Introduction to Network Virtualization Technologies in Future Internet Research

Kiyohide Nakauchi
NICT, Japan
nakauchi@nict.go.jp

Asia FI Summer School
August 26, 2010
Future Internet / New Generation Network

Network Virtualization

Social requirements for network
- High speed, Diversity
- High availability & quality
- Secure, Energy saving
- Action to unknown problem

Limitations of existing technology
- Complexity by added functions
- Limit of performance improvement

Next Generation NW (NXGN)

Collaboration of Industry, Academia and Government

International Competition and Collaboration
- US (FIND, etc.)
- EU (FP7, etc.)
- Asia

Collaboration
- NWGN Promotion Forum

JGN, JGN2, JGN2plus

Toward the New Testbed

Y2010 (Next Generation) → Y2015 (New Generation)

Current Network → Next Generation Network (NXGN) → New Generation Network (NWGN)

2010/8/26 K.Nakauchi, NICT
Outline

- Network Virtualization
- View from the Network Testbeds
  - GENI
  - FIRE
- NICT’s Activities
  - CoreLab
  - VNode
  - Wireless testbed
- Research Agenda (challenges)
What is Network Virtualization?

- **Isolate** network & computational resources through virtualization

- Hold multiple independent and programmable logical (virtual) NWs

- Multiple network architectures and services on top of isolated logical NWs
  - Meta-architecture to enable multiple architectures
  - Testbed technology to test and support multiple concurrent experiments
  - Substrate for providing differentiated and user-specific QoS and services
Definitions

- **Isolation**
  - independent and not interfered with one another

- **Elasticity**
  - Flexible re-organization

- **Programmability**
  - Any format, any protocol, any data processing

- **Abstraction**
  - Flexible description and usage for a set of components/resources

- **Aggregation**
  - Dynamic combination of local/remote components/resources

Network Virtualization ≠ Overlay, VPN, VLAN, Cloud, …
Why Network Virtualization?

- **Migration** to new network architecture
  - Enable co-existence of traditional and new network architecture
  - **Sustainable** network architecture

- **Innovation platform**
  - Users can freely embed novel functions, protocols, or services into the routers

- **Potential for new service and business model**
  - VN Operators (VNO) can provide a unique network without physical infrastructure (similar to MVNO)
Standardization

- IRTF Virtual Networks Research Group (VNRG)
  - Charter
    - Identifying architectural challenges resulting from VNs
    - Addressing network management of VNs
    - Exploring emerging technological and implementation issues
  - Co-chairs: Joe Touch (USC/ISI), Martin Stiemerling (NEC Europe)
  - http://www.irtf.org/charter?gtype=rg&group=vnrg

- ITU-T Focus Group on Future Networks (FG-FN)
  - Identify potential study areas, a timeframe, and potential impacts of future networks, including network virtualization technology.
  - Chair: Takashi Egawa (NEC)
Outline

- Network Virtualization
- View from the Network Testbeds
  - GENI
  - FIRE
- NICT’s Activities
  - CoreLab
  - VNode
  - Wireless testbed
- Research Agenda (challenges)
GENI

- The programmable network testbed for Future Internet research
  - Derived from PlanetLab and slice-based facility
  - Designing and prototyping of control framework over hetero-testbeds
  - Deploying the testbed, collaborating with Internet2 and NRL
- Funded by NSF, Led by GPO
  - $12M for 29 projects (2008), $11.5M for 33 projects (2009), planned (2010)
  - $10.5M for 18 universities for campus deployment (2009)
- 60 PJs, grouped into 4 Clusters

Programmable & federated, with end-to-end virtualized "slices"

Common control framework across heterogeneous testbeds

src: http://www.geni.net/
GENI Prototype Examples

DRAGON core nodes
Mid-Atlantic Crossroads  WAIL, U. Wisconsin-Madison

ViSE, U. Mass Amherst  SPPs, Wash U.

