Finalizing Transition to the New Data Center at BNL

Imran Latif[1], *Shigeki* Misawa[1], and *Alexandr* Zaytsev[1,*]

[1]Brookhaven National Laboratory, Upton, NY 11973-5000, U.S.A.

Abstract. Computational science, data management and analysis have been key factors in the success of Brookhaven National Laboratory's scientific programs at the Relativistic Heavy Ion Collider (RHIC), the National Synchrotron Light Source (NSLS-II), the Center for Functional Nanomaterials (CFN), and in biological, atmospheric, and energy systems science, Lattice Quantum Chromodynamics (LQCD) and Materials Science, as well as our participation in international research collaborations, such as the ATLAS Experiment at Europe's Large Hadron Collider (LHC) at CERN (Switzerland) and the Belle II Experiment at KEK (Japan). The construction of a new data center is an acknowledgement of the increasing demand for computing and storage services at BNL in the near term and enable the Lab to address the needs of the future experiments at the High-Luminosity LHC at CERN and the Electron-Ion Collider (EIC) at BNL in the long term. The Computing Facility Revitalization (CFR) project is aimed at repurposing the former National Synchrotron Light Source (NSLS-I) building as the new data center for BNL. The construction of the new data center was finished in 2021Q3, and it was delivered for production in early FY2022 for all collaborations supported by the Scientific Data and Computing Center (SDCC), including STAR, PHENIX and sPHENIX experiments at RHIC collider at BNL, the Belle II Experiment at KEK (Japan), and the Computational Science Initiative at BNL (CSI). This paper highlights the key mechanical, electrical, and networking components of the new data center in its final configuration as used in production since 2021Q4 and gives an overview for the extension of the central network systems into the new data center and the migration of a significant portion of IT load and services from the old data center to the new data center carried out in 2021-2023, with expected completion of the main phase of the gradual IT equipment replacement and migration from the old data center into the new one set to the end of FY2023 (Sep 30, 2023).

1 Introduction

The Computing Facility Revitalization (CFR) project at Brookhaven National Laboratory (BNL) is aimed at repurposing the former National Synchrotron Light Source (NSLS) building as a new data center for the Scientific Data and Computing Center (SDCC). The

* Corresponding author: alezayt@bnl.gov

new data center's construction was finished in 2021Q3 and it has become operational in 2021Q4 and available to host compute, disk storage and tape storage equipment for all groups and collaborations supported by the SDCC including, but not limited to: the ATLAS experiment [1, 2] at the LHC accelerator at CERN (European Center for Nuclear Research) for which SDCC Facility hosts the USATLAS Tier-1 site [3], the STAR, PHENIX and sPHENIX experiments at the RHIC collider [4] at BNL, the Belle II experiment [5] at the High Energy Accelerator Research Organization (KEK) [6] in Japan for which SDCC Facility hosts the US Belle II Tier-1 site, and Computational Science Initiative [7] at BNL for which SDCC Facility hosts the BNL Institutional Cluster [8]. The period of migration of IT equipment and services to the new data center started with installation of new core network equipment and the new tape library pair for BNL ATLAS Tier-1 site in the new data center in 2021Q3 (within U.S. fiscal year (FY) 2021, covering Oct 2020 to Sep 2021 period). The transition process continued in FY2022 with deployment of the first sPHENIX tape library pair as well as installation of new compute and disk storage systems on the floor of the new data center that brought the new datacenter to the level of 54 active storage and CPU racks by the end of FY2022 period, which included a set of racks serving beamlines of National Synchrotron Light Source (NSLS) II [9] at BNL as well.

The scaling of information technology (IT) load deployment continued into FY2023 period with projected number of active storage and CPU racks on the floor of new datacenter to 106 racks by the end of the period. The process of scaling of IT load of the new datacenter in FY2022 and FY2023 allowed the reduction of the occupied rack count of systems deployed the old data center by 135 racks to take place as projected for the end of FY2023 period, out of which 20 racks were physically moved to the new datacenter and the rest – retired and replaced by the new systems, following the normal hardware refresh cycle. This allows the retirement of the oldest areas of SDCC data center environment to be retired in October 2023, reducing the total active storage and CPU rack count in the old datacenter below 85 starting from FY2024. Furthermore, a series of network interventions was performed in FY2023 aiming at finalizing the transition from central network components deployed in the old data center to those of the new datacenter becoming the primary source of internal connectivity for the entire SDCC data center environment. The state expected to be reached by the SDCC datacenter environment by end of FY2023 is expected to mark the end of the main phase of the transition period from the old data center to the new one and the beginning of the period of normal operations for the upcoming years (FY2024-FY2027) during which a stable configuration for central network systems is expected to be maintained and no large scale migration of equipment and services expected to occur between the parts of the old datacenter remaining in production and the new datacenter.

