

# MIT Tier2 site report



Max Goncharov  
March 23, 2015

Y. Iiyama L.Di Matteo, C.Paus, M.Tiernan

USCMS Tier2 workshop



- **Hardware Infrastructure**
  - location
  - cluster hardware
  - network configuration
- **Cluster configuration**
- **Expansion plans**
- **Projects in development**



## MIT Tier2 at Bates (20 miles from main campus)

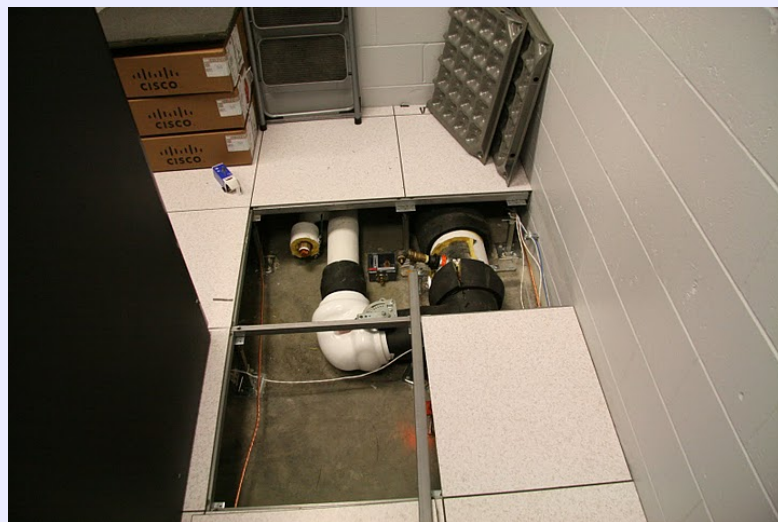
- Cooling, power, infrastructure managed by Bates personal
- Supported 24/7/52
- UPS backup for all servers (that are not worker nodes) (2 racks)
- 30 water cooled racks, rack – 47 U and 10 kW
- Network managed by MIT IS&T

## Cluster hardware

- Nodes – Dell 2U multiple core with 6/8/11 disks for storage (Hadoop)
- Dell PowerEdge 2950, R710, and R720XD

