BNL Site Report

Zhihua Dong (Presenter. site manager)

Chulwoo Jung (site architect)

Costin Caramarcu (site architect)

Apr 30, 2020

BROOKHAVEN NATIONAL LABORATORY

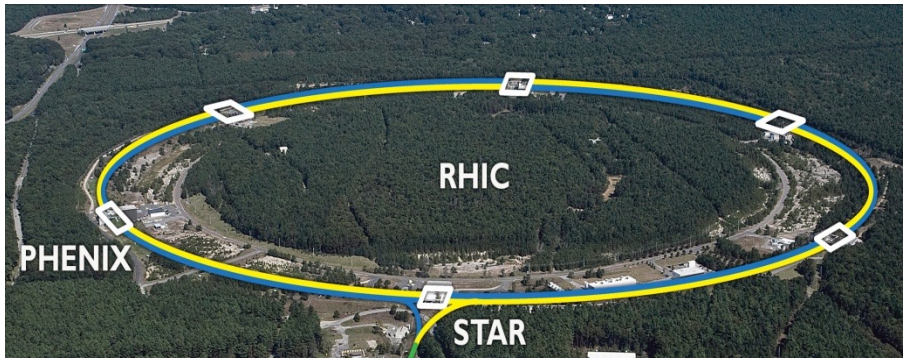
U.S. DEPARTMENT OF ENERGY

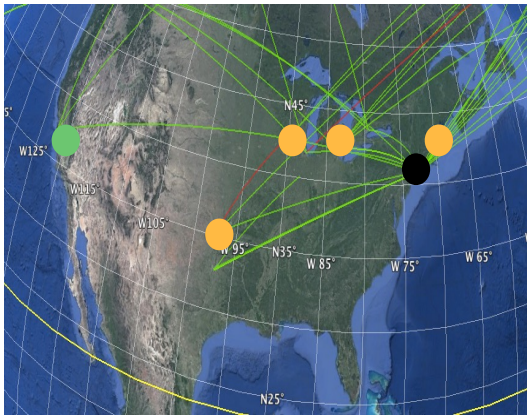
Scientific Data and Computing Center (SDCC)

- Support for various programs:
  - RHIC, LHC ATLAS, BER ARM, LSST, DUNE, LQCD,RIKEN, BES, Center for Functional Nanomaterials(CFN), National Synchrotron Light Source(NSLS) II, National Nuclear Data Center, Simons Foundation,…
- ~2000 users from >20 projects
- Staff
  - 38 full-time regular members
  - 1 current openings

SDCC support for HEP experiments

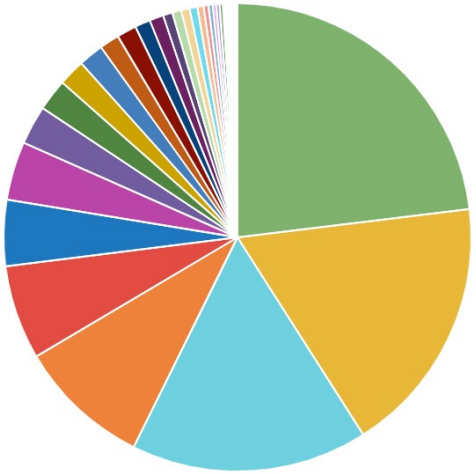
- **The RHIC Tier 0 (21 years)**
  - Store and process data from RHIC experiments
  - Provide analysis means for ~1200 users
  - Long term data preservation
  - Simulation resources for future programs (sPHENIX & EIC)
- **The US ATLAS Tier 1 & Tier 3**
  - ~25% of ATLAS Tier 1 computing capacity worldwide
  - Store RAW data from LHC and from simulation
  - Distribute data to the 4 US Tier 2 sites + analysis site (SLAC)
  - **Analysis center for US physicists**
    - From 41 institutions (incl. 4 Nat. Labs)
    - 600 physicists, 190 PhDs
- **A Belle II data center outside Japan**
  - **Initial operations began on Oct. 2017**
  - **Data taking began in Fall 2018**





