

RIPE

Enhancing the Internet
Architecture:

LISP- Locator/Id Separation protocol

Luigi Iannone



Tutorial Road Map

- Why LISP???
- LISP Principles
- LISP Data Plane
 - RFC 6830
- LISP Control Plane
 - LISP-MS (RFC 6833)
 - LISP-NERD (RFC 6837)
 - LISP+ALT (RFC 6836)
 - LISP-DDT (draft-ietf-lisp-ddt-03.txt)
- LISP-Versioning
 - RFC 6834
- LISP-Interworking
 - RFC 6832

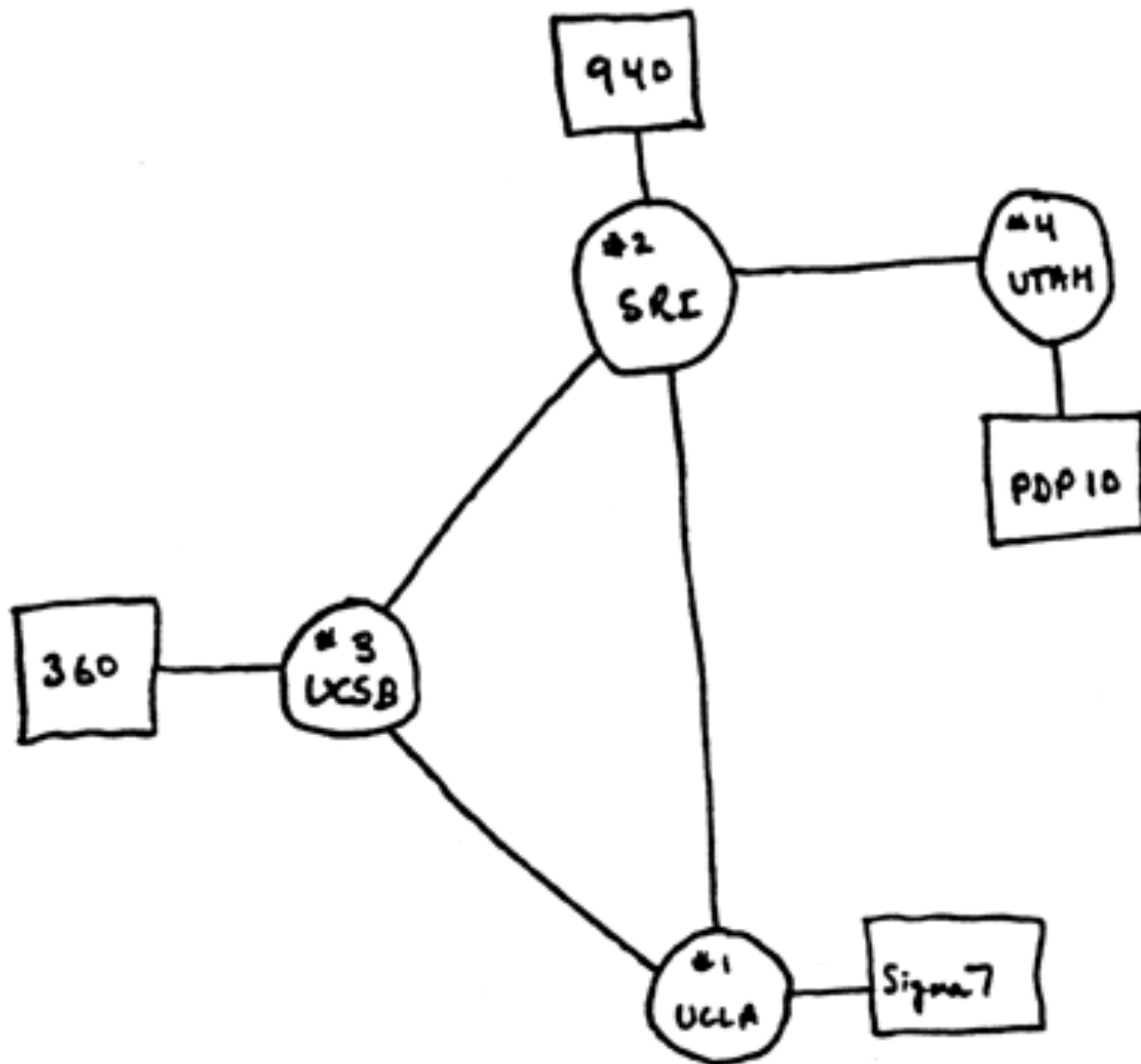
RIPE

A large graphic element on the left side of the slide. It features the word "RIPE" in a bold, teal, sans-serif font. To the right of the text are two vertical white bars of different heights and two horizontal teal bars of different lengths, creating a stylized cross or plus sign.

Why LISP ???



The origins: back in 70s



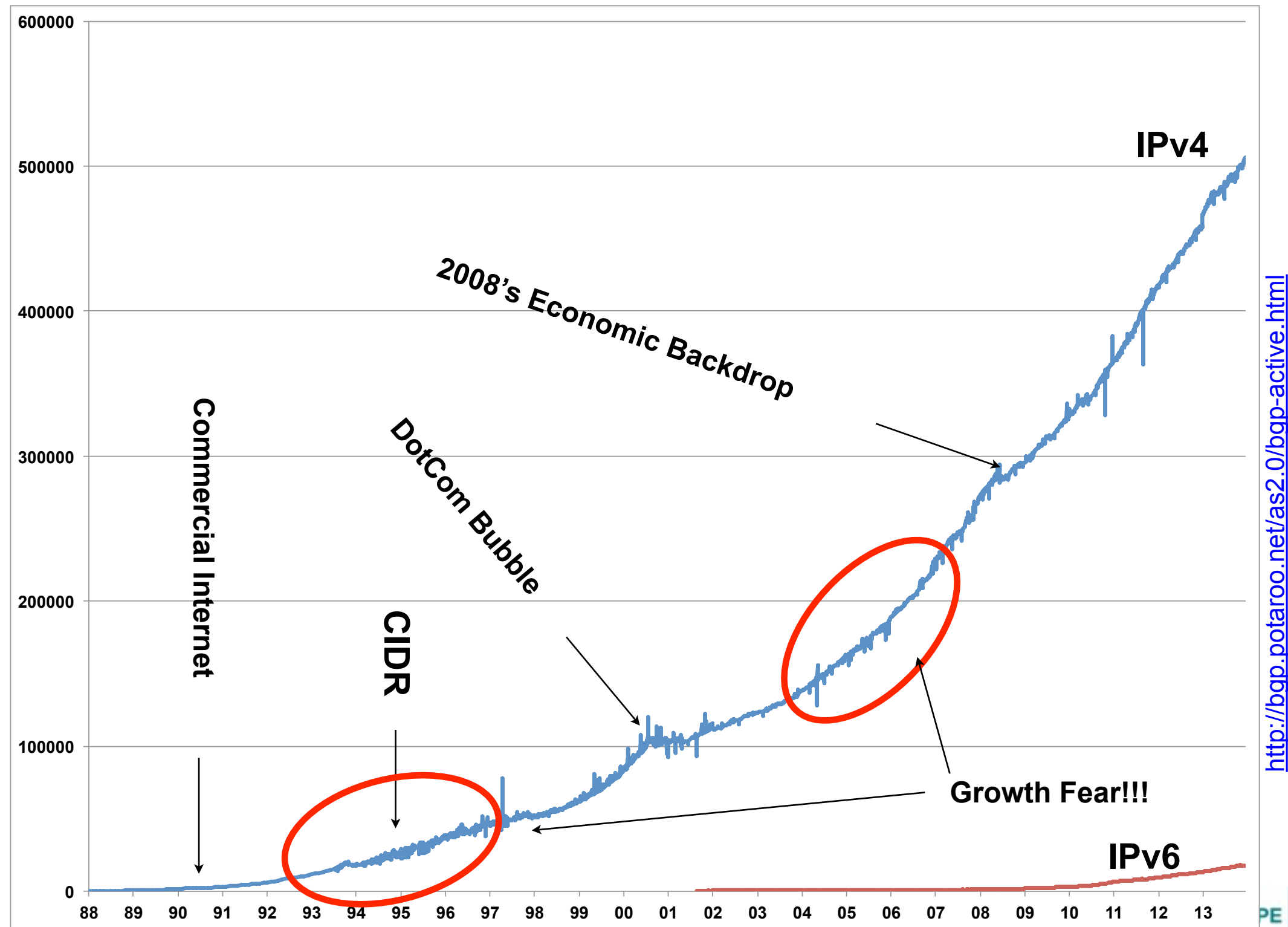
Things Happen When You Wear
ELEGANT



Do you ever heard of HOSTS.TXT ???

- In the 70s HOSTS.TXT contained all translation from hosts names to IP addresses
- Same approach as routing today:
 - everybody knows everybody
- In the early 80s it was a nightmare to maintain
- ... and then came DNS (Domain Name System)
 - On demand
 - Distributed
 - Hierarchical
 - in other words scalable

BGP = HOSTS.TXT ?????

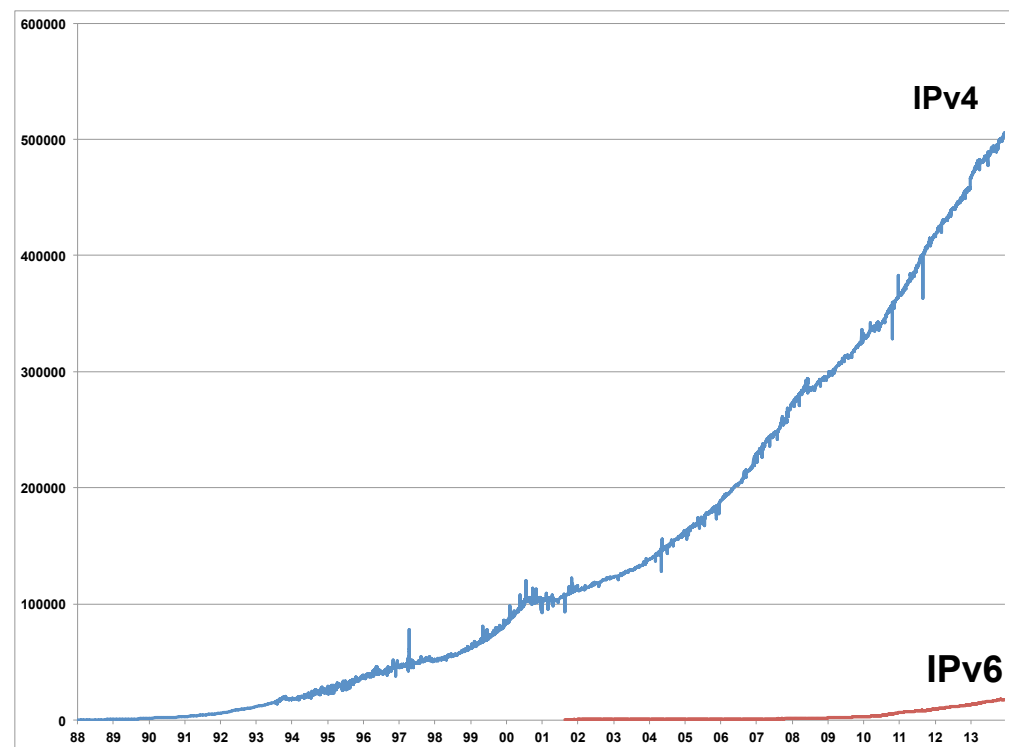


<http://bgp.potaroo.net/as2.0/bgp-active.html>

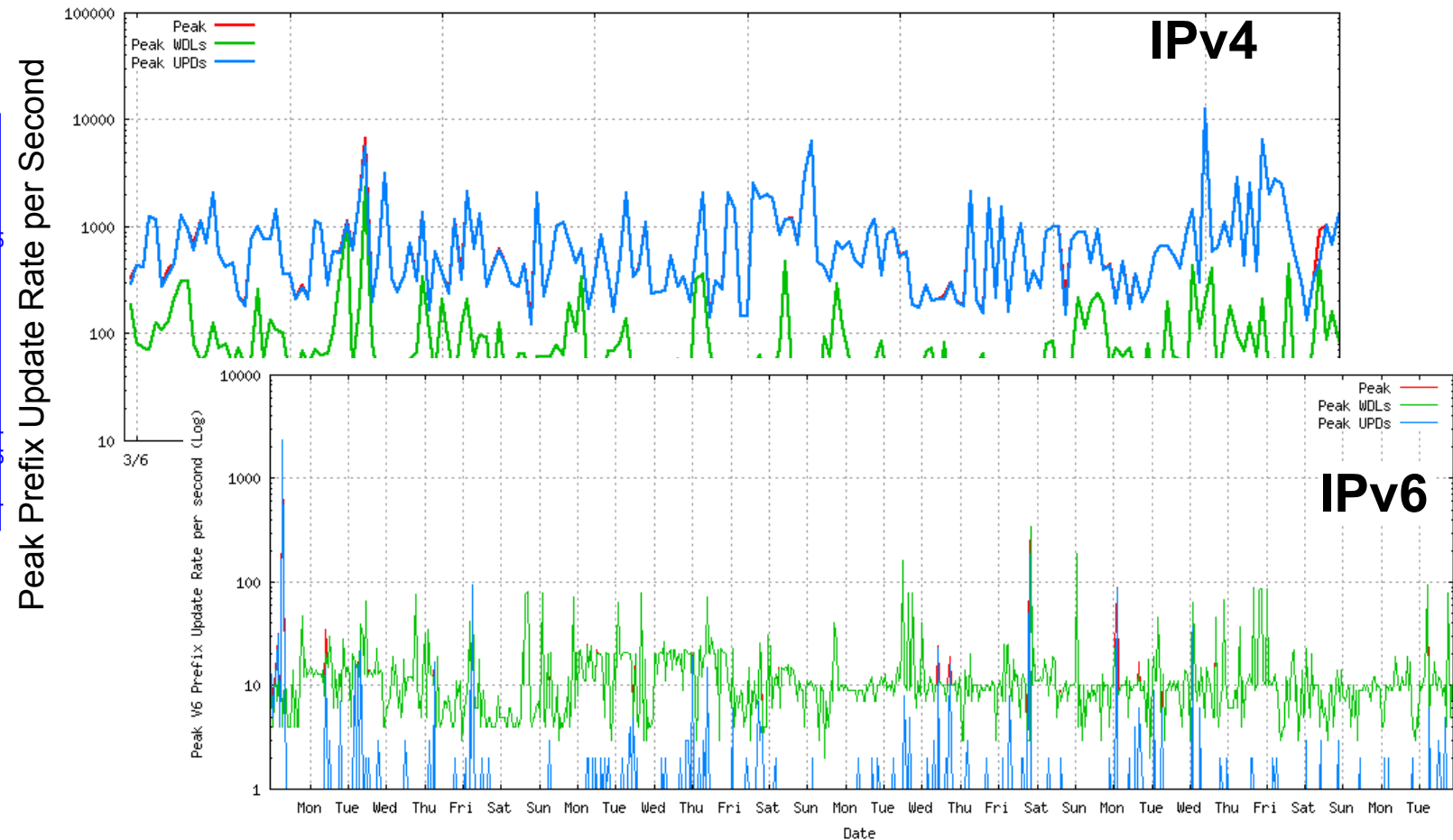


Root of the BGP's FIB inflation problem

No... is not just natural Internet growth



<http://bgp.potaroo.net/as2.0/bgp-active.html>



BGP Forwarding Information Base (FIB) and Churn Explosion:

PI (Provider Independent) prefix assignment

Multi-homing

Traffic-Engineering

Security

Remember the YouTube incident?

Churn can have peaks of thousands per seconds

Churn increases the need processing power



So where the problem comes from???

*“Addressing can follow topology
or*

topology can follow addressing. Choose one.”

Rekhter's Law

The overloaded IP address Semantic

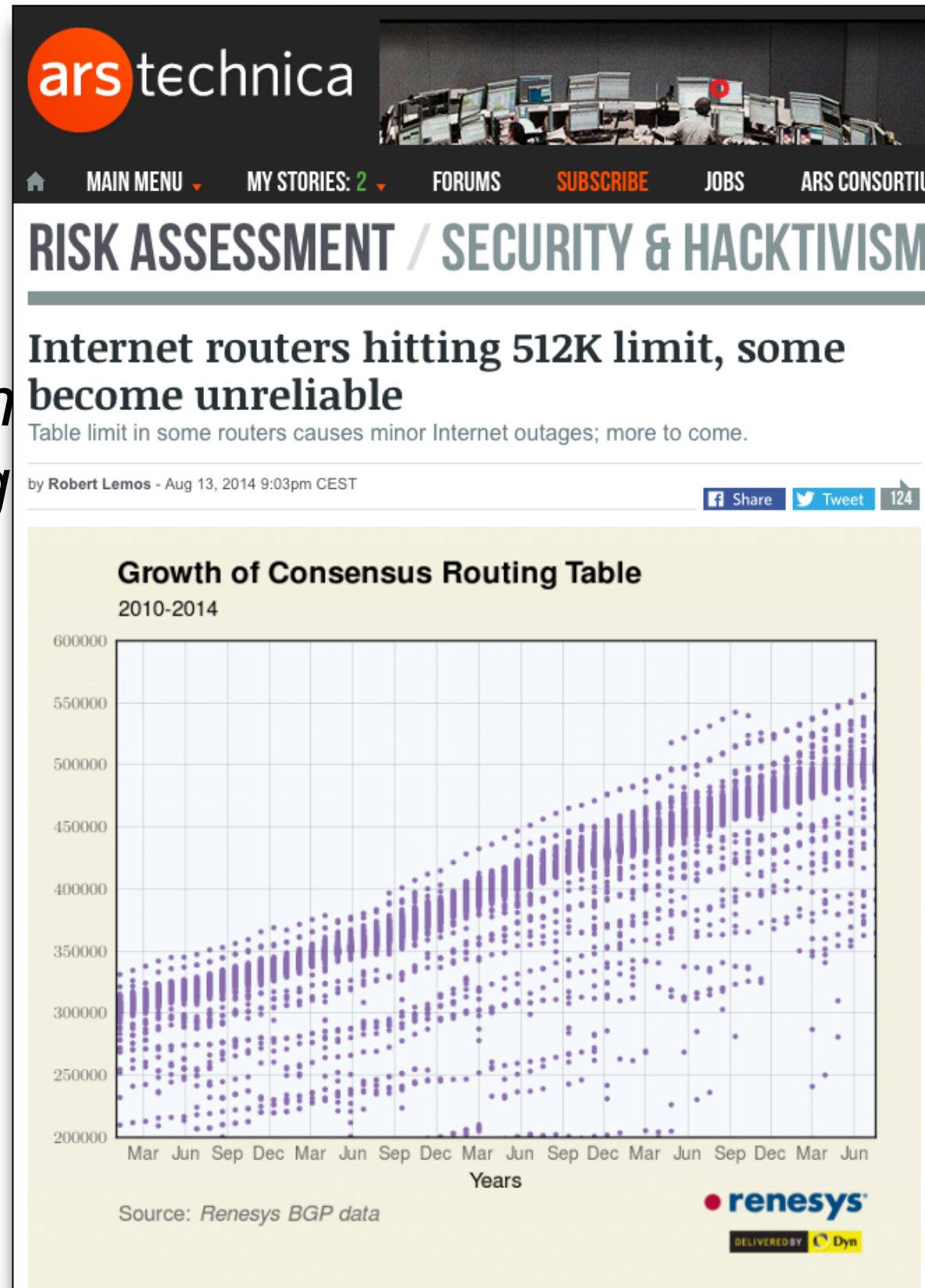
- An IP Address tells:
 - Who you are
 - Hi! I am *Luigi Iannone RIPE 70....*
 - Where you are
 - ...and you can find me at *Luigi Iannone RIPE 70*
- This design was OK in the 70s-80s
 - Because was easier to implement
 - Because the Internet was a small academic network of networks



Internet's Scaling Issues

- D. Meyer, L.
Routing and

*"It is common
addressing*



B Workshop on
September 2007.

*Internet routing and
ing problems."*

From “*Divide and Conquer*”

....to “*Split and Scale*”

“The Research Group has rough consensus that separating identity from location is desirable and technically feasible. However, the Research Group does NOT have consensus on the best engineering approach to such an identity/location split.”

From RFC 6115: Recommendation for a Routing Architecture

Along with a plethora of proposals:

Locator/ID Separation Protocol (LISP)

Routing Architecture for the Next Generation Internet (**RANGI**)

Internet Vastly Improved Plumbing (**Ivip**)

Hierarchical IPv4 Framework (**hIPv4**)

Name Overlay (**NOL**) Service for Scalable Internet Routing

Compact Routing in a Locator Identifier Mapping System (**CRM**)

Layered Mapping System (**LMS**)

Two-Phased Mapping

Global Locator, Local Locator, and Identifier Split (**GLI-Split**)

Tunneled Inter-Domain Routing (**TIDR**)

Identifier-Locator Network Protocol (**ILNP**)

Enhanced Efficiency of Mapping Distribution protocols in Map-and-Encap Schemes (**EEMDP**)

Evolution

Name-Based Sockets

Routing and Addressing in Networks with Global Enterprise Recursion (**IRON-RANGER**)

Hierarchical Architecture for Internet Routing (**HAIR**)



RIPE

A large graphic element on the left side of the slide. It features the word "RIPE" in a bold, teal, sans-serif font. To the right of the text are two vertical white bars of different heights and two horizontal teal bars of different lengths, creating a stylized cross or plus sign.

LISP Principles



Core/Edge Separation

Who are all those prefixes?

