



---

# DOE-HEP Networking & Grids

Craig E. Tull

*DOE Office of High Energy Physics*

May 27, 2005

International ICFA Workshop

Daegu, Korea

Craig.Tull@science.doe.gov  
ICFA Workshop; Daegu, Korea



# DOE's Office of Science

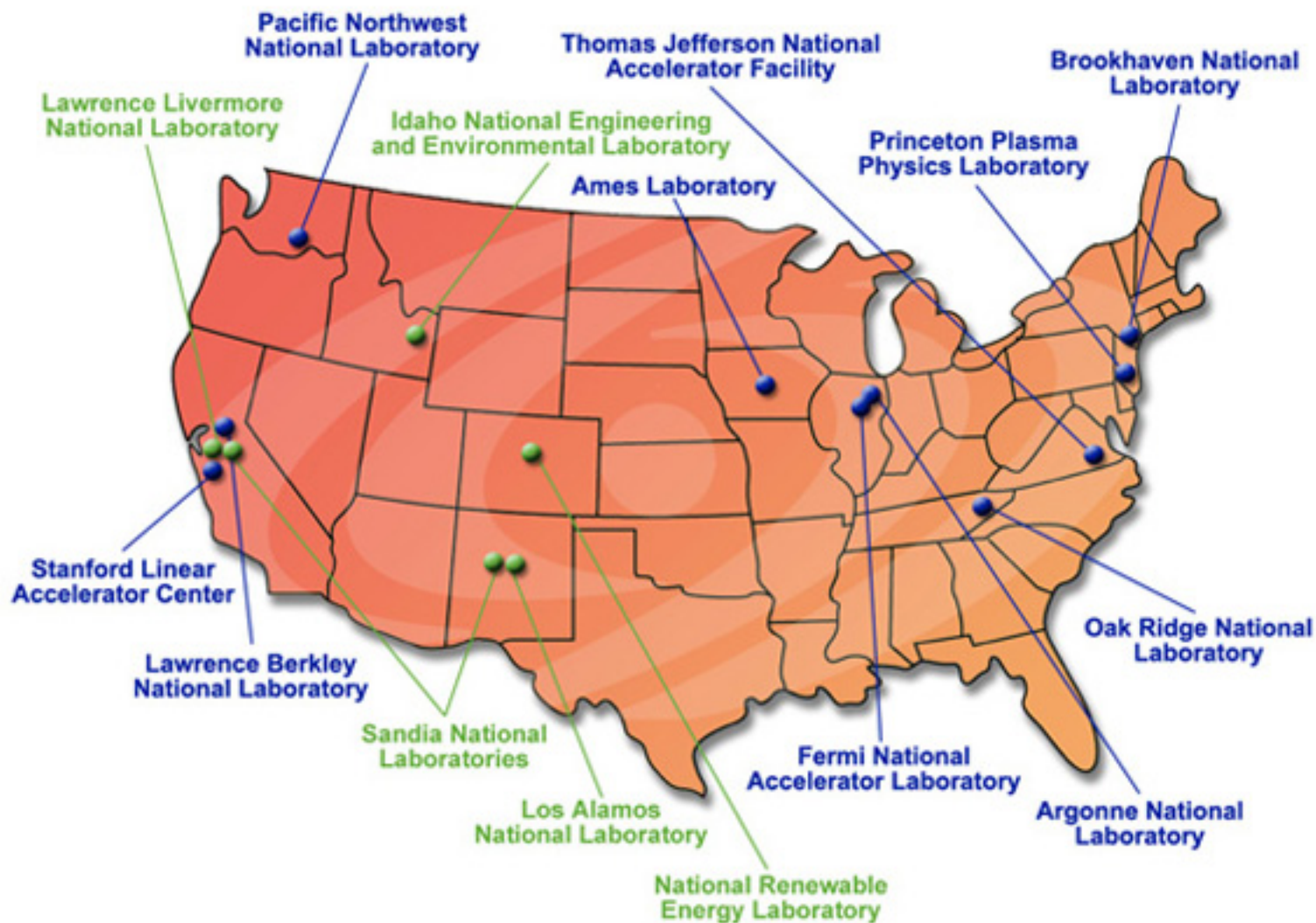


- "The Office of Science is the single largest supporter of basic research in the physical sciences in the United States, providing more than 40 percent of total funding"
- Advanced Scientific Computing Research
- Basic Energy Sciences
- Biological and Environmental Research
- Fusion Energy Sciences
- **High Energy Physics**
  - ATLAS, BaBar, CDF, CMS, D0
  - ILC, JDEM, neutrinos, ...
- Nuclear Physics

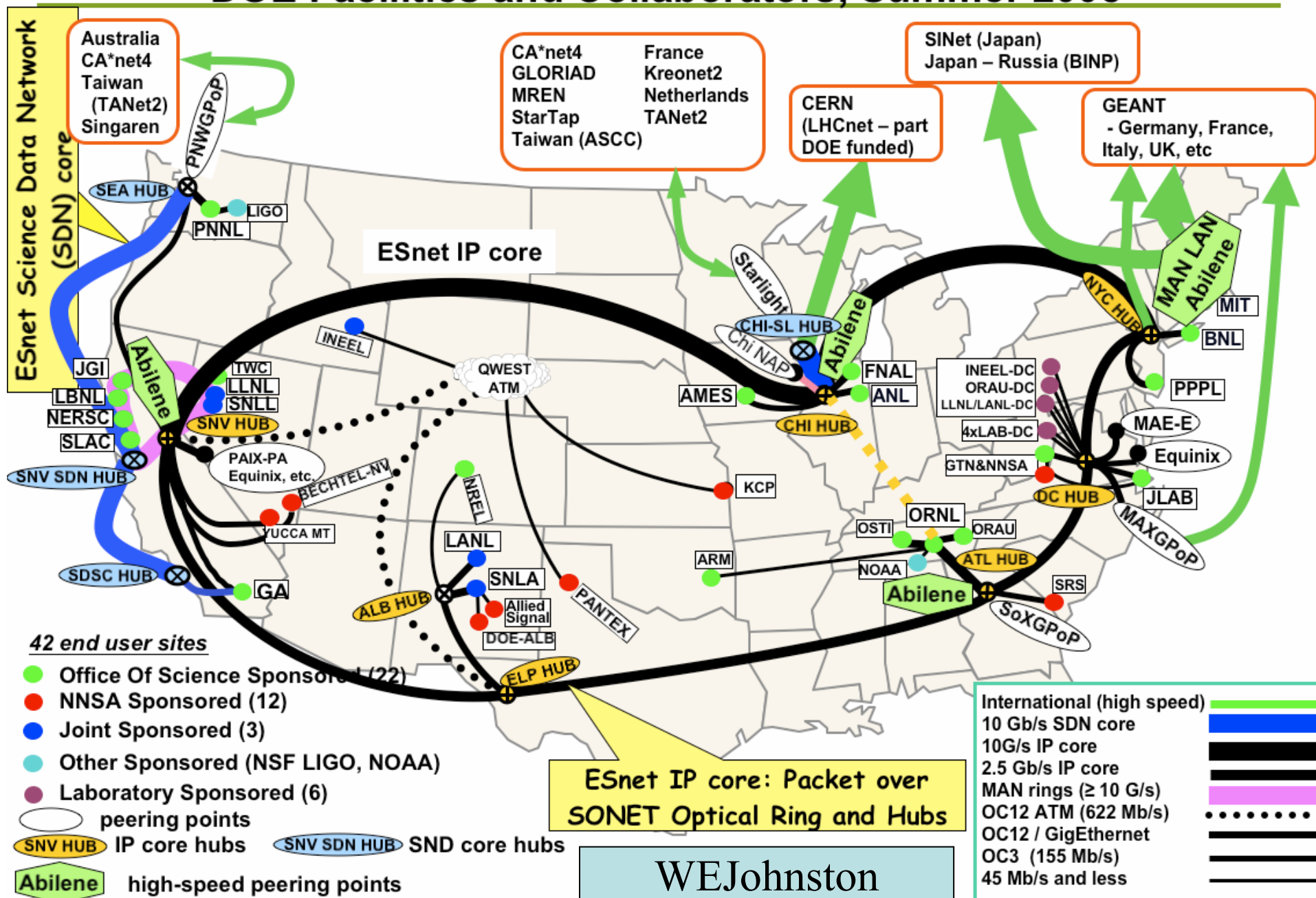
Craig.Tull@science.doe.gov  
ICFA Workshop; Daegu, Korea



