FP7 Support Action - European Exascale Software Initiative

DG Information Society and the unit e-Infrastructures

Addressing the Challenge of Exascale

European Exascale Software Initiative EESI

Towards Exascale roadmap implementation

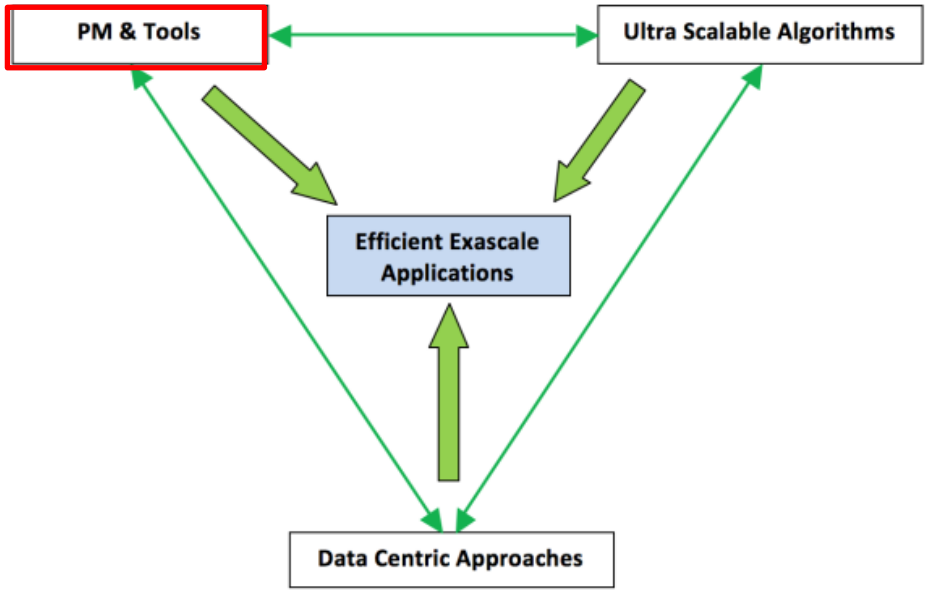
EESI2 – Recommendations

Holistic Approach for Extreme Heterogeneity Management of Exascale Supercomputers

Giovanni Erbacci, CINECA

Holistic Approach for Extreme Heterogeneity Management of Exascale Supercomputers

- Exascale requires heterogeneity at an unprecedented level
- Power is the main architectural constraint
- The energy efficiency and the ability to handle large amount of data are primary concerns to reach exascale
- Huge impact on applications

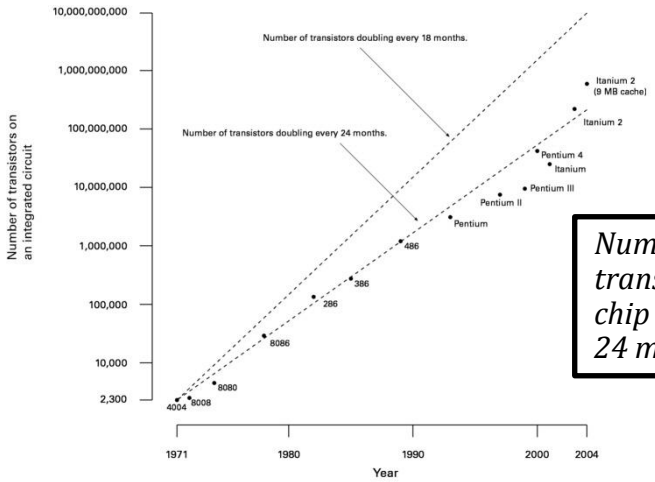


Massively complex heterogeneous supercomputers must provide HW level energy-efficiency and performance.

It is fundamental to support programmability, efficiency and productivity of tools and applications, energy aware, at Exascale

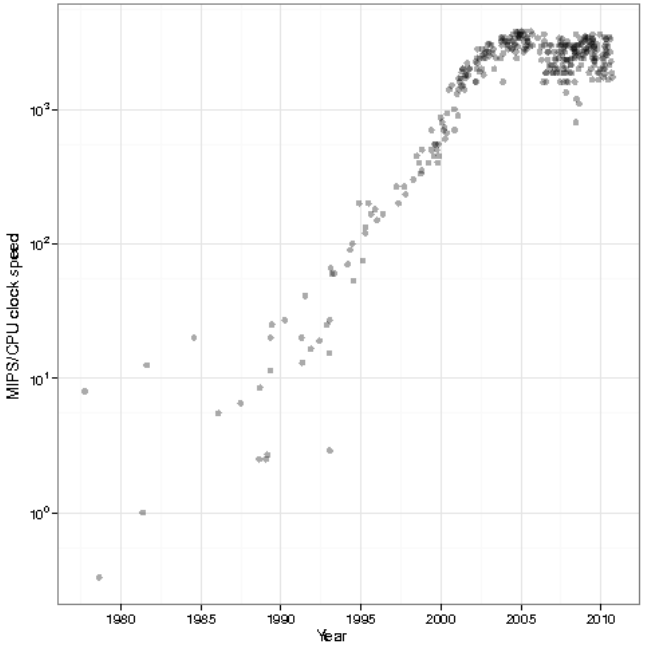
HPC Trends

Moore's Law



Number of transistors per chip double every 24 month

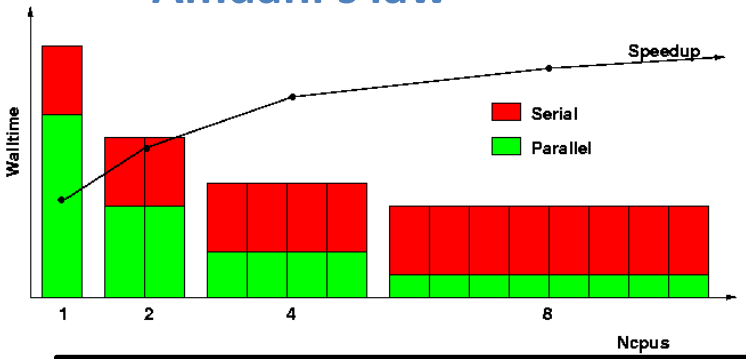
Dennard scaling law (downscaling)



