

Internet a propojování sítí

Adam Golecky

sut.sh.cvut.cz

29.4.2014



NIX.CZ

Kdo/co je Internet?

- Poskytovatelé (Content) obsahu

vs

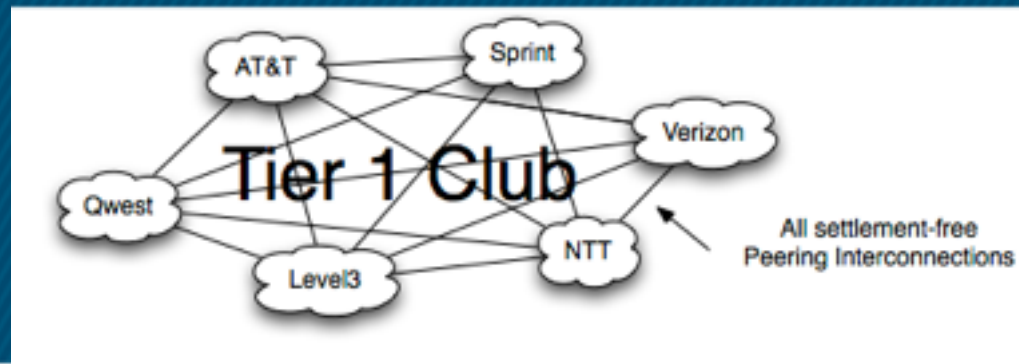
Operatori (ISP)

- Rozdělení na Tier 1 vs Tier 2 etc
- Regionální vs Globální operátoři

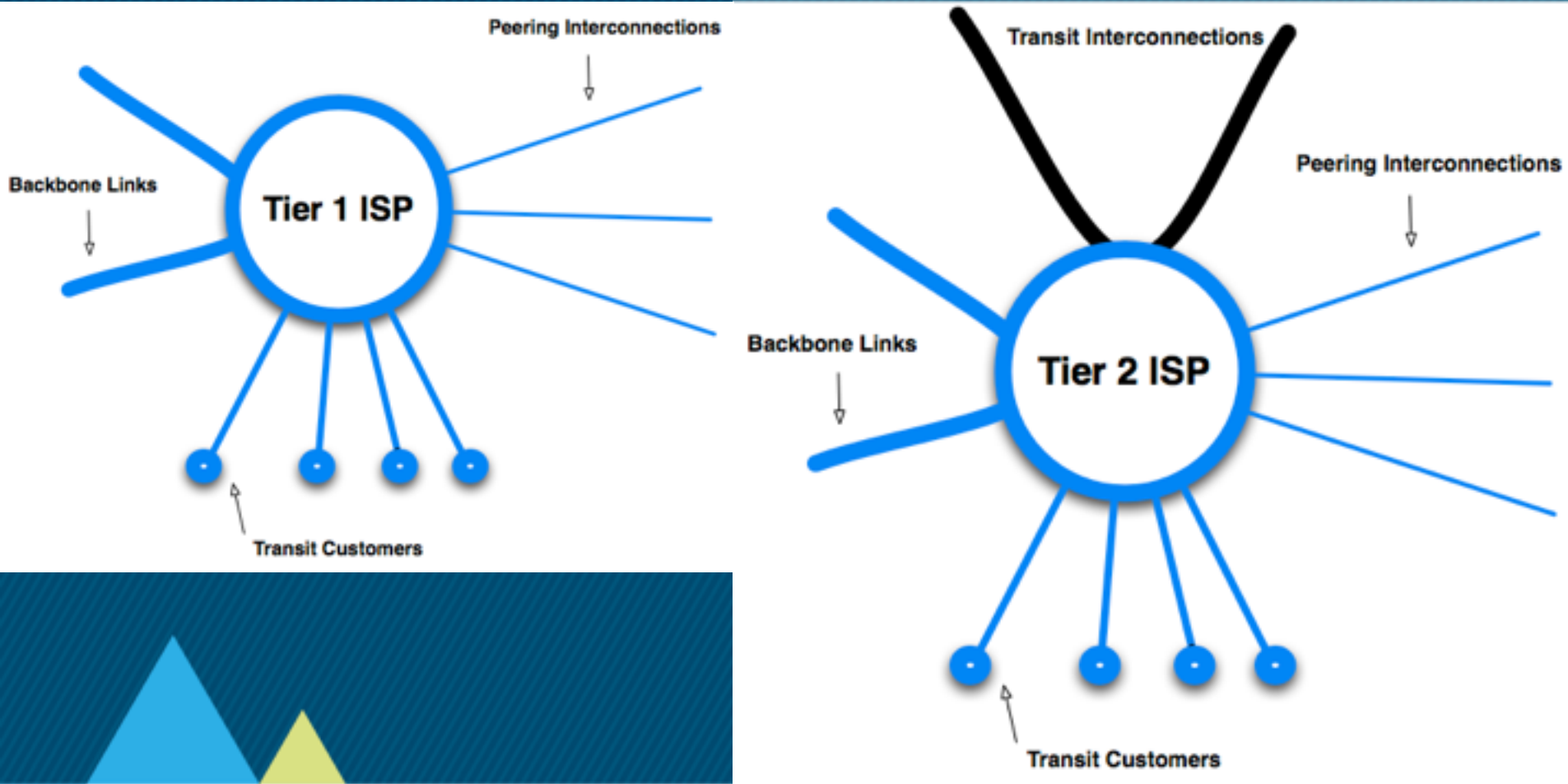


Kdo/co jsou Tier 1?

- “Final club”
- Globální sítě
- Bezplatně se propojují pouze s jinými T1
- Tier2 a T3 jim platí za přístup do jejich sítí
- http://en.wikipedia.org/wiki/Tier_1_network



Tier 1 vs Tier 2, 3....



Content Delivery Network

- sítě pro rozložení zátěže pro distribuci
- zlepšení UX (user experience)
- snížení nákladů na distribuci
- Akamai, Amazon, Limelight
- např.: distribuce aktualizací nebo streaming videa



Modely nákup tranzitu

- fixní cena za port
 - konstantní OPEX
- 95th percentile
 - OPEX dle skutečné “spotřeby”

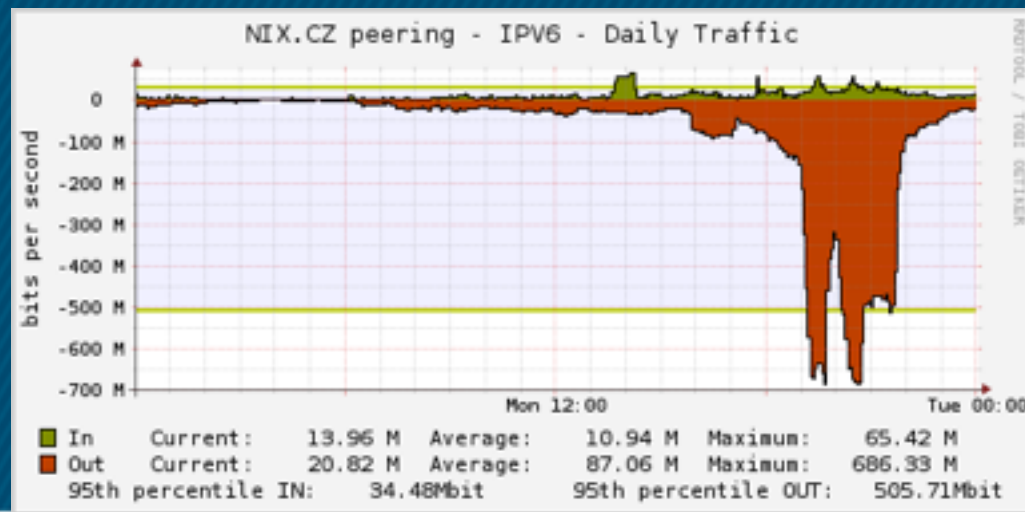
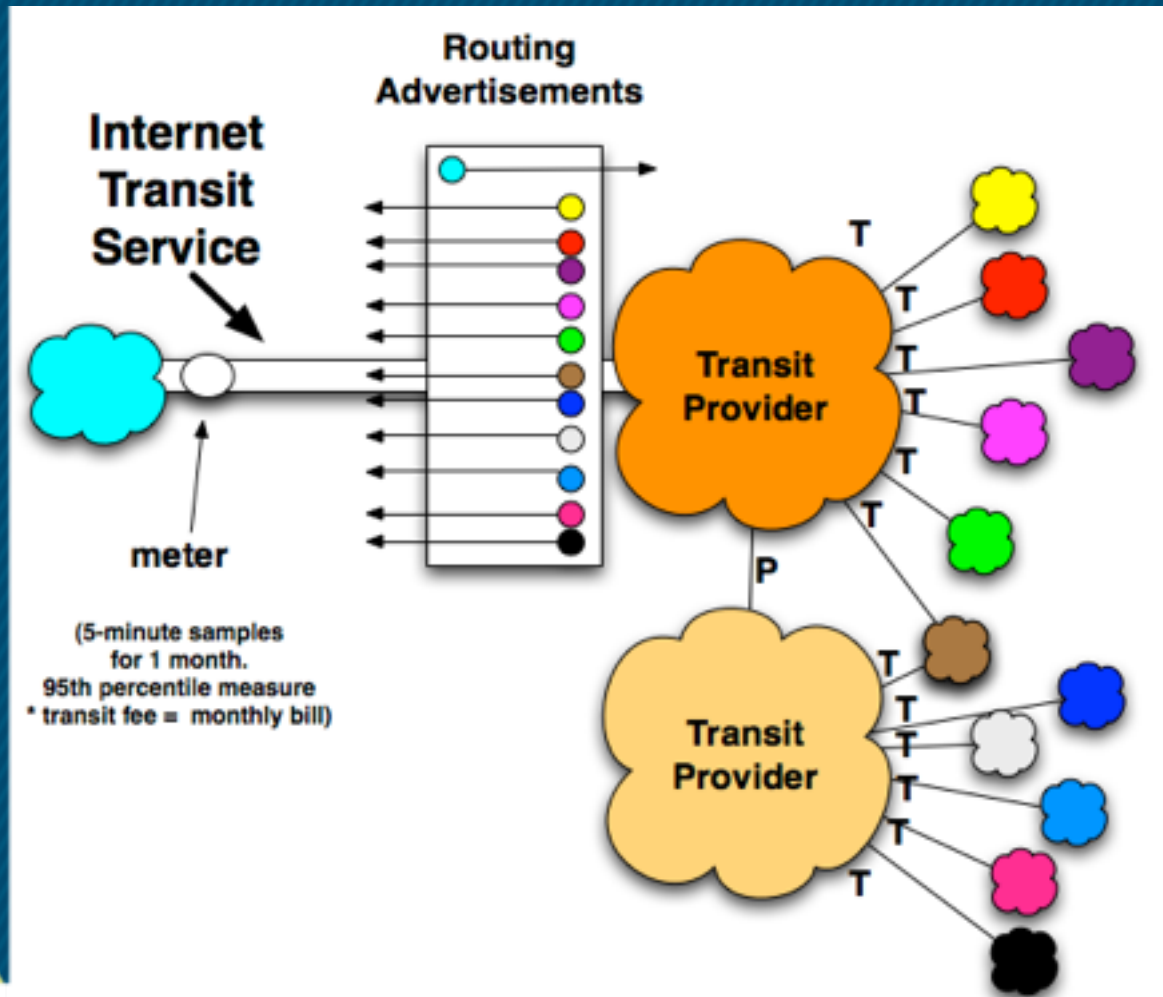


