



HPC application services providing portal

Y. L. Li, B. T. Cheok & S. K. Chin
Institute of High Performance Computing
Singapore

Abstract

High performance computing (HPC) is widely used nowadays. But the investment in HPC is a large amount both in hardware/software purchasing and maintenance. So the market has a desire in looking for application services providers (ASP) to provide the HPC services. The providers set up all the required HPC infrastructures. Users can subscribe and remotely access the high-end resources via Internet through a low-end computer with a browser. It makes HPC services available at anytime, anyplace and on any device. This helps greatly in reducing individual HPC investment and makes HPC resources widely available. Based on this desire in the market, an HPC application services providing portal has been set up in the Institute of High Performance Computing (IHPC), Singapore.

This paper first briefly reviews the existing ASPs in the market. Then it describes the features of the HPC application services providing portal in IHPC. It also presents the three-tier distributed computing architecture for extending HPC resources from high-end servers to low-end desktop computers. The HTML-based Unix shell command tools and data transferring tools are also designed because of the security limitations. Finally, the implementation of the portal is given and conclusions are drawn on it.

1 Introduction

HPC is widely used in fields such as computer-aided engineering (CAE), computational fluid dynamics (CFD) and life science. HPC needs expensive supercomputers and software packages. The investment is not easily available to many small-to-medium sized companies. Even the investment is not a problem to the big consortiums, but quite often, the investment is only sparingly used in some particular projects. After the projects, maybe the investment in dedicated

hardware and software packages is idle and not helpful to other suitable projects. So the market has very strong desire in looking for HPC application services providers. Whenever the expensive HPC resources are required, the suitable HPC application services provider(s) can be found.

The easy Internet accessibility and wide bandwidth technology bring the desire into realization. To cater the desire, an HPC application services providing portal has been built in the Institute of High Performance Computing (IHPC), Singapore. Users need only do a registration on the portal and send their particulars and requirements of HPC resources to the portal administrators. Upon receiving the registrations, system administrators inside prepare user accounts for them after auditing their backgrounds. Then the registered users are informed and can access the HPC resources from the portal via Internet using low-end desktop computers.

This paper discusses all the technologies for setting up this HPC services providing portal. Before discussions of the technologies, a brief review of the available ASPs in the market is given.

2 Review of existing ASPs

Traditionally in the software distribution industry, Applications Developers deliver their software packages to customers in a packaged, shrink-wrapped distribution model. In the Internet age, a new business mode, ASP, is introduced and being under rapid development. The vendors host their applications on the server. Users then rent this application either by running what they need "live" over the Internet, downloading the application to their desktops on an as-needed basis, uploading their files to the ASP server where the analysis is performed, or combining these operations systematically to accomplish a complex task. The most prominent ASPs in the market are for CAE analysis and simulation because these applications are sold in very high price, but at the same time they are applications that may be used only sparingly [1].

Table 1. Typical ASPs.

Vendors	Functionality and URLs
Alibre	Internet based CAD service providing http://www.alibre.com
MSC.Software	Web-based FEA simulation with eVisualNastran http://simulate.engineering-e.com
PTC	Web-based collaboration for enterprise PDM with Windchill http://www.ptc.com
Tech-Net	Web-based CAE consulting http://www.tecnetinc.com
Theorem Solutions	Web-based data translation http://www.cadverter.com

Among the existing ASP portals for CAD and CAE applications, the one run by MSC.Software is highlighted. The URL of the portal is <http://simulate.engineering-e.com>. This portal was set up by MSC.Software to

provide the most sophisticated software applications in an online environment. The focus is to provide premium tools in simulation, analysis and CAD that are essential to the engineering community. Instead of having to purchase new equipment to handle the immense power of these applications, users can run these applications anytime, from almost any computer connected to the Internet. For other ASP portals, the functions and models are nearly similar and are left out here for concision. They are listed in Tab. 1. Readers can surf their web sites to explore more.

All the portals listed are run by the software vendors and are used for their software packages outsourcing. Though software outsourcing is also one functionality of our HPC services providing portal, yet it also outsources the high-end supercomputers in IHPC. Its subscribed users include CAD, CAE and CFD engineers, and each user has his individual requirements even on the same application. At the same time, users must have confidence in the security of their data on the third party's servers. These render the portal special features and add more factors to be considered in construction.

