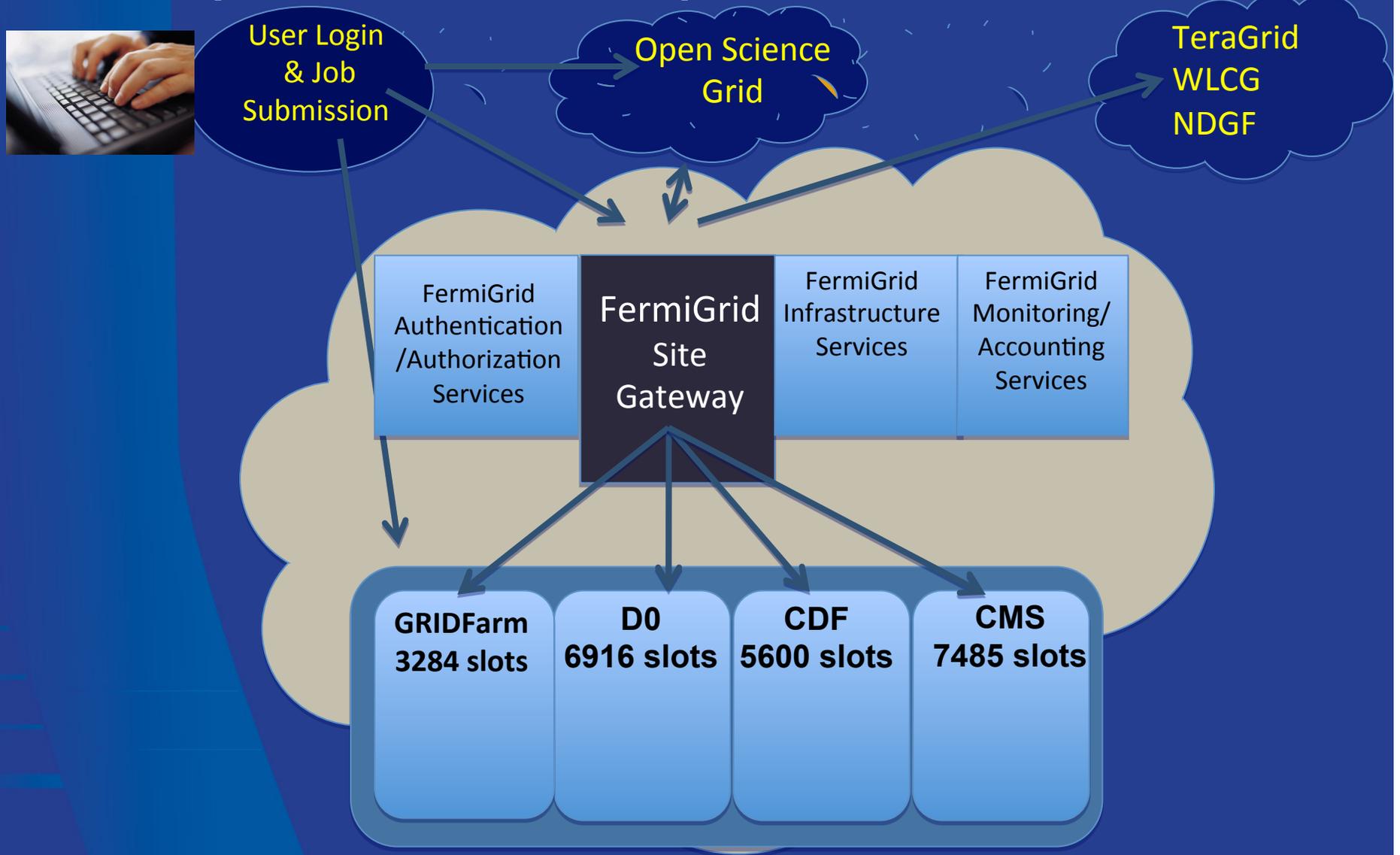




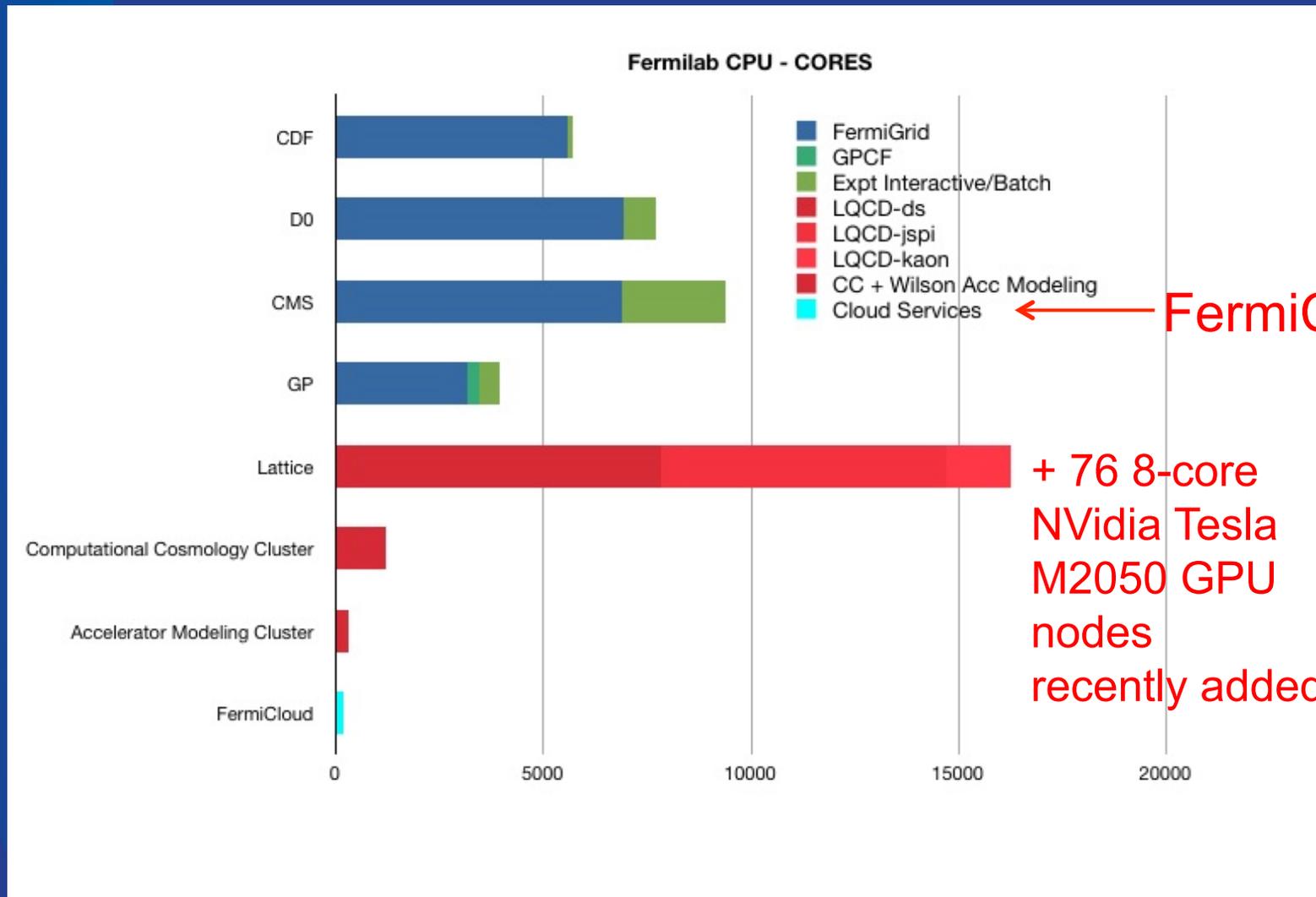
Grids 10 years later: now what? (One view from the US)

VICTORIA WHITE (VICKY)
ASSOCIATE DIRECTOR FOR COMPUTING SCIENCE AND TECHNOLOGY AT FERMILAB
AND CHIEF INFORMATION OFFICER

Campus Grid example – FermiGrid



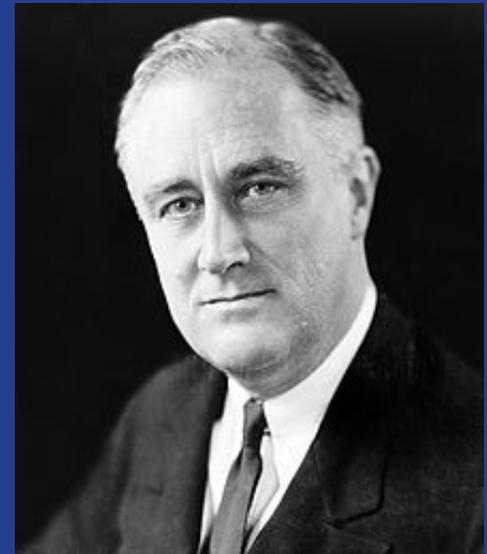
Fermilab core count – HPC & HTC



My speech: a brief personal perspective

Be sincere; be brief; be seated.

