Analysis Computing Facility @ SLAC
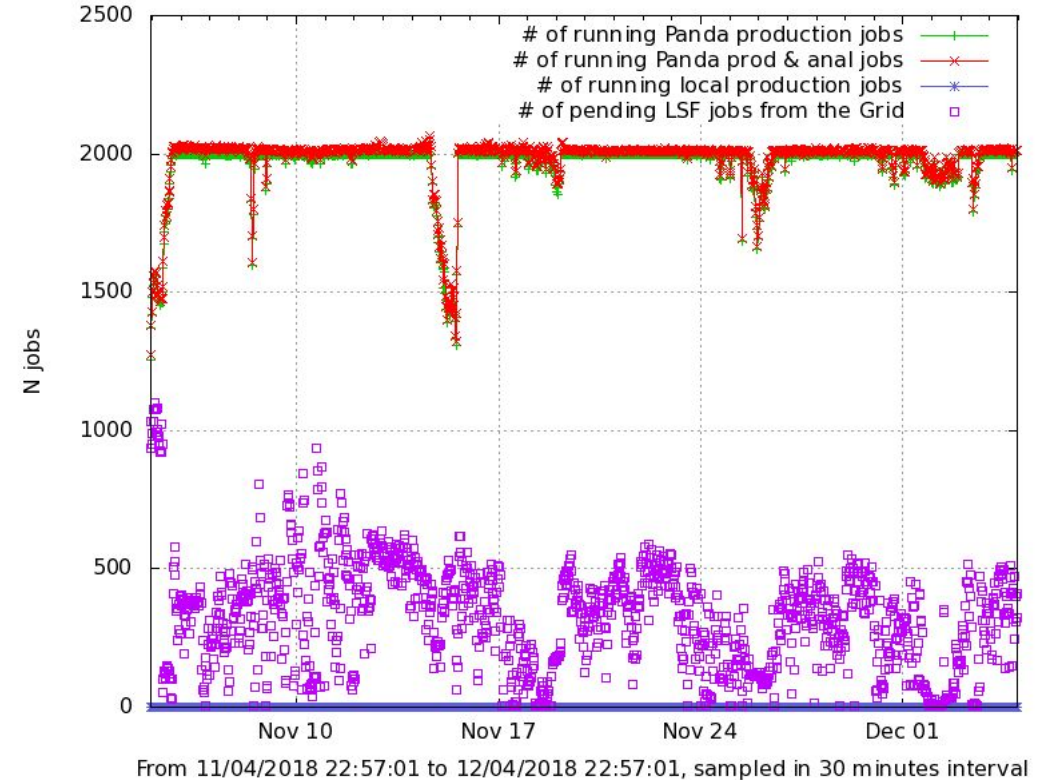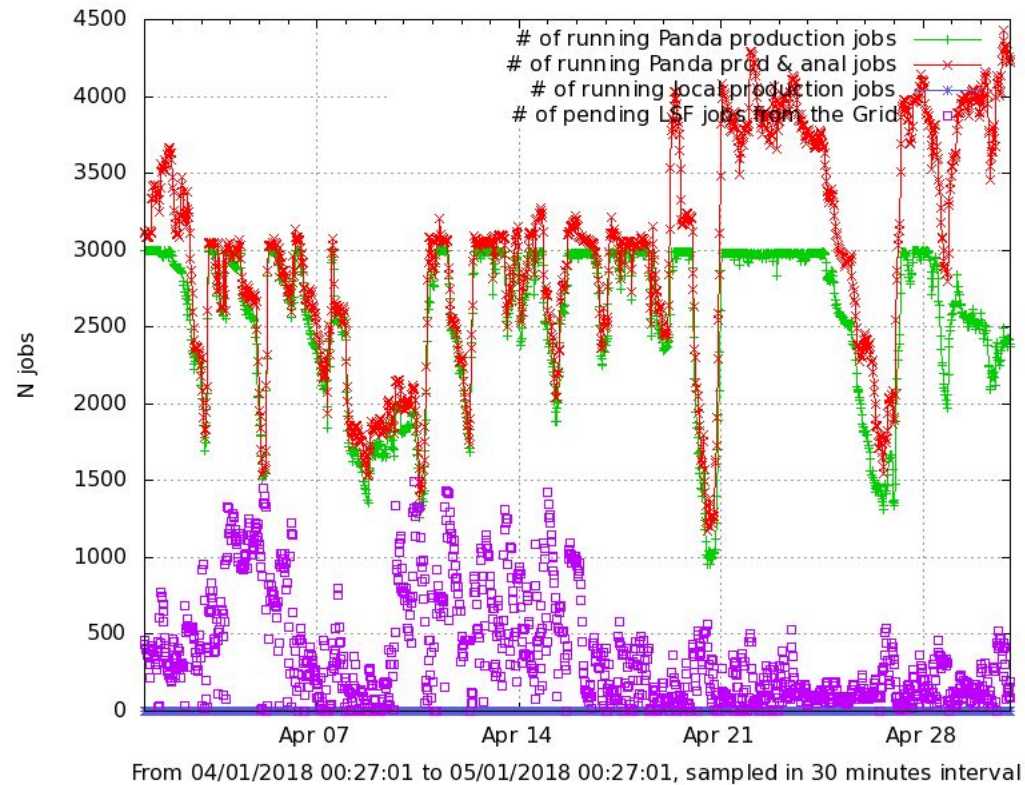
Wei Yang

