



Motivace

Idea

Budoucnost

Demo

Vladimír VESELÝ
ivesely@fit.vutbr.cz



Nová směrovací architektura



AKTUÁLNÍ STAV

Motivace

Idea

Budoucnost

Demo

- ◆ Jeden plochý číselný systém
- ◆ IP adresa zastává dvě funkce
 - ◆ Identifikuje zařízení
 - ◆ Lokalizuje jeho polohu v síti
 - ◆ Adresování odráží topologii
- ◆ Před 20 lety se neřešilo/nebyly problémem jako dnes:
 - ◆ Mobilita
 - ◆ Multihoming
 - ◆ Růst **default-free zone (DFZ)** směrovací tabulky
 - ◆ Neagregovatelné štěpení prostoru IP adres



PROBLÉMY

◆ Multihoming

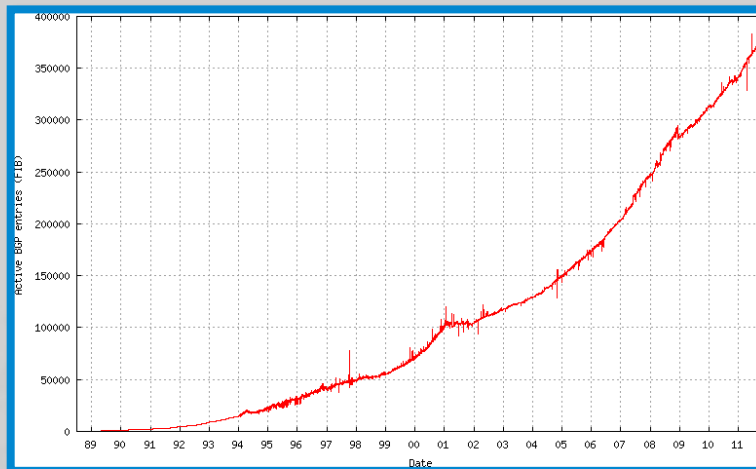
- ◆ *Co když má zařízení 2 a více síťových karet?*
- ◆ ...pak IP adresa = bod připojení do sítě!

◆ Mobilita

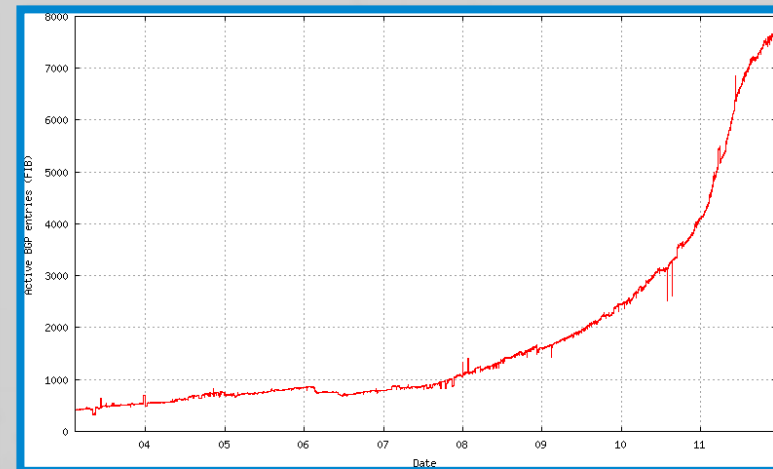
- ◆ *Co když přecházíme od jednoho providera k druhému?*
- ◆ ...pak musíme většinou přečíslovat!

◆ Růst DFZ

- ◆ Rozpor mezi best-effort delivery modelem a množstvím informací v DFZ



IPv4



IPv6

Motivace

Idea

Budoucnost

Demo



Motivace

Idea

Budoucnost

Demo

CÍLE

- ◆ Omezení růstu DFZ směrovacích tabulek
- ◆ Zamezení vzniku neagregovatelných bloků adres
- ◆ Provider independent adresní plán
- ◆ Multihoming bez BGP
- ◆ Mobilita
- ◆ Traffic engineering na vstupu



IDEA

Motivace

Idea

Budoucnost

Demo

Oddělme od sebe dvojí funkcionality IP adresy!

◆ Locator/ID Separation Protocol (LISP)

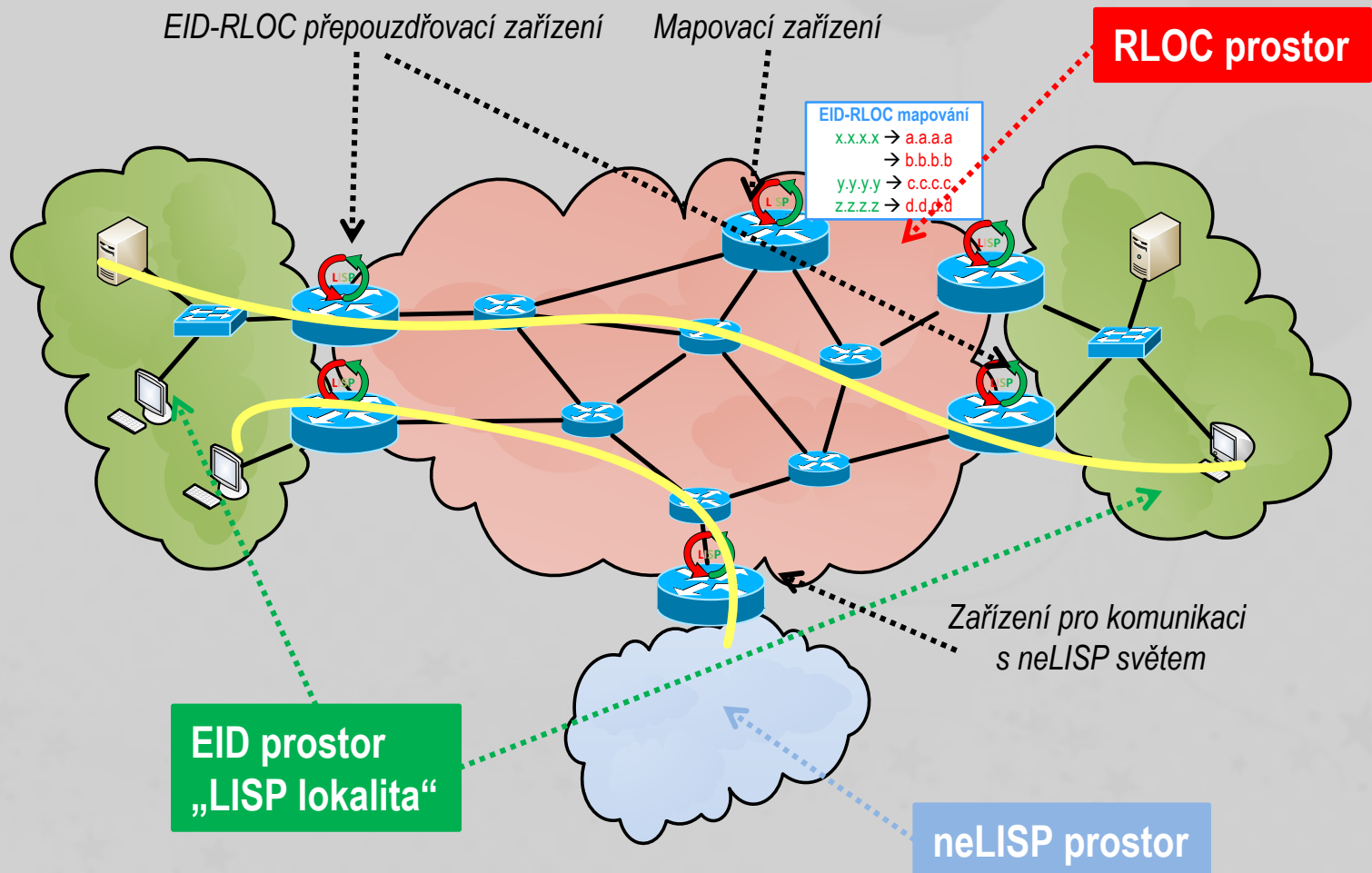
- ◆ Identifikaci zajišťuje **Endpoint Identifier (EID)** adresa
- ◆ Připojení do sítě – lokalizaci – pak **Routing Locator (RLOC)** adresa

Motivace

Idea

Budoucnost

Demo





Motivace

Idea

Budoucnost

Demo

VÝHODY

- ◆ Bez změn konfigurace koncových stanic
- ◆ Bez změn v DNS
- ◆ Nezávislé na použitém adresním prostoru
- ◆ Designová podpora mobility
- ◆ Designové přechodové mechanismy



