We4.1.1

G-lambda: Coordination of a Grid Scheduler and Lambda Path Service over GMPLS

Michiaki Hayashi[1], Tomohiro Kudoh[2] and Yasunori Sameshima[3]

[1] KDDI R&D Laboratories Inc.
[2] National Institute of Advance Industrial Science and Technology (AIST)
[3] NTT Network Innovation Laboratories

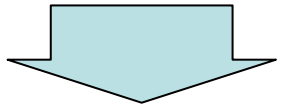
Sep. 27, 2006

Agenda

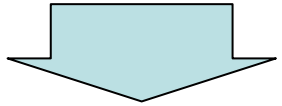
- Background
- Network service invocation
- G-lambda project
- Architecture overview
- GNS-WSI: Network service interface
- Nation-wide Grid networking demonstration

Background

- Growth of the data traffic: Computers connected to the network
  - Grid technology is an emerging framework for computing applications
  - Today, network service provider must aware of computing applications



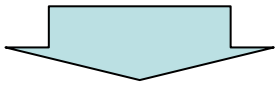
- Grid computing application's requirements and assumptions
  - Topological flexibility in geographically-distributed environments
  - Awareness of application requirements (Bandwidth, Latency, etc.)
  - Support of advance reservation as other Grid resources
  - Resource control/management based on Web services framework



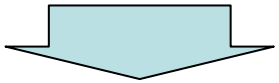
- Network service for Grid
  - Inter-working between Grid scheduler/middleware and network
  - Definition of network service interface
  - Network resource management techniques

Network service invocation by Grid users

- Service invocation models discussed in OGF (Open Grid Forum)
  - Direct invocation model
    - The client is directly attached to the transport network and is itself a member of the service signaling



    - Control plane interaction between Grid users and network service provider
    - Standard-based control plane protocols (MPLS, GMPLS, etc) do NOT support advance reservation
  - Indirect invocation model
    - The client invokes transport network services using proxy



    - Networks need to provide resource management service to Grid middleware.
    - Resource management and service interface are the key issues.

- Network service
  - Direct invocation model fits for user's full controllability of networks
    - Control protocols need enhancements to be a external service interface
  - Indirect invocation model fits commercial network services better
    - A standard open interface between Grid middleware and network resource manager is required, but has not established yet.

Examples of network service strategy for users

Control plane-based strategies

- Control plane techniques in standard bodies
  - IETF MPLS/GMPLS allows users to request network service on-demand by peering with PE or overlaying carrier network
    - IETF L1-VPN framework explicitly supposes user interface both with signaling-based and management system-based (management system interface has not been defined)
  - OIF UNI 1.0/2.0 signaling allows overlay connection by overlaying carrier network
- Just in time (JIT) signaling
  - Fast (one-way or “tell&go”) signaling scheme for lambda/OBS path provisioning
- Grid enabled GMPLS (G^2MPLS)
  - Enhance GMPLS for user-control of heterogeneous Grid resources

Vertical integration strategies

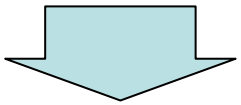
- Network description language (NDL)
  - Network resources are described by RDF and found based on semantic Web technology
- User Controlled Light Path (UCLP)
  - UCLP software allows users to manage the network by configuring virtual private network
- Vertical integrated optical testbed for large applications (VIOLA)
  - Meta scheduler coordinates network resource using WS-Agreement with client
- Enlightened
  - HARC acceptor provides services for user side with XML and phase-commit procedure
- G-lambda
  - Defines “GNS-WSI” between Grid middleware and network resource middleware

Related activities in standard bodies

- OGF
  - Telco-CG (Community Group)
    - Discussing to create informational document of the CG
    - How to provide dependable large-scale Grids
    - How to use Grid technologies to improve own operations (e.g., billing, events' analysis, and modeling)
    - How to provide Grid managed services to adopt new business models
  - GHPN-RG (Grid High Performance Networking-Research Group)
    - Established GFDs regarding network services for Grids
    - Network service use cases is under discussion
    - Discussion on Grid service interface spec is just started…
- ETSI
  - Established working group on Grid and Grid networking
    - http://portal.etsi.org/portal_common/home.asp?tbkey1=GRID
  - Handling standardization of middleware/protocol and interoperability

Activities in standard bodies (contd.)

- ITU-T
  - Workshop NGN and Grids will be held with OGF in 23-24 Oct. (this month)
    - http://www.itu.int/ITU-T/worksem/grid/programme.html
  - The workshop will discuss;
    - Additional features required to be considered in ITU-T's NGN Release 2
    - Impact of NGN on Grids
    - Scenarios for telcos



- Efforts of OGF and Grid technology itself started to be considered not only for research community but also commercial or telecom industry
  - Is Grid understood still as a research one?
  - Is Grid understood same as OGF expectation?
  - Is Grid evolved in the discussion of other standard bodies?
- Transformation point for wide adoption to telecom field

G- lambda project overview

- Joint project of AIST, NICT, NTT and KDDI R&D labs.
- G-lambda project has been started in December 2004.
- The goal of this project is to establish a standard web services interface (GNS-WSI) between Grid resource manager and network resource manager provided by network operators.