DieselNet, U. Mass Amherst

ORBIT, Rutgers WINLAB

src: http://www.geni.net/
Building GENI Meso-scale Prototype

OpenFlow
- Stanford
- U Washington
- Wisconsin
- Indiana
- Rutgers
- Princeton
- Clemson
- Georgia Tech

WiMAX
- Stanford
- UCLA
- UC Boulder
- Wisconsin
- Rutgers
- Polytech
- UMass
- Columbia

ShadowNet
- Salt Lake City
- Kansas City
- DC
- Atlanta

OpenFlow Backbones
- Seattle
- Salt Lake City
- Sunnyvale
- Denver
- Kansas City
- Houston
- Chicago
- DC
- Atlanta

Juniper MX240 Ethernet Services Router
- NEC WiMAX Base Station
- Cisco 6509 Switch
- NEC IP8800 Ethernet Switch

HP ProCurve 5400 Switch
- Arista 7124S Switch

2010/8/26 K.Nakauchi, NICT

src: http://www.geni.net/
Cluster B: PlanetLab Control Framework

- **PlanetLab**
  - Clearinghouse, CM
  - 800+ nodes
  - VINI (virtual topologies)
- **Enterprise GENI (OpenFlow)**
  - GENI VLANs on enterprise nets
- **SPP Overlay Nodes**
  - PL-based Programmable router
- **GUSH Tools**
  - Experiment design tools
- **Provisioning Service**
  - Slice & experiment management tools
- **Mid-Atlantic Crossroads**
  - Regional network with VLAN control plane
- **GpENI**
  - Regional network with sliceable optics & routers

src: http://www.geni.net/
PlanetLab and OpenFlow

- Virtualization using Linux Vserver
  - Auditing service
  - Monitoring services
  - Brokerage services
  - Provisioning services

- Virtualization using FlowVisor, a network virtualization layer

Virtual Machine Monitor (VMM)

- Linux kernel (Fedora Core)
  - Vservers (namespace isolation)
  - Schedulers (performance isolation)
  - VNET (network virtualization)
Cluster E: ORBIT Control Framework

- **ORBIT**
  - WINLAB wireless testbed resources (400+ AP nodes)
  - NICTA (Australia) wireless outdoor traffic testbed

- **WiMAX**
  - Open, programmable WiMAX base station

src: http://www.geni.net/
Virtualization of WiMAX BS

- Isolation by exploiting IEEE802.16e common packet scheduler
- Mapping IEEE802.16e service classes with GENI slices

**GENI Open API**
- Maximum Bit-Rate, Burst-Rate
- Minimum Tolerable Sending Bit-Rate
- Maximum Tolerable jitter
- Minimum Delay
- Scheduling Type (e.g., UGS, rtPS, nrtPS)
- Frequency of Operation
- Rate / Power requirements

**Base Station State**
- Radio resources (UL and DL)
- Time slots (UL and DL)
- Downlink burst profile
- Uplink burst profile
- Frequency
- Power
- Rate

src: http://www.geni.net/
Outline

- Network Virtualization
- View from the Network Testbeds
  - GENI
  - FIRE
- NICT’s Activities
  - CoreLab
  - VNode
  - Wireless testbed
- Research Agenda (challenges)
FIRE

- **Pan-Europe federated testbed facility** for Future Internet research
  - Network virtualization is one of the key features
  - Deployed on GEANT2 and NREN
  - Service testbeds on network testbeds
- Funded by EC in FP7
  - 40M € for 14PJs (Call 2, 2008), 50M € for 17PJs (Call 5, 2010)

http://cordis.europa.eu/fp7/ict/fire/
OneLab2

- Operate PlanetLab Europe
  - PlanetLab extension (measurement, wireless, …)
- http://www.onelab.eu/

- Develop a SFA-based federation mechanism
  - Federate w/ PlanetLab, PLJ, ORBIT, G-Lab, etc

Over 140 nodes at over 70 sites in Europe
A site contributes two nodes in exchange for ten slices
Each node provides Linux VServer virtual machines
Deploy Juniper’s virtualization-capable routers

http://www.fp7-federica.eu/
Panlab II

- Pan-Europe laboratory
  - [http://www.panlab.net/](http://www.panlab.net/)

- Teagle: a federation framework
  - non-slice-based testbeds and non-VM-based slices are also handled
  - Service level description
  - [http://www.fire-teagle.org/](http://www.fire-teagle.org/)

Figure 9. VCT tool view.

Figure 10. Configuration page for a PLC node in Teagle.
Why Federation?

