PDSF User Meeting

- PDSF performance (UGE+SLURM)
- announcements
- AOB
- SLURM+CHOS evaluation
aggregated load on PDSF interactive nodes

https://portal-auth.nersc.gov/pdsf-mon/
Jan Balewski, NERSC

PDSF User Meeting

**CPU aggregated over month**

UGE fully loaded

SLURM not used (all load generated by Jan)

http://portal.nersc.gov/project/mpccc/balewski/pdsf3Load/latest/

plot generated: 2017-07-10 11:58:42

SLURM+CHOS

SLURM+Shifter
queue(s) load

UGE - looks fine

SLURM - no queue limits
(no monitor vs. time)

<table>
<thead>
<tr>
<th>Partition</th>
<th>Nodes per Job</th>
<th>Physical Max Cores per Job</th>
<th>Walltime per Job</th>
</tr>
</thead>
<tbody>
<tr>
<td>shared-chos</td>
<td>1</td>
<td>1-32</td>
<td>2 days</td>
</tr>
<tr>
<td>shared</td>
<td>1</td>
<td>1-32</td>
<td>2 days</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Partition</th>
<th>Total vCores [nt1]</th>
<th>Total Nodes</th>
<th>Default Mem per vCore (MB)</th>
<th>Max Mem per vCore (MB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>shared-chos</td>
<td>420</td>
<td>7</td>
<td>1000</td>
<td>2000</td>
</tr>
<tr>
<td>shared</td>
<td>176</td>
<td>4</td>
<td>1000</td>
<td>2000</td>
</tr>
</tbody>
</table>
# Aggregated utilization of UGE

<table>
<thead>
<tr>
<th>user \ timeRange</th>
<th>sum last 30_days</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-all-users : owner</td>
<td>251.1 (cpu*year), 1342121 jobs, wallT frac=1.000, CPU/wallT=0.74</td>
</tr>
<tr>
<td>alice : project</td>
<td>113.8 (cpu*year), 99351 jobs, wallT frac=0.453, CPU/wallT=0.70</td>
</tr>
<tr>
<td>atlas : project</td>
<td>16.6 (cpu*year), 386661 jobs, wallT frac=0.066, CPU/wallT=0.93</td>
</tr>
<tr>
<td>dayabay : project</td>
<td>35.2 (cpu*year), 105573 jobs, wallT frac=0.140, CPU/wallT=0.36</td>
</tr>
<tr>
<td>dybprod : owner</td>
<td>0.0 (cpu*year), 0 jobs, wallT frac=0.000, CPU/wallT=0.00</td>
</tr>
<tr>
<td>hack : owner</td>
<td>3.3 (cpu*year), 5687 jobs, wallT frac=0.013, CPU/wallT=0.97</td>
</tr>
<tr>
<td>lux : project</td>
<td>3.5 (cpu*year), 4195 jobs, wallT frac=0.014, CPU/wallT=0.95</td>
</tr>
<tr>
<td>lz : project</td>
<td>2.3 (cpu*year), 181794 jobs, wallT frac=0.009, CPU/wallT=0.92</td>
</tr>
<tr>
<td>majorana : project</td>
<td>9.4 (cpu*year), 141214 jobs, wallT frac=0.037, CPU/wallT=0.74</td>
</tr>
<tr>
<td>star : project</td>
<td>69.8 (cpu*year), 415768 jobs, wallT frac=0.278, CPU/wallT=0.92</td>
</tr>
<tr>
<td>staremb : owner</td>
<td>0.0 (cpu*year), 0 jobs, wallT frac=0.000, CPU/wallT=0.00</td>
</tr>
</tbody>
</table>

## June average

- Served 3.1k vCores
- Capacity 3.4k vCores
- CPU/WallT 0.74

## May average

- Served 3.1k vCores
- Capacity 3.4k vCores
- CPU/WallT 0.83
Mendel utilization - snapshot

