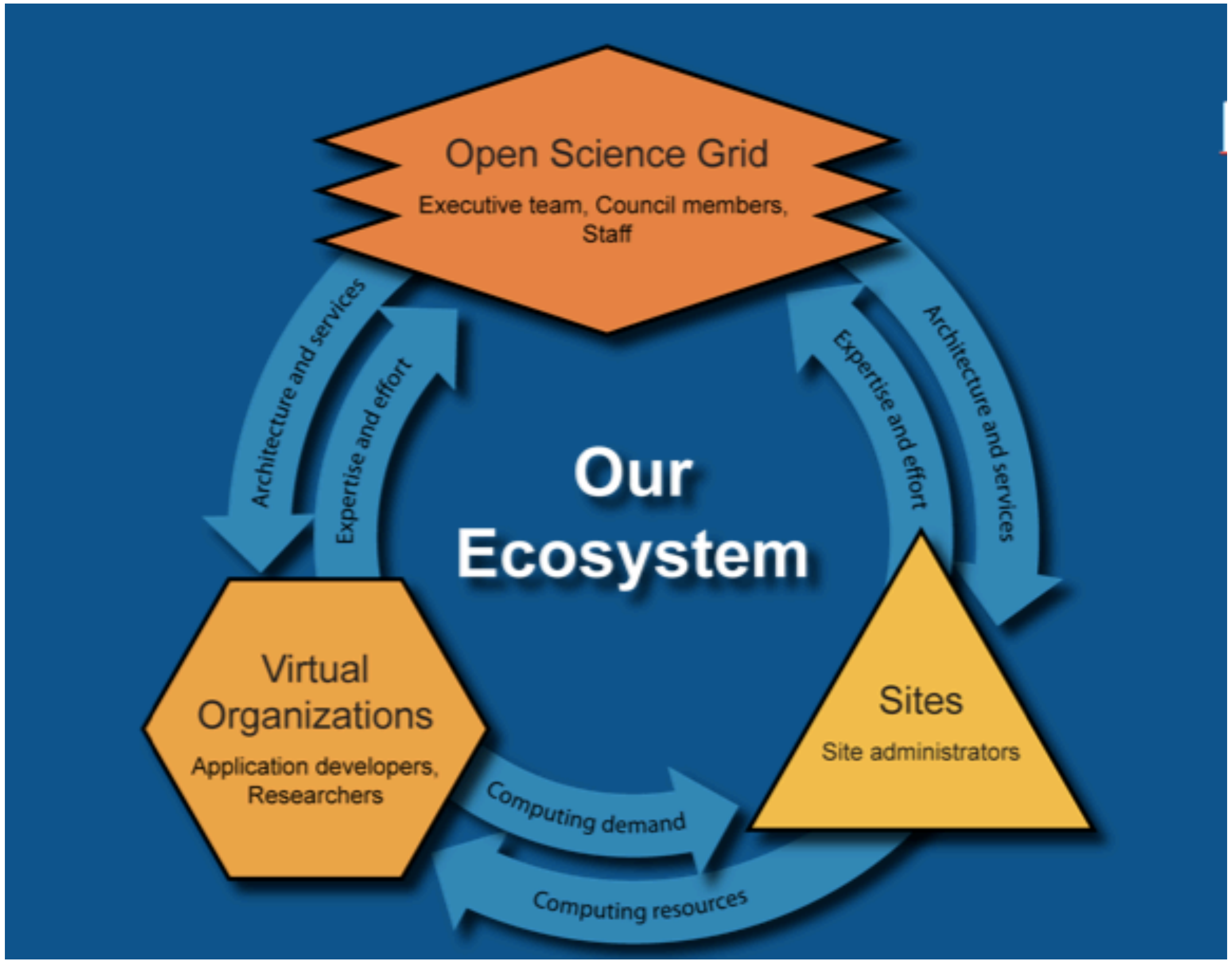
Open Science Grid

LATBauerdick/Fermilab

The OGS Ecosystem

Mission: *The Open Science Grid aims to promote discovery and collaboration in data-intensive research by providing a computing facility and services that integrate distributed, reliable and shared resources to support computation at all scales.*

- ✦ OSG Consortium
  - ★ sites/resources providers, science communities, stakeholders
- ✦ OSG Project
  - ★ staff, deliverables, operations
- ✦ Satellite Projects
  - ★ extensions, loosely coupled



OSG Resources

- Resources accessible through the OSG are contributed by the community.
  - Their autonomy is retained.
  - Resources can be distributed locally as a campus infrastructure

