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

HPCC Staff

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

- 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: Nural Akchurin (Dept. Chair), Richard Wigmans, Igor Volobouev, Sungwon Lee, 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

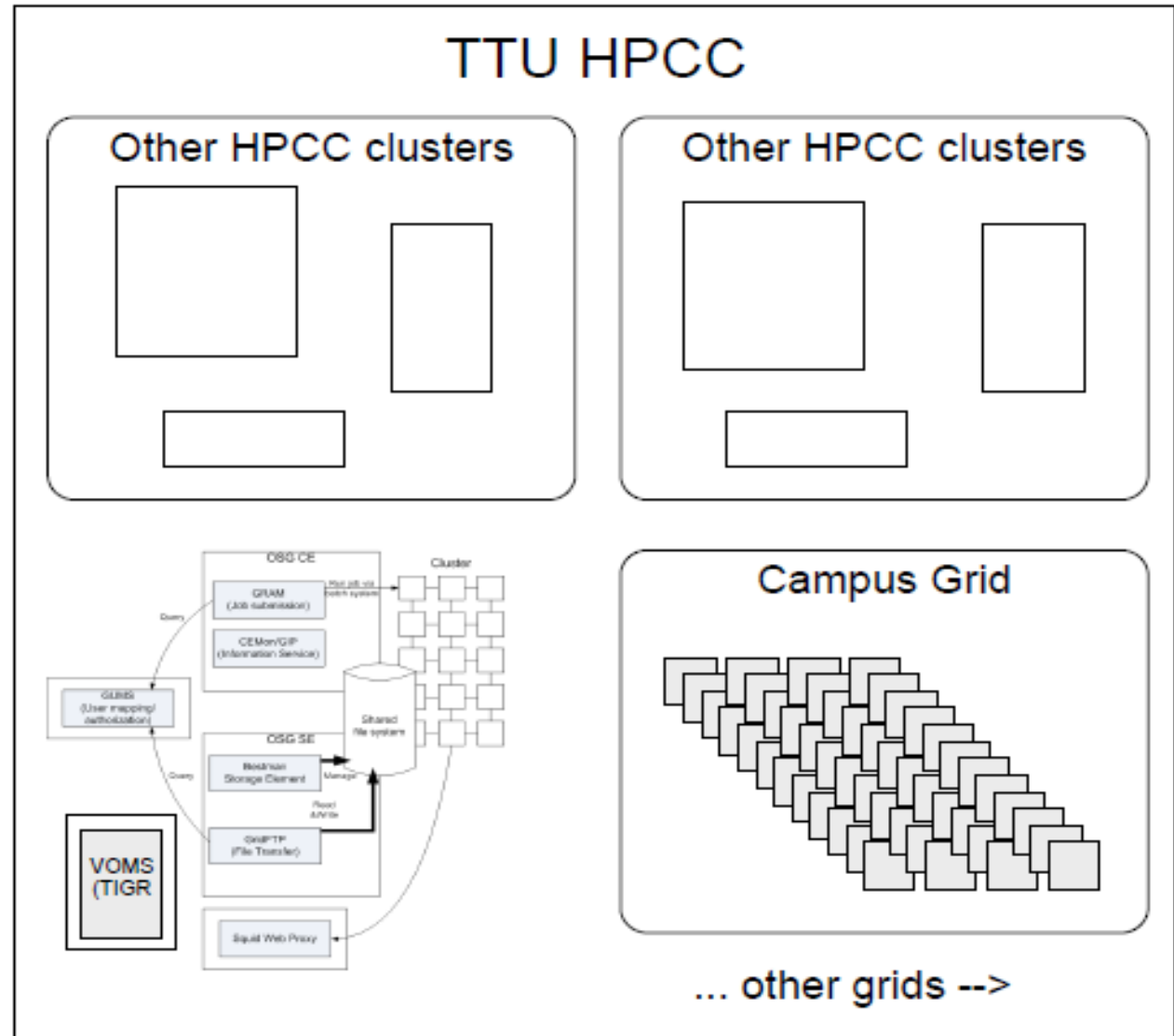- 15% of the total grants to TTU are generated by research groups supported by HPCC

