Networks in ATLAS

Shawn McKee / University of Michigan

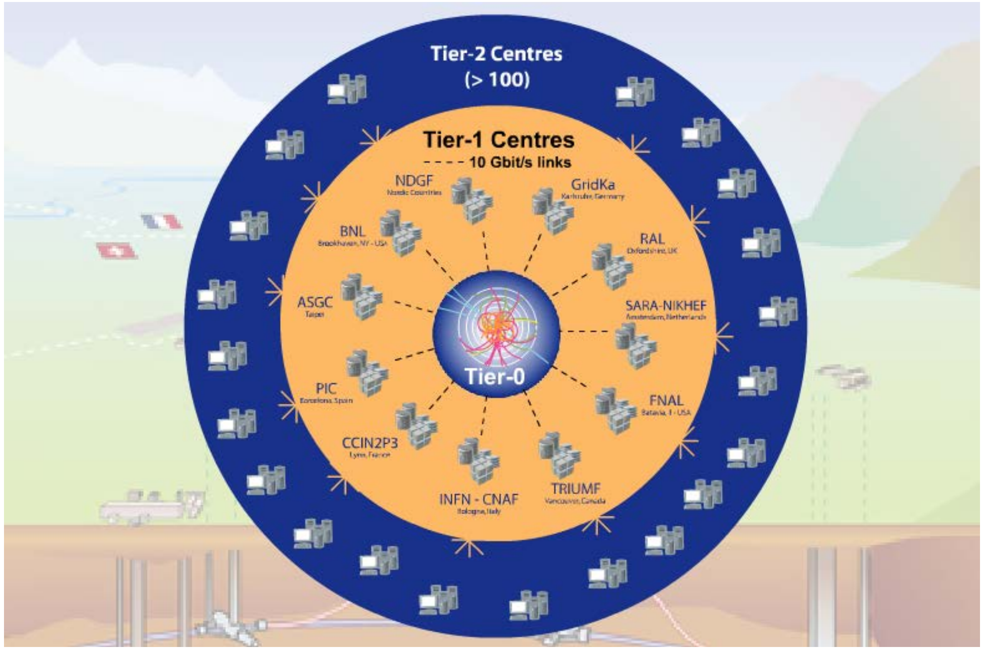
for the ATLAS Collaboration

CHEP 2016, San Francisco, October 10-14, 2016

Why Networking?

- Why talk about networking for ATLAS (a high-energy physics collaboration)?
  - Of course we know we need, and heavily utilize, the network but what are the concerns?
  - To-date the main need has been to better support diagnosing, localizing and repairing network problems
- I will review our status and recent activities
- Then I will cover how networking is evolving and what may be changing in mid-to-long-term

Distributed Computing in ATLAS



ATLAS Computing Model :

11 Clouds : 10 T1s + 1 T0 (CERN)

Cloud = T1 + T2s + T2Ds

T2D = multi-cloud T2 sites

2-16 T2s in each Cloud

PanDA Workload Management System

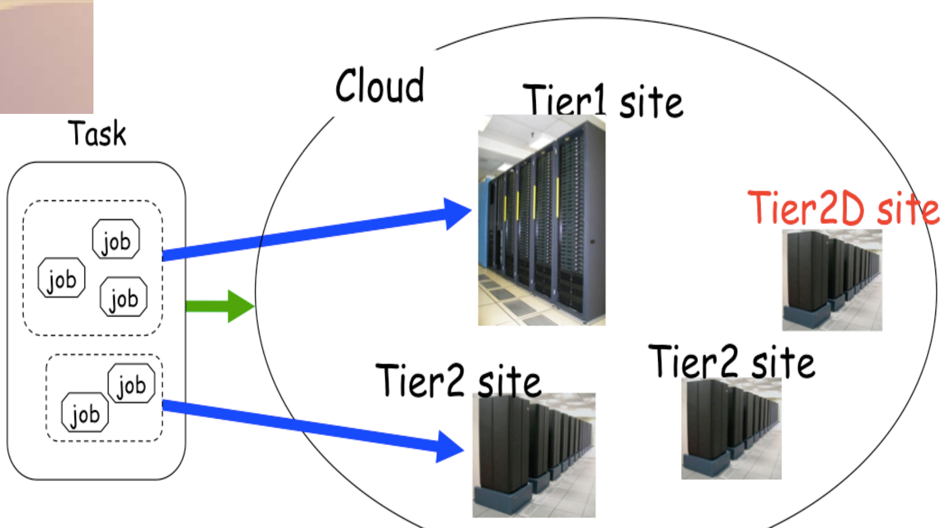
Task → Cloud :Task brokerage

Jobs → Sites :Job brokerage

- Basic unit of work is a job:
  - Executed on a CPU resource/slot
  - May have inputs; Produces outputs
- JEDI – layer above PanDA to create jobs from ATLAS physics and analysis 'tasks'

