NESUS COST IC1305
Network for Sustainable Ultrascale Computing

COST Action IC1305:
Network for Sustainable Ultrascale Computing Systems (NESUS)

2015
Krakow, September 11th 2015

Prof. Jesus Carretero
Nesus Action Chair
University Carlos III of Madrid
Spain

ARCOS

University Carlos III of Madrid

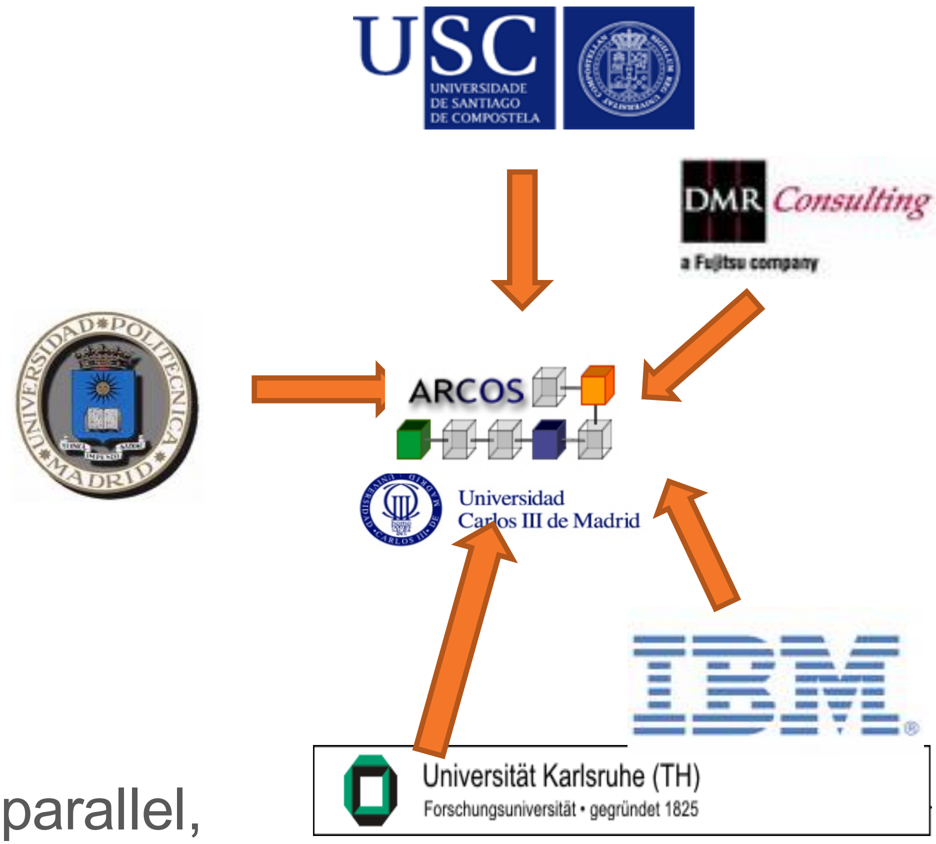
- ❑ Created in 1989.
- ❑ Centers:
  - Social Sciencies and Law School
  - Humanities and Journalism School
  - Engineering School.
    - Computer Science & Engineering Department
      - Research group: Computer Architecture and Systems (ARCOS)

Leganés

Madrid, Spain

ARCOS Research Group

- Created in 1999.
  - Leader: Jesus Carretero
- Staff:
  - 2 Professors
  - 4 associate professors
  - 3 assistant professors
  - 5 researchers
  - 12 PhD Students.
- Goals:
  - Applied research on large-scale parallel, distributed systems (refactoring, runtimes and I/O).

