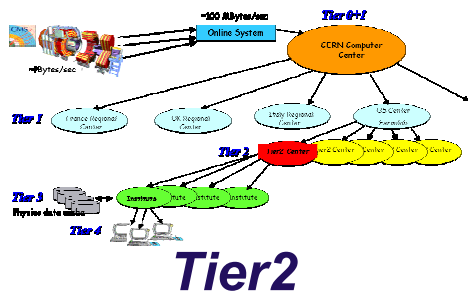


Experiences at the Tier2

Ian Fisk

LCG Launching Meeting, CERN

March 12, 2002

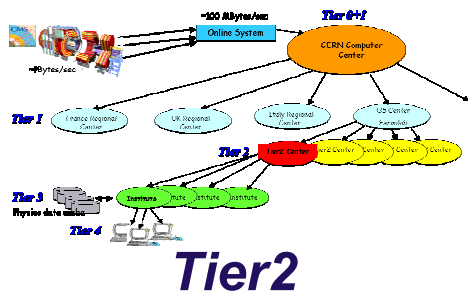


MONARC and Grids

- ◆ MONARC Model calls for a hierarchical computing model
 - ➔ 1/3 of computing resources are located at Tier2s
 - ➔ Specific kinds and volumes of data is stored at each type of center
 - ➔ Specified responsibilities

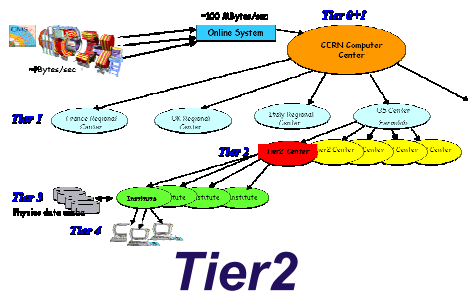
- ◆ A number of Grid models call for greater data transparency
 - ➔ Smart schedulers (moving jobs to data or data to jobs)
 - ➔ Global Data Replica Catalogs
 - ➔ Access to very large datasets from the desktop

- ◆ Basically a lot of promises
 - ➔ Spend today talking about the Tier2 experience so far.



Challenges at the Tier2 Centers

- ◆ Planned to run with very small staffs
Need to find ways to work more efficiently
- ◆ Generally planned for universities and research institutes without large IT staffs and resources common to CERN
Need to rely on prototyping and development of innovative technologies
- ◆ Limited budgets
Need to find ways of leveraging Grid, University, and Lab resources
- ◆ In the final system represent 1/3 of the total computing resources.

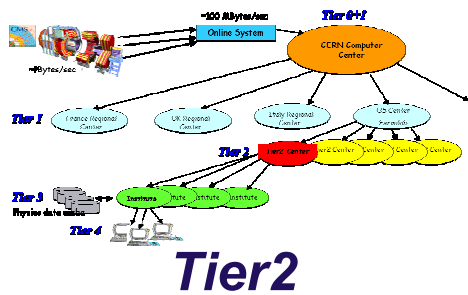


Benefits of the Tier2 Centers

- ◆ Bring computing resources close of user communities
 - ➔ Provide dedicated resources to regions (of interest and geographical)

- ◆ Leverage Additional Resources, which exist at the universities and labs
 - ➔ Reduce computing requirements of CERN.
 - ➔ Help meet the LHC Computing Challenge

- ◆ Provide diverse collection of sites, equipment, and expertise for development and testing
 - ➔ Provide much needed computing resources

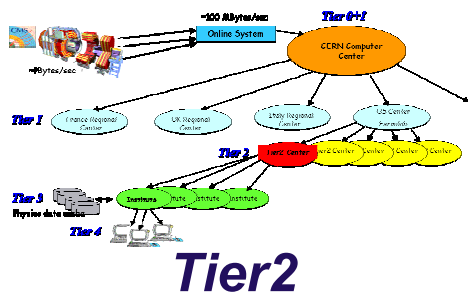


Staffing at Tier2 Centers

- ◆ As an example, US-CMS plans for greater than 30 people to staff the Tier1 facility at Fermilab
 - ➔ 5 US Tier2's get 2 people each (originally 1.5)

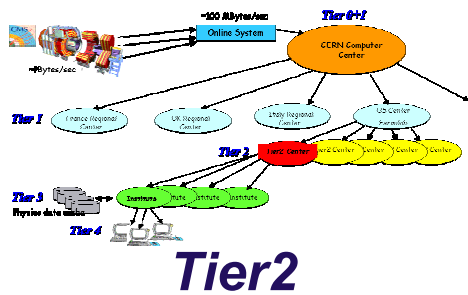
Problem: How do you run a center with greater processing power than ?? With only two people?