National Institute of Advanced Industrial Science and Technology
AIST

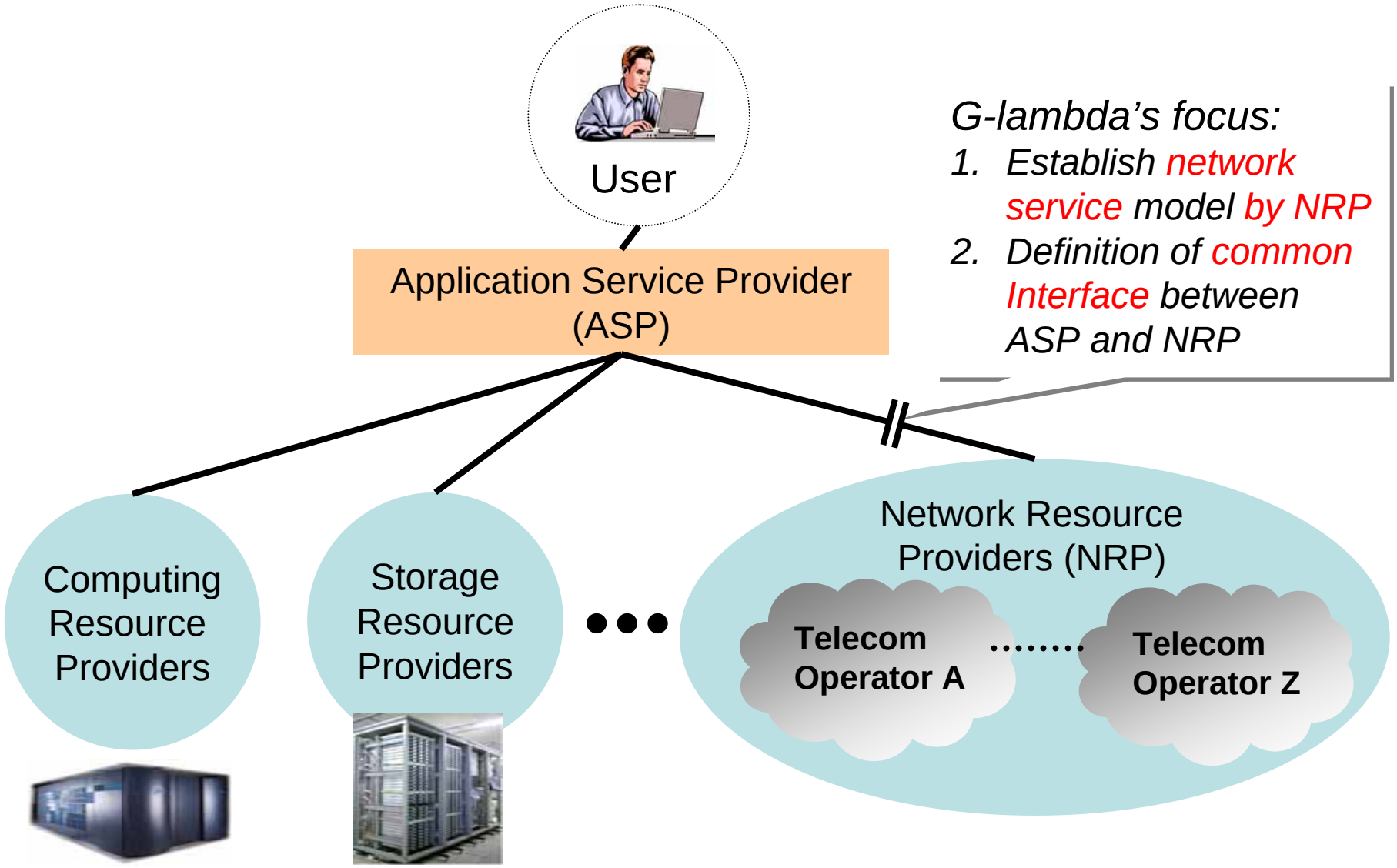
NTT

NiCT National Institute of Information and Communications Technology

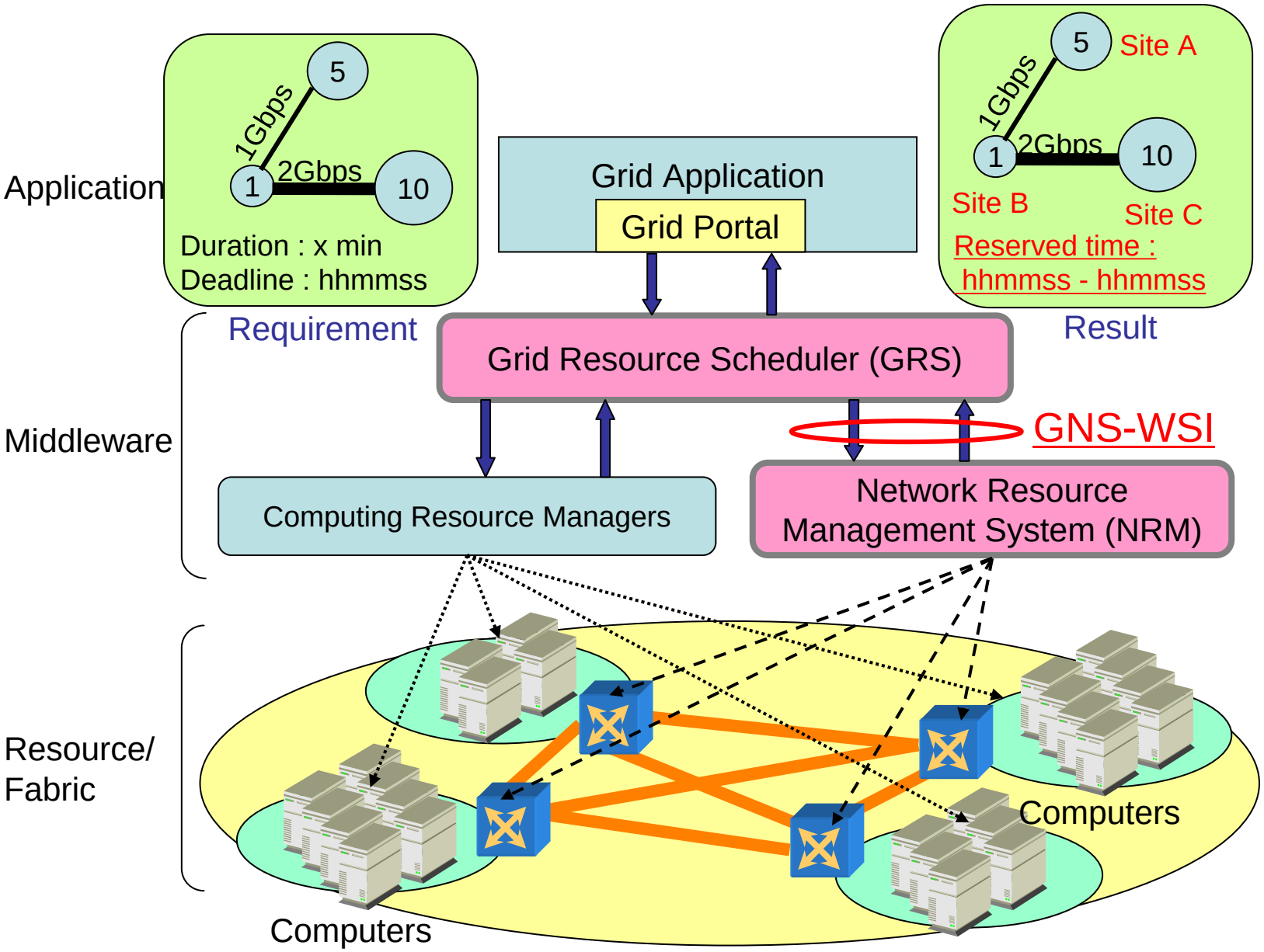
JGNII

KDDI
KDDI R&D LABS

An example of service model



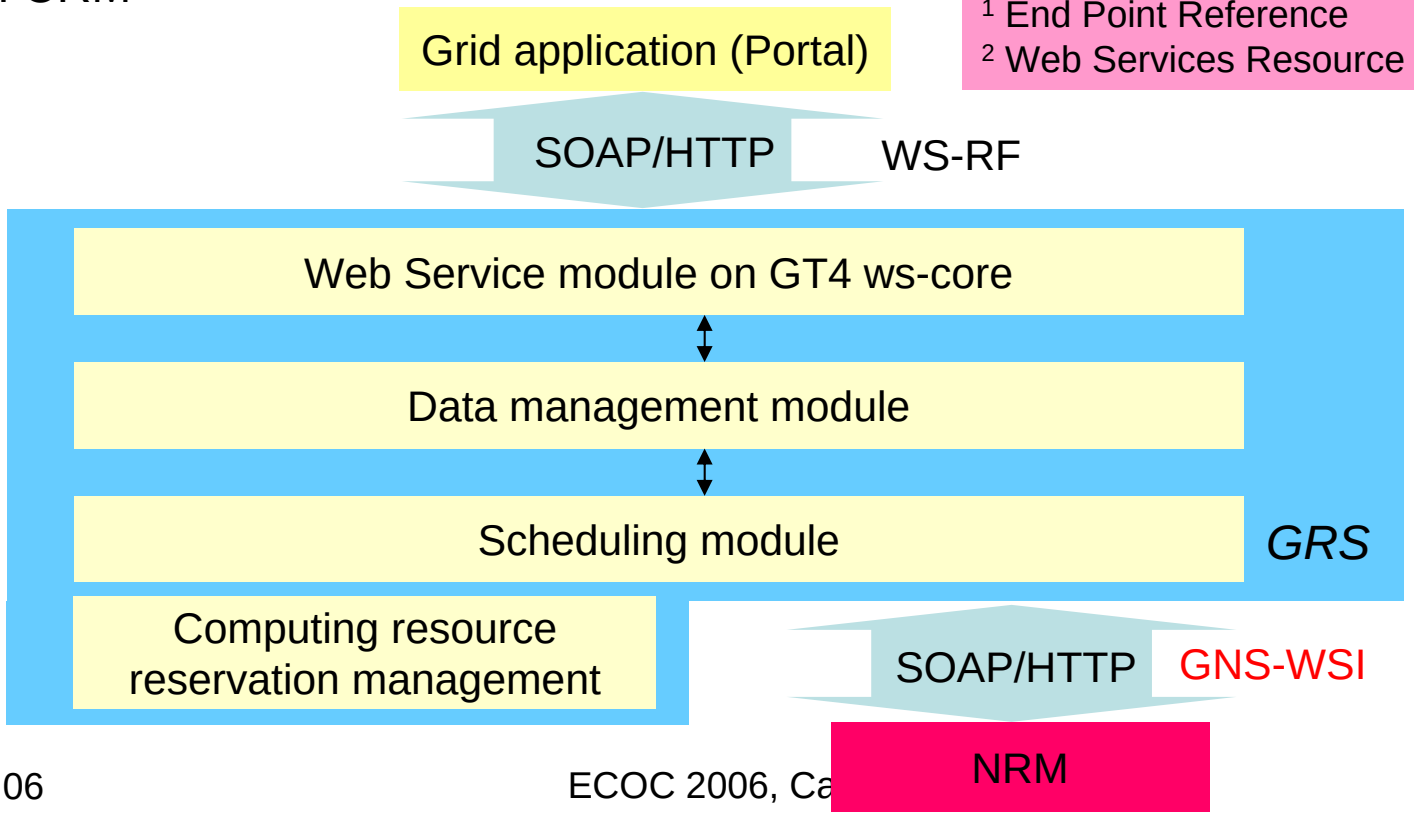
G-lambda: Architecture



GRS

- Developed by AIST
- Web services module: Provides reservation service to Grid portal
  - Reservation resource is managed as EPR[1] based on WSRF[2]
- Data management module: Stores reservation resource information
- Scheduling module: Negotiates resource reservation with a NRM and CRMs
  - GNS-WSI allows the network resource negotiation
- Computing resource reservation management: Manages reservation state in each CRM

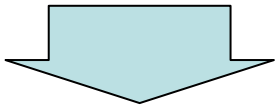
[1] End Point Reference
[2] Web Services Resource Framework



Co-scheduling operation of GRS

Input conditions from Grid portal