2 SDCC Data Center Environment

2.1 Overview

The old (BNL Campus building 515 based) 1,940 m^2 SDCC data center dates from the 1960's, with some additions made in 2009. It is a Tier I or "non-redundant" data center as defined by the Uptime Institute's Tier classification system [10]. A detailed description of the old data center and its capabilities is given in [11].

The new (BNL Campus building 725 based) SDCC data center, being built in the shell of the former NSLS light source building (see Figure 1), is a Tier III class data center that meets the RAS requirements of the SDCC Facility, which requirements are primarily driven by the service level agreements set in the U.S. commitments to the Worldwide LHC

Computing Grid, a global collaboration of computing centers supporting computing for the LHC [12, 13]. All critical data processing equipment will be supported by a fully redundant infrastructure (N+1) that is concurrently maintainable (i.e., without facility shutdown). The data center is also fully self sufficient, capable of operating without utility power or BNL campus chilled water for prolonged periods of time (provided that the diesel generator group is getting refueled). The data center design targets a PUE of 1.2, and the seasonal variation of the monthly averaged PUE is expected to be in the 1.17–1.28 range given the climate conditions at BNL geographical location, so the real world PUE should be well below the 1.4 maximum for new data centers mandated by the U.S. Federal Government executive order in effect at the time of project definition [14]. The initial buildout of the new data center as delivered by the CFR project is capable of supporting 3.6 MW of combined IT load provided with full battery UPS protection and 3.5 MW of diesel backup capability (covering both IT load and PUE component). When fully built the new data center will be able to support 9.6 MW of IT load provided with full battery UPS protection and 14 MW of diesel backup capability (covering both IT load and PUE component with N+1 redundancy), six 18-frame tape libraries, and 478 standard 42U, 19 in. wide, up to 1200 mm deep equipment racks. Total IT floor space is roughly 1,600 m^2 (17,000 ft^2). More details on the design features and capabilities of the new datacenter are also given in [11].

2.2 Layout

The IT equipment in the new data center is split between three separate rooms: a dedicated Tape Library Room, a Network Room, and the Main Data Hall (MDH) for compute and storage equipment. Separate rooms are used as the power, cooling and fire suppression systems are different for the three rooms. Figure 1 shows the location of these rooms in the former NSLS building. All support equipment, except cooling towers, high voltage switch gear and diesel generators, are located within the building shell in the areas surrounding the IT rooms.

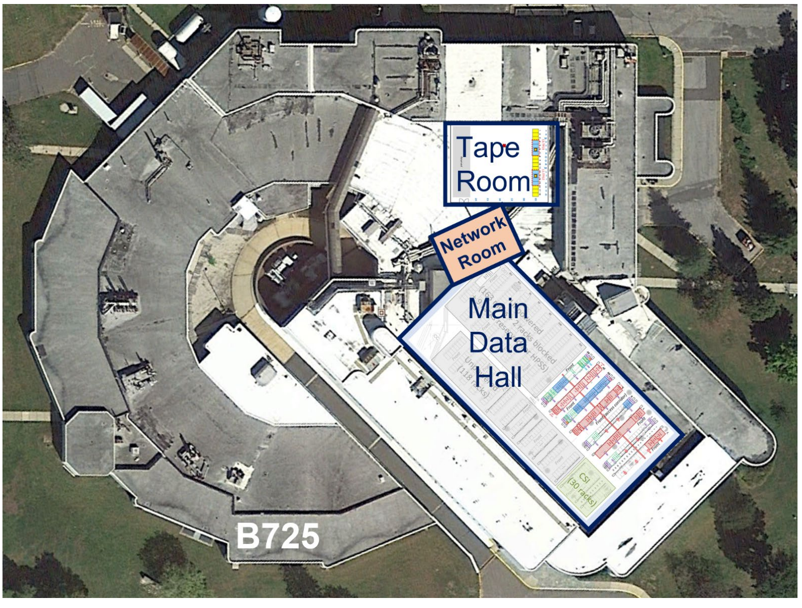

Fig. 1. New Data Center – Layout within former NSLS building.

