Foundations of data-intensive science: Technology and practice for high throughput, widely distributed, data management and analysis systems (version 3)

William Johnston
Senior Scientist and Advisor
Eli Dart, Brian Tierney, Jason Zurawski, Chin Guok, and others
ESnet
Lawrence Berkeley National Laboratory

2014 Technology Exchange
Indianapolis, Indiana
October 26-30, 2014
Why are data-intensive applications issues important to ESnet?

- The U.S. Department of Energy’s Office of Science (“SC”) supports about half of all civilian R&D in the U.S. with about $5B/year in funding (with the National Science Foundation (NSF) funding the other half)
  - Funds some 22,000 PhDs and PostDocs in the university environment
  - Operates ten National Laboratories and dozens of major scientific user facilities such as synchrotron light sources, neutron sources, particle accelerators, electron and atomic force microscopes, supercomputer centers, etc., that are all available to the US and Global science research community, and many of which generate massive amounts of data and involve large, distributed collaborations
  - Supports global, large-scale science collaborations such as the LHC at CERN and the ITER fusion experiment in France
  - www.science.doe.gov
DOE Office of Science and ESnet – the ESnet Mission

- ESnet - the Energy Sciences Network - is an SC facility whose primary mission is to enable the large-scale science of the Office of Science that depends on:
  - Multi-institution, world-wide collaboration
  - Data mobility: sharing of massive amounts of data
  - Distributed data management and processing
  - Distributed simulation, visualization, and computational steering
  - Collaboration with the US and International Research and Education community

- “Enabling large-scale science” means ensuring that the network can be used effectively to provide all mission required access to data and computing

- ESnet connects the Office of Science National Laboratories and user facilities to each other and to collaborators worldwide
HEP as a Prototype for Data-Intensive Science

- The *history of high energy physics (HEP) data management and analysis anticipates many other science disciplines*
  - Each new generation of experimental science requires more complex instruments to ferret out more and more subtle aspects of the science
  - As the sophistication, size, and cost of the instruments increase, the number of such instruments becomes smaller, and the collaborations become larger and more widely distributed – and mostly international
  - These new instruments are based on increasingly sophisticated sensors, which now are largely solid-state devices akin to CCDs
    - In many ways, the solid-state sensors follow Moore’s law just as computer CPUs do: The number of transistors doubles per unit area of silicon every 18 mo., and therefore the amount of data coming out doubles per unit area
    - the data output of these increasingly sophisticated sensors has increased exponentially
    - Large scientific instruments only differ from CPUs in that the time between science instrument refresh is more like 10-20 years, and so the increase in data volume from instrument to instrument is huge
HEP as a Prototype for Data-Intensive Science

High Energy Physics estimates of their data transport requirements

- may predict data transport needs a bit sooner than it actually show up,
- but the volume projections (and therefore eventual transport needs) tend to be accurate

Data courtesy of Harvey Newman, Caltech, and Richard Mount, SLAC and Belle II CHEP 2012 presentation
HEP as a Prototype for Data-Intensive Science

- What is the significance to the network of this increase in data?
- Historically, the use of the network has tracked the size of the data sets used by HEP

“HEP data collected” 2014 estimate (green line) in previous slide
HEP as a Prototype for Data-Intensive Science

- Two major proton experiments (detectors) at the LHC: ATLAS and CMS
- ATLAS is designed to observe a billion \((1 \times 10^9)\) collisions/sec, with a data rate out of the detector of more than 1,000,000 Gigabytes/sec (1 PBy/s)
- A set of hardware and software filters at the detector reduce the output data rate to about 25 Gb/s that must be transported, managed, and analyzed to extract the science
  - The output data rate for CMS is about the same, for a combined 50 Gb/s that is distributed to physics groups around the world, 7x24x~9mo/yr.
The LHC data management model involves a world-wide collection of centers that store, manage, and analyze the data.

A Network Centric View of the LHC
(one of two detectors)

<table>
<thead>
<tr>
<th>Tier 1 centers</th>
<th>Tape 115 PBy</th>
<th>Disk 60 PBy</th>
<th>Cores 68,000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tier 2 centers</td>
<td>0</td>
<td>120 PBy</td>
<td>175,000</td>
</tr>
</tbody>
</table>

(WLCG 2012)

This is intended to indicate that the physics groups now get their data wherever it is most readily available.
The PanDA jobs, executing at centers all over Europe, N. America and SE Asia, generate network data movement of 730 TB/day, ~68Gb/s.

PanDA manages 120,000–140,000 simultaneous jobs (PanDA manages two types of jobs that are shown separately here.)

- It is this scale of data movement going on 24 hr/day, 9+ months/yr, that networks must support in order to enable the large-scale science of the LHC.
HEP as a Prototype for Data-Intensive Science

- ESnet has been collecting requirements for all DOE science disciplines and instruments that rely on the network for distributed data management and analysis for more than a decade, and formally since 2007 [REQ]

  - In this process, certain issues are seen across essentially all science disciplines that rely on the network for significant data transfer, even if the quantities are modest compared to project like the LHC experiments.