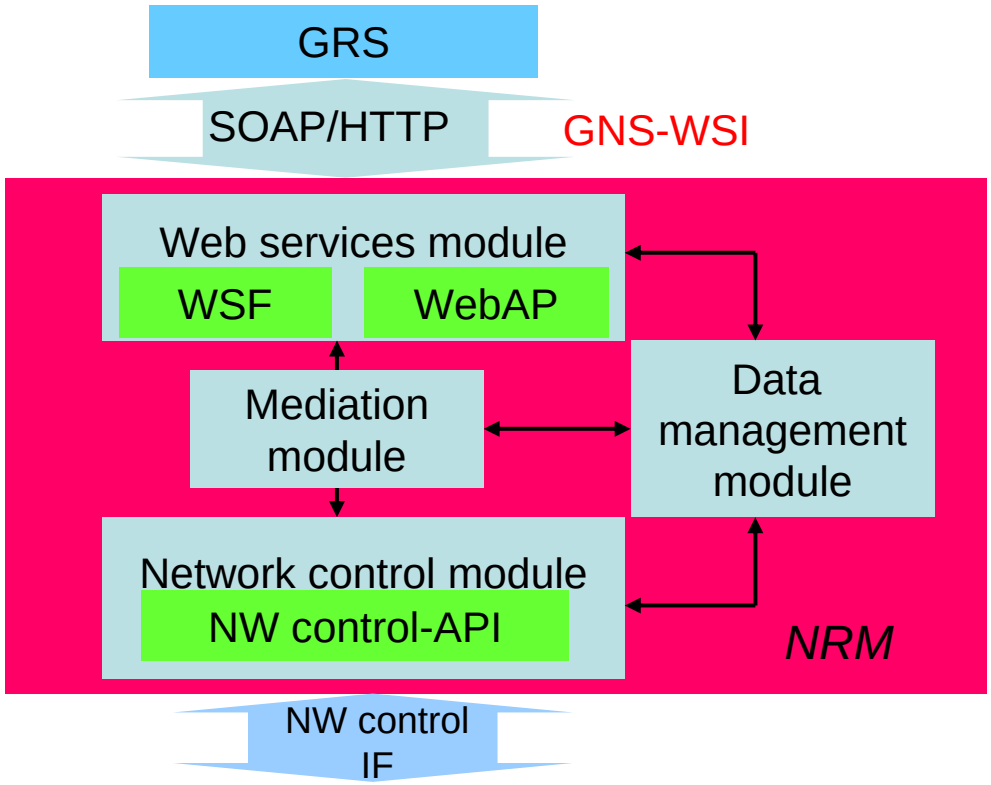
- Required time duration of a job execution
- Deadline of the job completion
- Required resources to carry out a job
  - Computing resource: The number of CPUs
  - Network resource: Bandwidth between cluster sites



- Scheduling module searches available resources under the specified conditions
  - Discovery of available resource from resource managers
  - Resource finding with the depth first search scheme
- Developing an effective scheduling scheme is future work

NRM

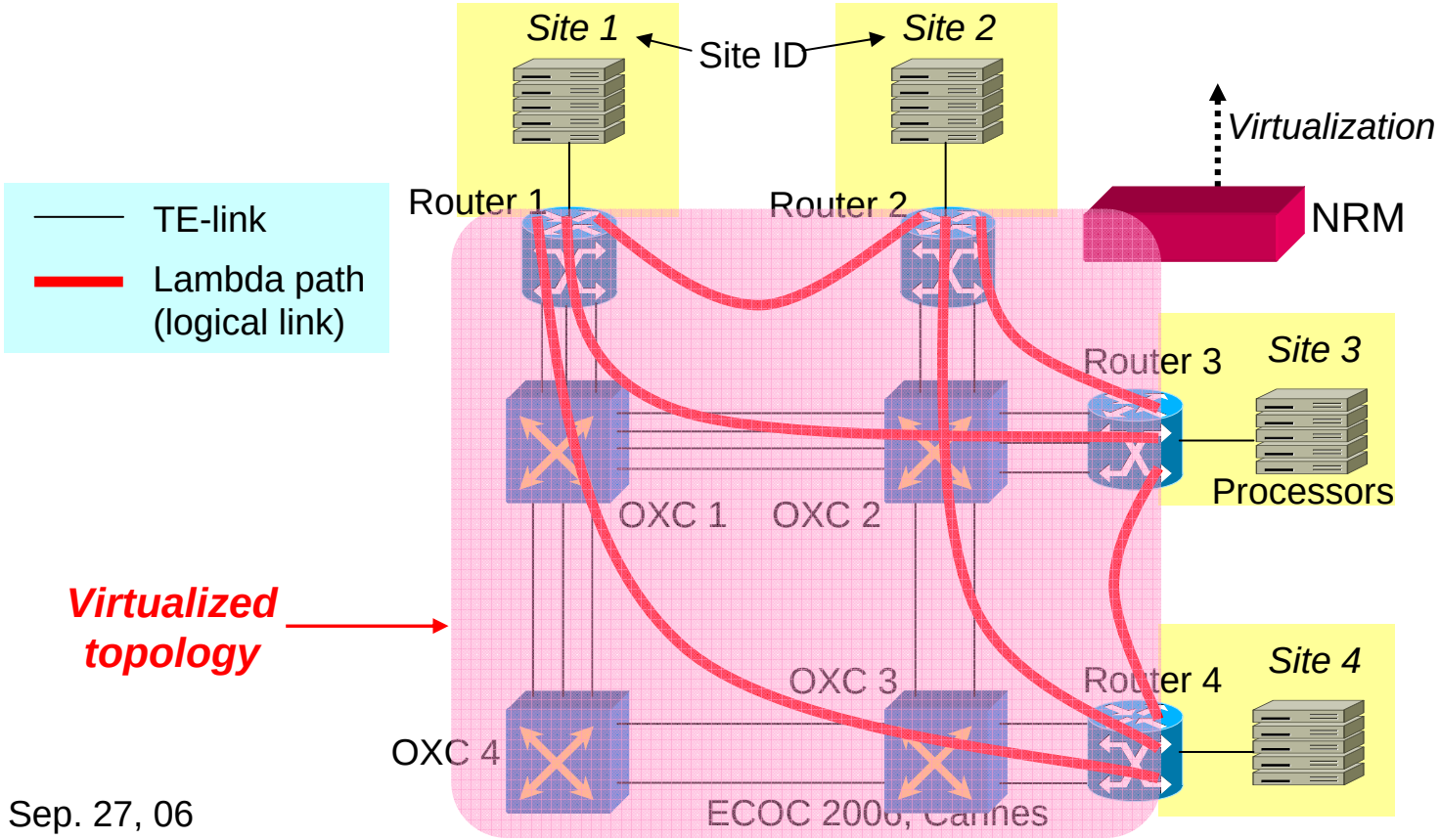
- Developed by KDDI R&D Laboratories Inc.
- Web services module: Provides network service for GRS
  - Polling-based service over GNS-WSI
  - Service operation and procedure defined by the WSDL
- Mediation module: Scheduling and virtualization of network resources
- Network control module: Control and monitoring of the network
  - Set up and monitor end-to-end paths through network control IF