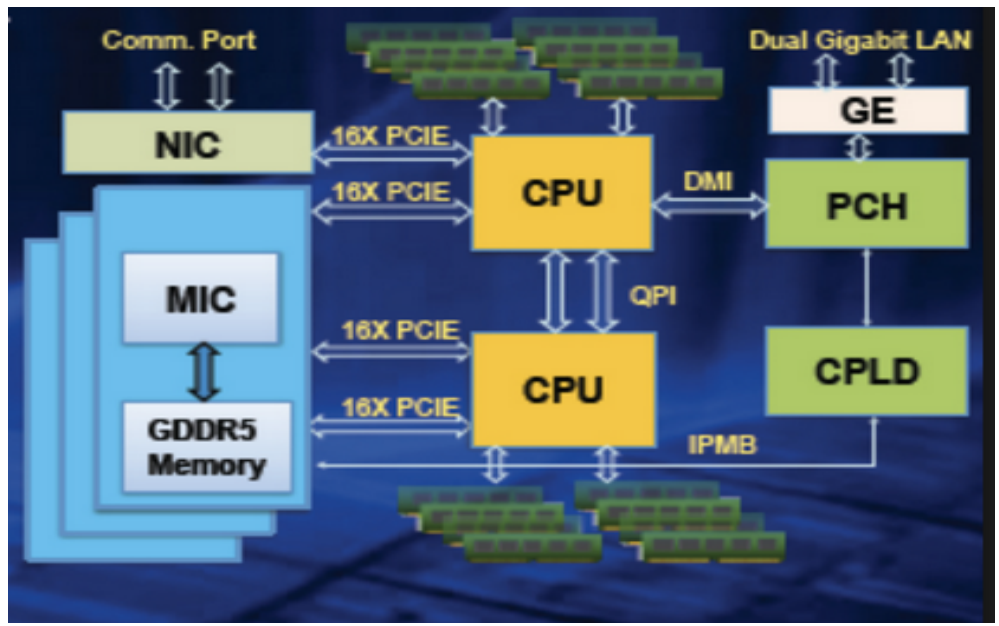


Action Background

- There is a major research effort around the world towards:
  - Exascale (PRACE, EESI, HP-SEE, IESP, JESI)
  - Large scale virtual systems (XSEDE, FutureGrid, Grid5000).
  - Big data solutions (BIG, EIOW, BDEC)
- Efforts are mostly separated
- However, computing convergence and platform convergence are foreseen:
  - Capability (HPC) and capacity (HTC)
  - Clusters and clouds

Top 500: Tianhe 2

- World faster supercomputer, developed in China.
- 16000 computing nodes with 2 Intel Ivy Bridge Xeon
  - Each with 3 Intel Xeon Phi
- 55 PFlops, 18 MWats



Green 500: L-CSC

- World most power efficient supercomputer, developed in Germany.
- 160 servers: 2 Ivy-Bridge + 4 AMD FirePro GPU
  - 1600 cores per node
  - 5.2 GFlops/W, 57 Kwats
  - Energy-efficient software design

- 20 top systems in Green 500 are based on accelerators

Extreme-scale datacenters

- ❑ Huge clusters with X00K computers
- ❑ Distributed storage + back end
- ❑ E.g. Council Bluffs Google datacenter

Expected HPC systems characteristics ranges in 2020

	Petascale system (2012)	Exascale / data center (2020)	Petascale / departmental (2020)	Terascale / embedded (2020)
Number of nodes	[3-8] x10^3	[50-200] x10^3 (20x)	[50-100]	1
Computation (Flop/s & Instructions)	10^15	10^18 (1000x)	10^15	10^12
Memory Capacity (B)	[1-2] x10^14	> 10 ^17 (1000x)	> 10^14	> 10^11
Global Memory bandwidth (B/s)	[2-5] x10^14	> 10^17 (1000x)	> 10^14	> 10^11
Interconnect bisection bandwidth (B/s)	[5-10] x10^13	~10^16 (1000x)	~10^13	N/A
Storage Capacity (B)	[1-10] x10^15	>10^18 (1000x)	> 10^15	> 10^12
Storage bandwidth (B/s)	[10-500] x10^9	> 10 x10^12 (1000x)	> 10 x10^9	> 10^6
IO operations/s	100 x10^3	> 100 x10^6 (1000x)	> 100 x10^3	> 100
Power Consumption (W)	[.5-1.] x10^6	< 20 x10^6 (20x)	< 20 x10^3	< 20

