LQCD/LSST/PanDA: integration update

P.Svirin, S.Panitkin

LQCD

- **BNL:**
  - **QGP calculations:**
    - input data: 255 configurations ~13 TB
    - output data so far: ~176 GB
    - 255 configurations*6 sets each = 1330 jobs
    - job walltime: ~9-12 hours
    - problems: sometimes payloads were going to wrong GPU types, although batch description was correct
  - Preparing for next campaigns

- **Thomas Jefferson Lab:**
  - Preparations for their campaigns on their own
  - Installed their own instances on NERSC and JLab cluster
  - Doing experiments using CLI tools for PanDA

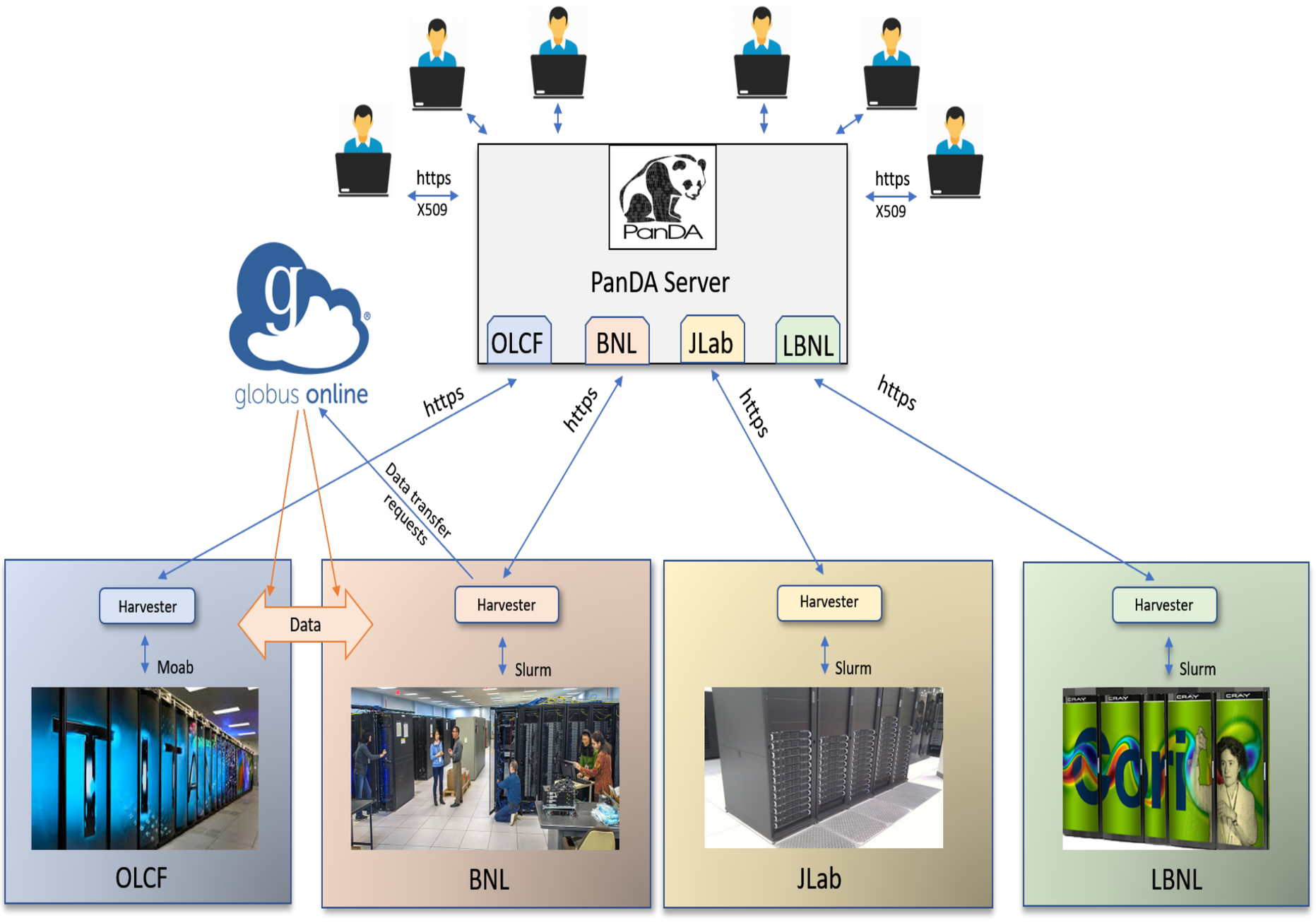
JLab payloads

Attempt#		Task ID			d:h:m:s	d:h:m:s			
18995 Attempt 10	Daniel Trewartha / Gluex	2	failed	2018-09-19 18:36	0:2:38:45	0:0:01:44	09-19 21:26	ANALY_TJLAB_LQCD	364

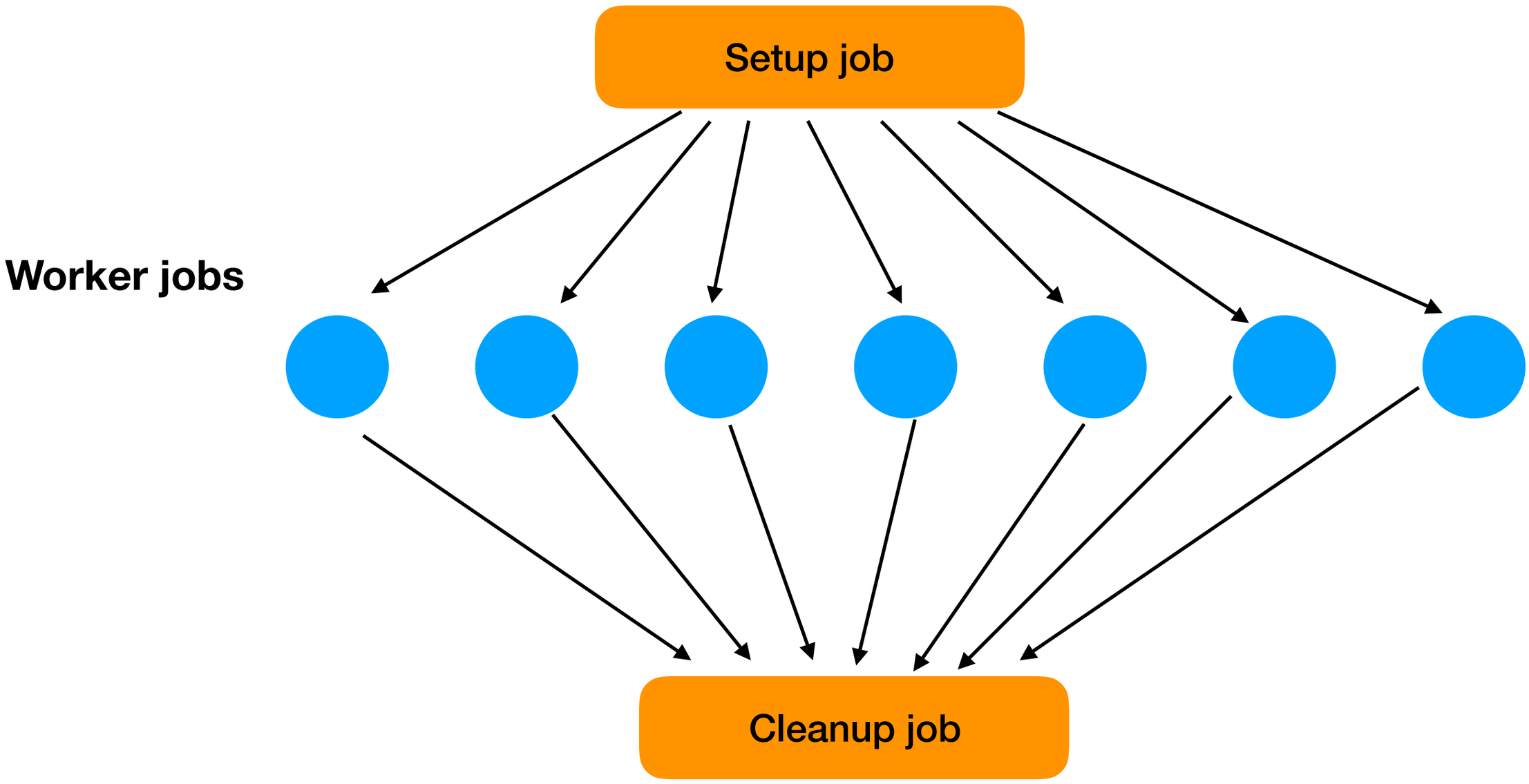
Error description	**pilot:** 146376 tmp_ZVnty+ phi spectrump 256 FAILED 1:0
Pilot error code	1008
Pilot error message	146376 tmp_ZVnty+ phi spectrump 256 FAILED 1:0

