High Performance and Secure Grid Network Middleware

July 21, 2004

Osaka University
Shinji Shimojo
NAREGI Software Stack

Grid-Enabled Nano-Applications

- Grid Programing
  - Grid RPC
  - Grid MPI

Grid Visualization

Grid PSE

Grid Workflow

Super Scheduler

Distributed Information Service

(Globus, Condor, UNICORE ➔ OGSA)

Grid VM

High-Performance & Secure Grid Networking

SuperSINET

NII

IMS

Research Organizations

etc

Computing Resources

National Research Grid Initiative
Overview of Research and Development

SuperSINET Grid Concept
“A Grid is a collection of distributed computing resources over network that appear to an user or an application as one large virtual computing system”

Our Research Group
Network Function Infrastructure Group (Measurement, Management and Control)
Communication Protocol Infrastructure Group
Secure Grid (PKI) Infrastructure Group
Grid Networking Issues

1. The network infrastructure which prevents degrading the throughput of whole grid applications by network delay and network fault is required.
2. It is necessary to carry out network resource scheduling as well as computing resource scheduling.
3. The network design and deploy technique for complicated grid networking is necessary.
4. The optimized communication protocol which allows grid application to utilize high-speed networks is required.
R&D Project Organization

R&D on networking technologies for providing/realizing network infrastructure for NAREGI.

Network Function Infrastructure  Kyushu institute of Technology
- Develop multi-point real-time measurement and visualization system
- Develop Grid network management and control system based on virtual organization operation policies

Communication Protocol Infrastructure  Osaka University
- Design and analysis of a communication protocol optimized for Grid
- Develop a parallel distributed network simulator for large-scale network

Secure Grid Infrastructure  NEC
- Policy definition for an authentication across organizations
- Development of an authentication for UNICORE/Globus
- Development of VO management cooperation functions
# R&D & Standardization Roadmap

<table>
<thead>
<tr>
<th>2004</th>
<th>2005</th>
<th>2006 -</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>High performance network middleware</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Network Information</td>
<td>Network Reservation</td>
<td>Advanced NW services</td>
</tr>
<tr>
<td><strong>Network service interfaces</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Measurement based Network control function</td>
<td>Network resource Scheduling function</td>
<td>Adaptive next generation NW infra function</td>
</tr>
<tr>
<td><strong>Network service Functions</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gigabit IP Network (MPLS, MPLS-VPN)</td>
<td>Optical Network</td>
<td></td>
</tr>
<tr>
<td><strong>Target Network Infrastructure</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Contribute use-case draft and advanced network service interfaces

GGF: Grid High-Performance Networking RG
Our Goal of Grid Networking Group

➢ To develop a system of measurement, management and control for adaptively using and assigning network resources in order to avoid resource contentions and to cost-effectively maintain quality of the network for the grid computing
Research Overview of Network Function Infrastructure Group

Develop a multi-point real-time flow measurement system and network management and control system using measurement data.

- Network Measurement
  - Multi-point real-time measurement and inference of fine-grained network-internal traffic and resource status

- Network Monitoring & Predictive
  - Develop distributed real-time flow measurement system with visualization tool

- Application Program Interface For Grids
  - Dynamic bandwidth, prioritization control and QoS routing based on virtual organization operation policies, and network resource scheduling

- NW Management and Control
  - Develop grid network management and control system based on VO operational policies

- Network Control & Network Analysis

Our Approach of Grid Networking Group

- Measurement based network resource control
  - Monitoring various grid application traffic status (such as traffic flow, RTT, and so on)
  - Real-time, fine-grained, and on-demand flow measurement

- Operating network control policy considering virtual organizations
  - Management of network resources based on virtual organizations
  - Sharing of network resources refer to virtual organization operational policies
Current System: Measurement Components

- Since the flow characteristic is grasped, attributes such as the amount flows, duration of flows and gaps of packets composing a specific flow can be measured
- Highly precise measurement of micro-second is proposed
- Distributed analysis processing in order to perform heavy flow identification process on real time

(1) Packets are mirrored to a distribution device

(2) Top N-byte packet are collected and sent to capturing device in parallel

(3) Every capturing device identifies flows by pattern matching

(4) Manager maintains flow definitions and stores collected date counted
Current System: Management Components

- Grid resource management considering virtual organizations (Network resources, Grid computing, etc)
- APIs (for evaluation) allow grid applications and grid middleware to obtain request network resources
- Dynamic bandwidth control and QoS routing based operational policies

Network resource analysis based on VO operational policies and using measurement info.

Network Equipment info., Network topology map info., grid node info., VO operational policies.

Dynamic configuration such as bandwidth, prioritization, and QoS routing.
Entire R&D Prototype

Grid Application / Grid Middleware

R&D prototype

VO manage GUI
VO operation Policy DB
VO network Manage DB

Network resource scheduling
Network resource analysis
Network utilization manager

Reservoir API
Monitoring API
Inter-working API
measurement DB

Grid network management and control system

Network control comp.
Network control comp.

