ETRI Virtualized Programmable Platform and ProtoGENI Integration

Myung-Ki SHIN
ETRI

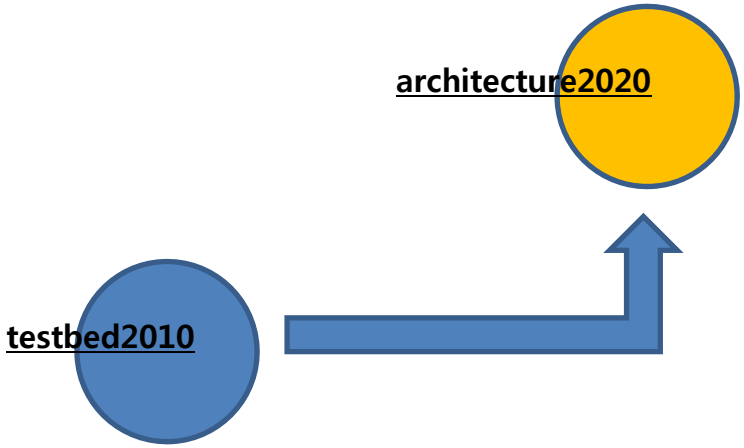
Future Internet Winter Camp 2010@SNU, Korea
Feb. 23 2010

Why Virtualized Programmable Platform ?

- The current Internet architecture is under serious reconsideration and people started thinking about alternatives.
  - Redefining Internet architecture requires many challenged works
- It's necessary to support a variety of the new different architectures to accommodate the heterogeneity of Future Internet (FI).
  - **A common means should be provided to accommodate the new heterogeneous architecture research and experiments in a shared infrastructure.**

Two Goals

- Future Internet Testbed as a short-term solution for architecture experiments
  - Running multiple experiments simultaneously in a shared experimental testbed
  - E.g., GENI
- Future Internet Architecture as a long-term solution for the future Internet
  - Virtualization, programmability, and federation will be an integral part of Future Internet Architecture
  - E,g, CABO and FP7 projects …



FiRST

Future Internet Research
for Sustainable Testbed

Future Internet Research for Sustainable Testbed

Our FIRST@ETRI Platform

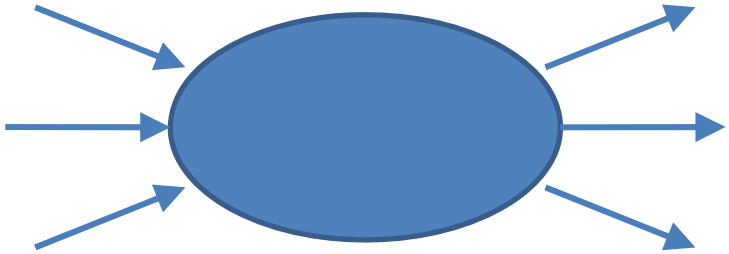
- Virtualized programmable routers
  - Researcher-defined "Silver-based Virtual Routers"
- FIRST@ETRI Platform
  - Programming APIs for Researchers
  - Common platform interface
- FIRST@ETRI Capabilities
  - End-to-end Slice Operations on our own ATCA-based, hardware platform
  - US GENI Compatible Control Framework
  - Allocate Resource Specification (Rspec) to Sliver and Link
  - Change Rspec during Operation
  - Modify Experiment Topology during Operation

Sliver-based Virtual Routers

Sliver = Packet Processing Power +
Multiple Input Ports +
Multiple Output Ports
(with Wireless capability)

〓 Virtualized Packet Processable H/W Platform which is loadable any kind of layer three^{+} packet processing software

= Router



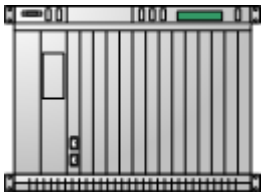
Small

Huge

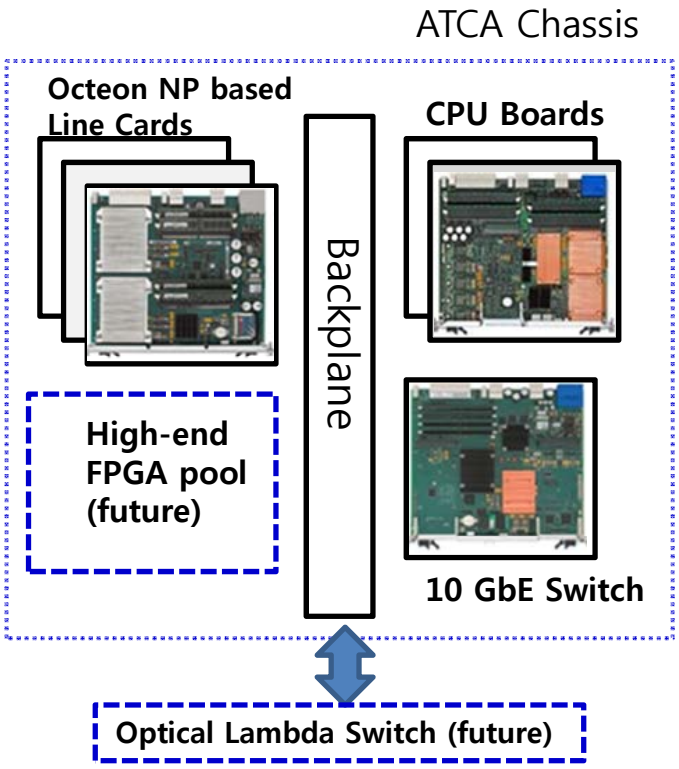
Router H/W Bare Machine ::=

- PC with Ethernet Interface,
- PC with NetFPGA,
- One board Router (Pizza box),
- Multi-board Single-Chassis Router,
- Multi-Chassis Router

