

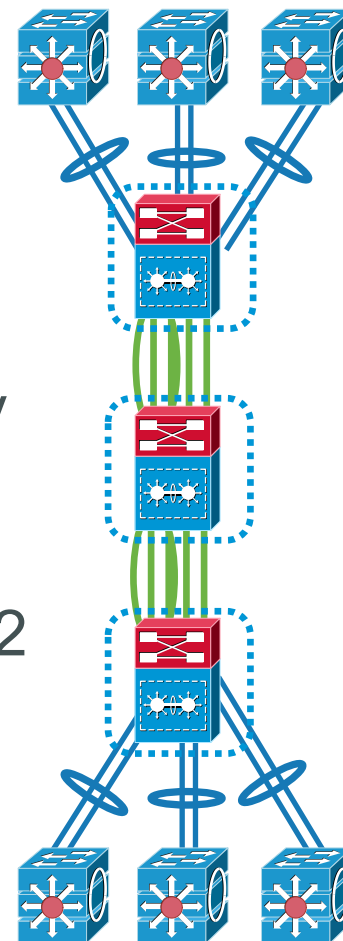
Návrh vysoce dostupných campus sítí

ARCH1/L2

Pavel Křižanovský, CCIE #11457

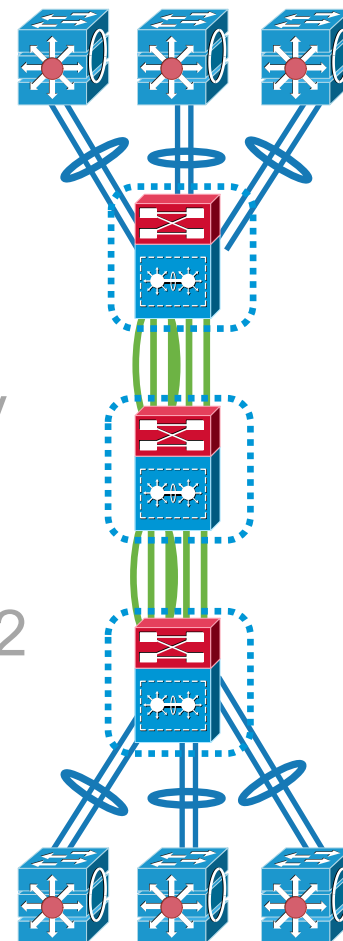
Program

- Proč vysoce dostupné sítě ?
- Tradiční návrh campus sítí
- Alternativy řešení přístupové a distribuční vrstvy
- Aspekty redundantního návrhu
- Case Study – Campus síť pro LOH Londýn 2012
- Očekávané trendy v campus sítích
- Shrnutí



Program

- **Proč vysoce dostupné sítě ?**
- Tradiční návrh campus sítí
- Alternativy řešení přístupové a distribuční vrstvy
- Aspekty redundantního návrhu
- Case Study – Campus síť pro LOH Londýn 2012
- Očekávané trendy v campus sítích
- Shrnutí



Jaké jsou nároky na dostupnost sítě?

- Architektura podnikových sítí se vyvíjí s ohledem na tyto požadavky:
 - Požadavky aplikací a služeb
Podnikové IS, prostředky pro spolupráci, video...
 - Standardy odvětví/podniku/organizace:
Finanční
Zdravotnické
7x24x365 globální přístup
...
 - Očekávání uživatelů: Být stále připojen

Vyžaduje strukturovaný a spolehlivý návrh infrastruktury

Globální přístup

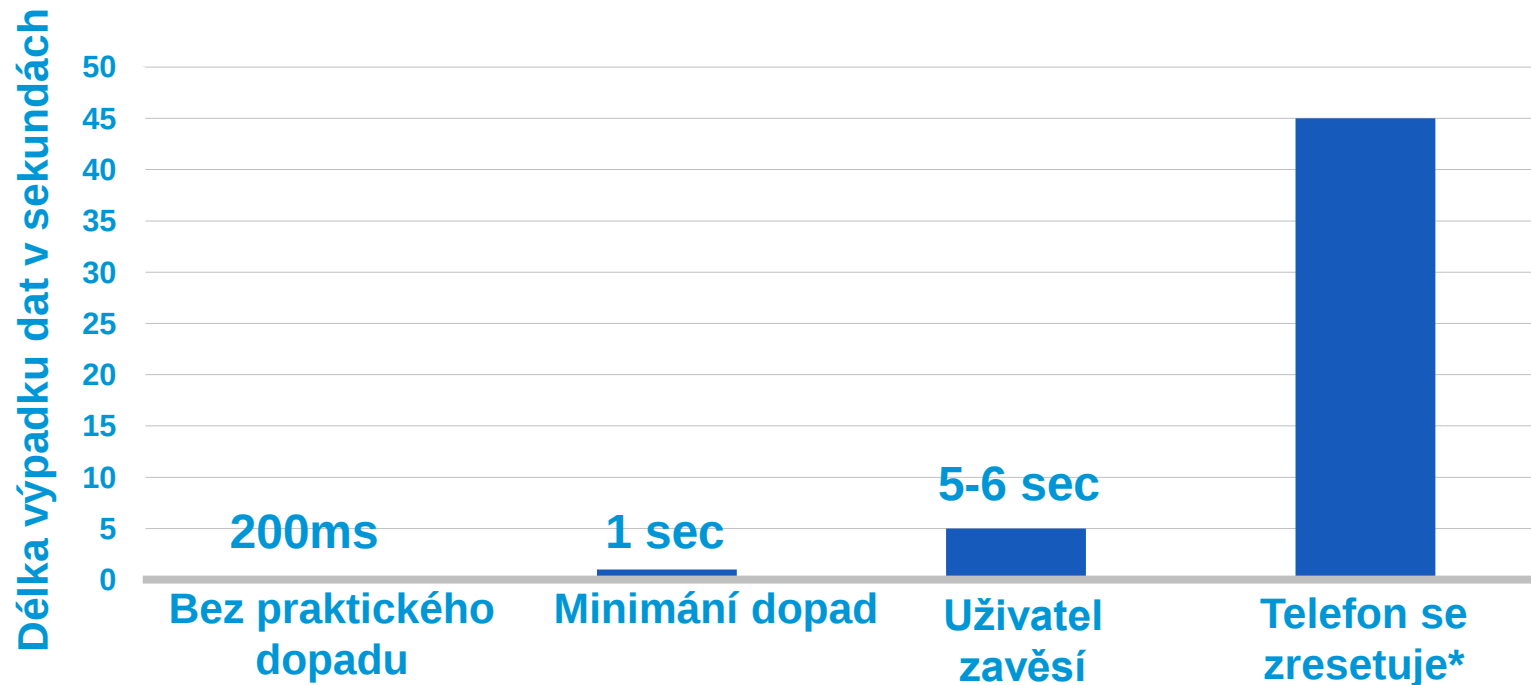


Spolupráce a komunikace v reálném čase



Jak výpadek ovlivňuje hlasovou komunikaci ?

- Požadavky na dostupnost UC: Více než pět devítek
- Nutno uvažovat dopad na subjektivní pocit z hlasové komunikace



* Doba, po které se IP telefon zresetuje je proměnná a závisí na signalizačním protokolu (SCCP, SIP) a na aktuálním stavu hovoru (aktivní, vyzvánění apod.).

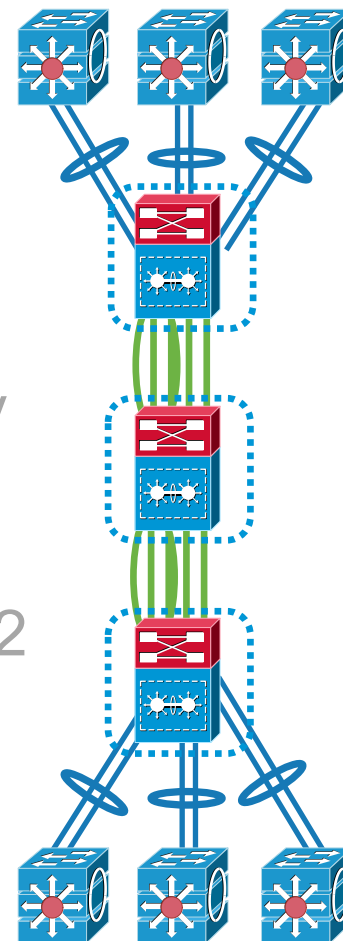
Jak výpadek ovlivňuje video komunikaci?

- Požadovaná síťová SLA se pro tradiční videokonference a HD teleprezenční systémy velmi liší

Metric	TelePresence			Traditional Video Conferencing
	Target	Threshold 1 (Warning)	Threshold 2 (Call Drop)	
Latency	150 ms	200 ms	400 ms	400-450 ms
Jitter	10 ms	20 ms	40 ms	30-50 ms
Loss	0.05%	0.10%	0.20%	1%
BW	2.5 - 12.6 Mbps + overhead			384 or 768 kbps + overhead

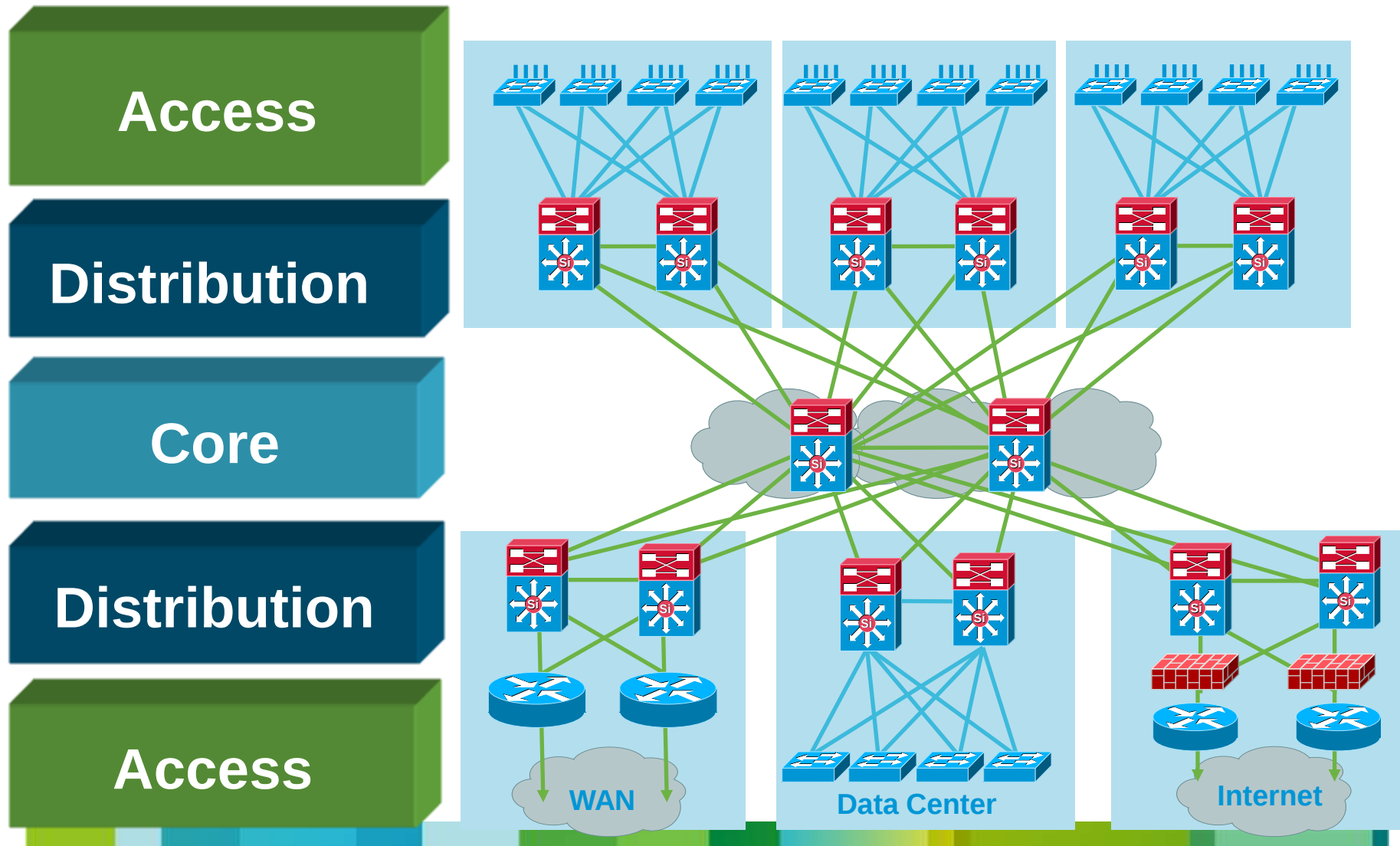
Program

- Proč vysoce dostupné sítě ?
- **Tradiční návrh campus sítí**
- Alternativy řešení přístupové a distribuční vrstvy
- Aspekty redundantního návrhu
- Case Study – Campus síť pro LOH Londýn 2012
- Očekávané trendy v campus sítích
- Shrnutí



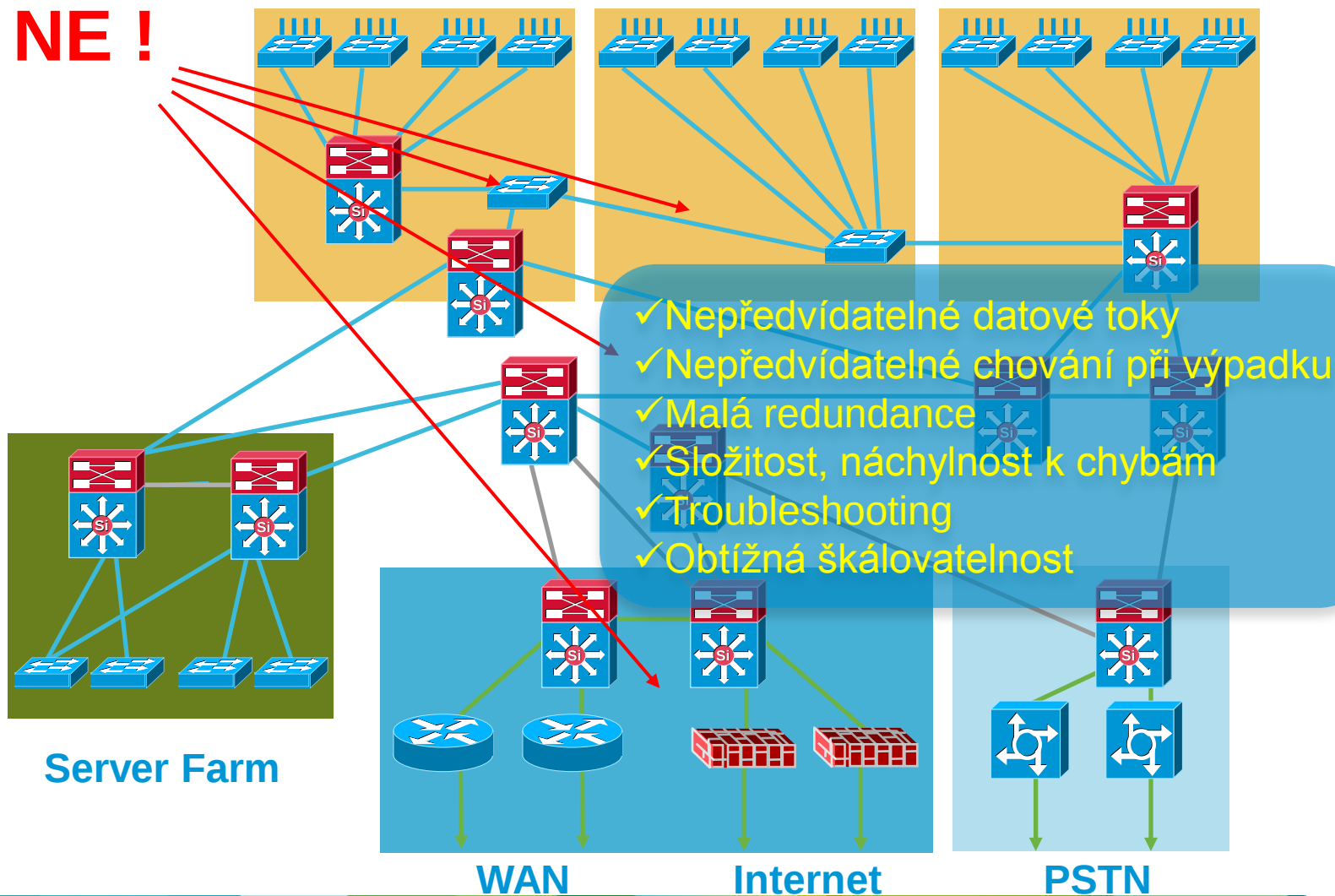
Návrh vysoce dostupné Campus sítě:

