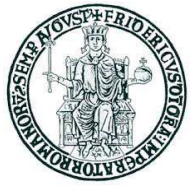
Cloud and Datacenter Networking

Università degli Studi di Napoli Federico II
Dipartimento di Ingegneria Elettrica e delle Tecnologie dell'Informazione DIETI
Laurea Magistrale in Ingegneria Informatica

Prof. Roberto Canonico

OpenFlow



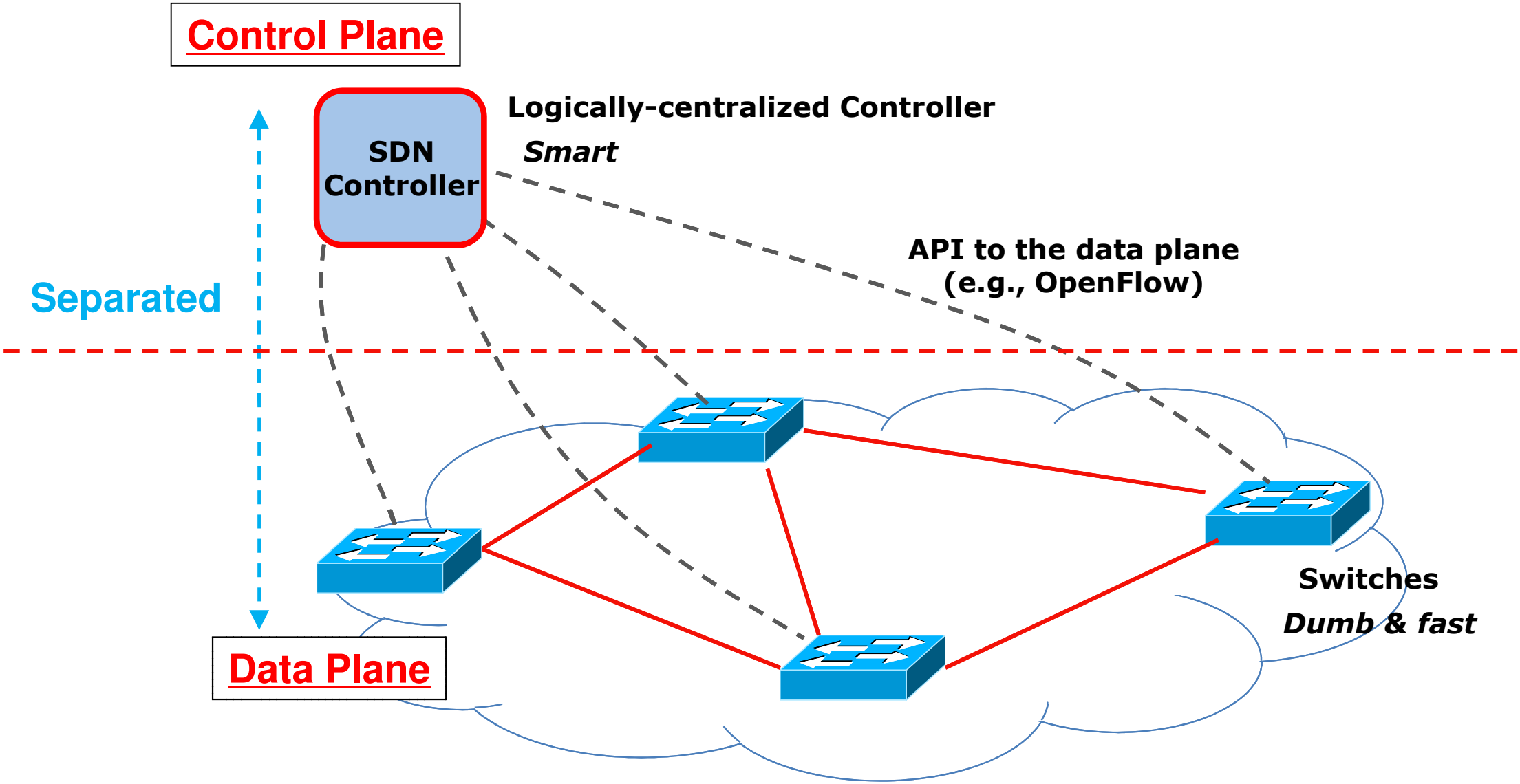
Lesson outline

- OpenFlow
- Credits for the material:
  - Jennifer Rexford
  - Nick McKeown
  - Srini Seetharaman
  - Scott Shenker

SDN Concept

- Separate control plane and data plane entities
  - Network intelligence and state are **logically centralized**
  - The underlying network infrastructure is **abstracted** from the applications
- Remotely control network devices from a central entity
- Execute or run control plane software on general purpose hardware
  - Decouple from specific networking hardware
  - Use commodity servers
- Expected advantages:
  - Ability to innovate through software
    - Overcome the “Internet ossification problem”
  - Cost reductions through increased competition, hardware commoditization and open-source software
- OpenFlow is the most popular implementation of the SDN paradigm

Software Defined Networking (SDN)



- A logically centralized "Controller" uses an open protocol to:
  - Get state information **from** forwarding elements (i.e. switches)
  - Give controls and directives **to** forwarding elements

What is OpenFlow

- OpenFlow is an *open* API that provides a standard interface for programming the data plane of switches
- OpenFlow assumes an SDN network model, i.e. separation of control plane and data plane
  - The datapath of an OpenFlow Switch consists of a Flow Table, and an action associated with each flow entry
  - The control path consists of a controller which programs the flow entry in the flow table

- But, SDN is not OpenFlow
  - OpenFlow is just one of many possible data plane forwarding abstraction

- Openflow standardization
  - Version 1.0: December 2009
  - Version 1.1: February 2011
  - OpenFlow transferred to ONF in March 2011
  - Version 1.5.0 Dec 19, 2014

