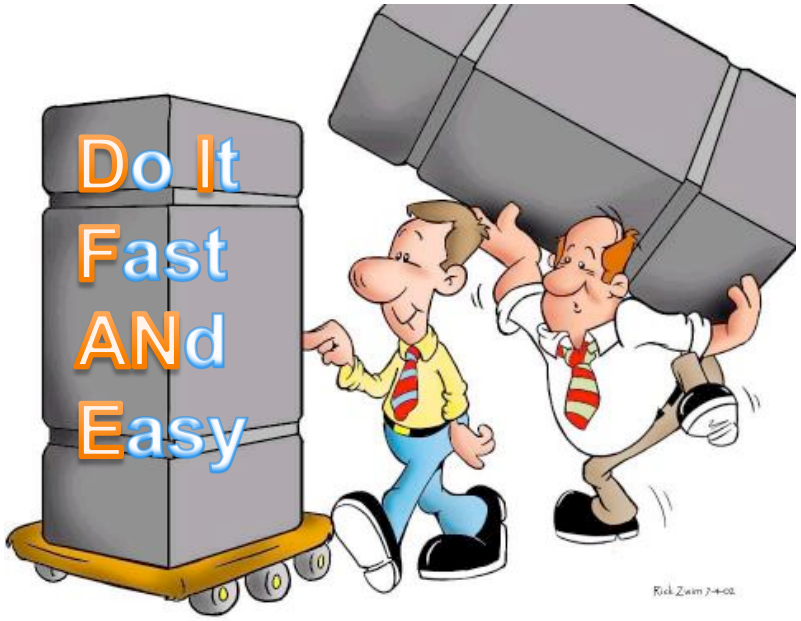
Scalable Flow-Based Networking with DIFANE

Minlan Yu

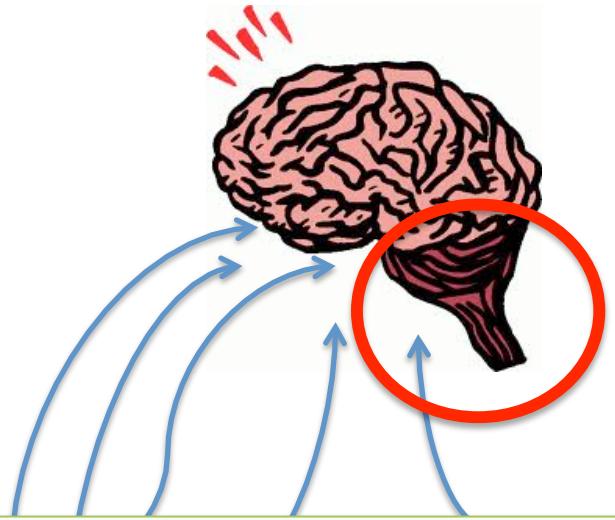
Princeton University



Joint work with Mike Freedman, Jennifer Rexford and Jia Wang

What's DIFANE?

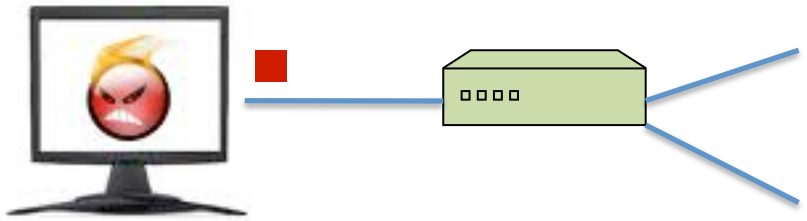
- Traditional enterprise networking
  - Hard to manage
  - Limited fine-grained policy
  - Scalability remains a challenge

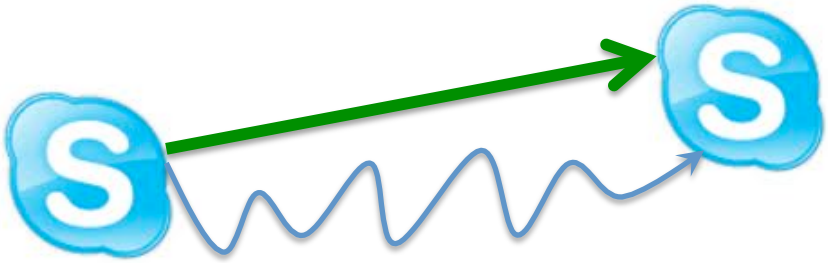


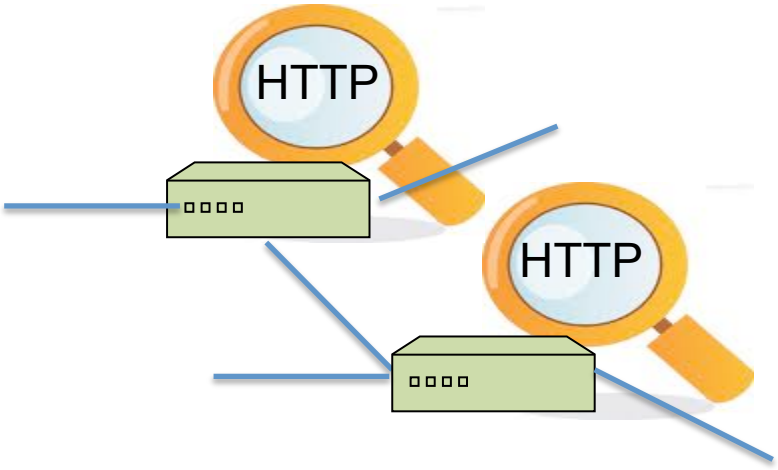
DIFANE:
A scalable way to apply fine-grained policies in enterprises

Flexible Policies in Enterprises

- Access control
  - Drop packets from malicious hosts
- Customized routing
  - Direct Skype calls on a low-latency path
- Measurement
  - Collect detailed HTTP traffic statistics

