LISP – Új generációs hálózati architektúra

Nagy Tibor
Cisco Systems
tinagy@cisco.com

Agenda

- Problem Statement
- Architectural Concepts
- LISP Data Plane
- LISP Control Plane
- Interworking LISP Sites and Legacy Sites
- LISP Security and Management
- LISP Use Cases
- LISP Implementation Status
- References
- Q & A

Problem Statement

- What provoked this?

  BGP has been “holding the Internet together” for close to two decades now.

  October 2006 workshop convened by the Internet Architecture Board concluded that: "***routing scalability is the most important problem facing the Internet today and must be solved***“

  Routing scalability includes the size of the DFZ RIB and FIB, and has implications on both RIB/FIB growth and routing convergence times.

  `RFC 4984`

  More info on problem statement:

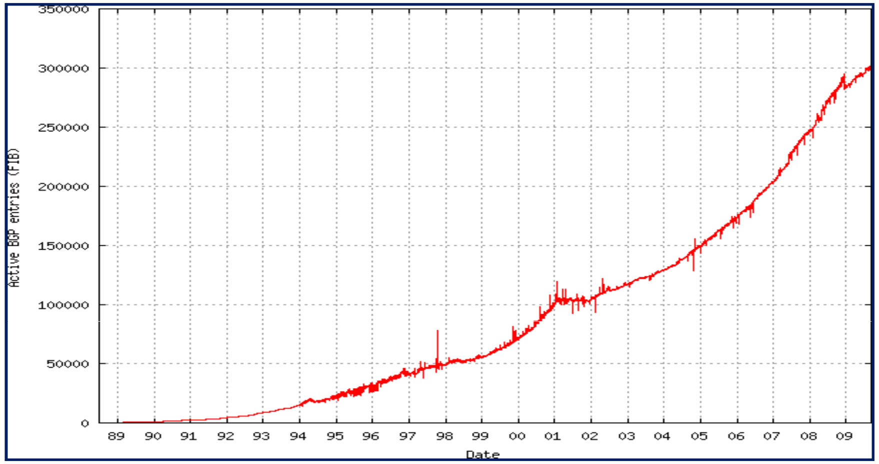
  `http://www.vaf.net/~vaf/apricot-plenary.pdf`

- First and foremost - scale the Internet

LISP Overview
The Problem Statement

RFC4984

- LISP originally conceived to address Internet Scaling



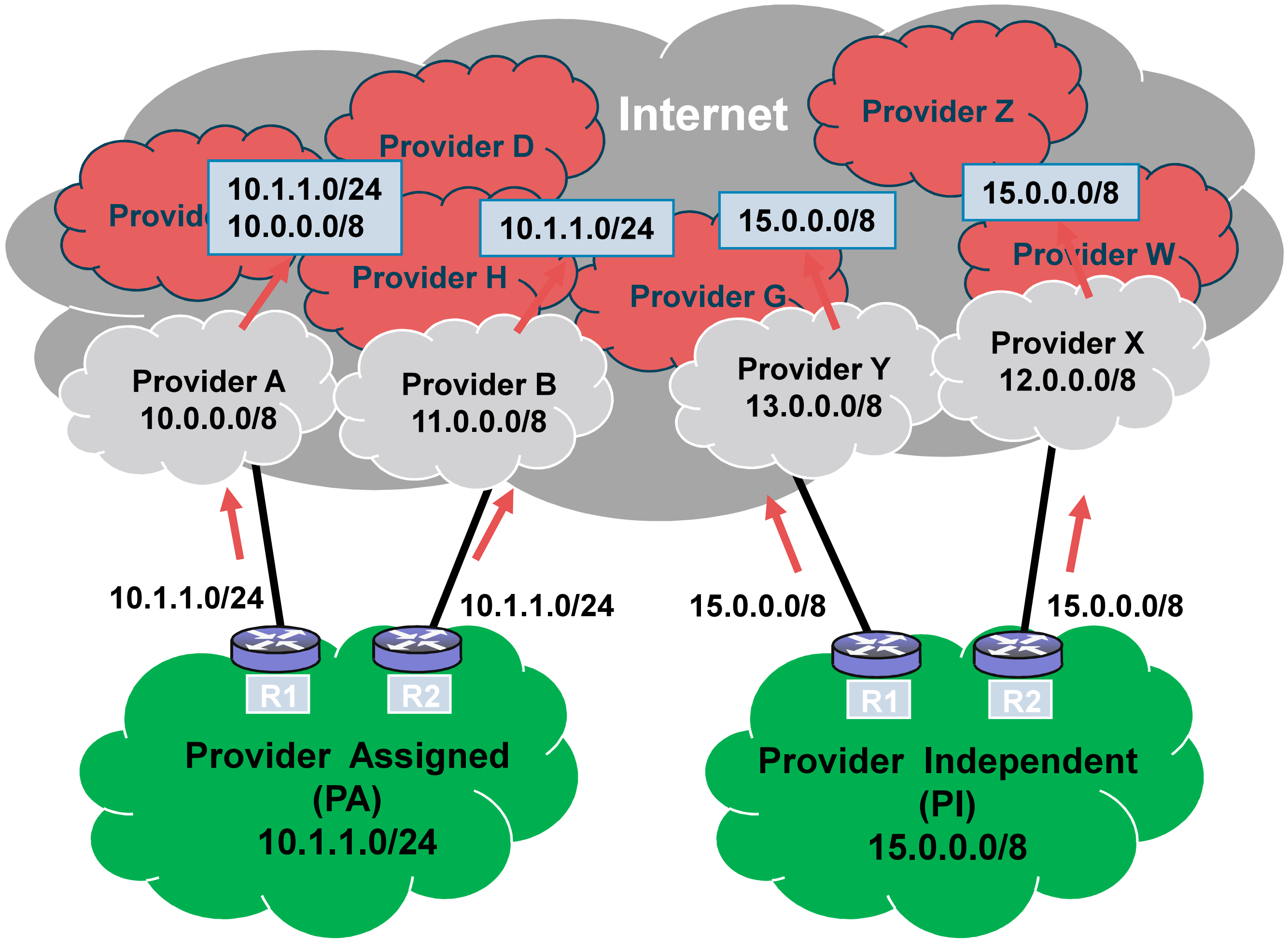
IP addresses **today** denote both **location** and **identity**

Overloading address function makes efficient routing system impossible; pollutes DFZ

IPv6 does not fix this problem

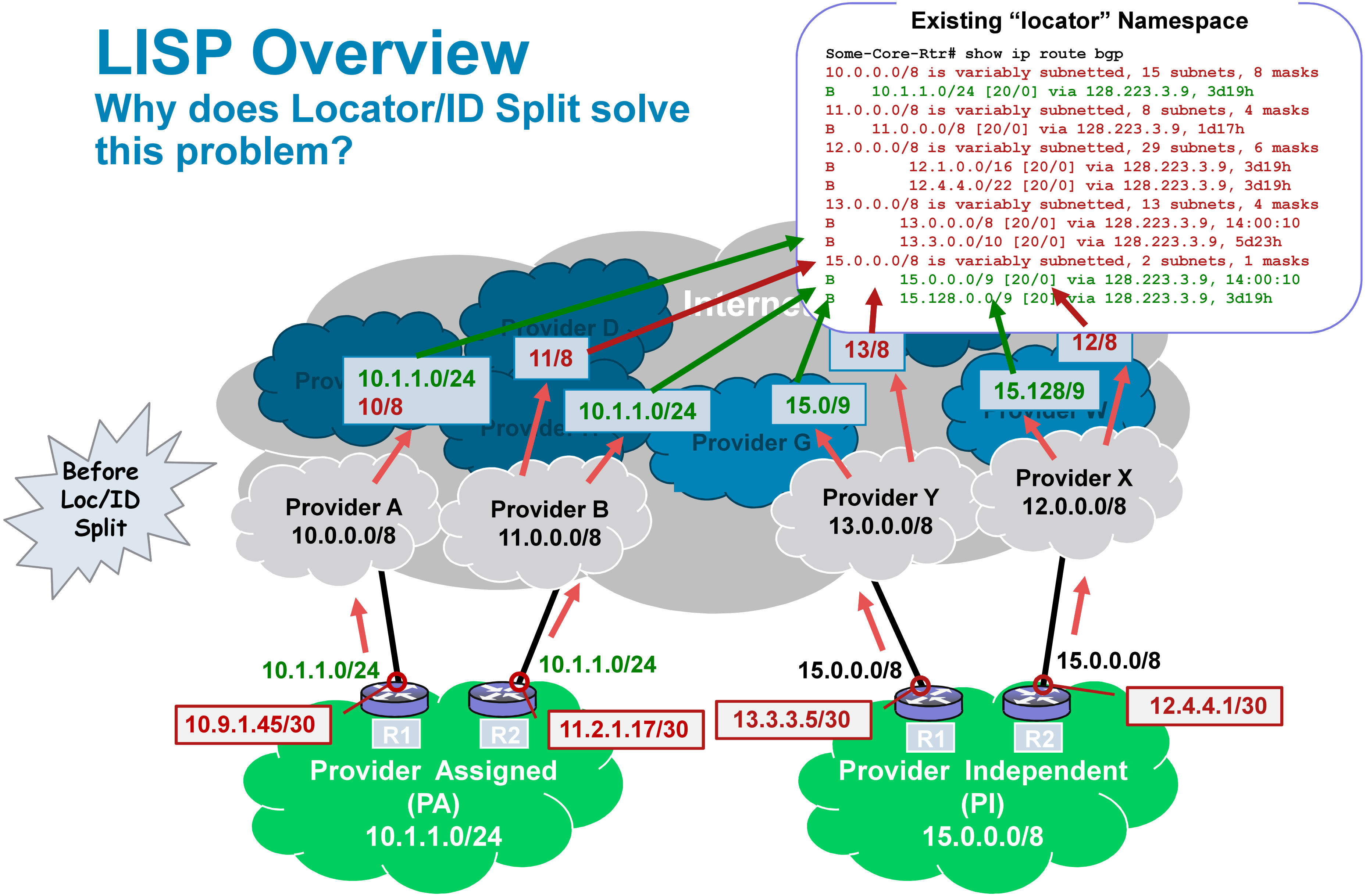
- LISP creates two namespaces: **EID** and **RLOC**
- LISP addresses other "problem spaces" as well
  - Multi-homing without the need for BGP, and with "ingress" TE
  - Mobility (handset, server, virtual computing)
  - IPv4/IPv6 co-existence

