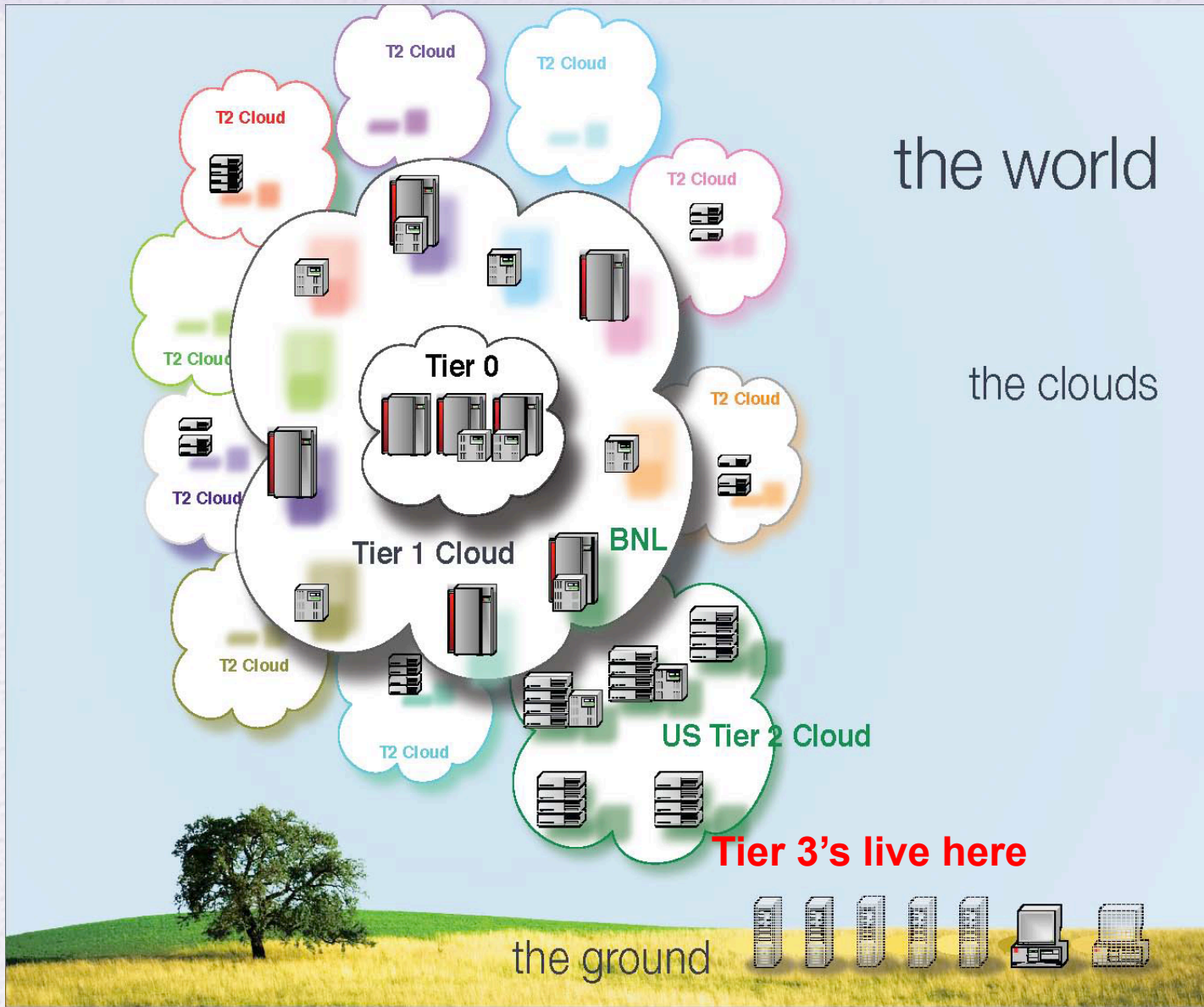




Atlas Tier 3

Doug Benjamin
Duke University
On Behalf of the Atlas Collaboration

Atlas Tier 3's



What is a Tier3?

- Working definition
 - “Non pledged resources”
 - “Analysis facilities” at your University/Institute/...
- Tier3 level
 - The name suggests that it is another layer continuing the hierarchy after Tier0, Tier1s, Tier2s...
 - Probably truly misleading...
 - 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)**

What is a Tier3?

