IESP AND APPLICATIONS

IESP BOF, SC09

Portland, Oregon

Paul Messina

November 18, 2009

Outline

- "Scientific Challenges" workshops
- Applications involvement in IESP workshops
- Applications role in IESP

Purpose of DOE workshops

- To identify grand challenge scientific problems in [research area] that can exploit computing at extreme scales to bring about dramatic progress toward their resolution.
- The goals of the workshops are to
  - identify grand challenge scientific problems […] that could be aided by computing at the extreme scale over the next decade;
  - identify associated specifics of how and why new high performance computing capability will address issues at the frontiers of […]; and
  - provide a forum for exchange of ideas among application scientists, computer scientists, and applied mathematicians to maximize the use of extreme scale computing for enabling advances and discovery in […].

Science Workshop Series

- Climate, November 6-7, 2008
- HEP, December 9-11, 2008
- Nuclear Physics, January 26-28, 2009
- Fusion Energy Sciences, March 18-20, 2009
- Nuclear Energy, May 11-12, 2009
- BES, August 13-15, 2009
- Biology, August 17-19, 2009
- NNSA, October 6-8, 2009

Process used

- Workshops were organized jointly by US DOE's office of Advanced Scientific Computing Research and other DOE program offices
- Workshop chair(s) worked with relevant DOE program offices and colleagues to identify key areas to cover
- Four – six panels defined, panel co-chairs recruited
- White papers for each panel drafted and posted in advance of workshop
- Priority Research Directions (PRDs) identified by each panel
- Panels populated by domain science experts as well as mathematicians and computer scientists, including some international
- Observers from other agencies and math and CS community invited to each workshop, including some international

Priority Research Direction (use one slide for each)

Scientific and computational challenges

Brief overview of the underlying scientific and computational challenges

Summary of research direction

What will you do to address the challenges?

Potential scientific impact

What new scientific discoveries will result?

What new methods and techniques will be developed?

Potential impact on SCIENCE DOMAIN

How will this impact key open issues in SCIENCE DOMAIN?

What's the timescale in which that impact may be felt?

Climate PRDs for Model Development and Integrated Assessment

- How do the carbon, methane, and nitrogen cycles interact with climate change?
- How will local and regional water, ice, and clouds change with global warming?
- How will the distribution of weather events, particularly extreme events, that determine regional climate change with global warming?
- What are the future sea level and ocean circulation changes?

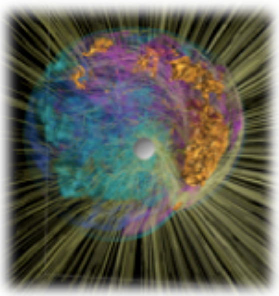
Climate PRDs for Algorithms and Computational Environment

- Develop numerical algorithms to efficiently use upcoming petascale and exascale architectures
- Form international consortium for parallel input/output, metadata, analysis, and modeling tools for regional and decadal multimodel ensembles
- Develop multicore and deep memory languages to support parallel software infrastructure
- Train scientists in the use of high-performance computers.

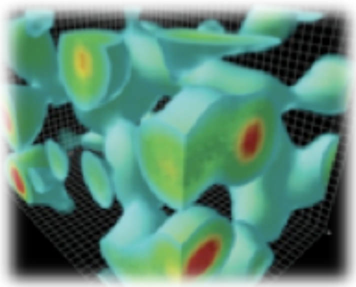
Exa-scale Computational Resources

(slide courtesy Martin Savage)

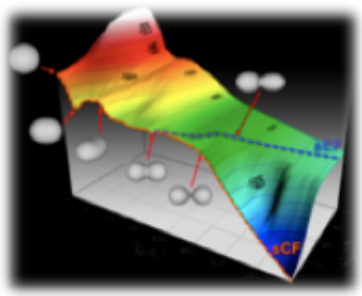
- Meeting structured around present Nuclear Physics areas of effort