What Pollutes the Internet



LISP Overview

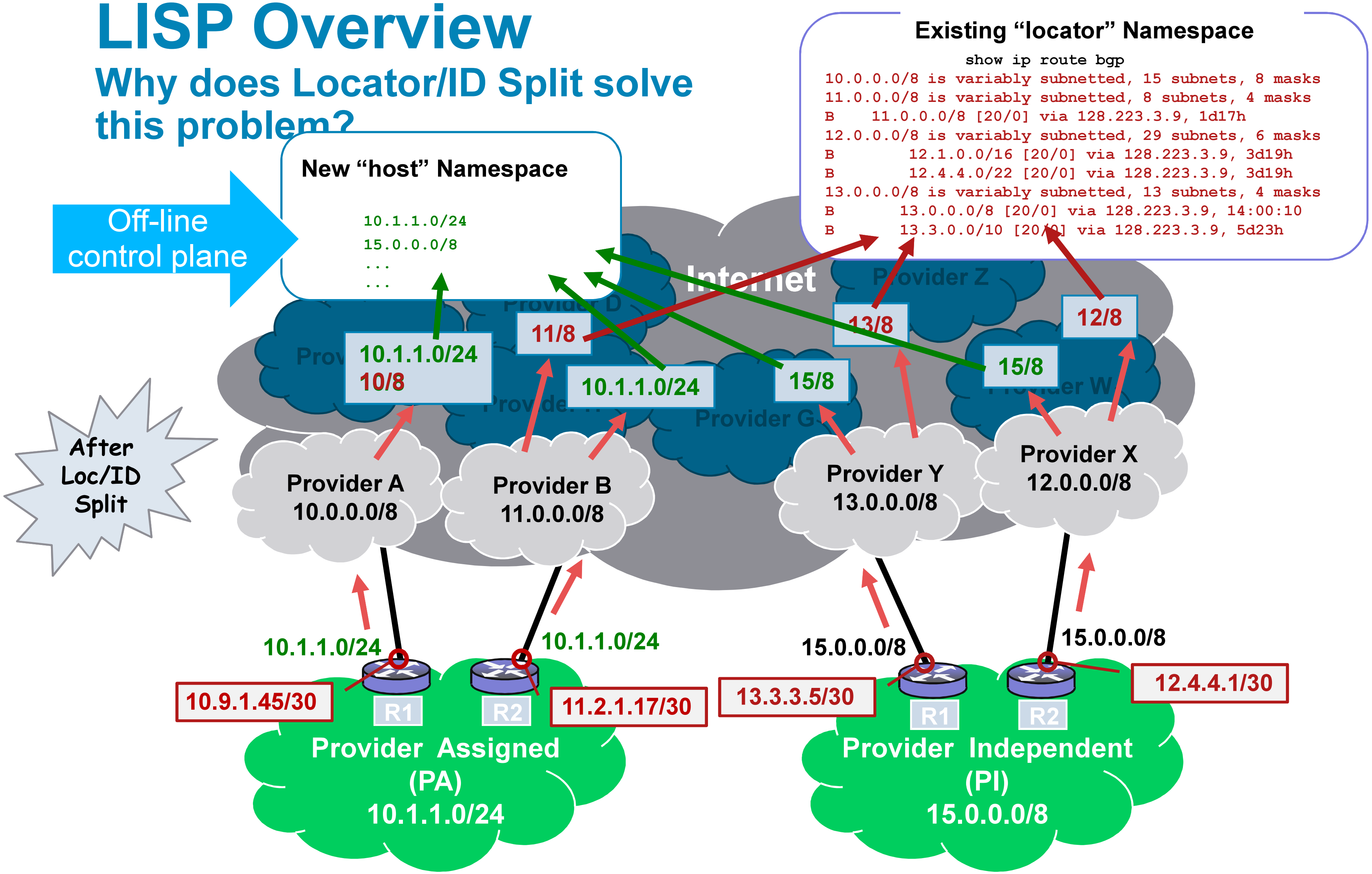
Why does Locator/ID Split solve this problem?



- Addresses at sites, both PA and PI, can get de-aggregated by multi-homing
- Aggregates for infrastructure addresses (e.g. CE-PE links) get advertised as well

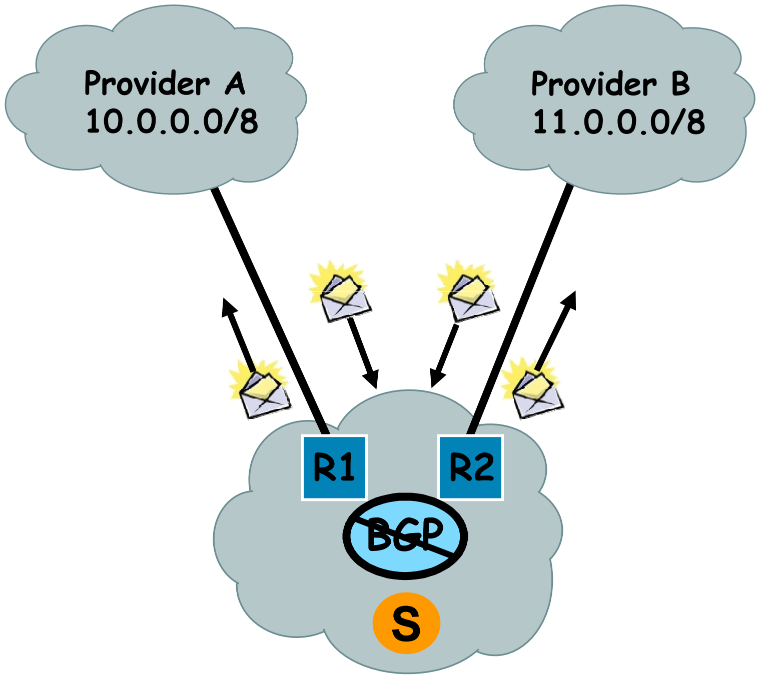
LISP Overview

Why does Locator/ID Split solve this problem?



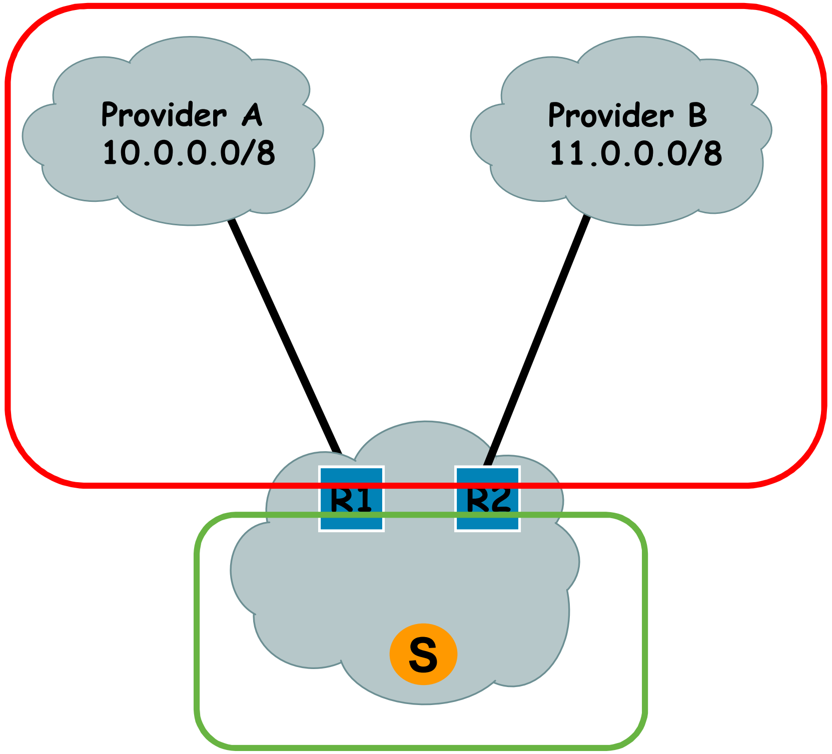
- Addresses at sites, both PA and PI, can get de-aggregated by multi-homing
- Aggregates for infrastructure addresses (e.g. CE-PE links) get advertised as well

Foster Growth in Multi-Homing



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

Foster Growth in Multi-Homing



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

Separating (or Adding) an Address

Changing the Semantics of the IP Address

ID & Location

IPv6: 2001:0102:0304:0506 1111:2222:3333:4444

Locator — ID

ID & Location

IPv4: 209.131.36.158 10.0.0.1

Locator — ID

If PI, get new locator
If PA, get new ID

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
- By allowing **Locators** to change…
  - 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

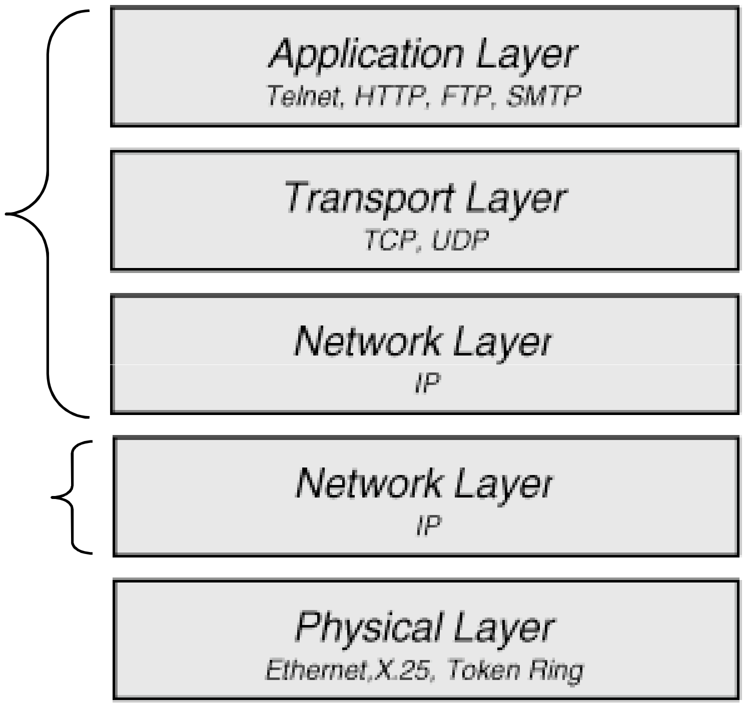
What Is LISP?

