ComputationWorld 2010, Lisbon, November 21 - 25, 2010

Distributed
European
Infrastructure for
Supercomputing
Applications

Advances in Computational Sciences: From HPC to Grids to Clouds

Wolfgang Gentzsch
EU DEISA Project & Board of Directors of OGF
gentzsch at rzg.mpg.de

IARIA

Content

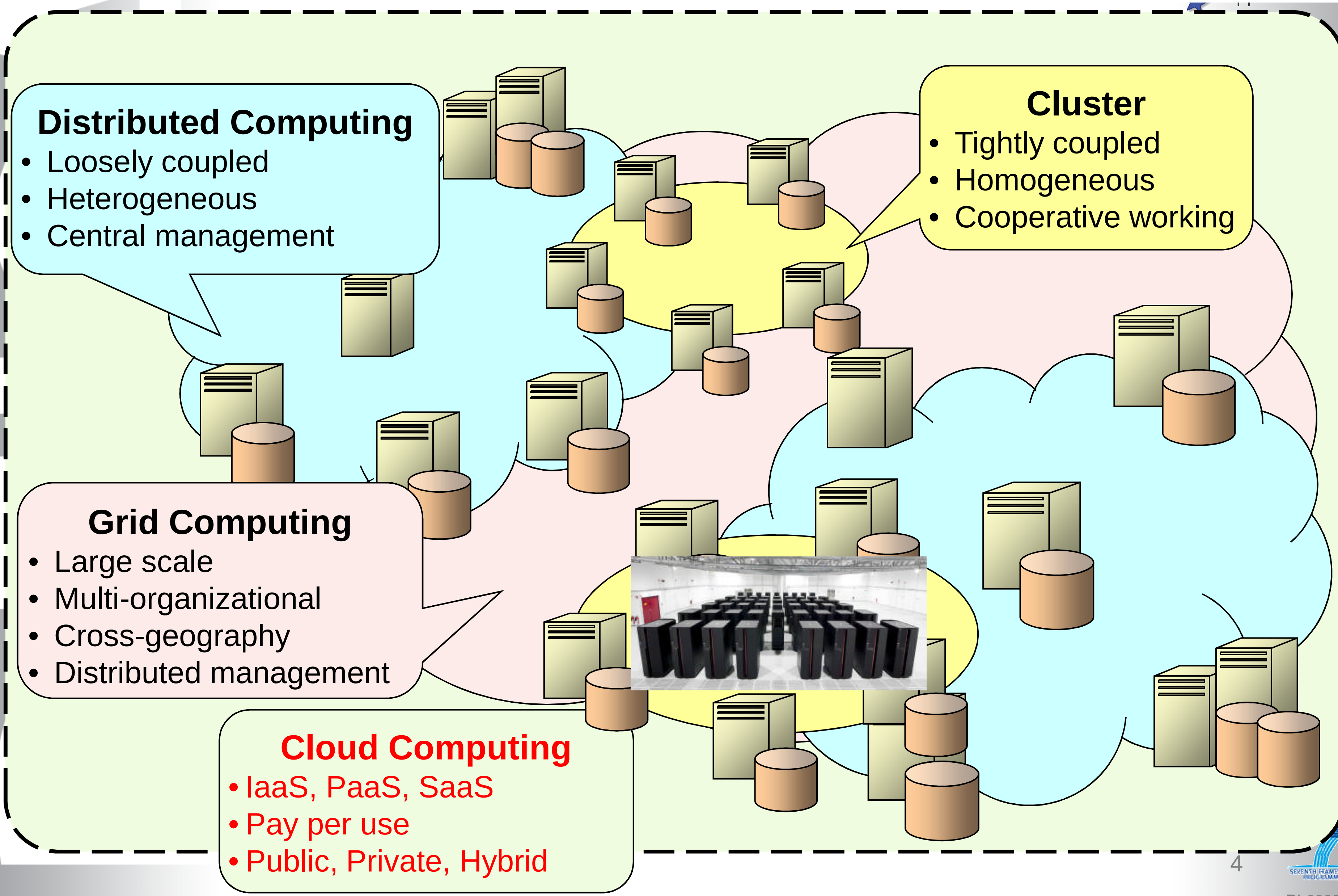
- Performance requirements of scientific applications
- Components: HPC Centers, Grids, and Clouds
- Example: The DEISA Ecosystem for HPC Applications
- Cloud Computing
- HPC in the Cloud
- Applications in the Cloud
- Challenges in the Cloud
- Conclusions

Science Apps: Performance Requirements

Scientific Field (numbers in Teraflop/s)	2005-2007	2007-2009	2010
Climate and Earth System Research	20	50-100	>500
Geophysics	1	10-100	>1000
Nanostructure Physics	1	10-50	>200
Solid-State Physics	1	50-100	>1000
Computational Fluid Dynamics	2.5	25-100	>1000
Astrophysics	10	50-100	>500
Elementary Particle Physics and Physics of Hadrons and Nuclei	30	100	>1000
Materials Science	10	50-100	>500
Theoretical Chemistry	3	25-125	>300
Soft Matter	3	30	>200
Biophysics and Bioinformatics	3	15-80	>1000
Plasma Physics	10	50	>500

A. Bode, W. Hillebrandt, and Th. Lippert: Scientific Case for the German Government, 8/2005

Terminology



HPC Centers and Clusters - still our bread & butter -

Distributed European Infrastructure for Supercomputing Applications

- HPC Centers are **service providers**, for past 35 years
- IT Service: Computing, storage, applications, data, etc
- Serve (local) research, education, and industry
- Very professional: to end-users, they look (almost) like Cloud services, if compared with Amazon Cloud definition: easy, secure, flexible, on demand, pay per use, self serve)
- Challenges: peta/exa, software, scalability, multicore, GPUs, Green Computing, connected to Grids & Clouds,...

RI-222919

Grids

1998: The Grid: Blueprint for a New Computing Infrastructure:

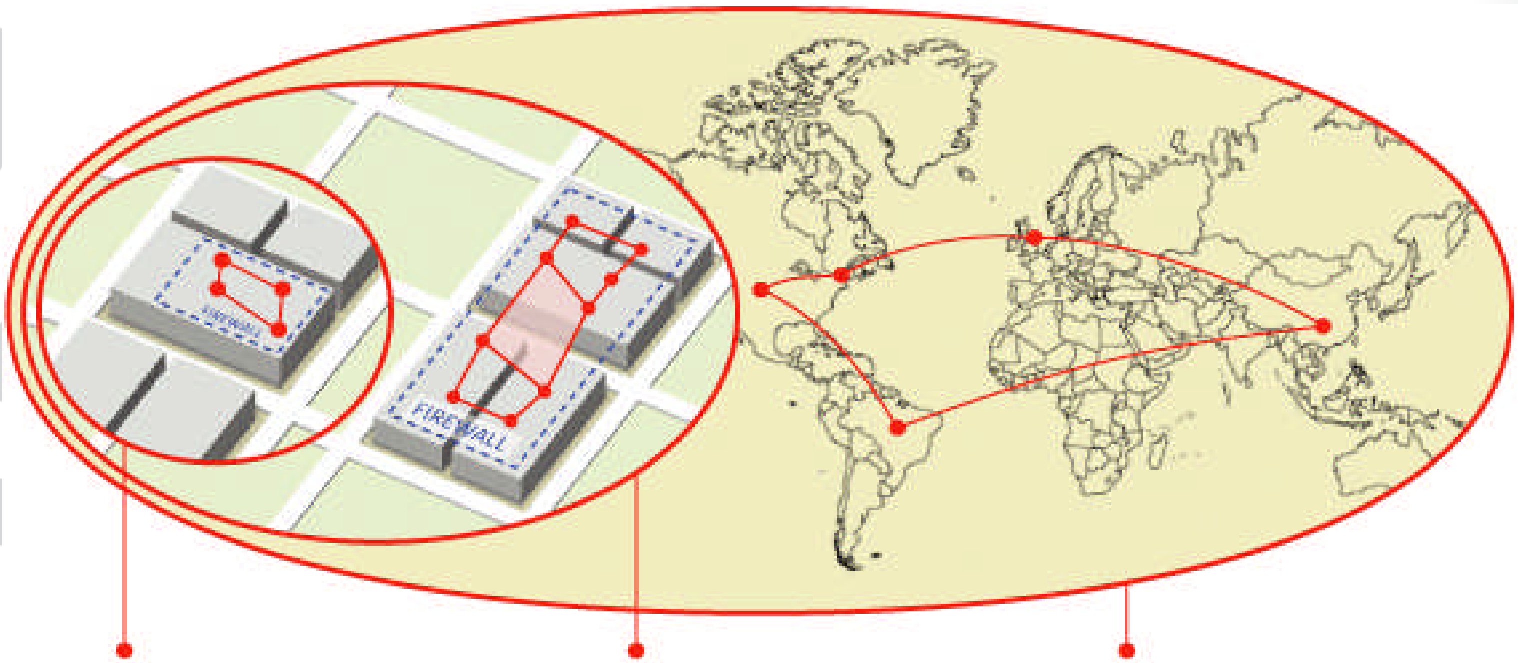
"... hardware and software infrastructure ... **dependable, consistent, pervasive, inexpensive access** to high-end computational capabilities."

2002: The Anatomy of the Grid:

"... coordinated resource **sharing** and problem solving in dynamic, multi-institutional **virtual organizations**."

