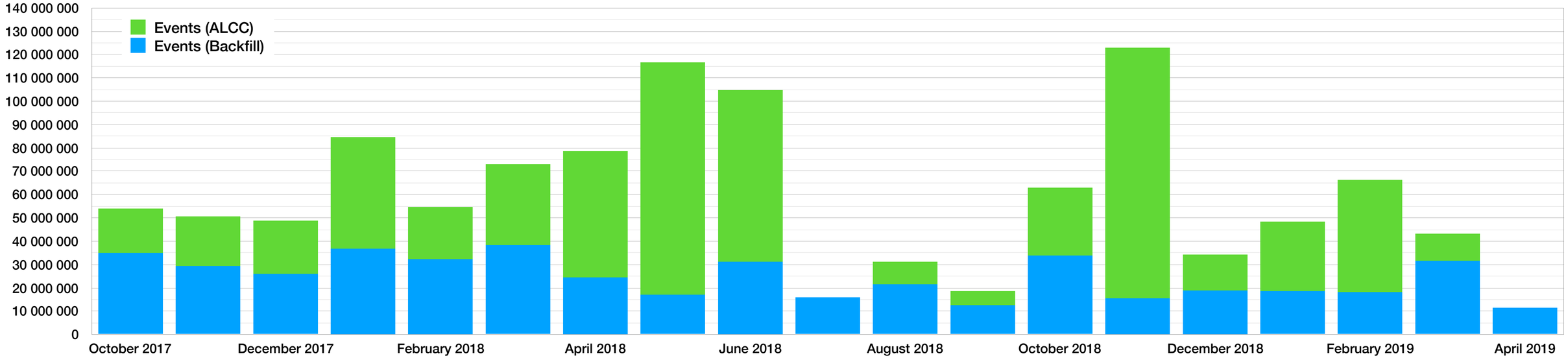
ATLAS @ OLCF

Danila Oleynik

BigPanDA Technical Interchange Meeting.

