C-MUG
Cisco Mid-Atlantic User Group

CISCO PARTNER
Gold Certified

Introduction to LISP

(not (the (programming (language))))

Eric Stuhl

Introduction

- **Eric Stuhl – CCIE 16349**

  Routing & Switching

  Security

- **Verticals:**

  Healthcare **Enterprise**

  Financial **Academic**

  Federal **Local**

Slides provided by Cisco

Agenda

- Problem Statement
- Architectural Concepts
- Unicast and Multicast Data Plane Operation
- Mapping Database Mechanism
- Locator Reachability
- Interworking LISP Sites and Legacy Sites
- Security and Management
- Implementation and Deployment Status
- Spec References
- Q & A

Problem Statement

- **What provoked this?**

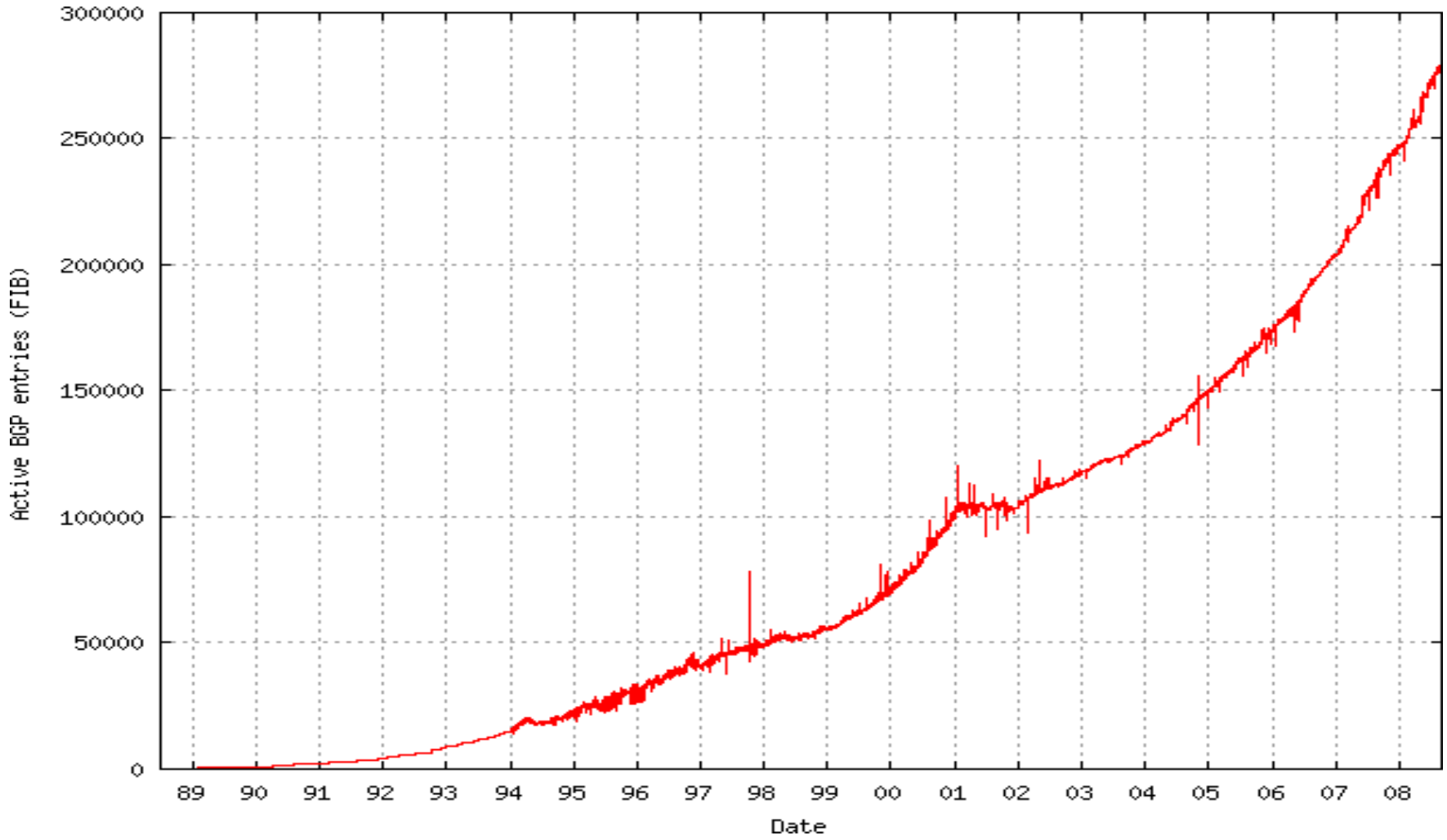
> *"… routing scalability is the most important problem facing the Internet today and must be solved …"*

Internet Architecture Board (IAB)

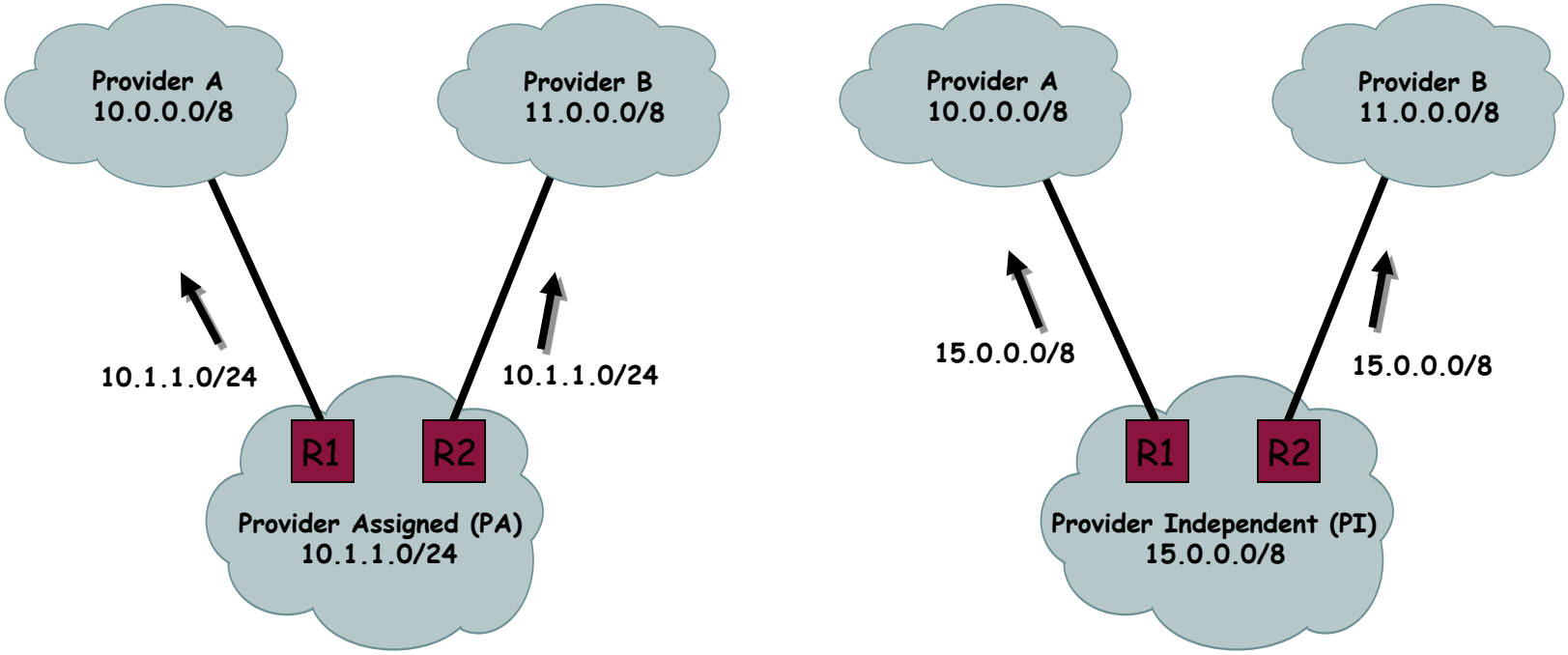
October 2006 Workshop (written as RFC 4984)