- From: ETP4HPC Strategic Research Vision

More complex computing scenarios

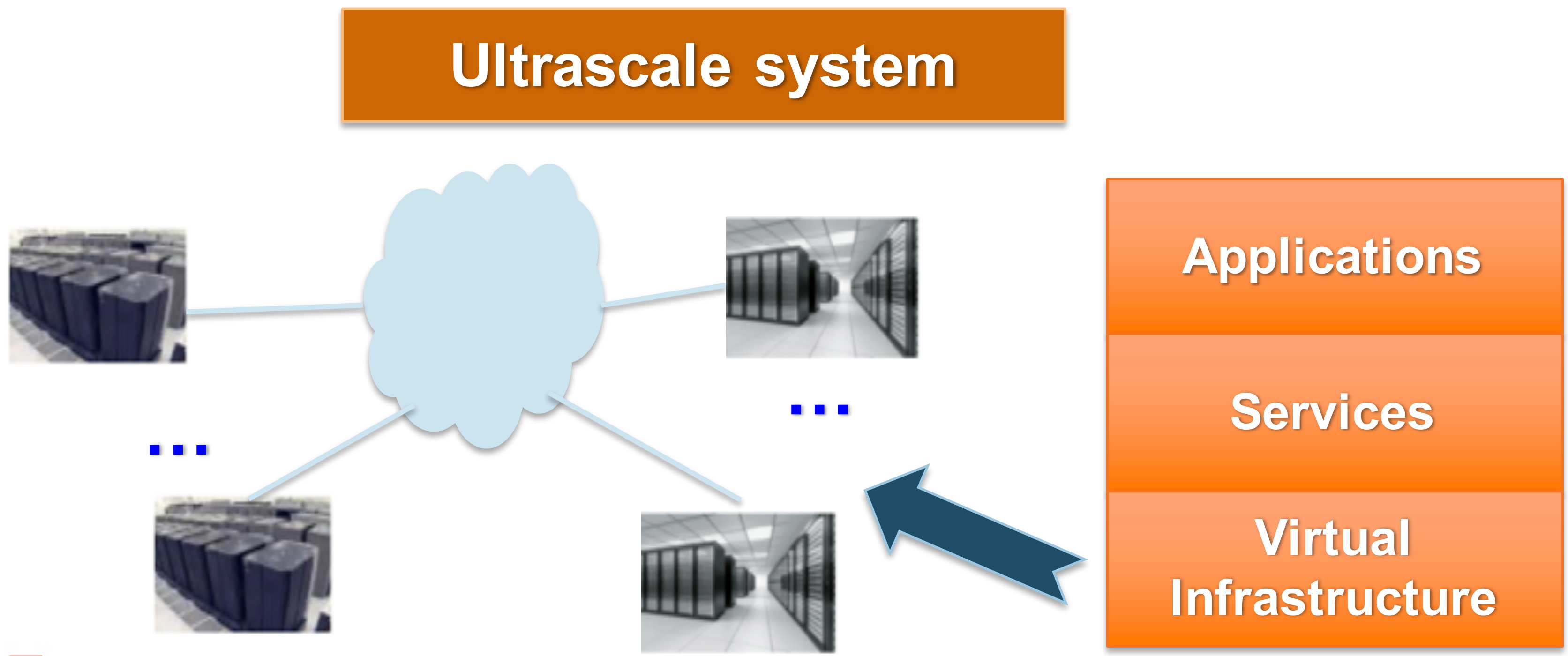
- High-performance computing (HPC)
  - heavily focused on compute-intensive applications;
- High-throughput computing (HTC)
  - focuses on using many computing resources over long periods of time to accomplish its computational tasks;
- Many-task computing (MTC)
  - aims to bridge the gap between HPC and HTC by focusing on using many resources over short periods of time;
- Data-intensive computing (DIC)
  - heavily focused on data distribution, data-parallel execution, and harnessing data locality by scheduling of computations close to the data.

Crossing domains

- Many applications have mixed features:
  - HPC + Data intensive: simulations
  - Many tasks + HPC: multi-scenario simulations
  - Many-tasks + data intensive: workflows
  - HTC + data intensive: data analysis
- How are we going to provide flexible and powerful architectures:
  - ¿Extreme-scale parallel systems?
  - ¿Through software-defined systems?
  - ¿Relying on virtualization: cloud computing?

Ultrascale systems

- Ultrascale computing systems (UCS)
  - Big-scale complex system integrating parallel and distributed computing systems, that cooperate to provide solutions to the users at unprecedented scale.



The problem …

- As the scale and complexity increase in UCS, **sustainability will become a major challenge**

- Sustainability not only means energy, but all factors that wiil allow the system to be adopted and maintained.