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

7
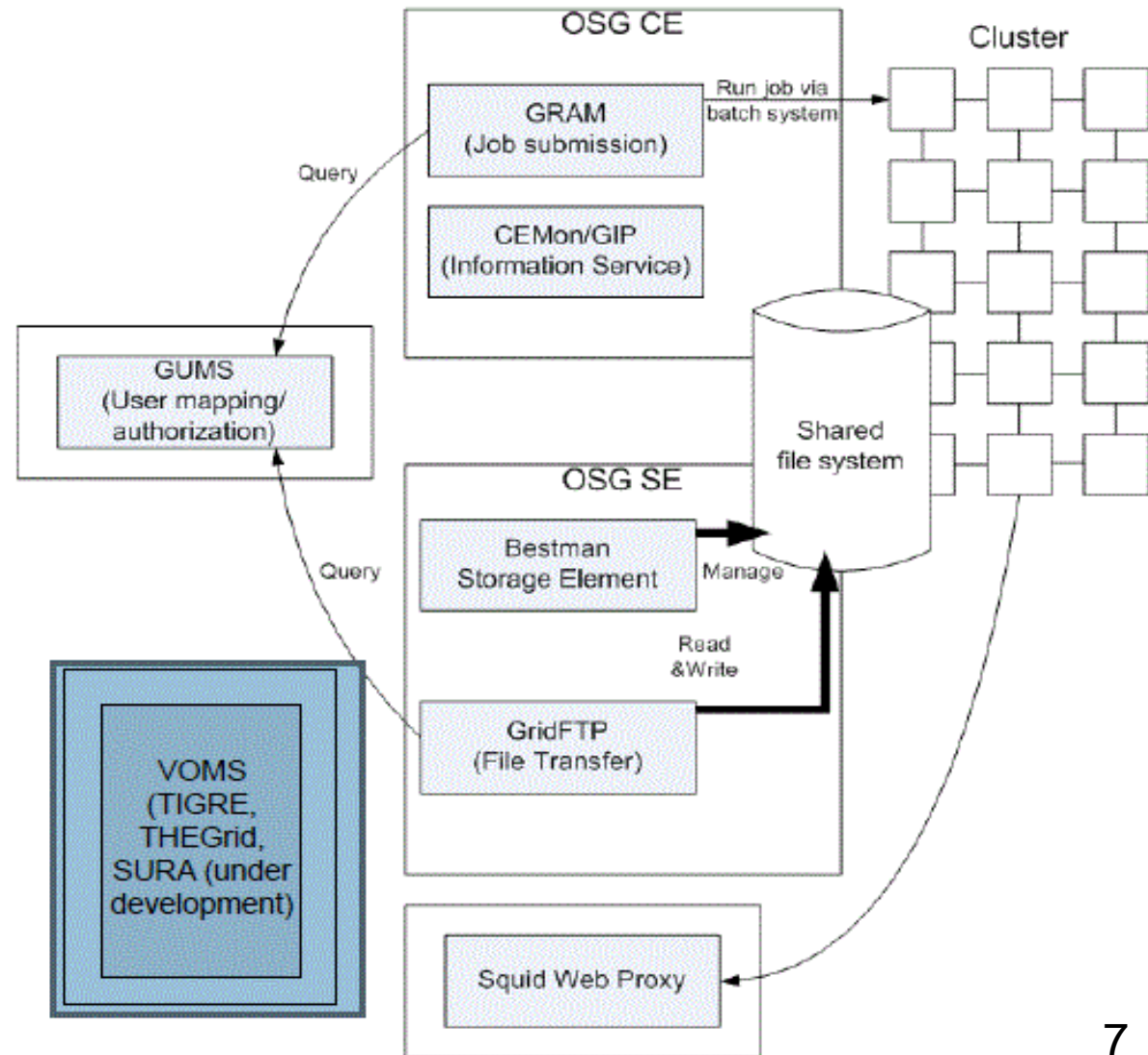
Our T3 exists and is supported in the context of other clusters in the TTU High Performance Computing Center

TTU Tier-3 Current Status and Capabilities

CPU power:

- 128 Xeon core available to cms queue + 16 core interactive login (HEP group)

- Idle time access to remaining additional 96 cores

Storage:

- 261.4 TB dedicated HEP disk space, 6 TB general use

- PhEDEx production T3_US_TTU instance

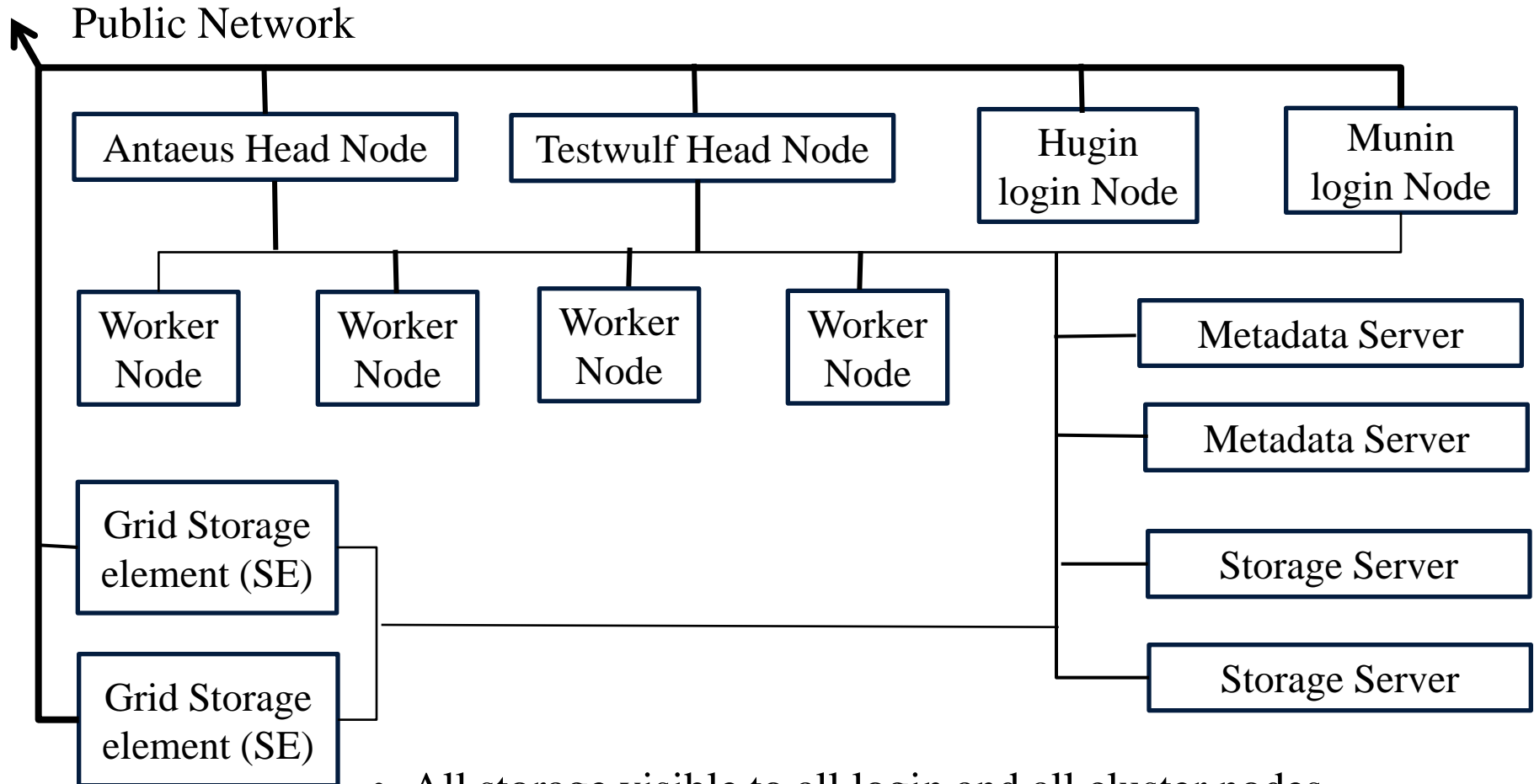- SRMv2 interface (BeStMan) to cluster storage

Batch system:

- SGE

Other resources:

- Idle-time access to campus cycle-scavenging grid and other clusters

- ITB resource (TTU-Testwulf), GUMS, and TIGRE VOMS server

Worker node and storage layout

Public Network

Antaeus Head Node	Testwulf Head Node	Hugin login Node	Munin login Node

Worker Node	Worker Node	Worker Node	Worker Node

Metadata Server

Metadata Server

Grid Storage element (SE)

Storage Server

Grid Storage element (SE)

Storage Server

- 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

Grendel (primary parallel computing resource)

