A Complex Engineering Design Problem Solving Environment Grid Portal

Hee-Khiang Ng, Yew-Soon Ong, Bu-Sung Lee, Ming Tang
Nanynang Technological University, Nanyang Avenue , Singapore
mhkng@ntu.edu.sg, asysong@ntu.edu.sg,ebslee@ntu.edu.sg,mtang@pmail.ntu.edu.sg

Terence Hung
Institute of High Performance Computing,, Singapore Science Park II, Singapore 117528
terence@ihpc.a-star.edu.sg

Abstract: A Complex Engineering Design Problem Solving Environment (CED PSE) portal targeted at computational science and engineering (CSE) users is described in this paper. The portal offers engineers and scientists with seamless access to the PSE's range of domain specific computational tools, information and resources acting as a single point of access to the Complex Engineering Design support environment. It effectively uses the grid as a vast and powerful computation engine that extends the engineers and scientist's desktop to heterogeneous resources. In particular, we focus on providing a standard computational workflow, seamless integration of CED tools and provide visualization tools to support the engineers and scientists throughout the entire CED design and analysis life cycle.

1 INTRODUCTION

Over the last decades, the process of design in engineering and science has been transformed by the introduction of massive computing power and advances in Internet technology, computational sciences and intelligence. There has been a move away from paper-based systems towards computer modeling and simulations that involve the use of highly sophistical computer-aided design packages, high-fidelity analysis packages, meshing packages, complex optimization tools and others. The move has led us to the development of increasingly advanced products such as the space telescope, large airliners and space shuttles. Nevertheless, this recent move towards computer modeling and simulations bring with it the need to access more complex and specialized software libraries, massive or rare data sets, and the unique contributions of their colleagues in other disciplines, locations, and organizations. Those engaged in design research are focusing on these issues and how they can be taken forward. Recently, there has been a 'paradigm shift' in engineering design systems towards complex engineering design problem solving environment (CED PSE), a new approach in working with science and engineering design problems. PSEs are defined as problem oriented computing environments that support the entire assortment of scientific computational problem-solving activities ranging from problem formulation to algorithm selection to simulation execution to solution visualization.

In CED, the engineers use computer aided design (CAD) tools to perform the design and evaluates this via numerical simulation. He/She then enters a design-evaluate-redesign process until the design is satisfied based on his/her knowledge and experiences. A CED PSEs should provide an environment that simplifies access and the integration of heterogeneous mix of analysis and design resources that will assist and support collaborations among engineers and scientist in solving the design problems they have to undertake. It must be configurable by the design team to suit each new task or suit the personal preferences of each individual designers and scientists. The CED PSE must be able to handle all the requirements of a design team and integrate between the stages of the entire design life cycle [1].

To provide a standard entry point and seamless access to the sea of services and resources provided for engineers and scientists in science and engineering design, a CED PSE Grid portal [2] is proposed in this paper. A web-based portal will provide engineers and scientists access to information and computing resources via a standard browser. A CED PSE Grid portal will provide more than standard portal services including job management, file management and resource management across the distributed grid resources.

The recent advances in Grid [2] fuelled the research and development of grid portals activities rapidly in recent years, with two major categories of portal architecture design, namely Generic portal and Domain Specific Problem Solving Environment Portal. A generic grid portal provides users simple job submission, job monitoring, file management and resource selection to a grid of heterogeneous distributed resources. The NPACI Hot Page [3], IPG Launch Pad [4], and the ClearingHouse [5] portal in the Nanyang Technological University are showcases of generic grid portals. PSE portal takes a variety of architectures, from a centralized web server, like the Ecce[6], Cactus Portal [7] or the XCAT Science Portal [8], which integrated the user's desktop/laptop with a built-in server. This emerging Grid technology intends to enable collaborative computing of an unprecedented scale by leveraging compute resources, storage devices, instrumentations, high bandwidth network infrastructure as well as visualization systems that spans across geographically distributed locations and organizations. The benefits are numerous: the ability to tap on more compute power and to achieve better utilization of resources; and the possibility of establishing cross-border research collaboration in real-time.

