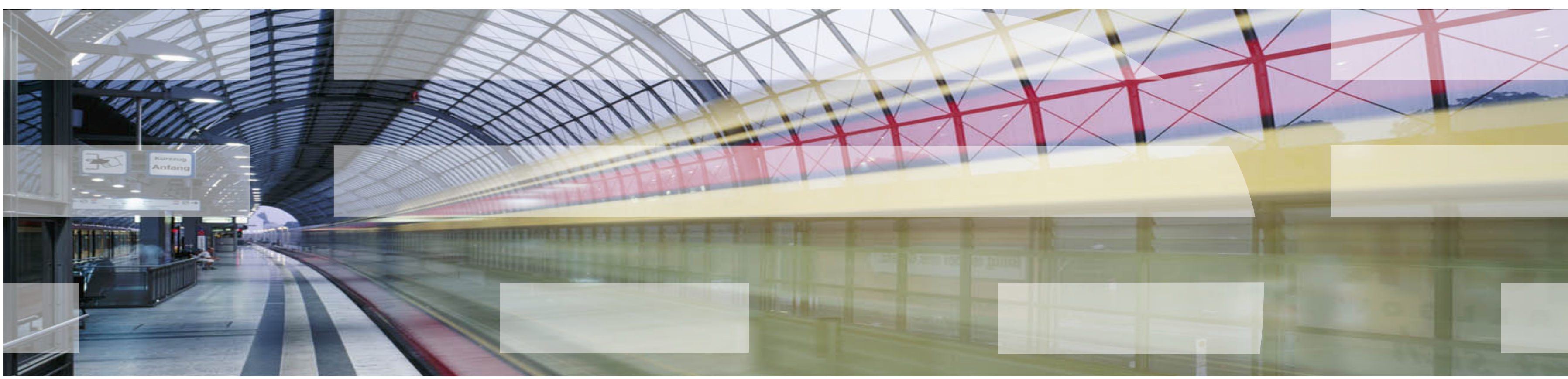
Connecting HPC and High Performance Networks for Scientists and Researchers

ESnet
ENERGY SCIENCES NETWORK

GÉANT

INTERNET2®

SC15 Austin, Texas
November 18, 2015



Agenda

- Welcome and introductions
- BoF Goals
- Overview of National Research & Education Networks at work Globally
- Discuss needs, challenges for leveraging HPC and high-performance networks
- HPC/HTC pre-SC15 ESnet/GEANT/Internet2 survey results overview
- Next steps discussion
- Closing and Thank You

BoF: Connecting HPC and High Performance Networks for Scientists and Researchers

- Goal:
  - Have an interactive conversation between HPC participants, the Research and Education (R&E) networking community, and the scientists and researchers we serve.
  - Work to solve problems and challenges scientists and researchers experience.

- Outcomes:
  - Have a continuing conversation between HPC, R&E networking community, and scientists and researchers to develop solutions.
  - Discover areas to develop best practices for serving the HPC and network organization end users: science and research collaborations.
  - Publish a report on our findings to be shared with the community.

INTERNET2 NETWORK INFRASTRUCTURE TOPOLOGY

OCTOBER 2014

INTERNET2 NETWORK BY THE NUMBERS

17	JUNIPER MX960 ROUTERS SUPPORTING LAYER 3 SERVICE
34	BROCADE AND JUNIPER SWITCHES SUPPORTING LAYER 2 SERVICE
62	CUSTOM COLLOCATION FACILITIES
250+	AMPLIFICATION RACKS
15,717	MILES OF NEWLY ACQUIRED DARK FIBER
8.8	TBPS OF OPTICAL CAPACITY
100	GBPS OF HYBRID LAYER 2 AND LAYER 3 CAPACITY
300+	CIENA ACTIVEFLEX 6500 NETWORK ELEMENTS
2,400	MILES PARTNERED CAPACITY WITH ZAYO COMMUNICATIONS IN SUPPORT OF THE NORTHERN TIER REGION

