CORE
A Network Embedded Test-bed for Future Network Architecture Research

Aki Nakao

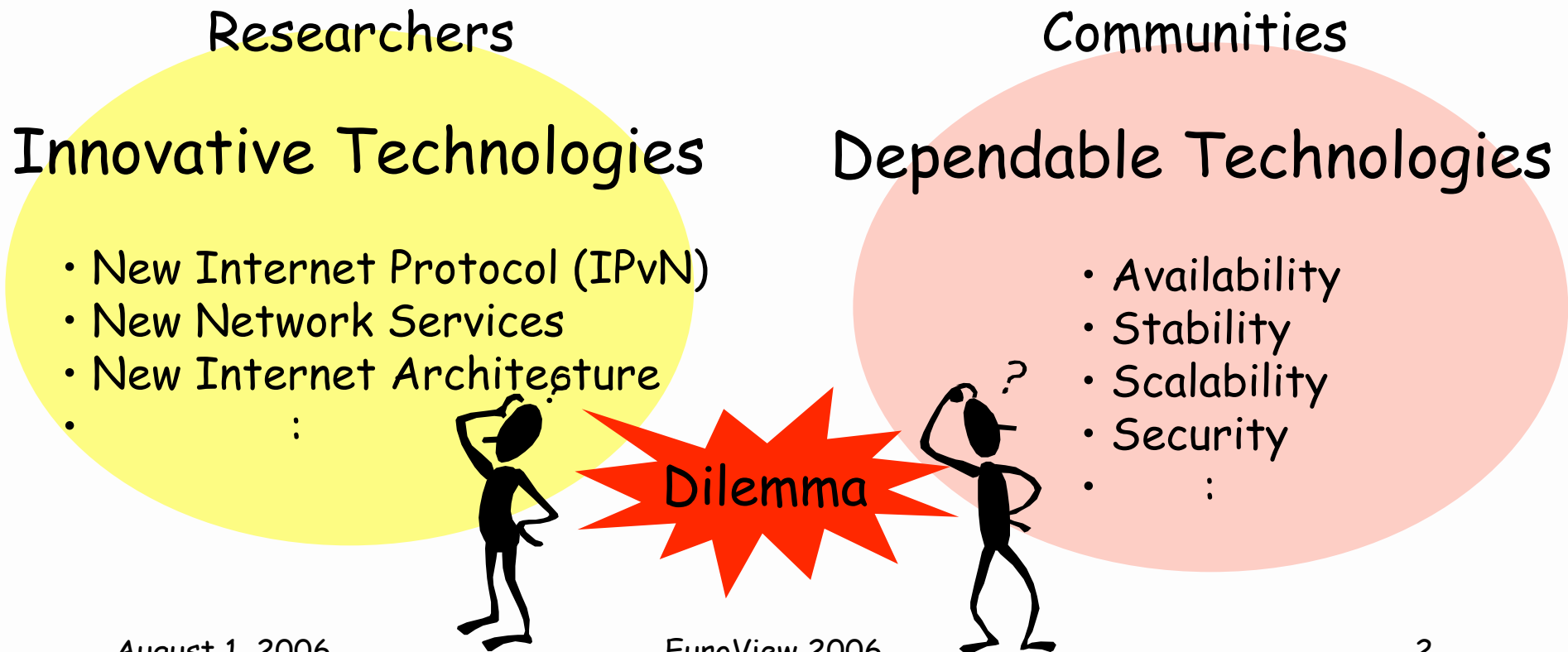Associate Professor

The University of Tokyo

"Ossified Internet"

...successful and widely adopted technologies are subject to ossification, which makes it hard to introduce new capabilities...

(U.S. National Research Council Report 2001)

Researchers

Innovative Technologies

- New Internet Protocol (IPvN)
- New Network Services
- New Internet Architecture
- :

Communities

Dependable Technologies

- Availability
- Stability
- Scalability
- Security
- :

?

Dilemma

PlanetLab

- Planetary-scale overlay network to the rescue
- Network experiments w/o changing the Internet
- 700 nodes span across 30 countries / 330 sites
- 600 slices (projects) running on top of PL
- A prototype for GENI

Network Embedded Test-beds

- PlanetLab
- GENI
- Private PlanetLab (PL-based test-beds)
  - OneLab (France)
  - EverLab (Israel)
  - PlanetLab@EPFL (Switzerland)
  - PlanetLab@Utokyo (Japan)
  - :

Private PlanetLab

- A tool to build our own "private" PlanetLab
  - pl_box: planetlab in a box
  - MyPLC: my little planetlab

PlanetLab Node

"slice"

The Internet

PLC (PlanetLab Central)

