US ATLAS Tier1 Report

Chris Hollowell <hollowec@bnl.gov>

US ATLAS Computing Facility Meeting 3/16

BROOKHAVEN NATIONAL LABORATORY | Scientific Data and Computing Center

U.S. DEPARTMENT OF ENERGY

New Datacenter (CFR) Update

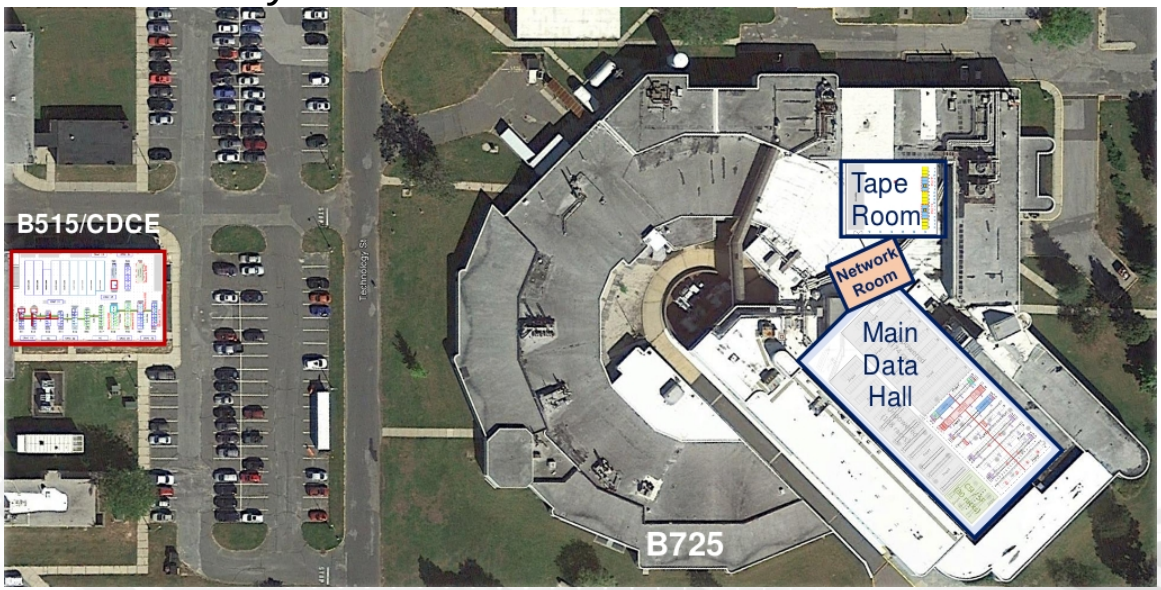
Construction on our new datacenter, in support of HL-LHC computing and storage requirements, commenced in May 2019

Anticipate completion by end of the fiscal year – 10/20

New tape library installed by 12/20

Most existing equipment won't be moved to the new datacenter: will remain in building 515/BCF datacenter until it's retired

