Old Wine in New Bottles

Key Security Technologies in Data Center Networks for Cloud Computing

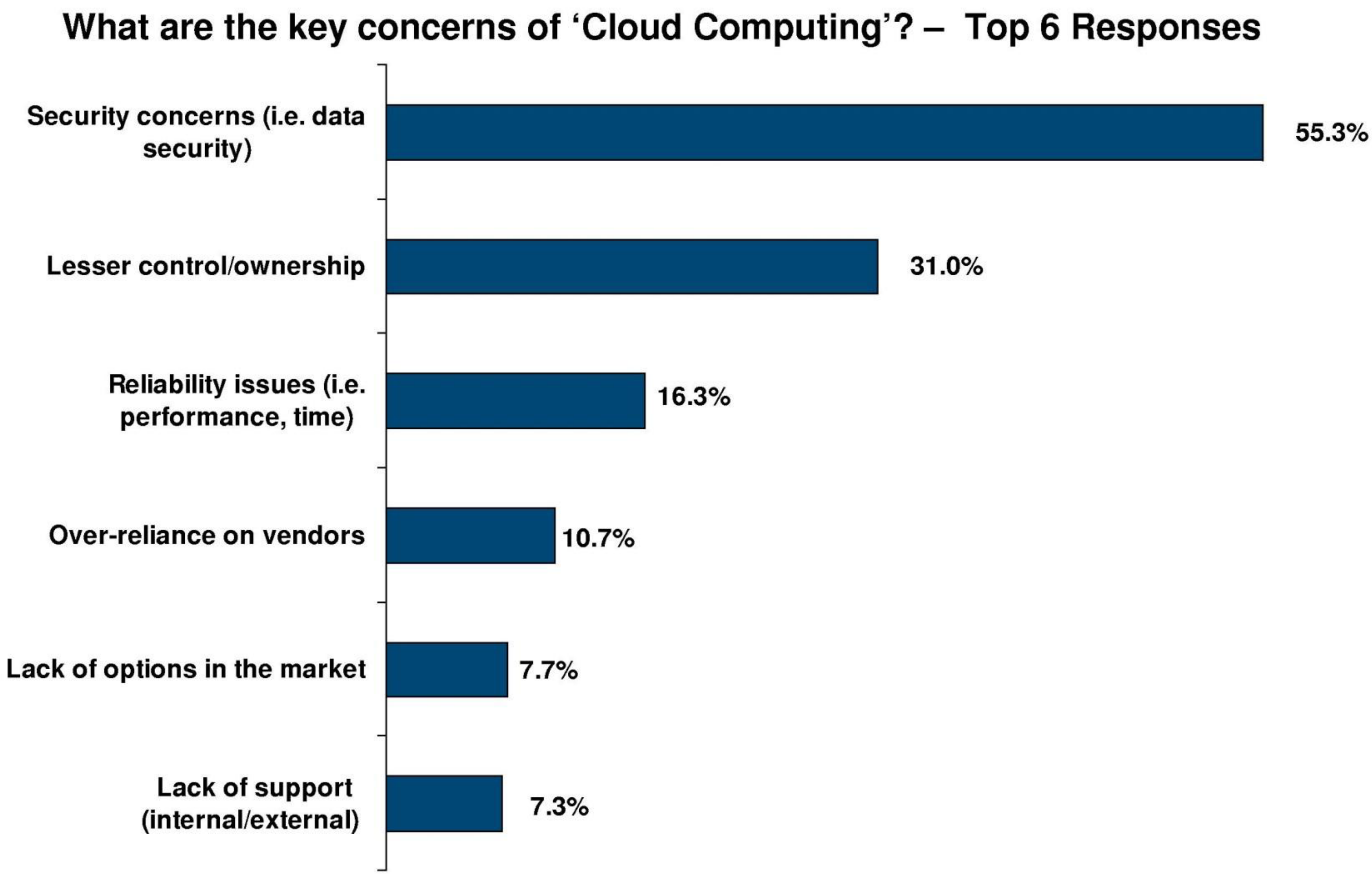
Jun Li
Tsinghua University

With many help from my students Yaxuan Qi, Fei He, Baohua Yang

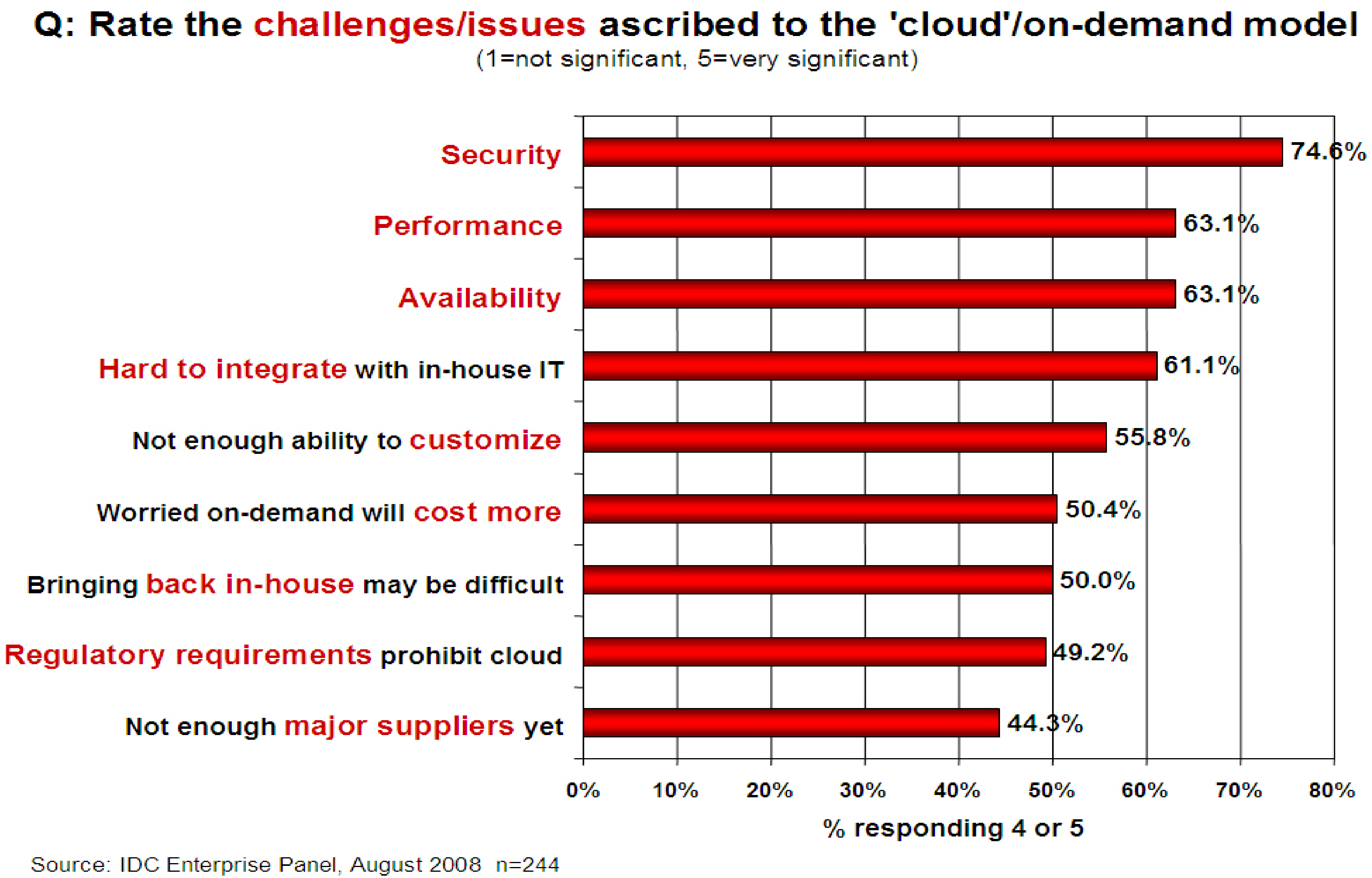
Outline

- Cloud Labels the “New Bottles”
  - Security Concerns of Cloud
  - DCN is Key to Cloud Performance
  - Topology and Algorithm Matters
- Cloud DCN inside the New Bottle
- “Old Wine” with a New Taste

Frost & Sullivan's Survey



IDC's Survey


Q: Rate the challenges/issues ascribed to the 'cloud'/on-demand model
(1=not significant, 5=very significant)
Security
Performance
Availability
Hard to integrate with in-house IT
Not enough ability to customize
Worried on-demand will cost more
Bringing back in-house may be difficult
Regulatory requirements prohibit cloud
Not enough major suppliers yet
74.6%
63.1%
63.1%
61.1%
55.8%
50.4%
50.0%
49.2%
44.3%
0%
10%
20%
30%
40%
50%
60%
70%
80%
% responding 4 or 5
Source: IDC Enterprise Panel, August 2008 n=244

Ref. [1], etc.

Gov Going Cloud with Reservations

- Governments are using cloud computing to reduce IT costs and increase capabilities
  - US government GSA (General Services Administration) now offers a portal for cloud computing services
  - By 2014, over $1 billion of the US federal IT budget would be devoted to cloud computing
- Governments have serious hurdles to overcome
  - Public perception of the secure processing of citizens' personal information in cloud computing infrastructures
  - Legal and regulatory obstacles which prevent many eGov applications from moving to cloud

The Facts about Cloud DCN

- Cloud Data Centers are the big “new bottles”

- Typical data center lifecycle is 10 to 15 years
- Data center instance: Costs in billion range with >100K servers
- By 2010, >15M of servers installed in US, and >$45B to power and cool servers
- 10GE (40, 100 GE later) fiber from servers to ToR, and then to switch fabric
- Traffic volume between servers to entering/leaving data center is 4:1
- Majority of flows are small and each machine has <10 flows in >50% of time (only <5% time with >80 flows)

Ref. [2]

Network Perspective of Cloud Computing

Performance(cloud computing) = ∑{performance(network connections)} + ∑{performance(IT resources)}

Cloud networking is centered around cloud data centers:

- Resource networking inside DCN
  - **Network of resources**: servers and storages
  - Connects the IT elements together to create data centers
- Access networking to DCN
  - **Network of clients**: direct and pervasive connect
  - Allows users to access the applications running in those data centers
- Federation networking between DCN
  - **Network of clouds**: private and public clouds

Most of challenges are within Data Center Networking (DCN)

Impacts of DCN Virtualization

- Topology-independent Service Assignment
  - Decouple service assignment in DC from network topology to support evolutional service deployment
- Location-independent Server Addressing
  - Decouple server's location in DC from its address to enable seamless server migration

