US Atlas Tier 3 Overview

Doug Benjamin
Duke University

Purpose of a Tier 3

- Tier 3's (and experiment computing in general) are tools to aid the Physicists in their work
  - Work – analyzing the data to make measurements and scientific discoveries
  - The computing is a tool and a means to an end
- Tier 3 Productivity
  - The success of the Tier 3's **will** measured by
    - **The amount of scientific output**
      - **Papers written**
      - **Talks in conferences**
      - **Students trained (and theses written)**
      - **Not in CPU hours or events processed**

What is a Tier3?

- Working definition
  - "Non pledged resources"
  - "Analysis facilities" at your University/Institute/...
- Tier 3 level
  - The name suggests that it is another layer continuing the hierarchy after Tier0, Tier1s, Tier2s...
  - Probably truly misleading... (except for political reasons some places in Europe really want to be small Tier 2's)
  - Qualitative difference here:
    - **Final analysis vs simulation and reconstruction**
    - **Local control vs ATLAS central control**
    - **Operation load more on local resources (i.e. people) than on the central team (i.e. other people)**

Tier 3 Types

- Tier 3's are non pledged resources
  - Does not imply that they should be chaotic or troublesome resources though
- Atlas examples include:
  - Tier 3's collocated with Tier 2's
  - Tier 3's with same functionality as a Tier 2 site
    - There is a very strong push in Europe that All Tier 3's be grid sites and supply some cycles to Atlas
  - National Analysis facilities
  - Non-grid Tier 3 (Tier 3g) (most common for new sites in the US and should be through out Atlas)
    - Very challenging due to limited support personnel

Tier 3: interesting features

- Key characteristics (issues, interesting problems)
  - Operations
    - Must be simple (for the local team)
    - Must not affect the rest of the system (hence central operations)
  - Data management
    - Again simplicity
    - Different access pattern (analysis)
      - I/O bound, iterative/interactive
      - More ROOT-based analysis (PROOF?)
      - Truly local usage
    - "Performances"
      - Reliability (successful jobs / total)
      - Efficiency (CPU/elapsed) → events read per second

Tier 3

- Of course the recipe *Tier 3 = (small) Tier2* could make sense in several cases
  - *This seems to be how Europe thinks things should be.*
- But in several other cases:
  - Too heavy for small new sites
    - "Human cost"
    - The new model is appealing for small Tier2-like centre as well
- Once again it appears the US Atlas is diverging from Atlas (we are out in front again)
- In <u>all</u> cases:
  - We got data! The focus is more and more on doing the analysis than supporting computing facilities ;)