3 Factors to be considered in building the portal

The portal is accessed by various users through Internet. Users have different requirements on its webtops, software/hardware configurations and security. This requires the portal can be customized for the users to cater their individual favorites and subscription. At the same time, users also worry about the security of their data on the portal. So security is another important factor. Security relates to the users' confidence to their data on the portal servers. It influences users' confidence in using the portal. This section discusses webtops customization, applications customization and security of the portal.

3.1 Webtop customization

This section discusses users' webtop customization based on their favorites and subscriptions.

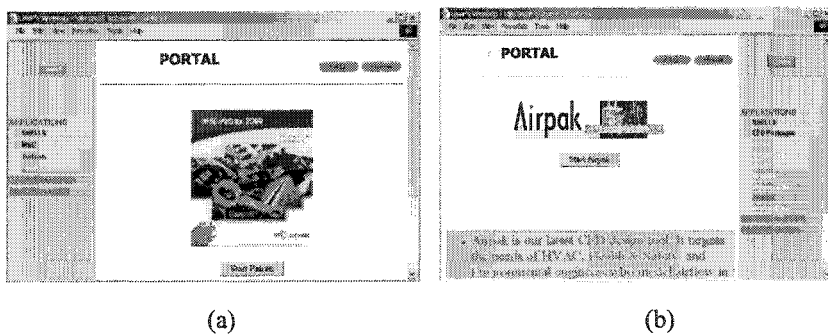


Figure 1. Customized users' webtops

Webtop is the interface through which users launch applications hosted on the high-end supercomputers. Its layout can be customized according to individual's favorites and subscription. Webtop customization requires each user's webtop is dynamic pages. Using dynamic pages, the layout of a webtop can be customized according to the user's favorites and its content can be customized based on his subscription. Dynamics pages can be achieved using server side programming technologies of Java servlets [2] and JSP [3]. Fig. 1 shows two customized webtops. The left one is for a user who subscribed CAE applications of MSC.Software and the right one is for a user who subscribed CFD packages. Their layouts are different.

3.2 Applications customization

Even users have interest in the same application, but their requirements and usage of the same application may be different. Most CAD, CAE and CFD applications include different modules for different purposes. There is no reason to give all the users the same configuration even for the same application. The application must be configured according to users' subscription and need. This can be easily solved if the application can be configured at the operating system (OS) level that which user has the authority to access its different modules. But unfortunately, most applications still have no this option. Once the application is installed, all users can launch anything has been installed. To solve this problem, the application is installed in different machines or in different directories on the machines. Each installation has different configuration of the same application. Different user is given different installation of the application.

Subscribed users often require the hosted application can supply them tools or batch commands to ease their work, especially tools with domain knowledge. Using the tools, users can do design analysis or simulation easier with knowledge wrapped inside. Some of the tools actually are small expert systems [4] developed based on the hosted applications. For example, using PCL in Patran® and ACUMEN® programming interface, some useful tools such as parameterized analysis and intelligent analysis tools can be developed for the hosted application of Patran® [5].

Applications customization gives the portal lots of attraction to the users because they found the customized applications provide them many additional useful tools with design and analysis knowledge.

3.3 Security

Security is an important usage issue in the HPC application services provider portal. The portal is to provide the users with the three most important elements of computer security: confidentiality, integrity and availability [6].

- **Confidentiality:** Confidentiality refers to the authorized access to users' data that passes through or resides on the servers.
- **Integrity:** Integrity is about the writing of data. It means that users' data will not be deleted or modified by anyone without permission. This means users will be able to access their data as they have left it and in whole.

- **Availability:** Availability is about being able to use the service when needed and without undue delay. Availability ensures that denial-of-service attacks are not possible.

To achieve these security elements for the portal, the following methods are applied.

3.3.1 Users' backgrounds auditing

Users' backgrounds auditing ensures all the registered users are legal users and can be traced. The HPC application services provider of IHPC reserves all the rights to audit the users' background and projects they need the HPC resources. And the HPC resources are strictly used only for the registered users and approved projects. This ensures the registered users will not abuse the HPC resources.

3.3.2 User ID and password protection

Access to the HPC application services providing portal is controlled by a user authentication method. User names and passwords are encrypted and stored in a secure, firewall protected database. When a user login to his account from a browser, his password and user name are never sent together, and they are sent after encryption of secured socket layer (SSL) protocol.