[1] Web Services Framework
[2] Web Application Server
[3] Application Program Interface

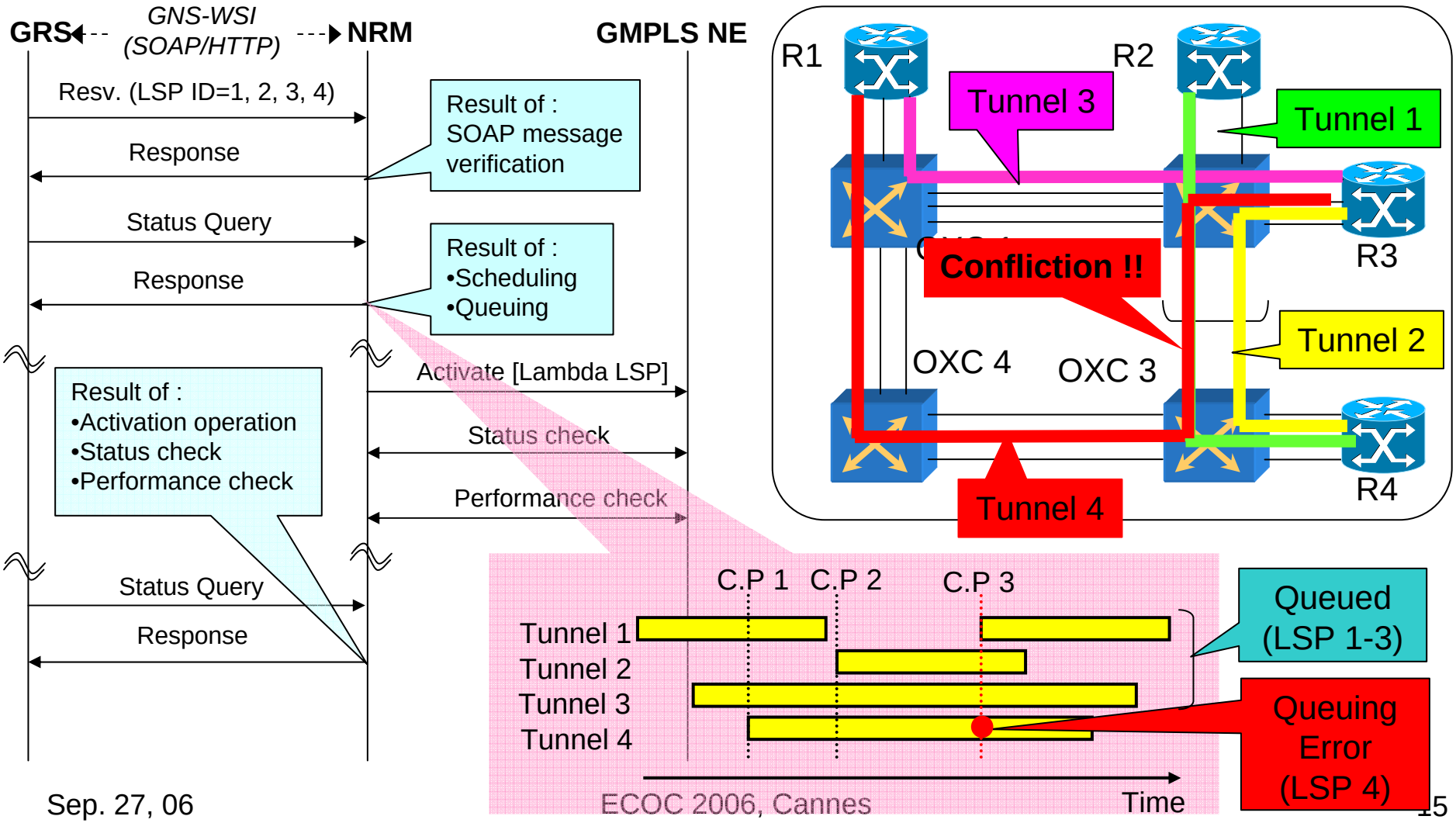
Network resource virtualization

- Network resource virtualization of NRM
  - Creation of policy-based virtualized topology using lambda-LSPs
    - Masking detailed network configuration (TE-links, routers, OXCs, etc.)
    - Resource scheduling based on the virtual links
  - Resource specification using string-based site ID
    - Masking IP addressing
- Role of GMPLS protocol in this model
  - Allows NRM to control multi-layer network by simply accessing Ingress node
  - No direct GMPLS inter-work with users



Network resource scheduling

- Advance reservation requirements
  - End-points, bandwidth, reservation time, network delay, fault recovery
  - Network resource identification by GRS using LSP ID



