- Logistics

for CMS/ATLAS sessions coffee will be served outside rooms 214/216

lunch can be ordered in the cafeteria on the main floor. The cafeteria is cash only, and there is an ATM in the register area.

There is reserved eating space on the second floor (marked in the program).

dinner: dinner

and transportation are included in registration, but people had to select the dinner option in order to take advantage of this. The bus will be leaving at 6:45PM sharp from Best Western (NOT THE CONFERENCE CENTER). Please bring your badge (which will be marked with a sticker). We might be able to accommodate additional people at the restaurant, but they will have to arrange transportation and let us know as early as possible (please email ahm2011@sbgrid.org, or talk to us at the registration desk in AM).

Welcome

Its been great year

The LHC Took Data to Discovery;

Non-LHC Usage - 50%; Non Physics - 20%

SBGrid used 6 million CPU hours

OSG, Boston, 2011

It's another excellent program at the All Hands Meeting

Kudos to Piotr,
his Program Committee,
Leaders of the Workshops,
Presenters

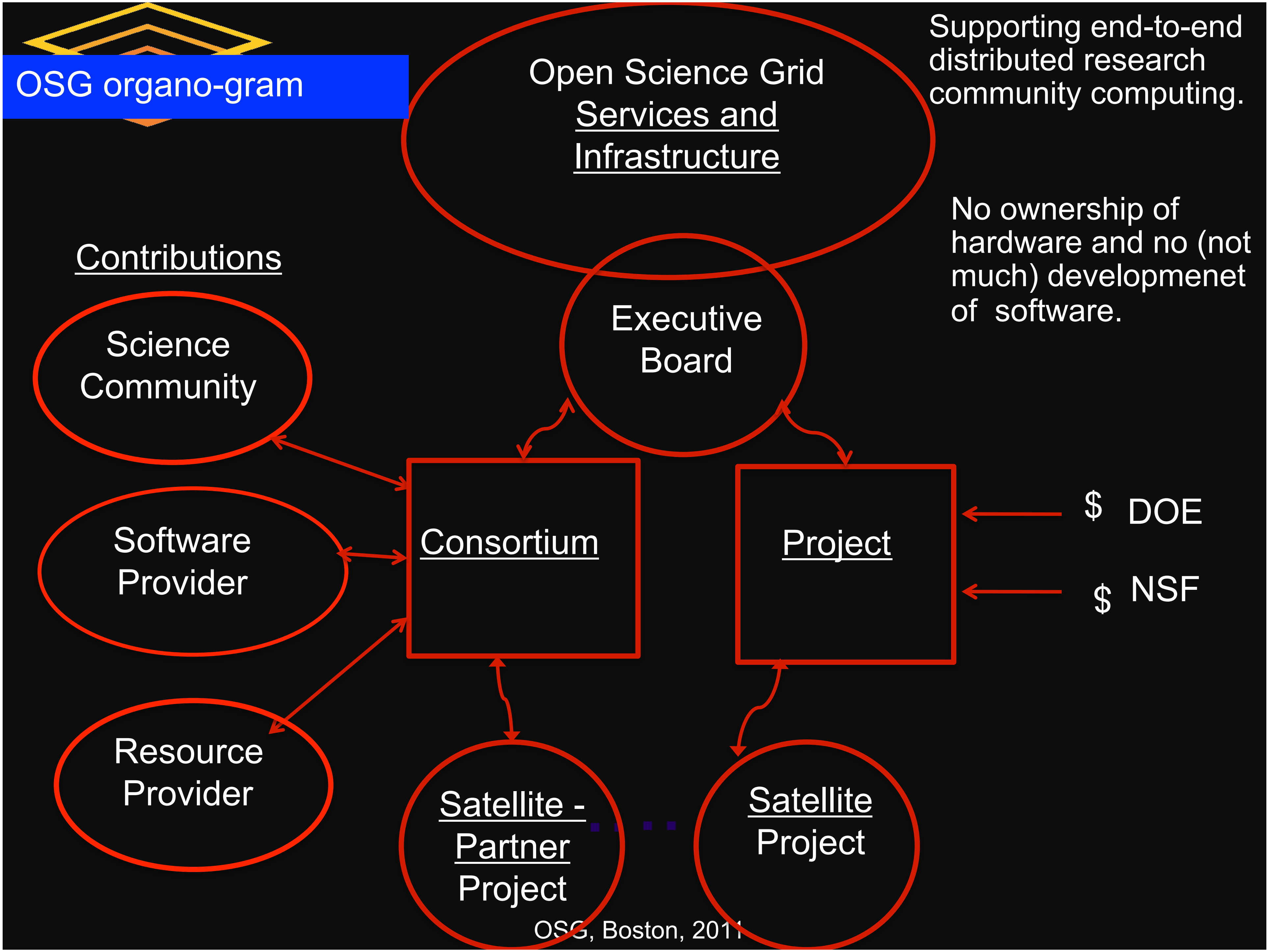
Look around for old(er) and new(er) faces.

Executive Team & Area Leads
are In & available to talk

OSG "Staff" here this week

Michael Ernst	Jose Caballero	Maxim Potekhin	John Hover
Miron Livny	Tim Cartwright	Greg Thain	Alain Roy
Scot Kornenfeld	Robert Engel	Britta Daudert	Kent Blackburn
Dan Fraser	Robert Gardner	Marco Mambelli	Suchandra Thapa
Tanya Levshina	Doug Strain	Steve Timm	Gabriele Garzoglio
Mine Altunay	Miriam Boon	Chander Sehgal	Jim Weichel
Alain Deximo	Elizabeth Chism	Robert Quick	Kyle Gross
Scott Teige	Ewa Deelman	Mats Rynge	Alex Sim
Jeff Dost	Derek Weitzel	Frank Wuerthwein	Igor Sfiligoi
Anand Padmanabhan			

You are the Backbone

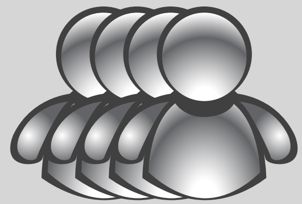


OSG organo-gram
Open Science Grid Services and Infrastructure
Supporting end-to-end distributed research community computing.
No ownership of hardware and no (not much) developmenet of software.
Contributions
Science Community
Software Provider
Resource Provider
Executive Board
Consortium
Project
$ DOE
$ NSF
Satellite - Partner Project
Satellite Project
OSG, Boston, 2011

OSG's Community-Focused Architecture

Communities...
(LHC, LIGO, SBGrid, Campuses...)

... have **users**,
application developers,
site administrators,
researchers, students etc:



... may use
custom **services** in
addition to OSG's:

Job management
Data management
...

...may provide diverse
resources via one
or more autonomous sites:

Compute Interface	Storage Interface
Site Fabric	System Status Accounting Security

OSG

Consultation Services:
- Community
- Security
- Architecture
- Education/Training

Operational Services:
- Front-line support
- Security
- Info/monitoring
- WLCG Interaction
- IT Services

Software Services
- Software Distribution
- Software Evaluations
- Support

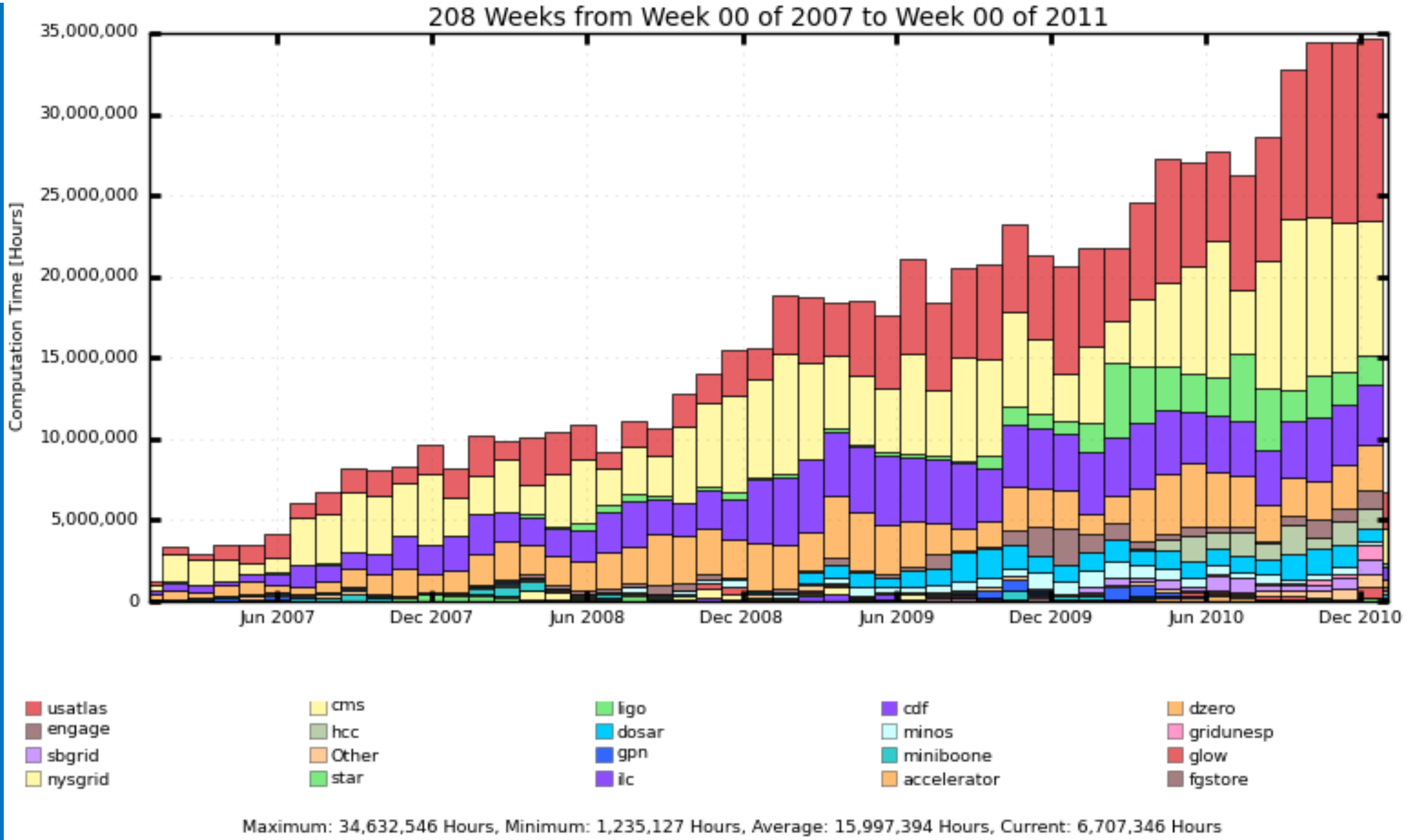
There is a sharing of software, operational services,
and knowledge between the communities
and OSG in each of these areas.

