BNL Site Report

Tony Wong

April 20, 2018

Scientific Data and Computing Center (SDCC)

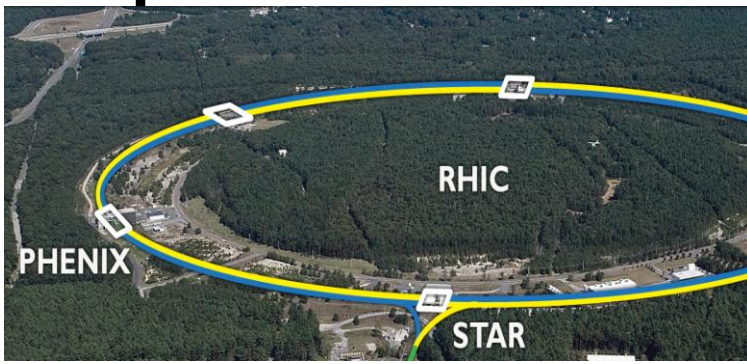
- Support for various programs:

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

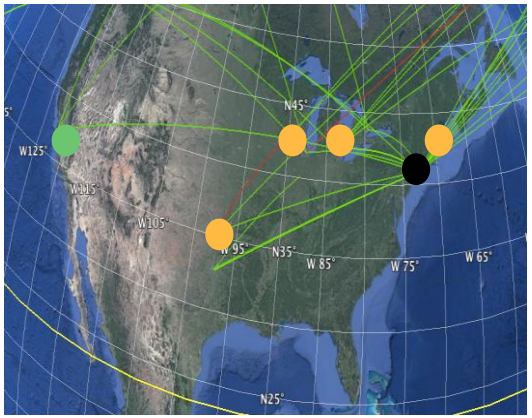
- ~1500 users from 20 projects (<10 to 100‘s users/project)
- Staff
  - 35 full-time regular members
  - 2 students
  - 4 summer students + 2 more full-time regular members soon

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
- **The Belle II data center outside Japan**
  - **Initial operations began on Oct. 2017**
  - **Data taking expected in Fall 2018**



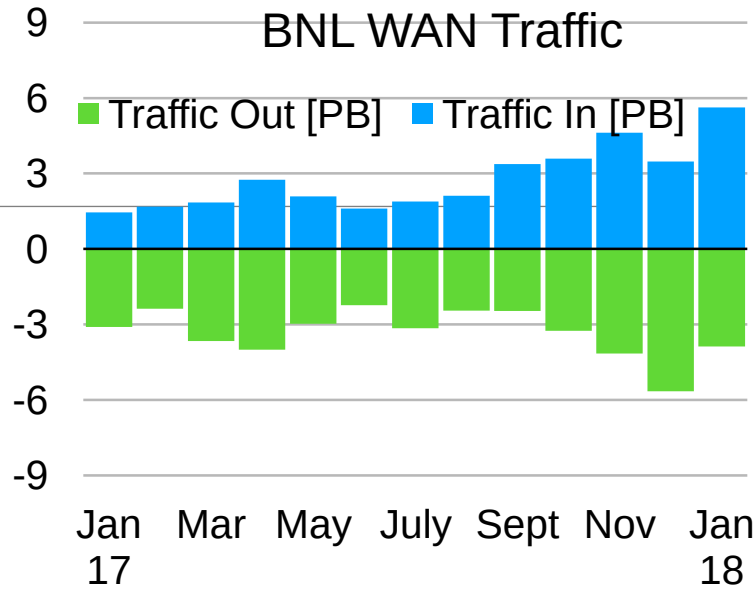
SDCC in numbers

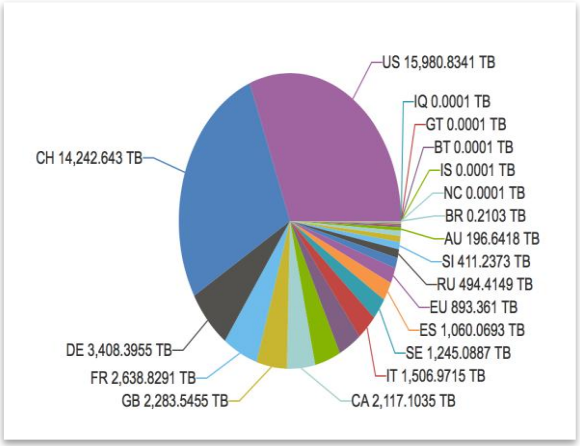
- 90+k CPU cores — 4 PFlops
  - 3 HPC Institutional Clusters (GPU, KNL, Skylake)
- ~90 PB of disk storage
  - Central and distributed storage systems
- 130+ PB of tape storage
  - Largest HPSS tape library in the US, 3rd worldwide
- 2x100 Gbps connection to ESnet
  - Onsite ESNet support