Quotes: Ian Foster, Carl Kesselman, Steve Tuecke

Grids

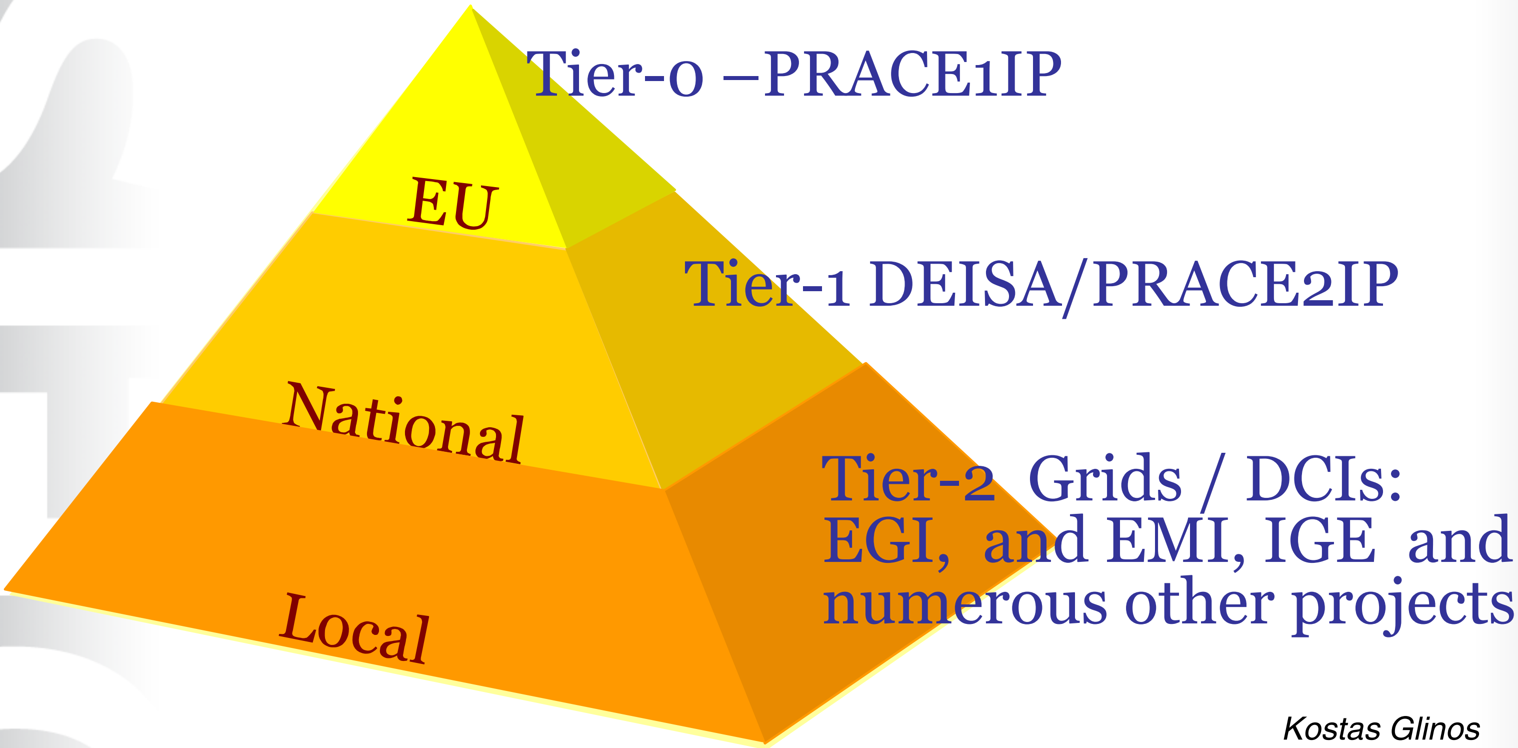


Departmental Grids

Enterprise Grids

Global Grids

European HPC Eco-System



Kostas Glinos

European Commission, 2010

IARIA

Distributed
European
Infrastructure for
Supercomputing
Applications

Example :

The DEISA Ecosystem for HPC Grand-Challenge Applications

HPC Centers in the Grid

DEISA:
Distributed European Infrastructure for Supercomputing Applications

DEISA: Vision and Mission

Distributed
European
Infrastructure for
Supercomputing
Applications

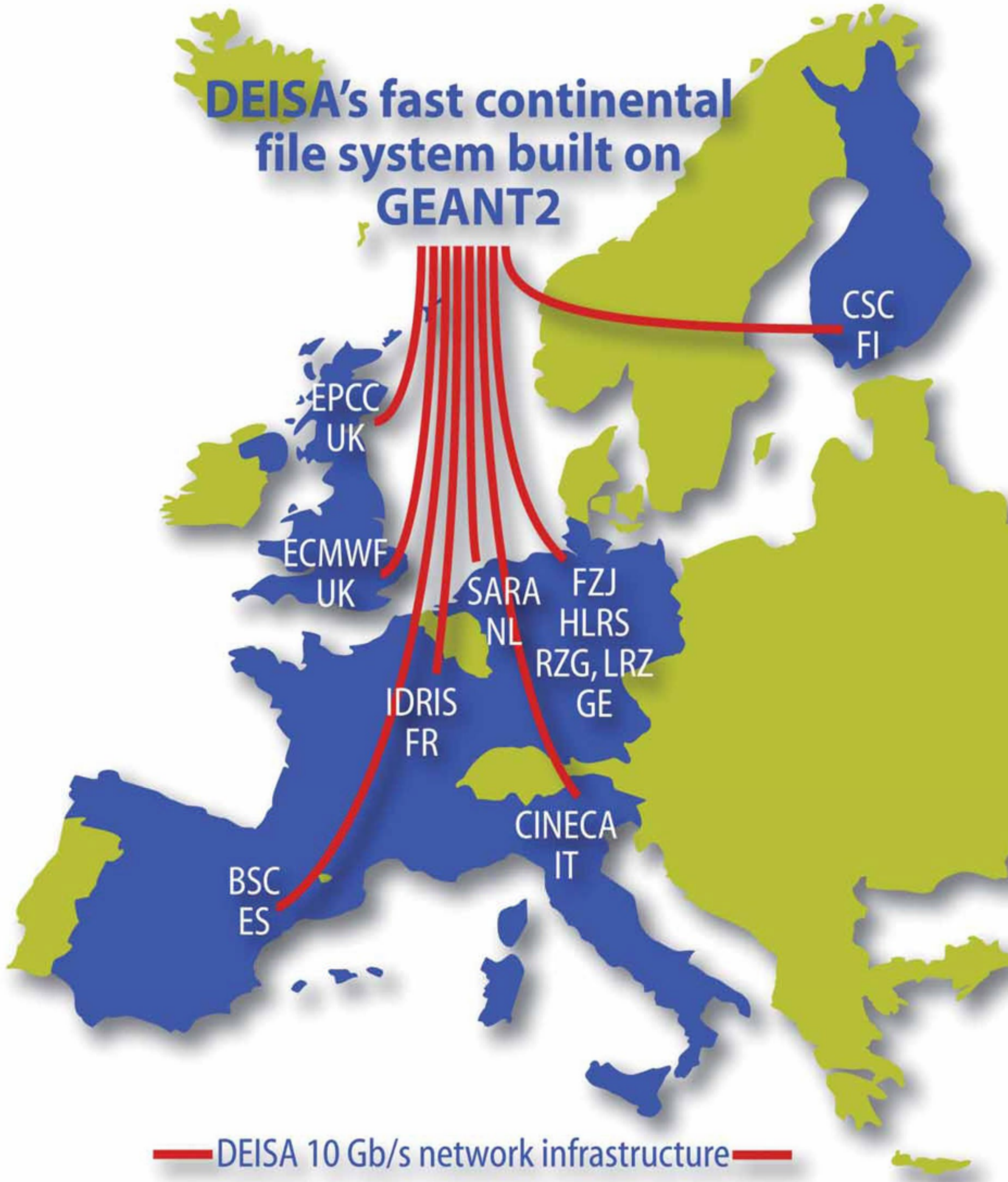
Vision:

Persistent European **HPC ecosystem** integrating Tier-1 (Tflop/s) centres and European Tier-0 (Pflop/s) centres.

Mission:

Enhance Europe's capability in computing and science by **integrating most powerful supercomputers** into a European HPC e-infrastructure.

Built European Supercomputing Service **on top of existing national services**, based on the deployment and operation of a persistent, production quality, distributed supercomputing environment with continental scope.



DEISA

Distributed European Infrastructure for Supercomputing Applications

Six years of operation

Most powerful European Supercomputers for most challenging projects

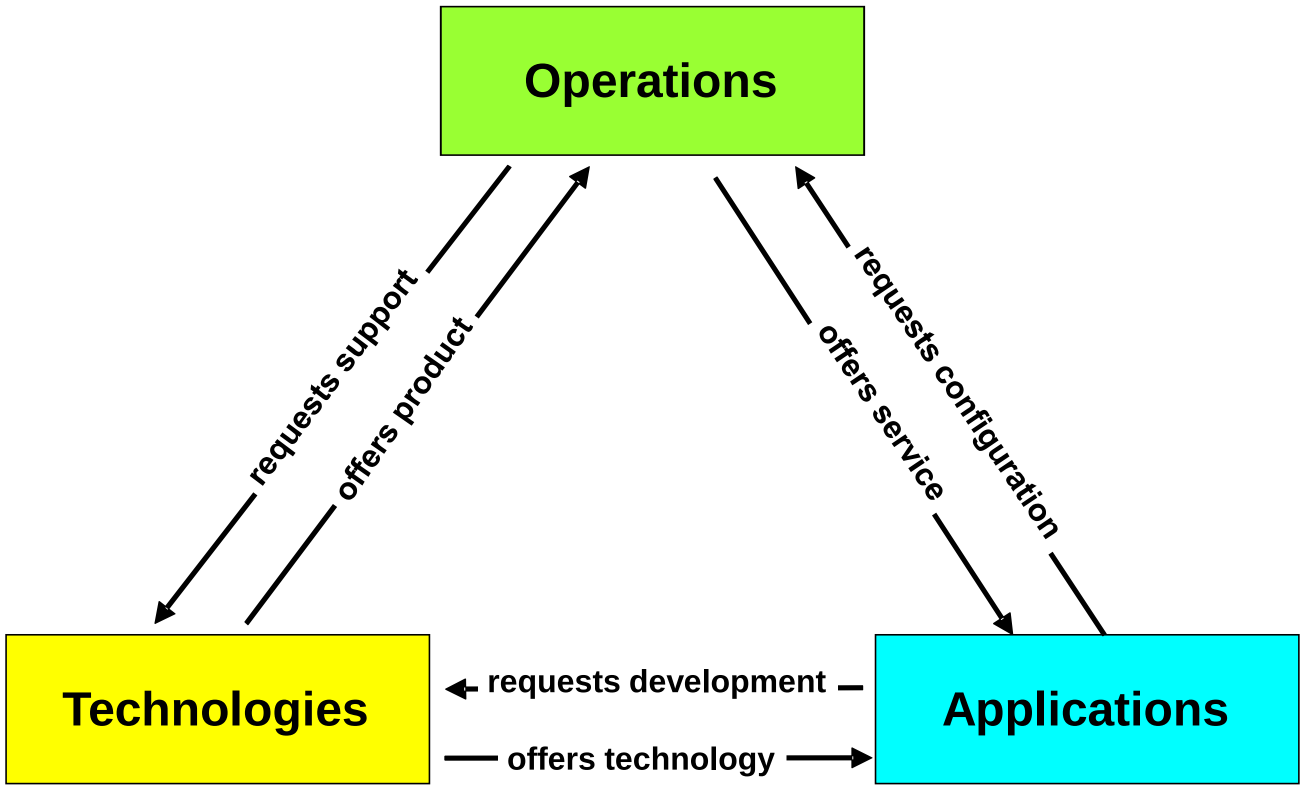
Grand Challenge projects performed on a regular basis

