

1.0 Goals

One goal of the documentation alignment initiative is to review the documentation available for users, engagement, site, VO and security administrators and training schools and identify areas where we can make more consistent, complete and aligned information [G1]. The [bracketed] items in this section are used as identifiers to verify coverage of the goals of the initiative.

Another goal is to improve the structure and maintenance of the TWiki web site itself [G2]. The existing TWiki-based documentation is in a mixed state. Some *TWiki webs* contain documentation from working groups in various developmental stages and degrees of quality. Some have been carefully reviewed and of are relatively high quality.

A third goal is to provide ways for dynamic comment and editing integrated into the information structure in order to enable the readership of the pages to make immediate contributions when they find missing, incorrect or misleading information as they actually read and use the web pages [G3].

The following subsections itemize problems with the current documentation system that will be addressed by this initiative.

1.1 Processes for Creation and Review of Document

In the ReleaseDocumentation web, we have a framework for assigning “Responsibles” and assigned “Reviewers”. This is accomplished with TWiki directives at the bottom of the document, and an automated summary page giving a snapshot. This framework should be reviewed, and processes around creation and reviewing better defined.

- Create process descriptions and associated policies [G4]

1.2 “Documentation Hub” and Common Framework

The Documentation hub (i.e. the Documentation TWiki web) is poorly organized and its maintenance is not sustainable since it points into documents from various webs and sometimes without respecting context of the source. The organization of this web needs to be properly factorized and put into a sustainable framework. It needs to reach into other webs, but in reliable way so that navigations to those dependent webs can be done with confidence. One way would be to provide in each web a tag that represents its quality. With Anne’s reassignment as iSGTW editor the ownership of this page has been orphaned.

- Find an owner or deprecate the page.
- Remodel Documentation web [G5]
- Define a system for tagging “certified” documents within webs [G6]

- Create common framework to align certified documents across TWiki webs [G7]

1.3 Annotating existing documents

Often times readers and users of the documentation will find problems or have suggestions feel comfortable editing the documents (especially new users, or users quickly working through an issue who may not have time to carefully make changes). We need an easy way for users of documentation to add comments without necessarily making significant changes to the material.

- Implement an user annotation tool/system that users can easily leave comments and includes: comment, name, date [G3]

1.4 Ongoing maintenance of documentation

The most difficult problem with any documentation system is how to keep the documentation contents maintained and up to date. The above processes should produce a good set of documentation, but there needs to be an ongoing process for updating the content to capture the current architecture, design, software, and procedures in OSG.

- Define a process that will encourage and verify that the contents of the documentation are up-to-date. [G8]
- Define plan for removing stale documents [G9]
- Define any policies or rules about removing stale documents [G10]

1.5 Management of the Migration

Once all the above is settled for the common framework, annotation tools, and policies a migration of various documents in various TWiki webs into the certified framework needs to be managed. There should be an initial big put into the system from all sources, especially the following webs [G11]:

- Documentation
- ReleaseDocumentation
- Security
- Storage
- Education and Outreach

2.0 Documentation Alignment Project Approach

This plan for this project consisted of the following steps which also define the basic structure of the remainder of this document:

1. Understand and Document the Current System – We must understand what is there in order to change it. (Section 3)

2. Interview User Representatives – The system exists to meet their needs, so we need to understand those needs, how they find and use the documentation, and what they like and would like to see improved. (Section 4)
3. Interview Providers – The providers of documentation produce the existing result based on their goals, incentives, tools, systems, knowledge, and direction. (Section 5)
4. Review literature on wikis – This step was added during the project since the major system at the center of OSG documentation is a wiki. A selected set of studies of wikis was reviewed to distill useful information that applies to OSG. (Section 6)
5. Convene Team to Develop Desired Documentation Structure & Processes – Given a summary and analysis of what was learned in the first four steps, a team of experts, closely associated with the documentation system was used to refine the analysis and develop the system and process to be used in the future. (Section 7)
6. Project for Migration – Once the team produced the common framework, annotation tools, and policies, a plan for the creation of the new/changed system and the subsequent migration of various documents into the new system was developed. (Section 8)

3.0 The Current Documentation System

We begin by an analysis of the current documentation system. We are including the entire system and all its suppliers (documentation writers) and customers (documentation users) so that any enhancements we implement can be done in a way that complements (or, at least, does not harm) other parts of the system.

2.1 Documentation Users

The users of OSG documentation fall into the following groups:

1. End Users – Scientists who are using OSG to do their work. Although most in this group are supported by people in a VO application layer, some of those in the Engage and At Large (OSG) VOs are more directly supported by OSG documentation.
2. System Administrators – People who are installing and administering the systems and services at various OSG sites. These are the primary and most important users. The documentation needs of a System Administrator new to OSG differ from an OSG experienced System Administrator.
3. VO Administrators – People who administer a VO's appearance on the grid and adds users and assigns roles. This group includes people who are managing the VO information and registration in the OIM.
4. VO Managers – Managers in VOs who are interested in understanding and/or influencing the current status, performance, or upcoming changes in OSG.