Schéma nákupu transitu



Příklad z běžného života



Vhodné přirovnání jsou Aerolinie, kde každá aerolinka představuje jednotlivého ISP (operátora)

A cestující nahrazují data přenášená internetem

Se stoupajícím provozem přestane
stačit kapacita nástupní haly



Řešením je **neutrální bod**, kde cestující
mohou samostatně přestoupit

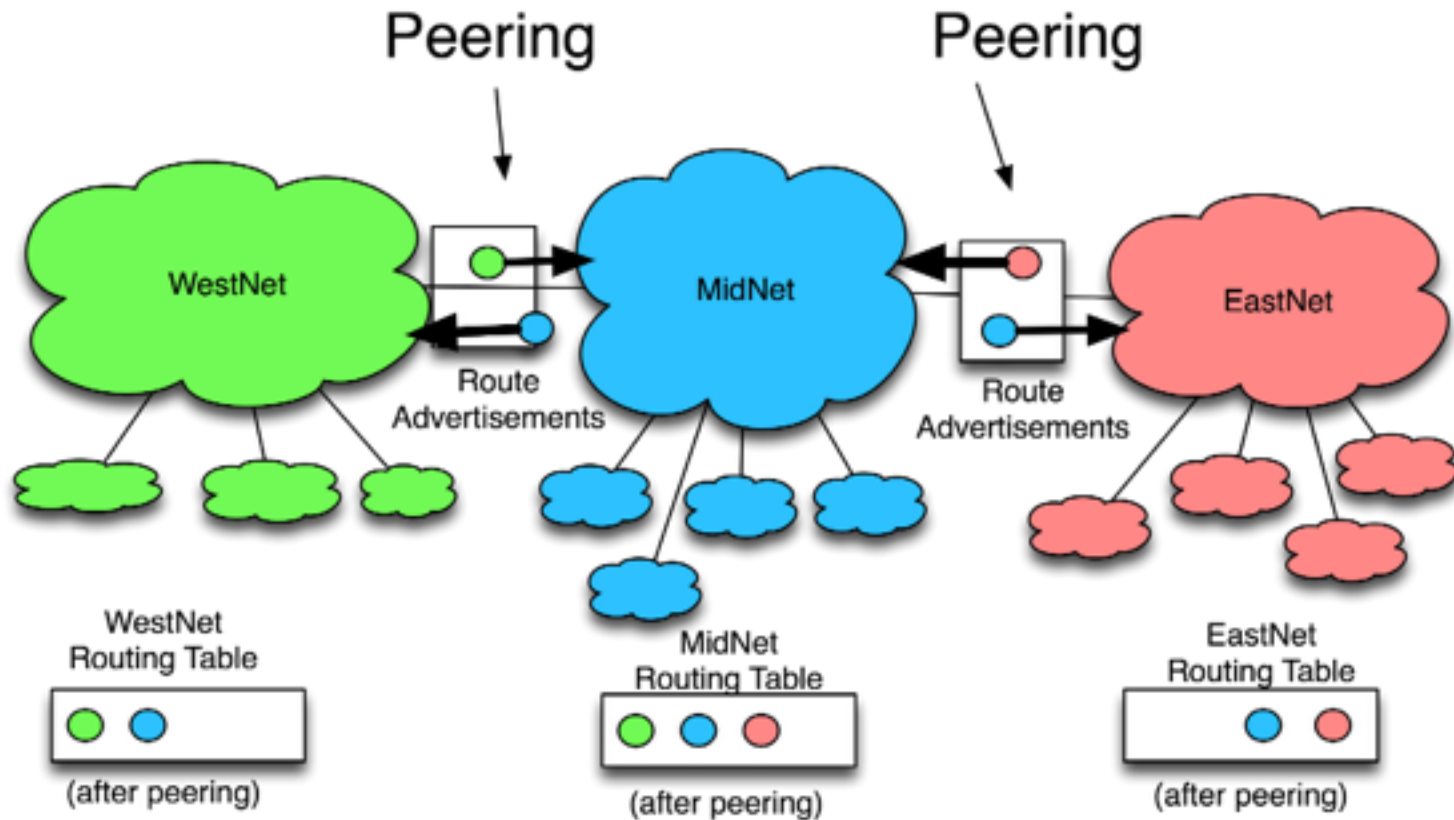


Motivace peeringu

- Redukce nákladů na transit
- Snížení latencí
- Zkrácení cest
- Zvýšení stability sítě



Peering free, paid...



Příklad z blízkého východ

[Abort / Remove](#) cr01.dub01.pccwbtn.net 84.233.221.50

traceroute ip 84.233.221.50

Tracing the route to Gi0-3.dxb-003-access-3.interoute.net (84.233.221.50)

1 ge5-0-1.var01.dub01.pccwbtn.net (63.218.176.66) 0 msec 0 msec 0 msec

2 pos4-6.cr03.ldn01.pccwbtn.net (63.218.176.38) 136 msec 136 msec 136 msec

3 TenGE11-2.br02.ldn01.pccwbtn.net (63.218.12.146) 136 msec 136 msec 136 msec

4 xe-11-1-1.lon21.ip4.tinet.net (77.67.94.153) 136 msec 136 msec 136 msec

5 xe-11-3-0.par72.ip4.tinet.net (141.136.111.246) 144 msec

xe-2-2-2.par72.ip4.tinet.net (141.136.111.250) 148 msec 144 msec

6 interoute-gw.ip4.tinet.net (77.67.75.238) 148 msec 144 msec 144 msec

7 ae1-0.mrs-001-score-1-re0.interoute.net (217.118.118.74) 156 msec 156 msec 160 msec

8 Gi0-1.mrs-boi-access-2.interoute.net (217.118.118.86) 160 msec 160 msec

Gi0-3.mrs-boi-access-2.interoute.net (217.118.118.82) 160 msec

9 so-1-0-0-0.dxb-003-access-1-re1.interoute.net (84.233.221.41) [MPLS: Label 299776 Exp 0] 260 msec 260 msec 264 msec

10 ge-0-0-0-0.dxb-003-access-2-re1.interoute.net (84.233.221.30) 260 msec 260 msec 260 msec

11 Gi0-3.dxb-003-access-3.interoute.net (84.233.221.50) 264 msec * 260 msec

Query Complete



Peering policy

- **Open**
- **Selective** - incumbent, nebo “národní” operátoři
- **Restrictive** - Využívané pro $T_2 > T_1$
- **Closed** - Typické pro T_1



Příklad “tvrdých” peeringových podmínek

- **Not** have **been** a **customer** of service for at **least 1 year**;
- have a **European footprint**, with presence in **5 countries** where NET also has presence and able to interconnect to NET in at least **3 locations** using (1, 10 or 100) GE;
- have a **non-European** footprint and able to interconnect to NET in at least **2 US locations**;
- Meet a **balanced traffic ratio** between its network and NET's network between **1:3 and 3:1** (inbound/outbound); Exchange a **minimum of 5 Gbps** sustained peak traffic with NET's network (number subject to change);
- Exchange a **maximum of 3 Gbps per location** where peering is established over a public internet exchange;
- Operate a professionally managed **24x7 NOC**
- Agree to actively cooperate to resolve security incidents and other operational problems;
- Demonstrate and enforce strict filtering policies to prevent route leaks;
- Show good faith efforts to facilitate communication regarding network maintenance with regard to the traffic exchange;
- Not abuse the peering relationship by engaging in activities such as but not limited to: pointing a default route or otherwise forwarding traffic for destinations not explicitly advertised, resetting next-hop, selling or giving next hop to others.