IN SUPPORT OF U.S.UCAN

NETWORK PARTNERS

ciena

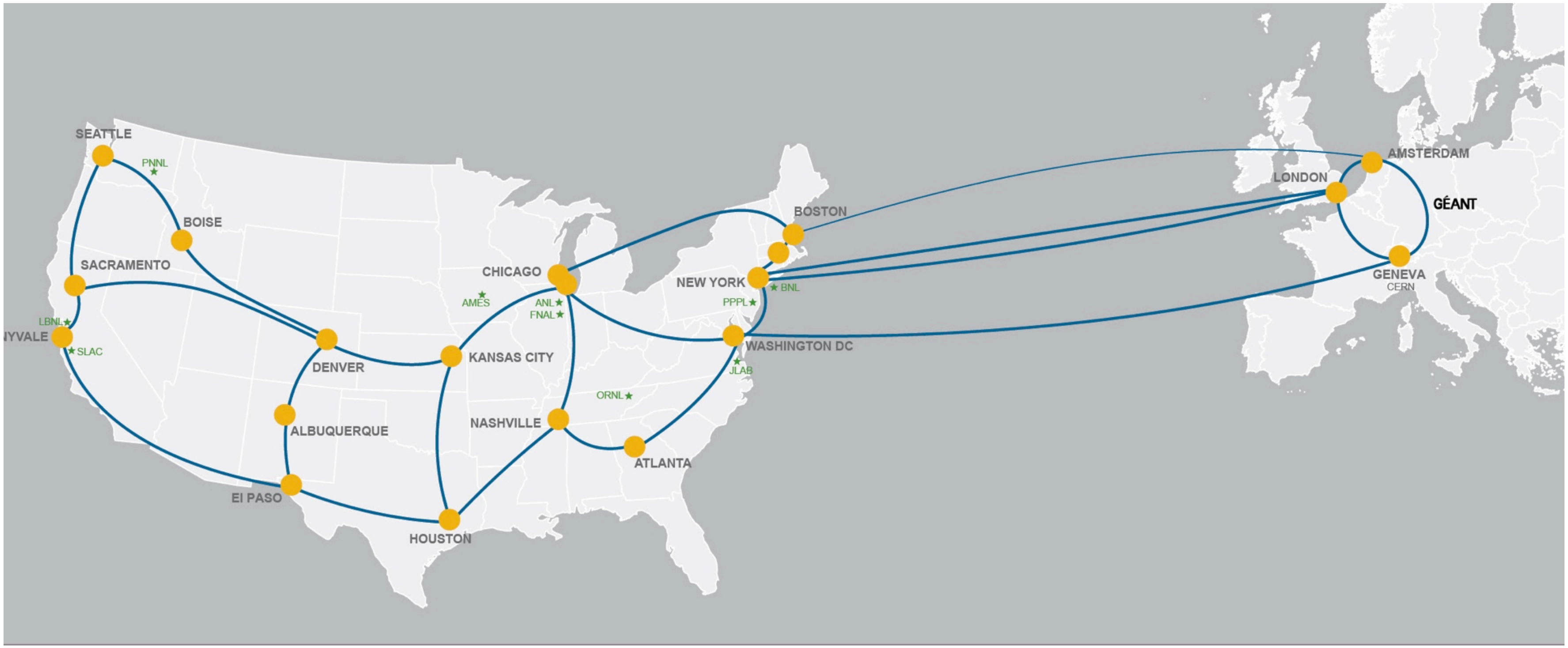
cisco

INDIANA UNIVERSITY

infinera

JUNIPER NETWORKS

Department of Energy's international research network

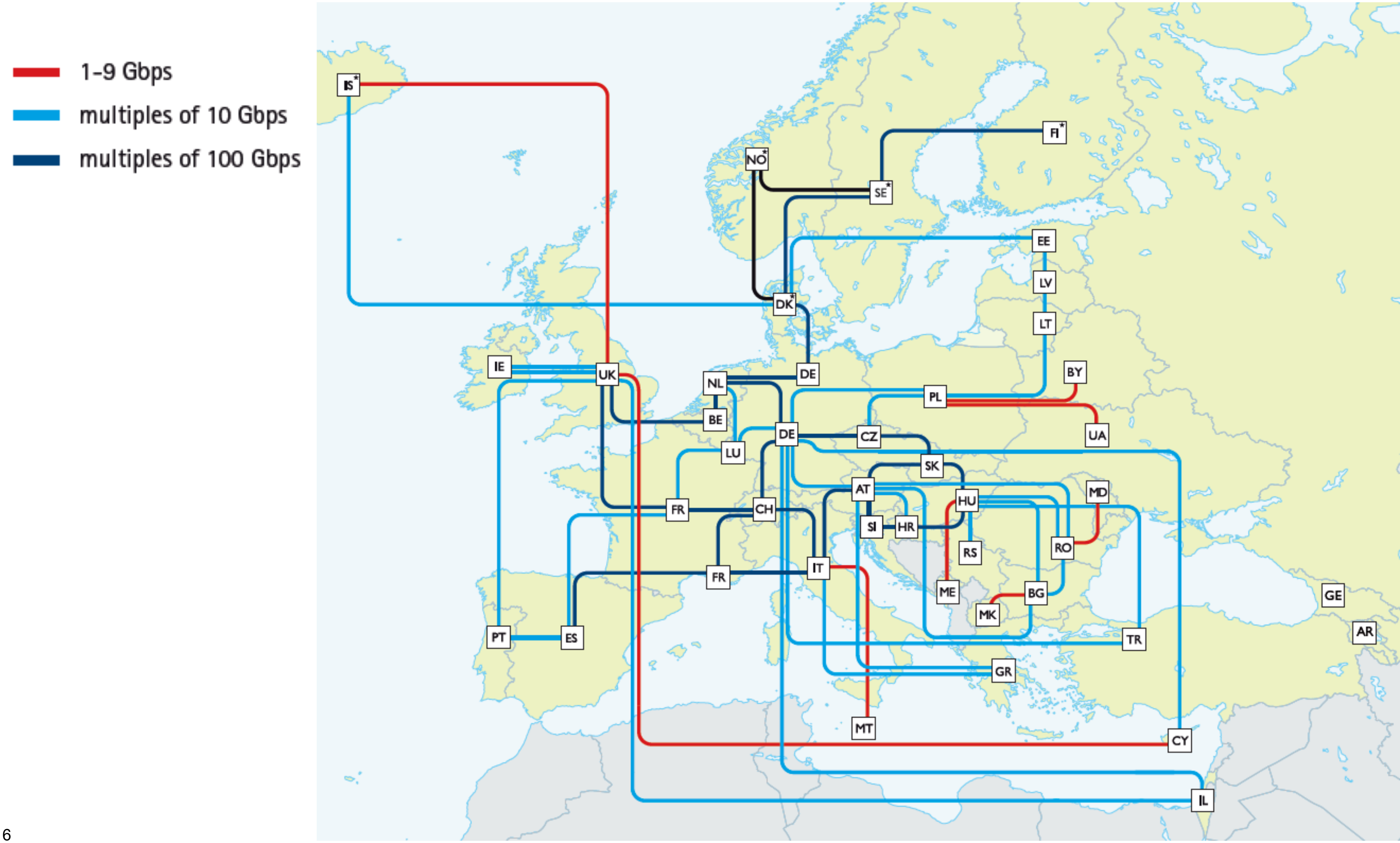


ESnet
ENERGY SCIENCES NETWORK

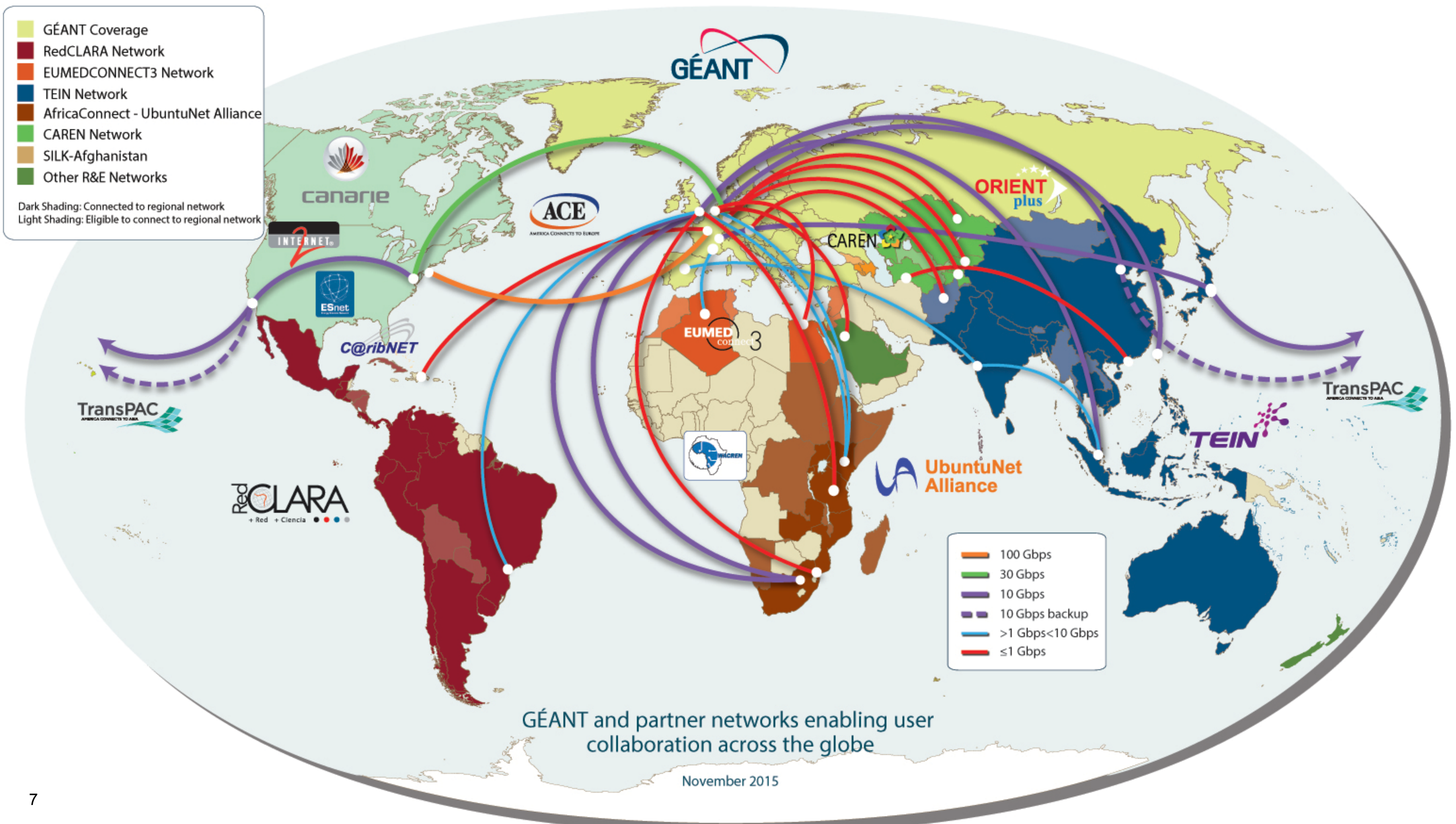
★ Department of Energy Office of Science National Labs

Ames	Ames Laboratory (Ames, IA)
ANL	Argonne National Laboratory (Argonne, IL)
BNL	Brookhaven National Laboratory (Upton, NY)
FNAL	Fermi National Accelerator Laboratory (Batavia, IL)
JLAB	Thomas Jefferson National Accelerator Facility (Newport News, VA)
LBNL	Lawrence Berkeley National Laboratory (Berkeley, CA)
ORNL	Oak Ridge National Laboratory (Oak Ridge, TN)
PNNL	Pacific Northwest National Laboratory (Richland, WA)
PPPL	Princeton Plasma Physics Laboratory (Princeton, NJ)
SLAC	SLAC National Accelerator Laboratory (Menlo Park, CA)