# DOE Institutions



# ESnet High-Speed Physical Connectivity to DOE Facilities and Collaborators, Summer 2005





# A New ESnet Architecture

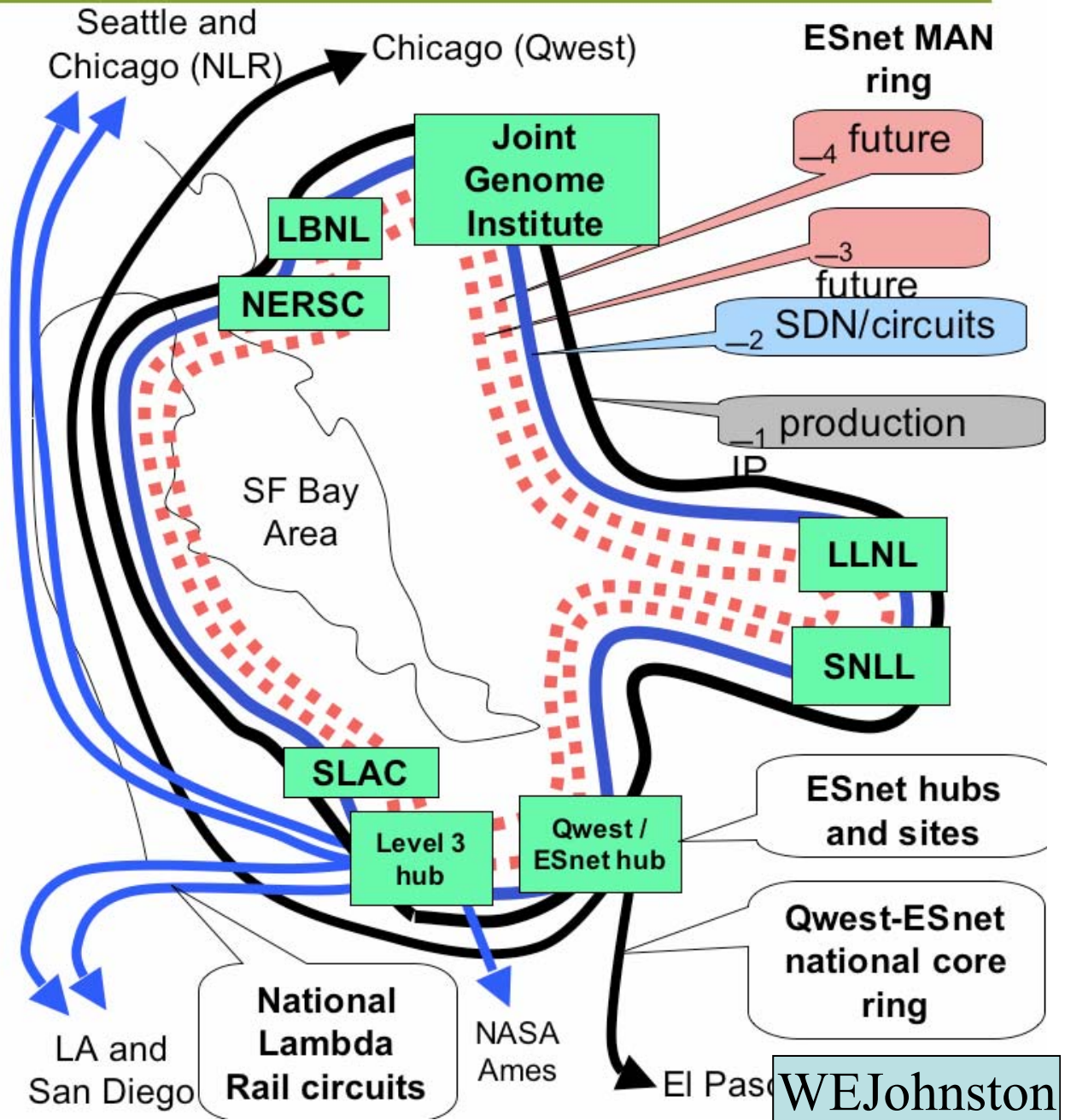
---

- In response to SC requirements, ESnet management has proposed a new architecture.
- Goals
  - Full redundant connectivity for every site
  - High-speed access for every site (at least 10 Gb/s)
- Three part strategy
  - Metropolitan Area Network (MAN) rings provide dual site connectivity and much higher site-to-core bandwidth
  - A Science Data Network (SDN) which provides very high national bandwidth with moderately high availability for affordability.
  - A highly reliable IP core (like the current ESnet core) with good bandwidth, but very high availability (~ 99.999%)

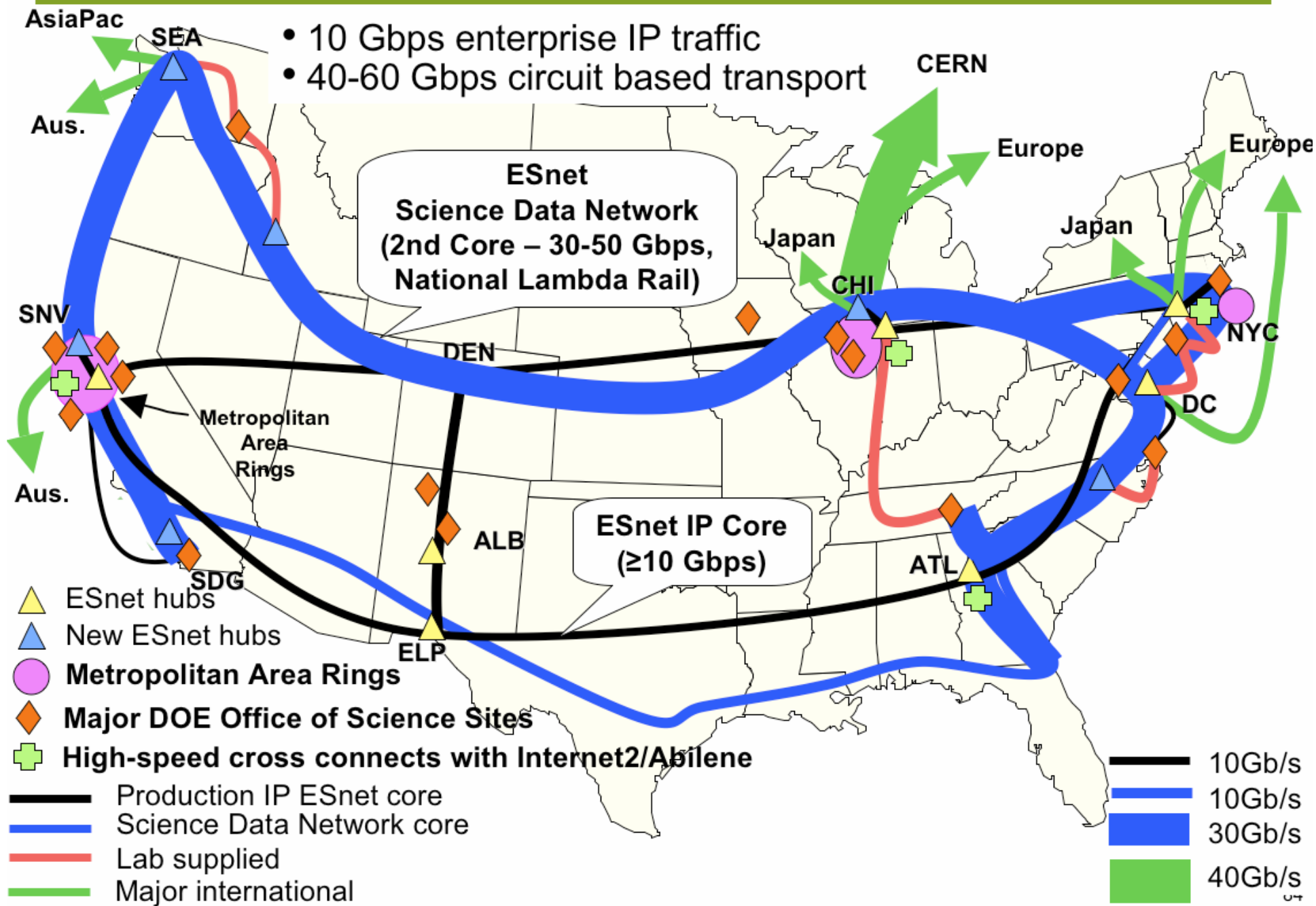


# The First ESnet MAN: SF Bay Area (Sept., 2005)

- 2 λs (2 X 10 Gb/s channels) in a ring configuration, and delivered as 10 GigEther circuits
- Dual site connection (independent “east” and “west” connections) to each site
- Will be used as a 10 Gb/s production IP ring and 2 X 10 Gb/s paths (for circuit services) to each site
- Qwest contract signed for two lambdas 2/2005 with options on two more
- One link every month - completion date is 9/2005