ZAPOUZDŘOVÁNÍ

- ◆ LISP hlavička (36 B pro IPv4 nebo 56 B pro IPv6)
- ◆ UDP hlavička
 - ◆ Cílový port 4341 je určen pro datový provoz
 - ◆ Cílový port 4342 je určen pro kontrolní provoz
- ◆ Vnější mapovaná IP či jiná síťová hlavička

Motivace

Idea

Budoucnost

Demo

IPv4 Outer Header:
Router supplies
RLOCs

UDP:

**LISP
Header:**

IPv4 Inner Header:
Host supplies
EIDs

0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1
Version		IHL		Type of Service				Total Length																							
Identification										Flags		Fragment Offset																			
Time to Live				Protocol (17)				Header Checksum																							
Source Routing Locator																															
Destination Routing Locator																															
Source Port (xxxx)																Dest Port (4341)															
UDP Length																UDP Checksum															
N		L		E		V		I		Flags		Nonce/Map-Version																			
Instance ID/Locator Status Bits																															
Version		IHL		Type of Service				Total Length																							
Identification										Flags		Fragment Offset																			
Time to Live				Protocol				Header Checksum																							
Source EID																															
Destination EID																															

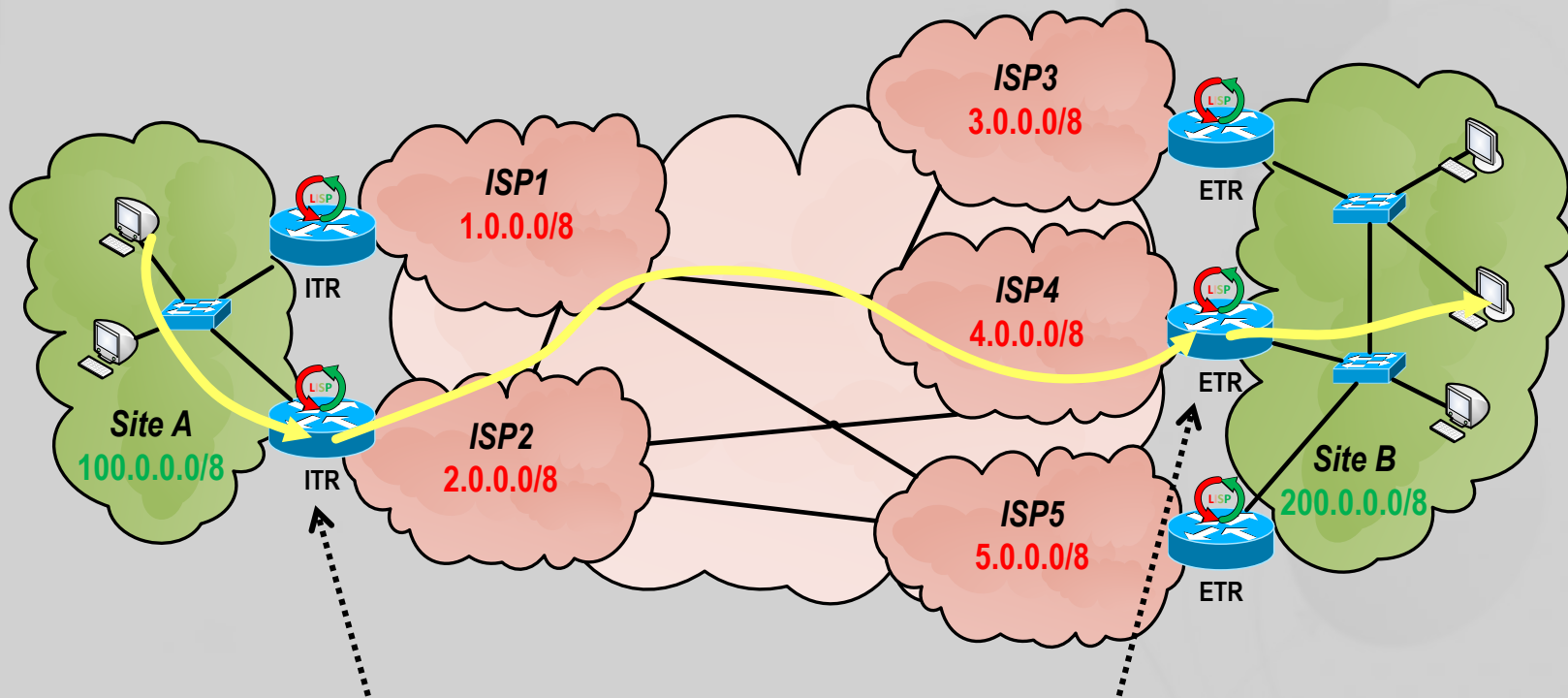
KOMPONENTY: ITR A ETR

Motivace

Idea

Budoucnost

Demo



Ingress Tunnel Router (ITR)

- ◆ Vstupní bod do RLOC prostoru
- ◆ Zapouzdřuje původní provoz LISPem
- ◆ Uchovává si mapovací cache

Egress Tunnel Router (ETR)

- ◆ Výstupní bod z RLOC prostoru
- ◆ Odpouzdřuje LISP a doručuje původní traffic EID hostům

- ◆ Duální funkcionality ITR a ETR



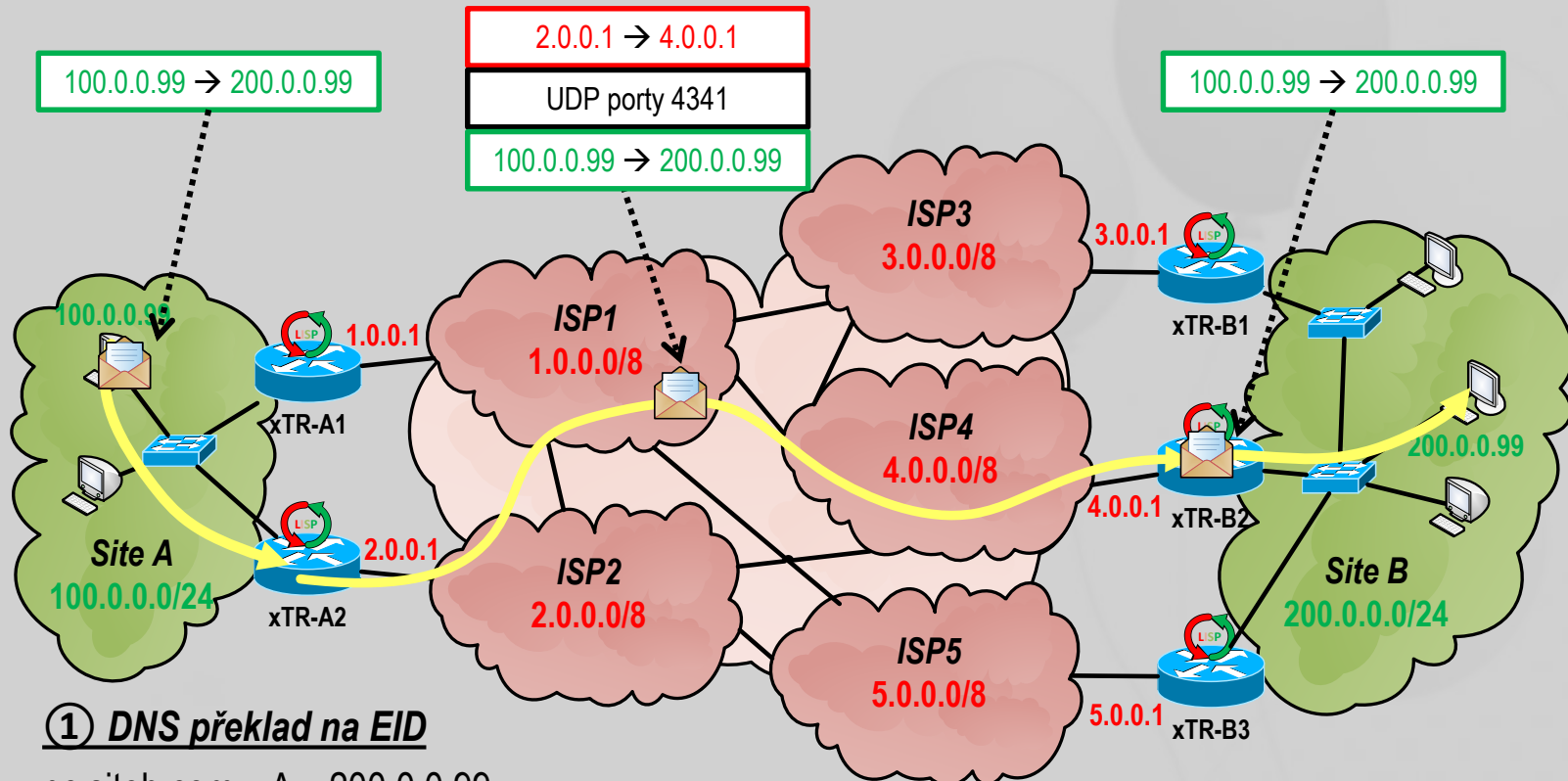
UKÁZKA: UNICASTOVÉ SMĚROVÁNÍ

Motivace

Idea

Budoucnost

Demo



① DNS překlad na EID

pc.siteb.com A 200.0.0.99

② Vyhledání v map-cache

EID-prefix: 200.0.0.0/24

RLOCy:

3.0.0.1, priorita 254, váha 50 = xTR-B1

4.0.0.1, priorita 1, váha 100 = xTR-B2

5.0.0.1, priorita 254, váha 50 = xTR-B3

③ Zapouzdření na ITR

Jako cílová IP vybrán

RLOC xTR-B2 kvůli
nejnižší prioritě

④ Odpouzdření na ETR



MAPOVÁNÍ

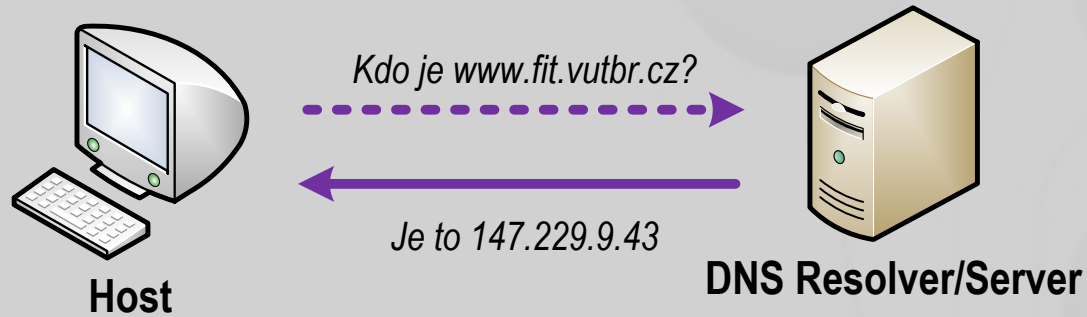
- ◆ Jak zjistit RLOC k EID?
- ◆ Analogie k DNS

Motivace

Idea

Budoucnost

Demo



- ◆ V případě LISPu





MAPOVÁNÍ

- ◆ Zprávy
 - ◆ *Jak dát vědět ostatním o EID v mé LISP oblasti?*
 - ◆ **Map-Register**
 - ◆ *Jak se zeptám na konkrétní mapování?*
 - ◆ **Map-Request**
 - ◆ *Jak je mi na můj dotaz odpovězeno?*
 - ◆ **Map-Reply**
 - ◆ **Negative Map-Reply**
- ◆ Distribuovaná hierarchická databáze

Motivace

Idea

Budoucnost

Demo

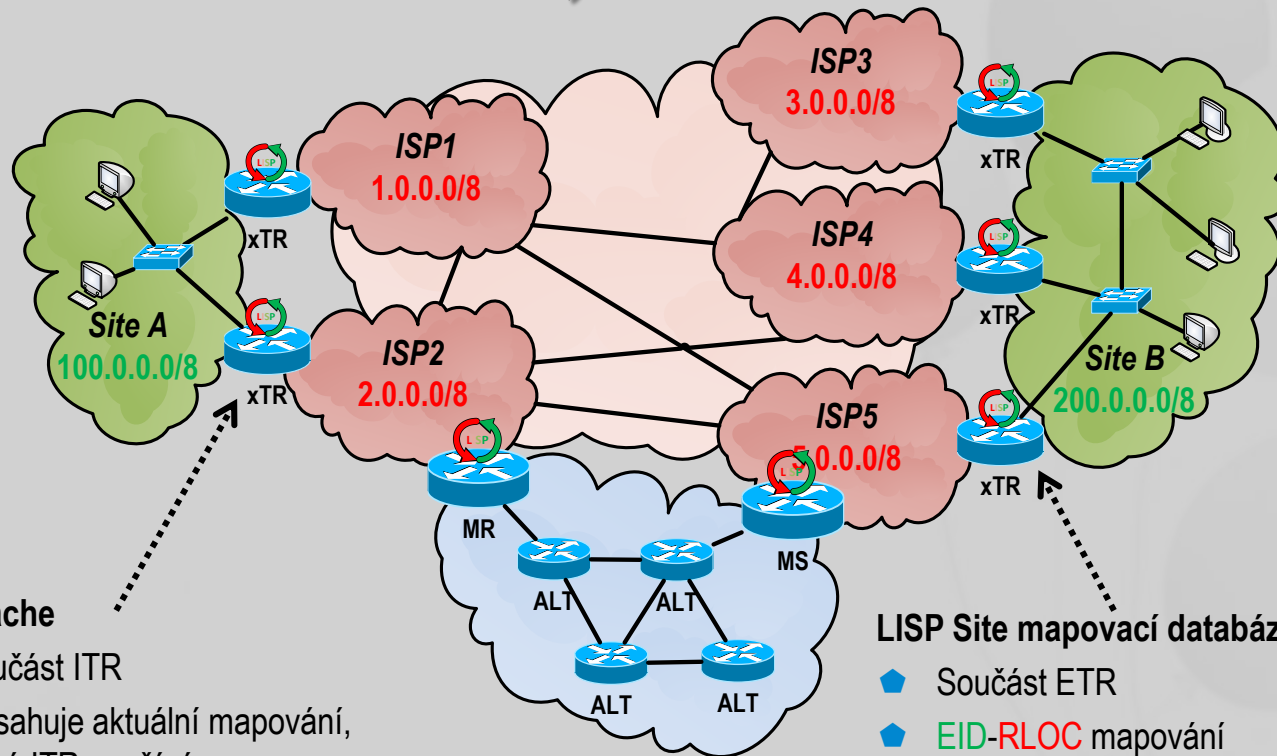
KOMPONENTY: MS, MR A ALT

Motivace

Idea

Budoucnost

Demo



Map Cache

- Součást ITR
- Obsahuje aktuální mapování, které ITR používá
- Aktualizována **Map-Reply** zprávami od MS či ETR
- Musí se podřídít informacím v **Map-Reply**, tj. prioritě, váhám, stavu **RLOCů** a TTL)

LISP Site mapovací databáze

- Součást ETR
- **EID-RLOC** mapování
- ETR autoritativně zodpovídá za **EID**
- Zasílá **Map-Reply** odpovědi ITR
- Na základě informací v **Map-Request**, může měnit přístupové politiky

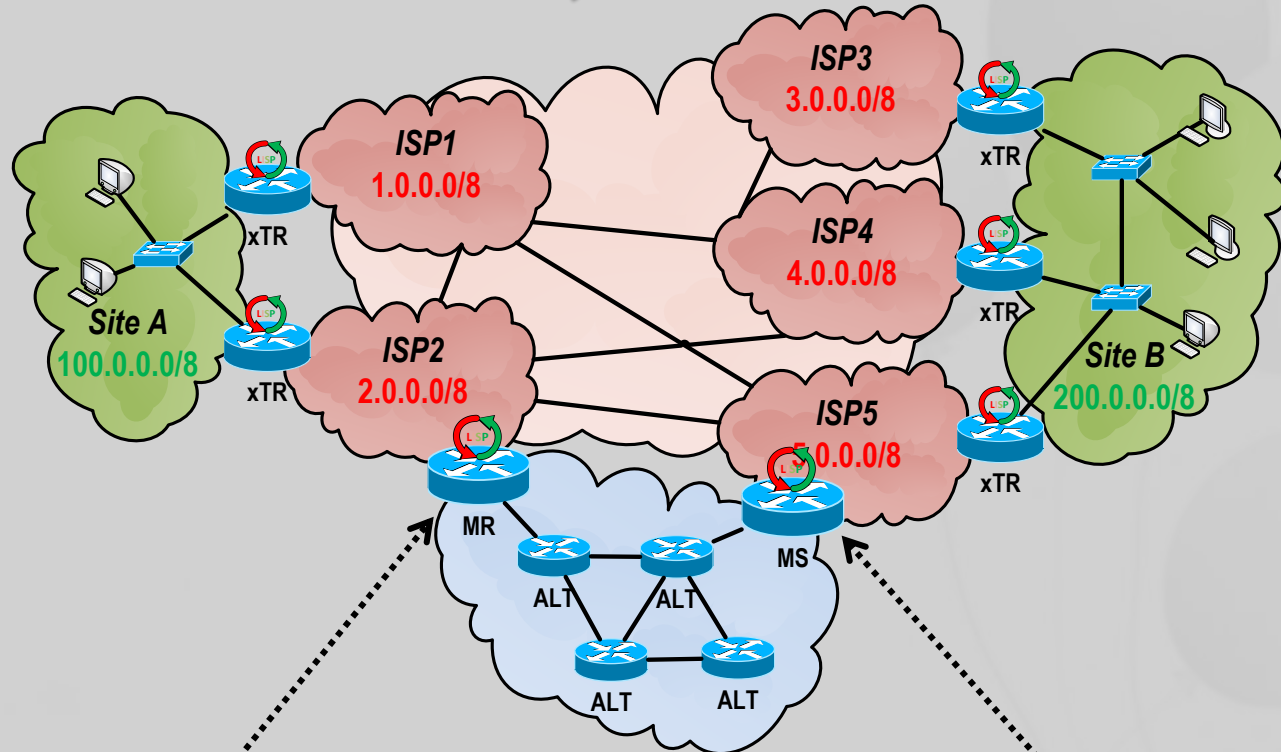
KOMPONENTY: MS, MR A ALT

Motivace

Idea

Budoucnost

Demo



Map-Resolver (MR)

