



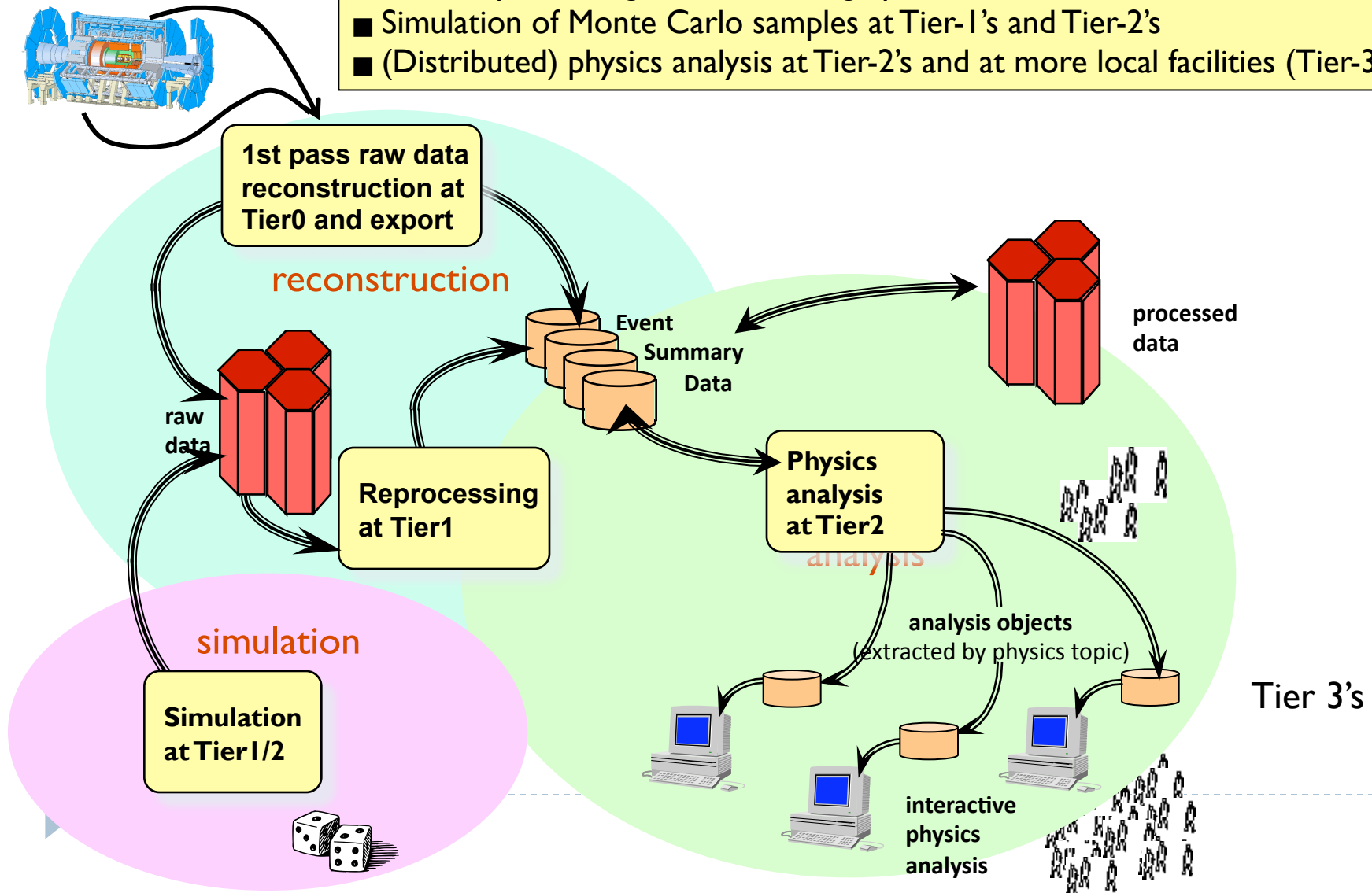
Atlas Tier 3 Plans

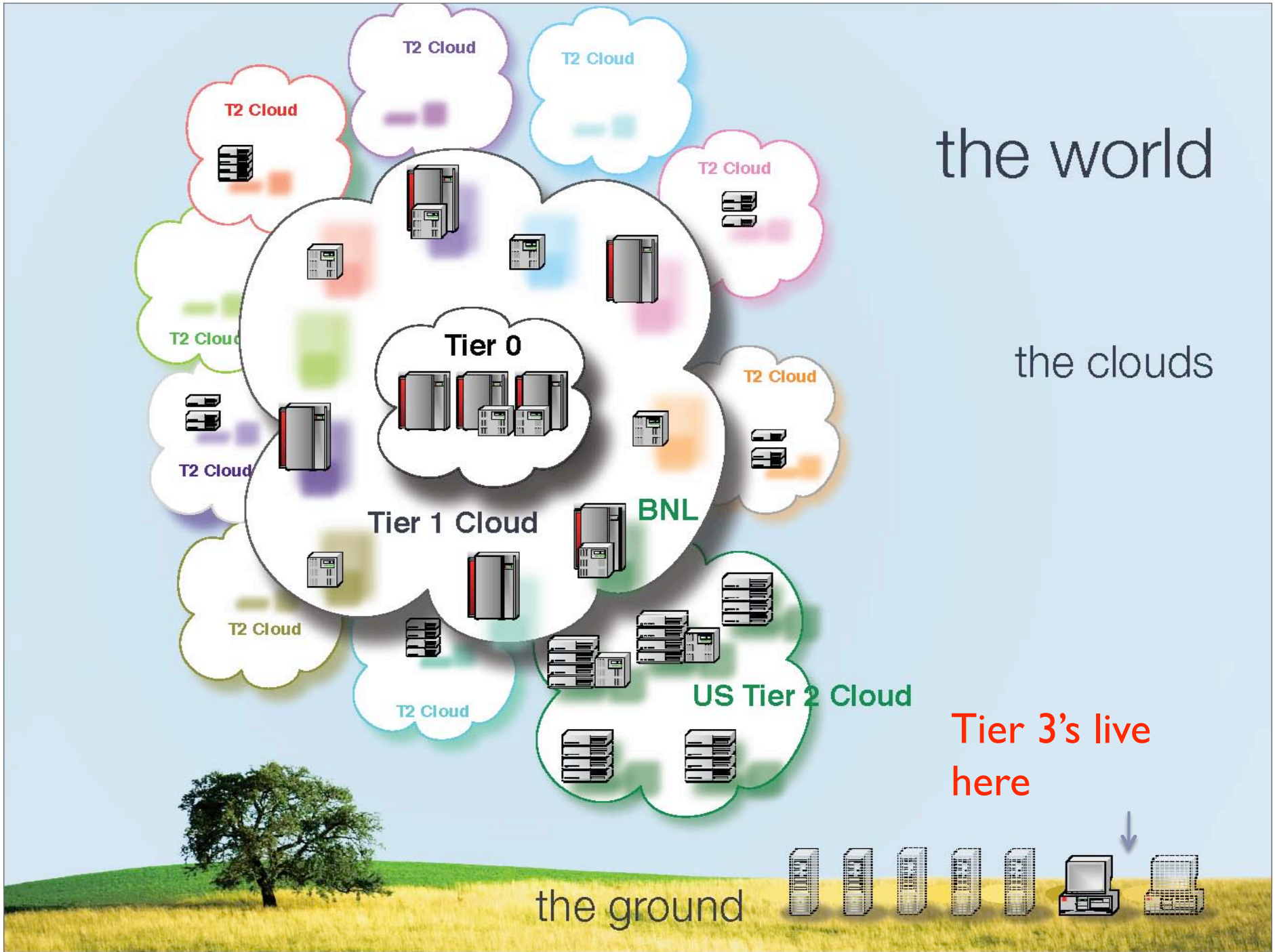
Doug Benjamin
Duke University

Simplified View - Atlas Computing Model

4 main computing operations according to the Computing Model:

- Initial processing of Raw data at CERN Tier0 - data export to Tier-1's/Tier-2's
- Data re-processing at Tier-1's using updated calibration constants
- Simulation of Monte Carlo samples at Tier-1's and Tier-2's
- (Distributed) physics analysis at Tier-2's and at more local facilities (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

Tier3: 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

- Of course the recipe $Tier3 = (small) Tier2$ 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:
 - Got the first collisions! The focus is more and more on doing the analysis than supporting computing facilities ;)

What is the difference between a Tier2 and a Tier3?

