

MultiJob pilot on Titan. ATLAS workloads on Titan

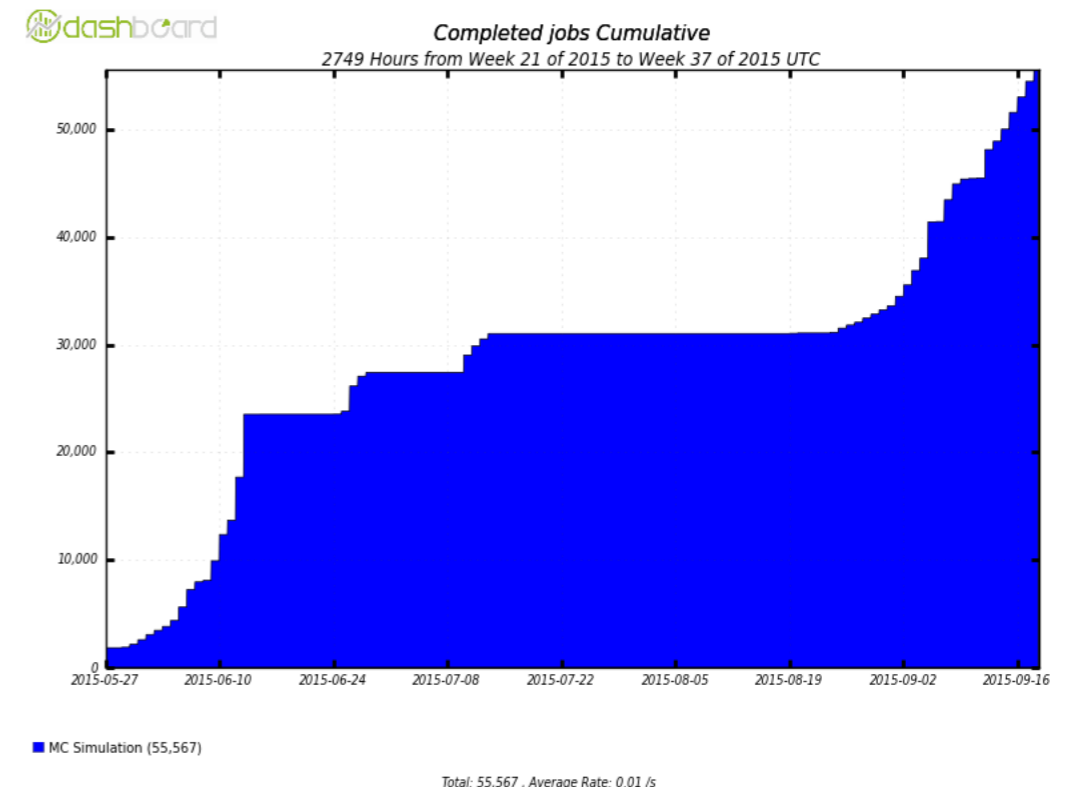
Danila Oleynik (UTA), Sergey Panitkin (BNL)
US ATLAS HPC. Technical meeting
18 September 2015

Overview

- Current status
- Workflow and issues
- ATLAS workloads on Titan

MultiJob Pilot on Titan

- First version deployed at the end of May 2015
- Validation starts June
 - ~50 000 Completed jobs
 - 2 simultaneous pilots
 - Only backfill mode



