Architectures for the Future Networks and the Next Generation Internet

Washington University in Saint Louis
Saint Louis, MO 63130
Jain@cse.wustl.edu
http://www.cse.wustl.edu/~jain/

Tutorial at MILCOM 2010, November 2, 2010

Overview

1. Why Next Gen?
2. Internet 3.0
3. Content Centric Networks
4. Challenged Networks
5. Routing Architectures: Open Flow, ID-Locator Split Proposals
6. Green Networking
7. Next Generation Testbeds

Future Internet Projects

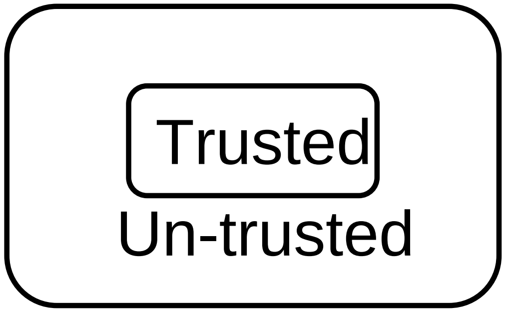
- In 2005 US National Science Foundation started a large research and infrastructure program on next generation Internet
- Q: How would you design Internet today? Clean slate design.
- "Future Internet Design" (FIND): 48+ projects
    - Stanford, MIT, Berkeley, CMU, …
    - "An Architecture for Diversified Internet" at WUSTL
- "Global Environment for Networking Innovations" (GENI): 29+ projects
- European Union: 7th Framework program
- Japan: AKARI (A small light in the dark pointing to the future)
- China, Korea, Australia, …20+ countries

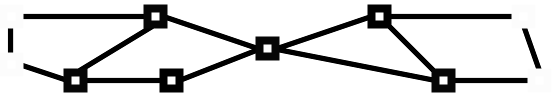
Ref: S. Paul, J. Pan, R. Jain, "Architectures for the Future Networks and the Next Generation Internet: A Survey," Accepted for publication in Computer Communications, July 2010, 72 pp., http://www.cse.wustl.edu/~jain/papers/i3survey.htm

Key Problems with Current Internet

1. Security:
   Fundamental architecture design issue
   Control+Data are intermixed
   Security is just one of the policies.

2. No concept of ownership
   (except at infrastructure level)
   Difficult to represent organizational, administrative hierarchies and relationships. Perimeter based.
   Ÿ Difficult to enforce organizational policies

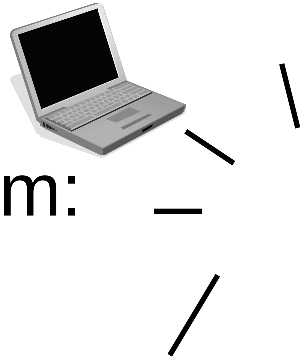




Realms

Problems (cont)

3. Identity and location in one (IP Address)
   Makes mobility complex.
4. Assumes live and awake end-systems
   Does not allow communication while sleeping.
   Many energy conscious systems today sleep.
5. No representation for real end system: the human.



Ref: R. Jain, ``Internet 3.0: Ten Problems with Current Internet Architecture and Solutions for the Next Generation,'' Proceedings of Military Communications Conference (MILCOM 2006), Washington, DC, October 23-25, 2006

Names, IDs, Locators

Name: John Smith

ID: 012-34-5678

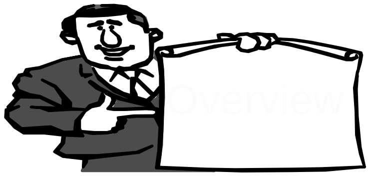
Locator:
1234 Main Street
Big City, MO 12345
USA

- Locator changes as you move, ID and Names remain the same.
- Examples:
  - Names: Company names, DNS names (Microsoft.com)
  - IDs: Cell phone numbers, 800-numbers, Ethernet addresses, Skype ID, VOIP Phone number
  - Locators: Wired phone numbers, IP addresses

Future Internet: Areas of Research

1. New architectures
2. Security
3. Content Delivery Mechanisms
4. Delay Tolerant Networking
5. Management and Control Framework
6. Service Architectures
7. Routing: New paradigms
8. Green Networking
9. Testbeds

Ref: S. Paul, J. Pan, R. Jain, "Architectures for the Future Networks and the Next Generation Internet: A Survey," Accepted for publication in Computer Communications, July 2010, 72 pp., http://www.cse.wustl.edu/~jain/papers/i3survey.htm



2. Internet 3.0

‰Internet 3.0: Next Generation Internet

‰Internet Generations

‰Organizational Representation

‰User- Host- and Data Centric Models

‰Policy-Based Networking Architecture

‰Multi-Tier Object-Oriented View

‰Virtualization

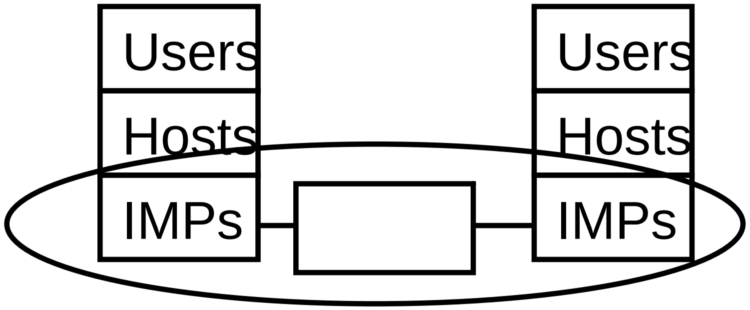
Internet 3.0: Next Generation Internet

- Internet 3.0 is the name of the Washington University project on the next generation Internet
- Goal 1: Represent the commercial reality of distributed Internet ownership and organization
- Goal 2: Develop a clean slate architecture to overcome limitations of the current internet
- Goal 3: Develop an incremental approach to implement the architecture

Internet Generations

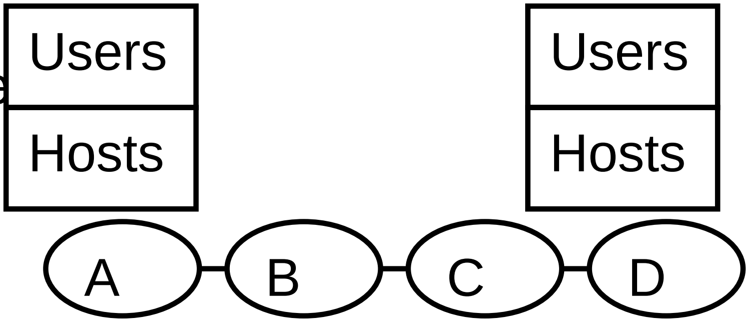
‰ Internet 1.0 (1969 – 1989)

- ¾ Single ownership Ÿ Trust
- ¾ complete knowledge
- ¾ Algorithmic optimality Ÿ RIP



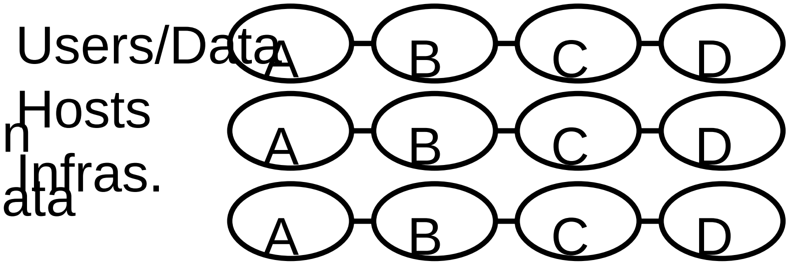
‰ Internet 2.0(1989–2009) Commerce

- ¾ Multiple ownership of infrastructure
  Ÿ Distrust,Security
- ¾ No knowledge of internal topology and resources
- ¾ Policy based routing Ÿ BGP

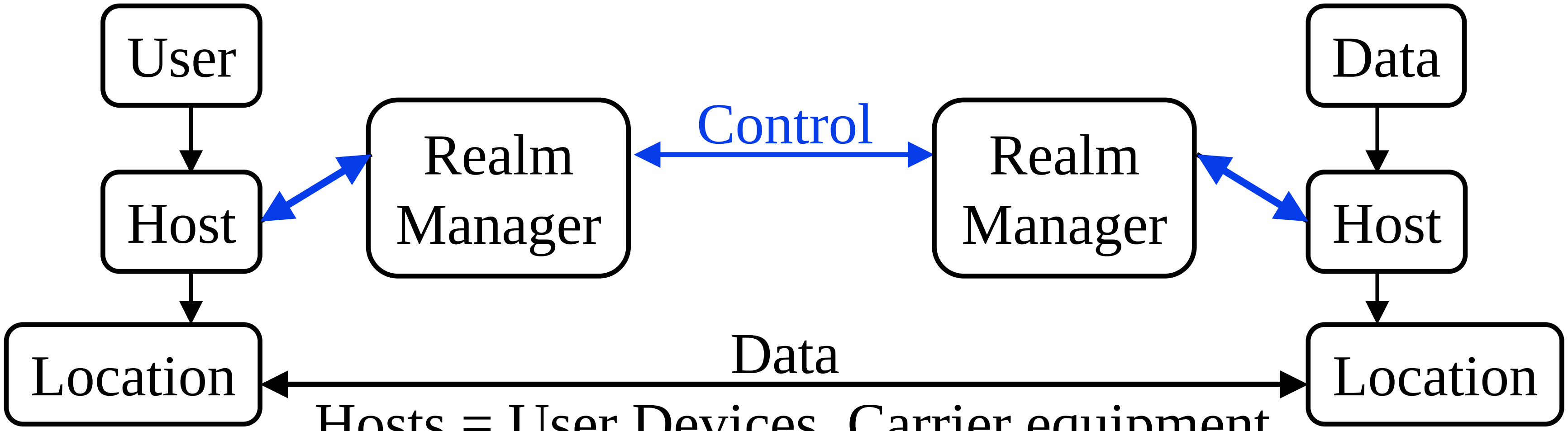


‰ Internet 3.0(2009–2029) Commerce

- ¾ Users, Content, Host ownership
- ¾ Requirements, Service Negotiation
- ¾ Mobility of users and distributed data



Policy Based Networking Architecture



Hosts = User Devices, Carrier equipment

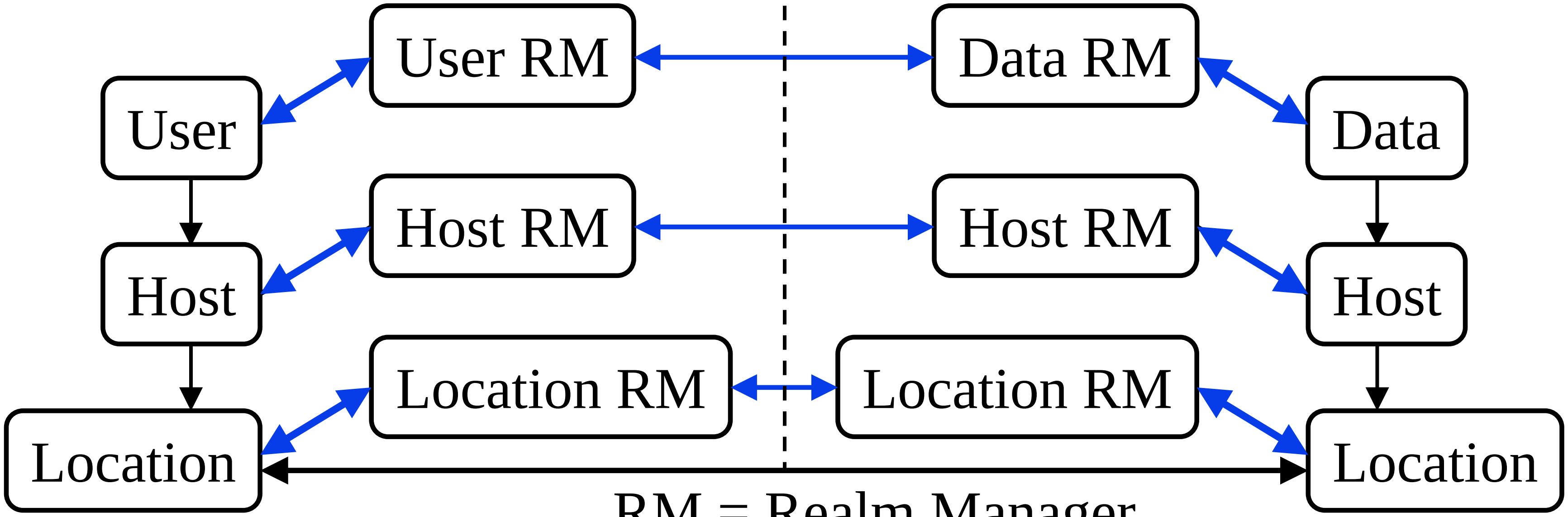
Realm managers (RM): Many organizational functions

- Allow **mobility**, multi-homing, location privacy
- **ID-Locator**: Resolve current location for a given host-ID
- Policy Monitoring. Conformance to Contracts. Troubleshooting.
- Enforce policies related to authentication, authorization, privacy
- Proxy services enabling hosts to sleep ⇒ Energy-aware networking

2. Intelligence in the network ⇒ Suitable for the masses

- Ref: http://www.cse.wustl.edu/~jain/papers/milsa.htm

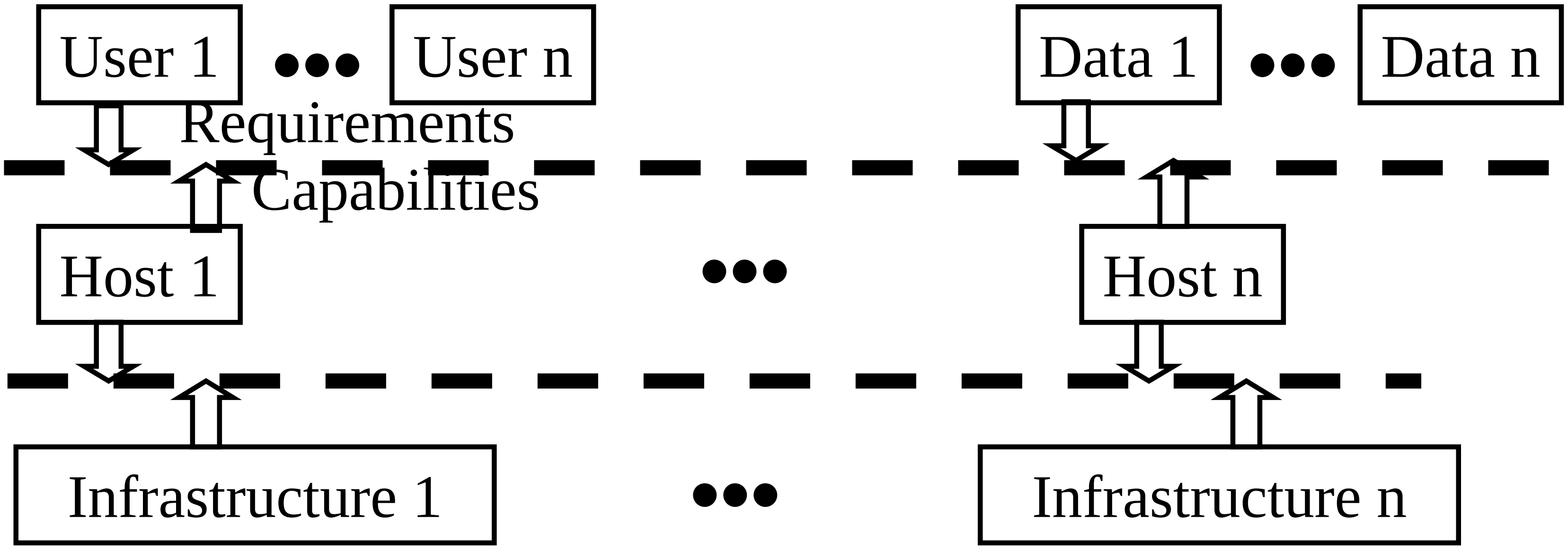
3-Tier Model of Internet 3.0



RM = Realm Manager

- Both Users and data need hosts for communication
- Data is easily replicable. All copies are equally good.
- Users, Hosts, Infrastructure, Data belong to different realms (organizations).
- Each object has to follow its organizational policies.
- Ref: http://www.cse.wustl.edu/~jain/papers/bcs08.htm

