Finalizing Construction of a New Data Center at BNL

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

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

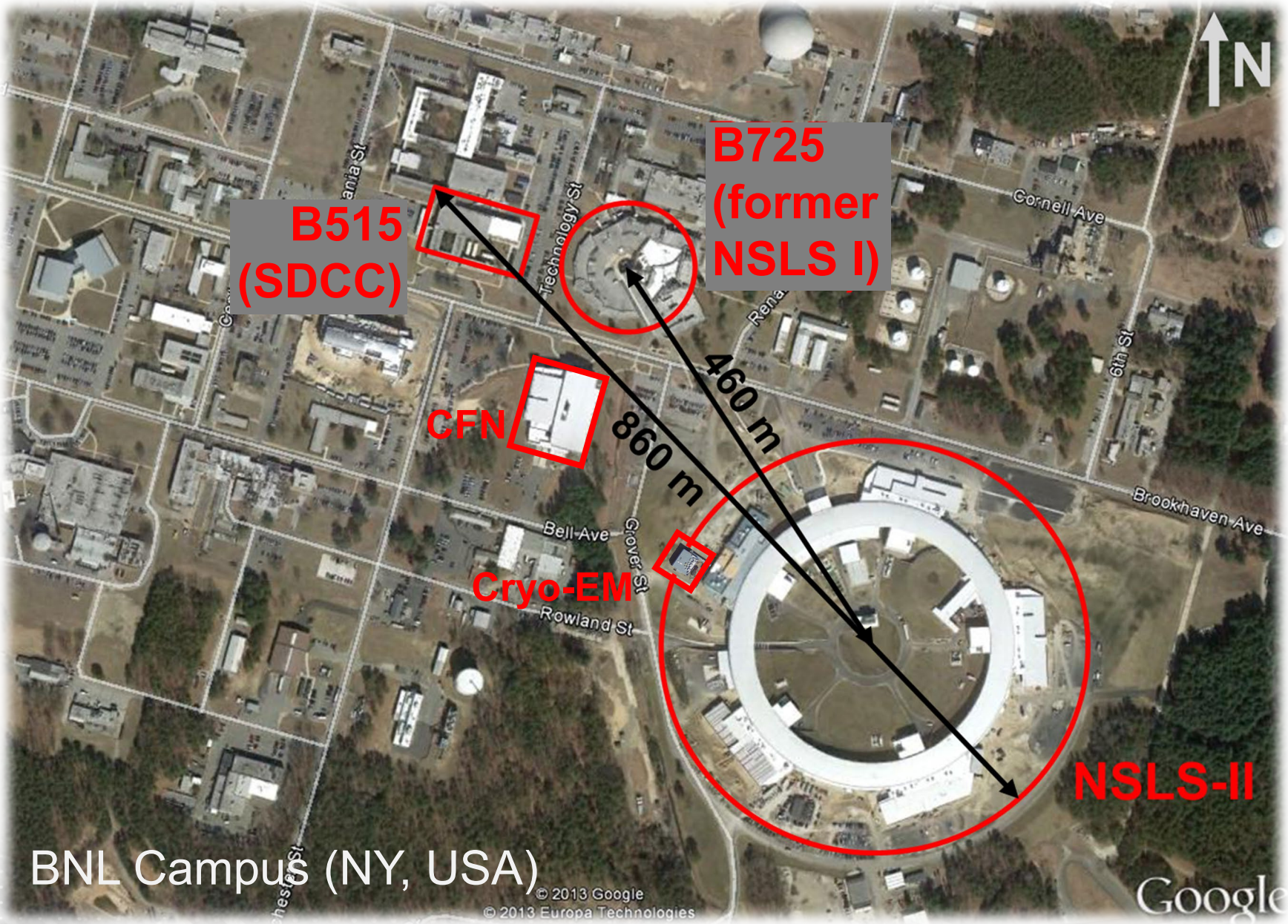
[*] alezayt@bnl.gov

25th International Conference on

Computing in High-Energy and Nuclear Physics (vCHEP2021)

B515 Data Center

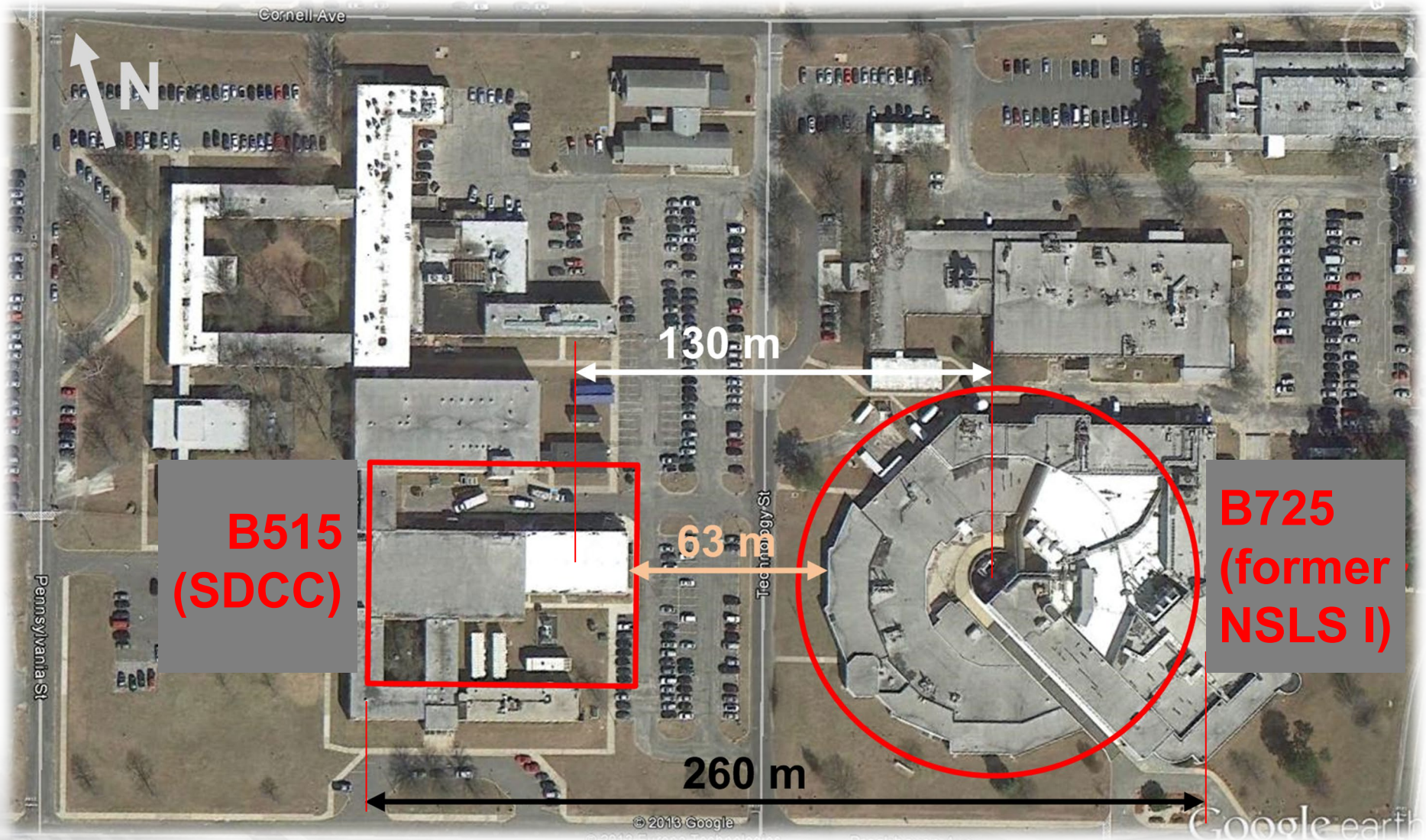
- Scientific Data and Computing Center (SDCC) Facility is currently operating a single B515 based data center which is a 1.5 MW scale air-cooled general purpose facility hosting 300 racks of equipment and 10 tape silos as of 2021Q2:
  - HTC and HPC computing systems (on the scale of 2500 nodes)
  - DISK storage (on the scale of 90 PB)
  - Robotic tape storage (on the scale of 200 PB)
- Serving multiple international collaborations and research communities:
  - STAR, PHENIX and sPHENIX at RHIC (BNL) with two on site counting houses: one active (STAR CH) and another one expected to be reconnected to SDCC in 2021Q3 (sPHENIX CH)
  - ATLAS Experiment at the LHC (ATLAS Tier-1 Site)
  - Belle II Experiment at KEK (Belle II Tier-1 Site)
  - National Synchrotron Light Source (NSLS) II at BNL, currently with 29 active beamlines (expected to scale up to 60+ beamlines in the future)
  - Center for Functional Nanomaterials (CFN) at BNL
  - BNL Computational Science Initiative (CSI) research groups and test labs
  - Cryo-EM Research Facility at BNL (currently in the process of joining in)


