

Přenos dat, počítačové sítě a protokoly

Libor Polčák

Software Defined Networking

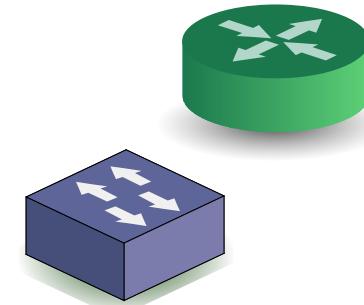
Obsah

- Motivace
- Předchůdci SDN
- Typická architektura SDN
 - ◆ Kontrolér, datová a řídící vrstva
 - ◆ OpenFlow
 - ◆ Nástroje
- Související technologie, virtualizace sítí

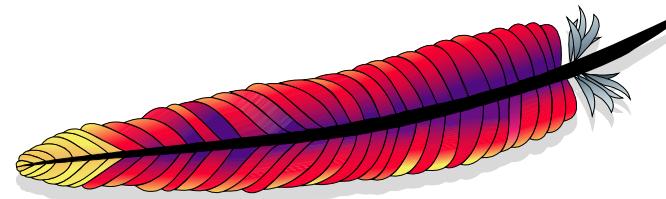
Motivace pro Software Defined Networking

Síťová zařízení

- HW a SW pevně svázány
 - ◆ Pevně daná funkcionalita
 - ◆ Uzavřenost inovacím
- Složitý a negranulární upgrade
- Operační systém milióny řádků
 - ◆ VLAN
 - ◆ Směrovací protokoly
 - ◆ Málo používané technologie
 - ◆ ...



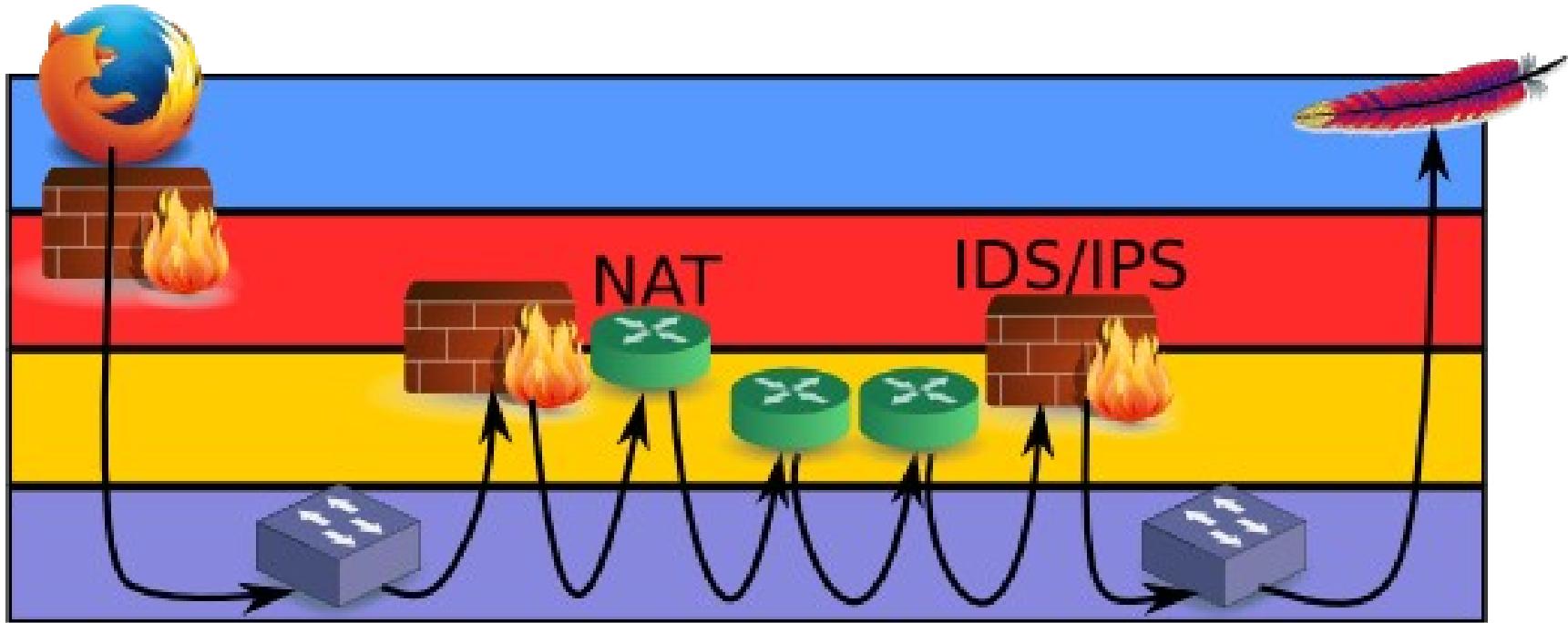
Konvergence sítí, priorita aplikací



Tradiční sítě



Sítě v praxi

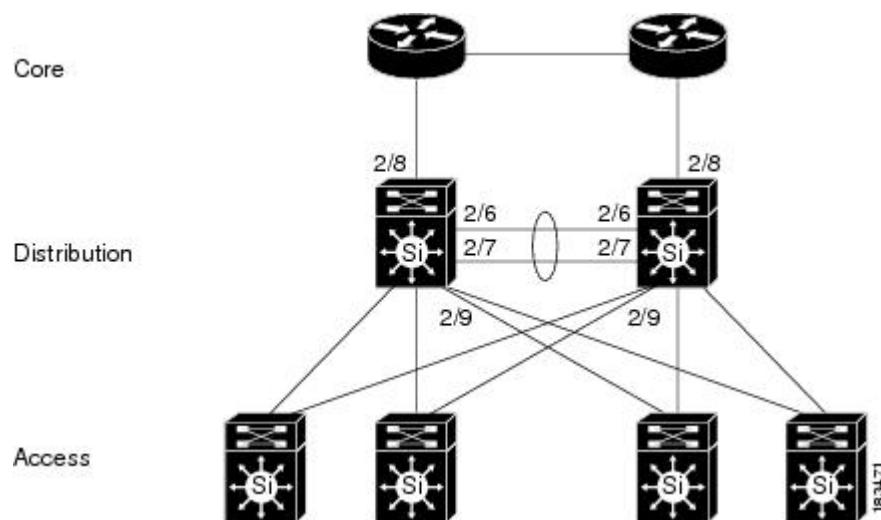


Sítě v praxi

- Směrovač
 - ◆ Firewall, IDS/IPS (L3-L4, příp. L7)
 - ◆ NAT (L3, L4, příp. L7)
 - ◆ Application Gateway (L7)
- Přepínač
 - ◆ IGMP snooping (L3)
 - ◆ DHCP snooping (L7)



Redundance, smyčky

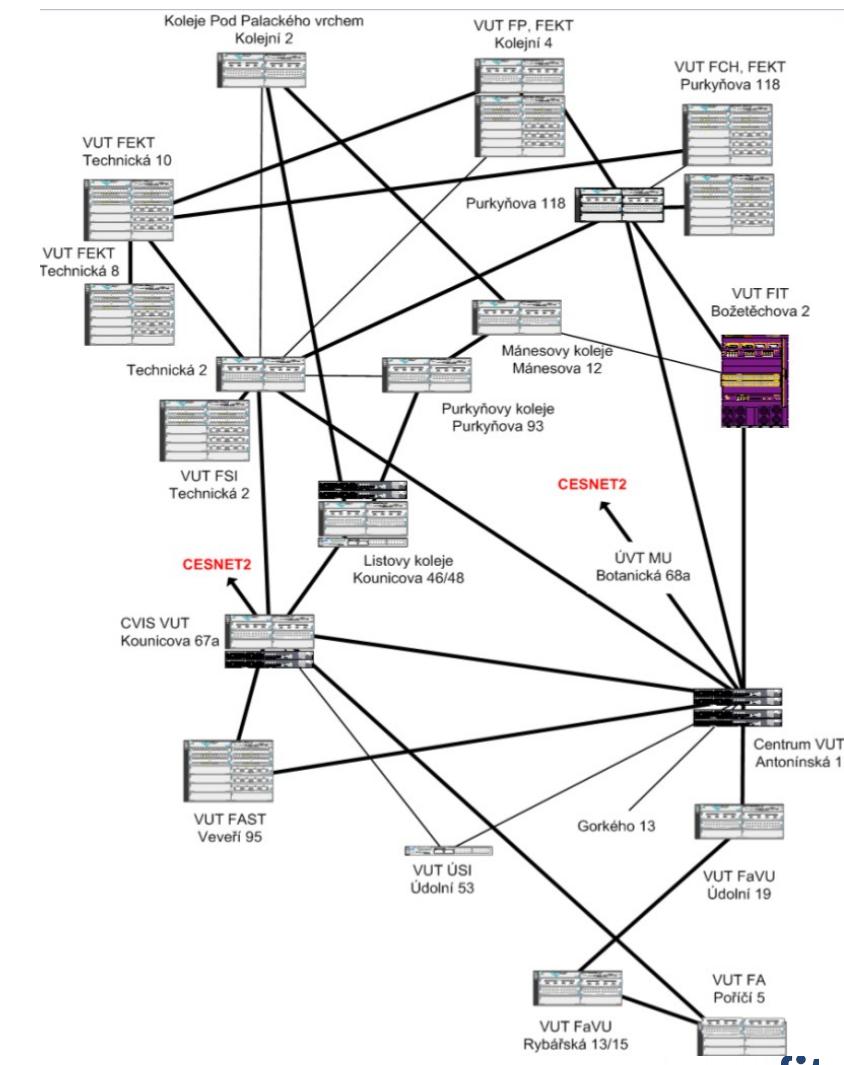
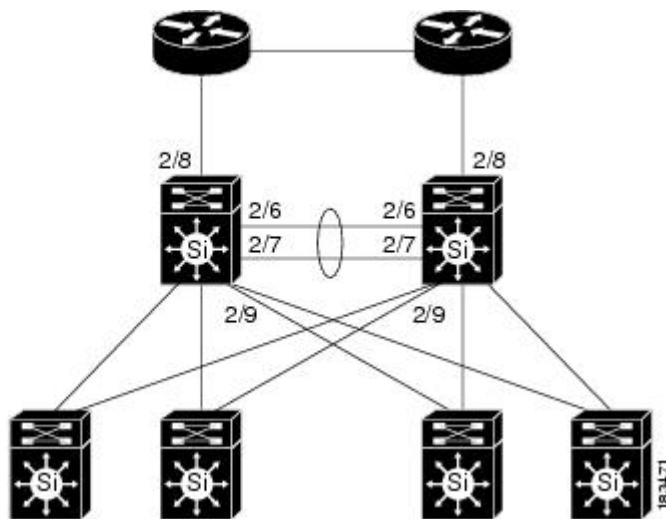