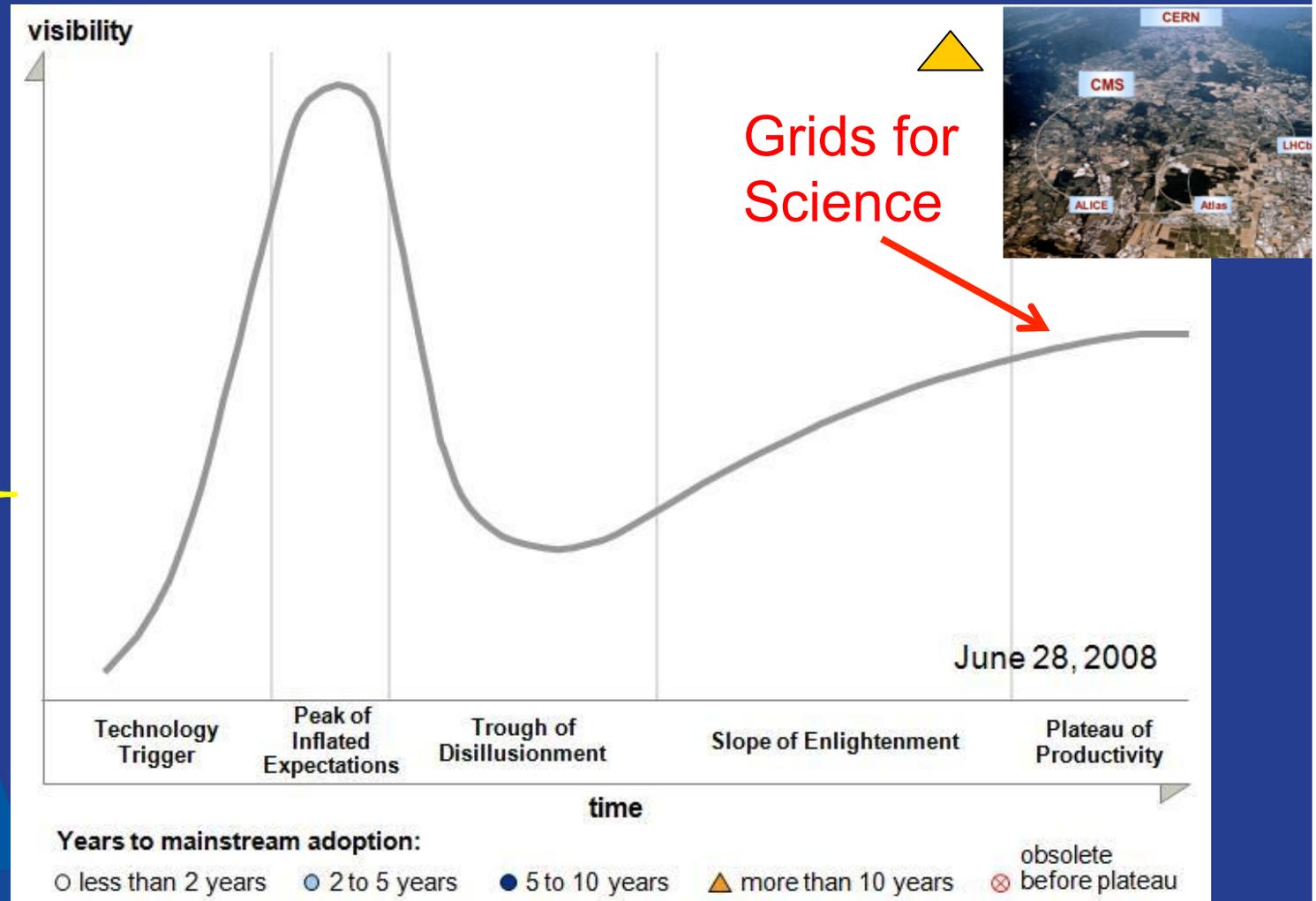
- Franklin D. Roosevelt, on speechmaking



From Grid and funding hype to useful sustainable Grids for science

Big Science

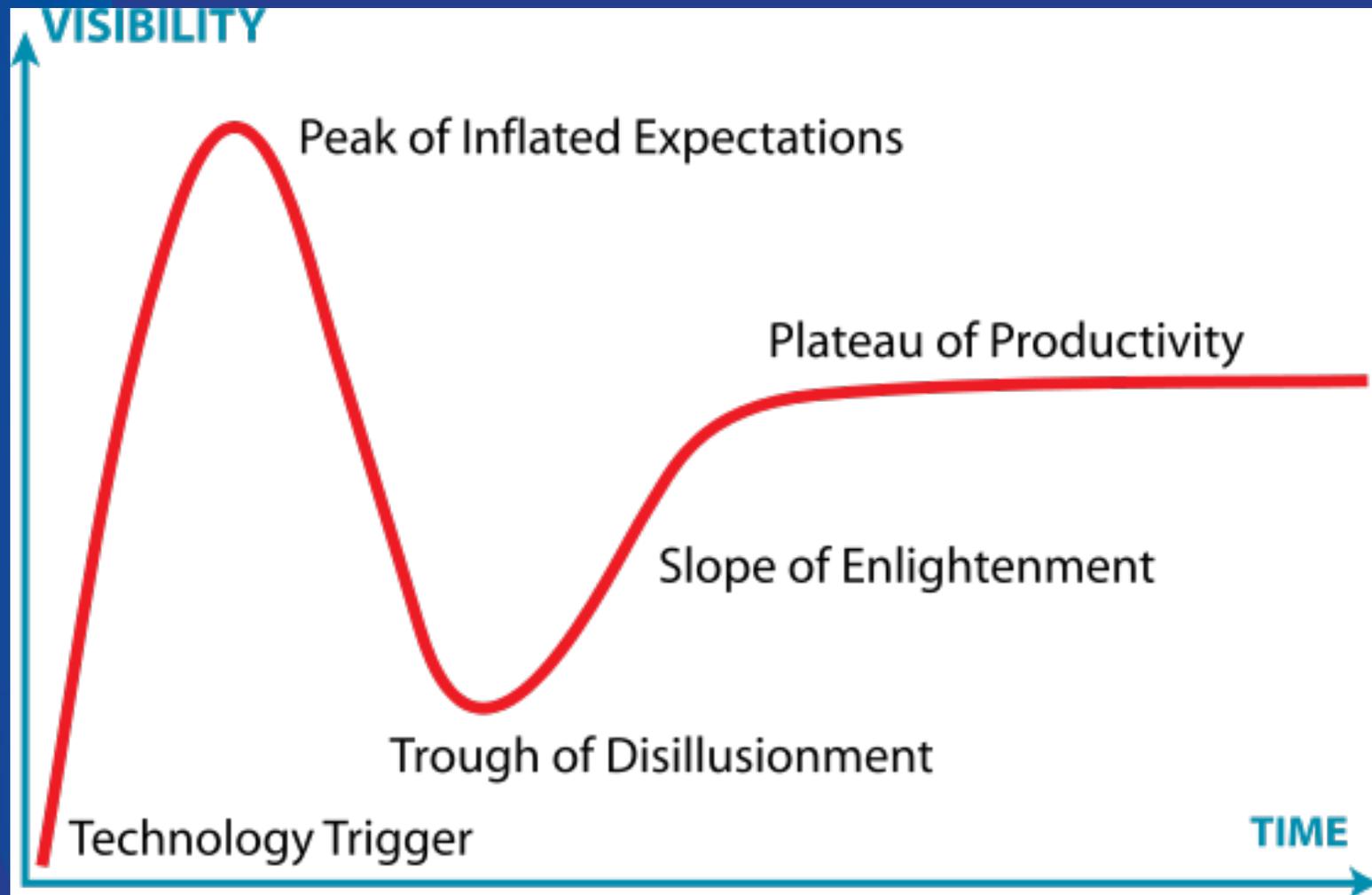
Huge datasets
Grid: Blueprint for a New Computing
Costly computing and storage
Foster, Kesselman
Excellent networks