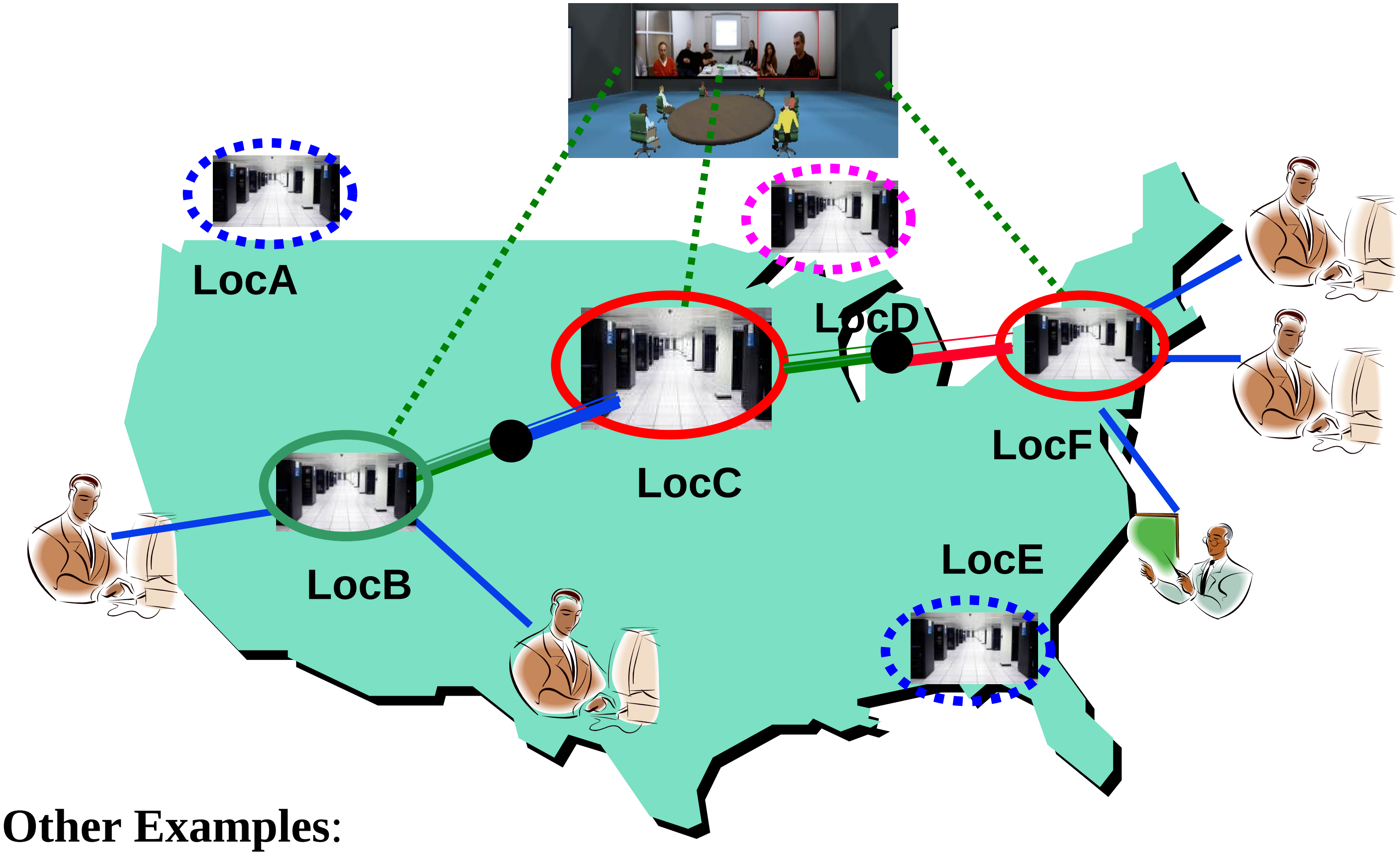
Multi-Tier Object-Oriented View



- Objects have capabilities and provide services. Higher tiers specify the requirements
- Tier service broker (shown by dotted line) composes a service – can negotiate with multiple realms in that tier
- Higher tier may not/need not find details of lower tiers

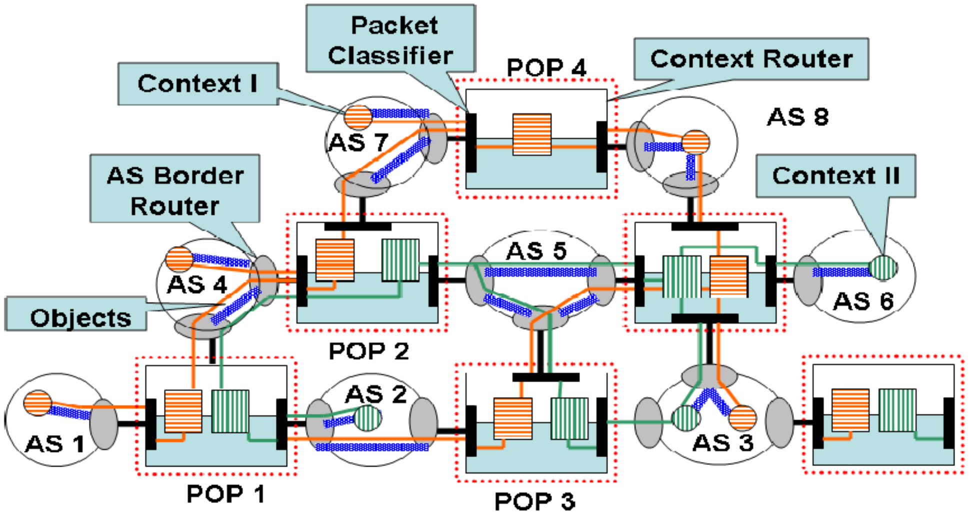
Allows creating requirement specific networking context

Cloud Computing: Network as a Service



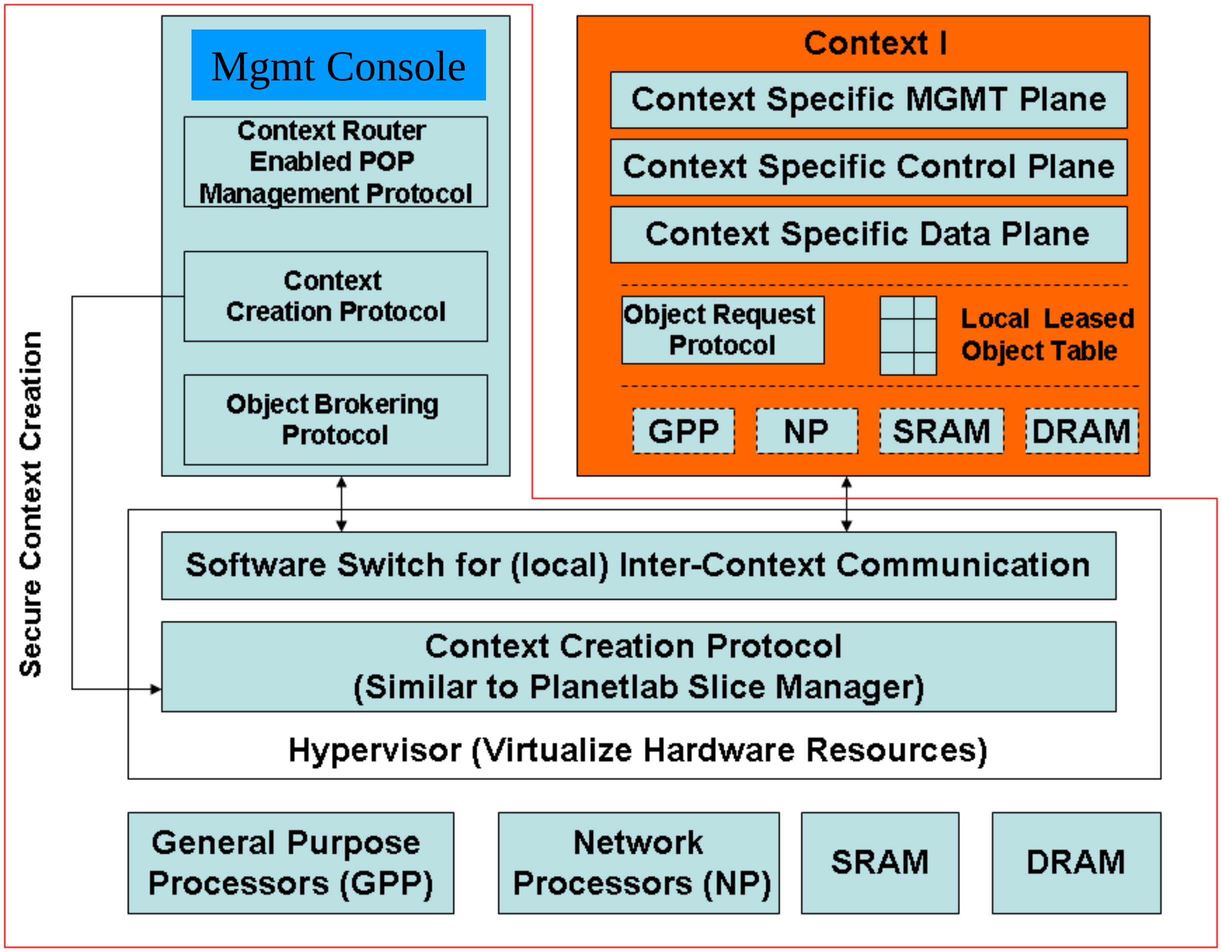
- **Other Examples**:
  - National Security: Infrastructure vs. national boundaries

Internet 3.0 POP/IXP Design



- ❑ Points of Presence (PoPs) at AS boundaries connect to other service providers. Internet Exchange Points (IXPs)
- ❑ Internet 3.0 POP's/IXPs enhanced with Context Routers
  - ➢ "Red" and "Green" Application Contexts
  - ➢ Overlays Programmable at the Infrastructure Tier
  - ➢ AS (Realms) advertize "objects" (or services)
  - ➢ Overlays setup Application Specific Topology

Context Router Design



Secure Context Creation

Mgmt Console

Context Router Enabled POP Management Protocol

Context Creation Protocol

Object Brokering Protocol

Context I

Context Specific MGMT Plane

Context Specific Control Plane

Context Specific Data Plane

Object Request Protocol

Local Leased Object Table

GPP

NP

SRAM

DRAM

Software Switch for (local) Inter-Context Communication

Context Creation Protocol
(Similar to Planetlab Slice Manager)

Hypervisor (Virtualize Hardware Resources)

General Purpose Processors (GPP)

Network Processors (NP)

SRAM

DRAM

Virtualization

- Allows multiple overlays on a single substrate
- Allows nodes to treat an overlay as a native network
- Provides isolation ⇒ multiple architectures, Partitioned Control
- Allow testing diverse routing protocols and service paradigms
- Better architectures will attract more users and become main line
- Allows diversified services while utilizing economies of scale in the substrate components
- Virtualization over IP networks
  ⇒ Not suitable for experiments at lower layers

Ref: T. Anderson, L. Peterson, S. Shenker, J. Turner, ``Overcoming the Internet Impasse through Virtualization," Computer, Volume 38, Issue 4, pp 34-41, April 2005.

Internet 1.0 vs. Internet 3.0: Design

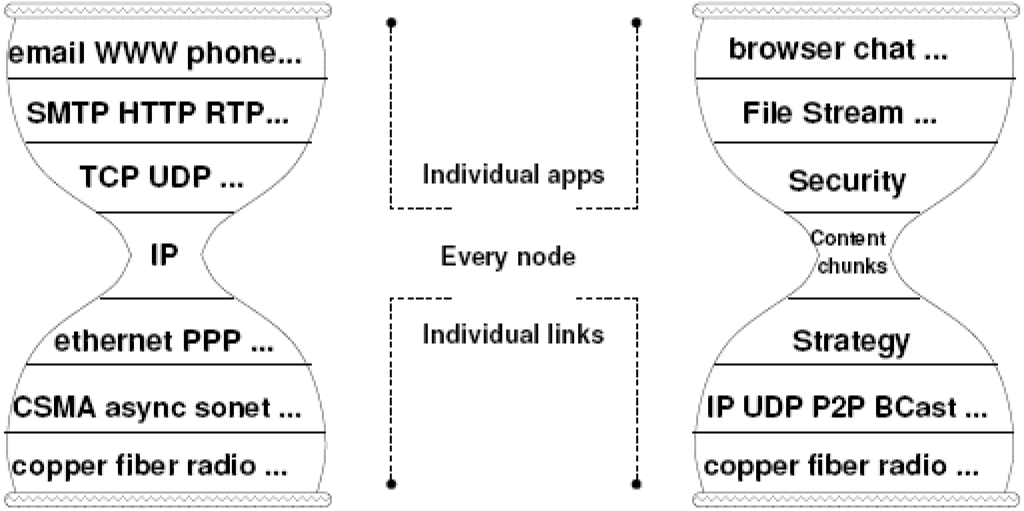
	Design Issue	**Internet 1.0 Solution**	**Internet 3.0 Solution**
1	Resource allocation	Algorithmic Optimization	Policy based
2	Intelligence	Manual/applications	In the network
3	Connections	Host-Host	User-Data (Hosts are intermediate systems)
4.	Ownership	Single=> Single Tier	Commercial Reality => Multi-Tier
5	Information	Complete knowledge of all tiers	Only service API's are disclosed
6	Mobility	Host mobility	Multi-tier mobility (User/data/host)
7	Multi-homing	Host multihoming	Multi-tier multihoming (User/Data/Host)
8	Virtualization	Network virtualization	Multi-Tier virtualization

3. Content Centric Networks

- Content-Centric Networks (CCN)
- CCN Packets
- CCN Capable Routers Operation
- CCN Security

Content-Centric Networks

- IP cares about “**Where**”: forward packets from A to B
- Users care about “**What**”: Movie X
- Replace “packets” with “Data Objects” or “Interests” (requests)
- Replace “Addresses” with “Names of Objects”



email WWW phone...
SMTP HTTP RTP...
TCP UDP ...
IP
ethernet PPP ...
CSMA async sonet ...
copper fiber radio ...

