Grid computing: yesterday, today and tomorrow?

Dr. **Fabrizio Gagliardi**

EMEA Director

External Research

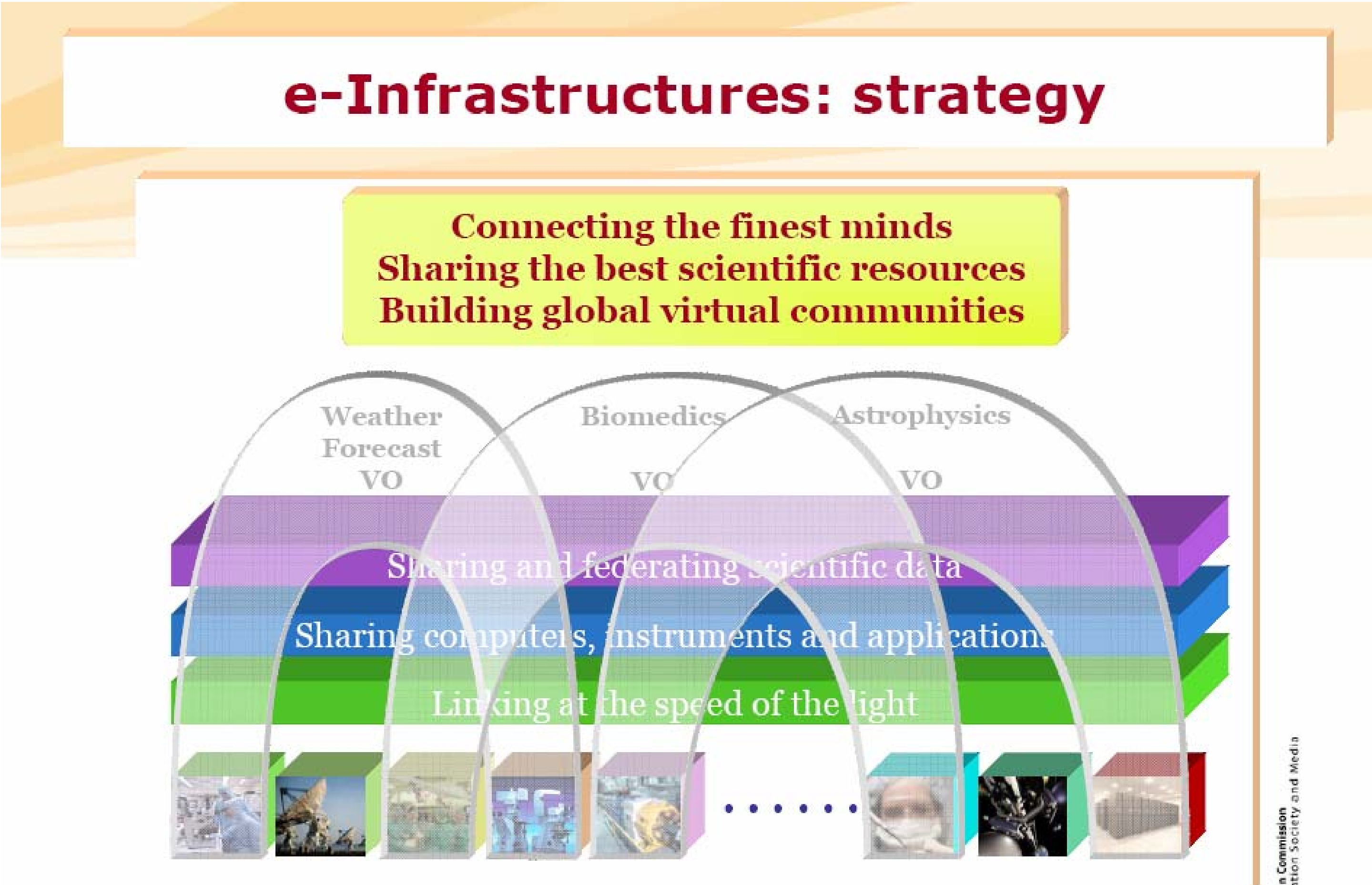
Microsoft Research

Cracow Grid Workshop 2008
Cracow, October 14th

Outline

- Yesterday and today:
  - Achievements in the area of e-Infrastructures and Grid computing
  - Examples beyond e-Science
  - Issues : Complexity, Cost, Security, Standards
- The future:
  - Cloud Computing, Virtualisation, Data Centers, Software as a Service, Multi-core architectures, Green IT
- Conclusions