Last state change	to failed at 2018-09-19 14:21
Transformation exit code	None (non-zero exit code from job payload)
Attempt number	10 of a maximum 2
Output destination	local
Job parameters	{"nodes": 1, "outputFile": "/volatile/users/danielt/forDaniel/szscl21_24_256_b1p50_t_x4p300_um0p0856_sm0p0743_n1p265/kpi.3half/000_A1p /szscl21_24_256_b1p50_t_x4p300_um0p0856_sm0p0743_n1p265.out1000b\n.gz", "command": "cd /volatile/users/danielt/szscl21_24_256_b1p50_t_x4p300_um0p0856_sm0p0743_n1p265 /kpi.3half/000_A1p/\nif ($?SLURM_SUBMIT_HOST != 0) then\n echo \"SLURM_JOBID = $SLURM_JOB_ID\"\n echo \"SLURM_JOBNAME = $SLURM_JOB_NAME\" \n echo \"SLURM_WORKDIR = $SLURM_SUBMIT_DIR\"\n echo \"SLURM_NODEFILE\"\n $SLURM_JOB_NODELIST \| sort -u\nendif\n\nscurce /dist/intel/parallel_studio_xe_2016.3.067/psxevars.csh intel64\n\nset file_prefix=\"szscl21_24_256_b1p50_t_x4p300_um0p0856_sm0p0743_n1p265\"\n\n#/home/edwards/qcd/data/redstar/analysis /szscl21_24_256_b1p50_t_x4p300_um0p0856_sm0p0743_n1p265/kpi.3half/000_A1p\n#set particle=${PWD:h:t}\n#set mom=${PWD:t}\n\nset param_file=\"run_redstar_npt.knl.ini.pl\"\nset arch=\"12s\"\nset particle=\"kpi.3half\"\nset mom=\"000_A1p\"\n\nset run_script=\"/home/edwards/qcd/git/devel/scripts_rge/run/redstar/run_redstar_npt.knl.pl\"\n\nset spectro_prefix=$file_prefix\nset list=${spectro_prefix}.list\n\nif (! -d \"/volatile/Spectrum\") then\n echo \"Lustre is not present\"\n exit 0\nendif\n\nset newseqno=1000b\n\n\n##set out = ${file_prefix}.out${newseqno}\nset poodoo = \"/lustre/volatile/users/danielt/forDaniel/szscl21_24_256_b1p50_t_x4p300_um0p0856_sm0p0743_n1p265/kpi.3half/000_A1p/\"\nset param_file = $poodoo$param_file\n#set poodoo = \"/lustre/volatile/users/danielt/contractionTest/NF2+1/${file_prefix}/dt_temp/${particle}/${mom}\"\nmkdir -p $poodoo\npushd $poodoo\npopd\nset out = \"${poodoo}${file_prefix}.out${newseqno}\"\n/bin/rm -f $out ${out}.gz\necho \"\" > $out\n\necho \"Cfg = $newseqno\"\necho \"Cfg = $newseqno\" >> $out\n\nif ($?SLURM_SUBMIT_HOST != 0) then\n echo \"SLURM_JOBID = $SLURM_JOB_ID\" >> $out\n echo \"SLURM_JOBNAME = $SLURM_JOB_NAME\" >> $out\n echo \"SLURM_O_WORKDIR = $SLURM_SUBMIT_DIR\" >> $out\n echo \"SLURM_JOB_NODELIST\" >> $out\n $SLURM_JOB_NODELIST | sort -u >> $out\nendif\n\necho \"Starting chroma script at time= \" `date`\necho \"Starting chroma script at time= \" `date` >> $out\necho $run_script >> $out\ntime $run_script $param_file $arch $newseqno >> $out\nset tmp_status=$status\n\nif ($tmp_status != 0) then\n echo \"Some error in run in $out\"\n exit 1\nendif\necho \"Ending chroma script at time= \" `date`\necho \"Ending chroma script at time= \" `date` >> $out\necho pwd\n\ngzip -f9 $out\n\nexit 0", "name": "d60dd976-21fd-40e2-ae73-0e3969027f1f\n", "walltime": "12:00:00"}
Batch ID	146376