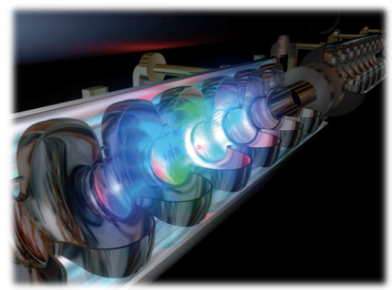
Nuclear Astrophysics



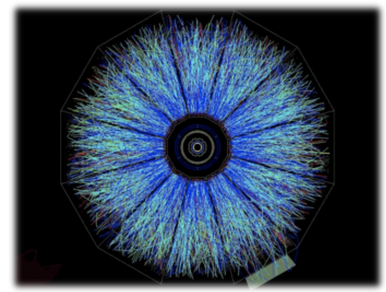
Cold QCD and Nuclear Forces



Nuclear Structure and Reactions



Accelerator Physics



Hot and Dense QCD

- Exa-scale computing is REQUIRED to accomplish the Nuclear Physics mission in each area
- Staging to Exa-flops is crucial :
  - 1 Pflop-yr to 10 Pflop-yrs to 100 Pflop-yrs to 1 Exa-flop-yr (sustained)

Nuclear Energy materials modeling

- Applications of high performance (peta-scale and exa-scale) computing carry along both the burden and the opportunity of
  - improved uncertainty evaluations,
  - margins quantifications, and
  - reliable predictions of materials behavior.
- Exascale computing will enable
  - simulations of trillions of atoms over seconds or days
  - simulations of complex, coupled physics and chemistry of reactor materials.

Recurring Topics in the Workshops -- Applications

- Exascale is needed: “the science case is clear”
- Predictive simulations
  - develop experimentally validated, predictive capabilities that can produce accurate and robust simulations
  - Ensemble simulations for V&V, UQ
- Multiphysics simulations
- Data volumes will be huge
  - Observations and simulation results

Recurring Topics in the Workshops – CS & Applied Mathematics

- Multiphysics and multiscale algorithm and software coupling
- Algorithms and software that deal with millions of cores and even more threads
- Data handling
- Interoperability
- Workflow issues
- Fault tolerance

For more information

- http://extremecomputing.labworks.org/

Role of Applications in IESP

- Identify applications needs in algorithms, software tools, programming frameworks
- Work with selected applications to co-design the software environment
  - and architectures (but not as part of IESP)

Applications considered
(not comprehensive, not ordered)

- Climate change
- Meterology
- Materials science
- Biology
- Plasma physics/fusion
- Geophysics
- Fluid dynamics
- Structural mechanics
- Electromagnetics
- Aerodynamics
- Combustion
- Lattice quantum chromodynamics
- Biophysics
- Astronomy/cosmology
- Molecular dynamics
- Video processing
- Chemistry
- Nuclear engineering/fission
- Epidemiology
- Nanotechnology/microelectronics
- Emergent sciences (e.g., social, networks, etc.)

Applications serve as Co-Design Vehicles

- Technology drivers
  - *Advanced architectures with greater capability but with formidable software development challenges*
- Alternative R&D strategies
  - *Choosing architectural platform(s) capable of addressing PRD's of Co-Design Vehicles on path to exploiting Exascale*
- Recommended research agenda
  - *Effective collaborative alliance between Co-Design Vehicles , CS, and Applied Math with an associated strong V&V effort*
- Crosscutting considerations
  - *Identifying possible common areas of software development need among the Apps that serve as co-design vehicles*
  - *Addressing common need to attract, train, and assimilate young talent into this general research arena*

4.3.1 Co-design Vehicles: Priority Research Directions

Criteria for Consideration

(1) Demonstrated need for Exascale

(2) Significant Scientific Impact in: basic physics, environment, engineering, life sciences, materials

(3) Realistic Productive Pathway (over 10 years) to Exploitation of Exascale

Summary of Barriers & Gaps

What will co-design vehicles do to address the barriers & gaps in associated Priority Research Directions (PRD's)?

Potential Impact on Software

What new software capabilities will result?

What new methods and tools will be developed?

Potential impact on user community (usability, capability, etc.)

