Chair for Network Architectures and Services
Department of Informatics
TU München – Prof. Carle

Software Defined Networking

Cornelius Diekmann

Contents

Motivation

Mastering Complexity and Extracting Simplicity

A thought experiment

- How would you develop a tool to manage data on an USB stick? On a bare metal machine!

1) Write Input/Output function in assembly
2) Write some operating system kernel in C that calls your assembly routines. Implement a file system.
3) Add an operating system kernel API that allows to start userland processes that can access the file system
4) Write userland tools such as `cat`, `cp`, `mv`, to manage the data on on your USB stick.

Mastering Complexity and Extracting Simplicity

The thought experiment, conclusion.

- Assembly
  - Low level machine language, you must be a of master complexity
- Languages such as C
  - Easier to program the hardware than in Assembly
  - Is translated to Assembly by a compiler
- Operating systems
  - Provide abstraction of hardware, processes, file systems, …
- Languages such as C or Java
  - You can simply write your program
  - The operating system manages (most of) your resources
  - Some environments even manage your memory for you

Mastering Complexity and Extracting Simplicity

Another thought experiment

- How **do** you manage a network on the link layer?

1) Configure all the forwarding tables
2) Configure all the Access Control Lists
3) … set up state in every device ...
4) Check the connectivity, test, test, `ping`, `traceroute`, `nmap`, …. debug, fix, test, …

Compared to the first thought experiment, this is like writing everything in Assembly and debugging it by manually, inspecting all the memory!

Mastering Complexity and Extracting Simplicity

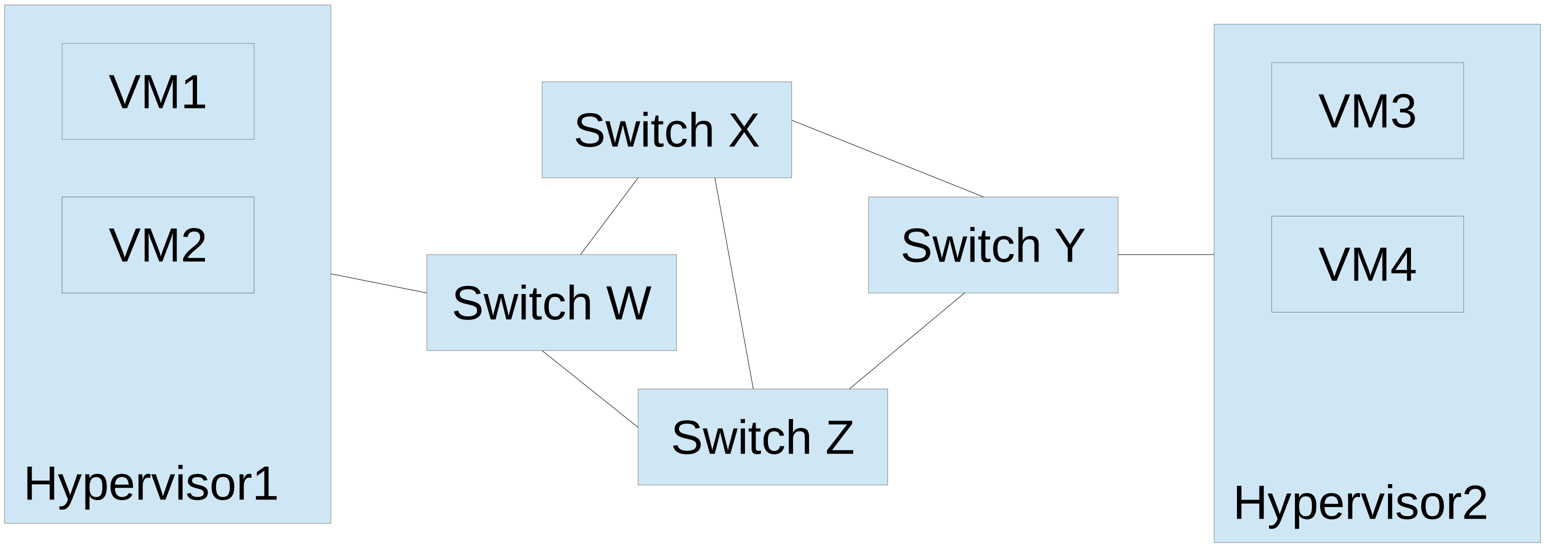
How is link layer connectivity computed?

- E.g. Spanning Tree Protocol (IEEE 802.1D)
- A distributed protocol that deals with distributed state
- May not result in the global optimal solution
- Defines its own protocol format, ...
- Must deal with packet loss, …

- Where are the abstractions?
- How do we influence the resulting connectivity structure?
- Can't we manage it centrally?

A possible SDN use-case

- In your datacenter, you know your traffic flows. It is your datacenter!
- How can you optimize your traffic flows?
  - VM1 to VM3 should flow via W → X → Y
  - VM2 to VM 4 should flow via W → Z → Y



Conclusion - The Problem

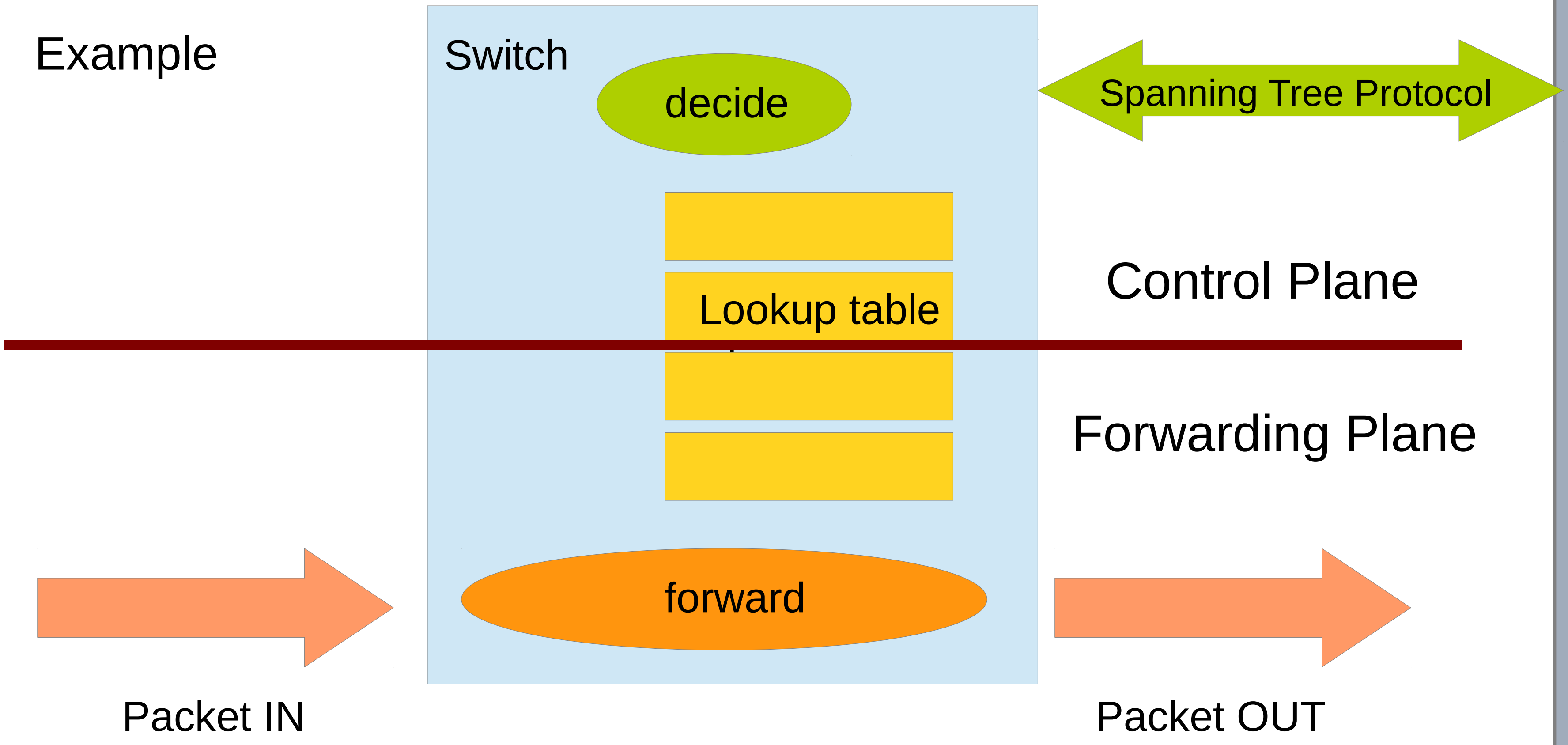
- The use cases
  - A centrally managed network
  - With scenario-specific requirements
    i.e. plugging together some switches is not enough, we have specific requirements that should be implemented by the switches
- The problem:
  - No abstractions or layers
  - No easy-to-use high-level APIs
  - No comfortable way to centrally manage your network
- The idea to solve the problem
  - Software defined networking

Preliminaries

Forwarding Plane and Data Plane

- Forwarding plane: Forwards packets
  - E.g. according to rules
- Control plane: Makes the decision what to do with packets
  - E.g. sets up forwarding plane rules

Example



Software Defined Networking

The Idea

A brave idea: Software Defined Networking (SDN)

- What is SDN?