Load on UGE worker nodes 2017-07-10_12.01

SLURM nodes are idle
Mendel RAM/task usage - snapshot

RAM utilization, mc01=r,12=b,13=g,15=y  2017-07-10_12.01
XrootD availability - snapshot (1/2)
XrootD availability - snapshot (2/2)

Xrd file access probability, sample 100-files/node, end=2017-07-10 08.39
### /project(a) utilization - snapshot

http://portal.nersc.gov/project/star/jthaeder/diskUsage/overview/indexExt.html
https://my.nersc.gov/data-mgt.php

```
pdsf6 $ prjquota dayabay
---------- Space (GB) ----------       ------------- Inode --------------
Project   Usage    Quota  InDoubt  Usage    Quota  InDoubt
----------   ---------  ---------  -------  ----------  ----------  ----------
      dayabay  755275   870400      2  130338118  150000000        280
```

```
pdsf6 $ prjquota dayabay
---------- Space (GB) ----------       ------------- Inode --------------
Project   Usage    Quota  InDoubt  Usage    Quota  InDoubt
----------   ---------  ---------  -------  ----------  ----------  ----------
      dayabay  708577   716800      0   2292229   10000000        0
```

```
pdsf6 $ prjquota majorana
---------- Space (GB) ----------       ------------- Inode --------------
Project   Usage    Quota  InDoubt  Usage    Quota  InDoubt
----------   ---------  ---------  -------  ----------  ----------  ----------
      majorana  37225    40960      0   2294117   40000000        0
```

```
pdsf6 $ prjquota majorana
---------- Space (GB) ----------       ------------- Inode --------------
Project   Usage    Quota  InDoubt  Usage    Quota  InDoubt
----------   ---------  ---------  -------  ----------  ----------  ----------
      majorana  56540    61440      1  4096124   10000000        280
```
Announcements

Bi-weekly office hours  12:30 -2:30pm  
Thursday,  July 20, August 3,  59-4016-CR

PDSF user meeting
   ●  Tuesday, August 8, 11am - 12pm, 59-3034-CR

SC Meeting: July 24
Outages :
Cori maintenance: this Wednesday
Edison major upgrade: July 17-31
Outage for Quarterly Maintenance: October 10, 2017
Stress-test of SLURM at PDSF

Please ask your users to test SLURM+CHOS at scale.

