Niagara Update

Scott Northrup

SciNet
University of Toronto

January 25, 2018

SciNet compute canada | calcul canada

Niagara

Design Criteria

- "LP" last of Stage 1 Systems (GP1,GP2,GP3)
- Designed for Large (MPI) Jobs
- Competitive RFP Process
  - LPBM (HPCG, WRF, NAMD, NEK5000, MiniDFT(QE), SPECMPI2007)
  - Energy Efficiency (Flops/Watt)
  - Network Design
  - Storage Design
  - Deployment Plan
- 11 original bidders, 5 shortlisted
- Lenovo bid chosen

Niagara

System Specifications

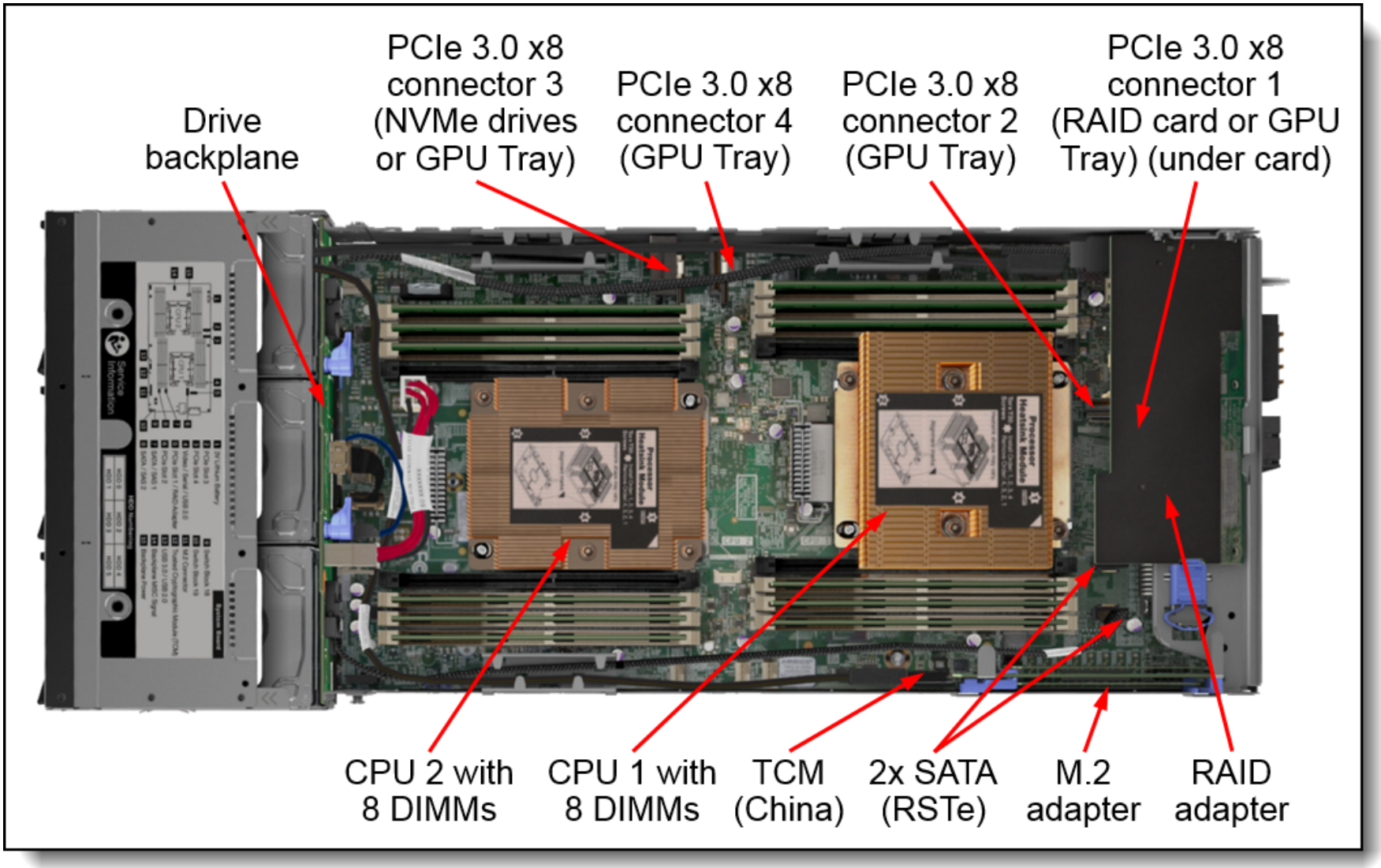
- 1500 nodes (2x20 core Intel Skylake @ 2.4GHz)
- 21 compute, 3 IB, 4 storage, 2 management racks
- 60,000 cores total
- 192GB RAM per node
- EDR Infiniband (Dragonfly+)
- 5PB Scratch, 5+2 PB Project (GPFS)
- 256 TB Burst Buffer (Excelero/GPFS)
- Rpeak of 4.61PF (GPC: 312TF, BGQ: 839TF, Graham: 2.6PF, Cedar: 3.7PF)
- Rmax of 3.0 PF
- 685 kW