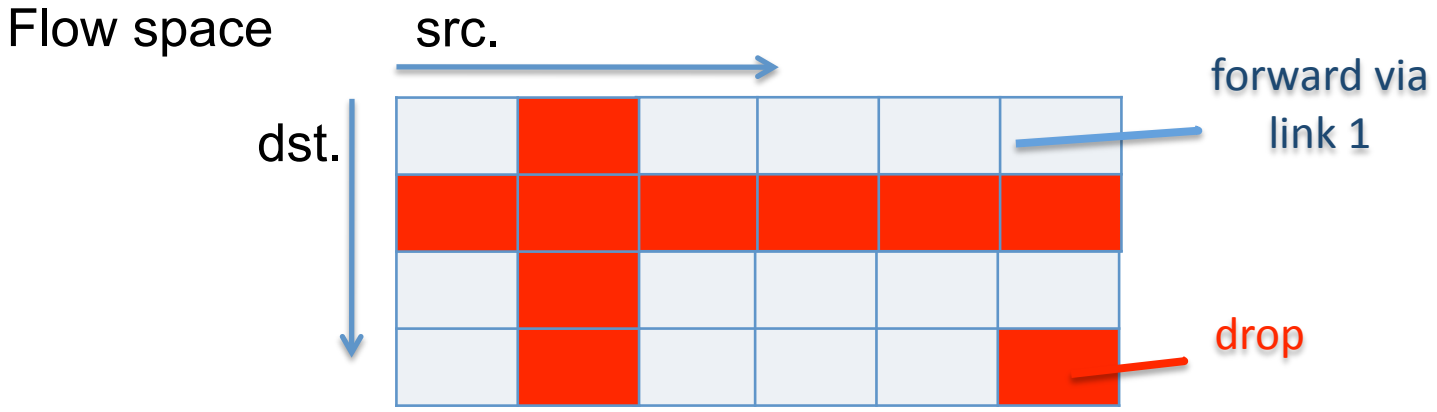






Flow-based Switches

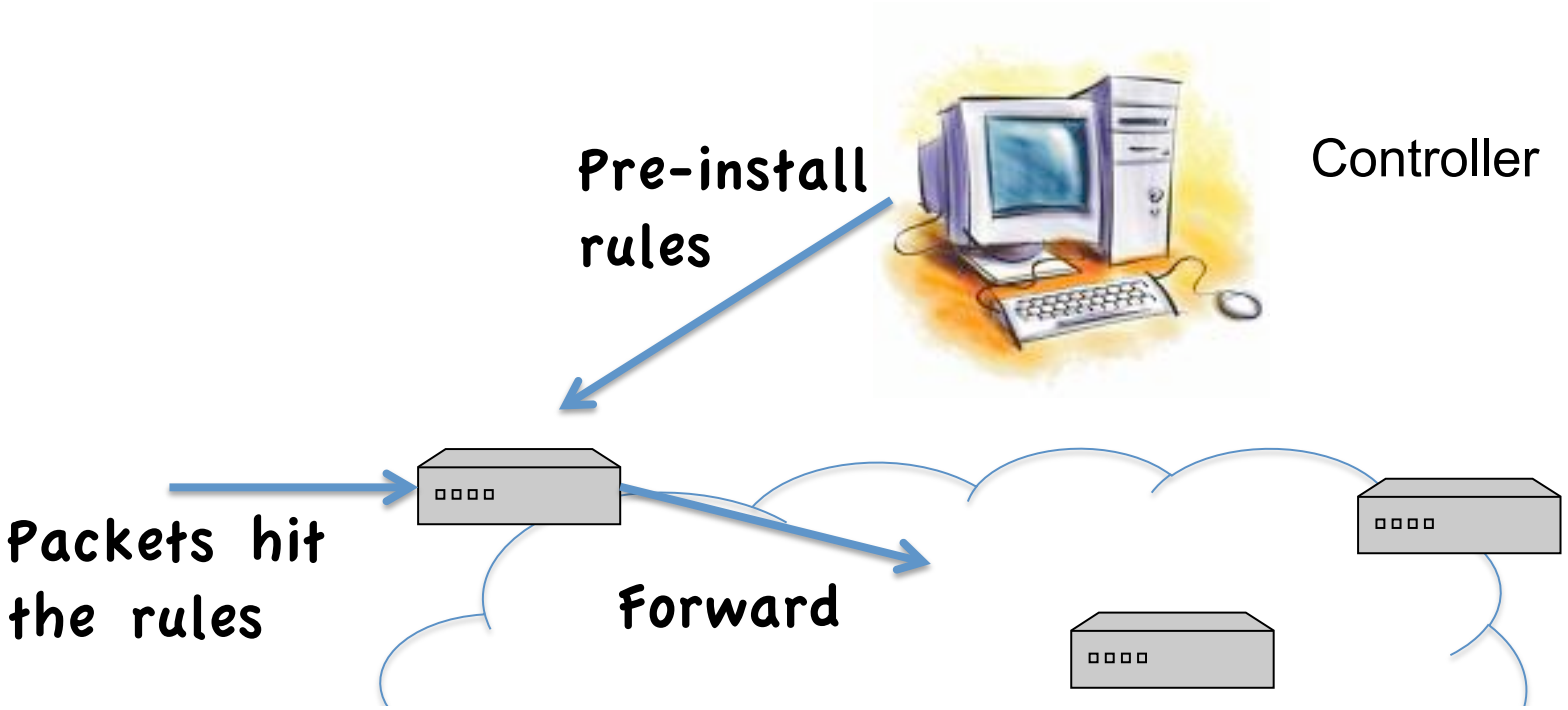
- Install rules in flow-based switches
  - Store rules in high speed memory (TCAM)
- Perform simple actions based on rules
  - Rules: Match on bits in the packet header
  - Actions: Drop, forward, count