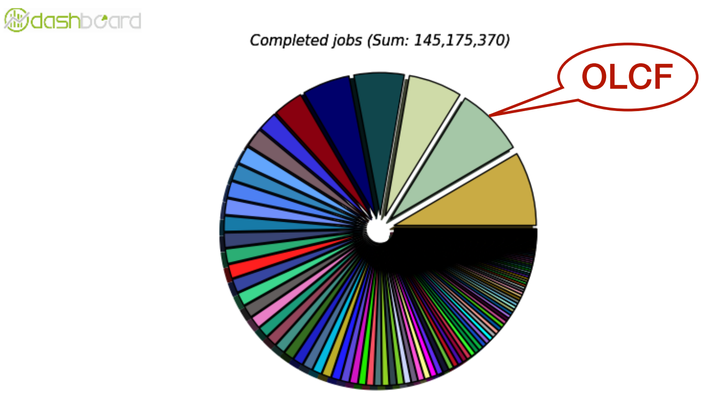
Apr. 24, 2019 BNL

Overview of ATLAS production at OLCF



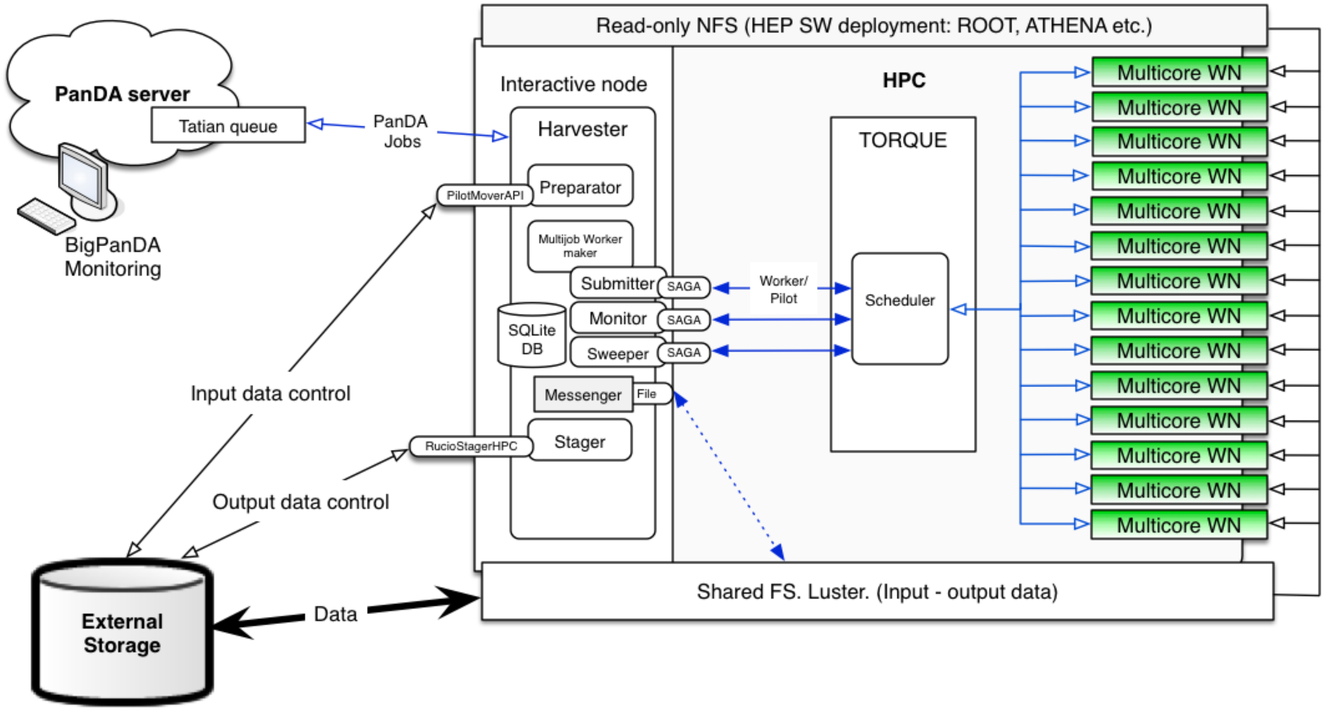
- In the period from October 2017 till the end of April 2019 about 70B events were simulated by ATLAS computing infrastructure. **1.1B events were simulated by OLCF (~1,5% of all simulations, in Top 20 sites)**.

- ~145M simulations jobs were completed through PanDA WMS during this period. ~11,3M jobs completed in OLCF: 7.8% from all sim jobs; Site #2 by completed sim jobs in ATLAS (CERN-PROD #1 with 8.4%)



PanDA components deployment at OLCF

- The newest generation of PanDA components are deployed at OLCF:
  - **Harvester** - resource-facing service deployed on the edge node of HPC, to take care about:
    - intercommunication with PanDA and LRMS
    - Data Management: pre-staging of input data and stage-out of output
  - Harvester extended with the possibility of the shaping of payload according to the available resources (“backfill”)
  - **Pilot2** takes care of the execution of particular PanDA job.
    - Pilot2 uses "HPC workflow” plug-in to perform execution without external network connectivity
    - Some, Titan related functionality are placed to corresponded plug-in



