

# Big Data and SDN

Shinji Shimojo  
NICT/Osaka U  
2012.12



SDN

an application of SDN is overlay

PIAX overlay network platform

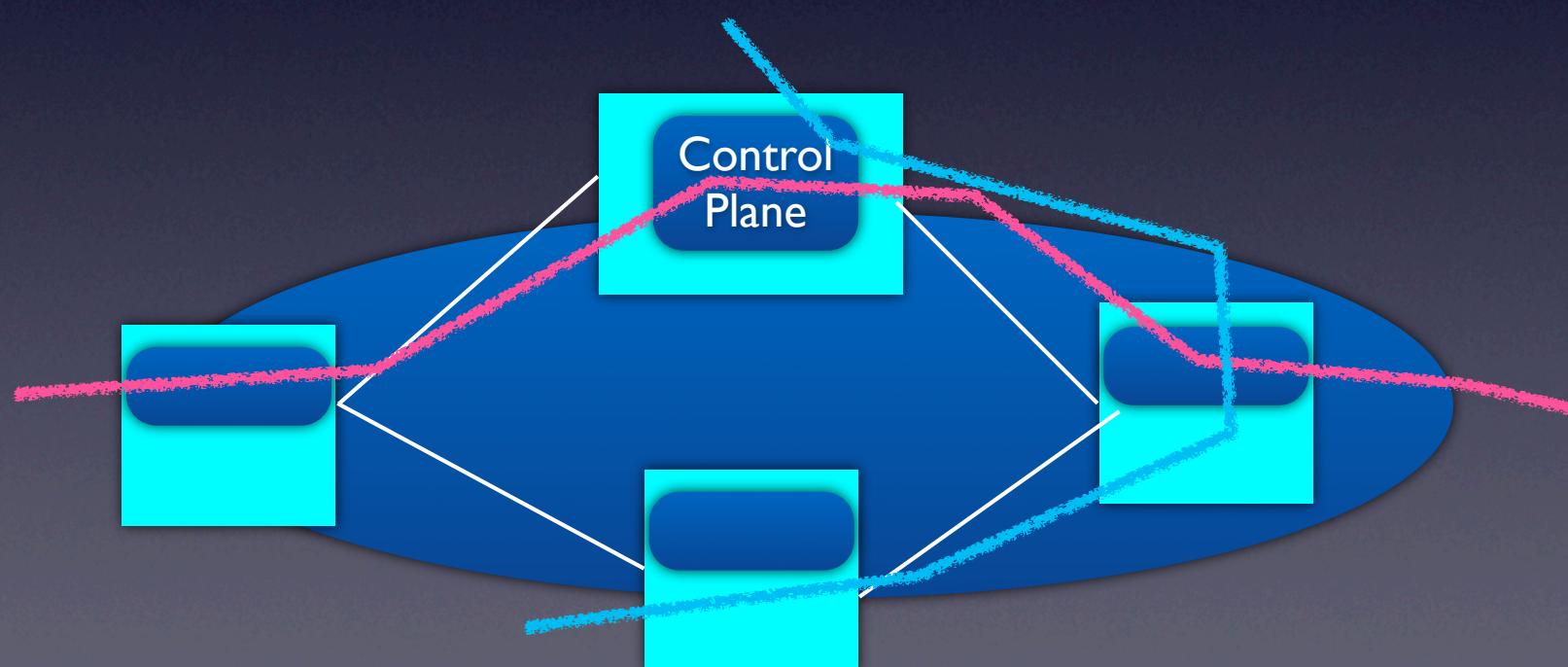
How does it help Big Data?

# New Generation Network Technology

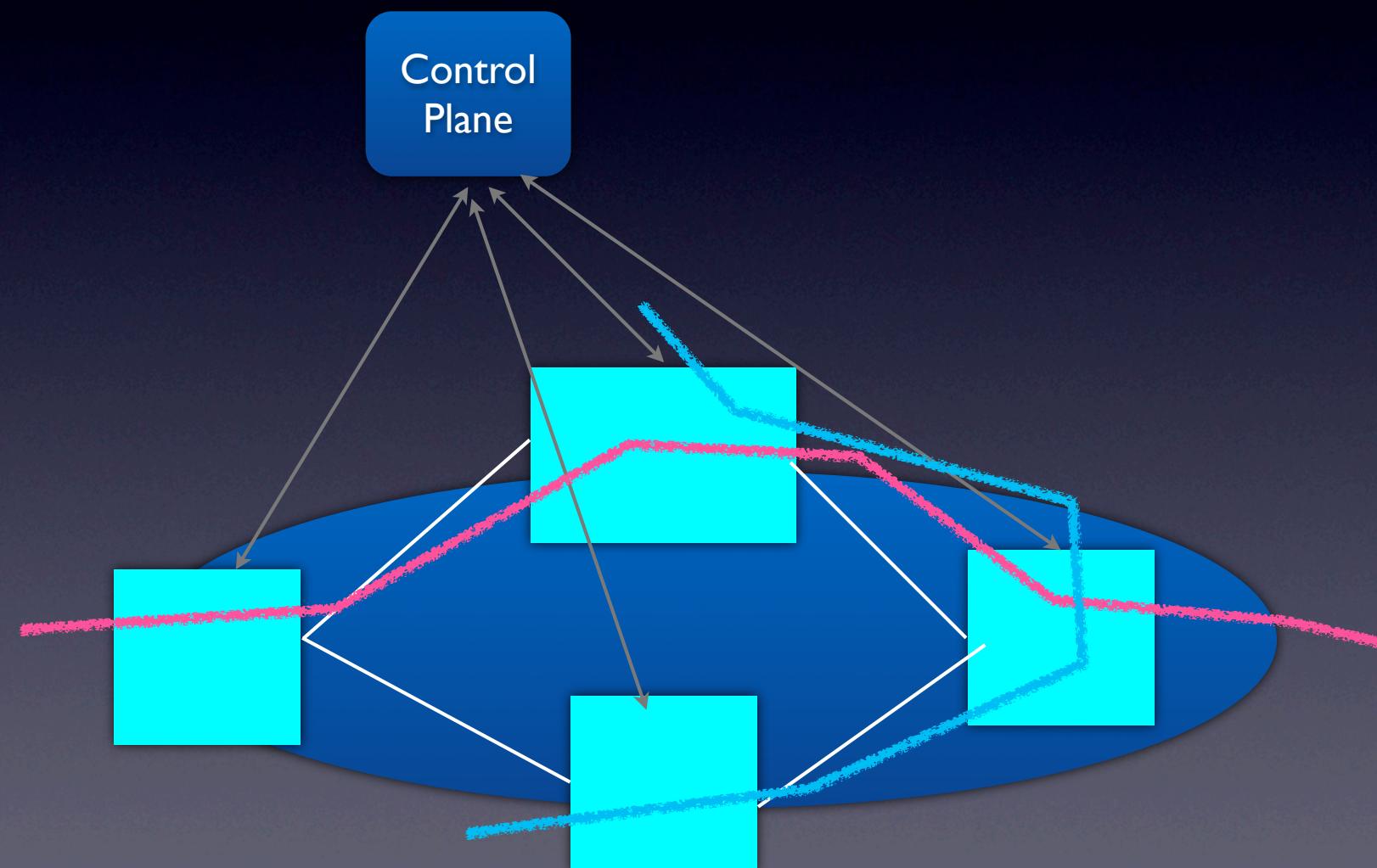
- Network Virtualization
- Programability on Network
- ID/Locator Separation
- Contents Oriented Network
- In Network Processing
- Optical Networking
- Wireless Networking