Challenges of Policy-Based Management

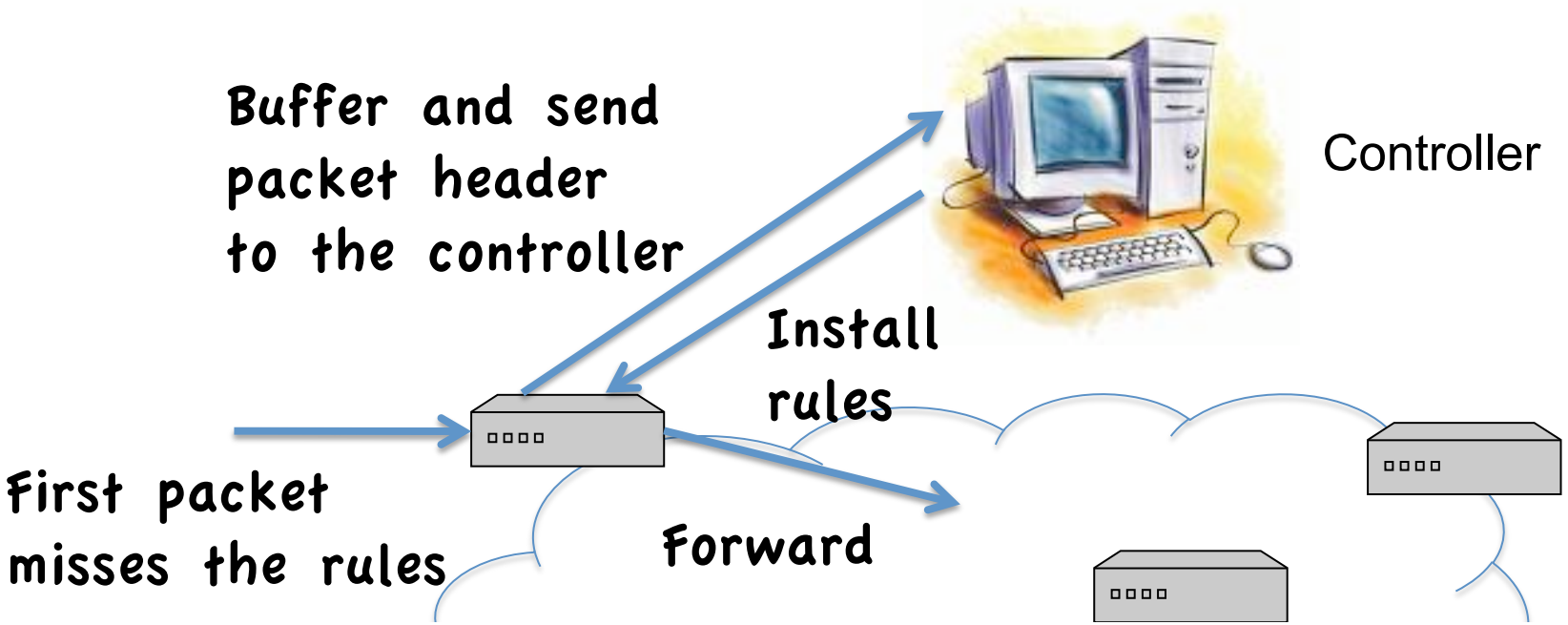
- Policy-based network management
  - Specify *high-level policies* in a management system
  - Enforce *low-level rules* in the switches
- Challenges
  - Large number of hosts, switches and policies
  - Limited TCAM space in switches
  - Support host mobility
  - No hardware changes to commodity switches

Pre-install Rules in Switches



- Problems:
  - No host mobility support
  - Switches do not have enough memory

Install Rules on Demand (Ethane, NOX)



- Problems:
  - Delay of going through the controller
  - Switch complexity
  - Misbehaving hosts

DIFANE: Combining Proactive & Reactive

Features \ Install rules	Proactive	Reactive(Ethane)	DIFANE

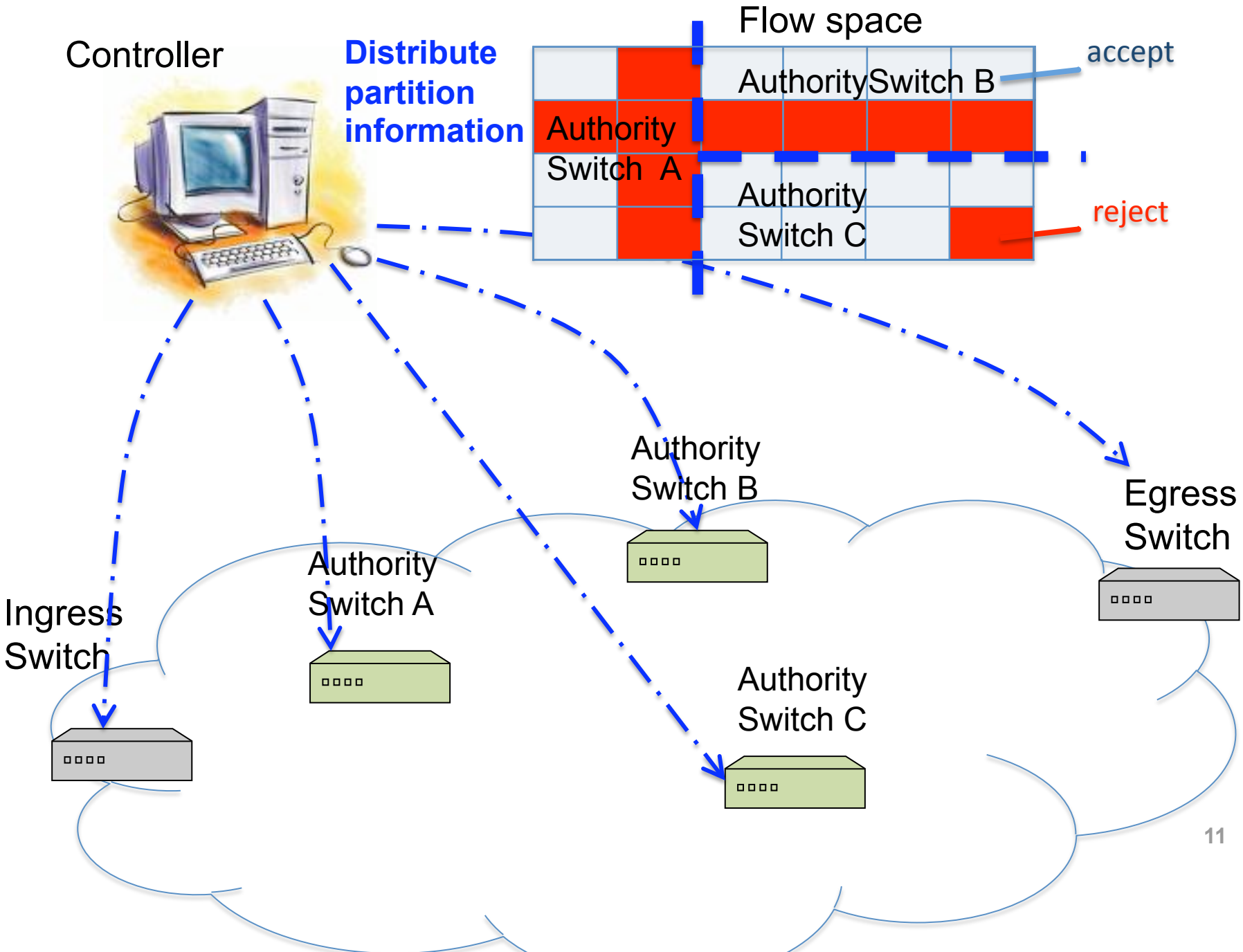
DIFANE Architecture (two stages)

DIstributed Flow Architecture for Networked Enterprises

Stage 1

The controller *proactively* generates the rules and distributes them to authority switches.