GNS-WSI: Service parameters

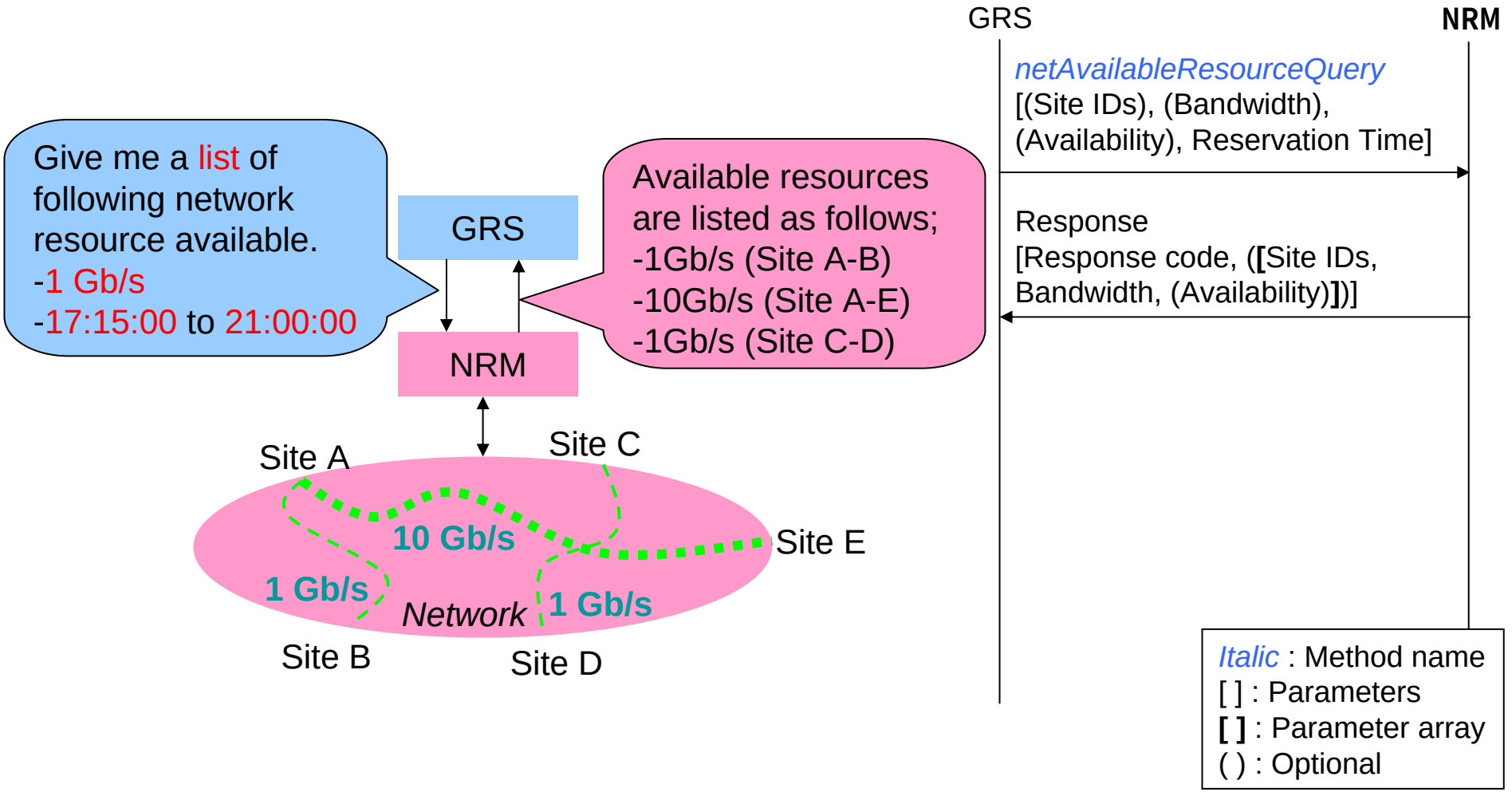
Parameter	Usage	Value	Remarks
Site ID	ID to specify A and Z points	String	Name or ID of sites
Bandwidth	Bandwidth of the resource	Positive integer (kbit/s)	
Reception ID	ID managed by NRM for each request	Integer (-2^{32} ～ $2^{32}-1$)	
LSP ID	Resource identification in each request	Integer (-2^{32} ～ $2^{32}-1$)	Reception ID-unique
Latency	Latency between end points	Positive integer (msec)	
Availability	Network protection of network resource	Integer (-2^{32} ～ $2^{32}-1$)	0 = Un-protected 1 = Protected
Reservation time	Start time and end time of the reservation	xsd:dateTime	YYYY-MM-DDTHH:MM:SSZ
Response code	Process result of messages over GNS-WSI	Integer (-2^{32} ～ $2^{32}-1$)	Included in SOAP response of NRM
Status code	Status of the network resource	Integer (-2^{32} ～ $2^{32}-1$)	

An example XML exchanged through GNS-WSI

```
<requirements>
    <network
      aPoint=“Tokyo"
      zPoint=“Osaka"
      startTime="2006-09-07T04:15:00Z"
      endTime="2006-09-07T06:15:00Z“
      bandwidth="1000000"
      latency="10"/>
</requirements>
```

