BNL Site Report

Tony Wong

April 26, 2019

BROOKHAVEN NATIONAL LABORATORY

U.S. DEPARTMENT OF ENERGY

Scientific Data and Computing Center (SDCC)

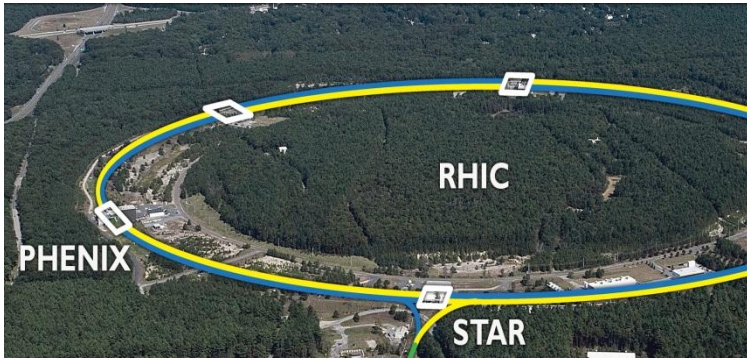
- Support for various programs:

RHIC, LHC ATLAS, BER ARM, LQCD, RIKEN, BES Center for Functional Nano Materials, National Synchrotron Light Source II, National Nuclear Data Center, Simons Foundation,…

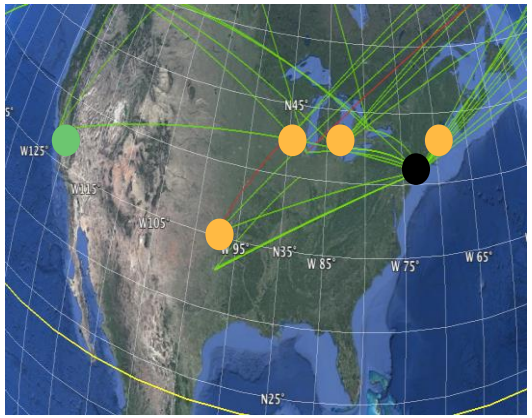
- ~1700 users from 20 projects (<10 to 100‘s users/project)
- Staff
  - 36 full-time regular members
  - 1 post-doc
  - 5 summer students
  - 2 current openings

24/7 availabi

SDCC support for HEP experiments

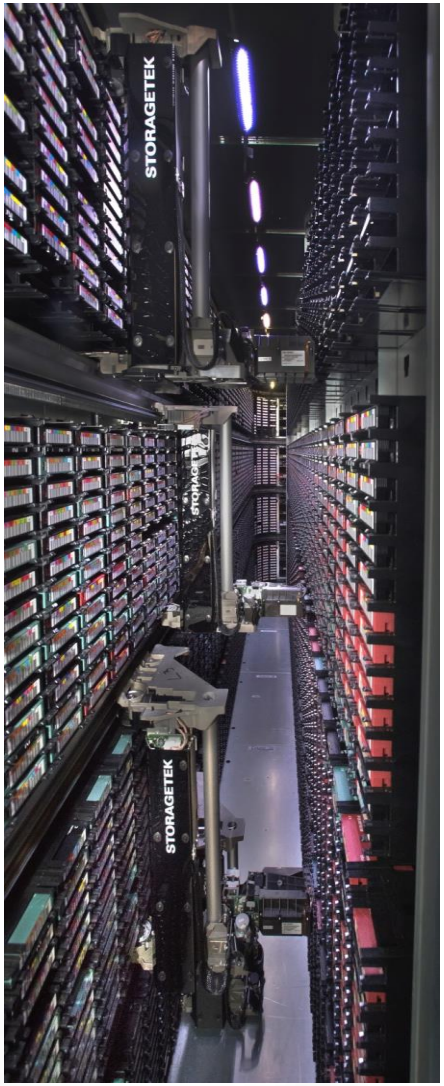


- **The RHIC Tier 0**
  - Store and process data from RHIC experiments
  - Provide analysis means for 1'200 users
  - Long term data preservation
  - Simulation resources for future programs (sPHENIX & EIC)
- **The US ATLAS Tier 1**
  - ~25% of ATLAS Tier 1 computing capacity worldwide
  - Store RAW 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 institutes (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**



SDCC Resources Summary

- 90+k CPU cores — 4 PFlops
  - 3 HPC Institutional Clusters (GPU, KNL, Skylake)
- ~80 PB of disk storage
  - Central and distributed storage systems
- 160+ PB of tape storage
  - Largest HPSS tape library in the US, 3rd worldwide
- 2x100 Gbps connection to ESnet
  - Onsite ESNet support



New BNL Data Center

- Existing data center is full
- Construction of new data center began last month
  - Approximately 3x times more floor space and electrical power with room to expand if needed
  - Higher PUE (power utilization efficiency) – mandated by DOE
  - Any new LQCD-accessible systems post-2021 would be housed in the new data center

New BNL Data Center

Proposed solution: Constructing the new datacenter in B725 in FY19-21, migrating all spinning disk storage and compute to it in FY21-23; leaving the B515 datacenter reduced to just one area (CDCE) as a tape storage room

