

# LISP tutorial

Damien Saucez  
Olivier Bonaventure

# Agenda

- Introduction
- Locator/Identifier Separation Protocol (LISP)
  - LISP in a nutshell
  - LISP data-plane
  - LISP control-plane
  - The reachability problem
  - Evaluation
- Conclusion

# Introduction

# The complementary roles of IP addresses

- The IP addresses currently used by endhosts play two complementary roles
  - **Identifier role:** the IP address identifies (with port) the endpoint of transport flows
  - **Locator role:** the IP address indicates the paths used to reach the endhost
    - these paths are updated by routing protocols after each topology change

# The Locator/Identifier Separation

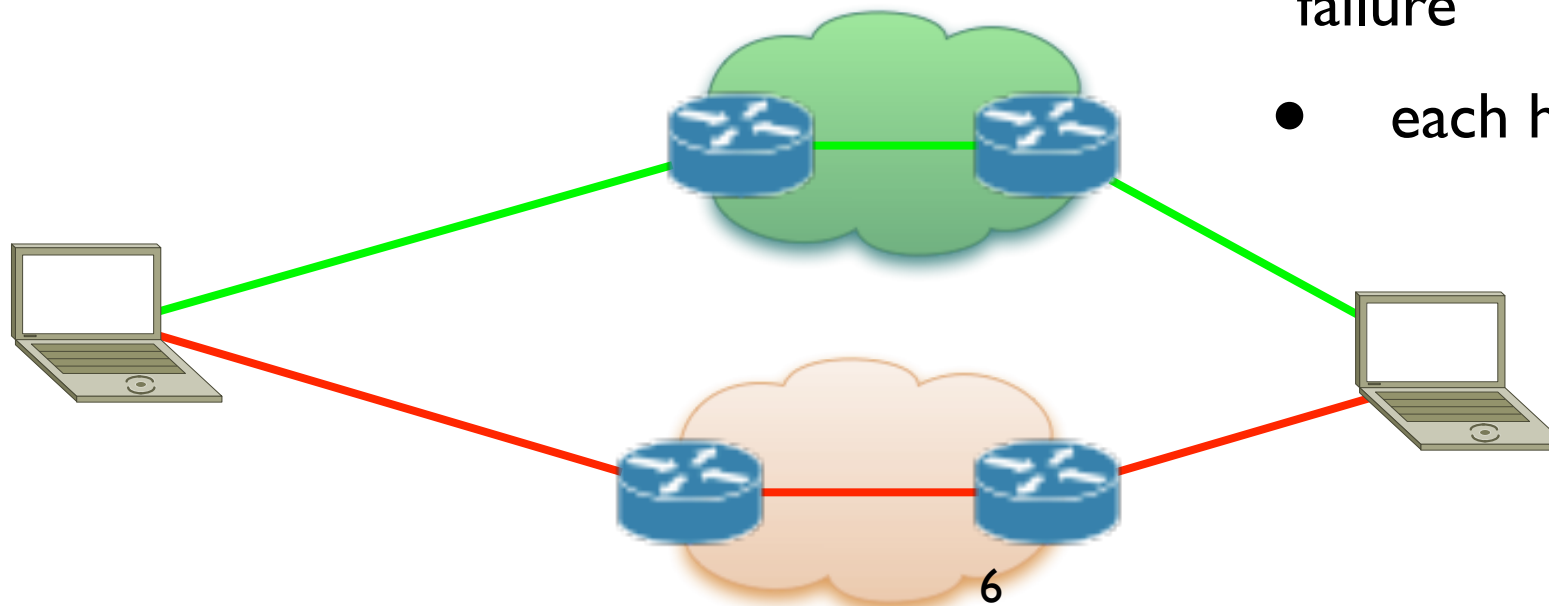
- Today, changing the locator means changing the identifier, breaking the pending flows
- Separating the locator and the identifier roles to avoid breaking the flows
  - Host-based approach
  - Network-based approach

# Host-based Loc/ID split

Transport layer

IP Routing sublayer

- Roles
  - Translates the packets so that
    - Transport layer always sees only the host identifier
    - IP Routing sublayer sees only locators
  - Manages the set of locators
  - Securely switches from one locator to another upon move or after link failure
  - each host maintains some state



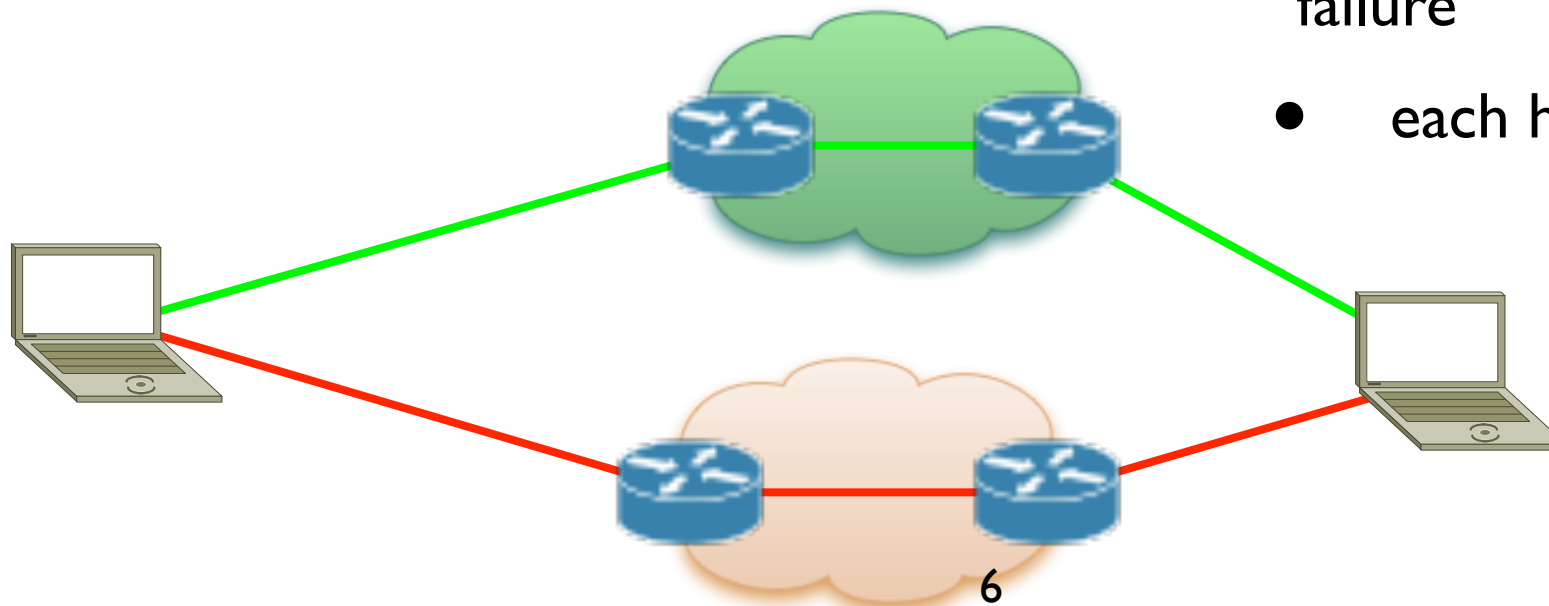
# Host-based Loc/ID split

Transport layer

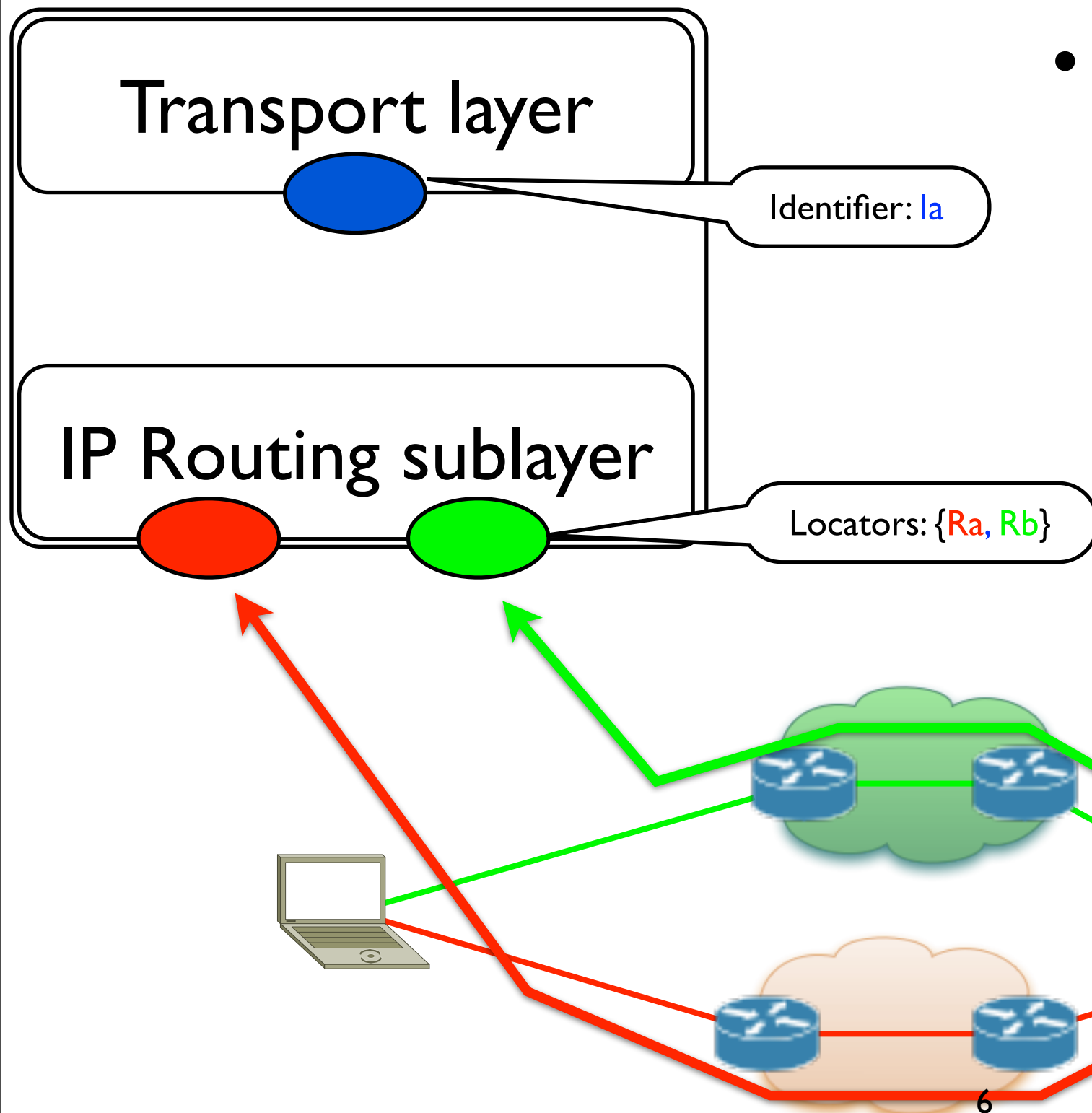
Identifier: *la*

IP Routing sublayer

- Roles
  - Translates the packets so that
    - Transport layer always sees only the host identifier
    - IP Routing sublayer sees only locators
  - Manages the set of locators
  - Securely switches from one locator to another upon move or after link failure
  - each host maintains some state

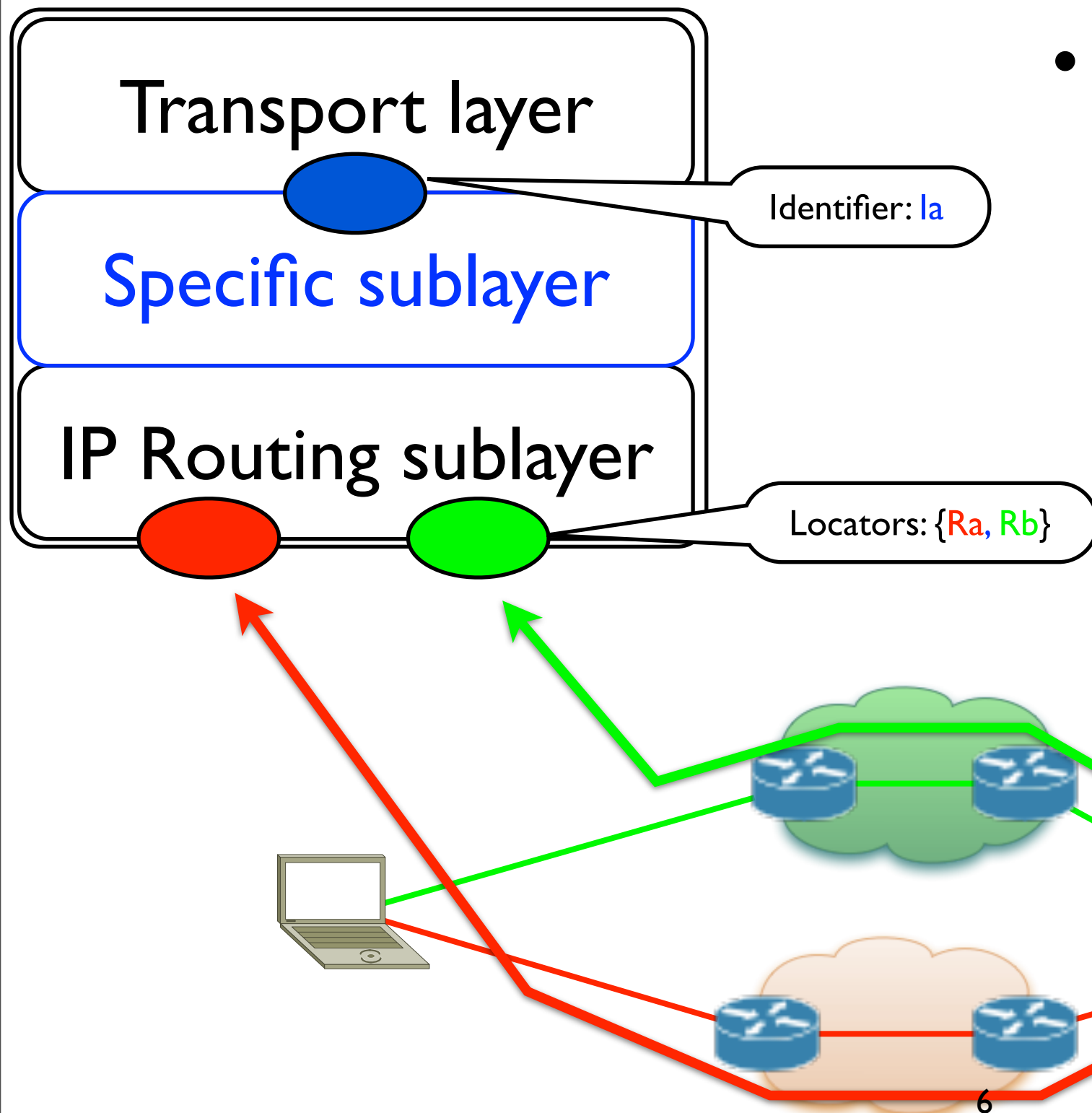


# Host-based Loc/ID split



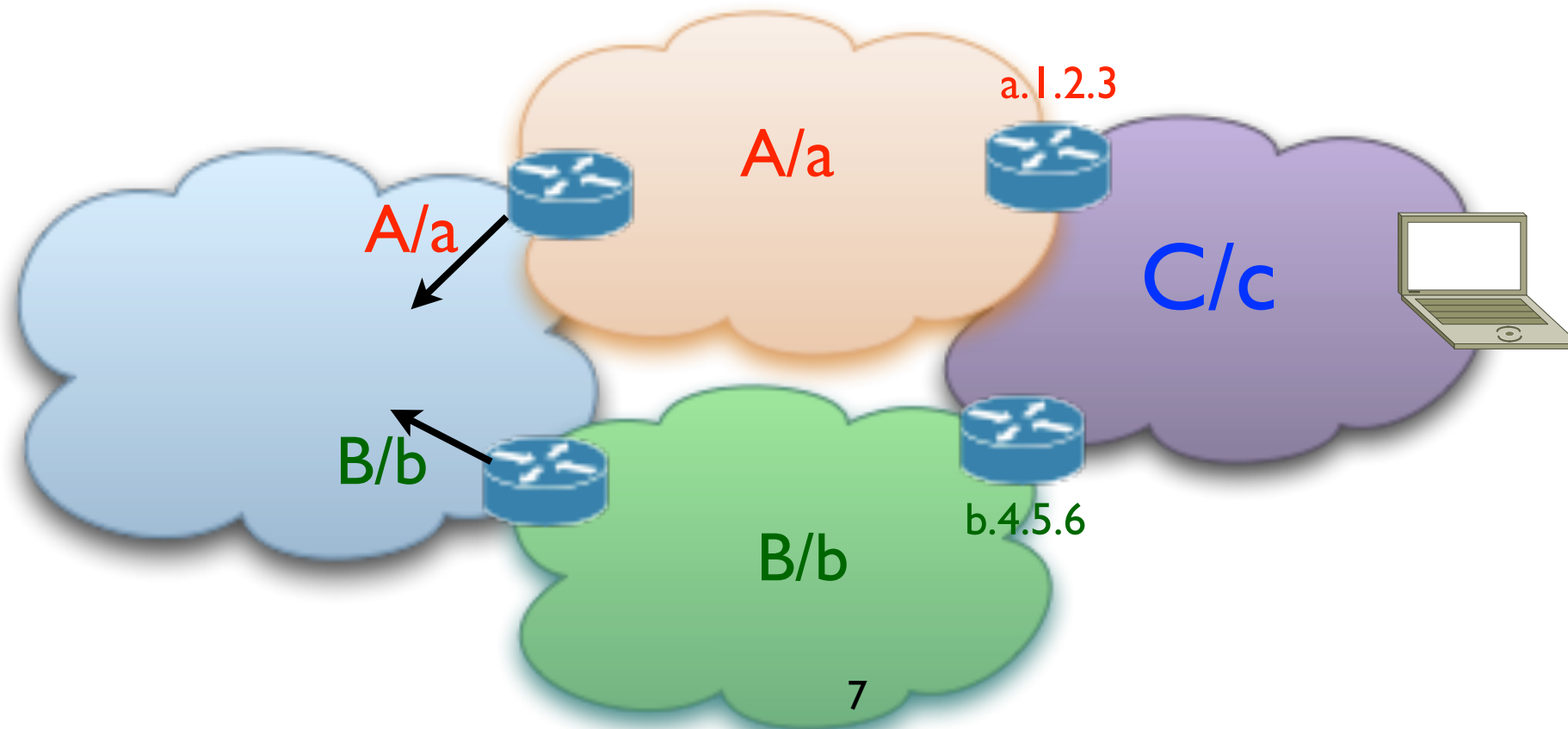
- Roles
  - Translates the packets so that
    - Transport layer always sees only the host identifier
    - IP Routing sublayer sees only locators
  - Manages the set of locators
  - Securely switches from one locator to another upon move or after link failure
  - each host maintains some state

# Host-based Loc/ID split



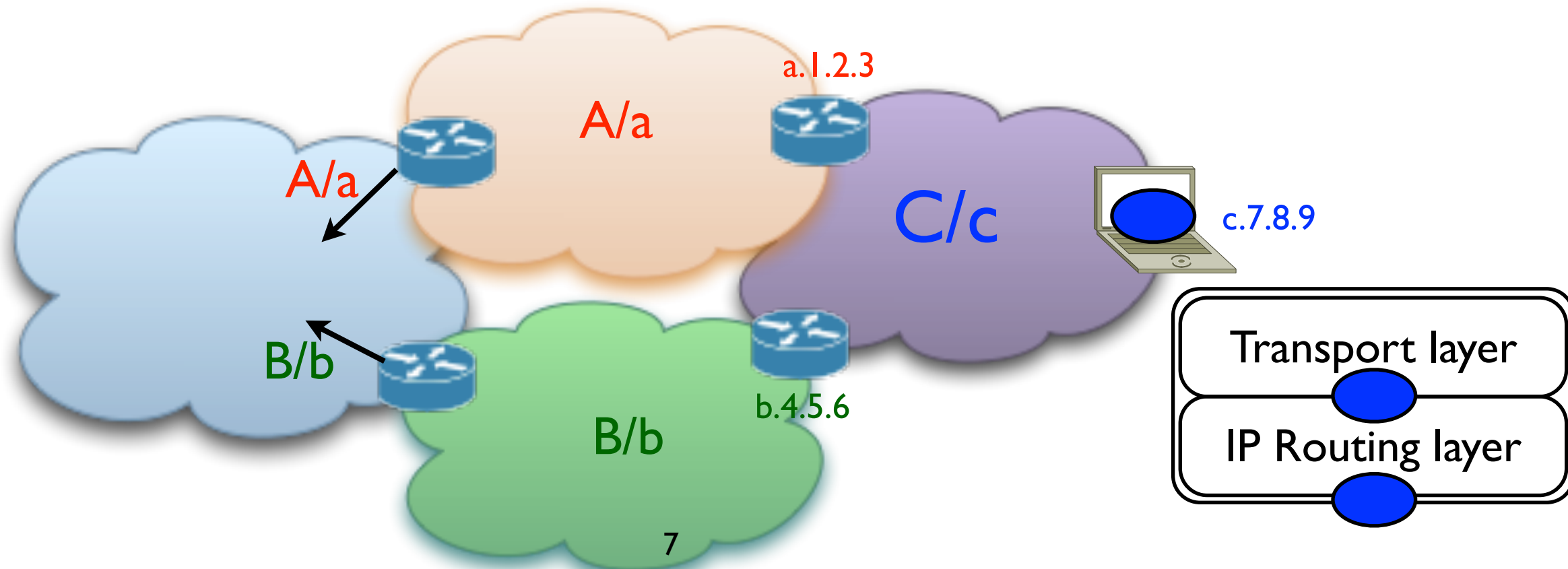
- Roles
  - Translates the packets so that
    - Transport layer always sees only the host identifier
    - IP Routing sublayer sees only locators
  - Manages the set of locators
  - Securely switches from one locator to another upon move or after link failure
  - each host maintains some state

# Network-based Loc/ID split



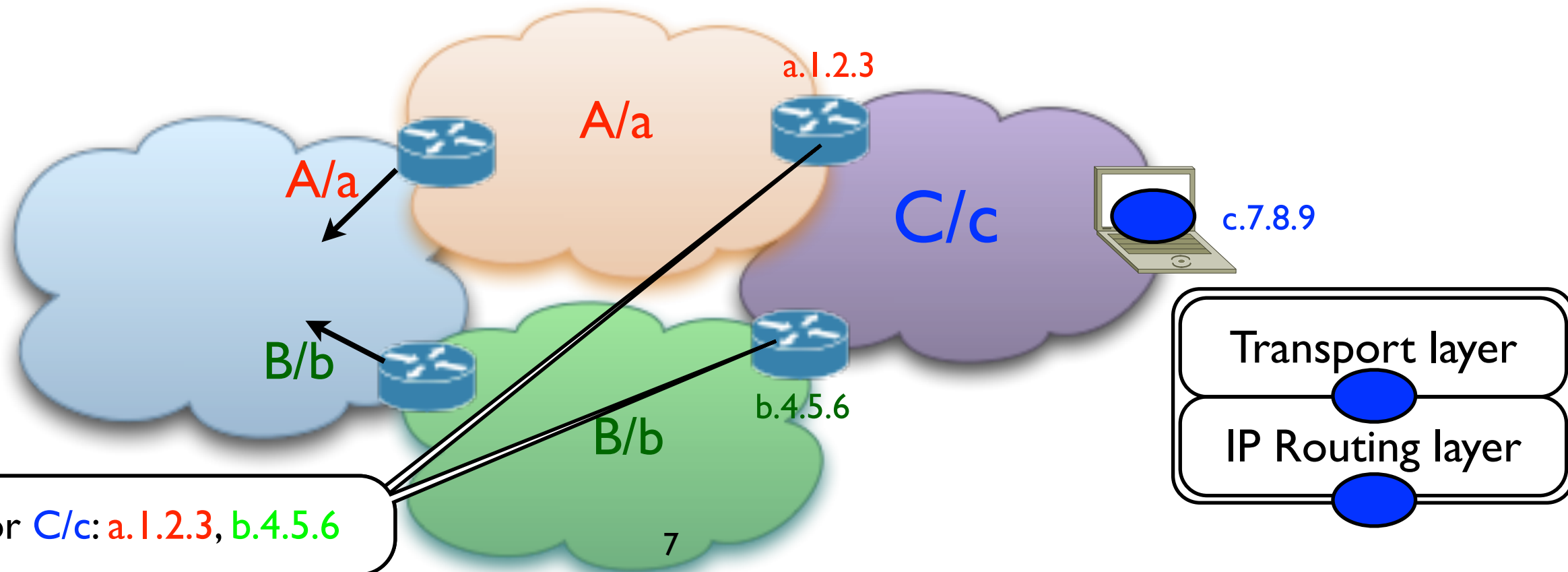
# Network-based Loc/ID split

- Host's IP stack unchanged
- Each host has one stable IP address
- used as identifier
- **not globally routable**



# Network-based Loc/ID split

- Each edge router owns
  - globally routed addresses used as locators
- **Mapping mechanism** is used to find locators associated to one identifier
- Packets from hosts are modified before being sent on Internet
- Host's IP stack unchanged
  - Each host has one stable IP address
  - used as identifier
  - **not globally routable**



# Locator/Identifier Separation Protocol (LISP)

# LISP in a nutshell

# LISP Main Design Goals

- Minimize required changes to Internet
- No end-systems (hosts) changes
- Be incrementally deployable
- No router hardware changes
- Minimize router software changes

# Alternative deployments?

- LISP's first intention is to improve the global Internet
- What about using LISP in:
  - enterprises environments (VPNs)
  - datacenter

# The Locator Identifier Separation Protocol (I/2)

- Define a router-based solution where current IP addresses are separated in two different spaces:
  - **EndPoint Identifiers (EID)**
    - identify end-hosts
    - non-globally routable
    - hosts in a given site are expected to use EIDs in the same prefix
  - **Routing Locators (RLOC)**
    - attached to routers (router interfaces)
    - globally routable

# The Locator Identifier Separation Protocol (2/2)

- Follows the Map-and-Encap principle
  - A **mapping system** maps EID prefixes onto site routers RLOCs
  - Routers **encapsulate** the packets received from hosts before sending them towards the destination RLOC
  - Routers **decapsulate** the packets received from the Internet before sending them towards the destination hosts

# Terminology

- **Ingress Tunnel Router (*ITR*)**: a router which accepts a packet containing a single IP header. The router maps the destination address of the packet to an RLOC and prepends a LISP header before forwarding the encapsulated packet.
- **Egress Tunnel Router (*ETR*)**: a router which accepts a LISP encapsulated packet. The router strips the LISP header and forwards the packet based on the next header

# Terminology

- **EID-to-RLOC Database:** a globally distributed database that contains all known EID-prefix to RLOC mappings
- **LISP Cache:** EID-to-RLOC Database stored at the ITR
- **LISP Database:** EID-to-RLOC Database stored at the ETR

