LQCD payloads/PanDA status update

P. Svirin, S. Panitkin

BigPaDA Technical Interchange Meeting, BNL
April 25, 2018

Introduction

- **Work with LQCD groups is part of the SCiDAC-4 project**
- **Currently we are working with groups from BNL (S. Mukherjee) and TJlab (R. Edwards)**
    - **Different codes, different resources, different schedules, etc**
- **The goal is to integrate computing resources used by the LQCD groups with PanDA and provide support and tools to use the resources effectively**

- **Currently we use PanDA server at EC2 for this project**

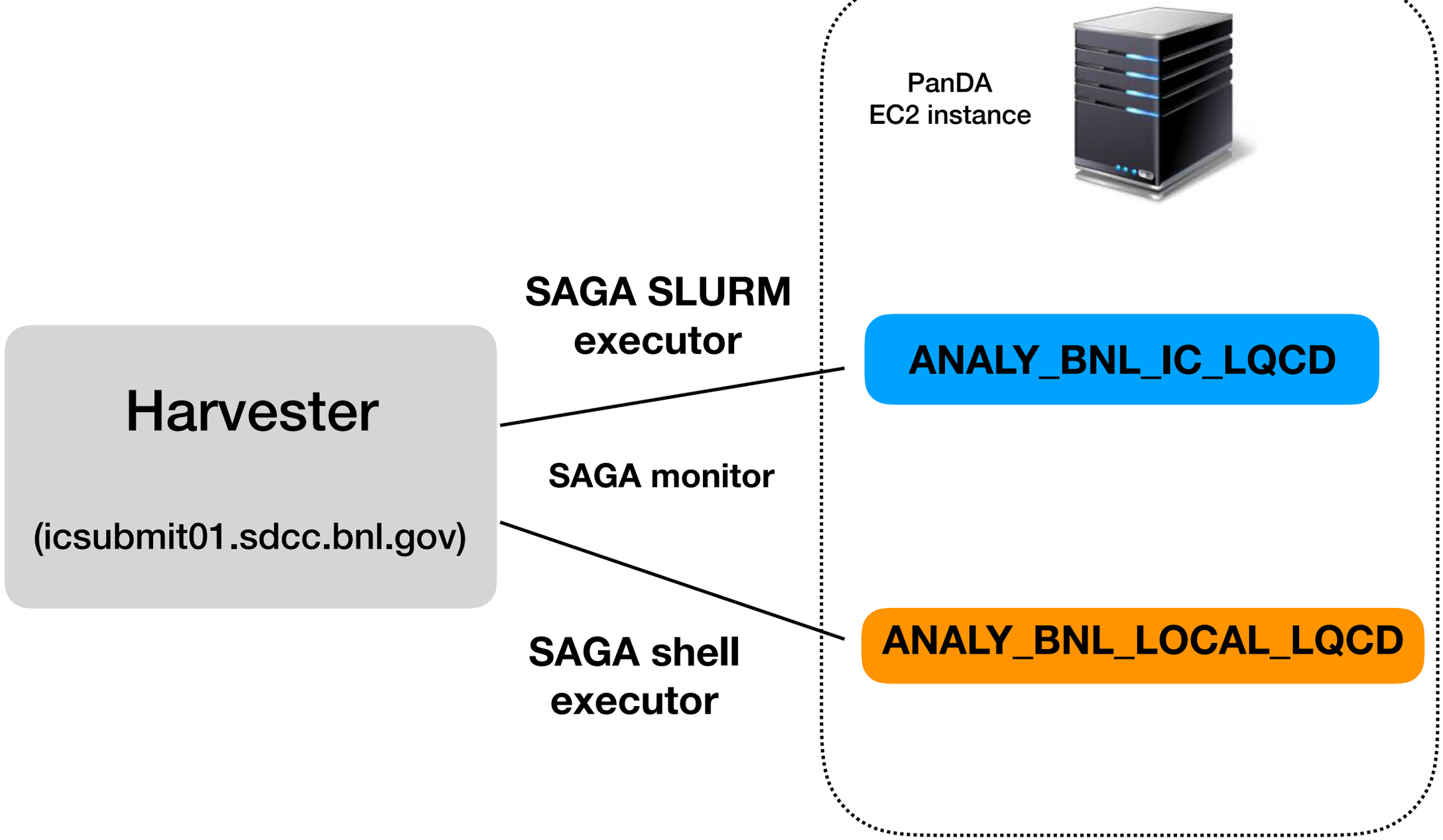
Currently available resources

- Harvesters that for LQCD payloads configured on front nodes of:
  - Titan
  - BNL Institutional GPU cluster
  - Thomas Jefferson Laboratory
  - NERSC