24/7 availability

Covid-19 Researches at BNL



	total
Brookhaven National Laboratory	30.85 Mil
Fermi National Accelerator Laboratory	23.93 Mil
Syracuse University	21.90 Mil
University of Chicago	12.28 Mil
Massachusetts Institute of Technology	8.72 Mil
University of Nebraska	6.10 Mil
University of Michigan	5.41 Mil
University of Texas Arlington	3.59 Mil
American Museum of Natural History	2.94 Mil
University of California San Diego	2.46 Mil
University of Wisconsin	2.32 Mil
Vanderbilt University	1.84 Mil
Rutgers University	1.80 Mil

Computational project details

- PDBQT Docking Study
- Molecules Docking Study
- Search and Mining System for COVID-19-related Literature via Natural Language Processing
- Natural Language Processing: Neural-network-based language model
- Natural Language Processing: Keyword searching
- Neural Fingerprint Method for Chemical Compound Characterization
- Simplified Molecular-Input Line-Entry System (SMILES) Searching
- Drug and Vaccine AI/ML Toolkit
- Neural Fingerprint Method
- ExaLearn Exascale Computing Project COVID-19 Response
- KBase (Predictive Biology)

(https://gracc.opensciencegrid.org/dashboard/db/covid-19-research?orgId=1)

- Most of current BNL SDCC computing resource usage for covid-19 from Open Science Grid
- Other BNL participation in covid-19 projects can be found in https://www.bnl.gov/science/COVID-working-group.php#projects

SDCC resource for High Throughput Computing

Providing our users with ~2,000 HTC nodes:
~80,000 logical cores
~890 kHS06

73 new Supermicro SYS-6019U-TR4 1U servers brought online in July 2020
- Dual Intel Xeon Cascade Lake 6252 CPUs @ 2.4 GHz (96 log. cores total)
- 12 x 16 GB (192 GB total) DDR4-2933 MHz RAM
- 4 x 1.8 TB SSDs
- 1U form factor
- 1140 HS06/node = ~83 kHS06 total

All nodes running Scientific Linux 7 for some time
- SL6 Singularity containers provided to experiments which still require this OS

New Cascade Lake-based Supermicro 6019U-TR4 Servers

SDCC support High Performance Computing

Currently supporting 5 HPC clusters

Institutional Cluster (IC)
216 HP XL190r Gen9 nodes with EDR IB
- 2x Intel Broadwell Xeon E5-2695v4 CPUs (36 cores total)
- 256 GB RAM (DDR-2400)
- 2x K80 or P100 GPUs

Skylake Cluster
64 Dell PowerEdge R640 nodes with EDR IB
- 2x Intel Skylake Xeon Gold 6150 CPUs (36 cores total)
- 192 GB RAM (DDR4-2666)

KNL Cluster
142 KOI S7200AP nodes with dual rail Omnipath Gen.1 interconnect
- 1x Intel Xeon Phi 7230 CPU (256 log. Cores total)
- 192 GB RAM (DDR4-1200)

ML Cluster
5 HP Proliant XL270d Gen10 nodes with EDR IB
- 2x Intel Xeon Gold 6248 CPUs (40 cores total)
- 768 GB RAM (DDR4-2933)
- 8x V100 GPUs

New HPC cluster for NSLS2
30 1U nodes with EDR IB
- 2x Intel Cascade Lake Xeon Gold 6252 (48 cores total)
- 768 GB RAM (DDR4-2933)
- 12 of the hosts with 2x V100 GPUs

