

# **Specific Interview Comments on Documentation**

## **TABLE OF CONTENTS**

<b>GENERAL COMMENTS AND GENERAL INSTALLATION</b>	<b>2</b>
<b>LOCATING AND SEARCHING INFORMATION</b>	<b>8</b>
<b>DECISION MAKING</b>	<b>10</b>
<b>CE INSTALLATION</b>	<b>11</b>
<b>WORKER NODE</b>	<b>13</b>
<b>BESTMAN</b>	<b>14</b>
<b>VOMS/GUMS</b>	<b>15</b>
<b>GIP</b>	<b>16</b>
<b>TROUBLESHOOTING/PROBLEM RESOLUTION/HELP</b>	<b>17</b>
<b>DIRECTORIES (RSV PROBE)</b>	<b>19</b>
<b>SECURITY</b>	<b>20</b>
<b>EDUCATION - USER</b>	<b>24</b>
<b>WEBSITE</b>	<b>27</b>

## **Specific Interview Comments on Documentation**

### **General Comments and General Installation**

- Site planning
  - If changed, adding list for small, medium, or large site (incl things to start with and what would have to be added as you grow would be useful)
  - Performance for each service (what needs to be separate, what can run together if you plan to grow) (eg. VOMS on CE or not) Implied in diagrams but don't know if he should configure it for ultimate size or start with small diagram and modify it as site is grown. How about a table by size that itemizes servers and certificates required.
  - Useful to have list of what certificates for each of the services (eg. Service cert for RSV probes)
- Need views for
  - VO managers, sys admins, end users, computational scientists, PIs
  - Need to understand context of information (not an encyclopedia) – gets lost in trail of hyperlinks on TWiki so don't even understand where he started
- Low level technical stuff like what ports need to be open get overlooked – needs to be in documentation (also permissions). Install from root or not, etc.
- Can't get information on high level architecture of the system
  - Major components picture (can't install everything on one machine)
  - Small site vs large site (they wanted to set up 100 member labs with a few machines)
- No guidance on creating captive accounts (E.g. Is it a good idea to have a special account for condor, bestman, xrootd or can you use the general daemon account?)
- Question: Is documentation for own code stack or for the community? If for a community, it needs to be a portal.
  - OSG only is fine for big VO's but to branch out to other science disciplines, documentation needs to broaden to support the community

## Specific Interview Comments on Documentation

- Engage wants to make it inclusive as a science community. Duke has a campus cluster to possibly set up virtual OSG site. Nuclear physics people need to pull in to large amounts of data.
- Need Docs and procedure orientated toward small or very small resource pools
  - Most difficulty came in bringing up their resources (admin view for small and very small sites)
  - One of the issues with SBGrid is that they not follow documentation directly since they had a 3 node site (by following the documentation, they tried to load all the services and got into trouble)
- Might be good to have doc pages targeted at groups moving from the Engage VO to becoming their own VO, possibly with their own (small scale) resources.
- Consistently use bold or italic for filenames/directories allows quick scanning for key info (use a standard like manuals with 2 pages of standards)
- Can't be done with 2 month crash effort. A software project has a lead – who is the documentation lead? (Needs at least a ½ time person with this assignment.)
- Want quick install to be no more than 5 pages
  - - For example, the "Overview of OSG Configuration" has about a page worth of introduction text, including a description of the steps configure-osg.py takes (what it chmods, what it configures, etc). This is not necessary to know in order to install the CE.
  - - As another example from the configuration section, the "-v" parameter for configure-osg.py is explained 3 times.
  - - The section "Included topic: Gip Configuration" is mostly errata notes for PBS/Condor/GUMS publishing. We could just include a link saying "For advanced PBS, Condor, or GUMS publishing recommendations, click here"
  - - In the "Listing Services" section of the CE install, an example is given for output from vdt-control for a CE and the output from a VOMS/VOMRS node; the latter is not needed.