Individual apps
Every node
Individual links

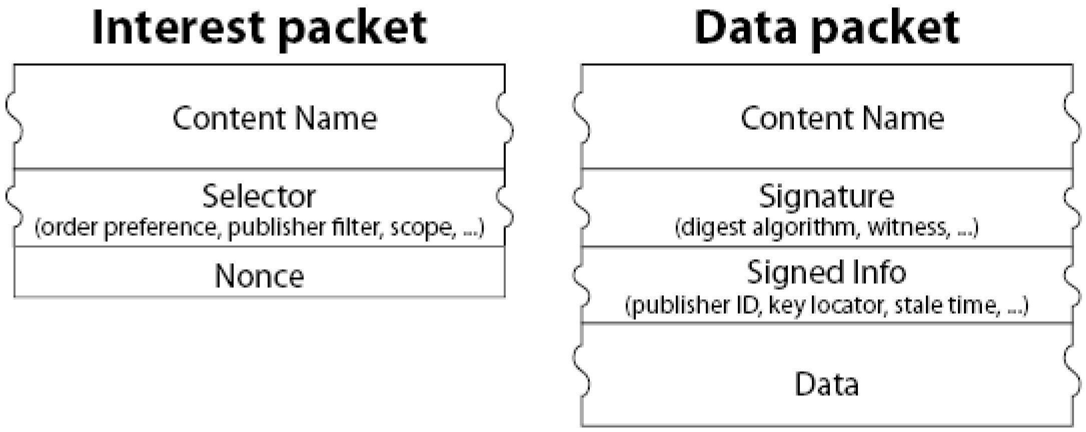
browser chat ...
File Stream ...
Security
Content chunks
Strategy
IP UDP P2P BCast ...
copper fiber radio ...

[JAC09]

Ref: [JAC09] V. Jacobson, et al, “Networking Named Content,” to appear in CoNEXT 2009, December 2009

CCN Packets

- Interest Packets: Request for Data
- Data Packets: Signed Data
- Longest prefix match is used as in IP addresses
  http://www.cse.wustl.edu/~jain/talks/ftp/in3_video matches
  http://www.cse.wustl.edu/~jain/talks/ftp/in3_video/V00/S00



Interest packet

Content Name

Selector
(order preference, publisher filter, scope, ...)

Nonce

Data packet

Content Name

Signature
(digest algorithm, witness, ...)

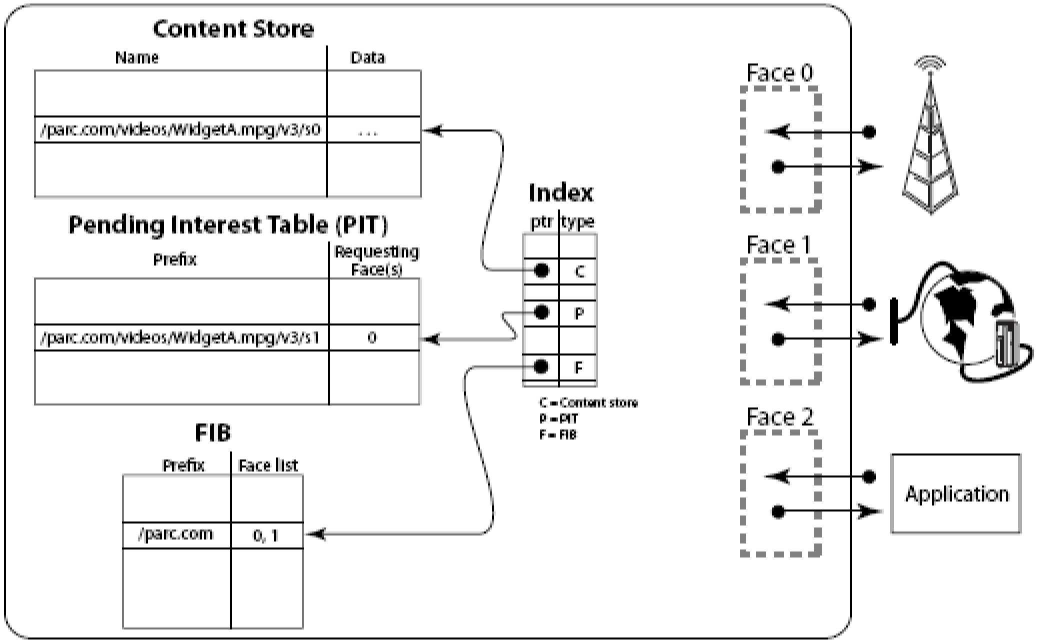
Signed Info
(publisher ID, key locator, stale time, ...)

Data

[JAC09]

CCN Capable Routers Operation

- **Content Store**: Local cache of data
- **Pending Interest Table (PIT):** Recent requests forwarded
- **Forwarding Information Base (FIB):** Known data locations
- **Faces**: Requesting processes and hardware interfaces



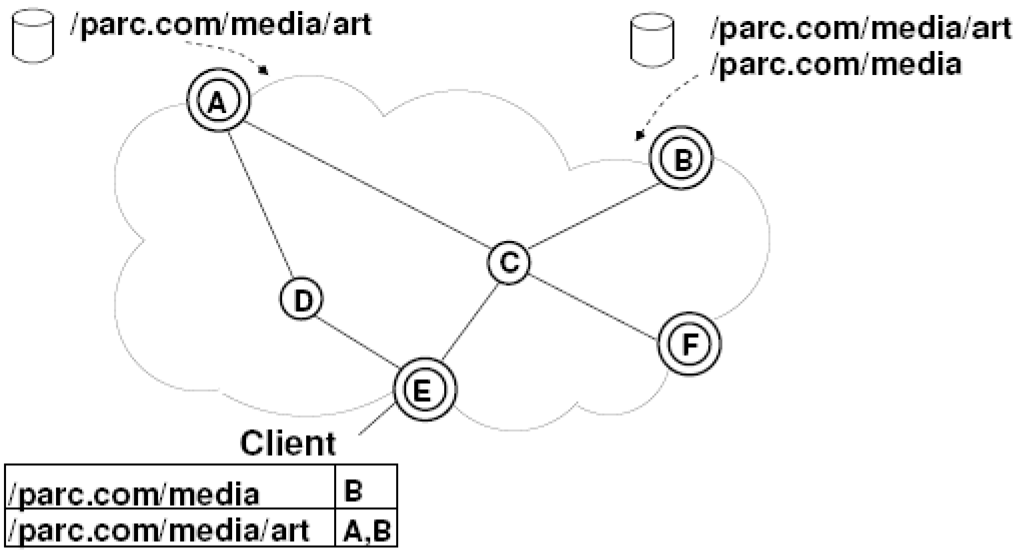
[JAC09]

Routers Operation (Cont)

- ❑ Applications send “Interest” in data X
- ❑ Router looks up in local store and sends if found
- ❑ Router looks up in PIT, if entry already exists (someone requested it recently) , adds the interest, face to the same entry
- ❑ Router looks up in FIB, if entry exists (data location is known), a PIT entry is made and the interest is multicasted to all faces in the FIB entry and the FIB entry is removed
- ❑ If there is no FIB entry, interest is discarded (router does not know how to get the data)
- ❑ When data arrives, Content Store match ⇒ duplicate, discard
  PIT match ⇒ Forward to all faces
  FIB match ⇒ No PIT ⇒ Unsolicited ⇒ Discard
- ❑ Data providers register their data ⇒ Creates FIB entries

CCN Security

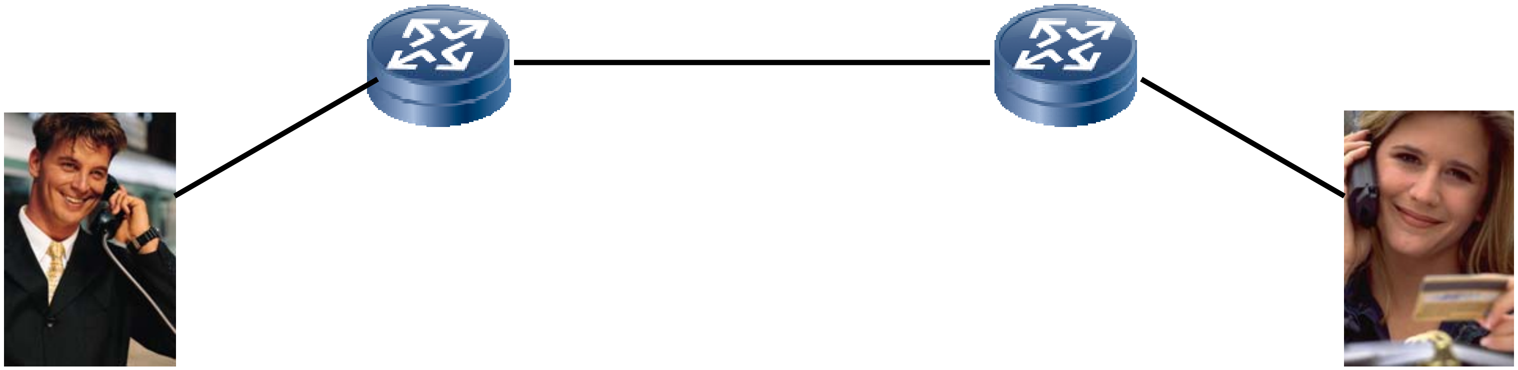
- ❑ Data-Centric Security ⇒ Protections travel with the data
- ❑ All data is signed
- ❑ Data can be replicated or moved
- ❑ All data is versioned and is immutable once in the system
- ❑ IP and CCN routers can coexist. Public domain code available.



/parc.com/media	B
/parc.com/media/art	A,B

VOIP over CCN

- On-demand publishing: Data is produced only when some wants to connect
- Callee’s phone registers a service
- Caller looks for the service



Ref: V. Jacobson, et al, “VoCCN: Voice over Content-Centric Networks,” ACM ReArch 2009, Rome, Italy.

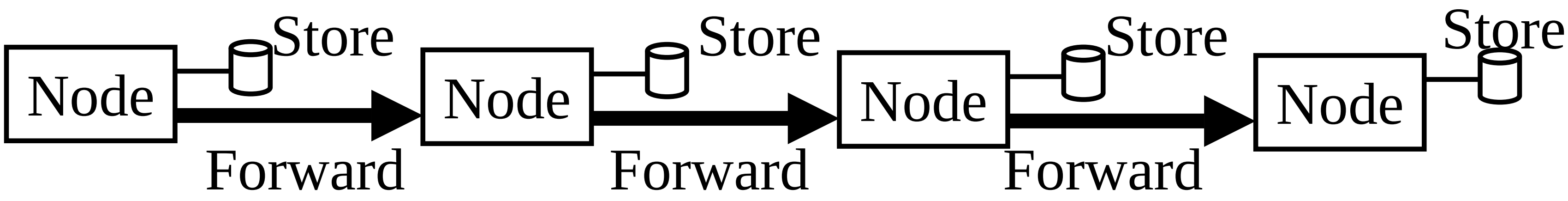
4. Challenged Networks

Challenged ⇒ End-to-end not continuous
E.g., DTN, Sensor, Ad-hoc, …

- Delay/Disruption Tolerant Net (DTN)
  - Bundle Protocol
  - Bundle Delivery Options
  - Known Issues with Bundle Protocol
- Other Challenged environments

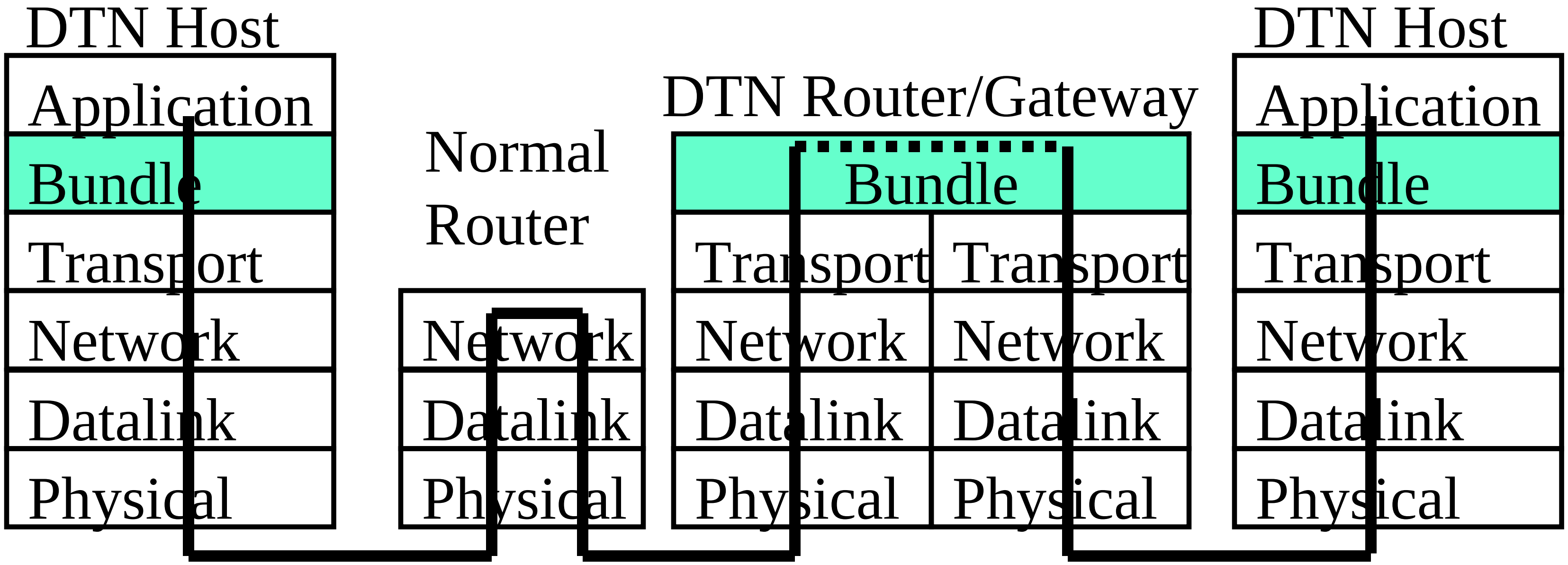
Delay/Disruption Tolerant Net (DTN)