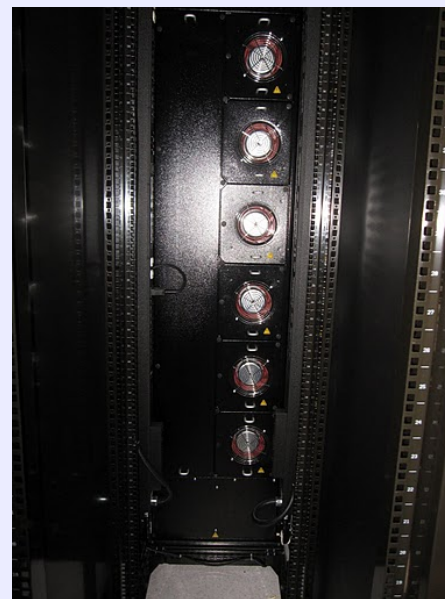
## Water Cooling

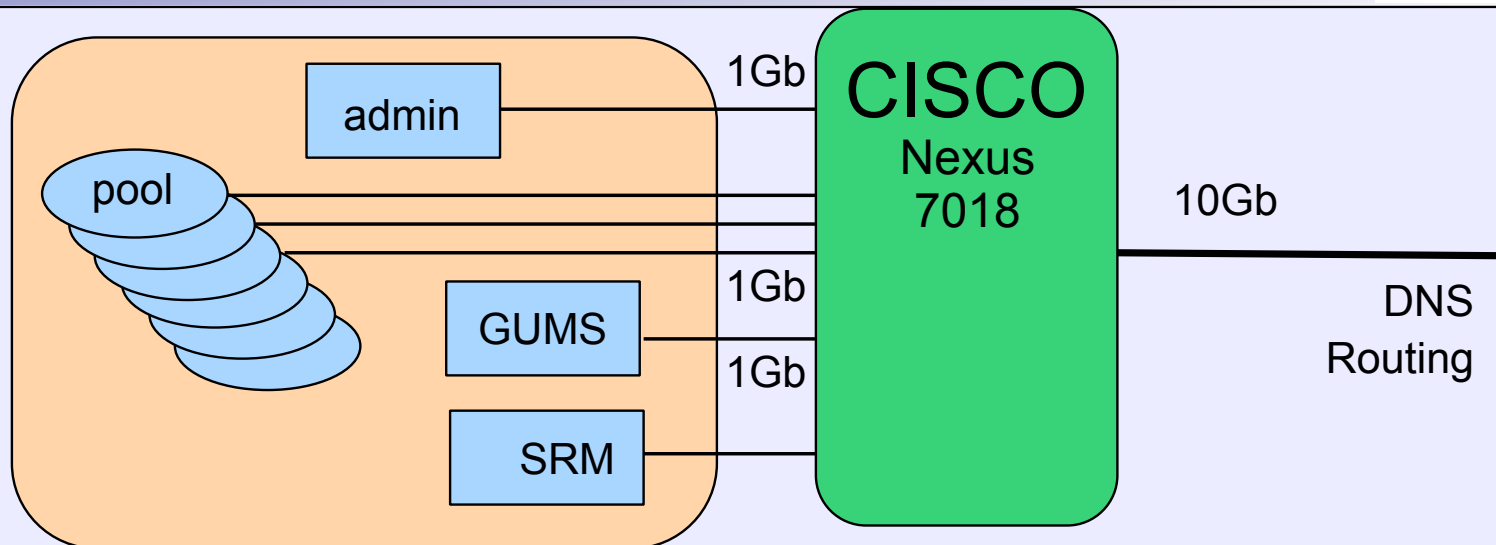
- infrastructure on the ground floor
- interlock in case of significant leak



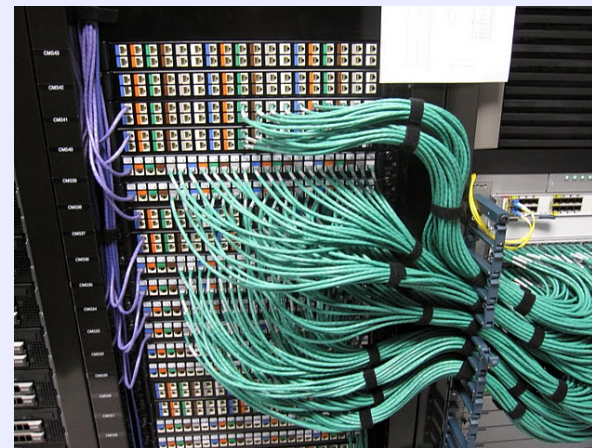
## Water Cooled Racks

- 47 U, 10 kW of power
- 90% filled, room for expansion
- 2 UPS racks for central servers
- UPS support is on the ground floor