## Specific Interview Comments on Documentation

- - Configuring WS-GRAM authentication is covered twice in the document -- if we're repeating ourselves, I suspect the document is getting too long.
- No clear place to start if I am an At Large VO user that does not have a VO organization with instructions.
- Certificates are a big problem - If you accept a bad cert (Get a expired root cert) – doesn't fail until you try to import it.
  - Add a trusted root certificate then it works
- Check out webdav twiki plugin to allow editing by MS applications.
  - <http://twiki.org/cgi-bin/view/Plugins/WebDAVPlugin>
- Examples of commands to run that include path they used in installation rather than a variable path that would be where I am installing
- Inconsistencies in how each author writes same thing (\_\_\_VO\_\_\_) – copied exactly and substituted vo name. There was a space in the documentation that causes command to fail
  - Maybe use [installpath]
  - Documentation does not use the path variables consistently – VDT is much better than OSG documentation on this
  - Specific example in Bestman install on CE:
    - <VDT\_LOCATION> and later <CE\_VDT\_LOCATION>??
- Can't tell if a page is up to date, can only go by TWiki modification date – hard one is 8 months old, is it still up to date.
  - How about saying that it was written for a certain version
- Biggest issue – Is documentation applicable to the version he is working with. (compare with docs and readmes that come with new software)
- Aren't good checkpoints – many places that can go wrong
  - VDT location not correct

## Specific Interview Comments on Documentation

- Packman doesn't have default directory – goes to current directory! - This needs a warning in the documentation since it is non standard for installers
- Not clear on install command execution what the result should be.
  - Sometimes silent is ok (should not be default)
  - Sometimes get a list of error messages that can be ignored
  - When you have an unexpected condition that does a retry that then spits out stack traces, etc.
- Can use links to explain decision, etc. and shorten doc for experienced people
- An author and editor have not sat down together and reviewed the install document
- He writes his own CE install doc for workshops he has taught
  - Starts with checklist of silly things that will trip you up.
- Site specific info is where he has the most trouble
- Want most important things in doc with depth
  - E.g. GIP section with batch system info
  - WSGram section tells you how to configure and do performance tweaks (Performance should be a link)
- An ideal document for installation would have more interactivity and feedback, e.g. don't go on unless last step was successful
- Direct you through given the choices you make (tunnel like)
- Rochester's experience is that TWiki is too hard to use – so no one is using it for students
  - They prefer DocuWiki, Drupal (supports public and private comments) and MediaWiki
  - Using Wysiwig on Drupal so it is very easy for people to update
- Need high level view of resources, #, where, etc. Need plots and maps
- Most improvements in OSG documentation have been in the system admin area (provider)

## Specific Interview Comments on Documentation

- From user side, there is less improvement
- Some pages present info in a low level way rather than have multiple pages that cover the high level also
- Conflicting info (can get 4 different answers)
  - Core info on TWiki
  - OSG EDU team (independent of TWiki) – Globus centric and not best way for OSG people
  - OSG mailing info (osg-sites, etc.)
  - GOC info
  - (Has examples of different advice from TWiki, EDU, and osg-sites mailing)
- Is there a way to automatically process all the pages looking for common problems – like hardcoded urls rather than wikiword references?
- Need to have a way to locate all pages whether they are referenced or not. For example:  
<https://twiki.grid.iu.edu/bin/view/ReleaseDocumentation/StorageElementAdmins> - can be found using a search (OSG\_DATA) of the TWiki but appears to be off a depreciated Documentation Table page and no longer needed/valid?
- Navigation – Hard to get to Engage now from website
- Difficult to write
- Very difficult to keep it up to date
- Need requirements, standards, and templates
- Getting user feedback is very useful
- User can put notes on what they find wrong – much preferable to emails, etc.
- Email capture (automatically on a TWiki page)
  - Drawback – a special email address to monitor – a cron job scans incoming email to a particular address every 15 minutes and uses the subject to select the email. You have to send the first email to a special address and the subject must be unchanged.
  - Supports TWiki markup also although plain text is usual
  - Requires a particular email address for each TWiki page.