The European Commission strategy for e-Science

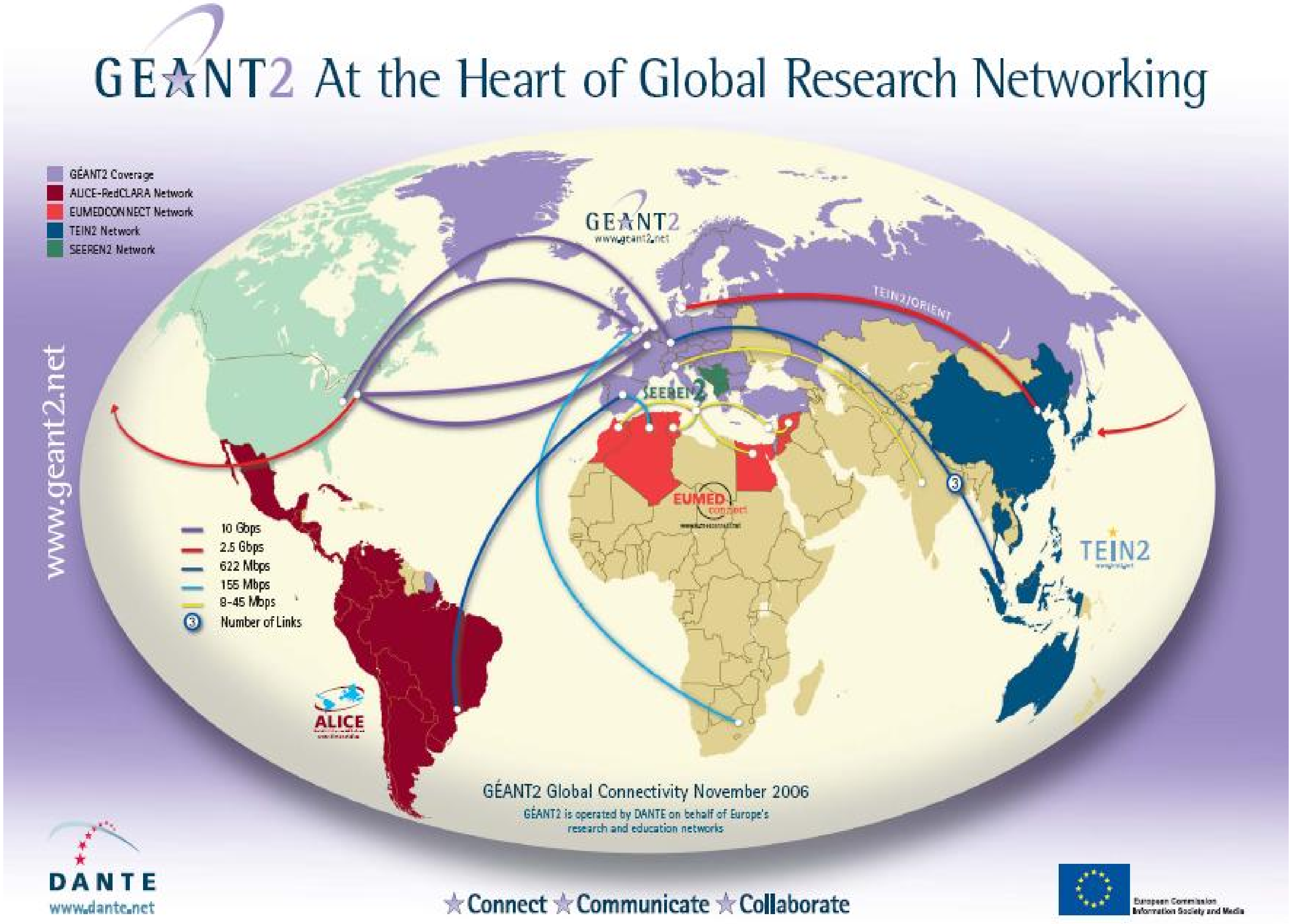


Mario Campolargo, EC, DG INFSOM, Director of Directorate F: Emerging Technologies and Infrastructures

