

TeraGrid Collaboration Framework

Abstract

The TeraGrid is a nationally distributed cyberinfrastructure that provides leading-edge computational and data services for scientific discovery through research and education. TeraGrid is a collaborative project with multiple independent awards supporting and integrated enterprise that relies on clarity of shared objectives, expectations, process, and policy for effective operation. This document describes the ways in which the TeraGrid partners work together to achieve a common vision of enabling and accelerating scientific discovery using advanced computational cyberinfrastructure.

Table of Contents

1	Background	2
2	TeraGrid Organization and Coordination	3
2.1	Roles and Responsibilities of Participating Organizations	3
2.1.1	Grid Infrastructure Group (GIG)	3
2.1.2	Resource Providers (RP)	4
2.1.3	Joint responsibilities of GIG and Resource Partners	4
2.2	Coordinating Groups and Interactions	4
2.2.1	TeraGrid Forum (TGF)	4
2.2.2	Working Groups	6
2.2.3	Requirements Analysis Teams (RATs)	6
3	TeraGrid-Wide Policy Adoption	7
3.1	Distributing and reacting to policy proposals	7
3.2	Reaching actionable conclusions	7
3.3	Documentation of Policy Finalization	8
4	TeraGrid Operational Activities routine decision-making	8
5	Resource Requirements	9
6	Scaling	9
7	Security Considerations	9
8	Changes to the TeraGrid Policy Collaborative Framework	9
9	Related Documents	9
	Policy Drafts in Discussion	10
10	Acknowledgments	10
11	Author Information	11
12	References and Notes	12
13	Consensus Summary	13
14	Appendix A: List of current voting members of the TGF	14

1 Background

TeraGrid is a leading-edge cyberinfrastructure project funded by the National Science Foundation, comprising multiple resource provider (RP) partners and a Grid Infrastructure Group (GIG) that provides common services and software and coordinates key TeraGrid functions.

The TeraGrid project was originally funded based on proposals submitted to NSF in late 2004, with governance and policy addressed in the Grid Infrastructure Group (GIG, [1]) proposal based on instructions from the National Science Foundation within the solicitation [2] of the GIG and Resource Provider (RP) proposals. Later funding has come through other programs, including the Track 2 and HPCOPS solicitations. Initial descriptions of TeraGrid policy management process can be found in the 2006 TeraGrid Program Plan [3], submitted to NSF in December 2006.

This TeraGrid Collaboration Framework document builds upon the governance and decision-making processes from these various sources and makes modifications to increase clarity, and incorporate experience gained through the use of the original processes. This document outlines the basic roles of TeraGrid participating organizations, and the process for making joint decisions necessary to provide a national cyberinfrastructure, such as developing and modifying policy.

2 TeraGrid Organization and Coordination

2.1 Roles and Responsibilities of Participating Organizations

TeraGrid participating organizations (partners) include individual Resource Providers and the Grid Infrastructure Group (GIG). While TeraGrid aspires to present a single (thus fully integrated) presence to the national research community, it is administratively composed of multiple, independent awards (RPs) to many different institutions and one integrated award (GIG) funding staff at many institutions. One of the key purposes of this TeraGrid Collaboration Framework is to help the TeraGrid as a whole to deliver advanced cyberinfrastructure services to the national research community, and to ensure that the various independent grantees are each able to effectively deliver the services and outcomes identified in their respective agreements with the NSF. The intended outcome of this is the best possible delivery of high-end capabilities and services to the US research community.

With this service delivery as a goal, the TeraGrid Collaboration Framework has three overarching goals:

1. Enable the TeraGrid to manage services, implement new services, and react to the nation's cyberinfrastructure needs by enabling the finalization of actionable decisions within a predictable and reasonable time period, while allowing for some variation in implementation due to the particular roles and characteristics of participating organizations and the services they offer.
2. Ensure that when responsiveness is required among institutions participating in the TeraGrid in order for the TeraGrid to function effectively, that such responsiveness occurs.
3. Provide a coordinated mechanism via which the TeraGrid interacts with the NSF, advisory committees, and the national cyberinfrastructure user community (current and potential users).

