

Joint USATLAS-USCMS HPC blueprint meeting: 6. April 2018

present in Black Hole room at FNAL:

Dirk Hufnagell
Chris Jones
Dave Mason
Eric Vaandering
Elizabeth Sexton-Kennedy
Stephan Lammel

present on Vidyo:

Anthony Tiradani
Sergey Panitkin
Taylor Childers => others in the room ?
Xin Zhao
Brian Bockelman
Burt Holzman
Paolo Calafiura

Didn't talk about XSEDE, would need Rob Gardner to participate

ATLAS use of ORN and Titan

- Long running: backfill mode on Titan
 - custom Panda pilots, submit to holes in batch system
- Trying to switch to Harvester
 - Also desired by Titan
- Exploring singularity images
 - Containers help with IO load on Lustre shared filesystem
 - Reduces Ops effort (also on the Titan side)
 - Titan kernel not supporting all singularity features
 - Needs custom images

- Data Management
 - One time passwords, no way to interact (easily) with storage from outside
 - Globus Online, 3 day authentication lifetime, renewal painful
 - Even Globus Online Rucio support wouldn't help
 - Edge service that pulls/pushes data from/to ATLAS storage works better
 - Recently switched to Harvester
- Backfill and Allocation mix
 - Backfill in 2017 : 170M core hours
 - ALCC allocation July 2017 to June 2018 : 80M core hours
 - Backfill considered to be opportunistic: 'icing on the cake'
 - Allocation is for planned work
 - Backfill is not something that can be counted on
 - It doesn't scale to multiple users (each additional user would reduce the draining holes that backfill tries to fill)
 - Backfill only available at large sites that primarily run very large jobs, at NERSC for instance there are too many users with too many different job sizes and wall times to leave large holes

Some general statements

- Central image building in terms of recipes, not so much in terms of images
 - Titan needs custom images because of kernel
 - Other HPC might need other images
 - Different architecture HPC definitely will need other images
 - Ops load should be ok as long as same recipes can be followed
 - Just need more hardware resources for building the images
- Backfill can still be useful
 - Would need a way to access
 - Submitting 'scavenger' jobs
 - Submitting to special backfill queue
 - No special "hole fitting" logic, but as long as these jobs are short we should be able to benefit
 - Question to our WM systems how we support such short jobs
- Singularity 2.4 supports squashfs images (useful for common image building infrastructure)