- **First and foremost - scale the Internet**

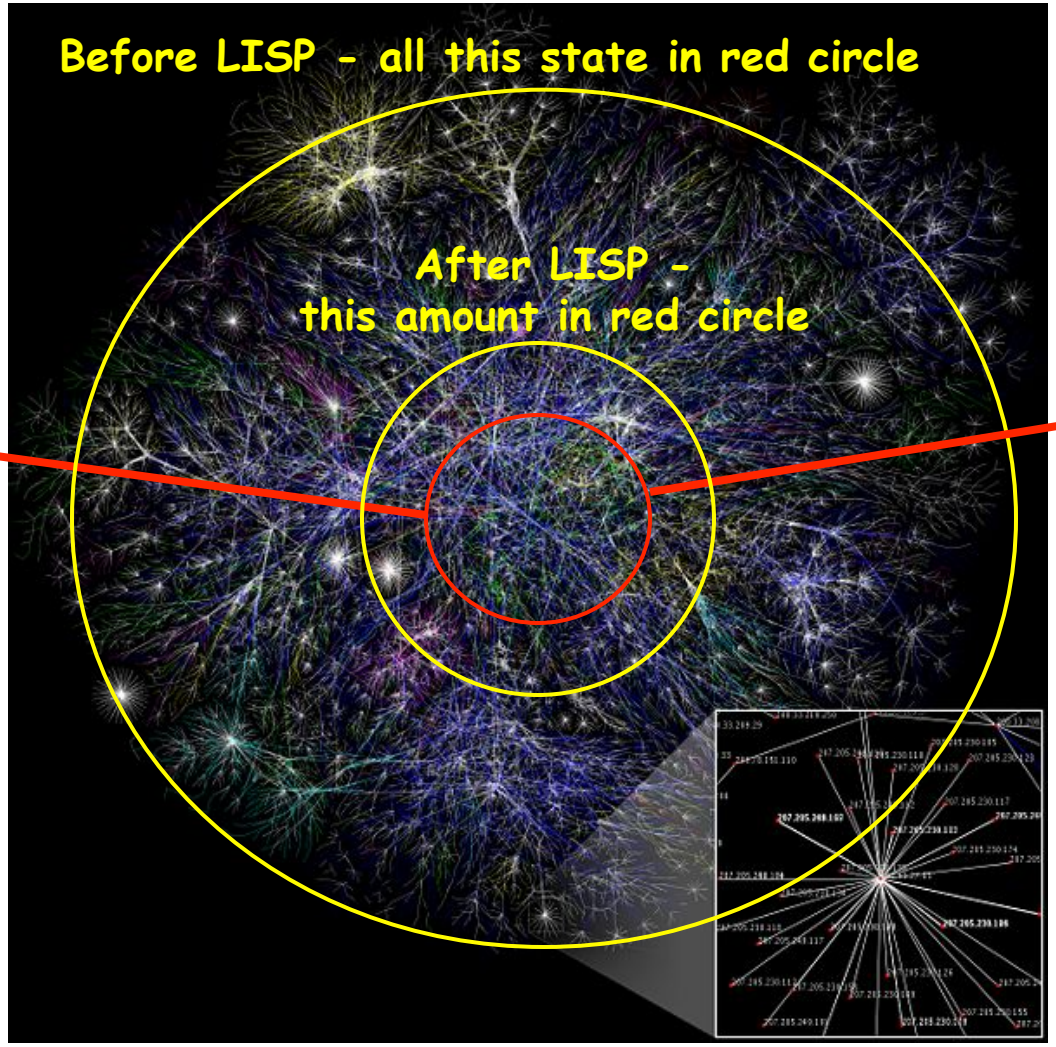
Scaling Internet Routing State



What Pollutes the Internet



Routing Table Size Problem

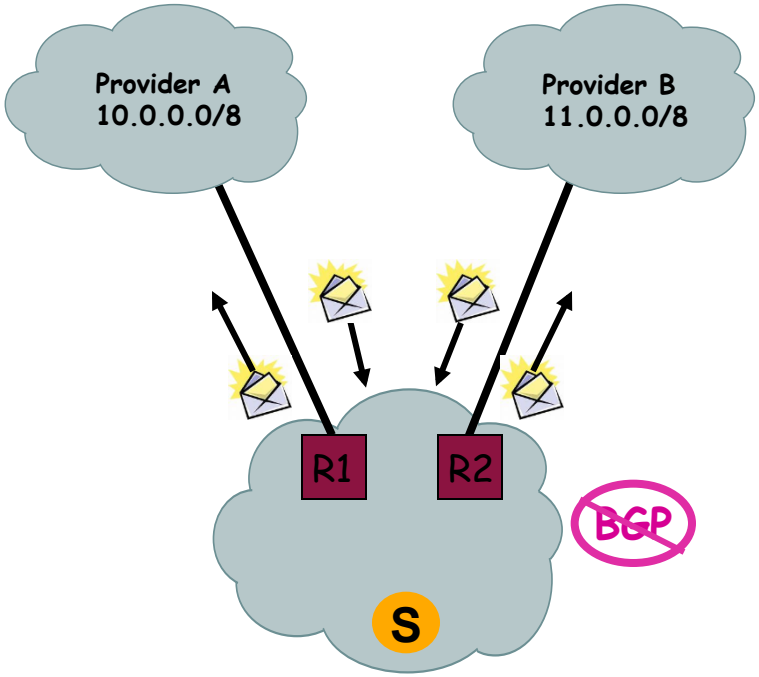


10^7 routes

10^4 routes

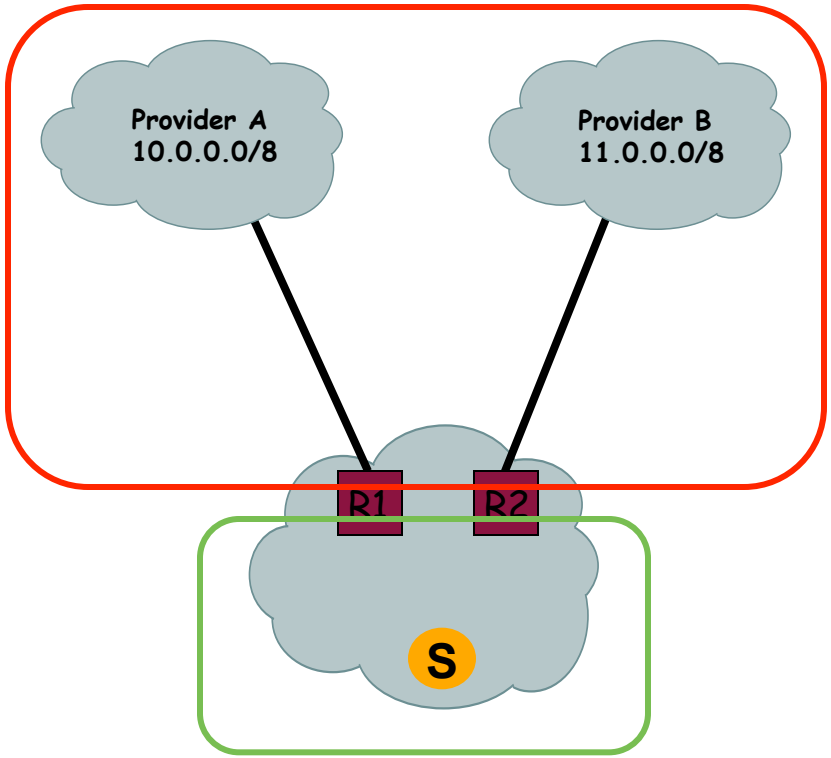
A 16-bit value!

Foster Growth in Multi-Homing



1. Improve site multi-homing
   a. Can control egress with IGP routing
   b. Hard to control ingress without more specific route injection
   c. Desire to be low OpEx multi-homed (avoid complex protocols, no outsourcing)
2. Improve ISP multi-homing
   a. Same problem for providers, can control egress but not ingress, more specific routing only tool to circumvent BGP path selection

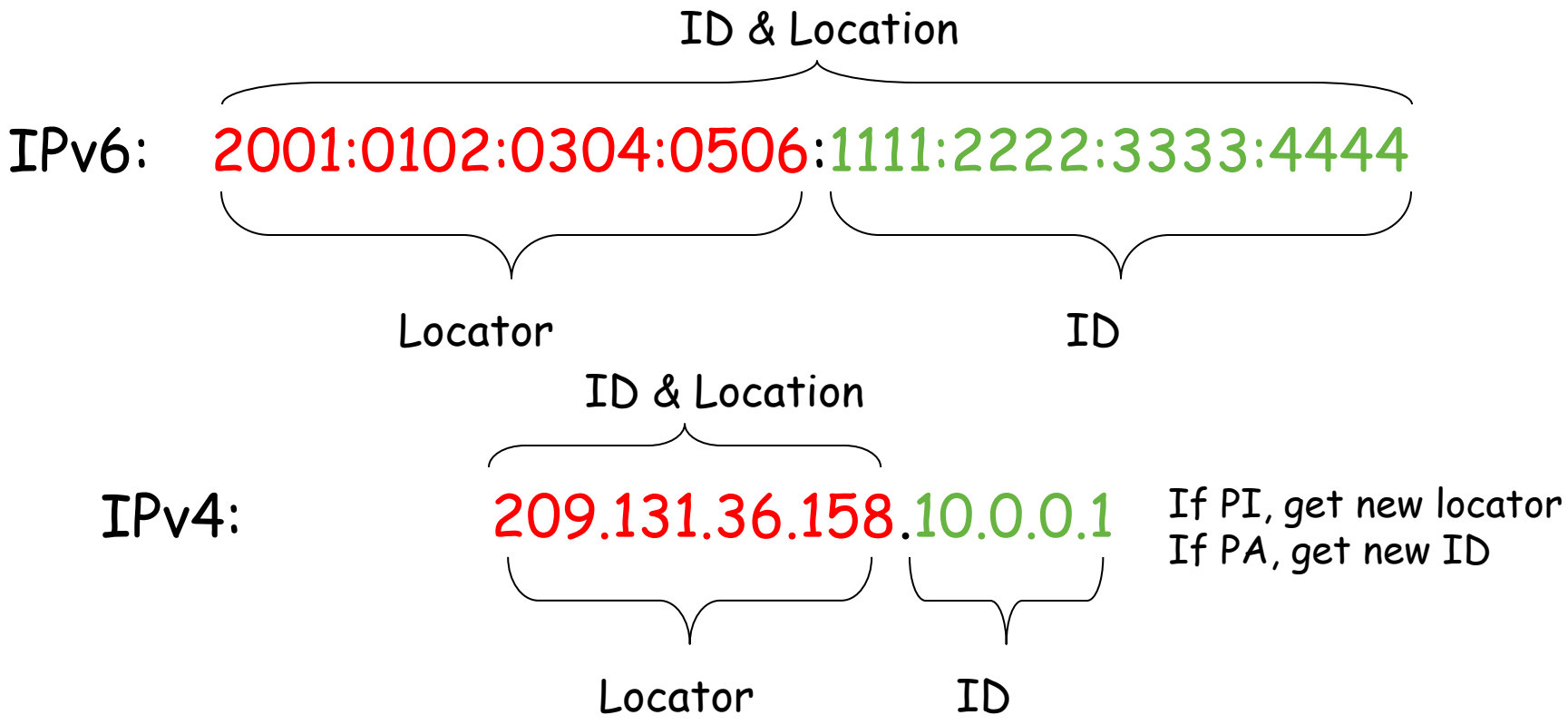
Growth in Multi-Homing



3. Decouple site addressing from provider
   a. Avoid renumbering when site changes providers
   b. Site host and router addressing decoupled from core topology
4. Add new addressing domains
   a. From possibly separate allocation entities
5. Do 1 through 4 and reduce the size of the core routing tables

Separating (or Adding) an Address

Changing the Semantics of the IP Address



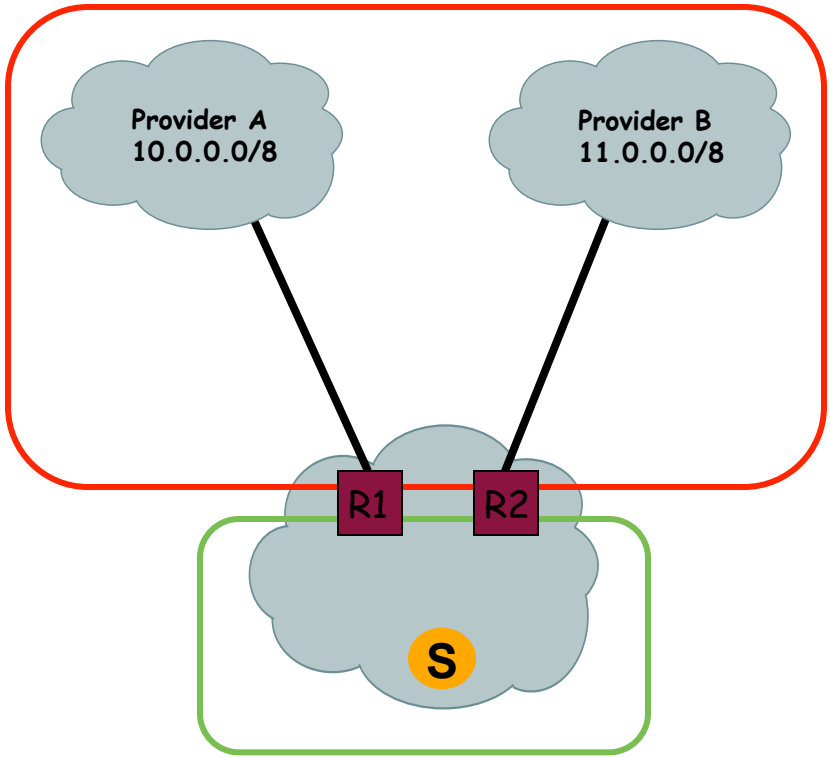
Why the Separation?

