15th ANNUAL WORKSHOP 2019

HPC NETWORKING IN THE REAL WORLD

Jesse Martinez

Los Alamos National Laboratory

March 19th, 2019

LA-UR-19-22146

ABSTRACT

- Introduction to LANL High Speed Networking
- Production Needs/Challenges
- Administration and Monitoring
- Future

INTRODUCTION INTO LANL HPC HIGH SPEED NETWORKING

LANL HPC SUPERCOMPUTERS

- **ATS-1 (Trinity)**
  - Advanced Technology Systems
  - ~19,000 compute nodes
  - 2 PB RAM
  - 42 PFlops
  - Cray Aries Network
- **CTS-1 Systems**
  - Commodity Technology Systems
  - ~60-1500 compute node clusters
  - ~30-200 TB RAM
  - ~.5-2 PFlops
  - Intel Omni-Path (OPA) / Mellanox EDR
- **Shares Storage Systems**
  - Home/Projects (NFS)
    - >TB (Ranges)
  - Scratch (Lustre)
    - ~5PB-72PB
  - Campaign Storage (MarFS)
    - 30PB
  - Archive (HPSS)
    - >PB (Ranges)

- **LANL HPC has very diverse networks and technologies for its users**
  - Ranges from old legacy hardware, new commodity systems, to new advanced technology systems (ATS-1) such as Trinity
- **Workloads ranging from MPI/Parallel based codes to GPU/on-node computations**
- **User datasets into the TB storage range (and growing quickly to PB!)**
  - Need to be able to move data quickly and reliably to ensure continued production use and on-going research
- **Storage solutions evolving quickly to handle not just data sizes/capacities but parallel workloads and metadata access**

DIVERSE FABRICS

- **Mellanox InfiniBand Networks**
  - Cluster High Speed Network
    - Needed for MPI/UCX low latency communication
  - SAN Network
    - Bandwidth capabilities but most importantly metadata access (dedicated systems for metadata, such as Lustre MDS)
  - Dedicated Subnet Manager systems
    - Used for Monitoring purposes
  - Topologies ranging from full fat trees to Dragonfly topologies
    - SAN made up of “Damselfly Topology” for Lustre Fine Grain Routing
- **Intel Omni-Path Networks**
  - Cluster High Speed Network
    - MPI/PSM2 low latency communication on Intel hardware/fabrics
    - Various fat-tree solutions (full fat tree, 2:1 oversubscription)
  - Dedicated Fabric Mangers
- **Cray Aries Networks**
  - Used in advanced technology systems

PRODUCTION DESIGN

STORAGE AREA NETWORK (SAN) DESIGN

- **All super computers have shared file system resources (except for high end, which have dedicated)**
  - Need to provide capability for moving data between shared storage tiers (on system memory -> burst buffers -> scratch -> archive)
  - Data movers need to be able to communicate between most of these tiers to help in data movement for user's workloads
- **SAN needs to be able to handle bandwidth needs**
  - High Speed Network (HSN) needed for bandwidth and latency (metadata)
  - Fabric needs to be designed in a way to allow for expansion and at the same time, being cost effective
    - Chassis solution would simplify everything, but not cheap
  - Different types of storage systems will live and be accessed on SAN in different ways (account for dynamic usage)

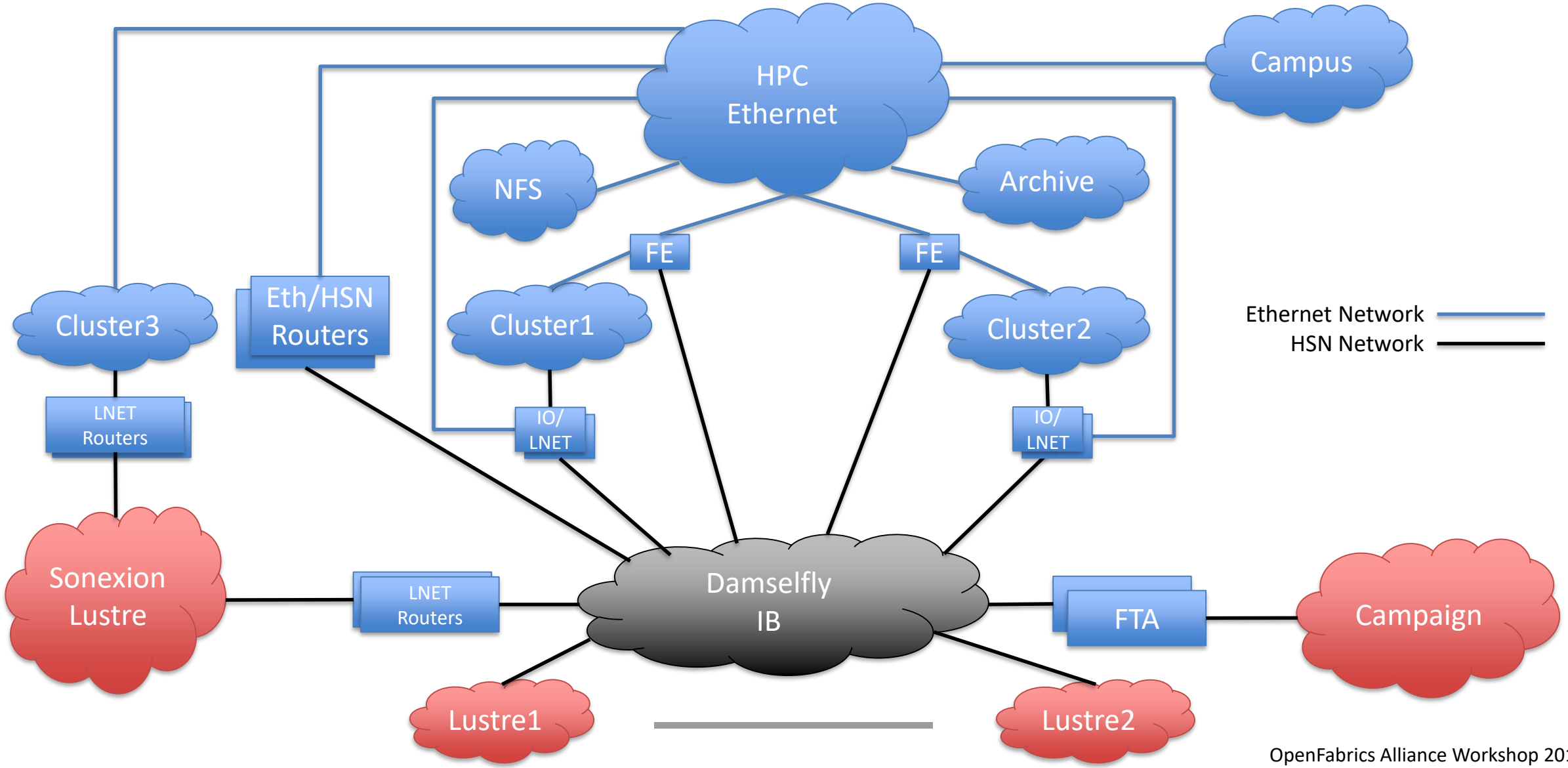
Storage Tiers
Memory
Burst Buffers (Flash)
Scratch (Disk)
Campaign (Disk)
Archive (Disk/Tape)