/plc

Linux

PlanetLab Architecture

MyPLC
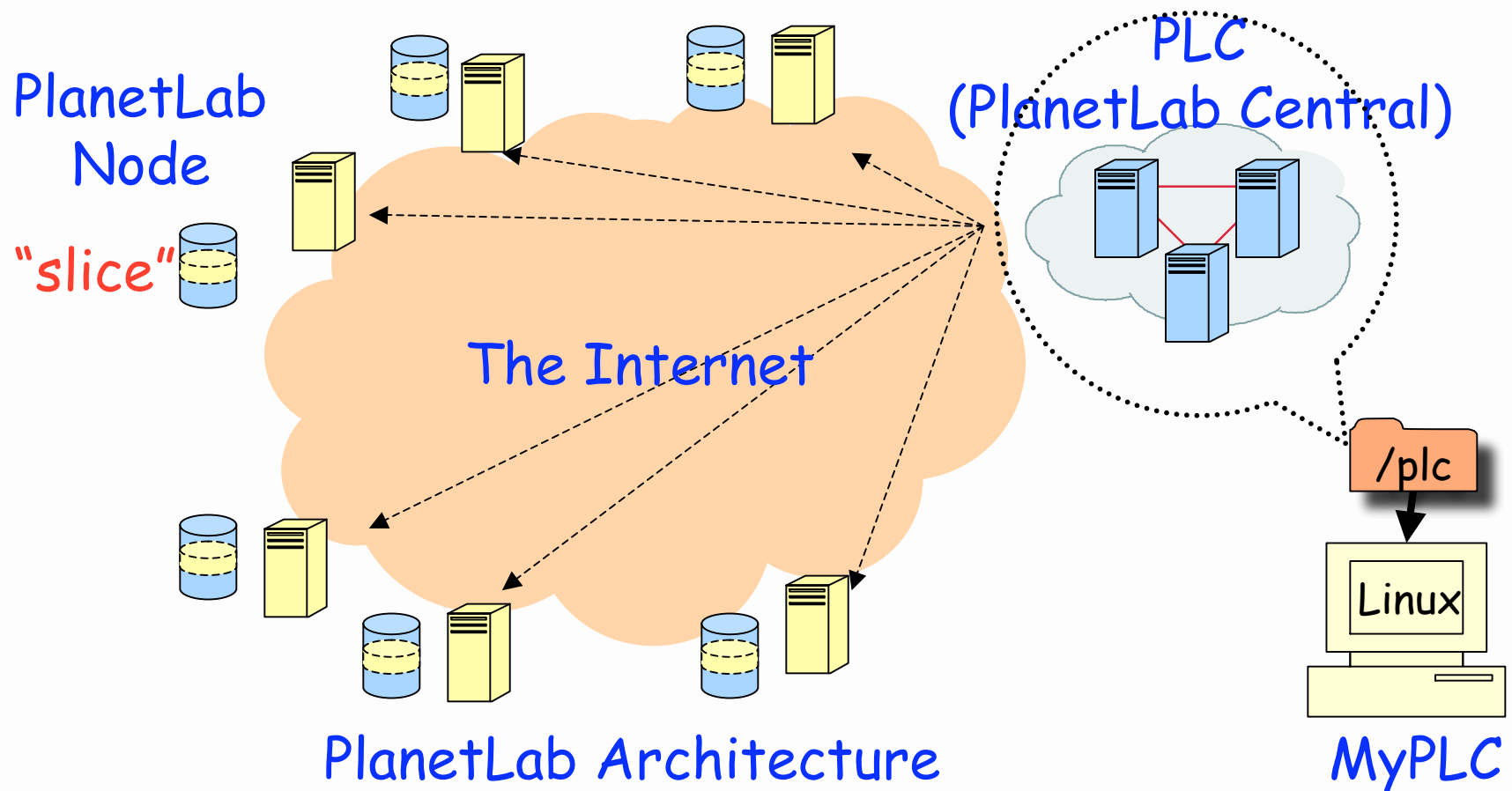
CORE

- **C**ollaborative **O**verlay **R**esearch **E**nvironment
  - Planning to build a domestic overlay test-bed
  - Based on Private PlanetLab tool
  - Provision resources for local & mission critical experiment
- Features we would like to have...
  - PoP/Core collocation (nodes "inside" network)
  - Resource provisioning
  - "Network Slicing"
- Other features
  - Custom h/w, wireless subnets
- Federation, peering, and collaboration
  - Public PL, Other Private PL, and GENI

CORE's Scope

- Target overlay research
  - Not just on distributed system apps
  - More on network core architectures
- Catalyze domestic/global research
  - CORE tailored for targeted researchers
  - Utilize both private/public environments
- Moving towards GENI-like environment
  - More exercises/discussions for a GENI-like environment
- Contributions to public PlanetLab
 - Sandbox for disruptive technologies

| PlanetLab | ⟶ | GENI |

Timeline ⟶   CORE ⟶ CORE?
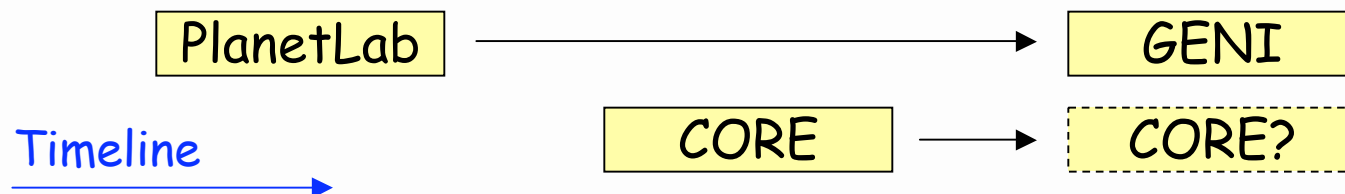
Nodes "Inside" Network

- **More cross-road nodes important in CORE**
  - Current PlanetLab has many edge nodes
  - Overlay consisting of "core" nodes more valuable
- **Overhead hindering network research**
  - ECE bounce
    - Overhead: doubles latency & loss, halves bandwidth
  - End system forwarding overhead
    - There exists a set of common overlay operations
    - Implement better overlay forwarding mechanisms on nodes inside network
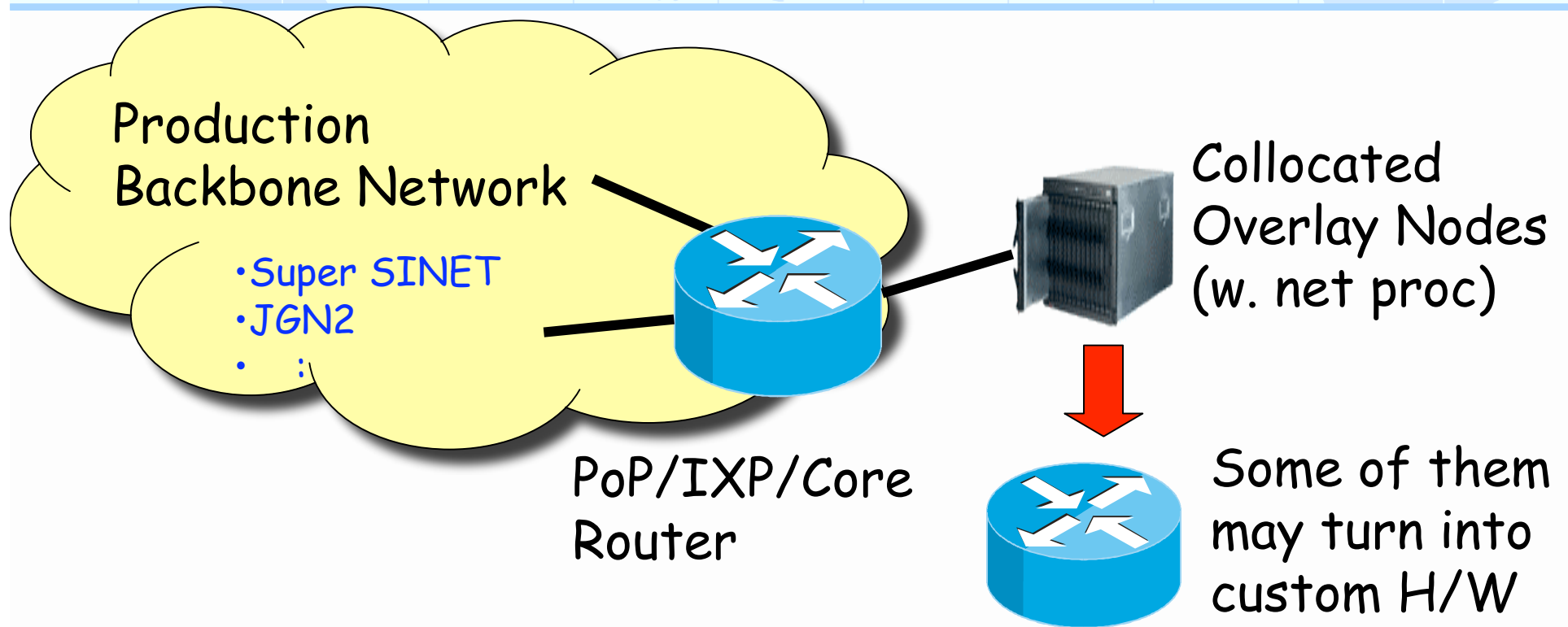
ECE Bounce

- Packet traversal overhead
    - Repeat ECE (Edge-Core-Edge) bounces
   - Inside nodes would reduce the overhead

