Large distributed systems such as Computational and Data Grids require that a substantial amount of monitoring data be collected for various tasks such as fault detection, performance analysis, performance tuning, performance prediction, and scheduling. Many techniques are currently available and others are being developed for collecting and forwarding this data. Currently available architectures implement centralized ontology based knowledge management systems for resource utilization. There is a need to develop a more flexible, scalable and more important is lightweight middleware architecture for efficient resource utilization in dynamic semantic Grid environment. In this paper we have proposed a distributed ontology based knowledge management architecture for resource utilization in Grid monitoring middleware. We have used a Consumer/Producers model, which is able to combine information and publish it. The major component of the system is the “selector” which is able to find the best knowledge base to get specific monitoring data. We suggest a distributed approach to Knowledge Management based on ontology for fault tolerance. Using multi-agent system we have designed architecture for negotiating policies and authorization in Grid resource utilization. We are carrying out the implementation of the proposed system. Our preliminary results show that despite some performance constraints in our implementation, the concept can lead toward resource management in adaptive and dynamic systems like semantic Grid. We argue that the proposed system will provide strong foundation that lead toward autonomic computing realization environment and semantic grid monitoring middleware.

**Key Words:** Computational Grid, Data Grid, Ontology, Semantics, Knowledge Management, multi-agent system