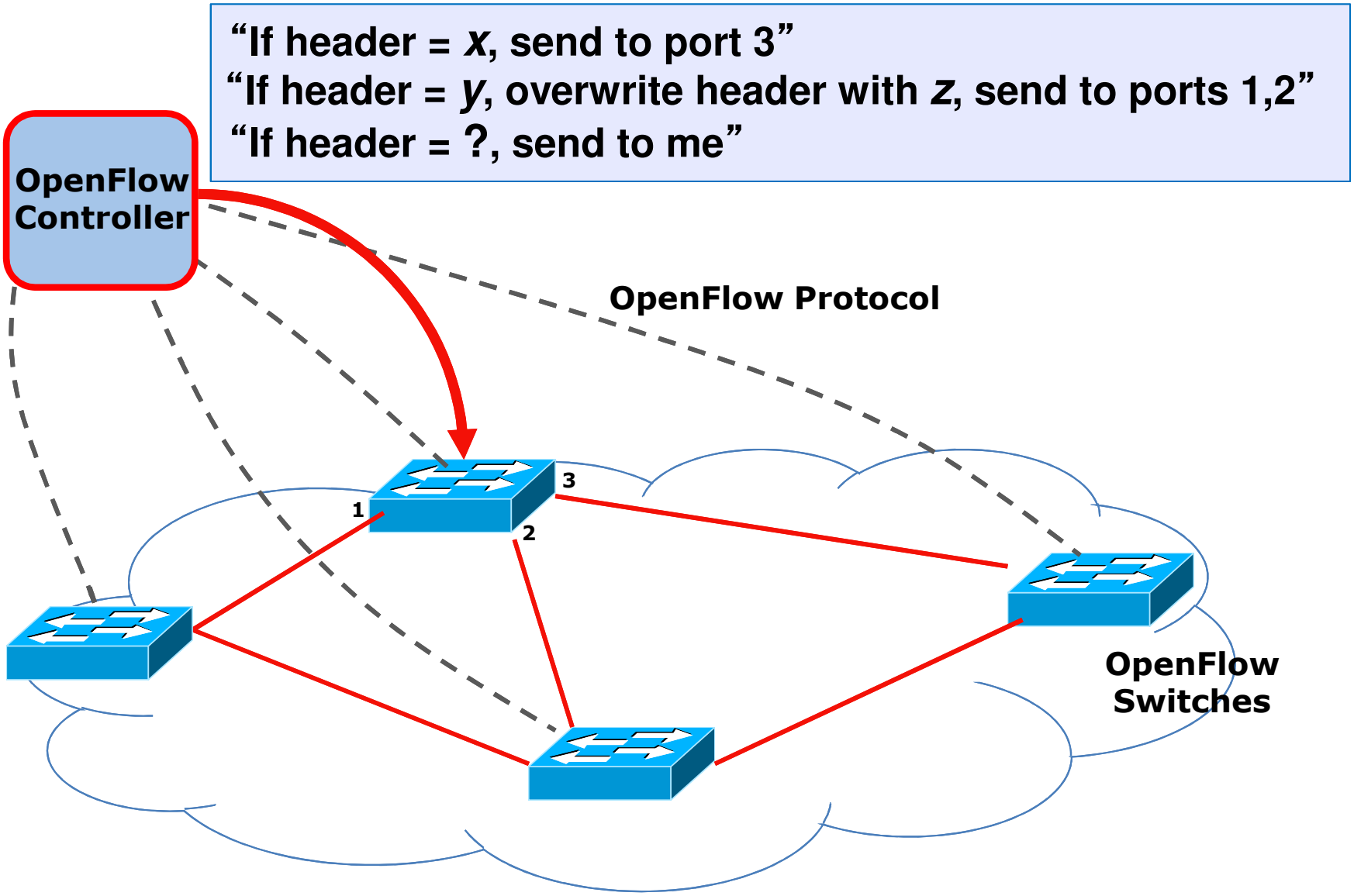
OpenFlow Consortium

http://OpenFlowSwitch.org

- Goal
  - Evangelize OpenFlow to vendors
  - Free membership for all researchers
  - Whitepaper, OpenFlow Switch Specification, Reference Designs
  - Licensing: Free for research and commercial use