NSLS2 Cluster

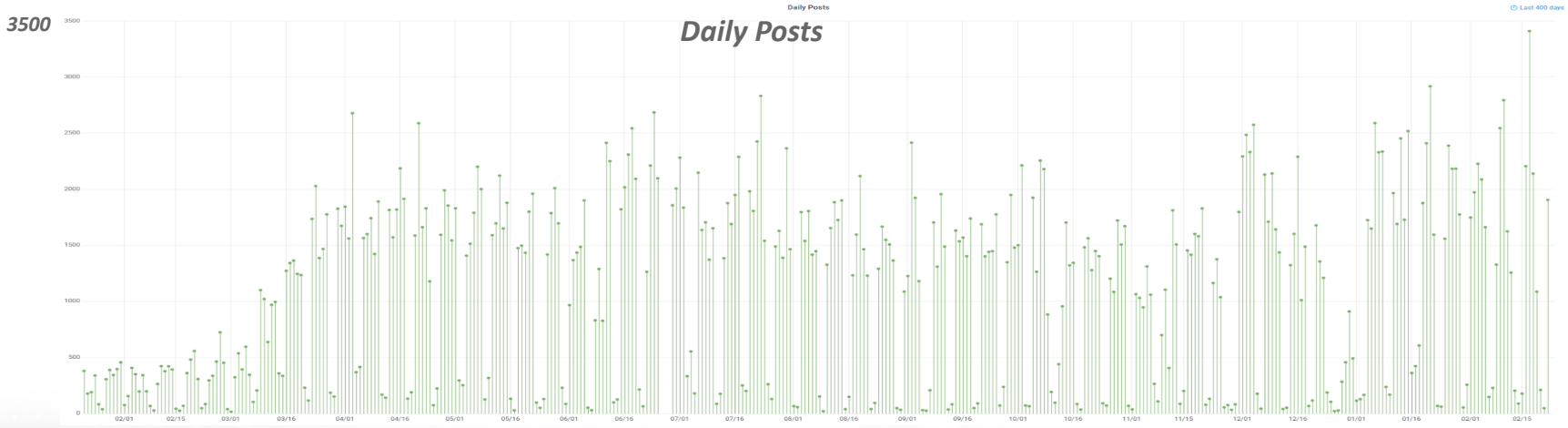
Storage

- Disk:
  - GPFS: 7 filesystems, 14PB
  - Lustre: 7PB
  - dCache: 55PB
  - XROOTD: 11PB

- HPSS:
  - ~202.6 PB accumulated data
  - Newly acquired TSM TS4500 tape library and data mover are in production since Apr 2021
  - Upgrade to HPSS v9.x in Aug 2021
  - Evaluation for new 20K-slot tape library
  - Relocation to new data center

Other tools and services

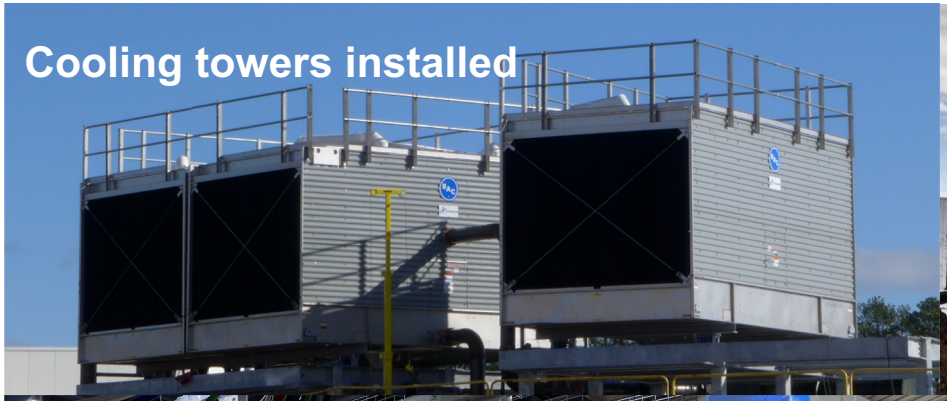
- BNL Box: 253 users ~3TB since production
- Invenio
- Indico
- Jupyter on both HPC/HTC
- Discourse
- Renana instance on staff k8s
- Gitea
- MatterMost
- …

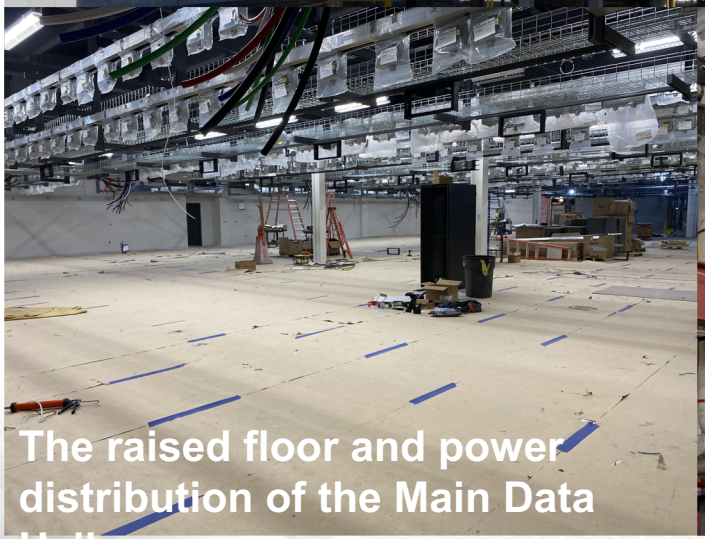


Mattermost Chat usage after lockdown (February 2020 - March 2021)

