



Data Center Operations and Networking

Imran Latif, Alex Zaytsev

ATLAS Pre-Scrubbing Review - June 27, 2022



@BrookhavenLab

Data Center Operations

What are the duties of the group

SDCC Data Center Ops team (I. Latif (0.15), J. McCarthy (0.25), E. Garcia (not in ATLAS WBS), P. Metz (0.30 starting from FY23) to ensure the following:

- The overall delivery, installation and decommissioning and commissioning of IT equipment in both 515 and 725 data centers at BNL
- The team is ultimately responsible for overall facility operations of the data centers ensuring safety, and minimized downtime to the scientific data and computing equipment
- Liaison with BNL Facilities and Operation (F&O) and supervisor sub-contractors in charge of maintenance
- Pro-active deployment of Rear Door Heat Exchangers (RDHx) units, cabinet power distribution units (CDU) and associated monitoring & controls
- Maintain the configuration of data center assets, layout and inventory using the Data Center Infrastructure Management (DCIM) system
- Report IT electricity kWh usage to the utilities for billing, report the PUE value to DOE (1.4 peak, 1.25 annual average expected by CFR design)
- Provide technical expertise and oversight on preventative and break-fix maintenance of mission critical power & cooling infrastructure at SDCC data centers in buildings 515 and 725

Achievements for ATLAS over last year

- Successful commissioning of the new state of the art B725 data center for production in 2021Q3-4
 - ATLAS beneficial occupancy started in Jul 2021
 - Critical production systems were on the floor by Nov 2021
- Deployment of new IBM TS4500 tape library pair for ATLAS in B725 data center
 - 1 HPSS mover rack is deployed in the Main Data Hall (MDH) for this library pair
- Migration of 11 CPU racks (2017-2020 purchase) from the old data center B515/CDCE to the Main Data Hall (MDH) in B725
 - Done in two interventions performed in 2021Q4 and 2022Q1
 - Only a temporary reduction of available CPU (**no service disruption**)
 - Both performed as pre-scheduled announced interventions
- Deployment of late-FY21 & FY22 purchased equipment in B725
 - 1 CPU rack + 4 JBOD DISK racks from FY21 purchase
 - 4 CPU racks + 2 JBOD DISK racks are to be added from FY22 purchases
- *Any deployment of new equipment for ATLAS will now be in the new data center B725 only (already applied to FY21 and FY22 purchases)*

B725 Data Center: Main Data Hall (MDH)



1st fully populated CPU row (rack frames are delivered pre-configured and retired with equipment in them)

Storage / infrastructure rows (all rack frames are pre-deployed)

Challenges related to future scaling of power and cooling infrastructure for B725 data center

- BNL Core Facility Revitalization (CFR) project has delivered the B725 datacenter with 3 power systems upfront (3.6 MW of IT power protected by UPS)
- Two 1.7 MW Diesel Generators deployed for B725 data center as of 2022Q1, covering up to 1.2 MW of IT power in B725 with N+1 redundancy in the generator group (or 2.4 MW without redundancy) - **all covered by the CFR project**
 - By FY24, two additional 1.7 MW Diesel Generators would be needed to be installed to ensure redundancy and reliability of the B725 data center (one in FY23 and another one - in FY24) to cover up to 3.6 MW of IT power in B725 at the same level of redundancy - **to be covered by non-program funds; request for funding has been made under BNL GPP / Lab strategic funds**
 - The need for adding the 4th power (and cooling) system covering additional 1.2 MW of IT power for B725 datacenter is still anticipated in FY25-26 period - **to be covered by non-program funds; request for funding has been made under BNL GPP / Lab strategic funds**
- The addition of both two more Diesel Generators and the 4th power system to B725 is the matter going beyond ATLAS Tier-1 in scope and thus no additional cost is to be inflicted on ATLAS while addressing these matters