- Number of Active ASes: 48349
- Number of Origin Only ASes: 41274 (85%)
- Average entries per Origin AS: ~11
- Roughly ~454 000 Prefixes are Stub Networks

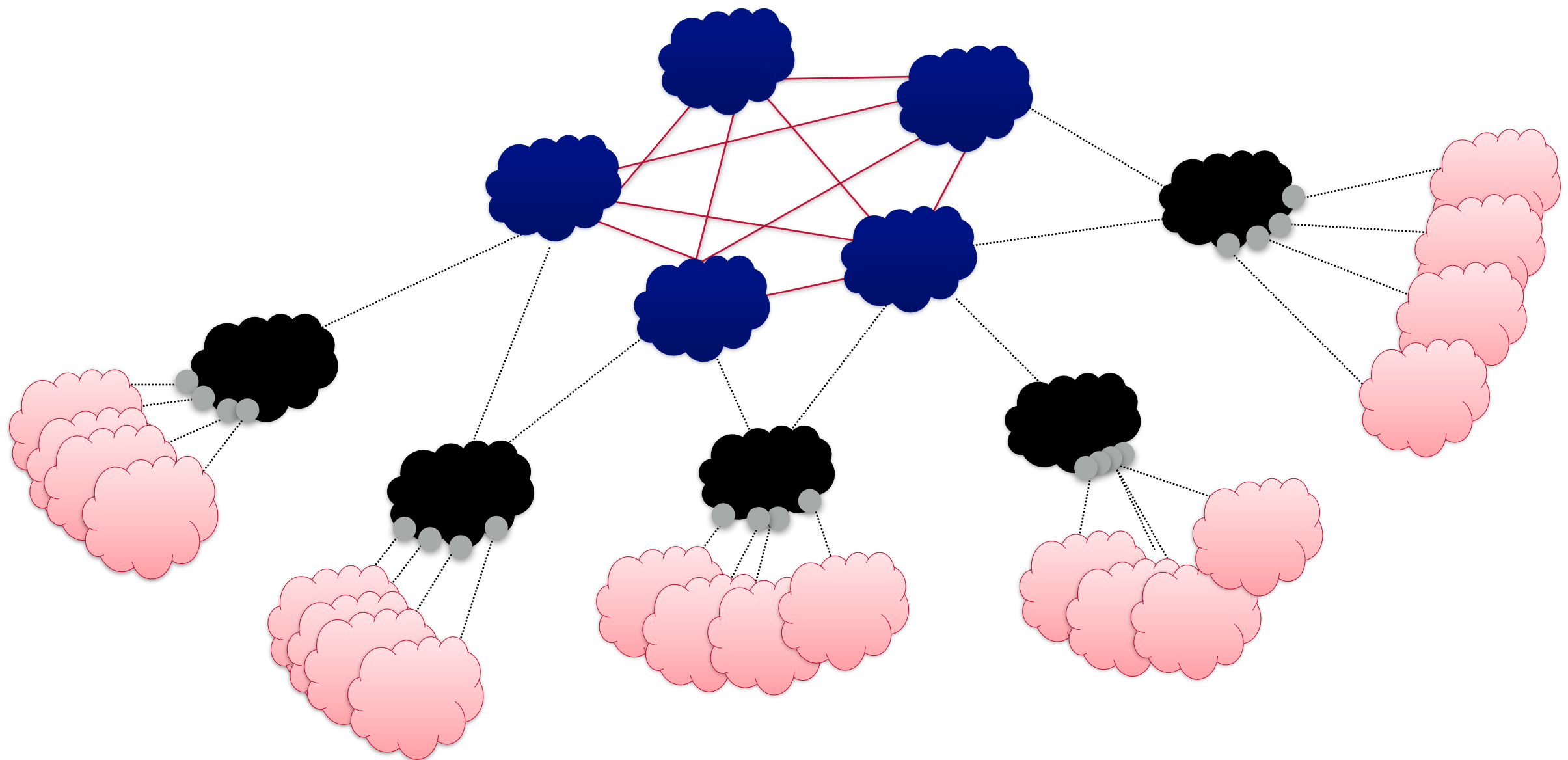
<http://bgp.potaroo.net/as2.0/bgp-active.html>

85%

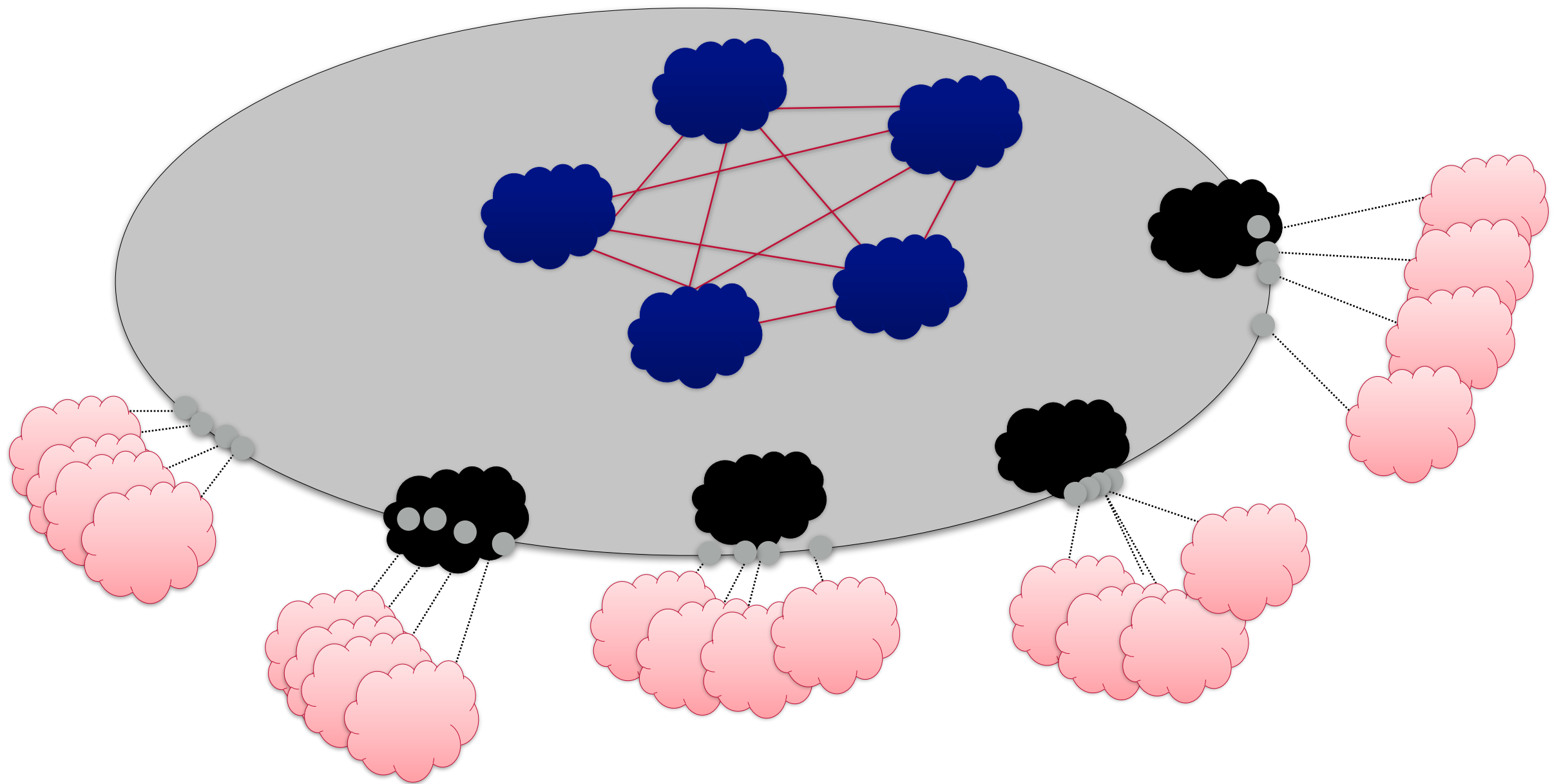


The Separated Core

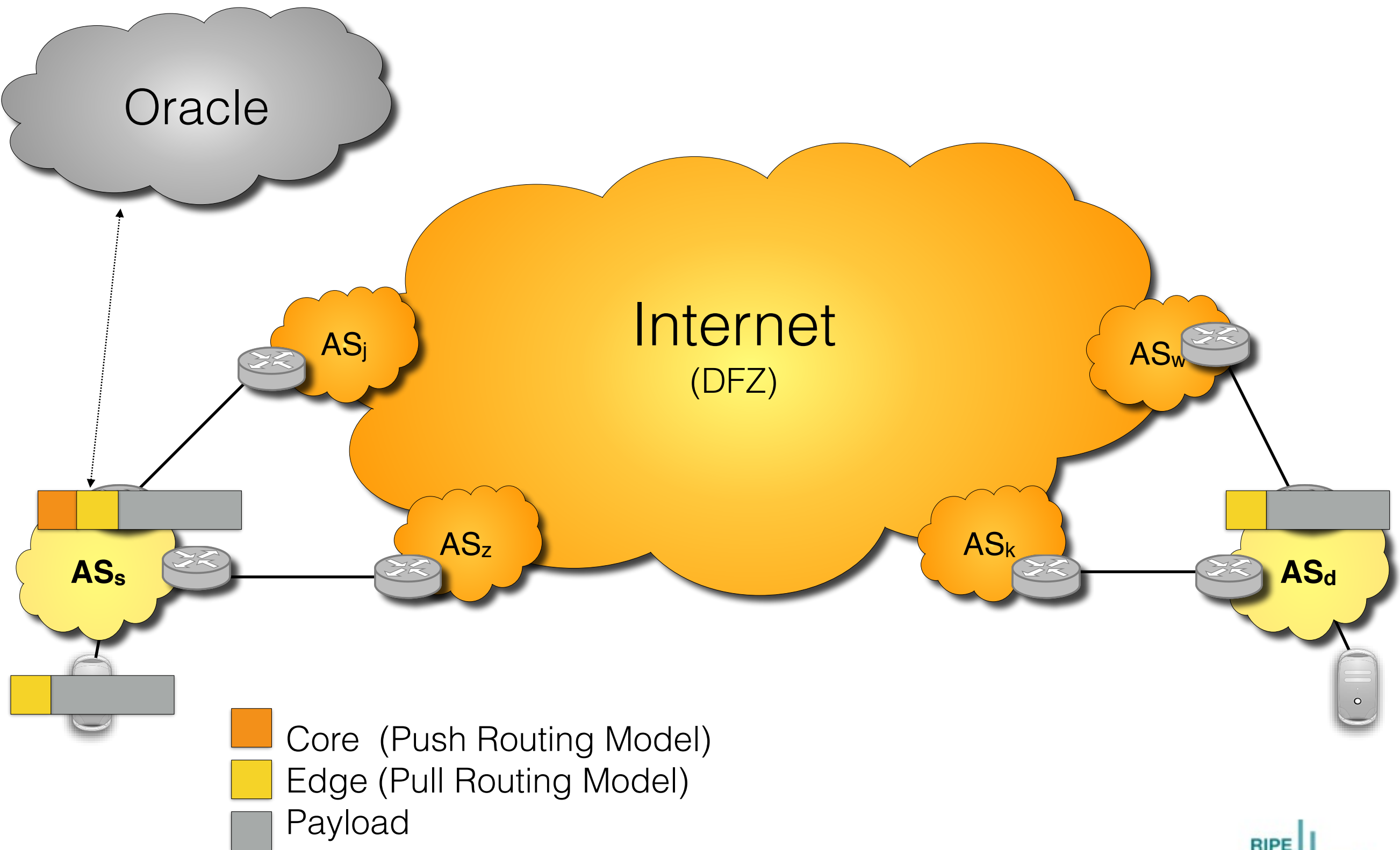
● Internet Routing Location



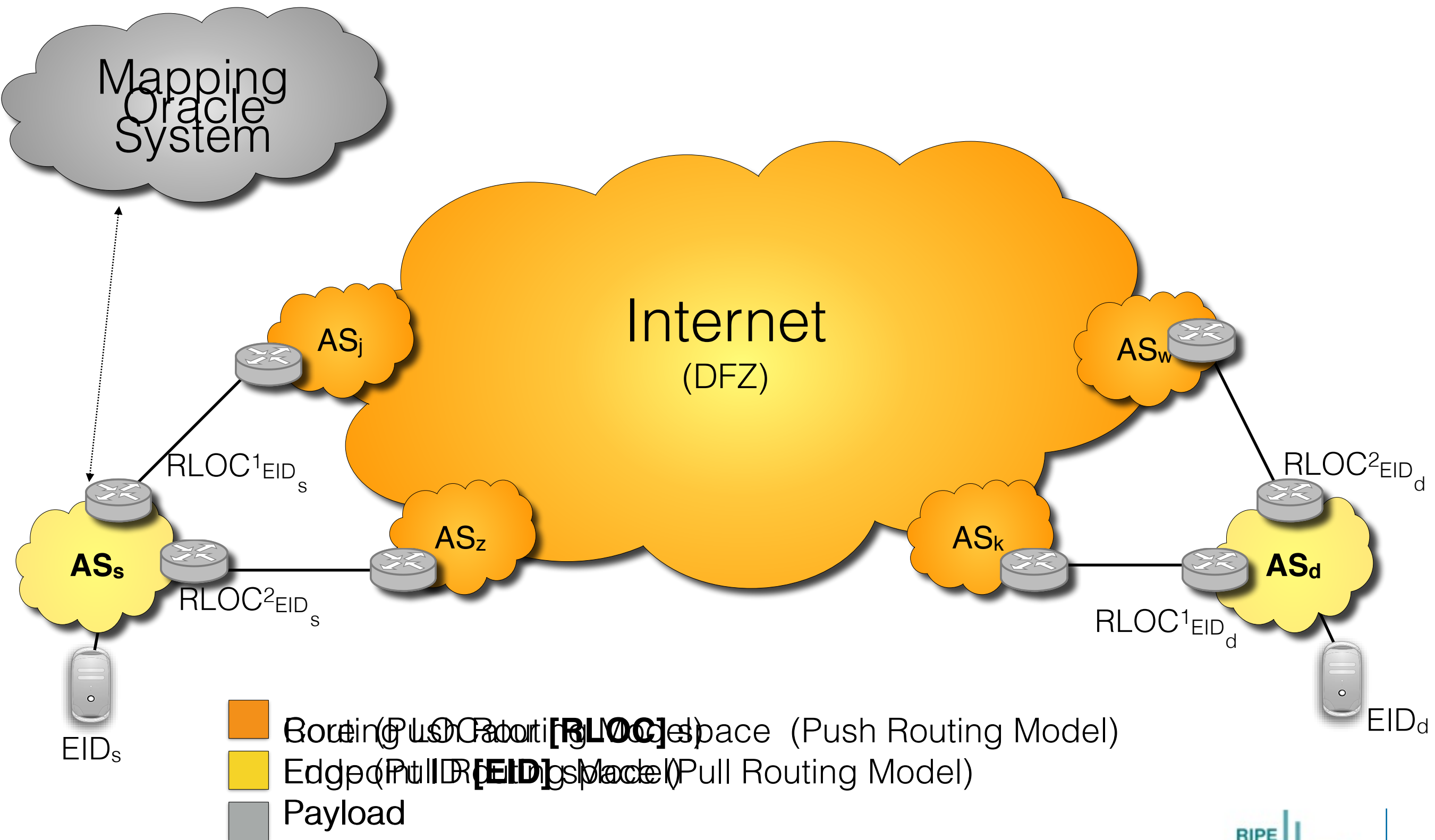
The Separated Edge



Packets in Core/Edge Separation



From Core/Edge Separation to LISP



RIPE

A large graphic element on the left side of the slide. It features the word "RIPE" in a bold, teal, sans-serif font. To the right of the text are two vertical white bars of different heights and two horizontal teal bars of different lengths, creating a stylized cross or plus sign shape.

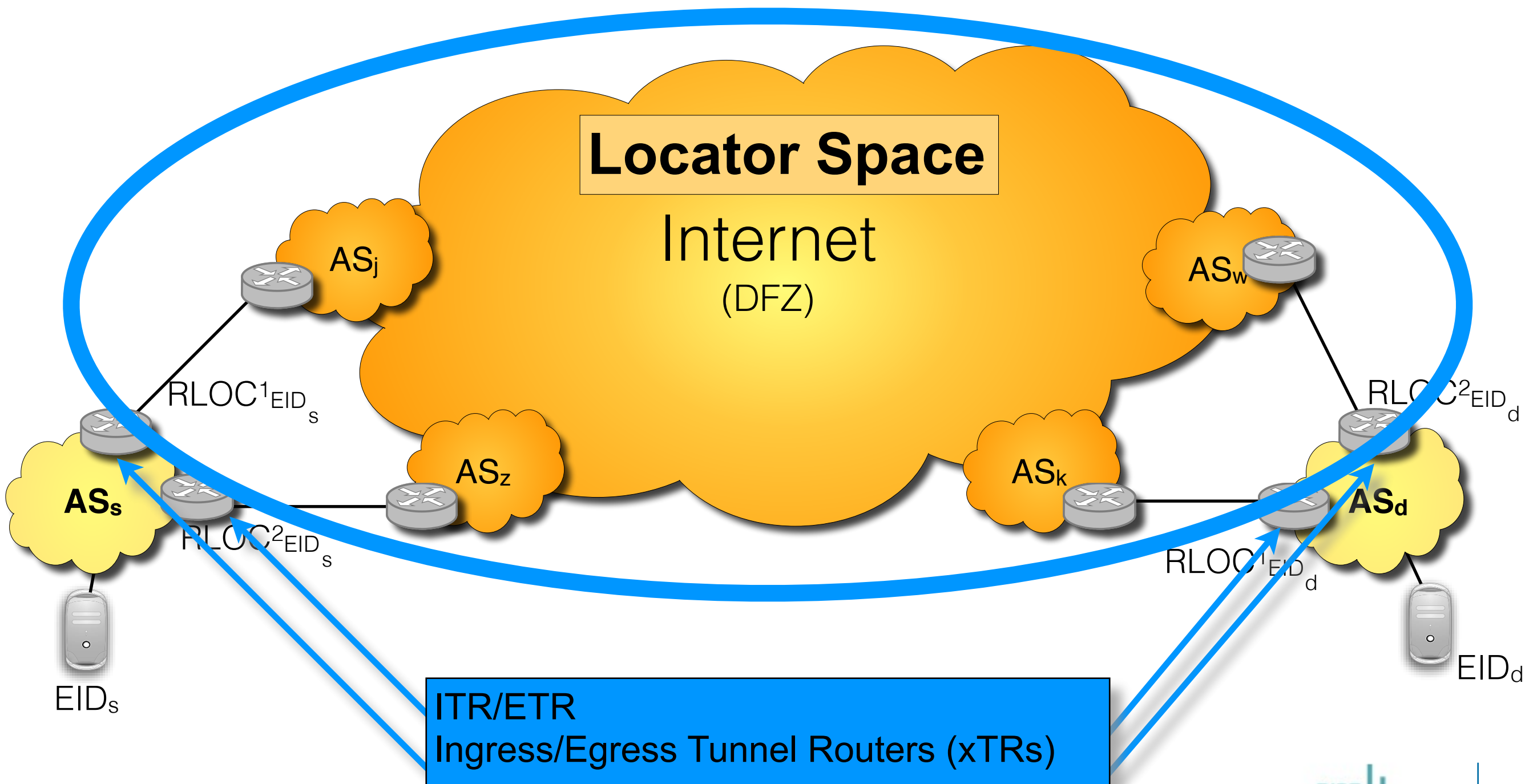
LISP Data Plane

RFC 6830



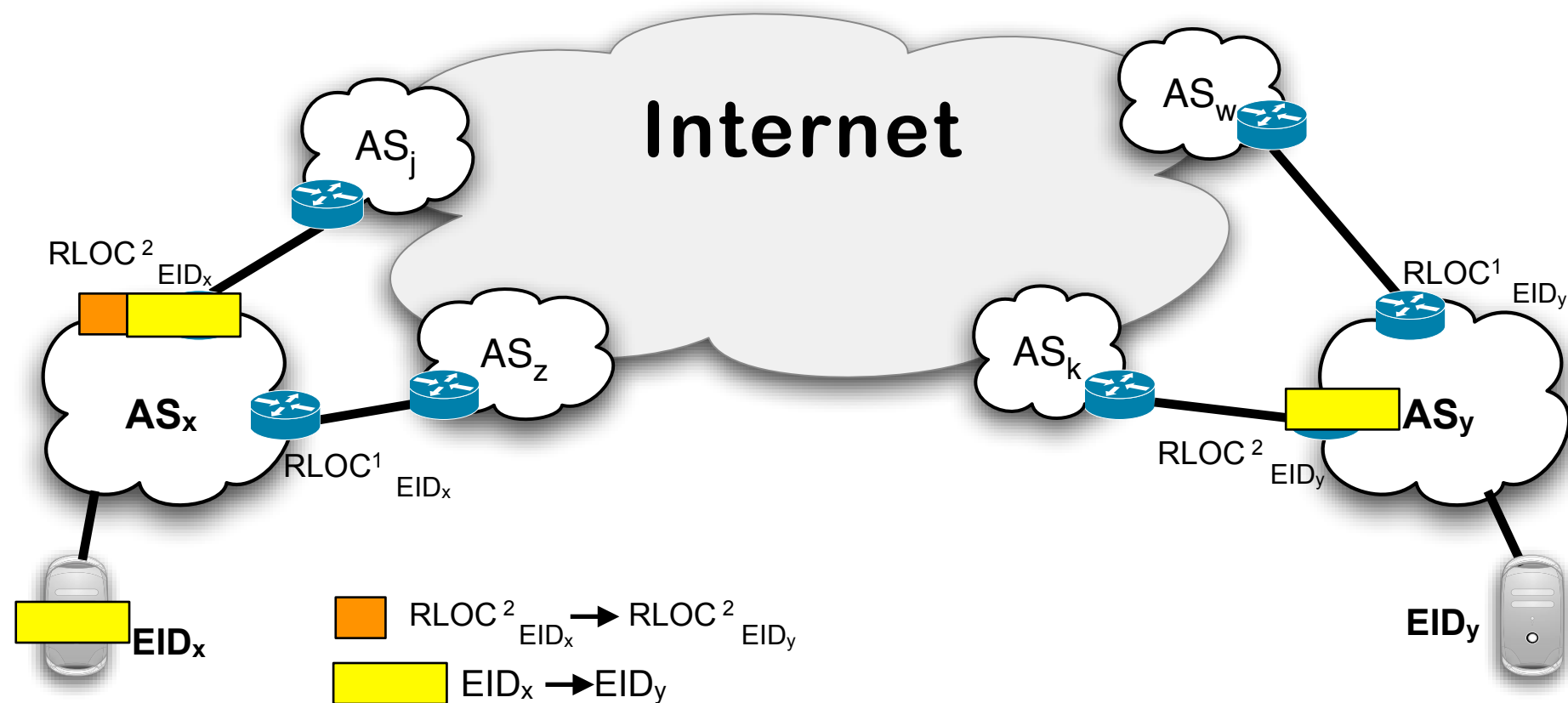
At the border of ID and Location

Bindings between **Identifier Space**: Mappings



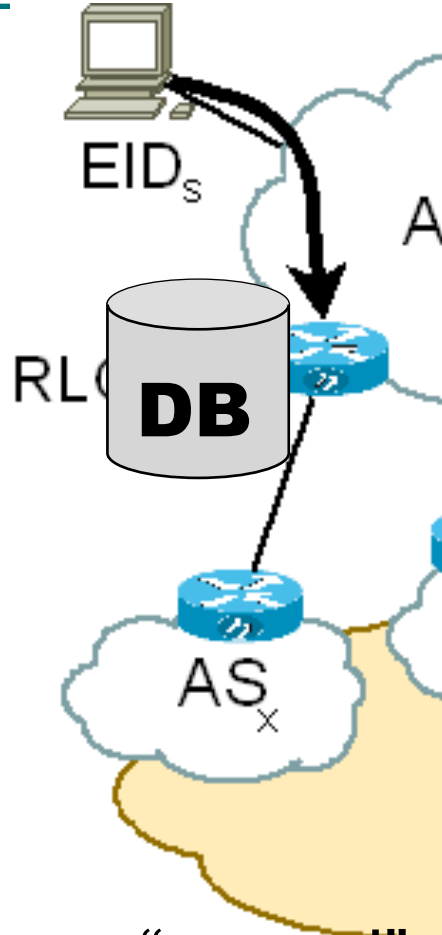
Life of a Packet in a LISP World..

