D3.3 Working Group report on Industrial and Engineering Applications: Energy, Transports

CONTRACT NO	EESI 261513
INSTRUMENT	CSA (Support and Collaborative Action)
THEMATIC	INFRASTRUCTURE

Start date of project: 1 JUNE 2010 Duration: 18 months

Name of lead contractor for this deliverable: GENCI

Authors: Jean Claude ANDRE (Cerfacs) ; Philippe RICOUX (Total)

Contributors to the report:, Jean-Yves BERTHOU (EDF), Henri CALANDRA (Total), Eric CHAPUT (EADS), Tanguy COURAU (EDF), Fabian DUDDECK (Münich Univ.), Christian HASSE (Freiberg Univ.), Norbert KROLL (DLR), Jacek MARCZYK (Ontonix), Jean-Daniel MATTEI (EDF), Keld NIELSEN (ENI), Demetrios PAPAGEORGIOU (Imperial College), Heinz PITSCH (Aachen Univ.), Thierry POINSOT (Cerfacs), Ulrich RUDE (Erlangen Univ.), Mark SAWYER (Edinburg Univ), Ali TABBAL (Alstom), Chuan-yu WU (Birmingham Univ.)

Name of reviewers for this deliverable: Stephane Requena (GENCI)

Abstract: This deliverable reports on the synthesis of work and main outcomes identified by the working Group dedicated to Industrial and engineering applications

Release N° 1

Due date of deliverable:	Mai 31, 2011
Submission date:	June 01, 2011

Internal release N° 2

Due date of deliverable:	November 30, 2011
Submission date:	December 13 2011
Publication date:	December 22 2011

Project co-funded by the European Commission within the Seventh Framework Programme (FP7/2007-2013)		
Dissemination Level		
PU	Public	X

Table of Contents

1. Introduction

HPC Use in Industry

Industry is in the midst of a new, 21st century industrial revolution driven by the application of computer technology to industrial and business problems. HPC already plays a key role in designing and improving many industrial products — including automobiles, airplanes, pharmaceutical drugs, microprocessors, computers, implantable medical devices, golf clubs, and household appliances — as well as industrial-business processes (e.g., finding and extracting oil and gas, manufacturing consumer products, modeling complex financial scenarios and investment instruments, planning store inventories for large retail chains, creating animated films, and forecasting the weather).

HPC users typically pursue these activities with *virtual prototyping and large-scale data modeling* – that is,using computers to create digital models of products or processes and then evaluating and improving the design of the products or processes by manipulating these computer models. Given their broad and expanding range of high-value economic activities, HPC users are increasingly crucial for industrial and business innovation, productivity, and competitiveness.

Recognition of HPC's Importance for Industrial Innovation

Political and organizational leaders are increasingly recognizing HPC's crucial value for driving innovation and competitiveness. Commenting recently on the EU Innovation Union, launched in October 2010, Robert-Jan Smits, Director General for Research and Innovation of the European Commission, noted that:

> *"…research and innovation are key strands of the Europe 2020 strategy. Stark figures confront this ambition to use knowledge as a driver for sustainable growth. Albeit with large internal variations, Europe consistently spends less than 2 per cent of GDP on research and development, only two-thirds of that in the US and a little more than half the Japanese figure. Meanwhile, China's investment is growing year by year and will be on a par with Europe in a few years. The EU Innovation Union Scoreboard tells a similar story: a big innovation gap with Japan and the US, with China (not to mention India and Brazil) quickly catching up."*

Other examples:

- In his 2006 State of the Union address, U.S. President George W. Bush promised to trim the federal budget, yet urged more money for supercomputing. President Obama also mentioned supercomputing prominently in his 2011 State of the Union address.
- In 2009, Russian President Dmitry Medvedev warned that without more investment in supercomputer technology, Russian products "will not be competitive or of interest to potential buyers."
- In June 2010, Rep. Chung Doo-un of South Korea’s Grand National Party echoed that warning: “If Korea is to survive in this increasingly competitive world, it must not neglect nurturing the supercomputer industry, which has emerged as a new growth driver in advanced countries.”

2. Executive summary

Questions and issues to address:

- for efficiently reaching goals: Which effort ? Where to spend the money?

- the roadmap must be implemented in several industries

Roadmaps are different depending upon particular industries:

For some domains, Exaflop **capacities** are needed to solve industrial problems:

- Aeronautics:

  Improved predictions of complex flow phenomena around full aircraft configurations with advanced physical modeling and increased resolution, multi-disciplinary analysis and design, real time simulation of maneuvering aircraft, aerodynamic and aero elastic data prediction, Exaflop not the ultimate goal, need at least Zetaflop for LES of complete aircraft. Many "farming" applications (almost independent simulations)

- Seismic, O&G:

  Largely embarrassingly parallel, major problems are programming model, memory access and data management, need Zetaflop for full inverse problem

- Engineering:

  Optimization, Monte Carlo type simulations, LES/RANS/DNS …

  (main problem: equations themselves, physics, coupling, …)

For these domains, production problems will be solved by **"farming" applications**.

- Combustion and Gasification:

  turbulent combustion/gasification modeling, LES in large scale reactors, Coupling

  turbulent combustion modeling, LES in large scale reactors, Coupling multi-physics, Exaflop for Combustion at the right scale, Multi-cycle engine (Weak scalability)

- Nuclear Energy:

  steady CFD calculations on complex geometries , RANS, LES and quasi-DNS type calculations under uncertainties, Monte Carlo Ultimate Neutronic Transport Calculation (Time dependant & Multiphysics coupling)

- Other domains:

  multi-Fluids Flow, Fluid Structure Interactions, Fluidized beds (particle simulations, stochastic PDE, Multi-scale approach, physics, fluid in particles

Common main issues to be addressed:

- At the level of the simulation environment:

  - Unified Simulation Framework and associated services: CAD, mesh generation, data setting tools, computational scheme editing aids, visualization, etc.
  - Multi-physics simulation: establishment of standard coupling interfaces and software tools, mixing legacy and new generation codes
  - common (jointly developed) mesh-generation tool, automatic and adaptive meshing, highly parallel
  - Standardized efficient parallel IO and data management (sorting memory for fast access, allocating new memory as needed in smaller chunks, treat parts of memory that are rarely/never needed based on heuristic algorithms, …)

- At the level of codes/applications:
  - New numerical methods, algorithms, solvers/libraries, improved efficiency
  - coupling between stochastic and deterministic methods
  - Numerical scheme involving Stochastic HPC computing for uncertainty and risk quantification
  - meshless methods and particle simulation
  - Scalable program, strong and weak scalability, load balancing, fault-tolerance techniques, multi-level parallelism (issues identified with multi-core with reduced memory bandwidth per core , Collective communications, Efficient parallel IO)
  - *Development of standards* programming models (MPI, OpenMP, C++, Fortran, …) handling multi-level parallelism and heterogeneous architecture
- Human resources, training (what level?)

3. Goals of the WG 3.1, Industrial & Engineering applications

EESI WG3.1 main goal is to propose a **European vision and roadmap** to address the challenge of Exascale computation and simulation for **Industrial and Engineering applications, mainly energy and transport.**

This was done by involving a large numbers of HPC European industrial and academic "experts" in domains such as **Oil & Gas, Aeronautics, Train, Automotive, Chemical engineering, Power generation, Combustion, Propulsion & Engine flows,** investigating what are the real needs for "many cores" industrial applications and the main issues to be solved for efficient use of Exaflop computers.

must be coherent for all applications of industrial and engineering domains listed above.

From that overview, the working group identified a roadmap and timing, along with the strengths and weaknesses of Europe in these fields. Key questions were, what are the topics and challenges that Europe may and must develop and what are the ones that may better be addressed from outside Europe?

4. Industrial & Engineering Applications Scientific Expectations

4.1 Introduction

The present roadmap was prepared by a large panel of experts in numerical simulation and "many cores" applications, which covers a broad spectrum of industrial needs in the domains of energy and transportation.

For energy activities, the following companies are represented in the experts panel: Oil & Gas with TOTAL and ENI, Power generation with EDF, Aeronautics with EADS/Airbus and DLR, Train with ALSTOM, Automotive with BMW, Neutronics with EDF, Hydraulics with EDF, Chemical units and chemical Engineering with TOTAL.

Academic experts are also involved in the working group in order to cover all scientific approaches and methods implemented in such industrial applications:

- Propulsion Engine and combustion with Aachen University (G), CERFACS (F), Freiberg University (G),
- Particle simulation for chemical units with Erlangen University (G) and Birmingham University (UK),
- Multifluids flows with Erlangen University (G), Imperial College (UK), Edinburgh University (UK),
- Computer Aid Engineering with Ontonix (Poland),
- And, of course, CFD with most of the industrial and academic experts.

Here is the table of experts involved in the working Group 3.1 :

Name	Affiliation, Country	Domain of expertise
Henri CALANDRA	TOTAL, F	Geophysics, Oil and gas
Keld NIELSEN	ENI, I	Multiphase flows, Oil and gas
Thierry POINSOT	CERFACS, F	Combustion, CFD
Eric CHAPUT	EADS/Airbus, EU	Flight physics, Aeronautics
Demetrios PAPAGEORGIOU	Imperial College, UK	Applied math., Physics, Multifluids
Ulrich RUDE	Erlangen univ., G	HPC, Particle dynamics
Jean-Daniel MATTEI	EDF, F	CFD, Hydraulics
Christian HASSE	BMW, Freiberg Univ., G	Propulsion, engine flows & combustion, Automotive, Gasification
Heinz PITSCH	Aachen Univ., G	Propulsion, Engine flows,
Norbert KROLL	DLR, G	CFD, Aeronautics
Tanguy COURAU	EDF, F	Neutronics, Nuclear industry
Ali TABBAL	ALSTOM, F	Surface transportation, Trains
Chuan-yu WU	Birmingham Univ., UK	Chemical Engineering
Mark SAWYER	Edinburgh Univ., UK	HPC,
Jacek MARCZYK	Ontonix, I	Computer Aid Engineering

Jean-Claude ANDRE	CERFACS, F , Vice Chair	HPC, CFD, Energy
Philippe RICOUX	TOTAL, F, Chair	Oil &Gas, Chemical Eng., Petrochemicals

With such a panel, largely distributed and balanced between industry and academy, covering multidisciplinary approaches (physics, HPC, numerical analysis, ...), we think that the conclusions of the working group will concern all energy and engineering applications, and will be coherent and applicable in these sectors.

The primary focus of this report is on industrial and other engineering problems in relation to future exascale systems, and it is important to note the following:

- Industry has computing challenges that can be as large and as daunting as those faced in the fundamental sciences. Unlike national governments, industrial and engineering firms typically cannot justify purchasing the largest versions of contemporary and future-generation HPC systems. They therefore require access to the large government-sponsored systems. If they do not secure this access, they will have nowhere to run their most important problems. Not having this access would hamper industrial innovation and competitiveness.
- Where the science or competitive time constraints don't support large single runs of a problem, smaller , concurrent iterative runs often allow users to home in on productive solutions. Long-established examples include stochastic modeling in the financial services sector and parametric modeling in the design engineering realm. Iterative solutions are no less valuable simply because they are less scalable than the highest-performing HPC codes. And while iterative solutions run in parallel may have smaller computing requirements than the highest-scaling codes, in the aggregate (multiple runs) they can still require considerable time on HPC resources.
- It follows from the preceding point that small and medium-size enterprises (SMEs), as well as larger companies, would benefit from gaining access to high-end HPC systems.