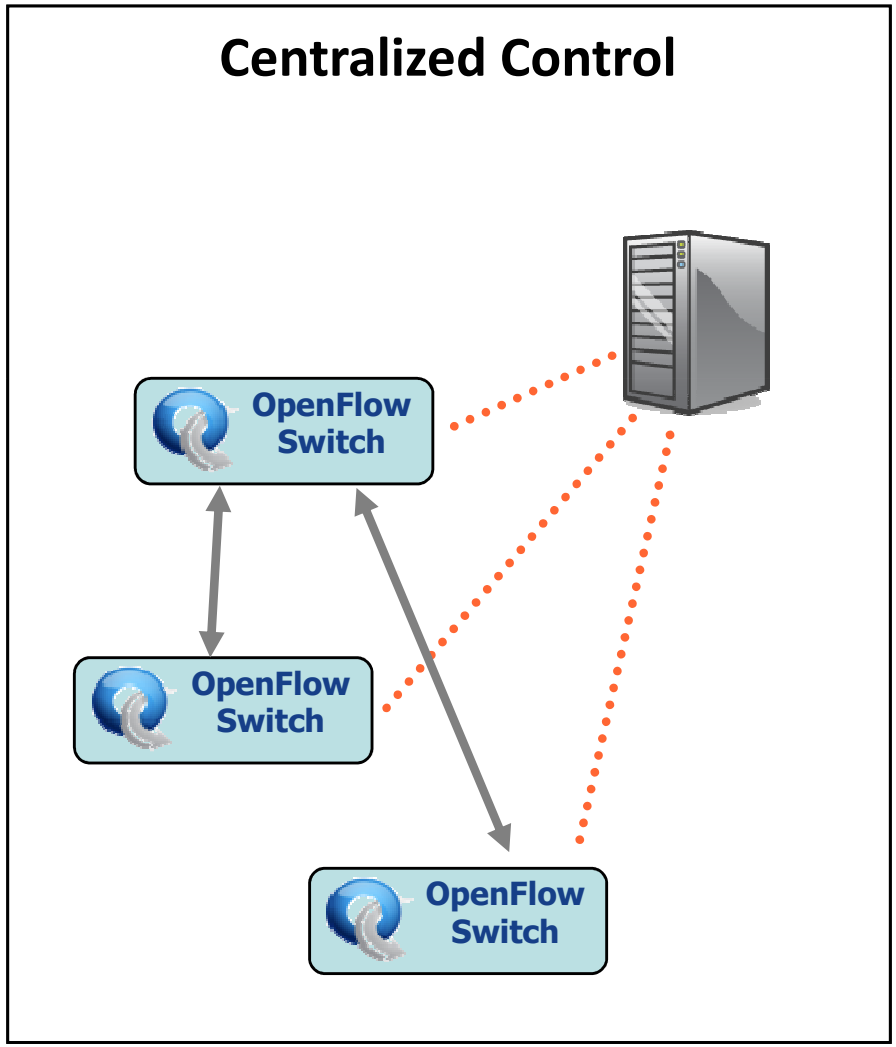
OpenFlow network model

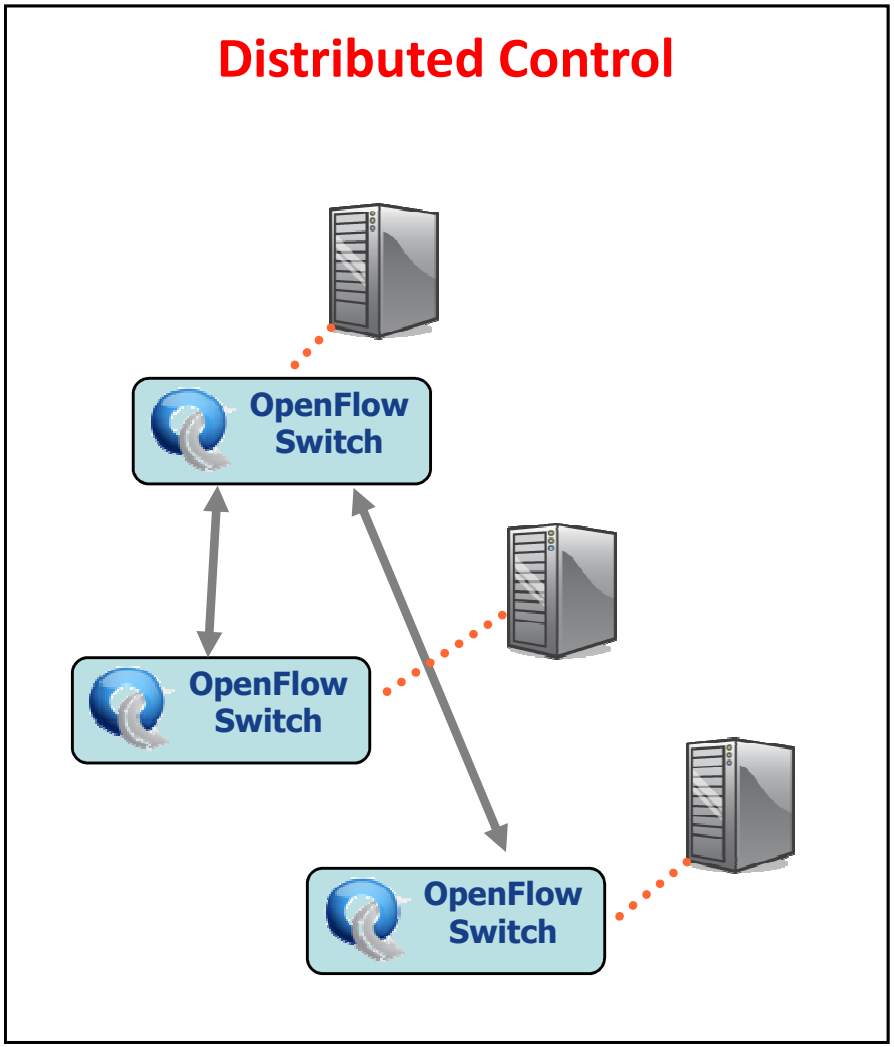
- The OpenFlow controller instructs switches about how they should process packets



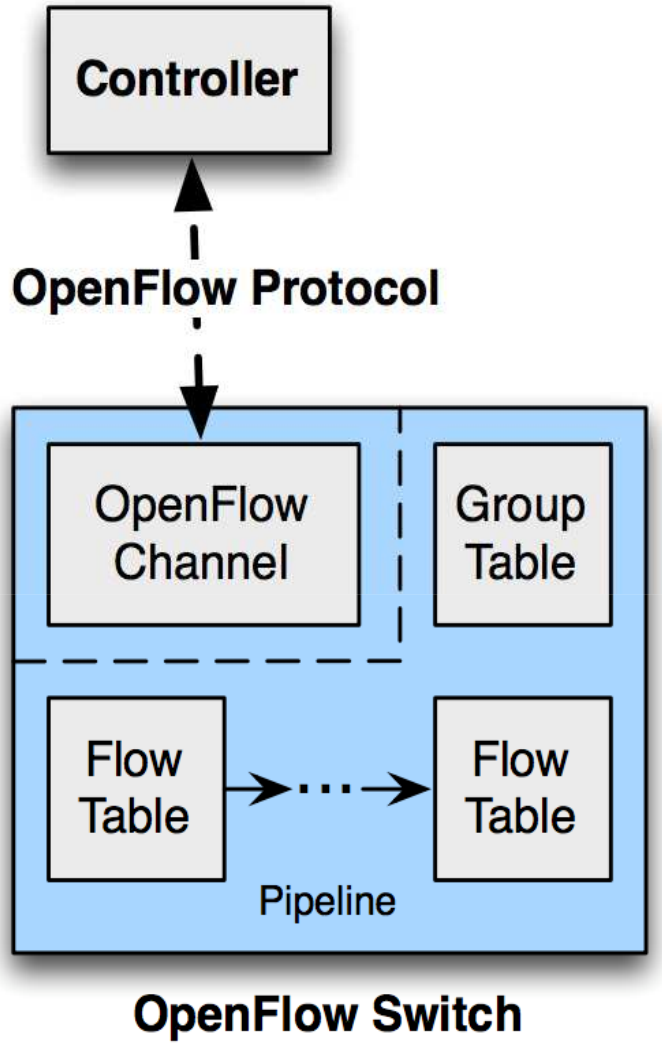
OpenFlow: centralized vs. distributed control

- Both models are possible with OpenFlow
  - Distributed control to reduce switch-controller latency and to avoid performance problems and a single-point-of-failure





OpenFlow switch: components



In current OpenFlow switches, Flow Tables are implemented by leveraging existing hardware components such as TCAMs (ternary content-addressable memory)

OpenFlow Switch specifications

- OpenFlow 1.0 (TS-001) – December 2009
  - https://www.opennetworking.org/images/stories/downloads/sdn-resources/onf-specifications/openflow/openflow-spec-v1.0.0.pdf
- OpenFlow 1.1 (TS-002) – February 2011
  - https://www.opennetworking.org/images/stories/downloads/sdn-resources/onf-specifications/openflow/openflow-spec-v1.1.0.pdf
- OpenFlow 1.2 (TS-003) – December 2011
  - https://www.opennetworking.org/images/stories/downloads/sdn-resources/onf-specifications/openflow/openflow-spec-v1.2.pdf
- OpenFlow 1.3.0 (TS-006) – June 2012
  - https://www.opennetworking.org/images/stories/downloads/sdn-resources/onf-specifications/openflow/openflow-spec-v1.3.0.pdf
    - OpenFlow 1.3.1 (TS-007) – September 2012
    - …
    - OpenFlow 1.3.5 (TS-023) – April 2015 [LINK]
- OpenFlow 1.4.0 (TS-012) – October 2013
  - https://www.opennetworking.org/images/stories/downloads/sdn-resources/onf-specifications/openflow/openflow-spec-v1.4.0.pdf
    - OpenFlow 1.4.1 (TS-024) – April 2015 [LINK]
- OpenFlow 1.5.0 (TS-020) – December 2014
  - https://www.opennetworking.org/images/stories/downloads/sdn-resources/onf-specifications/openflow/openflow-switch-v1.5.0.pdf
    - OpenFlow 1.5.1 (TS-025) – April 2015 [LINK]