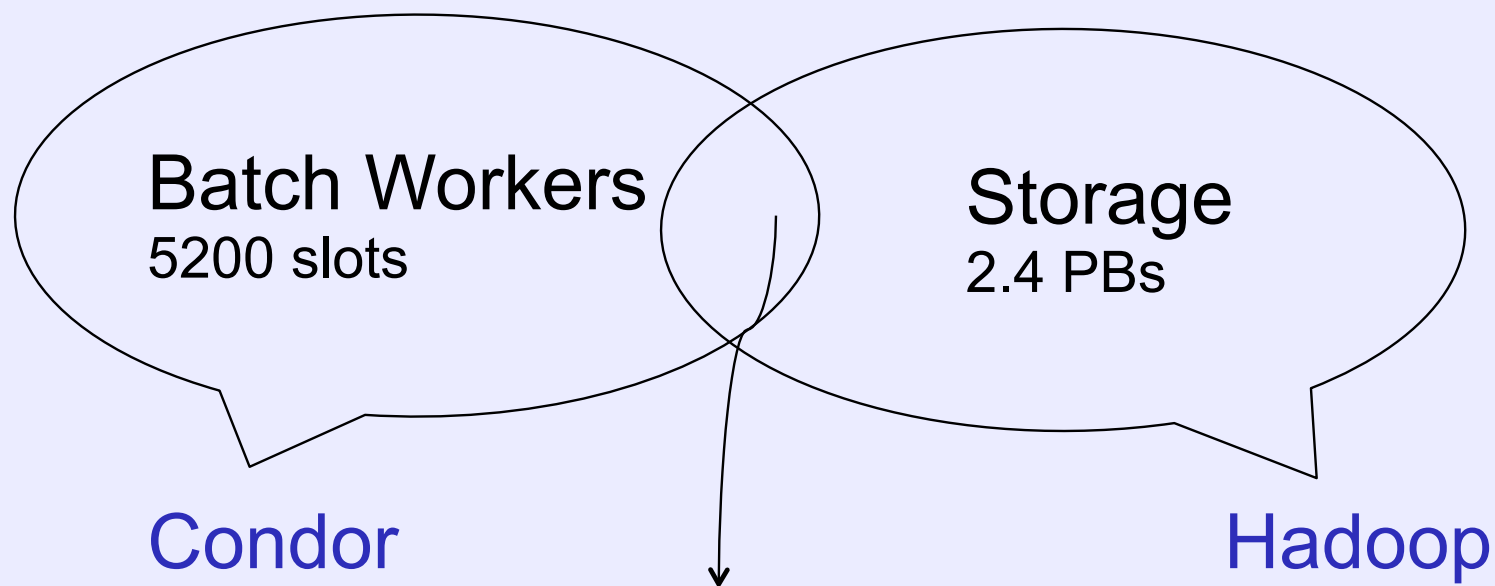




- Network managed by MIT IS&T
- 10 Gb Fiber link to the outside world
- CISCO is on separate UPS power
- Machines can talk at 1Gb through copper links
- Some worker nodes serve as gridFTP doors