# The big picture

2001:DB8B::5678

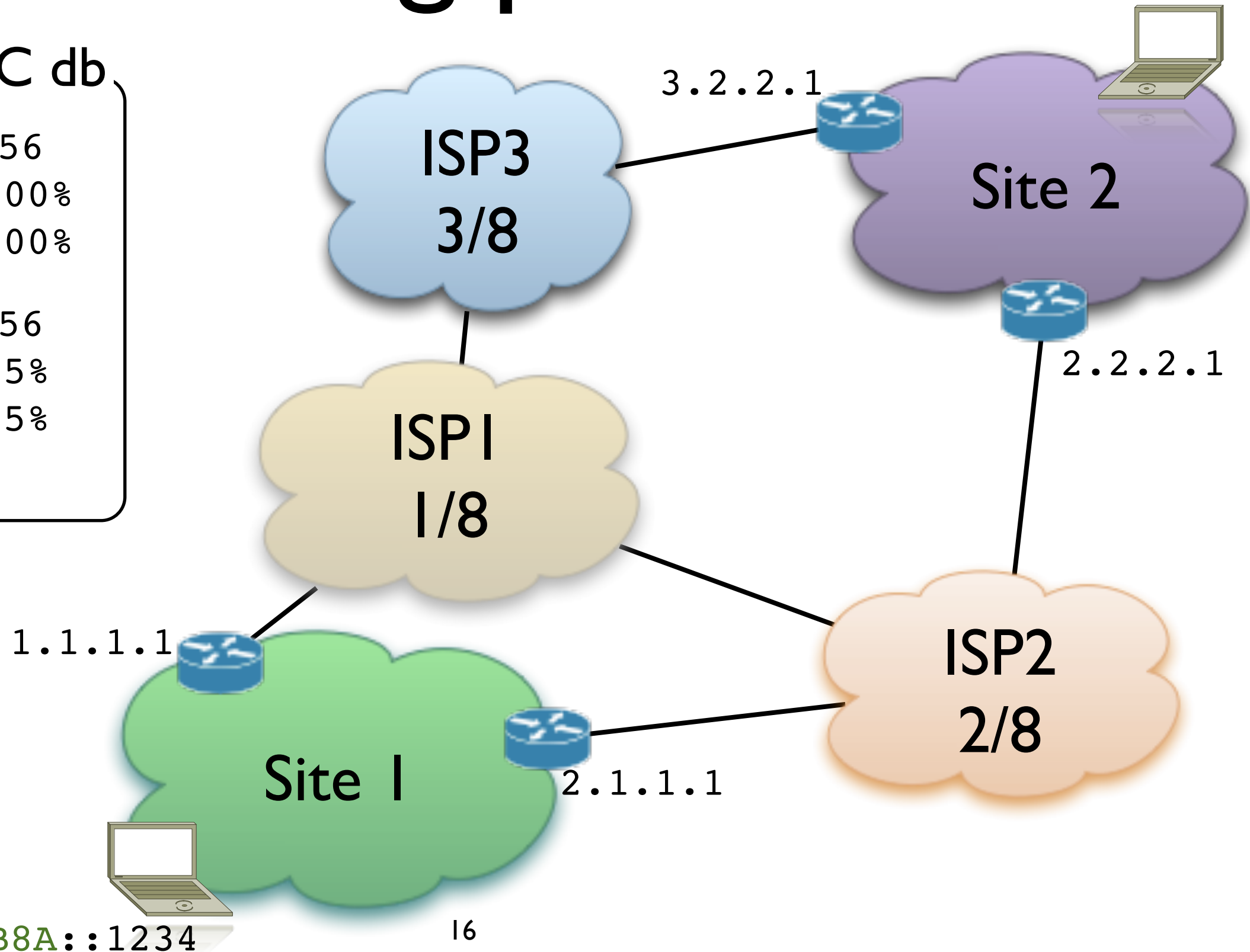
## EID-to-RLOC db

2001:DB8B::/56

- 3.2.2.1 1 100%
- 2.2.2.1 2 100%

2001:DB8A::/56

- 1.1.1.1 1 75%
- 2.1.1.1 1 25%



# The big picture

2001:DB8B::5678

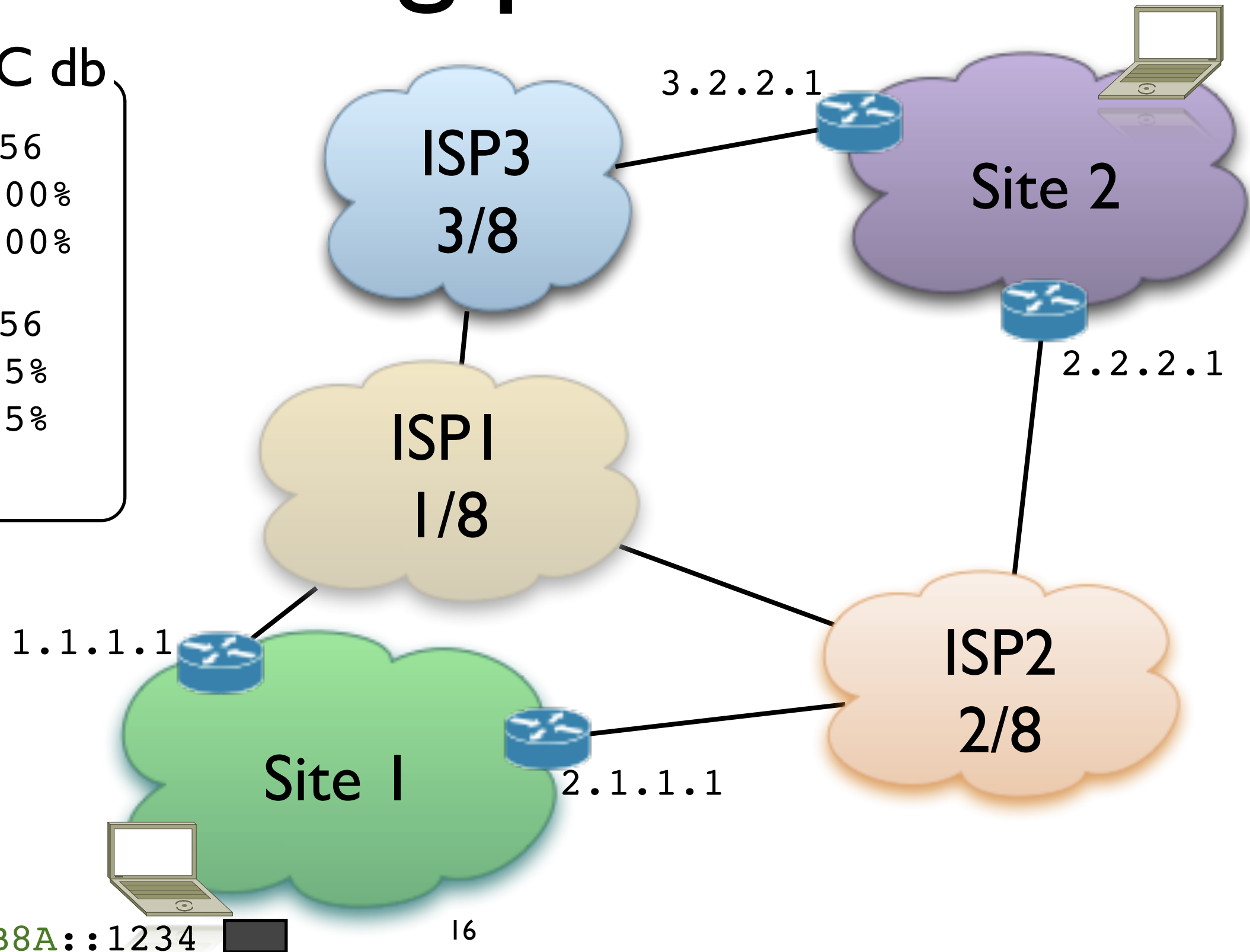
## EID-to-RLOC db

2001:DB8B::/56

- 3.2.2.1 1 100%
- 2.2.2.1 2 100%

2001:DB8A::/56

- 1.1.1.1 1 75%
- 2.1.1.1 1 25%



# The big picture

2001:DB8B::5678

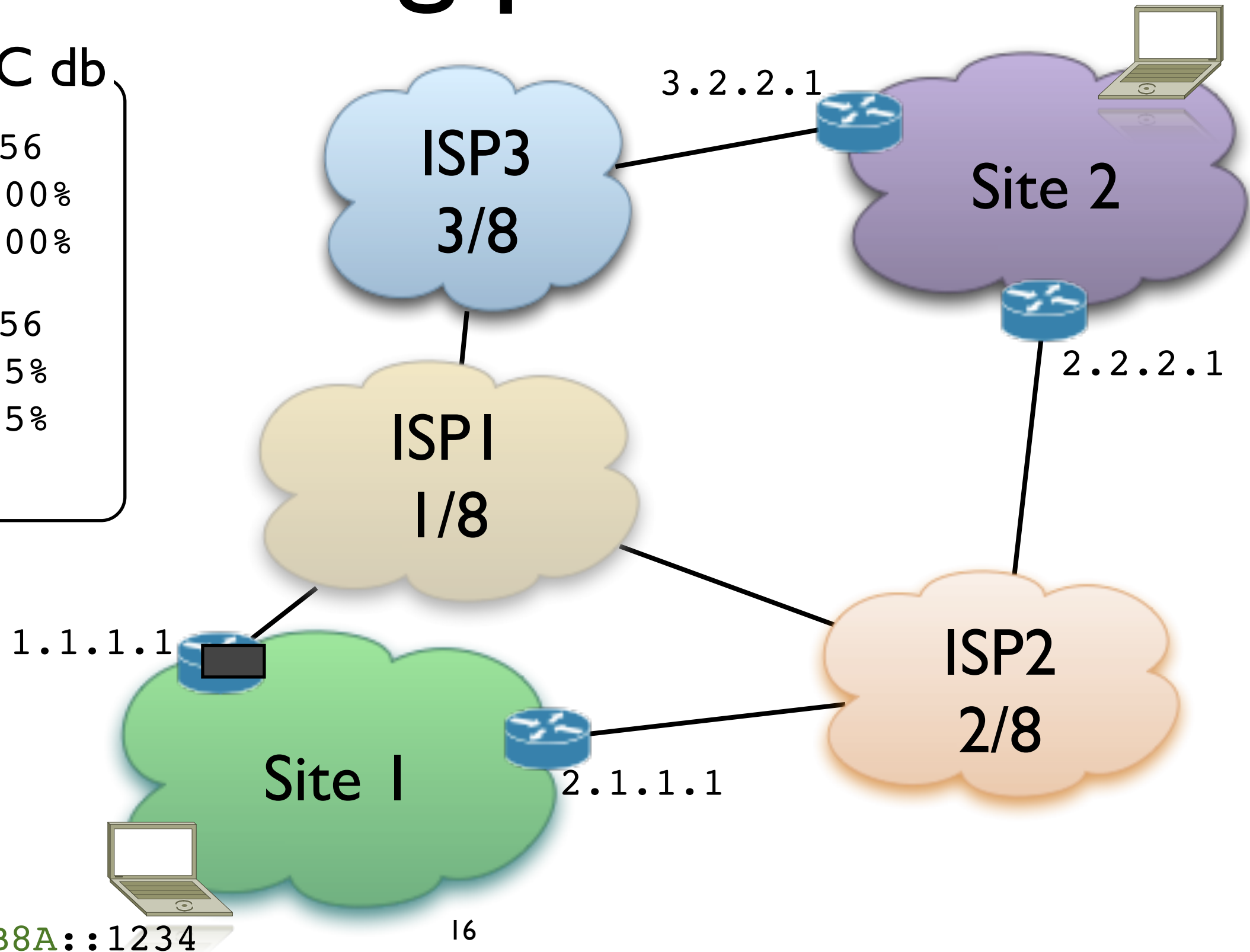
## EID-to-RLOC db

2001:DB8B::/56

- 3.2.2.1 1 100%
- 2.2.2.1 2 100%

2001:DB8A::/56

- 1.1.1.1 1 75%
- 2.1.1.1 1 25%



2001:DB8A::1234

# The big picture

2001:DB8B::5678

## EID-to-RLOC db

2001:DB8B::/56

- 3.2.2.1 1 100%
- 2.2.2.1 2 100%

2001:DB8A::/56

- 1.1.1.1 1 75%
- 2.1.1.1 1 25%

Map-Request:  
2001:DB8B::5678?

1.1.1.1

2001:DB8A::1234

ISP3  
3/8

ISP1  
1/8

Site 2

ISP2  
2/8

Site 1

2.1.1.1

3.2.2.1

2.2.2.1

# The big picture

2001:DB8B::5678

## EID-to-RLOC db

2001:DB8B::/56

- 3.2.2.1 1 100%
- 2.2.2.1 2 100%

2001:DB8A::/56

- 1.1.1.1 1 75%
- 2.1.1.1 1 25%

## Map-Reply:

2001:DB8B::/56

3.2.2.1 1 100%

2.2.2.1 2 100%

1.1.1.1

Site 1

ISP3  
3/8

ISP1  
1/8

Site 2

ISP2  
2/8

3.2.2.1

2.2.2.1

2.1.1.1

2001:DB8A::1234

16

# The big picture

2001:DB8B::5678

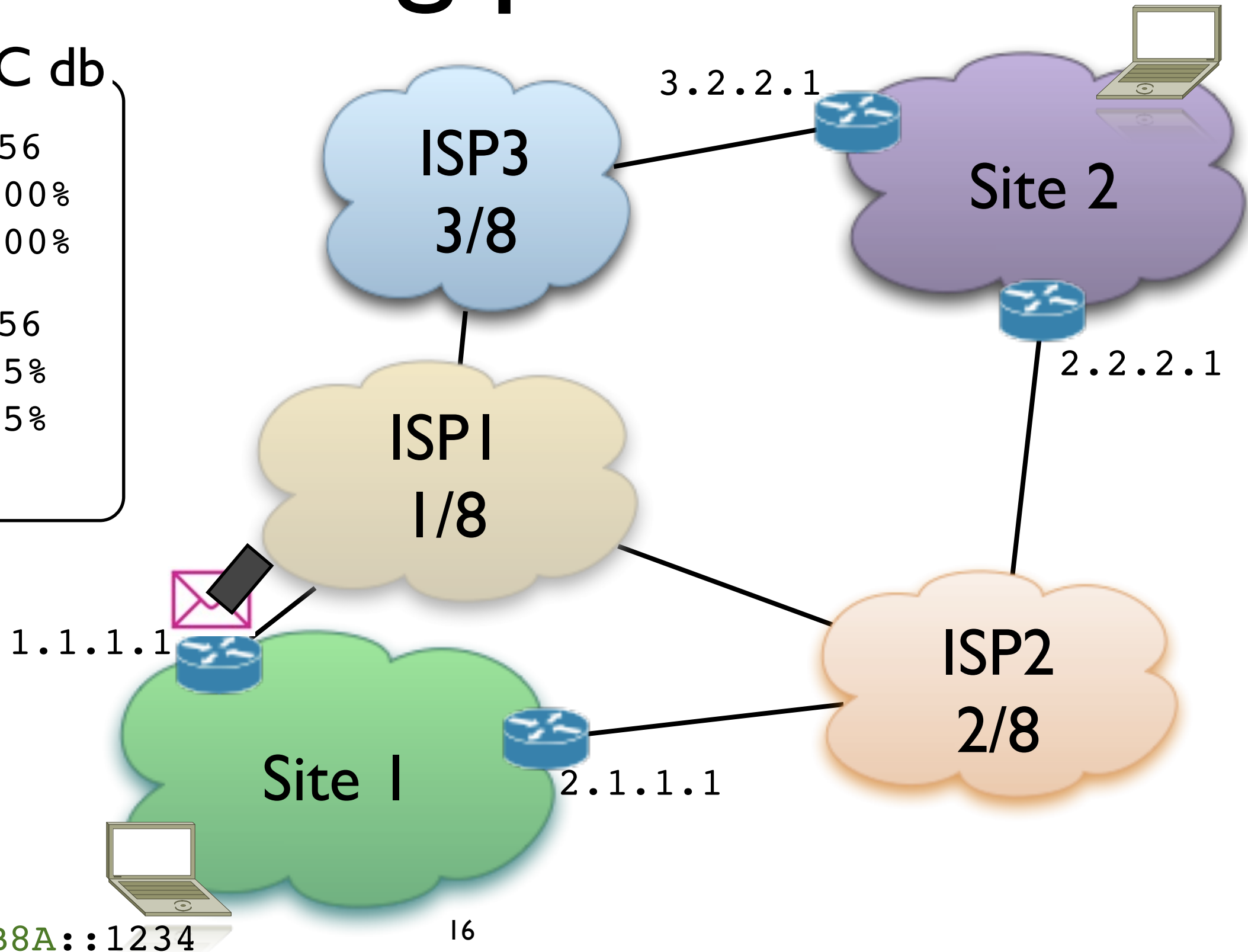
## EID-to-RLOC db

2001:DB8B::/56

- 3.2.2.1 1 100%
- 2.2.2.1 2 100%

2001:DB8A::/56

- 1.1.1.1 1 75%
- 2.1.1.1 1 25%



2001:DB8A::1234

16

# The big picture

2001:DB8B::5678

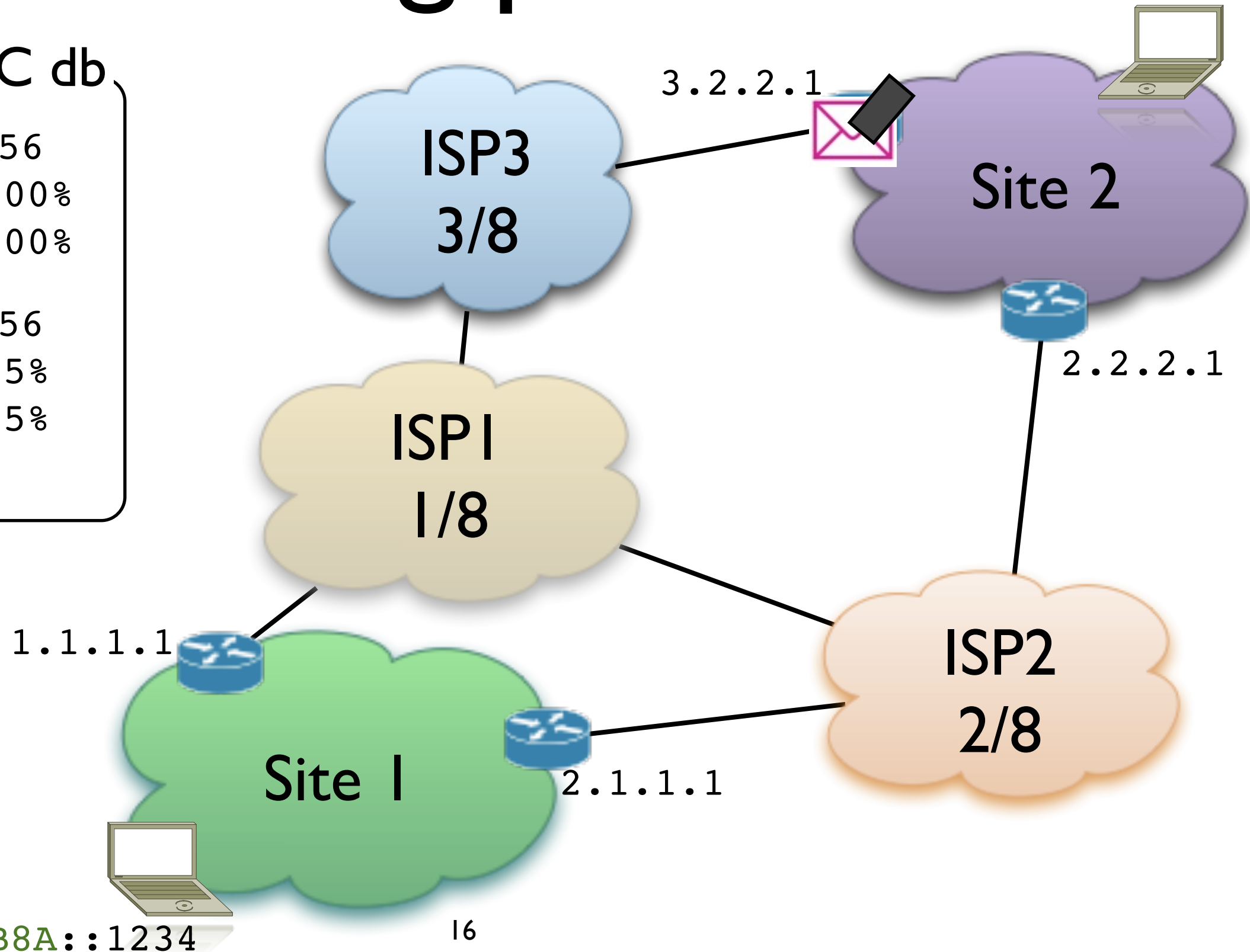
## EID-to-RLOC db

2001:DB8B::/56

- 3.2.2.1 1 100%
- 2.2.2.1 2 100%

2001:DB8A::/56

- 1.1.1.1 1 75%
- 2.1.1.1 1 25%



# The big picture

2001:DB8B::5678

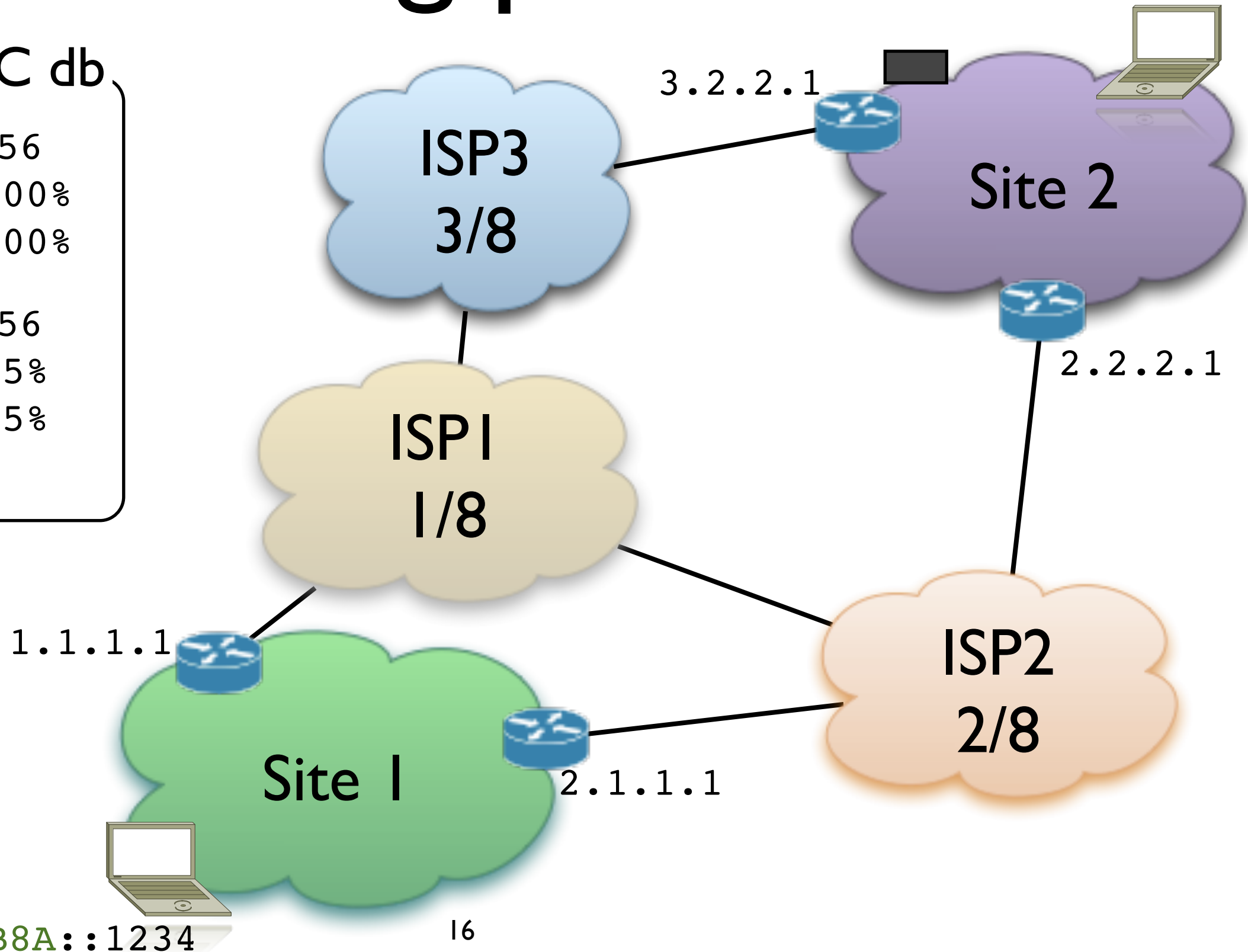
## EID-to-RLOC db

2001:DB8B::/56

- 3.2.2.1 1 100%
- 2.2.2.1 2 100%

2001:DB8A::/56

- 1.1.1.1 1 75%
- 2.1.1.1 1 25%



# The big picture

2001:DB8B::5678

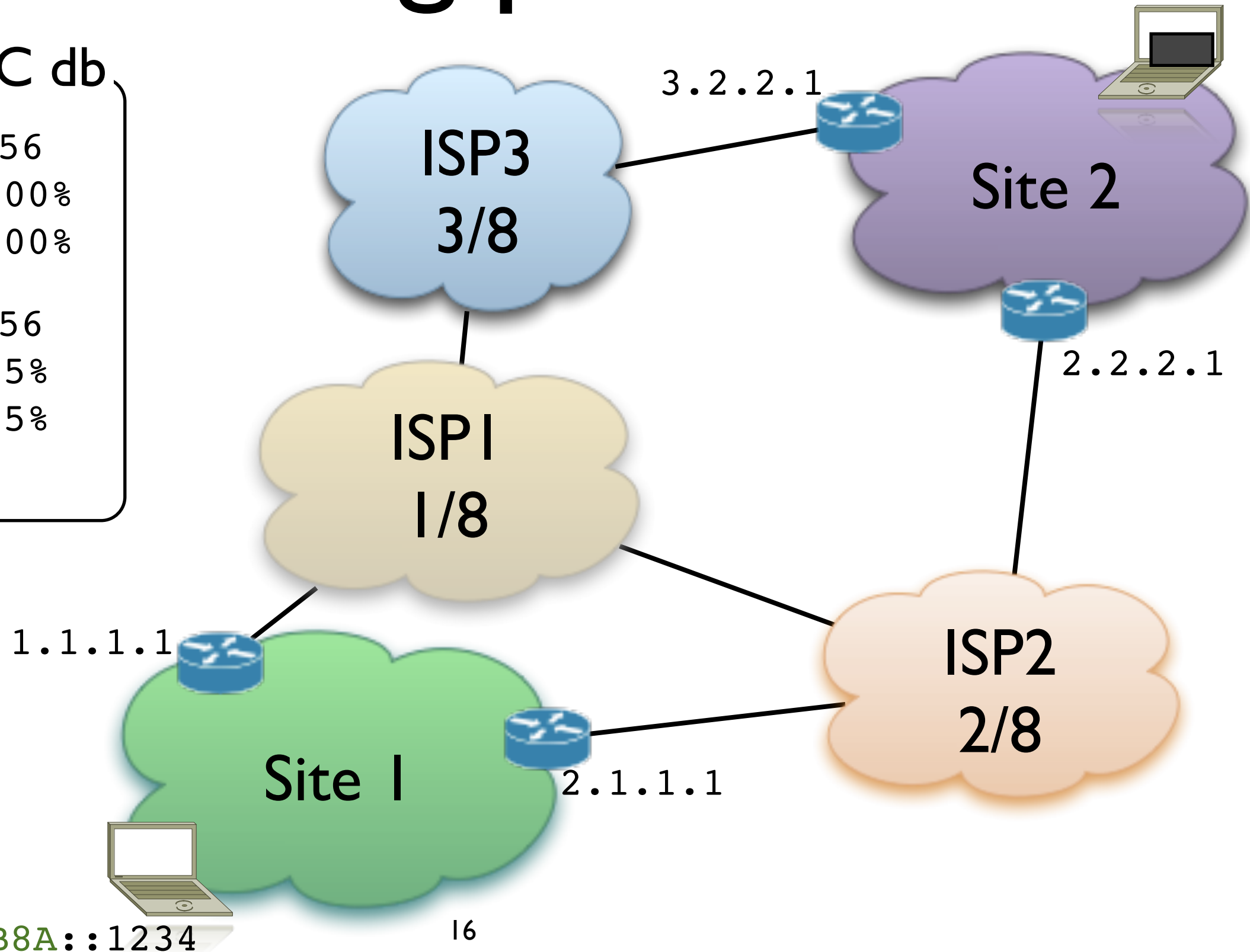
## EID-to-RLOC db

2001:DB8B::/56

- 3.2.2.1 1 100%
- 2.2.2.1 2 100%

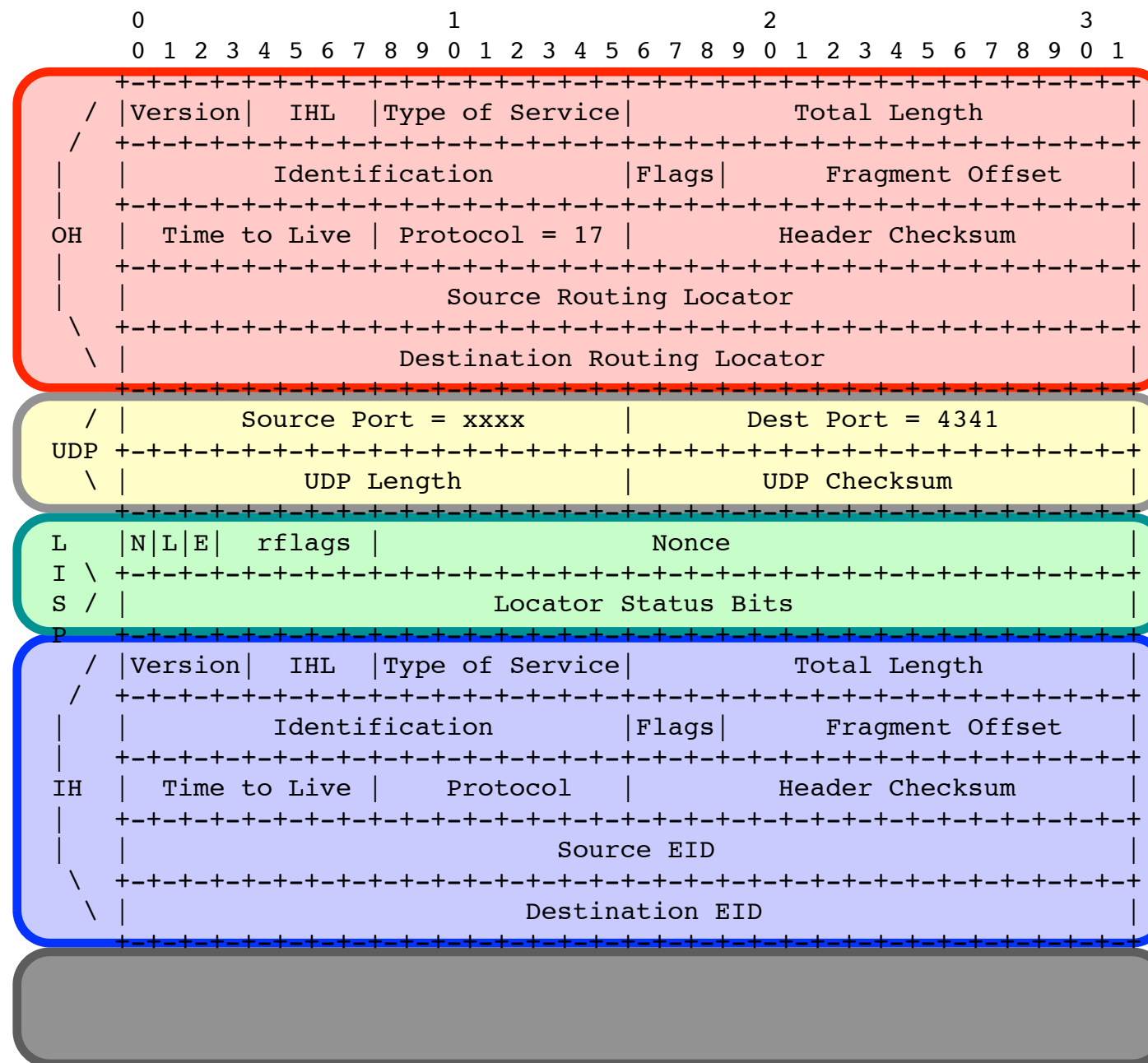
2001:DB8A::/56

- 1.1.1.1 1 75%
- 2.1.1.1 1 25%



# LISP Data-plane

# Data-plane packets



Locator part

Presentation

Control Flags

Identifier Part

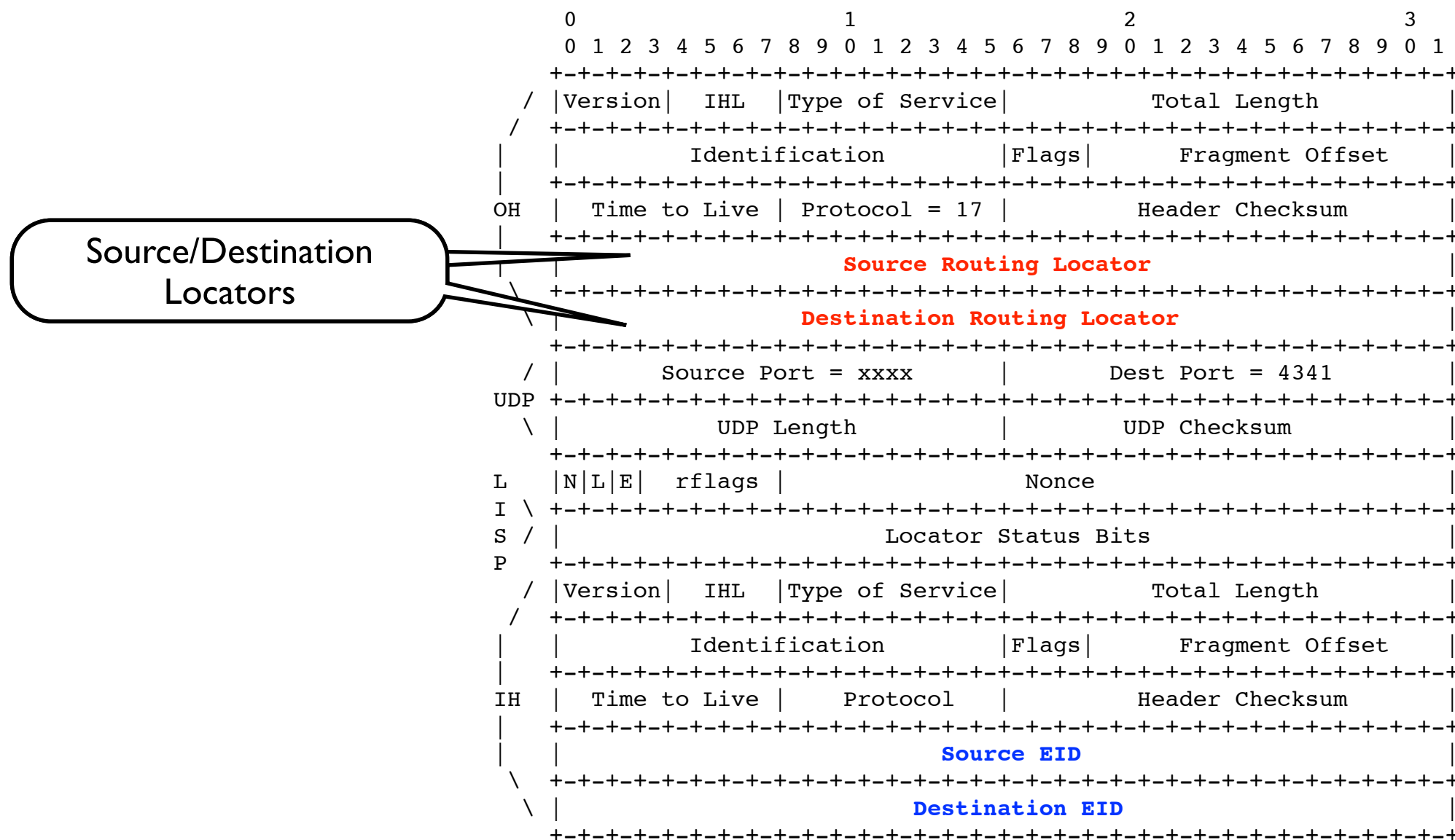
Payload

(IP(UDP(LISP(IP))))