- Level of Indirection allows us to:
  - Keep either ID or Location fixed while changing the other
  - Create separate namespaces which can have different allocation properties
- By keeping IDs fixed
  - Assign fixed addresses that never change to hosts and routers at a site
- You can change Locators
  - Now the sites can change providers
  - Now the hosts can move

Some Brief Definitions

- IDs or EIDs
  - End-site addresses for hosts and routers at the site
  - They go in DNS records
  - Generally not globally routed on underlying infrastructure
  - New namespace
- RLOCs or Locators
  - Infrastructure addresses for LISP routers and ISP routers
  - Hosts do not know about them
  - They are globally routed and aggregated along the Internet connectivity topology
  - Existing namespace

Multi-Level Addressing



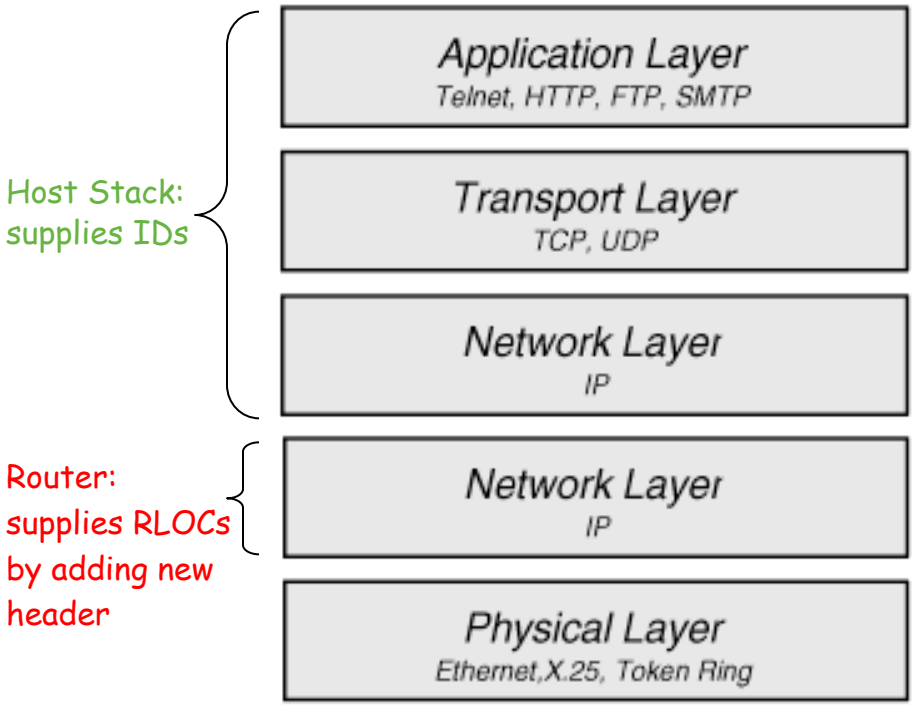
RLOCs used in the core

EIDs are inside of sites

What Is LISP?

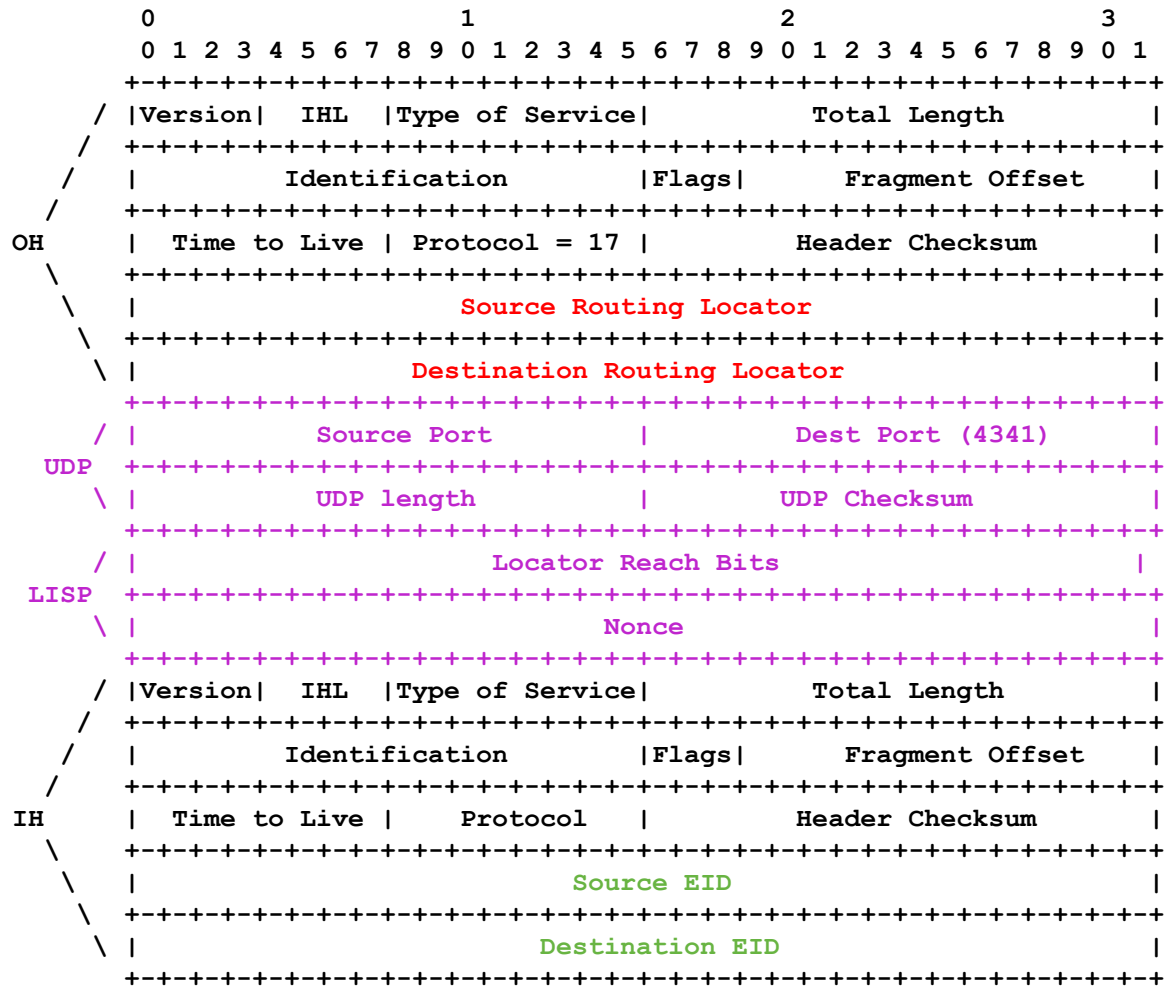
- Locator/ID Separation Protocol
- Ground rules for LISP
  - Network-based solution
  - No changes to hosts whatsoever
  - No new addressing changes to site devices
  - Very few configuration file changes
  - Imperative to be incrementally deployable
  - Support for IPv4 and IPv6 EIDs and RLOCs

What Is LISP?



"Jack-Up" or "Map-n-Encap"

draft-farinacci-lisp-12.txt



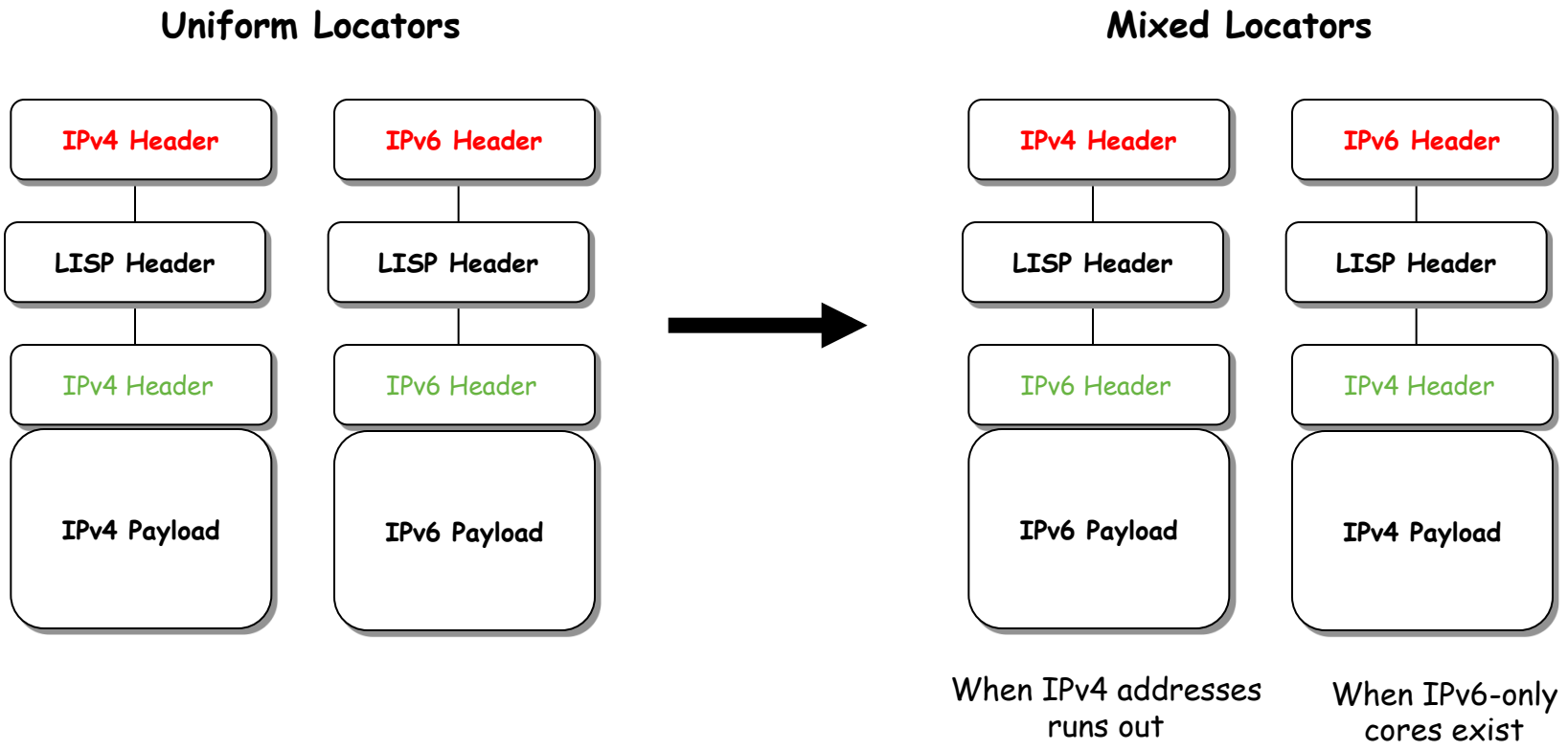
LISP and MTU

- LISP encapsulation increase the forwarded packet size
  - IPv4 – 36 bytes
  - IPv6 – 56 bytes
- Other tunneling/encapsulation protocols do the same
  - GRE, IPSec, IPinIP,
  - etc.
- Solutions for handling MTU and fragmentation issues with tunnels/ encapsulations are well documented
  - Stateful or Stateless
  - Ensure packets don't fragment
  - Allow packets to fragment
  - Drop packets

LISP and MTU

- Practical MTU on the Internet is 1500 bytes
  - Most of the core supports 4470 or 9162 bytes
  - Hosts assume “effective MTU” of 1500 bytes
- When using tunneling mechanisms, prepending headers could make packet sizes > 1500 bytes
  - Larger packets are better for efficiency purposes
- Network layer fragmentation is not performance-efficient
  - Decapsulating tunnel routers need reassembly buffers
  - Packet loss causes long buffer holding periods

LISP for IPv6 Transition



Legend: EIDs -> Green, Locators -> Red

What is the LISP Data-Plane?

