Grid Community Characteristics and their Relation to Grid Security

Markus Lorch, Dennis Kafura

Department of Computer Science
Virginia Polytechnic Institute and State University
{mlorch, kafura} @vt.edu

Abstract

The size, dynamics, composition and similar characteristics of Grid Communities constitute important data for Grid security requirements gathering and analysis. Collaborative Grid Communities are especially important as they constitute an important part of grid usage modes and demonstrate the need for more advanced Grid security solutions very clearly. This document reports the results of a survey conducted in the Fall of 2002 among members of the Grid community as to understand the needs of grid user and grid application developer communities today, provide information on and typical modes-of-use, and elicit requirements for future grid security systems.

1. Introduction and Motivation

The size, dynamics, composition and similar characteristics of Grid Communities constitute important data for Grid security requirements gathering and analysis. Collaborative Grid Communities are especially important as they constitute an important part of grid usage modes and demonstrate the need for more advanced Grid security solutions very clearly. This document reports the results of a survey conducted in the Fall of 2002 among members of the Grid community as to understand the needs of grid user and grid application developer communities today, provide information on and typical modes-of-use, and elicit requirements for future grid security systems.

The salient results of this survey are summarized immediately below. Appendix A provides a table with the survey questions and the answers provided by the respondents. The survey results are also available online at http://zuni.cs.vt.edu/grid-security. The online version leverages Virginia Tech's survey tool (http://survey.vt.edu) and allows for more advanced analysis of the data by visualizing answer interdependencies. In addition the results are available as a data file from the author.

2. Summary of Survey Results

The survey was announced and distributed at the sixth meeting of the Global Grid Forum. Most (77%) of the 39 respondents classified themselves as grid researchers and developers.

The survey showed that collaborations often cross organizational lines, underscoring the significance of inter-organizational trust. The responses showed a high level of communication and collaboration with people from outside the respondent's department (90%), outside their company or organization (82%), and outside their country (69%). The likelihood of sharing grid resources with collaborators decreased with larger distance: most share resources with collaborators from outside their department (77%), about two-thirds (67%) with people from outside their company or organization and less than half with others from outside their country (46%).

The majority of today's Grid Communities tend to be relatively small in size. Most (59%) respondents were members of a grid community. The community size was distributed as follows: 60% fall in the category of 1 to 25 members, 17% have 26 – 100 members, no membership was reported for communities of size 101 to 200, and 23% are large communities with more than 200 members. These communities consist equally of single projects (49%) as well as for multiple projects (49%); only one correspondent was a member of a community that had a lifetime of only a single task. Most respondents (71%) that are member of grid communities answered that they make use of less than 10% of the total number of community resources, which in 64% of all cases was between 1 and 50 resources. Only 16% of the communities shared more than 50 resources.

The plurality of grid users decide for themselves with whom they want to collaborate and what rights should be allocated to their collaborators. The communities of the respondents are equally divided into those using a structured hierarchical paradigm for administration and those using an ad-hoc manner, where membership is achieved through joining and contributing. In more than half (54%) of the cases of ad-hoc communities, privileges are delegated and assigned in a hierarchical manner through a community administrator, which may be evidence of the lack of support for efficient management of privileges in ad-hoc grid collaboration scenarios. The plurality of respondents (41%) decide for themselves with whom they share resources, of which 37.5% have to contact an administrator in order to grant the corresponding access rights.

Forty-six percent of communities use a resource administered authorization mechanism such as the Grid Security Infrastructure (GSI), 18% use community administered authorization mechanisms such as the experimental Community Authorization Service (CAS) or the Virtual Organization Membership Service (VOMS), a small fraction (5%) uses distributed authorization e.g. as in AKENTI and another 5% were not sure what they use. Eight percent responded that they leverage other systems, such as DCE and Kerberos.

