RHIC Computing Facility (RCF)

Eric Lançon

RHIC S&T Review

September 2019

BROOKHAVEN NATIONAL LABORATORY

U.S. DEPARTMENT OF ENERGY

BROOKHAVEN SCIENCE ASSOCIATES

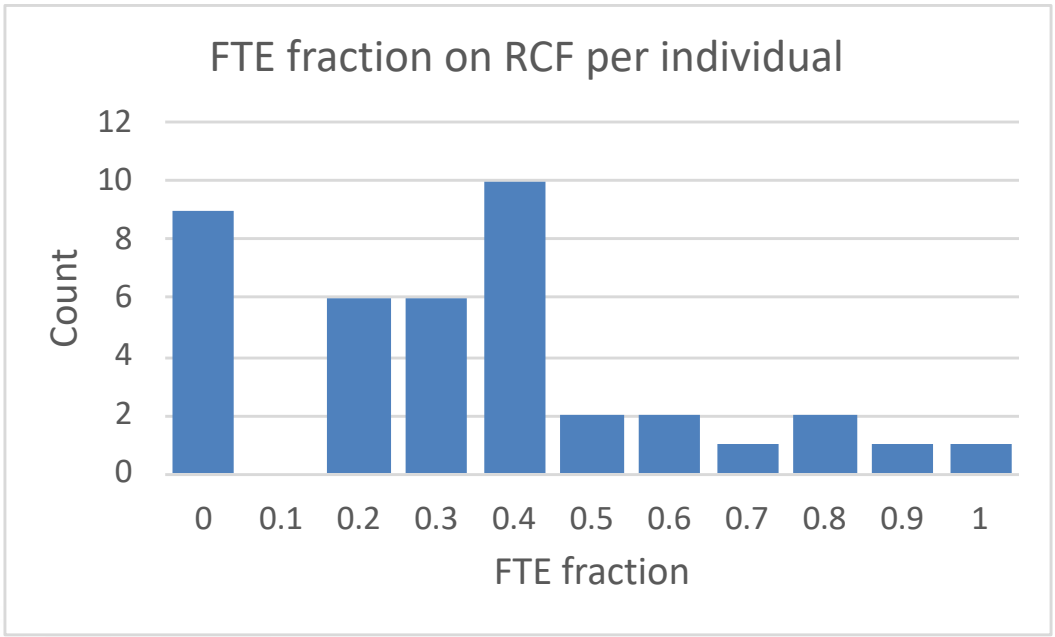
Outline

- **RCF since last review**
  - RCF & CSI
  - Performances
  - Optimization of resources
  - Modernisation
- **New Data Center**
  - Overview
  - Transition
- **sPHENIX**
  - Preliminary resources estimate

RCF & SDCC

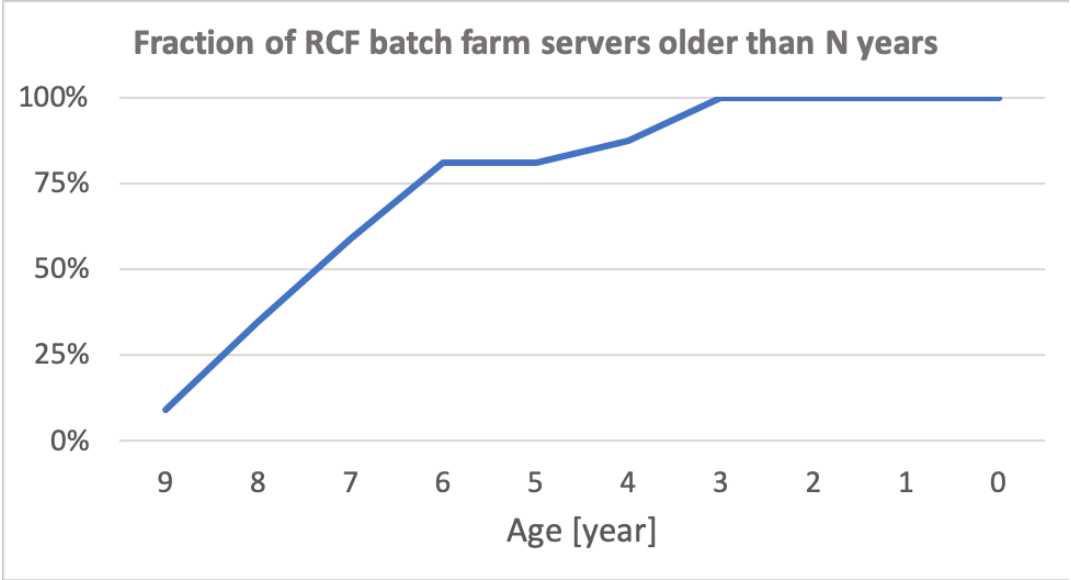
- BNL Scientific Data and Computing Center (**SDCC**) is part of BNL's Computational Scientific Initiative (**CSI**)
- It includes **RCF** from Physics Department
  - 13.6 FTE
  - Provides resources and support NP projects
- SDCC also supports
  - HEP : ATLAS, Belle II, LSST, Dune,…
  - Lab & outside : CFN, ARM, Simons Foundation, NNDC, LQCD, NSLS II,…
  - 39 individuals (31 previous S&T review)
- Sharing of resources, procedures, expertises etc…
- **Operation oriented with limited effort for R&D and development**

39 individuals - working in average 35% for RCF



Since last S&T: modernization & optimization

- CPU & storage capacity kept at constant level
  - Aging servers on the batch farm: 50% of older than 7 years
  - Budget did not allow for 4/5 years renewal cycle
- While maintaining & improving infrastructure leveraging from synergies with HEP and CSI projects



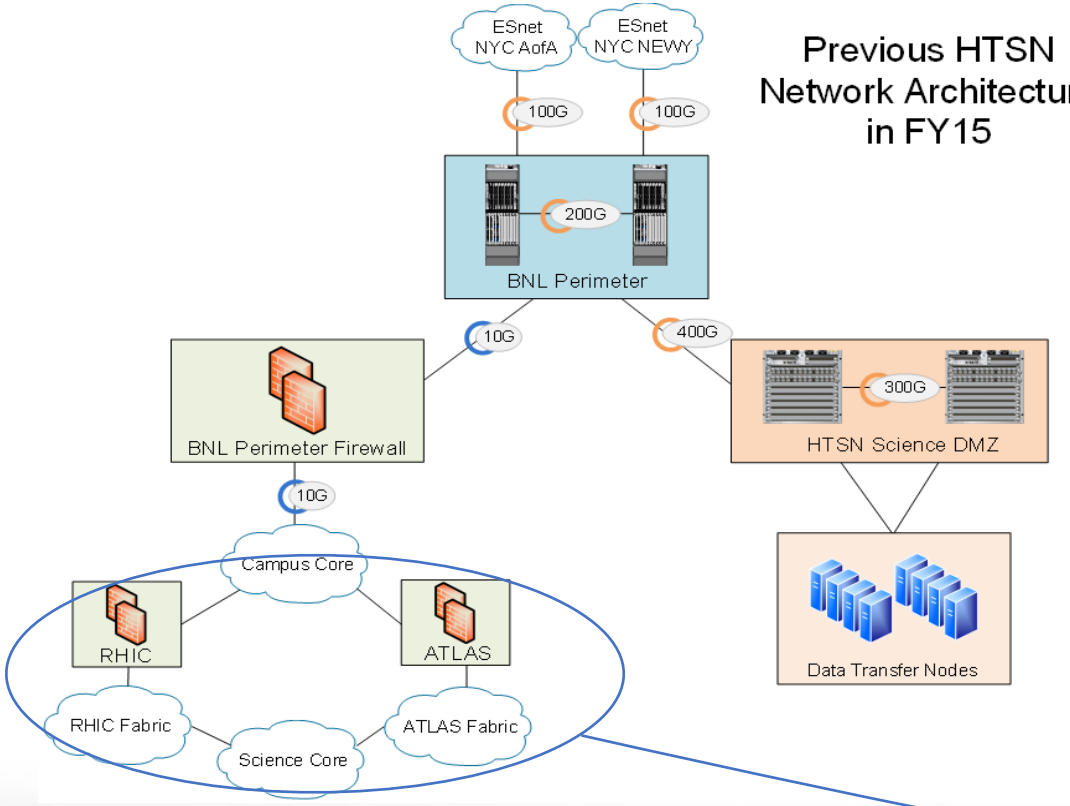
Since last S&T: modernization & optimization