A network in which the control plane is physically separate from the forwarding plane

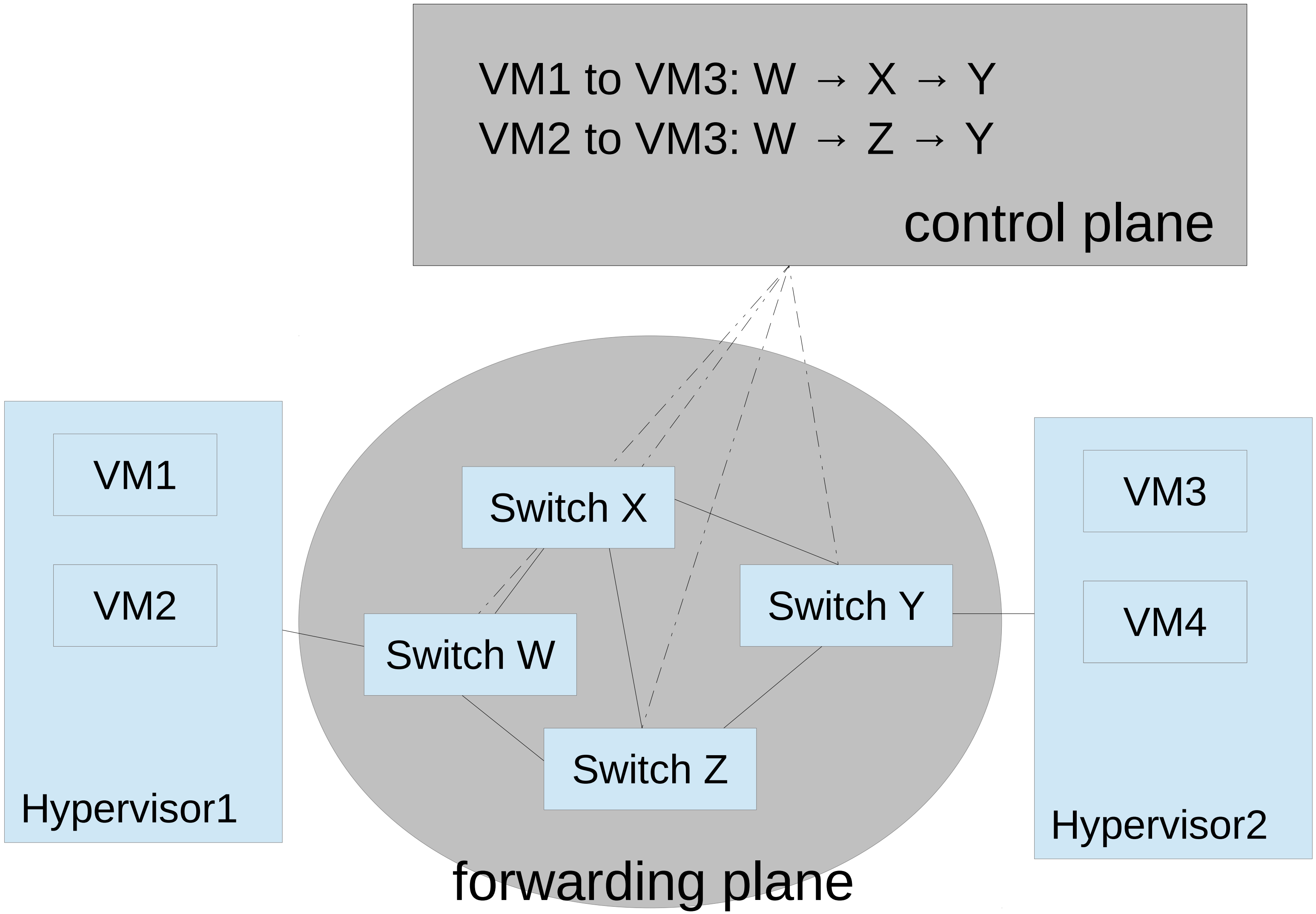
and

A single control plane controls several forwarding devices.

[Keown13]

- Forwarding plane: Forwards packets
  - E.g. according to rules
- Control plane: Makes the decision what to do with packets
  - E.g. sets up forwarding plane rules

Example



VM1 to VM3: W → X → Y
VM2 to VM3: W → Z → Y
control plane
VM1
VM2
Hypervisor1
Switch X
Switch W
Switch Y
Switch Z
forwarding plane
VM3
VM4
Hypervisor2

SDN Benefits

- Why the term `Software Defined'?
  - The control plane is just software.
- Abstraction
  - No distributed state, there is a **global network view** centrally at the control plane
  - No need to configure each forwarding plane device manually. Everything can be **managed centrally** at the control plane
  - Simple forwarding plane device configuration. A forwarding plane device model (like a **high-level API**) can be used to configure the devices. No need to develop a separate protocol, deal with packet loss, integrity of transferred data, distributed state, ...

SDN Benefits: No distributed state

- At a central point with a global view is programmed
  - Complex protocols such as the Spanning Tree Protocol are no longer necessary
  - A simple Dijkstra algorithm suffices
  - Globally optimal solutions can be computed
- Complexity is removed from the control plane

