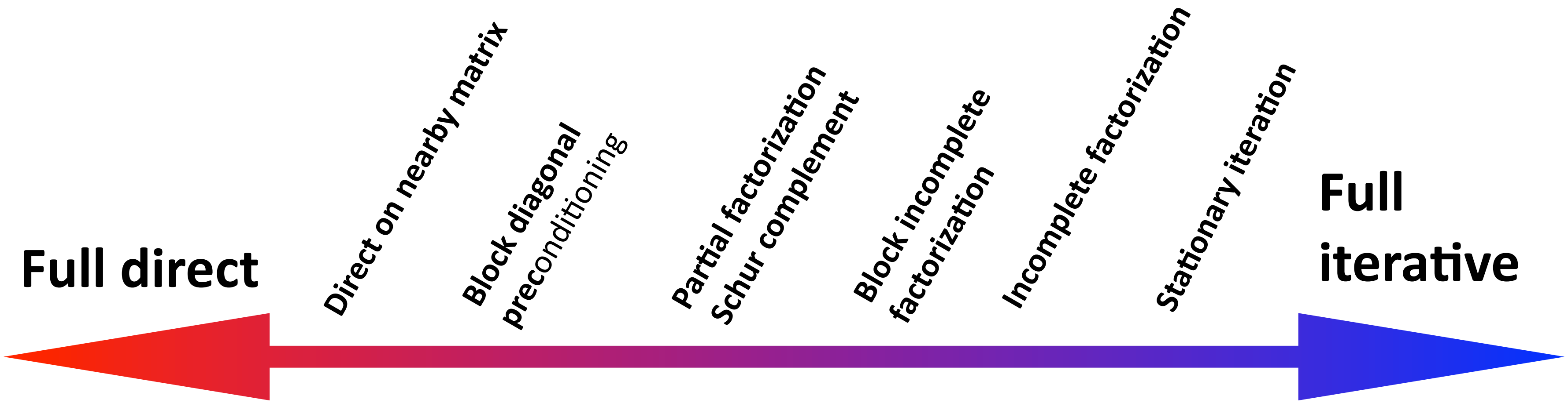
Associated project-teams: numerical mathematicians

CASTOR [INRIA Sophia Antipolis – Méditerranée]
Control, analysis and Simulations for tokamak research

POMDAPI [INRIA Paris - Rocquencourt]
Environmental modeling, optimization and programming models

C2S@Exa Inria Project Lab
Computer and Computational Sciences at Exascale
Focus: numerical linear algebra and fast solvers

Goal: solve $Ax = b$, where A is large and sparse



Direct
- ✓Robust/prescribed accuracy for general problems
- ✓BLAS-3 based implementations
- ✓Memory/CPU prohibitive for large3D problem
- ✓Limited weak scalability

Iterative
- ✓Problem dependent efficiency/monitored accuracy
- ✓Sparse computational kernels
- ✓Less memory requirements and possibly faster
- ✓Potential high weak scalability

C2S@Exa Inria Project Lab
Computer and Computational Sciences at Exascale
Focus: numerical linear algebra and fast solvers

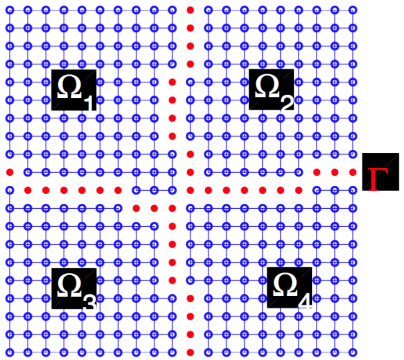
Sparse direct solver (PaStiX)
HiePACS project-team

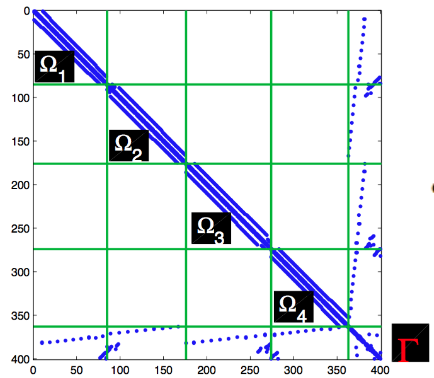
- ✓ Numerical features
  - LLT , LDLT , LU factorization with supernodal implementation
  - Static pivoting + refinement: CG/GMRES
  - 1D/2D block distribution + full BLAS3
  - Simple/double precision + float/complex operations
- ✓ Implementation features
  - MPI/Threads implementation (SMP/Cluster/Multicore/NUMA)
  - Dynamic scheduling inside SMP nodes (static mapping)
  - Support external ordering library (PT-Scotch/METIS)