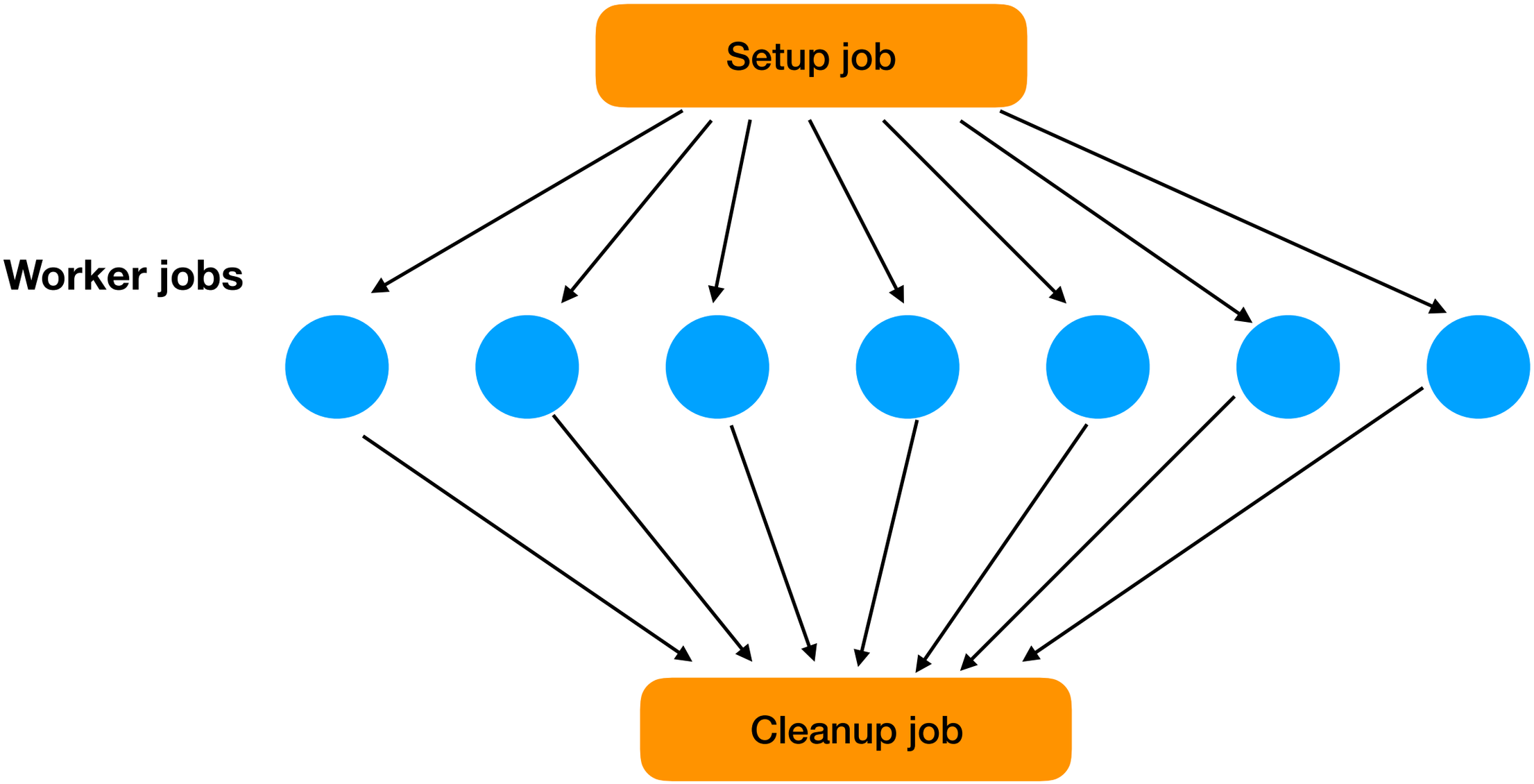
Production at BNL

- Input data: 127 data sets, each 51 GB, packed; total: 6.5 TB
- Every input is analyzed with 7 sets => 889 different worker jobs
- It is possible to have up to 50 submitted jobs in “long” queue
- Scratch disk space is limited to 1 TB, thus, it is possible to have up to 8 input data sets simultaneously
- LQCD uses PASCAL GPUs, walltime per each job: 12 hours