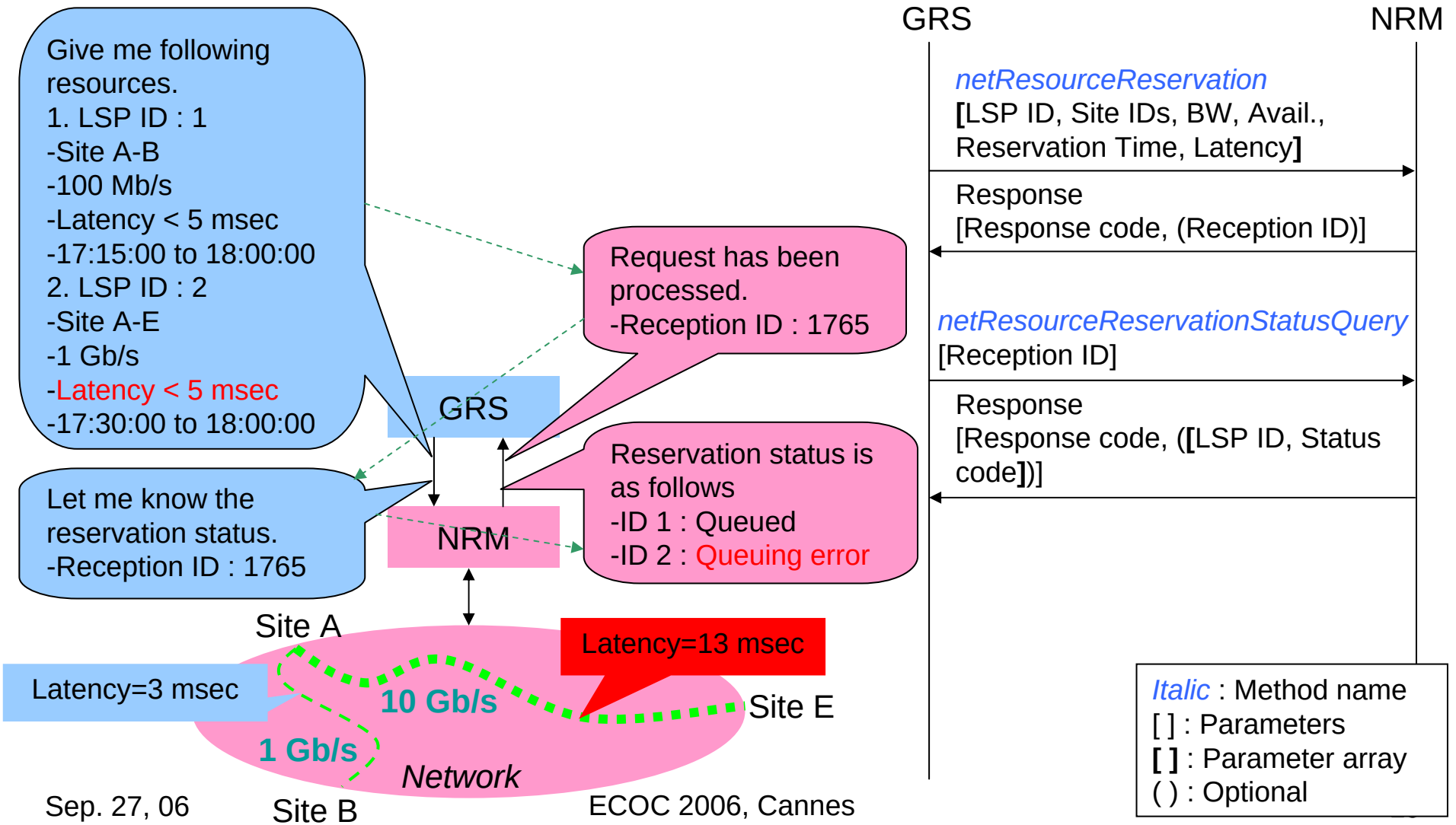
GNS-WSI: Information services

■Discovery of available resources
SOAP method : *netAvailableResourceQuery*



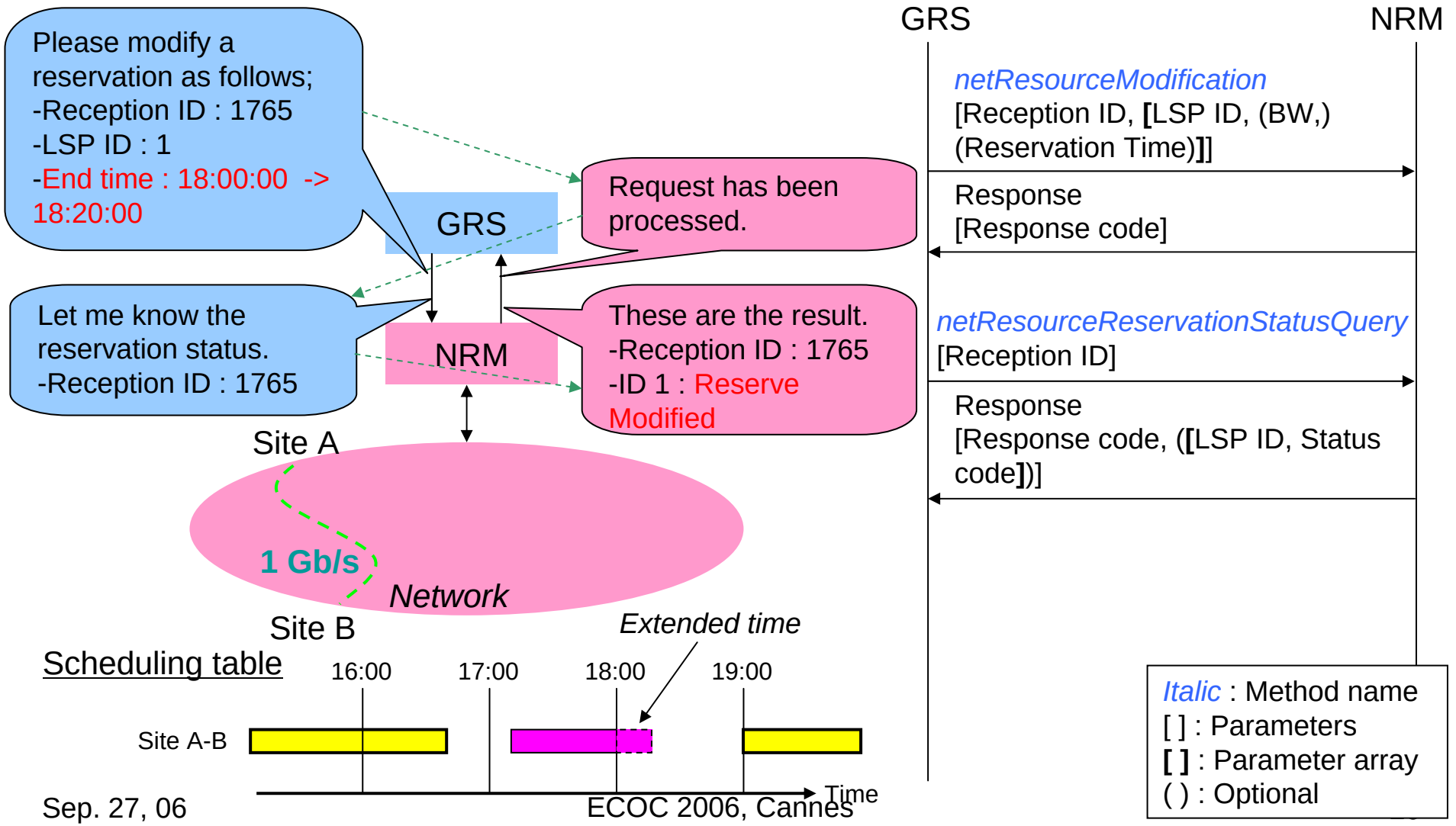
GNS-WSI: Advance reservation service

■Reservation of network resources in advance of the job execution
SOAP method : *netResourceReservation*



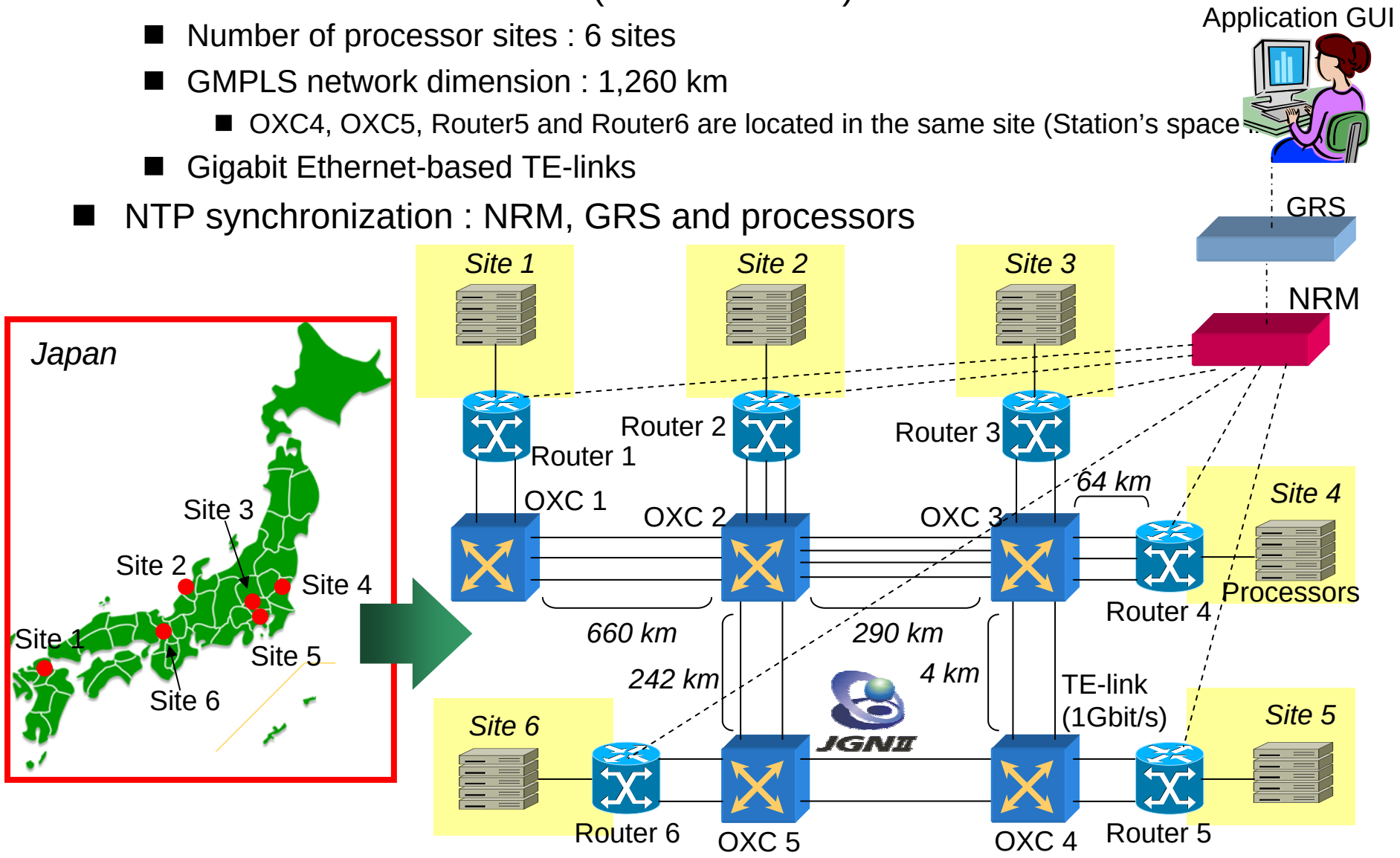
GNS-WSI: Reservation modification service