4.2 Main challenges

This chapter goes in more details about scientific contents covered by industrial and engineering applications relative to energy and transportation:

- **Aeronautics**: full Multidisciplinary Design and Optimization (MDO) of large-scale aeronautical systems, CFD-based noise simulation, real-time CFD-based in-flight simulation: the digital aircraft
- **Structure calculation**: design new composite compounds, deformation …
- **Special chemistry**: molecular dynamics, atom to continuum (multi scale) simulation for macro parameters estimation in catalyst (or/and electrochemistry, electronics, …), surfactants, tribology, interfaces, nano-systems …
- **Energy:** computations with smaller and smaller scales in larger and larger geometries: turbulent combustion in closed engines and opened furnaces, explosion in confined area, power generation, hydraulics, nuclear plants …
- **Oil and gas industries**: full 3D inverse waveform problem (seismic), oil & gas reservoir modeling (including compositional oil model, thermal, heavy oil, Enhanced Oil Recovery (chemical EOR), data assimilation, …) , multiphase flows in porous media at different scales (pore level (nm) to homogenized cells (cm)), transient multi-fluid pipe flow and phase equilibrium, process plant design and optimization, CO2 storage, …
- **Engineering (in general)**: multi-scale CFD, multi-fluids flows, multi-physics modeling, complex systems modeling, computer aid engineering, Mixed Integer linear and non-linear programming including robust optimization, stochastic optimization, …

4.3 Scientific and Computational key issues by sector

The energy and transportation industries, and more generally engineering industries, are looking for multi-petaflopic and exaflopic computing resources: not only the next supercomputers themselves but also the ways to know how to program on 10^5-10^6 cores machines.

Industry is in fact expecting the **best coupling between Architecture / Algorithm / Application**, in order to address and solve on exaflop systems crucial issues in energy and transportation, along with other economically and socially important engineering challenges.

Clearly the development of new algorithms and their implementation must be much better focused on the requirements that are posed by current and future high performance architectures. A closer look reveals that the sources of performance degradation can be classified in three categories that are necessarily interlinked:

- current and future **computer architectures** and their hardware implementation prohibit traditional implementation styles from performing acceptably;
- the traditional numerical algorithms are based on an **mathematical analysis** that does not help to improve performance in many parallelizations, but often rather obscures the needs of efficient machine execution;
- the **software implementation and the supporting tools** that are in use today often fail to exploit the performance potential of the underlying hardware.

In all three categories, exascale computing requires major additional research efforts to be performed in a coordinated fashion, with improved interdisciplinary interaction between mathematics, computer science, and the application domains.

Research in hardware and architecture development is largely outside the scope of this part of this software-focused report. However, the efficient implementation of computational methods and software presupposes a thorough understanding of architectural trade-offs and the future trends, such that existing and new algorithms can be adapted to the upcoming challenges. This includes in particular a better understanding and better analysis of the effects of deep memory hierarchies, latencies and bandwidth limitations. Such bottlenecks may occur both in the network fabric and in the inter-node memory transfer. An improved analysis also requires considering power consumption and fault tolerance effects. In these areas more research is needed.

In the second category, more research is necessary to design hardware-aware numerical algorithms and to revise existing algorithms such that their performance can be improved for existing and future computer systems. Currently, few numerical algorithms have been adapted for the challenges of exascale level computing.

In a general manner, experts consider that the following items are research domains where Europe has to put money for reaching the goals of exaflopic applications:

- **ultra-scalable algorithms**: The goal must be to design algorithms suited to run on hundreds of thousands of cores and beyond
- **overcoming the "memory wall":** Tuning and optimizing algorithms for deep memory hierarchies. Especially on manycores systems, the reduced memory bandwidth is identified as a major source of inefficiency.
- **latency-tolerant algorithms**. Collective communication is becoming a more and more important source of inefficiency. If their implementation is not redesigned significantly to address exascale capacities, new algorithms will have to be designed and implemented.
- exploiting **SIMD** features and **GPU**-like architectures. The current trends indicate that new architectures will provide more fine-grained parallelism in the form of wider SIMD-vector units, and fine-grained many-core acceleration, similar to that provided by current GPU cards.
- **communication-reducing- and communication-avoiding algorithms.**
- **power-efficiency:** designing algorithms for optimizing power consumption

- **algorithmic fault-tolerance:** exploring fault tolerant algorithms. A new generation of scalable, fault tolerant algorithms has to emerge.
- **new numerical methods and solvers:** these are critical issues for all engineering domains where new applications will become feasible by the power of exascale computers. This includes, e.g., numerical methods for solving more complex wave equation formulations, multiphase flows, adaptive methods for heterogeneous platforms, hybrid solvers, advanced numerical acceleration techniques, multi-grid, better and simpler pre-conditioners are key issues that have to be addressed.
- **parallel IO:** Efficient input/output for post-processing, possibly allowing interaction with the simulation, and scalable management of large data sets is also mandatory.
- **stochastic HPC computing, uncertainty and risk quantification:** Probabilistic quantification of the risks and uncertainties affecting a best-estimate model generates a whole new domain of applied science, linking probabilistic, numerical analysis as well as physics and decision-theory. Beyond the traditional Monte-Carlo sampling, a number of uncertainty propagation and probabilistic simulation algorithms have been and have to be developed, such as accelerated sampling (importance sampling, particulate methods etc.), reliability techniques, stochastic developments (e.g. chaos polynomials) and response surface techniques, ... The coupling between stochastic and deterministic methods is also a key for solving stochastic optimization problems.
- **data driven computing, inverse problems:** The challenges come with the need for inverse probabilistic techniques, as the observable data to calibrate model variability generally comes on parameters different from the model inputs. Closely related is the need for a general coupling between stochastic optimization and simulation in order to achieve robust design or operational management strategies, with challenging mathematical implications that are only partially solved under existing expectation-maximization or stochastic dynamic programming algorithms (typically limited to close to Gaussian/linear behavior).

Over all scientific issues, the working group draws attention to the fact that **roadmaps will be different, depending on the specific industry:**

4.3.1 Aeronautics

The impact of computer simulation in aircraft design has been significant and it continues to grow. Numerical simulation allows the development of highly optimized designs and reduced development risks and costs. Boeing, for example, exploited HPC to go from having to design 77 physical prototype wings for the 757 aircraft, to only 11 prototype wings for the 787 "Dreamliner" plane. HPC usage saved the company billions of euros. Aircraft companies are now heavily engaged in trying to solve problems such as calculating maximum lift using HPC resources. This problem has a nearly insatiable appetite for computing power and, if solved, would enable companies designing civilian and military aircraft to produce lighter, more fuel-efficient, environmentally friendlier planes.

The ***limitations of today's numerical tools*** reduce the scope of innovation in aircraft development, keeping aircraft design at a ***conservative level***.

To meet the challenges of future aircraft transportation (Greening the Aircraft), it is *indispensible to be able to flight-test a virtual aircraft with all its multi-disciplinary interactions in a computer environment and to compile all of the data required for development and certification with guaranteed accuracy in a reduced time frame.*

In parallel, future aircraft concepts require deeper basic understanding in areas such as turbulence, transition and flow control to be achieved by dedicated scientific investigations.

The roadmap for approaching the digital aircraft vision includes the following major simulation and optimisation challenges:

Improved physical modelling for highly separated flows

This requires full unsteady Navier-Stokes computation of complex aircraft configurations at the border of the flight envelope using advanced turbulence models (hybrid RANS/LES or LES) as well as accurate integrated prediction tools.

Since this type of simulations is exploring ***weak scaling***, *Exascale computers will push LES towards increased resolution quality and industrial relevant problems.*

But, due to the enormous resolution requirements both in space and time (10^{11} - 10^{12} grid points for Reynolds number Re=$3x10^7$), ***LES for full aircraft will most likely not be feasible with an Exascale computer. So Exascale applications in this area are not the final goal.***

Real-time simulation of aircraft in flight

Simulation of a manoeuvring aircraft requires the coupling of the aerodynamic, structural mechanics, aero elastic and flight-mechanic disciplines based on high-fidelity methods within a multi-disciplinary massively parallel simulation environment.

The main issue for this type of simulations is turn-around time.

Thus ***strong scaling*** is important when using massively parallel computer platforms.

But, since most of the legacy CFD codes scale up to $O(10^3)$ – $O(10^4)$ cores on typical flow problems, ***an Exascale computer*** with $O(10^8)$ – $O(10^9)$ cores ***will address*** the computation of ***several maneuvers simultaneously***.

Once more, Exascale systems will help to solve industrial problems by task-<u>farming</u> applications, that is, parcelling out jobs or sub-jobs to available computational resources.

Aerodynamic and aeroelastic data production

This requires methods to calculate in the shortest possible time frame aerodynamic loads and derivatives of the aircraft for every imaginable flight situation. The data are used for the structural layout of the aircraft, the flight simulator data base and the development of the flight control system. The required input is currently obtained mainly from experiments or from simple empirical methods. The use of time consuming CFD methods to obtain the necessary data has been, up to now, impossible due to the large volume of data required.

Based on current assumptions concerning the scalability and numerical efficiency of legacy CFD codes, more than 5000 cases can be run on an Exascale computer as <u>farming</u> applications over night.

Noise source and impact

This requires the full development of noise source mechanisms, acoustic radiation and noise impact simulation tools which compute acoustic disturbances on top of aircraft flow. Since resolution in space and time is one of the major issues (***weak scaling***), computer resources in the range of ***Exascale will significantly improve the understanding of noise source and impact for industry relevant configurations.***

Multidisciplinary aircraft design

This requires a fully coupled simulation of the flow around a parameterized aircraft configuration and surface shapes covering a reactive structural model within a sophisticated optimization process.

The coupled large-scale simulations will run ***multiple times, leading to <u>farming</u> applications on Exascale systems.***

The computational power needed to cope with full aircraft simulation and optimization as described above is 10^7 – 10^8 higher than today's capability. Such an increase in sustained speed and respective memory will enable the aircraft industry to develop improved aircraft technology and optimize their product from economical and ecological point of view.

A general underpinning challenge is likely to be posed by the expansion of future parallel systems to $O(10^8)$ – $O(10^9)$ cores.

For fundamental turbulence simulations requiring billions of grid points, scaling may be anticipated as a continuation of past trends.

For legacy CFD codes which have to deal with complex geometries and order of 100 - 500 million grid points, scaling becomes more of a question.

This may require renewed attention to numerical methods and software engineering.

In addition, for challenges like multi-physics or multi-disciplinary simulations the diverse modelling required makes the scaling issue for Exascale systems a research topic.

Furthermore, attention to extracting more of the theoretical peak performance of a massively parallel computer is an area of significant potential gain for simulation performance.

4.3.2 Seismic, Oil & Gas

The petroleum industry is strongly motivated to increase the efficiency of its processes, especially in Exploration and Production and to reduce risks by the deployment of high performance computing. Typical steps in the business process are, **geo-science** for identification of oil and gas in the underground, development of reservoir modeling, designing of facilities for the cultivation of hydrocarbons; drilling of wells and construction of plant facilities; operations during the life of the fields; and eventually decommissioning of facilities at end of production.

