

1. NASA IPG Launch Pad

1.1 Introduction

The NASA IPG Launch Pad is a web-based user portal to the Information Power Grid. The Launch Pad will provide IPG users with the ability to:

- view the status of Grid resources
- prepare, submit, and track jobs
- manipulate files at remote locations
- customize the environment to create your own portal view

This will be accomplished by developing user friendly interfaces to underlying services that enable access and utilization of the Grid.

Our use of the term "user portal" is in reference to the concept of a web based user interface that provides access to a set of resource, specifically, the NASA Information Power Grid. This easy-to-use interface will provide access to underlying functions that enable the user to make use of Grid resources.

1.2 Collaboration

This project is being done in collaboration with the NSF PACI programs at SDSC and NCSA. The purpose of this collaboration is to provide a common set of functionality upon which site specific portals can be built. Toward this effort we have contributed a Collaboration Web Site, suggestions on required functionality, and act as a test site for the resulting Grid Portal Development Kit (GPDK).

1.3 Technology

The IPG Launch Pad user interface will be developed using current web technologies such as JavaScripts. The underlying functionality will be provided using Java Servlets via Java Server Pages, which access the Grid Portal Development Kit (GPDK). User preference information will be maintained using a JDBC interface to a database.

1.4 Security

The IPG Launch Pad will provide information about Grid resources, and user jobs. Some of this information will be restricted to those who have authorization to use the resources. In addition, authorization to submit jobs to the IPG will be limited to those who have the proper authorization. The IPG currently uses Globus Certificates to authenticate a user's access to Grid resources. The IPG Launch Pad will authenticate a user via a delegated certificate using the MyProxy technology which allows the users to designate the duration of the delegation. This will be described later in this document. At this point we will simply make a distinction between information which is secure, requiring logging into the portal as a user, and information that is

not secure, which can be viewed by logging into the portal as a guest. Utilization of Grid resources requires further authentication via MyProxy.

1.5 Browser Availability

In compliance with NASA direction, we will make all attempts to keep the web pages accessible to the two major browsers, as well as browsers equipped for the handicapped.

2. Functions of the Launch Pad

As mentioned above, the NASA IPG Launch Pad will provide a variety of functions that will enable the users to perform their tasks. These functions will include:

1. User Session Functions (login, MyProxy, session data)
2. Resource Status (information services)
3. Job Processing (job creation, submission, and tracking)
4. File Manipulation (remote copy, list files, etc.)
5. Customizable Features
6. Problem Reporting/Tracking
7. Documentation
8. Account/Certificate Acquisition

Not all of these functions will be handled by the GPDK but, will be handled by other services, such as the Remedy Web Interface. In addition, the last two functions will be accessible outside the portal, prior to login. The following sections will discuss these functions in more detail.

2.1 User Session Functions

In order to distinguish between authorized usage and casual (guest) usage, we will track user sessions. This use of sessions refers to the concept described by Java Servlets with regard to scope. By making use of "session" scope we can provide the ability to log in, or create a Portal Session, which can be ended by logging out, and destroying the Portal Session.

User session functions roughly cover four areas:

1. Session Creation - user login
2. MyProxy Delegation - myproxy passphrase authentication
3. Session Destruction - user logout
4. User Session Data

2.1.1 Session Creation

In order to access secure information and perform portal functions such as job submission, a user must create a user session by logging in to the Launch Pad. This is the way the login bean works at this time. We may wish to discuss this further. Questions that come to mind here are:

What information might we show someone who can log in from whom we would not also require a delegated certificate proxy?

If the answer is nothing, or very little, then the question might be asked, why make the user enter two password/phrases?

2.1.2 MyProxy Delegation

MyProxy is a delegation process for the Globus X.509 Certificate. A user creates a proxy certificate using the **myproxy-init** command on one of the Grid hosts, designating a time frame for which the proxy will exist. Once this has been done the user can authenticate themselves for a delegated proxy on a web server that is registered with the MyProxy Server. They can gain access to Grid resources by entering the appropriate passphrase, and designating the amount of time they want the session to be active. If they enter the correct passphrase, they receive a certificate and can use the portal functions that access Grid resources. Failure may occur if they either enter the wrong passphrase, or their proxy has expired.