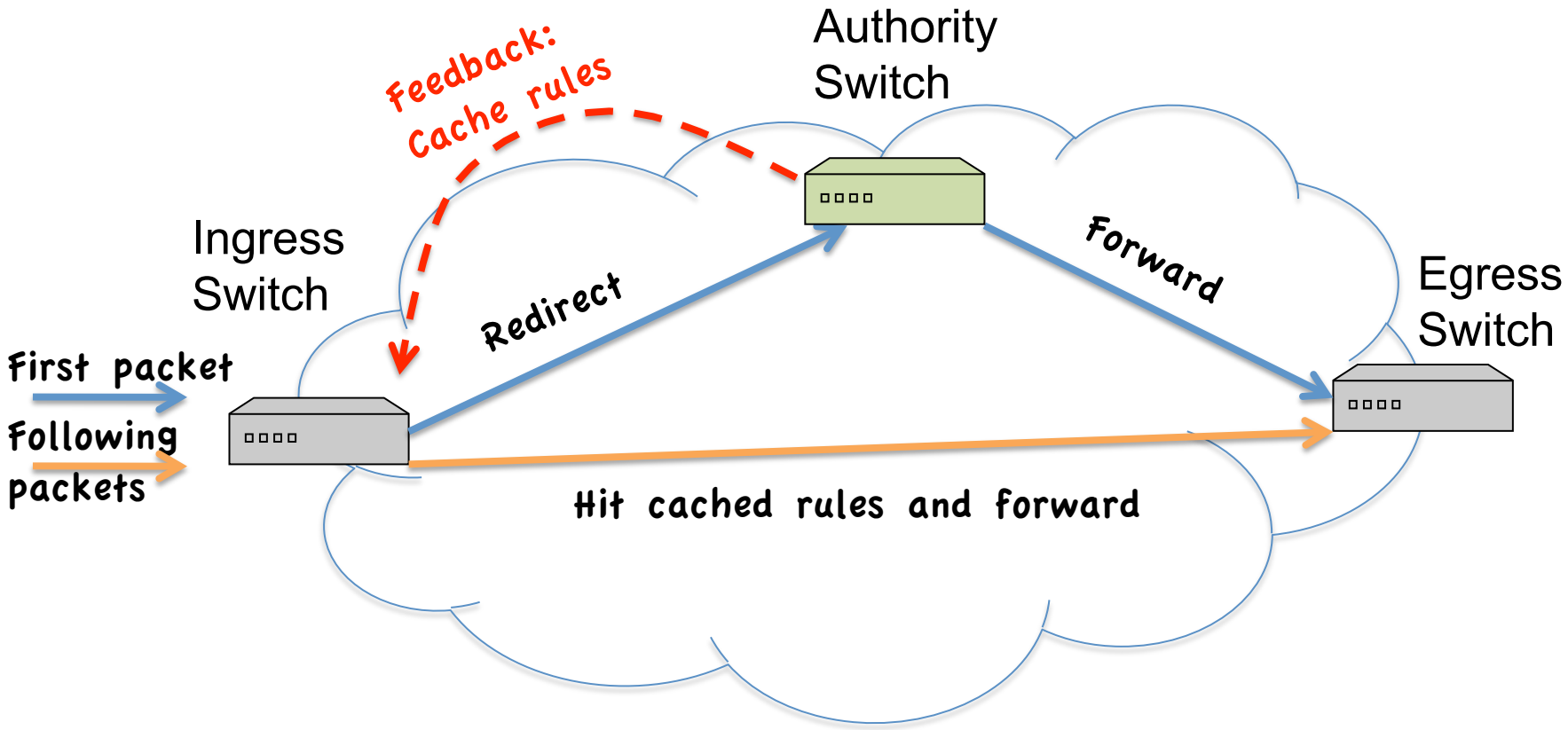
Partition and Distribute the Flow Rules



Stage 2

The authority switches keep packets always in the data plane and *reactively* cache rules.