Production at BNL: setup



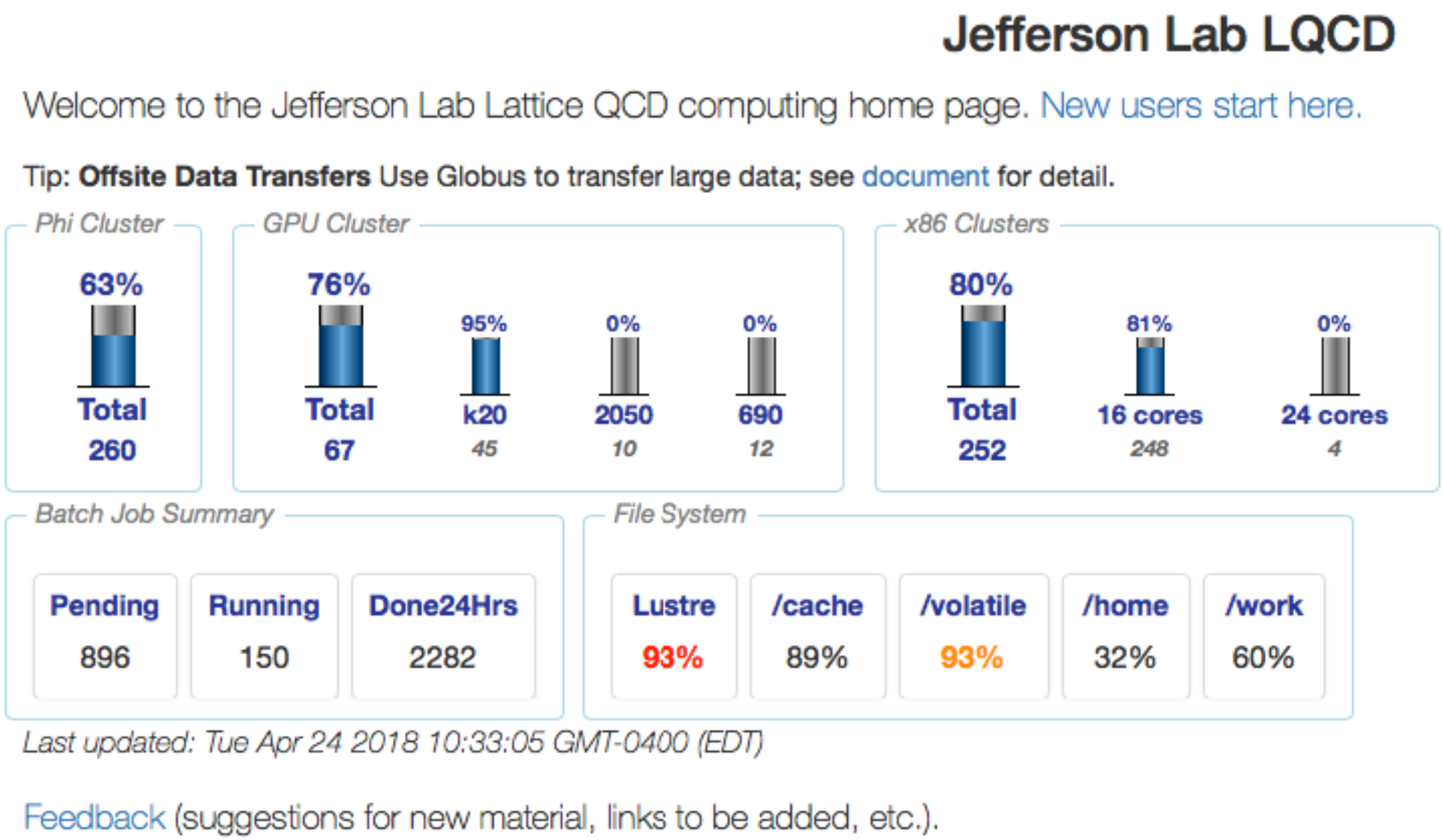
Production at BNL: workflow



Two-level payloads management:

- **per workflow**
- **per worker job (resubmit individual worker if it fails, cleanup can be called even if workers fail several times)**

Thomas Jefferson Lab



Thomas Jefferson Lab

- Harvester was installed and configured by TJLab team using our instructions

- Now they are working on their own interface with custom logic to local batch system

- Workloads differ from the ones used by the BNL group

LQCD projects at TJLab with PanDA

- Quark line contractions (3-6 months):
  - O(10^6) of independent single node jobs, walltime: 2 hours
  - validation script for jobs output to be run, ~30 minutes, may be run per each job or per bunch
  - Input data: gauge fields configurations, O(1000) files,~2GB each
  - Output: ~100 MB per job
  - X86 (Infiniband) cluster will be used
- Perambulation calculations:
  - 3 TJLab clusters+NERSC will be used with dynamic balancing between them
  - O(10^6) jobs + same validation script
  - Job sizes will be larger and depends on cluster (~160 cores) per job, linear scaling with cores change (cores count must be a multiple of 16)
  - Input and output data characteristics: same as above

Summary and next steps

- Harvesters installed at 4 sites
    - Also an instance of Harvester was installed on SummitDev front node, was tested with dummy jobs. LQCD groups do have payloads for Summit - part of the early science program
- Production started at BNL Institutional Cluster
- Globus online data transfers from OLCF have been tested with Globus Online tools in Harvester
- Both teams (TJLab and BNL) will be provided with necessary documentation and client tools (to be presented during the talk “Harvester Beyond ATLAS”) so they can manage payloads themselves
- CHEP2018: a poster about PanDA/LQCD integration

Next steps

- Finish production run at BNL - May?
    - Move PanDA@EC2 to a new VM - as soon as there is a break in LQCD production
- Get Thomas Jefferson Laboratory clusters fully integrated with panda and run production. May?
- Production on Titan if LQCD receives an allocation again
- Test Harvester on Summit as a part of Early Science program - July -August