2.1.3 Session Destruction

To end a session the user selects the log out button. This calls another java bean that destroys the session bean, thereby ending the session and all access to the Grid resources. Even if the user forgets to log out, their session will time out based on the interval they selected when they acquired the Proxy Certificate.

2.1.4 User Session Data

User session data refers to two kinds of information. The first kind of information is the "user preferences" data obtained from some source to be determined (initially, probably flat files on the server). The second type of data is information the user specifies during the session, such as resource specifications for job submission, and other data related to the users activities during this session. Some of this information may be retainable. This will be detailed later in the section on user customizable information, and file manipulation section.

2.2 Information Services

This Section of the IPG Portal provides access to a variety of information about the resources of the IPG. In part, the information services functionality will provide information such as the information currently provided by the "Hot Page" and associated pages. In addition, this piece of the portal pie will allow a user to select a subset of the Grid resources to be used in a job request. The *querybean* from the GPDK will provide information obtained from the Globus MDS. Additional scripts may be necessary to collect information which is only obtainable from the Job Scheduling systems in use on the various hosts.

Information services may be broken down into a few basic categories:

1. host information
2. job information
3. possibly user account information

2.2.1 Host Information

A variety of information about compute hosts may be useful to help users determine which resources they wish to use. The following is a list of most of the resources that can be tracked. Some of the information will come from the batch scheduler being used on a particular host.

1. Hardware
 - a. System make and model
 - b. Number of CPUs
 - c. Amount of memory
 - d. Amount of available storage
2. Scheduled interruptions
3. Recent load average
4. Message of the day (/etc/motd)
5. Batch Queue configuration
6. Disk usage
7. Status of key IPG components
8. Software information

2.2.2 Job Information

As with host information, a lot of information about jobs can be obtained from the system and batch queueing system, and made available to portal users. The following information is available:

1. Batch job identifier
2. Execution Host
3. Memory (used and requested)
4. Walltime (used and requested)
5. NCPU requested

2.2.3 User Account Information

To the extent that user accounting information can be obtained, this information can be made available to the user that owns the information. Exactly what information will be widely available is yet to be determined.

2.3 Job Submission

One of the primary purposes of the IPG Launch Pad is to be able make use of Grid resources, in other words, execute programs on the Grid. This process may be broken down into a few steps:

1. Review Resources
2. Construct Job
3. Launch Job
4. Track Job

2.3.1 Review Resources

One process a scientist may wish to perform is to review the Grid resources to see what is available, for example, what kind of systems, how many CPU's they have, how much memory, how much mass storage is available for running a job, etc. This will be available via the functionality described in section 2.2, *Information Services*. Access to this service will be a part of the *Job Submission* service.

2.3.2 Construct Job

The steps of constructing a job are:

1. Select or create a job file
2. Select resources
3. Specify command to execute

These services are handled by utilizing Java Beans that are part of the GPDK. In addition, interaction with the *File Management* section will be necessary to obtain previously created jobs, and to save the current job if desired.

This service will be handled by the *JobBean*, and *JobSubmissionBean*, along with a *JobInfoBean* that will contain job information used by the others. This section of the Portal will provide the ability to construct jobs, or retrieve job files from Grid hosts, modify them, and submit jobs to the Grid as one might from any of the hosts on the Grid. This interface works in conjunction with the *user configurable parameters* outlined in Section 2.6, "User Portal Customization."

Utilizing information obtained by the Resource Status services (Sec. 2.2), the user may compare features of the various systems to determine which they wish to use. This information will be accessible from the Job Submission Pages of the Portal. The user may then select Grid resources that will be used to construct a job, or they may upload a job file from a specified Grid host.