SDN=  
Software Defined Network

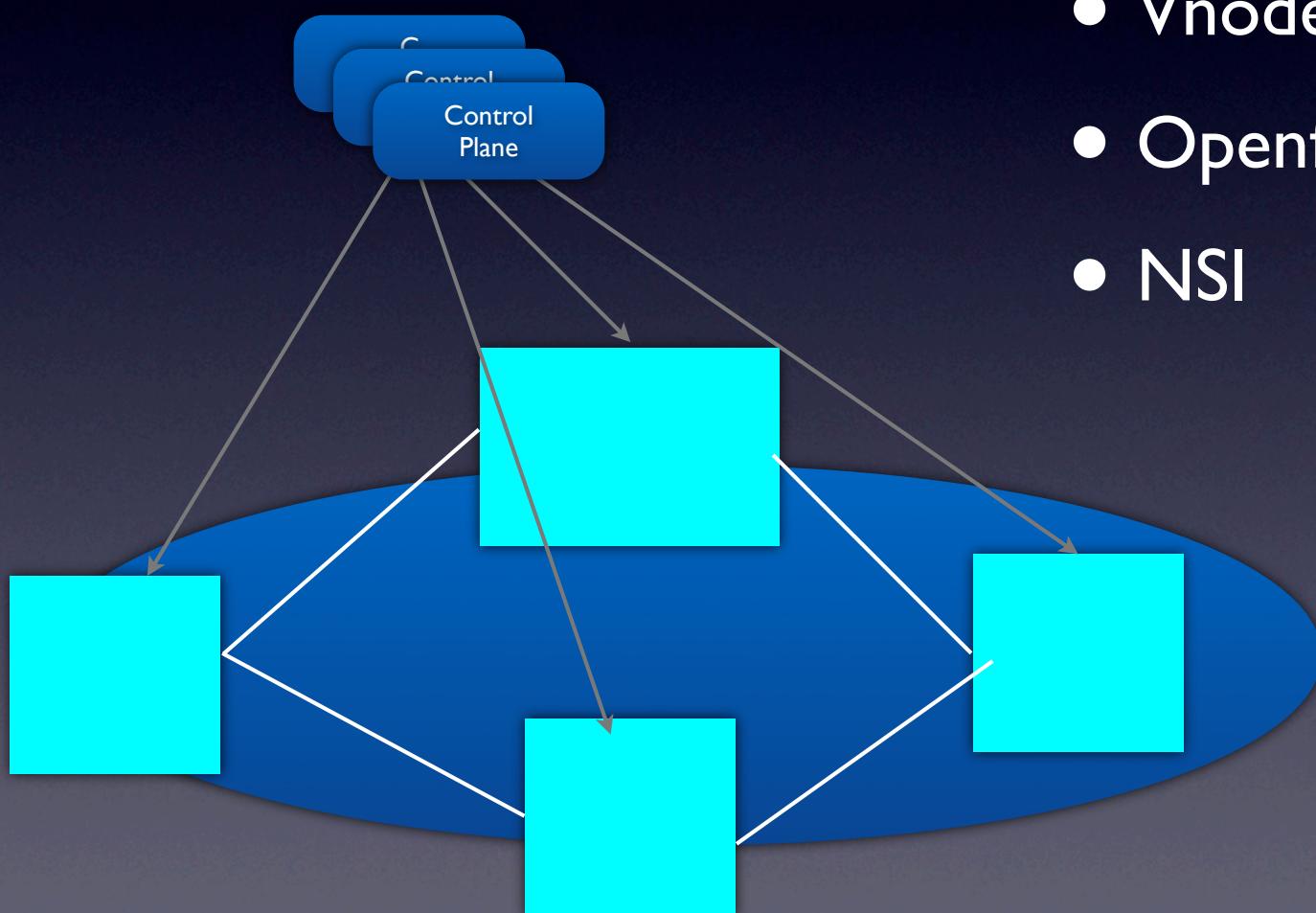
# Software Defined Network



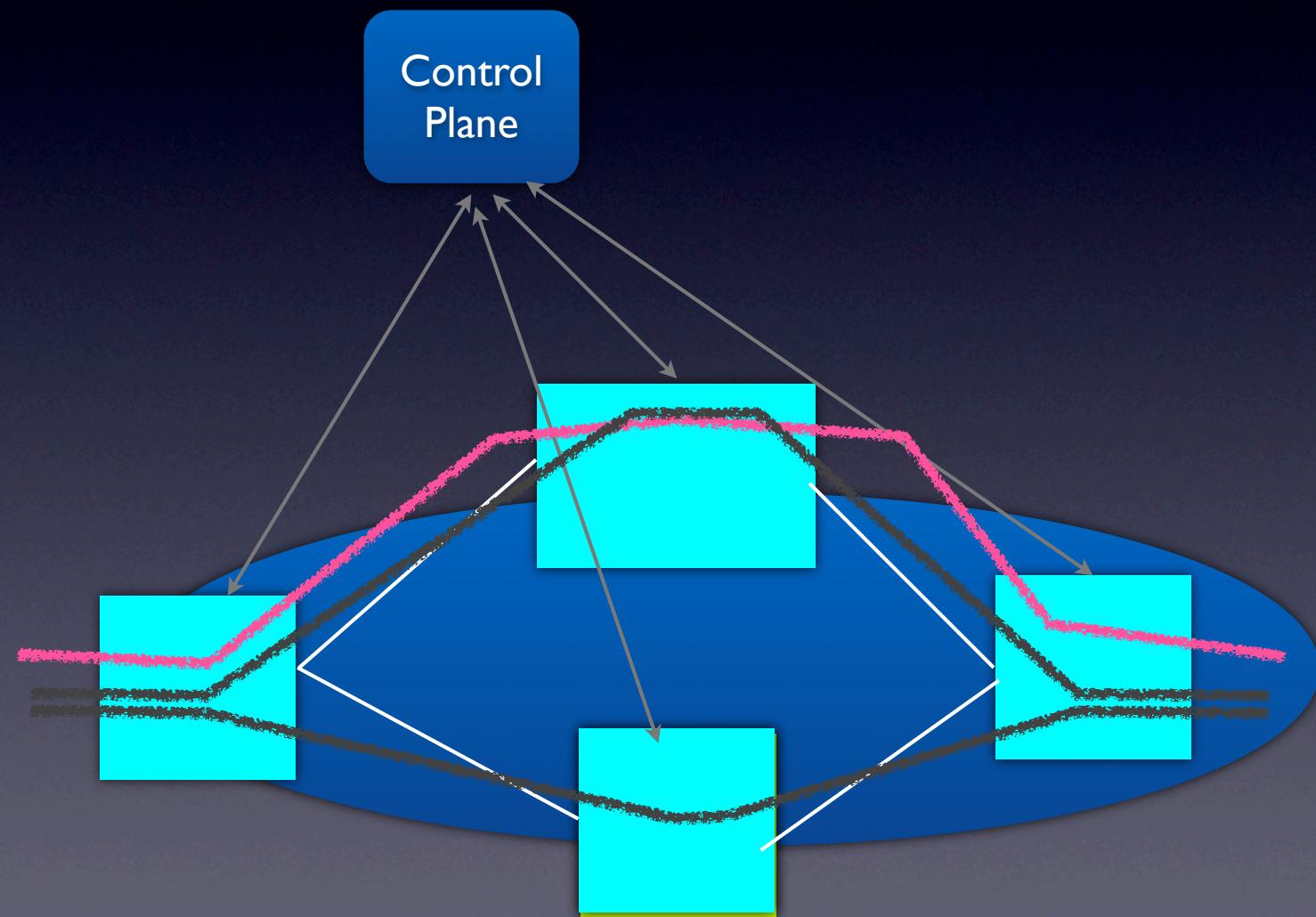
# Software Defined Network



# Software Defined Network



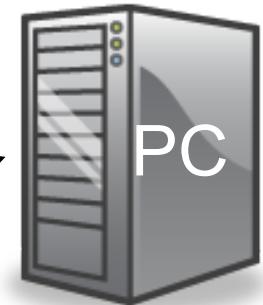
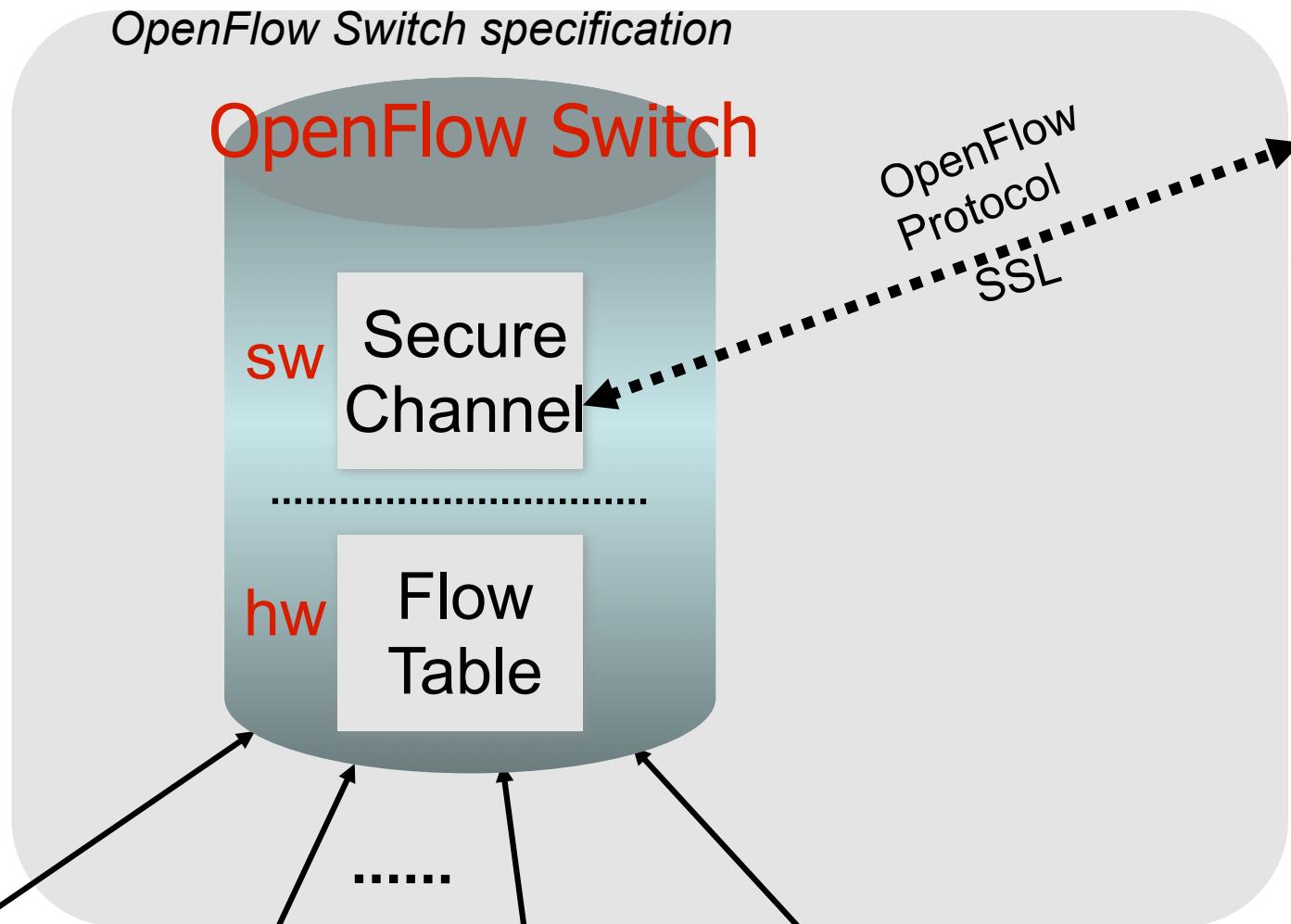
# path aggregation for maintenance