Strukturovanost, Modularita, Hierarchie

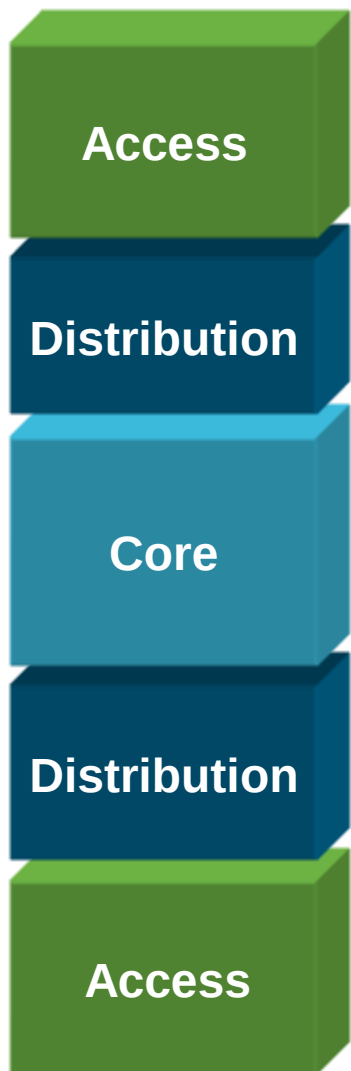


Návrh vysoce dostupné Campus sítě

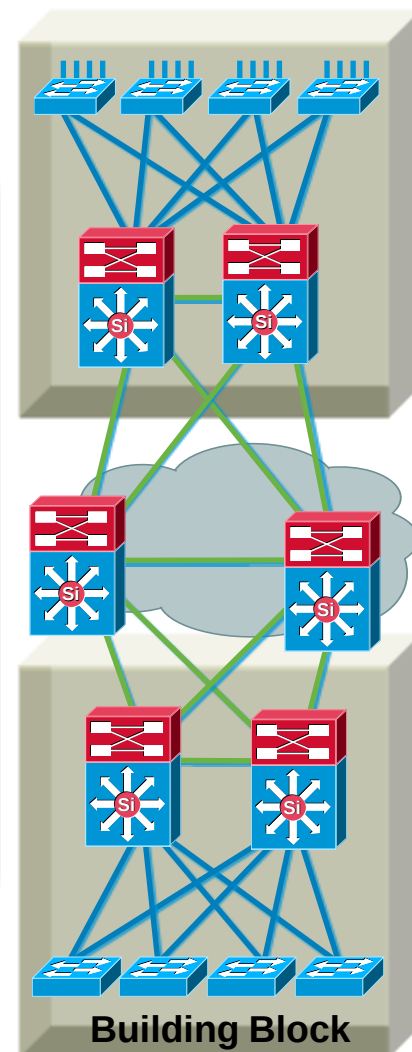
Takto NE !



Hierarchický síťový návrh je základ !



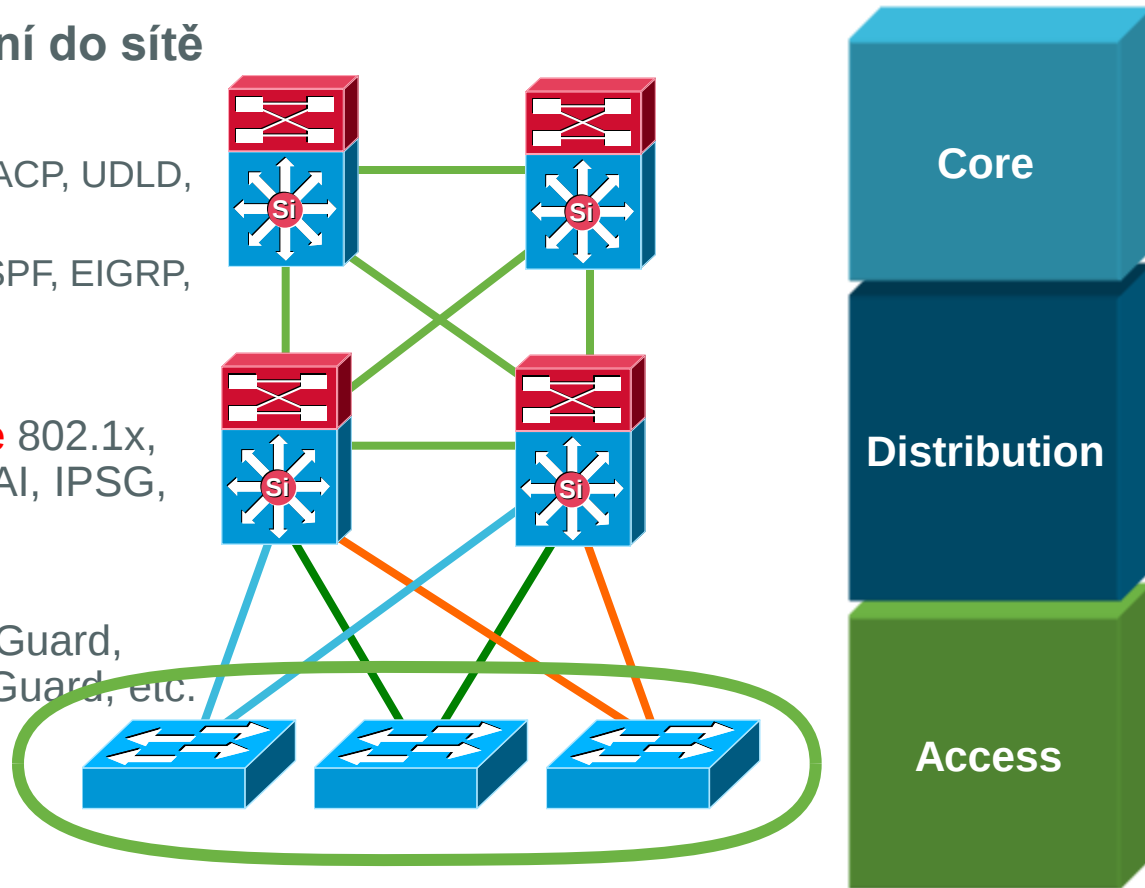
- Hierarchie — každá vrstva má specifickou úlohu
- Modulární topologie —stavební bloky
- Jednoduché škálování, snadná pochopitelnost, spravovatelnost a troubleshooting
- Vytváření malých fault domén— jasné ohraničení a izolace potenciálních problémů
- Load balancing a redundancy
- Deterministické datové toky



Access Layer - přístupová vrstva

Pokročilé vlastnosti, nikoliv pouze základní konektivita

- Připojení koncových zařízení do sítě
- **Inteligentní síťové služby:**
 - L2: PVST+, Rapid PVST+, PAgP/LACP, UDLD, FlexLink
 - L3 DTP, broadcast suppression, OSPF, EIGRP, IP multicast, IGMP snooping
 - QoS, trust boundary
- **Integrované bezpečnostní funkce** 802.1x, port security, DHCP snooping, DAI, IPSG, atd.
- Spanning tree toolkit: PortFast, UplinkFast, BackboneFast, LoopGuard, BPDU Guard, BPDU Filter, RootGuard, etc.
- Automatic phone discovery, Power over Ethernet, etc.
- **EnergyWise, MediaNet, ...**

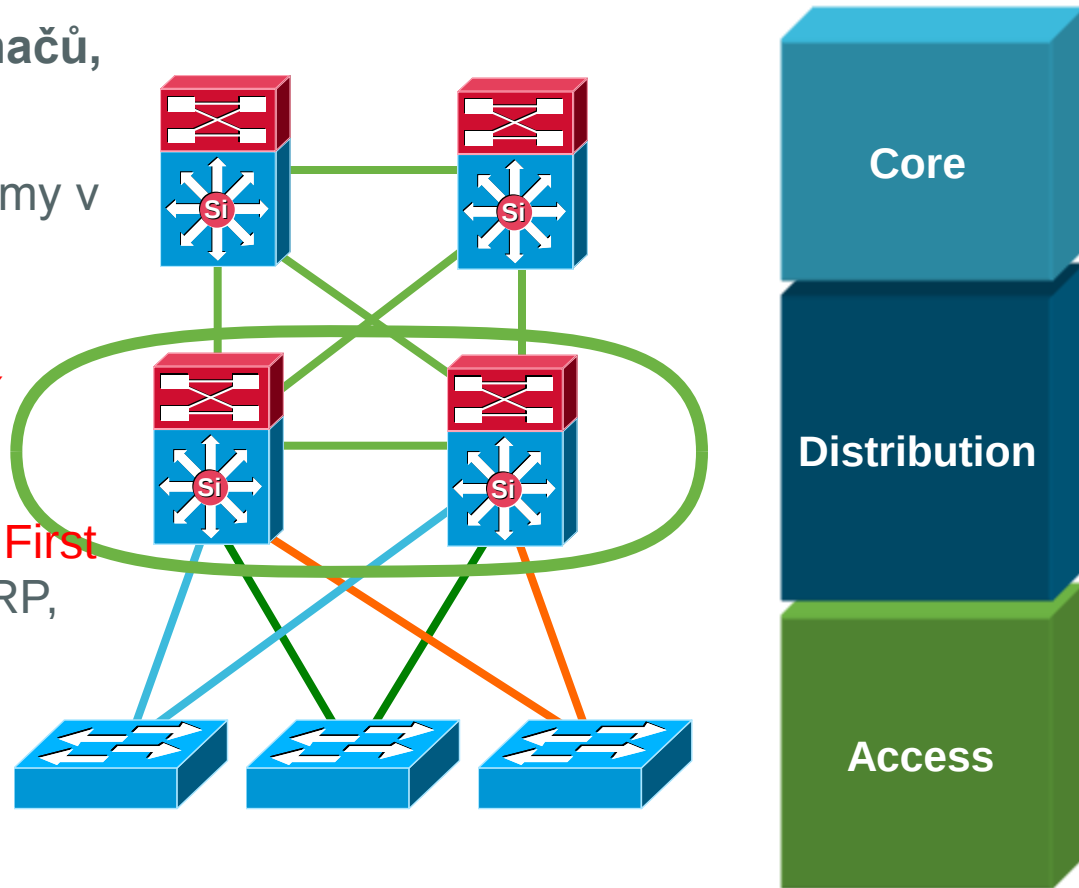


Viz přednáška “Pokročilé služby moderních LAN sítí”, NET3 ,15:00

Distribution Layer – distribuční vrstva

Konvergence, QoS, bezpečnost, vysoká dostupnost

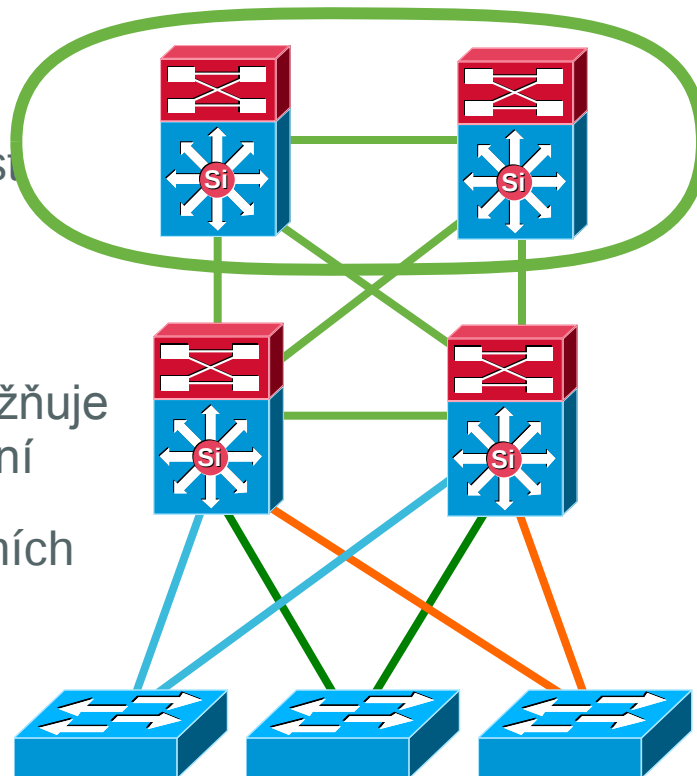
- Agregace přístupových přepínačů, propojení do core vrstvy
- Ochrana core vrstvy před problémy v přístupové vrstvě
- Sumarizace Route, rychlá konvergence, **rozdělování zátěže** mezi redundantní cesty
- Redundance na prvním skoku – **First Hop Redundancy protokoly**: HSRP, GLBP
- **Bezpečnostní** politika - ACL



Core Layer

Škálování, vysoká dostupnost a rychlá konvergence

- Páteř celé sítě — propojení distribučních bloků
- **Výkon, stabilita**, jednoduchost (méně je více)
- Dedikovaná core vrstva umožňuje jednoduché budoucí škálování
- Design nezávislý na konkrétních technologiích

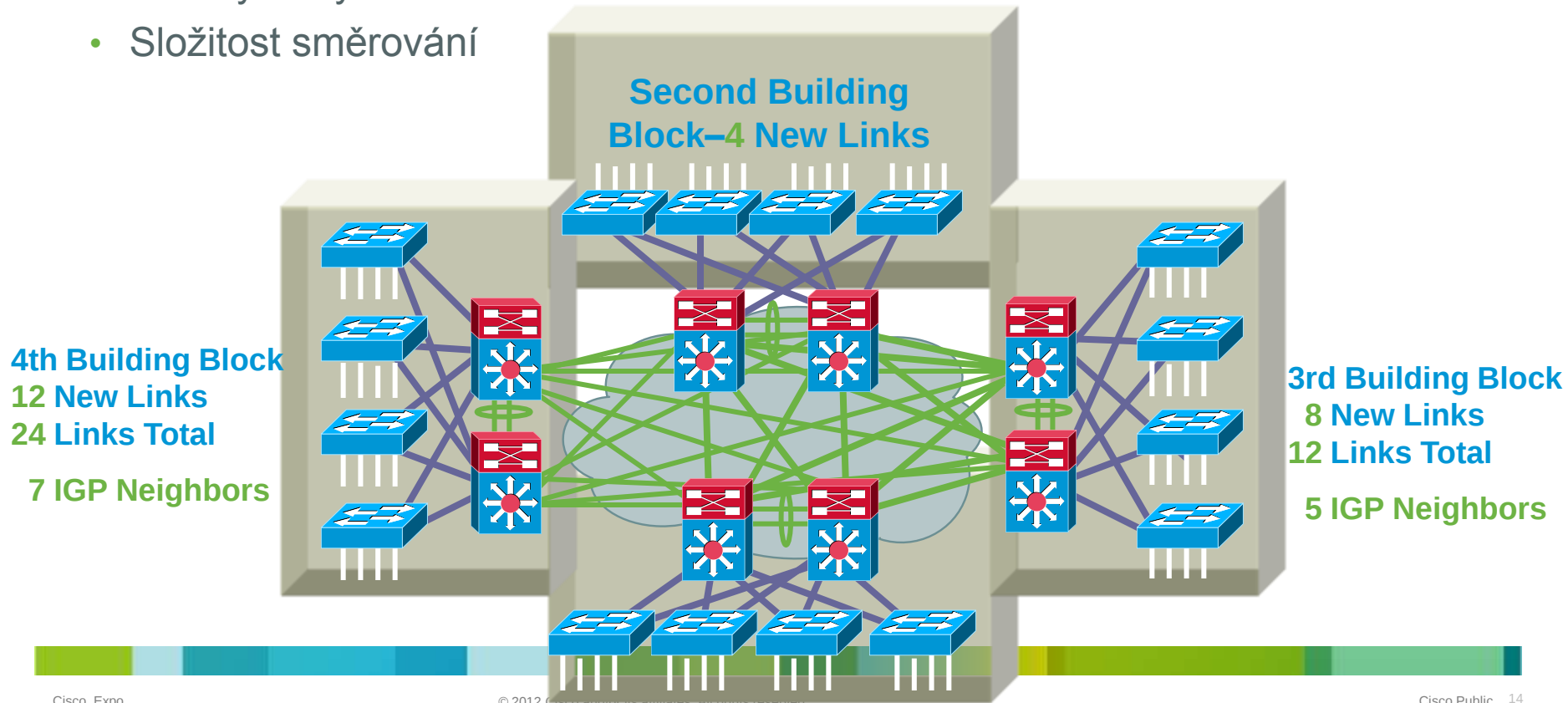


Potřebuji Core vrstvu ?