Once a job is constructed, it may be launched to the Grid. Status may then be retrieved using features of the Resource Status services outlined in section 2.2.

2.3.2.1 Select or Create a File

An option will be provided that makes use of the "File Manipulation" service to list and select a file to use as the job file. An alternative will be to create a new one by filling in the form for required resources.

2.3.2.2 Select Resource

The selected resources will be saved in a *JobInfoBean* associated with the current session. Initially, resources will either be loaded into the resource selection form from a file that gets loaded or an empty form will be available to create a new job specification. A later stage of the Job Creation page will make resources dependent upon each other, and upon known limits of selected systems. The following are some of the resources that may be identified:

1. Host Selection
2. Memory Requirements
3. Number of CPUs
4. Time required in walltime
5. Location to return standard output
6. Location to get standard input

2.3.2.3 Command Specification

The final step prior to launching the job is to identify the command to be executed. This may be a specific executable, or a command script that is required to execute several commands on the Grid host.

Co-allocation of Jobs

This future feature allows the user to specify that the job run on multiple hosts.

2.4 File manipulation

File manipulation is a functionality that will be useful from the job submission section, as well as a stand alone section from which a user might wish to manage files on various IPG hosts. This basic functionality should be flexible enough to work in conjunction with other sections. This service will be able to perform the following actions:

1. List files on specified host
2. Find a file on a specified host
3. Upload files to the web portal server

4. Download files from the web portal to specified locations
5. The ability to copy files between two IPG Hosts
6. Move files, i.e. change the name of files
7. Copy and remove, i.e. copy files from host "A" to host "B" and then delete the file on host "A".
8. Deletion, or remove a file from a specified host
9. Limited Editing of files
Ability to edit a file. This should be able to access the file from one host, edit it on the web server, and save it on some host other than the web server (choice of filename too). This will be limited to simple editing as might be done using the *textarea* feature of forms.

2.5 User Portal Customization

This section of the IPG Portal provides the mechanisms for a user to customize certain features, look and feel, and set filters that limit the information presented by the Portal.

1. Auto Generate list of hosts, let user choose from the list of all hosts
2. Establish defaults for commonly used forms
3. Eventually, color scheme
4. Eventually, compose screen layout for PSE's

2.6 Problem Reporting

This section of the IPG Portal provides the ability to submit a problem ticket to the problem tracking system.

1. ability to contact support online
2. enter a problem ticket, route to correct support organization

2.7 Documentation

This section of the IPG Portal addresses documentation provided. This will be broken down into a few areas:

1. **Reference Manual** - This document should provide information such as man pages. It should also document portal interfaces in order for others to be able to create their own portal/PSE.
2. **Quick Start Guide** - This document should provide minimal information. It should provide just enough to get a user started. A user should be able to step through this document and have the end result be output from a simple job.

3. **Portal level spot documentation/help** - These documents should provide portal specific information needed to complete the task intended by the portal. It should probably provide more detailed descriptions of form entries.

2.8 Account/Certificate

In an effort to simplify the process of obtaining a user account on a Grid related system, an on-line request system is being developed. The process of obtaining a Certificate will still require the Certificate Authority Manager to review the request to obtain authorization from the appropriate authority. The idea is to eliminate some of the confusing steps required to get a certificate for access to the grid.

The following list includes the initial capability and several additional features that may be automated as time permits.

1. **Request** - This interface should simplify the initial request for an account on the Grid.
2. **Renewal** - This interface should simplify the renewal of an account on the Grid.
3. **Allocation/Update** - Ability for users to request more hours with the request going directly to the appropriate PI. Additionally, ability to validate the request and increase the allocation by "authorized people".
4. **Tracking** - Ability for a user or a PI to view usage for those allocation entities to which they have access.

3.0 Deliverables

This project will be done in phases. The initial phase is to get some of the basic functionality working in order to demonstrate this capability for a Level One Milestone in September. To this end, the first deliverable is this document. The next deliverable will be a working prototype Launch Pad that can submit a job to the grid. The follow-on phases will be progressive enhancements to the Launch Pad.