■Modification of queued reservation or on-going reservation
SOAP method : *netResourceModification*



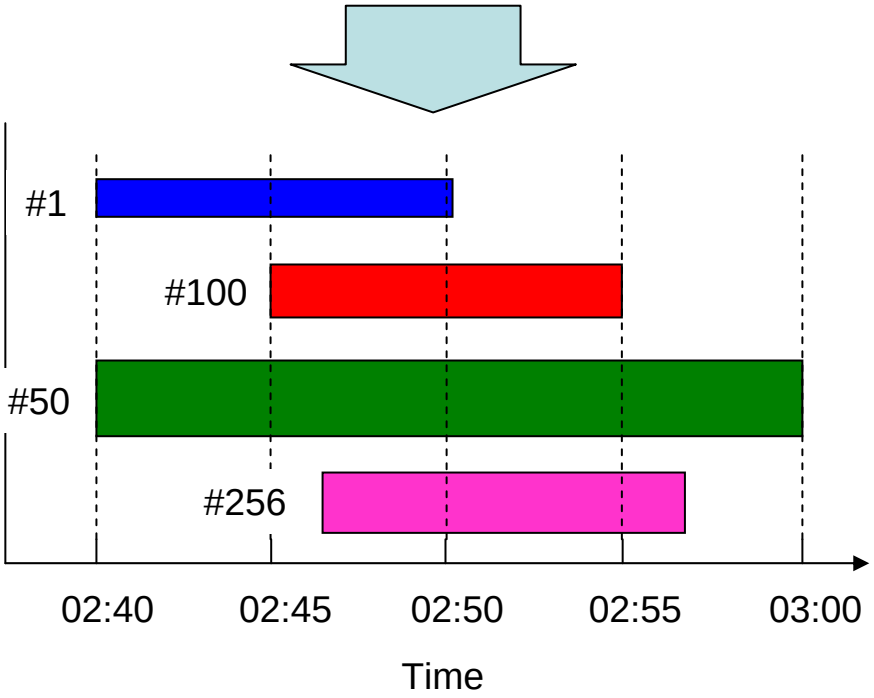
Demonstration network

- Nation-wide GMPLS network (JGN II test bed)
  - Number of processor sites : 6 sites
  - GMPLS network dimension : 1,260 km
    - OXC4, OXC5, Router5 and Router6 are located in the same site (Station's space)
  - Gigabit Ethernet-based TE-links
- NTP synchronization : NRM, GRS and processors