Top-level Europe-wide application enabling

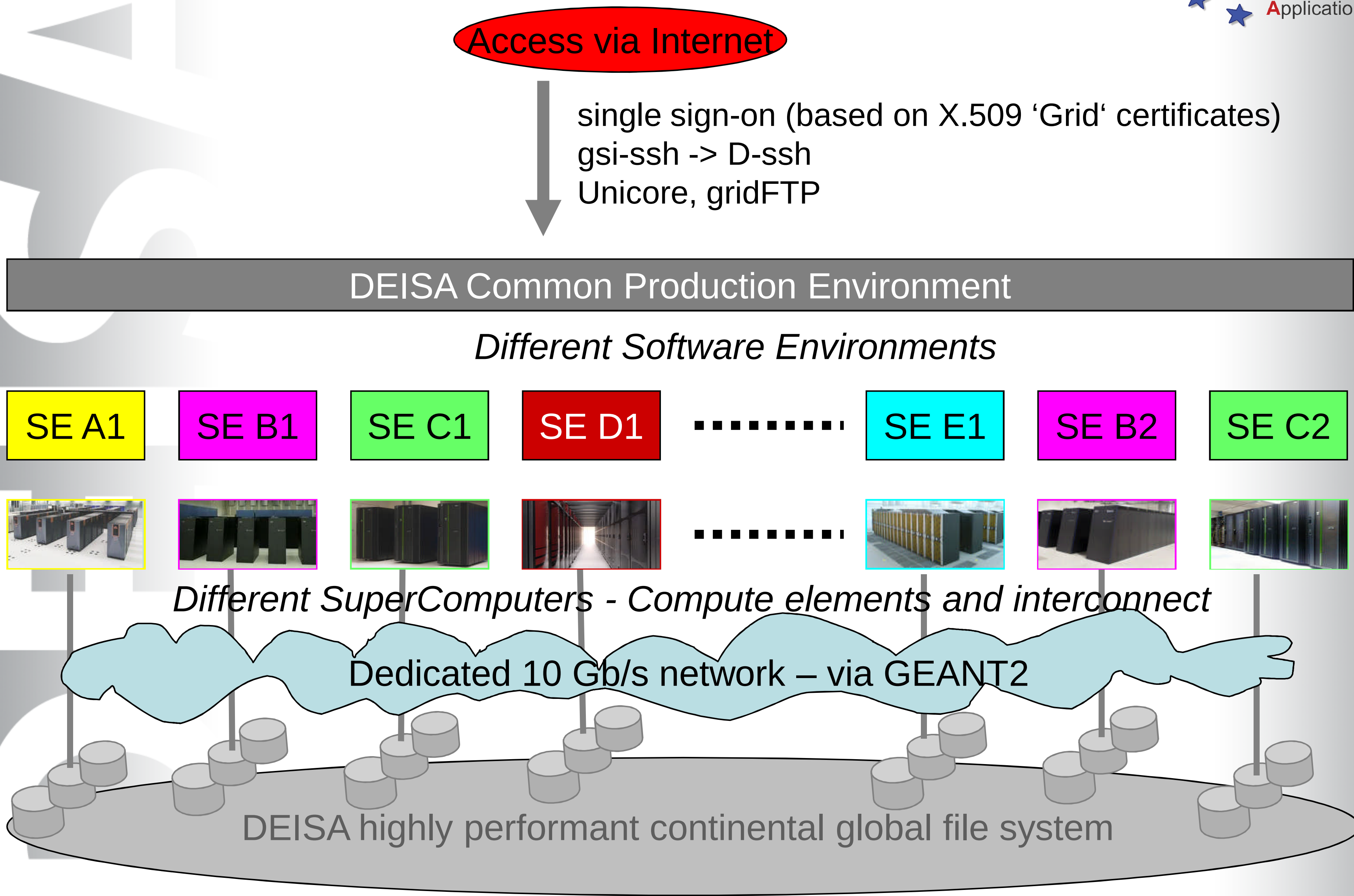
Virtual Science Community Support

RI-222919

Categories of DEISA services



Unified Access and Use of HPC Resources



RI-222919

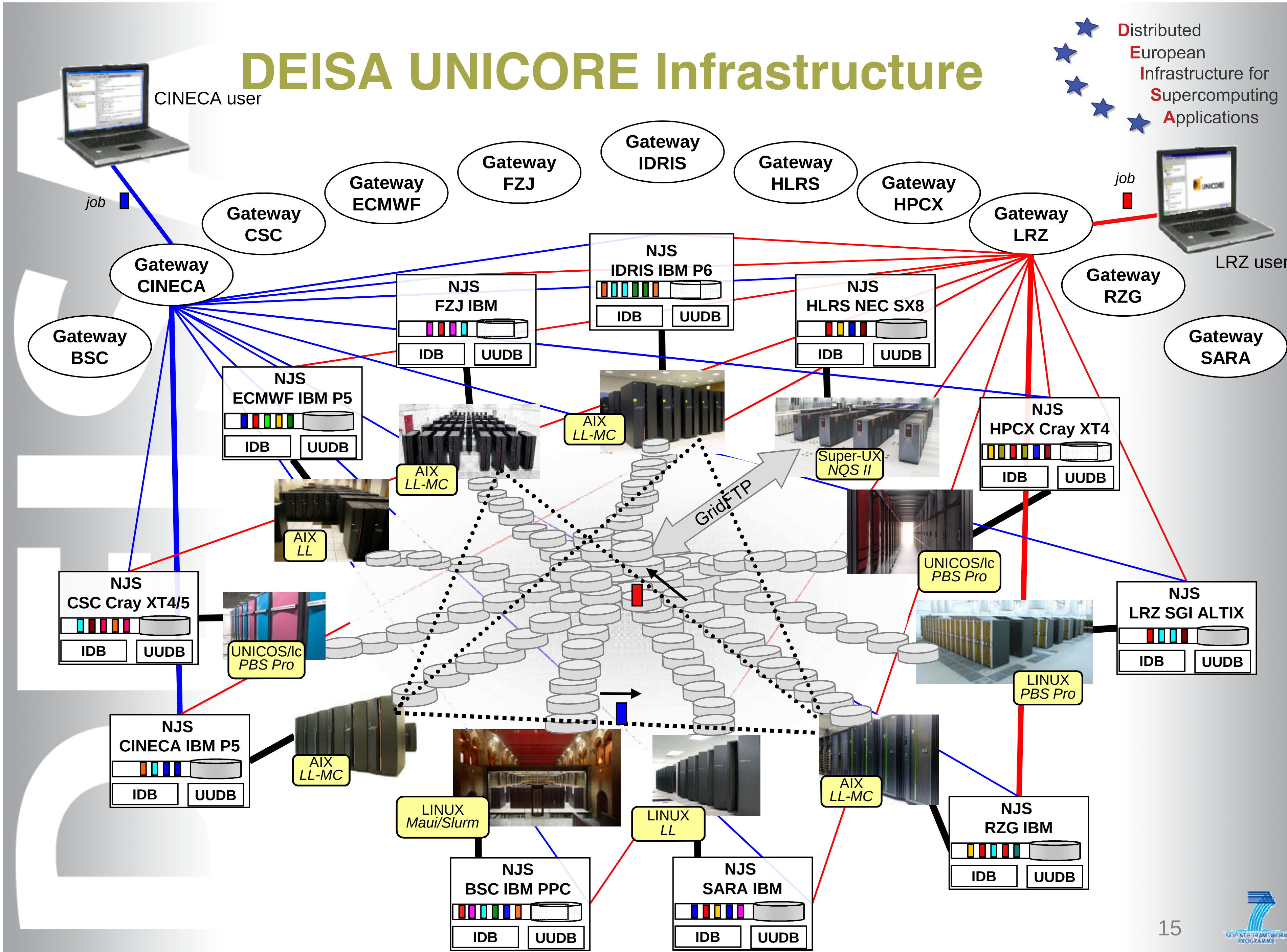
DEISA Service Layers

Distributed
European
Infrastructure for
Supercomputing
Applications

Multiple ways to access	Workflow managemnt	Common production environmnt	*Presen-tation layer*
Single monitor system	Job rerouting	Co-reservation and co-allocation	*Job manag. layer and monitor.*
Data staging tools	Data transfer tools	WAN shared File system	*Data manag. layer*
Unified AAA	DEISA Sites	Network connectivity	*Network and AAA layers*