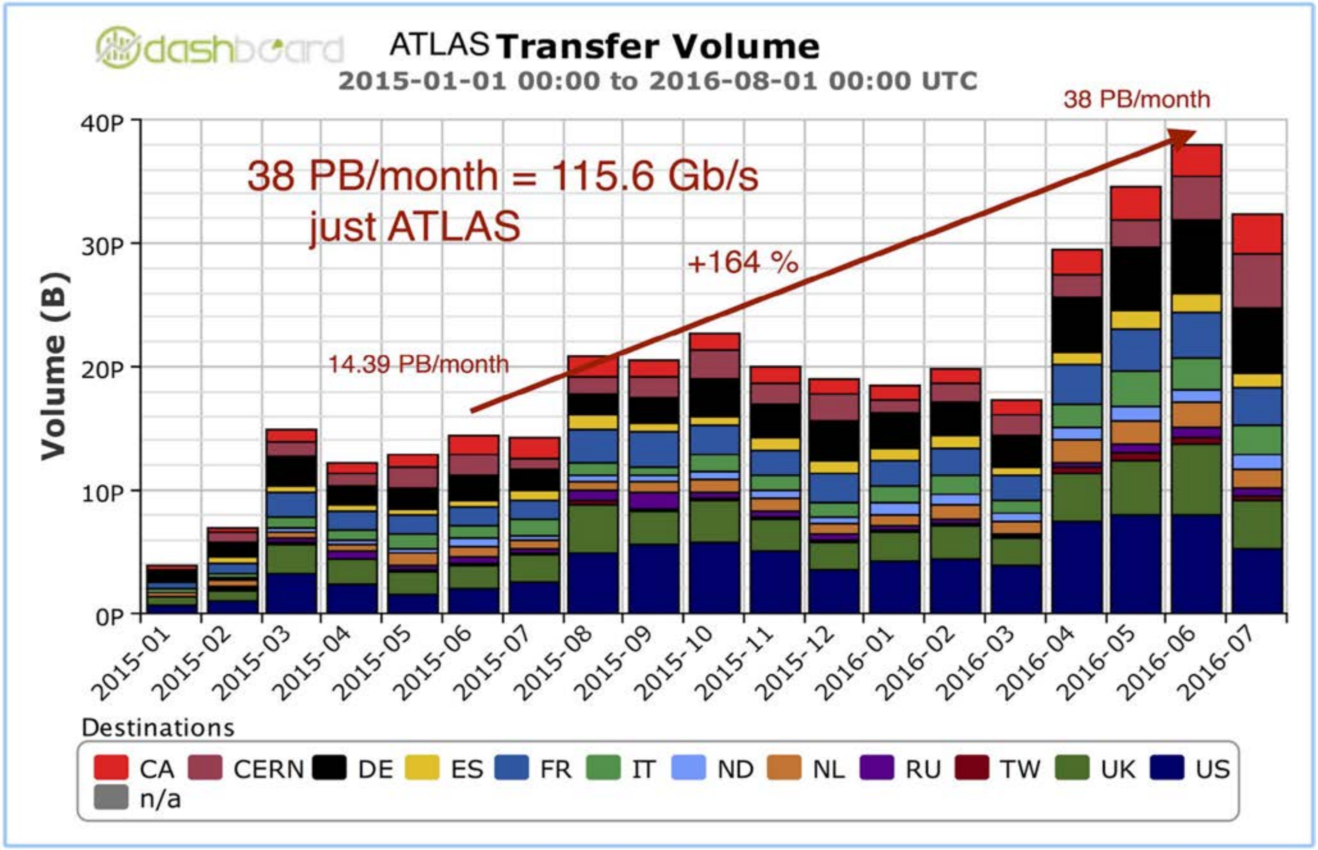
Current scale – one million jobs per day

The network ties this all together!



- 11 Clouds

Network Use in ATLAS



ATLAS (and LHC in general) has been transferring an exponentially increasing amount of data since startup. This trend is likely to continue and is driven by increasing data volumes, more capable infrastructures and the excellent networks supporting our needs.