Therefore addressing the LHC issues is a useful exercise that can benefit a wide range of science disciplines.
Foundations of data-intensive science

• The capabilities required to support HEP’s scale of data movement involve hardware and software developments at all levels:
  1. The underlying network
     1a. Optical signal transport
     1b. Network routers and switches
  2. Data transport (TCP is a “fragile workhorse” but still the norm)
  3. Network monitoring and testing
  4. Operating system evolution
  5. New site and network architectures
  6. Data movement and management techniques and software
  7. New network services
  8. Knowledge base
  9. Authentication and authorization
  10. Data integrity during transport

• Technology advances in these areas have resulted in today’s state-of-the-art that makes it possible for the LHC experiments to routinely and continuously move data at ~150 Gb/s across three continents
Foundations of data-intensive science

• This talk looks briefly at the nature of the advances in technologies, software, and methodologies that have enabled LHC data management and analysis
  ➢ The points 1a and 1b on optical transport and router technology are included in the slides for completeness but I will not talk about them. They were not really driven by the needs of the LHC but they were opportunistically used by the LHC.
  ➢ Much of the reminder of the talk is a tour through ESnet’s network performance knowledge base (fasterdata.es.net)
    – Also included are
      • the LHC ATLAS data management and analysis approach that generates and relies on very large network data utilization
      • and an overview of how R&E network have evolved to accommodate the LHC traffic
1) Underlying network issues

At the core of our ability to transport the volume of data that we must deal with today, and to accommodate future growth, are advances in optical transport technology and router technology.

ESnet has seen exponential growth in our traffic every year since 1990 (our traffic grows by factor of 10 about once every 47 months).

We face a continuous growth of data to transport.

Projected volume for Feb 2015: 63.0 PB
Actual volume for Feb 2014: 14.0 PB
We face a continuous growth of data transport

• The LHC data volume is **predicated to grow 10 fold** over the next 10 years

➢ **New generations of instruments** – for example the Square Kilometer Array radio telescope and ITER (the international fusion experiment) – will generate more data than the LHC

➢ In response, ESnet, and most large R&E networks, have built **100 Gb/s (per optical channel)** networks
  
  – ESnet's new network – ESnet5 – is complete and provides a **44 x 100Gb/s (4.4 terabits/sec - 4400 gigabits/sec)** in optical channels across the entire ESnet national footprint
  
  – Initially, one of these **100 Gb/s channels** is configured to replace the current **4 x 10 Gb/s IP network**

• What has made this possible?
1a) Optical Network Technology

- Modern optical transport systems (DWDM = dense wave division multiplexing) use a collection of technologies called “coherent optical” processing to achieve more sophisticated optical modulation and therefore higher data density per signal transport unit (symbol) that provides 100Gb/s per wave (optical channel).
  - Optical transport using dual polarization-quadrature phase shift keying (DP-QPSK) technology with coherent detection [OIF1]
    - dual polarization
      - two independent optical signals, same frequency, orthogonal
    - two polarizations → reduces the symbol rate by half
    - quadrature phase shift keying
      - encode data by changing the signal phase of the relative to the optical carrier
      - further reduces the symbol rate by half (sends twice as much data / symbol)
    - Together, DP and QPSK reduce required rate by a factor of 4
      - allows 100G payload (plus overhead) to fit into 50GHz of spectrum
    - Actual transmission rate is about 10% higher to include FEC data
      - This is a substantial simplification of the optical technology involved – see the TNC 2013 paper and Chris Tracy’s NANOG talk for details [Tracy1] and [Rob1]
ESnet5’s optical network uses Ciena’s 6500 Packet-Optical Platform with WaveLogic™ to provide 100Gb/s wave

- 88 waves (optical channels), 100Gb/s each
  - wave capacity shared equally with Internet2
    - ~13,000 miles / 21,000 km lit fiber
    - 280 optical amplifier sites
    - 70 optical add/drop sites (where routers can be inserted)
      - 46 100G add/drop transponders
      - 22 100G re-gens across wide-area

Geography is only representative.
1b) Network routers and switches

- Routers also use the latest in high-speed electronics
- ESnet5 routing (IP / layer 3) is provided by Alcatel-Lucent 7750 routers with 100 Gb/s client interfaces
  - 2 Tb/s backplane
  - 100 Gb/s throughput per interface
    - IP, MPLS, Ethernet services
    - 64,000 queues per module, 8 queues per subscriber
    - 3,000,000 routes (ESnet routers currently have about 500,000)
- In ESnet continental U.S. network
  - 17 routers with 100G interfaces
    - several more in a test environment
  - 59 layer-3 100GigE interfaces
  - 8 Lab-owned 100G routers
  - 7 100G interconnects with other R&E networks at Starlight (Chicago), MAN LAN (New York), and Sunnyvale (San Francisco)
Geography is only representational

ESnet Winter 2014/15

ESnet's Extension to Europe

SUNN

ESnet PoP/hub locations
ESnet managed 100G routers
R&E network peering locations – US (red) and international (green)
commercial peering points

Routed IP 100 Gb/s
Routed IP 40 Gb/s
Express / metro / regional

SUNN

EEX
ESnet’s Extension to Europe
2) Data transport: The limitations of TCP must be addressed for large, long-distance flows

Although there are other transport protocols available, TCP remains the workhorse of the Internet, including for data-intensive science

- Using TCP to support the sustained, long distance, high data-rate flows of data-intensive science requires an error-free network

- Why error-free?
  - TCP is a “fragile workhorse”: It is very sensitive to packet loss (due to bit errors)
    - Very small packet loss rates on these paths result in large decreases in performance
    - A single bit error will cause the loss of a 1-9 KBy packet (depending on the MTU size) as there is no FEC at the IP level for error correction
      - This puts TCP back into “slow start” mode thus reducing throughput

- TCP’s 16 bit checksum is too weak and too low in the stack to protect large data transfers end-to-end (see point 10 on data integrity, below)
Transport

• The reason for TCP’s sensitivity to packet loss is that the slow-start and congestion avoidance algorithms that were added to TCP to prevent congestion collapse of the Internet
  – Packet loss is seen by TCP’s congestion control algorithms as evidence of congestion, so they activate to slow down and prevent the synchronization of the senders (which perpetuates and amplifies the congestion, leading to network throughput collapse)
  – Network link errors also cause packet loss, so these congestion avoidance algorithms come into play, with dramatic effect on throughput in the wide area network – hence the need for “error-free”
Transport: Impact of packet loss on TCP