Networking & B515/B725 Transition

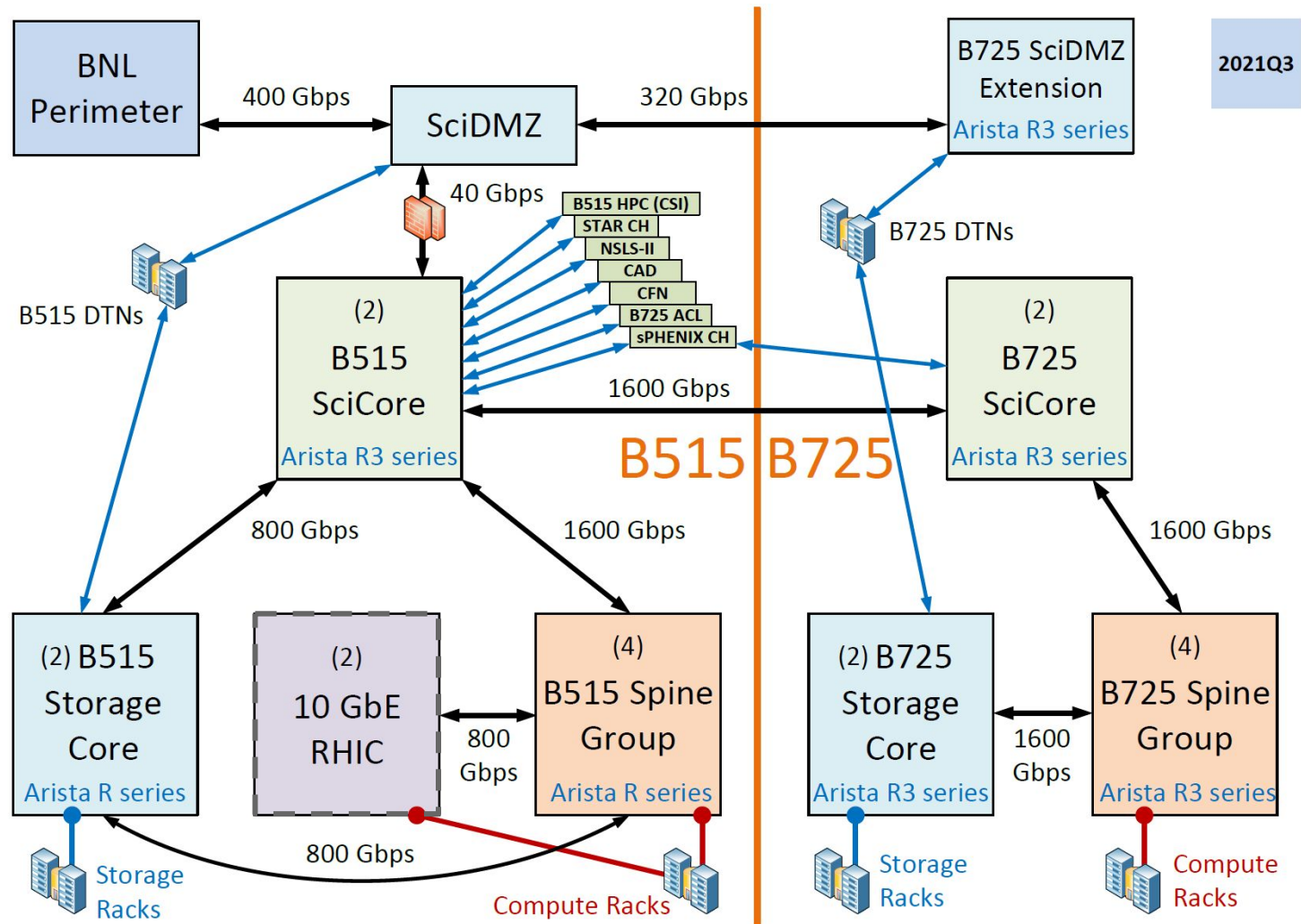
Transition to B515 and B725 datacenter inter-operations period

Network operations are handled primarily by 4 people: M. Lukasczyk (0.20), F. Burstein (0.30), N. Fontana (0.10), A. Zaytsev (0.35), as well as ESnet engineer on site.