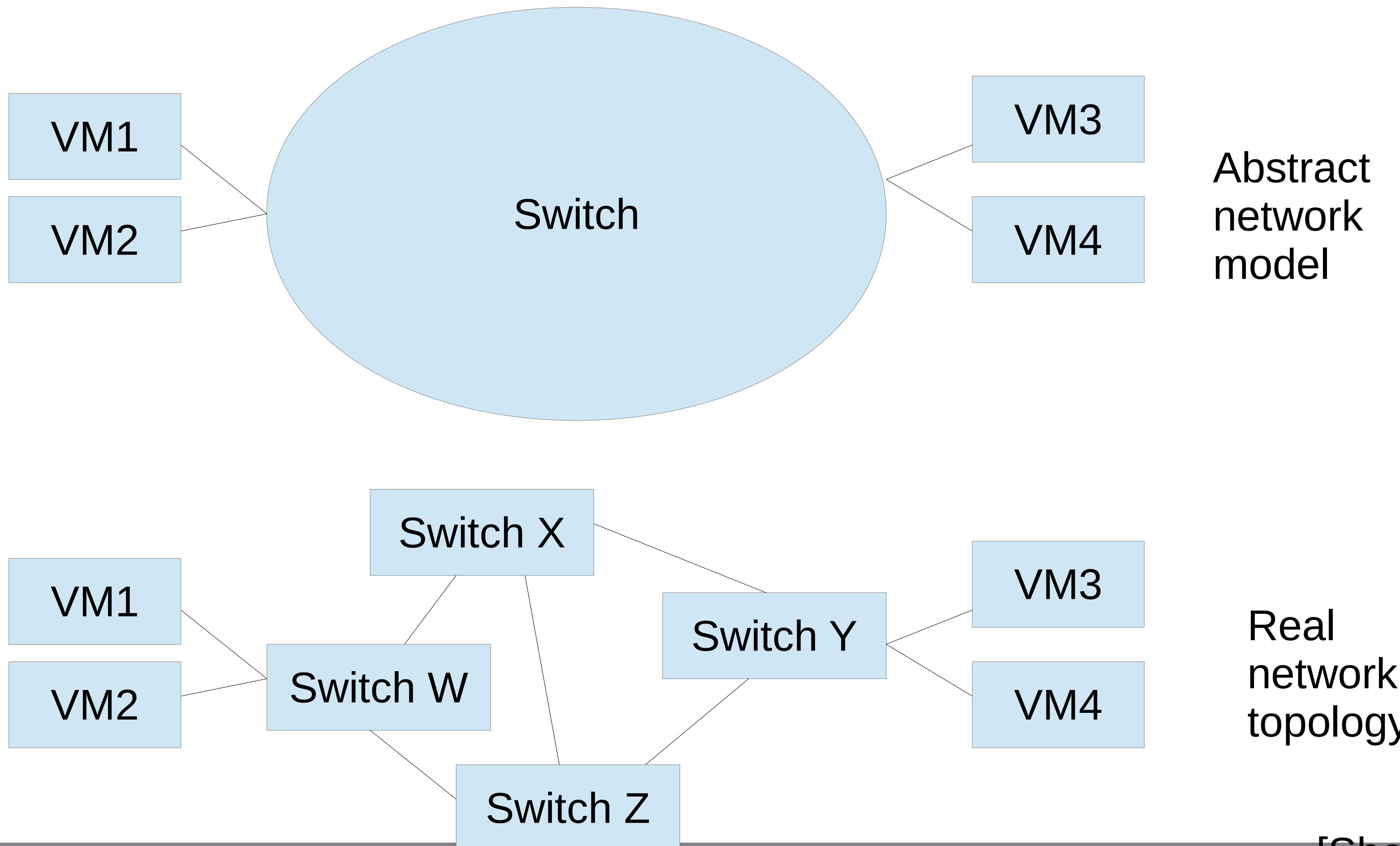
SDN Benefits: Specification Abstraction

[Shenker11]

- Control program **should** express desired behavior
- It **should not** be responsible for implementing that behavior on physical network infrastructure
- Natural abstraction: simplified model of network

SDN Benefits: Specification Abstraction - Example

Access Control
Desired behavior: VM1 cannot talk to VM3



[Shenker11]

SDN Benefits: Forwarding Abstraction

- Control plane needs **flexible** forwarding model
- Abstraction **should not** constrain control program
  - Should support whatever forwarding behaviors needed
- It **should** hide details of underlying hardware
  - Crucial for evolving beyond vendor-specific solutions

[Shenker11]

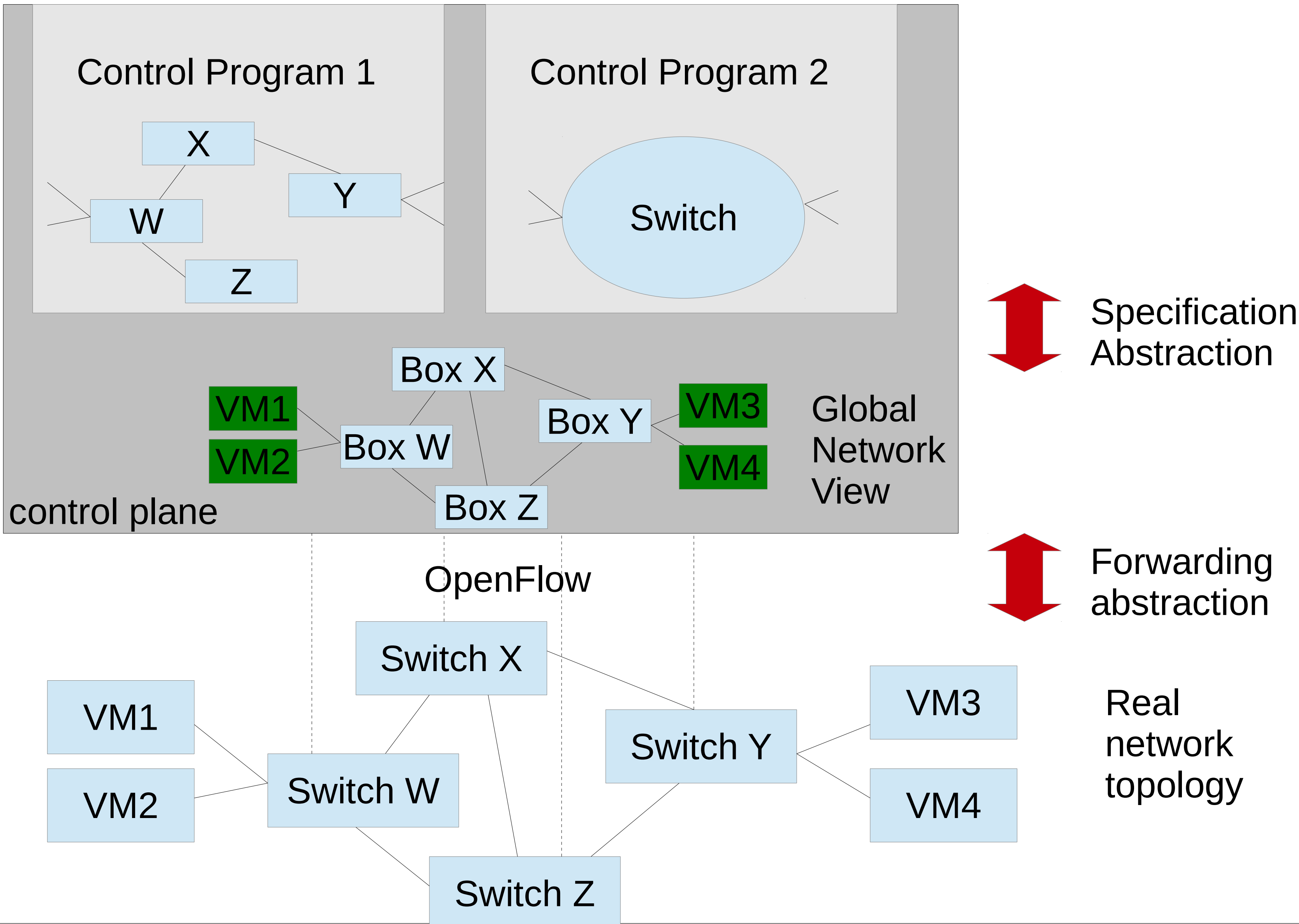
- The same interface for switches from different vendors

- Current standard and realization: OpenFlow

Sum up: Abstractions

- State abstraction
  - Global network view
- Specification abstraction
  - High level API to express desired behavior
- Forwarding abstraction
  - Simple model of forwarding `boxes'

SDN: The Big Picture



SDN: The Big Picture

Bottom-Up

- All forwarding plane devices (switches) are connected to the single control plane
- A common standard (OpenFlow) provides an abstraction such that all forwarding plane devices can be uniformly managed
- Instead of distributed state in the forwarding plane devices, one global network view is available
- Appropriate abstractions (specification abstraction), depending on the desired view, can be utilized to preprocess the global network view
  - E.g. one-big-switch for access control,
  - Complete topology and link speeds for spanning tree calculation
  - Complete topology and link utilization for load balancing
- The control program finally only defines the desired behavior

SDN: The Big Picture

Top-Down

- The control program defines the desired behavior
- The specification abstraction is reversed and maps the control program's output to the global network view / the global state
- A common standard is used to configure the forwarding devices according to the global state

Implementation

A Network Operating System
And
OpenFlow

Disclaimer

We are discussing an idea and scientific prototypes.

There currently exists no network operating system that is as widely used and comparable to common operating systems such as Linux.

Recommended reading (**1 screen page**):
Lee Doyle, *The return of the network operating system (NOS)*, Network World (US), Jan 2013
http://news.idg.no/cw/art.cfm?id=564A6F1B-EF7B-6318-5F43DCBF4BADE856

Recap: What is a `normal' Operating System?

An operating system

- Manages the hardware resources
  - Coordinate access to shared hardware resources.
    E.g. if there is only one printer, print *document1* first, then *document2*, don't try to print them interleaved
  - Manage the hardware
    E.g. put hard disk to sleep if idle, put packets from the NIC to memory, ...
- Ennobles the hardware
  - E.g. you can access /dev/sda as if it were a simple block device. You don't have to care about whether it is a SSD, HDD, or raid system
- Provides a standardized API to the hardware resources
  - E.g. you normally don't open /dev/sda, you have a file system to store and access data

What is a Network Operating System?