3.3.3 System physical security

The HPC application services providing portal's equipment, including all servers and data storage equipment, are hosted at the machine room owned and operated by IHPC. They all exercise the highest security. Advanced security features such as camera-controlled access, video monitoring and card key access are provided for protection. The facility is equipped with fault-tolerant power systems and is fire and earthquake ready. Only authorized personnel are permitted to access the locked cabinets in which the servers are contained. The IHPC data storage system utilizes RAID 5 disk fault handling that allows data to be available even during a partial drive failure. These measures can ensure not only availability, but also the confidentiality and integrity of the portal.

4 The three-tier architecture of the portal

The HPC application services providing portal is built around a three-tier architecture model. It includes the client devices tier, middle tier and application servers tier as shown in Fig. 2.

4.1 The front tier

The front tier includes the client devices. It is a piece of hardware that can communicate with the middle tier using a web browser. They are thin clients and there is no special application installation and configuration for running the applications hosted on the application server tier. The only requirement is a Java enabled web browser.

4.2 The middle tier

The middle tier includes the broker hosts. A broker host is a piece of hardware running a web server and an application broker. The web server generates documents that are sent to client devices. Java applets embedded in these documents allow direct communication between client devices and the application broker server. The applets use specific protocols, such as TCP/IP, X protocol or even Adaptive Internet Protocol (AIP), which is a protocol can adapt to the client devices being used and the applications being delivered, to provide optimum performance. The web server also has the responsibilities of negotiating with application servers to authenticate the users and keeping track of running applications even after users have logged out, so that they can be resumed later.

The applications broker responses the communication between the application servers and the web server. It generates the compressed data stream goes through the application servers and client devices using a dedicated protocol. The compressed data stream was finally interpreted through the Java applets downloaded to the client web browser. In this way, the application's user's interface is transferred from the application server to the clients. Users do their work in the transferred interface in their local clients.

Here the web server is Apache [7] plus Java Servlet engine of Jakarta Tomcat [8]. The application broker is Tarantella[®] Enterprise 2 from SCO [9].

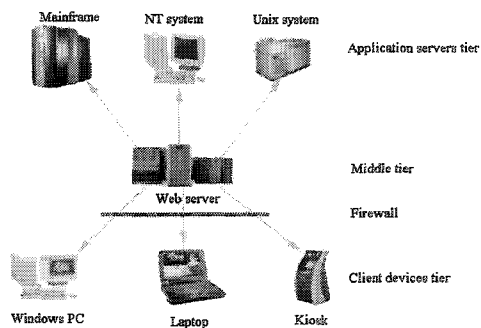


Figure 2. The three-tier architecture model

4.3 The third tier

The application servers compose the third tier. An application server runs users' applications. When a user clicks a link on his webtop, the request is sent to the web server via the browser. Then the web server passes the request to the application broker and the application broker starts the application on an appropriate application server. The data generated from the application server is compressed and redirected by the application broker from the application server to the client device. This actually is the key mechanism in the portal regarding

how to launch an application and display it remotely through Internet on the client devices.

5 Web-based tools design

When users access the hosted HPC applications, they need to send shell commands to manage their running jobs and files on the application servers frequently. It is easy work in the local machine to use shell commands. But it changes here because users access the hosted applications remotely.

Because of security issues, users are not allowed to use any shell terminal window. They have no permission using ftp (file transfer protocol) tools for transferring files between their local PCs and the high-end HPC servers either because ftp port opening in the high-end machines is a big security hole to them. To avoid any attack from outside, only ports for normal HTML files, SSL data transfer and other high ports(>1024) for the application broker are open in the firewall. Only data streams using HTML and application broker protocol are allowed to go through. Because the data stream using the application broker protocol is confidential to the third party vendors and only the dynamic HTML data generating and processing can be controlled using server side programming technology such as Java servlets and JSP. So all the shell commands and file transferring tools can only be provided in HTML-based way.

5.1 HTLM-based shell command tools

There are only the most frequently used shell commands provided in the HTML-based way. For examples, *ps* and *kill* are among the most frequently used Unix commands and they are moved to a web-based way. To realize these commands, server-side programming technology of Java servlets and JSP is applied. Running of the tools is shown here. Whenever user clicks the “List Process” button in Fig. 3, her current running processes are displayed with the process ID and other information as shown in Fig. 4. Based on the process ID listed, the user can kill the process by key in its ID into the form and pressing the “kill” button.