2.1.1 Grid Infrastructure Group (GIG)

The Grid Infrastructure Group (GIG) is a distributed organization coordinated through the University of Chicago and funded through a cooperative agreement with NSF. The GIG is organized in functional "areas" with an "area director" (AD) overseeing the work of the areas and reporting to the GIG Director. ADs and GIG staff are funded through the University of Chicago GIG award. The GIG ADs and staff include staff from most TG partners, who are funded through sub-awards.

GIG-funded staff members provide a number of crucial common services such as software integration, networking, and operations. An important role of the GIG is to coordinate a variety of activities performed by staff from the participating institutions, such as user support, outreach, and project management. ADs are responsible for facilitating various project coordination and communication structures such as working groups and requirement analysis teams, workshops, and other meetings.

Another role of the GIG is to provide a useful and effective interface layer enabling the TeraGrid Resource Providers to effectively present well-defined resources that they are contributing to the TeraGrid for use by the national research community.

Executing these roles requires several key personnel in GIG leadership positions (e.g. GIG Director, Director of Operations, ADs, etc.). The GIG also charters and proposes working groups

and requirements analysis teams (see below) to help develop and provide integrated services and capabilities.

2.1.2 Resource Providers (RP)

TeraGrid Resource Providers design, deploy, and operate the resources that make up the TeraGrid facility, including computational, storage, and visualization systems, and network connections. Resource providers work with the GIG to develop and provide coordinated services to the user community using standard capability kits and service interface definitions, and enabling a broad spectrum of usage modalities ranging from traditional batch supercomputing on individual resources to workflow that use multiple resources to simultaneous co-scheduling of multiple resources. Where feasible (and this may change with time), TeraGrid partners try to make the software environment common across participating resources under the guidance of the Common User Environment working group.

2.1.3 Joint responsibilities of GIG and Resource Partners

The GIG and the Resource Providers are jointly responsible to the national research community and NSF for achieving the goals set out for the TeraGrid. The GIG and Resource Providers are obligated to be responsive to requests from one grant-receiving entity to another. In the case of sub-awards, the responsibilities of the awardee and sub-awardee are defined by the NSF Grant Program Guide, Cooperative Services Agreements with the NSF, and memoranda of understanding between the awardee and sub-awardees.

2.2 **Coordinating Groups and Interactions**

Several types of groups are used to facilitate coordination, planning, and operations. The *TeraGrid Forum* is a formal decision-making group that provides leadership and high-level coordination for the TeraGrid project. Working groups (WGs) and Requirement Analysis Teams (RATs) are used for day-to-day coordination and planning, respectively.

Periodic meetings are also used for coordination, planning, and operation. These include weekly project-wide, open Access Grid status and communication meetings, quarterly face-to-face meetings of the TeraGrid Forum and management staff, and an annual open conference. Periodic focused workshops are used as needed for detailed planning.

2.2.1 TeraGrid Forum (TGF)

The TGF is the high-level body responsible for providing the leadership for the RPs and GIG to work together effectively to achieve the TG vision. The TGF facilitates building the consensus needed to develop and implement TeraGrid policies among interdependent partners. The TGF is not formally an executive body since it does not have authority over independent cooperative agreements; however, as the one body comprising the PIs of each of those CAs, it is agreed upon by TG partners to be the formal body for facilitating decisions (through discussion, with a goal of consensus) as well as sharing information.

The TGF facilitates the coordination and collaboration among the RP and GIG team regarding operational and policy issues common to RPs, GIG services, and functions of the GIG including policy, technical approaches and roadmaps. The TGF convenes as appropriate via distributed communication mechanisms (e.g. Access Grid, teleconferences, email) and meets in-person as part of the quarterly project meetings.

The TGF is comprised of one representative (PI) from each of the RPs and one representative from the GIG. Each TGF member shall appoint a single “proxy” member who may act in the place of the TGF member and may participate in all TGF discussions and meetings. TGF meetings are generally open to all TG and NSF staff. Sub-awardees of the RPs and GIG are to be represented in the voting process via the institution that is the prime on the NSF grant, but they are welcome and encouraged to participate in TGF meetings and discussions.