B515
(SDCC)
B725
(former
NSLS I)
CFN
Cryo-EM
460 m
860 m
NSLS-II
N
Technology St
Cornell Ave
6th St
Brookhaven Ave
Bell Ave
Grover St
Rowland St
BNL Campus (NY, USA)
© 2013 Google
© 2013 Europa Technologies
Google

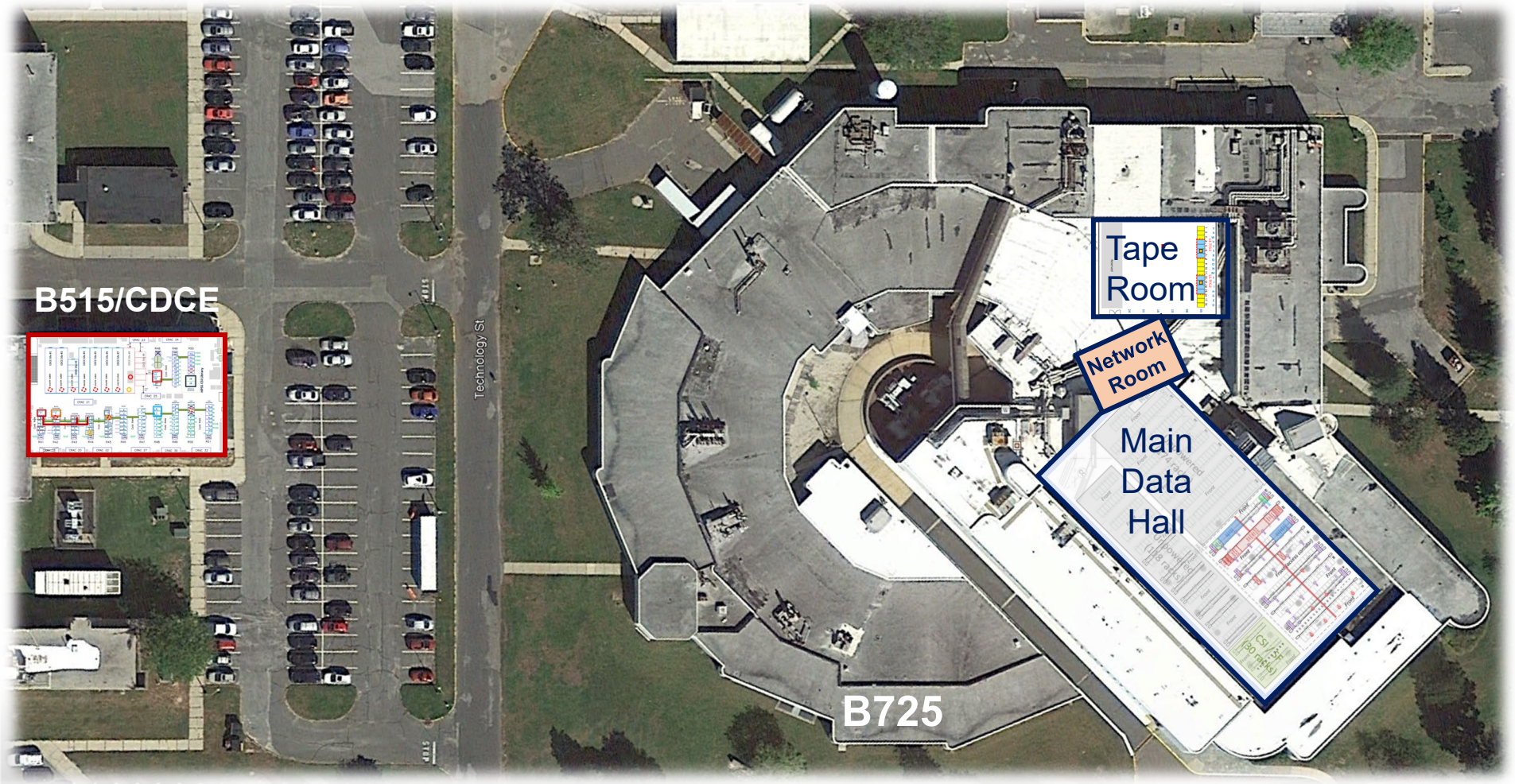
The Path Forward

- The existing SDCC B515 data center is a highly non-uniform facility aggregating the history of several decades of infrastructure solutions, with none of its areas providing the feature set needed for addressing the future points of growth associated with:
  - sPHENIX Experiment at RHIC beginning the data taking in (US)FY23 and gradually scaling up in FY23-25
  - STAR Experiment at RHIC continuing taking data in FY23-25
  - ATLAS Experiment at High Luminosity LHC (HL-LHC) starting from FY27
  - Scaling of the NSLS-II Facility to 60 active beamlines
- Furthermore, there are the following challenges to address driven by the need to increase IT equipment power density and efficiency of operations:
  - Increasing the power density of HTC CPU racks up to 20 kW/rack (current limit is about 12 kW/rack)
  - Scaling up the CSI HPC systems and increasing the power density of HPC racks up to 30 kW/rack (current limit is about 15 kW/rack)
  - Scaling the combined Facility power profile beyond 2 MW of IT load while protecting all its IT payload from the site-wide power outages and allowing the payload to remain operational during the prolonged site-wide power outages
  - Reduce overall PUE of the SDCC Facility below 1.4