Platform H/W Specification

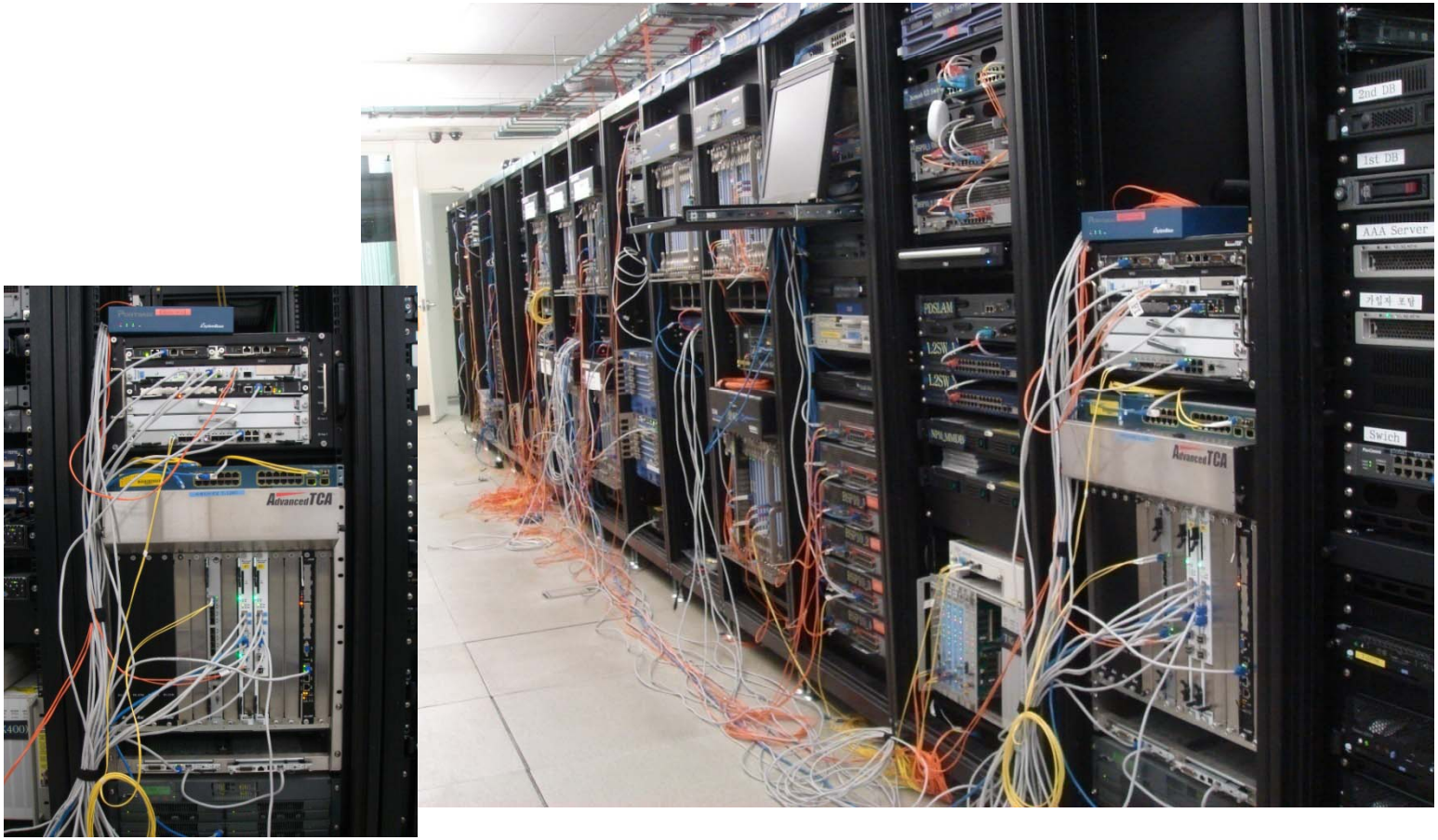


- **COTS blades – ATCA**
- **Octeon Processor**

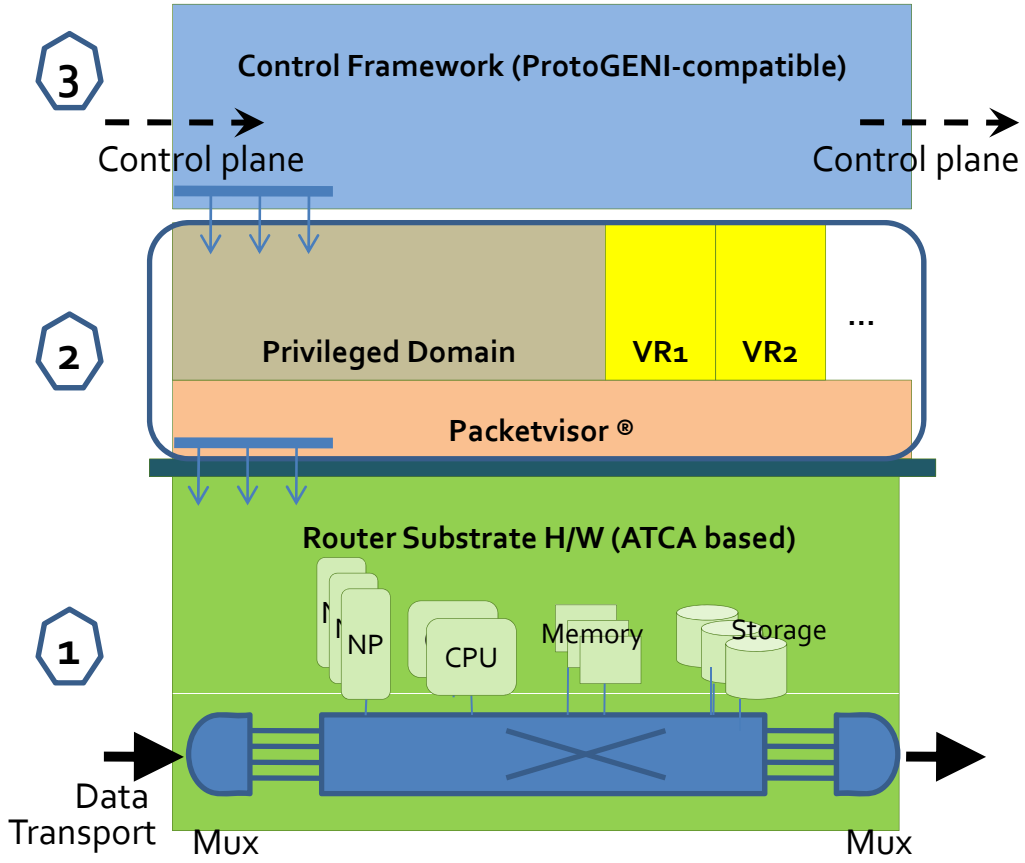


Cards	Specification
Line Card	- Dual Octeon NP 5860 - 2x 10GbE, 10x 1GbE
Ethernet-switch Card	- 16-slot 10GbE and 100/1000Base-T fabric switch - More than 100Gbps of external connectivity - Non-blocking Layer 2 switching
Processor Card	- Intel Xeon dual core - Dual 1GbE Ethernet controller - 2x 10/100/1000Base-T

FIRST@ETRI Platforms

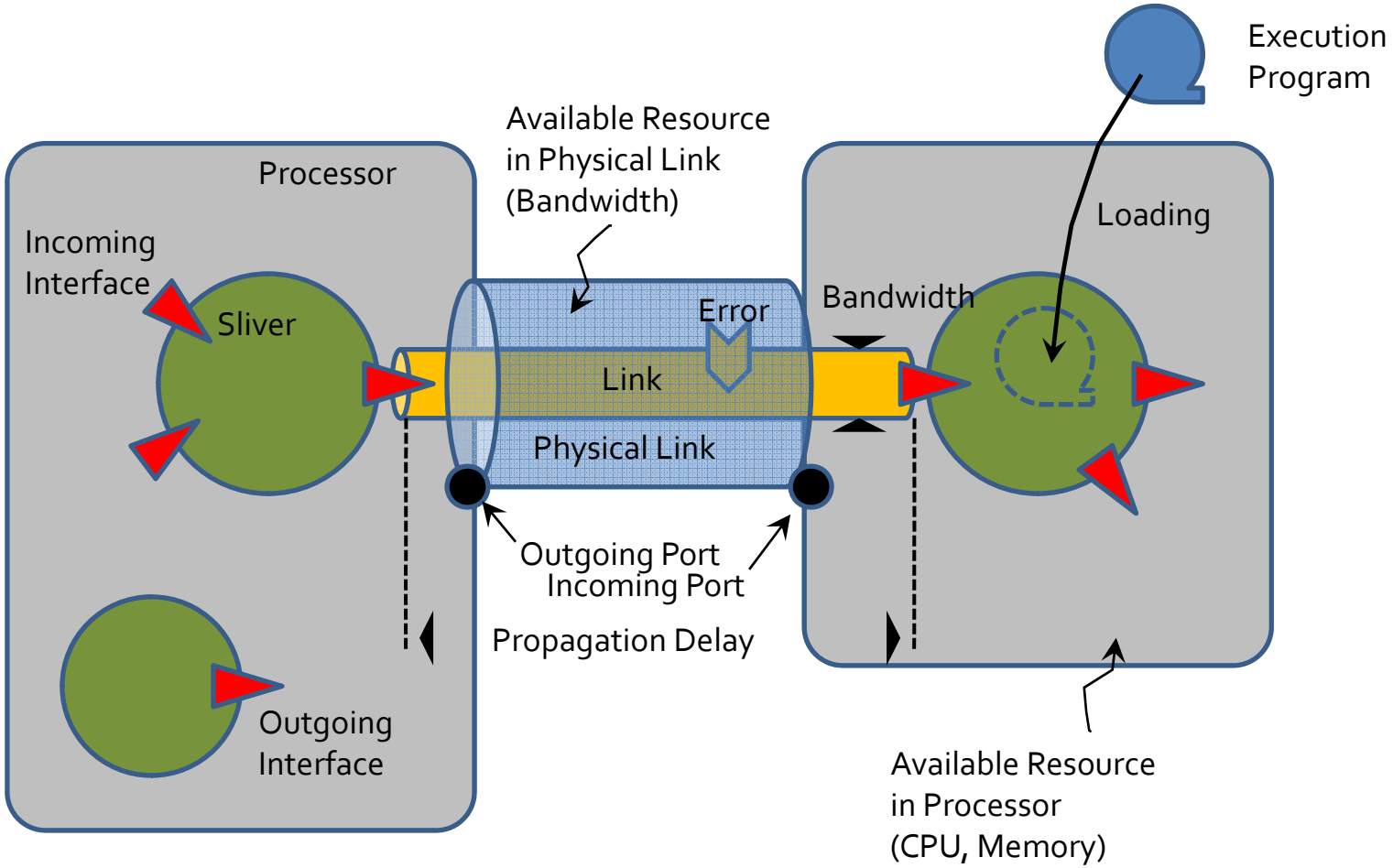


Three Layers of Our Platform

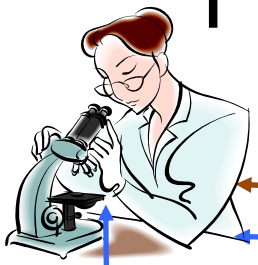


- **Control Framework**
  - International Federation
  - APIs
    - Clearinghouse APIs
    - Slice APIs
    - Node Manager APIs
  - Resource Negotiation
- **Virtualization**
  - Number of Max VMs
  - Resource Isolation
  - Scale-up Virtualization
  - Sliver Migration
- **Substrate**
  - Resource Specification
  - Resource Allocation
  - Resource Monitoring

Open Substrate Architecture



FIRST@ETRI Control Framework



1. GetCredential: SA issues self credential authenticating user to perform actions

5. AllocateSlice: User creates a new slice and receives a credential granting control over the slice

4. SelectedResourceInfo:User submit esource selection result

3. RequestResourceInfo: User submits credentials and send request to each AM for detail resource lists (Rspecs)

8. RequestTicket: User selects components, creates Rspec. If request is granted, the AM signs the request and returns a t

10. RedeemTicket: User redeems the ticket causing the sliver to be created.

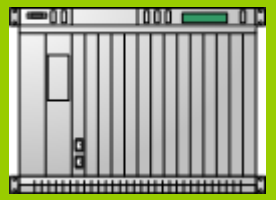
11. StartSliver: Client requests sliver to be brought to running state