# OpenFlow Switching

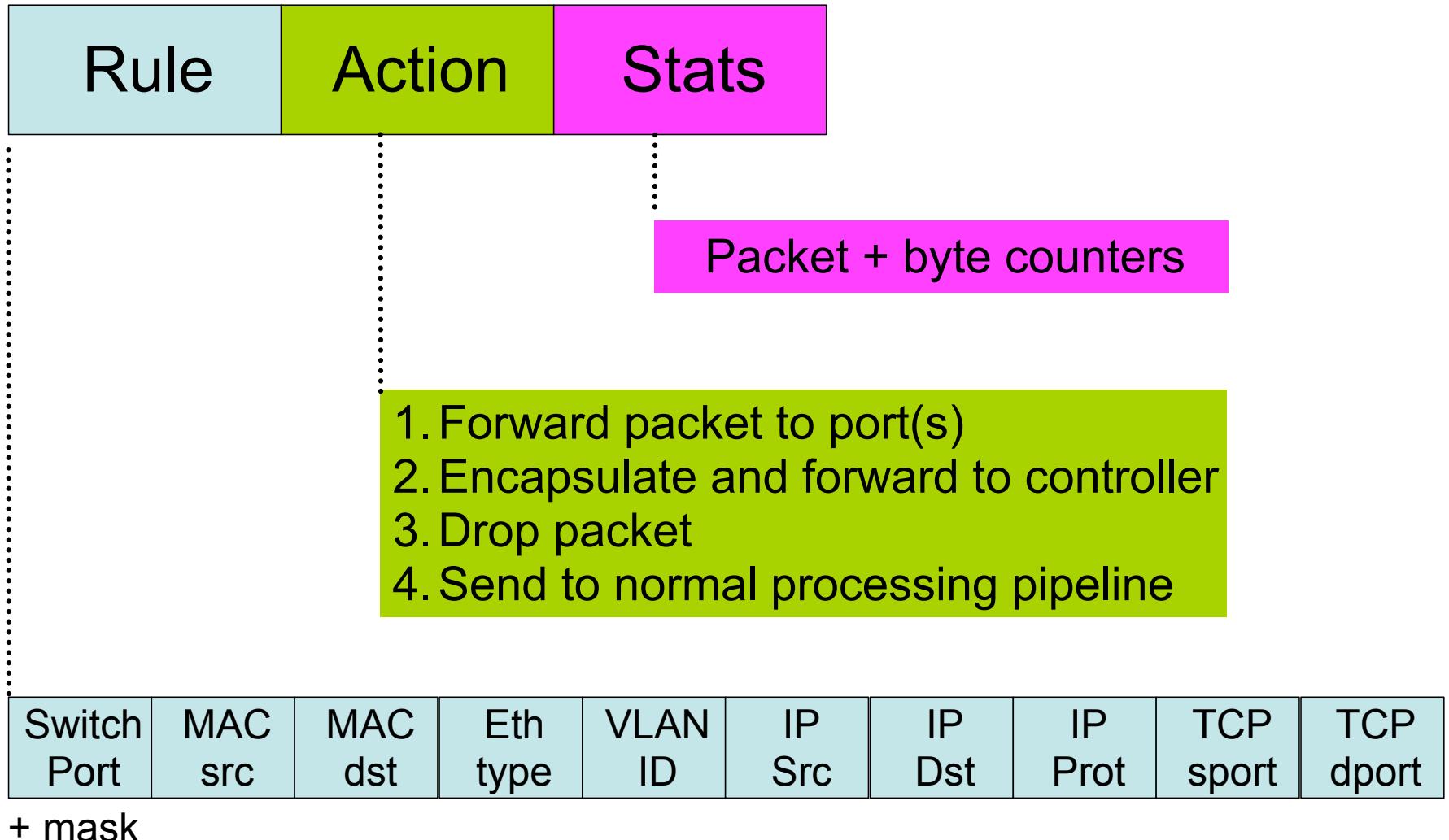
Controller

*OpenFlow Switch specification*



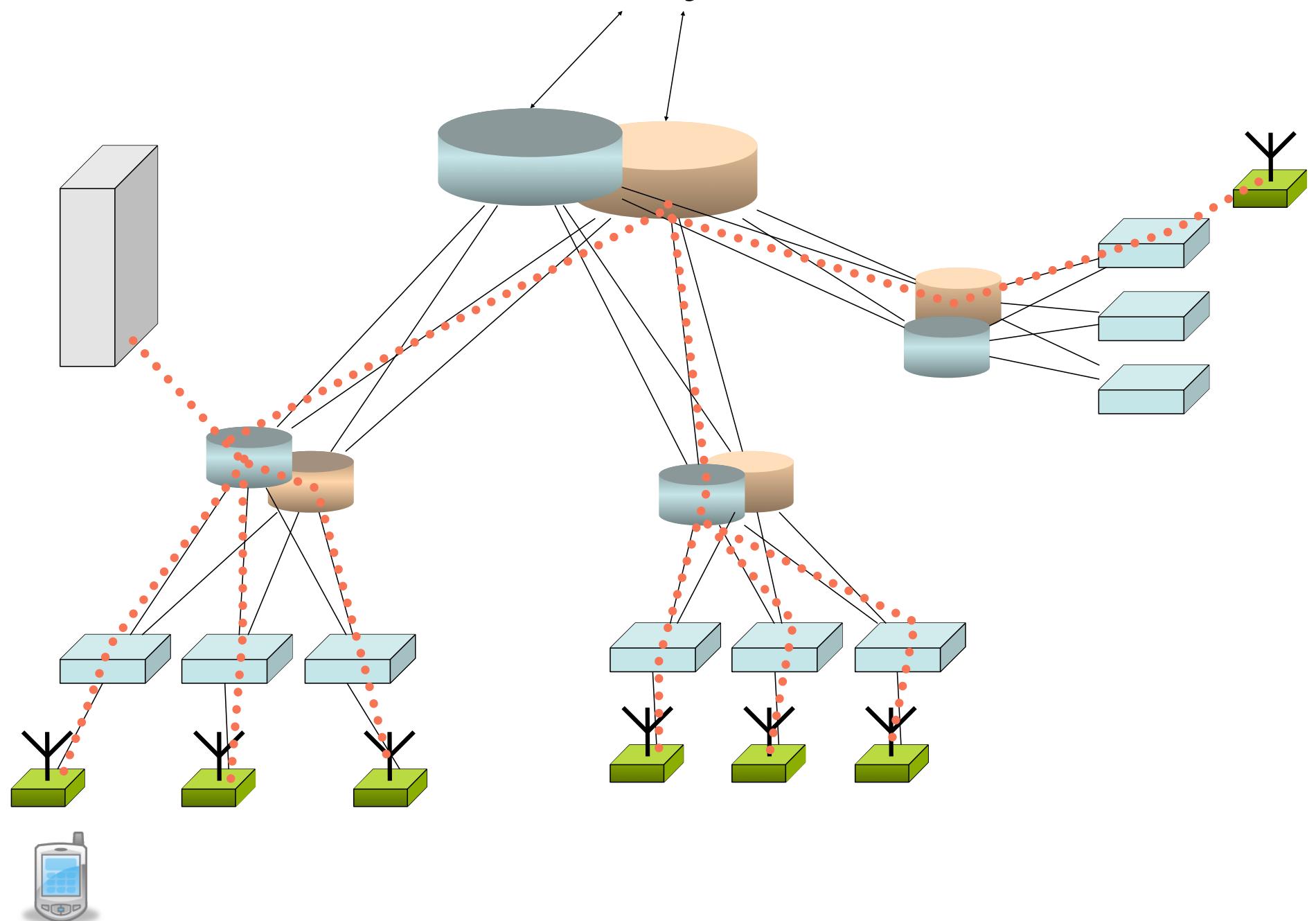
# Flow Table Entry

## “Type 0” OpenFlow Switch

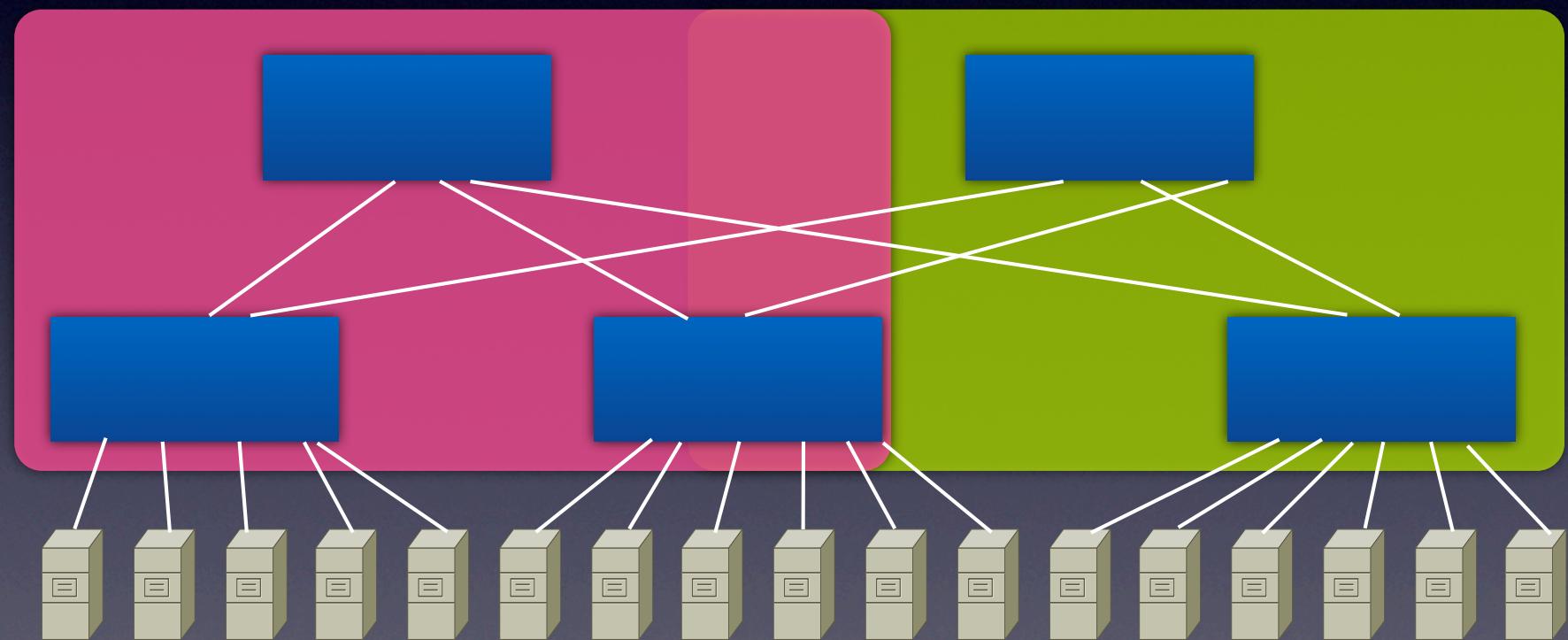


# Example Experiment at the flow level

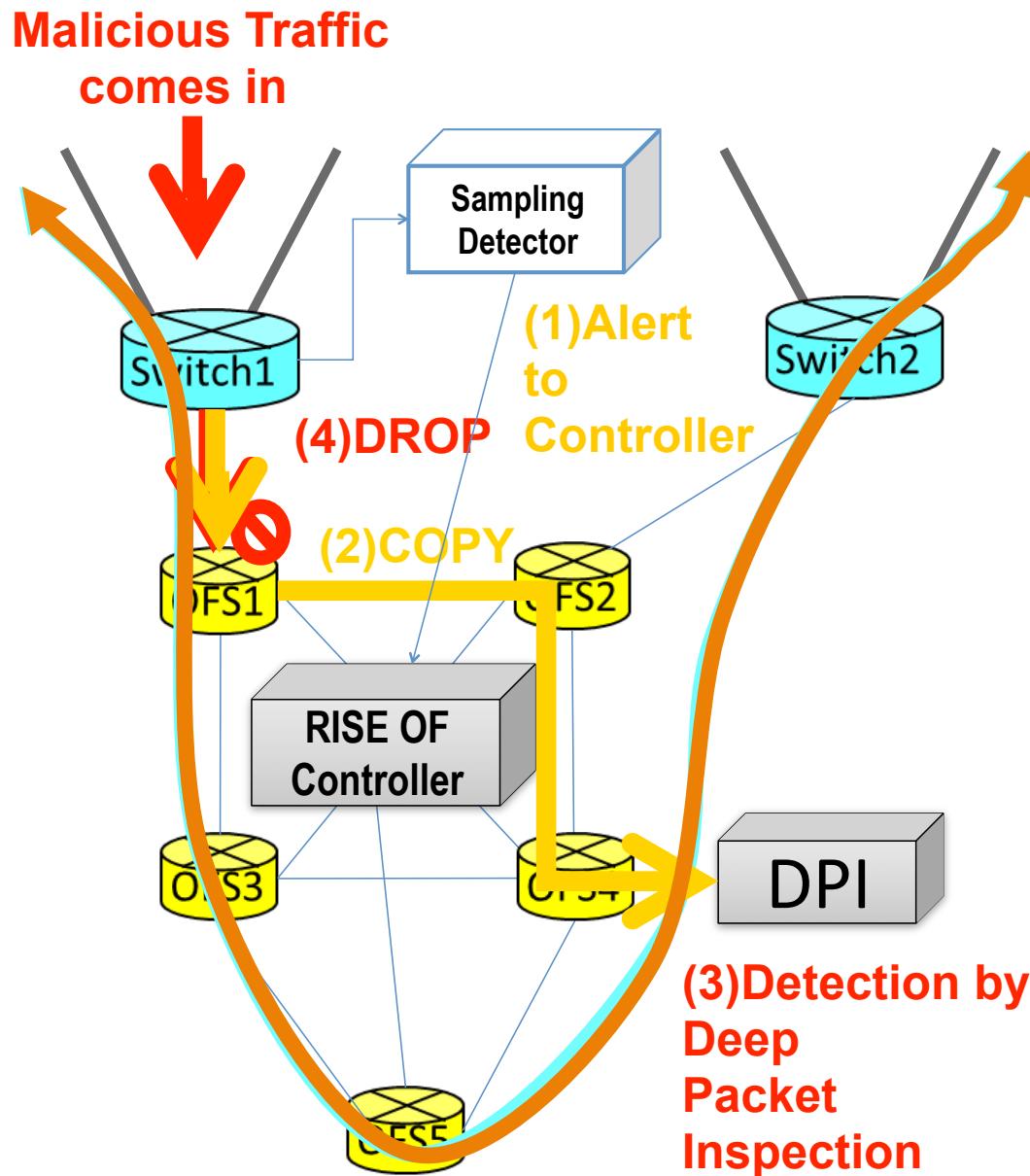
*Mobility*



# Data Center network



# Interop Tokyo 2012 Demonstration

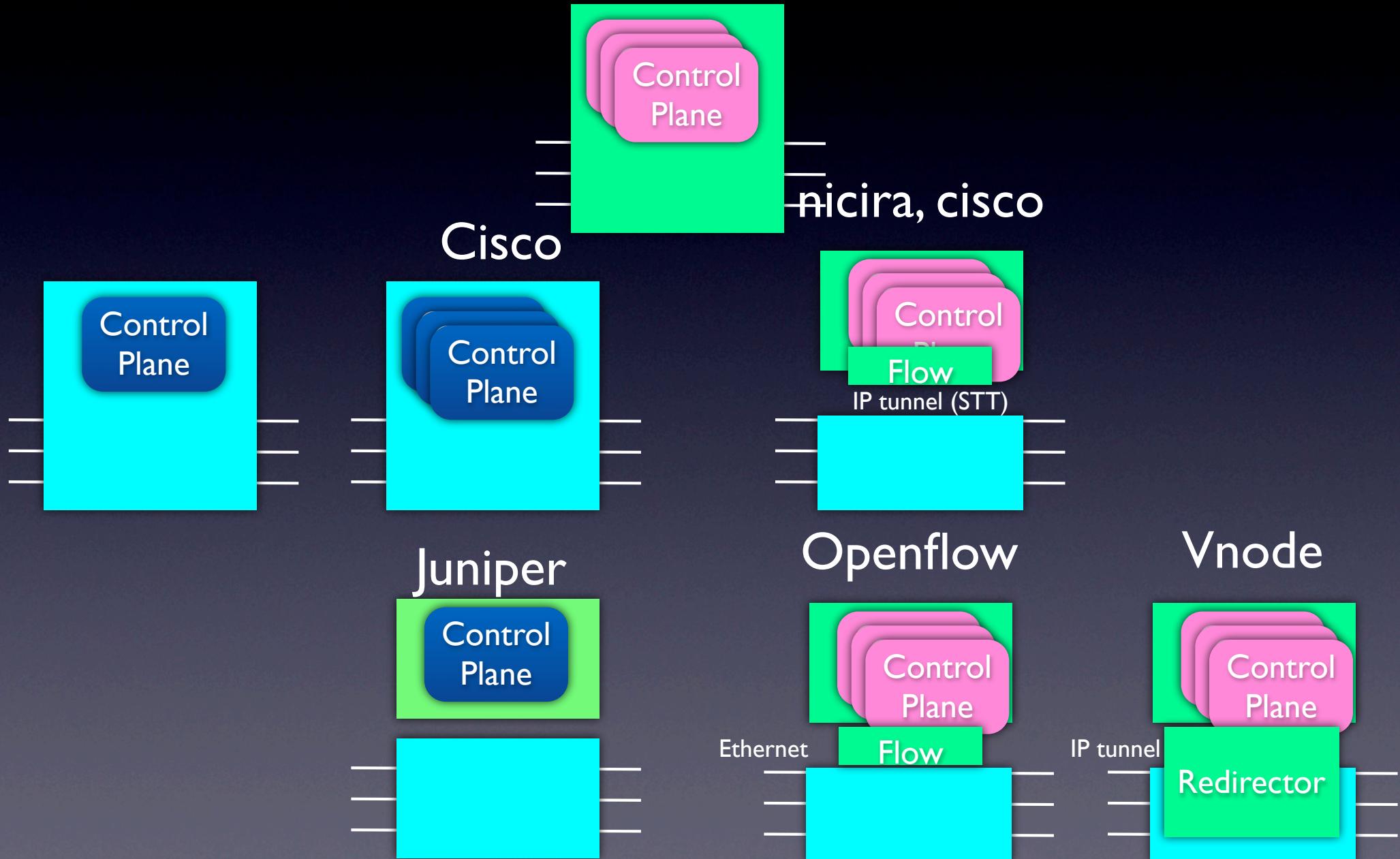


## Demonstration Scenario:

- Malicious traffic comes in, and sampling detector finds suspicious traffic flow
- Detector sends alert to RISE OpenFlow Controller