- Independently from the magic recipes for building and operating a successful ATLAS Tier3, we should consider the following directions:
 - **Maintaining grid services vs using grid clients**
 - **Tier3 as an independent layer** (with respect of the To/T1/T2 infrastructure)
- With the goal to guarantee:
 - Simpler and smoother (central) operations
 - Development point of view: Tier3s can be seen as an incubator of new solutions for the whole system
 - Note that the adoption and the integration of new technology is a major piece of work and we need deep understanding of new technologies before deploying them in the To/T1/T2 level

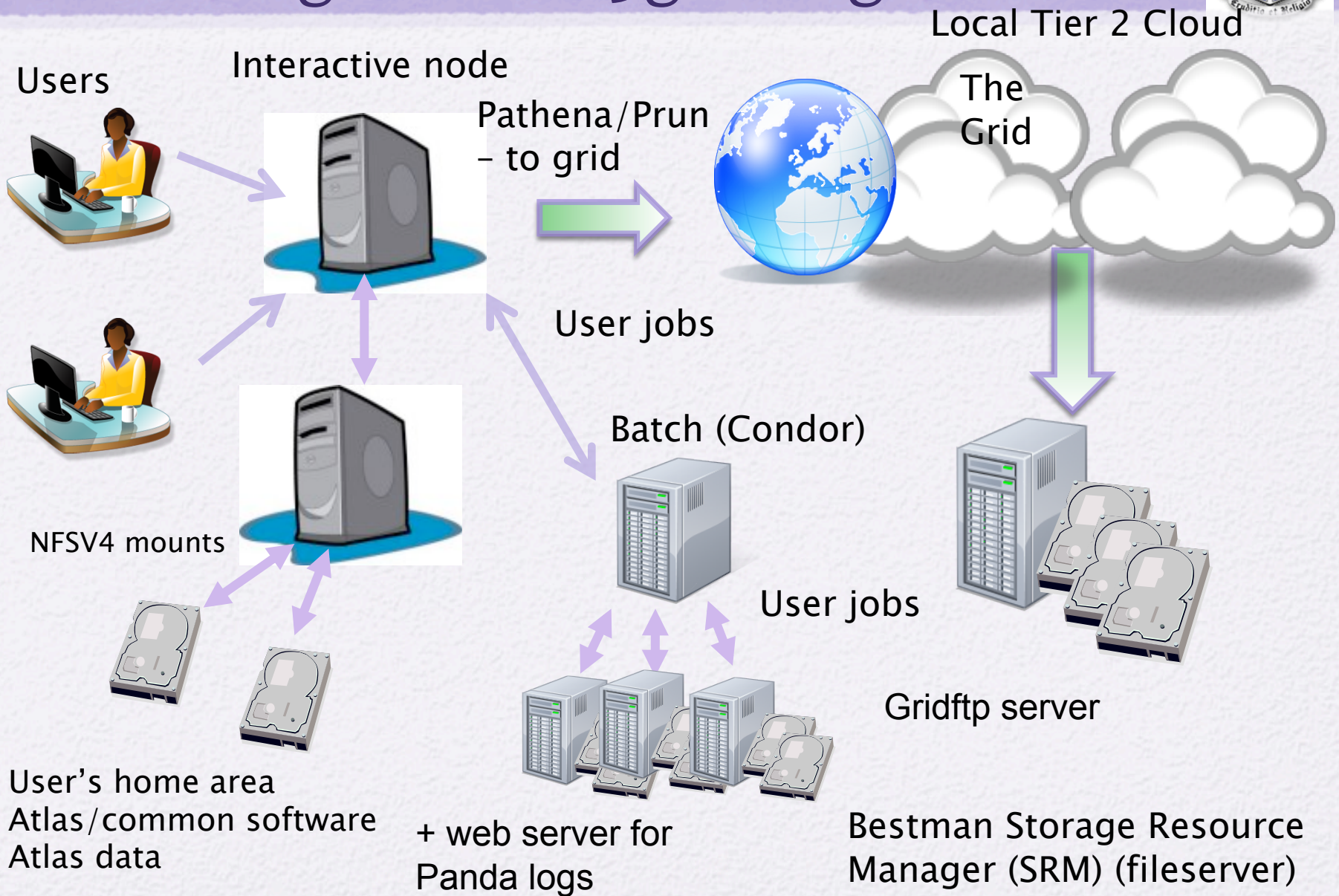
6 Tier 3 working groups

- **Distributed storage (Lustre/Xrootd/GPFS)**
- **DDM-Tier3 link**
- **Tier 3 Support**
- **PROOF Working Group**
- **Software / Conditions data Working Group**
- **Virtualization working group**

3-month time scale, Chaired by ATLAS persons

Open to experts (also from outside the collaboration)