- Topology-independent Policy Enforcement
  - Decouple security policy in DC from network topology by flow-based redirection to provide borderless security
- Location-independent Private Access
  - Decouple confidential access in DC from physical clustering by Virtual Private Network (VPN) to create virtual private cloud (VPC)

The Two “L”s in Cloud Networking

- The two specific variables likely to determine success in data center networking are the "two 'L's"
- Loss
  - All network protocols have to protect against data loss through retransmission of corrupted information as it takes time, and loss of an information packet is particularly critical with storage protocols because of the risk of creating a corrupted file or leaving a storage device in a bad operating state
- Latency
  - Latency is a special problem in data center and storage networks because it accumulates quickly across the tens of millions of operations involved

Performance Focuses for DCN

- Topology: Flat and Reliable
  - Latency accumulates in networks largely in proportion to the number of interfaces a packet transits, and each switch that handles packets poses a risk of loss, in addition to contributing to the total delay → It's better to reduce the number of intermediate interfaces, and that means reducing the number of switches
  - A few very large switches will provide better performance than several layers of smaller ones, but concentrating switching into a few devices could increase failure risk too → It's important for the switches to have the highest possible MTBF and also that the components be redundant and support automatic failover
- Algorithm: Fast and Scalable
  - Flow identification, distribution, and screening is now employed almost everywhere, from TOR to fabric, and from generic processor to dedicated chips → Acceleration is required at algorithmic level to adapt the evolution of hardware platforms and the trend of flow-based switching

DCN Key Security Technologies

- It all boils down to the basic technologies, and at the core of these innovations is flow-based network processing
- Foremost, DCN security topology change leads to new architecture
  - Ethane (SIGCOMM 07)
  - PSwitch (SIGCOMM 08)
  - DIFANE (SIGCOMM 10)
- At the same time, DCN security demands higher performance
  - Packet Classification
  - Pattern Matching
  - Flow Scheduling and Traffic Management

Outline

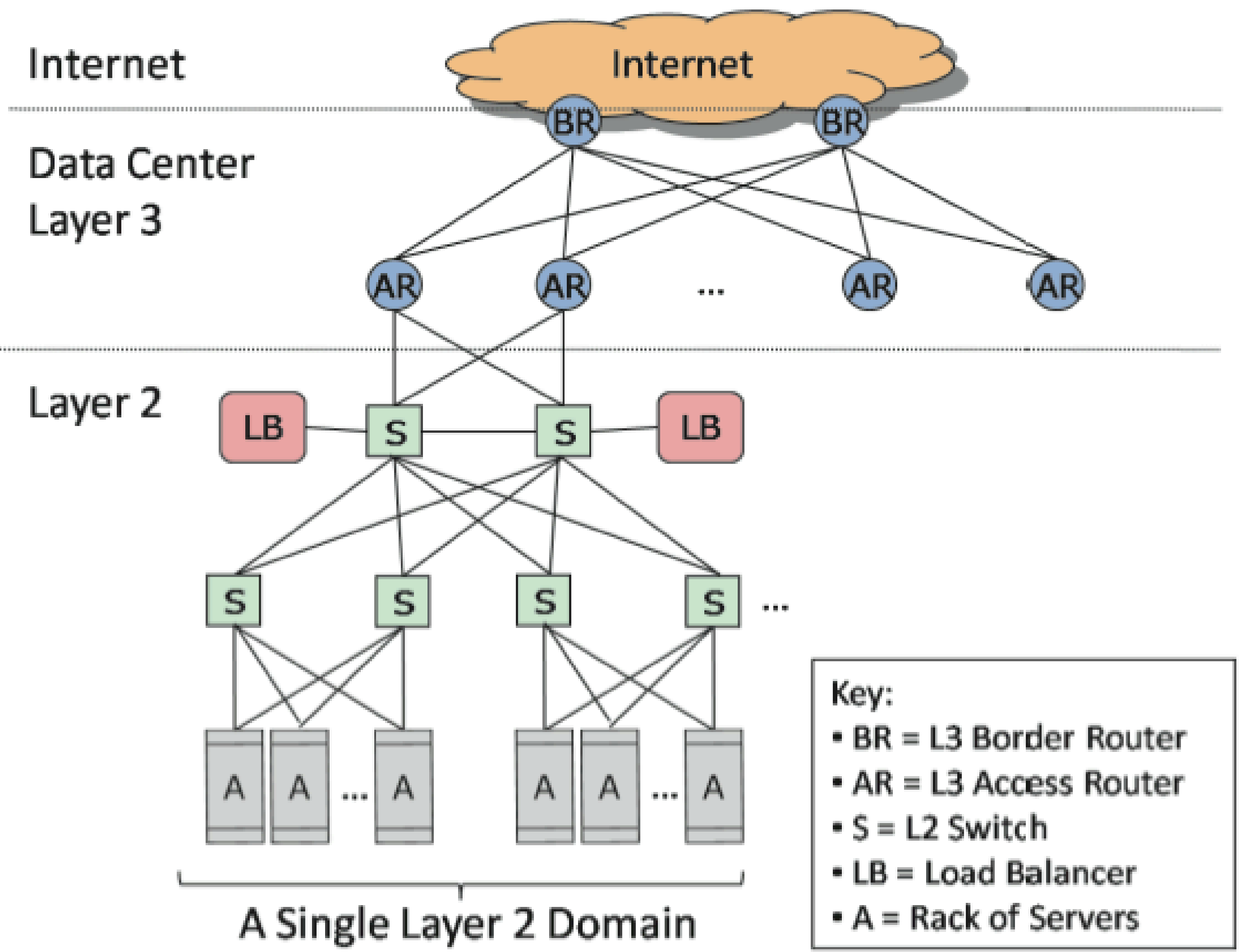
- Cloud Labels the "New Bottles"
- Cloud DCN inside the New Bottle
  - DCN Architecture is Transforming
  - DCN Security's Topology Changed
  - Ethane, PSwitch, and DIFANE
- "Old Wine" with a New Taste

Conventional DCN Architecture

Core

Aggregation

Access

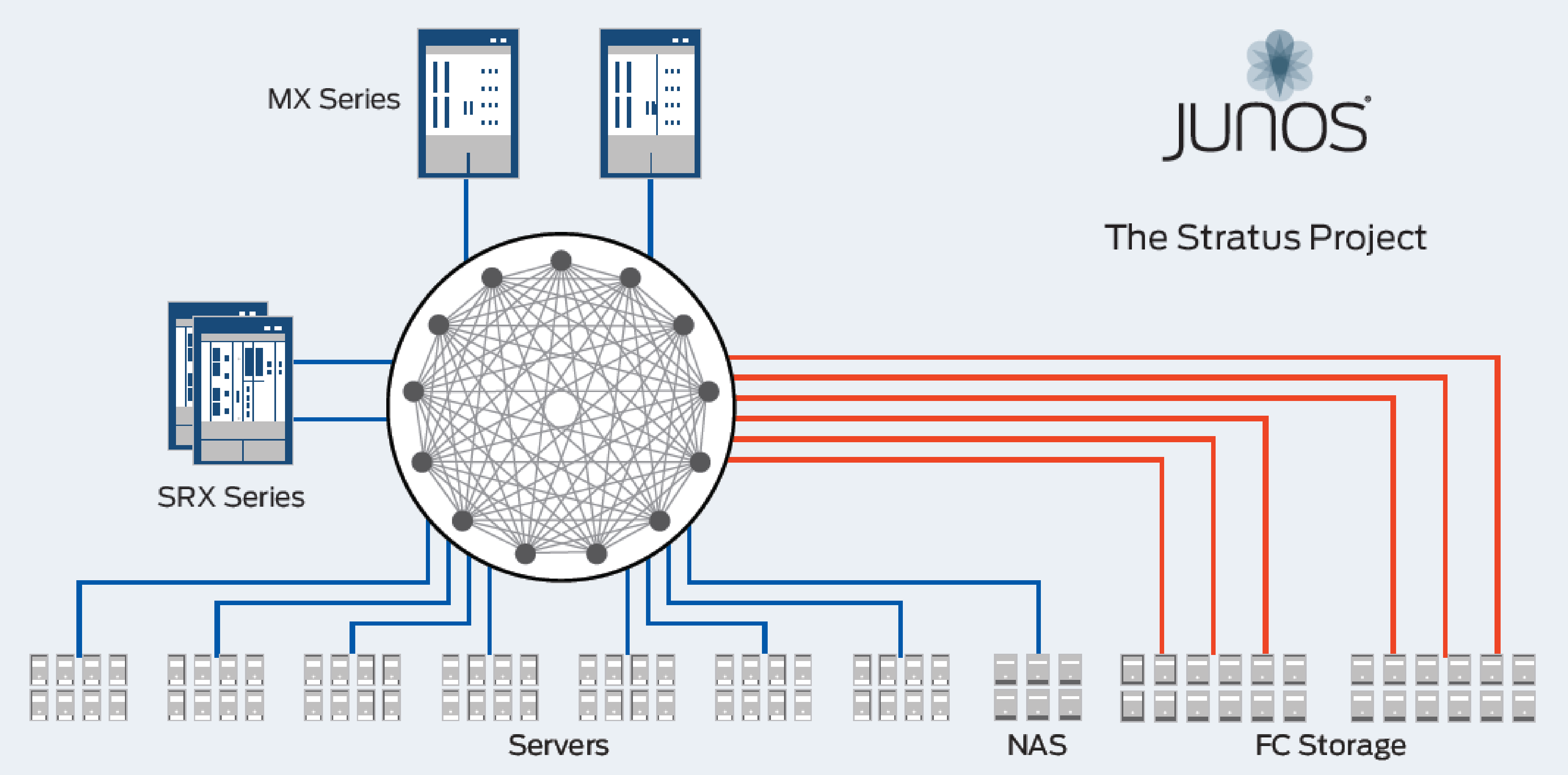


Courtesy of MSR

Ref. [3]

Juniper's Stratus

- Consolidate siloed systems and collapse inefficient tiers into a single fabric for any-to-any connectivity, which is flat, lossless, and delivers very low latency



Ref. [4]