19

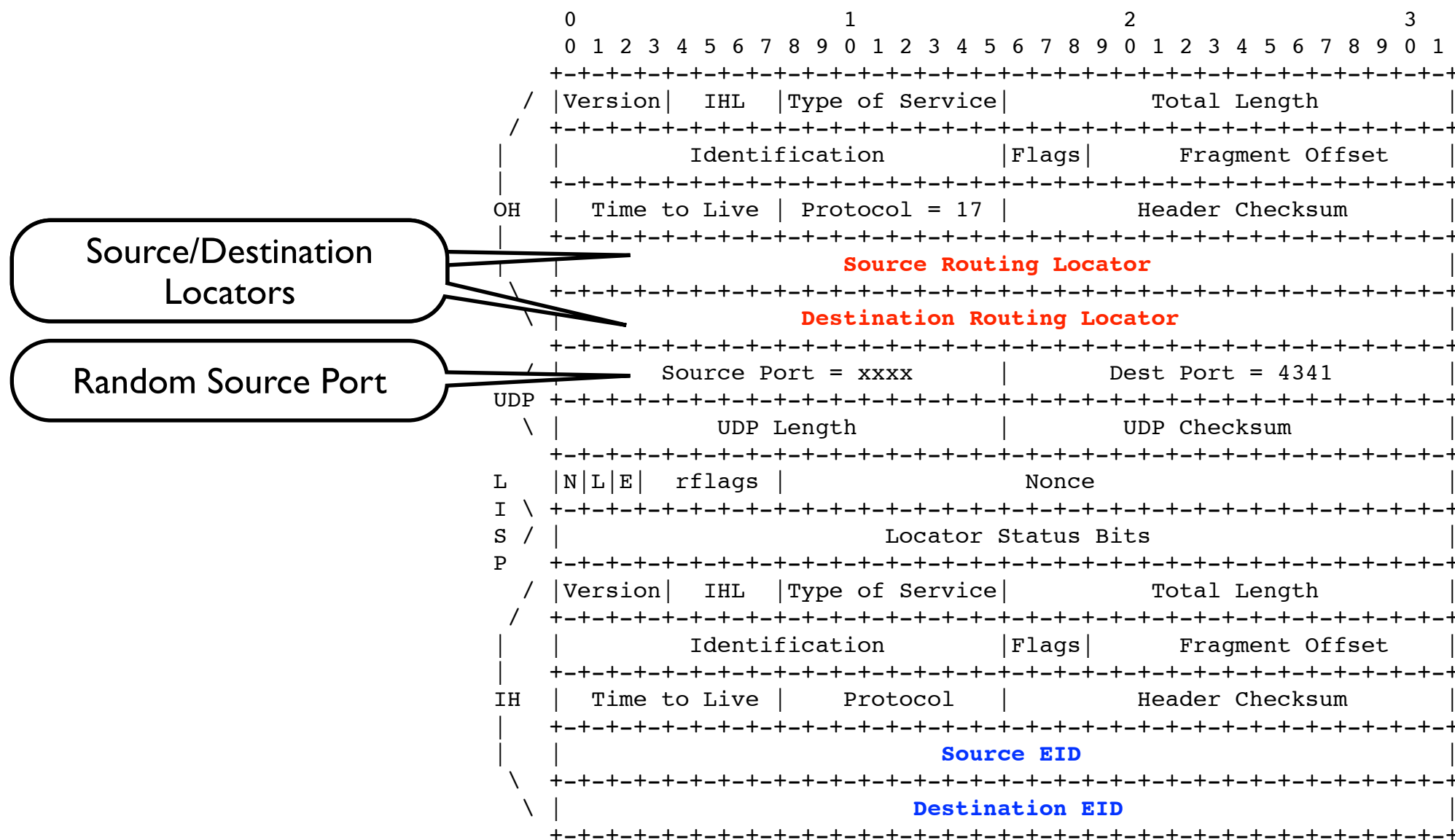
# Header details

(IP(UDP(LISP(IP))))



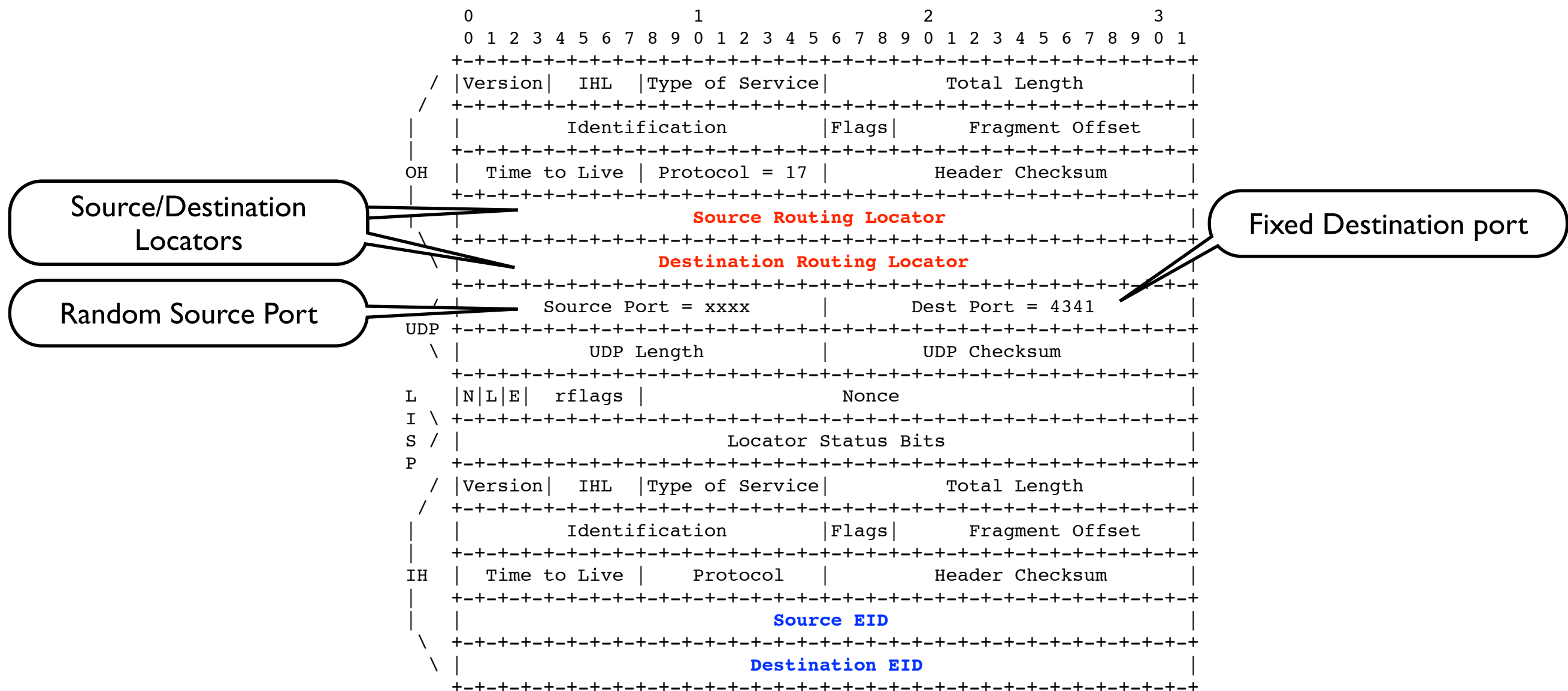
# Header details

(IP(UDP(LISP(IP))))



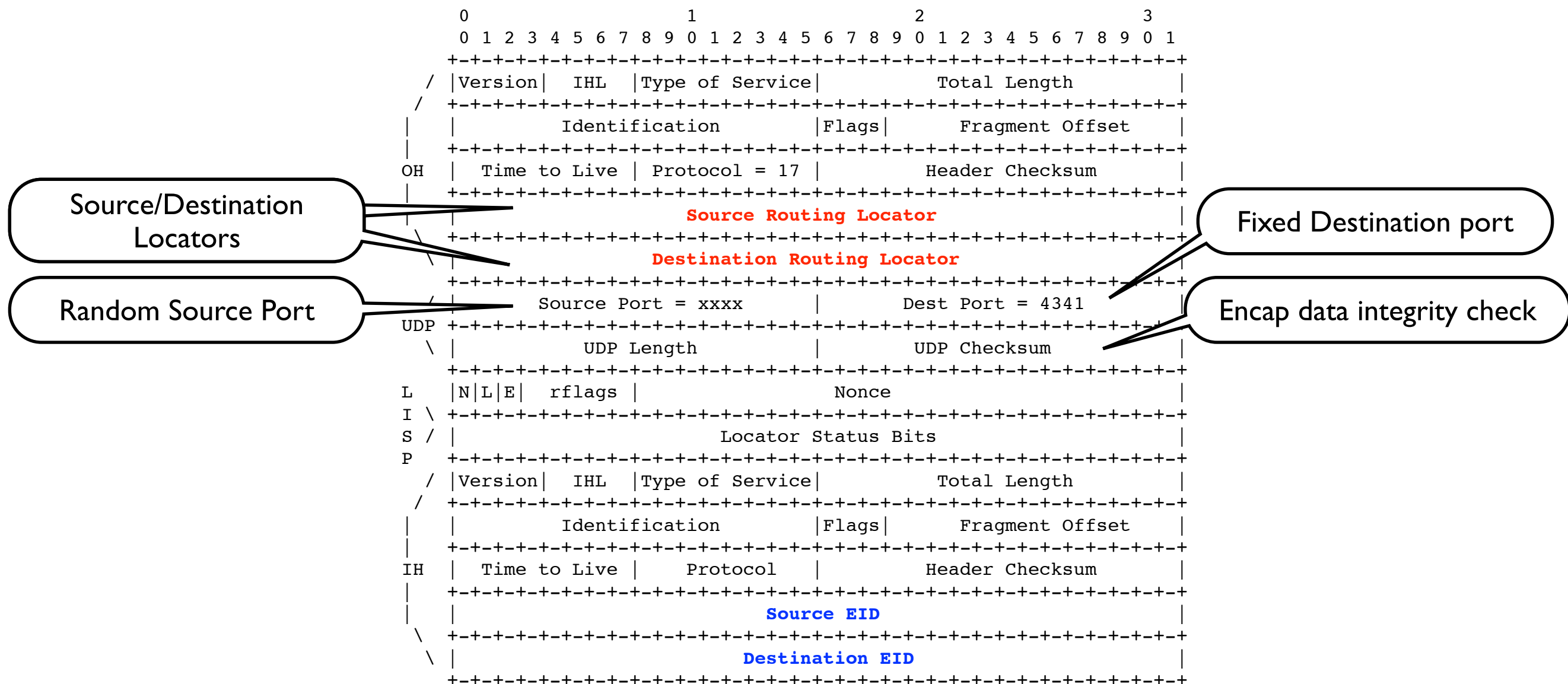
# Header details

(IP(UDP(LISP(IP))))



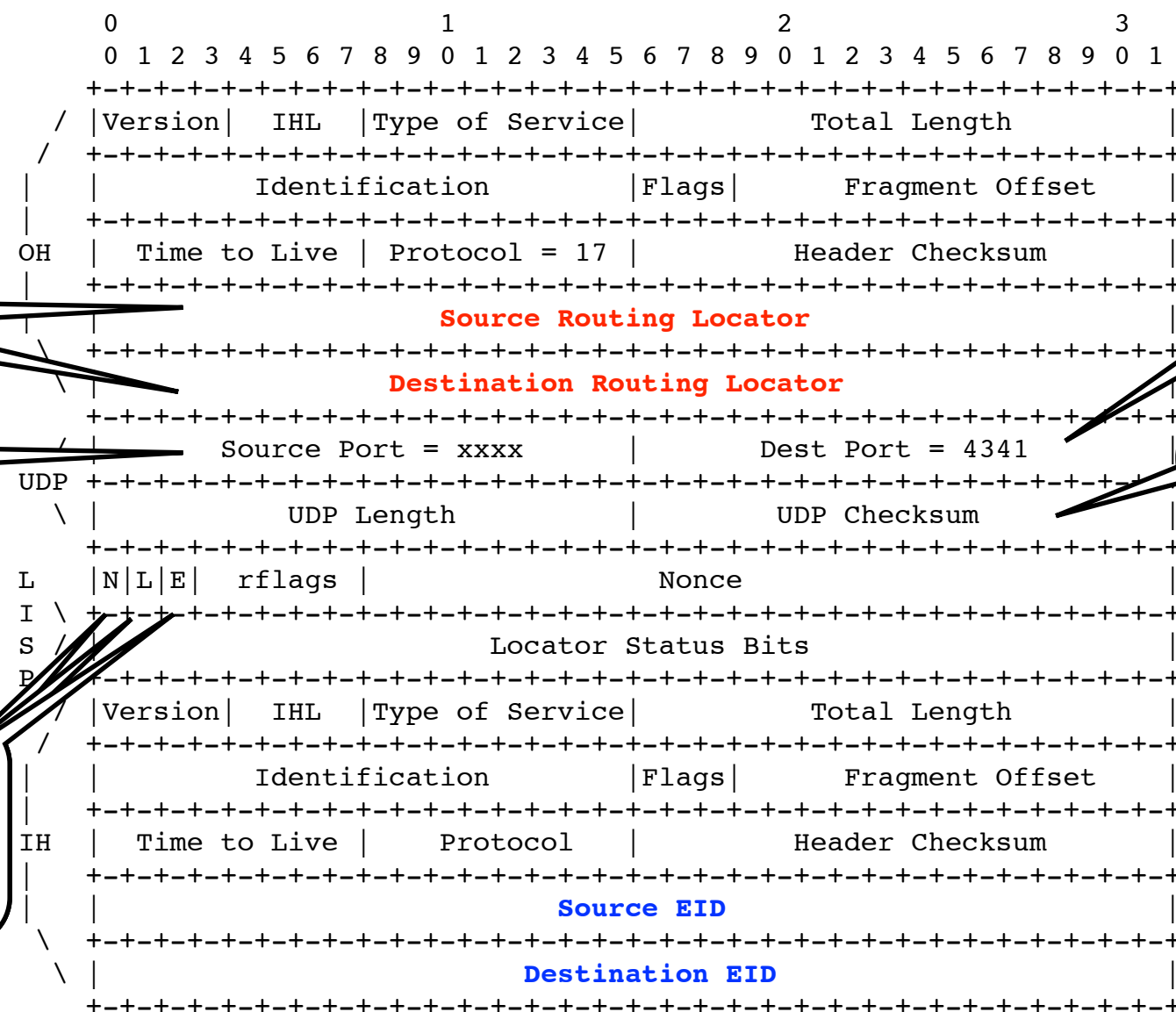
# Header details

(IP(UDP(LISP(IP))))



# Header details

(IP(UDP(LISP(IP))))



Source/Destination Locators

Random Source Port

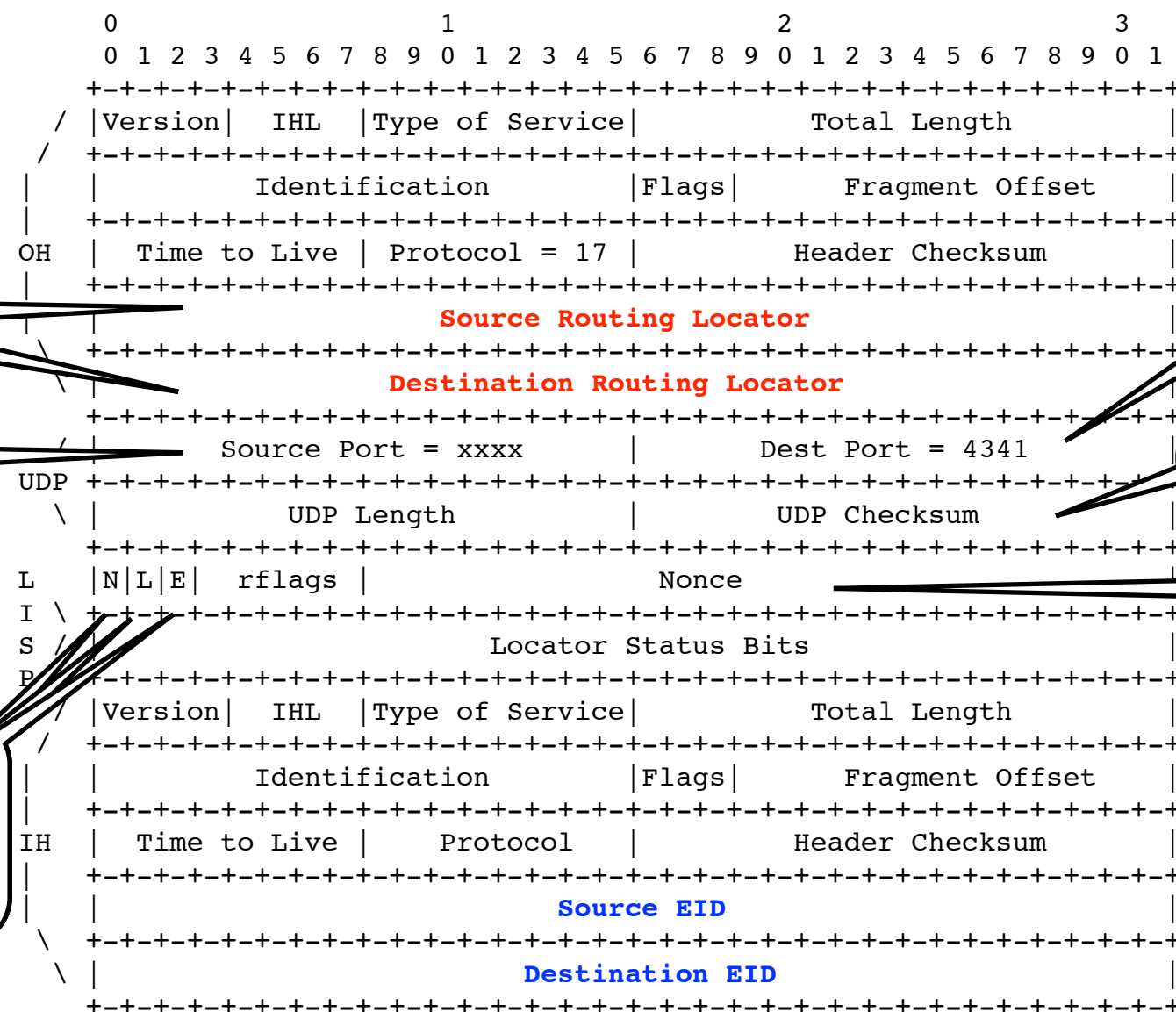
Fixed Destination port

Encap data integrity check

N: Nonce present bit  
L: Locator Status Bits present bit  
E: Echo-Nonce request bit

# Header details

(IP(UDP(LISP(IP))))



Source/Destination Locators

Random Source Port

Fixed Destination port

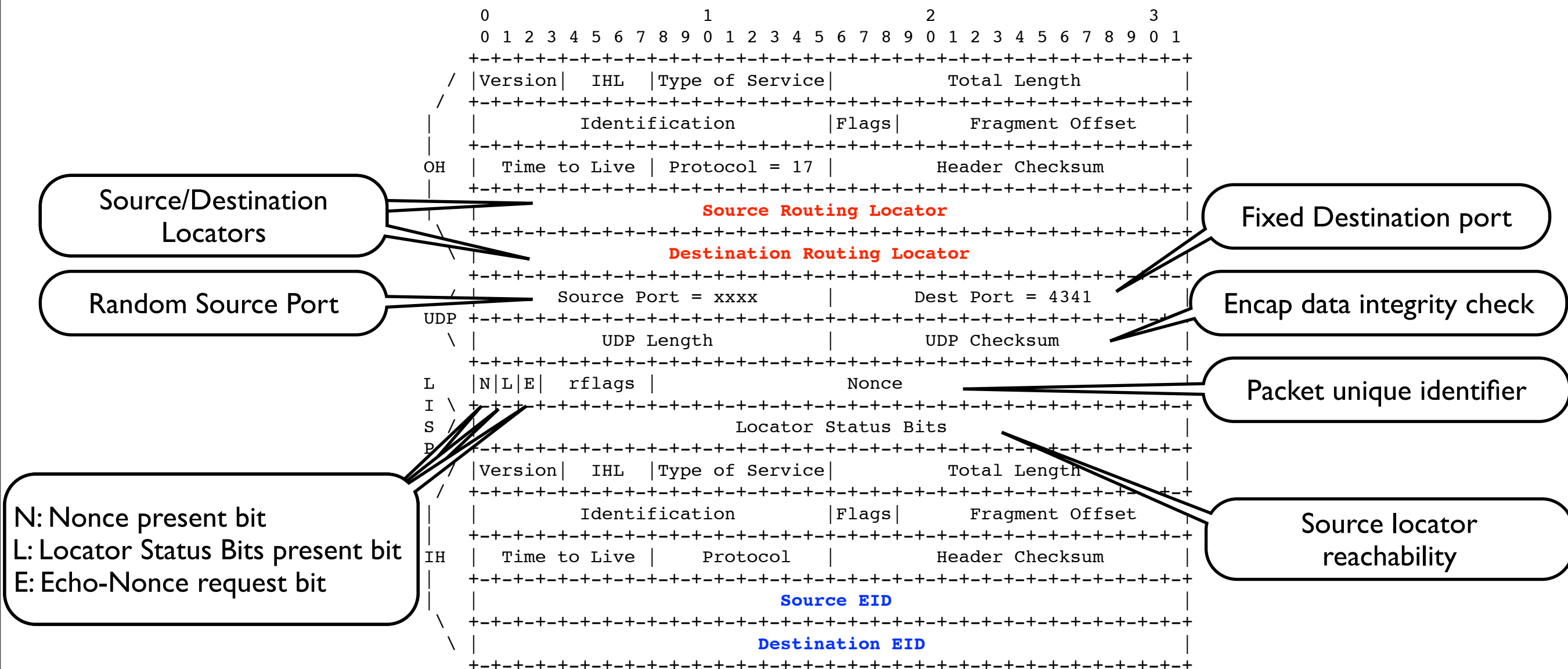
Encap data integrity check

Packet unique identifier

N: Nonce present bit  
L: Locator Status Bits present bit  
E: Echo-Nonce request bit

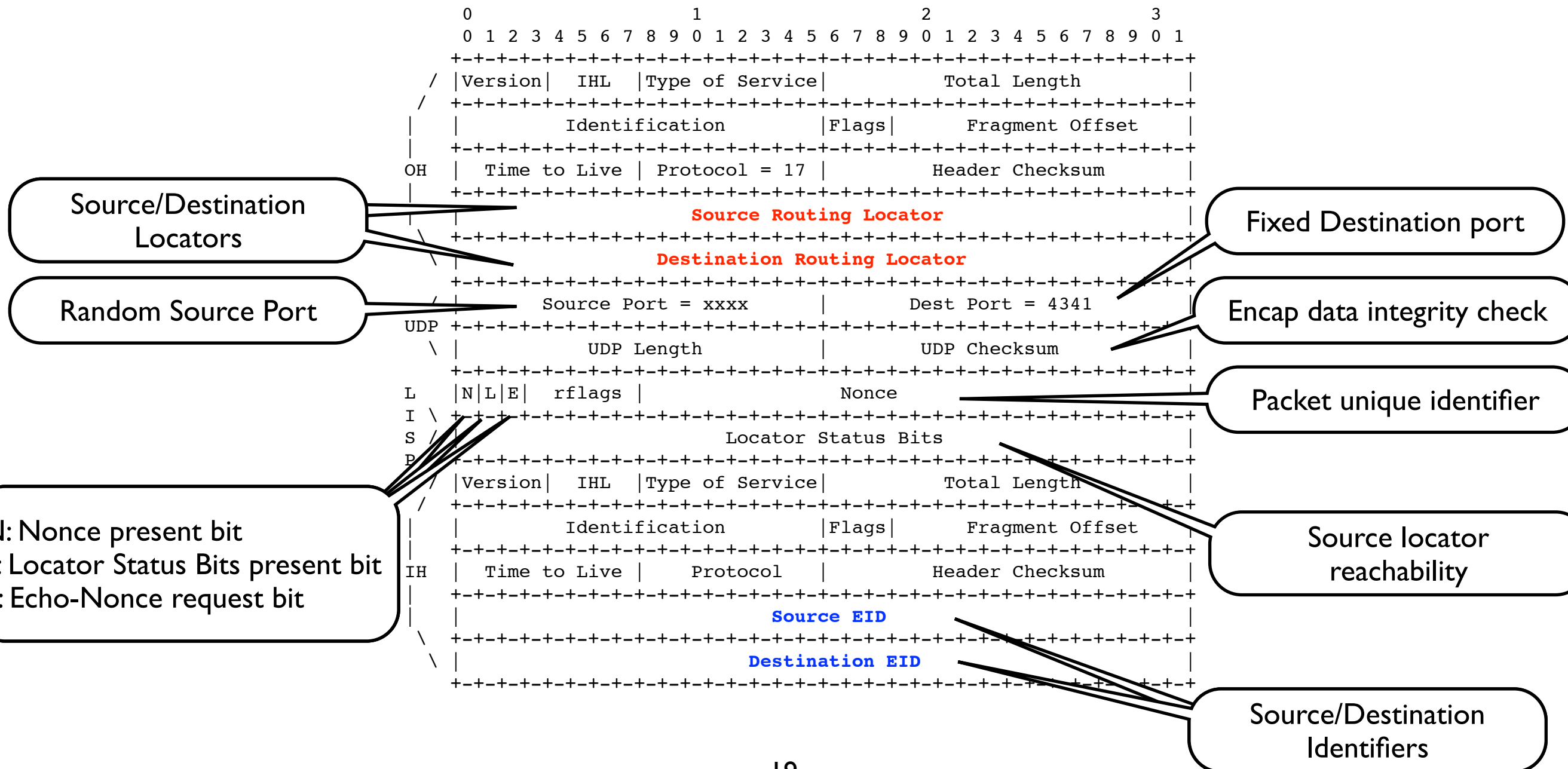
# Header details

(IP(UDP(LISP(IP))))



# Header details

(IP(UDP(LISP(IP))))



# Src/Dst Locators

- Source locator: IP address of the ITR that performed the encapsulation
- Destination locator: IP address of one ETR responsible for the destination EID's prefix
- Locators can be in IPv4 or IPv6

# Src/Dst EID

- Source EID: IP address of the source end-host
- Destination EID: IP address of destination end-hosts
- EIDs can be in IPv4 or IPv6
- AFI(RLOC) can be different AFI(EID)
  - LISP can be an IPv4/IPv6 transition mechanism (but does not support XAFI)

# LISP is over UDP

- UDP to traverse firewalls/NAT, limit the impact of ECMP hashing on reordering...
- Source port is random
  - but per-flow source port is recommended
- Destination port is fixed to 4341
- Checksum is important if IPv6 RLOCs

# Nonce

- Uniquely identify a packet by setting a nonce while encapsulating (at the ITR)
- nonce is optional,  $N=1$  if present
- Used as passive reachability test when  $E=1$
- ITR asks the ETR to start the Echo-Nonce algorithm

# Echo-Nonce Algorithm

- xTR  $I$  wants to know if the RLOC it uses to reach ETR  $E$  is reachable:
  - generate a nonce  $n$  when encap to  $E$
  - set  $E=I$
- Next time the  $E$  sends a packet to  $I$ , is set the nonce to  $n$
- If  $I$  receives the nonce within a given time, it considers the RLOC reachable, otherwise  $E$  is considered unreachable

# Locator Status Bits

- A vector of 32 bits (L bit set to 1 if present)
- Each **source** locator is mapped to one position in the vector

```
if locator_status_bit(i) = 1
    RLOC i is reachable
else
    RLOC i is not reachable
```
- How to determine the position in the vector?
  - Explicit Position (each RLOC has an explicit position in the locator status bits)
  - Locator Status Bits = non-reachability bloom filter
- How to set the bit? What is its meaning?

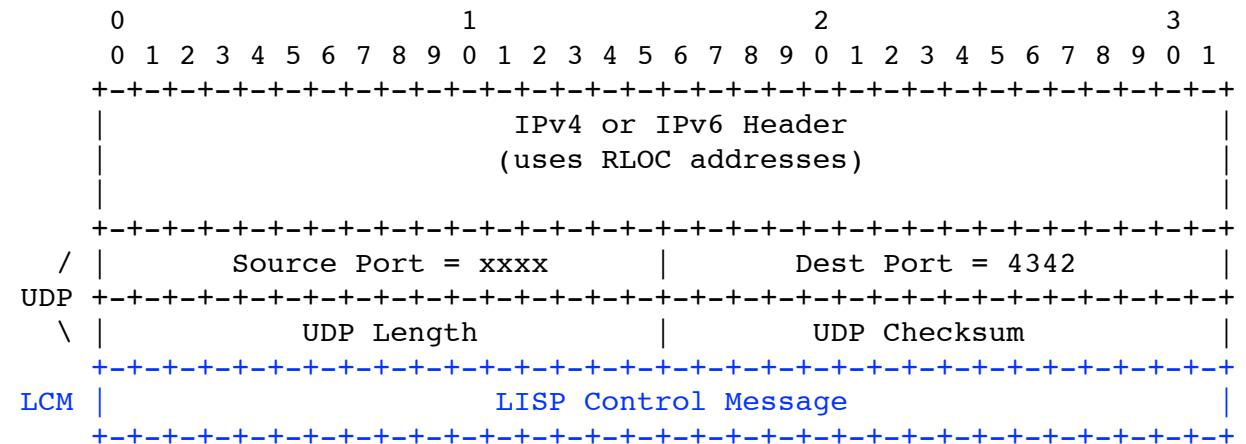
# LISP Control-plane

# LISP mapping

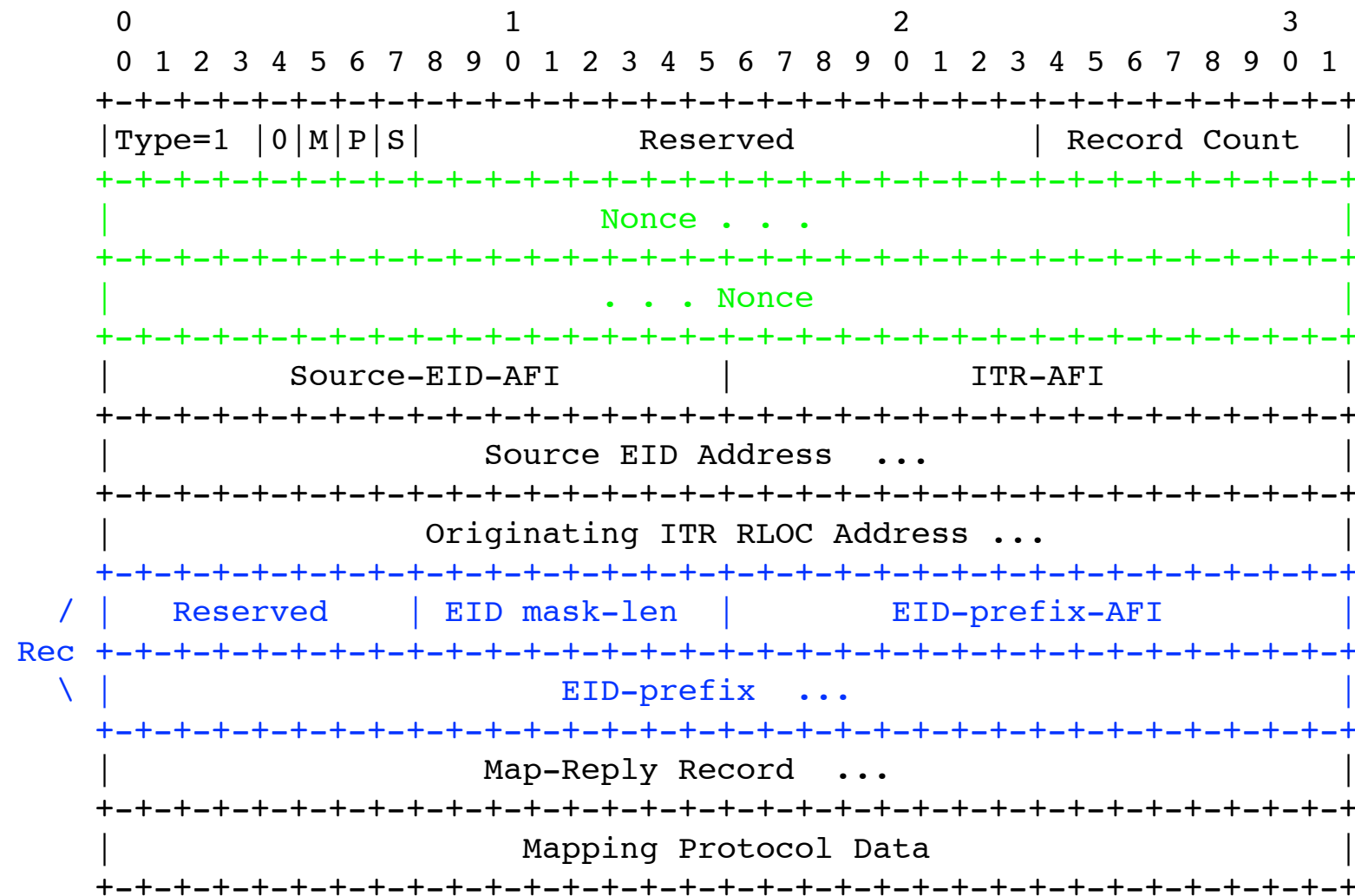
- Possible models for the mapping mechanism
  - Push model
    - LISP ETR routers receive from a protocol to be designed the mapping tables that they need to use to map EIDs onto RLOCs
  - Pull model
    - LISP ETR routers refresh their mapping table by querying the mapping mechanism each time they receive a packet whose mapping is unknown
  - Hybrid models
    - Push is used to place “popular” or “important” mappings on LISP ETR routers and they query for the less important mappings