At the same time, trying to make sure all data is intact and secure!

CHALLENGES

- **SAN technology require different network technologies**
  - NFS/Archive today on Ethernet
  - Scratch/Campaign on IB
  - Cluster Network on OPA/IB/Aries
  - What will be the next system?
- **How can we simplify data movement between storage tiers?**
  - Data Transfer Nodes
    - Need to bridge or route between all the storage tiers/types of networks
    - Simplify enough for users to be able to move data between the tiers easily
      - Requires tools/resources and systems capable of handling data movement
- **Redundancy/Failure options**
  - Lustre Fine Grain Routing (FGR) allows for OSS failure modes and reduce need for full fat-tree fabrics (isolates traffic for data movement)
  - Multiple IO/LNET routers provide multiple paths and access to FGR groups from all systems
  - Ethernet / HSN routers used to bridge the main to types of networks we have today
    - Used to move data between external resources and HPC systems as well as FTA data movement between storage tiers

NETWORK INTEGRATION



ADMINISTRATION AND MONITORING

FABRIC MONITORING AND ADMINISTRATION

- **IBMon**
  - Monitors error and performance of IB fabrics
    - Gathers counters from PerfMgr in OpenSM
  - Detects errors on fabric
- **HSNMon**
  - Monitors error and performance of OPA fabrics
  - Gathers counters from FM / opareport tools
  - Queries fast fabric tools for hardware information
- **Dead Gateway Detection (DGD)**
  - Monitors communication between compute clusters and their IO devices
  - Verifies connectivity between IO devices and SAN networks (IB, Ethernet)
  - Updates routes on computer clusters/SAN when IO non-operational
    - ECMP Routes
    - LNET Routing
    - OSPF Routes

- **Multiple Networks to manage and monitor**
  - Separate subnet fabric manager systems allow for integrating monitoring software
- **Monitoring**
  - IBMon - LANL
  - HSNMon - LANL
  - Dead Gateway Detection (DGD) - LANL
  - Lightweight Distributed Metric Service (LDMS) - Sandia
- **Alerts network administrations and 24/7 tech support staff of hardware issues to resolve**
- **Overall data collections allow for real time usage of fabrics**
- **Performance analysis allows correlation between user job submissions and fabric health counters**
  - Splunk and Grafana used to detail analysis of network performance in conjunction with user jobs
  - Mapping of routes through fabric used to identify bottlenecks/congestion

PREPARING FOR THE FUTURE

NEXT HIGH SPEED TECHNOLOGIES?

- **Storage solutions getting faster and more robust**
  - Need to design SANs to handle existing infrastructure but at the same time make room for other technologies needed
  - Ultra-Scale Research Center identifying ways to make use of different types of network technologies and storage solutions to get the performance and scales that we need, leveraging existing technologies today and expanding on use
- **Integrating Network Technologies**
  - LANL efforts to have network infrastructure support computing systems for future design
  - EMC3 (Efficient Mission-Centric Computing Consortium) working with vendors to look at integrating network components into helping the growth of computing technologies
- **New systems looking at existing capabilities**
  - Existing clusters meant for Hadoop/Container research only have Ethernet internal (look at soft RoCE?)
  - Storage leverage RDMA capabilities more efficiently
- **Can we merge many of these systems together?**
  - Separate routers and FTAs have to bridge many networks
  - How can we easily design a network to encompass all of these needs?

CONCLUSION

- **LANL HPC is ever growing and expanding fast!**
- **New technologies provide new opportunities/challenges**
  - Difficult to stay ahead of the curve of new systems, but designing networks efficiently allows to stay on top of the need for bandwidth and reliability
- **Efficient modeling techniques allow us to map usage of the networks and plan for future designs of user needs**
  - LANL making strides into providing more monitoring capabilities for administrators

It's never the network! Except for when it is…

15th ANNUAL WORKSHOP 2019

THANK YOU

Jesse Martinez

Los Alamos National Laboratory