Future LQCD computing infrastructure



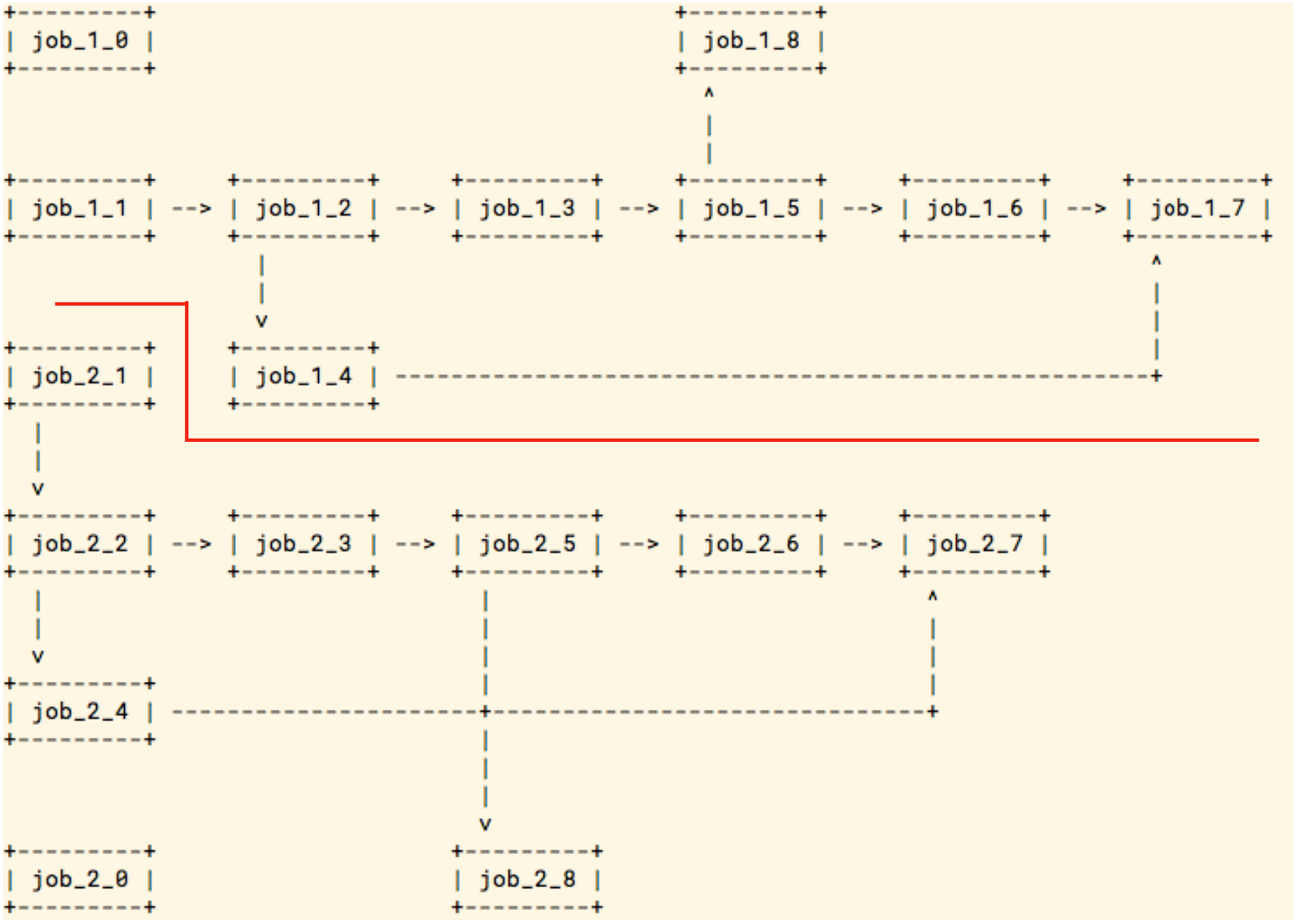
Production at BNL: workflow



Workflow template

```
name: wf_{{wf_no}}
jobs:
  job_{{wf_no}}_0:
      nodes: 1
      walltime: “00:30:00"
      resubmissions: 3
      queuename: "ANALY_BNL_LOCAL_LQCD"
      command: |+
        sleep 90
  job_{{wf_no}}_1:
      nodes: 1
      walltime: “00:30:00"
      resubmissions: 2
      continue_after_last_fail: “yes”
      queuename: "ANALY_BNL_LOCAL_LQCD"
      command: |+
        sleep 90
...
sequence:
  job_{{wf_no}}_1: job_{{wf_no}}_2
  job_{{wf_no}}_2: job_{{wf_no}}_3, job_{{wf_no}}_4
  job_{{wf_no}}_3: job_{{wf_no}}_5
  job_{{wf_no}}_4: job_{{wf_no}}_7
  job_{{wf_no}}_5: job_{{wf_no}}_6, job_{{wf_no}}_8
  job_{{wf_no}}_4: job_{{wf_no}}_7
  job_{{wf_no}}_6: job_{{wf_no}}_7
```