The TGF approves the creation of requirement analysis teams (RATs) and all working groups (WGs) for the project (which are chartered and submitted for approval by the GIG). Recommendations of RATs are reviewed and shared with the TGF who may then vote on the recommendations. It is the Chairperson’s duty, insofar as the funding and facilities are available, to see that the recommendations of the TGF are carried out.

On issues agreed in advance by consensus as requiring votes, the TGF acts as the voting body required to make decisions for the project. The goal is always consensus, but to facilitate rapid progress, 2/3 majority votes can approve all new measures with the exception of election of positions (i.e. the TGF Chairperson). However, any issue requiring voting will require one week (either one week by email, or one week advance notice for an in-person/in-call vote) in order to provide sufficient time for all TGF members to review the issue.

2.2.1.1 TeraGrid Forum Chair

The TGF is coordinated by the TG Chair. [This role may also be held by two co-chairs, but is hereafter referred to as the Chair, representing the position itself rather than the one or two individuals holding it]. The Chair serves as the leader for the TeraGrid Forum and thus coordinates the overall activities and responsibilities of the TGF. The position is not formally an executive position, though it is filled by election to provide leadership in bringing the team together to discuss and make decisions on high-level issues important to the success of the project.

The TGF Chair position is elected by simple majority vote of the TGF. (In the case of persons agreeing to be co-chairs, they are voted on as a pair.) The TGF Chair serves a 12-month term, which are renewable by election. Any TGF member may call for a vote of no-confidence, with a 2/3 vote required to remove the TGF Chair prior to the end of their term.

The Chair sets the meeting agendas and runs the TGF meetings (weekly or bi-weekly calls, quarterly in-person meetings).

The Chair coordinates the TeraGrid policy development/revision process and works closely with the GIG PI and RP PIs to encourage implementation of approved policies.

The Chair coordinates efforts by each RP PI and the GIG PI to develop the TeraGrid annual report and program plan, prepare for the annual review and coordinate review responses. The Chair will work with the TG-wide project manager and the TG project management working group in this activity.

The Chair serves as an advocate and may represent the TeraGrid at conferences, workshops, and other events (though this role is not limited to the Chair). The Chair also engages in discussions and serves as the representative of the TeraGrid with other grid organizations and vendors. The Chair reports on meetings and recommendations to the TGF who then act on the recommendations.

The Chair also coordinates with and reports results and recommendations from advisory groups and review panels to the TGF. It is the Chair’s duty, insofar as the funding and facilities are

available, to see that the recommendations of the advisory boards and review panels are fairly evaluated.

2.2.2 Working Groups

Ongoing coordination of TeraGrid activities and functions takes place in standing “working groups” that are led by GIG and/or RP staff members. Working groups play a vital role in communication and coordination within the TeraGrid project. Members of a working group both participate in the group and ensure that their local management and co-workers are informed of the discussions and appropriately consulted for decisions made in the working group.

Each working group is chartered by the Director of the GIG, and approved by the TeraGrid Forum. Leadership of the WG is assigned to a TG project member and reports to an appropriate GIG area director.

Working groups are intended to be persistent groups with representatives from each RP. While each RP must provide a point of contact for each working group, participation is optional in working groups with the exception of:

- User Services (coordination of user support across TeraGrid)
- Security (coordination of security policy, practice, and incident response across TeraGrid)
- Software (coordination of software deployment on TeraGrid RP resources)
- Accounting (coordination of all processes related to the allocation of TeraGrid resources)
- Project Management (coordination of all project management processes across TeraGrid)
- Common User Environment (coordination of the user interface and experience across all TeraGrid resources)

Active participants in a working group are designated by each RP PI to the working group chair and provide response and appropriate action on issues relevant to the participant’s site within roughly one week. For informational purposes, any TG staff member may subscribe to working group e-mail lists. In addition, certain WGs may also include TG users or other relevant at-large members.

2.2.3 Requirements Analysis Teams (RATs)

The Requirements Analysis Teams (RATs) fill several project needs. First, there are many issues that must be addressed, or planned, but which do not squarely fall within any particular working group. Second, it is often advantageous to form a small group of experts rather than use a large group, and most working groups contain at least several staff from each TeraGrid institution (and thus are very large).

