G-lambda: An Interface for Bandwidth Reservation from Applications and Middleware

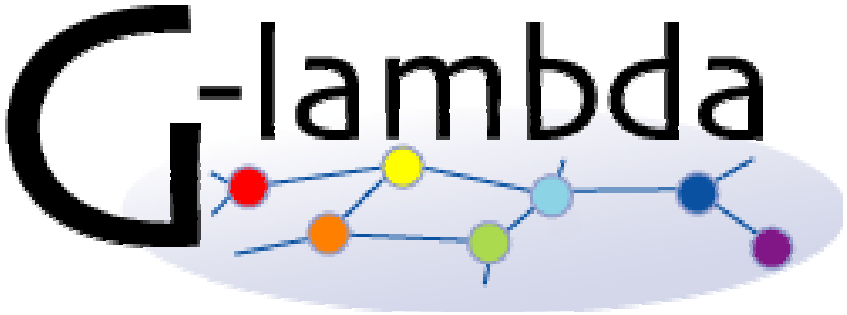
Tomohiro Kudoh[1], Michiaki Hayashi[2], Akira Hirano[3],

Shuichi Okamoto[4], Atsuko Takefusa[1], Takahiro Miyamoto[2],
Yukio Tsukishima[3], Tomohiro Otani[4], Hidemoto Nakada[1],

Hideaki Tanaka[2], Atsushi Taniguchi[3], Yasunori Sameshima[4],

1. National Institute of Advanced Industrial Science and Technology (AIST)

2. KDDI R&D Laboratories,
3. NTT Network Innovation Laboratories,
4. National Institute of Information and Communications Technology (NICT)

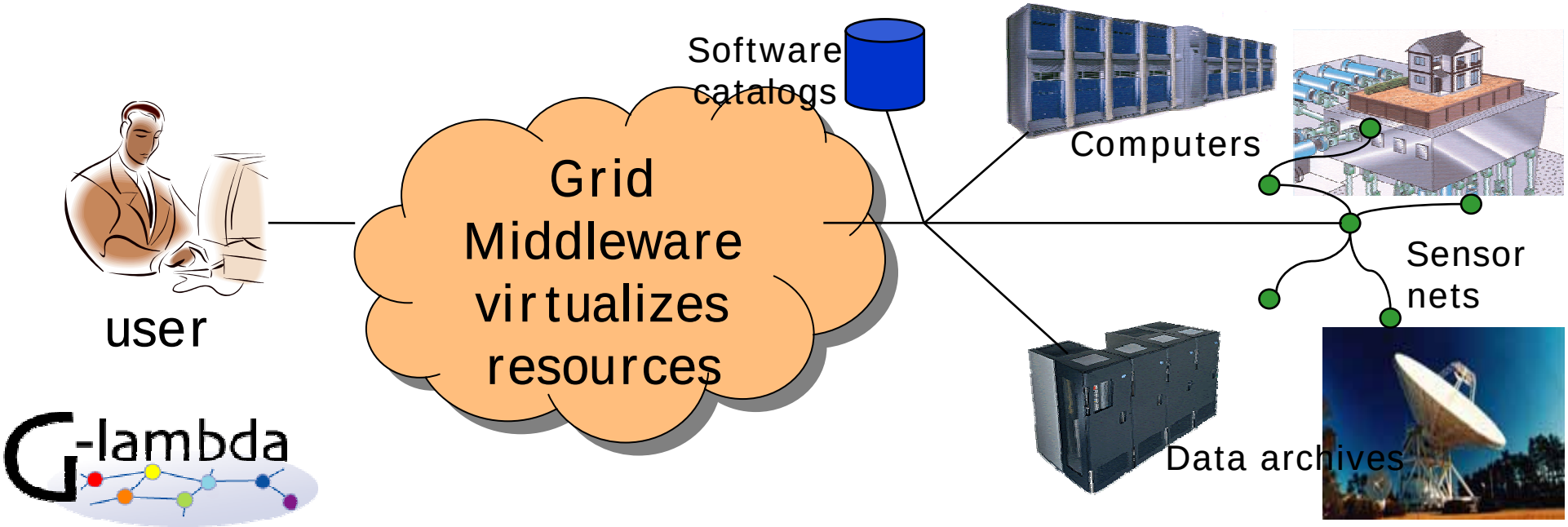


Outline

- **In-advance bandwidth reservation and Grid**
- **G- lambda project: GNS-WSI and architecture**
- **Experiment on a single-domain network**
- **Architecture for multi-domain network**
- **Experiment on a multi-domain network**

In-advance bandwidth reservation and Grid

- Grid provides a single system image to users by virtualization of service infrastructure such as computing, data and network resources from multiple domains.
- Users do not care about actual resources they are using. Grid middleware (such as planner, broker and scheduler) coordinates resources and provides virtual infrastructure.

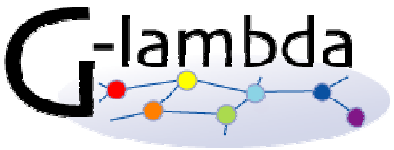


Network service for Grid

- To realize such virtual infrastructure for Grid, resource management is one of the key issues.
- Grid middleware should allocate appropriate resources, including network resources, according to user's request.
- Network resource manager should provide resource management service to Grid middleware.

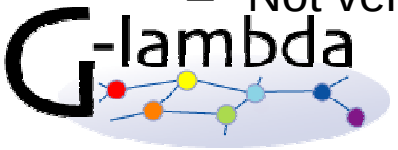
Network Service

- A standard open interface between Grid middleware and network resource manager is required, but not yet established.



Timing of resource provisioning

- On demand provisioning
  - Common in traffic engineering of network
  - End time of resource provisioning is undetermined
- Batch scheduling
  - Most of the schedulers for computing resources use the batch model
    - Make a queue of jobs with priority, and execute jobs in the order
    - Good for resources managed by a single scheduler
- In-advance reservation/scheduling
  - For resources provided by multiple providers, advance reservation is suitable
    - Each provider can control its own resources with a reservation table
  - Existing protocols such as GMPLS does not support advance reservation.
    - The routing function of GMPLS assumes on-demand provisioning of paths
  - Not very common for computing resources too.

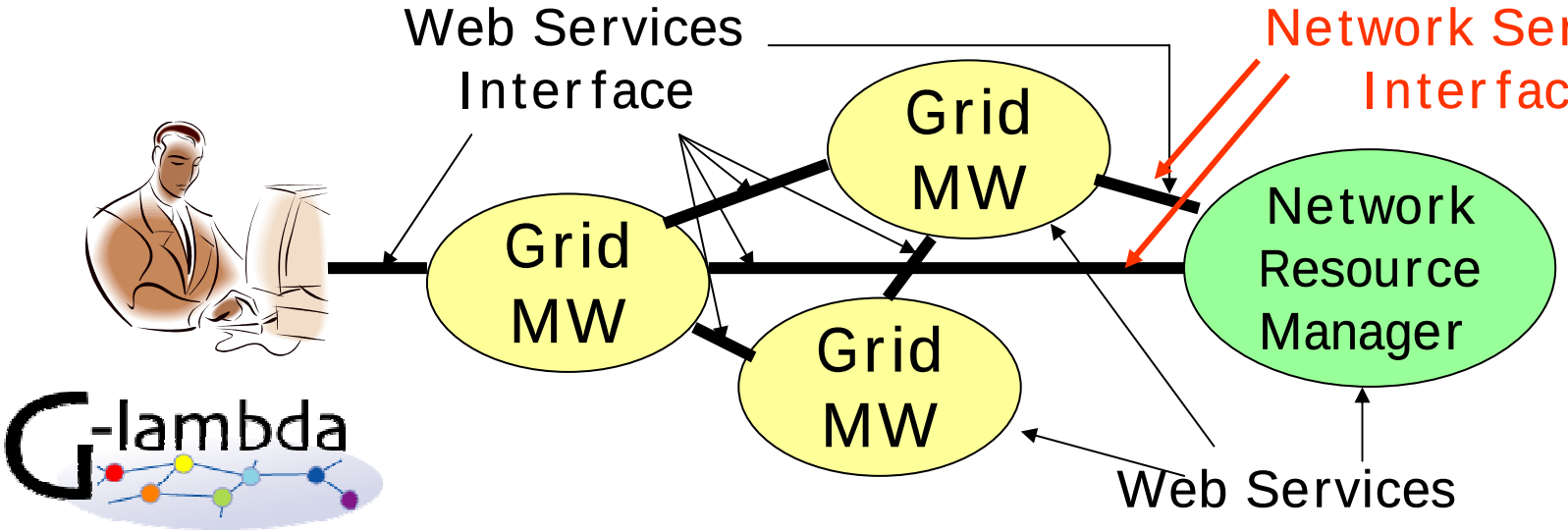


Design of the network service interface

- Web Services
  - Grid is being built based on Web Services technology
  - Network service should be provided as a “Web Services”.
- SLA support
  - Bandwidth, latency etc.
- In-advance reservation
  - Reserve bandwidth

What is “Web Services”?

- Application components which can be accessed thorough open standard web protocols (XML, SOAP, etc.).
- Web Services interface enables interaction between application components
  - Very high level interoperability among the components.
- A standard Web Services based open interface between Grid middleware and network resource manager is required



Outline

- In-advance bandwidth reservation and Grid
- **G- lambda project: GNS-WSI and architecture**
- Experiment on a single-domain network
- Architecture for multi-domain network
- Experiment on a multi-domain network

G- lambda project overview

- Joint project of KDDI R&D labs., NTT, NICT and AIST.
- 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 R&D LABS

The G-lambda Team

National Institute of Advanced Industrial Science and Technology
AIST

- Tomohiro Kudoh
- Hidemoto Nakada
- Atsuko Takefusa
- Yoshio Tanaka
- Fumihiro Okazaki
- Satoshi Sekiguchi
- Hiroshi Takemiya
- Motohiko Matsuda
- Seiya Yanagita
- Katsuhiko Okubo