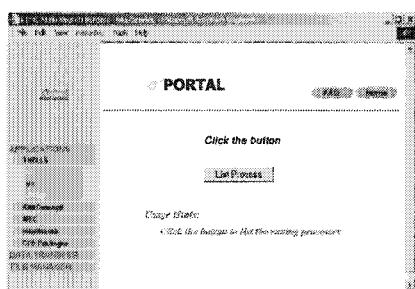
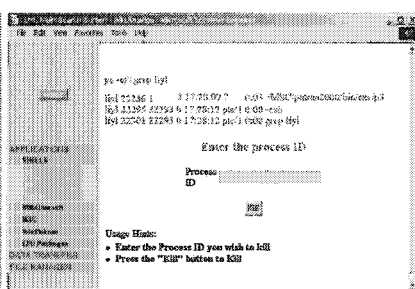
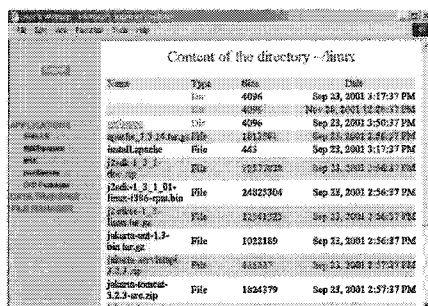
Figure 3: Interface of “*ps*” command

Figure 4: Current running processes

Besides the Unix comands of *ps* and *kill*, other frequently used comands like *zip*, *unzip* and file managing commands such as *ls*, *rm*, *rmdir*, *mkdir* etc. are also provided in HTML-based way. Fig. 5 shows the running result of “*ls*” command.

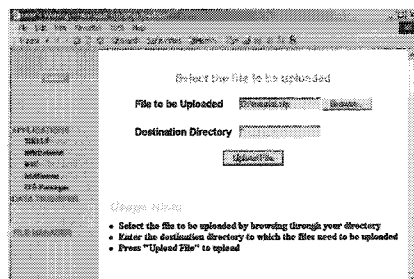


Name	Type	Size	Date
.	Dir	4096	Sep 23, 2001 1:17:37 PM
..	Dir	4096	Nov 16, 2001 12:06:37 PM
apps	Dir	4096	Sep 23, 2001 3:50:37 PM
apache_1.3.2.4.tar.gz	File	1811261	Sep 23, 2001 3:50:37 PM
install.apache	File	443	Sep 23, 2001 3:17:37 PM
jdk-1.3.1	Dir	7237228	Sep 23, 2001 3:50:37 PM
jdk.zip	File	2482304	Sep 23, 2001 3:50:37 PM
jdk-1.3.1_01	Dir	1234323	Sep 23, 2001 3:50:37 PM
jdk-1.3.1_01.tar.gz	File	1021180	Sep 23, 2001 3:50:37 PM
jdk-1.3.1_01.zip	File	433333	Sep 23, 2001 3:50:37 PM
jdk-1.3.1_01.tar.gz	File	1804379	Sep 23, 2001 3:50:37 PM

Figure 5: Listing files

5.2 HTML-based file transferring tools

There are frequent files uploading and downloading between the clients and application servers. But in this HPC services providing portal, using ftp for data transferring is prohibited by the firewall because of security issues. So HTML-based data uploading and downloading tools are designed.



Select the file to be uploaded

File to be Uploaded

Destination Directory

Change files:

- Select the file to be uploaded by browsing through your directory
- Enter the destination directory to which the files need to be uploaded
- Press "Upload File" to upload

Figure 6. File uploading

Files uploading through HTML is achieved using HTML form-based uploading [2]. Form-based uploading (also known as browser-based uploading) is the process of sending files from a client machine to the web server with a browser via an HTML form. This HTML form must have a special attribute, ENCTYPE="multipart/form-data", and contain one or more items of the type <INPUT TYPE=FILE> through which files are selected for uploading. The web server use a server component to capture files uploaded through a browser. This component is a Java servlet in the web server. After retrieves the file(s) selected in the form, the servlet also puts the uploaded file(s) to a specified destination. Fig. 6 shows the interface of HTML based file uploading. User presses the “Browser” button to locate the file to be uploaded in the local directory. Once the

file is selected and the destination directory is specified, user need only press on the "Upload File" button to upload the selected file to the specified destination.