In contrast to Working Groups, RATs are not persistent bodies and are typically small teams that may not have representatives from all RPs. They may be more likely to draw significantly on external expertise as well (e.g. TG users). A RAT is defined by a charter with a clearly defined statement of the issues to be addressed, a set of outcomes, and a schedule of milestones to be accomplished in a period of 6-12 weeks. Generally a RAT will create a recommendation or set of recommendations.

RATs are chartered by the GIG and approved by the TeraGrid Forum and, like working groups, are assigned to a TG project member. RATs are typically staffed with volunteers from RP sites and the GIG team, and potentially at-large members with appropriate expertise for the specific purposes of the RAT. RAT members are expected to spend a significant amount of time doing

the work described in the RAT charter, and often this involves a level of effort not readily sustainable on an ongoing basis such as in a working group.

3 TeraGrid-Wide Policy Adoption

Because of the complexity of the TeraGrid, changes in policy and new policy implementation are particularly important. It is essential to have time to reflect on and refine new policies and proposed changes to policy. At the same time, it is essential for the functioning of the TeraGrid to be able to move rapidly in order to meet the cutting-edge needs of the US research community and priorities set by the NSF. With a large distributed organization it is neither realistic nor desirable to have action possible only when there is unanimity. However, when there is clear critical mass, it must be possible to move forward effectively while respecting the reality that not every subunit within the TeraGrid will implement every new decision or every aspect of every new policy.

TeraGrid policy development and planning related to system changes, policy implementation, or other activities may begin in a variety of venues including working groups, RATs, the GIG management team, the TeraGrid Forum, individual participants, or advisory structures. Standard document templates for policy [5] and significant decisions, changes, or activities [6] are used to facilitate clear communication and necessary context and provenance information for the purposes of policy and planning. These templates include several required sections, including security impact, resource requirements (RP and GIG impact such as operations, staffing, etc.), and scaling (impact on new resources and/or RPs)

3.1 Distributing and reacting to policy proposals

Any proposal put forth for consideration by the TeraGrid Forum will be proposed formally by one of the designated voting representatives of the TeraGrid Forum (or designated proxy). This may be done directly by distribution to the TeraGrid Forum members, or via submission to the TeraGrid Forum Chair and then dissemination by the Chair. Proposals put forth for formal consideration will be distributed by email to the TeraGrid RPF mailing list and by posting the document in the TeraGrid Forum Wiki in the policy document portion of the hierarchy.

Within two weeks of mailing new or revised policies to the TGF, each TGF member (or proxy) shall respond with proposed modifications, proposed alternative policy, and an indication of level of support for the proposed policy overall (with one of four responses being the essential summary: strongly agree, agree, disagree, strongly disagree). [This is not a 'vote' but a means of measuring overall team support for the policy at this time.]

3.2 Reaching actionable conclusions

TeraGrid is primarily a consensus-driven system. The consensus development around a given policy will be documented for future reference both to capture any concerns and modifications related to those concerns and to capture the rationale behind, and context in which the policy was created. The process of cycling through responses should take no more than two weeks from start to finish, and an original proposing organization may at their discretion call for a vote on a proposal at the end of two weeks, to be completed in person, via phone, or by email over a period of no more than 7 business days after making the request to call for a vote.

The consensus voting process used for adoption of TG-wide measures will be by 2/3 majority vote of the member institutions.. A decision will be declared, and a policy adopted, whenever at least 2/3 of the TGF institutions have cast a vote approving a policy.

All votes will be announced five business days in advance. Exceptions to this will be allowed by agreement of all TGF voting members. In the case of a vote taken at a meeting or via a phone conference, voting institutions that will not be present are required to submit vote to chair to be kept in confidence until the vote; failure to do so will constitute an abstention and that entity will not be counted as voting on that particular matter.

For decisions involving policy or other documented actions, a table shall be included with the document recording the position of each TGF member and the date of consensus. All other decisions will be recorded in standard meeting minutes.

Where necessary for consensus, exceptions or RP-specific modifications can be made for unique local circumstances or policy conflicts for a particular RP. These exceptions are noted in the consensus summary of the policy or decision document.