- Sustainability in UCS should be the result of leveraging several cross-layer aspects to face complexity:
  - Programmability, Data management, Resilience, Energy efficiency, Scalability, …

The idea

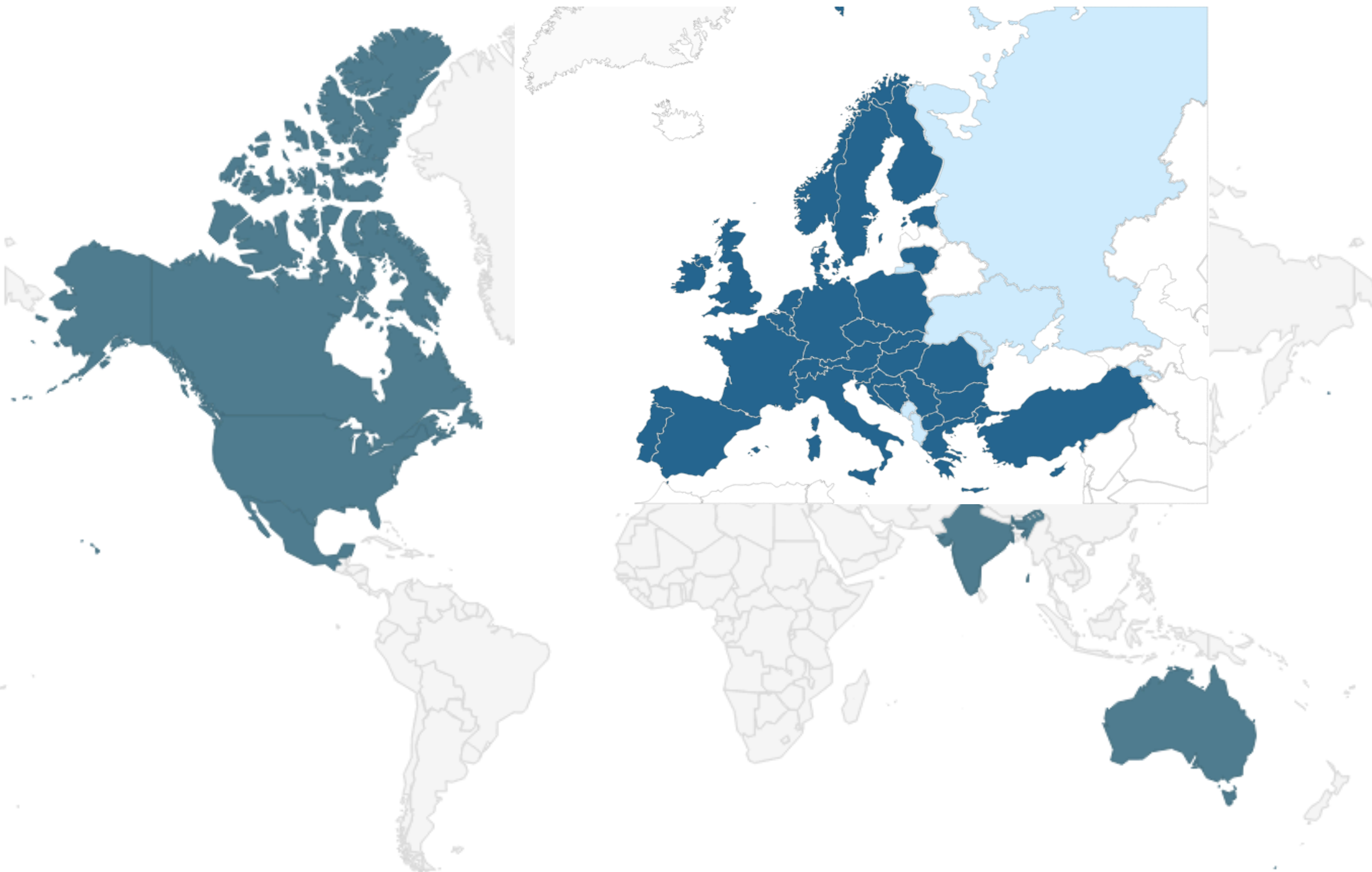
- Sustainability is a holistic goal, not only energy
  - Need of metrics to express sustainability
- We need important breakthroughs for UCS
  - Or they will not be sustainable
  - System software is crucial to have sustainable future systems -> ¿Software-defined systems?
- Mechanisms should be valid for different computing models and architectures
  - HPC, HTC, DIC, workflows …
  - Clusters, clouds, grids, …
- NESUS action will be instrumental to research sustainability in UCS with many other institutions
  - Definition of services for system layers

Aim of the Action

- To coordinate efforts for proposing realistic solutions addressing major challenges of **building sustainable Ultrascale Computing Systems (UCS)** with a collaborative approach.