C2S@Exa Inria Project Lab
Computer and Computational Sciences at Exascale
Focus: numerical linear algebra and fast solvers

Algebraic sparse linear solvers - HiePACS project-team

Partition the adjacency graph of the sparse matrix (Scotch)





$$\mathcal{A}^{(i)} = \begin{pmatrix} \mathcal{A}_{\mathcal{I}_i\mathcal{I}_i} & \mathcal{A}_{\mathcal{I}_i\Gamma_i} \\ \mathcal{A}_{\Gamma_i\mathcal{I}_i} & \mathcal{A}_{\Gamma\Gamma}^{(i)} \end{pmatrix}$$

Local calculation of Schur complements (MUMPS, PaStiX) and precondtioning operator (Magma, MUMPS, PasTiX)

$$\mathcal{S}^{(i)} = \mathcal{A}_{\Gamma\Gamma}^{(i)} - \mathcal{A}_{\Gamma_i\mathcal{I}_i}\mathcal{A}_{\mathcal{I}_i\mathcal{I}_i}^{-1}\mathcal{A}_{\mathcal{I}_i\Gamma_i} \qquad \mathcal{M} = \sum_{i=1}^{N} \mathcal{R}_{\Gamma_i}^{T}(\bar{\mathcal{S}}^{(i)})^{-1}\mathcal{R}_{\Gamma_i}$$

- ✓ Parallel hierarchical implementation: HIPS, MaPHyS
- ✓ Similar approaches: PDSLin (LBNL), ShyLU (Sandia) France Berkeley fund, EA FAST-LA (LBNL, Stanford)

C2S@Exa Inria Project Lab
Computer and Computational Sciences at Exascale
Focus: numerical linear algebra and fast solvers

Algebraic sparse linear solvers - SAGE project-team

DGMRES and AGMRES: adaptive coarse grid in preconditioned GMRES

KSPDGMRES: adaptive deflated GMRES (available in PETSc)

KSPAGMRES: adaptive augmented GMRES (soon available in PETSc)

GPREMS: domain decomposition method, parallel multiplicative Schwarz

SIDNUR: domain decomposition method, Schur complement type with Neumann-Neumann preconditioning and coarse grid correction

