Routing in Future Internet

2010. 10

Deokjai Choi

Outline

- What is Routing?
- Why Future Internet?
- Routing Problems for FI
- Proposed Solution

What is Routing?

- Routing : Routing refers to the process of choosing a path over which to send packets.(source to destination)
- desirable properties: correctness, simplicity, robustness, stability, fairness, optimality
- what optimize?
  - Mean packet delay
  - network throughput

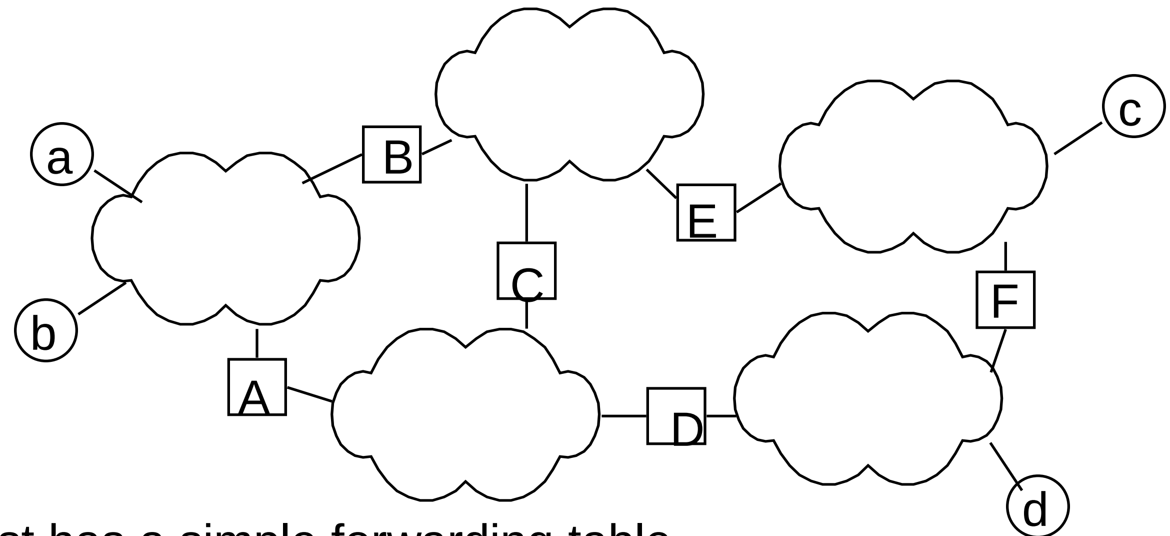
Routing

- How to construct routing tables
- Routing - Determine end-to-end paths
- Forwarding - Transmit packets according to routing table

Routing

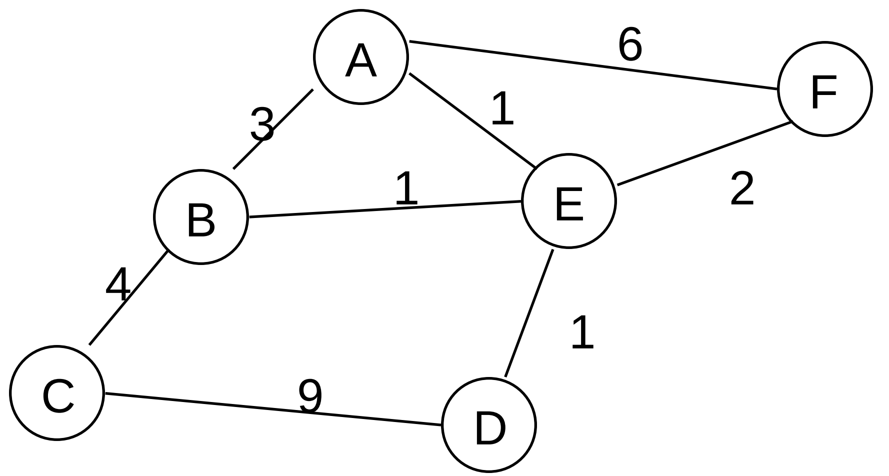
- Network layer supports routing over internet consisted of multiple physical networks
  - Form a logical network
  - Router
  - If possible, a packet should be routed over the shortest path between source & destination

IP Routing



Each host has a simple forwarding table
Router has a larger forwarding table
Case 1: Host *a* --> Host *b*
Host *a* should know that host *b* is in the same physical network
How?
Case 2: Host *a* --> Host *c*
Host *a* relay datagram to router A or B
IP only routes the datagram to router E or F

Routing



Problem: Find the lowest cost path between any two nodes

Under dynamic network changes

Why Future Internet?

- 2000s Internet becoming
  Social Infrastructure
- Problems

2000s Internet becoming Social Infrastructure

Internet population: one billion

Broadband Internet

Wireless and Mobile Internet

Personal Website

Convergence (Internet, Telephone, Television, Movie,...)

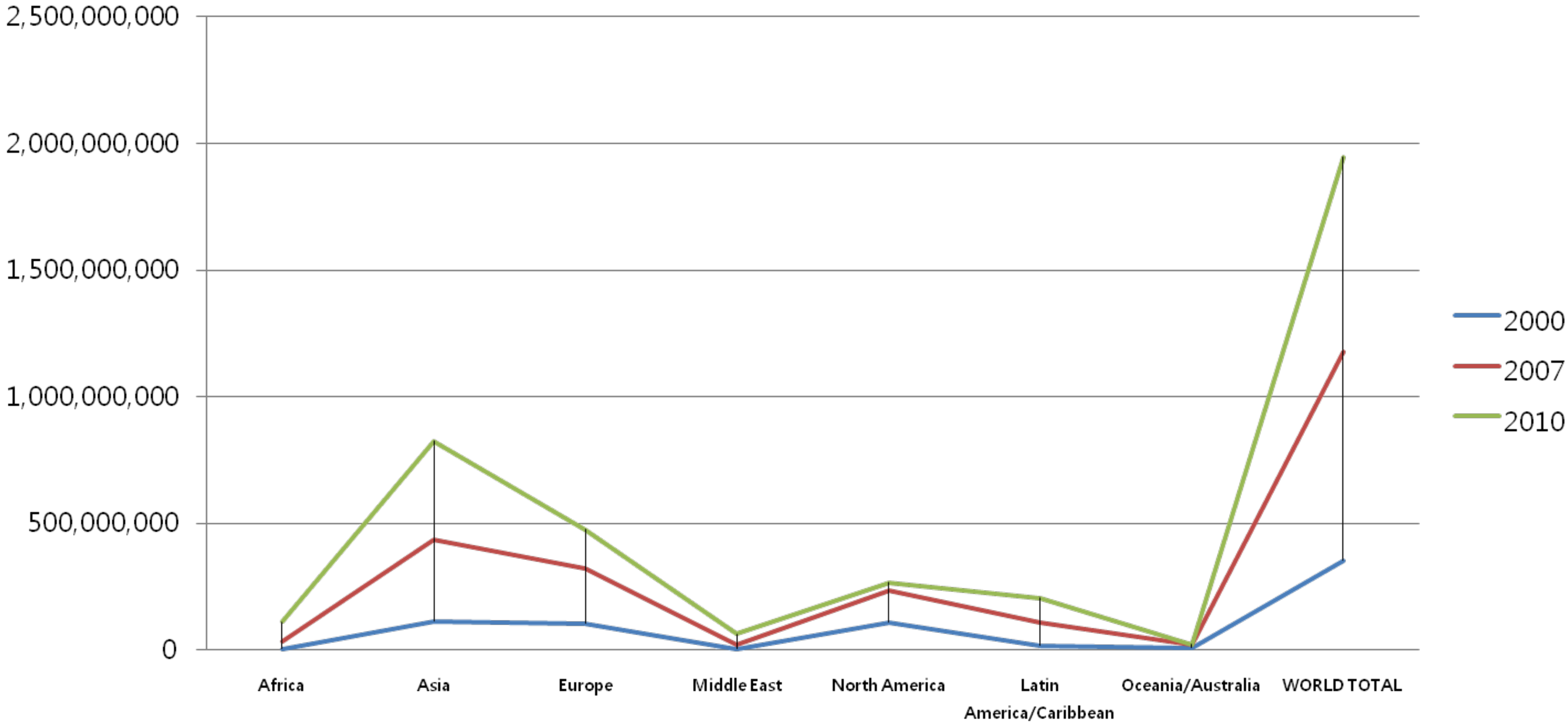
Negative Side Effects (spam, virus, privacy,...)

Northeast Asia as one of leading Internet regions

Internet Population

WORLD INTERNET USAGE AND POPULATION STATISTICS								
world Regions	Population (2010 Est.)	Internet Users			Penetration (%Population)	Growth 2000-2010	Users% of Table	
		2000	2007	2010			2010	2007
Africa	1,013,779,050	4,514,400	34,000,000	110,931,700	10.9 %	2,357.3 %	5.6 %	3.5%
Asia	3,834,792,852	114,304,000	437,000,000	825,094,396	21.5 %	621.8 %	42.0 %	36.9%
Europe	813,319,511	105,096,093	322,000,000	475,069,448	58.4 %	352.0 %	24.2 %	27.2%
Middle East	212,336,924	3,284,800	20,000,000	63,240,946	29.8 %	1,825.3 %	3.2 %	2.7%
North America	344,124,450	108,096,800	233,000,000	266,224,500	77.4 %	146.3 %	13.5 %	18.9%
Latin America/Caribbean	592,556,972	18,068,919	110,000,000	204,689,836	34.5 %	1,032.8 %	10.4 %	9.3%
Oceania/Australia	34,700,201	7,620,480	19,000,000	21,263,990	61.3 %	179.0 %	1.1 %	1.5%
WORLD TOTAL	6,810,909,759	353,367,012	1,175,000,000	1,945,252,836	28.7%	444.8%	100%	100%

Internet Users Change Chart



Wireless / Mobile Internet

Internet with Computer : 500 millions

Internet with Mobile Phone: 400 millions

Remark : Mobile phones : 2.5 billions
Internet Users : 1 billion

Remark : In 2010~2020, 80% of the Internet usage are mobile.

Negative Side Effects/Social Issues

Virus

Spam

Privacy

Intellectual Property

Future Internet (~2020)

Current Status

Internet was created for research community(~1970s).

One billion people are using the Internet now.
- One trillion machines are expected in future.
- Five billion users needs to be connected.