- RISE Controller copies suspicious flow to DPI machine
- RISE Controller drops the flow

# Evolution of router architecture

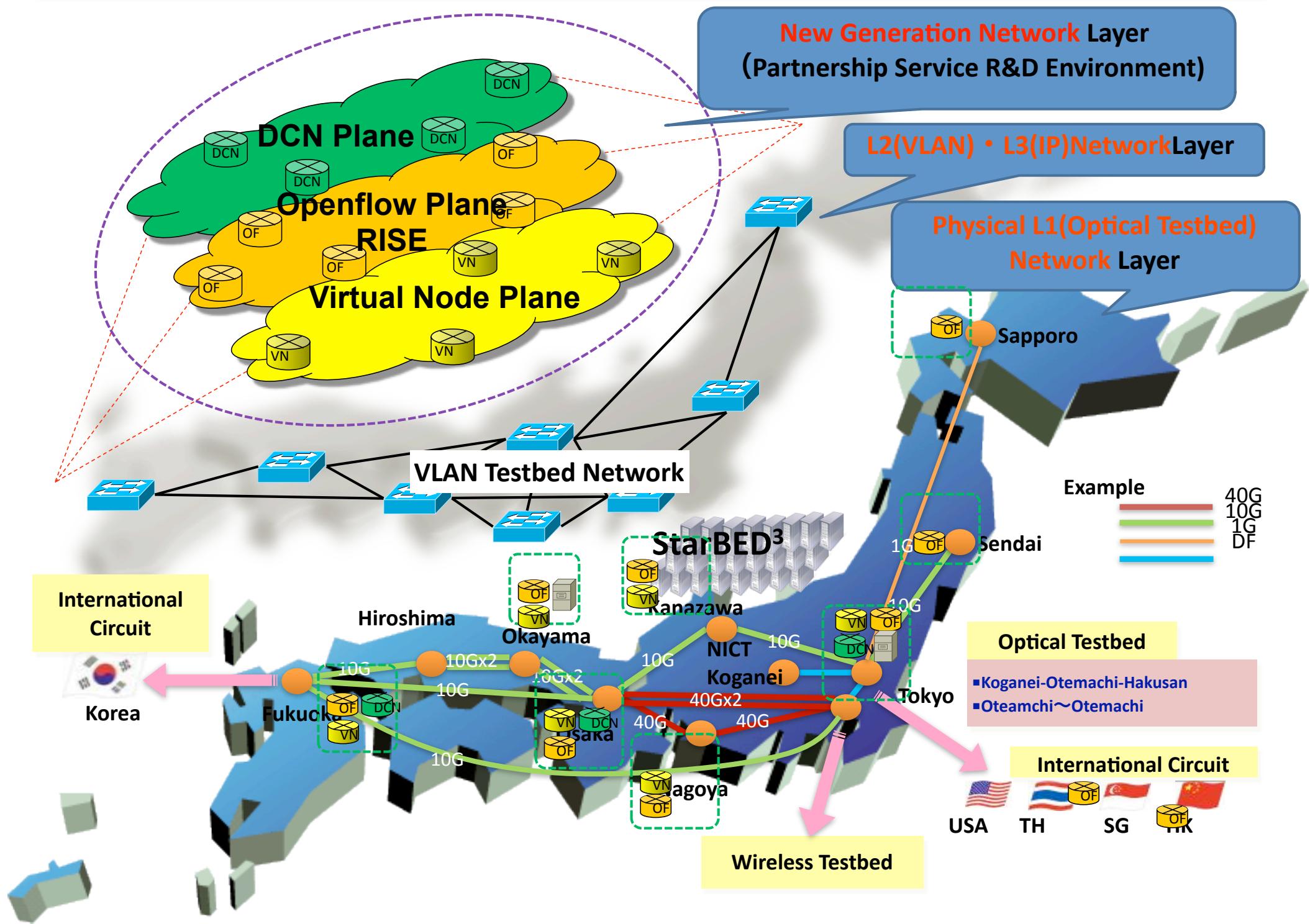
Vyatta, midokura, Junoshere



# SDN makes network design easy

- SDN gives you
  - programability of network design
  - central control of all network device
  - network virtualization
- Network Operation and configuration becomes centralized.
- Effective sharing of resources brings reduction of CAPEX&OPEX.
- Advanced network control such as
  - path aggregation for maintenance
  - selective use of network appliance
  - dynamic path addition for scale out