- Přijímá **Map-Requesty** od ITR
- Přeposílá je do ALT
- Odpovídá pomocí **Negative Map-Reply** v případě, že je dotazováno neLISP oblast

Alternate Topology (ALT)

- Agreguje **EID** prefix a udržuje alokační politiku
- **EID** jsou oznamovány skrz BGP v GRE tunelech
- Přeposílá **Map-Requesty** s dotazem k RLOCu na EID

Map-Server (MS)

- **EID-RLOC** mapování na všech ETR
- ETR autoritativně zodpovídá za EID
- Zasílá **Map-Reply** odpovědi ITR
- Na základě informací v **Map-Request**, může měnit přístupové politiky



UKÁZKA: REGISTRACE EID

3.0.0.1 → 5.0.0.255

UDP porty 4342

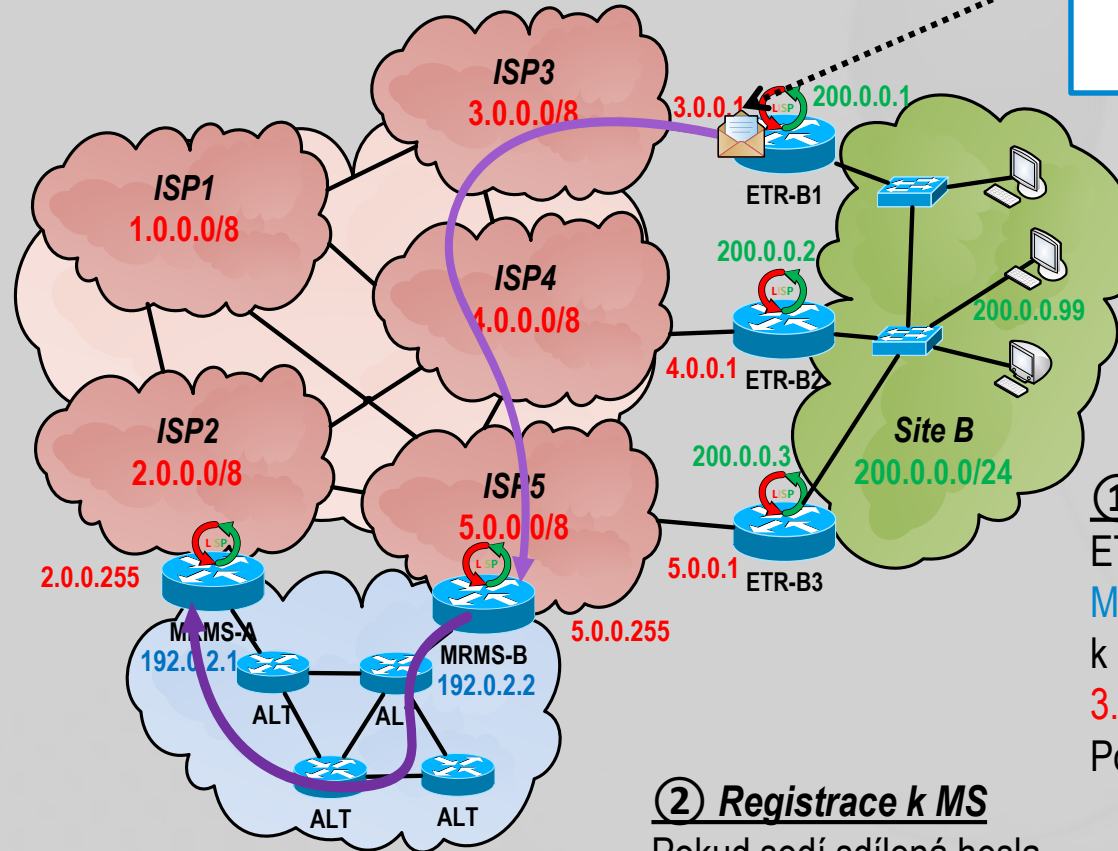
Map-Register

SHA-1 hash

lokální, 3.0.0.1, pri 254, váha 50

4.0.0.1, pri 1, váha 100

5.0.0.1, pri 254, váha 50



① Oznámení EID

ETR generuje zprávu

Map-Register, kde

k prefixu 200.0.0.0/24 patří

3.0.0.1, 4.0.0.1, 5.0.0.1

Posílá se periodicky co 1 min

② Registrace k MS

Pokud sedí sdílená hesla,
pak se zavede o RLOCích k
danému EID

Timeout po 3 minutách

③ MS propagueje do ALTu

BGP v GRE

Přijato dalšími MR

Motivace

Idea

Budoucnost

Demo



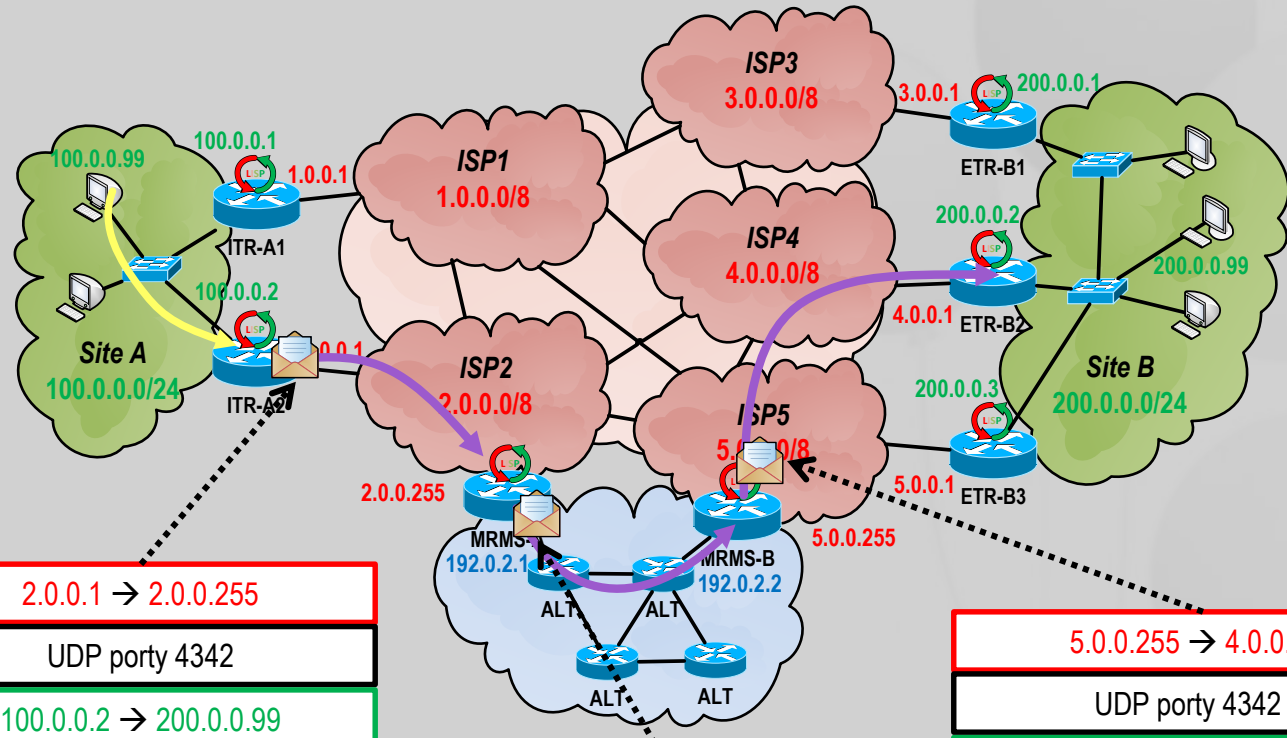
UKÁZKA: MAPOVACÍ DOTAZ

Motivace

Idea

Budoucnost

Demo



2.0.0.1 → 2.0.0.255
UDP porty 4342
100.0.0.2 → 200.0.0.99
UDP porty 4342
Map-Request nonce ITR 2.0.0.1 starající se o 100.0.0.0/24 se ptá, jaký RLOC náleží 200.0.0.99?