Ethane: Addressing the Protection Problem in Enterprise Networks

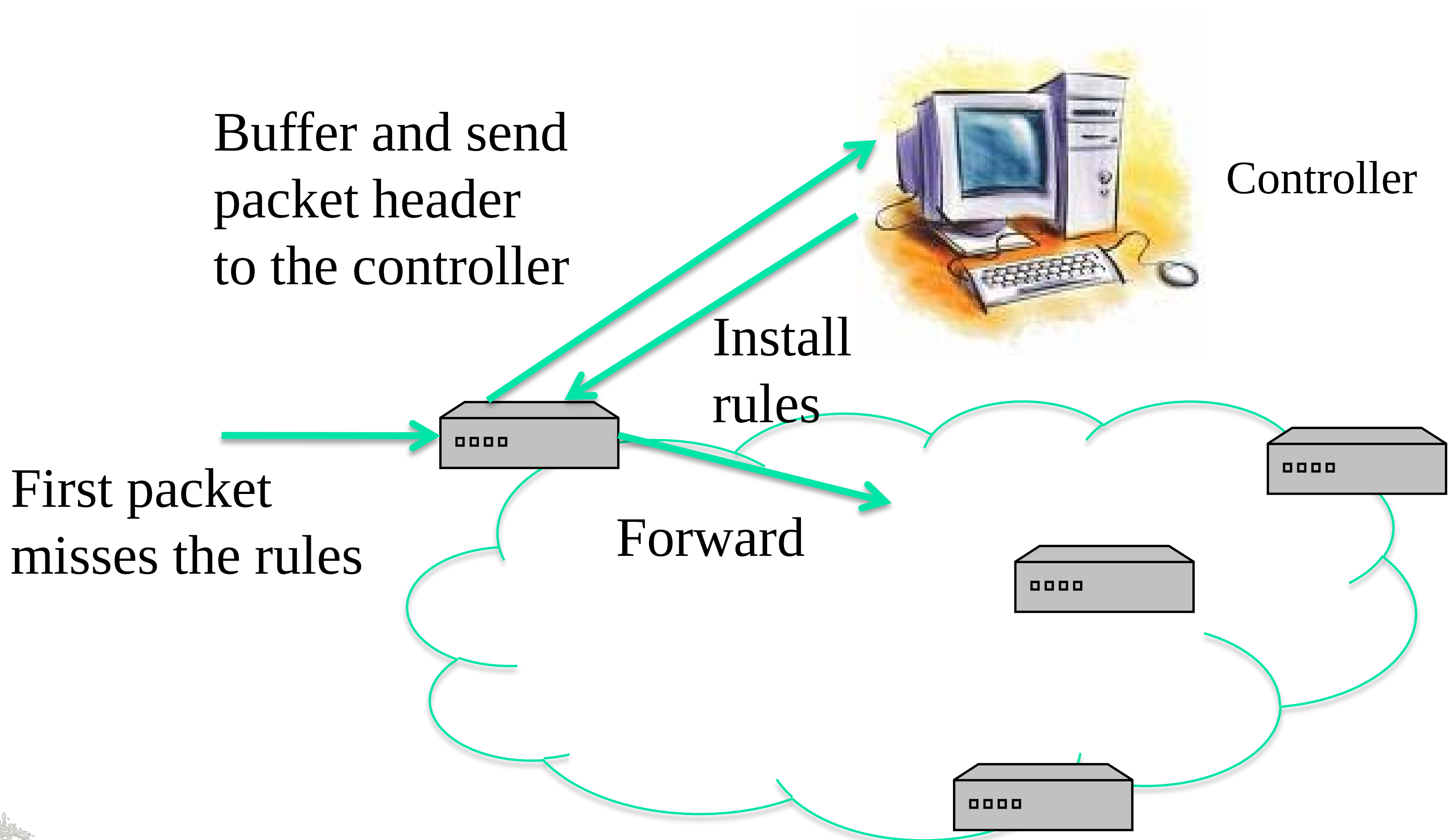
Martin Casado
Michael Freedman
Glen Gibb
Lew Glendenning
Dan Boneh
Nick McKeown
Scott Shenker
Gregory Watson

Presented By: Martin Casado
PhD Student in Computer Science,
Stanford University

casado@cs.stanford.edu
http://www.stanford.edu/~casado

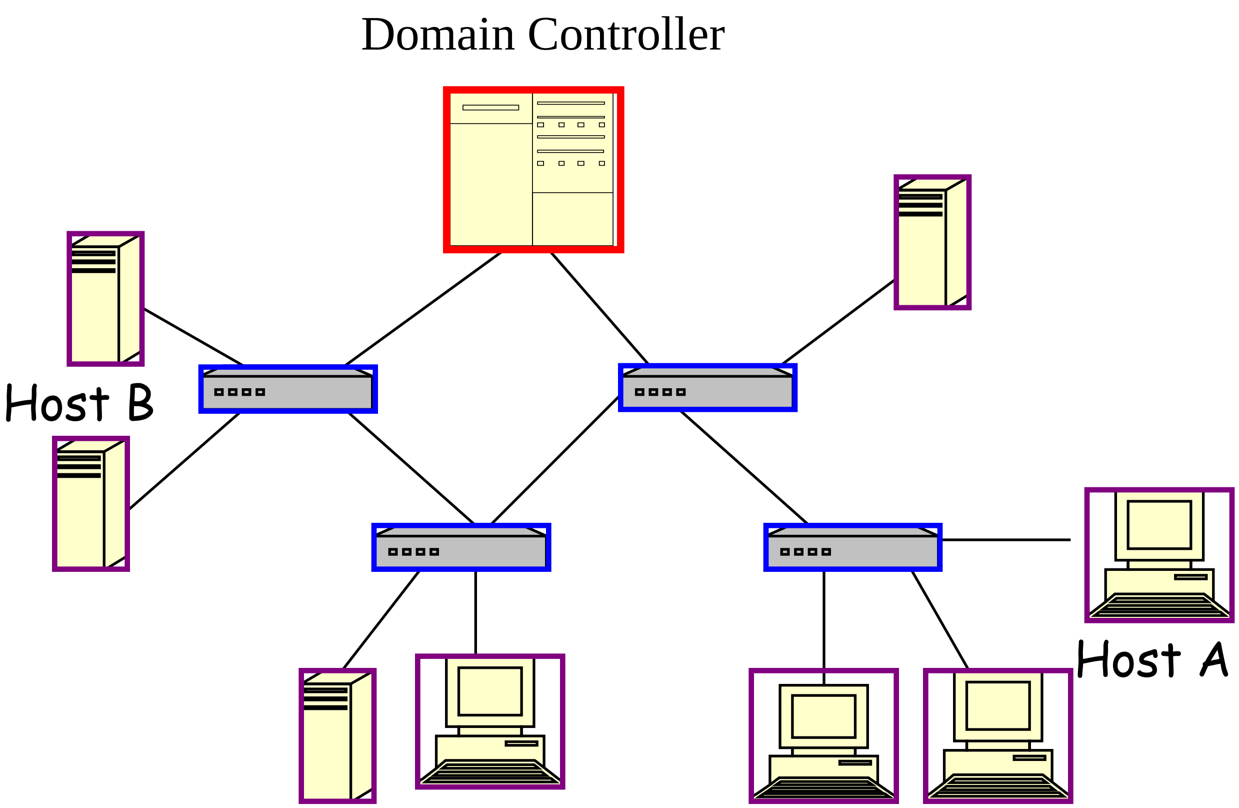
Ref. [6]

Install Rules on Demand



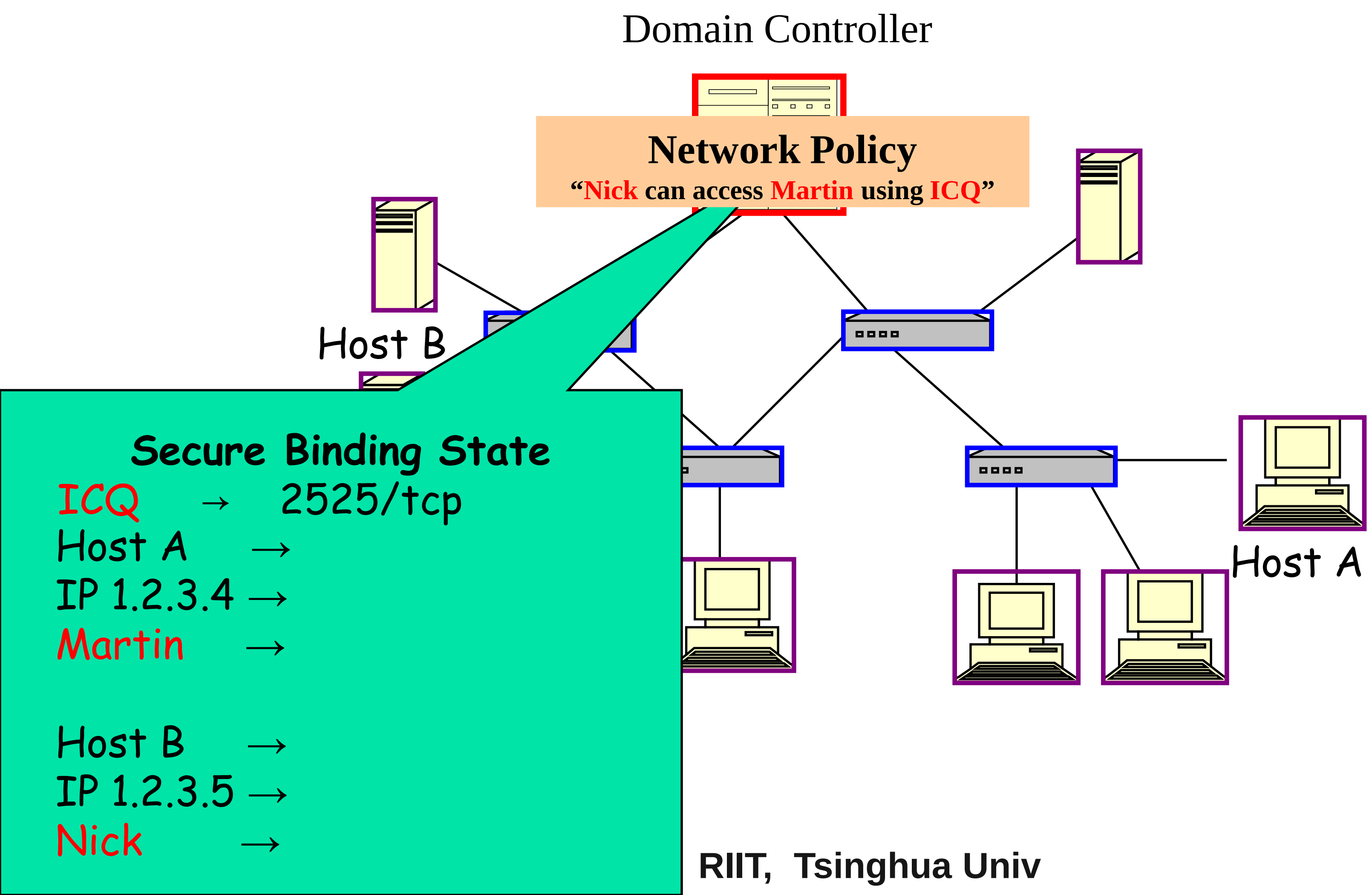
Ref. [4]

