

Minutes of the HiPCAT Meeting
06 June 2006
Author: A. Kerstens

HiPCAT Attendees:

- Borries Demeler (UTHSCSA)
 - Erik Engquist (UH)
 - Nick Grishin (UTSW)
 - Mark Huang (UH)
 - Steve Johnson (TAMU)
 - Lee Panetta (TAMU)
 - Andre Kerstens (UTEP)
 - Chuck Koelbel (Rice)
 - Josten Ma (UH)
 - Jerry Perez (TTU)
 - Karl Schultz (TACC)
 - Alan Sill (TTU)
 - Phil Smith (TTU)
 - Warren Smith (TACC)
 - Michela Taufer (UTEP)
 - Kiran Thyagaraja (Rice)
 - Jae Yu (UTA)
 - Sergio Zapata (UTEP)
 - Ravi Vaddipali (TTU)
-
- Phil opens the meeting at 1 pm MST.
 - Election for a new director and associate director:
 - People don't want a new committee to elect the new director/associate director of HiPCAT.
 - Because Michela has to leave at 1:45 MST Phil asks her to give an update on the status of BOINC.
 - Michela introduces BOINC:
 - David Anderson held a talk on BOINC at UH a couple of months ago and had a discussion with some faculty there on how UH cpu cycles can be donated to BOINC projects. Andre asks Josten whether he is aware that David Anderson was at UH. He says he and other people in his group are not.
 - BOINC is an open-source infrastructure for task distribution on a volunteer grid.
 - Andre and Michela are talking with Alan Sill how we can integrate Boinc in TIGRE. We also had a telephone conf. with David Anderson already.
 - Alan says he is happy with the BOINC package and will experiment further.
 - Michela asks everyone to download the docking@home client as soon as our project is online in the fall and help solve problems we experience on client's computers.
 - Warren gives an update on the tigre project:

- The TIGRE portal has gone on line at <http://tigre.hipcat.net>
- TIGRE members are working with Borries Demeler to get his application working on the TIGRE infrastructure.
- The team is working on packaging applications for TIGRE and is currently looking at the packman tool.
- Alan adds that they will start working on getting the listed resources hooked up to the system. They also will start working on improvements of the TIGRE web presence.
- Warren mentions that TIGRE is looking for more applications.
- Josten mentions that a new person has been hired for the tigre project (Mark). Mark started working for the project on the first of June.
- UTEP update:
 - Andre mentions that the RADC has now been officially opened by the president of UTEP and that Sergio and himself are installing schedulers on Sacagawea and Star.
 - He mentions that Michela has been given the Outstanding Young Researcher Award from the college of engineering (lots of applause :-) and that she and another faculty in bioinformatics received a \$100,000 ARP grant that will be used to built up a campus wide grid for computation of rna structures. Andre mentions that since this seems very similar to TIGRE, there might be a possibility of collaborating in the future.
 - Phil asks whether all the systems are now up and running with the new A/C. Sergio answers that all the systems are up and the room has now approx. 22 ktons of cooling power.
- TACC update:
 - Lonestar has been upgraded to 650 nodes (3 GB per node).
 - Phil asks whether 3 GB per node is sufficient. Karl answers that you never have enough memory of course, but that they will actually start upgrading the memory soon as well.
 - Somebody asks what kind of Infiniband hardware they have: Karl says it is from Cisco.
- UTSW update:
 - Nick Grishin visited TTU a while back did a presentation for the Bioinformatics group. He mentions that he had a great day and many interesting discussions.
 - His small cluster is being upgraded with new nodes every 3 or 4 months or so. By now the system has become very heterogeneous (oldest machine is 3 years).
 - UTSW's administrative systems are going to be upgraded soon.
- UH update:
 - Mark says that there is nothing to report for UH.
- TTU update:
 - TTU is putting in a new test and development system that people can use for compiling and testing before moving to the big cluster.
 - They are currently evaluating the Moab cluster suite from ClusterResources.
- Rice update:

- The Cray XD1 is in full production.
- The Mathworks distr. computing toolbox has been installed on the XD1.
- They are organizing the TeacherTech conference for the summer. Phil asks what level the teachers are: Chuck answers that they are high school and college level.
- A system has been implemented that provides 4 hour time slices on the Cray system. Phil asks whether users are happy with this. Kiran answers that this is hard to say, but generally there are mixed reactions. 4 hours might be too short and they are thinking of increasing the slice to 8 hours. Users can generally schedule as many processors as they want in their assigned time slice.
- UTHSCSA update:
 - The renovation of the biomolecular structure lab has been completed.
 - The Opteron cluster has been reconfigured to use Rocks 4.1 instead of SuSE Linux, but there are some problems during periods of high load on the filesystem.
 - The TIGRE software stack has been installed on this cluster so that it becomes a resource.
 - Set up bioinformatics of excellence center in SA. Positions available there.
 - They released a couple of new versions of the Ultrascan software.
 - About 10 students will come over to UTHSCSA in the summer for the NIH B-SURE program.
 - Question: What happens exactly on the cluster under load? Borries answers that an rsync copy of big files on the same filesystem corrupts files under high load. Perhaps the SATA drives might be the problem: one parity drive in the Raid 5 might not be enough (maybe get 2?).
 - Borries heard that switching off noatime might help? This usually is good for performance, but probably won't help for the file corruption.
- UT Arlington update:
 - The Tier 2 cluster is nearly online, I hope. The Ibrx file system has been installed and configured; we have cleared up the electrical infrastructure issues, and now have a functioning PBS batch system. We are in the process of configuring the OSG software stack so that we have a functioning gatekeeper/gridftp host.
 - There are some additional steps before we are running ATLAS production. This includes the setup and configuration of at least an additional dedicated gridftp server and the installation/configuration of a DQ2 server.
 - The Tier 2 cluster is comprised of 160 compute nodes, 8 front-end nodes, and 6 nodes supporting our distributed file system. The compute nodes, Dell SC1425's, are dual 3.2 GHZ Xeon EM64T, 4GB RAM, and 160GB SATA storage machines. The front-end machines are Dell 2850's with dual 3.2GHz Xeon EM64T, 8GB RAM, and 2x 73GB SCSI drives in RAID1 configuration. The nodes supporting the distributed files system are similar to the compute nodes with the addition of a Fiber Channel HBA.
 - We have 16TB of usable capacity in a local SAN, provided by DataDirect Networks. The system uses 82 250GB SATA drives connected to a redundant storage controller/FC interface. The storage is made available to the compute

nodes through the use of the Ibrix Fusion file system, which provides a global distributed file system.

- We are using Platform Rocks (v3.3.0-1.2) as the cluster management system which is RHEL 3 based. This provides us with the ability to manage installation and monitoring of the compute nodes.
- Geology and Environmental Science department researchers are working on implementing their simulation software on our DPCC cluster. Hopefully this completes will begin during the summer.
- TAMU update:
 - Steve says that there is nothing to report.
- Phil asks whether there is any new business. Nobody answers so there must be none!
- Phil adjourns the meeting at 1:45 MST.