- >100 sites
- >70,000 cores accessible
- >30 research communities

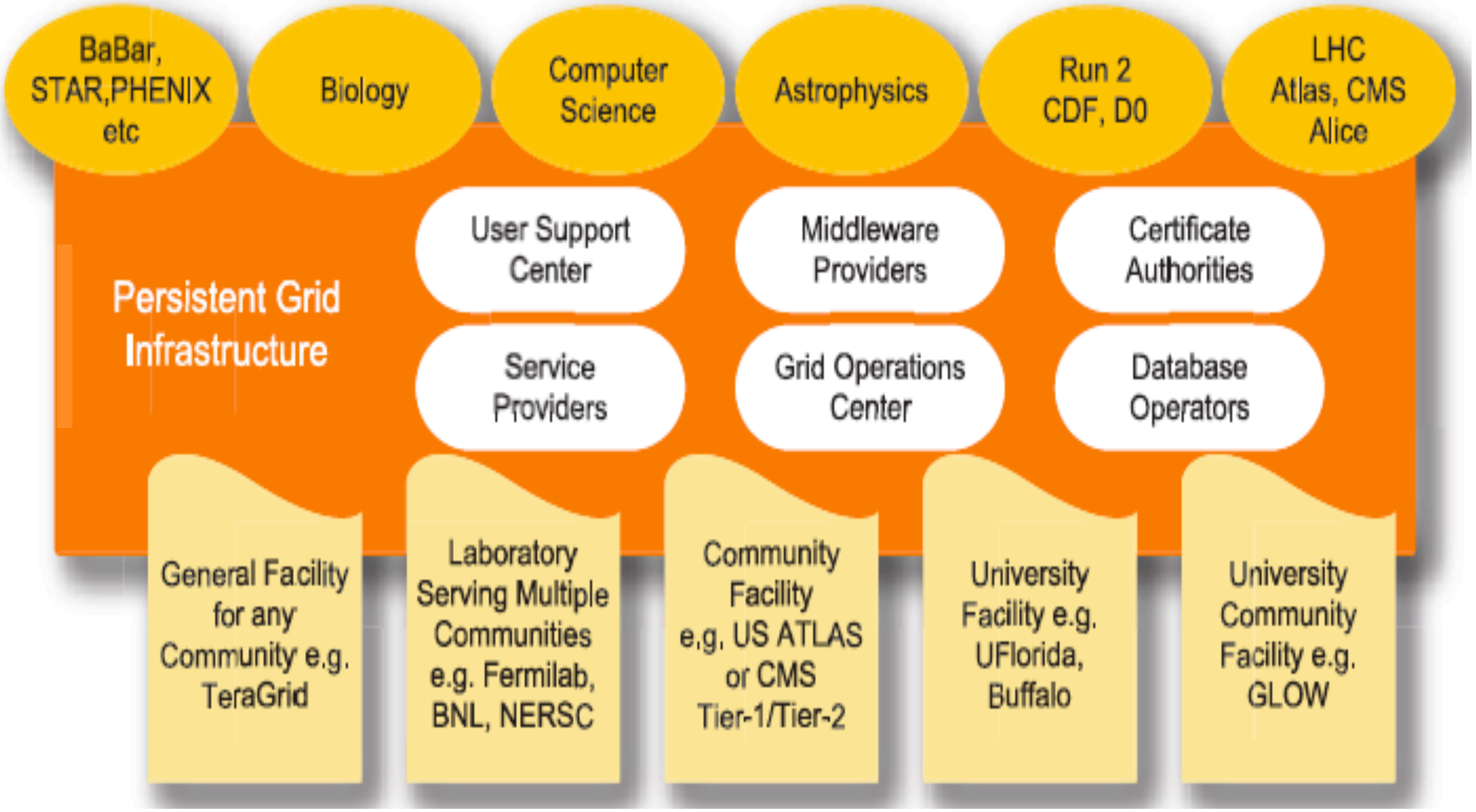
OSG Provides Common Services

Open Science Grid

Virtual Organizations

Common Services & Software

Resources: Sites, Campuses, Clouds

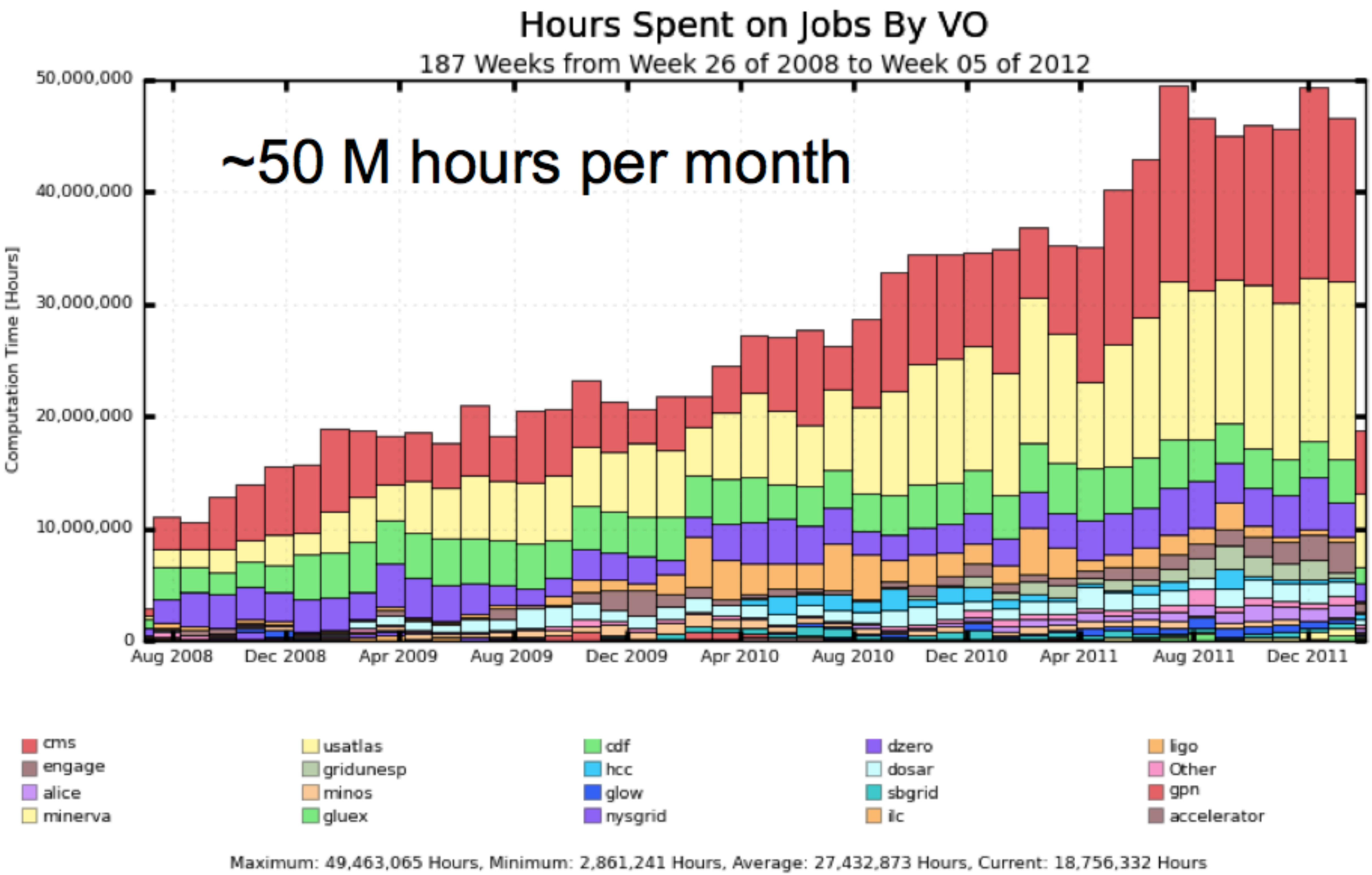


- ★ OSG services and infrastructures run by Operations and Security teams
- ★ OSG Software is packaged, tested, distributed as RPMs through the VDT
- ★ OSG User Support and Engage incubator to get new users started
- ★ OSG Blueprint Document defines architecture and principles

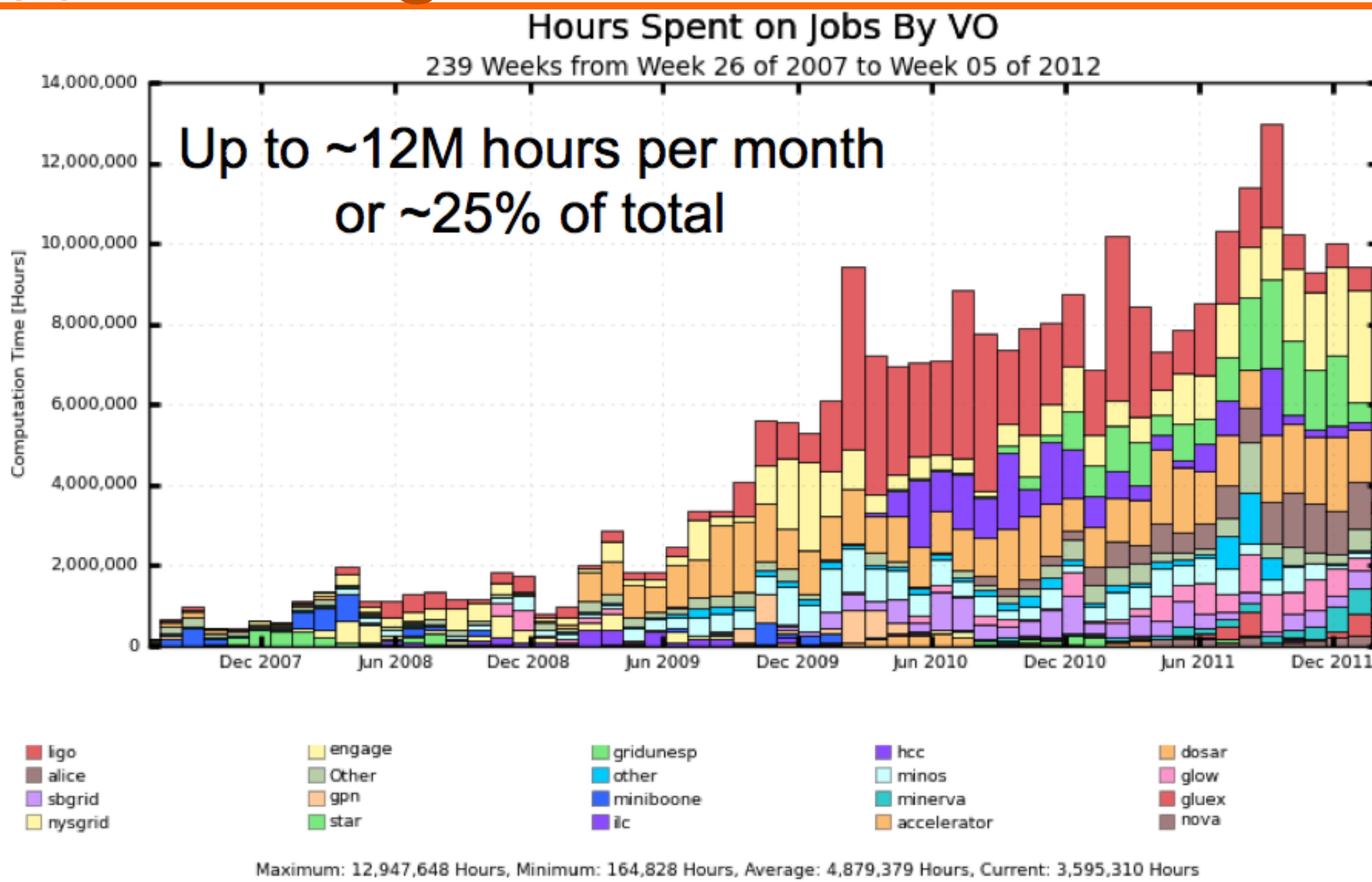
OSG Continuing for Another 5 Years

- We successfully made the case for continuing the OSG!
  - Very strong support from funding agencies, across DOE and NSF
  - Need continued contributions from stakeholders and success of satellite projects — continue to engage pushing satellites!
- Besides focus on physics and the huge momentum of the LHC, there is a broad spectrum of other science applications making use of OSG
  - Very significant NSF OCI funding means stronger emphasis on Campus Infrastructures, partnerships with other NSF funded projects: XSEDE
- Continue interfaces to peer infrastructures in Europe and elsewhere.
  - WLCG, EGI, national infrastructures in UK, France, Italy, Germany etc.
  - Worldwide continue partnerships in South America, Asia and Africa
- OSG goal to extend to science communities that can profit from DHTC
  - very successful approach of “submit local, compute global”
  - scientists learn operate in local batch environment, then “overflow” into OSG
  - great number of sciences are potential DHTC users, thereby enabling new levels of research —several exciting examples exist already in OSG!