New Datacenter (CFR) Construction

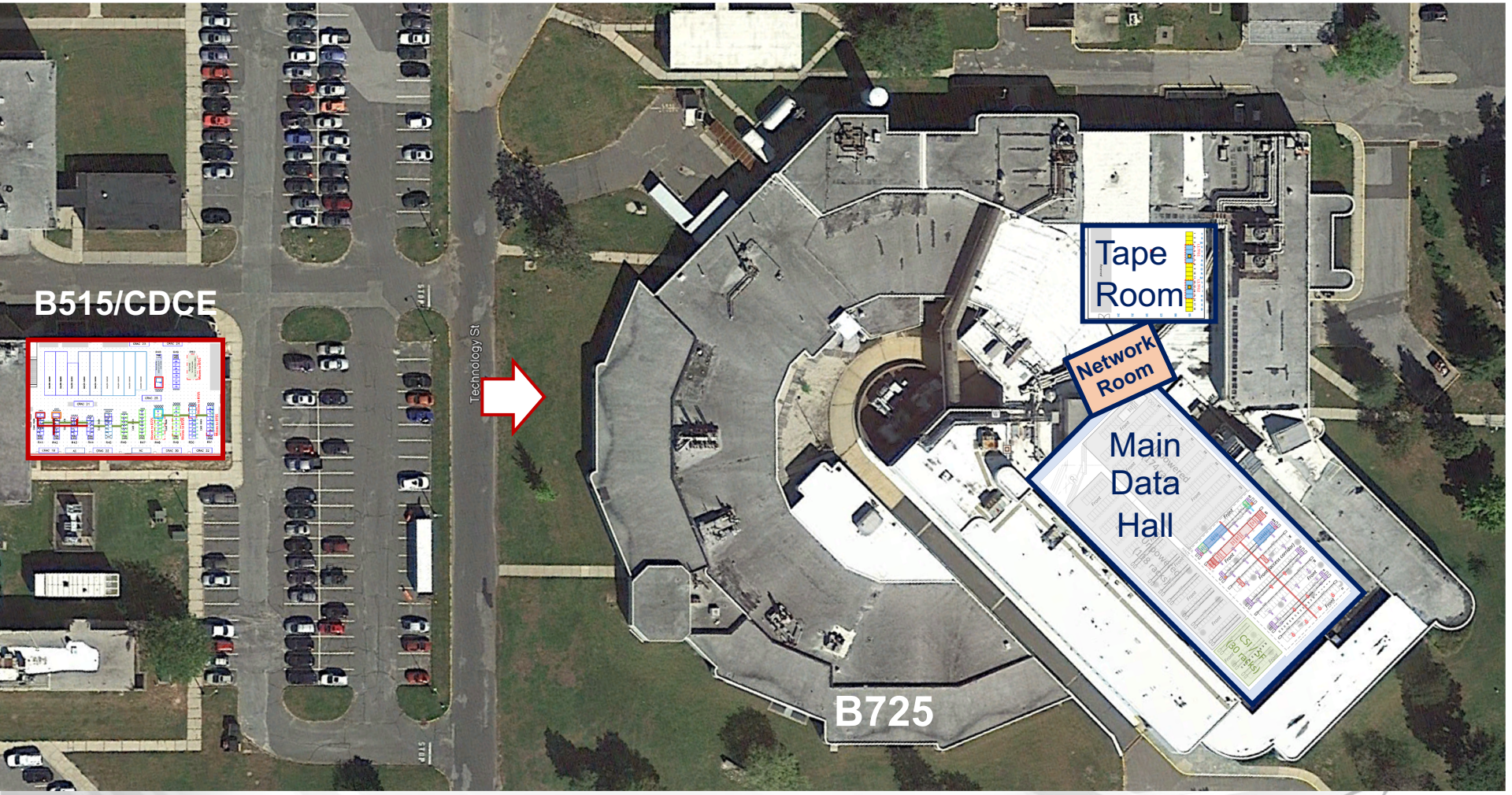
July 2020 - Mar 2021: construction is going ahead after 3 months of delay in 2020Q2 due to COVID-19. The early occupancy of datacenter is expected to begin in June 2021 (network equipment deployment starting Apr)