Some Definitions

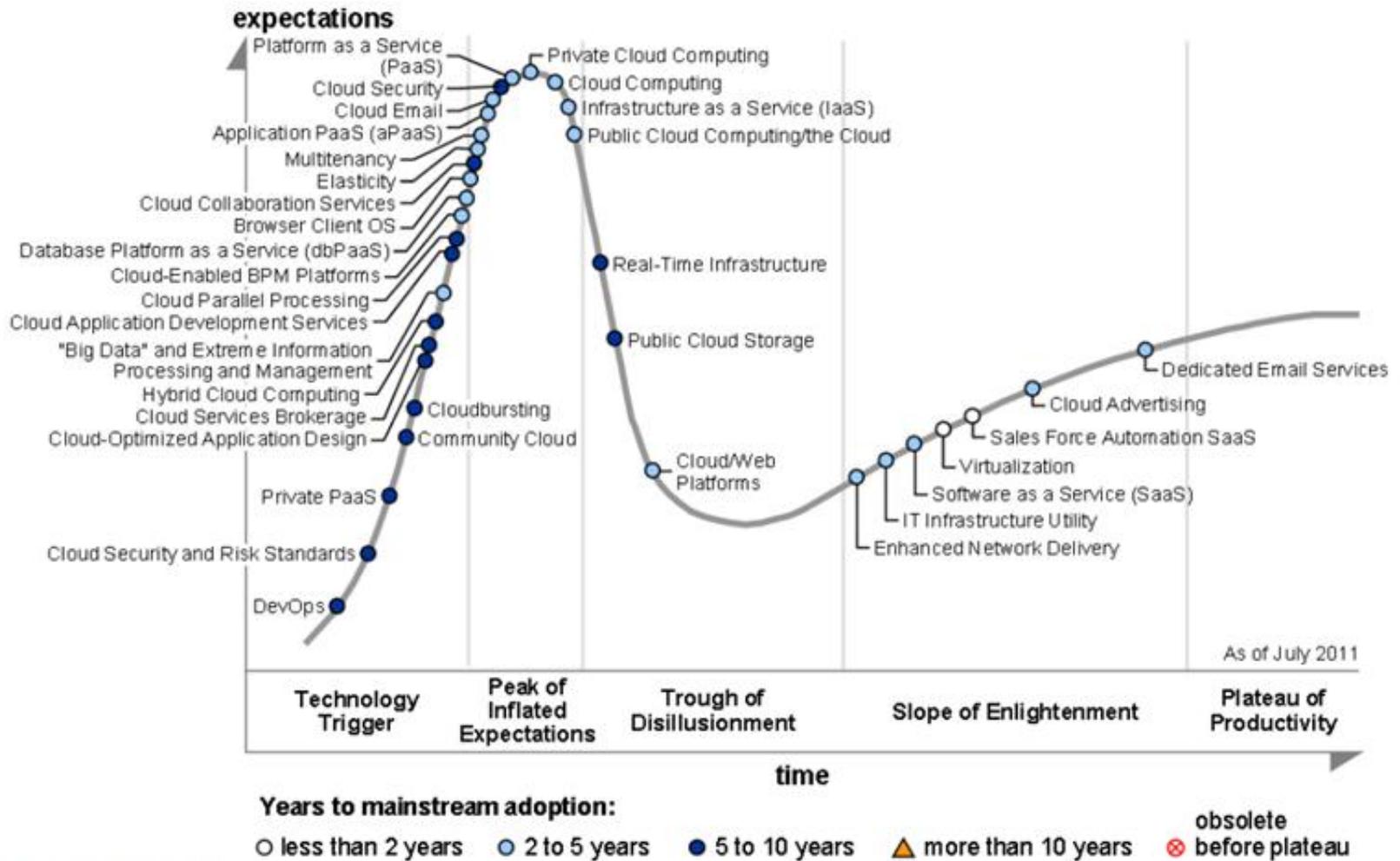
- Hype
 - exaggerated publicity; hoopla.
 - an ingenious or questionable claim, method, etc., used in advertising, promotion, or publicity to intensify the effect.
 - swindle, deception, or trick.
- Gartner
 - Gartner is an information technology research and advisory company providing technology related insight. Research provided by Gartner is targeted at CIOs and senior IT leaders in industries that include government agencies, high-tech and telecom enterprises, professional services firms, and technology investors. Gartner clients include large corporations, government agencies, technology companies and the investment community.