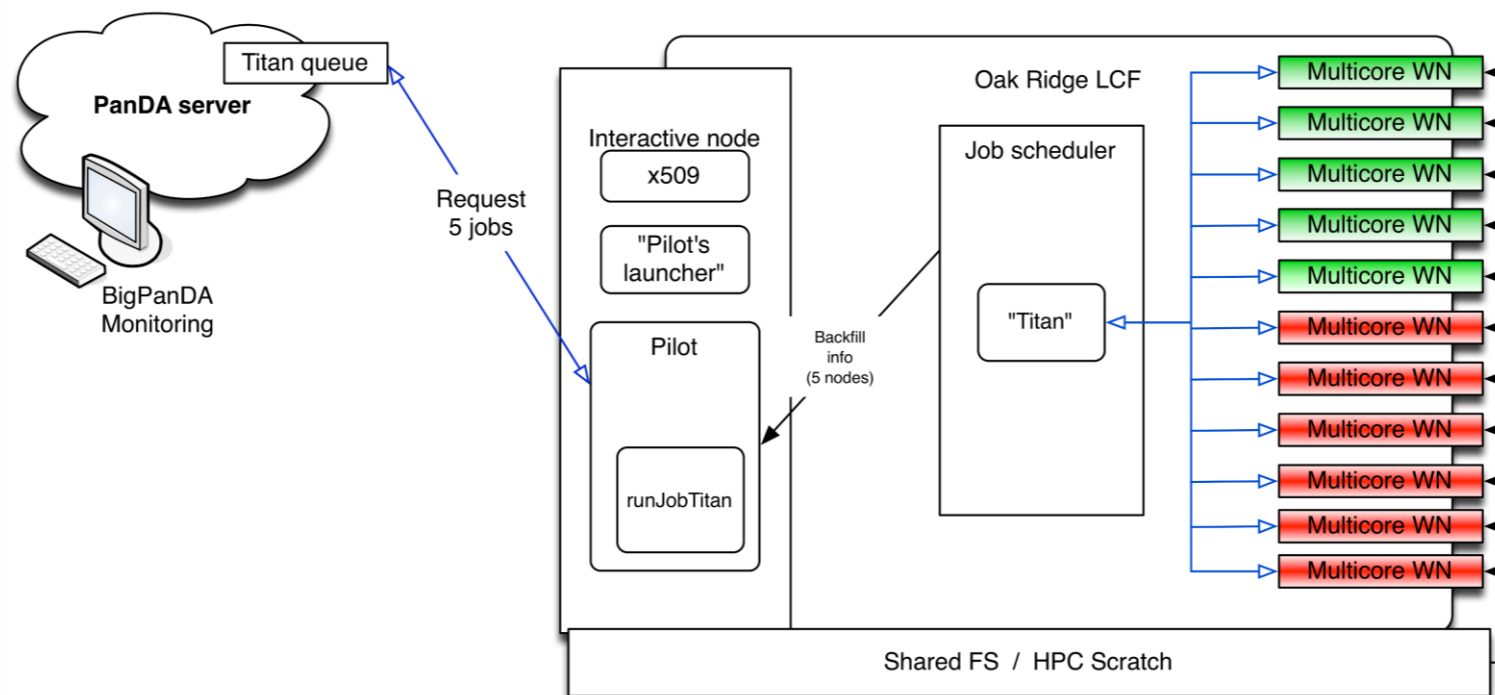
2015-06	754172	Increasing loading from 15 to 300 PanDA jobs in bunch.
2015-07	146250	Have no steady stream of jobs form PanDA
2015-08	114474	Have no steady stream of jobs form PanDA
2015-09	697337 824668	Steady stream of jobs from 15 to 300 jobs in bunch, two Pilots.

How to be efficient with backfill

Be fast and smart enough to grab resources, before other ~~bastards~~ users steal them and while time gap is enough for execution of payload.

- Minimize start up time as much as possible
- Minimize post processing procedures – to be ready for next round