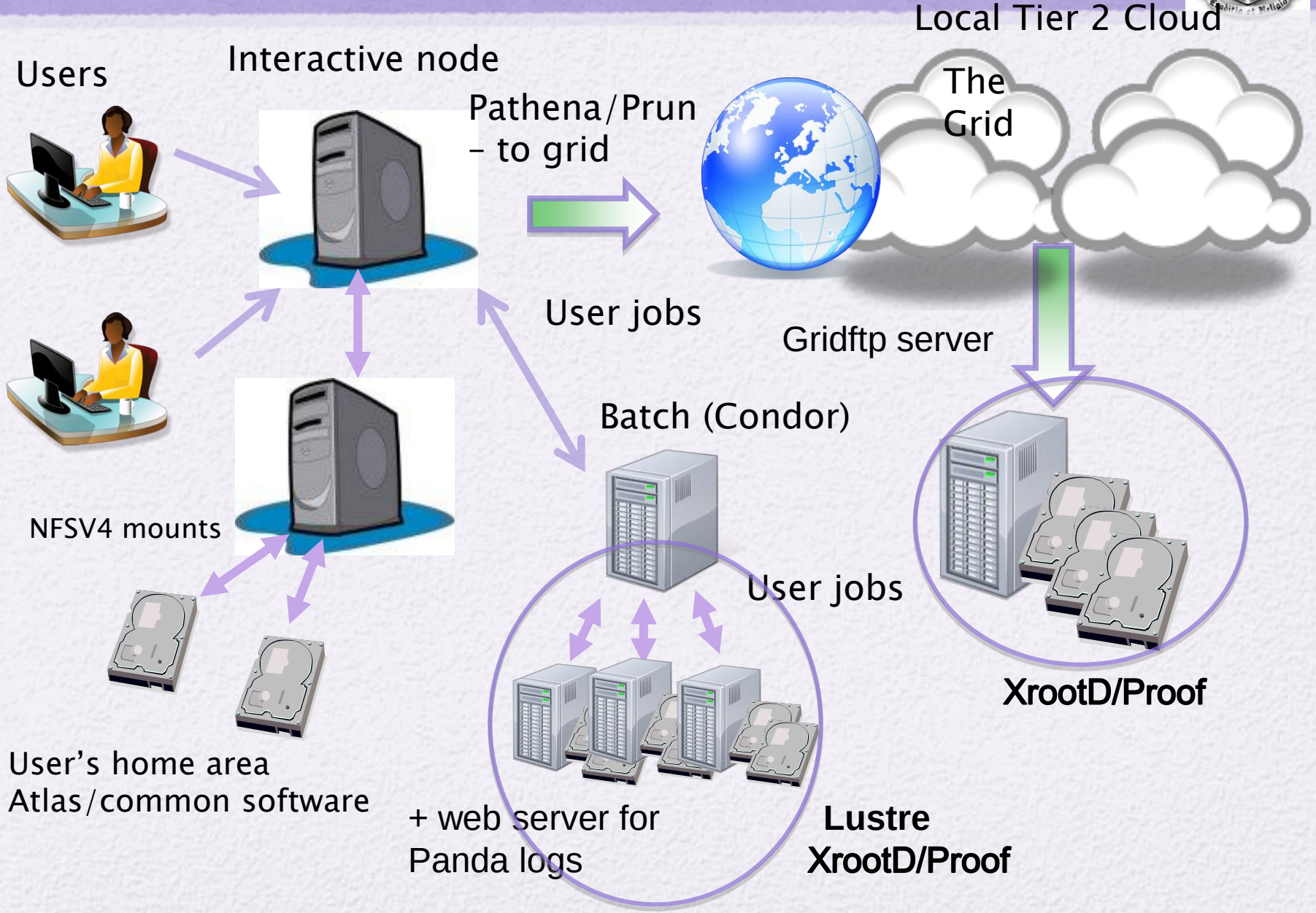
Tier 3g design/Philosophy

- Design a system to be flexible and simple to setup (1 person < 1 week)
- Simple to operate - < 0.25 FTE to maintain
- Scalable with Data volumes
- Fast - Process 1 TB of data over night
- Relatively inexpensive
  - Run only the needed services/process
  - Devote most resources to CPU's and Disk
- Using common tools will make it easier for all of us
  - Easier to develop a self supporting community.

Tier 3g

- Interactive nodes
- Can submit grid jobs.
- Batch system w/ worker nodes
- Atlas Code available
- Client tools used for fetch data (dq2-ls, dq2-get)
  - Including dq2-get + fts for better control
- Storage can be one of two types (sites can have both)
  - Located on the worker nodes
    - Lustre/GPFS (mostly in Europe)
    - XROOTD
  - Located in dedicated file servers (NFS/ XROOTD)

Tier 3g configuration



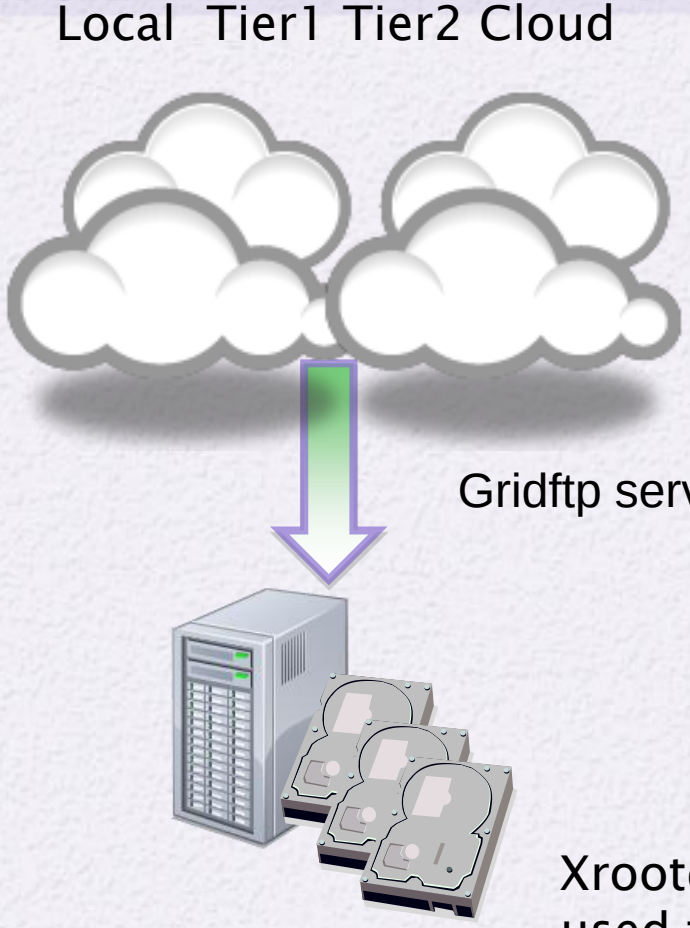
How data comes to Tier 3g's

Two methods
• Enhanced dq2-get (uses fts channel)

