ABSTRACT

Grid is a dynamic framework that enables sharing and selection of loosely coupled heterogeneous resources over geographically distributed locations. Resource sharing and dynamic application execution environment provision are the primary challenges of a grid environment. Resource sharing is a serious issue in grid since the participating resources belong to different administrative domains. Besides, these resources have multiple usage allocation policies in different virtual organizations. These resource usage policies express the desired sharing rules and sharing relationships of the resource providers. Since these rules and relationships are framed based on the nature, locality and business hours of an organization, they should not be violated during the operation of grid. If they are violated, the resource providers may not contribute the accepted resource share and cannot maintain the accepted service level in other virtual organizations. But a mechanism to give assurance to the resource providers that their local resources are never overrun is still lacking in the literature. Hence, it is mandatory to express the resource usage policies. In addition, these policies must be integrated with the grid meta-scheduler to enforce the sharing rules and relationships. Also, it is necessary to guarantee that the service levels required by the applications are provided till the end of its execution. In this thesis, an approach called policy management system is proposed and implemented to manage the resource usage policies. Also, a service level agreement framework is proposed to create, monitor and enforce the accepted service levels in grid environment. Furthermore, the policy management system and the service level agreement
framework are integrated with our own grid meta-scheduler to realize a controlled grid resource sharing and its effects are studied.

The existing grid infrastructure is not flexible enough to provide the applications’ execution environments. The execution environments in grid resources cannot be modified dynamically as per the applications’ requirements. This makes potential grid resources to unfit for the application execution. In addition, the application execution in virtual organizations is not isolated from the physical organizations. Integration of virtualization with grid environment solves the above said problems as it creates the customized application execution environment over the existing grid and provides isolated application execution. This integration leads to a birth of additional scheduling scenarios since the resource management system has to consider both the physical and virtual resources during the selection of resources for applications execution. To handle these additional scheduling scenarios, this research work proposes deviation based resource scheduling algorithm to classify the resources against job requirements: overmatch (plug-in), exact-match (Exact) and under-match (subsume). If a job request has resources in exact or plug-in region then a single resource provider is enough to execute the job. If a job has resources only in subsume region, the job may not be run by any conventional meta-schedulers. This thesis explores this region and proposes new scheduling scenarios to use the resources in this region to run a job. This approach is simulated and examined against the job scheduling rate and scheduling time. From the simulation, it is observed that the proposed approach schedules more jobs than the conventional meta-scheduler with reduced scheduling time and improved throughput.