Redundance, smyčky

Core

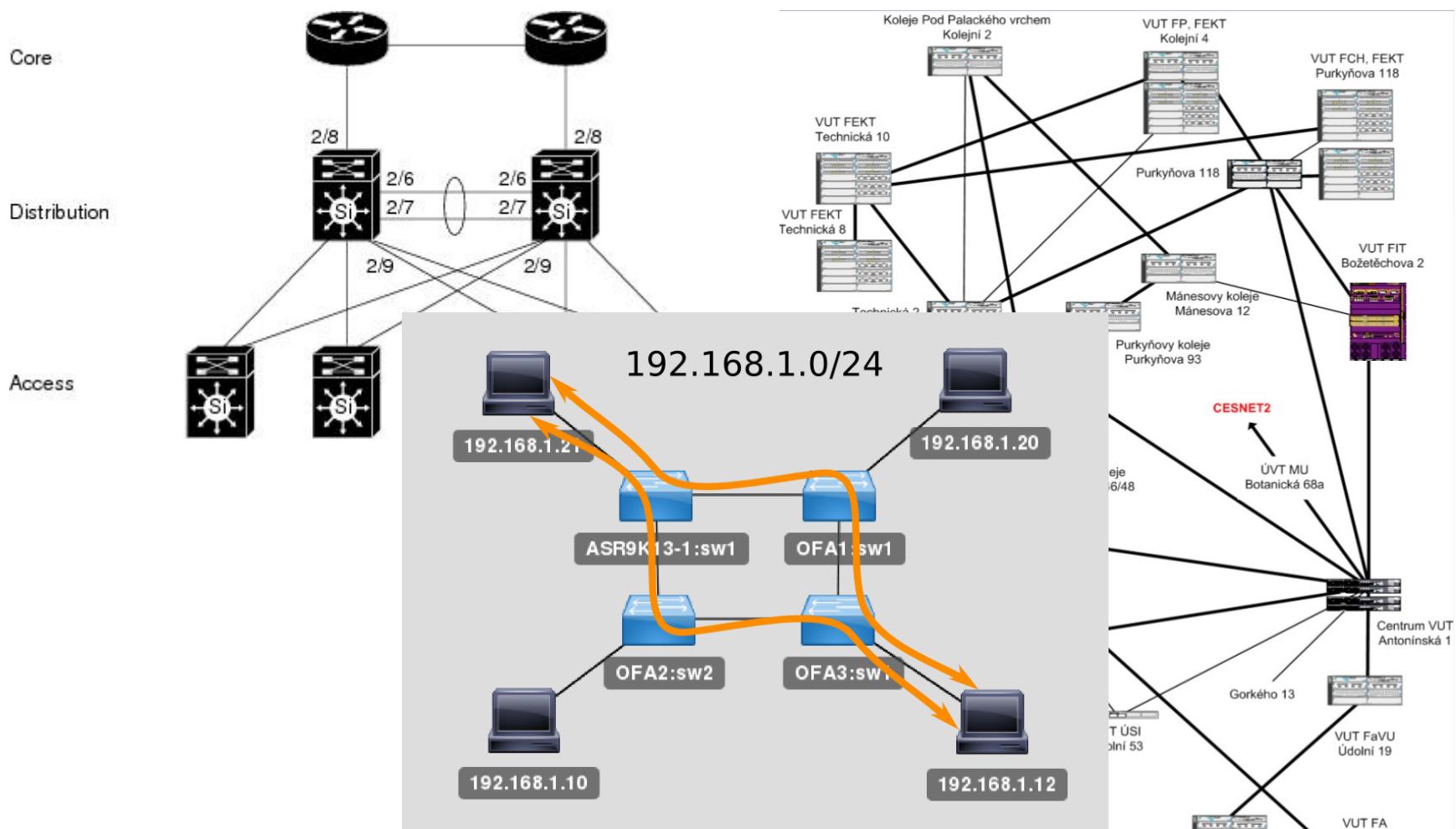
Distribution

Access



nesat.fit

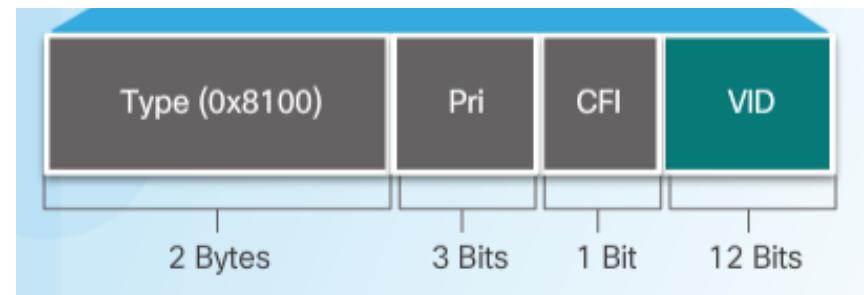
Redundance, smyčky



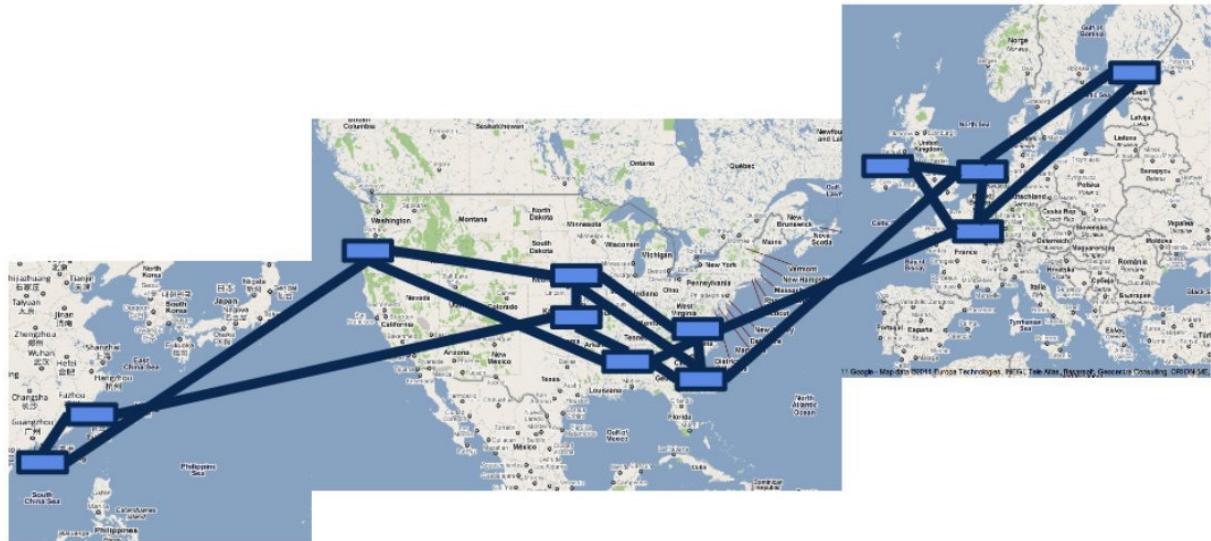
- Problém STP

VLAN: IEEE 802.1Q

- Pouze 4096 VLAN
- Co když máme desetitisíce zákazníků?



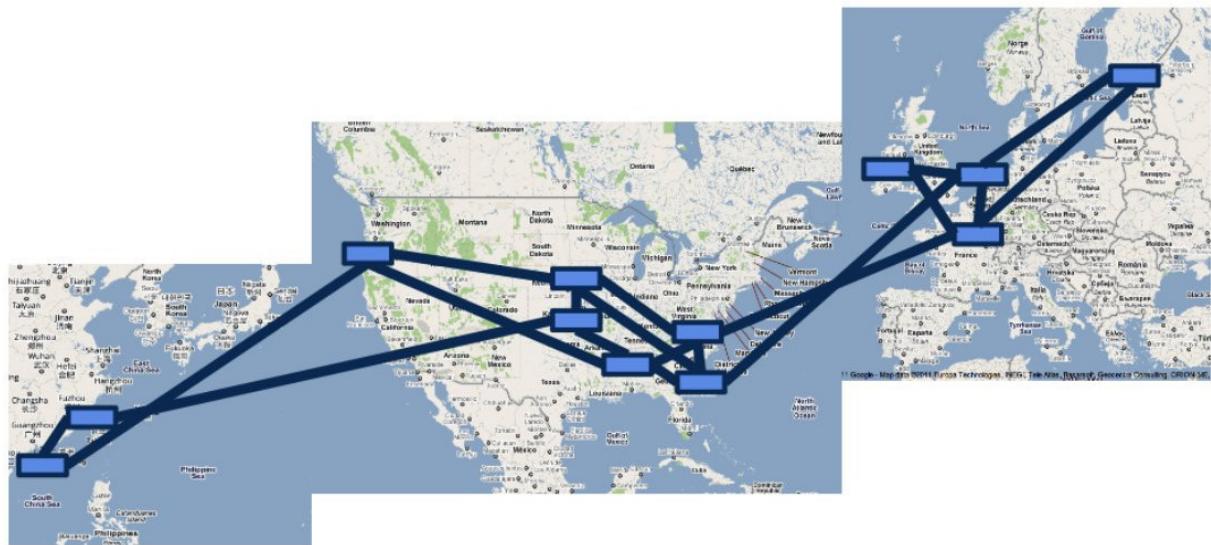
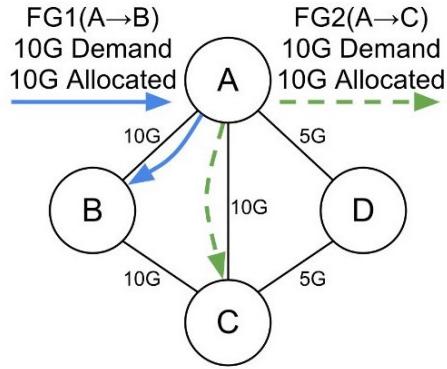
Datová centra/cloud



B4 worldwide deployment (2011).

Sushant Jain, Alok Kumar, Subhasree Mandal, Joon Ong, Leon Poutievski, Arjun Singh, Subbaiah Venkata, Jim Wanderer, Junlan Zhou, Min Zhu, Jon Zolla, Urs Hözle, Stephen Stuart, and Amin Vahdat. 2013. B4: experience with a globally-deployed software defined wan. In Proceedings of the ACM SIGCOMM 2013 conference on SIGCOMM (SIGCOMM '13). ACM, New York, NY, USA, 3-14.

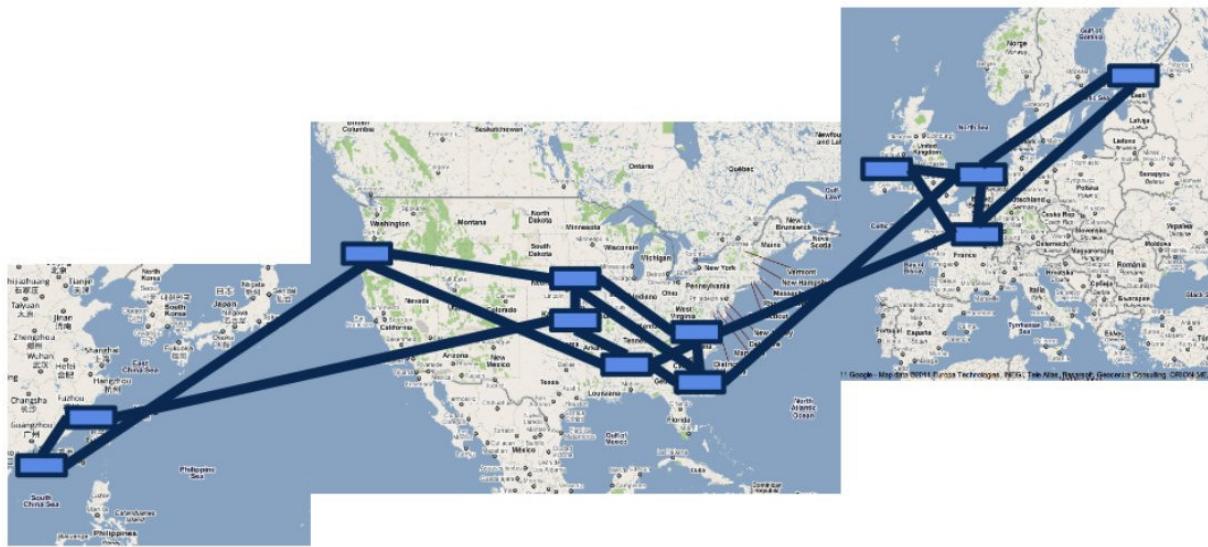
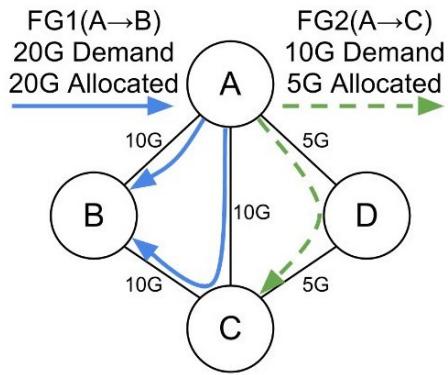
Datová centra/cloud



B4 worldwide deployment (2011).

Sushant Jain, Alok Kumar, Subhasree Mandal, Joon Ong, Leon Poutievski, Arjun Singh, Subbaiah Venkata, Jim Wanderer, Junlan Zhou, Min Zhu, Jon Zolla, Urs Hözle, Stephen Stuart, and Amin Vahdat. 2013. B4: experience with a globally-deployed software defined wan. In Proceedings of the ACM SIGCOMM 2013 conference on SIGCOMM (SIGCOMM '13). ACM, New York, NY, USA, 3-14.

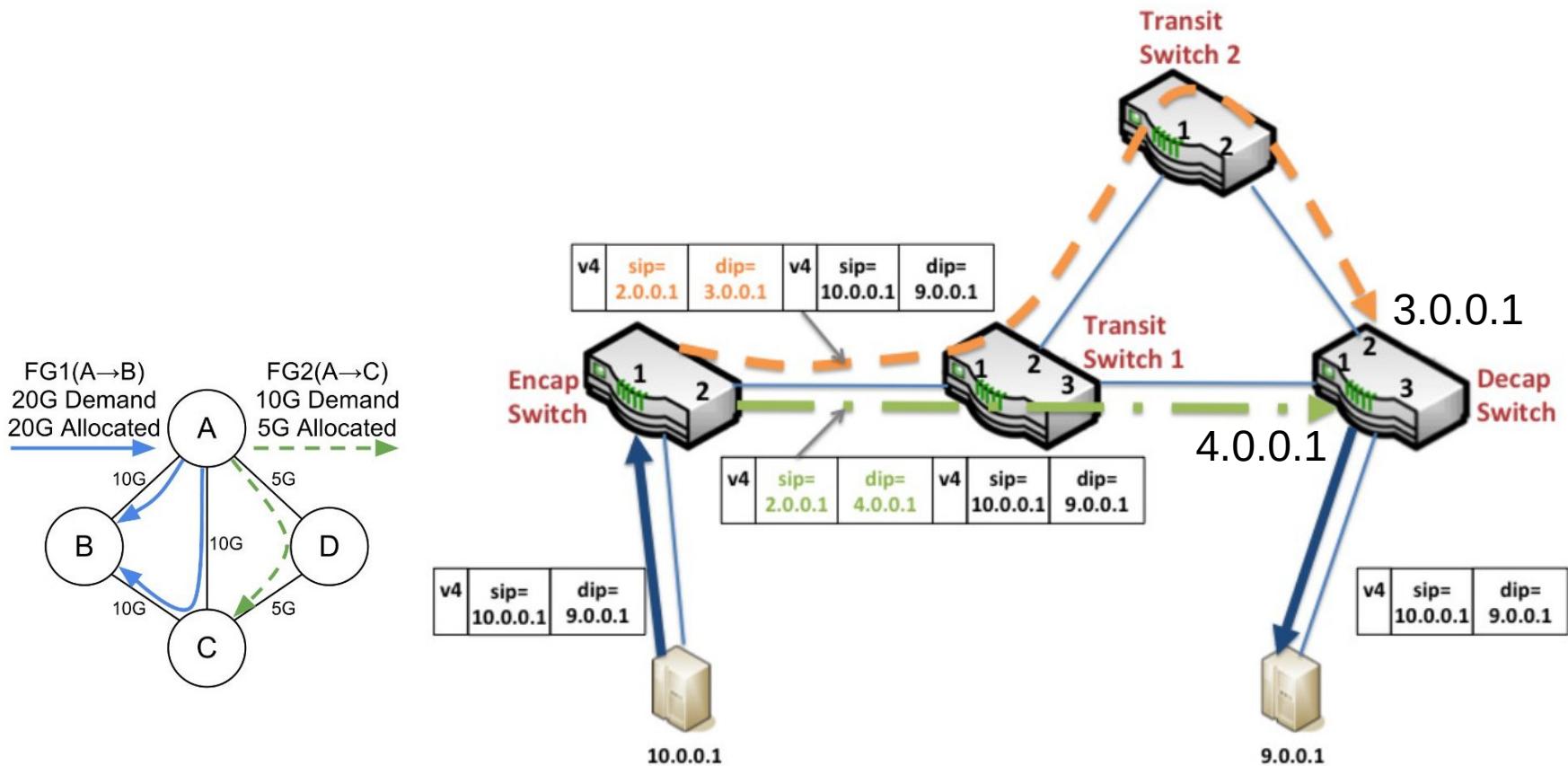
Datová centra/cloud



B4 worldwide deployment (2011).