NiCT National Institute of Information and Communications Technology

- Shuichi Okamoto
- Tomohiro Otani
- Yasunori Sameshima

KDDI R&D LABS

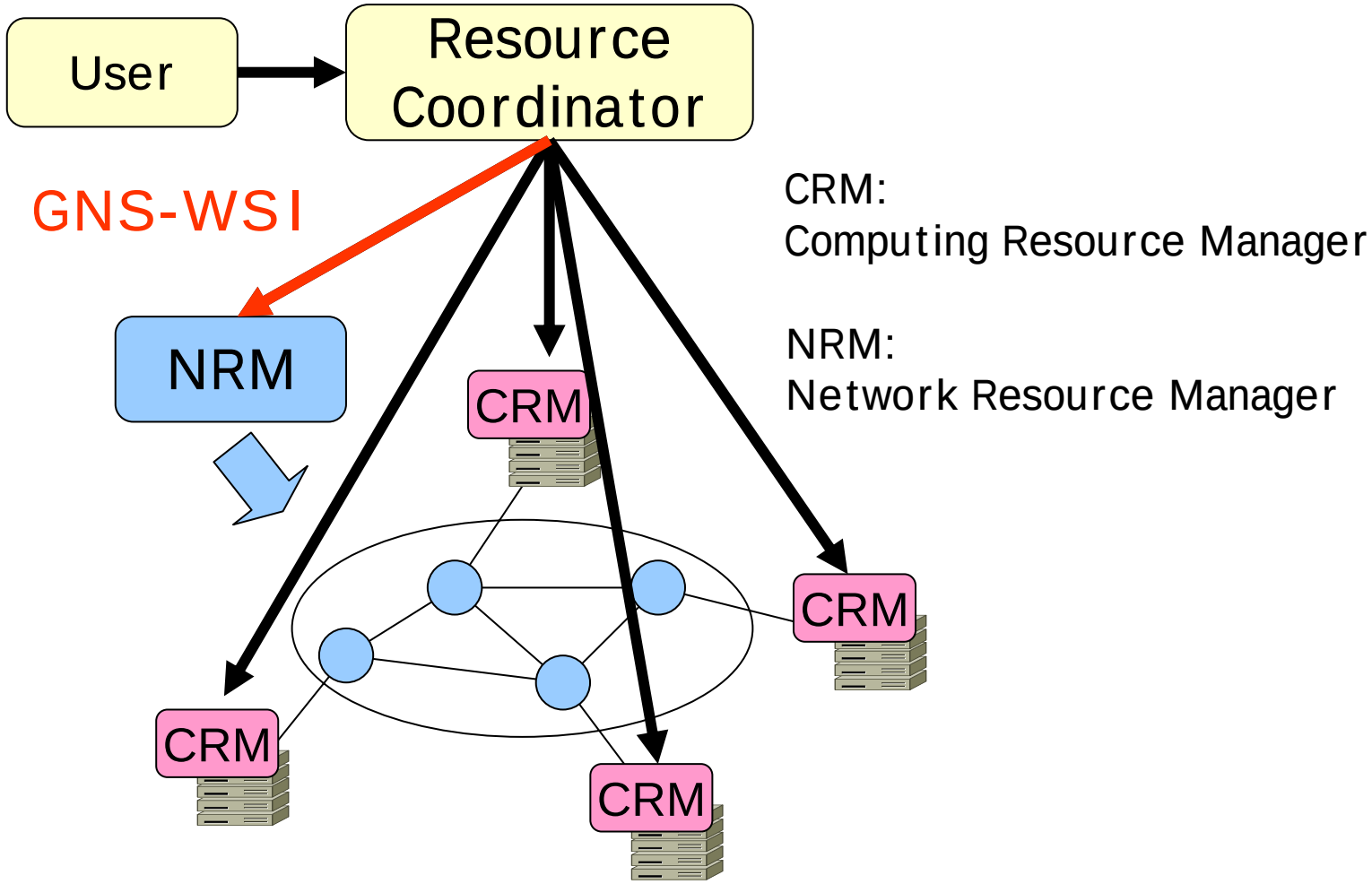
- Masatoshi Suzuki
- Hideaki Tanaka
- Tomohiro Otani
- MunefumiTsurusawa
- Michiaki Hayashi
- Takahiro Miyamoto

NTT

- Akira Hirano
- Yasunori Sameshima
- Wataru Imajuku
- Takuya Ohara
- Yukio Tsukishima
- Atsushi Taniguchi
- Masahiko Jinno
- Yoshihiro Takigawa

G-lambda

System architecture

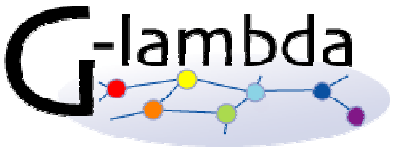


Two application components interact through GNS-WSI

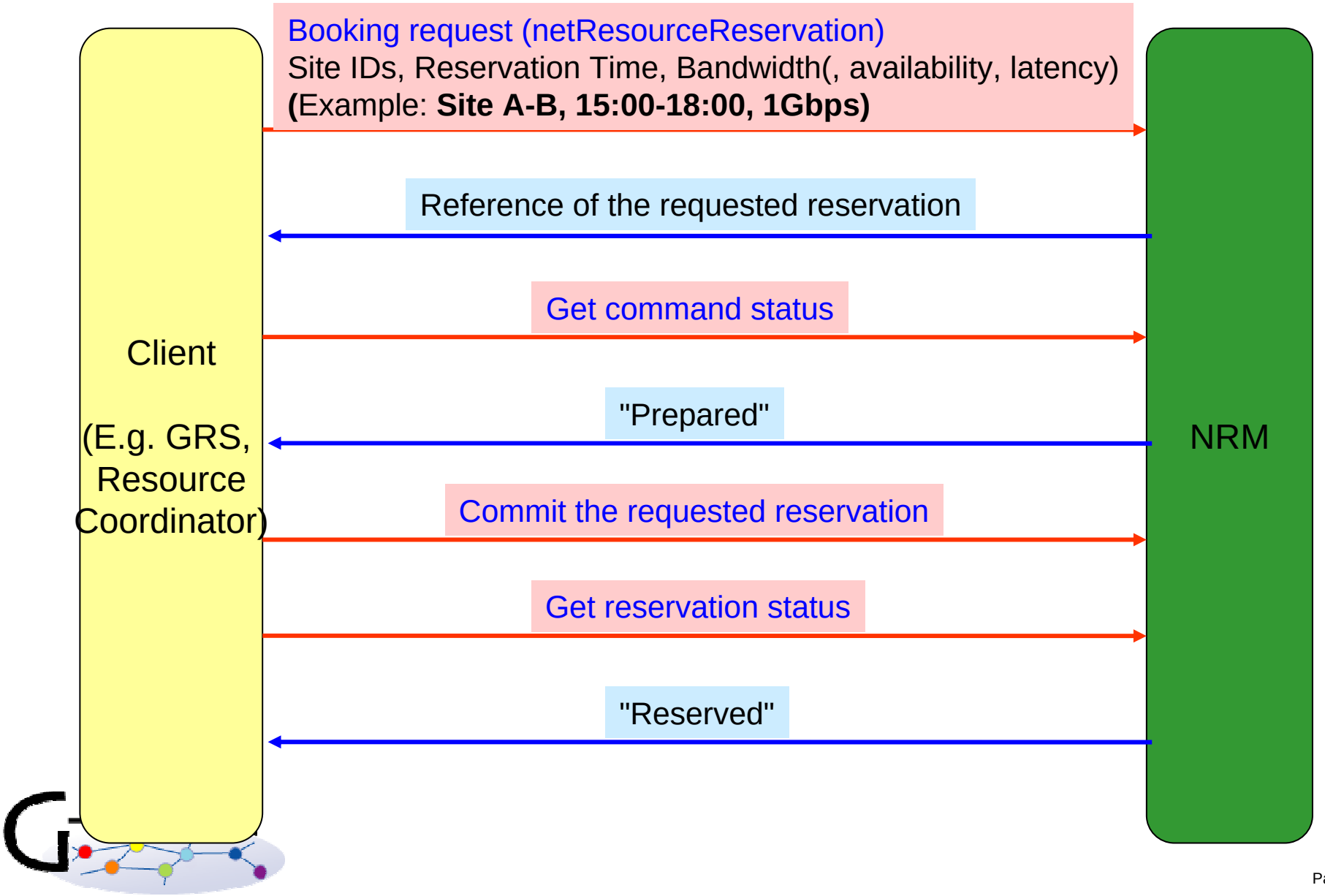
- Global Resource Coordinator
  - According to users' request, reserves computing and network resources (lambda paths) in advance
- Network Resource Manager
  - Responses to the requests from GRS through GNS-WSI
  - Manages reservation database
  - Hides detailed implementation. Provide required bandwidth between end points. (Path virtualization)
  - When the reserved time arrives, activate paths

GNS-WSI (Grid Network Service / Web Services Interface)

- Grid Network Service-Web Services Interface
- Interface to realize advance reservation of bandwidth
- Based on the Web Services interface technology
- Can be used for inter-domain coordination
- Polling-based operations
  - Advance reservation of a path between end points
  - Modification of reservation (i.e. reservation time or duration)
  - Query of reservation status
  - Cancellation of reservation
- GNS-WSI2
  - WSRF(Web Services Resource Framework) based interface
    - GT4 (Globus Toolkit 4) Java WS Core http://www.globus.org/toolkit/
  - 2-phase commit



An example XML exchanged through GNS-WSI



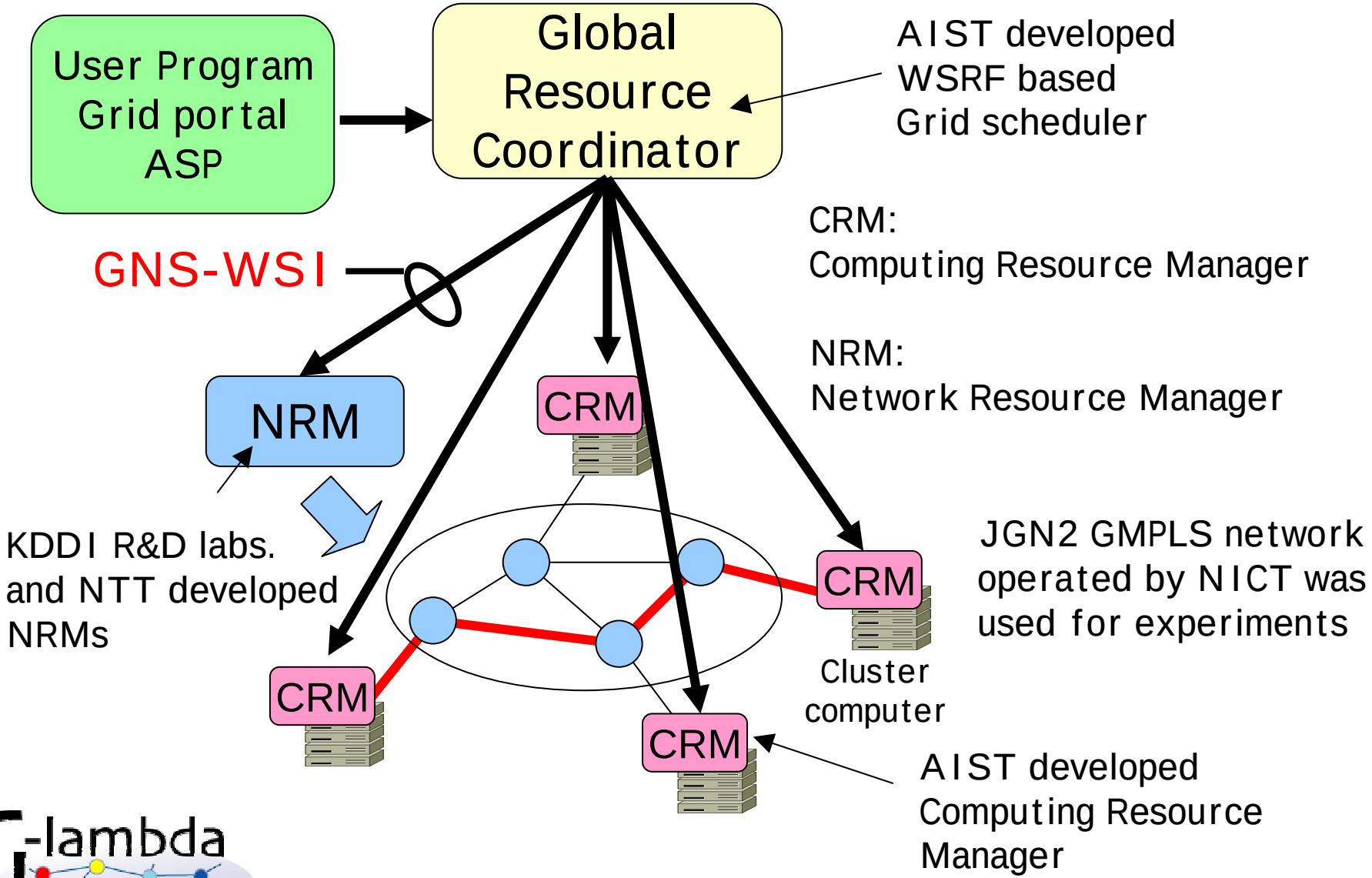
Service Parameters

Parameter	Usage	Value	Remarks
Site ID (APoint, ZPoint)	ID to specify A and Z points	String	Name or ID of sites
bandwidth	Bandwidth of the resource	Positive integer (kbit/s)	
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 (startTime, endTime)	Start time and end time of the reservation	xsd:dateTime	YYYY-MM-DDTHH:MM:SSZ
localUsername	user name of certificate	String	GT4 GSI
reservationStatus	status of reservation	String	Created/Reserved/Activated/Released/Error
commandStatus	status of each command	String	Initial/Prepared/Committed/Aborted
resourceStatus	status of network resource	String	*Available* / *NotAvailable*