B515 & B725 Data Centers



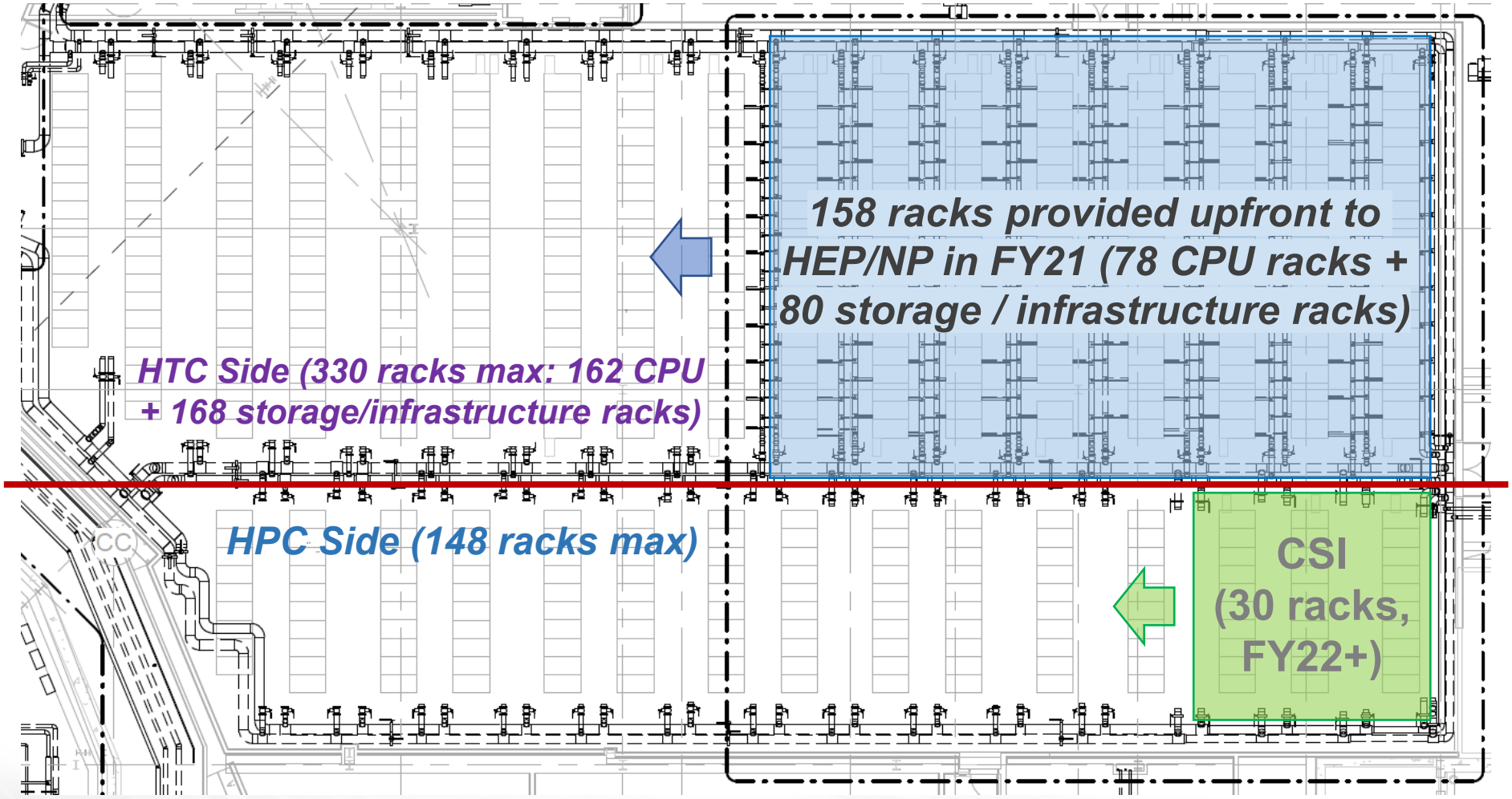
New data center is designed & being constructed for the SDCC Facility in B725 (former NSLS I building) in FY19-21 period, with migration of most of the DISK and all the CPU resources (primarily via gradual hardware refresh process) to the new data center from the existing B515 data center in FY21-23



BNL Core Facility Revitalization (CFR) Project: Design & Construction of B725 Data Center

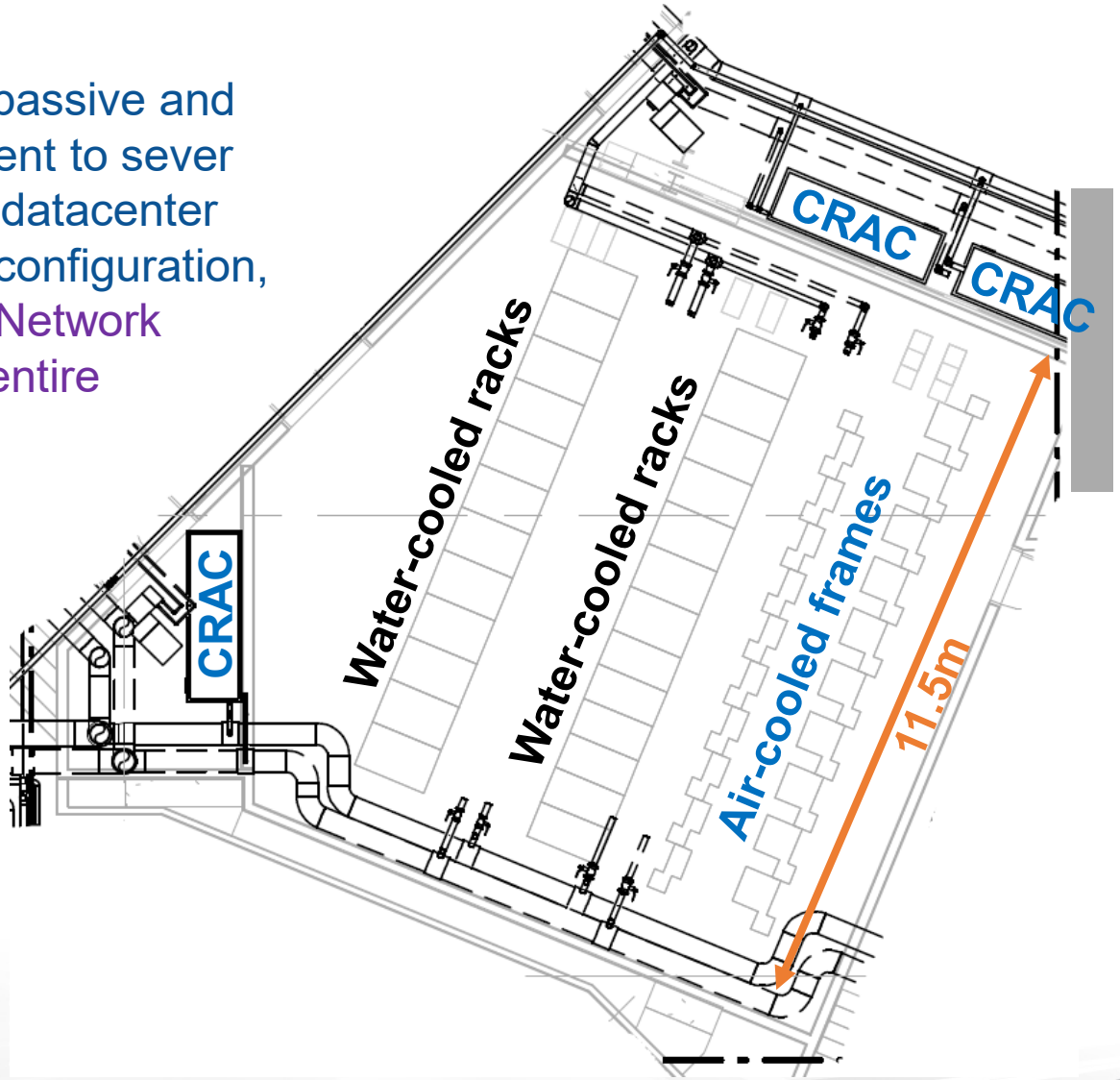
- Main design features of the SDCC B725 data center:
  - A single large data hall for CPU & DISK resources **(Main Data Hall)** divided into 2 aisles:
    - HTC (RHIC/ATLAS/Belle II) aisle: 16 rows of ~20 racks each + one row of 16 racks
    - HPC (CSI): 14 rows of 10 racks each plus one row of 8 racks)
    - 478 rack positions / up to 9.6 MW of IT load (all UPS/diesel protected) in a fully built out configuration, N+1 redundancy for all infrastructure components
  - 188 rack positions to become available starting from 2021Q3
    - 158 rack positions for RHIC/ATLAS/Belle II with 2.4 MW of power/cooling available
    - 30 rack positions for CSI with 900 kW of power/cooling available
    - 2.4 MW of diesel generator backup power (IT load) available
    - Unlocking 290 remaining rack positions will require construction of additional electrical rooms, installation of additional power distribution and UPS equipment, chillers, cooling towers and diesel generators (in 1.2 MW of usable IT power increments).
  - Dry-pipe/pre-action double interlock sprinkler system for fire suppression in the MDH
  - Inergen based fire suppression in the Tape Room and the Network Room
  - Standard APC 42U racks (600mm wide, 1070 mm deep (HTC compute) or 1200 mm deep (high capacity JBOD storage)) are to be used across the entire floor of B725 MDH:
    - All equipped with watercooled rear-door heat exchangers with chilled water supplier from under the raised floor (nothing but water is distributed under the raised floor in the MDH)
    - Isolation valves on the row-level and individual rack level on the water pipes
    - Zoned drainage system in the concrete floor
  - 3 level of overhead cable/power distribution: power/busbar, fiber tray with mini-racks attached, copper tray with RJ-45 patch panels attached