GÉANT European Topology

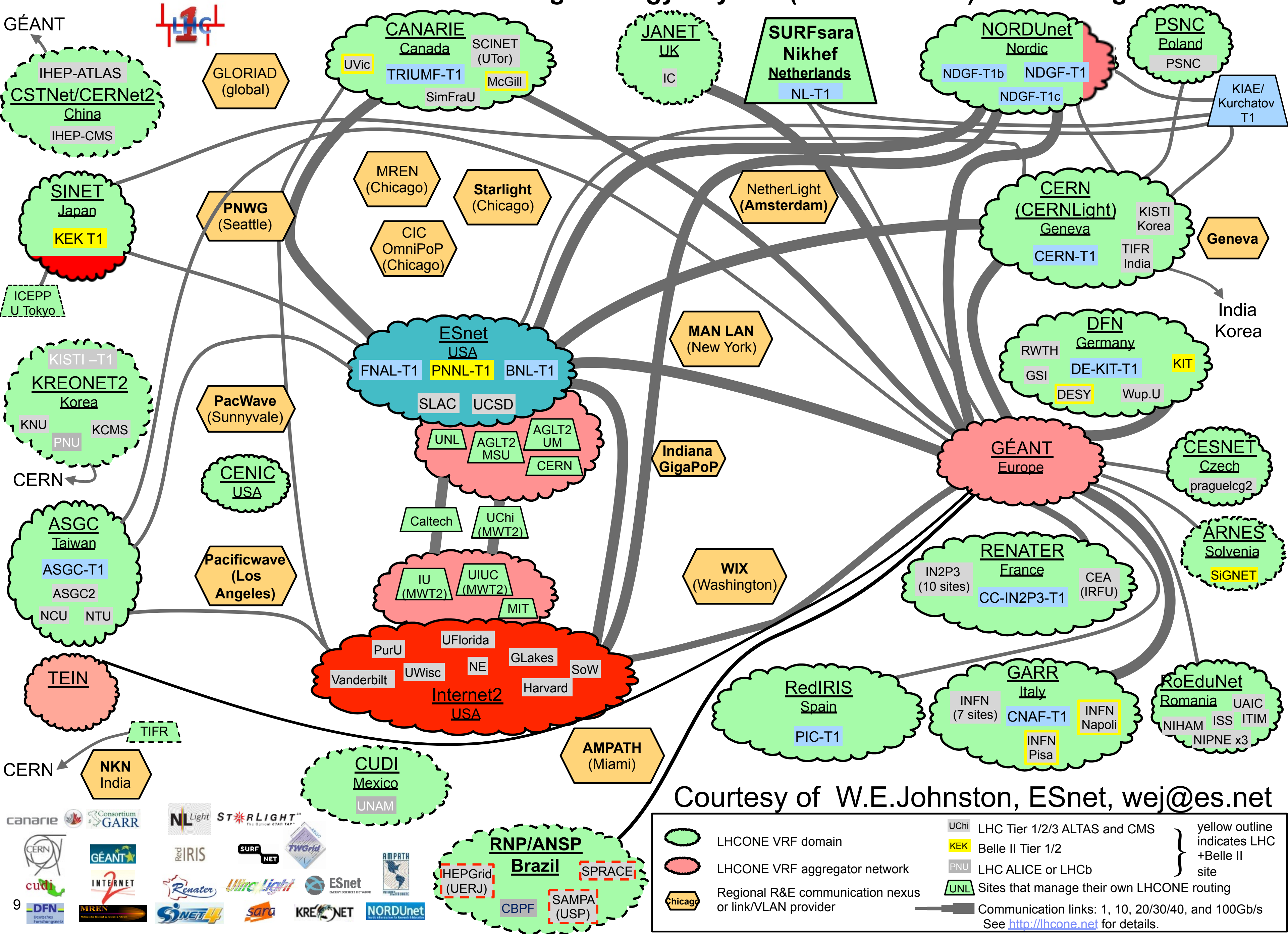


GÉANT / ESnet / Internet2 global connectivity



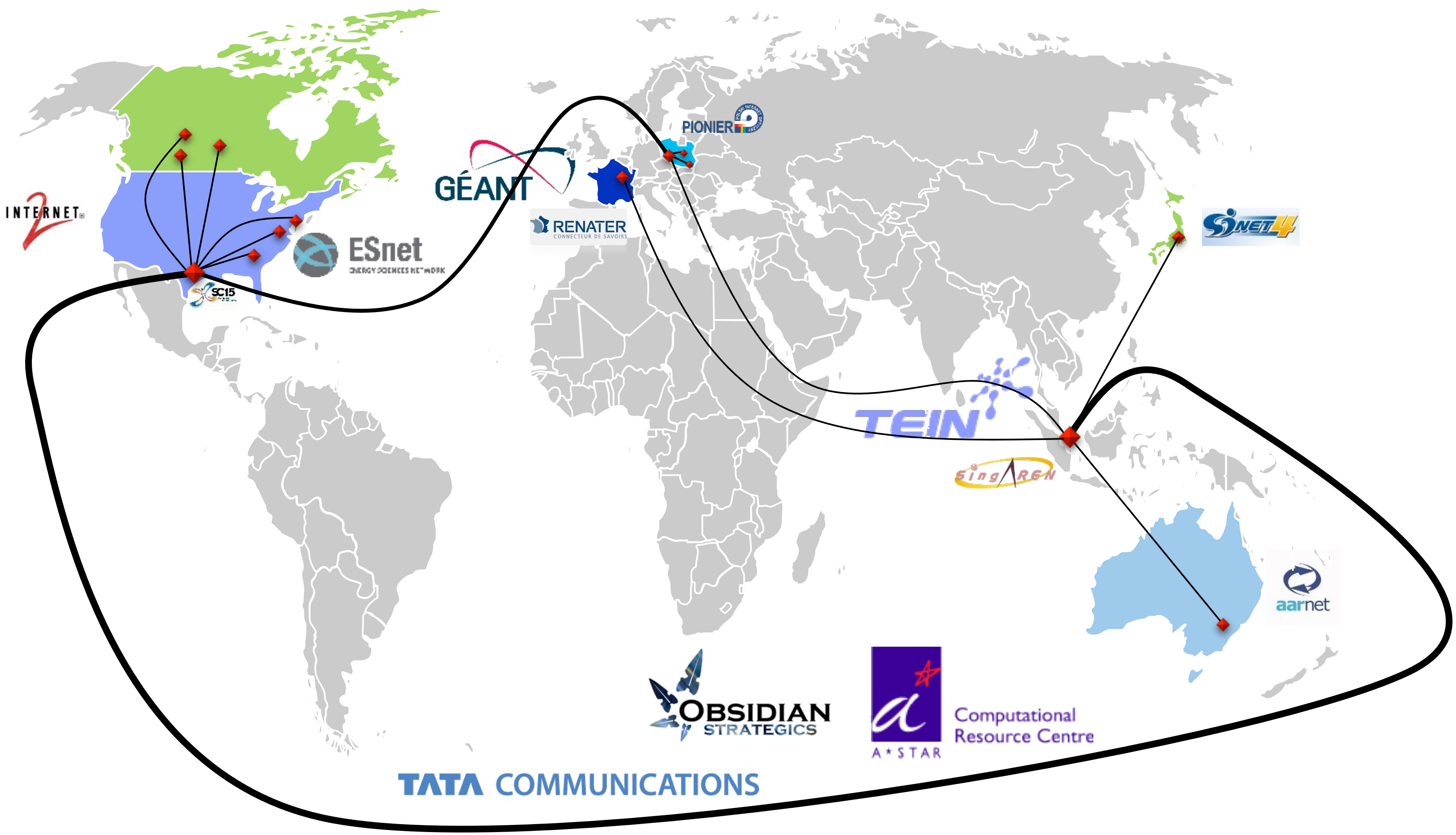
Examples from the R&E networking collaborations

LHCONE: Global infrastructure for High Energy Physics (LHC & Belle II) data management



GEANT reference : InfiniCortex

- InfiniBand over the WAN: connect HPC centers together to enable research at a global scale
- A “Galaxy of Supercomputers” scattered across multiple continents

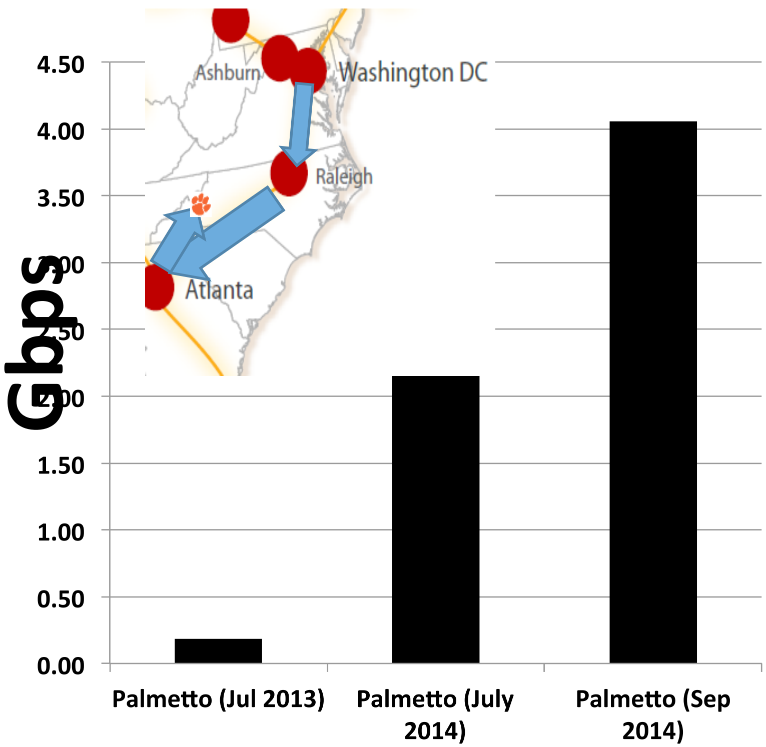


Optimizing Genomic Data Transfers Across Internet2

"The massive genomic data transfer rate increases across Internet2 enabled by a ***long list of collaborators*** *are helping my research because I can* **scale up my systems biology HPC workflows** *and* ***download raw data sets, process, and delete****, thus freeing up my very finite disk space allocation. Through an NSF award, we are extending our methods to genome databases."* - Dr. F. Alex Feltus, Associate Professor of Genetics & Biochemistry at Clemson University and CEO of Allele Systems LLC.

"We are seeing 38x data transfer speed improvements and getting better."