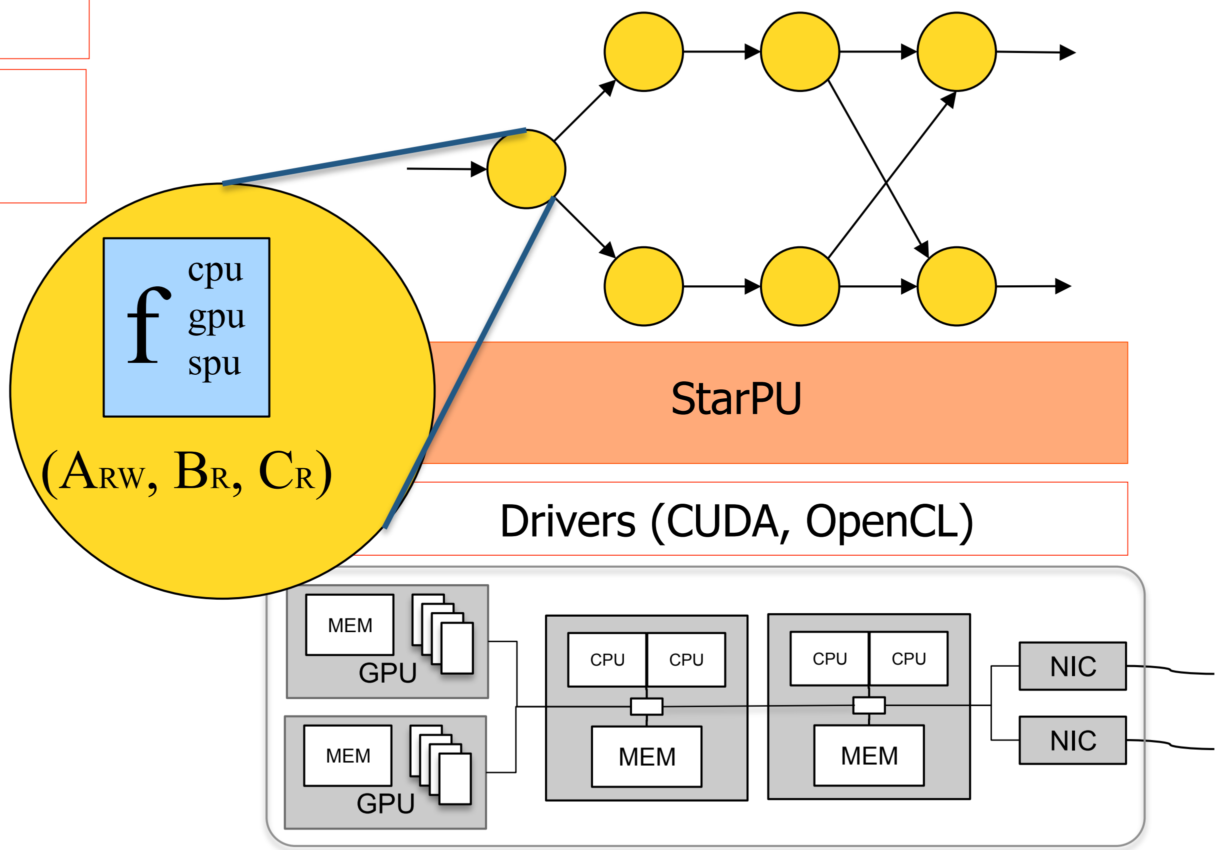
C2S@Exa Inria Project Lab
Computer and Computational Sciences at Exascale
Focus: exploiting heterogeneous architectures

The StarPU runtime system - RUNTIME project-team

HPC Applications

Parallel compilers | Parallel libraries

- ✓ Rationale: Scheduling tasks over heterogeneous machines
  - ✓ CPU + GPU + ...PU = *PU
  - ✓ Auto-tuning of performance models
  - ✓ Optimization of memory transfers, prefetching, overlap
  - ✓ Leverage theoretical scheduling algorithms
  - ✓ OpenCL API available



C2S@Exa Inria Project Lab
Computer and Computational Sciences at Exascale
Focus: exploiting heterogeneous architectures