Workflow & Issues 1

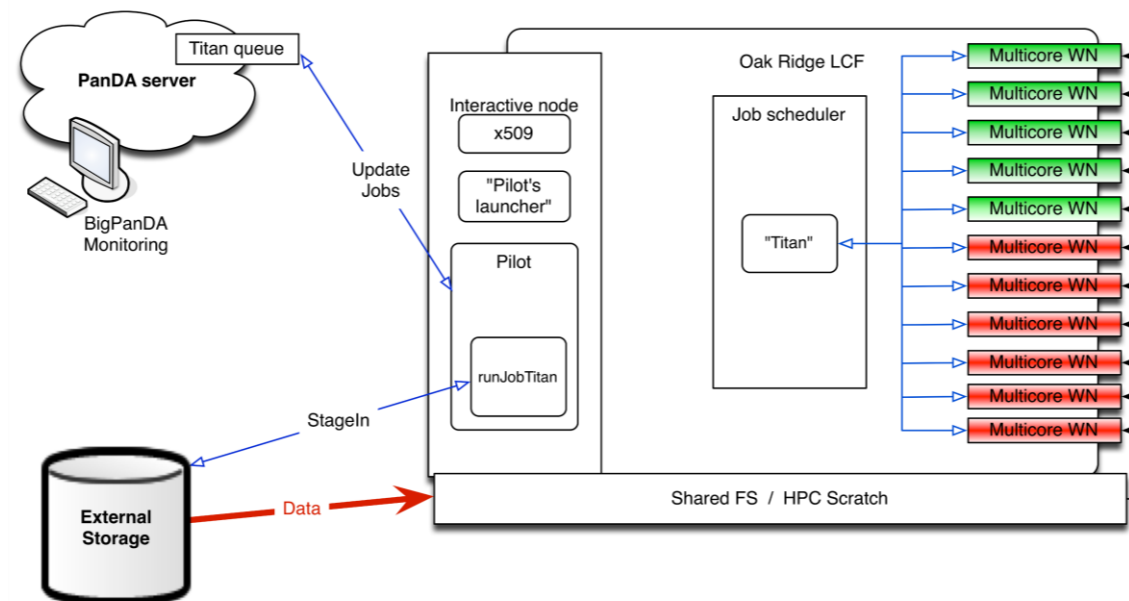


Pilot requests information about available resources (backfill)

According to available resources, pilot requests jobs from PanDA server

- Sequential requests works good for small chunks (one request took about 1 sec.), but for chunks with few hundred jobs it take significant time.
- Simultaneous requests gave about 10-15% of benefit due to execution time each particular request increasing in depending of number of parallel process
- Finally bulk function on PanDA server was implemented - now request for 300 jobs tooks about 15-20 sec.

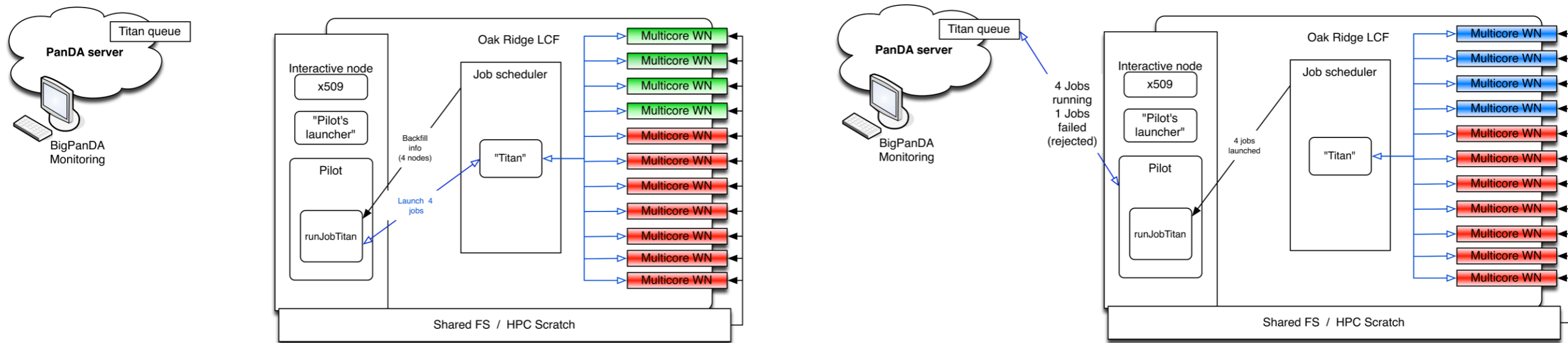
Workflow & Issues 2



Pilot stage-in and prepares environment before launching Jobs updates with initial info (space report etc.)