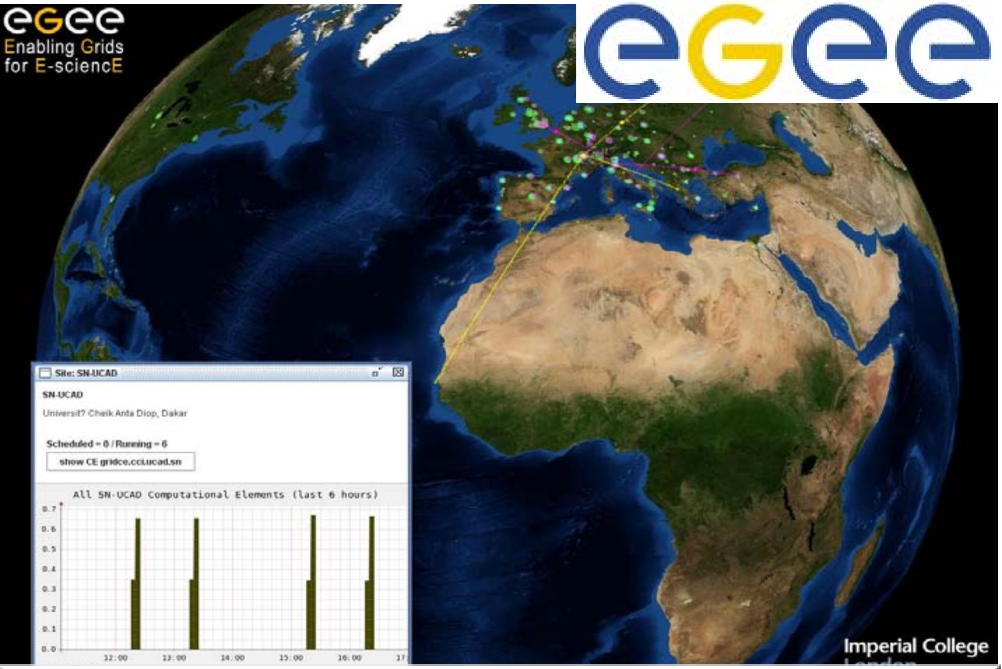
e-infrastructure

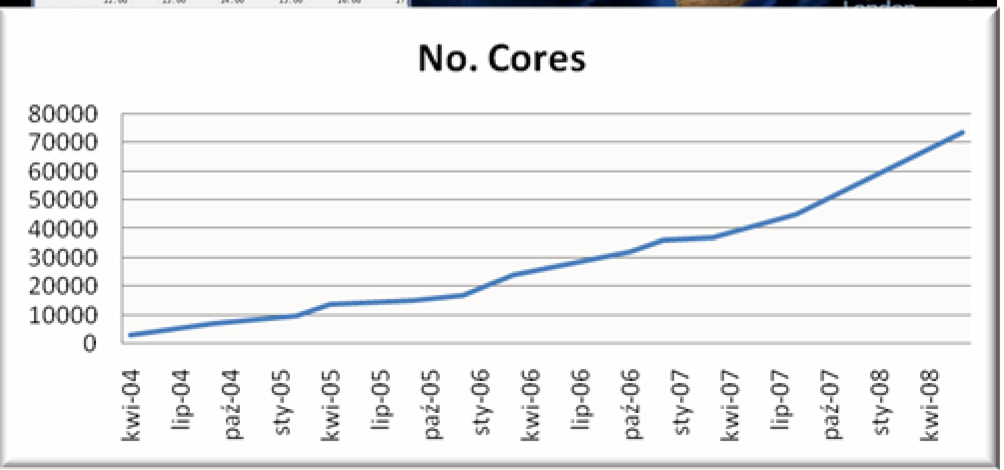
http://cordis.europa.eu/fp7/ict/programme/events-20070524_en.html

