Activities:

In August the group had a phone meeting with Charles Catlett (Executive Director of the TeraGrid project, ANL) and Dane Skow (TeraGrid Deputy Director, ANL). The phone meeting was an opportunity for the WG members to ask questions on TeraGrid to our two guests.

During the meeting several members pointed out their need for tools that allow interactive data analysis and visualization. From our conversation with Catlett and Skow as well as from the experience of some of our members who already use TeraGrid for their simulations, we have understood that currently TeraGrid does not support such interactive tools.

We would like to better understand the tools that are currently included in TeraGrid for visualization and data analysis. Therefore in our next phone meeting we plan to invite Kelly Gaither (Associate Director, TACC) to discuss with us on tools for grid-enabled terascale visualization. The tentative data of our next meeting is end of September - beginning of October.

The minutes of the phone meeting in August is reported below. A copy of the minutes was forwarded to: Grace Peng (NIH/NIBIB), Miriam Heller (NSF/OCI), Charles Catlett, and Kelly Gaither.

---

Multi-scale Modeling Consortium - Working Group 5
Minutes of the phone meeting on 08/15/06

Author of the minutes: Michela Taufer
Starting time: 2 PM (MDT)
Meeting duration: 1 hour
Participants of the meeting:
Michela Taufer (WG 5 member, UTEP)
Patricia J. Teller (WG 5 member, UTEP)
Elizabeth Gardner (TeraGrid user, UTEP)
Andre Kerstens (WG 5 member, UTEP)
Charles Catlett (Executive Director of the TeraGrid project, ANL)
Dane Skow (TeraGrid Deputy Director, ANL)
Ching-Long Lin (WG 5 member and TeraGrid user, UIOWA)
Wilfred Li (WG 5 member, SDSC)
Dan Haworth (WG 5 member, PSU)
George Karniadakis (WG 5 member and TeraGrid user, BROWN)
Jim Brasseur (WG 5 member, PSU)
Bill Hazleton (WG 5 member, FHCRC)
George Kovacic (WG 5 member, RPI)

Goals of the meeting: Discussion on TeraGrid, its potentials, and its challenges.

Report of the phone meeting:

All: Short presentation of the participants and their projects in terms of their science goals and the computer infrastructures used to pursue these goals.
Catlett: Introduction of the TeraGrid project using pdf slides available at: http://www.teragrid.org/about/presentations.html. TeraGrid integrates hardware and software resources across the USA. Most machines are clusters.

NSF is supporting extending the computing infrastructure of the TeraGrid. There will be a solicitation for either one big $30M machine or two smaller machines $15M each before the end of this year. This will significantly extend the size of TeraGrid. Between 2005-2006, NSF will award grants to maintain / evolve the TeraGrid in terms of user support, system support, coordination of software and other infrastructures across the TeraGrid sites.

It is possible to apply for access to these resources through an on-line proposal application. Development awards might range from 10,000 to 30,000 CPU hours. Proposals can target different uses of the TeraGrid resources:
1. To allocation computation on a specific TeraGrid machine
2. To allocation computation on several machines across the TeraGrid

Three possible ways to use the TeraGrid within a development award:
1. Search the grid for one specific machine that allows you to scale up your computations on top of your own infrastructures (e.g., work of Lin at UIOWA).
2. More aggressively extend applications to work across several TeraGrid sites (e.g., work of Karniadakis at Brown).
3. Build workflows of computations and data handling across different storage and computation sites.

What kind of help from the TeraGrid project can scientists receive?
1. On-line material, e.g., tutorials
2. Several web portals are work in-progress. They aim to allow an easy access to the TeraGrid infrastructures
3. User-support group that can be contacted through the email address help@teragrid.org

A first recommended step is to send an e-mail to this group to present your application and ask for advise; a second step would be to submit a short development proposal to request CPU hours on the TeraGrid infrastructures.

If you are planning to apply for grants within NSF or NIH, no financial support is required for the use of TeraGrid in that grant’s budget.

Brasseur: Post-processing of data after simulations and their visualization are important components playing a significant role in the analysis of simulations results. These operations need to be easy and fast. Are there resources on TeraGrid that facilitate this post-processing?

Catlett: The data within TeraGrid are moved using dedicated networks. There is a distributed file system for data analysis at SDSC that is mounted at ANL and NCSA. However the issue of how to facilitate scientist in accessing their data interactively is still an open question.

Karniadakis: No interactive system for analysis and visualization is available now except for some smaller facilities at TACC and SDSC. You are required to download your data from the TeraGrid storage to your local storage and locally work on it.

Catlett: Agree on this current weakness of the TeraGrid. He will be discussing this issue with NSF officials in the future; it would be good if all of the IMAG members will take this up with their respective officers as well.

Brasseur: Most of the time local storage and networks are not suitable for the post-processing of large amount of data that have to be moved from the TeraGrid storage to these local storage resources.

Teller: Is there any intention from NSF to support research to address this issue?
Catlett: Most of the funds are currently dedicated to increase the size of the TeraGrid project resource-wise. It is a good idea to schedule a second meeting with the TeraGrid experts who are dealing with post-processing and visualization.

All: This is a very interesting possibility that we want to consider for a next meeting of the working group.

At this point the meeting was adjourned. The next meeting will be scheduled at the beginning of October.

Conclusions:

The TeraGrid project seems very tempting for several projects in the working group because of the potential of access to huge amounts of computing power. Support for the migration of applications and the application process to apply for use of the TeraGrid resources are in place.

Participants agreed on the fact that remote post-processing and visualization are extremely important in science but at this point the TeraGrid resources that deal with these issues are not suitable for the applications that the working group are working on. The participants urge NSF and other agencies such as NIH (several projects that are using the TeraGrid are NIH-funded) to address this need through grant solicitations and research support in the next future.

We will continue the conversation about post-processing and visualization with the TeraGrid team in the next working group meeting. In particular we will invite experts from the TeraGrid visualization group to discuss with them the needs that we have when dealing with our applications and what kind of solutions are now available as well as what kind of innovative tools and infrastructures could improve our work if available in the next future.