Cyberinfrastructure User Advisory Committee (CUAC)

Initial Discussions and Recommendations
September 2006

Status: Draft Informational

This draft is being circulated among CUAC facilitators prior to circulating to TeraGrid PI's and CUAC members for comment.

Abstract

The first CUAC meeting was held during TeraGrid'06 in June 2006. At this organizational meeting the CUAC was grouped into three subgroups, each of which was assigned a TeraGrid PI as facilitator. The subcommittees were each asked to review a subset of the recommendations from the January 2006 TeraGrid Site Review as an exercise to both acclimate the CUAC members with the project and provide initial feedback.

The three groups were asked to meet via teleconference during July/August 2006 and to report out by teleconference during the quarterly TeraGrid management meeting in September 2006. Below are notes from the September report-out discussion, followed by notes from each of the group discussions.

Contents

cec@uchicago.edu

Summary Discussion

During the quarterly TeraGrid management meeting (September 13, 2006) a teleconference was held with nearly all CUAC members on the telephone and nearly all of the TeraGrid PI's present at the meeting. The following are notes from this summary.

Group A Summary Report and Discussion

Group A commented on the TeraGrid's response to January review team's recommendation that TeraGrid establish a "standard" workflow system for all users. They felt that although a standard may be premature, the issue of workflow was important and TeraGrid was in a position to affect what standards emerged. Dennis Gannon (IU) explained that there were several dozen popular tools used in the scientific community, each suited for a particular type of workflow and/or scientific discipline. Thus for TeraGrid to "force" a standard would in fact run counter to the requirements of the user community. However, TeraGrid is focused on providing the basic services required by these various workflow systems, and does provide a default workflow system for users who do not have a preferred system.

Group A also commented as follows:

- A portability document to help guide moving code from platform X to platform Y would be useful to many users.
- Data visualization is a major bottleneck for many users, and collaborative visualization tools for remote users would be particularly useful. Visualization workshops would also be a welcome and necessary service for users.
- It would be good for TeraGrid to keep data and report on the user of data collections and the movement of data between resources/sites.
- Given that TeraGrid is only funded through 2010, it is not clear what the user strategy should be for long-term data archiving.

The group also discussed the issue, raised by the January review team, of balance between TeraGrid's "Deep" and "Wide" initiatives. Some of the group felt that the "Deep" services (helping users with large-scale codes, for instance) were "under control" and that the challenge for TeraGrid today is to engage new communities. But it was also noted that with new petascale machines coming into TeraGrid there would need also to be an emphasis on helping users with their "deep" needs on these new systems.

Group B Summary Report and Discussion

Group B discussed a high-level strategy for the CUAC in the context that most members, as users, are not technical experts but rather users of TeraGrid services. As a result, the

group elected to focus on the user experience rather than the particular technical recommendations contained in the January review panel report. In particular, the group focused on the experience of new TeraGrid users in order to identify barriers and provide TeraGrid management with advice on how these might be reduced from the user point of view.

The group identified several barriers for new users:

- Allocation requests are fairly speculative in that it is not well-known what the probability is that a new PI will receive an allocation, and therefore it is unclear what the likely return on time investment will be in preparing the request.
- New users need to understand whether, and how, use of TeraGrid might accelerate their particular work.

The group also noted that the Science Gateways concept is excellent from the standpoint of addressing ease of use, and recommended that TeraGrid identify best practices. Nancy Wilkins-Diehr (SDSC, Area Director for Science Gateways) updated the team on the status of the Science Gateway Primer, which is intended to provide exactly this type of guidance.

The group suggested outreach to non-NSF funded researchers (e.g. NIH). Charlie Catlett noted that this is happening, and that NIH-funded researchers currently make up over 10% of TeraGrid usage.

Group B also recommended a "tiered" system for allocations, targeting the experienced users toward the largest capability systems and using the "smaller" systems for new users.

Group C Summary Report and Discussion

Group C was provided with an update regarding TeraGrid activities, in the eight months since the January review, in areas related to outreach, education, training, and communications. The group noted the following:

- The science highlights booklet being prepared for SC06 is an excellent idea.
- The systematic program undertaken by TeraGrid PI's to discuss TeraGrid with each of the NSF directorates and divisions has been very good. The group encouraged continued efforts to engage individual program officers at NSF.
- In terms of engaging the community, the group encouraged TeraGrid to explore the use of Wikipedia technology for documentation and training, as a way to harness the creativity of the entire community.
- The group noted that training and updates for users is a critical activity, and more services and events are needed in these areas. Online (self-paced) tutorial materials in particular were suggested as an important service for users.
- The work being undertaken in streamlining authorization (recent workshop on attributes-based authorization) was highly encouraged as a concrete way to reduce barriers for users.