Inter-working API

Visualization GUI

Prediction manager
Multi-pint measurement

Measurement DB

Multi-point Real-time Flow measurement system

Capturing device
Capturing device

Network

Research Plan of Communication Protocol Infrastructure Group

Develop a communication protocol optimized for Grid and a network performance evaluation method

Design a communication protocol optimized for Grid having scalability and compatibility with existing TCP

Establish a methodology for analyzing and evaluating a large-scale Grid network

Establish a methodology for designing a large-scale Grid network with strict SLA (Service Level Agreement)
Research Overview of Communication Protocol Infrastructure Group

Communication protocol optimized for Grid

Performance evaluation and improvement of GridFTP
- Modeling and performance analysis of GridFTP
- Simulation experiments for validating mathematical analysis
- GridFTP improvement

Design and analysis system for a large-scale Grid network

Distributed simulator of large-scale networks on Grid
- Large-scale simulation model division algorithm design
- Prototype implementation based on PDNS

Development of a communication protocol

Development of a design and analysis system
Modeling an Entire Network and GridFTP

TCP

represents the total packet rate

delay

explicitly model feedback loops

delay

delay

delay

delay

the sum of all outputs from previous models

distribute router output to other models

TCP

TCP

delay

delay

TCP

delay

RED router

Numerical Result for GridFTP

Number of TCP connections required for full link utilization

Network bandwidth: 1.0 [packet/ms]
(1) Perform steady state analysis. 
Predict the steady-state traffic volume on each transmission link

(2) Obtain multiple cuts based on the predicted traffic volume. 
Utilize the minimum cut algorithm in graph theory

(3) Choose one from multiple cuts that minimizes parallel simulation overhead 
Cut capacity is smallest and total traffic volumes are balanced between network submodels

(4) Repeat the above network division procedure 
Repeat steps (2) and (3) for the network submodel with the largest total traffic volume.
Example of Network Model Division

before

after

Research Plan of Secure Grid Infrastructure

- Develop a security model for Grid based on PKI and realize authentication across organizations and VO management.
- Development and Operation of authentication service for UNICORE and Globus.
- Development of certification authority (CA), registration authority (RA) and authentication policy based on the basic assurance level defined by GGF.
- Research & Development of authentication mechanism across policy domains to be proposed to GGF.
A Security Model of Grid Communication Platform

- User Proxy
- Resource
- Process
- Other Auth. Policy Domains

- User
- Create
- JOB Request
- Cert
- CSR
- RA
- CA
- Collaborate

- Network Function Infra.
- Secure Grid Infra.
Software Stack of Grid Communication Platform

- **Command User Interface**
- **Web User Interface**
- **User administration cooperation functions**
- **XKM S**
- **RA: Registration Functions**
- **Web Service Interface**
- **AICA** (existing Certificate Authority Free Software)
- **CP/CPS**
  - Auth. Policy (single domain)
  - Auth. Policy Extension (multi-domains)
- **NAS (NAREGI AUTHENTICATION SERVICE)**
- **NW Infrastructure**

- Development in FY 2003
- Development in FY 2004-2005
- Development in FY 2006-
NAREGI-CA V1.0 Architecture

- Collaboration with Grid and Web Services encrypted mail system

CA Server
- LDAP Server
- aicrlpub
- aicad

RA Server
- aienroll
- airad
- enroll (IIS ISAPI)
- enroll (apache CGI)

LDAP
- email
- LCMP

HTTP
- WEB

users

CA management tools
- aica
- certworker

PKI utilities
- certview
- certconv

• Collaboration with Grid and Web Services encrypted mail system

Already developed
NAREGI developed software
Open source

National Research Grid Initiative
VO (Virtual Organization)

- A set of individuals and/or institutions defined by such sharing rules form (OGSA)
- Study relationship VO and RO including identification rule.
  - ID Federation might be required
  - PKI Policy vs VO Policy

---

**Step for R&D**

2004
- VO & RO Modeling
  - Study of identification and related information, etc.
- A relation with a PKI policy is considered.
- The contents of VO policy are examined and extracted.

2005
- The service for VO policy is considered, designed and developed.
Collaborations within NAREGI

Super-Scheduler
Grid Distributed information Services

The notification of network monitoring info. API, scheduling Info. API, network resource reservation API to the super-scheduler and cooperation with account management

High-performance & secure Grid networking

- Specify Grid application types which requires reservation of communication quality
- Verification of network control functions on real nano-applications.

Grid-Enabled Applications

National Research Grid Initiative
Summary

- NAREGI will aim at High Performance Managed Network based on fine grain distributed measurements and VO based control.
- Analysis of High Performance Transport protocols are performed.
- Design tools for grid network will be developed.
- Federation of multiple VO in terms of security will be considered.