Toward critical/social infrastructure
- Water
- Electricity
- Road
- Internet / Phone / Television

Problems

Scalability (Users, Bandwidth)
Security / Trust
Mobile / Wireless
Management
(Semantic Overhead on IP)
(Engineering)
"The Other Billions"

What will be happening in 10 years

- New network technology.
  - Wireless
    - Mobility
    - Dynamic capacity allocation
    - Dynamic impairments
  - Advanced optics
    - Dynamic capacity allocation (again!)
- New computing paradigms
  - Embedded processor, sensors, everywhere
- Whatever computing is, that is what the Internet should support.
  - The Internet grew up in a stable "PC" time.

Problem Statement (1/4)

1. Basic Problems

1.1. Routing Failures and scalability

- The problems have been examined as being caused by mobility, multi-homing, renumbering, PI (provider independent) routing, IPv6 impact, etc. on the current Internet architecture.

1.2. Insecurity

- As current communication is not trusted, problems are self-evident, such as the plague of security breaches, spread of worms, and denial of service attacks.

1.3. Mobility

- Current IP technologies was designed for hosts in fixed locations, and ill-suited to support mobile hosts.
- Mobile IP was designed to support host mobility, but Mobile IP has problems on update latency, signaling overhead, location privacy, etc.

Problem Statement (2/4)

1. Basic Problems

1.4. Quality of Service

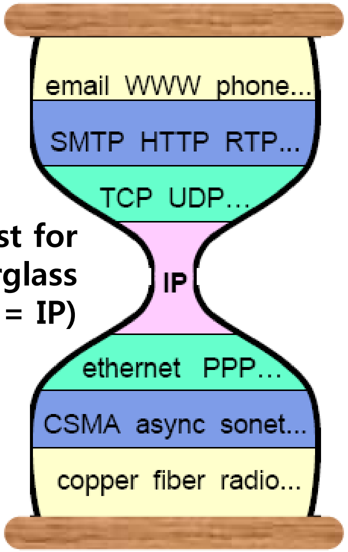
- Internet architecture is not enough to support quality of service from user or application perspective.
- It is still unclear how and where to integrate different levels of quality of service in the architecture.

1.5. Heterogeneous Physical Layers and Applications

- Recently, IP architecture is known as a "*narrow waist or thin waist*".
- Physical Layers and Applications heterogeneity poses tremendous challenges for network architecture, resourc allocation, reliable transport, context-awareness, re-configurability, and security.

1.6. Network Management

- The original Internet lacks in management plane.



Narrow Waist for Internet Hourglass (Common Layer = IP)

Source : Steve Deering, IPv6 :addressing the future

Problem Statement (3/4)

1. **Basic Problems**

   1.7. Congestive Collapse

   Current TCP is showing its limits in insufficient dynamic range to handle high-speed wide-area networks, poor performance over links with unpredictable characteristics, such as some forms of wireless link, poor latency characteristics for competing real-time flows, etc.

   1.8 Opportunistic and Fast Long-Distance Networks

   Original Internet was designed to support always-on connectivity, short delay, symmetric data rate and low error rate communications, but many evolving and challenged networks do not confirm to this design philosophy.

   – E.g., Intermittent connectivity, long or variable delay, asymmetric data rates, high error rates, fast long-distance communications, etc.

- **1.9. Economy and Policy**

  The current Internet lacks explicit economic primitives.

  There is a question of how network provider and ISP continue to make profit.

Problem Statement (4/4)

2. Problems with Original Design Principles

2.1. Packet Switching

– Packet switching is known to be inappropriate for the core of networks and high capacity switching techniques (e.g., Terabit).

2.2. Models of the End-to-End Principle

– The Models of the end-to-end principle have been progressively eroded, most notably by the use of NATs, which modify addresses, and firewalls and other middle boxes

2.3. Layering

– Layering was one of important characteristics of current IP technologies, but at this phase, it has inevitable inefficiencies.
– One of challenging issues is how to support fast mobility in heterogeneous layered architecture.

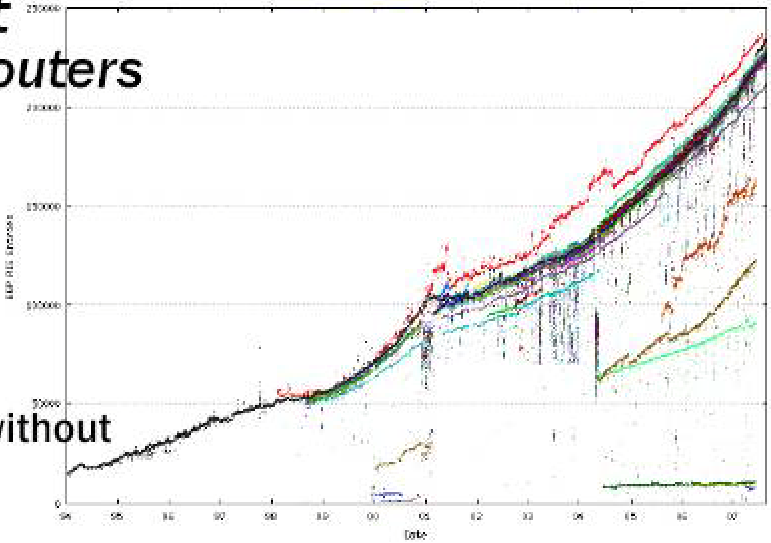
Routing Problems for FI

- Scaling Problem
  - IPv4 growth explosively
  - PI desire from users: destroys topology based address aggregation
  - Widespread of multihoming: destroys topology based address aggregation
- Usage Pattern change: Host Oriented -> Data (content) oriented
- Other Approaches
  - User Empowerments
  - Mgmt

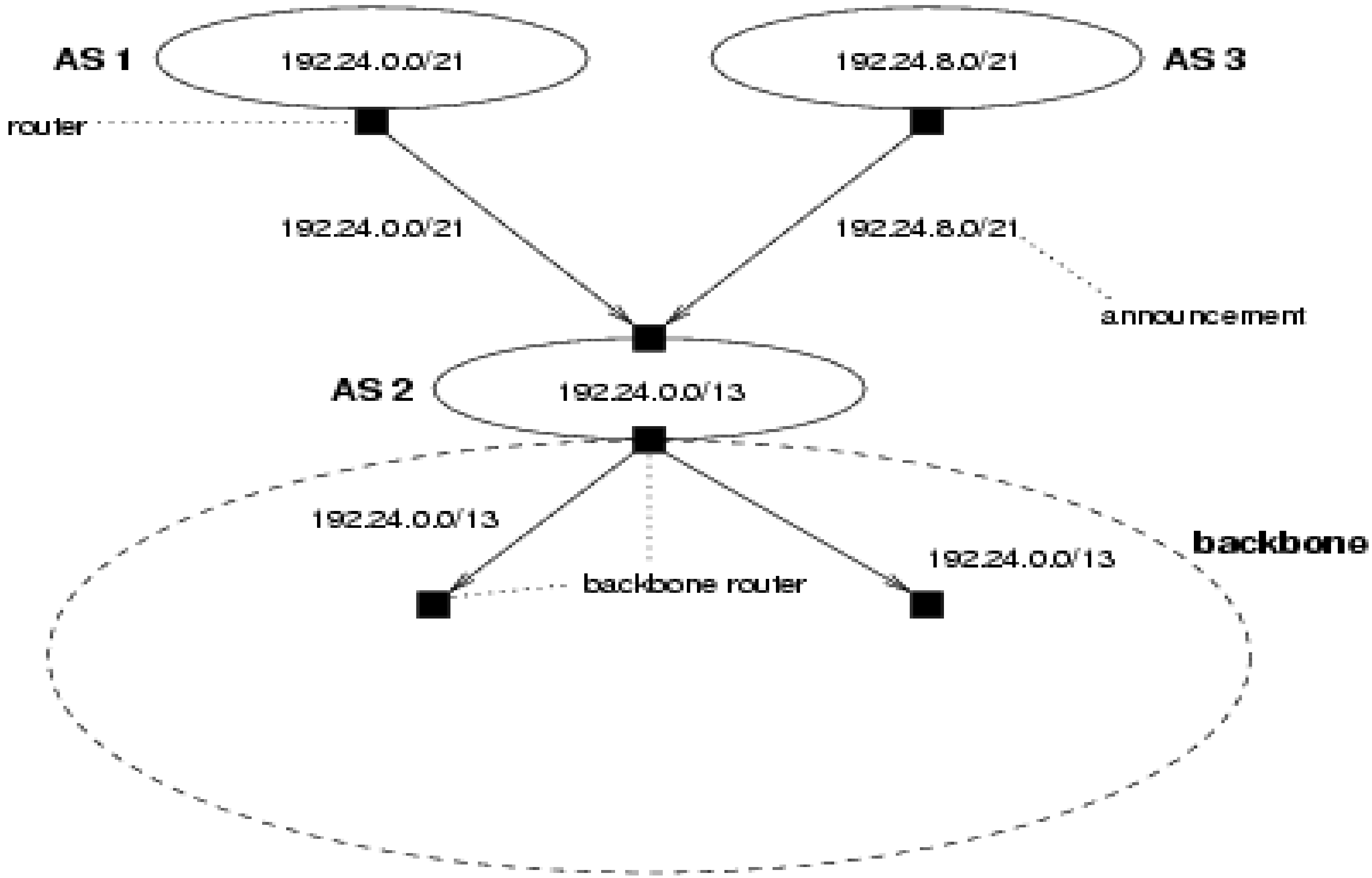
Routing Scalability

- *"Increasing rate of unaggregatable routing entries is so fast that the development speed of high-end hardware for core routers will not meet the requirements…"*
  - IETF IAB workshop on routing and addressing, 2007
- Hierarchical aggregation is broken
  - Provider-Independent addressing
    - Sites want to be able to change providers without renumbering
  - Site Multi-homing
    - Even if PI addressing is not used, multi-homing injects more-specific routes from one provider to another
  - Traffic Engineering
    - Providers inject more-specific routes to influence the behavior of the routing system, in order to control various traffic patterns