5. Students – Students in various OSG courses, including on-line courses, use the documentation system for course schedules, previous course presentations, lab work, or for the courses themselves.
6. OSG Internal – Internal OSG groups access documentation and use the documentation system to plan, manage, and document their activities and results of their meetings. They also use the documentation system to store the status of projects, various reports, papers, studies, etc.
7. VO Application Layer – Developers of the VO application layer customized for the scientists in the VO. They use OSG/VDT interface documentation in developing their application layer. For most VOs, these developers provide the End User interface and documentation for their users.
8. Potential OSG Members – University and science groups that are interested in OSG and who use the documentation to understand OSG, including the effort, benefits, and other implications of their joining.
9. Public – Everyone else, including the general public, use some part of the documentation system (e.g. starting with the OSG Web site).
10. Funding Agencies – The DOE and NSF may access OSG documentation, especially information associated with funded project status or the metrics that OSG has used to measure its progress.

Representatives from the most of above groups were interviewed to provide input on their view of the OSG documentation system. Interviewee comments could be roughly categorized into: a) structural or information organization, b) ability to locate information, c) quality and accuracy, and d) specific feedback on one or more documents. Each interview began with a prepared set of questions specific to each of the groups; but, depending on the experiences of the interviewee, the discussion was quite varied even within a particular group.

As interviewing progressed, some common themes began to emerge. These were captured as a tentative set of improvement ideas that might address the issues that had emerged.

2.2 Documentation Suppliers

The suppliers or writers of documentation are in various functional areas that can be organized into the following groups:

1. Administration – This includes all the people responsible for the day to day operation of the grid services whether or not they are part of the Grid Operations Center (GOC). These include the GOC, Information Services, Interoperability, Storage, Resource Selection, Support Centers, Monitoring, and the Integration Testbed.
2. Functional Groups – These include Security, Education, Engagement, Metrics & Reporting, Virtual Organizations (VO), and the VO User Group.
3. Management – This includes the Executive Team, the OSG Council, and Project Management.

4. Software and Tools – These include software suppliers to OSG that provide:
 - a. Virtual Data Toolkit (VDT),
 - b. Accounting (Gratia)
 - c. OSG Information Management (OIM)
 - d. Information Service - CEMon, Berkeley Database Information Index (BDII), Generic Information Provider (GIP)
 - e. Resource Selection Service (ReSS)
 - f. Resource Service Validation (RSV)
 - g. Storage Element (SE) – dCache, Bestman
 - h. Compute Element (CE) – Grid Resource Allocation and Management (GRAM), GridFTP, Condor-G, Grid User Management Service (GUMS), Web Caching service (Squid)
 - i. Log Collection & Distribution – Syslog-ng
 - j. Monitoring – Monalisa
 - k. VO Management – VO Management Systems (VOMS) and VO Management Registration Service (VOMRS)
 - l. OSG User Software – VDT-Client, Condor-G, Globus client tools
 - m. Software distribution – Pacman
 - n. OSG Information - MyOSG
5. External Information – This includes External Publicity and OSG Communication and Outreach.

2.3 Documentation Repositories

To round out the physical parts of the system, the following is a list of the various repositories of OSG documentation, with a few comments on the content and the owner's name in parentheses:

1. OSG TWiki – The main documentation repository and presentation tool at <https://twiki.grid.iu.edu/bin/view/Main/WebHome>. At the beginning of this project, it contained the following Webs with the owner in parentheses:
 - a. Accounting – (Philippe Canal) Gratia
 - b. ArchivedDocumentation – (None)
 - c. Auditing – no longer used (None)
 - d. Blueprint – Meeting notes for blueprint, high level brainstorming (Ruth Pordes)
 - e. CampusGrids – Oriented toward new users (Sebastian Goasguen)
 - f. CommunitySupport – Originally registered volunteer community support but mailing list is all that remains (None)

- g. Council – Internal business of the governing body that is publicly available (Kent Blackburn & Paul Avery)
- h. Deployment – Now points to management, needs to be deleted (None)
- i. Documentation – Originally the entry point for OSG technical documentation. (Anne Heavey) Needs new owner, is out of date.
- j. EdgeServices – Service information on service not implemented (Abashek Rana)
- k. Education – Needs better alignment with the rest (does it have links to more detailed info) (Alina Bejan, also Ben Clifford, Hose Cabellero) BNL – America’s initiative
- l. Engagement – Engagement VO information (John McGee - Sebastien Goasguen, Chris Green -1/2 time, Mats Rynge)
- m. InformationServices – External project to provide Generic Information Provider – dynamic info on the sites like cpus, storage, etc (Burt Holtzman-CMS)
- n. Integration – VO level testing of new software releases. The integration team produces the release documentation (Rob Gardner, Suchandra Thapa)
- o. Interoperability – Interoperability with other grids but only EGEE one is alive. (Bert Holzman, Keith Chadwick – Teragrid gateway?)
- p. Main – GOC owns the TWiki (Kyle Gross)
- q. Management – Management meetings, etc. (Chander Sehgal & Ruth Pordes)
- r. MeasurementsAndMetrics – Internal working documents (Brian Bockelman)
- s. MonitoringInformation – Documentation on the information monitoring system. Also pointed to from opensciencegrid.org page. (Rob Quick)
- t. NetworkTechnicalGroup – No longer active. (Shawn McGee)
- u. OSGRA – OSG Registration Authority (Mine Altunay)
- v. OSGReports – Management reports (Chander Sehgal)
- w. Operations – Operations information including infrastructure, services, ticketing, and SOPs (Rob Quick)
- x. ReleaseDocumentation – Focused on Site Administrator (Rob G)
- y. ResourceSelection – Technical information for resource selection (Parag Mhashilkar – Fermi)
- z. Sandbox – for testing /name to access (None)
- aa. Security – Procedures, awareness, certificates, working groups (Mine Altunay)
- bb. SiteCoordination – Meeting notes, etc. (Rob G.)