New datacenter is being designed & constructed for the SDCC Facility in B725 in FY19-21 period, with migration of most of the spinning disk storage and all of the compute capability (mostly via gradual HW refresh process) to the new datacenter from the existing B515 based datacenter to happen in FY21-23

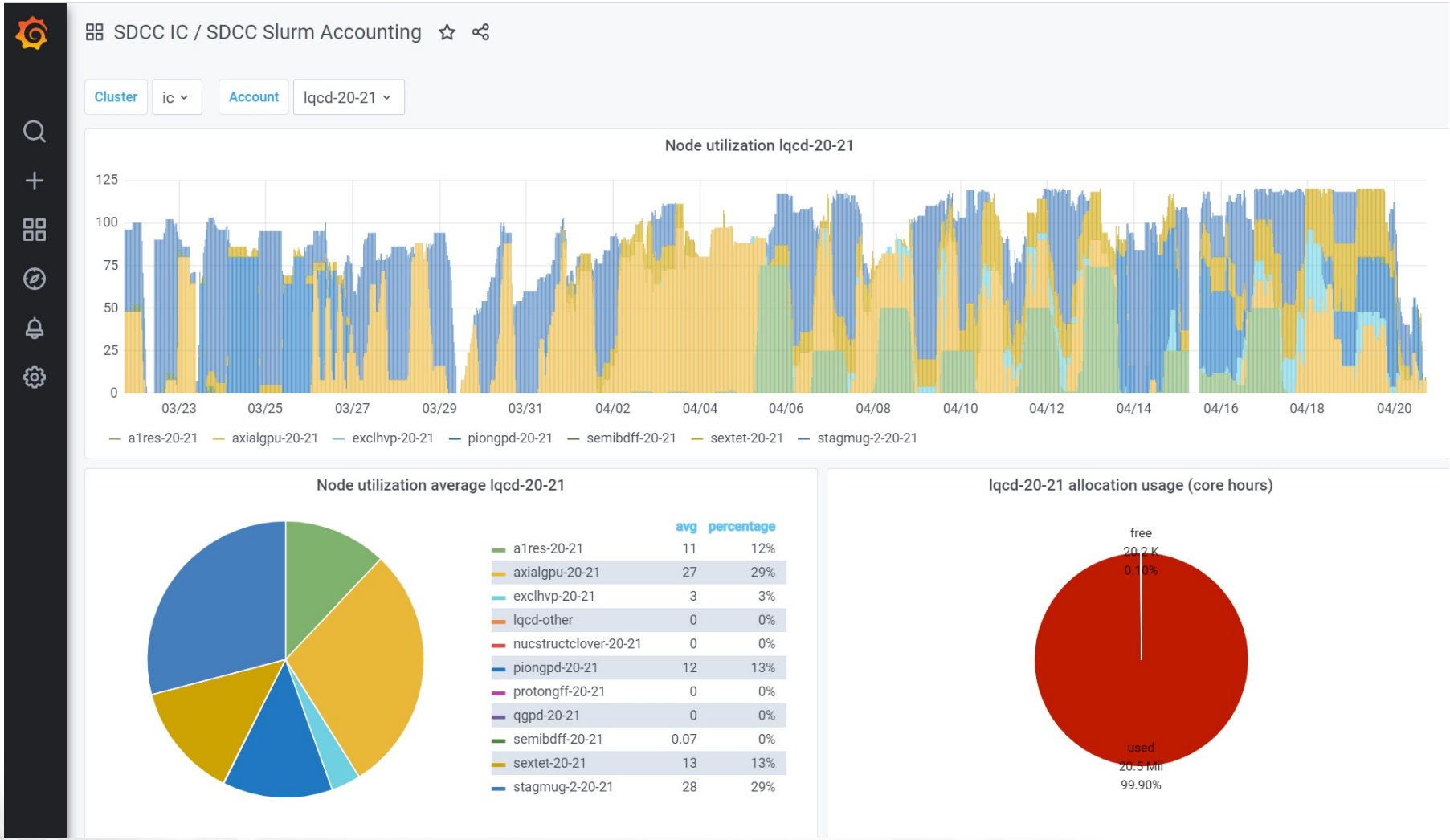
3x more floor space and power Room to grow if need Higher PUE Rear Door Heat Exchanger for racks