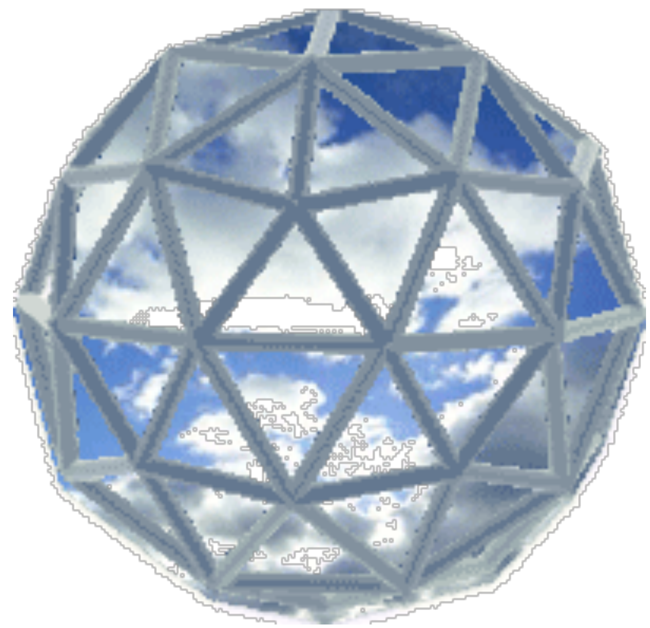
How will this realistically impact the research advances targeted by co-design vehicles that may benefit from exascale systems?

What's the timescale in which that impact may be felt?

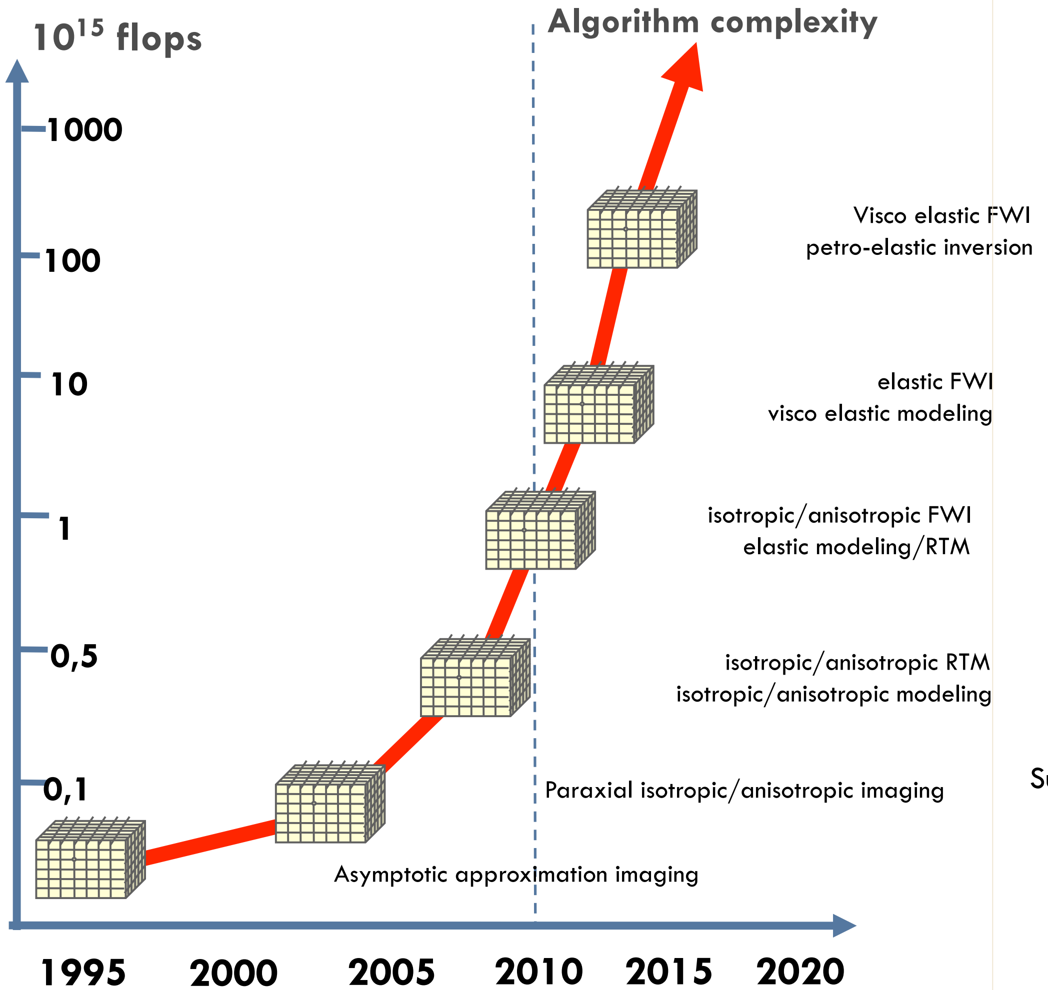
Computational challenges at the exascale