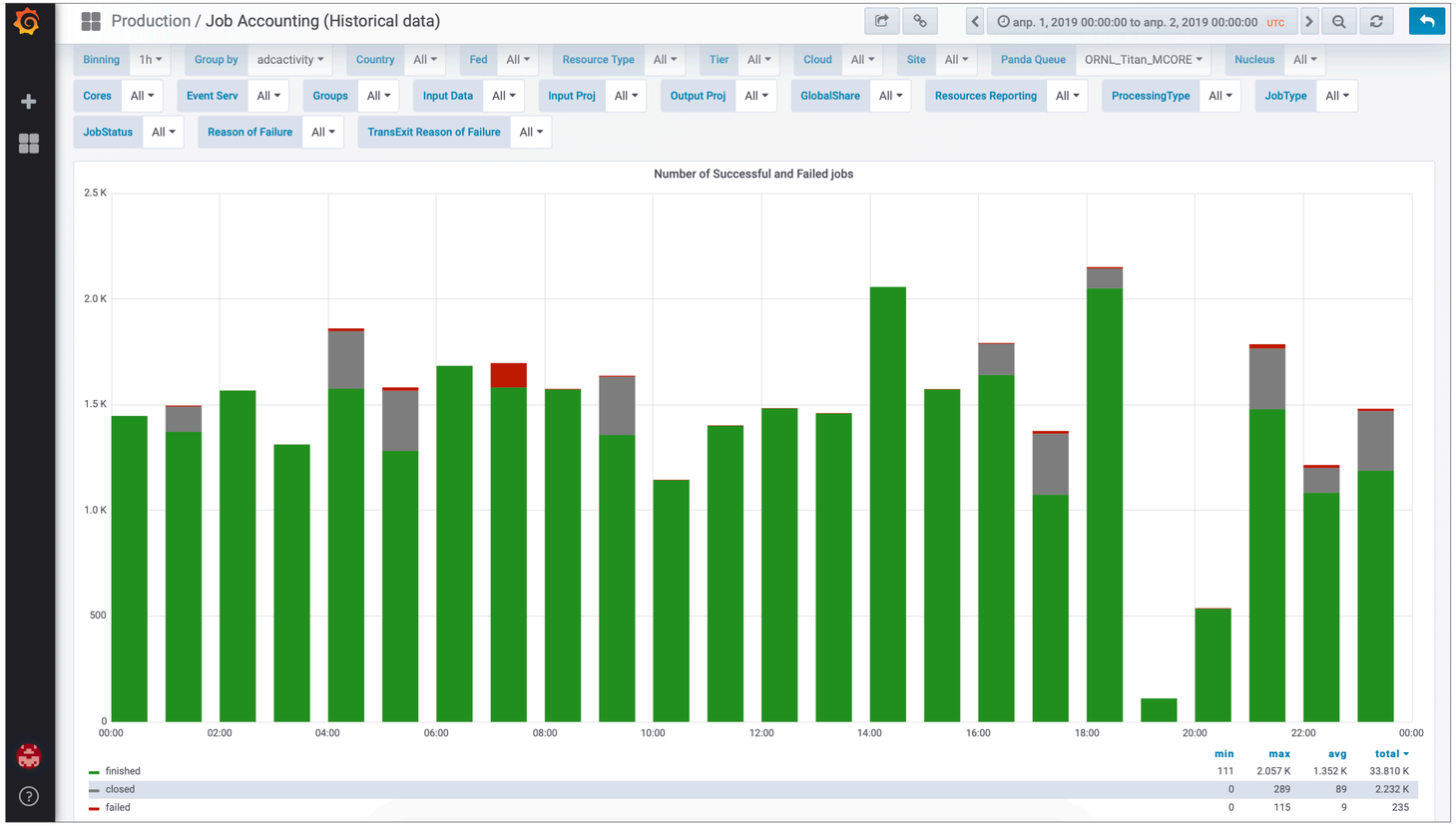
- Two Harvester instances at OLCF deployed on interactive DTNs (DTN35, DTN38)
  - One instance serves for ‘backfill’
  - Another serves for ALCC allocation

ATLAS SW deployment at OLCF

- We did not reach success with using of CVMFS on Titan. Size of ATLAS repository works against on-time synchronisation with shared FS at OLCF
- On first steps of production of ATLAS data at OLCF was discovered, that architecture of ATLAS SW is quite IO intensive. Usage of a lot of Python modules causes high metadata loading on startup)
  - ATLAS repository was placed to NFS, to avoid degradation of Lustre even on a small scale. We still work on a small scale, since unpleased IO were observed on NFS too
- Another option was using of Singularity containers with ATLAS software, but Singularity was not stable enough to work on production scale on Titan
  - We very hope to have better support of Singularity on Summit

