Software Defined Networking (SDN)

NETWORKWORLD FROM IDG

INSIDER Sign In

NETWORK INTELLIGENCE

By Zeus Kerravala, Network World | MAY 25, 2017 10:58 AM PT

About

Zeus Kerravala is the founder and principal analyst with ZK Research, and provides a mix of tactical advice to help his clients in the current business climate.

OPINION

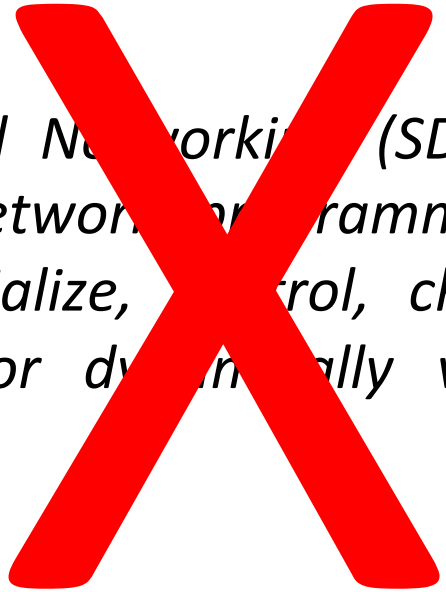
Cisco to network engineers: Get comfortable with software. It's here to stay

In this digital software-driven world, where companies must move with speed, software skills are now a must for network engineers

What is SDN?

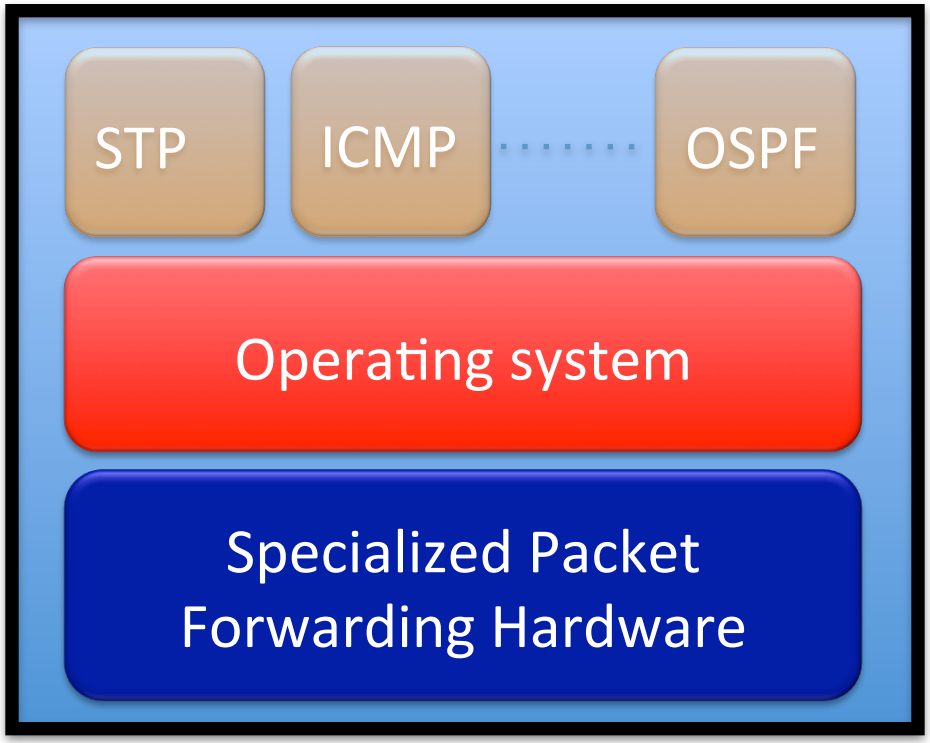
Software-Defined Networking (SDN) refers to a new approach for network programmability, that is, the capacity to initialize, control, change, and manage network behavior dynamically via open interfaces. [RFC7426]

Software-Defined Networking (SDN) refers to a new approach for network programmability, that is, the capacity to initialize, control, change, and manage network behavior dynamically via open interfaces. [RFC7426]



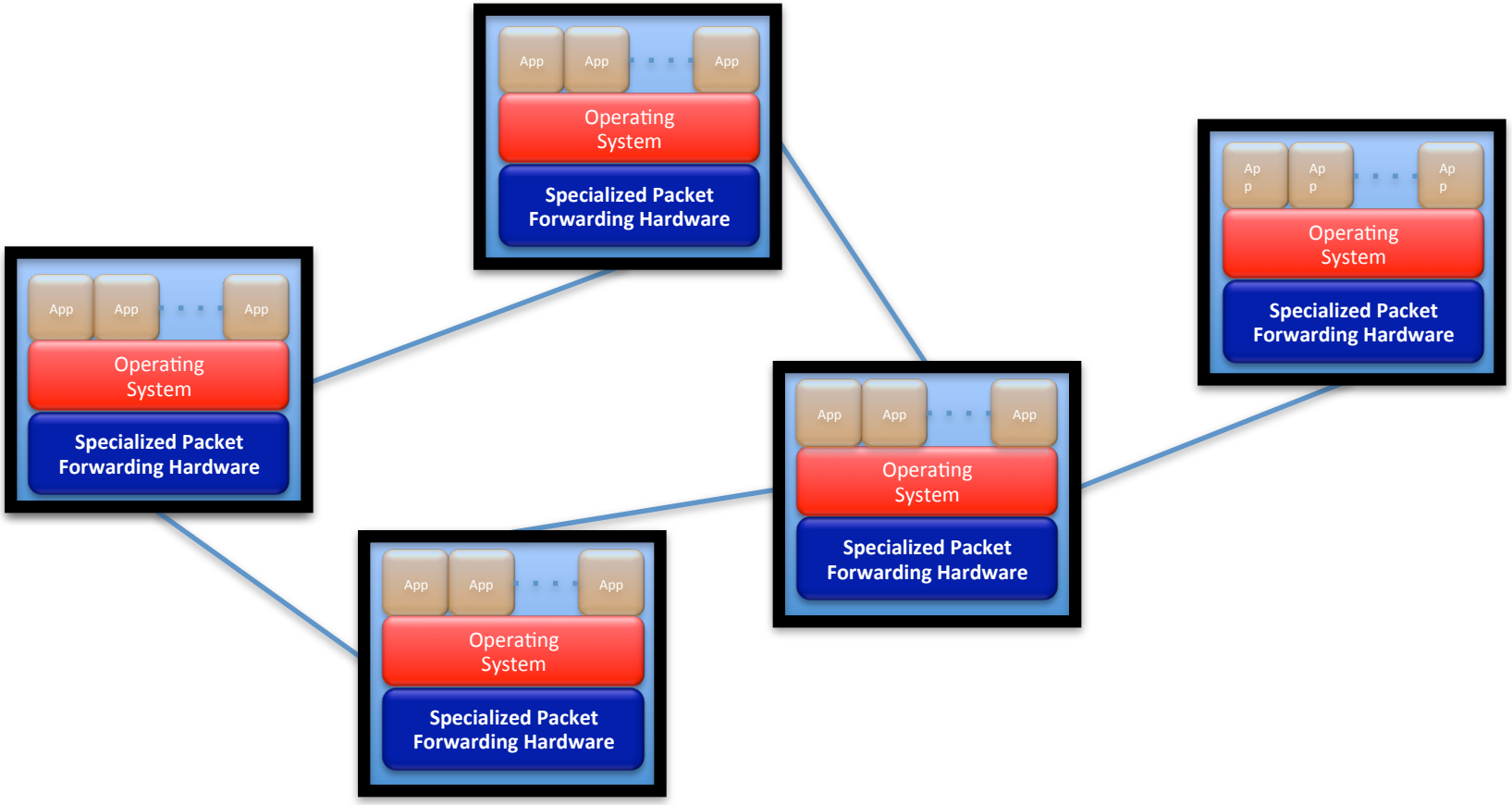
A different way of thinking about networks

Typical network node



Typical network

All nodes are equal
Peer-to-peer protocls



Protocolli peer-to-peer

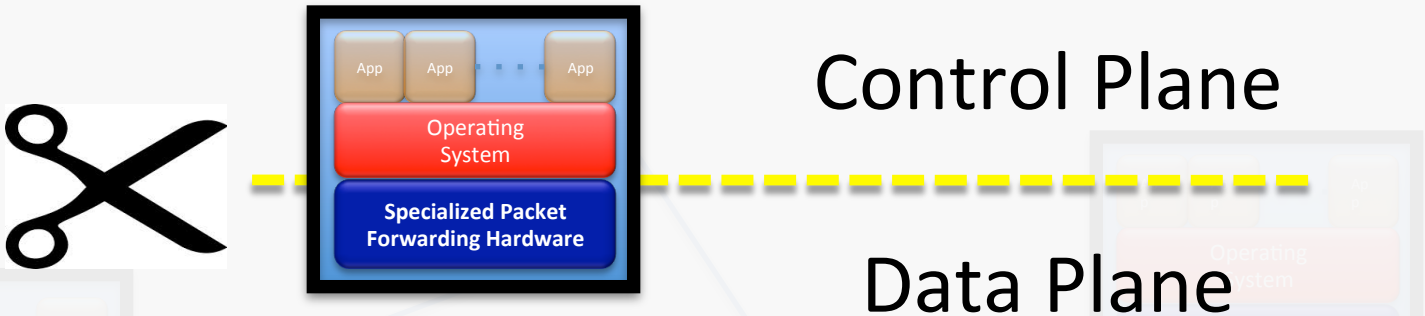
Peer-to-peer are excellent:

- They can be easily extended
- robust
- scalable (think about the Internet!)

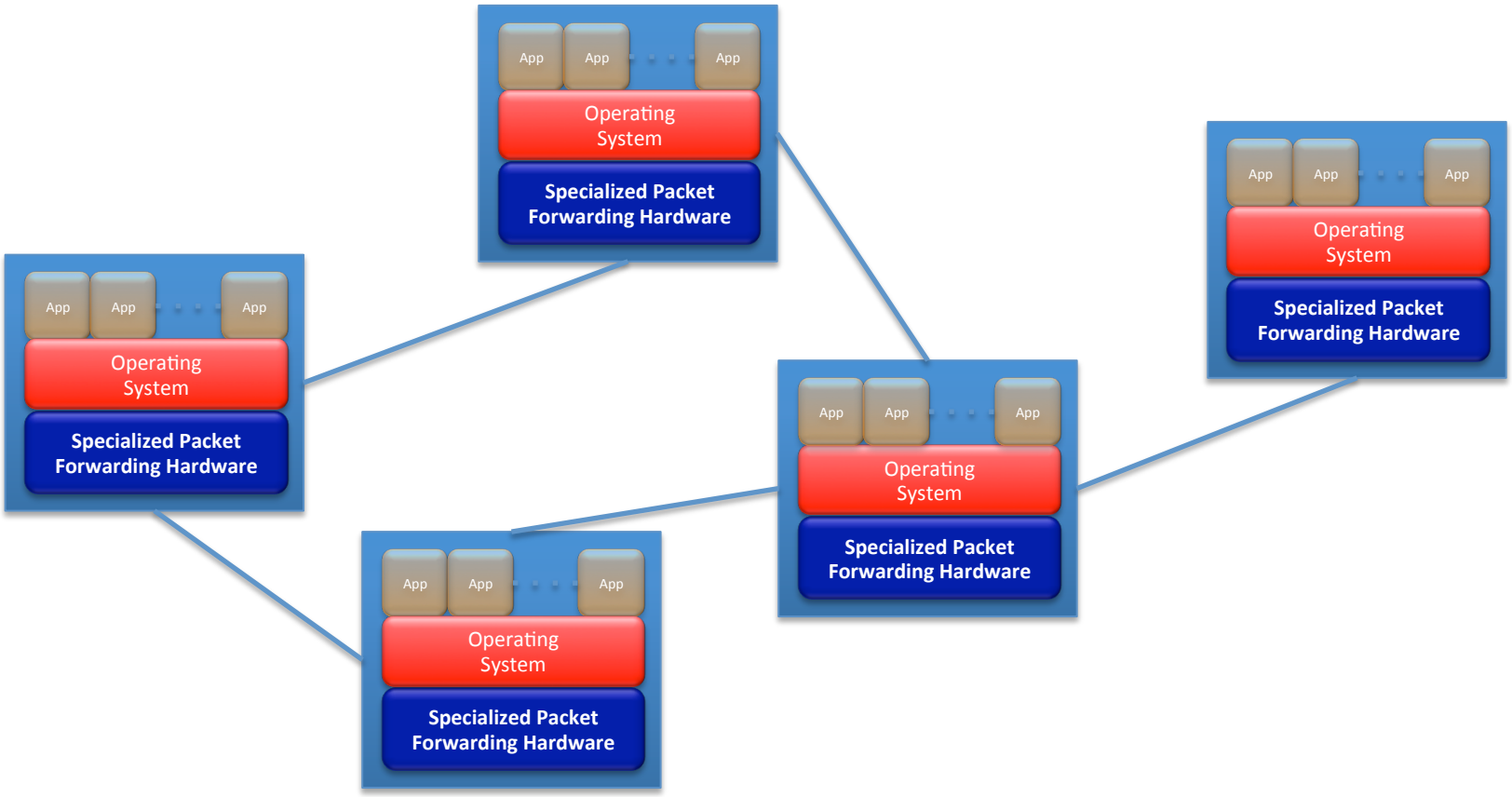
But…

- They are quite “expensive” to run
- problems are difficult to localize
- very very difficult to update and innovate