- Intermittent Connectivity
- Long or variable delay
- Asymmetric data rates
- High Error Rates



Bundle Protocol

- End-to-end protocol for DTNs
- Runs between application and transport
- Ties together heterogeneous regions
- Bundles can be fragmented or reassembled on the way



Bundle Protocol (Cont)

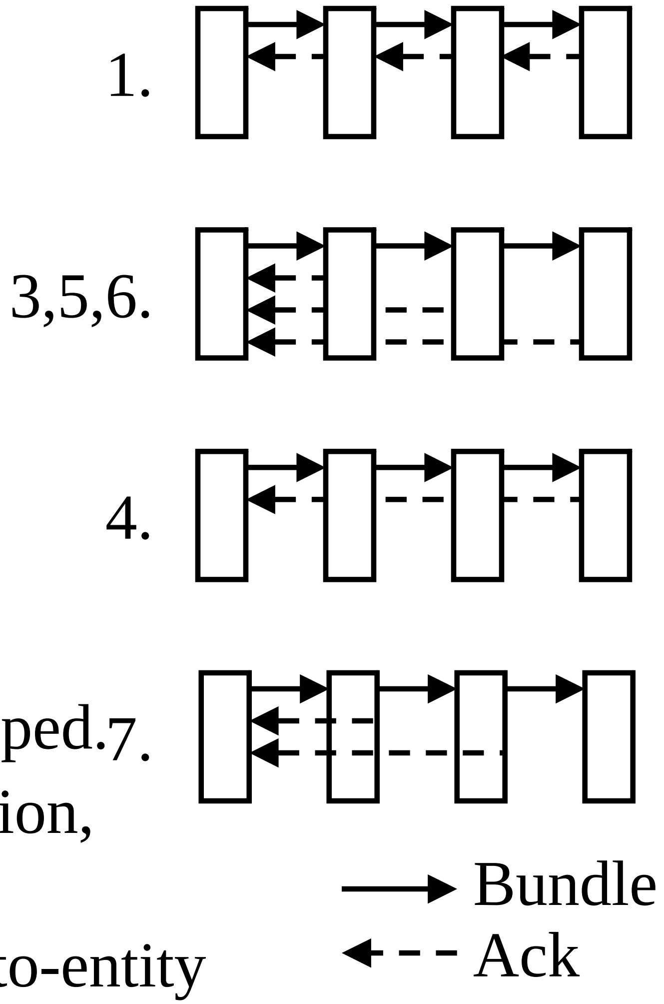
- Three Types of DTN Nodes: Hosts, Routers, Gateways (Convergence layer adapter)
- IPv6-like extension headers are used in bundle
- Custody transfer of bundles
- A node releases storage when someone else accepts custody
- End-to-End reliability can be achieved by requesting return receipt

Ref: [RFC4838, RFC 5050]

Bundle Delivery Options

1. Custody transfer requested
2. Source custody acceptance required
3. Report on bundle delivery
4. Ack by application requested
5. Report on bundle reception
6. Report on custody acceptance
7. Report on bundle forwarding
8. Reporting on bundle deletion

- Priority of Delivery: Bulk, normal, exped.
- Security: Confidentiality, Authentication, error detection
- Notifications can be to another reply-to-entity



Known Issues with Bundle Protocol

Specification in Development. Not yet defined:

- Operations for forming bundle protocol addresses
- Means for key exchange and security associations
- QoS mechanisms
- How to bind IDs to addresses
- Multiple different convergence layers
  ⇒ Multiple naming schemes ⇒ How to map?
- How to exchange routing information among regions?

Ref: L. Wood, W. Eddy, P. Holliday, ``A Bundle of Problems," IEEE Aerospace conference, Big Sky, Montana, March 2009.

Other Challenged Environments

- **Delay/Fault tolerant Mobile Sensor Networks** (DFT-MSN) [1]: Adjust sleep and multicast for power conservation.
- **Postcards from the edge** [2]: Post office nodes cache data for mobile nodes.
- **Disaster day after networks** (DANs) [3]: Prioritized data transmission over network islands
- **Selectively Connected Networks** (SCNs) [4]: Proxy wakes up nodes only when necessary.

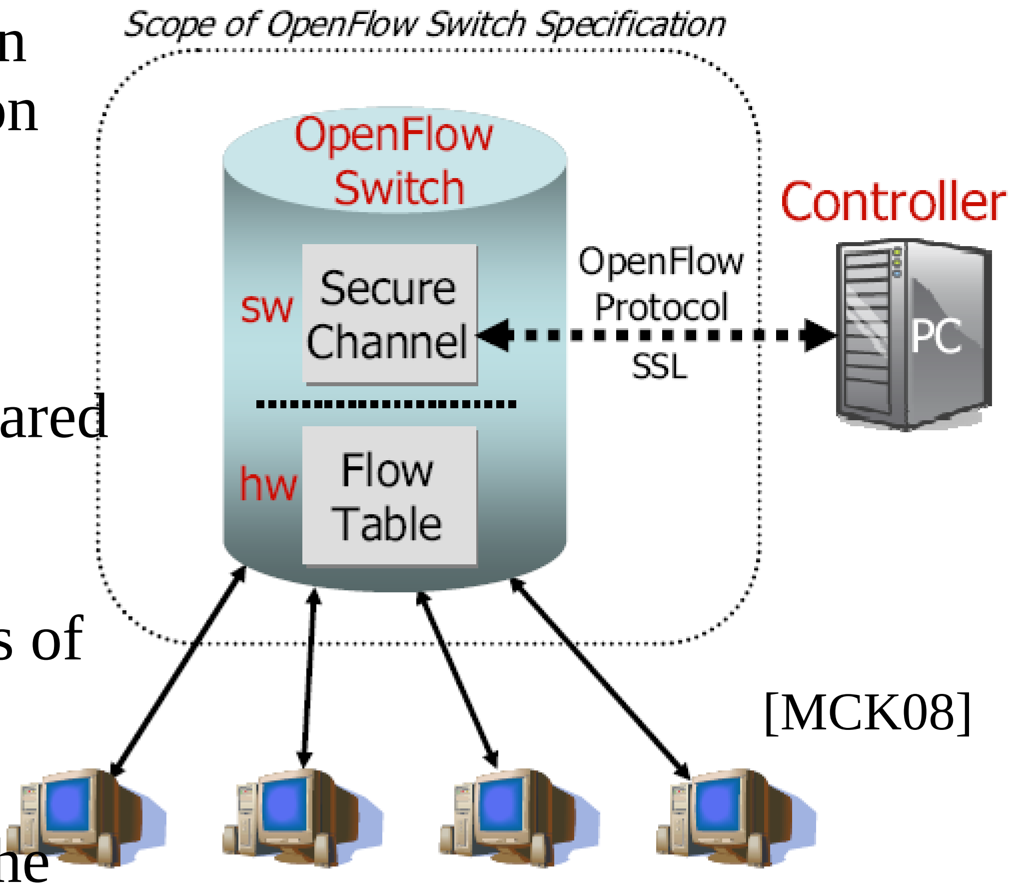
Ref: [1] Y. Wang and H. Wu, ``Delay/Fault-Tolerant Mobile Sensor Network (DFT-MSN): A New Paradigm for Pervasive Information Gathering," IEEE Transactions on Mobile Computing, Vol 6, No 9, pp 1021-1034, 2007.
[2] R. Yates, D. Raychaudhuri, S. Paul, et al., ``Postcards from the Edge: A Cache-and-Forward Architecture for the Future Internet," NSF NeTS FIND Initiative, http://www.nets-find.net/Funded/Postcards.php
[3] H. Luo, R. Kravets, T. Abdelzaher, ``The-Day-After Networks: A First-Response Edge-Network Architecture for Disaster Relief," NSF NeTS FIND Initiative, http://www.nets-find.net/Funded/DayAfterNet.php
[4] M. Allman, K Christenson, B. Nordman, V. Paxson, ``Enabling an Energy-Efficient Future Internet Through Selectively Connected End Systems," HotNets-VI, 2007.

5. Routing Architectures

- ❑ OpenFlow
- ❑ ID-Locator Split
    - ➢ Host Identity Protocol: HIP
    - ➢ ID Locator Split via Core-Edge Separation
    - ➢ LISP Protocol Details
    - ➢ MILSA

OpenFlow

- ❑ Allows researchers to run experimental protocols on production networks
- ❑ Each router/switch has a flow forwarding table
- ❑ Forwarding table is prepared by a central controller
- ❑ Vendors do not need to expose internal workings of their switches
- ❑ No need to program switches. Just program the central controller.

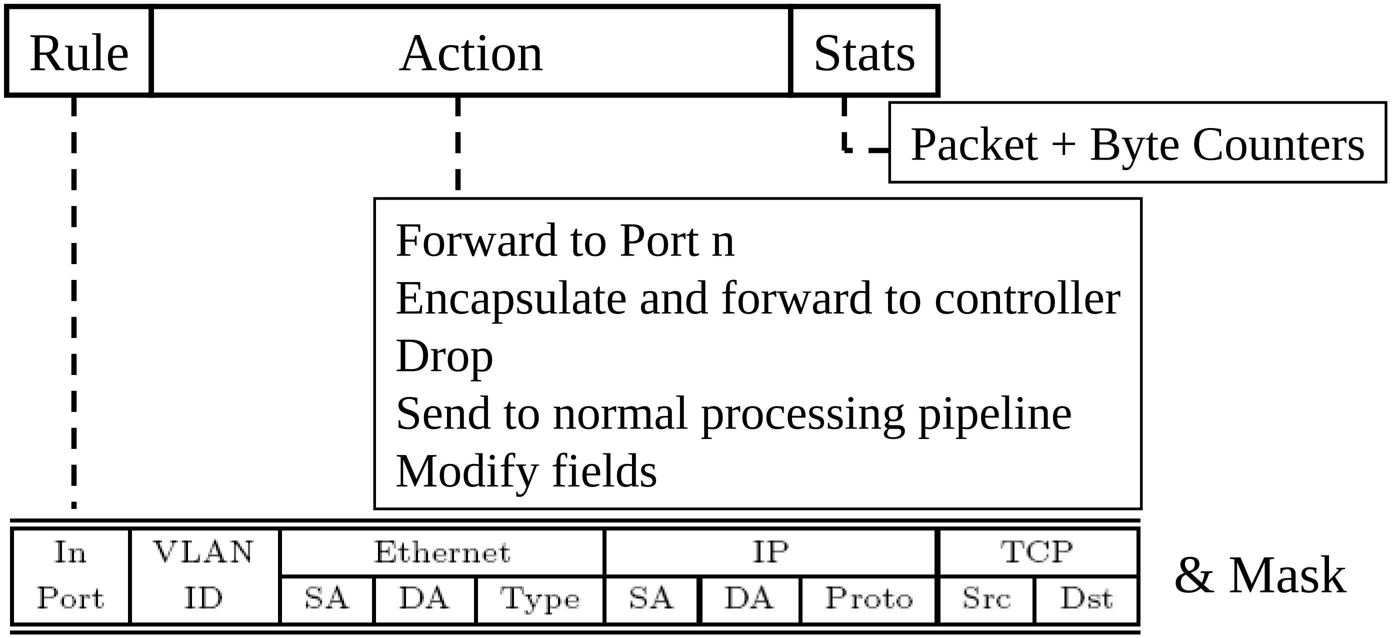


[MCK08]

Ref: [MCK08] ``OpenFlow: Enabling Innovation in Campus Networks," OpenFlow Whitepaper, March 2008

OpenFlow (Cont)

- Three Components:
  - Flow table: How to identify and process a flow
  - Secure Channel: Between controller and the switch
  - Open Flow Protocol: Standard way for a controller to communicate with a switch



Packet + Byte Counters

Forward to Port n
Encapsulate and forward to controller
Drop
Send to normal processing pipeline
Modify fields

In Port	VLAN ID	Ethernet			IP			TCP	
		SA	DA	Type	SA	DA	Proto	Src	Dst

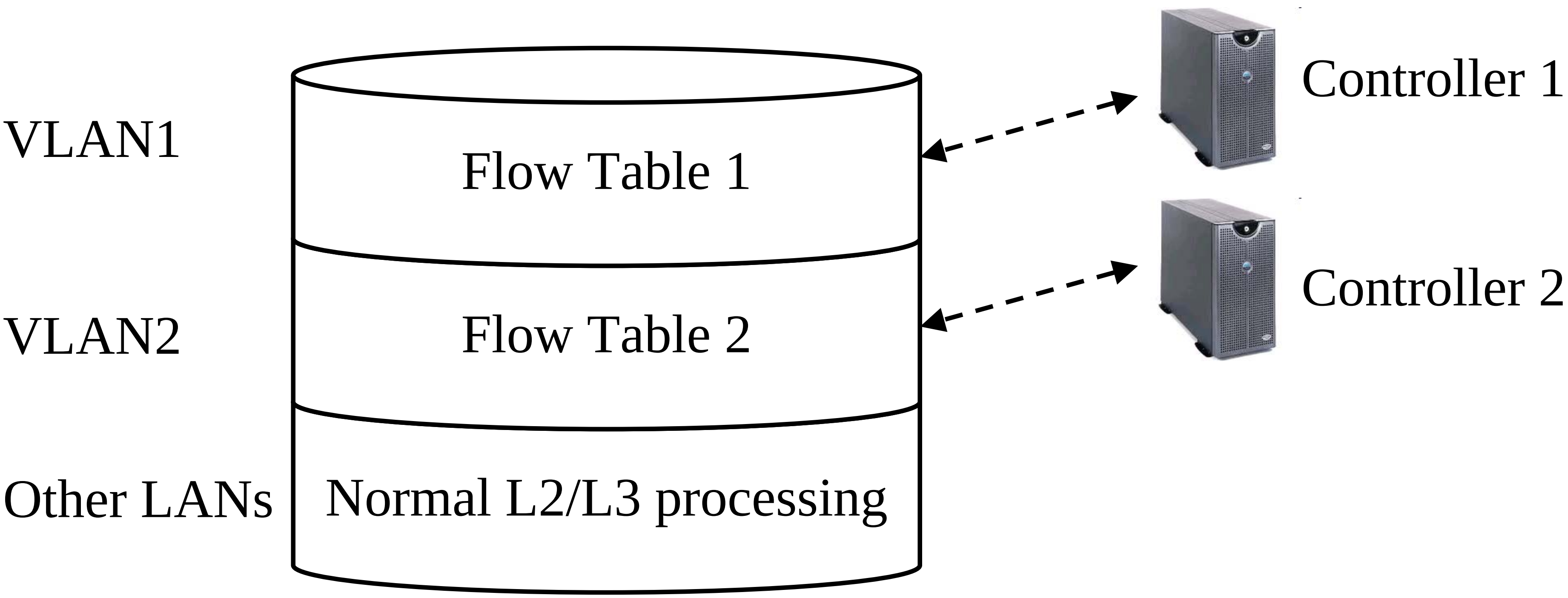
& Mask

OpenFlow (Cont)