•Data subscription
  •SRM/gridftp server part of DDM Tiers of Atlas

Bestman Storage Resource Manager (SRM) (fileserver)

•Sites in DDM ToA will tested frequently
•Troublesome sites will be blacklisted (no data) extra support load



Data will come from **any** Tier 2 site

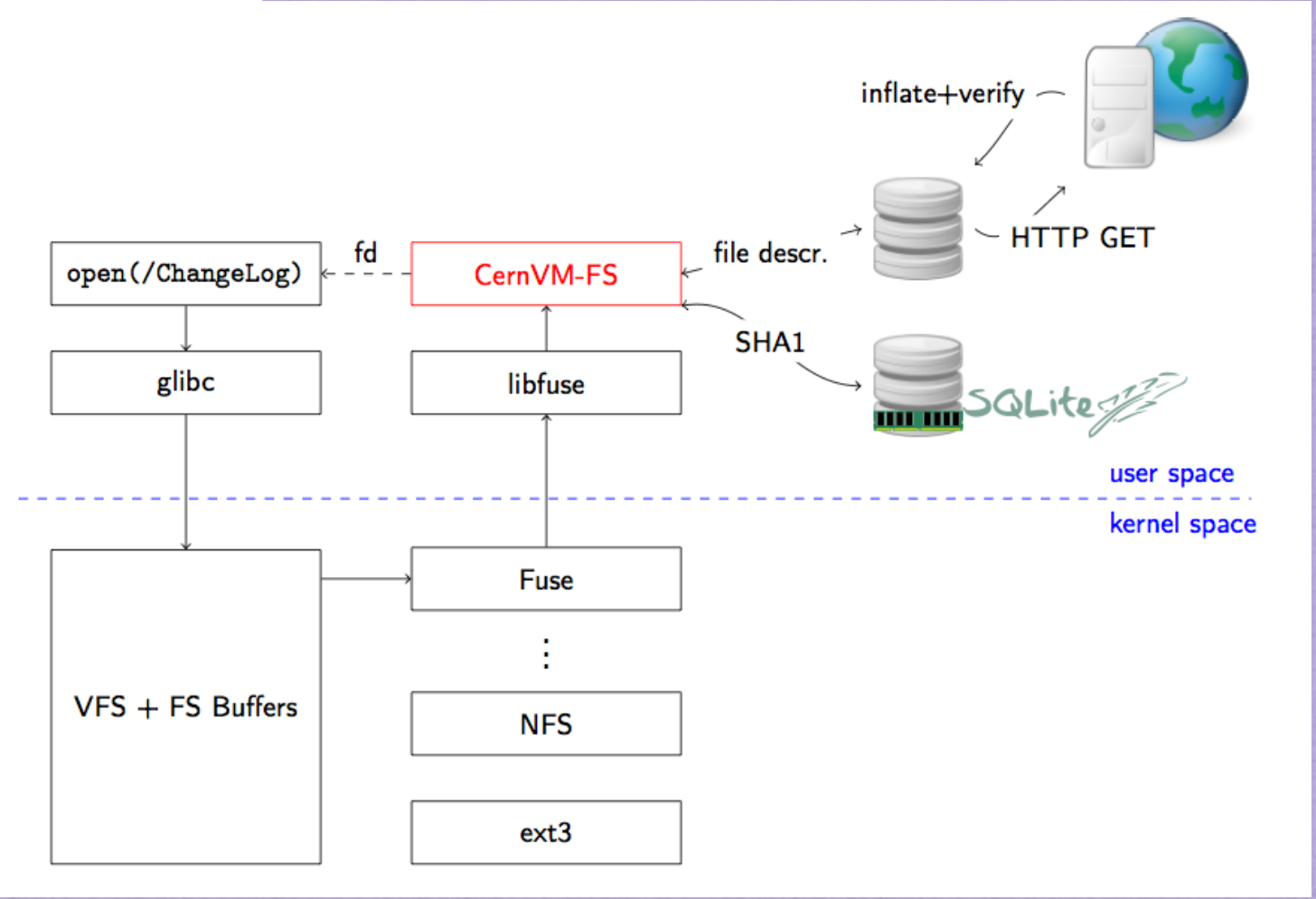
Xrootd/ Proof (pq2 tools) used to manage this

