# ATLAS ROOT I/O pt 2

- Atlas Hot Topics (with reference to CHEP presentations)
- Big data interlude (not ATLAS )
- ROOT I/O Monitoring and Testing on ATLAS
- Atlas feature requests /fixes

Wahid Bhimji

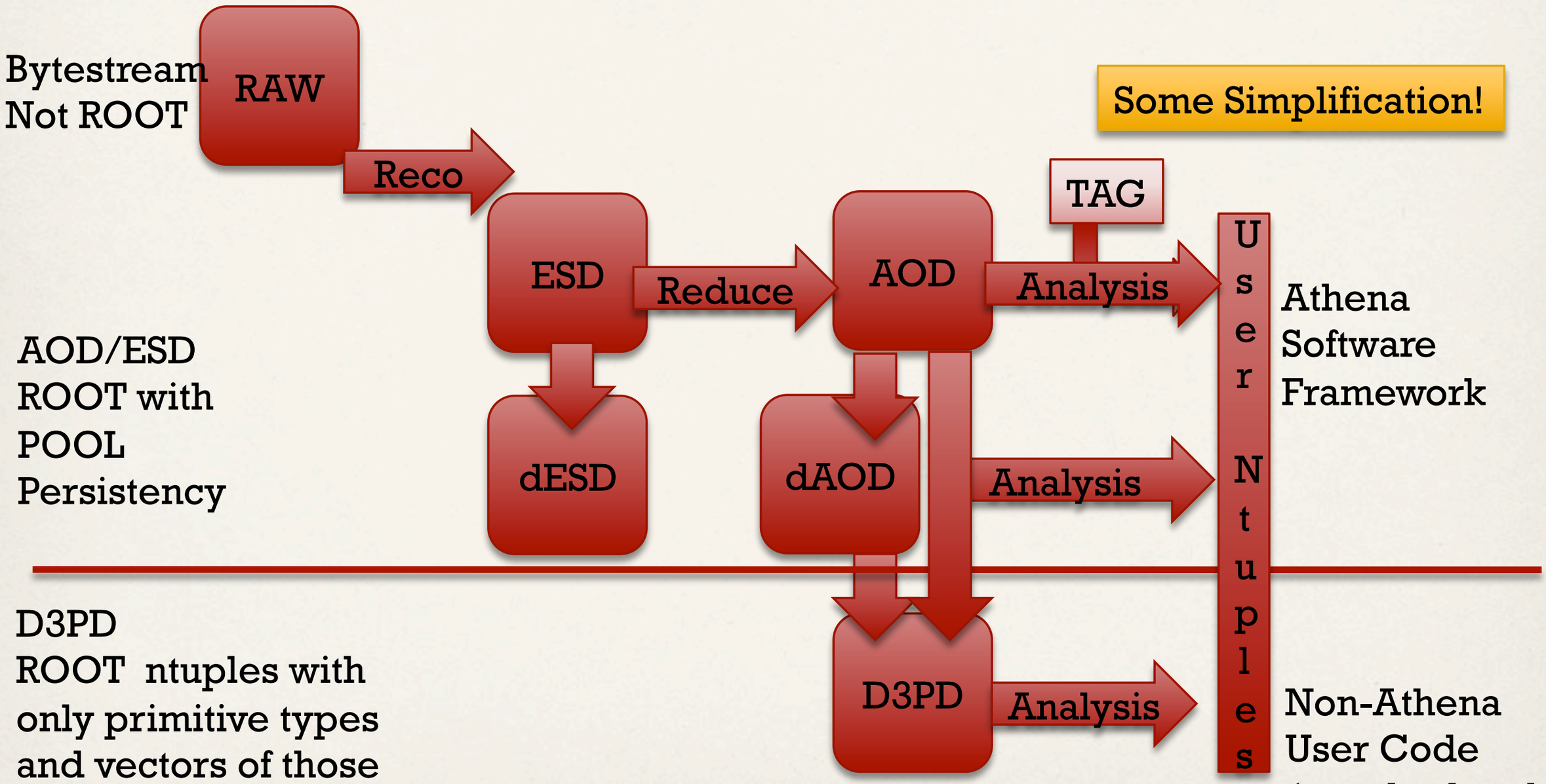
---

*6th Dec 2013 ROOT IO Workshop*

# Hot topics

# Old (current) ATLAS data flow

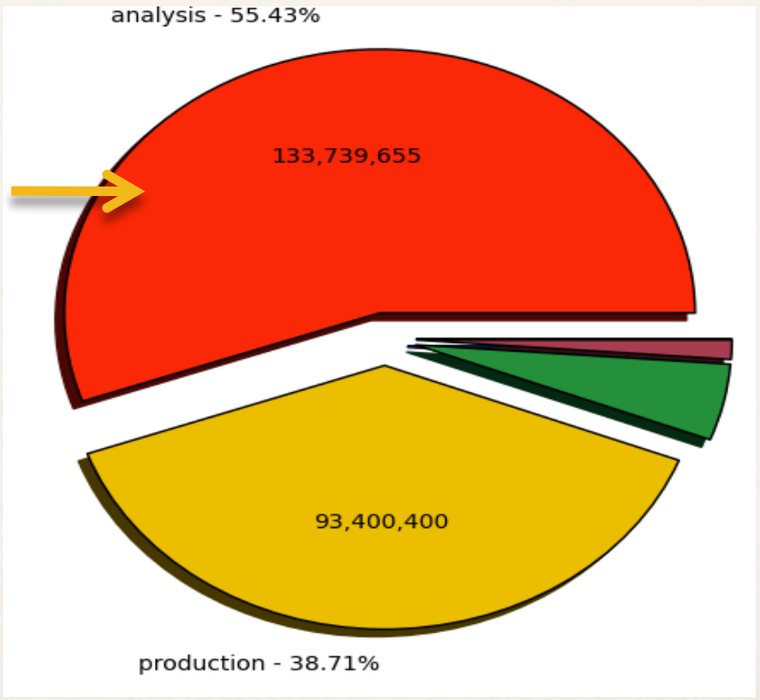
See my talk - at last CHEP



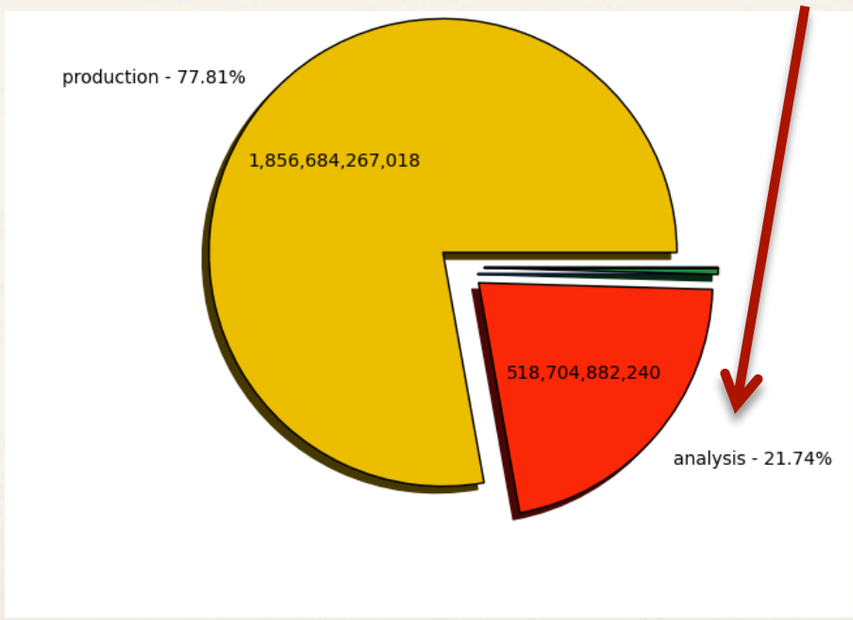
# A problem with this is

- Much heavy-IO activity is not centrally organised
- Run using users own pure-ROOT code - up to them to optimise

By no. of jobs, Analysis =55%



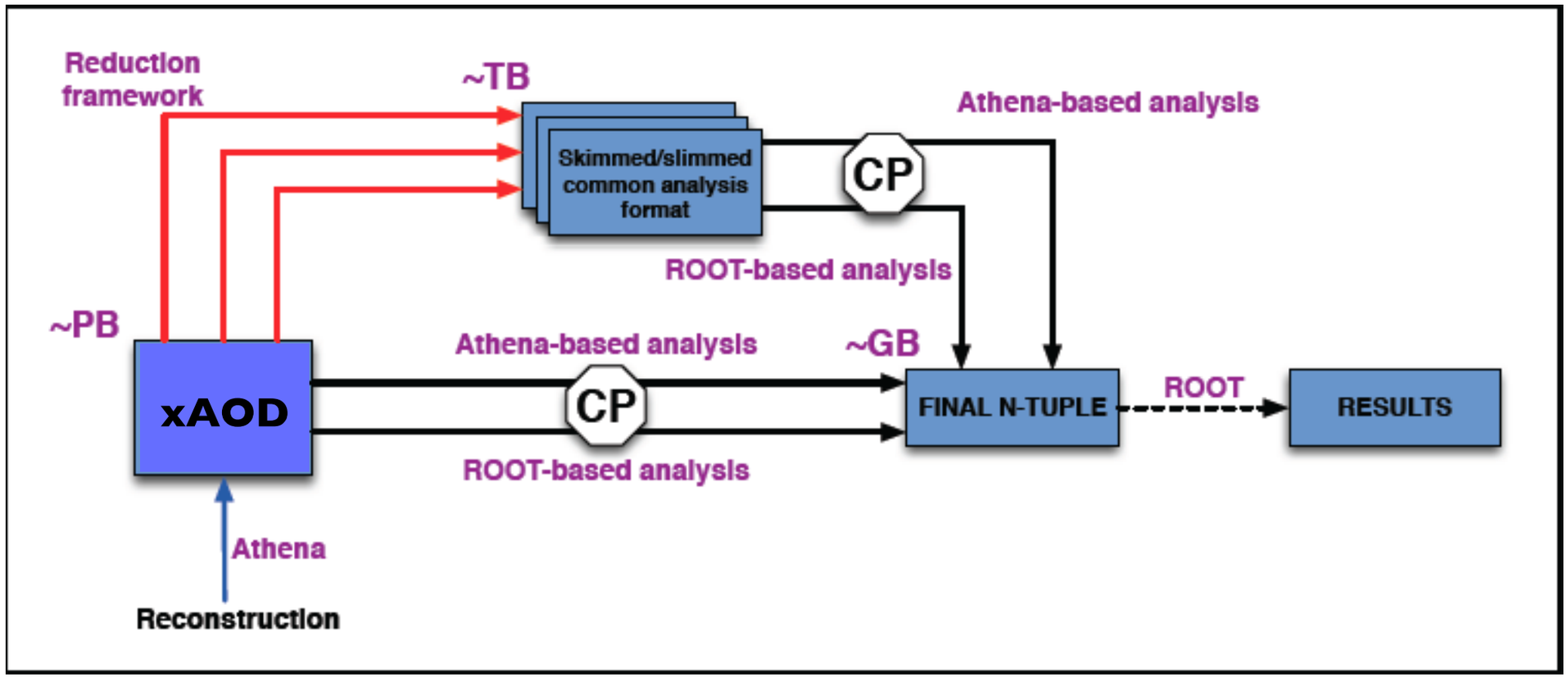

By wallclock time, Analysis=22%




ATLAS ROOT-based data formats - CHEP 2012

# New (future) ATLAS Model

see Paul Laycock's CHEP talk




- xAOD easier to read in pure ROOT than current AOD.
- Reduction framework centrally controlled and does heavy lifting
- Common analysis software

# New (future) ATLAS Model

- Of interest to this group
  - New data structure: the xAOD
    - Opportunities for IO optimisations that don't have to be in each users code: reduction framework and common analysis tools
- A step on the way: NTUP_COMMON
  - Previously many physics groups have their own (large) D3PDs overlapping in content - using more space than need be
  - So new common D3PD solves that issue. But has a huge number (10k+) of branches. Not an optimal solution - will go to xAOD.