RI-222919

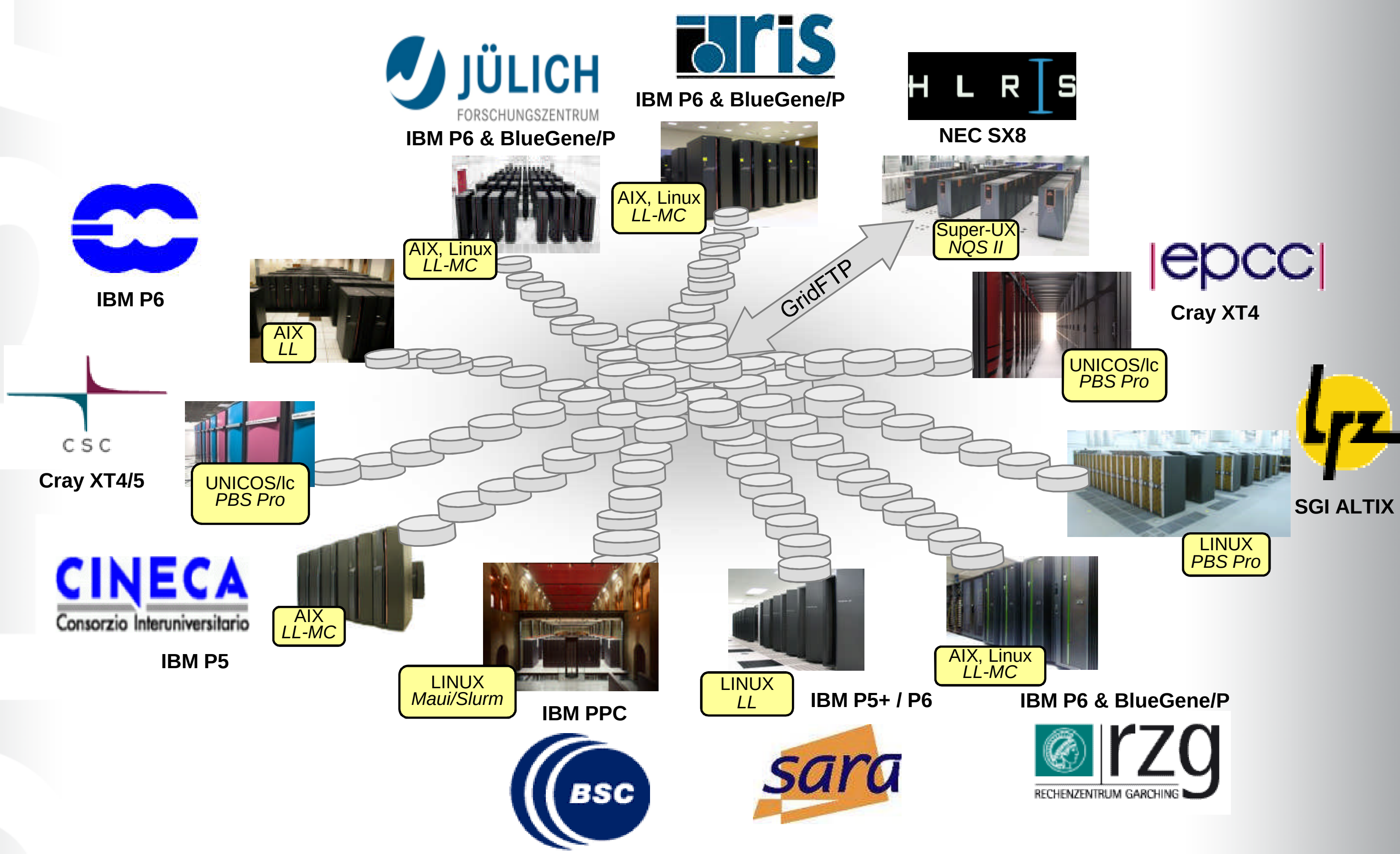
DEISA UNICORE Infrastructure



RI-222919

DEISA Global File System

Distributed
European
Infrastructure for
Supercomputing
Applications

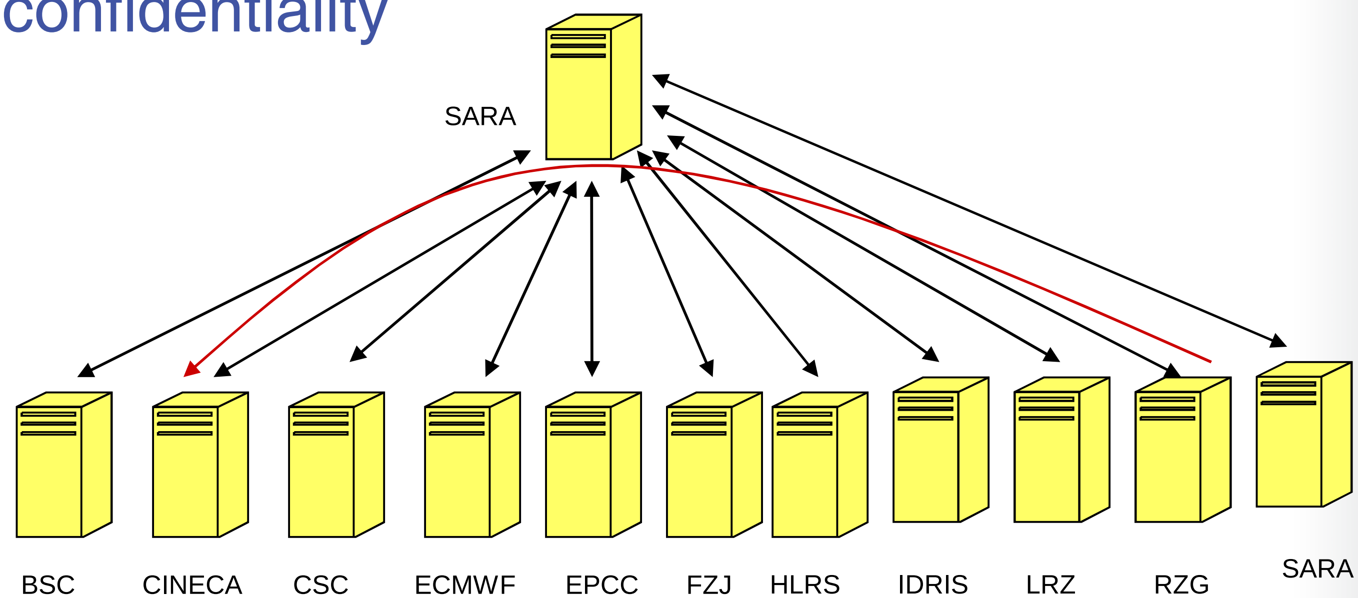


Global transparent file system based on the Multi-Cluster General Parallel File System (MC-GPFS of IBM)

Management of users in DEISA

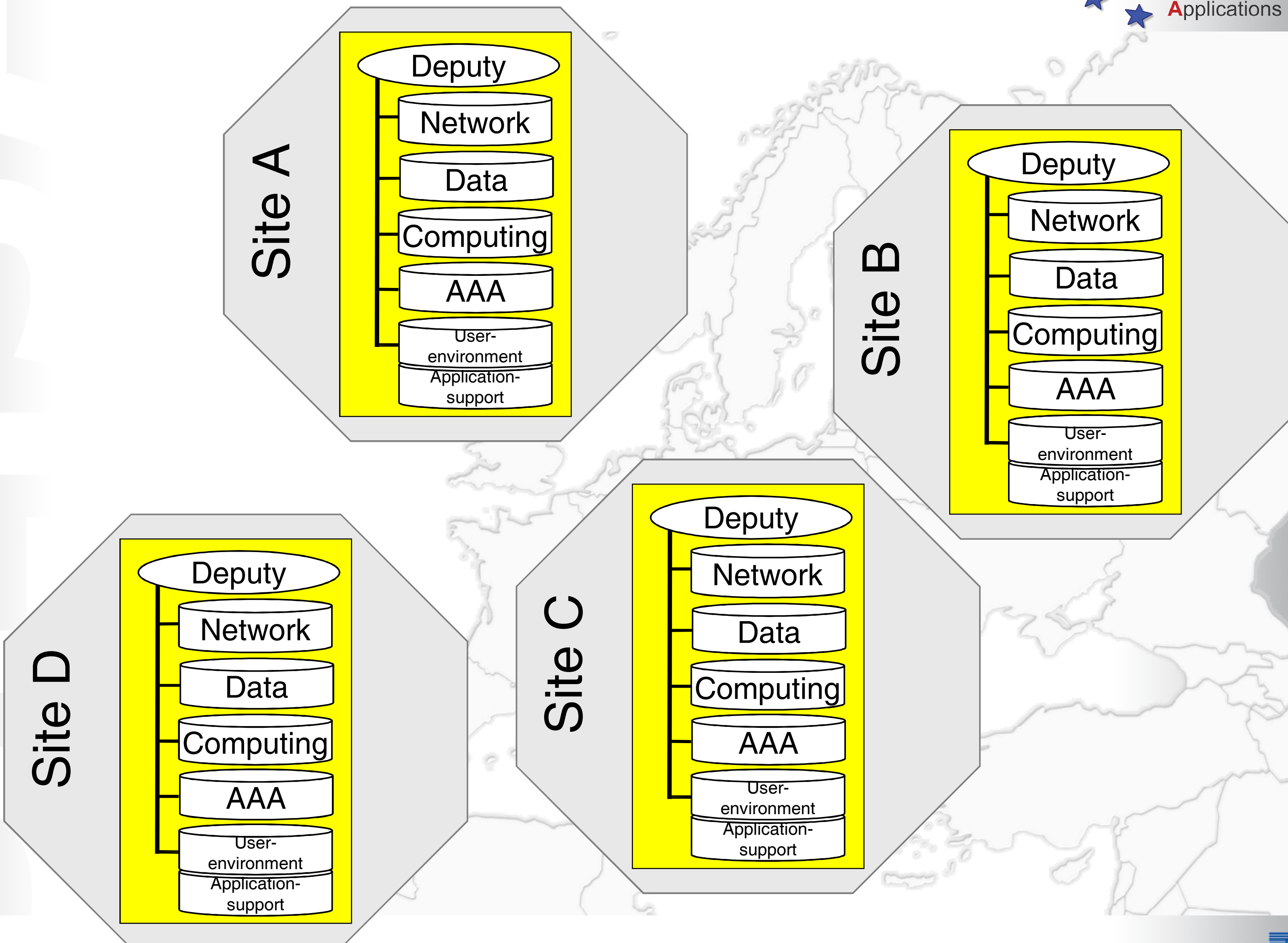
Distributed European Infrastructure for Supercomputing Applications

- A dedicated LDAP-based distributed repository administers DEISA users
- Trusted LDAP servers are authorized to access each other (based on X.509 certificates) and encrypted communication is used to maintain confidentiality