- The central network systems of B725 data center were activated in 2021Q3; the activation and testing of the 1.6 Tbps B515/B725 inter-building link were performed and it entered production at the same time
- Both data centers are now operating as a single facility network-wise since 2021Q3 and the transition to this mode was performed completely transparently for ATLAS operations, since all preparations for it were made earlier during Dec 2020 Facility-wide network intervention
 - SDCC SciZone perimeter firewall and the central switched for the SciDMZ still resides on B515 side, as well as ESnet perimeter as of Jun 2022 - **the migration of these components is to B725 is expected in FY23 - early FY24**
 - The ATLAS Tier-1 CPU, DISK and part of services deployment currently reside on both sides of the SciZone, and expected to gradually migrate from B515 to B725 by means of hardware refresh only by FY25 (**no physical move of equipment needed**)
- **The central network infrastructure of SDCC remain unified and consolidated across all programs served by the Facility for maximum efficiency and reduction of cost for hardware and operations, and planned to stay that way henceforth**

B515 and B725 data center central network infrastructure layout (current state)

Configuration reached in 2021Q3 and operated since



B725 Network Room



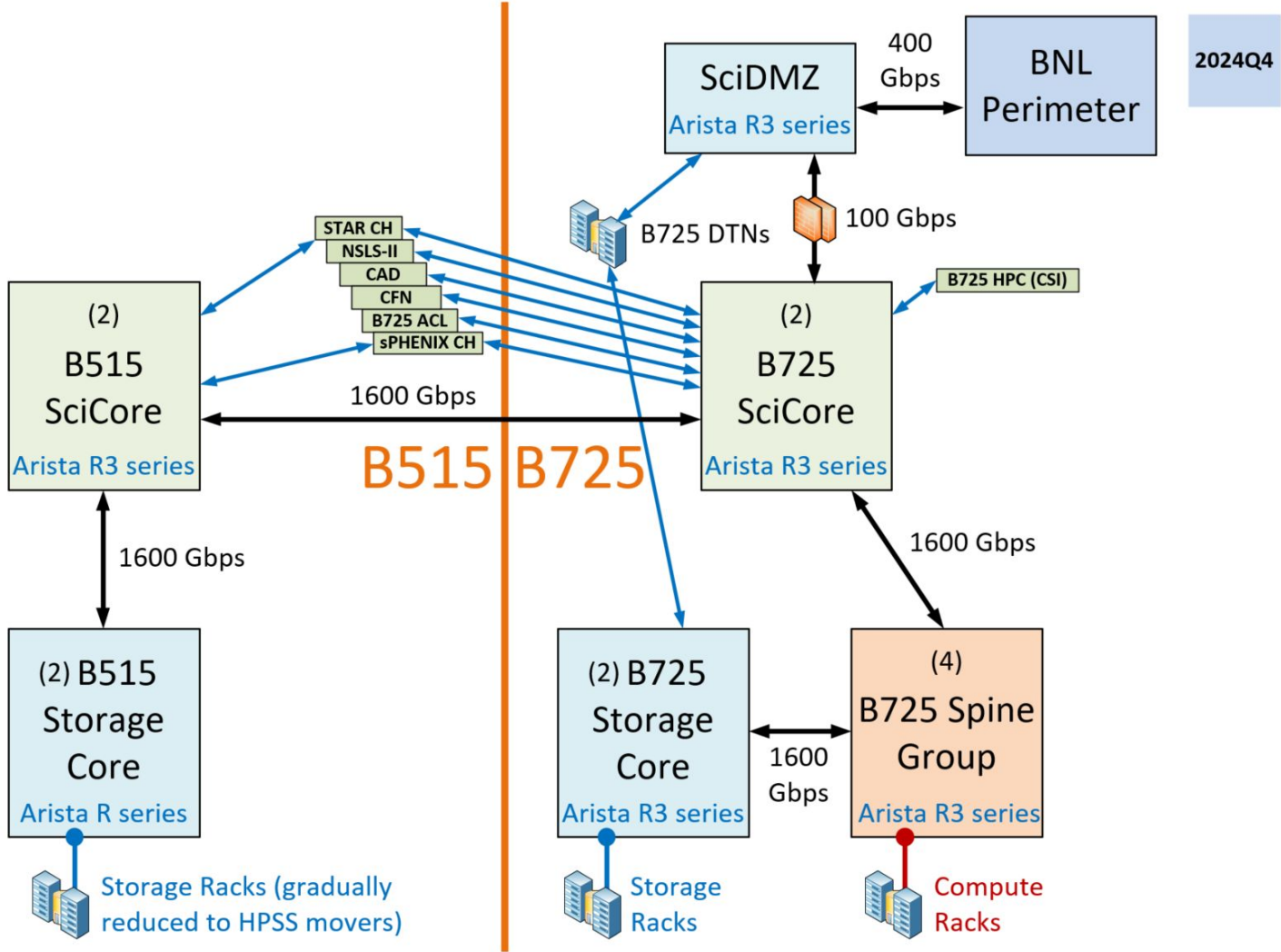
10x Arista 7508R3 400 GbE enabled chassis: 48x line card slots total, populated with 36-port 100 GbE linecards so far