Asked about their beliefs about the security features provided by grid software in collaborative scenarios, 53% answered that existing grid security solutions do not provide adequate services for collaborative grid communities. The reasons given ranged from the lack of an underlying threat model to the complexity and expense of inter-site trust relationships that are currently required. 59% believe that existing grid security solutions impose too much administrative overhead for efficient collaborative use and. 47% thought that the level of security provided by existing solutions (e.g. encryption strength, security of the protocol, non-repudiation guarantees) is not sufficient, however 38% indicated that they believe it is sufficient. The plurality (48%) of respondents did not believe that maintaining a public key infrastructure (PKI) for grid services is too expensive vs. only 20% that doubted the adequacy of maintaining a PKI.

Mechanisms for a user to manage his own privileges and credentials seem to be insufficient. This was the opinion of 23 respondents (59%), 26% were undecided about this issue and only 10% thought the mechanisms were sufficient (5% did not provide an answer). On the question of whether real world trust relationships can be modeled adequately using existing grid security solutions most respondents (56%) disagreed. The reasons given revolve mostly around immature trust models, the observation that collaborators trust each other but not their respective institutions, and the issue of users having multiple roles. Asked specifically if trust relationships between collaborators can be correctly modeled 46% indicated their disagreement, where as only 28% could agree to this question and 23% were undecided. One of the comments given was

stressing the need to "*push privilege management down to the individuals, which poses a challenge to the tools and plumbing*".

3. Summary

Summarizing the survey responses, it seems apparent that existing, first generation security solutions are adequate only for a subset of today's use-cases in grid environments. These mechanisms focus on traditional solitary and clearly defined collaborative scenarios and require additional functionality to provide for the full expressive range necessary for more secure dynamic collaborations and least-privilege access schemes. It is our conclusion, that one of the biggest hindrances to the efficient functioning of grid communities is the administrative overhead and lack of user controlled privilege management.

Appendix A – Survey Questions and Responses

The results consist of 39 responses provided between Oct. 14th and Nov. 27th, 2002.

Part 1: Personal demographics, computer usage and community membership

Your relation to grid computing

3 (8%) grid user

9 (23%) grid project manager

30 (77%) grid systems researcher / developer

1 (3%) other

How long have you been involved with grid technologies

5 (13%) 6 months or less

5 (13%) 6 - 12 months

17 (44%) 13 - 24 months

11 (28%) longer than 24 months

1 (3%) no answer

Do you communicate and collaborate frequently with people from

35 (90%) outside of your department

32 (82%) outside of your institution / company

27 (69%) outside of your country

Do you typically share resources with collaborators from

30 (77%) outside of your department

26 (67%) outside of your institution / company

18 (46%) outside of your country

Who decides with whom you can share your resources (e.g. your data files, your programs)?

16 (41%) you yourself

11 (28%) your manager / department head

4 (10%) your site administrator

5 (13%) I do not share resources

3 (8%) other

0 (0%) no answer

Who grants access to resources that you share with others?

15 (38%) you yourself

4 (10%) your manager / department head

13 (33%) your site administrator

5 (13%) I do not share resources

2 (5%) other:

0 (0%) no answer

Are you a member of a community that is based on grid technology (e.g. a virtual organization)

16 (41%) not a member of a grid based community

13 (33%) a member of 1 community

6 (15%) a member of 2-4 communities

4 (10%) a member of 5 or more communities

0 (0%) no answer

If the respondent answered "not a member of a grid based community" to the above question she was asked to skip the remaining questions in part one of this survey and proceed with part two.

The following six questions refer to the community the respondent is/was most engaged with:

What is the size of the community you are most engaged in

9 (23%) 1 - 12 members

6 (15%) 13 - 25 member

2 (5%) 26 - 50 members

1 (3%) 51 - 75 members

1 (3%) 76 - 100 members

0 (0%) 101 - 150 members

0 (0%) 151 - 200 members

6 (15%) more than 200 members

14 (36%) no answer

How is membership in this community defined?

12 (31%) structured, clearly defined membership

11 (28%) ad-hoc, one becomes a member by joining and contributing

2 (5%) other

14 (36%) no answer

What is the lifetime of this community?