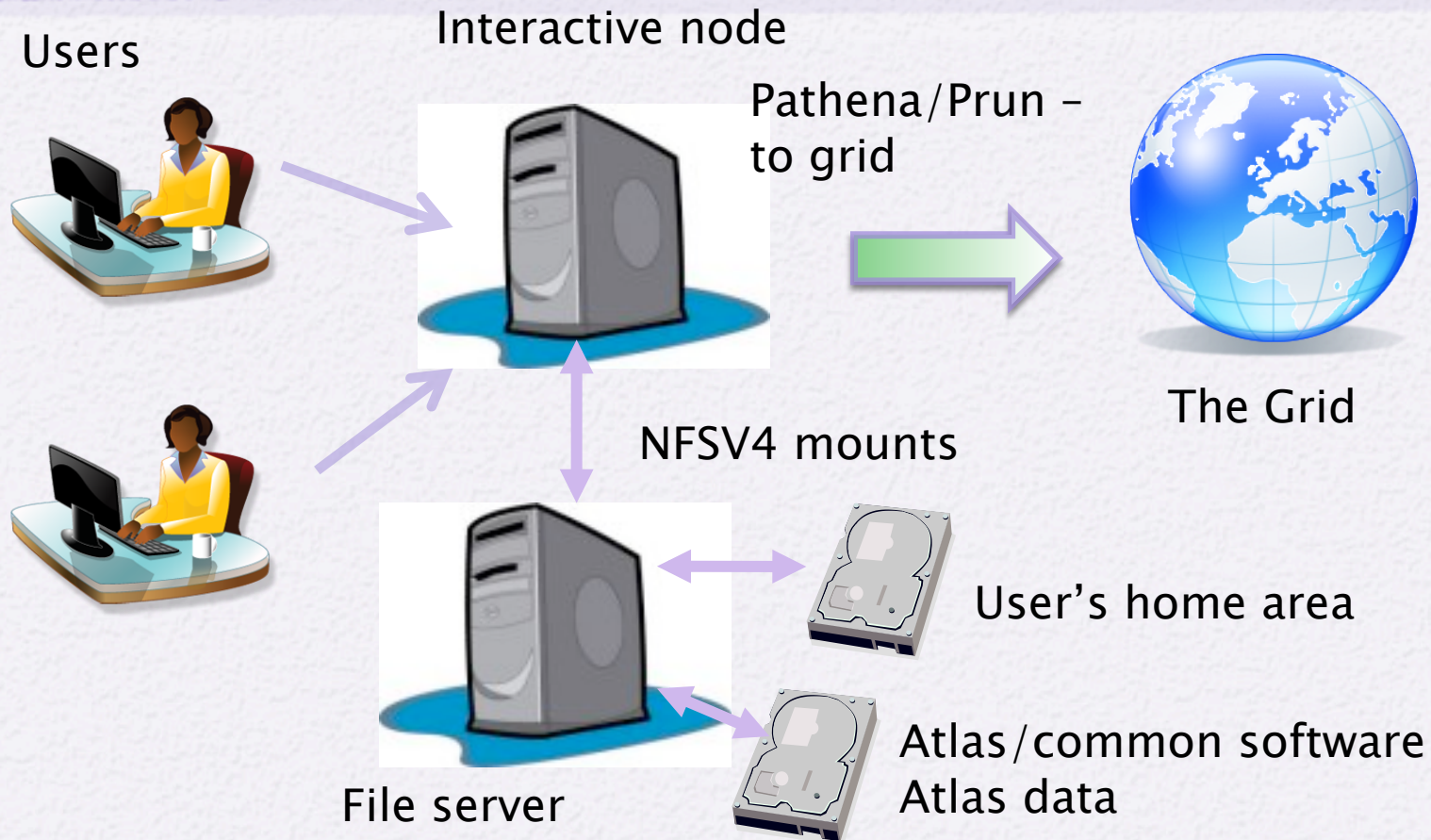
Non grid - Tier 3g configuration



Tier 3g design/Philosophy

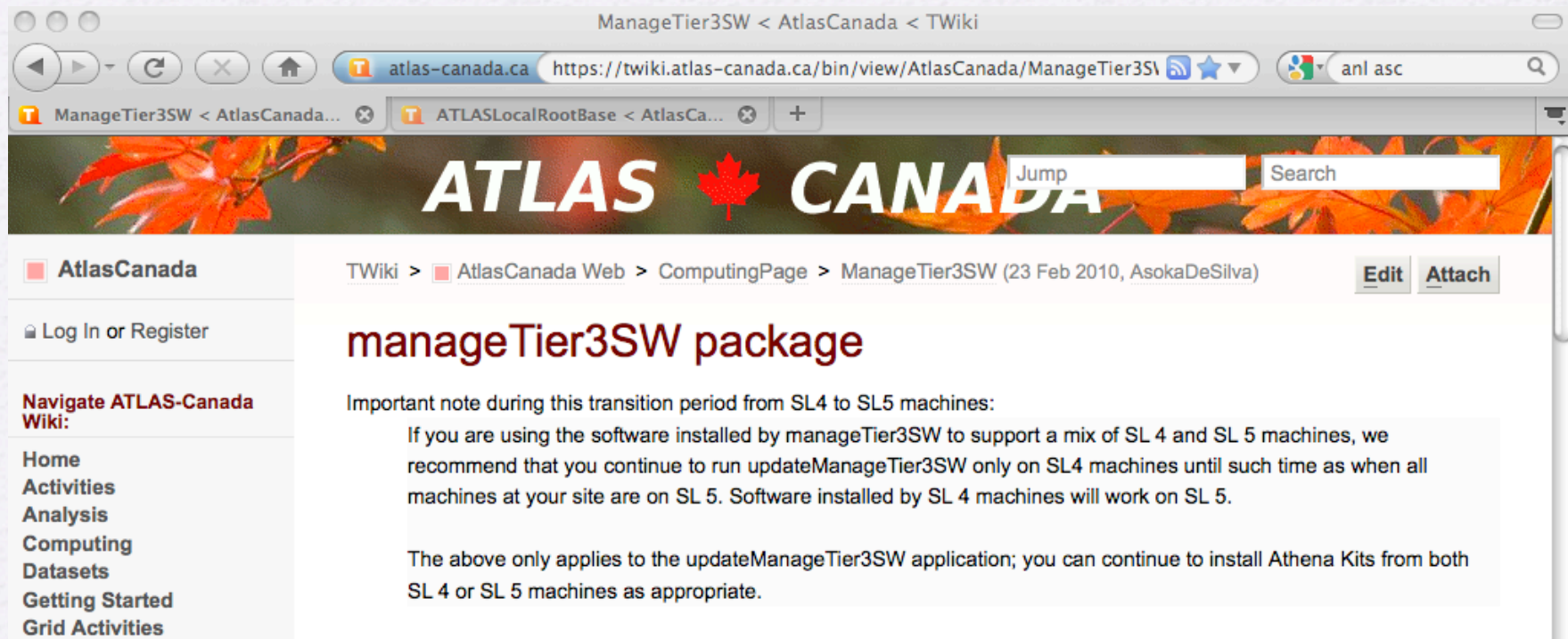
- Design a system to be flexible and simple to setup (1 person < 1 week)
- Simple to operate - < 0.25 FTE to maintain
- Scalable with Data volumes
- Fast - Process 1 TB of data over night
- Relatively inexpensive
 - Run only the needed services/process
 - Devote most resources to CPU's and Disk
- Using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

Tier 3g – Interactive computing



Common User environment (next slide)
 Atlas software installed (two methods)
 manageTier3SW
 Web file system CVMFS