- On a 10 Gb/s LAN path the impact of low packet loss rates is minimal
- On a 10 Gb/s WAN path the impact of low packet loss rates is enormous (~80X throughput reduction on transatlantic path)

Throughput vs. increasing latency on a 10 Gb/s link with 0.0046% packet loss

(see http://fasterdata.es.net/performance-testing/perfsonar/troubleshooting/packet-loss/)
Transport: Modern TCP stack

- A modern TCP stack (the kernel implementation of the TCP protocol) is important to reduce the sensitivity to packet loss while still providing congestion avoidance (see [HPBulk])
  - This is done using mechanisms that more quickly increase back to full speed after an error forces a reset to low bandwidth

![Graph showing TCP Results]

Note that BIC reaches max throughput much faster than older algorithms (from Linux 2.6.19 the default is CUBIC, a refined version of BIC designed for high bandwidth, long paths)

RTT = 67 ms
Transport: Modern TCP stack

- Even modern TCP stacks are only of some help in the face of packet loss on a long path, high-speed network.

![Graph showing Throughput vs. Increasing Latency with .0046% Packet Loss (zoom of tail)](image)

3) Monitoring and testing

The only way to keep multi-domain, international scale networks error-free is to test and monitor continuously end-to-end to detect soft errors and facilitate their isolation and correction

- perfSONAR provides a standardized way to test, measure, export, catalogue, and access performance data from many different network domains (service providers, campuses, etc.)

• perfSONAR is a community effort to
  - define network management data exchange protocols, and
  - standardized measurement data formats, gathering, and archiving

- perfSONAR is deployed extensively throughout LHC related networks and international networks and at the end sites (See [fasterdata], [perfSONAR], and [NetSrv])
  - There are now more than 1000 perfSONAR boxes installed in N. America and Europe
The test and monitor functions can detect soft errors that limit throughput and can be hard to find (hard errors / faults are easily found and corrected).

Soft failure example:
- Observed end-to-end performance degradation due to soft failure of single optical line card

- Why not just rely on “SNMP” interface stats for this sort of error detection?
  - not all error conditions show up in SNMP interface statistics
  - SNMP error statistics can be very noisy
  - some devices lump different error counters into the same bucket, so it can be very challenging to figure out what errors to alarm on and what errors to ignore
    - though ESnet’s Spectrum monitoring system attempts to apply heuristics to do this
  - many routers will silently drop packets - the only way to find that is to test through them and observe loss using devices other than the culprit device
The value of perfSONAR increases dramatically as it is deployed at more sites so that more of the end-to-end (app-to-app) path can characterized across multiple network domains

- provides the only widely deployed tool that can monitor circuits end-to-end across the different networks from the US to Europe
  - ESnet has perfSONAR testers installed at every PoP and all but the smallest user sites – Internet2 is close to the same
  - There are currently more that 1200 perfSONAR systems deployed world-wide

- perfSONAR comes out of the work of the Open Grid Forum (OGF) Network Measurement Working Group (NM-WG) and the protocol is implemented using SOAP XML messages
4) System software evolution and optimization

Once the network is error-free, there is still the issue of efficiently moving data from the application running on a user system onto the network

- Host TCP tuning
- Modern TCP stack (see above)
- Other issues (MTU, etc.)
- Data transfer tools and parallelism
- Other data transfer issues (firewalls, etc.)
4.1) System software tuning: Host tuning – TCP

- “TCP tuning” commonly refers to the proper configuration of TCP windowing buffers for the path length.
- It is critical to use the optimal TCP send and receive socket buffer sizes for the path (RTT) you are using end-to-end.

Default TCP buffer sizes are typically much too small for today’s high speed networks:
  - Until recently, default TCP send/receive buffers were typically 64 KB.
  - Tuned buffer to fill CA to NY, 1 Gb/s path: 10 MB
    - 150X bigger than the default buffer size.
System software tuning: Host tuning – TCP

• Historically TCP window size tuning parameters were host-global, with exceptions configured per-socket by applications
  – How to tune is a function of the application and the path to the destination, so potentially a lot of special cases

➢ Auto-tuning TCP connection buffer size within pre-configured limits helps

➢ Auto-tuning, however, is not a panacea because the upper limits of the auto-tuning parameters are typically not adequate for high-speed transfers on very long (e.g. international) paths
System software tuning: Host tuning – TCP

Throughput out to ~9000 km on a 10Gb/s network
32 MB by (auto-tuned) vs. 64 MB by (hand tuned) TCP window size

Effect of Window Size on Throughput vs. Increasing Latency

Roundtrip time, ms
(90 ms corresponds roughly to San Francisco to London)
4.2) System software tuning: Data transfer tools

- Parallelism is key in data transfer tools
  - It is much easier to achieve a given performance level with multiple parallel connections than with one connection
    - this is because the OS is very good at managing multiple threads and less good at sustained, maximum performance of a single thread (same is true for disks)
  - Several tools offer parallel transfers (see below)

- Latency tolerance is critical
  - Wide area data transfers have much higher latency than LAN transfers
  - Many tools and protocols assume latencies typical of a LAN environment (a few milliseconds)
    - examples: SCP/SFTP and HPSS mover protocols work very poorly in long path networks
  - Disk Performance
    - In general need a RAID array or parallel disks (like FDT) to get more than about 500 Mb/s
System software tuning: Data transfer tools

- Using the right tool is very important

- Sample Results: Berkeley, CA to Argonne, IL
  RTT = 53 ms, network capacity = 10Gbps.

