LISP: What and Why

RIPE Berlin – May, 2008

Vince Fuller
(for Dino, Dave, Darrel, et al)

http://www.vaf.net/prezos/lisp-ripe-long.pdf

Agenda

- What is the problem?
- What is LISP?
- Why Locator/ID Separation?
- Data Plane Operation
- Finding Mappings – LISP+ALT
- Incremental Deployability
- Implementation and Testing status

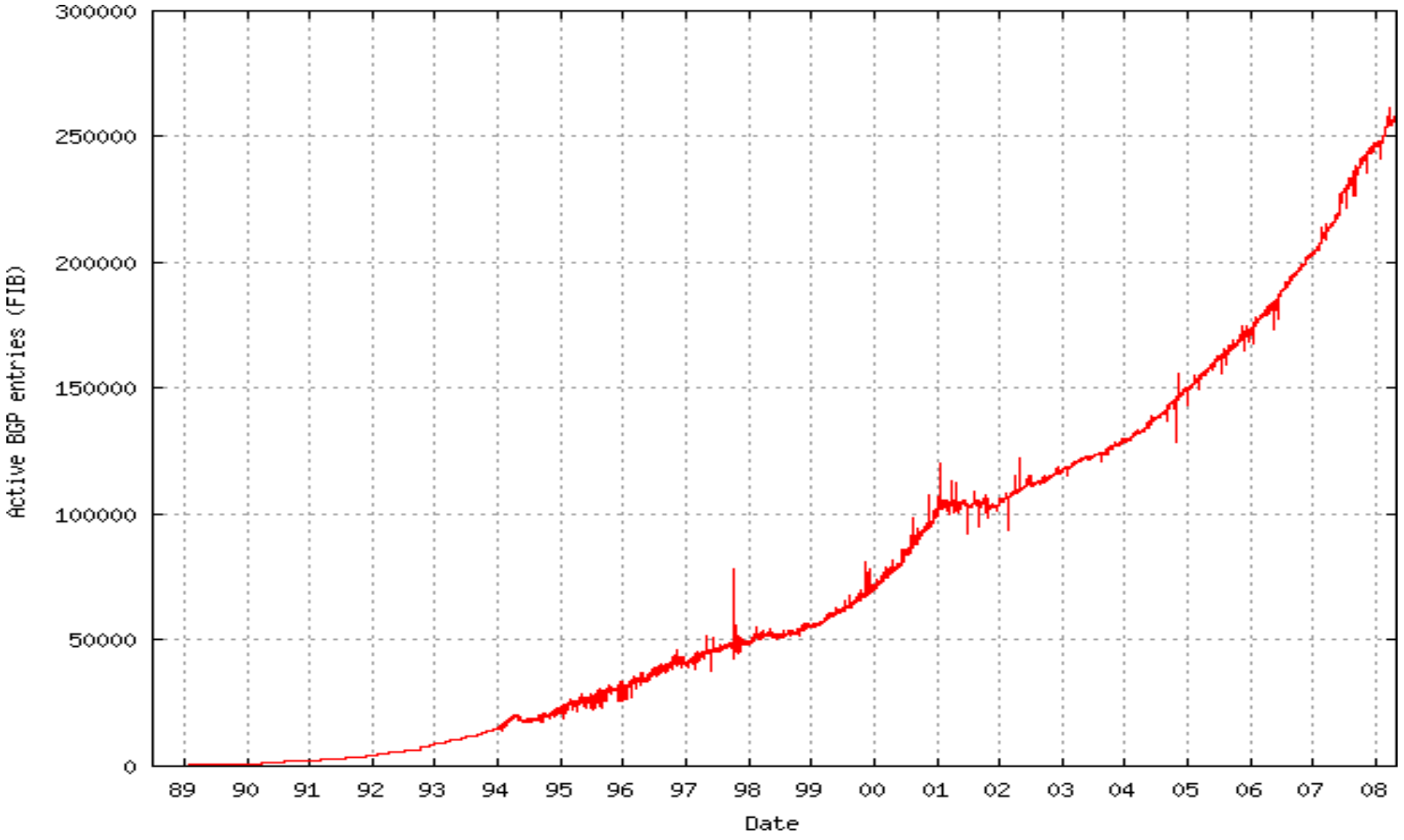
What is LISP?