## CPU/Storage Mix Model



Hardware Overlap – 99%  
Condor Slots/Worker Node ~ 8/12/16/24



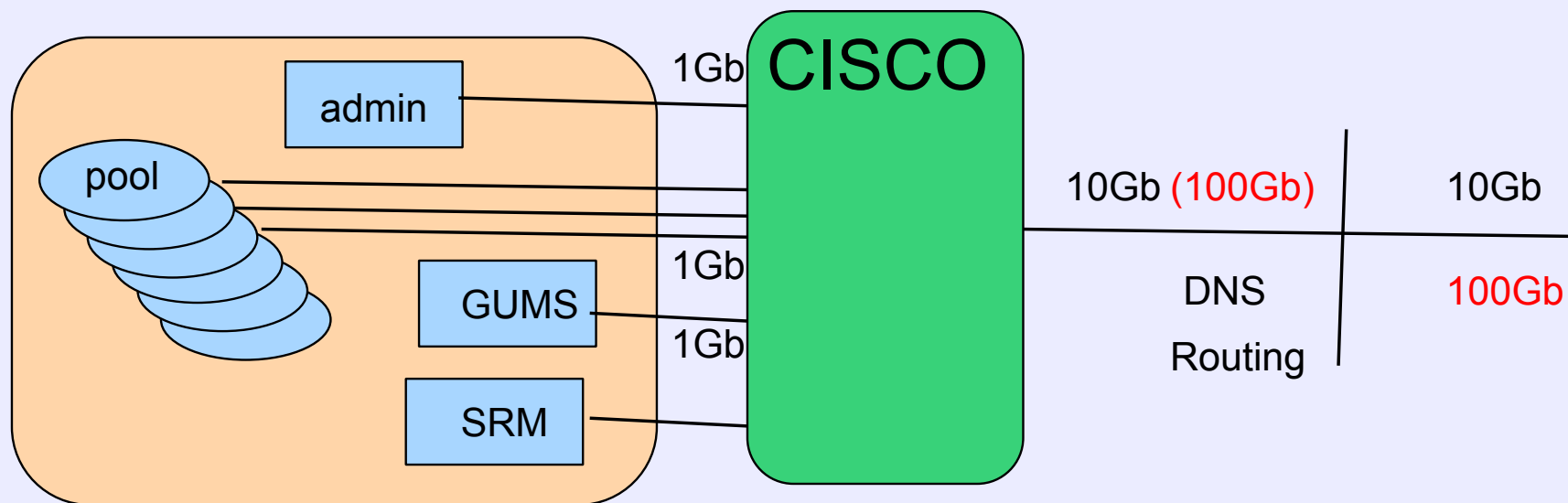
## Since last meeting a year ago

- Expanding the cluster hardware using same model of mixing CPU/Storage
- Last year:
  - storage expanded by 0.5 PBs for the total of 2.4 PBs
  - added ~960 condor slots for the total of 5200 (will be ~5500)
  - upgraded ~90 machines to double memory and hence double 4 GB memory Condor slots
- Central services upgrade
  - all WNs switched from SL5 to SL6, started to use Puppet
  - HTCondor, OSG 3.2, Condor 8.2, ...



- Network switch to 100 Gb
- Separation of local users
- Smart cache and Cinderella