OSG Providing Computational Resources to Science

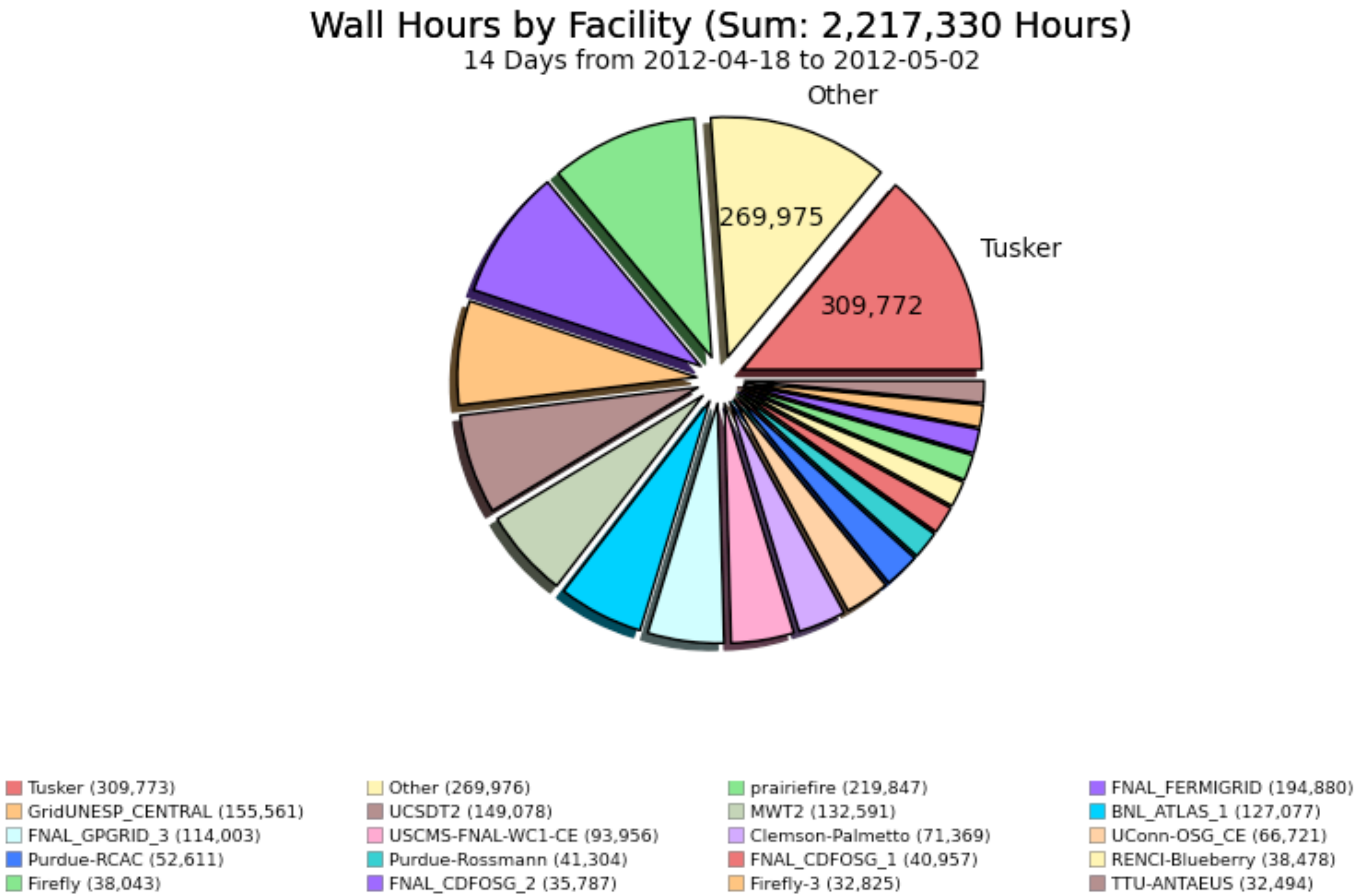


25% of CPU Cycles on OSG go to non-HEP!



- large fraction going to non-physics
- example: protein processing becoming a very large user of the OSG
  - e.g. at Nebraska HCC with CPASS, CS-Rosetta, and Autodock.

Opportunistic Use: Contributions from Sites



- (from Derek Weitzel, UNL http://derekweitzel.blogspot.com)

New opportunities with OSG and XSEDE working together

- XSEDE combines “high-end” services, in particular HPC centers
  - partners with OSG to allows XSEDE users access to the OSG HTC
- OSG became a XSEDE Service Provider
  - XSEDE science users get allocations through the XSEDE request and allocation process
  - OSG now part of that
  - 2,000,000 CPU hours on OSG allocated, 3 XSEDE allocation requests re-directed toward OSG
  - makes OSG DHTC resources available to HPC user community