overlay routing

Internet routing
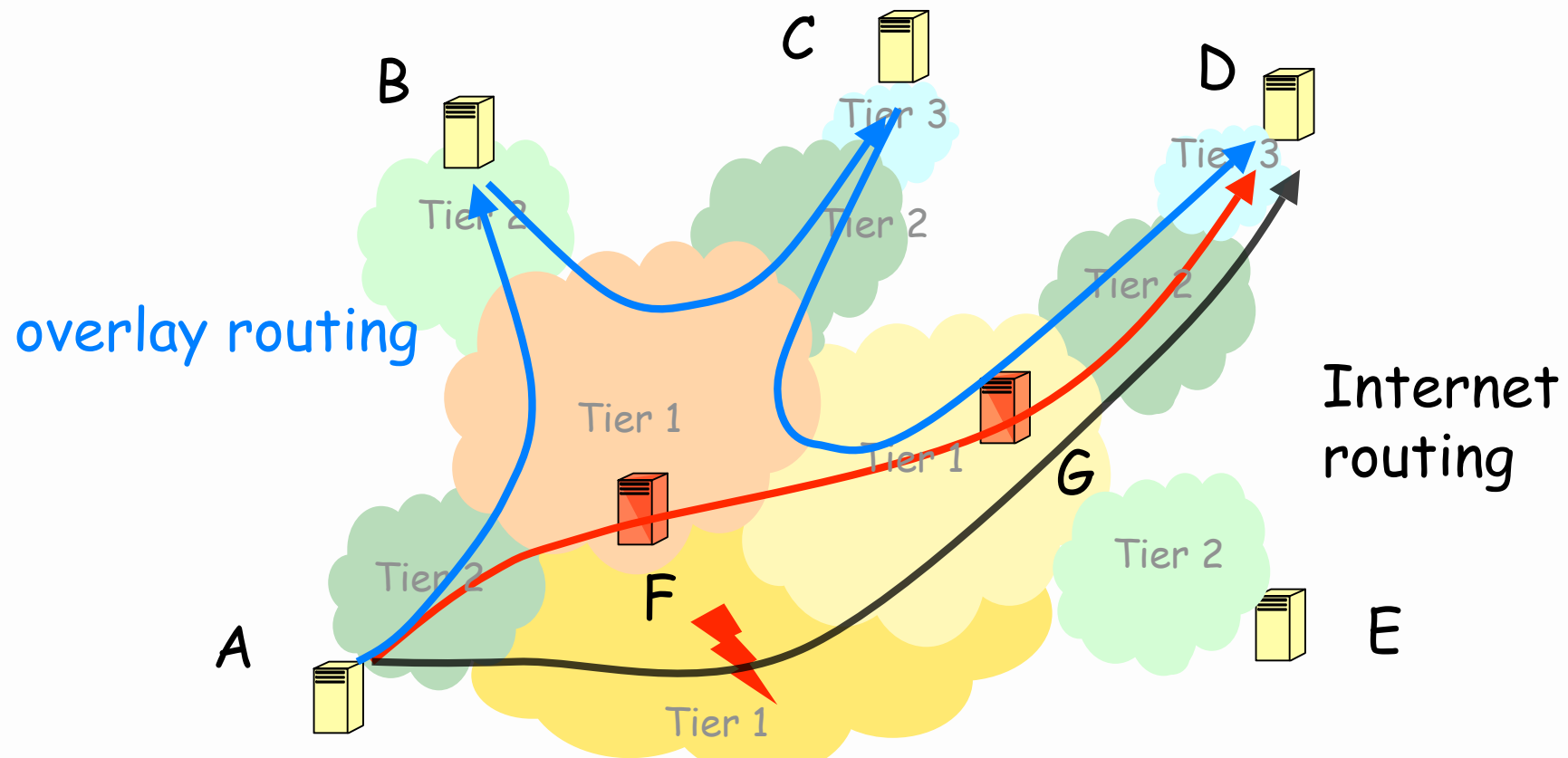
Nodes "Inside" Network

- **More cross-road nodes necessary**
  - Current PlanetLab has many edge nodes
  - Overlay consisting of "core" nodes more valuable
- **Overhead hindering network research**
  - ECE bounce
    - Overhead: doubles latency & loss, halves bandwidth
  - End system forwarding overhead
    - There exists a set of common overlay operations
    - Implement better overlay forwarding mechanisms on nodes inside network

PoP/Core Collocation

Production
Backbone Network

- Super SINET
- JGN2
- :

PoP/IXP/Core
Router

Collocated
Overlay Nodes
(w. net proc)

Some of them
may turn into
custom H/W

✕ replacing production routers with *active routers*
✓ may need GENI approach in future (using JGN2)
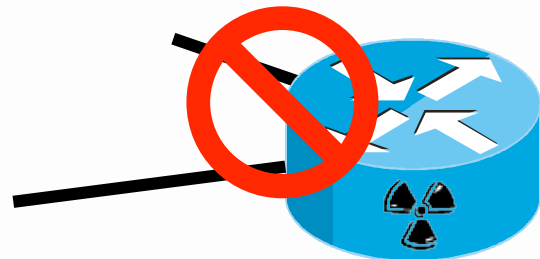
Custom Hardware

- Incrementally upgrading overlay nodes
  - PC server
  - PC server with network processors
  - PC server (w/net.proc) + Customizable Routers
  - Cluster of these
- Discussion
  - Abstraction to share additional capabilities
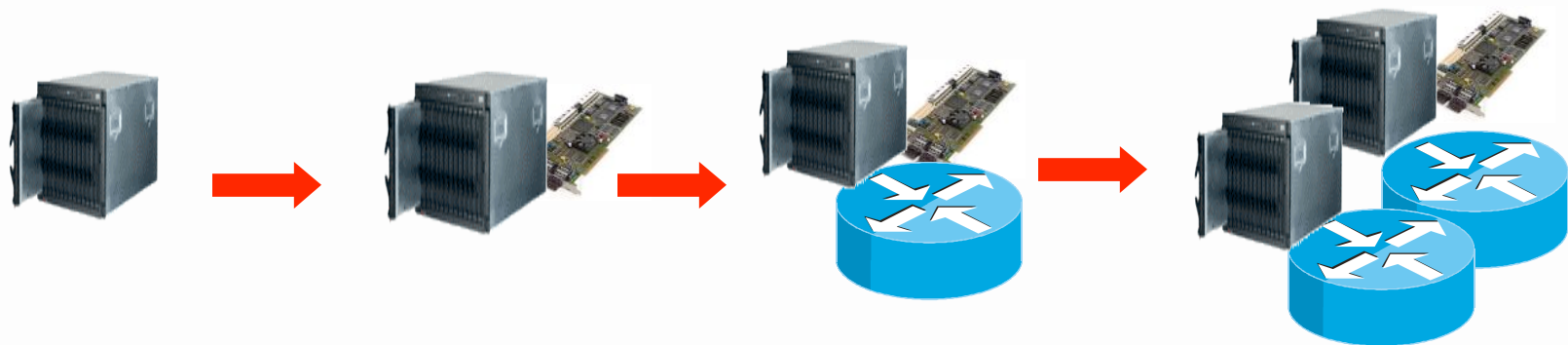
Different "Slice" Implementations

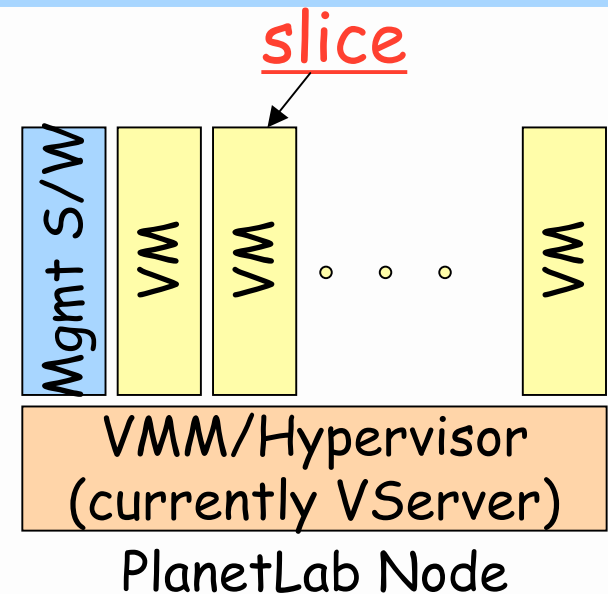- "Slice" needs to achieve...
  - Resource Provisioning
  - Flexibility
- Ideal VMM
  - Optimized in network I/O
  - Fast migration (load/unload/snapshot)
  - Single guest OS per H/W may be fine
  - Scalability may be resolved by clustering?
- Feasibility Study
  - Xlice: Xen Domain-U Slice
  - VMWare

slice

Mgmt S/W | VM | VM | o o o | VM

VMM/Hypervisor
(currently VServer)

PlanetLab Node

Slicing A Network

- PlanetLab Slice
  - A nice abstraction for sharing resources
    - CPU/Storage/Bandwidth at end hosts
- "network slicing"
  - A part of CORE may consist of lambda BB
    - Assign a "lambda" to a "slice"
    - Differentiated service model
  - Another "dual use" model
    - Experiments in the wild (PlanetLab)
    - Reproducible experiments (Emulab)