1 (3%) short (e.g. a single task)

12 (31%) medium (e.g. a single project)

12 (31%) long (e.g. multiple projects)

14 (36%) no answer

How are privileges delegated or shared among community members

11 (28%) hierarchical, through assignment by a community administrator

11 (28%) peer to peer, through direct delegation by any entity who is authoritative (e.g. you can directly assign access permissions to one of your files to another community member)

2 (5%) other

15 (38%) no answer

In average, what percentage of the total number of shared resources (e.g. computational nodes, databases) available to you do you use frequently?

18 (46%) less than 10%

2 (5%) 10 - 24%

2 (5%) 25 - 49%

0 (0%) 50 - 74%

1 (3%) 75 - 100%

2 (5%) I don't know

14 (36%) no answer

What is the absolute number of this subset of resources?

12 (31%) less than 10 resources

4 (10%) 11 - 50 resources

0 (0%) 51 - 100 resources

0 (0%) 101 - 200 resources

4 (10%) 201 or more resources

5 (13%) I don't know

14 (36%) no answer

The following four questions relate to all the communities the respondent is a member of:

What is the average size of all the communities your are engaged in

10 (26%) 1 - 12 members

3 (8%) 13 - 25 members

3 (8%) 26 - 50 members

1 (3%) 51 - 75 members

0 (0%) 76 - 100 members

0 (0%) 101 - 150 members

0 (0%) 151 - 200 members

8 (21%) more than 200 members

14 (36%) no answer

What is the average lifetime of the communities you are a member of?

2 (5%) short (e.g. a single task)

12 (31%) medium (e.g. a single project)

11 (28%) long (e.g. mulitple projects)

14 (36%) no answer

Please indicate which category of grid mechanisms are being used by the communities you are a member of:

12 (31%) synchronous communication tools (e.g. AccessGrid)

7 (18%) shared workspace environments

14 (36%) shared access to data files

10 (26%) indexing and information services

9 (23%) legacy tools like email lists and text based chats

7 (18%) other

What type of authorization mechanisms are being used in the communities you are a member of?

18 (46%) resource administered authorization (e.g. GSI)

7 (18%) community administered authorization (e.g. CAS / VOMS)

2 (5%) distributed authorization (e.g. Akenti)

2 (5%) I don't know

3 (8%) other

Part 2: Impact of grid tools on communities and security related questions

The next set of questions deal with the impact of advanced collaborative grid tools, such as multimedia synchronous communication tools, shared workspaces and collaborative problem solving environments.

Do you believe that such tools foster grid communities with close personal interactions among its members?

2 (5%) Strongly Disagree

0 (0%) Disagree

9 (23%) Undecided

18 (46%) Agree

10 (26%) Strongly Agree

0 *(0%) no answer*

Why do you believe so:

1. 10 years of VIC/VAT etc had little real impact
2. Our Users had these capabilities before the word "grid" became hot.
3. While collaborative tools help with widely distributed groups, face-to-face meetings are still very important.

4. Definitely helps when the members are geographically separated since it is not always the case that a phone or an email will do the job. One needs an option of both a synchronous and an asynchronous mode of communication as well as an easy way to share digital data.
5. they certainly help collaborative working but we are a long way off using them effectively and certainly not as effectively as face to face scenarios
6. yes - we've seen interaction patterns shift towards that seen between local researchers
7. Ability to share is a priority for Scientific growth
8. our uses had these capabilities before the word "grid" became hot
9. 10 years of VIC/VAT etc. -> little real impact
10. i use it, they use it

Do you believe that such tools improve community efficiency (less communication and synchronization overhead) when compared to traditional tools like email lists and chat rooms?