PIs plan discussed
At SC Meeting
June 9, 2017

Continue until Mendel is decommissioned
SLURM+CHOS: 60 root4star/node works

Load on SLURM worker nodes

2017-07-03_12.15 mc1532 $ top ibn1
load average: 60.07, 59.52, 49.50

<table>
<thead>
<tr>
<th>PID</th>
<th>USER</th>
<th>PR</th>
<th>NI</th>
<th>VIRT</th>
<th>RES</th>
<th>SHR</th>
<th>S</th>
<th>%CPU</th>
<th>%MEM</th>
<th>TIME+</th>
<th>COMMAND</th>
</tr>
</thead>
<tbody>
<tr>
<td>1072</td>
<td>balewski</td>
<td>20</td>
<td>0</td>
<td>937m</td>
<td>643m</td>
<td>71m</td>
<td>R</td>
<td>100.0</td>
<td>0.5</td>
<td>23:07.02</td>
<td>root4star</td>
</tr>
<tr>
<td>1822</td>
<td>balewski</td>
<td>20</td>
<td>0</td>
<td>943m</td>
<td>649m</td>
<td>71m</td>
<td>R</td>
<td>100.0</td>
<td>0.5</td>
<td>22:19.00</td>
<td>root4star</td>
</tr>
<tr>
<td>3424</td>
<td>balewski</td>
<td>20</td>
<td>0</td>
<td>972m</td>
<td>677m</td>
<td>71m</td>
<td>R</td>
<td>100.0</td>
<td>0.5</td>
<td>22:55.48</td>
<td>root4star</td>
</tr>
<tr>
<td>1107</td>
<td>balewski</td>
<td>20</td>
<td>0</td>
<td>942m</td>
<td>653m</td>
<td>70m</td>
<td>R</td>
<td>99.7</td>
<td>0.5</td>
<td>40:38.25</td>
<td>root4star</td>
</tr>
<tr>
<td>1448</td>
<td>balewski</td>
<td>20</td>
<td>0</td>
<td>958m</td>
<td>670m</td>
<td>71m</td>
<td>R</td>
<td>99.7</td>
<td>0.5</td>
<td>23:10.59</td>
<td>root4star</td>
</tr>
<tr>
<td>2526</td>
<td>balewski</td>
<td>20</td>
<td>0</td>
<td>937m</td>
<td>642m</td>
<td>71m</td>
<td>R</td>
<td>100.0</td>
<td>0.5</td>
<td>23:02.94</td>
<td>root4star</td>
</tr>
<tr>
<td>30918</td>
<td>balewski</td>
<td>20</td>
<td>0</td>
<td>939m</td>
<td>645m</td>
<td>71m</td>
<td>R</td>
<td>94.5</td>
<td>0.5</td>
<td>30:26.97</td>
<td>root4star</td>
</tr>
<tr>
<td>31158</td>
<td>balewski</td>
<td>20</td>
<td>0</td>
<td>960m</td>
<td>669m</td>
<td>71m</td>
<td>R</td>
<td>94.5</td>
<td>0.5</td>
<td>23:00.85</td>
<td>root4star</td>
</tr>
<tr>
<td>31873</td>
<td>balewski</td>
<td>20</td>
<td>0</td>
<td>951m</td>
<td>660m</td>
<td>71m</td>
<td>R</td>
<td>94.5</td>
<td>0.5</td>
<td>39:59.76</td>
<td>root4star</td>
</tr>
<tr>
<td>32236</td>
<td>balewski</td>
<td>20</td>
<td>0</td>
<td>987m</td>
<td>690m</td>
<td>71m</td>
<td>R</td>
<td>94.5</td>
<td>0.5</td>
<td>23:51.42</td>
<td>root4star</td>
</tr>
</tbody>
</table>

7*60=420 root4star

Jan Balewski, NERSC
PDSF User Meeting
STAR : optimized DB use at cold start

Cold start of 300 1-core tasks
sbatch --array=1-300 my.Job.sl

Case 1:
  sleep $((\$RANDOM \% 60))
  root4star (bhla....)

Case 2:
  sleep $((\$RANDOM \% 600))
  root4star (bhla....)