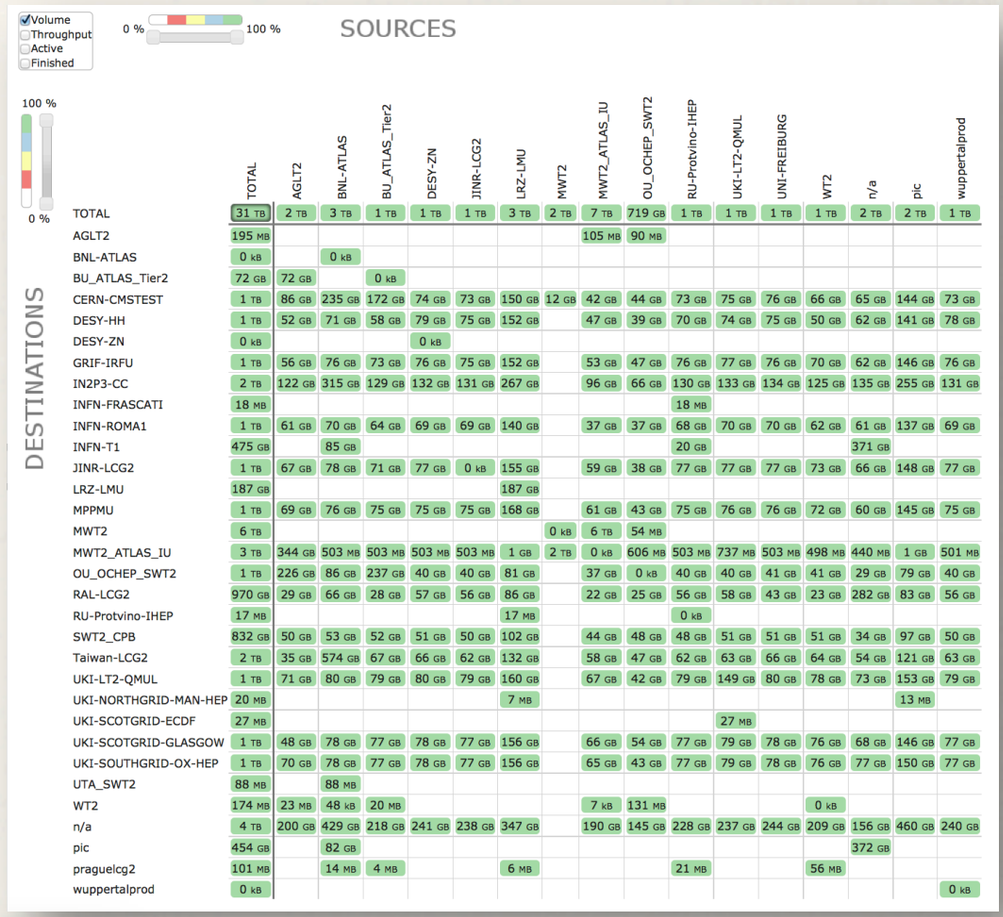
# ATLAS Xrootd Federation (FAX)

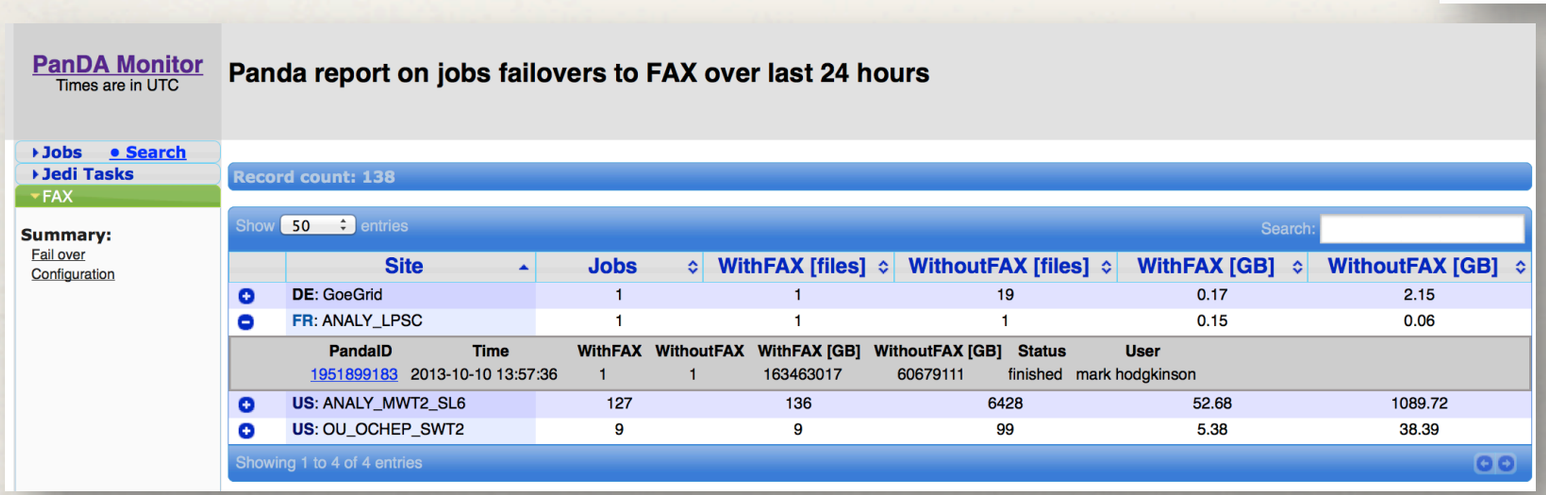
[see Ilija VUKOTIC's CHEP talk]