You are the Backbone

Usage & Growth

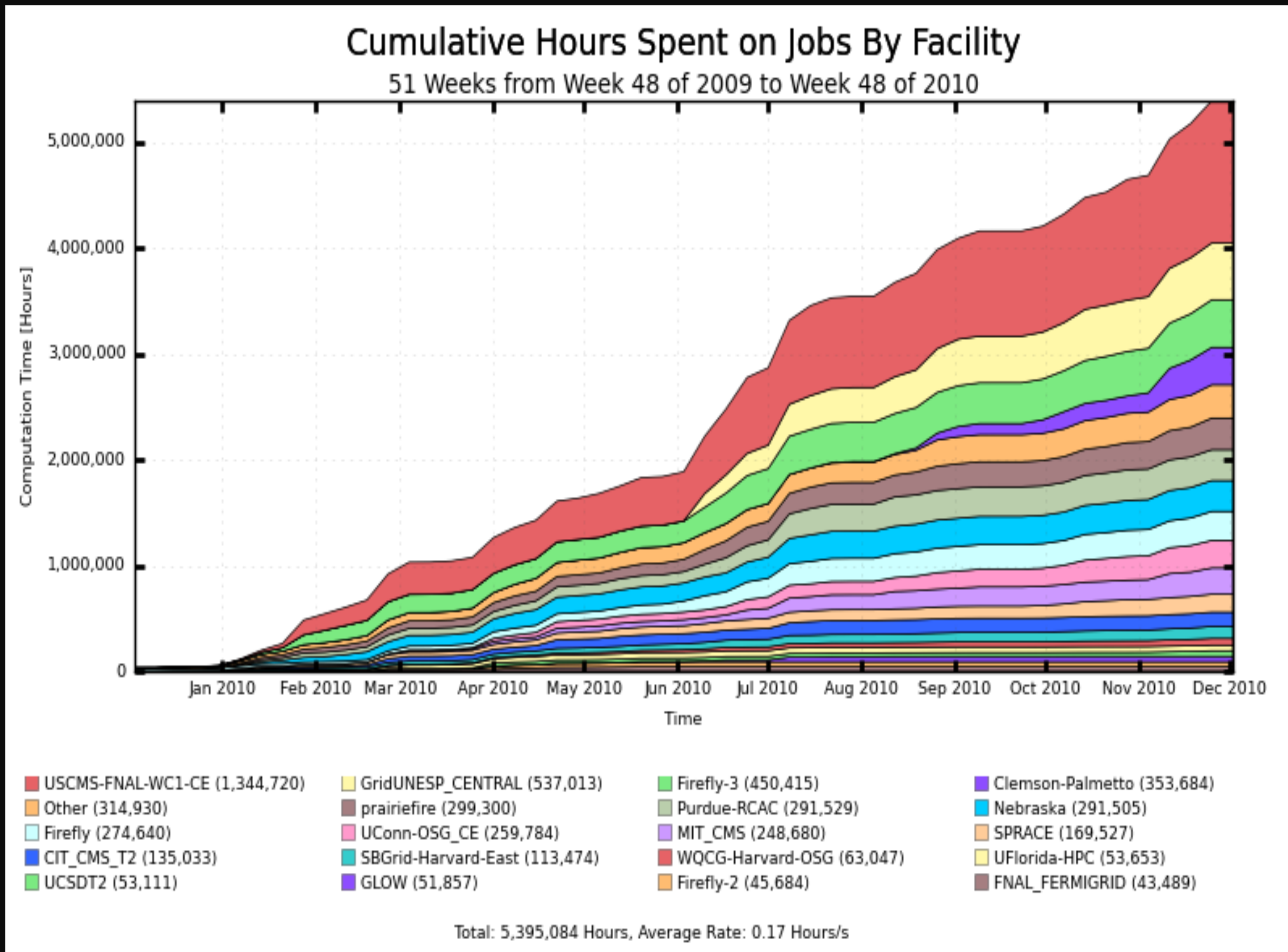
Campuses & Tier-3s

Satellites



Daily Usage of >1,000,000 hours/day
20 Communities showing usage (~4 are multi-science)

SBGrid –
5 million as SBGrid + 1 million in Engage



OSG Operations has handled >10,000 Tickets!

From Rob Quick, echoed by the ET:

While there is nothing special about ticket 10000 itself, it is very typical of the type of ticket seen by the Ops Team, it is a good place to reflect on the number of issues addressed day in and day out for the past five years.

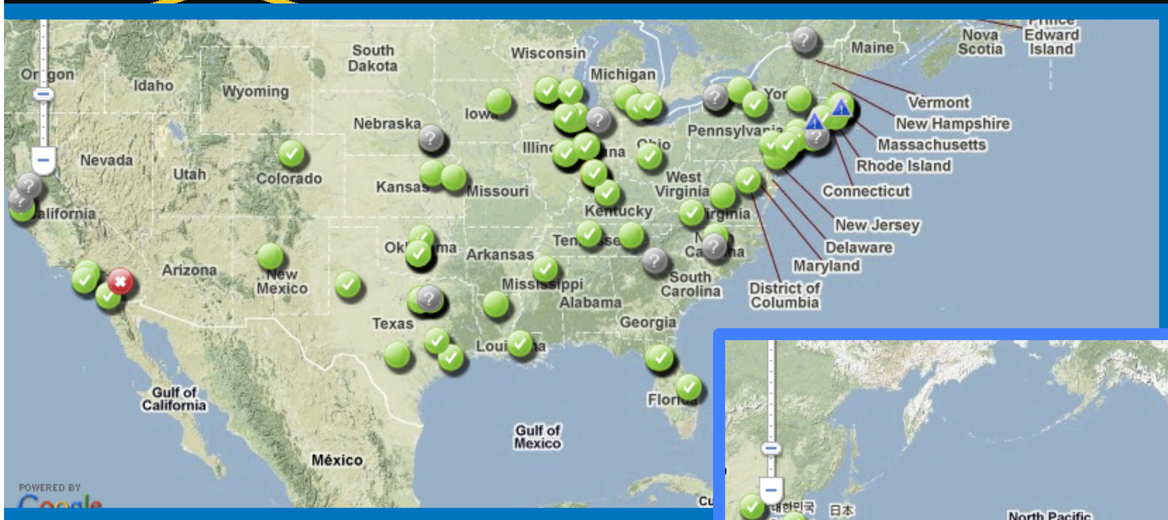
It is also a good place to point out the time, effort, and commitment given on a daily basis from both OSG Operations and the OSG Community as a whole to resolving problems faced by its users and resources providers.