- TCAMs are used to match the fields
- Controller forwards the packets correctly as the mobile clients move
- Can handle non-IP networks
- OpenFlow Consortium is developing OpenFlow Switch Specification.
- Reference designs for Linux, Access points (OpenWRT), and NetFPGA (hardware)
- Combined packet and circuit switching
- Software defined networks

Ref: [MCK08], OpenFlowSwitch.org

Software Defined Networks



VLAN1

Flow Table 1

VLAN2

Flow Table 2

Other LANs

Normal L2/L3 processing

Controller 1

Controller 2

Reactive and Proactive Operation

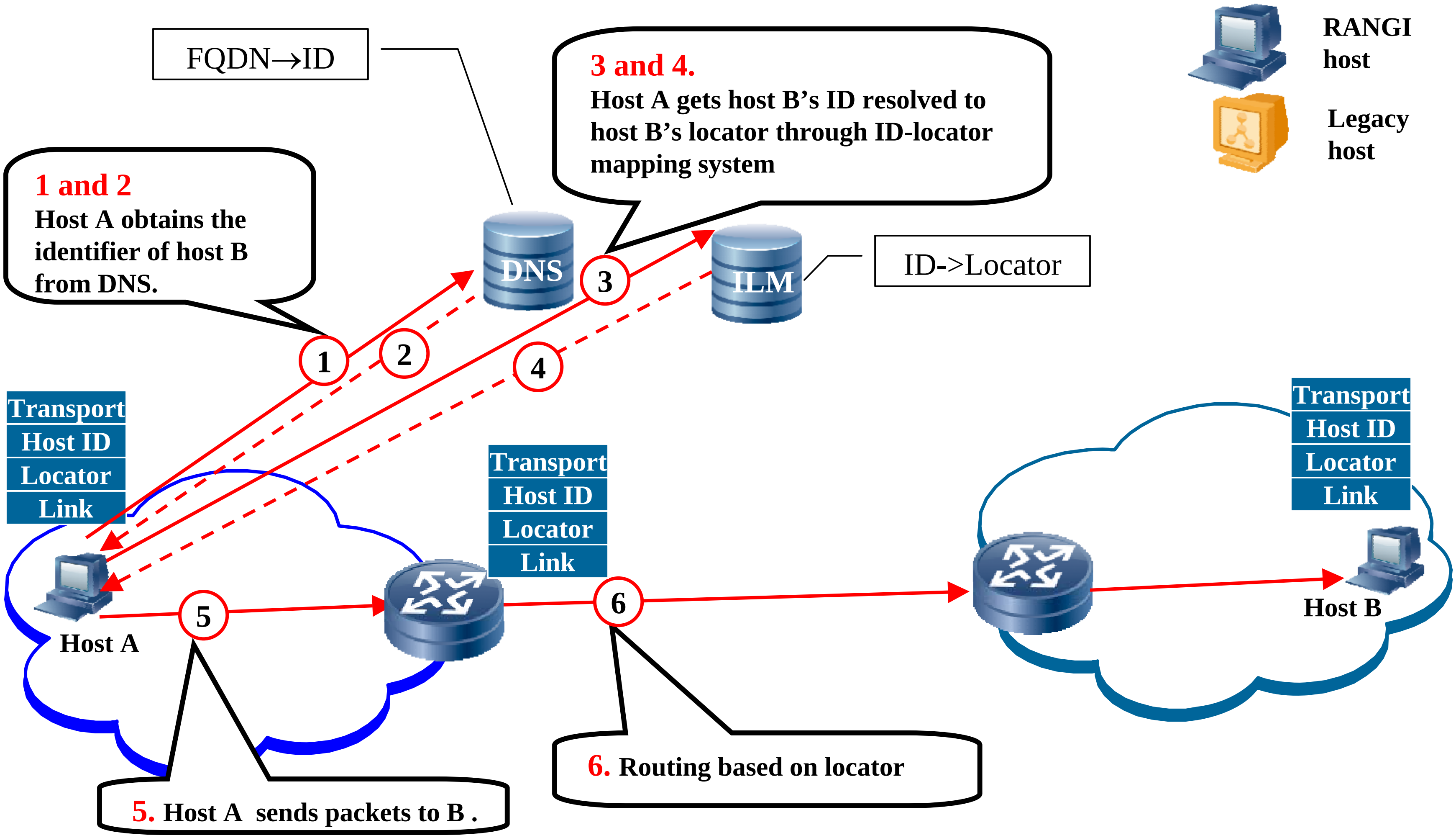
Proactive	Reactive
Switch flow tables pre-populated by the controller	First packet of the flow triggers new flow entries
No flow setup time	Flow setup time
Loss of control connection does not affect operation	Limited operation if control connection lost
Many entries never triggered	Efficient use of flow table entries

OpenFlow allows both models

Current Limitations of OpenFlow

- ❑ Use all switch features (vary with products)
- ❑ New packet formats (non-IP, non-Ethernet, …)
- ❑ Non-flow based applications
- ❑ Setup delay on new flows

ID-Locator Split



1 and 2
Host A obtains the identifier of host B from DNS.

3 and 4.
Host A gets host B's ID resolved to host B's locator through ID-locator mapping system

5. Host A sends packets to B .

6. Routing based on locator

ID-Locator Split (Cont)

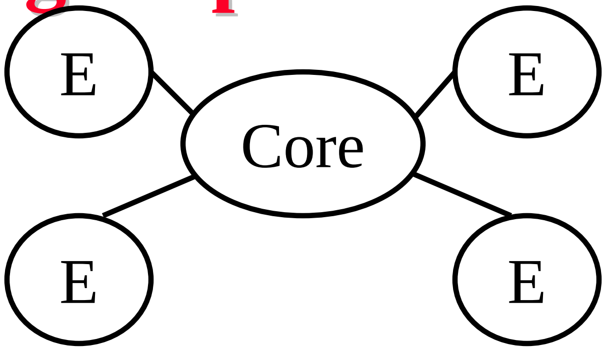
- Allows hosts to move
- Allows entire organizations to move
  Allows organizations to change providers
- No need to use "Provider Independent (PI)" addresses
- Provider Aggregatable (PA) addresses are preferred since they result in shorter BGP tables
  ⇒ Scalable
- Several proposals for host-based ID-locator split:
  HIP, Shim6, I3, and HI3
- All hosts have ID and global locators
- Allow mobility, multihoming, renumbering

HIP

- Host Identity Protocol
- 128-bit Host ID tag (HIT)
- TCP is bound to HIT. HIT is bound to IP address in the kernel
- Uses flat cryptographic based identifier
- Two Methods:
  - Locator registered using Update packets to DNS
    ⇒ Does not allow fast mobility
  - Use rendezvous servers
    ⇒ Does not adhere to organizational boundary
- Requires changes to end hosts

Ref: R. Moskowitz, P. Nikander and P. Jokela, ``Host Identity Protocol (HIP) Architecture," IETF RFC4423, May 2006.

ID Locator Split via Core-Edge Separation

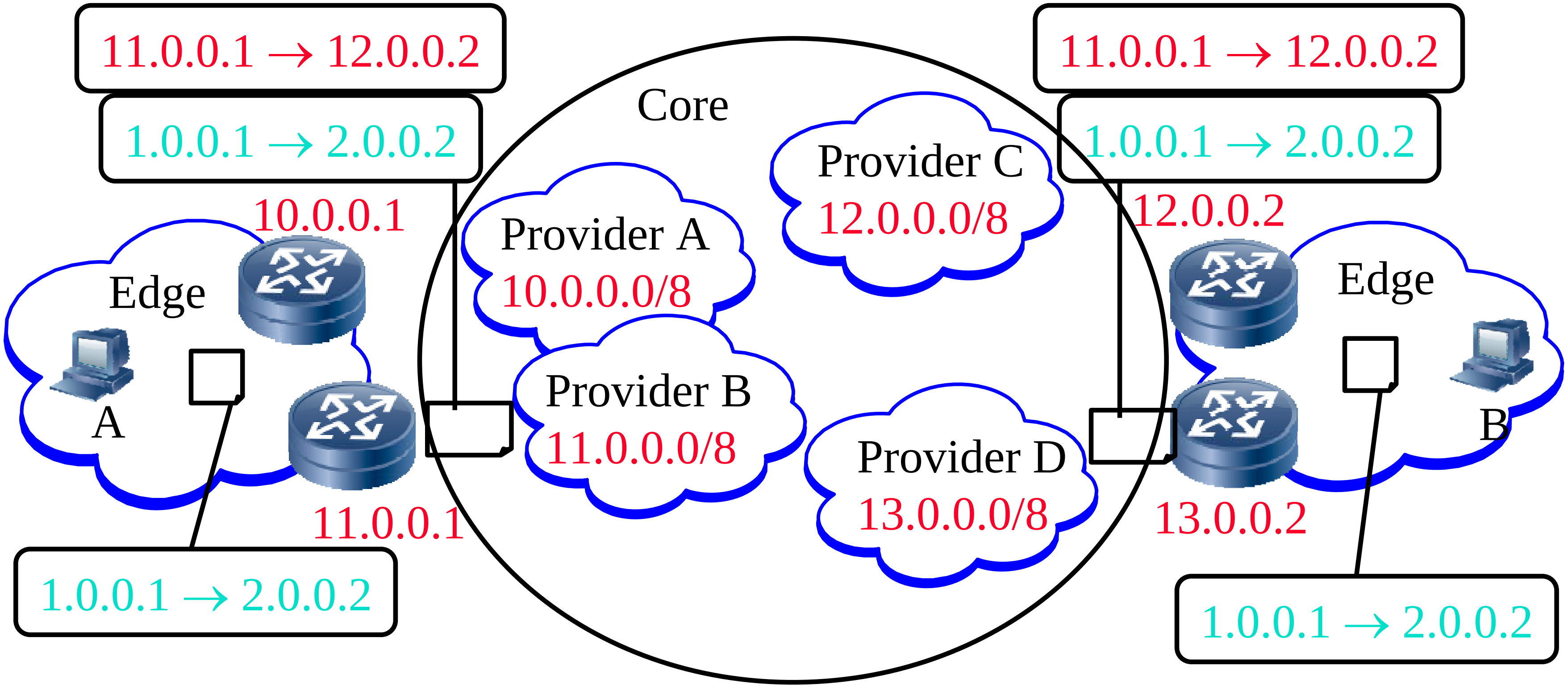


- Edge routing is based on ID.
  Core routing based on global locators
- Requires no changes to end hosts.
  Preferred solution for router vendors.
- Proposals: LISP, IVIP, DYNA, SIX/ONE, APT, TRRP
- LISP: Locator-ID Separation Protocol
- Uses IP-in-IP tunneling
- Edge routers translate IDs to routing locators and encapsulate
- No changes to core routers
- No centralized ID to locator mapping databases
- No support for host mobility, multihoming, traffic engineering

Refs: [LIT09, MEY07, RRG08]

Locators	EID	IP Data

LISP Protocol Details



EIDs: Green
Locators: Red

DNS: B→2.0.0.2

Mapping Entry:
EID-Prefix: 2.0.0.0/8
Locator Set:
12.0.0.2, priority 1, weight 50
13.0.0.2, priority 1, weight 50

Ref: *www.nanog.org/meetings/nanog41/presentations/* ***lisp****-nanog-abq.pdf*

MILSA

- Mobility and Multihoming Supporting ID-Locator Split Architecture
- New layer 3.5
- Distributed ID-locator mapping service
- New IDs = Secure Flat+Hierarchical
  ⇒ No PI addresses in the global routing
- Proxy edge routers can translate IDs to locators
  ⇒ Allows both router based and host based transition
- Incremental deployment

Refs: J. Pan, S. Paul, R. Jain, et al., ``MILSA: A Mobility and Multihoming Supporting Identifier Locater Split Architecture for Next Generation Internet," GLOBECOM 2008, http://www.cse.wustl.edu/~jain/papers/milsa.htm

6. Green Networking

- Information and Communication Technology: Energy Stats
- Effect of Networking
- Network Component Design
- Performance and Sleep States
- Rate Adaptation
- Wireless Mobile Networking

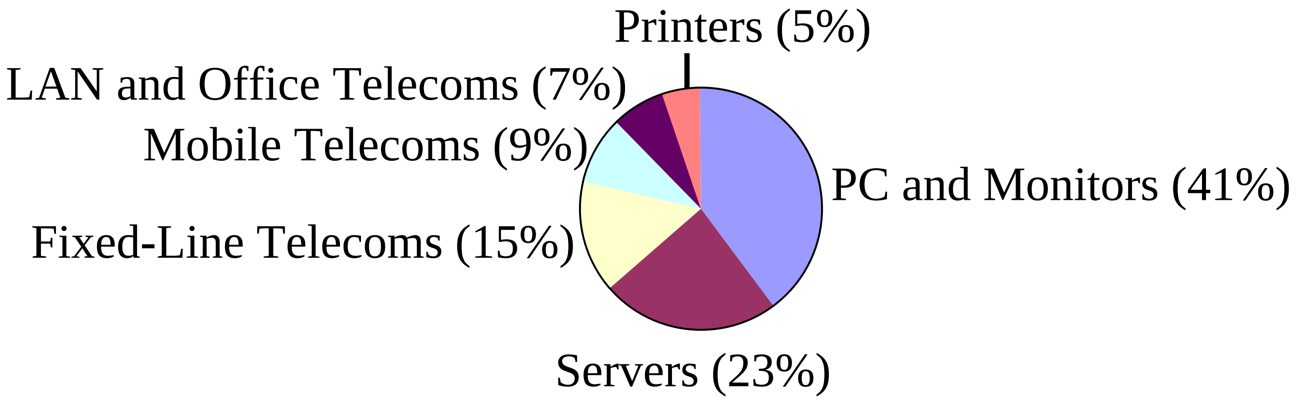
Info and Comm Technology Energy Stats

1. ICT produces 2-3% of GHG mostly through consumption of electricity produced by coal powered generator
2. ICT emissions equal to entire aviation industry
3. ICT emissions are doubling every 4 years ⇒ fastest sector
4. One small computer server = one SUV with 15 miles/gallon
5. Energy is 2nd highest data center costs (10% rising to 50%)
6. Cooling accounts for 20-50% of the total power consumption
7. Idle servers consume 50 to 80% of power at maximum load