- 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?

Host Stack:
supplies IDs

Router:
supplies RLOCs by adding new header

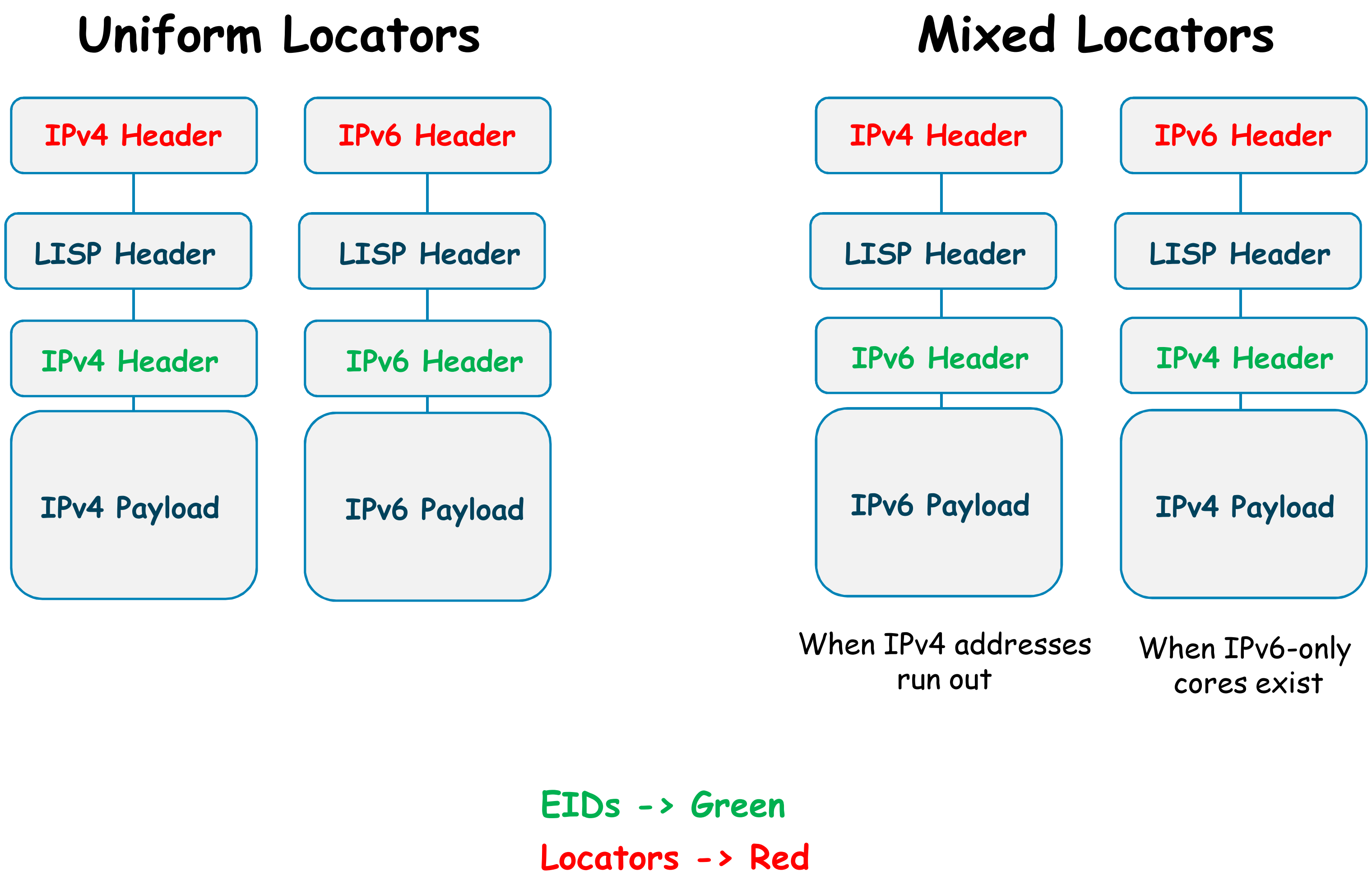


“Jack-Up” or “Map-n-Encap”