e-Infrastructure achievements: Research Networks

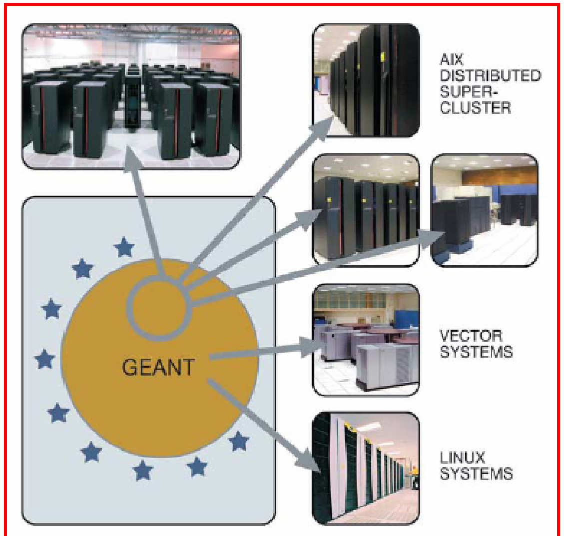


e-Infrastructure HPC achievements: EGEE and DEISA





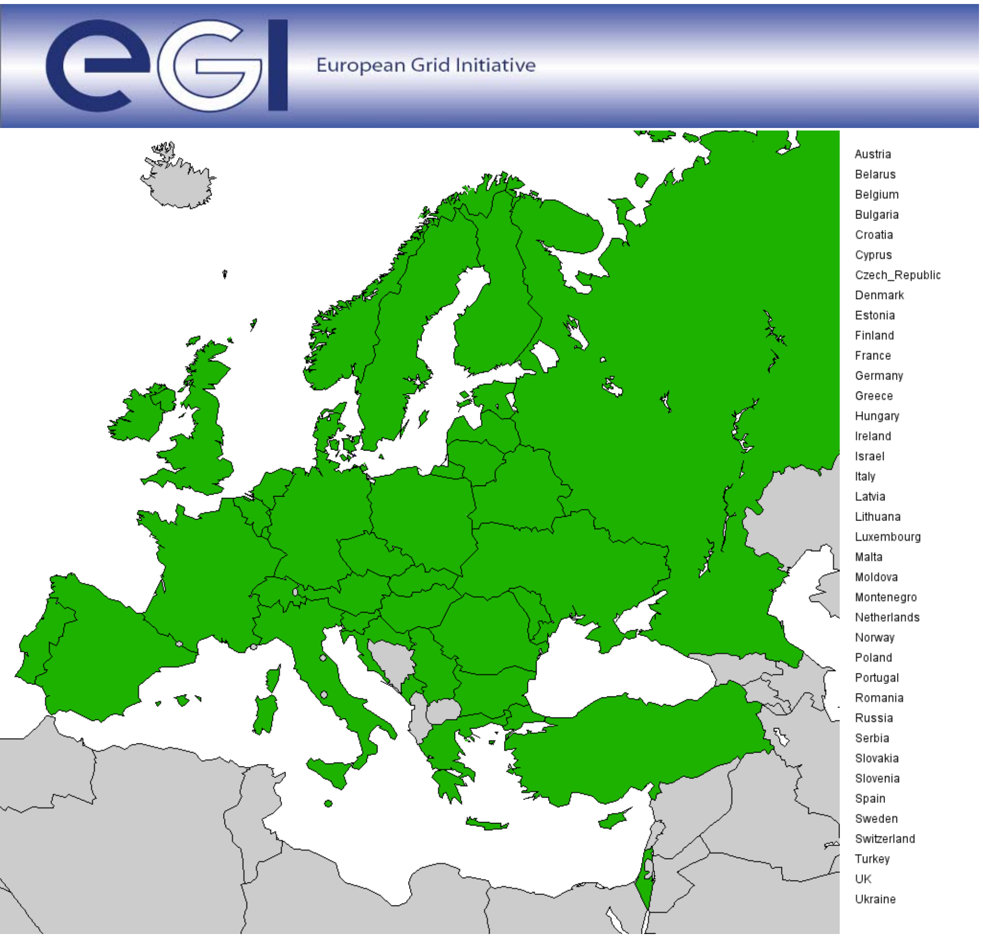
DEISA

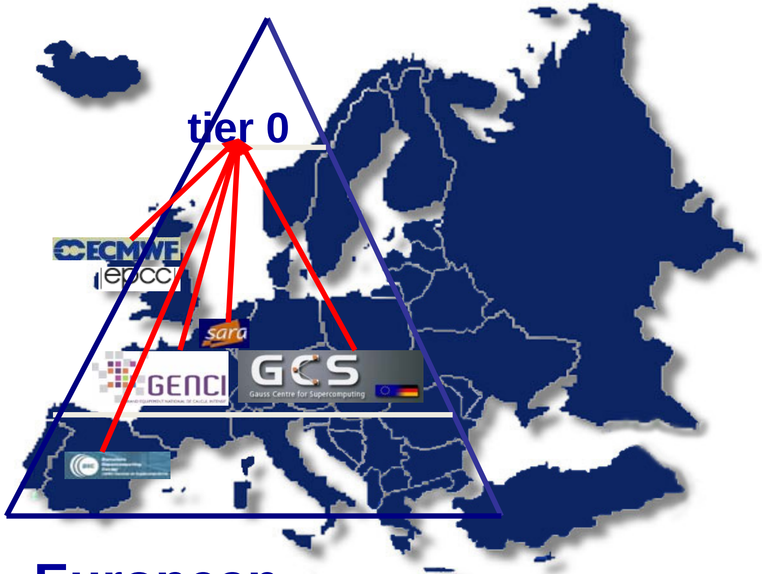


The DEISA supercomputing Grid

Distributed
European
Infrastructure for
Supercomputing
Applications

e-Infrastructure HPC next steps: EGI and PRACE





European Ecosystem

In summary:
Grid achievements for e-Science

- Grid for e-Science: mainly a success story!
  - Several maturing Grid Middleware stacks
  - Many HPC applications using the Grid
    - Some (HEP, Bio) in production use
    - Some still in testing phase: more effort required to make the Grid their day-to-day workhorse
  - e-Health applications also part of the Grid
  - Some industrial applications:
    - Early deployment mainly in different EC projects

Achieving global e-Science

Session convener

Participant

Courtesy EGEE Project Office

Grid achievements beyond e-Science

- Grid beyond e-Science?
  - Slower adoption: prefer different environments, tools and have different TCOs
    - Intra grids, internal dedicated clusters , cloud computing
  - e-Business applications
    - Finance, ERP, SMEs and Banking!
    - New economic and business models
  - Industrial applications
    - Energy, Automotive, Aerospace, Pharmaceutical industry,
      Telecom
  - e-Government applications
    - Earth Observation, Civil protection:
    - e.g. The Cyclops project

Examples beyond e-Science

CitiGroup (Citigroup Inc., operating as Citi, is a major American financial services company based in New York City) adopted Grid computing