Extending OSG across the Campuses

Open Science Grid

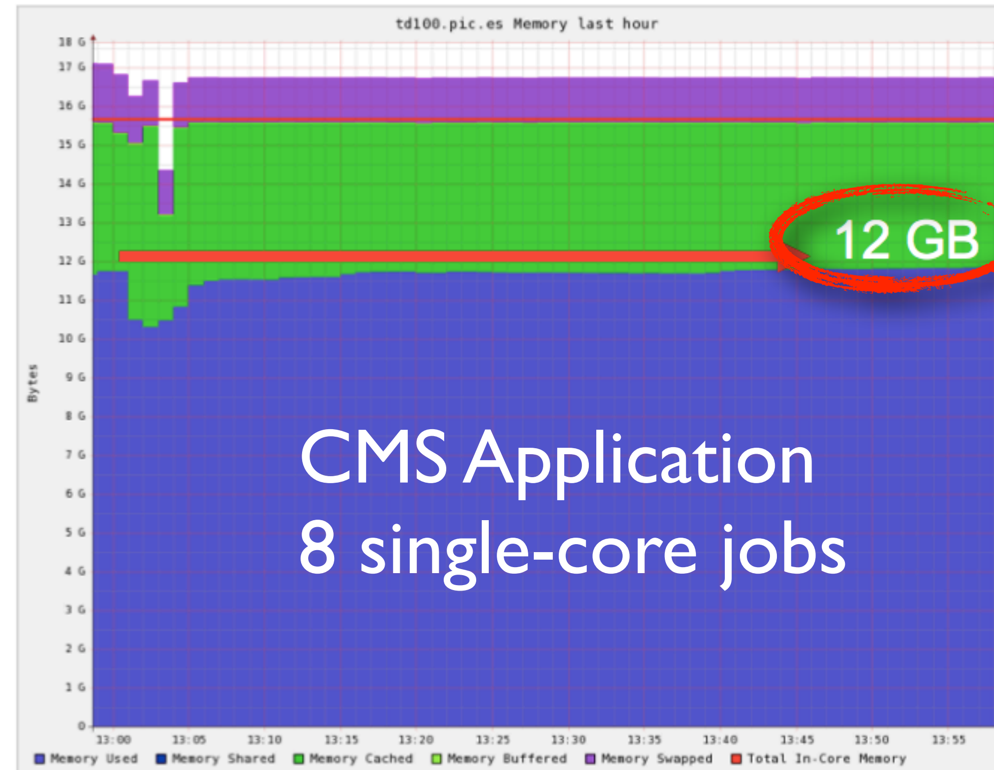
- Enable local cross-campus sharing of resources.
- Currently ~8 campus infrastructures – mostly LHC Tier-2s.
  - ★ US LHC Tier-3s are "beachheads" for their campus.
- software that will help users create a larger, more inclusive campus grid —> Dan's talk
- other challenges to be addressed, like federated ID management
  - ★ Moving in and out of campuses with a single identity

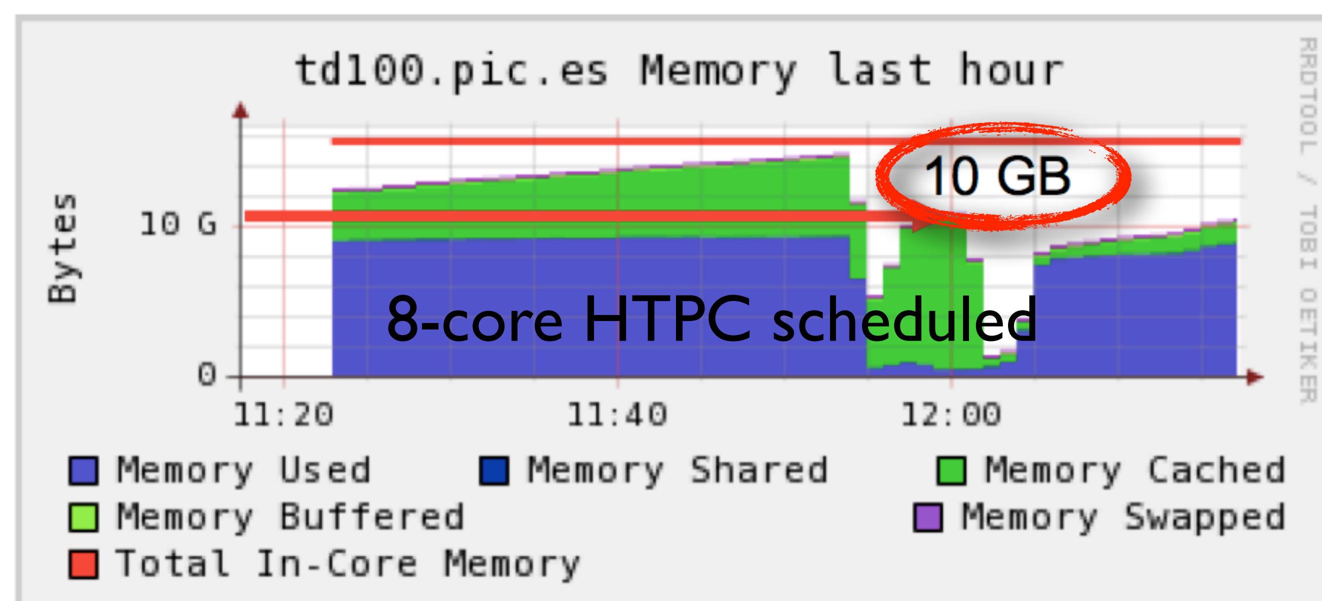
OSG Transition to "Overlay" Infrastructure

- Very successful transition to Job Overlay infrastructure (GlideinWMS)
  - ~50% of OSG cycles through glideinWMS, plus 25% through Atlas PanDA WMS
- tremendous boost to OSG usability, usage and effectiveness
  - Easier for VOs to access cycles, global queue to implement priorities
  - Site failures are not seen by the end user
  - Avoid direct grid submission overheads
- OSG "Glidein Factories" at IU, CERN and UCSD
  - factory receives requests from the VO frontend and submits glideins to requested sites using Condor-G
  - Knowledge how to submit to individual sites stored in factory configuration
  - Factory Operators perform routine maintenance on the Factory as well as monitor glideins to ensure they are running on sites without error
  - ~12 Communities served by OSG Glidein Factories
- Job Overlays provide avenue for flexible provisioning of compute resources
  - GlideinWMS releases support access to clouds, in particular Amazon EC2
  - e.g. purchased and open Cloud resources —need to solve policy issues

