BNL Site Report

Chulwoo Jung (Presenter, co-site architect)
Zhihua Dong (site manager)
Costin Caramarcu (co-site architect)
Tony Wong

May 1, 2020
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
  • 36 full-time regular members
  • 3 current openings
Covid-19 Researches at BNL

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

24/7 availability
SDCC Resources Summary

- 90+k CPU cores — 4 PFlops
  - 4 HPC Institutional Clusters (GPU, KNL, Skylake, ML)
- 688 GPUs
  - K80, P100, V100
- ~80 PB of disk storage
  - Central and distributed storage systems
- 165+ PB of tape storage
  - Largest HPSS tape library in the US, 3rd worldwide
- 2x100 Gbps connection to ESnet
  - 100Gbps redundancy
  - Onsite ESNet support
BNL Data Center (CFR)

- Existing data center is full
- Construction of new data center
  - Began last May,
  - Demolition phase close to finish …halted.
  - Was expect to be online On Feb 2021, but could be delayed…
- 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
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.
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 “Frances”) with 144 nodes (64 cores each) interconnected with dual-rail Intel OPA
  3. Skylake cluster: 64 nodes (36 physical cores each) with single-rail Infiniband EDR
  4. Machine Learning Cluster: 5 EDR connected nodes, each have 8 Nvlink connected V100 GPU
  5. Start In production since January 2017
     - >300 registered users

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
    • LQCD only
      • https://monitoring.sdcc.bnl.gov/pub/allocation/lqcd.html
      • After loading module lqcd, Command line “lquota” for same information
### BNL SDCC LQCD Projects Usage Sumary

#### Institutional Cluster

**Sky Core Hours**

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Account</th>
<th>Start Date</th>
<th>End Date</th>
<th>Allocation</th>
<th>Usage</th>
<th>Usage(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annie-IC</td>
<td>lqcd-19-20</td>
<td>2019-07-01</td>
<td>2020-06-30</td>
<td>6,728,385</td>
<td>4,783,845</td>
<td>71.10%</td>
</tr>
<tr>
<td>Project</td>
<td>Original SPC Allocation</td>
<td>Adjusted SPC Allocation</td>
<td>Usage</td>
<td>Progress(%)</td>
<td>Remain</td>
<td>30Day Usage</td>
</tr>
<tr>
<td>1</td>
<td>stagmug-2-19-20</td>
<td>1,034,714</td>
<td>425,017</td>
<td>1,459,731</td>
<td>1,609,524</td>
<td>110.26%</td>
</tr>
<tr>
<td>2</td>
<td>semibdf-19-20</td>
<td>691,747</td>
<td>273,454</td>
<td>905,201</td>
<td>857,264</td>
<td>88.82%</td>
</tr>
<tr>
<td>3</td>
<td>nucstructlover-19-20</td>
<td>842,865</td>
<td>170,303</td>
<td>1,013,188</td>
<td>740,004</td>
<td>73.04%</td>
</tr>
<tr>
<td>4</td>
<td>axialp-19-20</td>
<td>697,560</td>
<td>296,936</td>
<td>994,496</td>
<td>676,647</td>
<td>68.04%</td>
</tr>
<tr>
<td>5</td>
<td>sextet-19-20</td>
<td>546,422</td>
<td>156,075</td>
<td>702,497</td>
<td>604,166</td>
<td>86.00%</td>
</tr>
<tr>
<td>6</td>
<td>class-c-19-20</td>
<td>10,765</td>
<td>0</td>
<td>10,765</td>
<td>615</td>
<td>5.72%</td>
</tr>
<tr>
<td>7</td>
<td>qgopt-19-20</td>
<td>1,490,742</td>
<td>0</td>
<td>1,490,742</td>
<td>206,626</td>
<td>10.71%</td>
</tr>
<tr>
<td>8</td>
<td>UnAllocated</td>
<td>-1,510,527</td>
<td>1,593,273</td>
<td>82,766</td>
<td>0</td>
<td>0.00%</td>
</tr>
</tbody>
</table>