Year of (measuring) the 99.9X% operations services

Availability Metrics

Between Feb 5, 2011 00:00:00 UTC and Mar 7, 2011 00:00:00 UTC

Resource	Service	Availability	Reliability
GRATIA-OSG			OSG Production Resource Group
GRATIA-OSG-ITB	Gratia Collector	99.56%	99.56%
GRATIA-OSG-PROD	Gratia Collector	98.58%	98.58%
GRATIA-OSG-TRANSFER	Gratia Collector	99.42%	99.42%
GOC_BDII			OSG Production Resource Group
GOC_BDII_1	OSG BDII	99.98%	99.98%
	WLCG Interoperability BDII	99.63%	99.63%
GOC_BDII_2	OSG BDII	100%	100%
	WLCG Interoperability BDII	100%	100%
GOC_OIM			OSG Production Resource Group
GOC_OIM	OIM	99.95%	99.95%
GOC_RSV_Collector			OSG Production Resource Group
RSV_Collector	RSV Collector	99.7%	99.7%
GOC_Software_Cache			OSG Production Resource Group
GOC_Software_Cache_1	Software Cache	99.72%	99.72%
GOC_Software_Cache_2	Software Cache	99.95%	99.95%
GOC_Ticket			OSG Production Resource Group
GOC_Ticket_1	GOCTicket	99.77%	99.77%
GOC_Ticket_2	GOCTicket	99.95%	99.95%
MyOSG			OSG Production Resource Group
MyOSG_1	MyOSG	99.7%	99.7%
MyOSG_2	MyOSG	99.91%	99.91%
OSG_Display			OSG Production Resource Group
OSG_Display_1	OSG_Display	97.87%	97.87%
OSG_TWiki			OSG Production Resource Group
OSG_TWiki	Twiki Server	99.81%	99.81%



95 OSG 1.2.X resources

8 are 1.2.18, 6 are 1.2.17

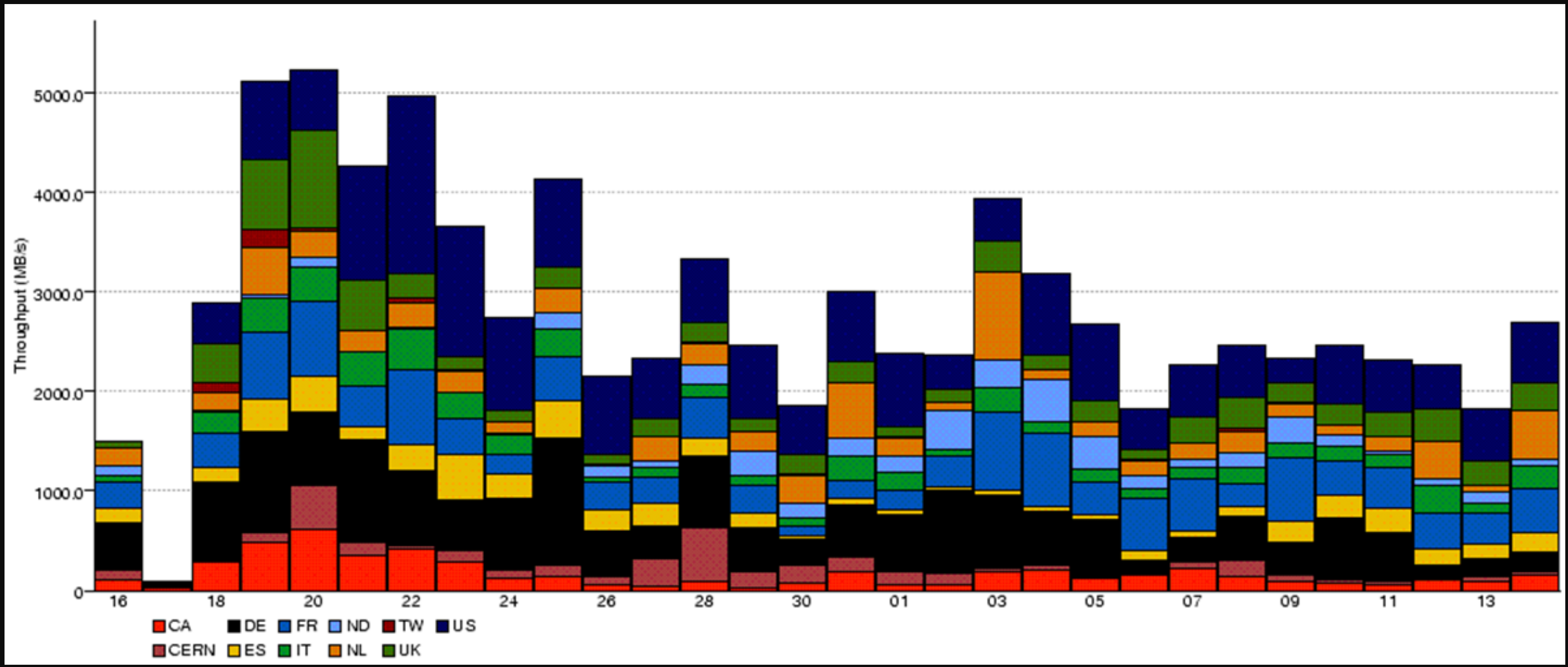
3 OSG 1.0.X resources

2 OSG 1.0.0 resources

1 OSG 0.8.0 resources

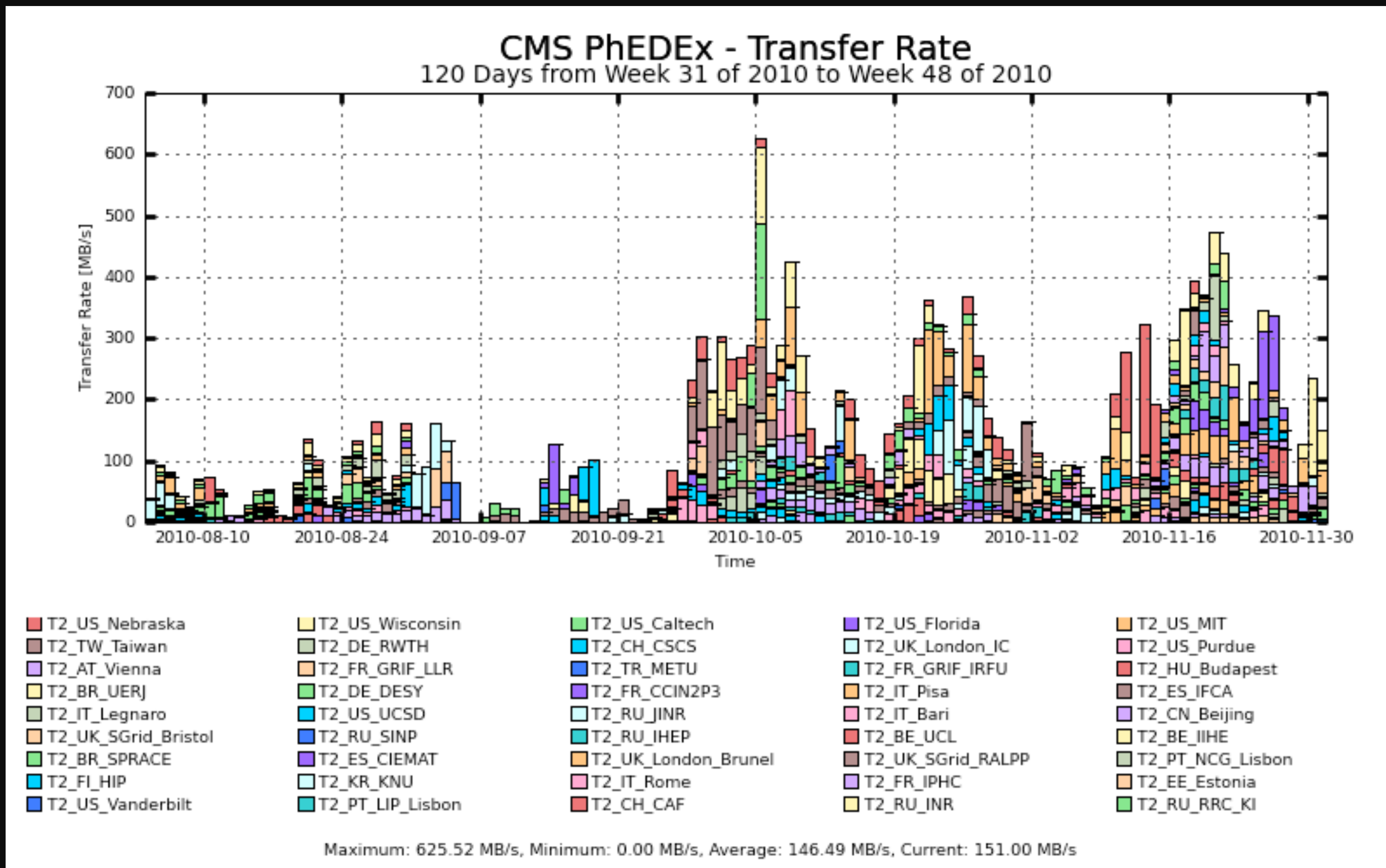
Data Moved!

Recent ATLAS Cloud Dataflows



Last 30 days of ATLAS data-flows by Tier-1 “Clouds”
Rates are consistently above 2GB/sec even while LHC is down.