# LISP Mapping messages

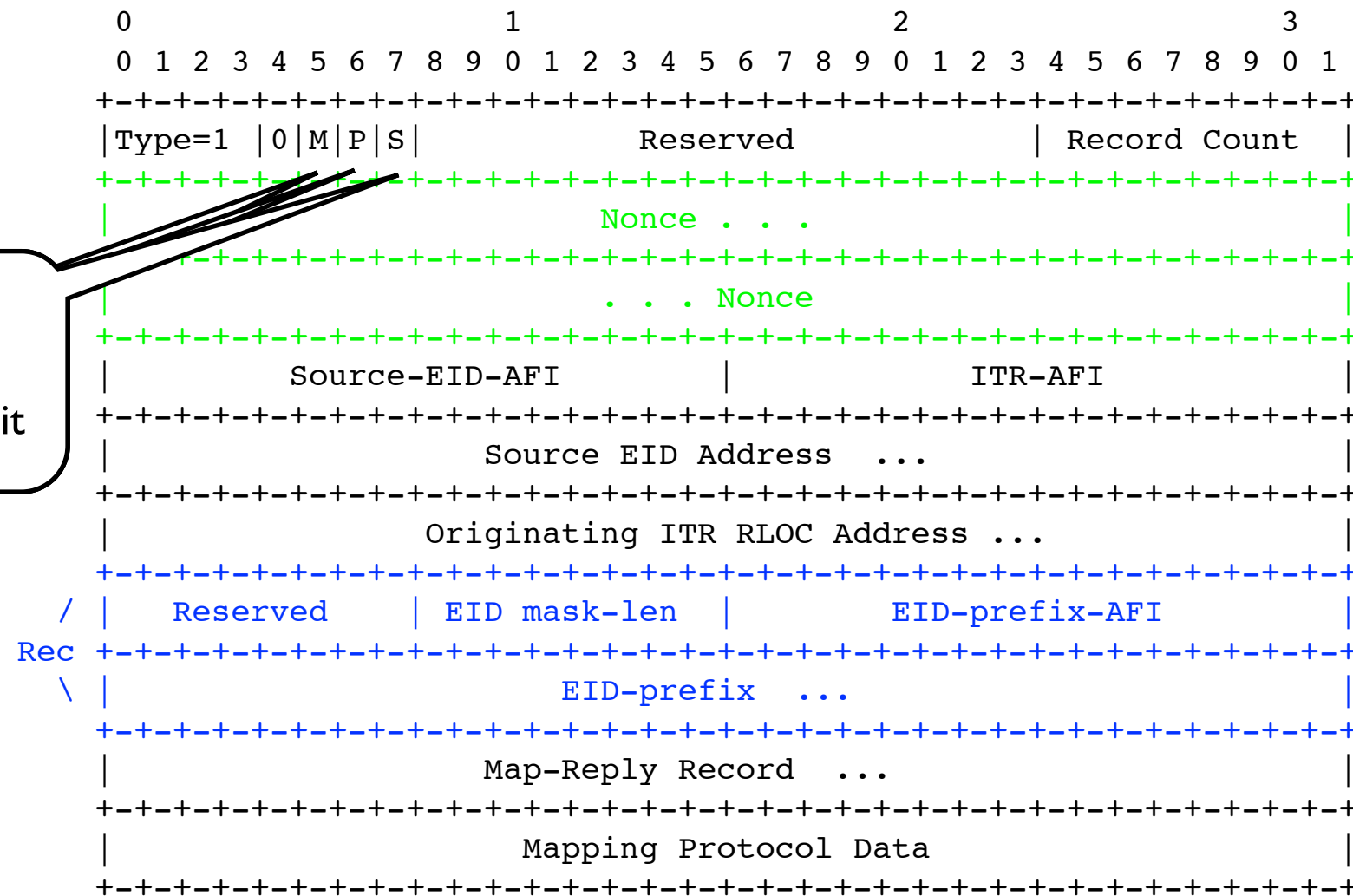
- Map-Request
  - request for a mapping
- Map-Reply
  - provides the mapping requested by a Map-Request
- Map-Register
  - an ETR informs the mapping system that it is responsible for an EID prefix
  - provides the RLOCs where the mapping can be found



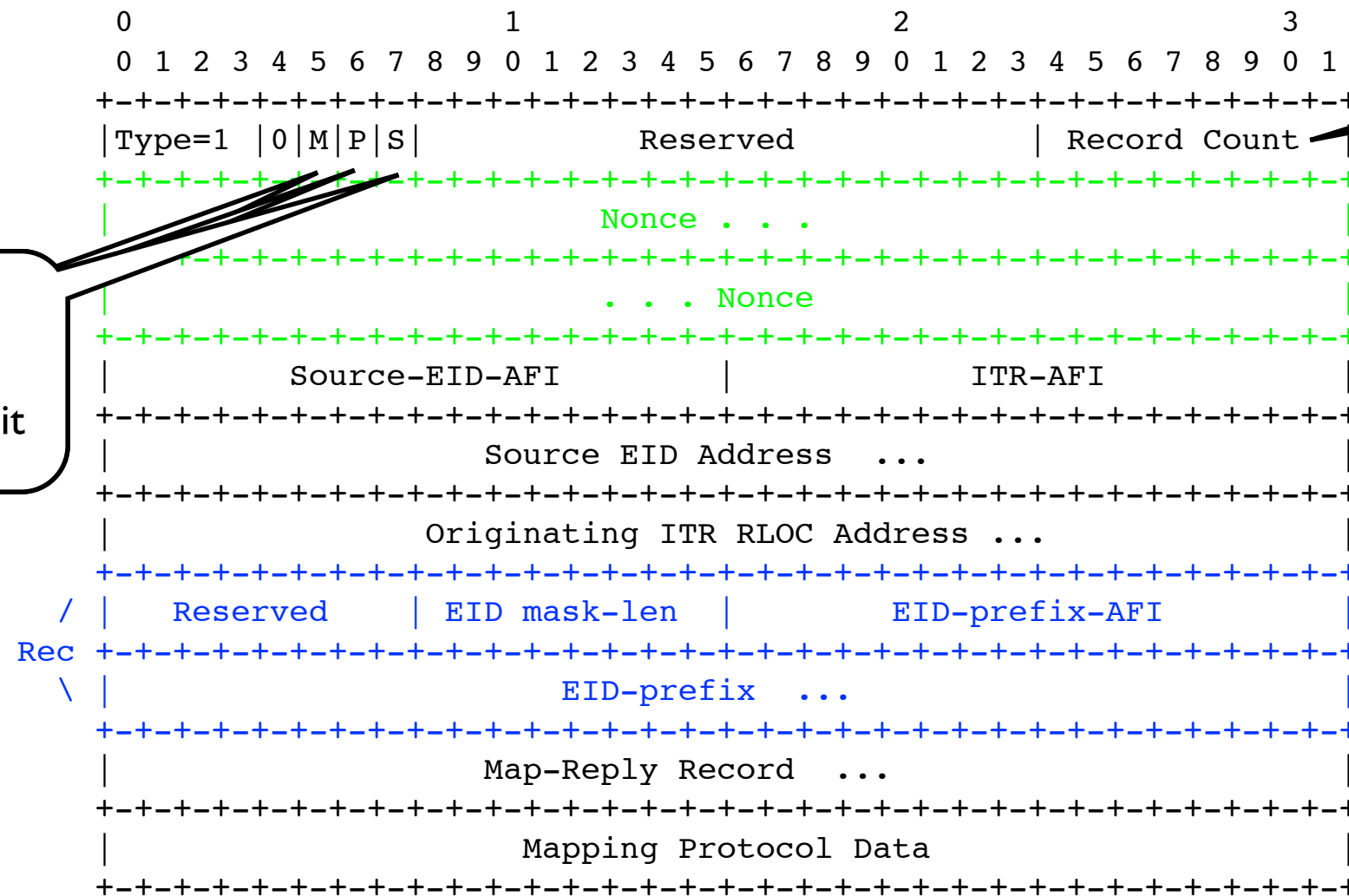
# Map-Request



# Map-Request



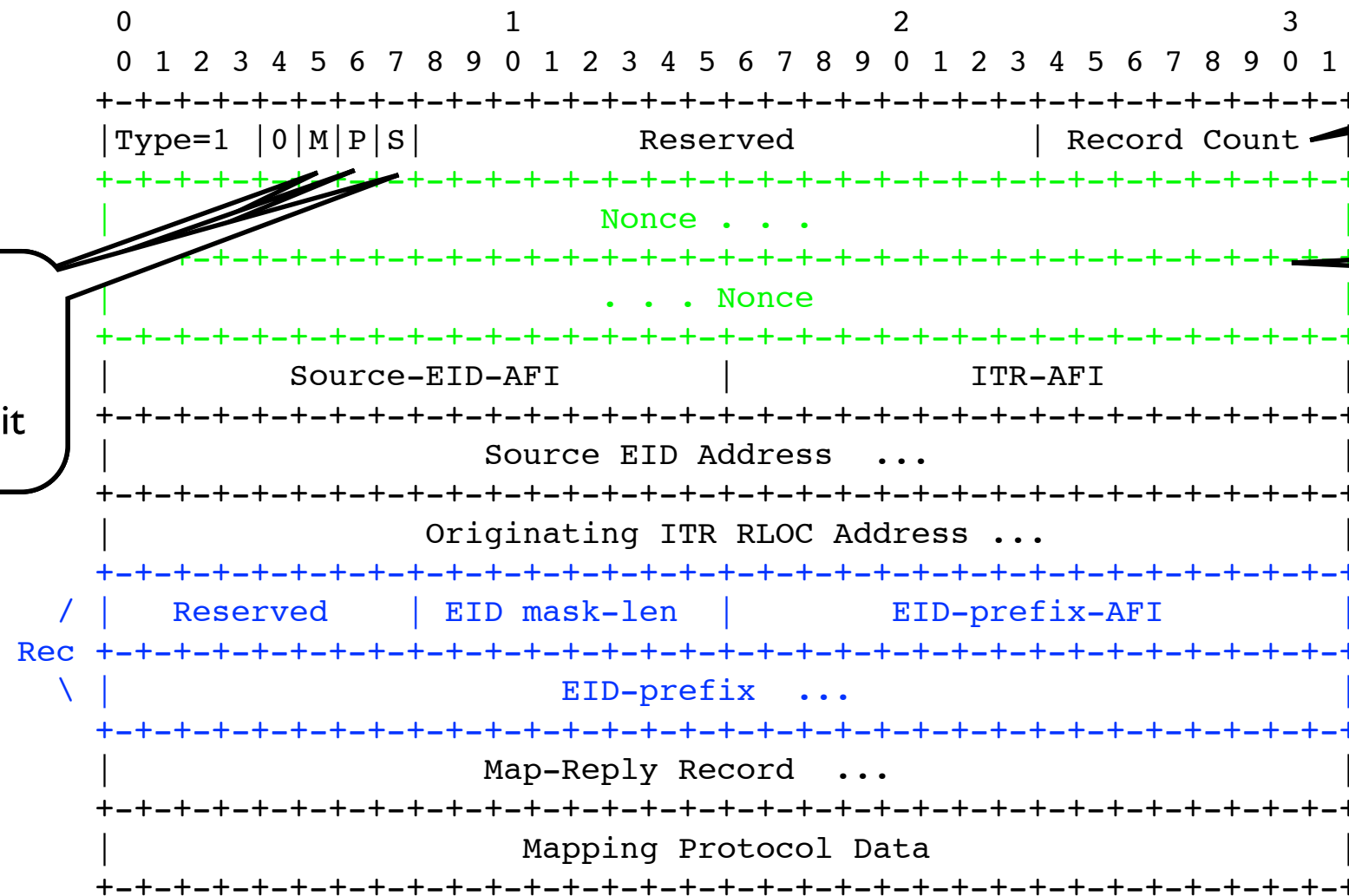
# Map-Request



Number of **records** in the request

M: Map-Reply  
P: probing bit  
S: Solicitation bit

# Map-Request

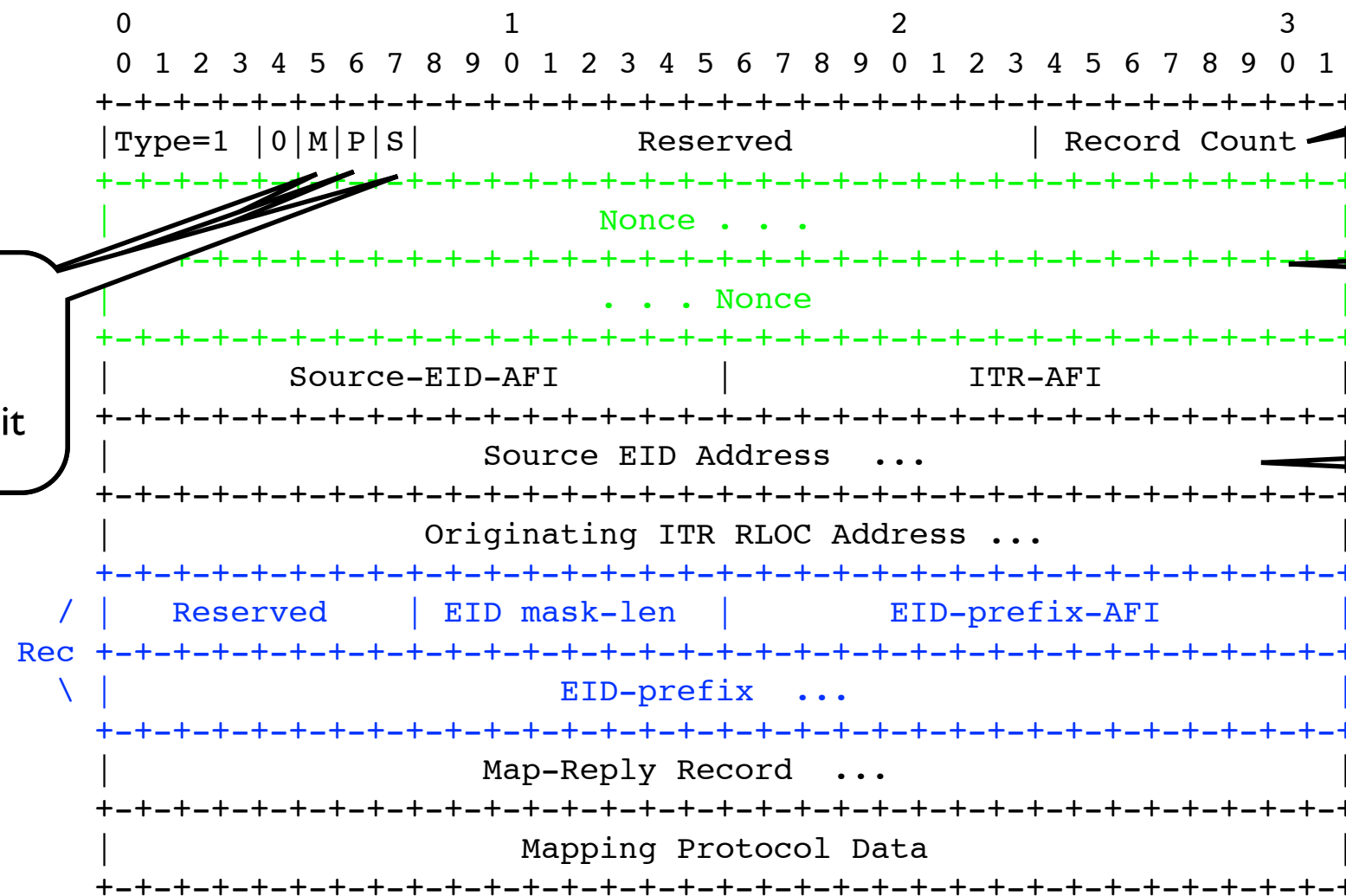


Number of **records** in the request

64-bits **Nonce**, uniquely identifying the request

M: Map-Reply  
P: probing bit  
S: Solicitation bit

# Map-Request



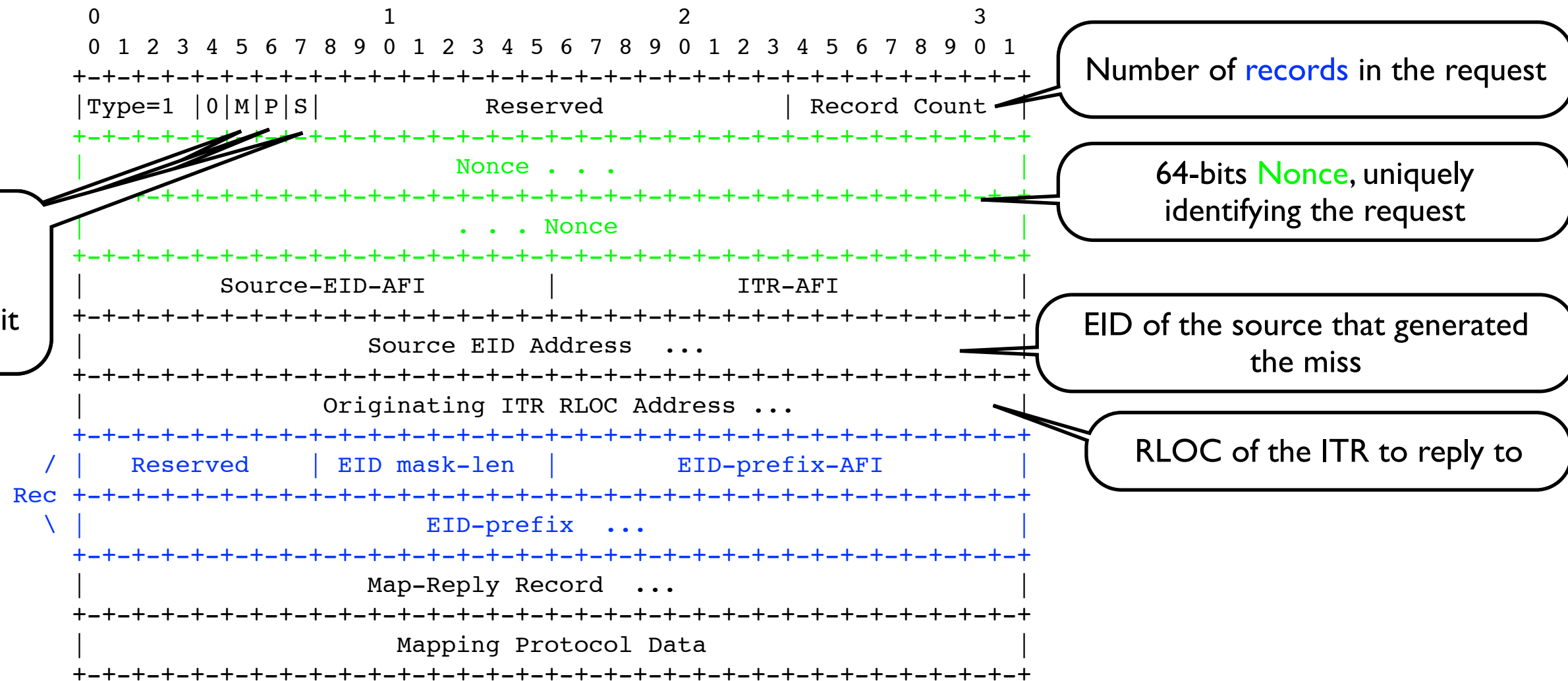
Number of **records** in the request

64-bits **Nonce**, uniquely identifying the request

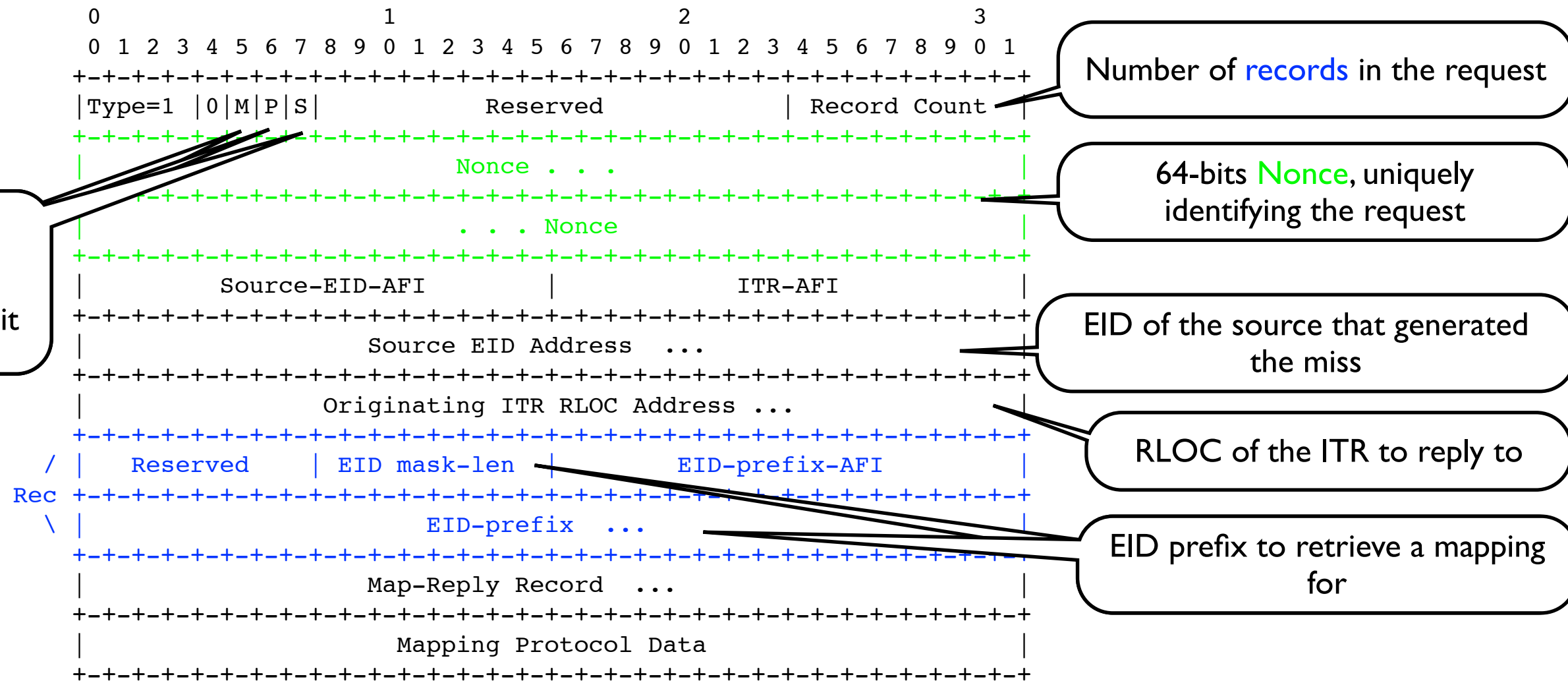
EID of the source that generated the miss

M: Map-Reply  
P: probing bit  
S: Solicitation bit

# Map-Request



# Map-Request

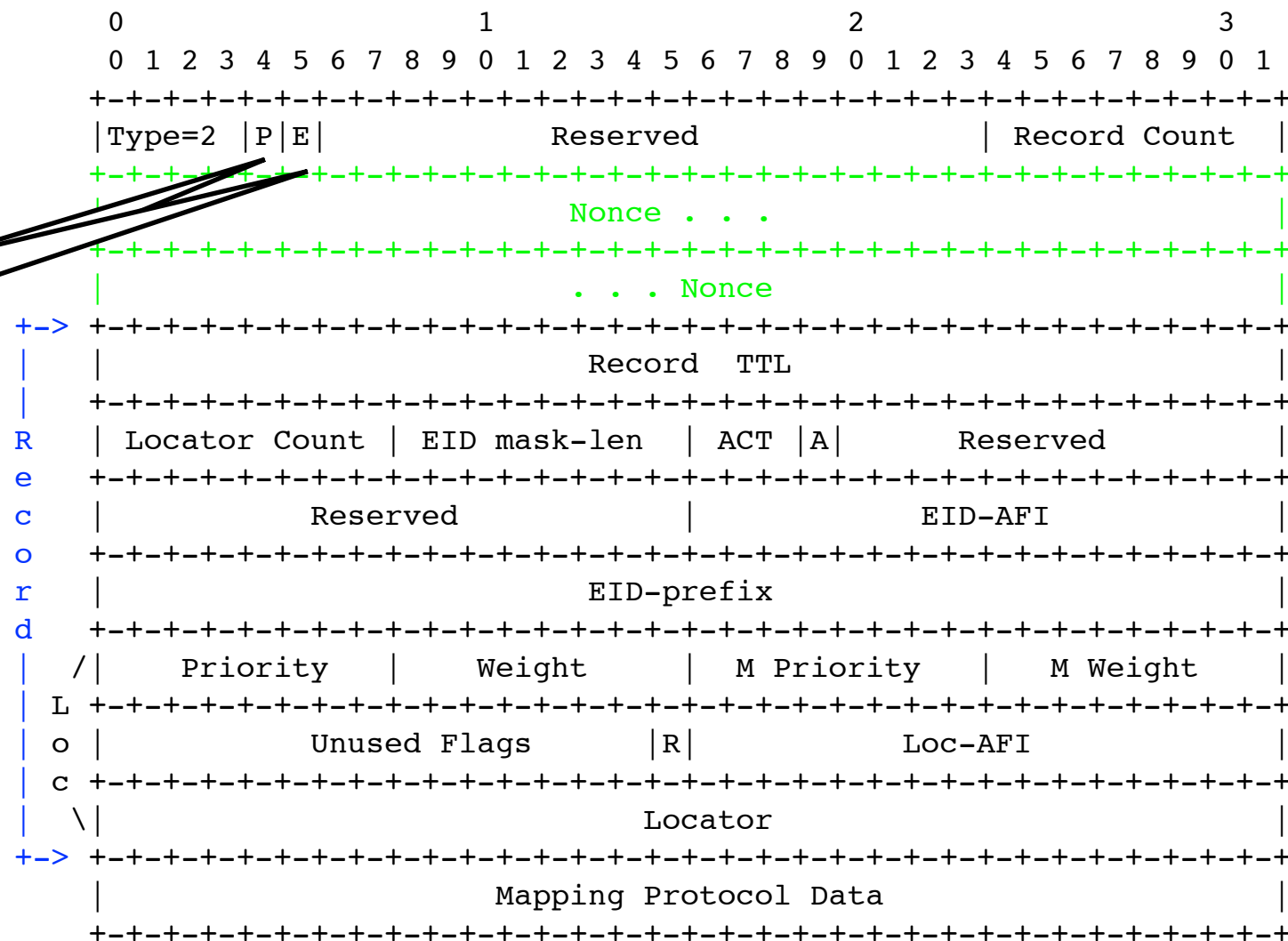


# Map-Reply

[illegible]

# Map-Reply

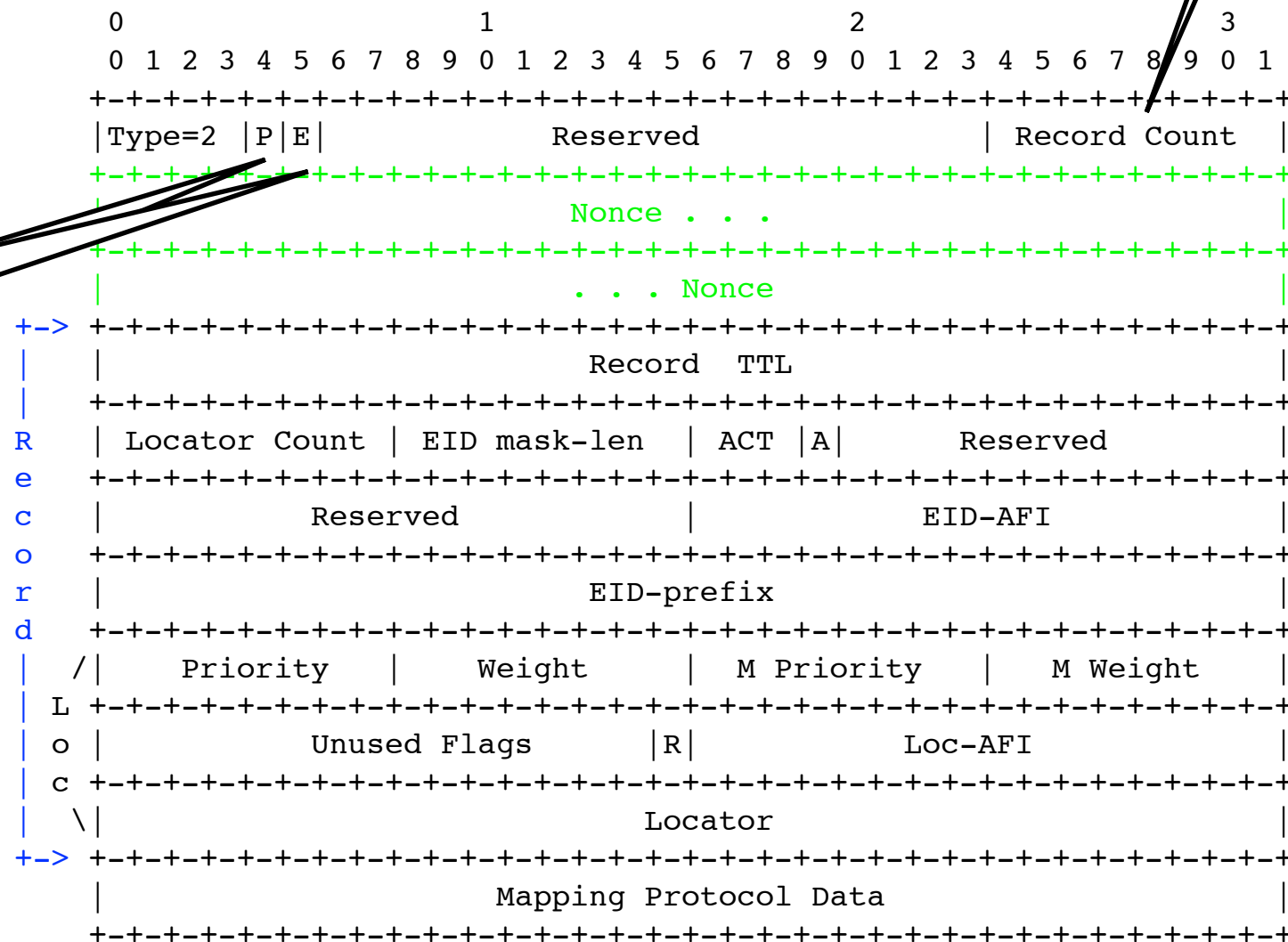
P: probing bit  
E: Echo-nonce  
capable