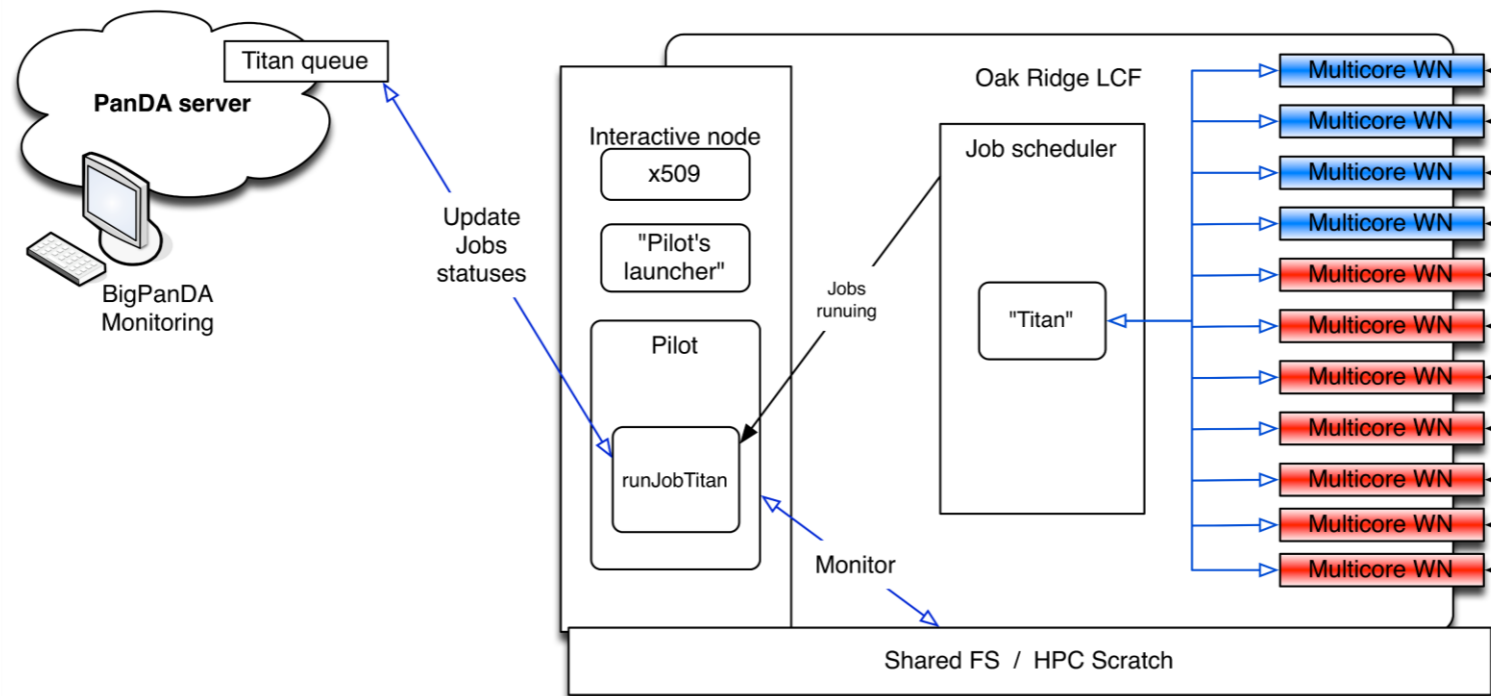
- Stage-in of one file take ~15-20 seconds;
- Stage-in procedure was modified to decrease number of real remote downloads. It takes now for 300 jobs about 300 seconds, it's very good, but will be great to decrease this time as well.
- Since update of one job on PanDA server consume about 1 sec. I try to minimize number of updates of PanDA server before job really launched. For example, 'space report' was disabled - and now this information come with regular job state update.

Workflow & Issues 3



Request available backfill and adjust number of jobs for launching.
Not fitted jobs will be failed and sub status of jobs will be updated, as only MPI job will be launched.