- 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
  - Address family agnostic

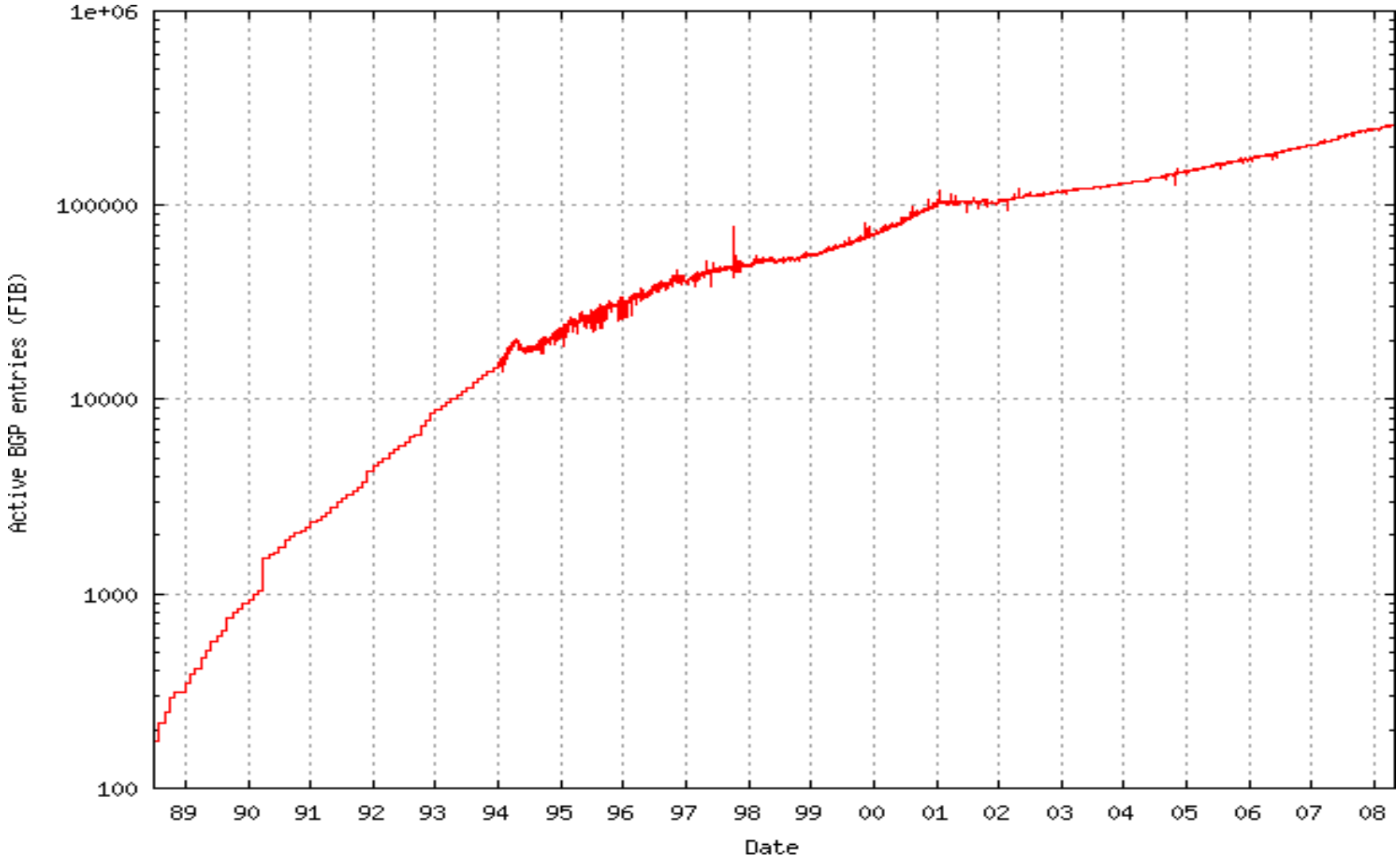
Problem statement

- There are reasons to believe that current trends in the growth of routing and addressing state on the global Internet may cause difficulty in the long term
- The Internet needs an easier, more scalable mechanism for multi-homing with traffic engineering
- An Internet-wide replacement of IPv4 with ipv6 represents a one-in-a-generation opportunity to either continue current trends or to deploy something truly innovative and sustainable
- As currently specified, routing and addressing with ipv6 is not significantly different than with IPv4 – it shares many of the same properties and scaling characteristics
- **More at: www.vaf.net/prezos/rrg-prague.pdf**

Scaling of Internet routing state



Scaling of state - Log plot



What is ID/Loc separation?

- Instead of IP addresses, two numbering spaces:
- Endpoint Identifiers (EIDs): hierarchically assigned to sites along administrative lines (like DNS hostnames)
  - do not change on devices that remain associated with the site; think "PI" but not routable
- Routing Locators (RLOCs): assigned according to network topology, like "PA" address assignments
  - Locators are aggregated/abstracted at topological boundaries to keep routing state scalable
  - When site's connection to network topology changes, so do the locators – aggregation is preserved

What Provoked This?

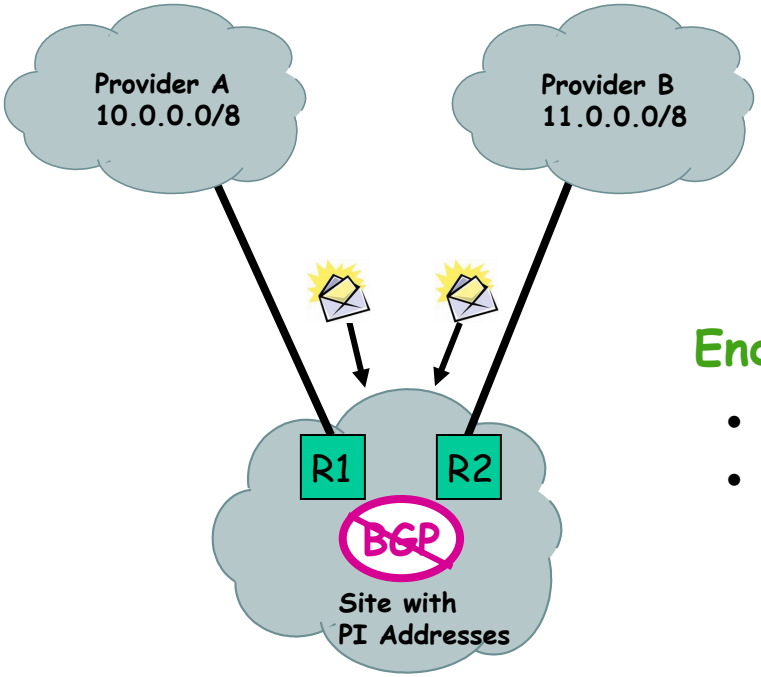
- Stimulated by problem statement effort at the Amsterdam IAB Routing Workshop on October 18/19 2006
  - Report published as RFC 4984
- More info on problem statement:

  www.vaf.net/prezos/rrg-prague.pdf

Why the Separation?

- The 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 for hosts and routers at a site
- You can change Locators
  - Now sites can change providers
  - Now hosts can move

What Features do I get?