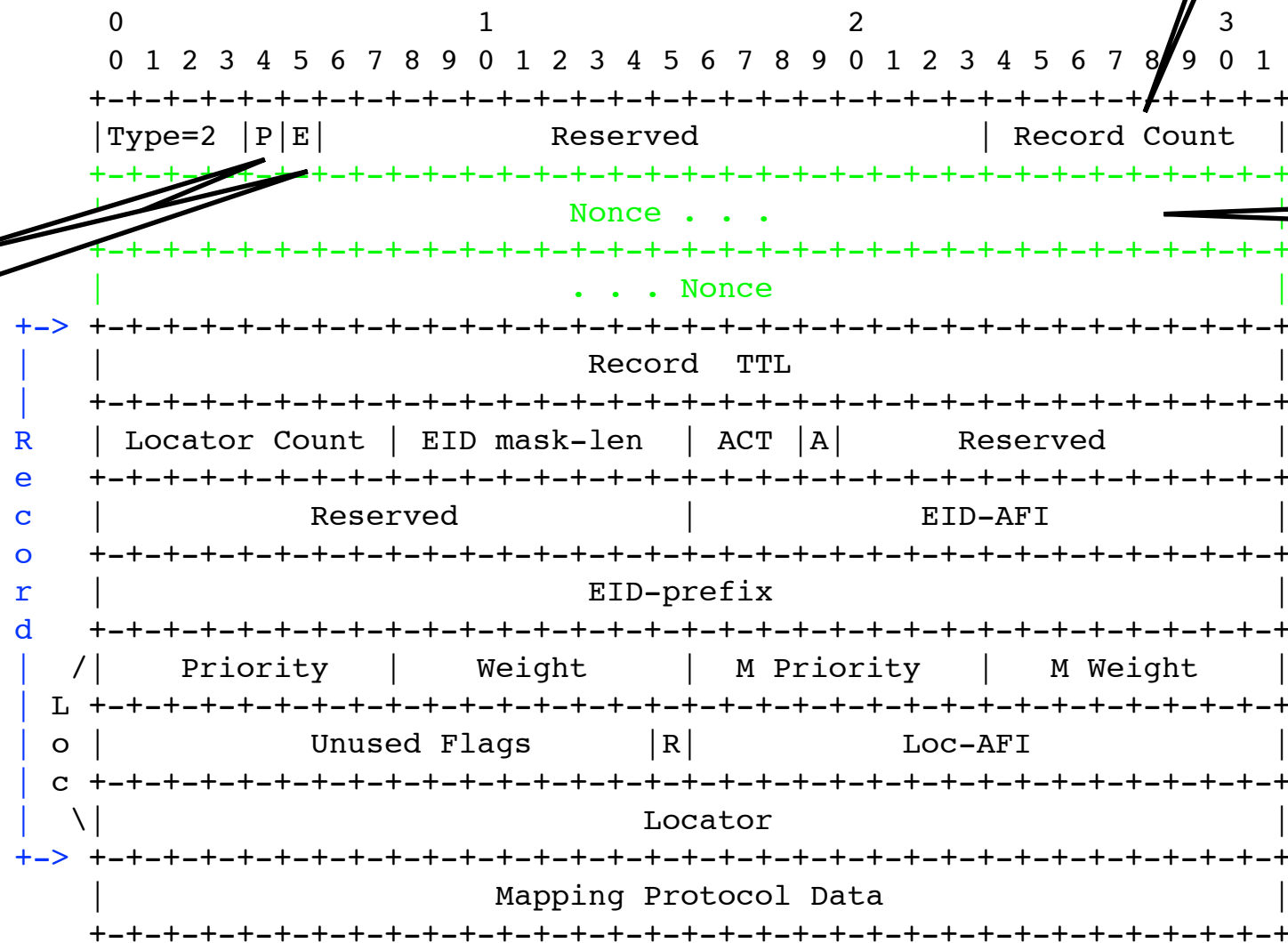
# Map-Reply

Number of **records** in the request

P: probing bit  
E: Echo-nonce  
capable



# Map-Reply

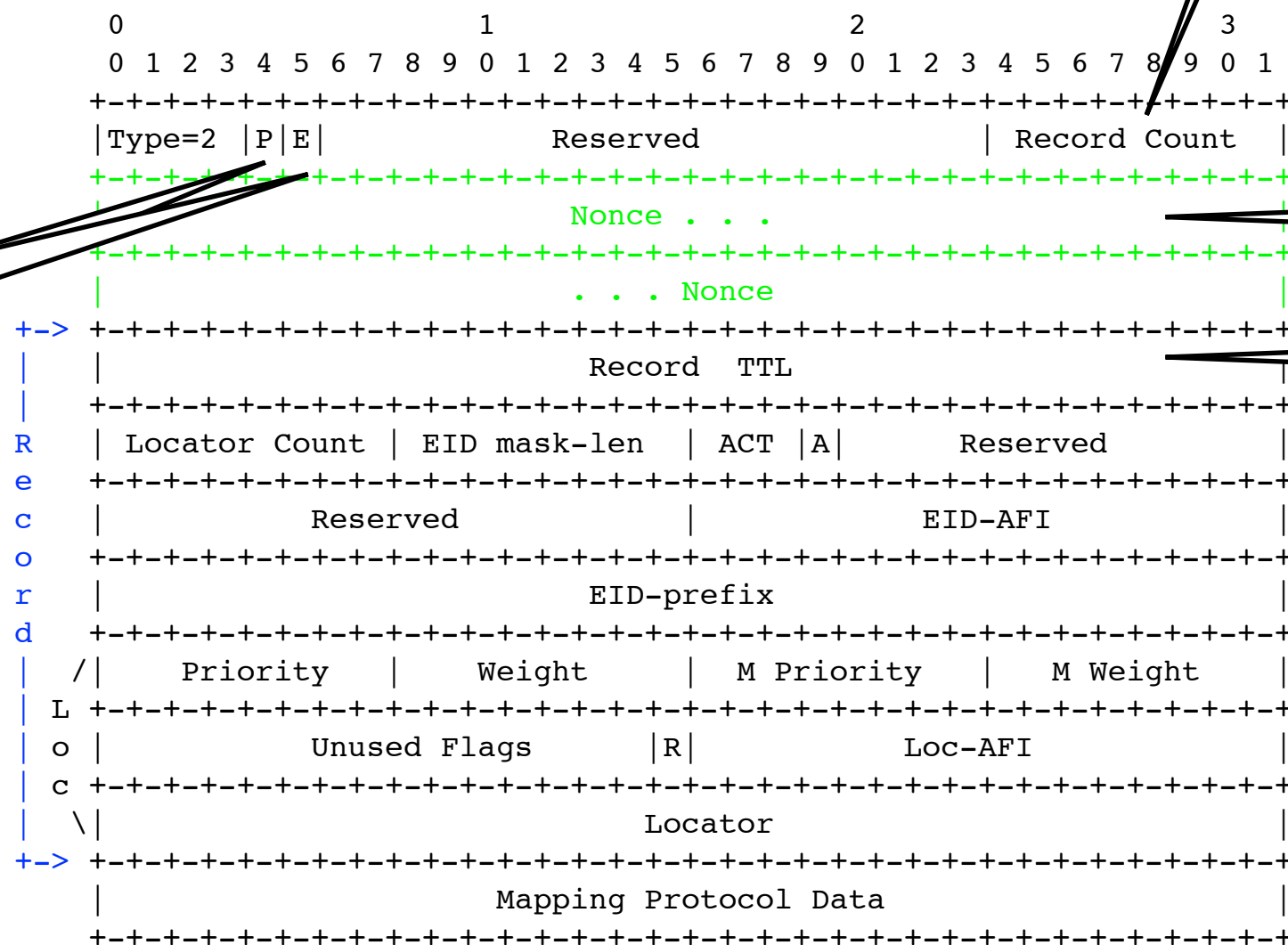


P: probing bit  
E: Echo-nonce capable

Number of **records** in the request

Copied from the Map-Request

# Map-Reply



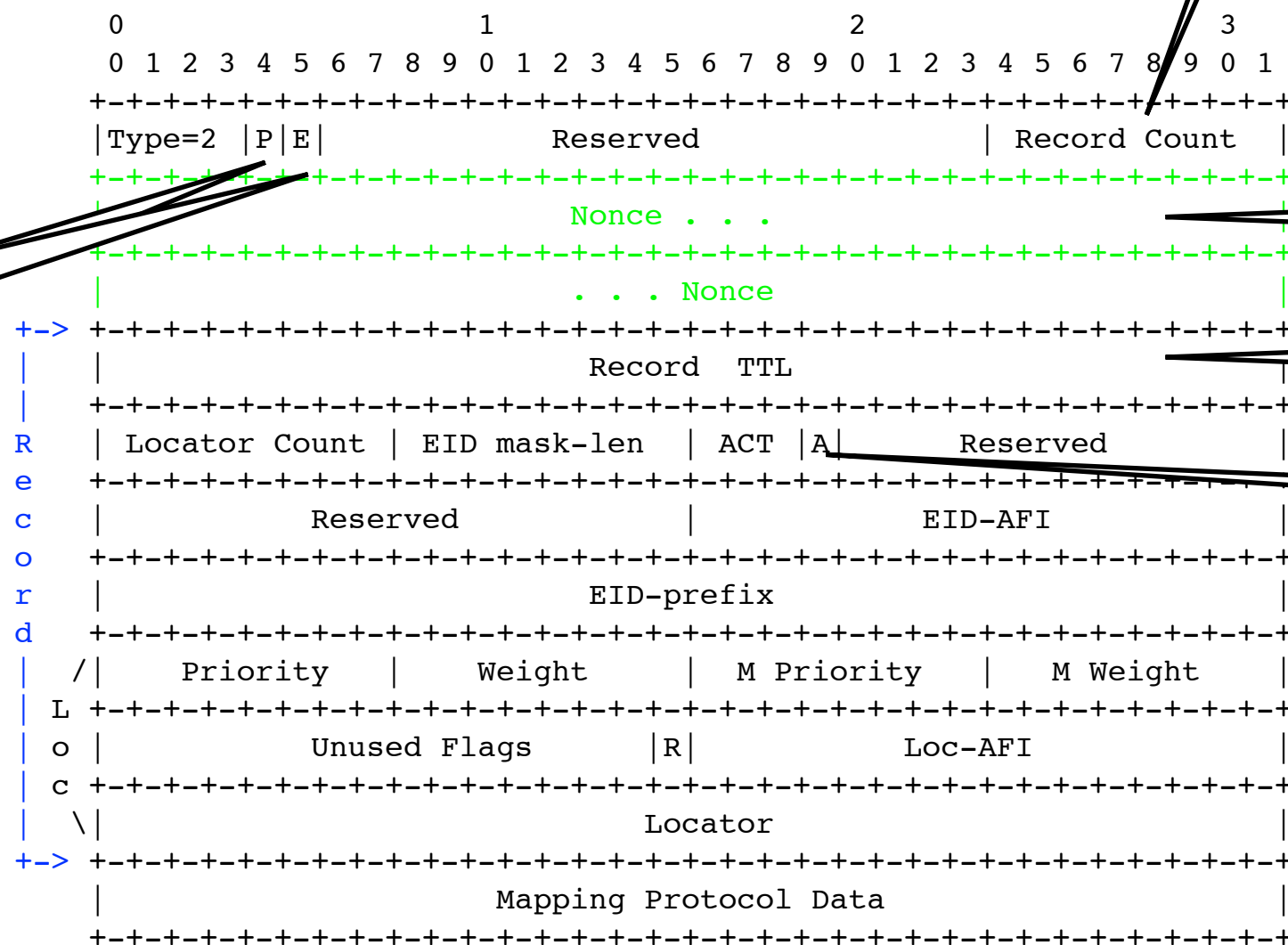
P: probing bit  
E: Echo-nonce  
capable

Number of **records** in the request

Copied from the Map-Request

Lifetime of the record in min

# Map-Reply



Number of **records** in the request

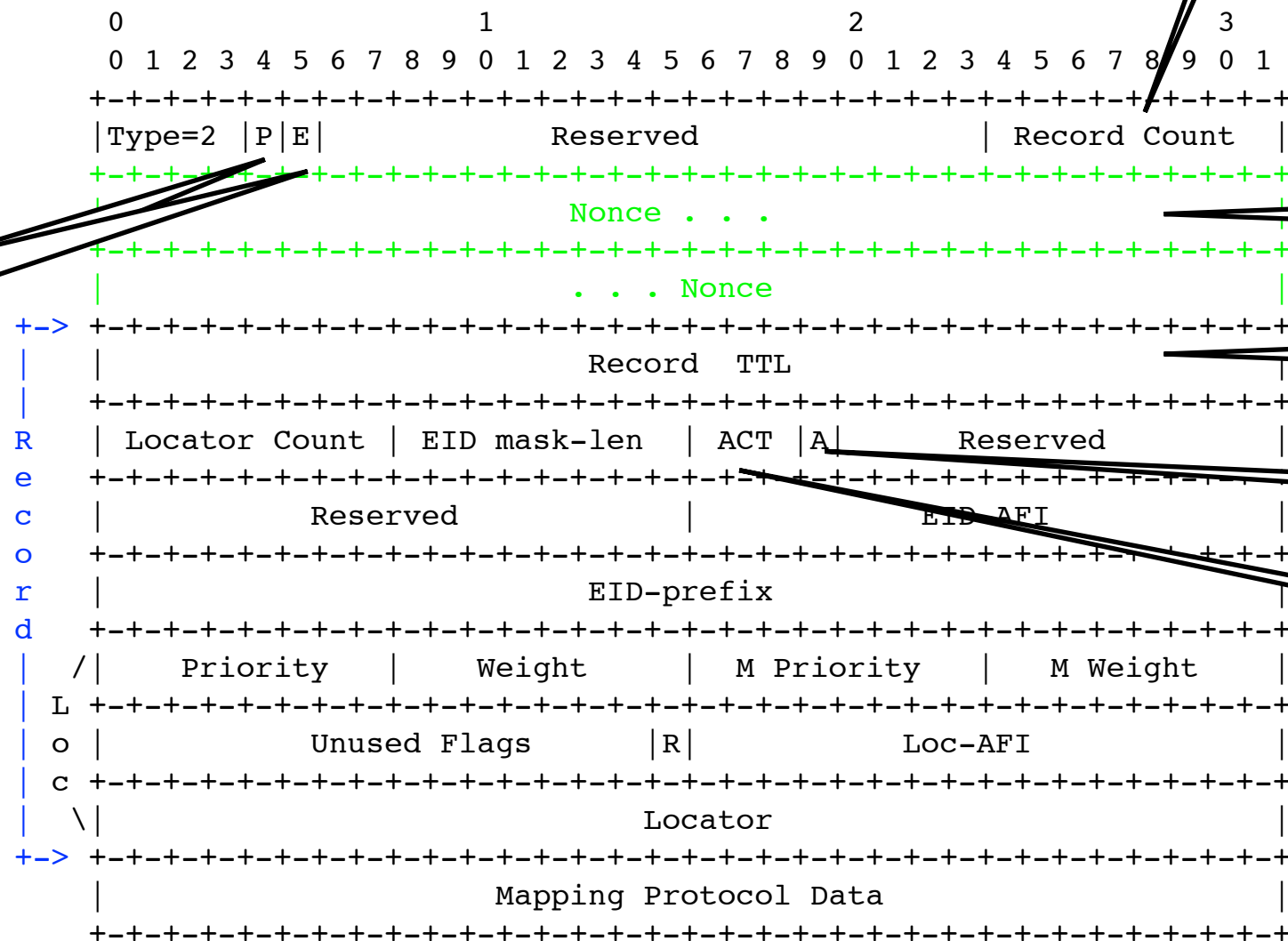
Copied from the Map-Request

Lifetime of the record in min

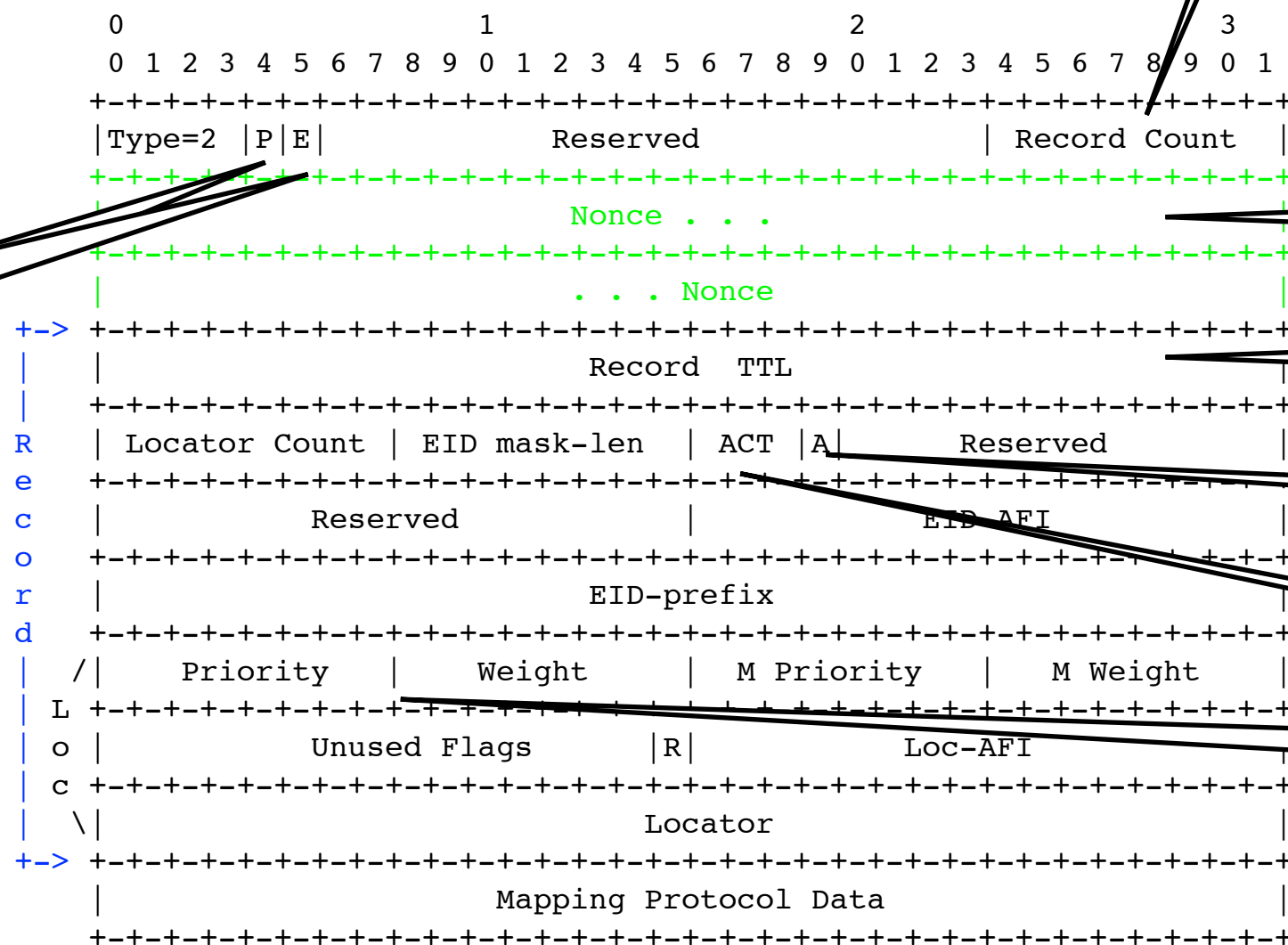
Authoritative bit

P: probing bit  
E: Echo-nonce  
capable

# Map-Reply



# Map-Reply



P: probing bit  
E: Echo-nonce capable

Number of **records** in the request

Copied from the Map-Request

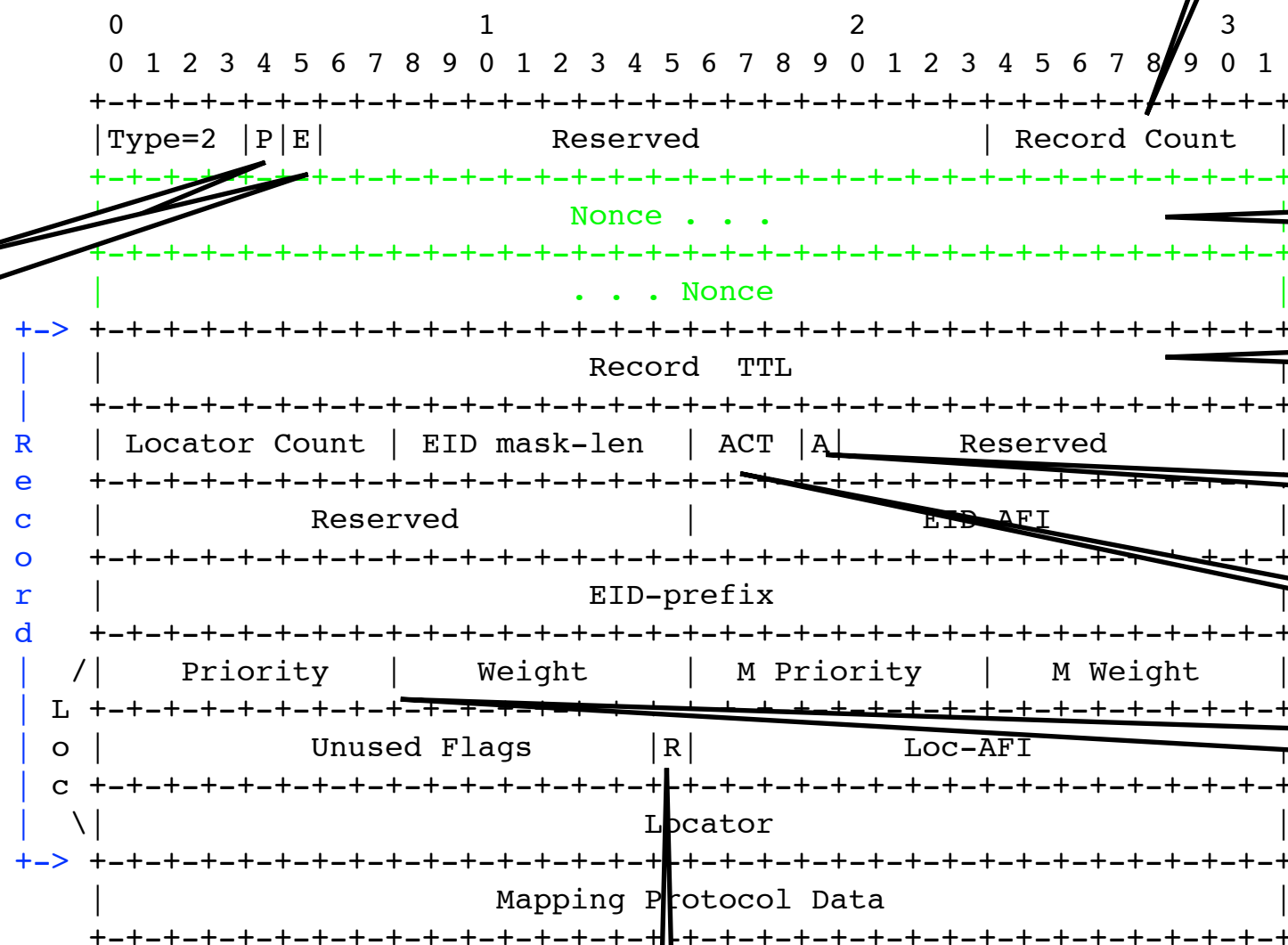
Lifetime of the record in min

Authoritative bit

Negative reply action

Priority : RLOCs with lower priority are preferred. If several have same priority, load balance among them  
Weight : percentage of traffic to this RLOC when load balancing is active. (M for multicast)

# Map-Reply



Number of **records** in the request

P: probing bit  
E: Echo-nonce capable

Copied from the Map-Request

Lifetime of the record in min

Authoritative bit

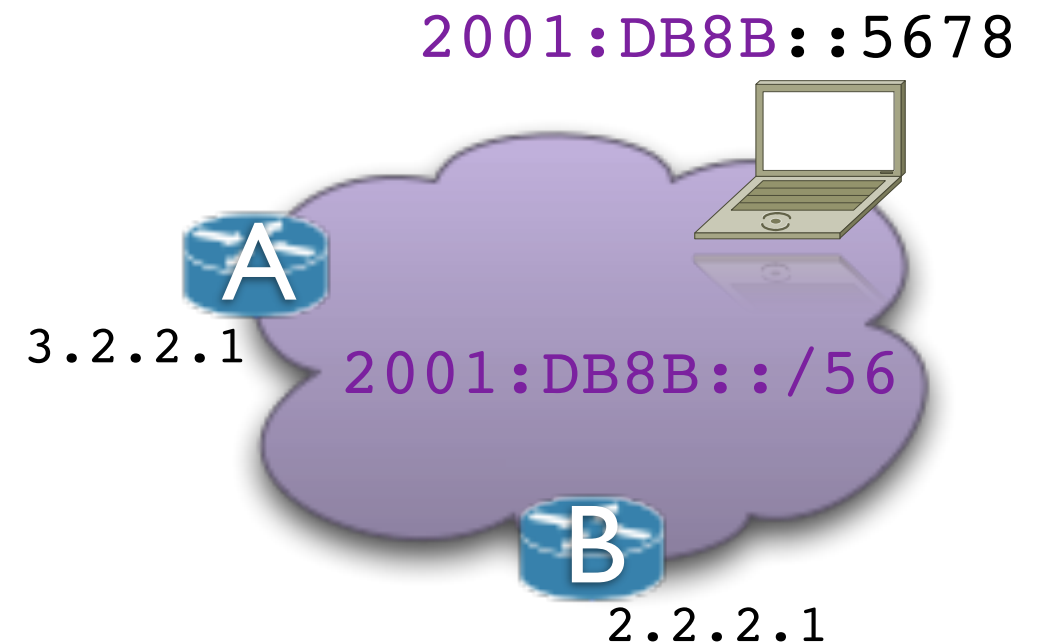
Negative reply action

Is record is reachable from responder's viewpoint ?

Priority : RLOCs with lower priority are preferred. If several have same priority, load balance among them  
Weight : percentage of traffic to this RLOC when load balancing is active. (M for multicast)

# How to control incoming traffic?

- LISP site can control incoming traffic with Weight and Priority



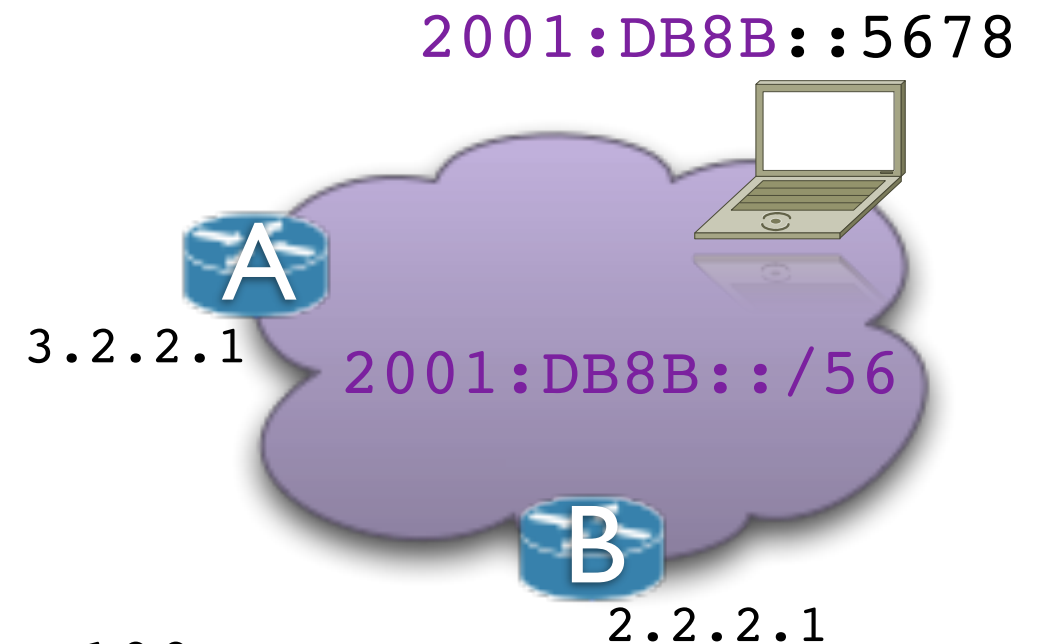
# How to control incoming traffic?

- LISP site can control incoming traffic with VWeight and Priority

- A Primary, B Backup

`2001:DB8B::/56`

- `3.2.2.1, prio: 1, weight: 100`
- `2.2.2.1, prio: 99, weight: 100`



# How to control incoming traffic?

- LISP site can control incoming traffic with VWeight and Priority

- A Primary, B Backup

`2001:DB8B::/56`

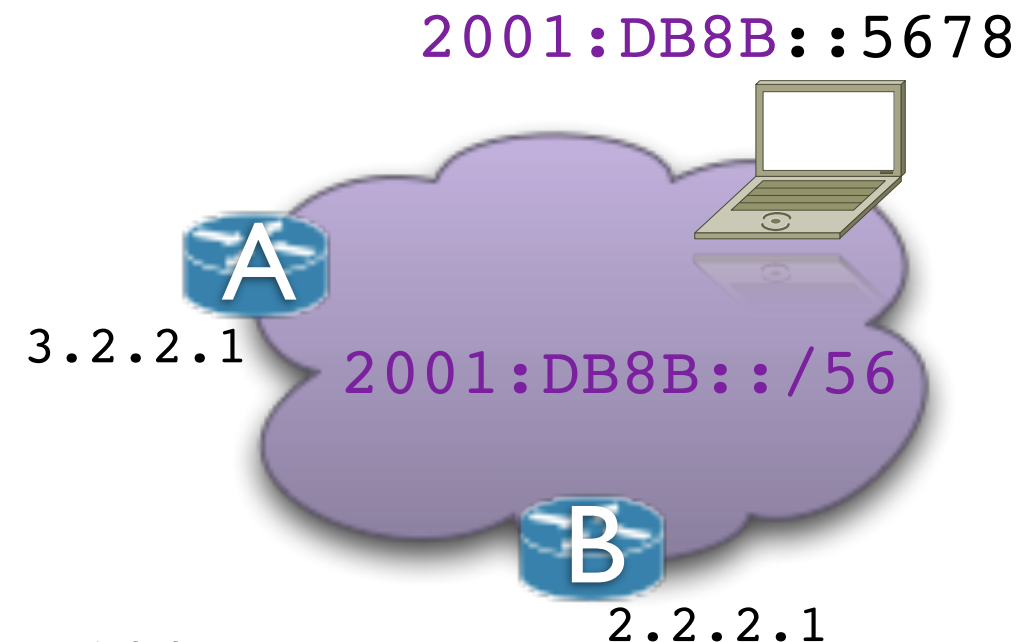
- `3.2.2.1`, prio: 1, weight: 100
- `2.2.2.1`, prio: 99, weight: 100

- A 60%, B 40%

`2001:DB8B::/56`

- `3.2.2.1`, prio: 1, weight: 60
- `2.2.2.1`, prio: 1, weight: 40

LISP ITR will load balance layer 4 flows by using hash as in ECMP



# How to design the deployment?

- Should we allow more specific EID prefixes?
- How EID must be allocated?
  - Per interface? Per hosts? Per service? Per site?
- How to build the mappings?
  - 1 RLOC per EID? Several RLOC per EID?