CMS T2 to T2


CMS PhEDEx - Transfer Rate
120 Days from Week 31 of 2010 to Week 48 of 2010
Transfer Rate [MB/s]
0
100
200
300
400
500
600
700
2010-08-10
2010-08-24
2010-09-07
2010-09-21
2010-10-05
2010-10-19
2010-11-02
2010-11-16
2010-11-30
Time
T2_US_Nebraska
T2_TW_Taiwan
T2_AT_Vienna
T2_BR_UERJ
T2_IT_Legnaro
T2_UK_SGrid_Bristol
T2_BR_SPRACE
T2_FI_HIP
T2_US_Vanderbilt
T2_US_Wisconsin
T2_DE_RWTH
T2_FR_GRIF_LLR
T2_DE_DESY
T2_US_UCSD
T2_RU_SINP
T2_ES_CIEMAT
T2_KR_KNU
T2_PT_LIP_Lisbon
T2_US_Caltech
T2_CH_CSCS
T2_TR_METU
T2_FR_CCIN2P3
T2_RU_JINR
T2_RU_IHEP
T2_UK_London_Brunel
T2_IT_Rome
T2_CH_CAF
T2_US_Florida
T2_UK_London_IC
T2_FR_GRIF_IRFU
T2_IT_Pisa
T2_IT_Bari
T2_BE_UCL
T2_UK_SGrid_RALPP
T2_FR_IPHC
T2_RU_INR
T2_US_MIT
T2_US_Purdue
T2_HU_Budapest
T2_ES_IFCA
T2_CN_Beijing
T2_BE_IIHE
T2_PT_NCG_Lisbon
T2_EE_Estonia
T2_RU_RRC_KI
Maximum: 625.52 MB/s, Minimum: 0.00 MB/s, Average: 146.49 MB/s, Current: 151.00 MB/s

The year of the Tier-3s

Help us make a complete list?

Analysis needs of the US LHC experiments

Leading to changes in the Data Models…

"Any Data, Anytime, Anywhere"

Otherwise known as remote I/O?

BELLARMINE-ATLAS-T3	brown-cms
DukeT3	FSU-HEP
GLOW-ATLAS	FLTECH
WISC-ATLAS	Rice
WISC-ATLAS	osu-cms
IllinoisHEP	Princeton_ICSE_T3_CMS
IllinoisHEP	rutgers-cms
IllinoisHEP	TAMU_BRAZOS
IllinoisHEP	UCD
SMU_PHY	UCLA_Saxon_Tier3
SMU_HPC	UColorado_HEP
SMU_HPC	UCR-HEP
Tufts_ATLAS_Tier3	umd-cms
UJ-OSG	UMissHEP
	uprm-cms
	Vanderbilt
	Omaha

US LHC management has stated *"It is vital to the LHC program that the present level of service continue uninterrupted for the foreseeable future, and that all of the services and support structures upon which the LHC program relies today have a clear transition or continuation strategy."*

Satellites

CI-TEAM

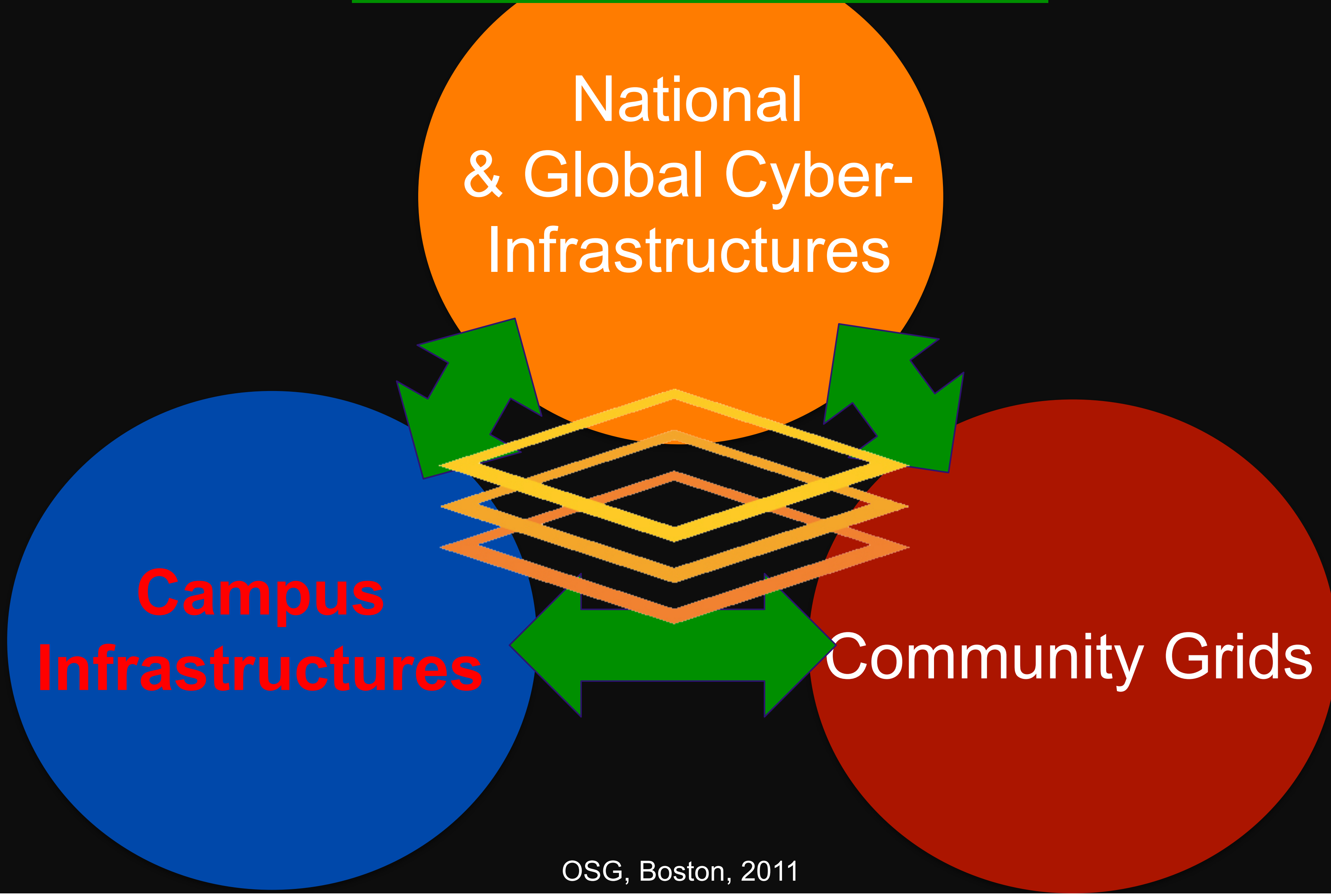
CorralWMS

High Throughput Parallel Computing (HTPC)