- **Major infrastructure modernizations**
  - New Identity and Access management system: a single account for all services (first step towards Federated Identity)
  - Scientific computing network isolated from BNL internal perimeter (Science DMZ, high performance network for science applications)
- **Optimization of resources**
  - Shared computing batch pool
  - Unified storage administration
- **New ‘services & tools’ group**
  - Upgrade to modern collaborative tools
  - Anticipation of future projects needs
- **While providing high level services for STAR and PHENIX**

Network re-design

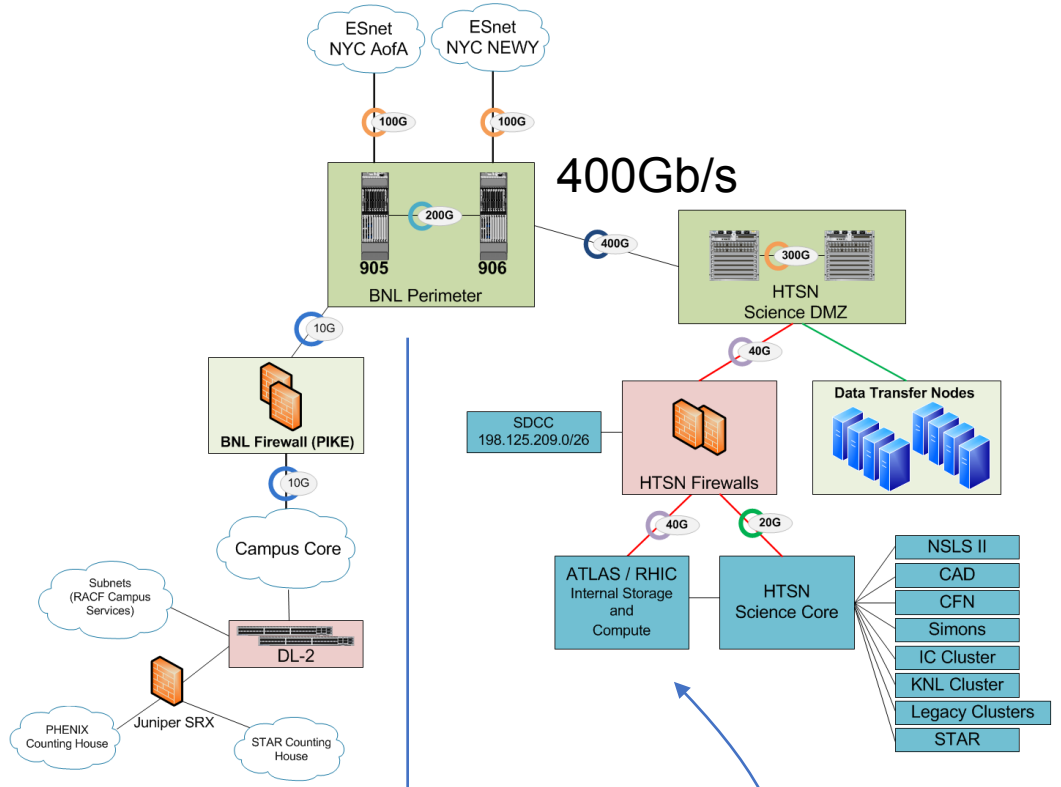
Full implementation of the DMZ concept

Before

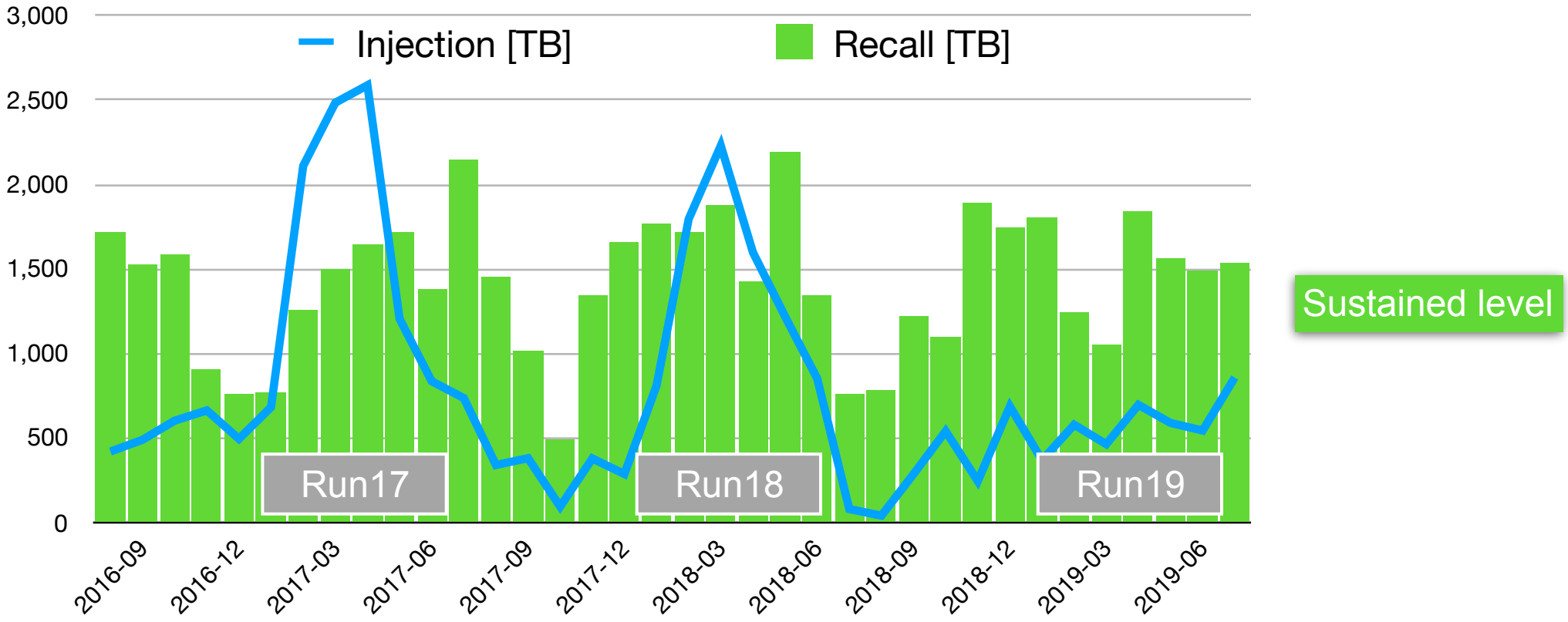