- ❑Ref: An inefficient truth, http://www.globalactionplan.org.uk
- ❑Ref: HTTP://UCLUE.OM/INDEX.PHP?XQ=724
- ❑Ref: http://www.ee.unimelb.edu.au/people/rst/talks/files/Tucker_Green_Plenary.pdf
- ❑Ref: http://www.nanog.org/mtg-0802/levy.html
- ❑Ref: Http://esdc.pnl.gov/

Effect of Networking

- Networking devices account for about 15% of a data center's total energy consumption
- Future Broadband Internet is expected to consume 5% of all electricity
- Networking uses 8W/port, 3 ports/server
  $\Rightarrow$ 75% servers, 15% storage, 10% network [EPA Aug 2007]



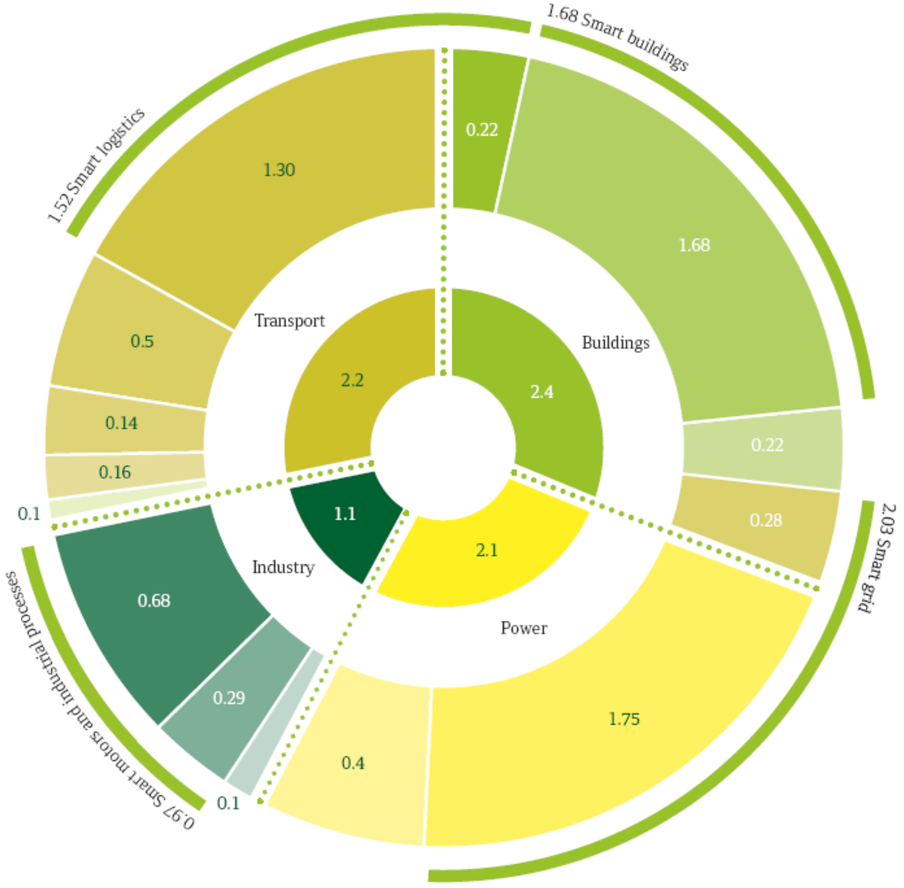
Ref: N. Chilamkurti, et al, "Green Networking for Major Components of Information Communication Technology Systems," 25th September 2009, EURASIP Journal of Wireless Communications and Networks

Networks: Green Enabler

- Use to reduce travel, Video Conference, Web meetings
- Allow remote working
  ⇒ Reduce desks, office space, commuting
- Use for remote sensing and energy control
- SOAK: Smart Operational Agriculture tool Kit
  ⇒ Weather forecast and sensors (ground moisture/temperature/wind/rainfall) to control irrigation)
- Distance learning courses rather than telecommuting

ICT's Enabling Effect

- ICT can help reduce (7.8 $GtCO_2e$) = 5 × ICT’s carbon footprint



Ref: SMART2020, “Enabling the low carbon economy in the information age,” 2008

Network Component Design

- Energy efficient components
- Efficient power adapters
- Eliminate fans or adjust based on needs
- Auto hibernate unused ports, devices
- Adjust signal strength based on cable length
- Challenge: Power over Ethernet
  - 15.4W/Port 384 port switch ⇒ 5.9kW
  - Increasing to 56W/Port ⇒ Larger battery backup
- Can save 2/3 by turning of VOIP phones 16 hours/day

Performance and Sleep States

- ❑ Processors have multiple sleep states
  ⇒ Power off different subcomponents
- ❑ Sleep states help when equipment is idle
- ❑ Disable ports, line cards, and switches when idle
- ❑ Ports are awakened when packets arrive
- ❑ A centralized power controller monitors the network and controls all ports
- ❑ Edge devices transmit packets in bursts
- ❑ Dynamically set the forwarding capacity of a port based on load: 1 Gbps, 100 Mbps, 10 Mbps, disabled
  ⇒ Energy Efficient Ethernet (IEEE 802.3az standard)

Why Network Devices Cant Sleep?

- ❑ Broadcast packets: ARP. Continuous Hellos?
- ❑ 1 to 3 packets per second
- ❑ DHCP leases expires
- ❑ A proxy service (running in a middle box firewall, DSL router) can help handle many of these tasks
- ❑ Broadcast traffic in office:

ARP	NBNS	IPX	NBDGM	LLC	ANS	RPC	BOOTP	NTP	Other
46%	23%	10%	6%	3%	3%	3%	2%	1%	3%

- ❑ Multicast traffic in office:

HSRP	SSDP	PIM	IGMP	EIGRP	Other
60%	25%	6%	5%	2%	2%

NBNS = Netbios Name Service
NBDGM = Netbios Datagram Service
HSRP=Hot-Standby Router Protocol
SSDP=Simple Service Discovery Protocol

Ref: S. Nedevschi, et al, “Skilled in the art of being idle: reducing energy waste in networked Systems,” NSDI 2009.

Rate Adaptation

- Processors have multiple performance States
  ⇒ Different frequencies
- Performance states help when the equipment is active
- Adapt rate of individual links by load
- 10G devices need more power than 1G than 100M
- Lower voltage and lower frequency
  ⇒ Change speed by queue sizes
- Reduce energy consumption by 50%

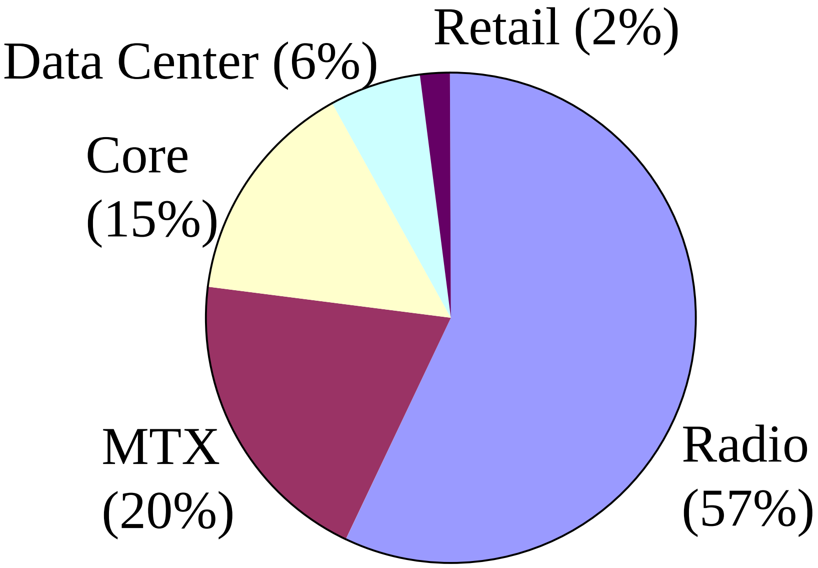
Component	100Mbps	1GBps
CPU	6W	6W
Interface	2W	12W
Other	9W	9W

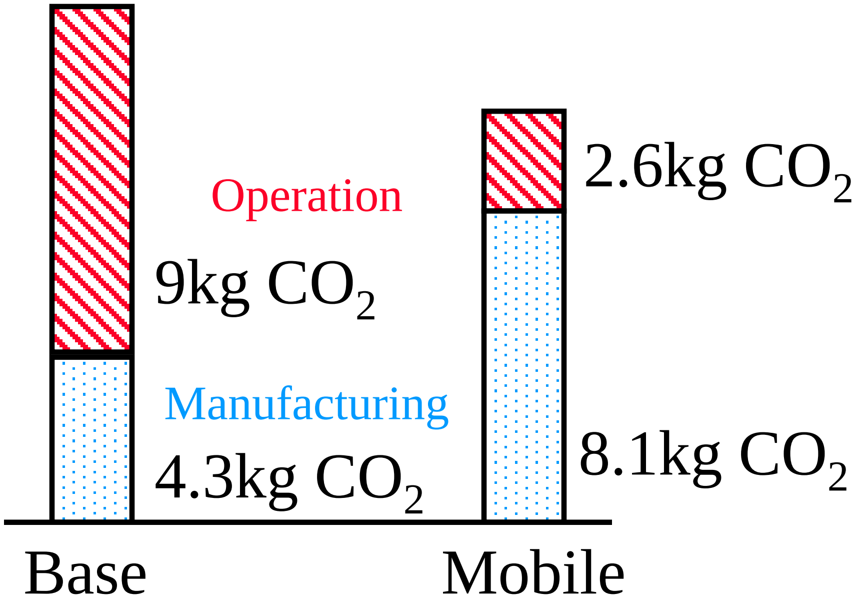
Ref: [Ned08]

Ref: W. Feng, IPDPS 2009

Wireless Mobile Networking

- 57% of the operator's electricity use is in radio access
- Base stations use most of the energy during operation
- Subscriber devices use most of the energy in manufacturing

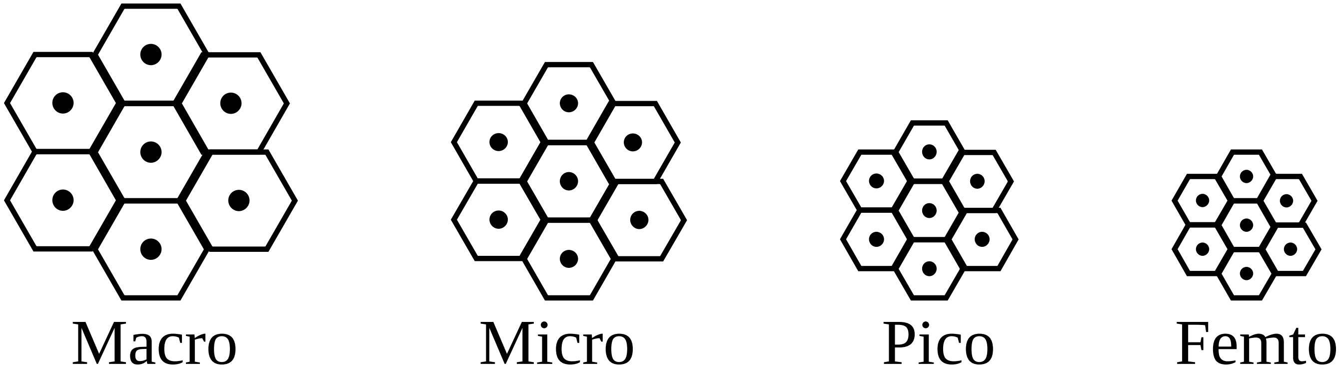




Ref: Tomas Edler, "Green Base Stations – How to Minimize CO2 Emission in Operator Networks," Ericsson, Bath Base Station Conference 2008

Wireless (Cont)

- Exponential Growth in Traffic ⇒ Reduce energy per data bit ⇒ lower power per cell
- Power consumption is a major issue in emerging markets
- Use renewable energy at base station
- Sleep techniques to allow power to scale with load
- Macro vs. Micro vs. Pico vs. Femto cells
  Energy is proportional to distance
  ⇒ operating vs. manufacturing energy tradeoff

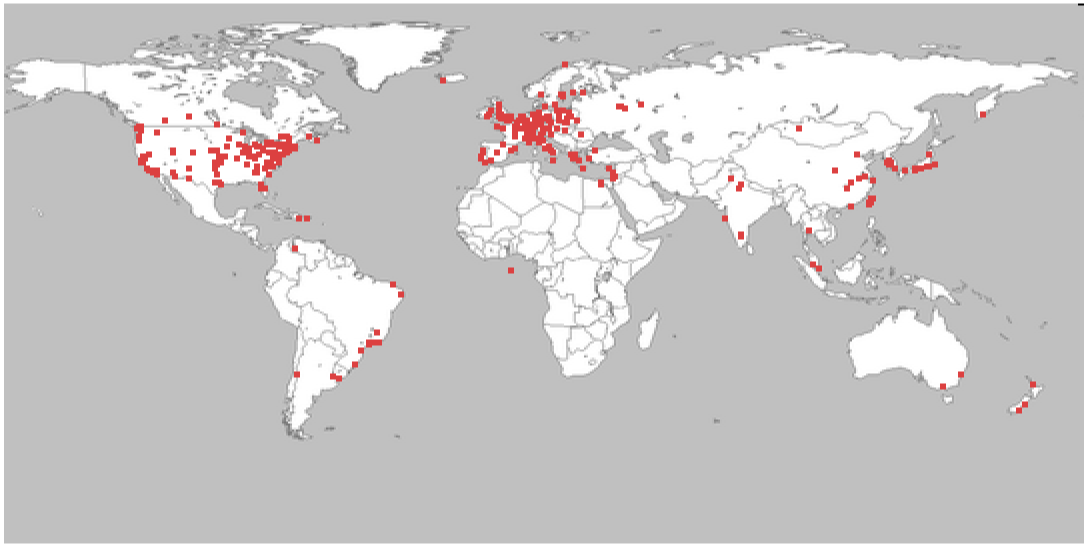


7. Next Generation Testbeds

- Past: PlanetLab, Emulab, VINI, OneLab
- Federation
- GENI, Requirements, Subsystems
- GENI Prototype Clusters
- Wireless Network Virtualization
- Supercharged PlanetLab Platform (SPP)
- FIRE, FEDERICA
- AKARI