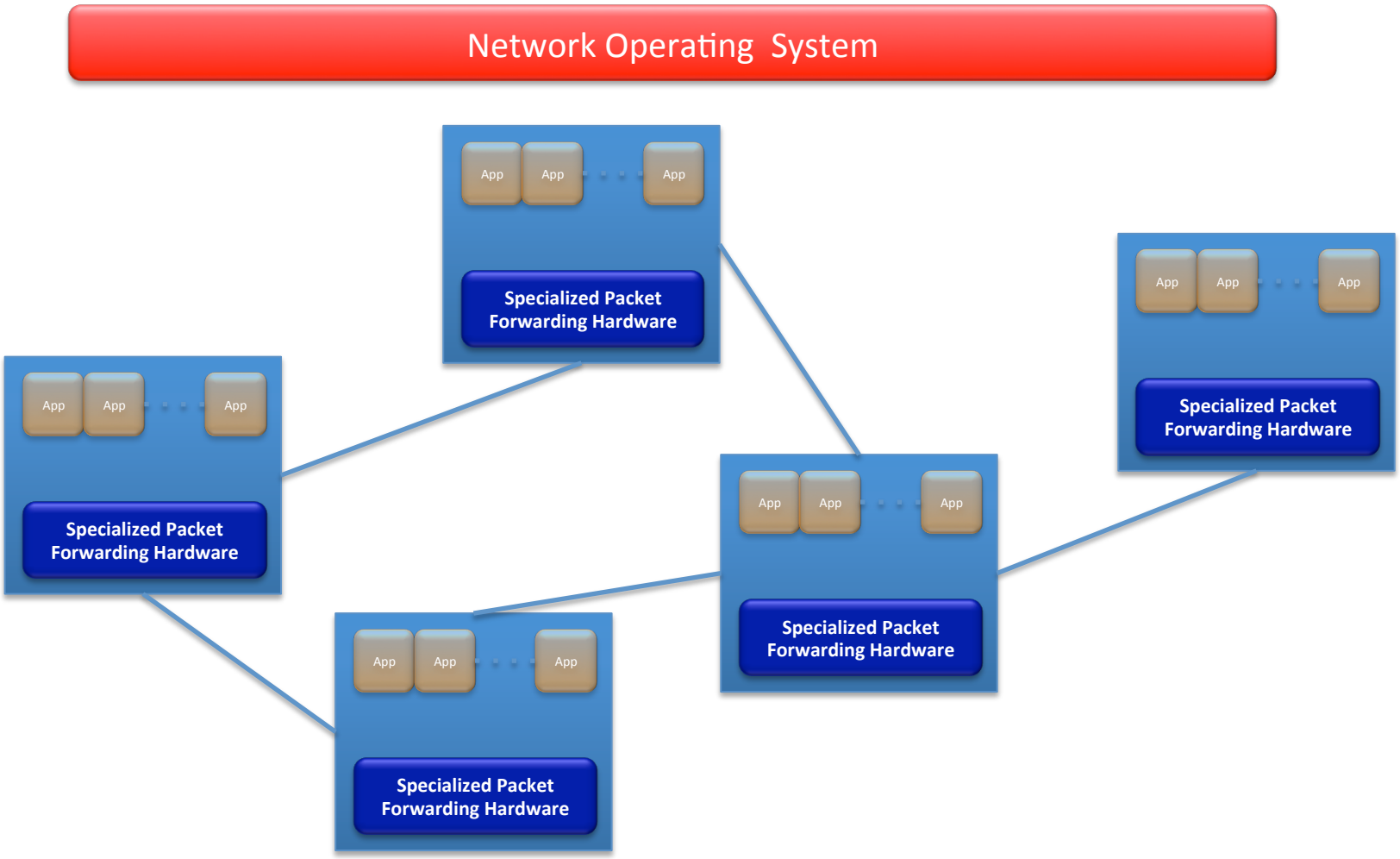
Separate Data Plane from Control Plane



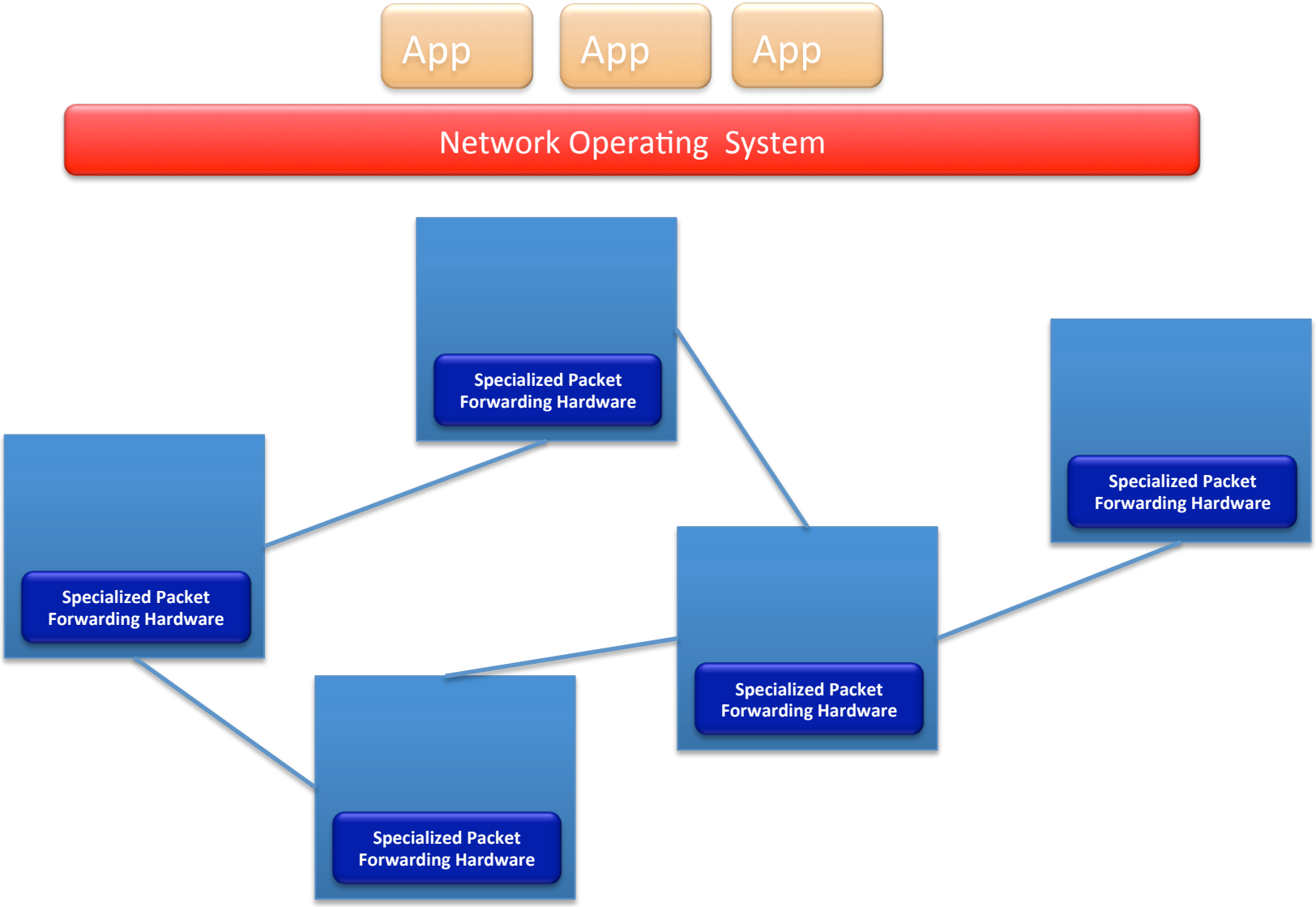
SDN moves network functionalities in a Network Operating System



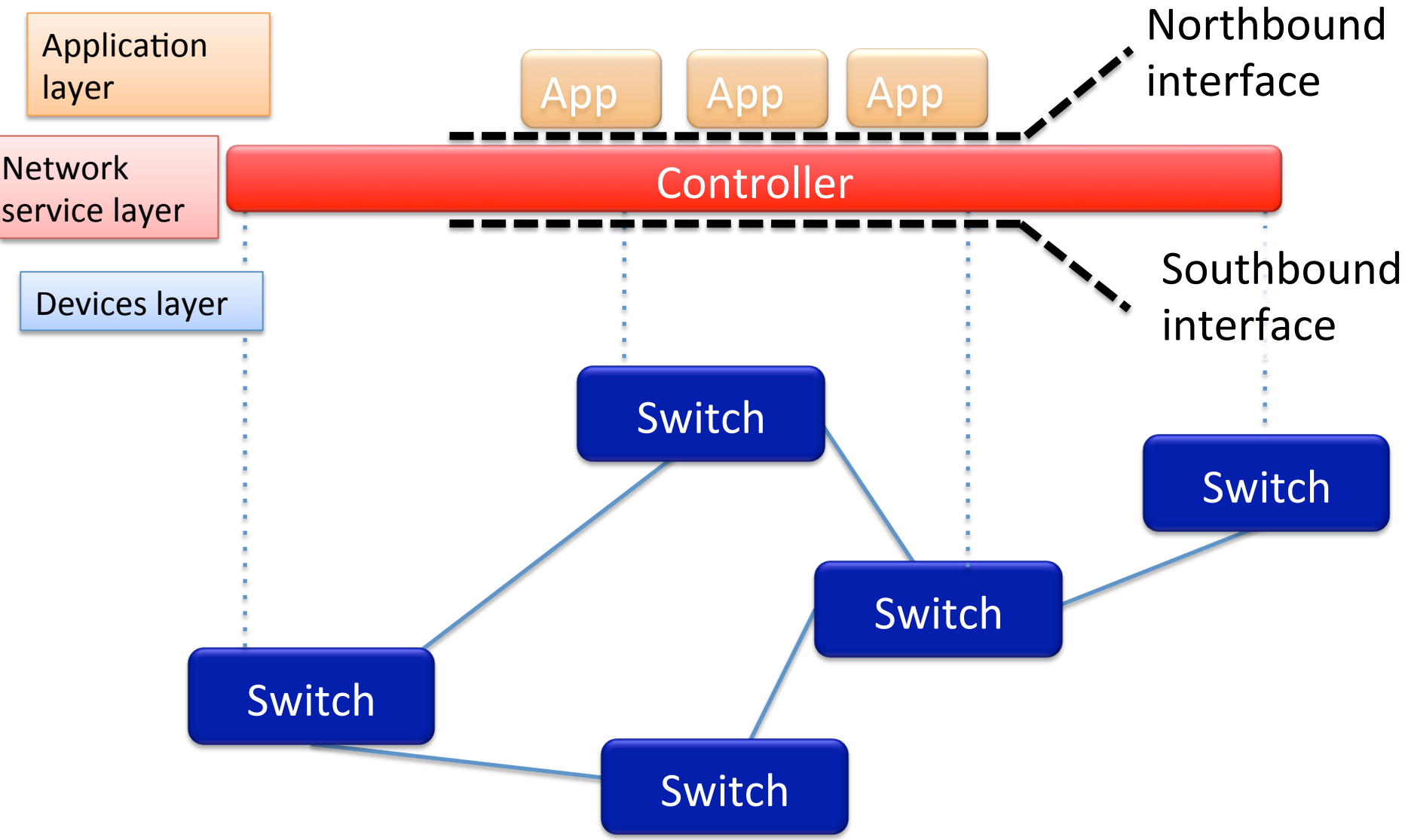
SDN moves network functionalities in a Network Operating System



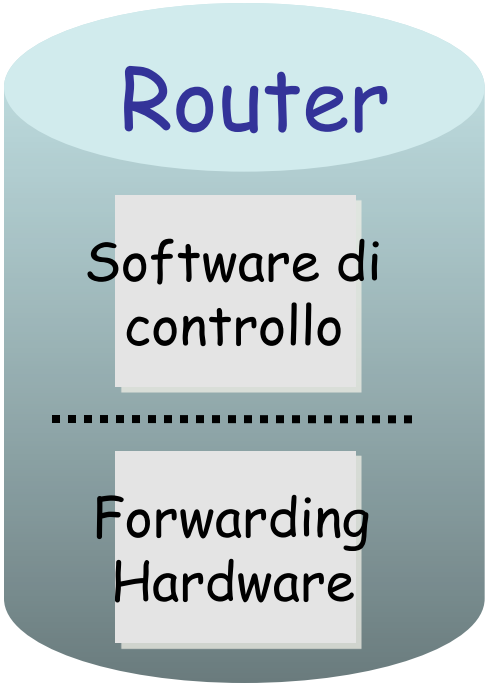
SDN moves network functionalities in a Network Operating System