Previous HTSN Network Architecture in FY15

After

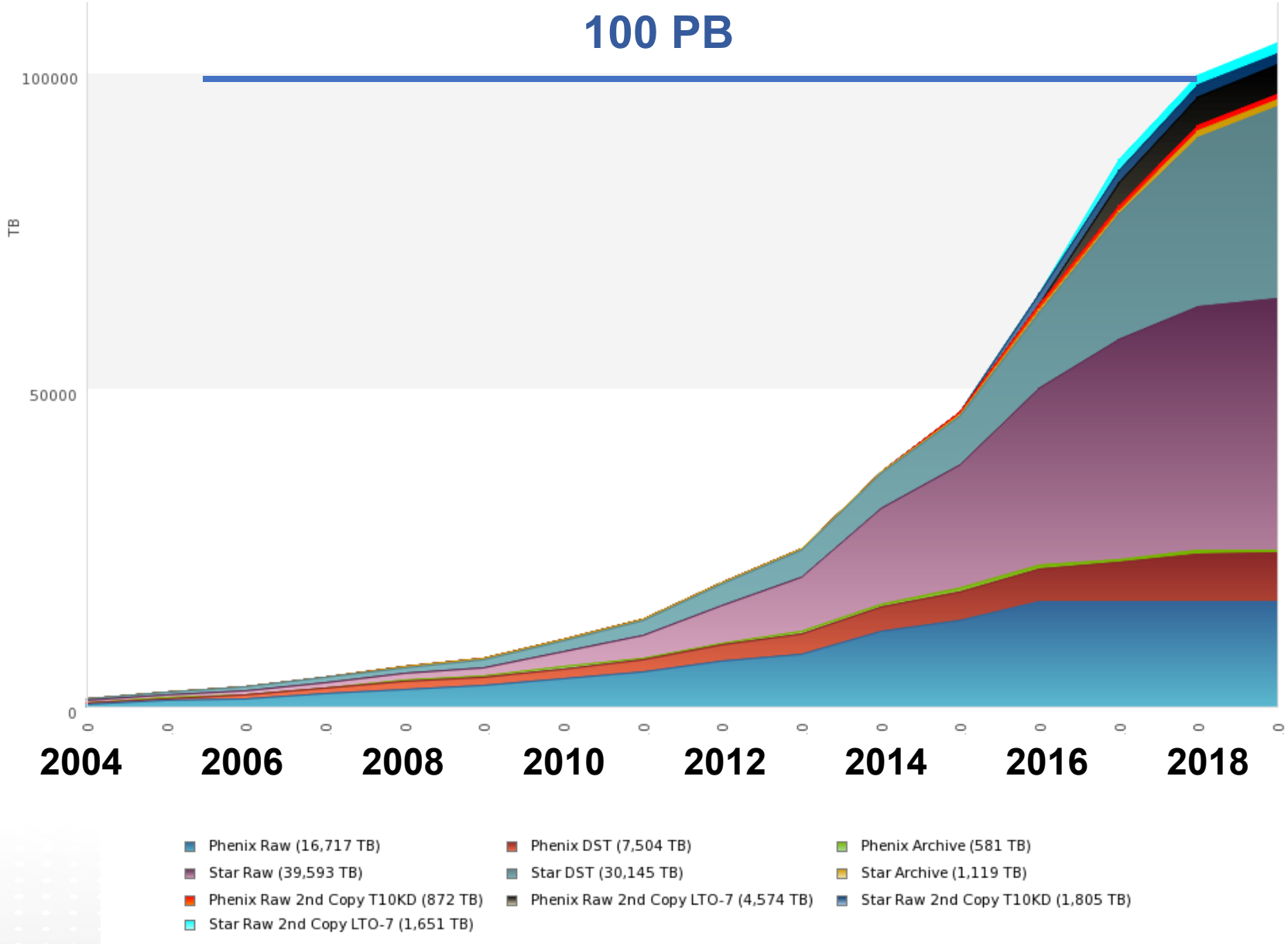


Usage of tape system by STAR



RHIC reached 100PB of data on tape in 2019

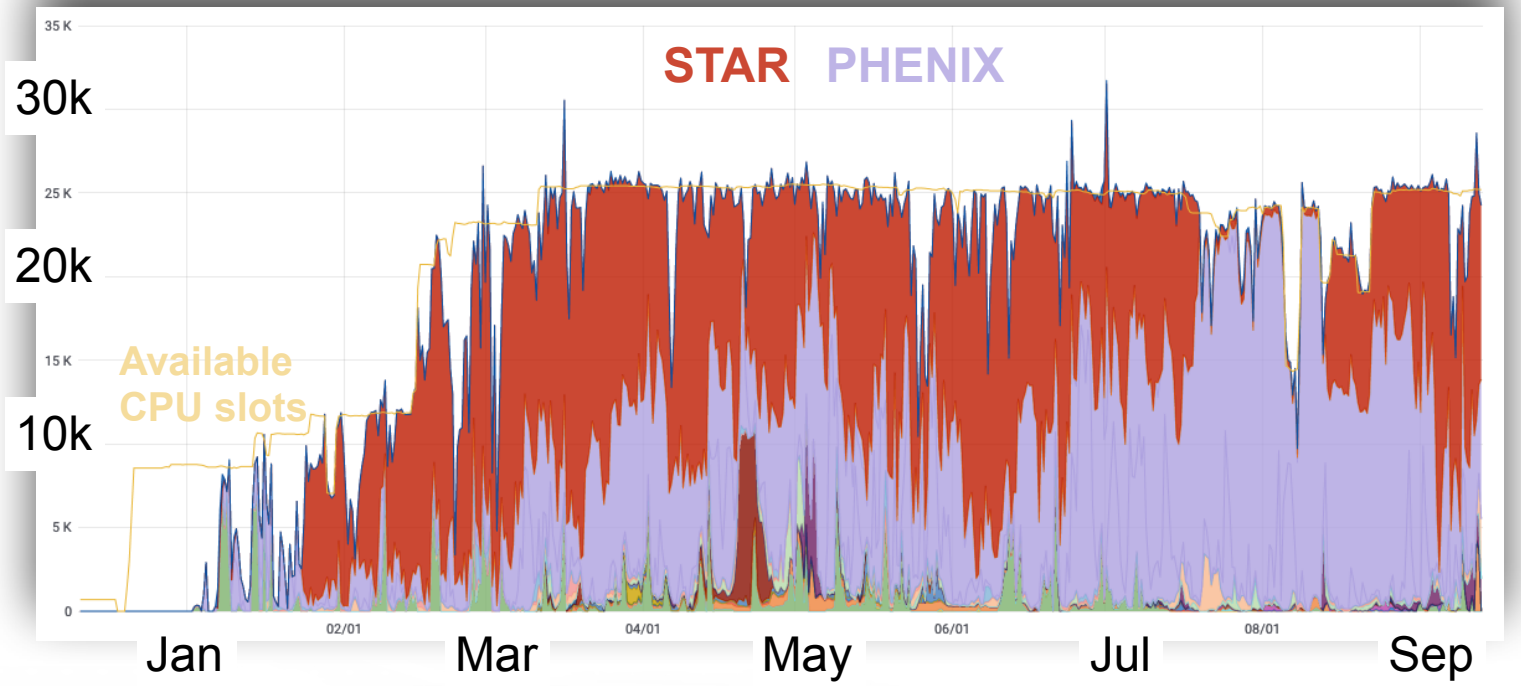
Two copies of RAW data made when tapes are ‘repacked’ to newest media generation



A common batch farm — The shared pool

- Optimisation of computing ressources
  - A single batch farm for STAR & PHENIX (2 farms before) and other projects
  - Better utilisation of available CPU cycles
- Standardisation of hardware configurations
  - No permanent storage on CPU nodes
- All of PHENIX & STAR recent equipment moved to shared pool
  - PHENIX moved almost all of their CPU to the shared pool (for the benefit of STAR)