OpenFlow datapath

- The OpenFlow specification defines three types of tables in the logical switch architecture
  1. A *Flow Table* matches incoming packets to a particular flow and specifies the functions that are to be performed on the packets
     - There may be multiple flow tables that operate in a pipeline fashion
  2. A flow table may direct a flow to a *Group Table,* which may trigger a variety of actions that affect one or more flows
  3. A *Meter Table* can trigger a variety of performance-related actions on a flow
- An OpenFlow switch process packets by associating them to ***flows***
- In general terms, a flow is a sequence of packets traversing a network that share a set of header field values
  - Curiously, this term is not defined in the OpenFlow specification

OpenFlow: Secure Channel (SC)

- SC is the interface that connects each OpenFlow switch to controller
- A controller configures and manages the switch via this interface
  - Receives events from the switch
  - Send packets out the switch
- SC establishes and terminates the connection between OpenFlow Switch and the controller using the procedures
  - Connection Setup
  - Connection Interrupt
- The SC connection is a TLS connection
  - Switch and controller mutually authenticate by exchanging certificates signed by a site-specific private key

OpenFlow: ports

- OpenFlow switches are connected through OpenFlow ports
  - Network interfaces to exchange packets with the rest of the network
- Types:
  - **Physical Ports**
    - Switch defined ports correspond to a hardware interface (e.g., map one-to-one to the Ethernet interfaces)
  - **Logical Ports**
    - Switch defined ports that do not correspond to a hardware switch interface (e.g. Tunnel-ID)
  - **Reserved Ports**
    - Defined by ONF 1.4.0
    - specify generic forwarding actions such as sending to the controller, flooding and forwarding using non-OpenFlow methods, such as normal switch processing

Ports - Reserved Port Types (Required)

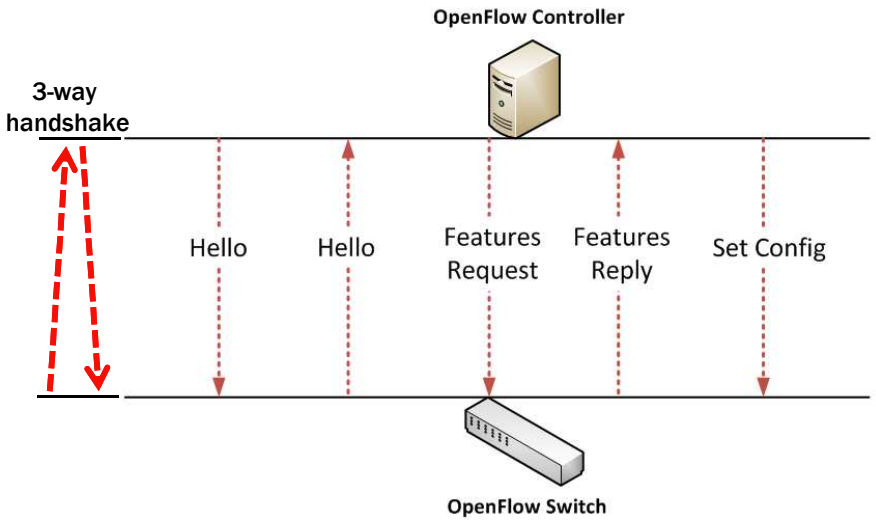
- ALL
  - Represents all ports the switch can use for forwarding a specific packets
  - Can be used only as output interface
- CONTROLLER
  - Represents the control channel with the OpenFlow controller
  - Can be used as an ingress port or as an output port
- TABLE
  - Represents the start of the OpenFlow pipeline
  - Submits the packet to the first flow table
- IN_PORT
  - Represents the packet ingress port
  - Can be used only as an output port
- ANY
  - Special value used in some OpenFlow commands when no port is specified
  - Can neither be used as an ingress port nor as an output port

Ports - Reserved Port Types (Optional)

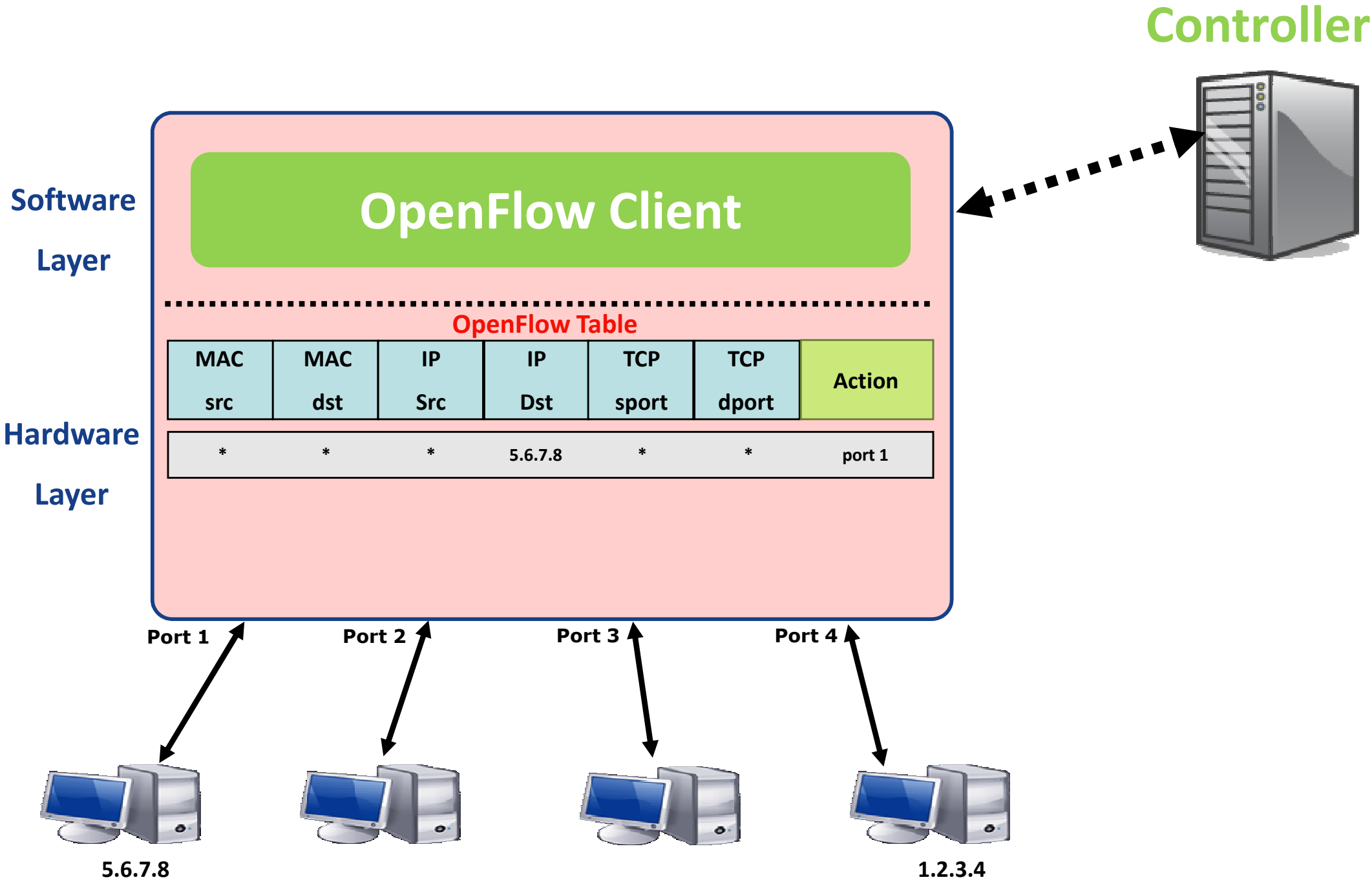
- LOCAL
  - Represents the switch's local networking stack and its management stack
  - Can be used as an ingress port or as an output port
- NORMAL
  - Represents the traditional non-OpenFlow pipeline of the switch
  - Can be used only as an output port and processes the packet using the normal pipeline
- FLOOD
  - Represents flooding using the normal pipeline
  - Can be used only as an output port
  - Send the packet out on all ports except the incoming port and the ports that are in blocked state