FIRST@ETRI

ID Issuer

Slice Authority

Aggregate Manager



6. CreateSlice: AM creates slice and get credential

2. GetComponentManagerInfo: Requests list of all AM registered with the CH

0. GetGID: ID Issuer allocates GID

ClearingHouse

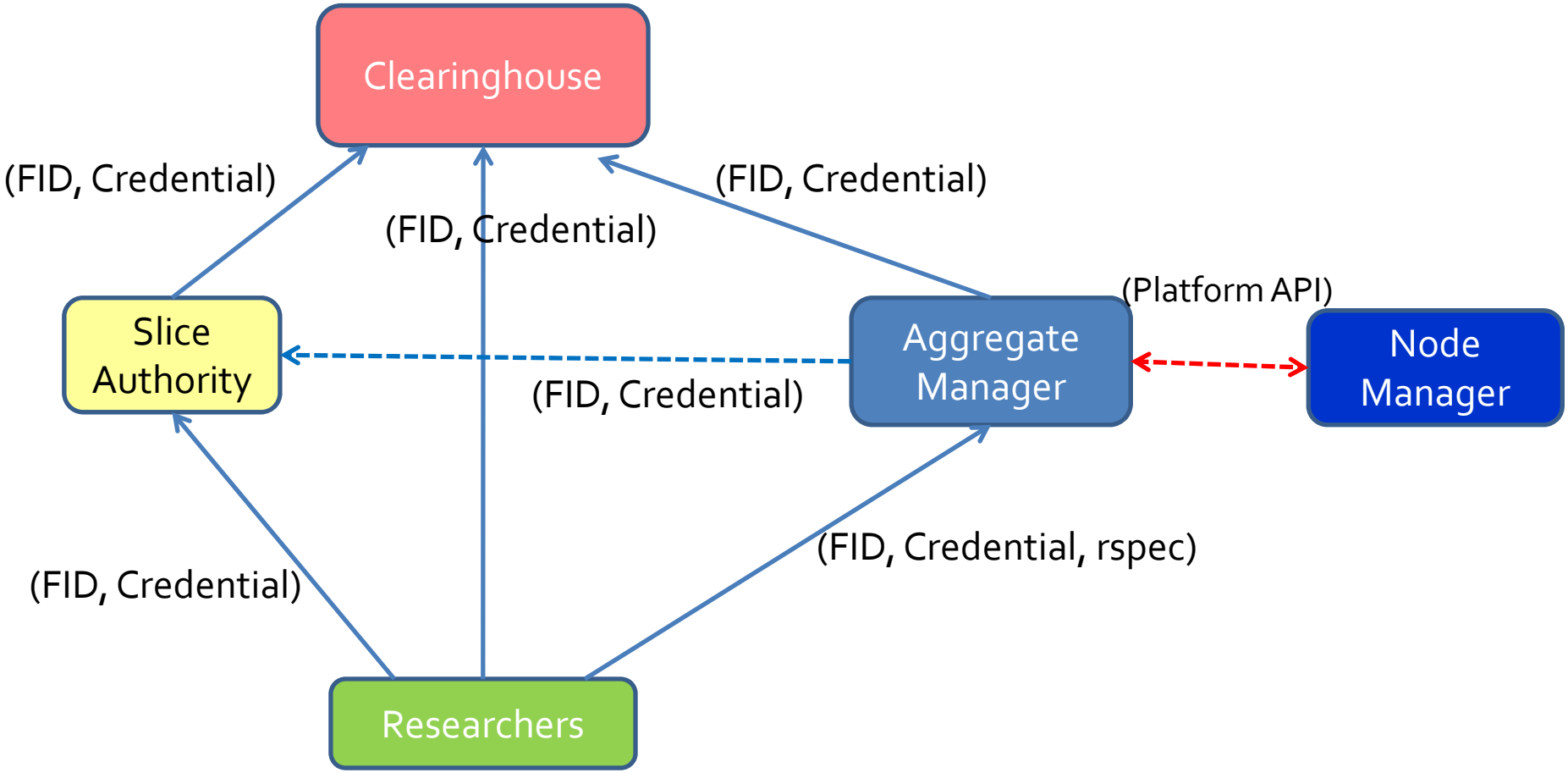
Resource Status Service

Slice & User Registry

7. Register: SA registers the user and the slice

9. AM sends copy of ticket to Slice Registry (who tracks resources in each slice).

XML_RPC relationships among control framework entities

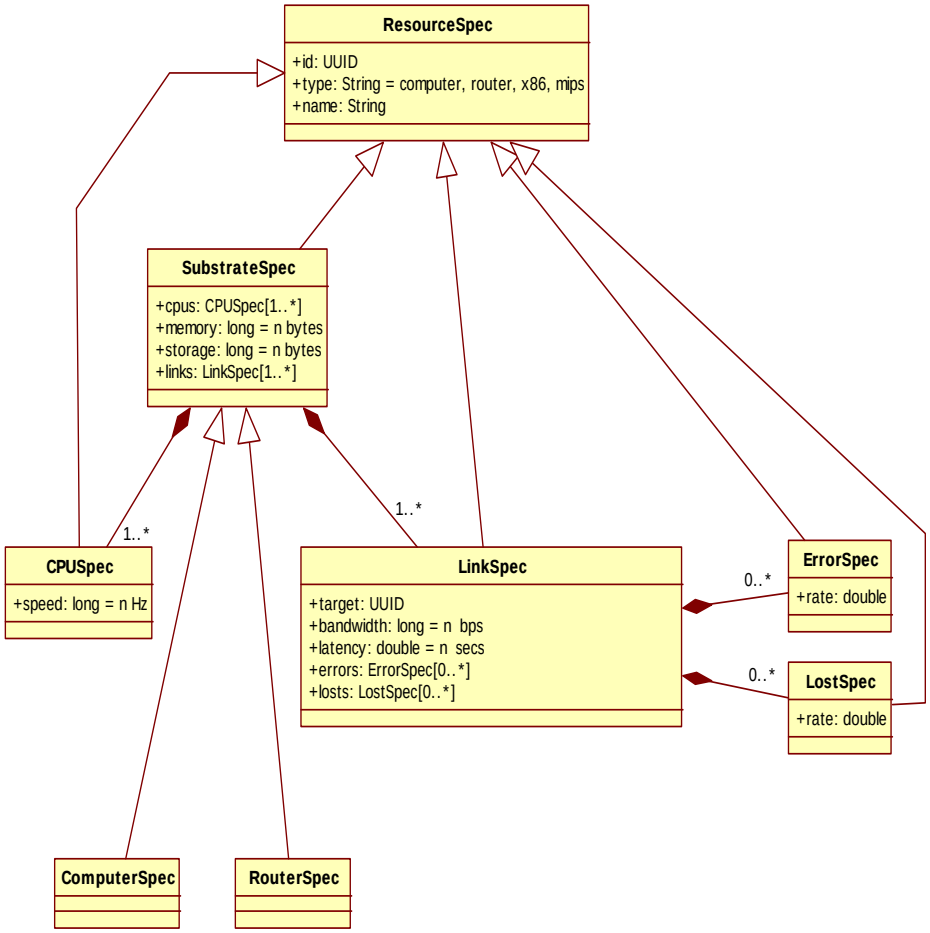


Programming APIs for Researchers

- Programming APIs for Researchers
  - E.g., To support hardware-based packet processing
    - work_request_sync() /* get_work */
    - send_packet_prepare() /* packet building */
    - send_packet_finish() /* packet sending */
    - …

Platform Interfaces

- E.g. Substrate APIs
  - allocateSliver/deallocateSliver
  - allocatePort/deallocte Port
  - allocateLink/deallocateLink
  - allocateProgram/de…
  - getSliverStatus
  - getPortStatus



FIRST@ETRI : Global Partners

- US GENI
  - Spiral-2 (2009.10. ~ 2010.9)
    - Integrating New Projects into the ProtoGENI Control Framework
      - University of Utah : ProtoGENI (Cluster C)
    - K-GENI: Establishment of operational linkage between GENI and ETRI/KISTI-Korea for international federation
      - Indiana University : GMOC
- EU FP7 (not fixed yet)
  - FEDERICA II (new capacities call)
  - GEYSERS (Network of the Future)
    - i2CAT (Spain)

ProtoGENI Project

- ProtoGENI Integration in Spiral-2
  - The objective is to build protoGENI control framework in ETRI platform and extend this framework for federated controls across national boundaries (e.g., GENI – Korea).
    - ETRI will develop a virtualized programmable router platform for its own, based on protoGENI control framework.
    - Since ETRI platform is an entirely different hardware platform from the one Emulab normally manages, ETRI plans to implement a ProtoGENI compatible Component Manager interface for ETRI platform.
    - ETRI researchers will be able to use a common API to create slices and allocate resources on any of protoGENI-based federated testbeds (Cluster C team) across national boundaries.

K-GENI Project

- Title
  - K-GENI : Establishment of operational linkage between GENI and ETRI/KISTI-Korea for international federation
- Principal Investigator Information:
  - PI: James G. Williams, Indiana University
  - Co-PIs: Myung Ki Shin-ETRI, Dongkyun Kim-KISTI
- Scope of the work
  - Provision a dedicated international connection between Korea and Indiana University in the US to facilitate an investigation into international federation strategies for operations between the GENI Meta-Operations Center, at Indiana University, and ETRII/KISTI-Korea.
  - Support tests for methods of interoperability between GMOC and the dvNOC system.
  - Develop an external networking report to help guide other GENI projects with future external connectivity.