Usage of the shared pool



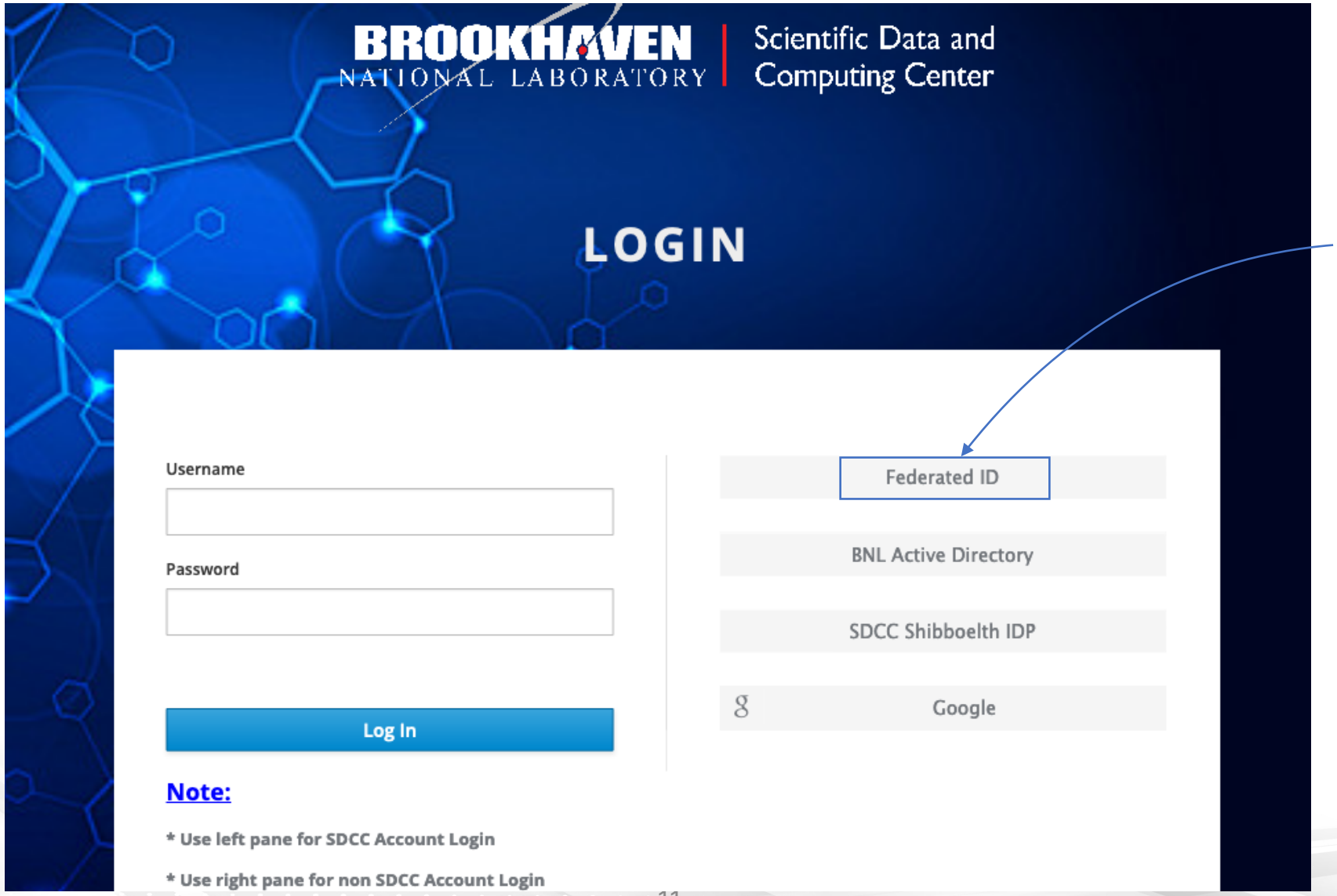
Support for collaborations

- **Modernization of collaborative tools and services for present and future RCF user communities**
  - Ticketing, project tracking
  - Repositories
  - Software development platform
  - Documentation
  - JupyterHub
  - BNLBox (cloud storage)
  - Documents publication

- **Active contributor to many community solutions**
  - Indico (with FNAL) for event handling
  - Invenio (with CERN) for document handling

- **Ongoing deployment of Federated Identity to access resource** (first DOE Lab?)

Access portal (prototype)