PlanetLab

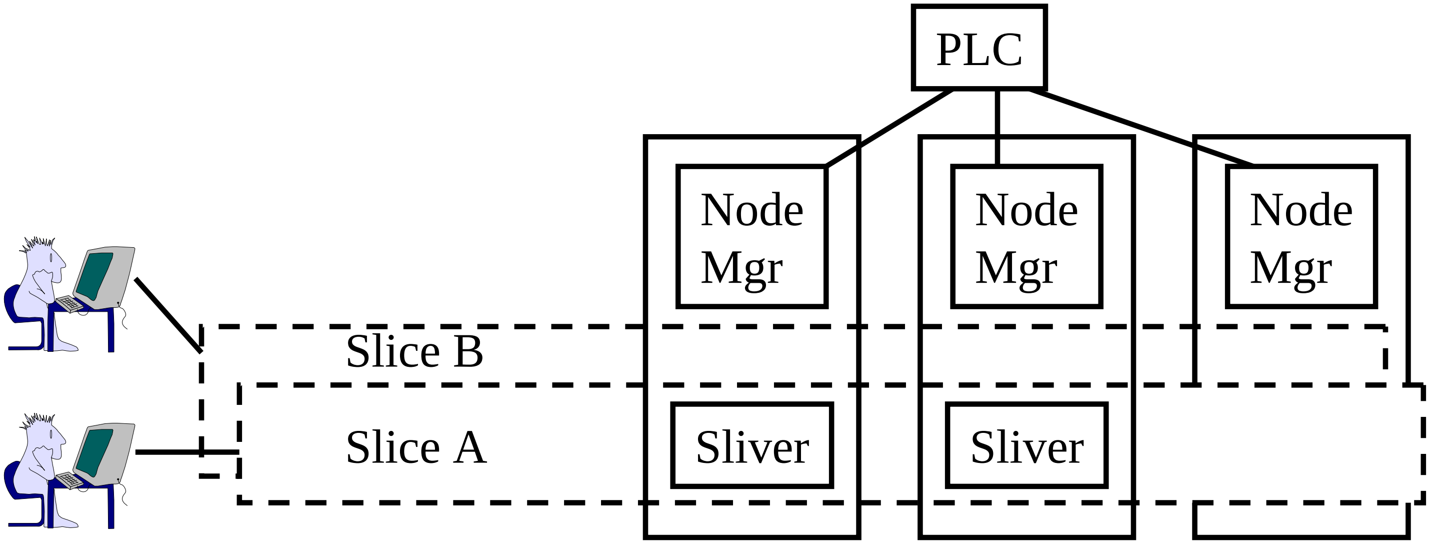
- Global networking research testbed
- 1055 nodes at 490 sites [Nov 2009]
- Researchers use it to experiment with new ideas on distributed storage, network mapping, peer-to-peer systems, distributed hash tables, and query processing



Ref: http://www.planet-lab.org/

PlanetLab (Cont)

- Linux virtual server software on Interned nodes
- **Slivers** = Piece of a resource
- Node manager (**NM**) manages the node's virtual servers
- Planet Lab Control (**PLC**) interacts with NM
- Experimenters request a "**Slice**" = slivers in various sites



Emulab

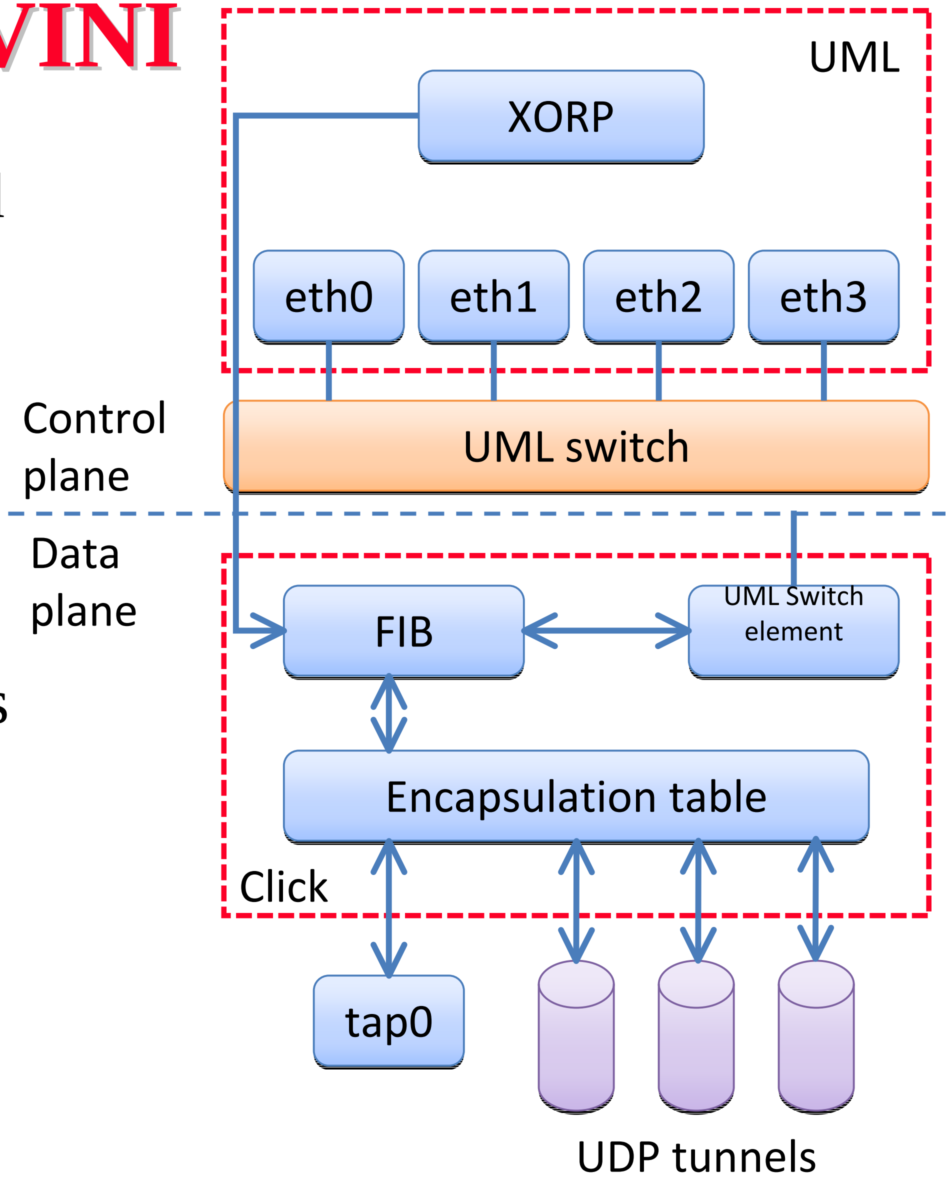
- Networking research testbed at University of Utah
- Available for public use for research and education
- Software implemented at two dozen sites around the world
- Allows simulated links and nodes in slices
  ⇒Allows fault studies
- Provides repeatability

[emulab.net]

Ref: http://www.emulab.net/

VINI

- ☐ A private PlanetLab testbed for Routing protocols
- ☐ Install User Model Linux (UML) on the slice
- ☐ Runs open source router code XORP on UML
- ☐ Flows through VINI routers also go over real routers
  ⇒ Not clean



Refs: [BAV06] and http://www.vini-veritas.net/

OneLab

- Federation of European PlanetLabs
- Enhanced monitoring infrastructure
- Wireless Testbeds
- IPv6 multihoming
- Emulation tools
- Deal with unstable connectivity

Ref: OneLab,http://www.onelab.eu

Federation

- Larger testbeds
- Testbeds for specialized resources such as access technologies
- Specialized research communities and cross-discipline
- Challenges:
  - Homogenization of diverse context
  - Interoperability of security protocols
  - Political or social-economic issues
  - Intellectual Property rights
  - Commercial and non-commercial interests

Ref: OneLab2 Whitepaper: ``On Federations…, January 2009,
http://www.onelab.eu/index.php/results/whitepapers/294-whitepaper-1-on-federations.html

GENI

- ❑ Global Environment for Network Innovations
- ❑ Dedicated shared substrate facility for large-scale experiments
- ❑ US National Science Foundation project
- ❑ Dedicated backbone links through LambdaRail and Internet2
- ❑ Diverse and extensible set of technologies

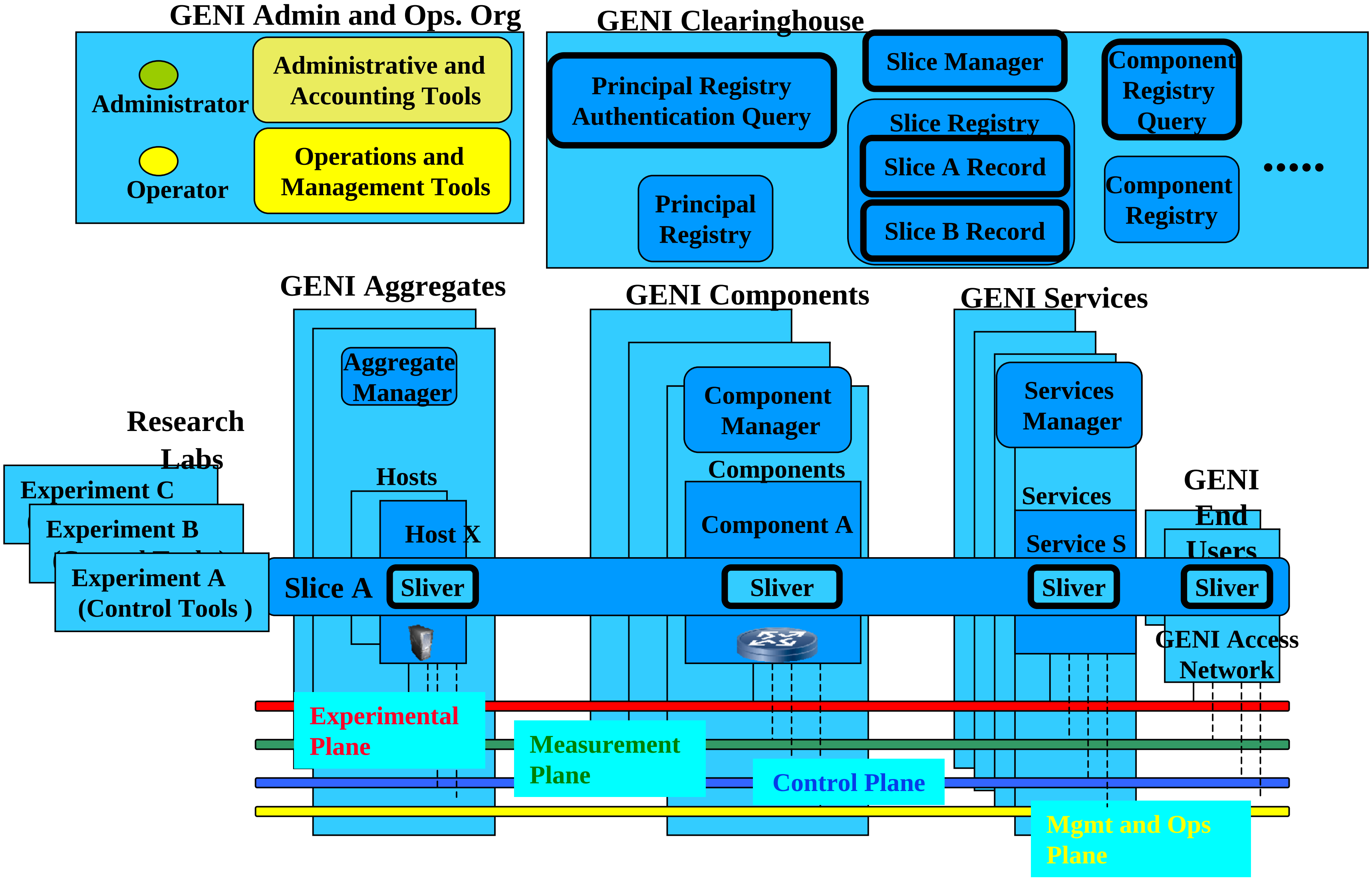
Refs: [GENI01, ON410]

GENI Requirements

- Sliceability: Sharing with isolation.
- Programmability: All components should be programmable
- Virtualization: Slicing via virtualization or space/time sharing.
- Federation: Combination of independently owned testbeds
- Observability: Allow specifiable measurement framework
- Security: Should not harm production Internet

Refs: [AND052, SHA05, CLA05, RAY05, BLU05, BELL05, KAA05]

GENI Subsystems



GENI Admin and Ops. Org

Administrator

Operator

Administrative and Accounting Tools

Operations and Management Tools

GENI Clearinghouse

Principal Registry Authentication Query

Principal Registry

Slice Manager

Slice Registry

Slice A Record

Slice B Record

Component Registry Query

Component Registry

•••••

Research Labs

Experiment C

Experiment B

Experiment A (Control Tools)

GENI Aggregates

Aggregate Manager

Hosts

Host X

GENI Components

Component Manager

Components

Component A

GENI Services

Services Manager

Services

Service S

GENI End Users

GENI Access Network

Slice A

Sliver

Sliver

Sliver

Sliver

Experimental Plane

Measurement Plane

Control Plane

Mgmt and Ops Plane

GENI Subsystems

- Component: A device that provides resources that are shared via virtualization.
- Aggregate: A set of components under a central control
- Clearinghouse: Centralized registry of principles, slices, and components
- Control Framework: Access control policies, trust/federation mechanisms
- Measurement Subsystem: Archival and retrieval of experimental data
- Administration and Operation: Incorporate new resources, identify mis-behaving resources
- Experimenter Tools and Services: Resource discovery, reservation, debugging, etc.

GENI Subsystems (Cont)

- GENI Generalized Control Framework
- Each component has a component manager.
- Each aggregate has an aggregate manager.
- Slice manager reserves slices for an experiment.
- Interfaces and messages between entities are defined.

Ref: GENI-SE-CF-RQ-01.3: GENI Control Framework Requirements, January 9, 2009, http://groups.geni.net/geni/attachment/wiki/GeniControlFrameworkRequirements/010909b%20%20GENI-SE-CH-RQ-01.3.pdf

GENI Prototype Clusters

Five Clusters in Spiral 1:

1. Trial Integration Environment with DETER (TIAD): Emulab based security experiments testbed
2. PlanetLab: Federate all slice-based substrates PlanetLab, Emulab, VINI, and GENI
3. ProtoGENI: Federation of Emulab testbeds, Enhanced Emulab Control
4. Open Resource Control Architecture (ORCA):
   Resource manager runs under the host operating system
   Uses virtualization to allocate containers
5. Open Access Research Testbed (ORBIT): Wireless testbed with emulated and real nodes

- ☐ Spiral 2: Improved instrumentation, tools for integration
- ☐ Spiral 3: Integration. Experimentation across clusters.