- Comments:
 - No concept of size (small Tier3 vs big Tier2...)
 - Tier3s can serve (and be controlled by) a subset of the ATLAS collaboration (local or regional users).
- Non-pledged resources does not mean uncontrolled or incoherent
 - Need to provide a **coherent** model (across ATLAS)
 - Small set of template to be followed while setting up a Tier3 for ATLAS users.
 - Coherent because:
 - Guarantee no negative repercussions on the ATLAS Grid (service overload, additional complex manual support load) by the proliferation of these sites

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**

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
 - National Analysis facilities
 - Non-grid Tier 3 (Tier 3g) (most common for new sites in the US and likely through 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

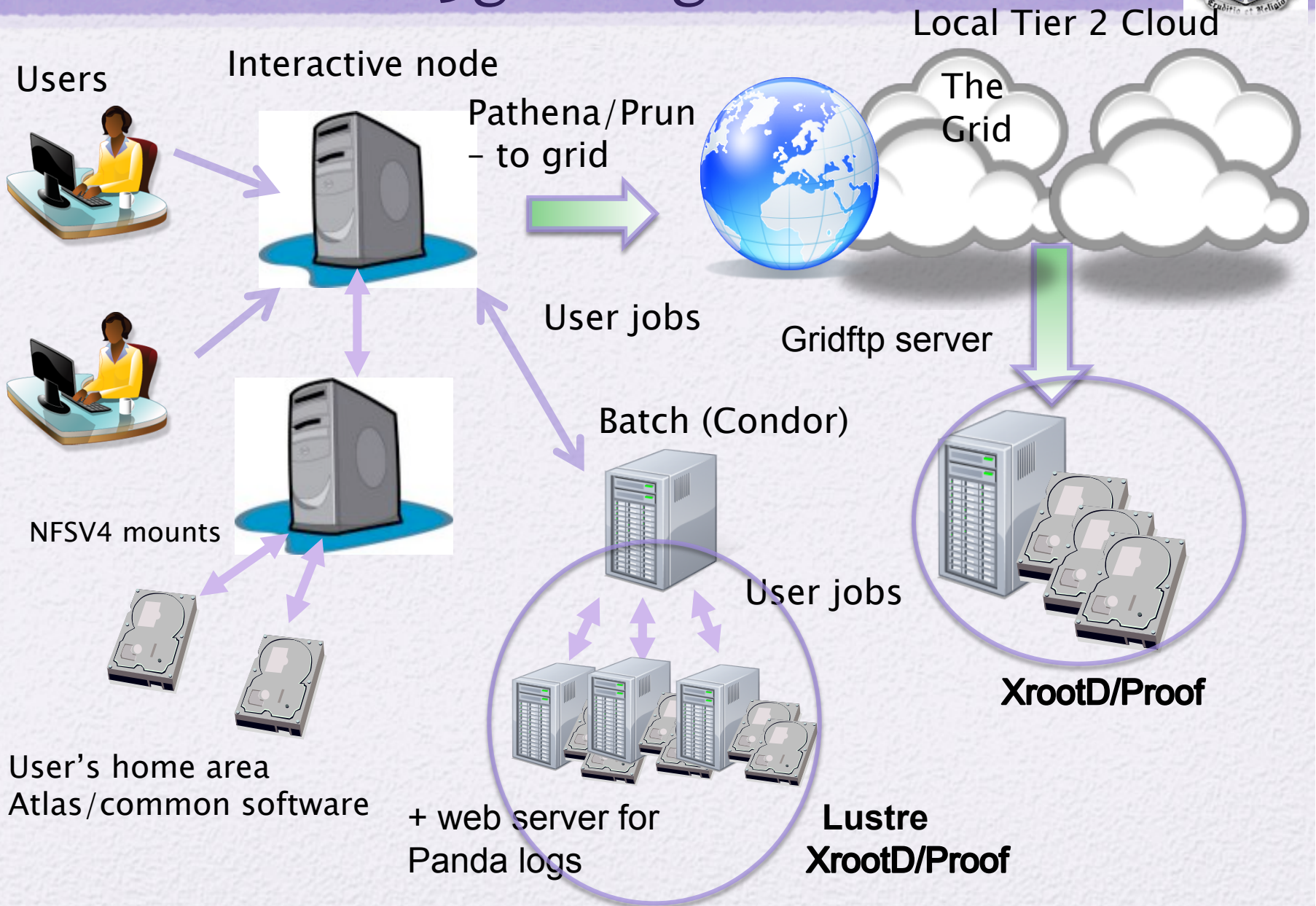
- Of course the recipe $Tier\ 3 = (small)\ Tier2$ could make sense in several cases
- But in several other cases:
 - Too heavy for small new sites
 - “Human cost”
 - The new model is appealing for small Tier2-like centre as well
- In all cases:
 - We got data! The focus is more and more on doing the 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.