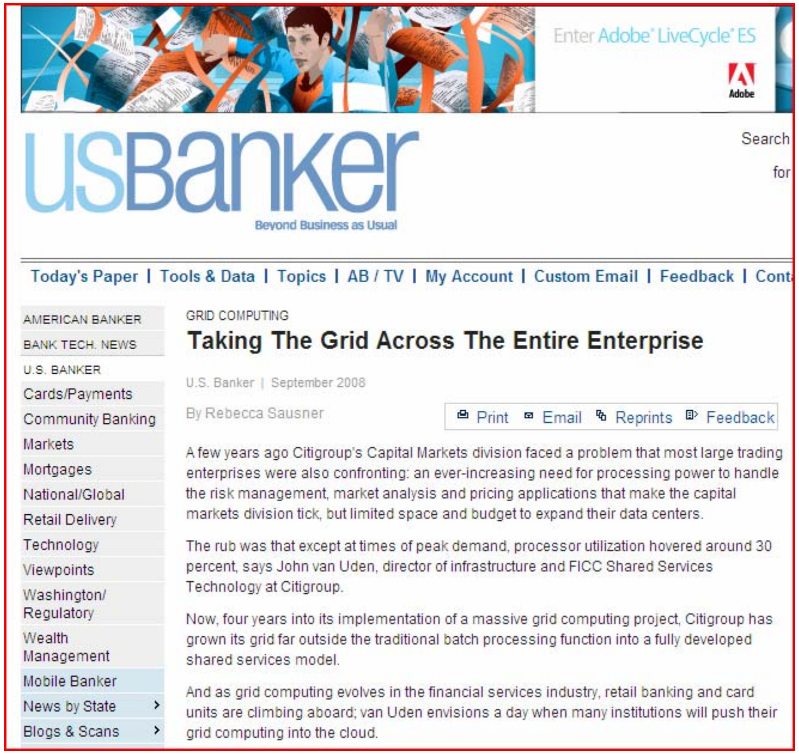
http://www.americanbanker.com/usb_article.html?id=20080825IXTFW8BS

- Citi chose **Platform Computing's Symphony grid product** to consolidate its computing assets into a single resource pool with increased utilization

- At Citi, since the grid was implemented, individual business units are charged for the processing power they use, creating **a shared services environment**

-Citi is now using near **20,000 CPUs** and there are periods of the day where the **utilization rate is 100 percent**

-Citi is planning of using the cloud in cases their data centers do not suffice (**overflow model** or cooperative data centers)



GRID COMPUTING

Taking The Grid Across The Entire Enterprise

U.S. Banker | September 2008

By Rebecca Sausner

Print Email Reprints Feedback

A few years ago Citigroup's Capital Markets division faced a problem that most large trading enterprises were also confronting: an ever-increasing need for processing power to handle the risk management, market analysis and pricing applications that make the capital markets division tick, but limited space and budget to expand their data centers.

The rub was that except at times of peak demand, processor utilization hovered around 30 percent, says John van Uden, director of infrastructure and FICC Shared Services Technology at Citigroup.

Now, four years into its implementation of a massive grid computing project, Citigroup has grown its grid far outside the traditional batch processing function into a fully developed shared services model.

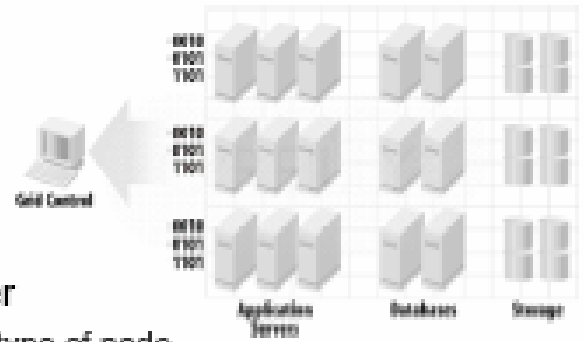
And as grid computing evolves in the financial services industry, retail banking and card units are climbing aboard; van Uden envisions a day when many institutions will push their grid computing into the cloud.

Grid achievements in industry

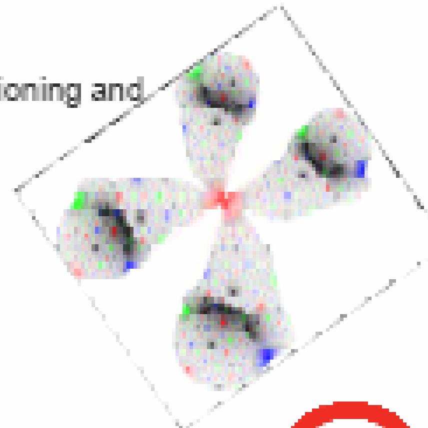
- IT Industry demonstrated interest in becoming an Grid infrastructure provider and/or user (intra-grids):
  - On-demand infrastructures:
    - Cloud and Elastic computing, pay as you go…
    - Data centers: Data getting more and more attention
  - Service hosting: outsourced integrated services
    - Software as a Service (SaaS)
    (e.g. Salesfoce.com services)
  - Virtualisation being exploited in Cloud and Elastic computing (e.g. Amazon EC2 virtual instances)
- "Pre-commercial procurement"
  - Research-industry collaboration in Europe to achiev leading-edge products
    - Example: PRACE building a PetaFlop Supercomputing Centre in Euro

The HPC view from ...the clouds!

Not All Clouds are Alike



- Flexible but complex: the Grid
  - Grids imply dynamic arrival/departure
  - Grids may include specialized nodes
- Cost-effective but confined: the Cluster
  - Clusters are typically monocultures: just one type of node
  - Applications may require tuning to a particular cluster size
- Responsive but repetitive: the Hypervisor
  - Virtualized servers can be quickly provisioned
  - Software stack within the virtual server retains issues of versioning and configuration
  - Virtual appliances demand monitoring/management
- Enterprise cloud computing implies API leverage
  - Immediate focus on function; immediate delivery of value
  - Concerns arise around perception of platform lock-in



salesforce.com

Courtesy Peter Coffee, Salesforce.com

Today and the future: Green IT, pay per CPU/GB virtualisation and/or HPC in every lab?