67% PUE efficiency increase in new datacenter: reduced TCO

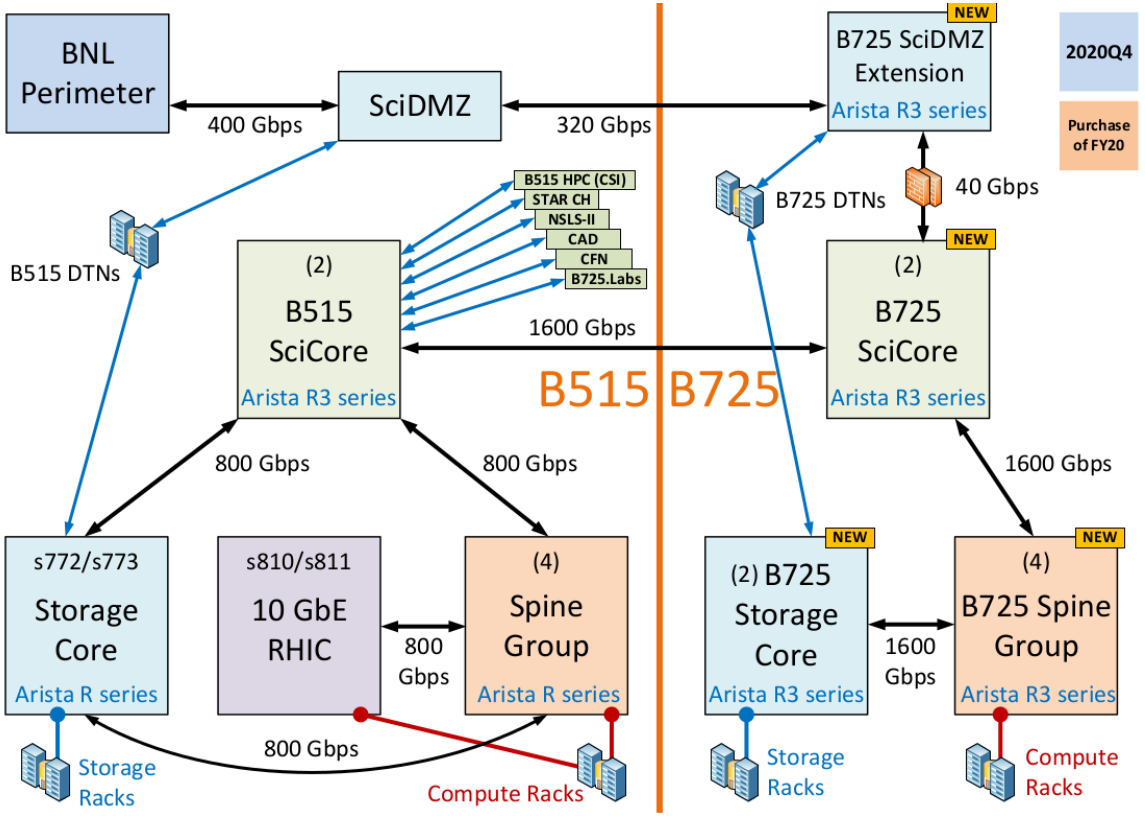


New SDCC Datacenter in Building 725

New Datacenter (CFR) Update (Cont.)

Fiber runs providing 1.6 Tbps now in place between old (515) and new (725) datacenters

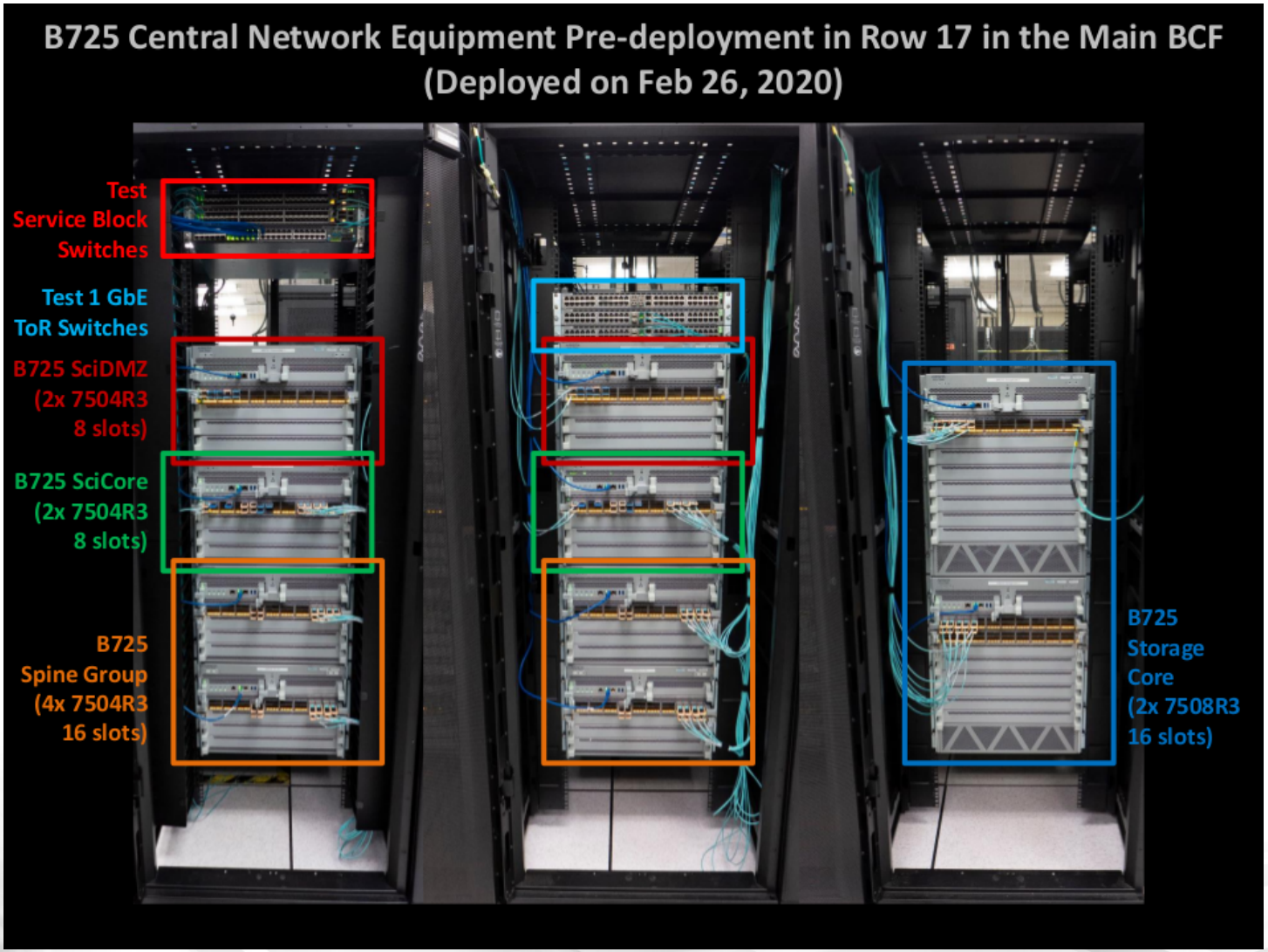
Arista core switchgear for B725 will be moved in 10/20
Already arrived onsite, and being tested