3.1 Milestones

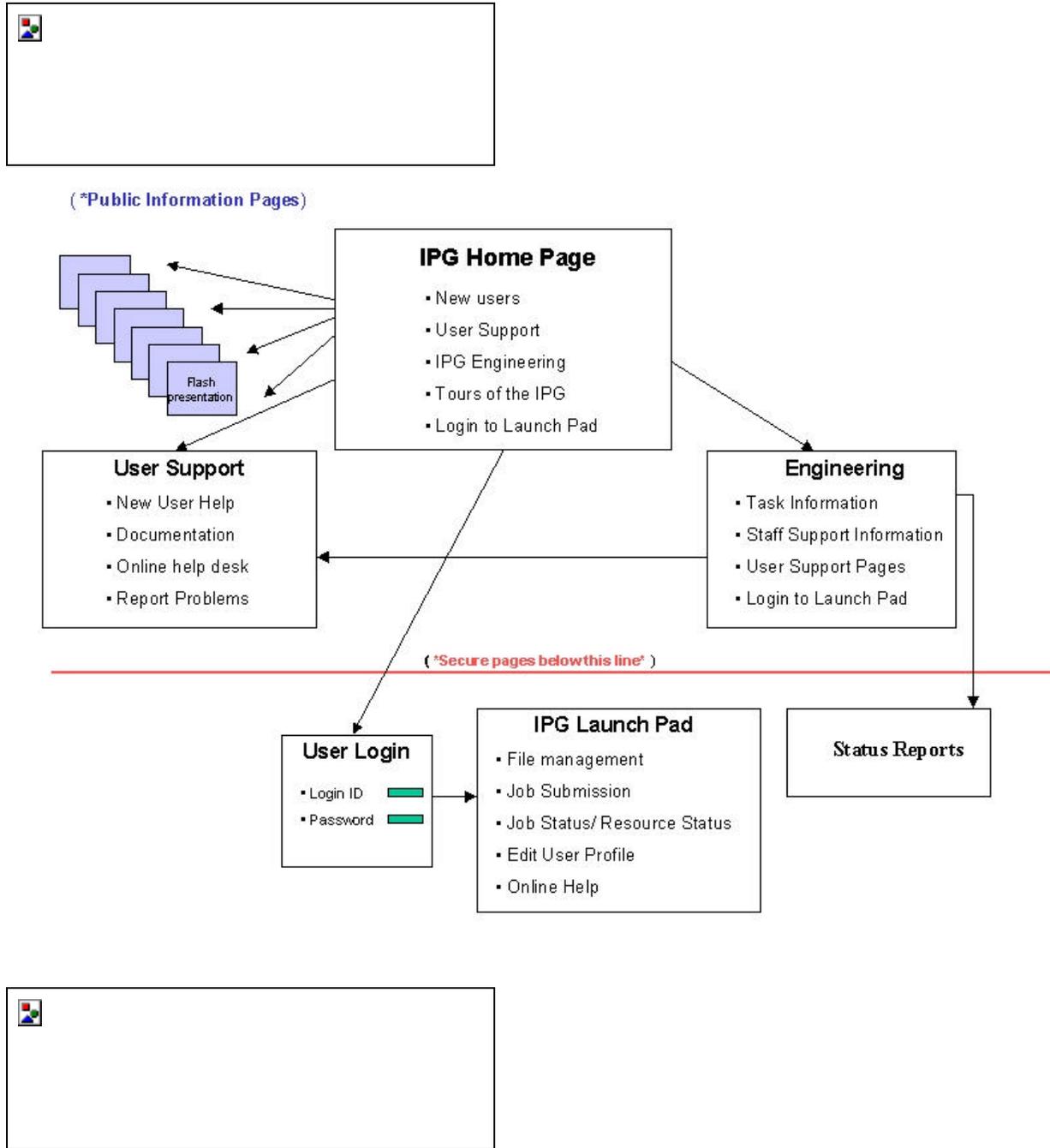
Date	Milestone	Done
06/29/00	Draft IPG Web Portal Design Document	
07/??/00	Working IPG Web Portal Design Document	
09/29/00	Demonstrate Prototype IPG Web Portal Capability	
11/15/00	Demonstrate IPG Web Portal Capabilities at SC2000	
03/15/00	Demonstrate fully functioning IPG Web Portal	