GENI Spiral 1+2 Project List

Project Name	Project Lead	Project Participants
1. CMUlab	Carnegie Mellon University	
2. D Meas, LEARN	University of Houston	Columbia University
3. Digital Object Registry	Corporation for National Research Initiatives (CNRI)	
4. CLOUD-CTL, DOME, ViSE	University of Massachusetts Amherst	
5. DTunnels	The Georgia Institute of Technology	
6. EnterpriseGENI, OpenFlow	Stanford University	Princeton University, University of California, Berkeley, Clemson University, Georgia Institute of Technology, Indiana University, Nicira Networks, Princeton University, Rutgers University, University of Wisconsin, University of Washington
7. GENI4YR	Langston University	
8. GMOC, netKarma, K-GENI	Indiana University	
9. GpENI	University of Kansas	Kansas State University, University of Nebraska-Lincoln
	The University of Missouri-Kansas City	
10. GushProto	Williams College	UC San Diego
11. INSTOOLS, ISM Infrastructure	University of Kentucky	
12. KANSEI, OTM	Ohio State University	Wayne State University
13. MAX	University of Maryland	
14. MeasurementSys	University of Wisconsin-Madison	Boston University, Colgate University
15. MillionNodeGENI, Security	University of Washington	
16. ORBIT, WiMAX	Rutgers University	UCLA, Los Angeles, CA; University of Colorado, Boulder, CO; University of Massachusetts, Amherst; University of Wisconsin, Madison, WI
	Columbia University, NY, NY	
	Polytechnic University of NYU, Brooklyn, NY	
17. ORCA/BEN	The Renaissance Computing Institute (RENCI)	Duke University
18. PlanetLab, Scaffold, Federation	Princeton University	Universite Pierre et Marie Curie (UPMC)
19. ProtoGENI	University of Utah	
20. PROVSERV	University of Arizona	
21. ERM	Columbia	
22. REGOPT	Pittsburgh Supercomputing Center (PSC)	
23. SECARCH, Distributed Identity	SPARTA, Inc.	
24. SPP	Washington University	
25. TIED	USC Information Sciences Institute	University of California, Berkeley
26. UB_OANets	SUNY Buffalo	
27. UMLPEN	University of Massachusetts Lowell	
28. CR-GENI	University of Colorado Boulder	Radio Technology Systems LLC, Rutgers University
29. CRON-T	Louisiana State University	
30. Design of Information Subs	MIT	
31. DSL, HIVE	UC Davis	Batelle, CA Labs
32. EXP-SEC	University of Alabama	
33. FPGA-RADIO	Clemson University	
34. GENI IMF	North Carolina State University	The Renaissance Computing Institute (RENCI), Columbia University
35. iGENI	Northwestern University	University of Illinois Chicago
36. LAMP	University of Delaware	Internet2
37. LEFA, Supercharged Planetlab	Internet2	Brown University
38. NLR	Cypress, CA	
39. Open--CIRRUS	HP Labs, Palo Alto	UCSD
40. OKGems	Oklahoma State University	
41. PIGEON-NET	Howard University	
42. PrimoGENI	Florida International University	
43. QUILT	The Quilt	
44. S3-GENI	Purdue University	HP Labs
45. SEC-POL	University of Illinois (NCSA)	
46. VMI	University of Alaska Fairbanks	

ProtoGENI Integration Issues

- Migration Roadmap
- API compatibility and Wrapper
- Clearinghouse
- Data Objects
- RSpec extension

Roadmap for Controls Integration

- **Step 1. National Domain**
  - **Holding our own clearinghouse**
  - **Maintain ProtoGENI-compatibility (Cluster C)**
- **Step 2. Domain (Clearinghouse) Federation**
  - **GENI (ProtoGENI, Utah) <-> ETRI Domain**
    - Engaged in an international federation trial with the GENI
    - Migrate our nodes and slices to an independent GENI authority and vice versa.

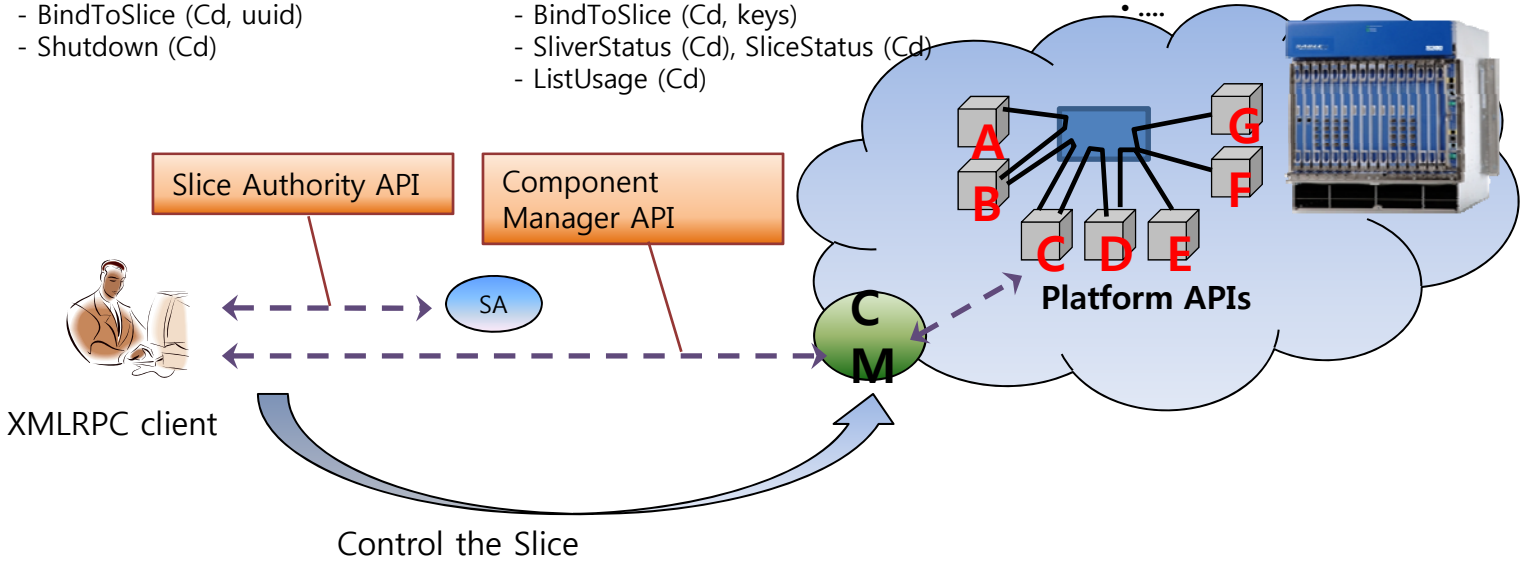
Control Integration and Wrapper

Slice Authority APIs
- **GenCredential** (Cd, uuid, type)
- Resolve (Cd, uuid, type)
- **Register** (Cd, hrn, type)
- Remove (Cd, uuid, type)
- DiscoverResources (Cd, rspec)
- GetKeys (Cd)
- BindToSlice (Cd, uuid)
- Shutdown (Cd)

Component Manager APIs
- Resolve (Cd, uuid, type)
- **DiscoverResources (Cd)**
- GetTicket (Cd, rspec)
- **RedeemTicket (ticket, keys)**
- UpdateSliver (Cd, rspec, keys)
- ReleaseTicket (ticket)
- **StartSliver (Cd)**
- DeleteSliver (Cd)
- DeleteSlice (Cd)
- GetSliver (Cd)
- BindToSlice (Cd, keys)
- SliverStatus (Cd), SliceStatus (Cd)
- ListUsage (Cd)

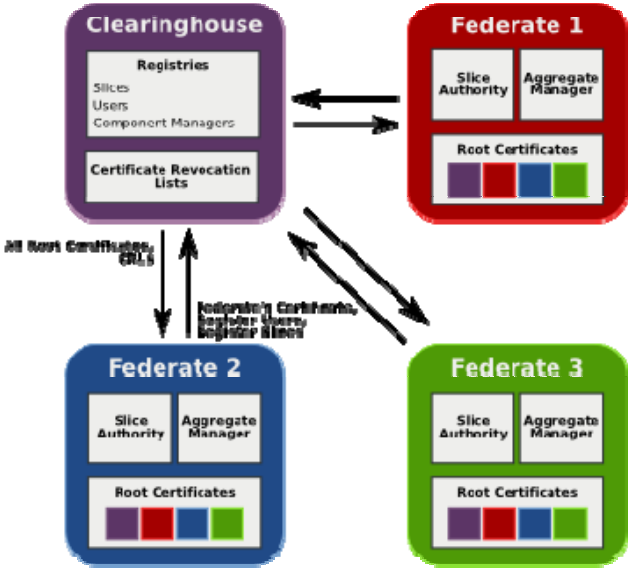
ProtoGENI mapper & wrapper

Platform APIs for Component Management
- •Create-Sliver
- •Define-Sliver-Rspec
- •Check-Sliver-Available-Rspec
- •Allocate-Sliver-Rspec
- •Delete-Sliver-Rspec
- •Delete-Sliver
- •Create-Interface
- •Define-Interface-Rspec
- •Check-Interface-Available-Rspec
- •Allocate-Interface-Rspec
- •Duplicate-Interface
- •Delete-Interface-Rspec
- •Delete-Interface
- ....



Clearinghouse- Operations

- Allows registration, deletion, and resolution of various principle objects (slices, users, slice authorities, aggregate/ components managers).
- Provides a list of all known aggregate/component managers (AM/CM).
- Exports a centralized "emergency shutdown" facility that can be used to terminate a misbehaving slice.
- Assists in the registration of new federates as they come on-line.
- Collects and distributes Certificate Revocation Lists (CRL) from all of the federates.
  - These CRL's are generated periodically by the federates so that others in the federation can be made aware of users that have been terminated or who have had to generate a new certificate (GID).



Clearinghouse - Issues

- Multiple clearinghouses
- The location (URL) of the clearinghouse is hardwired
- Authorities (SA/CM/AM) only can access the clearinghouse
  - Authorities can get a credential to access the clearinghouse in one of two ways;
    - Request a credential via XML-RPC. The credential is signed using the clearinghouse's GID (certificate).
    - Generate a self-signed credential. In this case, the authority uses its own certificate (GID) to sign the credential.
- Everything is stored in a mySQL DB, (at this time)

Data objects - Scope

- Operation of the control framework is carried out by the exchange of specialized data objects
  - Certificates (X.509)
    - To establish public key infrastructure.
  - Credentials (xml)
    - The means by which proof of privileges are conveyed.
  - Rspec
    - data interchange format
    - advertisements, requests, tickets, and manifests.