- 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

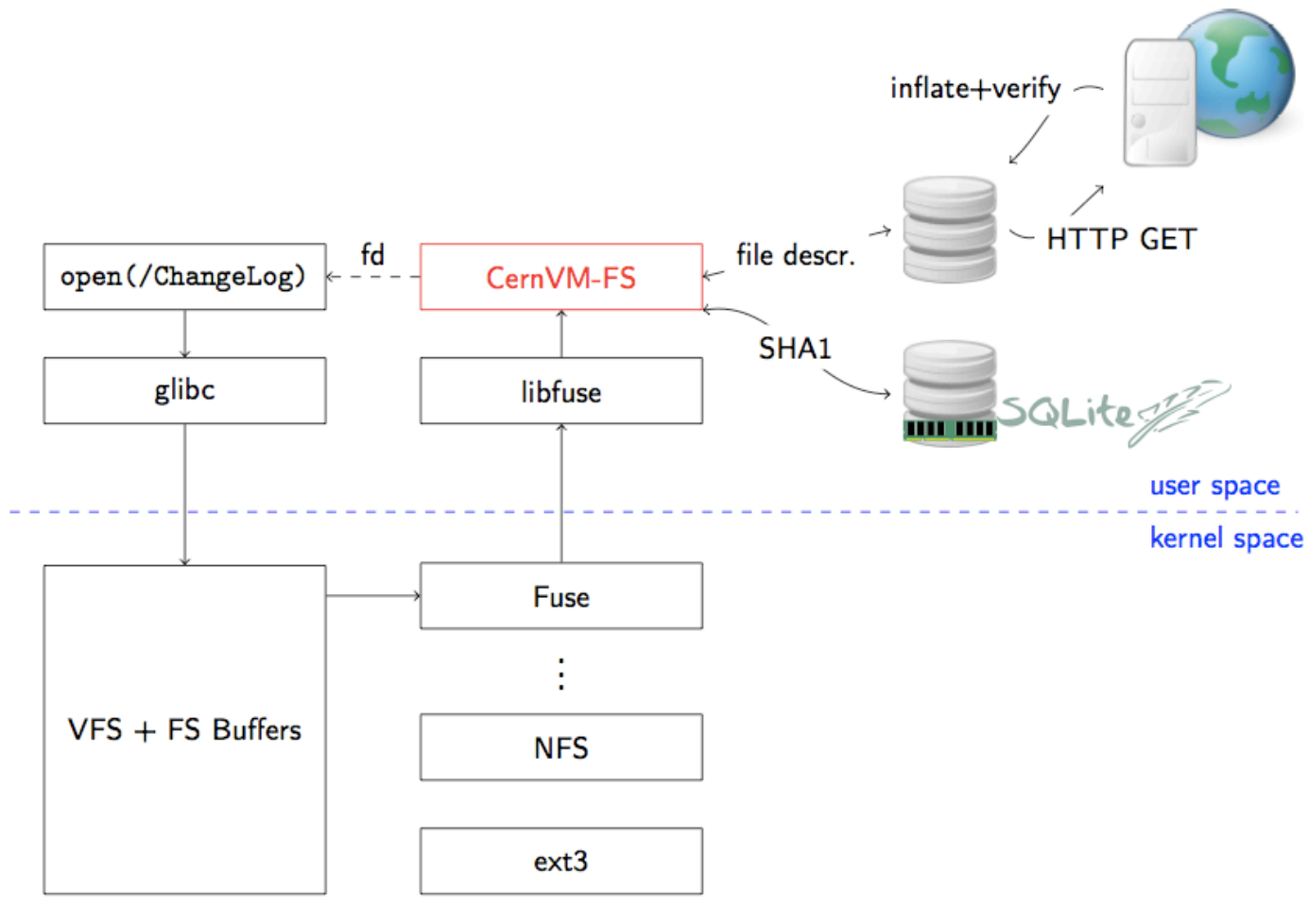
manageTier3SW package

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.

Well tested straight forward to use





NFS V4 vs CVMFS Comparison



Athena Compilations

Rik Yoshida (ANL)

Dell R710: 8 cores (16 hyperthreaded)

No. Simultaneous Condor jobs:	1	4	8	14
NFS4	7 min	15 min	60 min	
CVMFS2	7 min		8 min	11 min

ATLASLocalRootBase

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

- Can easily setup a tested suite of software needed for work in a Tier 3



The screenshot shows the ATLAS Canada Wiki page for the ATLASLocalRootBase package. The page header features the ATLAS CANADA logo with a red maple leaf and a search bar. The breadcrumb trail is: TWiki > AtlasCanada Web > ComputingPage > ATLASLocalRootBase (01 Mar 2010, AsokaDeSilva). The page title is "ATLASLocalRootBase package". The main content area lists several links: "ATLASLocalRootBase package", "Brief description and usage of the ATLASLocalRootBase package", "How-to" (with sub-links for "How do I install?", "How do I use a snapshot?", and "runKV"), "CernVM users", and "Release Notes". A highlighted box at the bottom of the content area contains the text "Brief description and usage of the ATLASLocalRootBase package". The left sidebar contains navigation links for AtlasCanada, Log In or Register, and a list of categories including Home, Activities, Analysis, Computing, Datasets, Getting Started, Grid Activities, Meetings, People in ATLAS Canada, User Services, and Travel To CERN.

Developed by Asoka DeSilva



Tier 3g User environment



- ANL ASC cluster configured with this User environment.

- User guide contains information

<https://atlaswww.hep.anl.gov/twiki/bin/view/UsAtlasTier3/Tier3gUsersGuide>

Add to your .bashrc file

```
export ATLAS_LOCAL_ROOT_BASE=/export/share/atlas/ATLASLocalRootBase
```

```
alias setupATLAS='source ${ATLAS_LOCAL_ROOT_BASE}/user/atlasLocalSetup.sh'
```

- Interactive use (to setup an athena version)

```
setupATLAS      # Some info output after this command
```

```
localSetupGcc --gccVersion=gcc432_x86_64_slc5 # Sets the compiler version
```

```
export ATLAS_TEST_AREA=<some area>/15.6.6 # defines your test area (note vers. #)
```

```
source /export/home/atlasadmin/temp/setupScripts/setupAtlasProduction_15.6.6.sh
```