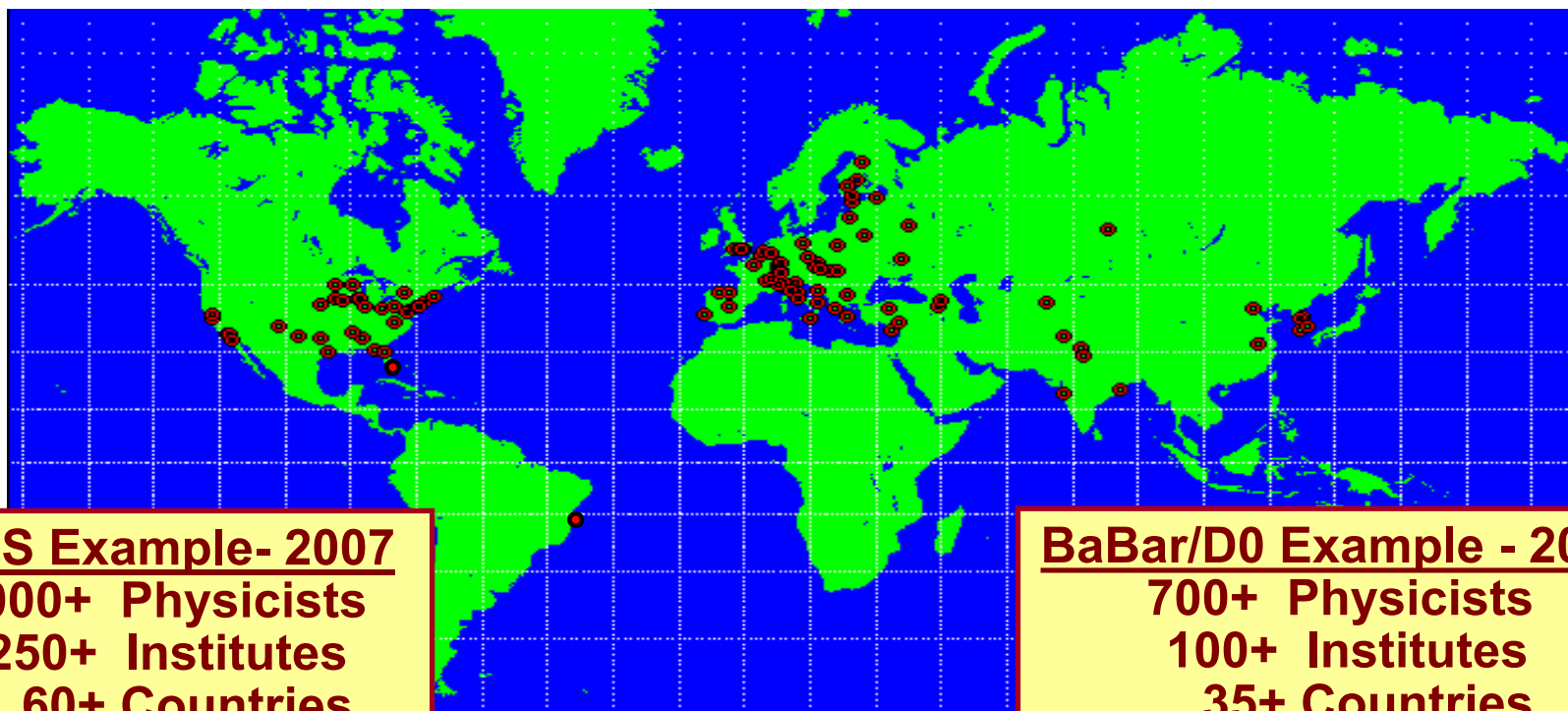
# ESnet Goal – 2007/2008





# LHC: Petascale Global Science

- Complexity: Millions of individual detector channels
- Scale: PetaOps (CPU), 100s of Petabytes (Data)
- Distribution: Global distribution of people & resources



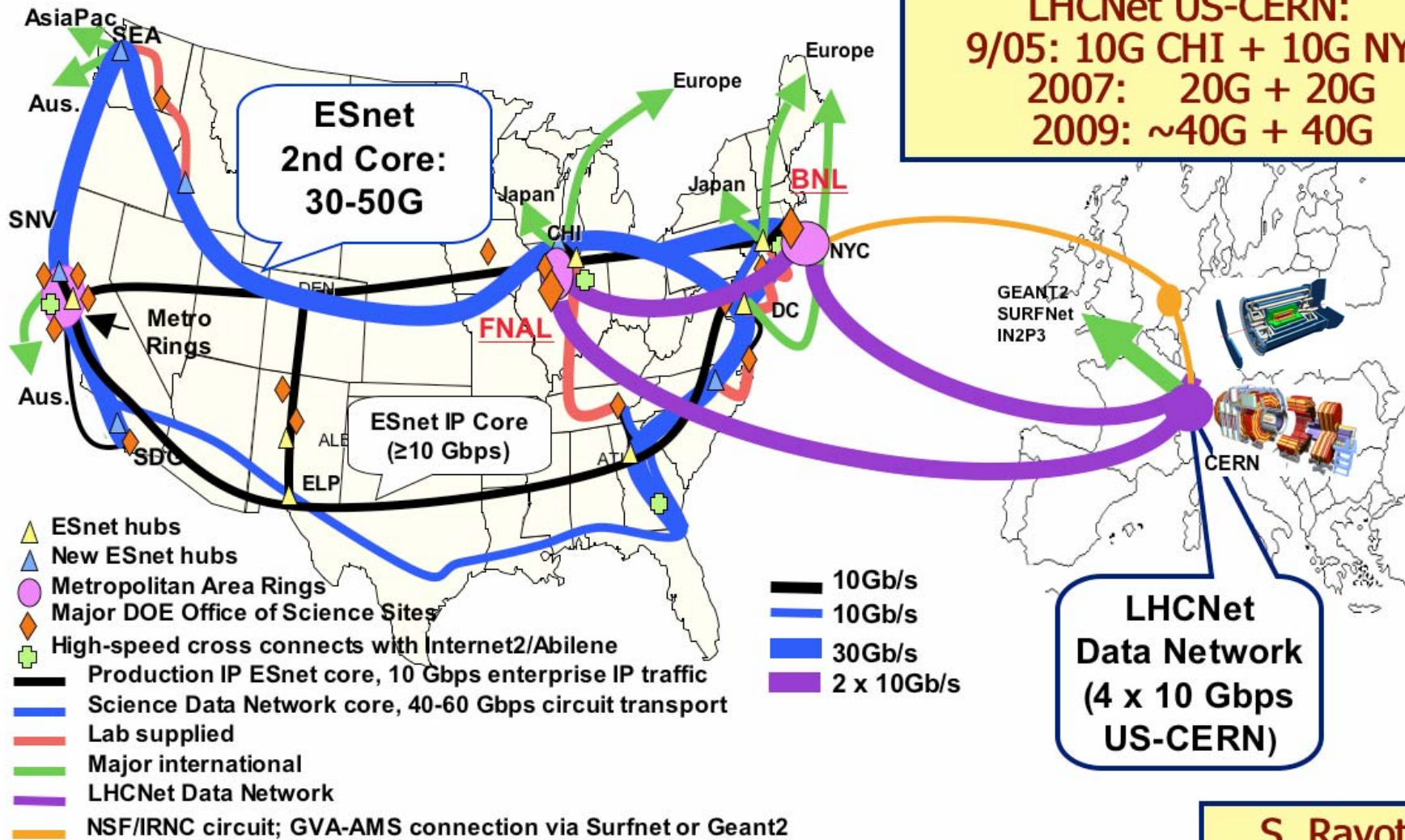
**CMS Example- 2007**  
5000+ Physicists  
250+ Institutes  
60+ Countries

**BaBar/D0 Example - 2004**  
700+ Physicists  
100+ Institutes  
35+ Countries



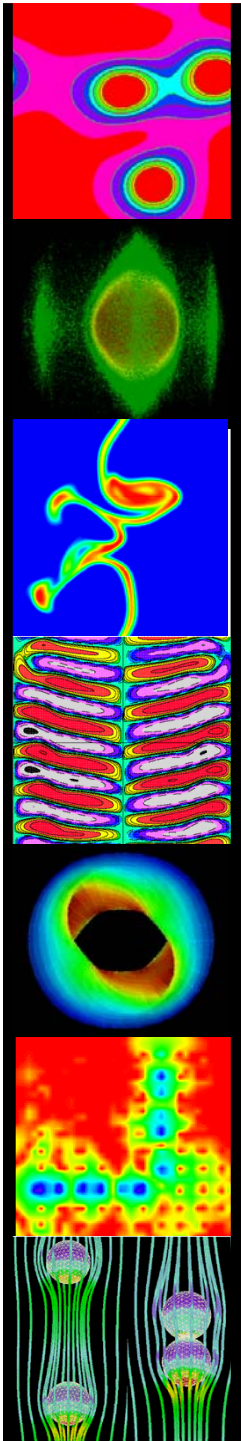


# LHCNet, ESnet Plan 2007/2008: 40Gbps US-CERN, ESnet MANs, IRNC



Craig.Tull@science.doe.gov  
ICFA Workshop; Daegu, Korea

# Scientific Discovery through Advanced Computing



- SciDAC is a \$60+M/yr pilot program for a “new way of doing science”
- first Federal program to support and enable “CSE” and (terascale) computational modeling and simulation as the third pillar of science (relevant to the DOE mission) along with theory and experiment
- spans the entire Office of Science (ASCR, BES, BER, FES, HEP, NP)
- involves all DOE labs and many universities
- builds on 50 years of DOE leadership in computation and mathematical software (EISPACK, LINPACK, LAPACK, ScaLAPACK, etc.)