B725 Data Center: Main Data Hall



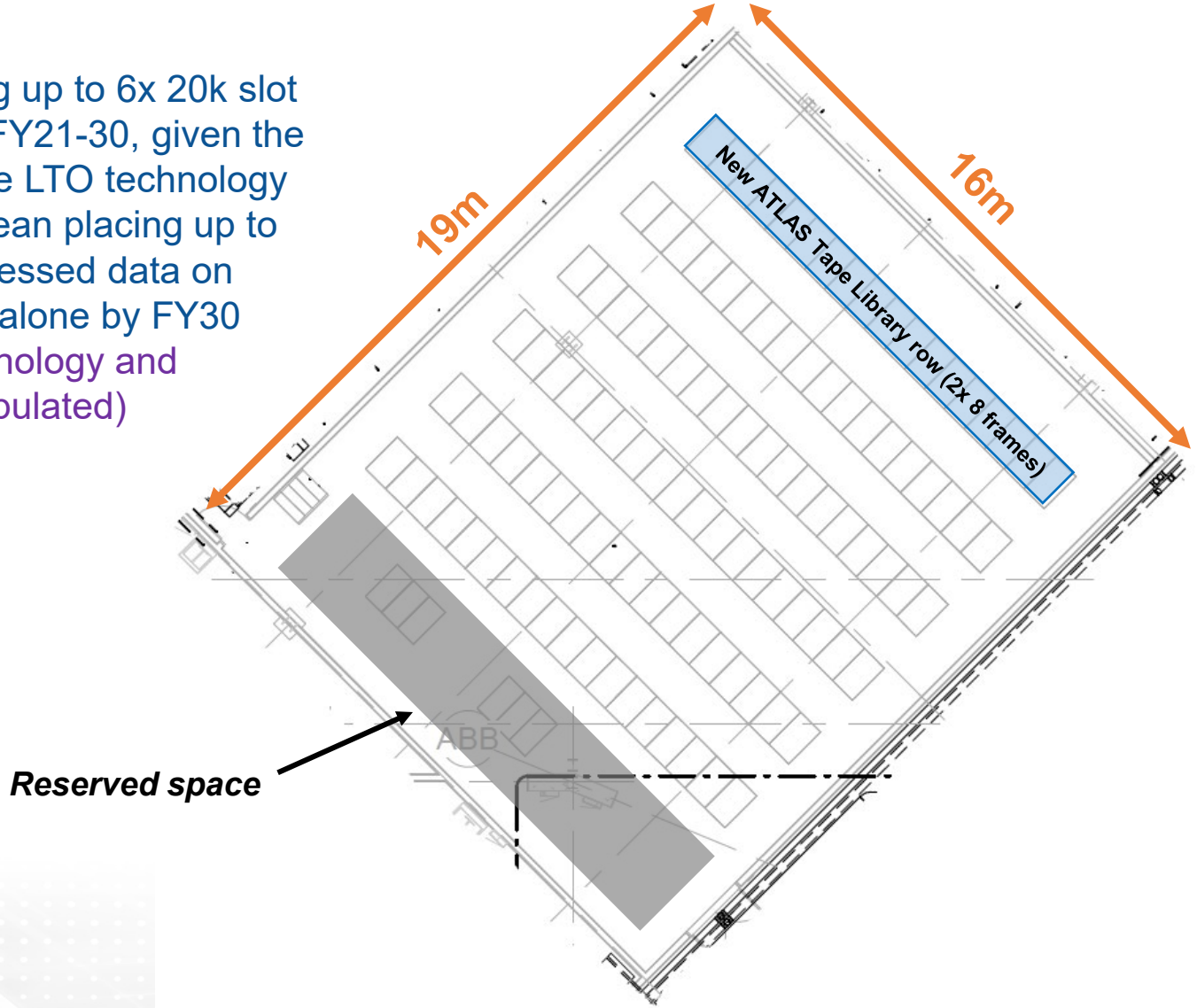
B725 Data Center: Network Room

Capable of hosting all passive and active network equipment to sever the fully built out B725 datacenter in 9.6 MW / 478 racks configuration, including ESnet / BNL Network Perimeter serving the entire BNL site, if required



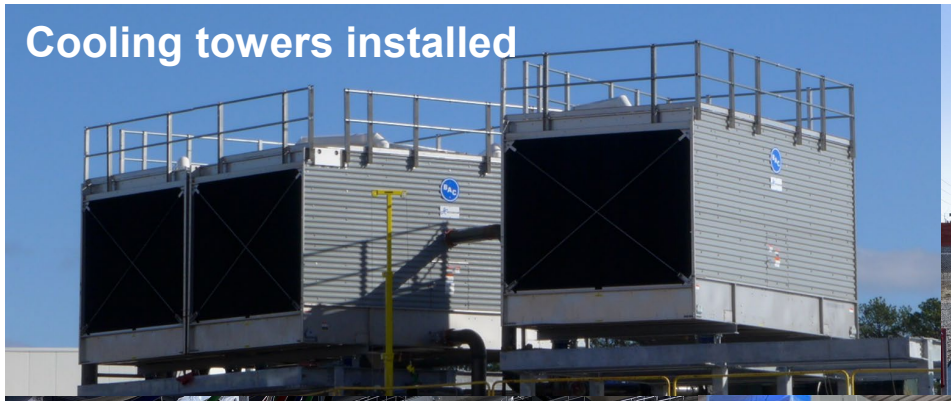
B725 Data Center: Tape Room

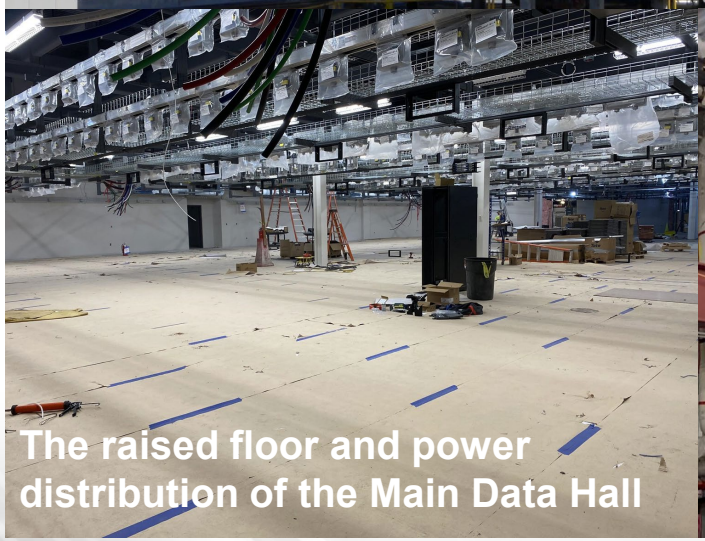
Capable of hosting up to 6x 20k slot libraries which in FY21-30, given the expectations of the LTO technology evolution could mean placing up to 11 EB of uncompressed data on TAPE in this area alone by FY30 (with LTO-12 technology and 4 libraries fully populated)