US ATLAS Computing Facility Meeting, Argonne

2018-12-05

Old WT2 Hardware and T2-like Operation

- 3500 cores and 3.1PB storage
  - Cap the production to 2000 slots to make priority (fair share) to users
  - About 2PB of storage are 8-9 year old. The rest are ~4 year old.
  - Two-tier storage put less stress on them – whole file streaming only
  - Hardware were purchased with Xrootd based storage in mind.

- HTCondor-CE was from an early R&D deployment
  - Update is needed (current HTCondor-CE doesn't support InCommon!)
  - many customizations need to be carry over

- No longer a Tier 2 means more focus on ATLAS users' needs, **in theory**
  - Grid-based operation is still smooth
  - No longer need to pay attention to OSG/WLCG monitoring requirement

Grid Production

From 04/01/2018 00:27:01 to 05/01/2018 00:27:01, sampled in 30 minutes interval

From 11/04/2018 22:57:01 to 12/04/2018 22:57:01, sampled in 30 minutes interval

Grid production (green line) helps to keep the environment – capped to 2000 slots

Between the green line and red line are Grid-based user analysis jobs

Cost and risk of running very old hardware

- **The main risk is losing data**
  - Especially at the back tier storage, ~8-9 year old hardware
    - One accident in November put one storage node (300TB) at risk
    - Spent two full days in machine room - any mistake will be irreversible!
  - Run out of spare parts, hard to find anyone selling them (+ no money)
    - All disks on the node are temporarily spare parts
- Proposal to the RAC committee
  - Significant reduction of available disk space, to improve reliability
  - Rebalance DATADISK and LOCALGROUP
    - This already started, DATADISK from 2.2PB to 1.4PB, can we go down to 500TB?
    - Boost LOCALGROUPDISK
- Can we use Xcache cluster to avoid R2D2 transfer
  - suggested to some users to use Xcache@NERSC. They seems to like it.
  - Much less risk of losing data
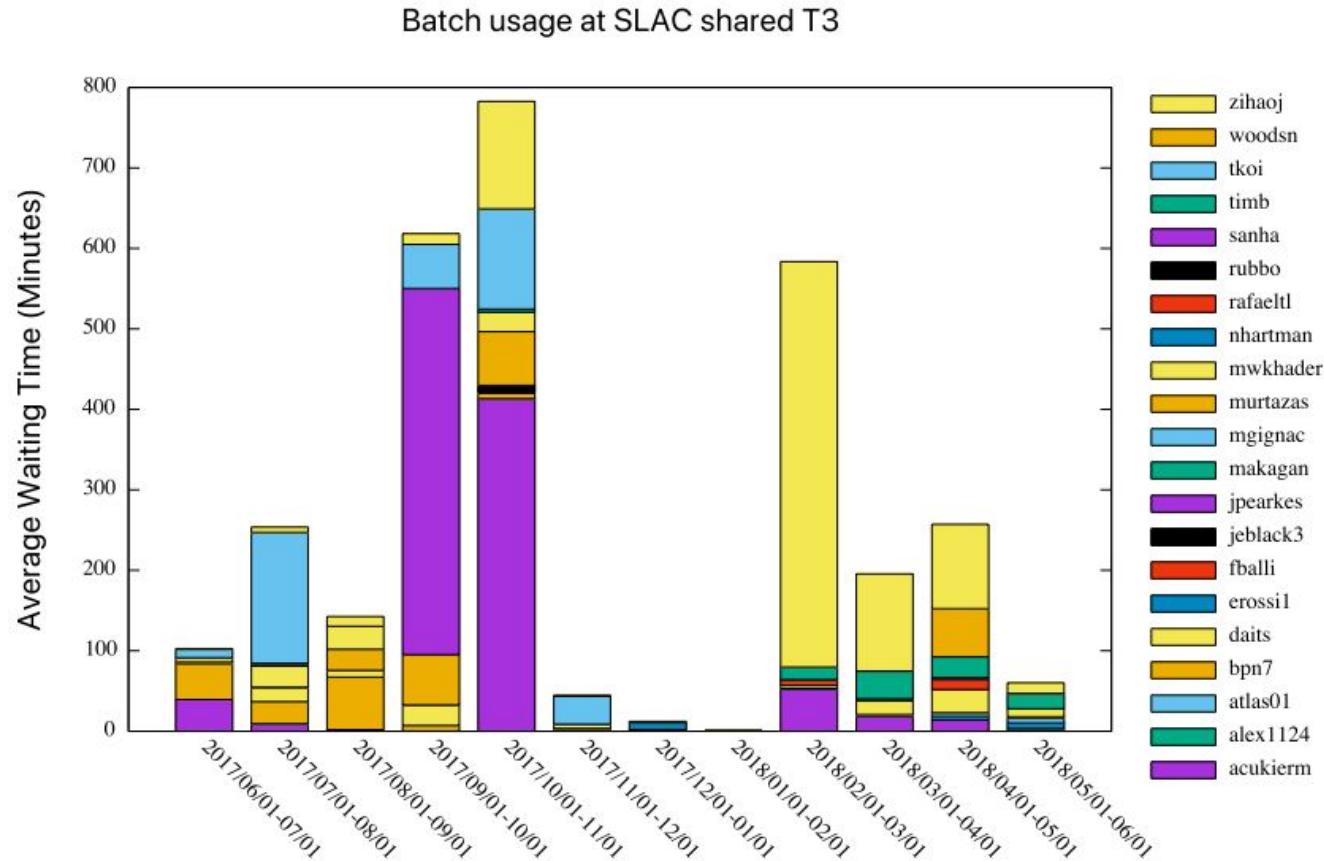
New Focus is User Analysis - the Resource