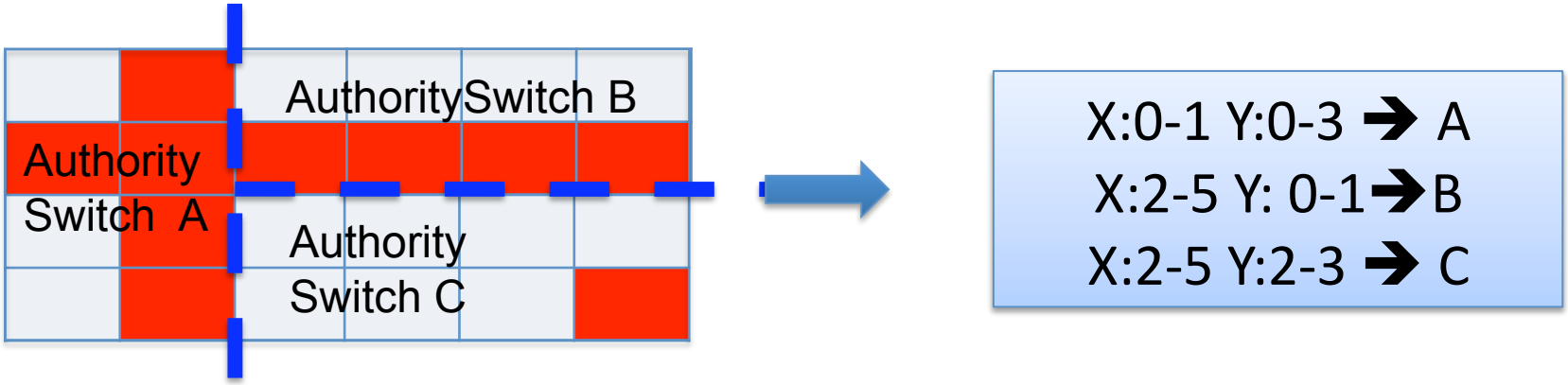
Packet Redirection and Rule Caching



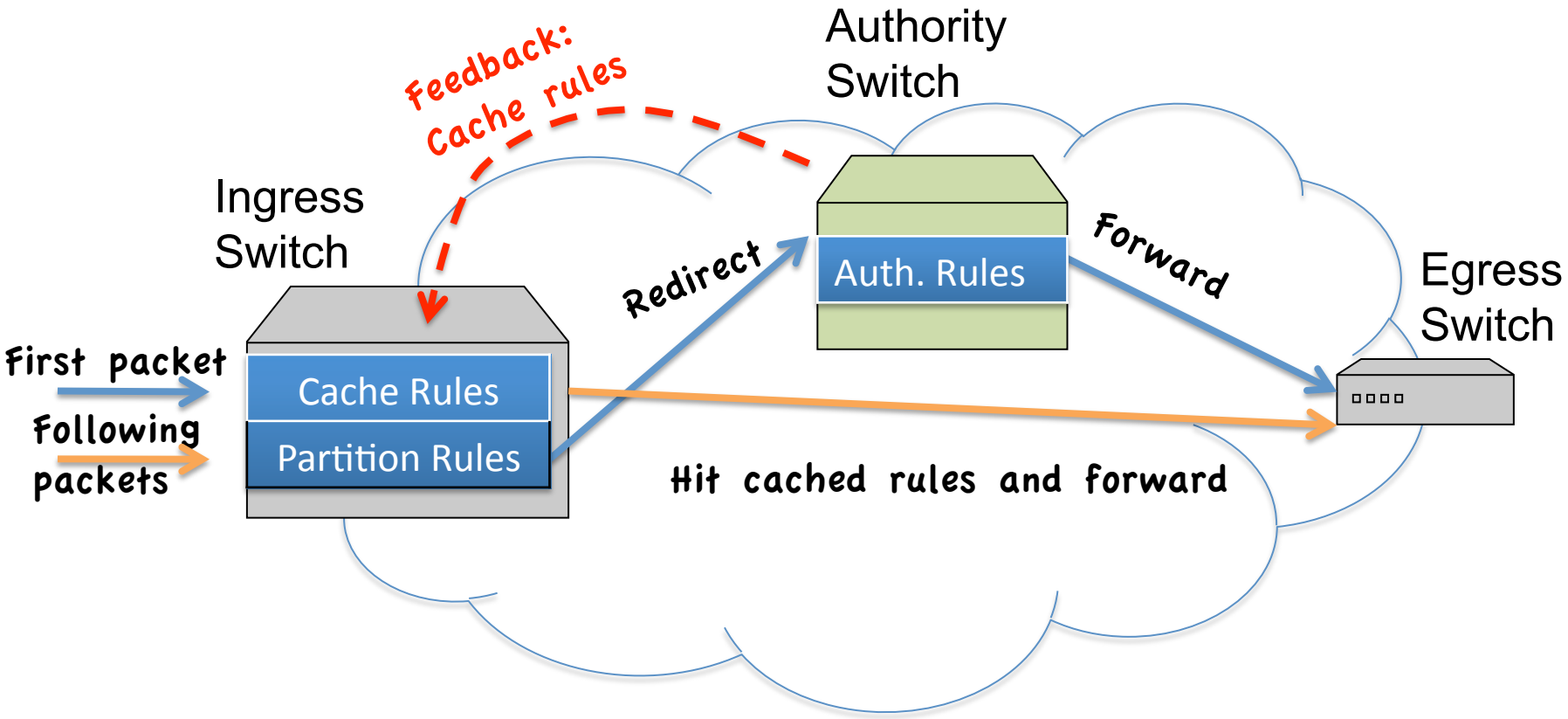
A slightly longer path in the data plane is faster than going through the control plane

Locate Authority Switches

- Partition information in ingress switches
  - Using a small set of coarse-grained wildcard rules
  - … to locate the authority switch for each packet
- Distributed directory service but not DHT
  - Hashing does *not* work for wildcards
  - Keys can have wildcards in arbitrary bit positions



Packet Redirection and Rule Caching



Three Sets of Rules in TCAM

Type	Priority	Field 1	Field 2	Action	Timeout
Cache Rules					
	…	…	…	…	…
Authority Rules	110	00**	001*	Forward	Infinity
	…	…	…	…	…
Partition Rules					
	…	…	…	…	…

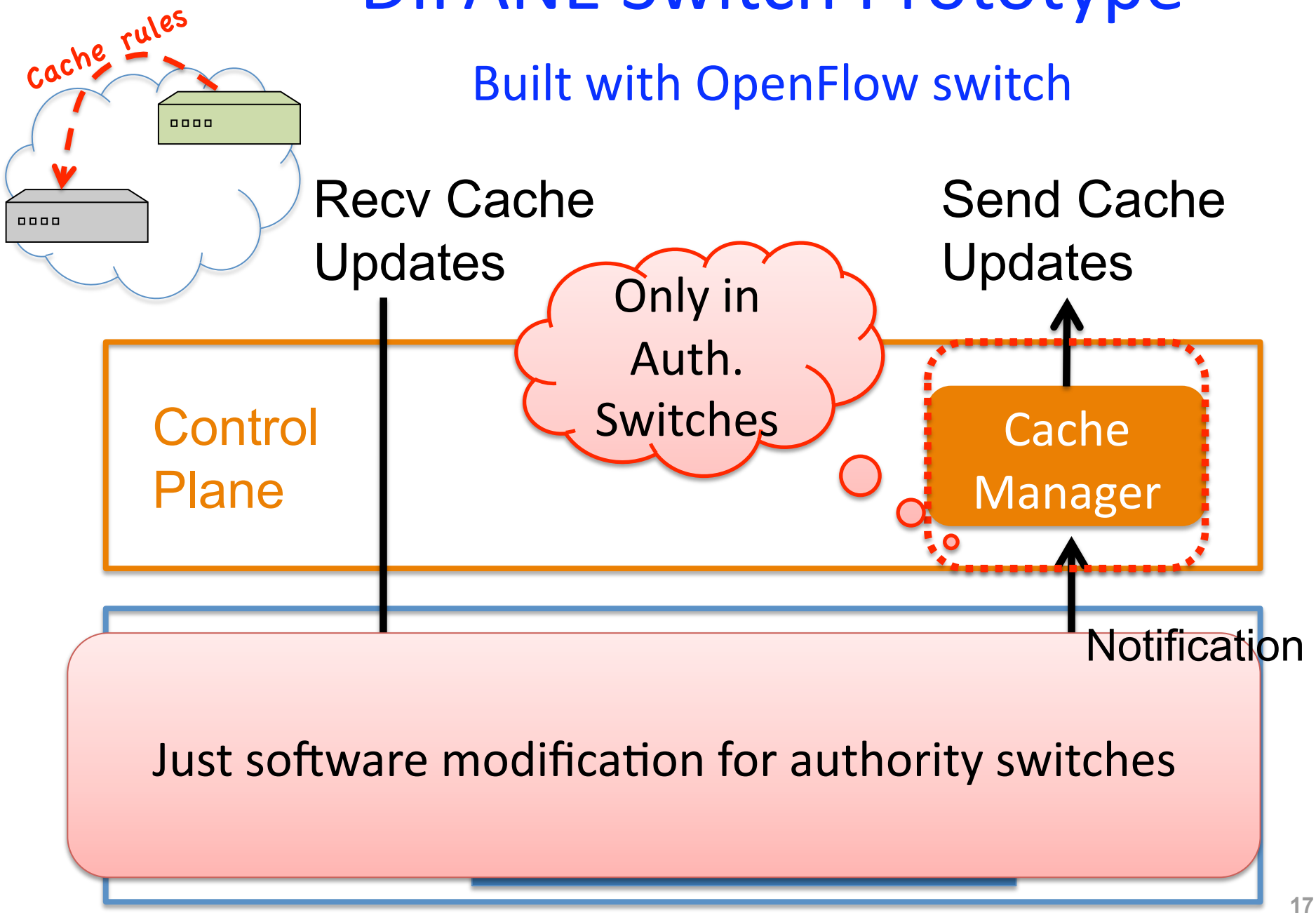
In ingress switches
reactively installed by authority switches