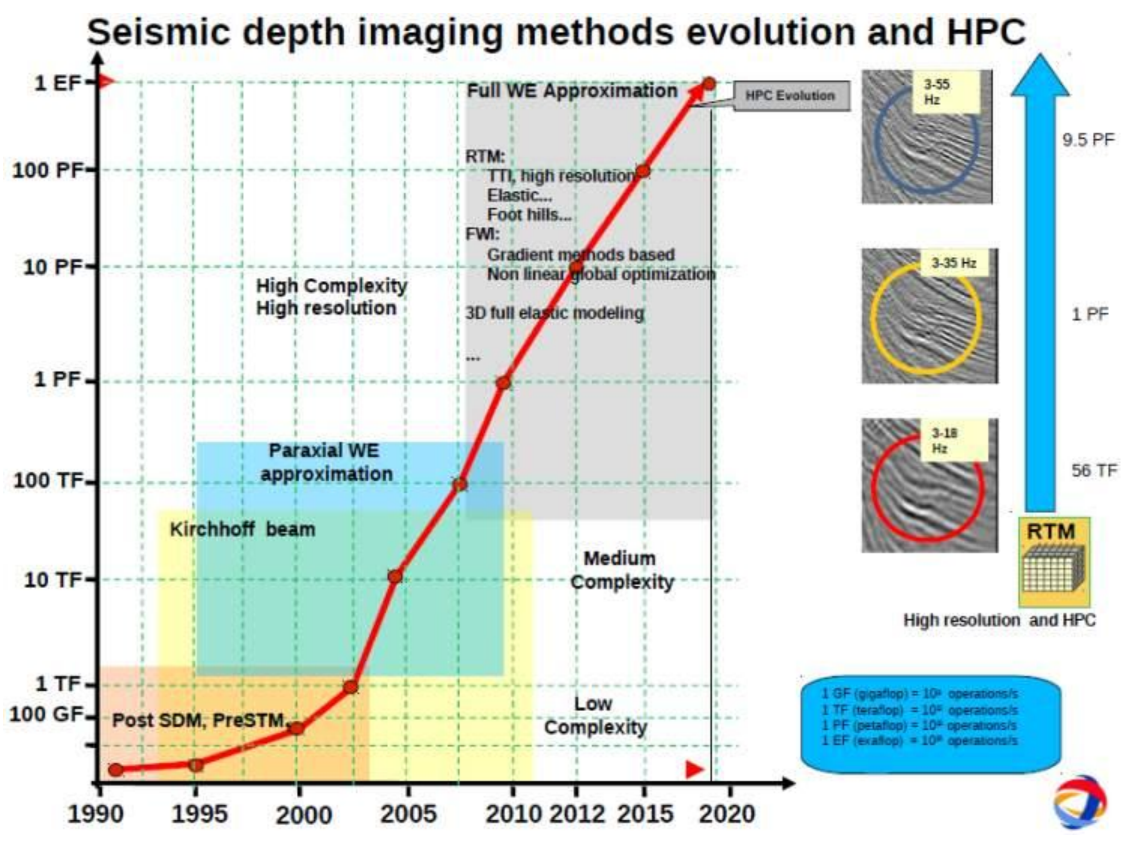
Geo-science analyses seismic data with numerical techniques for inverse problems. The economical impact of HPC is definitely high and the best possible tools are deployed.

As for aeronautics, once more, Exaflop is not the ultimate goal. The complete Inverse Problem Resolution of wave equation needs more computational resources.

The objective of this application is to produce from a seismic campaign the best estimation of the underground topography in order to optimize reservoir delineation and production by solving the Full Inverse Wave Equation.

This application is largely embarrassingly parallel, and the higher performing the HPC system is, the better the approximation of the underground topography.

A roadmap of the steps of this kind of approach, showing the different necessary and more complex methods of approximations of the physical reality (e.g. elastic, visco-elastic, etc), is given below with *courtesy of Total.*



For this application, it is essential to both define and implement new algorithms representing more accurately the physics of the problems to be solved, and also to deploy ever more powerful hardware.

Moving beyond geo-science, the other activities in the petroleum industry have aspects which are generally classified as system-of-systems-design and ***multiphase fluid dynamics***. Of these topics fluid dynamics require a significant effort in terms of computing.

In fact as for geo-science similar criteria are valid also for multi-fluid problems. ***Enhanced quality of simulations depend both on more appropriate physical models as well as on numerical methods*** and techniques, e.g. for bifurcation analysis. Physical scales are disparate, for instance in pipeline modeling where the diameter is measured in fraction-of-meter, whilst the length of the pipeline normally is measured in kilometers.

Moreover, chemical and thermodynamic modeling has to be considered too; vapor-liquid phases are to be calculated throughout the pipeline, it may be necessary with tracking of chemical species and for instance drag-reducers measured in ppm may have a significant impact on pressure drop between pipeline inlet and outlet.

So, ***multi-scale techniques*** are highly relevant as well as shear computing power.

In another hand, and it is crucial for ***heavy oil*** (or new oil resources) ***reservoir modeling***, ***multi-physics models*** must highly coupled for solving these very complex problems.

But, typically due to the equations (complex, non linear, stochastic, …), such problems should be examined as ***Farming Exaflopic applications, at least for the next 5 to 10 years.***

New discretizations of these different physical equations should be certainly be discovered but these domains will require one or more breakthroughs for an efficient use on Exascale systems.

Some techniques linked to Reservoir Modeling involving data assimilation, here called **History Matching,** will benefit much more at short term of supercomputing capacities.

System-of-systems-design has a wider scope and covers several engineering threads. Typically these threads are: Global optimization of plants including multi-disciplinary interactions, e.g. simulating full plant lifetime and control systems, emulation of operations and risk analysis. Solutions to these system-of-systems analysis tasks will typically be of the HPC-farming-class.

So, to sum up, major problems and issues are:

- use of standard programming model (MPI, Open MP etc.)cross-compiling – development on pc's, deployment on HPC's
- maintenance of legacy codes
- tools for test, verification and **validation** of (parallel) codes
- memory access
- data management
- task management and distribution
- efficient solvers
- numerical multi-scale techniques and methods
- efficient massively parallel coupling

These items will be described more in details in roadmap research priorities defined below.

4.3.3 Power Generation, Nuclear Plant

In this industrial domain, the objectives are multiple: first **Improvement of safety** and **efficiency of the facilities** (especially nuclear plants), and second optimization of maintenance operation and life span. This is one field in which physical experimentation, for example with nuclear plants, can be both impractical and unsafe. Computer simulation, in both the design and operational stages, is therefore indispensable.

Thermal hydraulic CFD Application Field.

Improvement of efficiency may typically involve mainly steady CFD calculations on complex geometries, while improvement and verification of safety may involve long transient calculations on slightly less complex geometries.

- *Study of flow-induced* loads so as to minimize vibration and wear through fretting in fuel assemblies may require from ***200 million to 2 billion cells*** per fuel assembly and to correctly account for both cross-flows in the core and walls around the core, at least one quarter of a core (over 100 assemblies) may need to be modeled***. Such studies on a quarter of a core need around 10 Pflops during few weeks. Exaflop systems will be useful to solve farming applications.***
- To *study flow-induced deformation in PWR cores*, a full core may need to be represented, at a slightly lower resolution, for an estimated grid size of ***at least 5 billion cells, which leads to runs on 100 Pflops during few weeks.***
- Detailed simulations designed to verify and increase safety may require full core simulations, and mesh sensitivity studies for these transient calculations may require unsteady calculations for meshes ***from 5 to 20 billion cells*** before 2020 ***which correspond to runs on 400 Pflops during few weeks.***
- To validate the models used for calculations such as the ones described above, as well as many others, running quasi-DNS type calculations on subsets of the calculation domain may be necessary.

  This will require meshes in the ***20-billion cell range by 2012*** (to study cross-flow in a tube-bundle, in a simplified steam generator type configuration), and running similar calculations for more complete calculation domains may require meshes well above ***100 billion by 2020***.

  Note that as safety studies increasingly require assessment of CFD code uncertainty, sensitivity to boundary conditions and resolution options must be studied, but turbulence models may still induce a bias in the solution.

  Doing away *with turbulence models and running DNS-type calculations* at least for a set of reference calculation would be a desirable way of removing this bias.

  Such studies will require multi-Exaflop capacities during few weeks.

Neutronics application field

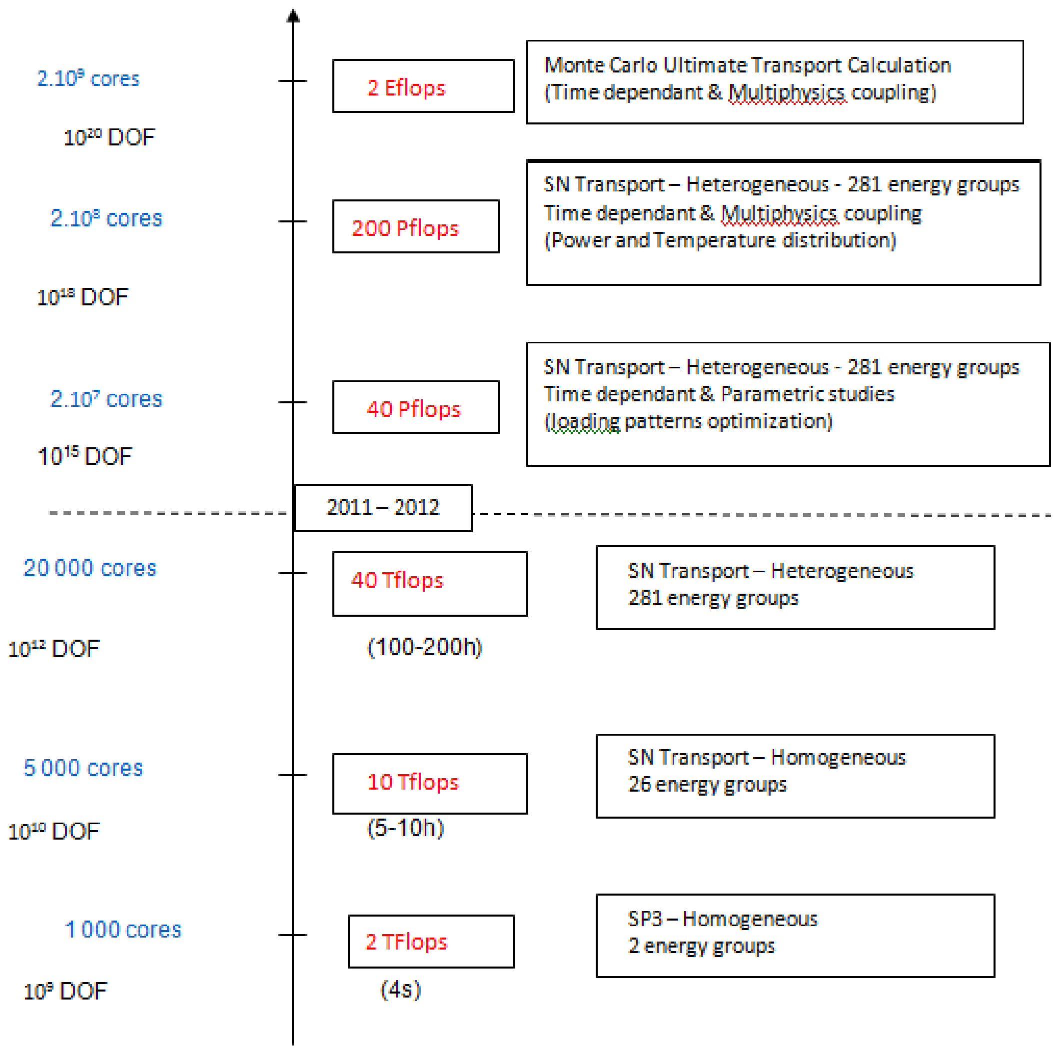
We have already established the value of running our neutronics codes on 100 Tflops / 30 000 cores computers which yields a much better understanding of the physics and in turn allows for: a better optimization of our power plants, potential operating margins, an increased safety and performance, a lower environmental impact and costs, and an extended lifetime.

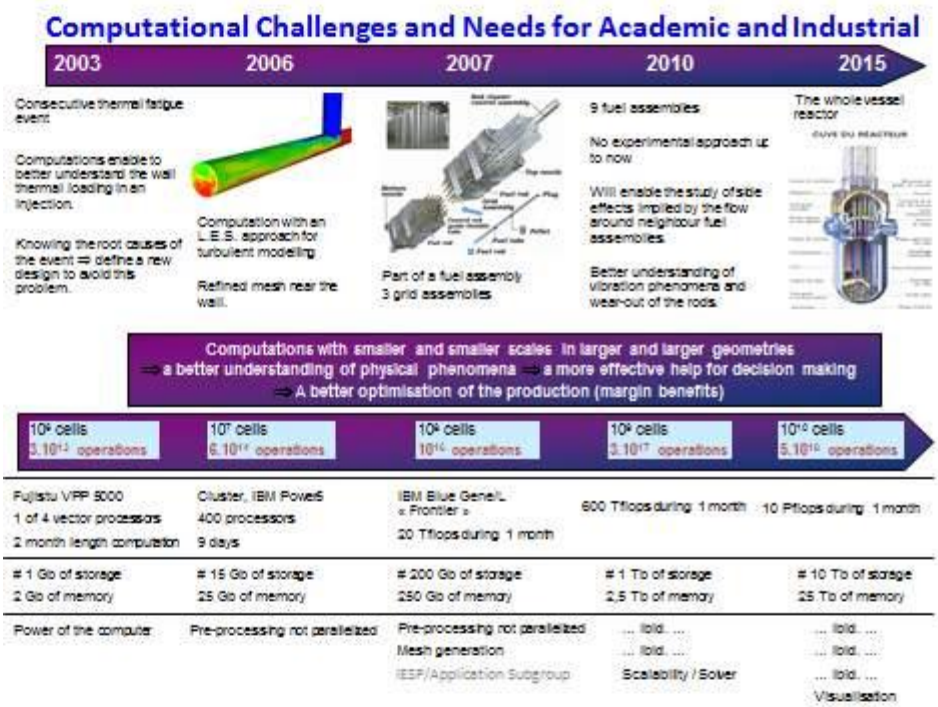
We therefore feel that Exaflopic software should not only be thought as a way of tackling daunting research problems. They also offer a mean to address the sometimes equally daunting challenges that stem from an industrial usage perspective. This includes the capability to model very complex, possibly coupled phenomena over extended spatial and time scale.