Otázka škálovatelnosti, složitosti a konvergence

Bez Core vrstvy

- Full-mesh distribučních bloků
- Nároky na fyzickou kabeláž
- Složitost směrování

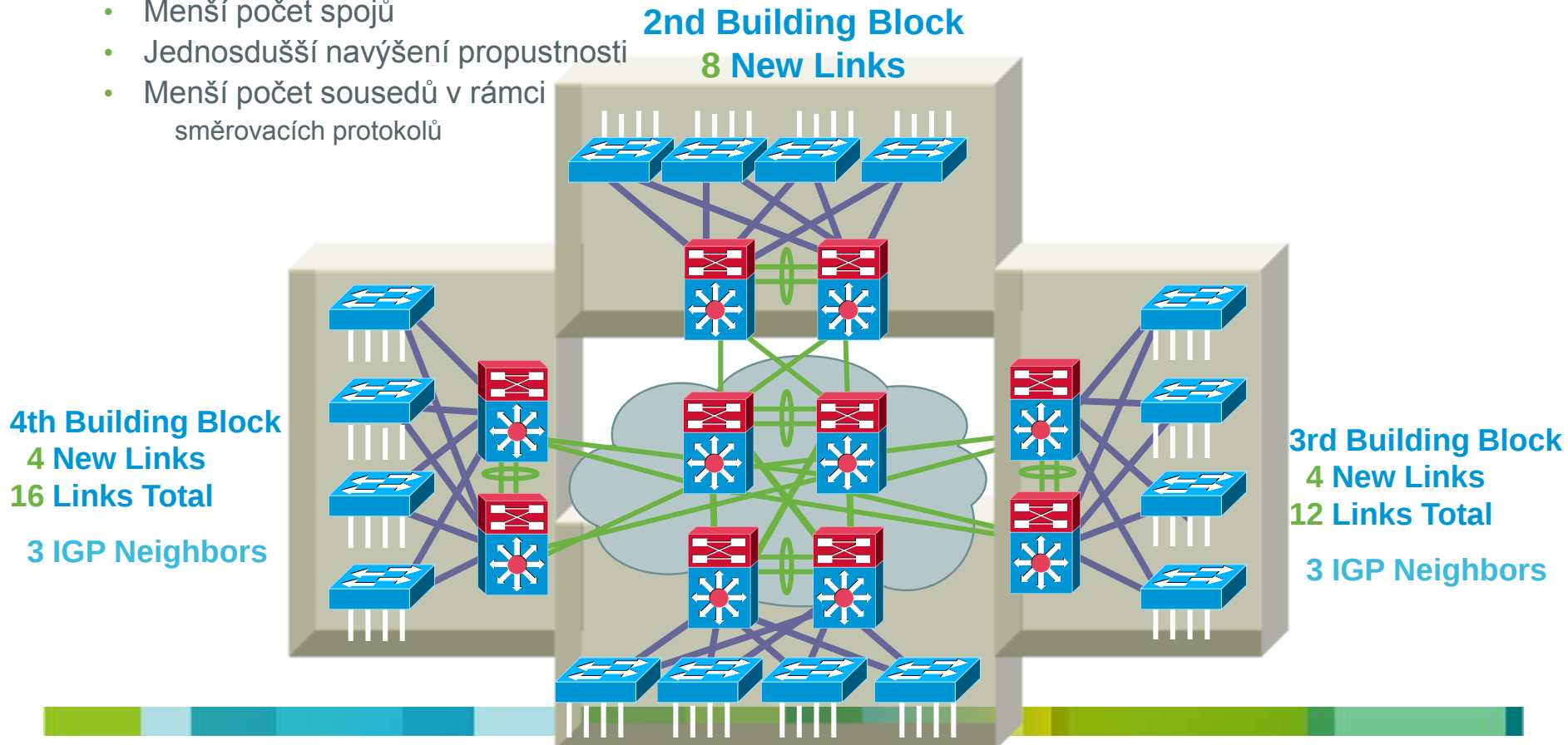


Potřebuji Core vrstvu ?

Otázka škálovatelnosti, složitosti a konvergence

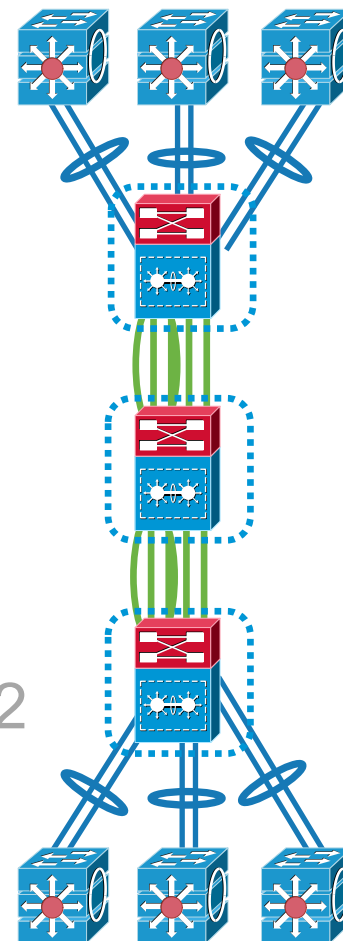
Dedikované Core přepínače

- Jednodušší přidání dalšího bloku
- Menší počet spojení
- Jednodušší navýšení propustnosti
- Menší počet sousedů v rámci směrovacích protokolů

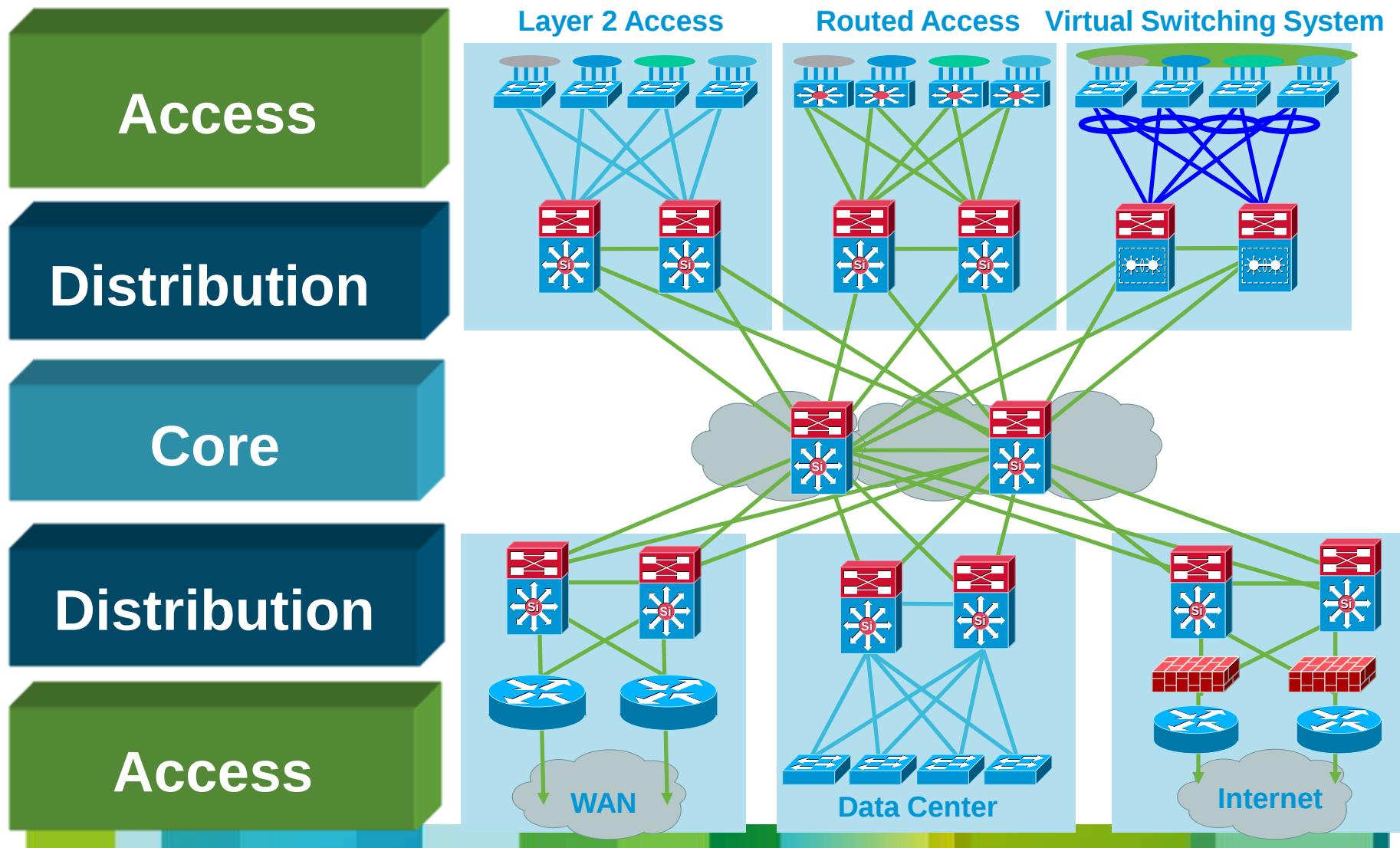


Program

- Proč vysoce dostupné sítě ?
- Tradiční návrh campus sítí
- **Alternativy řešení přístupové a distribuční vrstvy**
- Aspekty redundantního návrhu
- Case Study – Campus síť pro LOH Londýn 2012
- Očekávané trendy v campus sítích
- Shrnutí



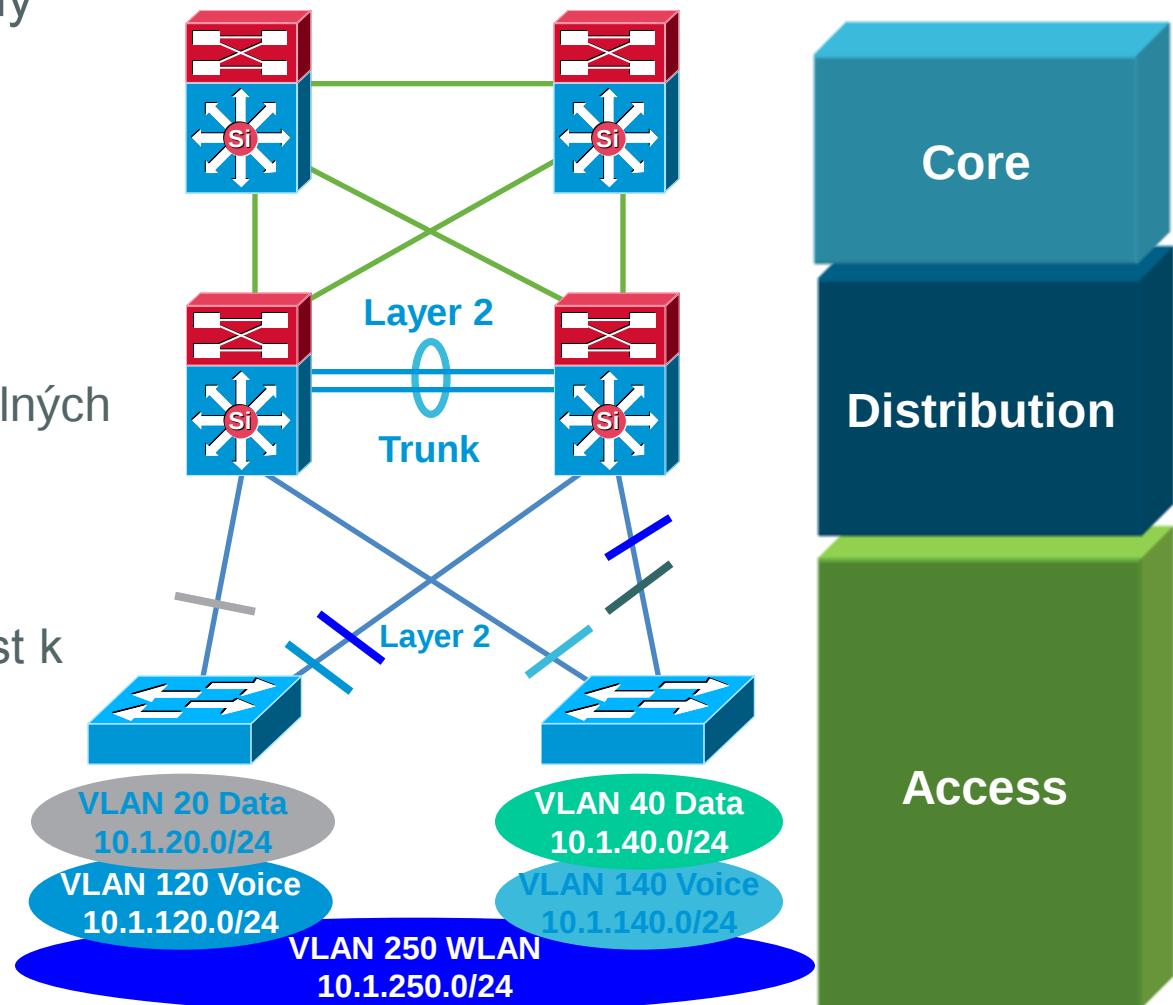
Různé alternativy řešení Access/Distribution



Layer 2 propojení v Distribution vrstvě

Layer 2 Access—Některé VLANy jsou rozprostřeny přes Access vrstvu

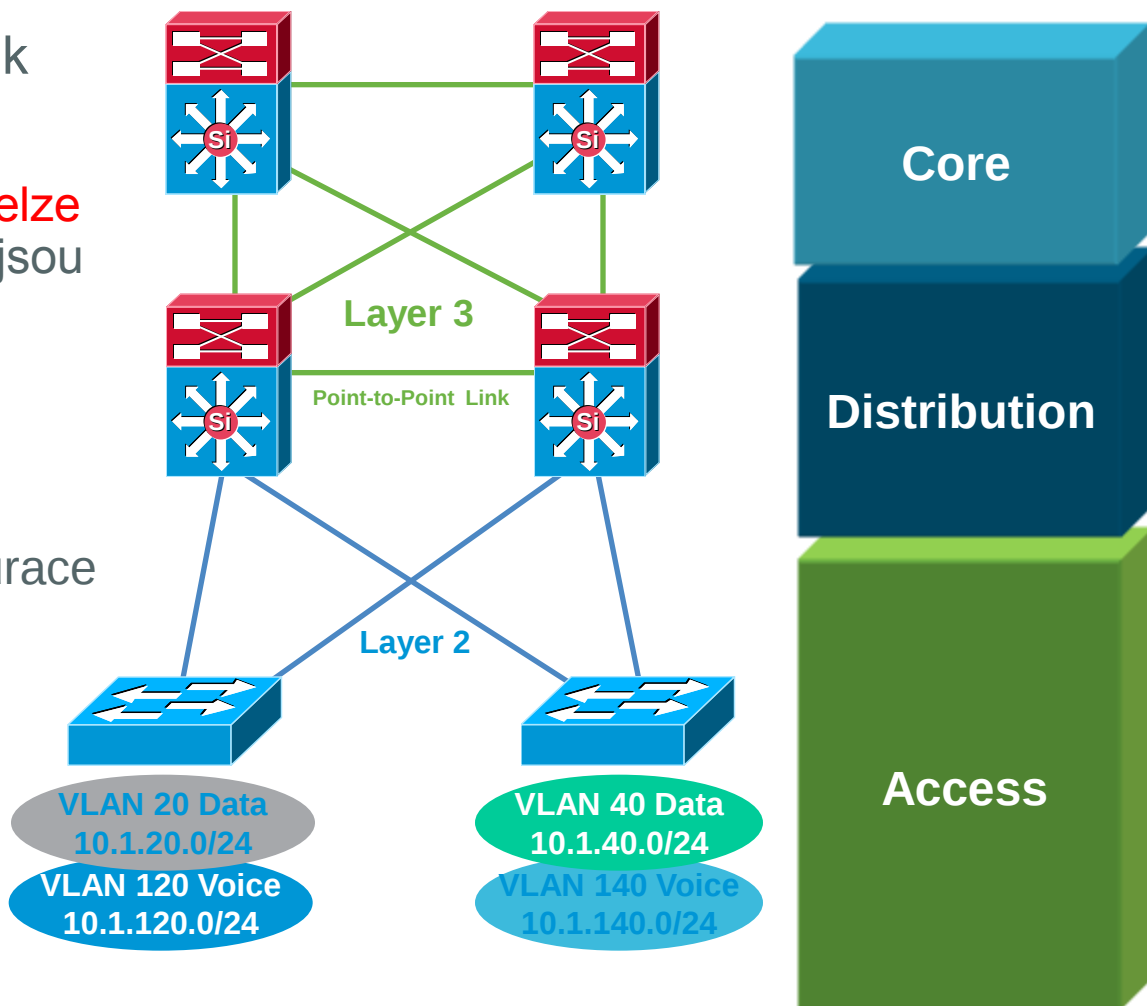
- Dlouholetou praxí prověřený design
- Typické problémy
 - Zákonnitý vznik L2 smyček
⇒ Nutnost nasazení STP
 - Konvergence** při výpadku
 - Blokování portů**, nevyužití plných rychlostí uplinku
 - Složitější multicast - dva směrovače v jedné VLAN
 - Celková složitost, náchylnost k chybám konfigurace