Who is Working on Networks for ATLAS

- There has been a small but long-term effort in the area of networks for LHC HEP (High Energy Physics)
  - Initial efforts started around the Internet2 HENP working group in 2001
  - The LHCOPN (and follow-on LHCONE) effort started in 2005 and focused on defining the LHC experiment networking needs and implementing them. It continues to meet twice per year.
  - USATLAS piloted perfSONAR in 2006, expanding to LHCONE in 2010 and WLCG wide in 2012
  - Open Science Grid(OSG) began a network focus area in 2012
    - OSG now provides a network service for WLCG/OSG, gathering perfSONAR metrics worldwide and making the available
  - WLCG has had a task-force and a working group in networks
    - perfSONAR deployment task-force which got ~250 perfSONAR toolkit innstances deployed globally in 2013-2014
    - WLCG Network and Transfer working group which organizes and maintains network and transfer data from perfSONAR and transfer data sources from 2015 to the present
  - A set of ATLAS collaborators working on network analytics

Importance of Measuring Our Networks

- End-to-end network issues are difficult to spot and localize
  - Network problems are multi-domain, complicating the process
  - Standardizing on specific tools and methods allows groups to focus resources more effectively and better self-support
  - Performance issues involving the network are complicated by the number of components involved end-to-end.
- **Network problems can severely impact ATLAS's workflows** and can take weeks, months and even years to get addressed!
- perfSONAR provides a number of standard metrics we can use
- Latency measurements provide one-way delays and packet loss metrics
  - Packet loss is almost always very bad for performance
- Bandwidth tests measure achievable throughput and track TCP retries (using Iperf3)
  - Provides a baseline to watch for changes; identify bottlenecks
- Traceroute/Tracepath track network topology
  - All measurements are only useful when we know the exact path they are taking through the network.
  - Tracepath additionally measures MTU but is frequently blocked

Current perfSONAR Deployment

http://grid-monitoring.cern.ch/perfsonar_report.txt for stats

249 Active perfSONAR instances
199 Running latest version (3.5)

95 sonars in latency mesh

- 8930 links measured at 10Hz
- packet-loss, one-way latency, jitter, ttl, packet-reordering

