Answers to the Enterprise Architect Magazine Query

1. What is grid architecture?

Let me start off with defining a Grid. However, it seems to me that various researchers have differing views on Grid computing mostly based on technologies or applications that they are developing and what they envision it to be. So, do I! My definition is as follows: “Grid is a type of parallel and distributed system that enables the sharing, selection, and aggregation of services of heterogeneous resources distributed across "multiple" administrative domains based on their availability, capability, performance, cost, and users' quality-of-service requirements”.

Like any distributed system, Grids need to address various issues and challenges including: security; autonomy; heterogeneity of resource access interfaces, policies, capability, pricing; data locality, dynamic variation in availability of resources, and complexity in creation of applications. Therefore, Grid follows a combination of hierarchical and decentralized architecture for resource management; and a layered architecture for implementation of various services.

2. How is it different from cluster computing? and from single system parallel systems?

A cluster is made up of multiple interconnected independent nodes that co-operatively work together as a single unified resource. Unlike Grids, cluster resources are owned by a single organisation and they are managed by a centralized resource management and scheduling system. That means all users of clusters have to go through a centralized system that manages allocation of resources to application jobs.

Actually, many Grids are constructed by using clusters or traditional parallel systems as their nodes. For example, the World-Wide Grid, used in evaluating the Gridbus technologies and applications, has many nodes that are clusters, which are located in organisations such as AIST-Japan, N*Grid Korea, University of Melbourne, and NRC Canada. Another example of Grid that contains clusters as its nodes is the NSF TeraGrid in the US.

3. How does it differ from or work with Web services?

Web services provide standard infrastructure for data exchange between two different distributed applications whereas, Grids provide an infrastructure for aggregation of high-end resources for solving large-scale problems in science, engineering, and commerce.

The recent trend is to implement Grid solutions using Web services technologies. For example, Globus 3.0 version is being implemented using Web services technologies. Within our Gridbus Project, we have implemented Grid technologies such as Grid Bank and Grid Market Directory using Web services technologies. In this context, we can safely say that low-level Grid services are instances of Web services.

4. Your research involves resource management and scheduling in a grid environment. Under what circumstances would an enterprise architect need to worry about these topics? Is there a certain system size or user community size at which resource management becomes a significant project?

When enterprises need to aggregate resources distributed within their organisation and prioritize allocation of resources to different users, projects, and applications based on their QoS (Quality of Service) requirements, they need to be concerned about resource management and scheduling. The user QoS driven allocation strategies enhance the value delivered by the utility.

The need for QoS based resource management becomes significant whenever more than one competing applications or users need to utilize shared resources.

5. Does Nimrod-G / Gridbus work with any grid solution or only specific solutions?

Yes, Nimrod-G developed at Monash University works with resources that are Grid enabled by low-level middleware systems such as Globus and Legion. Within the Gridbus project at the University of Melbourne, we are developing new resource allocation strategies for scheduling distributed data oriented applications on Global Grids and extending the grid economy paradigm to provide comprehensive support for service-oriented computing.

6. Can you give examples of existing solutions (Globus/IBM, United Devices, Data Synapse, etc.) that illustrate the utility and broker (or other) models? Are there certain applications that work better with each model?

Solutions provided by companies such as United Devices and Data Synapse act as brokers between applications and computational resources (desktop PCs) at the edge-of the network. Their architecture philosophy is hybrid as they follow both cluster and peer-to-peer computing model for resource management.

7. Do you have any comment on the grid solutions IBM announced in January and their partnerships with many of the most prominent peer-to-peer players?

The partnership between IBM and peer-to-peer players helps both parties. It enhances IBM's ability to provide high-performance computing solutions to low-budget enterprise customers. In addition, P2P players benefit from the high-profile that IBM offers, which is essential for their survival in tough economic times.

8. Do you have any comment on Gateway's alliance with United Devices and selling cycles from their demonstration PCs?

It is a great way of leverage economic potential of CPU cycles that otherwise would get wasted when the PCs are idle.

9. Do you have any comment on United Online's (Juno) changes to their user agreement permitting them to farm cycles from users' PCs?

It is a great way of helping customers—as it provides an economic incentive or discount on Juno services for sharing CPU cycles.