OpenFlow switch – Controller interactions

- An OpenFlow switch establishes a TCP connection to its Controller
  - An openFlow Controller by default listen on TCP port 6653 since OpenFlow 1.4.0
  - It used to be TCP port 6633 in previous OF versions
- Then the Controller starts an exchange of messages with the switch



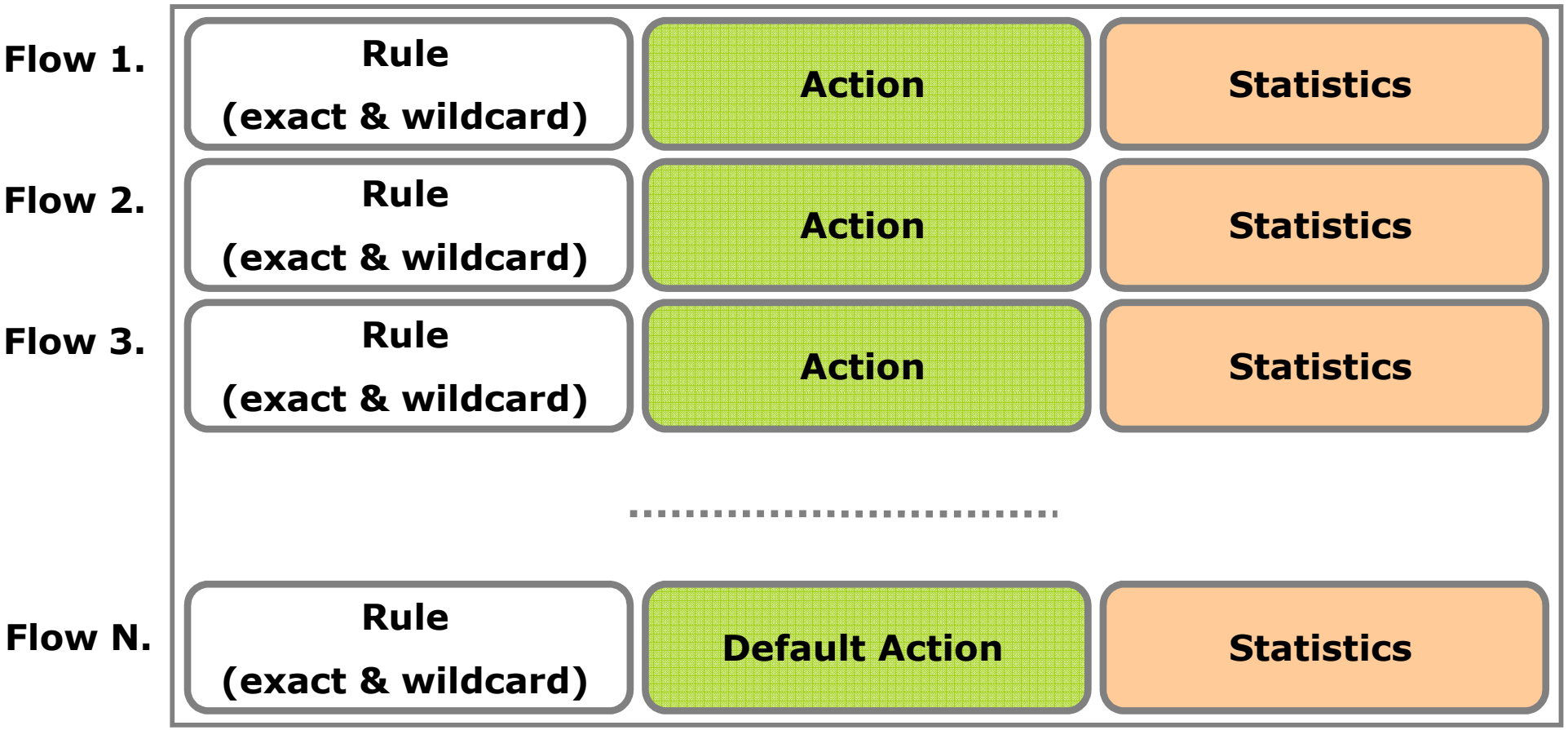
OpenFlow switching



OpenFlow switching

- The datapath of an OpenFlow Switch is governed by a Flow Table
- The control path consists of a Controller which programs the Flow Table
- The Flow Table consists of a number of ***flow entries***
- Each ***Flow Entry*** consists of
  - Match Fields
    - Match against packets
  - Action
    - Modify the action set or pipeline processing
  - Stats
    - Update the matching packets
- A Flow Table may include a **table-miss Flow Entry**, which renders all Match Fields wildcards (every field is a match regardless of value) and has the lowest priority (priority 0)

Flow Table


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

OpenFlow: packet processing

Match: 1000x01xx0101001x

Match arbitrary bits in headers:

- Match on any header, or new header
- Allows any flow granularity

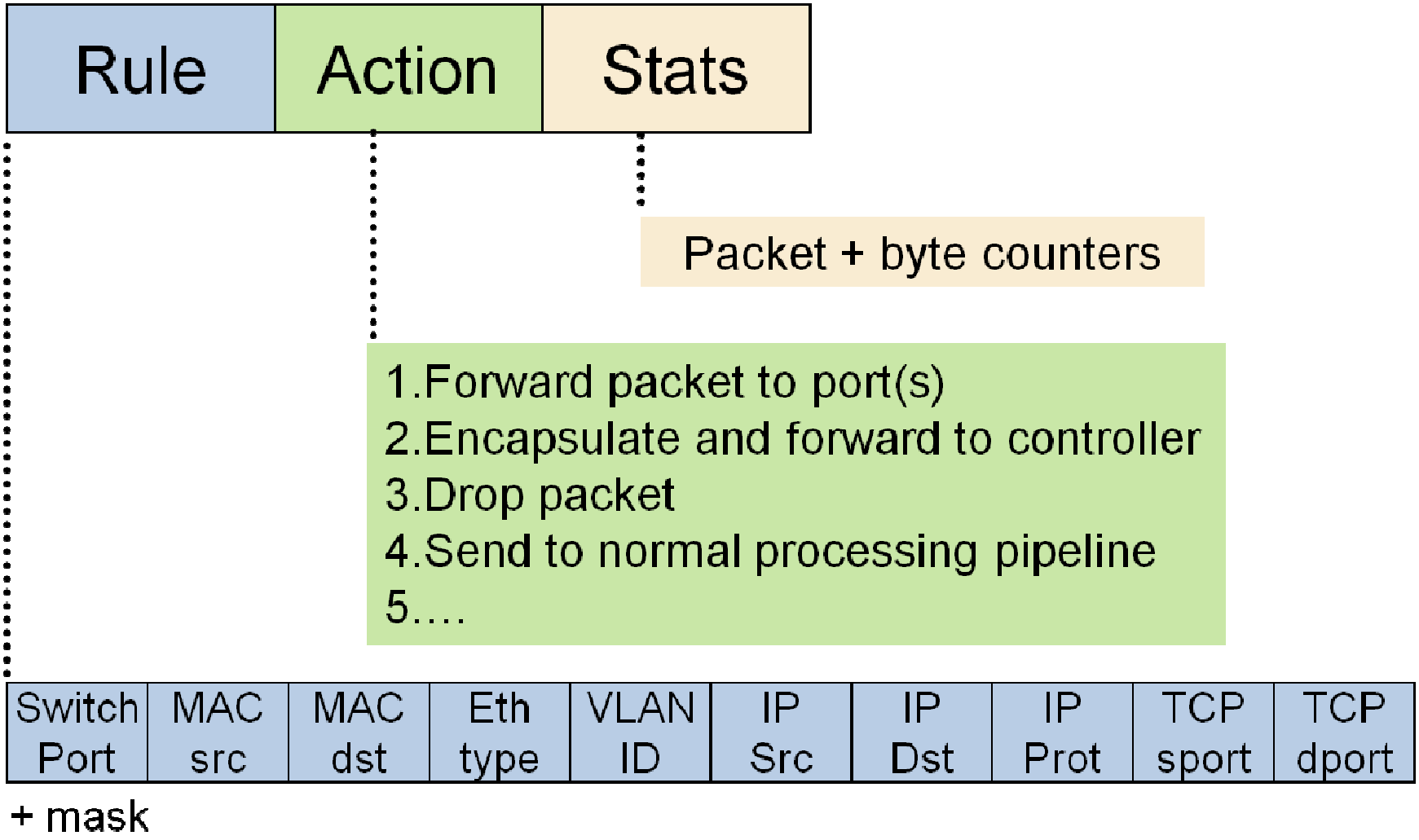
Action

- Forward to port(s), drop, send to controller
- Modify header
- …

OpenFlow actions

- Forward this flow's packets to a given port
  - This action allows packets to be routed
- Encapsulate and forward this flow's packets to a controller
  - This action allows controller to decide whether the flow should be added to the Flow Table
- Drop this flow's packets
  - This action can be used for security reasons, etc.
- Forward this flow's packets through the switch's normal processing pipeline (optional)
  - This action allows experimental traffic to be isolated from production traffic
  - Alternatively, isolation can be achieved through defining separate sets of VLANs
  - We can also treat OpenFlow as generalization of VLAN!
- Actions associated with flow entries may also direct packets to a ***group***
  - Groups represent sets of actions for flooding, as well as more complex forwarding semantics (e.g. multipath, fast reroute, and link aggregation)
  - As a general layer of indirection, groups also enable multiple flow entries to forward to a single identifier (e.g. IP forwarding to a common next hop)
  - This abstraction allows common output actions across flow entries to be changed efficiently

OpenFlow flow entry



Data-Plane: Simple Packet Handling

- Simple packet-handling rules
  - Pattern: match packet header bits
  - Actions: drop, forward, modify, send to controller
  - Priority: disambiguate overlapping patterns
  - Counters: #bytes and #packets

1. IP_src=1.2.*.*, IP_dest=3.4.5.* → drop
2. IP_src = *.*.*.*, IP_dest=3.4.*.* → forward to port 2
3. IP_src=10.1.2.3, IP_dest=*.*.*.* → send to controller

Overlapping rules !

OpenFlow examples

Switching

Switch Port	MAC src	MAC dst	Eth type	VLAN ID	IP Src	IP Dst	IP Prot	TCP sport	TCP dport	Action
*	*	00:1f:..	*	*	*	*	*	*	*	port6

Routing

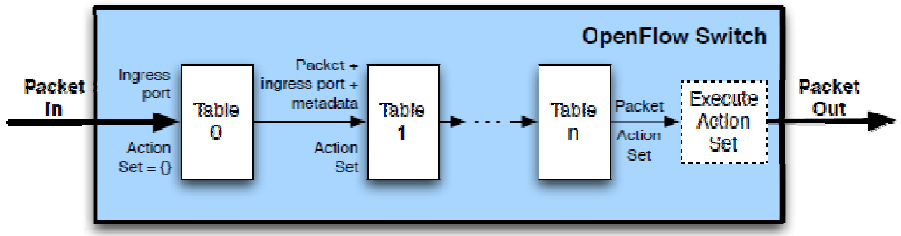
Switch Port	MAC src	MAC dst	Eth type	VLAN ID	IP Src	IP Dst	IP Prot	TCP sport	TCP dport	Action
*	*	*	*	*	*	5.6.7.8	*	*	*	port6

Firewall

Switch Port	MAC src	MAC dst	Eth type	VLAN ID	IP Src	IP Dst	IP Prot	TCP sport	TCP dport	Action
*	*	*	*	*	*	*	*	*	22	drop

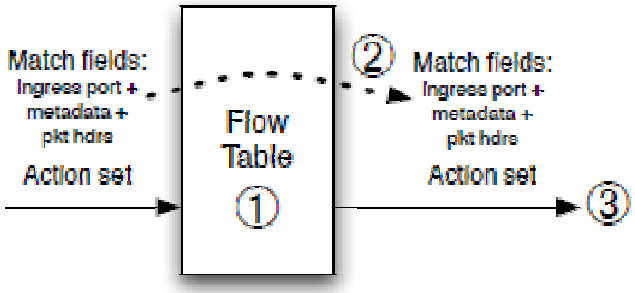
Flow Table pipelining (1)

- A switch includes one or more Flow Tables
- If there is more than one Flow Table, they are organized as a *pipeline*
- When a packet is presented to a Table for matching, the input consists of
  - the packet,
  - the identity of the ingress port,
  - the associated metadata value,
  - and the associated action set