<table>
<thead>
<tr>
<th>Tool</th>
<th>Throughput</th>
</tr>
</thead>
<tbody>
<tr>
<td>scp</td>
<td>140 Mbps</td>
</tr>
<tr>
<td>patched scp (HPN)</td>
<td>1.2 Gbps</td>
</tr>
<tr>
<td>ftp</td>
<td>1.4 Gbps</td>
</tr>
<tr>
<td>GridFTP, 4 streams</td>
<td>5.4 Gbps</td>
</tr>
<tr>
<td>GridFTP, 8 streams</td>
<td>6.6 Gbps</td>
</tr>
</tbody>
</table>

Note that to get more than about 1 Gbps (125 MB/s) disk to disk requires using RAID technology

- PSC (Pittsburgh Supercomputer Center) has a patch set that fixes problems with SSH
  - [http://www.psc.edu/networking/projects/hpn-ssh/](http://www.psc.edu/networking/projects/hpn-ssh/)
  - Significant performance increase
    - this helps rsync too
System software tuning: Data transfer tools

- Globus GridFTP is the basis of most modern high-performance data movement systems
  - Parallel streams, buffer tuning, help in getting through firewalls (open ports), ssh, etc.
  - The newer Globus Online incorporates all of these and small file support, pipelining, automatic error recovery, third-party transfers, etc.
    - This is a very useful tool, especially for the applications community outside of HEP
System software tuning: Data transfer tools

- Also see Caltech's FDT (Faster Data Transfer) approach
  - Not so much a tool as a hardware/software system designed to be a very high-speed data transfer node
  - Explicit parallel use of multiple disks
  - Can fill 100 Gb/s paths
4.4) System software tuning: Other issues

- Firewalls are anathema to high-speed data flows
  - many firewalls can’t handle >1 Gb/s flows
    - designed for large number of low bandwidth flows
    - some firewalls even strip out TCP options that allow for TCP buffers > 64 KB
  - See Jason Zurawski’s “Say Hello to your Frienemy – The Firewall”
    - Stateful firewalls have inherent problems that inhibit high throughput
      - http://fasterdata.es.net/assets/fasterdata/Firewall-tcptrace.pdf

- Many other issues
  - Large MTUs (several issues)
  - NIC tuning
    - Defaults are usually fine for 1GE, but 10GE often requires additional tuning
  - Other OS tuning knobs
  - See fasterdata.es.net and “High Performance Bulk Data Transfer” ([HPBulk])
5) Site infrastructure to support data-intensive science

The Science DMZ

With the wide area part of the network infrastructure addressed, the typical site/campus LAN becomes the bottleneck

- The site network (LAN) typically provides connectivity for local resources – compute, data, instrument, collaboration system, etc. – needed by data-intensive science
  - Therefore, a high performance interface between the wide area network and the local area/site network is critical for large-scale data movement

- Campus network infrastructure is typically not designed to handle the flows of large-scale science
  - The devices and configurations typically deployed to build LAN networks for business and small data-flow purposes usually don’t work for large-scale data flows
    - firewalls, proxy servers, low-cost switches, and so forth
    - none of which will allow high volume, high bandwidth, long distance data flows
The Science DMZ

➢ To provide high data-rate access to local resources the site **LAN infrastructure must be re-designed** to match the high-bandwidth, large data volume, high round trip time (RTT) (international paths) of the wide area network (WAN) flows (See [DIS])
  - otherwise the site will impose poor performance on the entire high speed data path, all the way back to the source
The Science DMZ

The ScienceDMZ concept:
The compute and data resources involved in data-intensive sciences should be deployed in a separate portion of the site network that has a different packet forwarding path

- Outside the site firewall – hence the term “ScienceDMZ”
- WAN-like router/switch technology
- Dedicated systems built and tuned for wide-area data transfer
- Test and measurement systems for performance verification and rapid fault isolation, typically perfSONAR (see [perfSONAR] and below)
- A security policy tailored for science traffic and implemented using appropriately capable hardware (e.g. that supports access control lists, private address space, etc.)

This usually results in large increases in data throughput and is so important it was a requirement for last round of NSF CC-NIE grants
The Science DMZ

(See http://fasterdata.es.net/science-dmz/ and [SDMZ] for a much more complete discussion of the various approaches.)
6) Data movement and management techniques

Automated data movement is critical for moving 500 terabytes/day between 170 international sites

- In order to effectively move large amounts of data over the network, automated systems must be used to manage workflow and error recovery
  - The filtered ATLAS data rate of about 25 Gb/s is sent to 10 national Tier 1 data centers
  - The Tier 2 sites get a comparable amount of data from the Tier 1s
    - Host the physics groups that analyze the data and do the science
    - Provide most of the compute resources for analysis
    - Cache the data (though this is evolving to remote I/O)
Highly distributed and highly automated workflow systems are central to data-intensive science

- The ATLAS experiment system (PanDA) coordinates/automates the analysis jobs and resources, and the data management
  - The resources and data movement are centrally managed
  - Analysis jobs are submitted to the central manager that locates compute resources and matches these with dataset locations
  - The system manages several million jobs a day
    - coordinates data movement of hundreds of terabytes/day, and
    - manages (analyzes, generates, moves, stores) of order 10 petabytes of data/year in order to accomplish its science

- The complexity of the distributed systems that have to coordinate the computing and data movement for data analysis at the hundreds of institutions spread across three continents involved in the LHC experiments is substantial
The ATLAS PanDA “Production and Distributed Analysis” system uses distributed resources and layers of automation to manage several million jobs/day.

ATLAS Tier 0 Data Center (1 copy of all data – archival only)

ATLAS Tier 1 Data Centers – 11 sites scattered across Europe, North America and Asia, hold a copy of all aggregate data and provide the working data set for analysis.

ATLAS analysis sites (e.g. 70 Tier 2 Centers in Europe, North America and SE Asia)

ATLAS production jobs

Regional production jobs

User / Group analysis jobs

Task Buffer (job queue)

Job Broker

Job Dispatcher

Data Service

Policy (job type priority)

PanDA Server (task management)

1) Schedules jobs, initiates data movement

2) DDM locates data and moves it to sites. This is a complex system in its own right called DQ2.