- Reasons:
  - **Concentrate research** on sustainability in ultrascale systems looking for integrated solutions to master the complexity.
  - **Federate a dispersed community** to eliminate overlaps by improving collaborations.
  - Increasing **awareness** on sustainability of ultrascale systems.
  - Impact **both** the IT and societal sides by:
    - Dissemination of best practices and experiences to achieve sustainable systems.
    - Providing sustainable ultrascale applications and benchmarks.

Consortium



- 45 countries
- >220 members
- 35% young researchers

Open to new members

Open to cooperation

Scientific goals

- Exploring new solutions for the **system software stack** (programming paradigms, runtimes, middlewares, resilience, data management, and energy models) and their application to enhance sustainability in UCS.
    - Understanding trade-offs and synergies to leverage all factors.
    - Considering new hardware and architectural solutions.

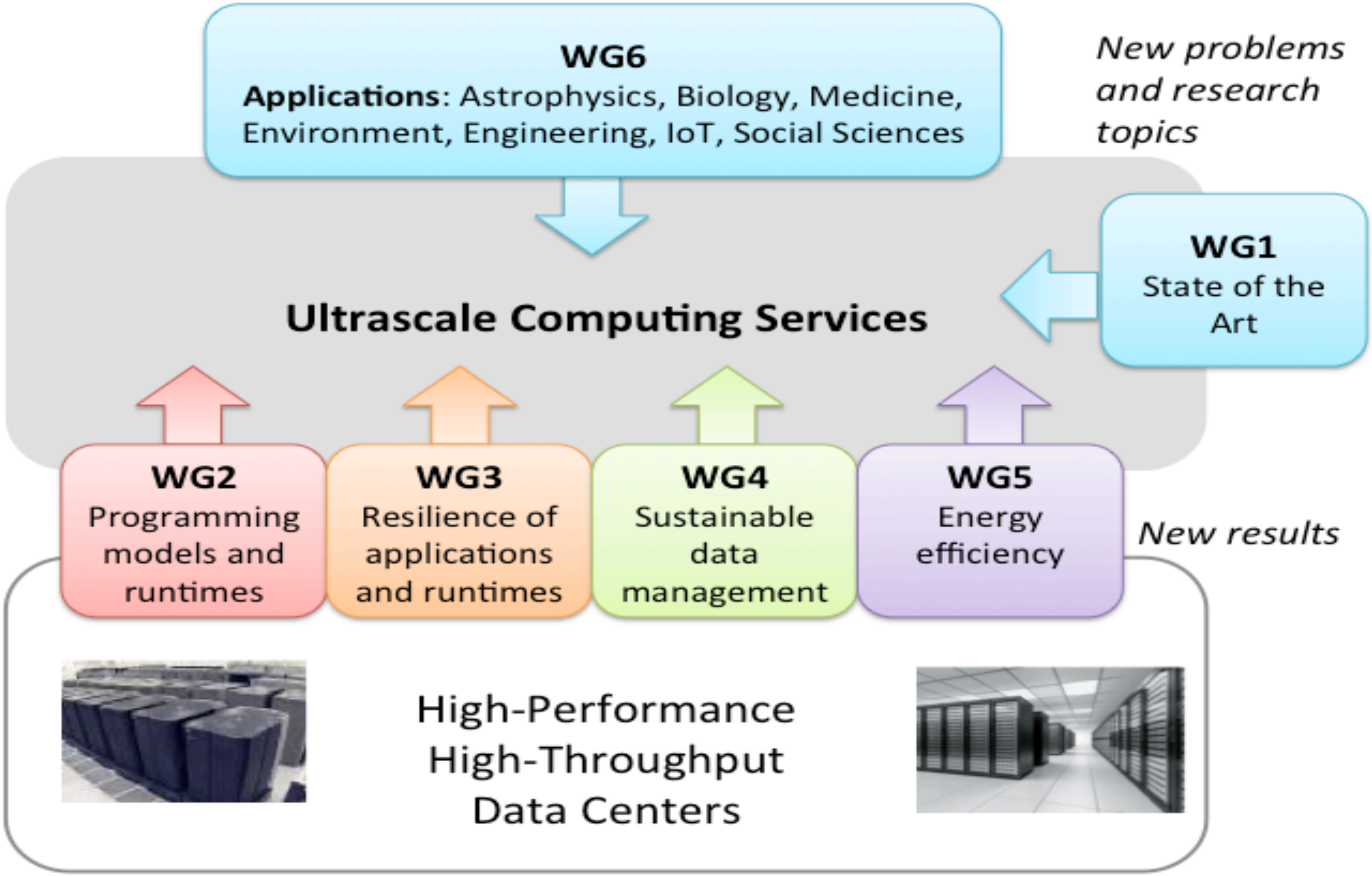
- Exploring **redesign and reprogramming** efforts for applications to efficiently exploit ultrascale platforms, while providing sustainability.