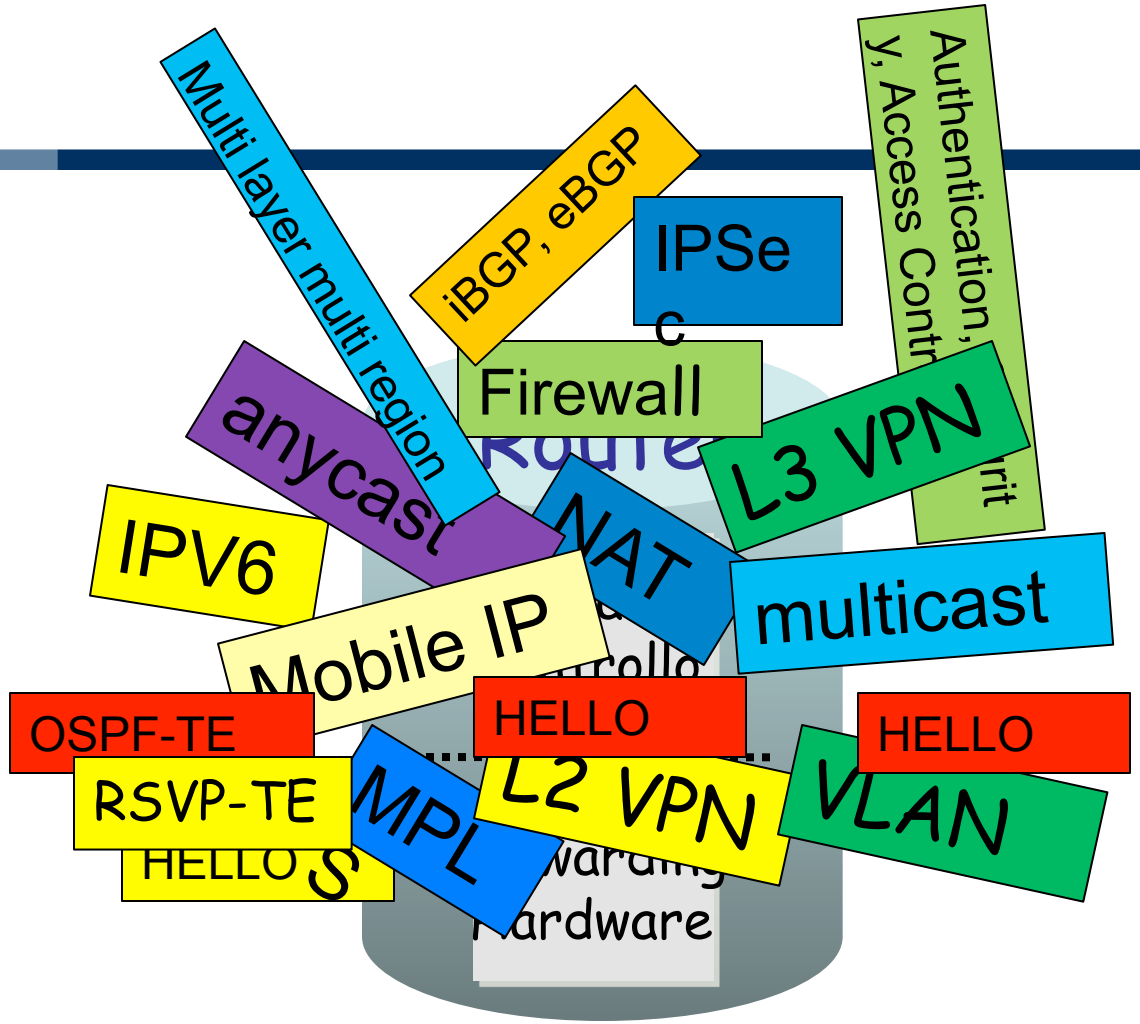


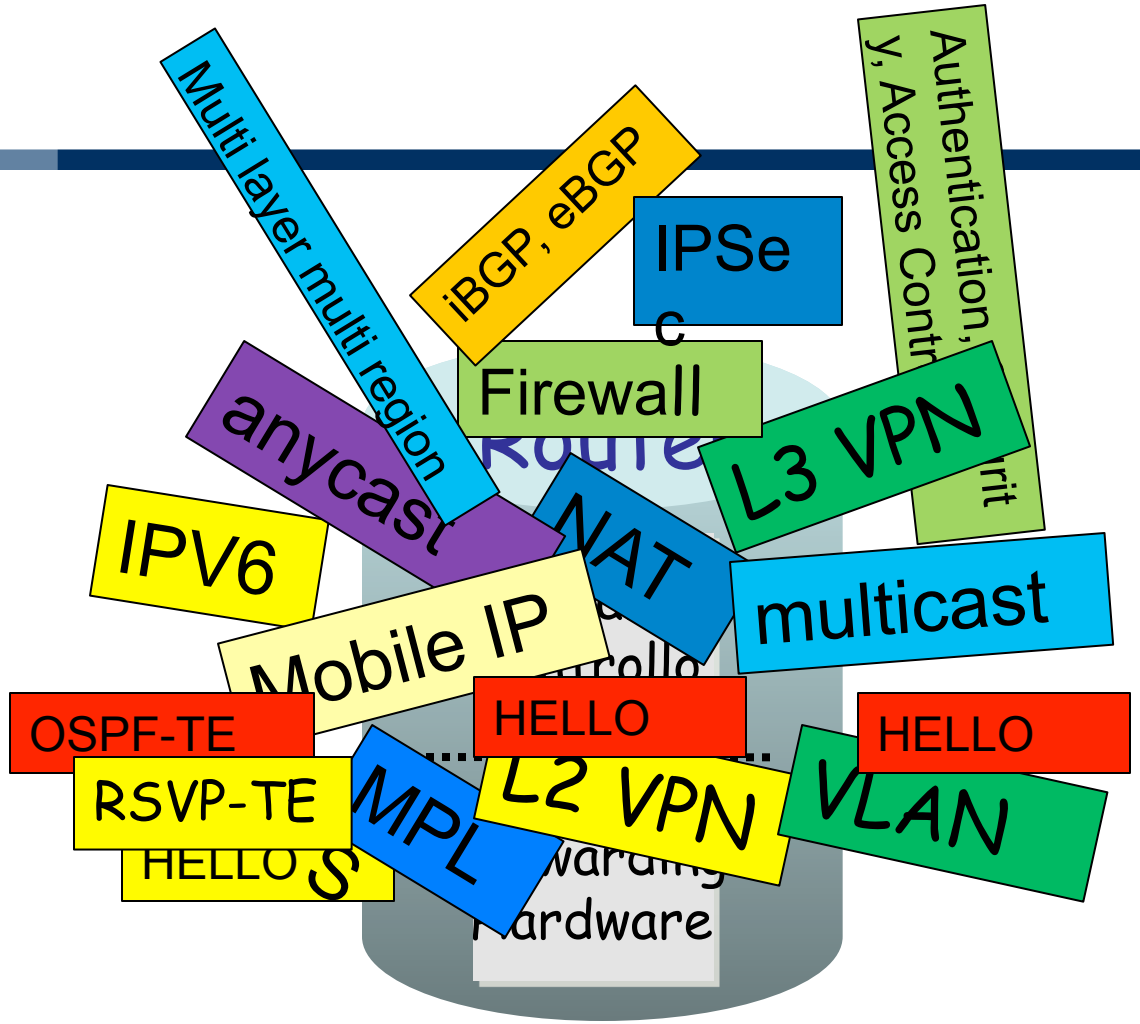
SDN stack: 3 layers, 2 interfaces



Why?


Router
Software di controllo
Forwarding Hardware


Multi layer multi region
iBGP, eBGP
IPSe
c
Authentication,
y, Access Contr
rit
Firewall
anycast
L3 VPN
IPV6
NAT
multicast
Mobile IP
OSPF-TE
HELLO
HELLO
RSVP-TE
MPL
L2 VPN
VLAN
HELLO
S
ardware

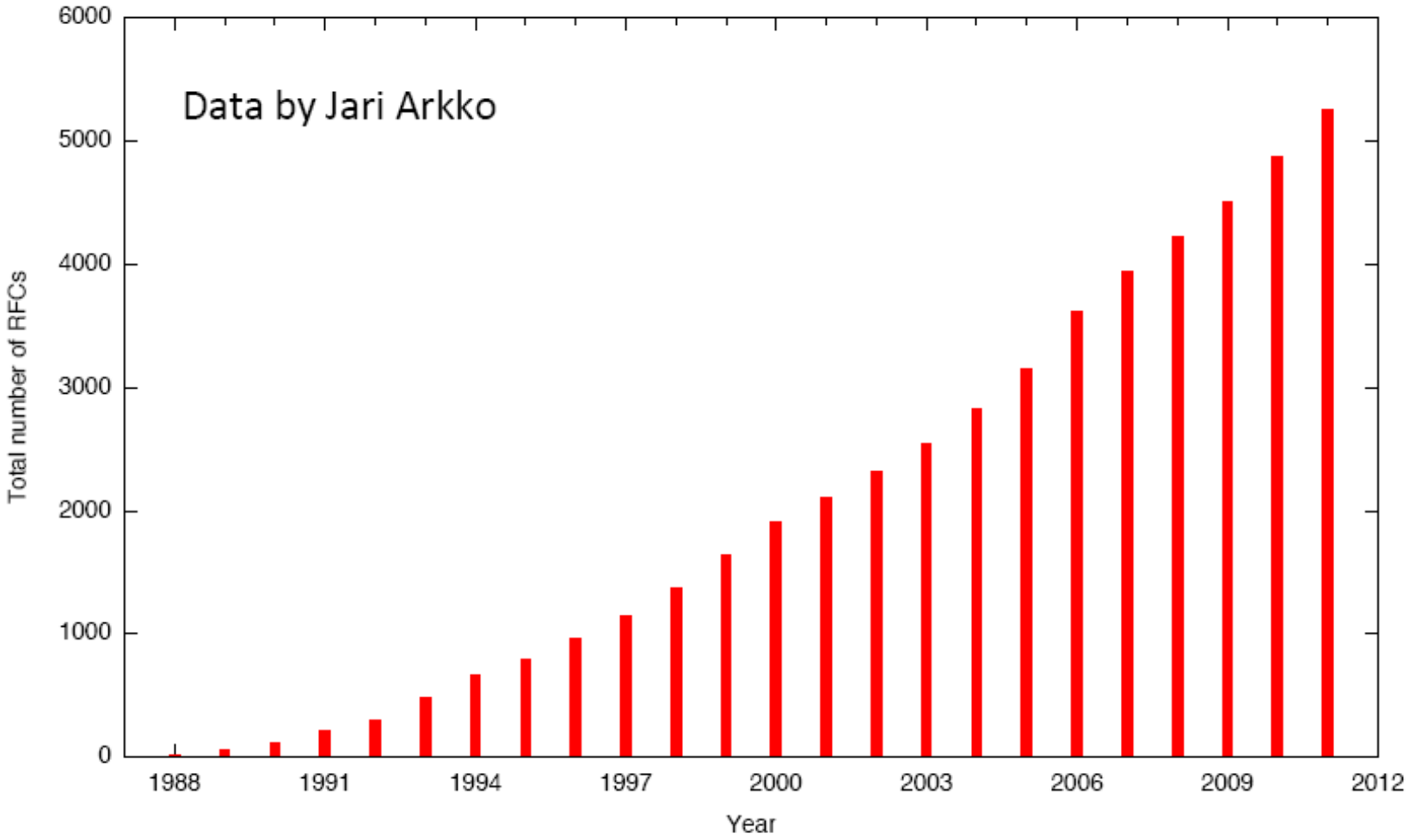


Infrastructure nodes do many things

OSPF, BGP, multicast, differentiated services,
Traffic Engineering, NAT, firewalls, MPLS, redundant layers, …

They should do less things WELL!

Number of published RFCs



Protocols implemented in a generic switch

(SLPP and MSTP)
- IEEE 802.1w – 2001 Rapid Reconfiguration for STP, RSTP
- IEEE 802.1Q – 2003 (formerly IEEE 802.1s) Multiple Instances of STP, MSTP
- EMISTP, Extreme Multiple Instances of Spanning Tree Protocol
- PVST+, Per VLAN STP (802.1Q Interoperable)
- Draft-ietf-bridge-rstpmib-03.txt – Definitions of Managed Objects for Bridges with Rapid Spanning Tree Protocol
- Extreme Standby Router Protocol™ (ESRP)
- IEEE 802.1Q – 1998 Virtual Bridged Local Area Networks
- IEEE 802.3ad Static load sharing configuration and LACP based dynamic configuration
- Software Redundant Ports
- IEEE 802.1AB – LLDP Link Layer Discovery Protocol
- LLDP Media Endpoint Discovery (LLDP-MED), ANSI/TIA-1057, draft 08
- Extreme Discovery Protocol (EDP)
- Extreme Loop Recovery Protocol (ELRP)
- Extreme Link State Monitoring (ELSM)
- IEEE 802.1ag L2 Ping and traceroute, Connectivity Fault Management
- ITU-T Y.1731 Frame delay measurements

Management and Traffic Analysis

- RFC 2030 SNTP, Simple Network Time Protocol v4
- RFC 854 Telnet client and server
- RFC 783 TFTP Protocol (revision 2)
- RFC 951, 1542 BootP
- RFC 2131 BOOTP/DHCP relay agent and DHCP server
- RFC 1591 DNS (client operation)
- RFC 1155 Structure of Management Information (SMIv1)
- RFC 1157 SNMPv1
- RFC 1212, RFC 1213, RFC 1215 MIB-II, Ethernet-Like MIB & TRAPs
- RFC 1573 Evolution of Interface
- RFC 1650 Ethernet-Like MIB (update of RFC 1213 for SNMPv2)
- RFC 1901, 1905 – 1908 SNMPv2c, SMIv2 and Revised MIB-II
- RFC 2576 Coexistence between SNMP Version 1, Version 2 and Version 3
- RFC 2578 – 2580 SMIv2 (update to RFC 1902 – 1903)
- RFC 3410 – 3415 SNMPv3, user based security, encryption and authentication
- RFC 3826 – The Advanced Encryption
- IEEE 802.1ag MIB
- Secure Shell (SSH-2) client and server
- Secure Copy (SCP-2) client and server
- Secure FTP (SFTP) server
- sFlow version 5
- Configuration logging
- Multiple Images, Multiple Configs
- RFC 3164 BSD Syslog Protocol with Multiple Syslog Servers
  - 999 Local Messages (critical's stored across reboots)
- Extreme Networks vendor MIBs (includes FDB, PoE, CPU, Memory MIBs)
- XML APIs over Telnet/SSH and HTTP/HTTPS
- Web-based device management interface – ExtremeXOS ScreenPlay™
- IP Route Compression

Security, Switch and Network Protection

- Secure Shell (SSH-2), Secure Copy (SCP-2) and SFTP client/server with encryption/authentication (requires export controlled encryption module)
- SNMPv3 user based security, with encryption/authentication (see above)
- RFC 1492 TACACS+
- RFC 2138 RADIUS Authentication
- RFC 2139 RADIUS Accounting
- RFC 3579 RADIUS EAP support for 802.1x
- RADIUS Per-command Authentication
- Access Profiles on All Routing Protocols
- Access Policies for Telnet/SSH-2/SCP-2
- Network Login – 802.1x, Web and MAC-based mechanisms
- IEEE 802.1x – 2001 Port-Based Network Access Control for Network Login
- Multiple supplicants with multiple VLANs for Network Login (all modes)
- Fallback to local authentication database (MAC and Web-based methods)
- Guest VLAN for 802.1x
- RFC 1866 HTML – Used for Web-based Network Login and ExtremeXOS ScreenPlay
- SSL/TLS transport – used for Web-based Network Login and ExtremeXOS ScreenPlay (requires export controlled encryption module)
- MAC Security – Lockdown and Limit
- IP Security – Trusted DHCP Server (RFC 3046 DHCP Option 82 with port and VLAN ID)
- Layer 2/3/4 Access Control Lists (ACLs)
- RFC 2267 Network Ingress Filtering
- RPF (Unicast Reverse Path Forwarding)
  - CA-97.28:Teardrop_Land-Teardrop and "LAND" attack
  - CA-96.26: ping
  - CA-96.21: tcp_syn_flooding
  - CA-95.01: UDP_service_denial
  - CA-95.01: IP_Spoofing_Attacks_and_Hijacked_Terminal_Connections
  - IP Options Attack
- Host Attack Protection
  - Teardrop, boink, opentear, jolt2, nextear, nestea, syndrop, smurf, fraggle, papasmurf, synk4, raped, winfreeze, ping –f, ping of death, pepsi5, Latierra, Winnuke, Simping, Sping, Ascend, Stream, Land, Octopus