In authority switches
proactively installed by controller

In every switch
proactively installed by controller

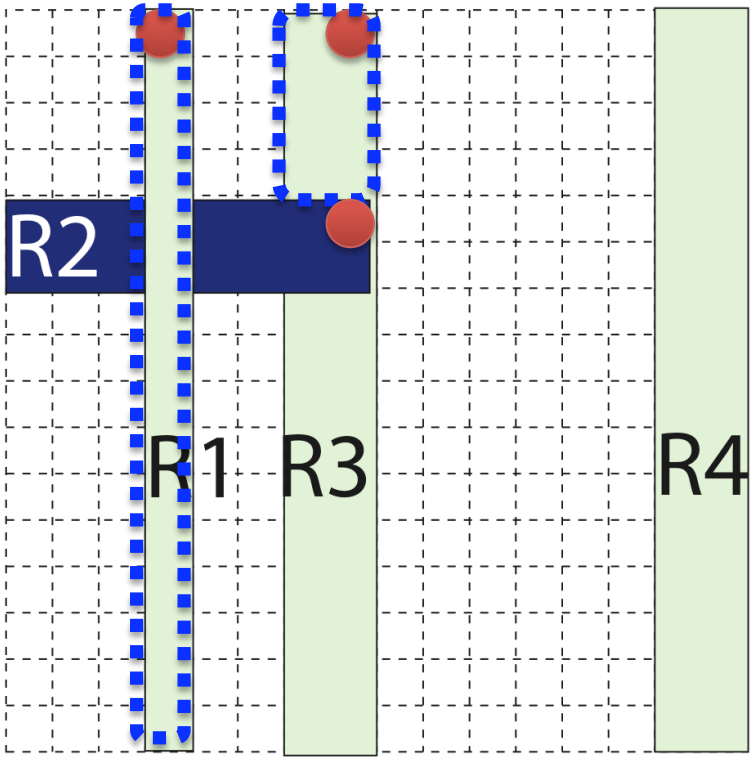
DIFANE Switch Prototype

Built with OpenFlow switch



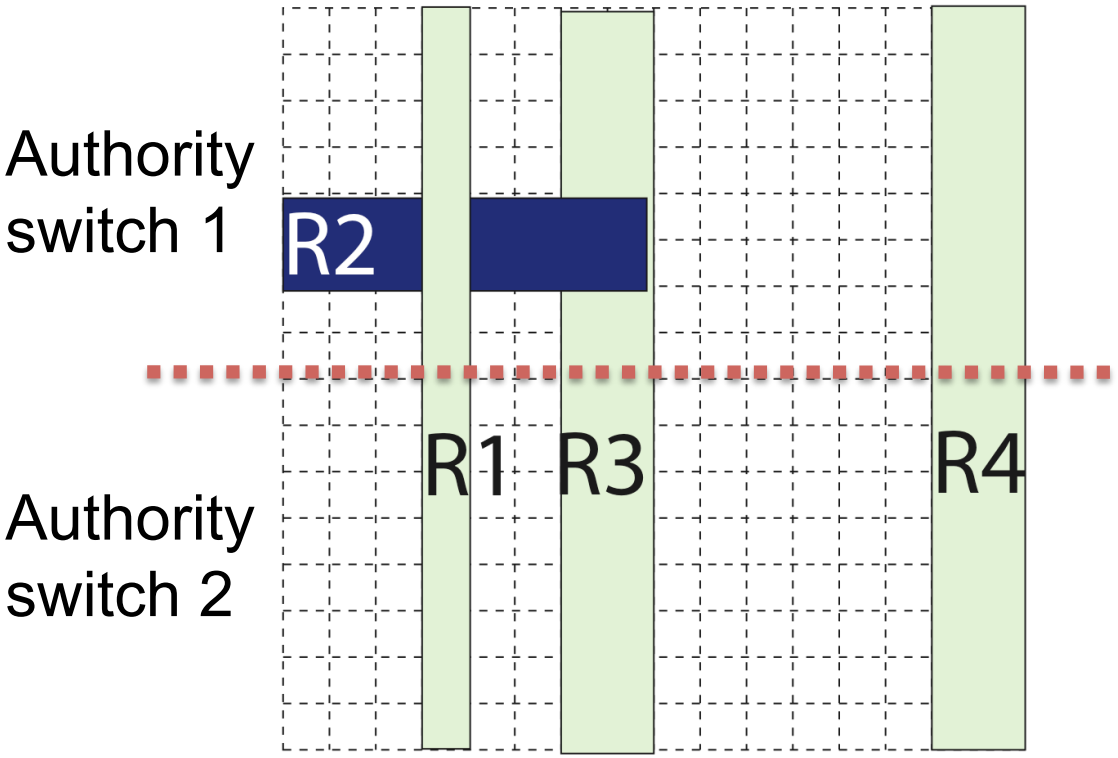
Caching Wildcard Rules

- Overlapping wildcard rules
  - Cannot simply cache matching rules



Caching Wildcard Rules

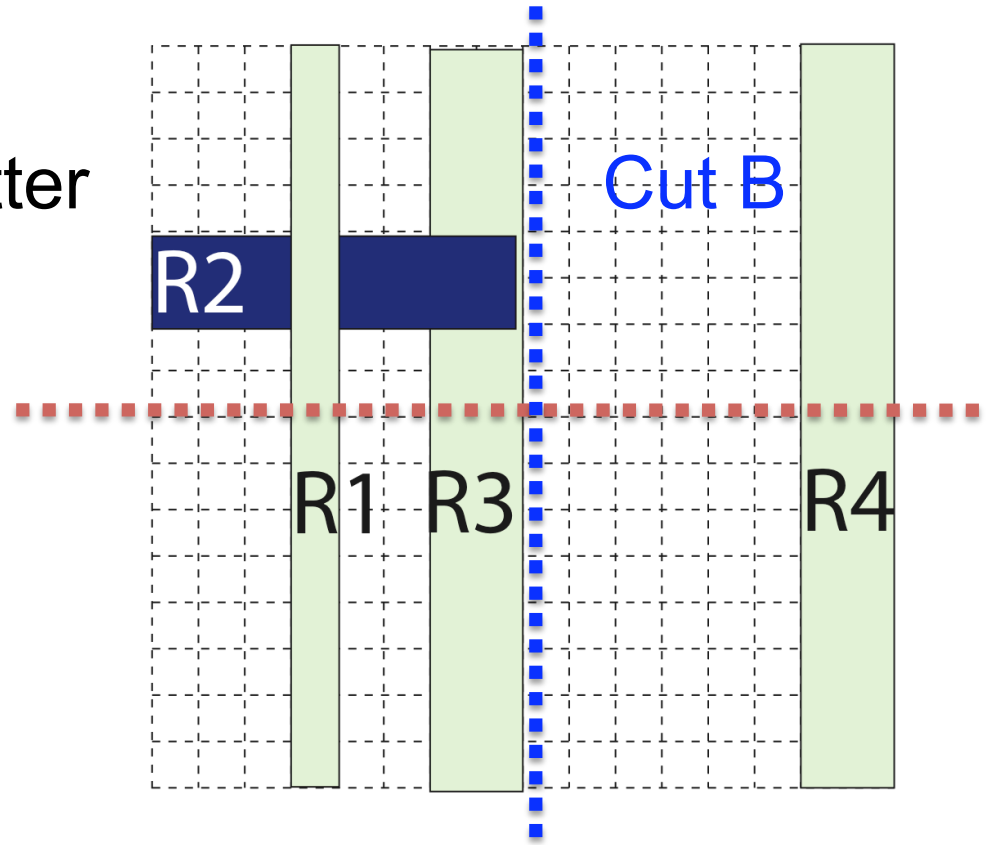
- Multiple authority switches
  - Contain independent sets of rules
  - Avoid cache conflicts in ingress switch



Partition Wildcard Rules

- Partition rules
  - Minimize the TCAM entries in switches
  - Decision-tree based rule partition algorithm