Encapsulation contd.

```
         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
        +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      / |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 (xxx)       |       Dest Port (4341)        |
    UDP +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      \ |           UDP length                  UDP Checksum            |
        +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      / |N|L|E|  rflags |                 Nonce                         |
   LISP +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      \ |                 Locator Status Bits                           |
        +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      / |Version|  IHL  |Type of Service|          Total Length         |
     /  +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    /   |         Identification        |Flags|      Fragment Offset    |
   /    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 IH     |  Time to Live |    Protocol   |         Header Checksum       |
   \    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    \   |                          Source EID                           |
     \  +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      \ |                        Destination EID                        |
        +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

LISP Supports IPv4 and IPv6



The LISP Data Plane

- Design for encapsulation and router placement
- Design for locator reachability
- Supports Unicast and Multicast Data Services

  Unicast Support

  – No changes at hosts, core routers

  – Minor changes at site routers

  Multicast Support

  – No changes at hosts, sites routers, core routers

  – Support PIM SSM, doesn’t preclude ASM & Bidir

  Supports separate Unicast and Multicast policies

LISP Data Plane Network Devices

- ITR - Ingress Tunnel Router

  Receives packets from site-facing interfaces and encaps to remote LISP site or natively forwards to non-LISP site

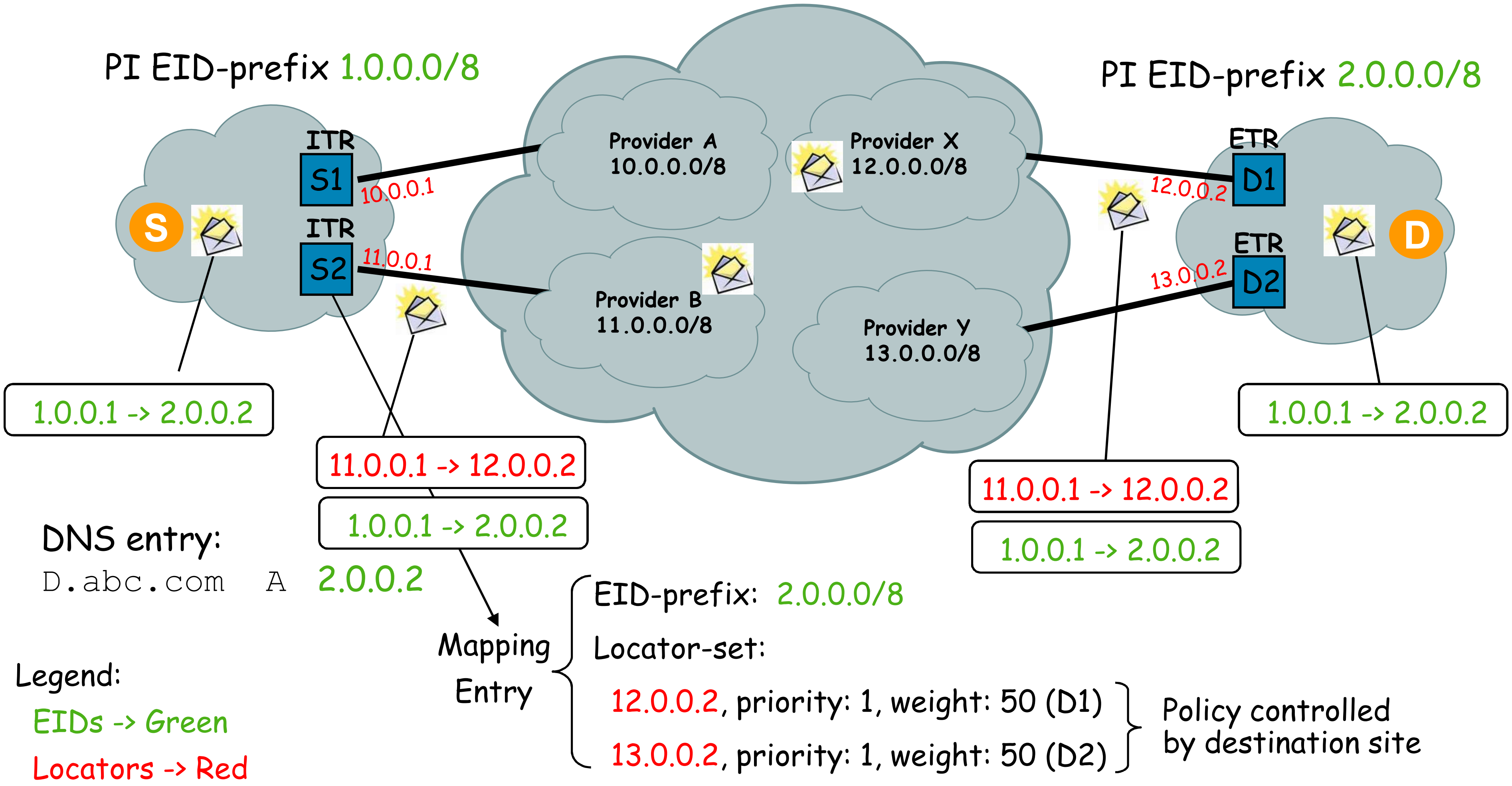
  Typically deployed as a CE device

- ETR - Egress Tunnel Router

  Receives packets from core-facing interfaces and decaps to deliver to local EIDs at the site

  Typical deployed as a CE device

Unicast Packet Forwarding Example



The LISP Control Plane

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

LISP Control Plane Network Devices

- Map-Server
  - Configures “lisp site” policy to authenticate which LISP sites can Register to it
  - Provides a service interface to the ALT, injects routes in ALT BGP when site Registers
  - Receives Map-Requests over the ALT and encaps them to registered ETRs
- Map-Resolver
  - Receives Map-Request which are encapsulated by ITRs
  - Provides a service interface to the ALT, decaps Map-Request and forwards on the ALT topology
  - Send Negative Map-Replies in response to Map-Requests for non-LISP sites

The LISP Control Plane

- Control Plane **EID** Registration

  Map-Register messages

  – sent by an ETR to a Map-Server to register its associated **EID** prefixes, and to specify the **RLOC(s)** to be used by the Map-Server when forwarding Map-Requests to the ETR

- Control Plane “Data-triggered” mapping service

  Map-Request messages

  – sent from an ITR when it needs a mapping for an **EID**, wants to test an **RLOC** for reachability, or wants to refresh a mapping before TTL expiration

  Map-Reply messages

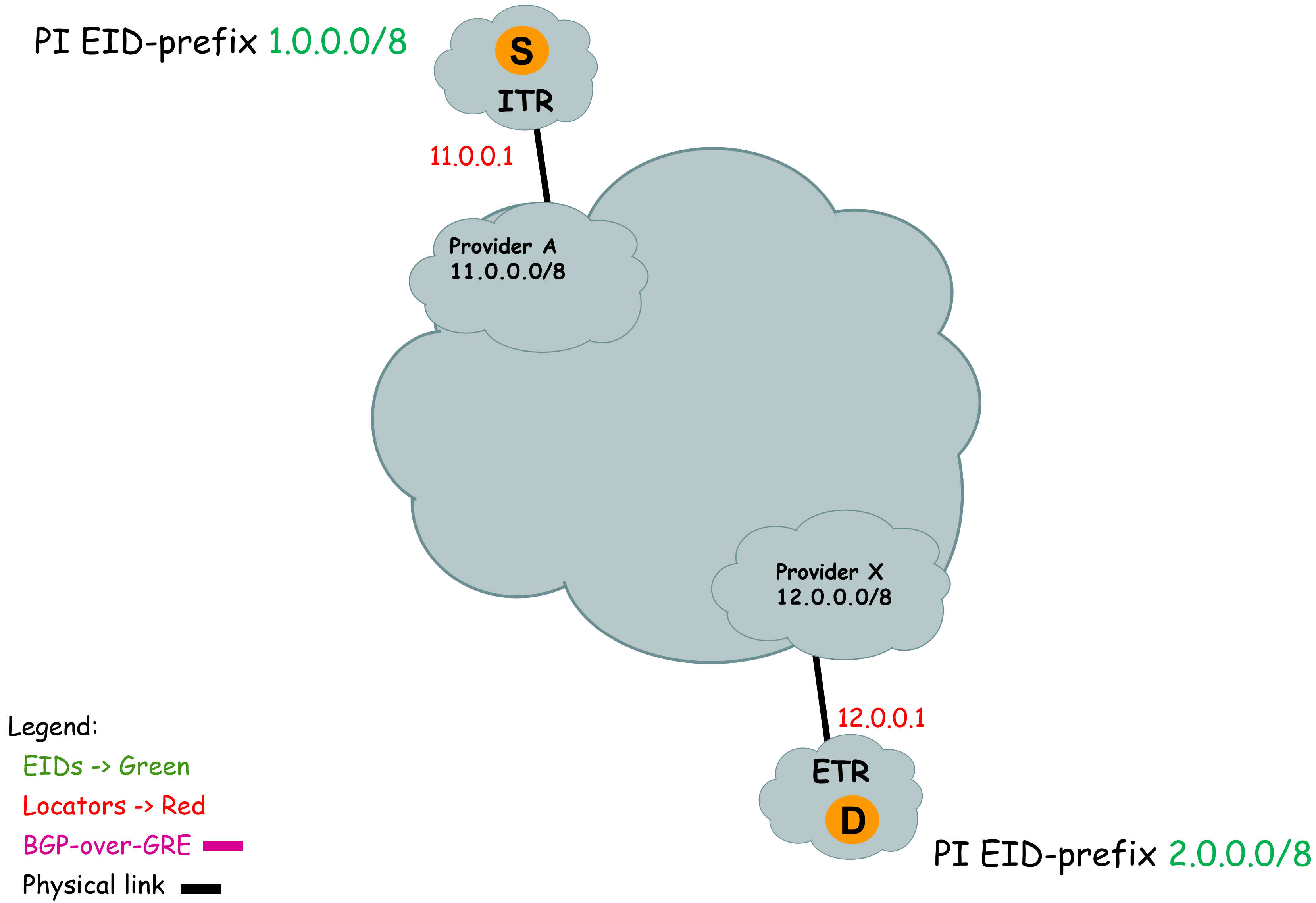
  – sent from an ETR in response to a valid map-request to provide the **EID**/**RLOC** mapping and site ingress Policy for the requested **EID**

LISP-ALT

- Map-Servers advertise EID-prefixes to ALT for scalability