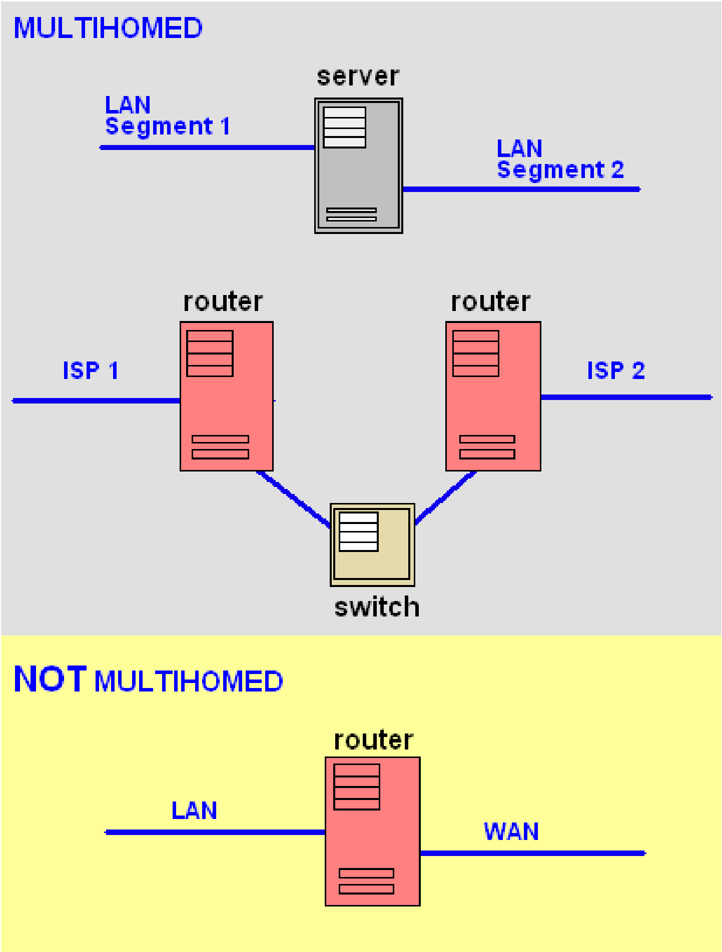
http://bgp.potaroo.net/



Prefix Aggregation



Multihoming



Proposed Solutions

- Scaling Problem
  - Separation address space: GRA for ISP and GDA for end networks
- Usage Pattern Change

The Separation of two address classes

- Address prefixes in the routing system should be topologically aggregatable, and aggregated when necessary to keep the table size under control.
- this desire of prefix aggregation runs into direct conflict with supporting end-site multihoming in the current routing system architecture.

IP address space – globally routable addresses(GRA) and globally deliverable addresses(GDA)

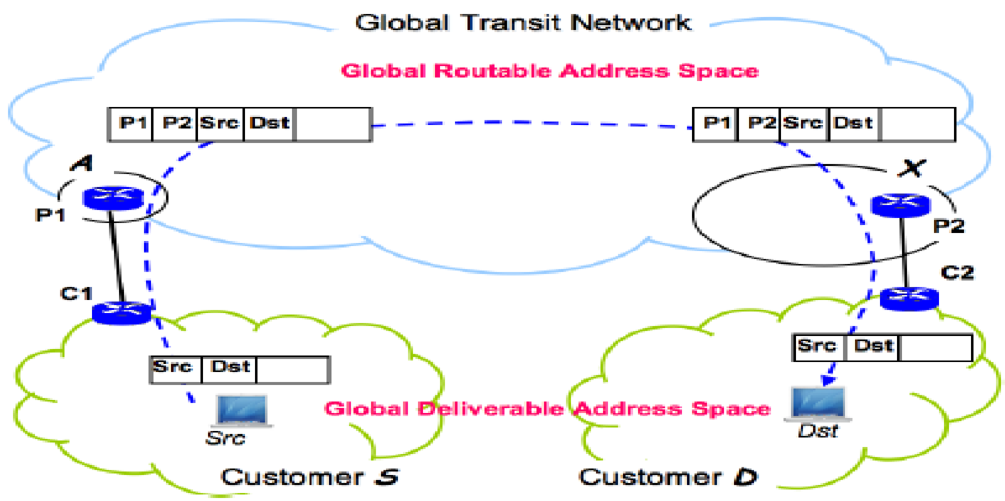


Figure 1: End-to-End Packet Delivery with Address Separation

Benefits from the separation

Routing Scalability and Stability

- Because of the separation of GDA from GRA, routing dynamics occurring inside end-sites or at the border (between end-sites and PNs) will no longer have an impact on the routing stability inside GTN.
- since the number of prefixes in the GTN is expected to be much smaller than the number of the prefixes in the routing system today, routing convergence would be substantially faster than that of today's BGP.

AS Number (ISP name)	Total Prefix	Transit Net. Pref (manual)	Transit Net. Pref (automated)
7018 (ATT)	1501	39	35
174 (Cogent)	930	21	19
1668 (AOL)	202	115	100
1239 (Sprint)	852	133	131
701 (Verizon)	4989	537	570
3549 (GBLX)	342	133	81
3561 (Savvis)	521	231	263
3356 (Level3)	514	50	99
209 (Qwest)	691	59	63

Table 1: Prefixes of some major ISPs

- The results show that, out of 209,549 prefixes in the global routing table, only 22,733, about 11%, belong to transit networks (although more were originated by them). Next we count the number of updates for transit network prefixes and end-site prefixes respectively during the month of August 2006. Out of 367 million updates from all the RouteViews monitors, only 57 million updates (15%) are for transit network prefixes.

Benefits from the separation

Site Multihoming and Traffic Engineering

- Once we separate end-sites to a separate address space (GDA), naturally the entire GDA address space becomes provider-independent. Customers may also want to fully utilize the parallel connectivities provided by multihoming.
- Since the address space separation between GDA and GRA introduces the need for a mapping function, we can utilize this mapping function for effective traffic engineering support.
- Customers can inject into the mapping record additional policy information to facilitate the selection of provider address among multiple alternatives.

Benefits from the separation

Security Enhancement

- Because our design puts all end hosts in an address space separate from that of backbone routers, all user data packets are encapsulated when they cross the backbone.

- Compromised hosts in the customer space no longer have direct access to the provider infrastructure.

- The encapsulation of end-user packets also makes it easy to trace attack packets back to the GTN ingress router even if they have spoofed source addresses, since the encapsulation header records the addresses of the GTN entry and exit routers.

Challenges

how to design scalable, secure and efficient mapping function, how to handle the failures between GRA and GDA, and how to conduct network measurement on the Internet backbone after the GRA and GDA separation.

The Mapping Function

given a destination customer address, it should return a destination provider address so that the packet can be encapsulated and forwarded across the Internet.

- Fast lookup: packets cannot be forwarded until the mapping is completed, so a fast lookup service is essential for good performance.
- Fast failure recovery: mapping entries should adapt quickly with changes.
- Resilience to abuses and attacks: mapping service can be a potential target for attacks. Updates to the mapping service or query replies from mapping service must be authenticated.

Challenges

Handling Border Link Failures

• Our proposed solution separate GRA and GDA address space, so that only topological changes in the GRA space, i.e. inside the global backbone, are handled by the global routing protocols.

• However, a link between an end-site D and its provider P is not part of the GRA routing space.
Thus when this link, or D's router at the other side of the link, fails, no routing update would be generated in the global routing system.
This can be viewed as an advantage as it provides the insulation of edge dynamics from the global routing system.

• At the same time this also introduces a challenge in assuring packet delivery, if the mapping function only reflects which providers connects to, but not whether the connectivity is up on a real time basis.

Challenges

Network Diagnosis

- The separation of GRA and GDA address space effectively presents end users a black box, which connects up all user networks but does not offer user networks any visibility or influence over the internal paths being used inside the transit backbone.

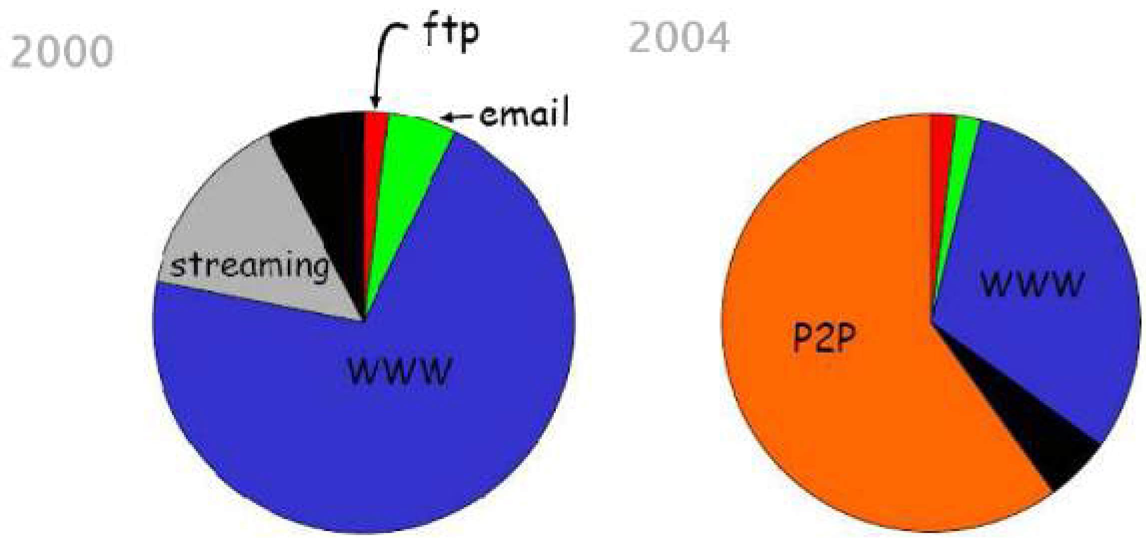
- End users can still measure the external behavior of this black box, detect any problems that affect their data delivery, and
move traffic between different access ISPs.