On top of that, ***uncertainty quantification and data assimilation are considered as key to industrial acceptance***, their associated computational needs that depend on the complexity of the model considered have to be met.

In terms of computing resources, projections are difficult to make because of the non linear behavior of iterative algorithms with respect to the degrees of freedom – and the number of processors. Additionally, new algorithms may have to be implemented to address the new types of numerical/physical problems within an evolving architecture.

A possible roadmap for Exaflopic neutronics computation is presented below with the courtesy of EDF:


2.10⁹ cores
10²⁰ DOF
2 Eflops
Monte Carlo Ultimate Transport Calculation
(Time dependant & Multiphysics coupling)
2.10⁸ cores
10¹⁸ DOF
200 Pflops
SN Transport – Heterogeneous - 281 energy groups
Time dependant & Multiphysics coupling
(Power and Temperature distribution)
2.10⁷ cores
10¹⁵ DOF
40 Pflops
SN Transport – Heterogeneous - 281 energy groups
Time dependant & Parametric studies
(loading patterns optimization)
2011 – 2012
20 000 cores
10¹² DOF
40 Tflops
(100-200h)
SN Transport – Heterogeneous
281 energy groups
5 000 cores
10¹⁰ DOF
10 Tflops
(5-10h)
SN Transport – Homogeneous
26 energy groups
1 000 cores
10⁹ DOF
2 TFlops
(4s)
SP3 – Homogeneous
2 energy groups


Computational Challenges and Needs for Academic and Industrial
2003
2006
2007
2010
2015
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.
Computation with an L.E.S. approach for turbulent modelling
Refined mesh near the wall.
Part of a fuel assembly
3 grid assemblies
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.
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)
10⁶ cells
10⁷ cells
10⁸ cells
10⁹ cells
10¹⁰ cells
10¹⁶ operations
3.10¹⁷ operations
Fujistu VPP 5000
1 of 4 vector processors
2 month length computation
Cluster, IBM Power5
400 processors
9 days
IBM Blue Gene/L
« Frontier »
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
IESP/Application Subgroup
... ibid. ...
... ibid. ...
Scalability / Solver
... ibid. ...
... ibid. ...
... ibid. ...
Visualisation

Electric Power Generation Overview

Many other applications exist beyond the ones mentioned: new generations of power plants, innovation in renewable energies and storage, protection against specific environmental threats (flood, heat wave, ...), customer's energy efficiency, development in home and building of technologies and services for energy efficiency, …

Several problems should be tackled for reaching these goals: CFD, Heat and Multi-fluids flows, Thermal hydraulic CFD, LES simulations, … ***modeling very complex systems, possibly coupled phenomena over extended spatial and time scales***, mixed with capacities like ***uncertainty quantification or data assimilation*** that are key to ***industrial acceptance.***

The challenge is particularly severe for **multi-physics, multi-scale simulation platforms** that will have to combine massively parallel software components developed independently from each others.

Another difficult issue is to **deal with legacy codes**, which are constantly evolving and have to stay in the forefront of their disciplines.

This will require ***new compilers, libraries, middleware, programming environments, languages, as well as new numerical methods, code architectures, mesh generation tool, visualization tool …***

Smart Power Grid Management

HPC is also needed to manage large "smart" power grids effectively, including real-time status monitoring and reporting, as well as disaster prevention and recovery.

4.3.4 Transportation, Automotive

General

The automotive industry is actively pursuing important goals that need exaflop computing capability or greater, including the following examples:

- Vehicles that will operate for 250000 kilometers (150000 miles) on average without the need for repairs. This would save automotive companies substantial money by enabling the vehicles to operate through the end of the typical warranty period at minimal cost to the automakers.
- Full-body crash analysis that includes simulation of soft tissue damage (today's "crash dummies" are inadequate for this purpose). Insurance companies especially desire this.
- Longer-lasting batteries for electrically power and hybrid vehicles.

For both aerodynamics, as for combustion, at least LES, and if possible DNS, simulations are required in an industrial scale and **Exaflop applications must be developed at the right scale,** according to weak scalability, … but these simulations must be coupled to all physics (flow, thermal, thermodynamic, chemistry, …) involved in the global transportation system.

So, coupled simulations involve **at least one legacy code** with

- Full scale, **multi-physics** configurations
- **Multiple runs** for optimization and parametric/statistical analysis

The global roadmap for this sector could be as following:

- *Individual performance and scalability of component codes*
  *Exaflop systems will mainly allow multiple runs, by **"farming" applications**, for "optimized" resolutions*
- *Overall performance of the multi-physics coupled system*
  *Once more, that leads to **farming applications**.*
- *Data management*

Combustion

Combustion is used to produce around 90 % of the earth energy and is essential for ground and air transportation, electricity production, industry applications or safety.

The central position of combustion in our world will not decrease before a very long time and science, especially simulations must allow combustion to continue ***with the lowest impact on climate and the highest efficiency***.

This can be achieved only through aggressive research on tools for combustion computations. Future combustion simulations will rely on three methodologies:

- ***Direct Numerical Simulation (DNS)*** are very high fidelity computations, ***without model for turbulence or chemistry.*** Because of their computational cost, they are performed in small cubic domains.

They are *the best workhorse today to reveal the internal structure of turbulent flames and understand their propagation, extinction, ignition, pollutant formation*.

In the next years, these methods will perform well on exascale machines and they will be mandatory for all groups willing to advance basic combustion science.

DNS is the topic of the Combustion CO design Center initiated in 2011 by J. Chen and J. Bell at Sandia National Laboratories in the USA. Multiple groups in Europe have the capacity to follow this evolution.

- ***Large Eddy Simulation (LES)*** use ***slightly simpler description of chemistry and turbulence models*** to compute combustion in larger domains, of realistic shapes, for example gas turbine chamber or piston engines. *LES have revolutionized the field of numerical combustion in the last twenty years* by bringing almost DNS-like capacities to real industrial systems. Today, industry relies and invests on LES to compute multiple phenomena which are beyond the capacities of existing 'classical' codes available in companies. ***For example, the ignition o***f a h***elicopter chamber at high altitude, the unstable modes of an in***dustrial gas turbine or the quenching of an ***aircraft engine can be addressed with LES on Tier1 (national wide) or Tier0 (PRACE European wide) machines.***

Similarly the ***cycle-to-cycle variations observed in piston engines (***which limit considerably multiple recent developments in this field) ***can be computed using modern LES tools*** *(see below for details)****). European groups are leaders in this field and the European codes for combustion LES are recognized as the*** most advanced.

- Finally, in many cases, computing only the reacting flow within the combustion chamber is not sufficient and ***multi physic / chemical phenomena must be coupled*** in the simulations: in a gas turbine, simultaneous simulations of the combustion chamber but also of the compressor (feeding the chamber) or of the turbine (fed by the chamber) are needed. For example, during ignition sequences of aircraft engines, a violent ignition in the combustion chamber can stall the compressor and lead to re-ignition failure at high altitude.

Similarly, the temperature of the walls of the chamber and the noise emitted by the combustor must be computed.

In this field, combustion LES coupled to other codes is needed. ***Running all these codes together on a parallel machine raises new challenges in the field of load balancing but also of simulation control*****.**

The proper framework for these multi-physics simulations is a <u>parallel coupler</u>, enabling to synchronize and distribute 3 or 4 codes on an exaflop machine.

Like for aerodynamic simulations, the challenge of combustion simulation is not only to perform a very expensive simulation of one combustor ***but to explore all possible combustor designs to maximize efficiency and safety while minimizing pollution by farming applications***.

Therefore optimization and ***uncertainty quantification*** are also crucial topics in this field.

The best combustion codes today can use up to 2 billion points, track ten to 30 species over hundreds of thousands time iterations. The typical partitions used on existing machines go from 4000 to 32 000 cores for production runs. In the next ten years, ***grid*** needed for industry applications will reach ***500 to 1000 billion points; they will be computed on 100 000 to 1 million cores.***

NB: Issues related with internal combustion are discussed below in the automotive section.

Automotive Applications

For the automotive area, three important applications, namely the CFD of internal combustion engine flows, the crash and thermal exchanges are discussed in the framework of Exascale.

Internal Combustion (IC Engine Combustion):

The general statements concerning combustion simulation above also apply for the special area of engine combustion simulation. Simulation of combustion and pollutant emissions in combination with engine test bench experiments is an established tool in the engine development process. The increasing complexity of future engine designs in the next decades e.g. due to worldwide variability in (bio-) fuels will lead to an almost exclusive use of simulation in the early development stages and to an extensive use in the later stages. ***This can only be achieved with the availability of reliable combustion simulation tools and their extensive use on exascale machines***.

The following four applications of virtual IC engine design will be most important:

- ***Large Eddy Simulation (LES) of single representative engine operation points*****.** Several physical phenomena can only be resolved using LES and a typical example is cycle to cycle variations, which limit the amount exhaust gas recirculation and therefore are critical to achieve the best fuel economy. Another phenomenon currently not fully accessible by simulation is the stochastic process engine knock, which highly depends on the available fuel. The accurate prediction of knocking to assess engine durability requires a large number of computed engine cycles (>100) and effects such as heating of engine parts, e.g. intake and exhaust valves, must be taken into account.

The wall clock time of such simulations must be less than a week *that leads to an estimated number of cores* ***larger than 20 000 per run, accessible with an Exascale system.***

- ***LES or RANS computations of whole engine*** maps with accurate pollutant predictions. In addition to the detailed LES of the most important engine operating points either defining fuel efficiency or engine durability described above, the evaluation of the remaining engine map (>500 additional operation points) must be computed.

The large number is due to the flexibility of the engine operating characteristics (boost pressure, injection timing, valve timing etc.). ***These calculations for one or a limited number of engine cycles*** are important for various areas such as catalyst design or electronic control unit calibration, ***but once more accessible with an Exascale system.***

- ***LES or RANS computations of long term engine operation effects*****.** An important issue is the change in engine performance during transient and long term operation. Typical examples are coking of valves or injectors leading to deteriorated engine performance. The predictions of such phenomena requires simulations of a large number of engine cycles (> 1 h engine operating time, >10^4 cycles).

To achieve an acceptable wall clock time of less than a week, a very strong scaling capability of the code is necessary. This code will be used only on exascale systems.

- ***Coupled CFD-FEM-Multi-body simulations***. The pressure on the piston couples the combustion simulation to the simulation of the mechanical system including piston, conrod, crank shaft and bearings. Only such a combined simulation will yield detailed information e.g. on vibrations of the coupled system and can prevent design changes late in the development cycle. Similar to ***coupling of different codes*** on the CFD side for combustion and acoustics, ***this approach requires a flexible interface to achieve simulation of such multi-physics problems***.

Crash:

The expected usage of Exascale applications in the field of crashworthiness is described in the following. On the one hand, the estimate is based on the current situation although the performance of the currently employed commercial software codes is rather different and the usage in the car industry is also not very homogeneous. On the other hand the forecast is done by projecting current trends into the future (ca. 2020). Finally there are some project areas which have to be addressed.