High Throughput Parallel Computing

- HTPC important for future applications on high-core-count nodes
  - already hugely important for LHC:
    - memory savings, future finer-grain parallelism etc
- HTPC being tested for OSG applications
  - CMS preliminary tests successful
  - test run w/ dynamic 8-core queue at Purdue
  - typical 20% memory savings critical for LHC application
- Condor capabilities in 7.7.8 will become very important
  - Automatic node draining, access to partial nodes
  - new capabilities need to trickle down to "production use cases"
  - OSG services need updating (information, testing, accounting)





Support For Data

- Current OSG services that support data – transport, access, archive, curation, provenance – across the end-to-end infrastructure are significantly deficient compared to the excellent support for processing!
  - in my mind, this potentially will be a big obstacle to scientific productivity
- The large communities have their own implementations/deployments of Data Services, new and small communities lack these.
  - the AAA ("any data, any time, anywhere") concept is a huge step forward, potentially also for non-LHC users: w/ AAA scientists have their data accessible everywhere they want to run via performant remote I/O!
- We still have to work out a strategy:
  - providing universal data access vs local storage provisioning
  - the AAA project could show the way towards generalizable Data Access/ Remote I/O services for use by other OSG communities
  - other Data Services – IRODS, Globus Online, Wide Area Lustre – will be usable on OSG and other resources to address particular needs.

OSG Eco-system at work:
New technologies enable new uses of OSG

- e.g. as reported by Dan Bradley of UW:
  - Isobel Ojalvo is a Wisconsin CMS student. She was in a crunch to get her analysis done. With the help of the [...] **AAA project**, Isobel was able to use the campus connection to the **glideinWMS system**, which wraps jobs (using the **parrot technology**) to give them remote access to CMS software using **CVMFS**. **Xrootd** provided remote access to the data to be analyzed. [...This way....] the Wisconsin campus and the OSG look much the same to Isobel.
  - Using this end-to-end solution, Isobel was able to immediately ramp up her resource usage **across 17 remote sites** on-the-fly, and **in a few hours, achieve the 24,000 hours of compute time** she needed. And, in addition, be the first CMS **opportunistic user of ATLAS Tier-2 sites** at the University of Chicago and Michigan.
  - **FermiGrid's squid server** [for cached database access] was confronted with a much higher rate of requests than previously sustained [...].
  - Progress all round!

CS Research Needs

- ★ OSG's existing capabilities are effective but basic and primitive
- ★ We need new levels of usability for portable, transparent services across an increasingly diverse and complex set of resource types, serving a broader range of science domains with growing diversity of scientific computing skills
- ★ The ongoing increases in size, complexity and diversity of both the science applications and the computer technologies present significant research and technical challenges.
- ★ Statically federated resources (Grids) need to be integrated with dynamically allocated resources (like clouds) causing new challenges for resource planning, acquisition and provisioning
- ★ OSG needs an influx of innovative frameworks and technologies in the areas of data, security, systems, workflows, tools and collaborative environments, working on issues like transparent usage, provisioning and management of resources, maximizing throughput and total benefit, improving robustness, managing identity information and trust, improving usability and integration
- ★ Innovation in these broad but inter-related areas can be only accomplished through a coordinated and collaborative CS research effort
- ★ see "Computer Science Research needed by the OSG" at http://bit.ly/Js0Qtk

OSG provides ideal deployment and testing opportunities for CS Researchers!

- Just one example: Network and Data Transfer Monitoring
  - ★ perfSonar-based network monitoring system between OSG sites
    - matrix of throughput or latency tests between each set of sites
    - history information collected in "dashboard database"
    - initiated by U.S. Atlas and I2, extended to all LHC sites and in future all OSG sites
  - ★ transfer-system level monitoring of success rates etc
    - assess success rate of transfers initiated between sites at the "middle-ware level"
    - data from FTS transfer system channels, or possibly GridFTP
  - ★ application-level monitoring of "transfer quality" between each set of sites
    - measures quality based on reliability of transfers and end-to-end throughput
    - based both on "production" data transfers and "debug" throughput measurements
- Through OSG, a wealth of information can be harvested
  - ★ by e.g. network provider for debugging or for SDN configuration,
  - ★ by network researcher, for novel scheduling, throughput optimization, ...?
    - e.g. can we learn from correlating network-level monitoring with performance of application-level data transfers, etc ...

perfSonar Monitoring: e.g. U.S. Atlas Tier-1 and Tier-2 sites

- Throughput (Gbps)

---	0	1	2	3	4	5	6	7	8	9
0:BNL (lhcmon.bnl.gov)	--- ---	0.86 0.85	0.38 0.86	0.78 0.78	0.83 0.88	0.75 0.51	0.86 0.90	0.71 0.73	0.49 0.60	0.75 0.46
1:AGLT2 (psmsu02.aglt2.org)	0.91 0.89	--- ---	0.94 0.94	0.93 0.89	0.93 0.83	0.93 0.94	0.91 0.91	0.92 0.84	0.69 0.58	0.90 0.65
2:AGLT2 (psum02.aglt2.org)	0.61 0.23	0.93 0.94	--- ---	0.21 0.29	0.91 0.90	0.68 0.66	0.17 0.14	0.18 0.15	0.00 0.09	0.36 0.28
3:MWT2 (iut2-net2.iu.edu)	0.91 0.90	0.92 0.85	0.92 0.92	--- ---	0.94 0.93	0.94 0.94	0.92 0.92	0.91 0.92	0.89 0.90	0.89 0.60
4:MWT2 (uct2-net2.uchicago.edu)	0.14 0.78	0.65 0.94	0.66 0.00	0.57 0.94	--- ---	0.66 0.94	0.44 0.66	0.43 0.72	0.00 0.91	0.51 0.00
5:MWT2-UIUC (mwt2-ps02.campuscluster.illinois.edu)	0.88 0.89	0.93 0.94	0.93 0.93	0.94 0.94	0.94 0.94	--- ---	0.92 0.00	0.93 0.93	0.93 0.00	0.90 0.00
6:NET2 (atlas-npt2.bu.edu)	0.92 0.84	0.62 0.59	0.00 0.60	0.00 0.91	0.61 0.70	0.00 0.92	--- ---	0.89 0.90	0.61 0.58	0.59 0.46
7:SWT2 (ps2.ochep.ou.edu)	0.82 0.87	0.92 0.93	0.92 0.92	0.46 0.61	0.90 0.88	0.93 0.93	0.91 0.91	--- ---	0.91 0.91	0.87 0.73
8:SWT2 (netmon2.atlas-swt2.org)	0.86 0.84	0.92 0.92	0.92 0.00	0.92 0.92	0.93 0.00	0.00 0.92	0.91 0.00	0.93 0.92	--- ---	0.89 0.62
9:WT2 (psnr-bw01.slac.stanford.edu)	0.55 0.84	0.74 0.76	0.75 0.89	0.00 0.88	0.81 0.00	0.00 0.68	0.45 0.21	0.80 0.00	0.72 0.23	--- ---