SLURM+CHOS
SLURM+CHOS : 60 atlas jobs per node

```
mc1530 $ top ibn1

NOTE the CPU time spread despite common start

top - 20:51:04 up 1 day, 5:28, 1 user, load average: 60.04, 51.24, 34.47

<table>
<thead>
<tr>
<th>PID</th>
<th>USER</th>
<th>PR</th>
<th>NI</th>
<th>VIRT</th>
<th>RES</th>
<th>SHR</th>
<th>%CPU</th>
<th>%MEM</th>
<th>TIME+</th>
<th>COMMAND</th>
</tr>
</thead>
<tbody>
<tr>
<td>10886</td>
<td>shapiro</td>
<td>20</td>
<td>0</td>
<td>1549m</td>
<td>930m</td>
<td>5788</td>
<td>R</td>
<td>100.0</td>
<td>0.7</td>
<td>10:29.10 runLBLttHAnalys</td>
</tr>
<tr>
<td>11934</td>
<td>shapiro</td>
<td>20</td>
<td>0</td>
<td>1549m</td>
<td>930m</td>
<td>5788</td>
<td>R</td>
<td>100.0</td>
<td>0.7</td>
<td>10:41.61 runLBLttHAnalys</td>
</tr>
<tr>
<td>12082</td>
<td>shapiro</td>
<td>20</td>
<td>0</td>
<td>1549m</td>
<td>930m</td>
<td>5788</td>
<td>R</td>
<td>100.0</td>
<td>0.7</td>
<td>10:30.19 runLBLttHAnalys</td>
</tr>
<tr>
<td>13307</td>
<td>shapiro</td>
<td>20</td>
<td>0</td>
<td>1549m</td>
<td>931m</td>
<td>7004</td>
<td>R</td>
<td>100.0</td>
<td>0.7</td>
<td>4:07.18 runLBLttHAnalys</td>
</tr>
<tr>
<td>13384</td>
<td>shapiro</td>
<td>20</td>
<td>0</td>
<td>1549m</td>
<td>930m</td>
<td>5788</td>
<td>R</td>
<td>100.0</td>
<td>0.7</td>
<td>10:13.53 runLBLttHAnalys</td>
</tr>
<tr>
<td>16841</td>
<td>shapiro</td>
<td>20</td>
<td>0</td>
<td>1549m</td>
<td>930m</td>
<td>5788</td>
<td>R</td>
<td>100.0</td>
<td>0.7</td>
<td>10:09.00 runLBLttHAnalys</td>
</tr>
<tr>
<td>16886</td>
<td>shapiro</td>
<td>20</td>
<td>0</td>
<td>1549m</td>
<td>930m</td>
<td>5788</td>
<td>R</td>
<td>100.0</td>
<td>0.7</td>
<td>10:10.67 runLBLttHAnalys</td>
</tr>
<tr>
<td>17192</td>
<td>shapiro</td>
<td>20</td>
<td>0</td>
<td>1549m</td>
<td>930m</td>
<td>5788</td>
<td>R</td>
<td>100.0</td>
<td>0.7</td>
<td>10:05.95 runLBLttHAnalys</td>
</tr>
<tr>
<td>20043</td>
<td>shapiro</td>
<td>20</td>
<td>0</td>
<td>1549m</td>
<td>990m</td>
<td>65m</td>
<td>R</td>
<td>100.0</td>
<td>0.8</td>
<td>1:51.90 runLBLttHAnalys</td>
</tr>
<tr>
<td>20049</td>
<td>shapiro</td>
<td>20</td>
<td>0</td>
<td>1549m</td>
<td>990m</td>
<td>65m</td>
<td>R</td>
<td>100.0</td>
<td>0.8</td>
<td>1:59.63 runLBLttHAnalys</td>
</tr>
<tr>
<td>22992</td>
<td>shapiro</td>
<td>20</td>
<td>0</td>
<td>1549m</td>
<td>974m</td>
<td>49m</td>
<td>R</td>
<td>100.0</td>
<td>0.8</td>
<td>2:15.51 runLBLttHAnalys</td>
</tr>
<tr>
<td>29130</td>
<td>shapiro</td>
<td>20</td>
<td>0</td>
<td>1549m</td>
<td>932m</td>
<td>8268</td>
<td>R</td>
<td>100.0</td>
<td>0.7</td>
<td>3:10.89 runLBLttHAnalys</td>
</tr>
<tr>
<td>30492</td>
<td>shapiro</td>
<td>20</td>
<td>0</td>
<td>1549m</td>
<td>930m</td>
<td>5788</td>
<td>R</td>
<td>100.0</td>
<td>0.7</td>
<td>10:35.90 runLBLttHAnalys</td>
</tr>
<tr>
<td>30698</td>
<td>shapiro</td>
<td>20</td>
<td>0</td>
<td>1549m</td>
<td>983m</td>
<td>58m</td>
<td>R</td>
<td>100.0</td>
<td>0.8</td>
<td>2:05.52 runLBLttHAnalys</td>
</tr>
<tr>
<td>32417</td>
<td>shapiro</td>
<td>20</td>
<td>0</td>
<td>1549m</td>
<td>983m</td>
<td>58m</td>
<td>R</td>
<td>100.0</td>
<td>0.8</td>
<td>2:10.79 runLBLttHAnalys</td>
</tr>
<tr>
<td>5696</td>
<td>shapiro</td>
<td>20</td>
<td>0</td>
<td>1549m</td>
<td>1.0g</td>
<td>100m</td>
<td>R</td>
<td>100.0</td>
<td>0.8</td>
<td>1:58.20 runLBLttHAnalys</td>
</tr>
</tbody>
</table>
```
user multi-thread jobs can coexist

