TEXAS TECH UNIVERSITY™

High Performance Computing Center

Texas Tech University (TTU) – Big Tier 3

OSG Site Administrators & CMS Tier 3 workshop

10th August, 2010
ACCRE, Vanderbilt University

Poonam Mane
Graduate Assistant TTU, HPCC

Outline and goals of this talk

- IT infrastructure
- Introduce people active in this area at TTU
- Describe range of resources available at TTU, and the fraction of those through the TTU high energy physics group for use in our Tier-3 CMS installation
- Seek to identify areas of applicability and potentially useful improvement

HPCC and IT infrastructure

- Research group + IT
- Collaboration is the key
- Emphasize on IT infrastructure for research
- Can be effective even with minimum owned personal resources
- IT department at TTU supports the user groups

TTU High Performance Computing Center

HPCC Staff

- Dr. Philip Smith (Sr. Director)
- Dr.James Abbott(Assoc. Director)
- Jerry Perez
- Srirangam Addepalli
- Dr.Alan Sill (Sr. Scientist)
- Jodi McMurray
- Dr. Per Andersen
- Dr. Ravi Vadapalli
- Huijun Zhu
- Poonam Mane

Purpose

- Facilitate high performance research computing
- Assist with PC to Linux migration and grid-based computing
- Provide consulting and assistance with use of HPC resources

TTU High Energy Physics

Faculty: Dr.Sungwon Lee (Dept. Chair), Dr.Richard Wigmans, Dr.Igor Volobouev, Dr. Nural Akchurin , Dr.Alan Sill (Adjunct Prof.)

Postdoc's: Efe Yazgan, Jordan Damgov

Students in TTU Tier-3 operations : Youn Roh, Chiyoung Jeong, Keng Kovitanggoon, Terence Libeiro, Poonam Mane

Students in offline CMS CSP shift : Cemile Bardak

Personnel based at TTU, FNAL, and CERN. Physics analyses carried out using a mix of resources, primarily those provided at the LPC and via CRAB.

TTU T3 is available principally to run CRAB jobs

Resources and capabilities

- Linux clusters for parallel and serial computing
- Large-scale Lustre-based data storage
- High Speed Networking
- Oracle Grid Engine (Sun Grid Engine)
- Local campus grid
- Nationwide, regional and state-wide Globus grid access
- Operation for past decade
- 15% of the total grants to TTU are generated by research groups supported by HPCC

Infrastructure and Support

Infrastructure

- Experimental Sciences Building (Main Campus
- Reese Center
- 12 miles apart, currently connected with 1Gb network that is being upgraded to 10Gb

Support

- TTU IT department
- HPCC

Reese Campus

Antaeus (OSG and CMS Tier-3 resource)

- Production Cluster
- Mixed dual-dual core + dual-quad core Xeons, 240 cores total, 2 GB memory per core, Gb backplane. (Some owned by other researchers)
- 261.4 TB dedicated HEP (CMS) + 6TB user storage interfaced via Lustre