The Main Data Hall (MDH), shown in Figure 2, is roughly 1,115 m^2 (12,000 ft^2) and is split into two areas: a High Throughput Computing (HTC) area and a High Performance Computing (HPC) area. The HTC area can host 16 rows of equipment, with 20 standard equipment racks per row. However, only 8 rows are provided with power and cooling pipes in the initial (37.5%) buildout. The HPC area consists of space for 15 rows of up to 10 standard equipment racks, of which 3 rows are energized in the initial buildout. The population of HTC area started in FY2021 and the HPC area – in FY2022.

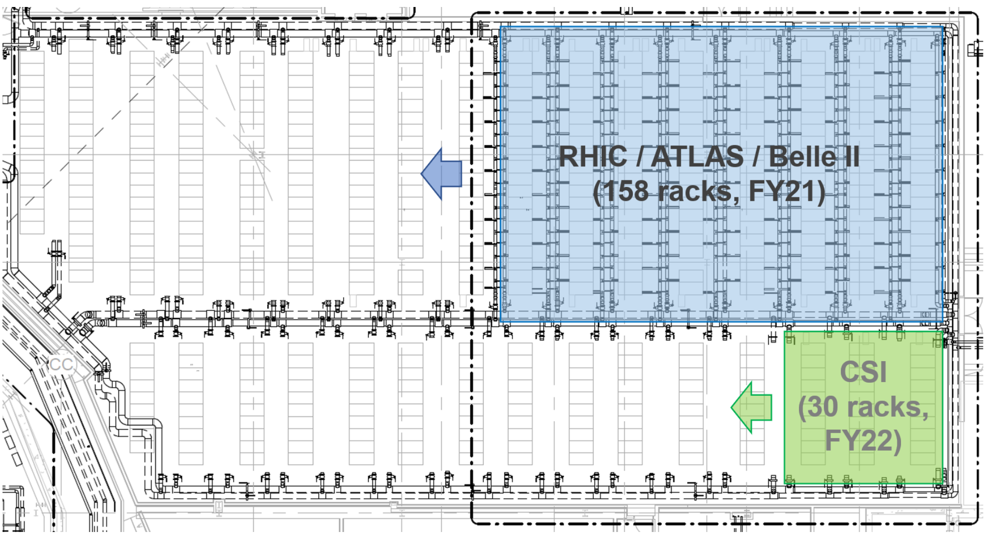

Fig. 2. New Data Center – Main Data Hall (MDH) floor plan.

2.3 Central Networks

The Ethernet networks of SDCC datacenter environment are based on the network core provided to each of the data centers involved, each core comprised of central router pair (Science Core), ECMP routing based Spine Group (set of 4 modular switches) serving connectivity to the (non-modular) top-of-the-rack (ToR) switches of CPU racks, and the Storage Core (a set of modular switches) serving connectivity to the storage racks. The SDCC datacenter environment is also hosting network perimeter equipment and ESnet [15] edge equipment serving WAN connectivity to the entire BNL site and specifically LHCOPN and LHCONE [16] connectivity to WLCG connected systems hosted in SDCC environment.

A high level network connectivity layout for the SDCC data center as of the beginning of transition between old and the new datacenters in 2021Q3 is shown in Figure 3. The interbuilding link allowing inter-operations between both data centers was initially configured with 1.6 Tbps of unidirectional aggregate bandwidth, which was determined to be sufficient for the transition period and expected to remain in this configuration for the duration of FY2024.

The state of central network systems acquired as a result of network interventions carried out between 2021Q4 and the end of FY2023 period are shown in Figure 4. Most notable changes here are introduction of the redundant BNL perimeter deployment with half of equipment remaining in the old data center ("Lab C" area) and another half placed in the new datacenter thus forming a redundant pair with building level redundancy, and also relocation of B515 Storage Core modular switch pair to the area of the old data center ("CDCE" area) that is expected to be kept in operations past the end of FY2023 period.