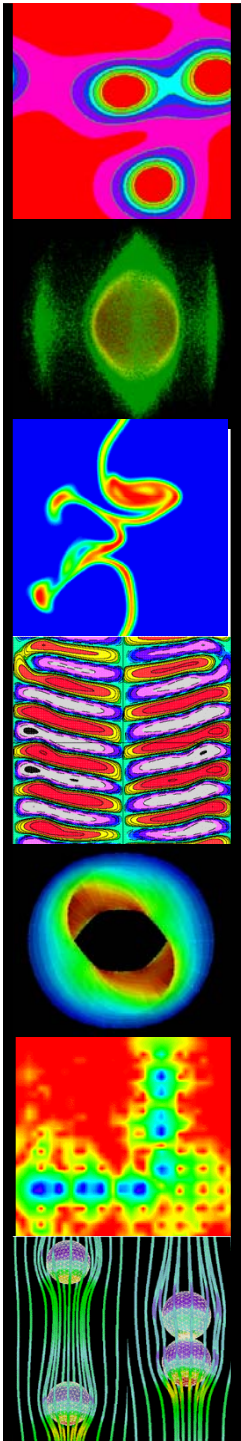
Craig.Tull@science.doe.gov  
ICFA Workshop; Daegu, Korea

# SciDAC Goals

- an INTEGRATED program to:
  - (1) create a new generation of scientific simulation codes that takes full advantage of the extraordinary capabilities of terascale computers
  - (2) create the mathematical and computing systems software to enable scientific simulation codes to effectively and efficiently use terascale computers
  - (3) create a collaboratory software environment to enable geographically distributed scientists to work effectively together as a *team* and to facilitate remote access, through appropriate hardware and middleware infrastructure, to facilities, data, and human resources

with the ultimate goal of advancing fundamental research in science central to the DOE mission

Craig.Tull@science.doe.gov  
ICFA Workshop; Daegu, Korea

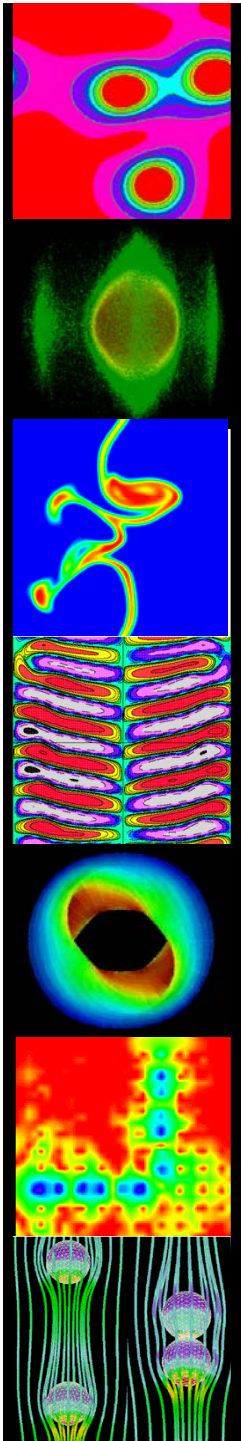




# HEP/NP SciDAC

- National Infrastructure for Lattice Gauge Computing
- Advanced Computing for 21st Century Accelerator Science and Technology
- Shedding New Light on Exploding Stars: Terascale Simulations of Neutrino-Driven Supernovae and Their Nucleosynthesis
- SciDAC Center for Supernova Research
- **The Particle Physics Data Grid (PPDG) Collaboratory Pilot (joint with the Office of Advanced Scientific Computing Research)**
- N.B. High Energy and Nuclear Physics office split into two separate office during SciDAC
  - Project Management and oversight now shared between two offices.

Craig.Tull@science.doe.gov  
ICFA Workshop; Daegu, Korea







# U.S. “Trillium” Grid Partnership



- Trillium = PPDG + GriPhyN + iVDGL
  - Particle Physics Data Grid: \$12M (DOE) (1999 – 2006)
  - GriPhyN: \$12M (NSF) (2000 – 2005)
  - iVDGL: \$14M (NSF) (2001 – 2006)
- Basic composition (~150 people)
  - PPDG: 4 universities, 6 labs
  - GriPhyN: 12 universities, SDSC, 3 labs
  - iVDGL: 18 universities, SDSC, 4 labs, foreign partners
  - Expts: BaBar, D0, STAR, Jlab, CMS, ATLAS, LIGO, SDSS/NVO
- Coordinated internally to meet broad goals
  - GriPhyN: CS research, Virtual Data Toolkit (VDT) development
  - iVDGL: Grid laboratory deployment using VDT, applications
  - PPDG: “End to end” Grid services, monitoring, analysis
  - Common use of VDT for underlying Grid middleware
  - Unified entity when collaborating internationally



Craig.Tull@science.doe.gov  
ICFA Workshop; Daegu, Korea



# Grid3 in the U.S.

---

- We built a demonstration functioning Grid in the U.S.  
Strongly encouraged by NSF and supported through iVDGL, GriPhyN, PPDG (DOE) as well as US-CMS, US-ATLAS, Ligo, SDSS and University of Buffalo, LBNL, some Biology applications, and more.... Based on a simple to install VDT package
  - Allowing diversity at different sites
- It worked and it stays working and it is being used by ATLAS and CMS for their data challenges
  - <http://ivdgl.org/Grid3>