Layer 3 propojení v Distribution vrstvě

Layer 2 Access—VLANy zůstávají vždy ve svém Access přepínači

- Odstranění L2 smyček, odstranění blokových linek
- Typické problémy
 - Ideální design VLAN často nelze dodržet, rozprostřené VLAN jsou vyžadovány
 - ⇒ Zde nelze využít
 - Vyšší spotřeba IP adres
 - Stále ještě celková složitost, náchylnost k chybám konfigurace



Layer 3 v Access vrstvě

VLANy zůstávají vždy ve svém Access přepínači

- Odstranění L2 smyček, odstranění blokových linek
- Konvergence daná konverencí L3 IGP protokolu, jednodušší multicast
- Typické problémy

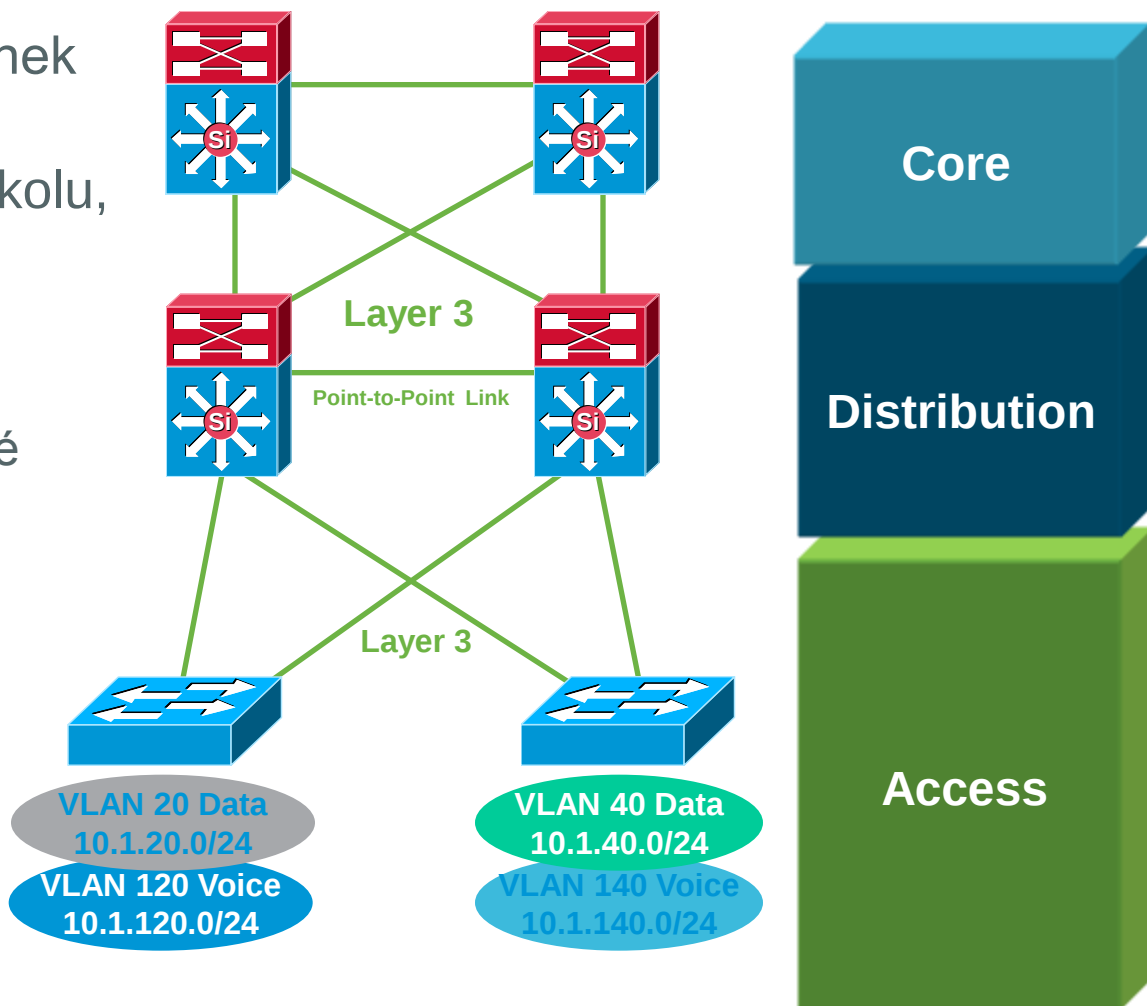
Ideální design VLAN často **nelze dodržet**, rozprostřené VLAN jsou vyžadovány

⇒ Zde nelze využít

Vyšší spotřeba IP adres

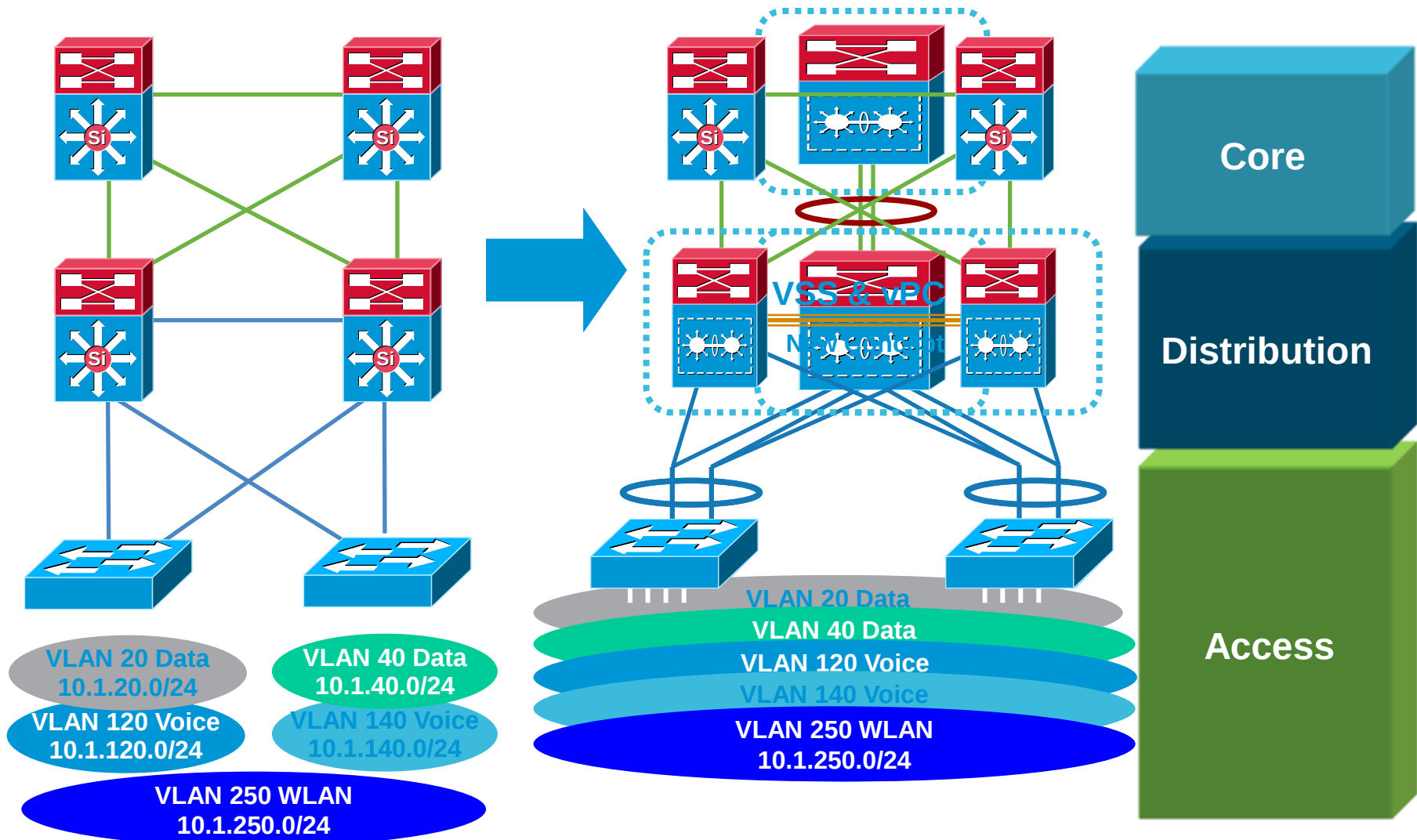
L3 přepínače v accessu – dražší HW, **licence**

Nelze použít RSPAN



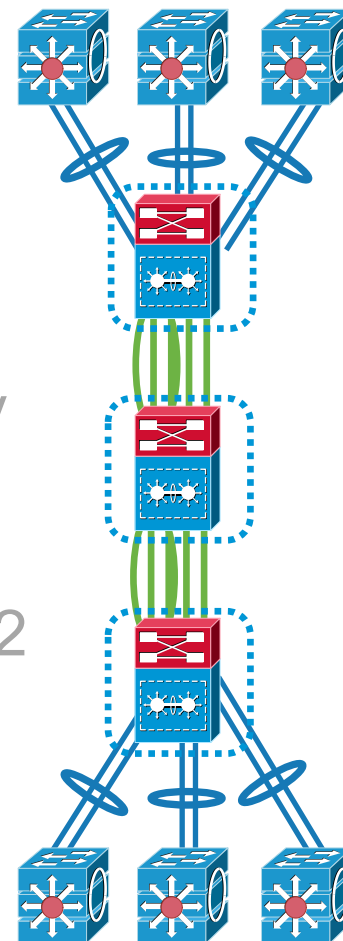
Jak využít všech výhod předchozích architektur ?

VSS (Virtual Switching System), VPC (Virtual Port Channel), apod...



Program

- Proč vysoce dostupné sítě ?
- Tradiční návrh campus sítí
- Alternativy řešení přístupové a distribuční vrstvy
- **Aspekty redundantního návrhu**
- Case Study – Campus síť pro LOH Londýn 2012
- Očekávané trendy v campus sítích
- Shrnutí



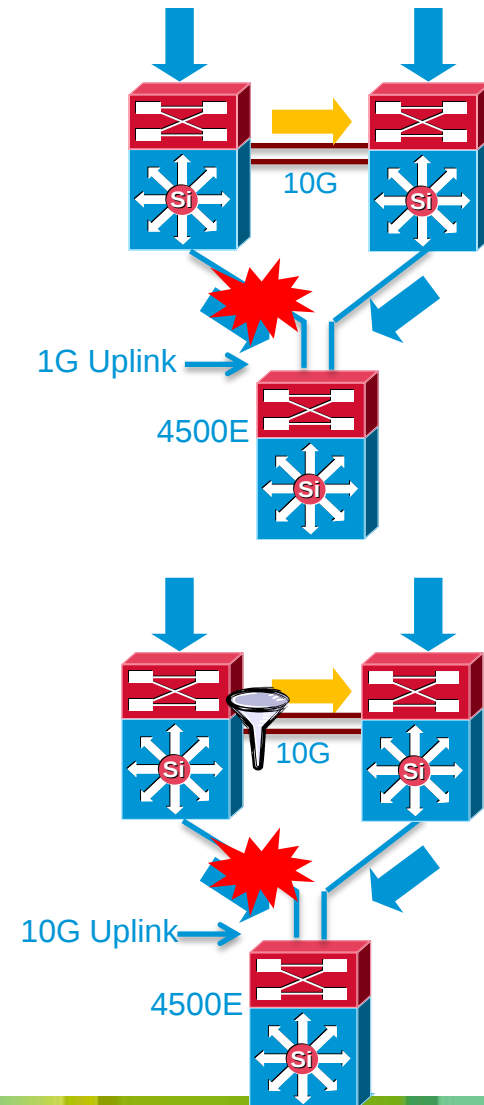
Redundantní připojení přístupové vrstvy

Modulární platforma s jedním supervisorem

- Uplinky ze supervisoru distribuujeme mezi switchy v přístupové vrstvě
- Nutno správně **dimenzovat propojení v rámci distribuční vrstvy** kvůli přesměrování provozu v případě výpadku spoje

Obzvláště při vysokorychlostním připojení z accessu

- Případně lze vytvořit i redundantní uplinky do každého switchu distribuční vrstvy



Redundantní připojení přístupové vrstvy

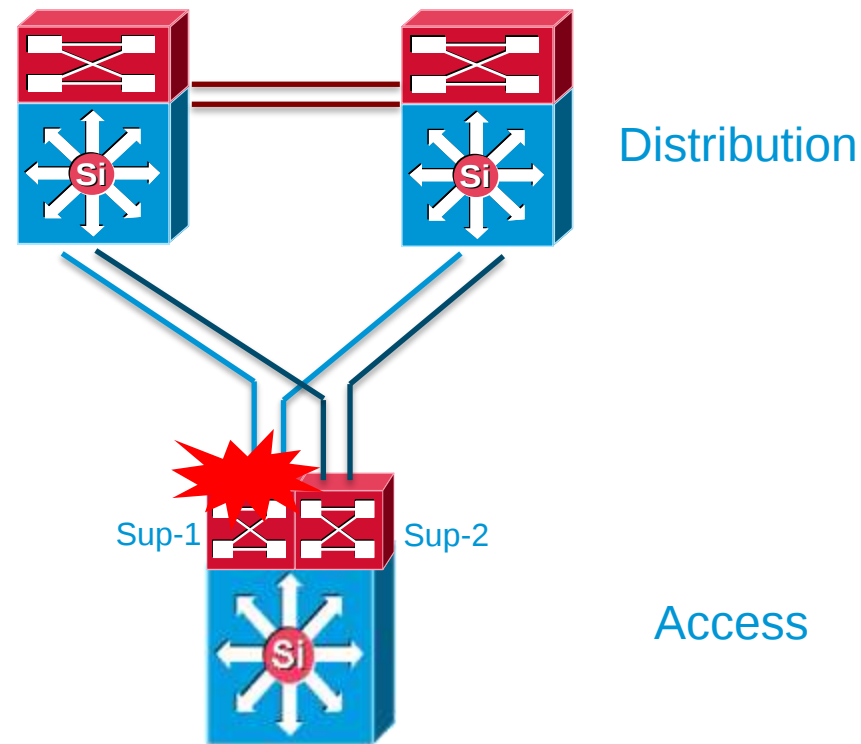
Modulární platforma se dvěma supervisory

- Non-stop komunikace díky redundantním supervisorům
Stateful SwitchOver
- Využití uplinku z každého supervisoru
- Využití dvou uplinků na každém supervisoru

Vyšší propustnost

Ale především minimalizace výpadku control plane, sub-second recovery při výpadku supervisoru