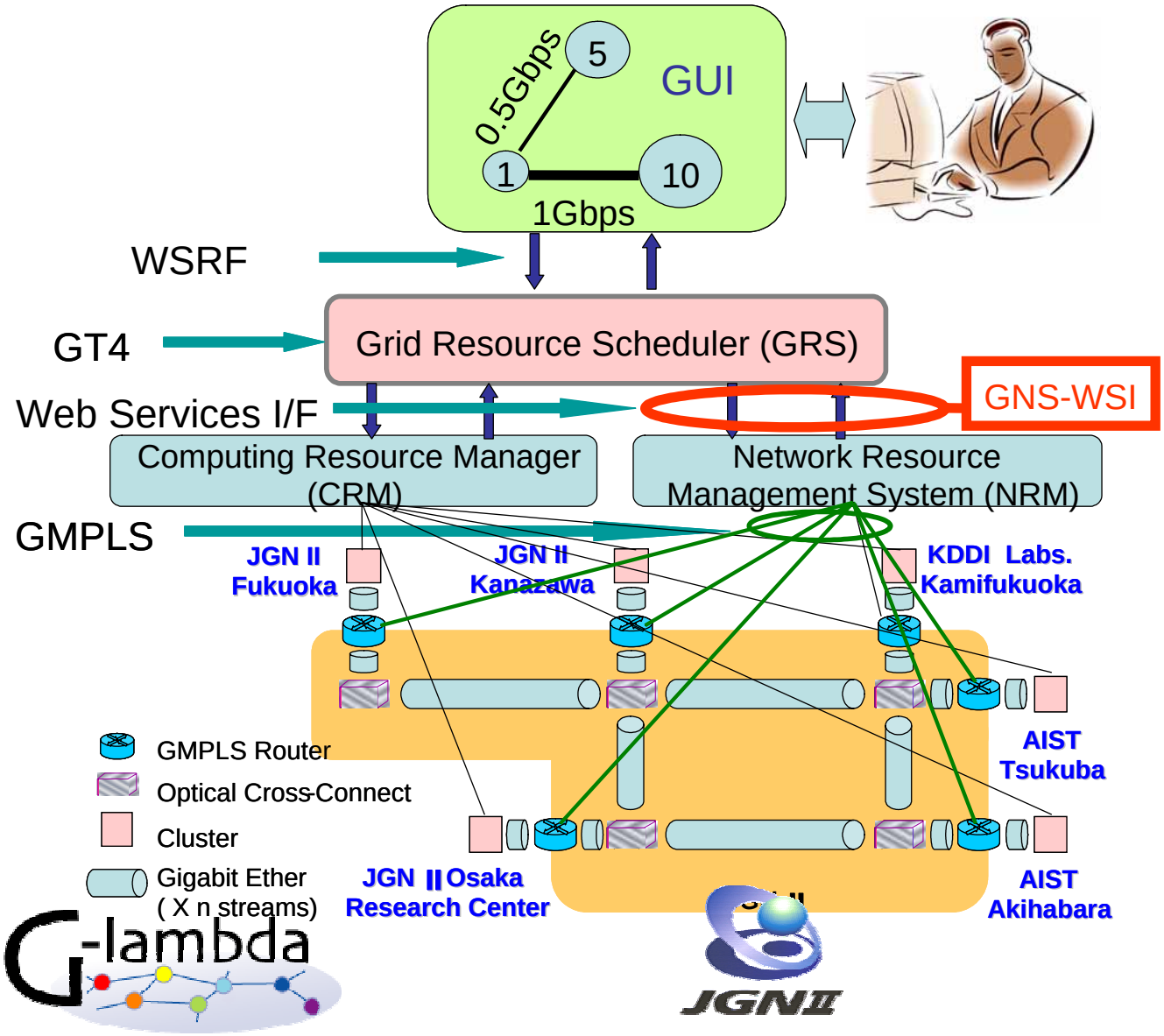
Outline

- In-advance bandwidth reservation and Grid
- G- lambda project: GNS-WSI and architecture
- **Experiment on a single-domain network**
- Architecture for multi-domain network
- Experiment on a multi-domain network

Experiment on a single-domain network



Overview of Demonstration at iGrid2005



① User requests service via GUI, specifying the required number of computers and the network bandwidth needed

②The computing resources and GMPLS network resources are reserved as the result of interworking between the GRS and NRM using GNS-WSI (Grid Network Service / Web Services Interface)

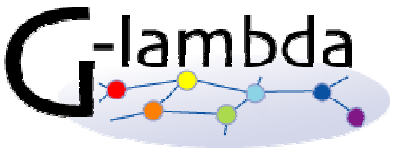
③ A molecular dynamics simulation is executed using the reserved computers and lambda paths. Ninf-G2 and Globus Toolkit 2 (GT2) are used at each cluster.

Grid Resource Scheduler (GRS)

- A Grid scheduler developed by AIST
  - Implemented using GT4 (Globus Toolkit 4)
- According to users' request, reserves computing and network resources (lambda paths) in advance
  - Accepts requests which specify required # of clusters, # of CPUs at each clusters, and the bandwidth between clusters.
  - GRS selects appropriate clusters by interworking between the NRM and multiple CRMs (Computing Resource Manager)

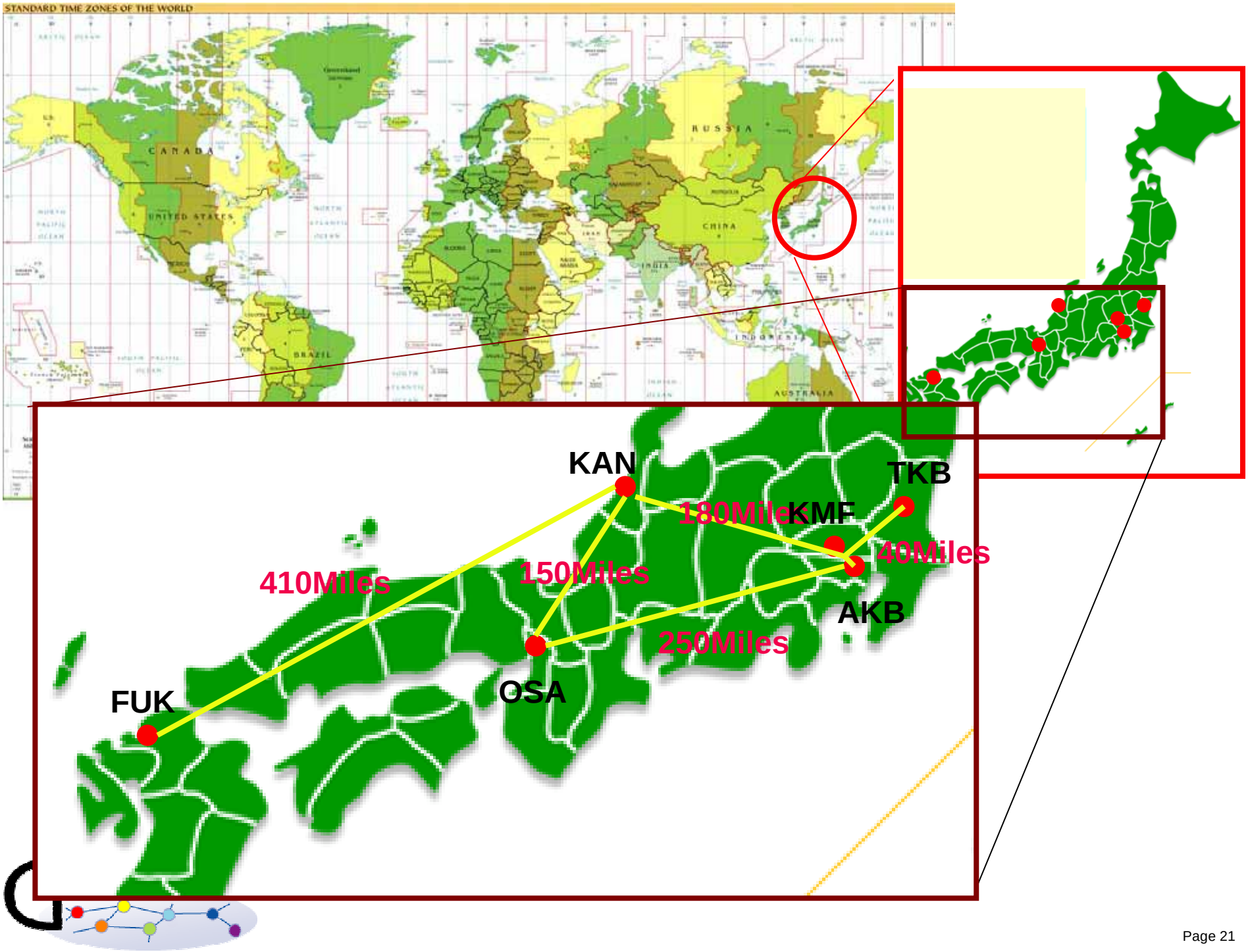
Globus Toolkit 4 (GT4)

- Globus Toolkit (GT) is one of most popular open source software toolkit for Grid.
- GT supports functions including communication, user authentication, resource management.
- Globus Toolkit 4 (GT4) is the latest version which uses Web Services technology