# Mapping Systems

# NERD: A Not-so-novel EID to RLOC Database

# NERD

- The only proposed push model
  - Composed of 4 parts
    - a network database format;
    - a change distribution format;
    - a database retrieval/bootstrapping method;
    - a change distribution method
- Principles
  - An authority computes the mapping database based on the stored registrations
  - The database signed by the authority is stored on servers
  - ITR poll regularly the database servers to update their own mapping database

# LISP+ALT

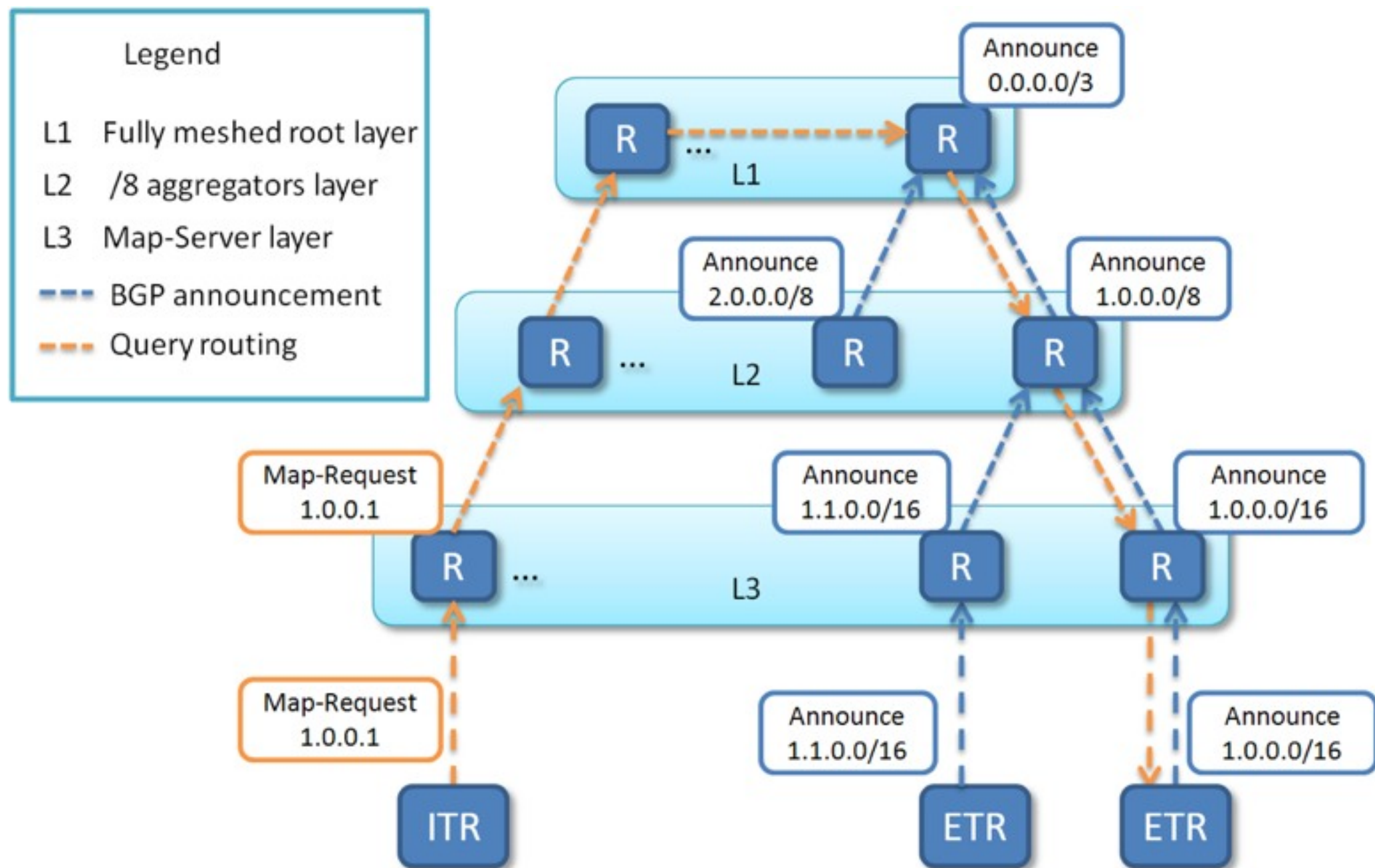
# LISP+ALT

- An pull model mapping system
- A mapping mechanism that relies on an alternate topology to distribute mapping requests to mapping servers
- LISP ITR routers sending mapping request messages to ALT routers
- ALT routers forward those mapping messages between themselves on an overlay topology built by using GRE tunnels

# LISP+ALT

- BGP announces **where** the mappings can be found
- Map-Requests are forwarded on the ALT
- Map-Replies are forwarded on the legacy Internet (directly sent to the ITRs' RLOC)
- **! BGP does not give the mappings !**

# LISP+ALT



# LISP+ALT issues

- Complex system with tunnels, BGP protocol (no discussion about policies), ...
- Still relies on lots of error-prone manual configuration
- Scalability will depend on whether aggregation will be possible
- If mapping requests are lost due to congestion, difficult to diagnose the problem or send them via another path
- Security needs to be studied

# LISP+ALT, big question

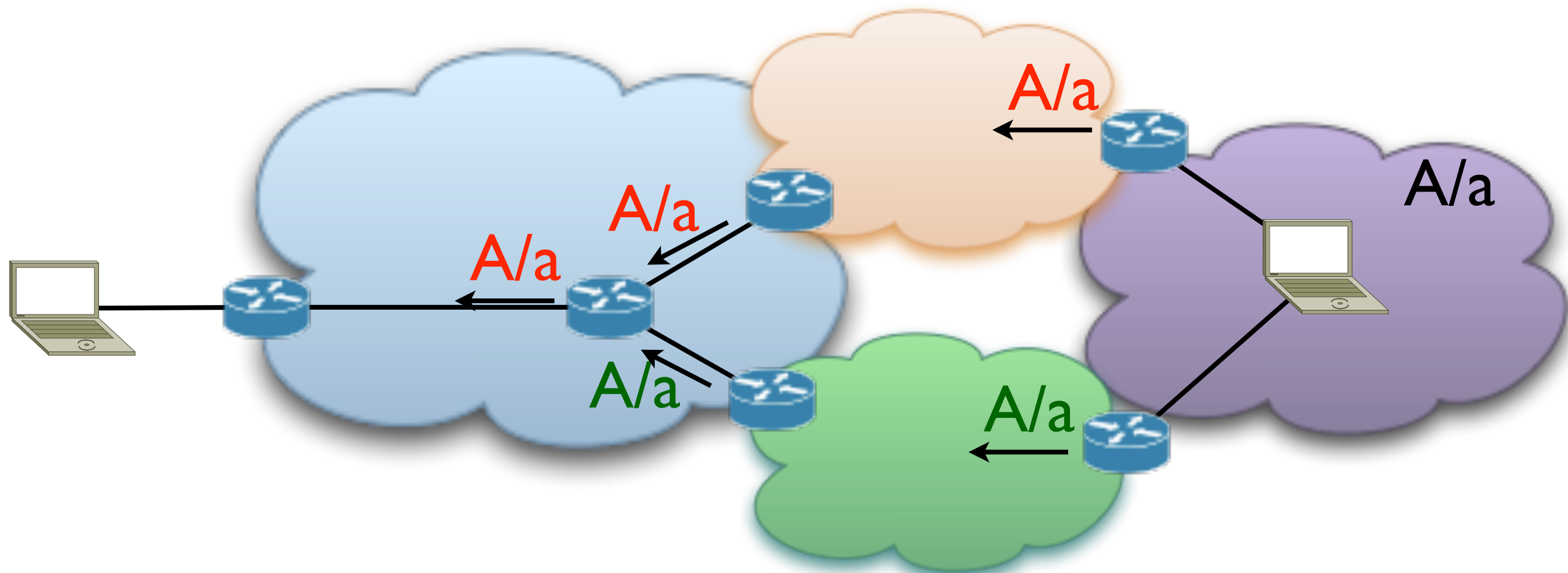
- How to deploy it?
  - aggregation vs recovery vs TE
  - how to cache the mappings?

# The reachability problem

# The reachability problem

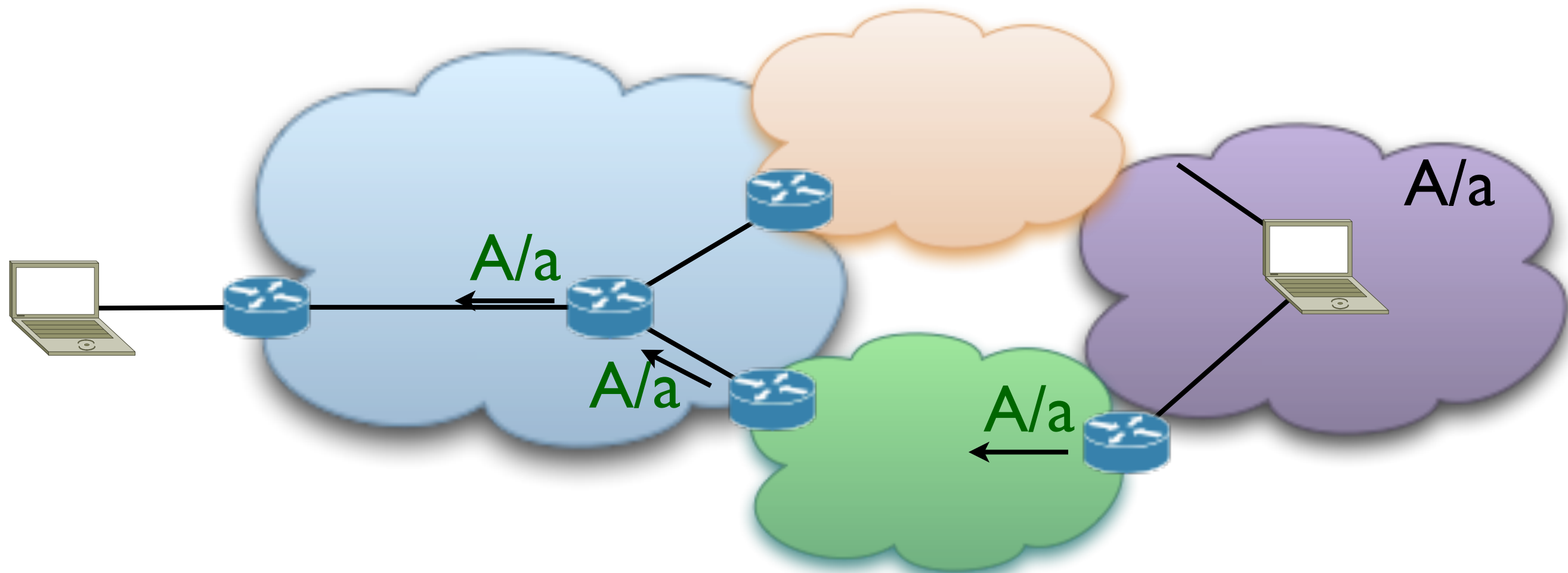
- Today, preserving the prefixes reachability is mainly performed locally
- In LISP, the legacy Internet is EID agnostic

# The reachability problem in today's Internet



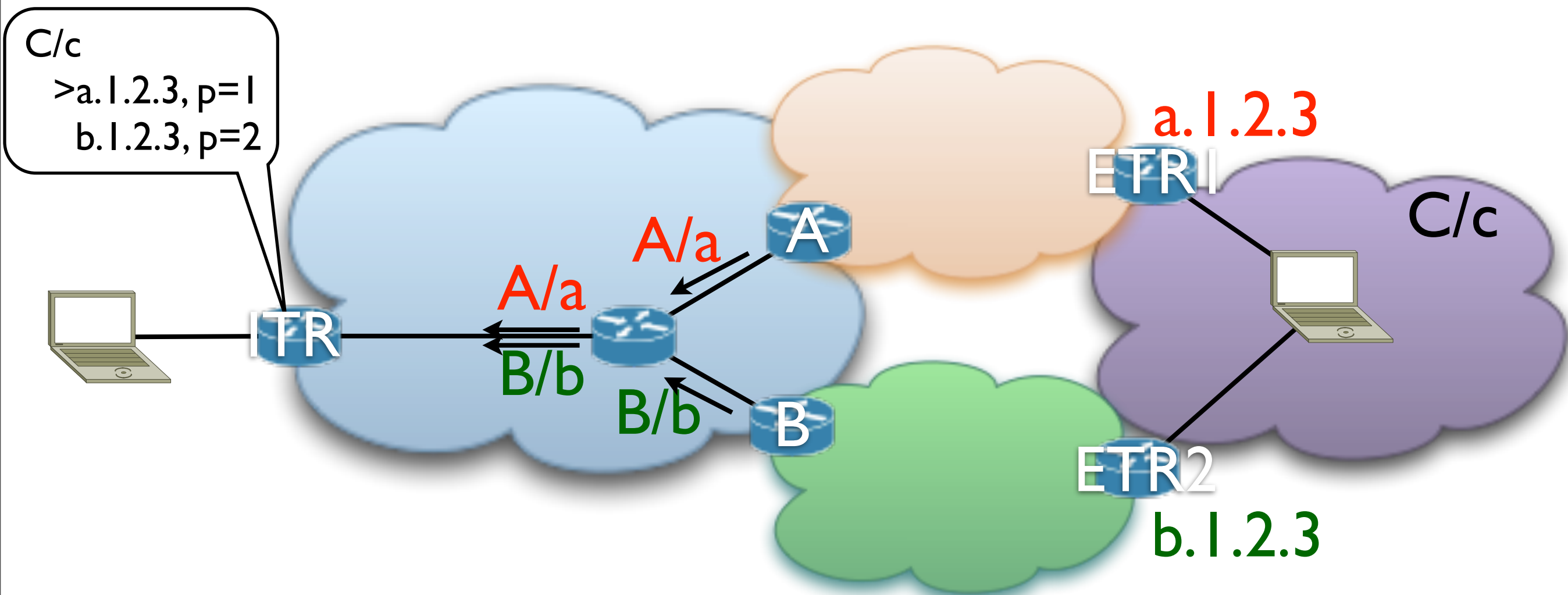
- In today's Internet, routing protocols converge after a link failure to ensure that multihomed prefixes such as  $A/a$  remain reachable

# The reachability problem in today's Internet



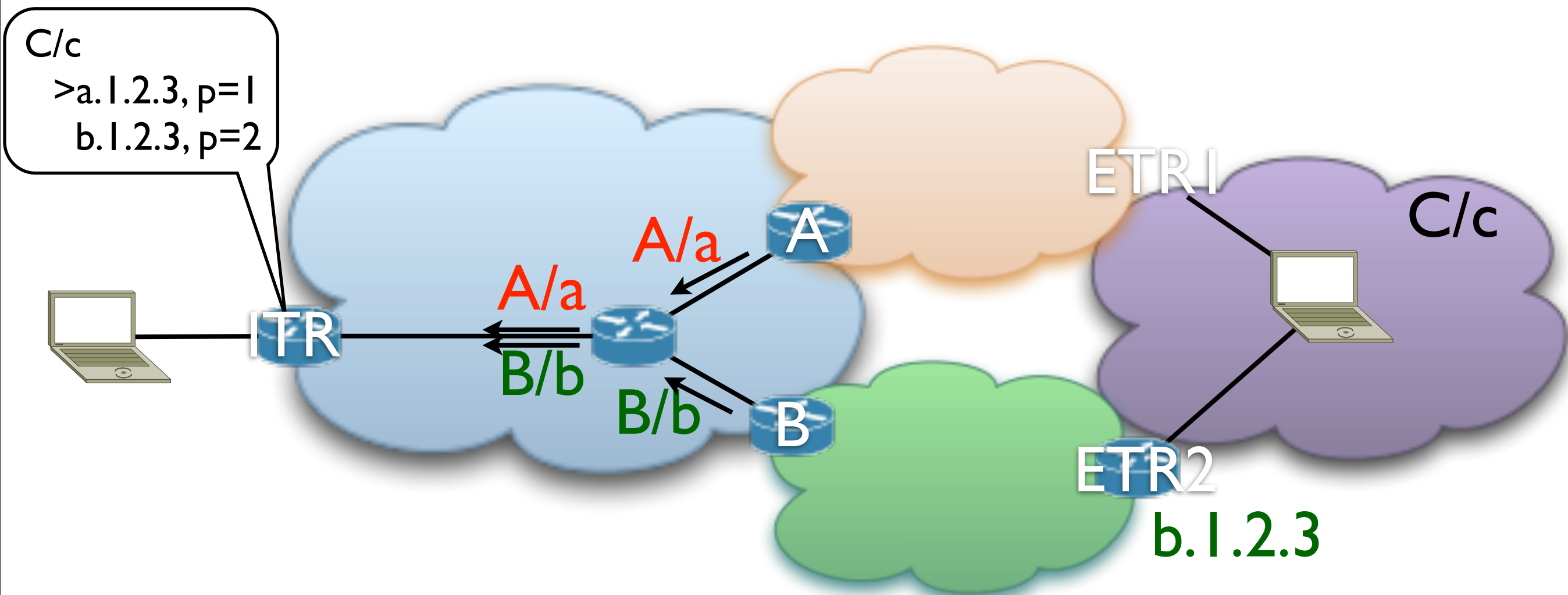
- In today's Internet, routing protocols converge after a link failure to ensure that multihomed prefixes such as  $A/a$  remain reachable

# The reachability problem in a LISP-based Internet



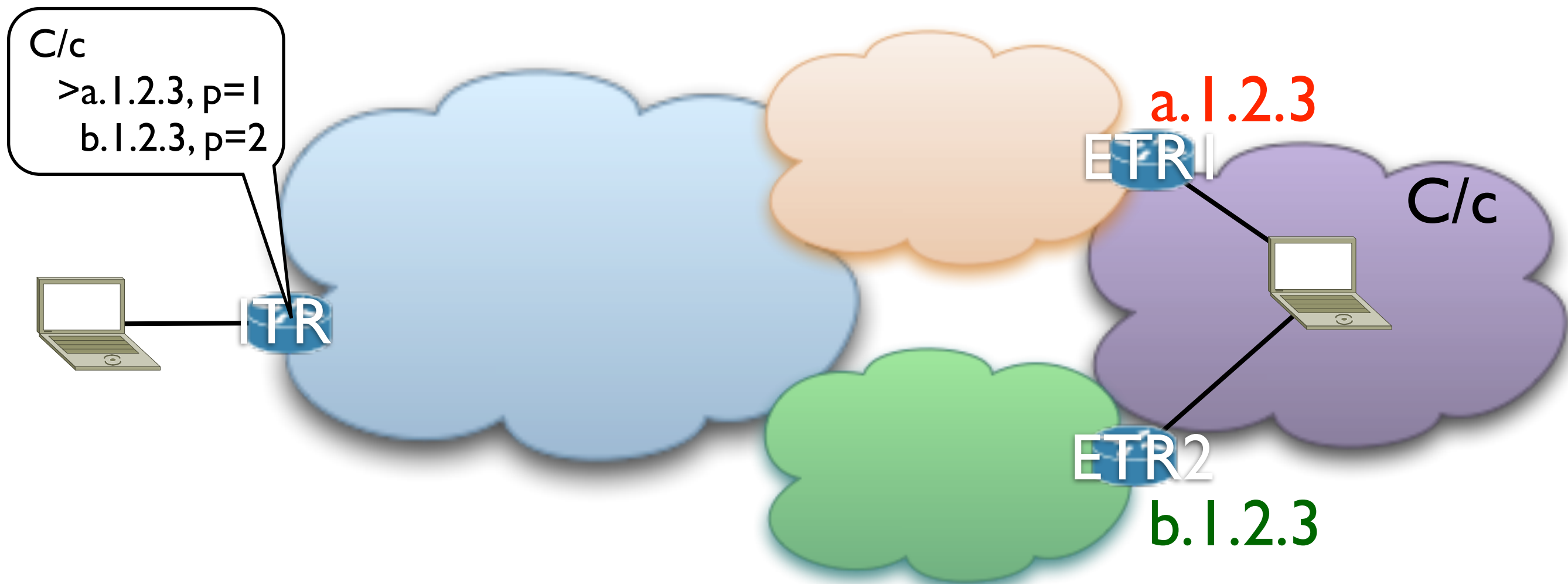
- Upon failure of ETR1, A continues to advertise A/a via BGP
- How can ITR notice that ETR1 failed and that ETR2 should be used instead?

# The reachability problem in a LISP-based Internet



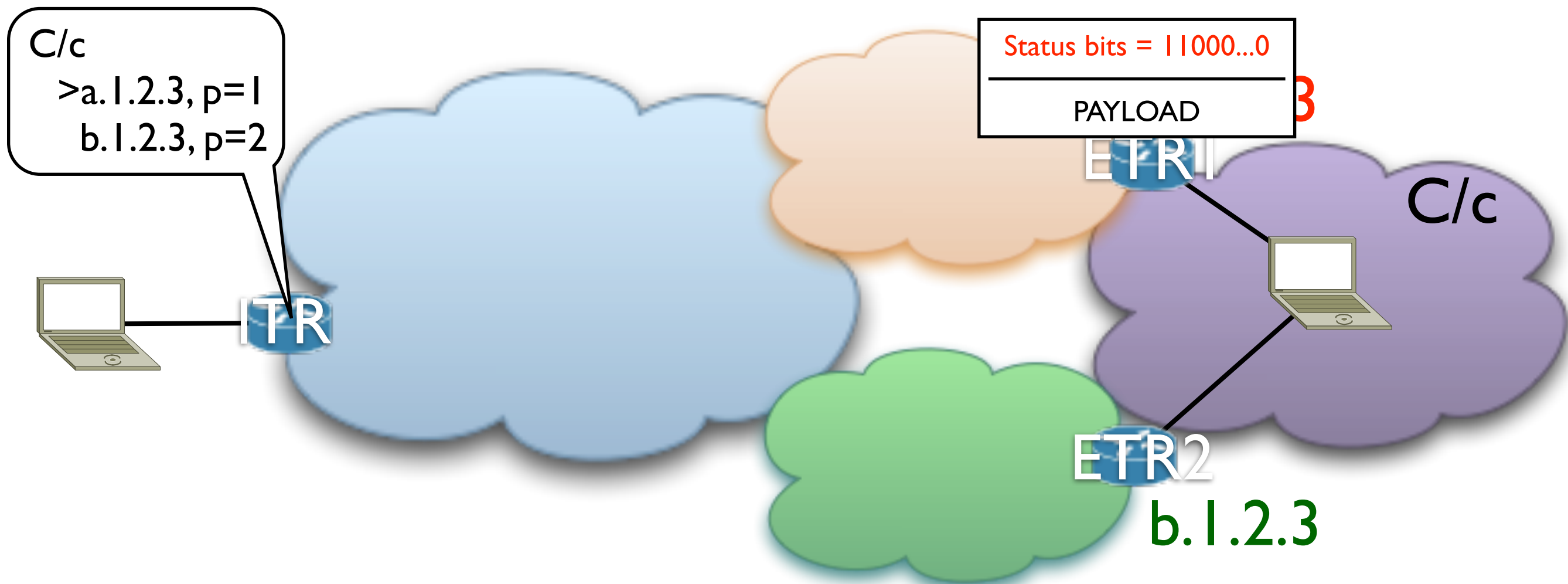
- Upon failure of ETR1, A continues to advertise A/a via BGP
- How can ITR notice that ETR1 failed and that ETR2 should be used instead?

# Solving the reachability problem with the locator status bits



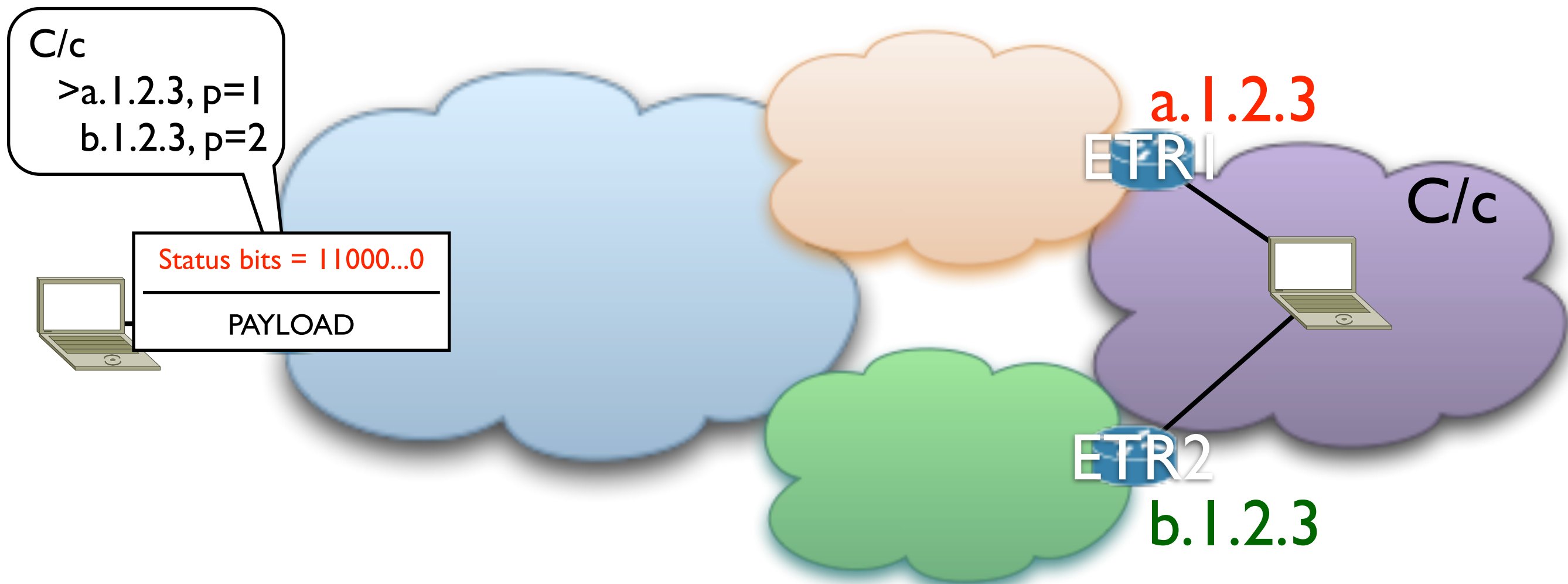
- ETR2 notices the failure and informs all ITRs to which it is sending LISP encapsulated packets by setting the reachability bit of ETR1 to 0

# Solving the reachability problem with the locator status bits



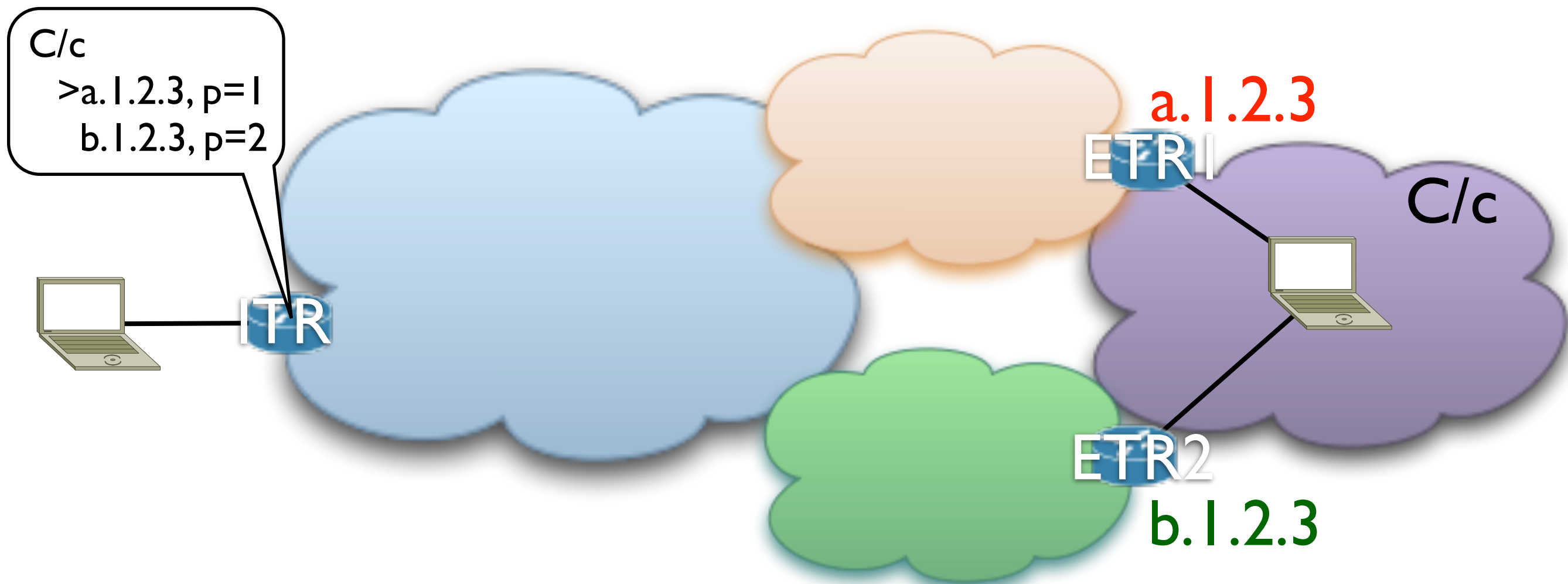
- ETR2 notices the failure and informs all ITRs to which it is sending LISP encapsulated packets by setting the reachability bit of ETR1 to 0

# Solving the reachability problem with the locator status bits



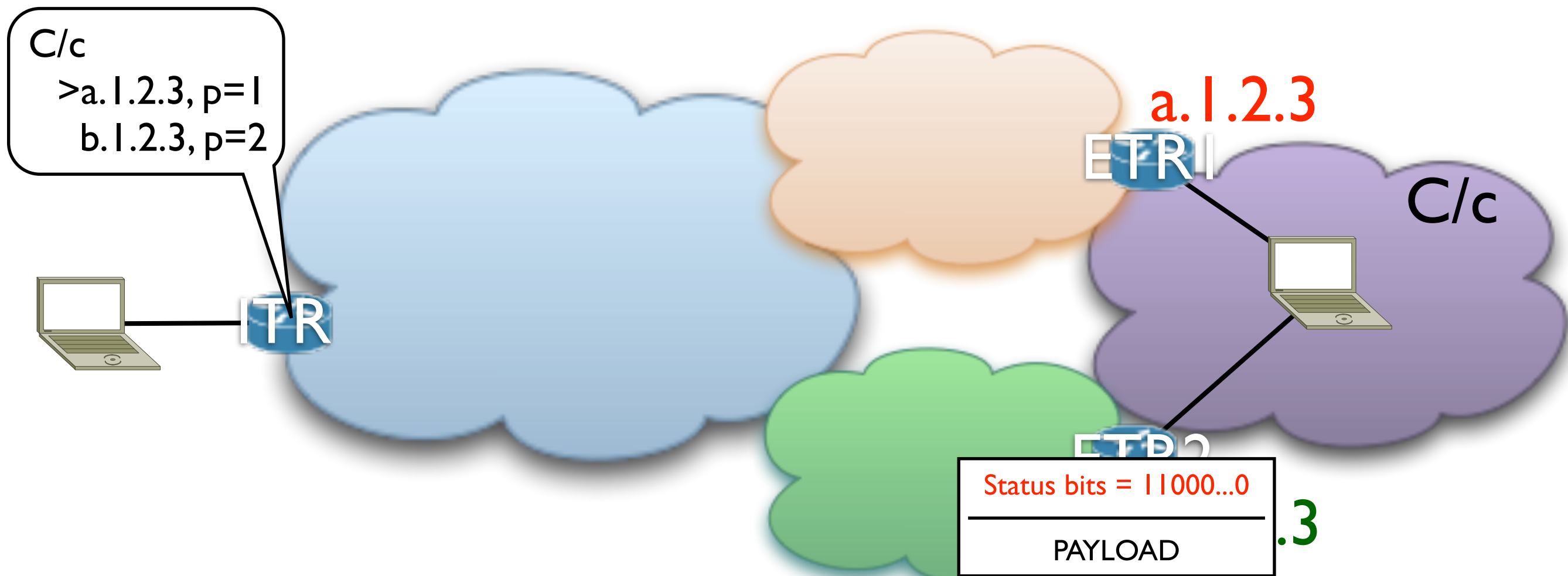
- ETR2 notices the failure and informs all ITRs to which it is sending LISP encapsulated packets by setting the reachability bit of ETR1 to 0

# Solving the reachability problem with the locator status bits



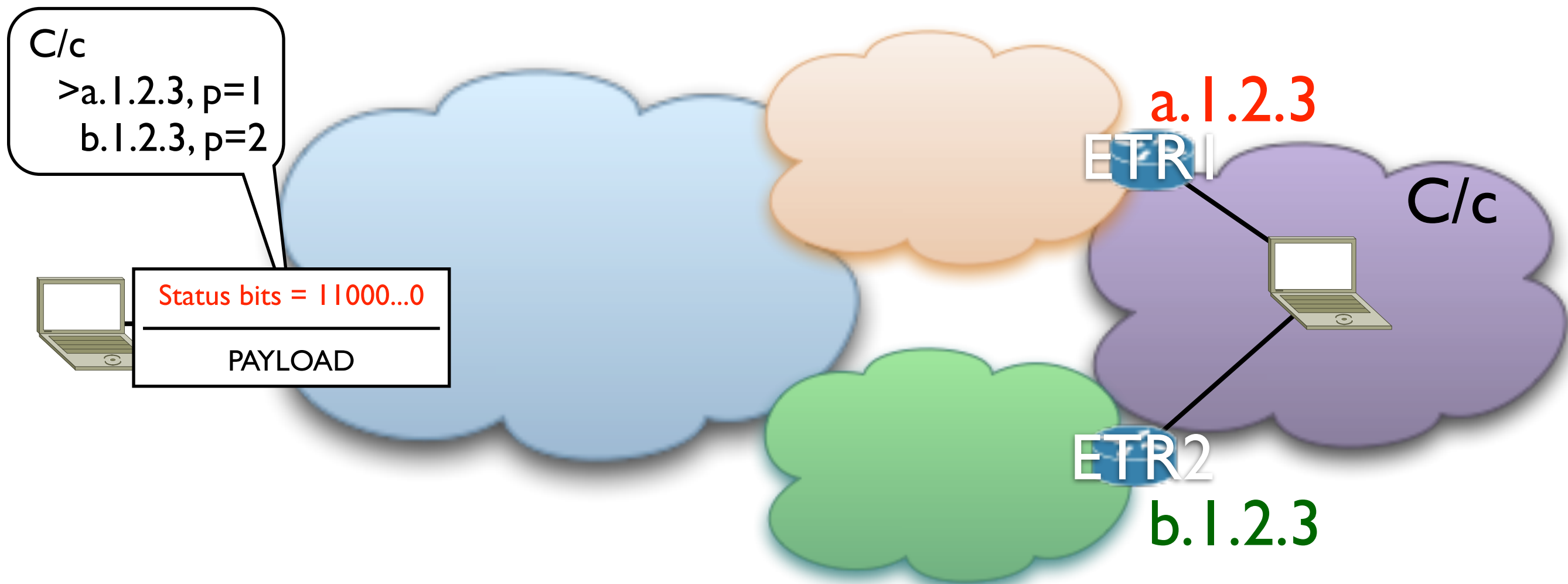
- ETR2 notices the failure and informs all ITRs to which it is sending LISP encapsulated packets by setting the reachability bit of ETR1 to 0

