GRID INFRASTRUCTURE

The grid infrastructure forms the core foundation for successful grid applications. This infrastructure is a complex combination of a number of capabilities and resources identified for the specific problem and environment being addressed.

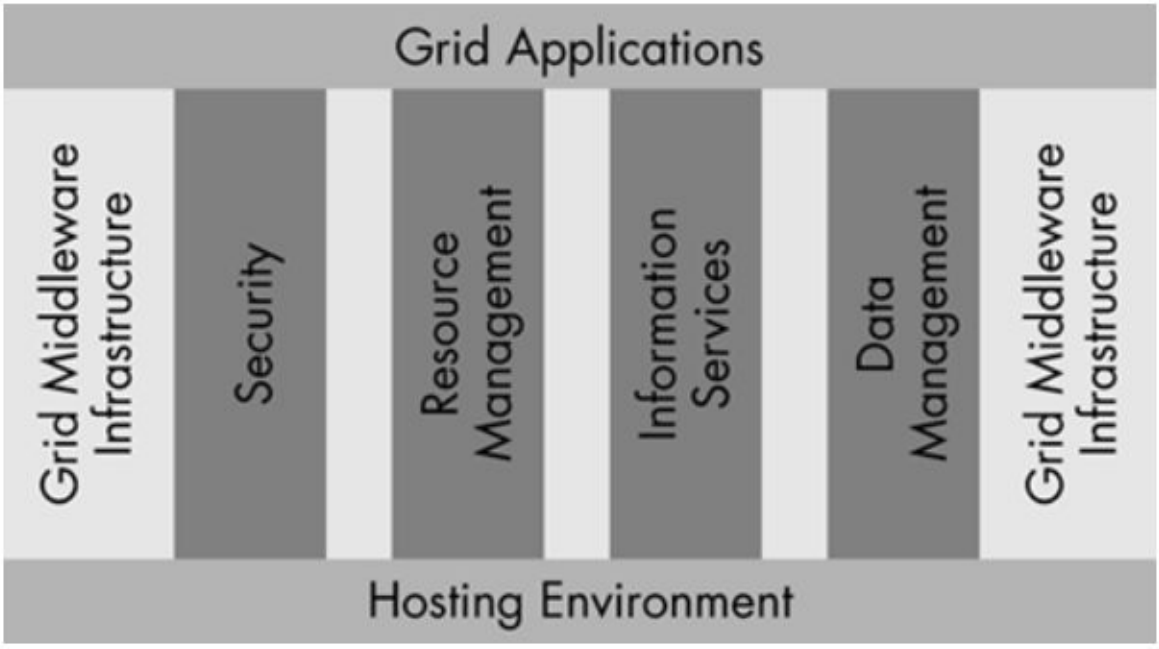

Figure 1.4. Grid middleware topic areas are becoming more sophisticated at an aggressive rate.

In general, a Grid Computing infrastructure component must address several potentially complicated areas in many stages of the implementation. These areas are:

- Security

- Resource management
- Information services
- Data management

Let us further examine the significance of each of these above components.

Security

The heterogeneous nature of resources and their differing security policies are complicated and complex in the security schemes of a Grid Computing environment. These computing resources are hosted in differing security domains and heterogeneous platforms. Simply speaking, our middleware solutions must address local security integration, secure identity mapping, secure access/authentication, secure federation, and trust management.

The other security requirements are often centered on the topics of data integrity, confidentiality, and information privacy. The Grid Computing data exchange must be protected using secure communication channels, including SSL/TLS and oftentimes in combination with secure message exchange mechanisms such as WS-Security. The most notable security infrastructure used for securing grid is the Grid Security Infrastructure (GSI). In most cases, GSI provides capabilities for single sign-on, heterogeneous platform integration and secure resource access/authentication.

The latest and most notable security solution is the use of WS-Security standards. This mechanism provides message-level, end-to-end security needed for complex and interoperable secure solutions. In the coming years we will see a number of secure grid environments using a combination of GSI and WS-Security mechanisms for secure message exchanges. We will discuss the details of security mechanisms provided by these standards later in this book.

Resource Management

The tremendously large number and the heterogeneous potential of Grid Computing resources causes the resource management challenge to be a significant effort topic in Grid Computing environments. These resource management scenarios often include resource discovery, resource inventories, fault isolation, resource provisioning, resource monitoring, a variety of autonomic capabilities,[5] and service-level management activities. The most

interesting aspect of the resource management area is the selection of the correct resource from the grid resource pool, based on the service-level requirements, and then to efficiently provision them to facilitate user needs.

Information Services

Information services are fundamentally concentrated on providing valuable information respective to the Grid Computing infrastructure resources. These services leverage and entirely depend on the providers of information such as resource availability, capacity, and utilization, just to name a few. This information is valuable and mandatory feedback respective to the resources managers discussed earlier in this chapter. These information services enable service providers to most efficiently allocate resources for the variety of very specific tasks related to the Grid Computing infrastructure solution.

Data Management

Data forms the single most important asset in a Grid Computing system. This data may be input into the resource, and the results from the resource on the execution of a specific task. If the infrastructure is not designed properly, the data movement in a geographically distributed system can quickly cause scalability problems. It is well understood that the data must be near to the computation where it is used. This data movement in any Grid Computing environment requires absolutely secure data transfers, both to and from the respective resources. The current advances surrounding data management are tightly focusing on virtualized data storage mechanisms, such as storage area networks (SAN), network file systems, dedicated storage servers, and virtual databases. These virtualization mechanisms in data storage solutions and common access mechanisms (e.g., relational SQLs, Web services, etc.) help developers and providers to design data management concepts into the Grid Computing infrastructure with much more flexibility than traditional approaches.