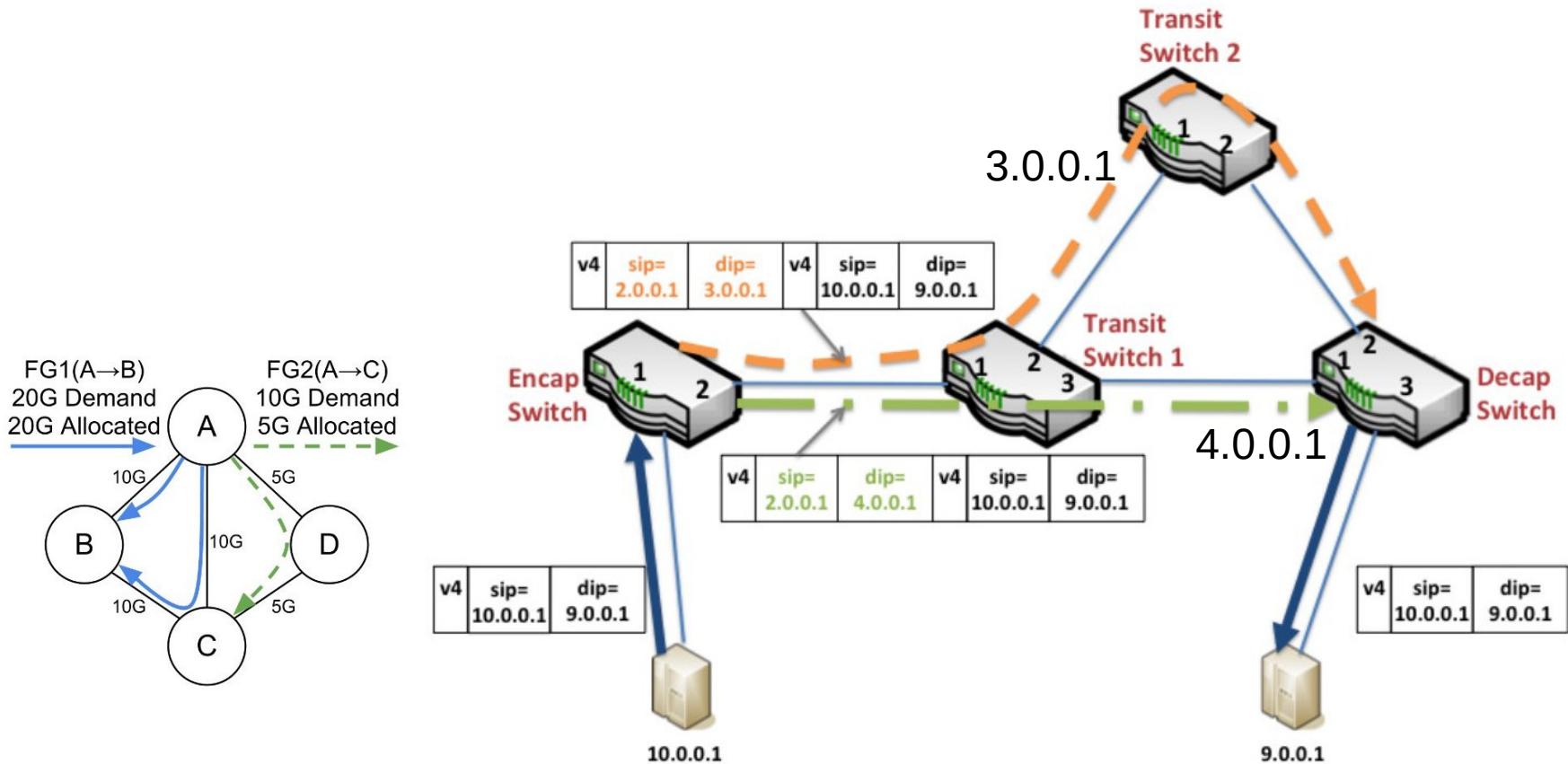
Sushant Jain, Alok Kumar, Subhasree Mandal, Joon Ong, Leon Poutievski, Arjun Singh, Subbaiah Venkata, Jim Wanderer, Junlan Zhou, Min Zhu, Jon Zolla, Urs Hözle, Stephen Stuart, and Amin Vahdat. 2013. B4: experience with a globally-deployed software defined wan. In Proceedings of the ACM SIGCOMM 2013 conference on SIGCOMM (SIGCOMM '13). ACM, New York, NY, USA, 3-14.

Datová centra/cloud



Sushant Jain, Alok Kumar, Subhasree Mandal, Joon Ong, Leon Poutievski, Arjun Singh, Subbaiah Venkata, Jim Wanderer, Junlan Zhou, Min Zhu, Jon Zolla, Urs Hözle, Stephen Stuart, and Amin Vahdat. 2013. B4: experience with a globally-deployed software defined wan. In Proceedings of the ACM SIGCOMM 2013 conference on SIGCOMM (SIGCOMM '13). ACM, New York, NY, USA, 3–14.

Datová centra/cloud



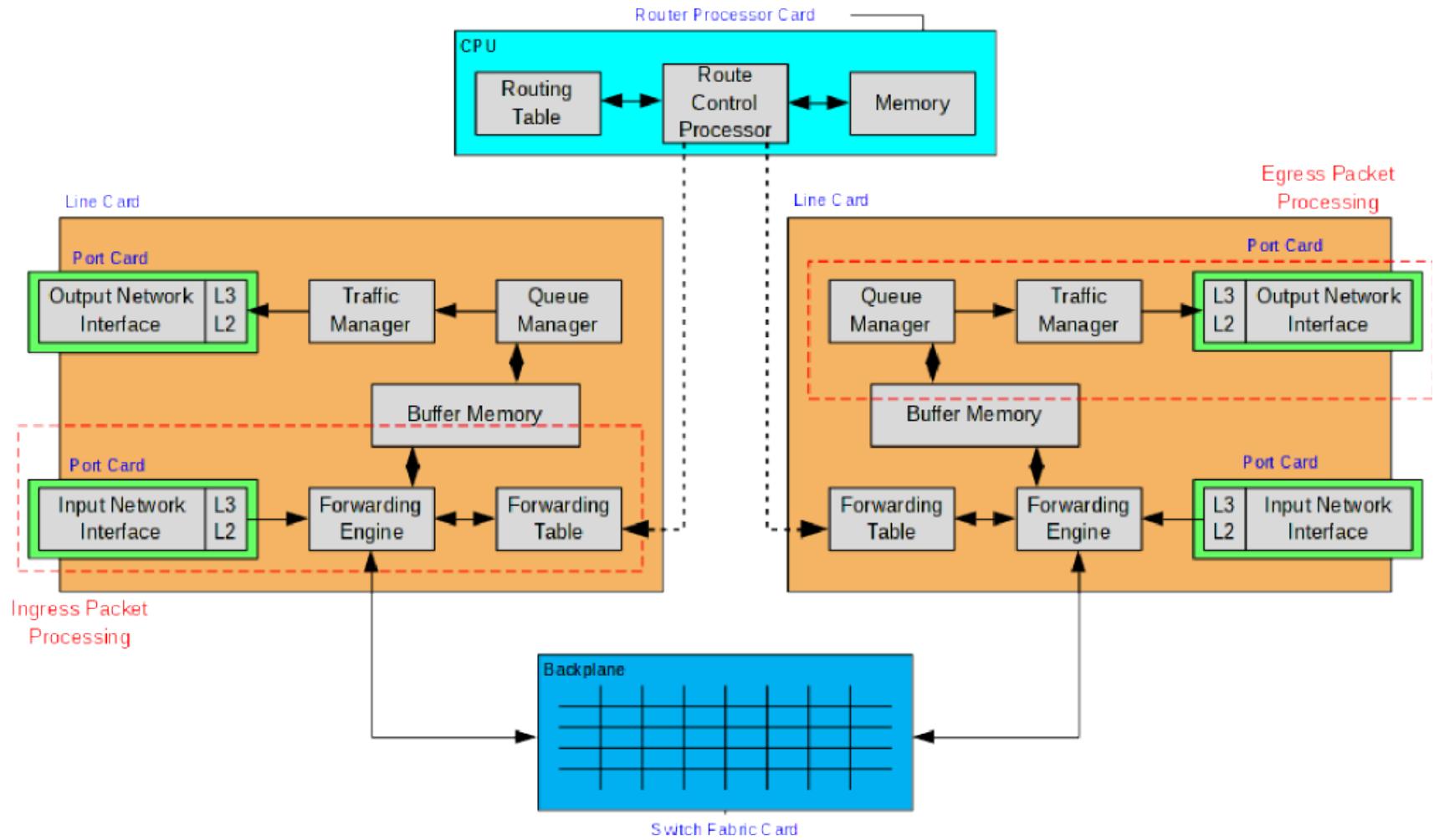
Sushant Jain, Alok Kumar, Subhasree Mandal, Joon Ong, Leon Poutievski, Arjun Singh, Subbaiah Venkata, Jim Wanderer, Junlan Zhou, Min Zhu, Jon Zolla, Urs Hözle, Stephen Stuart, and Amin Vahdat. 2013. B4: experience with a globally-deployed software defined wan. In Proceedings of the ACM SIGCOMM 2013 conference on SIGCOMM (SIGCOMM '13). ACM, New York, NY, USA, 3-14.

Shrnutí problémů

- Silná závislost na standardizovaných protokolech
 - ◆ Nepružné řízení síťových prvků
- Chtěli bychom vylepšit programovatelnost sítě
- Možnost inovace vs. Standardizace, implementace, rozšíření technologie

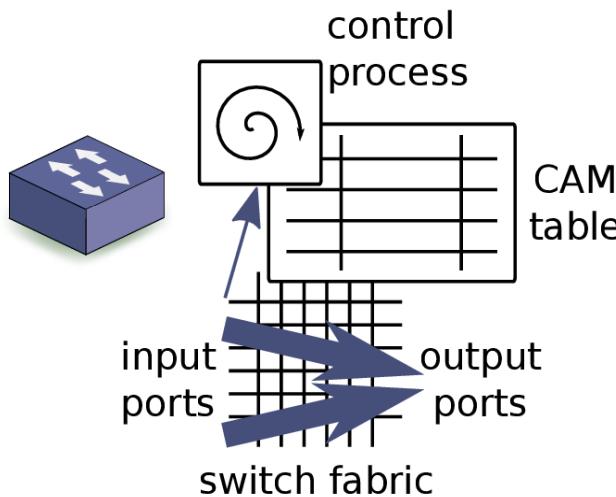
Jak vypadají aktivní prvky v síti?

Obecná architektura směrovače

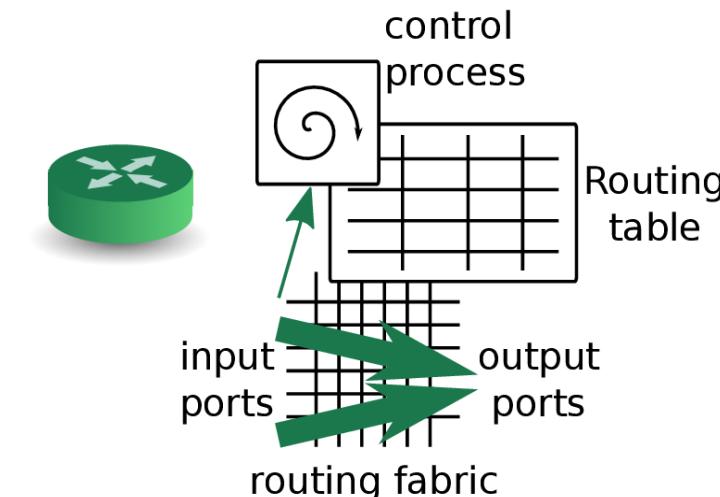


Směrovač/přepínač podrobněji

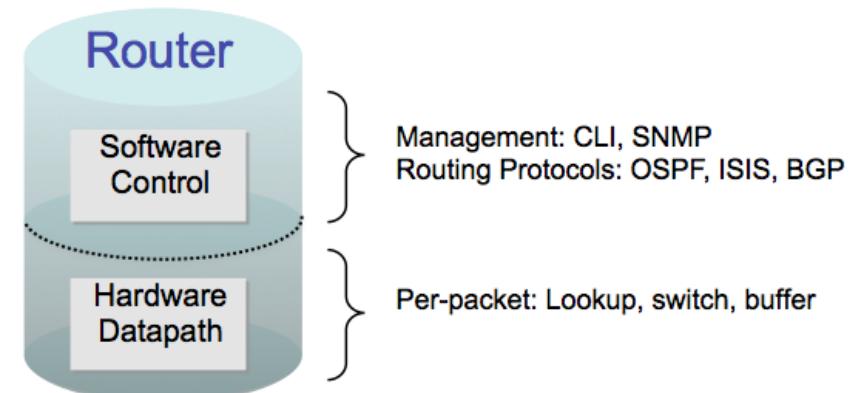
a) switch



b) router



- Požadavky na funkcionality
- Požadavky na rychlosť



Hledání vhodné architektury

Hledání vhodné architektury

Konfigurační nástroje

Hardwarová platforma

Hledání vhodné architektury

Aplikace

Abstrakce, otevřené
programovací prostředí (API)

Hardwarová platforma

Požadavky na vhodnou architekturu

- Hardwarová platforma
 - ◆ Rychlá, směrování, přepínání
- Programovací prostředí (API)
 - ◆ Podpora síťových protokolů
 - ◆ Řízení směrování
 - ◆ Inovace, nové strategie
- Aplikace
 - ◆ Jednoduché na vytváření

Aktivní sítě (Active networks)

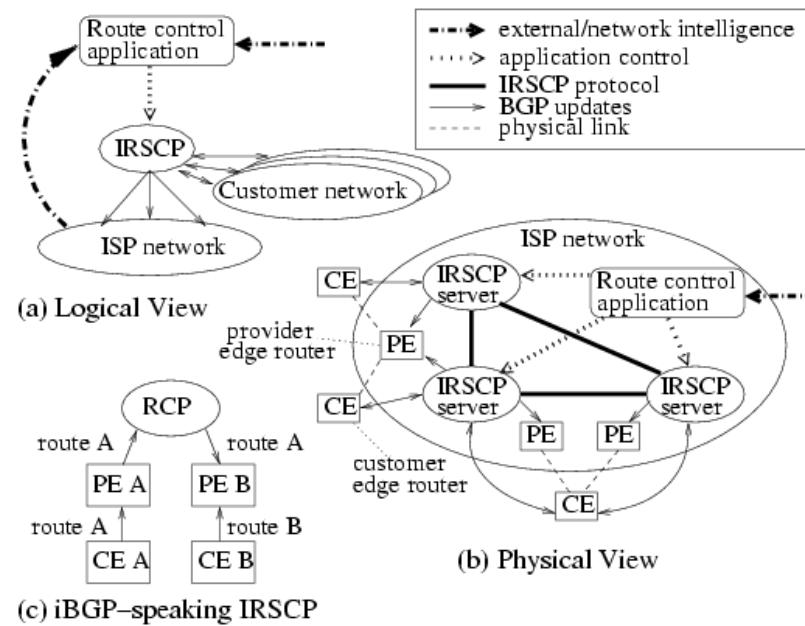
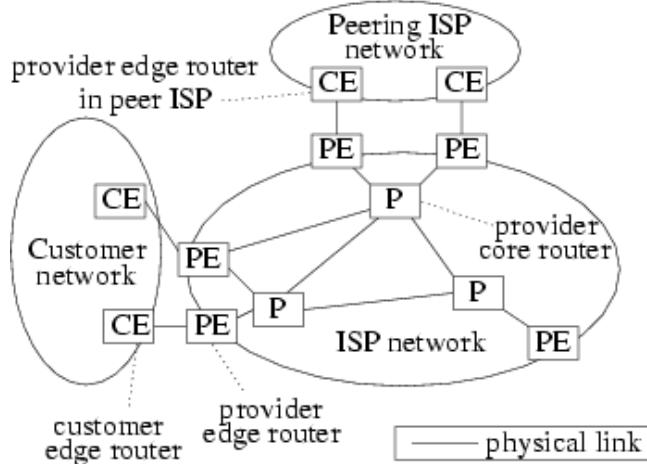
- Uživatelé vkládají kód, který vykonávají aktivní prvky po cestě
- Pakety → Kapsule (data+kód)



Tennenhouse, David L., David J. Wetherall. "Towards an active network architecture." DARPA Active Networks Conference and Exposition, 2002. Proceedings. IEEE, 2002.