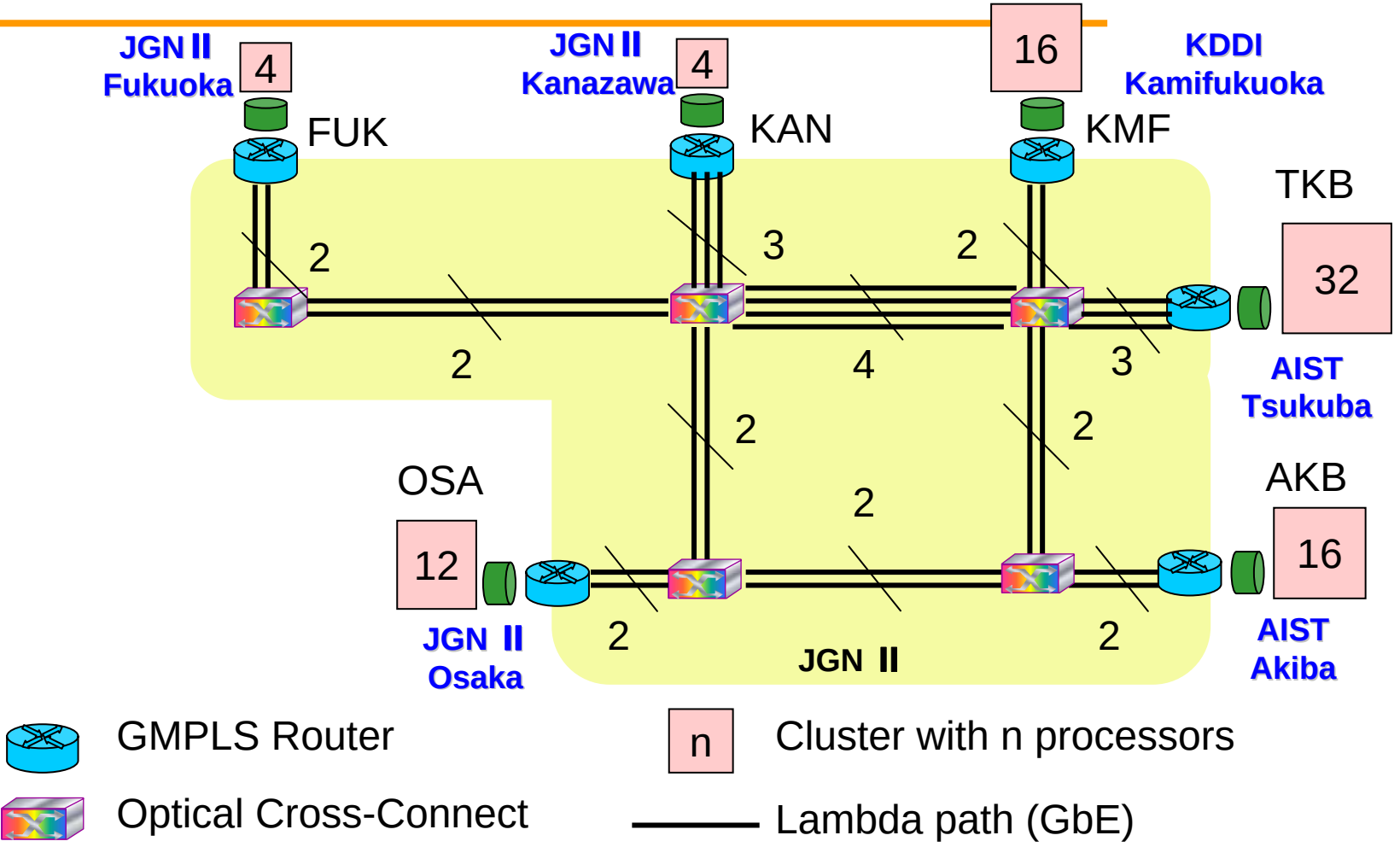


Network Resource Management System (NRM)

- NRM developed by KDDI R&D Labs. was used.
- Response to the requests from GRS through GNS-WSI
- Hide detailed path implementation. Provide a path between end points. (Path virtualization)
- Schedule and manage lambda paths. When the reserved time arrives, activate paths using GMPLS protocol.


STANDARD TIME ZONES OF THE WORLD
CANADA
UNITED STATES
RUSSIA
CHINA
BRAZIL
AUSTRALIA
KAN
TKB
KMF
AKB
OSA
FUK
180Miles
40Miles
150Miles
410Miles
250Miles

Demo Environment

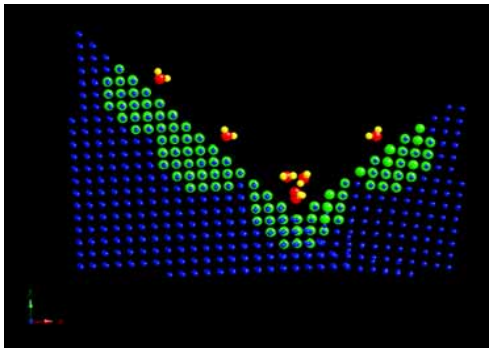


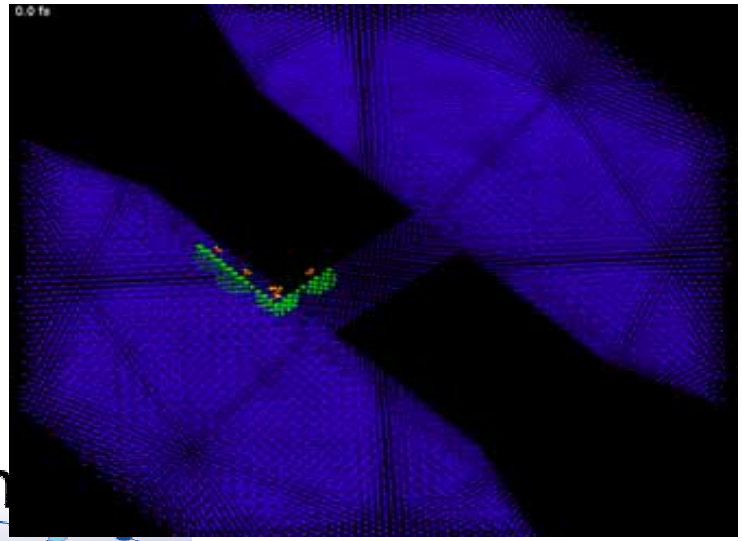
Clusters distributed over six locations in Japan are connected over GMPLS network test-bed deployed by JGN II

G-lambda

Overview of the Demo Application

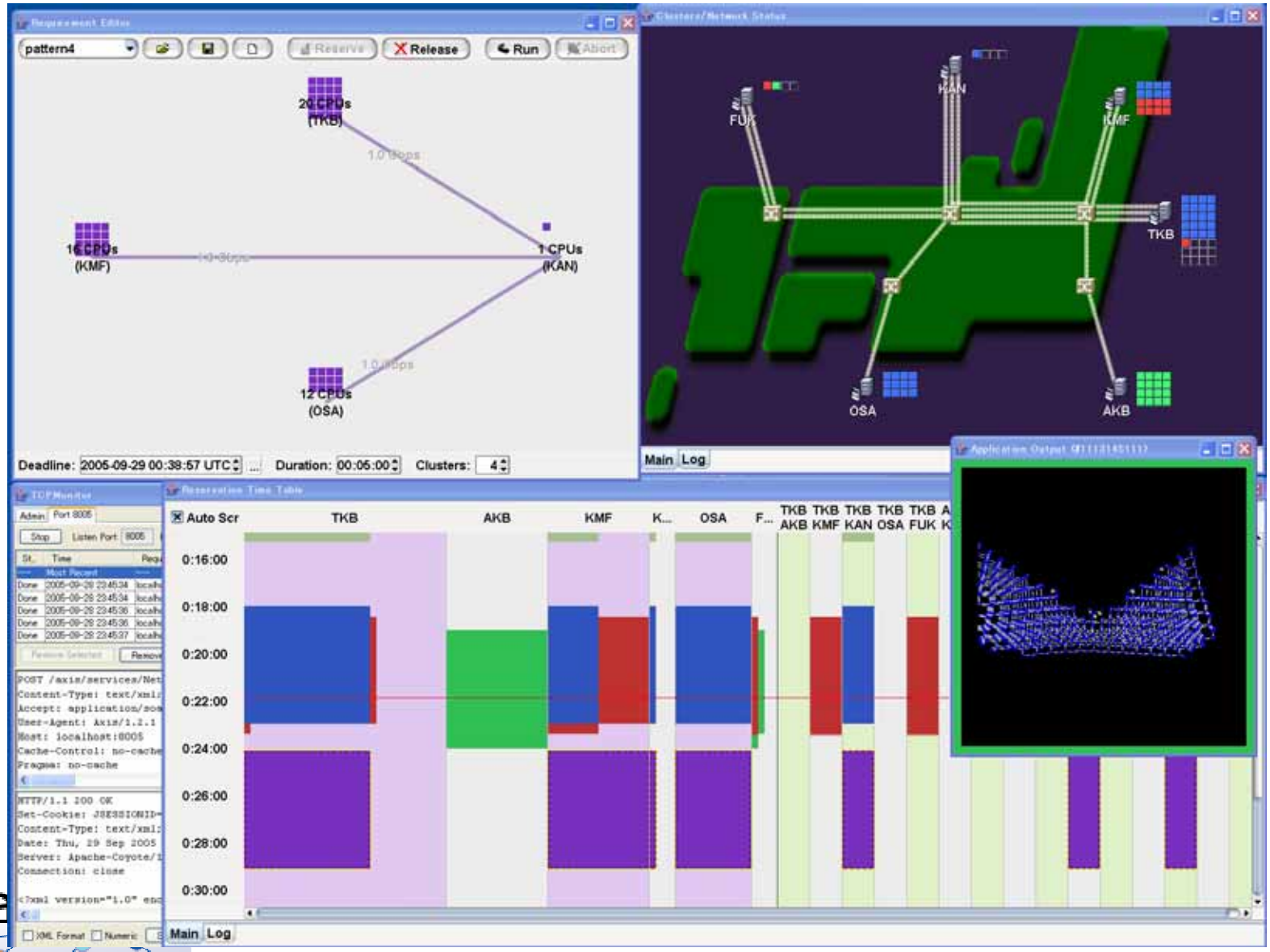
- A molecular dynamics simulation implemented with a Grid Middleware called Ninf-G2, that is developed by AIST, Japan
  - Ninf-G2 conforms the GridRPC API, a Global Grid Forum standard programming API for Grid
  - Uses Globus Toolkit 2 for job invocation and communication
- Simulation Scenario
  - Silicon and water reaction under stress





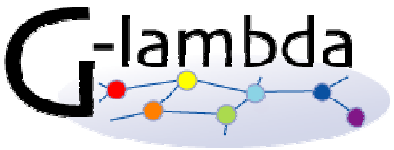
Global Grid Forum：
A standardization body for grid related technologies
Globus Toolkit：
Infra-ware for the Grid