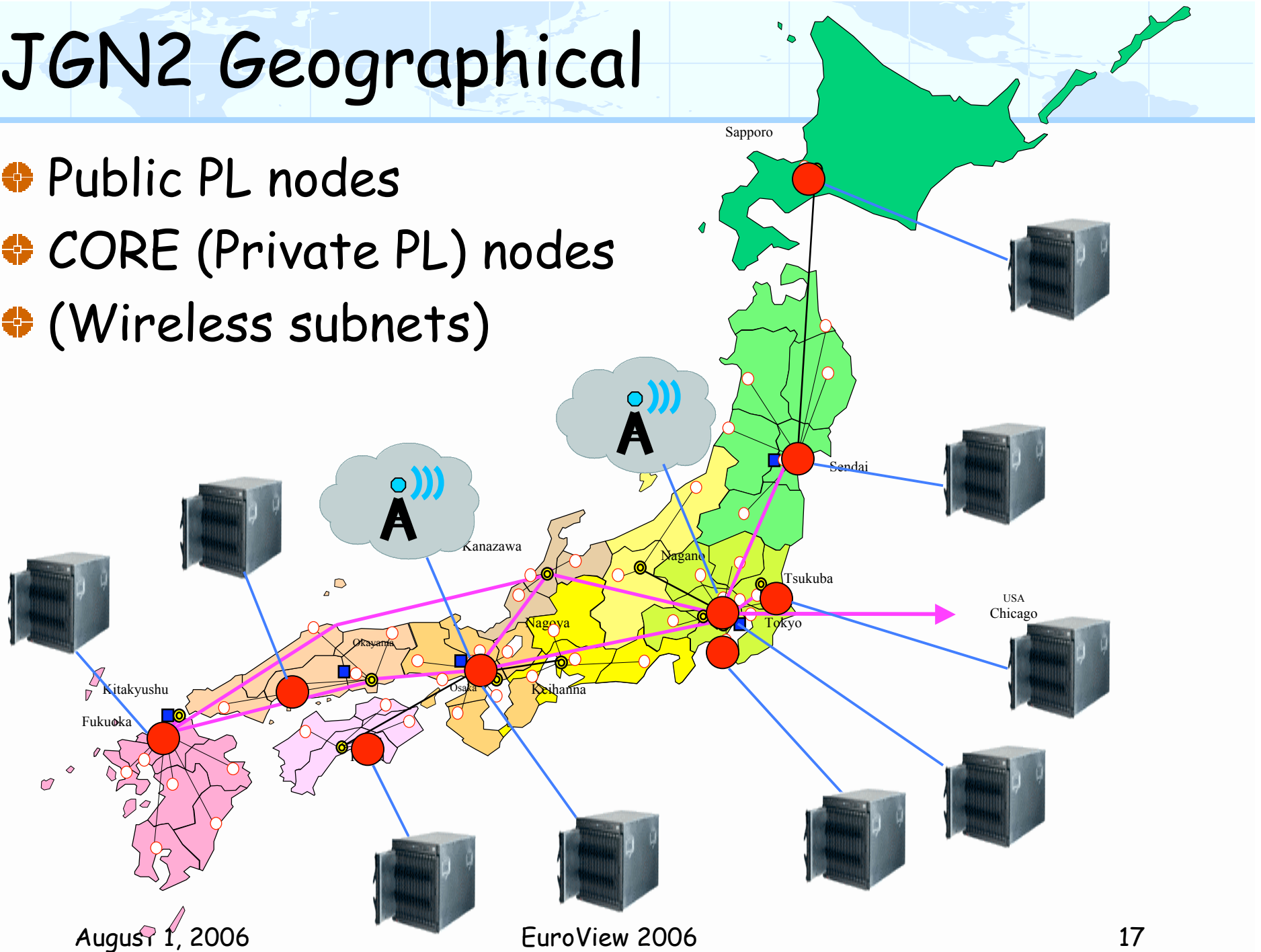# Wireless & SensorNet Subnets

- Private PL kernel to enable wireless & sensor network experiments
  - Wireless NICs support
  - Multi-homing support to enable access points
- Discussion
  - Resource Slicing
    - Resource scarcity on sensors (embedded systems)
  - Mobility/Security

Current plan on CORE

- National backbone networks (+Asian network)
  - JGN2 (JGN3 in future)
  - Super SINET (SINET-3 in future)
  - APAN
  - plus other national backbones available
- Hopefully 5-10 nodes near PoPs of national BBNs
  - PoP/IXP/core routers collocation
- Starting from a small setting…
  - 2-4 nodes will be deployed at a dozen places
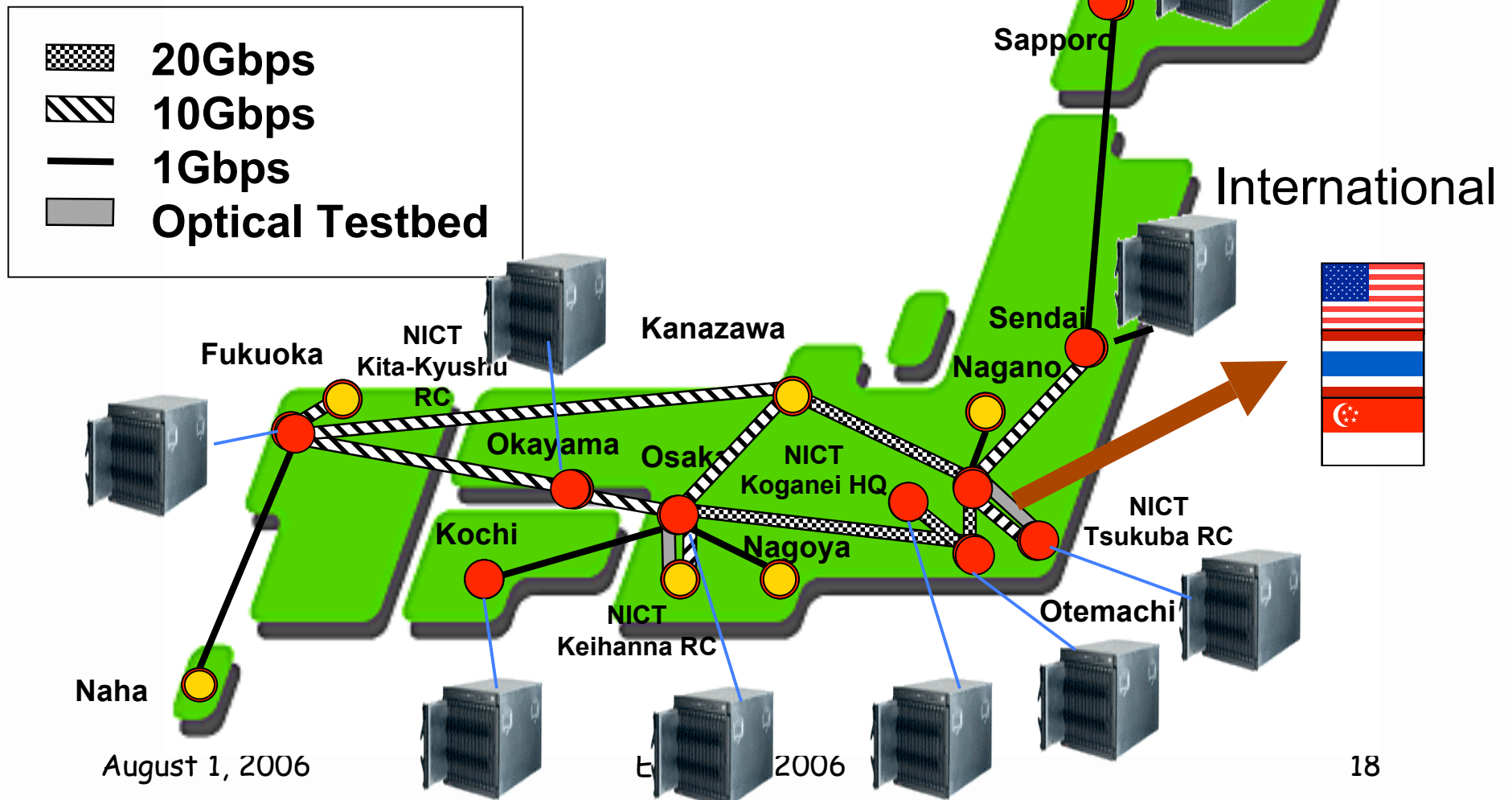 - The more funding, the more nodes :-)