The current situation is as follows: today, about 15 types of crash tests are virtually assessed with between 5 and 600 different assessments per type. The total number of crash simulations for one vehicle is in the range of 10 000-50 000.

The wall clock time for these computations is between 3 hours and 36 hours for a single run depending on the software used.

Currently most of the computations are done in parallel (8-64 cores).

Scalability tests have shown that up to 1024 cores may be reasonable today.

The model size for a full car ranges between 1.5 and 10 million finite elements. Crash software is mainly commercial; no appropriate open source is available.

Open source must be developed for Exaflopic systems using as farming applications.

- ***coupling*** is already very important; some companies perform a standardized mapping between manufacturing simulation and crash simulation.

This coupling approach will be improved in next applications.

- ***optimization and stochastic analysis*** is already used for some cases, which increases the amount of computations by a factor of ca. 1000. Shape and topology optimization starts to be used commercially.
- ***first multi-level computations*** are tested in research and in industrial applications the so-called sub-cycling is used where more detailed parts of the problem are treated on a specially dedicated group of cores.
- in general, crash simulation is already well embedded into a simulation ***data management*** system with automated pre- and post-processing including monitoring and coupling to other fields and functionalities.

Nevertheless this needs further improvement. The model generation still requires remarkable manual work.

For the 10-year perspective, the following main challenges have to be addressed:

- true virtual testing replacing some physical tests requiring reliable computations
- handling of a much higher complexity of finite element models (new materials, human models instead of dummies, etc.); these new materials require better and more efficient/stable algorithms. The human models have to be improved, stochasticity should be included.
- assuring that the overall computational wall clock time remains constant (ideally ca. 8 hours).
- addressing true multi-disciplinary and multi-physics simulations including optimizations and stochastic analysis; this will lead to a factor of >1000 for the required number of computations compared to the current situation and the necessity to embed all simulations in an overall simulation data management.
- establishing robust topology and shape optimization for crash including meta-modeling techniques for fast coupled multi-disciplinary analysis.
- multi-level simulations where some local effects (e.g. failure) are studied on the meso-level in parallel to the overall macro-computation, which might be realized based on hybrid parallelization schemes.

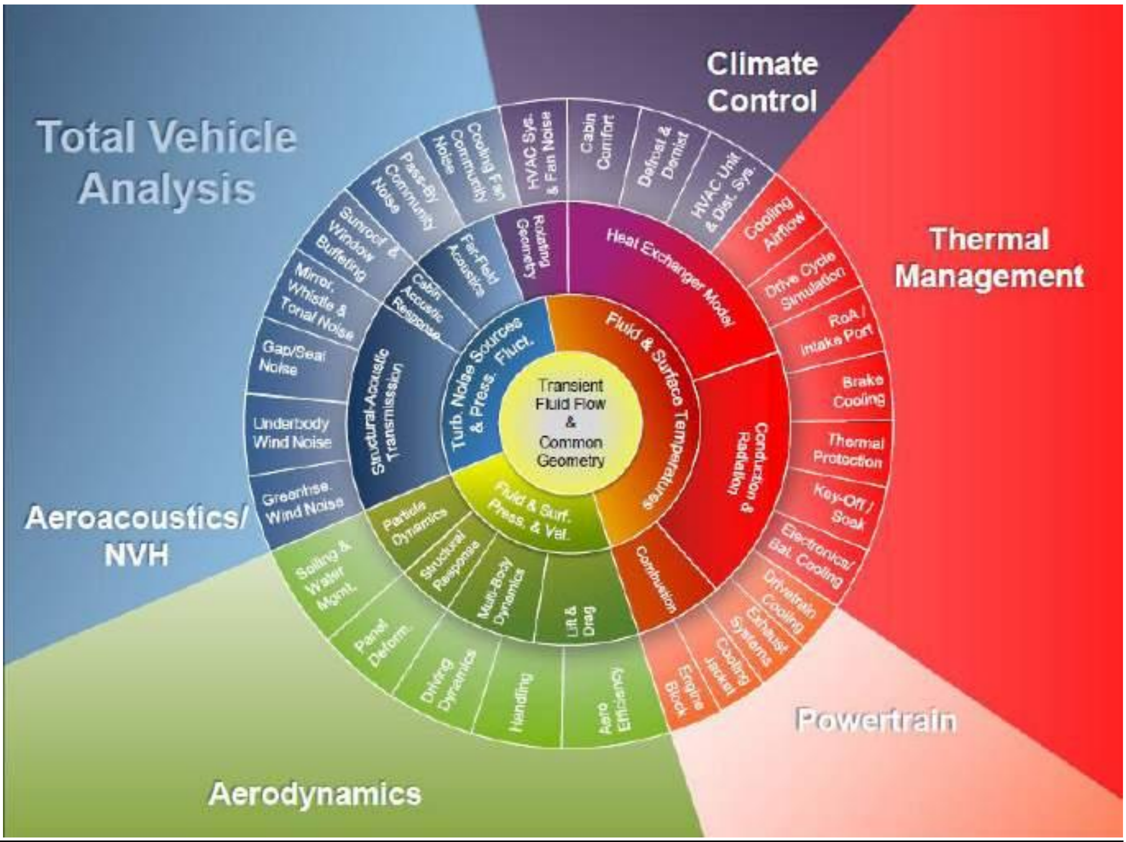
Addressing these challenges requires project in the following areas:

- due to the ***complexity and high non-linearity of crash simulation, it should be difficult to progress on a strong scalability***. A limit estimation could be between 64 (current standard) and 2000 cores for next crash simulations. ***Farming applications should run on Exaflop machine.***
- the ***memory*** is currently not an issue because of the explicit FE method, but it will be become more important especially because of the trend of coupled simulations where a high amount of data need to be mapped from manufacturing simulation to crash simulation.
- improving automated ***pre-processing*** (e.g. meshing of 3D objects, coupling between CAD and CAE, unified geometrical modeling by isogeometric analysis, parameterizations for sensitivity and optimization studies).

As for numerous other industrial applications in Energy domain, the following methods must be addressed for an ***Exaflopic crash test application***:

- More efficient algorithms for stochastic modeling.
- More efficient algorithms for shape and topology optimization.
- Establishment of a uniform approach for CAD and CAE (and other CAx).
- Improved material models for soft tissues (human model), composites, honeycomb structures, multi-material light weighting.
- Algorithms for multi-level analysis for composites and other new lightweight materials where a coupling between manufacturing and crash simulation is realized.
- Algorithms for multi-physics (especially for electric cars) and multi-disciplinary simulations.
- New techniques for parallelization to improve scalability (based on sub-cycling or other approaches).
- Robust meshing techniques for 3D modeling, which can be used during simulation to enable shape optimization and adaptive multi-level computation (for example for failure analysis). Adaptive meshing is still not realized.
- Fluid-structure interaction (simulation and optimization). Simulation of combustion and pollutant emissions in combination

Thermal exchanges



Heat exchange in a car and air conditioning are actually a large subject of competition. Very low air flows in the cabinet are more and more required.

Very complex CFD simulations, coupling several physics are necessary for achieving these goals, with specific meshes all around internal “obstacles”.

In another hand, the global energy efficiency is also linked to better heat exchange in the power train and will be completely controlled by a “super” management.

All these applications will be feasible thank to new computing capabilities.

The above figure shows how diverse are the potential simulations on thermal exchanges in the automotive industry, coupling fluid-structure simulation and several physics.

There are currently developments to establish mesh-less methods for thermal automotive applications and this might offer benefits in terms of parallelization. This could be an interesting alternative to established approaches using mesh based methods. For the latter, the definition and the dissemination of coupling tools and interfaces are considered crucial as a next step as discussed in the report in a separate section.

4.3.5 Chemical Engineering, General Engineering, Flows

In these different scientific domains, the main problem in running simulations on super computers is the **difficulty and often the impossibility to parallelize physical equations themselves**. For example, mechanical calculations (deformation, etc.) or viscous multifluids flows in porous media are governed by stochastic PDE (Partial Differential equations) in which discrete decomposition is not largely hierarchical.

So, in these domains, applications will be mainly “farming” ones.

Nevertheless, Exaflop systems will allow to reducing the importance of surrogate models and by this way to lead to more and more accurate approximation, and hopefully allow to **simulating phenomena at the right industrial scales.**

Exaflop systems will also permit a large development of **particle simulation,** for describing **multi scale interactions** fluids structure, or fluids solid suspension, interfaces, and not least **multi physics coupling** (as described above).

In fluid-solid suspensions simulation, as example, at the industrial process scale, on the order of a billion (10^9) or more particles may be suspended in a fluid. Hydrodynamics and chemical equations lead to solve very large matrices which can only be solved by an iterative manner on an exa-scale computer, using finite elements method or particle (Lattice Boltzmann) simulations.

A large number of industrial processes are produced in such reactors, Fluidized Bed Reactor, -in refining and petrochemicals plants, in biomass activities, in chemistry, in agro-industry, etc …- and an optimization of these processes is expected thanks to Supercomputer.

Chemical Engineering is clearly the major domain of multi scale simulation: a chemical reaction is produced in the pore of a catalyst (nano/micro scale), this leads to a transient move of atoms, and so a “micro” flow around porous grains, the non stochastic part of this flow will lead at the macro level a global quite uniform flow of blending products. The physics at the nano scale and at the macro scale

are different! Micro actions can lead to a macro behavior which is not known at micro scale. For example the viscosity is only a macro behavior of many fluids. At the nano scale, one can speak only friction forces. Nevertheless, a continuum function must adjust the different scales and in particular the meso-scale, that is the **upscaling function.**

The real problems of multi-scale simulations are to determine what is the meso-scale and to define the right upscaling function.

Large efforts should be done in this area, to lead at a relatively good understanding of the way from the nano to the macro scale. And new computing systems will help.

So, in these domains, applications will be mainly “farming” at short term but could be hero at mid term.

Other problem formulations describe fluid and solid physics in transient formulations. Such applications may be of a more local character and thus enable numerical methods like stream computing. This kind of method is suitable for also the later hardware designs with graphics card based computing, GPGPU’s.

Of course, data management, data assimilation and experimental validation, are, as for previous sectors, important items to be examined.

4.3.6 Computer Aid Engineering, Remarks on complexity

Rapidly growing complexity in all spheres of social life is seen as the major threat to sustainable development. This may be said of the economy as well as of the products of the manufacturing industry which are becoming increasingly sophisticated both from the standpoint of design as well as operation. Recent advances in mathematics have made it possible to actually measure complexity and treat it as a new attribute of systems just as energy, mass or temperature. With the newly developed metrics of complexity it is possible to envisage incorporating complexity in the design process of all sorts of products, ranging from software to aircraft, from satellites to automobiles or traffic systems. This is important for the following reasons:

- high complexity is a formidable source of fragility. This is intuitive. Highly complex systems and products have the capacity to deliver unexpected behavior. As systems become more complex they become less resilient and can fail unexpectedly. High complexity comes, therefore, at a very high price.
- high complexity is also reflected in the high costs of designing, engineering and operating new products. Using complexity as a design attribute/objective it is possible to design products and processes that are simpler hence less expensive and less risky.

In order to build new and advanced products it is necessary to involve in the design process multiple interacting disciplines. However, combining different disciplines and different scales increases complexity significantly. In order to keep things under control, it is necessary to keep an eye on complexity. Complexity, therefore, has the potential of impacting engineering quite significantly. Designing for lower complexity can lead to tremendous savings in terms of manufacturing costs and liability and therefore to reduce the impact of new products on the ecosystem. However, all this hinges on the ability to actually measure complexity and make it available to engineers and scientists. Today this is possible.

The computation of complexity and its participation in the design process is very expensive from a computational point of view. There are essentially two means of generating data which is necessary for the quantification of complexity:

- Monte Carlo and other meta-computing approaches such as DOE, Optimization, etc. (known as farming) whereas a certain number of independent simulations is run on a large number of processors, provide data which may be used to compute complexity. As computer models grow in size and detail, running hundreds or thousands of simulations can be prohibitive. Therefore, today this can be done only for simplified models.
- A single computer run may be used to provide data necessary for the computation of complexity and robustness. However, the simulation must be in the time-domain. When it comes to modeling phenomena such as life-sciences, climate, CFD, multi-physics or multi-scale phenomena, the computations are immensely CPU-intensive. This can only be achieved with compute power that is orders of magnitude beyond that of today.

