Integrating Networking into ATLAS

Shawn McKee(speaker), Marian Babik,
Hironori Ito, Mario Lassnig, Ilija Vukotic

for the ATLAS Collaboration

Track 8, Networks and Facilities

CHEP 2018, Sofia, Bulgaria, July 9, 2018

Integrating Networking into ATLAS

This presentation will focus on integrating networking into ATLAS tools, operations and workflow management in the following areas:

- **PanDA Networking**
- **ATLAS Network Analytics**
- **New and developing networking technologies**

We will provide some context to start, cover the three areas above and conclude with planned next steps

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? (We need to verify efficacy!!)**
    - 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

ATLAS Network Analytics

We have developed an analytics platform which hosts the net metrics

Monitoring

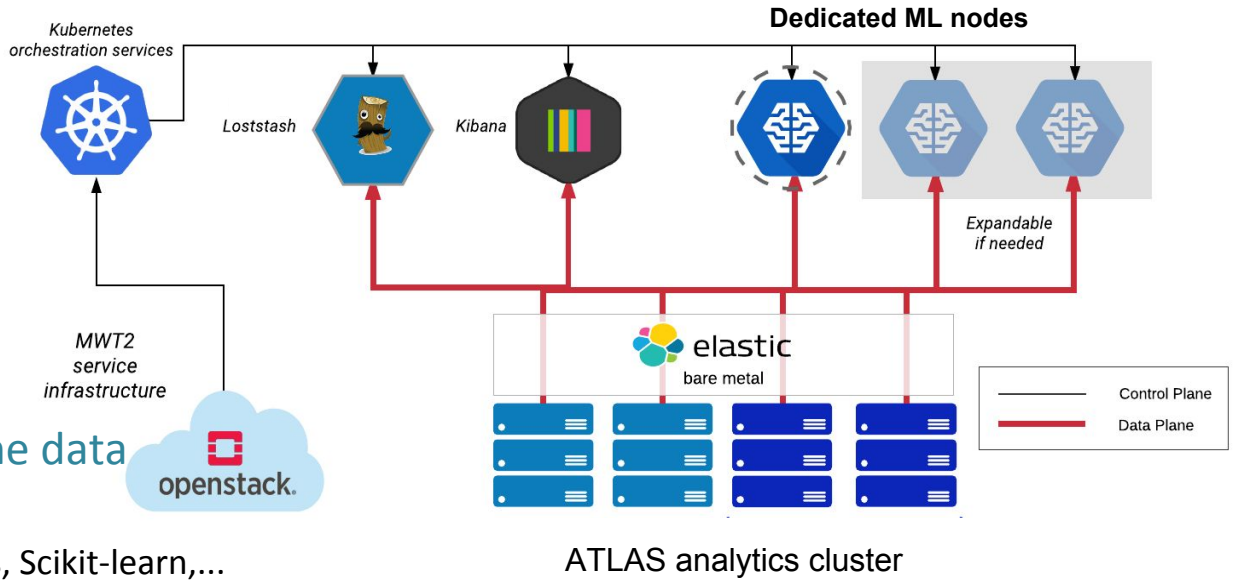
- Long retention
- Fast & reliable
- Has all data

Debugging issues

- Have raw data

Analysis

- Open remote access to the data
- ML platform
  - Python, JupyterHub, Keras, Scikit-learn,...
  - GPU support



ATLAS analytics cluster

Alarm & Alert: currently alerts available as self-subscribe

ATLAS Network Analytics(2)

Four sources of **complementary networking information** are in the platform:

- Perfsonar - link performance
  - One-way delay
  - Packet Loss
  - Path
  - Throughput
- ESnet
  - counters
- FTS
  - rate per file
  - number of active transfers
  - transfer queues
- HammerCloud direct copy tests
  - storage to worker node rates.