- Design for encapsulation and router placement
- Design for locator reachability
- Data-triggered mapping service
  - Map-Request messages
  - Map-Reply messages
  - Map-Register messages

Network Elements

New Network Elements

- Ingress Tunnel Router (ITR)

  Finds EID to RLOC mapping

- This is the map part of map-and-encap

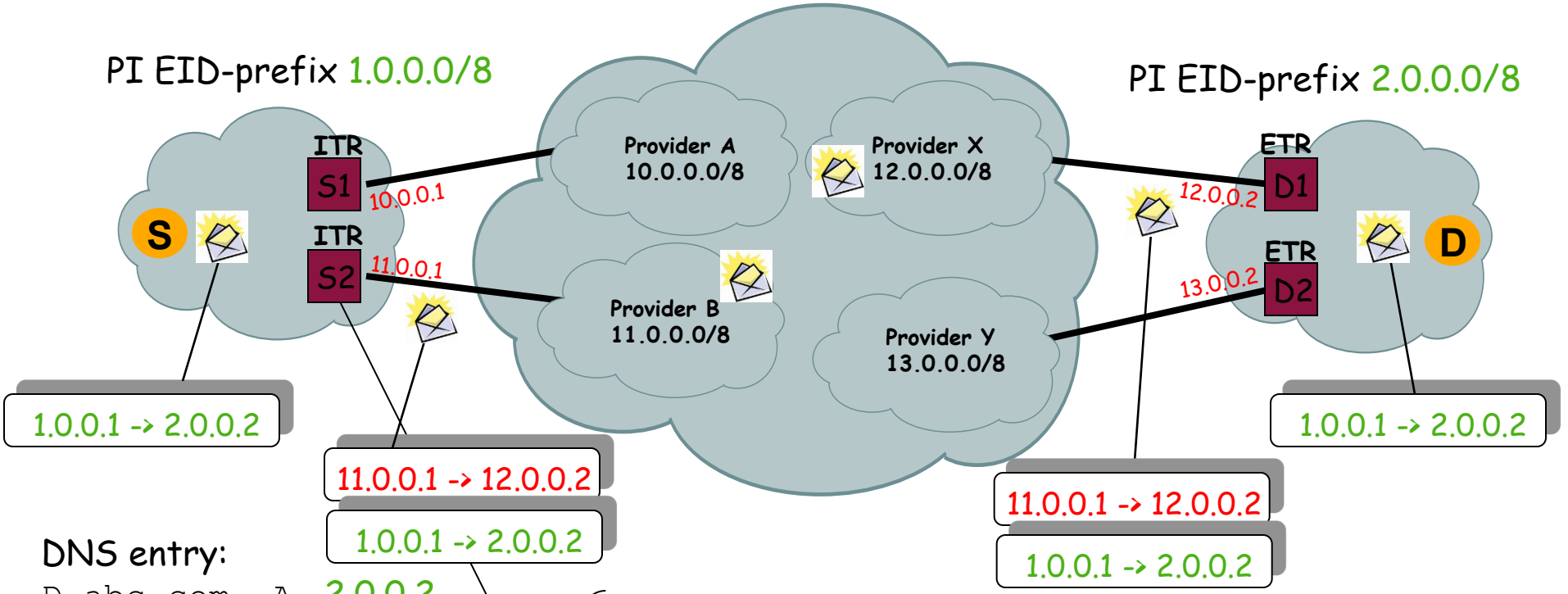
  Encapsulates to Locators at source site

- This is the encap part of map-and-encap
- Egress Tunnel Router (ETR)

  Authoritative for its EID to RLOC mapping

  Decapsulates at destination site

Unicast Packet Forwarding



DNS entry:

`D.abc.com  A  2.0.0.2`

Mapping Entry

EID-prefix: 2.0.0.0/8

Locator-set:

12.0.0.2, priority: 1, weight: 50 (D1)

13.0.0.2, priority: 1, weight: 50 (D2)

Policy controlled by destination site

Legend:

EIDs -> Green

Locators -> Red

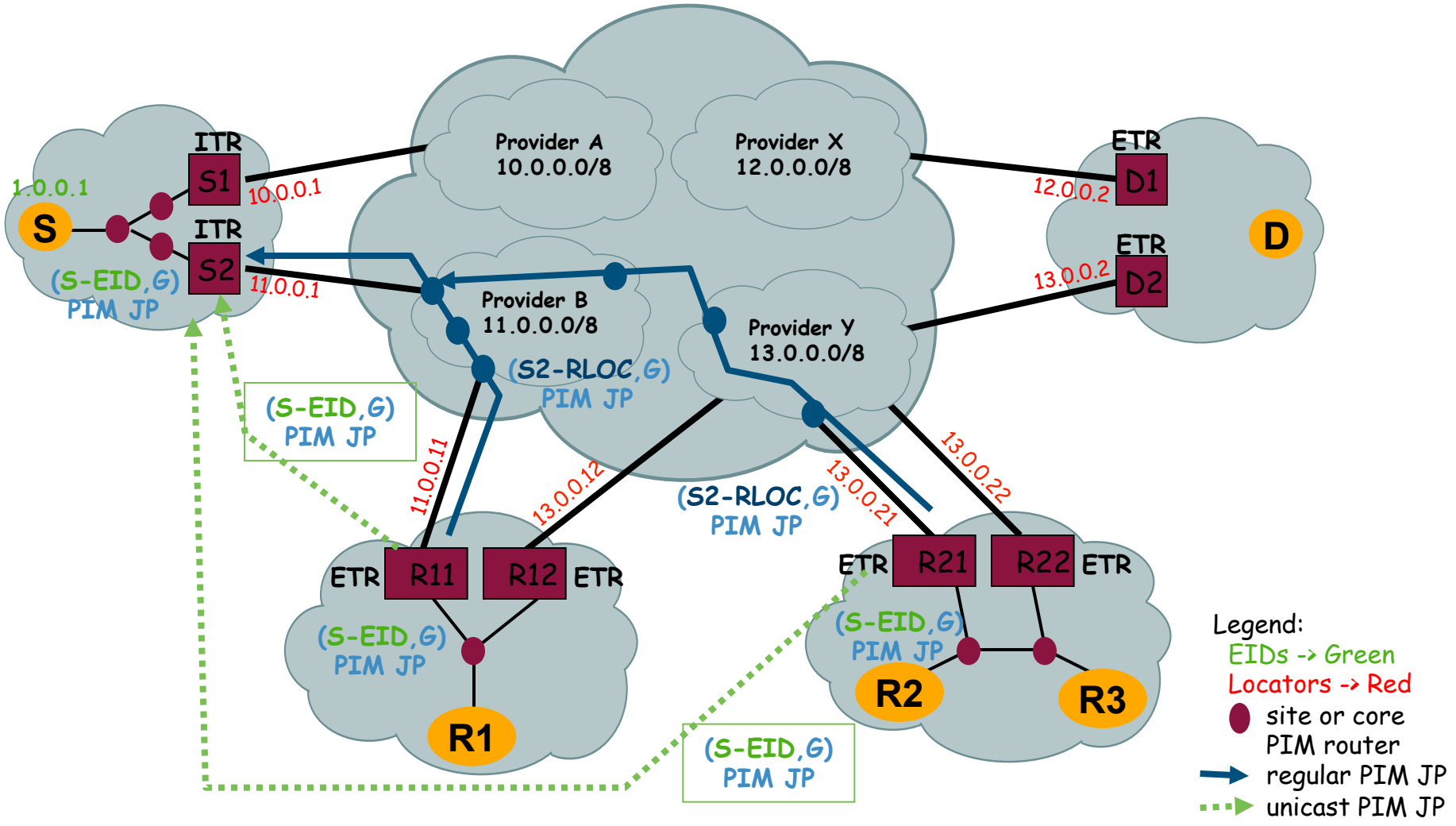
Multicast Packet Forwarding

- Keep EID state out of core network
- No head-end replication at source site
- Packets only go to receiver sites
- No changes to hosts, site routers, core routers
- Use existing protocols
- Support PIM SSM, don't preclude ASM & Bidir
- Have separate unicast and multicast policies

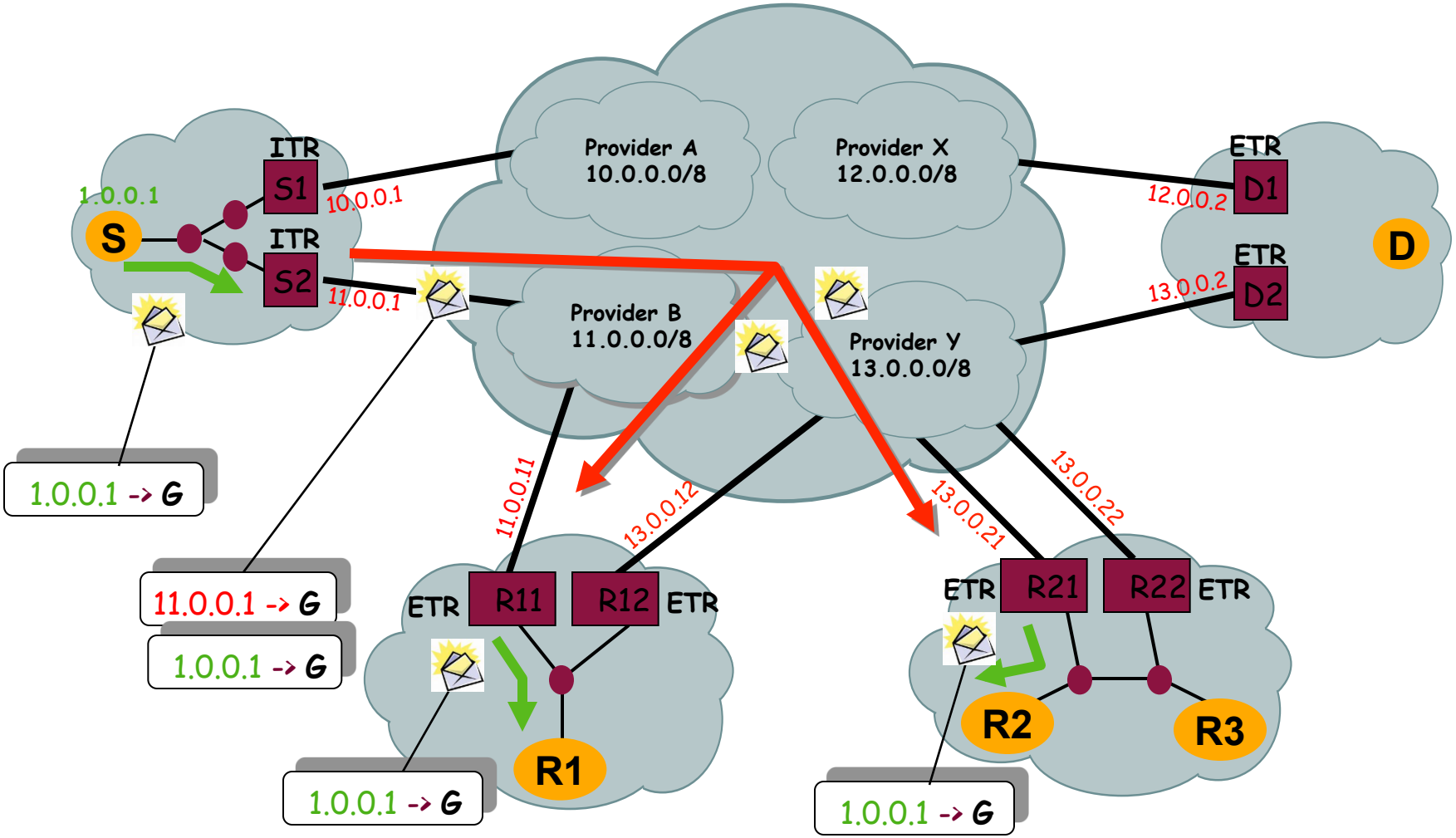
Multicast Packet Forwarding

- Group addresses have neither ID or Location semantics
  - G is topologically opaque - can be used everywhere
- (S-EID, G)
  - S-EID is source host
  - G is group address receivers join to
  - State resides in source and receiver sites
- (S-RLOC, G)
  - S-RLOC is ITR on multicast tree
  - G is group address receivers join to
  - State resides in core

Multicast Packet Forwarding



Multicast Packet Forwarding



What is the LISP Control-Plane?