## **Specific Interview Comments on Documentation**

- Each thread becomes a separate TWiki page.
- Project page
  - Proj Mgmt
    - Email records
      - Requirements review
      - Testing participation
- Difficult to edit status in TWiki table
- There is a WYSIWYG editor but it is disabled in OSG
  - Difficult to use because it doesn't keep up and you may have to refresh to see where you are at
- Concerned about the ongoing support and evolution of TWiki

## **Specific Interview Comments on Documentation**

### **Locating and Searching Information**

- Go to ReleaseDocument – everything is there
- Don't google for documents because you may get out-of-date document – stay in release documentation
- Have had mixed results in built in search – Sometimes uses jump box which seldom yields a result (since he didn't know the TWiki word)
- Sometimes get multiple versions when searching, dangerous if out of date version selected.
- It is poorly organized for someone with the role of system administrator
- Is not always easy to locate what you need to know
- Ordering of main TWiki page and other pages not ordered by importance and includes something for everyone
- For example, on doc hub, there are 4 links/wikis to packman
- Documentation should flow through process exactly what a new administrator (or any administrator) needs to know/do instead of being overwhelmed (e.g. maybe 6 major topics)
- Need a page focused on each type of user (by role)
- Needs to be specific to the topic of interest (installing a CE, for example)
- Easier to remember only OSG website and always start from there (TWiki address is complex.)
- I come to the TWiki and click on links until I find it. This sometimes takes several minutes because there is so much documentation and he can't find where the specific info is located.
- He never tries to use the search box in the TWiki because he heard it was "brain dead" (and I have to agree)
- He sometimes uses a Google search
  - In the CERN TWiki, google now cannot index info in CERN wiki and he heard at CHEP that this is a real problem now because people like to just google to find things
- It sometimes takes a few minutes for him to find what he is looking for because there is not a good way to search for it so



## **Specific Interview Comments on Documentation**

he has to click his way to the right document. It is not always clear where the specific [installation] information might be.

- Usually starts with OSG Website and uses left column to navigate
- Can't figure out how it works
  - Example: Wanted to find the Council org chart, didn't know where to look. Hard to find – Agreements, management documents, review documents, annual report
  - Note: I showed him the google search box on the twiki pages that makes these searches pretty quick.
- It is not easy to navigate because there aren't pages that are oriented around what I am looking for.
  - Need simple map
  - Get all kinds of detailed info but can't get broad level info

## **Specific Interview Comments on Documentation**

### **Decision Making**

- Choices during installation – no clear explanations why to do what?
- He has his own wiki on the side to do this
- Mostly did new installation like the one they had, so it wasn't a major issue
- Did change a couple of things to follow recommendations in documentation that differed from their current install
- Documentation was pretty good in this area and pretty consistent (same advice in other steps)
- Once you had to skip ahead 2 steps ahead (install creates a file that you later modify – if in other order, file would be correct)
- In one case, open up n ports for small site (small undefined, they did 100 ports, then discovered 16 jobs required 120 ports)
- Security/authentication (prima based or grid map) – documentation needs to provide guidance to decisions
- Tells you how to install but not what it is and parameters to decide (what is a small site)
- Install own batch system, if you don't have it, install condor first – should tell you to do it separately since you don't want it installed by OSG scripts
- NFS, CE file system organization, VDT locations, OSG app, OSG Data, User home, Batch system directory are example decisions that needs to be made with knowledge of OSG demands on system and its implications in the organization of the components.
- Data staging and where it is and how much to expect. Put on different file system or on a different disk because of contention

## Specific Interview Comments on Documentation

### CE Installation

- Difficult to get through (because of complexity)
- Configuration is complex – need to make decisions that you don't understand – e.g. need to know if you should install services even if you are not using them.
  - Use PBS (portable batch system)
  - Maybe setup condor even if not to be used?
- Installation of job managers (do full Condor install?)
- Don't understand implications of install options (especially given growth plans)
- Not told what directory permissions should be osg\_app, osg\_data, and vo home groups? Need to have list of directories and permissions (Arvind has mail on osg-sites with the current understanding – copied below)
- Configure condor settings – a problem he found caused a change in documentation – (a broad a statement on changes in condor config files but they weren't listed)
- Installed CE 3 times because didn't have things set in right order – Need firewall doc referenced in the before starting section
- Need to completely read Before Starting completely
- Has prepared his own notes page with his local install info – so now he can now install CE it in an hour
- Full Priv Pre Config – excellent
- Configure OSGAttributes is fine
- Exporting environment variables, etc. are in preparing for installation but you need them for each installation – now it is there as notes
- Isolate tech parts in 1 doc and put the higher level somewhere else (with a link to higher level if someone needs context)
- Core of doc should step you through every step and element of the install, not leaving anything out
- Cert infrastructure is missing – just a brief mention at end and it misses a step (If you execute the CE Install exactly as documented, you will not end up with a working CE!)