2 (5%) Strongly Disagree

4 (10%) Disagree

12 (31%) Undecided

13 (33%) Agree

7 (18%) Strongly Agree

1 *(3%) no answer*

Why do you believe so:

1. 10 years of VIC/VAT etc had little real impact
2. too complicated, non-uniform
3. The amount of time wasted configuring collaborative technologies is still very high compared to the value
4. I do not think the AccessGrid is less of an overhead compared to emails or chat rooms.
5. e-mail lists and chat rooms have specific and limited functionality - they have problems of their own. Communication and ways of working strategies are needed to define what technologies are used for what - playing to the strengths of each technology, but remembering that people need to communicate
6. why don't we use batch tools for all our desktop interactions with programs - synchronous interaction is important for certain tasks - problem solving, planning, debugging, exploring, ...
7. some truths can only come from rapid interchange of ideas
8. too complicated, non-uniform
9. 10 years of VIC/VAT etc. -> little real impact
10. i have now more communication

Do you believe that such tools provide a higher quality of interaction than traditional computer mediated communication tools?

2 (5%) Strongly Disagree

5 (13%) Disagree

7 (18%) Undecided

19 (49%) Agree

5 (13%) Strongly Agree

1 *(3%) no answer*

Why do you believe so:

1. more natural audio and video
2. 10 years of VIC/VAT etc had little real impact
3. we were fine with what we had
4. Existing tools provide very poor interactivity and communication quality
5. he potential is great but this is not realized yet - there are technology issues, cultural issues and behavioral issues to overcome
6. not sure what the comparison is with - A/V, shared screens, whiteboards, etc have existed for many years and they are improving steadily. Synchronous tools are good for different things than asynchronous tools.
7. can get nonverbal cues that reveal more than words
8. teleconference + showing an occasional powerpoint picture is adequate 99% of the time
9. we were fine with what we had
10. 10 years of VIC/VAT etc. -> little real impact
11. interactivity, desktop sharing

Do you believe that such tools will be useful for traditional online communities as well?

0 (0%) Strongly Disagree

2 (5%) Disagree

11 (28%) Undecided

18 (46%) Agree

6 (15%) Strongly Agree

2 *(5%) no answer*

The next set of questions deal with the applicability of existing grid security solutions to collaborative scenarios.

Do you believe that currently existing grid security solutions provide adequate security for collaborative grid communities?

6 (15%) Strongly Disagree

15 (38%) Disagree

9 (23%) Undecided

7 (18%) Agree

1 (3%) Strongly Agree

1 *(3%) no answer*

Why do you believe so:

1. Do not support VOs with no prior trust model in place
2. no realistic threat model underlies the tookits
3. Very large organizations need to create trust structure needed for sharing distributed high-value resources
4. I think the work on multiple credentials is definitely one of the solutions to a collaborative environment. The ability to play multiple roles is a necessity.
5. For highly secure collaboratories we need better solutions (like high assurance bridge CA's) for enabling and supporting inter-site trust relationships
6. too hard to use and too fragile to set up correctly (or requires too much work)
7. Depends what adequate means - not enough for commercially sensitive stuff, but fine for many scenarios
8. at scenarios that do not yet exist.
9. We are just begining to get set up, good for now, but the future???
10. too complicated, need finer grained delegation model
11. Weak storage of private keys
12. no realistic threat model underlies the toolkits
13. do not support VOs with no prior trust model in place
14. no support for ad-hoc collaborative scenarios, lack of flexibility

Do you believe that existing grid security solutions impose TOO MUCH administrative overhead for efficient collaborative use (e.g. require centralized administration)?

0 (0%) Strongly Disagree

11 (28%) Disagree

4 (10%) Undecided

17 (44%) Agree

6 (15%) Strongly Agree

1 *(3%) no answer*

Do you believe that existing grid security solutions DO NOT provide adequate service with respect to the level of security they provide (encryption strength, secure protocol, non-repudiation guarantee)?

0 (0%) Strongly Disagree

14 (36%) Disagree

6 (15%) Undecided

17 (44%) Agree

1 (3%) Strongly Agree

1 *(3%) no answer*

Do you believe that maintaining a public-key infrastructure for grid security is too expensive?