- Computer and data centers in energy and environmental favorable locations are becoming important
- Elastic computing, Computing on the Cloud, Data Centers and Service Hosting - Software as a Service, are becoming the new emerging solutions for HPC applications
- Many-multi-core and CPU accelerators are promising potential breakthroughs
- Green IT initiatives:
  - The Green Grid (www.thegreengrid.org) consortium (**AMD, APC, Dell, HP, IBM, Intel, Microsoft, Rackable Systems, Sun Microsystems and Vmware)**
  - IBM Project Big Green (**a $1 billion investment to dramatically increase the efficiency of IBM products)** and other IT industry initiatives try to address current HPC limits in energy and environmental impact requirements

Today and the future: Cloud computing and storage on demand

•Cloud Computing: http://en.wikipedia.org/wiki/Cloud_computing

•Amazon, IBM, Google, Microsoft, Sun, Yahoo, major 'Cloud Platform'

pote

ng and

rvice (S3)

•Sun Network.com www.network.com (1$/CPU hour, no contract cost)

•IBM Grid solutions (www.ibm.com/grid)

•GoGrid – a division of ServePath company (www.gogrid.com) Beta released (pay-as-you-go and pre-paid plans, server manageability)

EGEE cost estimation (1/2)

Capital Expenditures (CAPEX):

a. Hardware costs: 80.000 CPUs ~ in the order of 120M Euros (80-160M)
Depreciating the infrastructure in 4 years:30Meuros per year (20M to 40M)
b. Cooling and power installations (supposing existing housing facilities available)
25% of H/W costs: 30M, depreciated over 5 years: 6M Euros

Total: ~ 36M Euros / year (26M-46M)

Slide Courtesy of Fotis Karayannis

EGEE cost estimation (2/2)

Operational Expenditures (OPEX):

a. 20 MEuros per year for all EGEE costs (including site administration, operations, middleware etc.)

b. Electricity ~10% of h/w costs: 12M Euros per year (other calculations lead to similar results)

c. Internet connectivity: Supposing no connectivity costs (existing over-provisioned NREN connectivity)

***If other model is used (to construct the service from scratch), then network costs should be taken into account**

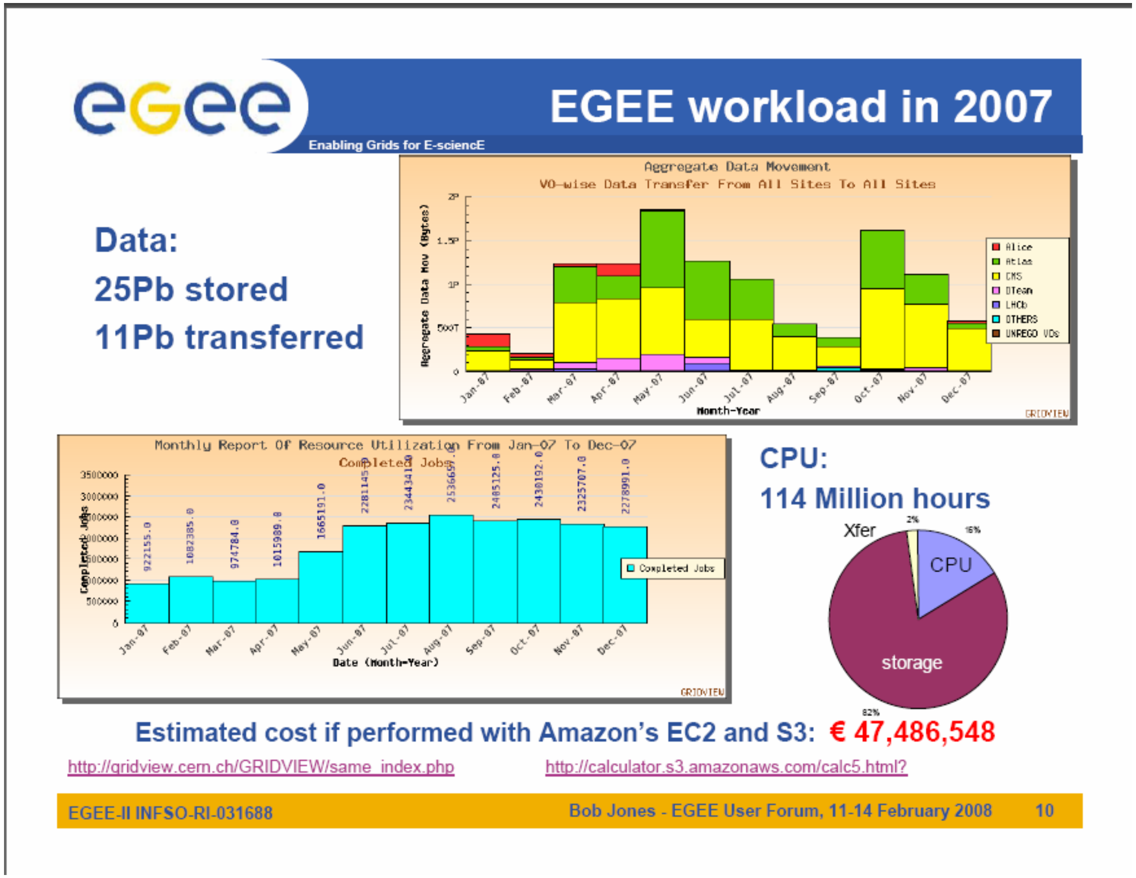
Total 32M / year

CAPEX+OPEX= 68M per year (58-78M)