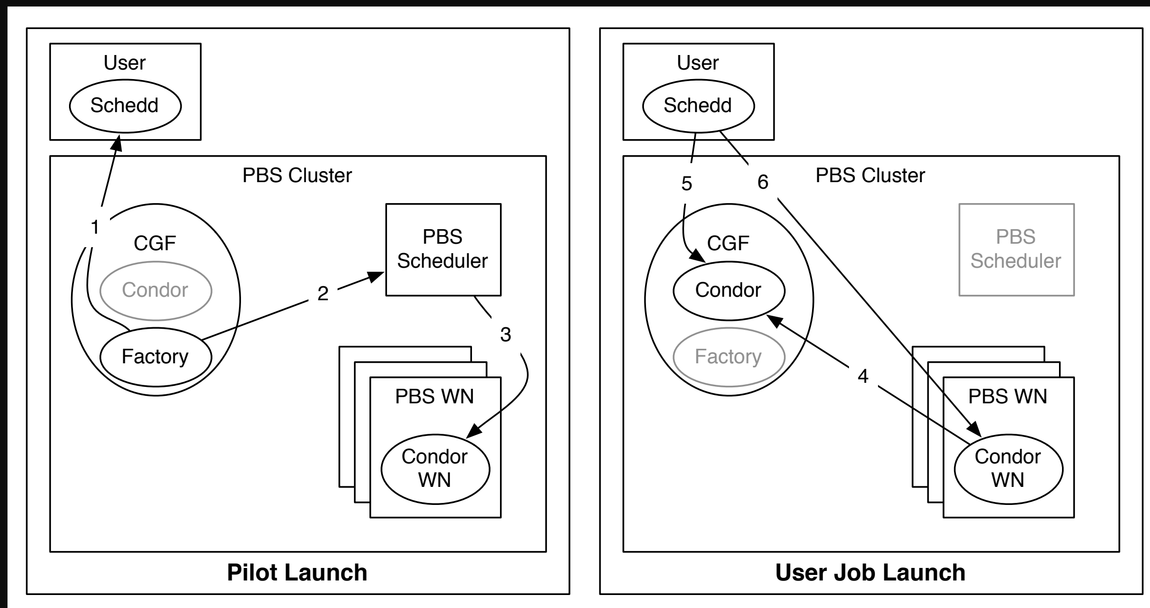
Advanced Network Intiative

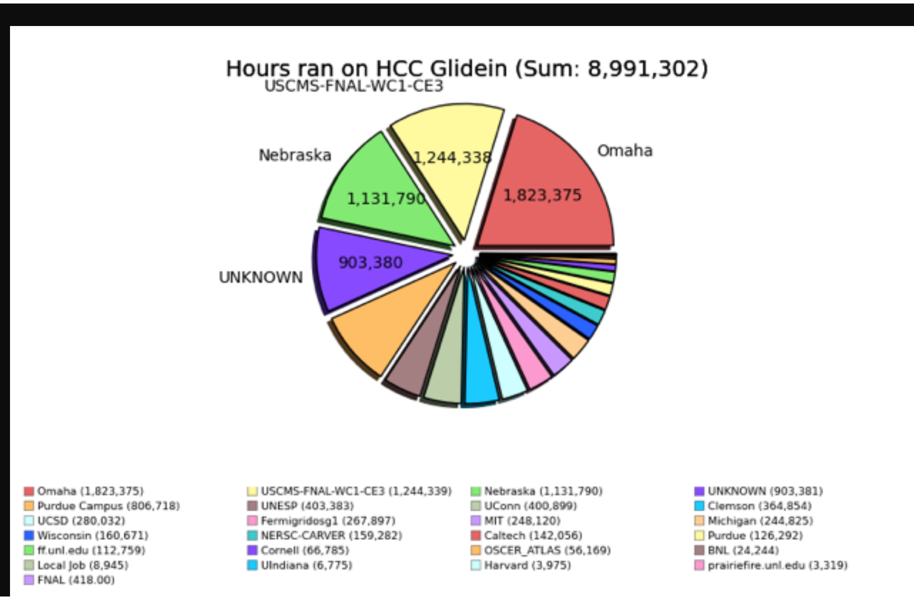
ExTENCI

Federated Autonomous CyberInfrastructures



Campus Overlays





Sustain and Extend

Short term

National directions

Proposals

Short Term

DOE HEP 1 Year Continued support for US LHC

March 2011-March 2012

Extend – don't just Sustain

Workshops..

Advanced Networks

Software Institutes

Campus Bridging

WLCG data management jamboree

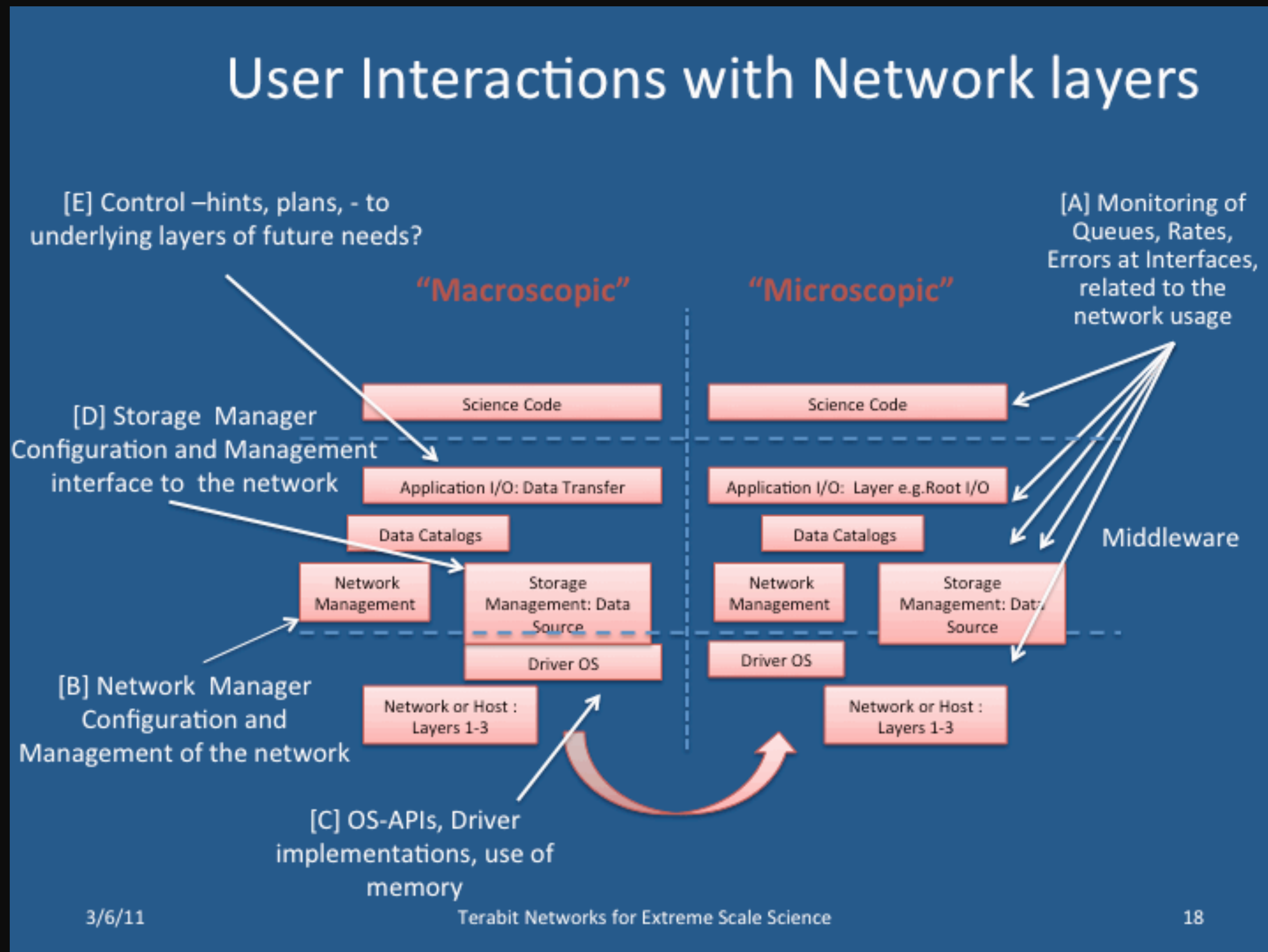
Etc

New Networks Coming!

LHCONE

100 Gigabit

Terabit….



Report from the Workshops on Distributed Computing, Multidisciplinary Science, and the NSF's Scientific Software Innovation Institutes Program

Miron Livny, Ian Foster, Ruth Pordes, Scott Koranda, JP Navarro

October 2010

3 Outcome – Recommendation 1

The most significant outcome of the workshops was the vision (and key attributes) for:

> "***A US Software Infrastructure Institute that provides a national center of excellence for community based software architecture, design and production; expertise and services in support of software life cycle practices; marketing, documentation and networking services; and transformative workforce***

The measure of success of such an institute should be the cost effectiveness (as measured in scholarly work) of our software infrastructure. This will be accomplished by a thriving and innovative software infrastructure community anchored by the Software Infrastructure Institute. The Institute will play a unique role by addressing organizational and life cycle elements not covered by infrastructure implementation and deployment projects that are driven by scientific objectives or technological trends.

4 Supporting Outcomes

To implement that vision attendees made the following more specific recommendations:

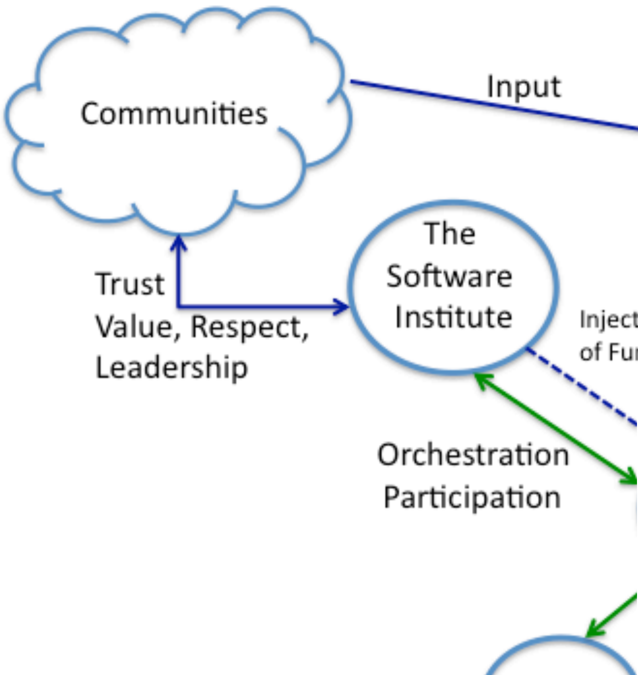
Recommendation 2. The Software Infrastructure Institute, together with its associated centers of excellence, should be **structured as an (Virtual) Organization** that provides a coherent and coordinated suite of high quality and dependable services that address the entire life cycle of distributed software infrastructure.