- Holistic approach to **manage the whole ecosystem**,
    - Important to understand how all the factors affect UCS sustainability -> sustainability metrics

Challenges in UCS

Programming environment	New programming models: Hierarchical models Many-task models Global memory models Data distribution and locality Awareness of data-movement cost	Emergence of new algorithms Energy-efficient algorithms Fault-tolerant algorithms Application code migration and re-writing High-productivity methods
System Software	Increased system heterogeneity Capability for virtualization Standardized APIs Scalability, modularity, robustness	Explosion of data volumes Awareness of data-movement cost Extensive system monitoring Performance models
System architecture	Heterogeneous platform architectures Power consumption I/O latency and bandwidth Concurrency and data locality Storage capacity	Extreme scale from sub-component to total system Resiliency , Reliability, Availability, Serviceability (RAS) Software-defined systems

Workplan


WG6
Applications: Astrophysics, Biology, Medicine, Environment, Engineering, IoT, Social Sciences
New problems and research topics
Ultrascale Computing Services
WG1
State of the Art
WG2
Programming models and runtimes
WG3
Resilience of applications and runtimes
WG4
Sustainable data management
WG5
Energy efficiency
New results
High-Performance
High-Throughput
Data Centers

WG 2. Programming models and runtimes

- Focus
  - Promoting new sustainable programming and execution models in the context of rapidly changing underlying computing architecture.
- Key objectives
  - Improving programmability.
  - Scale handling (optimal usage of resources, faults)
  - Dynamic adaptation to underlying computing architecture
  - Adaptations for data-centric programming models, resilience, and energy-efficiency

WG 3. Resilience of applications and runtime environments

- Focus
  - Innovative techniques to deal with hardware and system software failures or intentional changes within the complex system environment
- Key objectives
  - Monitoring and assessment of failures in Ultra-large-scale systems
  - Going beyond fail-stop errors to manage hard, transient, and failures in the SW stack
  - Understanding HW & SW dependencies and monitoring changes and their impact within complex systems.

WG 4. Sustainable data management

- Focus
  - Study data management lifecycle on scalable architectures in a synergistic approach to be able of managing huge data sets from application and real-world devices.

- Key objectives
  - Evolution of the storage I/O stack towards higher-levels of scalability and sustainability to cope with globalization of data
  - Improving the programmability of data management and analysis and enhancing data workload predictability.

WG 5. Energy efficiency

- Focus
  - Energy efficiency of ultrascale systems in front of other quality metrics

- Key objectives
  - Exploring energy sustainability in UCS and proposing new holistic models of energy consumption for UCS
  - Designing and studying energy aware components and applications
  - To explore the design of metrics, analysis, frameworks and tools for putting energy awareness and energy efficiency at the next stage.

