PRACE-3IP PCP

“Whole System Design for Energy Efficient High Performance Computing (HPC)”

Philippe Segers - GENCI

EAFIP Conference on Innovation Procurement

18th of October 2017 - Tallinn, Estonia

Outline:

- PRACE context: Europe within the global race for HPC
- Why Energy Efficiency matter ?
- PRACE PCP Goals and Process
- PRACE PCP Phase III early Results (project end December 2017)
- Summary & Conclusions

Simulation: th… …e

Traditional Scie…

1) Create a theo…
2) Perform exp…

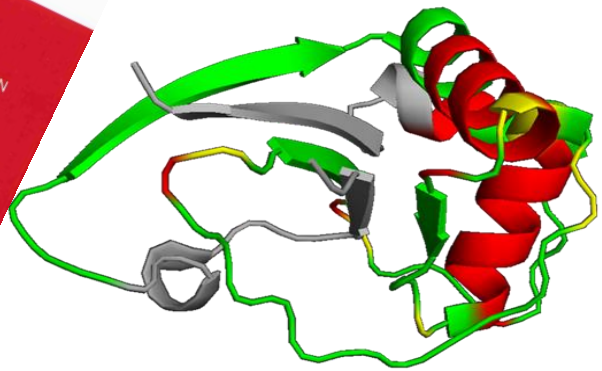
Constraints…

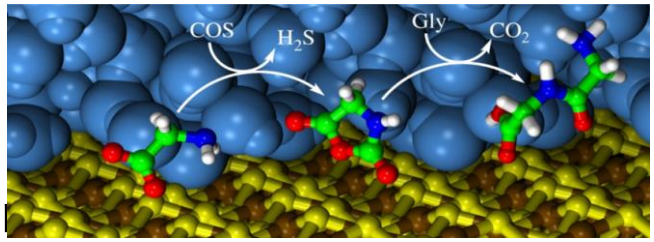
- Too difficu…
- Too expe…
- Too slo…
- Too d…