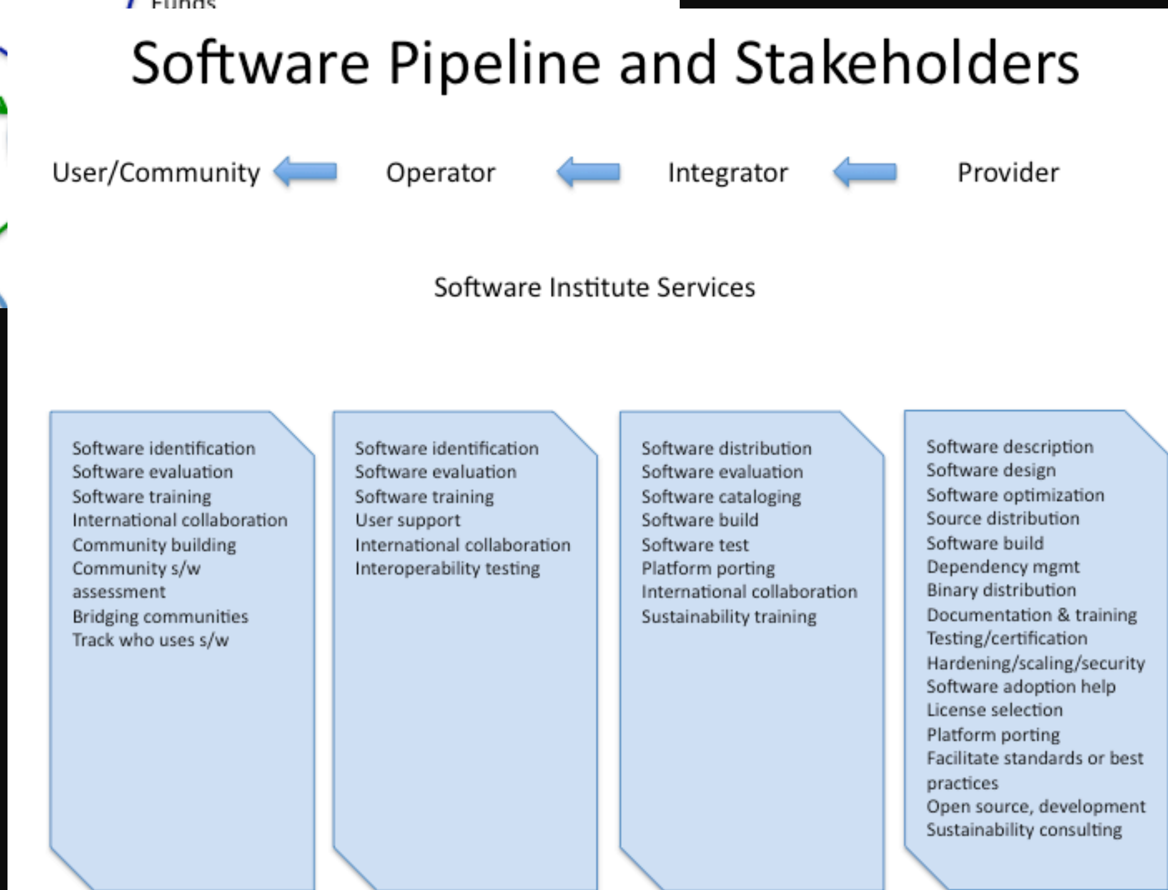
Recommendation 3. Services that advance the quality, adoption, and the longevity of the distributed software infrastructure should be provided by **teams of experts located at institutions with demonstrated leadership** in the areas of the service and strong commitment to the sustainability of the service. While we envision a distributed multi-layer structure for these services, we argue that these services need to be managed by a well-organized and highly respected team that can provide and sustain leadership in the complex and rapidly evolving area of distributed software infrastructure.

Recommendation 4. Quality must be the guiding principal for the services provided by the SII with accountability that is based on **independent quantitative impact assessment**. At all levels, allocation of effort and resources must be based on a professional and transparent ranking of impact and cost.

Recommendation 5. The funding model of the Institute must allow **quick turn-around time for requests to fund short-term projects** that address critical deficiencies in the Software Infrastructure that powers our science and research enterprise. The Institute must be able to quickly direct effort to a critical need. The Institute should adopt the metaphor of "supercomputer centers" in terms of assignment and allocation of the human resources needed to deliver the services to the DSI community.

Recommendation 6. The scope of the Institute must be to **offer software life cycle expertise and services** for distributed software infrastructure for a broad range of NSF programs, technology developers, and academic communities. The Institute must aim to improve the "accountability" in our software infrastructure enterprise by providing a recognized center of


Communities
Input
NSF
The Software Institute
Trust
Value, Respect, Leadership
Injection of Funds
Project Funds
Orchestration Participation


Software Pipeline and Stakeholders
User/Community
Operator
Integrator
Provider
Software Institute Services
Software identification
Software evaluation
Software training
International collaboration
Community building
Community s/w assessment
Bridging communities
Track who uses s/w
Software identification
Software evaluation
Software training
User support
International collaboration
Interoperability testing
Software distribution
Software evaluation
Software cataloging
Software build
Software test
Platform porting
International collaboration
Sustainability training
Software description
Software design
Software optimization
Source distribution
Software build
Dependency mgmt
Binary distribution
Documentation & training
Testing/certification
Hardening/scaling/security
Software adoption help
License selection
Platform porting
Facilitate standards or best practices
Open source, development
Sustainability consulting

Campus Bridging, CF21, CIF21

http://pti.iu.edu/campusbridging/

- The NSF should establish a national CI software roadmap. Through the Software Infrastructure for Sustained Innovation (SI^2) or other programs, the NSF should seek to systematically fund the creation and ongoing development and support of a suite of critical cyberinfrastructure software that identify and establish this roadmap, including CI software for authentication and access control; computing cluster management; data movement; data sharing; data, metadata, and provenance management; distributed computation / cycle scavenging; parallel computing libraries; network performance analysis / debugging; VO collaboration; and scientific visualization.
- The NSF should continue to invest in campus cyberinfrastructure through programs such as the Major Research Infrastructure (MRI) program, and do so in ways that achieve goals set in the Cyberinfrastructure Vision for 21st Century Discovery and a national CI software roadmap.

Strategic Recommendation to the NSF #3: The NSF should create a new program funding high-speed (currently 10 Gbps) connections from campuses to the nearest landing point for a national network backbone. The design of these connections must include support for dynamic network provisioning services and must be engineered to support rapid movement of large scientific data sets.

Strategic Recommendation to the NSF #4: The NSF should fund national facilities for at least short-term storage and management of data to support collaboration, scientific workflows, and remote visualization; management tools should include support for provenance and metadata. As a complement to these facilities and in coordination with the work in Recommendation #3, NSF should also fund the development of services for bulk movement of scientific data and for high-speed access to distributed data stores. Additionally, efforts in this area should be closely coordinated with emerging campus-level data management investments.

Strategic Recommendation to the NSF #5: The NSF should continue research, development, and delivery of new networking technologies. Research priorities funded by the NSF should include data intensive networks, sensor nets, networking in support of cyberphysical systems, geographically distributed file systems, and technologies to support long distance and international networking.

Strategic Recommendation to the NSF #6: The NSF should fund activities that support the evolution and

1st Tactical Recommendation

Tactical Recommendation to the NSF #1:The NSF should fund the TeraGrid eXtreme Digital program, as currently called for in existing solicitations, and should continue to fund and invest in the Open Science Grid.

Tactical recommendation to the NSF #2: The NSF should commission a study of current reward structures and recommendations about the reward structure – particularly as regards promotion and tenure for faculty – that would better align reward structures as perceived by individual faculty members with the type of large, collaborative virtual organizations that the NSF asserts are required for successful approaches to pressing, large scale scientific problems and transformative research.

Tactical Recommendation to the NSF #3: The NSF should support joint efforts with organizations such as the Association for Computing Machinery (ACM), the IEEE Computer Society, and/or Computing Research Association (CRA), to develop and maintain curriculum materials for undergraduate education in computer science and computational and data-driven science and engineering.

In its management of all of these programs, the NSF should make use of the Findings and Recommendations of this report and relevant Task Force on Campus Bridging workshop reports.

Recommendations to university leaders and the US higher education community

Strategic Recommendations to university leaders and the US higher education community

Strategic Recommendation to university leaders and the US higher education community #1: Institutions of higher education should lead efforts to fund and invest in university-specific, state-centric, and regional cyberinfrastructure to create local benefits (in research accomplishment and local economic development) and to aid the global competitiveness of the US and thus the long-term welfare of US citizens.

Strategic Recommendation to university leaders and the US higher education community #2: Every institution of higher education should have a plan, developed and endorsed at the highest level of its governance, for the establishment of a coherent cyberinfrastructure. Such a plan should have as one of its features a strategy for maximizing effective utilization of the institution's aggregate research

Security Our Future..

Proposal to NSF This Week for 5 Year Program of work

We present our plans to sustain and extend the OSG services for the next five years,

to transform the science and research computing landscape on our campuses through wide adoption of a new generation of DHTC technologies that support access to "any data, anytime, anywhere",

to an expanded set of job and data services via a single identity, and enable the transformation of our core stakeholders computing capabilities from petascale to exascale science.

Annual Goals in Sustaining the OSG:

a) Increase in CPU usage that is more than Moore's Law showing growth to meet the scaling needs of the users.

b) Full compliance with the operations service SLAs.

c) Doubling of accounted data movement to meet growth in data intensive science.

d) >=2 additional communities using DHTC services in production; >2 tutorials and documentation for new capabilities.

Transform computing on campuses through new DHTC technologies:

Year 1:

a) Deploy technology to account usage of users, jobs and data movement to campuses.

b) Release campus infrastructure software distribution Production Version 1.

Year 2:

a) Assess metrics that encapsulate a measure of adoption and increase in usage.

b) 20% increase in each of new, and usage by existing, users of campus technologies.

c) 2 production versions of campus software to extend the capabilities and scalability.

Year 3:

a) 20% increase in each of new, and usage by existing, users of campus technologies.

b) Transition of appropriate campus support services to sustained OSG operations.