Federated User Management

- SDCC moving towards accepting selected federated identity provider (IDP) for user management
  - First step towards Single Sign-On (SSO) with Multi-Factor Authentication (MFA)
  - InCommon and OneID– used at many universities and labs
  - SDCC user accounts still valid for now
  - Several issues still unresolved
    - Trust levels
    - Authorization concerns
    - Resources available to BNL users vs. non-BNL users
  - Several applications already enabled (Jupyter, Indico and Invenio)
  - Potentially beneficial to LQCD in the future—diminish the need to obtain a BNL appointment and an SDCC account
  - Evolving cyber-security policy to accommodate federated access to BNL resources

SDCC support for HPC

- Institutional Clusters
  1. CPU-GPU cluster (aka “Annie”) with 216 compute nodes (36 physical core Xeon Broadwell and 2 GPUs each: K80 & P100) inter-connected with dual-rail Infiniband EDR
  2. KNL cluster (aka “Francis”) with 144 nodes (64 logical cores) interconnected with dual-rail Intel OPA
  3. Skylake cluster: 64 nodes (36 physical cores each) with single-rail Infiniband EDR
- In production since January 2017
  - >250 registered users (150 last year)
  - Well utilized

MoUs (describing level of resources and services) organized with each user community

Monitoring

- Several tools available
  - Graphical interface here (authentication required)
    - https://monitoring.sdcc.bnl.gov/pub/grafana/
  - Accounting information
    - https://monitoring.sdcc.bnl.gov/pub/allocation/index1.html

Accounting

- Annie-IC
  - cfn
  - ms
  - proposal
  - lqcd-sky
  - lqcd-18-19
    - piondf-18-19
    - thermog-18-19
    - axialgpu-18-19
    - chispin-18-19
    - stagmug-2-18-19
    - nucstrucclover-18-19
    - semibdff-18-19
    - class-c-18-19
  - helpdesk-sky
  - csi
  - usatlas
- Frances-KNL

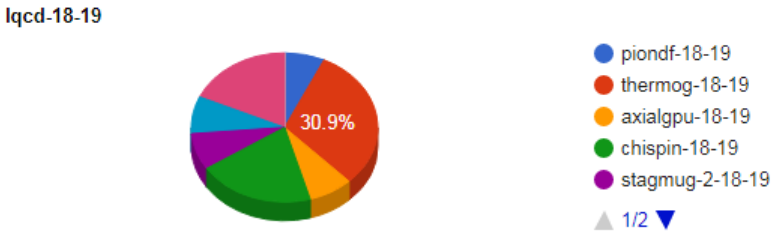
BNL SDCC Allocation Usage

(CPU Hours)

updated: 2019-04-24 00:02:31

Cluster	Account	Start Date	End Date	Allocation	Usage	Usage(%)
Annie-IC	lqcd-18-19	2018-07-01	2019-06-30	28,659,894	15,853,410	55.32%

	Project	Usage	Allocation	Remain	Used(%)	Pace
1	piondf-18-19	1,104,048	1,227,333	123,285	89.96%	109.88%
2	thermog-18-19	4,895,141	5,130,000	234,859	95.42%	116.56%
3	axialgpu-18-19	1,224,780	3,305,000	2,080,220	37.06%	45.27%
4	chispin-18-19	3,194,924	3,146,666	-48,258	101.53%	124.02%
5	stagmug-2-18-19	1,258,599	1,258,666	67	99.99%	122.14%
6	nucstrucclover-18-19	1,280,270	2,200,000	919,730	58.19%	71.08%
7	semibdff-18-19	2,893,464	2,833,333	-60,131	102.12%	124.74%
8	class-c-18-19	2,185	207,360	205,175	1.05%	1.29%
9	UnAllocated:	0	9,351,536	9,351,536	0.00%	0.00%



LQCD Access to SDCC Resources

- Initial agreement in 2016
- Current resources allocated
  - 800k node-hour allocation on CPU-GPU cluster
  - 694k node-hour allocation on KNL cluster
  - 561k node-hour allocation on Skylake cluster
  - 600 TB of GPFS storage
- Usage policy
  - Allocation valid for entire fiscal year
  - SDCC does not decrement unused allocation as a function of time, but allocation is increasingly “at risk” as we approach end of year when resource contention can become an issue

User Support

- Facility website is www.sdcc.bnl.gov
  - New accounts
    - Instructions on website
    - Usually ~24 hours to process after verification
    - Delays in account creation—see slide on federated user management
  - User support requests
    - Since January 2018, 280 tickets submitted and resolved (~92% within 2 business days)
    - currently 5 open tickets in the system
- Bi-weekly meetings between facility staff and program/experimental Liaisons
  - Agenda on https://indico.bnl.gov/category/169/
  - Remote access via BlueJeans—Minutes of meeting posted for those who cannot join in person or remotely

Recent Developments

- KNL used almost exclusively by the LQCD community
  - The Xeon Phi line has been discontinued by Intel
- BNL investigating future technologies such as quantum computing and FPGA's
- Globus endpoint @ the SDCC
  - Fast point-to-point data transfer mechanism
  - Available on Institutional Clusters
- Tape archival services available for LQCD
  - All hardware installed and tested
  - Initially 600 TB of tape storage with room to grow
  - Interface mechanism and documentation (including a Data Management Plan) available in early May 2019
- SDCC support for single-node (no low-latency interconnect) applications
  - Discussing solutions to adjust resources to meet demand and interest
  - Feedback from LQCD community?