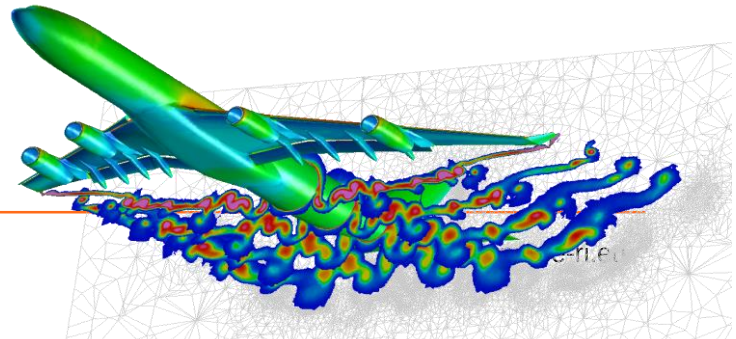
”Com…

3) U… …henomen…

… based o… …umerical methods

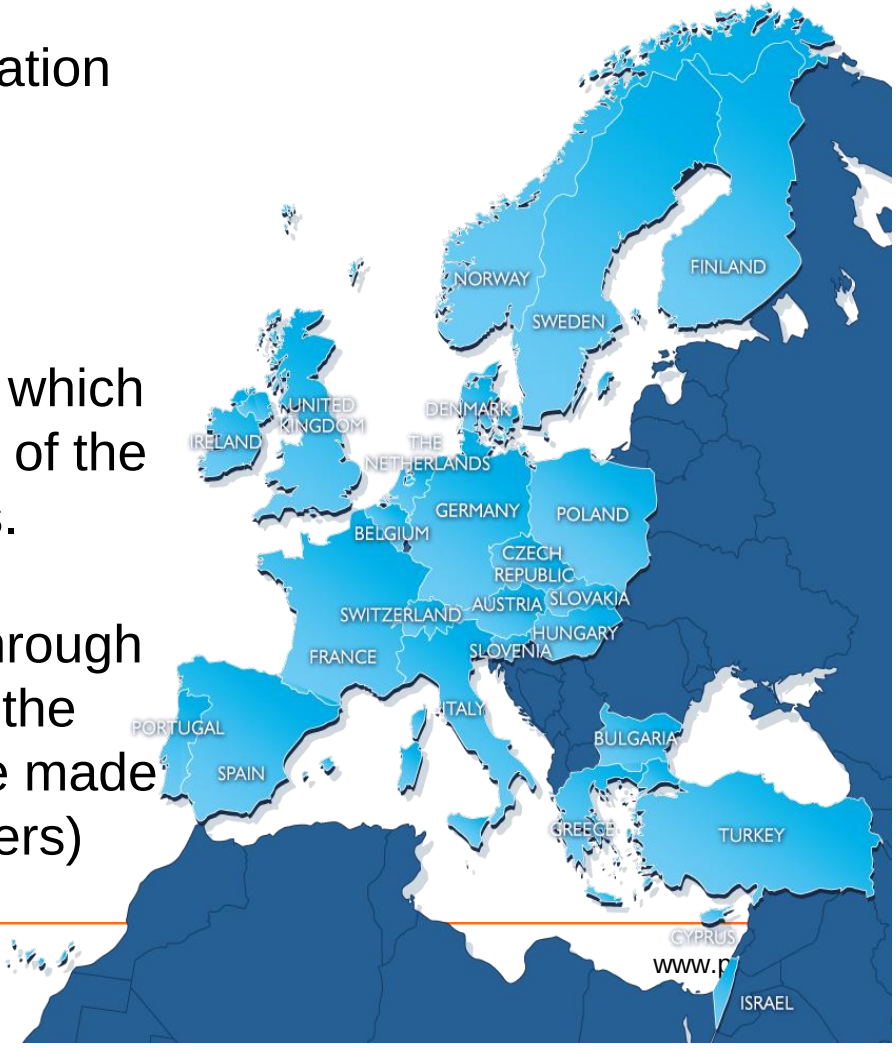






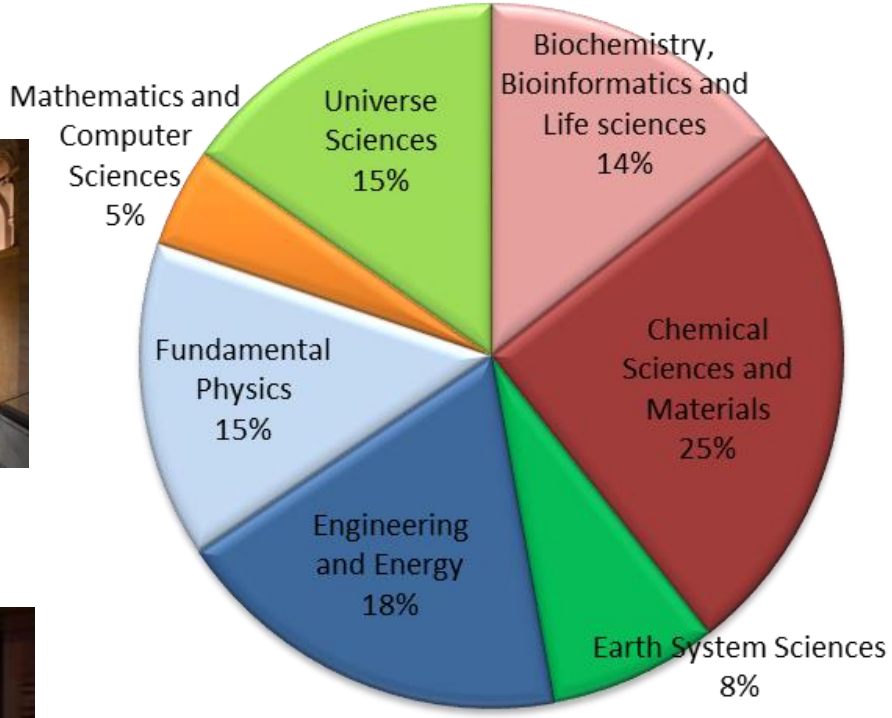
PRACE: Partnership for Advanced Computing in Europe