It is important to remark, however, that both of the above approaches have been demonstrated with relatively small models. The technology, therefore, is here. What is needed is compute power. Once complexity is incorporated in the design and engineering of new products and processes it will allow us to take the manufacturing industry to the next level, beyond the contemporary simplistic techniques of CAD and simulation which, in their present form, simply automate old ways of thinking. The social, economic and ecological impact of this technology is potentially immense.

5. Industrial & Engineering Applications Scientific Roadmap

5.1 Way to Efficiency of Large Scale Computational Software for Industrial Applications

The current basis of technical simulation software, as it is used by industry, suffers from a dramatically increasing performance gap. The performance that could be expected theoretically is very often not reached in practice by far. The current trends indicate that this problem will increase in the exascale environment, making coordinated research efforts in algorithms, application software, hardware, and systems software, of the utmost importance for the effective use of exascale systems.

Many existing algorithms are currently not implemented with optimal efficiency. Again, we expect that this will become even more dramatic when we move onwards to exascale systems. Additionally, it should be clear that the improvement potential is even larger, when more substantial adaptations of the data structures and more intrusive code restructurings

5.1.1 Key Issues of the Roadmap for Industrial Applications

As mentioned before, WG3.1 Experts consider that there will be only few “Hero” applications to push the limit, and for these, the real goal is more Zetaflop rather than Exaflop computers.

A majority of industrial applications corresponding to production problems are “farming” applications.

The first element of the roadmap is

Some flexibility in architecture in order to allow the possibility to compute**:**

several applications on the same large computer (computer center)

and, the same application on different architectures (true for farming only)

Then, **5 important items must be addressed:**

- Co-existence of (new) **common software platforms** and **legacy codes**
- Multi-physics simulation: establishment of **standard coupling** interfaces and software tools
- Development of an **automatic and highly parallel grid-generation** tool for Exascale applications
- **Standardized efficient data management**
- **Common numerical improvements (scalable algorithms)**

5.1.2 Unified Simulation Framework and associated services

Advancing individual solvers performance is not enough to bring high performance simulation to the end-user. Each community needs a much **broader set of tools including in one platform** in order to conduct industrial studies: CAD, mesh generation, data setting tools, computational scheme editing aids, visualization, etc.

Such platforms, like Salome developed by EDF in collaboration with CEA, or the FlowSimulator software developed by EADA and its research partners, must be developed (or adapted) for **massively parallel computing** and available for all users, with all integrated toolboxes. It is necessary for the following aims:

- reduce the cost of complex simulation platforms by **mutualizing a set of common tools**: pre and post-processing, calculation distribution and supervision etc.
- boost performance through **easy integration of multiple solvers for muti-physics studies** (via a common data model). This point is very important for achieving **standard and sustainable coupling** tools (see below).

Building a Unified Simulation Framework and associated services adapted to massively parallel simulation need following efforts:

- *Common data model for data exchang*e: designing a common data model and associated libraries for mesh and field exchange adapted to massively parallel computing would enable interoperability and the coupling of independent parallel scientific codes. High level operations on simulation data, such as mesh projection, data interpolation, could be implemented on top of this model.
- *Meshing:* Generating x10^{10} cells mesh as targeted in 2015, requires future meshing tools to provide parallel meshing, automatic hexahedral meshing, mesh healing, CAD healing for meshing and dynamic mesh refinement ***(see specific chapter grid generation).***
- *Parallel visualization tools:* Considering the volume of data that will be produced by Exaflop computers, end users are needed adapted parallel visualization tools and specific clusters to post-treat their simulation results. The international scientific community would benefit in focusing their research efforts in few software, (VISIT and Paraview seem two good candidates).
- *Remote and collaborative post-treatment:* the sheer volume of data produced by Petaflopic/Exaflopic calculations, storage and network limitations, and multi-sites teams make it necessary to further advance R&D on remote and collaborative multi-user visualization, parallel and distributed file systems.
- *Workflow and code coupling management*: new generation of workflow manager, such as YACS (EDF/CEA), have to be developed to handle parallel multi-physics coupling scheme (see specific chapter on coupling).

The platform involves at least one **legacy code** and the coupling using an external tool is as less intrusive in the legacy codes as possible.

5.1.3 Standard coupling interfaces and software tools

As already mentioned, for using common platform, standards of coupling must be defined and software (codes) must be developed with these standards in, of course, an efficient manner. **Open source multi physics coupling** for massively parallel applications are one of the key challenges of “Exaflop” software.

Couplers **YACS (EDF/CEA) and O-Palm (Cerfacs),** already existing, have to be improved for massively parallel systems, but seem to be good candidates.

As external couplers, they **are and will be not so intrusive** in legacy codes.

On the other hand, the advanced coupling algorithms for all multi-physic simulation will be shared in a dedicated algorithmic box in the platform. From an algorithmic point of view, the existing couplings are mainly explicit and semi-implicit (fixed point algorithm). Works must be performed to improve methods used in these algorithms (such as Newton-like algorithms).

5.1.4 Development of a Grid Generation Tool for Exascale Applications

Automatic grid generation and grid refinement has been recognized as a limiting factor in the simulation in a large number of technical applications. Especially with respect to future Exascale computations, either for very large grid of complex geometries or a large number of computations for different geometries, the availability of a flexible and highly parallel grid generation tool is of utmost importance.

Thus, it is suggested to jointly develop the next generation grid generation tool. The development process is divided in two phases, namely phase 1 from 2012-2015 (core development and selected technical applications) and phase 2 from 2016-2020 (establishing the software for wide use). Based on the currently available commercial and non-commercial technologies, initially in phase 1 it must be decided whether to work with a company or an university or a combination of both as a core team in the tool development.

For phase 1, a core development team is established. This core team consists of grid generations experts responsible for the tool itself on the one hand as well as a first user group. The user groups initially identify current bottlenecks in the grid generation process in their scientific area of expertise and continuously evaluate the developed software for their specific research applications. Two scientific areas are identified as ideal user groups for the first phase, namely aeronautics and internal combustion engines, respectively. Both research fields depend on high quality grids based on very different requirements. Aeronautics, very often, use static grids focusing on highly resolved boundary layers and wakes at very high flight Reynolds numbers for complex 3D geometries. On the other hand engine flows require meshes for moving geometries and strongly varying conditions such as combustion or spray formation. Both areas usually use hybrid meshes with hexahedral/prismatic and tetrahedral volumes but can also work with arbitrary elements. More details are given below for aeronautics and internal combustion engine applications. In the second phase, other important areas, e.g. hydraulic and thermal turbo machinery with moving frames of reference, are included in the scope.

Aeronautics

Automatic grid generation for the complete aircraft at flight Reynolds numbers including high-aspects ratio grids for boundary layers and wakes is a key issue for numerical simulation. The starting point for stating the challenges of unstructured grid generation are today's state-of-the-art methods using advancing-layer/advancing-front techniques based on unstructured, anisotropic, quad-dominant surface discretization, resulting in hexahedra-dominant near-wall layers with subsequent tetrahedral-based far-field resolution. The stretching ratio of the surface elements should increase where possible from currently 1:20 to 1:100 or even 1:1000, to decrease total grid point counts. A successful transition to the far-field can be achieved through the use of anisotropic tetrahedral, which can also be deployed to resolve expected free shear layer paths, i.e. viscous wakes of high-lift devices. A paradigm shift may be required away from a pure near-field O-grid topology towards a combination of O-, C- and H-type topologies, for being able to match the field discretization to the expected flow features, such as free shear layers or boundary layer interactions in sharp concave surface junctions. A-priori knowledge of the problem at hand should be accounted for to automatically generate adequate grids.

Internal combustion engines

Automatic generation of hybrid prism/tetrahedral grids with embedded hexahedral type regions for moving geometries is crucial for the simulation of engine in-cylinder flow and combustion. In order to simulate the boundary layer flow on the valves, the piston and on the liner, prism layers with high resolution in wall-normal direction are extruded from the surface. The transition to the far field is achieved using anisotropic tetrahedral. Especially for small valve lifts, for the valve gap region an embedded hexahedral mesh, which is properly aligned with the flow and the valve motion to minimize the numerical error, is required.

Due to the moving geometries piston, intake and exhaust valves, respectively, automatic remeshing (in communication with the CFD code) is required, which also includes different grid topologies. For example after exhaust valve closing, the exhaust port becomes a separate independent computational domain.

Adaptive mesh refinement and subsequent coarsening is required for several processes and flow regions. Especially important for internal combustion engines are free shear layers in LES, spray and mixture formation for direct injection and moving combustion fronts.

General requirement for the grid-generation tool

For both research applications (and this holds for most other applications such as gas turbines as well), adaptive mesh refinement should work using either a-priori knowledge from an existing solution or online based on the current results of the CFD solver. It is important that the grid generation software is able to process the numerical results in terms of physics for automatic refinement. This requires a flexible interface between simulation and grid generation software, which can be used by the simulation researcher to define criteria for grid resolution based on physical arguments instead of manually determining the necessary grid size. This requires the definition and implementation of a standard interface to CFD codes. It is important to note that many research areas will focus on LES, which requires meshes with low stretching ratios in contrast to RANS. The choice of the two fields aeronautics, which will use RANS also on Exascale systems, and internal combustion engines, which uses both LES and RANS, respectively, ensures that both grid generation strategies can be used for Exaflop computations.

Exascale CFD applications will rely on the large number of available cores. Since the corresponding grid sizes will increase by one to several orders of magnitude, the grid generation tool must also run efficiently in parallel (both CPU and RAM) in order to handle grids of that size.

The grid generation tool is expected to achieve a first mature level at the end of phase 1 in 2015. Thus, for phase 2 from 2016-2020 the tool will be made available for a larger number of groups from different areas (chemical engineering, structural mechanics, etc.) to further evaluate and improve the software.

5.1.5 Data Management

The experts in all industrial applications pointed out that exaflopic systems will give rise to very large quantities of data. This data can be generated as simulation output, but additionally increased amounts of data will be generated from sensors and automated measurement devices. All these data must be managed in massively parallel systems, requiring not only computation but also I/O parallelization and intermediate storage management. New concepts will be needed to handle and organize these data, including powerful new storage techniques, data analysis tools, and visualization techniques.

Data placement and memory access are two of present limits of many cores applications. So, for Exaflop runs, following items must be addressed:

- sorting memory for fast access
- allocating new memory as needed in smaller chunks
- splitting needed from unneeded memory
- treat parts of memory that are rarely/never needed based on heuristic algorithms
  - with compression
  - with writeout to disk
  - deletion of memory, giving reference to recalculation if ever needed
- optimize memory access for points not discretized in database supplying
  - built-in interpolation schemes
  - local database refinement (change in discretization/resolution)

Efficient data management libraries will be supported in a similar fashion as the software platform.

5.1.6 Common numerical improvements

Of course in the core of Exaflop codes all components must be optimized for a massively parallel use and by this way lead to the best possible scalability.

Remember that what is expected is the best coupling between architecture, algorithm, application, and so, following issues are important:

Multi-core with reduced memory bandwidth per core is identified for CFD applications and many others as an important source of inefficiency. Hardware or/and software tools should propose solutions to that hurdle.

Collective communications are becoming a more and more important source of inefficiency. If their implementation is not redesigned significantly to address Exascale capacities, new algorithms that do not rely on them will have to be designed and implemented.

Efficient **parallel IO** for post-processing, possibly allowing interaction with the simulation, and scalable management of large data sets is also mandatory.

Programming Exascale computers needs a **complete set of standard revisited software tools**:

- Efficient, “easy to use”, portable and **fault tolerant implementation** of Libraries, mathematical libraries /Languages (Fortran, C++, …) /**compilers for mixed parallelism** : MPI/OpenMP/”cuda like” languages
- Languages/compilers/performance analysis tools for achieving single core high performance, especially with accelerated compute units

New numerical methods and solvers are critical issues for all engineering domains. Numerical methods for solving more complex Wave Equation formulation, multiphase flows, mechanical structure, adaptive methods for heterogeneous platforms (may benefit all of computation/visualization/meshing), hybrid solvers, advanced numerical acceleration techniques, multi-grid, better and simpler pre-conditioner for CFD simulation are key issues that have to be solved.