Local conditions may require certain RPs to take a different position from the consensus on a particular issue. If an RP registers disagreement in the consensus summary, then the dissenting RP may “opt out” of implementing the consensus. A dissenting RP may not “opt in” by representing its alternate position as equivalent.

All TG Forum meeting minutes will be publicly posted, with the exception of the portions of the meetings in “executive session” as defined by the TGF Chairperson and approved by the TGF.

3.3 Documentation of Policy Finalization

All TeraGrid policy will be recorded in the form of a policy document within a numbered permanent document series, maintained by the GIG management team and available on a project website. Once a document has been published in the numbered series it may not be changed. Policy changes to an approved policy may be implemented through publication of a new policy document, indicating any prior documents that are voided by the new policy [7, 8, 9].

Upon finalization, a document will be added to a numbered TeraGrid document series that will be persistent and can be cited by other TeraGrid documents. By default the TeraGrid document series will be publicly available, though some documents may be kept internal for security or other reasons, at the discretion of the TGF Chairperson after consulting with the TGF.

A final document will include an assigned document number in the header field (upper left) and the status of the policy, including effective dates, will be included in the “Status” section on the cover page of the document.

A summary of the consensus outcome will be included as an appendix to the policy, indicating the names/affiliations of each partner representative, their agreement status and date provided.

4 TeraGrid Operational Activities routine decision-making

The TeraGrid will conduct normal decision-making and coordination activities other than policy adoption via email, calls, and in-person meetings. The TeraGrid is, as stated above, primarily a consensus-driven organization. For effectiveness in use of time, however, Robert’s Rules of Order will govern in-person and phone meetings, with an expectation that most motions made will pass by acclamation.

5 Resource Requirements

This policy framework requires minimal resources. The main impact is on the elected TGF Chair for her/his responsibilities, and GIG and RP management staff for consideration and timely response. Consideration of policies developed under this framework will likely require modest time investment from management and staff, but again, consideration of and ratification of policy should be part of current management scope, not additional effort.

Administrative support will be required for administering the processes described here, the policy document repository, and the requisite minutes.

6 Scaling

The roles and responsibilities of GIG and RP will have potential scaling issues based on the size of the group that is involved in reaching consensus. At present there are eleven RPs and this type of consensus should scale to at least twice this number, and with proper organization, well beyond that.

Impact on new resource providers will be in the form of effort required to participate in the TGF as well as the required working groups; however, this level of effort is low relative to the effort of providing a resource to the national user community.

7 Security Considerations

This document outlines process for policy-making and decisions. The document notes that all RPs must designate a contact person to participate in the Security Working Group (§2.2.2). The document is otherwise silent on security issues and has no other security implications.

8 Changes to the TeraGrid Policy Collaborative Framework

This is an evolving document and changes and additions are expected as TeraGrid matures and grows. Changes to this document will be through the TeraGrid Forum with input from the NSF through a simple majority vote of the TeraGrid Forum.

9 Related Documents

There are a few related documents that can be referred to for important relevant topics that are in addition to this collaborative agreement. These policies include:

- [TG-4](#): TeraGrid Security Memorandum of Understanding (MOU)
- [TG-5](#): TeraGrid Certificate Management and Authorization Policy
- [TG-6](#): TeraGrid Staff Roaming Grants Policy
- [TG-9](#): TeraGrid Community Software Areas
- [TG-10](#): TeraGrid Community Account Policy
- [TG-12](#): TeraGrid Logo Guidelines
- [TG-15](#): TeraGrid Persistent Storage Allocations Policy
- [TG-16](#): TeraGrid Storage Peer Review Option

Policy Drafts in Discussion

- [TGD-7](#): Draft TGD-7 - Integrate New RP
- [TGD-13](#): Draft TGD - Privacy Policy (still in initial draft)
- [TGD-14](#): Draft TGD - User Access Policy (still in initial draft)

10 Acknowledgments

This work was supported by the National Science Foundation Office of Cyberinfrastructure, grant number 0503697 “ETF Grid Infrastructure Group: Providing System Management and Integration for the TeraGrid.” Concepts outlined in this document have come from substantive input and feedback from Craig Stewart (IU), John Cobb (ORNL), Jay Boisseau (TACC), Richard Moore (SDSC), Phil Andrews (NICS/UTK), Fran Berman (SDSC), Mark Sheddon (SDSC), Kelly Gaither (TACC), Michael Levine (PSC), Ralph Roskies (PSC), Mike Papka (UC/ANL), Gary Bertoline (Purdue) and many other TeraGrid participants.