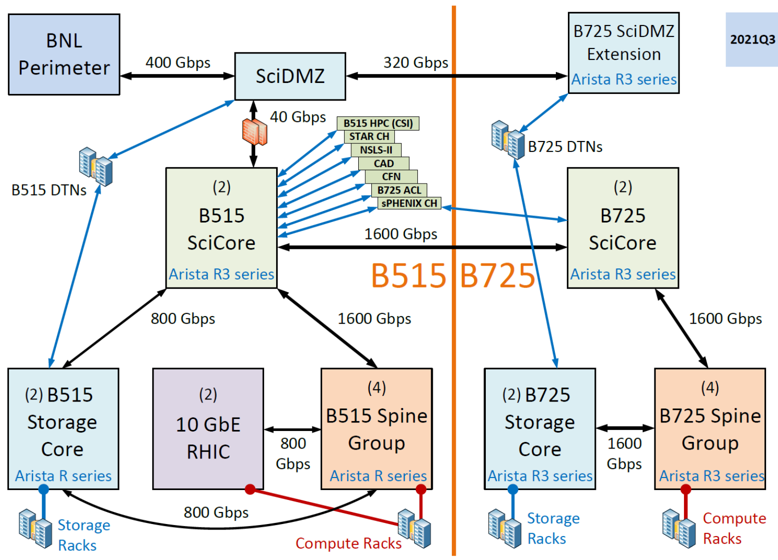

Fig. 3. High level diagram of network systems integration of the old data center (B515 on the left) and the new data center (B725 on the right) at the beginning of transition between old and the new datacenter in 2021Q3.

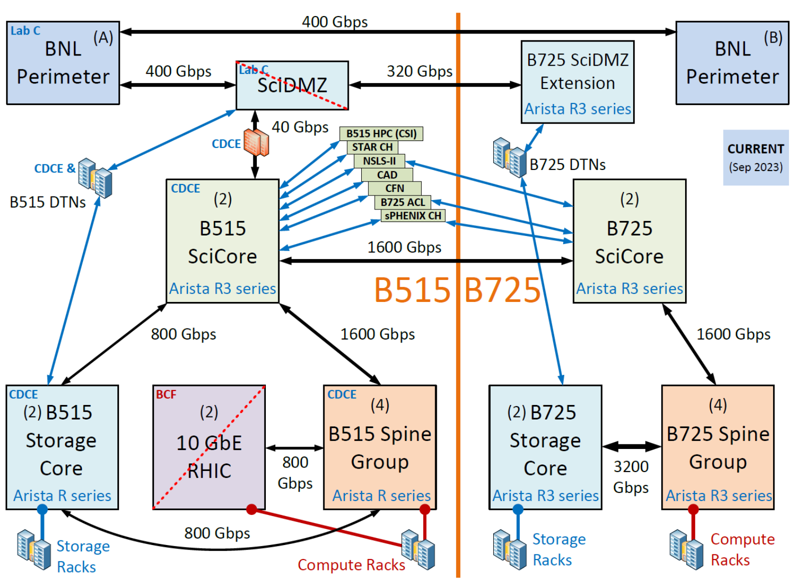

Fig. 4. High level diagram of network systems integration of the old data center (B515 on the left) and the new data center (B725 on the right) at the end of the main phase of the transition between old and the new datacenter in 2023Q3 (September 2023).

The state central network systems expected to be acquired by the end of calendar year 2023 is shown in Figure 5.

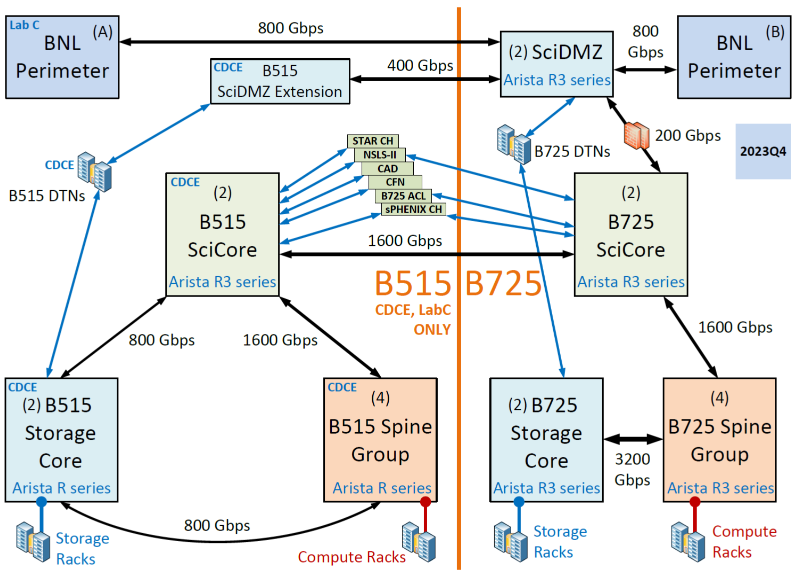

Fig. 5. High level diagram of network systems integration of the old data center (B515 on the left) and the new data center (B725 on the right) for the state projected for the end of calendar year 2023 (2023Q4).

