NAREGI middleware consists of security services, portal, client tools, programming libraries, resource and job management services, grid file system, management tools, and so on. User can describe his/her process workflow as a workflow job with GUI based Work Flow Tool (WFT) and submit it to NAREGI resource and job management services. To make job submission and data handling secure, NAREGI middleware uses GSI based security architecture.

NAREGI middleware supports heterogeneous computing environments (CPUs, OSs, batch schedulers), automatic resource brokering, advanced reservation based co-allocation, sharing resources with local submitted batch jobs (non grid batch job), Virtual Organization (VO), etc.
NAREGI Middleware Ver. 1.0
Architecture

Features

**Resource and Job Management**
- Automatic resource brokering
- Advanced reservation based co-allocation
- Resource sharing between grid and local batch jobs
- Bulk job submission
- Management tools
- GridMPI, GridRPC support
- Interoperation with EGEE/gLite (prototype)

**Data Grid**
- Grid wide file sharing
- Import and export files to grid file system
- Workflow data staging

**Security**
- CA/RA
- Proxy certificate renewal
- Authorization policy enforcement

**User Environment**
- Single sign-on
- Application registration and deployment
- Workflow editor
- Command line interface
- Grid visualization
- GridMPI, GridRPC libraries
- Communication tools for multi physics and multi scale jobs

**VO**
- VO policy enforcement
- Group and role support

Supporting Standards

**Common**
- WSRF/WSN
- OGSA-DAI
- JSDL

**Resource and Job Management**
- OGF/OGSA-EMS
- DMTF-CIM
- OGF/OGSA-RUS

**Data Grid**
- OGF/GFS

**Security**
- X.509
- GSI based security
- XKMS

**User Environment**
- ACS
- MPI-1(full compliant), MPI-2(compliant)
- OGF/RPC

**Install tools**
- RPM
- APT
NAREGI Resource and Job Management Services

NAREGI resource and job management services integrate distributed heterogeneous computing resources in a grid as a virtualized single computing environment. NAREGI client modules, such as a Work Flow Tool (WFT), can use any computing resources in a grid, if the conditions of job and resource are meet.

NAREGI resource and job management services consist of Super Scheduler (SS), GridVM, and Information Service (IS). SS is a meta scheduler of NAREGI grid. It allocates computing resource by match making of job requirements and resource properties, and submits the job into GridVM on the allocated computing resource. GridVM is the resource and job manager on each computing resource. GridVM submit input job to batch scheduler. If the job requires resource co-allocation, SS make reservation requests to GridVMs before submit the jobs. IS is the information repository of NAREGI middleware. It always collects and manages resource information and accounting information of computing resources in a grid.

Features:

- OGF OGSA-EMS based architecture
- OGF JSDL based job description
- DMTF-CIM based resource information model
- OGF OGSA-RUS based accounting
- Automatic resource brokering and job scheduling
- Advanced reservation based co-allocation
- Bulk job submission
- Work flow engine
- VO and local policy based access control
- Resource sharing between grid and local batch jobs
- Grid interoperability with EGEE gLite (prototype)
Resource and Job Management Architecture

Super Scheduler

NAREGI Super Scheduler (SS) is meta scheduler of NAREGI grid environment. The major features of SS are job workflow management, resource brokering, and advanced reservation based co-allocation. The architecture of SS is based on OGF/OGSA-EMS. It consists of four major services as figure shows. Job Management Service (JM) has workflow engine in its inside and orchestrates all other services of SS to submit each job activity in the input workflow job. Candidate Set Generator (CSG) match makes job activity’s requirements and resource properties then generates a list of candidate resources. Execution Planning Service (EPS) decides computing resources according to resource candidates list from CSG. Reservation Service (RS) reserves resources, if the job activity requires advanced reservation, and submits the job activities.

GridVM

NAREGI GridVM manages resource and job activity in computer resources. It consists of scheduler and engine. GridVM scheduler is located on the top of batch scheduler in the management server of each computing site (cluster). The features of GridVM scheduler are job manager and advanced reservation extension for WS-GRAM, resource local policy enforcement, information provider of resource properties and job accounting information. It also works as light weight virtual machine to hide the differences of supported batch schedulers. GridVM engine is located in each computing resource. The features of engine are executable launcher and sand boxing security.