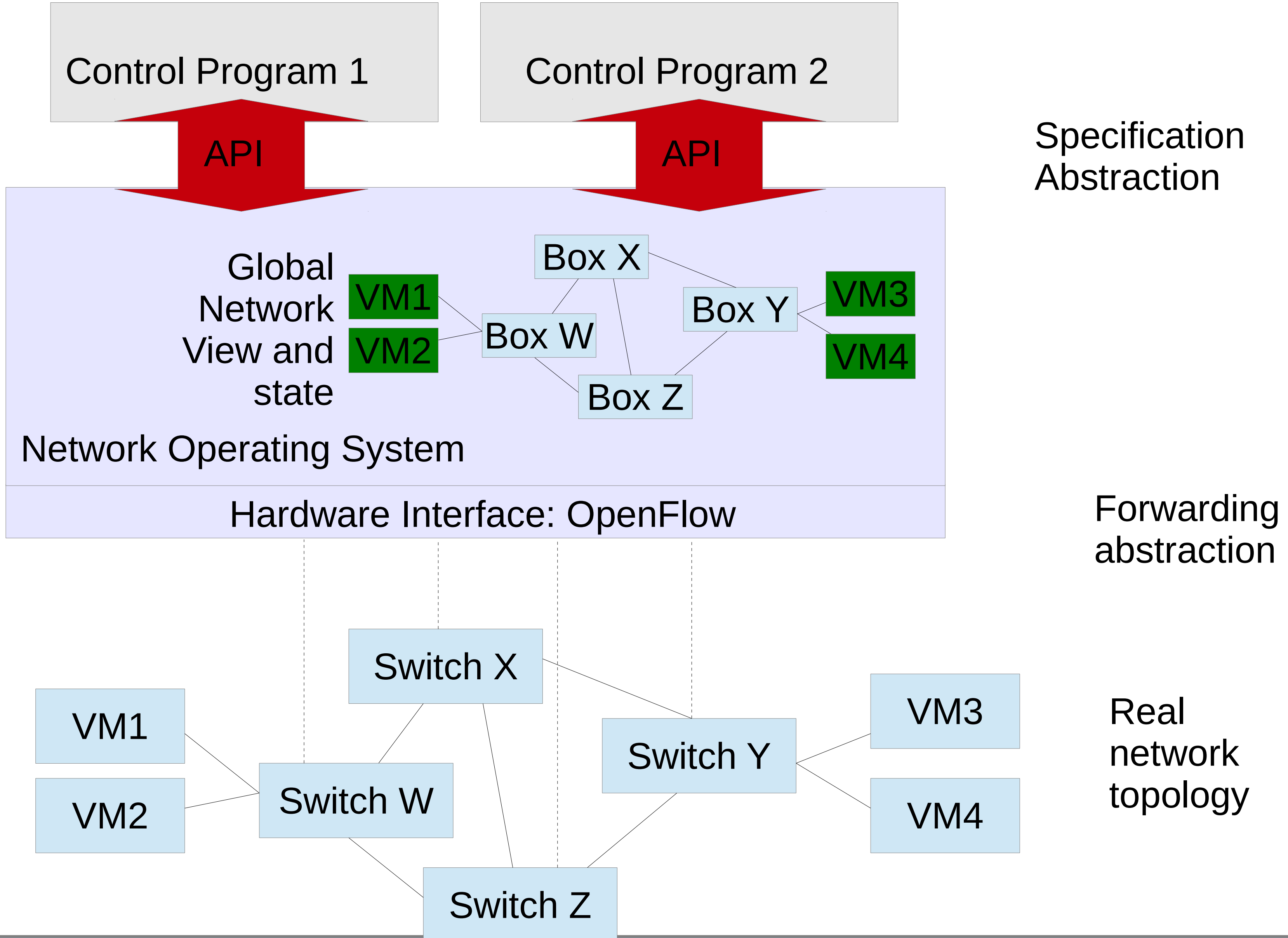
A network operating system

- Manages the hardware resources
  - E.g. all your switches in the network
- Ennobles the hardware
  - The global network view is a central place where all the state is stored and managed. As if there were no distributed state.
- Provides a standardized API to the hardware resources
  - Forwarding abstraction: Simple model of forwarding `boxes'
  - Specification abstraction: High-level API to express desired behavior

Operation of the hardware

- Recap: In a normal operating system
  - A device driver operates the hardware
  - For example via memory mapped areas, I/O instructions, ...
- In a network operating system
  - One needs to deploy the global network view to all the forwarding plane devices
  - The network operating system (control plane) is connected to all forwarding plane devices
  - Via a common protocol, the forwarding plane devices are programmed
  - This protocol is **OpenFlow**

Network Operating System: The Big Picture



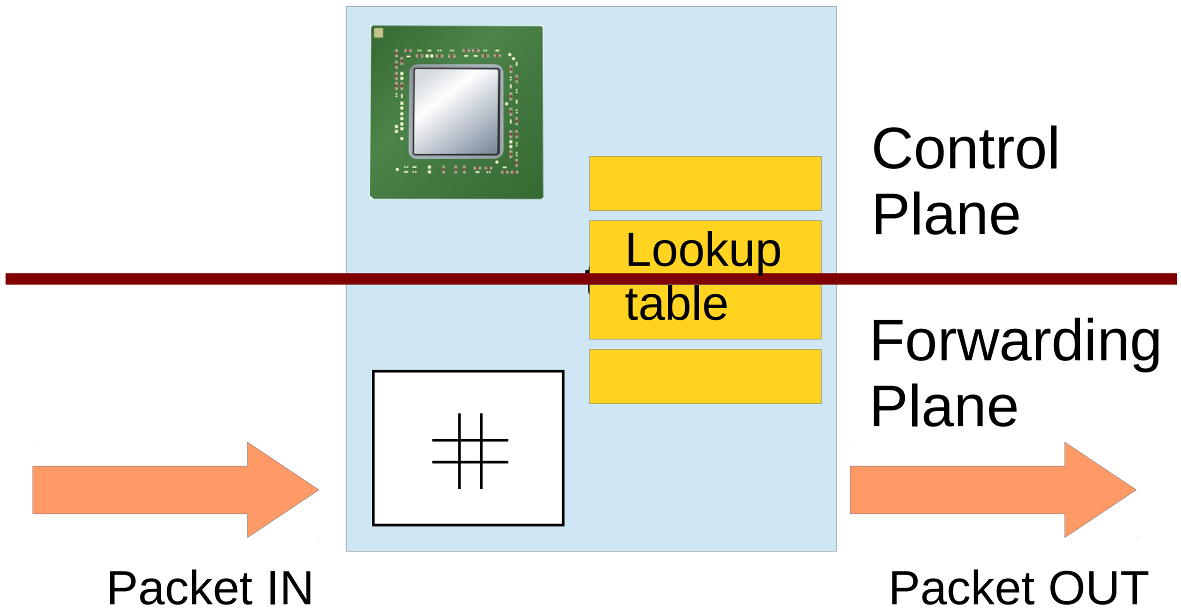
Implementation details

The hardware

Recap: today's common switches

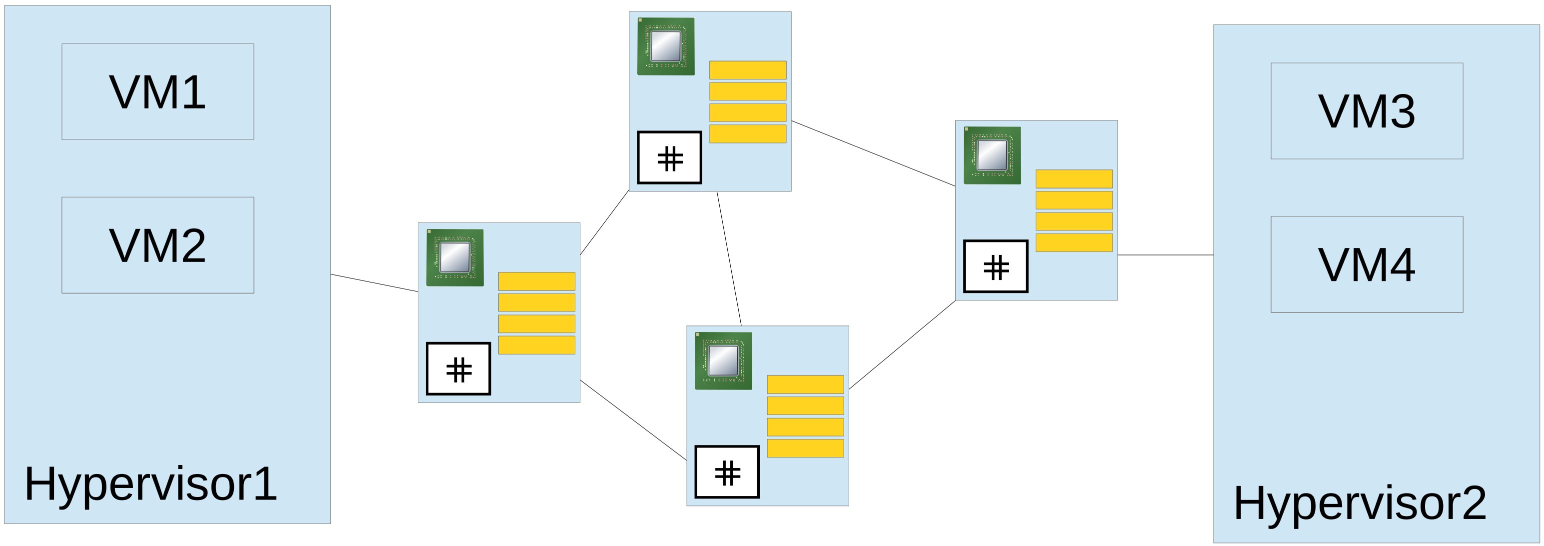
- Forwarding plane
  - Fast ASIC (application-specific integrated circuit)
  - I.e. special forwarding hardware
- Control plane
  - A (more or less) common CPU