## Specific Interview Comments on Documentation

- Why does a site admin have to decide which source of Cert Authorities – the VDT or OSG? Should be defaulted to one.
- Decisions need to be called out with info on how to make it. (Many can be made by default – like the Cert source.)
- Sometimes stuff taken out leaves doc not even flowing because of the number of editors – needs overall editor
- Likes green boxes on command lists
- Behind every good author is a good editor
- Some documents may have been over edited and may even have become a whole group of other TWiki pages
- Current CE install doc is not even current for 1.0 VDT
- Post CE install – should not be separate (Was the author worried that it was too long!!) You don't have to go to another book to get the last chapter!
- Experienced administrator needs on 10 pages of info for CE install with 25 not needed
- Sometimes get significant comments buried
  - Critical steps that would messed up
- The document is too long
  - - For example, the "Overview of OSG Configuration" has about a page worth of introduction text, including a description of the steps configure-osg.py takes (what it chmods, what it configures, etc). This is not necessary to know in order to install the CE.
  - - As another example from the configuration section, the "-v" parameter for configure-osg.py is explained 3 times.
  - - The section "Included topic: Gip Configuration" is mostly errata notes for PBS/Condor/GUMS publishing. We could just include a link saying "For advanced PBS, Condor, or GUMS publishing recommendations, click here"
  - - In the "Listing Services" section of the CE install, an example is given for output from vdt-control for a CE and the output from a VOMS/VOMRS node; the latter is not needed.

## **Specific Interview Comments on Documentation**

- - Configuring WS-GRAM authentication is covered twice in the document -- if we're repeating ourselves, I suspect the document is getting too long.

## **Worker Node**

- Need better documentation to get up on different machines (Windows, Mac, XP, Vista)
- Worker node client was good – likes option description with how to for each of the options
- Similar for other documents – Worker node Install - CA/CRL section will not work

## **Specific Interview Comments on Documentation**

### **BeStMan**

- Why Proposed Architecture rather than Architecture?
- Picture should show Gratia gridftp
- Configuration Decisions
- Nodes – give some info on server to load relationship or server to storage
- GUMS – give an OSG recommendation on GUMS – does it depend on the size of the site, the # VO's allowed to use it or what?
- Gratia gridftp-transfer – need to encourage installation and mention key benefits of installing probes
- Need a question explaining use (or not) of static space tokens
- BeStMan-gateway access to FS (why or why not)
- Downloading BeStMan-gateway
  - Mkgridmap – Is this if you are using GUMS, if so say so.
- Configuring GridFTP
- Start the instructions in a new paragraph from the “ignore the rest of this section.”
- Name of path on bestman gateway – got info from another admin that he copied and got to work – definition/explanation of the parameters were not clear

## Specific Interview Comments on Documentation

### VOMS/GUMS

- Why do you have to change the host file? (When she did this on the CMS Zen server, it broke it and had to be rebooted.)
- When she did "vdt-control -list", VOMS didn't show up
- The Install CA files info is out-of-date since the VDT no longer installs the certs. The question shouldn't even be asked (in the green box). Later the correction has been made and it says the certs aren't installed by the VDT
- The install script asks about setting up a cron job to update certificates. Then later, there is a section to set up a cron job to do the update. Why doesn't the install do it or is the second one there to allow you to change your mind. If so, it should be stated.
- VOMS web UI: the install no longer populates the database with a sample VDT VO, yet the UI requires at least one VO to function, so you have to add a sample VO anyway.
- No issue with VOMS install – no issue except a mail server (his error)
- GUMS also worked without a problem – good documents
- 1) GUMS configuration asks users to set VDT\_GUMS\_HOST environment variable before the pacman install. Unfortunately, it mentions this environment variable about 5 pages after it instructing the admin to perform the install
- 2) There's a note saying configure-osg.sh is available and should work. This is false.
- 3) The section titled "Steps to configure a simple site from scratch" does not leave you with a working configuration as you don't set a Subcluster or a SE
- 4) The WS-GRAM section needs to be cleaned
  - - Configuring authentication is covered. It links to pages that start with "This page is obsolete". As I mentioned above, this appears to mostly be a duplicate of the previous section on WS-GRAM auth
  - - Sudoers configuration is covered twice; at the beginning and the end of the document
  - - The following note (in bright red) is in the release documentation: "JeffPorter (Jan 2009) With VDT update XXX, these performance enhancements are

## **Specific Interview Comments on Documentation**

automatically enabled where possible. The one not managed is locating the persistent directory on local disk. The release documentation should be updated accordingly - when the release is ready." The performance enhancement section appears to be unnecessary?