- cc. SoftwareTools – Internal planning and notes of the Software Tools Group responsible for overseeing all software tools distributed by OSG (Alain Roy)
 - dd. Storage – Overview of OSG data storage with meeting notes and pointers to documents based on user role (Tanya Levshina - Fermi)
 - ee. SupportCenters – May not be needed, subsumed by operations, including support center meetings. (Doug Olsen - Berkeley)
 - ff. TWiki – Information on the TWiki itself.
 - gg. Troubleshooting – Group no longer active in this form. Historical, before Sept. 2008, information. (Shaowen Wang - UI)
 - hh. VirtualOrganizations – VO group goals, meetings, activities and membership information (Abhishek Rana)
2. GOC Website – <http://www.grid.iu.edu/> is home to the OSG Grid Operations Center and organizes links to their services, procedures, registration, and trouble ticketing. The secure OIM Website (<https://oim.grid.iu.edu/>) contains various views of the information in the OIM. There is also a new website, <http://myosg.grid.iu.edu>, providing a resource summary, real time resource status maps and tables, and GIP validation status. Operations also maintains a News and Announcements blog at <http://osggoc.blogspot.com> (Rob Quick)
 3. Measurements Website – <http://t2.unl.edu/gratia> contains tools to create graphs of data from the Gratia accounting information collected at each OSG site. It has many canned graphs that can be modified by the user. A more extensive set of graphs are available at <http://t2.unl.edu/gratia/xml> but there are no links to it from the main site. (Brian Bockelman)
 4. Integrated Digital Conference – OSG meeting management at <http://indico.fnal.gov/categoryDisplay.py?categId=77> used mostly by OSG groups having external meetings. It is used by management, education, and for major internal meetings. There are some links to OSG groups meeting records on the OSG TWiki. (Owner?)
 5. OSG Document Database – Part of the opensciencegrid.org website, this contains some portion of the official documents of OSG at <http://osg-docdb.opensciencegrid.org/>. Some documents are accessible by the public while most are only accessible by OSG members. (Marcia Teckenbrock)
 6. OSG Website - <http://www.opensciencegrid.org/> is both the public face of OSG and a selection of detailed OSG information. It includes research highlights from scientists doing research using OSG, technical detail about a particular component of OSG (e.g. GUMS), news, education, selected usage graphs, and documentation on the high level OSG policies. It also has links into the OSG TWiki. Finally, it hosts the OSG Document Database which is independent of the main website. (Marcia Teckenbrock)

7. OSG Engage VO Registration Site – Hosted at RENCI (<https://osg-engage.renci.org:8443/vomrs/Engage/vomrs?path=/RootNode&action=execute>), this site allows authorized users to register in the OSG Engage VO. (Owner?)
8. VDT Website – Home of Condor, this University of Wisconsin website (<http://vdt.cs.wisc.edu>) houses documentation on the OSG VDT and Condor. In addition, it includes links to documentation on various components provided by others and hosted at their respective sites. (Alain Roy)
9. E&T Website – Hosted at UC, the E&T Website only houses the online OSG Grid Tutorial (Alina Bejan)

2.4 The Complete OSG Information Environment

Combining the Documentation Users, Suppliers, and Repositories together captures the entire documentation system. After understanding the full set of users of this system, it is apparent that the word “documentation” doesn’t really capture what this system provides and that it is more than the “system”. A better word for “documentation” would be “information” since, in addition to documents, instructions, training, procedures, and policies, the OSG Information System also supplies real time information to some of the User Groups. For example, VO Managers are interested in the real time status of resources in the grid. Likewise, system administrators interrogate the OSG GOC Website for various status information. A better word for “system” would be “environment” since it is much more than the technical system that houses the information. Therefore, we will label the overall system The OSG Information Environment. The diagram in Figure 1 shows the major interrelationships in this Information Environment. The size of each node represents the relative size of that node in terms of people for Suppliers and Users or amount of information it contains for technical Systems.