Slide Courtesy of Fotis Karayannis

EGEE if performed with Amazon EC2 and S3

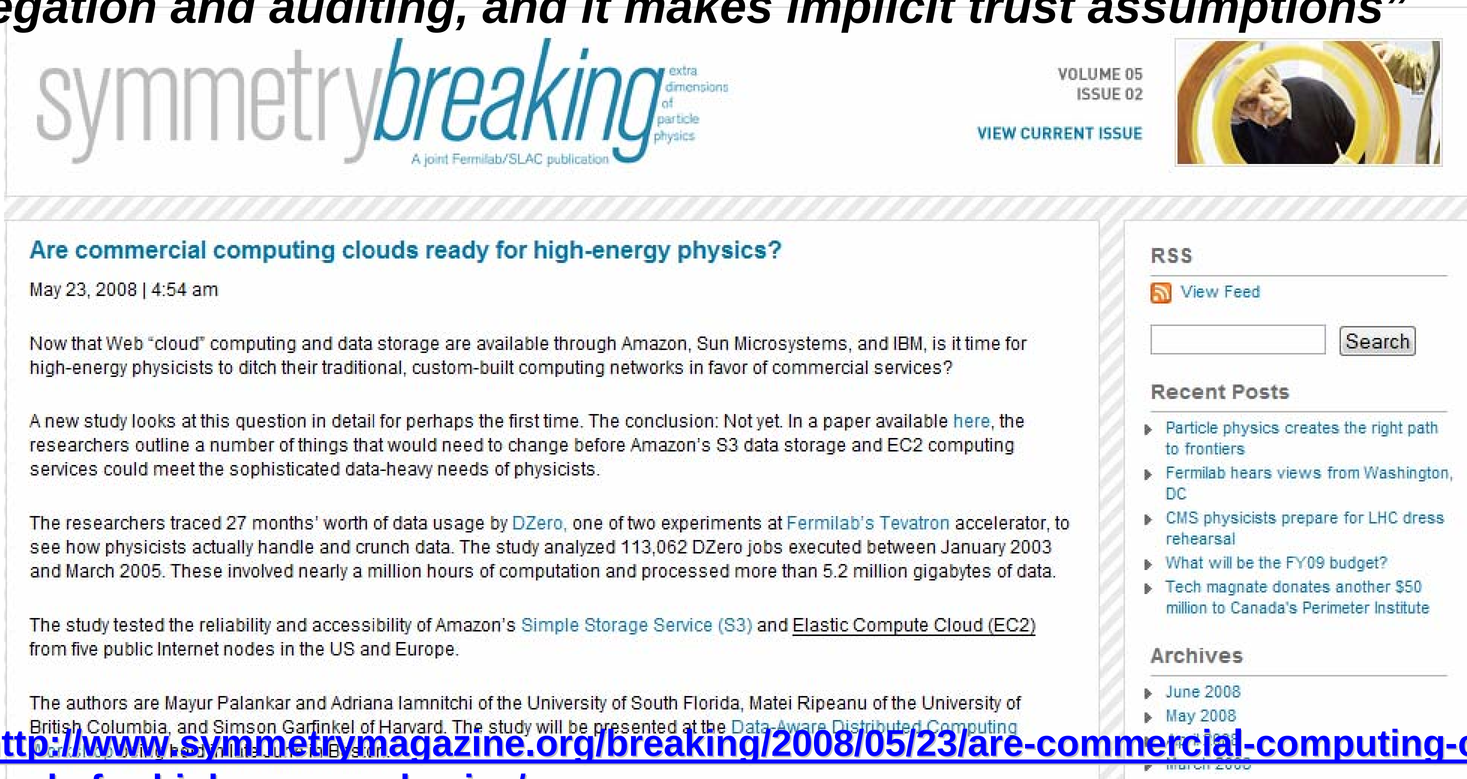
In the order of ~50M Euros, probably more cost effective of EGEE actual cost, depending on the promotion of the EC2/S3 service



Slide Courtesy of Bob Jones

Cloud mature enough for big sciences?

Probably not yet, as not designed for them; Does not support complex Scenarios: ***"S3 lacks in terms of flexible access control and support for delegation and auditing, and it makes implicit trust assumptions"***


symmetrybreaking extra dimensions of particle physics

A joint Fermilab/SLAC publication

VOLUME 05
ISSUE 02
VIEW CURRENT ISSUE

Are commercial computing clouds ready for high-energy physics?