Dáno prakticky nezměněnou topologií

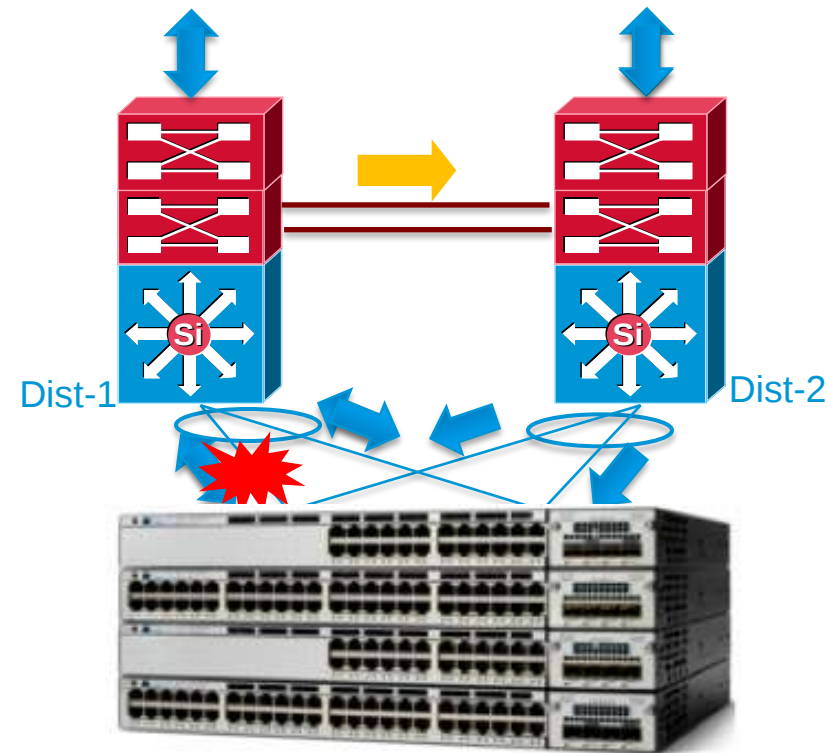


Redundantní připojení přístupové vrstvy

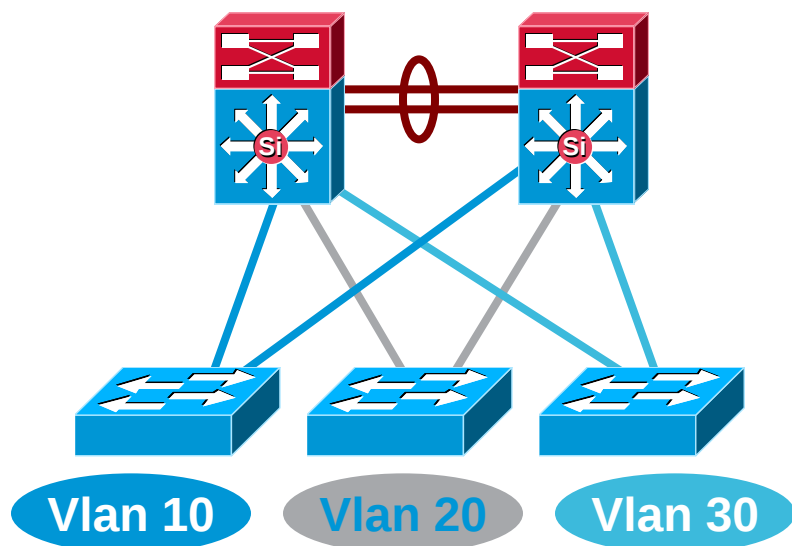
Fixní stohovaná platforma

Dual vs. Quad Uplink Network Design

- Fyzicky jsou stohované switche zapojeny v kruhu, logicky je to jedna control a management plane
- Uplinky je vhodné rozprostřít mezi členy stohu
- Pro případ výpadku uplinku je opět nutno plánovat
 - propojení mezi switchi distribuční vrstvy
 - a/nebo
 - více uplinků z access vrstvy
 - => nezmění se topologie, rychlejší recovery

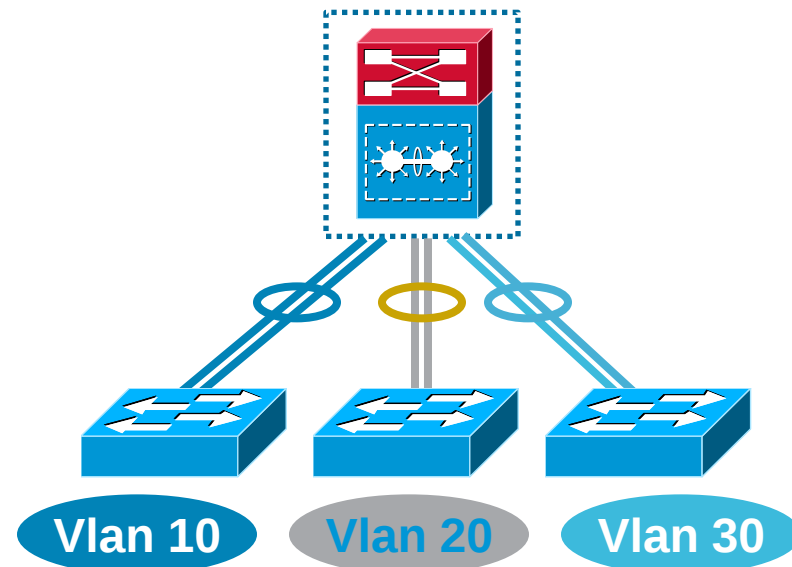


Tradiční a VSS model distribučního bloku



Tradiční návrh distribučního bloku

- Dva oddělené systémy
- Distribuovaný Data, Control a Management plane
- Detekce a recovery problému silně závisí na použitých L2 protokolech



VSS návrh distribučního bloku

- **Jeden** virtuální systém
- **Jednotný Control a Management plane**, distribuovaný Data plane
- **Deterministická** detekce a recovery

Tradiční návrh distribučního bloku

- Redundantní návrh se suboptimální topologií a vyšší mírou složitosti
- Nutnost stabilizovat topologii pomocí L2 mechanismů:

STP Primary and Backup Root Bridge

Rootguard

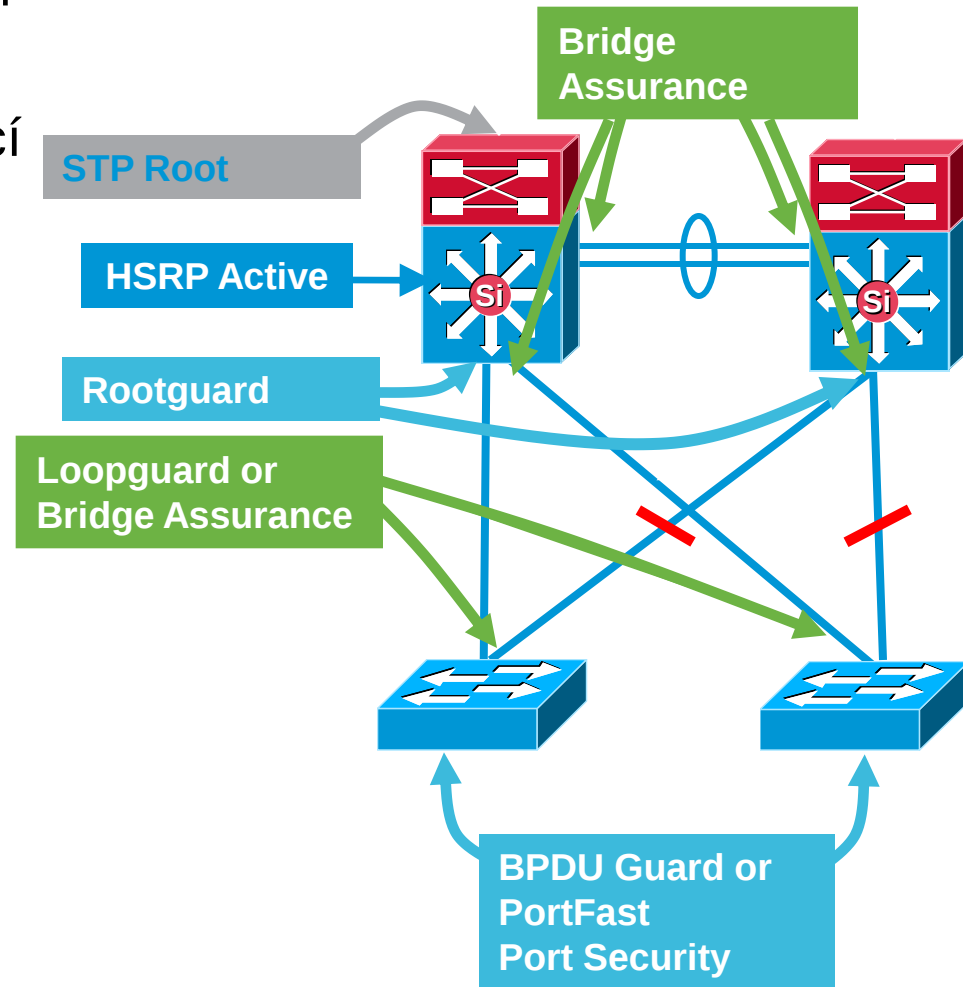
Loopguard or Bridge Assurance

STP Edge Protection

- Recovery lze ovlivnit pomocí

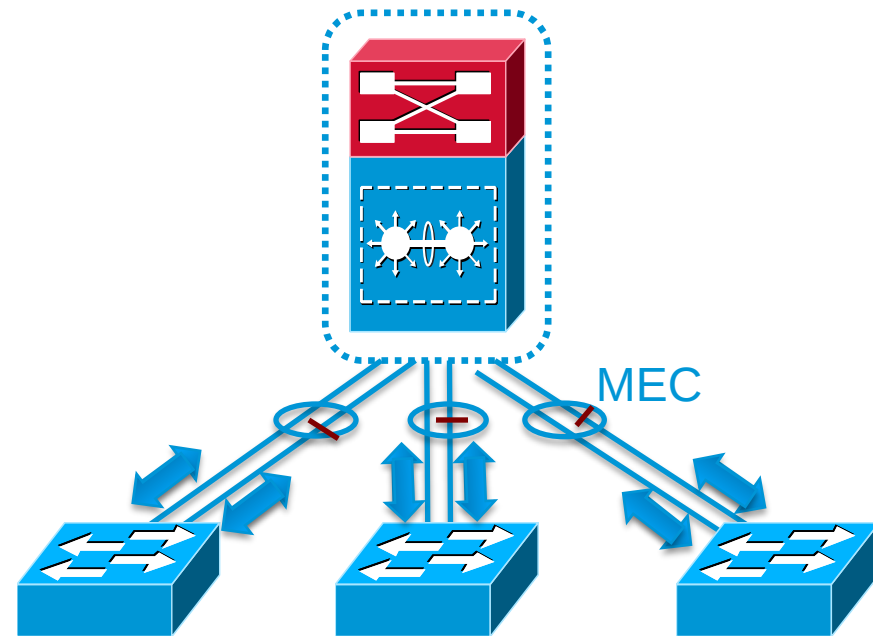
PVST/RPVST+

FHRP Tuning



Zjednodušení návrhu pomocí VSS: Spanning Tree

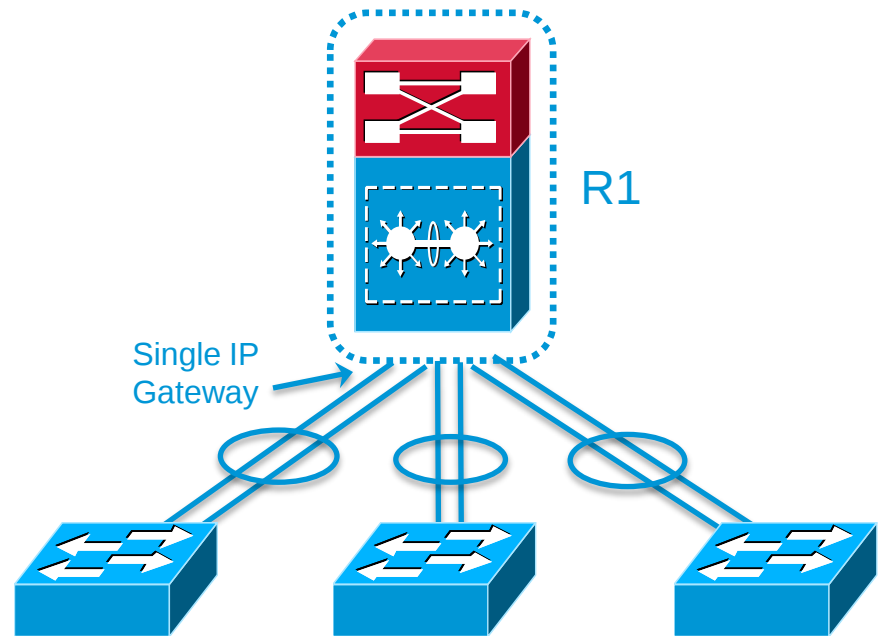
- V tradičním návrhu způsobí paralelní L2 cesty STP smyčku
- VSS díky Multichassis EtherChannelu umožňuje využít všechny dostupné linky
- Využití STP je stále doporučeno, ovšem pouze jako bezpečnostního mechanismu



— STP BLK Port
○ Loop-free L2 EtherChannel

Zjednodušení návrhu pomocí VSS: First Hop Resolution Protocol

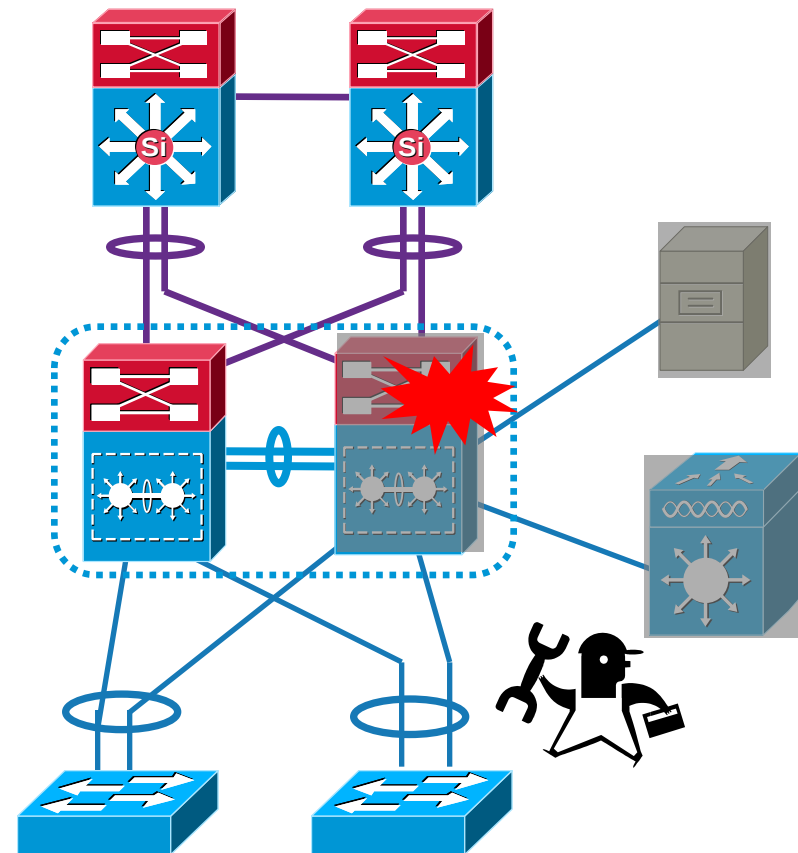
- Jedna logická Layer 3 brána
- Eliminace potřeby konfigurace FHRP (HSRP/GLBP/VRRP)
 - Nezávislost na FHRP timerech
- Deterministická sub-second konvergence
- Jednoduchá topologie, škálování, troubleshooting



Chování VSS v případě výpadku

Jeden supervisor v každém chassis

- Výpadek supervisoru znamená ztrátu jednoho chassis => snížení propustnosti VSS celku o 50%
- Některá zařízení mohou být z různých důvodů připojena pouze do jednoho chassis (single attach)
 - Servisní moduly
 - Servery
 - Geografické rozložení celého VSS chassis
- Takovýto výpadek supervisoru vyžaduje zásah obsluhy
 - Manuální proces instalace, zformování VSS z konfigurace druhé poloviny VSS
 - Nedeterministický čas výpadku



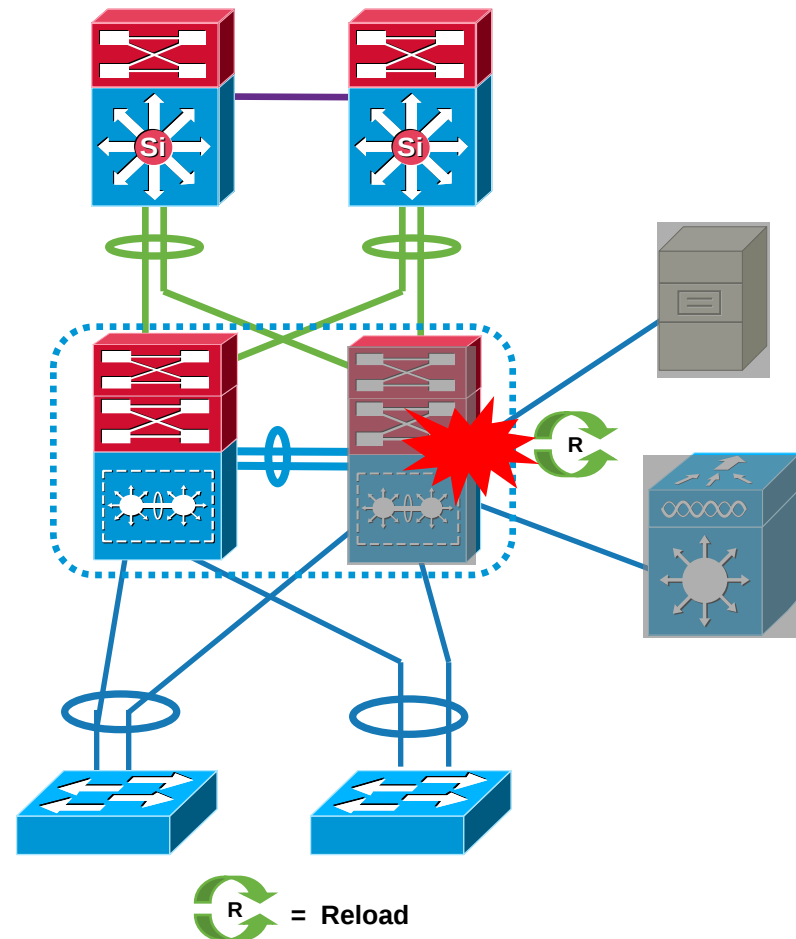
Chování VSS v případě výpadku

Dva supervisory v každém chassis (Quad-Sup) - Dnes

12.2(33)SXI4 a novější
Pouze Sup720

- Quad-Sup Uplink Forwarding
 - Druhý supervisor v každém chassis běží jako linecard se všemi porty aktivními
- Při výpadku aktivního supervisoru záložní supervisor automaticky
 - Není potřeba lidský manuální zásah
 - Deterministický výpadek