JGN2 Geographical

- Public PL nodes
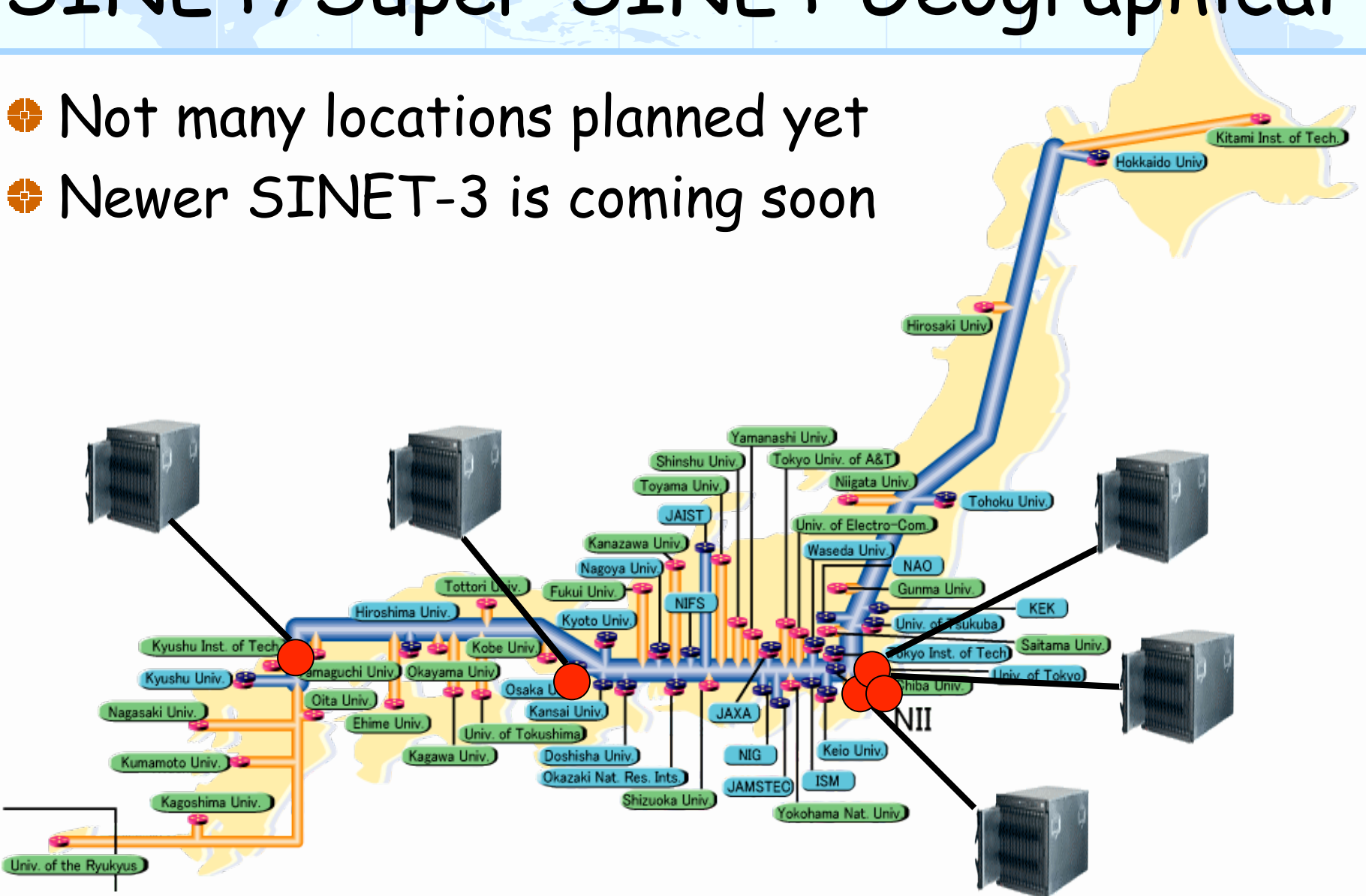- CORE (Private PL) nodes
- (Wireless subnets)