Privátní propoje vs Peeringové uzly

- Náklady na propoj
 - #portu = #peerů
 - Nutnost alternativní řešení
- Náklady na propoj
 - Náklady na porty IXP
 - Redundanci zajišťuje IXP

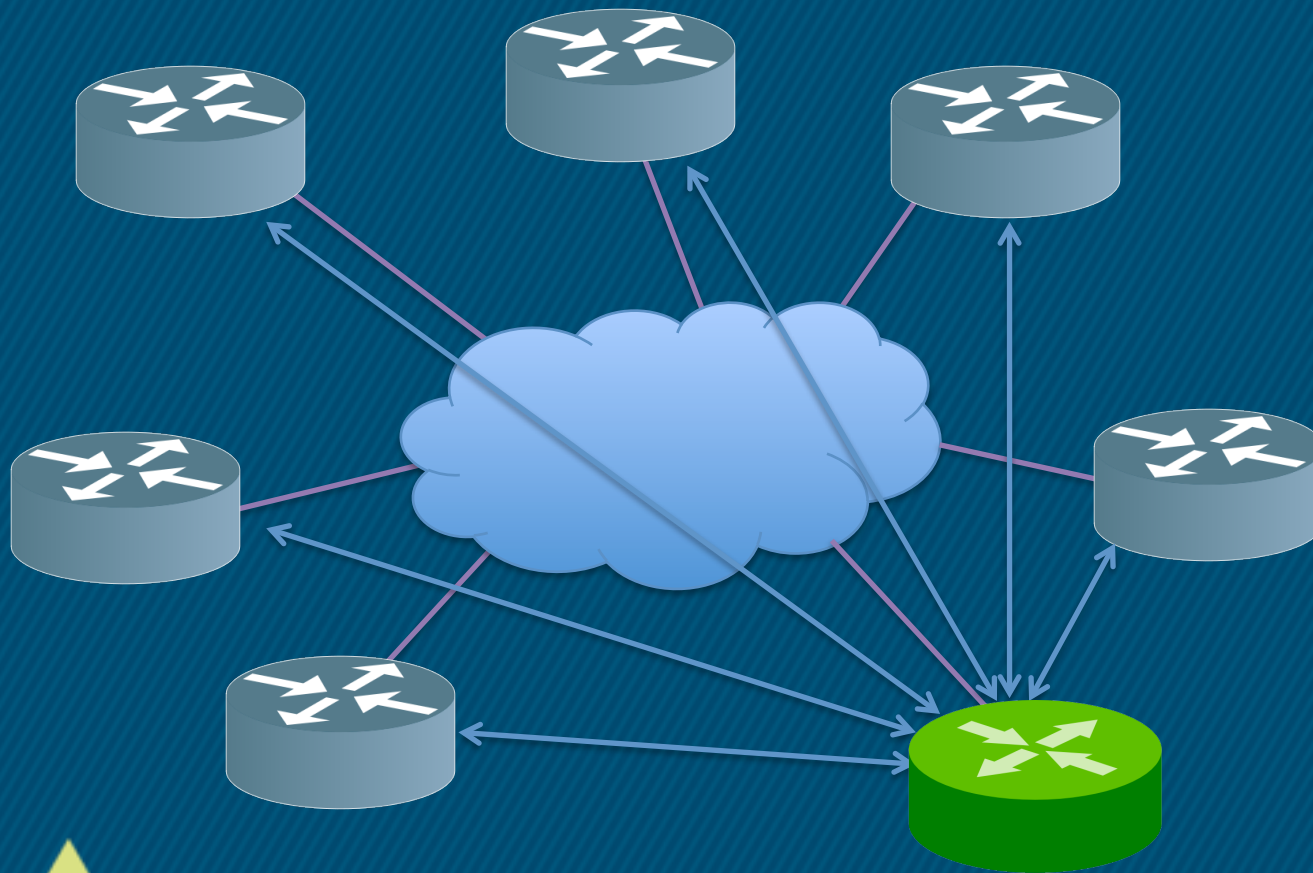


Peeringové uzly

- komerční
 - většinou navázané na další služby které prodává
 - tlak na oběh služeb provozovatele
- nekomerční (sdružení)
 - členové mají možnost určovat směrování IXP

Ideální stav

(z pohledu "zelené" sítě)

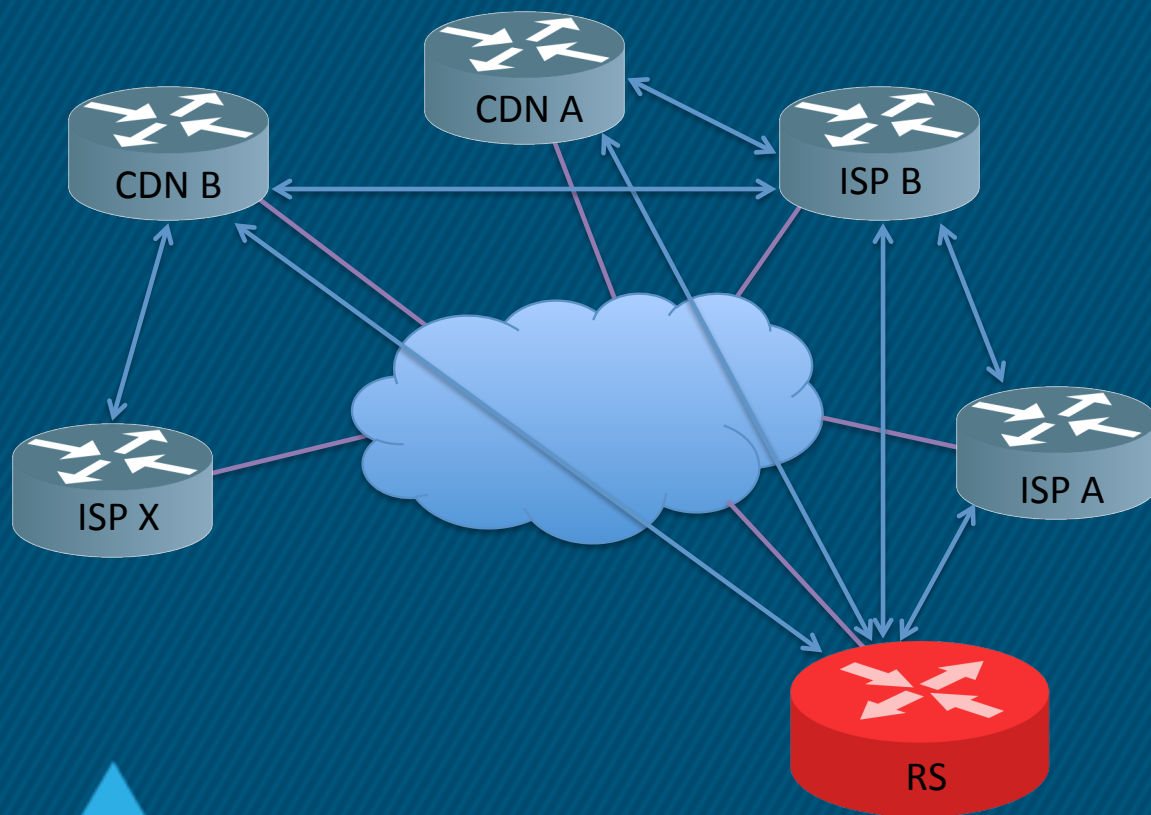


Route servery v peeringových uzlech

- Zjednodušení vstupu nových sítí do IXP
- Snížení zátěže pro router (jedna BGP session namísto mnoha desítek...)