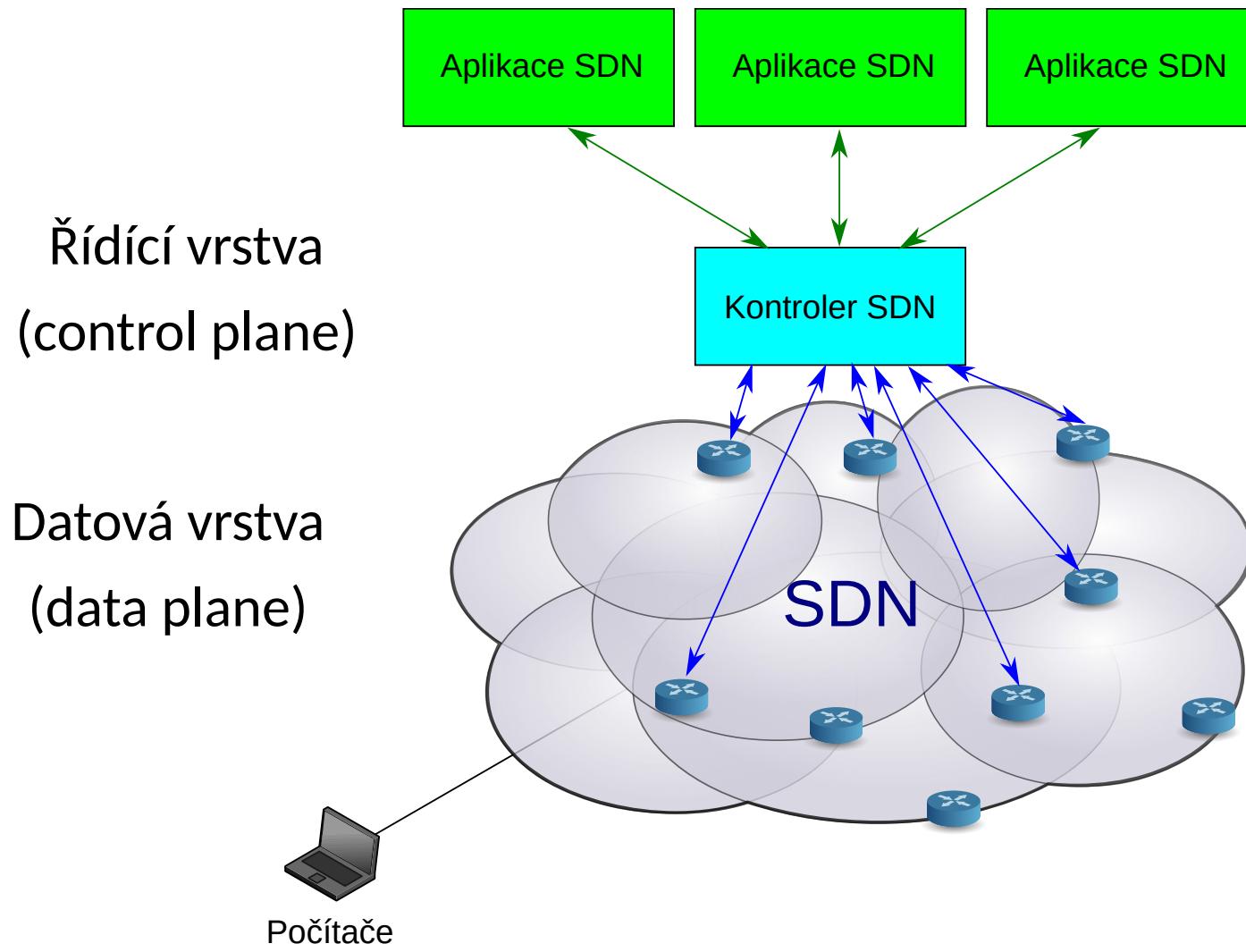
Řízení toků

- Přístupy jako Forwarding and Control Element Separation, Routing Control Platform, SoftRouter, Path Computation Element, Intelligent Route Service Control Point
- Oddělení data plane a control plane



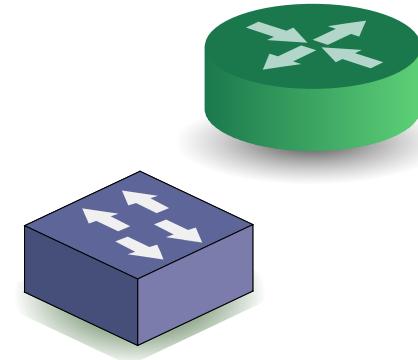
Software Defined Networking

Architektura SDN

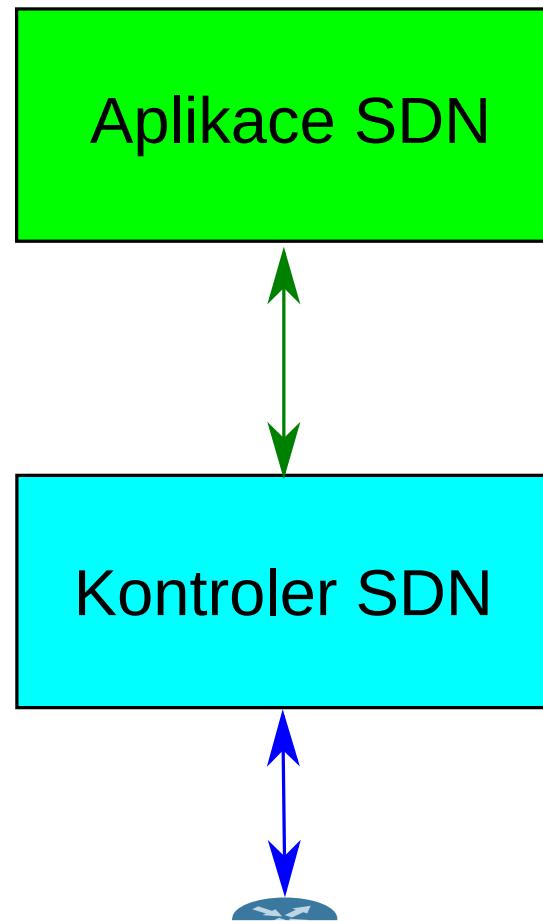


Síťová zařízení – datová vrstva

- Rychlé přenosy dat
- Přepínač SDN
 - ◆ Abstrakce od specifické činnosti
 - Směrovač, přepínač, firewall apod.
- Řízená kontrolérem



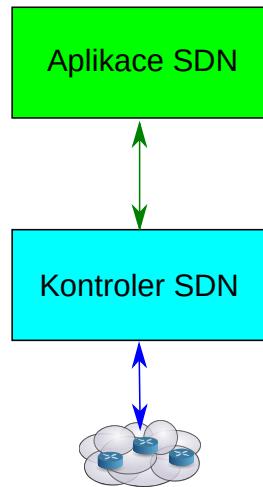
Řídící vrstva SDN



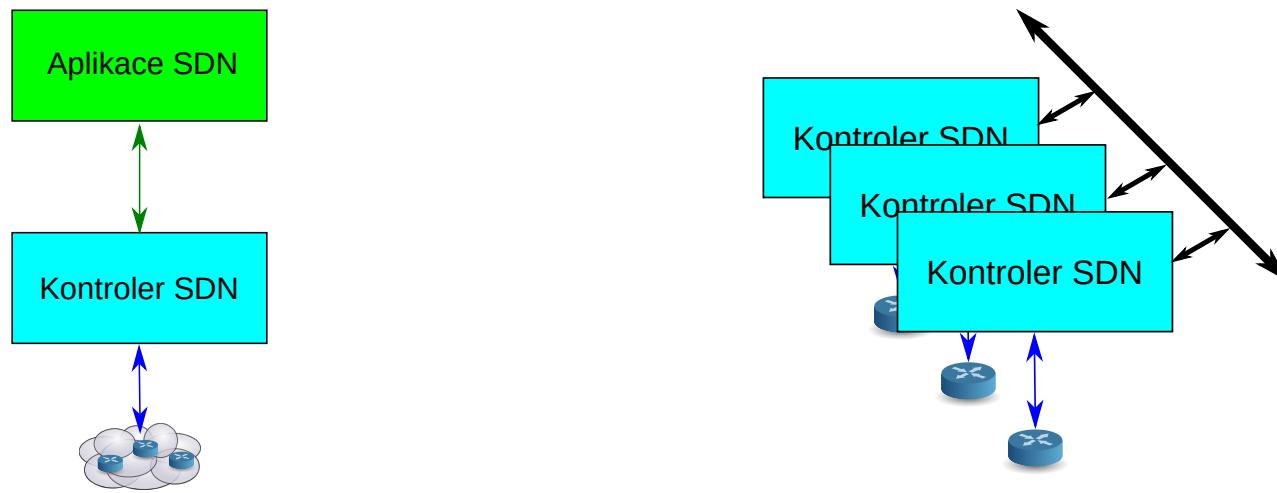
Vnitřnosti kontroléru

- Správa síťových uzlů
 - ◆ Koncové počítače připojené k síti
 - ◆ Topologie
- Správa toků
- Statistiky
- Směrování
 - ◆ Algoritmy pro výpočet cesty