Proposed Solutions

- Scaling Problem
  - Separation address space: GRA for ISP and GDA for end networks
- Usage Pattern Change
  - Host Centric to Data (Service, Content) Oriented

Content-oriented Network [1/2]

- Current addressing of Internet is Host-centric
  - designed to identify Hosts
- Current usage of Internet is Data-centric
  - Overwhelming use of today's networks is for a machine to acquire named chunks of data



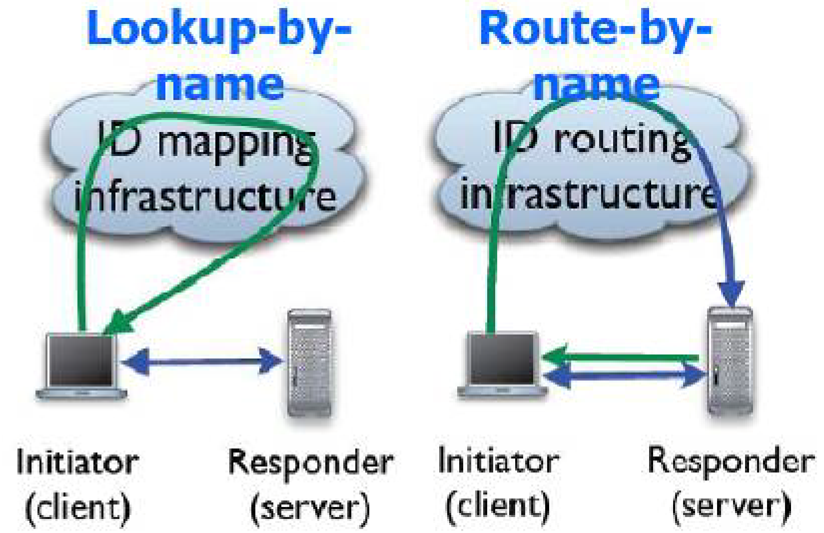
"The Internet is Flat", Don Towsley, Keynote Speech of Infocomm 2007

Content-oriented network [2/2]

- Content-oriented network
  - A network whose messages are driven by the content of the messages, rather than by explicit addresses assigned by senders
- Basic identifying elements for communication
  - Current Internet
    - Sender Hosts
    - Receive Hosts
  - Content-oriented Network
    - Content of data
    - (Providers publishing the content)
- *"Content-oriented Networking as a Future Internet Infrastructure: Concepts, Strengths, and Application scenarios"*
  - Technical session 7: "Data Oriented Architecture", 6/20

Route-by-name Paradigm

- Resolution between identifier and locator
- Resolution Models
  - Lookup-by-name
    - maintain a distributed database to response a query to find a locator
    - DNS
  - Route-by-name
    - find a locator with name while routing
    - DONA (TRIAD)
- Resolution based on route-by-name approach can be more appropriate for Future Internet

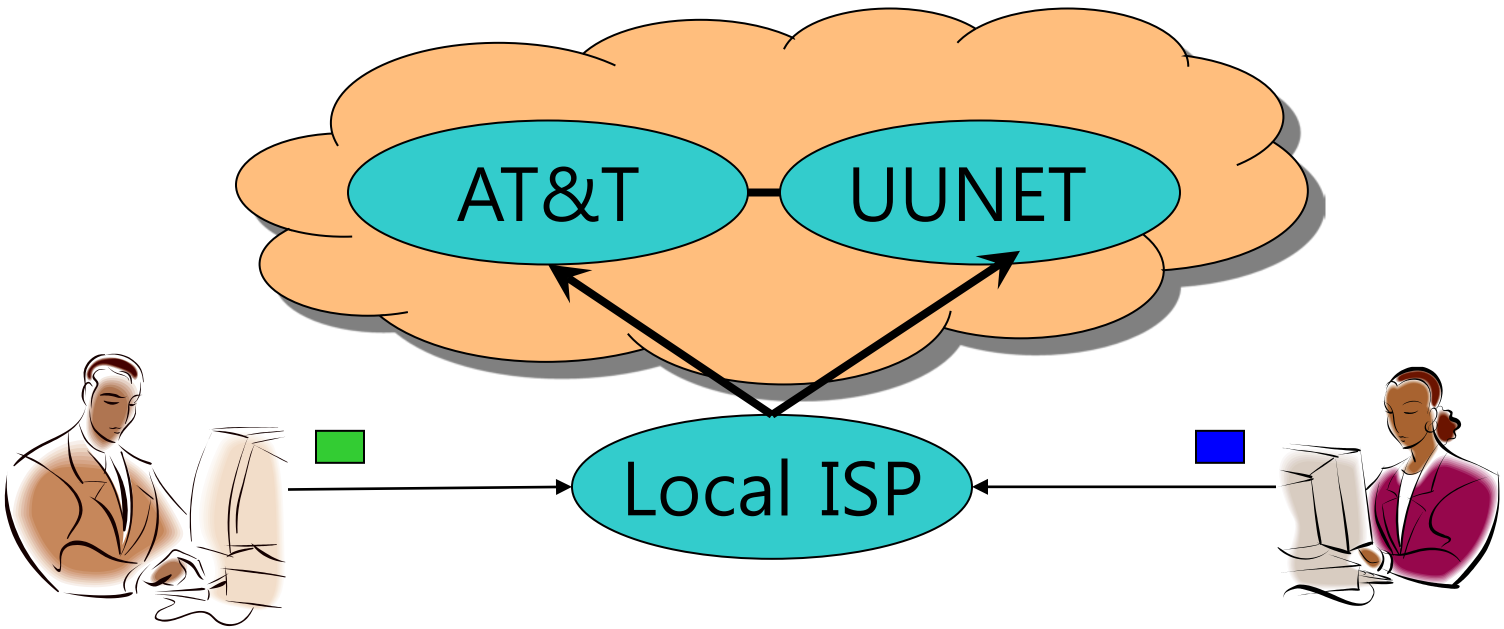


	Lookup-by-name	Route-by-name
Extensibility	Bad	Good
Routing Efficiency	Good	Poor
Robustness	Poor (single POF)	Good

Proposed Solutions

- Scaling Problem
  - Separation address space: GRA for ISP and GDA for end networks
- Usage Pattern Change
  - Host Centric to Data (Service, Content) Oriented
- Other Approaches
  - User Empowerment
  - Routing Management System
  - OpenFlow

We Want to Let Users Choose Domain-Level Routes



- Our hypothesis:
  - User choice stimulates competition.
  - Competition fosters innovation.
- Validation requires market deployment.
- NIRA: the technical foundation.

Central Ideas of NIRA

- Built on earlier ideas of explicit routing, up/down routing.
- Defines efficient representation of explicit route for common case.
  - Assuming today's generally tree-shaped inter-domain topology, with providers and customers
  - "Core" in the center.
- Strict provider-rooted hierarchical addressing

System Components of NIRA

- Addressing
- Route discovery
  - Topology Information Propagation Protocol (TIPP)
  - A user learns his addresses and topology information (static) and perhaps route availability (dynamic)
- Name-to-Route mapping
  - Name-to-Route Lookup Service (NRLS) – an enhanced DNS service
  - A user learns destination's addresses and optional topology information.
  - Combining information from TIPP and NRLS, a user is able to select an initial route.

Routing Mgmt System
by Roch Guerlin (Univ. of Pennsylvania)

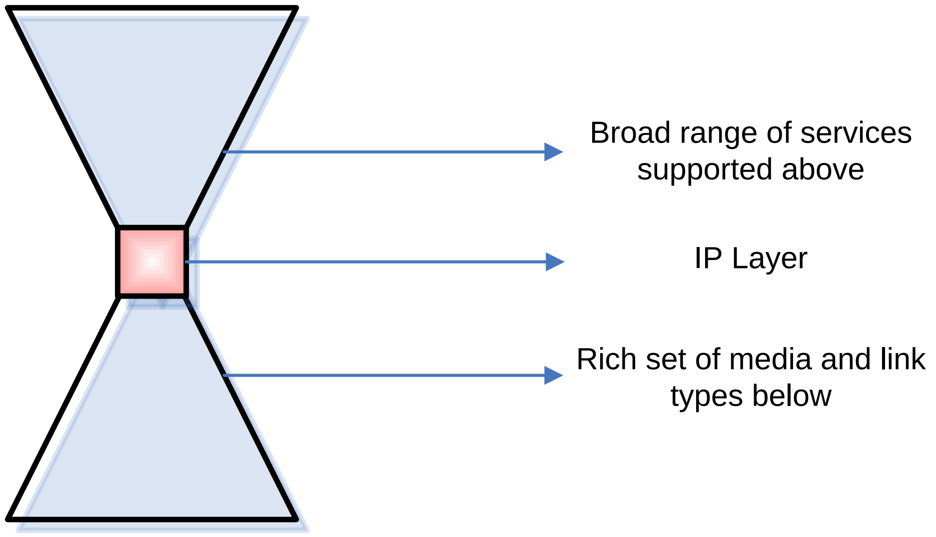
- Introduction

- Why we need Manageability in RS?
  - Manageability Challenges
  - Key Tenets of Manageability

- A Strawman Proposal for an Architectural Framework

Introduction

- Current Internet successful “hourglass” design choice



- Distributed routing decision making process

Introduction

- Recently, the old structure has some problems:
  - More advanced services are being deployed
  - Best-effort service may not sufficient for real-time apps.
  - Distributed decision making process → difficult to detect, pinpoint and fixing routing problems.
  - …
- More centralized solution are being advocated

Introduction

- In this work, we will:
  - Develop a generic framework for specifying details that should be present in design of any management solution for routing systems (RS).
  - Center around a number of specific problems associated with both existing and new routing systems.
  - Our proposed framework will be refined and validated using the GENI facilities.

Why we need Manageability in RS?

- Manageability features:
  - Configuration
  - Benchmark and Trending
  - Problem Detection
  - Analysis and Diagnosis

  } Most important features

- Our goal: how manageability can be successfully incorporated into RSs

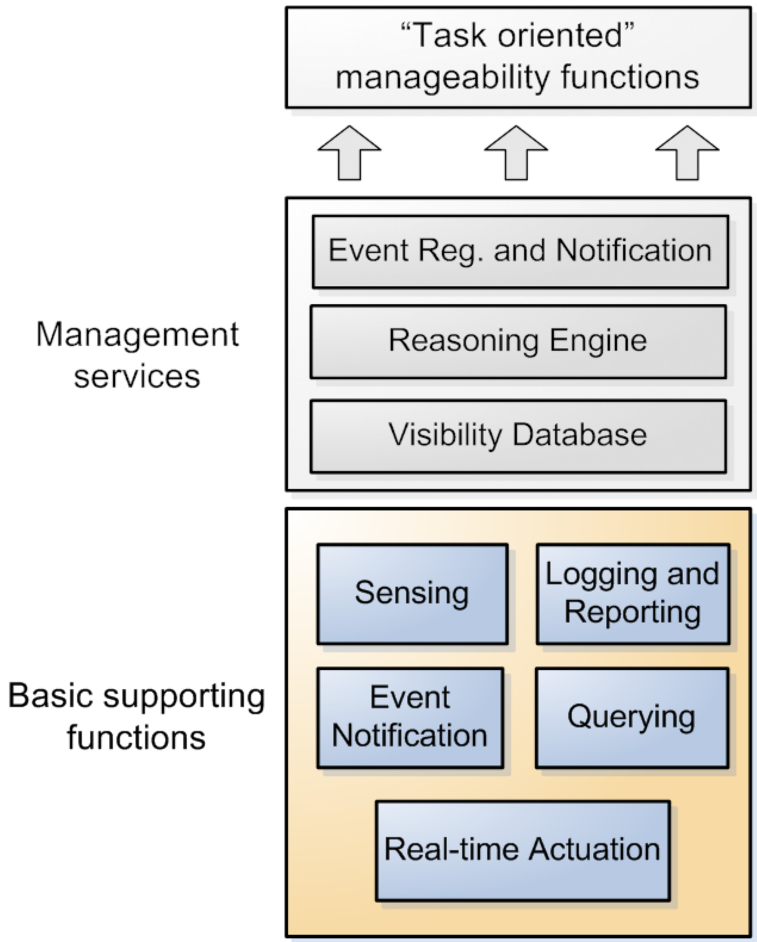
Why we need Manageability in RS?