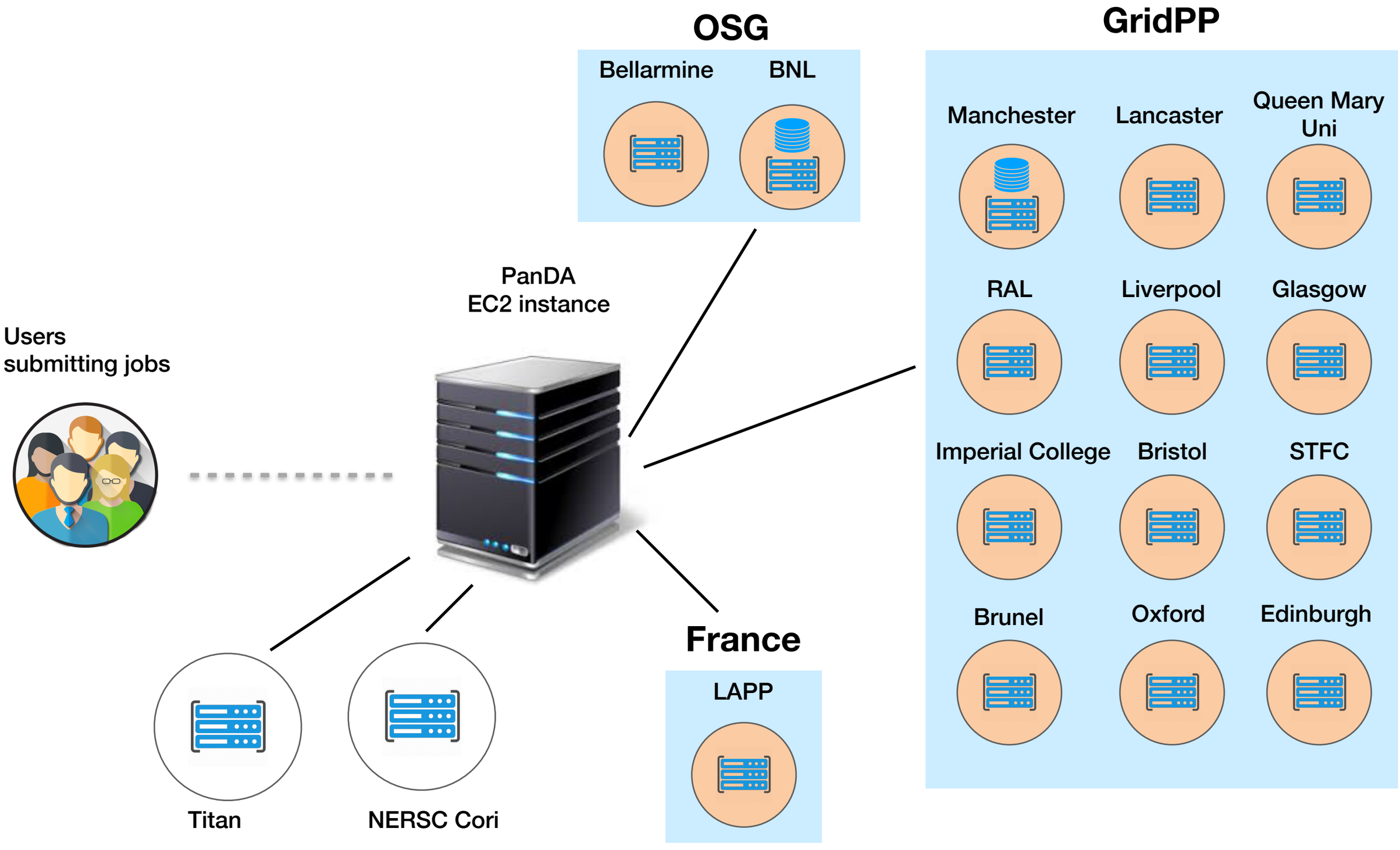
CLI tools: workflows



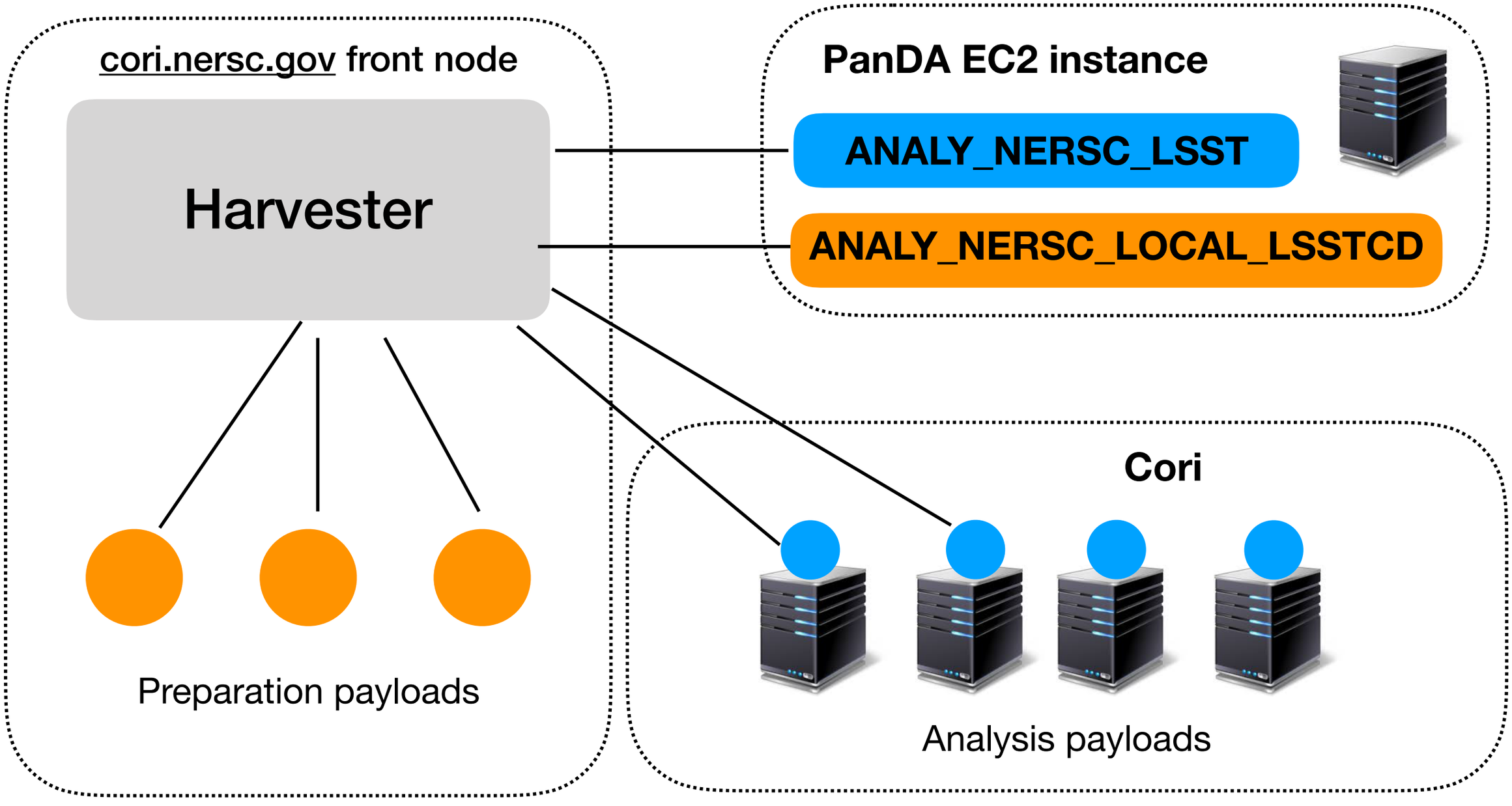
LSST/DESC

- OSG: 2 sites (BNL and Bellarmine)
- GridPP
    - 31 Grid endpoints on 12 sites configured for LSST in UK, 3 endpoints in France (LAPP Annecy)
        - Manchester alone has put online 3000 job slots which are not strictly assigned for LHC experiments.
- LSST software delivered via CVMFS
- Currently: able to acquire around 3,000 job slots without any allocations
- Storage for LSST now available: 7 European sites (~10 TB of transient data available, data is transferred to NERSC and removed from storages), 1 US (Astro storage @BNL: 200 TB)

Grid setup for LSST/DESC



PanDA Setup for NERSC



Current work and future plans

- Harvester:
  - correct propagation of error codes
  - job resubmission

- Data transfers for LQCD payloads:
  - Globus Online requires manual activation of endpoints
  - Any other way to exchange data with OLCF?

- Conferences/publications:
  - CHEP2018 poster: LQCD
  - CHEP2018 talk: LSST
  - CHEP2018 proceedings: joint LQCD/LSST/other experiments article

- Future plans:
  - test LQCD payloads on Summit via PanDA
  - use JEDI for generation of workflows?
  - GUI for defining workflows/following workflow progress?