# Solving the reachability problem with the locator status bits



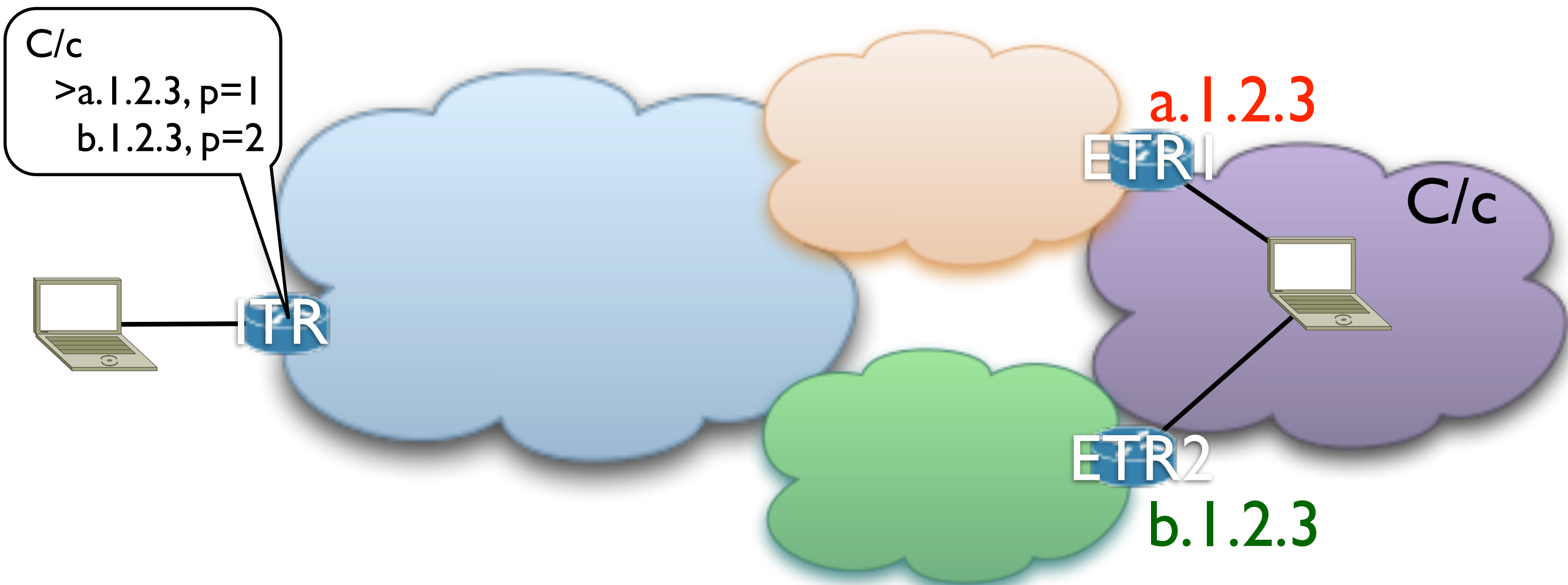
- ETR2 notices the failure and informs all ITRs to which it is sending LISP encapsulated packets by setting the reachability bit of ETR1 to 0

# Solving the reachability problem with the locator status bits



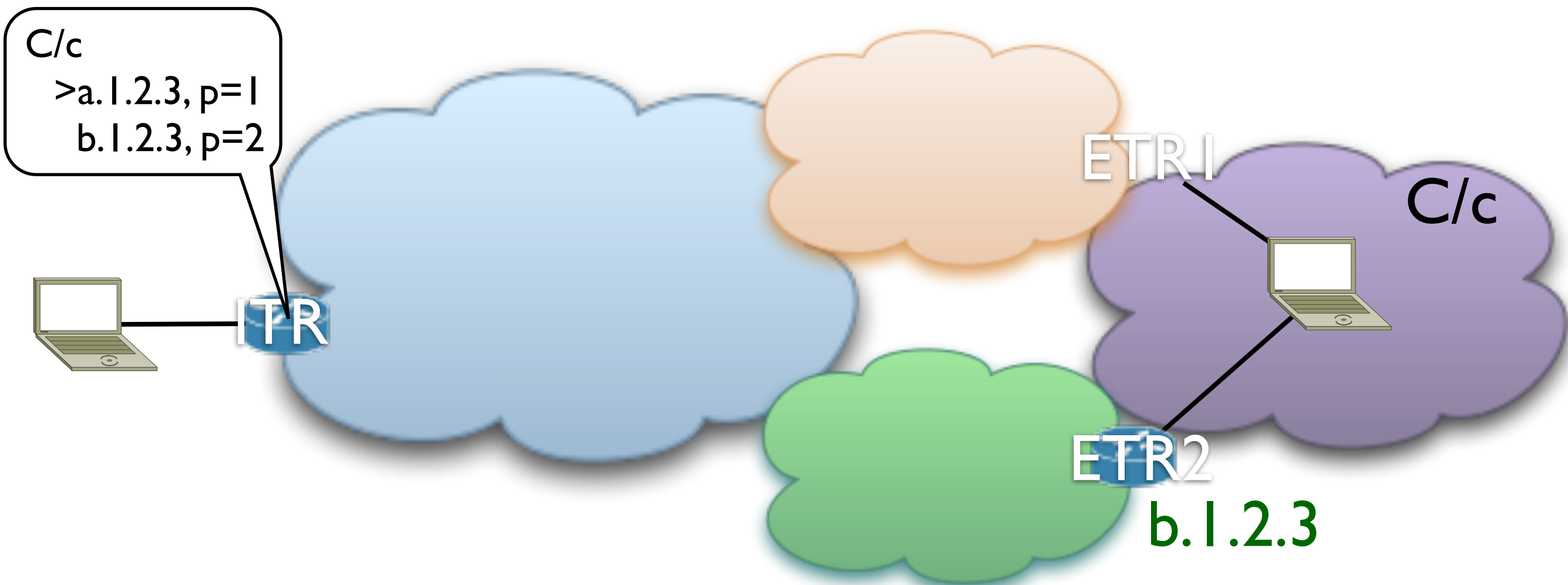
- ETR2 notices the failure and informs all ITRs to which it is sending LISP encapsulated packets by setting the reachability bit of ETR1 to 0

# Solving the reachability problem with the locator status bits



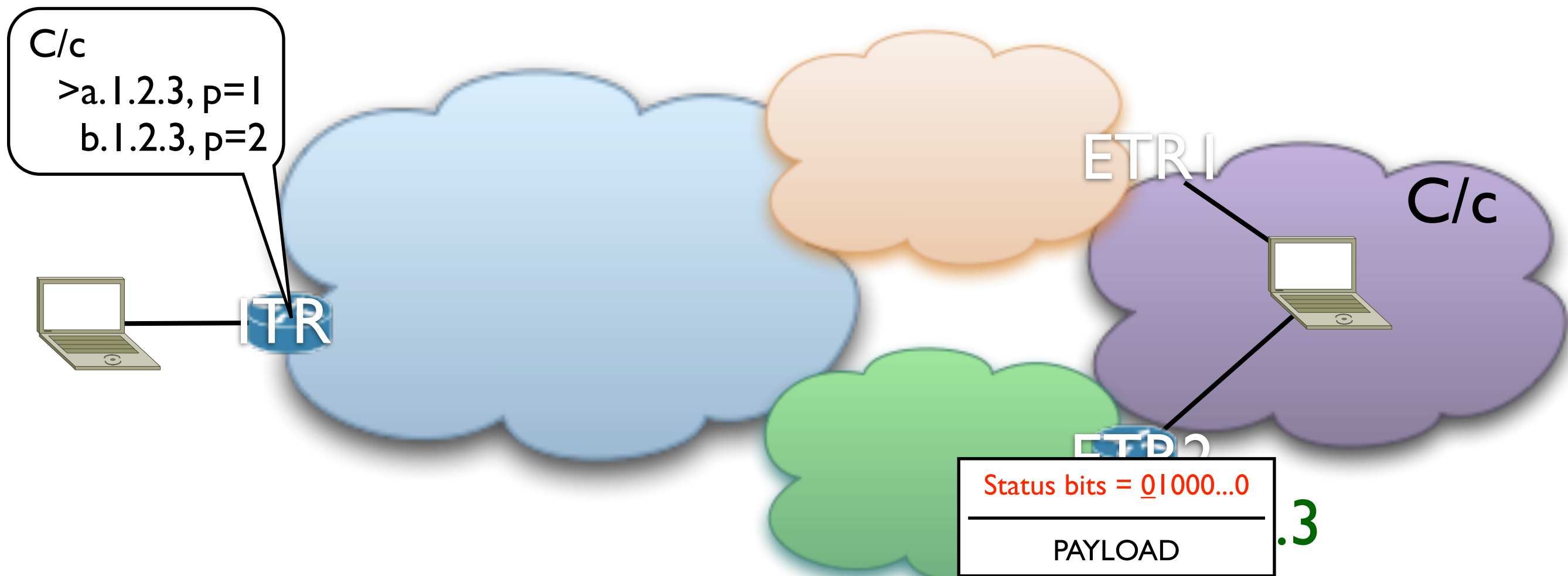
- ETR2 notices the failure and informs all ITRs to which it is sending LISP encapsulated packets by setting the reachability bit of ETR1 to 0

# Solving the reachability problem with the locator status bits



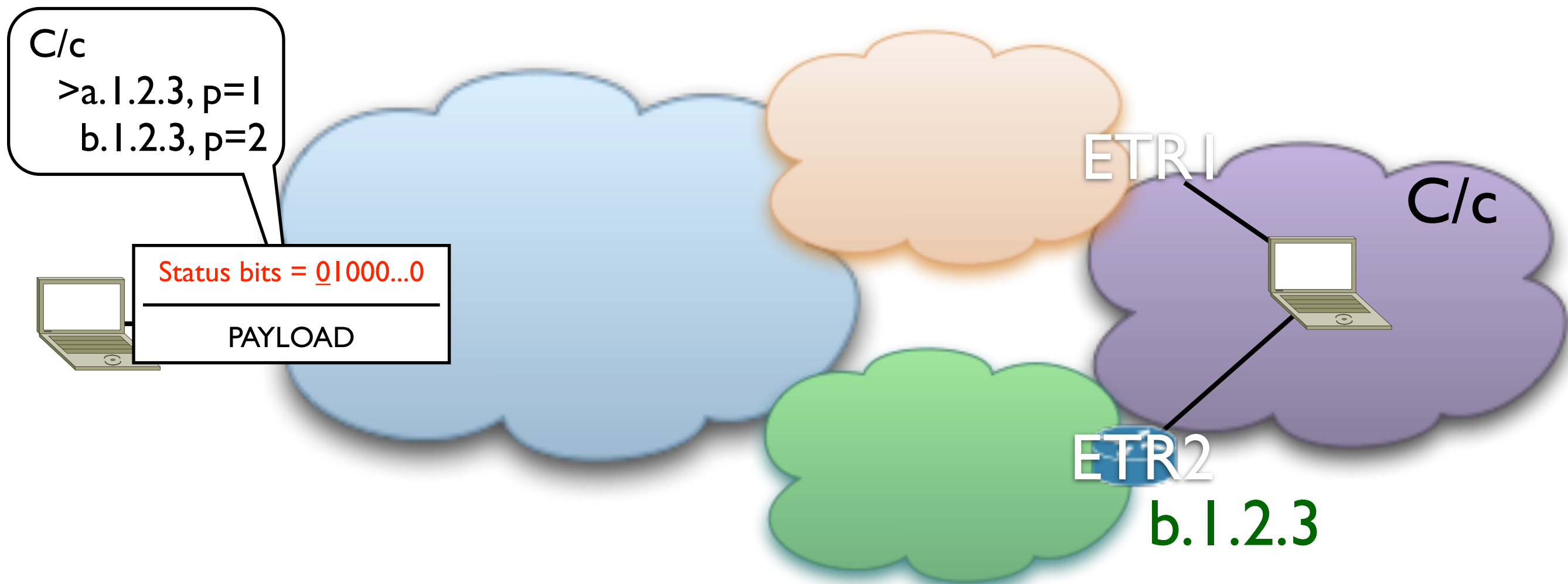
- ETR2 notices the failure and informs all ITRs to which it is sending LISP encapsulated packets by setting the reachability bit of ETR1 to 0

# Solving the reachability problem with the locator status bits



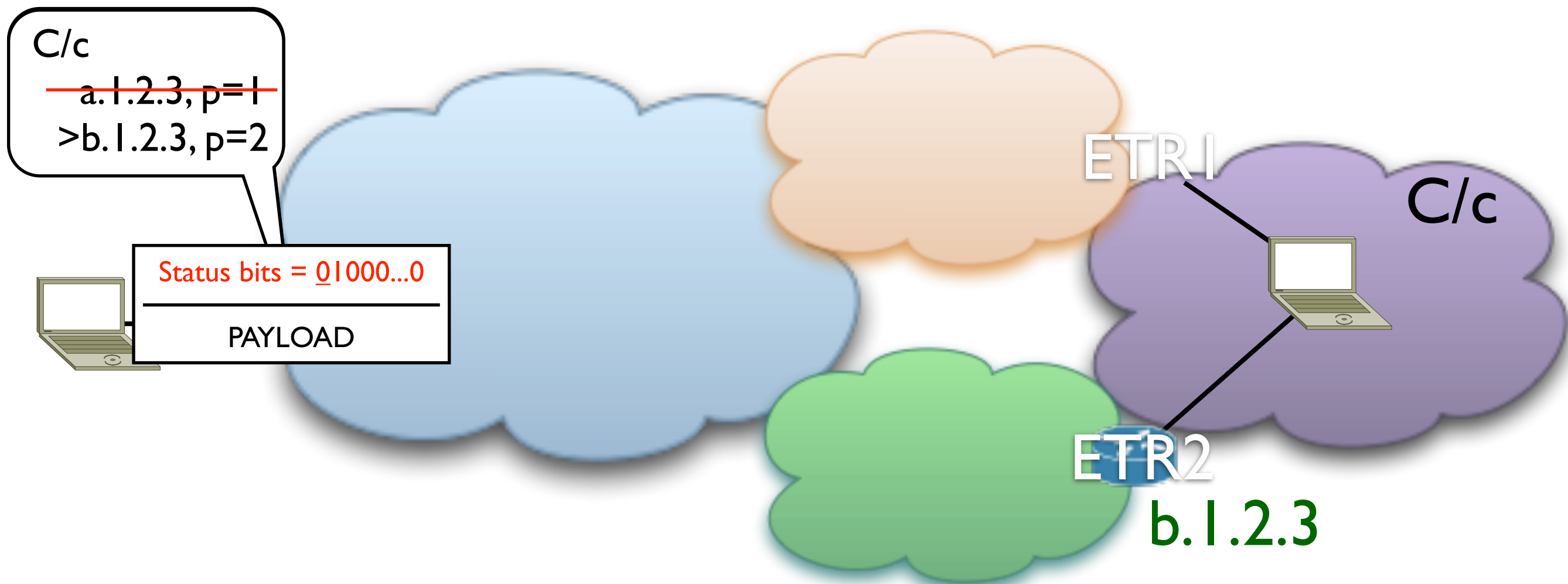
- ETR2 notices the failure and informs all ITRs to which it is sending LISP encapsulated packets by setting the reachability bit of ETR1 to 0

# Solving the reachability problem with the locator status bits



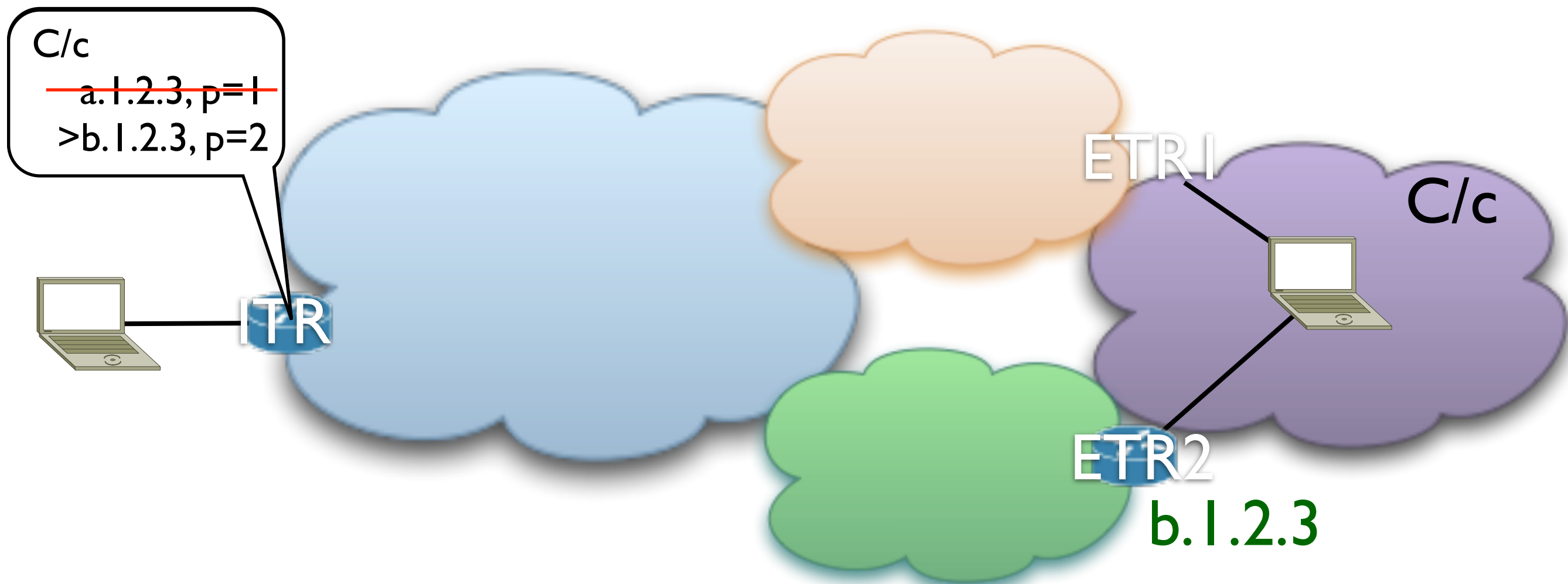
- ETR2 notices the failure and informs all ITRs to which it is sending LISP encapsulated packets by setting the reachability bit of ETR1 to 0

# Solving the reachability problem with the locator status bits



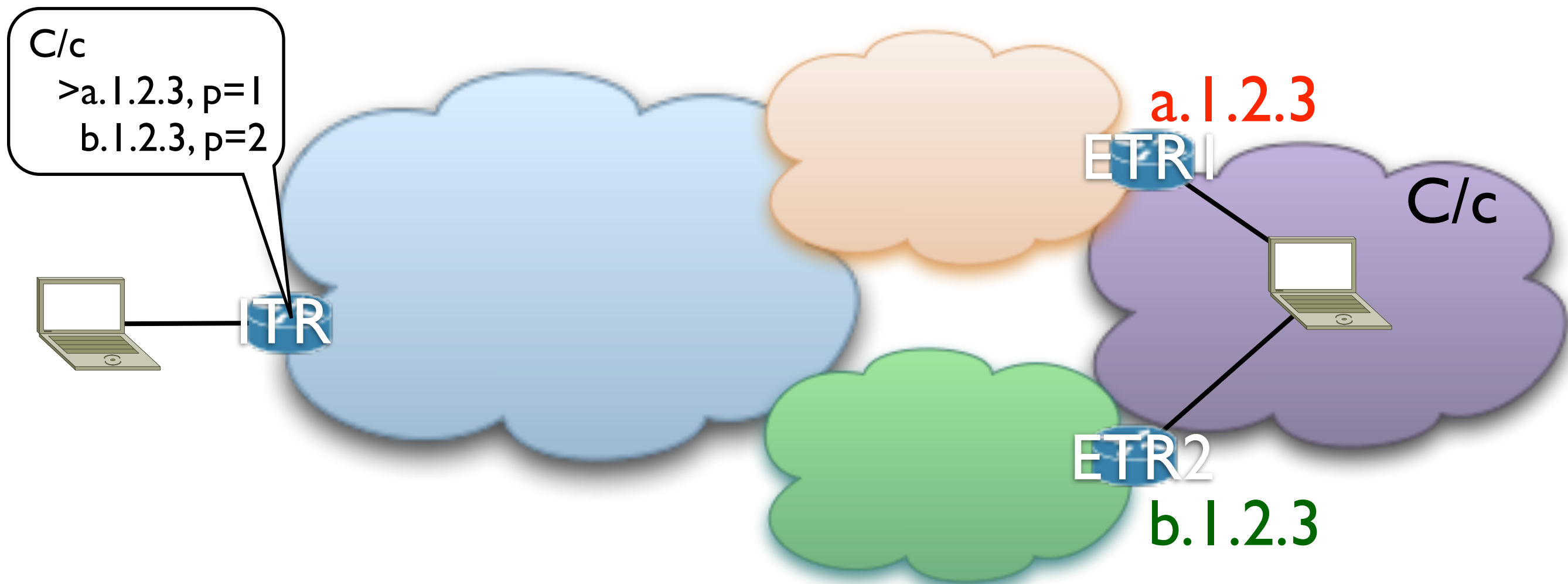
- ETR2 notices the failure and informs all ITRs to which it is sending LISP encapsulated packets by setting the reachability bit of ETR1 to 0

# Solving the reachability problem with the locator status bits



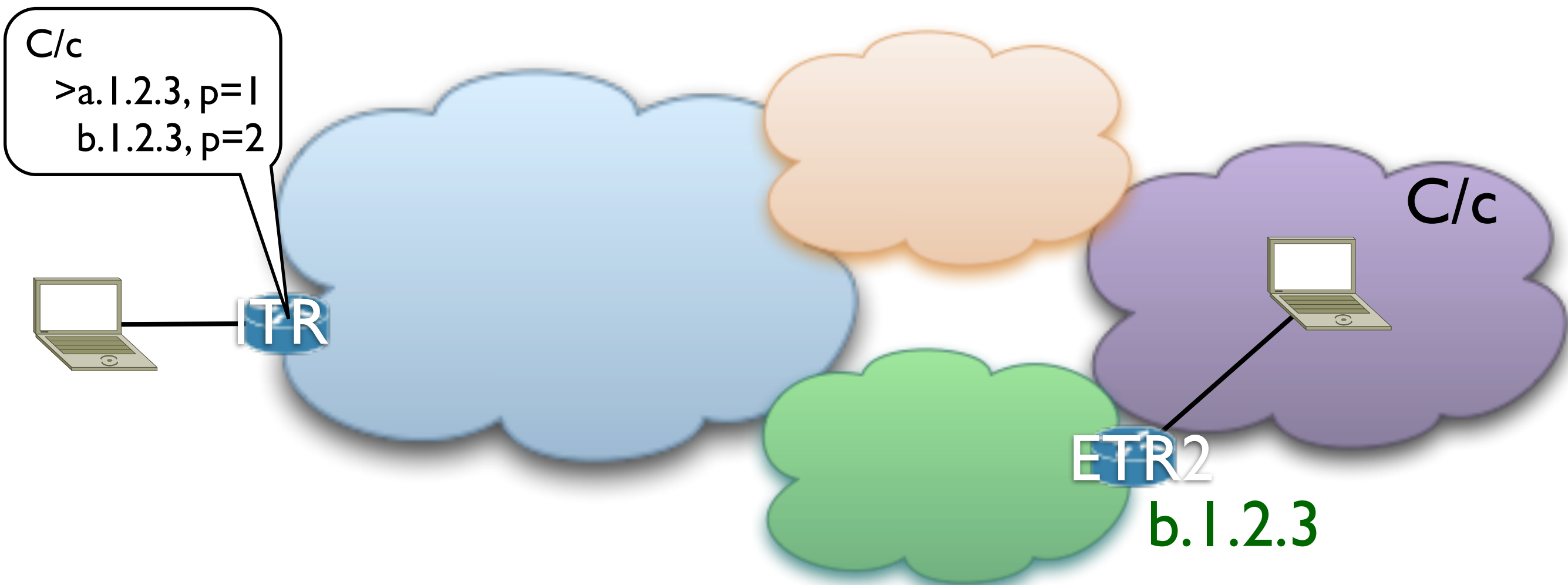
- ETR2 notices the failure and informs all ITRs to which it is sending LISP encapsulated packets by setting the reachability bit of ETR1 to 0

# Solving the reachability problem with the SMR bit



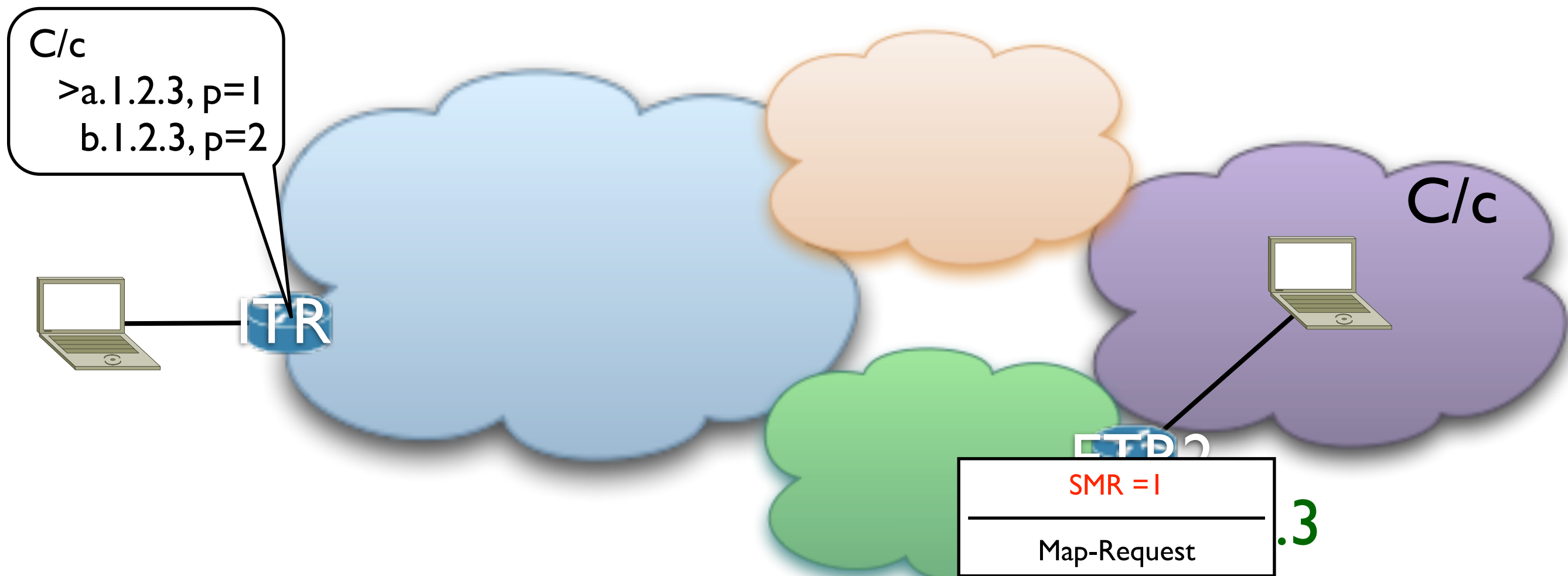
- ETR1 has been decommissioned and ETR2 wants that all the sites currently sending encapsulated data to itself update the mapping

# Solving the reachability problem with the SMR bit



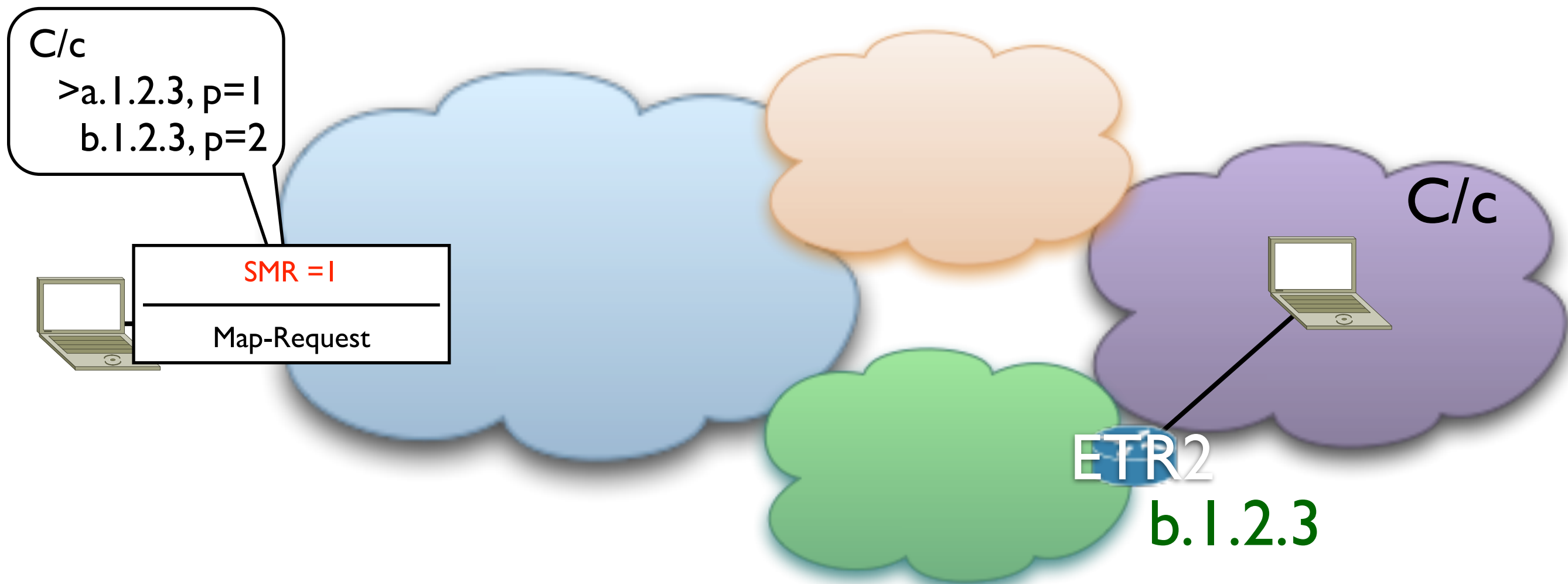
- ETR1 has been decommissioned and ETR2 wants that all the sites currently sending encapsulated data to itself update the mapping

# Solving the reachability problem with the SMR bit



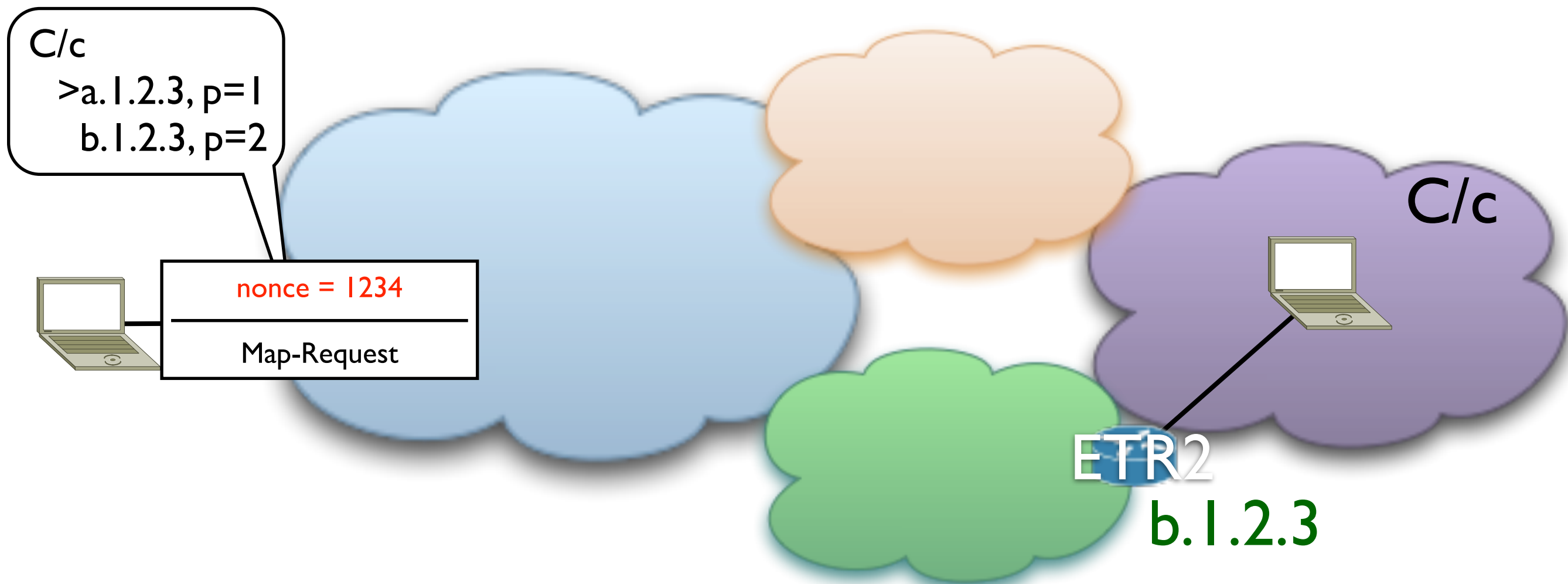
- ETR1 has been decommissioned and ETR2 wants that all the sites currently sending encapsulated data to itself update the mapping