Příklad propojení jednotlivých sítí v IXP

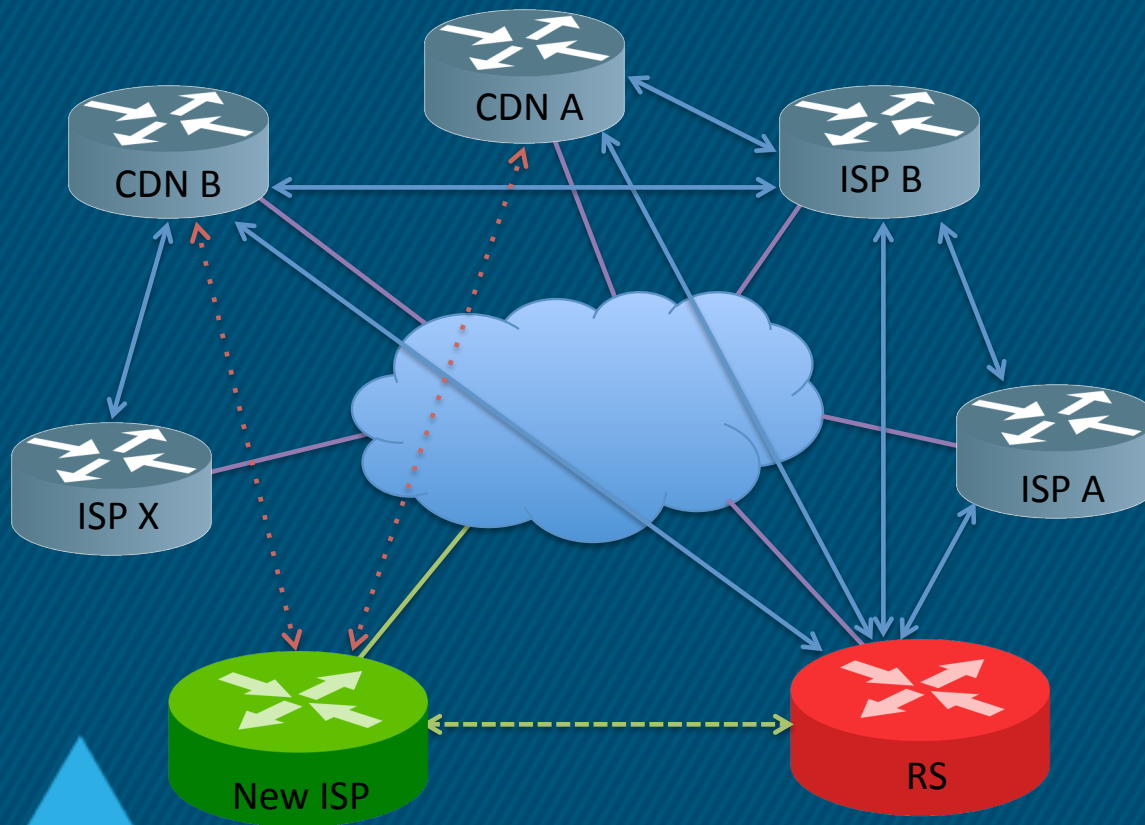


ISP A peeruje s ISP B a přes RS
CDN A a B

ISP B peeruje s ISP A CDN A a B

ISP X peeruje jen s CDN B

Nově připojená síť - navazování peeringu



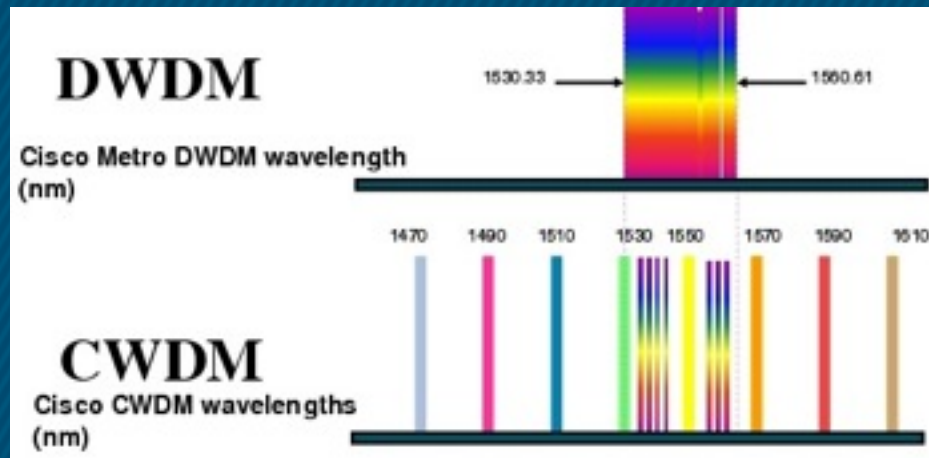
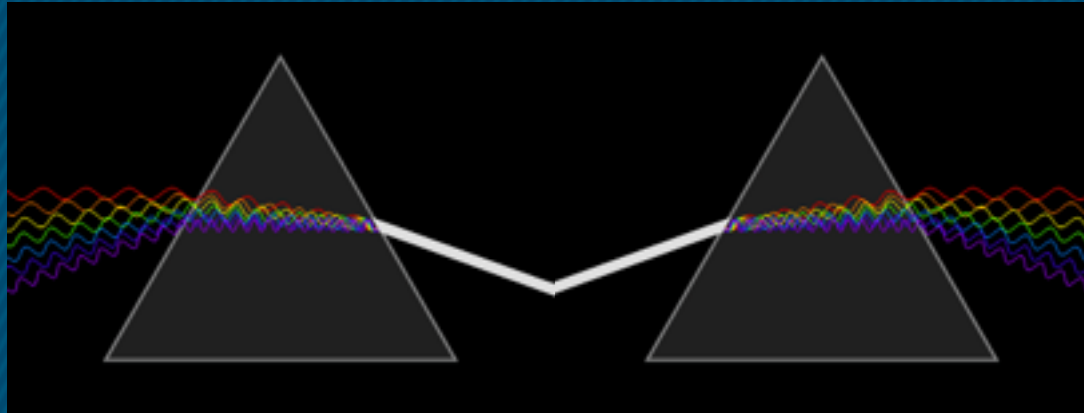
Nová síť může pro zjednodušení a zrychlení konfigurace navázat peering s ostatními přes RS

Vlnové multiplexy WDM

- Zvýšení kapacity FO
- Možnost 1vl. propojů
- Možnost budovat “alternativní” topologie



DWDM vs CWDM



Co je NIX.CZ?

- peeringové centrum
- více než 100 členů / zákazníků
- 5x telehouse v Praze
- datový tok více než 285 Gb/s



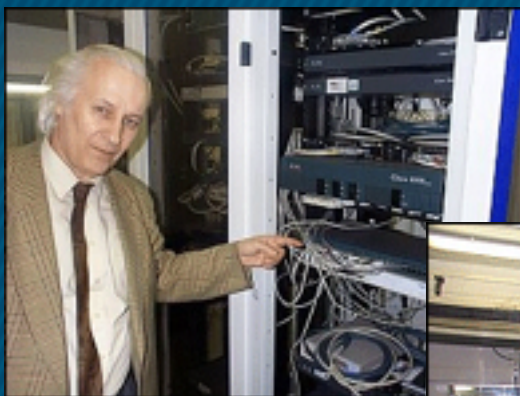
Jak to vše začalo?

- 30.8.1996 podepsalo 7 zástupců zakladatelskou smlouvu
- 1.10.1996
- 17.10.1996 proběhla 1.VH



1. Rok Českého peeringu

- Únor 1997 faktické zahajení peeringu



Éra 1999-2004

- 2000
 - 20 členů
 - 2x PoPy propojené 2x1GE
 - Kapacity přípojek 2-100Mbit
- 2002
 - 1. Zaměstnanec – Tomáš Maršálek
 - Překonání hranice 1Gbps
- 2003
 - Implementace IPv6
 - 2 nové PoPy, celkem 4
 - Vstup do Euro-IX

Éra 2005-2008

- 2005
 - Podpora 10GE
 - 5Gbps celkového trafficu
 - 45 členů
 - Na konci roku 10Gbps, nárůst +100%!
- 2006
 - 1. housované TLD v NIX.CZ (.eu, .be, .at)
 - Překonání hranice 20Gbps
- 2008
 - Překonání 50Gbps
 - Zahájení provozu Route Serverů
 - Vše na HW C6509E

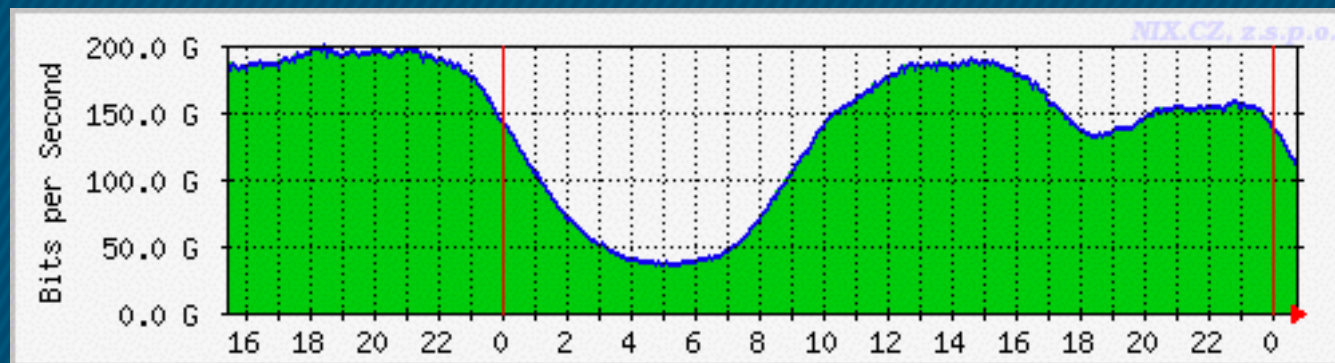
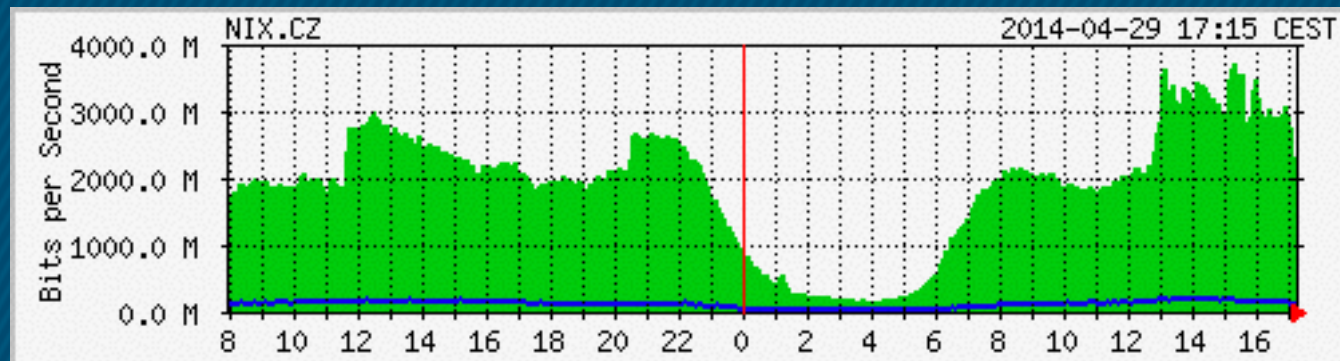
C6509E s příslušenstvím