The StarPU runtime system - RUNTIME project-team

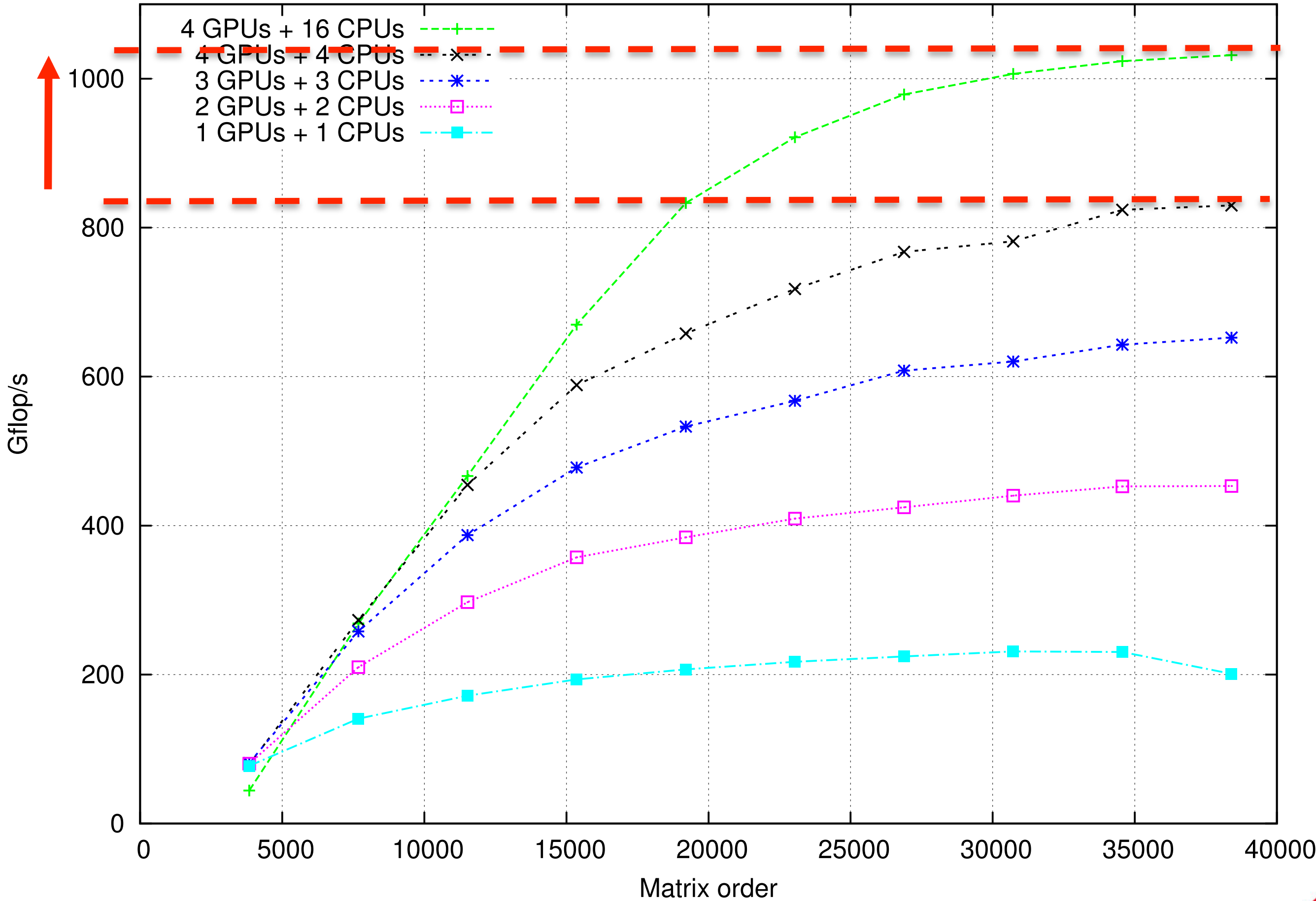
+12 CPUs ~200 Gflops

(although 12 CPUs alone ~150 Gflops)

Dense Linear Algebra Kernels, with UTK & Inria HiePACS

QR decomposition, MAGMA over StarPU

16 CPUs (AMD) + 4 GPUs (C1060)



C2S@Exa Inria Project Lab
Computer and Computational Sciences at Exascale
Prospective: numerical linear algebra and fast solvers