Niagara - Compute Nodes

Lenovo SD530 Node

- Intel Skylake 6148 Gold (2.4 GHz, AVX512)
- 192GB Ram (150 GB/s Memory Bandwidth)
- 3TFlops/node
- 100 Gb/s EDR IB
- stateless (diskless)

Niagara - Compute Nodes



Niagara - Compute Nodes

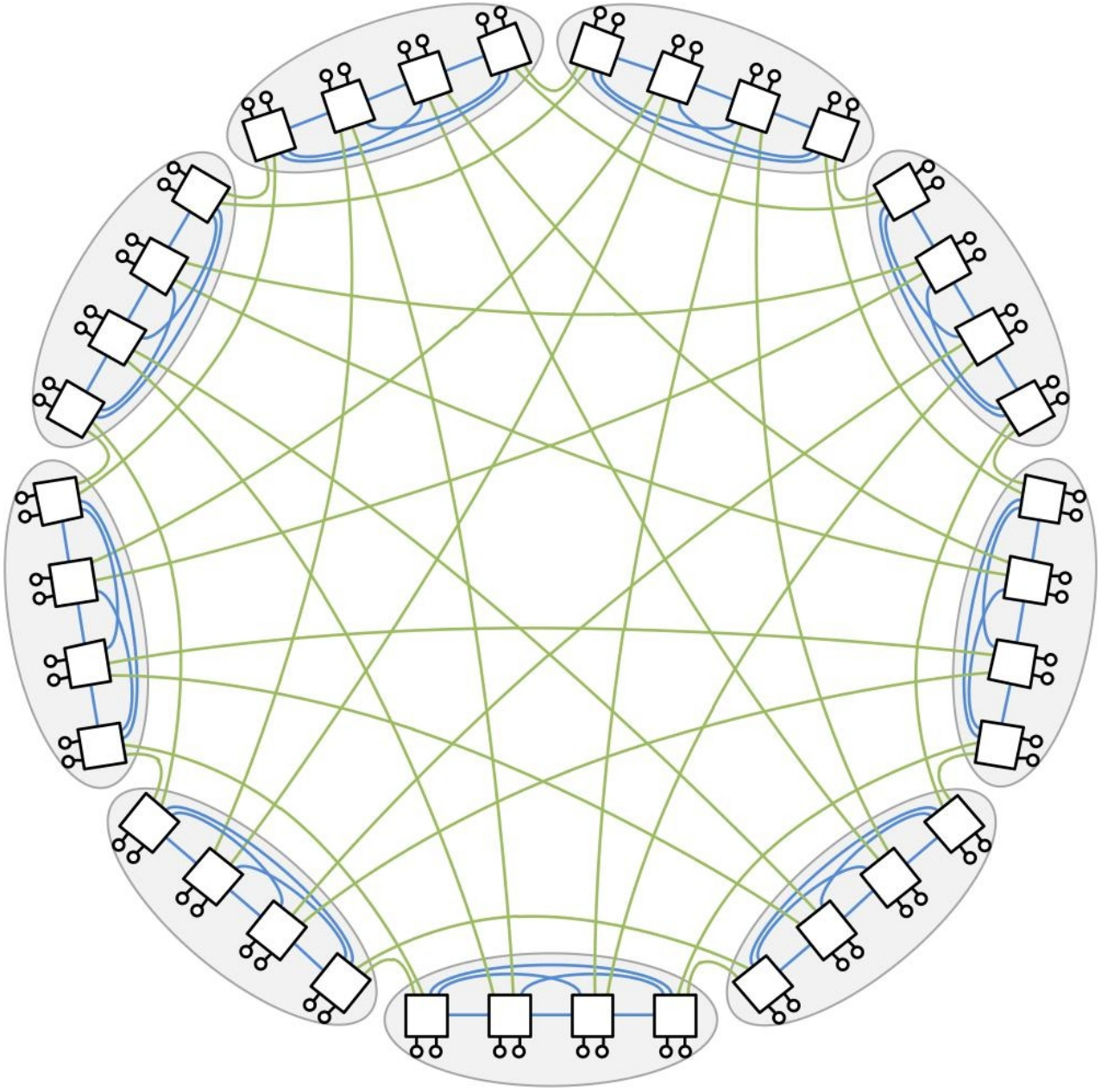
SciNet compute canada | calcul canada

Niagara - High Speed Network

Dragonfly+ Network Topology

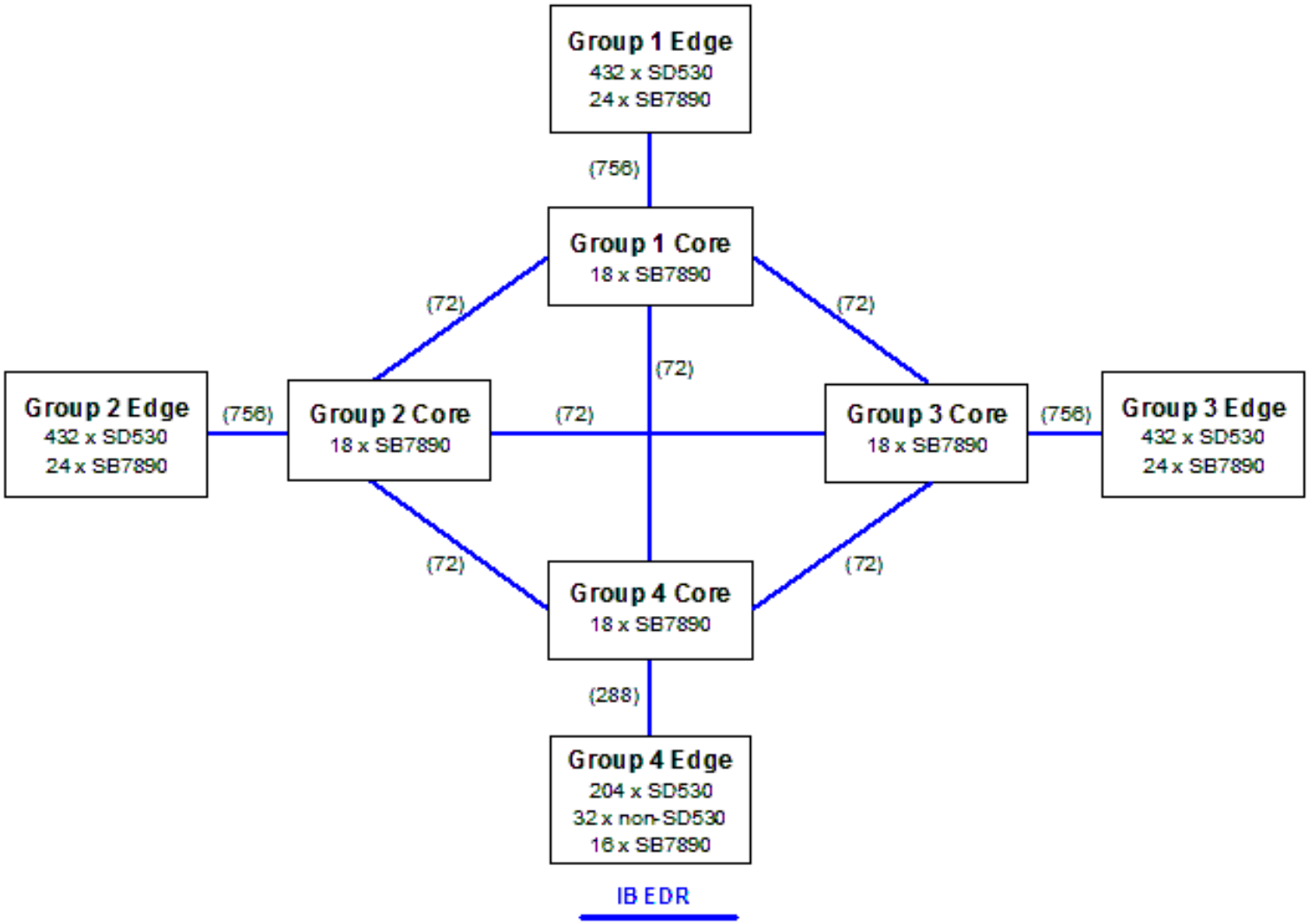
- Introduced by Kim et al (2008) aiming to decrease cost/diameter of network.
- Uses groups of high radix virtual routes to create a completely connected topology.
- Less expensive and more scalable than 1:1 Fat-tree with close to the same performance.
- Topology of Cray XC40/50 “Aries” network
- Requires only edge switches, no core-switches
- Adaptive Routing
- Congestion Control
- new for Infiniband (requires ConnectX-5, Switch IB2)