Background: Teleconference Discussion Notes

This section contains conference call notes provided by leaders from each of the three subgroups.

Group A Teleconference Notes (Provided by Ralph Roskies)

Collaborative software (R9-a) there was a sense that users would know about CVS servers, Bugzilla servers etc. and there was little need for TeraGrid to provide these services globally. Users would know how to invoke them themselves. Since there was so much to do in TeraGrid and limited resources, people felt that concentrating on this stuff was not the best use of their time.

Workflow Tools (R9-b). Workflow tools were perceived as potentially very useful. Maybe standardization was not so necessary (that was our response) but they felt that TeraGrid was uniquely suited to push for such standards. They seemed to suggest reconsidering our response.

End-to-End (network) performance tools (R12) end-to-end. Provided background that the networking community had tried to automate network tools, and that for many cases, a network expert was still required. The group also asked how often these kinds of problems arose, and the answer is 'not very often'. Given this, the group was less inclined to argue for a systematic effort to develop new tools to simplify this.

One suggestion that emerged was whether we could provide a document that outlined the portability issues from one platform to another. Suppose you have a code running on machine X and you want to move it to machine Y. What are the problems that have been encountered and how were they overcome?

They felt our answers to (R23) (standards) were fine.

Data (R16). One of the major points was that data visualization is still a major bottleneck for large data sets. This is tied to the questions of appropriate data analysis algorithms, but it was also felt that collaborative tools could be very useful in allowing people to view from their home institutions renderings on things like power walls located at the major centers. They stressed the use of visualization for discovery.

They wondered whether TeraGrid has workshops on visualization techniques.

The group felt that TG had done a good job in enabling rapid data motion between sites.

Publication and sharing of data was deemed very important. One example close to the interests of a CUAC member was LIGO data. Many groups will want to examine this data, do analysis on it etc.

They felt TG should report on the use of data collections, and on how much data was being moved from site to site.

One interesting point brought up was the persistence of data. They see TeraGrid as living from year to year. If TeraGrid is to serve as a data repository, what happens to data if TeraGrid funding gets cutoff.

In sum, they felt the committee recommendation addresses a serious need, and that the TG response was tentative.

Balance of deep and wide

There was spirited verbal discussion and even ensuing e-mail discussion. One participant felt that the centers had been doing 'deep' support for years, and that, while essential, this was the easy part and did not need an expanded effort. The real challenge which held the possibility for transformational change was engaging new communities. As he put it, he wanted to push "the centers to think broadly about how this technology can not only help me in my basic science but more broadly enable communities". That same person felt that the largest impact comes not from 2-10 million node hour allocations to a single PI but from the aggregate of small to medium size allocations. Another felt that the issues in deep support, especially with new architectures like Blue Gene and the petascale initiative, were very challenging and required concerted effort. Finally everyone seemed to agree that both 'deep' and wide' were important, but they disagreed on where more emphasis should be placed.

They saw little point in distributing codes across sites in real-time, unless there are unique required features at different sites. They viewed wide' as meaning reaching out to new communities (as opposed to using platforms at several sites), and felt this was very important. They also felt that the user portal might make it easier to move applications to particular sites and thus make it easier to pick the site with the more appropriate architecture rather than picking a site simply because users were familiar with it.

Group B Teleconference Notes (Provided by Phil Maechling)

This was the CUAC Group B first conference call. This call was a follow up to the initial formation meeting at the TeraGrid 06 Conference.

The call lasted one hour. One of the Group B members was traveling and did not participate although we had his approval to start our discussions without him. The discussions during the call focused on two primary issues: (1) CUAC's responsibilities and (2) general advice for TeraGrid Management.

During the call, some general recommendations emerged for the TeraGrid management. These recommendations are described below. Some of our recommendations relate to operation of the CUAC. Some of our recommendations are based on our user's perspective of the TeraGrid. We did not address the CUAC Group B panel recommendations during our call, and do not, at this time, have recommendations on those issues.

At this time, no follow-up call is scheduled. We will await either comments to our current recommendations, or further direction from TeraGrid management before re- convening the group. The scribe for this call was Philip Maechling and these notes have been reviewed by all the advisors that participated on the call.

What are the CUAC's responsibilities?

Within our group, there are significant uncertainties about the role of the CUAC with respect to OCI and TeraGrid management. During our discussion, we recognized the fact that TeraGrid is a panel reviewed project, and it gets advice from that panel. For several reasons, our CUAC group does not want to duplicate the function of the TeraGrid review panel. We recommend that TeraGrid management identify responsibilities for the CUAC that are distinct from the responsibilities of the review panel.