NCBI → CU::: 12 TByte DNA data in 5 hours and not 8 days



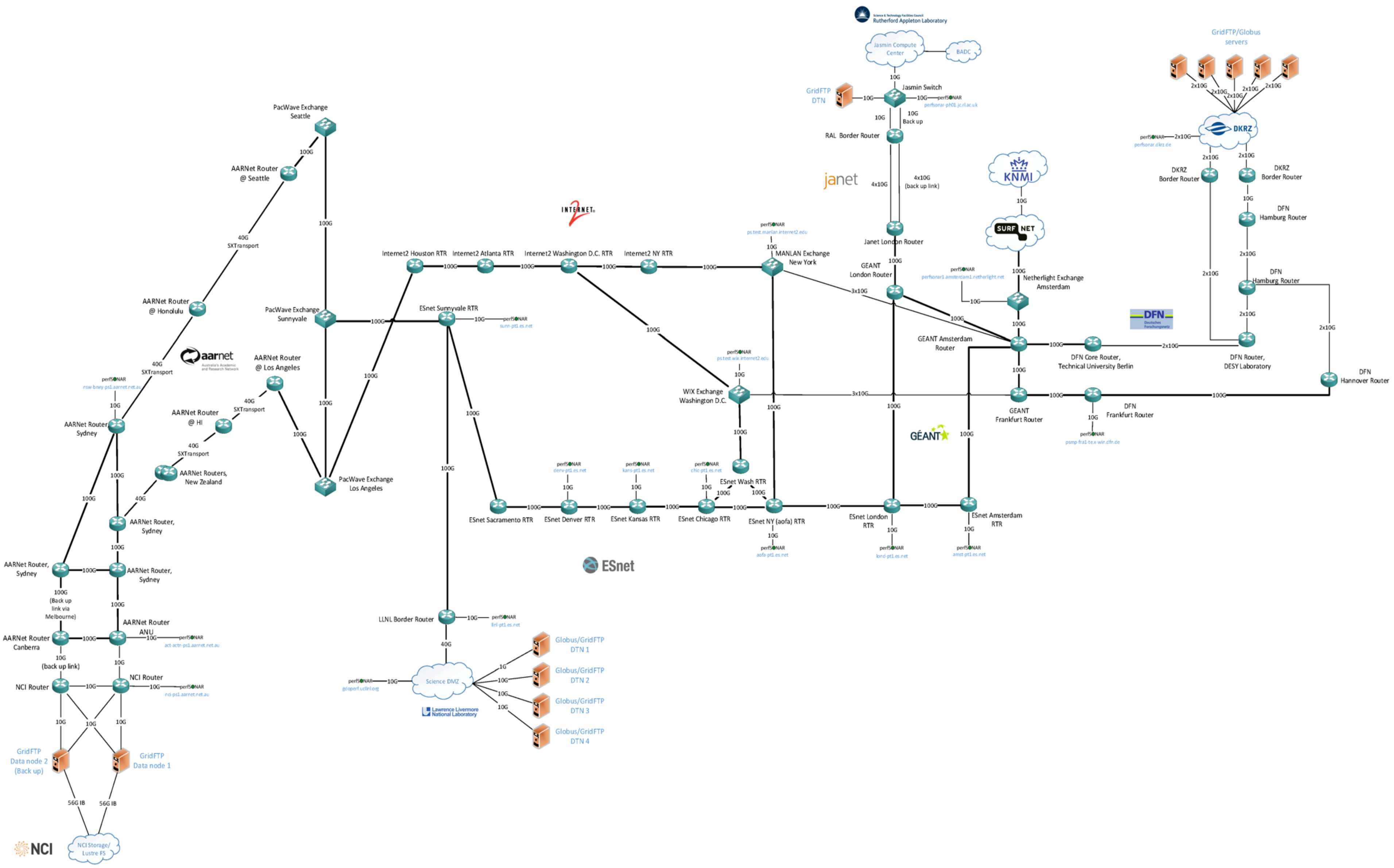
NSF DIBBS (1443030; S. Ficklin, PI) ::: Washington State, Clemson, UCONN, UT-Knox.

NCBI Tripal INTERNET2

Slide prepared by Alex Feltus, Clemson University

International Networking for Climate

International Networking for Climate



Agenda

- Welcome and introductions
- BoF Goals
- Overview of National Research & Education Networks at work Globally
- Discuss needs, challenges for leveraging HPC and high-performance networks
- HPC/HTC pre-SC15 ESnet/GEANT/Internet2 survey results overview
- Next steps discussion
- Closing and Thank You

Interactive discussion: Share needs & challenges & realities for researchers to leverage high-performance networks

- Low expectations of the network and lack of understanding of the state of the art of network
- HEP is pushing the boundaries, knowledge is high – digital divide with other domains.
- Element of culture
- Funding challenges
- Local, regional network cooperation – little grant support to do multi institutional collaboration (west va)
- Last mile problem – 100G connection to FLR, building on campuses have no fiber, no way to take advantage of 100G connection
- Brown – substantial network investment, communications breakdown
- ACI-REF – last mile problem = education, engage with researchers to enhance their science, how to have the conversation, how to elevate the conversation
- NOAA – convincing security team to open ports. . . ☺ cant get data in or out, or deploy science dmz
- What do you do if you want to put sensitive data in a science DMZ?
- Science DMZ – real world – constantly doing research on laptops and want to connect to research network, not sure what is on your laptop!
- The best advocates for networking are the scientists
- Attention span is so much , incentives need to be great
- What is the minimum amount of knowledge a scientist needs to be convinced of using new tools
- Technical challenge, no 100G NICs – communicating all the nerdy nobs and dials ☺

Interactive Discussion – Share best practices & success stories for research leveraging high-performance networks

- Embedding in a collaboration and viral nature of success stories
- If you want to have an impact on the science community, you need to have a team approach, tiered – domain scientists with computational science, interact with HPC professionals, in addition to embedding have to have team characteristics
- Central to notion of embedding is community building, make the people you embed as part of the community, one incentive – split the salary between IT and science, creates “skin in the game”
- Research computing group run through enterprise whose mission statement is not “performance”
- From user services HPC – viral nature of success stories among peers, may not be domain specific.
- NCAR – we’ve talked about this issue for awhile, still have problems fixing this. Hard nut to crack
- IU – performance engagement, trying to help researchers achieve better performance. How does a researcher know he/she is not getting good performance. “No one will ever call me b.c they don’t know they have a problem” – IU has people to help
- Better marketing? Networking at the speed of thought.
- New science requests – storage and network integrated into the proposal.
- Hiring personalities – who like solving problems, who can put themselves into the heads of the user
- NIH – outreach difficult, folks don’t know they are doing something out of the ordinary. People don’t know what to expect – speed test? Run on their server on their desktop. Perfsonar can give you a sense of optimal speed and what you are actually getting
- Simple tool to run some performance and you can see wha tyour performance is compared to others
- Database of success stories by research area

Agenda

- Welcome and introductions
- BoF Goals
- Overview of National Research & Education Networks at work Globally
- Discuss needs, challenges for leveraging HPC and high-performance networks
- HPC/HTC pre-SC15 ESnet/GEANT/Internet2 survey results overview
- Next steps discussion
- Closing and Thank You

Prominent themes from pre-SC15 survey participants:

“High Performance Computing / High Throughput Computing resource needs, challenges and best practices” (60 responses)