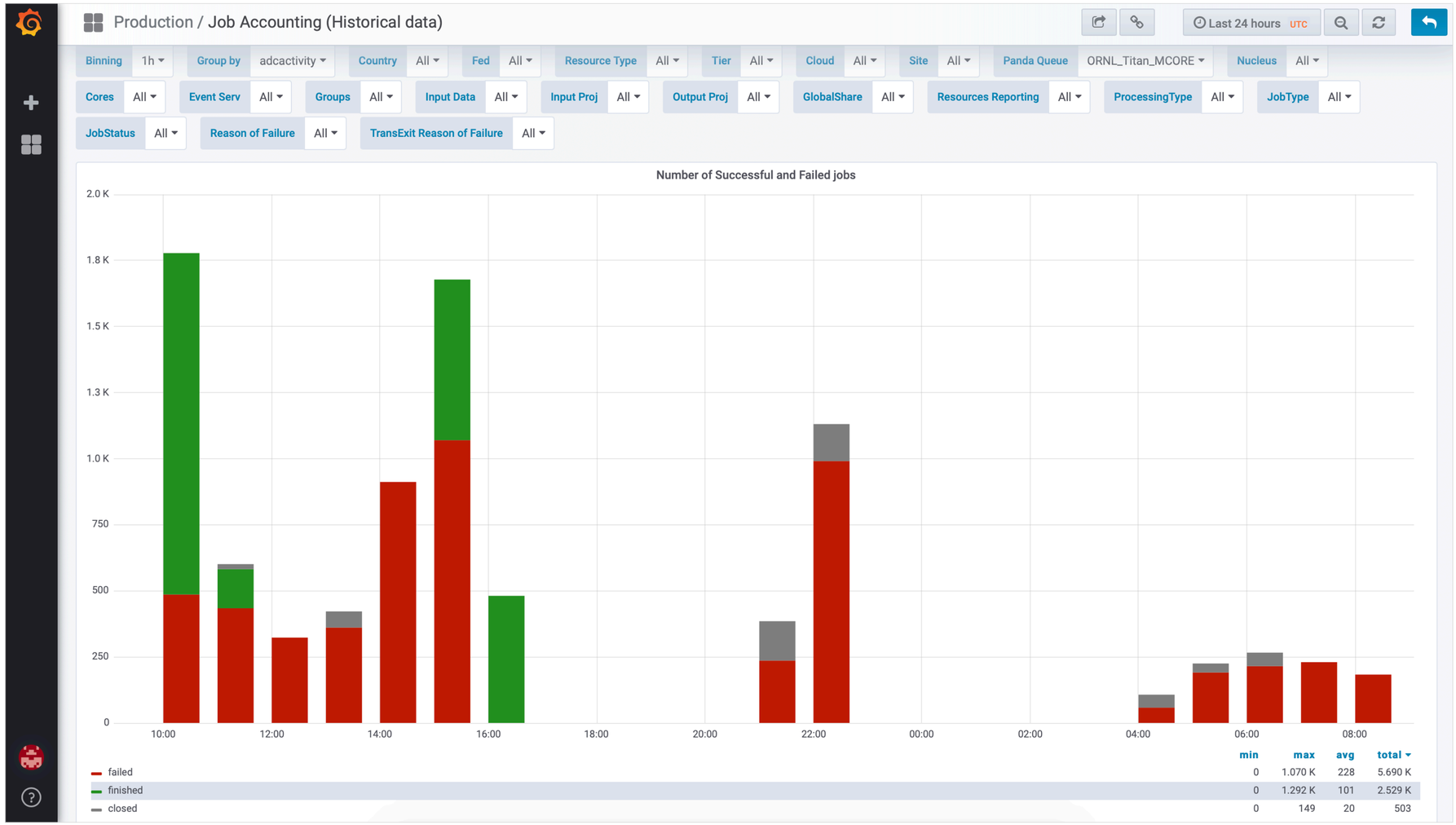
ATLAS operations at OLCF

- All middleware (PanDA stack) and ATLAS SW works very stable at OLCF.
- PanDA monitoring tools provide enough information about current state of the components. Accounting tools allows to understand bebaiovour over time.