Sapporo

Sendai

Kanazawa

Nagano

Tsukuba

USA
Chicago

Nagoya

Okayama

Tokyo

Kitakyushu

Osaka

Keihanna

Fukuoka

JGN2 Topological

- Countrywide coverage w. nodes connected via 10Gbps

Legend:
- 20Gbps
- 10Gbps
- 1Gbps
- Optical Testbed

Sapporo

International

Fukuoka

NICT Kita-Kyushu RC

Kanazawa

Sendai

Nagano

Okayama

Osaka

NICT Koganei HQ

NICT Tsukuba RC

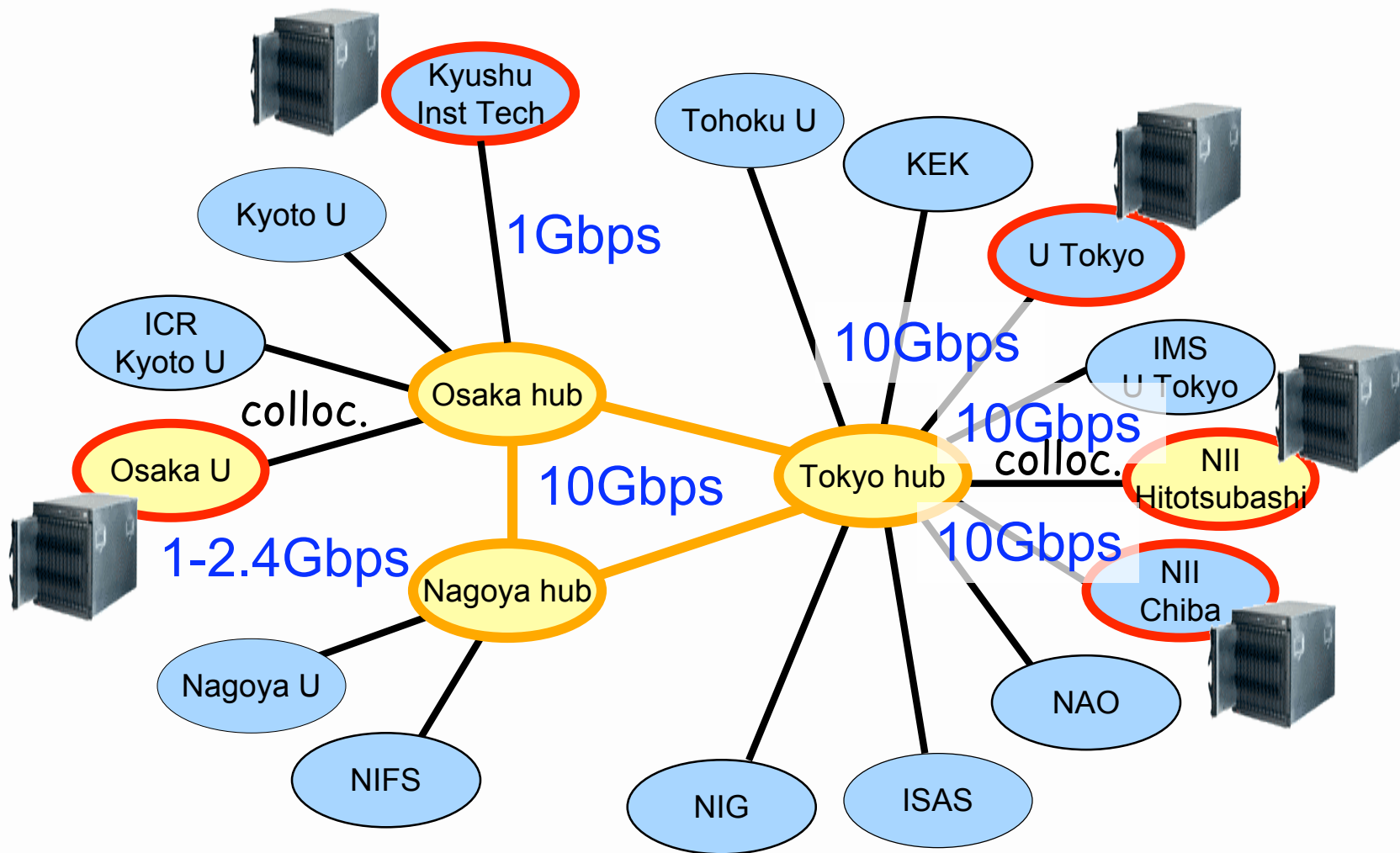Kochi

Nagoya

Naha

NICT Keihanna RC

Otemachi

SINET/Super-SINET Geographical

- Not many locations planned yet
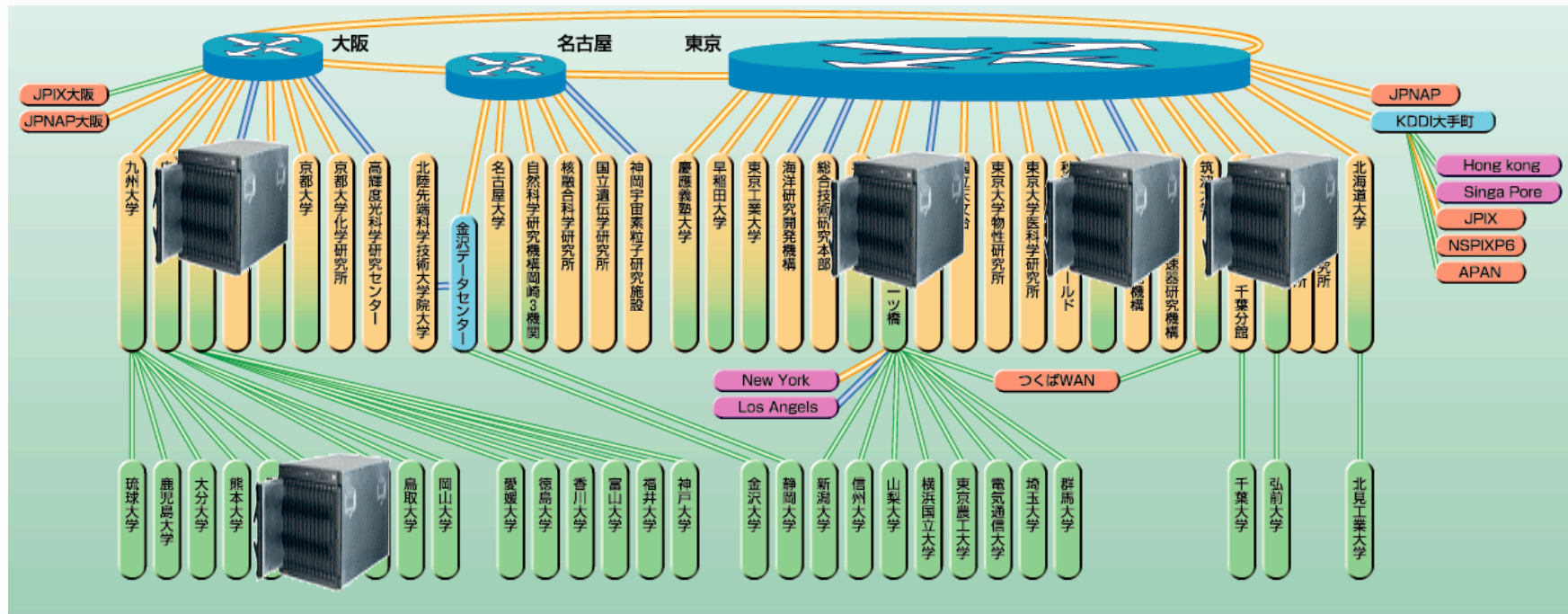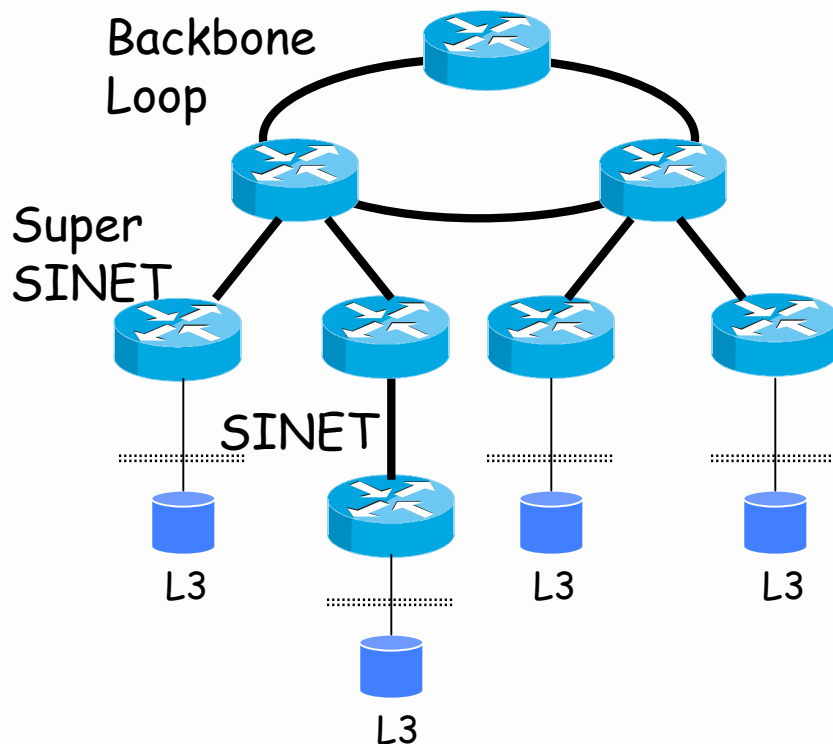- Newer SINET-3 is coming soon

SINET/Super-SINET Topological (1)

SINET/Super-SINET Topological (2)