Example



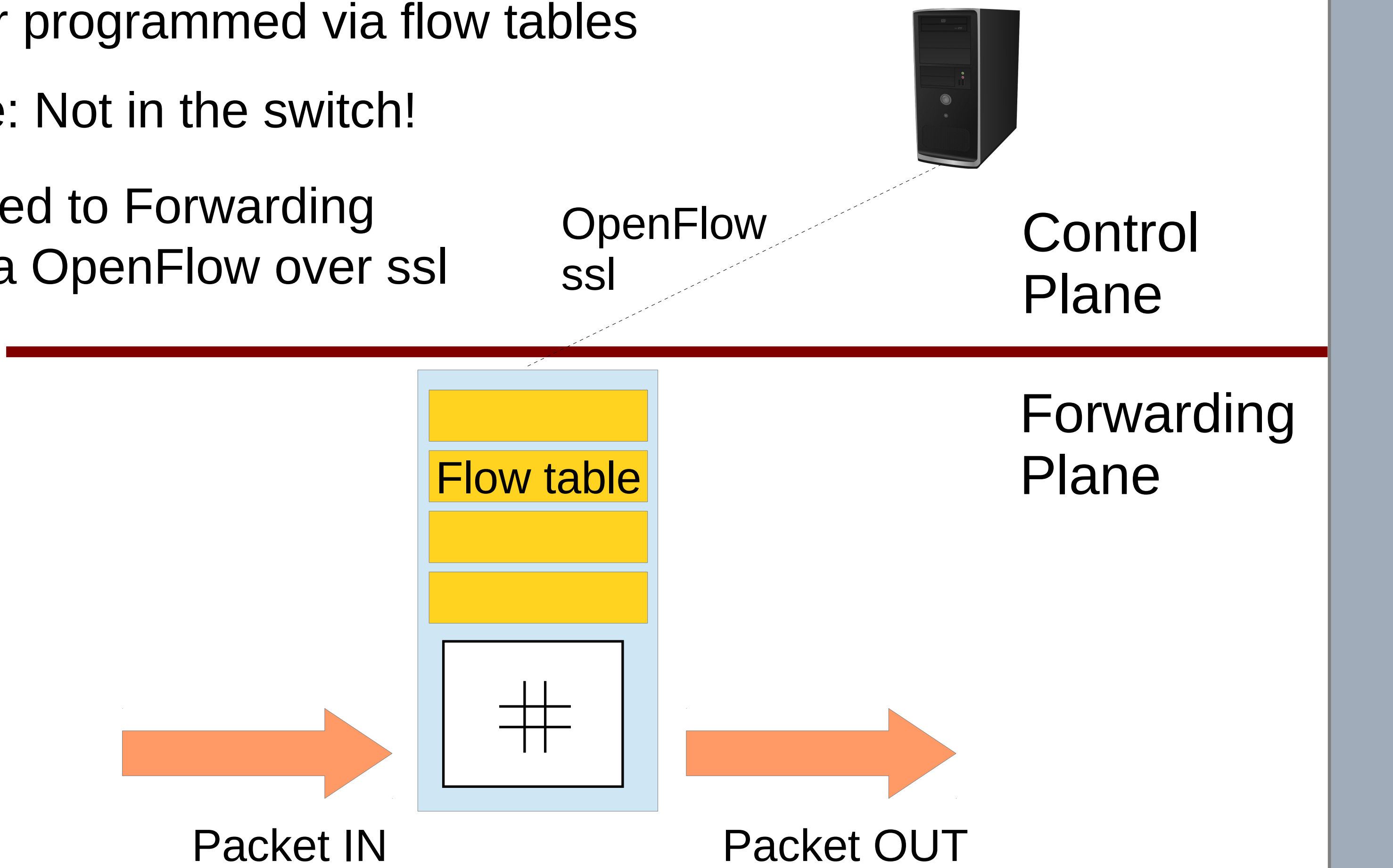
Recap: today's common switches

Forwarding plane and control plane distributed across the network



SDN switches

- Forwarding plane
  - Fast ASIC (application-specific integrated circuit)
  - I.e. special forwarding hardware
  - Behavior programmed via flow tables
- Control plane: Not in the switch!
  - Connected to Forwarding Plane via OpenFlow over ssl



An open switch - Discussion

A switch as open, programmable, forwarding-only platform

- Pros
  - Cheap, simple, fast but stupid devices
  - Allows innovation at software speed
  - Allows experimenting in real-world environments
  - Vendor independence
- Cons (possible vendor point of view)
  - Reveal switch internals
  - Open platforms lower the barrier-to-entry for new competitors
  - Opens the market, price pressure
  - Can sell less added value in their hardware
    Just stupid forwarding devices

An open switch - Discussion

Why not use commodity x86 PCs with Linux as open switches

- Pros
  - Open, available, well-tested
- Cons
  - Slow.
    Your memory bus is approximately completely jammed at 10 GB/s
  - Low port density.
    Did you recently see a PC with 100+ Ethernet ports?