- ALT Advertise EID-prefixes in BGP on an *alternate topology* of GRE tunnels
- An ALT Device can be:
  - 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

LISP Control Plane

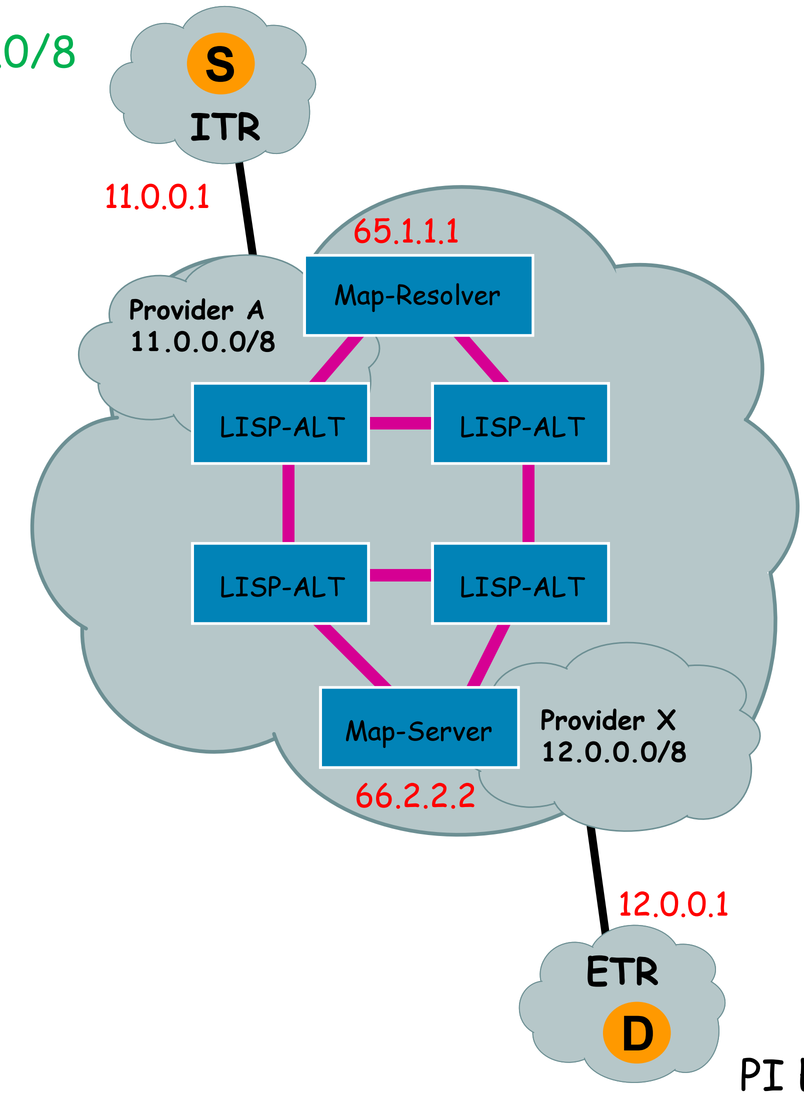
EID Topology



LISP Control Plane

Map-Resolver, Map-Server
and ALT Infrastructure

PI EID-prefix 1.0.0.0/8



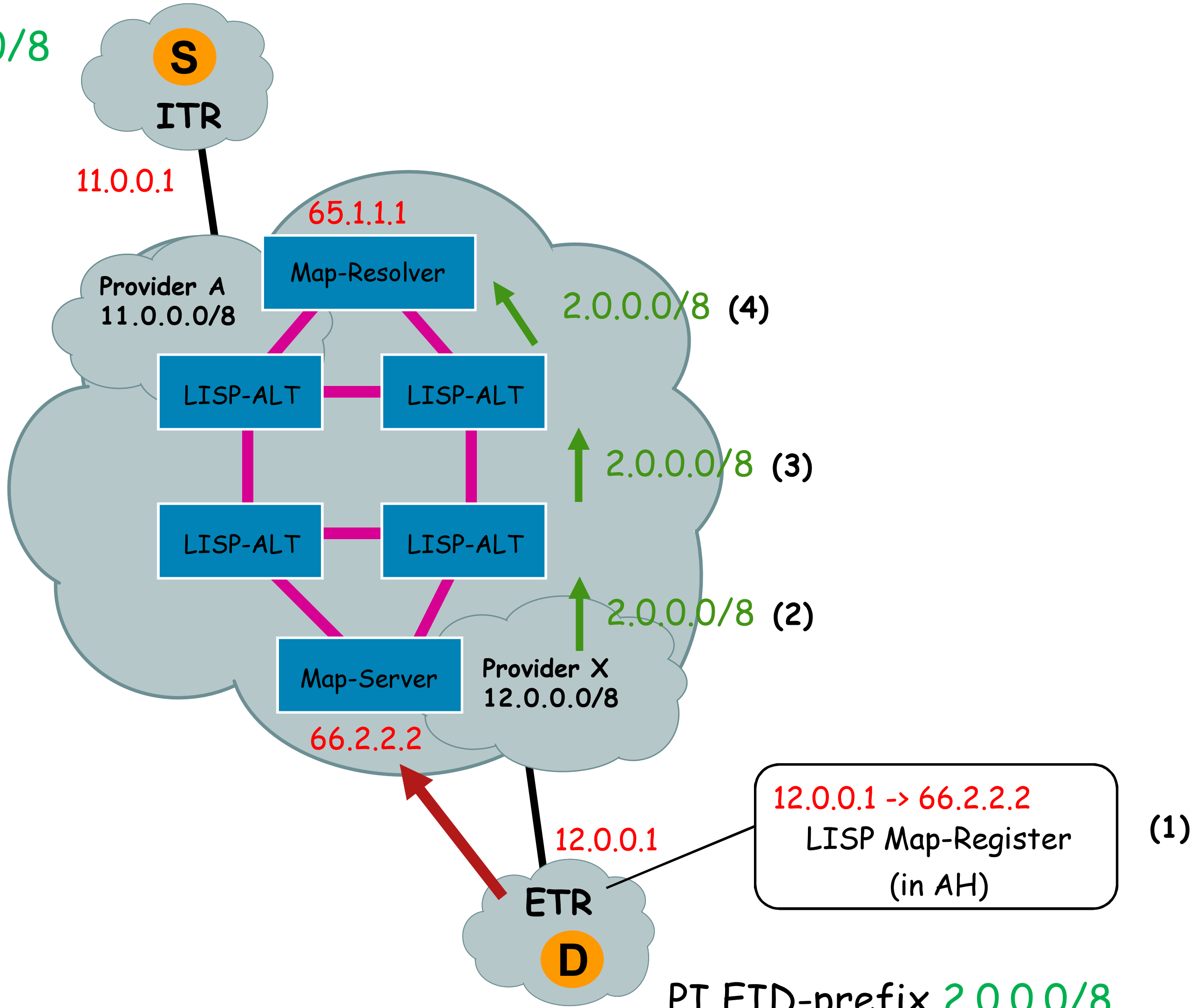
PI EID-prefix 2.0.0.0/8

Legend:
- EIDs -> Green
- Locators -> Red
- BGP-over-GRE
- Physical link

LISP Control Plane

[1] Map-Server Registration

PI EID-prefix 1.0.0.0/8



PI EID-prefix 2.0.0.0/8

Legend:

EIDs -> Green

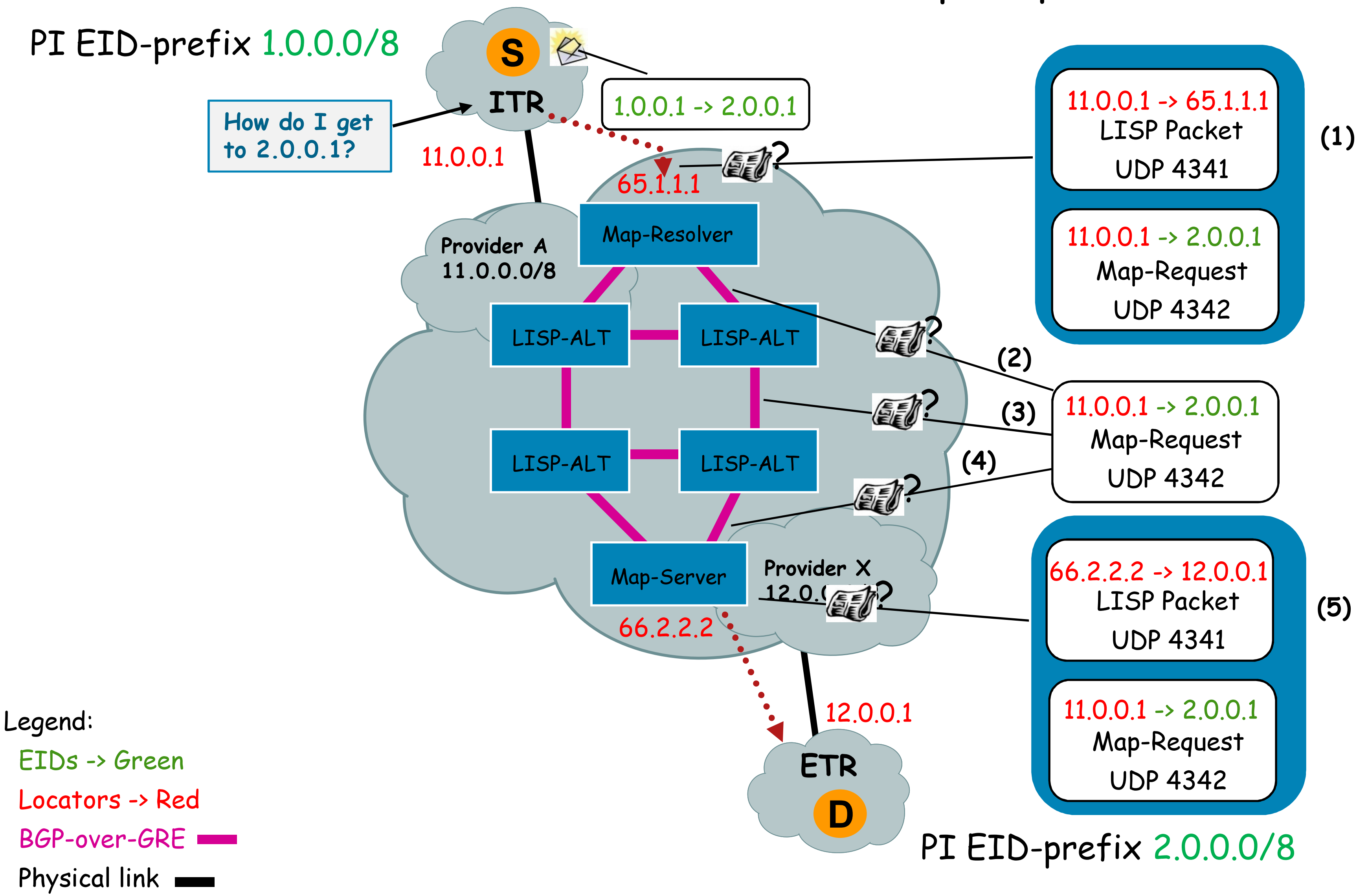
Locators -> Red

BGP-over-GRE

Physical link

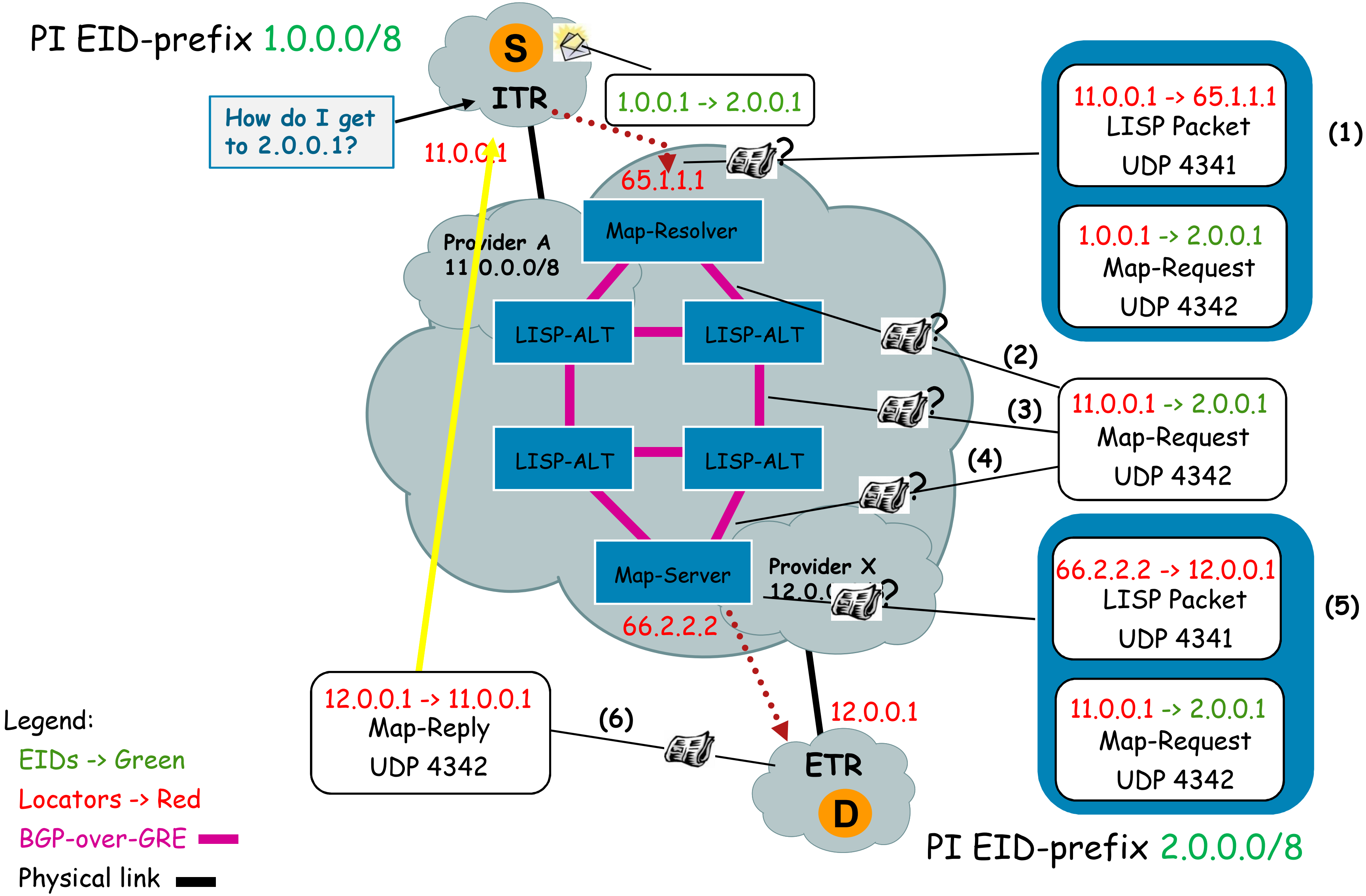
LISP Control Plane

[2] Data request Triggers Map-Request



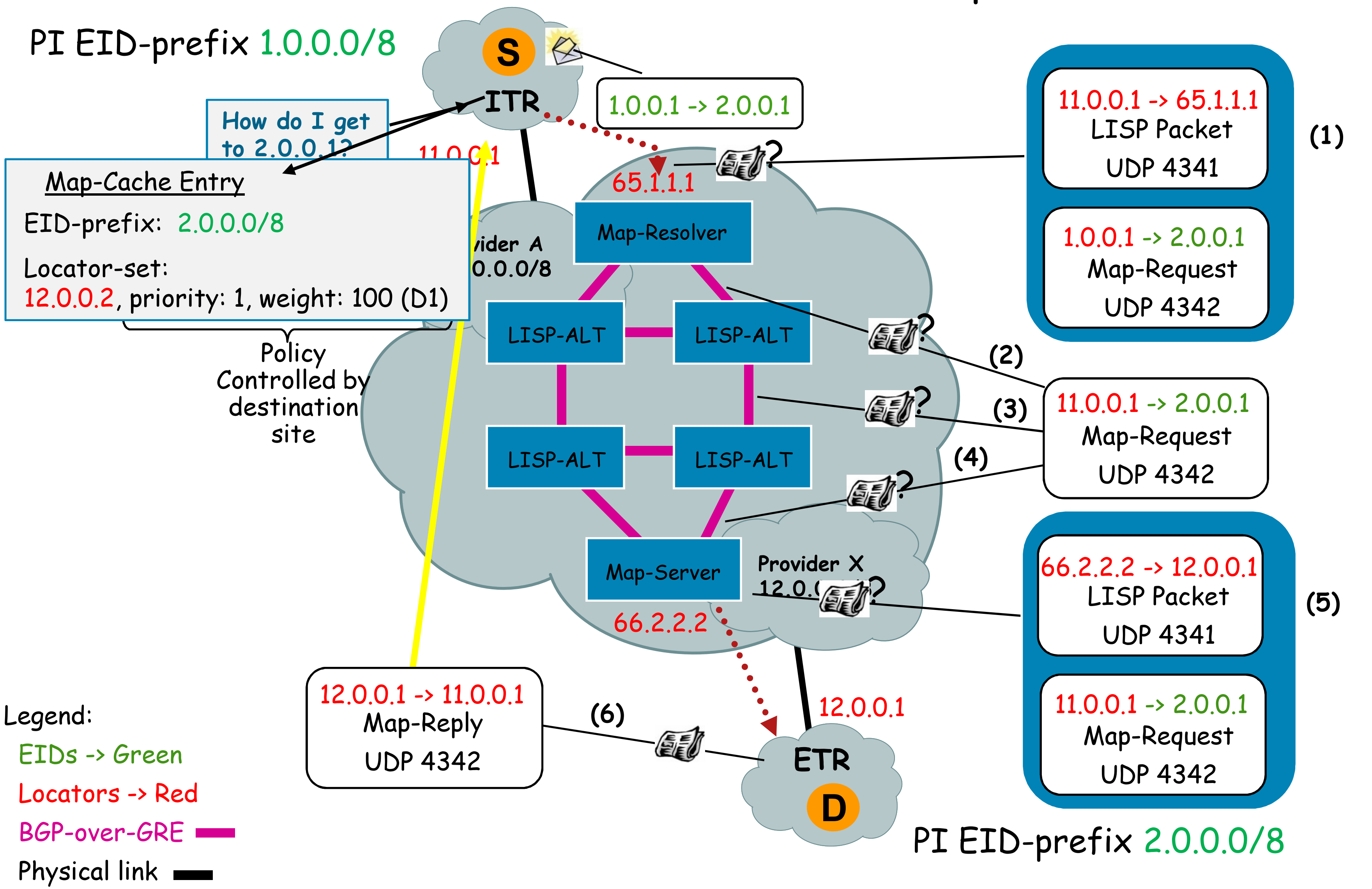
LISP Control Plane

[3] Map-Request Evokes Map-Reply



LISP Control Plane

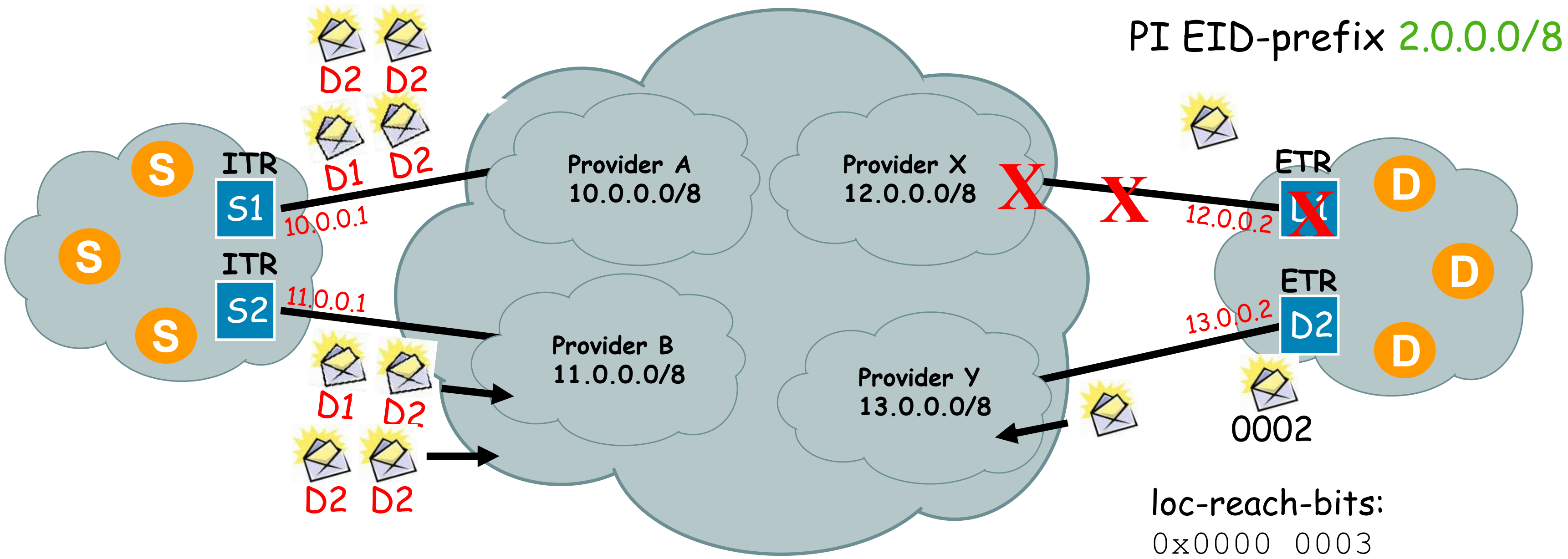
[4] Map-Cache Populated, data packets can flow



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:
  - Underlynig 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
- Use DPI heuristics?
- Use data-plane keepalives?
- 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)

13.0.0.2 priority: 1, weight: 50 (D2)

EIDs -> Green

Locators -> Red

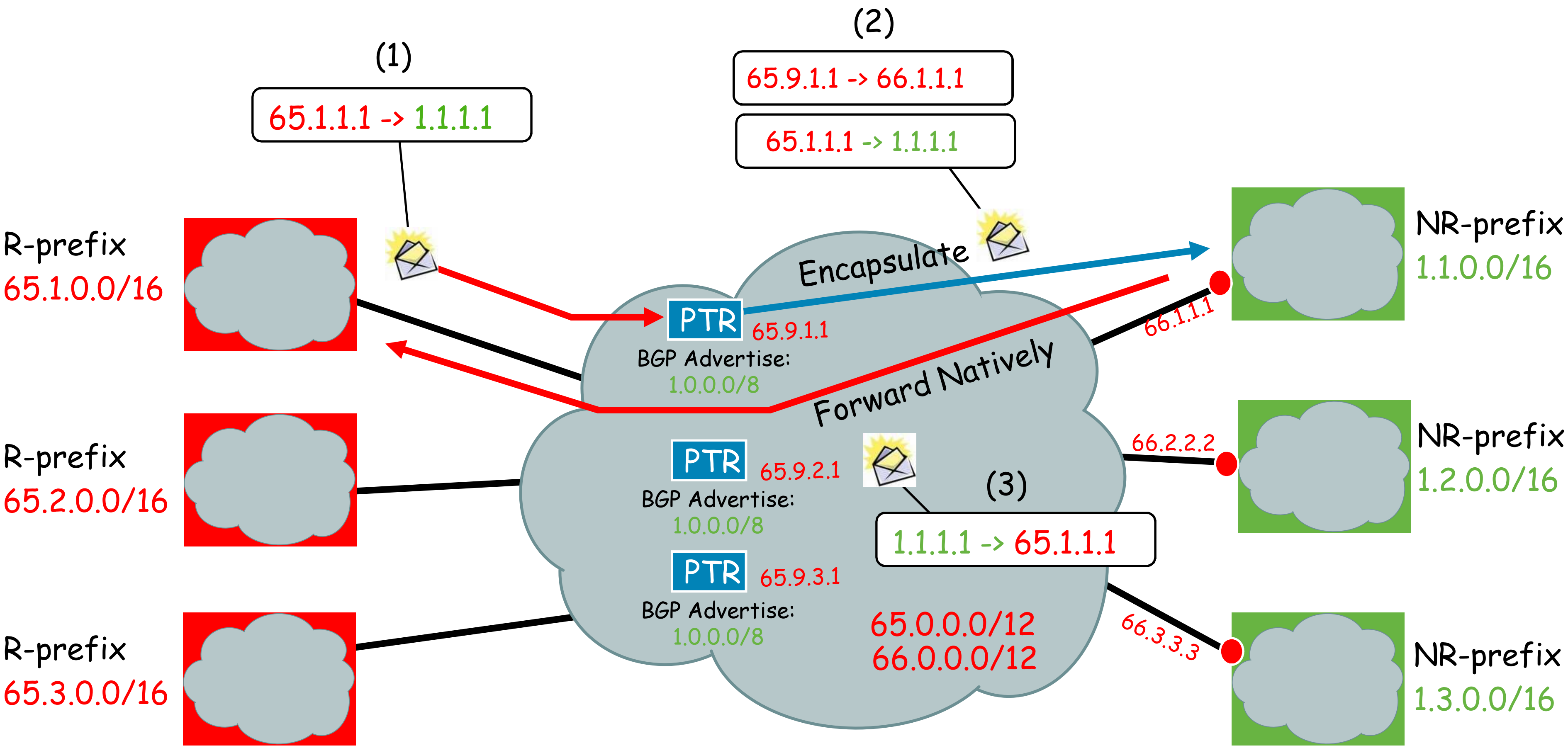
LISP Interworking

- 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
  - Infrastructure LISP network entity which receives packets from non-LISP sites and encaps to LISP sites or natively forwards to non-LISP sites
  - Creates a monetized service for infrastructure players

LISP Interworking

- Two important Interworking cases must be supported
  - **LISP site** to **non-LISP site**
  - **non-LISP site** to **LISP site**
- LISP Interworking allows LISP to be deployed incrementally