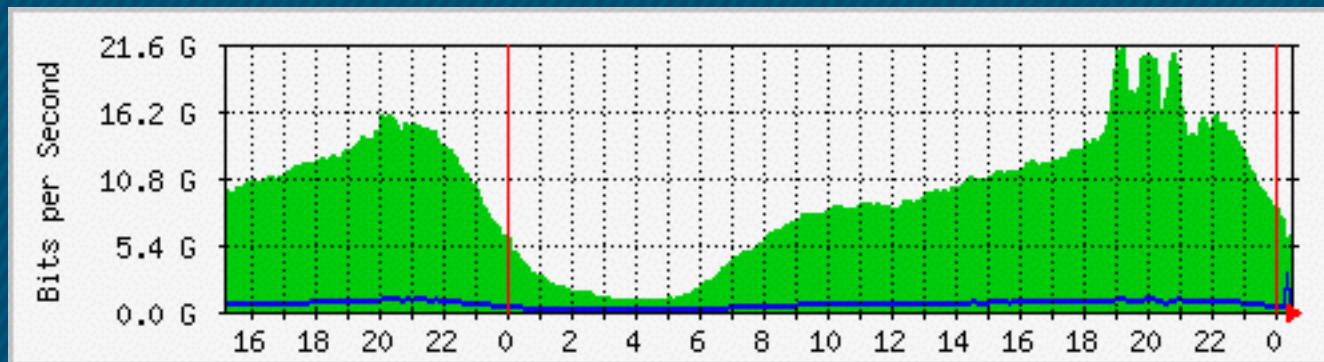
Éra 2009-2011

- 2005
 - Podpora 10GE
 - 5Gbps celkového trafficu
 - 45 členů
 - Na konci roku 10Gbps, nárůst +100%!
- 2006
 - 1. housované TLD v NIX.CZ (.eu, .be, .at)
 - Překonání hranice 20Gbps
- 2008
 - Překonání 50Gbps
 - Zahájení provozu Route Serverů
 - Vše na HW C6509E

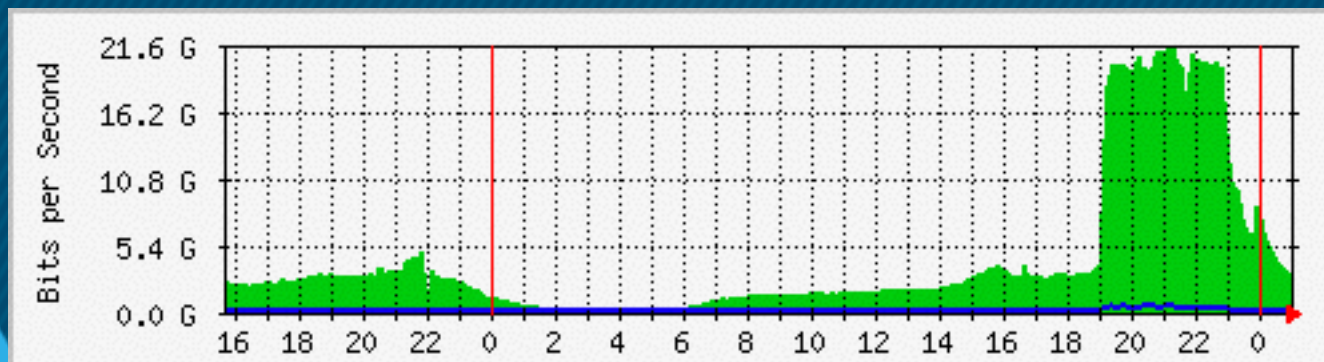
Statistiky



Statistiky



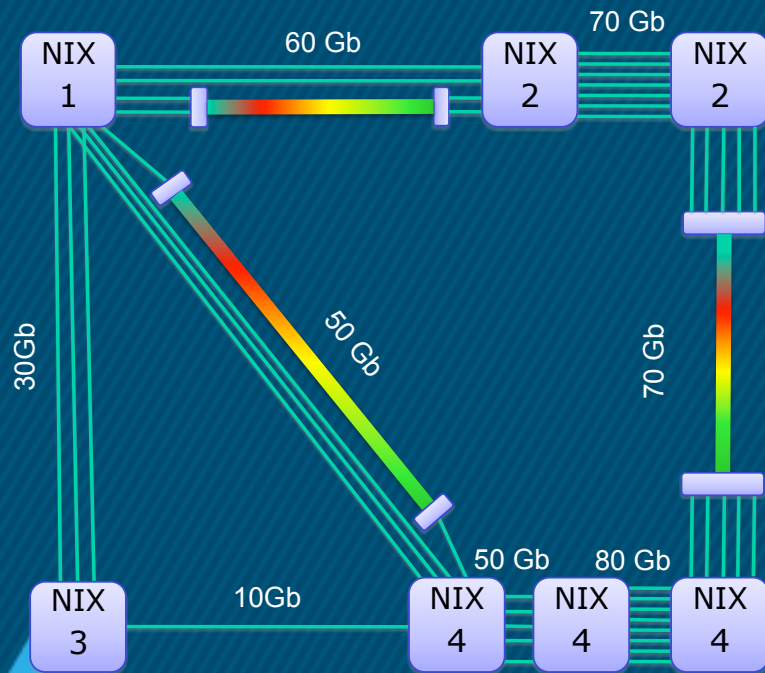
10.6.2013



18.10.2013

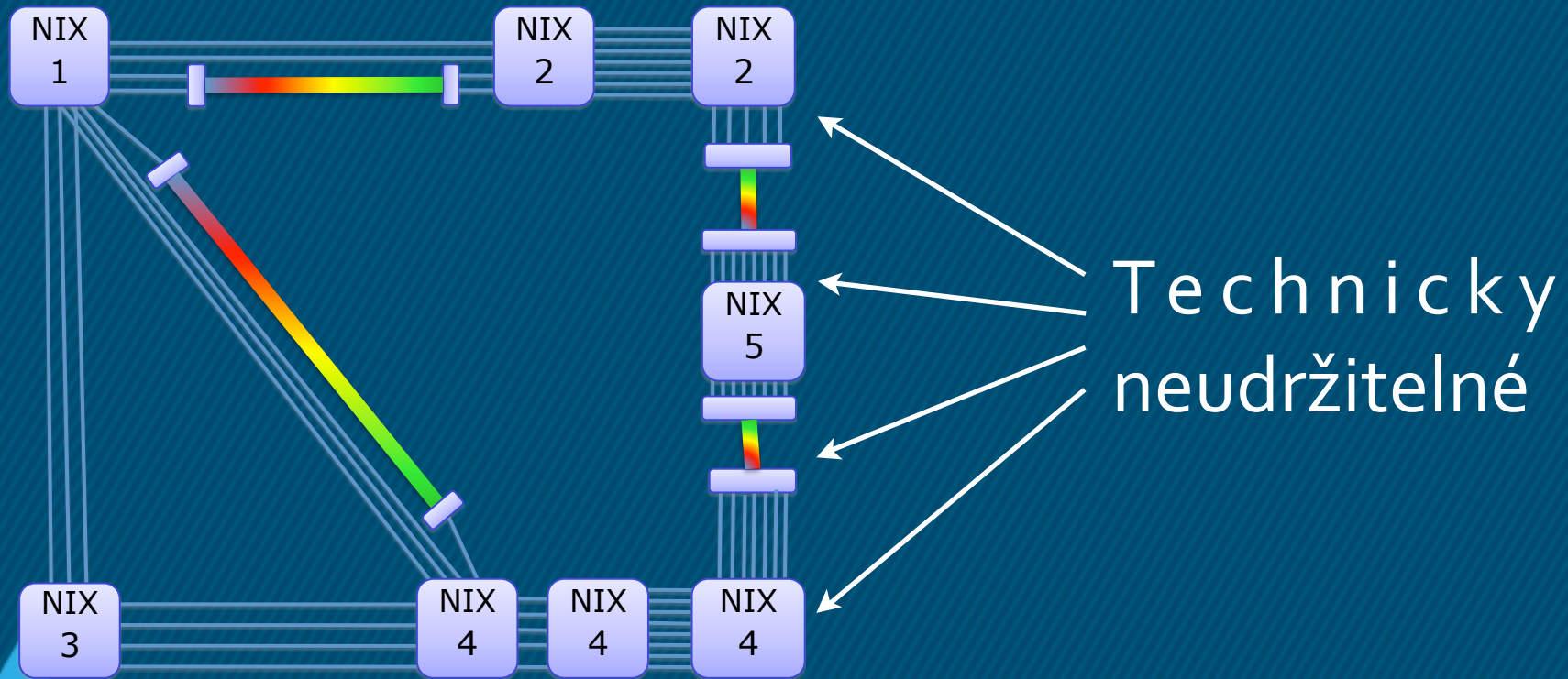
Migrace z kruhové topologie

Výchozí topologie NIX.CZ
(začátek 2010)