Ethane: High-Level Operation



Ref. [4]

Ethane: High-Level Operation

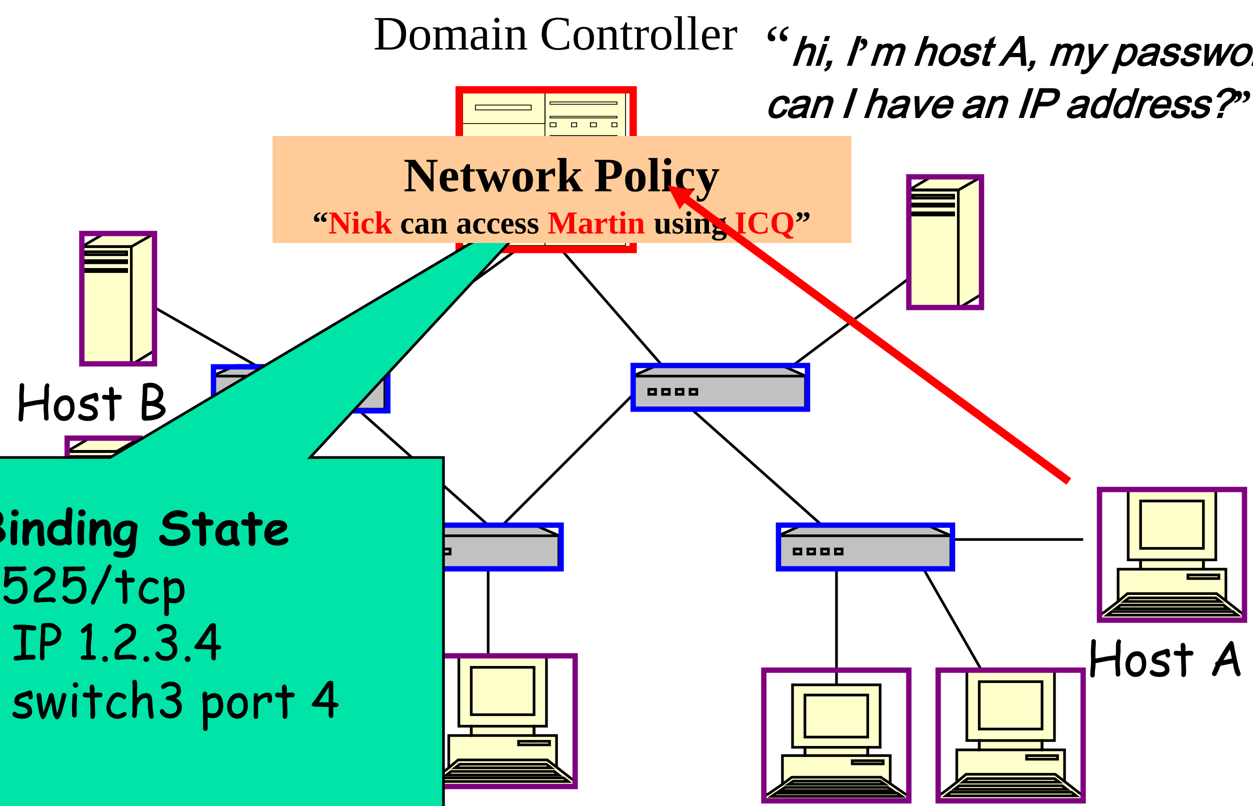


Ref. [4]

Ethane: High-Level Operation

Host Authentication

“hi, I’m host A, my password is … can I have an IP address?”



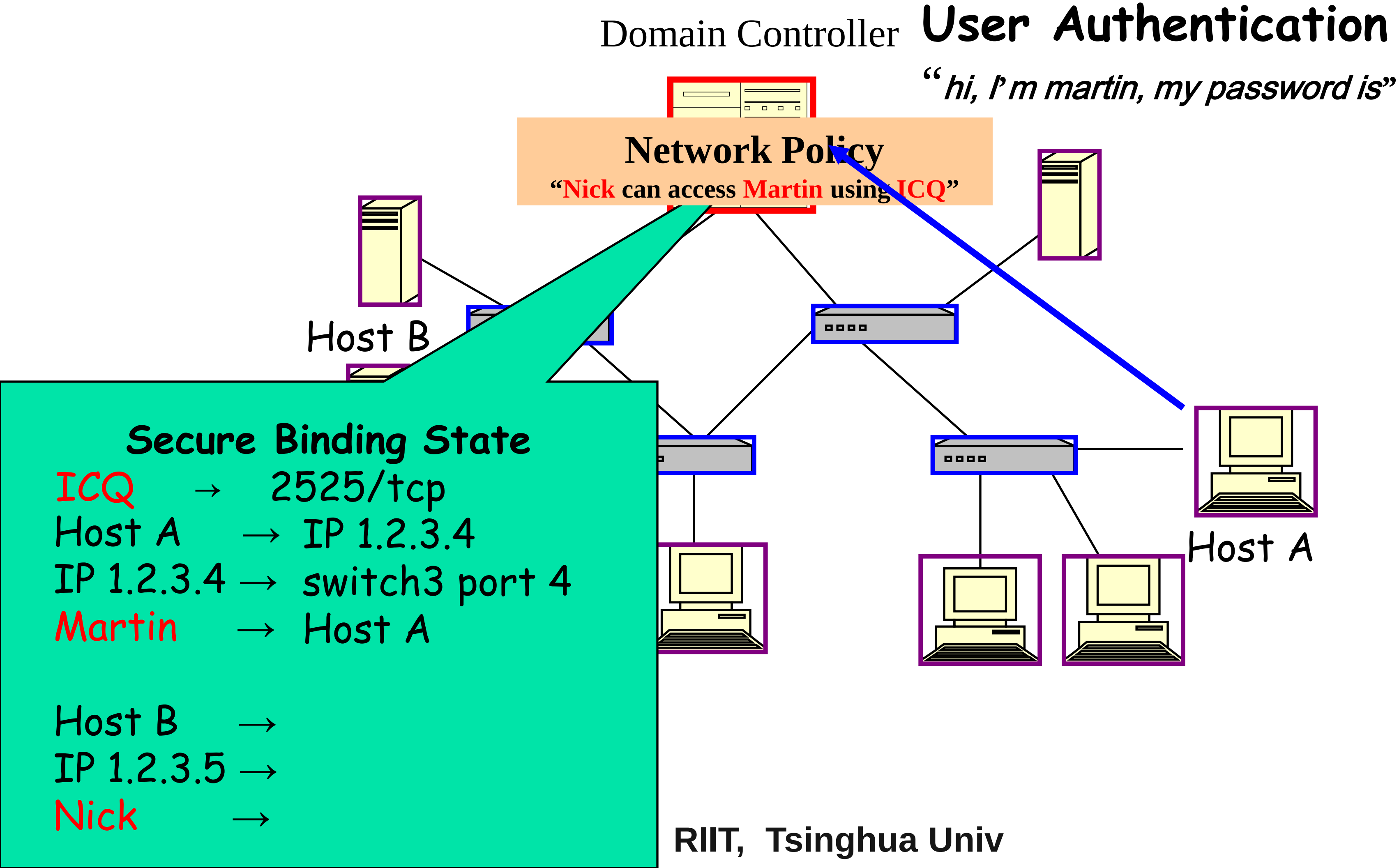
Secure Binding State

ICQ → 2525/tcp
Host A → IP 1.2.3.4
IP 1.2.3.4 → switch3 port 4
Martin →

Host B →
IP 1.2.3.5 →
Nick →

Ref. [4]

Ethane: High-Level Operation

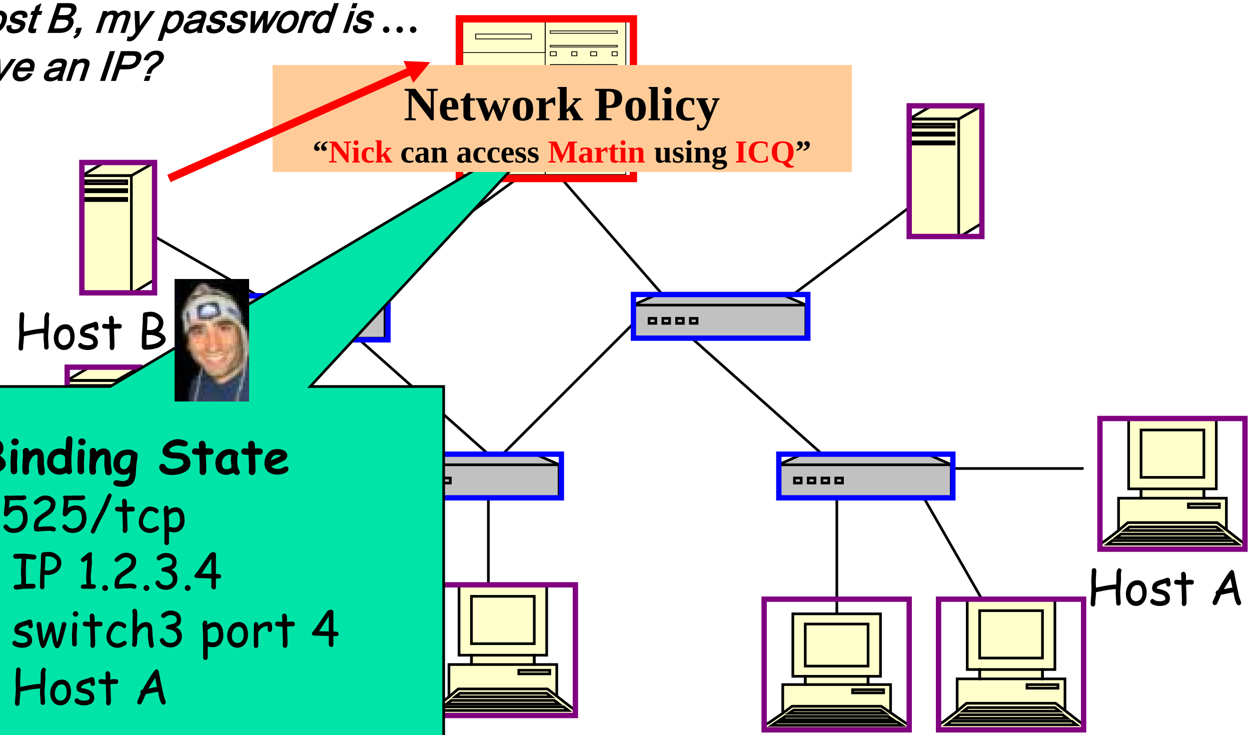


Ethane: High-Level Operation

Ref. [4]

Host authenticate Domain Controller

hi, I'm host B, my password is …
Can I have an IP?