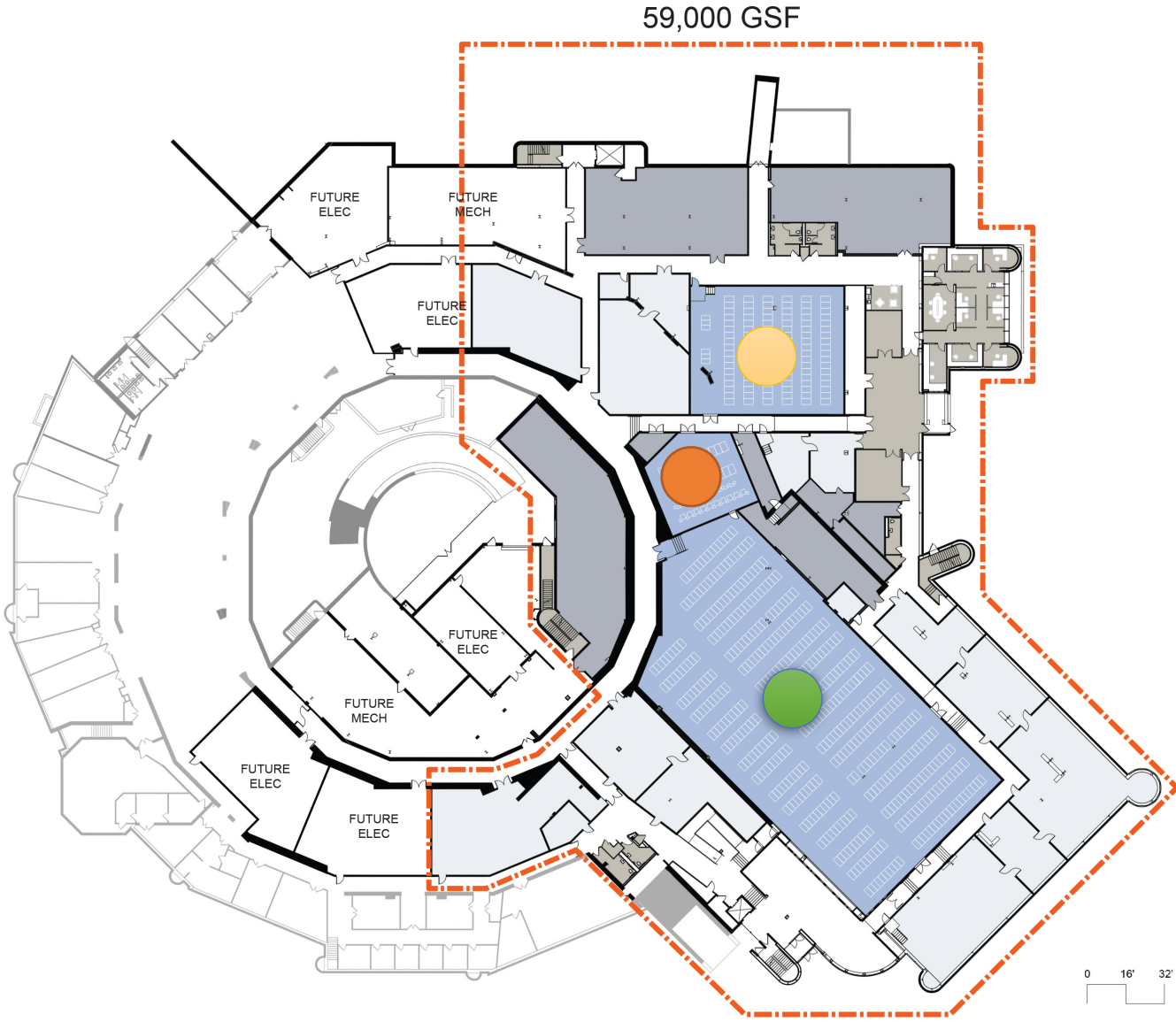


Federated ID account:
Lab/University
EDUGain
(CERN)
…

New Data Center

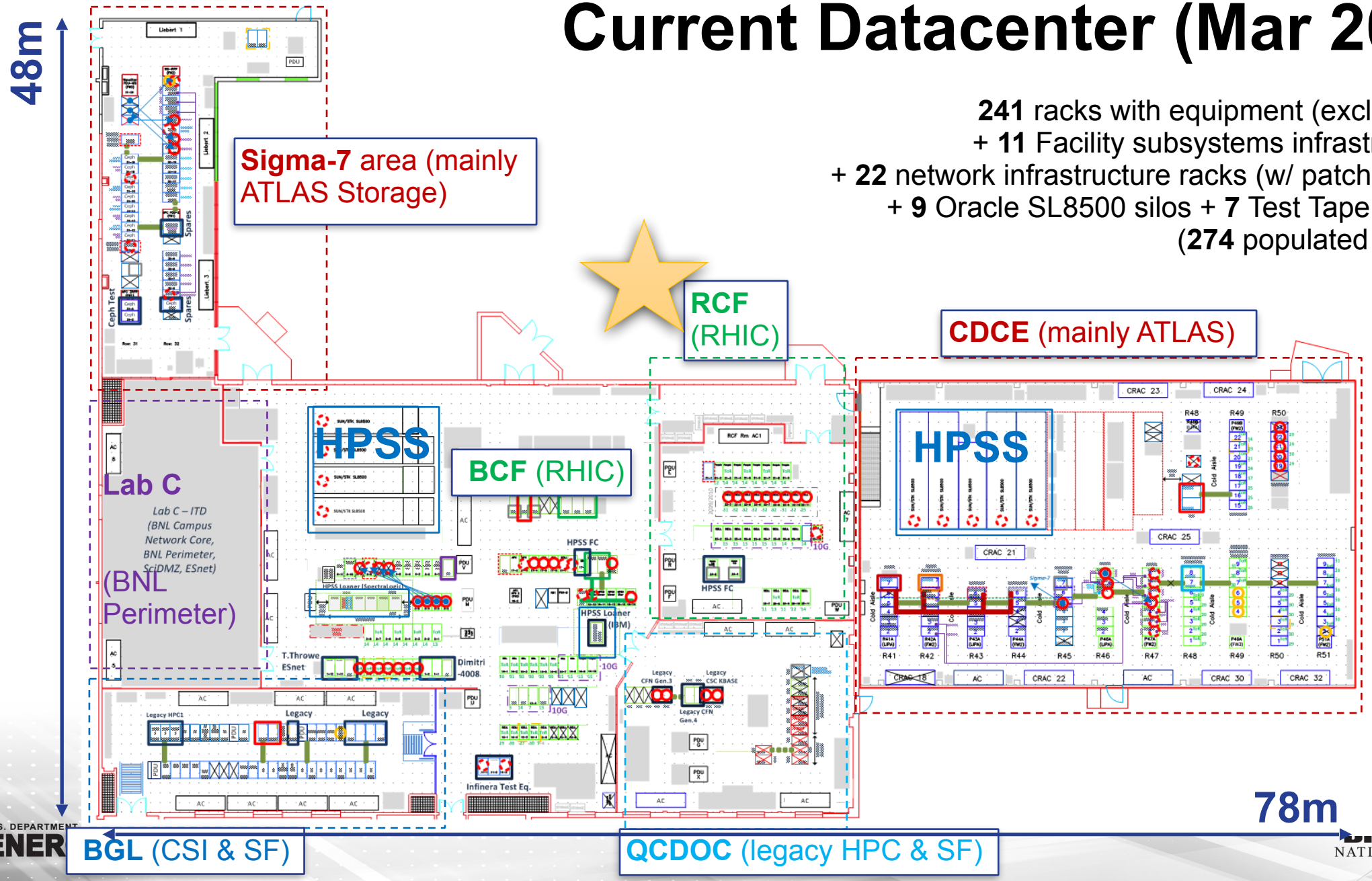
NSLS-I Building

- Achieved CD-3 May 30,2019
  (CD-3B Jan. 2019)
- Network available Fall 2020
- Tape room early 2021
- Main hall March 2021
- CD-4 end FY23



Current Datacenter (Mar 2019)

241 racks with equipment (excluding Lab C)
+ **11** Facility subsystems infrastructure racks
+ **22** network infrastructure racks (w/ patch panel racks)
+ **9** Oracle SL8500 silos + **7** Test Tape Library Silos
(**274** populated rack frames)



Musical chairs in the current data center

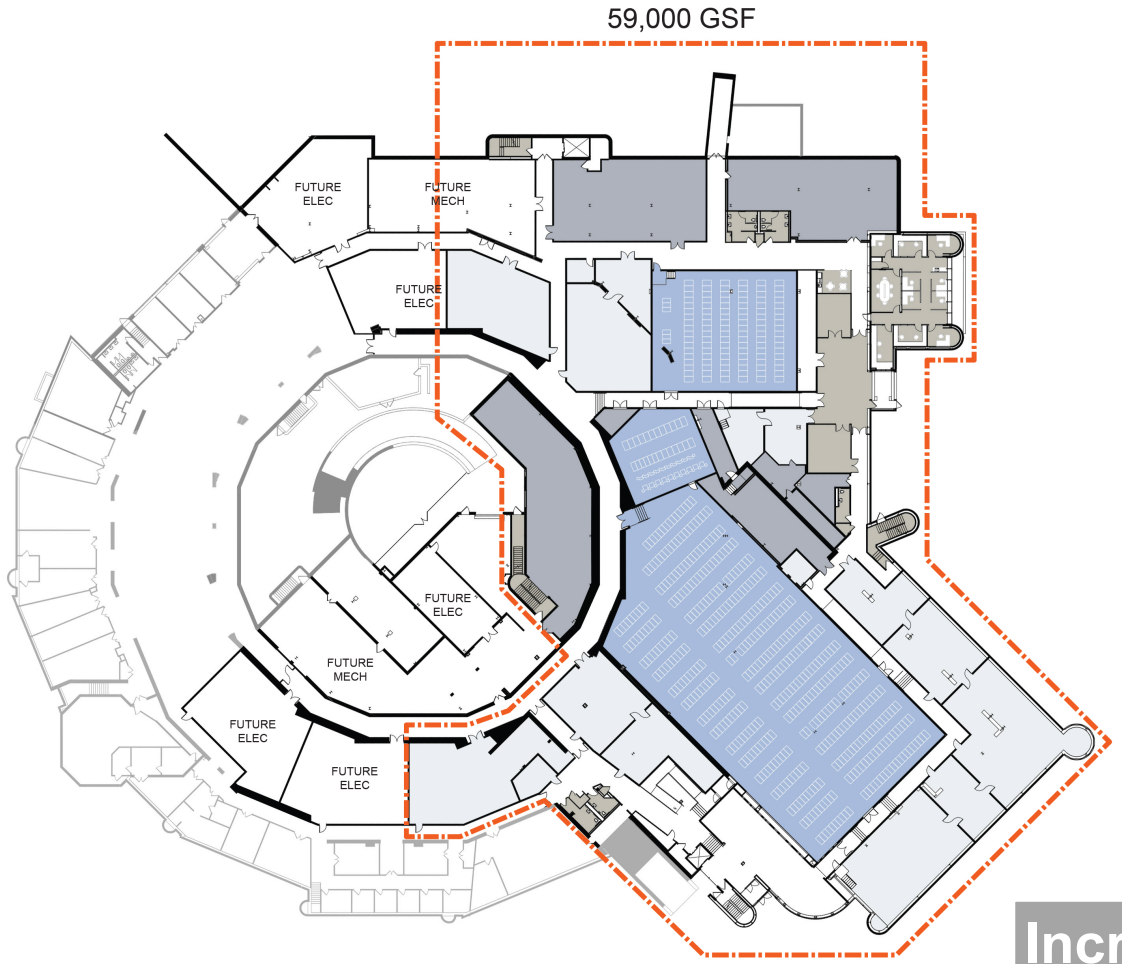
Retired equipment
Disks being sanitized

New equipment
Delivered in the main hall

Electrical work to comply with safety regulation

Scope of Work

Key Performance Parameters (KPPs):



Description of Scope	Threshold Value	Objective Value
Deliver identified Computing Facility IT power and emergency back-up power/ cooling capabilities	3.6 MW IT power, 1.2 MW emergency back-up capabilities	4.8 MW IT power, 3.6 MW emergency back-up capabilities
Renovate identified Gross Square Footage (GSF)	45,000 GSF	85,000 GSF

- **IT Power – 3.6MW**
  - **Provision for** concurrent maintenance and **Incremental growth** to facilitate support of BNLs scientific computing mission (Bypass System)
  - Highly redundant systems (Tier III, “N+1”)