3) Prepares the local resources to receive Panda jobs

4) Jobs are dispatched when there are resources available and when the required data is in place at the site.

Try to move the job to where the data is, else move data and job to where resources are available.

Site Capability Service

Job resource manager:
- Dispatch a “pilot” job manager - a Panda job receiver - when resources are available at a site
- Pilots run under the local site job manager (e.g. Condor, LSF, LCG,…) and accept jobs in a standard format from PanDA
- Similar to the Condor Glide-in approach

CERN

ATLAS detector

Tier 0 Data Center

Tier 2 Data Centers

Tier 0 Data Center (1 copy of all data – archival only)

DDM Agent

DDM Agent

Agent

Agent

Thanks to Michael Ernst, US ATLAS technical lead, for his assistance with this diagram, and to Torre Wenaus, whose view graphs provided the starting point. Both are at Brookhaven National Lab.)

Grid Scheduler

Pilot Job (Panda job receiver running under the site-specific job manager)
The PanDA jobs, executing at centers all over Europe, N. America and SE Asia, generate network data movement of 730 TBy/day, ~68Gb/s.

PanDA manages 120,000–140,000 simultaneous jobs.

(PanDA manages two types of jobs that are shown separately here.)

It is this scale of data movement going on 24 hr/day, 9+ months/yr, that networks must support in order to enable the large-scale science of the LHC.
In order to debug and optimize the distributed system that accomplishes the scale of the ATLAS analysis, years were spent building and testing the required software and hardware infrastructure:

- Once the systems were in place, systematic testing was carried out in “service challenges” or “data challenges”
- Successful testing was required for sites to participate in LHC production
Ramp-up of LHC traffic in ESnet

The transition from testing to operation was a smooth continuum due to at-scale testing – a process that took more than 5 years.

ESnet Accepted Traffic: Jan 2000 - Aug 2012
Petabytes/Month, Maximum Volume: 11.7 PB

- Traffic Accepted
- OSCARS Accepted
- Top 1000 Host-Host Accepted

Aug 2012: 11.7 PB
LHC operation
LHC turn-on
LHC data system testing
(est. of "small" scale traffic)
For sustained high data-rate transfers – e.g. from instrument to data centers – a dedicated, purpose-built infrastructure is needed

- The transfer of LHC experiment data from CERN (Tier 0) to the 11 national data centers (Tier 1) uses a network called **LHCOPN**
  - The LHCOPN is a collection of leased 10Gb/s optical circuits
- The role of LHCOPN is to ensure that all data moves from CERN to the national Tier 1 data centers continuously
  - In addition to providing the working dataset for the analysis groups, the Tier 1 centers, in aggregate, hold a duplicate copy of the data that is archived at CERN
• While the LHCOPN was a technically straightforward exercise – establishing 10 Gb/s links between CERN and the Tier 1 data centers for distributing the detector output data – there were several aspects that were new to the R&E community

• The issues related to the fact that most sites connected to the R&E WAN infrastructure through a site firewall and the OPN was intended to bypass site firewalls in order to achieve the necessary performance

➢ The security issues were the primarily ones and were addressed by

• Using a private address space that hosted only LHC Tier 1 systems (see [LHCOPN Sec])

  – that is, only LHC data and compute servers are connected to the OPN
The LHC OPN – Optical Private Network

N.B.

• In 2005 the only way to handle the CERN (T0) to Tier 1 centers data transfer was to use dedicated, physical, 10G circuits

➢ Today, in most R&E networks (where 100 Gb/s links are becoming the norm), the LHCOPN could be provided using virtual circuits implemented with MPLS or OpenFlow network overlays
  – The ESnet part of the LHCOPN has always used this— in fact this is what ESnet’s OSCARS virtual circuit system was originally designed for (see below)
  – As ESnet’s European extension comes on line later this year, the entire length of the CERN to BNL (US ATLAS Tier 1 Data Center) and CERN to FNAL (US CMS Tier 1) LHCOPN paths will be OSCARS virtual circuit based
Managing large-scale science traffic in a shared infrastructure

- The traffic from the Tier 1 data centers to the Tier 2 sites (mostly universities) where the data analysis is done is now large enough that it must be managed separately from the general R&E traffic.
  - In aggregate the Tier 1 to Tier 2 traffic is equal to the Tier 0 to Tier 1
  - (there are about 170 Tier 2 sites)

- Managing this with all possible combinations of Tier 2 – Tier 2 flows (potentially 170 x 170) cannot be done just using a virtual circuit service – it is a relatively heavy-weight mechanism.

- Special infrastructure is required for this: The LHC’s Open Network Environment – LHCONE – was designed for this purpose.
The LHC’s Open Network Environment – LHCONE

- LHCONE provides a private, managed infrastructure designed for LHC Tier 2 traffic (and likely other large-data science projects in the future)
  
  - The approach is a VRF-based overlay network whose architecture is a collection of routed “clouds” using address spaces restricted to subnets that are used by LHC systems
    - The clouds are mostly local to a network domain (e.g. one for each involved domain – ESnet, GEANT ("fronts" for the NRENs), Internet2 (fronts for the US universities), etc.
    - The clouds (VRFs) are interconnected by point-to-point circuits provided by various entities (mostly the domains involved)
  
  - In this way the LHC traffic will use circuits designated by the network engineers
    - To ensure continued good performance for the LHC and to ensure that other traffic is not impacted – this is critical because apart from the LHCOPN, the R&E networks are funded for the benefit of the entire R&E community, not just the LHC

- Uses the same security model as the LHCOPN – segregated systems in a “private” address space
LHCONE: A global infrastructure for the LHC Tier 1 data center – Tier 2 analysis center connectivity

End sites – LHC Tier 2 or Tier 3 unless indicated as Tier 1
Regional R&E communication nexus
Data communication links, 10, 20, and 30 Gb/s
See http://lhcone.net for details.
The LHC’s Open Network Environment – LHCONE