- Aim at providing a lxplus-like environment
  - With CVMFS, the environment at SLAC and CERN-lxplus are similar
- GPFS based home storage and personal data storage
  - 100GB/person home, 2TB/person data (up to 10TB) ← need input from WBS 5
  - 630TB GPFS storage (ATLAS owned)
  - Use it before it is decommissioned ! (~2021)
- User jobs have access to the WT2 DDM storage (xrootd storage)
  - SLACXRD_DATADISK, SCARTCHDISK, LOCALGROUPDISK
  - Can be the target of analysis job's output, or use R2D2
  - Xcache makes it easy to access remote data

Usage by Analysis Jobs (non-Grid)

Batch usage at SLAC shared T3

Batch usage at SLAC shared T3
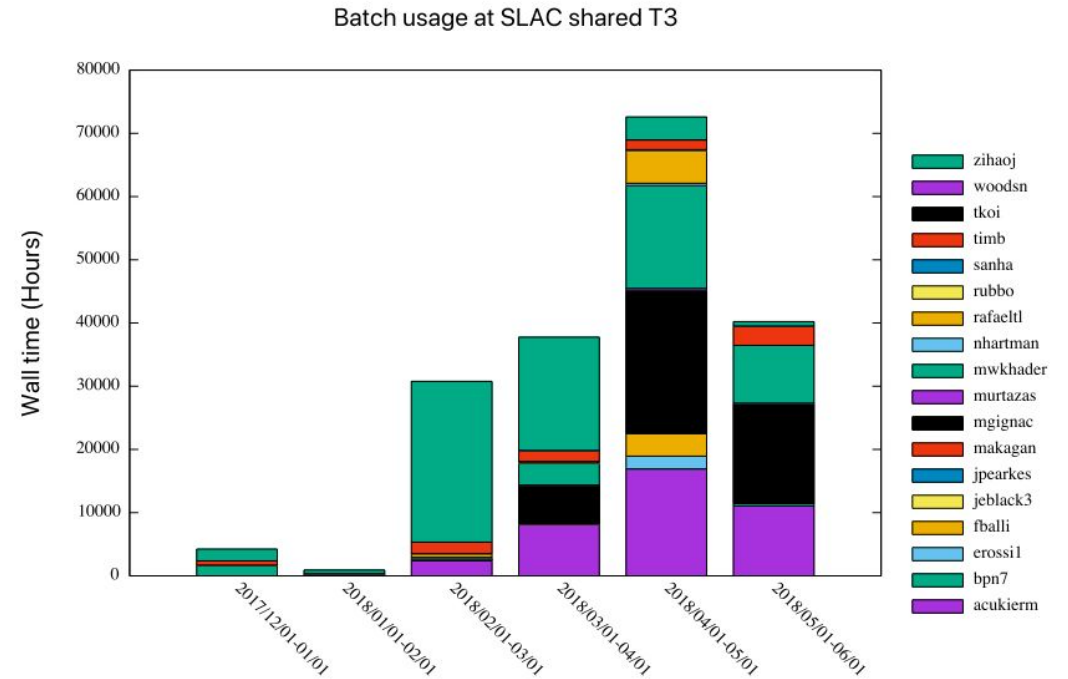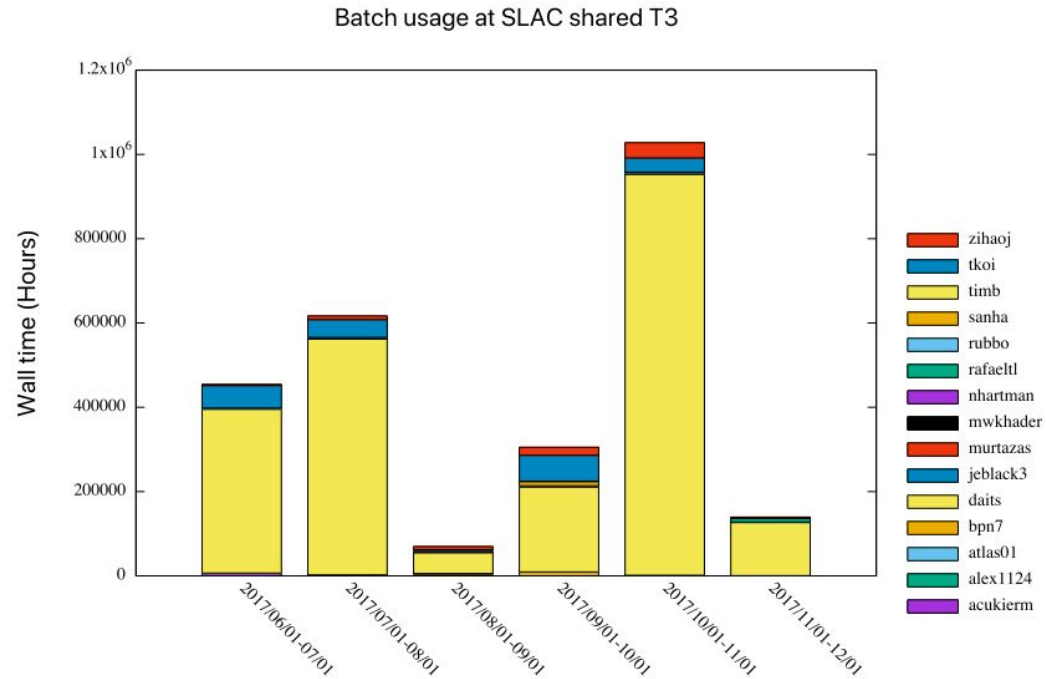
- **Active** Users change over time (students come and go/graduate)
- Sometimes usage is dominated by a few users (e.g. user "daits" in 2017)