- **PRACE** is an international not-for-profit association under Belgian law, with its seat in Brussels.
- **PRACE** has 25 members and 2 observers.
- **PRACE** is governed by the **PRACE** Council in which each member has a seat. The daily management of the association is delegated to the Board of Directors.
- **PRACE** is funded by its members as well as through a series of implementation projects supported by the European Commission. Computing resources are made available by a group of members (Hosting Members)



PRACE offering of core hours on 7 world-class machines

MareNostrum: IBM
BSC, Barcelona, Spain



JUQUEEN: IBM
BlueGene/Q
GAUSS/FZJ
Jülich, Germany

SuperMUC: IBM
GAUSS/LRZ
Garching, Germany

Hazel Hen: Cray
GAUSS/HLRS,
Stuttgart, Germany

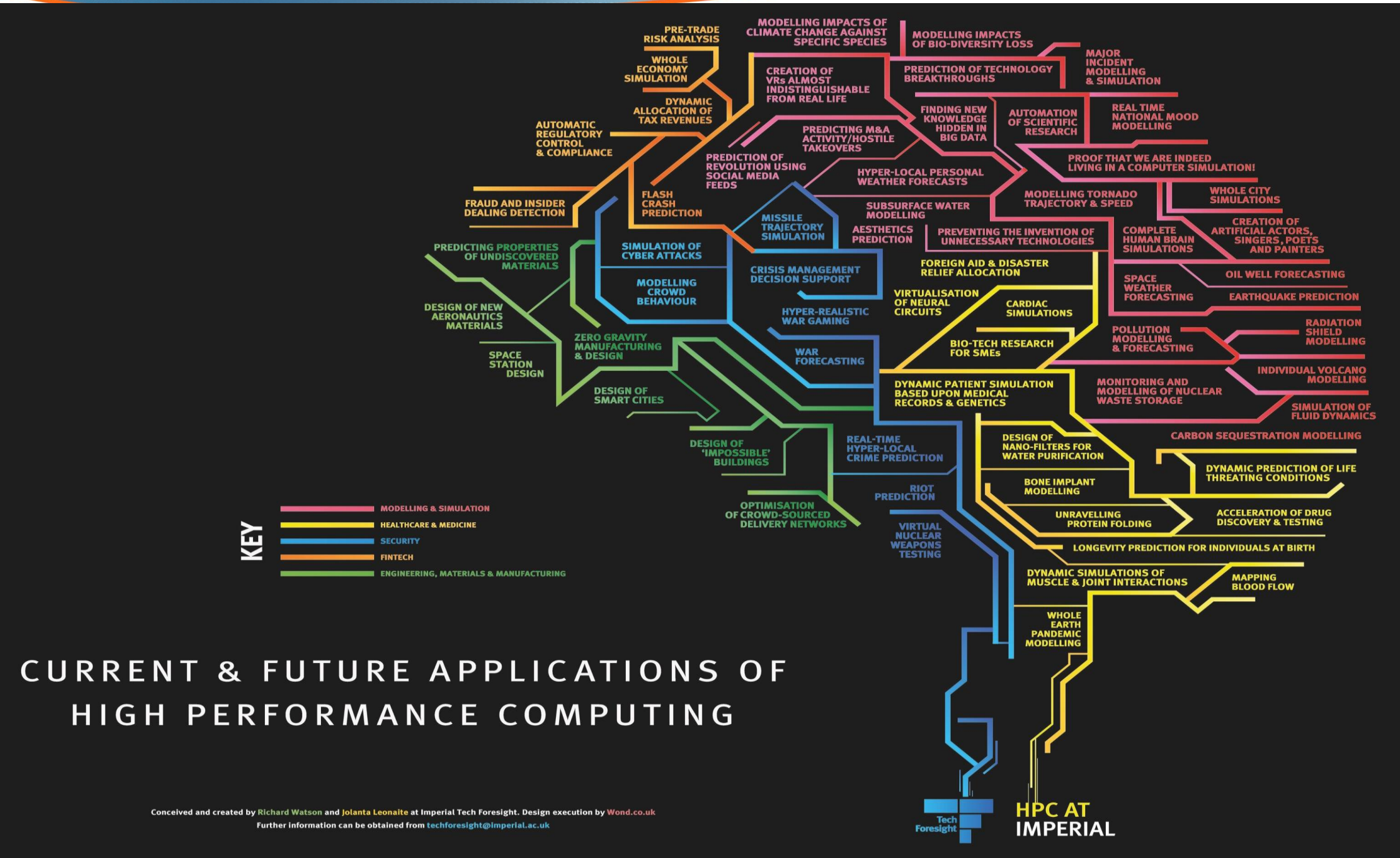
CURIE: Atos/Bull Bullx
GENCI/CEA
Bruyères-le-Châtel, France