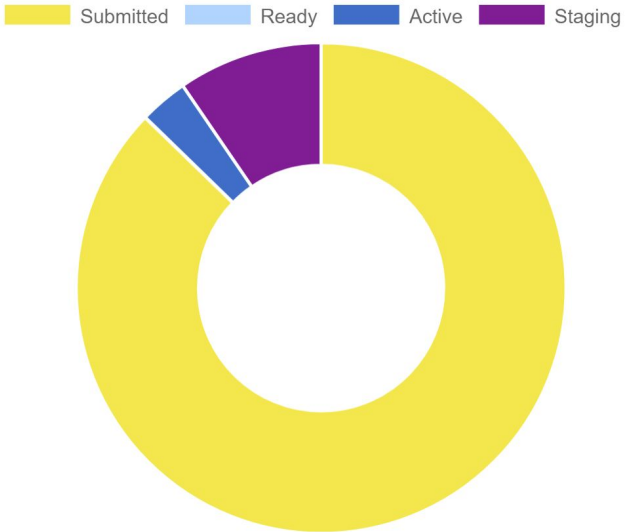
Underway: developing ways to derive and deliver best transfer rate estimates to job scheduling systems. http://costmatrix.slateci.net/

ATLAS Analytics Studies

- CPU benchmarking (relative ranking, what CPUs are best for the job)
- Data usage (what datasets are popular, what data collections are not needed, how to optimize derivations)
- IO studies (performance of different formats, ROOT options, storages)
- Site monitoring and optimization
- **File Transfer System (FTS) optimizer tuning, endpoint/link settings optimization**
- Job wall/CPU time efficiency, job brokering studies
- **Network anomaly detection**
- Local Cache simulations

Using Analytics: ATLAS FTS Delay

ATLAS is seeing issues in the amount of time it is taking to transfer data via ATLAS DDM.
Many sites have large queues of transfers in FTS.

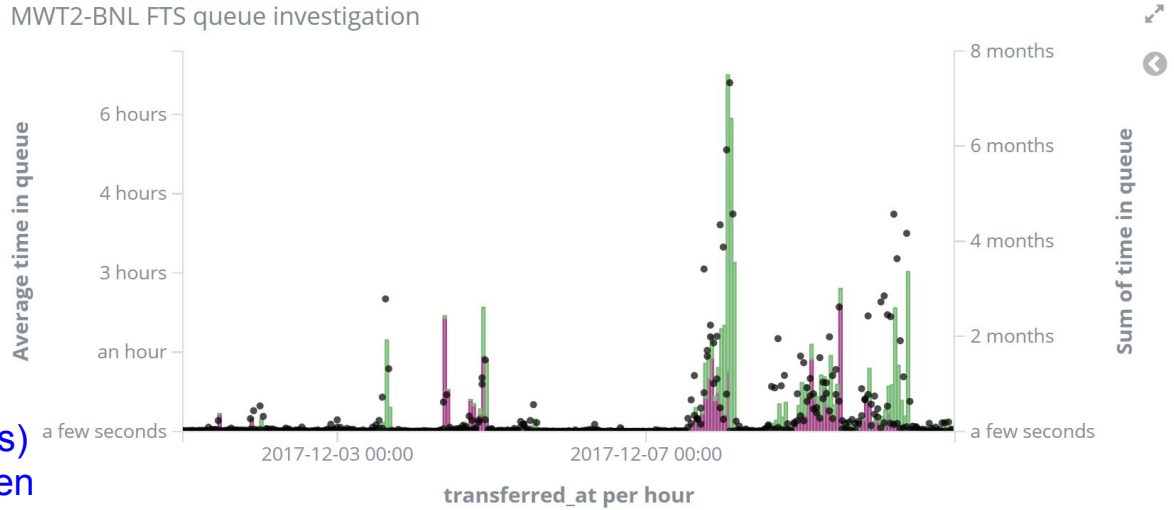


Example pie chart is from FTS
Active was typically ~5% compared to ~80% "queued" (active has a **cap**; queued varies)
This is a 6-hour window but seems almost independent of when we look or which FTS server we check…

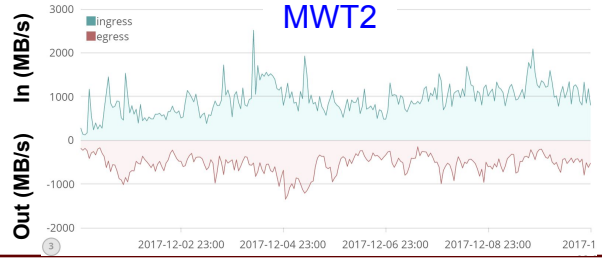
Jobs waiting more than 24 hours can be reassigned
Are we really always saturated with work for FTS?