- ◆ At US Tier2 centers we have been investigating cluster management software to reduce the personnel required to run the center



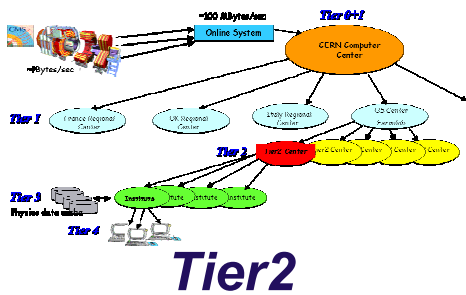
◆ US Tier2s have been investigating Rocks, which is a collection of open source tools for automating commodity clusters developed by the National Partnership for Advanced Computing Infrastructure.

- ➔ When new hardware is enabled the Rocks “frontend” node recognizes it and assigns tracking info
- ➔ A small partition is made on the primary disk, which contains a kernel
- ➔ This kernel can be used to completely reinstall the operating system.
- ➔ Selected pieces or the entire cluster can be upgraded, downgraded, or modified remotely in a matter of minutes.
 - Changing linux versions for a 50 node cluster is a 15 minute exercise requiring one command
- ➔ Progress can be monitored and intervention taken through a network keyboard/monitor port



Cluster Management

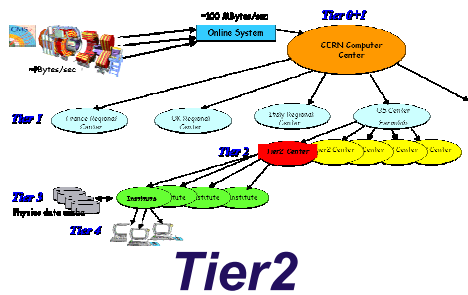
- ◆ Rocks has made basic system administration of the tier2 cluster a fraction of one person's time.
- ◆ Ensures that all computation nodes are identically configured
- ◆ Is nearly ideal for multi-experiment installations, where entire software configurations (including the operating) system may need to be swapped out.
- ◆ Configuration modules are stored as XML
 - ➔ Possible to create custom configurations for elements in the cluster is straightforward
- ◆ More information is available from
<http://rocks.npaci.edu>



Cluster Management Example

Normally at this point I reinstall a few nodes, but my cluster is currently busy with CMS event production, so a few screen shots.

- ➔ In CMS Tier2's prototypes already represent approximately 1/3 of the production capacity

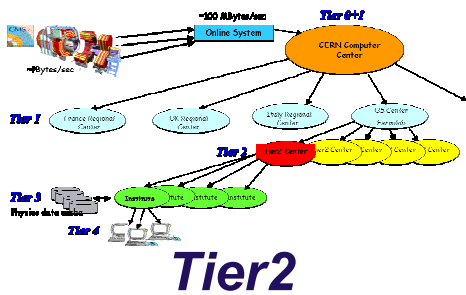


Dealing with lack of infrastructure

- ◆ Tier2 centers are not foreseen to necessarily have tape systems
 - ➔ All tier2 datasets are re-derivable but painful to do so
 - ➔ Already using Grid developed replication prototypes to replicate data to Tier1's and CERN. Have used GDMP for data replication.

- ◆ Networking a bottleneck
 - ➔ Working on optimizing what we have
 - ➔ Fortunately getting cheaper and faster

- ◆ Prototype efforts in US, Italy, UK, Russia, etc. very useful
 - ➔ Training and building expertise



Dealing With Limited Resources

- ◆ The Prototype effort has been supplemented with Grid, University and Lab funds
- ◆ A strong program of evaluating lower cost commodity hardware.