Gartner's basic Hype Cycle



And Cloud ?

Figure 1. Hype Cycle for Cloud Computing, 2011



Source: Gartner (July 2011)

Open Science Grid (OSG)



Open Science Grid

Jobs CPU Hours Transfers TB Transferred Status Map

Millions of Jobs/Month



24 Hours 30 Days 12 Month

OSG delivered across 103 sites

In the last 24 Hours

774,000	Jobs
1,618,000	CPU Hours
1,079,000	Transfers
1,006	TB Transferred

In the last 30 Days

16,559,000	Jobs
45,672,000	CPU Hours
43,084,000	Transfers
26,716	TB Transferred

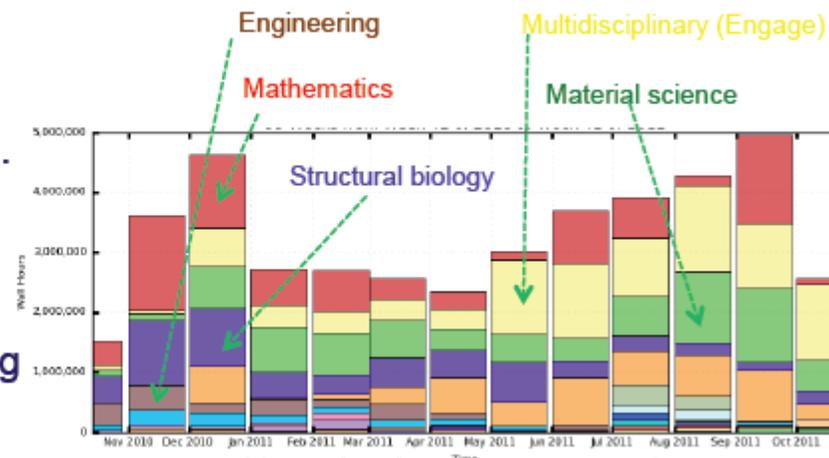
In the last Year

196,350,000	Jobs
497,208,000	CPU Hours
557,576,000	Transfers
279,836	TB Transferred

Each finished job on an OSG resource is reported to the central accounting system

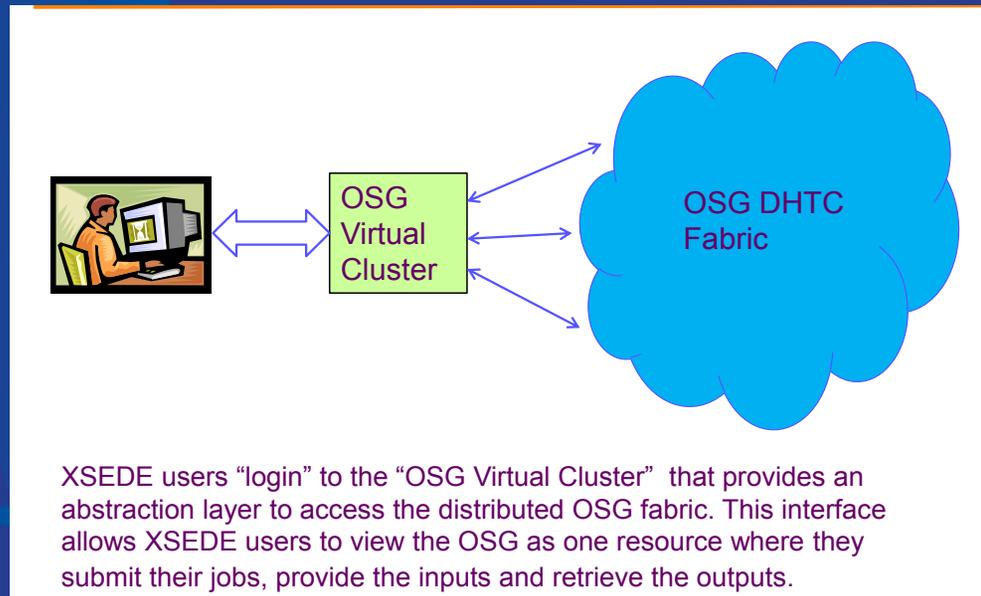


8% CPU cycles used for Non-Physics.
 25% CPU cycles used for Non-LHC
 Physics – Tevatron, LIGO,
 Non-LHC usage (25% + 8%) on
 resources user does not own – making
 more effective use of available capacity.