Security, Router Protection

- IP Security – DHCP enforcement via Disable ARP Learning
- IP Security – Gratuitous ARP Protection
- IP Security – DHCP Secured ARP/ARP Validation
- Routing protocol MD5 authentication

Security Detection and Protection

- CLEAR-Flow, threshold based alerts and actions

IPv4 Host Services

- RFC 1122 Host Requirements
- RFC 768 UDP
- RFC 791 IP
- RFC 792 ICMP
- RFC 793 TCP
- RFC 826 ARP
- RFC 894 IP over Ethernet
- RFC 1027 Proxy ARP
- RFC 2068 HTTP server
- IGMP v1/v2/v3 Snooping with Configurable Router Registration Forwarding
- IGMP Filters
- PIM Snooping
- Static IGMP Membership
- Multicast VLAN Registration (MVR)

IPv4 Router Services

- RFC 1812 Requirements for IP Version 4 Routers
- RFC 1519 CIDR
- RFC 1256 IPv4 ICMP Router Discovery (IRDP)
- Static Unicast Routes
- Static Multicast Routes
- RFC 1058 RIP v1
- RFC 2453 RIP v2
- Static ECMP
- RFC 1112 IGMP v1
- RFC 1587 OSPF NSSA Option
- RFC 1765 OSPF Database Overflow
- RFC 2370 OSPF Opaque LSA Option
- RFC 3623 OSPF Graceful Restart
- RFC 1850 OSPFv2 MIB
- RFC 2362 PIM-SM (Edge-mode)
- RFC 2934 PIM MIB
- RFC 3569, draft-ietf-ssm-arch-06.txt PIM-SSM PIM Source Specific Multicast
- draft-ietf-pim-mld-v2-01.txt
- Mtrace, a "traceroute" facility for IP Multicast: draft-ietf-idmr-traceroute-ipm-07
- Mrinfo, the multicast router information tool based on Appendix-B of draft-ietf-idmr-dvmrp-v3-11

IPv6 Host Services

- RFC 3587, Global Unicast Address Format
- Ping over IPv6 transport
- Traceroute over IPv6 transport
- RFC 5095, Internet Protocol, Version 6 (IPv6) Specification
- RFC 4861, Neighbor Discovery for IP Version 6, (IPv6)
- RFC 2463, Internet Control Message Protocol (ICMPv6) for the IPv6 Specification
- RFC 2464, Transmission of IPv6 Packets over Ethernet Networks
- RFC 2465, IPv6 MIB, General Group and Textual Conventions
- RFC 2466, MIB for ICMPv6
- RFC 2462, IPv6 Stateless Address Auto Configuration – Host Requirements
- RFC 1981, Path MTU Discovery for IPv6, August 1996 – Host Requirements
- RFC 3513, Internet Protocol Version 6 (IPv6) Addressing Architecture
- Telnet server over IPv6 transport
- SSH-2 server over IPv6 transport

IPv6 Interworking and Migration

- RFC 2893, Configured Tunnels
- RFC 3056, 6to4

IPv6 Router Services

- RFC 2462, IPv6 Stateless Address Auto Configuration – Router Requirements
- RFC 1981, Path MTU Discovery for IPv6, August 1996 – Router Requirements
- RFC 2710, IPv6 Multicast Listener Discovery v1 (MLDv1) Protocol
- Static Unicast routes for IPv6
- RFC 2080, RIPng
- RFC 1771 Border Gateway Protocol 4
- RFC 1965 Autonomous System Confederations for BGP
- RFC 2796 BGP Route Reflection (supersedes RFC 1966)
- RFC 1997 BGP Communities Attribute
- RFC 1745 BGP4/IDRP for IP-OSPF Interaction
- RFC 2385 TCP MD5 Authentication for BGP-4
- RFC 2439 BGP Route Flap Damping
- RFC 2918 Route Refresh Capability for BGP-4
- RFC 3392 Capabilities Advertisement with BGP-4
- RFC 4360 BGP Extended Communities Attribute
- RFC 4486 Subcodes for BGP Cease Notification message
- draft-ietf-idr-restart-10.txt Graceful Restart Mechanism for BGP
- RFC 4760 Multiprotocol extensions for BGP-4
- RFC 1657 BGP-4 MIB
- RFC 4893 BGP Support for Four-Octet AS Number Space
- Draft-ietf-idr-bgp4-mibv2-02.txt – Enhanced BGP-4 MIB
- RFC 1195 Use of OSI IS-IS for Routing in TCP/IP and Dual Environments (TCP/IP transport only)
- RFC 2763 Dynamic Hostname Exchange Mechanism for IS-IS
- RFC 2966 Domain-wide Prefix Distribution with Two-Level IS-IS
- RFC 2973 IS-IS Mesh Groups
- RFC 3373 Three-way Handshake for IS-IS Point-to-Point Adjacencies
- Draft-ietf-isis-restart-02 Restart Signaling for IS-IS
- Draft-ietf-isis-ipv6-06 Routing IPv6 with IS-IS
- Draft-ietf-isis-wg-multi-topology-11 Multi Topology (MT) Routing in IS-IS

QoS and VLAN Services

Quality of Service and Policies
- IEEE 802.1D – 1998 (802.1p) Packet Priority
- RFC 2474 DiffServ Precedence, including 8 queues/port
- RFC 2598 DiffServ Expedited Forwarding (EF)
- RFC 2597 DiffServ Assured Forwarding (AF)
- RFC 2475 DiffServ Core and Edge Router Functions

Traffic Engineering
- RFC 3784 IS-IS Extens for Traffic Engineering (wide metrics only)

VLAN Services: VLANs, vMANs
- IEEE 802.1Q VLAN Tagging
- IEEE 802.1v: VLAN classification by Protocol and Port

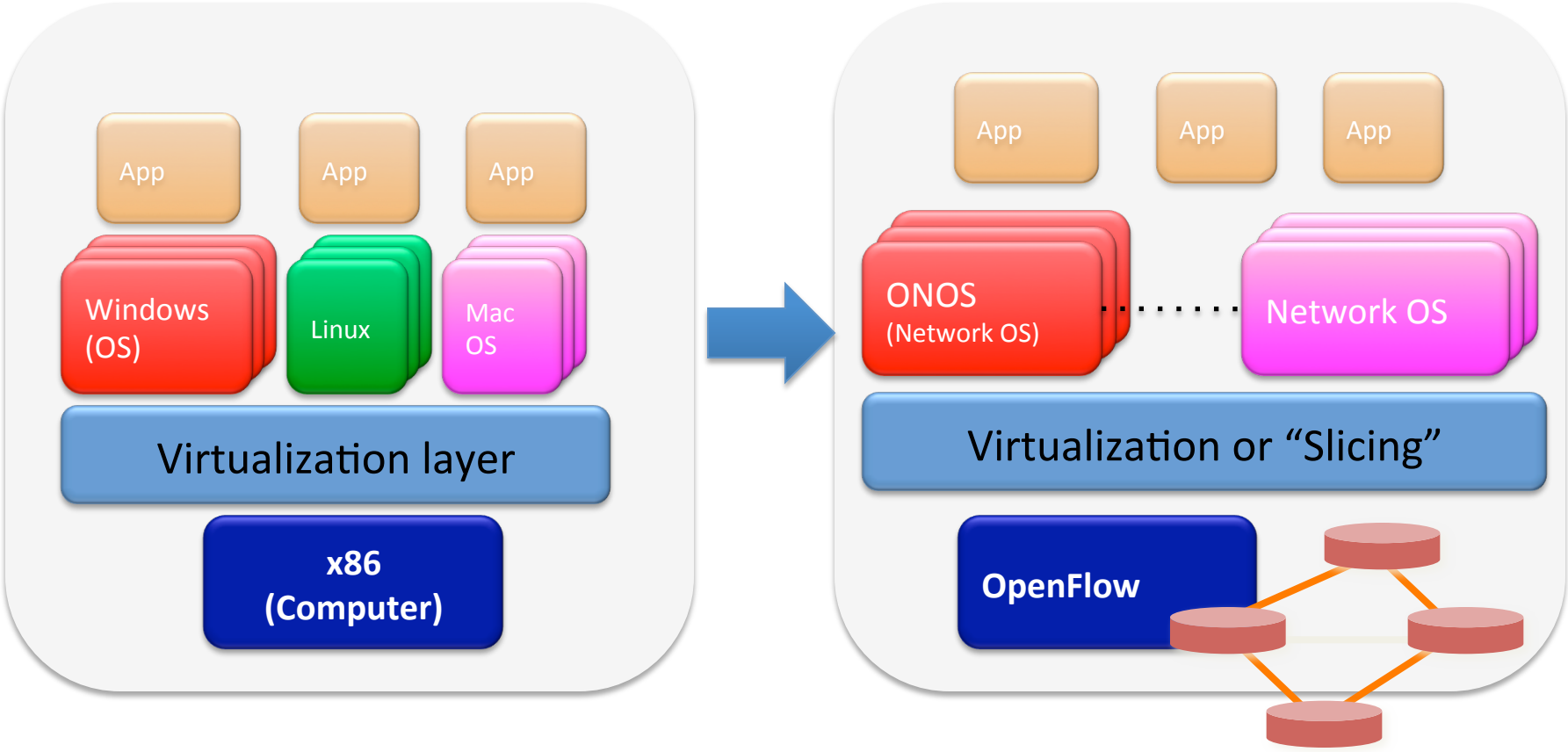
Advanced VLAN Services, MAC-in-MAC
- VLAN Translation in vMAN environments
- vMAN Translation
- IEEE 802.1ah/D1.2 Provider Backbone Bridges (PBB)/MAC-in-MAC

MPLS and VPN Services

Multi-Protocol Label Switching (MPLS)
Requires MPLS Layer 2 Feature Pack License
- RFC 2961 RSVP Refresh Overhead Reduction Extensions
- RFC 3031 Multiprotocol Label Switching Architecture
- RFC 3032 MPLS Label Stack Encoding
- RFC 3036 Label Distribution Protocol (LDP)
- RFC 3209 RSVP-TE: Extensions to RSVP for LSP Tunnels
- RFC 3630 Traffic Engineering Extensions to OSPFv2
- RFC 3784 IS-IS extensions for traffic engineering only (wide metrics only)
- RFC 3811 Definitions of Textual Conventions (TCs) for Multiprotocol Label Switching (MPLS) Management
- RFC 3812 Multiprotocol Label Switching (MPLS) Traffic Engineering (TE) Management Information Base (MIB)
- RFC 3813 Multiprotocol Label Switching (MPLS) Label Switching Router (LSR) Management Information Base (MIB)
- RFC 3815 Definitions of Managed Objects for the Multiprotocol Label Switching (MPLS), Label Distribution Protocol (LDP)
- RFC 4090 Fast Re-route Extensions to RSVP-TE for LSP (Detour Paths)
- RFC 4379 Detecting Multi-Protocol Label Switched (MPLS) Data Plane Failures (LSP Ping)
- draft-ietf-bfd-base-09.txt Bidirectional Forwarding Detection

Layer 2 VPNs
Requires MPLS Layer 2 Feature Pack License
- RFC 4447 Pseudowire Setup and Maintenance using the Label Distribution Protocol (LDP)
- RFC 4448 Encapsulation Methods for Transport of Ethernet over MPLS Networks
- RFC 4762 Virtual Private LAN Services (VPLS) using Label Distribution Protocol (LDP) Signaling
- RFC 5085 Pseudowire Virtual Circuit Connectivity Verification (VCCV)
- RFC 5542 Definitions of Textual Conventions for Pseudowire (PW) Management
- RFC 5601 Pseudowire (PW) Management

Technological Trend



Computer Industry

Network Industry

Hardware sublayer simple and stable, programmability, isolation and competition in upper layers