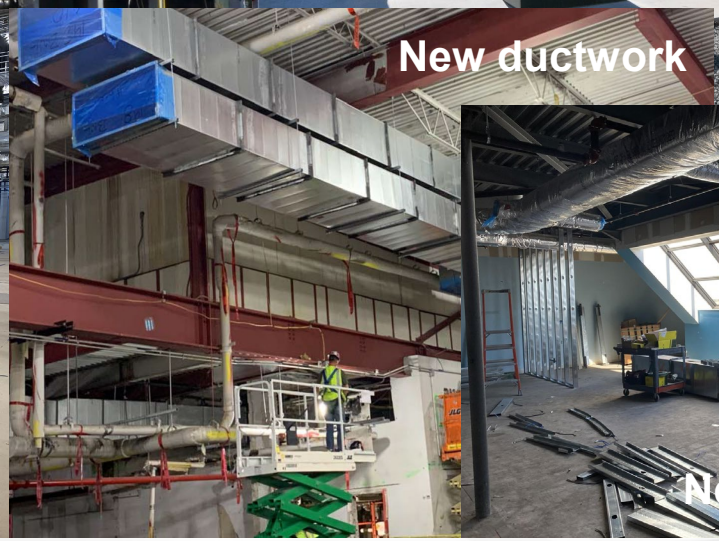


Status of B725 Datacenter Construction

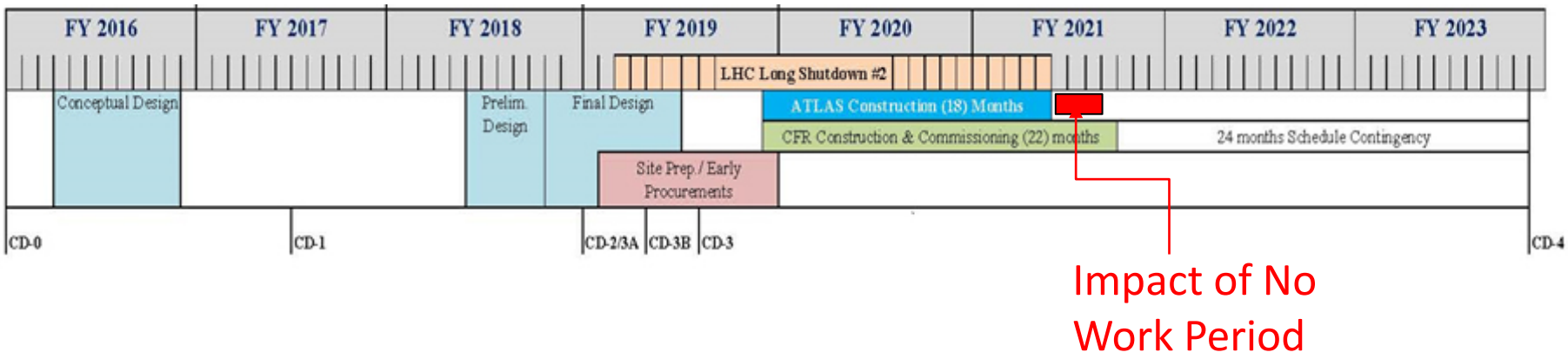
July 2020 – Mar 2021: construction is going ahead after 3 months of delay in 2020Q2 due to COVID-19. The early occupancy of B725 datacenter is expected to begin for ATLAS in June 2021 (network equipment deployment has started May 2021; occupancy for all tenants is expected to start in July 2021)