In this paper, our key focus is the proposal of the CED PSE Grid Portal, a novel e-design approach for engineers and scientists to tackle complex design problem. We believe it is an interesting challenge to provide a web-based PSE Grid portal catered to the specific needs of CED domain. We identify the components required to build a CED PSE, outlining the different types of services that we envisage to be part of the problem solving environment. This paper is organized as follows: Section 2 describes and outlines the typical complex engineering design life cycle. This entails various activities ranging from computer-based modeling, meshing, simulation, optimization and visualization. Section 3 presents the proposed CED PSE portal design while section 4 summarizes the main conclusions.

2 A TYPICAL COMPLEX ENGINEERING DESIGN LIFE CYCLE

The workflow of a typical CED life cycle is illustrated in Figure 1. It involves the processes of Computer Aided Design (CAD) modeling, Mesh generation, Quality check and healing, Load and boundary condition specification (LBC), Numerical analysis, Optimization, and Visualization. The CED life cycle is a process of a sequence of events that is dealt with often in series. This sequence is carried out more than once, and iterated through numerous refinement and investigations before finally entering the manufacturing phase. Usually the life cycle begins with the modeling of the product using CAD based tools, followed by numerical simulation to provide the design exploration process. Often preprocessing, numerical analysis, simulation and optimization consume and take up 80% of the entire CED life cycle. It is hence imperative for us to explore automated and distributed methods to speed up these processes. One aspect of preprocessing is related to meshing. In order to compute salient data (e.g. temperature, airflow, stress etc.) in non-continuous form, a mesh is needed to determine the coordinates at which the data is to be computed. Meshing is a tedious process that could take up a significant portion of the entire CED life-cycle. In addition, geometry quality check is done to determine if the geometry is 2D or 3D conforming. The check for 3D is very stringent. This check is needed to avoid problems during the numerical analysis stage. Automatic mesh generation tool takes in a geometry specification and automatically generates the desired mesh. Mesh could be 2D or 3D (surface and volume) depending on the requirement of the applications. There could be imperfection or problems during the automatic generation of the mesh. In this case, mesh healing technique is applied to repair the mesh.

Load and Boundary Condition Specifications is a simple user interface (UI) that allows for the specification of LBC data in various international standards such as Initial Graphics Exchange Specification (IGES) or Neutral File Format (NFF). Once the mesh and LBC data are ready, the next stage is numerical analysis. Here, various numerical techniques can be used to solve for the attributes under study. For instance, the analysis could solve for temperature profile in a building, stress coefficients in a structure or pressure on a surface. Typically, the numerical analysis process will output volume data or cost function values for further investigation and analysis. For example, volume data could be visualized whereas cost function can be used to drive iterative optimization work. In the optimization process, the cost function is use by the designers to aid decisions making on how to change the design, and this is often computationally intensive

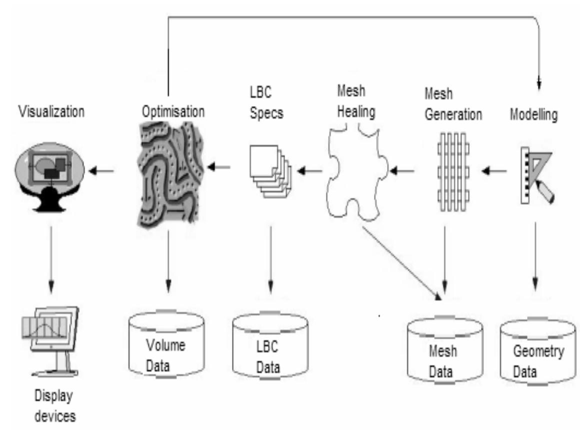

Figure 1 CED Life Cycle

Throughout the entire CED life cycle, many forms of data and files are created at different stages. The various design data that are produced and consumed at the different stages of the life cycle are also presented in Figure 1. These consist of the Geometry data (typically STEP or IGES format), Mesh data (typically IGES or NFF), Load and boundary condition data, and Volume data (e.g. temperature profile, airflow patterns etc.). The amount of simulation data could range from MB to GB in size and hence manual analysis by designers and scientists may not be practical. Therefore, visualization techniques are often used to present the large complex data into human-readable form. Thus, by using 2D and/or 3D graphical representation, engineers can gain significant insight into their domains, allowing them to better understand scientific and engineering phenomena. Furthermore, these data can also be archived for subsequent designs or designs process. Using computational intelligence and data mining techniques, constructive and vital information maybe learnt from these design archival data, which can assist the designers and scientists in future designs tasks.