Towards a solver software stack

Hybrid solvers

Sparse direct solvers

FMM
Fast Multipole Method

Dense linear kernels – $\mathcal{H}$ matrices

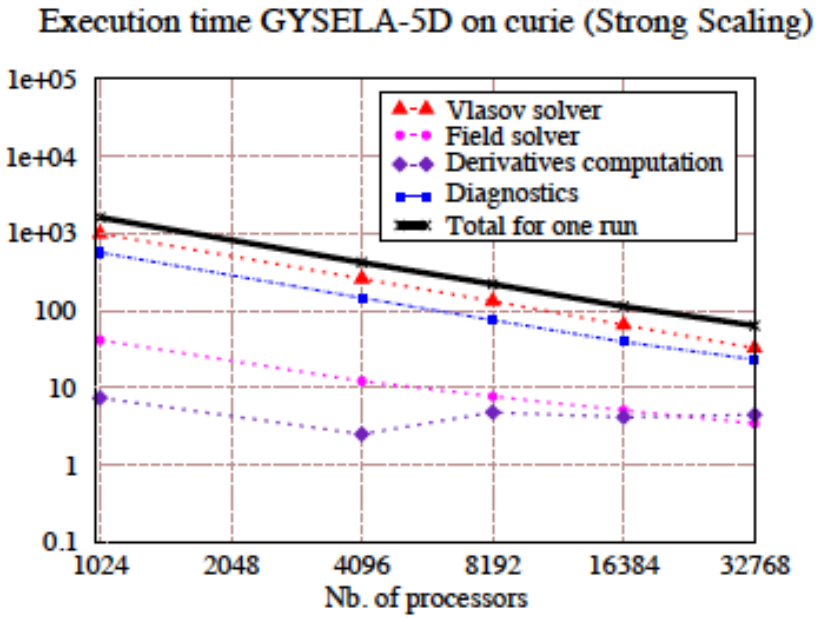
Runtime systems

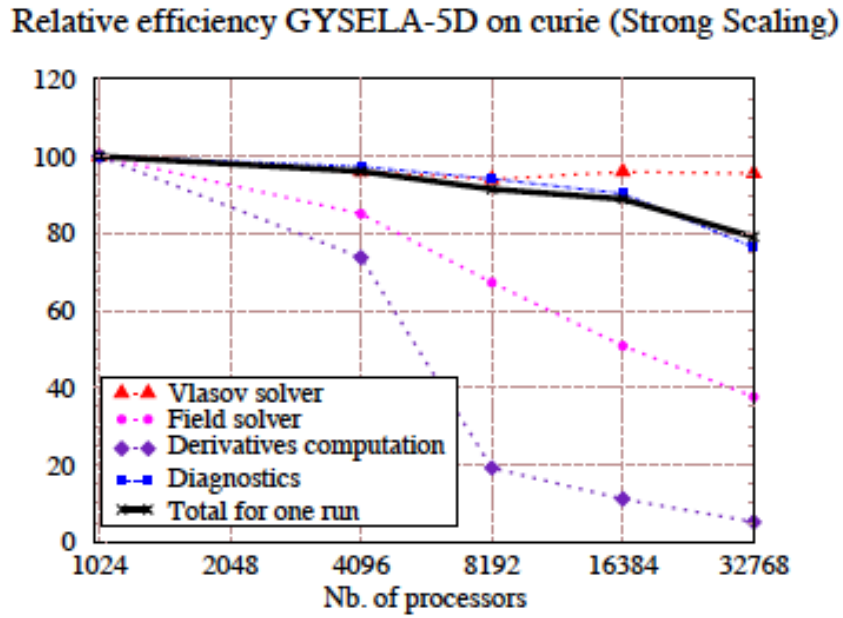
EA FAST-LA, MORSE – EADS/ASTRIUM/CRA

All the stack on parallel sparse linear algebra components

Ongoing ADT: MaPHyS@exa, HPC Collective

C2S@Exa Inria Project Lab
Computer and Computational Sciences at Exascale
Prospective: GYSELA





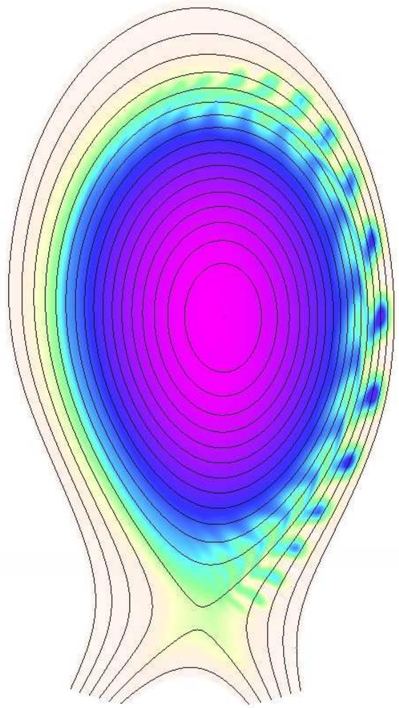
Current parallelization strategy
MPI + OpenMP (80% efficiency on 33k cores)

C2S@Exa objective: improve scalability up to and beyond 100k cores
Requires to deal with applied mathematics (numerical schemes and solvers) and computer science issues (parallel programming model and runtime system for exploiting heterogeneous architectures)

C2S@Exa Inria Project Lab
Computer and Computational Sciences at Exascale
Prospective: JOREK

Nb cores	24	48	96	192
Nb nodes	2	4	8	16
Steps (time in sec.)				
`construct_matrix`	6.9	3.8	2.0	1.3
`factorisation`	33	22	16	12
`gmres/solve`	1.9	1.6	0.8	0.7
`iteration time`	48	32	22	18
`rel. efficiency`	100%	75%	55%	33%

Fusion energy production in ITER requires the achievement of high pressure plasmas in high energy confinement mode (H-mode). This confinement mode is characterized by the formation of very steep plasma pressure profiles at the edge of the plasma that lead to periodic bursts of energy being expelled by the plasma (typically a small percentage of the total plasma energy) called ELMs (Edge Localized Modes)