Model Complexity

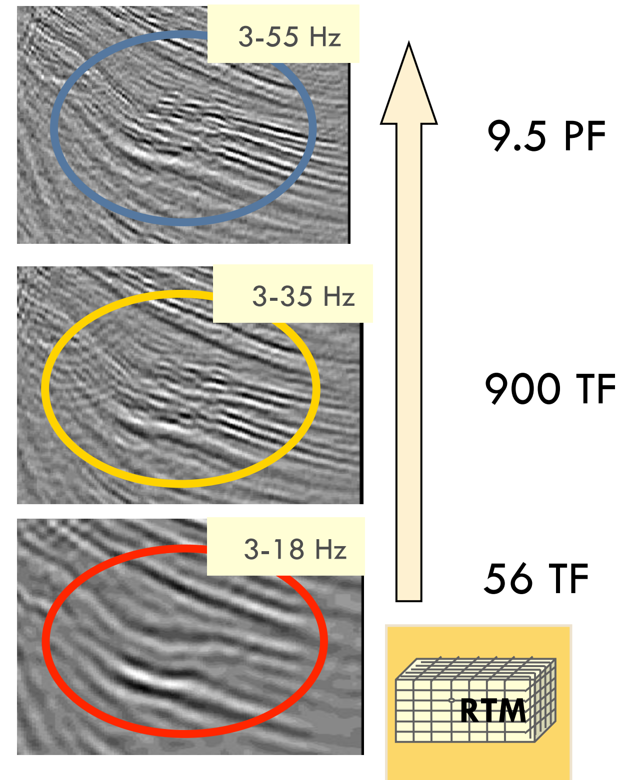
- The cloud feedbacks is the largest source of uncertainty in climate sensitivity estimates.
- Cloud Resolving Model (CRM) replaces the conventional convective and stratiform cloud parameterizations
- Global Cloud Resolving Model (GCRM as integration between Global Circulation Model and CRM) represents a global atmospheric circulation model with a grid-cell spacing of approximately 3 km, capable of simulating the circulations associated with large convective clouds.
- The major limitation is its high computational cost. Exascale architectures provide a solution to the last issue.



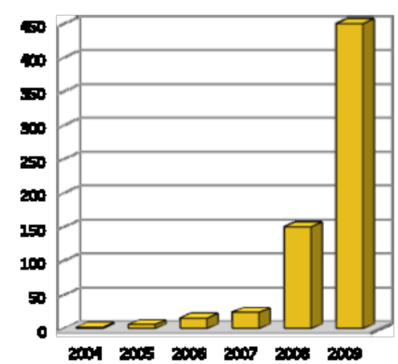
Industrial challenges in the Oil & Gas industry: Depth Imaging roadmap