(a) Packets are matched against multiple tables in the pipeline

- For Table 0, metadata value is blank and action set is null
- Each incoming packet is processed according to Flow Table entries
- A Flow Table entry may explicitly direct the packet to another Flow Table (using the Goto Instruction), where the same process is repeated again
- A flow entry can only direct a packet to a Flow Table number which is greater than its own flow table no.
  - Flow entries of the last Table of the pipeline cannot include the Goto instruction
- If the matching flow entry does not direct packets to another Flow Table, processing stops at this table. When pipeline processing stops, packet is processed with its associated action set and usually forwarded



① Find highest-priority matching flow entry

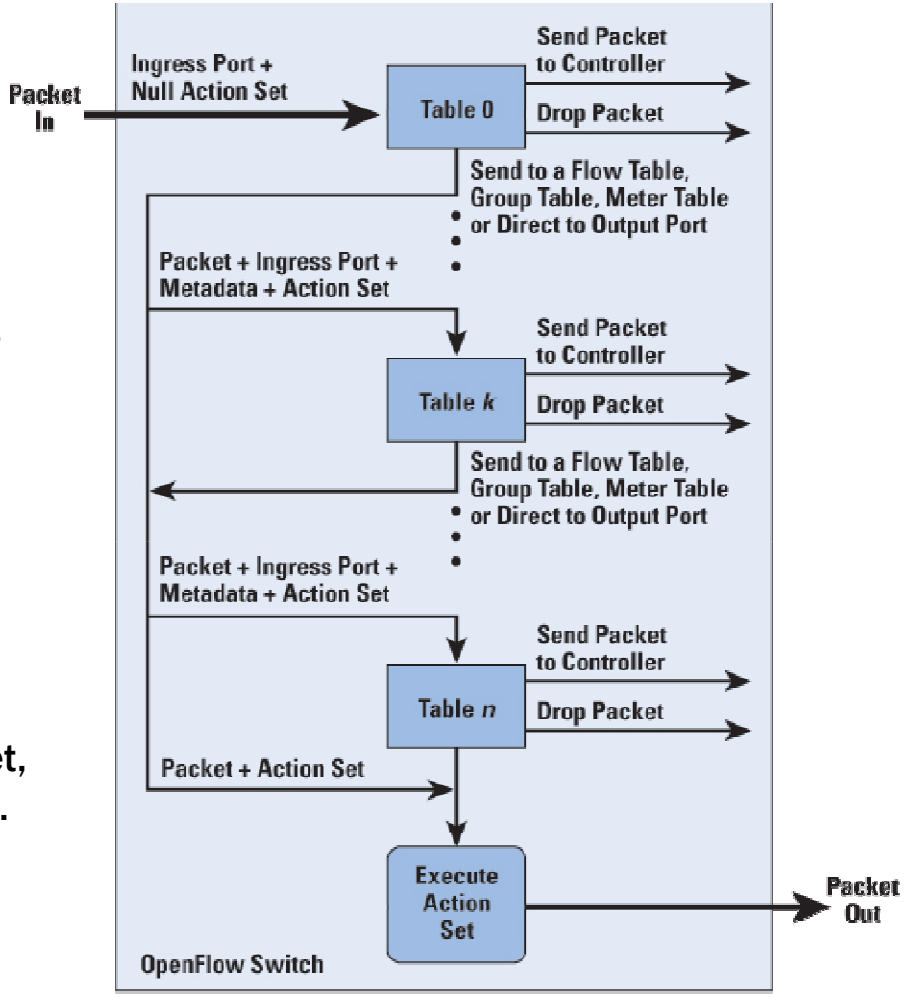
② Apply instructions:
- i. Modify packet & update match fields (apply actions instruction)
- ii. Update action set (clear actions and/or write actions instructions)
- iii. Update metadata

③ Send match data and action set to next table

(b) Per-table packet processing

Flow Table pipelining (2)

- At each table, find the highest-priority matching flow entry
  1. If there is no match on any entry and there is no table-miss entry, then the packet is dropped
  2. If there is a match only on a table-miss entry, then that entry specifies one of three actions:
     - Send packet to controller.
       This action will enable the controller to define a new flow for this and similar packets, or decide to drop the packet
     - Direct packet to another flow table farther down the pipeline
     - Drop the packet
  3. If there is a match on one or more entries other than the table-miss entry, then the match is defined to be with the highest-priority matching entry.
     The following actions may then be performed:
     - Update any counters associated with this entry.
     - Execute any instructions associated with this entry. These instructions may include updating the action set, updating the metadata value, and performing actions.
     - The packet is then forwarded to a flow table further down the pipeline, to the group table, or to the meter table, or it could be directed to an output port.
- If and when a packet is finally directed to an output port, the accumulated action set is executed and then the packet is queued for output



OpenFlow: Flow Routing vs. Aggregation

- Both models are possible with OpenFlow
  - Aggregated rules are necessary to cope with the hardware limit on number of entries imposed by current TCAMs

Flow-Based

- Every flow is individually set up by controller
- Exact-match flow entries
- Flow table contains one entry per flow
- Good for fine grain control, e.g. campus networks

Aggregated

- One flow entry covers large groups of flows
- Wildcard flow entries
- Flow table contains one entry per category of flows
- Good for large number of flows, e.g. backbone

OpenFlow: Reactive vs. Proactive management

- Both models are possible with OpenFlow

Reactive

- First packet of flow triggers controller to insert flow entries
- Efficient use of flow table
- Every flow incurs small additional flow setup time
- If control connection lost, switch has limited utility

Proactive

- Controller pre-populates (*a priori*) flow table in switch
- Zero additional flow setup time
- Loss of control connection does not disrupt traffic
- Essentially requires aggregated (*wildcard*) rules

OpenFlow software switches

- Open vSwitch: Open Source and popular
- Of13softswitch: User-space software switch based on Ericsson TrafficLab 1.1
- Indigo: Open source implementation that runs on Mac OS X
- LINC: Open source implementation that runs on Linux, Solaris, Windows, MacOS, and FreeBSD
- Pantou: Turns a commercial wireless router/access point to an OpenFlow enabled switch. Supports generic Broadcom and some models of LinkSys and TP-Link access points with Broadcom and Atheros chipsets

OpenFlow Controllers: first wave

Name	Lang	Platform(s)	License	Original Author	Notes
OpenFlow Reference	C	Linux	OpenFlow License	Stanford/Nicira	not designed for extensibility
NOX	Python, C++	Linux	GPL	Nicira	
POX	Python	Any	Apache	Murphy McCauley (UC Berkeley)	
Ryu	Python	Linux	Apache	NSRC	Component based design Supports OpenStack integration
Trema	Ruby, C	Linux	GPL	NEC	includes emulator, regression test framework
Floodlight	Java	Any	Apache	BigSwitch Networks	
RouteFlow	?	Linux	Apache	CPqD (Brazil)	Special purpose controller to implement virtual IP routing as a service