Cut B is better than Cut A



Handling Network Dynamics

Network dynamics	Cache rules	Authority Rules	Partition Rules
Policy changes at controller	Timeout	Change	Mostly no change
Topology changes at switches	No change	No change	Change
Host mobility	Timeout	No change	No change

Prototype Evaluation

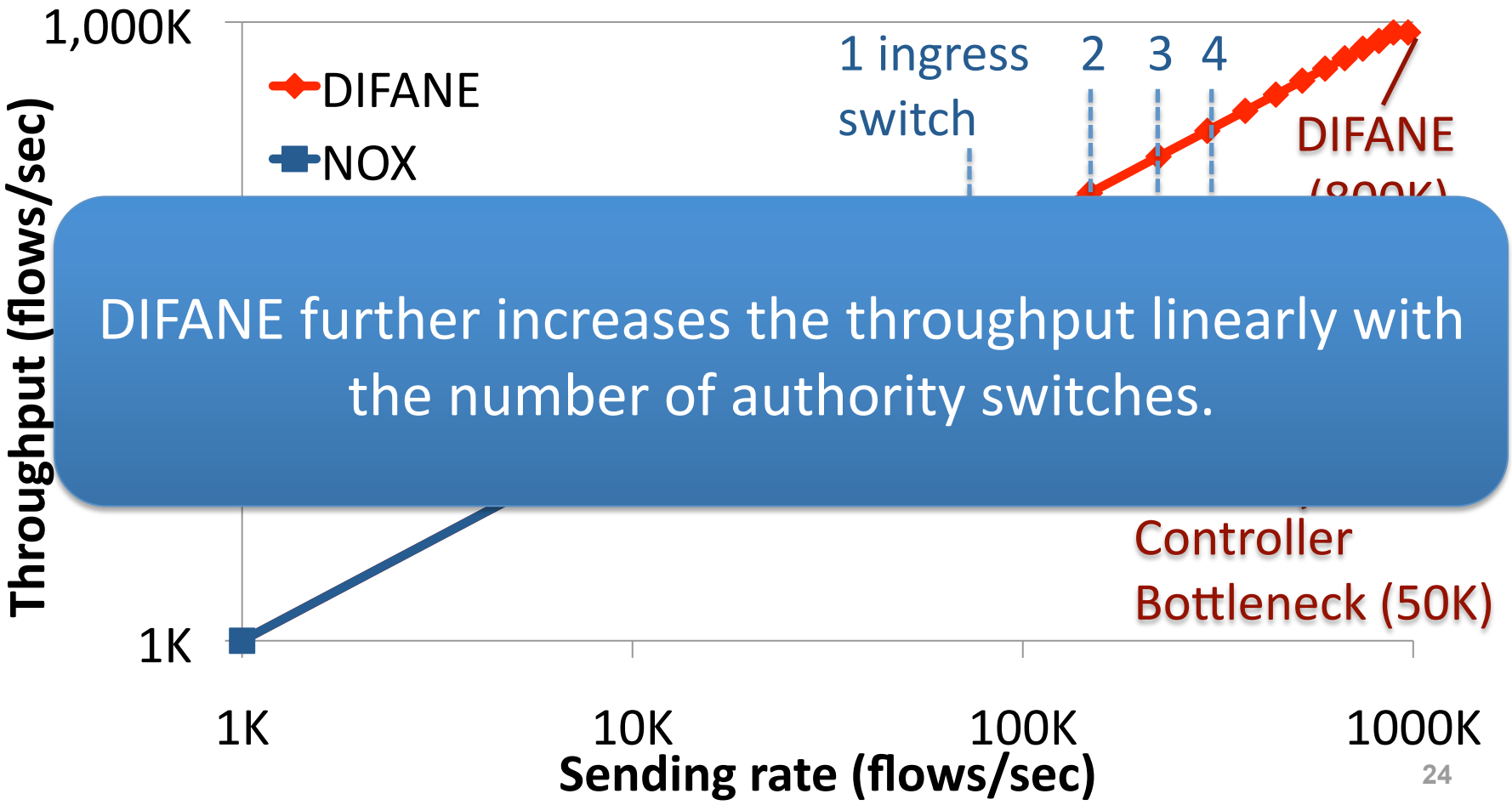
- Evaluation setup
  - Kernel-level Click-based OpenFlow switch
  - Traffic generators, switches, controller run on separate 3.0GHz 64-bit Intel Xeon machines
- Compare delay and throughput
  - NOX: Buffer packets and reactively install rules
  - DIFANE: Forward packets to authority switches