Atlas (6-thread Madgraph)

[kkrizka@pdsf6]$ sbatch --array=1-42 oneMG.slr

#SBATCH --mem 10G --ntasks=6
export WORKDIR=/global/project/projectdirs/atlas/kkrizka/janTmp/job${SLURM_JOBID}
mkdir -p $WORKDIR
cd ${WORKDIR}
sleep $(($RANDOM % 300 ))
cp /global/project/projectdirs/atlas/kkrizka/PROC_xia.tgz .
tar -zxf PROC_xia.tgz
CHOS=sl64 chos  $CODE_DIR/launch.sh

STAR (1-thread root analysis)

sbatch --array 1-500 starOne.slr

#SBATCH --ntasks=1
sleep $(($RANDOM % 300 ))
CHOS=sl64 chos $CODE_DIR/r4sTask_bfc.csh

SLURM+CHOS

Jan Balewski, NERSC
PDSF User Meeting 2016
SLURM : (bad) use patterns

$sbatch --array=1-500 myJob.slr

$cat Task.sh

# goto local scratch - good
cd $SLURM_TMP

# very bad:
cp -rp ~/codeDir .

# better:
cp /project/xyz/aa.tgz .
tar -zxf aa.tgz

# this will help to scatter at cold-start
sleep $((RANDOM % 30))

# never write to /tmp
SLURM - observed errors related to stuck IO

---
FAILED to unmount old image in this namespace, cannot continue.
FAILED to setup image.

---
FAILED to mount image /chos/common/nsg/images/66b12b5e69ca115ad8b792.squashfs (squashfs) on /var/udiLoopMount
FAILED to loop mount image: /chos/common/nsg/images/66b12b5e20f15ad8b792.squashfs

---
File "/usr/lib64/python2.6/threading.py", line 474, in start _start_new_thread(self.__bootstrap, ())
thread.error: can't start new thread

---
The compilation fails with the following output message:

---
cp: writing `/scratch/tmp.20588.0/job20588/89/unweighted_events.lhe.gz': Cannot allocate memory

---
Site pre-mount hook failed. Exiting.
FAILED to properly setup site modifications

---
The culprint is 'D'

<table>
<thead>
<tr>
<th>PID</th>
<th>USER</th>
<th>PR</th>
<th>NI</th>
<th>VIRT</th>
<th>RES</th>
<th>SHR</th>
<th>S</th>
<th>%CPU</th>
<th>%MEM</th>
<th>TIME+</th>
<th>COMMAND</th>
</tr>
</thead>
<tbody>
<tr>
<td>11284</td>
<td>xxxxx</td>
<td>20</td>
<td>0</td>
<td>111m</td>
<td>1052</td>
<td>752</td>
<td>D</td>
<td>0.0</td>
<td>0.0</td>
<td>0:06.00</td>
<td>cp</td>
</tr>
</tbody>
</table>
SLURM RAM/core ratio is 2 GB/task

#!/bin/bash
#SBATCH -t 55:00 --ntasks=1 --account atlas
#SBATCH --mem 3008m ⇐ # GB/task
#SBATCH -J atlas-chos -p shared-chos

pdsf8 $ sbatch -a 1-480 atlasOne.slr

Ticket INC0103391:
SLURM on PDSF is not accounting for available swap space

pdsf8 $ scontrol show job 28547_300
   TRES=cpu=1,mem=3008,node=1
  MinCPUsNode=2 MinMemoryNode=3008M
MinTmpDiskNode=0