Piz Daint: Cray XC 30
CSCS, Lugano, Switzerland

MARCONI: Lenovo
CINECA
Bologna, Italy

#3 at worldwide Top500 ranking

PRACE



CURRENT & FUTURE APPLICATIONS OF HIGH PERFORMANCE COMPUTING

Conceived and created by Richard Watson and Jolanta Leonaite at Imperial Tech Foresight. Design execution by Wond.co.uk

Further information can be obtained from techforesight@imperial.ac.uk

Tech Foresight

HPC AT IMPERIAL

Why Energy Efficiency matter ?

- HPC Energy wall (consumption must be kept bellow 20 MW)
  - Tianhe (former) #1 Top500: 34 PFlop/s @ 1900 Mflops/Watt
  - Extrapolate to 0.5 GW for 1 Exaflop/s : non sustainable !

1 GW = ½ EPR reactor or 125 windturbines and could cost 800 M€/year in 2020

- Energy efficient HPC technology ➔ Other Data Center (Big Data)
- World wide Data Centre electricity use forecast for 2020*:
  - 1031 billion kWh (equivalent to France + Germany + Canada + Brasil)
  - With a carbon footprint between 259 and 533 MtCO2e
- Market for energy efficient tech must be sustainable (beyond HPC)

* Source = SMART Global 2020 Make IT Green – Cloud Computing and its Contribution to Climate Change

Goals of the “Whole System Design for Energy Efficient HPC” PCP

- Fostering advances in energy efficiency (major TCO driver for HPC & Big Data)
- Energy wall for Exascale requires an O(100) increase of power efficiency
- Assessment of results through a pilot scalable to 100 PFlop/s

PRACE PCP Process and assessment methodology

- 3 phases competitive process:
  - Solution design (6 months, funding 10%)
  - Prototype development (10 months, funding 30%)
  - Pre-Commercial Pilot system (16 months, funding 60%)
- Assessment on “real” application benchmark from PRACE (suitable for PRACE 2)

Expected results and impact

- EU HPC supply industry (80% of R&D must be performed in EU) increase competitiveness
- EU HPC users (academia and industry) get early access to disruptive technology through PCP pilot
- EC and PRACE are learning by doing a new public procurement procedure, with high leverage effect
- PRACE procure Intellectual Property (IP) that paves the way toward sustainable Exascale:
  - IP is kept by HPC suppliers
  - Discount on future IP usage for PRACE members of the Group of Procurers of the PCP