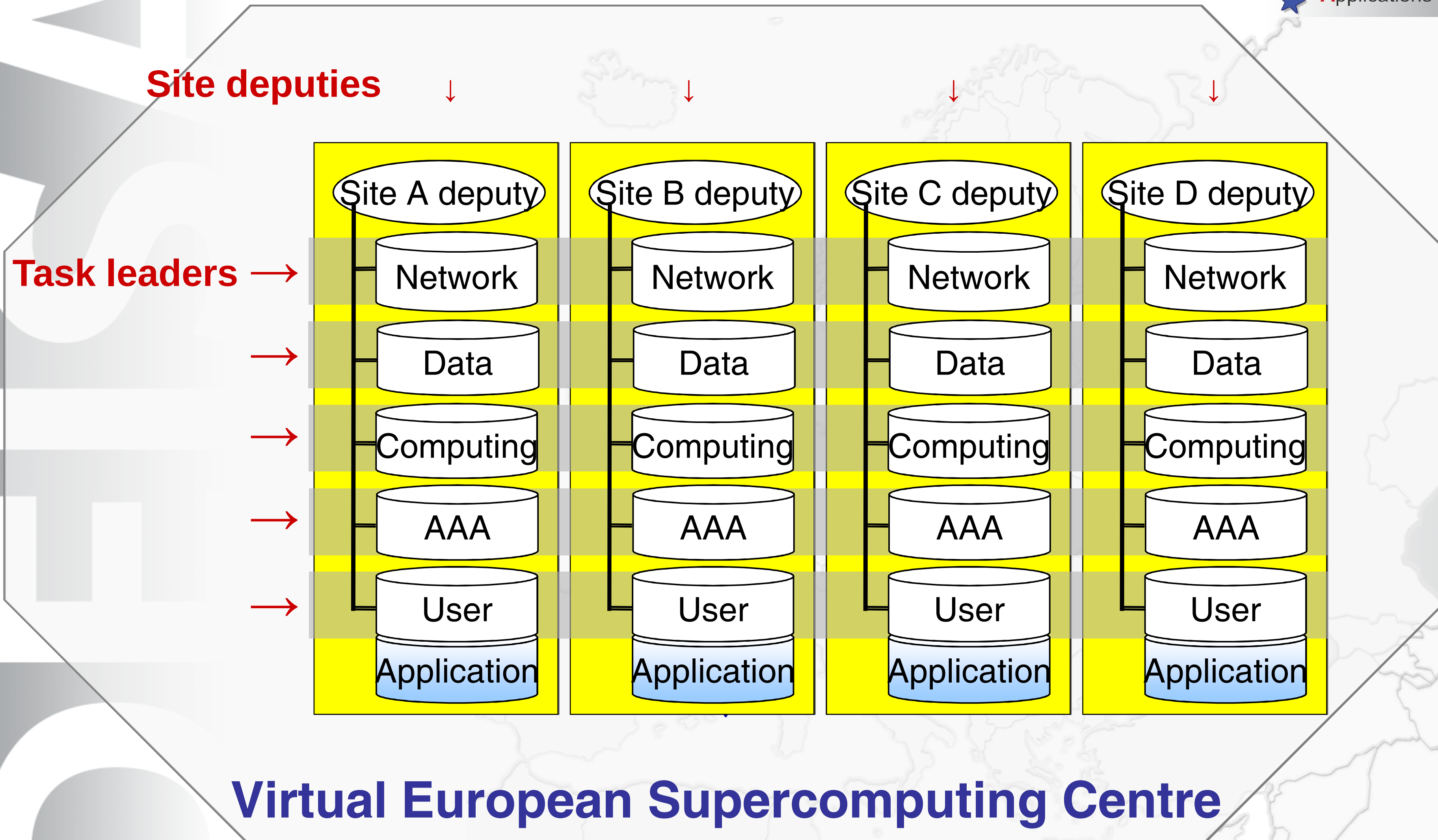


RI-222919

Federated Operation of DEISA

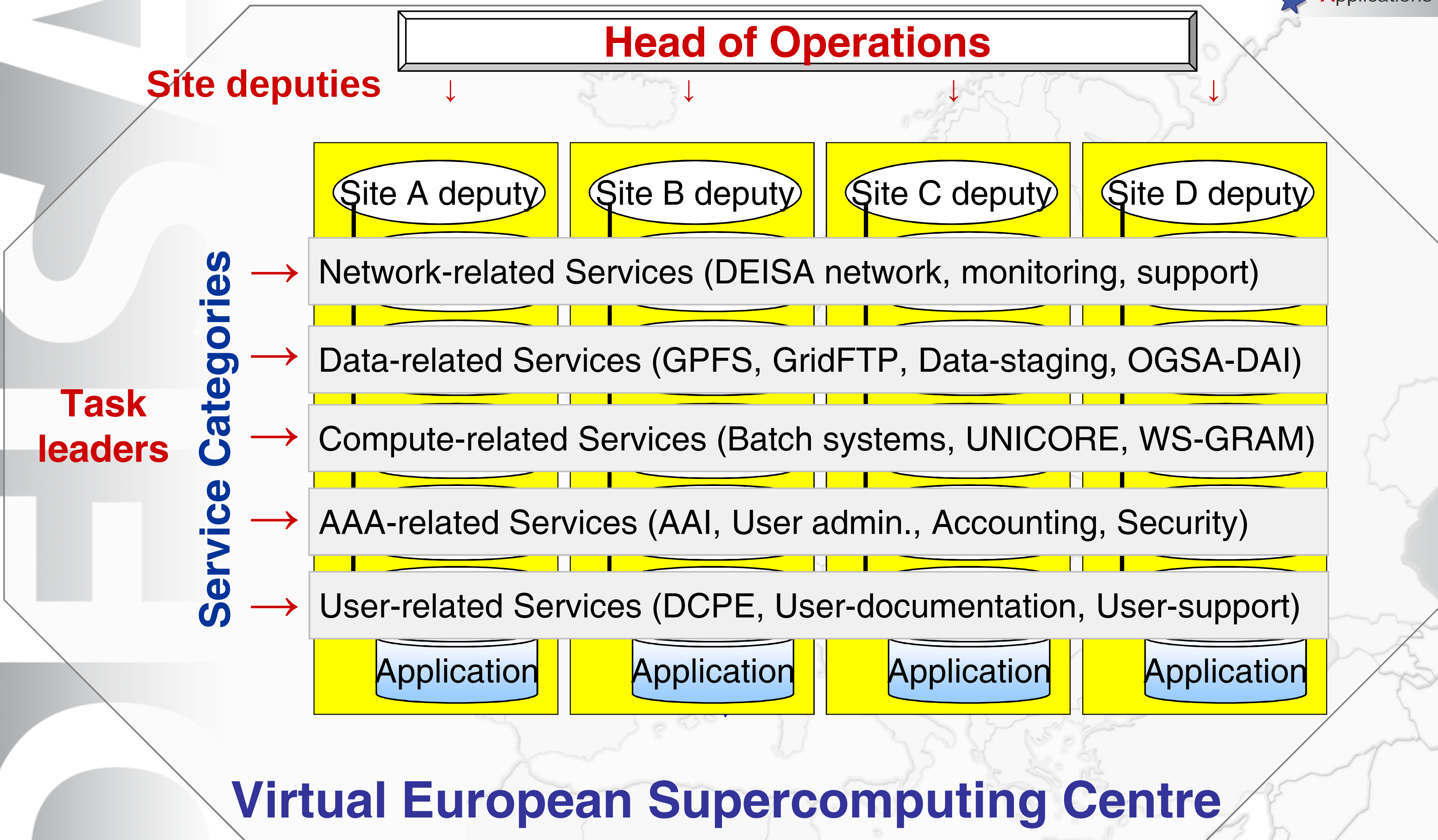


Federated Operation of DEISA



Federated Operation of DEISA

Distributed
European
Infrastructure for
Supercomputing
Applications



RI-222919

Cloud... as a Service

Cloud: dynamically scalable and virtualized resources provided as a service over the Internet

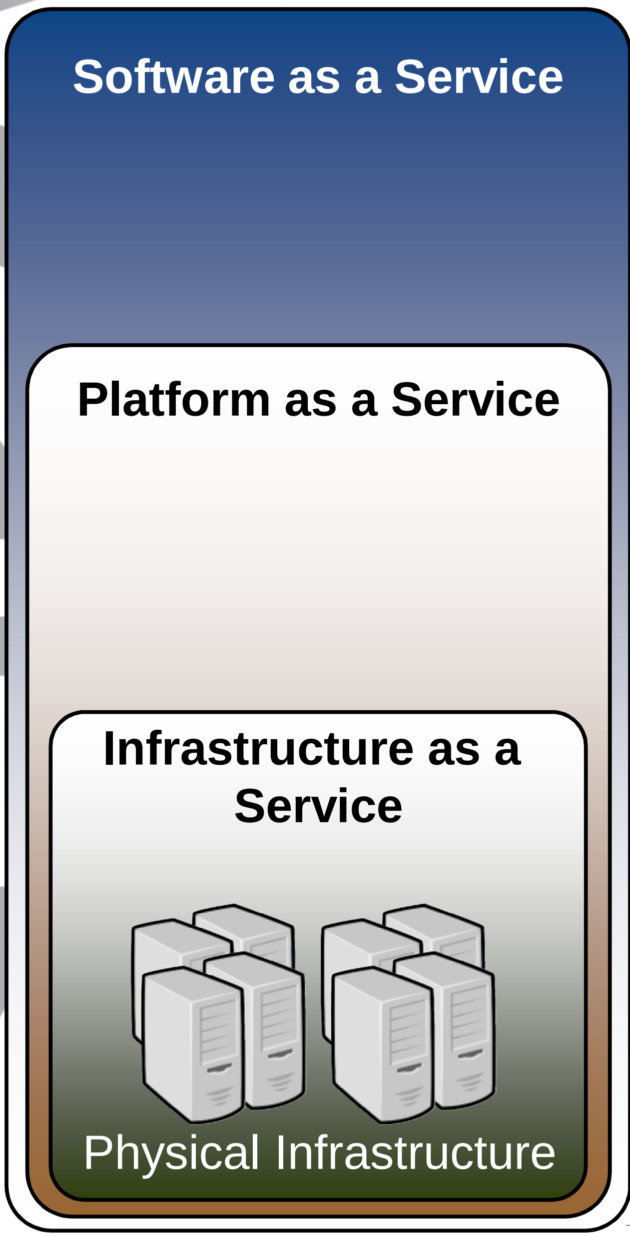
Infrastructure (IaaS)

Platform (PaaS)

Software (SaaS)

- Accessible online, anytime, anywhere
- Pay for what you use
- Available on demand
- Service Level Agreements
- Automated:
  - Scalability
  - Failover
  - Concurrency management

A Model for Delivering IT Capabilities



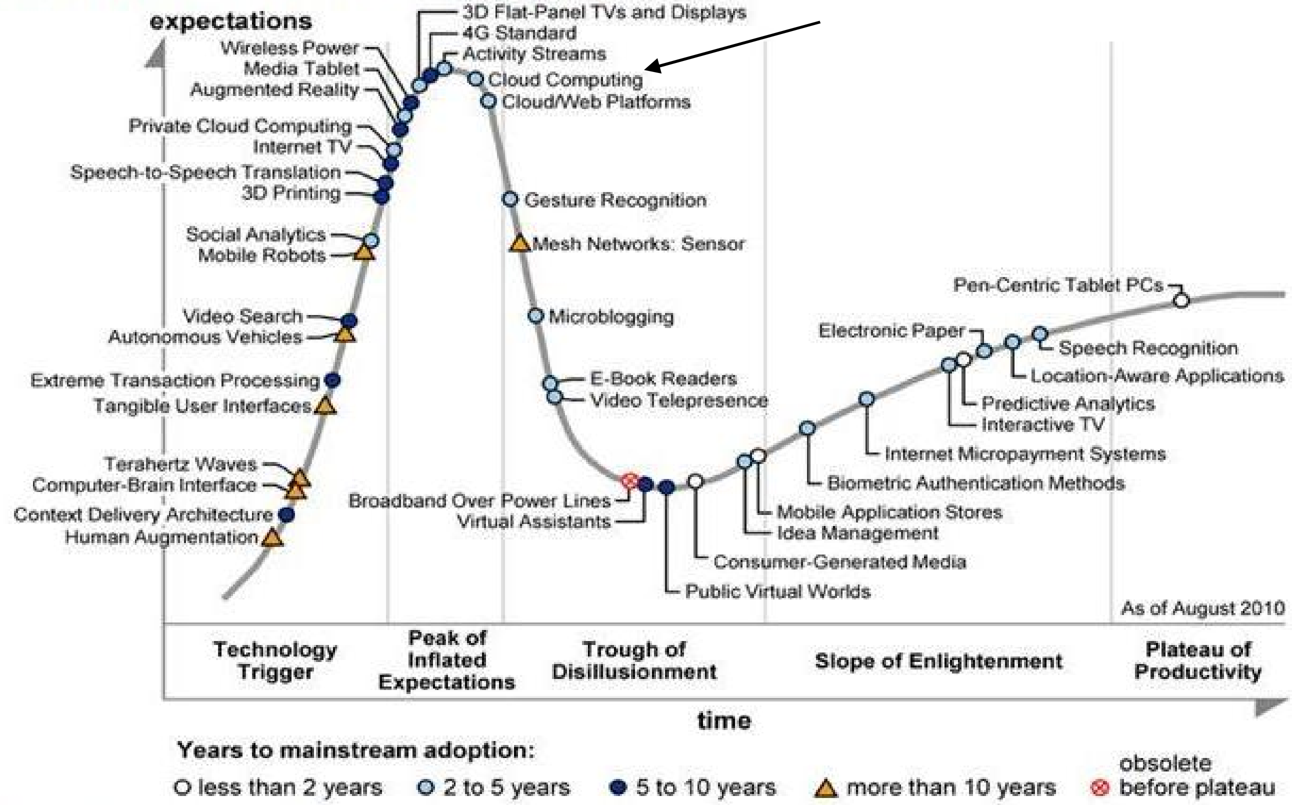
What	Who
On-demand access to any application	**End-user** (does not care about hw or sw)
Platform for building and delivering web applications	**Developer** (no managing of the underlying hw & swlayers)
Raw computer infrastructure	**System Administrator** (complete management of the computer infrastructure)