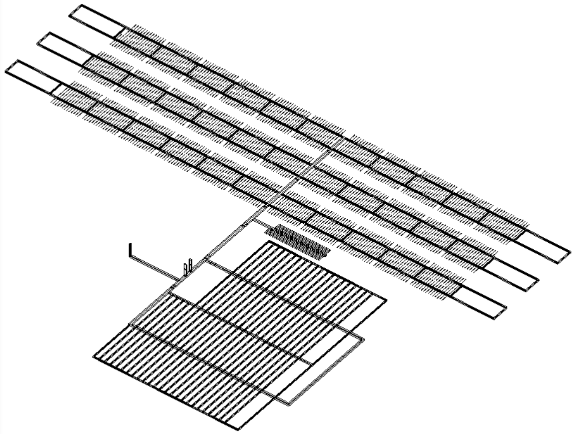
C2S@Exa objective: improve scalability up to and beyond 1k cores
Requires to deal with applied mathematics (numerical schemes and solvers) and computer science issues (parallel programming model and runtime system for exploiting heterogeneous architectures)

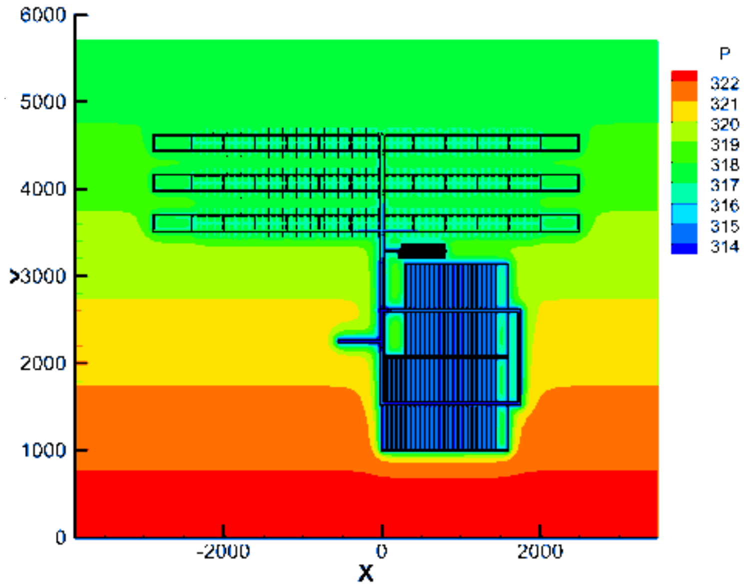
C2S@Exa Inria Project Lab
Computer and Computational Sciences at Exascale
General context: driving applications

Hydraulic and migration of solute transfer from waste packages to the geological medium

Up-to-date calculation

- ✓ 10 millions cells in structured mesh including all disposal cells and zone
- ✓ Many simplifications in the modeling (e.g. only one clay layer)
- ✓ Structured (Cartesian) grid approach
- ✓ High computing time (limit for sensitivity analysis)





C2S@Exa objective: improve scalability up to and beyond 1k cores

Requires to deal with applied mathematics (numerical schemes and solvers) and computer science issues (parallel programming model and runtime system for exploiting heterogeneous architectures)

Inria

C2S@Exa Inria Project Lab
Computer and Computational Sciences at Exascale
Prospective: satellite projects

ICT-2013.12.1 - Exascale computing platforms, software and applications (under negotiation)

- ✓ EXA2CT [IMEC, Belgium]

  EXascale Algorithms and Advanced Computational Techniques

  Brings together experts at the cutting edge of the development of solvers, related algorithmic techniques, and HPC software architects for programming models and communication

- ✓ DEEP-ER [Forschungszentrum Jülich, Germany]

  Dynamical Exascale Entry Platform - Extended Reach

  Addresses two significant Exascale challenges: the growing gap between I/O bandwidth and compute speed, and the need to significantly improve system resiliency

CNPq – INRIA project « HOSCAR » (starting date: 01/01/2012)
High performance cOmputing and SCientific dAta management dRiven by highly demanding applications
Brazilian partners: LNCC, COPPE/UFRJ, INFS/UFRGS, LIA/UFC