Math:
120 GB RAM/60 job-slots=2GB/task
7 nodes * 30 task=210
## SLURM - solved issues by CSG

<table>
<thead>
<tr>
<th>Issue</th>
<th>Priority 1=high</th>
<th>Impairies</th>
<th>Experiment</th>
<th>status</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Local SLURM_TMP</td>
<td>1</td>
<td>SLURM</td>
<td>ALICE, ATLAS, STAR</td>
<td>Implemented Jun 7, works</td>
<td>SLURM jobs running on PDSF should have local $SCRACH space reserved for each job and free after a job completes regardless if job was successful or crashed</td>
</tr>
<tr>
<td>Scalable CVMFS</td>
<td>3</td>
<td>Shifter</td>
<td>ALICE, ATLAS, LZ?</td>
<td>Impl. Jun 7, works</td>
<td>The limit on number of concurrent CVMFS connections should be 64/node, we have few/node, related to loopback limit</td>
</tr>
</tbody>
</table>

### Technical issues resolved:
- OK - Madgraph jobs run at scale on slurm+shifter make system erratic, ~half of jobs can't start, ticket: INC0103239 ⇐ see INC0103537
- Ok INC0103389, 2 slurm 32-core nodes do not accept any jobs if '-n2', mc1531 (chos partition), mc1528 (shifter partition)
- OK- INC0103453: zoo of slurm+shifter job crashes on PDSF ⇐ see INC0103537
- OK - INC0103455: node mc1526+mc1504 can't handle 60 cp operations from /global/home to $SLURM_TMP ⇐ see INC0103537
- OK - INC0103537 : SLURM+CHOS, exactly the same cp in D issue fro slurm+chos partition <= too much IO, reduce task/node and pread
- OK - INC0103567 : SLURM+CHOS, 90/210 jobs crash jobs trashed due to: slurm_script: fork: retry: Resource temporarily unavailable<= too much IO, reduce task/node and pread
- OK -INC0103754 : SLURM+CHOS, worker nodes in drained state, after heavy use <= OOM cause by malloc(10GB)
# SLURM - outstanding high priority issues

<table>
<thead>
<tr>
<th>Issue</th>
<th>Priority 1=high</th>
<th>Imparies</th>
<th>Experiment</th>
<th>status</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ganglia</td>
<td>2</td>
<td>SLURM</td>
<td>all</td>
<td>Known cause</td>
<td>Jobs running under SLURM are not reported on ganglia PDSF monitor page, ticket INC0103269</td>
</tr>
<tr>
<td>Ignored swap space</td>
<td>2</td>
<td>SLURM</td>
<td>all</td>
<td>discussed</td>
<td>SLURM is not seeing swap space on nodes, either partition (chos/shifter) INC0103391</td>
</tr>
<tr>
<td>OSG2</td>
<td>2</td>
<td>SLURM+grid jobs</td>
<td>ALICE, some STAR</td>
<td>Under eval.</td>
<td>deployment of 2nd OSG talking to SLURM on node mpdfsgrid02, will look like ...01. Jeff &amp; James</td>
</tr>
<tr>
<td>AliEn VOBox2</td>
<td>2</td>
<td>SLURM</td>
<td>ALICE</td>
<td>In progress</td>
<td>grid jobs submission, monitoring. Need 2nd system for SLURM</td>
</tr>
<tr>
<td>Bad tcsh</td>
<td>3</td>
<td>Shifter</td>
<td>STAR</td>
<td>Under eval.</td>
<td>SLURM + Shifter is not working on PDSF for users who have the default NIM shell set to tcsh, ticket INC0101446</td>
</tr>
<tr>
<td>Smooth image selection</td>
<td>3</td>
<td>shifter</td>
<td>all</td>
<td>Under eval.</td>
<td>A user logs into PDSF and their shell is run inside the preselected Shifter image without any action on their part.</td>
</tr>
</tbody>
</table>