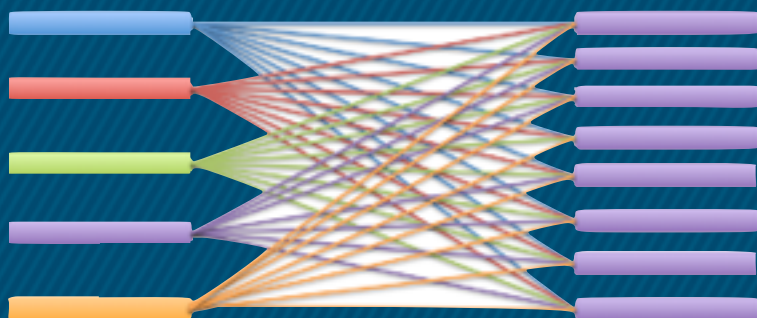


Kruhová topologie
vznikla postupným
připojováním nových
lokalit

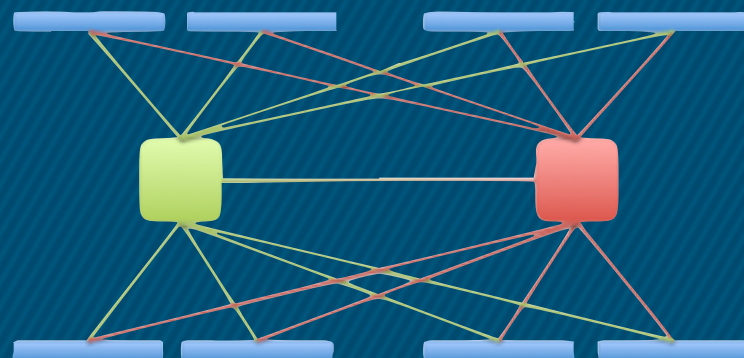
Kruhová topologie s další lokalitou



Možné varianty migrace

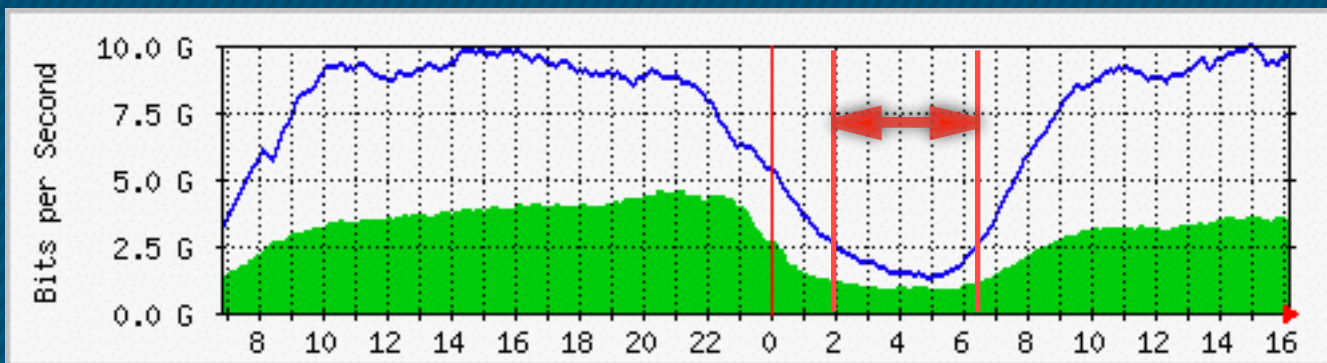


vs



Požadavky na přestavbu topologie

- Minimální výpadek
- Všechny změny provádět mimo “provozní špičky” => 2:00 - 6:30

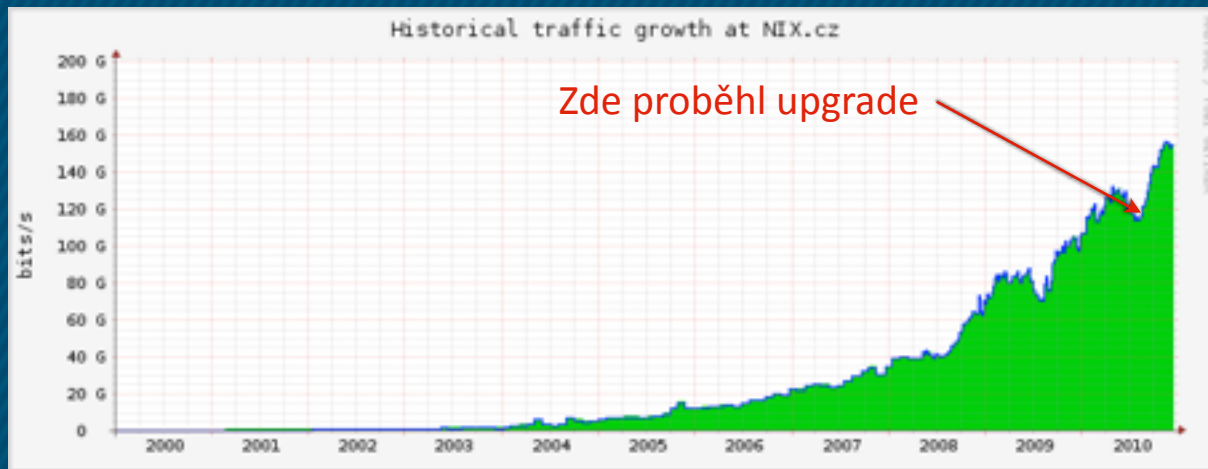


- V případě problému ponechat možnost “roll-backu”