→ Special forwarding hardware needed

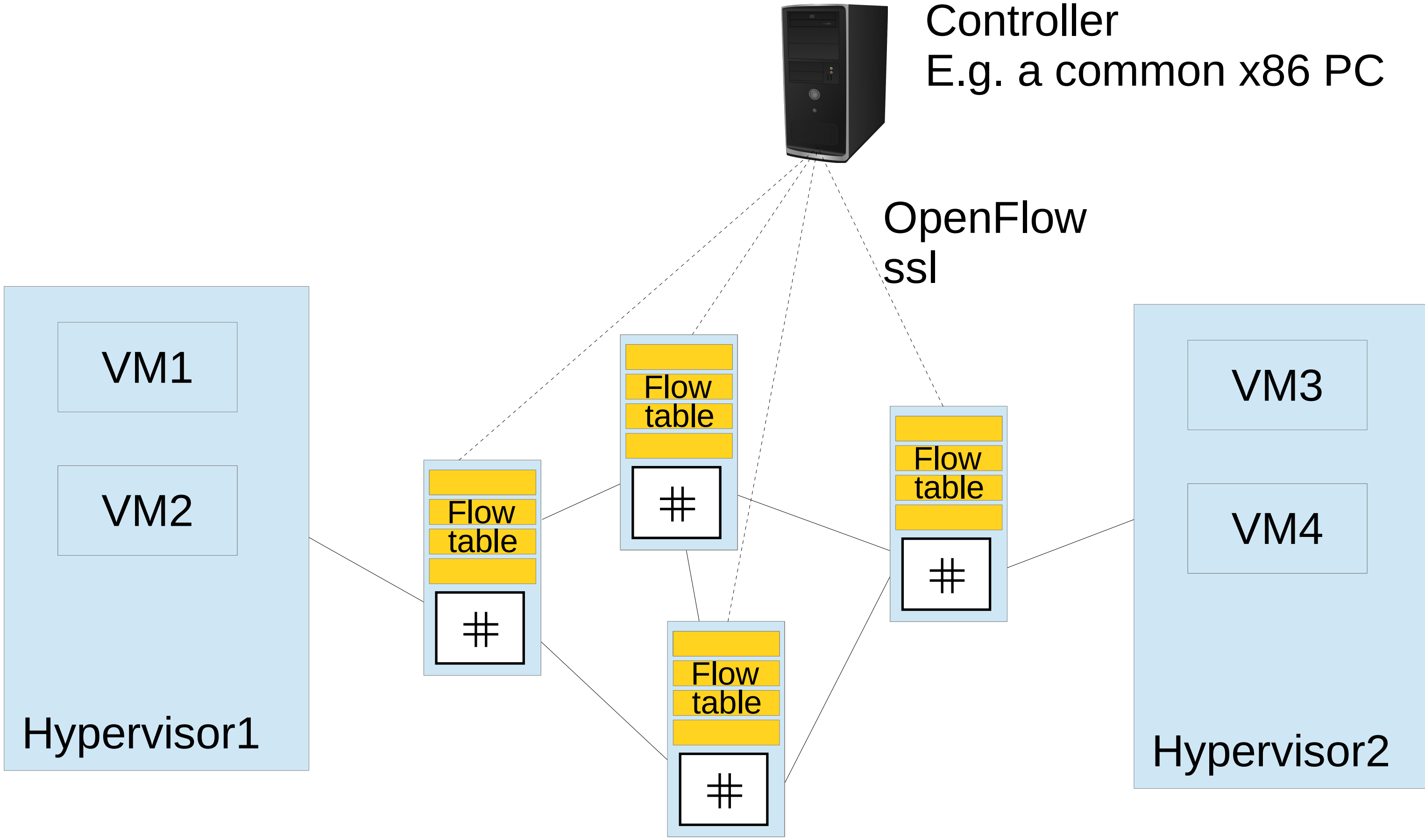
An open switch - Discussion

Why not use a commodity x86 PC as controller

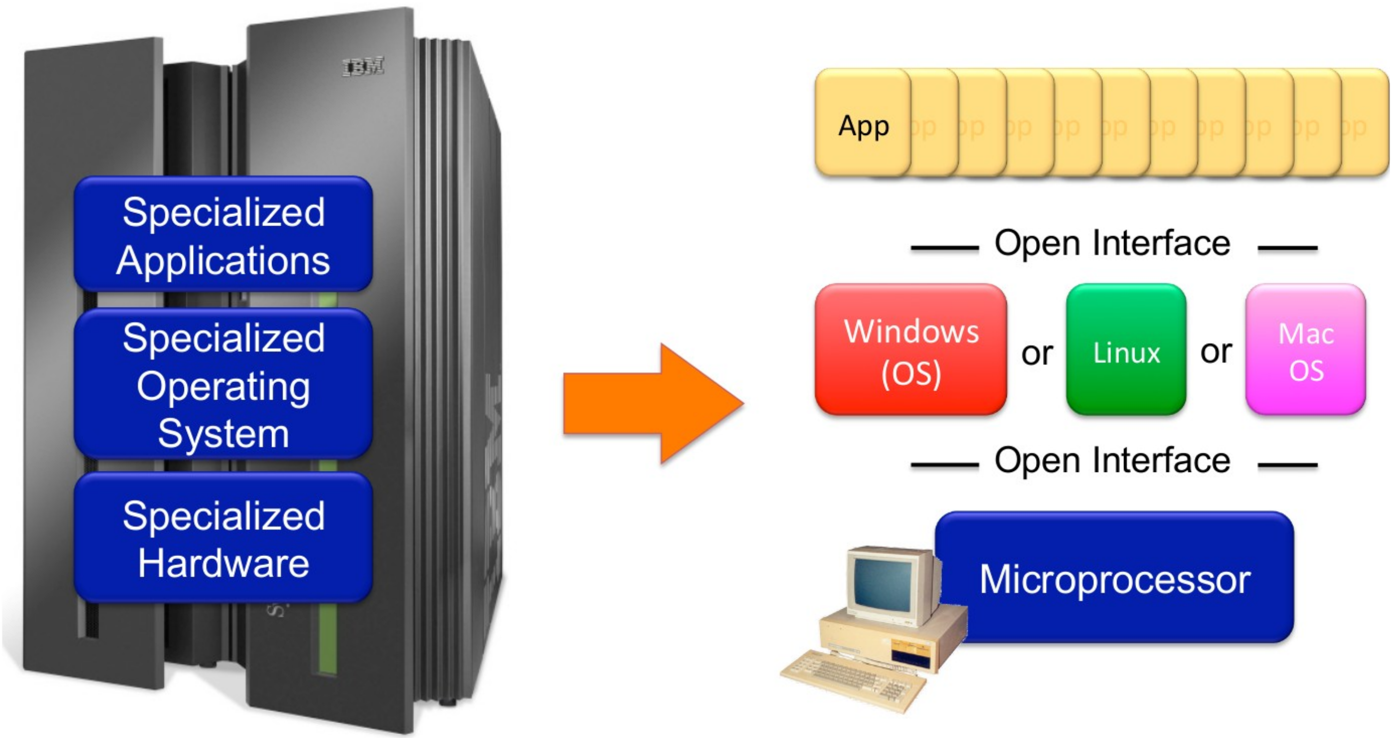
- Pros
  - Open, available, well-tested
- Cons
  - Reliability, but there are means to conquer this

→ You can use a simple x86 PC as your controller!

SDN Hardware: The Big Picture



A comparison: Progress in the Software Industry

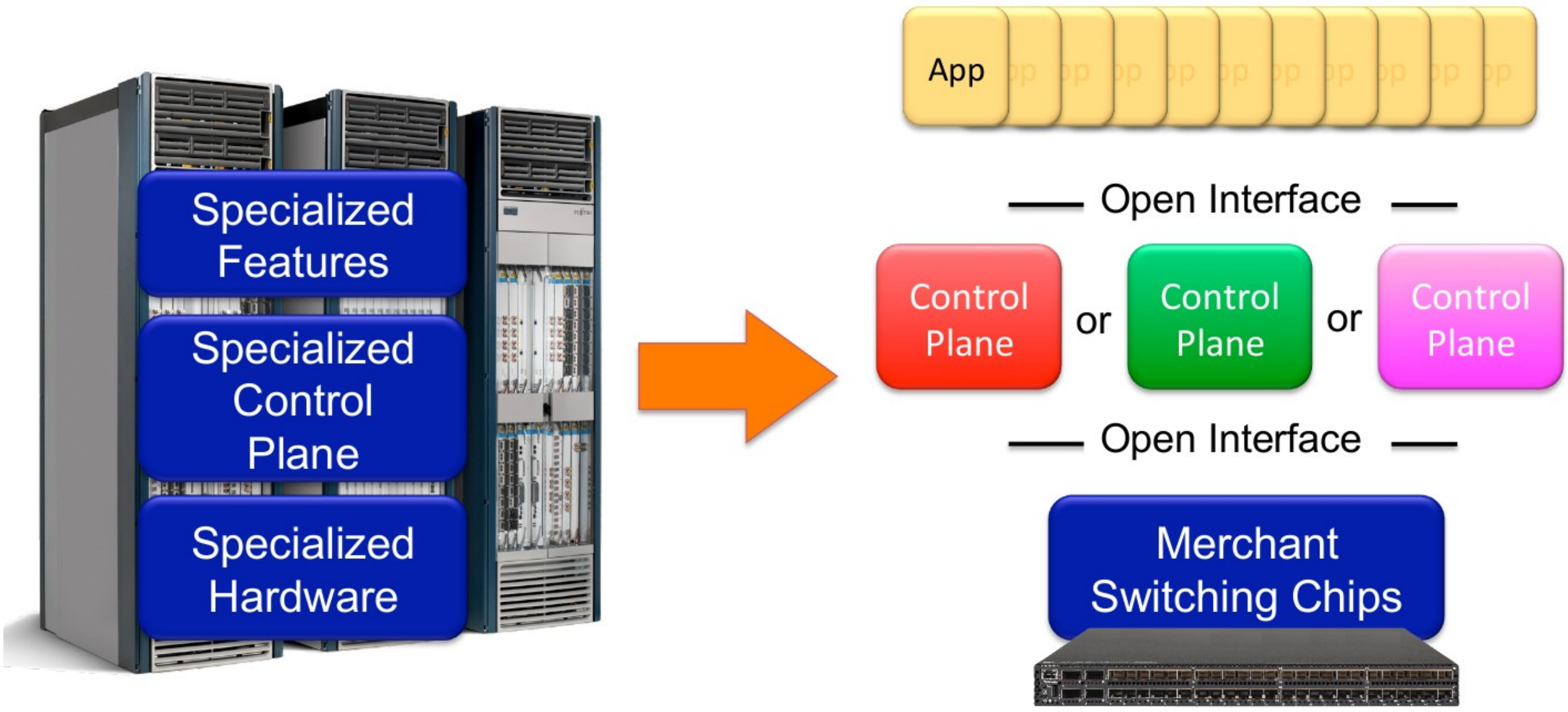


Vertically integrated
Closed, proprietary
Slow innovation
Small industry

Horizontal
Open interfaces
Rapid innovation
Huge industry

Slide borrowed from [Keown13]

The future of Networking?