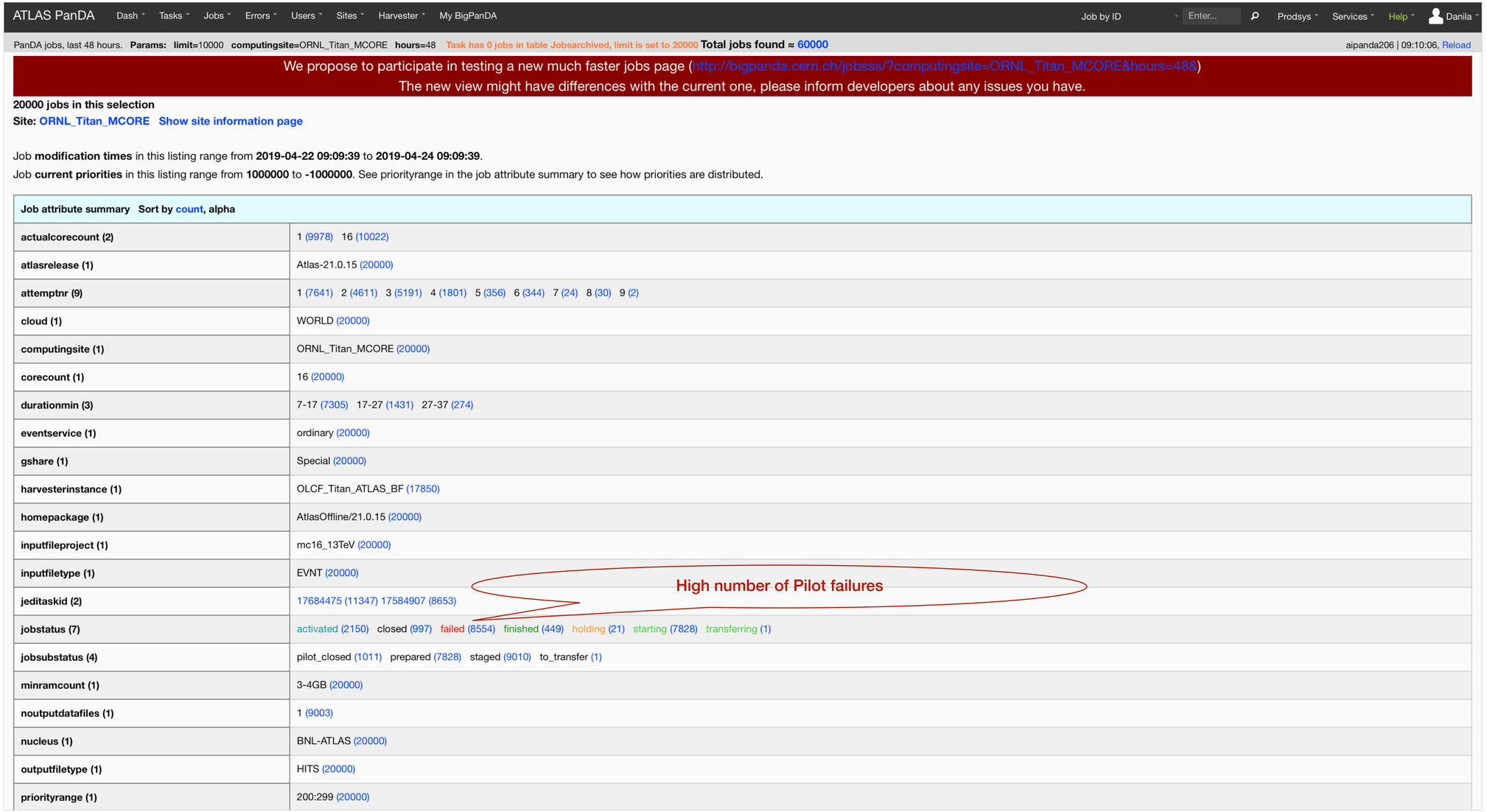
- Regular day on Titan, a lot of ‘green’, less red and grey: https://monit-grafana.cern.ch/dashboard/snapshot/4m5XNPKmSRMyuzC6LdXkOrFtUa5jmOWD