115 sonars in traceroutes mesh

- 13110 links
- hourly traceroutes, path-mtu

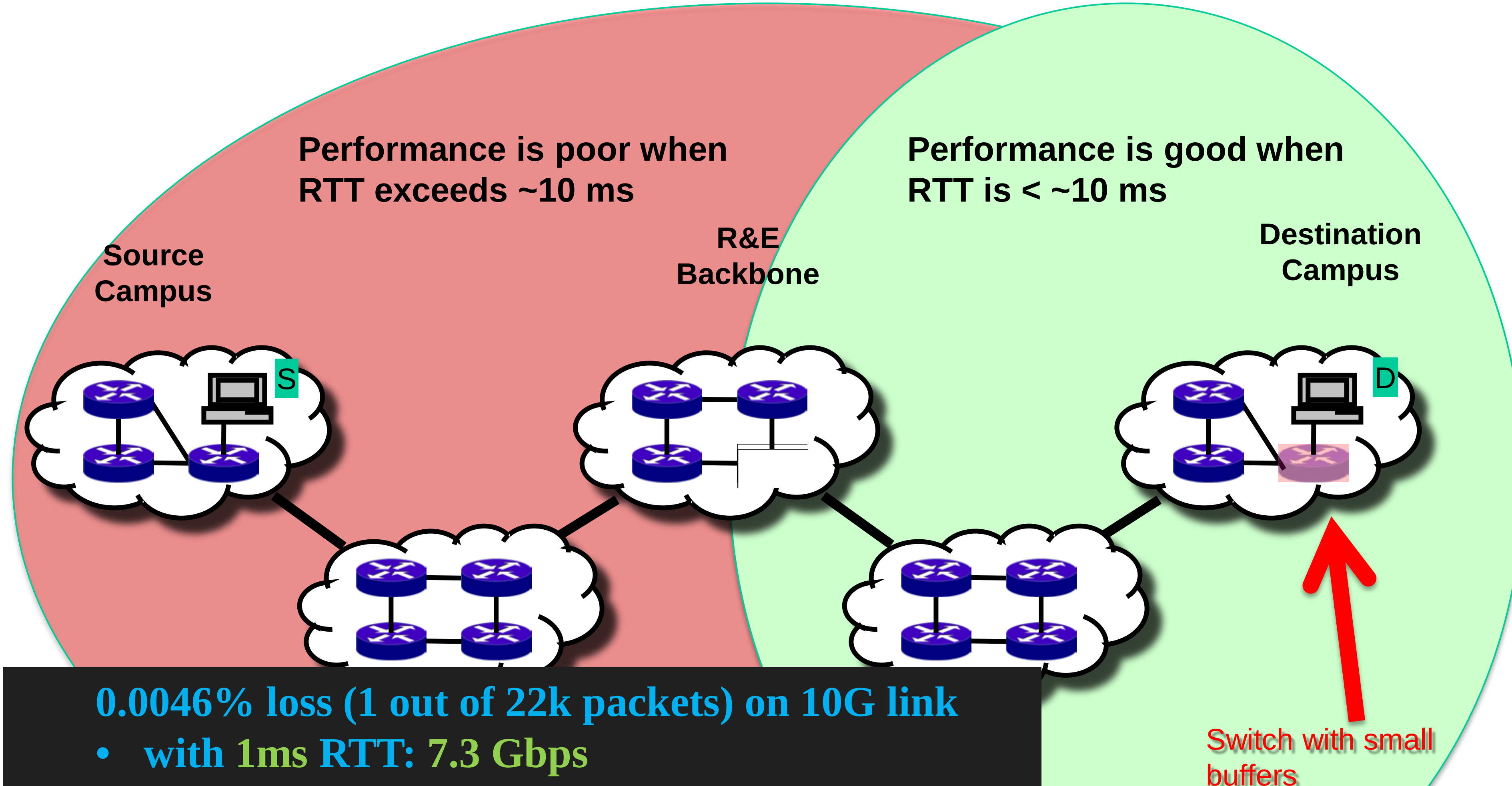
102 sonars in bandwidth mesh

- 10920 links (iperf3)

https://www.google.com/fusiontables/DataSource?docid=1QT4r17HEufkvnqhJu24nIptZ66XauYEIBWWh5Kpa#map:id=3