## Aggregating storage in global namespace with transparent failover

## Of interest to this group:

- WAN reading requires decent I/O.. And is working out OK for us (though not exposed to random users yet (except as a fallback))



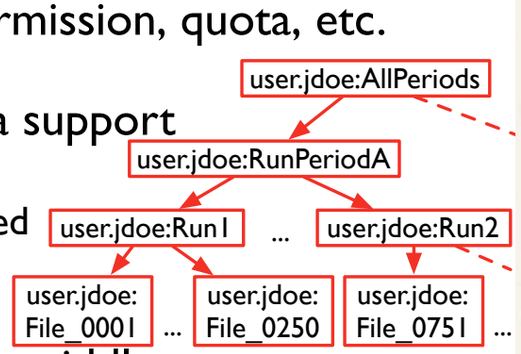
# Atlas Rucio

see Vincent GARONNE's CHEP talk

- Redesign of data management system

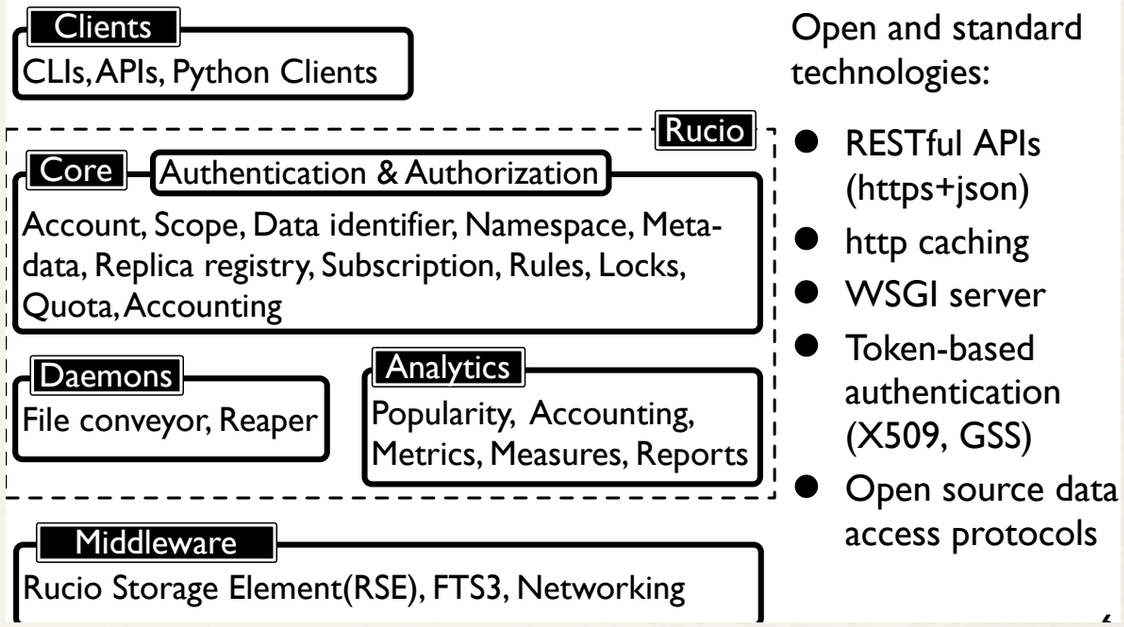
## Concepts - Highlights

- Better management of users, physics groups, ATLAS activities, data ownership, permission, quota, etc.
- Data hierarchy with metadata support
  - Files are grouped into datasets
  - Datasets/Containers are grouped in containers




- Concepts covering changes in middleware
  - Federations
  - Cloud storage
  - Move towards open and widely adopted protocols

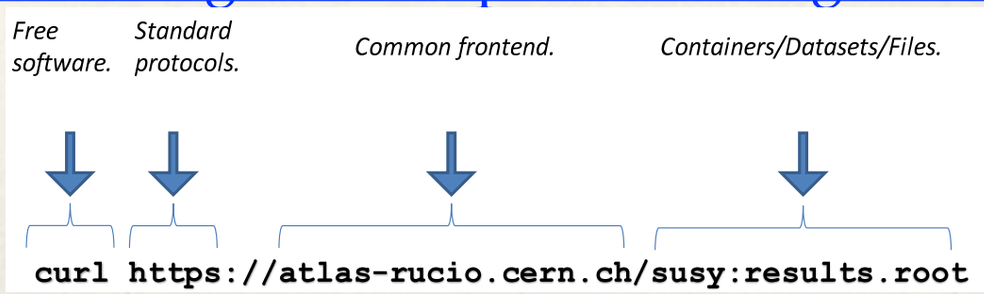
## Software Stack




## Of interest to this group:

see Mario Lassing's CHEP poster and lightening talk

- Supports container file (metalink) for multiple sources / failover.
- Possible http based "federation"



# Big data interlude (not ATLAS)

# Chep theme was “BiG data …”

- “Big data” in industry means Hadoop and its successors
- Number of CHEP presentations for physics event I/O (ie as well as the log mining / metadata use cases that came before))- e.g.:
  - EBKE and Waller: Drillbit column store
  - My “Hepdoop”poster; Maaike Limper’s poster and lightning talk
  - Many on using Hadoop processing with ROOT files
- Most don’t see explicit performance gains from using Hadoop
- My impression: scheduling tools and HDFS filessytem very mature data structures less so (for our needs) but much of interest from Dremel