Numerical scheme, more and more based on efficient “Nested” decomposition (hierarchical blocks decomposition) have to be implemented. Deep algebra libraries, like BLAS 3, are to be revisited and adapted to increase the level of parallelism of the discrete decomposition.

The coupling between stochastic and deterministic methods is also crucial for solving stochastic optimization problems. Dealing with more and more unknowns induces algorithmic convergence issues, requiring more computing iterations. This motivates the design of more efficient acceleration techniques. More generally speaking, a new generation of scalable, **fault tolerant algorithms** has to emerge.

Numerical scheme must involve Stochastic HPC computing for uncertainty and risk quantification. Probabilistic quantification of the risks and uncertainties affecting a best-estimate model generates a whole domain of applied science**, linking probabilistic, numerical analysis as well as physics and decision-theory.**

Beyond the traditional **Monte-Carlo** sampling, a number of uncertainty propagation and probabilistic simulation algorithms have been and have to be developed, such as **accelerated sampling** (importance sampling, particulate methods etc.), **reliability techniques** (FORM-SORM etc.), **stochastic developments** (e.g. **chaos** polynomials) and **response surface techniques**, yet still waiting for further development particularly regarding the challenges of low probability estimates for irregular response or high input dimension for sensitivity analysis/importance ranking or high-volatility time series.

Beyond **uncertainty propagation or risk computation**, even tougher challenges come with the need for inverse probabilistic techniques as the observable data to calibrate model variability generally comes on parameters different the model inputs, so that the identification of the extent of uncertainty affecting its input parameters requires the use of inverse techniques coupled with stochastic simulation. Closely related is the need for a general coupling between stochastic optimization and simulation in order to strike robust design or operational management strategies, with challenging mathematical implications that are only partially solved under existing Expectation-Maximization or stochastic dynamic programming algorithms (typically limited to close to Gaussian/linear behavior).

Bayesian settings are also bound to develop to better incorporate expert knowledge in a solid decision-theory foundation.

Beyond the development of the methods itself there are key implications on the way HPC is structured and used: Monte-Carlo sampling leads to straightforward massive distribution as the runs are all fully independent, but the other kinds of stochastic computing algorithms do need back-and-forth links between the various runs involved in exploring the stochastic space. Research must be done to improve them for massively parallel use.

6. Societal benefits, Societal, environmental and economical impacts

High Performance Computing (HPC) is a key enabler for economic growth in Europe today. It delivers a competitive edge to companies operating in the global marketplace allowing them to design and produce products and services that differentiate them from their competitors. A study in 2004 by IDC research of HPC users found that almost 100% indicated that HPC was indispensible for their business.

All of us experience the effects of HPC in our day-to-day lives, although in many (and probably most) cases we are unaware. We travel in cars and aeroplanes designed using modeling and simulation applications run on HPC systems so that they are efficient and safe. HPC is essential for ensuring that our energy needs are met. Finding and recovering fossil fuels require engineering analysis that only HPC can deliver. Nuclear power generation also relies heavily on HPC to ensure that it is safe and reliable.In the coming years HPC will have an even greater impact as more products and services rely on it.

Globally, nations are investing in HPC to tackle some of these issues. In the 1990s, the USA stood out as world leader in HPC, with Europe and Japan the other major players. Now, countries including India, Russia and China are undertaking ambitious HPC programs. According to the IDC report "An Agenda for Strategic European Leadership in Supercomputing" Europe has lost ground by 10% since 2007 in terms of HPC investment. Failure by Europe to increase its investment means that not only will it risk falling further behind the world leader, the USA, but worse, it may be threatened by emerging HPC powers.

Besides the above benefits that industry can draw from HPC to achieve increased innovation rates and competitivity, HPC is also a critical and essential technology as we address some of the societal challenges ahead, such as generating clean and efficient energy, predicting and mitigating against the effects of climate change, and ensuring safe and efficient travel. Clearly, one of the most important challenge facing our world is to design and provide clean and climate-friendly transportation systems and energy-producing technologies, which would not lead to as-large-as-today emission of CO_2 and other greenhouse gases (nitrous oxydes, …). HPC is there essential in two aspects: (i) it is the unique way to study and design new processes, as analogic mock-up systems are getting unaffordingly expensive, if not impossible to construct (as, *e.g.*, for the design of low-pressure stable combustion chambers, of more fuel-economic terrestrial and airborne vehicles, …); (ii) it is the only way to check the real impact of these new designs through the use of advanced climate models.

6.1 In the energy domain

- As every human activity, energy process (nuclear, thermal, hydro, …) needs to continuously improve its environmental impact (thermal discharge of nuclear/thermal power plants, long term geomorphology in rivers due to hydropower, water or air quality, …). Achieving this environmental risk assessment will also require the use of multi-physics (sedimentology, water quality), multi-scale (from local scale of near field to the large scale of a river basin), and complex multi-dimensional, time-dependent models (sometimes to simulate 5 years of "may-be" evolution.
- The main challenge facing nuclear industry is, today more than ever, to design safe nuclear power plants. This is presently being done for the so-called 3^{rd} generation (e.g., the French EPR), and is under active preparatory studies for the forthcoming 4^{th} generation (e.g., sodium- or gas-cooled fast-neutrons reactors). HPC and in particular Exaflopic computing will contribute to improve nuclear power plant design, involving the use of multi-physics, multi-scale, complex three-dimensional time-dependent models.- the easily accessible oil reserves

are decreasing rapidly, with an oil-peak forecasted to occur sometime around the middle of the century. Improving the efficiency of the search for new oil reservoirs, including non-traditional reservoirs, can only be done through very advanced wave-propagation models and image processing methods. Such methods are being considered by oil companies as giving raise to crucial competitive advantages. Equally, the production from non-traditional fields, like marginal or deepwater fields, is often characterized by oil and gas qualities, which are more difficult to cultivate. Flow from wells to production plants needs particular attention and engineering abilities; handling of produced fluids is more complex. Energy supply can only be safeguarded with extended capabilities in reservoir modeling and enhanced techniques for production; e.g. processing and transport of crude oil streams in harsh environments. Increased levels and quality of computer simulations enables engineers and scientists to turn potential risks into well managed opportunities for the energy consuming society.

- It is worth noting that HPC 3D-simulations are crucial to EDF for assessing the 10-year extension of the lifetime of nuclear power plants, such an extension representing hundreds of millions of euros of saving each.
  Another example concerns the costs of a drilling by oil companies, which amounts to approximately several tens of M$; avoiding an unproductive well, thanks to extensive HPC analysis of seismic data, is clearly worth the cost !

6.2 In the chemical engineering domain

Chemical industry heavily relies on oil as a carbon source. With the prospect of increasing oil prices and decreasing oil reserves, gasification (with its low CO2 footprint) of low grade coals and biomass providing synthesis gas for further chemical processing has become an important technology. Gasification with reactor sizes of several meters, residence times of 10-100 secs and physical processes occurring on micrometer and millisecond scales is a challenging multi-scale, multi-physics problem. Future (virtual) design and optimization of these reactors will be driven by numerical simulation. Developing for a large variety of different feedstock requires a large number of simulations. Each single full 3D and unstationary simulation requires efficient numerical models and the corresponding computational resources.

6.3 In the transportation domain

- Future air transport systems will have to meet the constantly increasing needs of European citizens for travel and transport, as well as the strong requirements to preserve the environment and quality of life. Within the ACARE (Advisory Council for Aeronautics Research in Europe) Vision-2020, ambitious goals have been set for air traffic of the next decades. These include a reduction of emissions by 50% and a decrease of the perceived external noise level by 10-20 dB. Continuous improvement of conventional technologies will not be sufficient to achieve these goals; however, a technological leap forward is required. Numerical based design and flight testing will be a key technology when aiming at more affordable, safer, cleaner and quieter and hence greener aircraft. The access to high performance computers in the Exascale, if not Zetascale, range is of utmost important. Considerable changes in the development processes will lead to significant reduction of development times while at the same time including more and more disciplines in the early design phases to find an overall optimum for the aircraft configuration. This enables the European aircraft industry to keep a leading role in world-wide competition, facing both an old challenge, *i.e.* competing with the US, and a new, rapidly-emerging one, *i.e.* keeping an innovation advantage over China.
- Designing efficient engines, motors and reactors, for airplanes and cars is critical for both the efficiency and safety of actual and future propulsion modes. This challenge is facing advanced European industries: how to use less energy for propelling the new vehicles presently under development, with the corresponding environmental challenge of emitting less greenhouse gases? How to design safe reactors, which could perform at low pressure without exhibiting dangerous instabilities?

7. Relationship with crosscutting issues

Industrial Energy and Transport applications on Exaflop machines share multiple common issues with the other EESI fields: parallel I/O, automatic mesh refinement, parallel domain decomposition, fast algebra solvers, compilers (including hybrid architectures), debuggers, etc… They also have a few specificities:

- Programmability: many codes in WG3.1 correspond to joint efforts of very large groups, disseminated in laboratories and companies but all participating to developments in the same code. Being able to program as simply as possible will be a key condition for success and a high-level API hiding parallelism will be useful.
- ***Load balancing***: this question is a central issue for codes to run on exaflop machines. It is linked to the problem of programmability: languages hiding parallelism will generally offer reduced performances on massively parallel systems. Finding a compromise between programmability and load balancing will be a key question for many codes developed in WG3.1.
- ***Fault tolerance***: ***for some specific applications*** fault tolerance strategies could be developed ***within the codes themselves*** as required by seismic industrial applications. But, *for a large number of codes, it makes more sense to handle these strategies in a common way, developing software which will handle issues* linked to fault tolerance *independently of users.*
- ***Power management***: this question should not drive the architecture of the hardware towards cores with small memories. Many codes targeted in WG3.1 are implicit and require significant memory sizes
- ***Data management***: Data placement and memory access are limits to be overshoot for quite all Exaflopic applications, such as optimize memory access for points.
- ***Performance optimization***: general purpose codes to optimize performances are needed for WG3.1. They should also be able to work for multiple codes coupled on the same machine (and not only for stand-alone codes).
- ***Reproducibility***: many phenomena studied in WG3.1 can exhibit chaotic behaviors. Repeating a simulation may lead to a different result because the phenomenon itself is chaotic. However, this might also indicate an error in the code itself. Reproducibility of results on future exaflop machines will be a mandatory condition for code development and validation. Most codes used in WG3.1 rely on very heavy management and validation of sources and being unable to repeat a simulation makes these exercises hazardous, constituting a significant risk for these projects.

Specifically, we suggest coordinated research programs in

- **hardware-oriented numerics**, including topics such as cache-aware algorithms, latency tolerant algorithms, fault-tolerance on the algorithmic level, power-efficient algorithms, massively parallel algorithms ($>10^6$ cores), GPU-acceleration of industrially relevant computational kernels, SIMD-vectorization, hybrid parallelization, etc.
- **software-engineering and performance analysis for exascale systems** (systematic performance modeling and performance prediction, tools and systems software, exascale enabled software frameworks)
- **new numerical techniques**, such as e.g. required by fully resolved multi-scale

8. International Context

For industrial applications, at the software and algorithm level, Europe has a lot of high-level research centers and institutes. Some of good candidates for Exaflop tools are already studied or in development in European companies or universities. European industrial strengths are developed in some detail below, but it can already be stressed that they are confirmed by the fact that a number of European industries do operate petaflopic-scale HPC systems, as it is thye case for ABB, AIRBUS, EDF, Shell, TOTAL,…

Europe Strengths are and will be certainly:

- Efforts on scalability ,

  Strong and weak scalability, **load balancing at the application level**, flexibility

  Numerical analysis, algebra, time scales, mainly thanks to mathematic academic research teams in Europe (Paris VI, Julich, Aachen, INRIA, Imperial College, …)
- Coupling multi physics codes with efficiency

  YACS, O- PALM, … in development at EDF/CEA and CERFACS.
- Stochastic HPC for uncertainties

  Robust optimization, uncertainties propagation in numerical simulation, numerical design at EDF, EADS, TOTAL, Saint Etienne Univ.,
- LES simulation improvements at CERFACS, EADS, IFP, Erlangen, Aachen, …
- European groups are leaders in the field of ***cycle-to-cycle variations observed in piston engines using modern LES tools*** and the European codes for combustion LES are recognized as the most advanced.
- DNS simulations could be developed by multiple groups in Europe.
- Simulation in Aeronautics leadership with EADS, DLR and academic partners
- Depth Imaging (seismic) in the world top with TOTAL and academic partners
- Numerical methods, algorithmic

Some of Europe strengths need to be reinforced:

- Automatic generation and dynamic mesh refinement
- New Particle simulation methods …
- Mesh generation
- Scientific libraries for fundamental algorithm
- Industrial end users and academics need to be better connected and contribute together to establish but also implement the Exascale roadmap.