Secure Binding State

ICQ → 2525/tcp
Host A → IP 1.2.3.4
IP 1.2.3.4 → switch3 port 4
Martin → Host A

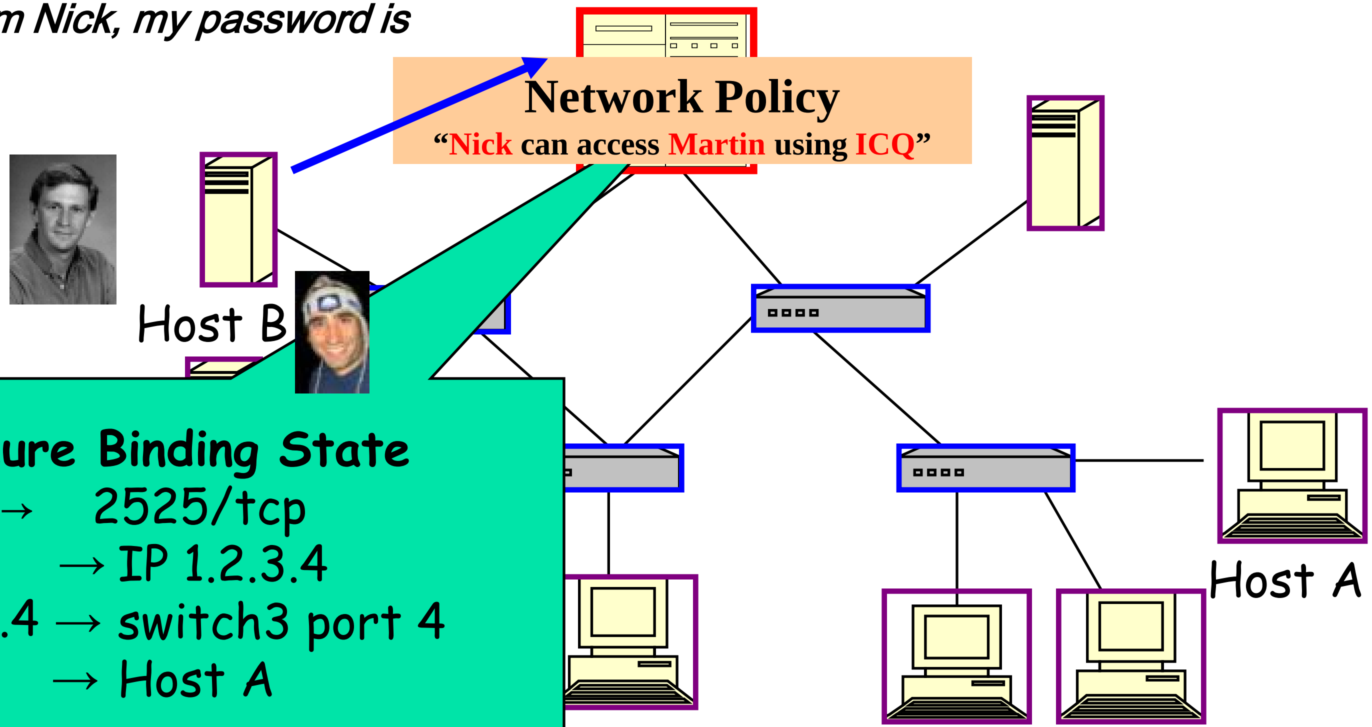
Host B → IP 1.2.3.5
IP 1.2.3.5 → switch 1 port 2
Nick →

Ethane: High-Level Operation

Ref. [4]

User authentication

hi, I'm Nick, my password is



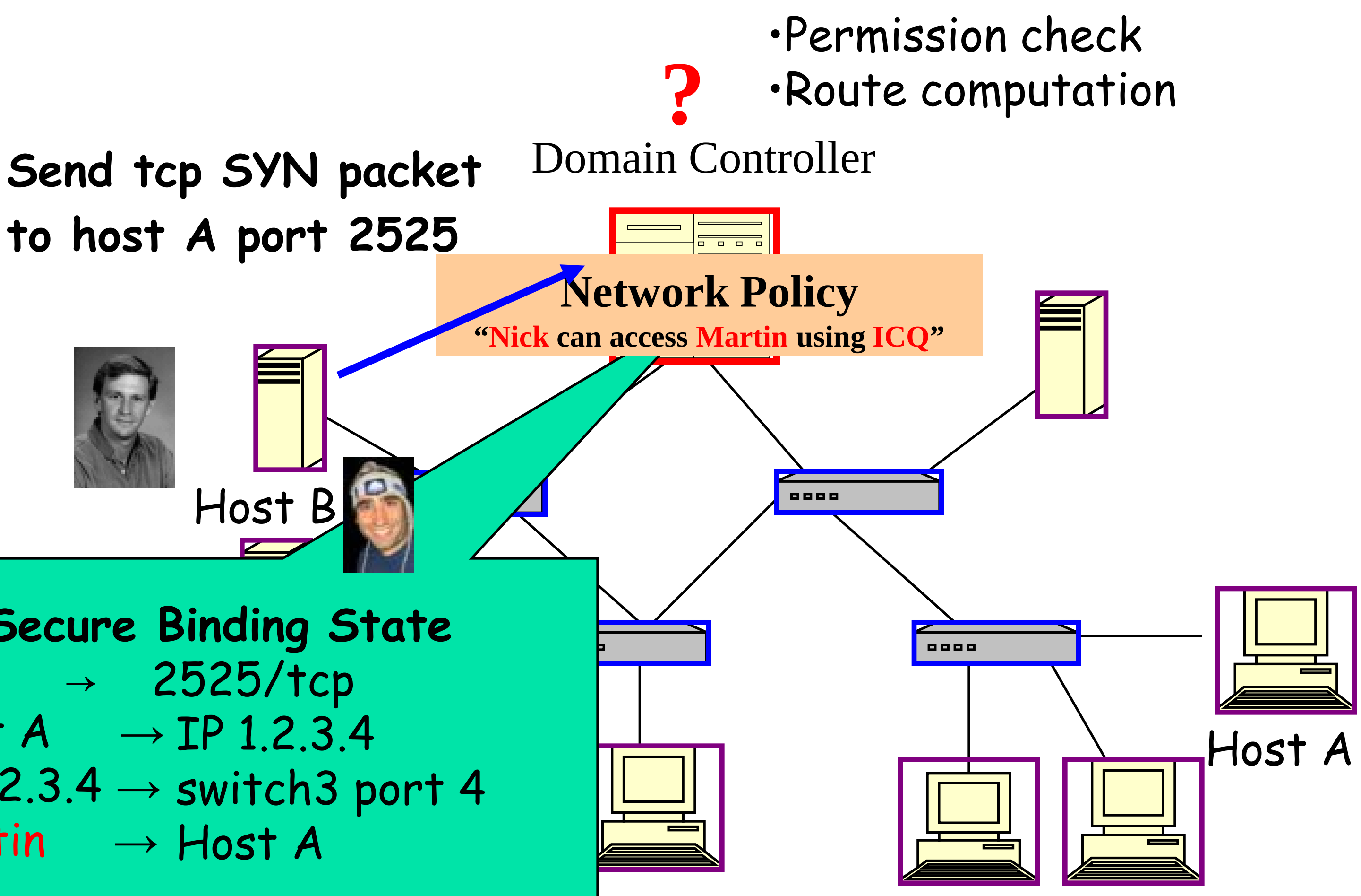
Secure Binding State

ICQ → 2525/tcp
Host A → IP 1.2.3.4
IP 1.2.3.4 → switch3 port 4
Martin → Host A

Host B → IP 1.2.3.5
IP 1.2.3.5 → switch 1 port 2
Nick → HostB

Ref. [4]