Organization of the PRACE PCP

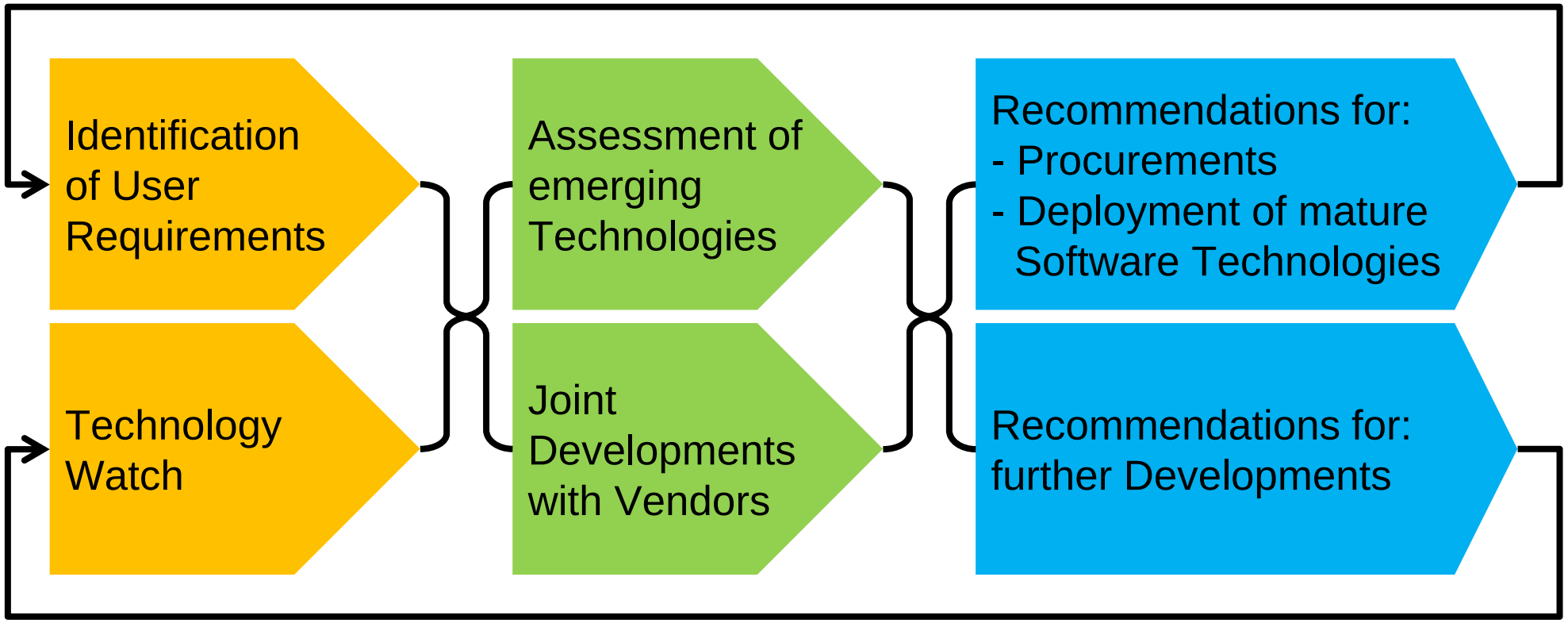
Procurers:

- Five PRACE-3IP partners (CINECA, CSC, EPCC, Juelich, GENCI) and PRACE aisbl as observer
- A GoP (Group of Procurers) was formed and contractually regulated by an agreement
- 9 M € Budget contributed by the procurers and EC (50/50)

Governance:

- Based on the GoP Committee as decision-making entity
- CINECA has been selected as the Procuring Entity
- Coordination between the Procuring Entity and the project assured

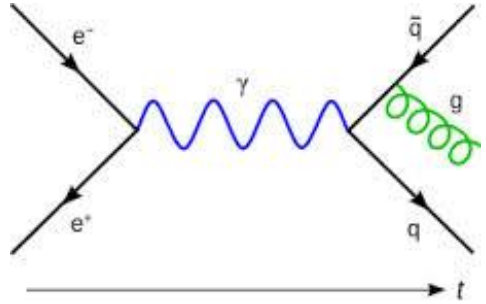
PRACE general prototyping expertise



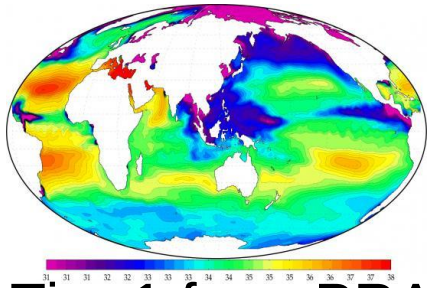
- Prototyping is a mandatory step in the selection and deployment of new technologies
- Prototyping is a vehicle for cooperation with technology providers