Vertically integrated
Closed, proprietary
Slow innovation

Horizontal
Open interfaces
Rapid innovation

Slide borrowed from [Keown13]

Programming the forwarding plane

OpenFlow

OpenFlow Switches

An OpenFlow Switch consists of three components

- A Flow Table
  - Associates an action with a matching flow table entry
  - E.g. Match(src=1 and dst=2) Action(Forward(Port4))
- A Secure Channel that connects the switch to the controller
  - SSL
- The OpenFlow Protocol
  - An open and standardized way for a controller to program the switch
  - I.e. set up the flow table entries

[OFwp08]

OpenFlow Actions

If a packet matches a flow table entry, the following basic actions can be performed

- Forward
  - Forward packet to a switch's given port(s)
  - Used to move packets through network
- Drop
- Encapsulate
  - Encapsulate packet and send it via the secure channel to the controller
  - The controller decides what to do
  - Sane default setting if packet does not match a flow table entry
  - Controller can install appropriate flow table entry after the first packet of a flow was sent to it

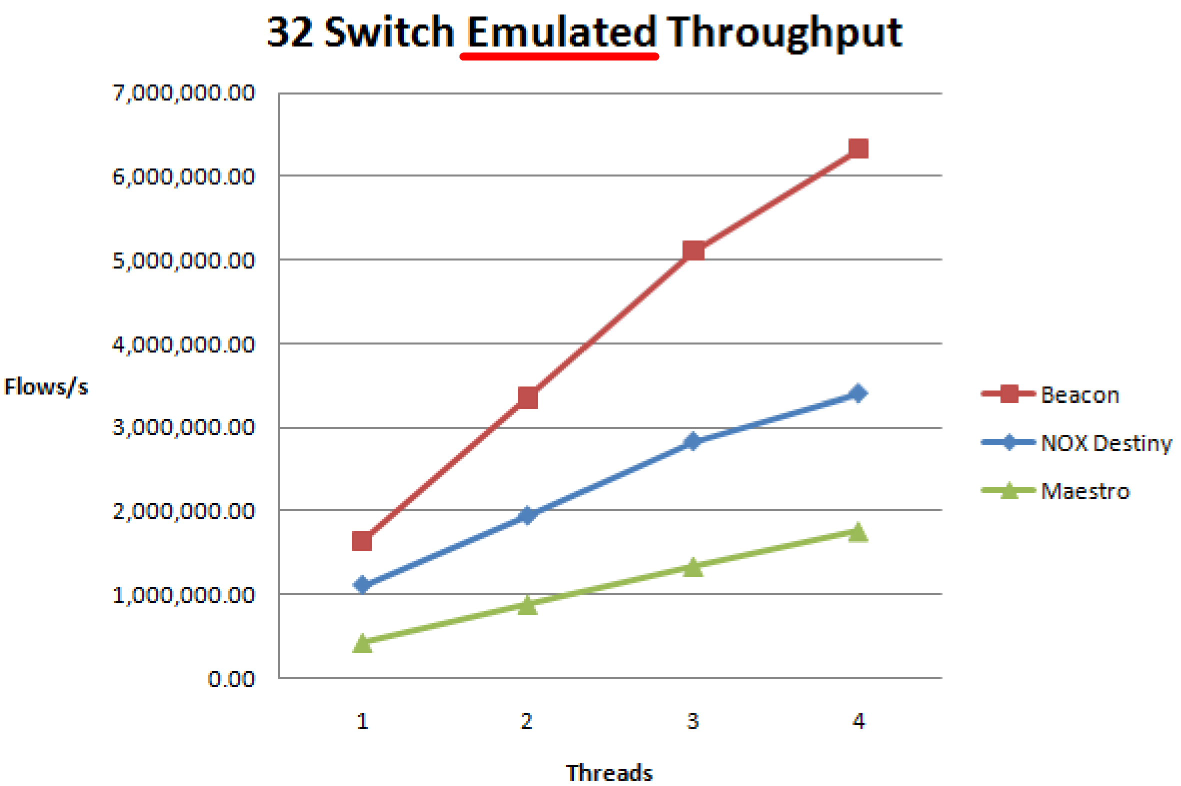
[OFwp08]

The Encapsulate Action

Is the encapsulate action a good choice? Discussion

- Observation
  - There may be many packets in a network but very few flows
  - A flow table entry is installed after the first packet of a flow has been observed
  - Afterwards, all packets that belong to the flow are forwarded by the switch directly
- Possible problems
  - When to delete old flow table entries?
  - How many flow table entries can a switch store? Is it enough?
  - What about attacks?
    Can an attacker send our arbitrary packets that match no flow table entry and thus congest the secure channel or overwhelm the controller?
    Think of port scanning!

OpenFlow Performance



CPU: 1 x Intel Core i7 930 @ 3.33ghz, 4 physical cores, 8 threads

RAM: 9GB

OS: Ubuntu 10.04.1 LTS x86_64

NOX, Beacon, and Maestro are controllers

Reminder

Always keep in mind: OpenFlow is just one tiny aspect of SDN and better alternatives may be thought of. But, you can buy OpenFlow switches!

Flow Table Entries

OpenFlow first generation

- Match
  - If a packet matches multiple entries, a priority decides the match
  - Only one match can apply to a packet but multiple actions can be performed per match
- Action
- Statistics
  - Byte and packet counter per flow

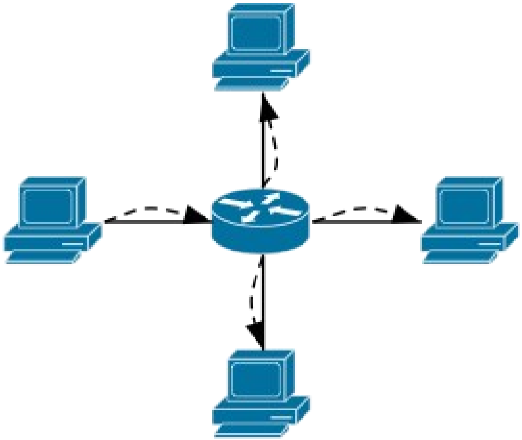
Match										Action	Statistics
In Port	VLAN ID	Ethernet			IPv4			TCP		Fwd Drop Encap	#pkts #bytes
		Src addr	Dst addr	Type	Src addr	Dst addr	Proto	Src port	Dst port		

Programming the Controller

A simple repeater

Forwards input packets out of all ports

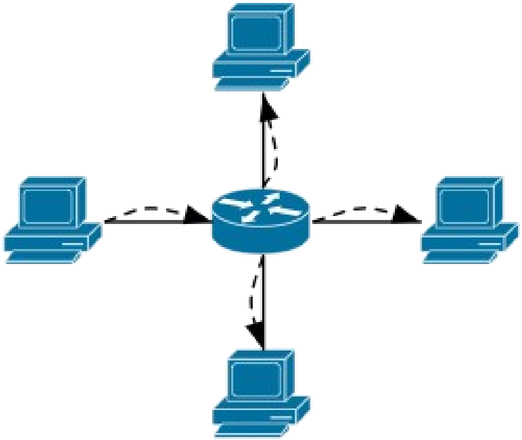
- We use a simple Ocaml controller library; in pythonized pseudo code
- We program the controller's `packet_in` callback which is called whenever an encapsulated packet is sent from the switch to the controller
- `send_packet_out` tells a switch (argument 1) to send out a packet



```
def packet_in (sw: switchId, pk: packetIn): unit =
  send_packet_out(sw,
    { output_payload = pk.input_payload;
      apply_actions = [Output AllPorts]
    }
   )
```

The examples are taken from
https://github.com/frenetic-lang/frenetic/wiki/Frenetic-Tutorial