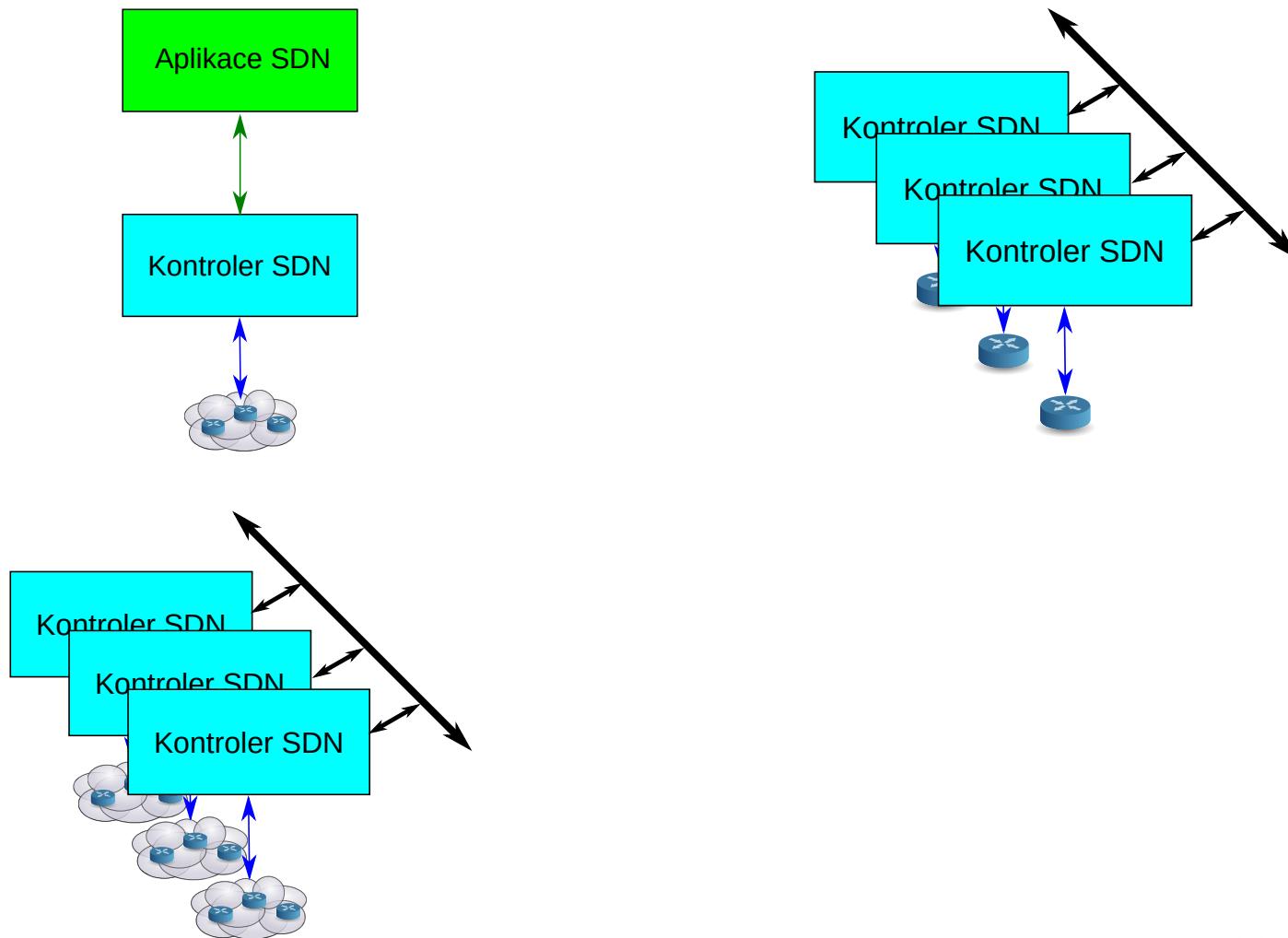
Řízení SDN



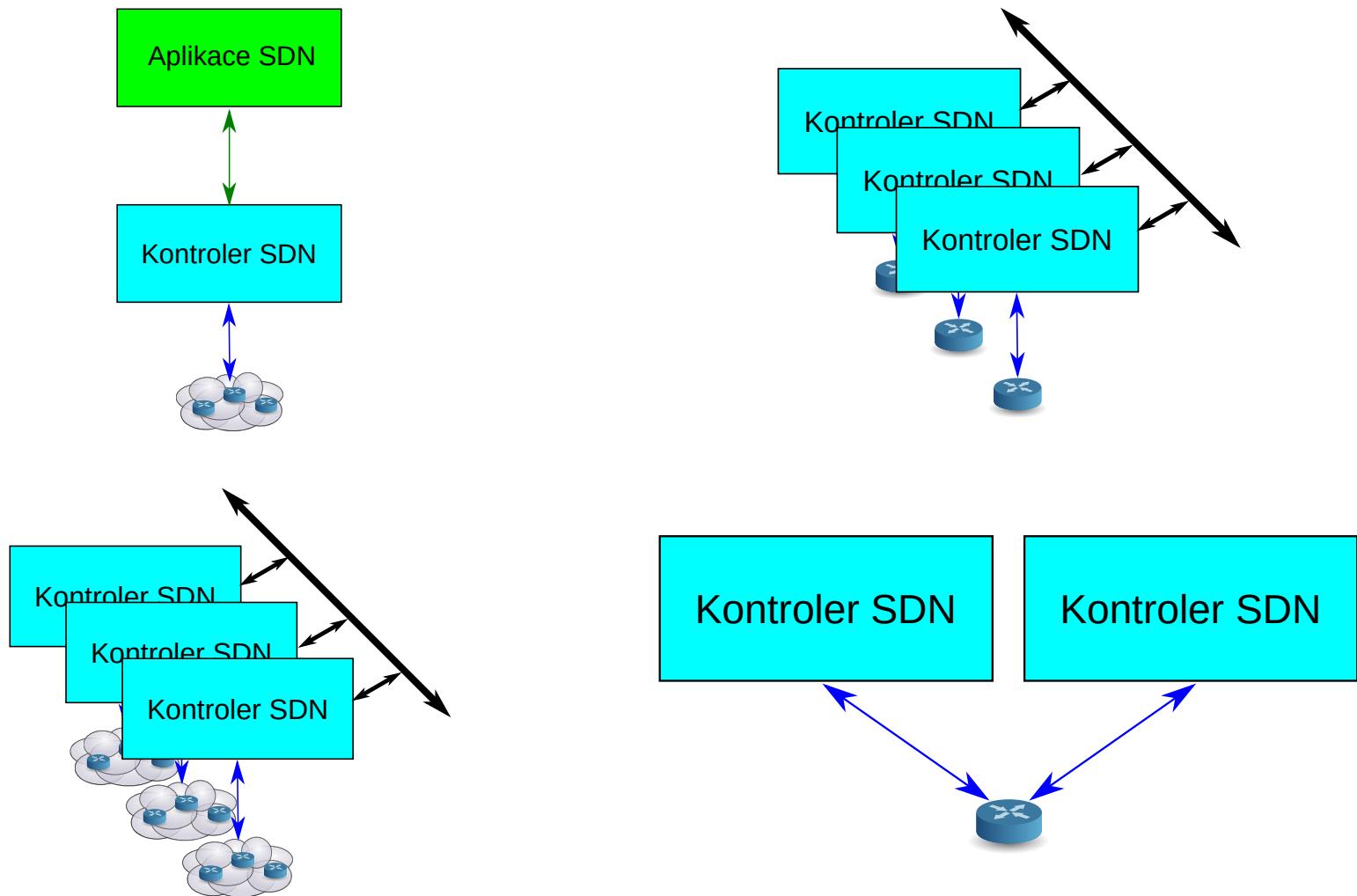
Řízení SDN



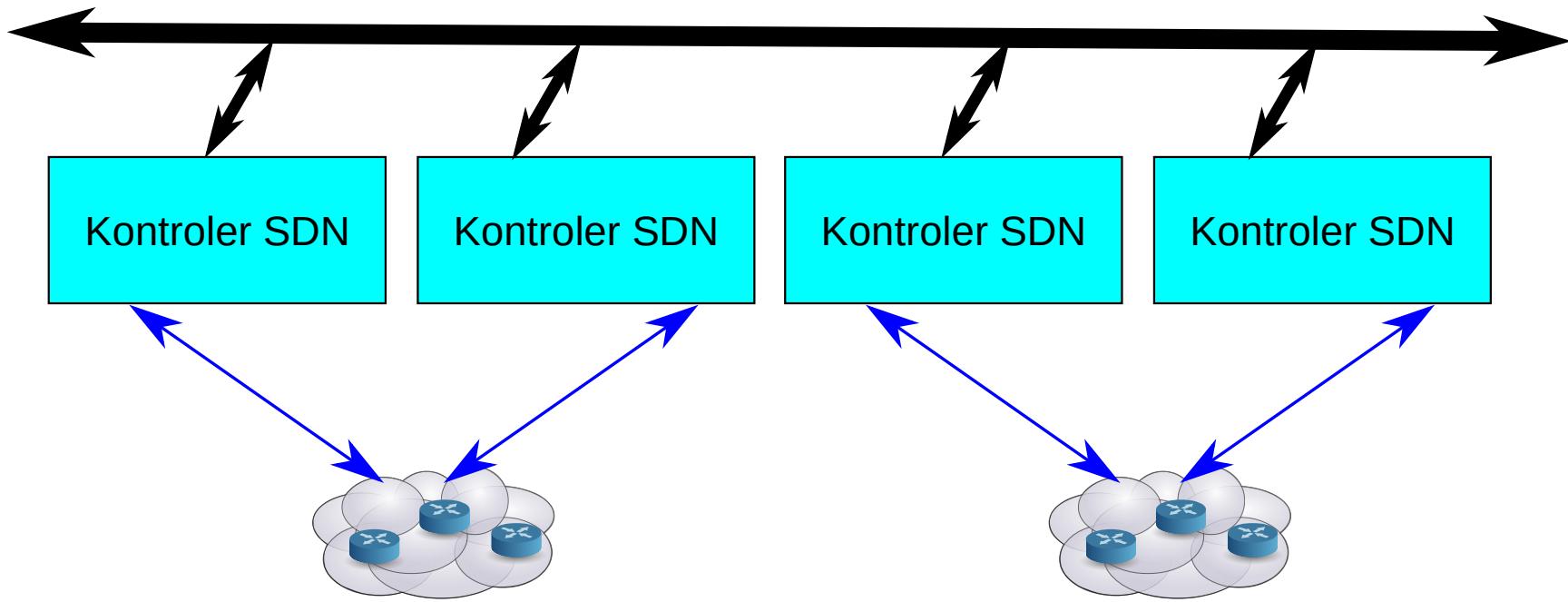
Řízení SDN



Řízení SDN

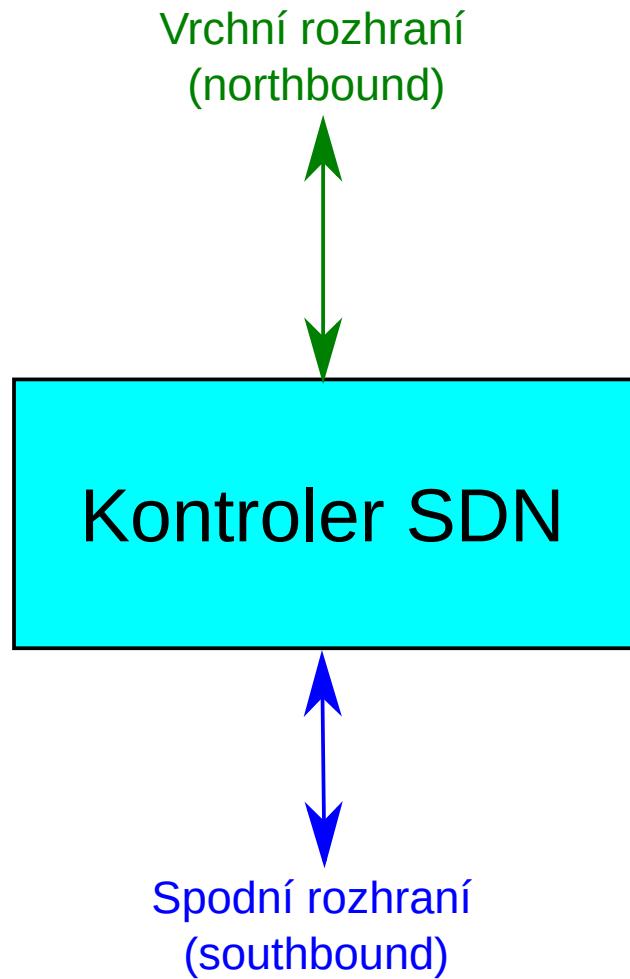


Řízení SDN

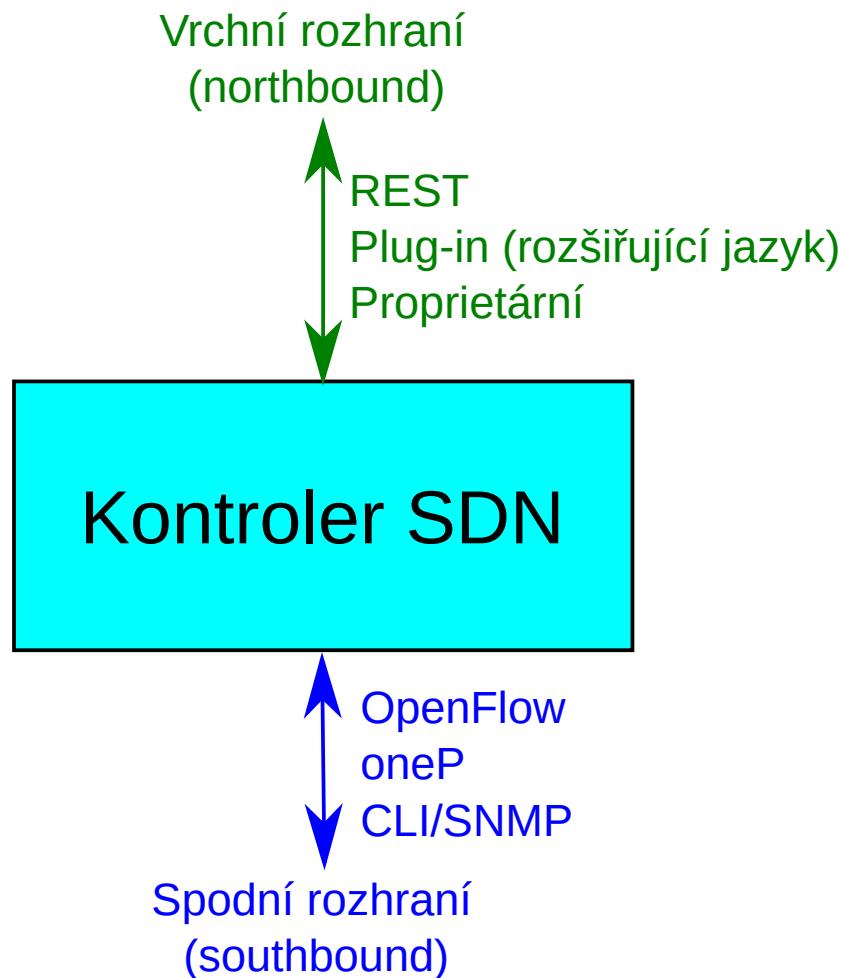


- Redundance
- Distribuované řešení
 - ◆ Koordinace
 - ◆ Centralizace control plane

Rozhraní kontroléru



Rozhraní kontroleru



Spodní rozhraní (Southbound)



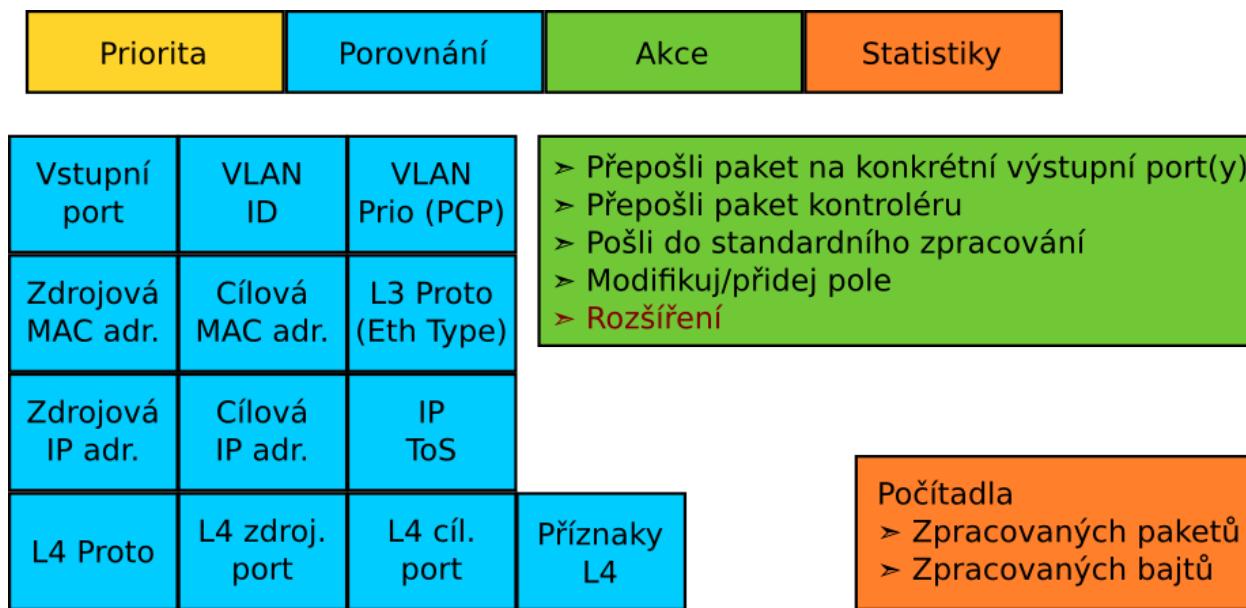
nesat.fit

Spodní rozhraní

- OpenFlow
 - ◆ Open Networking foundation
 - <https://www.opennetworking.org/technical-communities/areas/specification>
 - ◆ N. McKeown, T. Anderson, H. Balakrishnan, G. Parulkar, L. Peterson, J. Rexford, S. Shenker, and J. Turner: Openflow: enabling innovation in campus networks. Computer Communication Review, SIGCOMM, 2008.