Niagara - Dragonfly+ Toplogy



SciNet compute canada | calcul canada

Niagara - Dragonfly+ Toplogy



Niagara - Storage

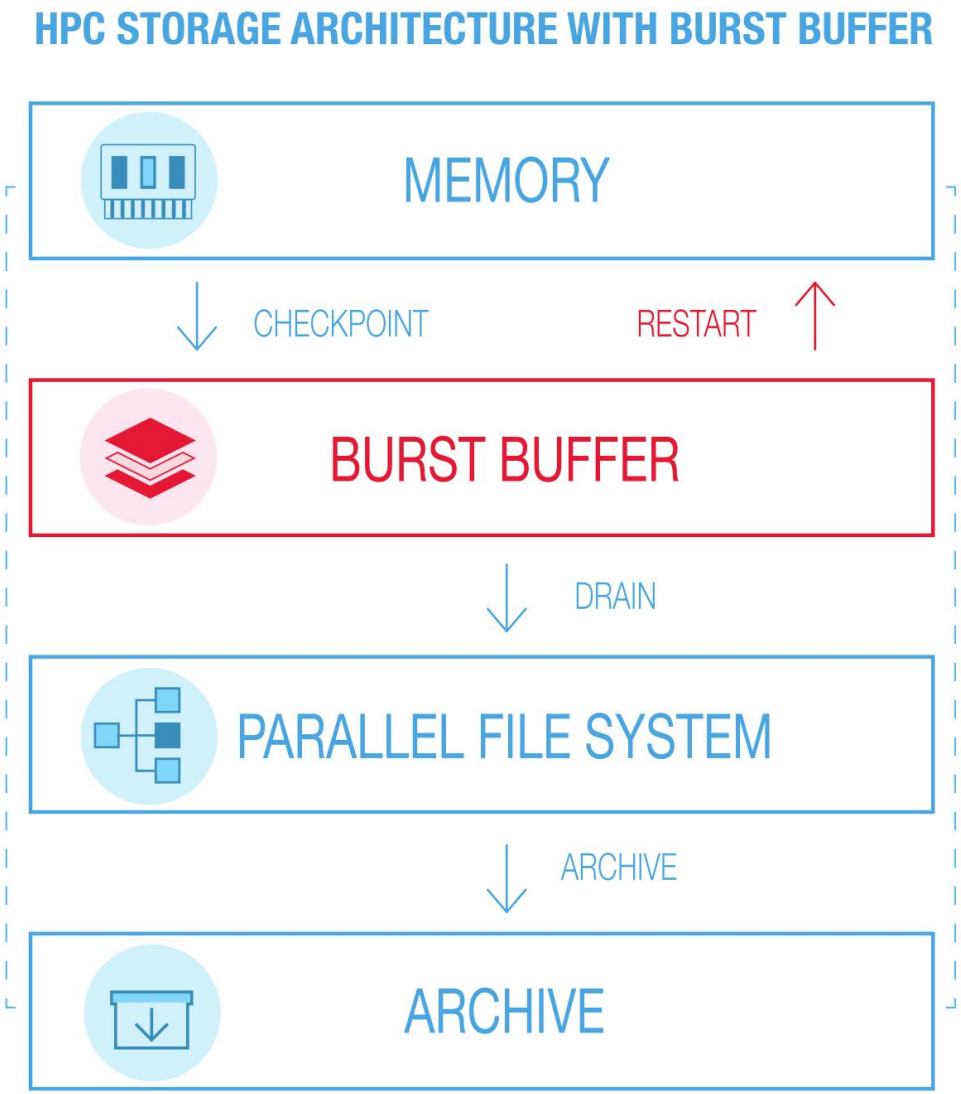
Storage

- 3x Lenovo DSS G260 (504x10TB)
- 10 PB Regular Disk
- 70-90 GB/s R/W
- Spectrum Scale (GPFS)