1. Energy & Environment and Healthcare & Life Sciences are emerging users/consumers of HPC and networking

2. Most respondents are using local HPC/HTC resources

3. IT departments are heavily involved with HPC/HTC

4. HPC data size is expected to double within next 2 years – currently working with petabytes or terabytes of data

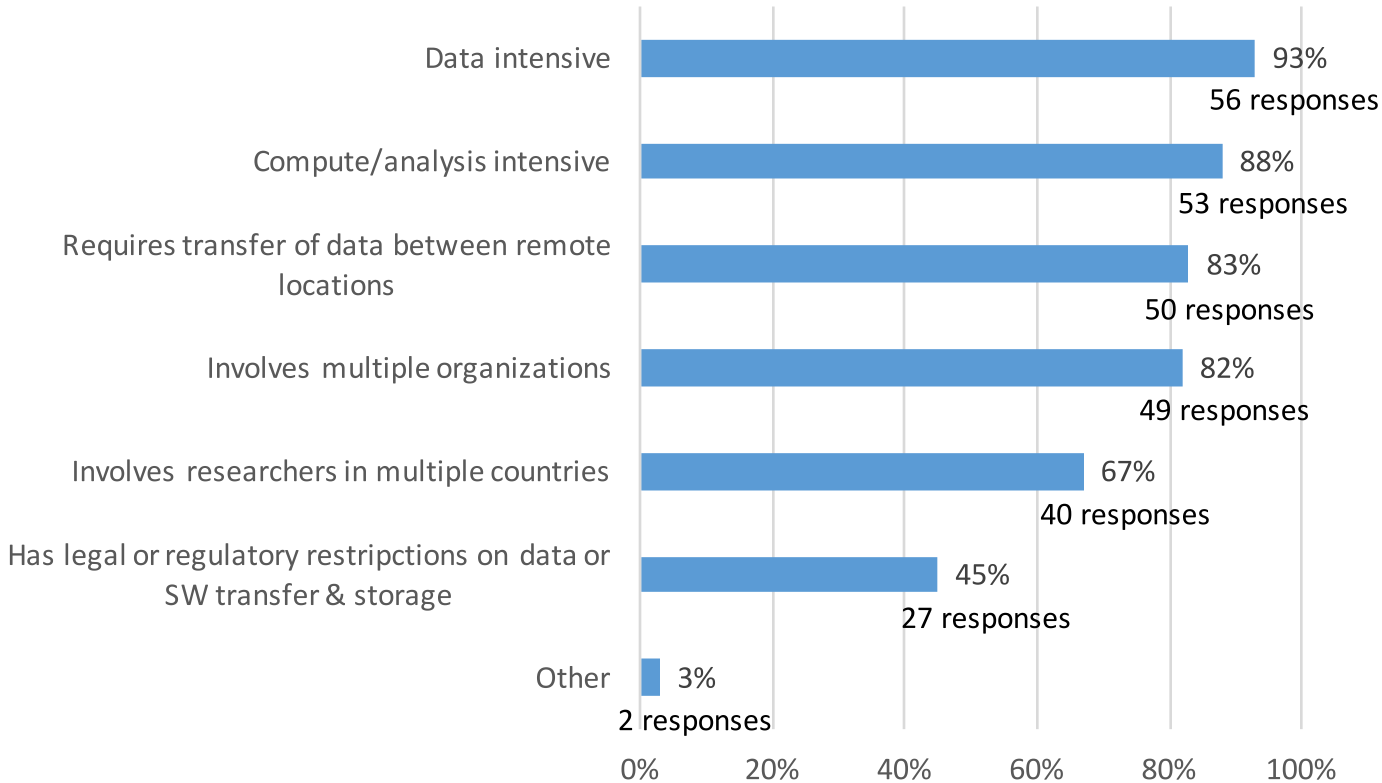
5. Difficulty with data transfer speeds amongst multiple locations and organizations

6. Speed and resources (storage, technical, analytical, tools, and human) are problematic

7. Globus is a popular choice in leveraging high performance networks for research computing

HPC/HTC needs revolve around sharing compute intensive data amongst multiple locations and organizations

Q3: How would you describe your research or the research you support? Choose all that apply



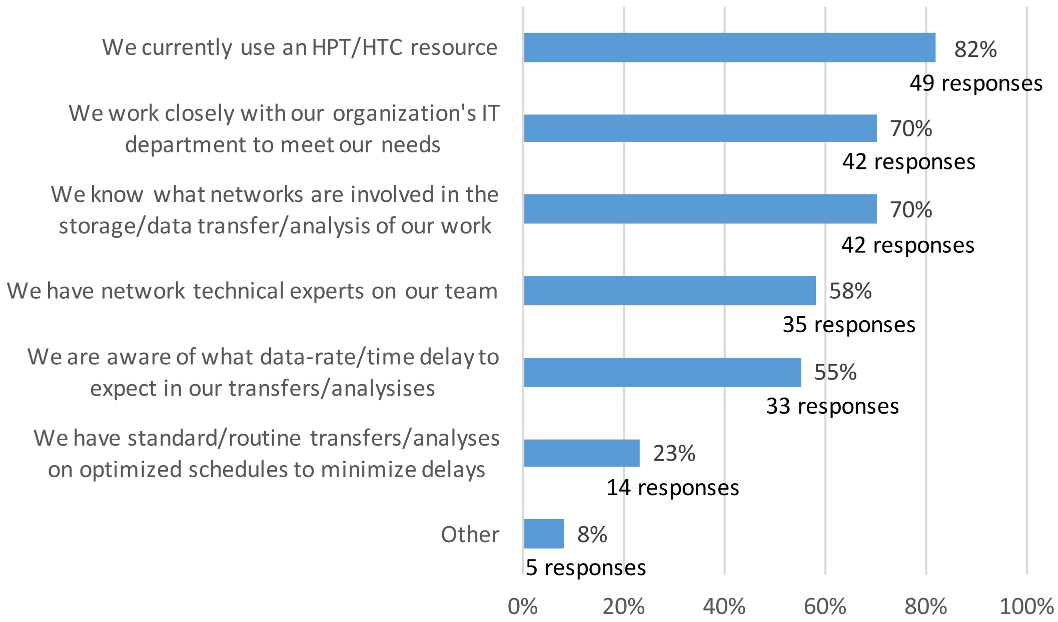
Other responses include:

- N/A
- Requires service provider carrier grade technology for the networks

N = 60

Most respondents are currently using an HPC/HTC resource, understand the networks involved with their data, and are closely tied to IT

Q4: How would you describe your HPC/HTC resource needs? Choose all that apply.



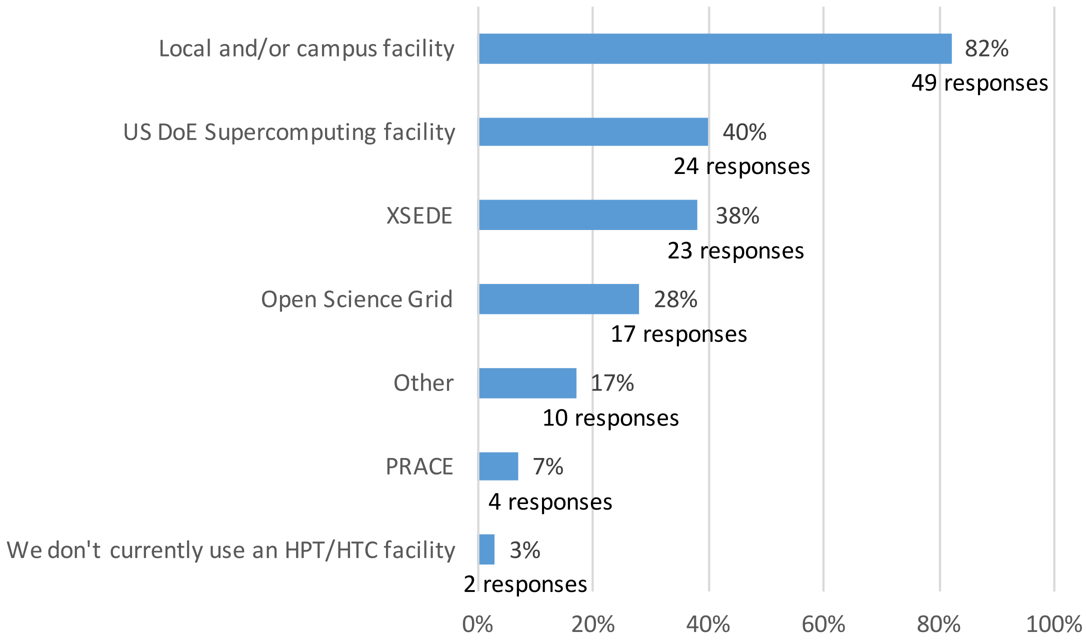
Other responses include:

- We are part of the IT department
- We make extensive use of Globus services
- We have our own HPC including the 7th fastest computer in the world
- I represent the needs of a large research university. The needs are fundamentally insatiable in some areas. There is an urgent need for a truly competitive market for research computing services to emerge.
- We run HPC and HTC resources

N = 60

Most respondents are currently using a local or campus HPC/HTC facility

Q5: Which HPC/HTC resource do you currently use? Choose all that apply.



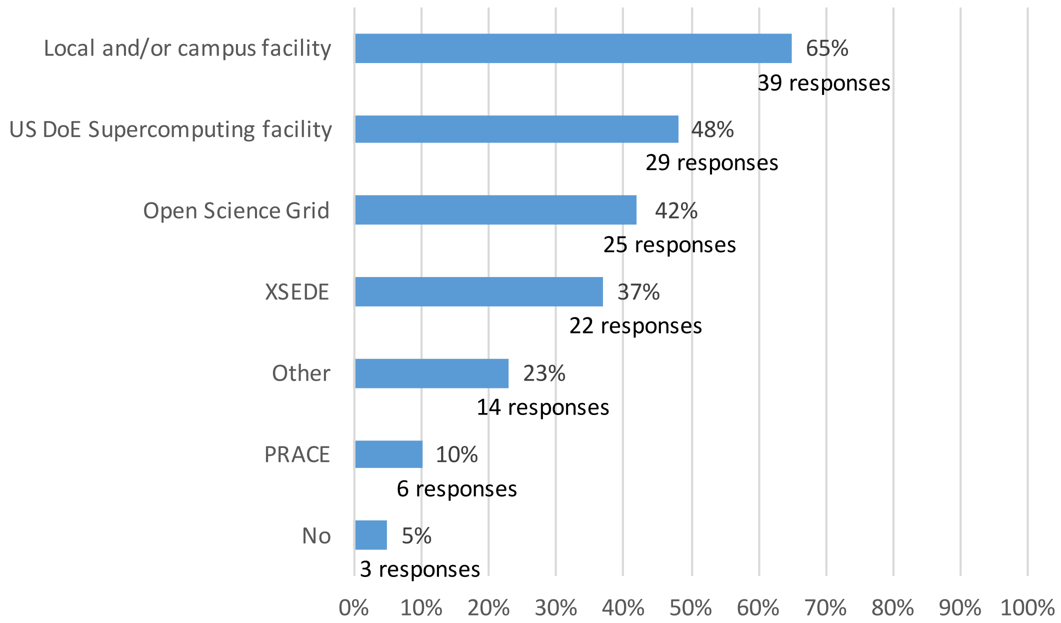
Other responses include:

- National e-infra resources (compute, storage, network)
- AWS
- TACC
- Statewide HPC resource
- European grid, national research cluster, national research cloud infrastructure
- Users we support use these resources when they can
- Network Infrastructure is in pale for up to 4x10 I2 connection to National Labs
- Compute Canada
- Other federal government facilities (NASA, NOAA), cloud providers (AWS, Softlayer)

N = 60

Most respondents would like to continue using a local facility or a US DoE supercomputing facility

Q6: Would you like to use an HPC/HTC resource? Choose all that apply.