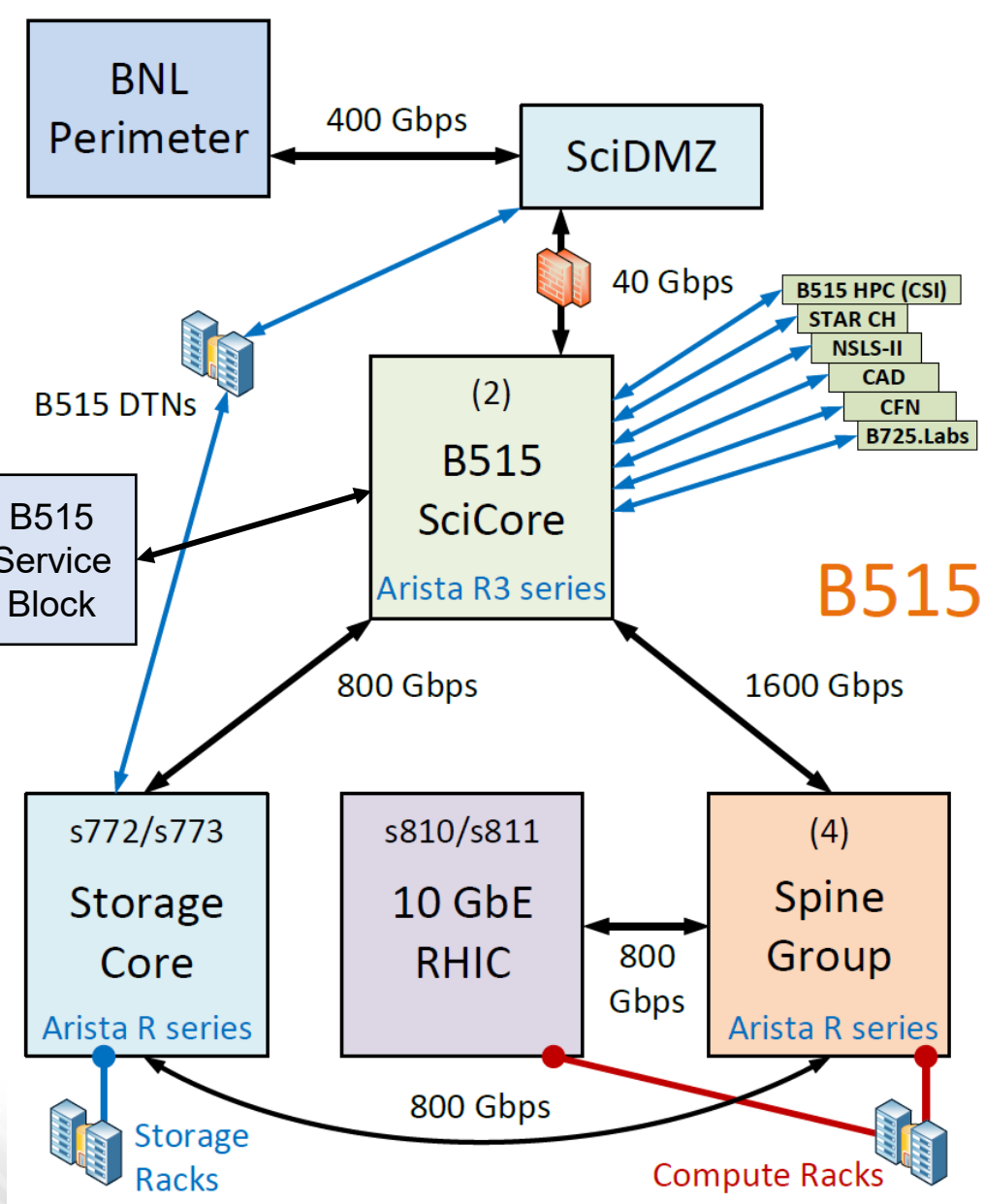


CFR Project and Impact of COVID-19



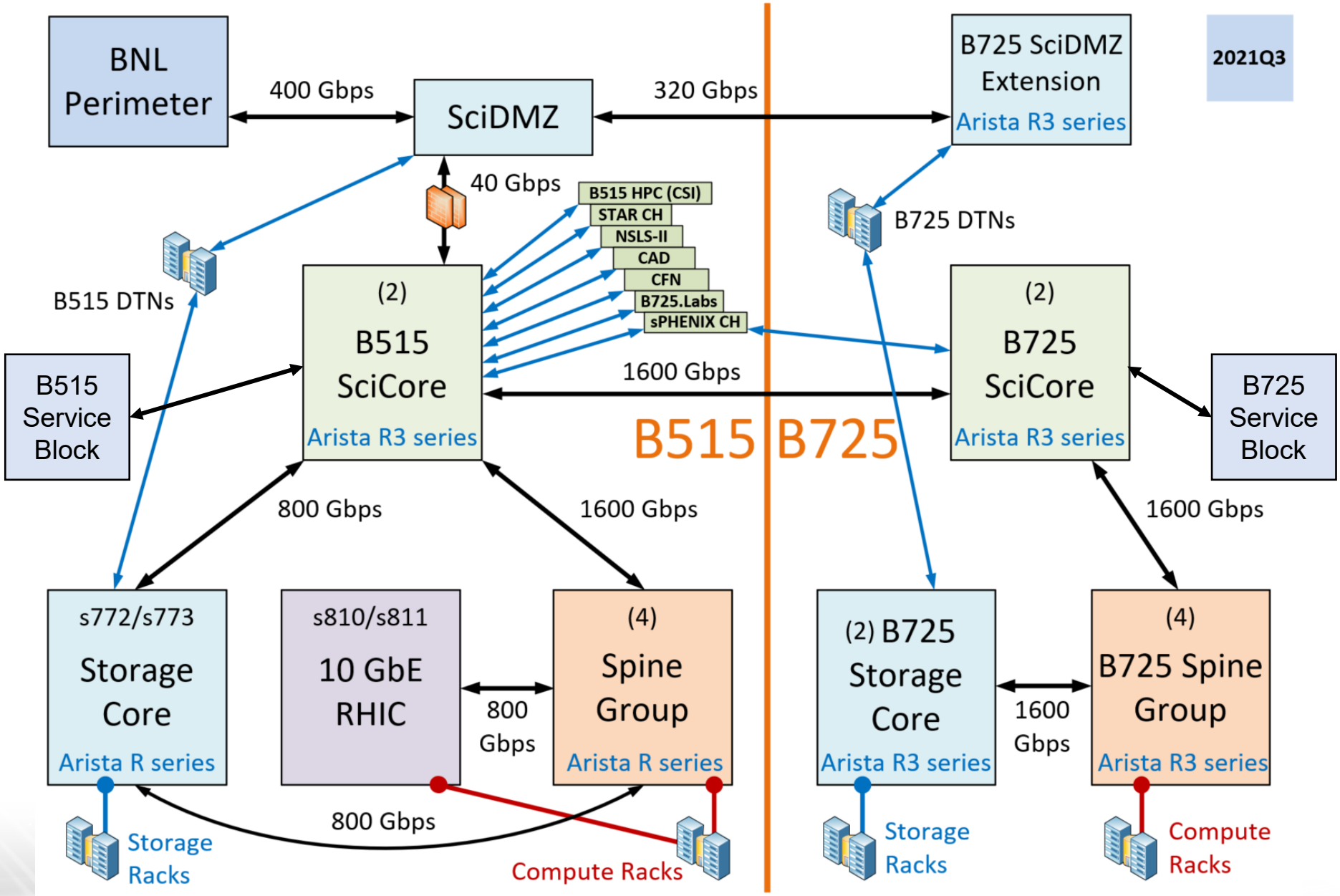
COVID-19 Schedule Impacts:

- 3 Months “No work” period (April, May, June 2020)
- Delays realized for equipment procurements (6-8 weeks)
- Delays realized for de-mobilization and re-mobilization as well as phased re-start of construction activity
- The early occupancy shifted from 2021Q1 to 2021Q3 as a result



2020Q4

The expansion of SDCC central network into B725 has started in 2021Q2 in order to have it ready for deploying IT payload on the floor of B725 in 2021Q3