- - There's troubleshooting information in the WS-GRAM section; this should be linked from the install document, but not contained in it.
- 5) The end of the document includes Reviewer information from 2007. We probably should just remove this.

## **GIP**

- GIP Config – Good except GIP Configuration File example was large and complex, but you only need a few things (need documentation about this since most must be deleted)



## Specific Interview Comments on Documentation

### Troubleshooting/Problem Resolution/Help

- OSG mailing list has helped resolve questions (osg-sites)
  - Sometimes get 5 different answers (e.g. permissions) which shows the differences in various sites (and maybe some confusion)
- Get help is poorly organized – it is help for a wide range of doc users, not directed at sys admins or anyone else (random collection)
- Lot of documentation on a-z steps, there is not much on when it is broken – common errors
  - Every time I update Bestman, it breaks
    - I get a different error each time for the same problem
    - Bestman server uses GUMS for user authentication, but it breaks this link when Bestman is updated
    - Missing some flags in install
    - 100 variables that are set in some scripts but it doesn't use how it is currently configured on an update. Other packages seem to do this better. Once it blew away current config file and replaced it with a default file!
  - One piece of software has a troubleshooting guide
    - Series of steps with results, but if the results don't occur, there was no info to help (like it assumed it all worked)
  - Even common errors are not covered
- I do remember having an issue with was finding the correspondence between the numerical status codes in the XML output of condor\_q and the more descriptive letter codes.
  - The information was not in the Condor Manual
  - He eventually found an email or news group posting that had a list of status codes, but he wasn't sure what release they applied to or if they were "official"

## **Specific Interview Comments on Documentation**

- It would be nice to have a button that would connect you to someone when a step failed
- When he had problems, he emailed GOC – not effective – was forwarded to developer, but they didn't respond. Now he use OSG storage email group and they help a lot because they were actively installing SEs
- Using ticket system is hard because it is a series of emails rather than editing/updating directly on the web

## Specific Interview Comments on Documentation

### Directories (RSV Probe)

So it seems to be the consensus so far is:

- [Required] OSG\_LOCATION
  - - must have 755 permissions,
  - - ownership is important -- root owned
  - - Ignore Sticky/UID bit
- [Required] OSG\_APP
  - - must have 777 permissions,
  - - ownership is important - root owned
  - - Ignore Sticky/UID bit
- - Based on what Iwona said, it sounds to me it's not possible to check
- for one or more permissions in a meaningful way ... and that we may be
- better off trying to write (and deleting) a file on this space. Thoughts?
- [Recommended] OSG\_DATA
  - - must have 777 permissions,
  - - ownership is important -- root owned
  - - Ignore Sticky/UID bit
  - - Does above comment about writing file to test instead of checking for
- permissions make sense in this context?
- [Required] OSG\_GRID
  - - must have 755 permissions,
  - - Ignore ownership
  - - Ignore Sticky/UID bit

## Specific Interview Comments on Documentation

### Security

- For me the most significant observation surrounds the movement to a new CERT infrastructure model without clearly outlining this in the documentation as well as the more general issue of more clear documentation on certificates and host certificates in general.
  - For the specific case of the new automated CA system the change to disable the CA infrastructure by default leaves me a bit confused. Why for example is not one option or the other a default? Is it merely due to indecision on whether or not VDT or OSG should be the official CERT? Part of the solution to the documentation problem to me would be to make a decision, probably in favor of the OSG, after all we are installing pacman -get OSG.
- The second is that the documentation needs a section on certificates and the certificate infrastructure. When I last installed a CE (about 3 weeks ago) the only mention of the need to configure certificates is in a blurb at the end of the installation script. Then when you try to start the infrastructure you get an error saying you need to run another script. This seems entirely backwards.
  - In the documentation after the install CE step there should be a clear section, not in a preparation for install document, but in the installation document outlining the steps to get the certificate infrastructure configured.
  - That is the main thing that stuck out for me though from my last install. What I can do is go through a CE install here, I have not had osg-gw-3.t2.ucsd.edu up for a while, and make detailed notes on any issues or problems in the documentation. That way things are fresh.
- GIP info may not work – site & VO specific info is a problem
- Have had problems with CRLs (making sure workers got CRLs up to date) – Cert source, versioning is confusing
- OSG no longer packages CRLs – need to understand ramifications of cert source choice
- GUMS install – current production GUMS server is old and when updated, the old xml config files did not work