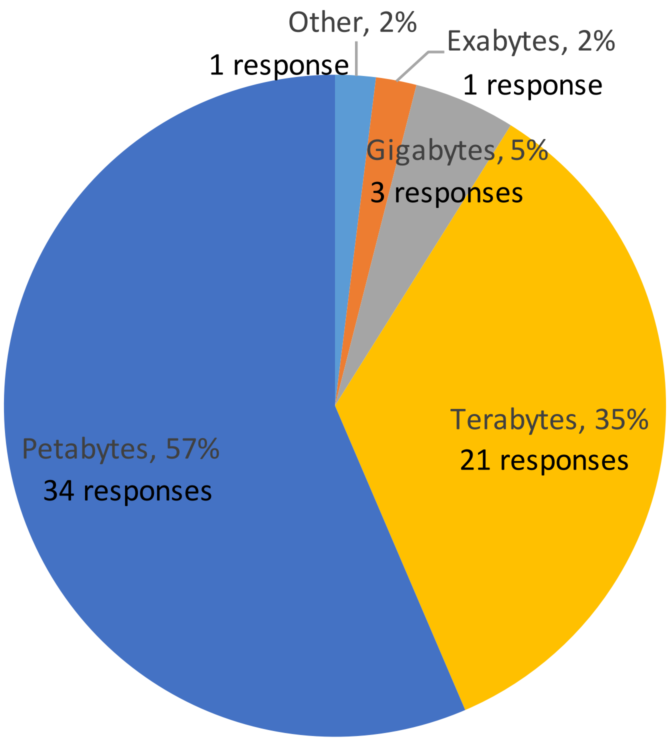
Other responses include:

- Private cloud/Boinc Volunteer computing
- We would like to assist our researchers in utilizing whatever HPC/HTC resource they need
- AWS
- AWS cloud
- I would like to help our researchers gain access to any/all of the resources available to them. The facilities checked above are ones of which I'm aware are being used
- Yes
- Network Infrastructure is in place for up to 4x10 I2 connection to National Labs
- Compute Canada
- The community I represent already uses all the listed resources, as well as commercial services
- We run HPC and HTC resources
- Already using per Q5
- Globus
- CloudLab
- Yellowstone

N = 60

57% of respondents are working with petabytes of data. 92% are working with data measured in at least terabytes.

Q7: What is the total volume of data you use and/or generate for your research (or the research you support)?

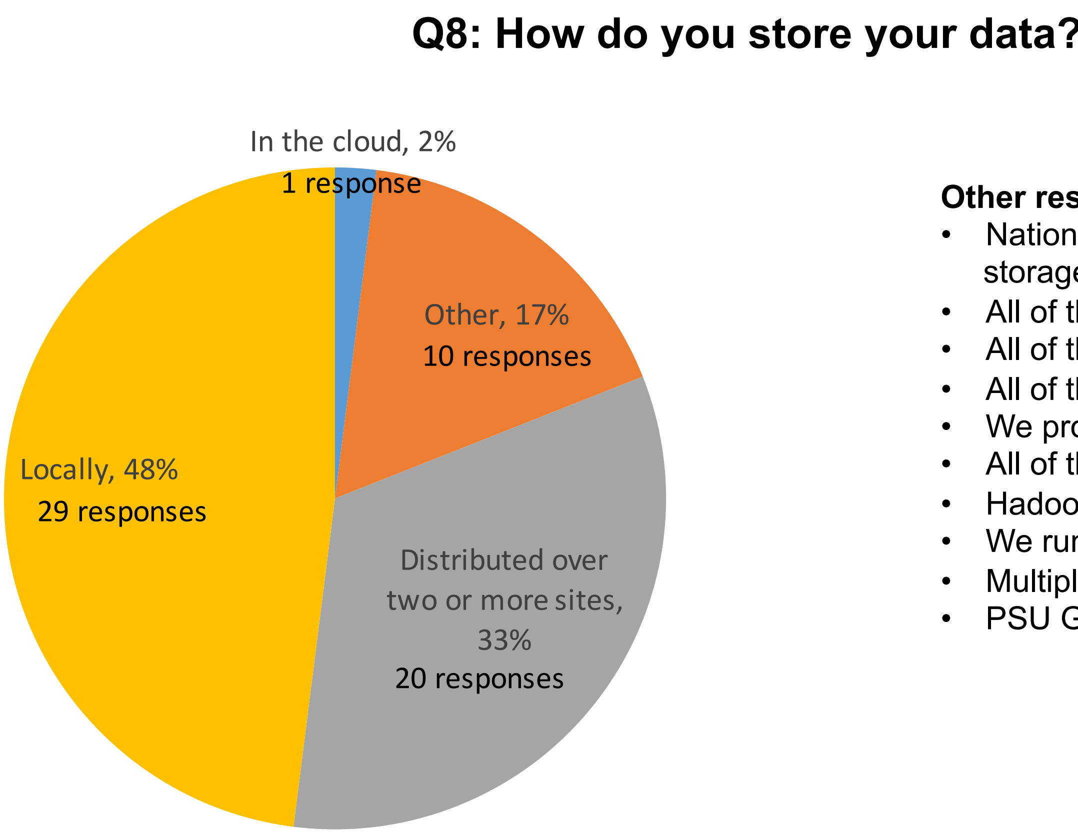


Other responses include:

- We run HPC and HTC resources

N = 60

Most respondents store their data locally, followed closely by distributed over multiple sites



Other responses include:

- National e-infra resources which include storage
- All of the above
- All of the above
- All of the above
- We provide cloud storage
- All of the above
- Hadoop HDFS cluster
- We run HPC and HTC resources
- Multiple platforms
- PSU GPFS storage

N = 60

Next Steps

- We can send you the survey, you will receive results if you participate
- Please fill out the survey and participate

https://www.surveymonkey.com/r/SC15BoF

- Report out v1 January 2016
- Lets keep the conversation going!

Thank you very much !
Domo arigato gozaimasu !
Muchas gracias !
Merci beaucoup !
Grazie mille !

ESnet
ENERGY SCIENCES NETWORK

GÉANT

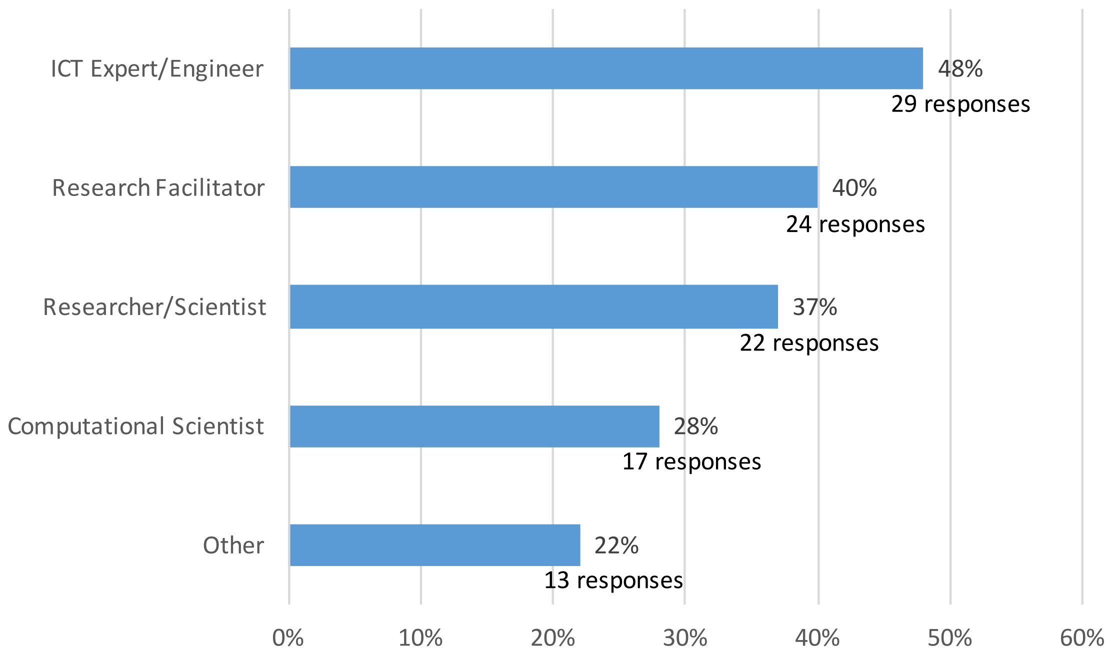
INTERNET2®

engage@es.net
cino@internet2.edu
businessdevelopment@geant.org

Appendix:
Survey Results
(as of Nov. 9 2015)