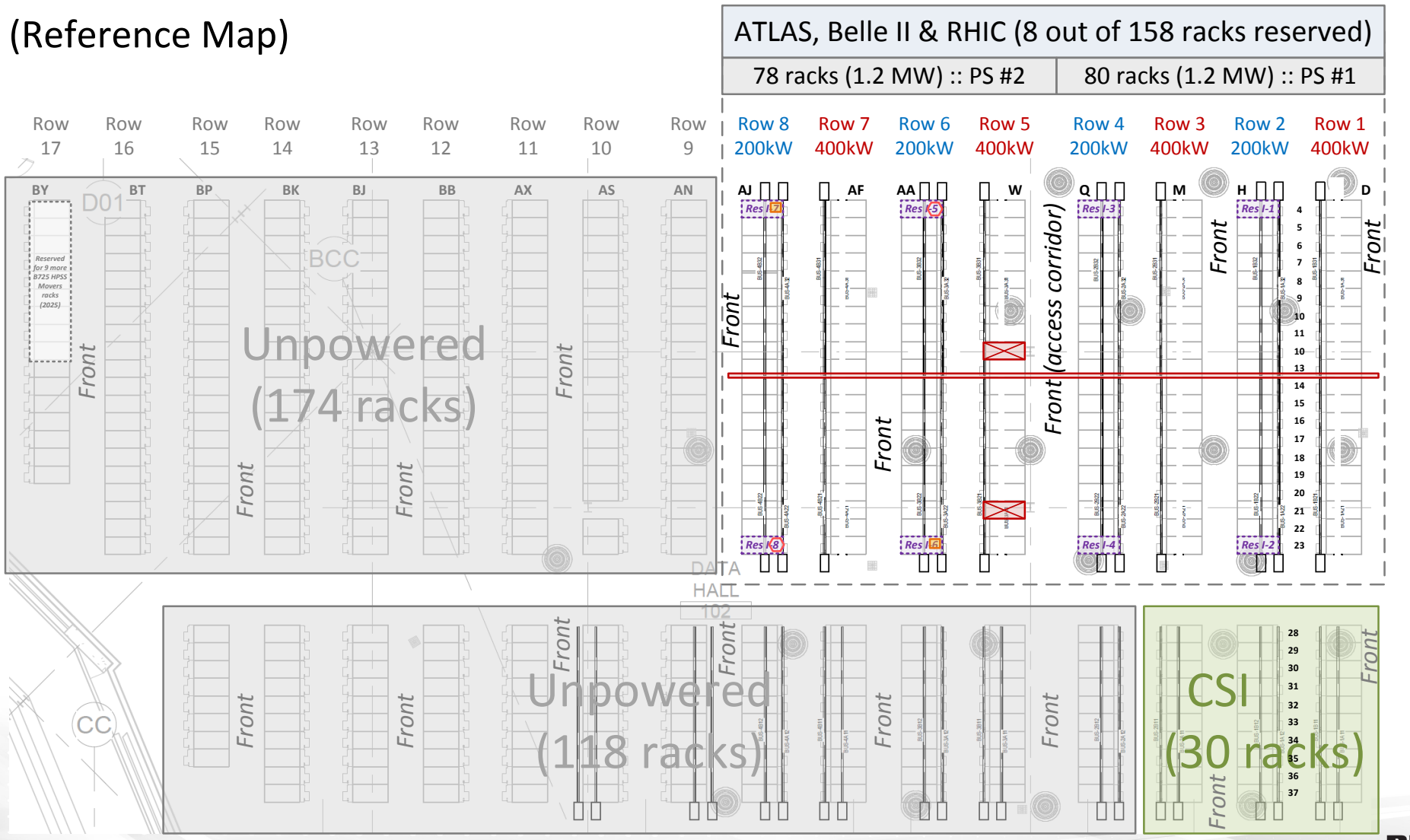
Increment by 1.2MW (~$6M)
Maximum capability ~10MW

Main data hall floor plan

For KPP threshold value (3.6 MW)

B725 Main Data Hall

(Reference Map)



Rack position is blocked by the column

Rack position is allocated

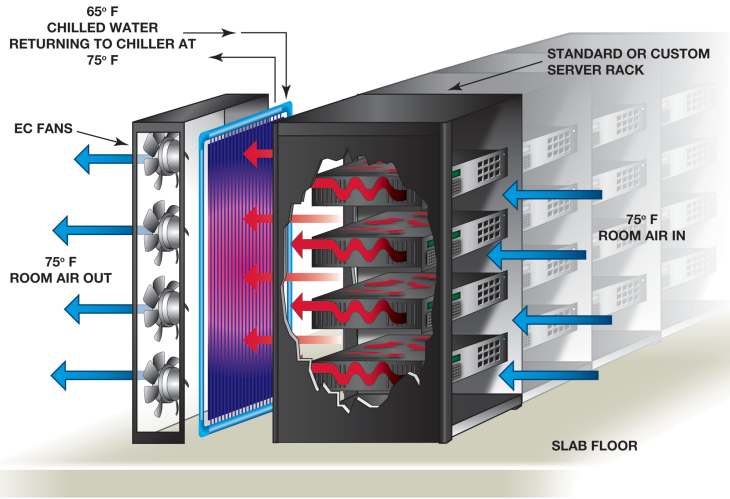
Rack position is reserved

Infrastructure rack allocation

GPS time sync servers location

High efficient Data Center

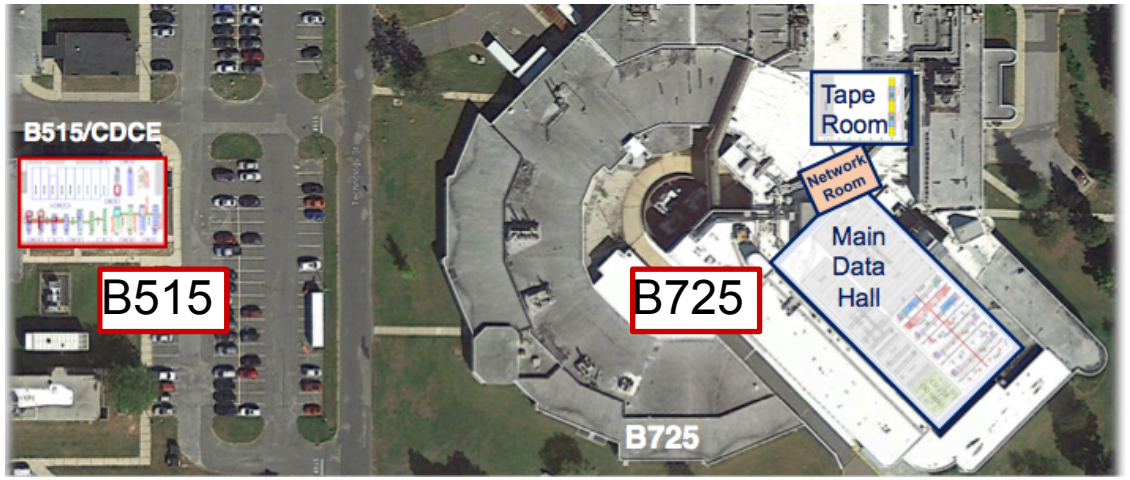
- **PUE** (Power Usage Effectiveness):
  - Less than **1.4** required (DCOI: Data Center Optimization Initiative)
  - Targeting **1.2 – 1.3** (Presently 2 ”+” in RCF)

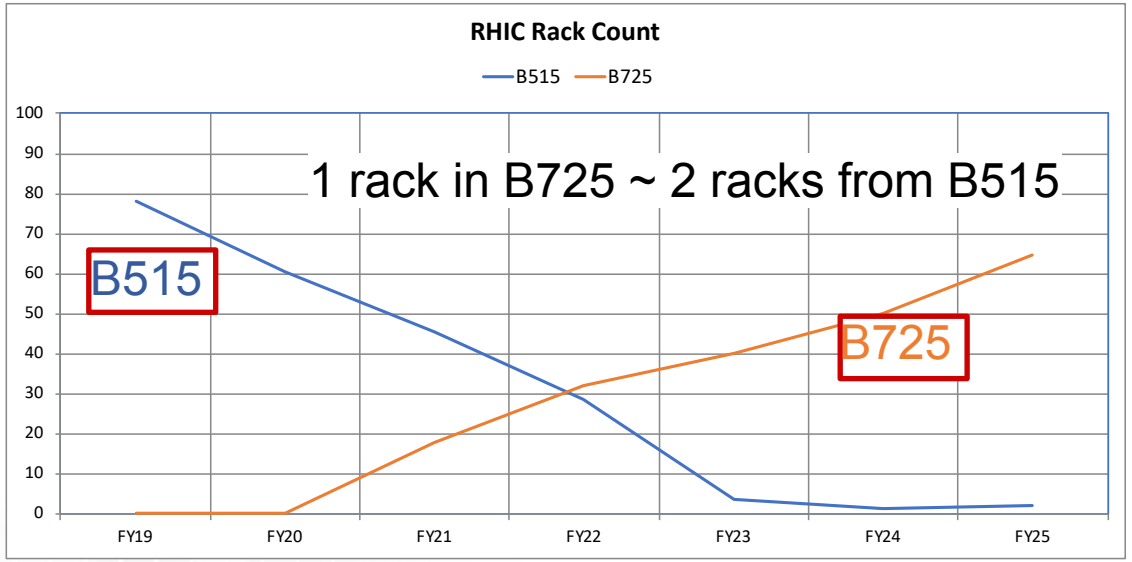


EXPLODED VIEW OF ACTIVE REAR DOOR
RACK COOLING SYSTEM

~~Migration~~ Transition

- Transition to new data center starts early FY21(beg. of LHC Run3)
- Equipement purchased in FY20+ installed directly in B725
- **No service interruption**
- 2021-2023: Operation between the 2 data centers
- Relocation of equipments;
  - CPU (only) move to B725
  - Storage stay in B515 until end-of-life
  - Existing tape libraries stay in B515
- Details in recent HEPiX presentation





sPHENIX

- Software & Computing review Sep. 5-6, 2019
  - Updated estimates of resource needs for data (& simulation) storage and processing
  - With breakdown by resource type
- **Iterating** with sPHENIX collaboration on needs and configurations