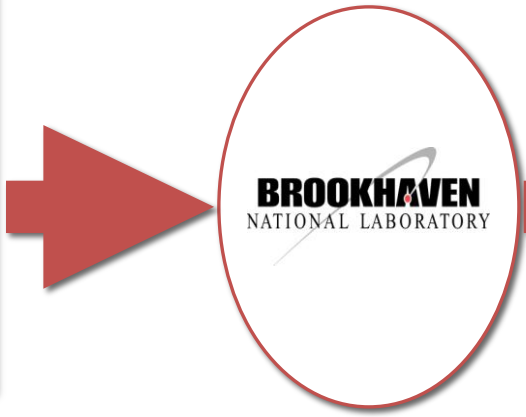
[1] http://www.hpss-collaboration.org/customersT.shtml

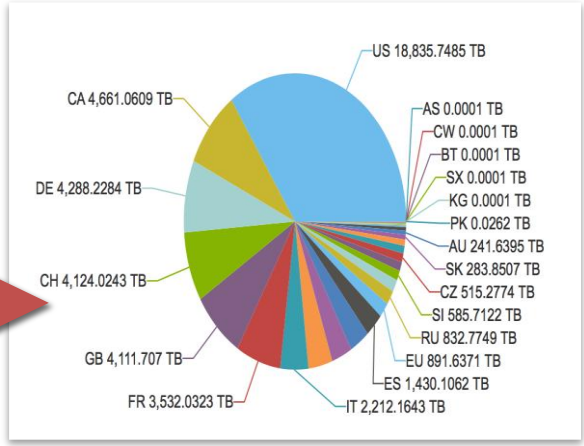
Data volume in 2017

- Transfer orchestrated by FTS (File Transfer System)
- Network performant routing
- Scheduling, etc..
- Data import : 30 PB
- Data export : 32 PB
- >500 PB (RHIC+ATLAS) processed, 1/3 by user analysis









