GRID ARCHITECTURE AND RELATIONSHIP

It is a known fact that in the technology of art that there are numerous well-defined and well-established technologies and standards developed for distributed computing. This foundation has been a huge success (to some extent) until we entered into the domain of heterogeneous resource sharing and the formation of virtual organizations.

Based on our previous discussions, grid architectures are defined as a coordinated, highly automated, and dynamic sharing of resources for a virtual organization. It is appropriate that we turn our attention at this stage toward the discussion regarding how these architecture approaches differ from the prior art of distributed technologies, that is, how the two approaches compliment each other, and how we can leverage the best practices from both approaches.

Our discussion will now begin to explore notions of the widely implemented distributed systems, including World Wide Web environments, application and storage service providers, distributed computing systems, peer-to-peer computing systems, and clustering types of systems.

World Wide Web

A number of open and ubiquitous technologies are defined for the World Wide Web (TCP, HTTP, SOAP, XML) that in turn makes the Web a suitable candidate for the construction of the virtual organizations. However, as of now, the Web is defined as a browser–server messaging exchange model, and lacks the more complex interaction models required for a realistic virtual organization.

As an example, some of these areas of concern include single-sign-on, delegation of authority, complex authentication mechanisms, and event correlation mechanisms. Once this browser-to-server interaction matures, the Web will be suitable for the construction of grid portals to support multiple virtual organizations. This will be possible because the basic platforms, fabric layers, and networking connectivity layers of technologies will remain the same.

Distributed Computing Systems

The major distributed technologies including CORBA, J2EE, and DCOM are well suited for distributed computing applications; however, these do not provide a suitable platform for sharing of resources among the members of the virtual organization. Some of the notable drawbacks include resource discovery across virtual participants, collaborative and declarative security, dynamic construction of a virtual organization, and the scale factor involved in potential resource-sharing environments.

Another major drawback in distributed computing systems involves the lack of interoperability among these technology protocols. However, even with these perceived drawbacks, some of these distributed technologies have attracted considerable Grid Computing research attention toward the construction of grid systems, the most notable of which is Java JINI.[1] This system, JINI, is focused on a platform-independent infrastructure to deliver services and mobile code in order to enable easier interaction with clients through service discovery, negotiation, and leasing.

Application and Storage Service Providers

Application and storage service providers normally outsource their business and scientific applications and services, as well as very high-speed storage solutions, to customers outside their organizations. Customers negotiate with these highly effective service providers on QoS requirements (i.e., hardware, software, and network combinations) and pricing (i.e., utility-based, fixed, or other pricing options).

Normally speaking, these types of advanced services arrangements are executed over some type of virtual private network (VPN), or dedicated line, by narrowing the domain of security and event interactions. This is oftentimes somewhat limited in scope, while the VPN or private line is very static in nature. This, in turn, reduces the visibility of the service provider to a lower and fixed scale, with the lack of complex resource sharing among heterogeneous systems and interdomain networking service interactions.

This being said, the introduction of the Grid Computing principles related to resource sharing across virtual organizations, along with the construction of virtual organizations yielding interdomain participation, will alter this situation. Specifically, this will enhance this utility model of application service providers and storage service providers (ASP/SSP) to a more flexible and mature value proposition.

Peer-to-Peer Computing Systems

Similar to Grid Computing, peer-to-peer (P2P) computing is a relatively new computing discipline in the realm of distributed computing. Both P2P and distributed computing are focused on resource sharing, and are now widely utilized throughout the world by home, commercial, and scientific markets. Some of the major P2P systems are SETI@home[2] and file sharing system environments (e.g., Napster, Kazaa, Morpheus, and Gnutella).

The major difference between Grid Computing and P2P computing is centered on the following notable points:

1. They differ in their target communities. Grid communities can be small with regard to number of users, yet will yield a greater applications focus with a higher level of security requirements and application integrity. On the other hand, the P2P systems define collaboration among a larger number of individuals and/or organizations, with a limited set of security requirements and a less complex resource-sharing topology.
2. The grid systems deal with more complex, more powerful, more diverse, and a highly interconnected set of resources than that of the P2P environments.

The convergence of these areas toward Grid Computing is highly probable since each of the disciplines are dealing with the same problem of resource sharing among the participants in a virtual organization. There has been some work, to date, in the Global Grid Forum (GGF) focused on the merger of these complimentary technologies for the interests of integrating the larger audience.

Cluster Computing

Clusters are local to the domain and constructed to solve inadequate computing power. It is related to the pooling of computational resources to provide more computing power by parallel execution of the workload. Clusters are limited in scope with dedicated functionality and local to the domain, and are not suitable for resource sharing among participants from different domains. The nodes in a cluster are centrally controlled and the cluster manager is aware of the state of the node. This forms only a subset of the grid principle of more widely available, intra/interdomain, communication, and resource sharing.