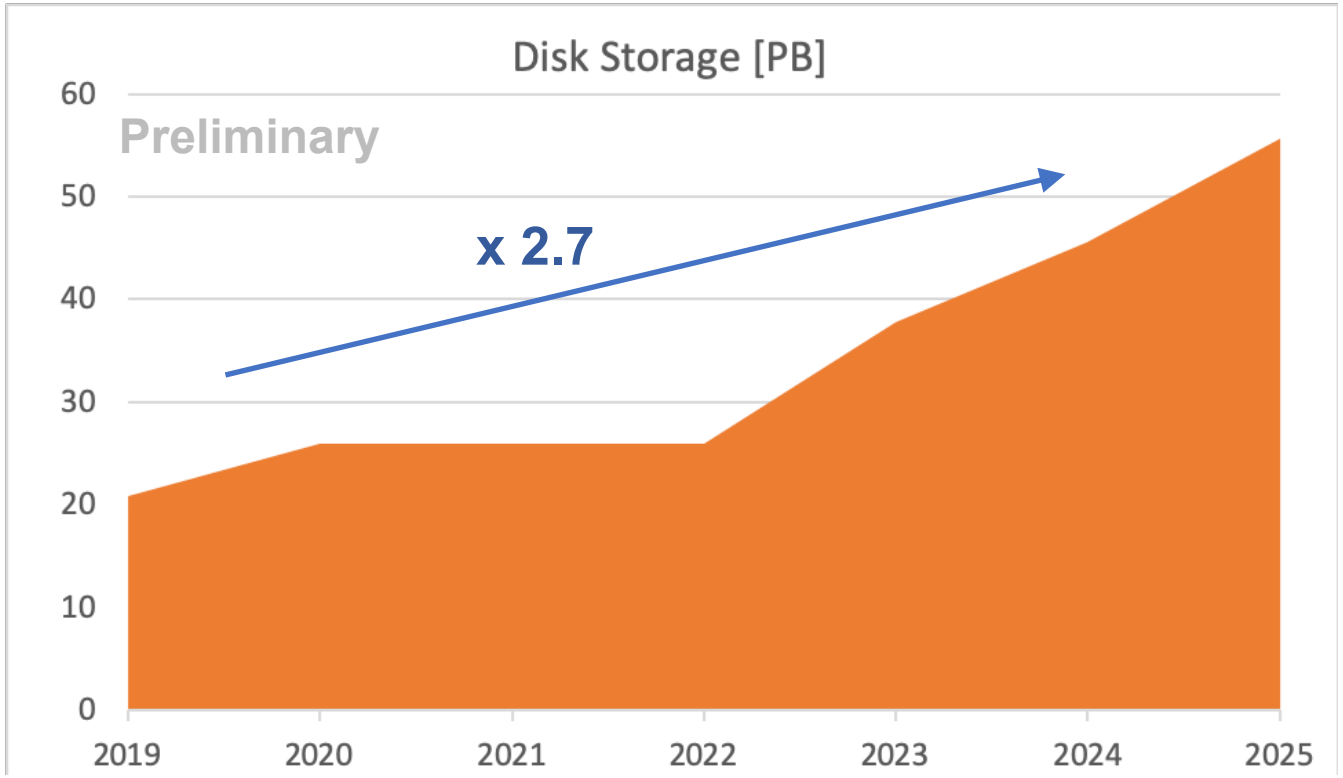
Data processing

Storage Needs	year-1	year-2	year-3
raw data (PB)	80	144	192
Disk (PB)	17	30	40
Tape (PB)	132	238	317
Cumulative Tape	132	369	686

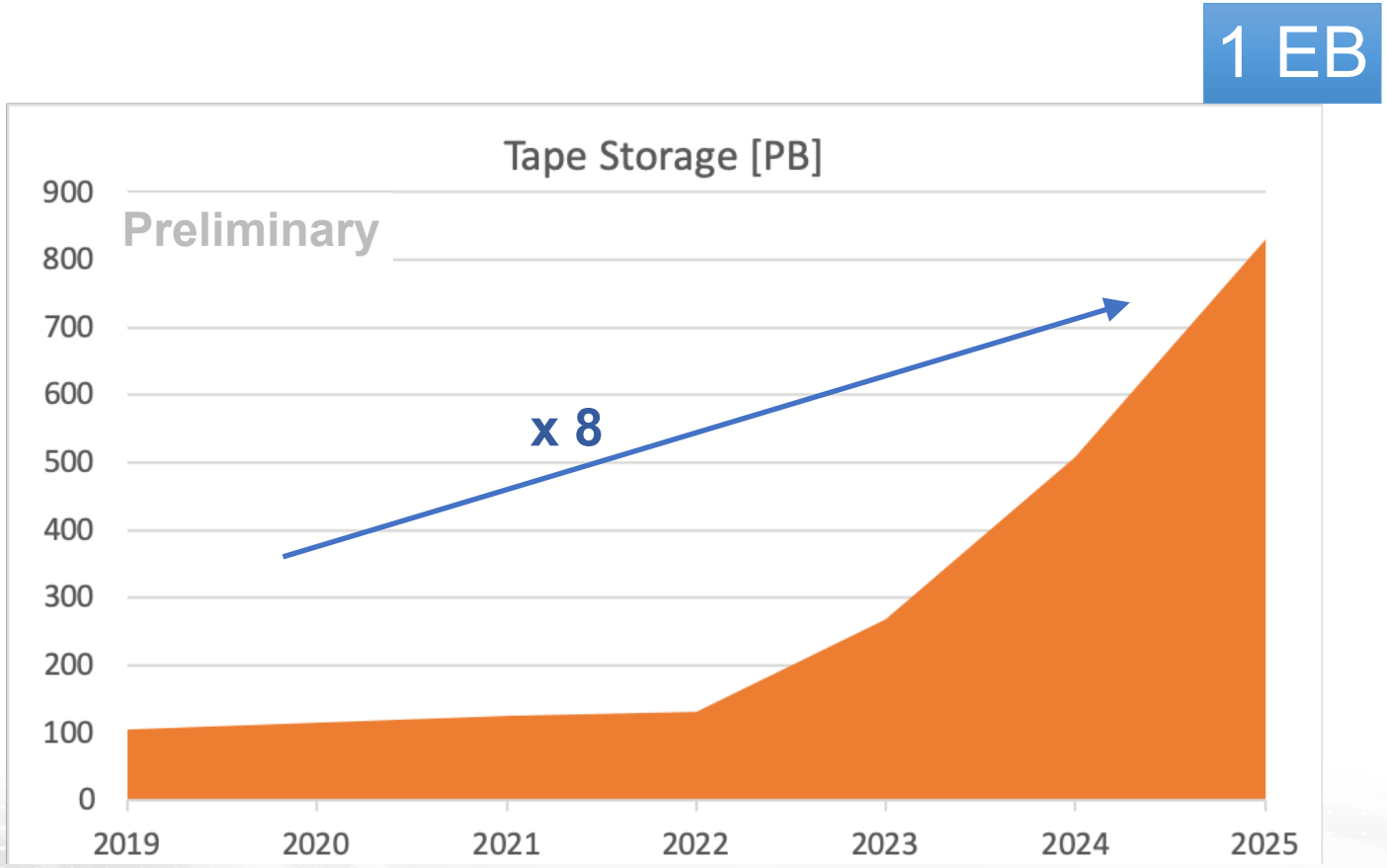
Simulation

MC Project/yr	2020	2021	2022	2023	2024	2025
Annual (PB)	5.2	5.2	1.2	3.2	2.0	3.2
Cumulative (PB)	5.2	10.4	11.6	14.8	16.8	20.0
Total CPU (10w)	60k	52k	2k	19k	41k	19k