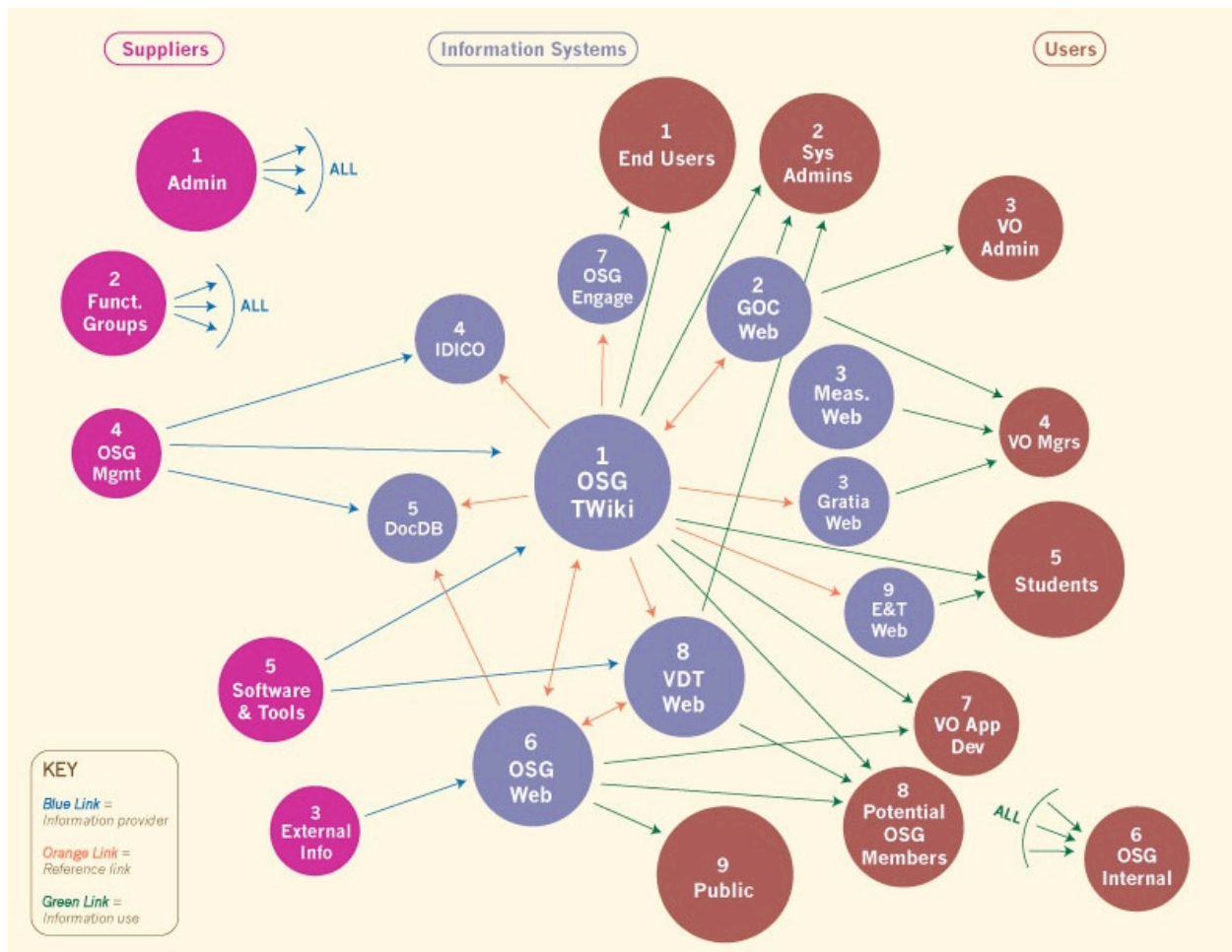


Figure 1 – The OSG Information Environment

4.0 Information User Perspectives

In analyzing an information system, the most important group to understand is the User because the system is there to serve those User's needs. Eighteen of the interviewee's were primarily users of the OSG information system.

In addition, existing studies of Grid systems were reviewed for applicable results. These studies were *The Report from the TeraGrid Evaluation Study, Part 1: Project Findings*ⁱ (Evaluation Study) and *Perspectives on Distributed Computing: Thirty People, Four User Types, and the Distributed Computing User Experience*ⁱⁱ (ANL Study).

4.1 Existing Studies

Although the primary target system of each of the two studies above was TeraGrid, both had a few findings that fit well with OSG. In the TeraGrid Evaluation Study, the use of project's wiki was analyzed. Interestingly, some interviewees expressed a desire for more centralized sources of information to be available since their wiki was used primarily to record team meeting notes, presentations, reports, procedures, and policies. This differs

from OSG's TWiki which is used as one of major central information repositories. Another complaint of the TeraGrid wiki was inconsistent recording of important working team meeting notes and decisions. The OSG TWiki may be used more consistently in this regard. Interviewees also reported challenges in finding documentation and information on the TeraGrid web site, which is again consistent with our user interviews. Since the Evaluation Study was a large project with a broad focus, they did a numerical summary of the number of times various barriers to use of the TeraGrid were mentioned and also recorded the areas of improvement that respondents believed would make the TeraGrid more useful. Documentation, support, and training was second (of 14) barriers and second in areas for improvement. For improvements, it was named within 1% of the top item and more than twice as often as all but one of the remaining 12.

The ANL Study was even broader and included users of both TeraGrid and OSG. Although respondents were sometimes evaluating documentation provided by their project staff rather than OSG, there were relevant findings that include:

- A1. Difficulty in locating and understanding documentation for Gridftp.
- A2. More engineering information is needed. (E.g. how big should a machine be to host a component, how big should the disks be, what should the network connectivity look like, and what security infrastructure, or other software, will be needed to support the service? Deployment and configuration details are of equal importance, particularly the host system technical requirements and scaling limitations.)
- A3. There is a lack of documentation about errors. (E.g. GRAM produces error codes that may require digging into the source code to determine the conditions that caused the error.)
- A4. There is a lack of troubleshooting documentation with better descriptions of what to do when things go wrong.