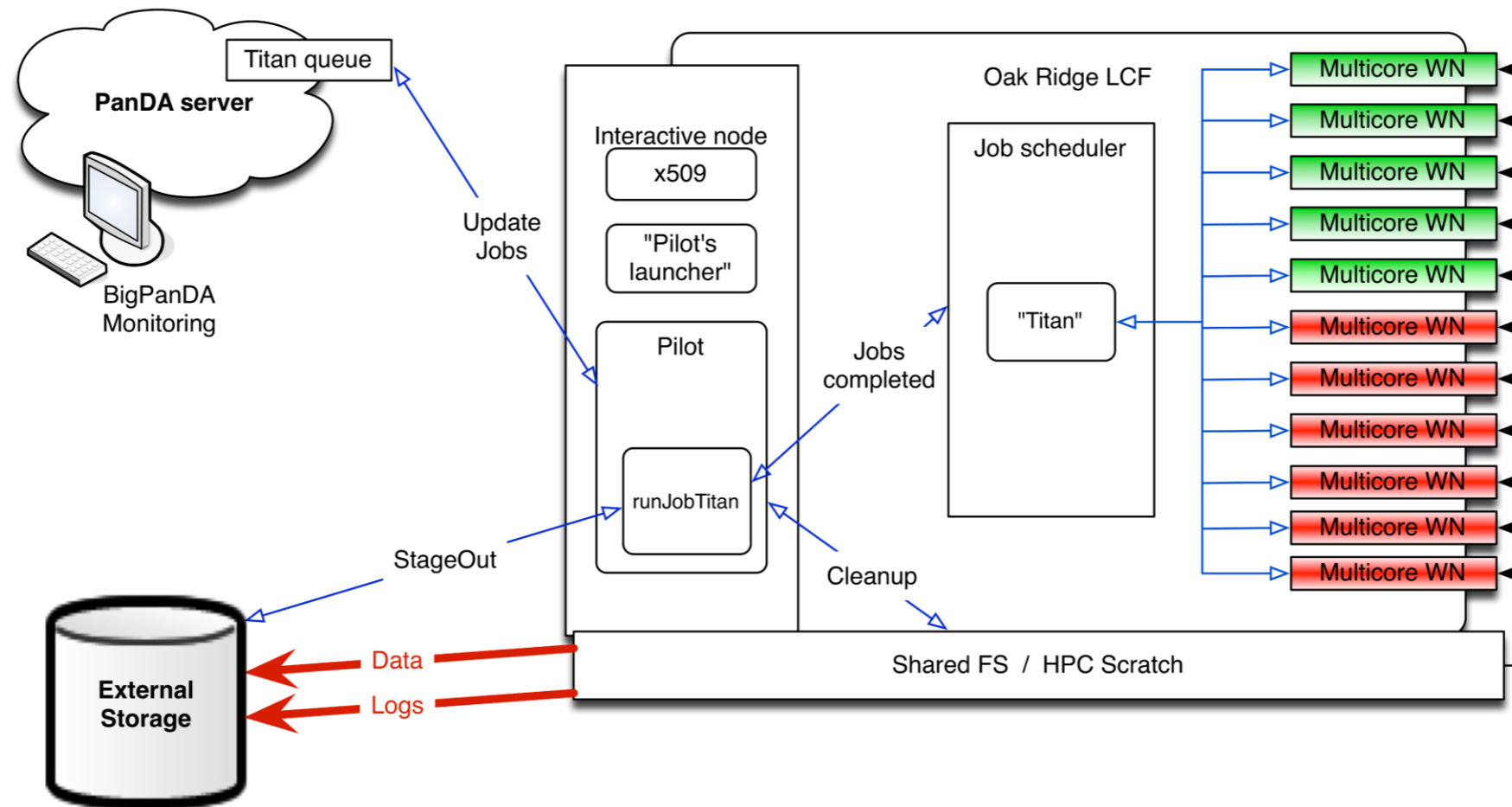
Workflow & Issues 4



Pilot monitors available space, output and sends heartbeat during payload execution

- Procedures, related with checking of space and outputs should be reviewed to avoid significant impacts to FS.

Workflow & Issues 5



Jobs completed, pilot perform stage-out and cleanup... and it may take up to few hours

Conclusion about workflow

We still got 'surprises' related with scalability.

There are still a lot of room for improvements, some of them may be achieved by optimization of code of Pilot, others by extension of architecture for example optimization of data management, pre-staging of data etc.