- We focus on two critical dimensions:
  - *Horizontal*: understanding how distributing the decision process that controls routing decisions affects its manageability.
  - *Vertical*: keep in mind that RSs do not operate in isolation (but depends on multiple components or layers)

Key Tenets of Manageability

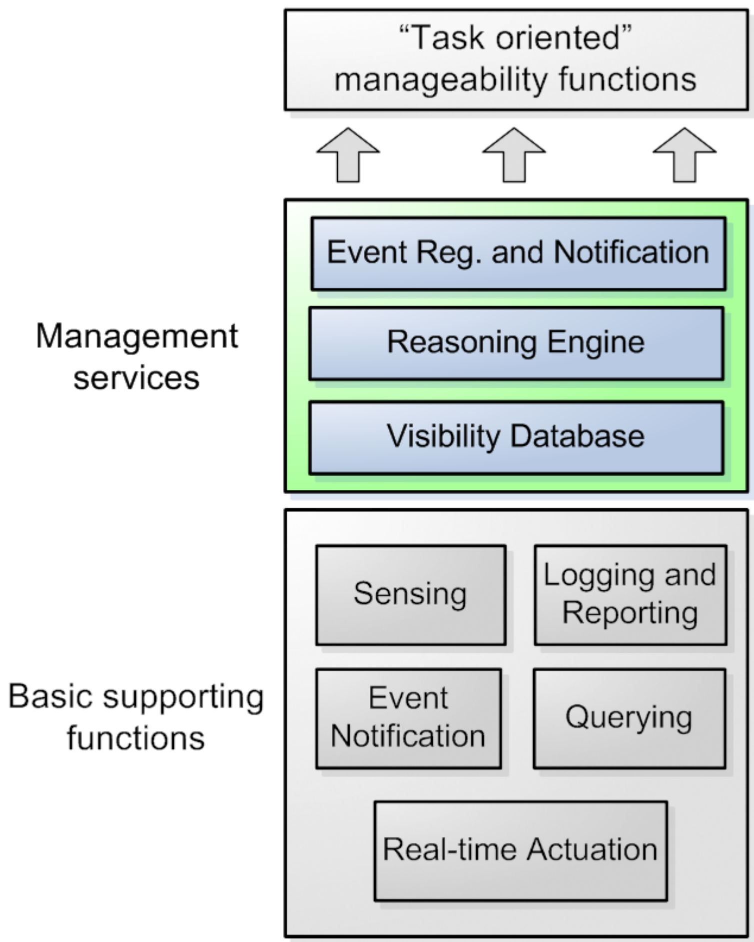
- Visibility
  - Ability to obtain information about routing state and knowledge of the routing decision making processes.
- Reasonability
  - Ability to analyze and reason about routing behaviors based on collected routing state information.
- Actionability
  - Ability to identify necessary changes in routing configuration, resources and operations.

Proposal for an Architectural Framework



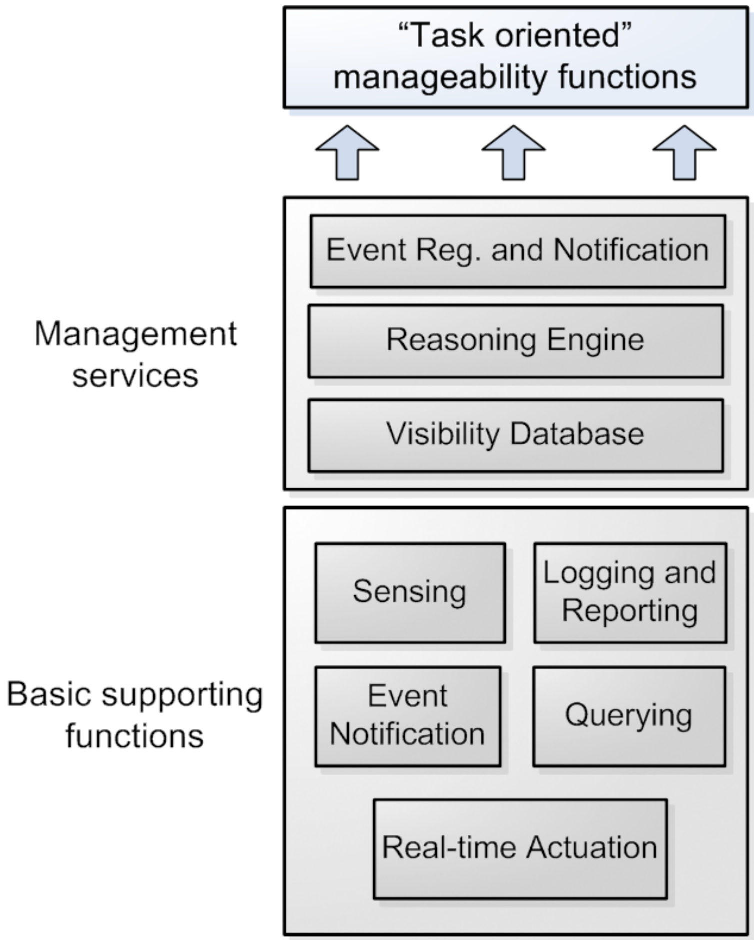
- Sensing: Monitor & detect changes in the network state
- Logging and Reporting: locally collect and record visibility information
- Event Notification: receiving report/notification regarding certain changes in network state
- Querying: Query a routing element for its information
- Real-time Actuation: Allow other entities to ask a routing element to execute certain actions

Proposal for an Architectural Framework



- Visibility Database:
  - Centralized repository for storing data collected from routing elements
- Reasoning Engine:
  - Consist of a set of tools and algorithms for analyzing network data and performing management functions
- Event Registration and Notification:
  - Provides network-wide service for entities and users to register and be notified of events of interest.

Proposal for an Architectural Framework



- "Task oriented" network-wide manageability support functions (within single network domain or across network domains)

Research Problems and Approaches

1. Modeling Routing Systems as Rule Systems
2. Manageable Distributed Computation Based Routing Protocol
3. Building Domain-Wide Integrated Management Systems
4. Building Network-Wide Management Services

OpenFlow

- Presented by Seungjun Seok at KRNet on 2010 June.

Innovations in Legacy Internet

- Problem with our network
  - Paths are fixed (by the network)
  - IP-only
  - Addresses dictated by DNS, DHCP, etc
  - No means to add our own processing
  - …

- Experiments we'd like to do new
  - Mobility management
  - Network-wide energy management
  - New naming/addressing schemes
  - Network access control
  - …

Future Internet

- Approaches to Future Internet
  - Clean-State Redesigns of Internet
  - Evolution of Internet
- Many Ideas and Technologies are being developed now
  - New Architectures
  - Protocols
  - New Wireless/Mobility
  - New Infrastructure/Switching
  - New Media/Applications

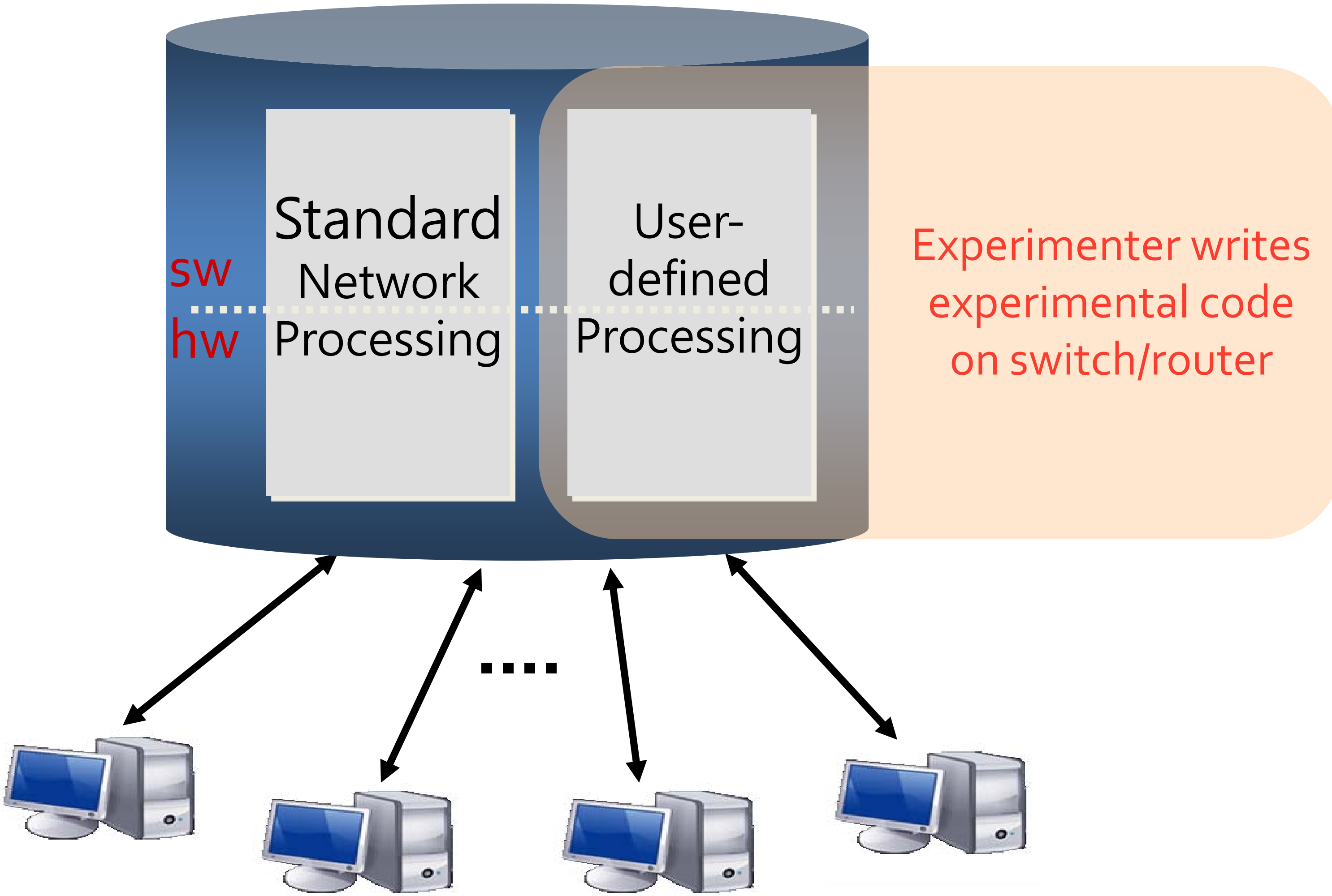
Why Internet Closed for Innovations?