Data objects - Issues

- ProtoGENI (GENI) compatible (R1.0/v2.0)
  - All objects are named by URNs
    - ProtoGENI : a number of conventions proposed by the GMOC for GENI identifiers (GID).
    - The same (or compatible) scheme should be designed
  - Each service is invoked via XMLRPC

Rspec - Scope

- Identifying resources
  - All nodes and links are identified by a UUID
  - There is also a "human readable" name field to aid readability
- Nodes
  - Node have types
  - Virtualization technology is included as a field
- Links
  - Links are point-to-point: LANs and other "full connectivity" environments (such at the Internet), a "LAN node" is created, and all members are linked to it.
  - Links have bandwidth, a type, etc.
  - Links endpoints reference Interfaces on Nodes
- Interfaces
  - Endpoint of a link
  - Named by node, plus an opaque interface name
  - In progress: Interfaces will be first-class entities, declared as part of the component they belong to
- Metadata
  - A "valid until" field
  - A "generated" time

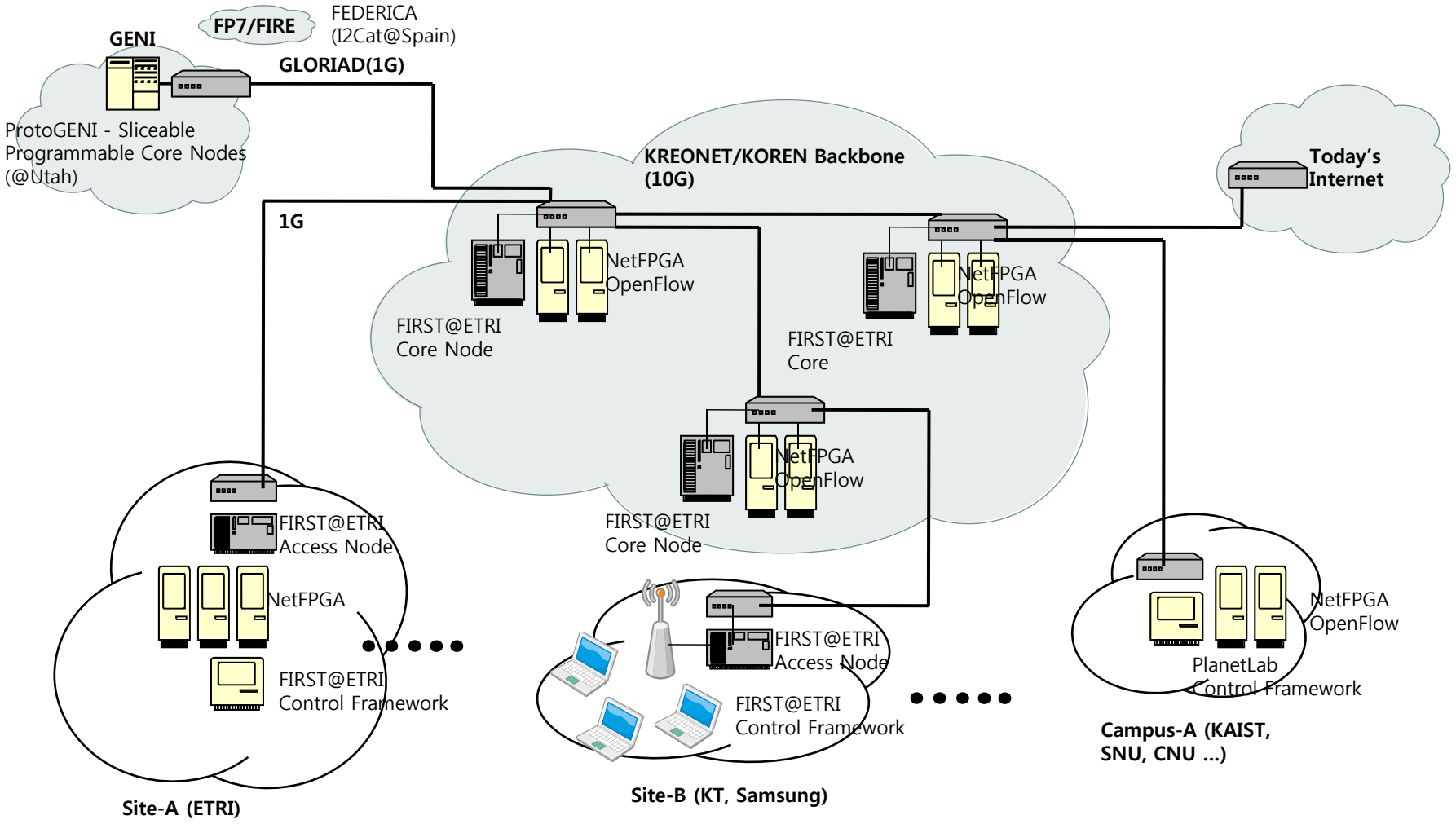
http://www.protogeni.net/trac/protogeni/attachment/wiki/RSpec/protogeni-rspec-common.rnc

Proposed URN scheme

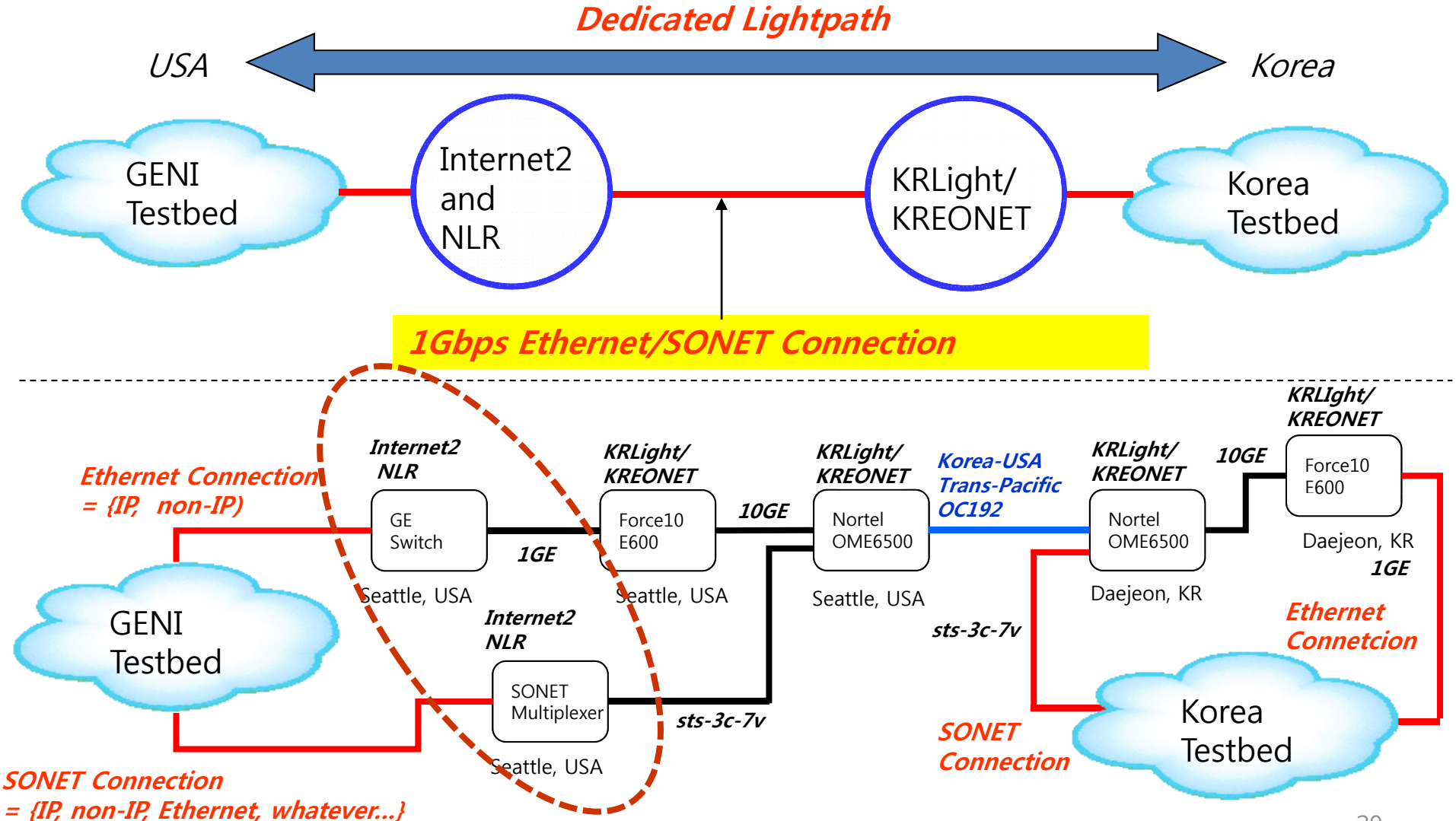
- urn:publicid:IDN+toplevelauthority+resource-type+resource-name

- E.g., Resource FIRST Identifier (FID)
  - User mkshin at the first testbed namespace
    - urn:publicid:IDN+first.kr+user+mkshin
  - FIRST tesbed node: platform1.etri.re.kr
    - urn:publicid:IDN+first.kr+node+platform1.etri.re.kr
  - Interface etho in FIRST tesbed node platform1.etri.re.kr
    - urn:publicid:IDN+first.kr+interface+platform1.etri.re.kr:etho
  - Slice mytestslice in the FIRST Korea slice authority
    - urn:publicid:IDN+first.kr+slice+mytestslice
  - The FIRST slice authority itself
    - urn:publicid:IDN+first.kr+authority+sa
  - Sliver 0002b33f75b7 in the FIRST component manager
    - urn:publicid:IDN+first.kr+sliver+0002b33f75b7
  - A ticket issued by the FIRST component manager
    - urn:publicid:IDN+first.kr+ticket+456