It is straightforward for data downloading through HTML links. The only requirement is that the dynamic HTML link of the file to be downloaded can be generated. This is easy by using server-side programming technology like Java servlet [2] or JSP [3]. Once HTML link for the file to be downloaded is generated, user can click the link to download it.

These web-based tools can overcome the security limitations in the portal and give users easy ways to manager their jobs and files.

6 Implementations

Using the technologies in this paper, an HPC application services providing portal has been set up in IHPC, Singapore. User subscribes the utility of it viva Internet. After the administrators in IHPC get the registration notification, they audit his backgrounds according to his registration information. Once he is verified as a legal user, the administrators can get his user account ready for access. Registered user can go to <http://www.e2-portal.com> and using the given user account and password to access his subscribed applications. An example is shown in Fig 7. The CFD engineer is using Hypermesh[®] remotely through web. But his application is running on the supercomputer in IHPC.

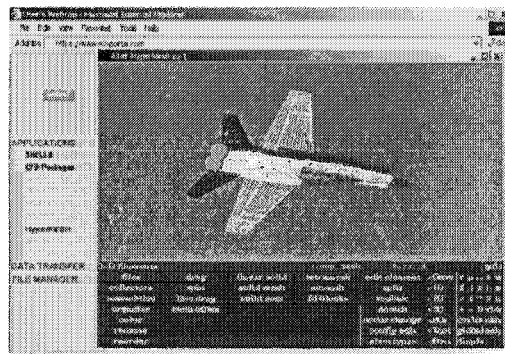


Figure 7. Running CFD application on web

7 Conclusions

The portal can provide the most sophisticated software applications in an online environment. Simulation, analysis and CAD tools essential to the engineering community are provided on the portal. Instead of having to purchase new equipment to handle the immense power of these applications, users can run these applications anytime, from almost any computer connected to the Internet. Instead of tying up their own computer for days or even weeks of computations, their jobs run on our servers. After computation, users can also ask experts in IHPC for advice to ensure the results, whatever it may be, is done correctly.

The advantages of the E-Engineering portal are numerous. For one, the software is more accessible because there is no large initial expense and therefore probably no need to submit lengthy proposals to upper management, followed by subsequent delays. Afterwards, less time is spent setting up, upgrading and maintaining the software as opposed to traditional methods. Finally, training is minimized in many cases.

Any way, the work is still at the starting point of the on-line application accessing. The work here can only enable users to access hosted applications at any time, any place on any device. Internet is a powerful tool for engineering community. It makes information sharing extremely easier and more available. Now, what the portal need is online real-time design, analysis information sharing. To realize this ambition, an on-line collaboration platform [10] is needed. This feature is go to be added to the HPC services providing portal.

References

- [1] Andrew, P., Douglas J.K.M. & Donald F.W., *High performance computing systems and applications*, Kluwer Academic: Boston, 2000.
- [2] Jason, H. & William, C., *Java servlet programming*, O'Reilly & Associates: Beijing and Cambridge, 1998.
- [3] Hans, B., *Java server pages*, O'Reilly: Beijing and Cambridge, Mass., 2001.
- [4] Nancy, A.B. and Tzvi, R., *Expert systems*, Industrial Engineering and Management: Atlanta, Ga., 1988.
- [5] MSC.Patran, MSC.Products, <http://www.mscsoftware.com/products/>, 2002.
- [6] Dieter, G., *Computer security*, Wiley: Chichester & New York, c1999.
- [7] The Apache HTTP Server Project, <http://httpd.apache.org/>, 2002.
- [8] Apache Tomcat, The Jakarta Site, <http://jakarta.apache.org/tomcat/>, 2002.
- [9] Tarantella, Products, <http://www.tarantella.com/products/>, 2002.
- [10] Dandavate, U., Steiner, D. & William, C., Working Anywhere: Co-design Through Participation, *Proc. of CoDesigning 2000*, eds. Stephen A.R. Scrivener, Linden J. Ball & Andrée Woodcock, Springer: London, New York, pp.101–110, c2000.