Ethane: High-Level Operation



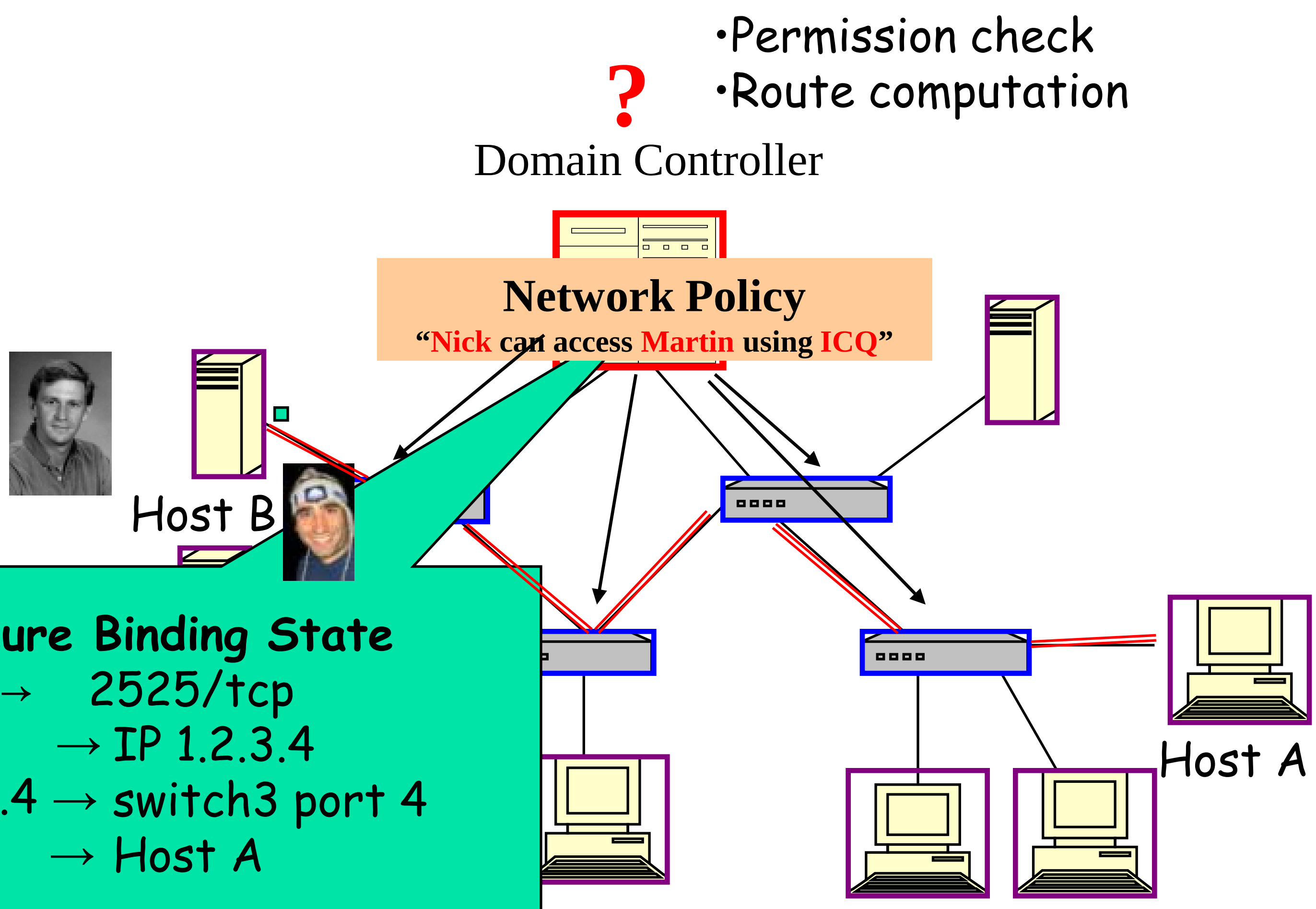
Secure Binding State

ICQ → 2525/tcp
Host A → IP 1.2.3.4
IP 1.2.3.4 → switch3 port 4
Martin → Host A

Host B → IP 1.2.3.5
IP 1.2.3.5 → switch 1 port 2
Nick → HostB

Ref. [4]

Ethane: High-Level Operation



Secure Binding State

ICQ → 2525/tcp
Host A → IP 1.2.3.4
IP 1.2.3.4 → switch3 port 4
Martin → Host A

Host B → IP 1.2.3.5
IP 1.2.3.5 → switch 1 port 2
Nick → HostB

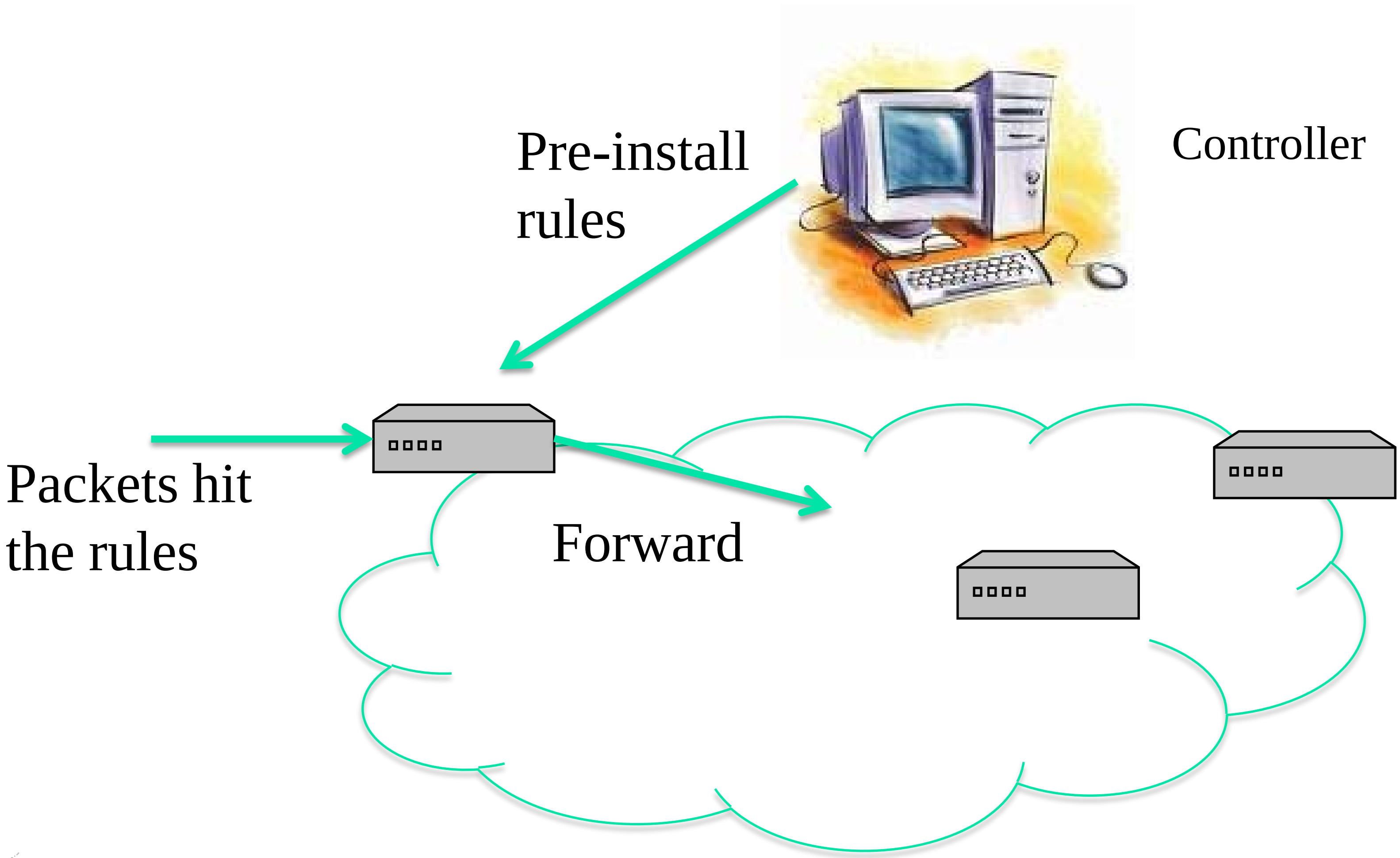
Ref. [5]

A Policy-aware Switching Layer for Data Centers

Dilip Joseph
Arsalan Tavakoli
Ion Stoica

University of California at Berkeley

Pre-install Rules in Switches



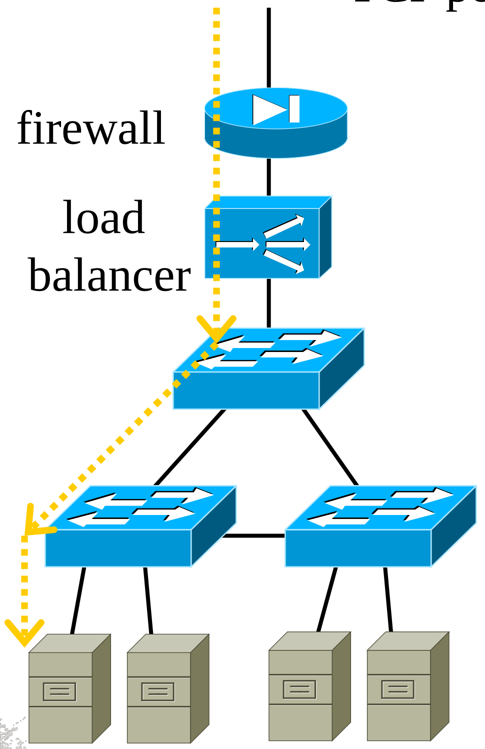
Policy-aware Switching Layer

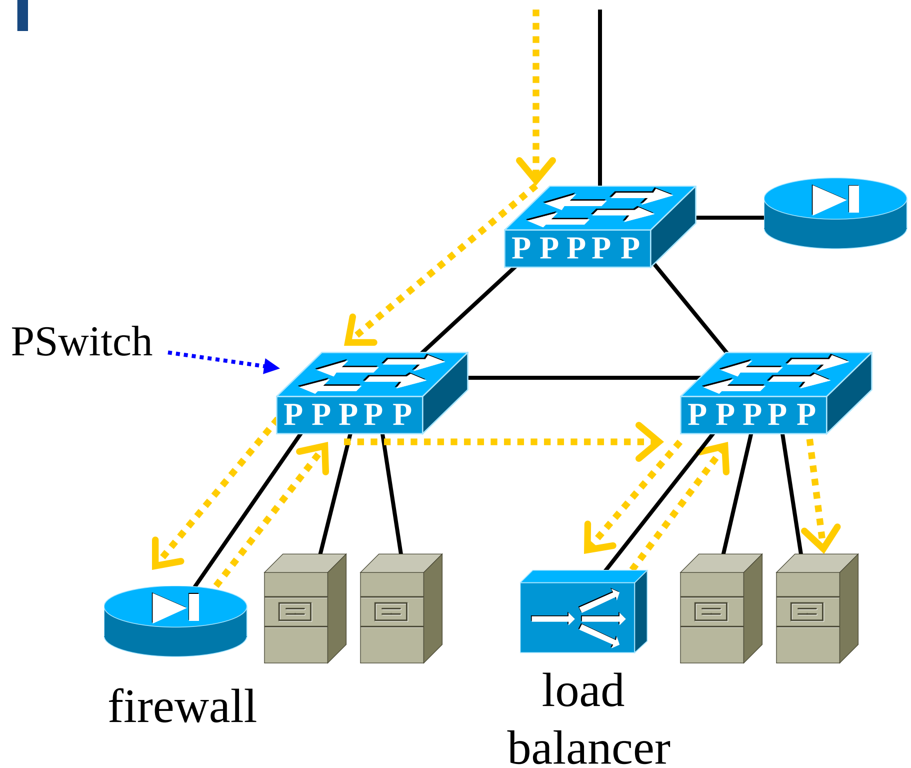
Ref. [5]