Bindings between ID and Locators: *Mappings*



$EID_x \rightarrow RLOC^1_{EID_x} \quad RLOC^2_{EID_x}$
 $EID_y \rightarrow RLOC^1_{EID_y} \quad RLOC^2_{EID_y}$

Map & Encap Operations: source side



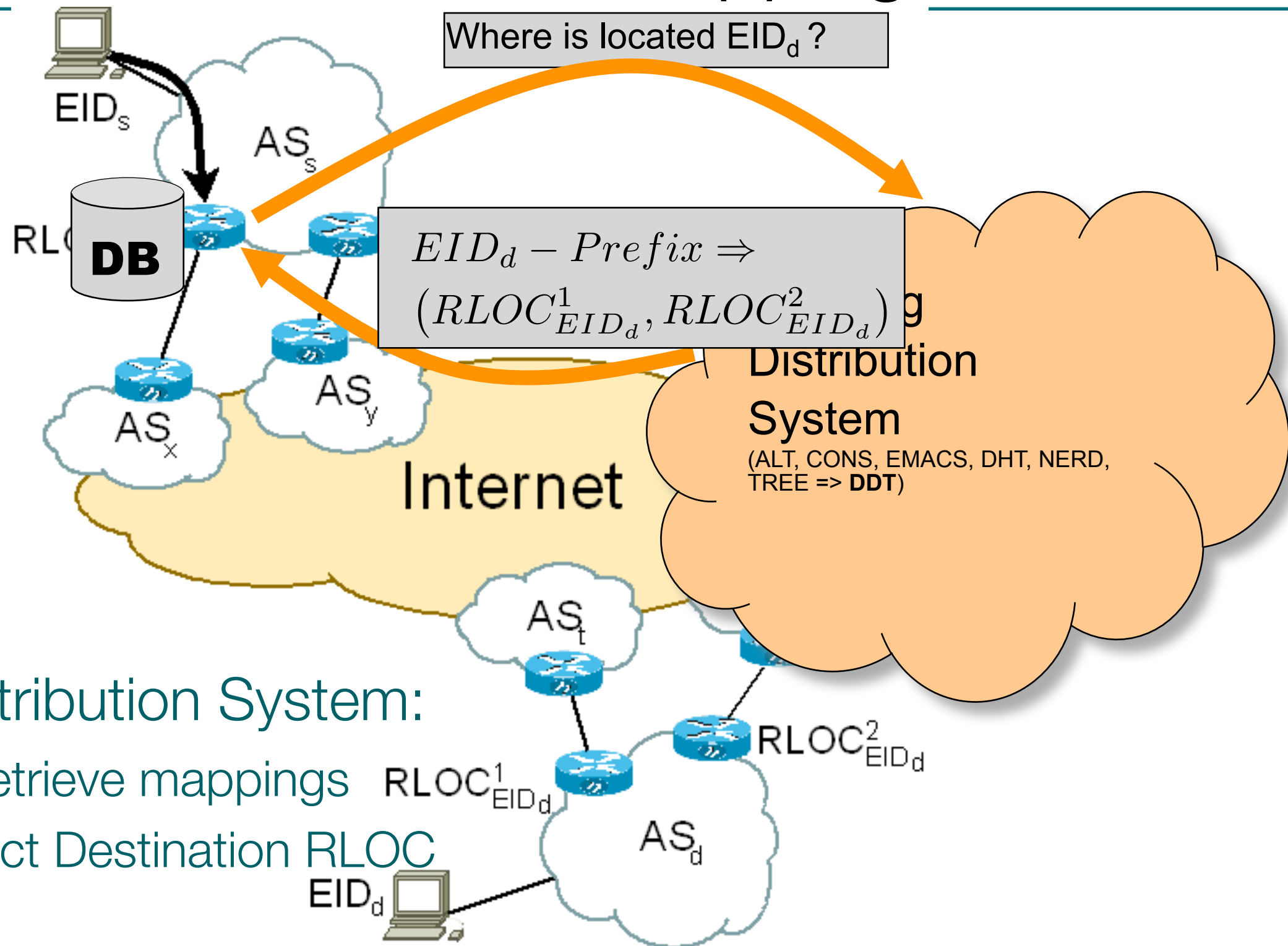
LISP-Database:

- Contains mappings “owned” locally
- Used to select source RLOC

$$EID_s - Prefix \Rightarrow (RLOC_{EID_s}^1, RLOC_{EID_s}^2)$$

0										1										2										3									
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								
+--+																																							

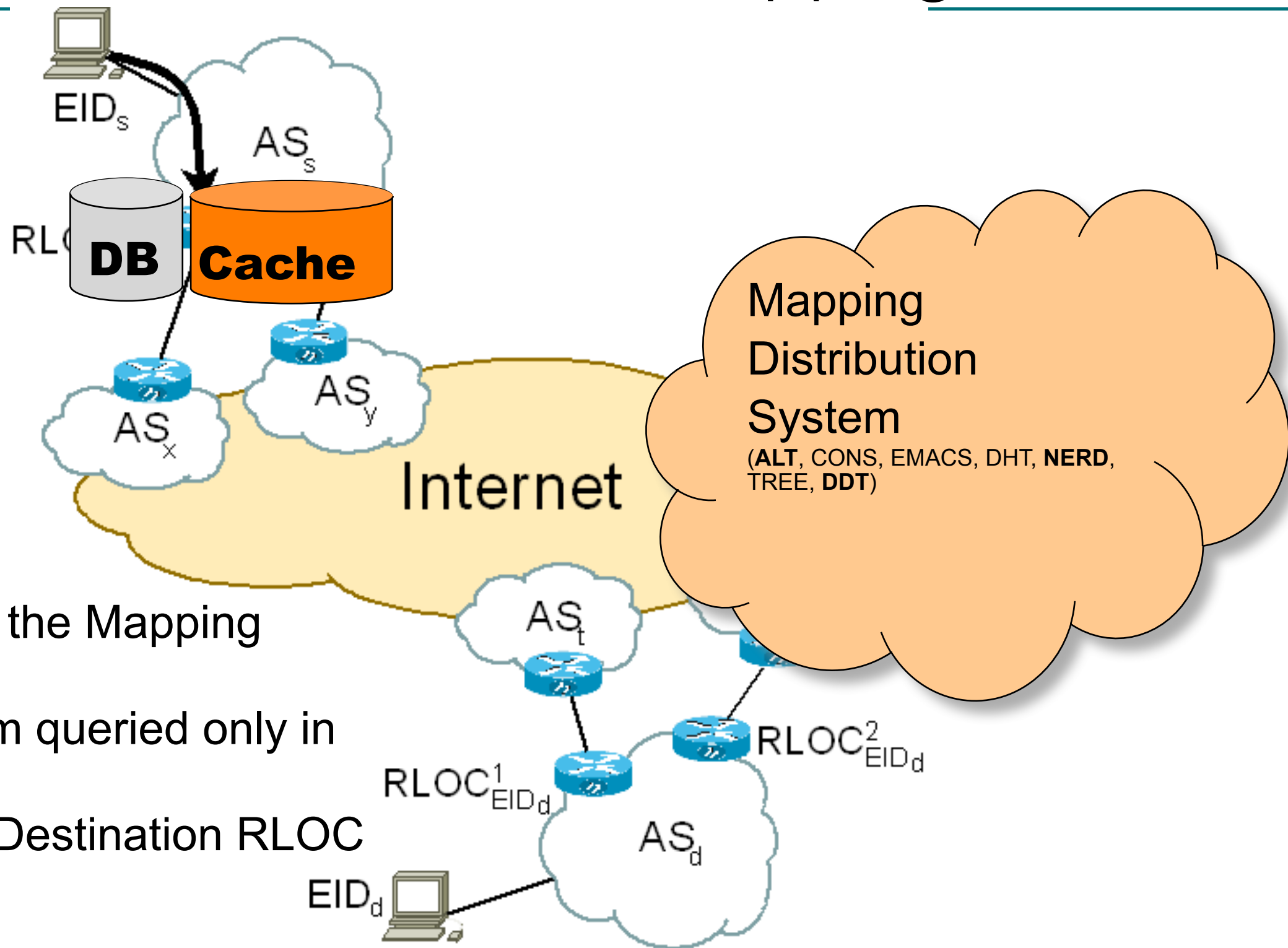
Where does LISP find the Mappings?



- Mapping Distribution System:

- Queried to retrieve mappings
- Used to select Destination RLOC

Where does LISP store the Mappings?



LISP-Cache:

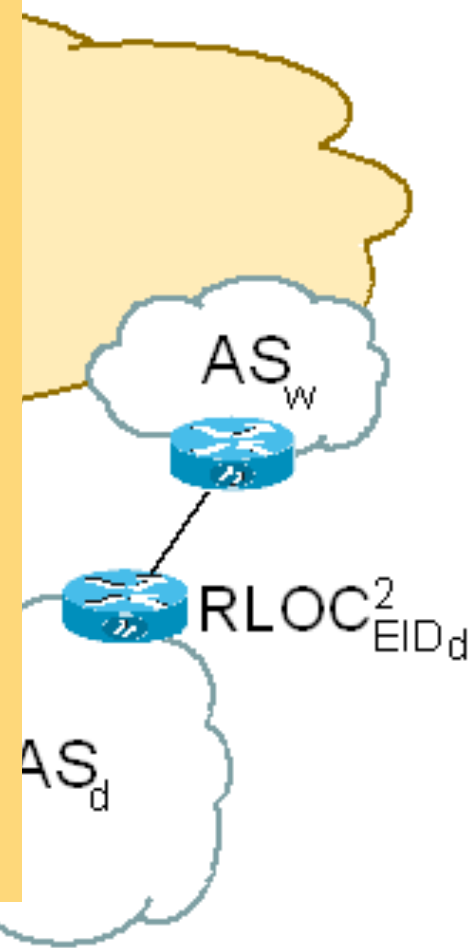
- Queried before the Mapping system
- Mapping system queried only in case of miss
- Used to select Destination RLOC

Map & Encap Operations: destination side

[illegible]

ding using RLOCs
Encapsulation)

ding using EIDs
nal IP packets)



Consistency Checks:

- Check DB: Am I the correct RLOC for the destination EID?

RIPE

A large graphic element on the left side of the slide. It features the word "RIPE" in a light teal, sans-serif font. To the right of the text are two vertical white bars of different heights and two horizontal teal bars of different lengths, creating a stylized cross or plus sign.

LISP Header

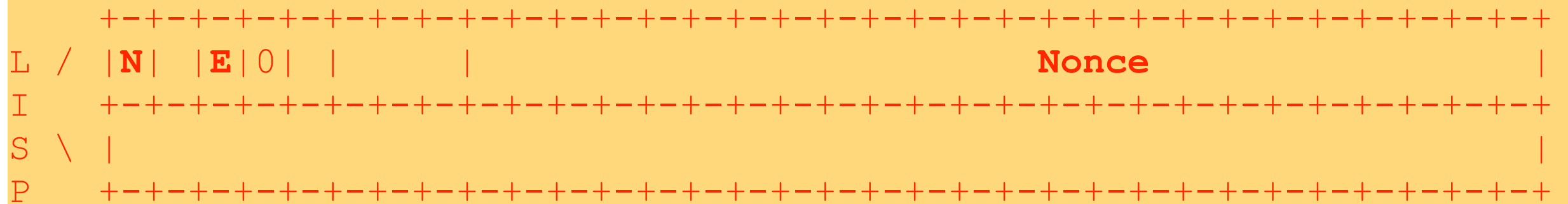
RFC 6830



LISP Header

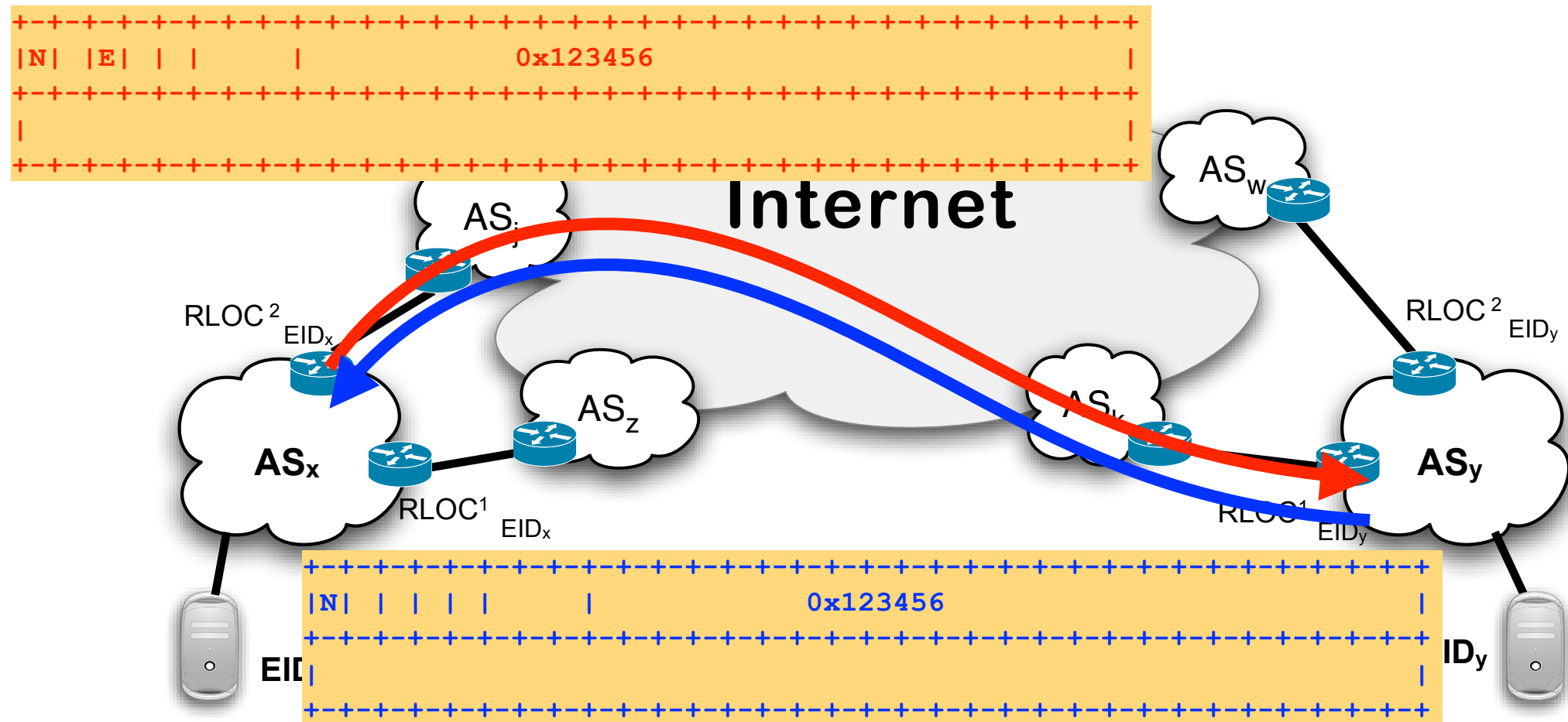
	0																1																2																3															
	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1																						
/	Version								IHL								Type of Service																Total Length																															
/																					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

LISP Header: Nonce

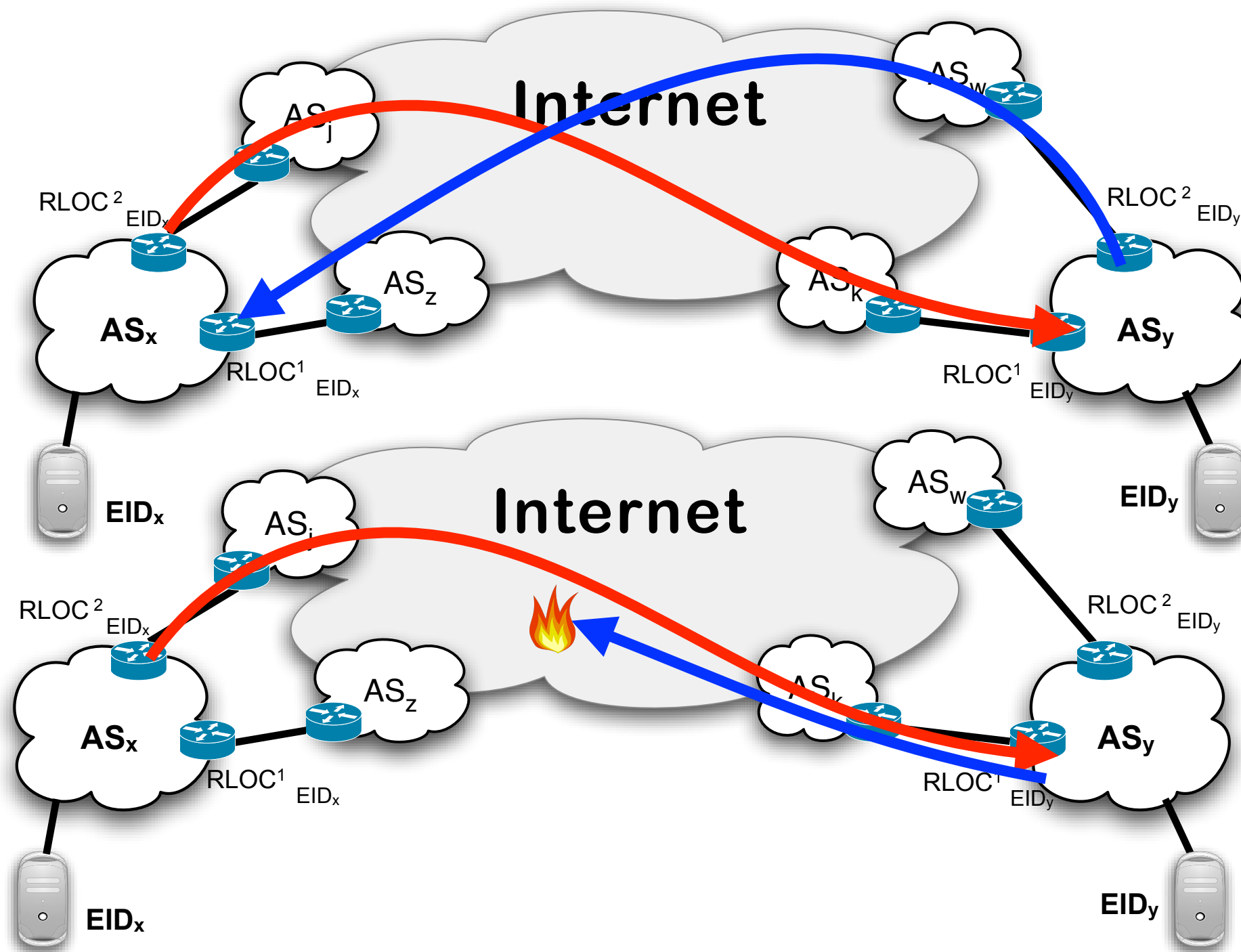


- N-bit: Nonce Present bit
 - V-bit must be 0
- E-Bit: Echo-Nonce Request Bit
- Echo-Nonce Algorithm:
 - E-Bit & N-Bit Set
 - Echo back the Nonce
 - Purpose: Reachability

LISP Header: Nonce in practice

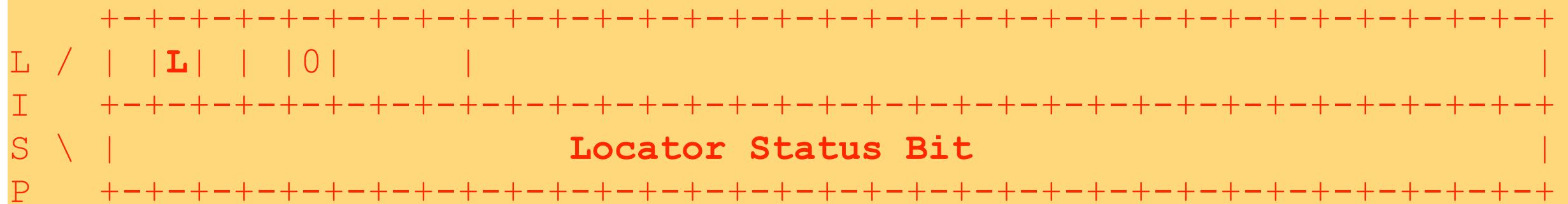


Echo-Nonce may be tricky....