May 23, 2008 | 4:54 am

Now that Web "cloud" computing and data storage are available through Amazon, Sun Microsystems, and IBM, is it time for high-energy physicists to ditch their traditional, custom-built computing networks in favor of commercial services?

A new study looks at this question in detail for perhaps the first time. The conclusion: Not yet. In a paper available here, the researchers outline a number of things that would need to change before Amazon's S3 data storage and EC2 computing services could meet the sophisticated data-heavy needs of physicists.

The researchers traced 27 months' worth of data usage by DZero, one of two experiments at Fermilab's Tevatron accelerator, to see how physicists actually handle and crunch data. The study analyzed 113,062 DZero jobs executed between January 2003 and March 2005. These involved nearly a million hours of computation and processed more than 5.2 million gigabytes of data.

The study tested the reliability and accessibility of Amazon's Simple Storage Service (S3) and Elastic Compute Cloud (EC2) from five public Internet nodes in the US and Europe.

The authors are Mayur Palankar and Adriana Iamnitchi of the University of South Florida, Matei Ripeanu of the University of British Columbia, and Simson Garfinkel of Harvard. The study will be presented at the Data-Aware Distributed Computing ...

RSS

View Feed

Search

Recent Posts

- Particle physics creates the right path to frontiers
- Fermilab hears views from Washington, DC
- CMS physicists prepare for LHC dress rehearsal
- What will be the FY09 budget?
- Tech magnate donates another $50 million to Canada's Perimeter Institute

Archives

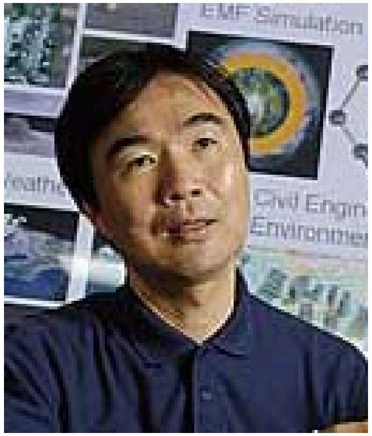
- June 2008
- May 2008
- March 2008

http://www.symmetrymagazine.org/breaking/2008/05/23/are-commercial-computing-clouds-ready-for-high-energy-physics/

http://www.csee.usf.edu/~anda/papers/dadc108-palankar.pdf

The future:

"To Distribute or **Not To Distribute**"



- Prof. Satoshi Matsuoka, TITech
- Keynote at Mardi Gras Conference Baton Rouge, 31 Jan 2008

- In the late 90s, petaflops were considered very hard and at least 20 years off …
- while grids were supposed to happen right way
- After 10 years (around now) petaflos are "real close" but there's still no "global grid"
- What happened:

→ ***It was easier to put together massive clusters than to get people to agree about how to share their resources***

→ ***For tightly coupled HPC applications, tightly coupled machines are still necessary***

→ ***Grids are inherently suited for loosely coupled apps or enabling access to machines and/or data***

- **With Gilder's Law*, bandwidth to the compute resources will promote thin client approach**

*** "***Bandwidth grows at least three times faster than computer power." This means that if computer power doubles every eighteen months (per Moore's Law), then communications power doubles every six months*

- **Example: *Tsubame* machine in Tokyo**

Multi-core architectures

- Computer CPUs have adopted multi-core architectures with increasing number of cores
  - 2-4 cores in PCs and laptops
  - 8-32 cores in servers, 64-80 cores under development
  - Intel announced a 6 core Xeon

- The trend is driven by many factors:
  - Power consumption, heat dissipation, energy cost, availability of high bandwidth computing at lower cost, ecological impact
- The entire software ecosystem will need to adapt including related applications

Conclusion (1/2)

- We are at a flex point in the evolution of distributed computing
  - **nothing new under the sun…!**
  - **Grid** has delivered an affordable HPC infrastructure
  - to scientists all over the world to solve intense computing and storage problems within constrained research budget (and often for social/political reasons) Grid computing
  - leveraging international research networks to deliver an effective and irreplaceable channel for international collaboration
  - This has also been effectively used by **industry**
  - to increase the usage of their HPC infrastructure and reduce Total Cost of Ownership (TCO)
  - Major issues with wide adoption of Grids have to do with:
  - Cost of operations, complexity, not a solution for all problems (latency, fine grain parallelism are difficult), reliability, security..

Conclusion (2/2)

- **Cloud computing** and **hosted services** are emerging as the next incarnation of distributed computing with some obvious additional advantages but not really designed with scientific applications in mind
- Many changes are happening in the basic underpinning technology (**parallel everywhere**!)
- New boundary constraints and very much energy are becoming the limiting factor to the otherwise still valid Moore's Law…
- If we will be able to harness the potential enormous power of parallel computing (not so good story so far) then we might be able to provide better computing solutions for research in energy and eventually better energy solutions for our computing needs!

Thanks

Thanks to the organizers for the kind invitation and to all of you for your attention

Contact me at:

Fabrig @ microsoft.com

Cracow '08 Grid Workshop