FY21 Planned Multi-Datacenter Network Architecture

New Datacenter (CFR) Update (Cont.)

All core switchgear for new datacenter purchased with BNL (not ATLAS) funds



ATLAS HTC Processing

New AMD EPYC Naples-based compute nodes brought online in 9/19

- 97 AS-1023US-TR4 Supermicro Nodes
  - 2x AMD EPYC 7351 CPUs
    - 64 log. cores total
  - 128 GB (16x8GB 2666 MHz DIMMs)
  - 4x4TB SATA 6 Gbps Drives
  - 4x1 Gbps NICs (only using one)
  - 1U form factor
  - ~775 HS06 per node = ~75 kHS06 total

First time ordering on a large scale from Supermicro – very positive experience

Systems working well – not a single hardware failure since they were brought online ~6 months ago

Rack of new EPYC-based Supermicro 1023US-TR4 Servers

ATLAS HTC Processing (Cont.)

FY2020 compute purchase bid out to vendors

Given the excellent pricing we obtained last year, we are again allowing vendors to bid either Intel or AMD configurations

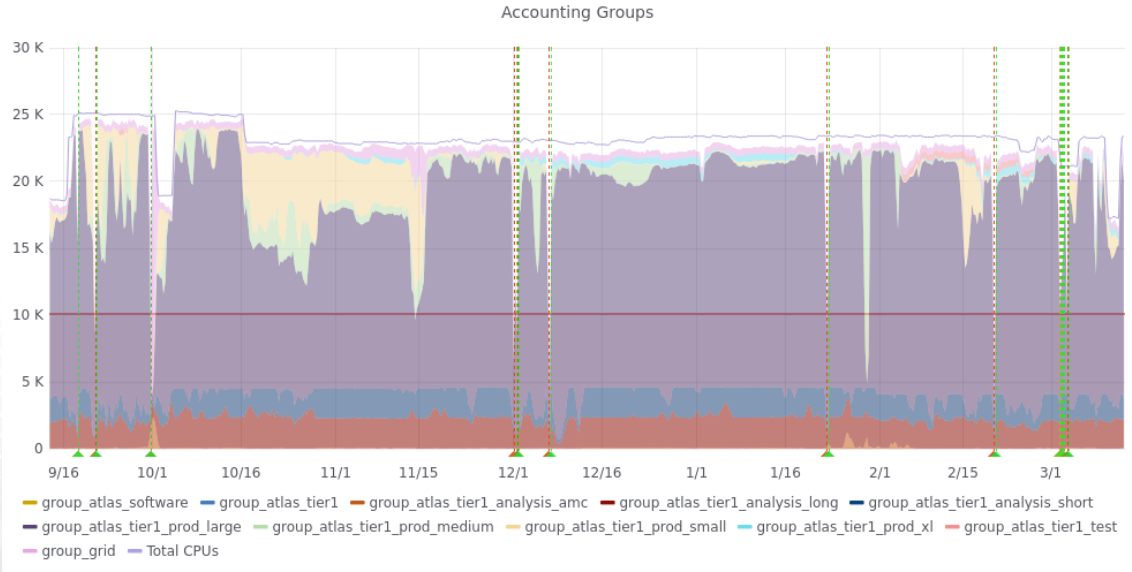
Configurations based on HS06 testing of eval hardware with various Rome and Cascade Lake CPUs

Allowing vendors to bid a single socket AMD (7452) configuration in the current RFP – single socket core, memory channel and PCI-e lane density in AMD Rome is impressive!