For efficient Exaflop industrial applications, Europe has following weaknesses

- Not enough participation in de definition of new standards for programming / compilers (MPI, OpenMP, C++, Fortran, …)
- Data management
- Europe is lacking of strong and long term coordinated efforts dedicated to Exascale,
- Europe needs to increase the ability of the European HPC eco-system to commercialize tools and services, increase the reliability of the European research production

Potential collaborations outside Europe, who's doing/deciding what?

The competitive context of industrial applications makes difficult the establishment of a collaborative framework. Disciplinary research teams are already internationals.

Existing funded projects and funding agencies

Nothing specific to add from this WG3.1 to the general item.

Sources of competitiveness for Europe

Nothing specific to add from this WG3.1 to the general item.

9. Agenda

For each of important issue of research in the proposed roadmap, experts suggest following steps and timing:

2012-2015:

- decision process concerning the software
- establishment of a core development group
- funding of a small number (at least one) of pilot projects and establishment of the **tool framework** described above , **grid generation tool, common solver software**

2015-2020:

- establish first projects with both the software platform and at least one legacy code for the specific application. Choose a few HPC applications with scientific and social impact
- establishment of a growing number of research projects based on the common platform especially for multi-physics applications requiring very high memory
- Review of the approach until 2020

For example, below timing and milestones for the development of grid generation tool

2012-2015: Phase 1: Development of the grid generation tool

- Decision concerning the most appropriate commercial or open-source software platform and establishment of a core development group for the grid generation tool
- selection of two user groups (aeronautics and internal combustion engines) to identify future needs and bottlenecks in the grid generation process
- joint development and evaluation of a first version of the new software tool
- standardization of the CFD interface
- Review of the approach in 2015 and decision for phase 2

2016-2020: Phase 2: Establishing the grid generation tool

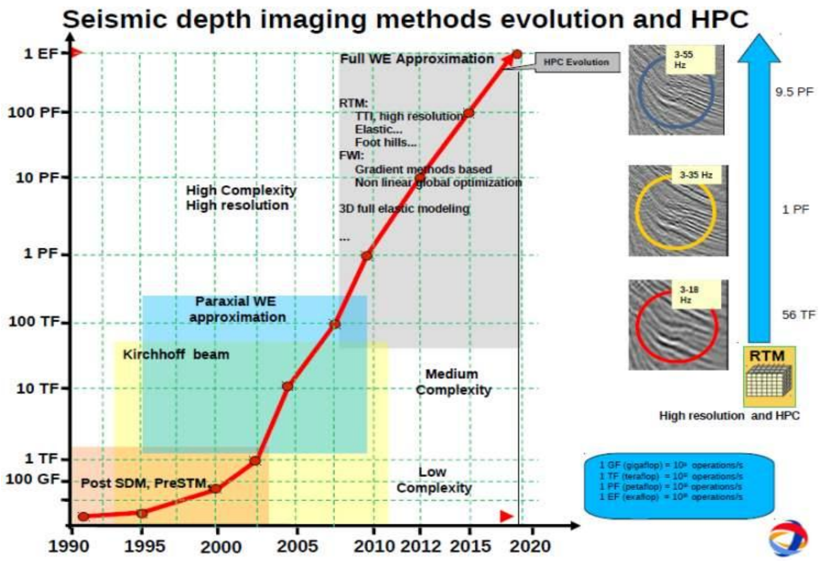
- based on the first mature version and the review at the end of phase 1, further research groups (hydraulic and thermal turbo machinery) are continuously invited to participate in tool evaluation and usage.
- Continuous improvement of the software based on user feedback done by the core development group
- Efficient parallelization for large grid sizes
- Simulations on Petaflop systems using the grid generation tool (especially for online grid generation and refinement)
- Public distribution of the software
- Simulation on prototype Exaflop machines

The grid generation tool is expected to achieve a first mature level at the end of phase 1 in 2015. Thus, for phase 2 from 2016-2020 the tool will be made available for a larger number of groups from different areas (chemical engineering, structural mechanics, etc.) to further evaluate and improve the software.

On the other hand, each industrial domain develops its own timing for the development of its many cores applications. Of course, these planning by industrial sector are completely linked to the development of research axes proposed in the roadmap, such as new linear algebra, new programming paradigm, new data management, new mesh generation, … :

For Oil & Gas:

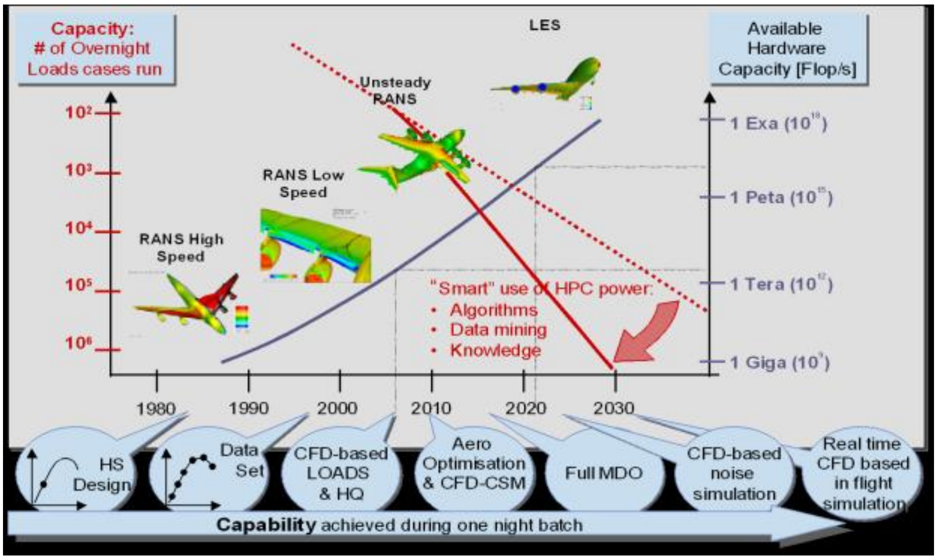
one can refer to schema already provided by TOTAL in the core of the document



We see clearly on this graph, the several steps from 2012 to 2020 in terms of methods of resolution of waves equation in relation with computing capacity increases: the complexity of the approximation is directly linked to computing capacities.

For aeronautics:

The following graph, from Airbus courtesy, could represent the necessary steps to “digital aircraft” with the fundamental methods as described in the core of the document. For this domain, experts remain that exaflop is not the ultimate goal and timing must be in concordance with zetaflop systems that means at least 2030.



For electrical power plant:

EDF, for example, presents in the core of the document its planning with such milestones: Full core simulations, and mesh sensitivity studies for these transient calculations may require unsteady calculations for meshes ***from 5 to 20 billion cells around 2016*** (before 2020) *which correspond to runs on 400 Pflops during few weeks.*

Running quasi-DNS type calculations on subsets of the calculation domain may be necessary. This will require meshes in the ***20-billion cell range by 2012*** (to study cross-flow in a tube-bundle, in a simplified steam generator type configuration), and running similar calculations for more complete calculation domains may require meshes well above ***100 billion by 2020***.

Such studies will require multi-Exaflop capacities during few weeks.

10. Needs of associated human resources and provisional costs

For each large item of proposed roadmap for Exascale Industrial Applications, experts tried to estimate what should be a “research” effort if EU wants to raise the technological bolts.

An important precision, in this document, only research effort is funded, and not the total development of complete industrial codes with debug, maintenance, …

Human Resources are given in Man * Year

Provisional costs are given in K€ with a yearly flat rate of 100k€/FTE

Roadmap issue	**Human Resources by year → 2015**	**Integrated (4 years) Provisional Costs 2012 → 2015**	**Human Resources by year 2016 → 2020**	**Integrated (4 years) Provisional Cost 2016 → 2020**
New Paradigm for programming model	20	8 000	30	15 000
Common Software Platform, legacy codes	30	12 000	80	35 000
Efficient Solvers, Numerical methods	25	10 000	90	45 000
Coupling Interfaces	20	8 000	35	17 500
Grid generation	30	12 000	40	20 000
Data management, Memory access	40	16 000	70	35 000
Fault Tolerance	45	18 000	45	22 500
Load Balancing	30	12 000	60	30 000
	240		450	
TOTAL in k€		**96 000**		**225 000**

11. Needs of education and training

The efficiency problem of numerical simulation software can be alleviated with providing suitable performance analysis tools, but it is obvious that the systematic training of code developers is even more important. The primary task is to educate developers of technical simulation codes in the design of hardware-aware algorithms and in the systematic analysis of the computational performance of their programs. This deficit will be detrimental to the practical industrial use of exascale systems.

Industry needs of course skilled personnel to build maintain and program exascale hardware. Human resources are the key element in grasping the competitive edge and added value of high performance computing, therefore education and training is mandatory for releasing the potential of the technology.

Generally education and training may be viewed in two distinct ways. The first way is the classical specialization on specific topics. For instance university courses on themes such as hardware design, compiler technology, numerical algebra etc. Courses at universities or research institutions are commonly organized in this way and are appropriate solutions for analysis or in-depth activities.

A second way to view education and training concerns a completely different aspect. The systems view. That is the capability to build up systems by joining different components or knowledge from different areas. The holistic view is typically necessary when solving real-world problems as found in industry. The implementation and deployment of exascale computing has been described with needs for “co-design” and “ecosystems”. These terms indicate the necessity for personnel with skills for systems -thinking and -building.

Classical education and training with focus on analysis and specific in-depth study is certainly both needed and appropriate. Here it is advocated, that it is highly desired to extend the formation with the complementary aspect of system -design and -engineering.

12. Communication plan, dissemination actions and identification of potential target public: R&D stakeholders, EC and national policy-makers,

Nothing specific to add from this WG3.1 to the general item.

13. Future Perspectives

As discussed in this report, the challenges facing European industries clearly show the needs for a continued but increased investment in HPC systems and development of software and applications codes. Transition to exascale requires a balanced effort between these various components. One way to achieve this is to gather around common objectives both the vendors, the software developers and the applications scientists and engineers. Such undertakings are presently being developed in the United States through the so-called "co-design centers" initiative. Europe has some experience in this field, *e.g.* in France where the Bull company and the defense department of the CEA have join efforts over the past ten years to construct systems and develop applications from the terascale to the petascale. There exist new avenues for such promising undertakings, by taking full advantage of the PRACE organization. PRACE indeed promotes the use of HPC for scientific applications by providing in a few countries high-performance systems which can then be accessed by all European scientists on a peer-review basis. It would be of highest value to fund and associate to PRACE a series of co-design centers, each of them addressing a particular domain (*e.g.,* for industrial applications: combustion, aerodynamics, seismics, materials, …).

The 3 present US co-design centers are dedicated to industrial physical domains, such as Combustion, Nuclear Plant and Materials.

They are clearly oriented to develop efficient many-cores applications much more than hardware definitions!

That also the recommendation of this WG for potential ones in Europe: multidisciplinary team focused on specific physics and mathematics for a specific industrial domain.