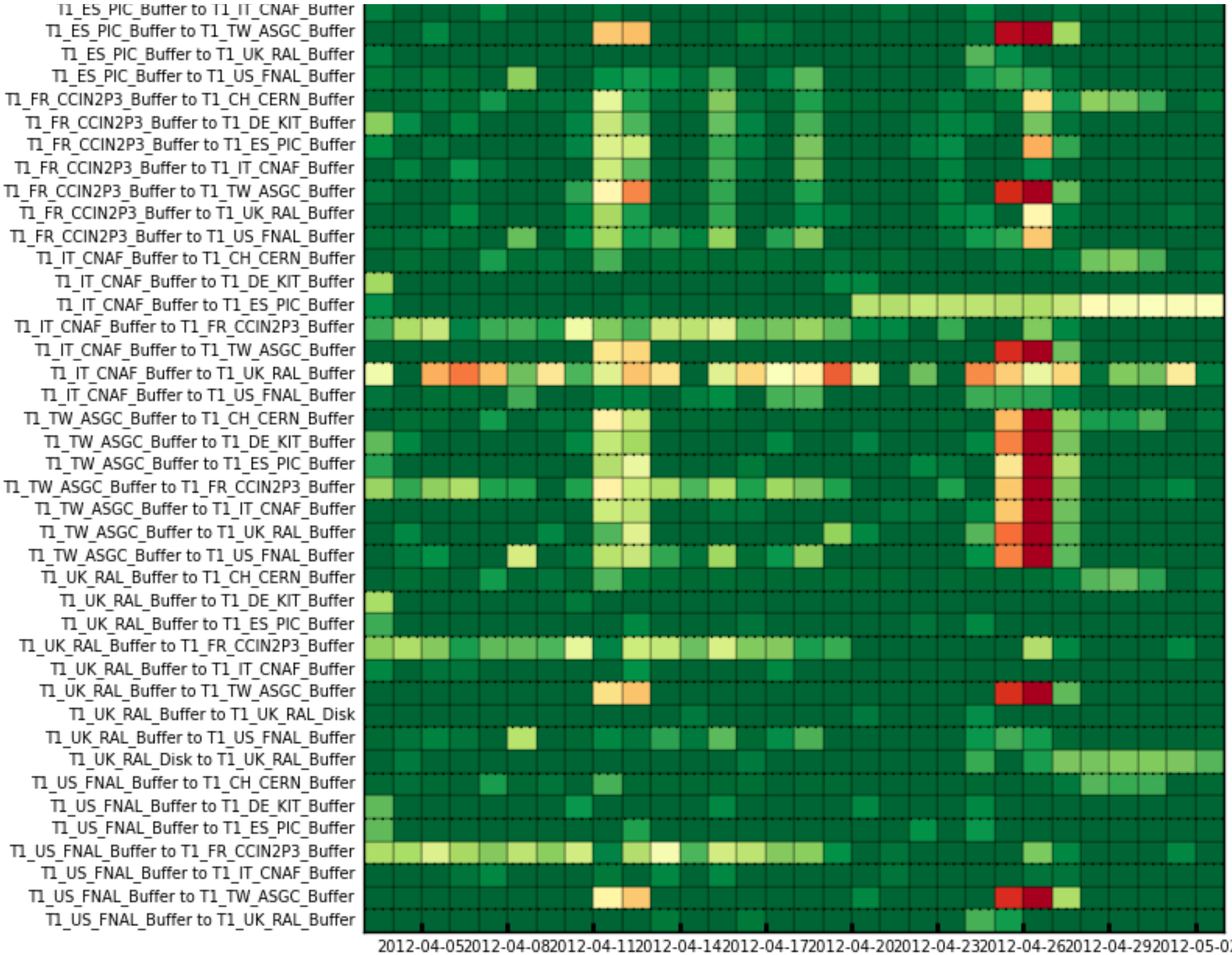
Latency, packet loss

---	0	1	2	3	4	5	6	7	8	9
0:BNL (lhcperfmon.bnl.gov)	0.0 0.0	0.0 0.0	0.0 0.0	0.0 0.0	0.0 1.0	0.0 0.0	0.0 0.0	0.0 0.0	0.0 0.0	0.0 0.0
1:AGLT2 (psmsu01.aglt2.org)	0.0 0.0	0.0 0.0	0.0 0.0	0.0 0.0	0.0 2.0	0.0 0.0	0.0 0.0	0.0 0.0	0.0 0.0	0.0 0.0
2:AGLT2 (psum01.aglt2.org)	0.0 0.0	0.0 0.0	0.0 0.0	1.0 0.0	0.0 2.0	0.0 0.0	0.0 0.0	0.0 0.0	0.0 0.0	0.0 0.0
3:MWT2 (iut2-net1.iu.edu)	0.0 0.0	0.0 0.0	0.0 0.0	0.0 0.0	0.0 1.0	0.0 0.0	0.0 0.0	0.0 0.0	0.0 0.0	0.0 1.0
4:MWT2 (uct2-net1.uchicago.edu)	0.0 4.0	8.0 8.0	8.0 9.0	9.0 0.0	0.0 0.0	9.0 8.0	13.0 8.0	12.0 11.0	8.0 11.0	9.0 0.0
5:MWT2-UIUC (mwt2-ps01.campuscluster.illinois.edu)	0.0 0.0	0.0 0.0	0.0 0.0	0.0 1.0	0.0 1.0	0.0 0.0	0.0 0.0	0.0 0.0	0.0 0.0	0.0 0.0
6:NET2 (atlas-npt1.bu.edu)	0.0 0.0	0.0 0.0	0.0 0.0	0.0 1.0	1.0 2.0	0.0 0.0	0.0 0.0	0.0 0.0	0.0 0.0	0.0 0.0
7:SWT2 (ps1.ochep.ou.edu)	0.0 0.0	0.0 0.0	0.0 0.0	2.0 1.0	0.0 0.0	0.0 0.0	0.0 0.0	0.0 0.0	0.0 0.0	0.0 0.0
8:SWT2 (netmon1.atlas-swt2.org)	0.0 0.0	0.0 0.0	0.0 0.0	1.0 0.0	2.0 3.0	0.0 0.0	0.0 0.0	0.0 0.0	0.0 0.0	0.0 0.0
9:WT2 (psnr-lat01.slac.stanford.edu)	0.0 0.0	0.8 0.0	0.0 0.0	1.0 0.0	3.0 1.0	0.0 0.0	0.0 0.0	0.0 0.0	0.0 0.0	0.0 0.0

FTS-level monitoring of transfer success rates, average and history

	TOTAL-	Austria+	Belgium+	Brazil+	China+	Estonia+	Finland+	France+	Germany	Hungary+	India+	Italy+	Poland+	Portugal+	Republic-	Russian-F	Spain+	Switzerlar	Taiwan+	Turkey+
TOTAL-	82 %	97 %	85 %	97 %	97 %	33 %	97 %	85 %	93 %	96 %	78 %	90 %	100 %	80 %	45 %	54 %	43 %	100 %	95 %	100
Austria+	92 %								100 %			100 %						100 %	100 %	
Belgium+	100 %								100 %			100 %						100 %	100 %	
Brazil+	91 %	100 %	100 %		100 %	0 %	100 %	80 %	100 %	100 %		100 %	100 %	100 %	100 %	60 %		100 %	100 %	
China+	100 %								100 %			100 %						100 %	100 %	
Estonia+	88 %	100 %	100 %	100 %	100 %		100 %	33 %	100 %	100 %		100 %	100 %	100 %		50 %		100 %	100 %	
Finland+	86 %	100 %	100 %	100 %	100 %	50 %		100 %	100 %	100 %		100 %	100 %	100 %	100 %	50 %		100 %	100 %	100
France+	83 %								46 %			59 %						97 %	100 %	
Germany+	88 %	100 %	100 %	100 %	100 %	38 %	100 %	71 %	100 %	100 %	100 %	100 %		100 %	0 %	100 %	100 %	100 %	100 %	