- Initial deployment coordinated by WLCG perfSONAR TF
- Commissioning of the network followed by WLCG Network and Transfer Metrics WG

Latency and packet loss matters



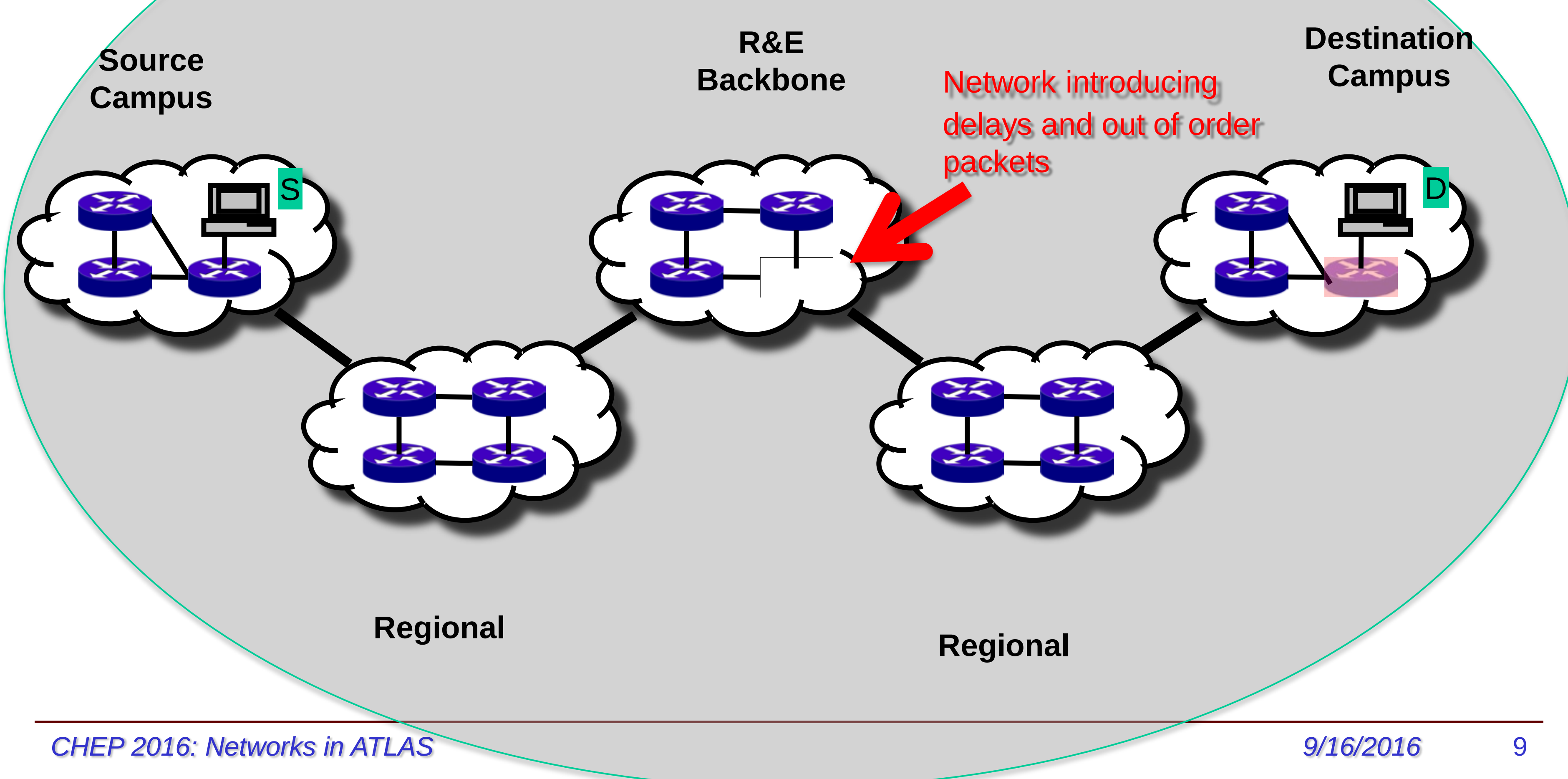
0.0046% loss (1 out of 22k packets) on 10G link