Courtesy: Ignacio Llorente

Gartner Hype Curve 2010

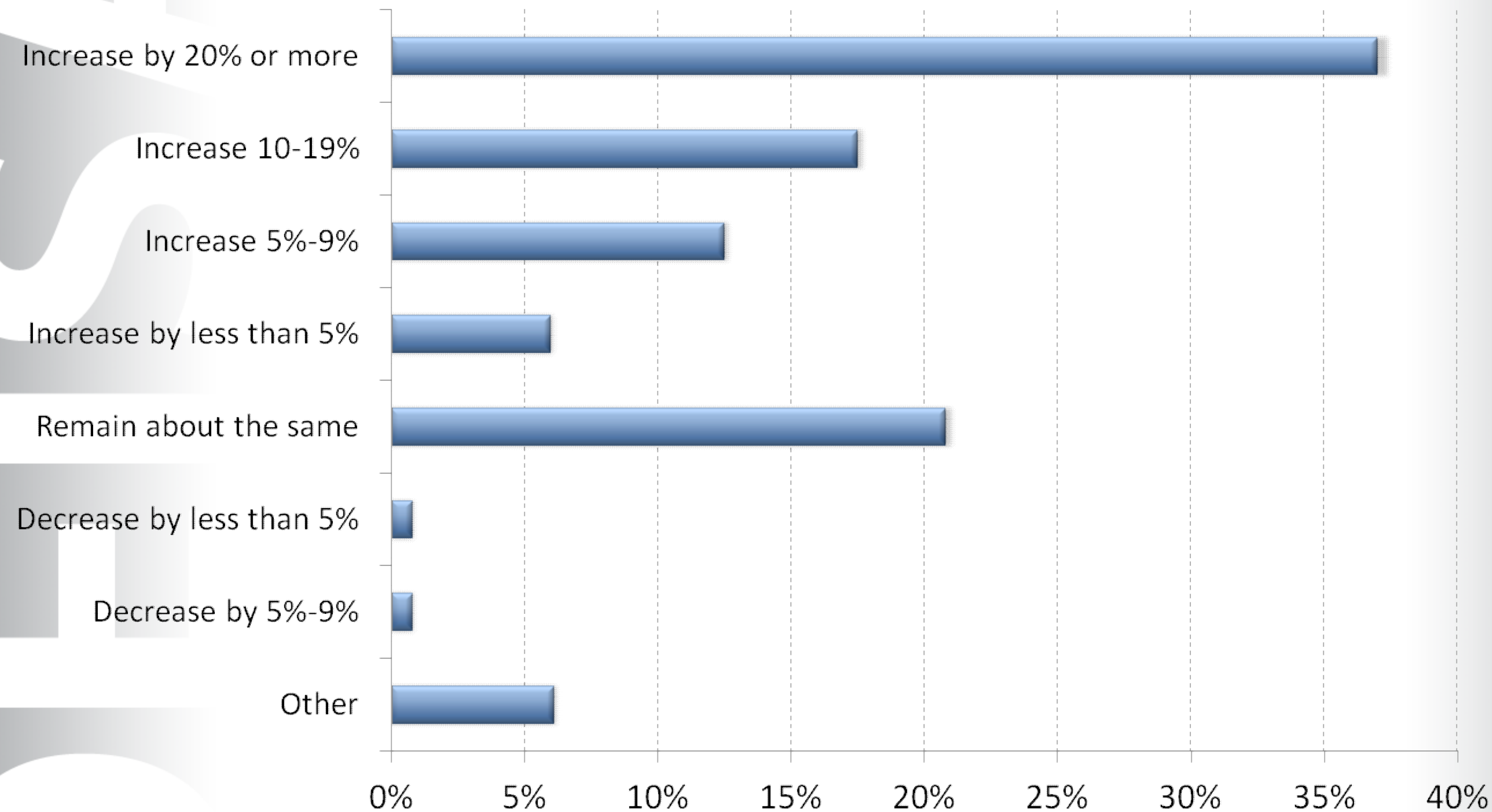
Distributed
European
Infrastructure for
Supercomputing
Applications

Figure 1 Hype Cycle for Emerging Technologies, 2010



Source: Gartner (August 2010)

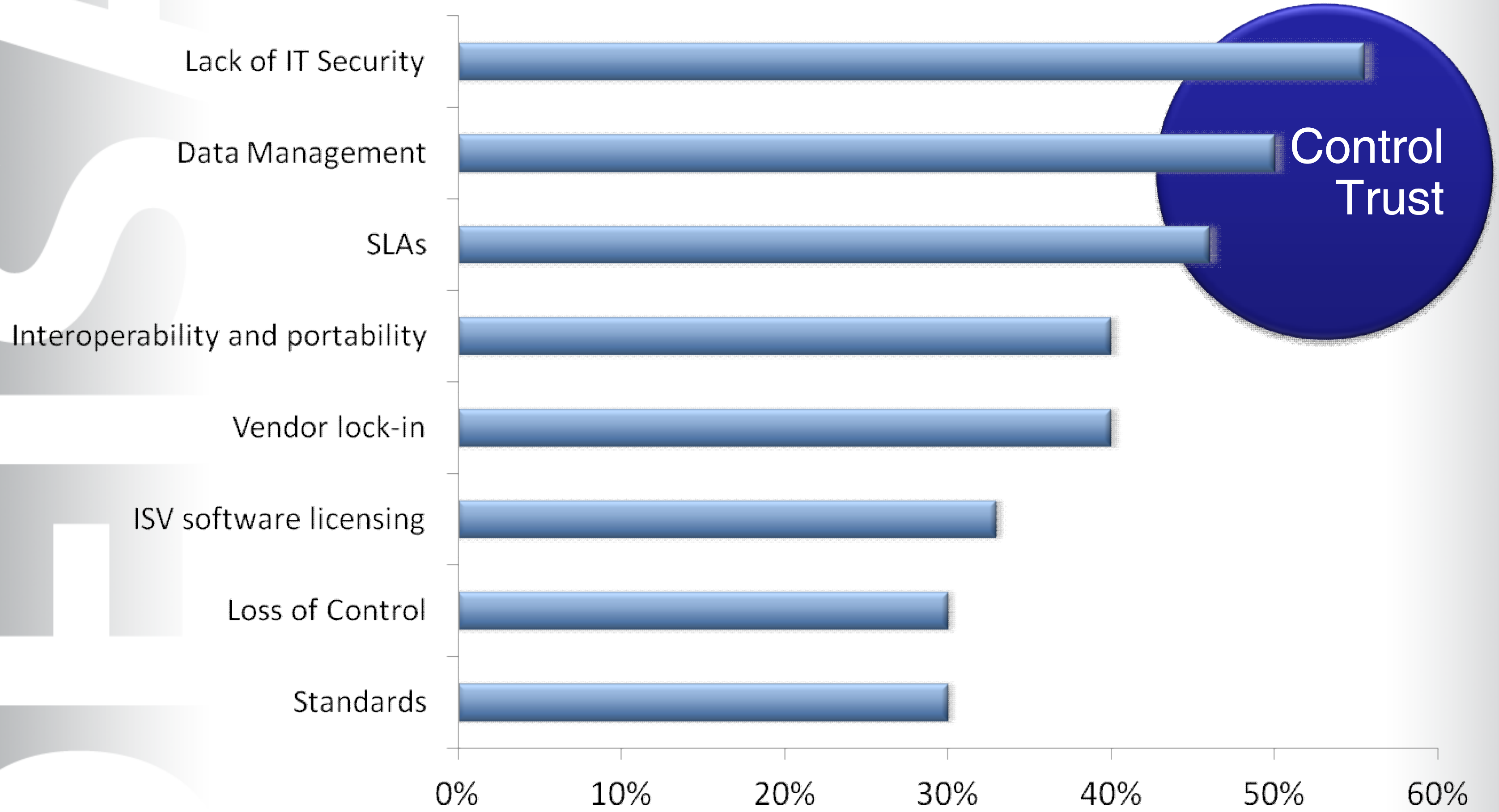
How will your budget for cloud computing change in 2011 compared with 2010



Source: John Barr, The 451 Group Cloud Adoption Survey 2010

Leading Inhibitors to Cloud Adoption – hint, it's not about the infrastructure!

Distributed
European
Infrastructure for
Supercomputing
Applications



Source: John Barr, The 451 Group Cloud Adoption Survey 2010

RI-222919

Grid versus Cloud

Why should my App run in the Grid ?

- Closer collaboration with colleagues (VCs)
- Mapping workflows to resources (plumbing)
- Suitable resources => faster/more/accurate processing
- Different architectures serve different apps
- Failover: move jobs to another system

. . . and why in the Cloud ?

- No upfront cost for additional resources
- CapEx => OpEx, pay-per-use
- Elasticity, scaling up and down
- Hybrid solution (private and public cloud)

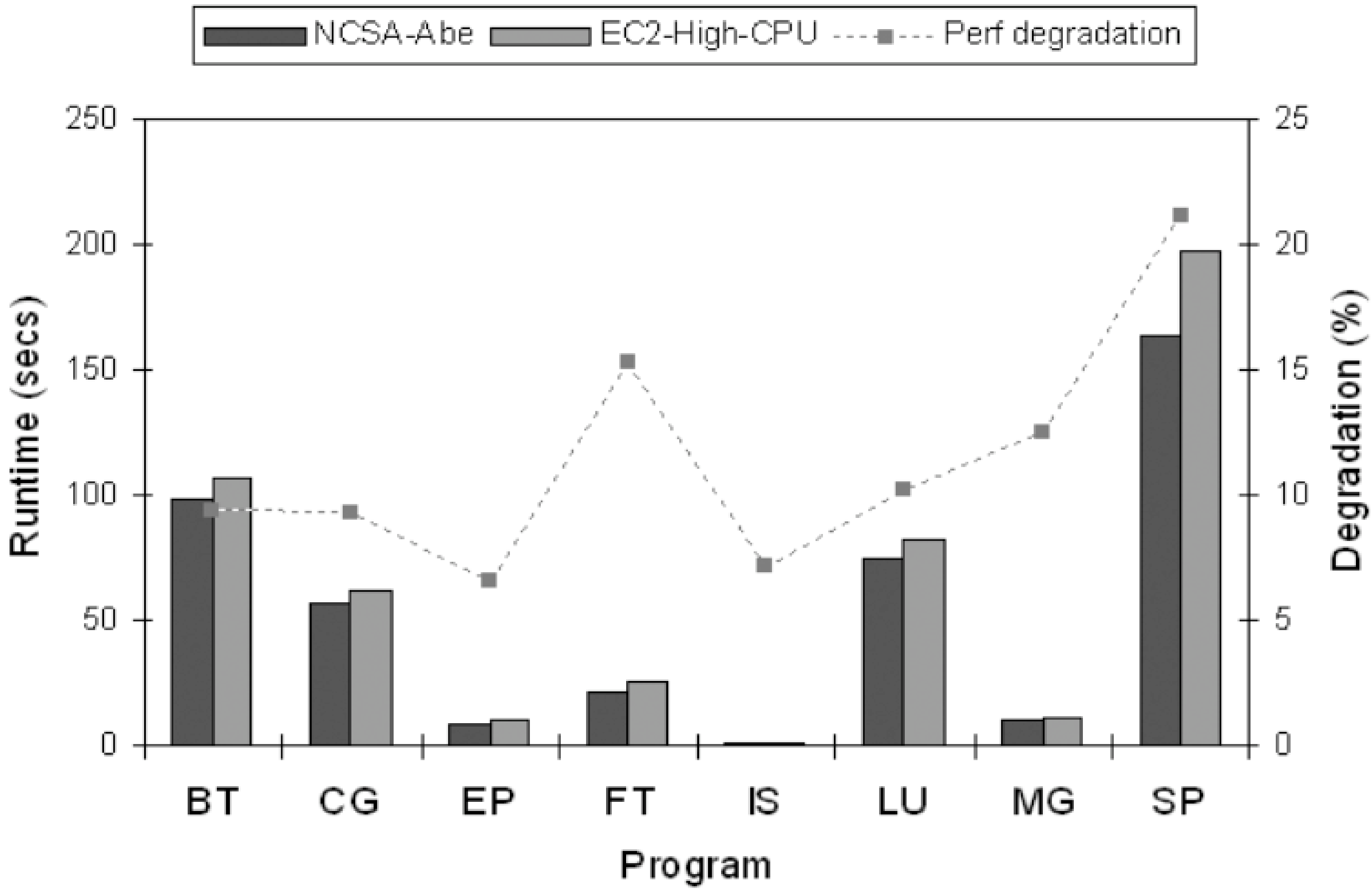


FIGURE 1. NPB-OMP (CLASS B) RUNTIMES ON 8 CPUS ON EC2 AND NCSA CLUSTER COMPUTE NODES. OVERLAID IS THE PERCENTAGE PERFORMANCE DEGRADATION IN THE EC2 RUNS.