Based on earlier discussions, our facilitator outlined the thinking of TeraGrid management on this issue. The TeraGrid management has suggested a sequence of CUAC activities in which the CUAC reviews the panel recommendations, and then at some point in the future, the CUAC helps to prioritize TeraGrid responses to panel recommendations. Our CUAC group did not reach a consensus on the reasonableness of this plan. However, there was concern by CUAC members on the amount of time and effort needed to evaluate the panel recommendations. The background and interests and time commitments of our CUAC group make it unlikely that all of us will be able to provide useful technical reviews of the TeraGrid issues. We recommend that TeraGrid management identify responsibilities for the CUAC that are appropriate for a group that includes both scientific users and computer scientists.

General Advice for TeraGrid Management

We began our group discussion at the "big picture" level. We began by assuming our role is to provide advice to TeraGrid management from a user's perspectives. We agreed to try to provide constructive advice. Each advisor then discussed some aspect of their use, or intended use, of the TeraGrid.

Several of the issues we discussed were related to TeraGrid allocations. Current access to the TeraGrid requires (at least perceived) extensive paperwork (development of an allocation request) that has no guarantee of resulting in access. New users are not likely to spend much time writing speculative allocation requests because they don't know whether the grant will be approved. In addition, even if the grant is approved, the TeraGrid may not

be useful to their research. There seems to be at least two issues here: One, how likely is an allocation request to be granted? Two, granted an allocation, will the TeraGrid resources help scientists perform their research?

To address the first issue, we recommend that TeraGrid management provides better information on how allocations requests are evaluated. The TeraGrid should also provide easy access to success rates for allocations requests.

To address the second issue, we recommend some mechanism that allows researchers to evaluate the usefulness of TeraGrid resources for their research. This may be some type of limited access mechanism for "testing" out the TeraGrid so they can decide whether writing an allocation request is worthwhile. Our facilitator pointed out that some of this is captured in the current allocation levels of development, MRAC, and LRAC. If development allocations are the existing TeraGrid mechanism that lets potential new users "test things out", then this second issue may be addressed by better explanations regarding the purpose and difficulty of getting development allocations.

One further allocation discussion related to matching users (or jobs) to TeraGrid resources. It seems that all classes of users, and all types of jobs, are combined together on the TeraGrid. However, the computing services needed by users and jobs vary widely. For example, small jobs submitted through science gateways may be competing with very large jobs submitted by terascale computing groups. At allocation request time, the TeraGrid might categorize users or jobs by "quality of services" needed by the request, such as number of nodes used per run, length of run-time, data storage, queue wait time. Users with more demanding requirements may need to meet more demanding allocation requirements. Then maybe some TeraGrid systems serve only the most demanding simulations, while other systems service new users and run jobs that can tolerate longer queue wait time. On our short phone call, we did not have time to consider the implications of this type of quality of service allocation system. However, we recommend that TeraGrid management consider the implications of assigning allocations to specific machines based on usage characteristics.

There are research communities that are not NSF-funded (e.g. NIH, DOE, university) that could benefit from TeraGrid resources. However, these communities have little knowledge of the TeraGrid resources. To make the TeraGrid more accessible to these non-NSF research groups, we recommend that the TeraGrid outreach efforts include non- NSF funded research communities.

While the TeraGrid provides a standard collection of software, some research groups use, or have developed, software tools not in the TeraGrid software collection. Some users don't view the TeraGrid as useful resource because it is not clear whether their tools will run on the TeraGrid. We recommend that the TeraGrid better identify what tools (e.g. Matlab) can be used on the TeraGrid and make it easier for users to determine whether their software will work on TeraGrid resources.

There was significant interest and support from our group for the concept of Science Gateways. One key perceived benefit to science gateways is that gateways will help scientific users do research and without having to learn high performance computing. If science gateways can make this possible, they are a powerful model. There was some discussion whether science gateways actually are providing science research results or whether they are large software developments without users. We recommend TeraGrid management provides the user community with some "best results and best practices" information on specific science gateways that meet the challenge of wide community usage together with science impact.

Group C Teleconference Notes (Provided by Scott Lathrop)

Outreach

- We used the TeraGrid '06 conference to share the TeraGrid vision and strategy with the community. Engaged over 200 people in tutorial sessions, and 450 people in the conference overall.
- Tom Finholt and Ann Zimmerman conducted a User Workshop to capture community feedback – they will report on this at the quarterly meeting Sept 13.
- We are preparing an updated TeraGrid brochure and a TeraGrid Science Highlights document in time for SC06
- Over 30% of the requests for over 1M service units on TeraGrid are from PIs supported by NIH
- We have been meeting with Program Officers across the various NSF directorates through which we are getting invitations to speak at community workshops; we have setup a speaker's bureau
- Charlie has started a TeraGrid blog (http://teragrid.typepad.com) as a method for keeping both TeraGrid staff and users updated on TeraGrid priorities and items of interest.