- Definition for the “mapping cache” and “mapping database”
- Design for a modular scalable mapping service
- Examples are: ALT, CONS, EMACs, NERD
- Map-Servers and Map-Resolvers

  Interface LISP sites to mapping database service
- User tools for querying the mapping database

Mapping Database vs Mapping Cache

- **LISP Mapping Database**
  - Stored in all ETRs of each LISP site, not centralized
  - Authoritative Map-Replies sent from ETRs
  - Hard to DoS attack
- **LISP Map Cache**
  - Map-cache entries obtained and stored in ITRs for the sites they are currently sending packets to
  - ITRs must respect policy of Map-Reply mapping data
  - TTLs, RLOC up/down status, RLOC priorities/ weights
  - ETRs can tailor policy based on Map-Request source

Building a Scalable Database Service *

- Need a scalable EID -> RLOC mapping service
  - 10^{10} entries
- The Internet has only 2 large databases
  - BGP - pushes all information everywhere
  - DNS - pulls data on-demand from servers
- Scaling techniques
  - BGP summarizes routing information where it can
  - DNS caches information when needed
- Choose your poison
  - Trading off `(state * rate)`
  - `state` will be large, `rate` will have to be small
- We have designed several mapping database protocols
  - Tradeoff push versus pull benefit/cost
  - Did I say it needs to be scalable to 10^{10} entries :-)

Mapping Database Designs

- NERD - pure push
  - Documented but deprecated
- EMACs - pure pull
  - EID-prefixes hash to multicast groups
  - Pull mappings by sending Map-Request on multicast tree
  - Documented but deprecated
- CONS - hybrid push/pull
  - Push EID-prefixes using link-state at each hierarchical level of alternate topology
  - Pull mappings
  - Documented and deprecated
- ALT - hybrid push/pull
  - Push EID-prefixes using BGP on alternate topology of GRE tunnels
  - Pull mappings
  - ALT has the most promise
  - We are deploying ALT

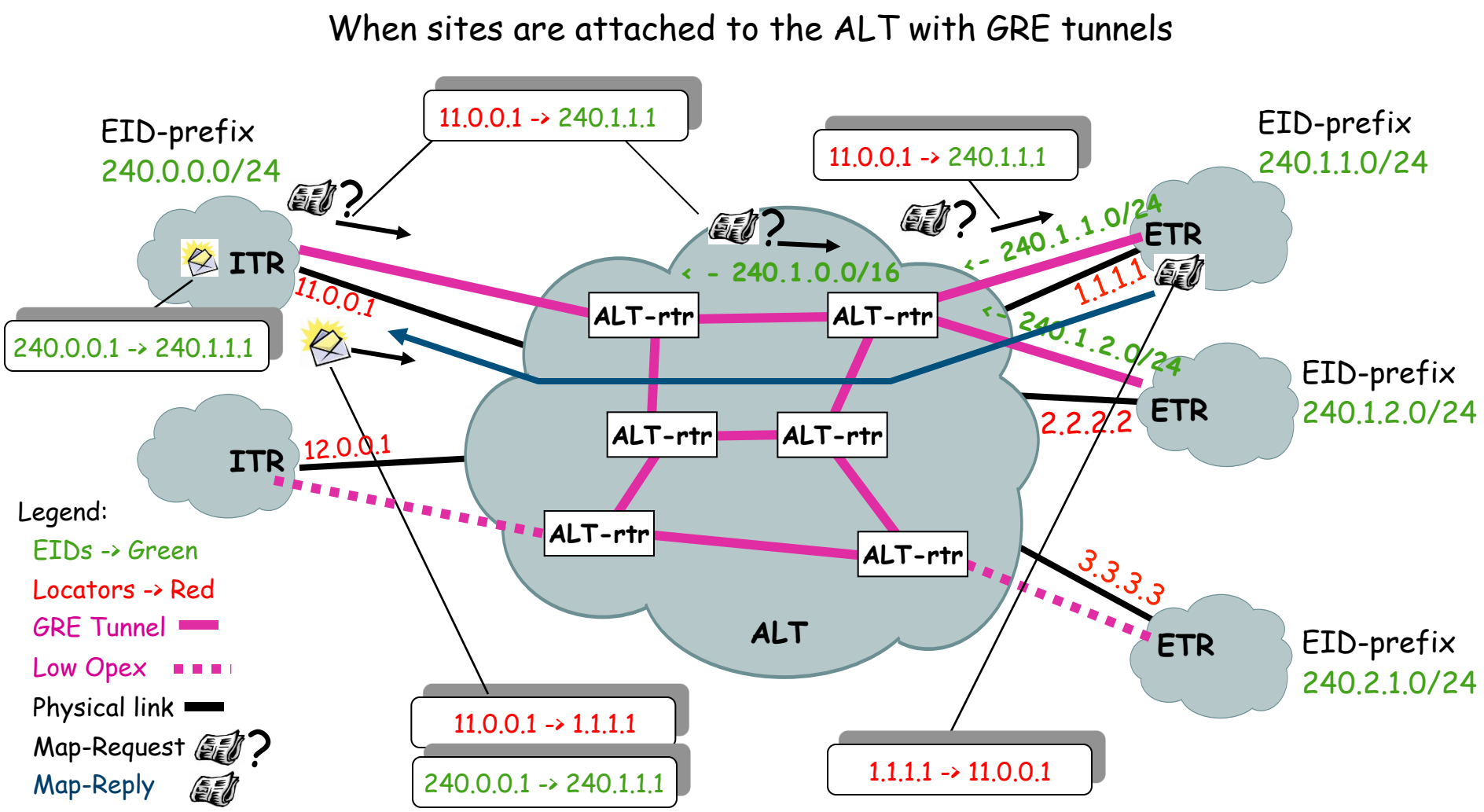
What Is LISP-ALT?

- Advertise EID-prefixes in BGP on an alternate topology of GRE tunnels
- An ALT Device is:
  - xTRs configured with GRE tunnels
  - Map-Servers
  - Map-Resolvers
  - Pure ALT-only router for aggregating other ALT peering connections
- An ALT-only device can be off-the-shelf gear:
  - Router hardware
  - Linux host
  - Just needs to run BGP and GRE

Service Interface for the Mapping Database

- ETRs register site EID-prefixes with Map-Servers
  - Securely with pair-wise trust model (no PKI needed)
  - Policy can be applied on Map-Servers before EID-prefix accepted into mapping service
- ETRs (at the site) are authoritative for their own database mappings
- When ALT is used, Map-Servers advertise EID-prefixes