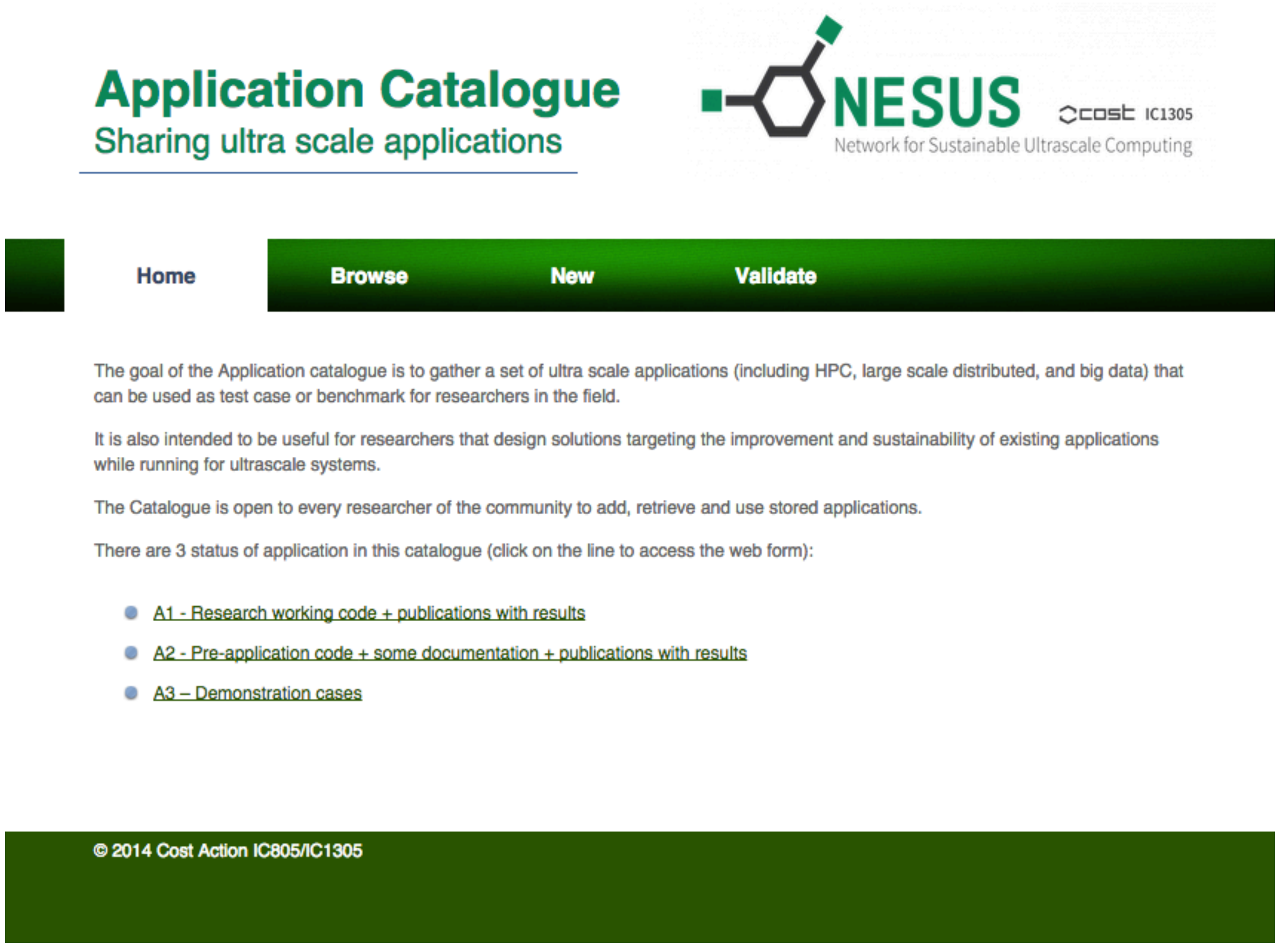
WG 6. Applications

- Focus
  - Identify algorithms, applications, and services amenable to ultrascale systems and to study the impact of application requirements on the sustainable ultrascale system design

- Key objectives
  - Categorization and selection a set of key applications with need for ultrascale computing
    - Evaluation of the needs of the selected applications concerning scalability, programmability, portability, resilience
  - Identification of computational patterns for expressing a higher level of abstraction at UCS
  - Estimation of the redesign and reprogramming effort for legacy applications