Non-physics use over past year

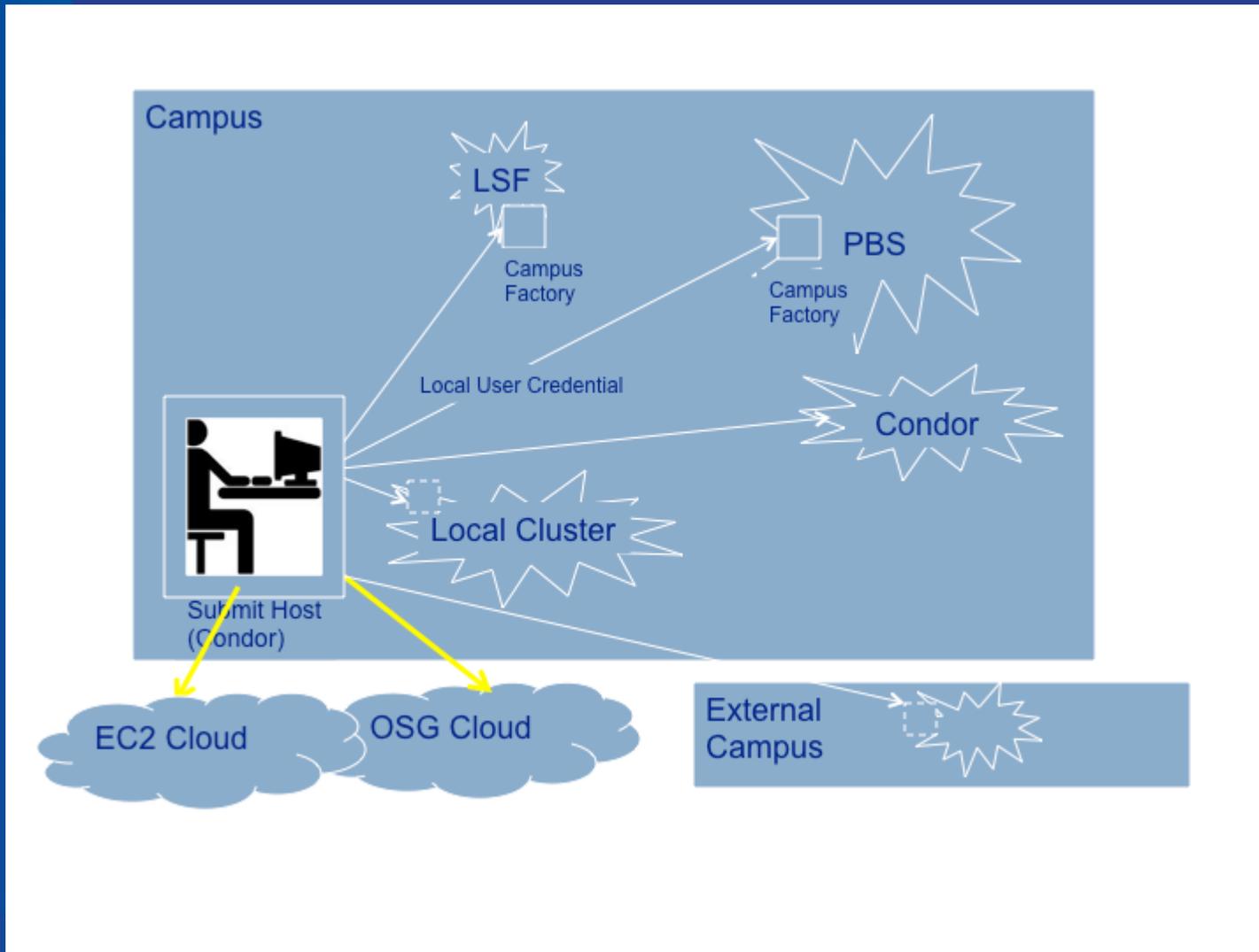
Supercomputing (HPC) meets Distributed Computing (HTC)



Teragrid => XSEDE

EXtreme **S**cience
And **EN**gineering
Discovery
Environment

Scientist at desk meets Campus Grid



Community not Technology



Open Science Grid in 2015



Sustainability (popular in the US)



But **sustaining software** used by multiple communities (such as Globus, Condor, GridFTP, Gratia, SRM, etc) is a problem and makes the software a victim of its own success

- Funding agencies want to fund innovation
- No economic model for fair-sharing the cost among many communities
- It's an issue we have to address globally

Grids => Clouds? or Grids + Clouds ?

Let's not throw out the baby
with the bathwater

Evolution not revolution

Build on successes

- The collaborations
- The sustainable software



Are Grids transparent, easy to use, simple to authenticate to, effortless to move data?

The answer is “not yet”

We would like it to be as easy and amazing as Google Earth



and as intuitive as social media

- Can't make it so by money investments alone
- But only by collaborating and evolving together
- **Better, more interoperable and easier to use security infrastructure is essential**

Lets keep building on the Grid framework

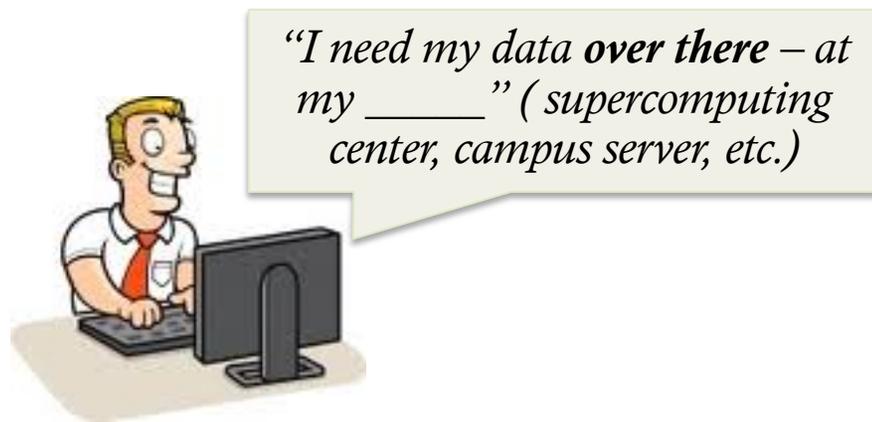
At this conference we are going to hear about a lot of building that is going on