There are a number of key issues in the CED life cycle that remains to be addressed. For example, it is possible to provide secure and seamless access to design packages, high-fidelity analysis packages, meshing packages, complex optimization tools and others, hence fully reducing the runtime incurred in these stages. It is clear that we must seek radical new ways to structure, and deploy the present CED life cycle so that the engineers can produce improve complex designs in faster time-scales and effective use of the advancing technology. To improve the CED PSE design process, in the subsequent sections we proposed with such an objective in mind some of the essential components in the CED PSE Grid portal.

3 PROPOSED CED PSE GRID PORTAL ARCHITECTURE

Figure 2 illustrates an overview of the portal. The portal is proposed to be java-based or via the use of Jetspeed framework, running on top of the Tomcat engine. It may be represented as a simple three-tier client server architecture containing three core services, namely Grid, Application and Database services.

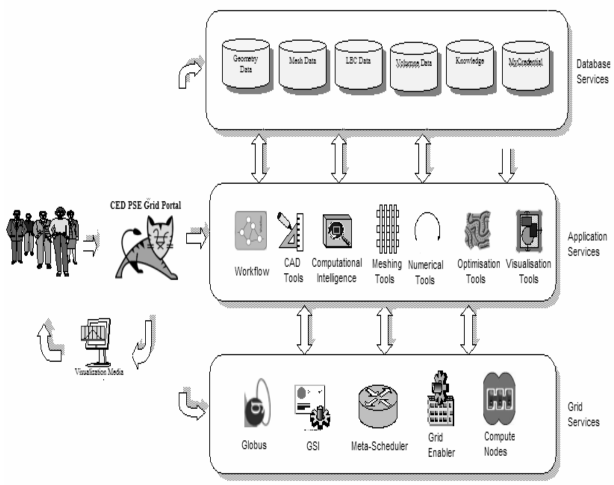

Figure 2 Architecture Overview

3.1 Grid Services (GS) Tier

The proposed CED PSE Grid portal Grid Services (GS) tier contains the basic setup of a grid environment. These may include the Globus middleware, grid security handling, meta scheduling, discovery, execution and GridRPC interfaces. Globus middleware provides the necessary interfaces to the Grid services and underlying compute resources. It provides a standard interface to manage and coordinate grid resources and activities. Using Globus Metacomputing Directory Services (MDS), we provide discovery and management of all resources such as grid compute nodes and software resources. We also identified Globus GridFTP to provide the seamless file transfer management in the grid resources necessary for the portal. To address single sign-on, interoperability with the local resource policies, the Grid Security Infrastructure Group (GSI) [9] is used. The GSI provides the underlying security architecture for the grid portal. In order to use any Grid resources, engineers and scientists only need to possess GSI credentials. This involves first the generation of user keys and user certificates signed by a trusted Certificate Authority (CA), and storing the user keys and certificates in the MyCredential repository. Next, the GSI credentials are created using these keys and certificates stored in the portal repository. Hence, we relieve the engineers from having to keep track of their GSI credential by offering to maintain their credential on their behalf.

In the present CED life cycle, it is increasingly obvious that one intuitive way to speed up the design cycle time significantly is via the mechanism of massively parallel systems. It is possible to incorporate parallelism in the preprocessing, analysis and optimization phases of the CED. The aggregation of computational facilities from various Grid initiatives will have thousands of compute nodes that would require powerful load mastering software to be developed. Hence, to attain immense parallelism and seamless access in the Grid environment, Metaschedulers will be provided to perform load balancing and soft/hard resources discovery for example, the discovery of compute clusters/nodes or specialized libraries required to carry out designing operation. The metascheduler is implemented as an application scheduler and has the responsibility to minimize execution time of CED applications. It finds suitable resources that span multiple clusters across geographical locations and schedules client's request. Further, it retrieves dynamic information about the grid and achieves effective mapping of requirements to resources.