→ Network switches

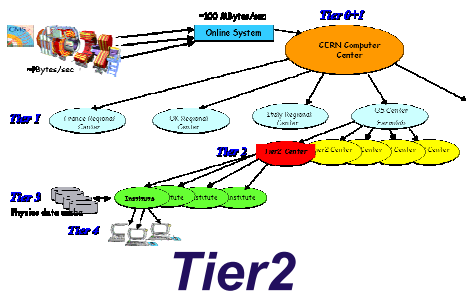
- Examined lower cost HP and Dell switches for small installations compared to higher priced Sisco gear
- Found for many small installations

→ RAID Arrays

- Studied very high end SCSI based RAIDs compared to inexpensive 3ware and promise IDE based arrays
- Performance gap rapidly closing still evaluating reliability

→ CPU's

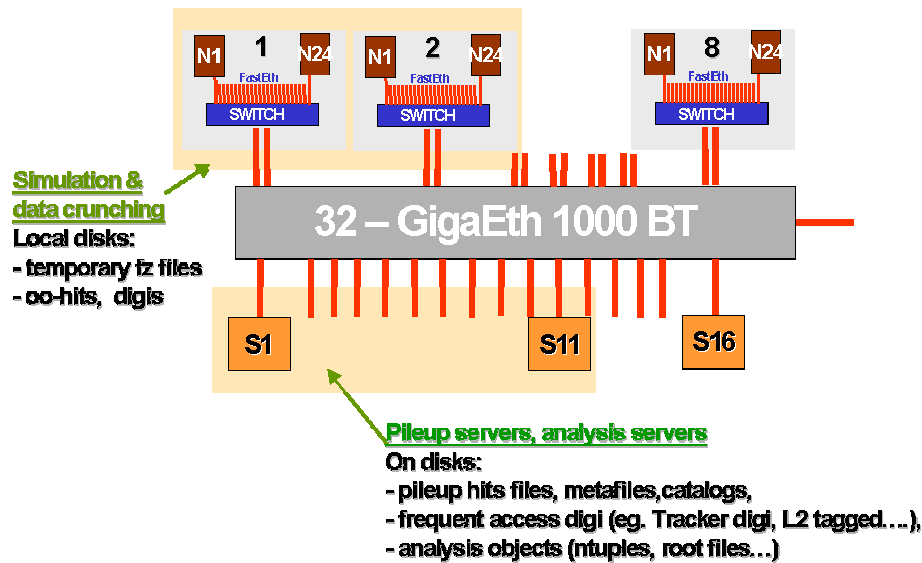
- Application based benchmarking program

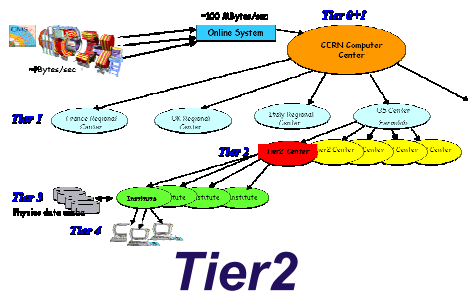


Prototype Development Continues

INFN Network lay-out optimized for CMS Event Production

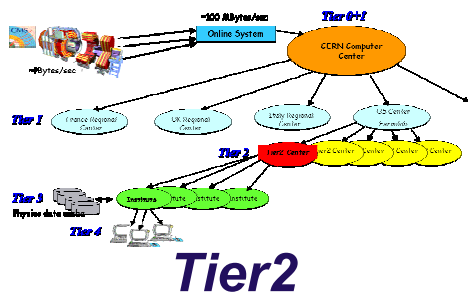
First Tier2 Prototype in US





Accomplishments

- ◆ Tier2's already completing computational contribution in some LHC experiments
 - ➔ Improving robustness of experiment code
 - ➔ Working on optimization in diverse environments
- ◆ Working with Grid projects to provide testbeds, evaluation, feedback, development
 - ➔ Using some grid software now
 - ➔ Helping to test and specify requirements
- ◆ Using Tier2 facilities and personnel to help prototype data access tools to achieve the transparency promised to remote users by the Grid.



For the Future

Tier2's

- ◆ will deliver on the promised computing resources
 - ➔ Will contribute more than they're funded for by taking advantage of existing resources
- ◆ will be the primary location for Grid development and evaluation.
- ◆ will bring computing resources close to user communities
- ◆ should change the ability of remote users to contribute
 - ➔ but is going to require a lot of work.