# Big data opportunities

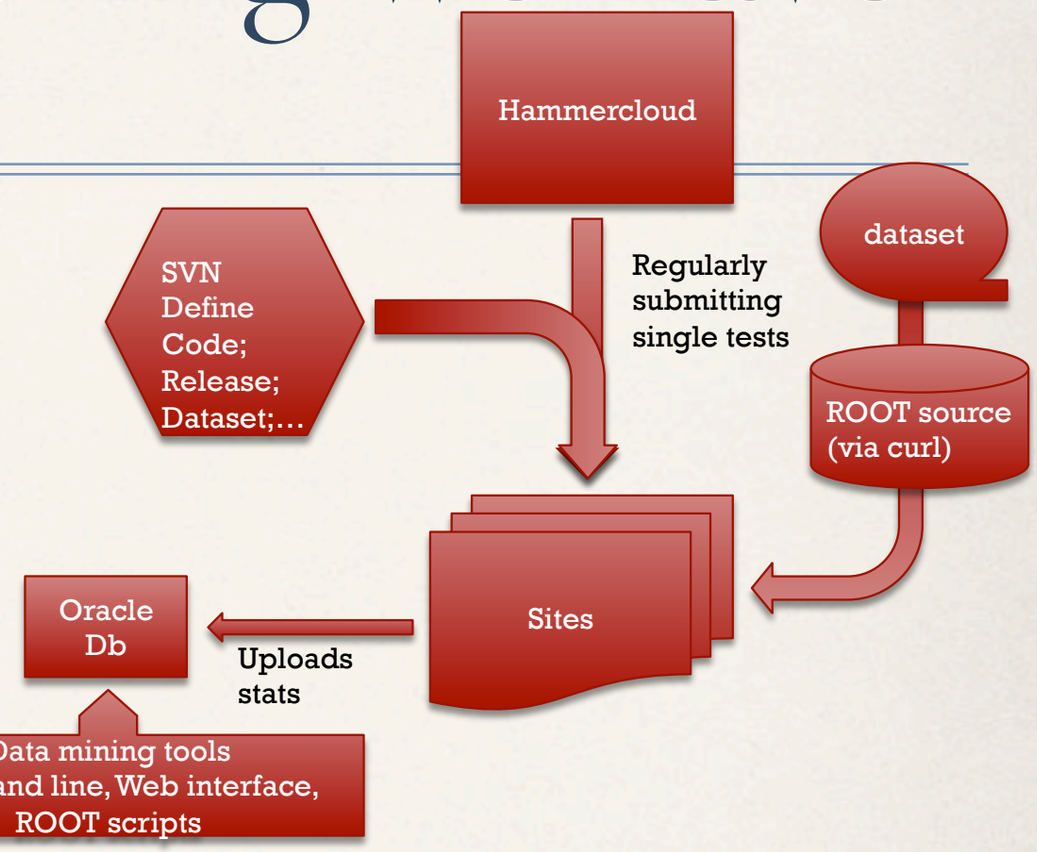
- Opportunities to benefit from growth of “big data”
  - “Impact” of our work
  - Sharing technologies / ideas. Gaining something from the others
- As Fons said parts of Dremel “sound pretty much like ROOT” but that should make us sad as well as proud.
- It would be great if we make ROOT usable / used by these communities and their products useable by us: some areas are Friend trees, ROOT modular “distribution”; chunkable ROOT files etc.
- I realize this requires manpower but surprising if the hype can’t get us some money for transferring LHC big data expertise to industry

