ICNWG
PROGRESS UPDATE & FUTURE ROADMAP

ESGF F2F Workshop,
Washington, DC, December 2016

Eli Dart, ESnet

2016 Progress & Accomplishments

- Update: Mary Hester has left ESnet
    - Joined SURFnet in the Netherlands
    - Looking forward to future collaboration

2016 Progress:

- perfSONAR deployment
    - Partial success (BADC, DKRZ, LLNL – tests partially functional)

- DTN deployment
    - Partial success
    - Software stack is not up in production on the DTNs

- Synda
    - Partial success
    - Testing done

2016 Missed Milestones

- DTN deployments (4x10G DTNs)
- Globus deployment, data availability via Globus
- Use of Synda by replication group?
- Performance targets

2017 Roadmap

- We have several things we still need to do
    - DTN deployments (in Science DMZ environments)
    - Synda + GridFTP/Globus as a fully-supported download method
    - Performance targets (4Gbps or 500MB/sec to capable sites)
    - User-facing documentation

- All of these require resource commitments
    - How do we get there?
    - Need people resources at sites

Additional Resources Needed

- Staff needed
  - Sysadmin/deployment resources at sites
  - Replication users to use resources and ensure everything is working

- Hardware (computing, storage, network) needed
  - 10G+ path from DTNs to 100Gbps backbones
  - High-speed storage access (capable data store for DTNs)

- Operational support needed
  - Sites need ongoing operational resources
  - Maintaining high-performance ESGF features appears starved for cycles
  - Significant work needed across the Federation (not just Tier1 sites)

History and Context

- Original ICNWG goals
    - 2014 - 4Gbps (500MB/sec) between data centers
    - 2015 - 8Gbps (1000MB/sec) between data centers
    - Deployment of perfSONAR for test and measurement

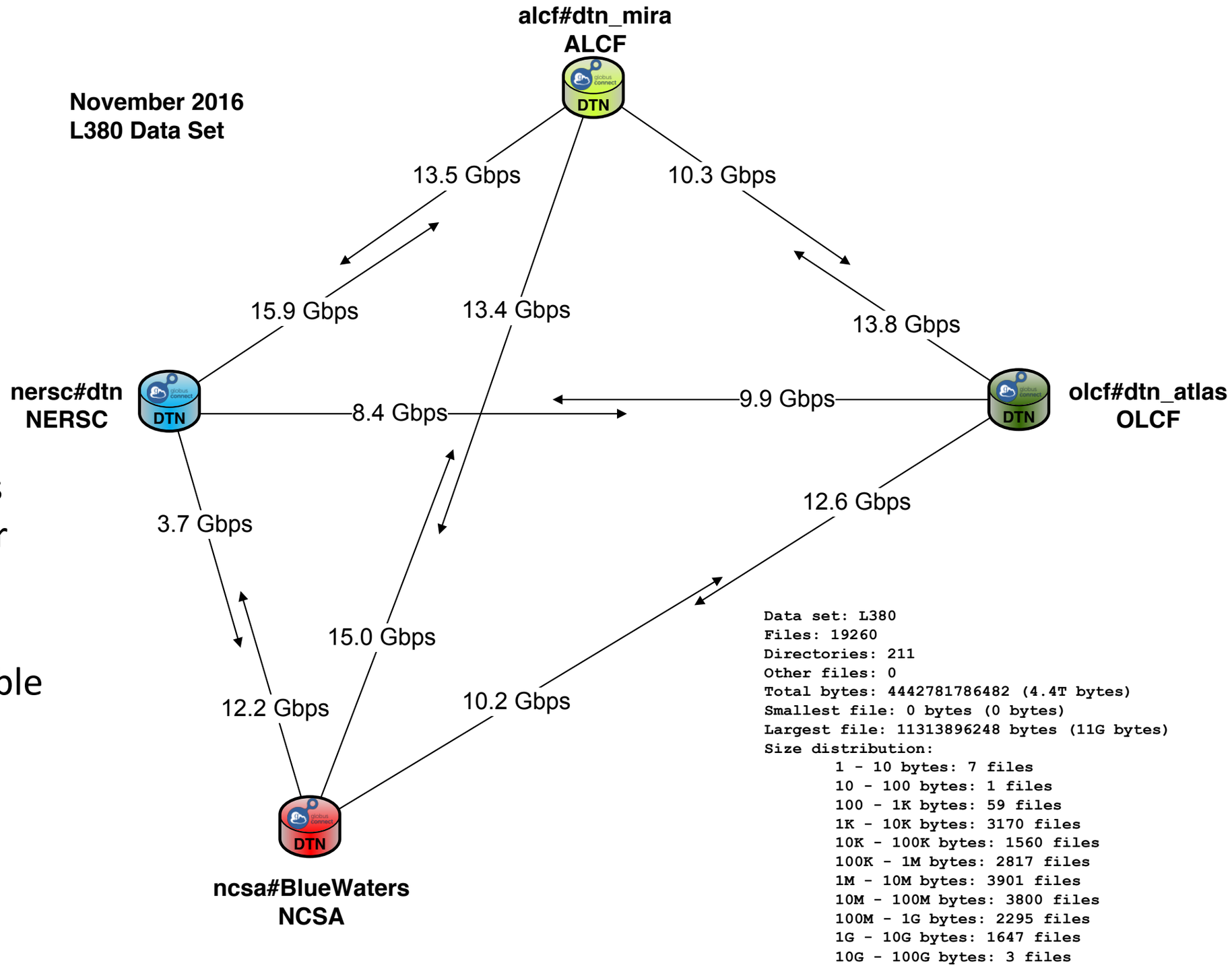
- Revised ICNWG goals
    - 2016 - 4Gbps (500MB/sec) between data centers
    - Deployment of perfSONAR for test and measurement

- Progress has been slow
    - Is this the right thing to be working on now?
    - Should we suspend the ICNWG for a time?

- How to serve the large-scale users?
    - Many scientists need facility-scale computing, with ESGF data
    - One user takes 10s of TB of CMIP5 data today – presumably more CMIP6 data
    - HTTP download doesn't work for these users – old architecture doesn't scale
    - What is the roadmap for serving these users?

High Performance Data Movement Is Possible

- Petascale DTN project
- Globus transfers between HPC facility DTN clusters
- Performance goal: 15Gbps for routine Globus transfer jobs
- Science DMZ model, Capable DTNs, Globus



Discussion Points

- Is this the right time to be doing this work?

- If so, how do we increase the rate of progress?

- How to serve the large-scale users?
    - Data from ESGF data centers to HPC facilities
    - Data from ESGF data centers to institutional HPC resources

- What can we learn from other fields/sites?
    - LHC experiments
    - HPC facilities

Thanks!

Questions?