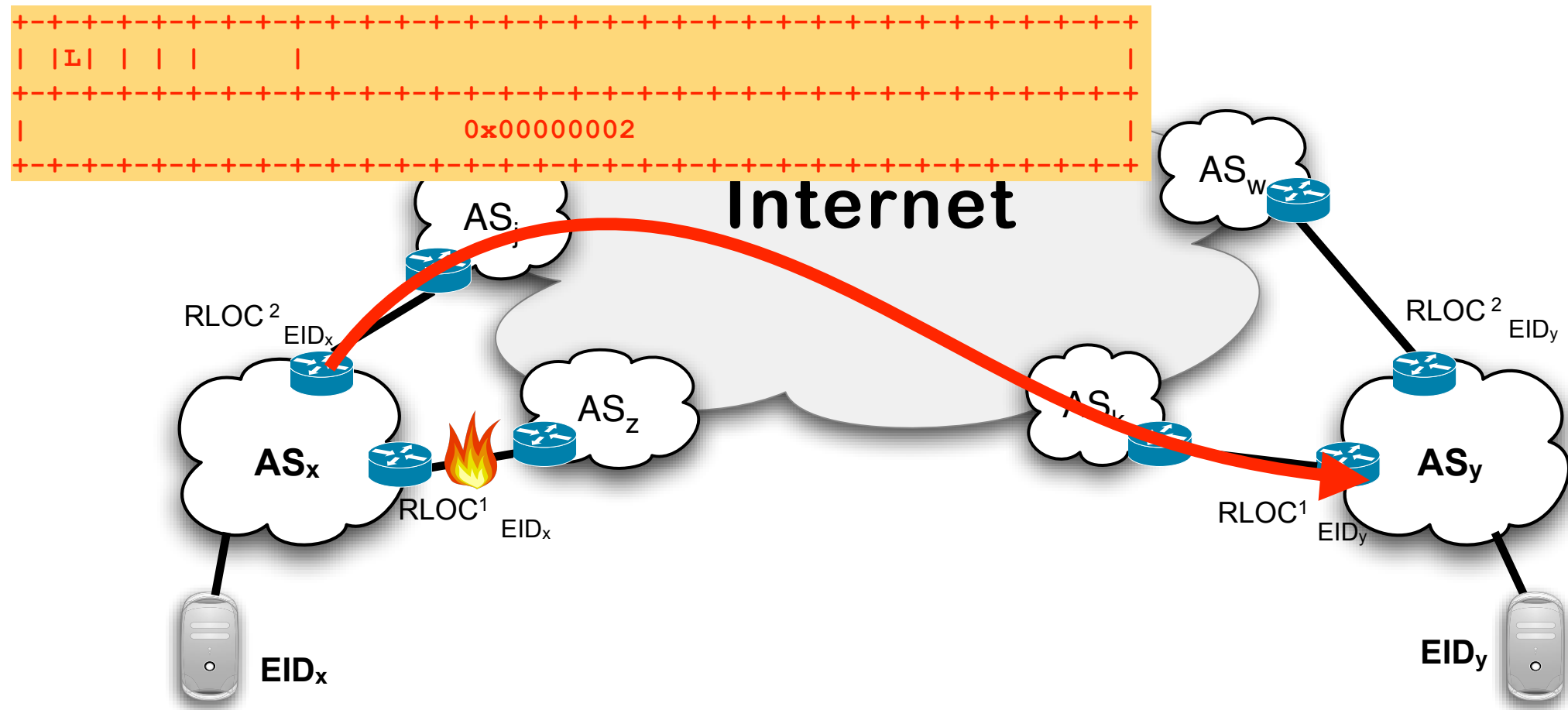
- How to understand the difference????

LISP Header: Loc-Status

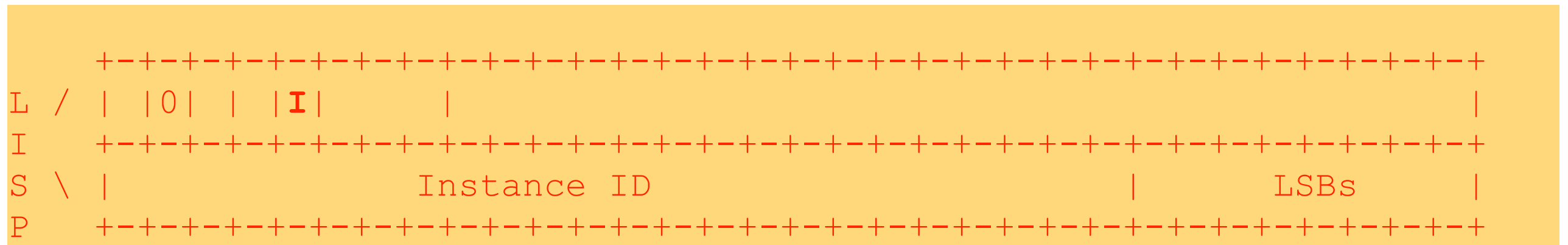


- L-bit: Locator-Status Present bit
 - I-bit must be 0
- Purpose: provide in-band information about RLOC status/reachability
 - How to define the status/reachability?

LISP Header: Loc-Status in practice



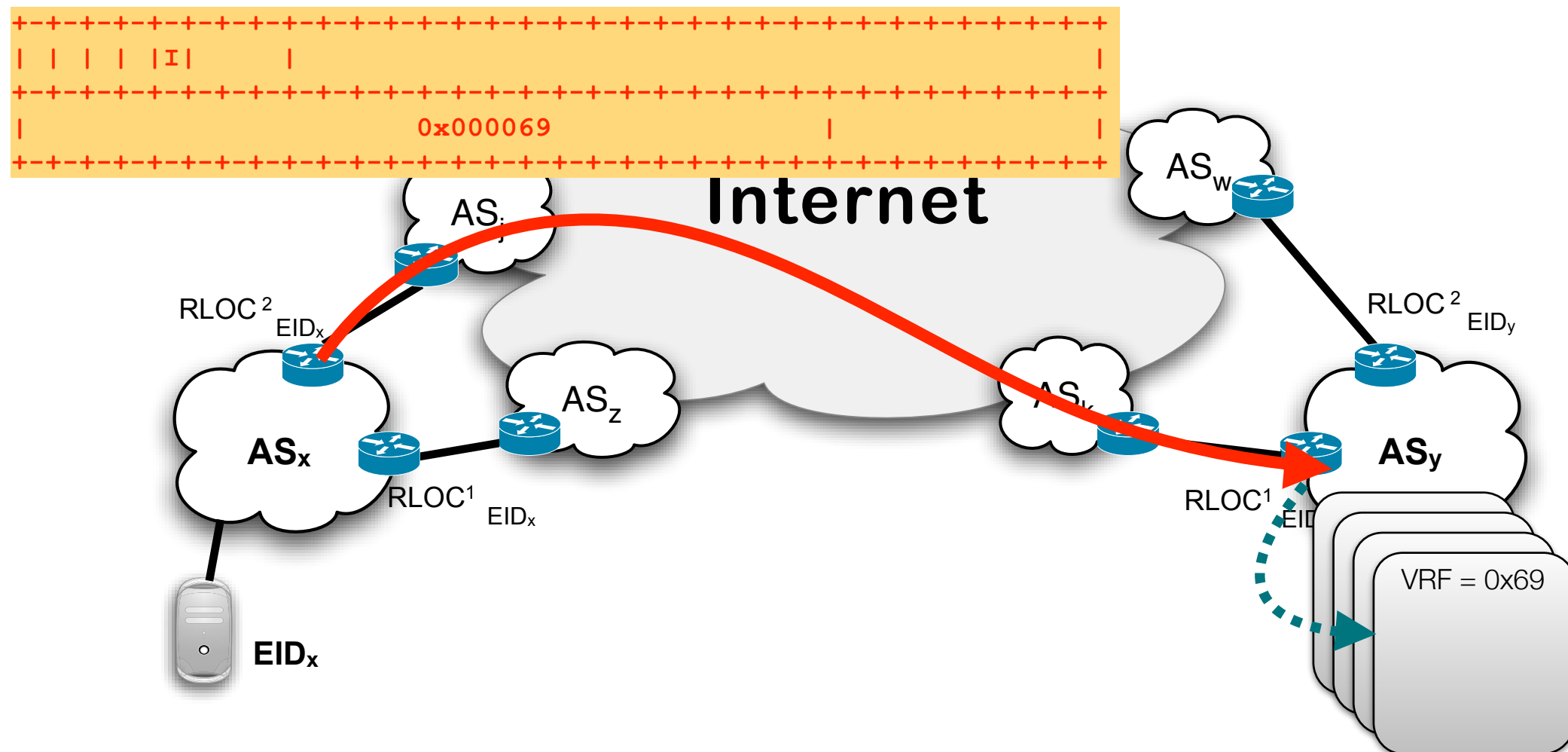
LISP Header: Instance ID



- I-bit: Instance ID Present bit
 - L-bit must be 0
- Second long-word split in two fields:
 - 24 high-order bits: Instance ID
 - 8 low-order bits: Loc-Status bits
- Purpose: Virtualization (VPN) and Segmentation



LISP Header: Instance ID in practice



- ETR uses IID as VRF selector

LISP Header: Versioning



- V-bit: Map-Version Present bit
 - E-bit must be 0
- RFC 6834 (later)

RIPE

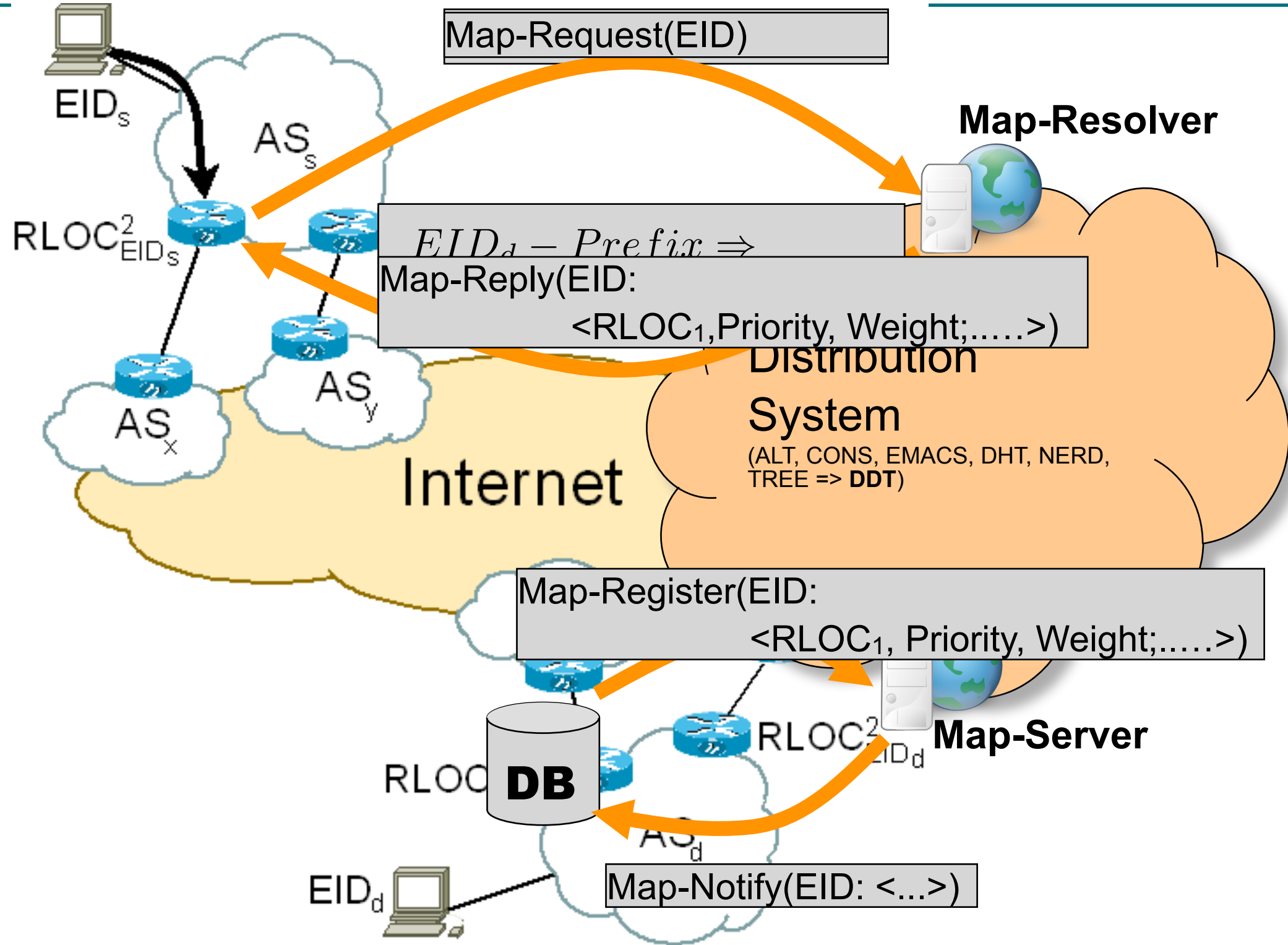
A large graphic element on the left side of the slide. It features the word "RIPE" in a light blue, sans-serif font. To the right of the text are two vertical white bars of different heights and two horizontal light blue bars of different lengths, creating a stylized cross or plus sign shape.

LISP Map-Server

RFC 6833

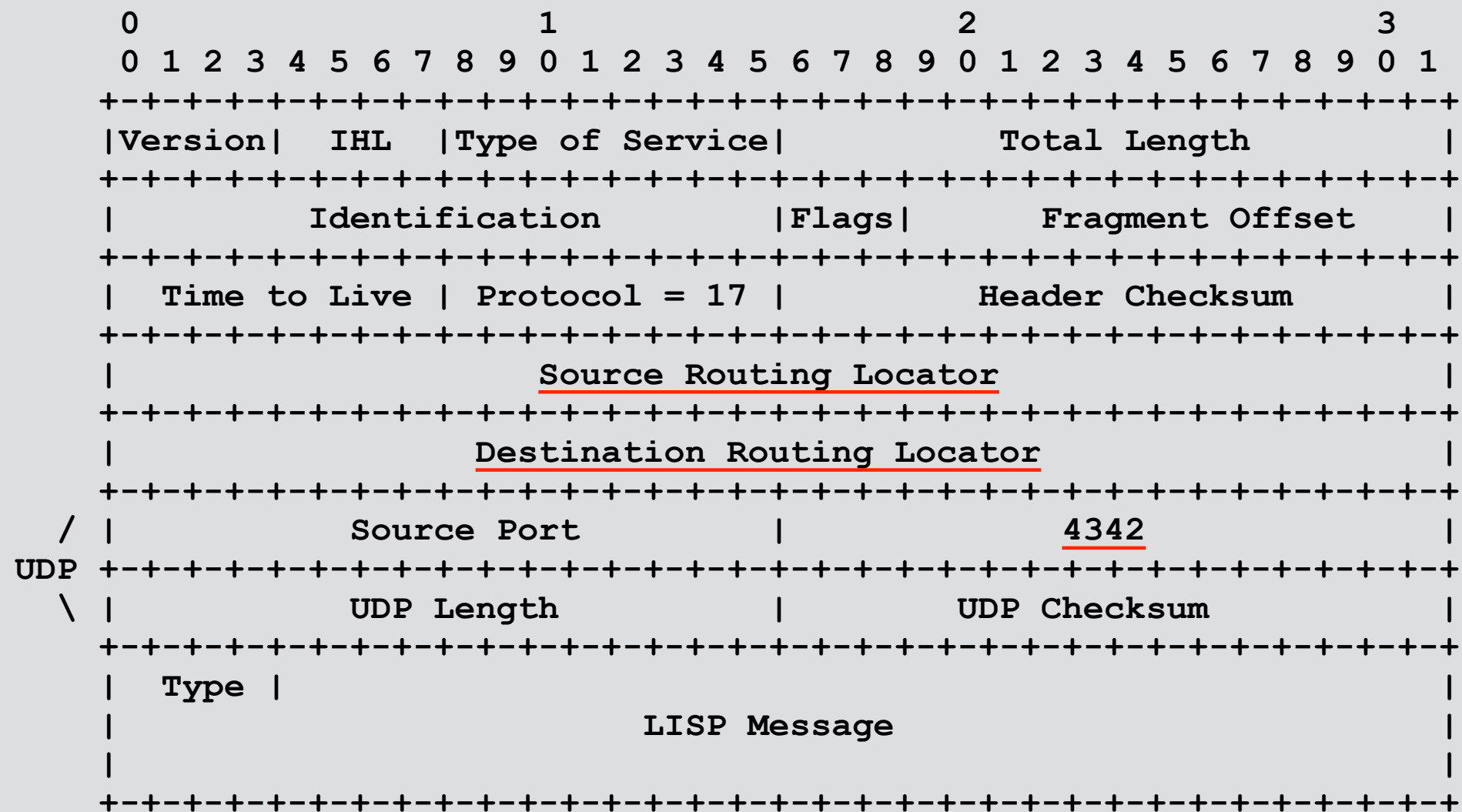


MS: A Mapping System API



LISP Control Message Outer Header

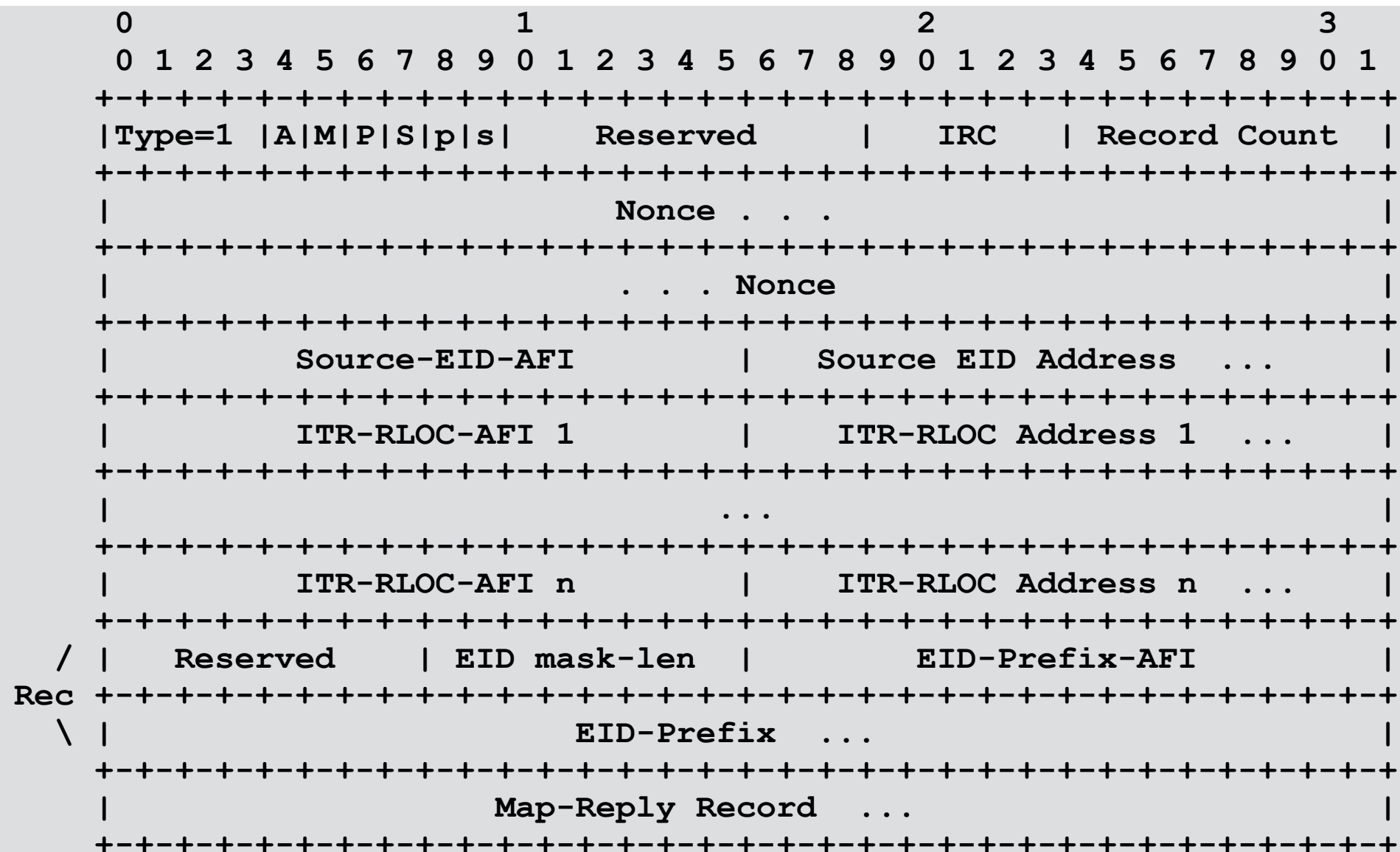
- RLOC space
- UDP Destination port 4342



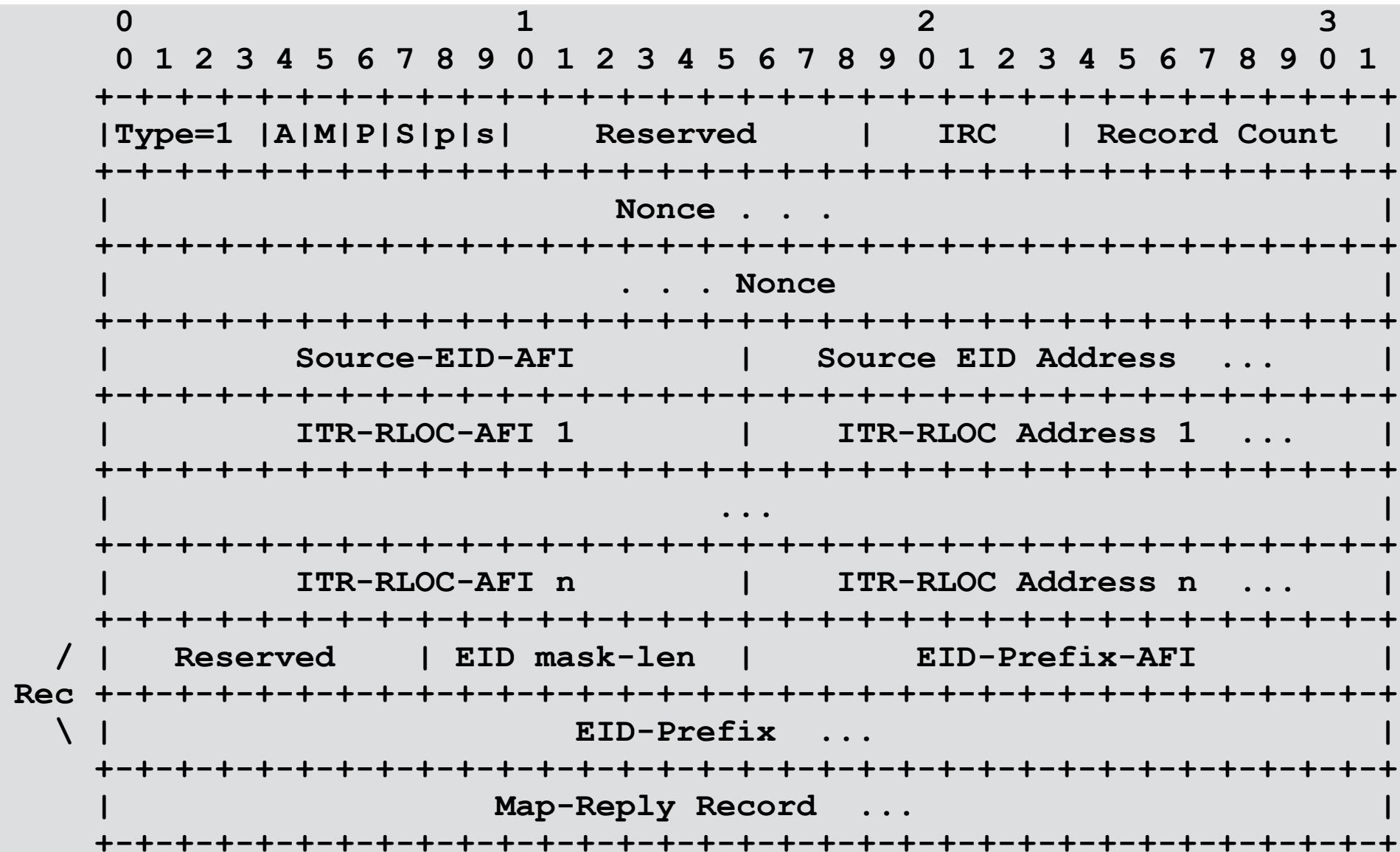
- Reserved: 0 b'0000'
- LISP Map-Request: 1 b'0001'
- LISP Map-Reply: 2 b'0010'
- LISP Map-Register: 3 b'0011'
- LISP Map-Notify: 4 b'0100'
- LISP Encapsulated Control Message: 8 b'1000'