1 Take middleboxes off-path

2 Separate policy from reachability

HTTP
TCP port = 80 | Firewall → Load balancer



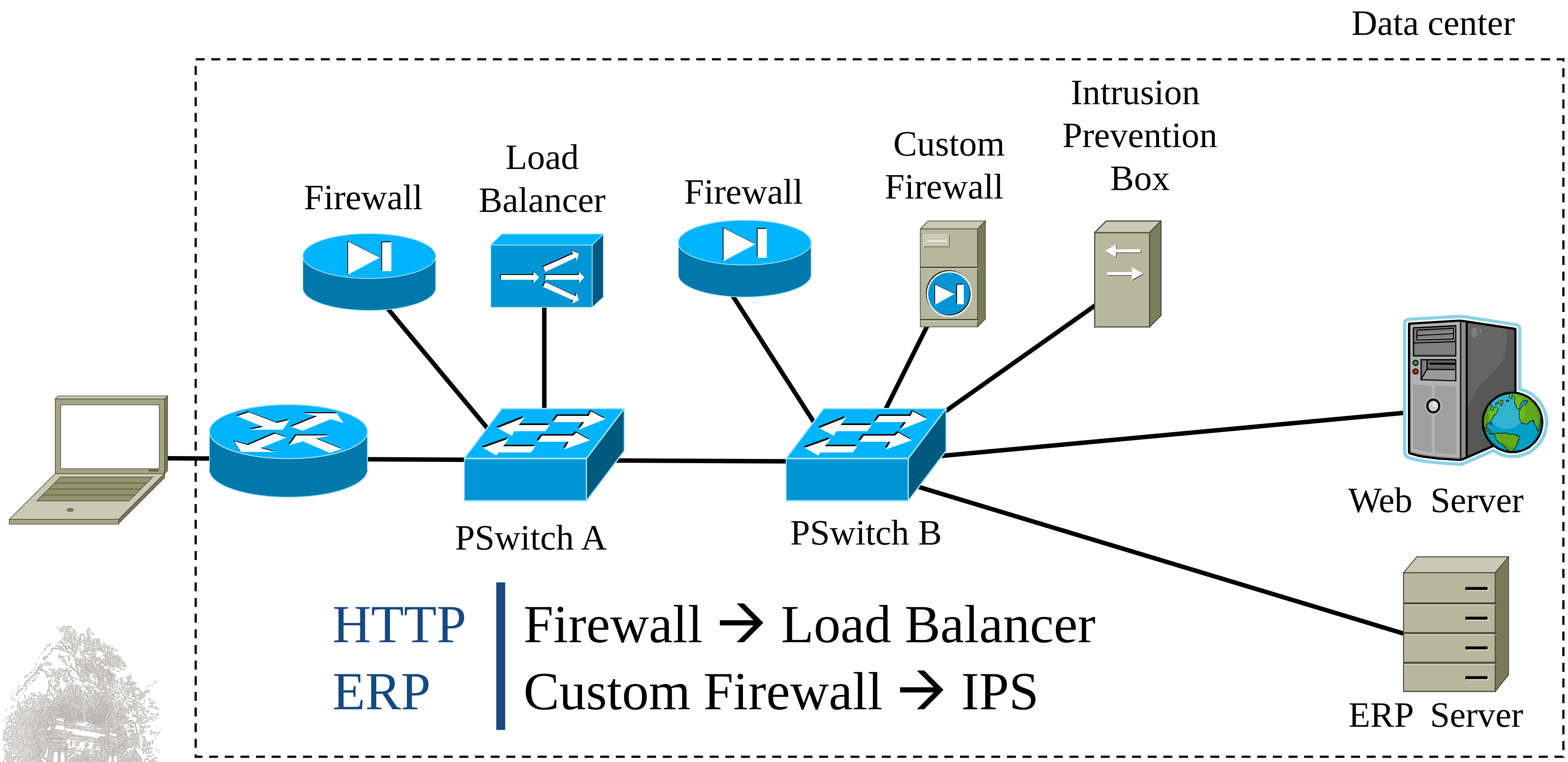


Existing mechanisms

Policy-aware switching layer

Ref. [5]

- Distributed forwarding
- Loadbalancing middleboxes
- Different policies for different traffic



Ref. [6]

Scalable Flow-Based Networking with DIFANE

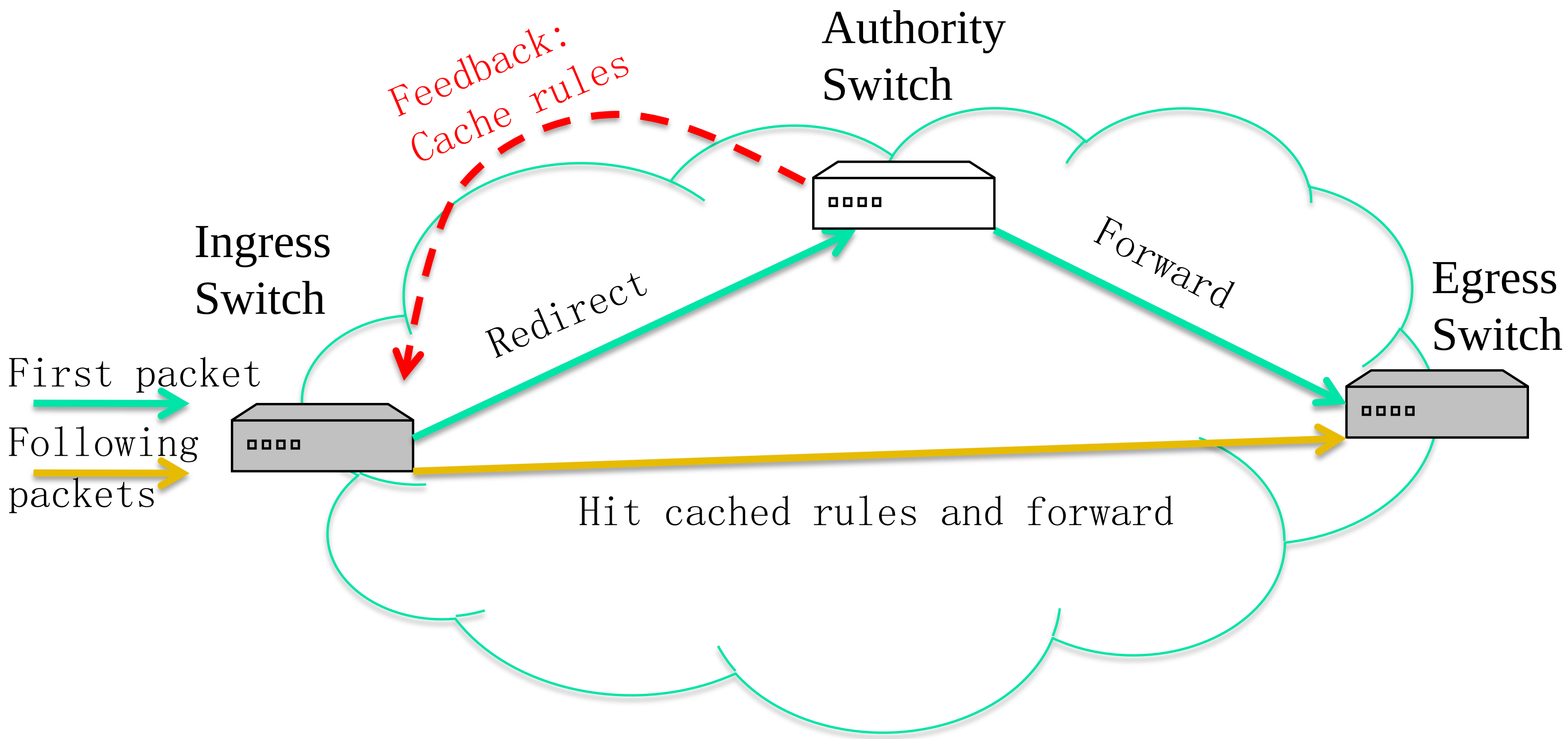
Minlan Yu

Princeton University

Joint work with Mike Freedman, Jennifer Rexford and Jia Wang

Ref. [6]

Packet Redirection and Rule Caching



A slightly longer path in the data plane is faster than going through the control plane

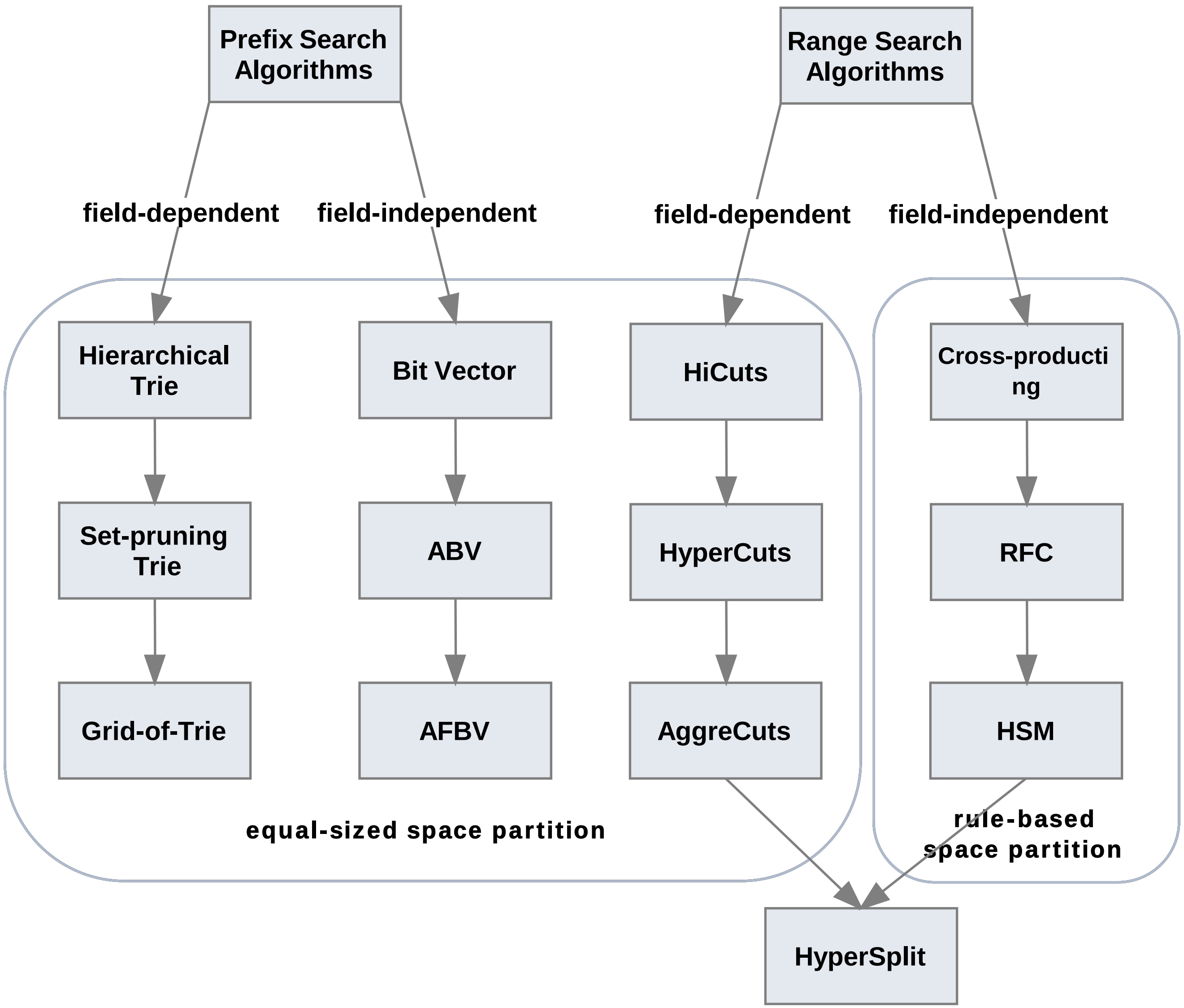
Outline

- Cloud Labels the “New Bottles”
- Cloud DCN inside the New Bottle
- “Old Wine” with a New Taste
  - Packet Classification
  - Pattern Matching
  - Flow Scheduling
  - Traffic Management