# JGN-X Network Infrastructure Overview



# What is the StarBED (\*BED) ?

- It's an Internet/Ubiquitous System Emulator/Simulator
  - Re-configurable cluster supporting various user requirements
  - Large scale
    - over 1000 PCs



# Our challenge

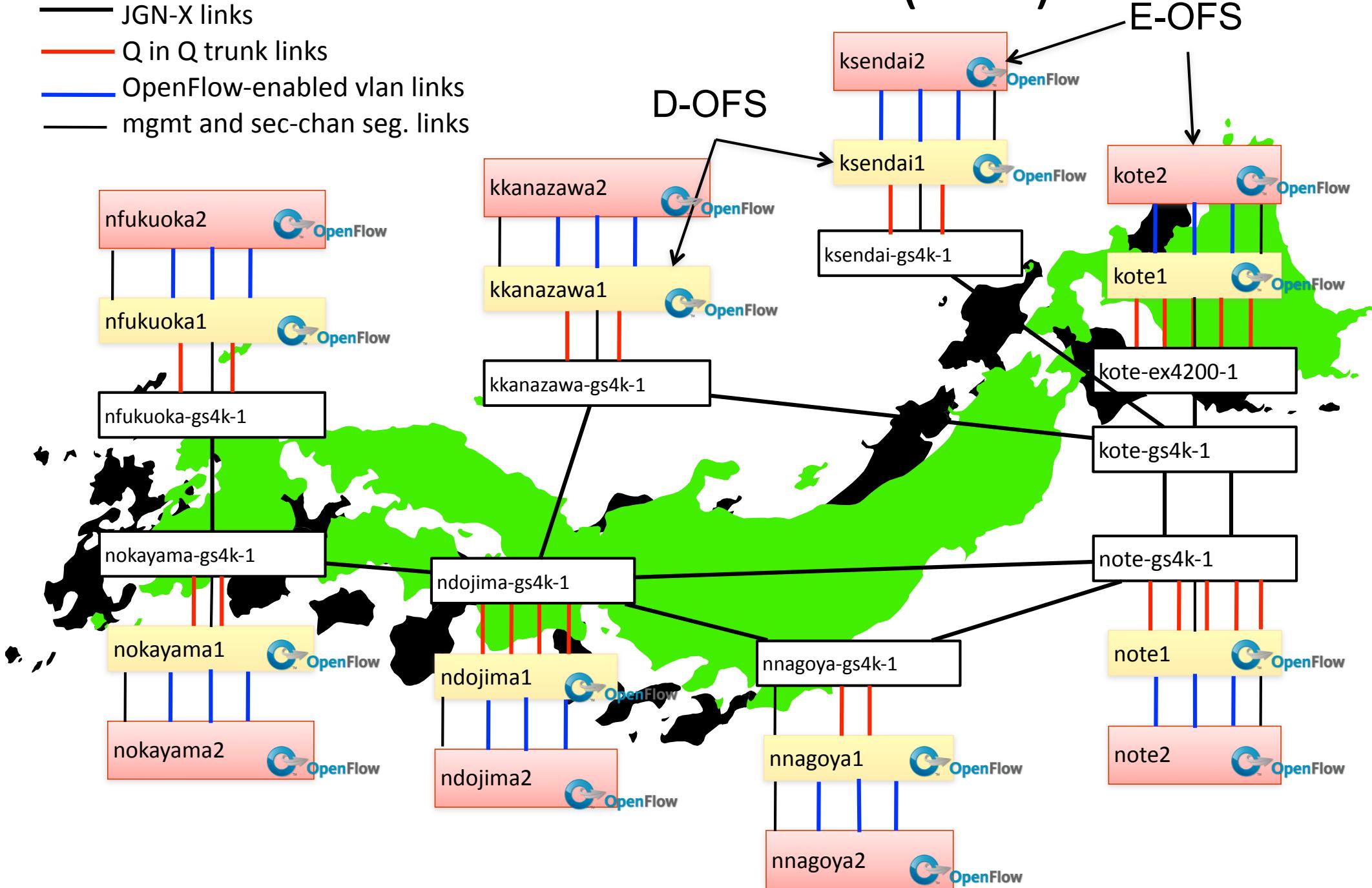
- Is SDN possible in wide area
  - Technology
  - Management
  - Business model

# RISE: OpenFlow/SDN Testbed

- A large-scale OpenFlow/SDN testbed on JGN-X
  - for researchers, students, engineers, operators, ...
- Challenges
  - Initially, demonstration of wide-area deployment of OpenFlow technology (since 2009)
    - Design of under-lay network architecture (tunneling)
    - Know-hows in shooting OpenFlow network troubles
  - Now, development of OpenFlow testbed with ...
    - User isolation that allows user-defined controllers
    - Interconnection of various testbed facilities such as network emulator, wireless/optical network testbed, cloud infrastructure, ...

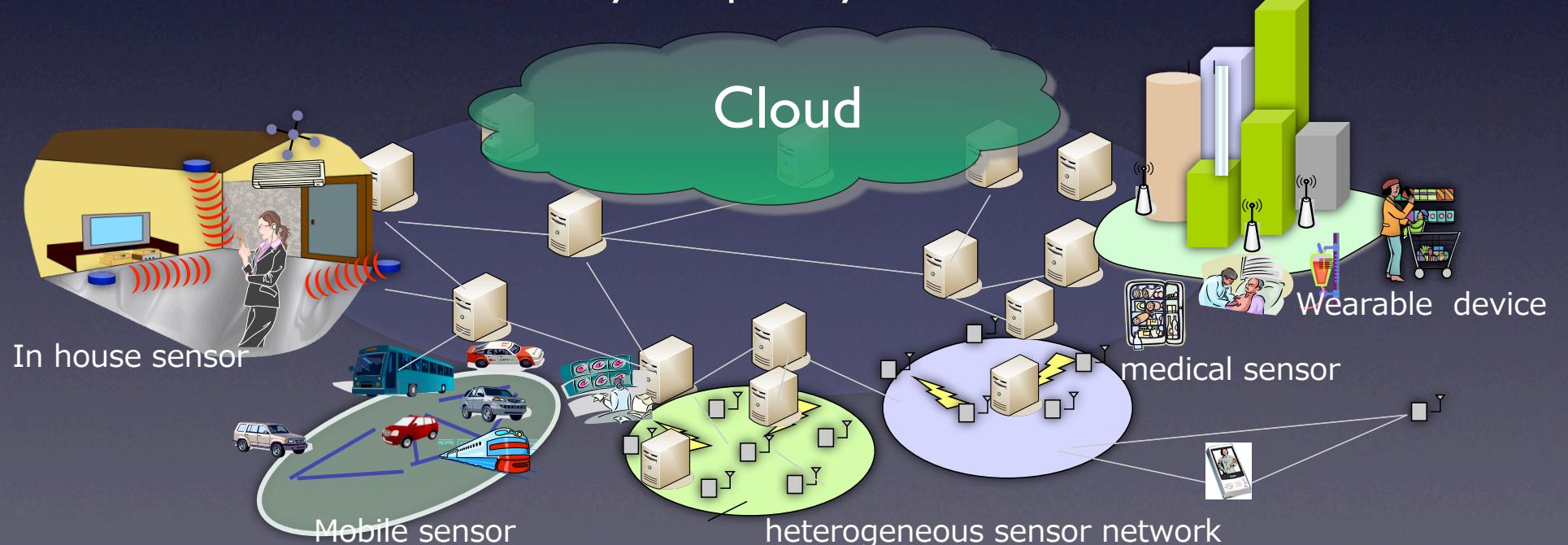
# RISE infrastructure (v1.0)

- JGN-X links
- Q in Q trunk links
- OpenFlow-enabled vlan links
- mgmt and sec-chan seg. links