Abstractions

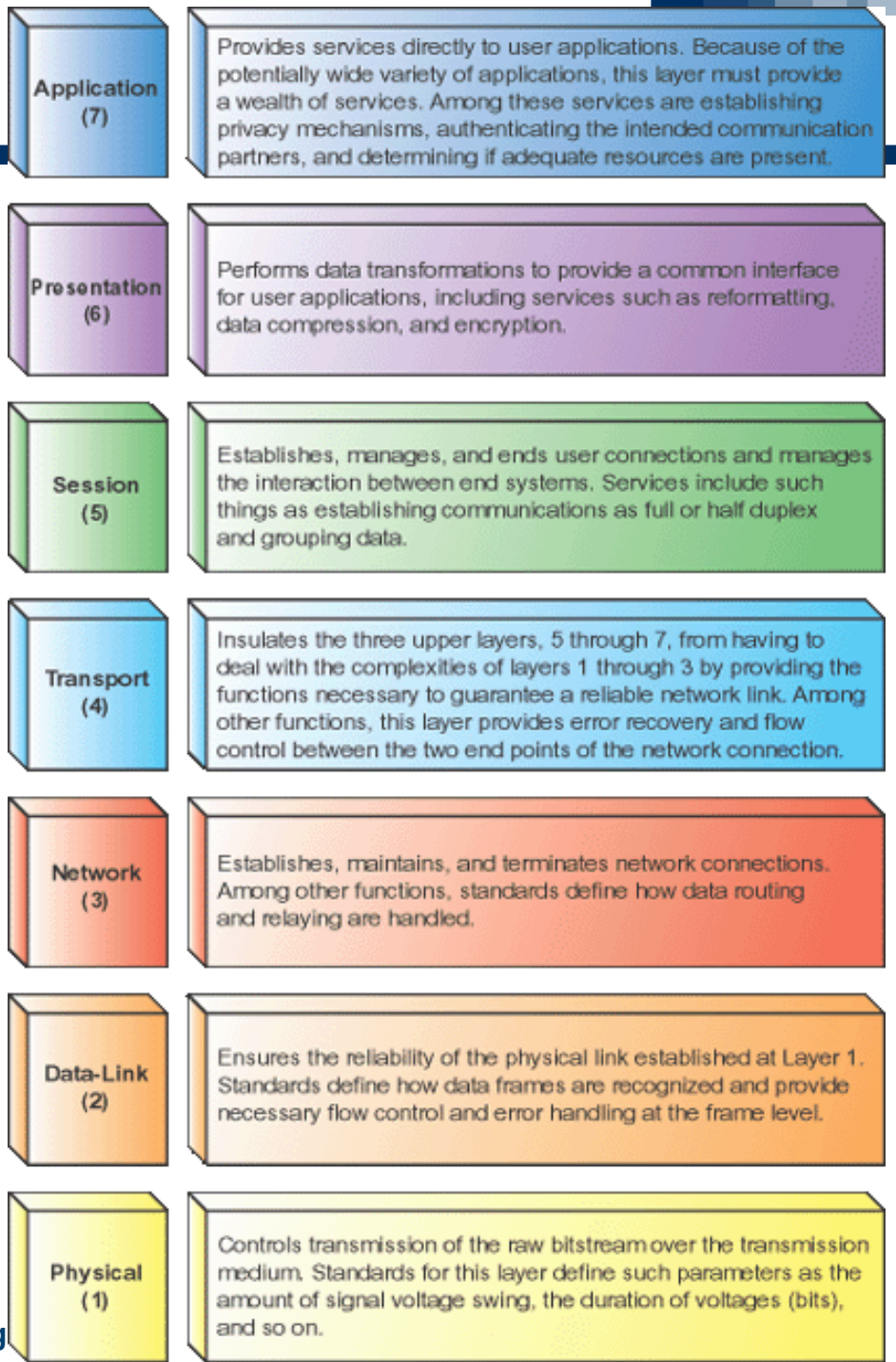
Abstractions allow programs easier to write and maintain

Data plane abstractions:

- the OSI stack

Control plane abstractions? They must be developed:

- Devices layer
- Network services layer



Devices Abstraction

Devices Abstraction

Current devices

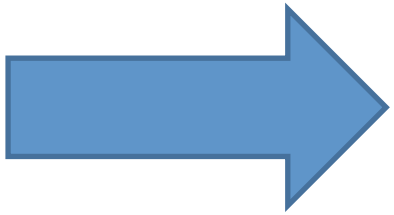
Router IP

Switch Ethernet

Firewall

NAT box

L4 switch



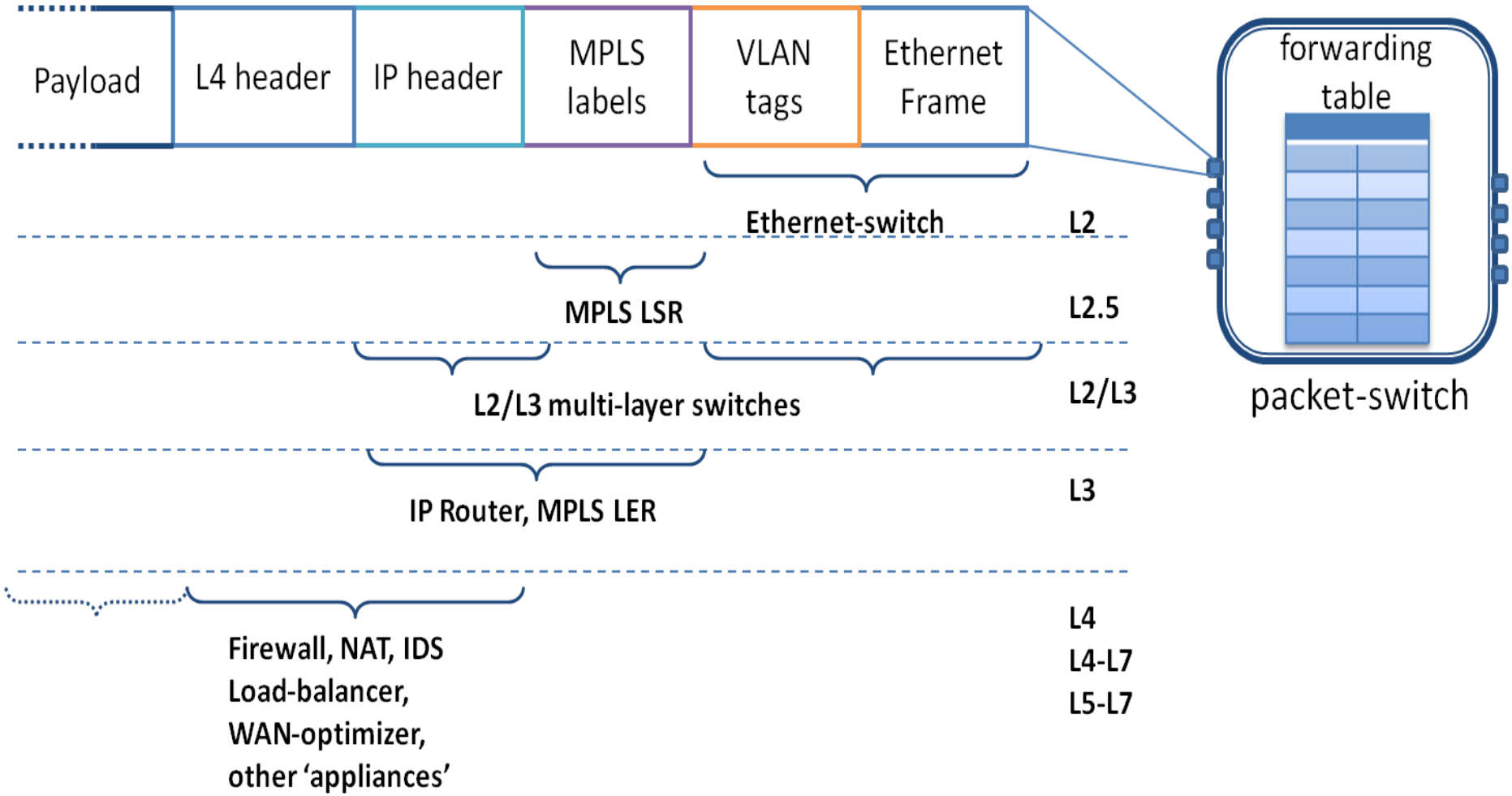
Abstract device

Flow table

The Flow Table abstraction is independent of the layer in which the device will operate

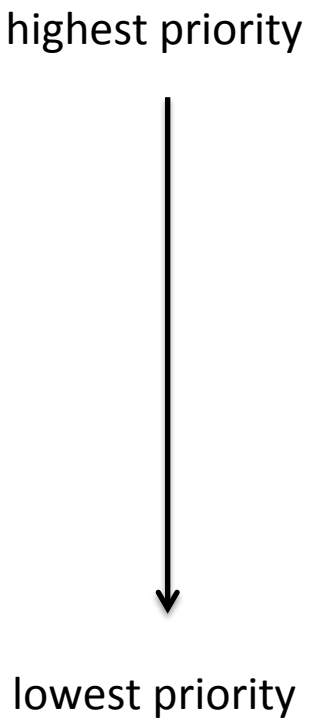
A flow is defined by a packet classification rule, based on the header values

Flow table abstraction



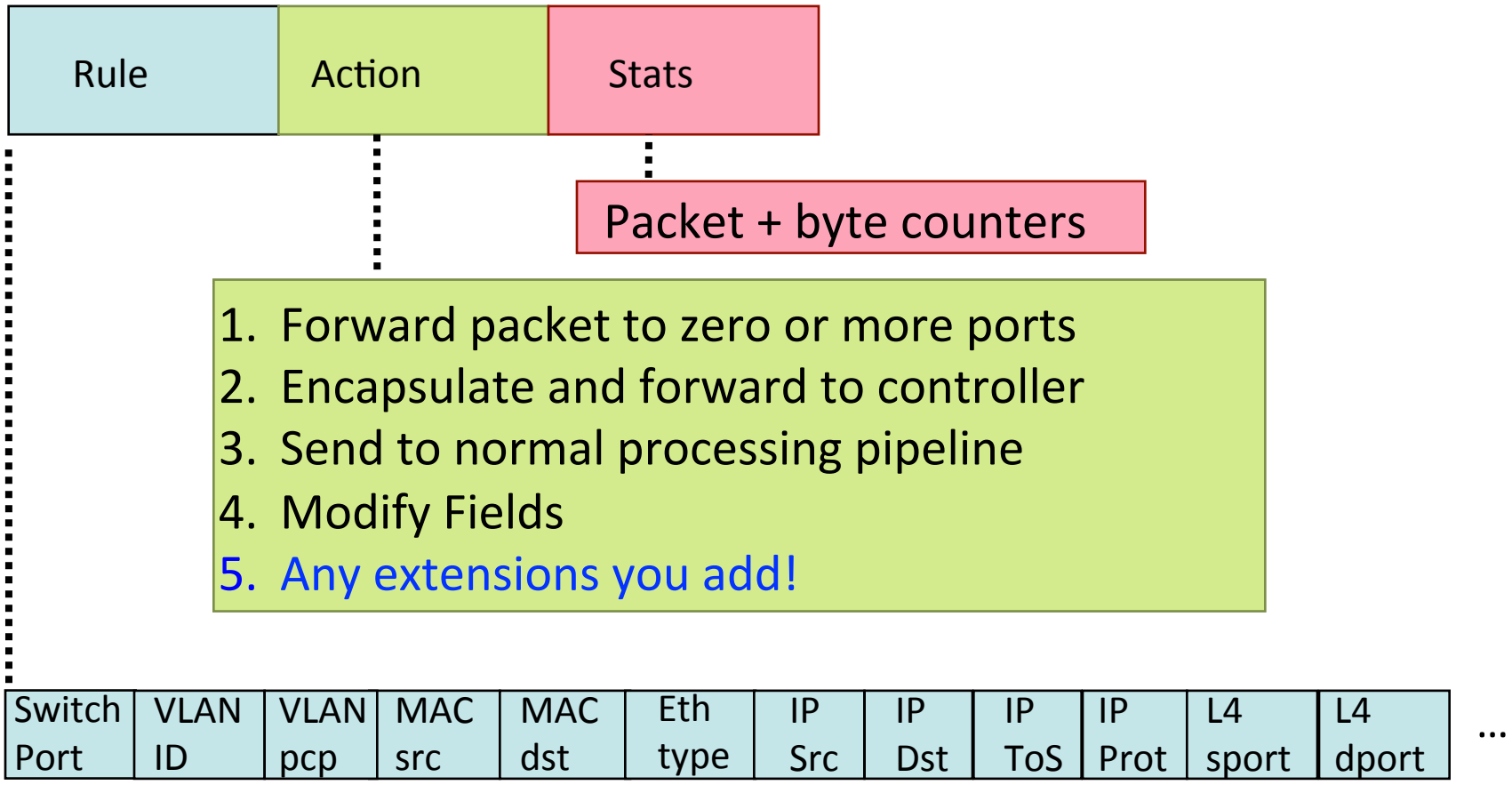
Astrazione a tabella di flussi

Rule	Action	Stats
Rule	Action	Stats
Rule	Action	Stats
Rule	Action	Stats
Default Rule	Action: Send to Controller	Stats