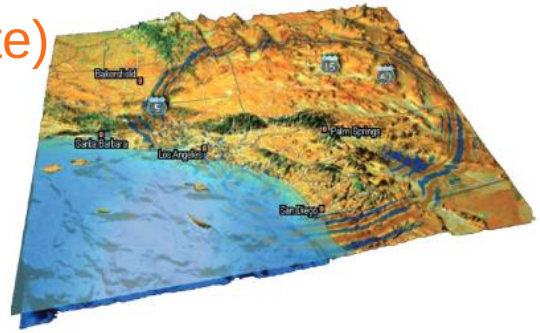
Implementation of the Benchmark Suite

- Criteria for Benchmark suite (BQCD, SPECFEM3D, QUANTUM E, NEMO) + LINPACK
  - Wide scientific domain coverage
  - Used in production in PRACE Tier0 systems
  - Energy-efficiency references available for contractors

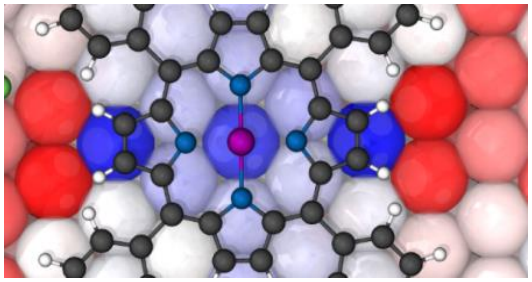


- Codes
  - From UEABS (Unified European Application Benchmark Suite)
  - PetaFlop scalability
  - Active development
  - Portability on hybrid machines





- Machines for validation: PRACE Tier-0 or Tier-1 from PRACE partners
  - CURIE and EURORA: x86_64 thin nodes (w & w/o GPU)
  - JUQUEEN: IBM BG/Q
  - SISU: Cray XC30



Summary of Main Technical Goals

Energy efficiency in whole system design

- Target energy efficiency in whole system design

Self-contained pilot system

- Allow for results to be tested in an HPC centre with real applications

Architecture suitable for PRACE

- Pilot system with 1 PFlop/s, scalable to 100 PFlop/s
- Capable of executing representative set of applications

Energy measurement capabilities

- Demonstration of improved energy-to-solution of real production applications

Energy efficient technology with sustainable market

- Aim for sustainable results

Overview Goals on Energy-to-Solution

	Reference	Atos/Bull	E4	Maxeler
BQCD	364.5	43.9	50.0	8.9
NEMO	169.6	46.8	74.0	30.4
Quantum Espresso	342.6	74.3	127.0	19.0
SpecFEM3D	1381.0	330.6	393.0	47.5

Units: kWh

Technical Results for Atos/Bull

Power-efficient compute devices: Intel Xeon Phi KNL

- High-level of parallelism, relatively low processor clock
- Energy-efficient high-bandwidth memory technology

Liquid-cooled PSU

- Enable fully liquid cooled design
- Allow for smaller PUE, enable full free cooling

High Definition Energy Efficiency Visualization

- Enable energy-consumption profiling of applications

Dynamic power optimisation

- Enable reduction of the energy consumption of application during execution

Technical Results for E4

Power-efficient compute devices: IBM OpenPower + NVIDIA Tesla GPUs

- Extremely parallel compute devices running at relatively low clock speed
- >97% of throughput of floating-point operations stemming from GPUs

Liquid cooled design based on Open Compute form factor

- Integration concept based on open standards
- Flexible cooling solution focussing on main components

Power/energy monitoring and power aware job scheduling

- Monitoring integrated into new power distribution design
- Machine learning approaches to power aware job scheduling

