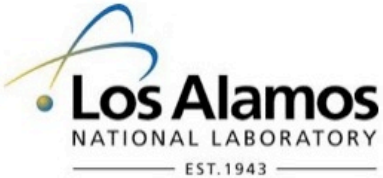

HPCOR 2014

Session Name: Data Transfer (session D2SD)

Co-Chairs: Andrew Cherry, Eli Dart

Contributors

- Curt Canada – LANL
- Andrew Cherry – ANL (co-chair)
- Bill Collins – SNL
- Eli Dart – ESnet (co-chair)
- Chris Fuson – ORNL
- Mark Gary - LLNL
- Cory Lueninghoener – LANL
- David Smith – LLNL
- Joel Stevenson - SNL

What WAN access is in place?

- Many labs/facilities have 100G now
  - ANL, NERSC, ORNL – ASCR HPC facilities
  - FNAL, LBL, LLNL, - Other labs/facilities
  - BNL, LANL, SLAC in process
- DTN capacity
  - Many sites have multiple 10G DTNs
  - Most 10G DTNs mount center parallel filesystem
- DTN tools
  - Globus Online
  - Globus-url-copy / command line GridFTP
  - BBCP
  - HPN-SSH
  - Classic SSH
- Performance monitoring
  - Most sites run perfSONAR
- Dedicated data paths for some programs (e.g. DISCOM)

What WAN access is needed?

- Larger DTN pools
  - Includes Globus improvements to more effectively utilize
- Multiple uses for DTNs – multiple pools?
  - Most WAN data transfers are POSIX file movement tasks
  - Streaming or other non-POSIX uses
  - Remote I/O (e.g. XRootD)
- Quality of service
  - Bandwidth management between DTNs and other network services
- Tools improvement
  - Better Globus/HPSS integration
  - Integration with QoS or other management/scheduling systems
- Security vs. performance is a continuing challenge
  - Compartmentalizing networks is labor-intensive
  - Firewalls cause performance problems compared to other packet filters

How do you handle data transfers In/Out of your facility today (e.g. do facility staff conduct data transfers or do users, etc)

- Tools – primarily Globus, but other tools used too
- Most transfers are user-initiated
  - Some transfers are done for users (e.g. physical – see below)
  - Center staff do operational transfers (e.g. for disaster recovery)
- Data transfer over the network is preferred
  - Physical transport of portable media sometimes occurs, but is labor-intensive
  - DTNs are preferred
- Portals use HTTP in many cases
  - Easy for users
  - Low performance
  - ESGF is one example, there are others

What drives the need for networking and what will HPC centers need to do to accommodate that need?

- Reduce the need to ship disks
- Data transfer between HPC facilities
  - Data preservation
  - Location of allocation
  - Data stored where it was generated, but accessed later
- Disaster recovery, data redundancy
  - Bigger pipes and higher performance endpoints would increase DR capability
  - Replication need scales with data set size (and therefore with HPC system capabilities)
  - Re-engineering firewall config for DR replication might help (e.g. Science DMZ model for internal DR operations)
- Data ingest/export: other facilities/experiments
  - Large scientific experiments (light sources, telescopes, etc.)
  - Large collaborations (Climate, cosmology, etc.)
- IPv6 to support some future experiments (e.g. CERN)

What are your major strategies and initiatives over the next 5 & 10 years? How do they affect staffing levels?

- Data sharing has to be addressed (out of self defense if for no other reason)
- IPv6 adoption
  - infrastructure changes
  - staff time (if no staff increase, then other projects may suffer)
  - Tools qualification/modification/support
- Keep pace with demand
- Getting common capability set deployed across facilities (see collaboration section)
  - Tool set
  - Policy
  - Bandwidth
  - perfSONAR
- Build multi-facility collaboration to improve operational outcomes
- Staffing – we don't expect new staff ☹

What are your current efforts and/or site configuration in this area?

- See previous section on per-site config
- Some sites working toward adopting Globus (see collaboration efforts for adoption of common toolset)
- Initial exploration of sharing (e.g. Globus sharing)
- 100G deployments continuing

What are your mandates and constraints?

- Security (especially in the context of sharing, but also in other contexts)
- Money is always a constraint
  - Issues of balance (compute/storage/DTN/network)
  - Need to identify issues that could be fixed for a given dollar amount
- Available network bandwidth vs. competing customers
  - Data transfer
  - data streaming
  - business operations
- Lack of requirements from users is a constraint (at least for productivity)
  - users are not always good at extrapolation
  - data generated vs. data products saved - not always easy to distinguish
- Mandates
  - **100% uptime, flawless performance, perfect reliability**
  - Transfer reliability/manageability (big driver for Globus adoption)
  - Ease of use for users
    - Users will complain
    - Reduced support load

How to do you forecast future needs and requirements?