Demonstration



Outline

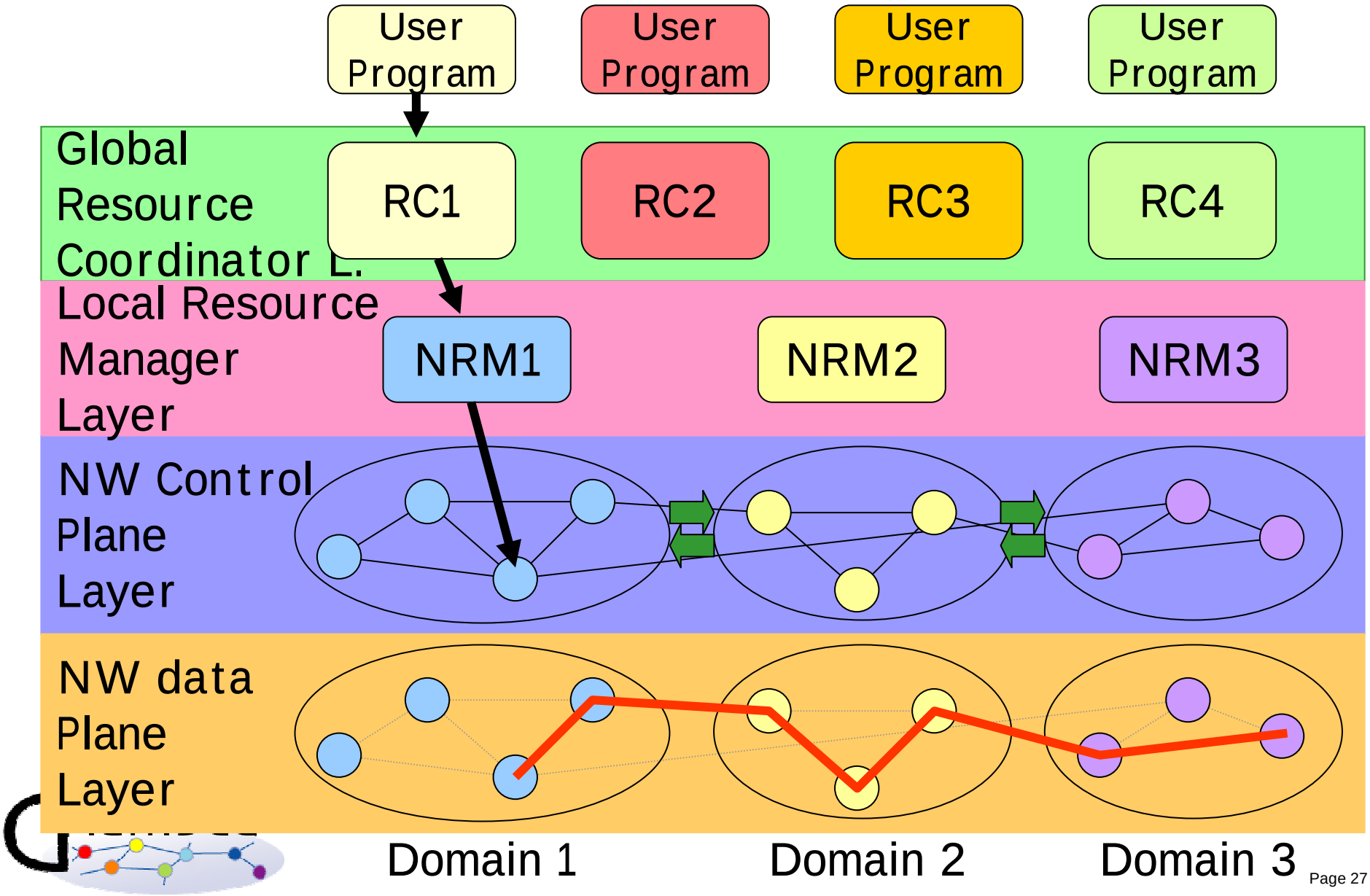
- In-advance bandwidth reservation and Grid
- G- lambda project: GNS-WSI and architecture
- Experiment on a single-domain network
- **Architecture for multi-domain network**
- Experiment on a multi-domain network



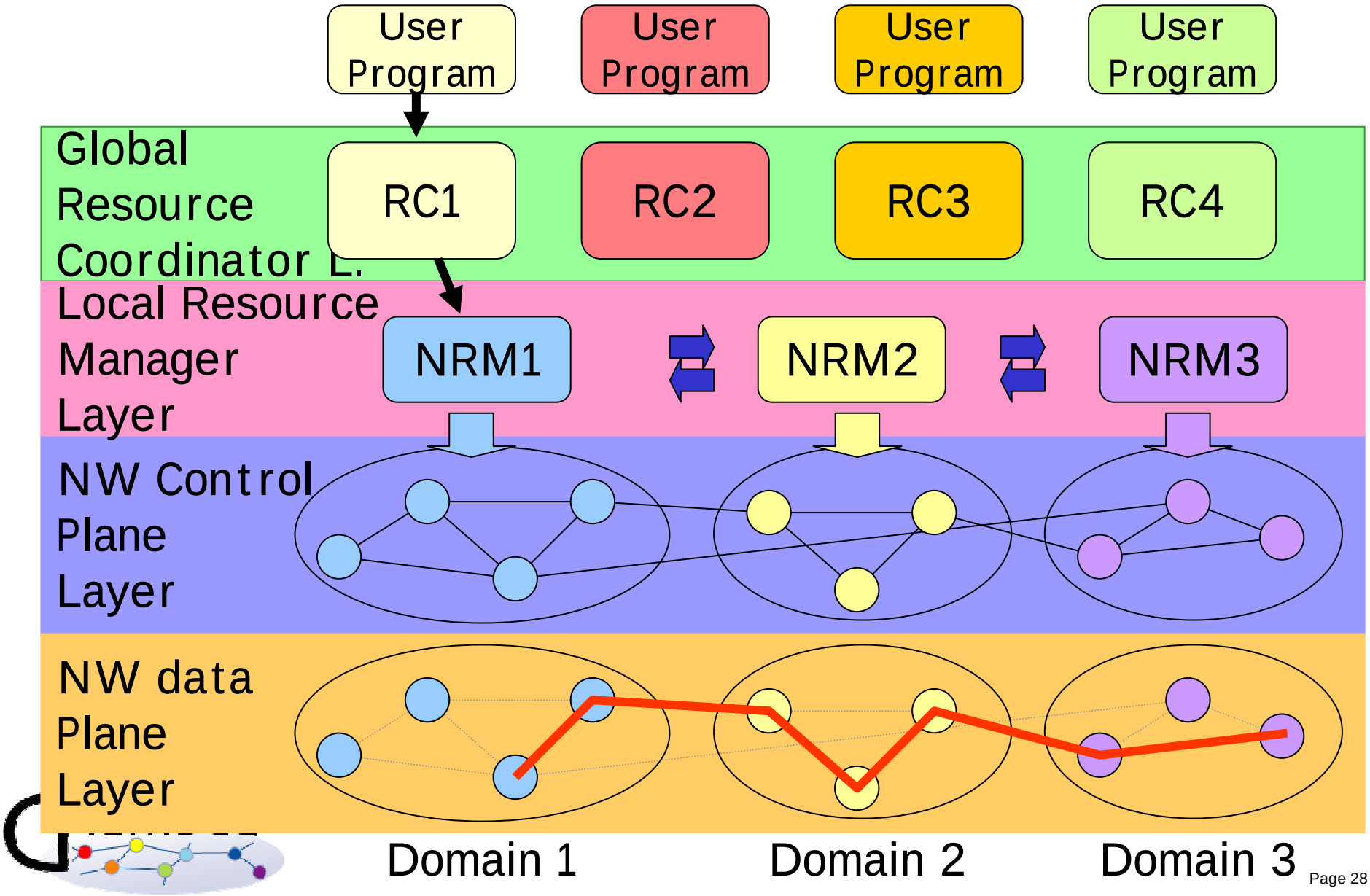
Architecture for multi-domain network

- Multi-domain
  - Network consists of multiple domains
  - Each domain is managed by its own NRM
- Bandwidth between end points in different domains
  - Coordination among domains is required
  - Three models of coordination

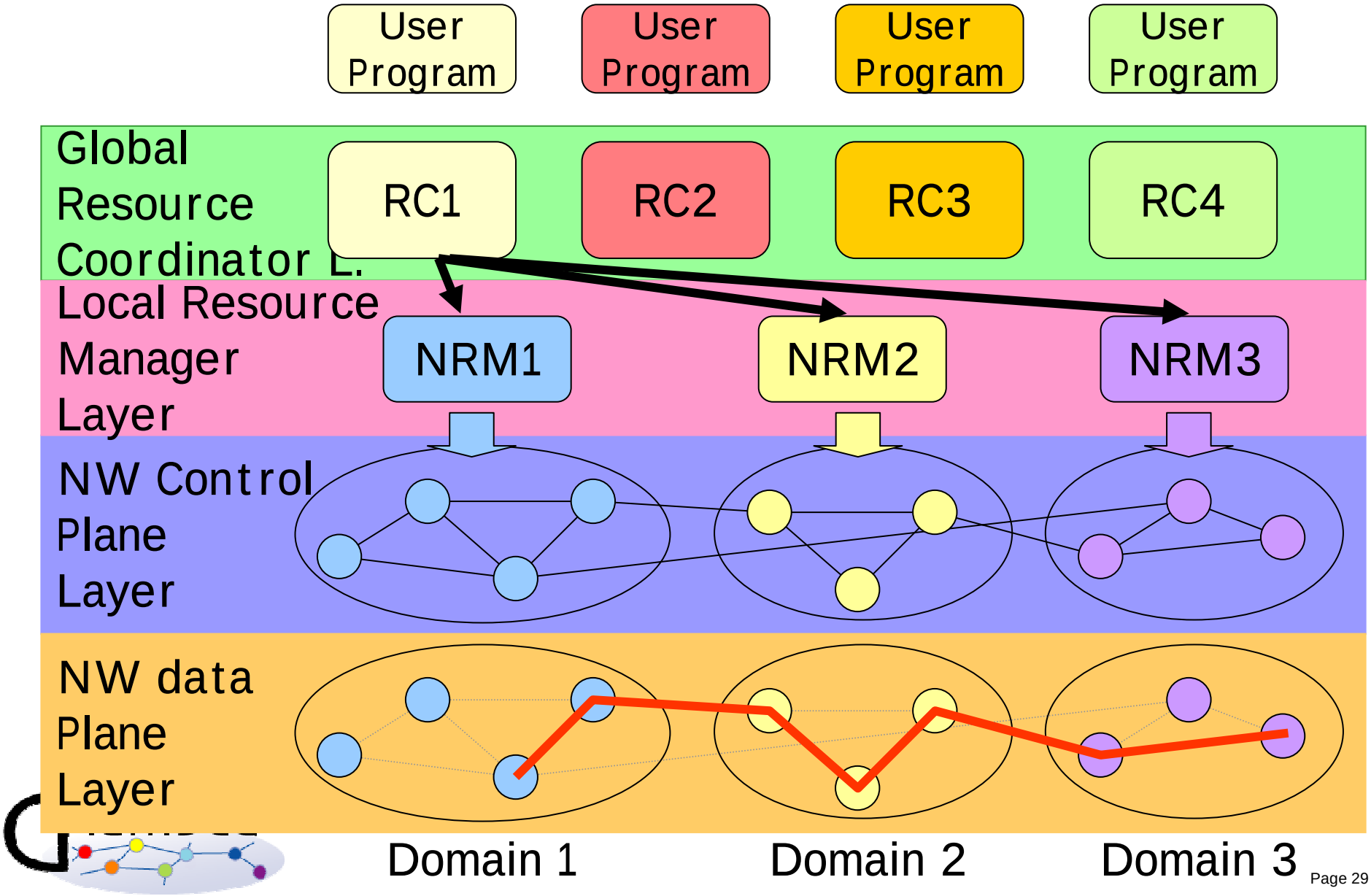
Three models of inter-domain coordination
(1) NW Control Plane Layer inter-working (ex. GMPLS E-NNI)



Three models of inter-domain coordination
(2) Local Resource Manager Layer inter-working



Three models of inter-domain coordination
(3) Global Resource Coordinator Layer inter-working