Not all of these results are directly applicable to OSG documentation, but notice the themes are the same as those from users of OSG documentation that are presented in the next section. Also note that documentation was consistently the second most critical issue in making Grid technologies more useful. The result of documentation deficiencies is loss of productivity of everyone using that documentation; so the costs, although hidden, are significant.

4.2 OSG Documentation User Interviews

A broad summary of these interviews would be that all the interviewee's had an overall positive view with many things that they liked about the documentation. This indicated to the author that the current system is doing a reasonable job in satisfying their needs and that incremental improvement of the system, rather than dramatic change, is called for.

Nevertheless, all the interviewees also had some issues with the documentation that, if it were improved in those areas, would increase their satisfaction and productivity. Specific recurring issues became themes that were categorized in six areas for improvement. This list identifies the important issues to be addressed in those six areas.

Information Organization:

- U1. Information needs to be better organized by user role: Almost all users believed the information presented tried to cover all users at once; and as a result, was not ideal for any user. Even the documentation directed at system administrators (for which the most documentation is provided) is among a clutter of information not directed toward them.
- U2. There is a lack of information at a broad architectural level: Although there are diagrams of the organization and interrelationship of the some of technical systems within OSG, there seems to be a lack of a set of diagrams of the overall set of system components and their relationships. These are needed to help people understand the broad picture and to then drill down to a subsystem a get more detailed architectural understanding. The diagrams that do exist are quite helpful and were commended by users. This deficiency causes a haphazard and slow learning process punctuated by mistakes caused by the lack of understanding.
- U3. The information needs to show context: Users get lost following links in the TWiki and have difficulty maintaining context. OSG information is not encyclopedic, but is very interrelated and the relationship is important to maintain.
- U4. Information varies from too detailed to not detailed enough: For example, procedural documentation needs to have the technical steps with higher level descriptions and decision help via links so the user is not lost in the shear volume of information in front of them at any one time.
- U5. Conflicting information exists depending on what entry point you take: This is the result of different views (the OSG EDU vs. TWiki) and different versions (osg-sites mail messages vs. TWiki documents).

Information Access/Searching:

- U6. Some documentation pages are very useful but have few links to them: For example, with so many acronyms, the Glossary of Terms would be very useful but appears only to be linked to from the main OSG website.
- U7. The TWiki Search box don't work at all as one would expect: For example search for "gratia installation" and it will find nothing on gratia installation. Using the Google search box on the TWiki home page also does not find anything on gratia installation and only finds things outside the TWiki.
- U8. The search box on the OSG website does not search the TWiki: Given that much of the information pointed to by the OSG website is in the TWiki, one would expect to be able to find it via the website search box. The user must understand the organization of the various repositories to do a search.

Decision Making:

- U9. The installation procedures require decisions for which the user does not have information required to make the decision: For example, many decisions are

based on the size of the site that is being installed but there aren't guidelines what is considered a small, medium, or large site.

- U10. There needs to be more suggested or recommended decisions: This may be caused by the fact that OSG wants to be neutral where there are multiple products that do similar things, but for many administrators of small VO's, an expert selection/decision that will work well for a small site is preferable to neutrality. For example Cert authority should have a default (pick either VDT or OSG).

Security:

- U11. This was one of the most often mentioned areas of difficulty. The security environment needs to be explained better, both for execution and data. Security for pool accounts needs explanation. The CA certificate chain and getting certs in the right place is difficult. Any way to check ownership, filenames, and access permissions would be useful.

Consistency:

- U12. There is inconsistent formatting and standards used by the documentation: This ranges from the use of variable names, to coloring, to font, to what is in boxes, etc.

Specific to Install Procedures:

- U13. Need a checklist of items (or a link to the checklist) to make sure background system setup is correct: This includes directory structures, correct system time, batch system installation, navigating to the correct directory, etc.¹ Packman doesn't have a default directory so it installs to the current directory. You better be in the right place and you get no warning about this strange behavior.
- U14. Things that need to be done for only the first install should be a link from an install document: Most uses of the document will not require the first install information so it should not be in-line in the install document.
- U15. CA distribution service should be mentioned in the ReleaseDocumentation TWiki:
- U16. The security certifications are common to any install: They should be written once and included by all separate installs where ever possible. This would also make it easier to skip over if the server already has services that require certs.
- U17. Each install TWiki page should list versions of the software it is applicable to: This is especially important if someone uses Google to find the page (and most do!). Often the procedure is written for the last major release and the user is installing some update whose version is beyond that.
- U18. Bestman installation appears to be a problem: The Bestman server uses GUMS for user authentication, but this link gets broken on an update. There are ~100 variables that are set in some scripts but the script doesn't use current

¹ See <http://hepuser.ucsd.edu/twiki2/bin/view/UCSDTier2/OSGCEWorkshopInstallNotes> as an example.

configuration information for an update. Even worse, in one instance the current config file was removed/replaced with a default file.