Nearly half of respondents classify their role as ICT Expert/Engineer, followed closely by Research Facilitator

Q1: What is your role? Choose all that apply

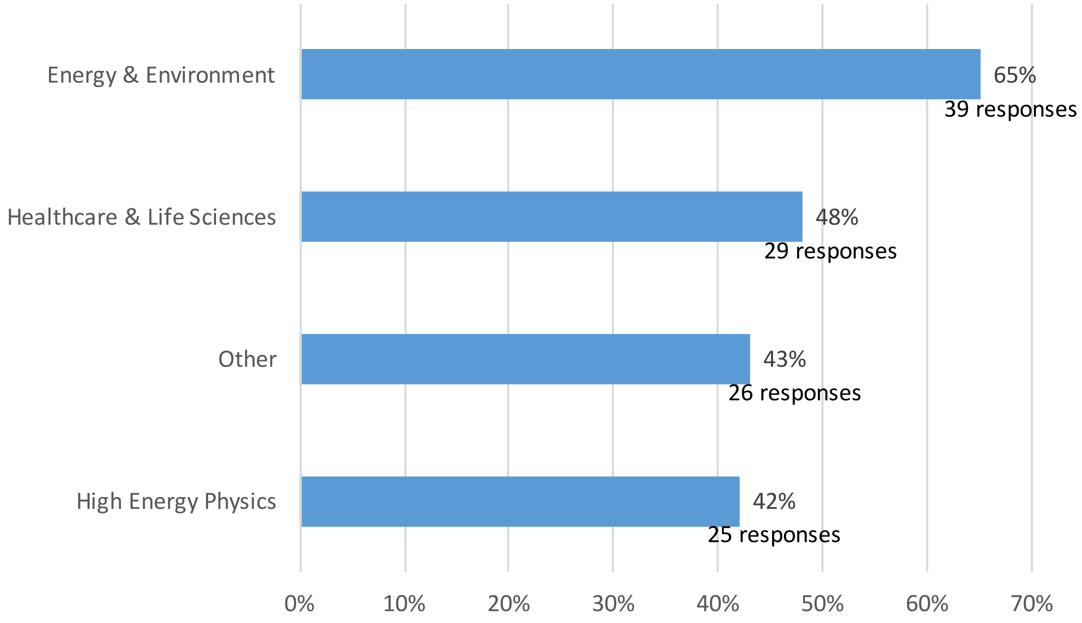


Other responses include:
- •Data scientist
- •CIO
- •Director (2)
- •Deputy CIO for Research – interface between central IT resources & our researchers
- •Research computing service provider
- •HPC user support
- •IT senior director for university research computing
- •Research support
- •Director of High Performance Computing
- •Research Computing Director
- •IT Project Manager
- •State coordinator for research cyberinfrastructure in higher education

N = 60

Energy & Environment and Healthcare & Life Sciences are the leading disciplines for HPC/HTC research areas

Q2: Please specify your area of research or areas of research that you support



Other responses include:

- Hurricane forecasting
- Social Science, humanities, physical sciences, engineering
- Transportation
- All science domains across all campuses
- Material sciences
- Financial services
- Modeling
- Big Data/Analytics, Internet of Things/Sensor Nets, Water Quality/ Environment
- All
- All Computational Science
- HPC
- Computer Science
- Chemistry, Materials, Macromolecular Science, Mechanical Engineering, Biomedical Engineering
- Engineering
- Hydrodynamic/offshore Engineering, marketing, economics
- Weather, Computational Chemistry, Material Science
- All engineering disciplines, all science disciplines, social sciences, humanities, architecture, and business
- Geophysics
- Humanity
- All sciences and engineering
- Computational chemistry & material science, virtual training & collaboration environments, IoT/CPS for freshwater research
- Genomic sequencing and Bioinformatics of large datasets
- Broad Academic

N = 60

Issues identified that impede research center around speed and proper resources

Q9: Is there any known problem that impedes your research activities? (e.g., data transfer to/from remote site is too slow, inconsistent or unpredictable, etc.)

1. Quality of Service for data transfers; conflicts between research and campus security needs.
2. Inter-campus data transfer and transfer to/from XSEDE
3. Yes, all of the above. Plus lack of security on administrative networks, would prefer private networks.
4. Need more aggregate bandwidth 10->100Gbps
5. Data transfer to/from remote site is too slow, inconsistent
6. No
7. Users are still developing data management strategies.
8. Data mobility between facilities remains a challenge. There are challenges around basic authentication and authorization as well as challenges around efficient services that use available bandwidth effectively.
9. Local storage limitations
10. Inefficiency when transferring a large number of small files
11. Need to establish policy for storage, access, and transport of Research Data
12. Not really
13. Existing ‘legacy’ software that expects data to be available on a locally mounted filesystem
14. Various “costs” associated with transient data placement and aggregation that would allow HTC solutions vs. simplicity of existing HPC solution
15. Data transfer rates are rarely sufficient for the volume of data
16. Yes
17. Always better to be faster
18. Too slow
19. Fileserver I/O speed are unpredictable due to million of files being updated. ZFS with SSDs alleviating this.
20. Complex user administration; complex system administration
21. Data transfer too slow, transfers inconsistent
22. "Last mile" problem me of limited data rates from labs to campus core network.
23. Lack of local network expertise to facilitate data movement
24. No
25. Robust data transfer
26. Not really
27. Budget for local infrastructure
28. No
29. Data share
30. High cost of data communications in and out of Saudi Arabia
31. Non-standard network configurations are unstable even with the greatest experts on the problem
32. Accessibility due to business firewalls
33. Mostly on campus networking issues in specific buildings here specific researchers are located
34. Data storage, transfer and archiving is not yet seamless
35. Regulatory and data custody issues
36. No
37. No
38. Local infrastructure is slow/obsolete
39. Network speed constraints; in the process of upgrading to 100Gb or better connections
40. Relative immaturity of master data management practice, lack of accurate global time synchronization, nascent internal state of semantic technologies and huge web of conflicting regulatory (domestic and international) requirements and constraints.
41. Works very well for large institutions; small sites have firewall/bandwidth/expertise limitations
42. Lack of super computing facility
43. Try to make it easier for new partners
44. The usual tuning issues
45. No
46. It could always be faster
47. Globus is not installed at all sites, or DTNs not configured properly at all sites
48. Transfer speed fluctuation
49. -
50. Workflow orchestration tools need to be made more mature
51. N/A
52. There is variability in transfer speed between sites. The data management work should be automated (data movement, archive to HPSS, standard analysis, documentation). While some sites have this capability and some are in progress, others have not started. There is a limitation in the amount of open data storage that can be utilized.
53. Data transfer too slow
54. Data transfer to/from remote site is too slow
55. Tools for interacting with remote data
56. Having the right people that can support the computational needs of the disciplines
57. Security is likely the biggest issue
58. Access to high speed storage
59. Technical support/staffing available to assist in optimal use of resources
60. Data Management between distributed locations

N = 60

Most respondents expect their data to at least double within the next 2 years

Q10: What is the expected growth rate of your data? (e.g., will double annually for 5 years)

