RESPONSE TO REFEREE COMMENTS ON

"THE LEGION GRID PORTAL"

Below are the authors' responses to comments from the referees. In most of the cases, the comments have been incorporated into the main body of the paper. However, in some cases, the authors feel that explanations for the comments would interfere with the flow of the paper. Those explanations are presented here.

Comment: *Box C2 says "Legion Grid Portal" but figure heading states the architecture of the Legion Grid portal. Which would imply that the whole figure is the legion grid portal. The authors must decide what the Legion Grid Portal is. Throughout the paper it seems to be used for a component as well as the architecture.*

This problem has been rectified. The general architecture is called The Legion Grid Portal. The specific component has been renamed General Portal Implementation.

Comment: *Re: II.A., At this point the reader asks is this just a wrapper around legion commands. If so, explain what legion commands are out there as many may not know about them. Point out where to get a comprehensive overview of all supported legion commands. A Table would help.*

A reference to the Legion manuals has been included. A full table of the Legion commands would distract the reader from the portal discussion.

Comment: *The question arises if I can list objects that are not owned by the user. This would indicate a violation in privacy. This must be pointed out.*

The text already mentions that users can access only those objects for which they have permissions to do so.

Comment: *Page 3: Users "should" : are you doing it or not? Be more specific.*

Clarified the wording.

Comment: *References to other information services such as SNIPE or MDS seem missing. Such tolls* [sic] *are provided for example by Hotpage. What about references.*

The inclusion of these services would certainly benefit the portal, but they are not necessary for functioning.

Comment: *Re: II.D, The term Amber is used but not defined.*

Corrected.

Comment: *Re: III, Major contradiction in first sentence: How can the Legion Grid portal contain a component that is called a Legion Grid Portal (C2). This is recursive and must not be possible in this case.*

Corrected, see above.

Comment: *Avoid the wording "special portals" why not problem specific portals, wouldn't that be more precise?*

Changed the wording to "specific portals".

Comment: *Page 3: I asked myself here how you deal with exceptions.*

The underlying grid infrastructure handles exceptions in a sophisticated manner. Describing exception-handling in Legion is outside the scope of this paper. Legion transmits exceptions in the form of error messages. The portal parses those error messages to display to the user.

Comment: *Page 3: The mix of PhP and CGI does not look like a well designed system to me. Why not just use php or cgi/ perl. It seems a hack to me. It almost looks like that because php was at hand you did php but kept other parts in cgi since you did not want to touch them.*

To an extent, yes, that was what we decided. The General Portal Implementation was written in Perl. When we decided to interface with legacy database systems, PHP seemed like a natural choice to access them. The mix of Perl and PHP is an example of the flexibility of the Legion Grid Portal.

Comment: *Page 3: I believe it may be important to not just have command line wrappers but also components that connect directly with other Grid services.*

We agree, as we discuss in Section II.C. Since we have not yet implemented such direct connections, we felt it premature to discuss them here.

Comment: *Page 3: The comment about Globus and the distributed filesystem is unclear. Instead explicitly describe who your portal can make use of Globus. Also I believe that many would be curious if you have such an integrated portal Legion/or Globus working. How does the architecture figure change? Isn't ther a more general architecture than the one you describe? How can we integrate with Hotpage.*

The basic architecture of the Legion Grid Portal does not change. Instead of invoking Legion command-line tools, we would invoke Globus command-line tools. We are not integrated with HotPage currently, but such an integration is possible (as discussed with the UI group at SDSC).

Legion provides a distributed filesystem for every grid. Globus does not. Therefore, if we adopted Globus as the underlying grid infrastructure, we would not be able to display a distributed filesystem using the Legion Grid Portal. The portal display could be modified in that case without changing the basic underlying functionality.

Comment: *Fig. 2: all over sudden the component C2 is now called CGI Perl Script.*

In this figure, we name all components with their implementation details. Hence, "Database" or "Run Script". The choice of "CGI Script" is consistent with that naming.

Comment: *Page 5: "perhaps by logging in" I do not understand this how this works and who* [sic] *you get credentials on a logging in basis?*

Legion provides a single sign-on mechanism to access a grid. The process is called "logging-in". The role of this process is to generate credentials for the user. The details of how Legion generates those credentials is outside the scope of this paper. From the perspective of the portal, it suffices to say that we can log in to a grid managed by Legion.