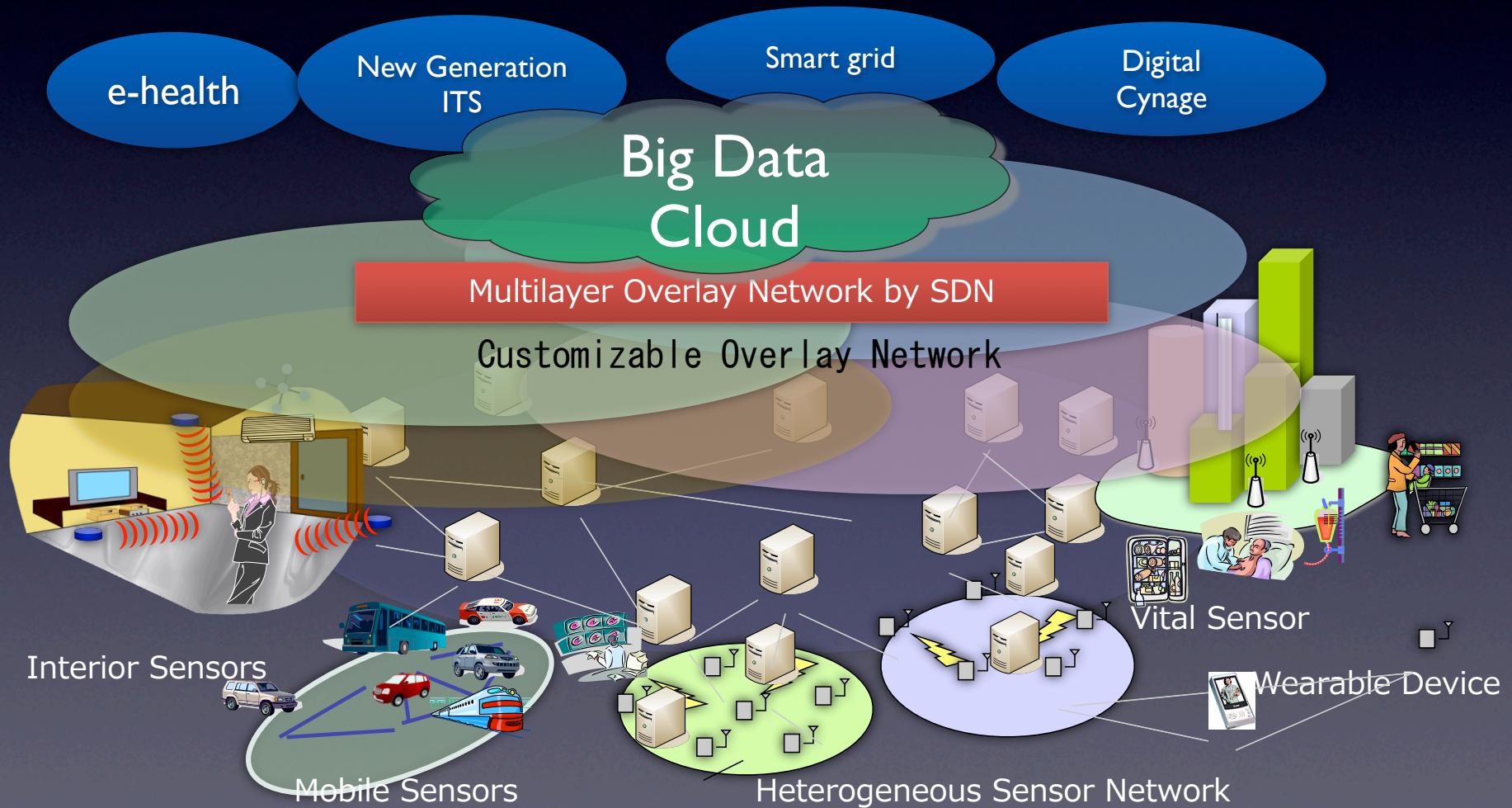


# Smart X = Cyber Physical System with Big Data

- IoT (Internet of Things)
  - billions of data sources, large scale data
  - mobility, charn
  - feature as a group or a set
  - a new form of security and privacy

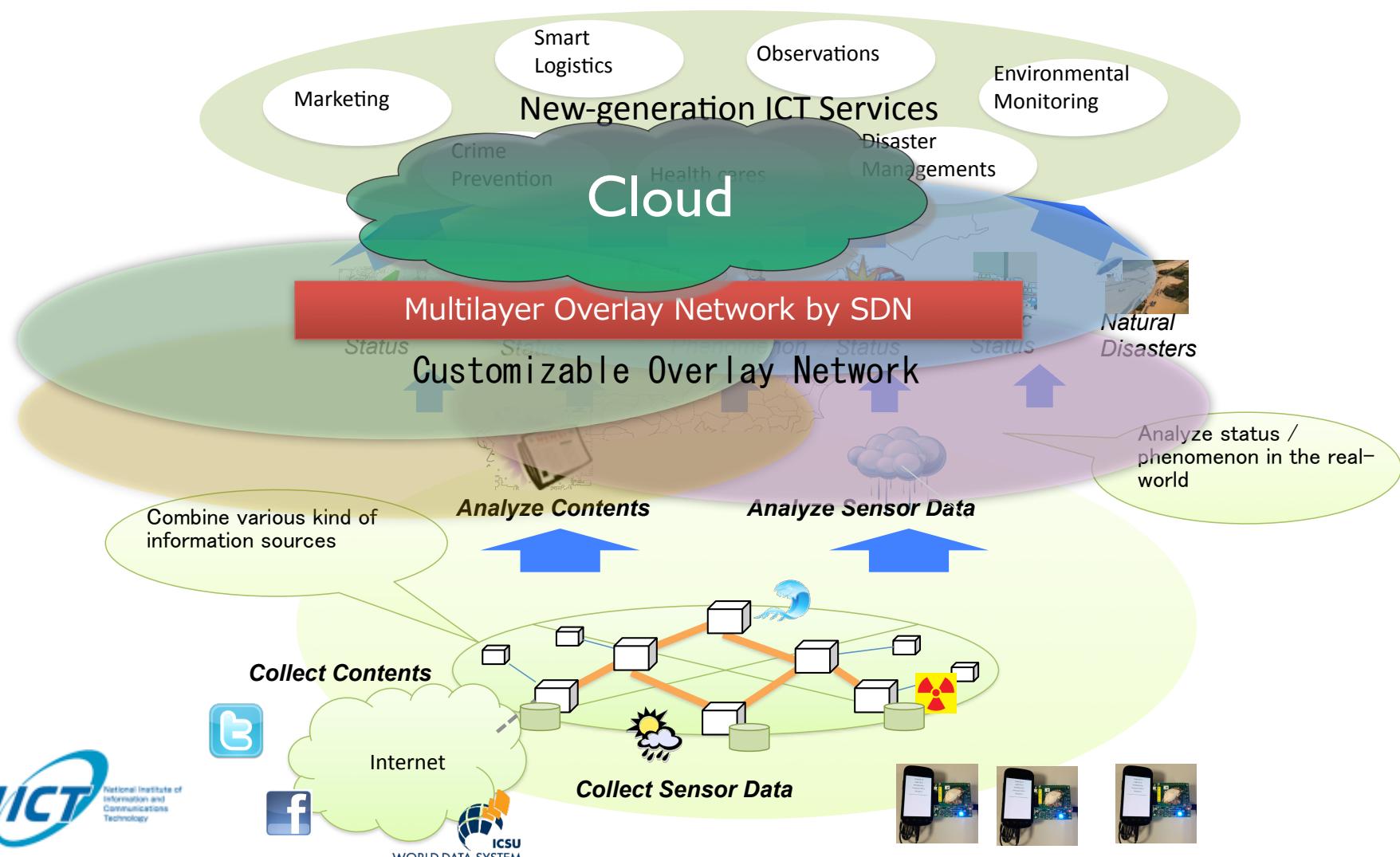


# Smart X = Cyber Physical System + SDN



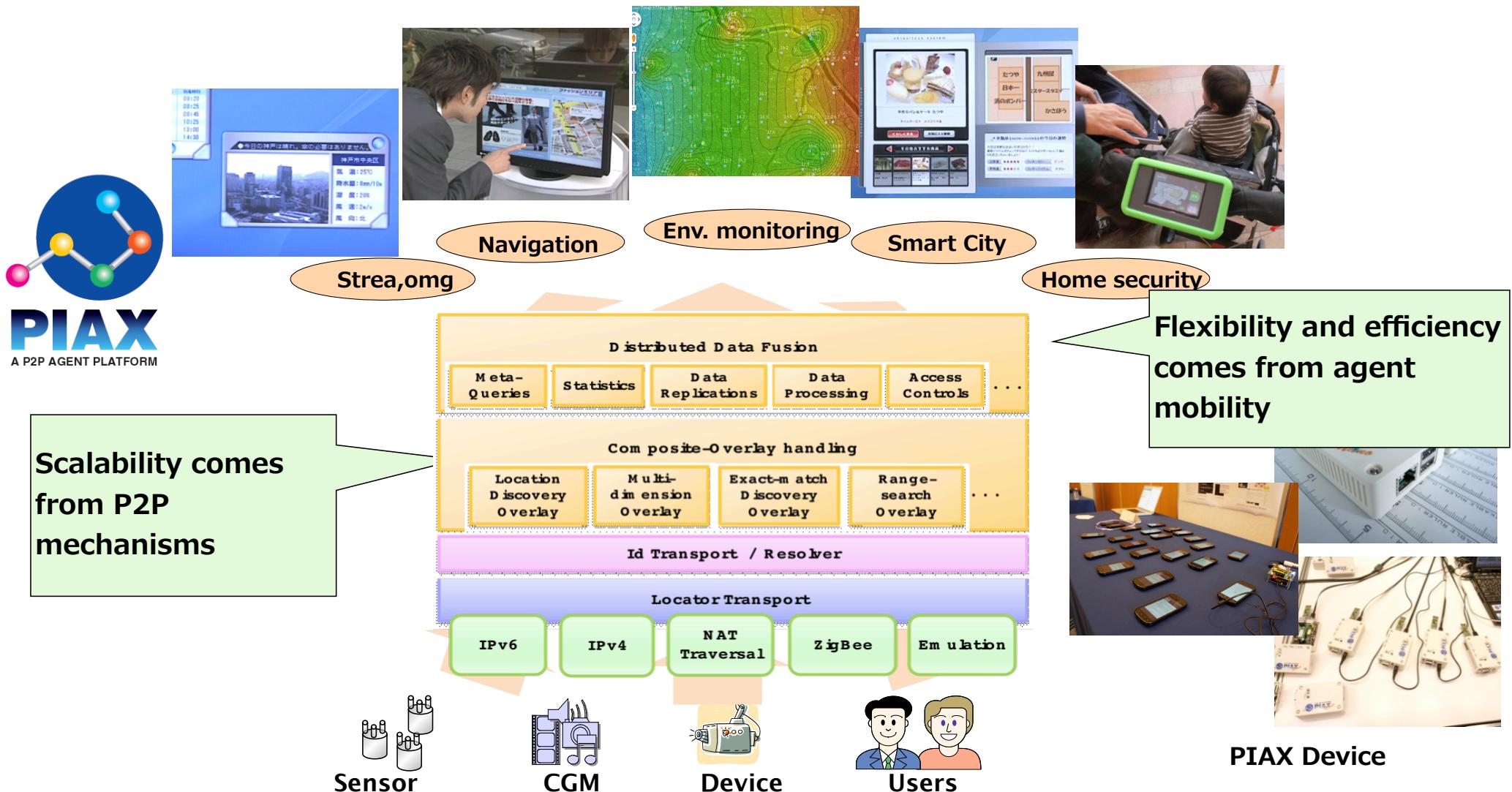
# Ongoing research project:CPS in MEXT Platform for the real-world analysis

The aim of the project is to provide a platform for real-world analysis using various data sources such as sensors, contents of the social network services, etc.