Časování nákupu a upgrade

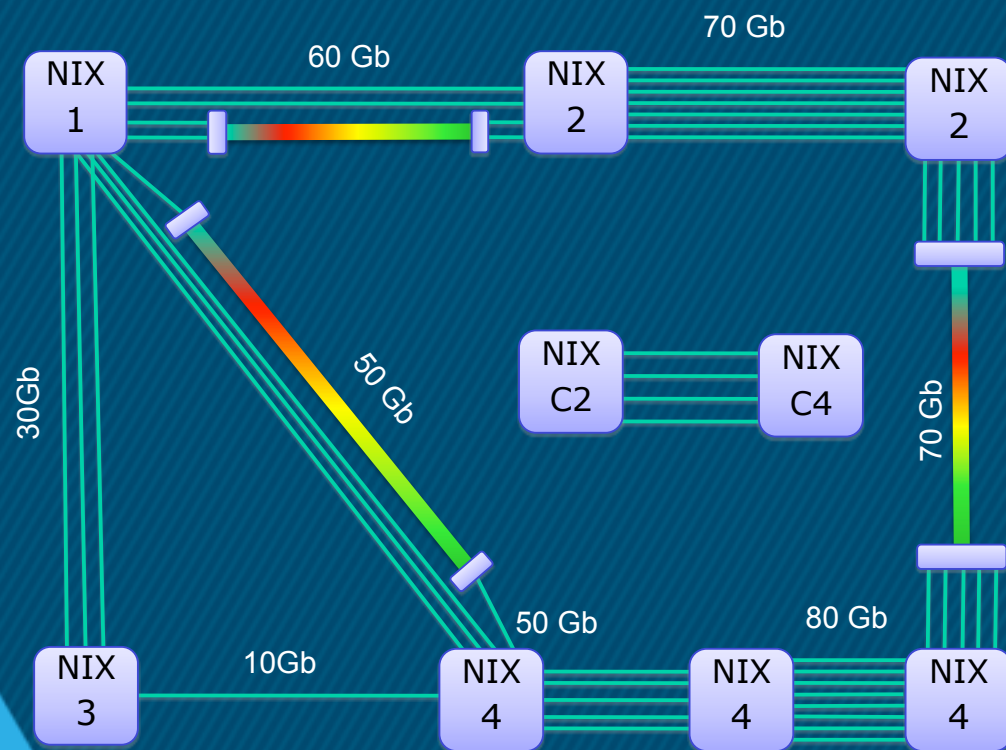


čas objednávky



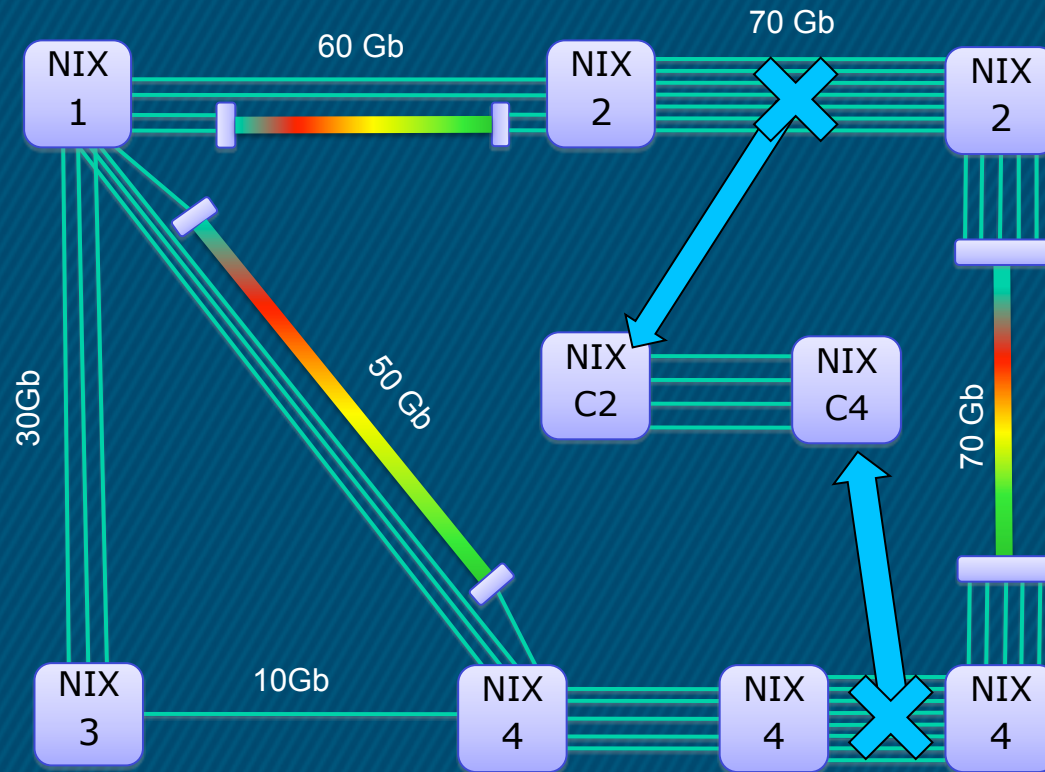
Zde proběhl upgrade

Původní topologie NIX.CZ

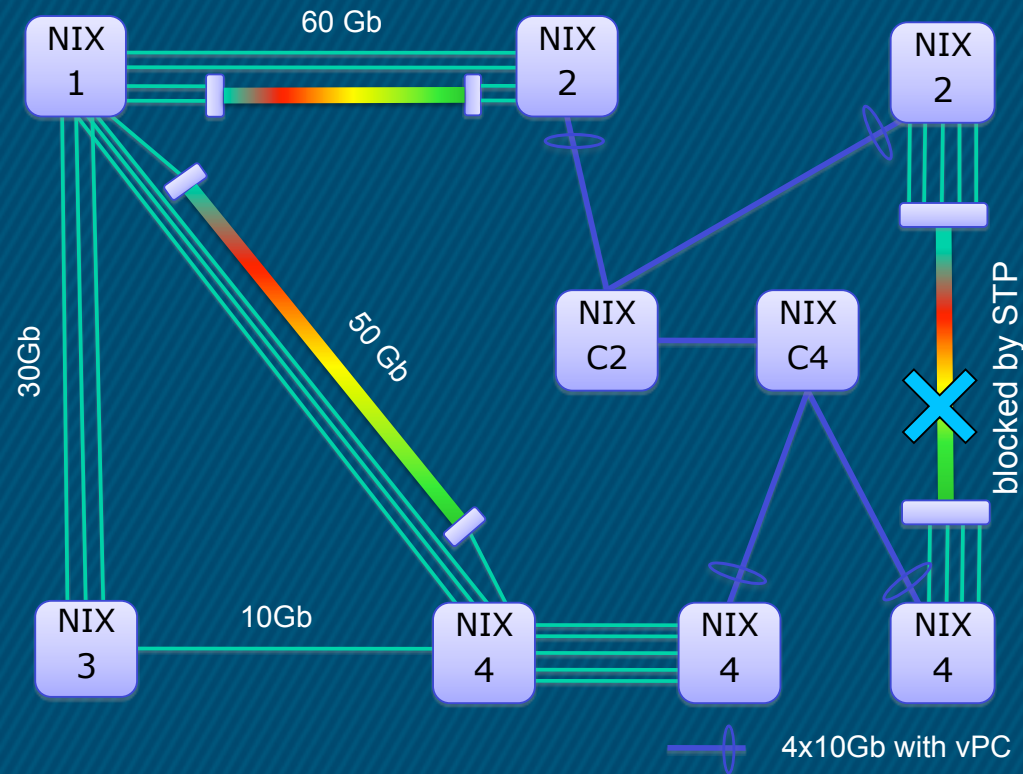


V této topologii
běžel NIX.CZ
do 26.8.2010

Migrace ring 2 star (1. krok)

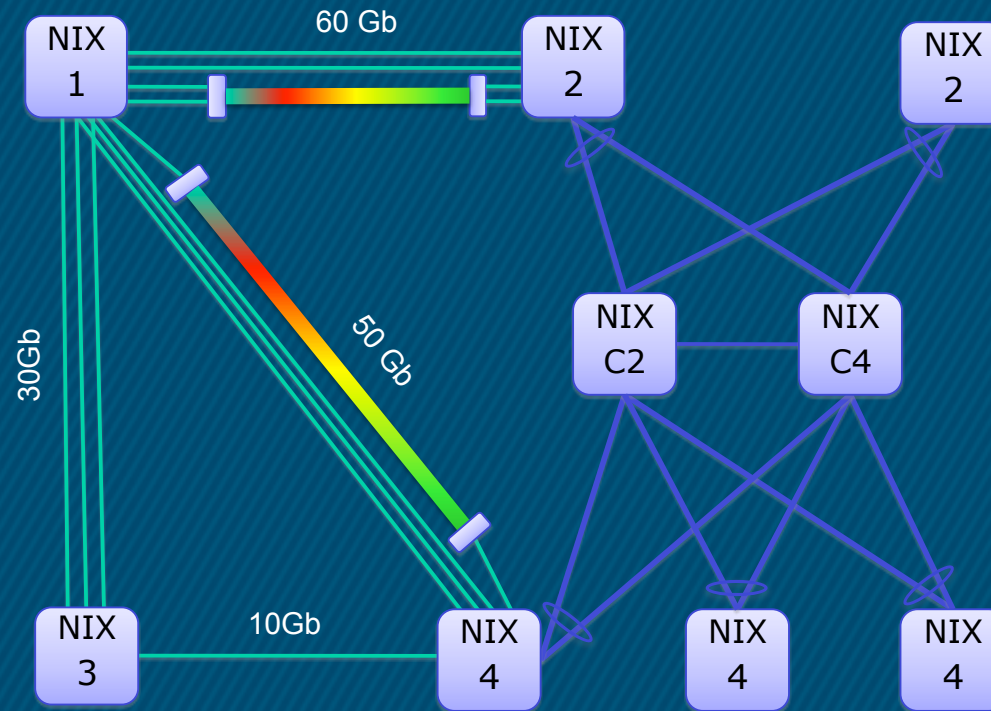


Migrace ring 2 star (průběh 1. kroku)



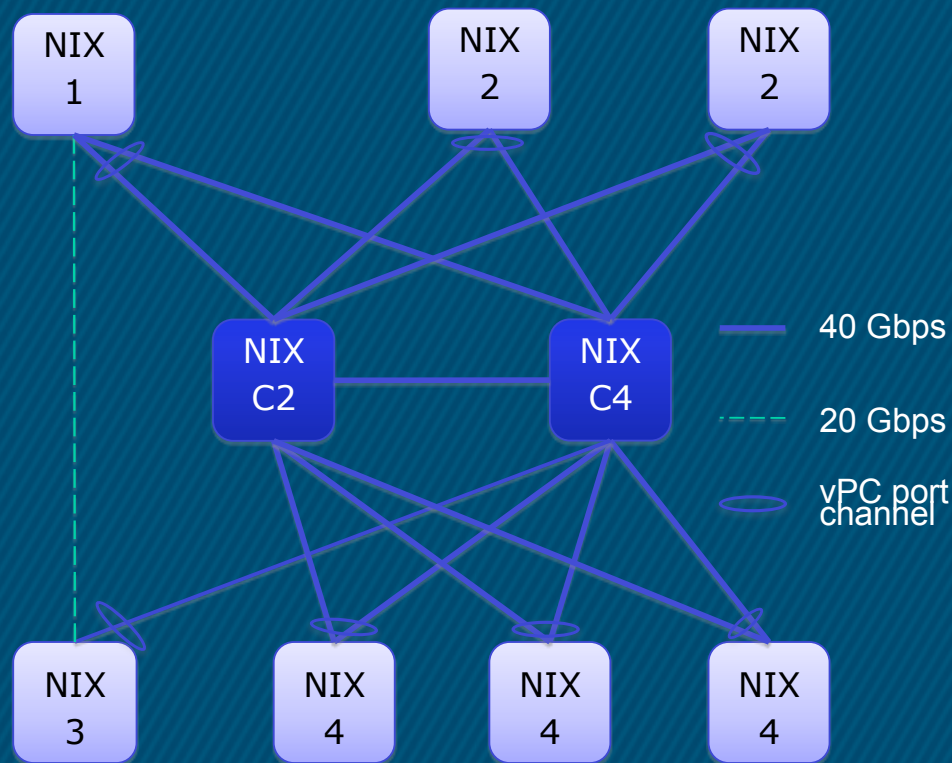
Migrace ring > star

(dokončený 1. krok)



