NAREGI/ Server Grid Resource Framework

Scalable middleware infrastructure for Server Grid will realize automated allocation of resources, scheduling of executing job, etc. Users in various research areas can run their HPC application programs as it is or slightly modified at federated large computing centers with the middleware deployed based on UNICORE, Globus, Condro.

NAREGI Middlewares

Grid-Enabled Nano-Applications
- Grid Programming
  - Grid RPC
  - Grid MPI

Grid Visualization
Grid PSE
Grid Workflow
Super Scheduler
Distributed Information Service
GridVM

(Globus, Condor, UNICORE → OGSA)

High-Performance & Secure Grid Networking

• Higher Throughput
• Higher Utilization
• Higher Reliability
• Higher Scalability

SuperSINET

Kyushu U
Kyoto U
IMS
NII
AIST, KEK
IMR, Tohoku U

Seamless resource coordination across many different research organizations.

Cross-Organizational Virtual Computer

National Research Grid Initiative
Scenario for Inter-site MPI Job Execution

Users can execute huge MPI jobs such as coupled simulations in nano-science, without being conscious of actual complicated configuration of the grid resources.

**Preliminaries**

A user has MPI programs for RISM-FMO coupled simulation.

Before its execution,

a. The user receives his/her own user certificate issued by the CA after user registration.

b. The user transfers and compiles his/her programs on the grid resources by using PSE.

c. The user describes the way of execution and the requirements of resources for the MPI job in a workflow by using WFT (Workflow Tool).

**Reservation**

1. The workflow is submitted to the Super Scheduler through WFT.

2. The Super Scheduler asks the Information Service to find adequate GridVMs for the MPI job described in the workflow.

3. The Super Scheduler negotiates with these GridVMs to reserve resources for job execution.

4. If these GridVMs agree with the Super Scheduler, they reserve resources under the Local Schedulers.

**Execution**

5. When the reservation time comes, an IMPI (Interoperable MPI) server is started.

6. After that, GridVMs co-allocate job queues with 64 CPUs on an SMP machine and with 128 CPUs on a PC cluster.

7. RISM is started on the SMP machine and FMO is started on the PC cluster.

8. IMPI server initiates and manages MPI communication.

**Monitoring**

9. The results of the MPI job can be visualized under and/or after the job execution by using Grid Visualization System.

10. After the job ends, its resource usage is recorded in Information Service.
RISM-FMO Coupled Simulation

Mediator-based grid middleware allows coupled simulations of nano applications to calculate the electronic structure of nano-scale molecule in solution.

RISM-FMO Coupled Simulation in Grid Environment

Simulation Result for Oligopeptide

Statistical high density in O molecular orientation of H$_2$O to the solute decreases electronic charge on H atom.

Electronic density of solute in FMO (in vacuum)

Solute in H$_2$O Solvent

Before SCF Calculation

NH at Cys

High density in O orientation

After SCF Calculation

Electronic charge on H atom

High density in O orientation

Electronic density of solute in FMO (in vacuum)
High Performance Grid Networking

Approach

High Performance Grid Networking group is researching and developing Grid Network Services, which work as a coordinator between network infrastructure and grid middleware. Grid Network Services allow grid middleware to reserve network resources and to provide network monitoring information. Network Resource Reservation Service controls network resources, such as bandwidth, priority flow control, packet shaping, based on network measurement results, job submission description information and WS-Agreement from Super Scheduler. Network Resource Information Service monitors network resources, such as utilization, round trip time, and provides network information for grid middleware in order to schedule and discover available grid resources.

Overview

![Diagram of Grid Resource discovery and Information Service]