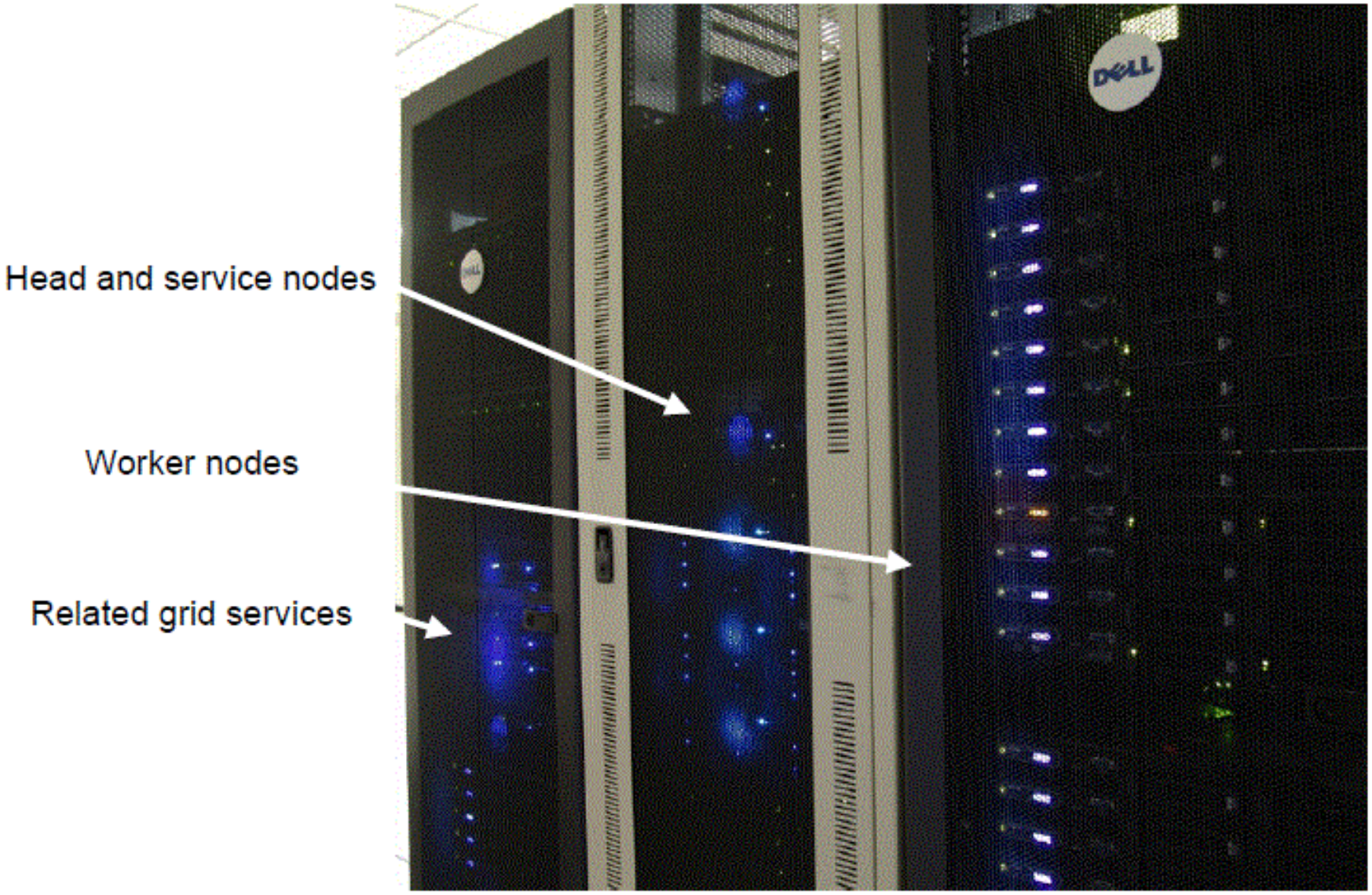
Special nodes

- **Hugin** – basic login node, Frontier Server
- **Munin** –basic login node, ITB storage element

Testwulf (ITB resource)