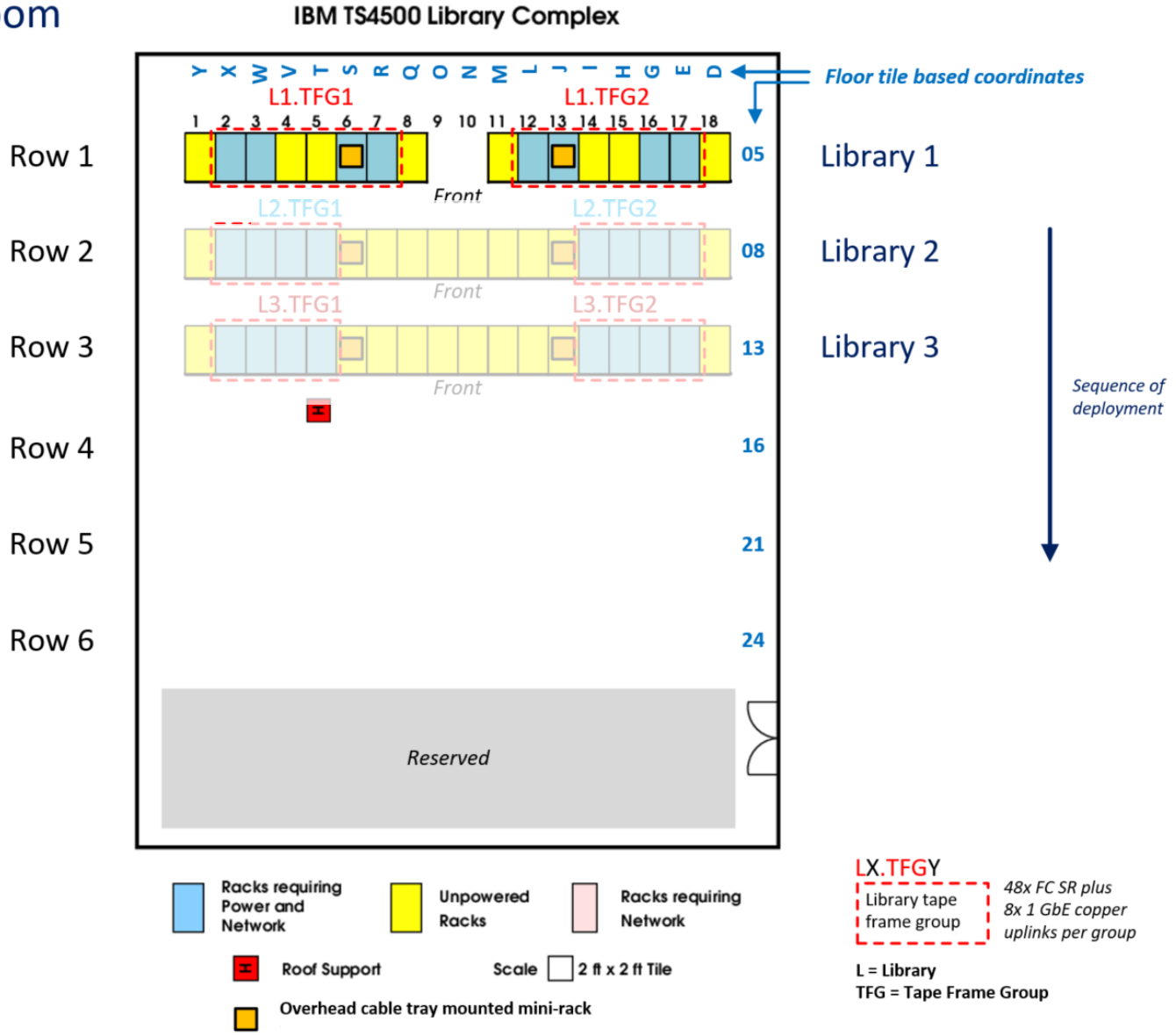

BNL Perimeter
400 Gbps
SciDMZ
320 Gbps
B725 SciDMZ Extension
Arista R3 series
2021Q3
40 Gbps
B515 HPC (CSI)
STAR CH
NSLS-II
CAD
CFN
B725.Labs
sPHENIX CH
B725 DTNs
B515 DTNs
(2)
B515 SciCore
Arista R3 series
(2)
B725 SciCore
Arista R3 series
1600 Gbps
B515 Service Block
B725 Service Block
B515
B725
800 Gbps
1600 Gbps
1600 Gbps
s772/s773
Storage Core
Arista R series
s810/s811
10 GbE RHIC
800 Gbps
(4)
Spine Group
Arista R series
(2) B725 Storage Core
Arista R3 series
1600 Gbps
(4)
B725 Spine Group
Arista R3 series
800 Gbps
Storage Racks
Compute Racks
Storage Racks
Compute Racks

B725 Tape Room

ATLAS (FY21)

sPHENIX (FY22)

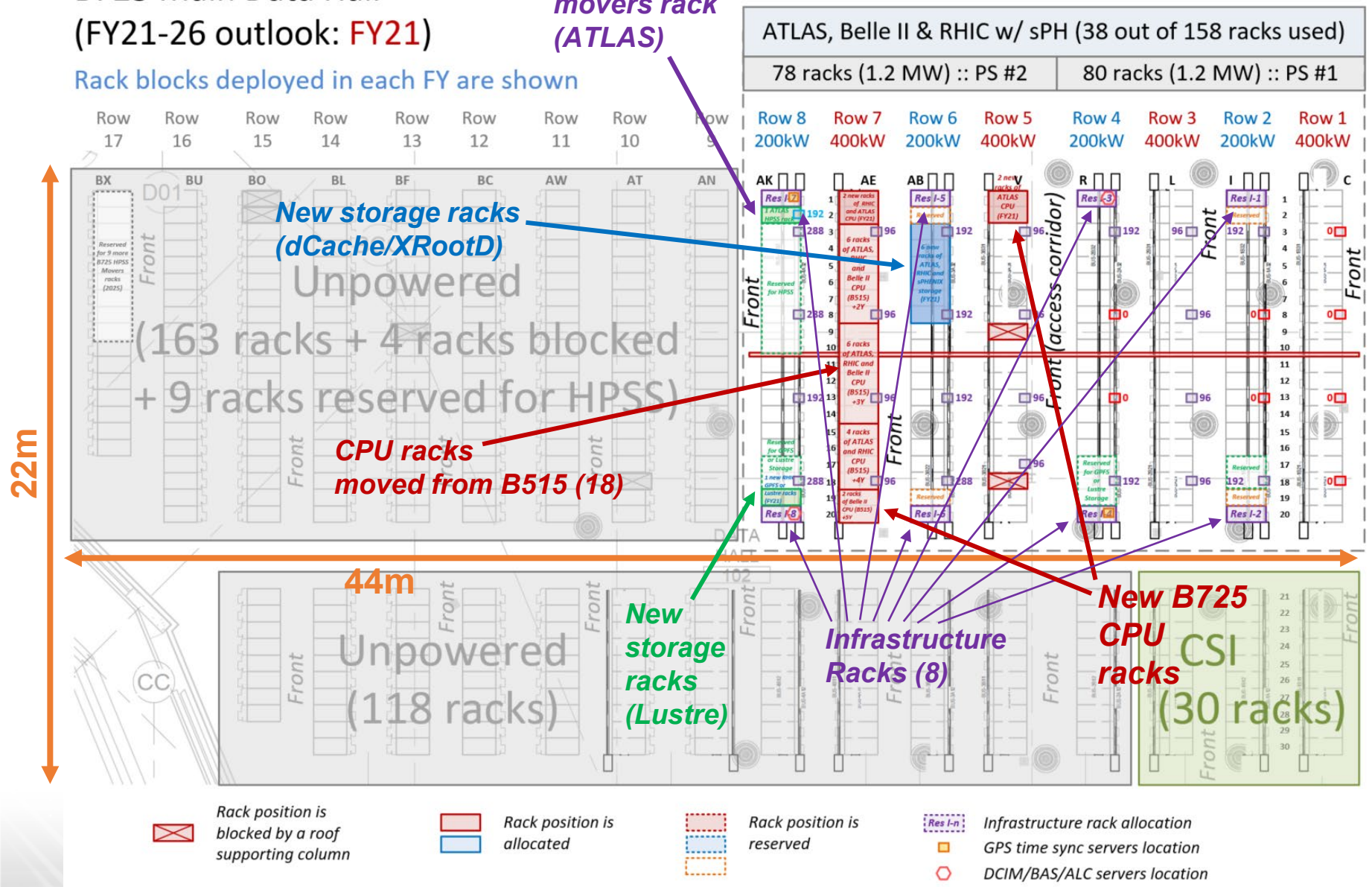
sPHENIX (FY24)