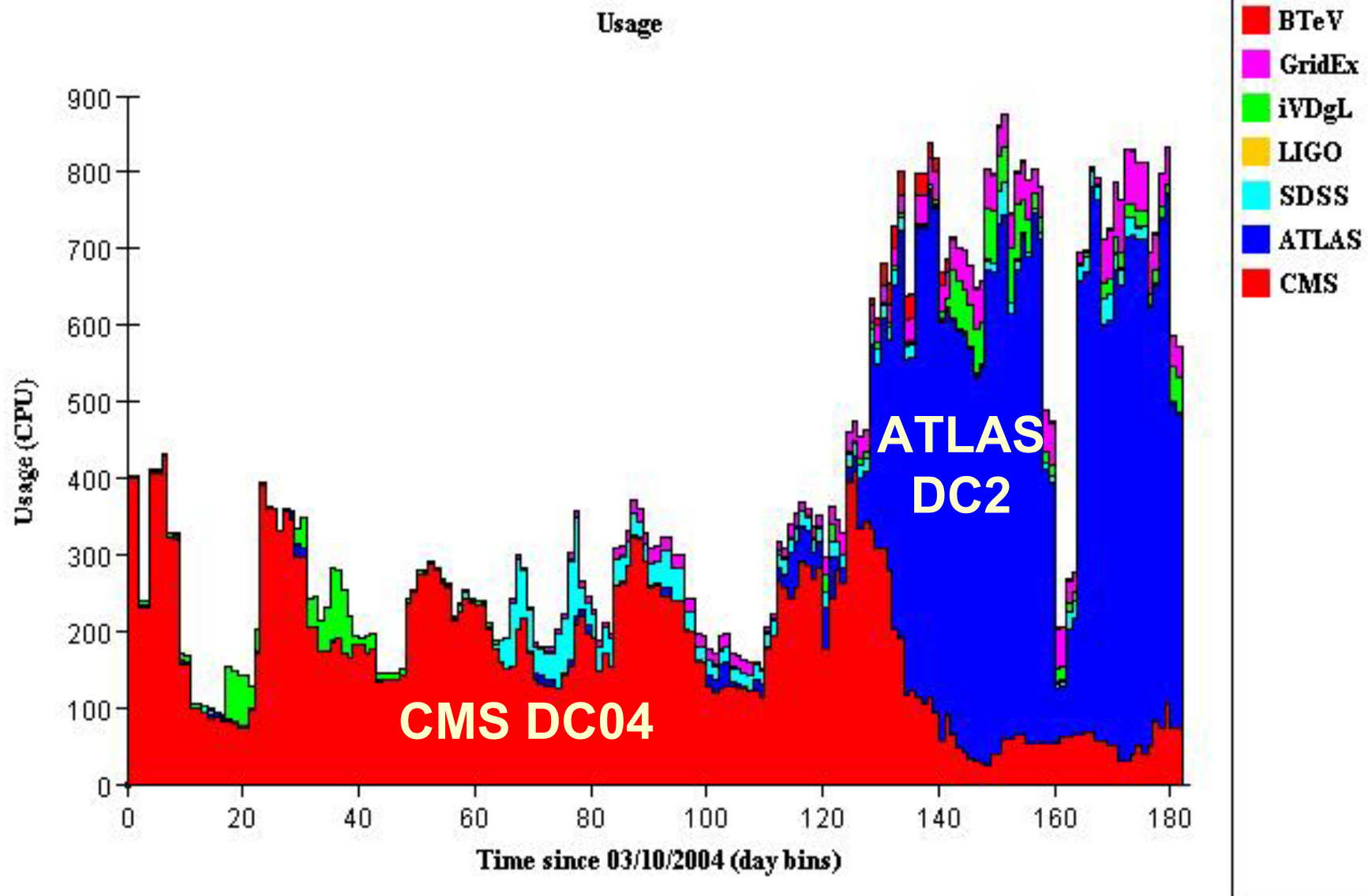


ICFA Workshop; Daegu, Korea





# Grid 3 Shared Usage





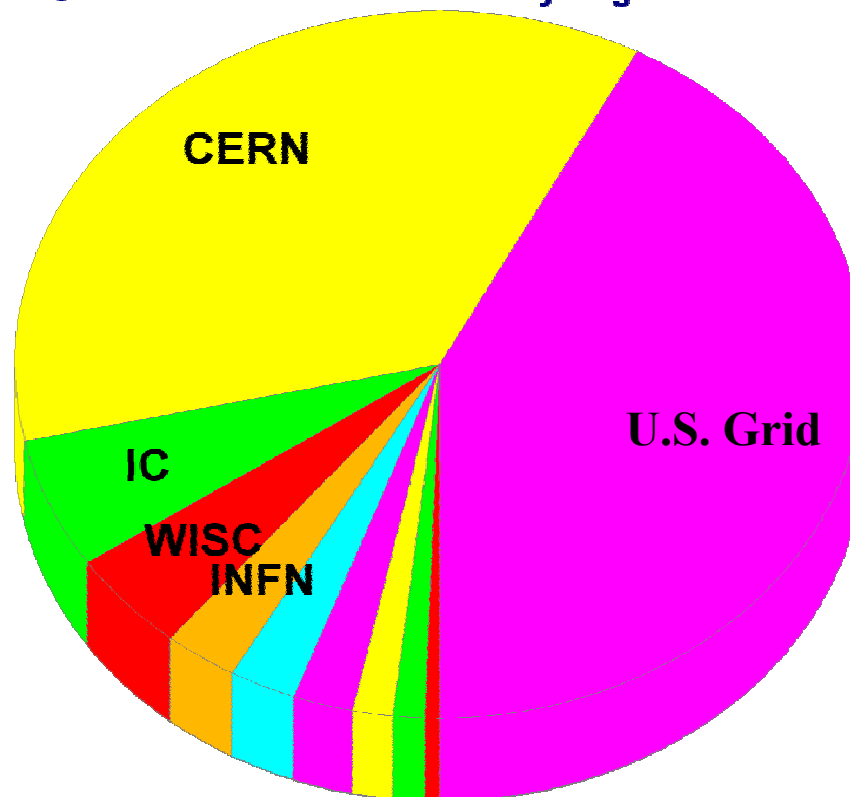


# CMS Use of U.S. Grids and European Regional Centers



- U.S. Grids largest contributor to CMS Event Simulation Production

Simulation Events Produced by Regional Center

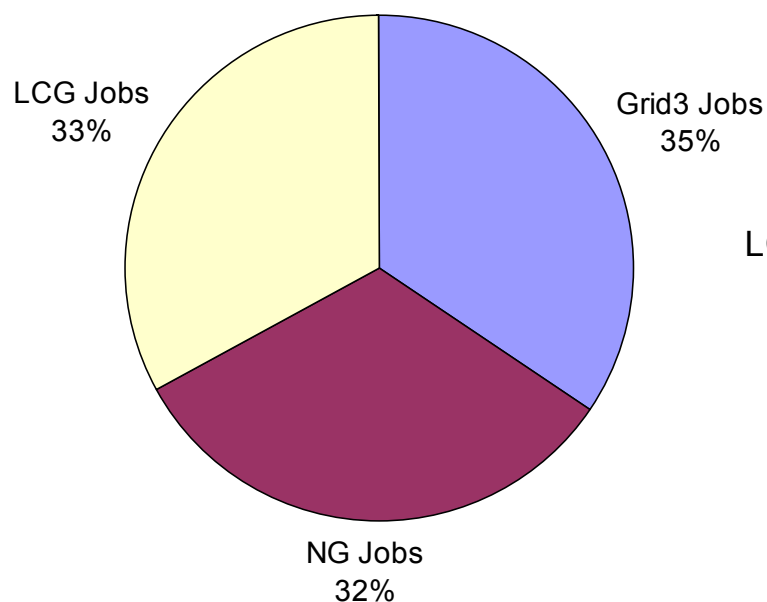


Craig.Tull@science.doe.gov  
ICFA Workshop; Daegu, Korea

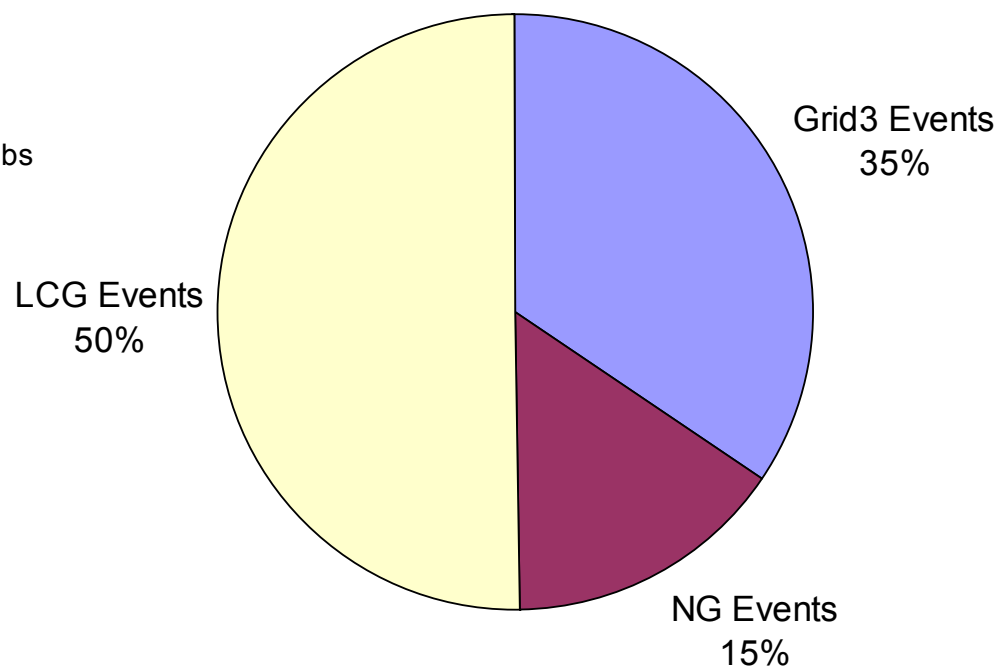


# ATLAS DC2 Jobs Per Grid

DC2 Jobs (till Jan 22, 2005)



DC2 Events Processed (till Jan 22, 2005)



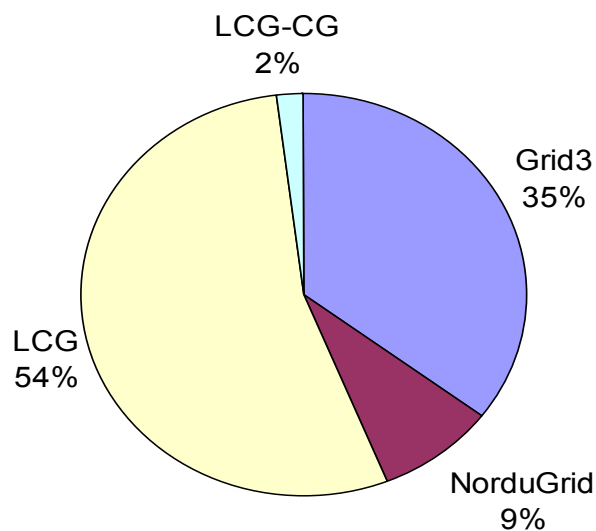


# ATLAS Rome Production Summary (Mar'05)

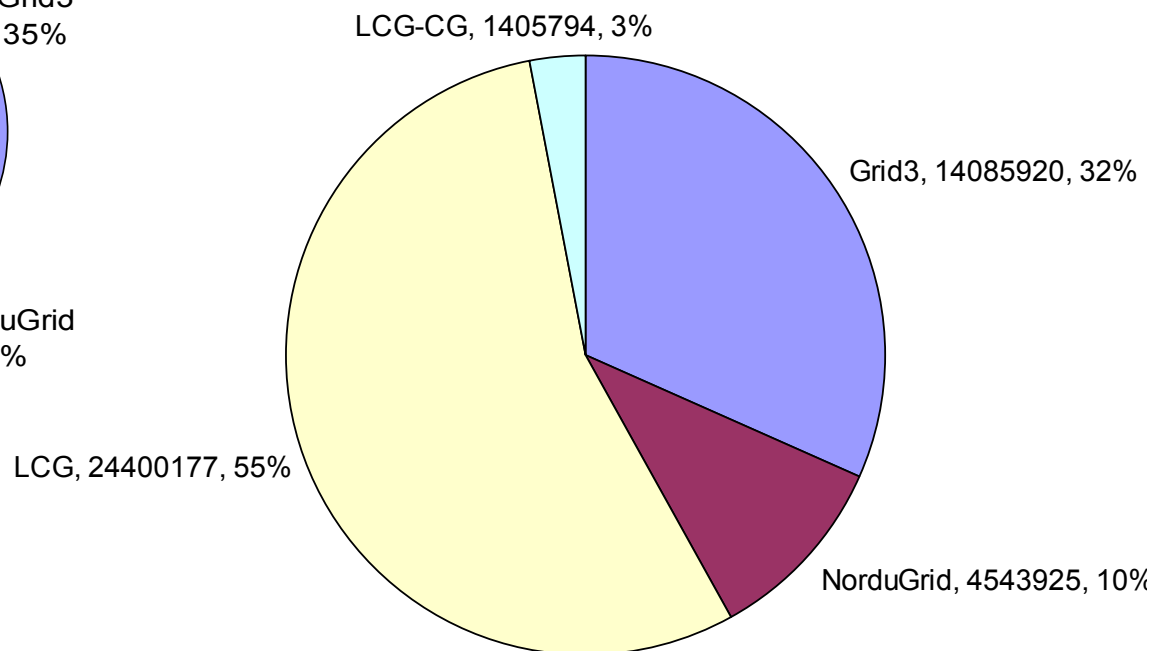


**14 million days of  
SI2000 CPU used on  
Grid3 in past 2 months**

Rome Total Jobs



Rome Total CPU SI2k-days



Only successful jobs included in these plots



# Open Science Grid Consortium



- What is it?
  - It is NOT a project (unlike EGEE, iVDGL, TeraGrid, etc)
  - It is a collaboration – a Consortium of many institutions. Universities and Labs, projects, experiments, middleware providers, campus Grids
    - Who want to leverage their efforts by joining together to build a sustainable infrastructure for physics and other sciences







# The Blueprint and Roadmap

---



- Agreement by architects on some fundamental goals
  - Aligned with and informed by EGEE architecture and design
- Illustrative Principles:
  - Common Interfaces not necessarily Implementations.
  - Autonomy of Sites and Resources
    - Overriding Policy Enforcement Points.
    - Move to “pull” model for Work to be done.
  - Organization based environment: services & management
    - Common services with VO-specific instantiations.
  - Mix of Assured and Opportunistic use.
    - Varied expectations on resource availability, performance and support.