- Test Cluster
- 2 worker nodes

Antaeus Cluster



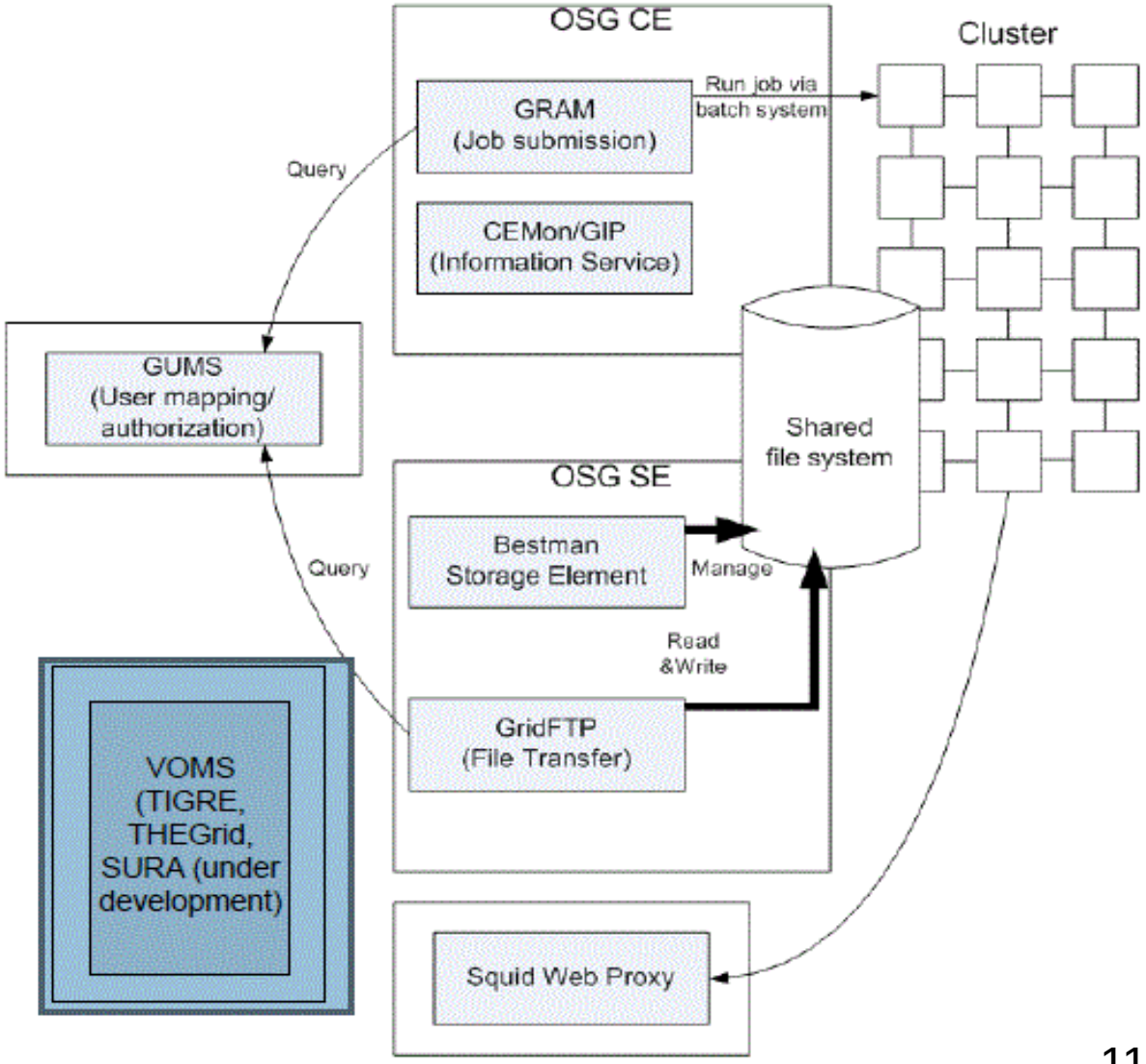
TTU Main Campus - HPCC Primary Linux Clusters

Grendel (primary parallel computing resource)

- Ranked 175 in the November 2009 Top 500 list, achieving 33.5 Teraflop/s max sustained and 40 Teraflop/s peak performance in LinPack test
- 420 dual slot quad core nodes with Intel(R) Xeon(R) CPU E5450 processors for a total of 3360 cores
- Each node has two Intel 5450 Quad Core 64 bit processors on a single board, as an SMP unit. Each node contains 16 GB of memory
- The core frequency is 3.0 GHz
- The core's are connected with DDR Infiniband, with a rating of 40.2 Tflop peak performance
- Interconnect: 4X DDR Non-Blocking InfiniBand
- A 80 TB Lustre file system runs over Infiniband