Algorithmic complexity Vs. corresponding computing power



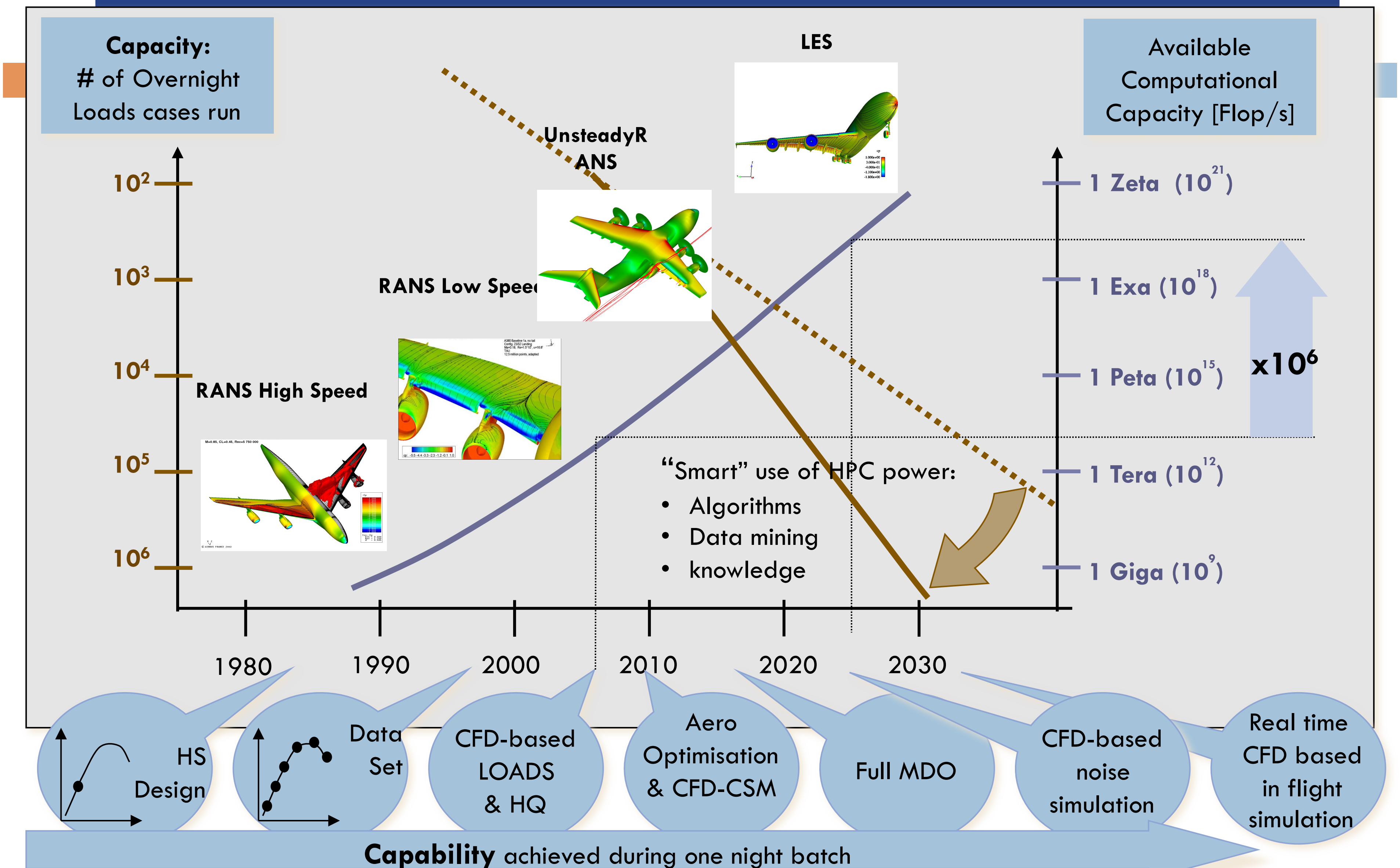
Substained performance for different frequency content over a 8 day processing duration



HPC Power
PAU (TF)

courtesy

High Performance Computing



Courtesy AIRBUS France

Computational Challenges and Needs for Academic and Industrial Applications Communities

2003 — 2006 — 2007 — 2010 — 2015

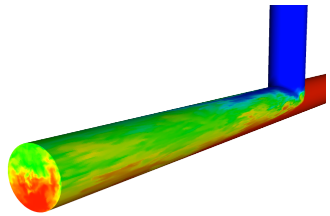
2003

Consecutive thermal fatigue event

Computations enable to better understand the wall thermal loading in an injection.

Knowing the root causes of the event ⇒ define a new design to avoid this problem.

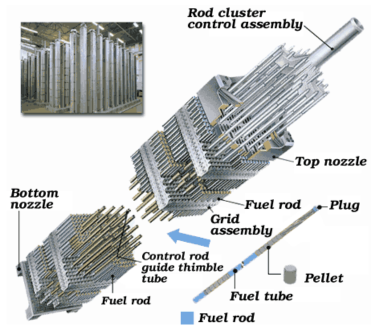
2006



Computation with an L.E.S. approach for turbulent modelling

Refined mesh near the wall.

2007



Part of a fuel assembly
3 grid assemblies

2010

9 fuel assemblies

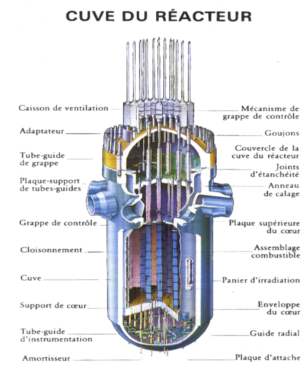
No experimental approach up to now

Will enable the study of side effects implied by the flow around neighbour fuel assemblies

Better understanding of vibration phenomena and wear-out of the rods.