# OSG Core Infrastructure Services and Middleware

---



- Core middleware packages based on VDT 1.3.6.
  - Added EDG/EGEE VOMS, LBNL DRM, Privilege Services (GUMS,Prima).
  - Integration and testing provided by multi-platform NMI support infrastructure.
- US CMS Tier-1 & Tier-2 Storage Elements accessible via SRM Grid interfaces
  - Support for opportunistic use of this storage by non-US CMS VOs.
- VO role based authorization through Gatekeeper and GridFTP standard callouts to Prima.
- Site based dynamic and role based account management and policies applied using GUMS.



# Towards OSG Deployment

---

- Maintaining Grid3 operations
  - In parallel with extending Grid3 to OSG
- OSG Integration Testbed (ITB) has 20 sites and ITB infrastructure has had 5 releases.
- ATLAS, CMS, STAR, CDF, D0, BaBar, FMRI, readying their applications & infrastructure for the common grid:
  - CDF run simulation on an OSG ITB site
  - D0 running re-reconstruction on US CMS Tier-1 Grid3.
  - STAR running on BNL and LBNL ITB sites
  - SLAC, FermiGrid, PDSF Facilities
- Deployment now started to support applications from ~10 VOs and transition from Grid3.

<http://osg.ivdgl.or/twiki/bin/view/Integration/WebHome>



# US DOE-HEP Partnerships

---

- DOE-ASCR
  - ESnet, SciDAC, Collaboratories
  - SciDAC II in 2006
- NSF
  - Abilene, Tier-2 centers, OSG, ...
  - ESnet/Abilene recently established high-speed interconnects and cross-network routing
- CERN/LCG/EGEE
  - LHCnet, Worldwide Grid Federation for LHC
  - Very important to establish & maintain agreements on interfaces to ensure interoperability

Craig.Tull@science.doe.gov  
ICFA Workshop; Daegu, Korea

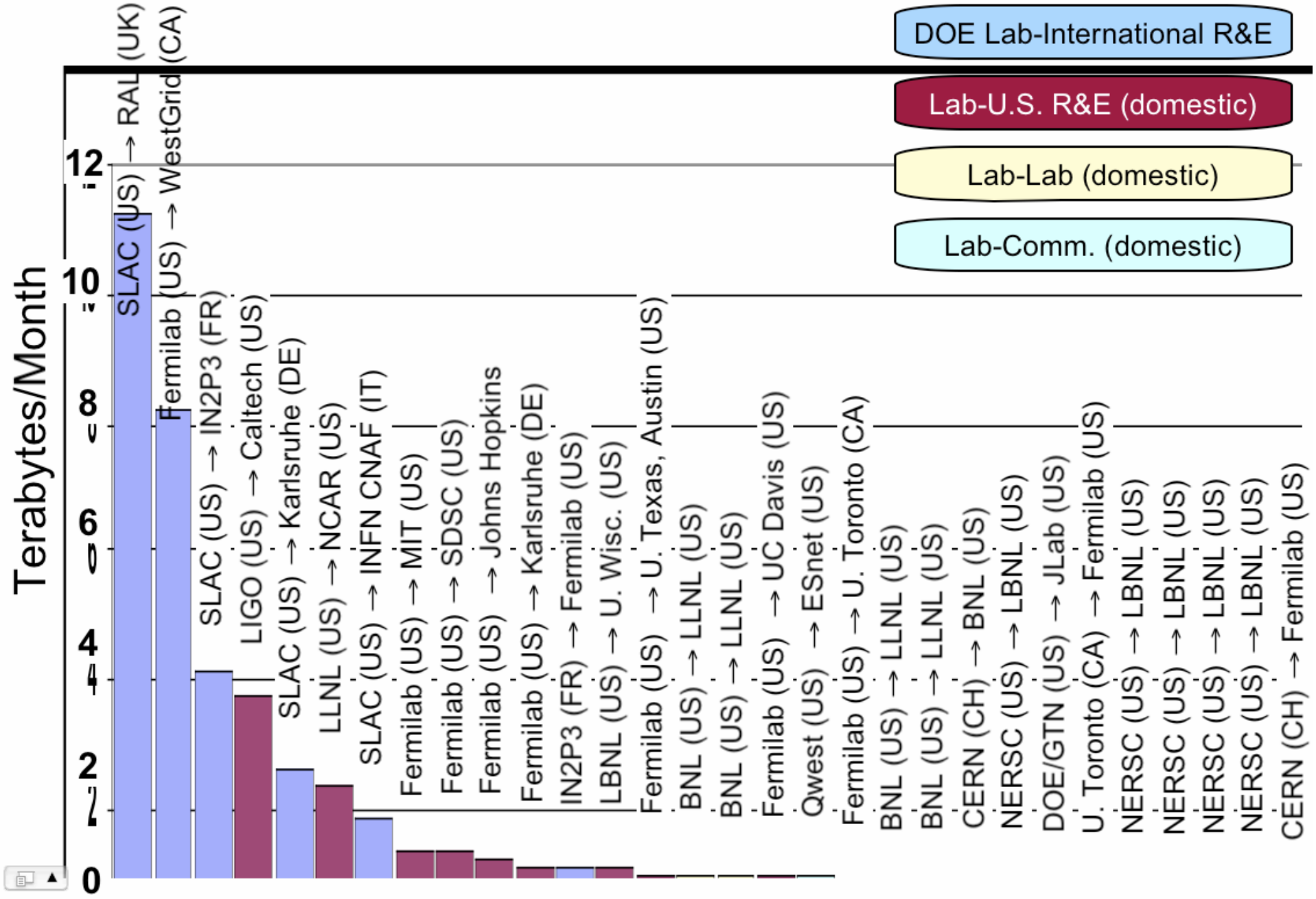




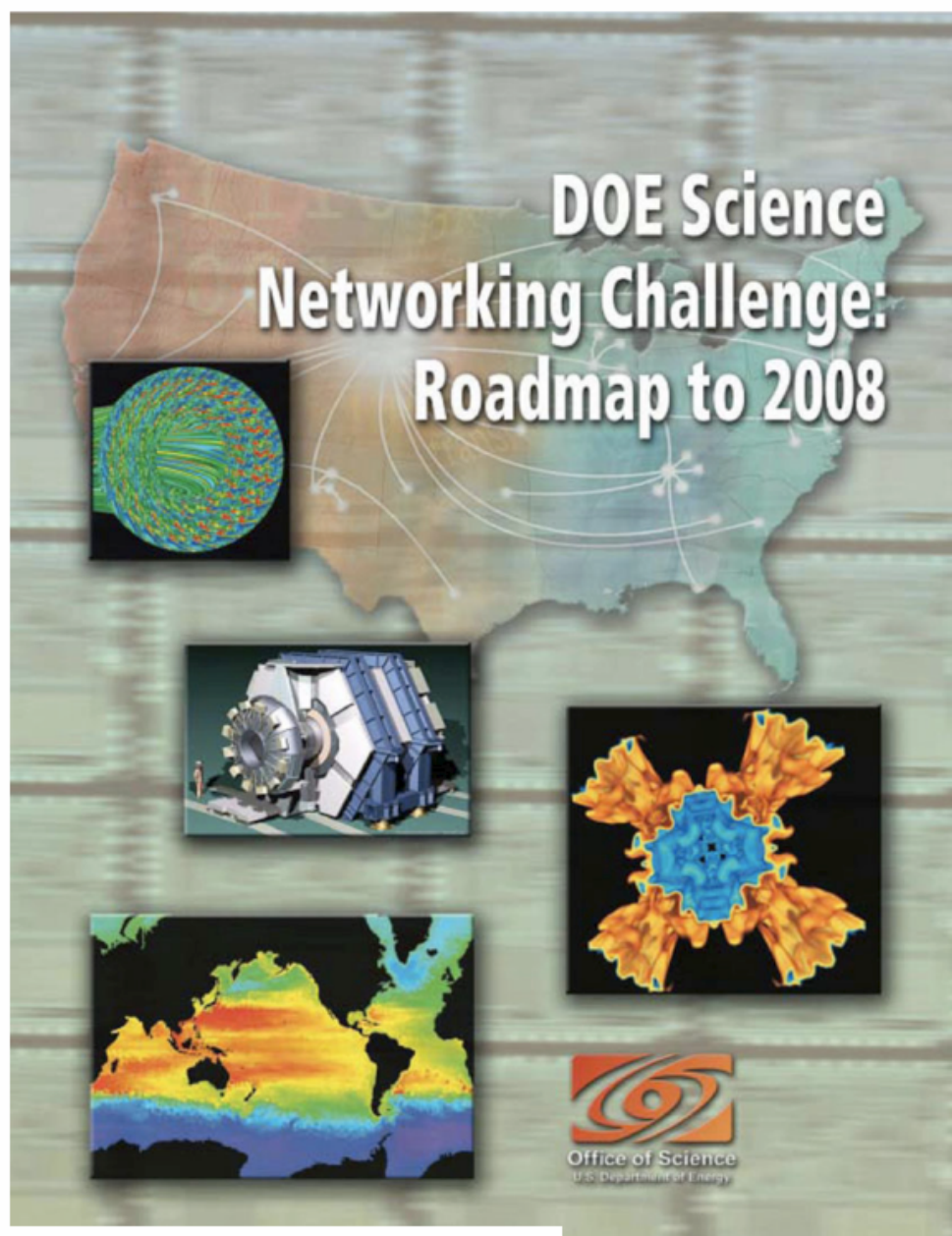
---

Craig.Tull@science.doe.gov  
ICFA Workshop; Daegu, Korea

# Source and Destination of the Top 30 Flows, Feb. 2005



# New Strategic Directions to Address Needs of DOE Science



Available at [www.es.net/#research](http://www.es.net/#research)