- LHCONE could be set up relatively “quickly” because
  - The VRF technology is a standard capability in most core routers, and
  - there is capacity in the R&E community that can be made available for use by the LHC collaboration that cannot be made available for general R&E traffic

- LHCONE is essentially built as a collection of private overlay networks (like VPNs) that are interconnected by managed links to form a global infrastructure where Tier 2 traffic will get good service and not interfere with general traffic

- From the point of view of the end sites, they see a LHC-specific environment where they can reach all other LHC sites with good performance

- See LHCONE.net
LHCONE is one part of the network infrastructure that supports the LHC

A Network Centric View of the LHC

<table>
<thead>
<tr>
<th>CERN → T1</th>
<th>miles</th>
<th>kms</th>
</tr>
</thead>
<tbody>
<tr>
<td>France</td>
<td>350</td>
<td>565</td>
</tr>
<tr>
<td>Italy</td>
<td>570</td>
<td>920</td>
</tr>
<tr>
<td>UK</td>
<td>625</td>
<td>1000</td>
</tr>
<tr>
<td>Netherlands</td>
<td>625</td>
<td>1000</td>
</tr>
<tr>
<td>Germany</td>
<td>700</td>
<td>1185</td>
</tr>
<tr>
<td>Spain</td>
<td>850</td>
<td>1400</td>
</tr>
<tr>
<td>Nordic</td>
<td>1300</td>
<td>2100</td>
</tr>
<tr>
<td>USA – New York</td>
<td>3900</td>
<td>6300</td>
</tr>
<tr>
<td>USA - Chicago</td>
<td>4400</td>
<td>7100</td>
</tr>
<tr>
<td>Canada – BC</td>
<td>5200</td>
<td>8400</td>
</tr>
<tr>
<td>Taiwan</td>
<td>6100</td>
<td>9850</td>
</tr>
</tbody>
</table>

The LHC Open Network Environment (LHCONE)

This is intended to indicate that the physics groups now get their data wherever it is most readily available.
7) New network services

Point-to-Point Virtual Circuit Service

Why a Circuit Service?

- Geographic distribution of resources is seen as a fairly consistent requirement across the large-scale sciences in that they use distributed applications systems in order to:
  - Couple existing pockets of code, data, and expertise into “systems of systems”
  - Break up the task of massive data analysis and use data, compute, and storage resources that are located at the collaborator’s sites
  - See https://www.es.net/about/science-requirements

- A commonly identified need to support this is that networking must be provided as a “service”
  - Schedulable with guaranteed bandwidth – as is done with CPUs and disks
  - Traffic isolation that allows for using non-standard protocols that will not work well in a shared infrastructure
  - Some network path characteristics may also be specified – e.g. diversity
  - Available in Web Services / Grid Services paradigm
The way that networks provide such a service is with “virtual circuits” (also called pseudowires) that emulate point-to-point connections in a packet-switched network like the Internet.

- This is typically done by using a “deterministic” routing mechanism:
  - E.g. some variation of label based switching, with the static switch tables set up in advance to define the circuit path
    - MPLS and OpenFlow are examples of this, and both can transport IP packets
  - Most modern Internet routers support this type of functionality

- Such a service channels big data flows into virtual circuits in ways that also allow network operators to do “traffic engineering” — that is, to manage/optimize the use of available network resources and to keep big data flows separate from general traffic
  - The virtual circuits can be directed to specific physical network paths when they are set up
**Point-to-Point Virtual Circuit Service**

- OSCARS is ESnet’s implementation of a virtual circuit service (For more information contact the project lead: Chin Guok, chin@es.net)
- Has been in production service in ESnet for the past 7 years, or so
- OSCARS received a 2013 “R&D 100” award

- OSCARS is fully compliant with the OGF’s NSI v2 specification
End User View of Circuits – How They Use Them

• How are the circuits used?
  – End system to end system, IP
    • Almost never – very hard unless private address space used
      – Using public address space can result in leaking routes
      – Using private address space with multi-homed hosts risks allowing backdoors into secure networks
  – End system to end system, Ethernet (or other) over VLAN – a pseudowire
    • Relatively common
    • Interesting example: RDMA over VLAN likely to be popular in the future
      – SC11 demo of 40G RDMA over WAN was very successful
      – CPU load for RDMA is a small fraction that of IP
      – The guaranteed network of circuits (zero loss, no reordering, etc.) required by non-IP protocols like RDMA fits nicely with circuit services (RDMA performs very poorly on best effort networks)
  – Point-to-point connection between routing instance – e.g. BGP at the end points: the “user” is a site
    • Essentially this is how all current circuits are used: from one site router to another site router
      – Typically site-to-site or advertise subnets that host clusters, e.g., LHC analysis or data management clusters
End User View of Circuits – How They Use Them

• When are the circuits used?
  – Mostly to solve a specific problem that the general infrastructure cannot
    • Most circuits are used for a guarantee of bandwidth or for user traffic engineering
Cross-Domain Virtual Circuit Service

• Large-scale science always involves institutions in multiple network domains (administrative units)
  – For a circuit service to be useful it must operate across all R&E domains involved in the science collaboration to provide end-to-end circuits
  – e.g. ESnet, Internet2 (USA), CANARIE (Canada), GÉANT (EU), SINET (Japan), CERNET and CSTNET (China), KREONET (Korea), TWAREN (Taiwan), AARNet (AU), the European NRENs, the US Regionals, etc. are all different domains

• LHCONE Architecture Working Group Point-to-Point experiment is pushing and debugging the technology for NSI multi-domain P2P circuits
Point-to-Point Virtual Circuit Service

Open Grid Forum’s Network Services Interface (NSI)