# ATLAS Monitoring and Testing

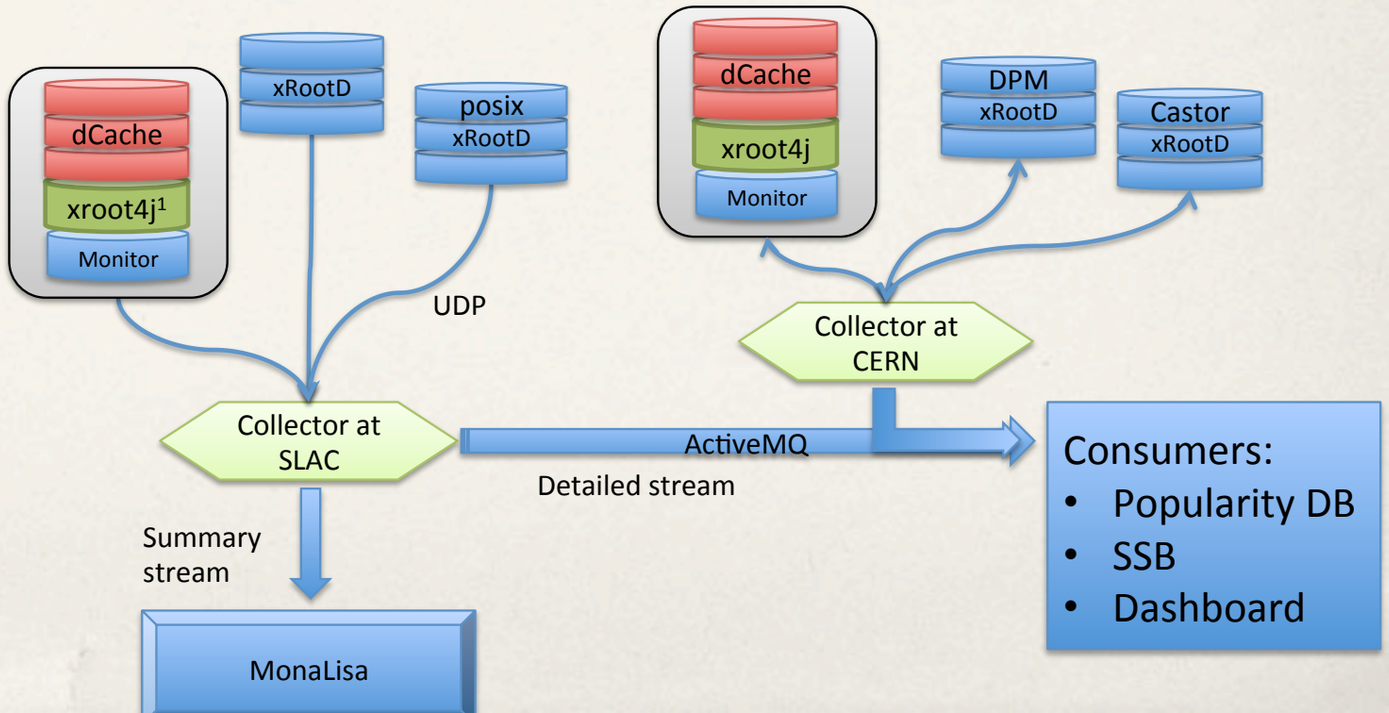
# What testing /monitoring we have

- Still have Hammercloud ROOT I/O tests (see previous meeting). Now testing also FAX /WAN

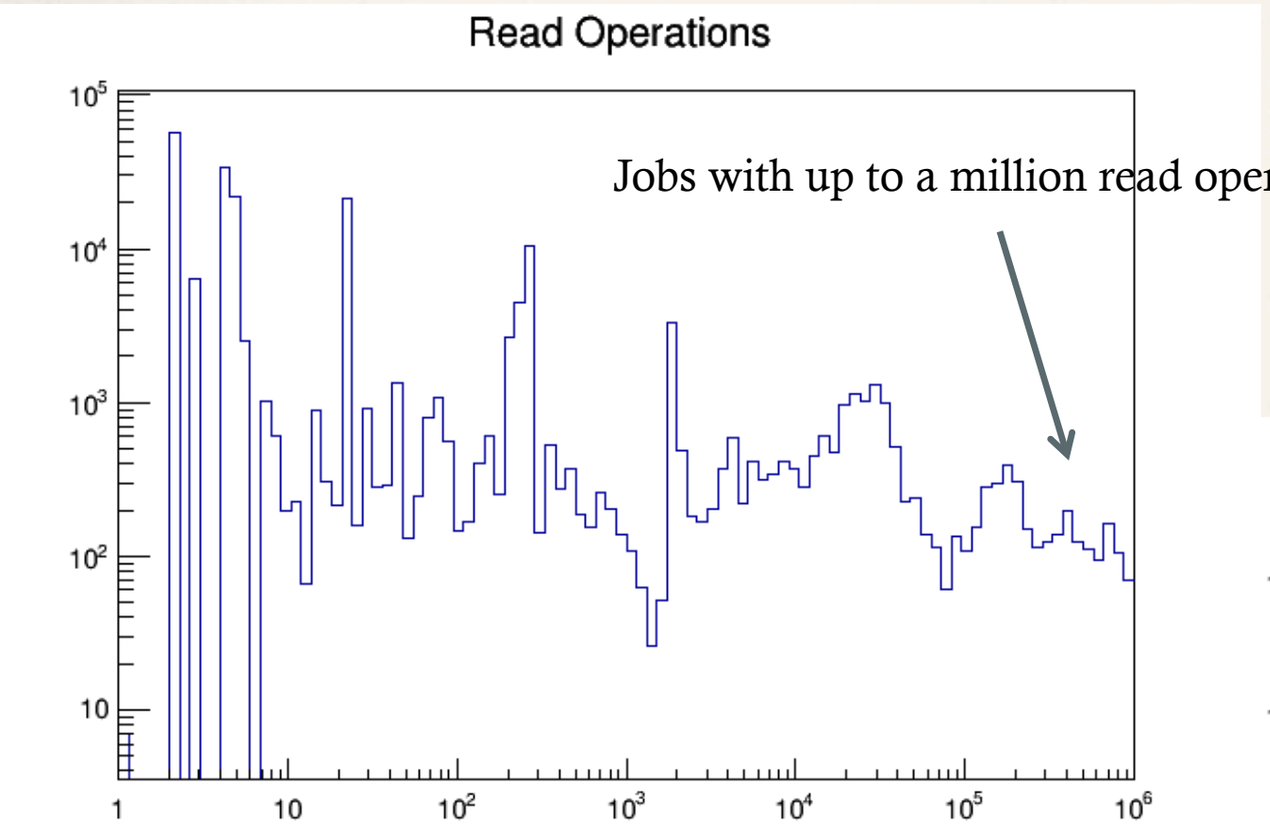
- ad hoc Hammercloud stress tests with real analysis codes (also for FAX)