## Specific Interview Comments on Documentation

- Then started GUMS install from scratch and it worked, then had problem with pool map factory
  - If someone has FQN, accounts can be mapped
  - CMS T2 has 2000 accounts, so users get mapped to one of those accounts via a SQL database
  - Went back to old mapping and moved table into new server
  - May have been better to have used configs and pool mappings by importing SQL database (maybe documentation on how to transfer old mapping would have helped)
- Another problem - Centos 5 (Linux), the default setting was wrong for his operating systems (GOC told him to get a wiki account and make the change himself)
- There is a name on top, but you can't tell who is responsible for the page
- GOC has a security procedures page
  - Confusing since the terms aren't the same as the information on the Security TWiki page
  - Should link to Security page to make sure it is up to date and uses correct terminology – probably the IncidentDiscoveryReporting page
- Installation TWiki
  - Has a security (certificate) page similar to Security TWiki page
- Trouble ticket needs a VO dropdown
- Need a security policy page that links to area policies so Mine can view (& audit) them
- Security Best Practices should be referred to in installation
- Security analysis – Probably needs a referral from sys admin documentation (from installation documentation or somewhere else?)
- CADistribution should be in release documentation
- CA distribution service should be mentioned in the ReleaseDocumentation twiki
  - Maybe have a security Service category containing:

## Specific Interview Comments on Documentation

- CA Distribution service
- How to get service & personal certificates
- OSGRA page has a requesting certificates paragraph. The links on the left are how to's that link to pages that are different than the pages in release documentation
- Idea: Maybe the installation page on getting and installing certificates should be referred to from the security page
- Personal certificate page on security TWiki page is very thorough and should be used by individual users. Would need link from appropriate user page(s)
- Major issue: The security certifications, etc. are common to any install. They really should be written once and included by all the separate installs, if possible. That also might make it possible to easily skip it if the server has already been set up for another service that required certs.
- Need to tell people about the security environment they are entering for OSG
- To stop access from a VO or add a VO, you have to edit an xml file in 10 different places and there is no documentation on this
- Understanding the data security
- Pool accounts and security
- CA certificate chain and getting certs in right place, etc. Any way to check ownership, filenames, access for these would be helpful
  - 80% of components use certificate verification of some sort (two way verification is even harder)
- Issues on cert authority changes (certs from OSG) – no automated procedure and now no documentation on setting up the cert system
- Have had 1 issue. terminable.uchicago.edu was a machine he first did his cert on and they asked him to use the same machine but it no longer exists. So he was unable to renew it.
- Had problem with certificates. Machines listed in documentation no longer there? Also when certs expire the original machine through which cert was issued is gone, they don't know how to proceed

## **Specific Interview Comments on Documentation**

- Teragrid has a very clear process for certificates. This certificate process needs improving in OSG documentation. Especially when certs have to be renewed.
- GOC has a security procedures page
  - Confusing since the terms aren't the same as the information on the Security TWiki page
  - Should link to Security page to make sure it is up to date and uses correct terminology – probably the IncidentDiscoveryReporting page
  - Even Mine was somewhat confused by the procedure because it doesn't appear to be the procedure that was agreed upon and the terminology is not consistent with that used by the Security team
- Installation TWiki
  - Has a security (certificate) page similar to Security TWiki page
- Interesting: Even a security expert had difficulty finding how to get service certificates on the Security TWiki