- U19. Troubleshooting is seldom covered: Only one piece of software has a troubleshooting guide. Most documents are a series of steps with expected results but no information to help if those results don't occur. Even common errors are not covered.
- U20. It is not clear on install command execution what the result should be: Sometimes silent is OK, sometimes you get a list of error messages that can be ignored, sometimes they can't be ignored. Some unexpected conditions cause a retry followed by stack traces.
- U21. It should be clear that you should install your own batch system: If you don't have a batch system, install condor first since you don't want it to be installed by OSG scripts.
- U22. Readme file of ~2000 lines contains recommendations that are not clear: Some of the recommendations are automatically done by the installation, some you have to do, and some you shouldn't do (prima?).

5.0 Information Provider Perspectives

The owners of major documentation systems (OSG website, TWiki, OSG DocDB, Education, Engagement) were interviewed. Several information providers and other OSG experts are part of the Improvement Team, and eleven interviews were completed in this group.

The consistent message from this group is that they believed they had the necessary authority to make improvements in what they provide. All of them also had some ideas of areas under their control that could be improved. The main impediments to production of better documentation were the priority of documentation, availability of time, lack of strong incentives (the job isn't fun); and, in some cases, unclear ownership.

Specific issues and suggestions mentioned by various information providers are:

- P1. Clear ownership is not established – providers often named things they contribute, but there was not a strong sense of responsibility for documentation. Contrast this with the feeling you would get in talking to someone who had written a piece of software or a paper that is to be published.
- P2. Clear, written standards for things like Installation Guides would be useful. One interviewee said she tends to copy CE documents and use them as a standard.
- P3. They would like to see a template for each type of document that followed the standard and had the TWiki code to implement it. None of the documenters liked to figure out the TWiki format to get something to look like they wanted (not that it was hard, it just takes time compared to WYSIWYG editors). Things like tables (with edit buttons) would be used more if they could simply copy one from a template.

- P4. They would like to get direct input from the users when there are issues of clarity, confusion, errors, or when they have suggestions.

6.0 Understanding Wikis

Given the centrality and importance of the OSG TWiki as a socio-technical system used by OSG for documentation, it is also important to have a base understanding of the characteristics of this relatively new technology and its impact on the providers of documentation. Ann Zimmerman of the School of Information at the University of Michigan (one of the authors of a previously cited study) has followed studies, some of which were done at UM, on what make wikis work. Following up after her interview, she identified four studies that were likely to be the most relevant to OSG. These sociological studies of the characteristics of successful wikis were reviewed to glean an understanding of the use of wikis and the incentives that make them successful. For our purposes, they will be named in a brief way that helps understand topic and context of the study. They were:

- Wikis for IT Documentationⁱⁱⁱ
- Member Motivation for Contribution^{iv}
- Viable Wikis^v
- Policy to Mediate Participation^{vi}

Wikis for IT Documentation describes Rice University's experience in using wiki technology to enable IT groups to provide technical documentation to users to solve problems of stale documentation, orphaned ownership of content, and increased demand for documentation from users. Their wiki has similarities with OSG's TWiki, containing directories, calendars, security information, server set up, technical procedures, tutorials, and minutes from meetings and events. They validated the importance of some of the structure and techniques that OSG already uses but several of their observations speak directly to issues OSG users report:

- L1. Accommodate the line of thinking of the users and offer information in chunks that satisfy this line of thinking. Try not to offer the information in chunks that correspond to our organizational chart or to our skill sets.
- L2. Make instructions around tasks with options for environment rather than around environments with the same task repeated for each.
- L3. Put attachments (presentations, spreadsheets, video training, etc.) with pages they are associated with rather than in a separate repository.
- L4. Student workers became some of the best contributors and editors because they understood IT and the customers of IT (since they were customers).
- L5. At least one or two people need to assume the role of "gardener" of each area in the wiki.
- L6. Management must take the position of, "If its not in the wiki, it doesn't exist."
- L7. Put directions in the wiki and reserve training courses for underlying concepts.

Member Motivation for Contribution studied the motivations that drove contributions to virtual community generated content systems. They identified fourteen motivational

categories that have been documented for these communities and identified the ones that studies have shown are a factor in wikis. The nine motivational factors in wiki contribution are: altruism, belonging, collaboration, egotism, knowledge, power, reciprocity, reputation, and self-esteem. The strongest motivation in their study of student use of a wiki was collaboration with evidence of egotism, reputation, power, and reciprocity. For OSG purposes, any structure, process, or technique that would enhance any of the fourteen motivations could improve the level of contribution; and therefore, the quality of the TWiki documentation.

Viable Wikis studied open access (public) wikis to distill out success factors in terms of policies, norms, user incentives, as well as technical, and structural features. Since this domain is quite different from the OSG TWiki instance, there are only a few of the factors that they identify that have application to OSG. Those factors are that:

- L8. the perceived status of authors/contributors is a function of their contribution,
- L9. the quality of any page is directly related to the number of contributors,
- L10. the quality is also related to the level of quality set and allowed by the owner/primary editor of the page, and
- L11. the quality is improved by the number of contributions by “novice” users rather than expert/registered users. These conclusions strongly support the need strong owners for pages and for the users of OSG documentation to have an easy way to contribute.