**June 3-5, 2003**

**Organized by the ESnet  
Steering Committee**

**(OSC program representatives)**

**Workshop Chair**

**Roy Whitney, JLAB**

**Report Editors**

**Roy Whitney, Larry Price, ANL**

**Workshop Panel Chairs**

**Wu-chun Feng, LANL**

**William Johnston, LBNL**

**Nagi Rao, ORNL**

**David Schissel, GA**

**Vicky White, FNAL**

**Dean Williams, LLNL**

- Developed a qualitative approach to achieve the science driven network requirements of the previous workshop



# Observed Drivers for ESnet Evolution



- The combination of
  - the exponential growth in ESnet traffic, and
  - observed large science data flows becoming a significant fraction of all ESnet trafficshow that the projections of the science community are reasonable and are being realized
- The current predominance of international traffic is due to high-energy physics
  - However, all of the LHC US tier-2 data analysis centers are at US universities
  - As the tier-2 centers come on-line, the DOE Lab to US university traffic will increase substantially
- High energy physics is several years ahead of the other science disciplines in data generation
  - Several other disciplines and facilities (e.g. climate modeling and the supercomputer centers) will contribute comparable amounts of additional traffic in the next few years



# ESnet's & U. S. and International Science

---



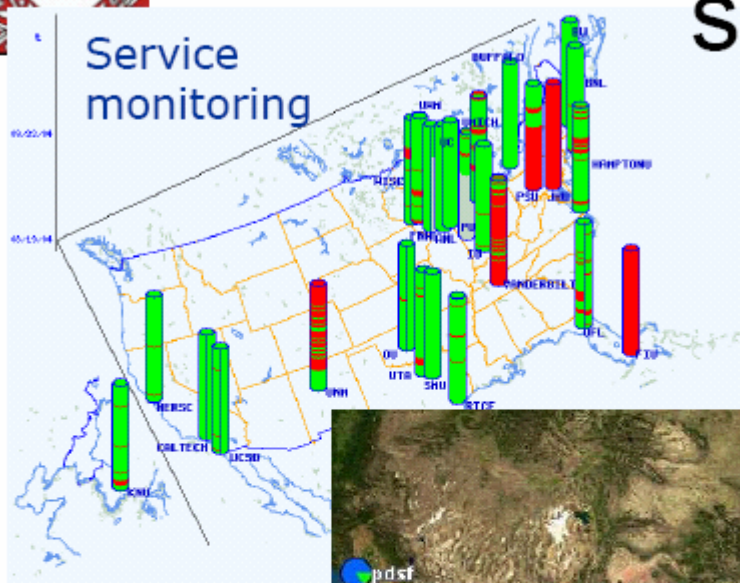
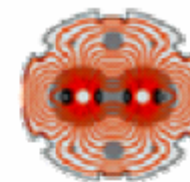
- ESnet and Abilene together provide most of the nation's transit networking for basic science
  - Abilene provides national transit networking for most of the US universities by interconnecting the regional networks (mostly via the GigaPoPs)
  - ESnet provides national transit networking for the DOE Labs
- ESnet and Abilene have recently established high-speed interconnects and cross-network routing
- Goal is that DOE Lab-Univ. connectivity should be as good as Lab-Lab and Univ.-Univ.



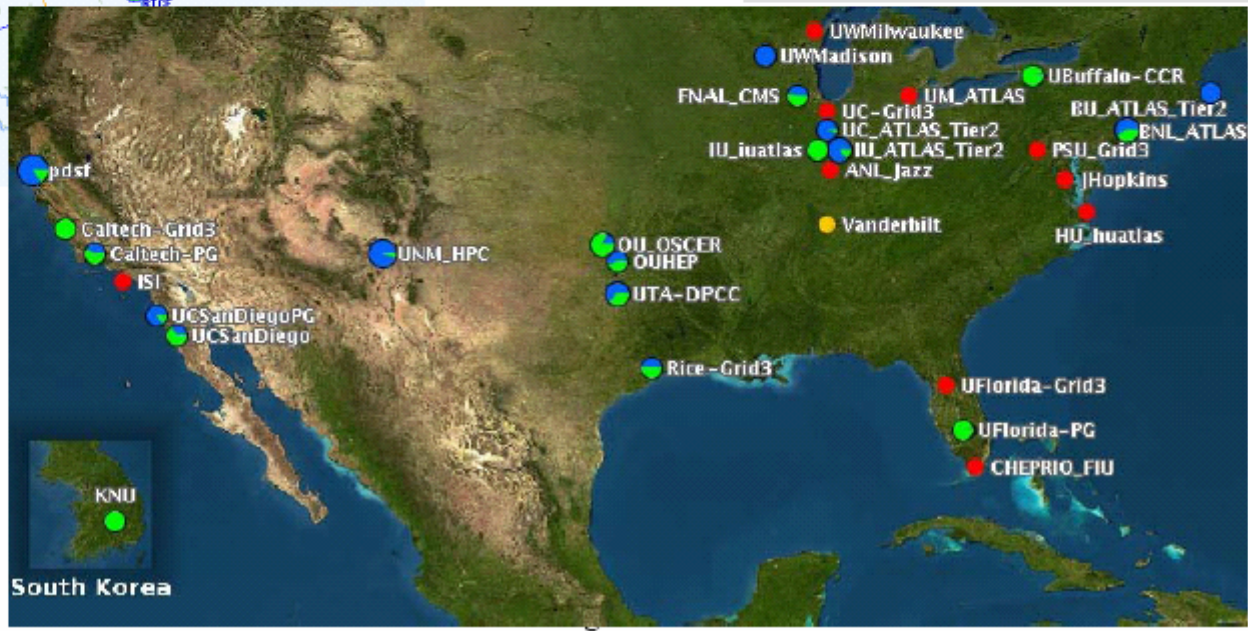


# Grid 3

## Grid3 – a snapshot of sites



**Sep 04**  
 •30 sites, multi-VO  
 •shared resources  
 •~3000 CPUs (shared)



21 February 2005

# ATLAS DC2: countries (sites)



Australia (1)

Austria (1)

Canada (4)

CERN (1)

Czech Republic (2)

Denmark (4)

France (1)

Germany (1+2)

Italy (7)

Japan (1)

Netherlands (1)

Norway (3)

Poland (1)

Slovenia (1)

Spain (3)

Sweden (7)

Switzerland (1)

Taiwan (1)

UK (7)

USA (19)