Map-Request



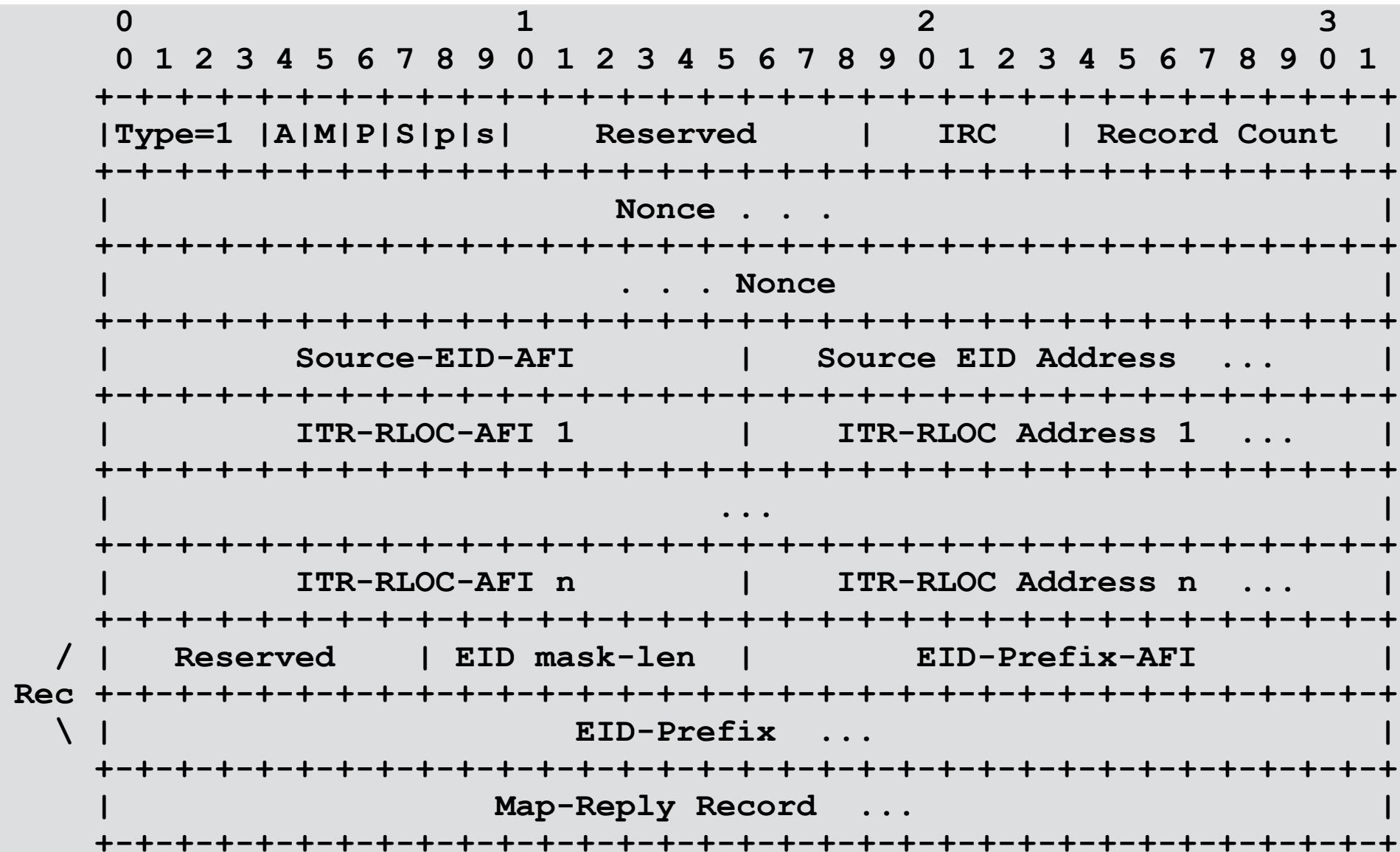
Map-Request: Flags



- A - Authoritative bit: Request authoritative answer.
- M - Map Data bit: Indicates Map-Reply record present
- P - Probe bit: Message to be used as locator reachability probe (sent directly to xTRs)
- S - Solicit Map Request: Solicit Map Request
- p - PITR bit: set when PITR sends the request (RFC 6832)
- s - SMR-invoked Map-Request: set when Map-Request has been solicited by SMR.



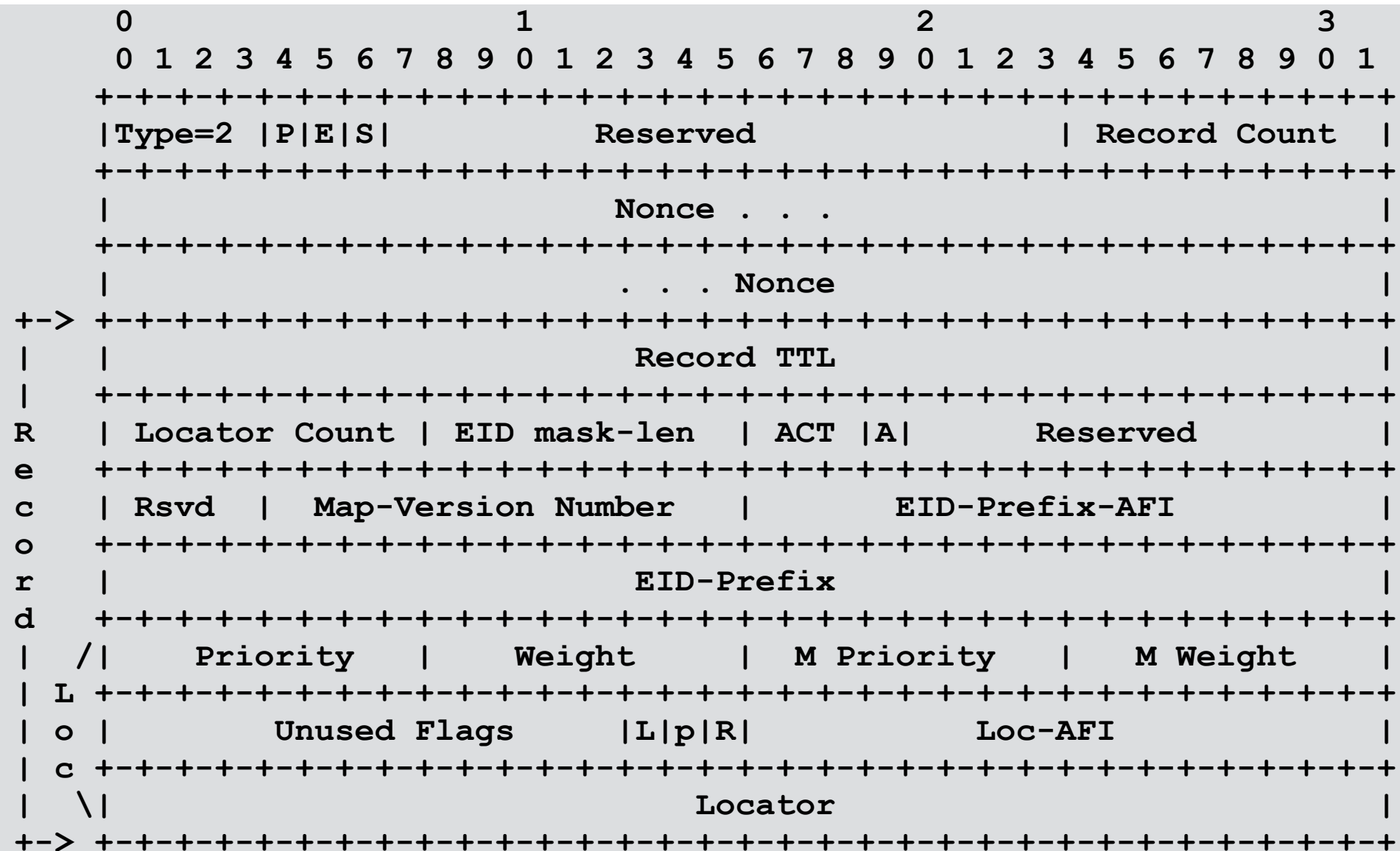
Map-Request: The not flags part.....



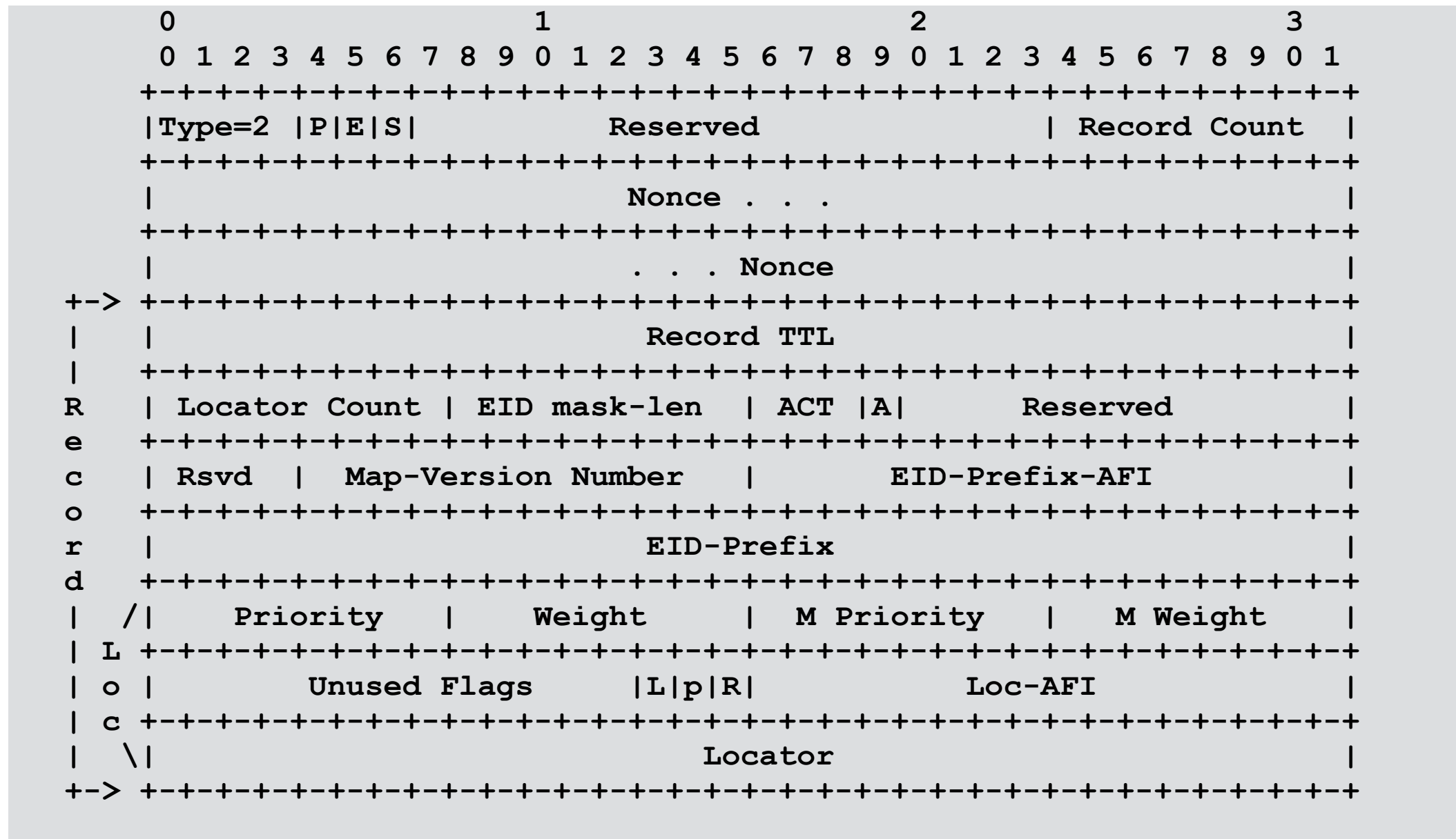
- IRC - ITR-RLOC Count: number of <ITR-RLOC-AFI, ITR-RLOC-Addr> present
- Record Count: Number of actual Requests in the message
- Nonce: Security Random value
- Source EID (AFI & Addr)
- Rec - AFI, Prefix Mask Length, EID prefix request



Map-Reply



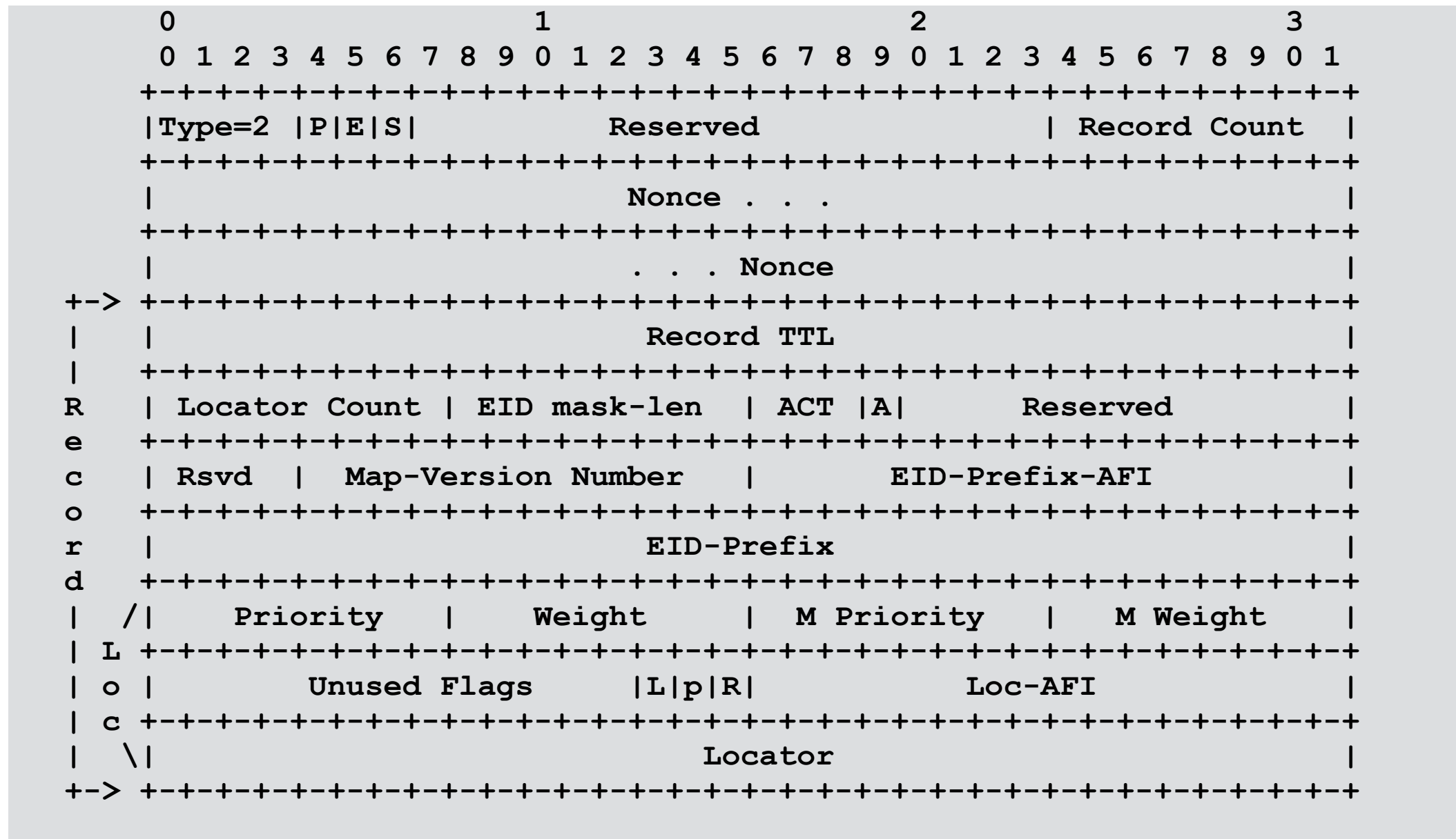
Map-Reply: Flags + Nonce + RecCount



- P - Probe bit: Message is a reply to a locator reachability probe (sent directly to xTRs)
- E - Echo None Bit: ETR is Echo-Nonce Capable
- S - Security bit: If present security material is appended to the packet
- Record Count: number of mappings
- Nonce: Security Random value



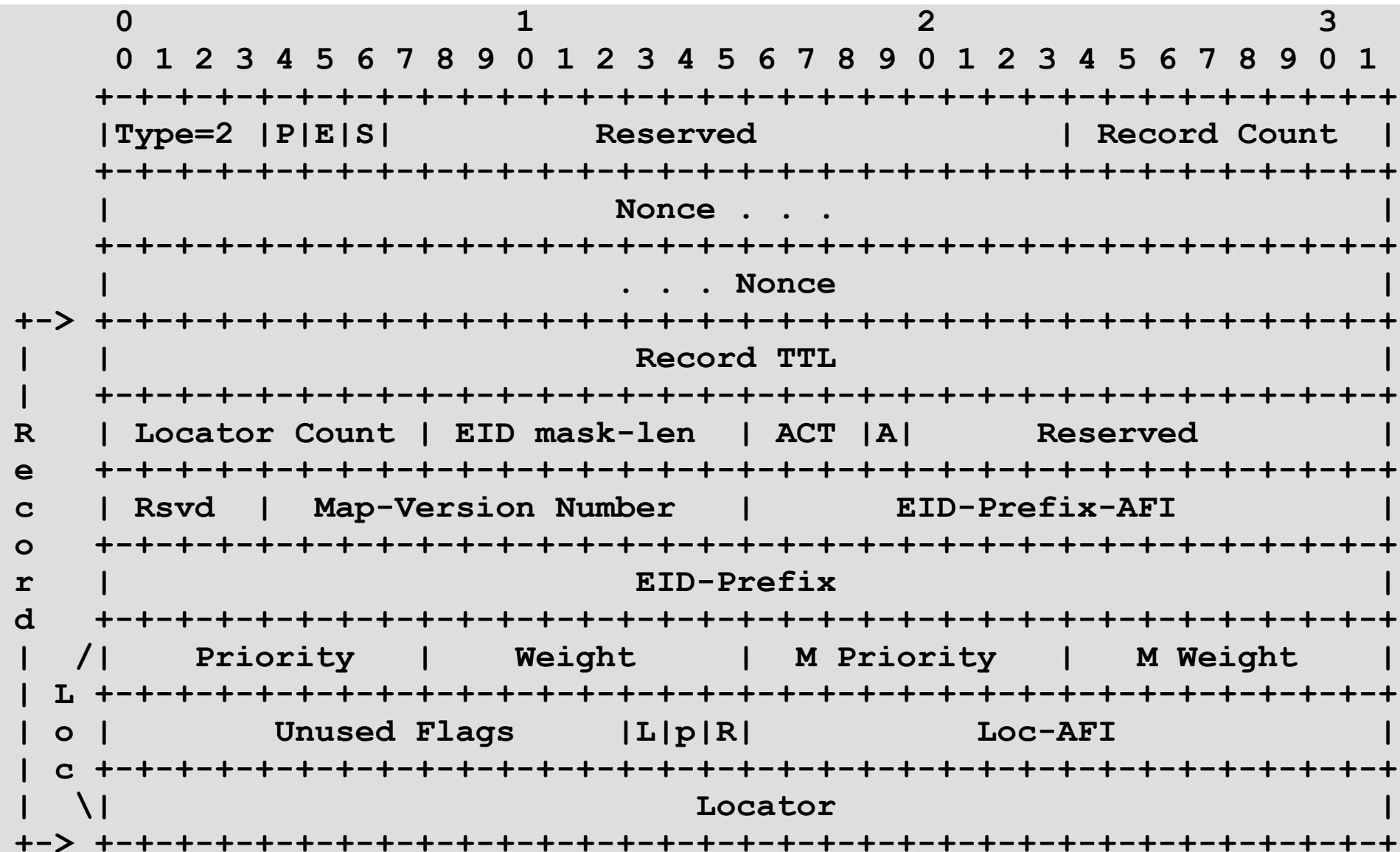
Map-Reply: Record



- TTL - Time To Live: How many minutes the mapping can be stored
 - 0: Remove
 - 0xFFFFFFFF: local decision
- Loc Count: Number of locators
- EID Mask Length
- ACT: Action bits (No-Action; Forward Native; Send-Map-Request; Drop)
- A - Authoritative Bit: set if the sender is an ETR