- To see what other software is available

```
showVersions
```

How data comes to Tier 3g's

US Tier 2 Cloud

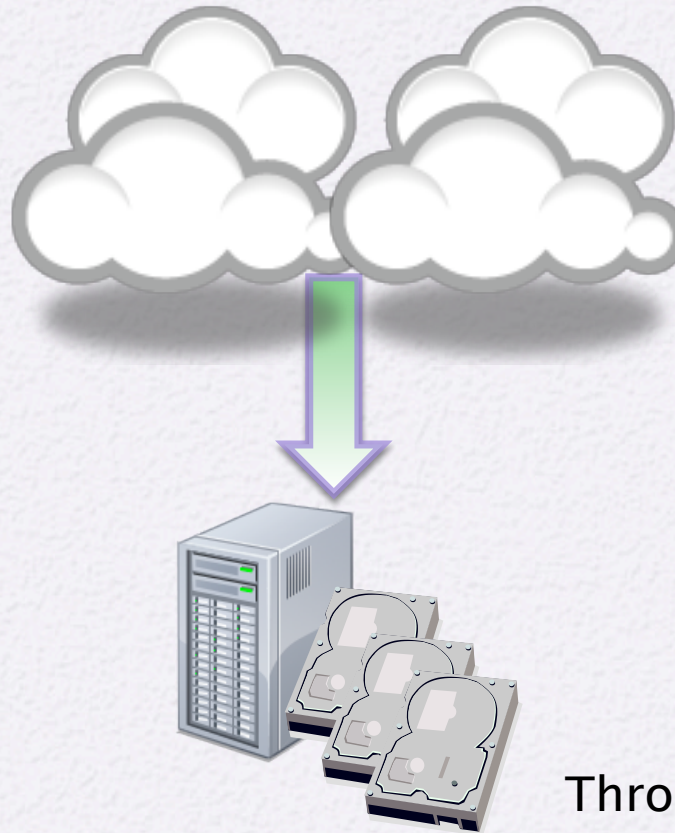
Two methods

- Enhanced dq2-get (uses fts channel) (available fairly soon)