#### Skylake Cluster

**Sky Core Hours**

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Account</th>
<th>Start Date</th>
<th>End Date</th>
<th>Allocation</th>
<th>Usage</th>
<th>Usage(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skylake</td>
<td>lqcd-sky-19-20</td>
<td>2019-07-01</td>
<td>2020-06-30</td>
<td>17,278,272</td>
<td>11,118,173</td>
<td>64.35%</td>
</tr>
<tr>
<td>Project</td>
<td>Original SPC Allocation</td>
<td>Adjusted SPC Allocation</td>
<td>Usage</td>
<td>Progress(%)</td>
<td>Remain</td>
<td>30Day Usage</td>
</tr>
<tr>
<td>1</td>
<td>semibdf-sky-19-20</td>
<td>4,500,000</td>
<td>2,414,093</td>
<td>6,914,093</td>
<td>5,208,518</td>
<td>75.33%</td>
</tr>
<tr>
<td>2</td>
<td>stagmug-2-sky-19-20</td>
<td>5,000,000</td>
<td>(2,731,934)</td>
<td>2,268,066</td>
<td>2,417,375</td>
<td>106.58%</td>
</tr>
<tr>
<td>3</td>
<td>etap-sky-19-20</td>
<td>1,500,000</td>
<td>300,000</td>
<td>1,800,000</td>
<td>625,947</td>
<td>34.77%</td>
</tr>
<tr>
<td>4</td>
<td>qgopt-sky-19-20</td>
<td>5,000,000</td>
<td>17,841</td>
<td>5,017,841</td>
<td>2,862,086</td>
<td>57.04%</td>
</tr>
<tr>
<td>5</td>
<td>class-sky-19-20</td>
<td>8,333</td>
<td>0</td>
<td>8,333</td>
<td>4,247</td>
<td>50.97%</td>
</tr>
<tr>
<td>6</td>
<td>UnAllocated</td>
<td>1,491,667</td>
<td>(221,728)</td>
<td>1,269,939</td>
<td>0</td>
<td>0.00%</td>
</tr>
</tbody>
</table>

#### KNL Cluster

**Sky Core Hours**

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Account</th>
<th>Start Date</th>
<th>End Date</th>
<th>Allocation</th>
<th>Usage</th>
<th>Usage(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frances-KNL</td>
<td>lqcd-knl-19-20</td>
<td>2019-07-01</td>
<td>2020-06-30</td>
<td>11,133,023</td>
<td>9,796,792</td>
<td>88.00%</td>
</tr>
<tr>
<td>Project</td>
<td>Original SPC Allocation</td>
<td>Adjusted SPC Allocation</td>
<td>Usage</td>
<td>Progress(%)</td>
<td>Remain</td>
<td>30Day Usage</td>
</tr>
<tr>
<td>1</td>
<td>qgopt-knl-19-20</td>
<td>6,756,000</td>
<td>164,205</td>
<td>6,920,205</td>
<td>7,142,674</td>
<td>103.21%</td>
</tr>
<tr>
<td>2</td>
<td>posnpr-knl-19-20</td>
<td>1,689,000</td>
<td>328,417</td>
<td>2,017,417</td>
<td>1,376,230</td>
<td>68.22%</td>
</tr>
<tr>
<td>3</td>
<td>ndbeta-knl-19-20</td>
<td>1,970,500</td>
<td>(492,622)</td>
<td>1,477,878</td>
<td>1,251,523</td>
<td>84.68%</td>
</tr>
<tr>
<td>4</td>
<td>class-c-knl-19-20</td>
<td>28,150</td>
<td>0</td>
<td>28,150</td>
<td>26,365</td>
<td>93.66%</td>
</tr>
<tr>
<td>5</td>
<td>UnAllocated</td>
<td>-928,850</td>
<td>1,618,323</td>
<td>689,373</td>
<td>0</td>
<td>0.00%</td>
</tr>
</tbody>
</table>
Accounting: [https://monitoring.sdcc.bnl.gov/pub/grafana/]
LQCD Access to SDCC Resources

- Current resources allocated
  - 578k node-hour allocation on CPU-GPU cluster
  - 309k node-hour allocation on KNL cluster
  - 480k node-hour allocation on Skylake cluster
  - 600 TB of GPFS storage
  - 600 TB of Tape Storage service (134TB used so far)

- Usage policy
  - SDCC does not decrement underused allocations as a function of time, but unused allocations are increasingly “at risk” as we approach end of year when resource contention can become an issue.
  - Opportunistic lower priority usage after allocation used up, when cluster have available resource.
USQCD data growth at BNL (134TB as of 4/20)
User Support

- Facility website is 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, 2018 to June 30, 2019, 200 tickets were submitted to ticket queues (majority from LQCD users) and 91% 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 BlueJeans—Minutes of meeting posted for those who cannot join in person or remotely
Recent Developments

- Globus endpoint @ the SDCC
  - Fast point-to-point data transfer mechanism
  - Available on Institutional Clusters (Endpoint ‘SDCC’)
    - https://www.racf.bnl.gov/experiments/sdcc/institutional-cluster/storage

- 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.
  - Information on USQCD webpage (https://www.usqcd.org/bnl/tape-archive.html)
    - Deletion or renaming creates dead space. Will count against USQCD.

- BNLBox service operational since December 2019 – file sharing and archiving (like CERNBox) between SDCC users
  - www.racf.bnl.gov/docs/services/cloud-storage/using-cloud-storage
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 establishing IDP under InCommon (independent from BNL Active Directory IDP)
• Some issues still unresolved
  • Trust levels
  • Resources available to BNL users vs. non-BNL users
• Some applications already enabled (Jupyter, Indico and Invenio)
• Potentially beneficial to LQCD users at BNL—SDCC account can be used to authenticate to external organizations
• Evolving cyber-security policy to accommodate federated access to BNL resources
Questions?