- some of it takes years to get enough people to adopt to make it standard
- some of it is evolutionary
- some of it is revolutionary
- all of it is worthwhile

The challenge: Moving big data *easily*



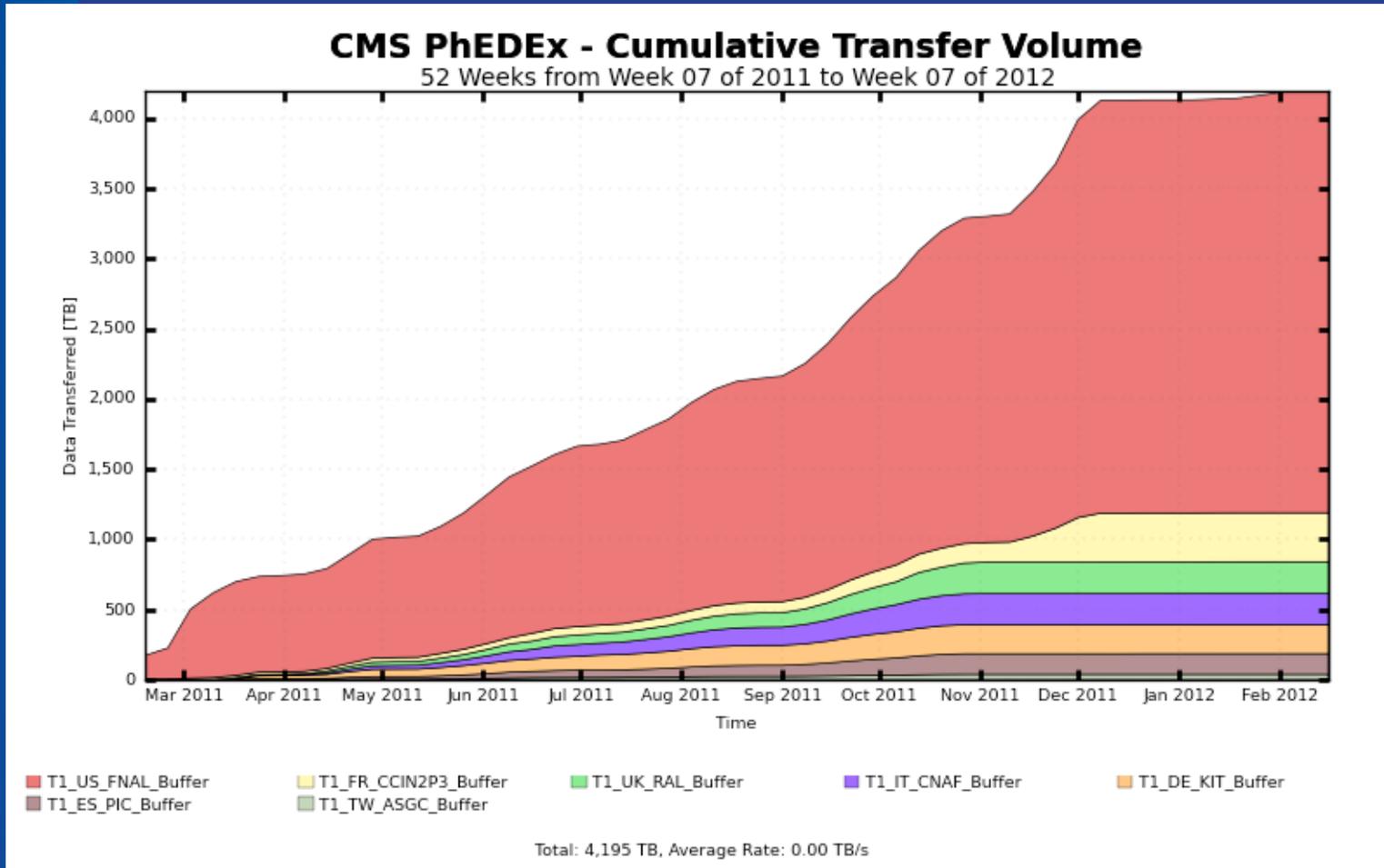
What should be trivial ...



... can be painfully tedious and time-consuming



LHC experiments do move data efficiently – e.g CMS PhedEx



2011: Transfers from CERN to T1 sites, total 4.195 TiB



globus online

Reliable File Transfer. No IT Required.

Sign Up

Sign In

Reliable, high-performance, secure file transfer.
Move files fast. No IT required.

+ WATCH A VIDEO

Globus Online in a nutshell



> GET STARTED

Sign up and get moving

2,160,719,309 MB
TRANSFERRED



Why Use Globus Online?