- PTRs allow LISP sites to see the benefits of ingress TE “day-one”

Interworking Using PTRs



Infrastructure Solution

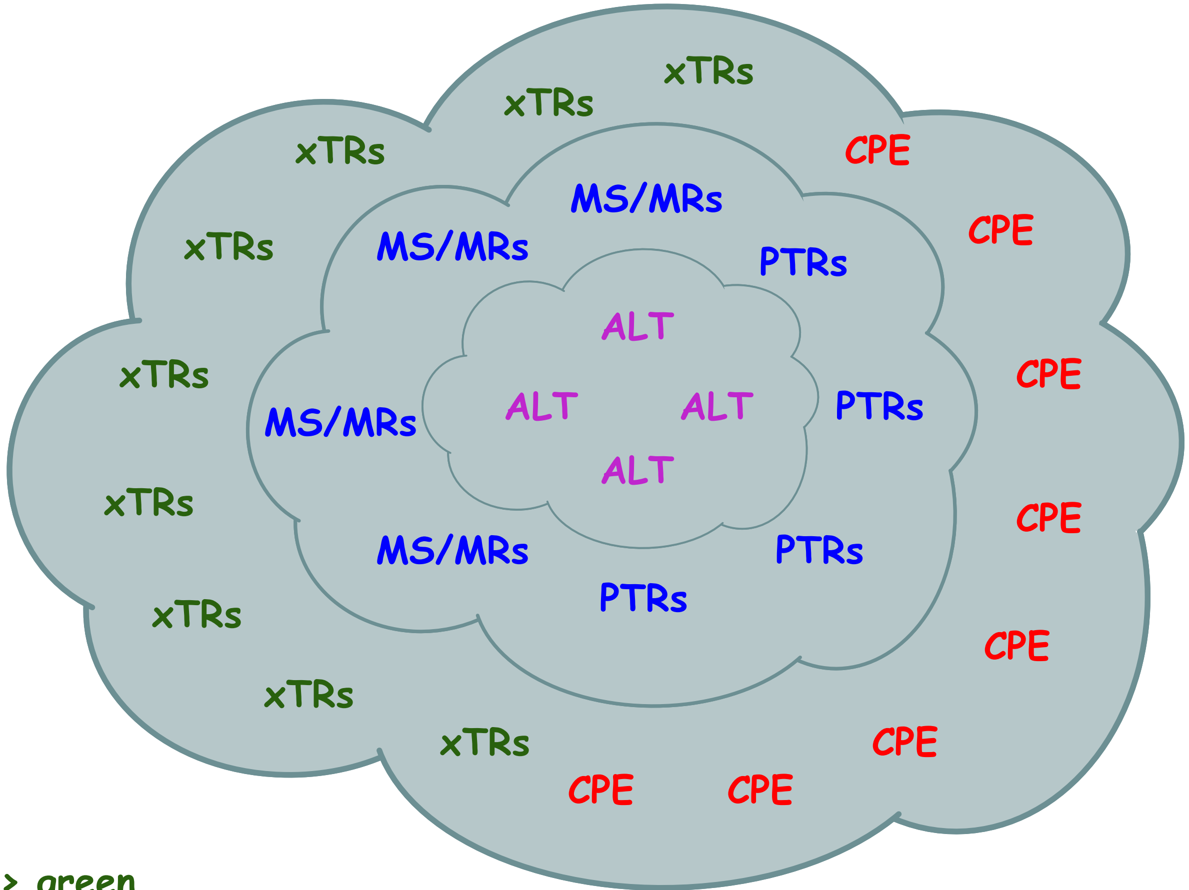
Legend:

LISP Sites -> Green (and EIDs)

non-LISP Sites -> Red (and RLOCs)

● xTR

The Whole Picture - LISP based Internet



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

Compelling Reasons for LISP (Summary)

LISP enables IP Number Portability

- With session survivability
- Hosts don't ever have to change IP addresses; No renumbering costs
- DNS "name -> EID" binding never changes

- LISP enables a "pull" vs "push" routing
  - OSPF and BGP are a push-models; routing stored in the forwarding plane
  - LISP is a pull-model; Analogous to DNS; massively scalable
- LISP is an "over-the-top" technology
  - Address Family agnostic
  - Incrementally deployable

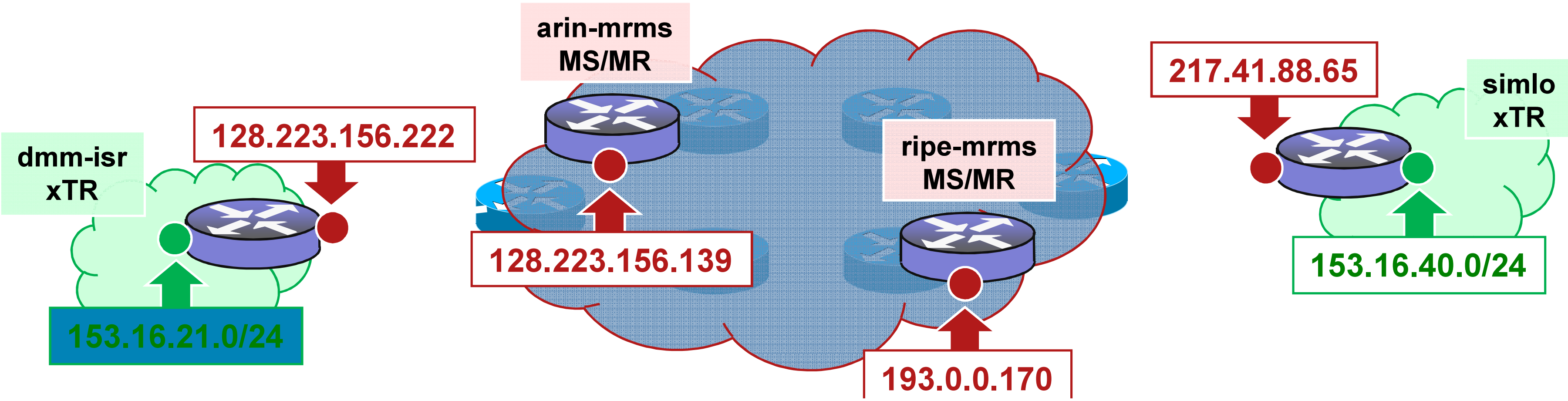
LISP enables:

- Improved OpEx of multi-homed sites by simplifying configuration overhead in comparison to BGP
- Improved utilization of upstream links by allowing for simple ingress traffic engineering
- Reduce or eliminate the need for renumbering when changing ISPs
- Control ISP expense associated with the ever growing default free zone prefix table size.

Some LISP Use Cases

1. Scales routing tables in Internet core
2. Supports low-opex site active-active multi-homing
3. Supports low-opex ISP active-active multi-homing
4. Avoids site renumbering with provider independence
5. Data Center mobility of Virtual Machines (VMs)
6. Data Center Server Load Balancing (SLBs)
7. A/V Truck Roll
8. L2 or L3 VPNs over Internet with or without parallelism
9. Hand-set mobility in localized regions
10. Better residential multi-homing
11. IPv6-only site connectivity over existing Internet
12. Movement/reallocation of Cloud Computing Resources

LISP Example
Configurations



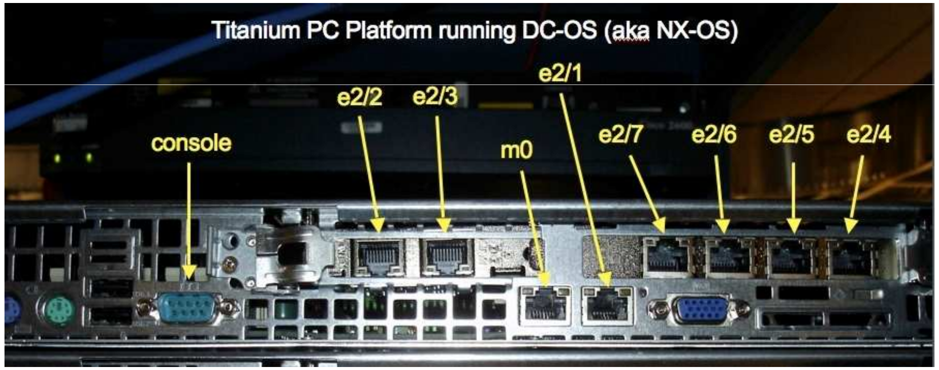