Transformation of our core communities computing capabilities to exascale science:

Year 1: a) 10% of resources support end-to-end capability to be schedulable as HTPC and simultaneously usable and available by HTC/ single processor job.

Year 2: a) 20% of resources support end-to-end capability to be schedulable as HTPC and simultaneously usable and available by HTC/ single processor job.

b) 90% of LHC workload and 10% of non-LHC workload supports remote I/O capabilities.

c) Data movement across the DHTC fabric of >750 Petabytes/ year.

Year 3: a) >50% of resources support end-to-end capability to be schedulable as HTPC and simultaneously usable and available by HTC/ single processor job.

b) 50% of non-LHC workload supports remote I/O capabilities.

c) Data movement across the DHTC fabric of >750 Petabytes/ year

Annual Goals in Sustaining the OSG:

a) Increase in CPU usage that is more than Moore's Law showing growth to meet the scaling needs of the users.

b) Full compliance with the operations service SLAs.

c) Doubling of accounted data movement to meet growth in data intensive science.

d) >=2 additional communities using DHTC services in production; >2 tutorials and documentation for new capabilities.

Transform computing on campuses through new DHTC technologies:

Year 1: a) Deploy technology to account usage of users, jobs and data movement to campuses.

b) Release campus infrastructure software distribution Production Version 1.

Year 2: a) Assess metrics that encapsulate a measure of adoption and increase in usage.

b) 20% increase in each of new, and usage by existing, users of campus technologies.

c) 2 production versions of campus software to extend the capabilities and scalability.

Year 3: a) 20% increase in each of new, and usage by existing, users of campus technologies.

b) Transition of appropriate campus support services to sustained OSG operations.

Transformation of our core communities computing capabilities to exascale science:

Year 1: a) 10% of resources support end-to-end capability to be schedulable as HTPC and simultaneously usable and available by HTC/single processor job.

Year 2: a) 20% of resources support end-to-end capability to be schedulable as HTPC and simultaneously usable and available by HTC/single processor job.

b) 90% of LHC workload and 10% of non-LHC workload supports remote I/O capabilities.

c) Data movement across the DHTC fabric of >750 Petabytes/year.

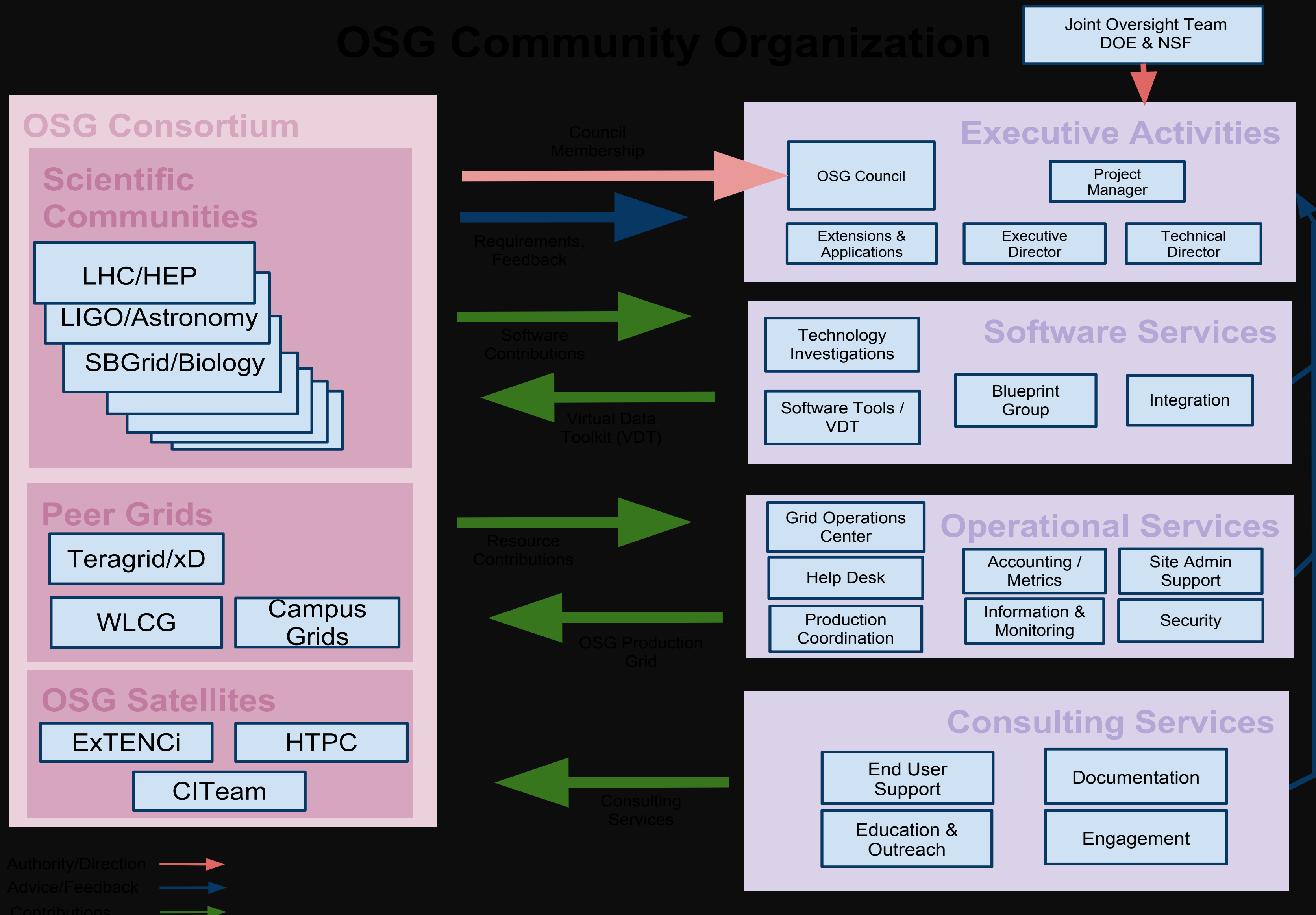
Year 3: a) >50% of resources support end-to-end capability to be schedulable as HTPC and simultaneously usable and available by HTC/single processor job.

b) 50% of non-LHC workload supports remote I/O capabilities.

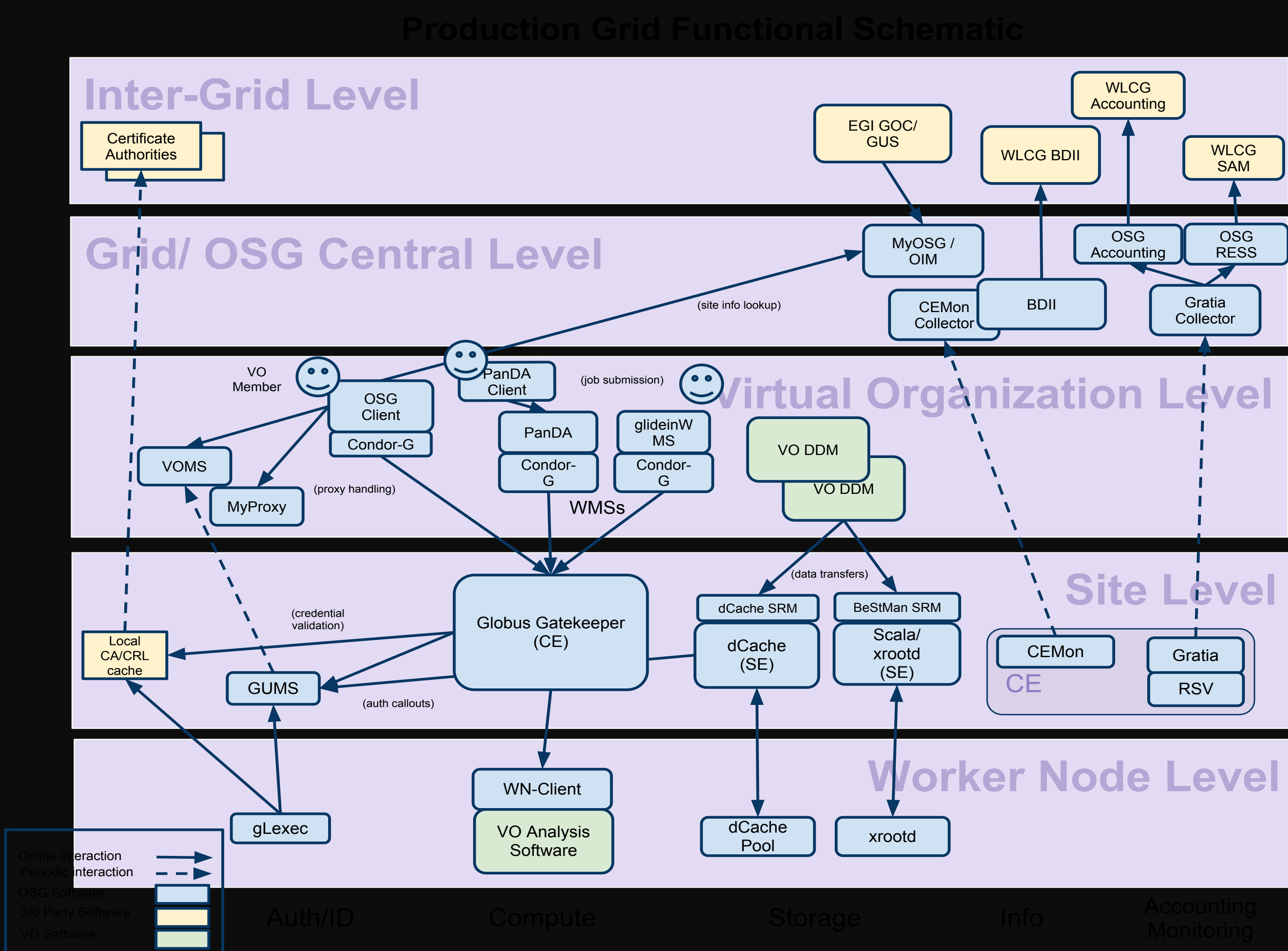
c) Data movement across the DHTC fabric of >750 Petabytes/year