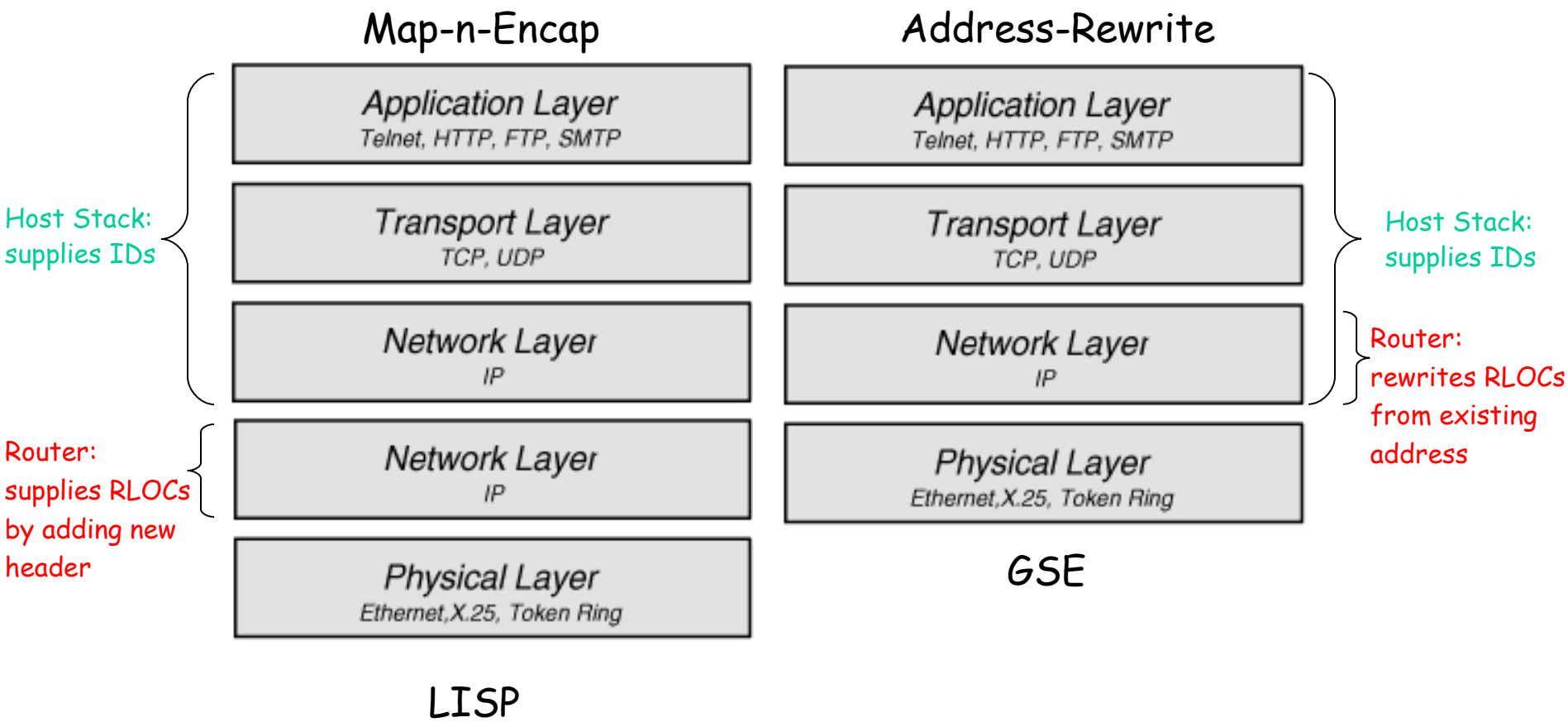
Lower OpEx for Sites and Providers

- Improve site multi-homing
- Improve provider traffic engineering
- Reduce size of core routing tables