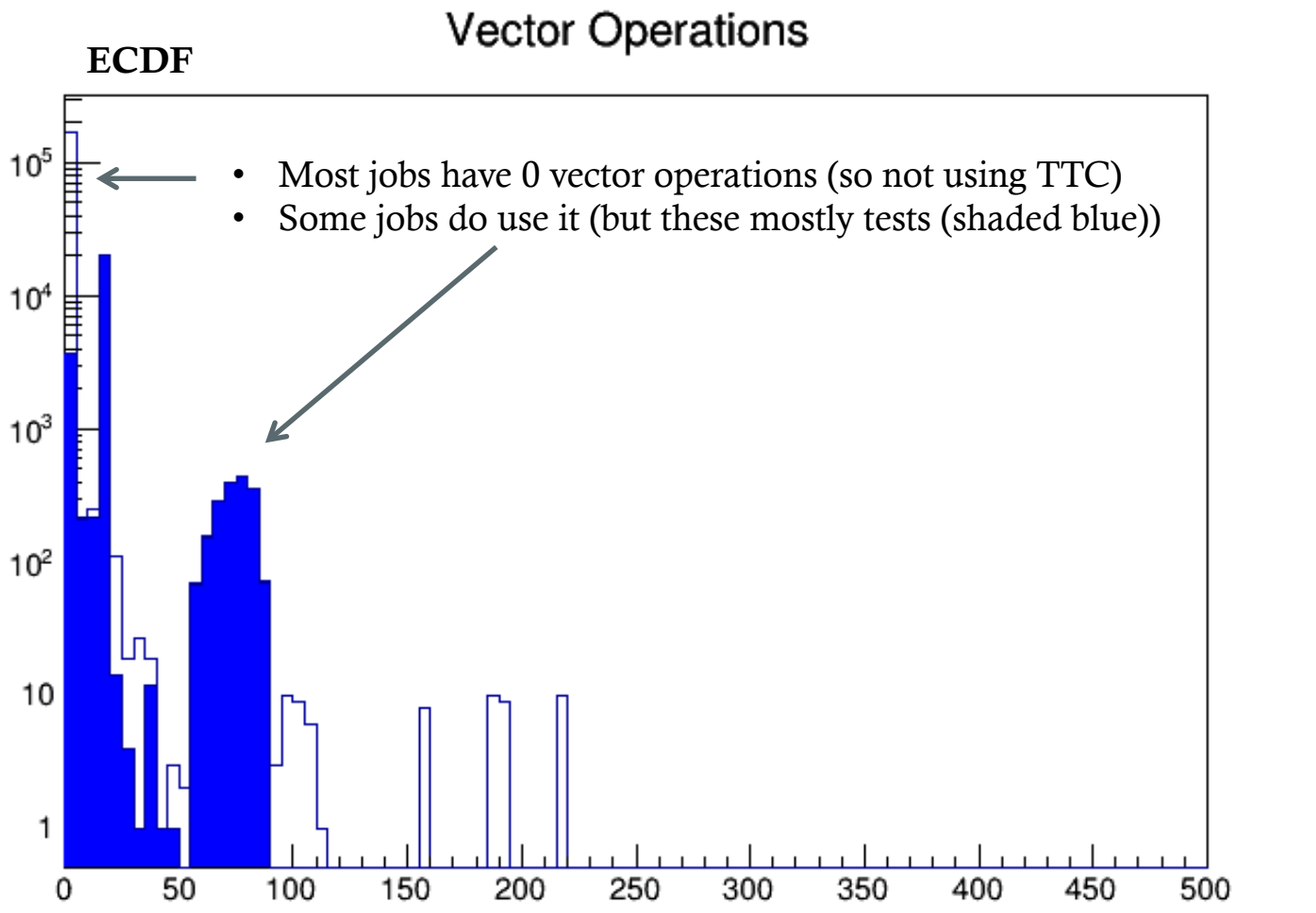

- But also now have server-side detailed xrootd records
  - for federation traffic
  - Also local traffic for sites that use xrootd
  - Regular monitoring but also can be mined