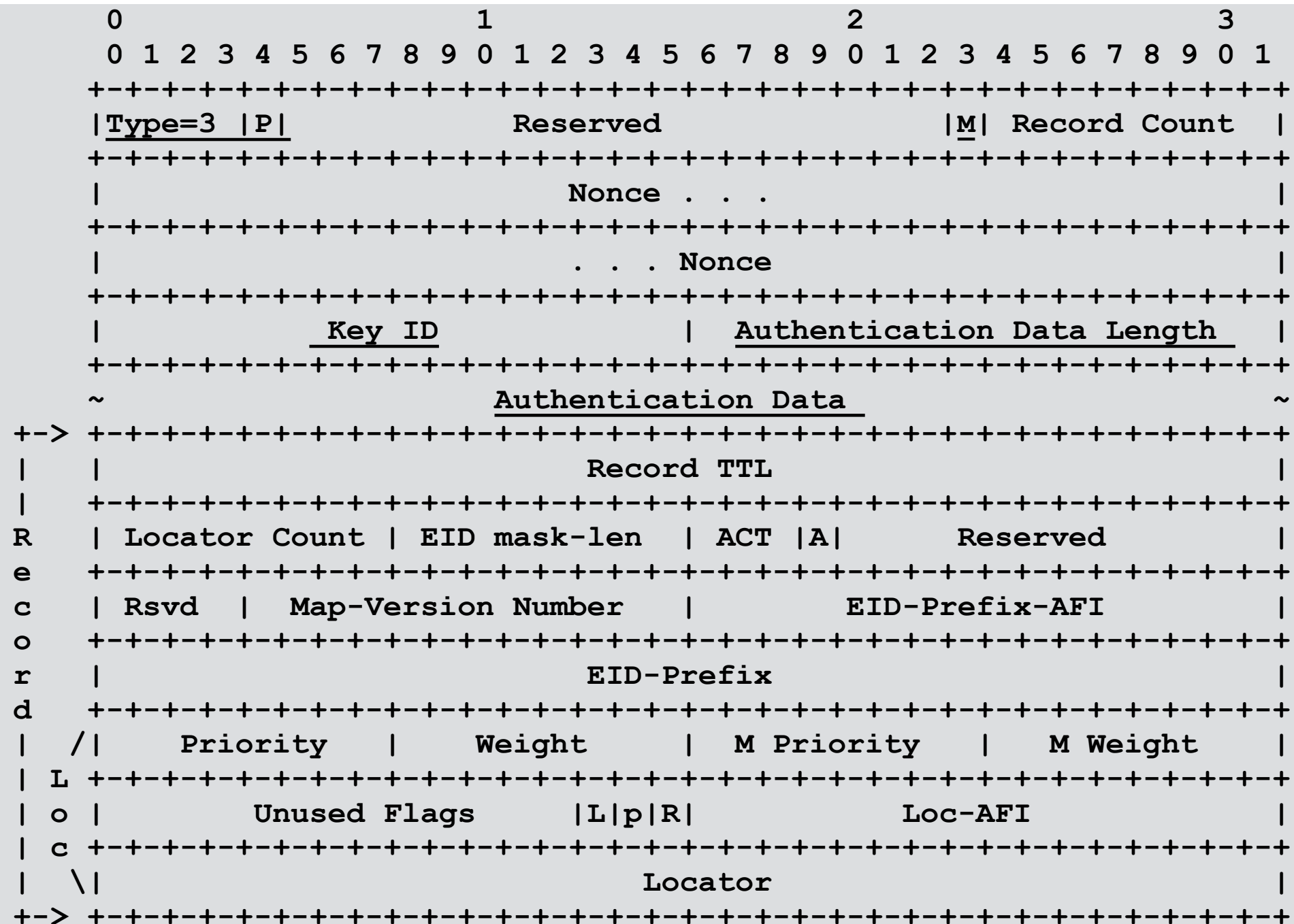
Map-Reply: Record



- Map Version Number
- EID (AFI & Prefix)
- Priority & Weight (also for multicast)
- L: Local Locator
- p: Probed Locator
- R: Route for Locator
- Locator (AFI & Addr)



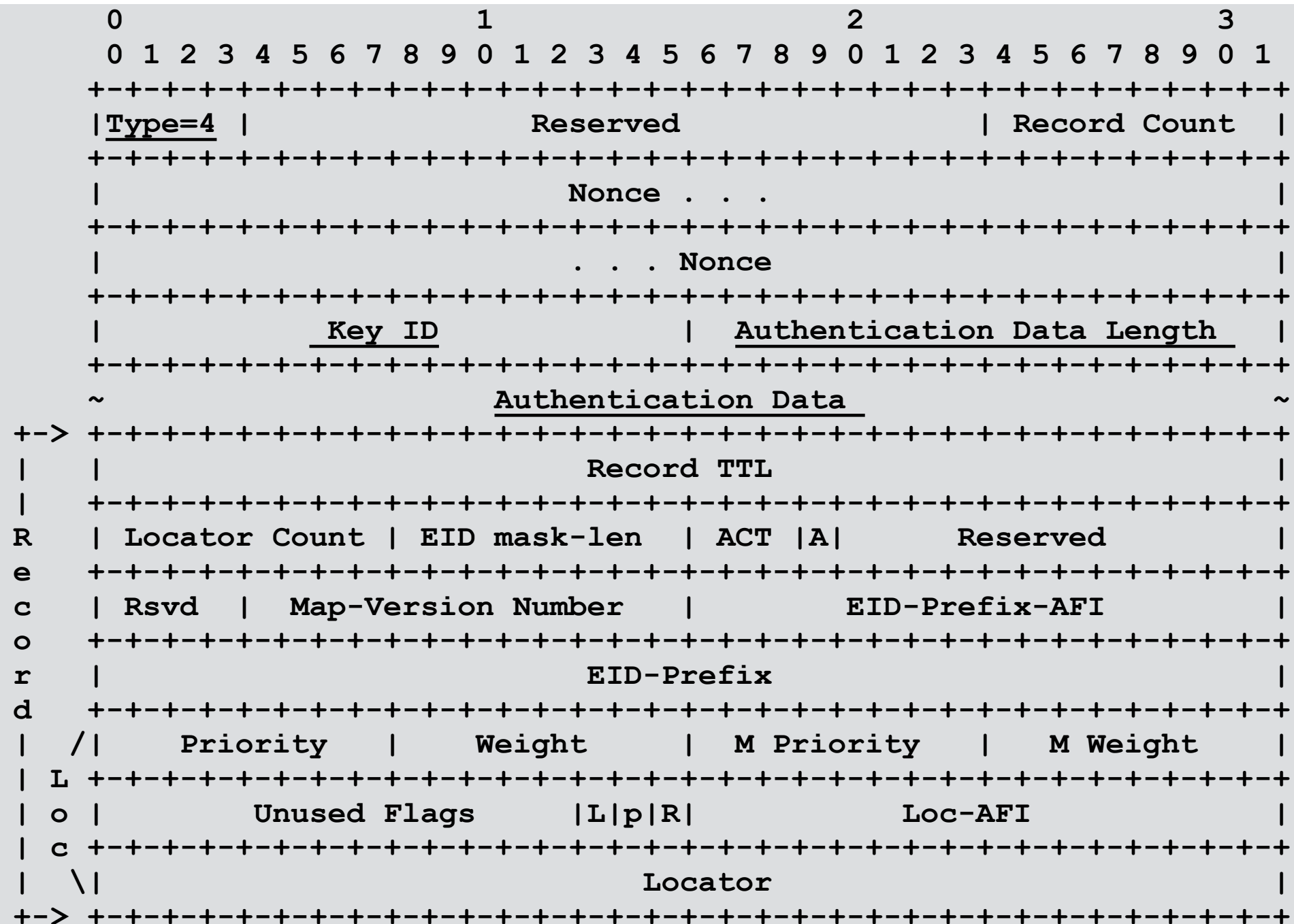
Map-Register



- P - Proxy Bit
- M - Map-Notify requested:



Map-Notify



LISP Control Plane

- LISP+ALT (RFC 6836)
- LISP-NERD (RFC 6837)
- LISP-DDT (draft-ietf-lisp-ddt-01.txt)

RIPE

A large graphic element on the left side of the slide. It features the word "RIPE" in a light teal, sans-serif font. To the right of the text are two vertical white bars of different heights and two horizontal teal bars of different lengths, creating a stylized cross or plus sign.

LISP NERD

RFC 6837



LISP-NERD

Not-so-novel Eid to Rloc Database

- Centralized approach based on one (or more) authority(ies)

- xTRs periodic update request
- Updates contain all mappings

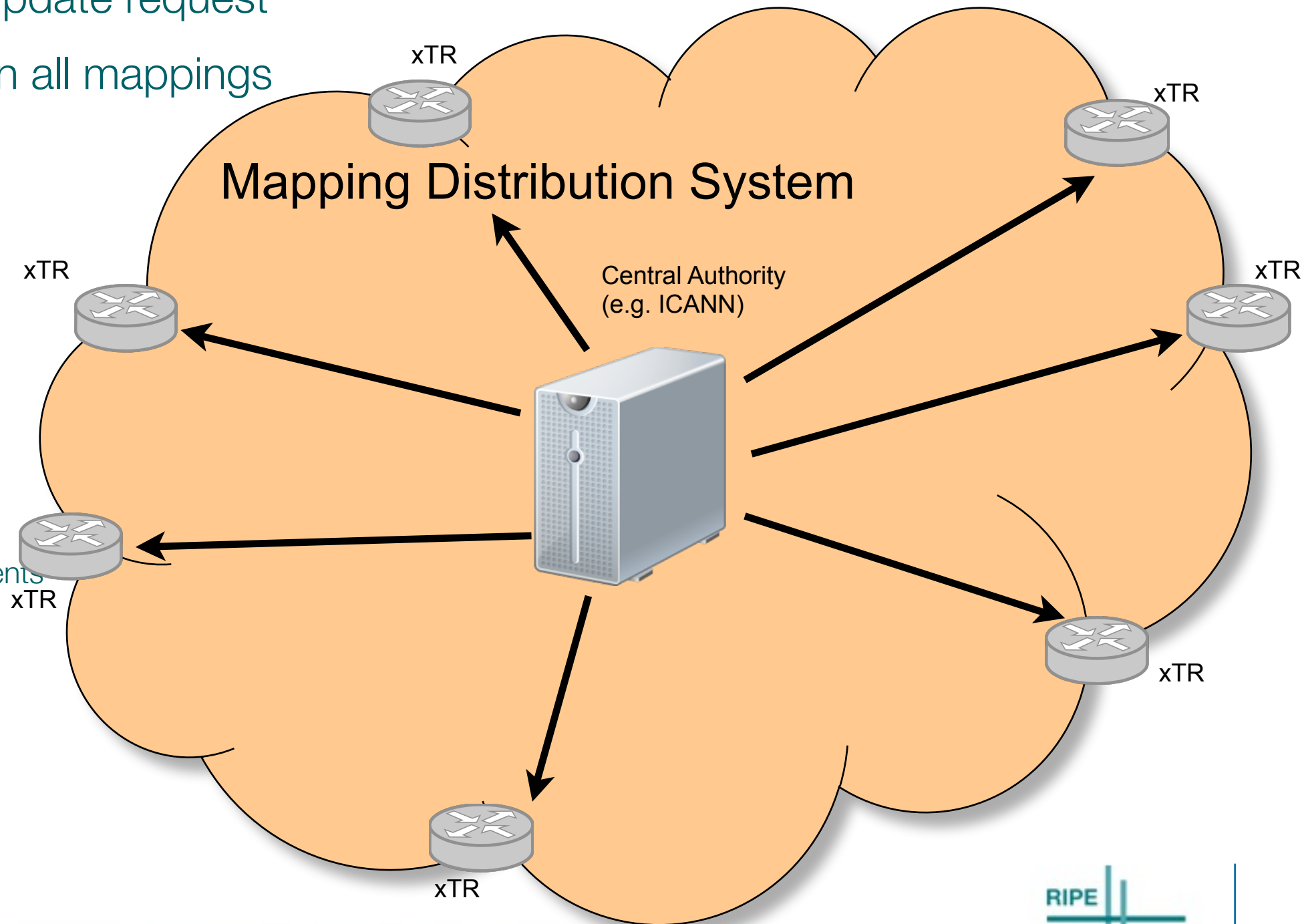
- Pros

- Simple

- Bootstrap
- Data Center
- Private Deployments

- Cons

- Not scalable
- Static



RIPE



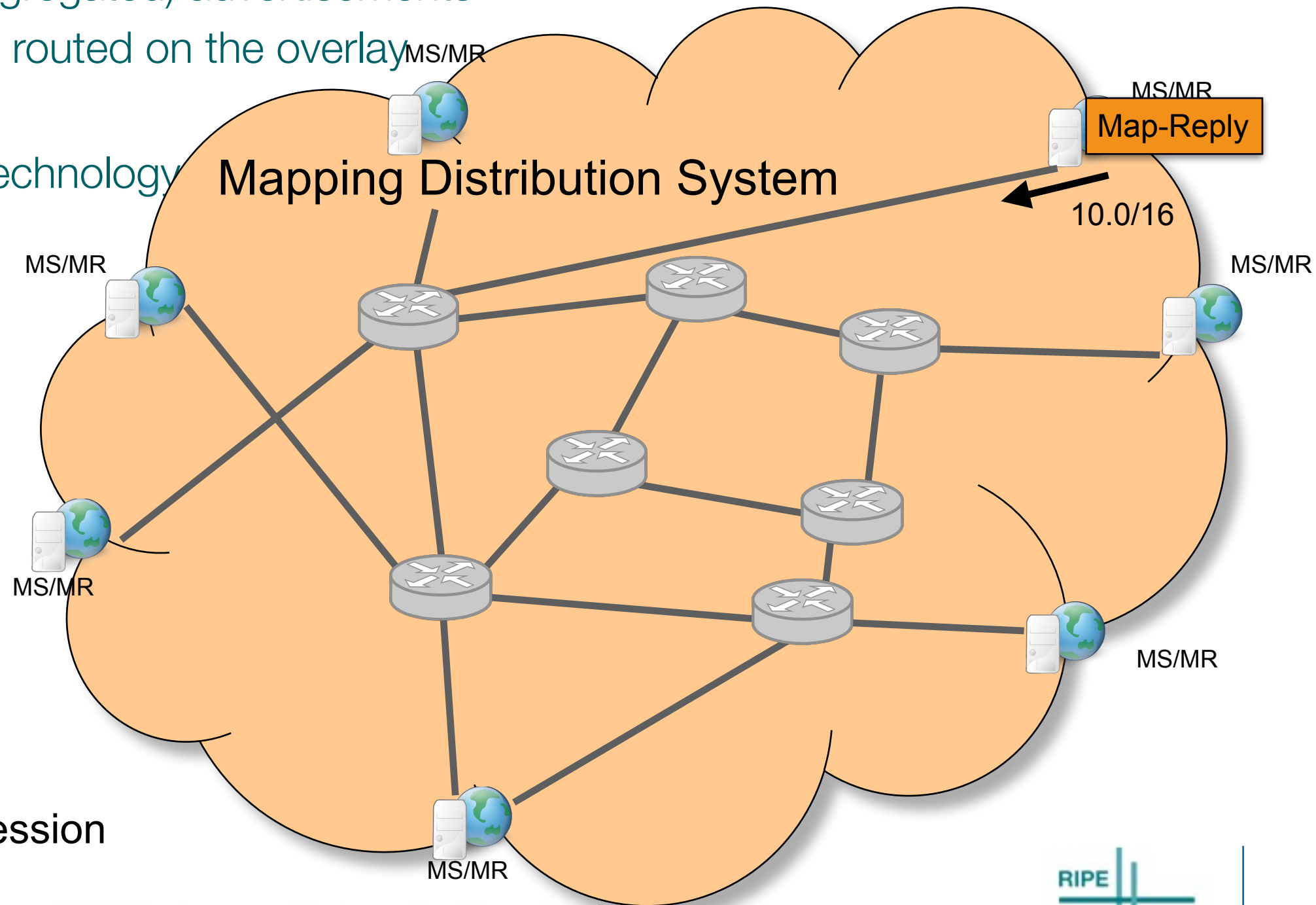
LISP+ALT

RFC 6836



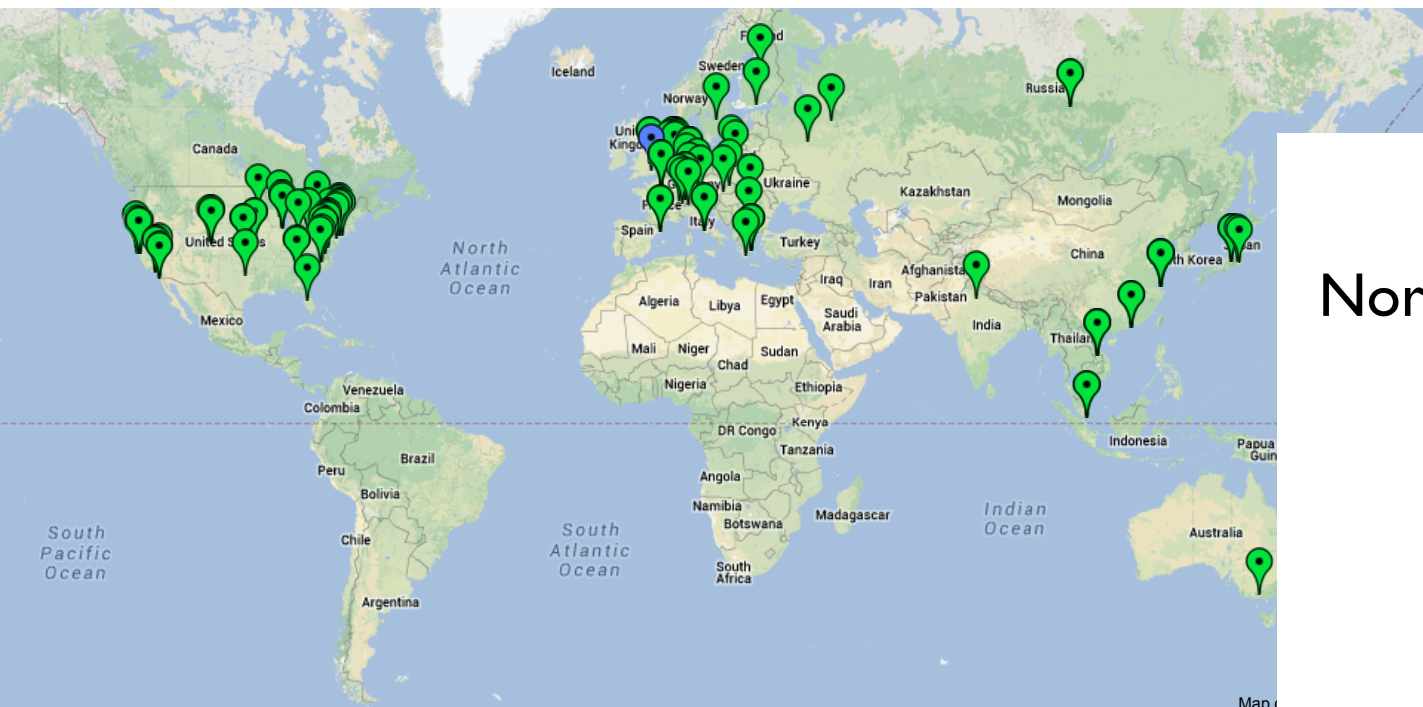
LISP+ALT

- BGP over GRE overlay
 - EID-Prefix (aggregated) advertisements
 - Map-Request routed on the overlay^{MS/MR}
- Pros
 - Well-known technology
- Cons
 - PITA



LISP+ALT Failure

- Used on an international testbed (www.lisp4.net)
 - Rapidly resulted in very cumbersome maintenance



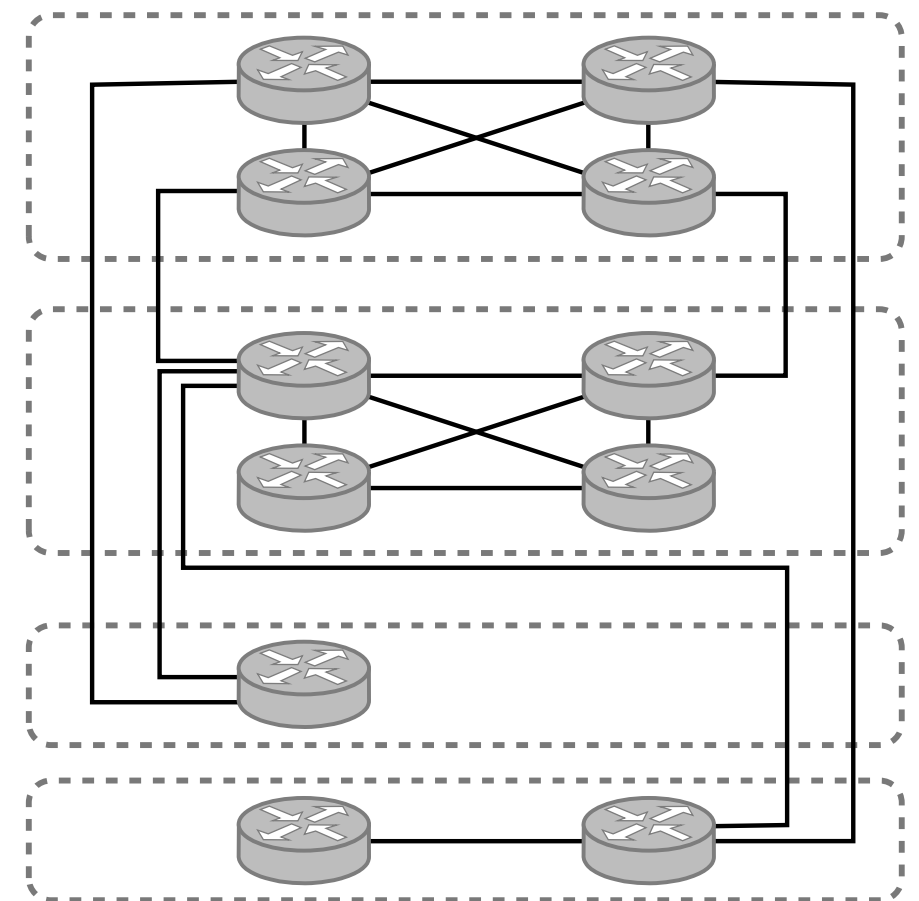
www.lisp4.net

North America

Europe

Asia

Latin America



LISP+ALT Topology early 2012



RIPE

A large graphic element on the left side of the slide. It features the word "RIPE" in a bold, teal, sans-serif font. To the right of the text are two vertical white bars of different heights and two horizontal teal bars of different lengths, creating a stylized cross or plus sign.