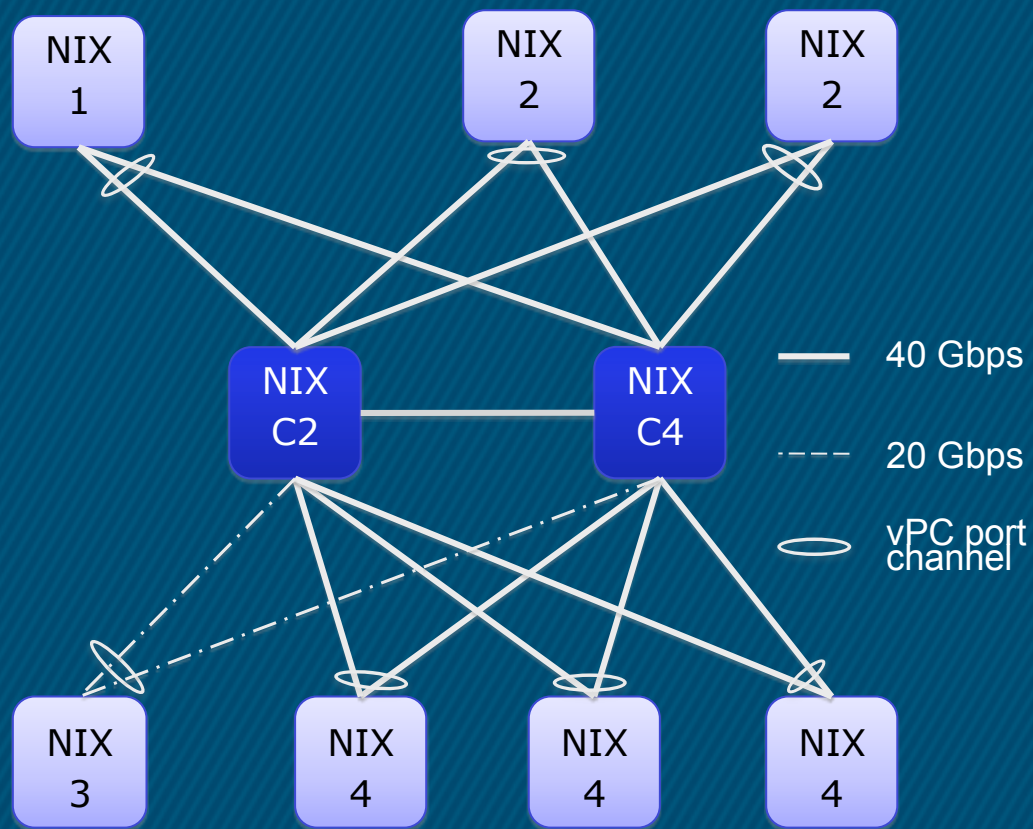
Migrace ring > star

(2. krok - 30.8.2010)



Cílový stav

(poslední krok v průběhu listopadu '10)

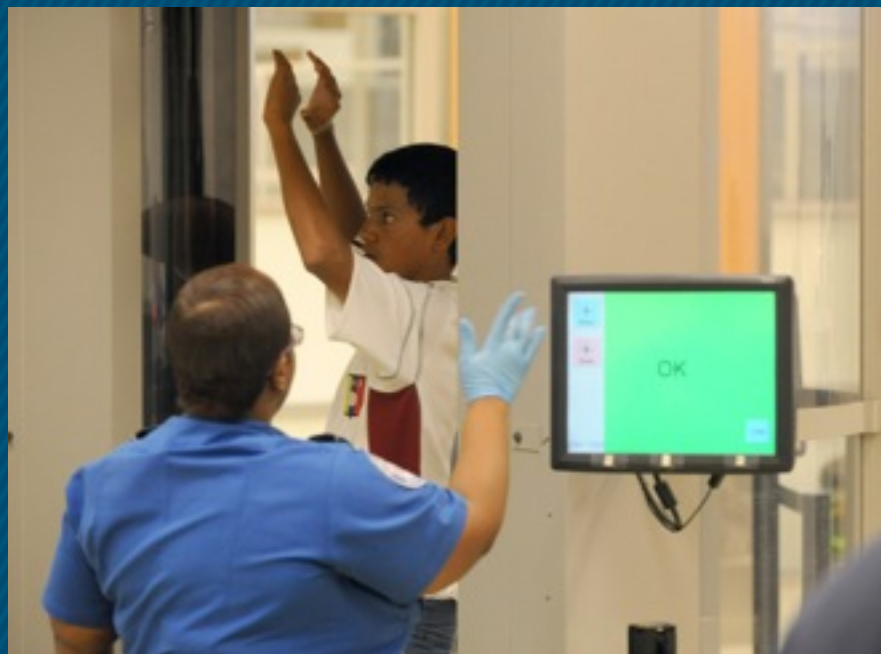


Bezpečná VLAN

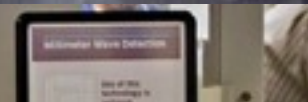
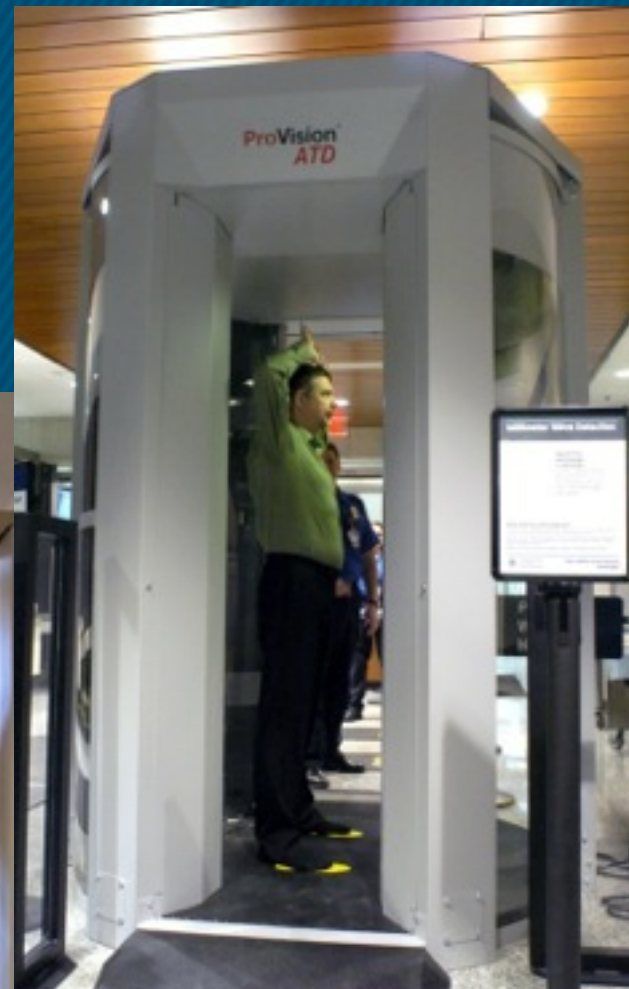
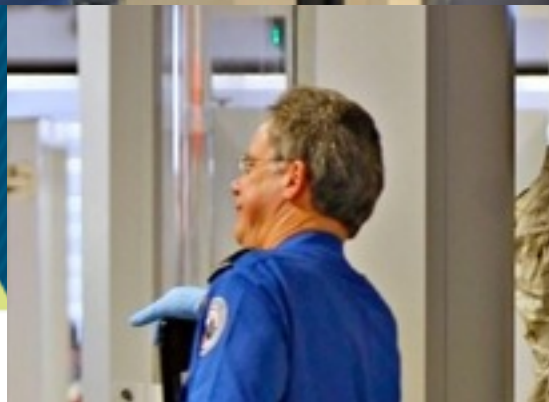
- Reakce na útoky v 2013
 - cíle: banky, on-line media, mob. op., seznam.cz
- Útok procházel NIX.CZ a z upstreamu firem (pravděpodobně z RF)
- Některé cíle/oběti použili “ostrovní režim”
- Dopad i na regionalní provoz



Bezpečná VLAN



ProVision



Bezpečná VLAN

- “Klub důvěryhodných” společností
- Připojky do Bezp VLAN i do peering VLAN
- Bezpečná VLAN — poslení instance
- “Ať se aspoň můžou lidi v CZ dostat do CZ bank”
- Vysoké vstupní požadavky



Bezpečná VLAN

technické podmínky

- BCP38/SAC004
 - RTBH filtering přes RS
 - Důraz na kvalitu a nové technologie:
 - DNSSec
 - IPv6
 - Redundance připojek
 - Monitoring
 - Control plane policy - RFC6192
 - DNS (UDP) amplification protection
- 

Bezpečná VLAN

organizační podmínky

- Smlouvy se zákazníky - spam, útoky etc
- Kontakt pro bezpečnostní incidenty 24x7 - ne IVR
- CSIRT team
- V NIX.CZ aspoň 6měsíců
- Dobrá pověst
- Prohlášení, doporučení, nevetovaný



Bezpečná VLAN

stav

- 6 zakládajících společností
- ACTIVE 24, CESNET (NREN), CZ.NIC, Dial Telecom, Seznam.cz and Telefónica Czech Republic
- Vlastní technická pracovní skupina
- Testování BFD
- Příprava RBTH



Dotazy a připomínky

ag@nix.cz