STEP 4
FY24

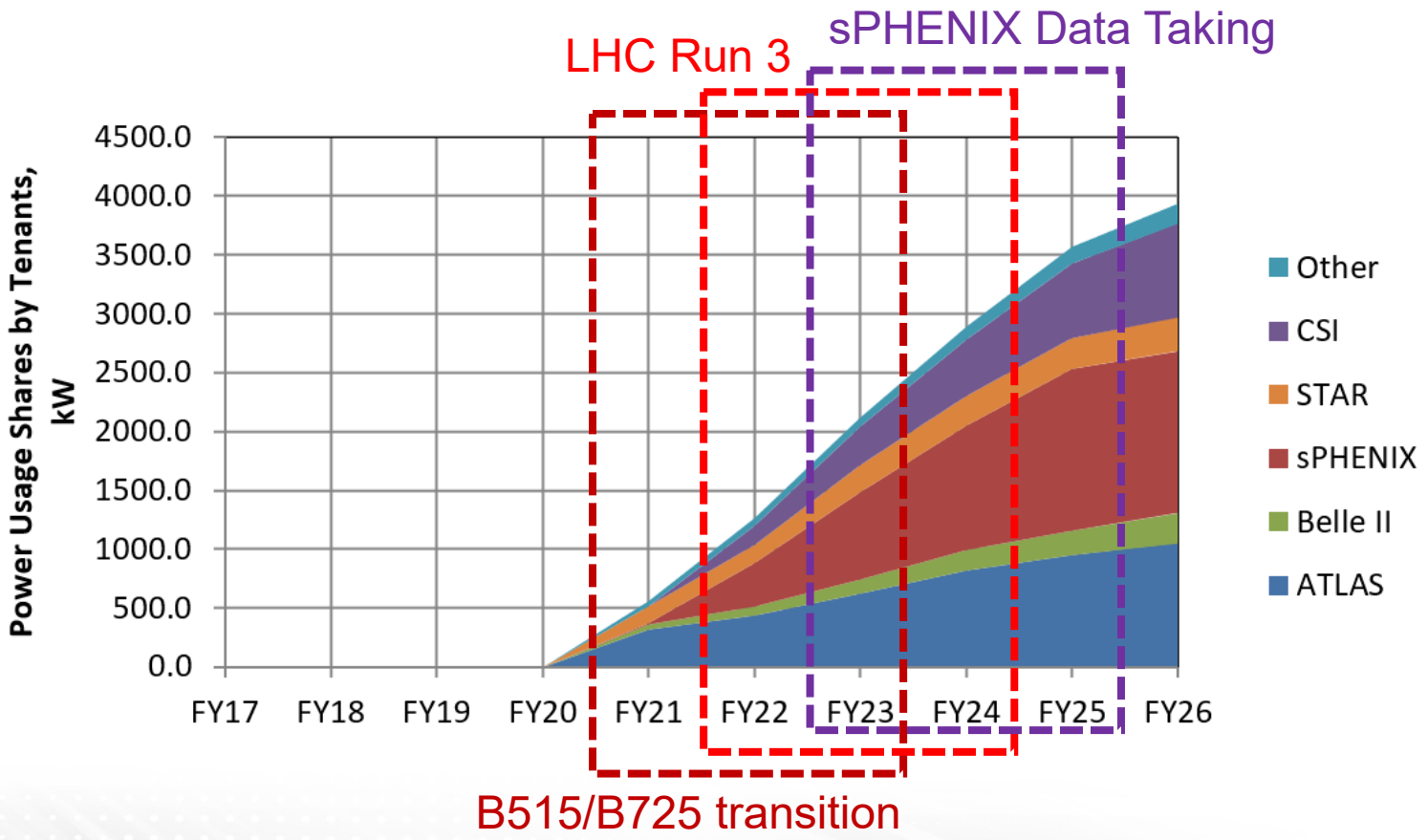
B725 Main Data Hall
(FY21-26 outlook: FY21)

Rack blocks deployed in each FY are shown



Expected B725 Data Center Scale-Up in (US)FY21-26

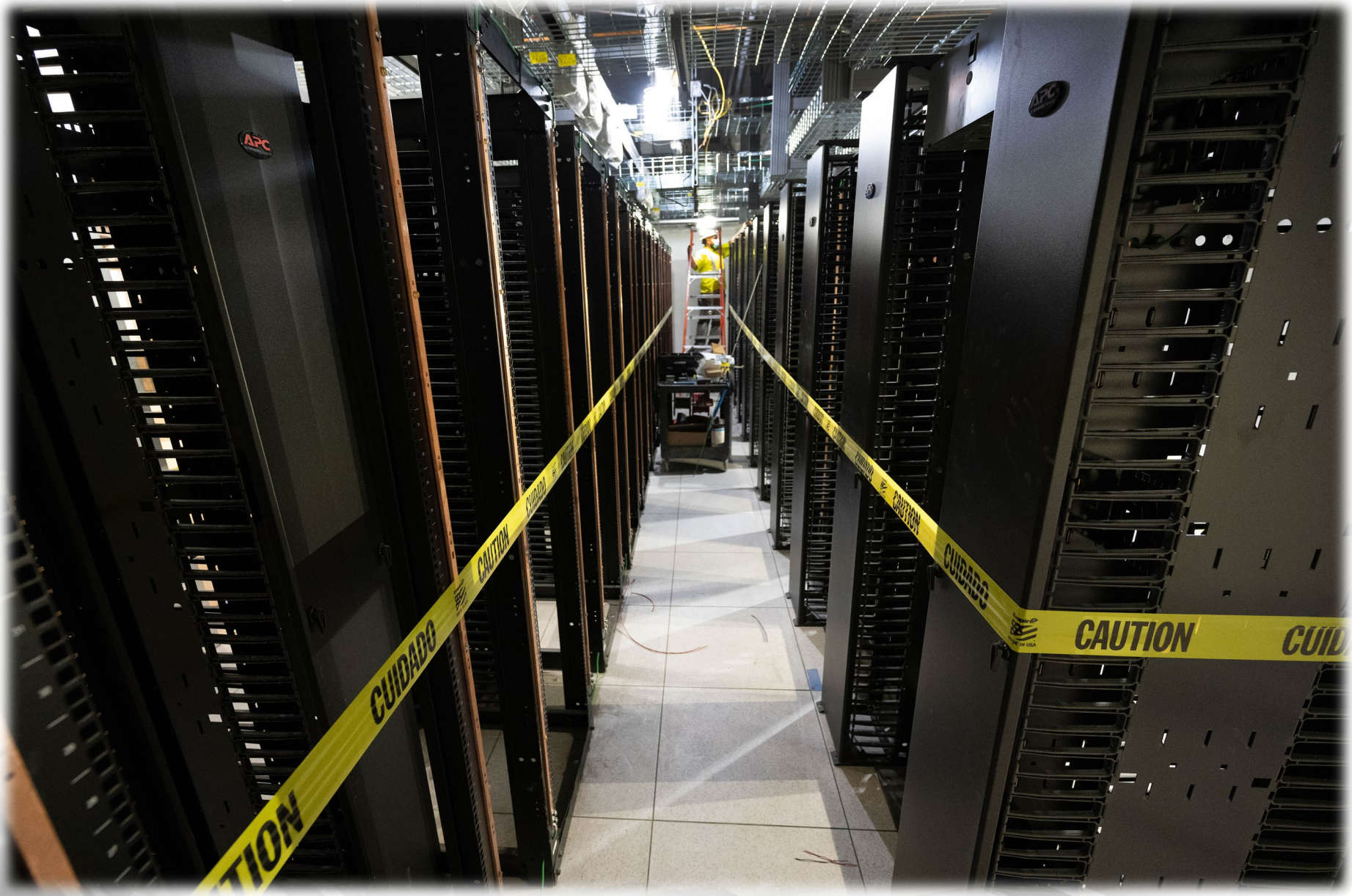
(Max. allocated IT Power w/o PUE component, kW)



Summary

- The CFR project has finished the design phase in the first half of 2019 and then entered the construction phase in the second half of 2019 which is currently projected to be finished in May-June 2021
- The existing B515 data center is already on track for consolidation into the CDCE area (newest built area on B515 side) and Lab C area (currently hosting BNL Network Perimeter) since 2018
- The occupancy of the B725 data center for ATLAS is expected to start in June 2021, and occupancy for all tenants – in July 2021
- A gradual migration of compute and storage capabilities between the B515 and B725 data centers is expected to occur in FY21-23 period with only a subset of existing CPU racks physically moving between two locations in 2021Q3
- B515 data center reduction to the CDCE and Lab C areas is expected to be completed by the end of FY23
- Scaling up of the B725 data center to 4.8 MW of total IT payload capacity is expected in FY25 to address the growing needs of HEP/NP experiments involved (sPHENIX at RHIC and ATLAS at the LHC/HL-LHC)

Questions & Comments



https://www.sdcc.bnl.gov