192.0.2.1 → 192.0.2.2 GRE
100.0.0.2 → 200.0.0.99
UDP porty 4342
Map-Request

5.0.0.255 → 4.0.0.1
UDP porty 4342
100.0.0.2 → 200.0.0.99
UDP porty 4342
Map-Request

① **Záznam není v cache**
Vygeneruje Map-Request

② **MR deleguje do ALT**
Zapouzdří se do GRE a
pošle se skrz ALT topologií

③ **MS deleguje na ETR**
Posílá se poslednímu
registrujícímu v EID-RLOC
mapovací databázi pro
danou lokalitu



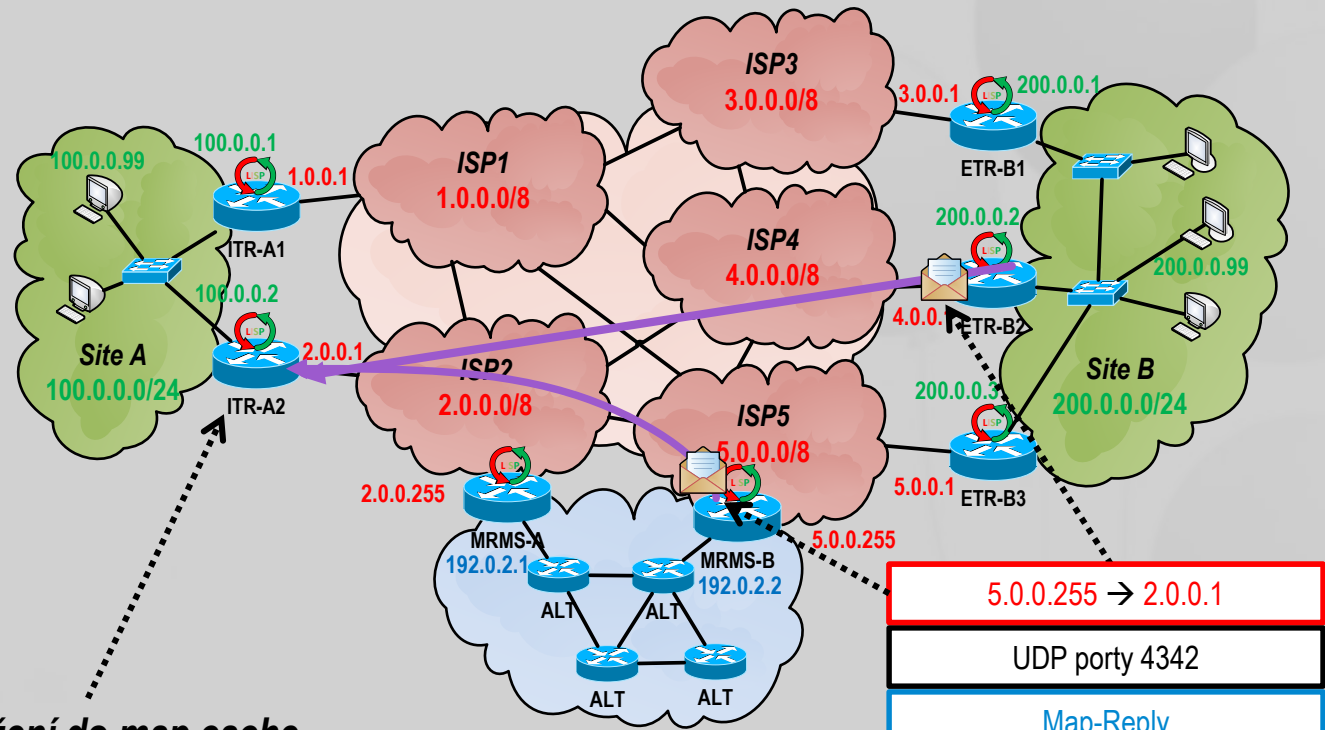
UKÁZKA: MAPOVACÍ ODPOVĚĎ

Motivace

Idea

Budoucnost

Demo



② Uložení do map cache

EID-prefix: 200.0.0.0/24

RLOCy:

- 3.0.0.1, priorita 254, váha 50
- lokální, 4.0.0.1, priorita 1, váha 100
- 5.0.0.1, priorita 254, váha 50

① Alternativní odpověď

MS odpoví ITR zprávou **Map-Reply** místo ETR, pokud mu to ETR během registrace EID přikáže (volba proxy-reply)

① Odeslání odpovědi

ETR odpoví ITR zprávou **Map-Reply**



PŘECHOD NA LISP

- ◆ *Okamžitý přechod je nemožný...*

- ◆ Je potřeba pořešit směřování mezi:
 - ◆ z LISP oblastí do neLISP oblastí
 - ◆ z neLISP oblastí do LISP oblastí

- ◆ Přechodové mechanismy
 - ◆ **Proxy ITR a Proxy ETR**
 - ◆ **LISP-NAT**
 - ◆ LISP-LISP používá standardní způsob překladu
 - ◆ LISP-neLISP překládá EID na RLOC přímo

Motivace

Idea

Budoucnost

Demo



KOMPONENTY: PITR A PETR

Motivace

Idea

Budoucnost

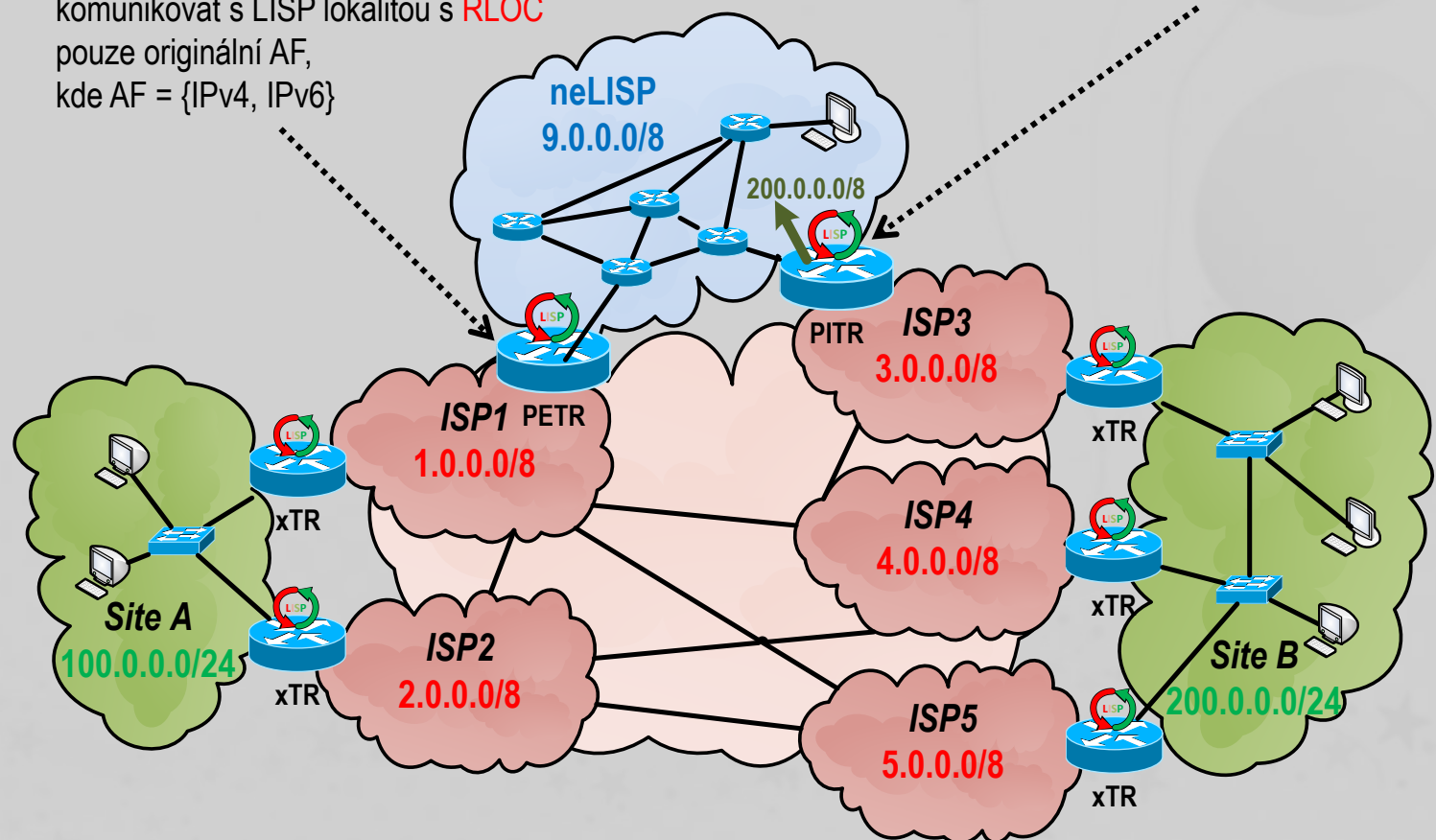
Demo

Proxy Egress Tunnel Router (PETR)