Pros and Cons of the three models

1. NW Control Plane Layer inter-working (ex. GMPLS E-NNI)
   - Pros: User do not have to care about “multiple domains”
   - Cons: GMPLS is an on-demand protocol and can not support advance reservation
   - Cons: Very close relationship between domains is required. May not be always possible for commercial service.
2. Resource Manager Layer inter-working
   - Pros: User do not have to care about “multiple domains”.
   - Cons: Requested NRM may make a reservation which is advantageous for the domain
3. Global Resource Coordinator Layer inter-working
   - Pros: User can control combination of domains
   - Pros: No under-layer interaction is required
   - Cons: User must have knowledge of inter-domain connection

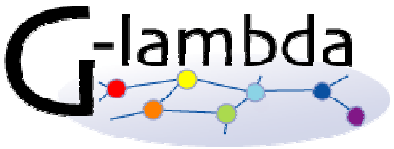
WE EMPLOYED THIS MODEL FOR INTER-DOMAIN CONNECTION

Outline

- In-advance bandwidth reservation and Grid
- G- lambda project: GNS-WSI and architecture
- Experiment on a single-domain network
- Architecture for multi-domain network
- **Experiment on a multi-domain network**

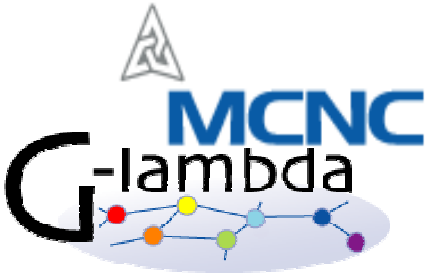
Experiment on a multi-domain network (GLIF 2006)

- G-lambda & Enlightened collaboration
- “Automated” interoperability between network and computing resources in two countries’ grid computing research testbeds is shown
  - the first such experiment of this scale between two countries
- Integrated computing and communication technology
  - Automated simultaneous in-advance reservation of network bandwidth between the US and Japan, and computing resources in the US and Japan
  - World’s first inter-domain coordination of resource mangers for in-advance reservation
    - Resource managers have different I/F and are independently developed



EnLIGHTened Project Overview

- Established in 2005, is a NSF seed-funded collaborative inter-disciplinary research initiative that seeks to research the integration of optical control planes with Grid middleware under both highly dynamic and advanced reservation application requests. Team: MCNC, LSU, NCSU, RENCI, Cisco, AT&T, Calient
- The focuses are on research and integration of cross-layer (applications, Grid resource co-scheduling, and optical network control plane) and interactions between Management , Control plane and Grid middleware.
- The goal of the Enlightened research project is establishing dynamic, adaptive, coordinated, and optimized use of networks connecting geographically distributed high-end computing and scientific instrumentation resources for faster problem resolution.



The Enlightened Team

MCNC

- Yufeng Xin
- Steve Thorpe
- Gigi Karmous-Edwards
- John Moore
- Carla Hunt
- Lina Battestilli
- Andrew Mabe
- Trevyn Leighton
- Ray Suitte
- Shane Rockriver
- Bonnie Hurst
- Avery Smith
- Syam Sundar
- Phil Misenheimer