- with 1ms RTT: 7.3 Gbps
- with 51ms RTT: 122Mbps
- with 88ms RTT: 60 Mbps (factor 80)

Packet ordering and jitter

At 70ms RTT on 10G link, 60 seconds test

- with 1% re-ordering, 0.2 ms jitter: 8.45 Gbps
- with 1% re-ordering, 1ms jitter: 1.1 Gbps

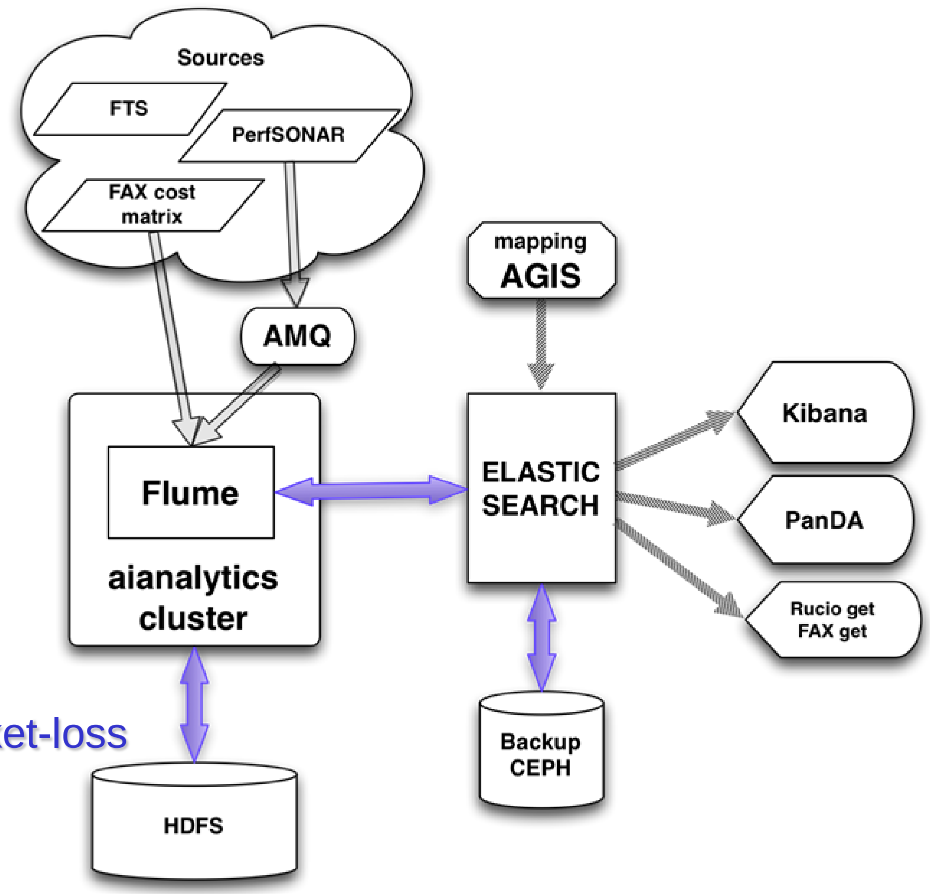


OSG and WLCG Network Efforts

- OSG is in its fifth year of supporting WLCG/OSG networking and is focused on:
  - Developing effective **Alarming and Alerting** for network problems
  - Supporting higher-level network services
  - Improving the ability to manage and use network topology and network metrics: Analytics Platform based upon ELK in use
  - Preparing for and integrating **S**oftware **D**efined **N**etworking
- The WLCG Network and Transfer Metrics working group has created a support unit to coordinate responses to potential network issues
  - Tickets opened in the support group can be triaged to the right destination
  - Many issues are potentially resolvable within the working group
  - Real network issues can be identified and directed to the appropriate network support centers
  - Documented at https://twiki.cern.ch/twiki/bin/view/LCG/NetworkTransferMetrics#Network_Performance_Incidents
  - Many issues resolved within hours of being reported mainly due to our ability to narrow down using perfSONAR

ATLAS Network Analytics

- Ilija Vukotic/U Chicago has led the effort to get network metrics into an analytics platform.
- This analytics service indexes historical network related data while providing predictive capabilities for network throughput.



Primary functions:

- Aggregate, and index, network related data associated with WLCG "links"
- Serve derived network analytics to ATLAS production, DDM & analysis clients
- Provide a generalized network analytics platform for other communities in the OSG
- Initial "Alarm" query prototyped and tested for Source-Destination paths with high packet-loss

- More details at: http://tinyurl.com/gt92zwb

PanDA and Networking

- PanDA is ATLAS's workload manager
  - PanDA automatically chooses job execution site
    - Multi-level decision tree – task brokerage, job brokerage, dispatcher
    - Also predictive workflows – like PD2P (PanDA Dynamic Data Placement)
  - Site selection is based on processing and storage requirements
    - **Why not use network information in this decision?**
    - Can we go even further – network provisioning?
  - Network knowledge useful for all phases of job cycle
- Network as resource
  - Optimal site selection should take network capability into account
    - We do this already – but indirectly using job completion metrics
  - Network as a resource should be managed (i.e. provisioning)
    - We also do this crudely – mostly through timeouts, self throttling
- Longer-term goal for PanDA
  - Direct integration of networking with PanDA workflow – never attempted before for large scale automated WMS systems

Playing with SDN in ATLAS

- Future networks won't just have larger capacity
- A group of people in the US from AGLT2, MWT2, SWT2 and NET2 are planning to explore SDN in ATLAS
  - Working with the LHCONE point-to-point effort as well
- The plan is to deploy Open vSwitch on ATLAS production systems at these sites (http://openvswitch.org/)
  - IP addresses will be move to virtual interfaces
  - No other changes; verify no performance impact
  - Traffic can be shaped accurately with little CPU cost
- The **advantage** is the our data sources/sinks become **visible** and **controllable** by OpenFlow controllers like OpenDaylight
- Follow tests can be initiated to provide experience with controlling networks in the context of ATLAS operations.
- Interest from UVic, KIT and SurfSARA in participating
- Possible partnership with ESnet/CORSA in ~Dec 2016 timeframe
- *For more details talk to Rob Gardner or Shawn McKee*

Network Evolution

- Historically the Wide-Area Network capacity has not always had a stable relationship compared to the data-center or end-node
  - In early days network links (on modems) significantly lagged the local speeds achievable within and between computers
  - The WAN technologies grew rapidly and for a while outpaced LAN and even local computing bus capacities
  - Today 100Gbps WAN links are the typical high-performance network speed but LANS are also in the same range.
    - In Fall 2015 I bought a 32 port 100G switch, 4 dual-ported 100G NICs, 4 dual-ported 50G NICs, 4 dual-ported 25G NICs and all cables for $18K
    - This summer I ordered a 100G NIC (Qlogic) for $350
- Today it is easy to oversubscribe our WAN links (in terms of $ of local hardware at many sites)
- Will our R&E network providers be able to keep up with our needs?
  - So far, not a problem….
  - CERN already tested 200Gbps waves
  - By 2020 800 Gbps waves will be available (assuming you buy the new hardware to support it)

R&E Networking

- High-Energy Physics (HEP) has significantly benefited from our strong relationship with Research and Education (R&E) network providers
  - To-date they have given us “infinite” capacity at relatively low (or no-direct) cost
    - **They have been able to continually expanded their capacity to overprovision their networks relative to our needs and use.**
- At the Terena network conference last spring, SKA (Square Kilometer Array) noted they will operate at data volumes **200xLHC scale** (https://tnc16.geant.org/core/presentation/721)
  - Besides Astronomy there are MANY science domains anticipating data scales beyond LHC: Health, Bioinformatics, Engineering…
- R&E network providers work closely with us in part because they view HEP as representative of future data-intensive science domains
  - HEP serves as the early prototype for such user communities
  - Network providers are **concerned** about what happens when there are **N** more HEP-scale science domains all wanting infinite capacity
    - **Perhaps we should be too!**