Information Service

Main feature of Information Service (IS) is the information repository of NAREGI grid environment. It holds resource information, job accounting information etc. Resource information is based on DMTF CIM schema and job accounting information based on UR/RUS. Registered information can be retrieved with OGSA/DAI grid standard. It will be used by not only NAREGI modules but also customer developed tools. To treat widely distributed grid resources, IS has layered architecture.
Grid Interoperation with EGEE/gLite

In near future, all grid middleware will interoperate with each other like many OSs of the day can communicate with other OSs with standard way. One of the goals of NAREGI middleware is realize grid interoperation. As the first step of grid interoperation, NAREGI project developed prototype of interoperation modules for EGEE/gLite in three areas, Grid Information, Data Management, and Job Submission. Since, both NAREGI middleware and EGEE/gLite use MyProxy and VOMS, they can use same proxy certificate.

- All of grid information can be retrieved by each of grid in its fashion WRT resource description schema, data format, query language, client API, ...
- Each information service in grid acts as an information provider for the other and translator embedded in the provider performs conversion between different schemas.
- NAREGI and EGEE gLite clients can access to both data resources (e.g., bi-directional file copy) using SRM interface.
- GridFTP is used as its underlying file transfer protocol.
- File catalog (metadata) exchange is planned.

- There are two goals for job submission interoperation.
  - Bilateral job submission with EGEE/gLite
  - Bilateral job submission with WS-GRAM based grids
- In the job submission between NAREGI and EGEE, the resource brokers can be used for the job requirements - resource matchmaking.
**VO Operations**

### Resource Sharing among VO Members

1. **Creation of a VO**
   - VO administrator sets up VOMS, Portal, Super Scheduler and Information Service to be used by VO members.
   - VO administrator assigns group/subgroups, differentiates role of the manager and registers the members to VOMS.
   - VO administrator applies for using computing resources as a local group to resource providers and the resource providers set local accounts of VO members, usage policy for the VO and mappings of local names to global names.

2. **Use of the Grid Resources**
   - A VO member signs on the Portal then VOMS server issues his proxy certificate with VO/group/role information.
   - PSE manages application programs shared by members of each VO/group and deploys them to computing resources.
   - SS retrieves information from IS and finds appropriate resources for execution of jobs submitted by VO members.
   - DataGrid transfers data files and GridVM manages submitted jobs and registers the resource usage records to IS.
   - Grid Operation Center will assist with operations and problem solving among VO and resource providers.
The NAREGI Data Grid Environment has been developed for enabling users to use distributed data resources (e.g., file servers, databases, experimental instruments, observatories, etc.) in the Grid environment. It consists of the Data Grid Resource Management System and the Data Transfer System.

- The Resource Management System provides data sharing services using a Grid File System that is a virtual file system with a global name space (i.e., the same path name independent of client and server locations) in the Grid environment. The Grid File System is based on the Gfarm (Grid Data Farm) file system and extended for having different file system views according to their VOs (Virtual Organizations).

- The Data Transfer System is in charge of file access and transfer operations. In cooperation with WFT (Work Flow Tool) and Super Scheduler, the Data Transfer System provides a data-staging functionality for transferring files between the Grid File System and computing nodes.
NAREGI middleware provides security services including authentication and authorization. NAREGI-CA is developed for Grid to realize certification services and deployed worldwide so far. Authorization is plugged in as a web service and used in the coordination with virtual organization management.

**NAREGI Middleware Architecture for Security**

- NAREGI middleware is based on the standard GSI and adopts VOMS as virtual organization management. New functions are available to extend proxy certificate validity periods and adapt authorization policies to managed resources.
  - Extension of proxy certificate validity periods
    - Renewal services of proxy certificates in accordance with job execution
  - Authorization services
    - Enable to adapt authorization policies based on VO by resource providers

**NAREGI Certification Service**

NAREGI Certification Authority is accredited by APGrid PMA, and a member of IGTF that realizes global trust relationships among Asia, Europe, and Americas.

NAREGI-CA is an open software to realize production-level CA operations, and installation records suggest its effectiveness.