1. Supervisor Failure event
2. Chassis reloads
3. In-chassis Standby now becomes VSS standby and chassis dataplane is active again

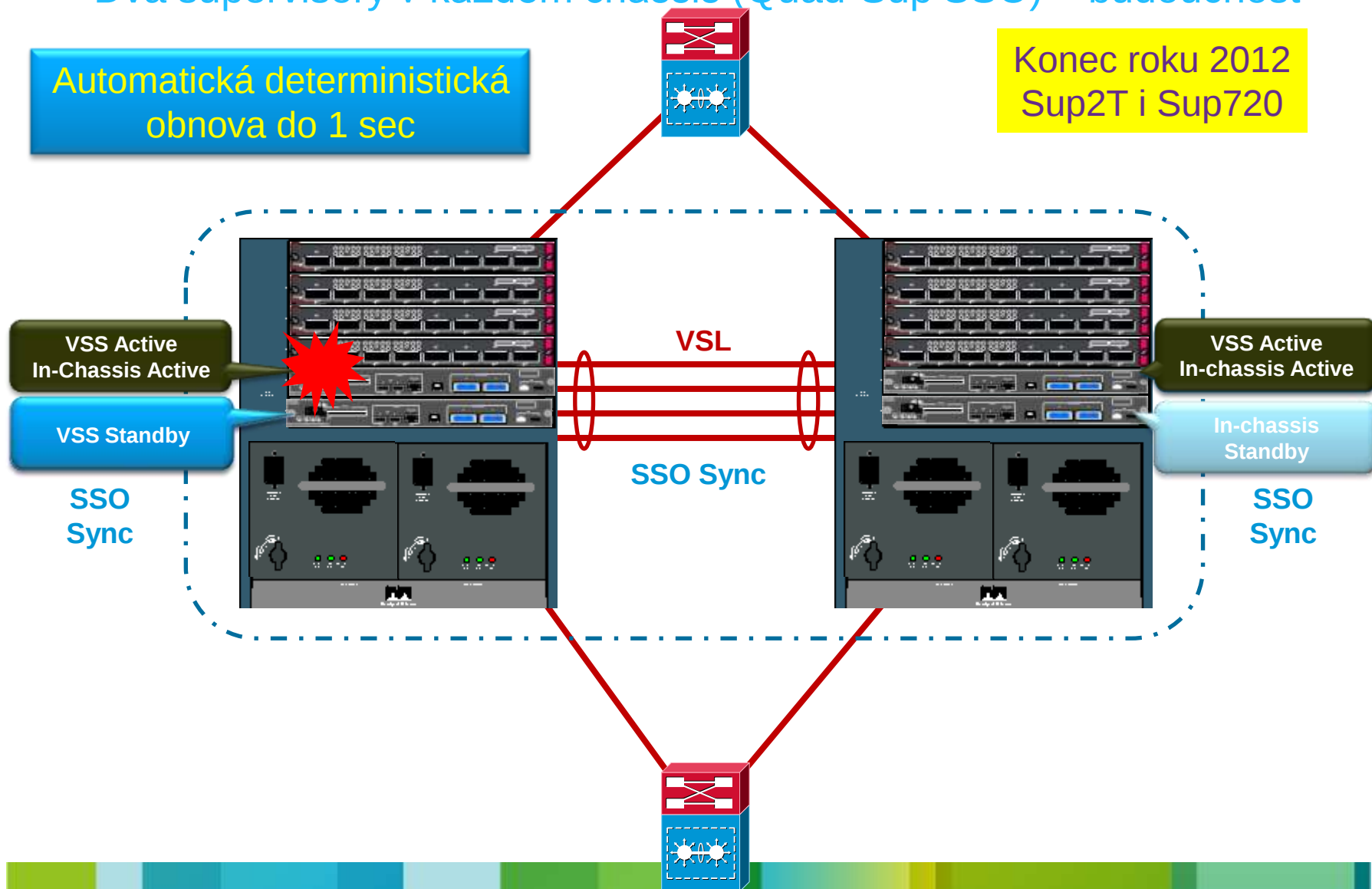


Chování VSS v případě výpadku

Dva supervisory v každém chassis (Quad-Sup SSO) – budoucnost

Automatická deterministická
obnova do 1 sec

Konec roku 2012
Sup2T i Sup720

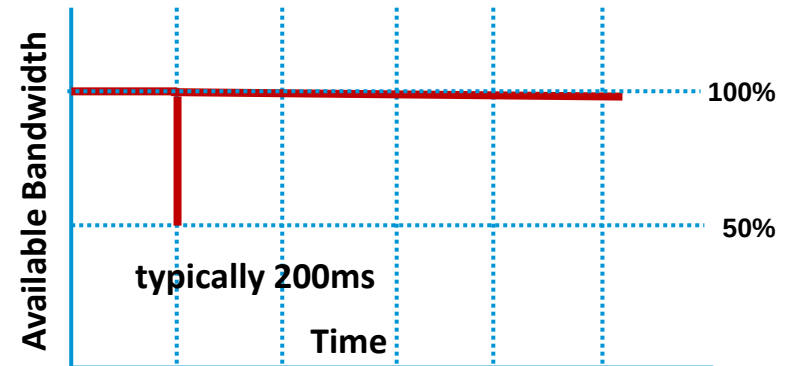


Chování VSS v případě výpadku

Porovnání

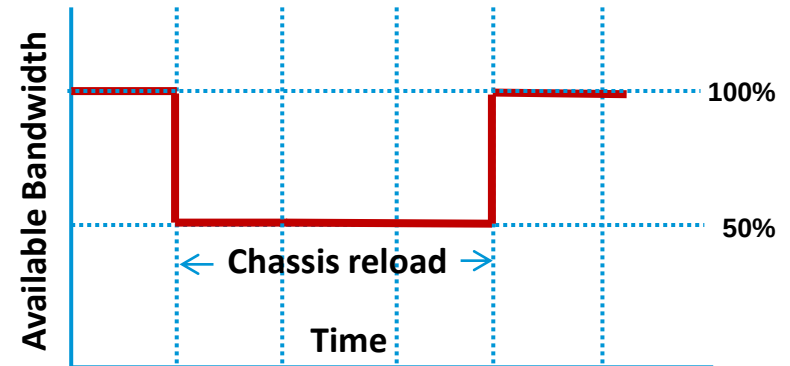
- Quad-Sup SSO

Plánováno na konec roku 2012

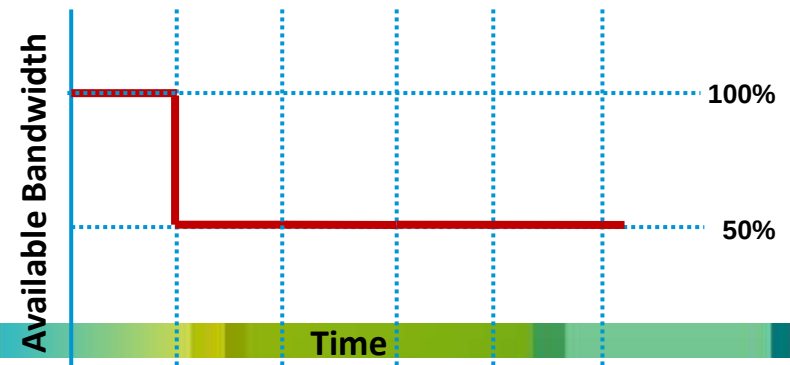


- Quad-Sup Uplink Forwarding

Pouze VS-Sup720-10G



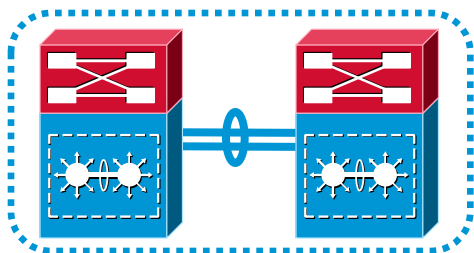
- Jeden supervisor per Chassis



VSS, vPC nebo Stacking ?

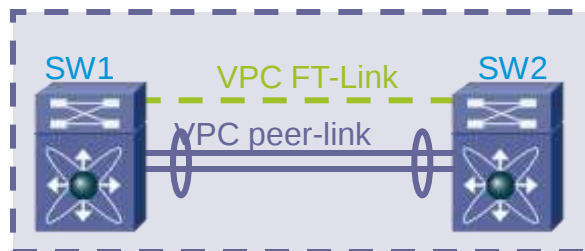
Virtual Switching System

- Jednotný control plane
- Jednotný management plane
- Redundantní supervisory v chassis
- Automatická synchronizace konfigurace portů
- Jednotná L3 doména (single SVI), FHRP není potřeba



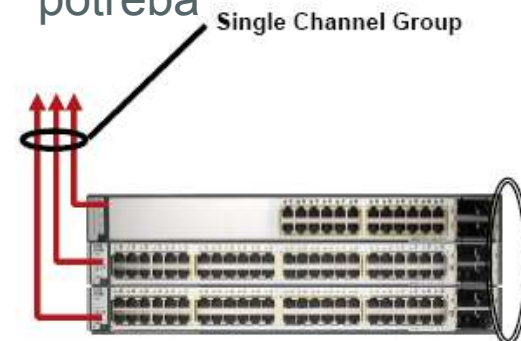
Virtual Port Channel

- **Separátní** control plane
- **Separátní** management plane se synchronizací VPC
- Redundantní supervisory per chassis s hitless SSO
- Manualní synchronizace konfigurace portů (DataCenterNetworkMgr)
- Vylepšený HSRP/PIM SVI pro active-active funkcionalitu



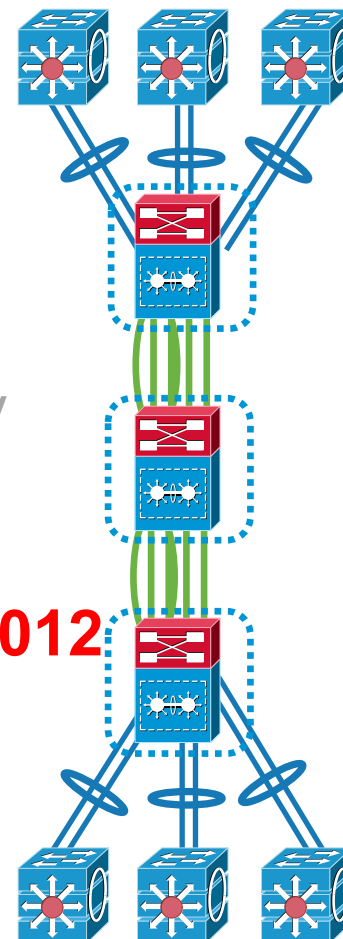
Stackwise+

- Jednotná control plane řízená Master Switchem
- Master switch řídí etherchannel
- Redundantní master switch per stack
- Automatická synchronizace konfigurace portů
- Stack se jeví jako jeden směrovač, FHRP není potřeba



Program

- Proč vysoce dostupné sítě ?
- Tradiční návrh campus sítí
- Alternativy řešení přístupové a distribuční vrstvy
- Aspekty redundantního návrhu
- **Case Study – Campus síť pro LOH Londýn 2012**
- Očekávané trendy v campus sítích
- Shrnutí