Outreach Feedback

- Highlighting the science impact is important
- Working with NSF directorates is important to re-enforce Dan Atkin's cross-cutting CI efforts within NSF
- A primer for new users is very important, with a clear vision and strategy for TeraGrid
- Lots of PIs still aren't aware of TeraGrid and what is has to offer
- Wikipedia is a good approach to get materials on-line and engage the community
    - You can build the wiki during the course
    - You can use wikipedia to improve documentation
    - Result is much better materials in the end
    - People contribute much and the spam issues are small
    - Use wikipedia for documentation, on-line primer, and watch the use grow

  - Quality has proven to be good in study by Nature
  - Fast tracking approach
  - Can setup your own wiki and cross-reference to the common wikipedia
  - Can create layers of documentation
    - Well supported and vetted materials
    - Documents in development through community contributions
  - Users can help determine and push where more work is needed, and reduce some of the load on the documentation staff
  - A single site with TeraGrid terminology would be very useful
  - Run a pilot with nay-sayers involved to watch and learn
    - Start with TeraGrid primer
    - Let it evolve based on user input
  - Roy would be glad to have a strategy discussion on use of wiki/wikipedia
  - Major risk is lack of an audience to make it a rich discussion and evolution
- Facebook is a good way to connect our friends' networks!

Education and Training

- An EOT page linked from the TeraGrid home page has been created with lists of TeraGrid workshops, institutes, and conferences as well as pointers to on-line tutorials, courses and other resources
- The on-line tutorials have proven to be very popular with the community
- We are working to balance live workshops, with Access Grid enabled events, with on-line self-paced learning opportunities
- SC07-09 Education Programs will focus on working with undergraduate faculty to integrate computational science and grid computing into the curriculum. A number of MSIs have been enlisted and we're looking for more institutions that want to participate.

Education and Training Feedback

- More training on all aspects is needed - new technologies, software systems, data handling, bio/bioinformatics, regular updates, etc. for new-comers and institutions new to TeraGrid (including MSIs)
- More tutorials are needed about TeraGrid
- Many useful on-line learning tools – like Breeze allow more interactivity and allow the learner to ask questions of the instructors, an improvement over talking heads approaches
  - Want access to slides
  - Allow writing and annotating
  - Layering representations
  - Instructor can see what students are marking up as most confusing or important
  - Stanford is experimenting with representations

- Breeze has been effective for live workshops
- CI Channel has been good for broadcast of events – encapsulating the materials in another document could prove to be very useful
- Emphasis is also needed on integration of TG resources into the curriculum
- A major need in the bio/bioformatics area
- GGF taking some good steps to foster EOT collaboration between EU and US – work of Malcolm Atkinson has been good

Ease of Use

- Recently held an Authentication, Authorization and Account Management workshop to reduce the challenges for people getting started and using TeraGrid resources
- Charlie has found that the Amazon AAA process is quite easy and quick

Ease of Use Feedback

- Amazon provides a good model for simplifying the Authentication, Authorization and Account Management processes
- Sensitivity is needed for writing software in relation to core programs

Acknowledgments

This report draws from notes and committee reports provided by Ralph Roskies (PSC), Scott Lathrop (UC/ANL), Fran Berman (SDSC), Phil Maechling (USC).

Appendix: CUAC Members and Groups

The CUAC was divided into groups in order to better support discussions (including ease of scheduling). The three groups are each facilitated by a TeraGrid leader.

GROUP A
Pat Teller (http://faculty.utep.edu/pteller)
P.K. Yeung (http://www.ae.gatech.edu/people/pyeung/)
Thomas Cheatham (http://www.chpc.utah.edu/~cheatham/)
Luis Lehner (http://www.phys.lsu.edu/faculty/lehner/)
Facilitator: Ralph Roskies (roskies@psc.edu)

GROUP B
Cathy Wu (http://pir.georgetown.edu/pirwww/aboutpir/wubio.html)
Phil Maechling (SCEC - http://www.usc.edu/directories/dept/scec.html)
Bennett Bertenthal (http://www.ccp.uchicago.edu/~bbertent/)
Gerhard Klimeck (http://dynamo.ecn.purdue.edu/~gekco/)
Facilitator: Fran Berman (berman@sdsc.edu)

GROUP C
Nora Sabelli (http://www.cilt.org/people/nsabelli.html)
Gwen Jacobs (http://neuron.montana.edu/faculty/gwenjacobs.html)
Alex Ramirez (http://conference.wcet.info/2004/program/speakerdetail.asp?id=5681)
Roy Pea (http://scil.stanford.edu/about/staff/bios/pea.html)
Facilitator: Charlie Catlett (cec@uchicago.edu)