Comment: *Page 5: "detailed knowledge about Legion" Which steps do not require detailed knowledge, which require me to get that knowledge, and where can I obtain that*

knowledge quickly. I like to see more details on the legion handler.

The details of each command in Legion are present in the Legion manuals (added reference to it). The handler simply parses the output and error that can occur in each case of running a Legion command.

Comment: *Page 5: second column: Are you confident that caching of proxys similar to the Globus MyProxy mechanism is secure? Your architecture must include a complete secure and separate machine behind a very secure firewall with nothing on it other than the cache. Using a cgi based mechanism probably forces you to this solution.*

Actually, it's the grid infrastructure that forces us to this solution. Incidentally, Globus has a similar mechanism, and the HotPage performs a similar task.

No, this solution is not perfectly secure. However, it's a reasonable start to a difficult problem.

Comment: *Page 5: I do not believe that the way the session id is created seems unbreakable.*

We disagree. The probability of guessing the session ID is 10^{-45}. The session ID is encrypted during transmission from the browser to the server. The ID is stored using as restrictive Unix permissions as possible on the server. Perhaps none of these solutions are secure for some pathological cases, but for the common and resonable cases, they are sufficient.

Comment: *Page 5: you use crontab, but what about running the server on windows.*

We have not considered running the server on Windows.

Comment: *Another point is that the use of html/cgi portals encourages the use on shared terminals. How can you make sure to remove your contents on them?*

We don't understand the question. Which contents? User-specific state information is saved by the portal on the server in separate directories so there is no overlap. Consequently, several users can access the portal concurrently without interfering with one another.

Comment: *Page 9. handlers are complex. Actually, many cs programs are complex, instead describe how it is done. Do you mean they use asynchronicity?*

Added some explanation. In keeping with the suggestion to not go into too much detail, we have not described a handler step-by-step.

Comment: *Page 11: by now I am asking myself why not have the whole portal done in php?*

It is possible, though pointless.

Comment: *Page 12, section Security: Though you try following the given GCE format, you have mentioned the security already so many times before that this section looks like a repetition. But I acknowledge the fact that it is difficult not to do so based on the template given. Maybe it is better not just to follow the template and try to better group certain descriptions based on a better architecture figure you provide.*

We regret the repetition. However, our objective was to follow the outline suggested to us as faithfully as possible.

Comment: *Page 13: Passing a password as part of a command line to the legion system that can be viewed with ps makes me very uncomfortable using the system. I recommend changing this while using or developing a different method.*

We agree. We are investigating methods to rectify this.

Comment: *Page 14: explain better what in implicit and explicit task is.*

An explicit create or run involves the user specifying the target host whereas an implicit creates involves letting the grid infrastructure select the host. This explanation is present in the section.

Comment: *Change the IEEE style to the concurrency style*

We expect to do so after all referee changes.

Comment: *Quality of the images is week* [sic]*, use vector graphics not gif.*

Added higher-resolution pictures.

Comment: *I applaud the author on the use of she instead of he! make sure if this is consistent with the whole paper.*

Thank you. We switch genders of users in different discussions.

Comment: *Figures should be reduced in final print. But good that I can read them here ! Reduce the size of the screenshots by not showing so much redundant information. Simply make the area displayed smaller.*

We have continued to provide visible and readable graphics.

Comment: *Page 17: describe more about the error modes. This seems important but leaves the reader puzzled.*

We're not sure what more is required. Any command can print errors, whether the command is a Unix/Windows command or a Legion command. The portal merely presents those errors in a pleasing manner.

Comment: *Page 17: running application in a non-blocking or asynchronous mode is a necessity in Grid programming.*

We agree.

Comment: *Missing: section about user requirements that motivate portal, how many users use the portal, , what lessons do you have learned while using the portal in real live* [sic].

These sections are dispersed in the paper in accordance with the outline given to us.

Comment: *The presentation does not cover sufficiently how the scheduling is done, how data transfers are performed (FTP, UNIX RCP, something else ?) and if transfers exploits security mechanisms.*

This paper does not discuss both of these issues because they are out of its scope; these functions are carried out by the grid infrastructure, not the grid portal. Both of these questions are answered in the several Legion papers available online (`http://legion.virginia.edu`).