US ATLAS Computing Facility Meeting @Argonne, 2018-12-05

Usage of Analysis Jobs (non-Grid), cont'd

Batch usage at SLAC shared T3

- Average waiting time:
  - Variate depend on user fair share usage, and # of submitted jobs
  - Not long in most cases

US ATLAS Computing Facility Meeting @Argonne, 2018-12-05

ACF operation 1

- Established a usable batch and storage environment for user analysis
  - Finally put user account creating process in good shape
    - Integrated into SLAC user account process
    - Otherwise, it will be a huge resource sink
    - Smooth, predictable outcome, though long turnaround time (normally days)
- Working with SLAC ATLAS Center
  - Most new account requests were forwarded to us by SLAC ATLAS Center
  - Improve documents, paying closer attention to their needs
  - Lesson learned:
    - most new users don't have CERN accounts
      - Do not put "ACF 101" doc in CERN Twiki, or protected SLAC web space
    - Need to how users finding out where to ask questions
      - I was asked to help debugging HLT code

US ATLAS Computing Facility Meeting @Argonne, 2018-12-05

ACF operation 2

- Batch accounting and usage monitoring is clearly missing
  - A wider problem for SLAC computing
    - IBM platform analytics for LSF proven to not worth the money
- Need guidance from US ATLAS
  - What to do with inactive users
  - What if a user want extra resource
    - batch priority, disk space, rucio quota
  - User requires on specific software environment
    - So far our approach is to facilitate such a requires and environment
- **Inadequate manpower to deal with required changes and workload**
  - Can not respond timely

Toward more interactive computing

- We need non-trivial resource in interactive pool
  - RHEL6 login pool: many small VMs (small in vCPU and memory)
  - CentOS 7 login pool: fewer, much larger VMs, with Singularity
- Traditional ROOT based analysis via remote X-windows is supported
  - Fast-X (free) doesn't seems to performance as good as NoMachine ($$)
- Jupyter Notebook/Hub is a new direction
  - Learning the experimenting at BNL T3
    - How to "elegantly" connect Jupyter to batch
  - Other SLAC community are also interested and working on that
    - Yee has extensive experience in R&D and support of Jupyter and ML computing for photo science at SLAC
- Support on software environment will become important
  - Just started, funded by WBS 5

US ATLAS Computing Facility Meeting @Argonne, 2018-12-05