Ref: GENI Spiral 1, http://groups.geni.net/geni/wiki/

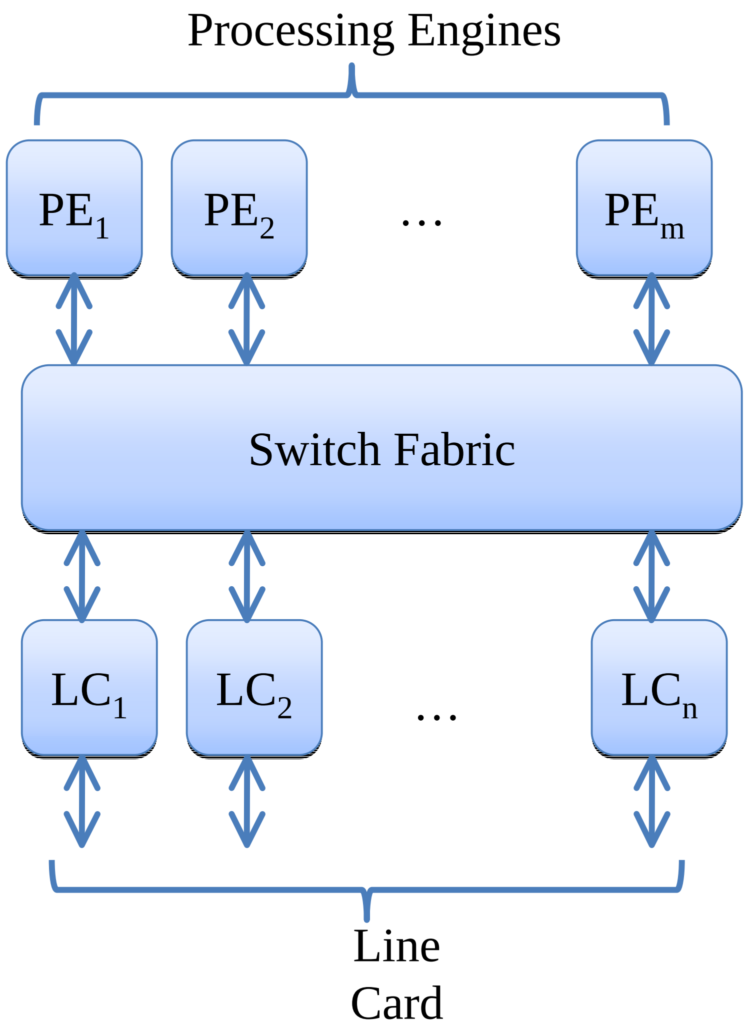
Wireless Network Virtualization

- Low bandwidth ⇒ Can't share a link ⇒ Partition the network for isolated experiments
- Coding, multiplexing are node specific
- Slicing via Space Division Multiple Access/FDMA/TDMA/combinations
- Can’t overprovision the spectrum
- Nodes of different types cannot substitute (WiFi vs. WiMAX)
- Sharing Techniques: FDMA, TDMA, FDMA+TDMA, Freq Hopping, CDMA
- Slicing Techniques: SDMA, SDMA+TDMA, SDMA+FDMA, SDMA+TDMA+FDMA

Ref: S. Paul, S. Seshan, ``GDD-06-17: Technical Document on Wireless Virtualization," GENI Design Document 06-17, September 2006, http://groups.geni.net/geni/attachment/wiki/OldGPGDesignDocuments/GDD-06-17.pdf

Supercharged PlanetLab Platform (SPP)

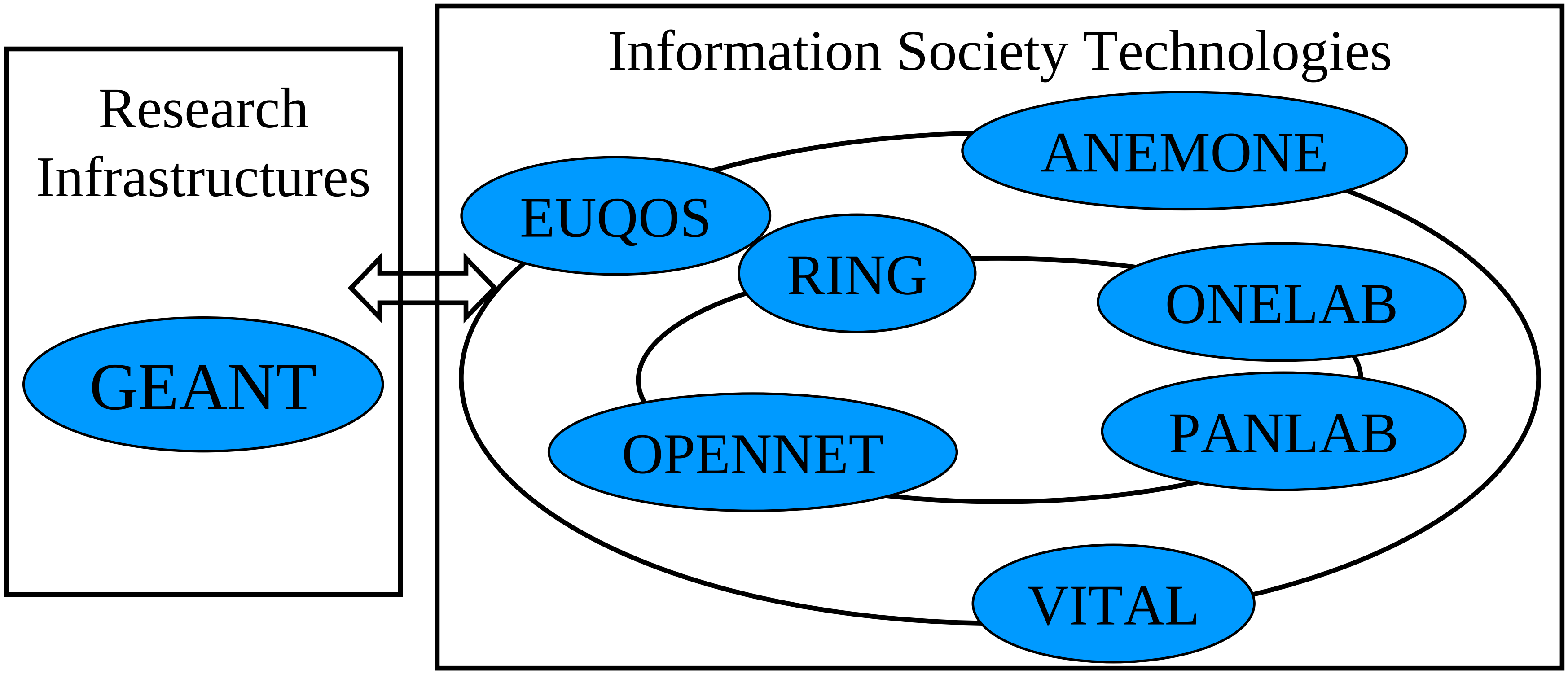
- Allows multiple virtual routers w different stacks
- Fast path for line speed packet forwarding
- Slow path for application specific processing
- Multiple meta-networks (routers, links) on a substrate
- 3 Components: Line cards, switching fabric, control proc
- Virtualizing line cards is difficult
- Processing Pool Architecture:
  No processing in line cards
  Simply switch to proc engines



Refs: [TUR06, TUR107, TUR207]

FIRE

- Future Internet Research and Experimentation
- Federate multiple existing testbeds in Europe
  ⇒ Provide a large multi-context research testbed



Research Infrastructures

GEANT

Information Society Technologies

ANEMONE

EUQOS

RING

ONELAB

PANLAB

OPENNET

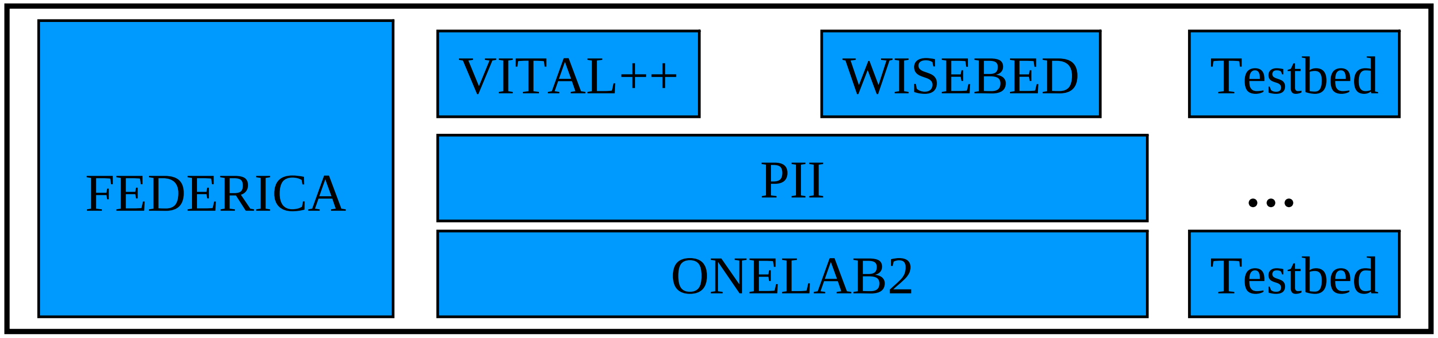
VITAL

FIRE (Cont)

- GEANT connects 30 national res and education nets (NRENs)
  - Onelab2: Develop economic incentive models for sharing
  - Panlab II (PII): Pan European Laboratory –
    A consortium of telecom providers.
    Federate test labs and testbeds for large scale experiments
  - VITAL++: P2P technology
  - WISEBED: Wireless Sensor Networks Testbed
- Implements IPv6, IP with QoS, Multicasting, Prioritized IP
- Production network like NSFNet, LambdaRail, Internet3

Refs:[ON411, ON426, ONE09, ON427, ON404, ON428, ON429, ON430, ON431]

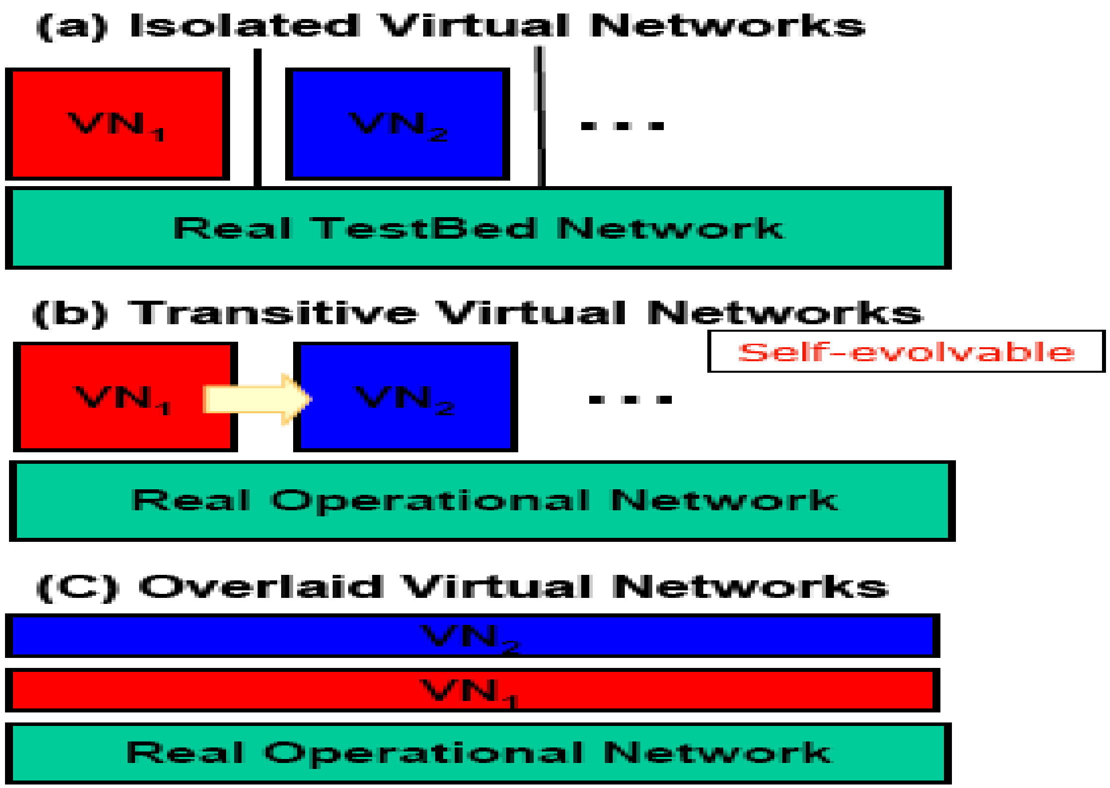
FEDERICA



- Federated E-infrastructure Dedicated to European Researchers Innovating in Computing network Architectures
- Goal: Test new paradigms at large scale
- End-to-end slice via virtualization like GENI
- Substrate is agnostic about protocol, services and applications Similar to Diversified Internet Architecture
- 12 PoPs of GEANT2 and NRENs connected via 1 Gbps
- 4 core + 8 on-core sites: Core connected via dedicated links Non-core connected via GEANT2 or NRENs or Internet
- Core allows BGP peering with Internet

Refs: [FED01, FED02], http://www.fp7-federica.eu/

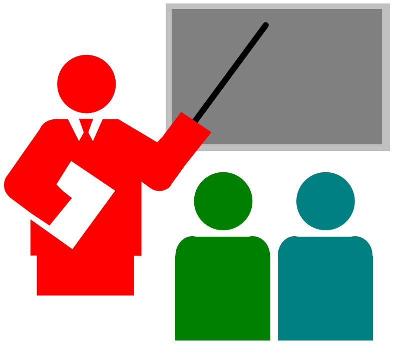
AKARI



- ❑ Research on optical switching, ID-Locator, Self-organizing control, and network virtualization
- ❑ Isolated/transitive/overlaid virtual networks

Refs: [ON408, HAR08]

Summary



1. NSF FIND program has funded a number of architectural component research programs.
2. GENI testbed consists of 5 clusters in Spiral 1
3. FIRE in Europe and AKARI in Japan are similar to GENI.
4. Internet 3.0 is an industry sponsored full architecture program

Top 10 Features of Next Generation Internet

1. Security
2. Mobility
3. User/Data-Centric: Network support of data objects
4. Easy to use: Self-organizing, better user control
5. Disruption Tolerant
6. Green: Proxy, Sleep Modes,
7. Services: Storage, Translation, Monitoring
8. Organizational Representation
9. Virtualizable to create Application Specific Context
10. Policy Enforcement

NSF FIA Winners

- ☐ **Named Data Networking**: CCN
  - ➢ Routing scalability, Fast forwarding, Trust models, Network security, Content protection and privacy
- ☐ **Mobility First**: Generalized Delay Tolerant Networking with self-certifying public key addresses
- ☐ **Nebula** (Latin for Cloud): Trustworthy data, control and core networking for cloud computing
- ☐ **eXpressive Internet Architecture (XIA):** Application programming interface (API) for communication, flexible context-dependent mechanisms for establishing trust

Ref: NSF Announces Future Internet Architecture Awards, August 27, 2010, http://www.nsf.gov/news/news_summ.jsp?cntn_id=117611