6 (15%) Strongly Disagree

13 (33%) Disagree

11 (28%) Undecided

6 (15%) Agree

2 (5%) Strongly Agree

1 *(3%) no answer*

Do you believe that the way existing grid security mechanisms allow you to manage your privileges and credentials is sufficient?

5 (13%) Strongly Disagree

18 (46%) Disagree

10 (26%) Undecided

4 (10%) Agree

0 (0%) Strongly Agree

2 *(5%) no answer*

Do you believe that the real world trust relationships between institutions that have collaborating entities can be modeled adequately using existing grid security solutions.

4 (10%) Strongly Disagree

18 (46%) Disagree

6 (15%) Undecided

8 (21%) Agree

2 (5%) Strongly Agree

1 *(3%) no answer*

Why do you believe so:

1. Neither X.509 not Kerberos have adequate semantics
2. grid-mapfile has all the power of assembly language
3. Although new trust models are evolving, they are not sufficiently mature for prime time yet.
4. Not all the users in the community have the same rights as the others in the community and a single person may have multiple roles to play. Both of these issues haven't been dealt with yet.
5. Some (not all) grid solutions are adequate with enough infrastructure work. For example, GSI via smart cards, using cross-certified CAs (or bridge CA's) and daily CRL checking should be adequate.
6. if tools are better ...
7. Trust is more than security!!!! - cultural elements - risk aversion/sharing/ trust in actions, behaviors and motives
8. it depends - VOs have been built without 'Grid' security (VPNs, commercial certs, etc.). Will we need more capable security models in the future - yes.
9. We need practice experience - a history first...
10. if you can't trust your own security then how can you trust someone else's
11. need more flexibility in the federation of credentials and selective exposure
12. Most often two individuals trust each other -- not their respective institutions!
13. "grid-mapfile" has all the power of assembly language ...
14. neither x509 nor kerberos have adequate semantics
15. firewalls prevent me from doing collaborations

Do you believe that the trust relationships between community members can be modeled correctly using existing grid security solutions?

4 (10%) Strongly Disagree

14 (36%) Disagree

9 (23%) Undecided

11 (28%) Agree

0 (0%) Strongly Agree

1 *(3%) no answer*

Why do you believe so:

1. Neither X.509 not Kerberos have adequate semantics
2. "can be"=yes, "will be"=no
3. I think trust models will evolve over the next several years
4. Right now all the security solutions seem to categorize people into the same class. Most communities are hierarchical in nature.
5. see above answer.....
6. The DOE2000 Diesel Collaboratory documented ongoing and dynamically changing exceptions to stated community security policies - there are no solutions that match their reality, just 80% solutions.
7. one has to push the privilege management down to the individuals, which poses a challenge on the tools and plumbing
8. "can be"- yes, "will be"-no
9. neither x509 nor kerberos have adequate semantics
10. have not used it yet. this is an issue that needs to be addressed differently

Do you have any comments or ideas related to this survey?

1. No
2. Excellent survey. just going through these questions made me realize that I havent given thought to a lot of finer details. I hope Markus will be putting up some results from this survey so that I know what the general consensus is for these issues.
3. High-dollar resources and data that would have high consequence if disclosed or modified will reside behind firewalls. Many user desktops are protected by an institutional firewall. I don't think the current grid technology adequetely addresses collaboration over extranets that operate through these firewalls.
4. Although we are not yet part of a grid based community we are hoping to be in the near future.
5. I find the survey odd - the security requirements for collaboration, and the usefulness of collaboration are independent of the size of data and computations - there have been a wide range of collaboratory projects/virtual organizations, with a wide range of security needs and uses/styles of collaboration. Is there any reason to expect that putting the word Grid in front of things will change the trust relationships between people or the effectiveness of how they collaborate? (Yes, the cost and generic nature of typical Grid resources require strong security, but does that change how collaborations occur? is it different than the need to protect the one paragraph statement of a patentable idea by people who would not consider themselves connect to Grids in any way?
6. Good luck!