Active tape storage : 2017 statistics

21 PB

25 PB

Archived

19M files

Retrieved

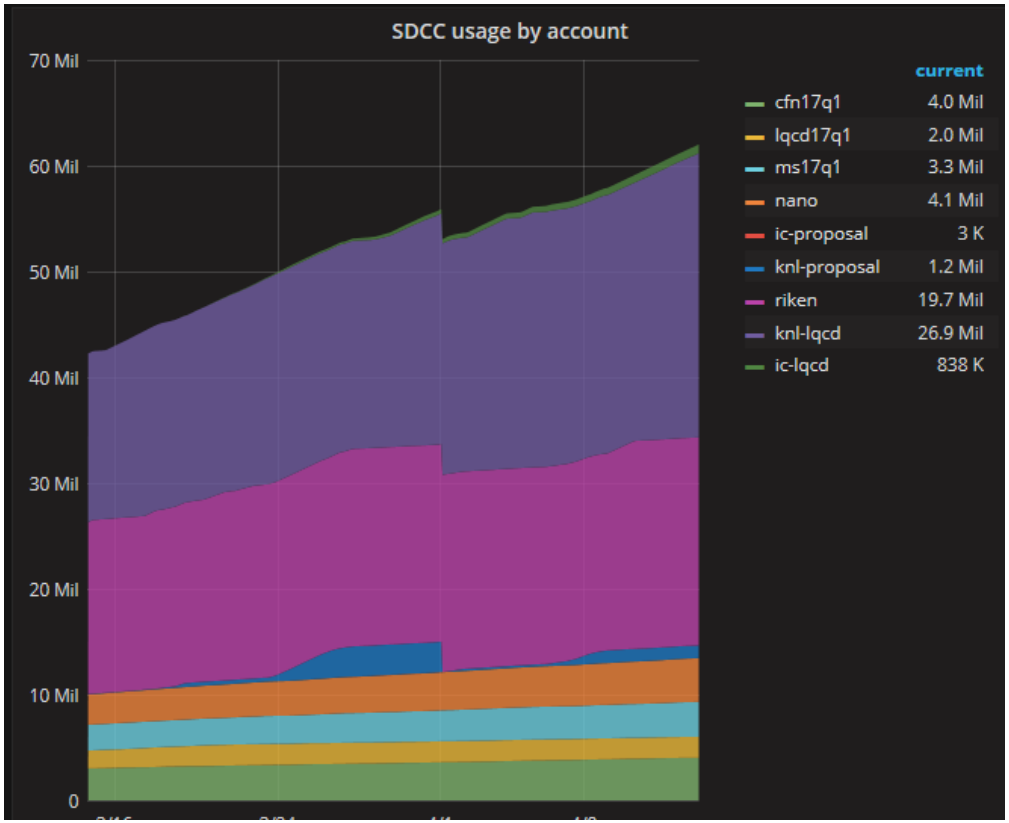
12M files

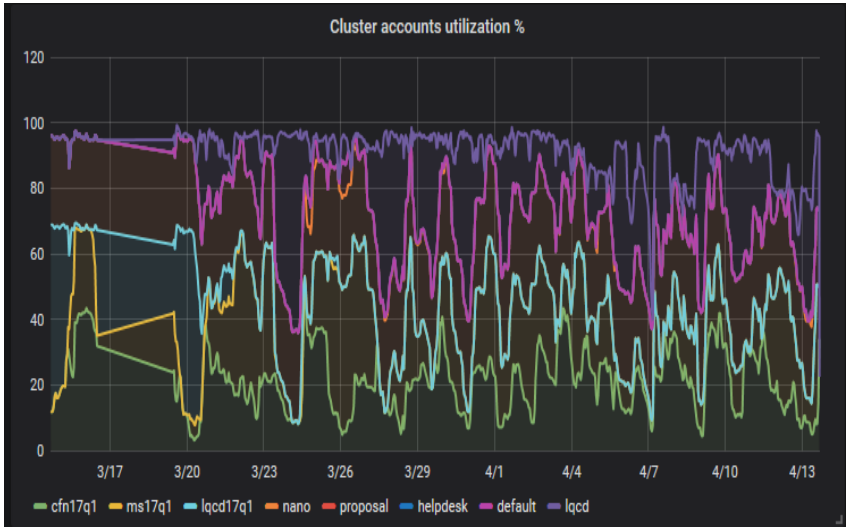
SDCC support for HPC

- Institutional Clusters
  1. CPU-GPU cluster (aka “Annie”) with 124 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
  - >150 registered users
  - Over 95% utilization

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

Monitoring





Monitoring

- Public → https://monitoring.sdcc.bnl.gov/pub/grafana
- Facility users (authentication required) → https://monitoring.sdcc.bnl.gov/grafana

Accounting

- Prototype being evaluated—available on BNL public page soon

BNL SDCC CPU Allocation (prototype)

(Core Hours)

Cluster	Account	Start Date	End Date	Total Allocation	Usage	Percentage Used
Annie-IC	lqcd	2018-03-15	2018-09-30	13,824,000	1,101,783	7.97%

	Project Name	project Usage	project allocation	Free	Percentage Used	Pace
1	axial-gpu	132,016	909,166	777,150	14.52%	82.56%
2	chispin	418,934	1,212,166	793,232	34.56%	196.50%
3	class-c	0	276,480	276,480	0.00%	0.00%
4	nucstrucclover	227,496	606,000	378,504	37.54%	213.45%
5	stagmug-2	180,145	303,000	122,855	59.45%	338.04%
6	thermog	143,190	2,811,666	2,668,476	5.09%	28.96%
7	Total:	1,101,781	6,118,478	5,016,697	18.01%	102.39%

Cluster	Account	Start Date	End Date	Total Allocation	Usage	Percentage Used
Francis-KNL	lqcd	2018-01-01	2018-12-31	105,431,040	30,438,063	28.87%

	Project Name	project Usage	project allocation	Free	Percentage Used	Pace
1	class-c	428,491	1,072,166	643,675	39.96%	134.70%
2	k2pipipbc	15,328,875	41,386,666	26,057,791	37.04%	124.83%
3	qcdqedta	14,680,696	21,617,777	6,937,081	67.91%	228.88%
4	Total:	30,438,062	64,076,609	33,638,547	47.50%	160.10%

LQCD Access to SDCC Resources

- Initial agreement (for 2017)
  - support BG-Q until Sept. 2017
  - 347k node-hour allocated on CPU-GPU cluster
    - 269k node-hours used in 2017 (23,712 jobs)
  - 150 TB of GPFS storage
- Revised agreement in 2018
  - 592k node-hour allocated on CPU-GPU cluster
  - 410k node-hour allocated on KNL cluster
  - 332k node-hour allocated on Skylake cluster
  - 400 TB of GPFS storage
- Usage policy
  - Allocation valid for entire calendar 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
  - After account creation, select “Access” on website for access instructions
  - There’s a ticketing system for users to request help—select “Get Help” on website
    - Since January 2017, 306 tickets submitted and resolved
    - currently 12 open tickets in the system
- Bi-weekly meetings between facility staff and program/experimental Liaisons
  - Agenda on https://indico.bnl.gov/category/169/
  - Select “RACF Liaison Meeting”
  - Remote access (video and audio) via BlueJeans—Minutes of meeting posted for those who cannot join in person or remotely

Relevant Facility News

- CPU-GPU cluster (aka “Annie”) expansion
  - From 124 to 162 nodes by late April 2018
  - Plan to further expand to 216 later in 2018
  - Meets demand by other programs
- Skylake cluster is expected to be delivered to BNL in May 2018—available to LQCD users in June
- HTC shared cluster available to suitable workloads, but must use HTCondor instead of Slurm for batch access
- Tape robot access mechanism available for archival storage