The core frequency and performance do not grow following the Moore's law any longer

Increase the number of cores to maintain the architectures evolution on the Moore's law

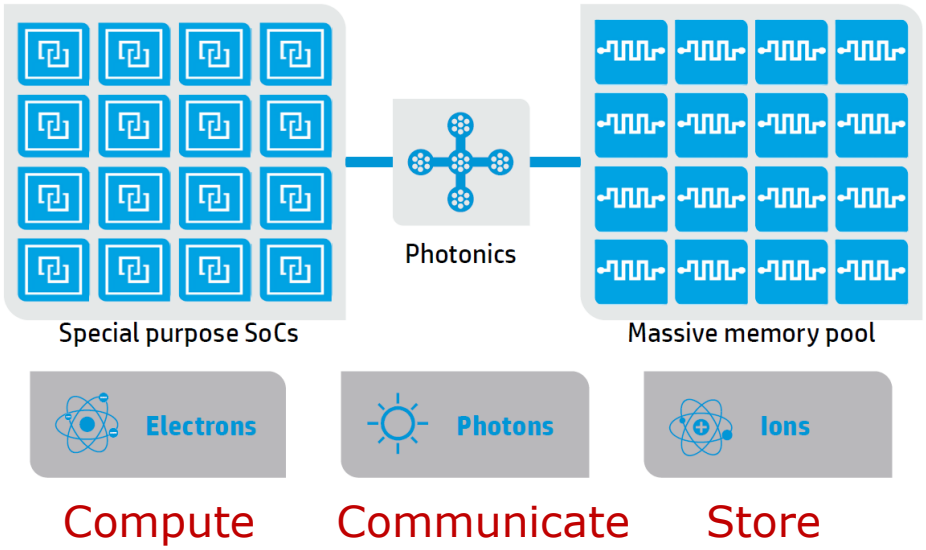
Amdahl's law



maximum speedup tends to $1 / (1 - P)$
P= parallel fraction

The upper limit for the scalability of parallel applications is determined by the fraction of the overall execution time spent in non-parallel operations.

HPC Trends toward Exascale



- SoC System on a Chip based servers
  - Less general-purpose, more workload focused
  - Dramatic reduction in power, cost, and space
- Non-volatile memories:
  - Breakthrough in storage and memory technology
- Integrated photonics
  - Dramatically improve memory bandwidth and many-core performance
  - Lower energy, Reduce power consumption

new

PM & Tools

Motivation:

Integrated heterogeneous supercomputers with high level energy-efficiency and performance are the way to Exascale

… but communication overheads as well as the increased complexity induced by the variety of design choices mine heterogeneity potential gain

- The entire **SW stack** (programming models, run-time, OS and system support software) **needs to be fully innovated** to support programmability and efficient performance/energy usage of the different resources and the computational models
- **HW support** for efficient communication, coherency and offload control **needs to be co-designed with the SW stack** to exploit the energy-efficiency and performance
- **Novel abstraction layers and optimization strategies** must be developed to cope with the extreme heterogeneity and parallelism of Exascale systems
- **Synergic integration with novel programming models strategies** and software engineering methods
- **Development of novel management and control APIs** which can be exploited by the programmer and by the runtime to deploy at Exascale the potential energy-efficiency of novel architectures in the different application domains

Holistic Approach for Extreme Heterogeneity Management of Exascale Supercomputers

PM & Tools

Design and develop new efficient HW/SW APIs for the integrated management of heterogeneous systems, near-data technologies and energy-aware devices, to enable exascale-ready applications.

Proposal: Fund R&D programs to foster the research and development of:

- HW/SW APIs to manage the complexity and the programmability gap inherent the extreme heterogeneous level of Exascale supercomputers
- Design strategies for scalable and efficient heterogeneous-aware exascale applications
- Scalable and efficient community scientific applications for exascale
- System software to support efficient usage of exascale heterogeneous supercomputers in production

Funding Description

Impact of the recommendation on the efficiency of future Exascale applications

- It is crucial to produce efficient **HPC and Data intensive applications** exploiting the Exascale architectures
- Research projects in this area are expected to improve the performance and **energy-efficiency scalability of the entire ecosystem** of supercomputers applications which includes different aspect of the **daily society** as well as **industrial competitiveness** and **social security**

Emergency for starting this work

- Energy efficiency is fundamental at Exascale level
- It is urgent to be ready in different application domains to **deploy the energy-efficiency** potentiality of novel architectures at Exascale
- Establish cooperation with the **activity of CoEs**

Impact on improving EU strength

- Handling **extreme heterogeneous exascale supercomputing** means being able of sustain **exascale ICT infrastructure**.
- This is in line with the **ETP4HPC strategic research agenda**
- Synergic integration with the other research pillars identified by the **EESI recommendations**

Level of R&D efforts to tackle this issue

10 M Euros over 3-4 years