Korea Testbed and Backbone

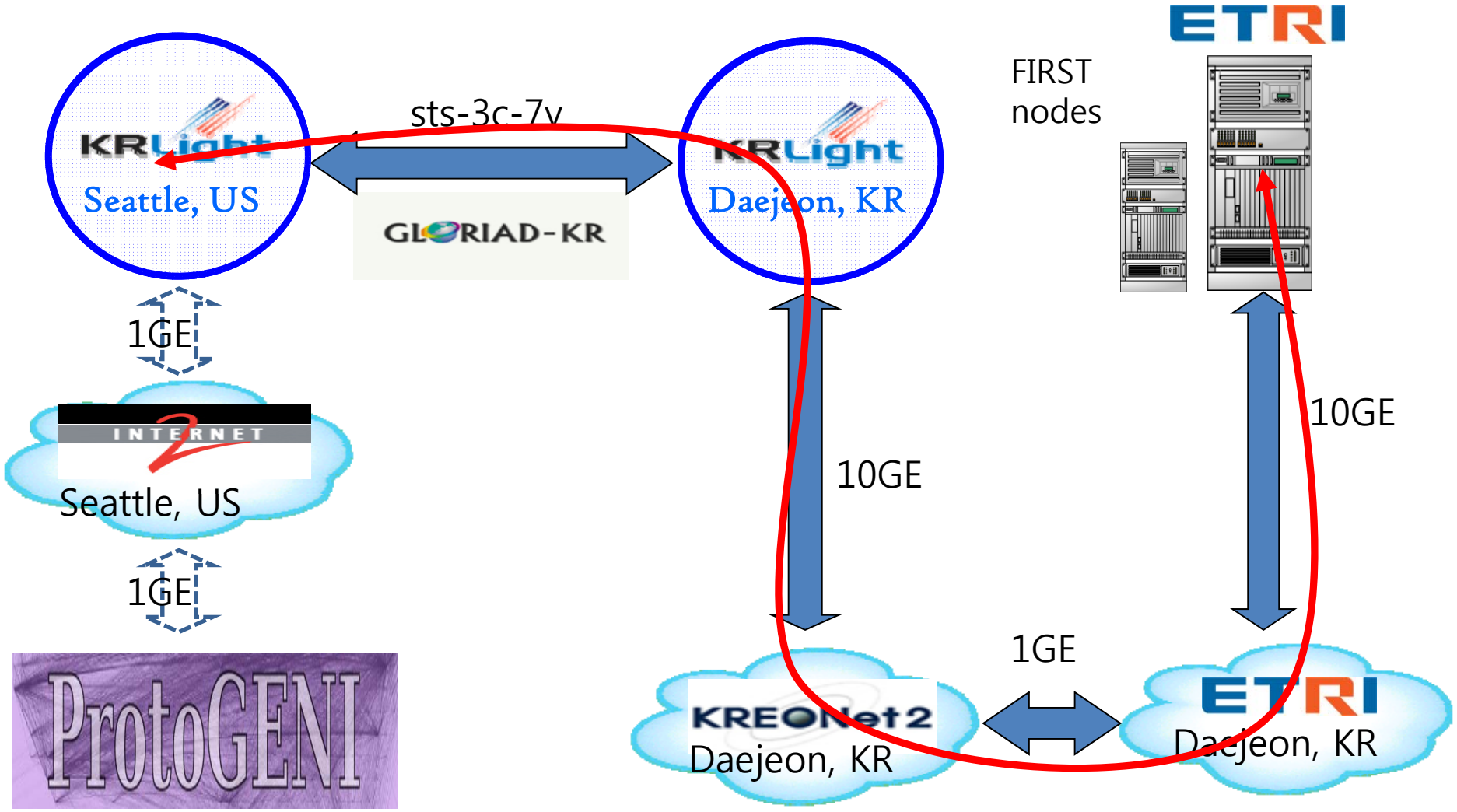


Korea-US GENI Connectivity



Contact : Dongkyun Kim@KISTI, mirr@kisti.re.kr

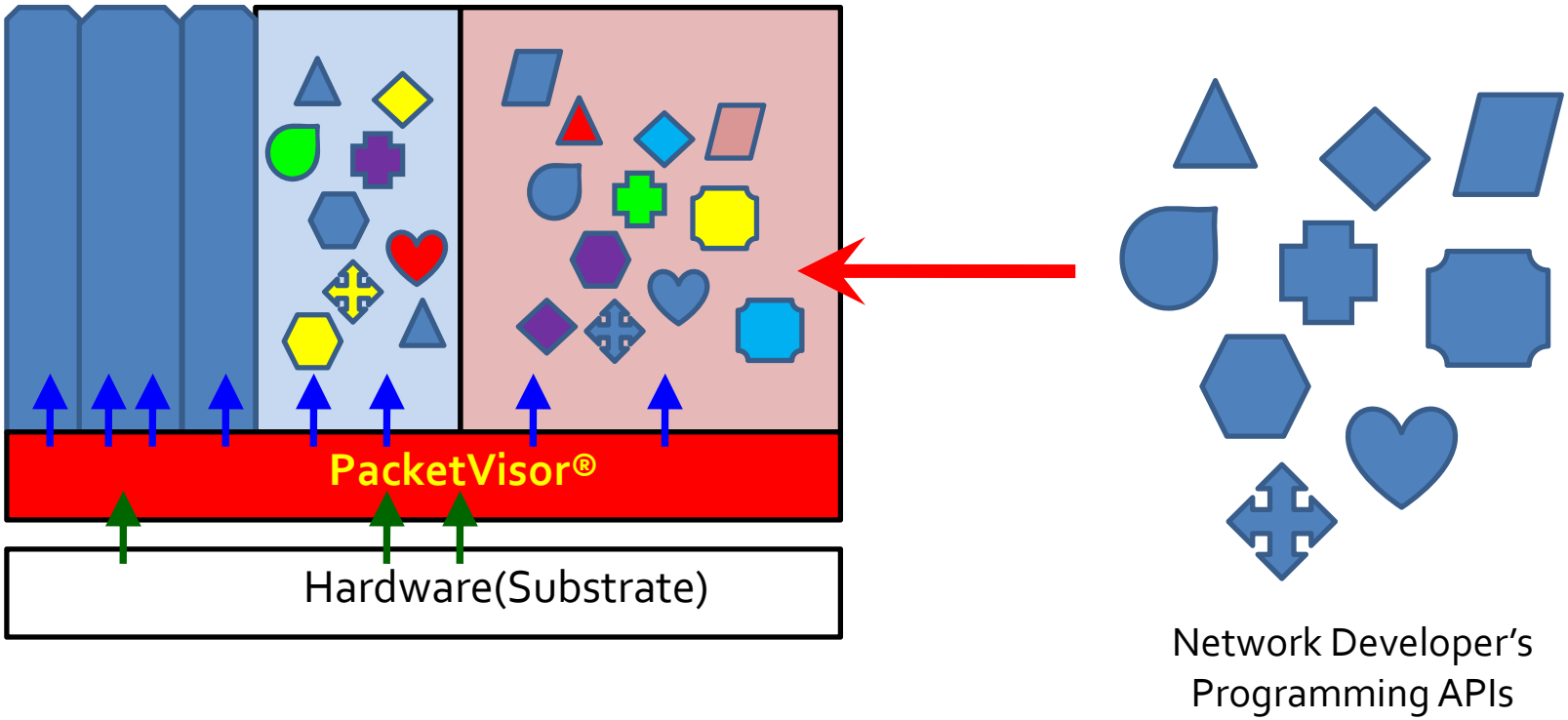
ETRI-ProtoGENI E2E Connections



Summary and Further Steps (1/3)

Network Architecture Implementation =

Concept + Combinational &
+ Sequential Logic of APIs
+ Minimum Programming of C



Summary and Further Steps (2/3)

- Go to our second goal …
  - **Quick reconfigurability**
  - **Scalability**
  - **Traffic Isolation**
  - **Performance**
  - **Security**
  - **Management**

Summary and Further Steps (3/3)

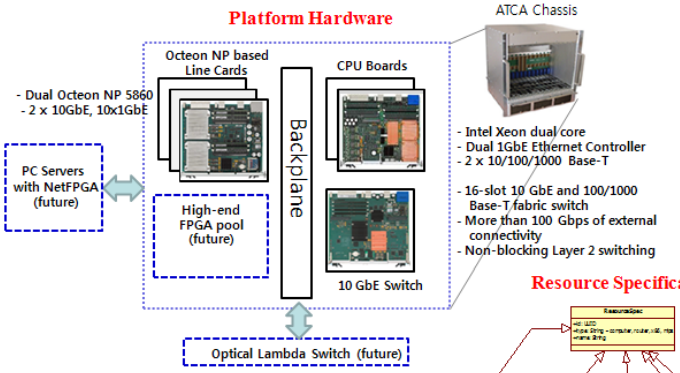
- Two or three core nodes will be installed soon in Korea FI backbone (KREONET, KOREN)
- ProtoGENI control integration
- International federation (ETRI-ProtoGENI 1Gbps End-to-End Connectivity)
  - Provisioned: ETRI (Korea) – Seattle (USA) Lightpath
  - Waiting for Cross-Connections in Seattle
    - Connectors: KRLight and Internet2 at Westin Bldg.
  - Next Steps
    - (Dedicated) Lightpath Extension to ProtoGENI in Seattle or Utah
    - VLAN Connections between ETRI and ProtoGENI