Access to an expanded set of job and data services accessible via a single identity:

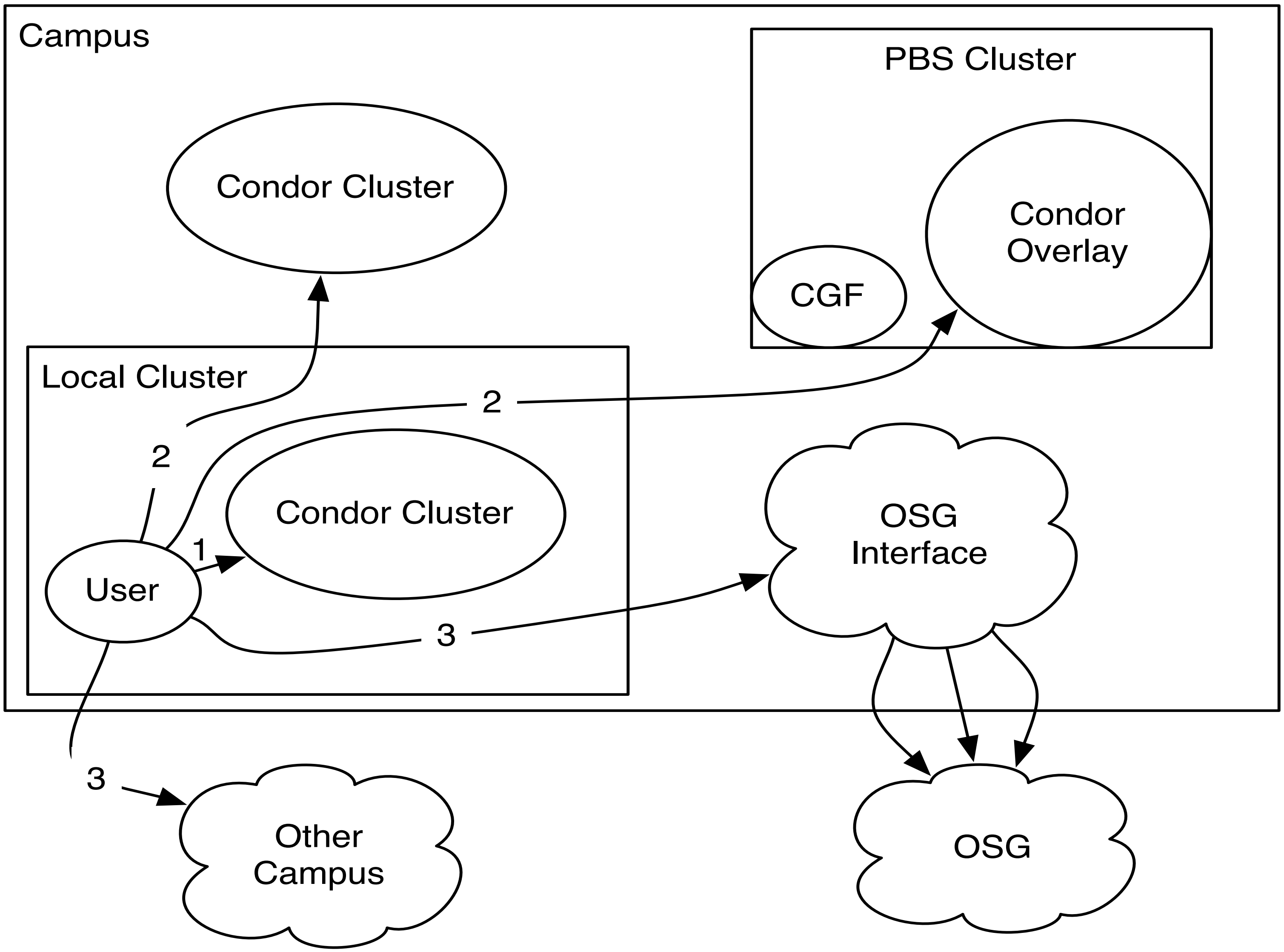
Check the new Blueprint - Brian Bockelman, John Hover



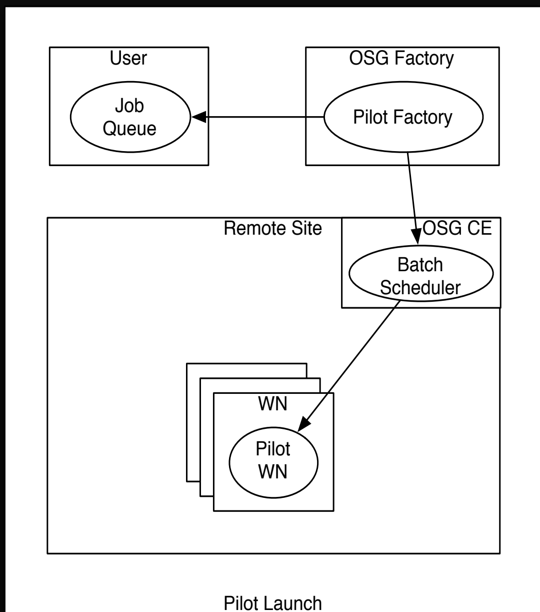
Production Grid

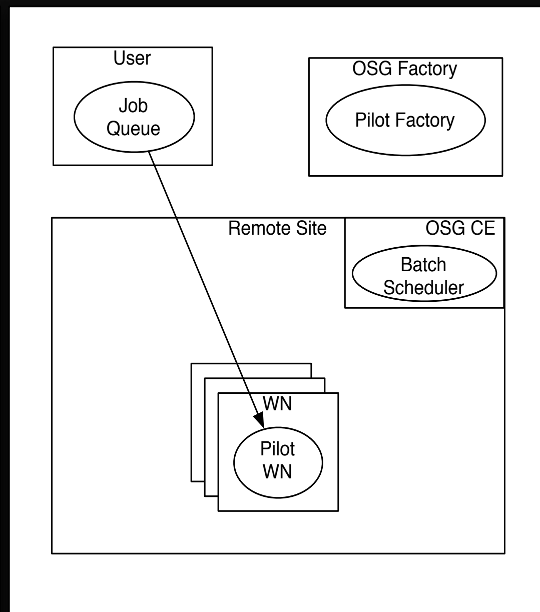


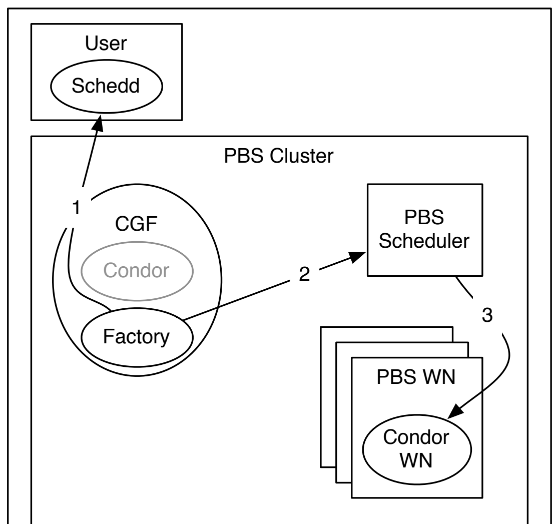
Campus Infrastructure



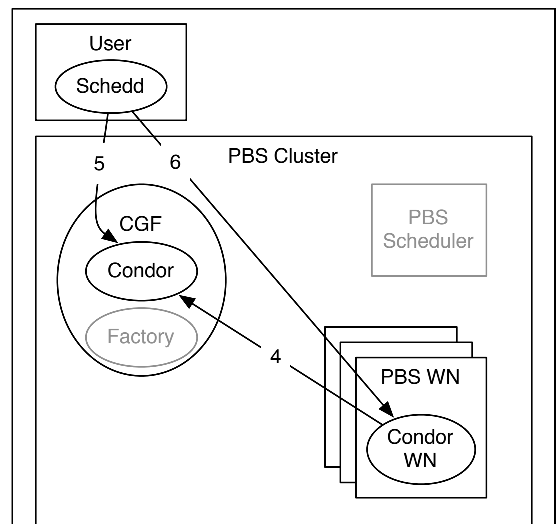
Pilots/Glideins
On Production and On the Campus







Pilot Launch



User Job Launch