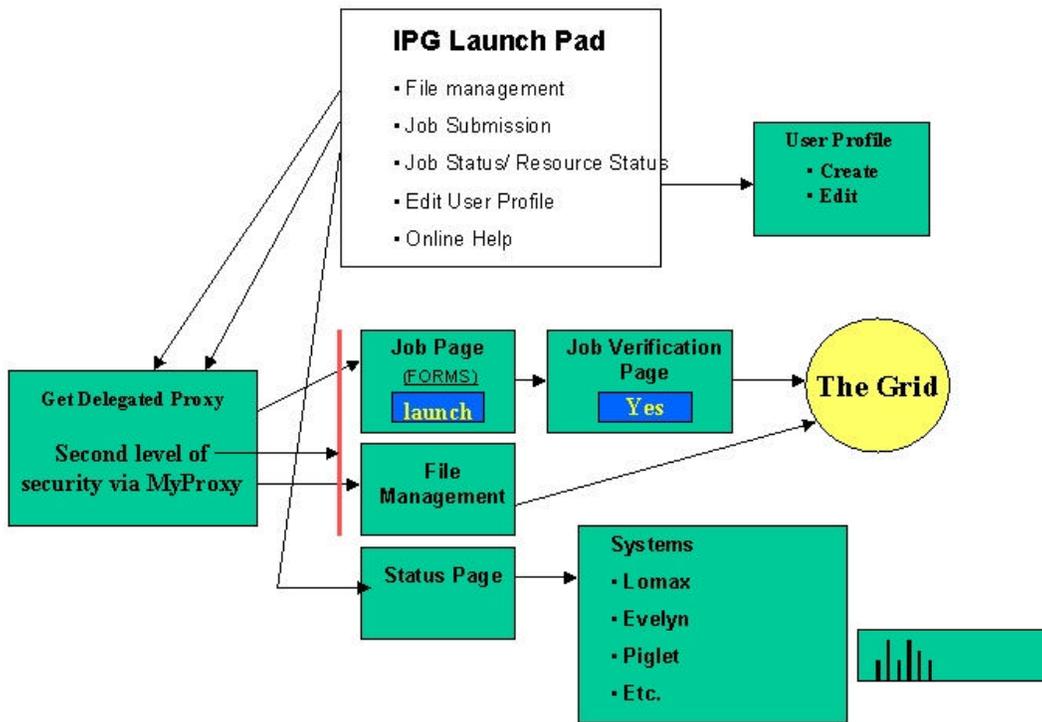
3.2 Future Plans

The long term plans are to make it possible for a user to customize the portal to fit their particular Problem Solving Environment needs. To this end we are keeping this requirement in mind while we design each component so that they can be "plugged-in" to a particular environment.

Appendix

The following graphics describe the overall page layout of the IPG web site and Launch Pad.





New User Page (formerly Getting Started)

- [What You Need to Do to Use the Grid](#)
- [How to Get an ID](#)
- [How to Get a Certificate](#)
- [Setting Up Your Environment](#)
- [Where to Get Help](#)
- [Some Information on IFC Systems](#)

User Support Page

- [Link to New User Page](#)
 - [Basic Usage Information](#)
 - [Getting a Proxy](#)
 - [Running Jobs From PBS](#)
 - [Running Jobs From LSF](#)
 - [Basic Globus Commands \(see Globus Quick Start\)](#)
 - [CORBA and the Grid](#)
 - [Condor and the Grid](#)
-