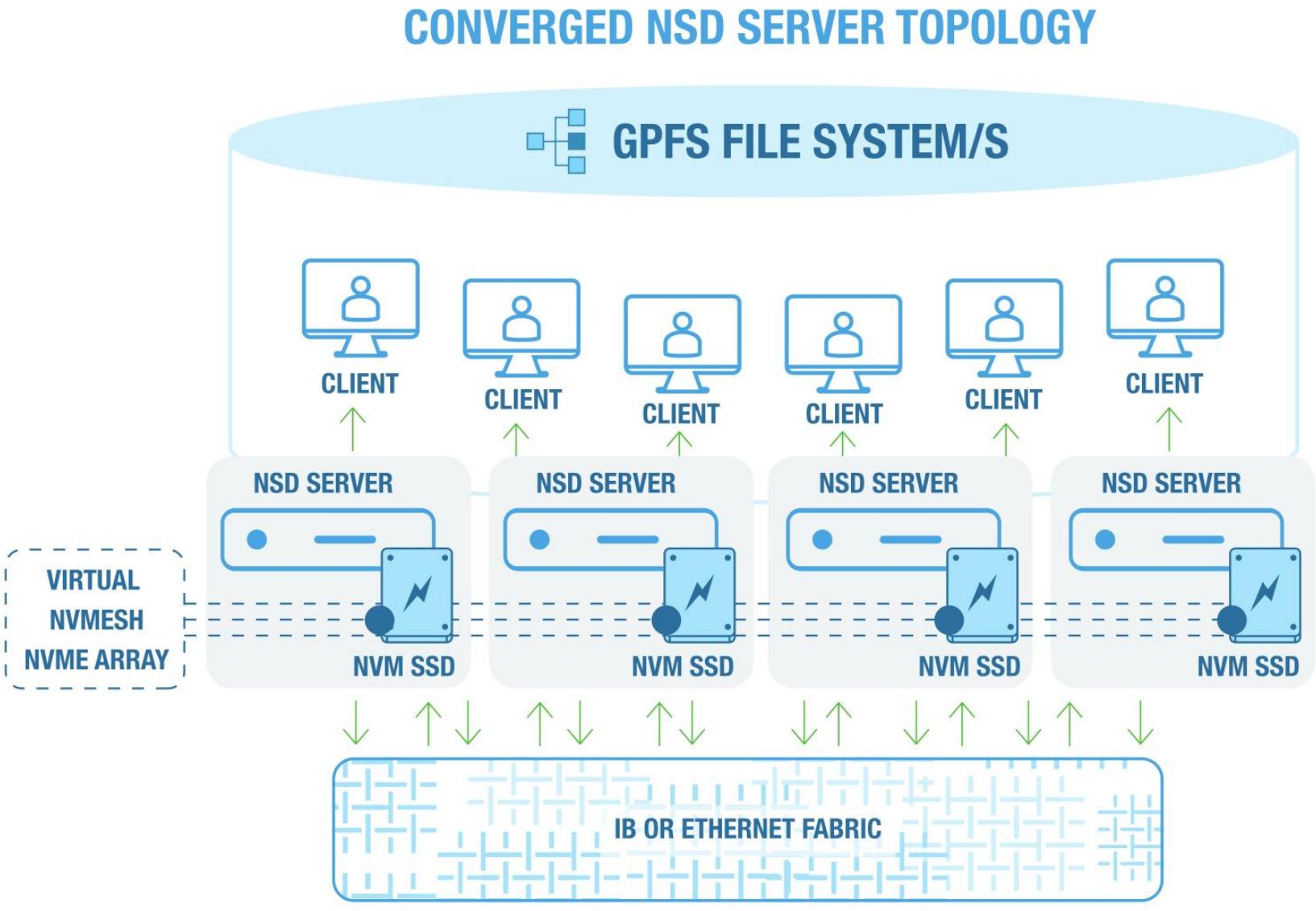
Burst Buffer

- 256TB burst buffer in Raid 1
- 10 nodes with (8x6.4TB NVMe SSD)
- Excelero NVMe Fabric
- 160 GB/s R/W
- very high IOPs performance
- Spectrum Scale (GPFS)