Policy to Mediate Contribution was a study of role of policy to mediate conflict during extended contributions to articles in Wikipedia. Useful guidelines mentioned in this study include:

- L12. “Be Bold” meaning if you notice an error, take action and fix it.
- L13. that well known (i.e. written) policies on contributions are critical to producing quality information. For the OSG TWiki, there are a set of policies that one might derive by study, but they aren’t well known or written down.

7.0 Documentation System Improvements

This section presents a synthesis of a set of key changes that need to be made to address a majority of the issues identified. Considering the large amount of information and ideas collected and summarized in the previous four sections, the challenge is to identify actions that will produce the most improvement, in the eyes of the users of OSG documentation, using a reasonable level of staff.

Documentation is almost exclusively produced in OSG by people who do it as a part of their normal work. This is a reasonable model, especially given the technical nature of the documentation and the relatively high technical expertise of the users of that documentation. One the other hand, this also means that no one has the documentation system as the primary focus of their work. This leads to a sequence of documentation improvement projects that produce some good results which then decay until there is

another effort to focus some effort on the system. The goals of the current project includes producing structure, framework, and processes that will lead to ongoing improvements and maintenance of documentation. The following changes or improvements are proposed to address the findings of this study in a way that improves ongoing maintenance. The bracketed numbers following each proposal are a list of the issues from the previous sections that are partially, or fully, addressed by that proposal. Also see Appendix 1 for a table that relates each of these changes to the project goals and issues identified in the study.

1. Establish true ownership and policies

- 1.1. Writers should feel the documentation they do belongs to them and reflects on them just like they would feel about a program they write. (Actually, much OSG documentation, like installation procedures, are programs executed by humans.) There should be a clear owner of each page (or set of pages) who is responsible for evolving the page as feedback is received or new features or releases are provided. The TWiki should show who owns the page rather than who happened to change it last so the ownership is clear to everyone. Pride is a strong motivator but it is hard to have any pride on your work if others don't know you did it. See Recommendation 4.2 for a way to reward good documentation. [G8, L5, L10]
- 1.2. Name an owner for the documentation architecture whose job it is to evolve the architecture resulting from this project over time. This owner would participate in deciding where new documentation should be placed and help to resolve differences in documentation that one sees depending upon the entry point into the system. This architect would also manage the organization and guidelines in recommendation 2. [G1, G2, G7, G8, G9, L5]
- 1.3. Over time, the Documentation Architect should identify and resolve differences among the Education, TWiki, and current best knowledge among the mailing lists. Start by linking to common security information from Education, Installation, and Operations. In cases where there are multiple approaches, links should be provided to the different views so a reader can evaluate the answers for his/her situation. [G1, U5, U6, L10]
- 1.4. Write down Documentation Policies and Procedures. Many can be derived from this list of improvements. [G4, G10, L12, L13]

2. Provide some guidelines for organizing, presenting, and searching for information, including:

- 2.1. Provide a view by role (with only one role on the page). Roles that should be included are Grid User, System Administrator, VO Administration, Student, and Joining OSG. [G2, G5, U1, L1]
- 2.2. Present the overall architecture of OSG technology that is now missing. Each major subsystem should have a link that presents more detail that includes the main components and their major interfaces. [G1, U2, U9]
- 2.3. Provide good engineering information for decision making, with suggested decisions for likely cases. [G1, A2, U9, U10]

- 2.4. Over time, add more error documentation and troubleshooting information [A3, A4, U19]
- 2.5. In procedure documents, put mainline flow inline, and provide reference to separate documents for preparation required, checklists (add them), decision information, and first time only steps. Include the software versions supported by each procedure. [G1, G4, A2, U4, U9, U10, U12, U13, U14, U16, U17, U20, U21, P3, L2, L3, L10]
- 2.6. Establish some standards and TWiki templates for procedures. Add a page that lists those standards. This is unlikely to be read by users of the documentation, but its existence will be recognized because the users will quickly see the consistency that results. [G7, U12, P1, P2, P3, L10]
- 2.7. Make search boxes work as one would expect, explain what they search, or remove them. [A1, U7, U8]
3. Create a review process that provides better personal incentives.
 - 3.1. How about trying a concept similar to Pair Programming? Think about what makes writing software so enjoyable. It is continuous creation and problem solving with nearly instant satisfaction when it works or a new problem to solve when it doesn't. It's not likely that we could make writing documentation as fun as writing programs, but maybe we can improve the satisfaction while spending little additional resource overall. The approach would be to assign a writer and reviewer (corresponding to the driver and observer) to do the documentation page together with the reviewer. With collaborative technology, the two people would not have to be in the same place. Maybe it would be difficult to change roles, as is done in Pair Programming; but the quick feedback and interaction would likely produce a much higher quality product as it does with Pair Programming. If the observers were selected from the group who uses the documentation, the flow, level of detail, clarity, and information for decision making might be dramatically improved. [G6, L8, L9, L11]
 - 3.2. Establish a standard review process for each major subset of information in the TWiki (e.g. for each section in the current Documentation page). This process should be modeled after the process being used by the ReleaseDocumentation web where the status of each document can be recorded by the reviewer. This would get most of the documented areas up to the higher standard set by the release documentation. [G9, U12]
 - 3.3. Use an annotation plugin that allows anyone, without any approval or signup, to leave a comment on the document itself. [G3, P4, L4, L9, L11, L12]
 - 3.4. How about providing a standard evaluation where any reader could click a radio button at the end of each document to rate it. These ratings could be numerically tallied to give a score for the document. You could have an award ceremony for the best 3 scores and the 3 most improved scores each year. This would add some friendly competition to make it more fun. [P4, L8]
4. Remaining issues