Hungary+	100 %								100 %			100 %							100 %	
India+	100 %							100 %				100 %						100 %	100 %	
Italy+	98 %	100 %	100 %	100 %	100 %	100 %	100 %	75 %	100 %	100 %		100 %	100 %	100 %	100 %	60 %	100 %	100 %	100 %	100
Pakistan+	86 %											100 %						100 %	100 %	
Poland+	100 %											100 %						100 %	100 %	
Portugal+	88 %								100 %			100 %						100 %	100 %	
Russian-Federation+	100 %								100 %			100 %						100 %	100 %	
Spain+	94 %								100 %			78 %						100 %	70 %	
Switzerland+	94 %	100 %	100 %	100 %	100 %	22 %	100 %	96 %	100 %	100 %	100 %	100 %	100 %	100 %	0 %	63 %	100 %		100 %	
Taiwan+	78 %	100 %			100 %	50 %	100 %	52 %	100 %		100 %	65 %		100 %	100 %	58 %	100 %	100 %	100 %	
Turkey+	93 %	100 %	100 %	100 %	100 %	0 %	100 %	50 %	100 %	100 %		100 %	100 %	100 %	100 %	60 %	100 %	100 %	100 %	
UK+	93 %	100 %	43 %	100 %	100 %	13 %	86 %	89 %	94 %	100 %	25 %	59 %	100 %	25 %	100 %	38 %	58 %	100 %	82 %	
USA+	75 %	86 %	100 %	92 %	88 %	36 %	100 %	91 %	93 %	67 %	100 %	95 %	100 %	100 %	88 %	48 %	32 %	100 %	95 %	100
Ukraine+	84 %	100 %	100 %	100 %	100 %	100 %	100 %	67 %	100 %	100 %		100 %	100 %		0 %	60 %		100 %	67 %	
n/a+	43 %	100 %	100 %	100 %	100 %	0 %	100 %	50 %	100 %	100 %		100 %	100 %	100 %	100 %	38 %	100 %	100 %	100 %	100

Application-Level end-to-end “Data transfer quality” and throughput



Closing Remarks

- The triumphant success of DHTC ideas at a huge scale in data-intensive science like at the LHC, and it's equally successful and flexible use for smaller applications proves we have something outstanding to offer to a very broad set of science communities
- OSG's unique strength is our presence in science and on the university campuses and at the labs
  - OSG has a broad presence on almost 100 campuses and close scientific collaboration of domain and computing experts
- With the renewed OSG project we will continue to provide an excellent environment that allows great science to happen across the OSG
- We are grateful for the outstanding support from the Condor community and would like to continue our very fruitful collaboration to address the next level of challenges!