Tier 3g challenges

- Need simplified way for Users to setup what they need to do work
  - Found a existing solution in Canada
  - Collaborate with Asoka de Silva (Triumf)
  - We test the US part of the software package
- Atlas constantly producing new software releases
  - Maintaining an up to date code stack much work
  - Tier 1 and Tier 2 sites use grid jobs for code installation
  - Need a transformative solution to solve problem
    - CernVM File system (cvmfs)

CernVM
SoFTware Appliance

CVMFS (v2)



Transformative technologies

- By their operational requirements non-grid Tier 3 sites will require transformative ideas and solutions
- Short term examples
- CVMFS (Cern VM web file system)
  - Minimize effort for Atlas software releases
  - Conditions DB
- Atlas recently officially request long term Support for CVMFS for Tier 3's
- Atlas is testing cvmfs for Tier 1's and Tier 2's also

Transformative technologies(2)

- Xrootd/Lustre
  - Xrootd allows for straight forward storage aggregation
  - Some other sites using Lustre
  - Wide area data clustering will helps groups during analysis (couple xrootd cluster of desktops at CERN with home institution xrootd cluster)
- Dq2-get with fts data transfer. – Robust client tool to fetch data for Tier 3 (no SRM required – not in ToA – a simplification)
- Medium/Longer term examples
- Proof
  - Efficient data analysis
  - Tools can be used for data management at Tier 3
- Virtualization / cloud computing

Atlas XROOTD Demonstrator project

- Last June at WLGC Storage workshop
  - Atlas Tier 3 proposed alternative method for delivering data to Tier 3 using confederated XROOTD clusters
- Physicists can get the data that they actually use
- Alternative and simpler than ATLAS DDM
  - In testing now
  - Plan to connect Tier 3 sites and some Tier 2 sites
- CMS working on something similar (Their focus is between Tier 1/Tier 2 – complimentary – we are collaborating)

Status of Tier 3g's in US

- Setup instructions have moved to CERN
  https://twiki.cern.ch/twiki/bin/view/Atlas/Tier3gSetupGuide
- ARRA funds have finally arrived
- People are making purchases and making choices
- Brandeis - First Tier 3g (other than ANL or Duke) setup with ARRA funds is almost completely up.
  - I need to finish:
    - setting up Tier 3 Panda (w/ Alden's help)
    - Gridftp servers
    - XrootD
    - Proof for managing the data
    - I am revising the instructions as I go
- Rik and I have advising sites as they request help
  - Willing to travel as needed.

Missing Pieces for Tier 3's

- Standard installation method for XROOTD
  - VDT promised a YUM repository (within a few months)
- Data set affinity (binding jobs to data location)
  - Charles is working on this
- Proof installation and configuration instructions
  - Important for dataset management and ultimately for data analysis (see Sergey's talk today about BNL Tier 3)
- Configuration management (Puppet)
  - Yushu will present the status today
- Standard Test Suite of "typical" Tier 3 jobs
  - Attila's talk covered some of this
- XROOTD Proxy
  - Collaborating with CERN to use a proxy for data severs behind a firewall or on a private network
- SVN repository for all code/instructions scripts

Conclusions

- Tier 3 computing important for data analysis in Atlas
- A coherent Atlas wide effort had begun in earnest
  - Situation now much less clear
- US Tier 3's are being designed according to the needs of the local research groups
- Striving for a design that requires minimal effort to setup and successfully run.
- Technologies for the Tier 3's are being chosen and evaluated based on performance and stability for data analysis
- Ideas from Tier 3 should be moving up the computing chain

Backup Slides

Atlas Code installation

- NFS file server
  - ManageTier3 SW package (Asoka DeSilva Triumf)

https://twiki.atlas-canada.ca/bin/view/AtlasCanada/ManageTier3SW

Well tested straight forward to use

Atlas Tier 3 Workshop

- Jan 25-26 2010
  - http://indico.cern.ch/conferenceDisplay.py?ovw=True&confId=77057
  - Organizers Massimo Lamanna, Rik Yoshida, DB
  - Follow on to activities in the US the year before
  - Showed the variety of Tier 3's in Atlas
  - Good attendance from all across Atlas
  - 6 working groups formed to address various issues
    1. Distributed storage(Lustre/GPFS and xrootd subgroups)
    2. DDM – Tier3 link
    3. Tier 3 Support
    4. Proof
    5. Software and Conditions DB
    6. Virtualization