G-lambda

cct

- Jon Maclaren
- Andrei Hutanu
- Lonnie Leger

renci
renaissance computing institute

SURA

NC STATE UNIVERSITY

- Savera Tanwir
- Harry Perros

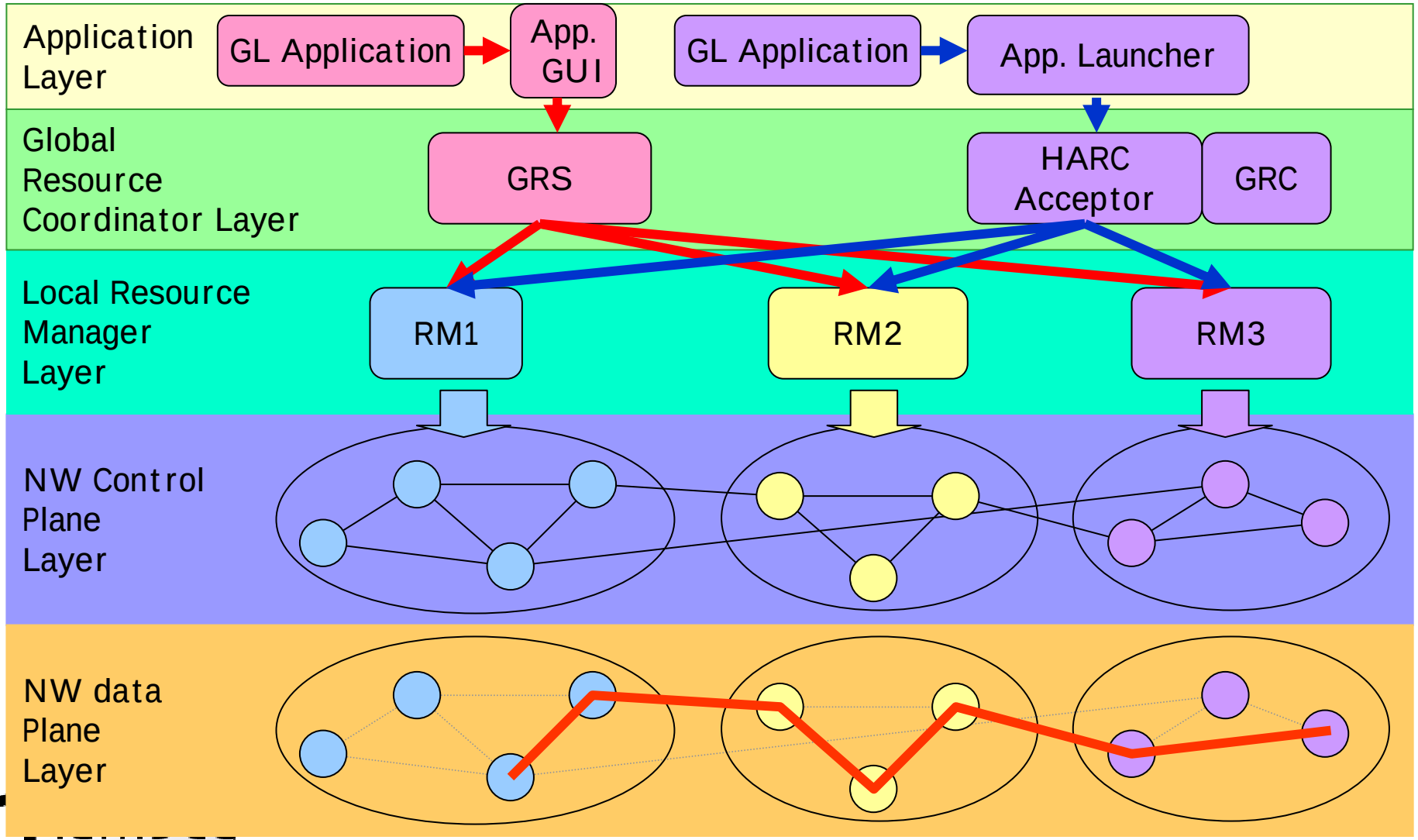
Calient NETWORKS

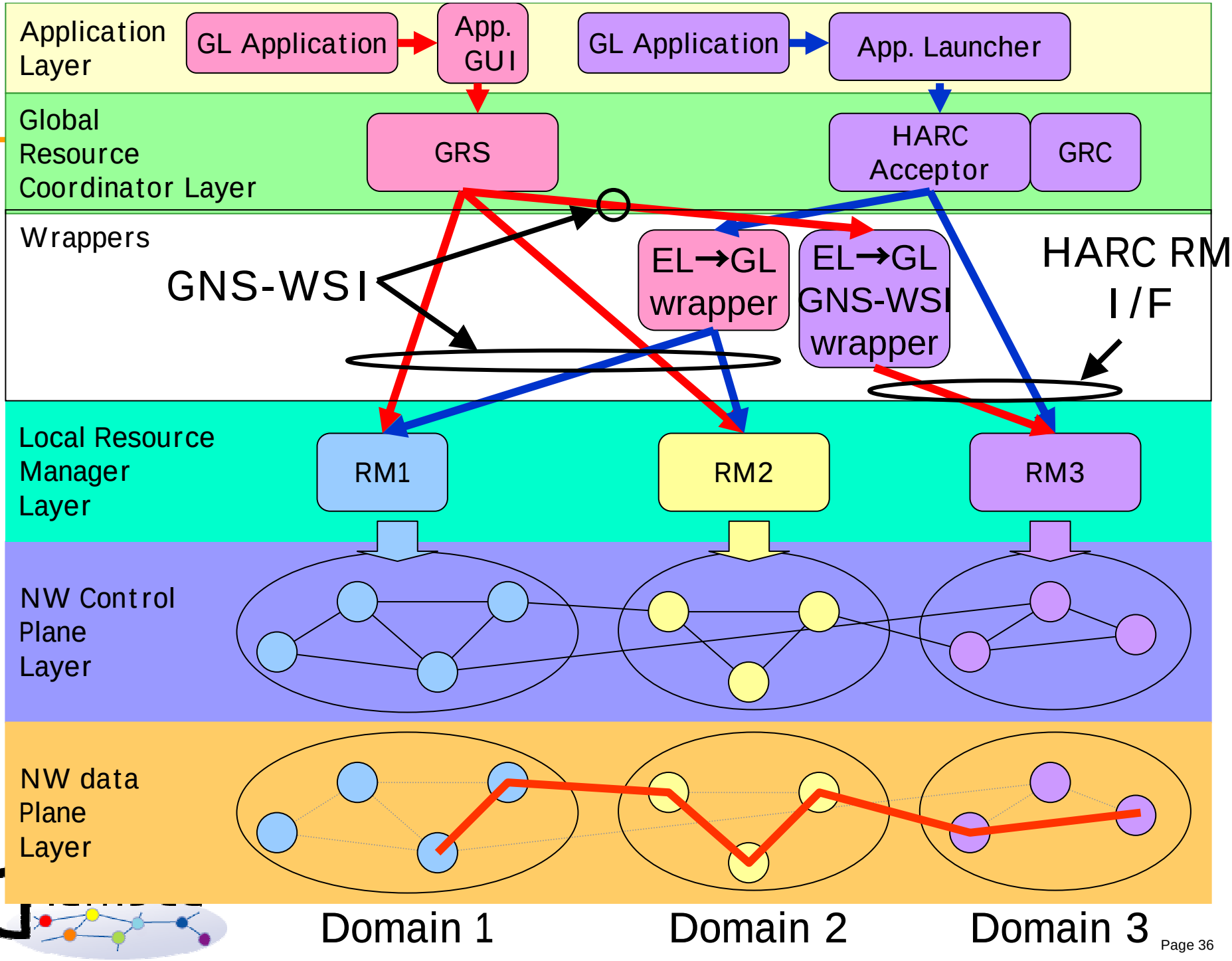
- Olivier Jerphagnon
- John Bowers

Cisco Systems
Empowering the Internet Generation

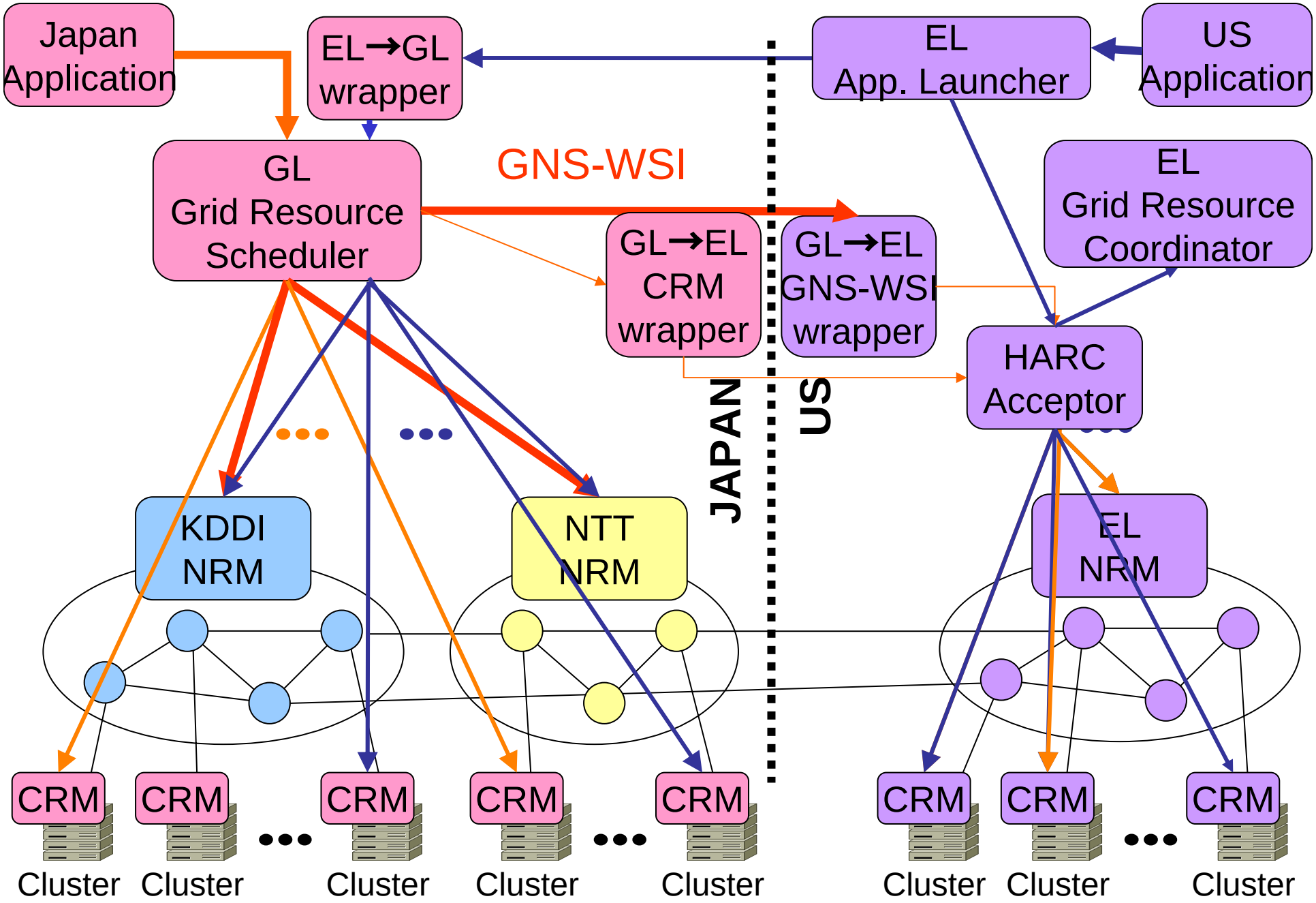
IBM

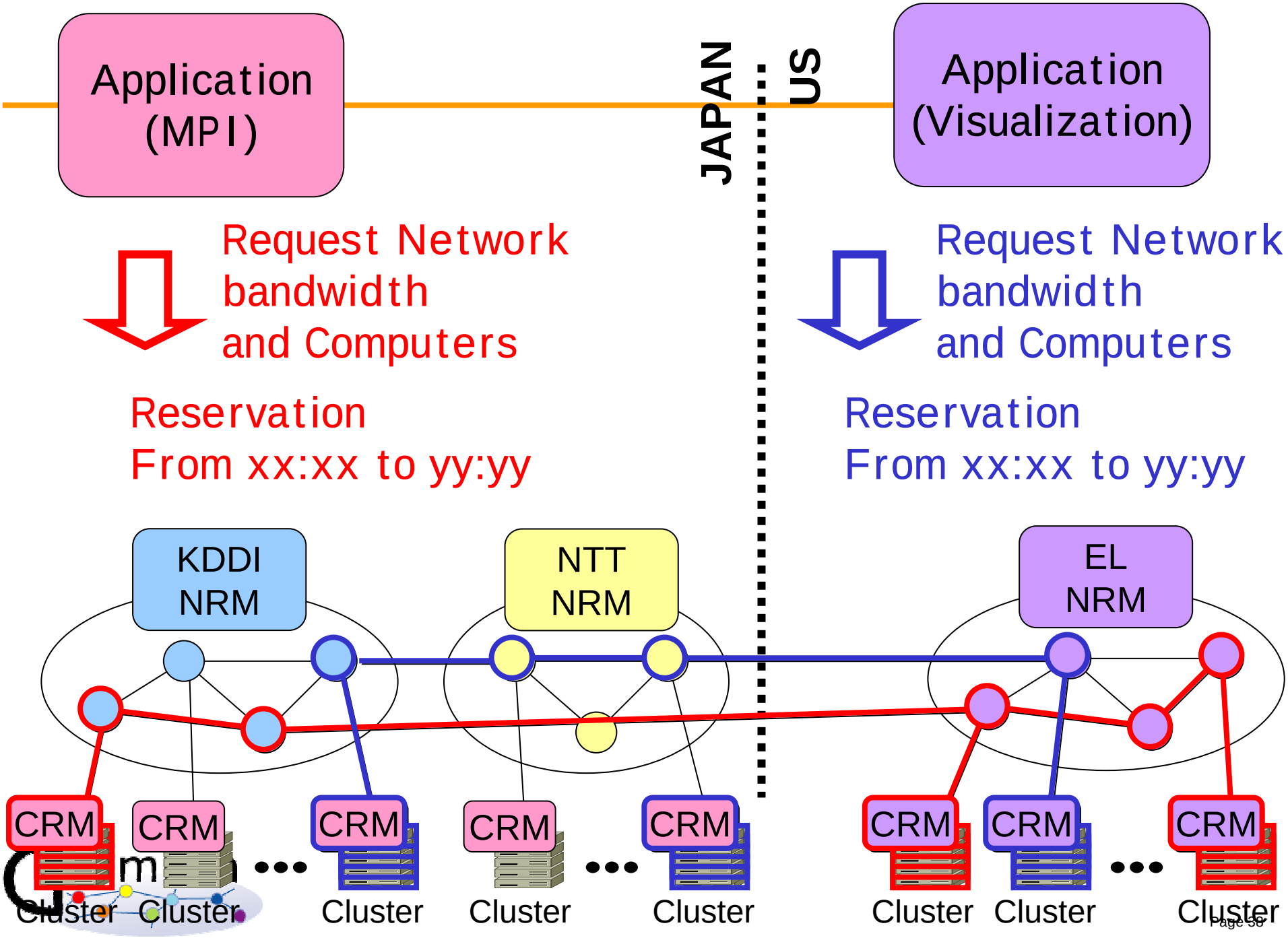
AT&T


Application Layer
GL Application
App. GUI
GL Application
App. Launcher
Global Resource Coordinator Layer
GRS
HARC Acceptor
GRC
Local Resource Manager Layer
RM1
RM2
RM3
NW Control Plane Layer
NW data Plane Layer
Domain 1
Domain 2
Domain 3

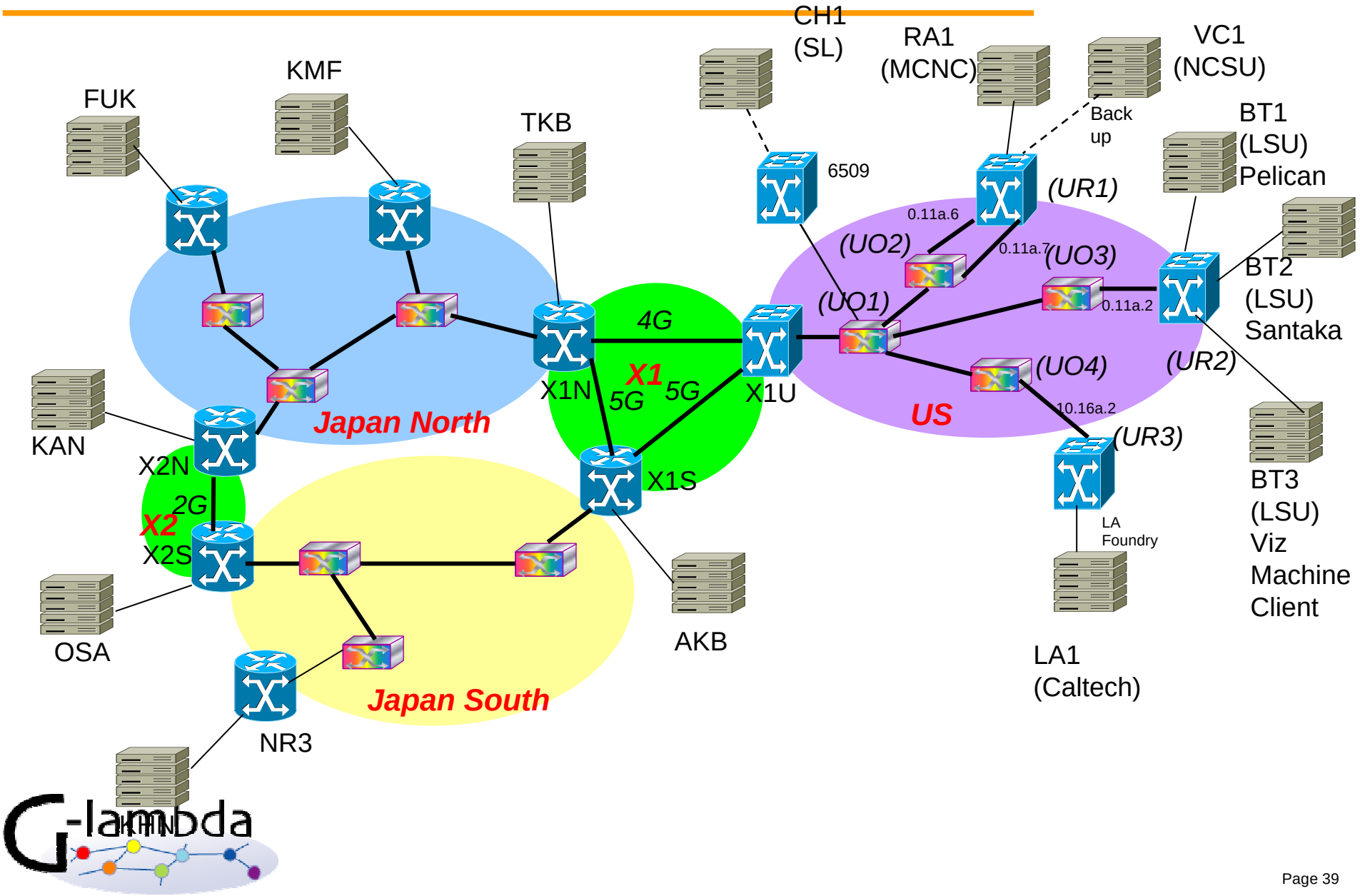

Application Layer
GL Application
App. GUI
GL Application
App. Launcher
Global Resource Coordinator Layer
GRS
HARC Acceptor
GRC
Wrappers
GNS-WSI
EL→GL wrapper
EL→GL GNS-WSI wrapper
HARC RM I/F
Local Resource Manager Layer
RM1
RM2
RM3
NW Control Plane Layer
NW data Plane Layer
Domain 1
Domain 2
Domain 3