- Umožňuje LISP lokalitám komunikovat s neLISP lokalitami při splnění uRPF podmínky
- Dovoluje LISP lokalitám s EID z jedné AF a RLOC pouze z opačné AF, komunikovat s LISP lokalitou s RLOC pouze originální AF, kde AF = {IPv4, IPv6}

Proxy Ingress Tunnel Router (PITR)

- Provádí ITR funkcionalitu pro neLISP oblasti, zapouzdřuje data z neLISP světa do LISP světa
- Propaguje do neLISP světa agregované EID prefixy





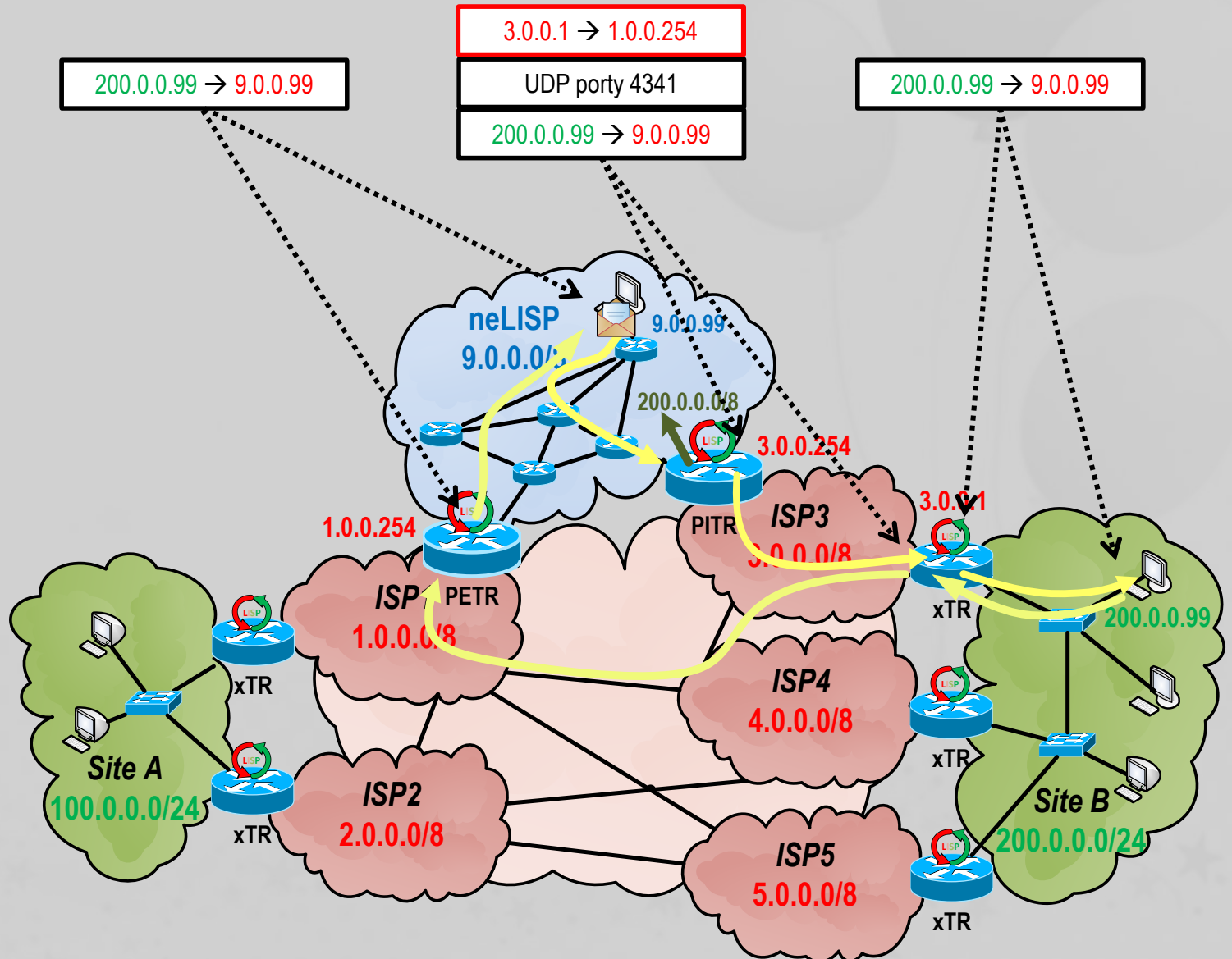
UKÁZKA: KOMUNIKACE S NELISP

Motivace

Idea

Budoucnost

Demo





NA CO SE TEDY LISP POUŽÍVÁ NYNÍ?

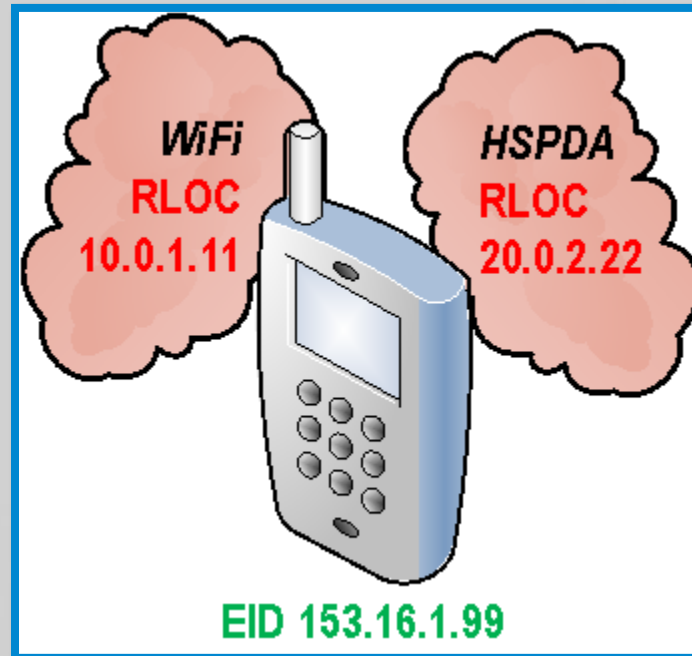
- ◆ Multihoming

Motivace

Idea

Budoucnost

Demo





NA CO SE TEDY LISP POUŽÍVÁ NYNÍ?

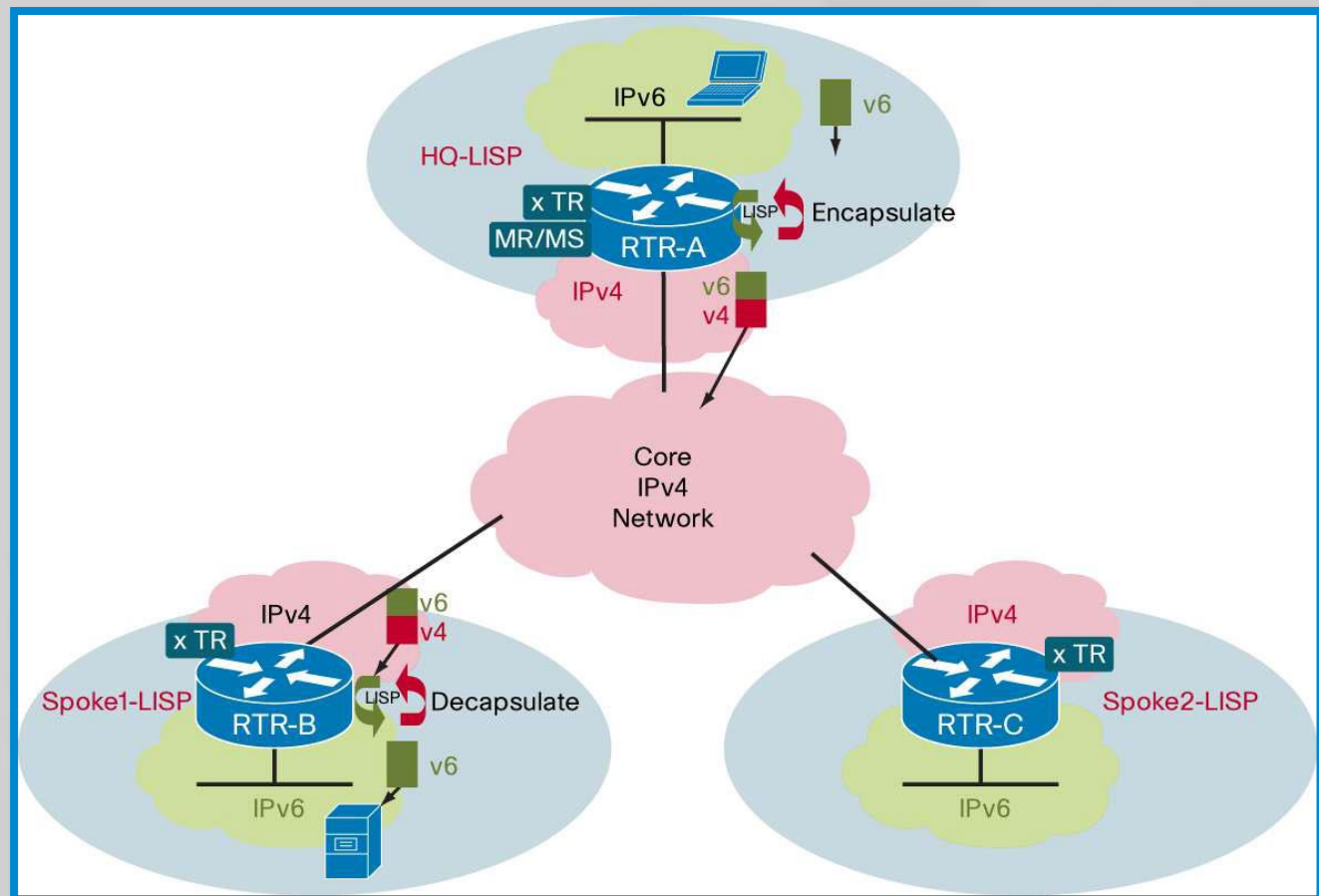
- ◆ Multihoming
- ◆ IPv6 tranzitní mechanismus