- Commercial Vendor won’t open software and hardware development environment
  - Complexity of support
  - Market protection and barrier to entry

- Hard to build my own
  - Prototypes are unstable
  - Software only contribution is Too Slow
  - Hardware/software: Fanout too small (need > 100 ports for wiring closet)

[The Stanford Clean Slate Program]

Experimenter's Dream
(Vendor's Nightmare)



[The Stanford Clean Slate Program]

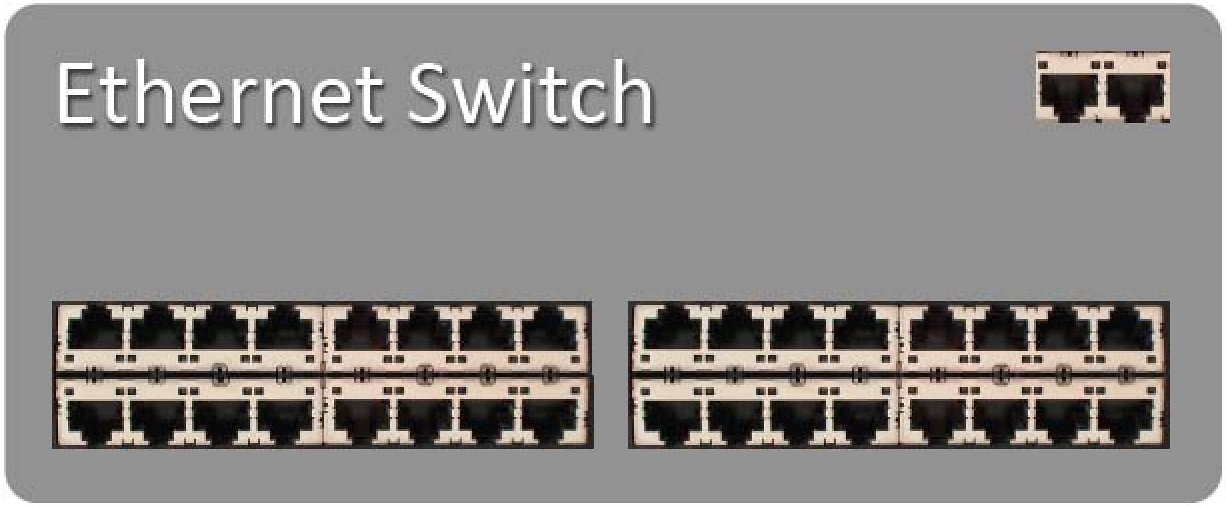
OpenFlow's Goal

- Put an open platform in hands of researchers/students to test new ideas at scale through production networks.
  - without requiring vendors to expose internal workings

- Bring Future Internet to legacy Internet
- An open development environment for all researchers (e.g. Linux, Verilog, etc)

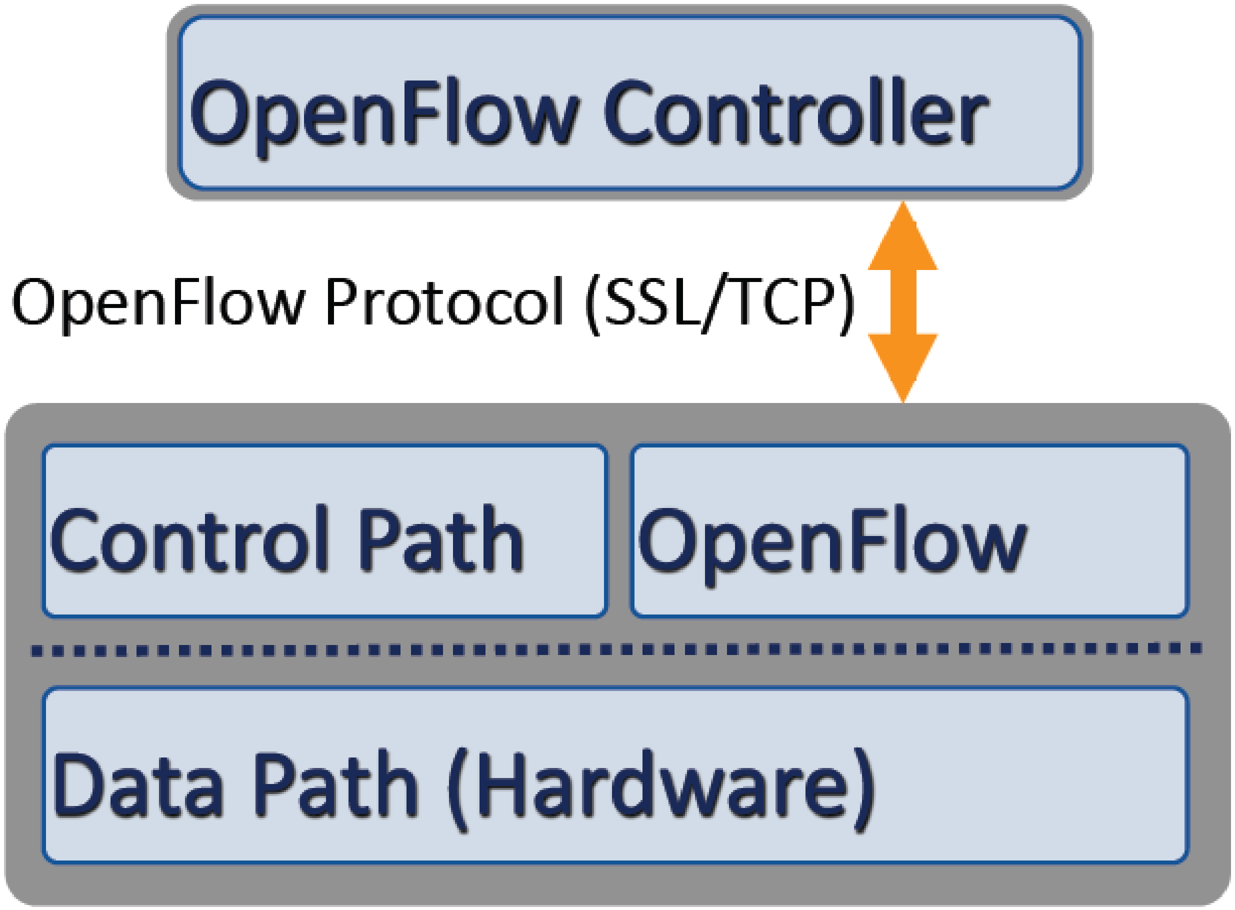
[The Stanford Clean Slate Program]