Delay Evaluation

- Average delay (RTT) of the first packet
  - NOX: 10 ms
  - DIFANE: 0.4 ms
- Reasons for performance improvement
  - Always keep packets in the data plane
  - Packets are delivered without waiting for rule caching
  - Easily implemented in hardware to further improve performance

Peak Throughput

- One authority switch; Single-packet flow



DIFANE further increases the throughput linearly with the number of authority switches.

Scaling with Many Rules

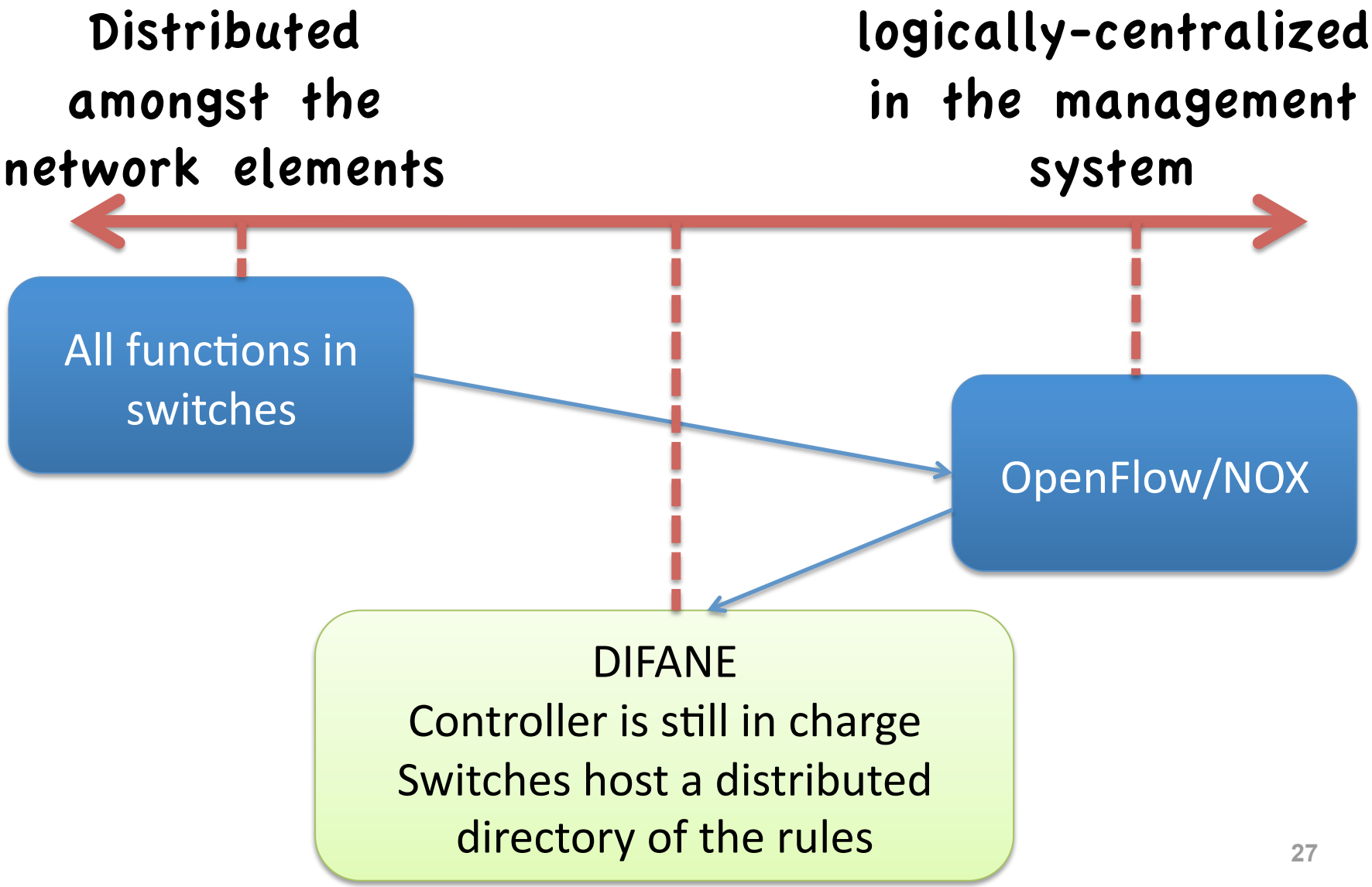
- How many authority switches do we need?
  - Depends on total number of rules

  … and the TCAM space in these authority switches

	Campus	IPTV
# Rules	30K	5M
# Switches	1.7K	3K
Assumed Authority Switch TCAM size	160 KB	1.6 MB
Required # Authority Switches	5 (0.3%)	100 (3%)

Stepping back …

Distributed or Centralized?



Thanks!