Motivace

Idea

Budoucnost

Demo



NA CO SE TEDY LISP POUŽÍVÁ NYNÍ?

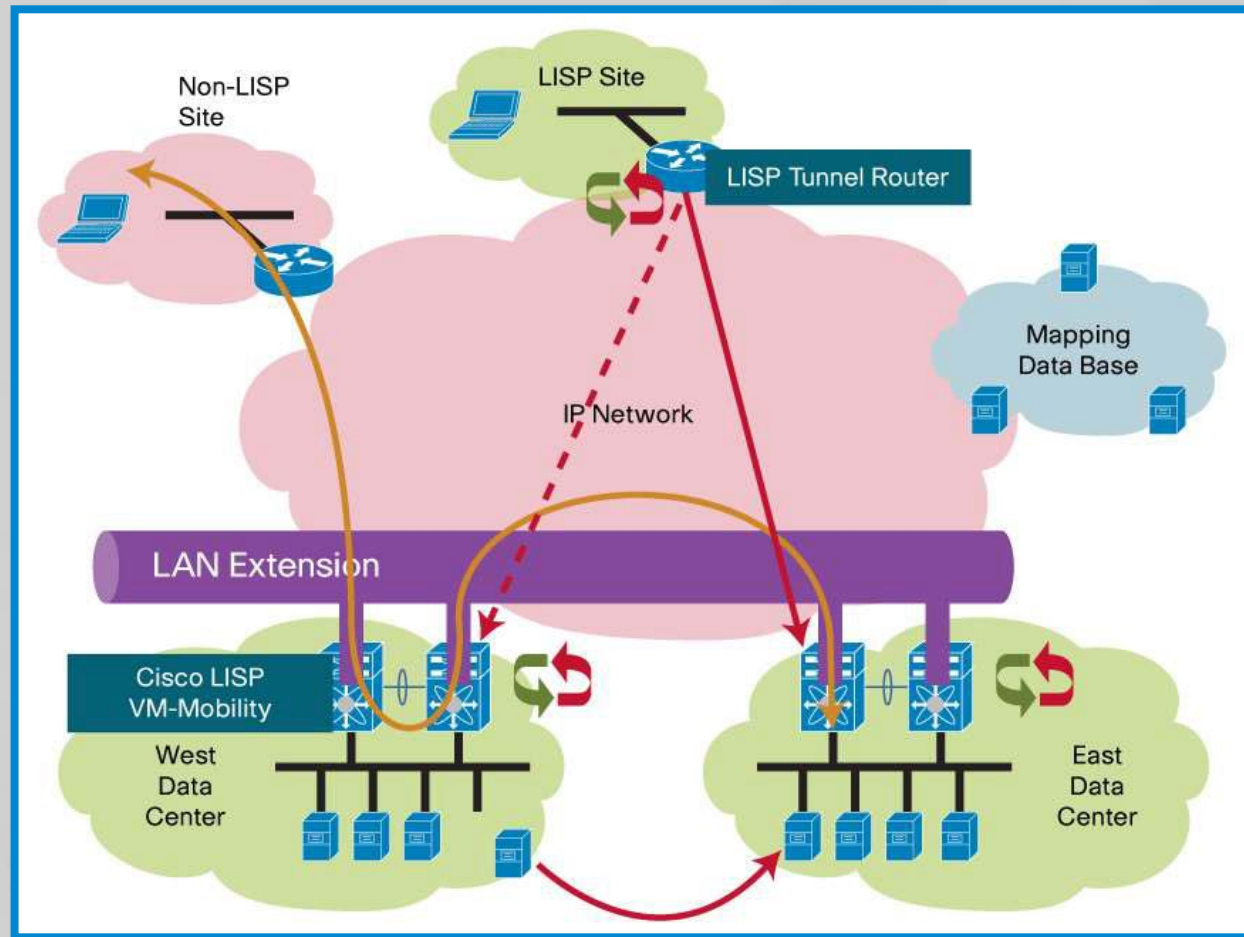
- ◆ Multihoming
- ◆ IPv6 tranzitní mechanismus
- ◆ Mobilita VM

Motivace

Idea

Budoucnost

Demo





Motivace

Idea

Budoucnost

Demo

AKTUÁLNÍ STAV LISPU

- ◆ Vyčleněné prefixy pro EID
 - ◆ IPv4 - **153.16.0.0/16**
 - ◆ IPv6 - **2610:00d0::/32**
- ◆ ALT infrastruktura používá číslování
 - ◆ IPv4 adres GRE tunelů 240.0.0.0/4
 - ◆ ASN 32656.x
- ◆ **LISP Beta Network**
 - ◆ běží více než 4 roky
 - ◆ v cca 32 zemích na 200+ místech
 - ◆ Google, Facebook, Cisco, Qualcomm, AT&T, Lufthansa, Microsoft, ...
- ◆ Implementace
 - ◆ Pro Cisco v IOS, IOS-XE, IOS-XR, NX-OS
 - ◆ Pro FreeBSD je OpenLISP
 - ◆ Pro Linux je LISPmob, Aless a OpenWRT
 - ◆ Pro Android je Gingerbread



KONEC...?

Motivace

Idea

Budoucnost

Demo

Current Internet is just unfinished demo!

John Day



STANDARDIZAČNÍ SNAHY

Motivace

Idea

Budoucnost

Demo

Lisp Status Pages

Int Area: Brian Haberman, Ralph Droms | 2009-Apr-28 —

Locator/ID Separation Protocol (Active WG)

Chairs:

[Joel Halpern](#) 




[Terry Manderson](#) 

[Drafts](#) | [Agendas](#) | [Minutes](#) | [Wiki](#) | [Issues](#) | [Charters](#) | [Jabber Room, Logs](#) | [List Archive](#) | List search:

Working Group Documents:

Document collections: [epub](#) [mobi](#)



<u>Draft name</u>	<u>Rev.</u>	<u>Dated</u>	<u>Status</u>	<u>Comments, Issues</u>
<i>Active:</i>				
draft-ietf-lisp-architecture	-00	2012-10-15	Active	
draft-ietf-lisp-ddt	-00	2012-10-15	Active	
draft-ietf-lisp-deployment	-05	2012-10-20	Active	
draft-ietf-lisp-introduction	-00	2012-10-15	Active	
draft-ietf-lisp-lcaf	-00	2012-08-30	Active	
draft-ietf-lisp-sec	-04	2012-10-12	Active	
draft-ietf-lisp-threats	-03	2012-10-16	Active	
<i>IESG Processing:</i>				
draft-ietf-lisp-eid-block	-03 <small>new</small>	2012-11-07	AD Evaluation::AD Followup	
draft-ietf-lisp-mib	-07	2012-10-16	AD Evaluation	
<i>RFC-Editor's Queue:</i>				
draft-ietf-lisp	-23	2012-05-05	RFC Ed Queue	
draft-ietf-lisp-alt	-10	2011-12-06	RFC Ed Queue	
draft-ietf-lisp-interworking	-06	2012-03-04	RFC Ed Queue	
draft-ietf-lisp-lig	-06	2011-09-09	RFC Ed Queue	
draft-ietf-lisp-map-versioning	-09	2012-03-01	RFC Ed Queue	
draft-ietf-lisp-ms	-16	2012-03-05	RFC Ed Queue	
draft-ietf-lisp-multicast	-14	2012-02-08	RFC Ed Queue	