- **Federation** is necessary when two or more organizations allocate some of their resources to implement a common service

  Many benefits
  - Diversity, realism (geography, technology)
  - Scalability (number of nodes, resources)
  - Multiplexing/aggregation (more efficient resource usage)
  - Diversity growth is limited (reduce complexity)
  - Creation of a global research community

  But also challenges
  - Manage complexity
  - Different communities
  - Legal and trust issues, policies & governance

FIRE Second Wave (Call5, 2010)

Coordination and Support Actions
- FIREBALL
- MyFire
- PARADISO 2
- FIREstation

Building the Experimental Facility and stimulating its use
- BonFIRE
- OFELIA
- Smart Santander
- Tefis
- CREW

Experimentally-driven Research
- CONECT
- Euler
- BONFIRE
- SPITFIRE
- SCAMPI
- CONVERGENCE
- NOVI
- HOBNET

- Integrated project
- Focused project
- Coordination & support actions
Outline

- Network Virtualization
- View from the Network Testbeds
  - GENI
  - FIRE
- NICT’s Activities
  - CoreLab
  - VNode
  - Wireless testbed
- Research Agenda (challenges)
Enabling net-virt via only S/W on COTS x86 machines

Motivation

- Performance
  - Network I/O
- Scalability
  - The number of slices
- Isolation
  - Resources must be isolated per slice
- Flexibility
  - Arbitrary kernel/network stacks/resources
- Code-Reusability
  - Catch up with the latest technology
  - Minimize software engineering effort

Largely Disregarded

http://www.corelab.jp/
CoreLab Design

- **Flexible virtualization**
  - Hosted Virtual Machine Monitor (KVM + QEMU)
  - Resource Container (OpenVZ, LXC)
- **VM Disk Image Transfer**: Multipath & Caching
- **Network Configuration**: Multihoming
- **Management Tools**: Reuse PlanetLab’s design (MyPLC)
  - Compatibility with PlanetLab (PLCAPI + CoreLab APIs)
- **OpenFlow** capability

<table>
<thead>
<tr>
<th></th>
<th>PlanetLab</th>
<th>CoreLab (KVM type)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sliver</td>
<td>Resource Container</td>
<td>Hosted VMM</td>
</tr>
<tr>
<td>Kernel</td>
<td>Shared</td>
<td>Separated</td>
</tr>
<tr>
<td>Guest OS</td>
<td>Various Linux Distros</td>
<td>Any OSes</td>
</tr>
<tr>
<td>IP &amp; Port</td>
<td>Shared</td>
<td>Shared (managed)</td>
</tr>
<tr>
<td>Privileged System Calls</td>
<td>Shared</td>
<td>Isolated</td>
</tr>
</tbody>
</table>

2010/8/26 K.Nakauchi, NICT
# Sliver Implementation

## Design Principles

<table>
<thead>
<tr>
<th>Resource Container</th>
<th>Hypervisor-Based</th>
<th>Host-Based</th>
</tr>
</thead>
<tbody>
<tr>
<td>Performance</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scalability</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Security Isolation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Isolation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flexibility</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Code Reusability</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Resource Container
- **Performance**: +
- **Scalability**: +
- **Security Isolation**: +
- **Isolation**: +
- **Flexibility**: -
- **Code Reusability**: disadvantageous

### Hypervisor
- **Performance**: +
- **Scalability**: +
- **Security Isolation**: +
- **Isolation**: +
- **Flexibility**: -
- **Code Reusability**: disadvantageous

### Hosted-VMM
- **Performance**: ?
- **Scalability**: ?
- **Security Isolation**: +
- **Isolation**: ?
- **Flexibility**: +
- **Code Reusability**: +

---

**CoreLab**

---

**NiCT**

---

2010/8/26

K.Nakauchi, NICT
CoreLab Deployment

- 12 sites / 24 nodes over JGN2plus + SINET + others
- Extending to Asia

Specifications:
- HP DL 580 (4U)
- Intel Xeon 16-core (quad 4-core)
- 128GB (Max 256GB) memory
- 300GB RAID5 (accommodate up to 254 slices)
Outline

- Network Virtualization
- View from the Network Testbeds
  - GENI
  - FIRE
- NICT’s Activities
  - CoreLab
  - VNode
  - Wireless testbed
- Research Agenda (challenges)
Motivations

Enabling net-virt via designing H/W based on production routers

- Learned lessons from PlanetLab and CoreLab
- Starting over and designing a new “slice” mechanism

- Our goals (benefits)
  (A) Accommodate multiple independent NWs (Meta architecture)
  (B) Enable in-network processing
  (C) Create app-specific NWs
  (D) Adaptive computational resources
  (E) Enable non-IP E2E communication
VNode Prototype

Programmer Part
(IA Serverx4 +ATCAx2+OpenFlow SWx2)

Fast-Path Network Processor Card

OpenFlow Switch

Redirector Part (AX6700+SMCx2)