Challenges related to ATLAS requirements (Networking & B515/B725 transition)

- **FY22** (retraction of all ATLAS systems from B515/BCF):
 - Performing the retirement of all remaining ATLAS equipment in B515 / BCF by the end of FY22 (no network connected equipment for ATLAS in the area henceforth)
 - Retiring all remaining ATLAS Tier-1 site CPU racks in B515 data center by in early FY23 (after the FY22 CPU purchased CPU racks are deployed for production)
- **FY23** (retraction of all ATLAS systems from B515/Sigma-7 and reduction of ATLAS deployment in B515 down to CDCE area - **effective completion of B515/B725 transition**):
 - Retiring all remaining ATLAS equipment in B515 / Sigma-7 by the end of FY23 (no network connected equipment for ATLAS in that area once it is done)
- **FY24-25** (reducing ATLAS deployment in B515 to a minimum):
 - Retiring all remaining ATLAS DISK equipment in B515 / CDCE by the end of FY25 (only ATLAS HPSS movers specific to 2x Oracle SL8500 tape libraries are going to remain in B515 for ATLAS past that point)

Transition to B515 and B725 datacenter: reaching the steady state in FY25

Only the HPSS movers serving Oracle SL8500 libraries are expected to remaining for ATLAS in B515 after FY25



Challenges related to ATLAS requirements (Network interventions / Eq. HW refresh)

- The network reconfiguration and upgrades are driven by the following overlapping processes in FY22-25:
 - SDCC SciZone perimeter firewall upgrade / physical move from B515 to B725 / HW refresh: the full HW refresh is postponed until FY26, while the BW capability upgrade from 40 Gbps to 160-200 Gbps is expected in FY23 - **to paid for by the programs, ATLAS share is 50%**
 - BNL Perimeter HW refresh (ESnet) - being prepared to in FY22 (400 GbE enabled edge router evaluation), purchases are expected to be executed in FY23, the equipment deployment for production, likely in early FY24 - **expected to be covered by non-program funds, the on-site evaluation of the new perimeter equipment is ongoing as of 2022Q2**
 - The physical migration of the SciZone Perimeter FW system and migration to the new BNL Perimeter are expected during Facility-wide ~10h long scheduled network intervention in 2023Q4 (early FY24, likely Dec 2023) - **the previous similar scale network intervention happened in 2020Q4, so the frequency of such interventions is well within the WLCG MoU limits**
- Network equipment remains most affected by the COVID induced market shortages in FY21-22, with lead times for components reaching out into **6 (DAC cables and NICs), 9 (most of Arista product line), 12 (Juniper product line) months** for the new orders placed in 2022Q1.
 - Thus, purchases of network equipment components are made well in advance to make sure that the effects of these lead times on CPU and DISK components are minimized (to the extent permitted by the full FY budget release timeframe within the FY period). **Currently, these issues are expected to continue into FY23 period.**

Backup Slides

ATLAS Equipment Deployment in B725/MDH as of Jun 2022 (22 racks total)

B725 Main Data Hall
(Jun 13, 2022)

ATLAS

Low Density (LD) Area (44 out of 158 racks populated)

78 racks (1.2 MW) :: PS #2

80 racks (1.2 MW) :: PS #1