Key Algorithm Development (I)

- Packet Classification
  - Many classical algorithms: HiCuts, RFC, BV, GT, etc.
  - Recent algorithm advancements: HyperSplit (INFOCOM 09), etc.
- State-of-the-art
  - 100 Gbps and beyond throughput
  - Essential for Stateful Inspection
  - Enables rule based search without caching session
  - Makes TOR packet filtering possible

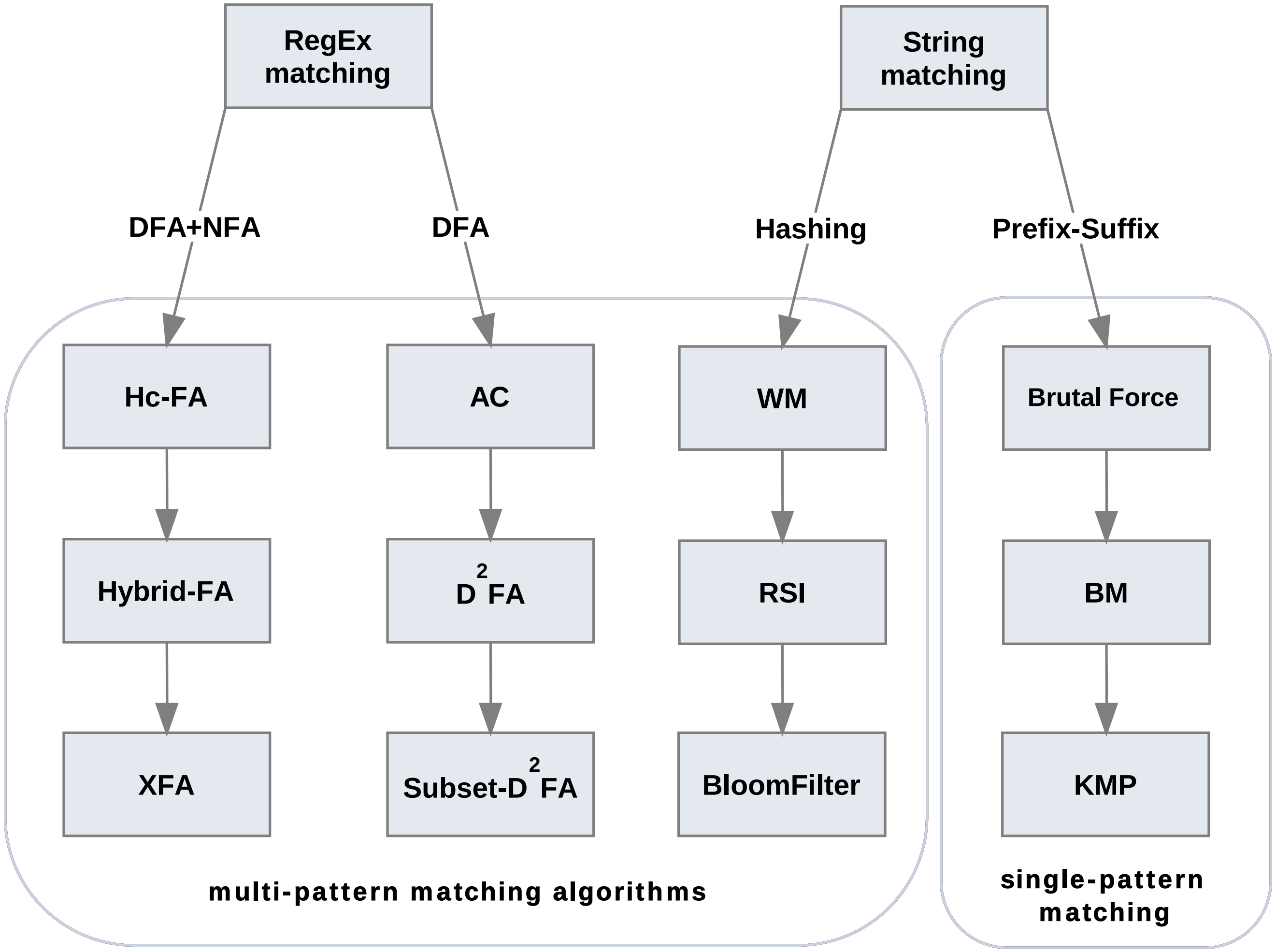
Packet Classification Taxonomy



Key Algorithm Development (II)

- Pattern Matching
  - Many classical algorithms: AC, BM, WM, D^2FA, etc.
  - Recent algorithm advancements: Subset-D^2FA (TBP), etc.
- State-of-the-art
  - 10Gbps and beyond throughput
  - Essential for Deep Inspection
  - Enables protocol identification at very high speed
  - Makes direct access hacking prevention possible

Pattern Matching Taxonomy



Key Algorithm Development (III)

- Flow Scheduling
  - Current industry standard: Equal-Cost Multi-Path (ECMP)
  - ECMP drawback: Static + Oblivious to link-utilization
- Hedera (NSDI 10)

  - Optimize achievable bisection bandwidth by assigning flows non-conflicting paths (Upto 96% of optimal bisection bandwidth, > 2X better than standard techniques)
  - Uses flow demand estimation + placement heuristics to find good flow-to-core mappings

Key Algorithm Development (IV)

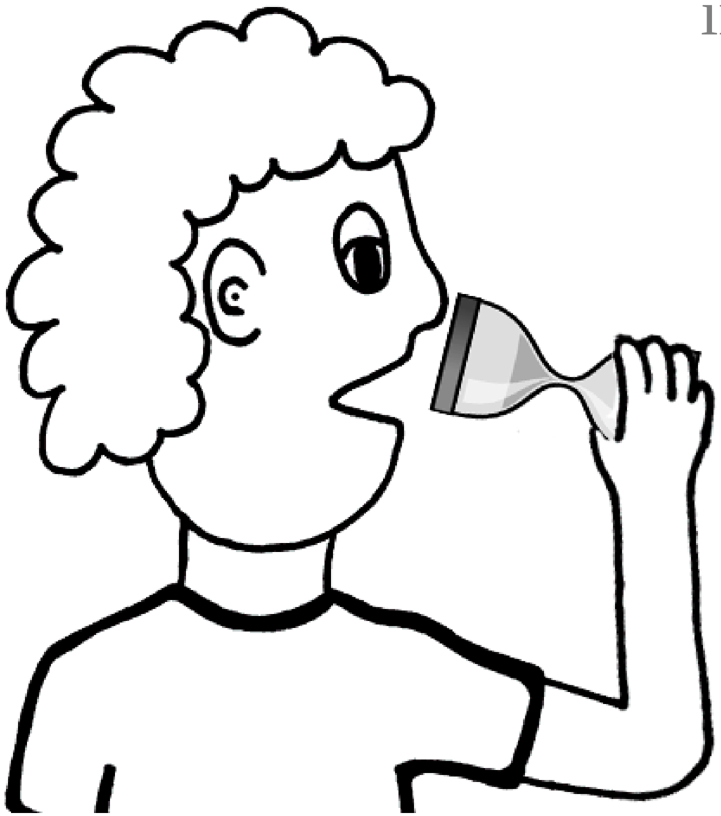
- Traffic Management
  - Previous work mostly on congestion control in end-to-end environment
  - Now approaches start to study global control mechanisms based on flow and protocol identification and measurement
- Besides NOX, recently DCTCP (SIGCOMM 10)
  - TCP-like protocol for DCN to achieve high burst tolerance, low latency, and high throughput with commodity shallow buffered switches
  - Queue length: about 1/10 of that of TCP

Summary

- Cloud DCN is key to overall cloud performance and security
- Cloud DCN requires flat architecture and demands high bandwidth, also to security solutions
- Cloud DCN security has different topology for security deployment, and key algorithms are still the same, but to be accelerated

Summary

Refined ***old wine*** will have better taste in ***new bottles*** with modern style!



Reference

1. Cloud Computing: Benefits, Risks and Recommendations for Information Security, European Network and Information Security Agency (ENISA) , Nov. 2009
2. T. Nolle, Cloud Networking: Inter-networking Data Centers and Connecting Users, Feb. 2010
3. The Cloud-ready Data Center Network, Juniper, May 2010
4. M. Casado, M. Freedman, G. Gibb, L. Glendenning, D. Boneh, N. McKeown, S. Shenker, & G. Watson, Ethane: Addressing the Protection Problem in Enterprise Networks, Stanford & UCB, Aug. 2007
5. D. Joseph, A. Tavakoli, & I. Stoica, A Policy-aware Switching Layer for Data Centers, UCB, Aug. 2008
6. M. Yu, M. Freedman, J. Rexford, & J. Wang, Scalable Flow-Based Networking with DIFANE, Princeton, Sept. 2010
7. Y. Qi, L. Xu, B. Yang, Y. Xue, & J.Li, Packet Classification Algorithms: From Theory to Practice, THU, Apr. 2009
8. Y. Qi, J. Fong, W. Jiang, B. Xu, J. Li, & V.Prasanna, Multi-dimensional Packet Classification on FPGA: 100 Gbps and Beyond, THU & USC, Dec. 2010
9. M. Al-Fares & S. Radhakrishnan, B. Raghavan, N. Huang, & A. Vahdat, Hedera: Dynamic Flow Scheduling for Data Center Networks, UCSD, Apr. 2010
10. M. Alizadeh, A. Greenberg, D. Maltz, J. Padhyey, P. Pately, B. Prabhakarz, S. Senguptay, & M. Sridharan, Data Center TCP (DCTCP), MSR & Stanford, Aug. 2010