OpenFlow

- Binární protokol
- V současné době verze 1.0-1.5
- Přepínač OF
- Pravidla:



Emulace různých síťových prvků

Přepínání

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

Emulace různých síťových prvků

Zpracování na úrovni mikro-flow

Switch Port	MAC src	MAC dst	Eth type	VLAN ID	IP Src	IP Dst	IP Prot	TCP sport	TCP dport	Action
port3	00:20..	00:1f..	0800	vlan1	1.2.3.4	5.6.7.8	4	17264	80	port6

Emulace různých síťových prvků

Firewall

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

Emulace různých síťových prvků

Směrování

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

Emulace různých síťových prvků

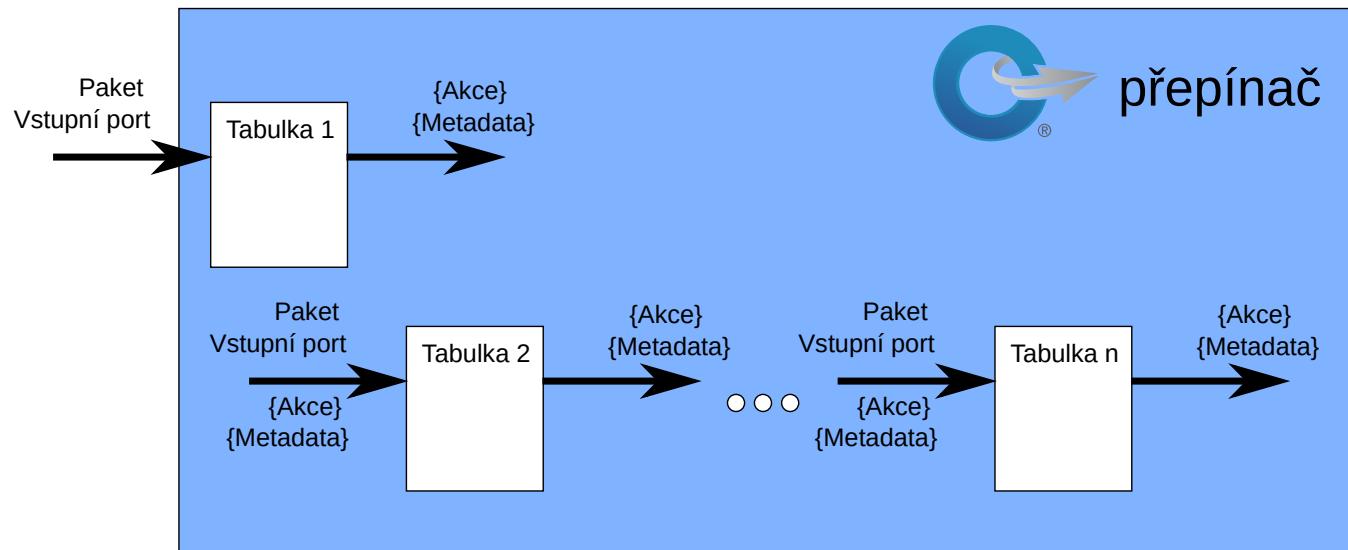
Multicast uvnitř VLAN

Switch Port	MAC src	MAC dst	Eth type	VLAN ID	IP Src	IP Dst	IP Prot	TCP sport	TCP dport	Action
*	*	00:1f..	*	vlan1	*	*	*	*	*	port6, port7, port9

Chování přepínače OF

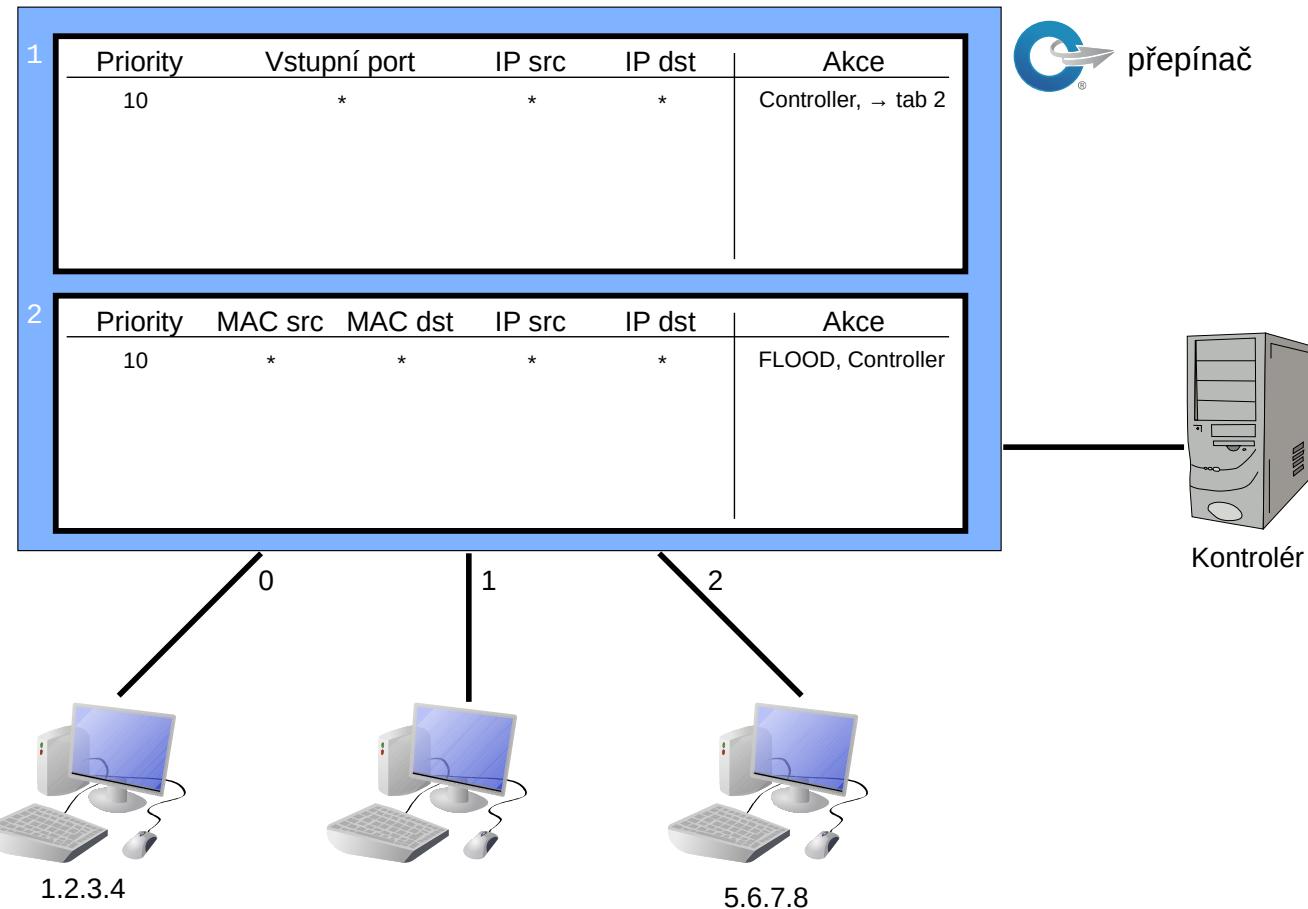
- Reaktivní
 - ◆ První paket neznámého toku přeposlán do kontroléru
 - ◆ Kontrolér vytvoří požadovaná pravidla
 - ◆ Režie spojená s komunikací
 - ◆ Typicky mikrotoky
 - Kontrolér získává detailnější informace o dění na síti
- Proaktivní
 - ◆ Pravidla předinstalována kontrolerem
 - ◆ Kontroler musí znát dopředu očekávané toky
 - Typicky agregace

Tabulky OF

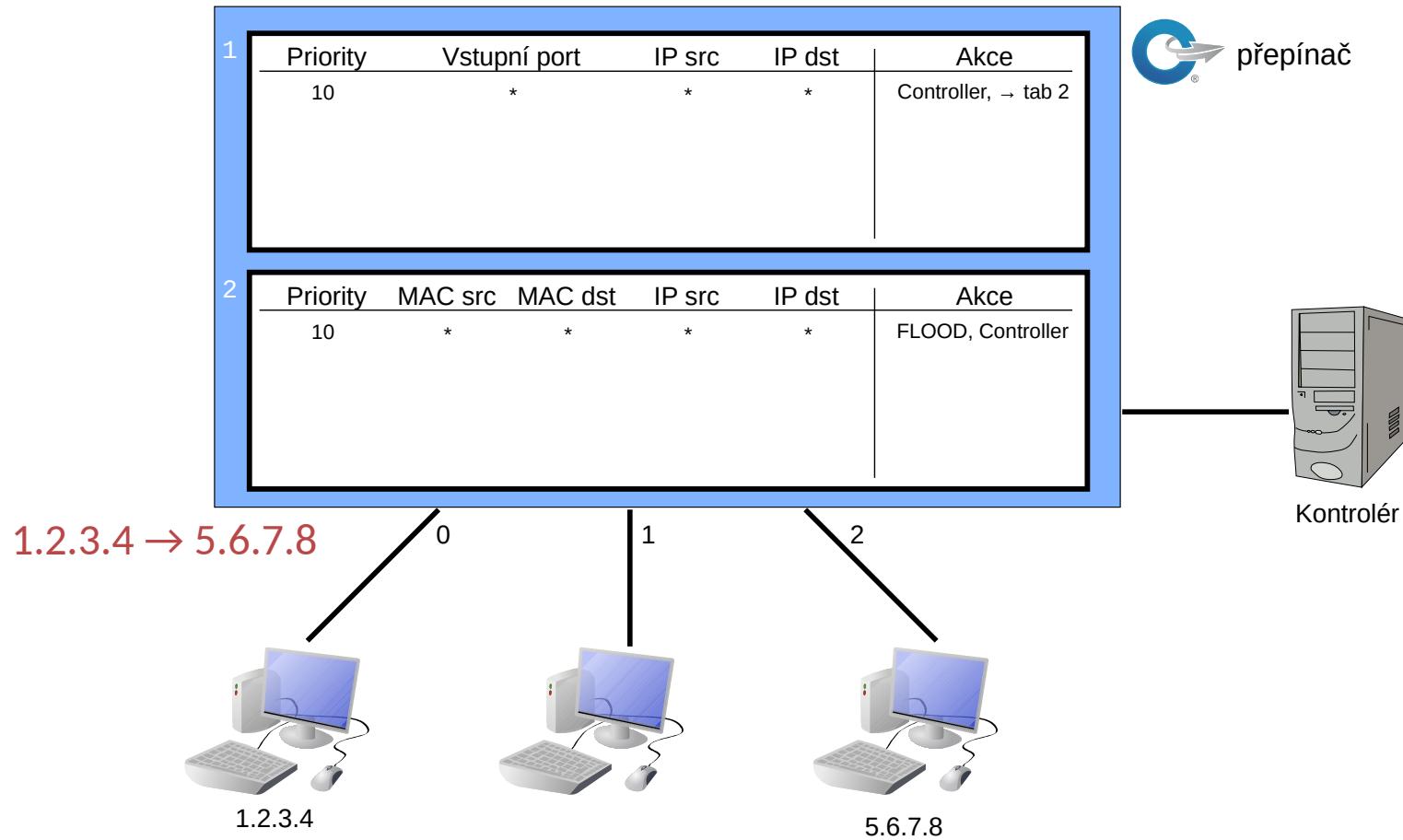


- 1) Nalezení pravidla s nejvyšší prioritou
- 2) Rozšiř množinu akcí/metadat, modifikuj paket
- 3) Volitelně aplikuj akce
- 4) Volitelně skoč do další tabulky

