

Analysis Computing Facilities

Wei Yang, William Strecker-Kellogg

Many names

- Shared Tier 3
- Share Physics Analysis Computing Facility (SPAR)
- Analysis Computing Facility (ACF)

They all refers to the same thing - computing resources for users

A new service in the US facility

US ATLAS Computing Facility has extensive experience operating T1 and T2s

- We also have Tier 3s, but they are mostly on their own (+ help from Doug)
 - Not an effective solution to address user analysis

Analysis Computing Facility is a new category of services

- **Because:** Grid based analysis can not meeting all user needs:
 - Steep curve to learn the Grid
 - Long tail in completion time and data movement time
- Users, not ATLAS production, is the focus at ACFs
 - So the range stretch from infrastructure to ATLAS software environment
 - Much less tolerant to site issues!
- WBS 2.3 is responsible up to infrastructure software (accurate?)
- Physics Support (WBS 5) is responsible for supporting ATLAS software
 - Dr. Shuwei Ye (BNL) and Dr. Yee-Ting Li (SLAC)
 - Remember there is also ATLAS Centers at ANL, BNL, LBNL and SLAC

Coming from Tier 1 and Tier 2

ACFs at BNL and SLAC inherit from Tier 1 and Tier 2s:

- Storage, CPU nodes
- Networking, DTNs, Login nodes
- Batch system, Grid Infrastructure, CVMFS
- Downsized production environment: config in AGIS, Panda, RUCIO

The above covers some of ACFs hardware and software needs

- But running old hardware is costly and risky
- ACFs workload sometimes put different emphasis on hardware
 - iops, GPU, etc.
 - backup?

User accounts

- Open a user account is part of the Labs process
 - No user group structures and PI
 - When user leaves ATLAS (or US ATLAS ...)
- Quick response to user account request?
 - very different (and slow) compare to NERSC
 - Also, have you heard of annual cyber-security training for NERSC users? ← no such thing!

Accounting

- RUCIO storage quota + Local site storage quota ← decided by who?
- CPU usage ← may not want a quota? But want fair usage among users
- Monitoring: know resource availability; know individual user usage
 - BNL did reasonably good job, SLAC has homework to do.

Remote access

- X-window is the standard today
- Web based remote access is increasing - see later

Login access

- A few login nodes funded by T1 and T2 infrastructure may not be enough
- Mixed with software environment requirement, or even GPU requirement.

Alternative data access

- Long tailing in R2D2 data transfer
- Xcache maybe a useful service at ACFs

Data sharing? ← do we have a standard service?

Container, Jupyter notebooks, ML capabilities

- Current hot topics, some are still in R&D phase
- Requirement and service are both stable AND frequent changing

Software management tools for non-CVMFS software?

- spack, container, etc.
 - Install something for user or help user install?

User reach out and training

- This needs to be done periodically
- Documentations!

User support

- Diagnose problem, suggest best solutions
- Are we ready for this task? manpower, ticketing system? Process?
 - Need more manpower for user support in context of new technologies!

Summary

ACFs is a new category of service

- Users are the focus
- Need to deal a different set of requests.
- Innovative analysis methods are growing
- Need both manpower and hardware resources to support user analysis
- Cooperation between BNL and SLAC ACFs may help improving utilization of our limited human resource.