Future Directions

- The WLCG efforts at CERN are being reorganized and this is an opportunity to chart future directions for the our networking efforts.
- We have a number of areas (projects; see next slide) we are considering and we need to understand where these efforts should be housed (Stay in WG, move to GDB, to LHCONE)
  - It is important to note there is currently very little manpower for networking (much, much less than computing and storage)
  - To undertake all our plans will require identifying new effort
- We are planning a Pre-GDB meeting on January 11-12th focusing on networking.

Summary

- We have a working infrastructure in place to monitor and measure our networks in use for ATLAS
- perfSONAR provides lots of capabilities to understand and debug our networks
- Work on new applications is underway
  - Notifications/alerting
  - Predictive capabilities
  - Current utilization and capacity planning
  - Evaluating network performance of commercial clouds
- It is in ATLAS's best interest to stay aware of how the network is evolving and what the future landscape may look like
  - Important to start thinking of the network as something we will eventually be able to program/integrate into our architecture(s)

Questions or Comments?

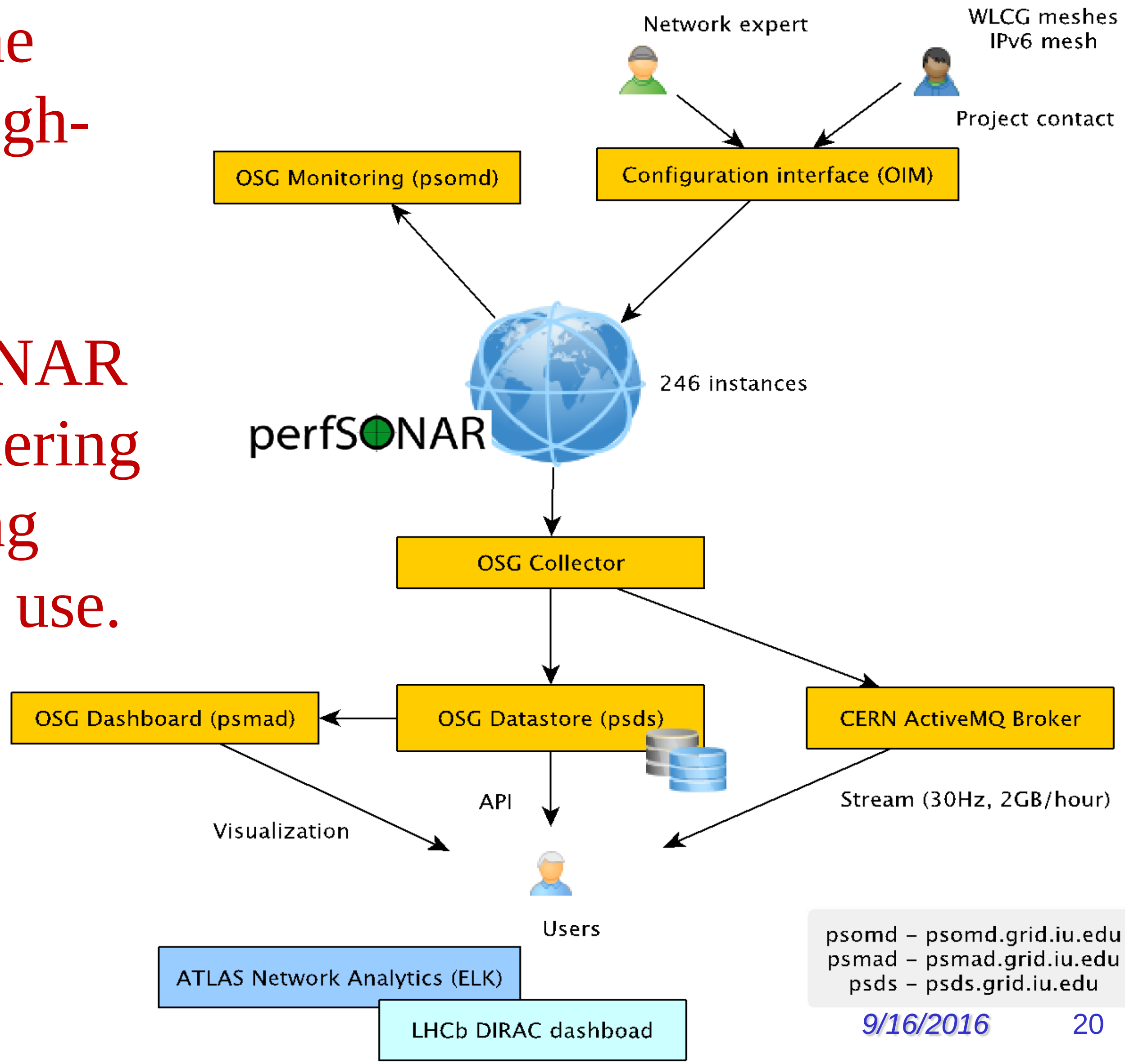
References

- Network Documentation https://www.opensciencegrid.org/bin/view/Documentation/NetworkingInOSG
- Deployment documentation for OSG and WLCG hosted in OSG
  https://twiki.opensciencegrid.org/bin/view/Documentation/DeployperfSONAR
- Measurement Archive (MA) guide http://software.es.net/esmond/perfsonar_client_rest.html
- Modular Dashboard and OMD *Prototypes*
  - http://maddash.aglt2.org/maddash-webui https://maddash.aglt2.org/WLCGperfSONAR/check_mk
- **OSG Production instances for OMD, MaDDash and Datastore**
  - http://psmad.grid.iu.edu/maddash-webui/
  - https://psomd.grid.iu.edu/WLCGperfSONAR/check_mk/
  - http://psds.grid.iu.edu/esmond/perfsonar/archive/?format=json
- Mesh-config in OSG https://oim.grid.iu.edu/oim/meshconfig
  - Being updated to a new standalone mesh-config application (ready for v3.6?)
- Use-cases document for experiments and middleware https://docs.google.com/document/d/1ceiNlTUJCwSuOuvbEHZnZp0XkWkwdkPQTQic0VbH1mc/edit

Back up slides

Overview of perfSONAR Pipeline

The diagram on the right provides a high-level view of how WLCG/OSG is managing perfSONAR deployments, gathering metrics and making them available for use.