- Core forms a single loop
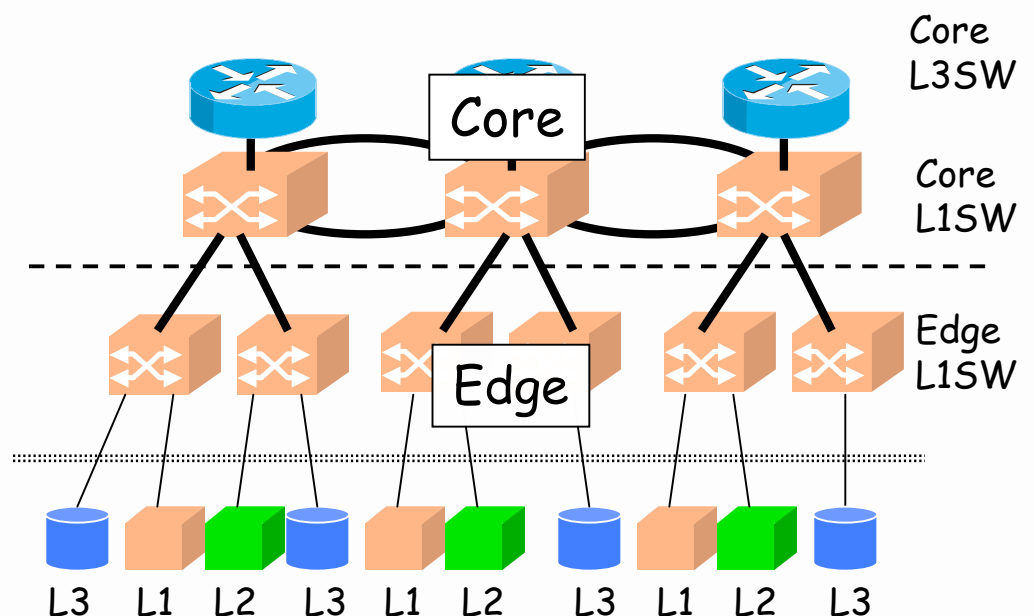- Edge forms a multi-level hierarchy

SINET-3 Structure (planned)

SINET/SuperSINET → SINET-3 (2007)

transition

Backbone Loop

Super SINET

SINET

L3   L3   L3

L3

L3

Core L3SW

Core

Core L1SW

Edge

Edge L1SW

L3   L1   L2   L3   L1   L2   L3   L1   L2   L3

- Similar topological appearance
- Two-level hierarchy (core/edge)
- Reliable/redundant core plane

SINET-3 Geographical (planned)

- Core routers hosted at data centers
- Edge routers located at Super SINET bases
- Nodes can be collocated at data centers

- Edge L1SW
- CoreL1/L3 SW
- 1G-10G
- 10G-40G
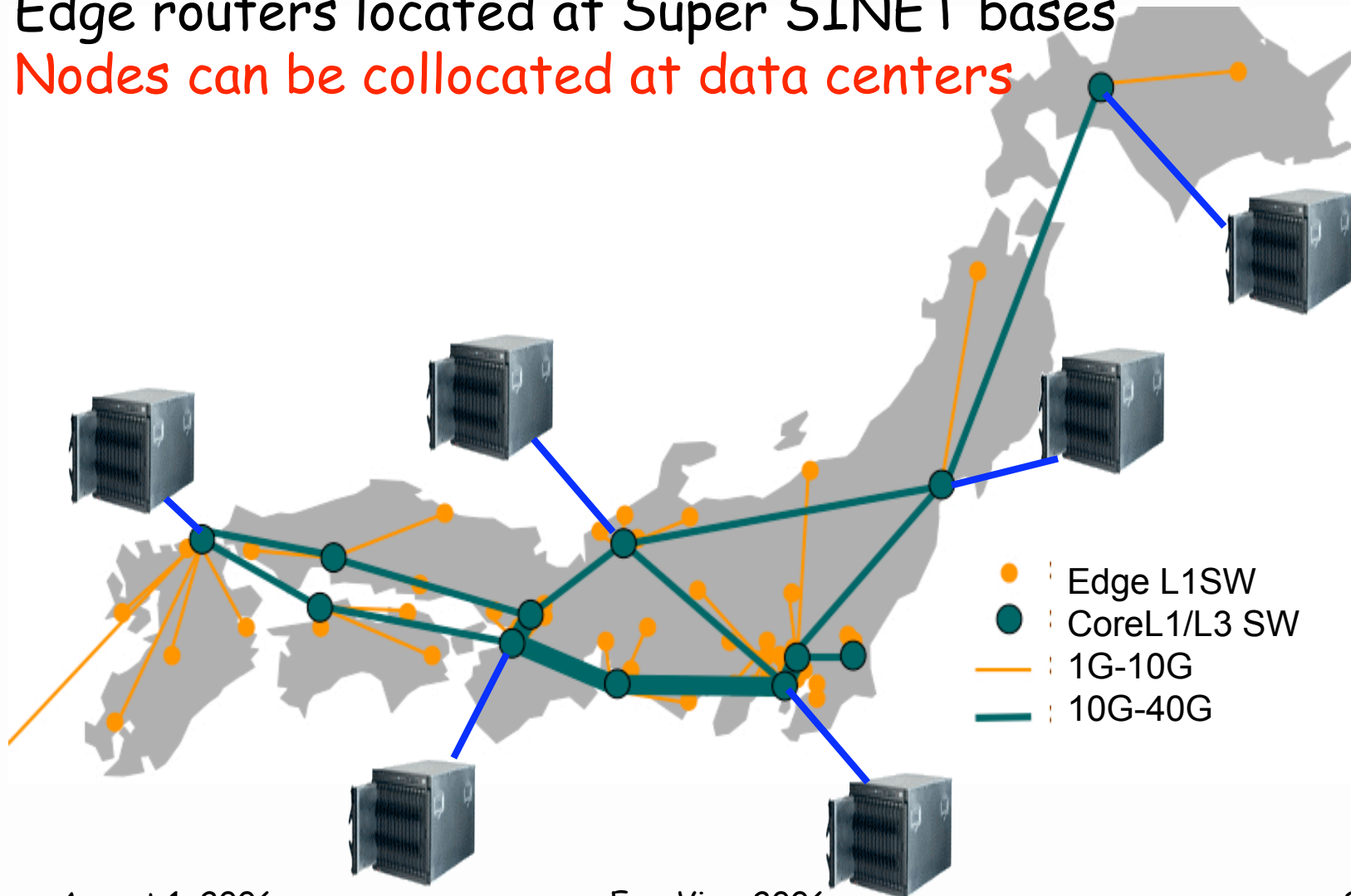
Target Network Services