Technical Results for Maxeler

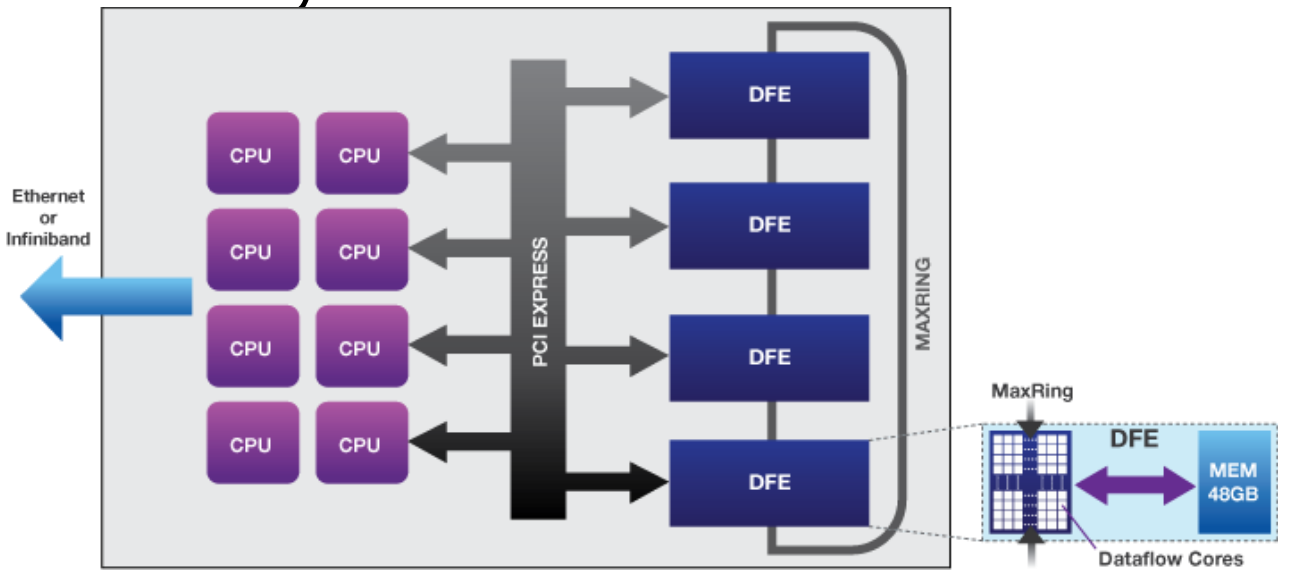
MAXELER Technologies
MAXIMUM PERFORMANCE COMPUTING

Approach based on newest generation of Data Flow Engine architectures

- MAX5 DFE based on Xilinx Ultrascale FPGAs → very low clock rate
- Project impacted MAX5 design

Main R&D focus: Demonstrate usability of architecture for HPC applications

- Application kernels implemented in MaxJ



Summary and Conclusions

First pan-European PCP on HPC, and it successfully made it to Phase 3

- Evaluation criteria focussed on PRACE real-world benchmark
- Energy-efficiency measurement methodology useful for other procurement (TCO evaluation)

Broad range of solutions offered

- Standard vs. exotic architectures: Xeon Phi KNL/GPU vs. data-flow engines in FPGA
- Energy measurement and tuning capabilities

Pilot systems are in the process of being deployed

- Facilitate evaluation in production environment and expose our community to new technologies
- But long process in comparison with a standard procurement

Solutions close to productisation

- Approaches well aligned with contractors' roadmaps while sharing R&D risks (and reward)

Contractors well positioned for follow-up procurements ➔ **warning PPI ahead !**

- Solutions pre-tested at leading supercomputing centres
- Contractors exposed to applications that continue to be relevant

High leverage opportunity for HPC & Big Data sector in Europe:

➔ If PCP budget match the cost to tackle selected technological locks

➔ Must produce useful "Pre-Commercial" innovation (no blue sky R&D)

THANK YOU FOR YOUR ATTENTION

www.prace-ri.eu