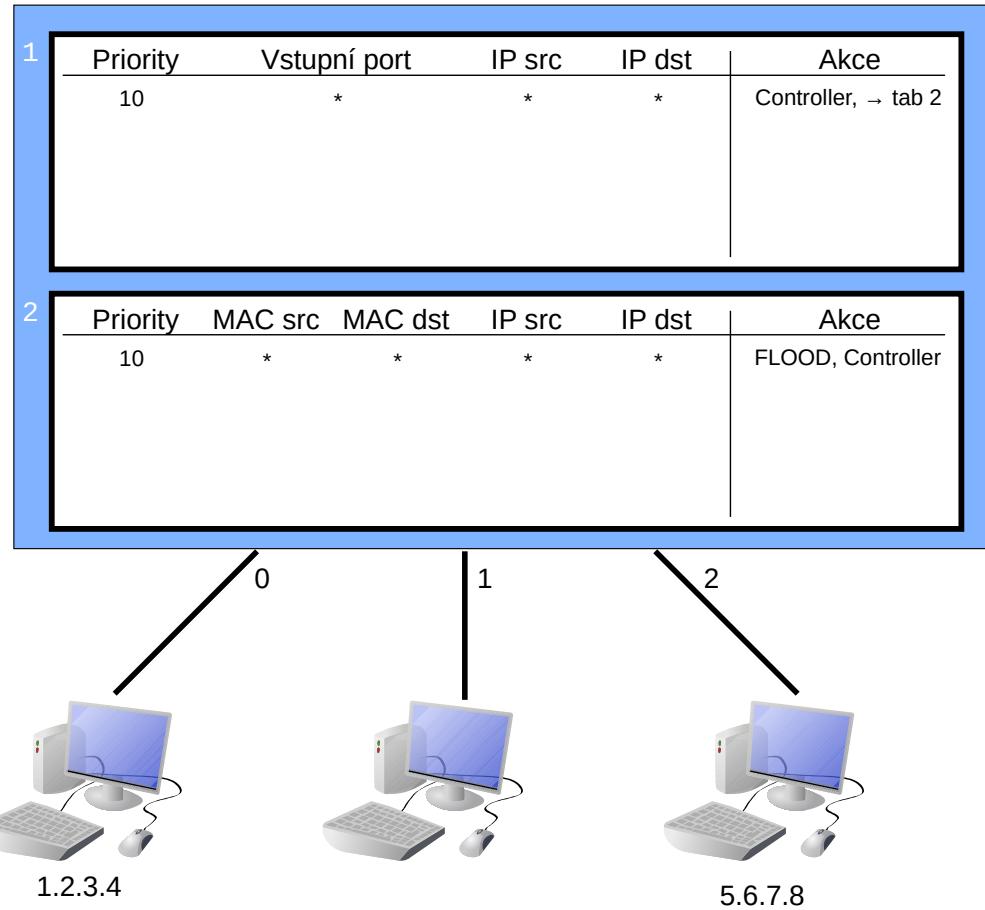
Příklad



Příklad

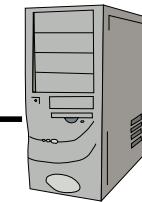


Příklad



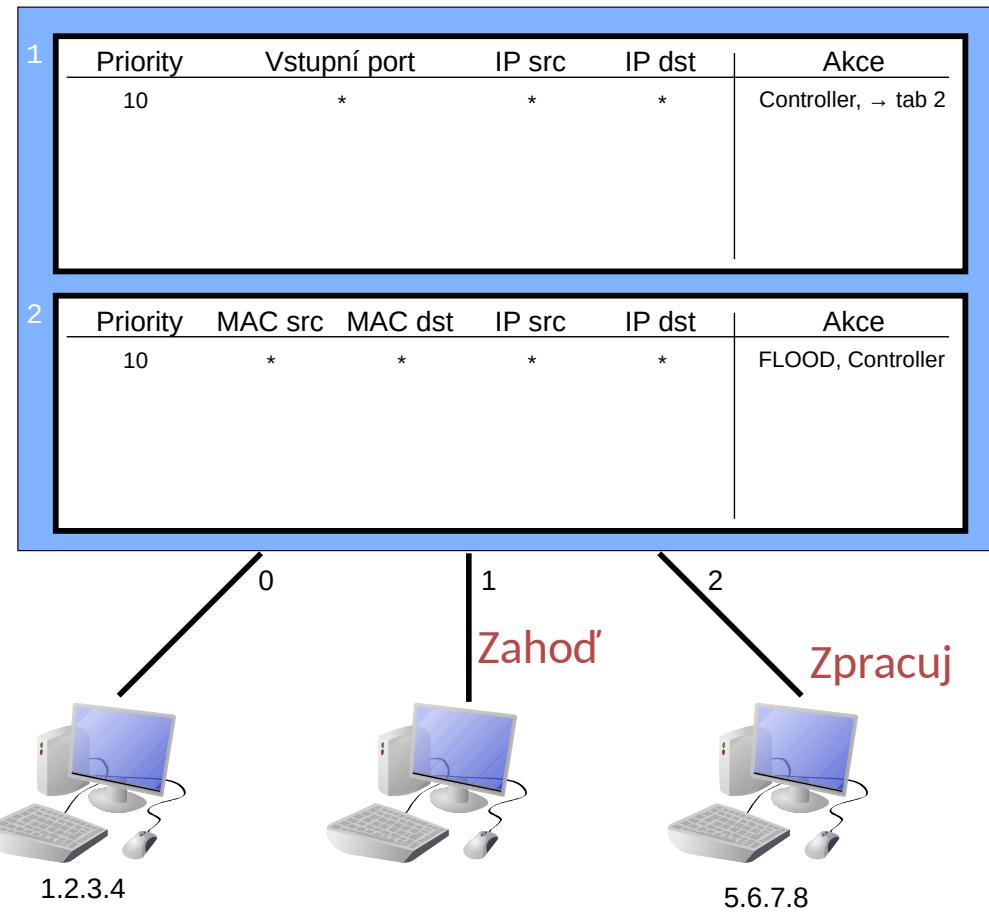
přepínač

Pošli na port 1, 2 a kontrolér



Kontrolér

Příklad



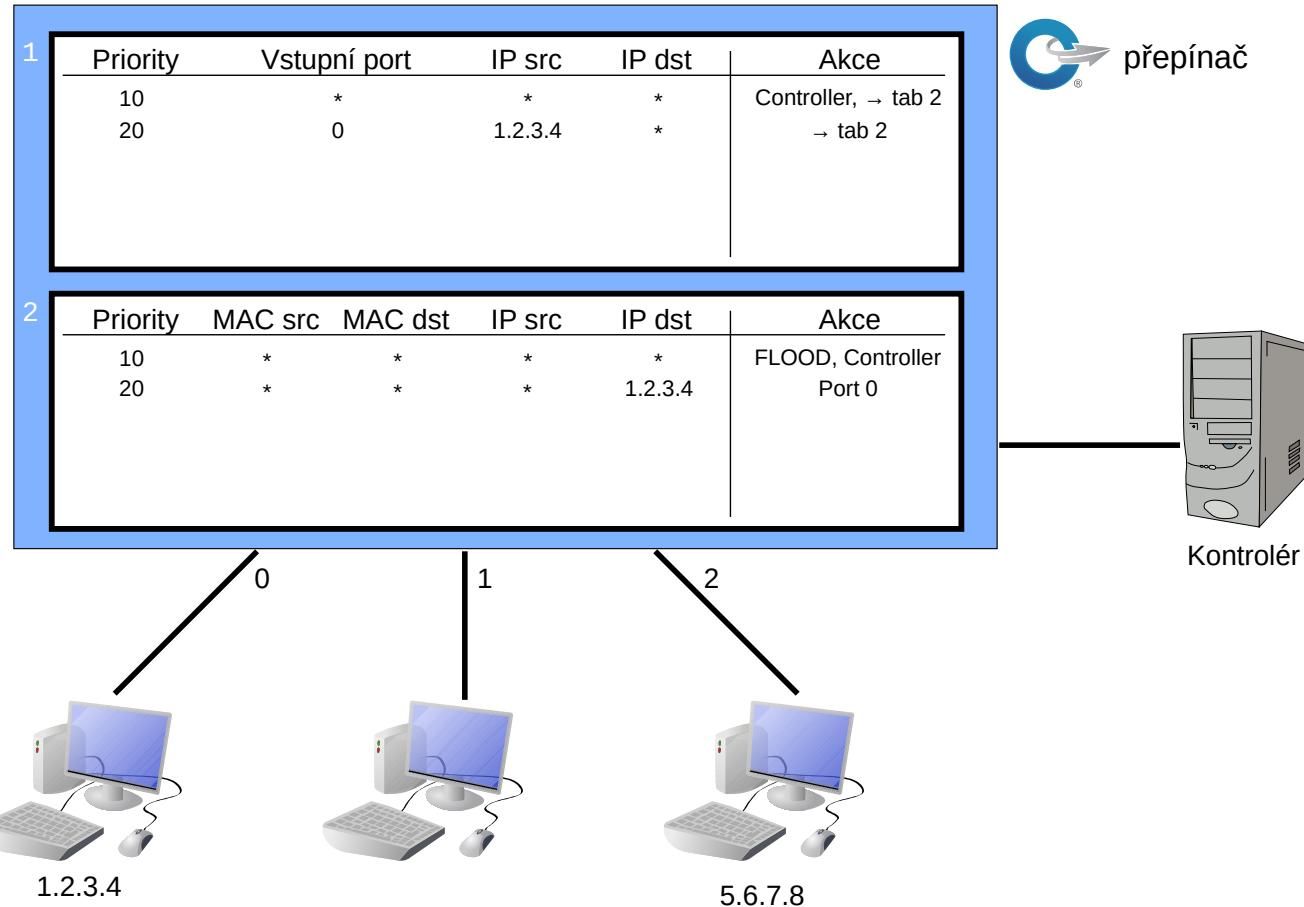
přepínač



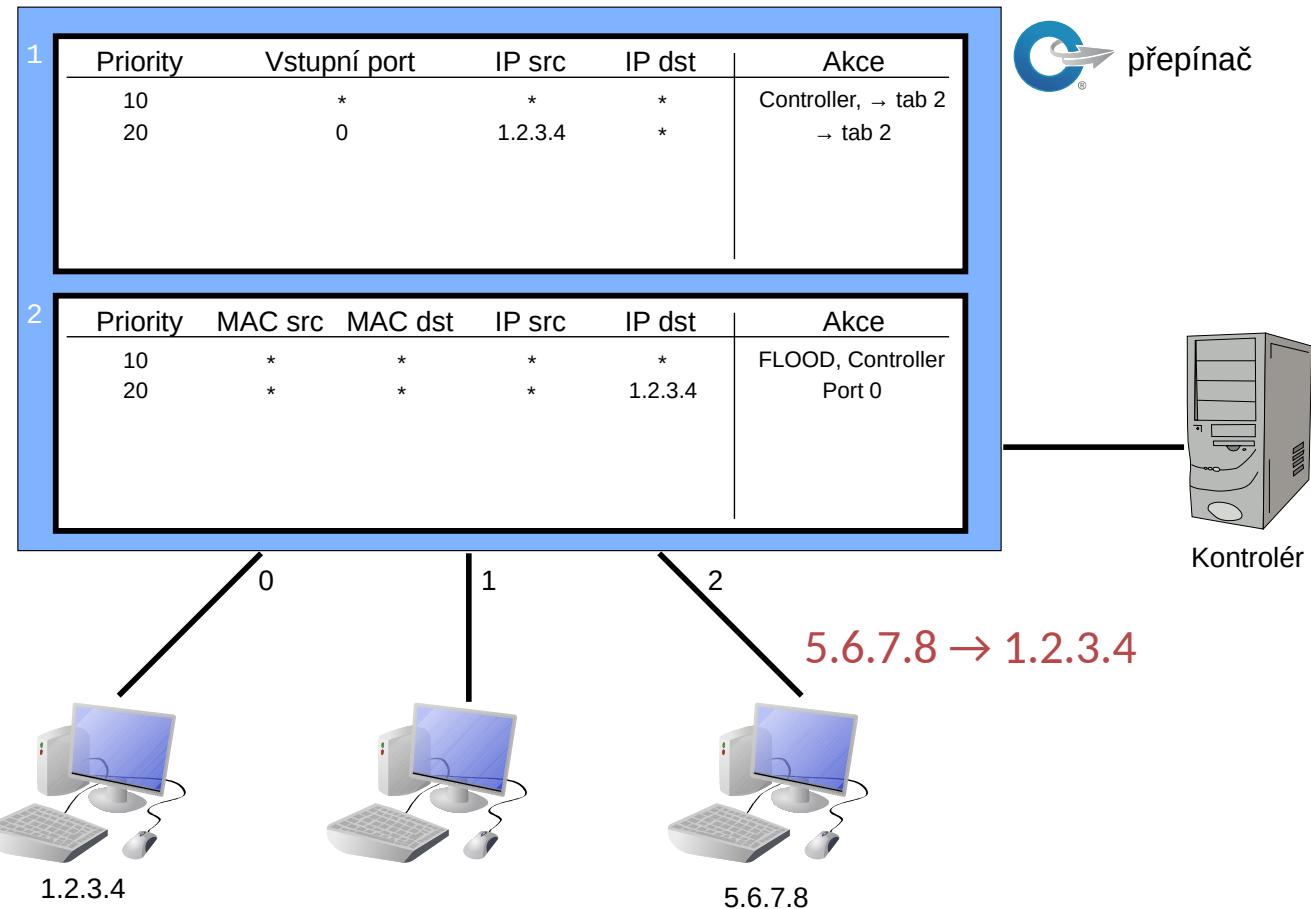
Kontrolér

Vlož nová pravidla

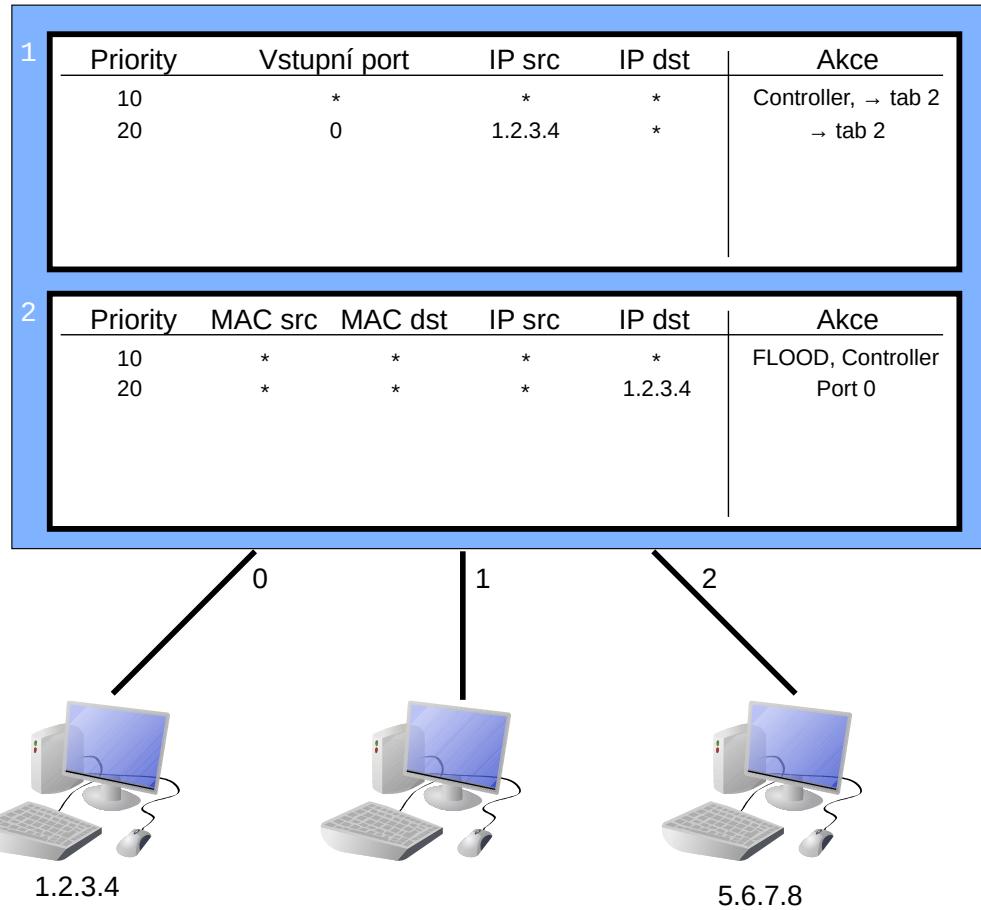
Příklad



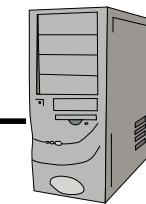
Příklad



Příklad

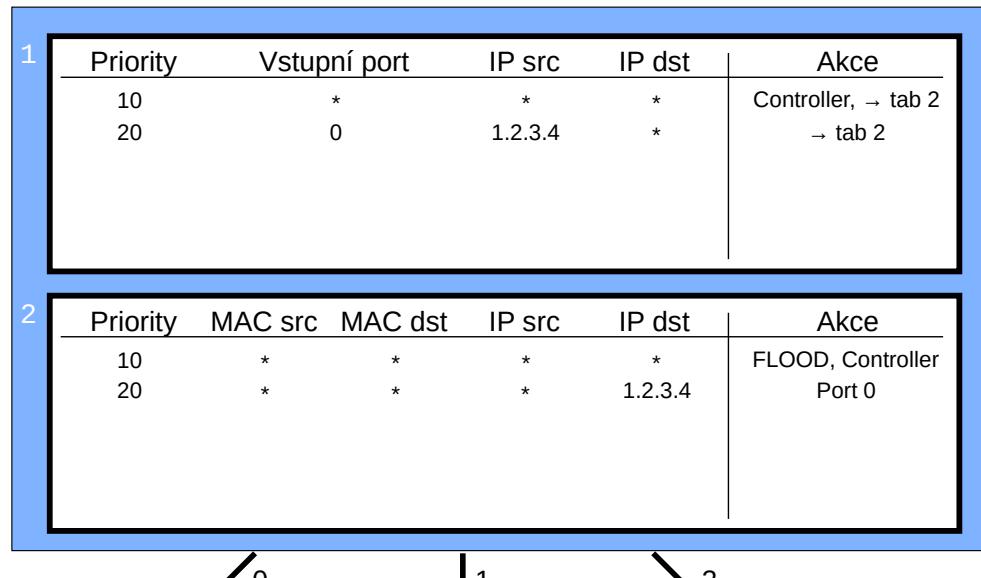


Pošli na port 0 a kontrolér

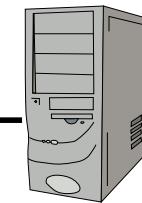


Kontrolér

Příklad



→ přepínač



Kontrolér

Kontrolce Vlož nová pravidla

Zpracuj

A photograph of a computer workstation. It includes a white desktop tower with two red indicator lights, a flat-panel monitor displaying a blue screen, a black keyboard, and a black mouse on a matching mouse pad.