OpenFlow Concept

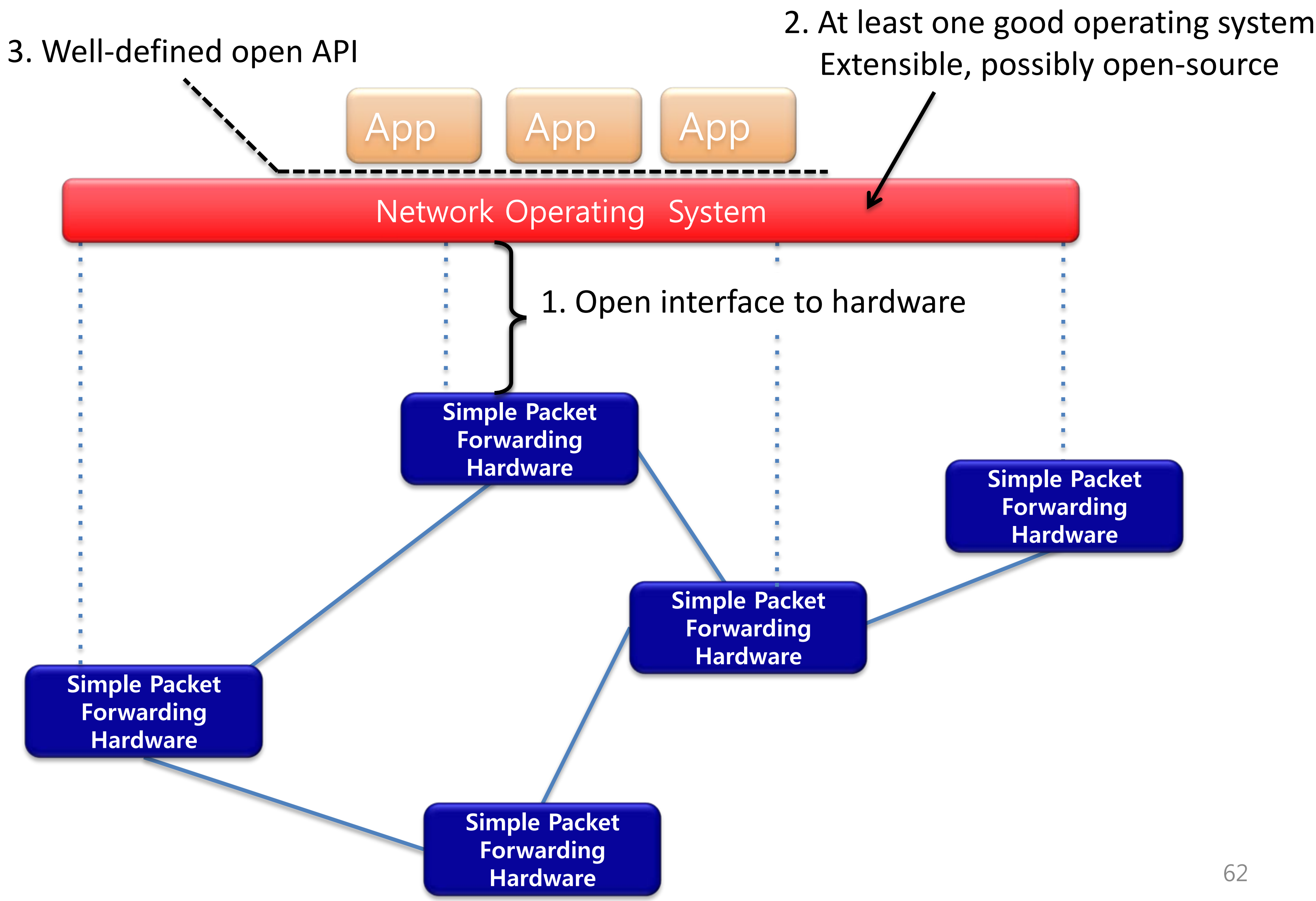


OpenFlow Concept

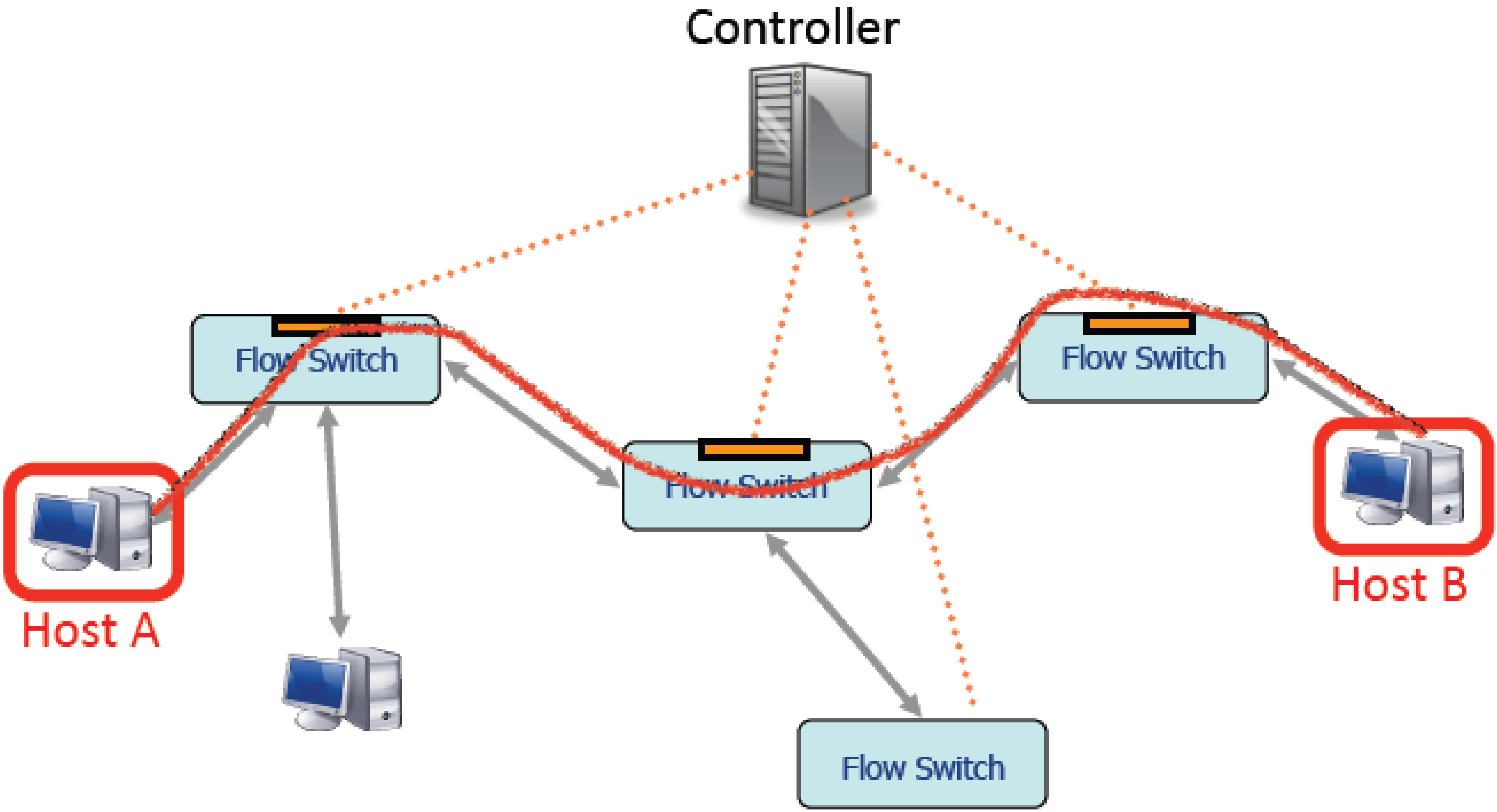
OpenFlow Concept



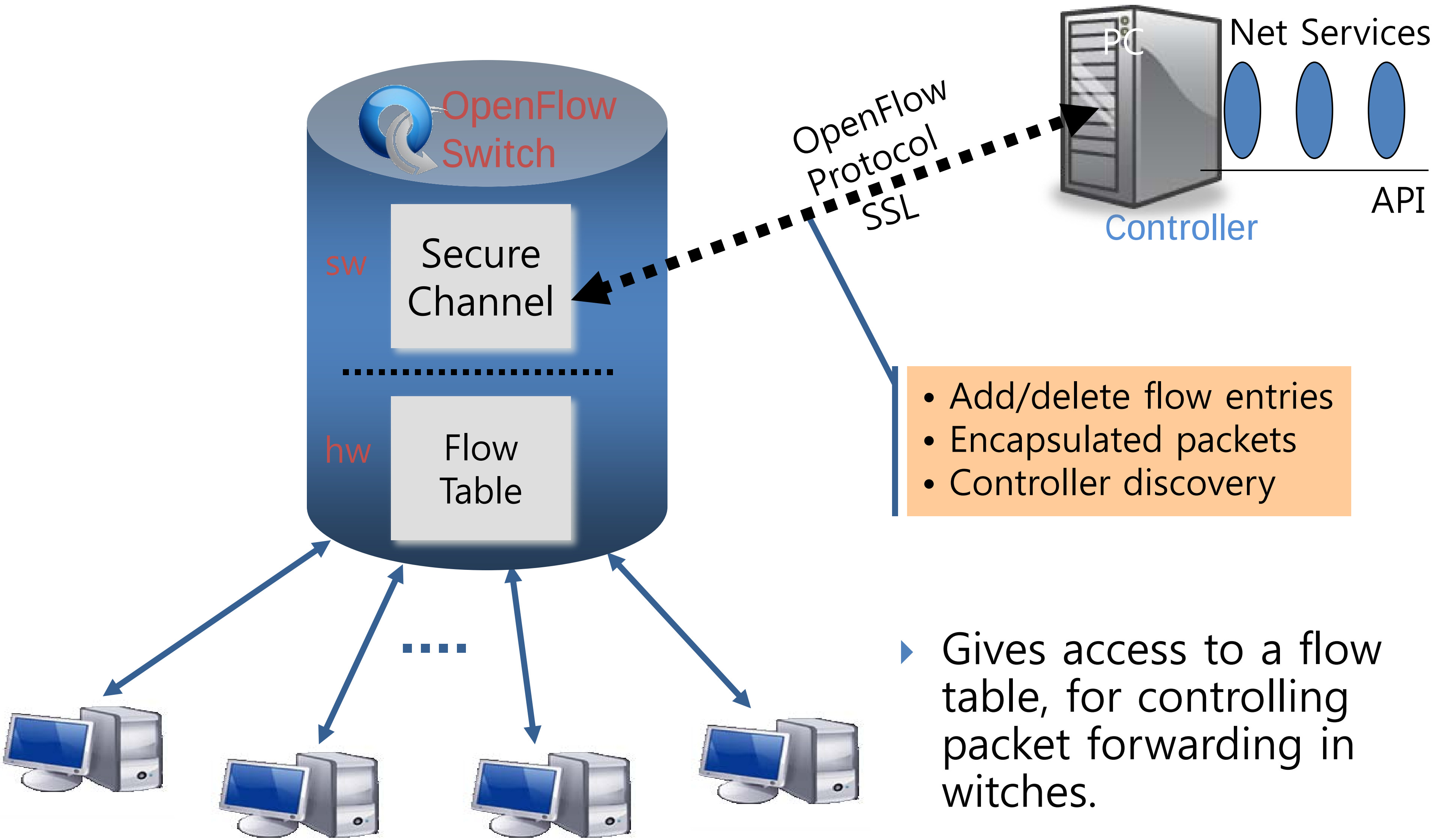
Architecturally what It Means



Network Configuration



OpenFlow Network Architecture



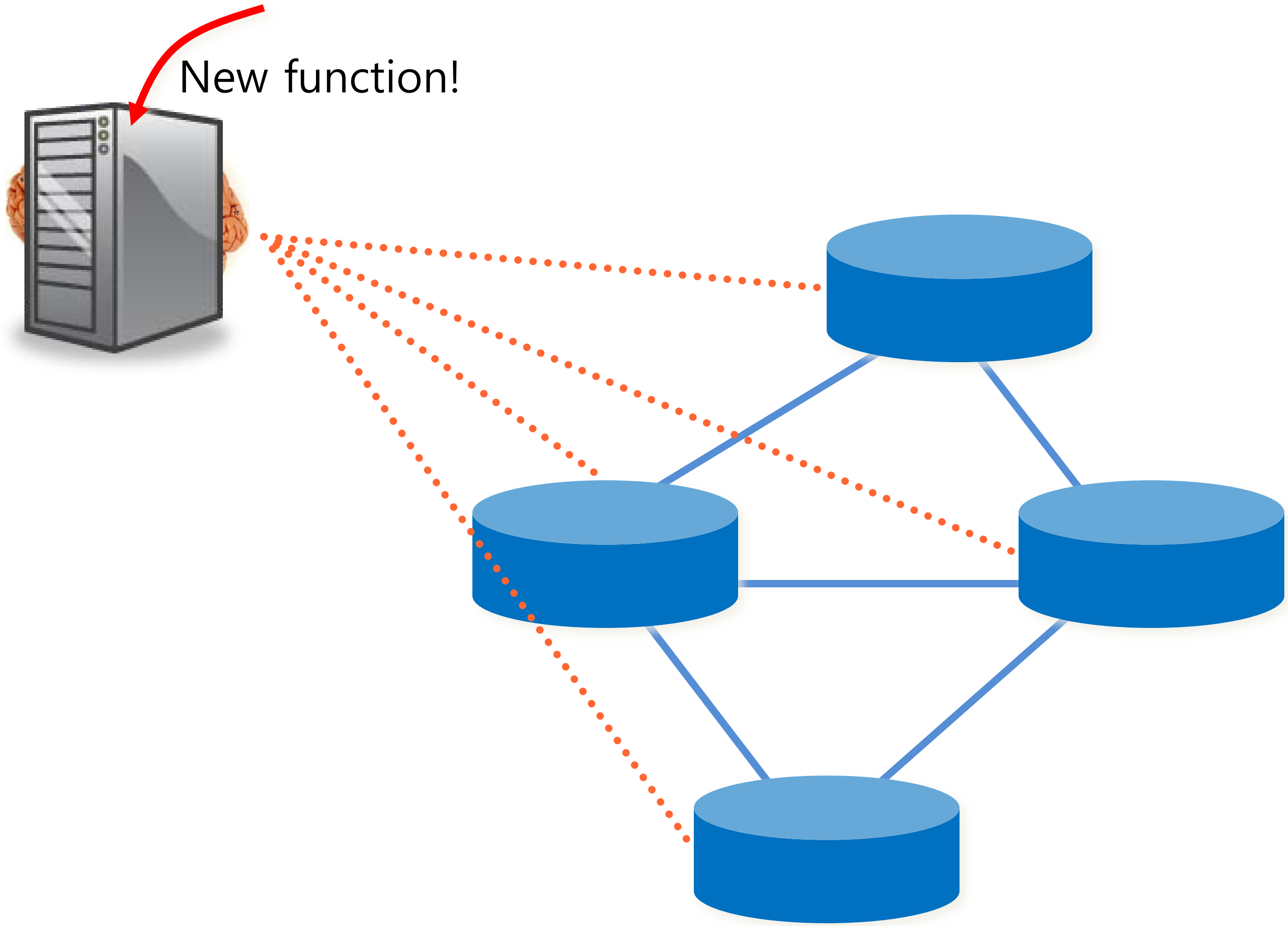
- Gives access to a flow table, for controlling packet forwarding in witches.

Example Network Services

- Static "VLANs"
- New routing protocol: unicast, multicast, multipath, load-balancing
- Network access control
- Mobile VM management
- Mobility and handoff management
- Energy management
- Packet processor (in controller)
- IPvX
- Network measurement and visualization
- …

Operation Step1: Separating Intelligence from Datapath

Operators, users, 3rd party developers, researchers, …

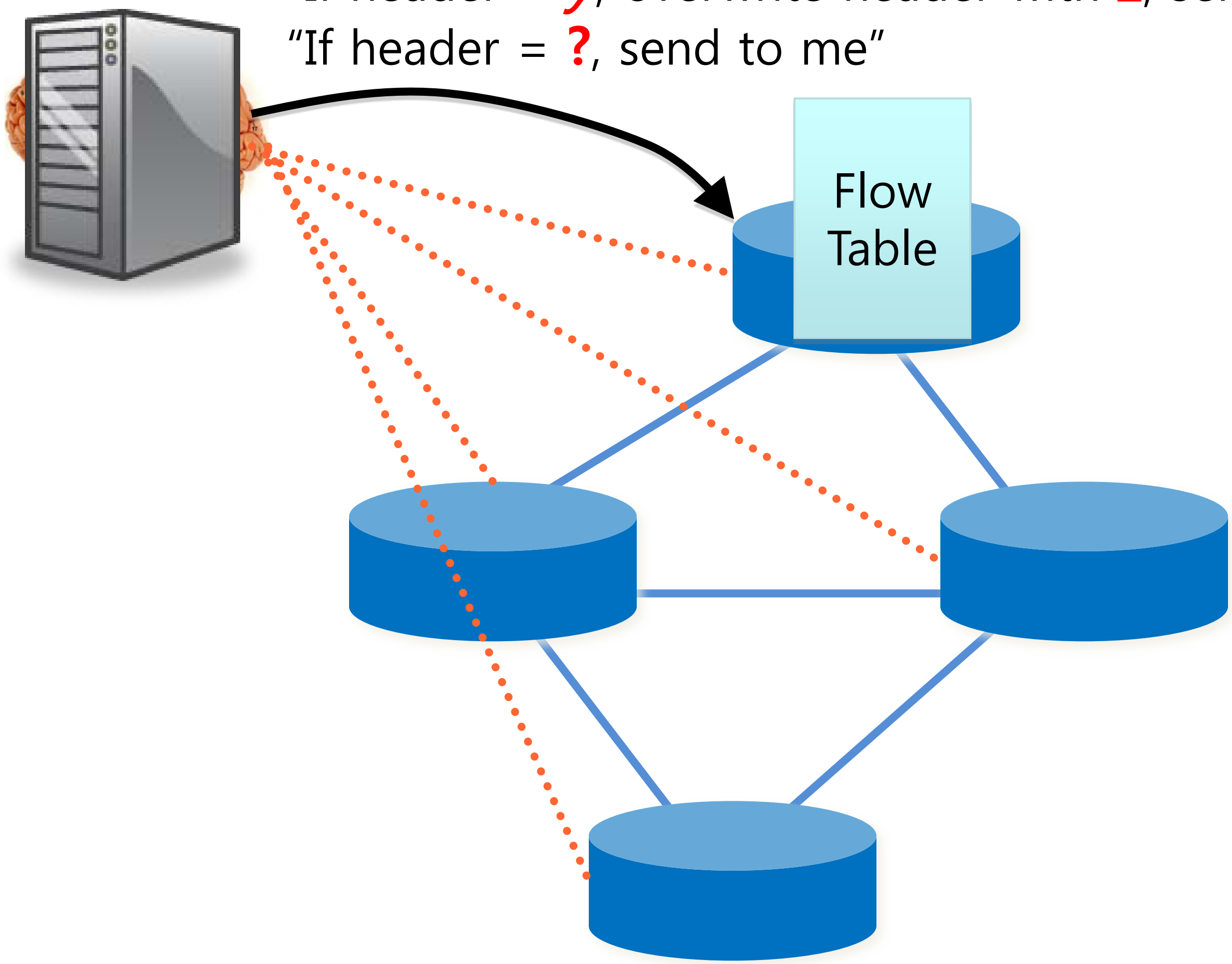


Operation Step2: Cache Decisions in Flow-based Datapath

"If header = *x*, send to port 4"

"If header = *y*, overwrite header with *z*, send to ports 5,6"

"If header = **?**, send to me"



Flow Table Structure

- Exploit the flow table in switches, routers, and chipsets

Flow 1.	Rule (exact & wildcard)	Action	Statistics