# From mining of xrootd records from Edinburgh site - all jobs

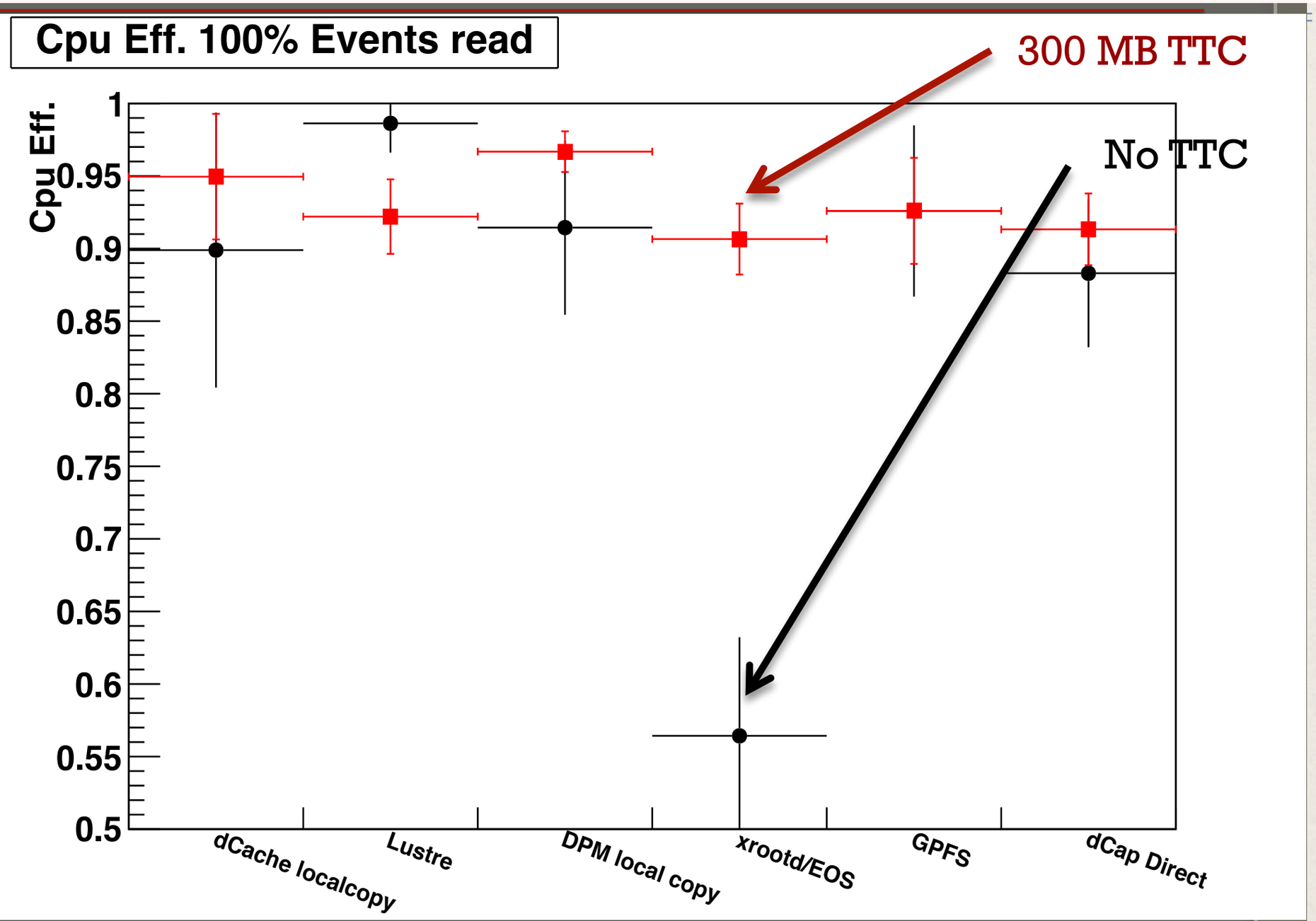



We need to be able to switch on TTreeCache for our users!



# Old slide from CHEP 2012 to remind of the impact of TTC

- TTreeCache essential at some sites
- Users still don't set it
- Different optimal values per site
- Ability to set in job environment would be useful



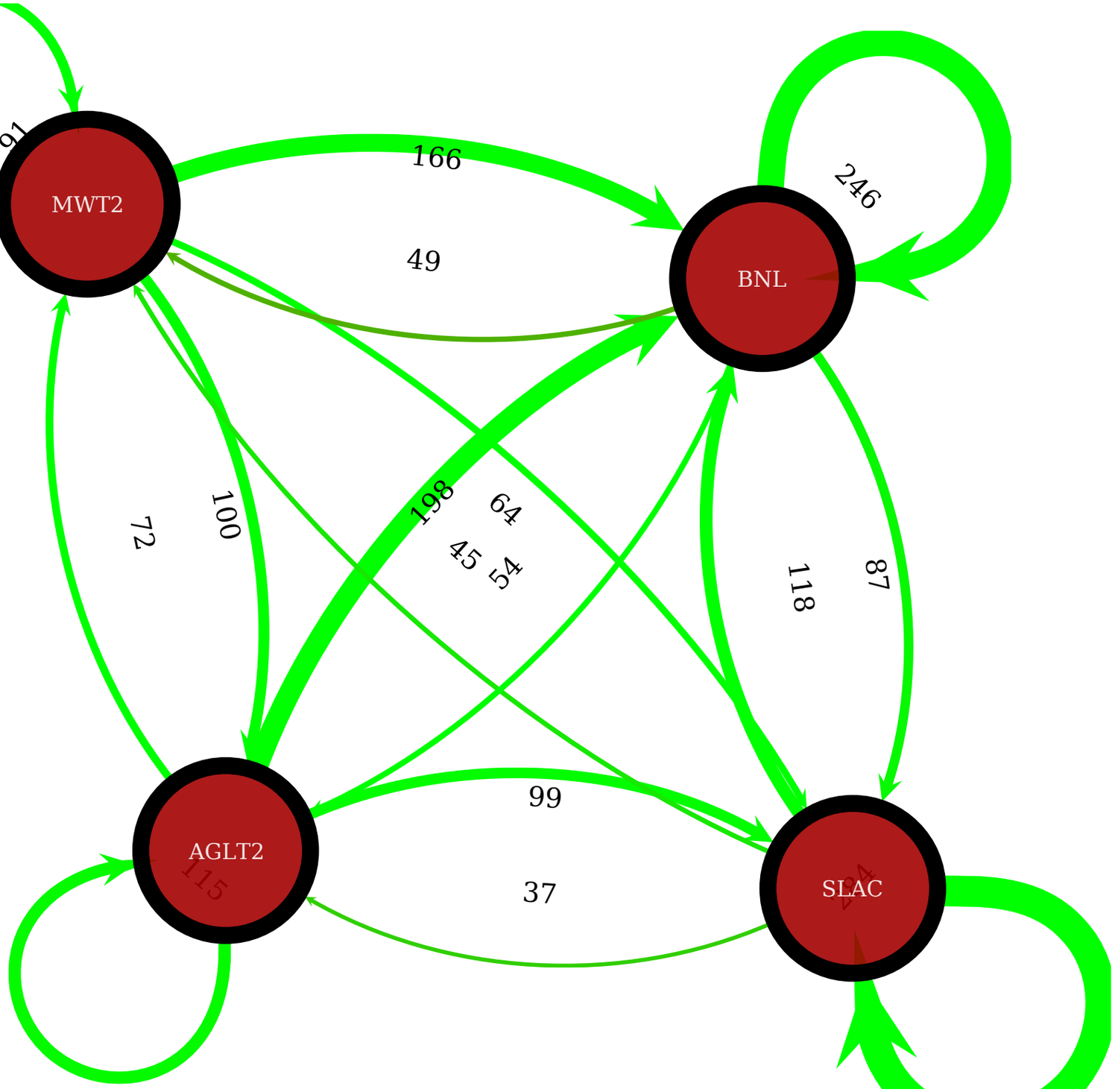
# Oct HC FAX stress tests - US cloud

[See Johannes Elmsheuser et al.'s CHEP poster]

Width and numbers: Event rate

Green = 100% success

Remote reading working well in these tests (which do use TTreeCache !)



# ROOT IO Feature Requests

- TTreeCache switch on/ configure in environ.: (including in ROOT 5.)
  - This is useful even if TTC on by default or in new framework.
  - Choices (multiple trees etc.) - but are there blockers?
- Support for new analysis model: generally for xAOD as it develops
  - Specific Reflex feature “rules for the production of dictionaries for template types with markup” in ROOT 6 (already on todo list I hear)
  - Advice on handling Trees with 5000+ branches.
- Planned http access would benefit from TDavixFile: is it in now?