Possible Future Project Areas

- **Title:** LHCONE Traffic engineering
- **Areas:** LHCONE, routing, debugging, network orchestration
- **Title:** LHCONE L3VPN Looking Glass
- **Areas:** LHCONE, monitoring, debugging
- **Title:** Integration of network and transfer metrics to optimize experiments workflows
- **Areas:** FAX/Phedex, Rucio, perfSONAR, DIRAC
- **Title:** Advanced notifications/alerting for network incidents
- **Areas:** WAN, Advanced Notifications/Alerting, perfSONAR, Hadoop/Spark
- **Title:** Network performance of the commercial clouds
- **Areas:** Clouds, WAN connectivity, WAN performance (perfSONAR), establishing and testing network equipment at the cloud provider (VPN)
- **Title:** Software Defined Network Production Testbed
- **Areas:** WAN, SDN, LHCONE/LHCOPN, Storage/Data nodes

Throughput predictions

- Throughput measurements are expensive so done at low frequency. Delays and packet loss rate are cheap.
- Idea is to use delays and packet loss rate to predict maximum possible throughput.

- Mathis formula is used to model impact of packet loss and latency on throughput
  - **Rate < (MSS/RTT)*(1 / sqrt(p))**
    - MSS – segment size
    - RTT – round trip time
    - p – packet loss
- Packet (re)ordering and jitter to be added as well

Example: Faster User Analysis

- First use case for network integration with PanDA
- Goal - reduce waiting time for user jobs
  - User analysis jobs normally go to sites with local input data
  - This can occasionally lead to long wait times (jobs are re-brokered if possible, or PD2P data caching will make more copies eventually to reduce congestion)
  - While nearby sites with good network access may be idle
- Brokerage uses concept of ‘nearby’ sites
  - Use cost metric generated with Hammercloud tests
  - Calculate weight based on usual brokerage criteria (availability of CPU resources, data location, release…) plus new network transfer cost
  - Jobs will be sent to the site with best overall weight
- Throttling is used to manage load on network

Cloud Selection

- Second use case for network integration with PanDA
- Optimize choice of T1-T2 pairings (cloud selection)
  - In ATLAS, production tasks are assigned to Tier 1's
  - Tier 2's are attached to a Tier 1 cloud for data processing
  - Any T2 may be attached to multiple T1's
  - Currently, operations team makes this assignment manually
  - Automate this using network information