A simple repeater - Discussion

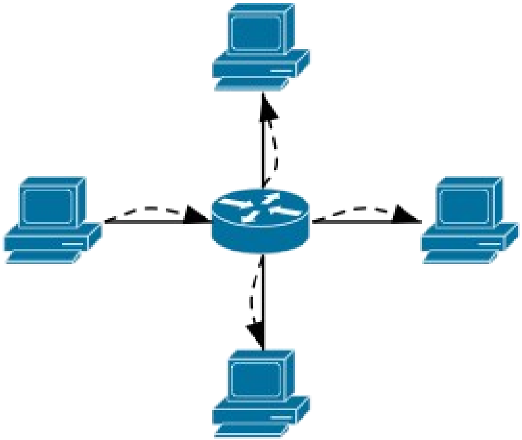


- This setup is extremely inefficient
- All packets that are received from the switch are forwarded to the controller.
- The controller decides the action
- No flow table entries are used, this is no better than using a common x86 PC as network switch!

```
def packet_in (sw: switchId, pk: packetIn): unit =
  send_packet_out(sw,
    { output_payload = pk.input_payload;
      apply_actions = [Output AllPorts]
    }
   )
```

The examples are taken from
https://github.com/frenetic-lang/frenetic/wiki/Frenetic-Tutorial

An efficient repeater



- We install a flow table entry when the switch is connected
- `add_flow(priority: nat,`
  `pattern: flow_table_match,`
  `action: flow_table_action)`

```
def switch_connected (sw: switchId): unit =
  send_flow_mod(sw, (add_flow 10 match_all [Output AllPorts]))

def packet_in (sw: switchId, pk: packetIn): unit =
  send_packet_out(sw,
    { output_payload = pk.input_payload;
      apply_actions = [Output AllPorts]
    }
   )
```

- Why is the `packet_in` kept?
  Assume you reboot the switch. Packets may arrive before the flow table entry is installed!

The examples are taken from
https://github.com/frenetic-lang/frenetic/wiki/Frenetic-Tutorial

An efficient repeater - Discussion

- The efficiency is still not very satisfying
- Incoming packets are flooded to all output ports
- This strategy resembles more to hubs than switches

- We will build a learning switch next
- Forwarding strategy:
  - When a packet with src MAC address *A* arrives at switch port *n*, we know that packets for device *A* should henceforth only be forwarded to port *n*.

A simple learning switch

```
/* a new dictionary that maps MAC addresses to switch ports */
KnownHosts: Map[MACaddr, Port] = new Map()


def packet_in(sw: switchId, pk: packetIn): unit =
  /* learn */
  KnownHosts[pk.MACaddr_src] =  pk.inPort

  /* forward packets */
  if KnownHosts.contains(pk.MACaddr_dst):
    out_port = KnownHosts[pk.MACaddr_dst]
    send_packet_out(sw, {
      output_payload = pk.input_payload;
      apply_actions = [Output (PhysicalPort out_port)]
    })
  else:
    /* unknown destination port, flooding to all */
    send_packet_out(sw,{
      output_payload = pk.input_payload;
      apply_actions = [Output AllPorts]
    })
```

The examples are taken from
https://github.com/frenetic-lang/frenetic/wiki/Frenetic-Tutorial

An efficient learning switch

```
/* a new dictionary that maps MAC addresses to switch ports */
KnownHosts: Map[MACaddr, Port] = new Map()


def packet_in(sw: switchId, pk: packetIn): unit =
  /* learn */
  KnownHosts[pk.MACaddr_src] =  pk.inPort

  /* forward packets and install flow table entries */
  if KnownHosts.contains(pk.MACaddr_dst):
    dst = pk.MACaddr_dst
    src = pk.MACaddr_src
    out_port = KnownHosts[dst]
    src_port = pk.inPort
    Match1 = {match_all with MACsrc = src and MACdst = dst}
    send_flow_mod(sw, (add_flow 10 Match1 [Output (PhysicalPort out_port)])
    Match2 = {match_all with MACSrc = dst and MACdst = src}
    send_flow_mod(sw, (add_flow 10 Match2 [Output (PhysicalPort src_port)])

    send_packet_out(sw, {
      output_payload = pk.input_payload;
      apply_actions = [Output (PhysicalPort out_port)]
    })
  else:
    … /* flooding as before */
```

The examples are taken from https://github.com/frenetic-lang/frenetic/wiki/Frenetic-Tutorial

An efficient learning switch - Discussion

- Once source and destination ports are known, flow table entries are installed

- Why are rules only installed if both ports are known? I.e.
  If A → B then Port(*m*)
  If B → A then Port(*n*)
- Imagine a packet with src MAC address *A* arrives at port *n* and one would install the rule
  If ***** → A then Port(n)
  Now, all packets, destined to *A* will be automatically forwarded by the switch. The controller will never see them.
  In particular, the controller will never learn the port for *B*

Programming Languages for Software Defined Networks

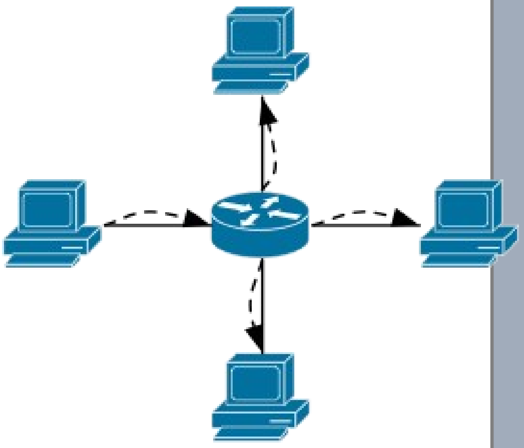
Introduction

Recall our hub example

- The controller code and the flow tables need to be written
- This is comparable to writing your C code and writing the same program in assembly again – by hand
- Programming languages for SDNs (such as frenetic) help out

```
let hub =
     if inPort = 1 then fwd(2,3,4)
  elsif inPort = 2 then fwd(1,3,4)
  elsif inPort = 3 then fwd(1,2,4)
  elsif inPort = 4 then fwd(1,2,3)
in hub
```

- This pseudo code can be compiled to a SDN controller **and** flow table entries



The examples are taken from
https://github.com/frenetic-lang/frenetic/wiki/Frenetic-Tutorial

Composition

- Composition of flow table entries is also a non-trivial task
- Imagine you want to count all packets with a srcIP *A,* using OpenFlow's statistics features (e.g. packet counter)
  If *srcIP* = *A* then *count*
- Also, you want to statically forward all packets with a dstIP *B* to Port 2
  If *dstIP* = *B* then *Port(2)*
- First statistics, then forwarding (sequential composition)
  If *srcIP* = *A* then *count*
  If *dstIP* = *B* then *Port(2)*
  Problem: all packets with srcIP A are counted and lost afterwards as the first matching rule **only** counts them. Recall: In the flow tables, packets can only match one rule.
- How do you apply both rules together (parallel composition)?
  If *srcIP* = *A* and *dstIP* = *B* then *count, Port(2)*
  If *srcIP* = *A* then *count*
  If *dstIP* = *B* then *Port(2)*

Composition

- Using composition operators, policies can be combined
  - Parallel composition `**|**'
  - Sequential Composition `**>>**'

[pyretic13]

- Assume we wrote a controller that collect statistics and one that does firewalling
- We want to collect the statistics in parallel with applying the firewalling. Afterwards, our learning switch should forward all packets the firewall let through
- Code:

```
( statistics | firewall ) >> learning_switch
```

Conclusion

Conclusion

- Software Defined Networking is an idea that focuses on
  - Centralized management
  - Abstractions
  - Innovation at software speed
- Implementation
  - Controller, specifies what we want
    - special programming languages (an active research area) provide easy-to-use means to program it
    - The controller runs on a
  - Network Operating System
    - provides easy-to-use API, keeps central state
    - and manages the hardware (forwarding plane) using
  - OpenFlow
    - an open standard which allows programming the forwarding plane devices

Bibliography

[Shenker11] Scott Shenker, *The Future of Networking, and the Past of Protocols.* Open Networking Summit 2011 http://www.youtube.com/watch?v=YHeyuD89n1Y

[Keown13] Nick McKeown, *Forwarding Plane Correctness*, Summer School on formal methods and networks, Cornell, 2013 http://www.cs.cornell.edu/conferences/formalnetworks/

[OFwp08] Nick McKeown et al., *OpenFlow: Enabling Innovation in Campus Networks,* ACM SIGCOMM Computer Communication Review, Apr 2008

[pyretic13] C. Monsanto et al., *Composing Software-Defined Networks*, 10th USENIX conference on Networked Systems Design and Implementation, 2013