# Solving the reachability problem with the SMR bit



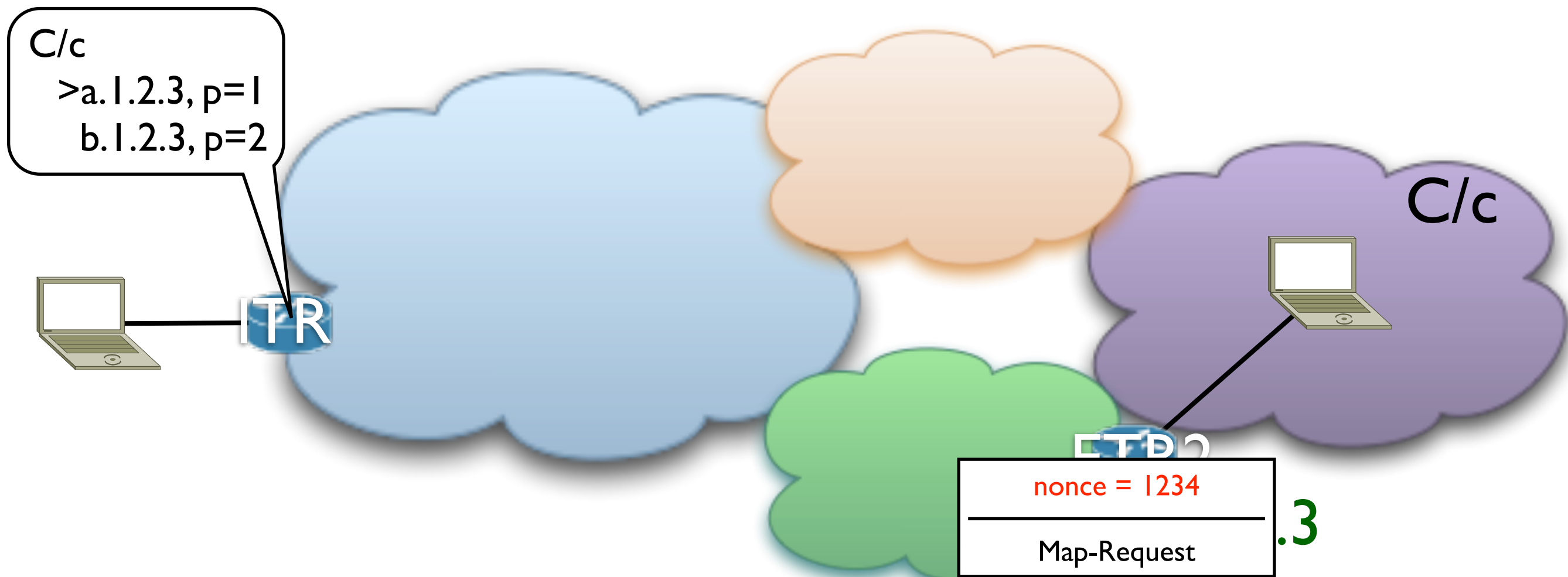
- ETR1 has been decommissioned and ETR2 wants that all the sites currently sending encapsulated data to itself update the mapping

# Solving the reachability problem with the SMR bit



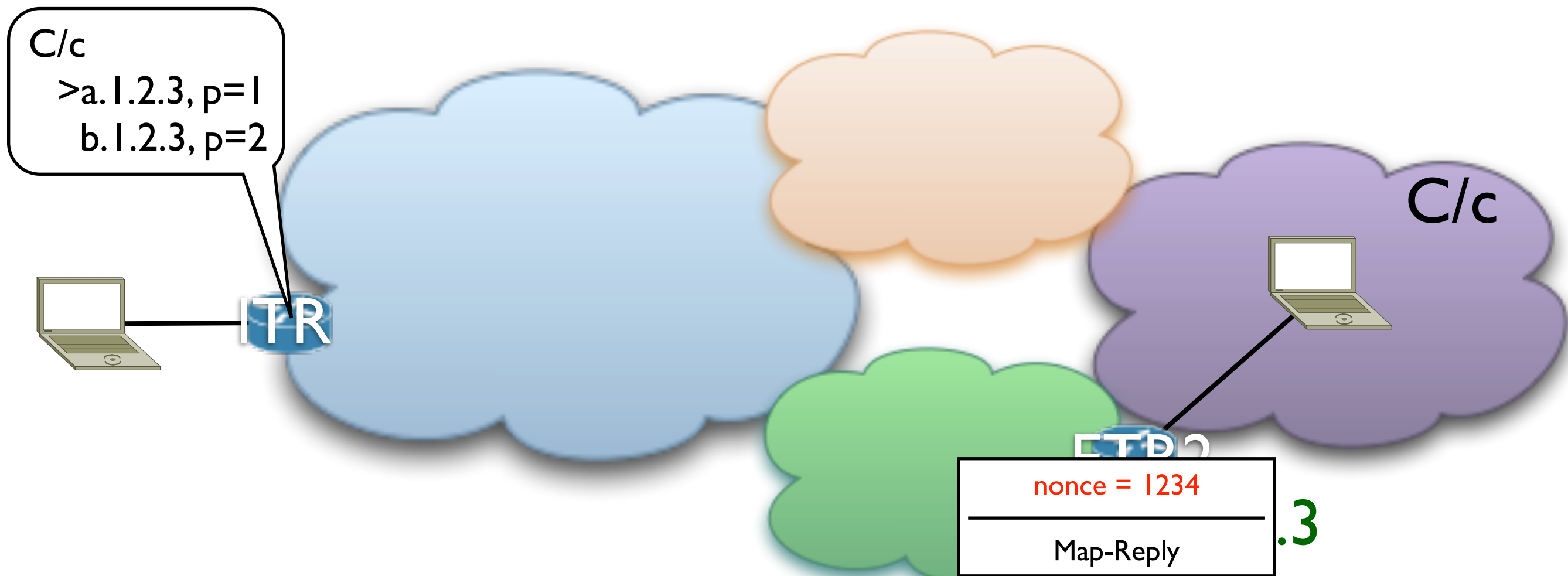
- ETR1 has been decommissioned and ETR2 wants that all the sites currently sending encapsulated data to itself update the mapping

# Solving the reachability problem with the SMR bit



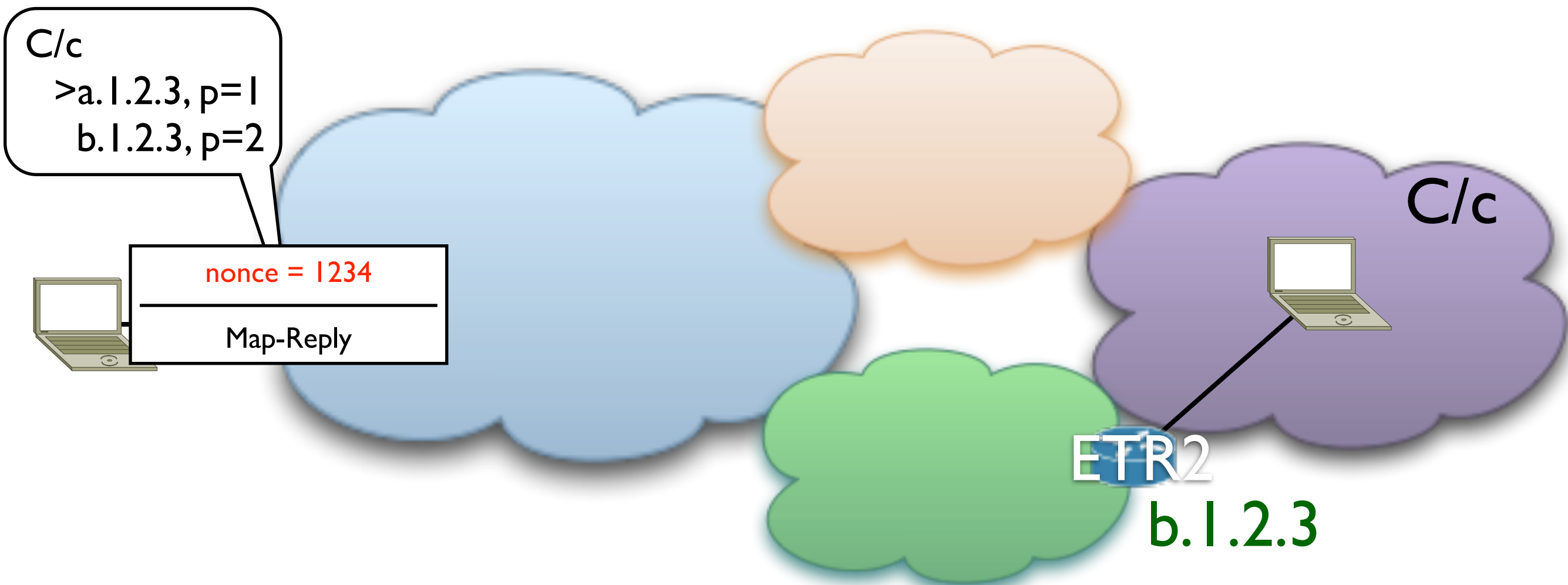
- ETR1 has been decommissioned and ETR2 wants that all the sites currently sending encapsulated data to itself update the mapping

# Solving the reachability problem with the SMR bit



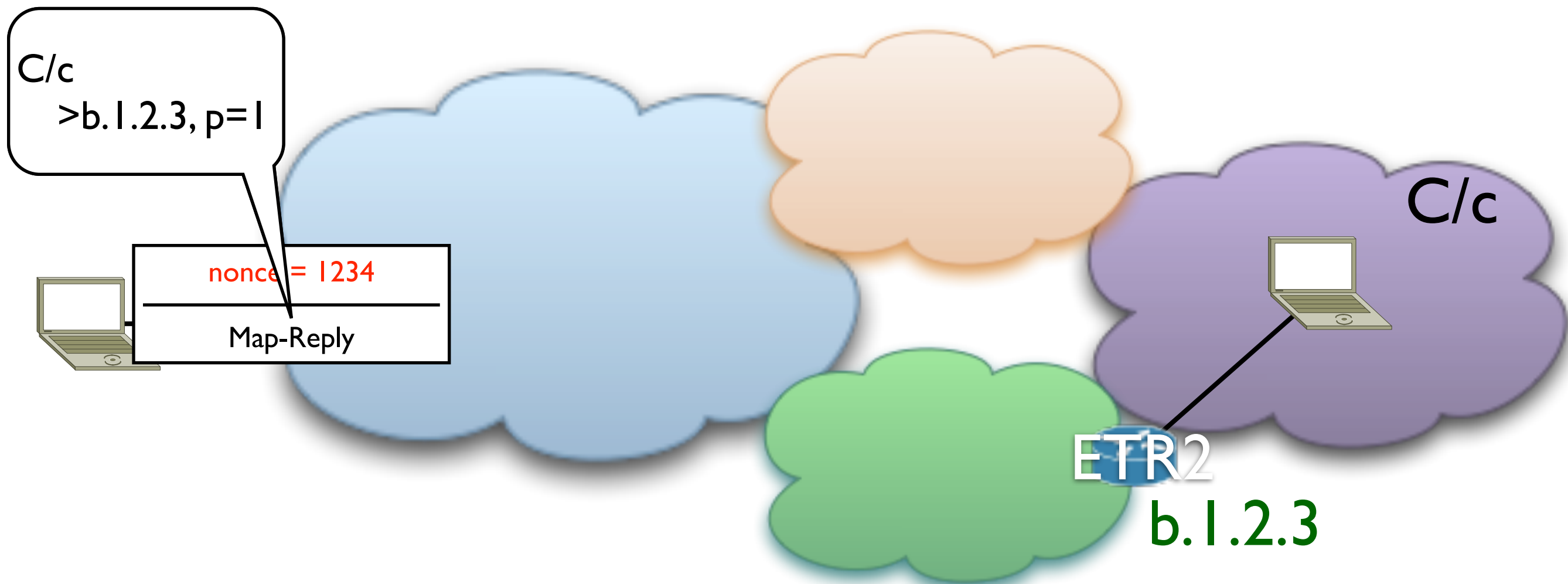
- ETR1 has been decommissioned and ETR2 wants that all the sites currently sending encapsulated data to itself update the mapping

# Solving the reachability problem with the SMR bit



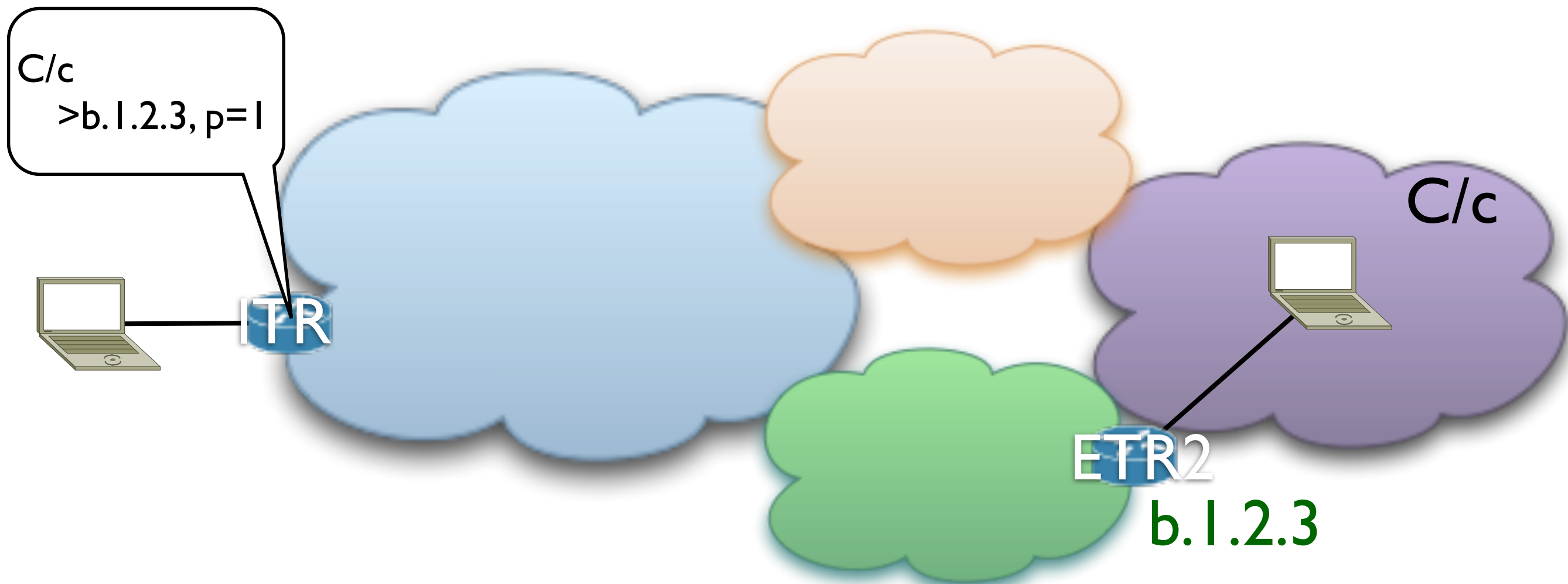
- ETR1 has been decommissioned and ETR2 wants that all the sites currently sending encapsulated data to itself update the mapping

# Solving the reachability problem with the SMR bit



- ETR1 has been decommissioned and ETR2 wants that all the sites currently sending encapsulated data to itself update the mapping

# Solving the reachability problem with the SMR bit



- ETR1 has been decommissioned and ETR2 wants that all the sites currently sending encapsulated data to itself update the mapping

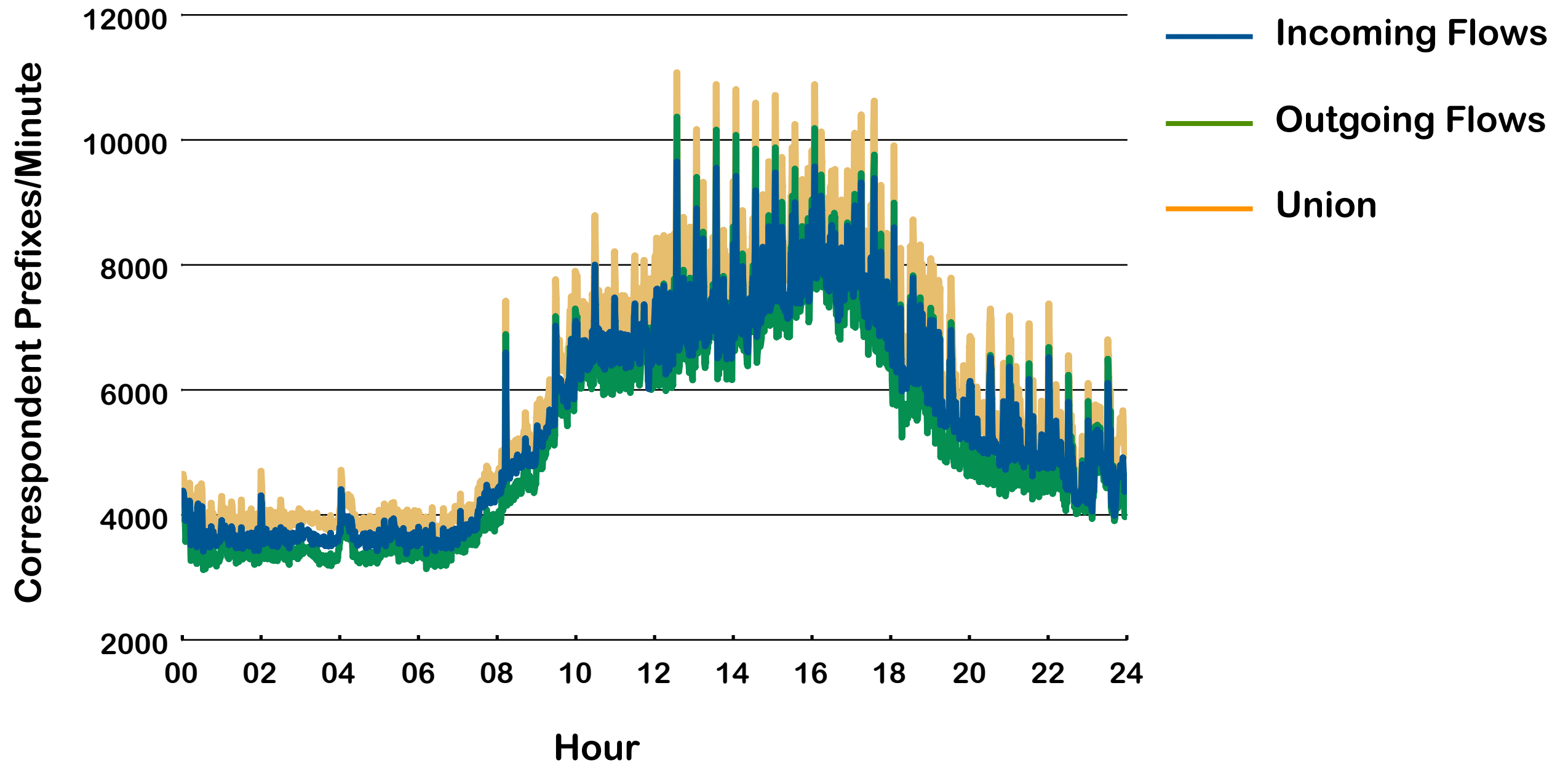
# Evaluation

# Evaluation of the cost of using LISP mappings

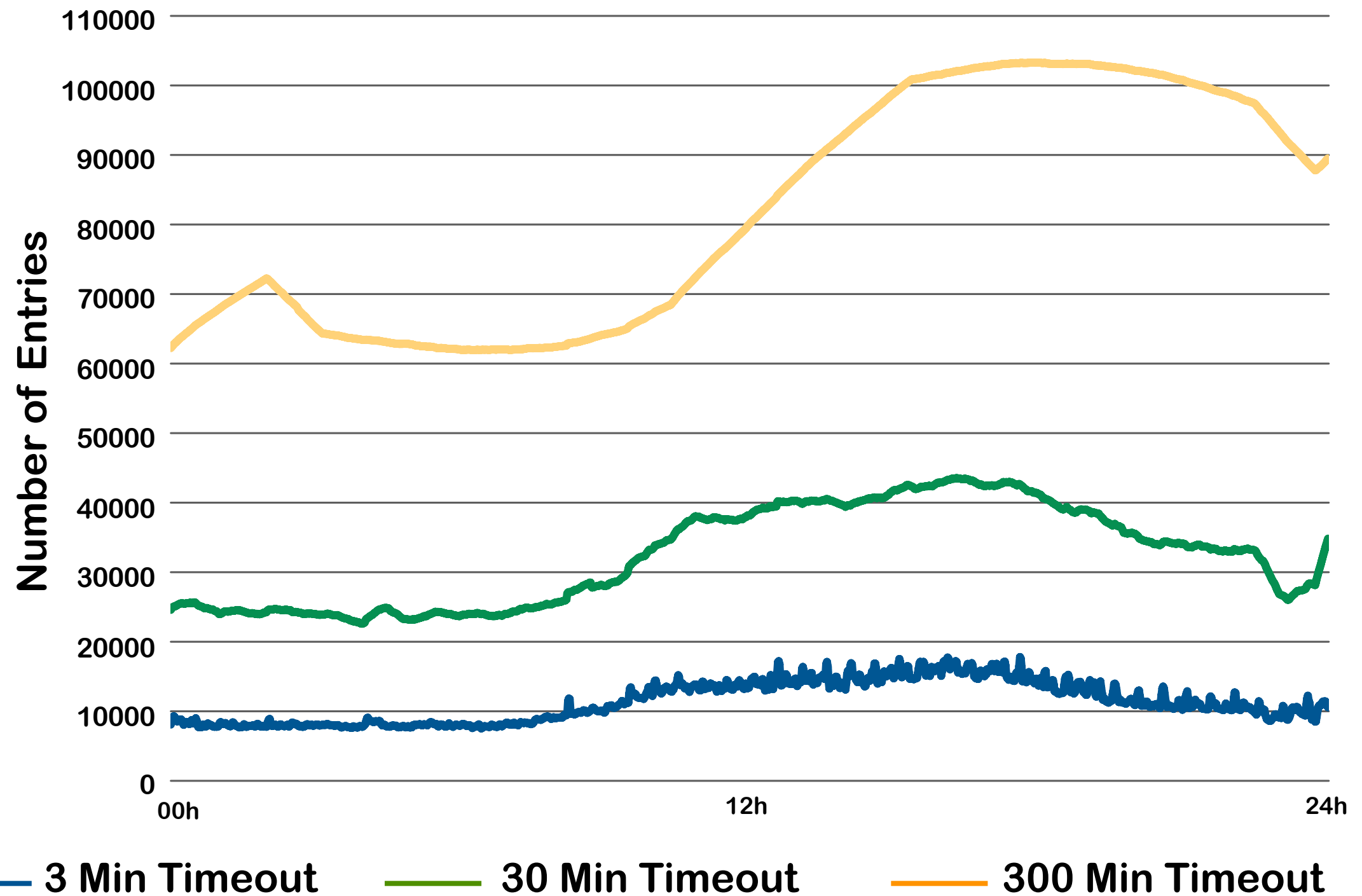
- Full Netflow (v7) on border router
- 1 Gigabit link to Belnet
- ~10000 users (/16 prefix block)
- Analysis: flow-tools + custom software
- /BGP Granularity of mappings
- iPlane data set



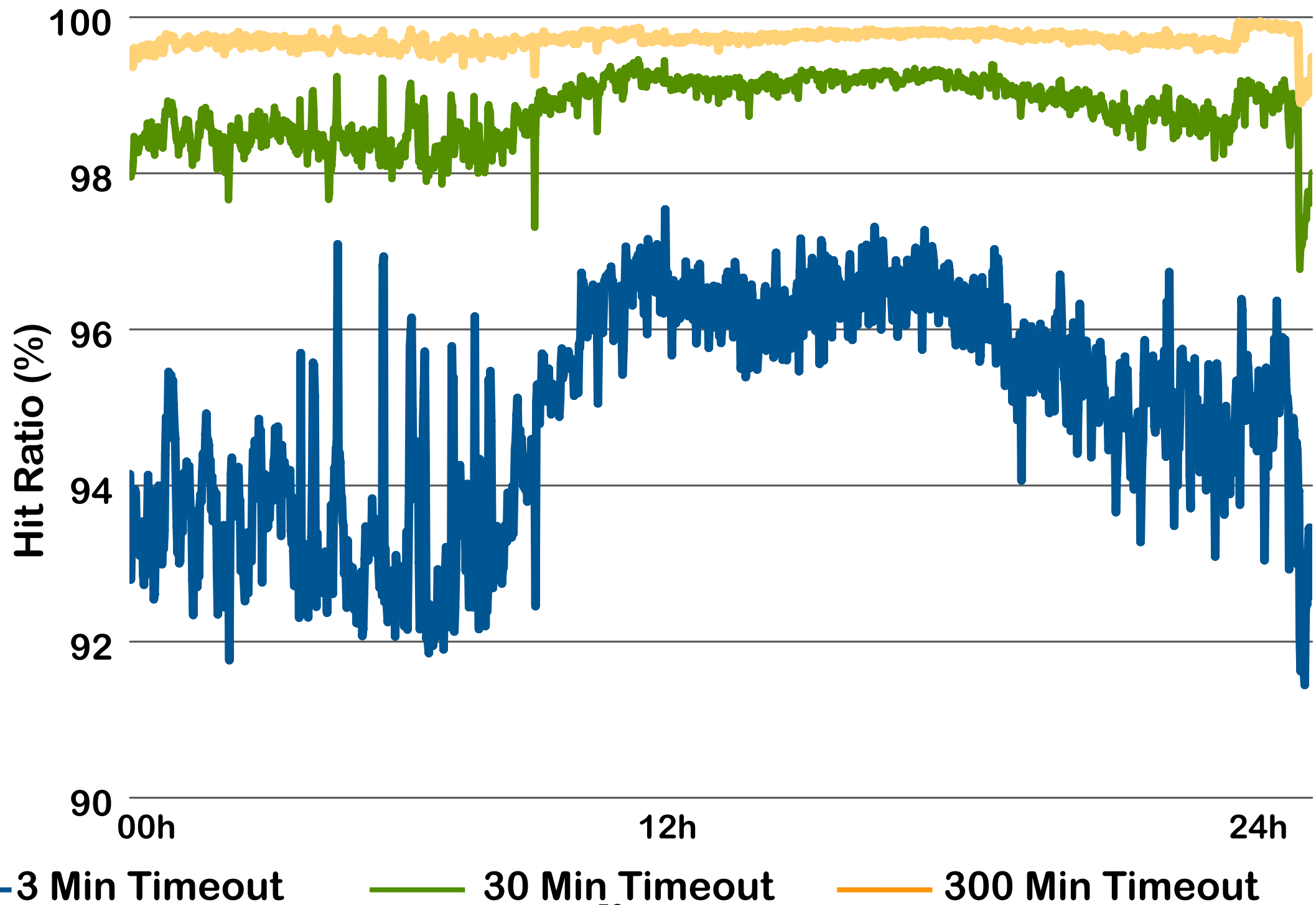
# Correspondent prefixes



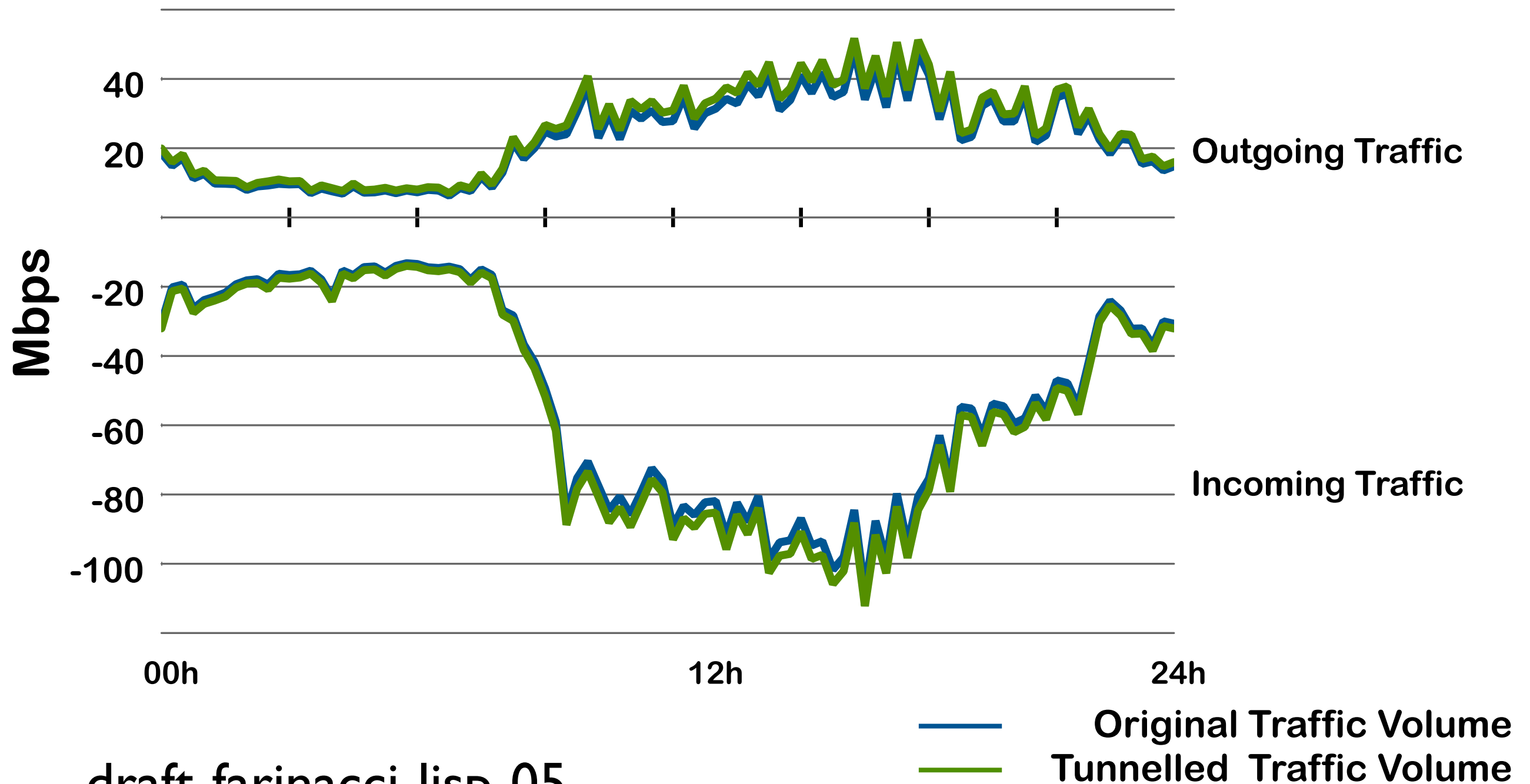
# LISP Cache size



# Hit ratio (full PULL)



# Traffic volume & tunneling overhead



# Conclusion

# Conclusion

- Network based Locator/Identifier separation proposal
  - limited impact on the router
  - limited impact on the traffic
- How to recover from failures?
- How to secure the whole system?