2.4 Data Center Migration

Migration to the new data center was being performed in a manner of a phased migration, with newly purchased equipment installed in the new data center starting from 2021Q3, and concurrent staged retirement of existing equipment in the old data center. Under such arrangements the set of equipment rack that needed to be physically moved from the old data center to the new one in the entire FY21-23 period consisted only of 20 CPU racks which corresponded to about 8% of the number of racks with IT equipment deployed in the old data center as of Feb 2021. To make the phased migration possible, the SDCC network was extended to the new data center in 2021Q2-3 as described above in a completely transparent way for the old data center operations. The main phase of the process of gradual transition to the new data center is expected to be completed by the end of FY2023 (end of September 2023) and the period of normal operations with stable central network configuration and gradual scaling of IT load in the new data center is expected to start from the beginning of FY2024 (October 2023).

3 Summary

The new data center at BNL substantially enhanced the capabilities of the SDCC Facility. From the perspective of SDCC customers, the new data center allowed SDCC to support more and higher power / higher density IT equipment and at the same time increased the facility robustness and resiliency to the infrastructure issues. Gradual migration of the majority of IT load and services has been carefully planned and performed in FY2021-

FY2023 period with minimal impact on uptime for the facility in general and the equipment hosted by SDCC in particular. The main phase of this transition is now on the path to be completed successfully by the end of FY2023 (end of September 2023). From the SDCC Facility perspective of the SDCC, it reduced operational complexity, maintenance burdens, energy consumption and simplified the installation of new IT equipment and allowed SDCC to meet the requirements of scientific collaboration involved up to and including FY2023. The new data center with its modular design and ability to support up to 9.6 MW of IT load in the full buildout provides the path forward for SDCC to continue addressing these requirements for the years to come.

References

[1] The ATLAS Collaboration *et al.,* The ATLAS Experiment at the CERN Large Hadron Collider, 2008 JINST 3 S08003: https://iopscience.iop.org/article/10.1088/1748-0221/3/08/S08003/meta

[2] The ATLAS Collaboration, "The ATLAS Experiment at CERN": https://atlas.cern

[3] USATLAS Tier-1 site at BNL: https://www.sdcc.bnl.gov/experiments/usatlas

[4] Brookhaven National Laboratory, "RHIC Relativistic Heavy Ion Collider": https://www.bnl.gov/rhic

[5] The Belle II Collaboration, "The Belle II Experiment at KEK (Japan)": https://www.belle2.org

[6] The Belle II Facility at KEK (Japan): https://www.kek.jp/en/Facility/IPNS/Belle2/

[7] Brookhaven National Laboratory, "Computational Science Initiative": https://www.bnl.gov/compsci

[8] BNL Institutional Cluster: https://www.sdcc.bnl.gov/information/institutional-cluster

[9] Brookhaven National Laboratory, “National Synchrotron Light Source (NSLS) II”: https://www.bnl.gov/nsls2/

[10] Uptime Institute, Tier Classification System: https://uptimeinstitute.com/tiers

[11] “Finalizing Construction of a New Data Center at BNL”. Proceedings of vCHEP2021 ("25th International Conference on Computing in High Energy and Nuclear Physics"), CERN, May 17-21, 2021. Imran Latif et al EPJ Web of Conferences 251, 02069 (2021).

[12] Worldwide LHC Computing Grid (WLCG): https://wlcg.web.cern.ch

[13] Worldwide LHC Computing Grid (WLCG), "Signed Memoranda of Understanding": https://wlcg.web.cern.ch/mou/signed

[14] U.S. Federal Government. (Mar 19, 2015), "Executive Order (EO) 13693, Planning for Federal Sustainability in the Next Decade": https://www.fedcenter.gov/programs/eo13693/

[15] DOE Energy Science Network: https://www.es.net

[16] “LHCOPN and LHCONE: Status and Future Evolution”. E Martelli and S Stancu 2015 J. Phys.: Conf. Ser. 664 052025.