Application and tool catalogues

- http://www.nesus.eu/catalogue/



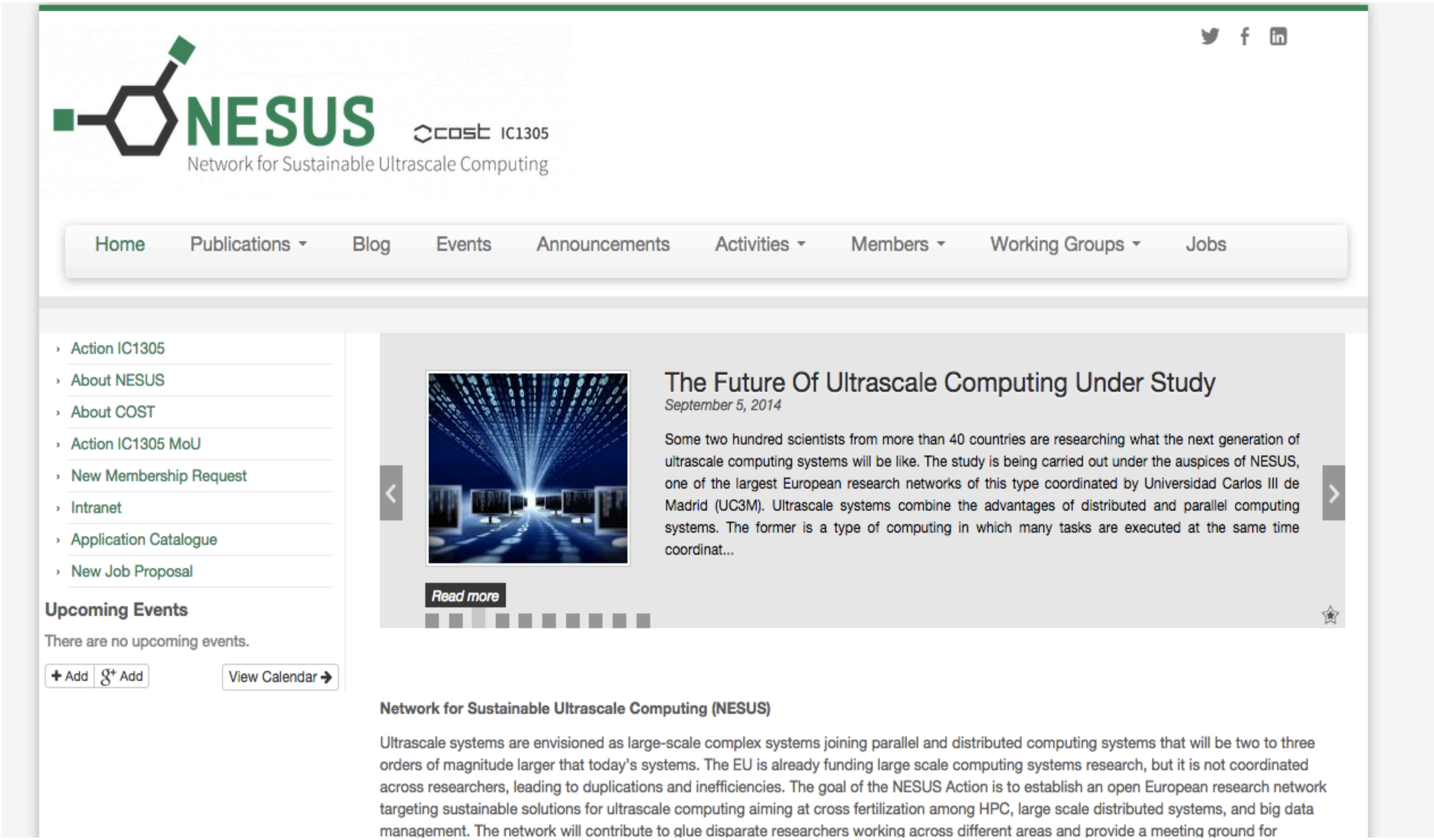
NESUS Activities

- Research coordination
- Working Group meetings
- Research stays
- Yearly workshop
- Training school
- PhD symposium

- Strong emphasis in cooperation
  - Join publications, tools, applications, …
  - With industry to solve real-world cases

NESUS
Network for Sustainable Ultrascale Computing

NESUS Web portal (nesus.eu)



ARCOS

NESUS COST IC1305
Network for Sustainable Ultrascale Computing

Thank you!

Prof. Jesus Carretero

www.nesus.eu

ARCOS