Ogni pacchetto viene classificato in base a una regola

Flow table abstraction

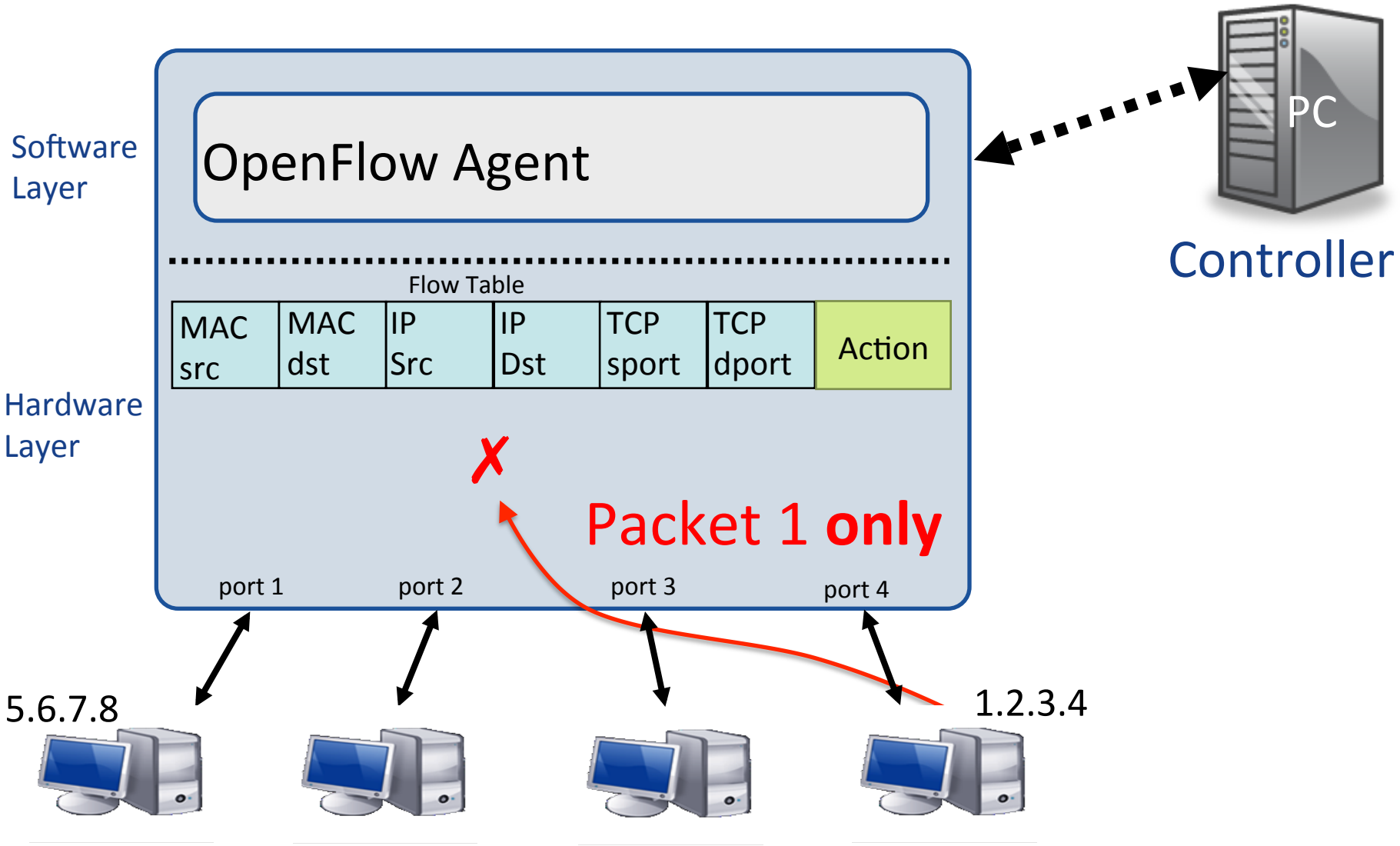


Matches can be exact or with a wildcard

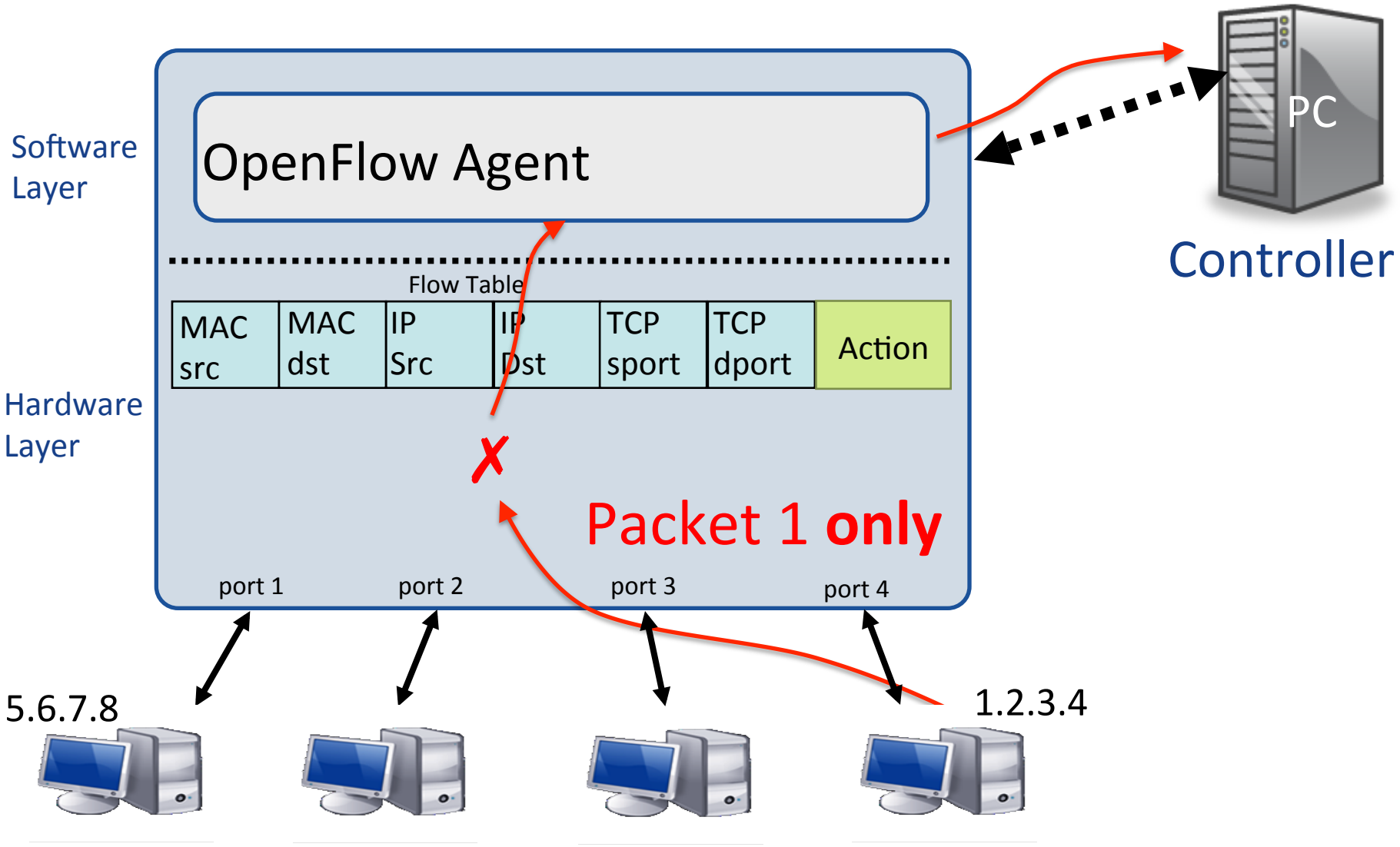
One table, many possible behaviors

Name	Port	MAC Src	MAC Dst	Eth Type	VLAN ID	IP Src	IP Dst	IP Prot	UDP/ TCP Sport	UDP/ TCP Dport	Action
Switchboard	p1	*	*	*	*	*	*	*	*	*	port2
Port Mirroring	p1	*	*	*	*	*	*	*	*	*	port2, port3
L2 Switching	*	*	00:1f...	*	*	*	*	*	*	*	port2
VLAN Switching	*	*	00:1f...	*	vlan3	*	*	*	*	*	port2
IP Routing	*	*		*	*	*	5.6.7.8/ 16	*	*	*	writeMAC, port2
Firewall	*	*	*	*	*	*	*	*	*	22	drop
Flow Switching	p3	00:20...	00:1f...	0800	vlan3	1.2.3.4	5.6.7.8	4	17264	80	port2