CVMFS client update to 2.7.0 underwav

Currently providing ~23.5k log. CPUs for the ATLAS T1

Recently enabled memory-based partition of slots to more efficiently pack jobs



ATLAS ML Jobs On The BNL Institutional Cluster

A number of ATLAS users interactively submitting jobs to our HPC cluster (both via shell and Jupyter)

15% increase in total utilization over the past month

ANALY_BNL_GPU_ARC PQ

6 IC nodes, 2 P100 GPUs on each host

Singularity (user namespaces enabled) on WNs

CVMFS installed

Proxy setup for outgoing HTTP connectivity

Central CERN Harvesters submit pilots to ARC-CE

Successful completion of ATLAS ML jobs – queue enabled and available

4654054766 Attempt 1 of 1	Doug Benjamin	3484 20665405	runcontainer	finished	2020-02-24 23:25:11	0:0:02:24	0:0:06:21	2020-02-24 23:40:41	US ANALY_BNL_GPU_ARC brokeroff testing	1000	0 (1)
	Job name: user.bdouglas.hp.test.20200224172414/.4654054766 #1										
	Datasets: **In:** user.aforti:user.aforti.btag.dl1.hp.optimisation.configs_hpscan || Rucio link **Out:** user.bdouglas.hp.test.20200224172414_out.json.297980496										
4653979810 Attempt 1 of 1	Xin Zhao	189 20664836	runcontainer	finished	2020-02-24 20:14:21	0:0:02:39	0:0:06:21	2020-02-24 20:28:22	US ANALY_BNL_GPU_ARC brokeroff testing	1000	0 (1)
	Job name: user.xzhao.hp.test.20200224211308/.4653979810 #1										
	Datasets: **In:** user.aforti:user.aforti.btag.dl1.hp.optimisation.configs_hpscan || Rucio link **Out:** user.xzhao.hp.test.20200224211308_out.json.297960536										

ATLAS Storage

dCache

21.95 PB data

Hardware RAID
- Hitachi HUS 130
- Hitachi G400

Software RAID
- Western Digital Ultrastar 102 JBODs

Running v4.2

Upgrade to v5.2 scheduled for 3/25 – 3/27 postponed

Postgres11 → Postgres12

Java 1.8 → Java 11

BNL T1 @ Run2 RPVLL Reprocessing

Ongoing reprocessing runs in Data Caraousel mode

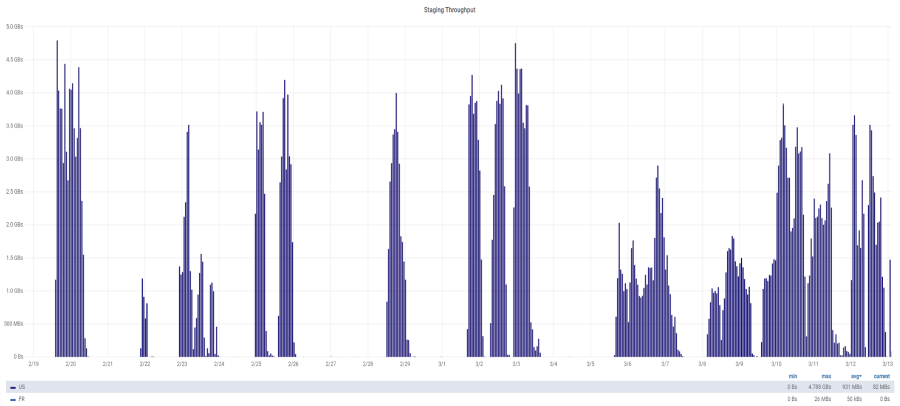
Data15~Data18 RAW data

The data volume for BNL: 2.9 PB, 110 datasets, 1.3 M files

Excellent perfromance @ BNL T1 – best of the T1s

Reached 3~4.5 GB/s staging throughput

We will further stress HPSS
in a dedicated exercise after the
reprocessing



Questions?

Thanks to the following people for contributing to the information presented:

Doug Benjamin, Eric Lancon, Jane Liu, Robert Hancock, Hiro Ito, Alex Zaytsev, Xin Zhao