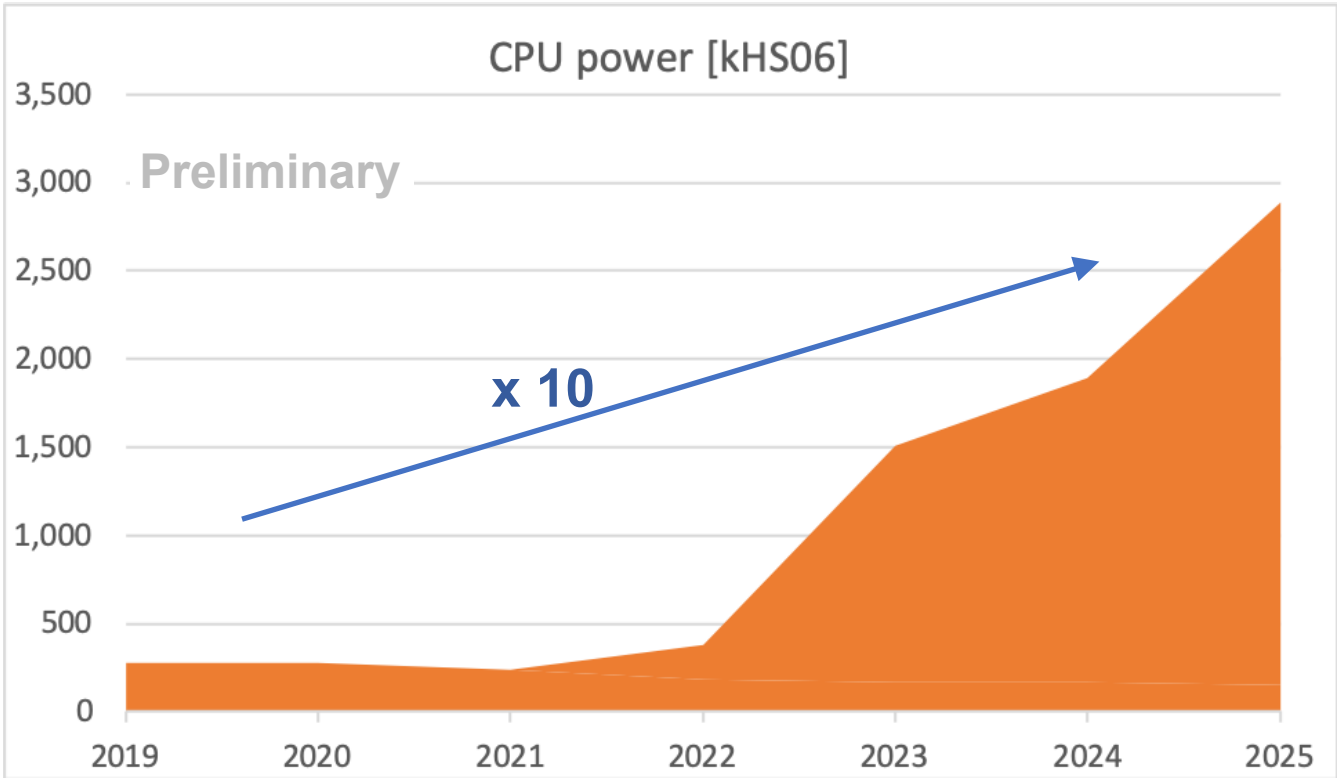
Disk storage evolution



Tape storage evolution



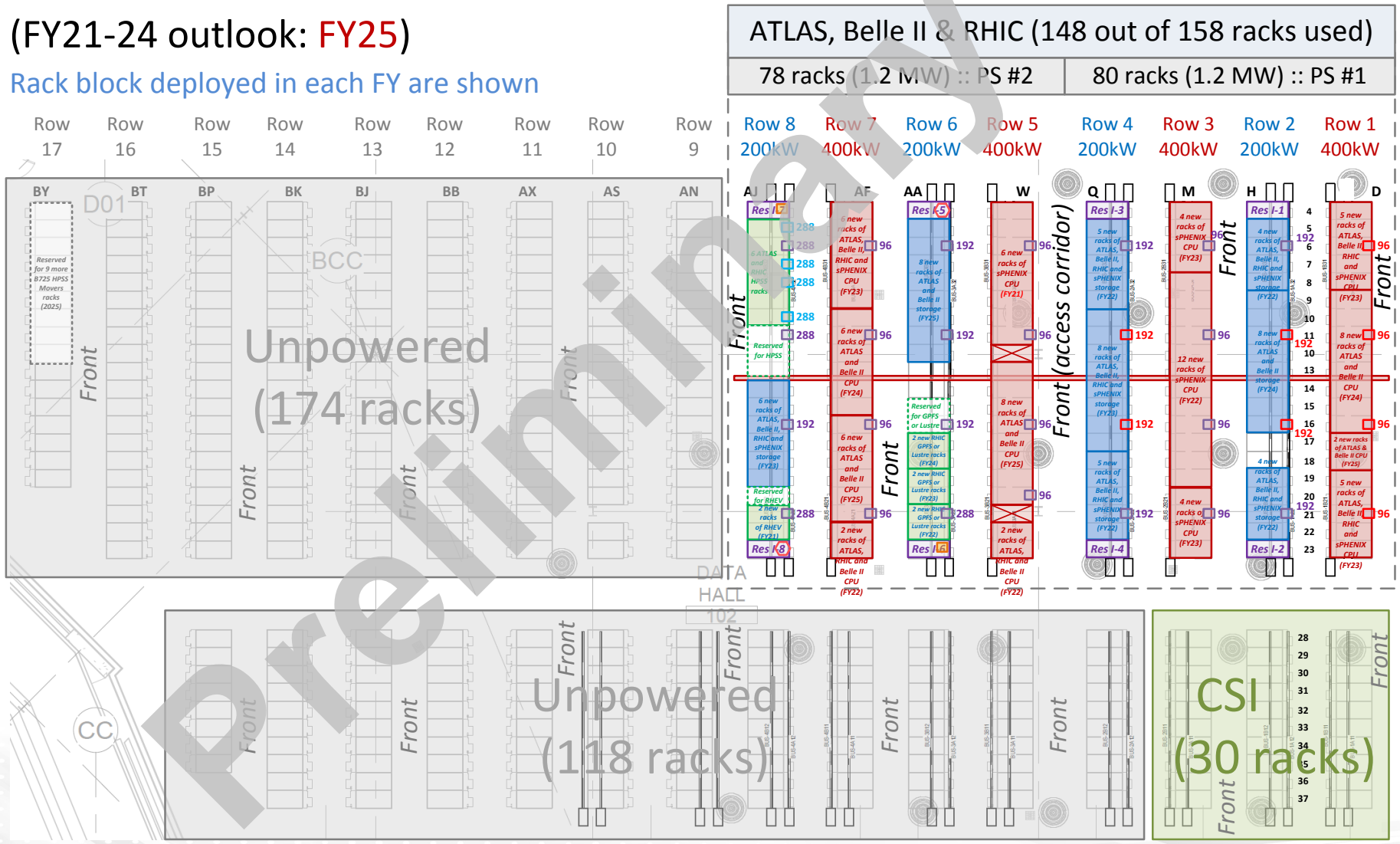
CPU evolution



Data Center Hall may get full depending on the hardware configurations (and needs of other programs)

B725 Main Data Hall
(FY21-24 outlook: FY25)

Rack block deployed in each FY are shown



Preliminary

Rack position is blocked by the column

Rack position is allocated

Rack position is reserved

Infrastructure rack allocation

GPS time sync servers location

DCIM/BAS/ALC servers location

Fig. A6

Summary & Outlook

- Excellent performance of RCF during recent RHIC runs
- Effort in modernization and deployment of collaborative tools for present and future RHIC projects
- Transition plan to new data center without interruption of service
- Iterating with sPHENIX to refine needs

Thank you

CFR – Schedule