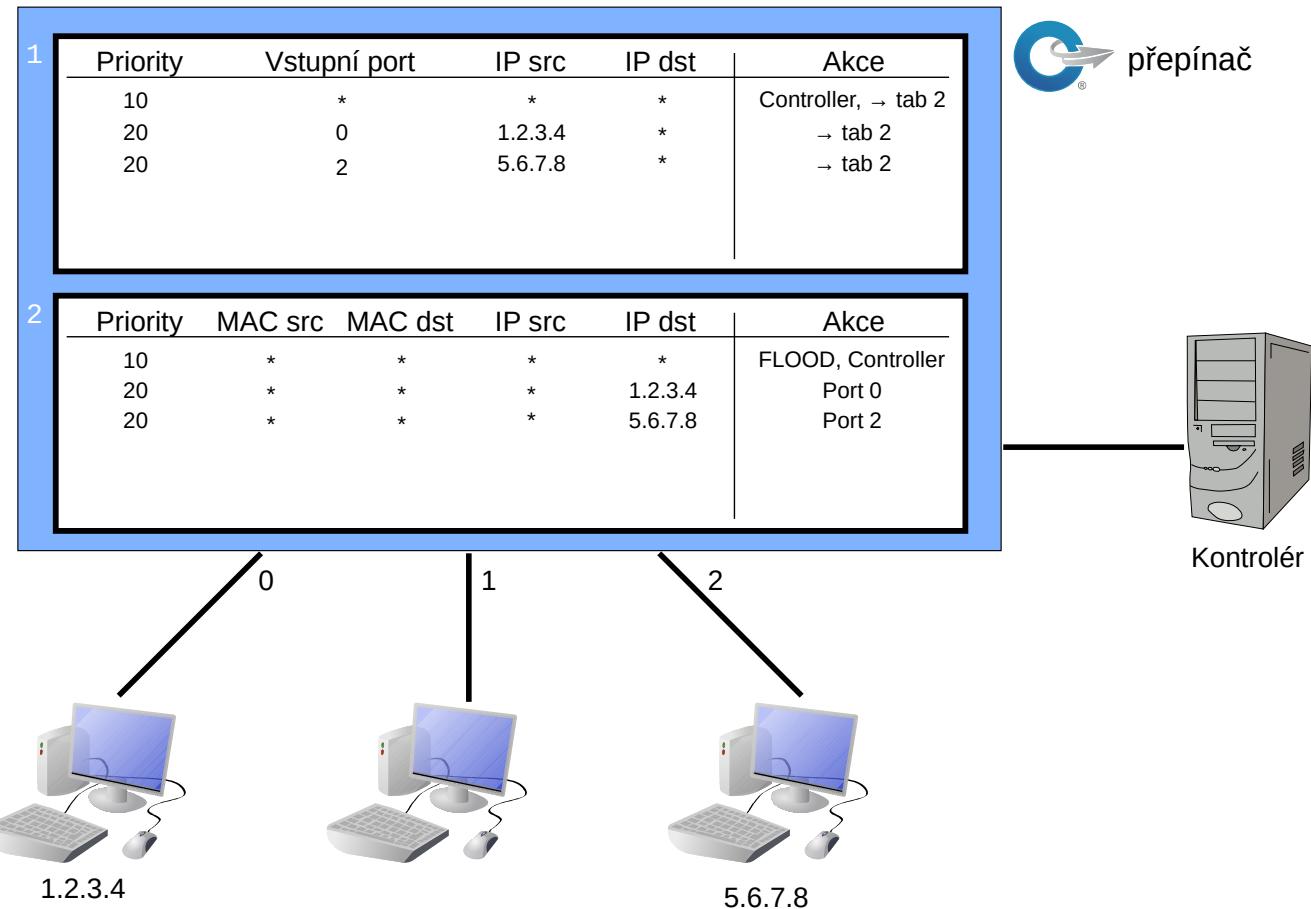
1.2.3.4

A white desktop computer setup featuring a monitor displaying a blue screen with a grid pattern, a white keyboard, and a white mouse connected by black cables.

A 3D rendering of a desktop computer system. It includes a white tower case with a blue light bar, a flat-panel monitor displaying a blue screen with a white curve, a black keyboard, and a black mouse on a matching mouse pad.

5.6.7.8

Příklad



Komunikace mezi 1.2.3.4 a 5.6.7.8 bez účasti kontroléru

OpenFlow 1.1

- Více tabulek
- Skupiny portů
 - ◆ Multicast, Multipath
- Podpora MPLS a VLAN
- Virtuální porty
 - ◆ Tunely
- Změna chování po přerušení spojení
 - ◆ Zrušena speciální tabulka
 - ◆ OF/Fallback na standardní přepínač

OpenFlow 1.2

- Rozšířené možnosti polí pravidel (TLV)
- Rozšířené možnosti přepisu polí
- Rozšíření metadat
- Rozšířené chybové zprávy
- Podpora IPv6
- Odstraněna specifikace parsování paketů

OpenFlow 1.3

- Změna specifikace inzerování možností (capabilities)
- Flexibilnější kontrola při nenalezení pravidla
- Podpora pro rozšiřující hlavičky IPv6
- Vlastní čítače pro každý tok

OpenFlow 1.4

- Rozšíření protokolu (TLV)
- Vylepšení podpory metadat pro vstupní pakety
- Podpora pro optické porty
- Změny týkající se podpory více kontrolérů
 - ◆ Mazání čítačů, změn provedených jiným kontrolérem
- Vylepšení ošetření přeplnění tabulek
- Kvaziatomické operace
- Změna použitého portu z 6633 a 976 na 6653

OpenFlow 1.5

- Tabulky pro výstupní porty
- Podpora ne-Ethernetových rámců (IP, PPP)
- Rozšířená práce se statistikami toků
- Akce *Copy-Field*
- Hledání pravidel podle příznaků TCP

OpenFlow - summarizace

- Nástroj umožňující inovaci
 - ◆ Samotné OpenFlow žádný problém neřeší
- Řízení toků na různých stupních granularity
 - ◆ Mikrotoky a *
- Proaktivní/reaktivní přístup
- Oddělení „control plane“ a „data plane“

Vrchní rozhraní (Northbound)

Aplikace nad SDN

- Vrchní rozhraní (NB) není standardizováno
 - ◆ Každý kontrolér implementuje po svém
 - ◆ Aplikace nepřenosné mezi kontroléry
- Typicky
 - ◆ Plug-in (Ryu, POX, POX+Pyretic, ODL+OSGi)
 - ◆ REST (ODL)
 - ◆ Jiný vhodný protokol

Distribuce aplikací

- HP SDN App Store
 - ◆ <https://www.hpe.com/us/en/networking/applications.html>
 - ◆ HP SDN Controller, OpenDaylight
- Komunitní portály (funkční?)
 - ◆ <http://sdnhub.org/>
- Součástí kontrolerů

Využití SDN

Využití SDN

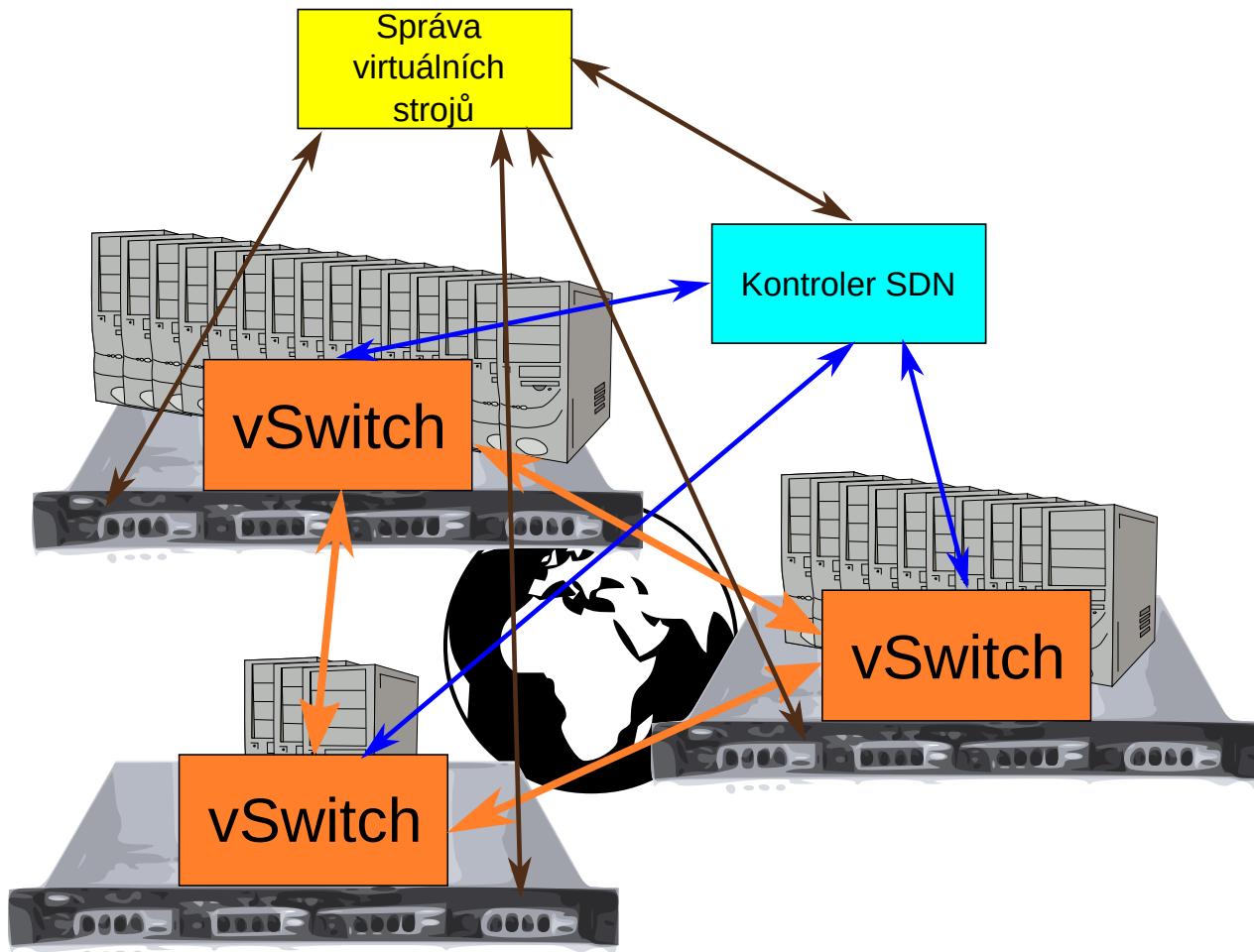
- Alternativní přístup ke správě sítě
- SDN je obecný nástroj
 - ◆ Samo o sobě neřeší žádný specifický problém
 - ◆ Programovatelnost aktivních prvků
 - Snadnost inovací
- Virtualizace sítí
 - ◆ Logické oddělení provozu přestože sdílí stejnou fyzickou infrastrukturu
 - ◆ Symbiotický vztah
- Network Function Virtualization
 - ◆ SDN jedna z možností řešení

Současný stav využití SDN

- Nasazení primárně v data centrech
- Zkušenosti z HP
 - ◆ Hybridní sítě, hybridní přepínače
 - ◆ Omezené možnosti tabulek (např. Počet pravidel s žolíkem)
- Nejednotnost výrobců
 - ◆ Např. Hybridní režim u zařízení Cisco (ships-in-the-night)
- Podpora jen u specifických zařízení

SDN v data centrech

- Jedno z uplatnění SDN, momentálně nejviditelnější



Network function virtualization

- Využití SDN switchů pro funkcionalitu middleboxů
 - ◆ Firewall
 - ◆ Load balancer
 - ◆ NAT
 - ◆ ...

Programovatelná „data plane“

- Programming Protocol-independent Packet Processors (P4)
 - ◆ programování protokolově nezávislých síťových procesorů
 - ◆ Platformě nezávislý popis → CPU, GPU, síťový procesor, FPGA

```
header_type ethernet_t {  
    fields {  
        dstAddr : 48;  
        srcAddr : 48;  
        etherType : 16;  
    }  
}  
  
header ethernet eth;  
parser ethernet {  
    extract(eth);  
    switch(eth.ethertype) {  
        case 0x8100: vlan;  
        case 0x9100: vlan;  
        case 0x800: ipv4;  
        case 0x86dd: ipv6;  
    }  
}
```

<http://p4.org/>

Související projekty

- Open Network Install Environment (ONIE)
 - ◆ <http://onie.opencompute.org>
 - ◆ Vlastní OS pro aktivní prvky
- Open Network Linux
 - ◆ <http://opennetlinux.org/>
 - ◆ Linuxová distribuce pro ONIE
- Open Network Operating System
 - ◆ <http://onosproject.org/>
 - ◆ Distribuovaný OS pro service providery
- OpenStack
 - ◆ <https://www.openstack.org/>
 - ◆ Software pro vytváření a správu cloudu

Shrnutí

Závěr

- Problémy, které se SDN snaží řešit:
 - ◆ Pevně daná funkcionalita sítí
 - ◆ Uzavřenost inovacím
- Architektura
 - ◆ Řídící/datová vrstva („control plane“, „data plane“)
 - ◆ Kontrolér a jeho rozhraní
 - OpenFlow
 - Netconf, Proprietární protokoly
- Problémy
 - ◆ Zpoždění při vkládání nových toků (reaktivní režim)
 - ◆ Velikost tabulek
 - ◆ Stabilita

Reference

- Open Flow
 - ◆ <https://www.opennetworking.org/sdn-resources/onf-specifications/openflow>
- Nástroje
 - ◆ <http://mininet.org/walkthrough/>
 - ◆ <http://www.opendaylight.org/>
- Literatura v knihovně FIT
 - ◆ Göransson Paul, Black Chuck: Software Defined Networks: A Comprehensive Approach, Morgan Kaufmann