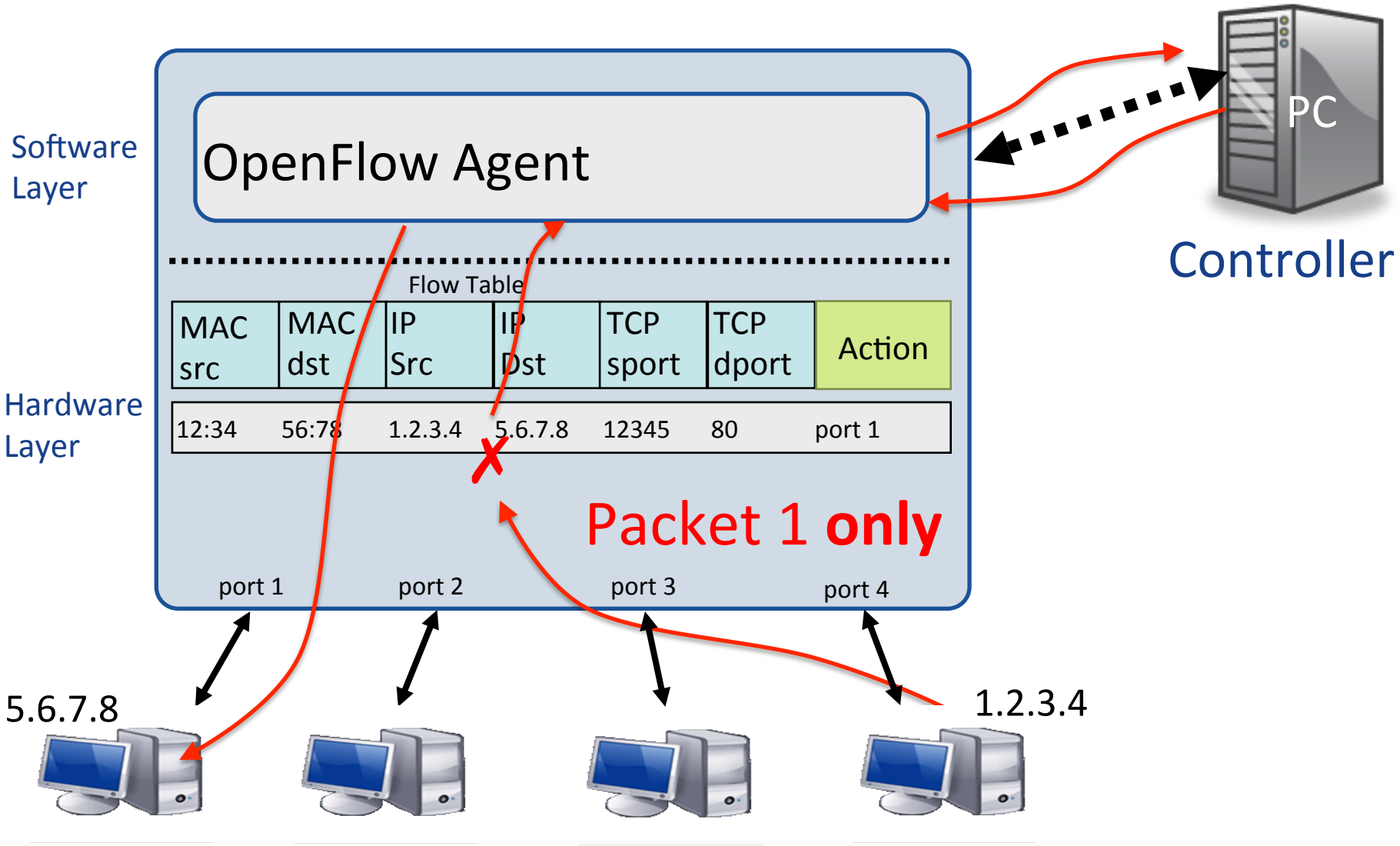
Reactive installation of Rules



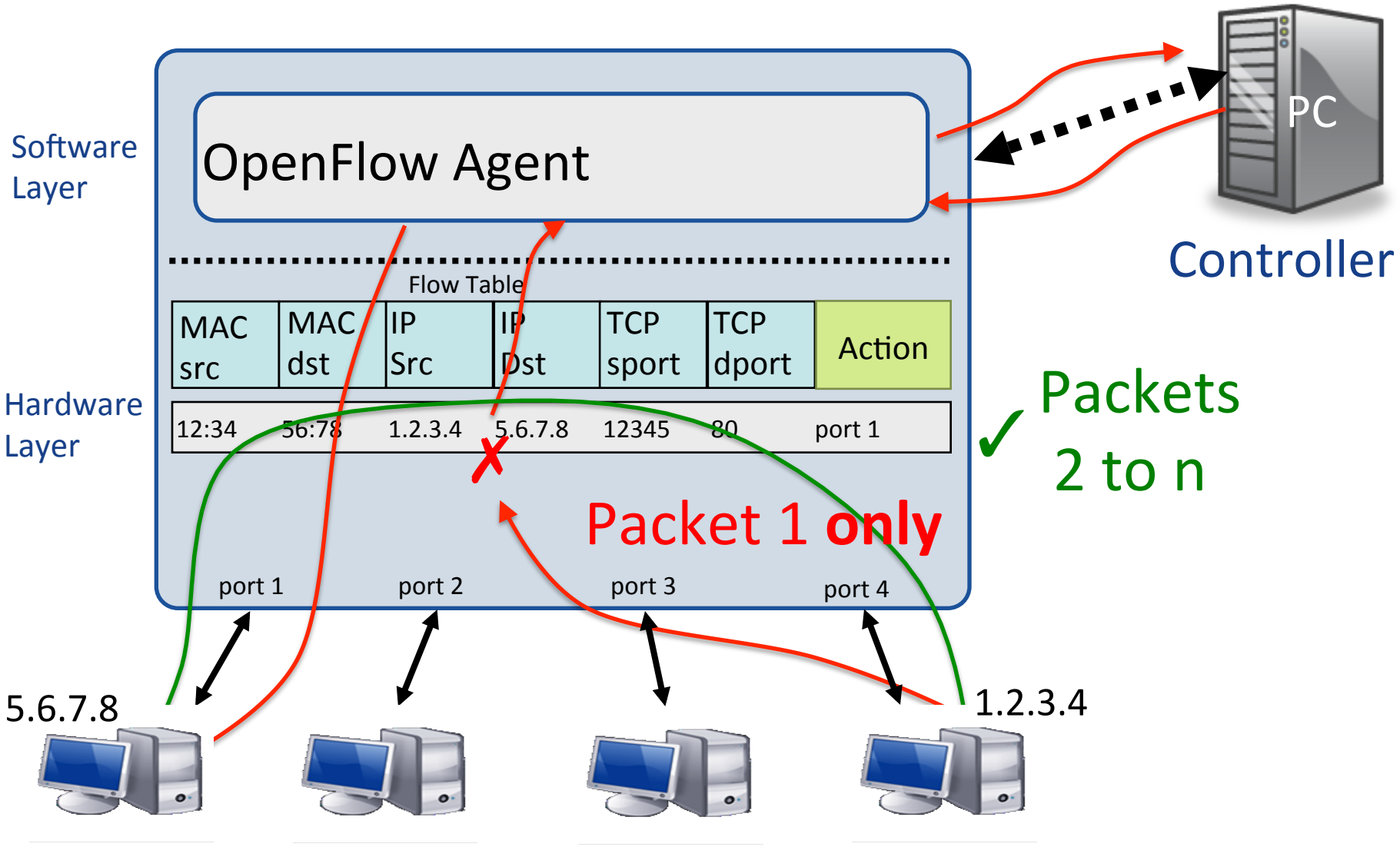
Reactive installation of Rules



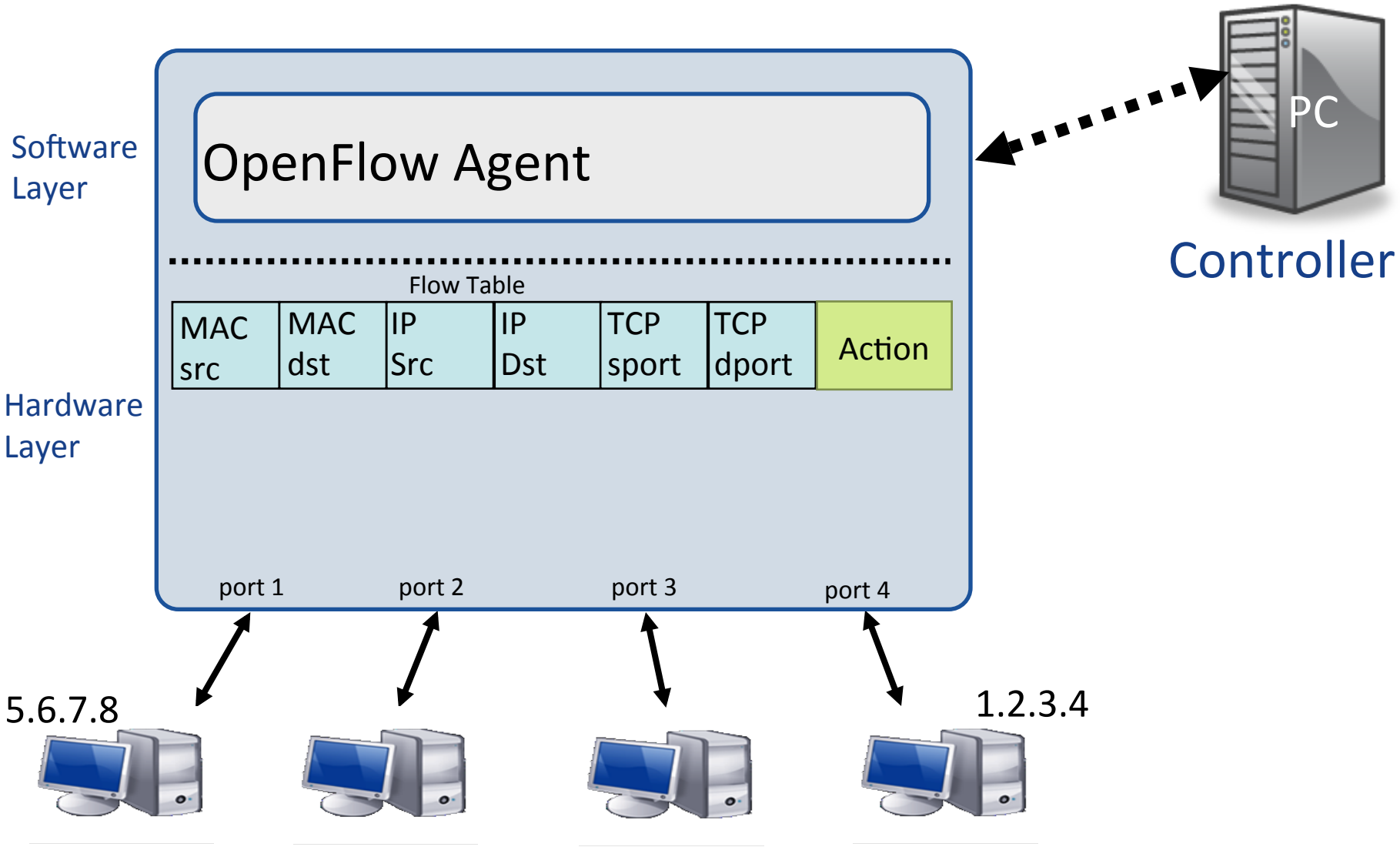
Reactive installation of Rules



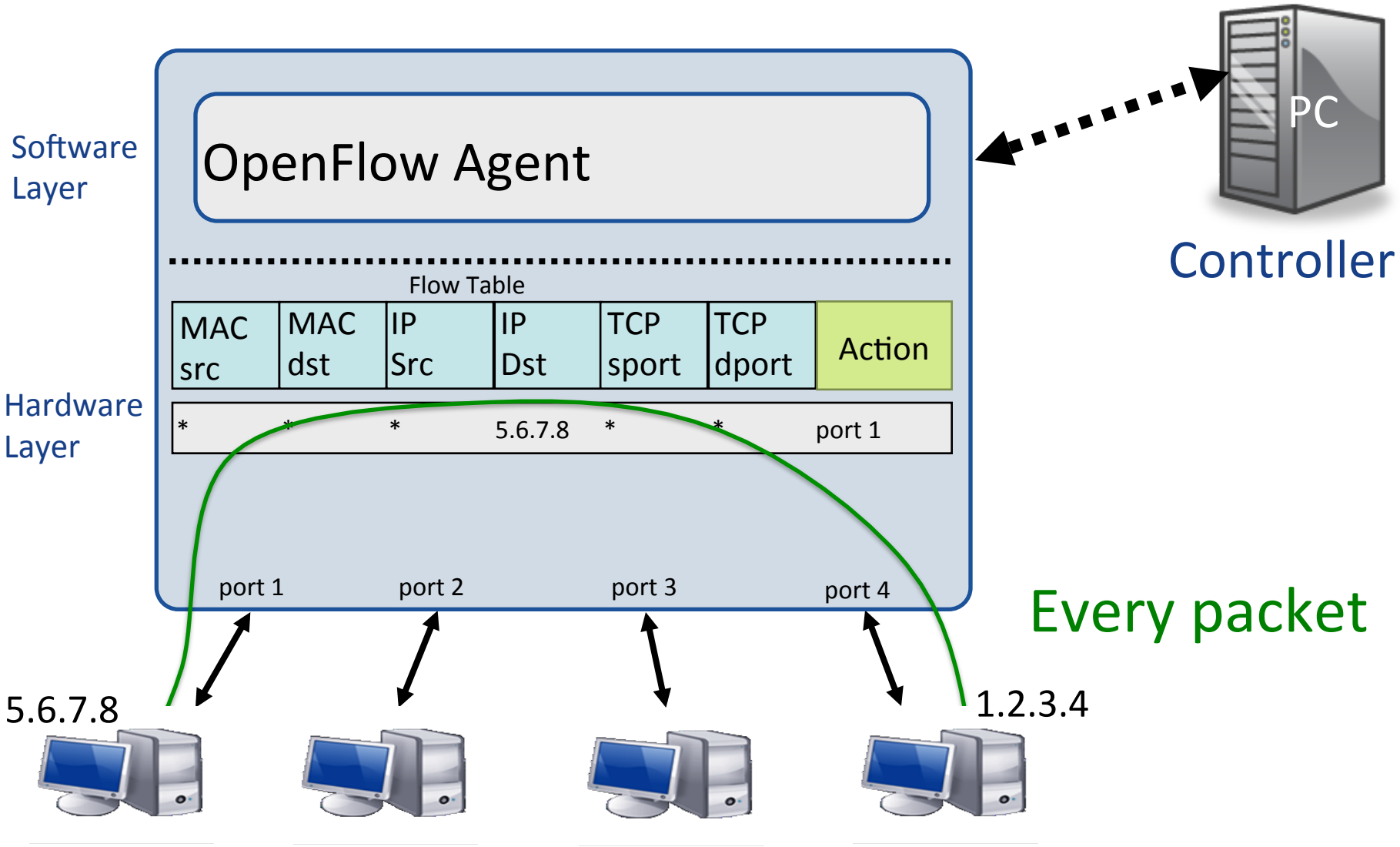
Reactive installation of Rules



Proactive installation of Rules



Proactive installation of Rules



Network Services Abstraction

Abstraction of Network Services

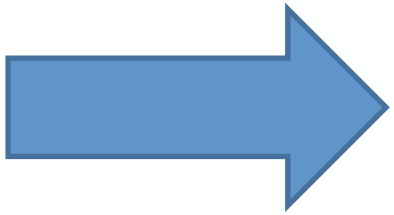
Current Services

Topology discovery

Path computation

State dissemination

Fault recovery

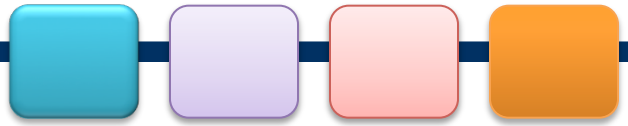


Abstract Services

Network Map

Intent-based Networking

SDN Abstractions



Network Applications
routing, access-control, mobility,
traffic-engineering, guarantees,
recovery, bandwidth-on-demand

Northbound API

Network Map

Control Plane

State Collection
State Dissemination &
Application Isolation

Built for Performance
Scale & Reliability

Southbound API

Tables for identifiers and actions

Flow table

Flows are combination of

L4
L3
L2.5
L2
L1

Softw ed Net

Example of network application
Classic imperative approach

Objective: establish a connection between Host 1 and Host 2

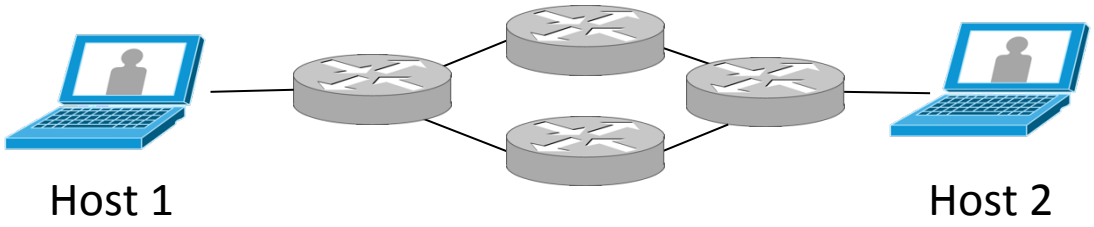
Host 1

Host 2

Example of network application
Classic imperative approach

Objective: establish a connection between Host 1 and Host 2

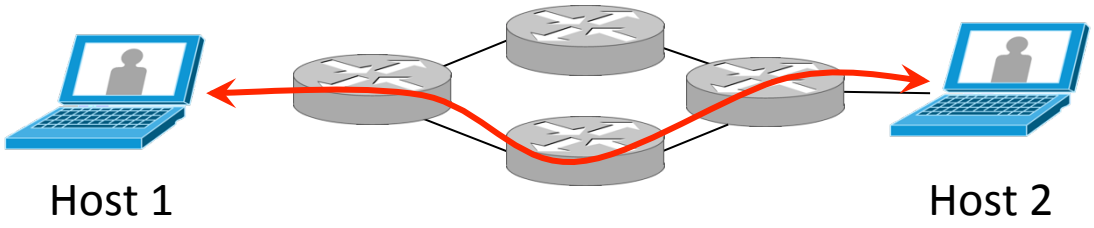
1. Discover network topology



Example of network application
Classic imperative approach

Objective: establish a connection between Host 1 and Host 2

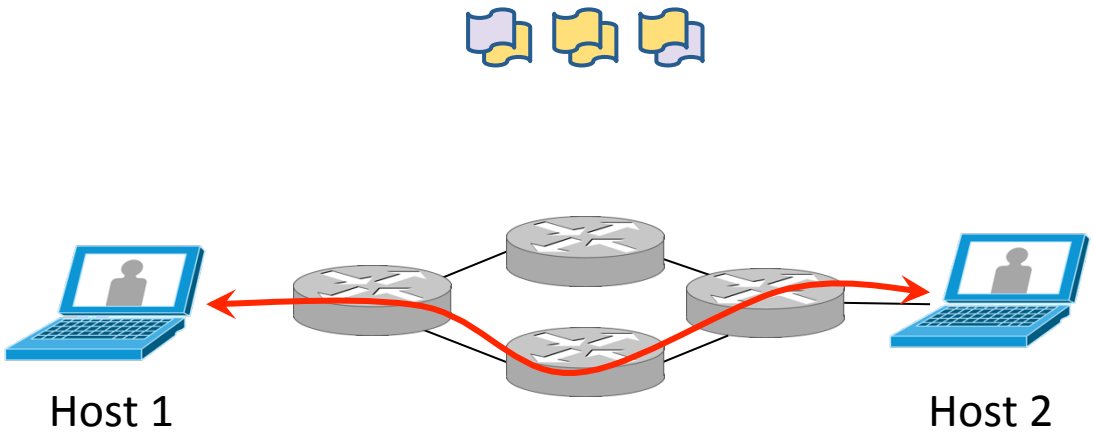
1. Discover network topology
2. Determine the path



Example of network application
Classic imperative approach

Objective: establish a connection between Host 1 and Host 2

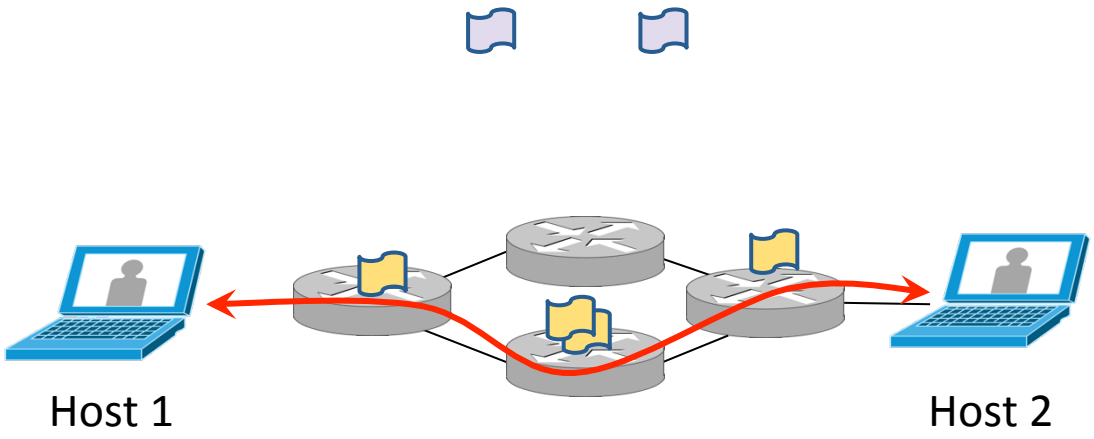
1. Discover network topology
2. Determine the path
3. Write the rules that define the flows and corresponding actions