Flow 2.	Rule (exact & wildcard)	Action	Statistics
Flow 3.	Rule (exact & wildcard)	Action	Statistics
		……	
Flow N.	Rule (exact & wildcard)	Default Action	Statistics

Flow Table Entry

Rule	Action	Stats

Packet + byte counters

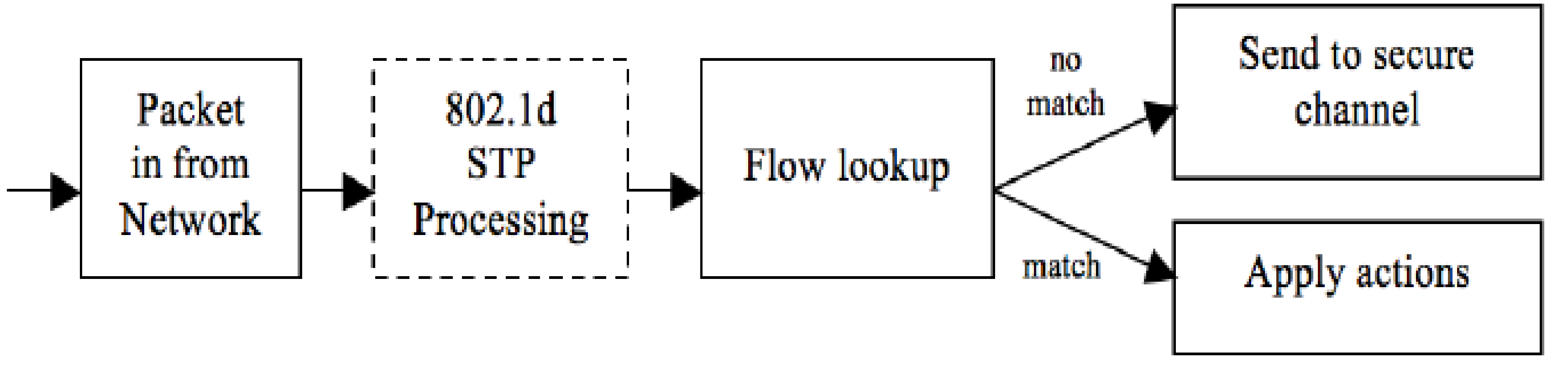
1. Forward packet to port(s)
2. Encapsulate and forward to controller
3. Drop packet
4. Send to normal processing pipeline

Switch Port	MAC src	MAC dst	Eth type	VLAN ID	IP Src	IP Dst	IP Prot	sport num	dport num

+ mask what fields to match

Packet Processing

- OpenFlow SW's Packet Processing
  - Search a matched entry of flow table with arriving packet's information



References

- A Framework for manageability in Future Routing System.pdf by Lixin Gao of Univ. of Mass
- Daniel Massey, A scalable routing system design for FI, IPv6'07, August 31, 2007, Kyoto, Japan.
- Jaeyoung Choi, Addressing in Future Internet: Problems, Issues, and Approaches
- Theus Hossmann, Implementing the Future Internet: A Case Study of Service Discovery using Pub/Sub in the ANA Framework
- Routing in a FI Architecture by Bob Braden, USC/ISI
- The Role of IP Address in the Internet Architecture by Lixia Zhang, UCLA at Asia Future Internet Summer School August 2008