- Testing is being coordinated in GLIF (Global Lambda Integrated Facility - an international virtual organization that promotes the paradigm of lambda networking)
- The LHCONE Architecture working group is conducting an experimental deployment in the LHCONE community

Network Service Interface in a Nut Shell

GEC 19, Atlanta, GA

Presenter: Chin Guok (ESnet)
Contributors: Tomohiro Kudoh (AIST), John MacAuley (ESnet), Inder Monga (ESnet), Guy Roberts (DANTE), Jerry Sobieski (NORDUnet)
1. "Network Service Interface" is a framework for inter-domain service coordination

Examples:
- Connection Service (NSI-CS)
- Topology Service (NSI-TS)
- Discovery Service (NSI-DS)
- Switching Service (NSI-SS)
- Monitoring Service
- Protection Service
- Verification Service
- Etc.
2. Designed for flexible, multi-domain, service chaining

Supports Tree and Chain model of service chaining

Fits in well with Cloud/Compute model of provisioning as well as Network/GMPLS model
3. Principles of Abstraction applied – to network layers, technologies and domains

Service Termination Points (STP) and Service Demarcation Points (SDP) are abstract and technology independent.
NSI Connection Service (v2.0)

- NSI is an advance-reservation based protocol
  - A reservation of a connection has properties such:
    - A-point, Z-point (mandatory)
    - Start-time, End-time (optional*)
    - Bandwidth, Labels (optional)
- A reservation is made in **two-phase**
  - First phase: availability is checked, if available resources are held
  - Second phase: the requester either commit or abort a held reservation
  - Two-phase is convenient when a requester requests resources from multiple providers, including other resources such as computers and storages
  - Timeout: If a requester does not commit a held reservation for a certain period of time, a provider can timeout
- **Modification** of a reservation is supported.
  - Currently, modification of start_time, end_time and bandwidth are supported

*NB: Restricted to PA policies*
NSI Service Type and Definition

- Introduction of Service Type and Service Definition removes the dependencies of service specification from the core NSI CS protocol.
- This allows the NSI CS protocol to remain stable while permitting changes to the services offered by NSA within the network.
- Abstraction of physical properties of the underlying data plane can be achieved by the Service Definition.

**Common service**

The providers need to agree among themselves the service they wish to offer to the customer. For example they may wish to offer an Ethernet VLAN Transport Service (EVTS). The service must be common to all providers and all providers must agree in advance a minimum service level that they are all able to meet.
NSI NSA Implementations

- **AutoBAHN** – GÉANT (Poznan, PL)
- **BoD** - SURFnet (Amsterdam, NL)
- **DynamicKL** – KISTI (Daejeon, KR)
- **G-LAMBDA-A** - AIST (Tsukuba, JP)
- **G-LAMBDA-K** – KDDI Labs (Fujimino, JP)
- **OpenNSA** – NORDUnet (Copenhagen, DK)
- **OSCARS** – ESnet (Berkeley, US)
OGF NSI Information

- OGF NSI Working Group Site
- NSI Project Page
  - [https://code.google.com/p/ogf-nsi-project/](https://code.google.com/p/ogf-nsi-project/)
- NSI Documents
- NSI Co-Chairs
  - Guy Roberts <guy.roberts@dante.net>
  - Inder Monga <imonga@es.net>
  - Tomohiro Kudoh <t.kudoh@aist.go.jp>
The Fasterdata Knowledge Base provides proven, operationally sound methods for troubleshooting and solving performance issues. For over 25 years, ESnet has operated an advanced research network with the goal of enabling the highest levels of performance for the Department of Energy (DOE) scientific community. During this time, our engineers have identified a common set of issues that hinder performance and we would like to share our experiences and findings in this knowledge base. Our solutions fall into five categories:

- Network Architecture, including the Science DMZ model
- Host Tuning
- Network Tuning
- Data Transfer Tools
- Network Performance Testing

as well as additional information and references contributed by the community.
The knowledge base

- http://fasterdata.es.net topics:
  - Network Architecture, including the Science DMZ model
  - Host Tuning
  - Network Tuning
  - Data Transfer Tools
  - Network Performance Testing
  - With special sections on:
    - Linux TCP Tuning
    - Cisco 6509 Tuning
    - perfSONAR Howto
    - Active perfSONAR Services
    - Globus overview
    - Say No to SCP
    - Data Transfer Nodes (DTN)
    - TCP Issues Explained

- fasterdata.es.net is a community project with contributions from several organizations
9) Authentication and authorization

- Authentication (AuthN) and authorization (AuthZ), collectively “AA” are critical in a multi-institution collaboration where resources are being shared.

- The LHC collaboration uses a PKI approach that has been developed and deployed over the past 15 years.

- Characteristics of the WLCG (Worldwide LHC Computing Grid) infrastructure include:
  - Multiple administrative organizations
  - Multiple service providers participate in a single transaction
  - Multiple authorities influence policy

And unless you can do AA in this sort of environment you cannot do the enormous data processing associated with global, data-intensive science.