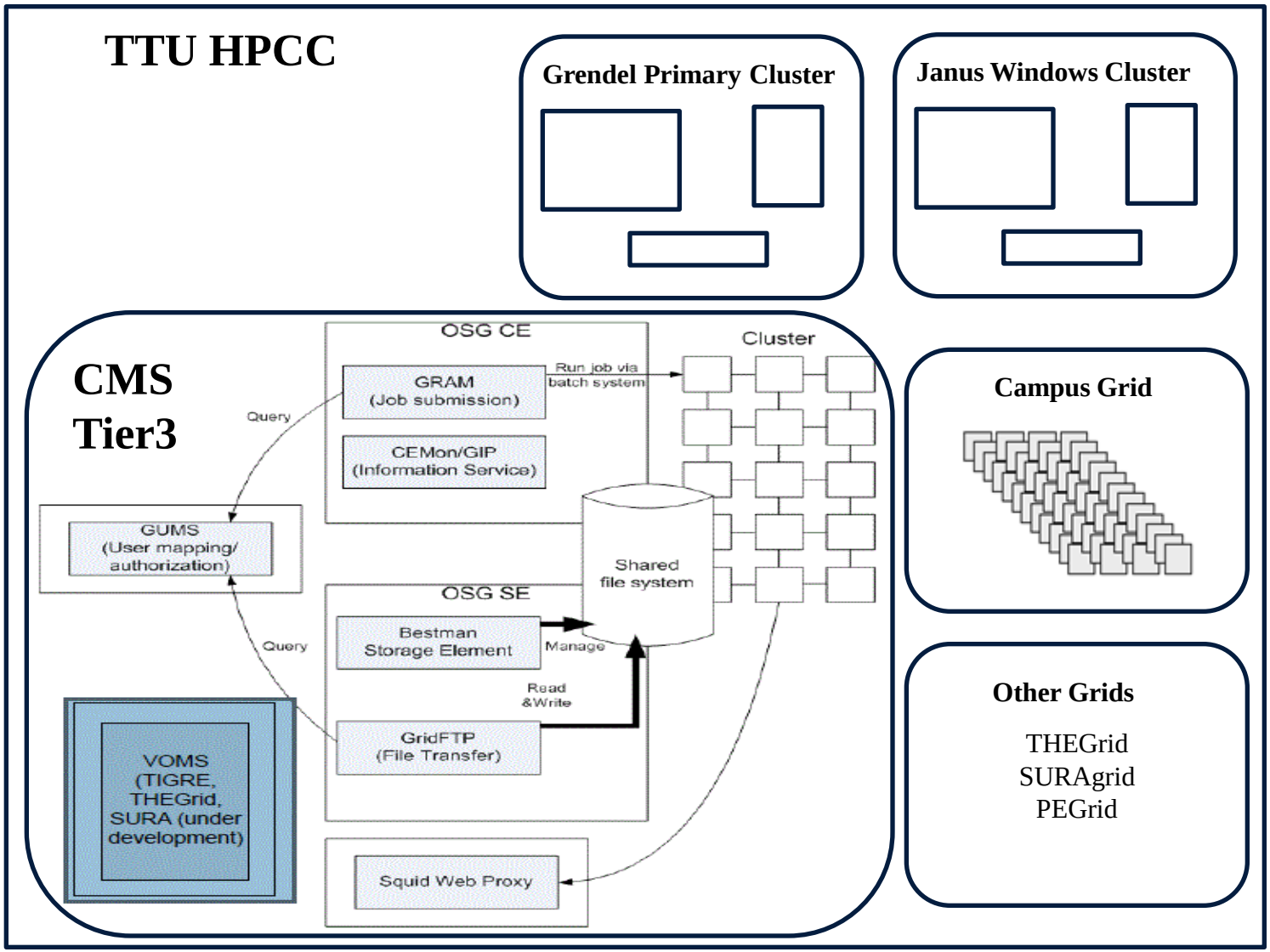
TTU Tier-3 site configuration

Basically the same as an OSG medium site, with additional services for other VOs



TTU Tier-3 site configuration

Our T3 exists and is supported in the context of other clusters in the TTU High Performance Computing Center



TTU Tier-3 Current Status and Capabilities

CPU power:

- Community Cluster – HEP, Chemistry group , HPCC
- Theoretically total 240 cores, two are down and out of warranty
- HEP – 144 cores,

  - 128 Xeon core available to cms queue + 16 core interactive login (HEP group)
- Chemistry – 64 cores, HPCC – 32 cores

Storage:

- 261.4 TB dedicated HEP disk space, 6 TB general use
- PhEDEx production T3_US_TTU instance
- SRMv2 interface (BeStMan-2.2.1.3.13) to cluster storage

Software & Batch system:

- Rocks 5.3 based on CentosOS 5.4
- Oracle Grid Engine (previously known as Sun Grid Engine)

TTU Tier-3 installation methods

CE and worker node base grid software:

- OSG-1.2.11 on CE, uses prima authentication
- CMS software (CMSSW) is automatically installed
- Monitoring: SAM, Gratia, RSV, MonaLisa and Ganglia,

GUMS:

- OSG supplemented by TIGRE VO

VOMS:

- Used to support TIGRE, THEGrid and SURA, PEGrid

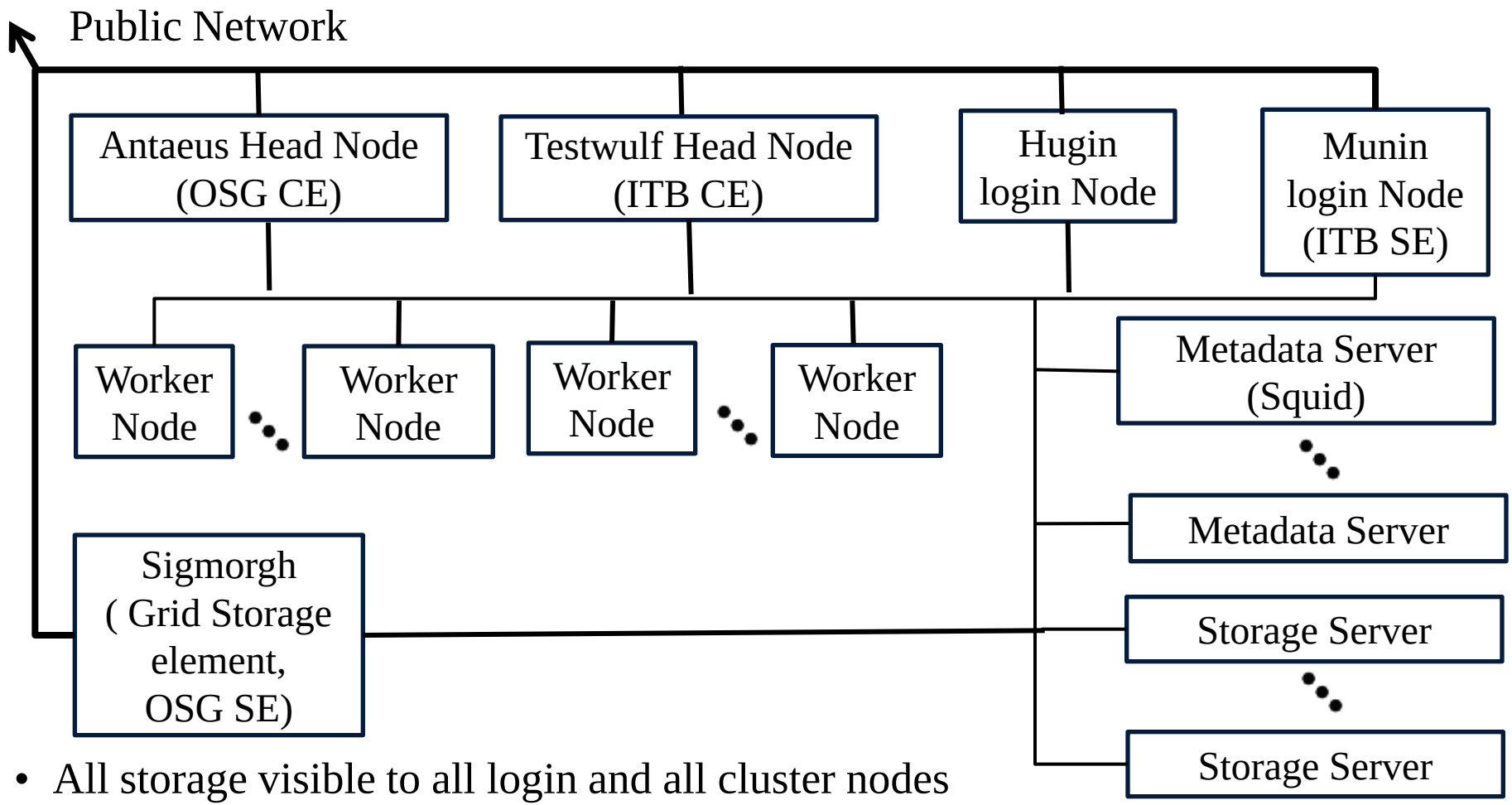
FroNtier and squid:

- CMS

Storage (BeStMan):