Ed Walker, Benchmarking Amazon EC2 for high-performance scientific computing, ;Login, October 2008.

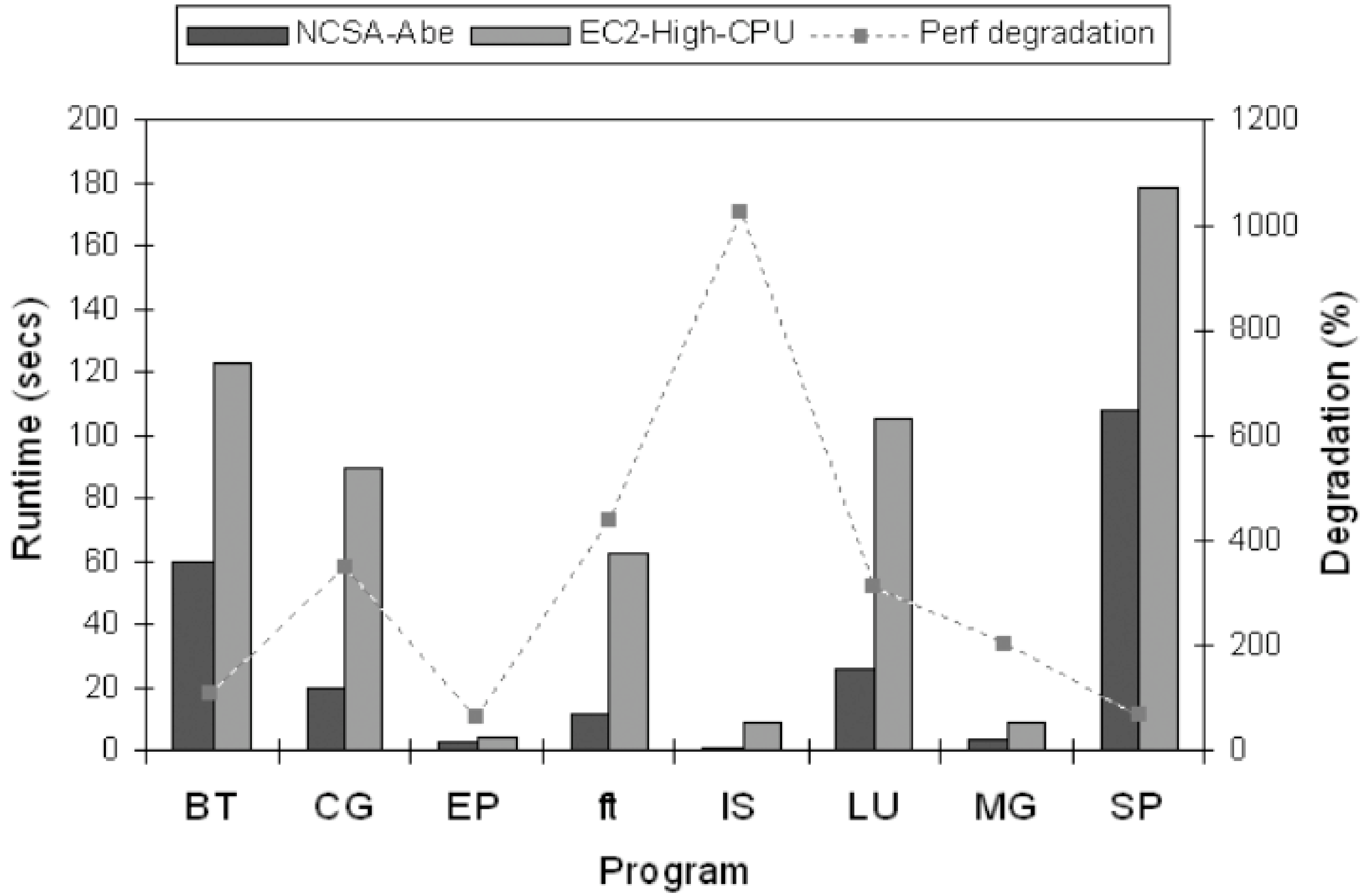
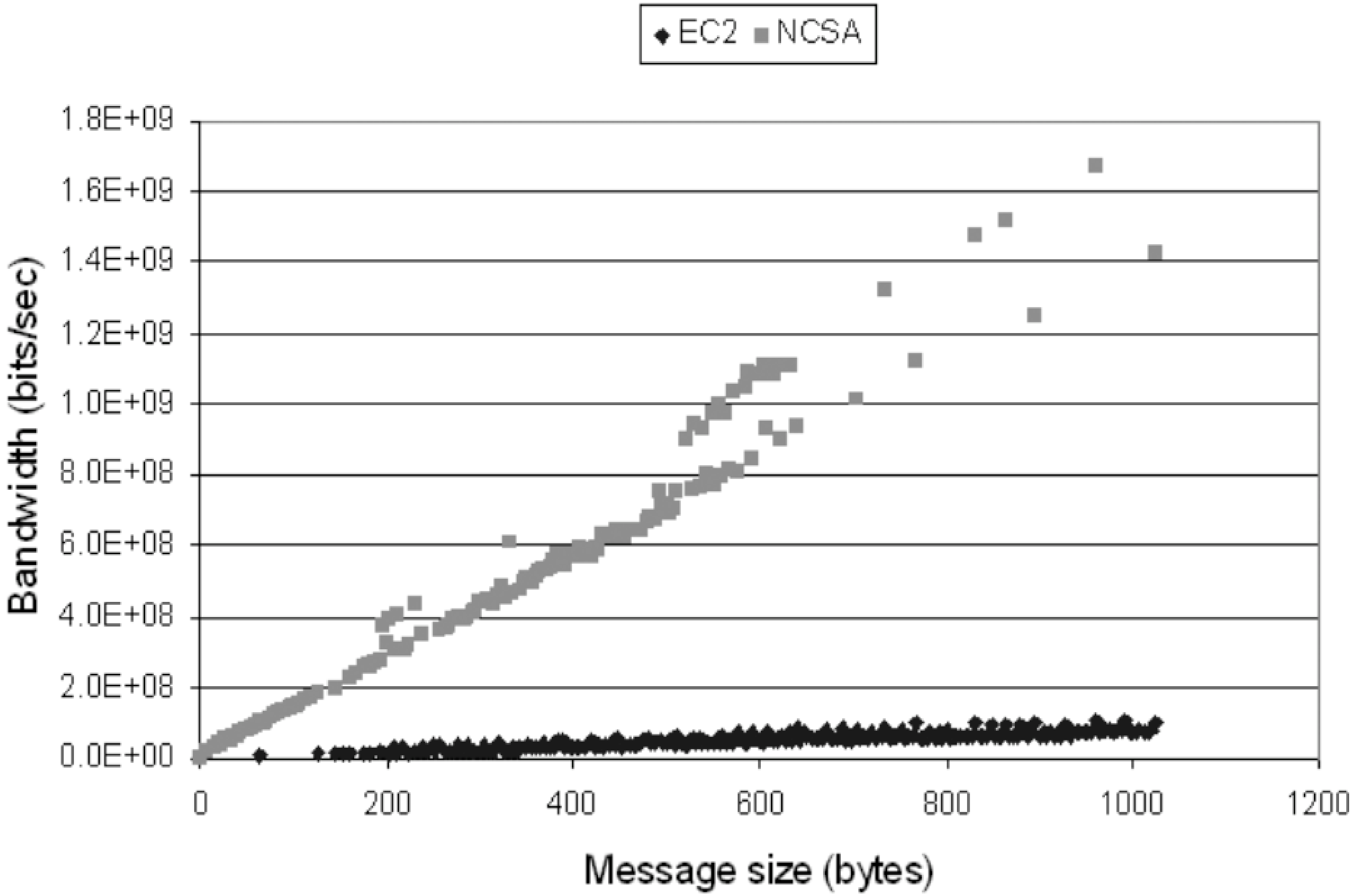
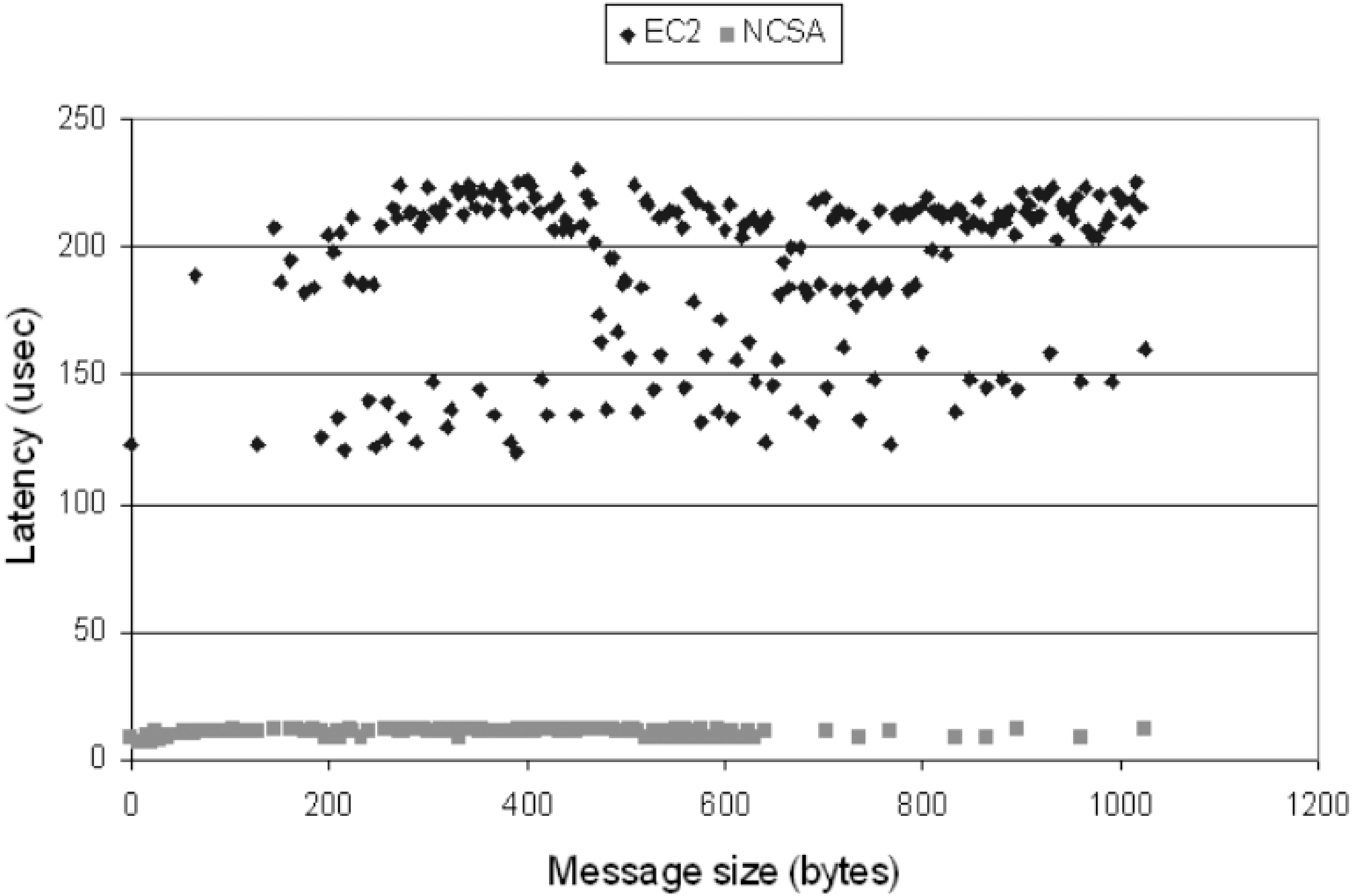


FIGURE 2. NPB-MPI (CLASS B) RUNTIMES ON 32 CPUS ON THE NCSA AND EC2 CLUSTER. BT AND SP WERE RUN WITH 16 CPUS ONLY. OVERLAID IS THE PERCENTAGE DEGRADATION IN THE EC2 RUNS.