## **Specific Interview Comments on Documentation**

### **Education - User**

- On-line tutorial reinforced the class work, but still unable to submit a job
- In workshop, he could submit a test job (the computers were set up and authorized)
- But using tutorial, he couldn't make the test environment work
- Doesn't recall what documentation he used for sure, but has used OSG website
- Really hard to find documentation he needed given the number of tools, etc.
- Many steps to get access to grid and does not have a clear idea of how the grid works
- If he had to use the grid, he would do it
- Given he is just interested, it is not worth the effort
- A working example of the whole process. You have step-by-step process, but it didn't work.
- Need to provide a student/fake account that allows you to just submit a test job to get the experience
- Surprised how long it took to get through it because of getting certificates, etc.
- The security steps were the hardest
- The actual job submission was no problem
- Has experience with Globus and Condor
- Testbed was being used in a class so there was no access for them
- OSG is easier, but was unavailable to my students when there is a GRID class in progress
- Help center support is better in Teragrid (simple ticketing system) – He was unaware that OSG has a ticketing system – he has just emailed people he knows. (Which works great, but his students need to be able to just submit a ticket and not have to know someone.)
- There are a few pages that merged OSG Education documentation onto user documentation that have caused issues



## Specific Interview Comments on Documentation

- <https://twiki.grid.iu.edu/bin/view/Documentation/UsingTheGrid> used to be a clean-cut practical guide. Over time, it has picked up (and mixed up) documents from Education's tutorials, without giving due thought to the target audience.
- Education tutorials are generally targeted toward individuals who are just stepping onto the grid, for a short transient time, and as a light-weight user. As a result, it often doesn't hurt if client-side tools are discussed which may otherwise not be suited to regular users.
- The first link on running a grid job is the one above. Do the education team in the grid school advocate running jobs in a different way from the "standard" best practice? That is currently what it looks like.
- On tech doc TWiki under "Run a job", there is not an agreement on how it should be done – until there is it can't be well documented
  - The issue is that there are many tools in the VDT that do the same job. Each has its place, but it is too complicated to expose directly to a user.
- Most info (Usage of Globus/RSL to end)
- Submit multiple jobs
  - 4 jobs against 1 resource
  - Can't scale up to 1000s (targeted at 1 resource)
  - No failure handling
- How to use compute and storage resources
  - Need short desc of Panda, swift, job router, etc. OSG-mm (layer on top of condor-g – lightweight and simple)
  - See <http://osgmm.sourceforge.net>
- One approach would be to look at different usage models
  - Simple
  - Data staging (or code changing)
  - Requirements parameters
    - Data
    - Network connectivity (Internet access while jobs are running - BOINK)

## Specific Interview Comments on Documentation

- Job independence (most osg jobs are independent)
  - Maybe 3-10 models (actually 2-3)
    - Self contained (small I/O)
      - Data with jobs
    - Intermediate (I/O 100 mb – 10gb per job)
      - Gridftp (pull/push)
      - Swift – Mike Wilde - probably best for small user
      - Panda - BNL
      - OSGmm
      - glideinWMS - Fermi
      - Jobrouter - Miron
      - Pegasus
    - Monster data (I/O > 10gb per job)
      - SRM
    - Dependent jobs (use Swift/Pegasus)
- Comments on Sys Admin Course
  - However I think we should have been much more prepared for this, slides should have been done more in advance and the schedule should have been tightened much more. It was too loopy goosy to my taste (but maybe effective, this is just my view).
  - There is now lost of similar content getting in indico and twiki from the various workshops, there is content by multiple folks and providing links to all these tutorials is rather confusing to me...which one do we look at, which set of slides...etc...I believe we need to be careful in having "forks" of documentations. there should be a well thought out set of slides that can be re-used, and depending on the timeframe you add or subtract module. Right now it still looks like new content each time.
- We now have specific content for each grid school – way too much for any reader – we need to create one we think is best instead of giving 4 examples written by different people

## **Specific Interview Comments on Documentation**

- Powerpoint templates are different – no OSG branding
- Improve TWiki page for Campus Grids
  - Explain the elevator speech – concise message of why
  - FAQs – Why share resources?, what about security?

## **Website**

- The Organization chart needs Dan Fraser listed as Production Coordinator
- Collaborative TWiki and Chat – If you select Chat, you can only login with no info about who should join the group and how to get signed up, etc.
- Brochures page – link to NSF... 2007 – only the 007 is underlined.
- VORS reference from main page and on Monitoring page