# PIAX: Overlay Agent Platform

- Fusion of overlay network and agent platform
- Middleware for highly distributed processing
- Open source <http://www.piax.org/>



## PIAX Large scale data management infrastructure

100,billion data processed in 1 million nodes

finding nodes in structured P2P

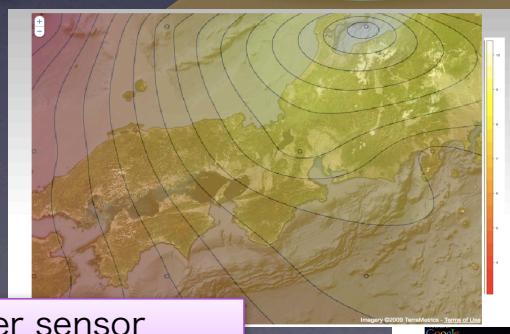
**finding nodes in structured P2P network based on POI**

Distributed nodes

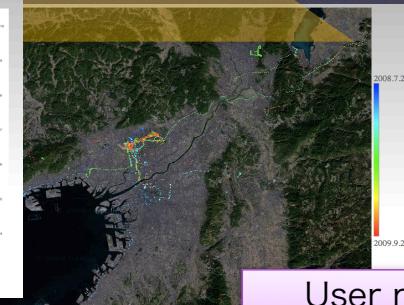


Retrieving data from the nodes

POI by area



Weather sensor



User motion data

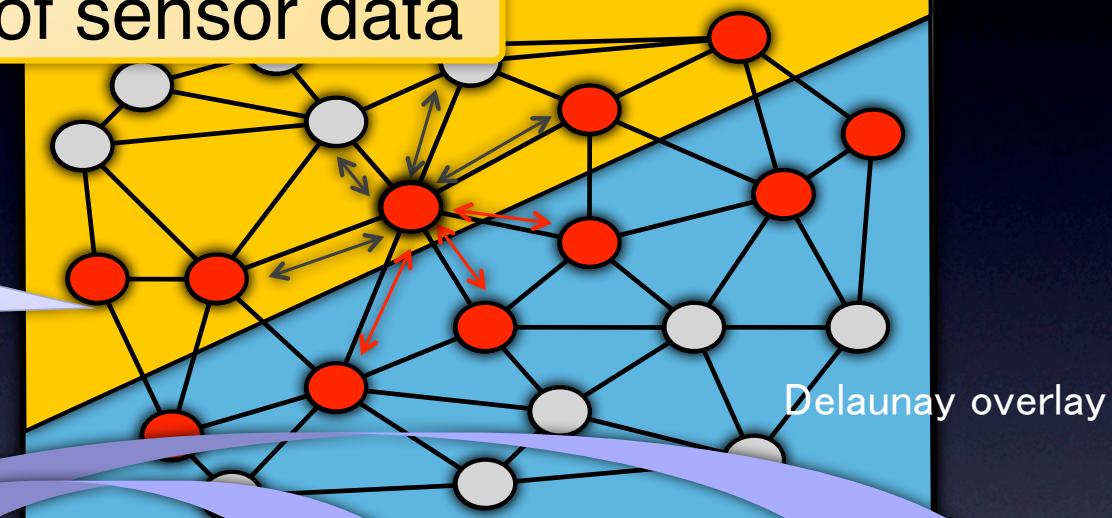
# Distributed Data fusion

- Efficient data retrieval and analysis from overlay network
- Control of data granularity based on the interest

## Macroscopic view of sensor data

Compare data with neighbors

Reducing no. of peer responding using comparison



Actual

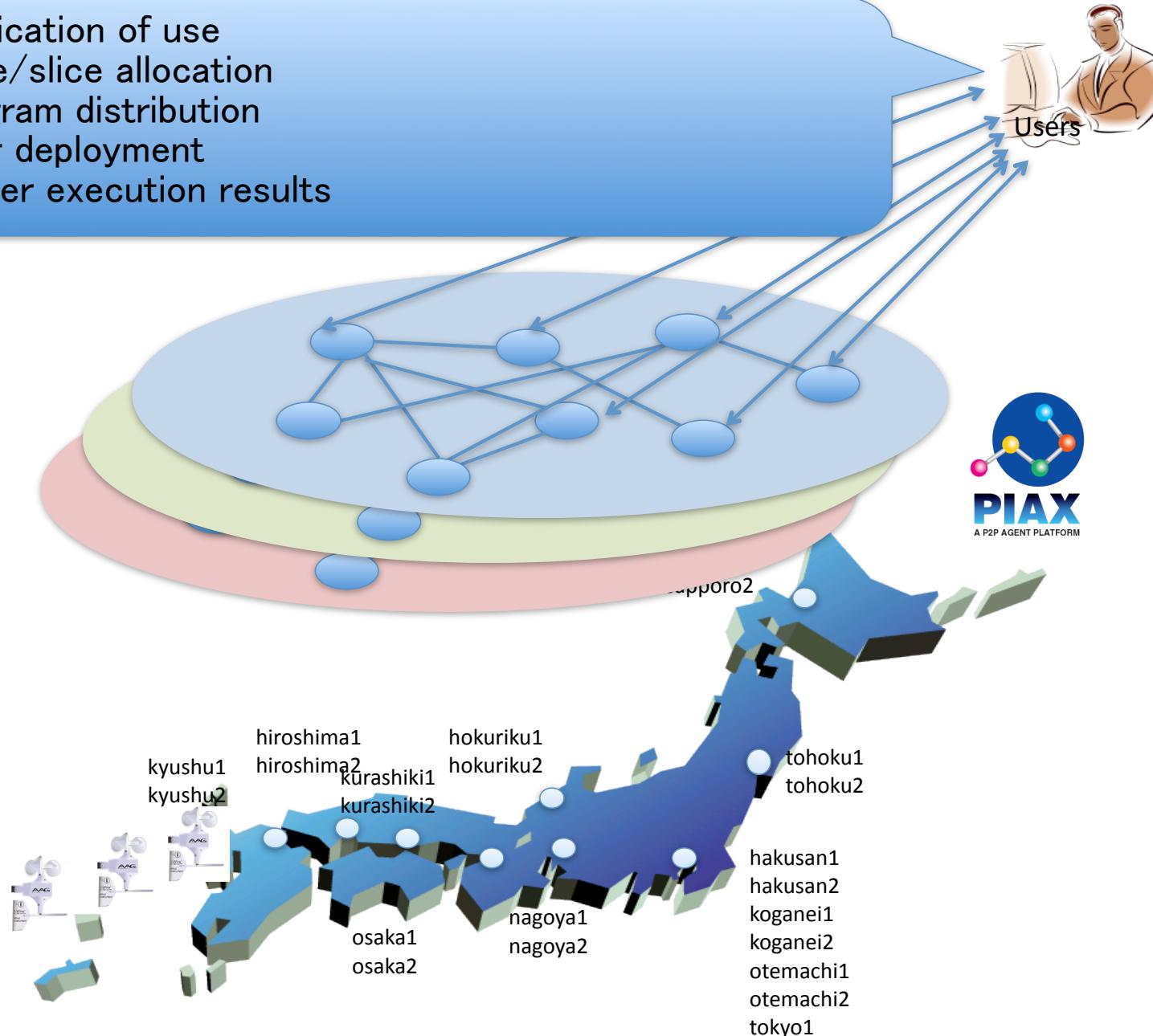
Random sampling  
(response peer : 1000/10000)

Our Method  
(Response peer : 257/10000)