- 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

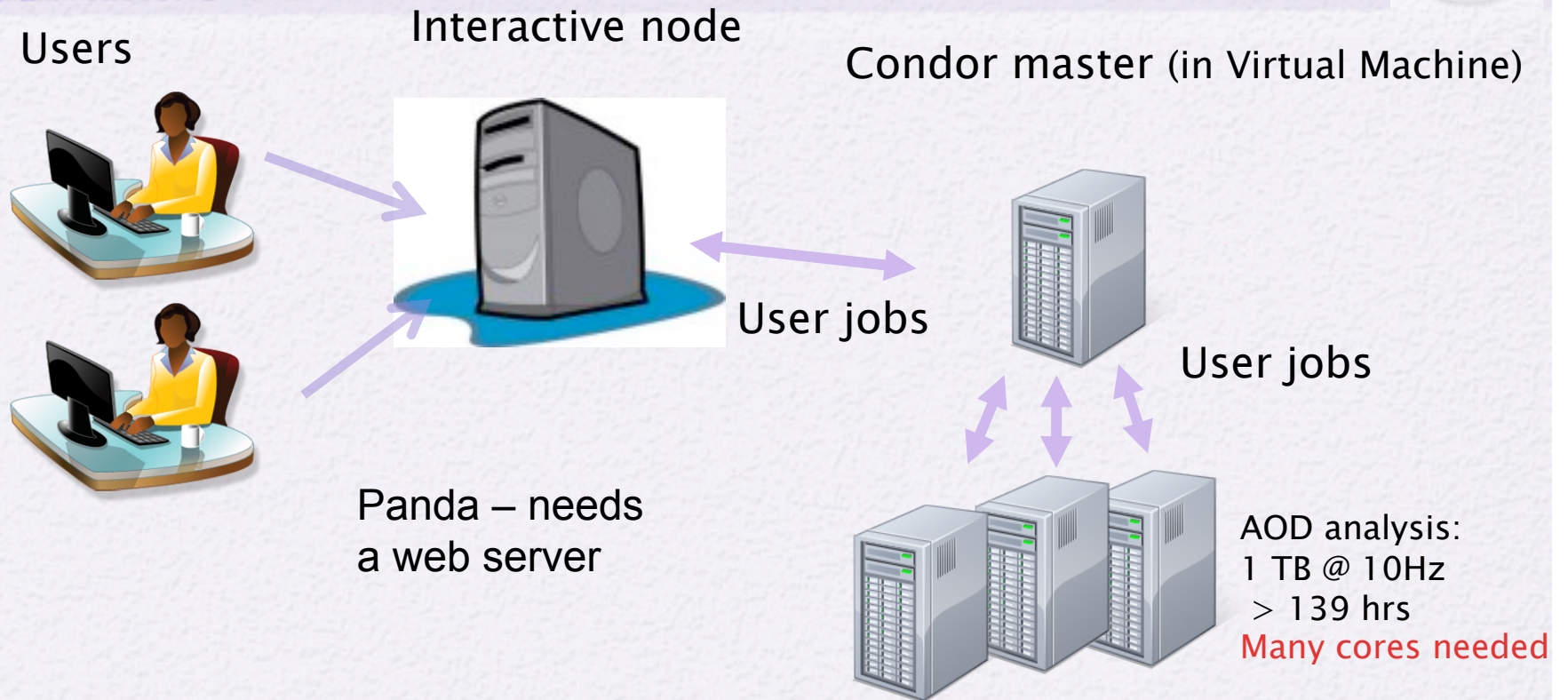


Data will come from **any** Tier 2 site

Throughput test with ANL SE - (> 500 Mb/s)

Shows \$1200 PC (Intel i7 chip/ X58 chipset/ SL5.3) can be a SE for a small T3.

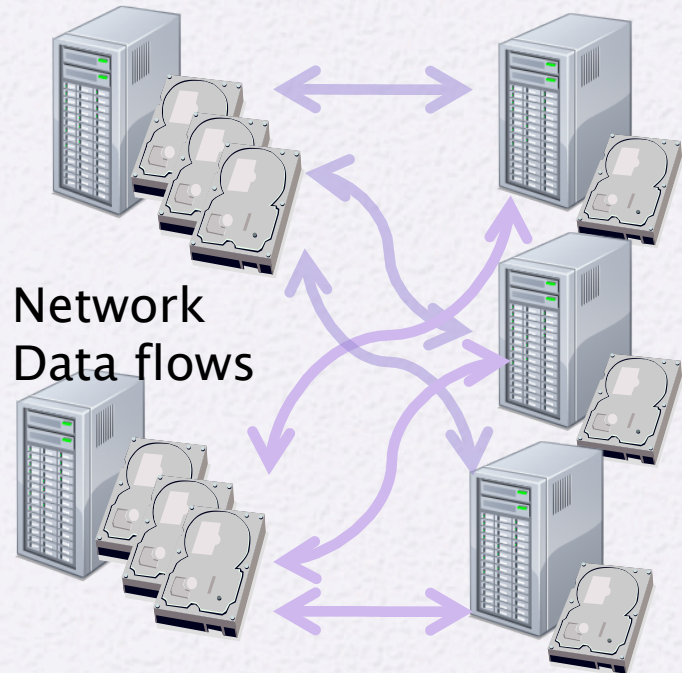
Tier 3g – Batch/ Distributed computing



- ✧ Common user interface to batch system simplifies users' work
- ✧ Panda being testing in Tier 3g (Duke and ANL ASC)
 - ✧ Torre is writing the instructions now
- ✧ ANL has developed such an interface **ARCOND**
 - ✧ Well tested on their system
 - ✧ Will need to be adapted for Xrootd storage

Tier 3g – Data storage options

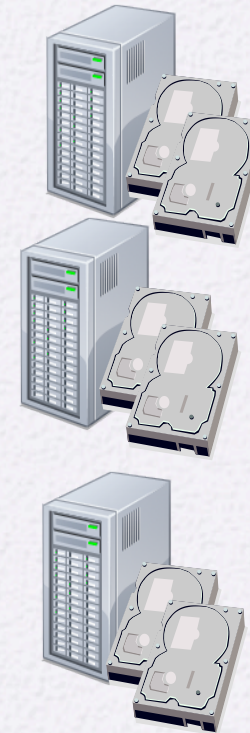
Storage on worker nodes



Network
Data flows

File servers

Worker nodes with
little local storage



XRootD can be used to manage either
type of storage

Ntuple processing in Batch

- Ran Sergei Chekanov's standalone ntuple analysis
 - ~20 M events in 100 files (random order)
 - Data in xrootd system on two nodes
 - 14 jobs at a time

Processing Node (Events/sec)			
Node containing data files	Avg rate per job	ascwrk0	ascwrk1
	ascwrk0	3480	4770
	ascwrk1	4786	5626

Network processing
 ganglia 100 MB/s
 No I/O wait seen
 Data node only serving data

ascwrk0 – 4 disk raid 5
 ascwrk1 – 6 disk software Raid 6

Local processing
 (effect of # disks seen)
 I/O wait seen on ascwrk0

Dell R710: 8 cores (16HT)

Number Simultaneous Condor Jobs	1	4	8	14
NFS4	11 min	12 min	14 min	19 min
Local Disk	15 min	13 min	14 min	19 min
XDR local disk	11 min	11 min	13 min	18 min
XDR* remote disk	14 min	16 min	43 min	

* 2 jobs out of 13 jobs had a read error

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