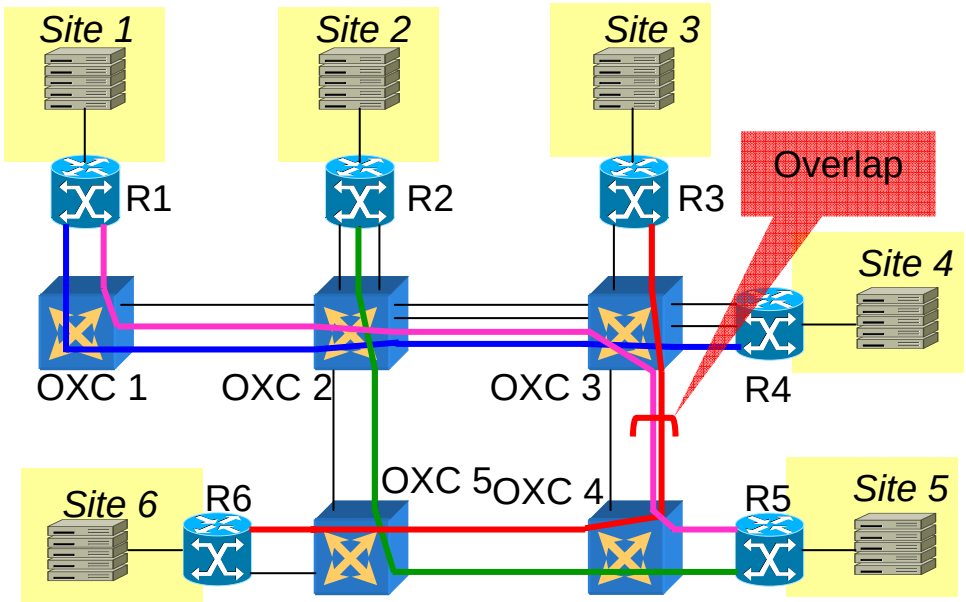


Advance reservation parameters

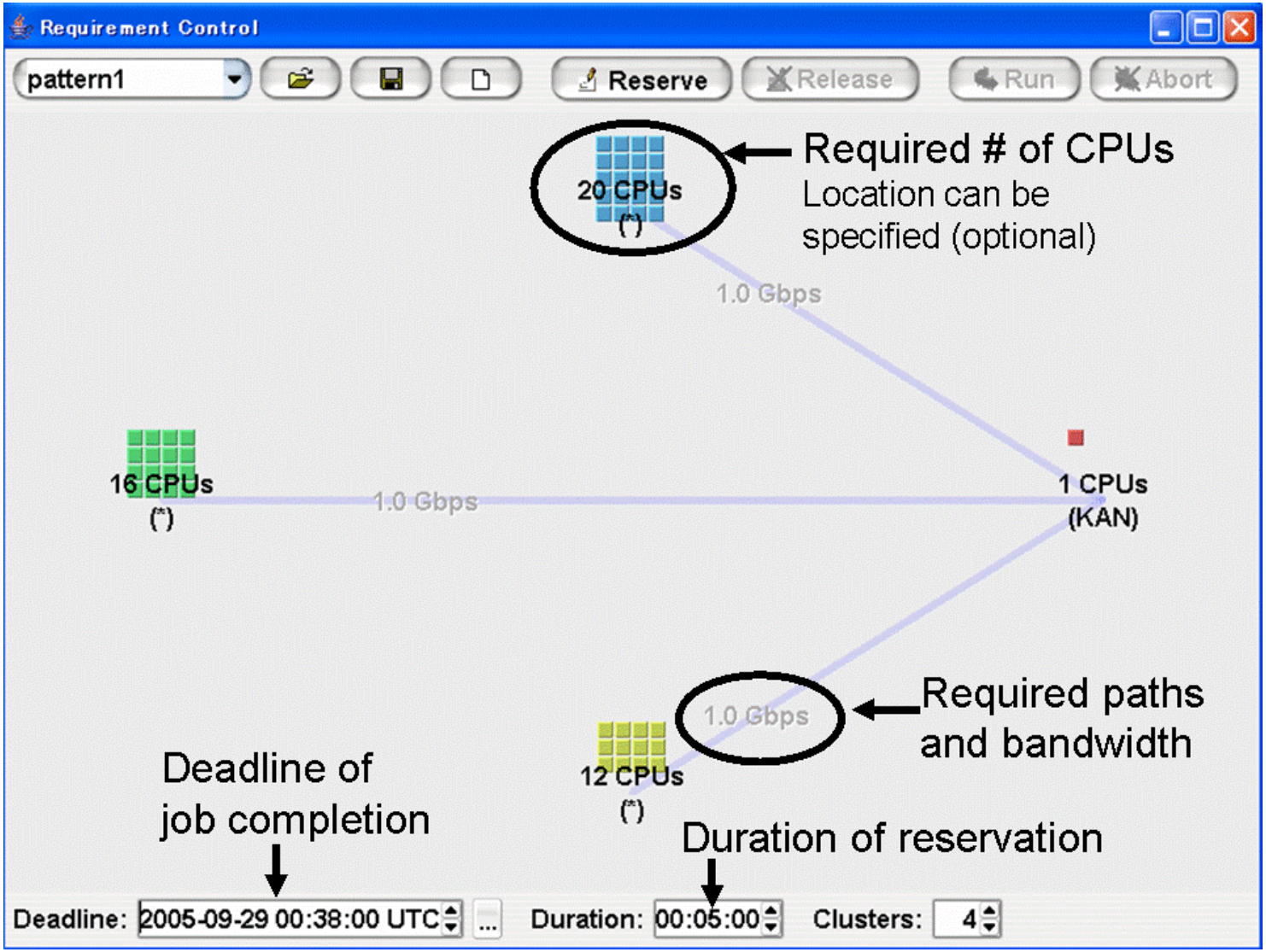
LSP ID #	End points	Bandwidth	Network restoration	Delay	Start time	End time
1	Site 1-4	500 Mbit/s	Unprotected	10 msec	02:40:00	02:50:10
100	Site 3-6	700 Mbit/s	Unprotected	10 msec	02:45:00	02:55:00
50	Site 2-5	1 Gbit/s	Unprotected	10 msec	02:40:00	03:00:00
256	Site 1-5	800 Mbit/s	Unprotected	15 msec	02:47:00	02:57:00

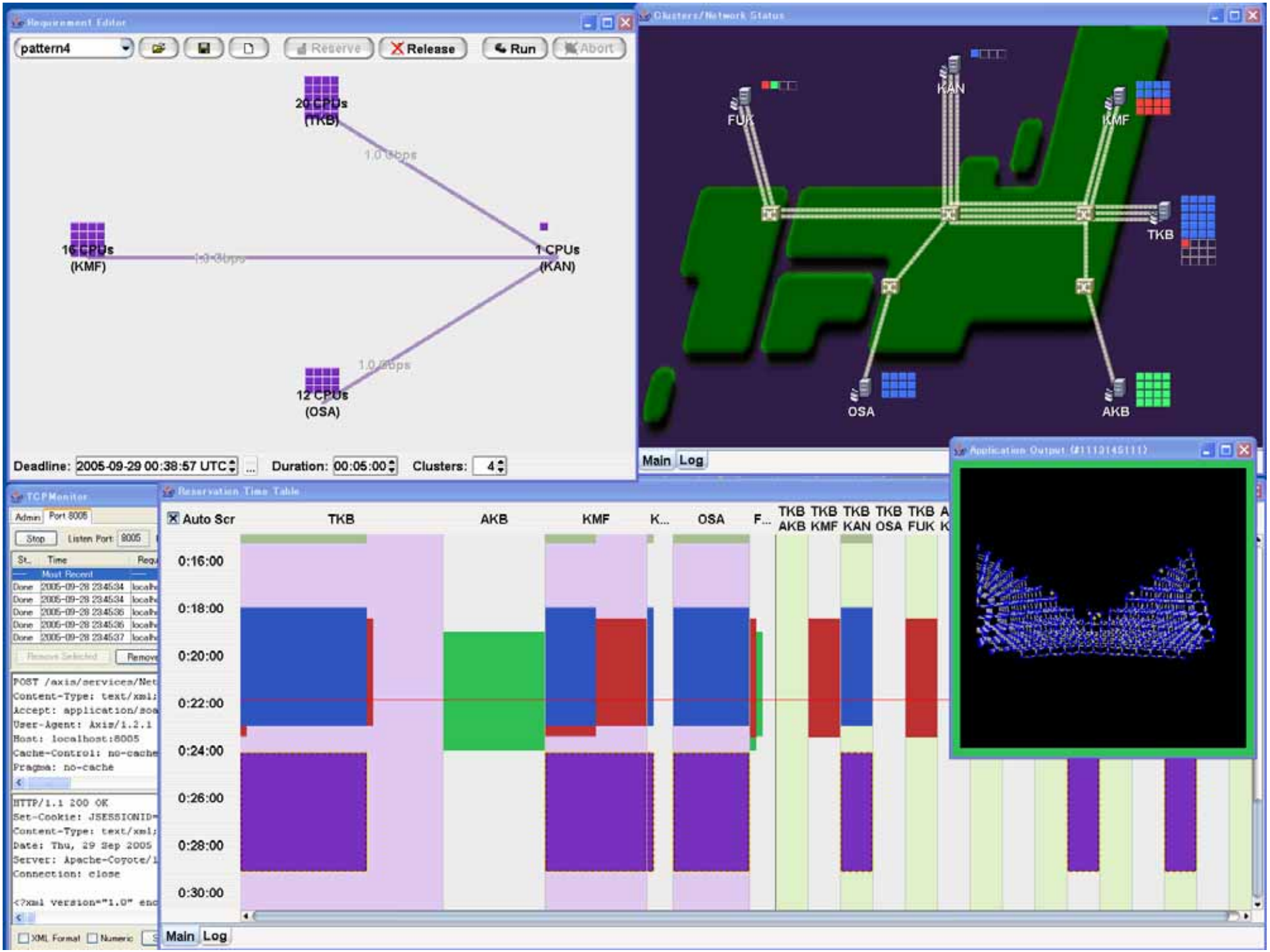


LSP route management of NRM



Grid portal: Requirement editor input




Requirement Editor
pattern4
Reserve
Release
Run
Abort
20 CPUs
(TKB)
1.0 Gbps
16 CPUs
(KMF)
1 CPUs
(KAN)
1.0 Gbps
12 CPUs
(OSA)
Deadline: 2005-09-29 00:38:57 UTC
Duration: 00:05:00
Clusters: 4
Clusters/Network Status
FUK
KAN
KMF
TKB
OSA
AKB
Main
Log
TCPMonitor
Admin
Port 8005
Stop
Listen Port: 8005
Reservation Time Table
Auto Scr
TKB
AKB
KMF
K...
OSA
F...
TKB AKB
TKB KMF
TKB KAN
TKB OSA
TKB FUK
0:16:00
0:18:00
0:20:00
0:22:00
0:24:00
0:26:00
0:28:00
0:30:00
Application Output

Conclusion

- GRS and NRM allows co-allocation of both computing and network resources
- Network service for Grid application is demonstrated by using preliminary interface definition of “GNS-WSI”
- Nation-wide Grid computing is successfully demonstrated

Next steps

- Inter-domain advance reservation and co-allocation
  - Collaboration with Enlightened project: “Inter-domain advance reservation of coordinated network and computing resources over the Pacific”
- Enhancement of GNS-WSI
- Promote standardization of interface to reserve bandwidth in advance

Thank you.

G-lambda project
http://www.g-lambda.net