Niagara - Burst Buffer - Excelero NVMesh



Niagara - Burst Buffer - Excelero NVMesh



SciNet compute | calcul canada

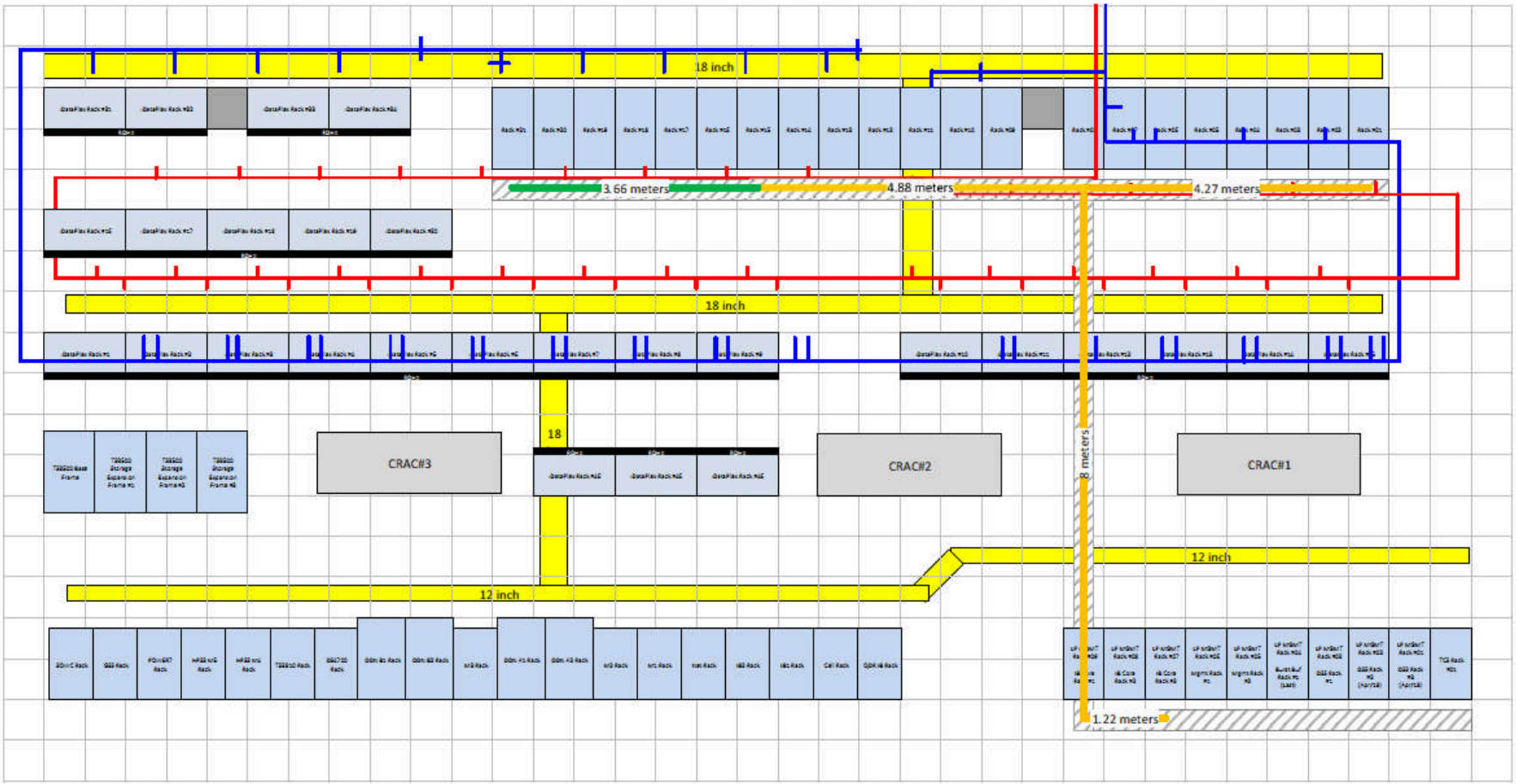
Niagara

Software Configuration

- Centos 7
- CC LDAP
- CC software stack available
- Slurm Scheduler
- target is large jobs (1024+ core)