```
interface Loopback0
 ip address 153.16.21.1 255.255.255.255
!
interface FastEthernet0/0
 ip address 128.223.156.222 255.255.255.0
!
interface FastEthernet0/0/0
 ip address 153.16.21.17 255.255.255.240
!
ip lisp database-mapping 153.16.21.0/24 128.223.156.222 priority 1 weight 100
ip lisp itr map-resolver 128.223.156.139
ip lisp itr
ip lisp etr map-server 128.223.156.139 key 6 #%$^%##
ip lisp etr
!
ip route 0.0.0.0 0.0.0.0 128.223.156.1
!
```

LISP Pilot Deployment

- LISP Interworking Deployed
  - Have LISP 1-to-1 address translation working
    - `http://www.translate.lisp4.net`
  - Proxy Tunnel Router (PTR)
    - IPv4 PTRs: Andrew, ISC, and UY
    - IPv6 PTRs: Dave (UofO), ISC, and UY
    - `http://www.lisp6.net` reachable through IPv6 PTR
    - `http://www.ptr.lisp4.net` reachable through IPv4 PTR
- Go type into your browser now: `http://www.lisp4.net`
  - Web server in LISP site at University of Oregon
  - Demonstrates "LISP-Interworking" in action - you at non-LISP site talking to a LISP site
  - It's in green because it's an EID!

LISP Pilot Deployment

- LISP Pilot Network Operational
  - Deployed for nearly 2 years
    - More than 32 sites across 7 countries
    - US, UK, BE, JP, UY, AU, DE
  - Uses the NX-OS Titanium Platform
    - IOS and OpenLISP platforms to be added
  - EID-Prefixes:
    - IPv4 – `153.16.0.0/16`
    - IPv6 – `2610:00d0::/32`
  - RLOCs:
    - Current site attachment points to the Internet
  - Network is <u>Dual-Stack</u>
    - Can carry IPv4 and IPv6 Map-Requests



LISP Initiatives

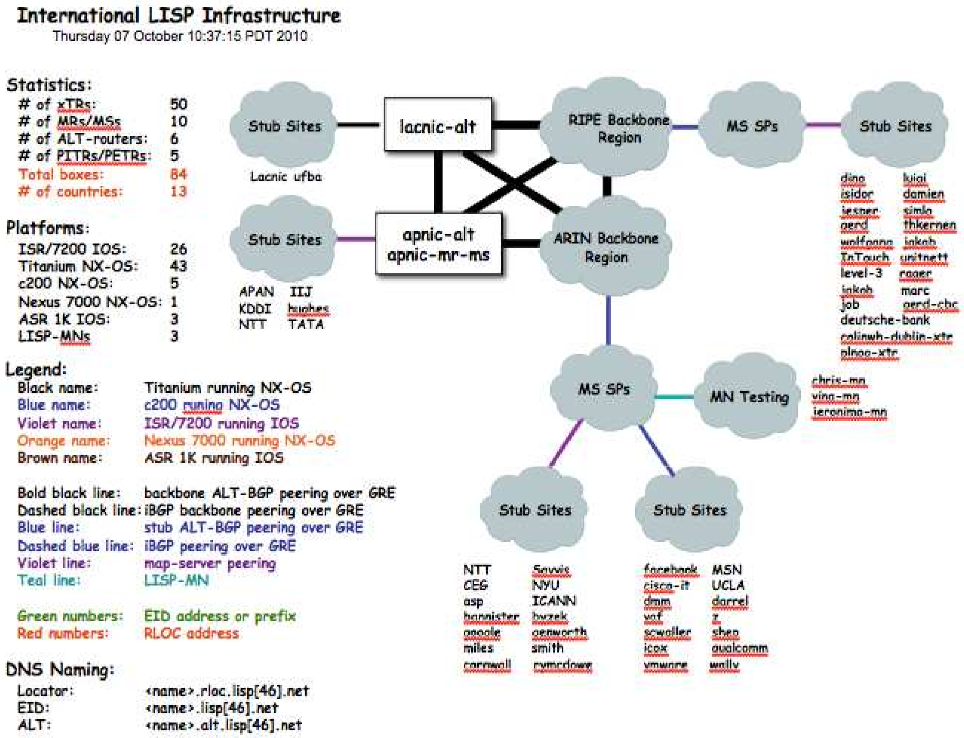
What's Cisco Doing in LISP?

- Cisco LISP Prototype Implementation
  - Started at Prague IETF, Mar 07; Deployed Pilot Network, July 07
  - Since then**, >220** releases of experimental code
- Cisco LISP Product Implementations
  - Phase 1 (December 24, 2009)
    - ISR, ISR-G2, 7200 (xTR)
  - Phase 2 (March 31, 2010)
    - ISR, ISR-G2, 7200 (xTR, PxTR, ALT) [**IOS 15.1(1)XB1**]
    - ASR 1000 (xTR, PxTR, ALT) [**IOS-XE 2.5.1**]
    - Nexus 7000 (xTR, PxTR, MS/MR) [**NX-OS 5.1(1.13)**]
    - UCS C200 (MS/MR) [**NX-OS 5.1(1.13)**]
  - Phase 3 (June 30, 2010)
    - **More LISP!**

Available Now!

- **External LISP Efforts**
  - FreeBSD OpenLISP
    http://gforge.info.ucl.ac.be/projects/openlisp/
  - Open Source LIG Diagnostic Tool
    http://www.github.com/davidmeyer/lig

LISP Initiatives
LISP Development Initiatives [2]



- Cisco-operated LISP Beta Network
  - >2.5 years operational
  - >60 sites in 10 countries
  - Built for experimentation
  - and Proof-of-Concept testing
- LISP Interworking
  - Proxy Ingress Tunnel Router (PITR)
    - – IPv4 and IPv6 P-ITRs deployed
    - – http://www.lisp4.net, http://www.lisp6.net (Univ of Oregon)
    - – http://www.lisp4.facebook.com (Facebook)

References

- **Locator/ID Separation Protocol (LISP) - draft-ietf-lisp-09; 11-Oct-2010.**
  http://tools.ietf.org/html/draft-ietf-lisp-09
- **LISP Map Server - draft-ietf-lisp-ms-06.txt; 18-Oct-2010.**
  http://tools.ietf.org/html/draft-ietf-lisp-ms-06
- **LISP ALT - draft-ietf-lisp-alt-04; 26-April-2010.**
  http://tools.ietf.org/html/draft-ietf-lisp-alt-04
- **LISP Interworking - draft-ietf-lisp-interworking-01; 26-April-2010.**
  http://tools.ietf.org/html/draft-ietf-lisp-interworking-00
- **LISP Multicast - draft-ietf-lisp-multicast-04; 11-October-2010.**
  http://tools.ietf.org/html/draft-ietf-lisp-multicast-04
- **LISP Mobility Architecture - draft-meyer-lisp-mn-05; 25-October-2010.**
  http://tools.ietf.org/html/draft-meyer-lisp-mn-05
- **LISP Internet Groper (LIG) specification - draft-ietf-lisp-lig-01; 11-October-2010.**
  http://tools.ietf.org/html/draft-ietf-lisp-lig-01
- **The Locator/ID split, its implications for the IP Architecture, and a few current approaches," D. Meyer, APRICOT 2007.** http://www.1-4-5.net/~dmm/talks/apricot2007/locid
- **"Report from the IAB Workshop on Routing and Addressing," D. Meyer, L. Zhang, K. Fall (editors).** http://www.ietf.org/internet-drafts/draft-iab-raws-report-00.txt
- **"Projecting Future IPv4 Router Requirements from Trends in Dynamic BGP Behavior," G. Huston, G. Armitage.** http://www.potaroo.net/papers/phd/atnac-2006/bgp-atnac2006.pdf
- **"BGP in 2008," G. Huston.** http://www.potaroo.net/presentations/2009-05-06-bgp2008.pdf