Backup slides

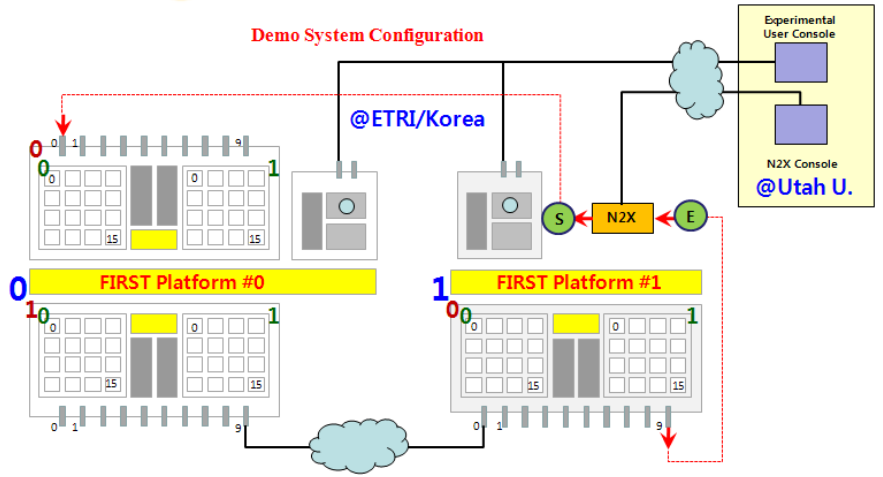
FIRST@ETRI Future Internet Virtualized Programmable Platform

ETRI is developing the high performance Future Internet Virtualized Programmable Platform based on Network Processors.
It has the ProtoGENI compatible Control Framework and the functions of dynamic resource management, dynamic topology management and program download.

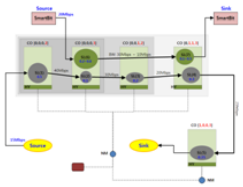
Platform Hardware



Demo System Configuration



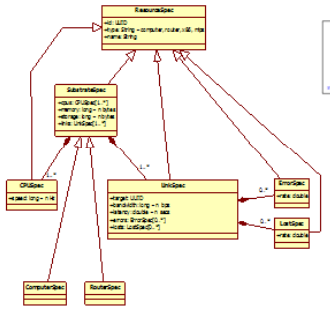
Example of Experiment



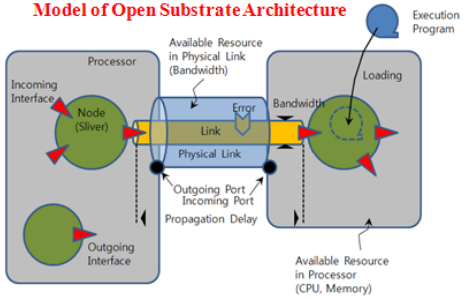
Demonstration Procedures

- Create Experiment Topology
- Allocate components to Virtualized Resources
- Allocate Rspecs to Sliver and Link
- Modify Experiment Topology during Operation
- Change Rspecs during Operation
- Create more than 2 slivers in a single CPU
- Download Network Program into Sliver
- Query CPU, Link Utilization Status

Resource Specification



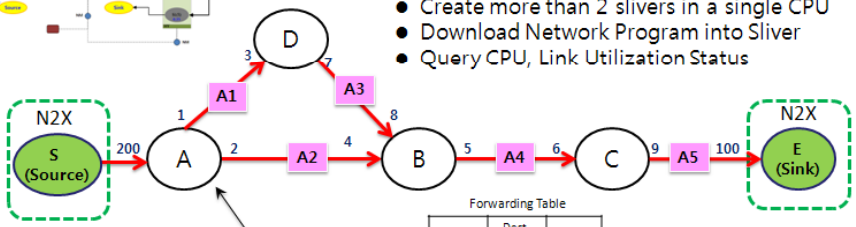
Model of Open Substrate Architecture



List of Open Substrate Interfaces

- allocateSliver/deallocateSliver
- allocatePort/deallocatePort
- allocateLink/deallocateLink
- allocateProgram/de...
- getSliverStatus
- getPortStatus
- getLinkStatus

Experiment Topology



Forwarding Table

SLV	Dest. SLV	Port
A	D	1
A	B	2
A	C	2
A	E	2
B	C	5
B	E	5
C	E	9
D	B	7
D	C	7
D	E	7

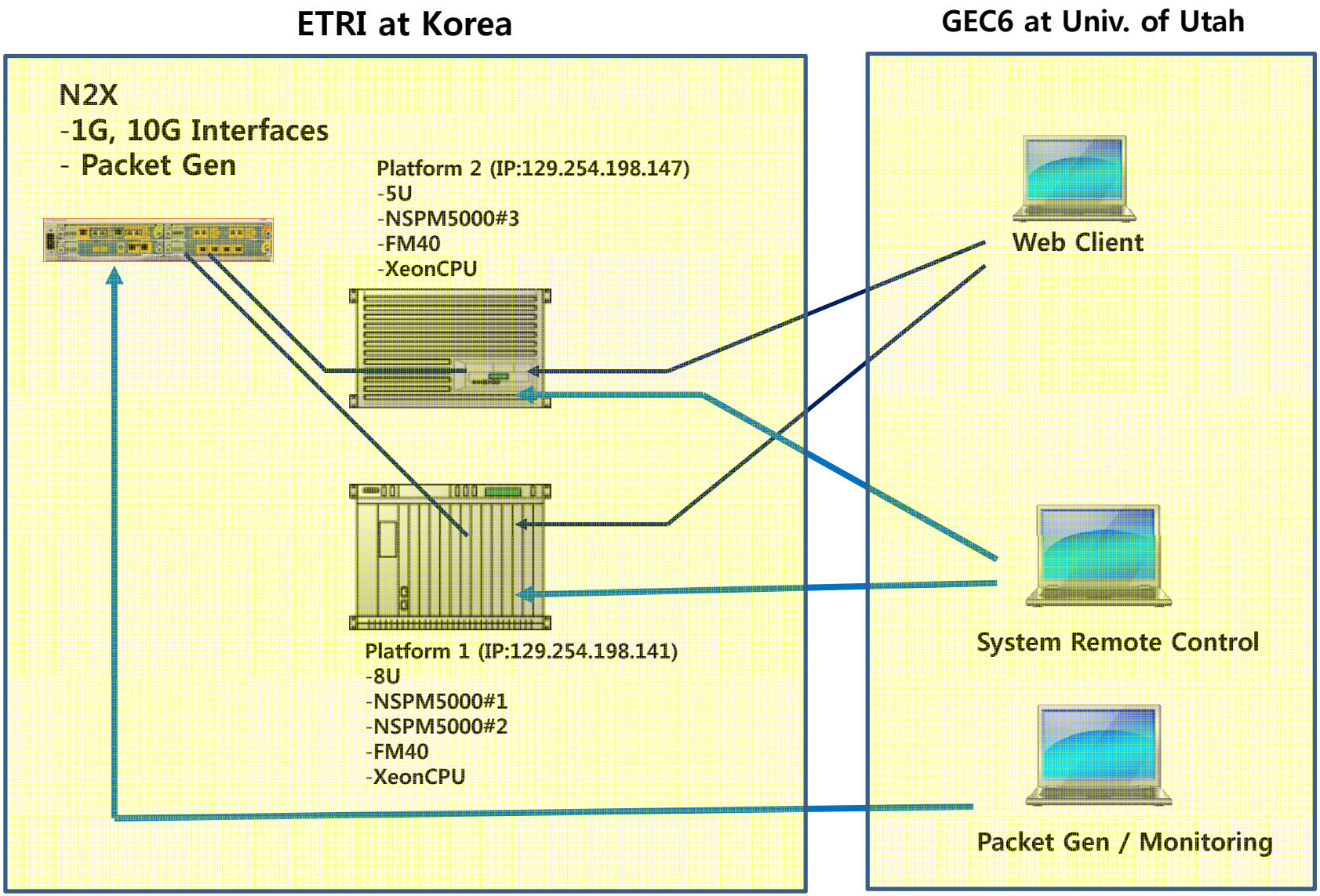
Link Table

Link	Src	Dst
A1	1	3
A2	2	4
A3	5	6
A4	7	8
A5	9	100

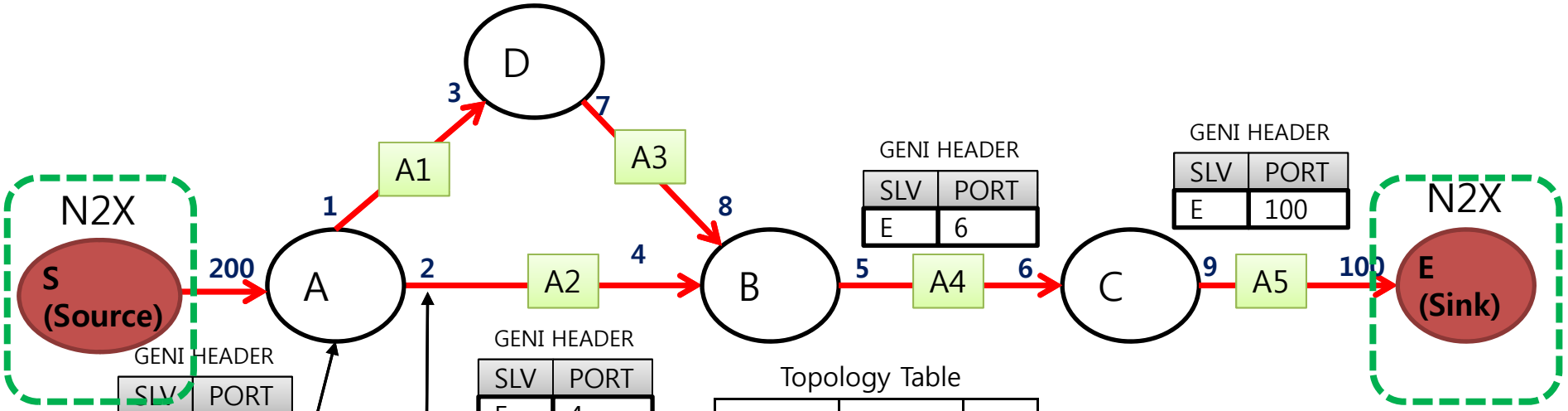
FIRST : Future Internet Research for Sustainable Testbed
ETRI : Electronics and Telecommunications Research Institute in Korea