- BeStMan-2.2.1.3.13

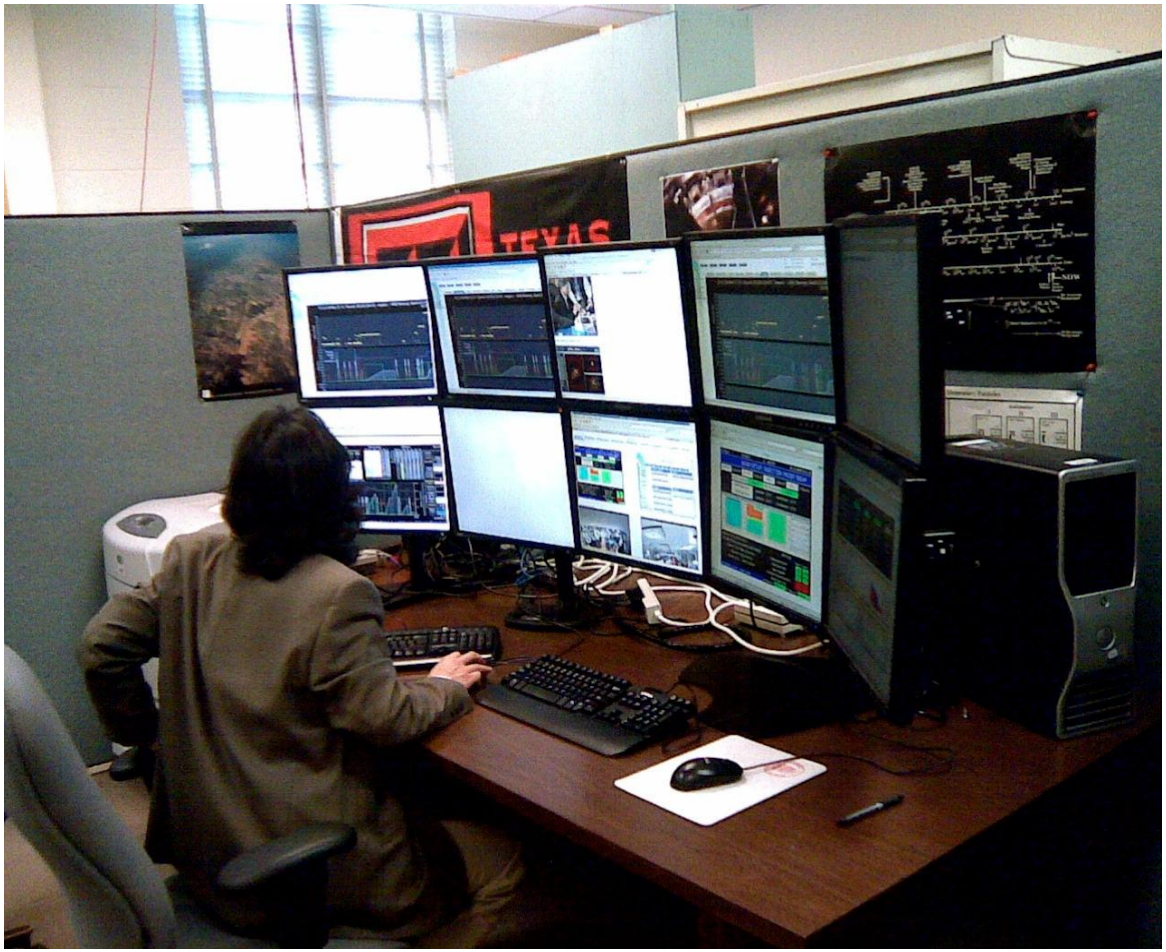
Worker node and storage layout



- All storage visible to all login and all cluster nodes
- Public network access through grid storage elements
- Logical partition of storage servers adjustable via Lustre

CMS center

- 12 monitor, two dual core system
- Periodic Shifts according to US times to support on global basis
- Official offline CSP CMS shifts
- Count towards requirement of CMS authorship



Way ahead

- Integrate - HEP office desktops with shared file system
- Use the Grendel's idle time to submit jobs as it is 20 times faster
- Virtualization – multiple copies of GUMS, Frontier db
- REDDnet (Research and Education Data Depot network) initiation
- Improve the CMS CSP facility

Current problems and general topics for discussion

- Queue advertisement

  - Control advertisement of queues to include only those available to grid users by VO, so that CMS jobs are not submitted to the wrong queues
  - Blacklist/whitelist parameters in the config.ini script helped

- Grid jobs in general & CMS jobs, do not clean up after execution. What are the best practices ?
- If all the jobs occupy all the queues, there no room for monitoring jobs. Hence, monitoring jobs should run at high priorities
- Pilot jobs have increased from UCSD

  - violation of grid certificate usage
  - certificate is assigned to each pilot job instead of assigning it to a person