Bad day on Titan



- Bad day on Titan, too much ‘red’. Feature investigation required -> BigPanDA monitor

Bad day on Titan: site status overview



ATLAS PanDA Dash Tasks Jobs Errors Users Sites Harvester My BigPanDA Job by ID Enter... Prodsys Services Help Danila

PanDA jobs, last 48 hours. **Params:** **limit**=10000 **computingsite**=ORNL_Titan_MCORE **hours**=48 Task has 0 jobs in table Jobsarchived, limit is set to 20000 **Total jobs found ≈ 60000** aipanda206 | 09:10:06, Reload

We propose to participate in testing a new much faster jobs page (http://bigpanda.cern.ch/jobsss/?computingsite=ORNL_Titan_MCORE&hours=48&)
The new view might have differences with the current one, please inform developers about any issues you have.

20000 jobs in this selection
Site: ORNL_Titan_MCORE **Show site information page**

Job **modification times** in this listing range from **2019-04-22 09:09:39** to **2019-04-24 09:09:39**.
Job **current priorities** in this listing range from **1000000** to **-1000000**. See priorityrange in the job attribute summary to see how priorities are distributed.

Job attribute summary **Sort by** count, alpha	
actualcorecount (2)	1 (9978) 16 (10022)
atlasrelease (1)	Atlas-21.0.15 (20000)
attemptnr (9)	1 (7641) 2 (4611) 3 (5191) 4 (1801) 5 (356) 6 (344) 7 (24) 8 (30) 9 (2)
cloud (1)	WORLD (20000)
computingsite (1)	ORNL_Titan_MCORE (20000)
corecount (1)	16 (20000)
durationmin (3)	7-17 (7305) 17-27 (1431) 27-37 (274)
eventservice (1)	ordinary (20000)
gshare (1)	Special (20000)
harvesterinstance (1)	OLCF_Titan_ATLAS_BF (17850)
homepackage (1)	AtlasOffline/21.0.15 (20000)
inputfileproject (1)	mc16_13TeV (20000)
inputfiletype (1)	EVNT (20000)
jeditaskid (2)	17684475 (11347) 17584907 (8653)
jobstatus (7)	activated (2150) closed (997) failed (8554) finished (449) holding (21) starting (7828) transferring (1)
jobsubstatus (4)	pilot_closed (1011) prepared (7828) staged (9010) to_transfer (1)
minramcount (1)	3-4GB (20000)
noutputdatafiles (1)	1 (9003)
nucleus (1)	BNL-ATLAS (20000)
outputfiletype (1)	HITS (20000)
priorityrange (1)	200:299 (20000)

High number of Pilot failures

- BigPanDA monitor. 'Site' status overview: https://bigpanda.cern.ch/jobs/?computingsite=ORNL_Titan_MCORE&hours=48&display_limit=100
- High number of Pilot failures discovered -> Harvester instances should be monitored

Bad day on Titan: Harvester instance overview

PanDA monitor | Dash | Tasks | Jobs | Errors | Users | Sites | Harvester | My BigPanDA | Job by ID | Enter... | Prodsys | Services | Help | Danila