- Ranked 175 in the November 2009 Top 500 list

- 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

Reese Campus

**Antaeus (OSG and CMS Tier-3 resource)**

- 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

Hugin – special node , basic login node, Frontier Server

Munin – special node, basic login node, ITB storage element

Network between Main Campus and Reese is upgraded to10Gb

Head and service nodes

Worker nodes

Related grid services
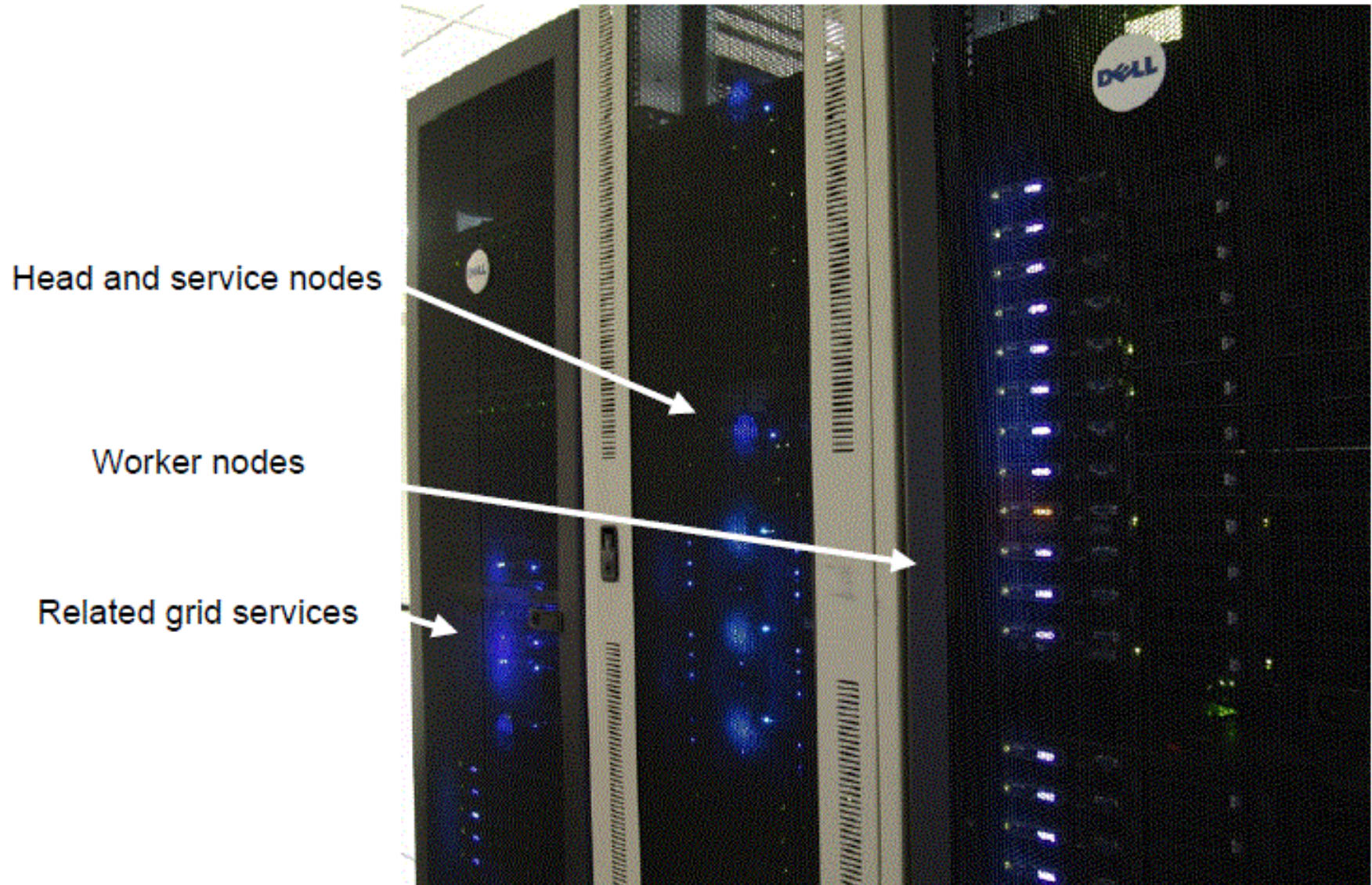
CMS center

- 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 CPS 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