LISP DDT

draft-ietf-lisp-ddt-03.txt



LISP-DDT: Delegated Database Tree

- DNS-Like logical hierarchy

- EID-Prefix more specific delegated
- Delegation is statically configured

- Pros

- Easy maintenance
- Economics
- Security

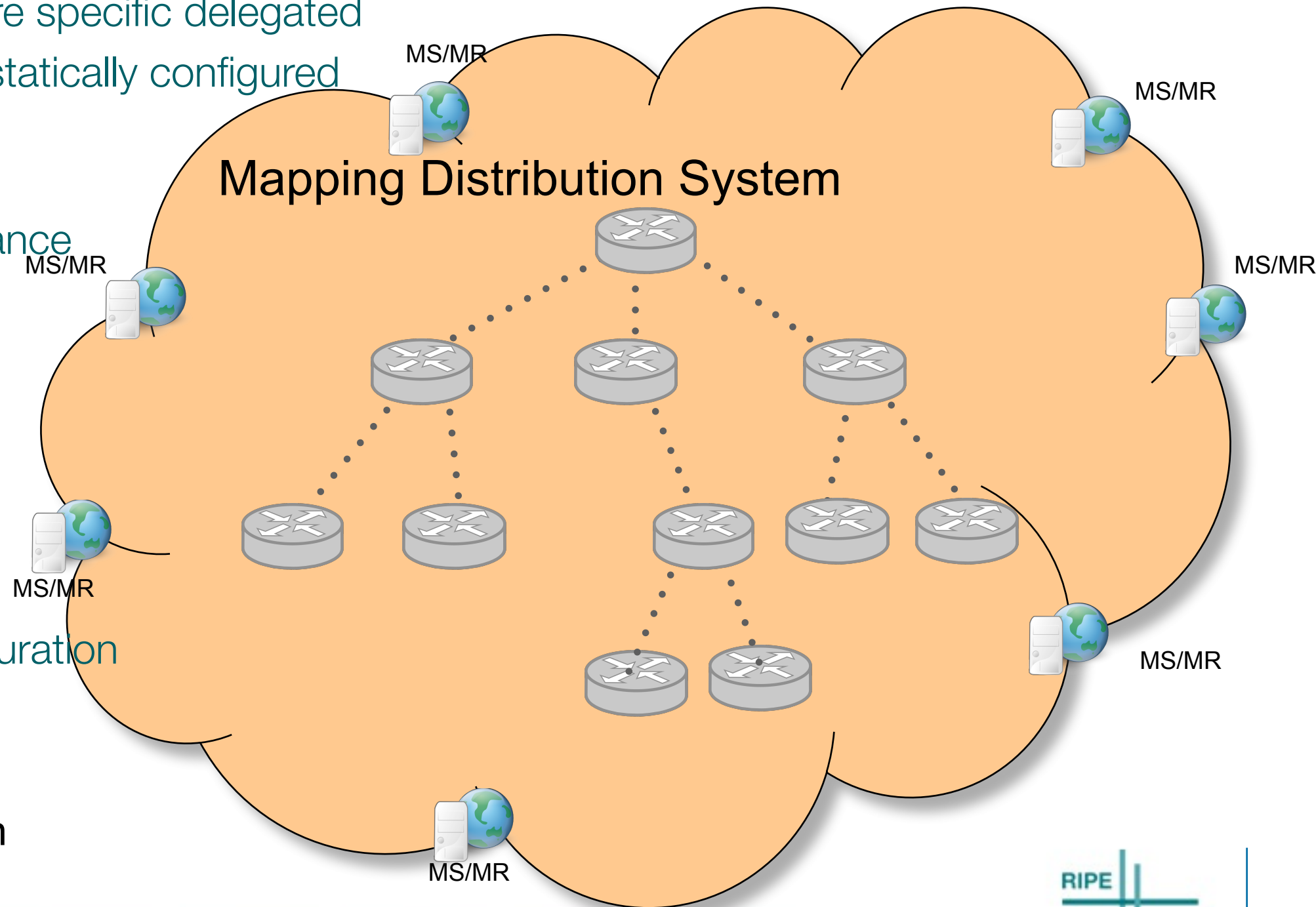
- Cons

- manual configuration
- static tree

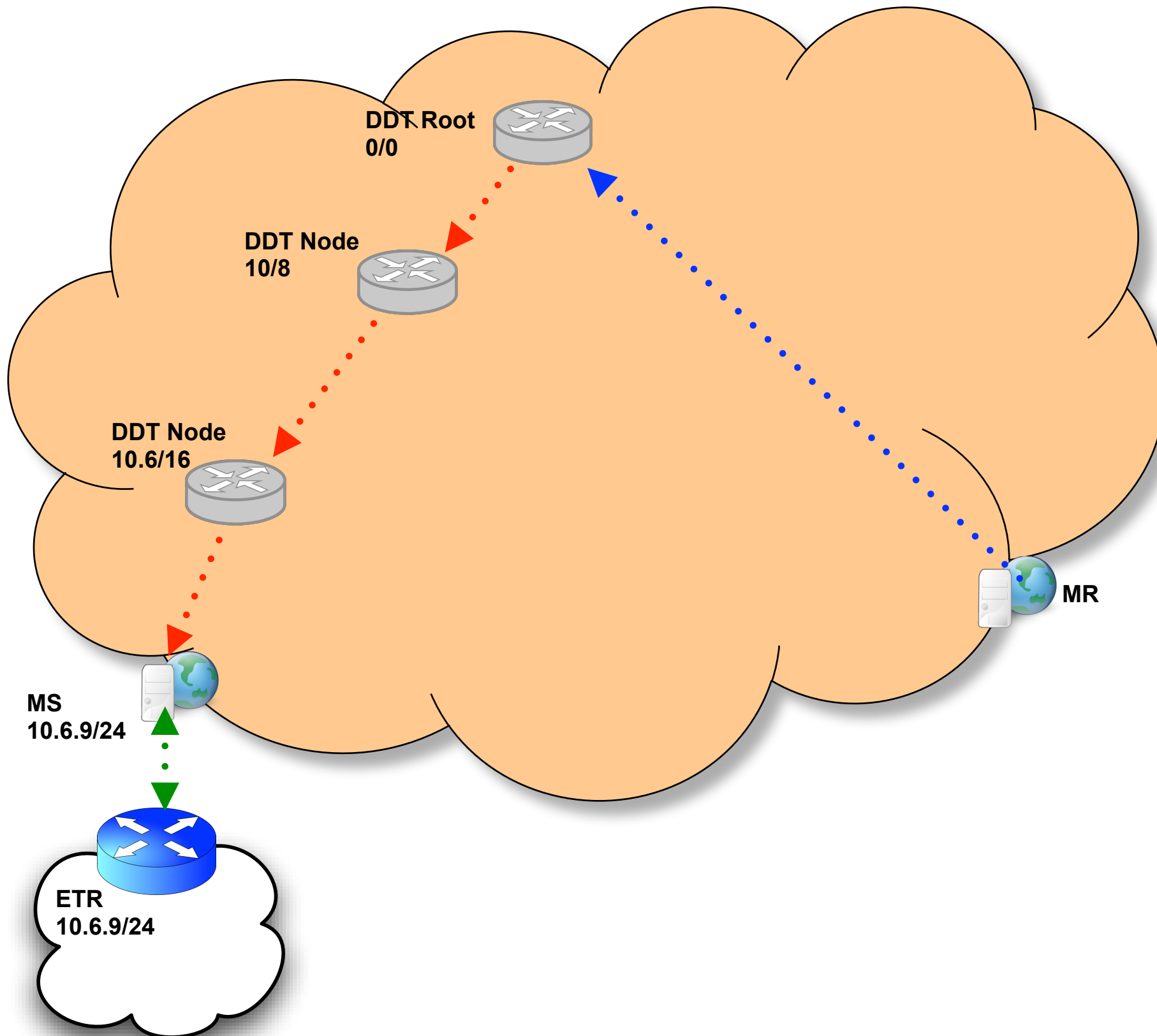


DDT Node

... Logical Delegation



LISP-DDT



.... **MR**

Static pointer to root

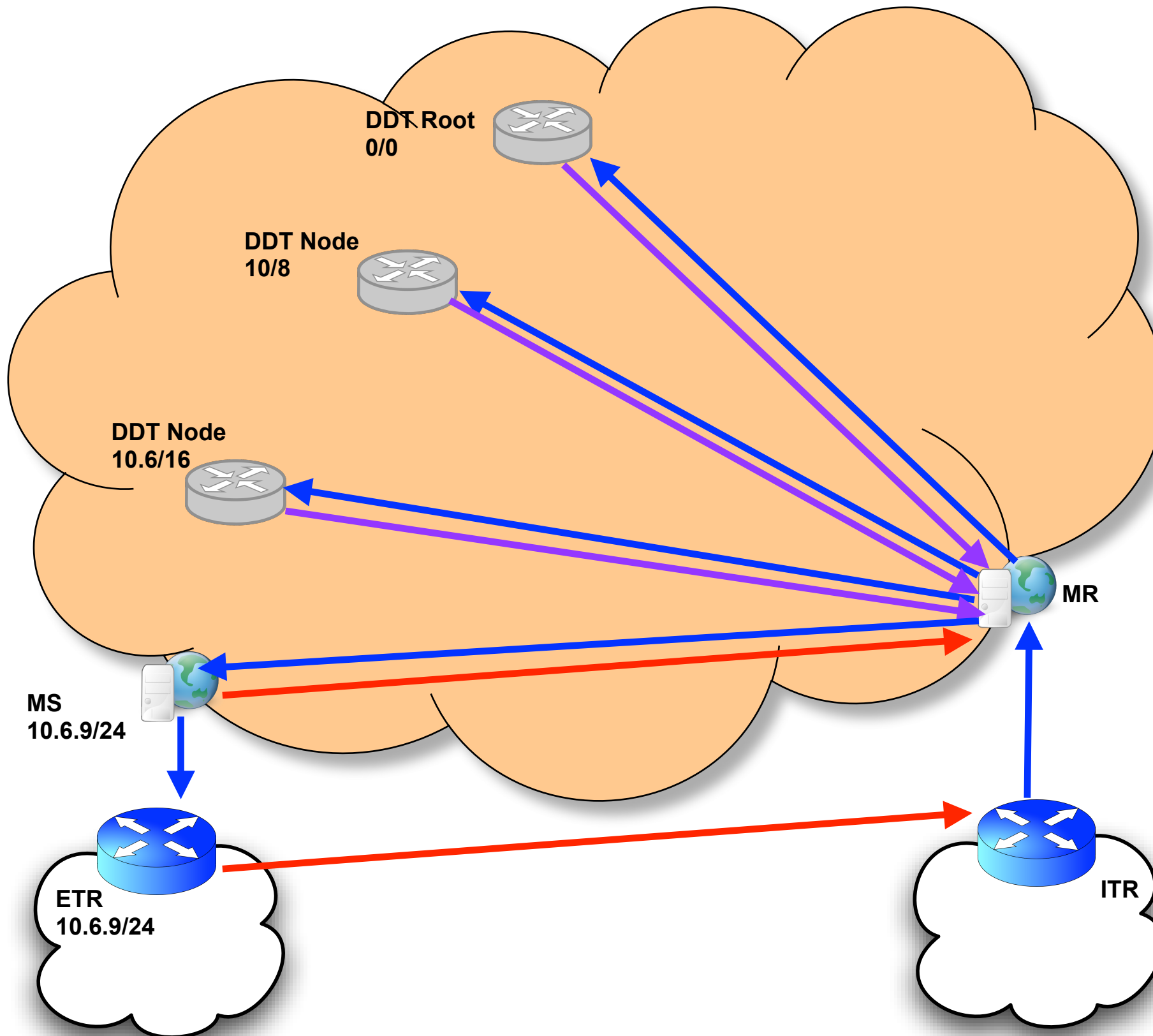
.... **DDT Nodes**

Static pointers to children
authoritative of more-
specific (leafs are MS)

.... **ETR**

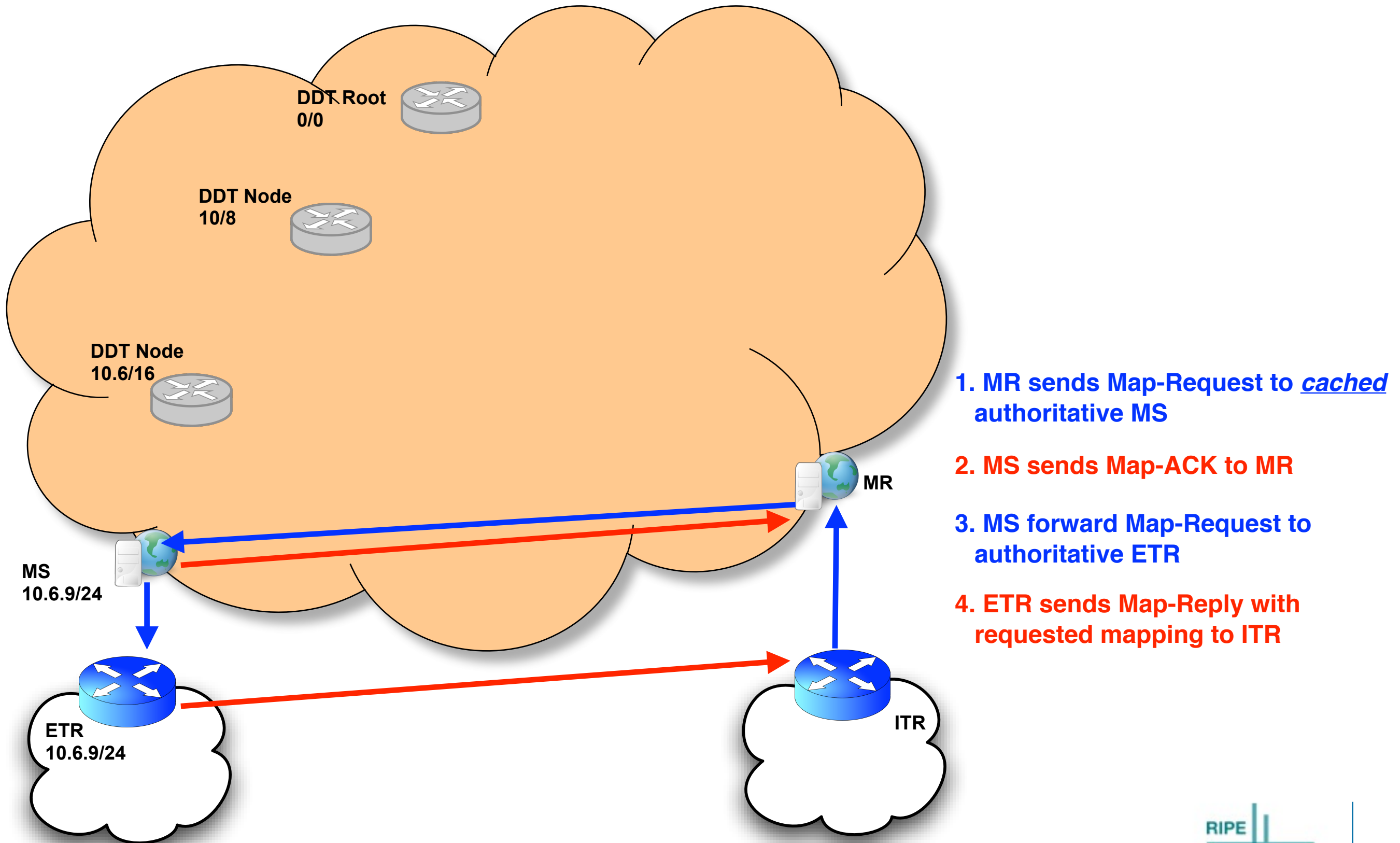
Registers to statically
configured MS

LISP-DDT: Mapping Retrieval (cold start)

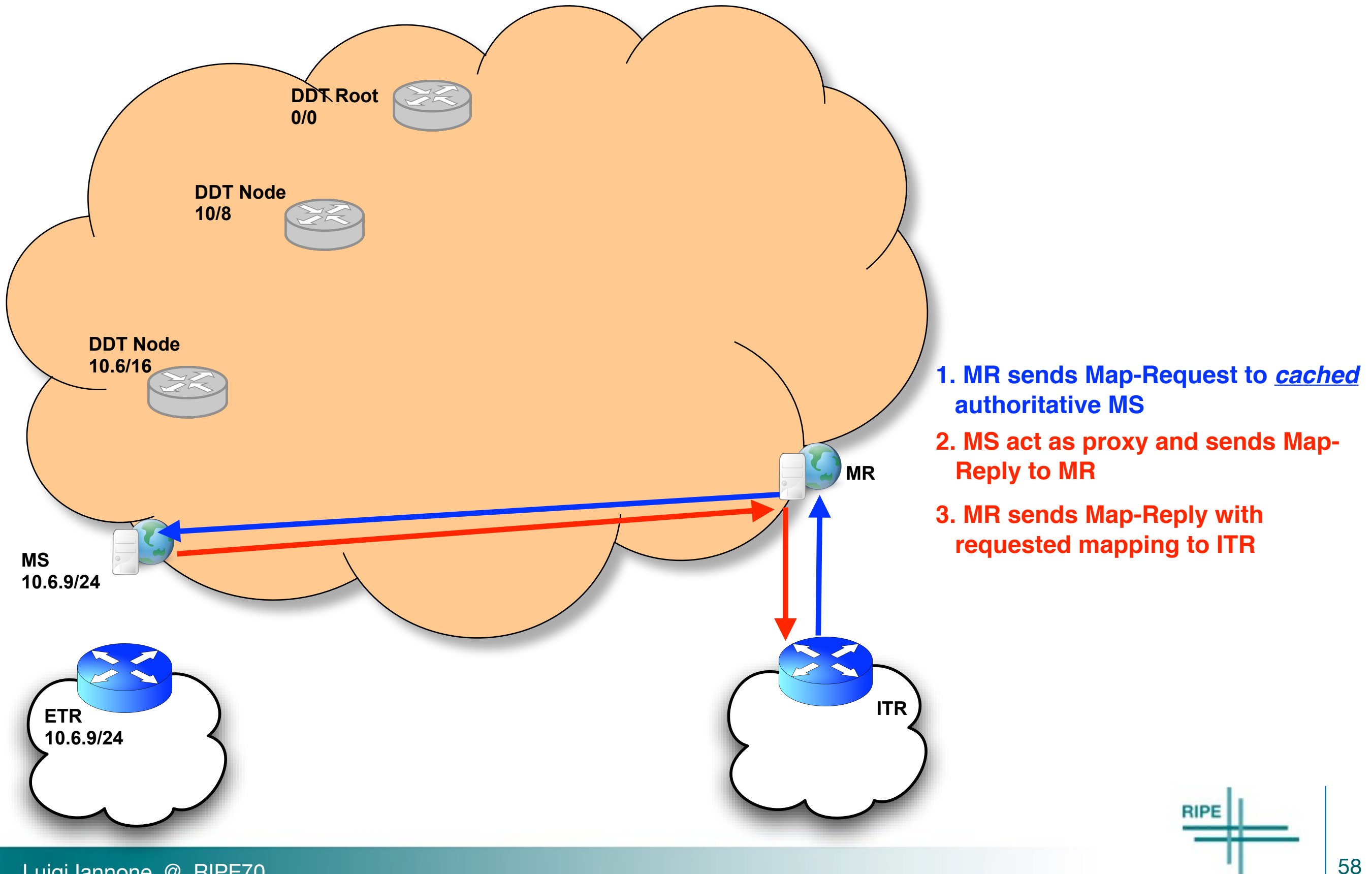


1. ITR sends Map-Request to its configured MR
2. MR sends Map-Request to configured Root Server
3. Root Server sends back Map-Referral to configured DDT Node authoritative on more specific
4. Repeat 2 and 3 until MS reached
5. MR sends Map-Request to authoritative MS
6. MS forward Map-Request to authoritative ETR
7. MS sends Map-Ack to MR
8. ETR sends Map-Reply with requested mapping to ITR

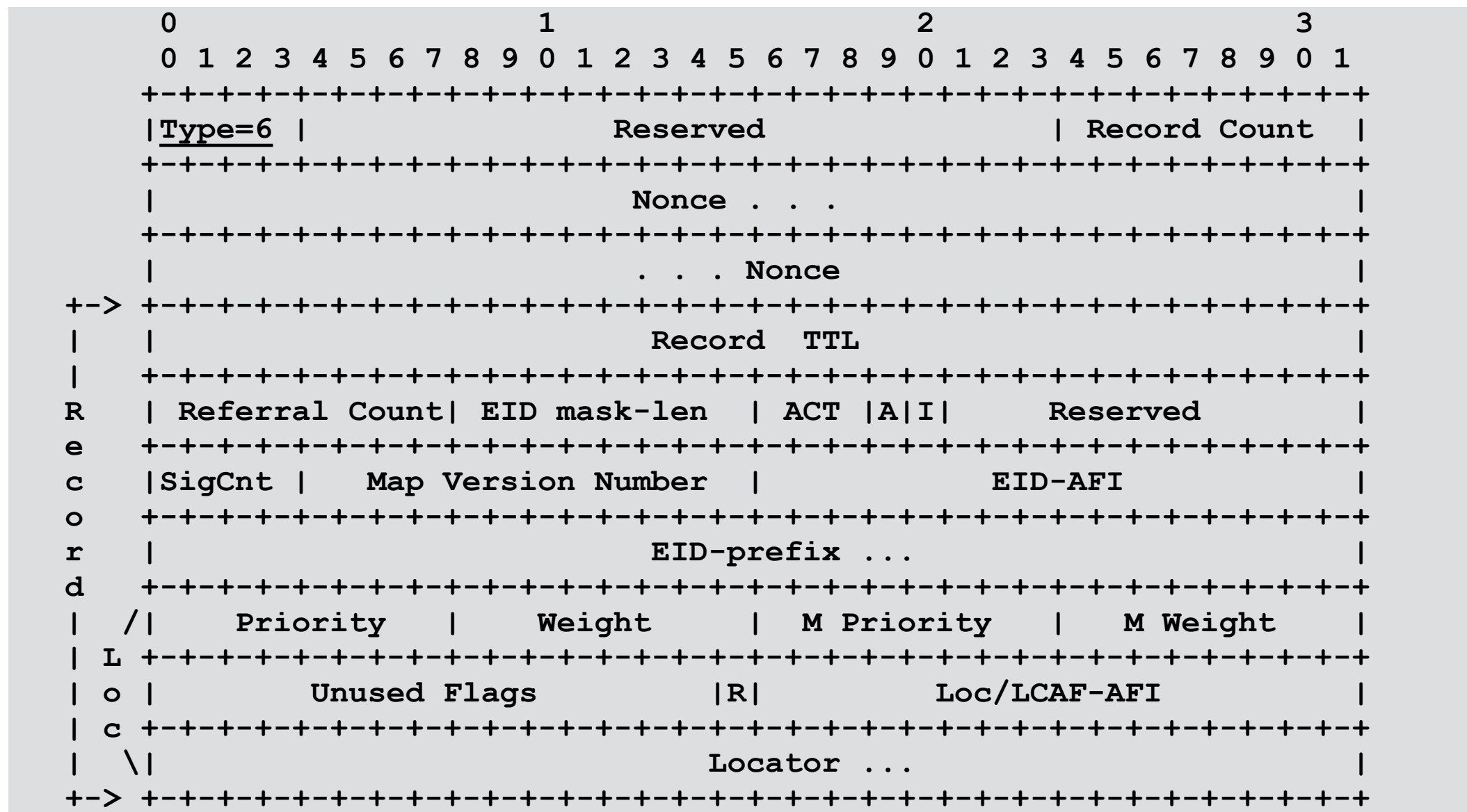
LISP-DDT: Mapping Retrieval (Regular Case)



LISP-DDT: Mapping Retrieval (Proxy Case)



Map-Referral



- Action Bits:

- Node Referral (0): sent by authoritative node
- MS-Referral (1): DDT-Node sending the referral is not a Map Server.
- MS-Ack (2): Sent by DDT Map Server
- MS-Not-Registered (3): Sent by a DDT Map Server with no ETRs registered
- Delegation-Hole (4): No child delegation for the EID. Also sent by a DDT Map Server
- Not-Authoritative (5): Sent by a DDT node not authoritative for the requested EID
 - May send incomplete referrals (bit I set)



RIPE

A large graphic element on the left side of the slide. It features the word "RIPE" in a light blue, sans-serif font. To the right of the text are two vertical white bars of different heights and two horizontal light blue bars of different lengths, creating a stylized cross or plus sign shape.