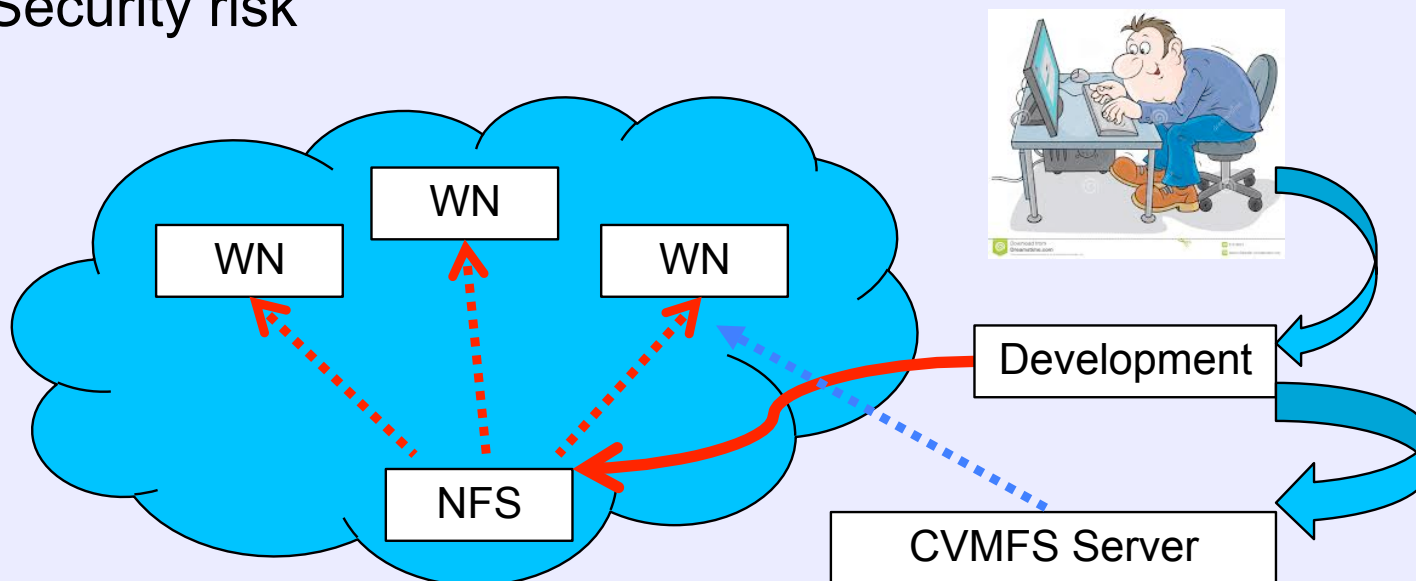


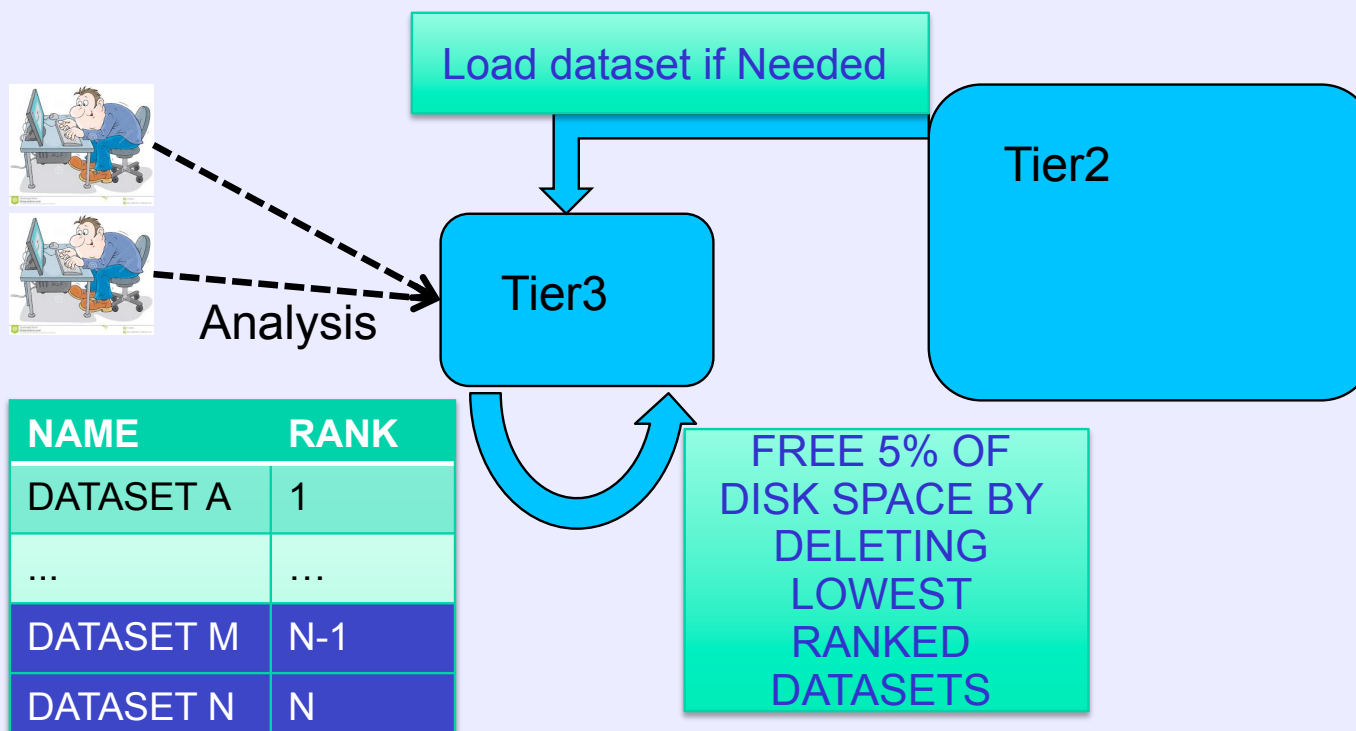


- Upgrade outside link to 100 Gb
- Upgrade Cisco Nexus switch to accept 100 Gb
- Decide if we put xrootd and gridFTP servers on 10 Gb links

## Local Users

- CMS HI community, Neutrino, and Dark Matter experiments
  - users get local accounts, can login on various machines
  - groups have dedicated machines that run Condor master and export libraries through NFS to worker nodes.
- Careless NFS use can/did hang the whole cluster
- Security risk





- Users run analysis jobs on Tier3
- Tier3 can't host all possible datasets, but Tier2 can
- Automated procedure to provide needed datasets
- Automated procedure to ensure there is free space on Tier3



- **Hardware expansion**
  - continue with mixing CPU and storage
  - rack space in host facility available for the nearest future
- **Network upgrade**
  - switch to 100Gb
- **Projects**
  - put local users onto CVMFS (before data taking starts)
  - SmartCache (done)