- 4.1. Some issues [U3, U11, U16, U20, U22, L6, L7] are not fully addressed by the above proposals. Some of these are specific items [U11, U15, U18] that need to be addressed in one document and can be addressed by the document owner. Others are from the literature [L6, L7] and are already being done to some extent by OSG. Two [U3, U22] should be reviewed by an expert to determine the possible improvement.

8.0 Migration Project Plan

To be developed.

Appendix 1 – Coverage Table

The following table shows the Goals or Issues identified in the project in the left column. The letter-number identifiers refer to the like numbered items in Sections 4, 5, and 6. The column headings are the numbers of recommended changes/actions (detailed in Section 7) developed to address each of the Goals and Issues. Cells are populated when the column action partially or fully addresses the Goal or Issue. The last column covers areas not addressed by the actions with recommendations on the disposition.

	Actions															Other	
Goals/ Issues	1.1	1.2	1.3	1.4	2.1	2.2	2.3	2.4	2.5	2.6	2.7	3.1	3.2	3.3	3.4	4.1	Reso- lution
G1		1.2	1.3			2.2	2.3		2.5								
G2		1.2			2.1												
G3														3.3			
G4				1.4					2.5								
G5					2.1												
G6												3.1					
G7		1.2								2.6							
G8	1.1	1.2															
G9		1.2											3.2				
G10				1.4													
A1											2.7						
A2							2.3		2.5								
A3								2.4									
A4								2.4									
U1					2.1												
U2						2.2											
U3																4.1	Expert
U4									2.5								
U5			1.3														
U6			1.3														
U7											2.7						
U8											2.7						
U9						2.2	2.3		2.5								
U10							2.3		2.5								
U11																4.1	Owner
U12									2.5	2.6			3.2				
U13									2.5								
U14									2.5								
U15																4.1	Owner
U16									2.5								
U17									2.5								
U18																4.1	Owner
U19								2.4									
U20									2.5								
U21									2.5								
U22																4.1	Expert
P1										2.6							
P2										2.6							

P3									2.5	2.6							
P4														3.3	3.4		
L1					2.1												
L2									2.5								
L3									2.5								
L4														3.3			
L5	1.1	1.2															
L6																4.1	Done
L7																4.1	Done
L8												3.1			3.4		
L9												3.1		3.3			
L10	1.1		1.3						2.5	2.6							
L11												3.1		3.3			
L12				1.4										3.3			
L13				1.4													
	1.1	1.2	1.3	1.4	2.1	2.2	2.3	2.4	2.5	2.6	2.7	3.1	3.2	3.3	3.4	4.1	
	Actions															Other	

Count 3 6 4 4 4 3 4 3 17 6 3 4 2 6 2 7

References

-
- ⁱ *Report from the TeraGrid Evaluation Study, Part 1: Project Findings*, Ann Zimmerman and Thomas A. Finholt, Collaboratory for Research on Electronic Work, School of Information, University of Michigan, www.crew.umich.edu, August 2008 (obtained from author).
- ⁱⁱ *Perspectives on Distributed Computing: Thirty People, Four User Types, and the Distributed Computing User Experience*, L. Childers, L. Liming, and I. Foster, Distributed Systems Laboratory Mathematics National Laboratory, Computation Institute, University of Chicago, September 30, 2008.
- ⁱⁱⁱ *(Wiki + ResTechs) = (Fresh documentation + Organic Knowledge Management + Training Materials + Good, Cheap Technical Writers)*, Niki Serakiotou, et. al., Rice University, SIGUCCS'08, October 19-22, 2008, Portland Oregon, USA, Copyright 2008 ACM 978-1-60558-074-6/08/10.
- ^{iv} *Understanding Member Motivation for Contributing to Different Types of Virtual Communities: A Proposed Framework*, Trevor D. Moore and Mark A Serva, Ph.D., University of Delaware, SIGMIS-CPR'07, April 19-21, 2007, St. Louis Missouri, USA, Copyright 2007 ACM 978-1-59593-641-7/07/0004.
- ^v *Viable Wikis Struggle for Life in the Wikisphere*, Camille Roth, Department of Sociology, University of Surrey, Guilford, UK, WikiSym'07, October 21-23, 2007, Montreal, Quebec, Canada, Copyright 2007 ACM 978-1-59593-861-9/07/0010.
- ^{vi} *Community, Consensus, Coercion, Control: CS*W or How Policy Mediates Mass Participation*, Travis Kriplean, Ivan Beschastrikh, David W. McDonald, Schott A Golder, GROUP'07, November 4-7, 2007, Sanibel Island, Florida, USA, Copyright 2007 ACM 978-1-59593-845-0/07/001.