GEC 6, Salt Lake City, Nov. 16 - 18, 2009

Demo Configuration@GEC6



Topology Example



Topology Table

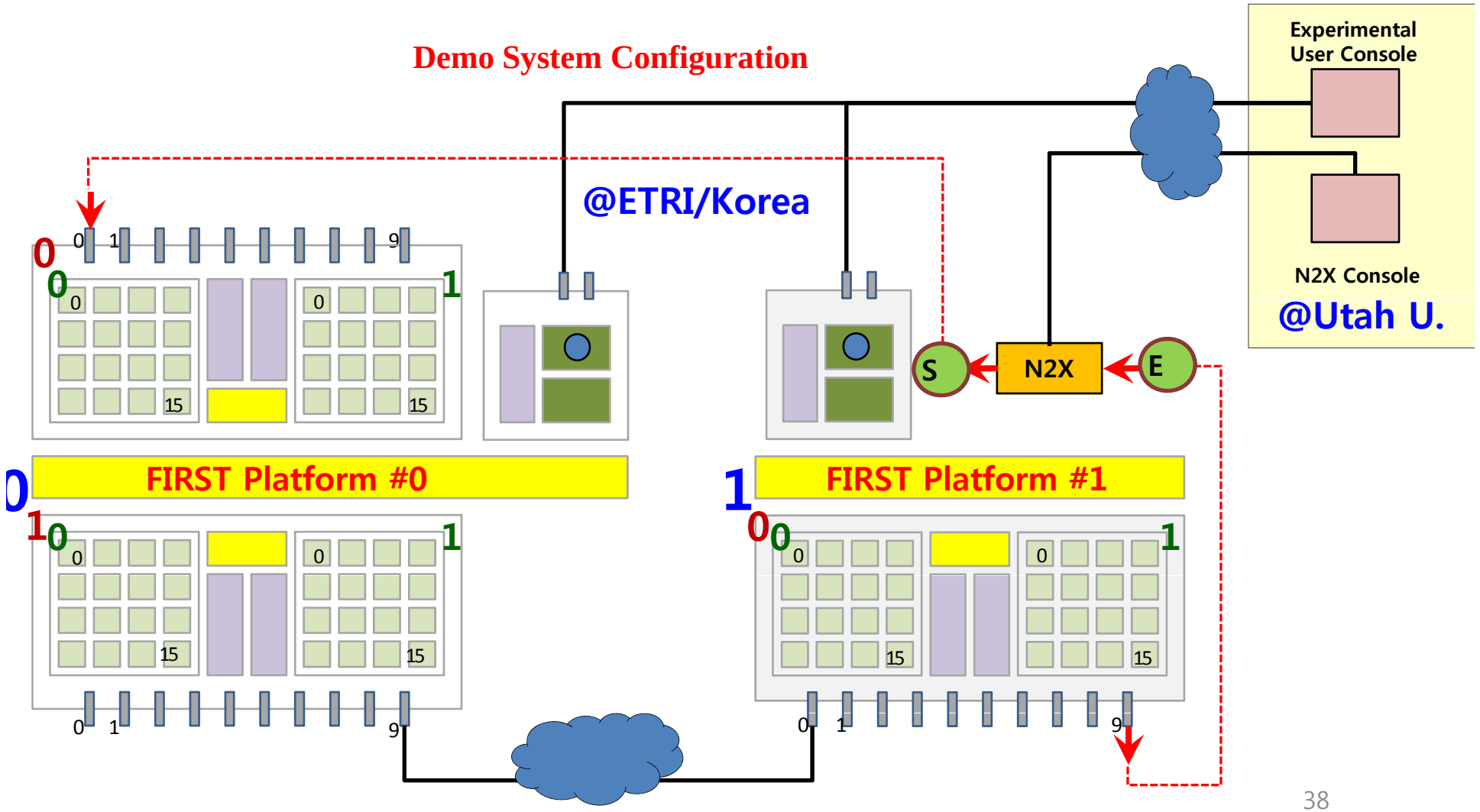
SLV Src (Hex/Dec)	SLV Dst (Hex/Dec)	Port
A(10)	D(13)	1
A(10)	B(11)	2
A(10)	C(12)	2
A(10)	E(14)	2
B(11)	C(12)	5
B(11)	E(14)	5
C(12)	E(14)	9
D(13)	B(11)	7
D(13)	C(12)	7
D(13)	E(14)	7

Every node has this table!!

Link Table

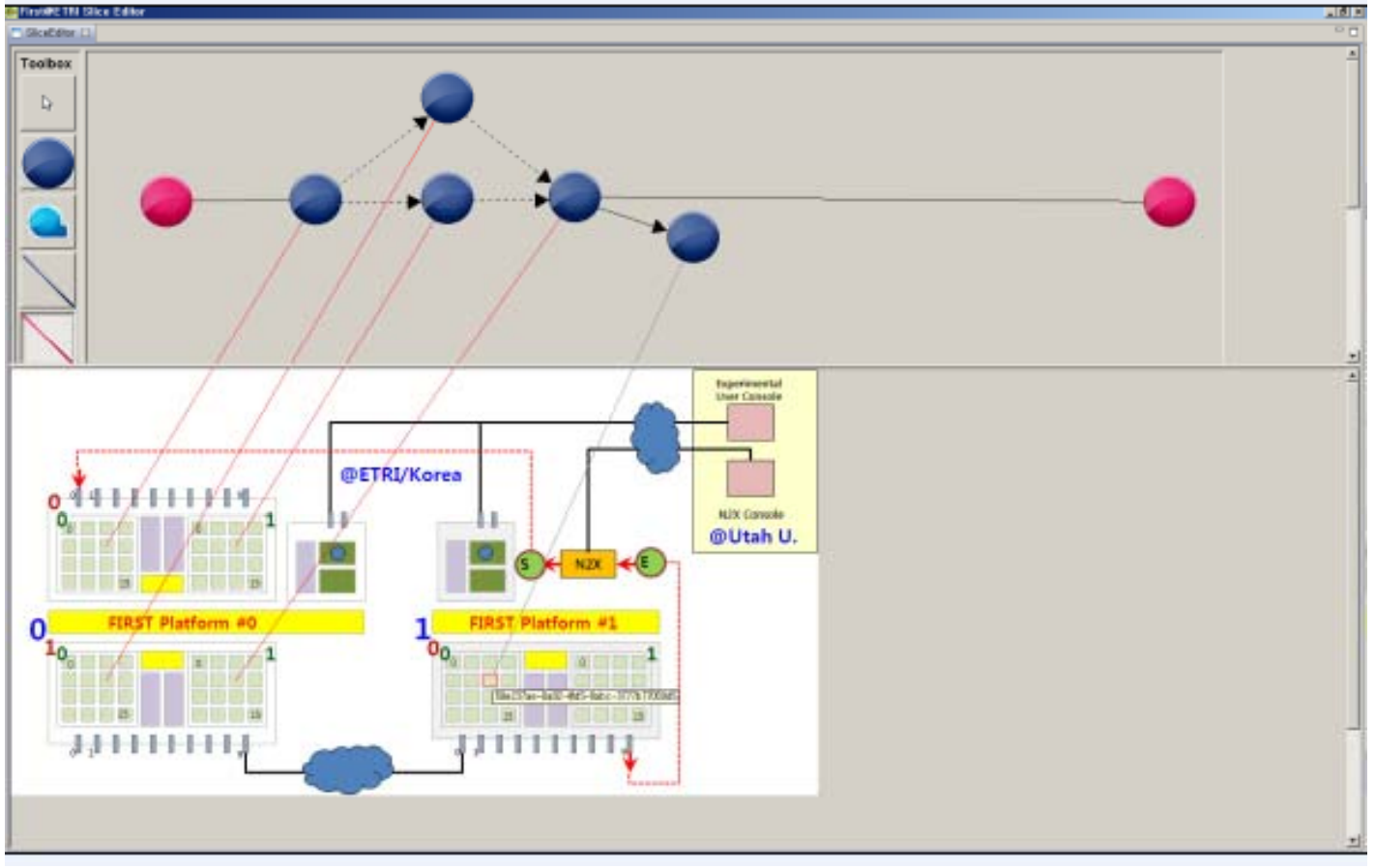
Link	Src	Dst	Src Mac	Dst Mac
A1(161)	1	3	00:??	00:??
A2(161)	2	4	00:??	00:??
A3(163)	5	6	00:??	00:??
A4(164)	7	8	00:??	00:??
A5(165)	9	100	00:??	00:??

Resource Allocation in Processors



DEMO@GEC6

6th GENI Engineering Conference, Salt Lake City



DEMO@GEC6

6th GENI Engineering Conference, Salt Lake City