- 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

Local Tier1 Tier2 Cloud

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)

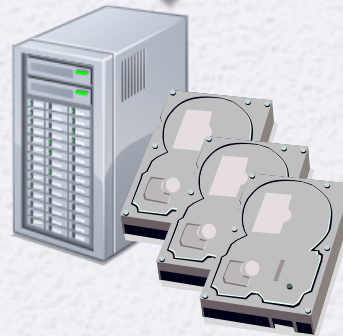
- Sites in DDM ToA will be tested frequently
- Troublesome sites will be blacklisted (no data) extra support load



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



Gridftp server



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 (Triumpf)
 - 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)**

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 help 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

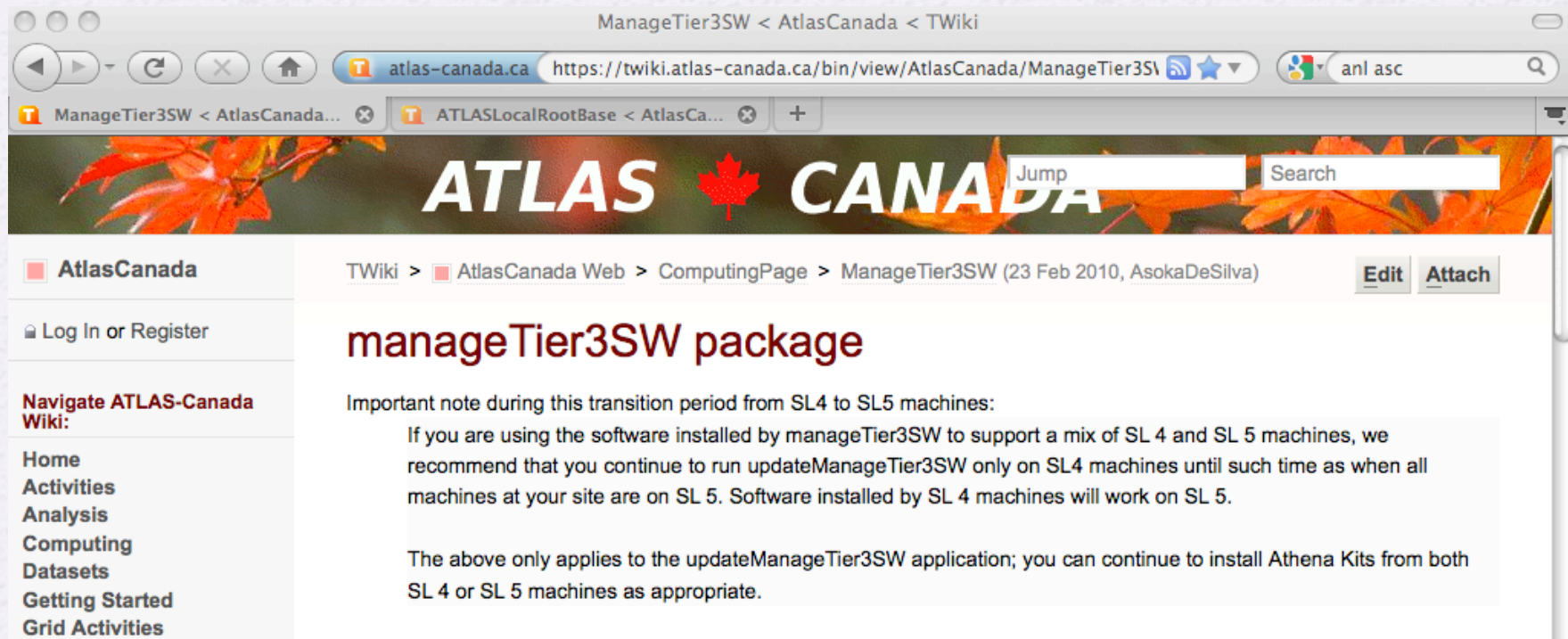
- Last June at WLGC Storage workshop
 - Atlas Tier 3 proposed alternative method for delivering data to Tier 3 using confederated XROOTD clusters
- Physicists can get the data that they actually use
- Alternative and simpler than ATLAS DDM
 - In testing now
 - Plan to connect Tier 3 sites and some Tier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

Conclusions

- Tier 3 computing important for data analysis in Atlas
- A coherent Atlas wide effort has begun in earnest
- 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 are moving up the computing chain

Backup Slides

- NFS file server
 - ManageTier3 SW package (Asoka DeSilva Triumph)
<https://twiki.atlas-canada.ca/bin/view/AtlasCanada/ManageTier3SW>