Monitoring

- Several tools available
  - Graphical interface here (**authentication required**)
    - https://monitoring.sdcc.bnl.gov/pub/grafana/
  - Accounting information (No authentication)
    - https://monitoring.sdcc.bnl.gov/pub/allocation
    - LQCD only
      - https://monitoring.sdcc.bnl.gov/pub/allocation/lqcd.html
      - After loading module lqcd, Command line “lquota” for same information

Accounting https://monitoring.sdcc.bnl.gov/pub/grafana/



Accounting https://monitoring.sdcc.bnl.gov/pub/allocation/lqcd.html

BNL SDCC LQCD Projects Usage Sumary

Institutional Cluster

(Sky Core Hours)

updated: 2021-04-20 00:03:35

	Cluster	Account	Start Date	End Date	Allocation	Usage	Usage(%)
	Annie-IC	lqcd-20-21	2020-07-01	2021-06-30	37,905,000	38,734,612	102.19%

	Project	Original SPC Allocation	Adjustment	Adjusted SPC Allocation	Usage	Progress(%)	Remain	30Day Usage	30Day BurnRate
1	stagmug-2-20-21	4,655,000	(734,289)	3,920,711	4,385,093	111.84%	0	1,394,821	35.58%
2	exclhvp-20-21	3,990,000	223,767	4,213,767	4,425,278	105.02%	0	131,273	3.12%
3	qgpd-20-21	3,657,500	40,207	3,697,707	3,669,330	99.23%	28,377	0	0.00%
4	piongpd-20-21	2,992,500	340,595	3,333,095	5,916,078	177.50%	0	494,588	14.84%
5	axialgpu-20-21	3,325,000	8,505	3,333,505	5,244,397	157.32%	0	1,435,172	43.05%
6	nucstructclover-20-21	4,555,250	208,679	4,763,929	3,929,348	82.48%	834,581	0	0.00%
7	sextet-20-21	3,657,500	8,329	3,665,829	2,420,691	66.03%	1,245,139	534,196	14.57%
8	protongff-20-21	4,821,250	(107,881)	4,713,369	4,439,285	94.18%	274,083	0	0.00%
9	a1res-20-21	2,992,500	4,321	2,996,821	2,470,989	82.45%	525,832	489,473	16.33%
10	semibdff-20-21	2,926,000	7,768	2,933,768	1,834,124	62.52%	1,099,644	6,766	0.23%
11	lqcd-other	0	0	0	0	0.00%	0	0	0.00%
12	UnAllocated:	332,500	2	332,502	0	0.00%	0	0	0.00%

Skylake Cluster

(Sky Core Hours)

updated: 2021-04-20 00:03:35

	Cluster	Account	Start Date	End Date	Allocation	Usage	Usage(%)
	Skylake	lqcd-sky-20-21	2020-07-01	2021-06-30	17,100,000	14,654,388	85.70%

	Project	Original SPC Allocation	Adjustment	Adjusted SPC Allocation	Usage	Progress(%)	Remain	30Day Usage	30Day BurnRate
1	gflow8f-sky-20-21	5,500,000	126,658	5,626,658	4,252,899	75.58%	1,373,759	657,726	11.69%
2	qgpd-sky-20-21	3,700,000	34,762	3,734,762	2,834,859	75.90%	899,903	210,778	5.64%
3	axialgpu-sky-20-21	1,100,000	(344,806)	755,194	474,172	62.79%	281,022	215,500	28.54%
4	stagmug-2-sky-20-21	4,000,000	140,087	4,140,087	4,245,882	102.56%	0	420,306	10.15%
5	semibdff-sky-20-21	2,800,000	43,299	2,843,299	2,846,576	100.12%	0	0	0.00%
6	UnAllocated:	0	0	0	0	0.00%	0	0	0.00%