See how easy file transfer can be



For HPC Resource Owners

Enable Globus Online for your users



For Developers

Integrate with Globus Online

Any data, Any time, Any where (AAA)

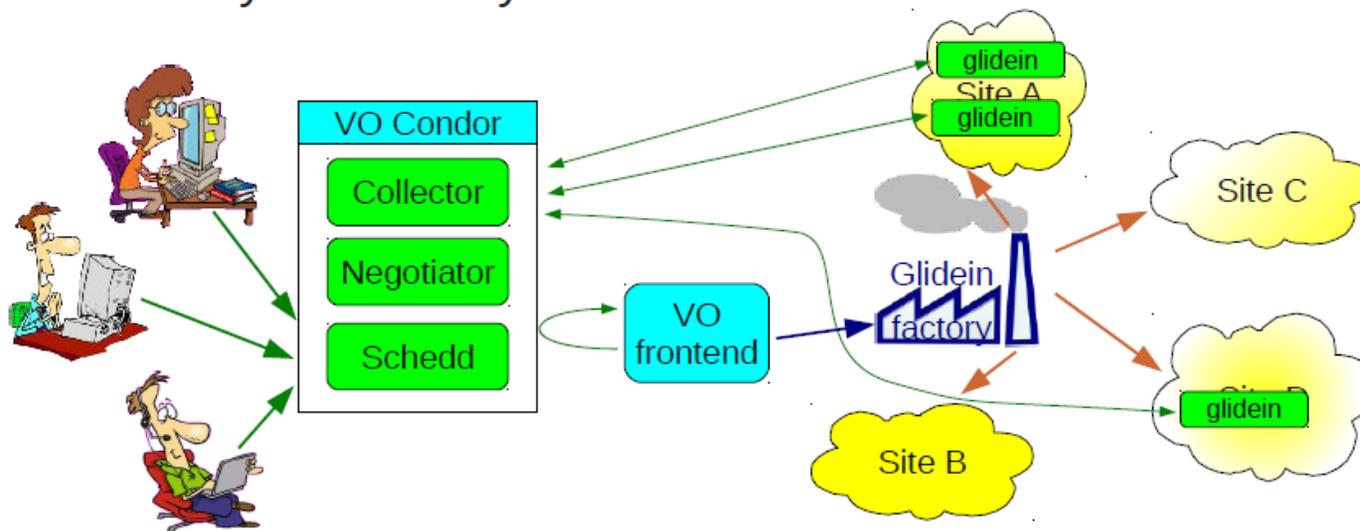
- (CMS driving)

- A network of redirectors to route data access requests from anywhere to the 'best' location with the requested data.
- Extensions to the Condor system to support co-scheduling of processing and data access resources.
- Improvements in the IO client libraries on the application side.
- Integration of the new technologies into the WLCG and OSG software stacks.
- **Deployment and operation of the new capabilities jointly by OSG and WLCG, becoming part of the global fabric of DHTC services that the two organizations jointly operate.**
- Support the use and adoption of the “Any Data, Anytime, Anywhere” technologies and service by the entire range of scientific communities served by the OSG.

Glidein workload management system (SaaS)

glideinWMS gets a step further

- Separates glidein submission logic from actual Grid submission of glideins
 - Only the factory sees the Grid



Building on existing structures
and innovating can go together



There's room for
both collaboration
and competition

- Grids and Clouds – we need both
- Grids for collaboration, sharing, communities, diversity
- Cloud computing for economies of scale, homogeneity
 - +to handle peaks



A 21st C science IT infrastructure strategy

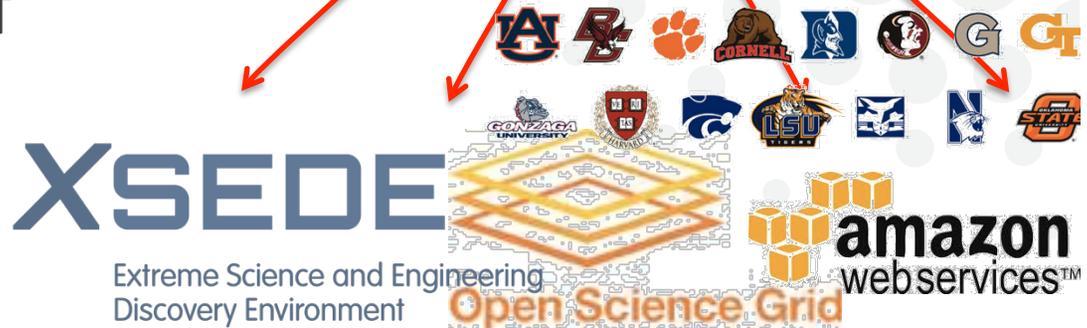


- To provide **more** capability for **more** people at **less** cost ...
- Create **infrastructure**
 - Robust and universal
 - Economies of scale
 - Positive returns to scale
- Via the creative use of
 - **Aggregation** (“cloud”)
 - **Federation** (“grid”)

Small and medium laboratories and projects

L L L L L L L L L L
L L P L P L L P L P L L P L L
L L L L L L L L L L

Research data management a
Collaboration, computation a
Research administration S



Success on a Worldwide scale

- If we can bring together people from all over the world (whether they be physicists, biologists, computer scientists, climate researchers or ...) and they
 - Want to be part of building the "cyber infrastructure" or Grid environments or "e-science environments" for the future
 - Actively participate
 - Get benefit from the collaboration

Then we will be succeeding

From my keynote speech at this conference in 2005

Conclusions

- We are succeeding
- We will continue to succeed if we work together and combine innovative ideas into solid working environments
 - That are still *far* from transparent and easy to use
- Sharing continues to be cost effective and opportunistic use of “other people’s resources” is a reality and a huge boost to some sciences and scientists
- Sustaining successes and software is a challenge
- Cloud computing is still maturing – but can be an effective and cost-effective component of the environment for supporting science



Thank you for listening

white@fnal.gov

QUESTIONS?