1. Annual doubling is a reasonable guess (but just a guess)
2. Order of magnitude growth as new 5-year multi-institutional project kicks off with both computational and data-intensive components
3. Increases by a third each year.
4. Doubles every 3-5 years
5. Double every 2 years
6. Doubles appx every year
7. N/A
8. We expect doubling every 3 years.
9. Double every three years
10. Will grow about 5X
11. Move from 1 Ped/yr to 1 Ped/mo in 3 yr, w/ 2 Ped/mo in 5 yr.
12. Haven't thought about that.---!!!!!!!
13. Unknown
14. Could double over next 2-3 years.
15. 35% year over year growth
16. Factors of 5 to 10
17. Double every two years
18. Expect raw data to double every year but archived data storage will be less
19. Double each year ; no end in sight
20. 150 TB / year for the next 3 years
21. Don't know
22. 50% annual increase.
23. Unknown but likely to double every year or so
24. Double every 3 years
25. Quadruple every 2 to 3 years
26. No idea, up to our users
27. Double in 2 years
28. No estimates
29. Double every two years
30. Will triple annually for the next 5 years
31. Petabytes per month in 5 years
32. Exponential
33. Doubling around every 12-18 months
34. It is more than doubling every year
35. Unknown
36. Double each second year
37. 150Gb a year
38. > doubling annually
39. Expected to double annually for foreseeable future
40. Circa 30% CAGR
41. 20% growth per year
42. Triple in 5 years
43. Lots
44. Unknown
45. 50%/year
46. 20% per year
47. Doubling every 1-2 years
48. Roughly double in three years
49. -
50. Double every year
51. N/A
52. Every 3-4 years the data amounts will quadruple. If significant amounts of storage are available we can start examining higher temporal and spatial resolution data, and thus, our data rates can grow
53. Ten to hundreds of petabytes
54. Double every 2.5 years
55. Will double every year
56. Higher than is comfortable
57. 35% annual
58. Double every six months -----???
59. Will double over next three years
60. Double within two years … not know farther out

N = 60

Globus is a popular choice in leveraging high performance networks for research computing

Q11: What success stories and/or best practices can you share for research leverage of high performance networks? (Optional)

1. Best Practice: Globus Online combined with a high-speed Science Network and Science DMZ
2. We don’t have one yet
3. PerfSonar, DMZ and Globus deployments
4. We have several projects that share observational, simulation datasets broadly. These span
5. Reprocessing all CrIS sounder data from mission start, making this generally available, statistical analysis and intercomparison of AIRS, CrIS, IASI sounder data with ECMWF model data
6. Combining the resources of the RON and UT System provides leverage to access I2, ESnet, and Community Networks to reduce payments to LEC/CLEC sources.
7. Use automated movement of data
8. CC-NIE award was instrumental but we still need a cyberinfrastructure engineer
9. Globus helps
10. Contact the experts supporting the infrastructure you use in case of problems or questions
11. Standardization and consolidation of resources
12. http://www.internet2.edu/research-solutions/case-studies/accelerating-genomic-research-advanced-networking-collaborations
13. Leveraged high speed SAN with GridFTP and 40GbE connections to 100GbE uplink
14. Science DMZ/perfSonar very helpful in moving large data sets
15. Send the analytics to the data, not vice versa.
16. I like the new technologies, we need it to make it for the next generation to excel.
17. Globus is a great tool for transferring large data sets
18. Part of calculation (eigenvector) is done in remote site on capability machine on one national Lab and transferred to capacity machine on another lab to use it to maximize the efficiency of each site.
19. When Globus transfers are not limited by the requirement of OSG certificates, they work great even if there is variability.
20. We are piloting a staging service optimized for transfer of data (using gridftp)
21. Globus online, with sharing, and better protocol to utilize network
22. Faculty led governance model.

N = 22

Other challenges center around storage, automation, and human resources to handle data

Q12: What else do you want to share as a need, challenge or solution for leveraging advanced technologies in research and science endeavors? (Optional)

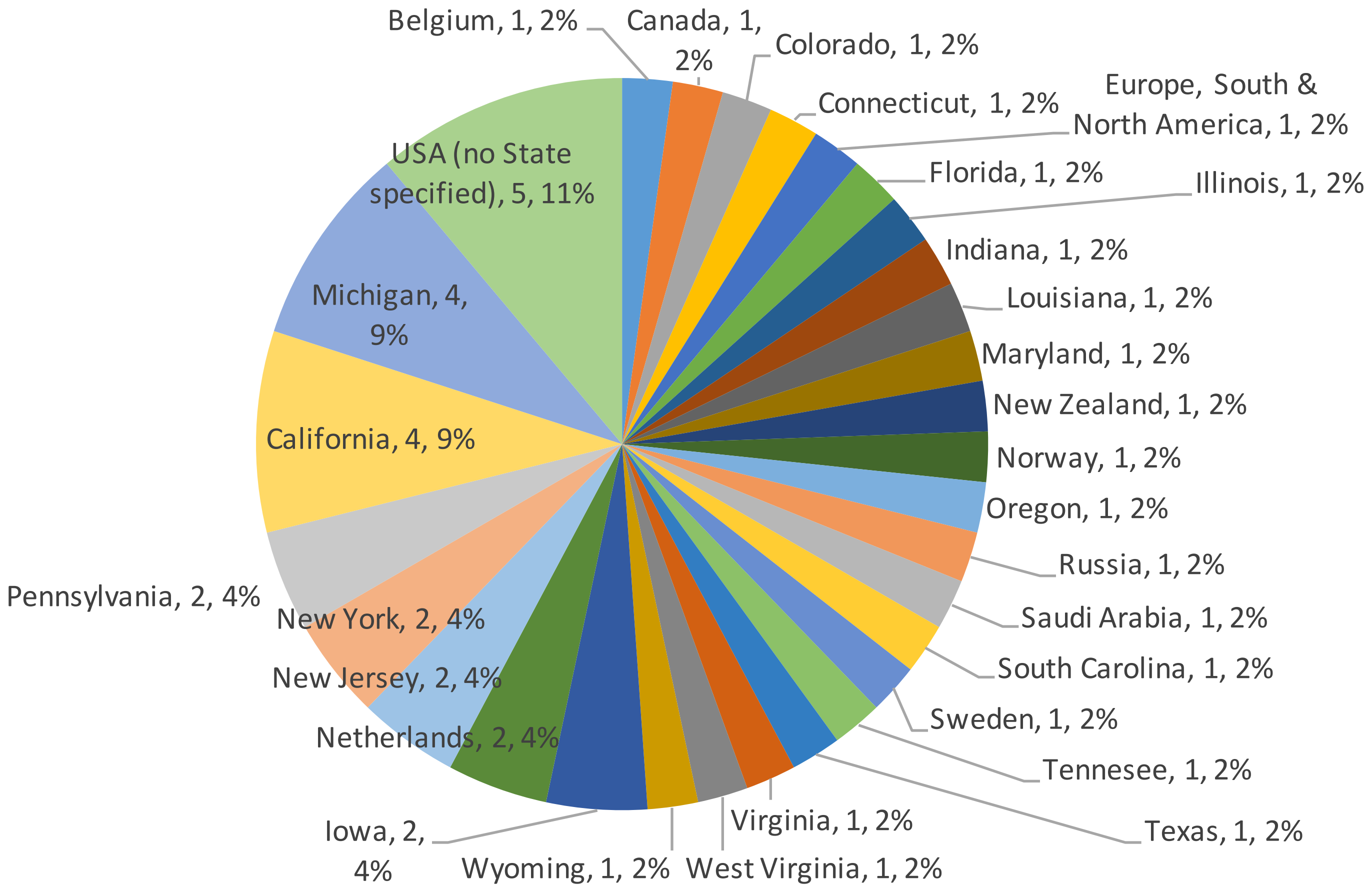
1. Challenge: Cost-effective preservation for large-scale data
2. Need for tools to help federate regional resources and inter-federate with national resources
3. Virtual machines and head nodes for intensive data analysis and customized solution provisioning and sharing (such as Galaxy instances or Globus instances)
4. Need additional identity management federation
5. There is a need investment in a hierarchy of resources. Over emphasis on large monopolistic enterprises (whether tax-payer or commercially funded) are troubling and a challenge to research needs. A healthy market place for computing probably needs the challenge of data lock-in to be solved more effectively. The current market for research computing services has many problems that distort the economics and inhibit competition. Networking can play an important role in creating a functioning market.
6. Cluster tools like slurm are usable but need refinement
7. Institutional resources in IT Network are necessary and not sufficient to support Research Lab requirements and Data Infrastructures. Must address HR/sourcing issues.
8. Challenge: getting expected performance from all the components of an emerging technology high speed network (e.g.. 40 Gb)
9. Research support resources are scarce, in particular providing last mile connectivity on campus.
10. More non-HPC researchers using HPC!!!
11. Local high speed science network
12. We do not have sufficient support staff to help researchers maximize use of HPC, GPU, Hadoop/Spark environments. I suspect this is a problem for many institutions.
13. Help the next Frontier
14. Data management expertise.
15. Research data management services
16. Human resource with both domain science and computer skills
17. The need is to automate the handling of data to reduce project cost and to have scalable data handling. This would allow handling 10-100x more data. Our project gets ~10-13% of total cycles on major machines. Given the last two items we are unlikely to need significantly more cycles so there is little ability to grow in simulation length. Where the growth can come from is having access to significantly more scratch and archival storage, we can output more data and get finer detailed view of our simulation.
18. Mid-tier computing using new architectures.
19. Need for movement of compute to where data rests

N = 19

Pre-SC15 Survey Respondents were primarily (73%) U.S.

Half of all respondents requested a copy of the results.

Q13: What country/state do you work in? (Optional)



N = 45