/1/ See “WLCG Authentication and Authorization (certificate infrastructure) and its use,” Dave Kelsey (STFC - Rutherford Appleton Lab, GB) at https://indico.cern.ch/event/289680/
Authentication and authorization

- The LHC community uses a single, interoperating, global infrastructure based on Public Key certificate technology
  - Trust is obtained by using a common set of community standards as the basis for issuing certificates /2/
  - A fairly small number of authorities issue these identity certificates for authentication
  - PKI authorization certificates convey permission to access compute and data systems

/2/ The IGTF (Interoperable Global Trust Federation) is the community-based mechanism for establishing trust in a community the size of the LHC (actually considerably larger because IGTF serves many science communities: 100,000 users in more than 1000 different user communities, 89 national and regional identity authorities, major relying parties include EGI, PRACE, ESEDE, Open Science Grid, HPCI, wLHF, OGF, ....)
  - The IGTF – through its members – develops guidance, coordinates requirements, and harmonizes assurance levels, for the purpose for supporting trust between distributed IT infrastructures for research.
  - For the purpose of establishing and maintaining and identity federation service, the IGTF maintains a set of authentication profiles (APs) that specify the policy and technical requirements for a class of identity assertions and assertion providers. The member PMAs are responsible for accrediting authorities that issue identity assertions with respect to these profiles.
  - Each of the PMAs will accredit credential-issuing authorities (the Certificate Authorities) and document the accreditation policy and procedures.
Authentication and authorization

- When the AuthN problem is solved you still have to address authorization

- Even a “single” collaboration like the LHC (actually several collaboration that are organized around the several detectors) you still have to allocate and manage resource utilization
  - There may be common jobs – e.g. the track reconstruction – that everyone has to have to do any analysis, so these get high priority on available resources
  - Different physics groups have different analysis approaches, and so the collaboration will allocate resources among competing groups
    - AuthZ certificates will be issued to groups or users to let them “draw” against (use) the resources (CPUs and storage) that are allocated to them
    - Accounting (who has used what portion of their allocation) is done centrally
  - In a widely distributed resource environment (the CMS and ATLAS collaborations each have some 70-100 participating institutions world-wide that provide resources) it is not practical for a given user to use his AuthN cert to log in to each system that he might have an allocation on
    - Proxy certificates are used for this purpose
    - Proxy certs carry the user’s identity for a limited period of time and are sent with a computing job to a remote system for authorization to access and use that system
Authentication and authorization

• One way or another, all of the issues must be addressed for widely distributed collaborations doing data-intensive science
  – The climate science community uses a different AuthN approach
    • They use OpenID in which home institutions certify identity and then institutions that trust each other accept the identity tokens from other institutions in a series of bilateral agreements

• AA is just one of a set of issues to be solved before large-scale, data-intensive collaboration is possible
10) Data integrity during transport

Undetected data corruption in the network is a potentially non-trivial problem

- In 2000, Jonathan Stone and Craig Partridge, following up on work by Vern Paxson, analyzed packet streams at four large sites over a period of several months [TCPCKS]

- Their conclusion was that the 16 bit TCP “checksum will fail to detect errors for roughly 1 in 16 million to 10 billion packets” – that is, there will be checksummed packets that contain an undetected error

- Where are undetected corrupted packets generated? “we found that just about any factor that could cause a bad checksum periodically does. We found buggy software, and defective hardware, and problems in both end-systems and routers.”

- Their conclusion was that if data integrity is important for particular data sets then “the safest thing to do is checksum the data as early as possible in the transmission path: before the data can suffer DMA errors or data path errors in the network interface”
Data integrity during transport

• Many of the sources of errors that Stone and Partridge describe are in software and hardware that one would expect to have improved over the past 15 years

• However, even assuming that we now operate much more toward the lower end of the undetected error range - $10^{-10}$ (1 in 10 billion) - there is still the possibility of between .01 and .1 undetected packet errors per 100 terabytes of data – that is, between one in ten and one in a hundred 100 TBy transfers will have an undetected error
  – A study at CERN, where the concern is silent corruption in RAID disk arrays, detected 1000 instances of silent corruption in 41 petabytes of data movement [Kelemen2007]:
    • “About $1.2 \times 10^{-9}$ of the data written to CERN’s storage was permanently corrupted within six months” [Clarke2010]
    • “silent corruptions [of data] are a fact of life” and must be dealt with

• This is still a concern in the community that jumbo-frame Ethernet usage is increasing the risk of undetected link errors even with 32 bit CRCs
  – In 2010 a draft Internet standard (not acted on) recommended stronger error detection codes be used [IETFDraft1]
Data integrity during transport

• When using GridFTP there is an option for generating and verifying strong, end-to-end checksums (“--verify-checksum”)
  ➢ if your data use/analysis is sensitive to occasional errors, then this end-to-end checksum option should always be used
    – There are other tools available for protecting data in storage systems – see [Kelemen2007] and [Clarke2010]
The Message

- A significant collection of issues must *all* be addressed in order to achieve the sustained data movement needed to support data-intensive science such as the LHC experiments.
  - But once this is done, international high-speed data management can be done on a routine basis.

- Many of the technologies and knowledge from the LHC experience are applicable to other science disciplines that must manage a lot of data in a widely distributed environment – SKA, ITER, ……
Infrastructure Critical to Science

• The combination of
  – New network architectures in the wide area
  – New network services (such as guaranteed bandwidth virtual circuits)
  – Cross-domain network error detection and correction
  – Redesigning the site LAN to handle high data throughput
  – Automation of data movement systems
  – Use of appropriate operating system tuning and data transfer tools

now provides the LHC science collaborations with the data communications underpinnings for a unique large-scale, widely distributed, very high performance data management and analysis infrastructure that is an essential component in scientific discovery at the LHC

• Other disciplines that involve data-intensive science will face most of these same issues
References


[fasterdata] See http://fasterdata.es.net/fasterdata/perfSONAR/

[HPBulk] “High Performance Bulk Data Transfer,” Brian Tierney and Joe Metzger, ESnet. Joint Techs, July 2010. Available at fasterdata.es.net/fasterdata-home/learn-more


[Jacobson] For an overview of this issue see http://en.wikipedia.org/wiki/Network_congestion#History


[LHCOPN Sec] at https://twiki.cern.ch/twiki/bin/view/LHCOPN/WebHome see “LHCOPN security policy document”

References


References


http://www.perfsonar.net/
http://psps.perfsonar.net/

[REQ] https://www.es.net/about/science-requirements/


(may be available at http://staffweb.cms.gre.ac.uk/~gm73/com-mag/COMG_20100701_Jul_2010.PDF )