KNL Cluster

(Sky Core Hours)

updated: 2021-04-20 00:03:15

	Cluster	Account	Start Date	End Date	Allocation	Usage	Usage(%)
	Frances-KNL	lqcd-knl-20-21	2020-07-01	2021-06-30	9,345,800	13,290,700	142.21%

	Project	Original SPC Allocation	Adjustment	Adjusted SPC Allocation	Usage	Progress(%)	Remain	30Day Usage	30Day BurnRate
1	ndbeta-knl-20-21	2,871,300	(913,194)	1,958,106	130,973	6.69%	1,827,133	26,141	1.34%
2	qcdqedta-knl-20-21	5,067,000	962,978	6,029,978	5,173,953	85.80%	856,025	724,330	12.01%
3	pdfa-knl-20-21	1,407,500	(1,407,500)	0	0	0.00%	0	0	0.00%
4	lsd4p6-knl-20-21	0	1,357,716	1,357,716	6,265,224	461.45%	0	1,721,981	126.83%
5	class-c-w0scale-20-21	20,000	0	20,000	903	4.51%	19,097	0	0.00%
6	class-c-omegamass-20-21	20,000	0	20,000	1,719,647	8,598.23%	0	778,568	3,892.84%
7	UnAllocated:	-40,000	(0)	-40,000	0	0.00%	0	0	0.00%

USQCD Access to SDCC Resources

Current computing resources allocated by LQCD ext. III

65 nodes(out of 216) allocation on GPU cluster. ~104% (as of 4/27)

Other users: CFN, MS, USATLAS, CSI….

30 nodes(out of 142) allocation on KNL cluster ~150%

Other Users: ECA

54 nodes(out of 64) allocation on Skylake cluster ~ 87%

LQCD only

Disk:

600 TB of GPFS disk storage

Tape Storage:

600 TB short term

Long Term Archive purchased 3PB tape

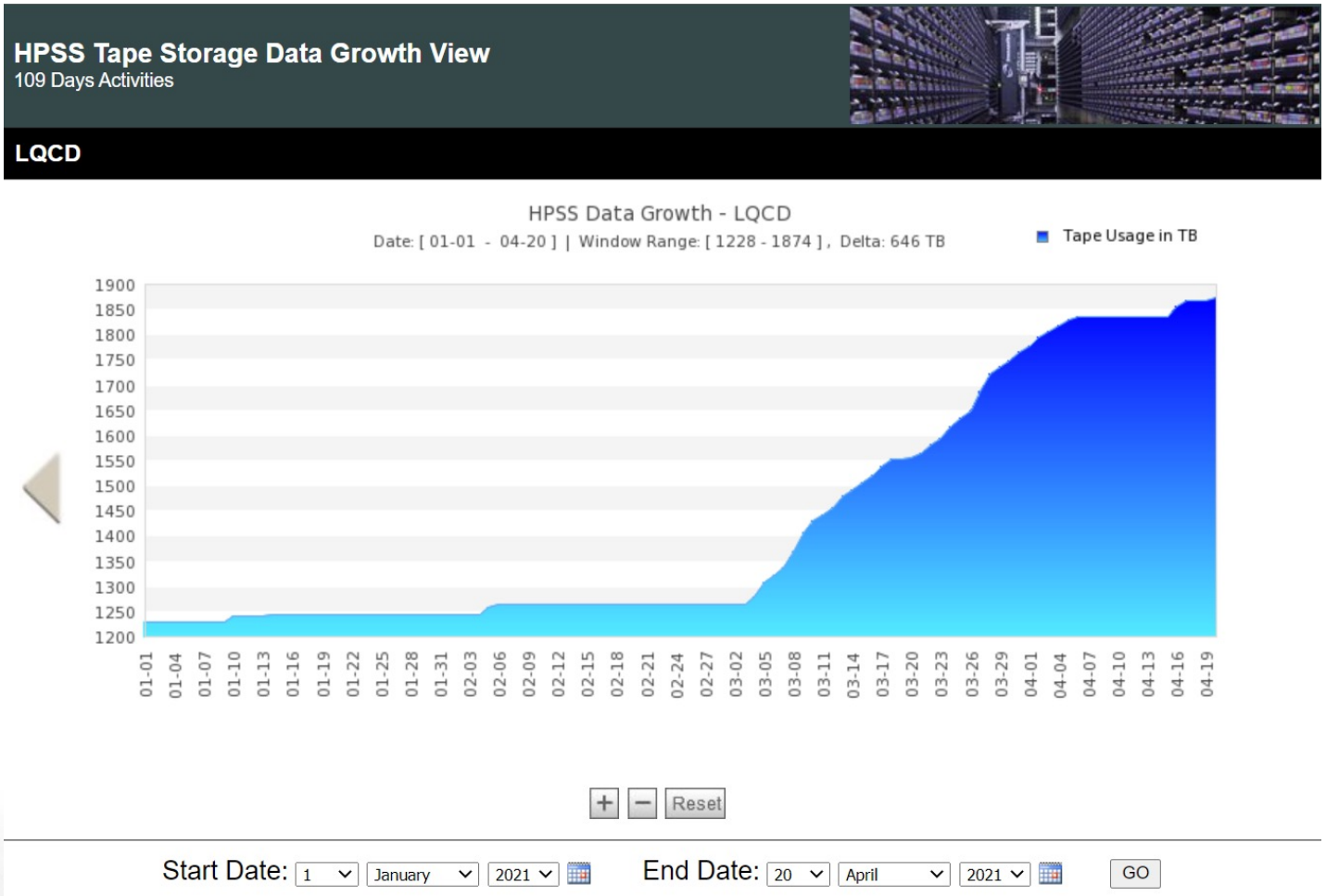
currently data 600+TB

Usage and policy :

- USQCD Jeopardy Policy applies at end of each month for sub-projects
- Scavenger/Opportunistic usage available when LQCD ext. III allocation finishes and when cluster have available resource . Jobs subject to preemption. Same for all experiments on IC.
- For the overall benefits, LQCD allowed lower priority usage after sub-project allocation used up . But these counted towards LQCD allocation.
- IC cluster :
  - Due to Undersubscription (~85%) and other experiments under use.
  - LQCD average 82 nodes usage since 7/1
  - Whole LQCD finished allocation usage on 4/21, while some sub-project allocations unfinished.

Project	Adjusted SPC Allocation	Progress(%)
stagmug-2-20-21	3,920,711	112.41%
exclhvp-20-21	4,213,767	106.12%
qgpd-20-21	3,697,707	99.23%
piongpd-20-21	3,333,095	183.33%
axialgpu-20-21	3,333,505	169.13%
nucstructclover-20-21	4,763,929	82.48%
sextet-20-21	3,665,829	74.89%
protongff-20-21	4,713,369	94.18%
a1res-20-21	2,996,821	83.17%
semibdff-20-21	2,933,768	62.61%

LQCD Tape data growth at BNL



User Support

- Facility website www.sdcc.bnl.gov .
  - New accounts
    - Instructions on website
    - Usually ~24 hours to process after verification
  - User support requests
    - SDCC policy is to respond within 3 business days. Majority is resolved within this period
    - In the July 1, 2019 to June 30, 2020, 165 tickets were submitted to ticket queues (majority from LQCD users) and 164 tickets were responded within 3 days.
      139 were resolved within 3 business days
- Bi-weekly meetings between facility staff and program/experimental Liaisons
  - Agenda on https://indico.bnl.gov/category/169/
  - Remote access via ZoomGov—Minutes of meeting posted for those who cannot join in person or remotely

Recent Developments

- Redesigned Web Page https://www.sdcc.bnl.gov with Drupal model
  - Authentication easily managed through usage of Keycloak
  - Enabled Drupal roles and page access, allowing for secure private pages based on user role and open public pages for all visitors
- NSLS2 computing activity in SDCC:

  30 node HPC cluster, 12 nodes Jupyter cluster ,

  New REHV with 7 dedicated hypervisors able to host100s of VMs, Apache Guacamole used as Remote desktop,

  Luster accessible via Globus/SFTP/SSHFS/SMB,

  Dedicated ssh gateway
- Next Generation of Institutional Cluster :
  - IC user Survey feed back suggest most people like x86 CPUs/Nvidia GPUs (code readiness)
  - Currently Benchmarking from vendor testbed and in house available computing nodes
  - Expect new IC online end of Year .

Questions?