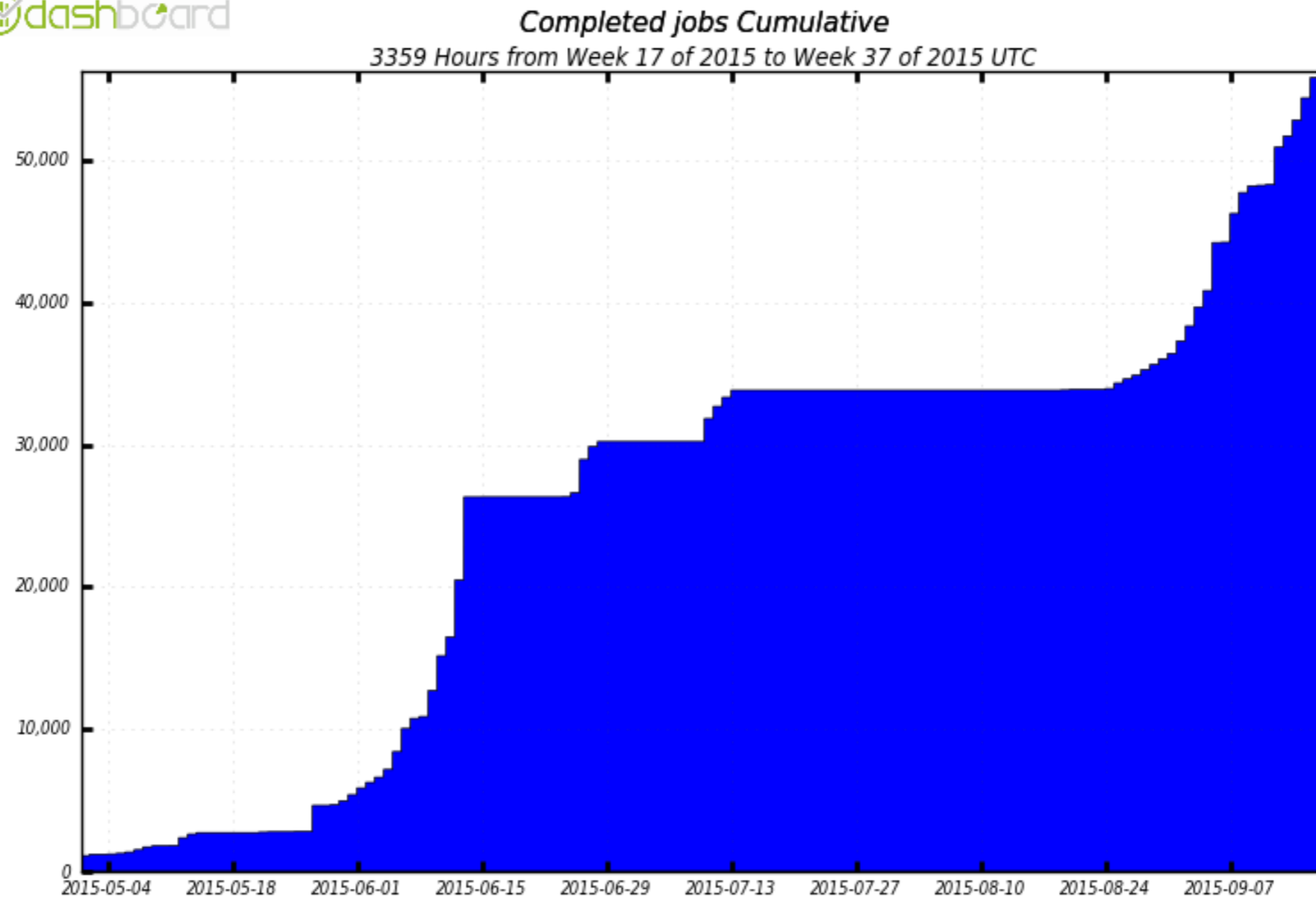
Update on Titan workloads

Sergey Panitkin

Validation of the data produced on Titan

- We ran validation jobs since May
 - Started with Rel. 19.2.3.8 based tasks
 - A set of physics validation pages generated
 - http://pagacova.web.cern.ch/pagacova/Technical_validation/BNLvsTitan/
 - No conclusive message. Some plots there are green, some are red
 - Looks like larger statistics are needed
 - Rel 19.2.3.7 based task started recently on Titan
 - Found a curious bug in the release's pacman file (ref. to older release). Fixed now.

Completed ATLAS jobs on TITAN



■ MC Simulation (56,279)

Total: 56,279 , Average Rate: 0.00 /s

From May 1 to Sept 17

ATLAS Simulations with Geant4 v10

- Installed rel. 20.3.2.1
- Ran on data suggested by Sasha
- Tests on events that caused crashes on other platforms.
- Interim conclusion: What crashes on Intel runs on AMD
- We would like to run a few tasks with this release to help with Geant4 v10 validation

ATLAS simulation release

- “Small” release fitting on RAM disk
 - ~12 GB in size right now
- Installed AthSimulationBase 1.0.0, then 1.0.1
 - Many Root 6 persistency related crashes
- Reported initial crashes to dev. team (Zach and John Chapman)
- After fixes and patches suggested by Vakho rel. 1.0.1 ran in a single threaded mode
- AthenaMP still crashes in the merging step
 - Identified as CLHEP related issue
 - Devs are working on it.
 - Waiting for the next release 1.0.2