ManageTier3SW < AtlasCanada < TWiki

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

ATLAS CANADA Jump Search

AtlasCanada TWiki > AtlasCanada Web > ComputingPage > ManageTier3SW (23 Feb 2010, AsokaDeSilva) Edit Attach

manageTier3SW package

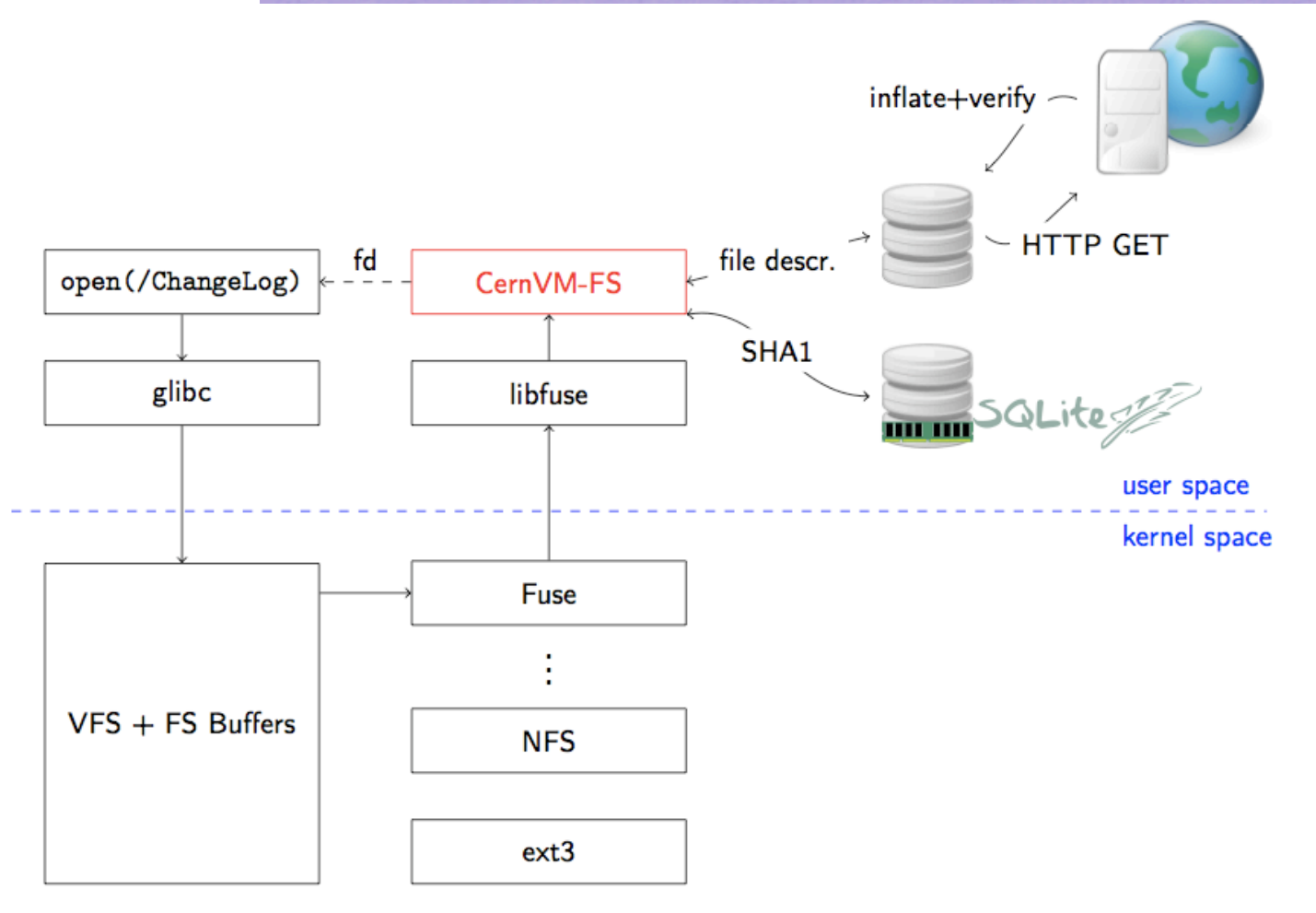
Important note during this transition period from SL4 to SL5 machines:

If you are using the software installed by manageTier3SW to support a mix of SL 4 and SL 5 machines, we recommend that you continue to run updateManageTier3SW only on SL4 machines until such time as when all machines at your site are on SL 5. Software installed by SL 4 machines will work on SL 5.

The above only applies to the updateManageTier3SW application; you can continue to install Athena Kits from both SL 4 or SL 5 machines as appropriate.

Home
Activities
Analysis
Computing
Datasets
Getting Started
Grid Activities

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