Harvester workers, last 24 hours — aipanda205 | 09:11:02, Reload

Commit stamp	28-03-2019 15:49:37 on OLCF_validation (by DanilaOleynik)
Computingelements (1)	null (132)
Computingsites (1)	ORNL_Titan_MCORE (132)
Description	test harvester instance at Titan
HarvesterID	OLCF_Titan_ATLAS_BF
Hostname	None
Jobs count	9417
Lastupdate	24-04-2019 09:05:34
Owner	Danila Oleynik
Resourcetypes (1)	null (132)
Software version	0.0.30
Starttime	17-04-2019 19:24:58
Statuses (3)	finished (18) failed (106) cancelled (8)

High number of failed workers. Problem dose not relate with ATLAS SW

Workers | Dialog messages | Worker stats | Jobs

Number of entries: 1000 Reload table

Show 20 entries — Search:

Workers submitted by harvesters

WorkerID	Jobs	Last update	Status	BatchID	NodeID	Queuename	Computingsite	Instance	Submit time	Start time	End time	Ncore	Error code	Stdout	Stderr	Batchlog	Resource type	Nativeexitcode	Native status	Diag message	Computing element
13048	72	2019-04-24 08:55:52	failed	4676342			ORNL_Titan_MCORE	OLCF_Titan_ATLAS_BF	2019-04-24 08:43:06	2019-04-24 08:43:19	2019-04-24 08:53:38	1152		Link	Link	-		271	Vacated	Job stopped by user (SIGTERM)	
13047	33	2019-04-24 08:55:52	failed	4676339			ORNL_Titan_MCORE	OLCF_Titan_ATLAS_BF	2019-04-24 08:35:41	2019-04-24 08:35:52	2019-04-24 08:53:26	528		Link	Link	-		271	Vacated	Job stopped by user (SIGTERM)	
13046	21	2019-04-24 08:55:28	failed	4676337			ORNL_Titan_MCORE	OLCF_Titan_ATLAS_BF	2019-04-24 08:31:29	2019-04-24 08:31:38	2019-04-24 08:53:26	336		Link	Link	-		271	Vacated	Job stopped by user (SIGTERM)	
13045	75	2019-04-24 08:32:26	failed	4676333			ORNL_Titan_MCORE	OLCF_Titan_ATLAS_BF	2019-04-24 08:15:48	2019-04-24 08:16:00	2019-04-24 08:30:44	1200		Link	Link	-		271	Vacated	Job stopped by user (SIGTERM)	

Error code, Native status and Diag message corresponded to preempted batch job

- BigPanDA monitor. Harvester instance overview: https://bigpanda.cern.ch/harvesters/?instance=OLCF_Titan_ATLAS_BF

Bad day on Titan: Harvester instance overview

PanDA monitor | Dash | Tasks | Jobs | Errors | Users | Sites | Harvester | My BigPanDA | Job by ID | Enter... | Prodsys | Services | Help | Danila

Harvester workers, last 24 hours — aipanda205 | 09:11:02, Reload

Commit stamp	28-03-2019 15:49:37 on OLCF_validation (by DanilaOleynik)
Computingelements (1)	null (132)
Computingsites (1)	ORNL_Titan_MCORE (132)
Description	test harvester instance at Titan
HarvesterID	OLCF_Titan_ATLAS_BF
Hostname	None
Jobs count	9417
Lastupdate	24-04-2019 09:05:34
Owner	Danila Oleynik
Resourcetypes (1)	null (132)
Software version	0.0.30
Starttime	17-04-2019 19:24:58
Statuses (3)	finished (18) failed (106) cancelled (8)

High number of failed workers. Problem does not relate with ATLAS SW

Workers | Dialog messages | Worker stats | Jobs

Number of entries: 1000 Reload table

Show 20 entries — Search:

Workers submitted by harvesters

WorkerID	Jobs	Last update	Status	BatchID	NodeID	Queuename	Computingsite	Instance	Submit time	Start time	End time	Ncore	Error code	Stdout	Stderr	Batchlog	Resource type	Nativeexitcode	Native status	Diag message	Computing element
13048	72	2019-04-24 08:55:52	failed	4676342			ORNL_Titan_MCORE	OLCF_Titan_ATLAS_BF	2019-04-24 08:43:06	2019-04-24 08:43:19	2019-04-24 08:53:38	1152		Link	Link	-		271	Vacated	Job stopped by user (SIGTERM)	
13047	33	2019-04-24 08:55:52	failed	4676339			ORNL_Titan_MCORE	OLCF_Titan_ATLAS_BF	2019-04-24 08:35:41	2019-04-24 08:35:52	2019-04-24 08:53:26	528		Link	Link	-		271	Vacated	Job stopped by user (SIGTERM)	
13046	21	2019-04-24 08:55:28	failed	4676337			ORNL_Titan_MCORE	OLCF_Titan_ATLAS_BF	2019-04-24 08:31:29	2019-04-24 08:31:38	2019-04-24 08:53:26	336		Link	Link	-		271	Vacated	Job stopped by user (SIGTERM)	
13045	75	2019-04-24 08:32:26	failed	4676333			ORNL_Titan_MCORE	OLCF_Titan_ATLAS_BF	2019-04-24 08:15:48	2019-04-24 08:16:00	2019-04-24 08:30:44	1200		Link	Link	-		271	Vacated	Job stopped by user (SIGTERM)	

Error code, Native status and Diag message correspond to preemped batch job

- BigPanDA monitor. Harvester instance overview: https://bigpanda.cern.ch/harvesters/?instance=OLCF_Titan_ATLAS_BF

Bad day on Titan: Harvester worker

PanDA monitor | Dash | Tasks | Jobs | Errors | Users | Sites | Harvester | My BigPanDA | Job by ID | Enter... | Prodsys | Services | Help | Danila

PanDA Harvester Workers Summary — aipanda207 | 09:24:09, Reload

Harvester worker info

Name	Value
harvesterid	OLCF_Titan_ATLAS_BF
workerid	13028
lastupdate	2019-04-24 06:23:54
status	failed
batchid	4676287
nodeid	
queuename	
computingsite	ORNL_Titan_MCORE
submittime	2019-04-24 06:09:25
starttime	2019-04-24 06:09:42
endtime	2019-04-24 06:23:05
ncore	256
errorcode	None
stdout	https://aipanda011.cern.ch:25443/cache/OLCF_Titan_ATLAS_BF__MPI_pilot_stdout_6890320
stderr	https://aipanda011.cern.ch:25443/cache/OLCF_Titan_ATLAS_BF__MPI_pilot_stderr_6890320
batchlog	
resourcetype	
nativeexitcode	271
nativestatus	Vacated
diagmessage	Job stopped by user (SIGTERM)
computingelement	
njobs	16
Correspondent Jobs	4321369606 4321369616 4321369618 4321369619 4321369604 4321369615 4321369608 4321369610 4321369613 4321369617 4321369612 4321369609 4321369614 4321369605 4321369611 4321369607
jobstatus	failed (16)
jobsubstatus	staged (16)

- BigPanDA monitor. Worker overview: https://bigpanda.cern.ch/harvesterworkerinfo/?harvesterid=OLCF_Titan_ATLAS_BF&workerid=13028

ATLAS @ OLCF Summary

- Collaboration with OLCF very valuable for ATLAS and ADC:
  - Pioneer project for integration of supercomputer into the distributed computing environment
  - Significant deliveries for the experiment
  - Ignition for enhancement of PanDA WMS and ATLAS Software
- Understanding of behaviour and specialities of HPC facilities
- High level of integration allows using the same operational tools, like ordinary grid sites
  - Expert knowledge is needed only for non-routine situations and development

BackUp

AES @ Titan

- ATLAS event service middleware Yoda was validated on Titan.
  - Harvester managed Yoda, and took care of some parts of post-processing
  - IO intensive workflow does not allow to scale up significantly
  - Any activity was suspended a couple of months ago in favour of stable delivery from regular jobs and low failure rate