Job submission or Grid Remote Procedure Calls will go through a standard interface to the metascheduler. Presently, majority of the operational scientific libraries and numerical solvers were developed using traditional languages. Hence one key task in the CED PSE portal is nonetheless to provide scientists an unified programming environment that can easily Grid-enable existing codes. Most existing Finite Element and Navier Stokes Computational Fluid Dynamic solvers, Optimization Packages used in structures, MEMS, antenna, optics and electromagnetic detailed design processes, were developed using traditional languages like Fortran and C. Much of its power with these languages stemmed from its ability to transparently encode scientific calculations, i.e. encode them in a way that exposes the underlying science to the inspection of the programmer and allows numerical solvers to be applied efficiently. Moreover, these high-quality libraries for solving design problems are often carefully tested and hence should be reused seamless within the Grid environment rather then turning to any form of re-venting of the wheels. Any requirements of involvement with low level programming tools such as CORBA , SOAP, XML and MPI by the scientists would be a real nightmare. Motivated by the ease of GridRPC [10] and aforementioned gaps of existing systems, we propose *a Unified GridRPC framework* as the middleware to Grid resources so to enable users to provide and to consume services easily in our Design Grid PSE. The framework is based on a provider and consumer architecture that implements scientific applications as web-services. GridRPC, which has recently been proposed by Grid Forum's Advanced Programming Models research group of the Global Grid Forum is seen as the standard Remote Procedure Call programming models, hence offering scientists and designers seamless access to remote computational resources through several interfaces for C, Fortran, Mathlab, and Mathematica. In addition, it promises to be effectively built on future Grid software based on Web services such as OGSA. As part of this project, we will help lead the Grid Community effort to create a standard client and to do a robust reference implementation of it. By building a *Unified GridRPC Framework*, we hope to assure that the community will be investing in a technology that will exhibit reasonable longevity.

3.2 Application Services (AS) Tier

The Application services tier includes the CED PSE workflow process and domain specific tools, such as CAD, meshing, simulation, optimizations and visualization tools. The application services are deployed onto the CED PSE Grid portal to provide a web-based interface to the various applications.

The workflow engine is designed to specify, execute, monitor and coordinate the flow of the sequential design process. This allows the engineer to formally represent and visualize the workflow processes using the graph-editing component of the workflow engine. In addition, the engineers can also retrieve workflow pre-defined template and modify it to suit personal needs. To provide personalized tracking capability, information about the design process for each engineer is organized and stored in the database. Furthermore, customization of design workflow and CED tools based on users' preferences are also provided as part of the CED management.

The Application services act as the service broker to CED applications, launching the CED tools. In the CED PSE Grid Portal, we have a collection of application sub grids, namely Computational Intelligence sub grid, CAD sub grid, Meshing sub grid, Optimization sub grid, Visualization sub grid, and Numerical sub grid. For example, the optimization sub grid provides a pool of optimization resources such as I-SIGHT and OPTIONS[11] that contains many diverse optimization methods for which the engineer can invoke. The meta scheduler discovers the resources, authenticated users based on GSI single sign-on, mapped the requirements to resources and invokes the requested CED tools.

Features that cater to engineers needs to add and upload new CED tools is also proposed. If necessary, these tools are transferred to the compute resources for compilation to generate the executables. Gridmake may be used in the CED PSE Grid portal to coordinate and manages this process. Furthermore, the newly supplied CED tools will be registered, and make available as new tools to the authorized consumers.

In CED PSE, large simulation and design data are produced repeatedly. Data of such scale is practically not amenable to manual analysis. Thus, in order to facilitate and enhance quick and effective analysis of scientific data, graphical presentation is desired. In this framework, we propose an interactive visualization environment where geometry data can be combined with volume data and where users can interact and navigate within the geometry space to explore the scientific data. To achieve this, we devised a "visualization container" that consists of a generic visualization program that can take in generic geometry and volume data in specific formats. The container provides a user-friendly interface for users to interact with the visualization services. For presentation to the users, the output of the visualization container is piped into SGI's VizServer software, which subsequently delivers pixels to the users via a light-weight client GUI. This approach pushes the visualization processing to a more powerful high end server while the client device can be freed up to do other work. Another benefit is the avoidance of large data file transfer from the compute server to the client machine. In this way, only data pertaining to the view area is being transferred (in pixel form) to the client.