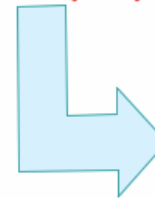
20 countries  
69 sites



13 countries; 31 sites



7 countries; 19 sites



SW\_week\_May\_2005

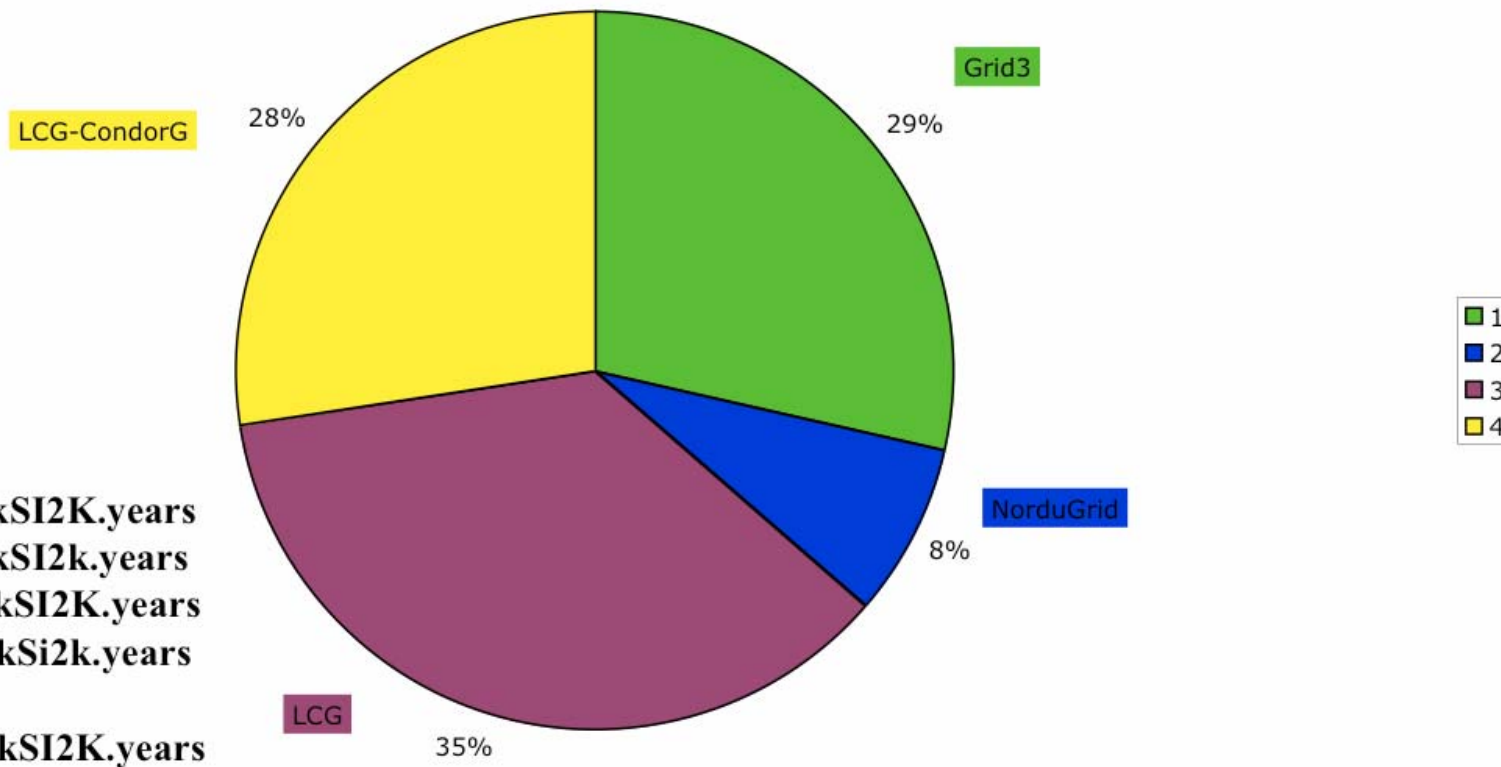
G. Poulard - CERN PH

2

# ATLAS Rome-production



## Rome production - CPU usage



Grid3: 80 kSI2K.years  
NordGrid: 22 kSI2k.years  
LCG: 101 kSI2K.years  
LCG-CG: 77 kSi2k.years  
Total: 280 kSI2K.years

(from ProdDB)  
SW\_week\_May\_2005

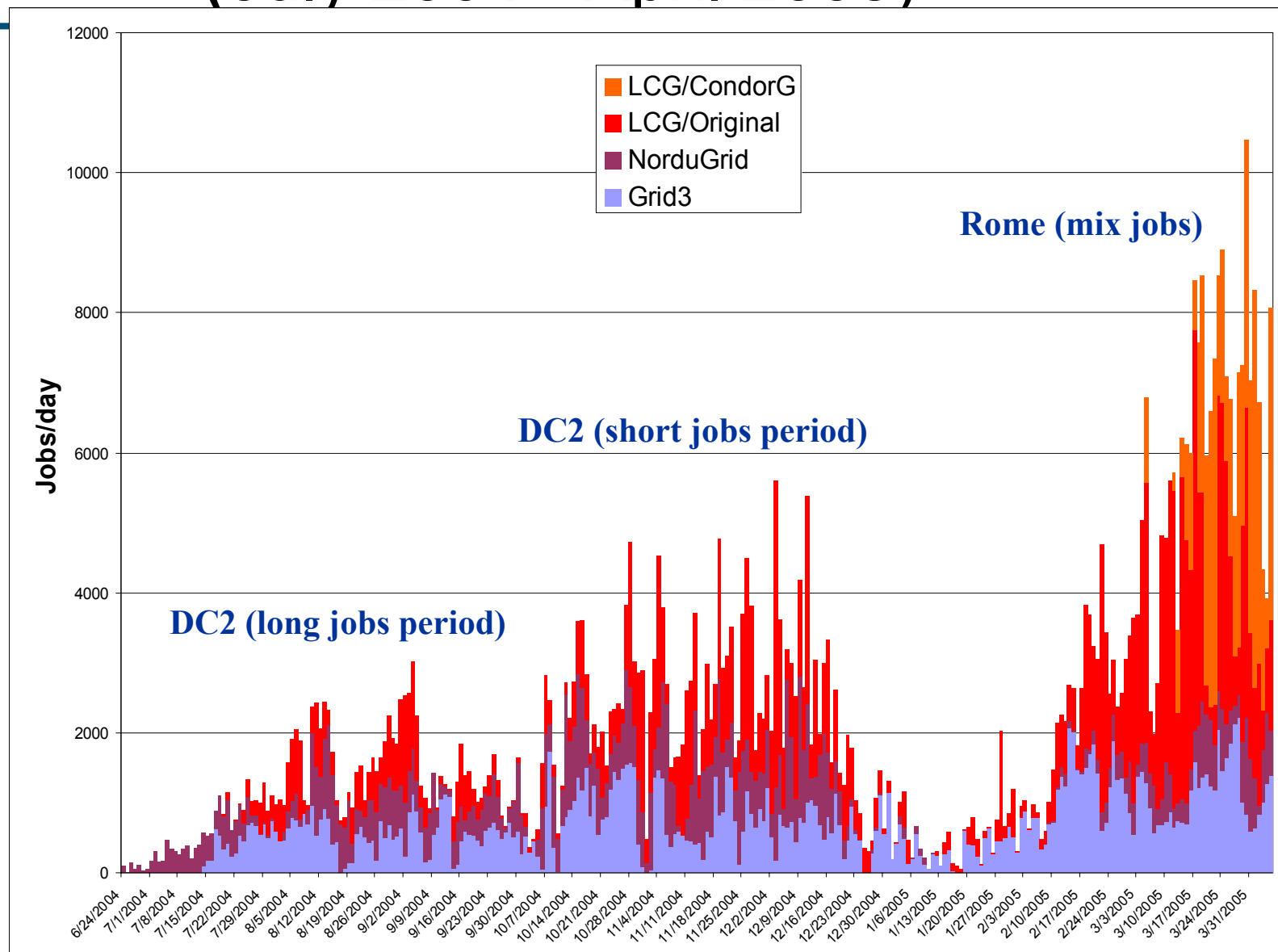
G. Poulard - CERN PH

6

Craig.Tull@science.doe.gov  
ICFA Workshop; Daegu, Korea



# ATLAS Production (July 2004 - April 2005)







# Open Science Grid Consortium



- What is it?
  - It is NOT a project (unlike EGEE, iVDGL, TeraGrid, etc)
  - It is a collaboration – a Consortium of many institutions. Universities and Labs, projects, experiments, middleware providers, campus Grids
    - Who want to leverage their efforts by joining together to build a sustainable infrastructure for physics and other sciences



# DC2 Phase I (June-Dec, 2004)

- LCG
  - Included some non-ATLAS sites, used the LCG-Grid-Canada interface
- NorduGrid
  - Scandinavian resources + sites in Australia, Germany, Slovenia, Switzerland
- Grid3
  - Used computing resources that are not dedicated to ATLAS

## ATLAS DC2 Production Status

### Overview of Grids

as of 2005-01-28 10:33:00

Grid	submitted	pending	running	finished	failed	efficiency
Grid3	27	10	215	122531	41385	75 %
NorduGrid	16			111108	66311	63 %
LCG	32	380	432	122688	215105	36 %
<b>TOTAL</b>	<b>75</b>	<b>390</b>	<b>647</b>	<b>356327</b>	<b>322801</b>	<b>52 %</b>

[Prodsys homepage](#)

Above statistics does not include Rome Production



---

Craig.Tull@science.doe.gov  
ICFA Workshop; Daegu, Korea



---

Craig.Tull@science.doe.gov  
ICFA Workshop; Daegu, Korea





# Open Science Grid Consortium



- What is it?
  - It is NOT a project (unlike EGEE, iVDGL, TeraGrid, etc)
  - It is a collaboration – a Consortium of many institutions. Universities and Labs, projects, experiments, middleware providers, campus Grids
    - Who want to leverage their efforts by joining together to build a sustainable infrastructure for physics and other sciences



# Open Science Grid Consortium

---



- Enables US-CMS and US-ATLAS to work together in a highly coordinated way
  - They can provide resources to ATLAS and CMS through the Open Science Grid as participants in the global “LHC Computing Grid”
- Seeing partners and interest in joining this Grid outside of Physics & outside of the US also
  - Korea
  - Brasil
  - Taiwan
  - ?
- Joining OSG isn't exclusive. We believe resource providers can be a member of many Grid infrastructures



# The Blueprint and Roadmap

---



- Agreement by architects on some fundamental goals
  - Aligned with and informed by EGEE architecture and design
- Illustrative Principles:
  - Common Interfaces not necessarily Implementations.
  - Autonomy of Sites and Resources
    - Overriding Policy Enforcement Points.
    - Move to “pull” model for Work to be done.
  - Organization based environment: services & management
    - Common services with VO-specific instantiations.
  - Mix of Assured and Opportunistic use.
    - Varied expectations on resource availability, performance and support.



# OSG Infrastructure - a Federated Grid of Grids

---



- Heterogeneous Infrastructure made coherent by a set of Interfaces, core Services and Policies.
- Campus, Facility, Experiment, (Commercial?), “Other” Grids present (existing) Resources and Services that can be shared across local and wide area Grids, as well as support direct non-grid based interfaces.
- Organizations & VOs present a transparent distributed execution and storage environment to the Users that spans across Grid and Facility boundaries.





---

Craig.Tull@science.doe.gov  
ICFA Workshop; Daegu, Korea



---

Craig.Tull@science.doe.gov  
ICFA Workshop; Daegu, Korea