11 Author Information

Charlie Catlett
University of Chicago / Argonne National Laboratory
cec@uchicago.edu
+1-630-252-7867

Dane Skow
University of Chicago / Argonne National Laboratory
dds@uchicago.edu
+1-630-252-8694

Craig Stewart
Indiana University
stewart@iu.edu
+1-812-855-4240

Ralph Roskies
Pittsburgh Supercomputing Center
roskies@psc.edu
(412) 268-4960

Gary Bertoline
Purdue University
Bertoline@purdue.edu
765-496-6071

John Cobb
Oak Ridge National Laboratory
cobbjw@ornl.gov

Jay Boisseau
University of Texas-Austin / TACC
boisseau@tacc.utexas.edu
512-475-9451

John Towns
University of Illinois / NCSA
jtowns@ncsa.uiuc.edu
+1-217-244-3228

12 References and Notes

- [1] Catlett, C. et al., Grid Infrastructure Group proposal, October 2004.
- [2] Letters from the National Science Foundation were issued to the TeraGrid team in August 2004 outlining the overall program and requesting proposals for the SMG (now GIG) from the University of Chicago and for RP support from all TeraGrid sites. The two letters- describing the SMG (GIG) and RP roles and proposal requirements, are referenced in this document as a single entity for simplicity because all referenced text is common to both letters.
- [3] TeraGrid 2006 Program Plan, submitted to NSF in December, 2005.
- [4] Thompson, K., Instructions (via email) on formation of the CUAC.
- [5] Catlett, C., "TeraGrid Policy Document Template and Required Information". TG-2, March 2006.
- [6] Stewart, C., Catlett, C. "TeraGrid Internal Proposal Template and Required Information." TG-3, March 2006.
- [7] The use of a numbered document series has proved to be useful in standardization and large-scale enterprise policy efforts such as the Internet Engineering Task Force (IETF, produces standards and policies for the Internet) and Global Grid Forum (GGF, produces standards and policies relevant to Grid computing). For more detail on the mechanics of such processes see [8] or [9].
- [8] Bradner, S., "The Internet Standards Process – Revision 3," RFC 2026, www.ietf.org, October 1996.
- [9] Catlett, C., "Global Grid Forum Documents and Recommendations: Process and Requirements," GFD-1, www.ggf.org, December 2001.

13 Consensus Summary

Partner	Name	Yea/Nay	Date
Grid Infra Group (GIG)	Matt Heinzl	Yea	9/9/2008
NCSA	John Towns	Yea	9/9/2008
Purdue University	Carol Song	Yea	9/9/2008
UC/ANL	Joe Insley	Yea	9/9/2008
Indiana University	Craig Stewart	Yea	9/9/2008
ORNL	John Cobb	Yea	9/9/2008
TACC	Jay Boisseau	Yea	9/9/2008
SDSC	Richard Moore	Yea	9/9/2008
PSC	Michael Levine	Yea	9/9/2008
NCAR	Rich Loft	Yea	9/9/2008
LONI/LSU	Dan Katz	Yea	9/9/2008
NICS/UTK	Phil Andrews	Yea	9/9/2008

14 Appendix A: List of current voting members of the TGF

Partner	Representative	Alternate
Grid Infra Group (GIG)	Matt Heinzl	Ian Foster
NCSA	John Towns	Tim Cockerill
Purdue University	Carol Song	Gary Bertolini
UC/ANL	Mike Papka	Joe Insley
Indiana University	Craig Stewart	Steve Simms
ORNL	John Cobb	Jeff Nichols
TACC	Jay Boisseau	
SDSC	Mark Sheddon	Richard Moore
PSC	Michael Levine	Ralph Roskies
NCAR	Rich Loft	Tom Bettge
LONI/LSU	Dan Katz	Charlie McMahon
NICS/UTK	Phil Andrews	Patricia Kovatch