- Routing Overlays/Virtual ISP
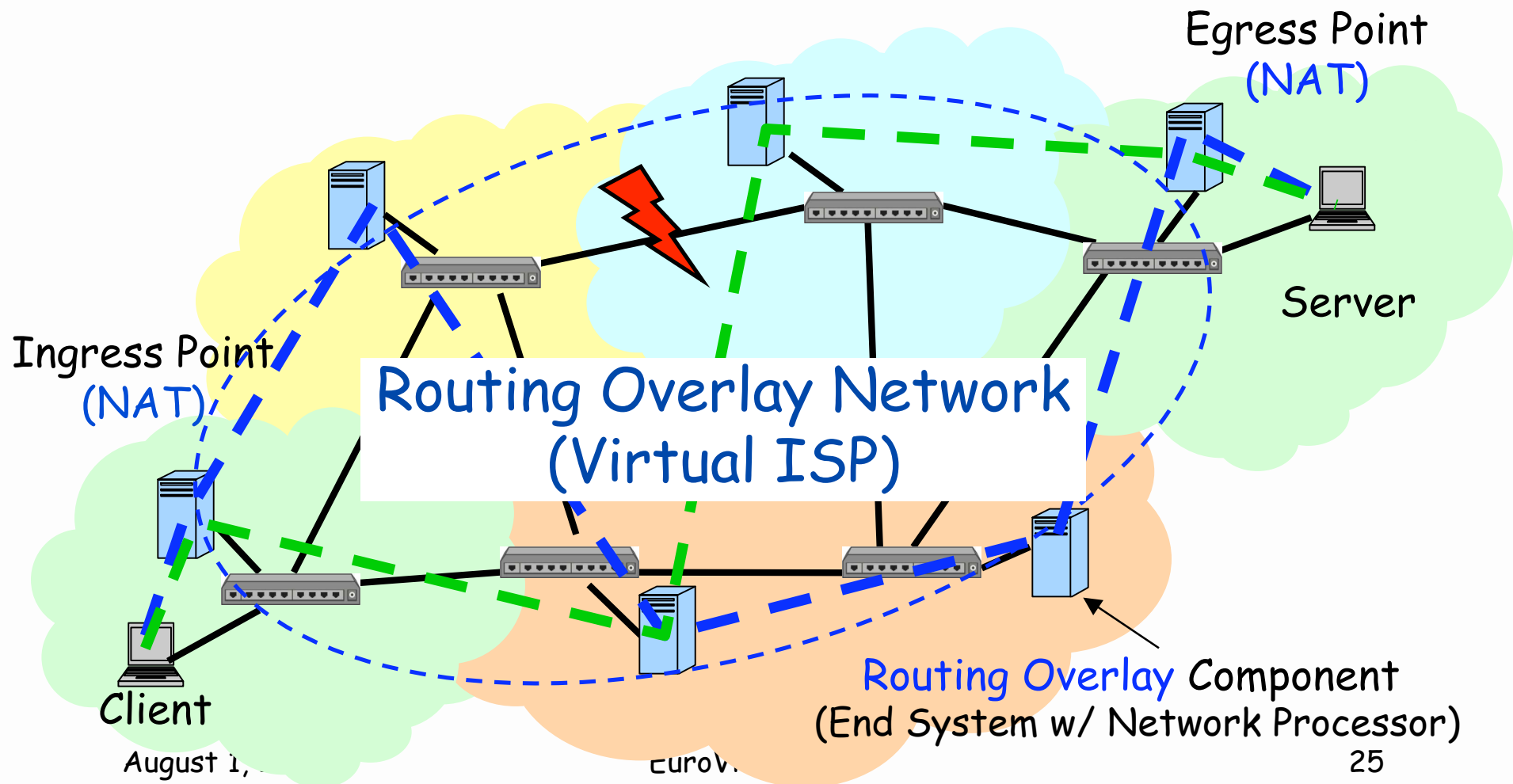 - IIAS/VINI/PLUTO
- Overlay Awareness
  - Underlay, Primitives
- Universal Access
- :

Routing Overlay / Virtual ISP

- Resilience (recover link failures quicker than BGP)
- Performance (optimize application specific metrics)

Egress Point (NAT)

Ingress Point (NAT)

Routing Overlay Network
(Virtual ISP)

Server

Client

Routing Overlay Component
(End System w/ Network Processor)

Production Routing Overlay

- **Overlay Routing** (Detour, RON, PLUTO, IIAS…)
  - **All the goodies**
    - Resilience, Performance, Flexibility, Security
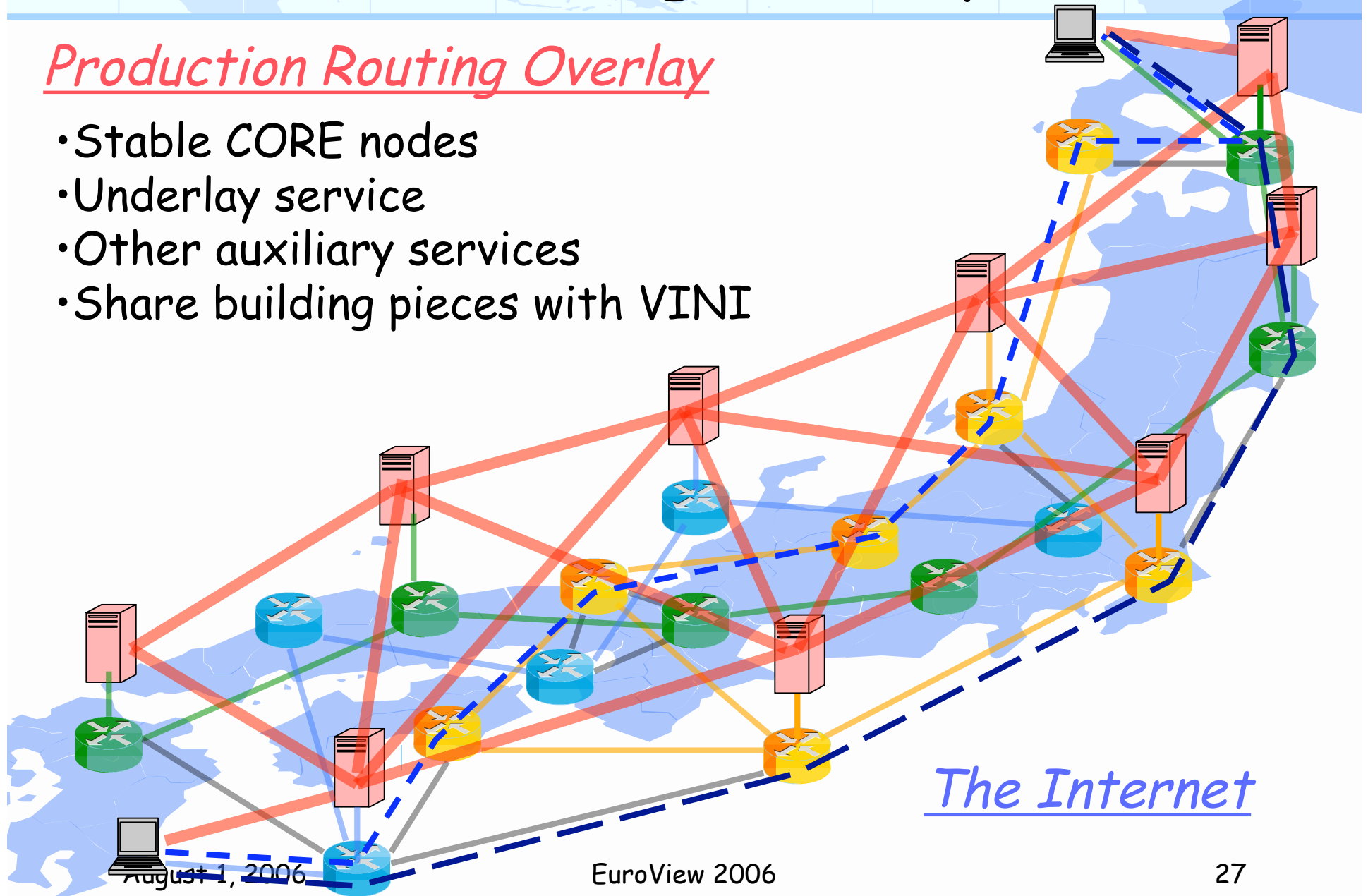 - **Commonly perceived obstacles**
    - Scalability, Addressability, Overhead, Ingress-egress discovery
- **CORE appropriate to explore these**

Production Routing Overlay

Production Routing Overlay

- Stable CORE nodes
- Underlay service
- Other auxiliary services
- Share building pieces with VINI

The Internet

Overlay Aware Network Architecture

- **Overlay would always exist in future**
  - The Internet is an overlay network
  - "Inevitable Ossification" with "Success"
- **Overlay awareness in network architecture**
  - "Underlay" to support overlay networks
    - E.g, exposing information on underlying network
  - What would be a minimal set of primitives?

Timeline

- 2006 PoP Collocation Deployment
  - Public PlanetLab / CORE (Private PlanetLab)
  - Xlice on CORE
  - Experiments with several target net services
- 2007~
  - Further extension of CORE
  - GENI collaboration
  - Wireless subnets/Custom hardware (maybe later…)

Conclusion

- **Experiences learned in PlanetLab**
  - Incremental approach promising
  - More network architecture oriented research using "inside" nodes
- **CORE shares similar goals with PL/GENI**
  - Continuously redesigning the Internet
  - Happy to collaborate with PL/GENI/others
  - Explore our own research space as well

Contact : nakao@iii.u-tokyo.ac.jp
core@planet-lab-jp.org