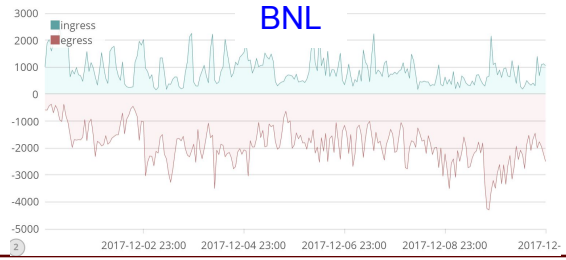
ATLAS FTS Delay

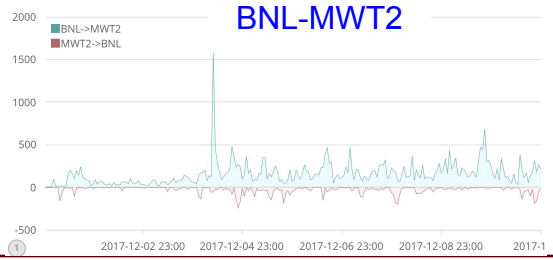
For an ATLAS-only link we see queues while not saturating anything: BNL-MWT2



Plots below
Network Traffic (MB/s)
Each site and between







Using Analytics: ATLAS Alarming

- Simple alarms for now based only on perfSONAR*
  - To make sure we have all the data, we set up alarms on issues with the data transport chain.
  - Alarm on **major** problems at individual sites
    - if packet loss from a site to more than 6 other sites is greater than 2%. While highly specific, this test is not sensitive enough.
    - Coming soon: alarms on large increases in latencies and significant changes in path.

* For now FTS data doesn't tell us much about network performance (limited by the FTS itself).

Making the Most of our Networks

- Much of our ATLAS infrastructure is NOT tuned to take the best advantage of the networks we currently have
  - There are a wide range of misconfigurations, non-optimal tunings and incorrect application, firmware and hardware settings that lead to inefficient use of our networks
  - We have a wealth of data now available and analyzable to help identify bottlenecks and poor performance.
- With this infrastructure we need to take the next step and work to improve ATLAS resources' ability to effectively utilize the network
  - Doesn't require SDN, new hardware or new networks but can have a large impact
  - **Challenge:** how best to organize this work (**what** and **who**)?

New/Developing Network Technologies

- **Future networks won't just have larger capacity but additional interactivity & programmability**
- The **HEPiX Network Function Virtualization working group** is exploring how best to position the LHC experiments to take advantage of new networking capabilities
  - There are ~monthly meetings with tutorials and best-practice discussions
  - Goal is a whitepaper by Spring 2019 to determine a work plan moving forward.
- Within ATLAS, three sites AGLT2, MWT2 and KIT are working to explore SDN technology
  - Each is deploying **Open vSwitch** on ATLAS production systems (http://openvswitch.org/)
  - IP addresses will be move to virtual interfaces
  - Traffic can be shaped accurately with little CPU cost
- The **advantage** of using Open vSwitch is that our data sources/sinks become **visible** and **controllable** by OpenFlow controllers like OpenDaylight
  - **BENEFIT**: Traffic shaping can result in significantly improved use of the WAN
  - Follow-on tests can be initiated to provide experience with controlling networks in the context of ATLAS operations. (Plan: work with HEPiX NFV to determine needed tests)

Future Directions

- Continue to improve analytics for pinpointing the type and location of network issues **automatically**
- Using existing measurements, **find** and **tune-up** problematic sites.
- Complete **at-scale production testing and evaluation** of Open vSwitch **traffic shaping** between ATLAS sites
- Work with **ATLAS Ops** to determine how to better automate network problem finding and solving once they have a hint of a problem:
  - **Example**: ATLAS Ops notices transfer rates are decreasing between two sites. How best to automate debugging to determine if the cause is network vs end-site?
- GEANT's Data Transfer Node (DTN) group is interested in working with ATLAS to test using GEANT's DTN nodes: both for network performance monitoring and to determine how best to optimize transfers
  - This should be integrated with the existing network metrics collection pipeline and related alerting/alarming work.

Summary

We have a number of activities underway to try to integrate the networking knowledge and visibility we are acquiring with ATLAS tools and operations.

We are able to leverage the work of **WLCG**, **OSG** and **HEPiX** to help reach our goals but still need to identify additional manpower with **ATLAS** to enable useful integration and evaluate its real impact.

Questions or Comments?

For Further Information

OSG networking documentation:
http://opensciencegrid.org/networking/

WLCG Network Throughput Support Unit:
https://twiki.cern.ch/twiki/bin/view/LCG/NetworkTransferMetrics#Network_Throughput_Support_Unit

HEPiX NFV Working Group Indico:
https://indico.cern.ch/category/10031/