Ed Walker, Benchmarking Amazon EC2 for high-performance scientific computing, ;Login, October 2008.



Ed Walker, Benchmarking Amazon EC2 for high-performance scientific computing, ;Login, October 2008.



Ed Walker, Benchmarking Amazon EC2 for high-performance scientific computing, ;Login, October 2008.

Loosely coupled problems are suitable for the Cloud

- Ensemble runs to quantify **climate model uncertainty**
- Identify **potential drug targets** by screening database of ligand structures against target proteins
- Study **economic model sensitivity** to parameters
- Analyze **turbulence dataset** from many perspectives
- **Numerical optimization** to determine optimal resource assignment in energy problems
- Mine collection of data from **advanced light sources**
- Construct databases of **chemical compounds** properties
- Analyze data from the **Large Hadron Collider**
- Analyze **log data** from 100,000-node parallel computations

Clouds and supercomputers: Conventional wisdom?

	Loosely coupled applications	Tightly coupled applications
Clouds/ clusters	✓	Too slow
Super computers	Too expensive	✓

Courtesy
Ian Foster

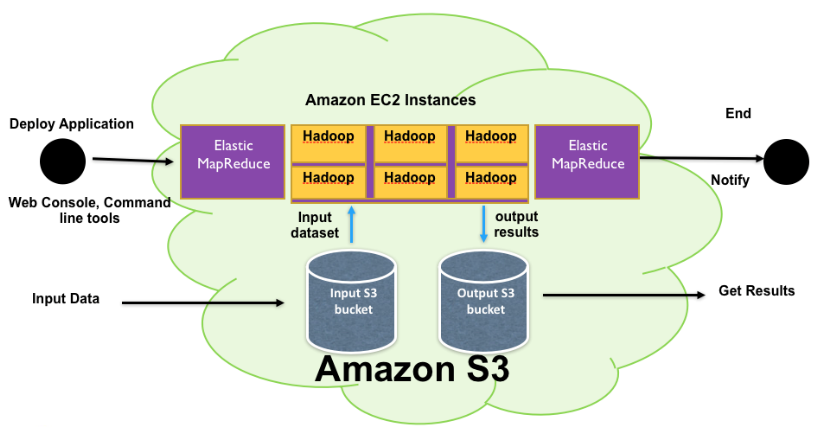
Finally:
Amazon July 2010
Introducing CCI Cluster Compute Instances

- New Amazon EC2 instance type
- Optimized for network intensive computing
  - Low latency
  - High bandwidth
  - New EC2 API: Placement groups
- Instance Configuration
  - 2 * Xeon 5570 (Intel "Nehalem" quad-core architecture)
    - 33.5 Elastic Compute Units
  - 23 GB 1333MHz DDR3 Registered ECC RAM
  - 1690 GB of local instance storage
  - 10 Gbps Ethernet interconnects CCI's

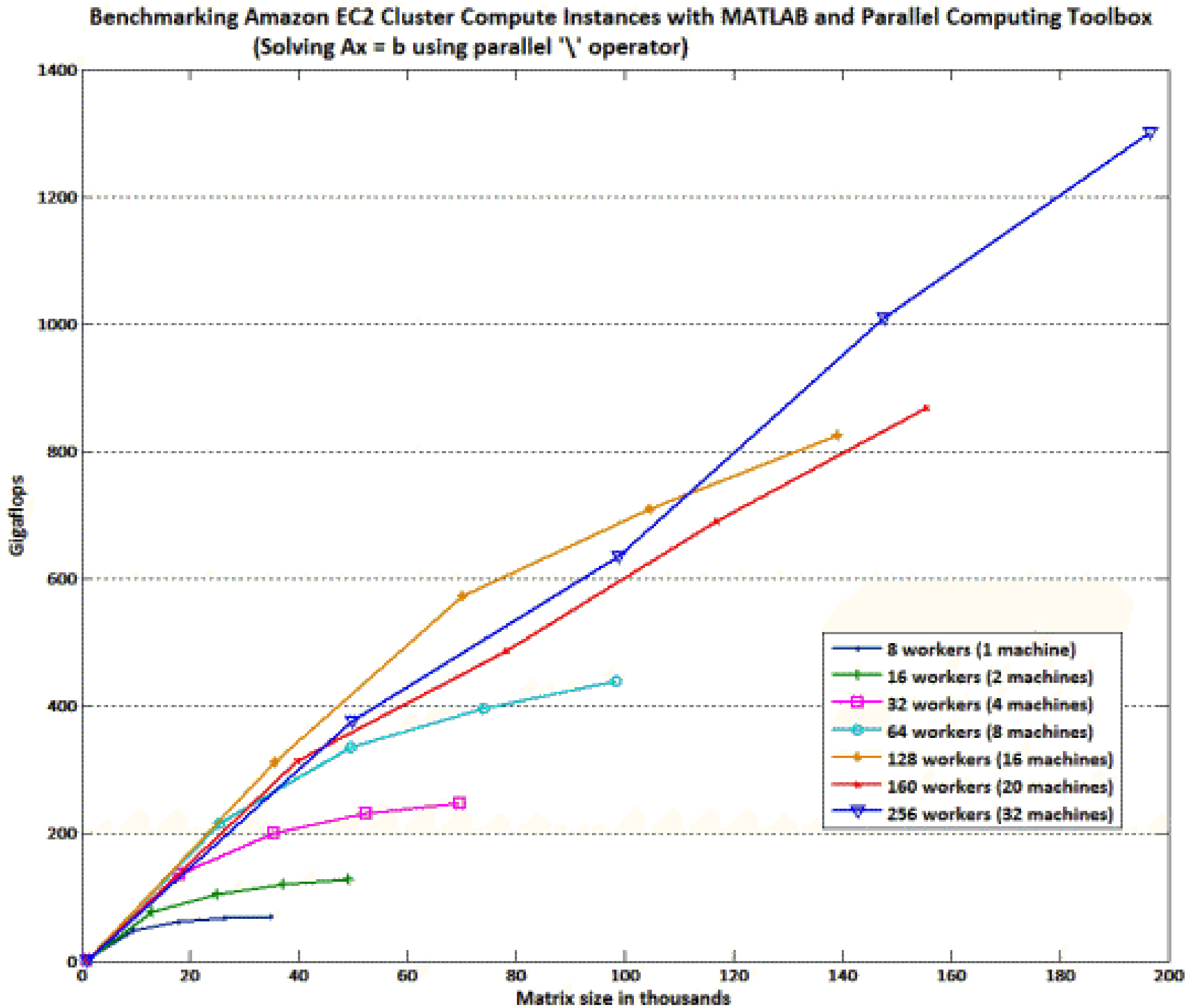
Some EC2 CCI Results

- Some applications can expect 10x better performance
- LNBL NERSC saw 8.5x compared to similar clusters on standard EC2 instances
- Linpack benchmark
  - 880-instance CC1 cluster
  - Performance: 41.82 Tflops

➔EC p 500
ran



MATLAB Ax = b Benchmark, Sept 2010 Results

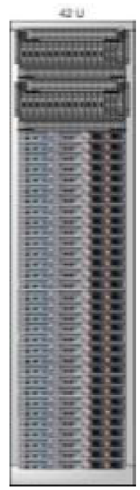


Finally: Pricing

- On-demand
  - $1.60/hr
  - Linpack: 880 CCIs (7040 cores)
  - 1.6*24*30*880 = $1M per month, $12M per year
- Reserved Instance
  - 1 yr: $4290 one time + $0.56/hr
  - 3 yr: $6590 one time + $0.56/hr
  - Linpack: 880 CCIs (7040 cores)
  - $4.3M per year
- In general, too expensive
- Best solution: fire-drill problems, additional resources are needed immediately, for a restricted period of time.
- It takes 6 months on average to procure, deploy and activate new (own) resources !
- All you need is some ‘Cloud Adapter‘ software (e.g. Oracle Grid Engine, Univa UniCloud, Eucalyptus, Nimbus, OpenNebula, etc)

But:

Case Study: Managed/Cloud Service

Without Mellanox InfiniBand	With Mellanox	Savings
 256 VMs 4 racks, 4U servers 4X proc, 16GB 16 edge switches 192 I/O cards $744K capital cost	 256 VMs 1 rack, 1U servers 4X proc, 16GB 2 I/O Directors 32 I/O cards $347K capital	**Cap Ex:** **$397,000** **Op Ex:** Floor space: **$54,000** Power: **$20,000** MAC*: **$104,000** **TOTAL SAVINGS** **$575,000** (32 servers over 3 years) *2 moves/adds/changes per server per year

Customer profile : Web-based travel transactions

Volume exceeding Amazon.com transactions

Courtesy: Gilad Shainer, Mellanox

Source: Xsigo

Challenges in the Cloud

Distributed
European
Infrastructure for
Supercomputing
Applications

- Many HPC algorithms have to be optimized towards underlying computing architecture for best performance
- In the past, on vector, parallel, grid, and dataflow systems, the system architecture was known to the user
- Then, hand-optimization was possible by restructuring the core algorithms
- In the cloud, user has no information about individual system components, overall architecture, heterogeneity, etc
- In the cloud, in addition, virtualization hides the physical architecture from the user, and introduces additional performance uncertainties

Lessons Learned & Recommendations

Distributed
European
Infrastructure for
Supercomputing
Applications

- HPC Clusters for number crunching hand-optimized codes
- Grids for collaborative computing and for ‘plumbing‘
- Fact is: Clouds are becoming the new Utility
- In the Cloud: no infos about system specifics, architecture, virtualization, network, etc...
- Successively, Clouds become inexpensive and will soon show better performance
- Engineering software ASPs will move from licenses-based models to service-oriented pay-per-use models.

Distributed European Infrastructure for Supercomputing Applications

Thank You for your attention

Gentzsch @ rzg.mpg.de