LISP Map-Version

RFC 6834



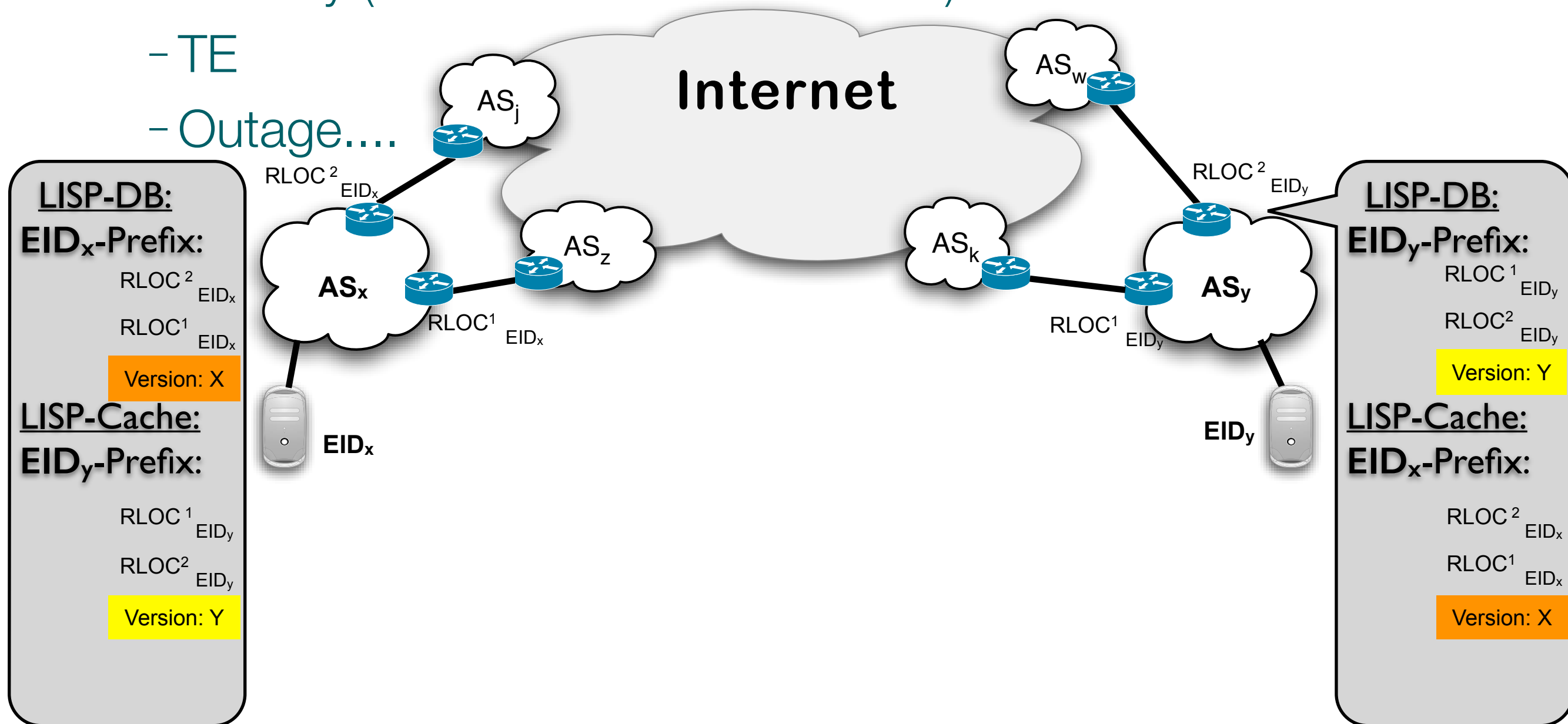
LISP Header: Versioning



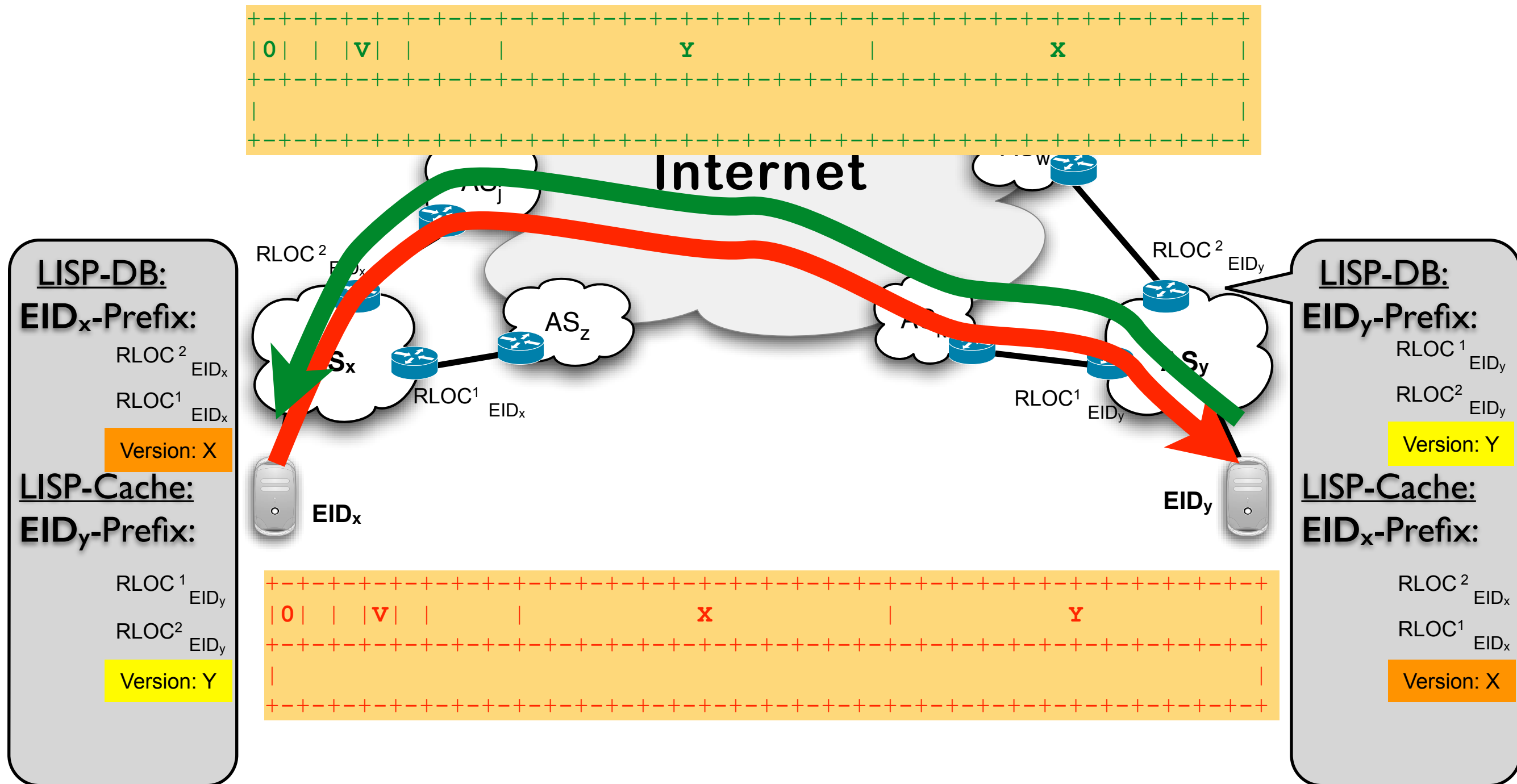
- V-bit: Map-Version Present bit
 - E-bit must be 0
- RFC 6834 (Now)
- LISP Map Versioning:
 - Bind a Version number to mappings
 - Embed Map Version in LISP header

The problem of Mapping “*Freshness*”

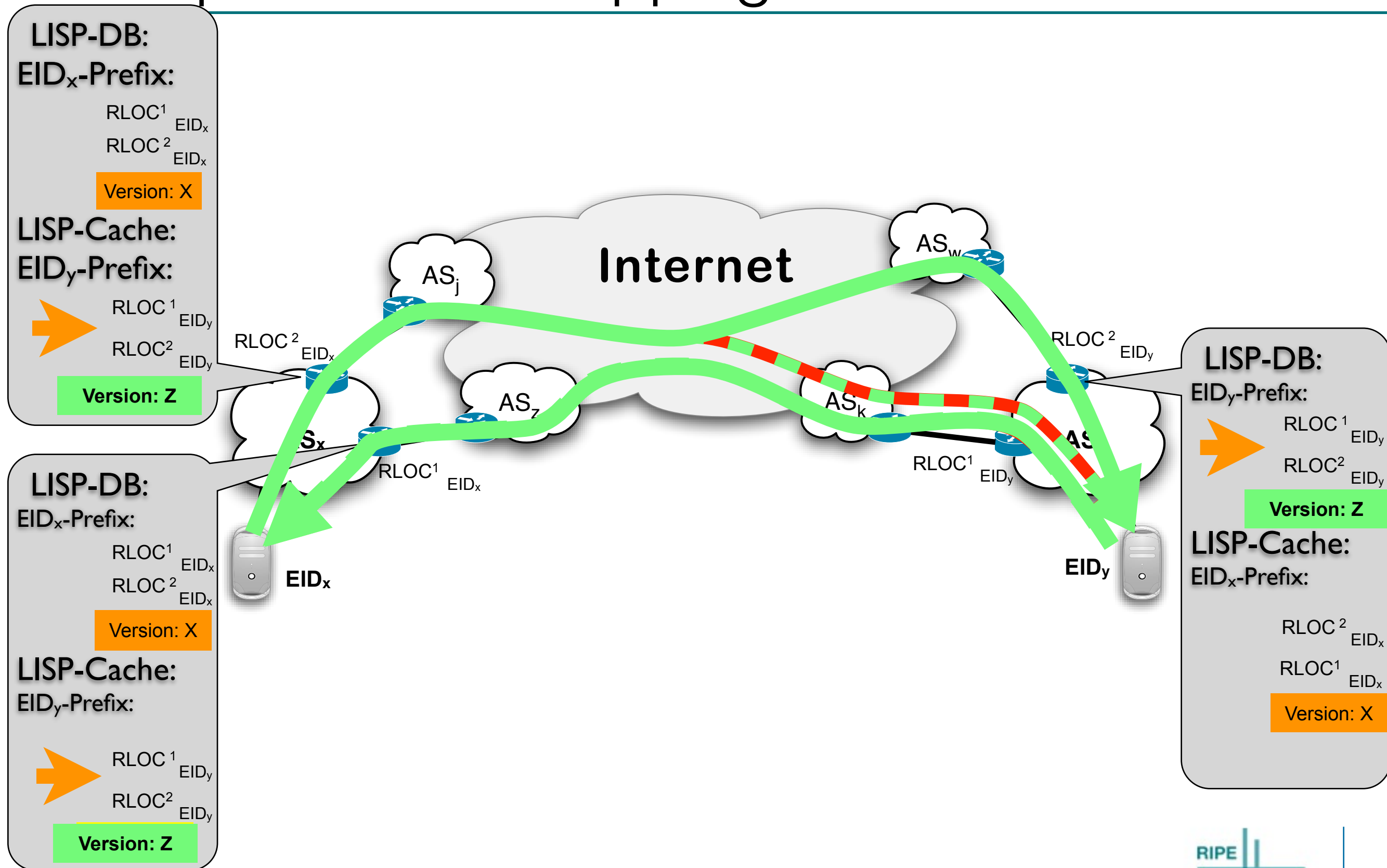
- Mappings can change due to:
 - Mobility (for some definition of it....)
 - TE
 - Outage....



LISP Versioning



The problem of Mapping “Freshness”



RIPE

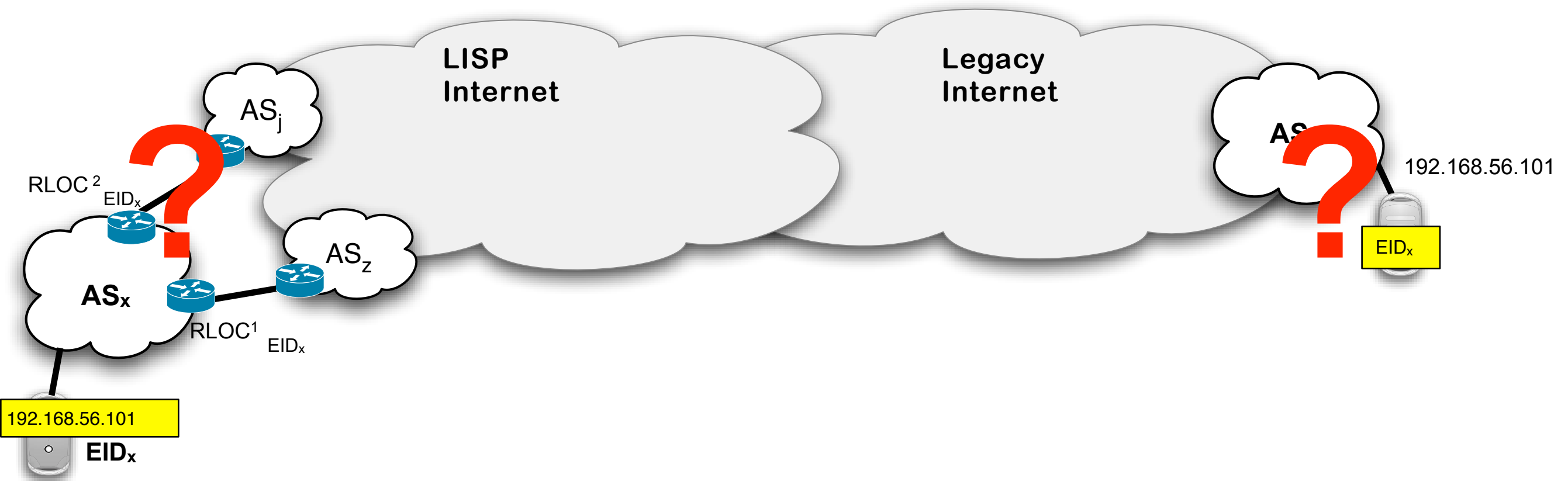


LISP Interworking

RFC 6832

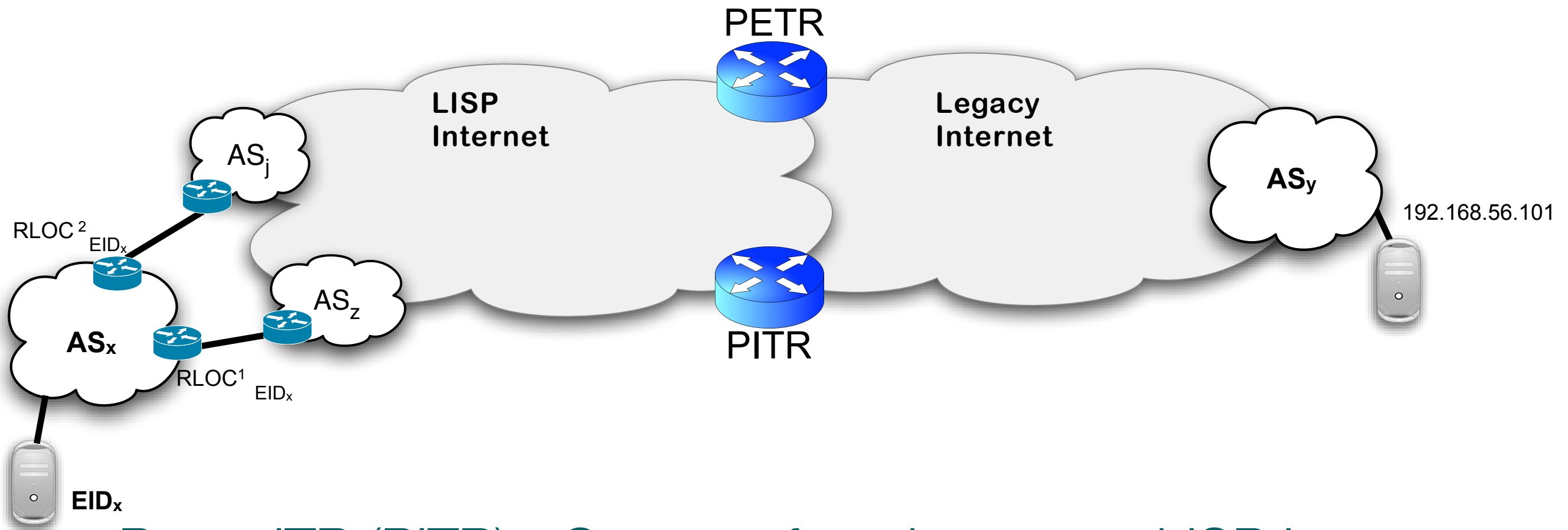


Interworking LISP and the Legacy Internet



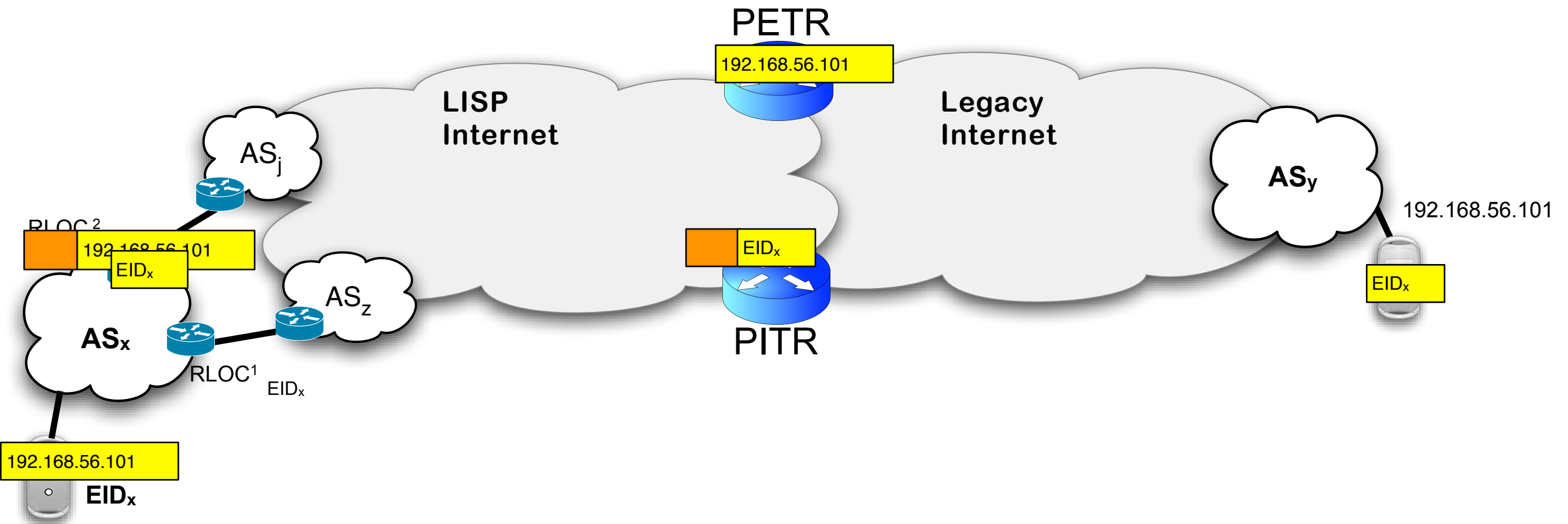
- How to forward packets from an ITR to a non-LISP site?
- How to forward packets from a non-LISP site to an EID?

Proxy-xTRs (PxTRs)



- Proxy-ITR (PITR) - Gateway from Legacy to LISP Internet
 - Announce large aggregates of EID-Prefixes into the legacy BGP
- Proxy-ETR (PETR) - Gateway from LISP to Legacy Internet
 - Could be a static configuration (similar to default route)
 - Could be obtained from Mapping System

PxTRs and asymmetric traffic flow



Few Pointers...

- www.lisp-lab.org
- datatracker.ietf.org/wg/lisp/
- www.openlisp.org
- www.lisp4.net

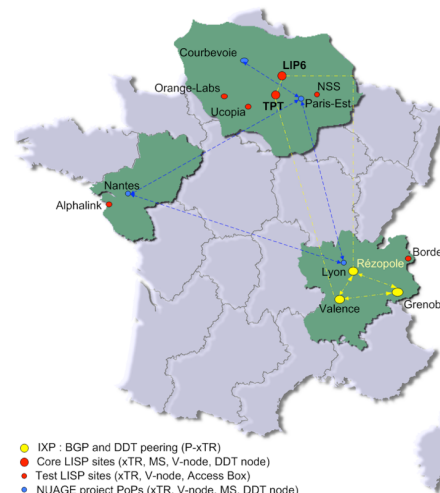


Coming Soon.....



Lisp-Lab

Building Today the Internet of Tomorrow



• IXP - BGP and DDT peering (P-xTR)
• Core LISP sites (xTR, MS, V-node, DDT node)
• Test LISP sites (xTR, V-node, Access Box)
• NUAGE project PoPs (xTR, V-node, MS, DDT node)



Questions?