Example of network application
Classic imperative approach

Objective: establish a connection between Host 1 and Host 2

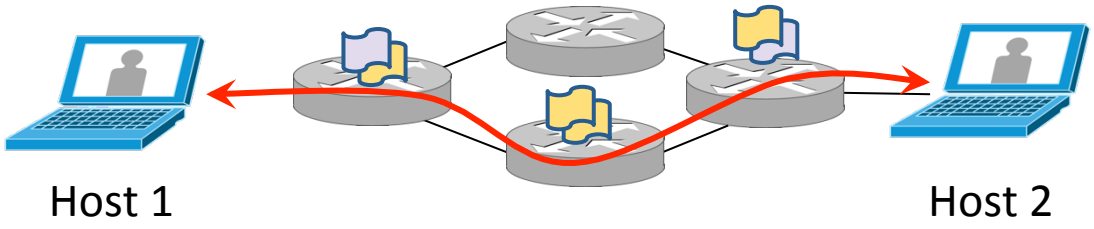
1. Discover network topology
2. Determine the path
3. Write the rules that define the flows and corresponding actions
4. Install rules on devices



Example of network application
Classic imperative approach

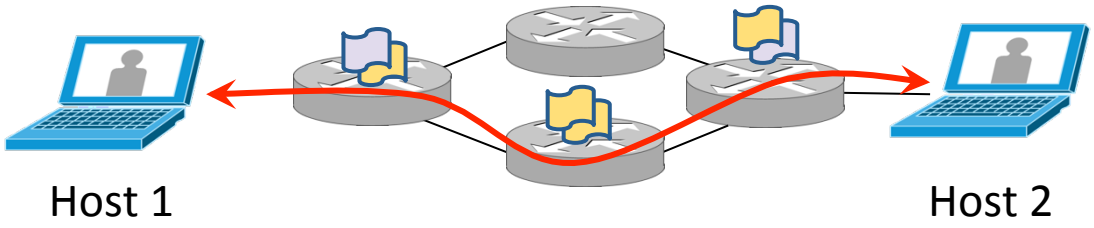
Objective: establish a connection between Host 1 and Host 2

1. Discover network topology
2. Determine the path
3. Write the rules that define the flows and corresponding actions
4. Install rules on devices



Example of network application Problems

This approach may fail in several ways

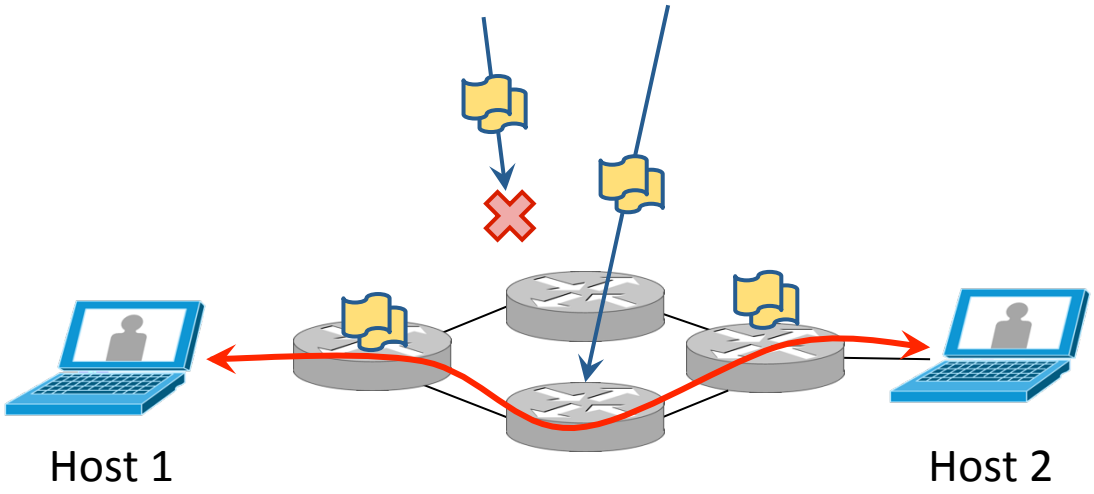


Example of network application Problems

This approach may fail in several ways

Missing rules, refused or cancelled

- Continuosly control that devices can be reached
- Guarantee that a consistent state is reached between two updates



Example of network application Problems

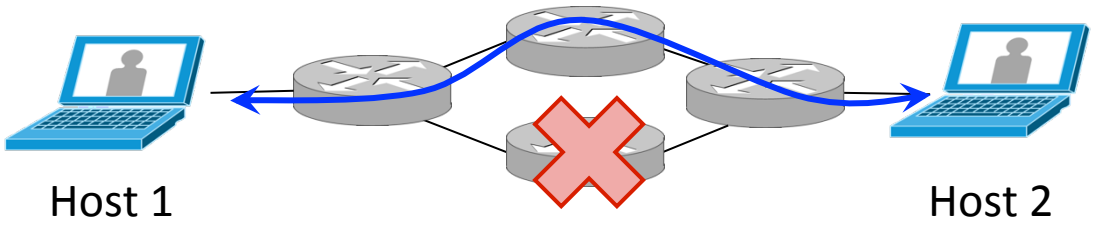
This approach may fail in several ways

Missing rules, refused or cancelled

- Continuously control that devices can be reached
- Guarantee that a consistent state is reached between two updates

Topology Modifications

- Listen to failure events from all devices and links
- Compute new paths and new flows



Programming Network Applications

Each application requires the calculation of routing paths, the installation of rules, the updating of state machines

In the event of failures, we risk inconsistent behavior

Bugs need to be fixed at various points in the network

Updating algorithms involving multiple applications is expensive

Difficult to resolve conflicts between applications

Programming Network Applications
Declarative Programming (intent-based networking)

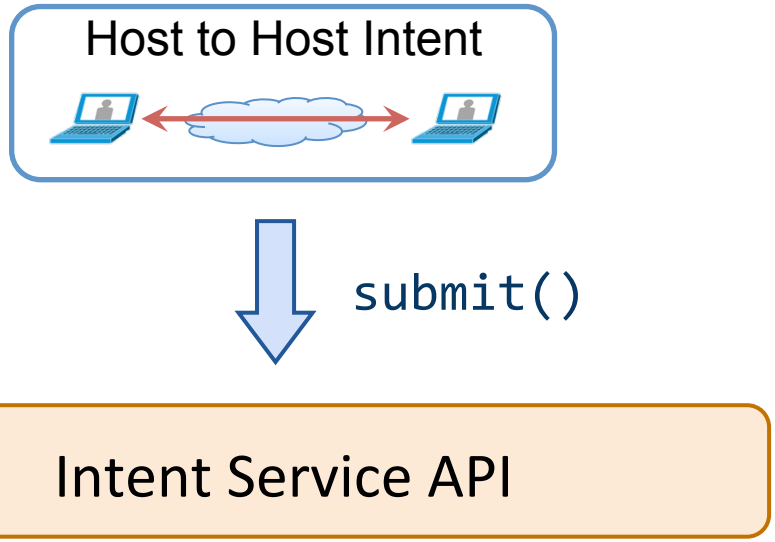
Network intentions are a high-level interface that describes *which result* you want to achieve and *delegates* how to get it to the network services layer

It hides the complexity of the network from applications

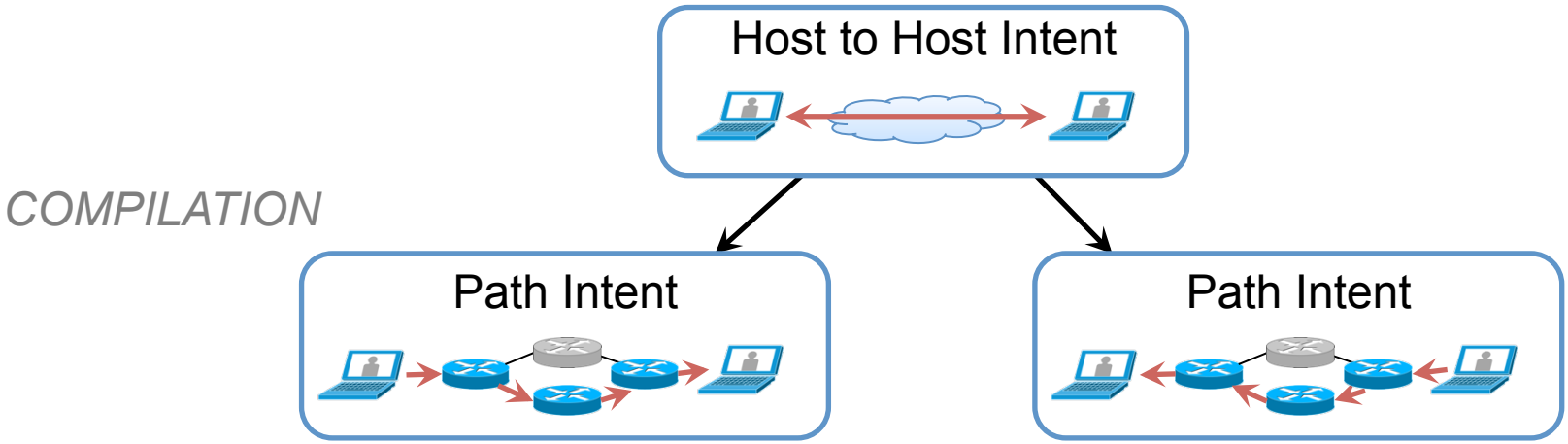
It guarantees the maintenance of the result even in the presence of *topology changes*

Intent Example

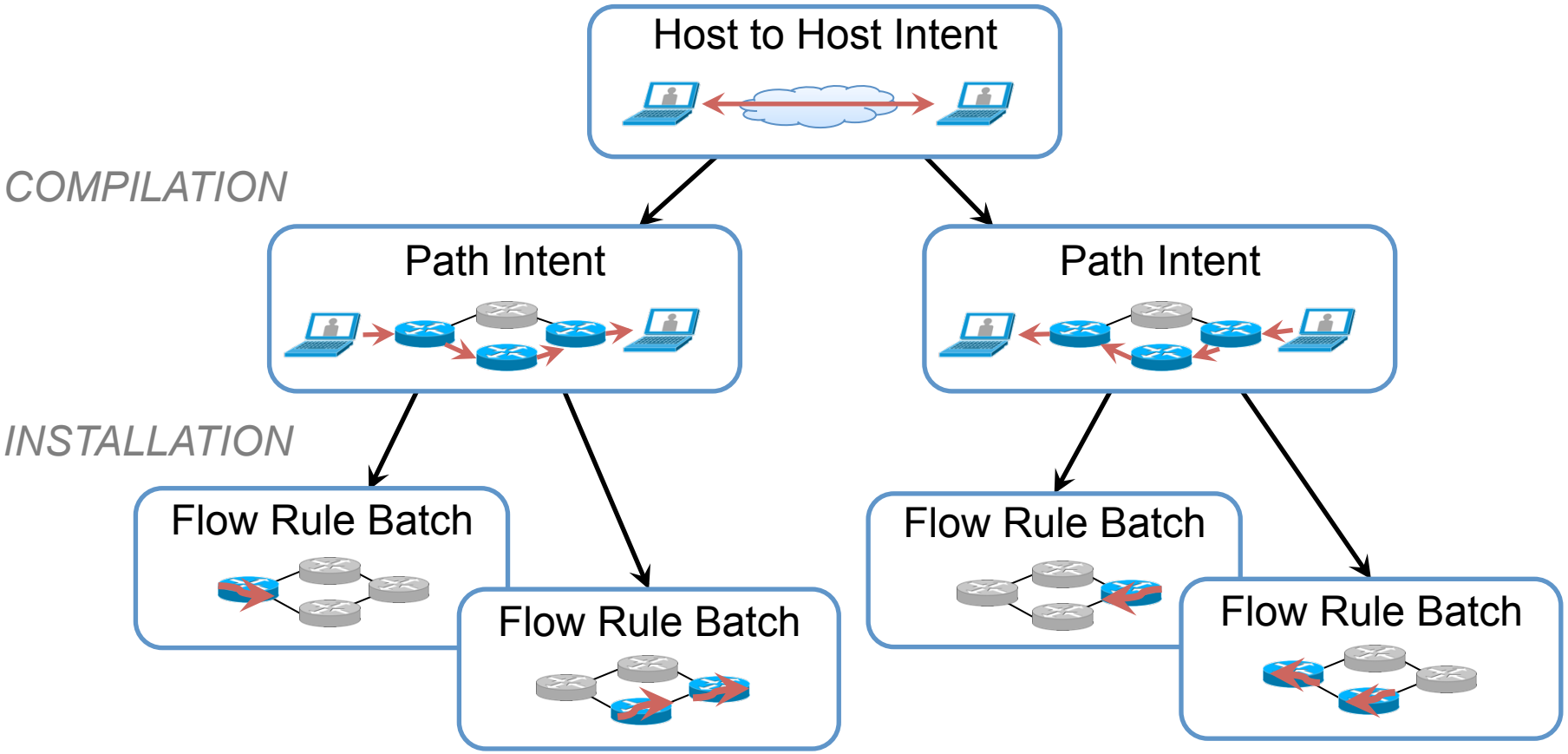
Intent Example



Intent Example



Intent Example



SDN in action

SDN in action

NETWORKWORLD FROM IDG

INSIDER Sign In

NETWORK INTELLIGENCE

By Zeus Kerravala, Network World | MAY 25, 2017 10:58 AM PT

About

Zeus Kerravala is the founder and principal analyst with ZK Research, and provides a mix of tactical advice to help his clients in the current business climate.

OPINION

Cisco to network engineers: Get comfortable with software. It's here to stay

In this digital software-driven world, where companies must move with speed, software skills are now a must for network engineers

Google's B4 Architecture