G-lambda/Enlightened middleware coordination diagram



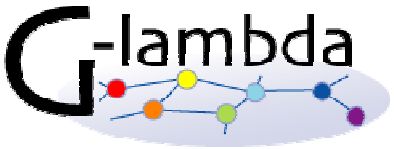

Application
(MPI)
JAPAN
US
Application
(Visualization)
Request Network
bandwidth
and Computers
Reservation
From xx:xx to yy:yy
Request Network
bandwidth
and Computers
Reservation
From xx:xx to yy:yy
KDDI
NRM
NTT
NRM
EL
NRM
CRM
CRM
CRM
CRM
CRM
CRM
CRM
CRM
Cluster
Cluster
Cluster
Cluster
Cluster
Cluster
Cluster
Cluster

Resource map



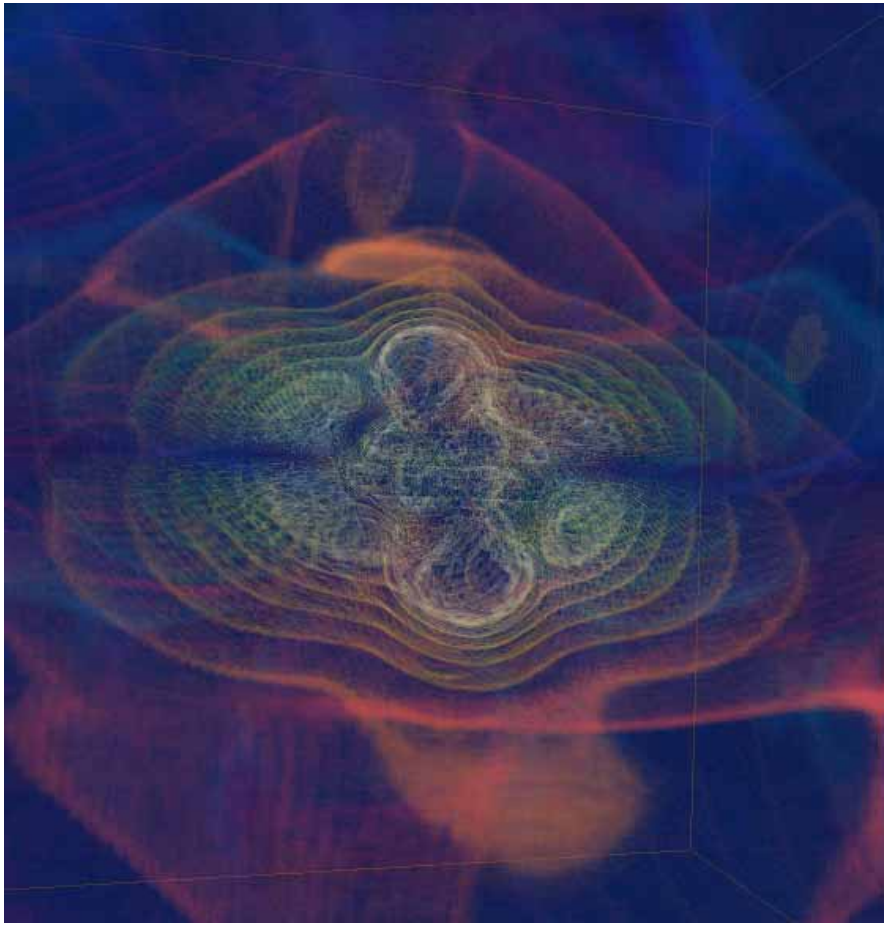
Demo overview

1. G-lambda makes a reservation. Reservation status will be shown.
2. Enlightened makes a reservation. Reservation status will be shown.
3. When the reserved time arrives, applications start running.
4. Activated paths and computing resources will be shown on MonALISA and RNDS.
   - Enlightened Viz client will show a blackhole.
5. (optional) reserve one more job from G-lambda



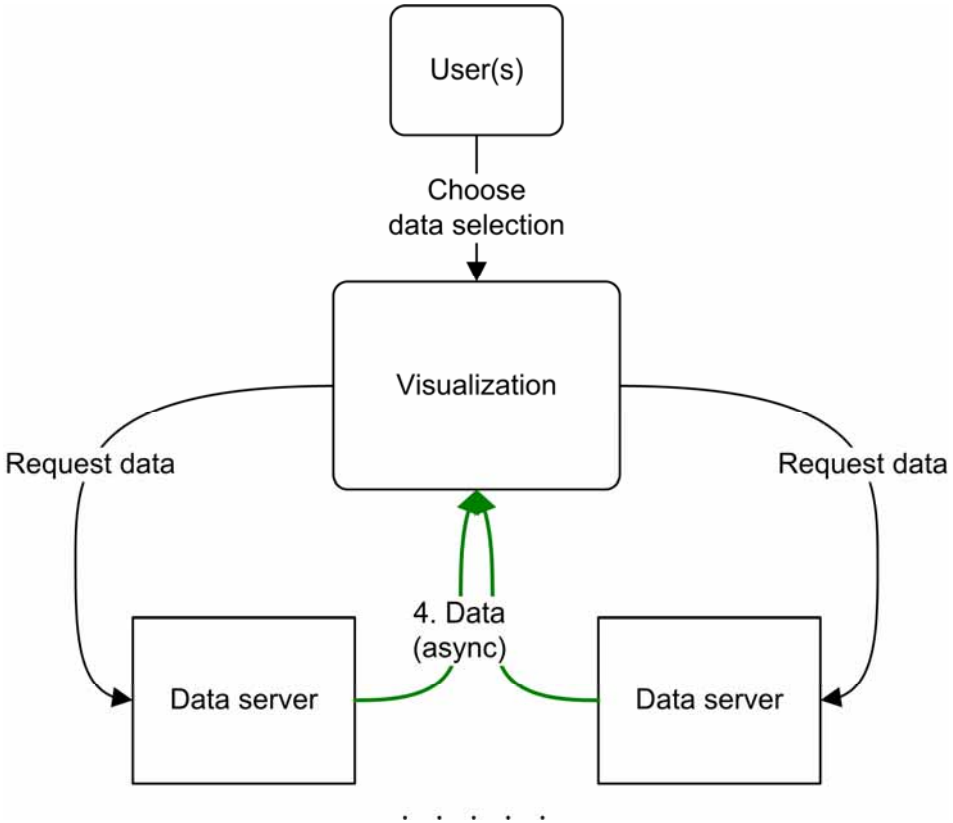
Enlightened: Visualization of remote data

- Data generated by remote simulation
- Here : a black hole simulation
- Need to explore and visualize the dataset
- Enhanced Amira visualization system to take advantage of optical networks



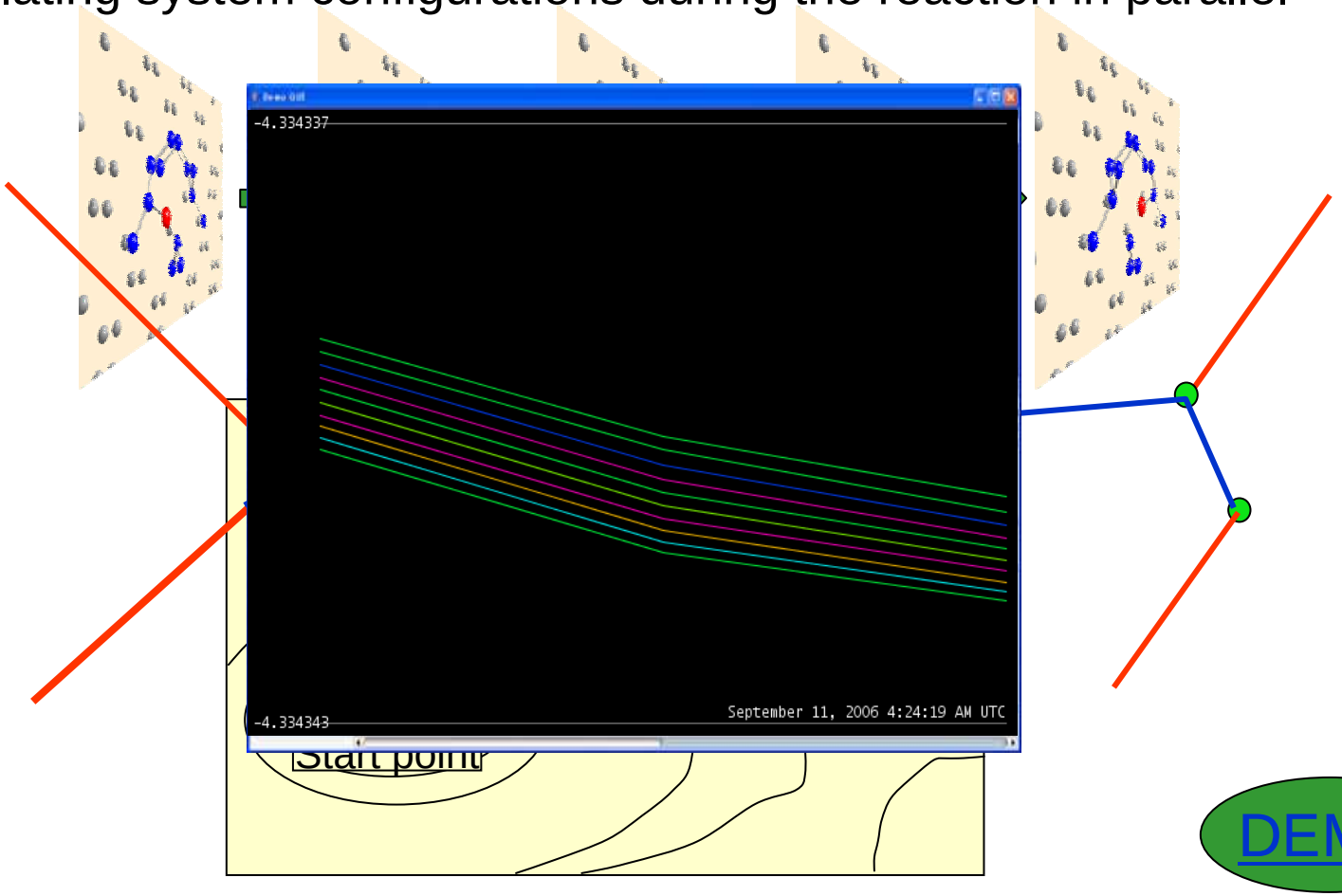
Enlightened: Distributed data server

- Data available at multiple sites
- Distribution can be beneficial (parallelism, caching options, executing simple operations)
- A distributed data server (using the optical networks) can be faster than the local disk



G-lambda: QM/MD simulation

- Surveying a chemical reaction path by Nudged Elastic Band method
  - calculating system configurations during the reaction in parallel



Thank you

G- lambda project

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