3.3 Database Services (DS) Tier

A typical CED life cycle involves many different forms of data and files. Hence, database service management of the portal data is thus essential. Here, the database services houses and manages all the data that pertains to the entire CED life cycle. At each stage of the CED cycle, different data such as geometry, mesh, LBC, volume data are generated. All these data are stored into the respective databases to enable task checkpoint, and hence capability to resume the task at the defined stage of the CED workflow. In addition, during the design stage, all these results from the computation can be stored in the knowledge database. The knowledge database provides these data to the Computational Intelligence to seed future design process and allow more informed choices about appropriate design strategies and hence greatly improved the future computational efforts. In addition, reporting tools such as crystal report can be used to generate useful reports from the pool of data. Besides CED data related management, DS also handle the portal security data, known as MyCredential repository. The repository managed all the user keys and certificates, which are retrieved to generate the GSI credential necessary for accessing the Grid resources. Another database services will be the AWS service repository, which maintains the AWS WSDL information and server URLs.

4 CONCLUSION

In this paper, we described a CED PSE portal to assist engineers and scientists to help them solve the design tasks they have to undertake. At the same time it allows the engineers more time for more qualitative task of complex engineering design instead of tackling the complexity of advance technology. We concentrated on providing a standard computational workflow and processes life cycle in the proposed CED problem solving environment, where analysis tasks can search and uses whatever available grid resources. Here, a number of independent analysis tasks can be run in parallel over one or more linked computational clusters. The results are then stored in database for future analysis and assist the designers in future design task. One chain of CED concern for the PSE Grid Portal is to achieve the virtualization and sharing of complex engineering design domain specific tools and applications. Thus, the engineers will be able to access and uses the tools for their problem solving, without regards to the location of these tools and applications. This allows a spectrum of tools to be used in a flexible manner, depending on the engineers' needs. In the future, we will like to see the CED PSE Grid portal evolves towards a e-Design market for CED.

5 REFERENCES

[1] Gregor von Laszewski, Ian Foster, Jarek Gawor, Peter Lane, Nell Rehn, Mike Russell. "Designing Grid-based Problem Solving Envionments and Portals," 34th Annual Hawaii International Conference on System Sciences (HICSS-34)-Volume 9 January 03 - 06, 2001.

[2] Foster and C. Kesselman, editors, "The Grid: Blueprint for a New Computing Infrastructure," Morgan Kaufman Publishers, 1999.

[3] NAPCI Hot Page, http://hotpage.npaci.edu.

[4] NASA IPG Launch Pad Portal, http://www.ipg.nasa.gov

[5] Nanyang ClearingHouse Portal, http://ntu-cg.ntu.edu.sg/ClearingHouse

[6] Karen Schuchardt, Brett Didier, Gary Black, "Ecce-A Problem Solving Environment's Evolution Toward Grid Services and a Web Architecture," Concurrency and Computation: Practice and Experience, Volume 14, Number 13-15, 2002

[7] Albert Einstein Institue, "Cactus," http://www.cactuscode.org.

[8] Sriram Krishnan, Randall Bramely, Dennis Gannon, Madhusudhan Govindaraju, Rahul Indurka, Aleksander Slominski, Benjamin Temko. "The XCAT Science Portal," Proceedings of SC2001.

[9] Ian Foster, Carl Kesselman, Gene Tsudik, Steven Tuecke, "A Security Architecture for Computational Grids," Proc. 5th ACM Conference on Computer and Communications Security Conference*, pp. 83-92, 1998*

[10] H Nakada, S. Matsuoka, K Seymour, J Dongarra, C Lee, H Casanova, "GridRPC: A Remote Procedure Call API for Grid Computing," Grid Computing - Grid 2002, LNCS 2536*, pp. 274--278, November 2002*

[11] OPTIONS: Design Exploration System, http://www.soton.ac.uk/~ajk/options/welcome.html