OpenFlow controllers: new generation

- OpenDayLight
- ONOS

ONOS: Architecture Tiers

Northbound Abstraction:

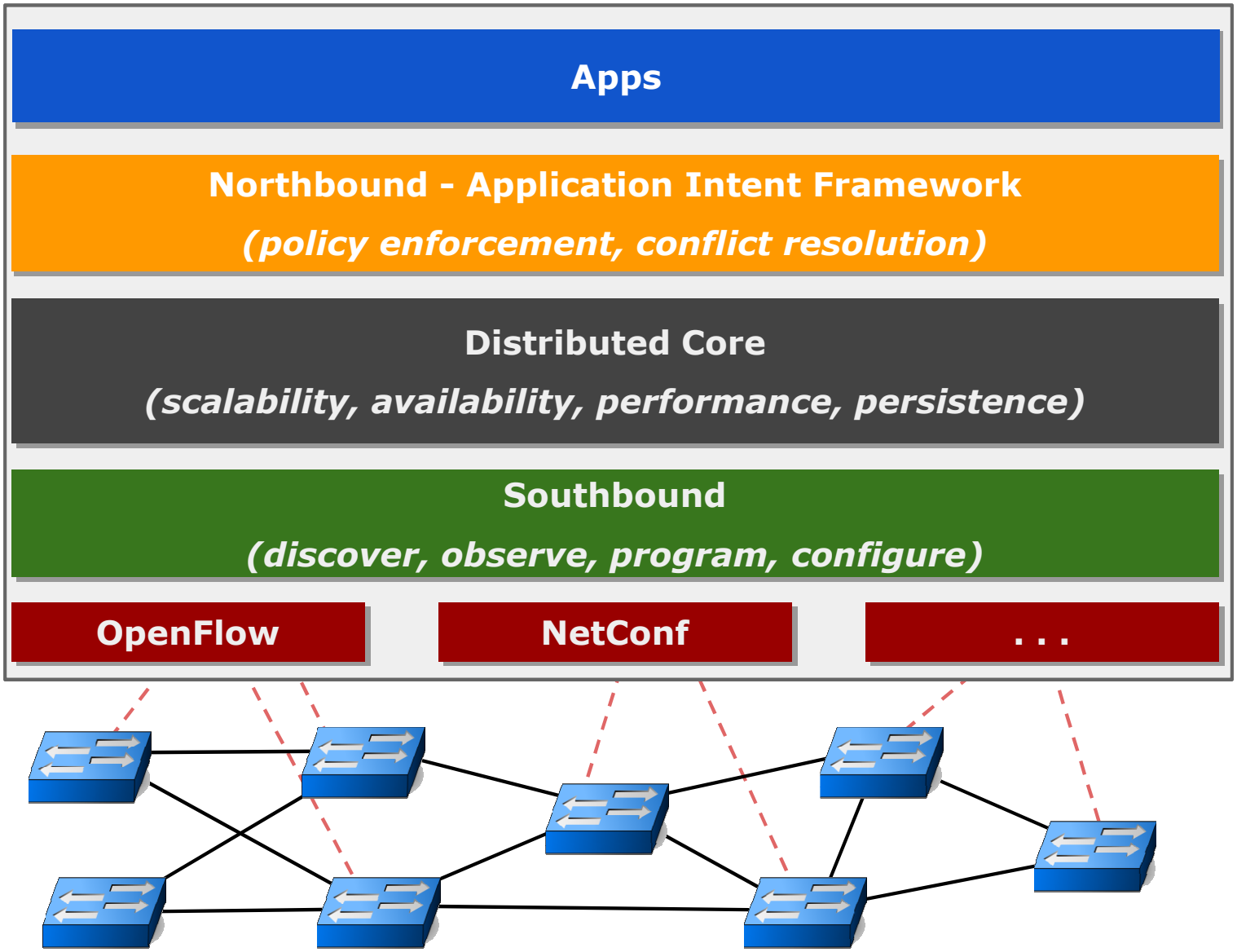
- network graph
- application intents

Core:

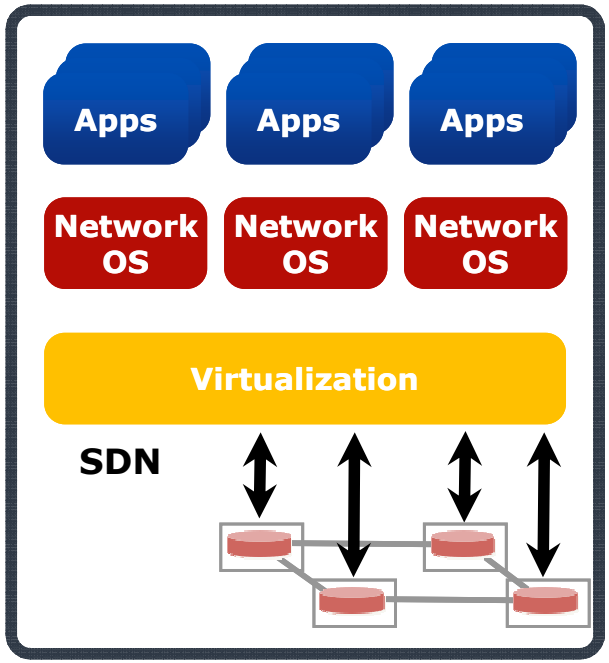
- distributed
- protocol independent

Southbound Abstraction:

- generalized OpenFlow
- pluggable & extensible



Virtualizing OpenFlow networks



- One of the goals of the SDN approach is to enable *Network Virtualization*, i.e. the possibility of creating and managing separately multiple logically-defined virtual infrastructures on top of a single shared substrate
- *FlowVisor* is a solution developed at Stanford University that allows network virtualization in the context of an OpenFlow network
- Network operators "delegate" control of subsets (*slices*) of network hardware and/or traffic to other network operators or users
- Multiple controllers can talk to the same set of switches
- FlowVisor is a software proxy between the forwarding and control planes of network devices
- FlowVisor intercepts OpenFlow messages from devices
  - FV only sends control plane messages to the Slice controller if the source device is in the Slice topology
  - Rewrites OF feature negotiation messages so the slice controller only sees the ports in its slice
  - Port up/down messages are pruned and only forwarded to affected slices
- Likewise, FlowVisor intercepts OpenFlow messages from controllers to preserve slice isolation

Network virtualization with OpenFlow and FlowVisor

- Slices are defined using a ***slice definition policy***
- The policy language specifies the slice's resource limits, flowspace, and controller's location in terms of IP and TCP port-pair
- FlowVisor enforces transparency and isolation between slices by inspecting, rewriting, and policing OpenFlow messages as they pass