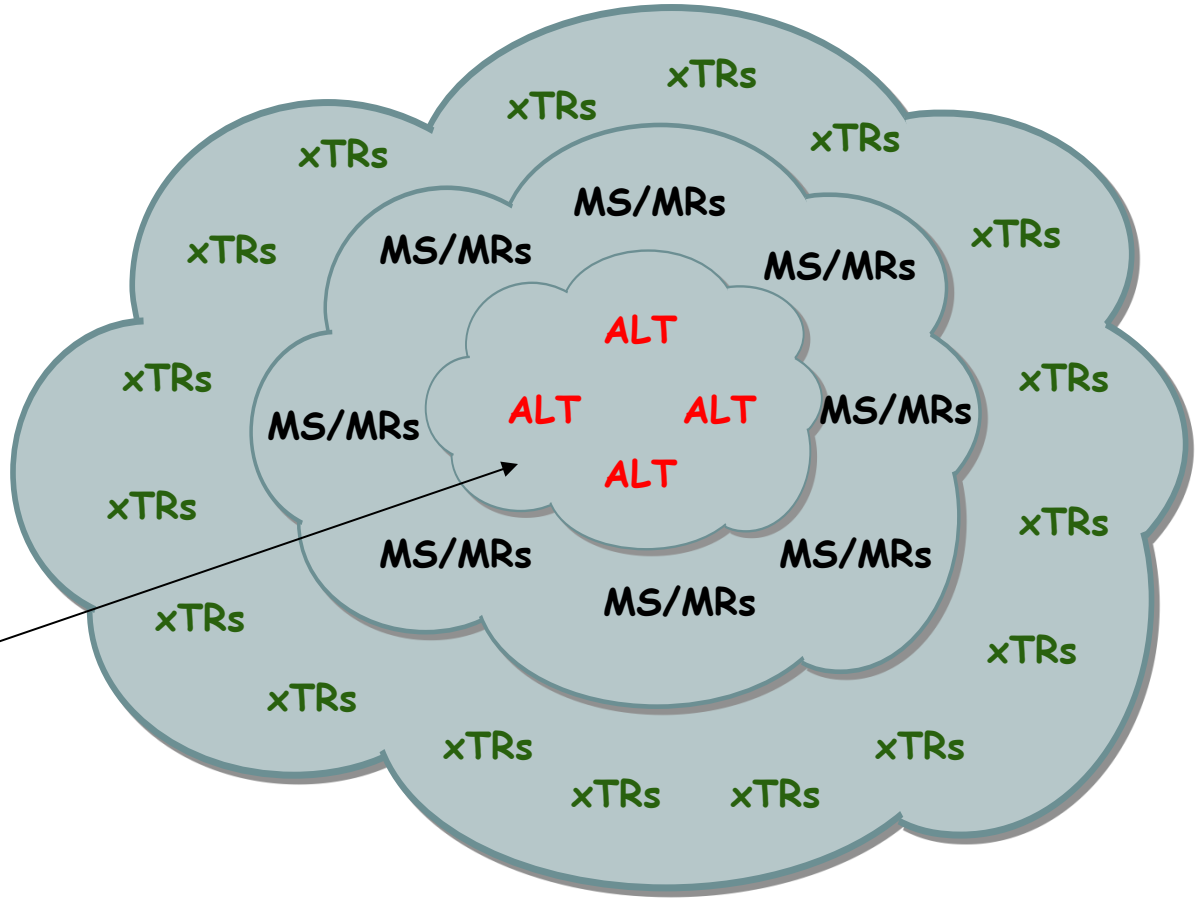
How LISP-ALT Works



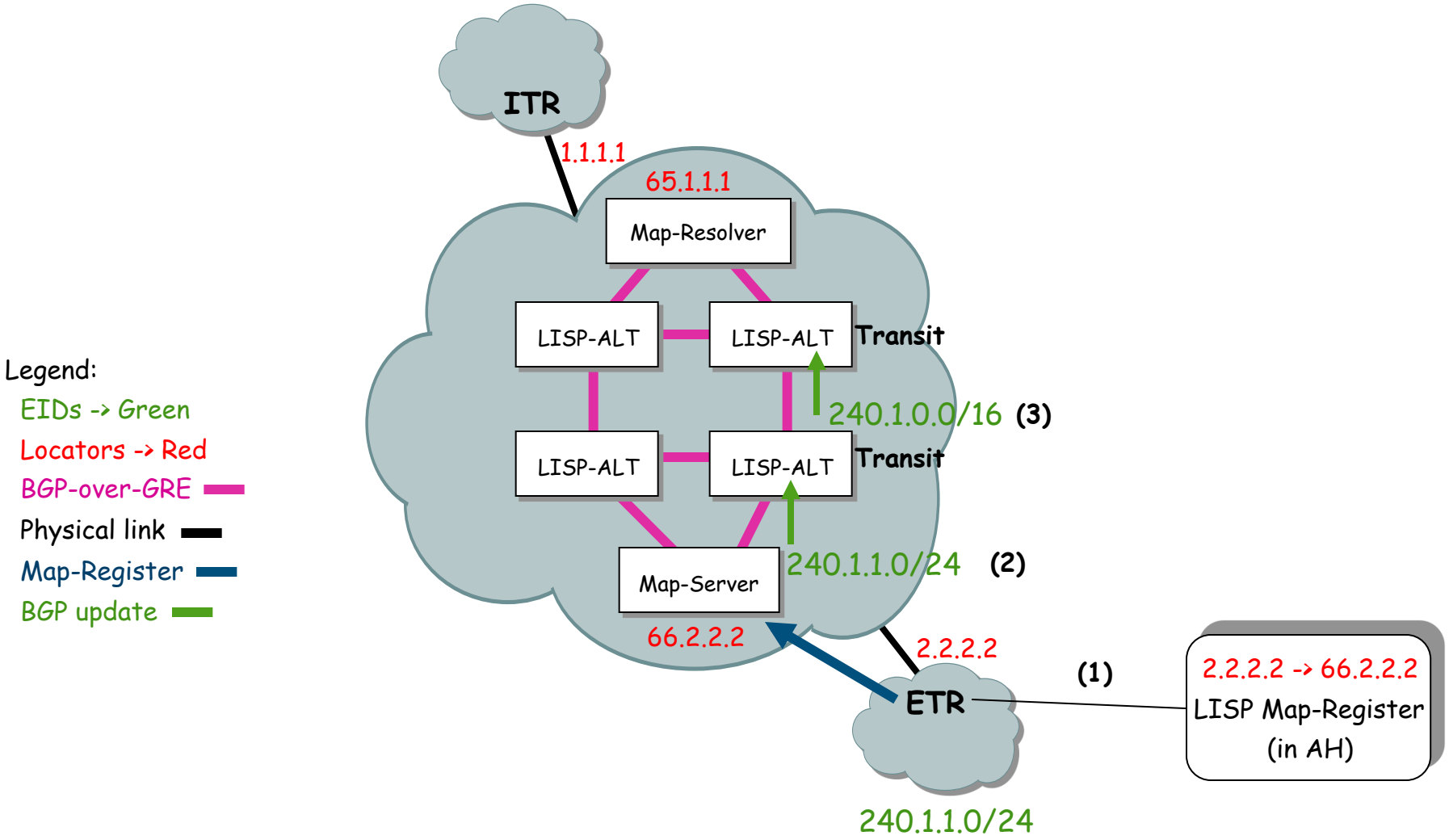
Modular Mapping Database Infrastructure

Legend:
- LISP Sites -> green
- 1st layer access infrastructure -> blue
- 2nd layer core infrastructure -> red

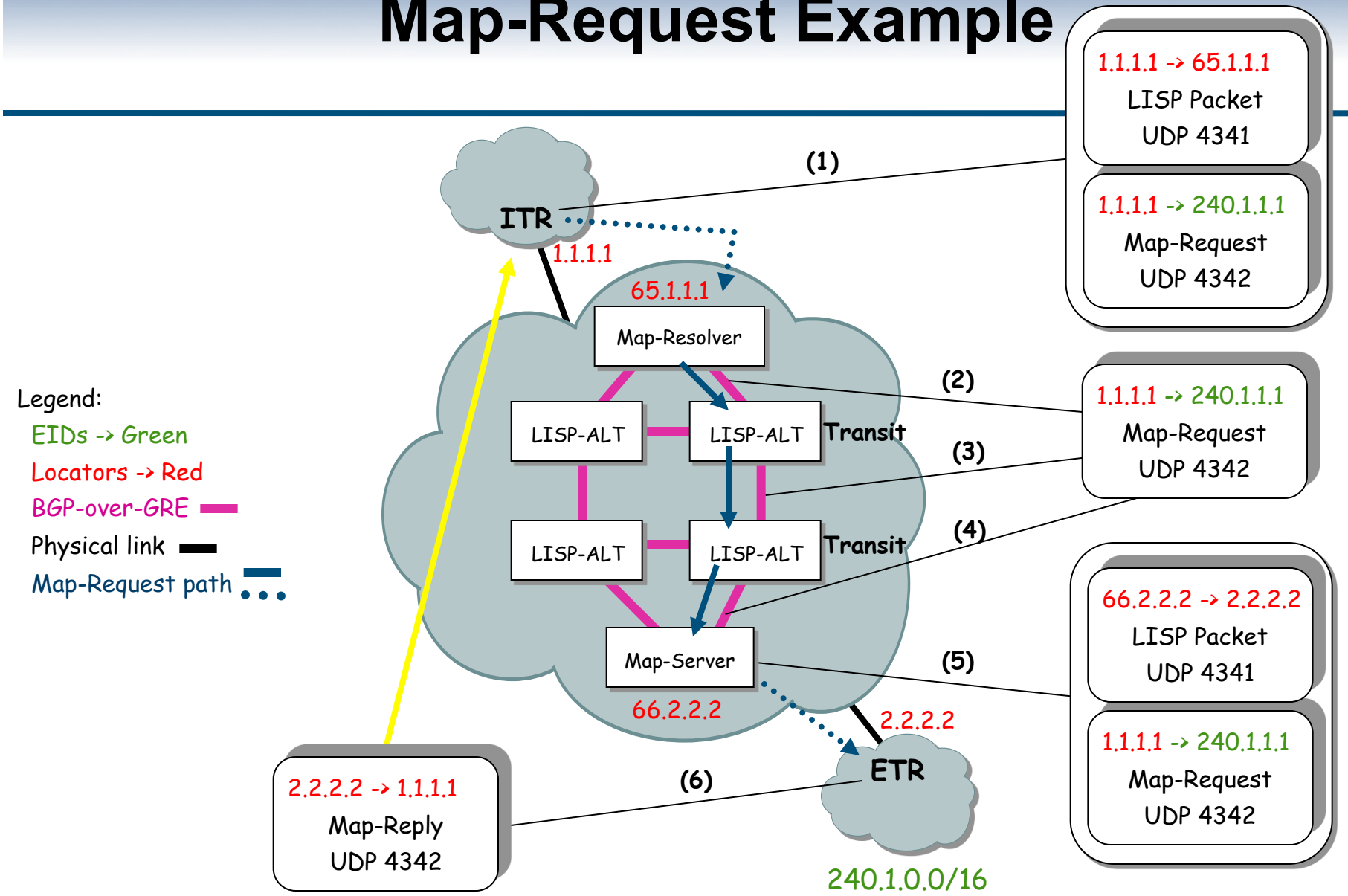
Want the ability to swap Mapping Database Infrastructure without changing sites



How Map-Server Registration Works



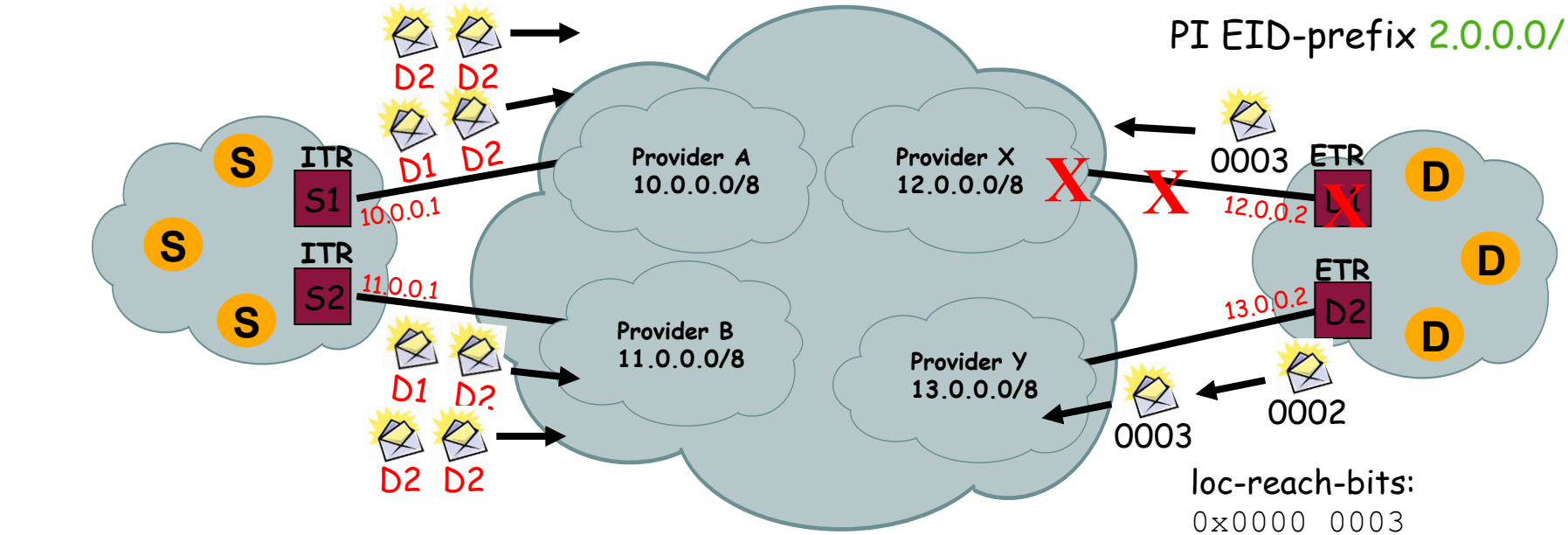
Map-Request Example



Locator Reachability

- When RLOCs go up and down
  - Don't want this reflected in mapping database -- keep be rate of database change very low
- Use following mechanisms:
  - Underlying BGP where available
  - ICMP Unreachables, when sent and accepted
  - Use data reception heuristics
  - Use loc-reach-bits in data packets and mapping data
- Don't use poll probing
  - Won't scale for the pair-wise number of sites and RLOC sets that will exist
- Data-plane locator reachability bits for certain classes of failures

How "loc-reach-bits" Work



Mapping Entry
- EID-prefix: 2.0.0.0/8
- Locator-set:
  - 12.0.0.2, priority: 1, weight: 50 (D1) -> ordinal 0
  - 13.0.0.2, priority: 1, weight: 50 (D2) -> ordinal 1

```
7654 3210
b'xxxx xxxx'
```