Motivace

Idea

Budoucnost

Demo

REFERENCE

◆ GoogleTech Talks

- ◆ <http://www.youtube.com/watch?v=WSI1RAIFU3s>
- ◆ http://www.youtube.com/watch?v=_bz4cRuAcak
- ◆ <http://www.youtube.com/watch?v=fxdm-Xouu-k>
- ◆ http://www.cisco.com/web/about/ac123/ac147/archived_issues/ipj_11-1/111_lisp.html

◆ IETF WG

- ◆ <https://datatracker.ietf.org/wg/lisp/charter/>

◆ Další informace

- ◆ Oficiální stránka projektu - http://lisp.cisco.com/lisp_tech.html
- ◆ LISP Beta Network - <http://www.lisp4.net> a <http://www.lisp6.net>
- ◆ LISP DDT Root - <http://www.ddt-root.org>
- ◆ LinkedIn - <http://www.linkedin.com/groups/LISP-3776183>



SCÉNÁŘ

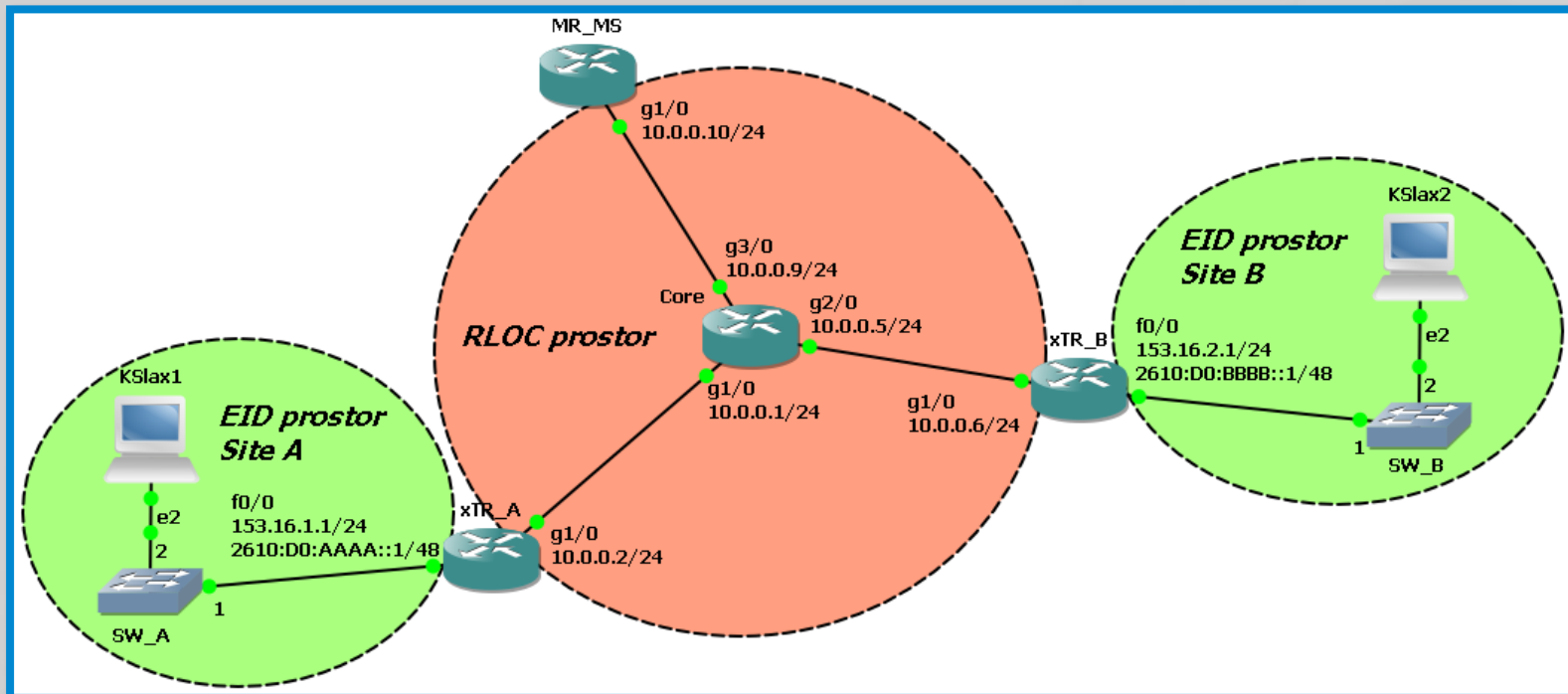
- ◆ GNS3
- ◆ C7200 s IOSem ADVIPSERVICESK9-M, verze 15.2(4)M
- ◆ Koncové stanice se SLAX 6
- ◆ IPv6 ping z jedné lokality na druhou

Motivace

Idea

Budoucnost

Demo





VÝSLEDKY

◆ LISP jako tranzitní mechanismus pro IPv6-IPv4

Motivace

Idea

Budoucnost

Demo

Filter: icmpv6

No.	Source	Destination	Time	Length	Protocol	Info
548	2610:d0:aaaa::1	2610:d0:bbbb::1	2012-11-23 11:41:59.345407	150	ICMPv6 Echo (ping)	request id=0x097c, seq=1
551	2610:d0:aaaa::1	2610:d0:bbbb::1	2012-11-23 11:42:01.345407	150	ICMPv6 Echo (ping)	request id=0x097c, seq=2
552	2610:d0:bbbb::1	2610:d0:aaaa::1	2012-11-23 11:42:01.395407	150	ICMPv6 Echo (ping)	reply id=0x097c, seq=2
553	2610:d0:aaaa::1	2610:d0:bbbb::1	2012-11-23 11:42:01.465407	150	ICMPv6 Echo (ping)	request id=0x097c, seq=3
554	2610:d0:bbbb::1	2610:d0:aaaa::1	2012-11-23 11:42:01.515407	150	ICMPv6 Echo (ping)	reply id=0x097c, seq=3
555	2610:d0:aaaa::1	2610:d0:bbbb::1	2012-11-23 11:42:01.585407	150	ICMPv6 Echo (ping)	request id=0x097c, seq=4
556	2610:d0:bbbb::1	2610:d0:aaaa::1	2012-11-23 11:42:01.635407	150	ICMPv6 Echo (ping)	reply id=0x097c, seq=4

Frame 548: 150 bytes on wire (1200 bits), 150 bytes captured (1200 bits)

- Ethernet II, Src: ca:01:1d:00:00:38 (ca:01:1d:00:00:38), Dst: ca:00:1d:00:00:1c (ca:00:1d:00:00:1c)
- Internet Protocol Version 4, Src: 10.0.0.2 (10.0.0.2), Dst: 10.0.0.6 (10.0.0.6)
 - Version: 4
 - Header length: 20 bytes
 - Differentiated Services Field: 0x00 (DSCP 0x00: Default; ECN: 0x00: Not-ECT (Not ECN-Capable Transport))
 - Total Length: 136
 - Identification: 0x0001 (1)
 - Flags: 0x02 (Don't Fragment)
 - Fragment offset: 0
 - Time to live: 254
 - Protocol: UDP (17)
 - Header checksum: 0x685c [correct]
 - Source: 10.0.0.2 (10.0.0.2)
 - Destination: 10.0.0.6 (10.0.0.6)
- User Datagram Protocol, Src Port: 1024 (1024), Dst Port: lisp-data (4341)
- Locator/ID Separation Protocol (Data)
- Internet Protocol Version 6, Src: 2610:d0:aaaa::1 (2610:d0:aaaa::1), Dst: 2610:d0:bbbb::1 (2610:d0:bbbb::1)
 - 0110 = version: 6
 - 0000 0000 = Traffic class: 0x00000000
 - 0000 0000 0000 0000 = Flowlabel: 0x00000000
 - Payload length: 60
 - Next header: ICMPv6 (0x3a)
 - Hop limit: 64
 - Source: 2610:d0:aaaa::1 (2610:d0:aaaa::1)
 - Destination: 2610:d0:bbbb::1 (2610:d0:bbbb::1)
- Internet Control Message Protocol v6

Standard input: <live capture in progress> Fi... Packets: 733 Displayed: 7 Marked: 0 Profile: Default