AX6708S

Service Module Card

VNodeManager
Outline

- Network Virtualization
- View from the Network Testbeds
  - GENI
  - FIRE
- NICT’s Activities
  - CoreLab
  - VNode
  - Wireless testbed
- Research Agenda (challenges)
NICT Wireless Testbed

- Make effective use and integration of recently developed prototypes and software tools
  - Network virtualization, Cognitive radio, Sensor/mesh networks, Cloud, Network operation,…
  - Joint project of research groups in NICT

Programmable Wireless BS (X86/Linux)

2010/8/26  K.Nakauchi, NICT
Outline

- Network Virtualization
- View from the Network Testbeds
  - GENI
  - FIRE
- NICT’s Activities
  - CoreLab
  - vNode
  - Wireless testbed
- Research Agenda (challenges)
Research Agenda (Challenges)

- **Virtual Node Architecture**
  - System design of highly flexible, isolated, and high performance virtual router
  - Resource management

- **Federation**
  - Design of resource description and API, policy management

- **Photonic**
  - Integration with lambda control (slice-lambda mapping)

- **Wireless**
  - Wireless link virtualization
  - Integration with software radio/cognitive radio technology

- **Green**
  - Design of green routing or transport mechanism
References

- GENI: http://www.geni.net/
- PlanetLab: http://www.planet-lab.org/
- OpenFlow: http://www.openflowswitch.org/
- EmuLab: http://www.emulab.net/
- Kansei sensor network: http://cast.cse.ohio-state.edu/kansei/
- ORBIT: http://www.orbit-lab.org/
- OMF: http://omf.mytestbed.net/
- 4WARD: http://www.4ward-project.eu/
- CoreLab: http://www.corelab.jp/
- Akihiro Nakao, “CoreLab and VNode” (Invited talk), GEC8, July 2010.
End. Thank you!

Introduction of NICT

- National public research institute on information and communications technology
  - Future network and testbed
  - Photonic
  - Wireless and mobile
  - Service and applications, etc
- Many foreign researchers
  - From the world

Question to nakauchi@nict.go.jp

http://www.nict.go.jp/
Backup
Cluster C: ProtoGENI Control Framework

- ProtoGENI
  - Clearinghouse, CM
  - Emulab resources
  - (370+ nodes)

- CMULab
  - Home Wireless APs
  - Emulab cluster
  - Wireless emulation testbed

- Instrumentation Tools
  - UK Edulab (compute/store)

- Measurement System
  - GIMS prototype

- Virtual Tunnels
  - Dynamic tunnel tools
  - BGP distribution tools

src: http://www.geni.net/
Cluster D: ORCA Control Framework

- **ORCA/BEN**
  - ORCA resource leasing software (SHARP)
  - Metro-Scale Optical Testbed (BEN)
- **Sensor/Actuator Network**
  - CASA (radar, video, weather sensors)
- **Kansei Sensor Network**
  - Wireless sensor network arrays
  - 3 federated sites each with ~100 sensor nodes
- **Vehicular Mobile Network**
  - Programmable nodes with radios on city busses

Source: [http://www.geni.net/](http://www.geni.net/)
Academic-Industrial Teams (Spiral 2)

Project Name
1. CMUlab, LEARN
2. D Meas, LEARN
3. Digital Object Registry
4. CLOUD-CTL, DOME, VISE
5. DTunnels
6. EnterpriseGENI, OpenFlow

Project Lead
- Carnegie Mellon University
- University of Houston
- Corporation for National Research Initiatives (CNRI)
- University of Massachusetts Amherst
- The Georgia Institute of Technology
- Stanford University

Project Participants
- Clemson University
- Georgia Institute of Technology
- Indiana University
- Nicira Networks
- Princeton University
- Rutgers University
- University of California, Berkeley
- University of Washington
- University of Wisconsin
- University of Nebraska-Lincoln
- UC San Diego
- University of Colorado, Boulder, CO
- University of Wisconsin-Madison
- Colgate University
- University of Wisconsin, Madison, WI
- Duke University
- Universite Pierre et Marie Curie (UPMC)
- University of California, Berkeley
- Radio Technology Systems LLC
- University of Alabama
- Clemson University
- North Carolina State University
- University of Delaware
- Internet2
- Cypress, CA
- HP Labs, Palo Alto
- Oklahoma State University
- Howard University
- Florida International University
- The Quilt
- Purdue University
- University of Illinois (NCSA)
- University of Alaska Fairbanks