*official partner of the
Olympic and Paralympic Games*

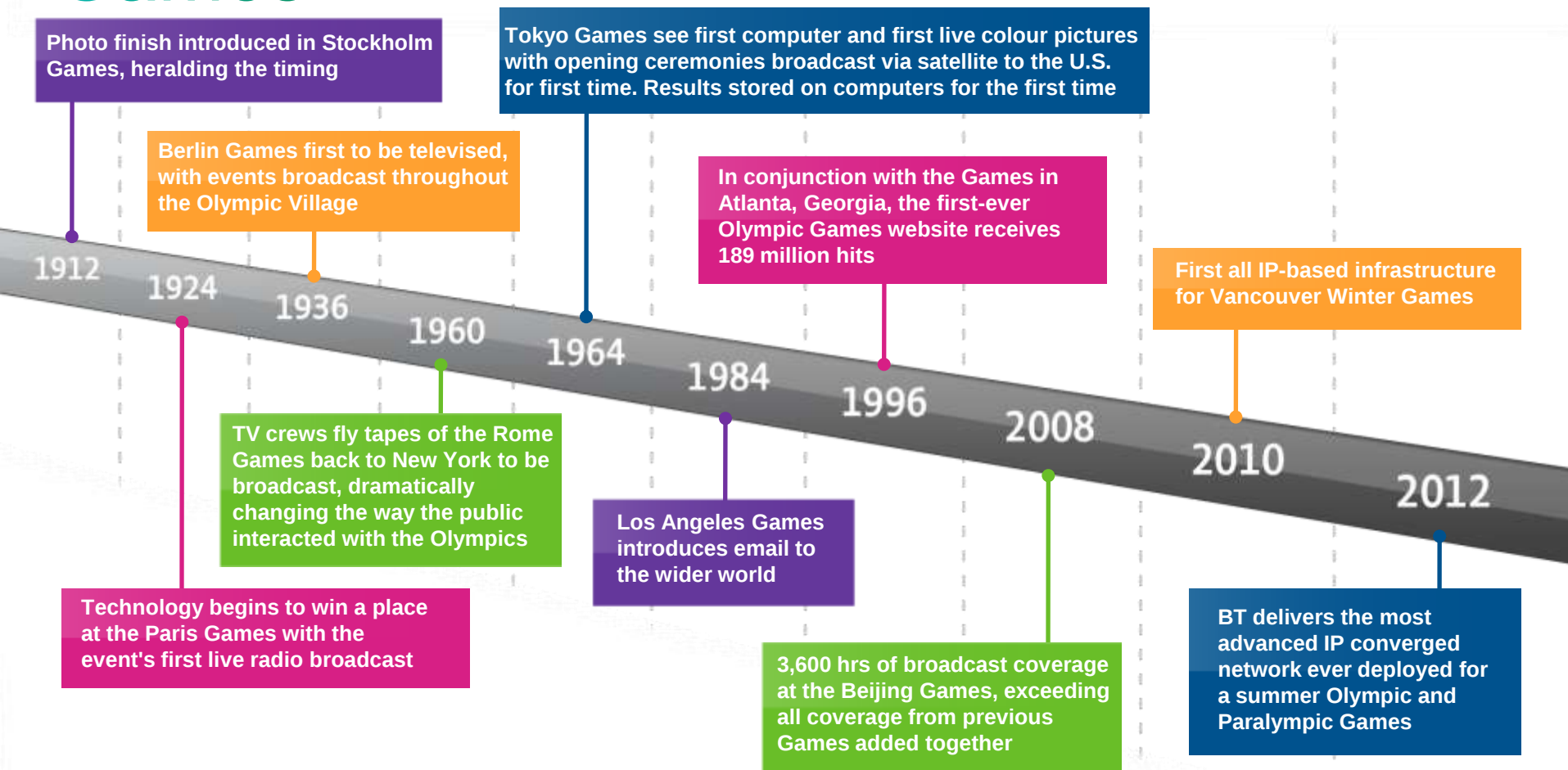
LOH Londýn 2012



www.ciscolondon2012.co.uk



Evolution of Technology through the Games



London 2012 communications services

80,000 connections across **94** locations

Up to **60Gb** of information carried each second

1,800 wireless access points

5,500m of internal cabling

16,500 telephone lines

14,000 mobile
SIM cards

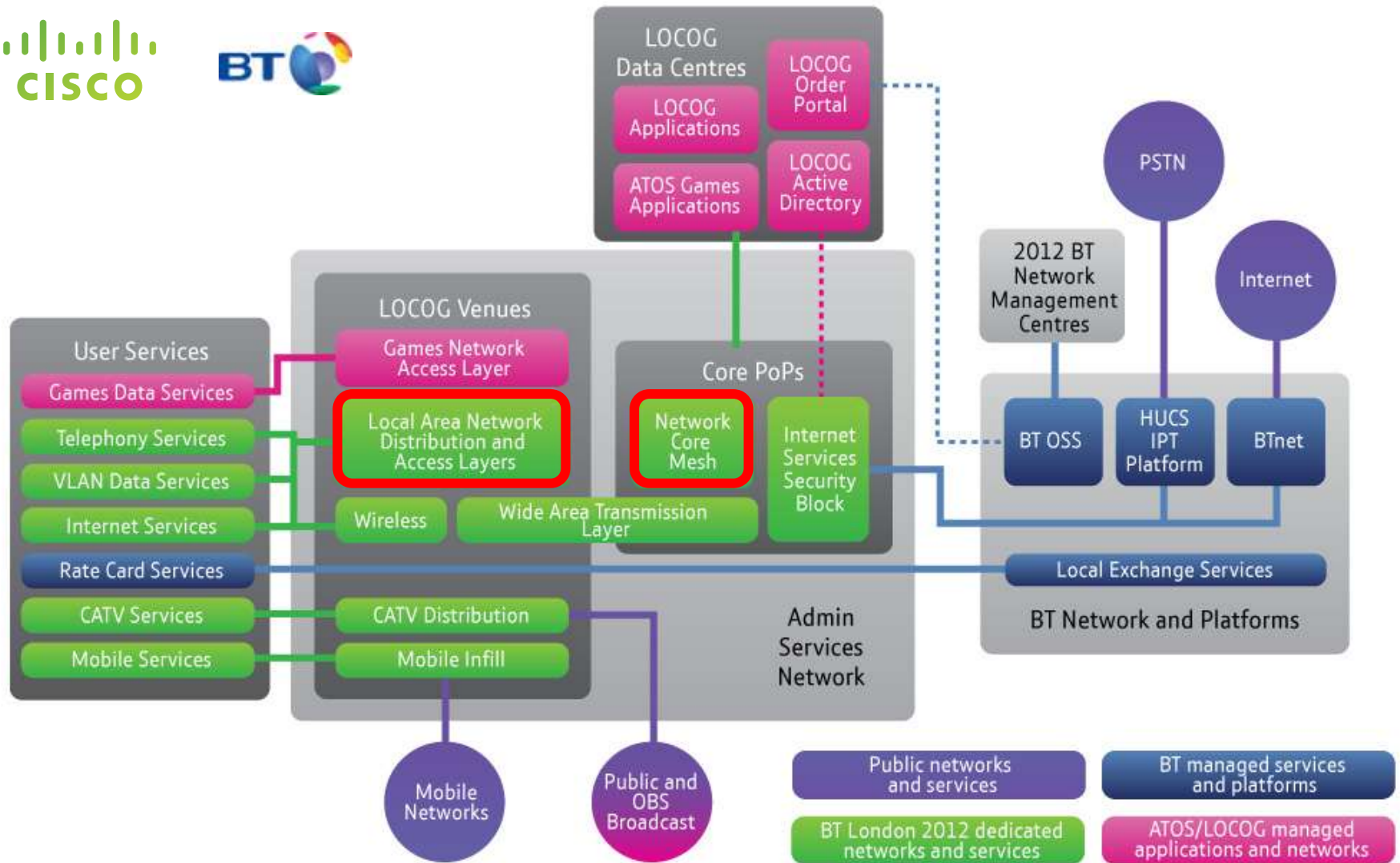
10,000 cable
TV outlets

642,000
man-hours

Over **800** people
on the ground at
Games' Time



London 2012 Communications Services Model



Network Architecture

Design priorities

All Core and Competition Venues are implemented using Cisco VSS

1. Reliability

Capable of surviving loss of core site and/or core route, with fast convergence.

2. Flexibility:

Dynamic solution able to cater for a myriad of service types, with different traffic requirements

3. Scalability:

Can support all likely bandwidth needs of the Games.

4. Security:

Hardware based features to segregate traffic, and provide infrastructure protection.

5. Manageability:

Proactively monitored with rapid fault resolution.

6. Cost-effectiveness:

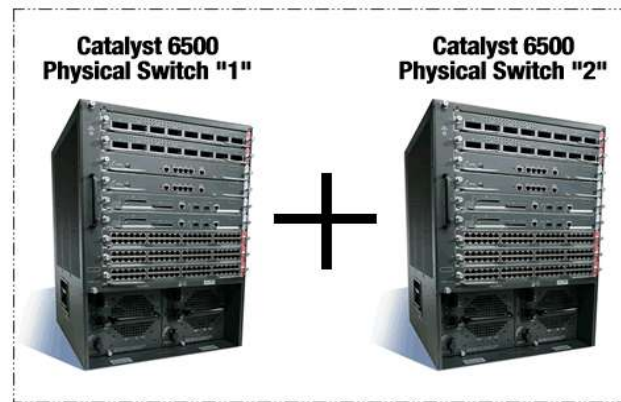
Standard, off-the-shelf products with multi-service software features within IOS

7. Sustainability:

VLANs and VRFs allow multiple services per box, rather than device per function

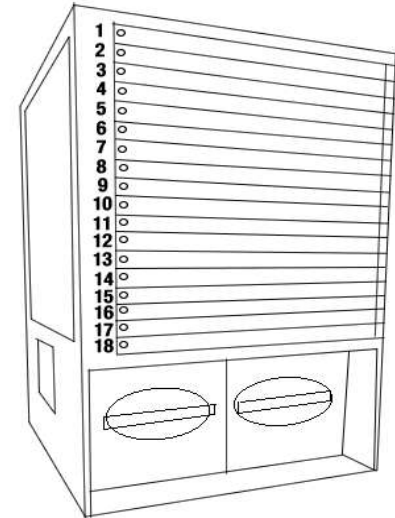
Network Architecture

All Core and Competition Venues are implemented using Cisco VSS



Virtual Switching System (VSS)

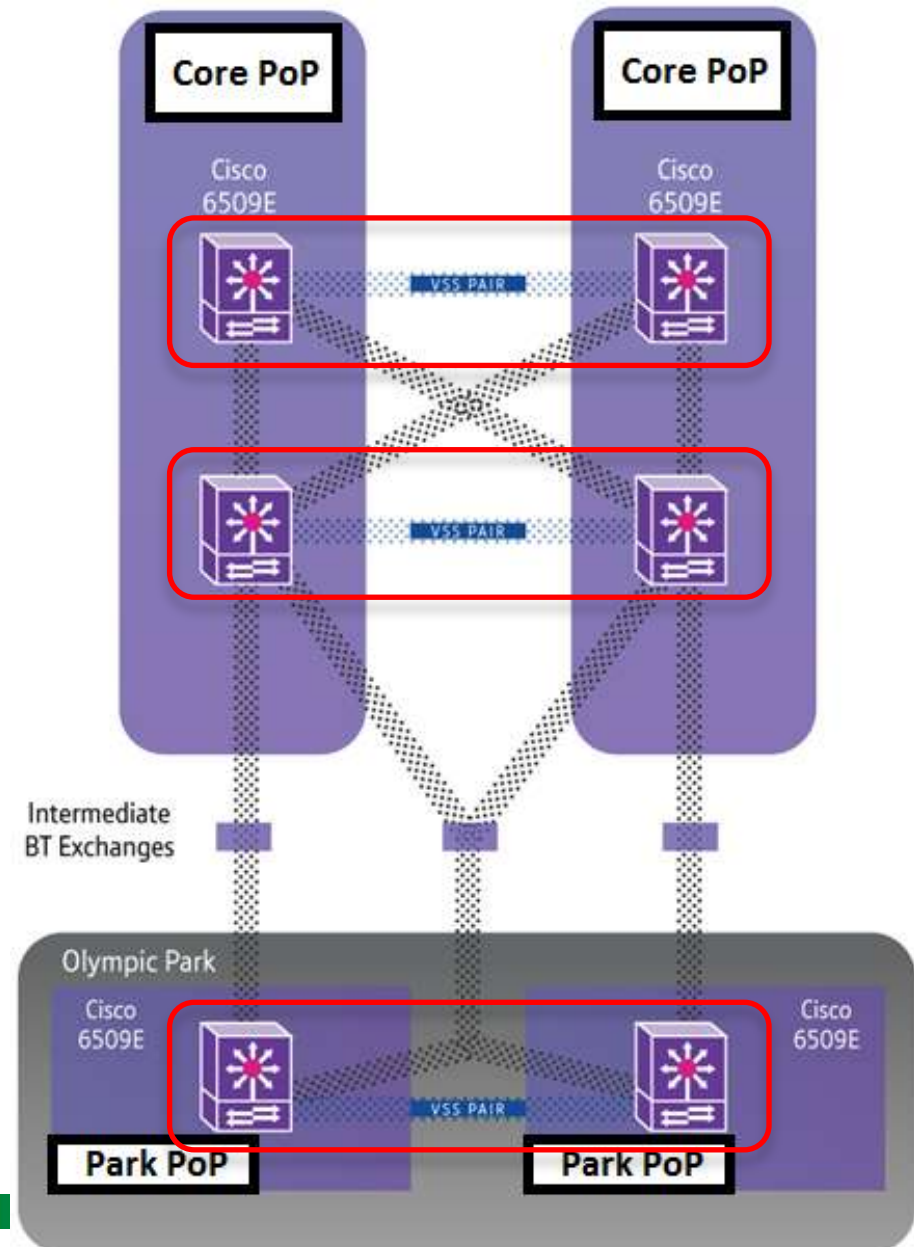
defines two physical Catalyst 6500 switches joined via a special link called a Virtual Switch Link (VSL) running special hardware and software that allows them to operate as a single logical switch



- Large Campus, Ethernet end to end
- Fully converged network, (voice, video and data),
 - comprehensive QoS - 10 traffic classes
- HW used:
 - Core/Distribution - Cisco 6509E with VS-Sup720 with VSS, 67xx with DFC
 - Access - Cisco 3750E

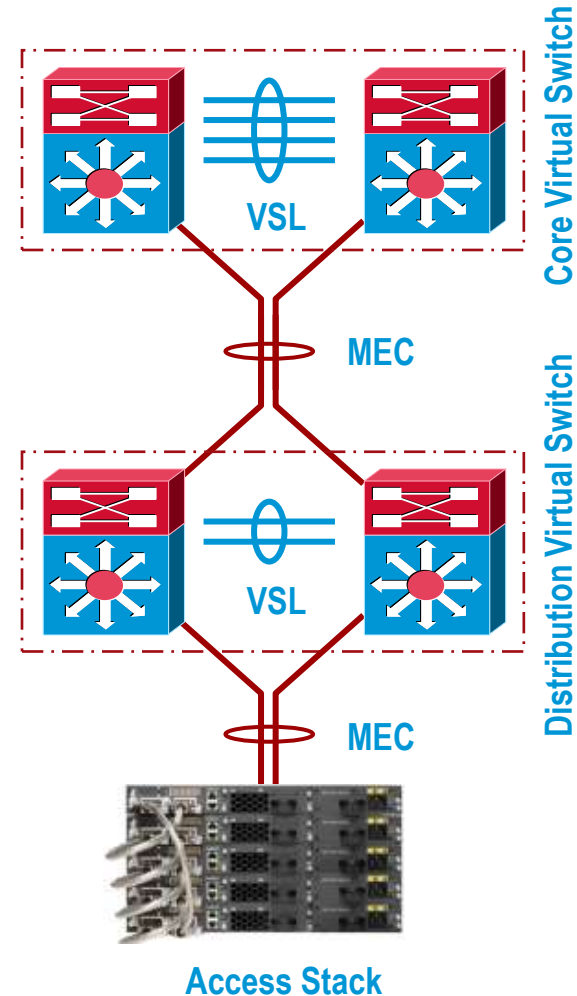
Network Core Mesh

- Core = 3 VSS pairs
- Competition Venues each connect to 2 Core PoPs
- Inter-PoP
 - physical routes redundancy
 - Nx10GE



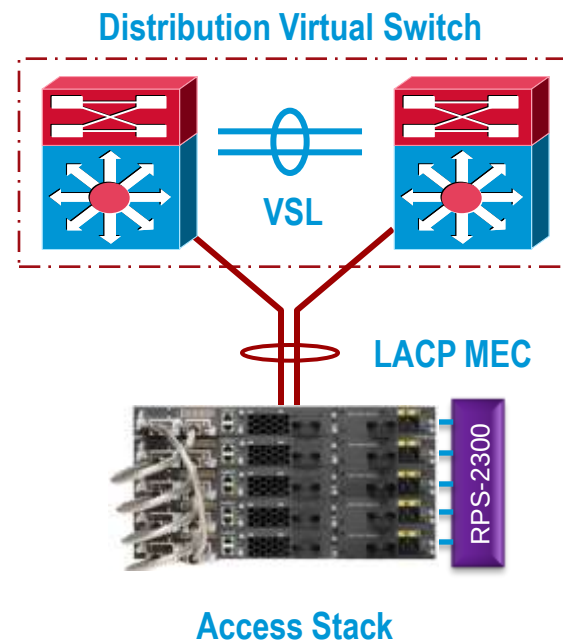
VSS in Core and Distribution

- **Active-Active Bandwidth utilisation**
 - Layer 2 Multi-Chassis Ether-Channels (MECs)
 - **very clean logical design**, with optimal use of bandwidth
- Loopless topology
 - RSTP(802.1w) deployed only for loop prevention
- VSS Dual-Active Detection – 2 of the 3 possible mechanisms deployed (Enhanced PAgP and IP-BFD)
- **UDLD deployed** to detect uni-directional links and prevent loss of traffic



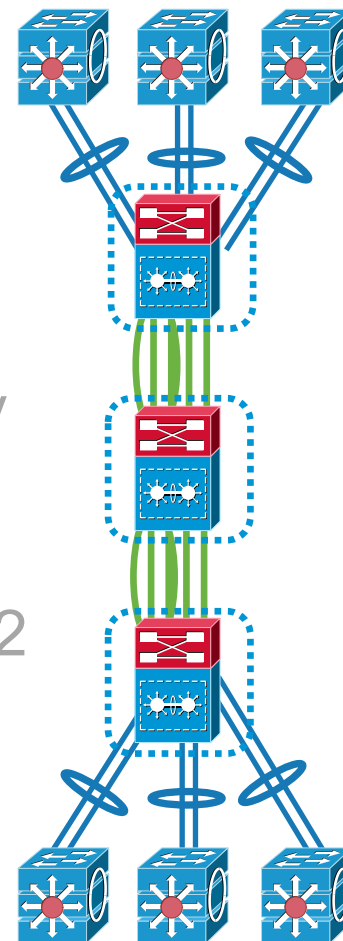
Competition Venue Access Layer

- 3750E 48 port PoE stacks at the user Access Layer, supporting:
 - Stackwise Plus ensures a single management point
 - 10/100/1000 Copper
 - Class 3 PoE for IP Phones and Wireless APs
 - 2 x 10G MEC uplink to Distribution VSS node
 - Limit of six switches per stack, RPS-2300
- 8 port 3560 PoE used for reach (e.g. Wireless APs)
- 3750-X series was introduced too late for consideration.



Program

- Proč vysoce dostupné sítě ?
- Tradiční návrh campus sítí
- Alternativy řešení přístupové a distribuční vrstvy
- Aspekty redundantního návrhu
- Case Study – Campus síť pro LOH Londýn 2012
- **Očekávané trendy v campus sítích**
- Shrnutí



Předpokládaný rozvoj

- Současný rozvoj v DC = budoucí rozvoj v campus sítích ???

Rozvoj DC – L2 span, vysoká hustota portů

Campus – PoE/EnergyWise, Wireless, MediaNet, TrustSec, ...