2015

The whole vessel reactor



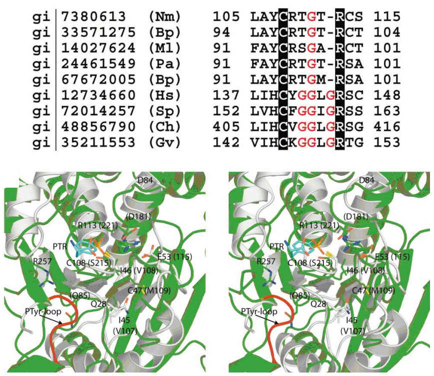
Computations with smaller and smaller scales in larger and larger geometries
⇒ a better understanding of physical phenomena ⇒ a more effective help for decision making
⇒ A better optimisation of the production (margin benefits)

2003	2006	2007	2010	2015
10^6 cells 3.10^{13} operations	10^7 cells 6.10^{14} operations	10^8 cells 10^{16} operations	10^9 cells 3.10^{17} operations	10^{10} cells 5.10^{18} operations
Fujistu VPP 5000 1 of 4 vector processors 2 month length computation	Cluster, IBM Power5 400 processors 9 days	IBM Blue Gene/L 20 Tflops during 1 month	600 Tflops during 1 month	10 Pflops during 1 month
# 1 Gb of storage 2 Gb of memory	# 15 Gb of storage 25 Gb of memory	# 200 Gb of storage 250 Gb of memory	# 1 Tb of storage 2,5 Tb of memory	# 10 Tb of storage 25 Tb of memory
Power of the computer	Pre-processing not parallelized	Pre-processing not parallelized Mesh generation	… ibid. … … ibid. … Scalability / Solver	… ibid. … … ibid. … … ibid. … Visualisation

From sequences to structures : HPC Roadmap

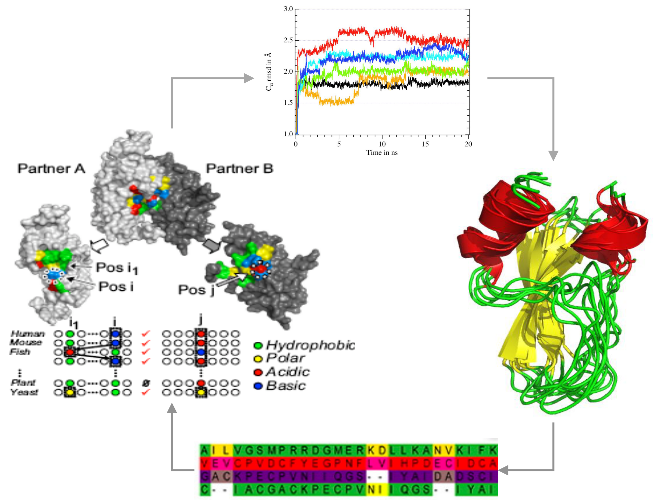
2009 | **2011** | **2015 and beyond**

Grand Challenge GENCI/CCRT

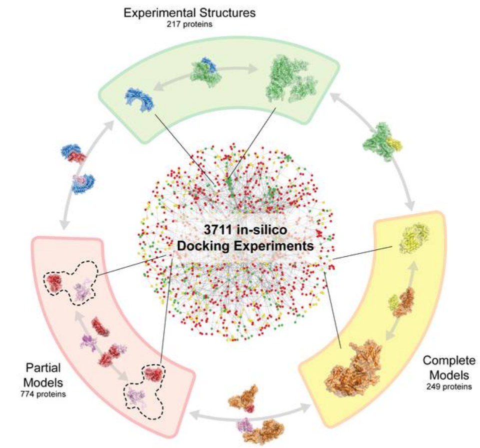


Proteins 69 (2007) 415

Identify all protein sequences using public resources and metagenomics data, and systematic modelling of proteins belonging to the family (Modeller software).



Improving the prediction of protein structure by coupling new bio-informatics algorithm and massive molecular dynamics simulation approaches.



Systematic identification of biological partners of proteins.

Computations using more and more sophisticated bio-informatical and physical modelling approaches ⇒ Identification of protein structure and function

2009	2011	2015 and beyond
1 family **5.10^3 cpu/~week**	**1 family** **5.10^4 cpu/~week**	**1 family** **~ 10^4*KP cpu/~week** **CSP : proteins structurally characterized ~ 10^4**
# 25 Gb of storage	# 5 Tb of storage	# 5*CSP Tb of storage
500 Gb of memory	5 Tb of memory	5*CSP Tb of memory

Summary

- Many application domains will benefit from usable exascale systems
- IESP is involving representatives from a number of those applications areas
- Some applications teams are eager to serve as co-design vehicles