Legend:
- EIDs -> Green
- Locators -> Red

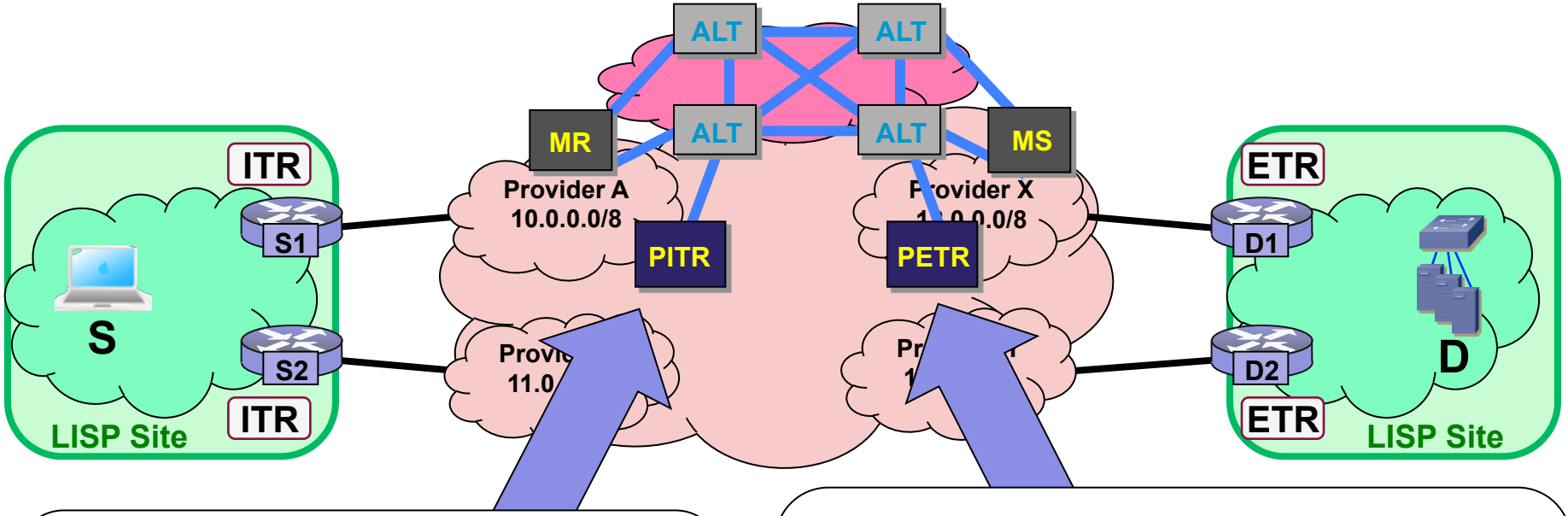
LISP Interworking

- LISP will not be widely deployed day-1
- Need a way for LISP-capable sites to communicate with rest of Internet
- Two basic Techniques
  - LISP Network Address Translators (LISP-NAT)
  - Proxy Tunnel Routers (PTRs)
- PTRs have the most promise
  - Creates a monetized service for infrastructure players

LISP Interworking

- These combinations must be supported
  - Non-LISP site to non-LISP site
    - Today's Internet
  - LISP site to LISP site
    - Encapsulation over IPv4 makes this work
    - IPv4-over-IPv4 or IPv6-over-IPv4
  - LISP-R site to non-LISP site
    - When LISP site has PI or PA routable addresses
  - LISP-NR site to non-LISP site
    - When LISP site has PI or PA non-routable addresses

LISP Interworking
Proxy Ingress/Egress Tunnel Routers (PITR/PETR)



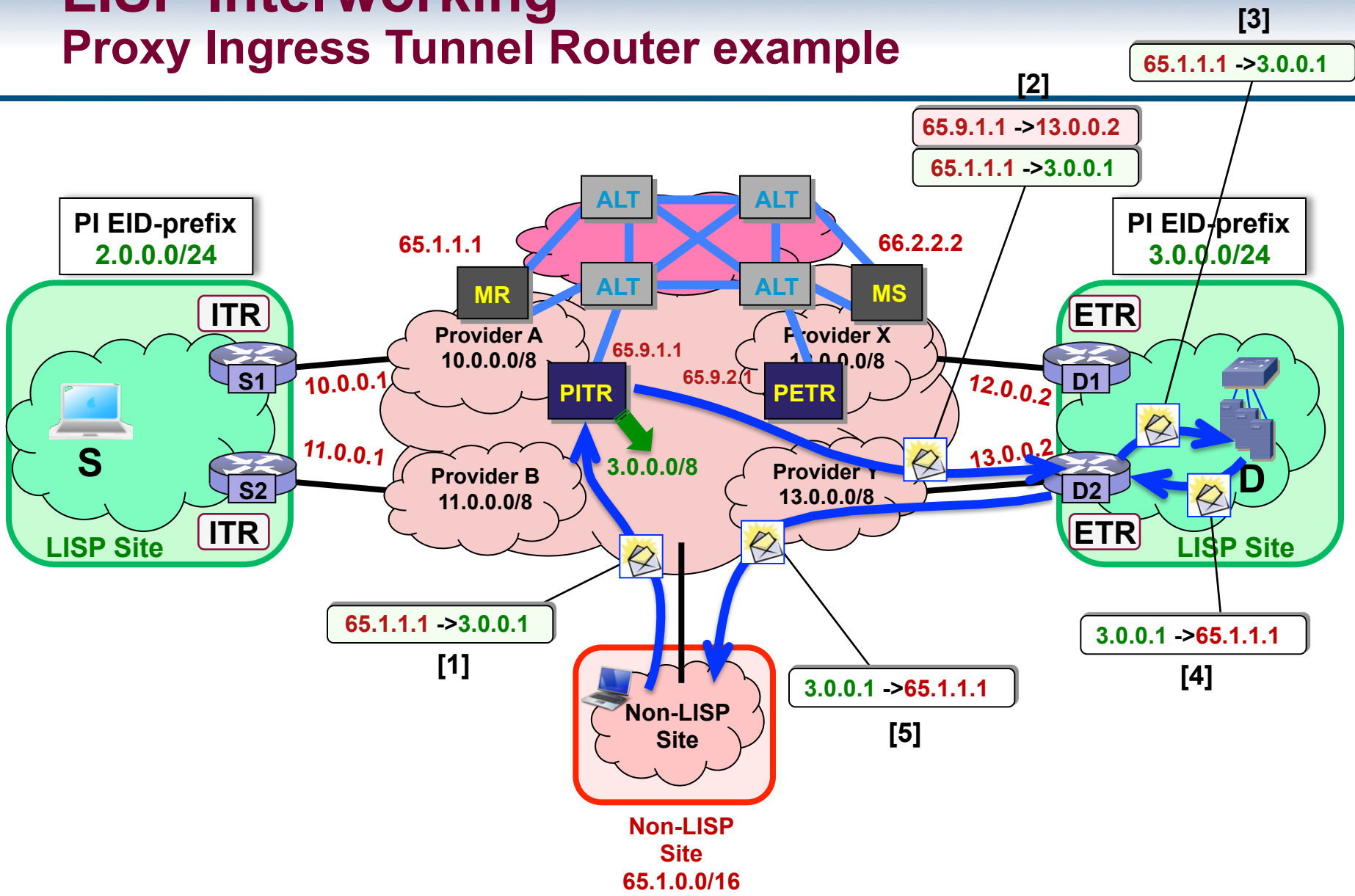
PITR – Proxy ITR

- Receives traffic from **non-LISP**sites; encapsulates traffic to **LISP sites**
- Advertises coarse-aggregate **EID** prefixes
- **LISP sites** see ingress TE “day-one”

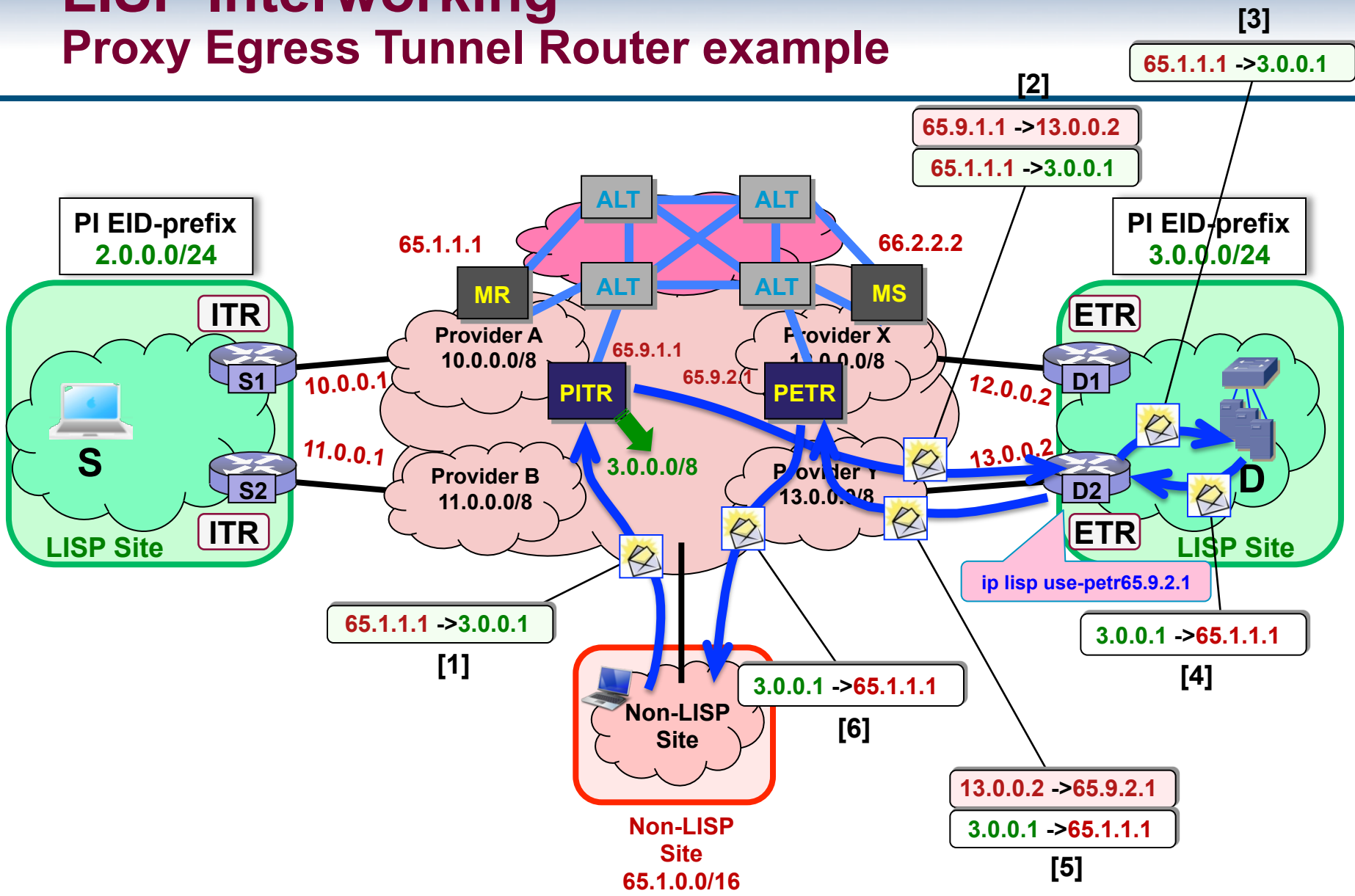
PETR – Proxy ETR

- Allows **IPv6 LISP**sites with **IPv4 RLOCs** to reach **IPv6 LISP**sites that only have **IPv6 RLOCs**
- Allows **LISP**sites with uRPF restrictions to reach **non-LISP**sites