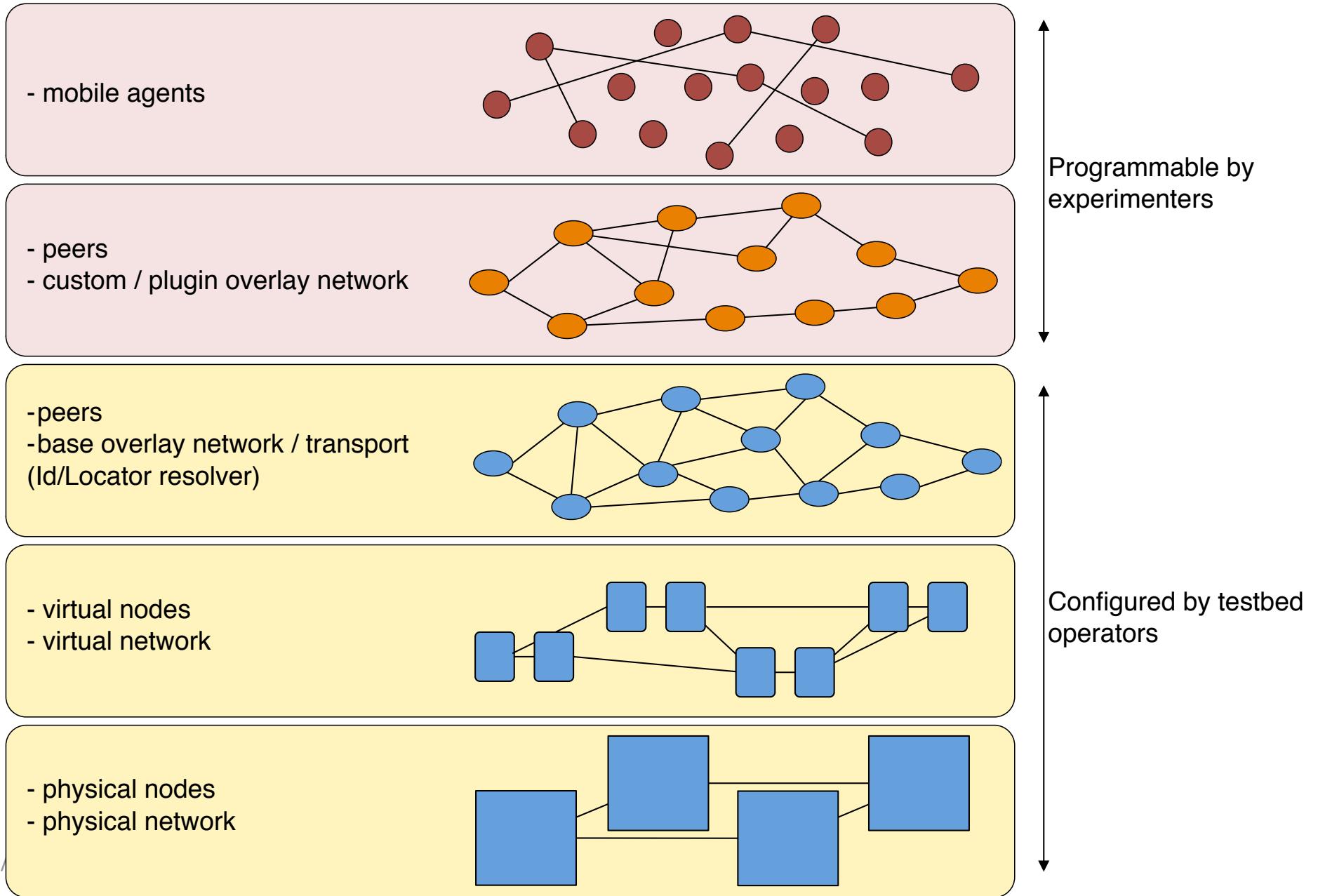
missing data

# PIAX Testbed on JGN2plus

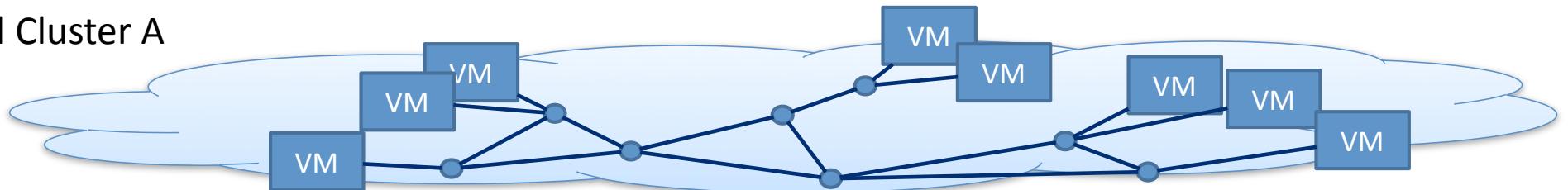
application of use  
node/slice allocation  
program distribution  
peer deployment  
gather execution results



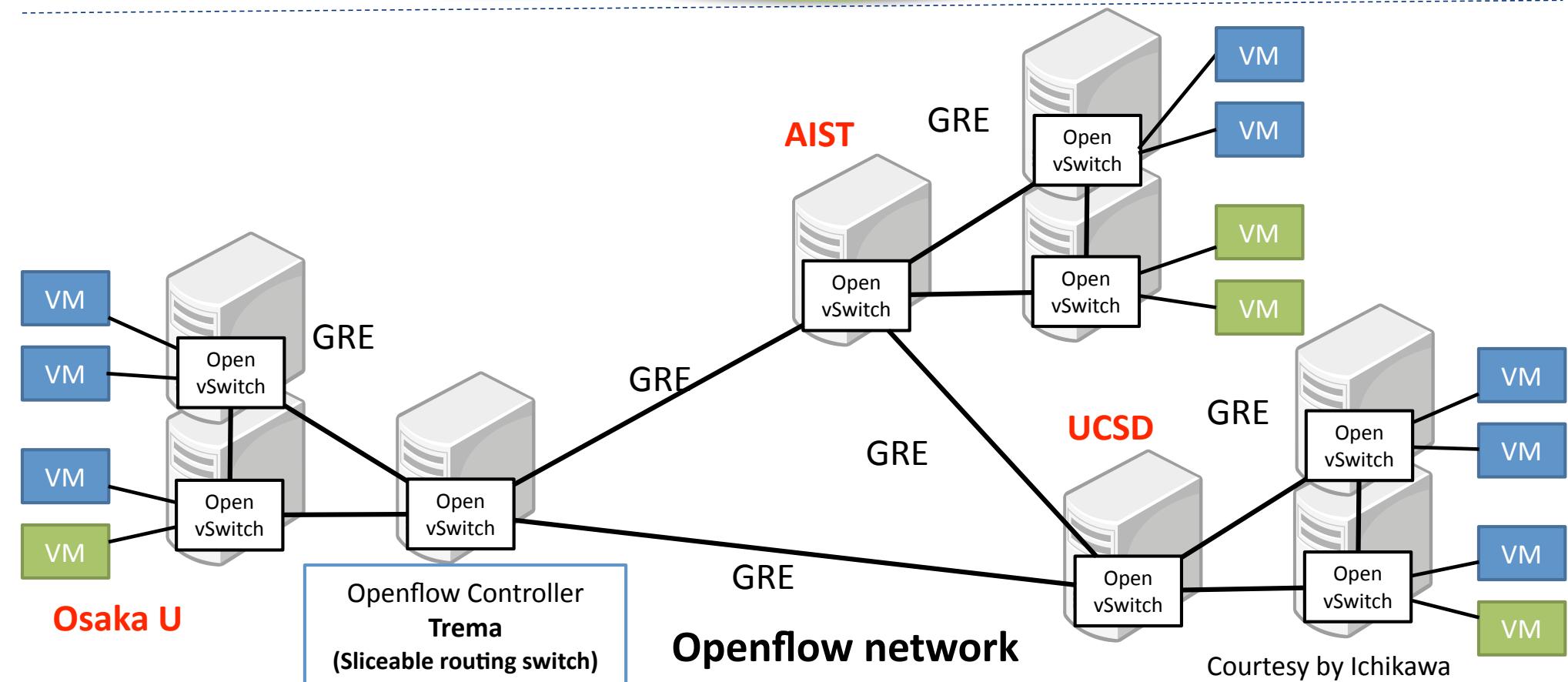
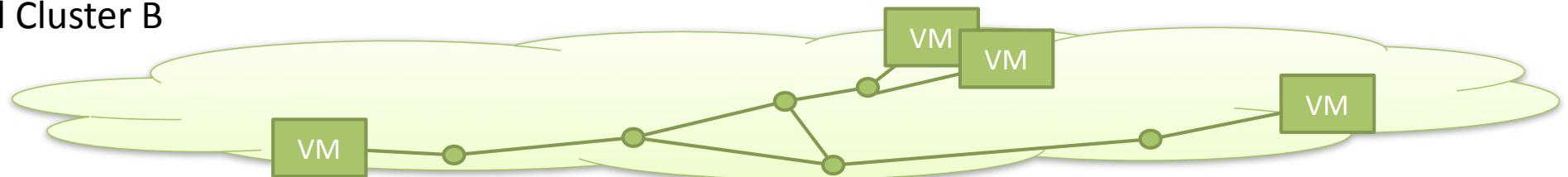
# Layers of the PIAX testbed



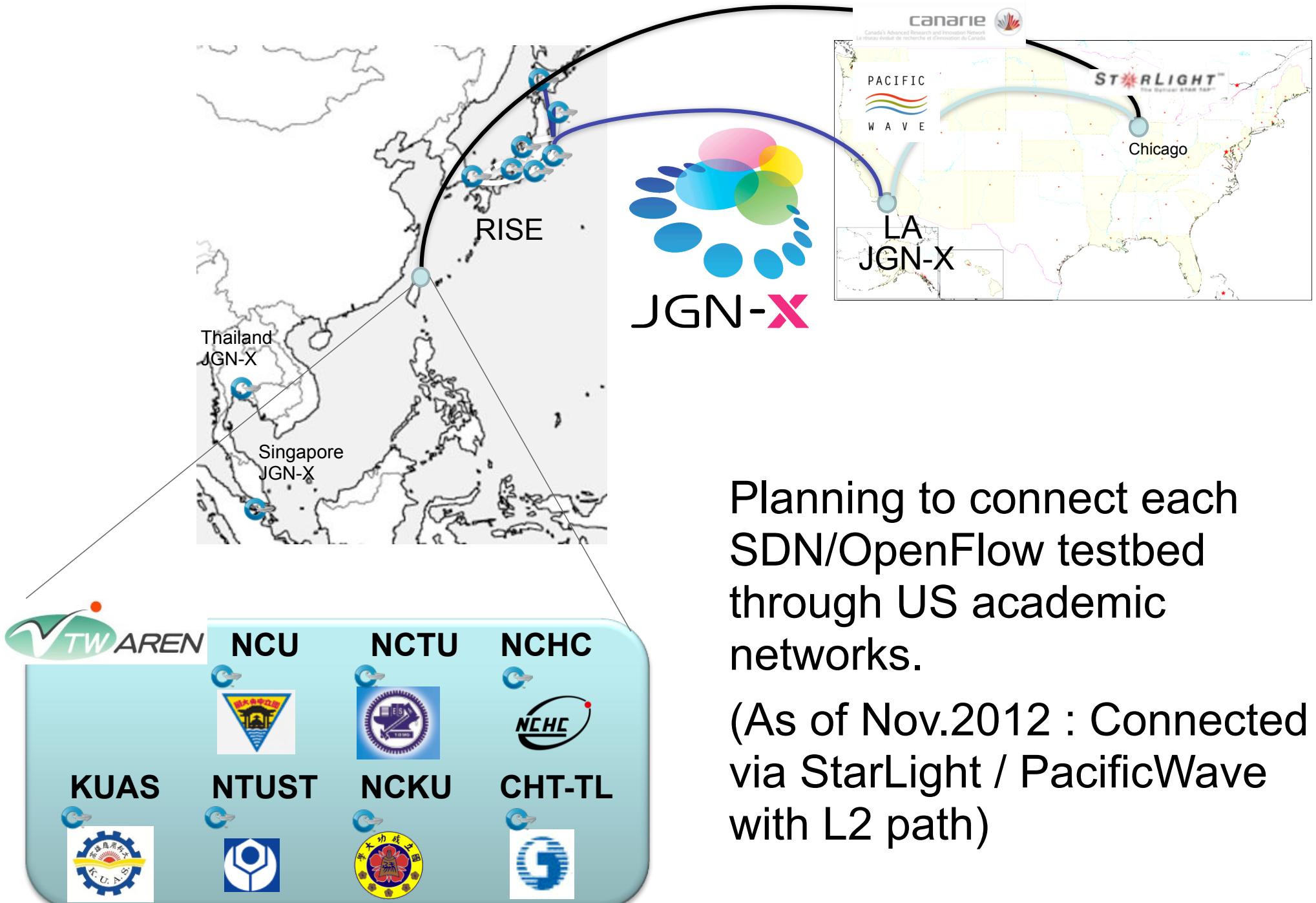
Virtual Cluster A



Virtual Cluster B



# Collaborating with TWAREN



# Global Testbed is the field



Thank you

Presented version is here!

<https://sites.google.com/site/sshimojo/talks>