- Formal programmatic vehicles for requirements analysis
  - ESnet requirements reviews - http://www.es.net/about/science-requirements/
  - NERSC requirements reviews - https://www.nersc.gov/science/hpc-requirements-reviews/
  - LANL - Every 3 years, big internal review (large-scale planning)
- Utilization metrics used for forecasting
  - Network usage
  - storage usage
  - User input
  - Proposal process

What opportunities exist for productive collaborations among DOE HPC centers?

- Regular testing of tools/capabilities
  - performance
  - interoperability
  - DTN matrix a la ALICE (http://alimonitor.cern.ch/speed/index.jsp?site=LLNL)
- Deployment of common toolset across facilities
- Build collaborations between system/network/ security administrators for troubleshooting and performance
- Collaborate between centers and with Globus for making better use of DTN pools that mount parallel filesystems (see above)
- Consistent tuning/best practice a la fasterdata.es.net

What are the biggest challenges and gaps between what you can do today and what will be required in 5 - 10 years?

- Data ingest and export from commercial or other sites without science network connectivity is a significant challenge
  - smaller universities
  - business
  - portable disk
  - boxes of LTO tape
- Security enclave partitioning takes a lot of cycles for networking folks
- LAN protocols perform poorly in the WAN (HPSS mover protocol, commercial backup solutions, etc)
- TCP loss behavior is an impediment to performance
- Gap between data we have and data we can move will increase
- Increased desire by users for sharing will present challenges
  - Security
  - Performance
  - Scale

What are the biggest challenges and gaps between what you can do today and what will be required in 5 - 10 years?

- Identity management issues
  - Federated identity would be a big help
  - Concerns about sharing data outside of the labs (how to track/control when necessary)
- Cloud computing and cloud storage are likely to challenge the labs in the future
  - Integrating with Amazon et. al.
  - Transfer between Lab infrastructure and Cloud
- Data location services will become more necessary
  - Some fields do this now (e.g. Climate)
  - Sloan Digital Sky Survey archive at Johns Hopkins sees a lot of traffic after data set made public
    - Challenges for WAN
    - Potential benefits for science

What are the biggest challenges and gaps between what you can do today and what will be required in 5 - 10 years?

- Increased dependence on WAN
  - can't disconnect for security issues anymore
    - compartmentalization is required
    - additional operational costs
- Balance between network capability and ability to use that network capability
  - Tools gaps
  - System gaps
    - underlying storage (e.g. per-DTN performance to parallel filesystem)
  - Data type and data set composition (file count, file size, etc)
- Network security appliances continue to impact performance (sometimes catastrophically)
  - Science DMZ model can help, but can be a policy challenge
- What comes after 100GE?
- Need to figure out how to best configure tools to use a DTN pool that mounts a parallel filesystem (striping across nodes, per-node parallelism, filesystem mount parameters, etc). (see collaboration section)

Describe some practices that you think are effective as well as lessons learned that would be helpful to other centers?

- Use of DTNs for data ingest/export
  - Single place to steer users
  - Single point of entry for network provisioning
  - Single place to deploy tools
- Adoption of Globus - best practice
  - easier for end users to deal with
- Project-based documentation that could be provided to external organizations to allow them to reason about data ingest/export when working with the Labs
- Elicit data transfer and sharing and storage requirements from users/projects
  - This is always a moving target, could always be done better
  - A bit of actionable information is better than trying to go for the complete answer
  - Users don't always know how to do this, so working with users is important
  - Better data management planning
- Including additional metrics in utilization monitoring for capacity planning

Top Findings

Opportunities

- Deployment and regular testing of a consistent toolset across facilities
- Data sharing across DOE facilities
- Share DTN configuration and practices (e.g. data transfer working group)

Best Practices

- Data ingest and export should be done using DTNs
- Globus Online as primary DTN toolset
- Collect data transfer and sharing requirements from projects

Challenges

- Managing TCP
- Security vs. performance
- Security vs. sharing
- Get out of business of ingesting data via physical media
- Identity management
- Effective use of networking capability
  - DTN to filesystem performance
  - Tools performance
  - Data set composition
- Cloud interaction