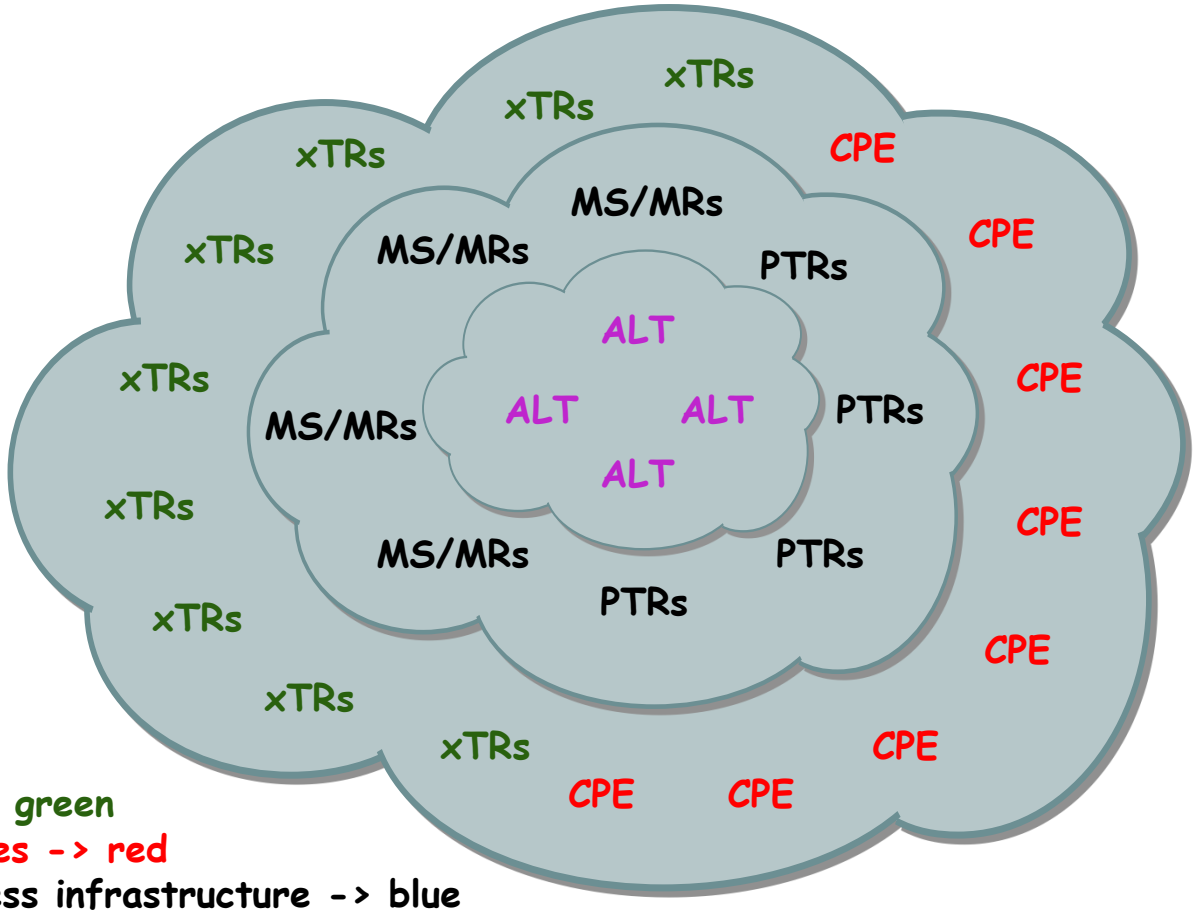
LISP Interworking
Proxy Ingress Tunnel Router example



LISP Interworking
Proxy Egress Tunnel Router example



The Whole Picture - LISP based Internet



Legend:

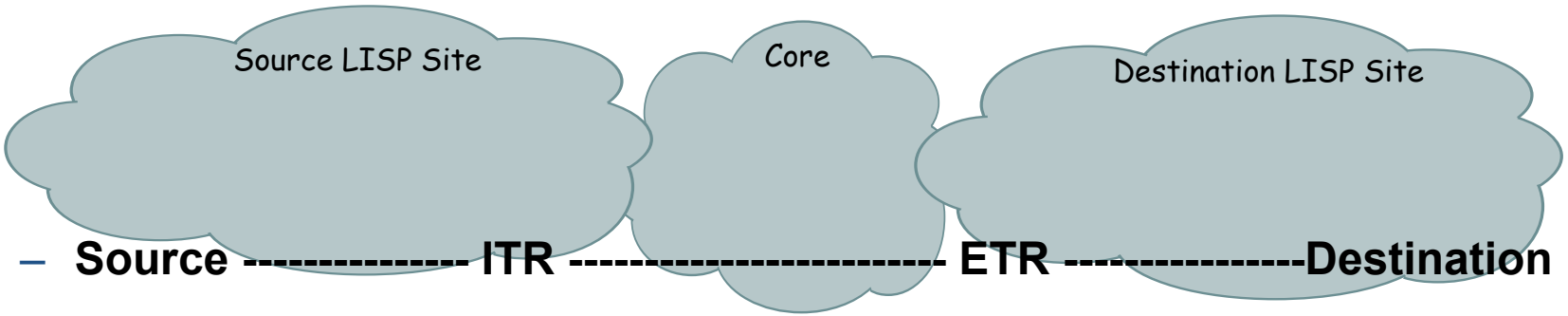
- LISP Sites -> green
- Non-LISP Sites -> red
- 1st layer access infrastructure -> blue
- 2nd layer core infrastructure -> violet

Security in LISP

- EID-prefixes are injected into the mapping system securely
  - Uses shared-key IPsec-AH
  - Using access control on map-server
- ITRs do not accept unsolicited Map-Replies
- ITRs accept Map-Replies only with nonces inserted in Map-Requests
- ALT can be secured with sBGP
- Map-Replies could carry public keys
  - So ITR can encrypt encapsulated data with ESP headers

Management of LISP

- **Traceroute from LISP site to LISP site**
  - **No tunnel mode - these aren't tunnels but encapsulating nodes**
  - **3-segment path**



- **Source ---------- ITR ---------- ETR ----------Destination**
- **<- EIDs -> <- RLOCs -> <- EIDs ->**

Management of LISP

- **LISP Internet Groper (lig)**
  - **Fetches a database mapping entry**
  - **Both router and host lig available**

```
titanium-dino# lig titanium-dmm.lisp4.net
Send map-request to 128.223.156.139 for 153.16.10.254 ...
Received map-reply from 128.223.156.134 with rtt 0.042518 secs

Map-cache entry for titanium-dmm.lisp4.net EID 153.16.10.254:
153.16.10.0/24, uptime: 00:00:01, expires: 23:59:58, via map-reply, auth
  Locator           Uptime    State  Priority/Weight  Packets In/Out
  128.223.156.134   00:00:01  up     1/100            0/0
```

Management of LISP

- **LISP Internet Groper (lig)**
  - **Verifies you have registered your own EID-prefix to the mapping system**

```
rutile# lig self
Send loopback map-request to 128.223.156.139 for 153.16.12.0 ...
Received map-reply from 207.98.65.94 with rtt 0.002839 secs

Map-cache entry for EID 153.16.12.0:
153.16.12.0/24, uptime: 00:11:12, expires: 23:59:57, via map-reply, self
  Locator       Uptime    State  Priority/Weight  Packets In/Out
  207.98.65.94  00:11:12  up     1/100            0/0
```

Management of LISP

- **LISP Internet Groper (lig)**
  - **Supports cross address-family**

```
titanium-dino# lig self6
Send loopback map-request to 193.0.0.170 for 2610:d0:2105:: ...
Received map-reply from 173.8.188.25 with rtt 0.231016 secs

Map-cache entry for EID 2610:d0:2105::/48, uptime: 00:00:01, expires: 23:59:58, via map-reply, self
  Locator             Uptime    State  Priority/Weight  Packets In/Out
  173.8.188.25        00:00:01  up     1/33             0/0
  173.8.188.26        00:00:01  up     1/33             0/0
  173.8.188.27        00:00:01  up     1/33             0/0
  2002:ad08:bc19::1   00:00:01  up     2/0              0/0
```

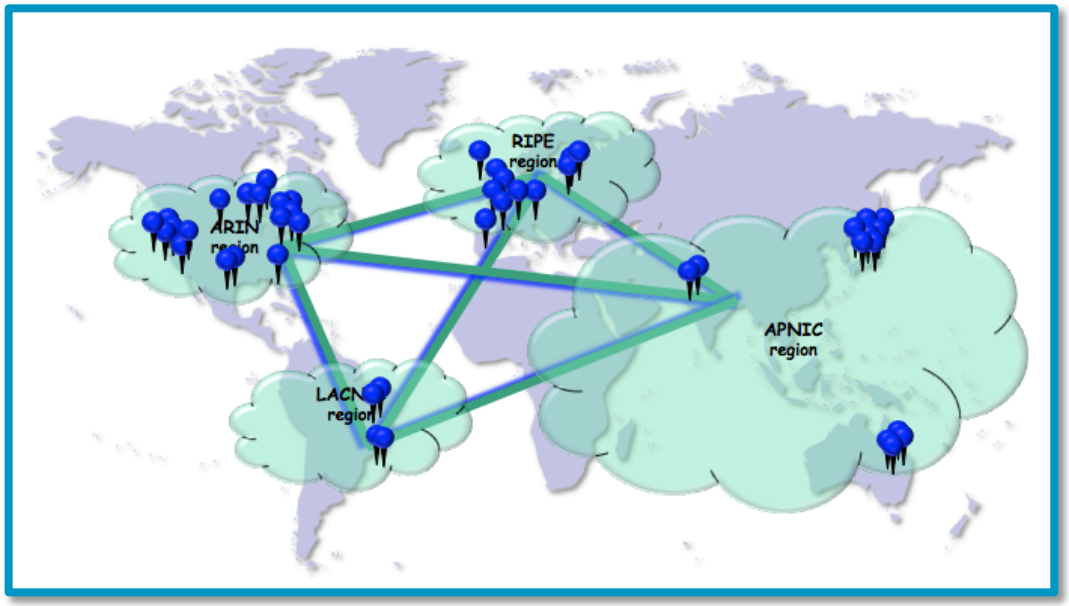
LISP Global Pilot Network
Deployed/Operational Experience

Cisco-operated

- > 3 years operational
- > 85+ sites, 13 countries

Six LISP implementations

- Cisco: IOS/IOS-XE, NX-OS
- FreeBSD: OpenLISP
- Two Linux implementations
- Android implementation (coming)



Who's Using LISP?

Company	IPv4 Sites	IPv6 Sites
CISCO	http://www.lisp4.net http://lisp4.cisco.com	http://www.lisp6.net http://lisp6.cisco.com
facebook	http://www.lisp4.facebook.com	http://www.lisp6.facebook.com
QUALCOMM		http://www6.eudora.com http://myvpn6.qualcomm.com
InTouch	http://www.lisp.intouch.eu/	http://www.lisp.intouch.eu/

LISP Standardization Effort
Open Design

Draft	Current Status	Next Steps/Target
LISP base protocol (draft-ietf-lisp-09)	WG Document Submitted: 10/11/2010	Experimental RFC by 3/31/2011
LISP+ALT (draft-ietf-lisp-alt-05)	WG Document Submitted: 10/18/2010	Experimental RFC by 3/31/2011
LISP Interworking (draft-ietf-lisp-interworking-01)	WG Document Submitted: 08/26/2010	Experimental RFC by 3/31/2011
LISP Map Server (draft-ietf-lisp-ms-06)	WG Document Submitted: 10/18/2010	Experimental RFC by 3/31/2011
LISP Multicast (draft-ietf-lisp-multicast-04)	WG Document Submitted: 10/12/2010	Experimental RFC by 3/31/2011
LISP Internet Groper (draft-ietf-lisp-lig-01)	WG Document Submitted: 10/12/2010	Several implementations (incl. open source) available
LISP Mobile Node (draft-meyer-lisp-mn-04)	Not WG Document Submitted: 10/25/2010	Three prototype implementations underway
LISP Canonical Address Format (draft-farinacci-lisp-lcaf-04)	Proposed for WG adoption Submitted: 10/14/2010	-04 update sent to WG list
LISP MIB (draft-schudel-lisp-mib-00)	Not WG Document Submitted: 8/16/2010	-00 update sent to WG list