SciNet compute canada | calcul canada

Niagara



SciNet compute | calcul canada

Ni

compute canada | calcul canada

Niagar

SciNet
compute | calcul
canada | canada

Niagara

compute canada | calcul canada

Niagara

compute canada | calcul canada

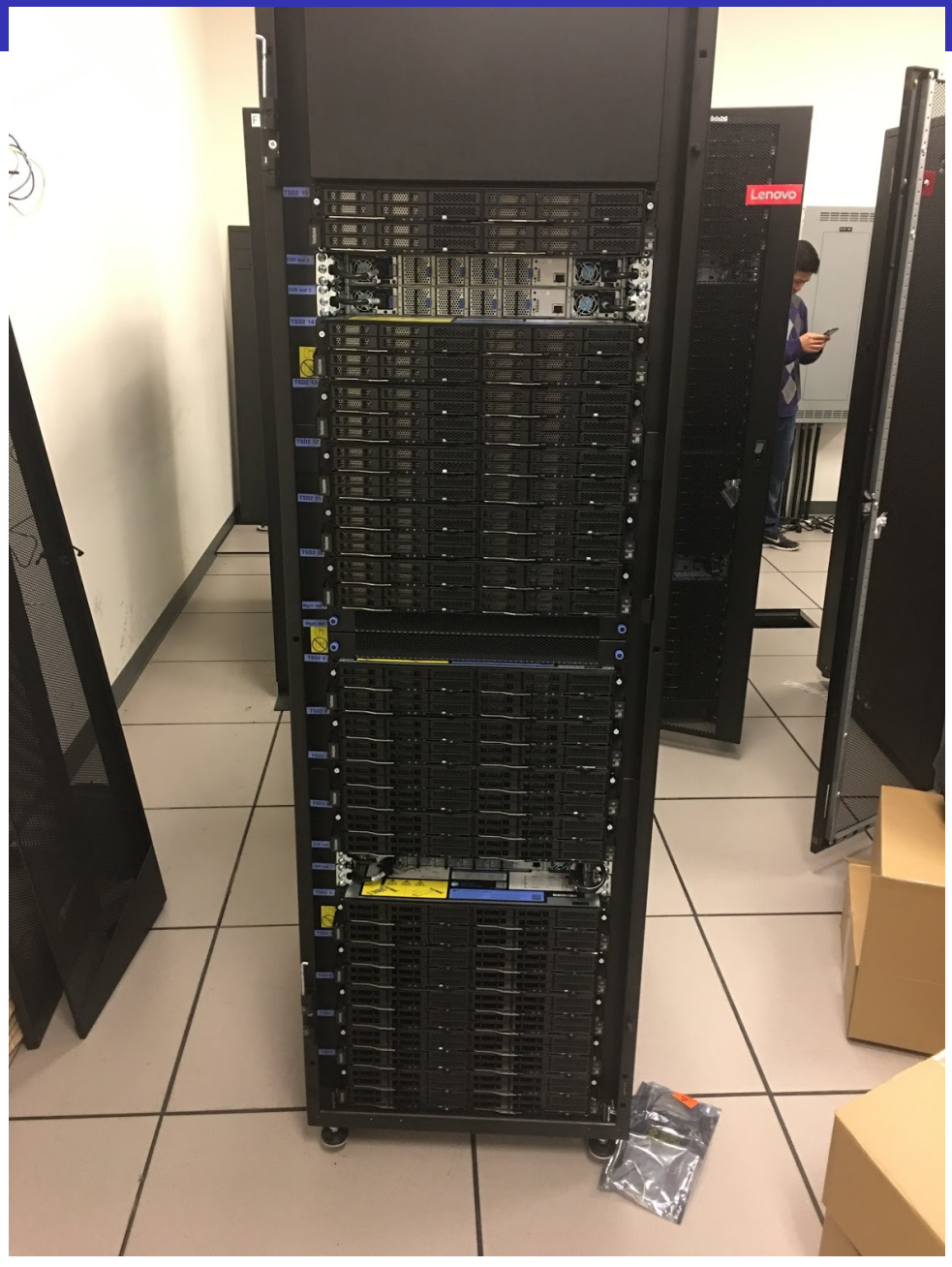
Niagara

SciNet compute | calcul canada

Niagara

SciNet
compute | calcul
canada | canada

Niagara



SciNet
compute | calcul
canada | canada

Niagara

SciNet
compute | calcul
canada | canada

Niagara

compute canada | calcul canada

Niagara

Niagara Install Time-Lapse

Thanks to Marcelo Ponce & Joseph Chen

SciNet compute | calcul canada

Niagara

Migration Details

- HPSS Archive (Tape) stays as is
- Existing GPC $HOME, $SCRATCH, $PROJECT will migrate
- 2018 RAC allocations are for Niagara

Niagara

Niagara Timeline

- RFP Process (Jan - Sep 2017)
- Negotiation & Contracts (Aug - Oct 2017)
- TCS Decommission Oct 2017
- 1/2 GPC Decommission Nov 2017
- Deployment (Dec 2017 - Feb 2018)
- Test/Config (March 2018)
- Final Storage (March 2018)
- Production (April 2018)