=> obdobné principy jako v DC plus další vlastnosti

Škálování L2 , změna L2 řídicího protokolu

L2 Multipathing

Totální eliminace Spanning Tree

=> Cisco Fabric path / TRILL

Spravovatelnost

Snižování počtu spravovaných logických zařízení

=> “Satelitní” přepínače

- Konvergence Wired & Wireless

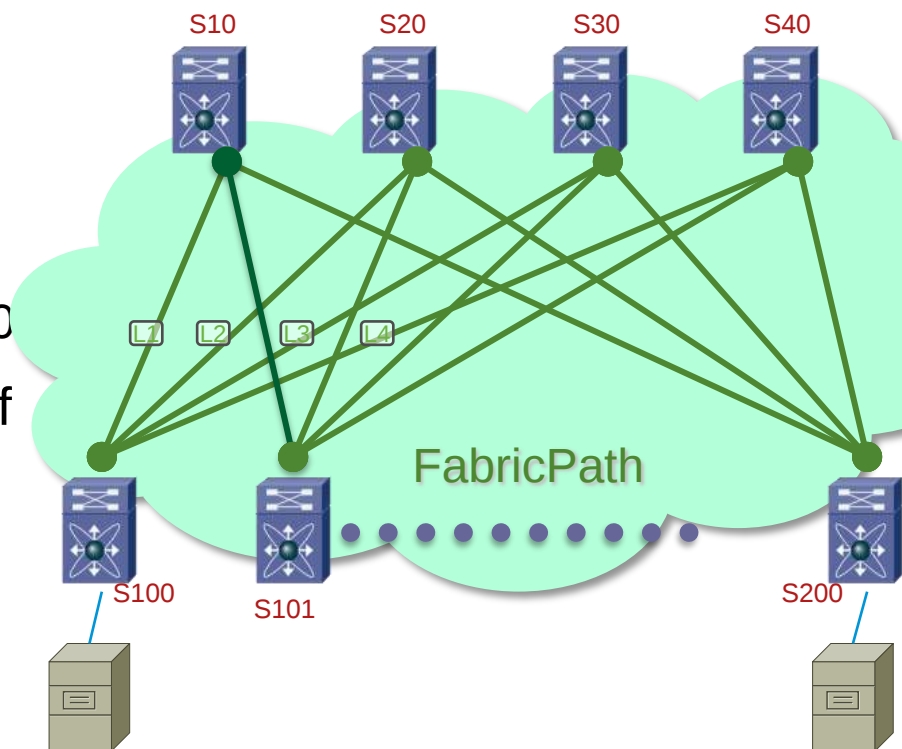
Viz techtoriál “Od spanning tree ke směrování na druhé vrstvě”, T-NET4, zítra14:30

Viz techtoriál “Jednotné řízení přístupu v bezdrátových a pevných sítích”, T-NET3,zítra 13:30

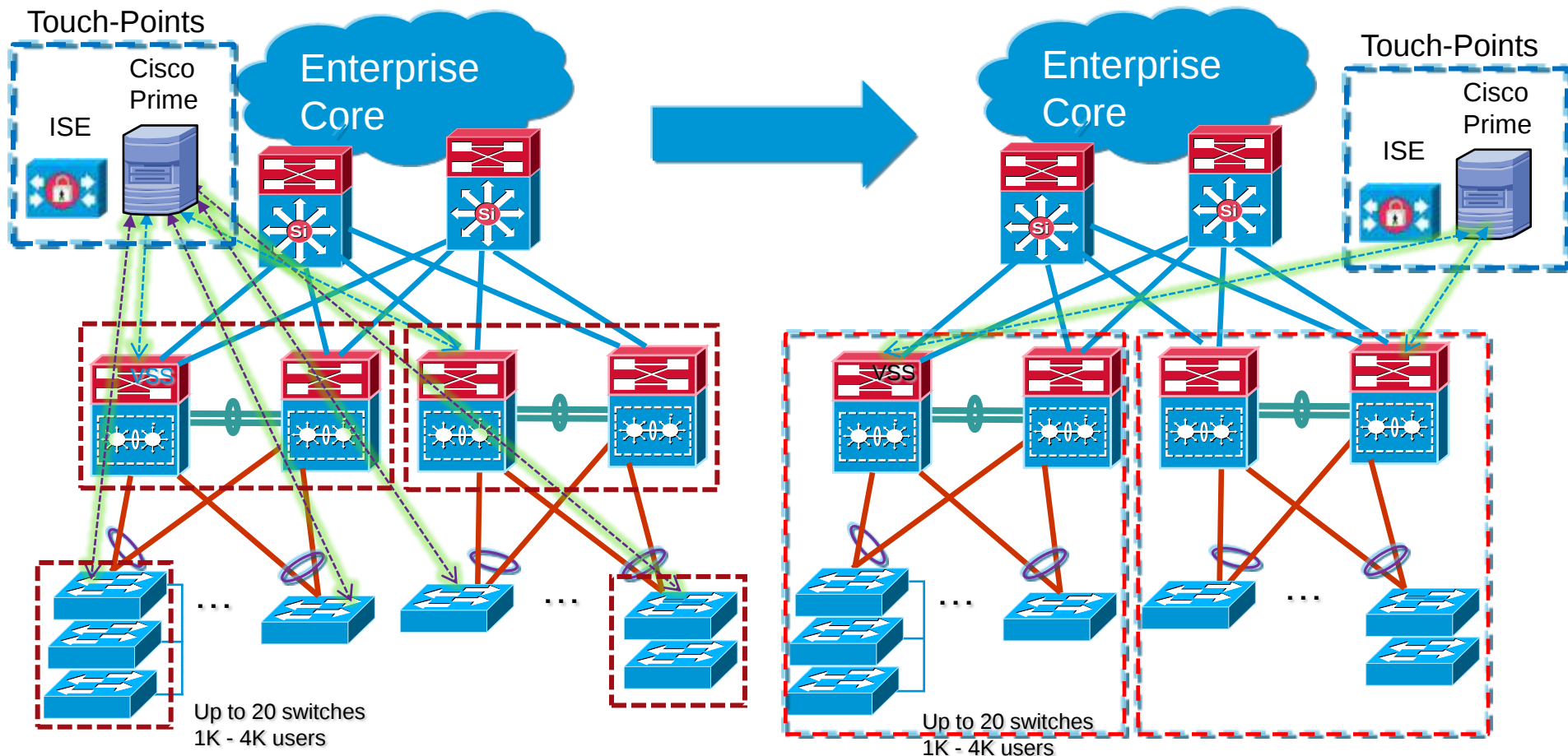
Další evoluční krok - FabricPath

Layer 2 Multipathing

- Konečné odstranění STP ze sítě
- Navýšení propustnosti L2 s využitím paralelních aktivních linek
- Cisco FabricPath
 - V současnosti pro použití v DC
 - Dostupný pro Nexus 7000 a Nexus 5500
- Transparent Interconnection of Lots of Links (TRILL) – RFC 6326
- Principy a protokoly známé z IP L3 multipathing jsou aplikovány na L2
 - Využití IS-IS
 - Jednoduchá konfigurace

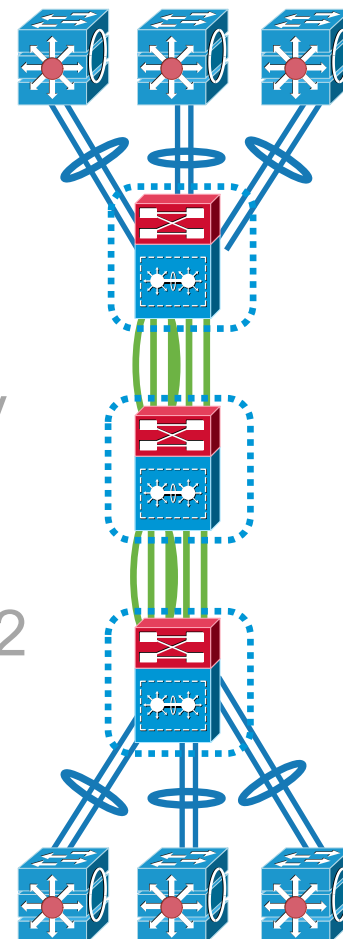


Snižování počtu spravovaných zařízení Satelity v Campus síti ?



Program

- Proč vysoce dostupné sítě ?
- Tradiční návrh campus sítí
- Alternativy řešení přístupové a distribuční vrstvy
- Aspekty redundantního návrhu
- Case Study – Campus síť pro LOH Londýn 2012
- Očekávané trendy v campus sítích
- **Shrnutí**



Shrnutí

- Nároky na dostupnost campus sítí prudce narůstají
- Při návrhu aplikovat hierarchický modulární design
 - Access – Distribution - Core
 - Redundance , škálovatelnost, deterministické chování, spravovatelnost
- Pokud možno využívat VSS/VPC/stacking
 - Vyzrálý a ověřený design
 - Pro vyšší dostupnost – Quad Sup
- Pokud tradiční design – dimenzování propojů, aplikovat STP toolkit, vyladit FHRP
- Do budoucna lze očekávat
 - Totální náhradu STP jako L2 řídicího protokolu pomocí L2 směrování
 - Zjednodušení správy – satelity
 - Wired-Wireless konvergence

Odkazy

- Design Zone for Borderless Networks

www.cisco.com/go/designzone/borderless

- Borderless Networks: Medium Enterprise Design Profile

http://www.cisco.com/en/US/docs/solutions/Enterprise/Medium_Enterprise_Design_Profile/MEDP.html

- Campus 3.0 VSS Design Guide

http://www.cisco.com/en/US/docs/solutions/Enterprise/Campus/VSS30dg/campusVSS_DG.html

- VSS Support Pages

www.cisco.com/go/support

[Cisco Catalyst 6500 Virtual Switching System Deployment Best Practices](#)

[Migrate Standalone Cisco Catalyst 6500 Switch to Cisco Catalyst 6500 Virtual Switching System](#)

[Troubleshoot Packet Flow in Cisco Catalyst 6500 Series Virtual Switching System 1440](#)

- VSS White Paper

http://www.cisco.com/en/US/prod/collateral/switches/ps5718/ps9336/white_paper_c11_429338.pdf

- Catalyst 6500 Configuration Guide

<http://www.cisco.com/en/US/partner/docs/switches/lan/catalyst6500/ios/12.2SX/configuration/guide/vss.html>

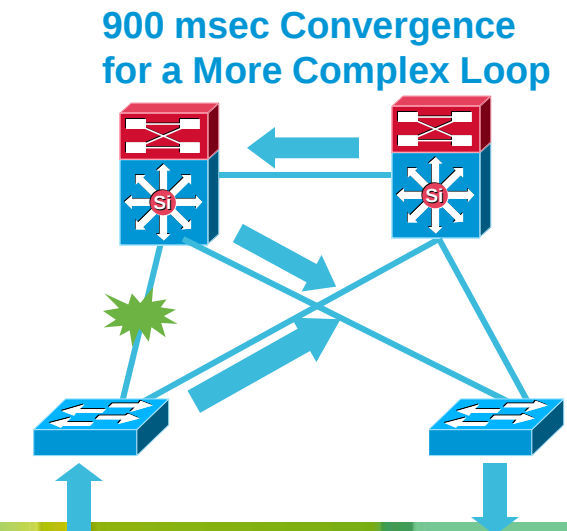
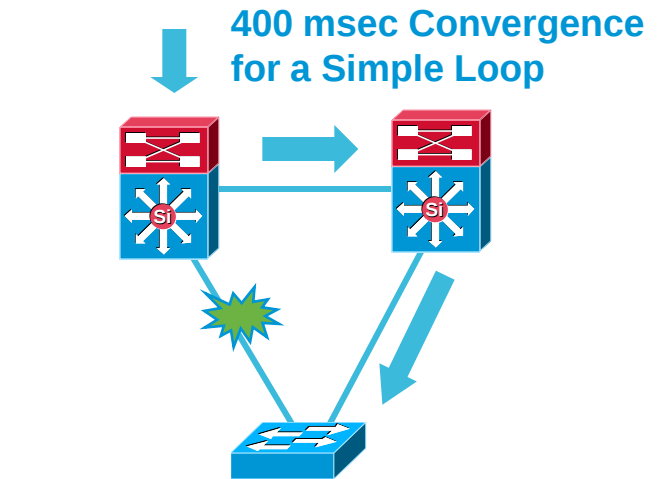


Doplňkové materiály



Convergence Is Longer with Complex Spanning Tree Loops and Older Protocols

- Time to converge is dependent on the protocol implemented 802.1d, 802.1s or 802.1w (all now a part of IEEE 802.1d 2004 spec)
- It is also dependent on:
 - Size and shape of the L2 topology (how deep is the tree)
 - Number of VLAN's being trunked across each link
 - Number of ports in the VLAN on each switch
- Complex Topologies Take Longer to Converge



Even with Faster Convergence from RPVST+ We Still Have to Wait on FHRP Convergence

VRRP Config

```
interface Vlan4
ip address 10.120.4.1 255.255.255.0
ip helper-address 10.121.0.5
no ip redirects
vrrp 1 description Master VRRP
vrrp 1 ip 10.120.4.1
vrrp 1 timers advertise msec 250
vrrp 1 preempt delay minimum 180
```

HSRP Config

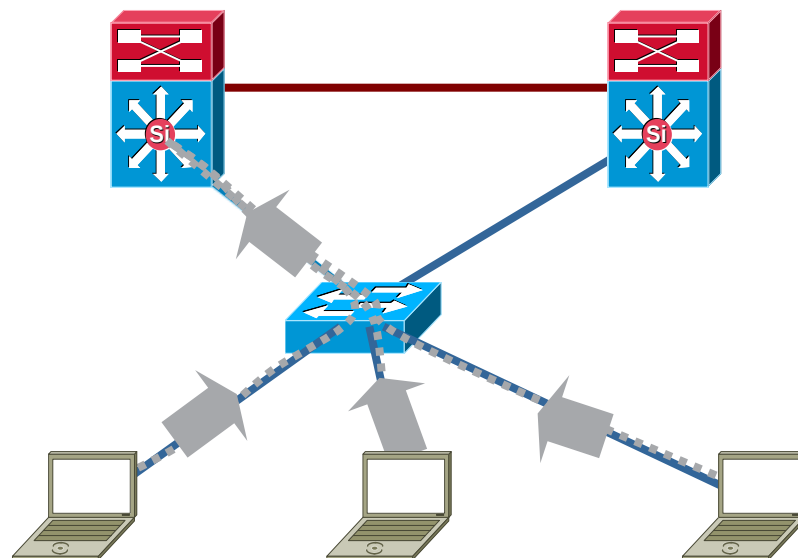
```
interface Vlan4
ip address 10.120.4.2 255.255.255.0
standby 1 ip 10.120.4.1
standby 1 timers msec 250 msec 750
standby 1 priority 150
standby 1 preempt
standby 1 preempt delay minimum 180
```

GLBP Config

```
interface Vlan4
ip address 10.120.4.2 255.255.255.0
glbp 1 ip 10.120.4.1
glbp 1 timers msec 250 msec 750
glbp 1 priority 150
glbp 1 preempt
glbp 1 preempt delay minimum 180
```

FHRP Active

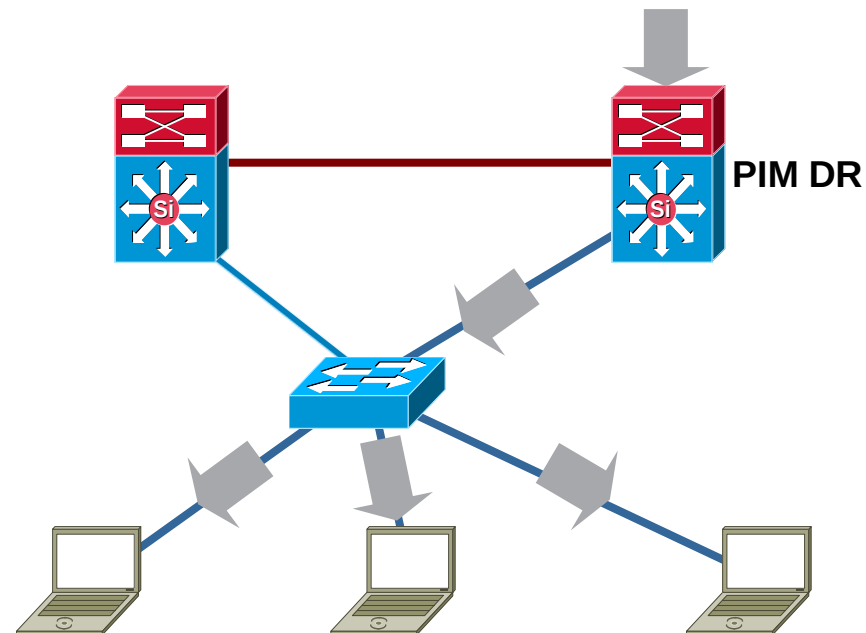
FHRP Standby



- GLBP offers load balancing within a VLAN
- For Voice, sub-second Hello timer enables < 1 Sec traffic recovery upstream
- Sub-Second protocol timers must be avoided on SSO capable network

PIM Needs Timer Tuning Too

- Multicast recovery depends on PIM DR failure detection in Layer 2 network
- PIM routers exchanges PIM expiration time in query message –
 - Default Query-Interval – 30 seconds
 - Expiration – Query Interval x 3
 - DR Failure Detection – ~90 seconds
- Tune PIM query interval to sub-sec as FHRP for faster multicast convergence
- Sub-second protocol timer must be avoided on SSO capable network



```
interface Vlan4
ip pim sparse-mode
ip pim query-interval 250 msec
```

Subsecond Protocol Timers and NSF/SSO

- NSF is intended to provide availability through route convergence **avoidance**
- Fast IGP timers are intended to provide availability through fast route **convergence**
- In an NSF environment dead timer must be greater than:

SSO recovery + Routing Protocol restart + time to send first hello