End Site Benefit

- Easier Transition to ipv6
- Change provider without address change

Map-n-Encap vs Address-Rewrite



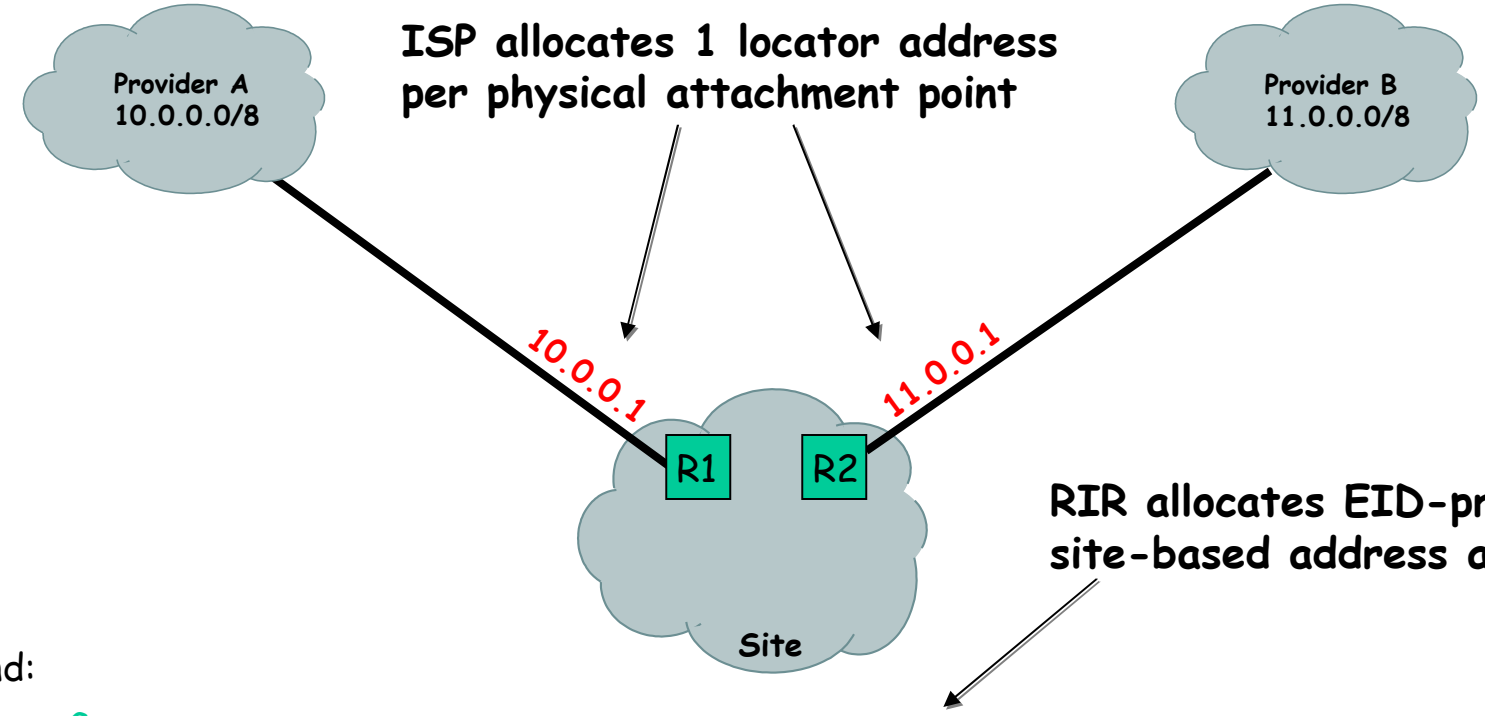
What is LISP?

- Data plane
  - Design for encapsulation and tunnel router placement
  - Design for locator reachability
  - Data-triggered mapping service
- Control plane
  - Design for a scalable mapping service
  - Examples are: CONS, NERD, ALT, EMACS

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
    - routable in site/local scope, so not "pure" EIDs
  - New namespace - essentially invisible to core routing/forwarding
- 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 routing/forwarding uses today

Different Address Allocation Authorities



Legend:

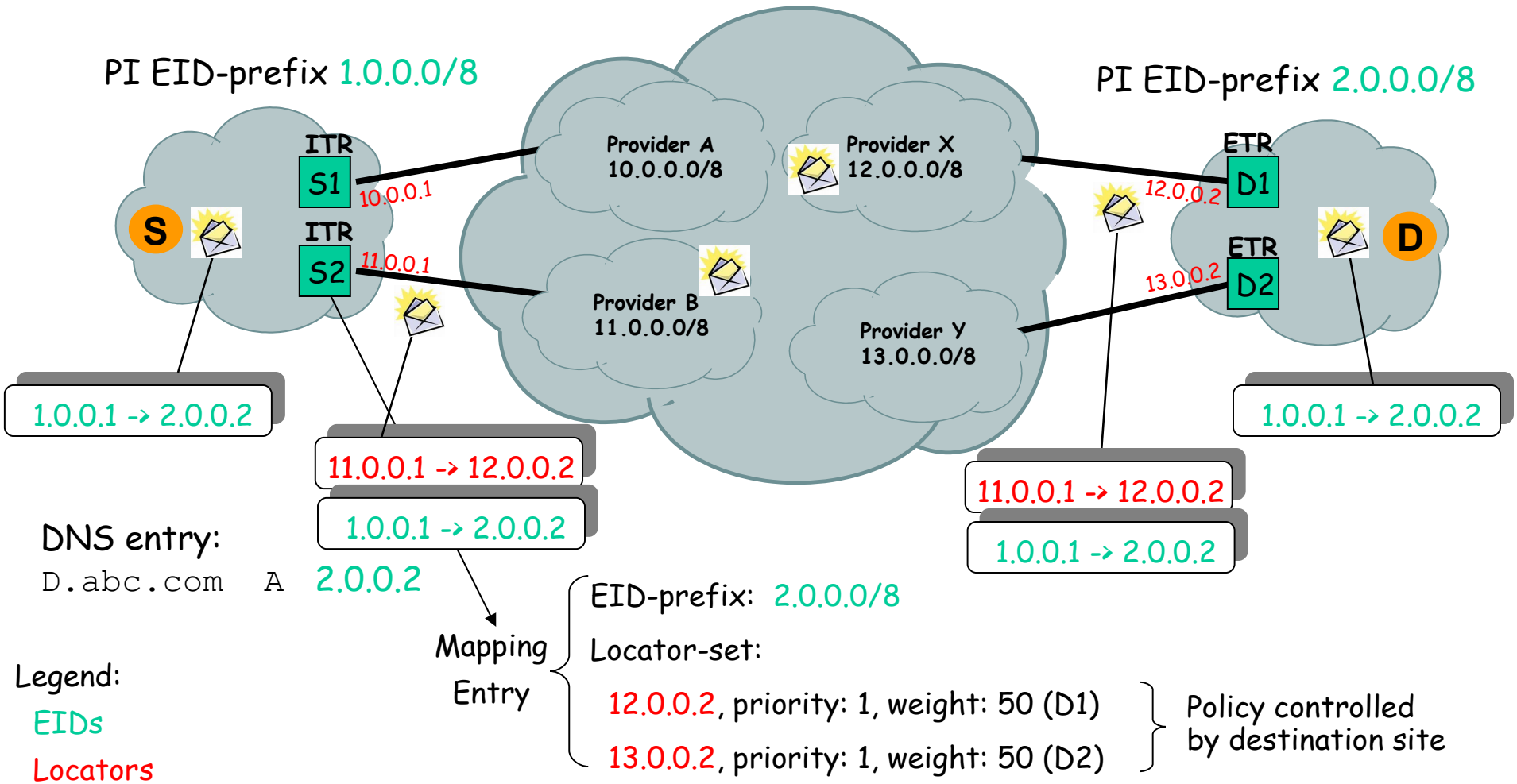
EIDs -> Green

Locators -> Red

New Network Elements

- Ingress Tunnel Router (ITR)
  - Finds EID to RLOC mapping
  - Encapsulates to Locators at source site
- Egress Tunnel Router (ETR)
  - Owns EID to RLOC mapping
  - Decapsulates at destination site

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:

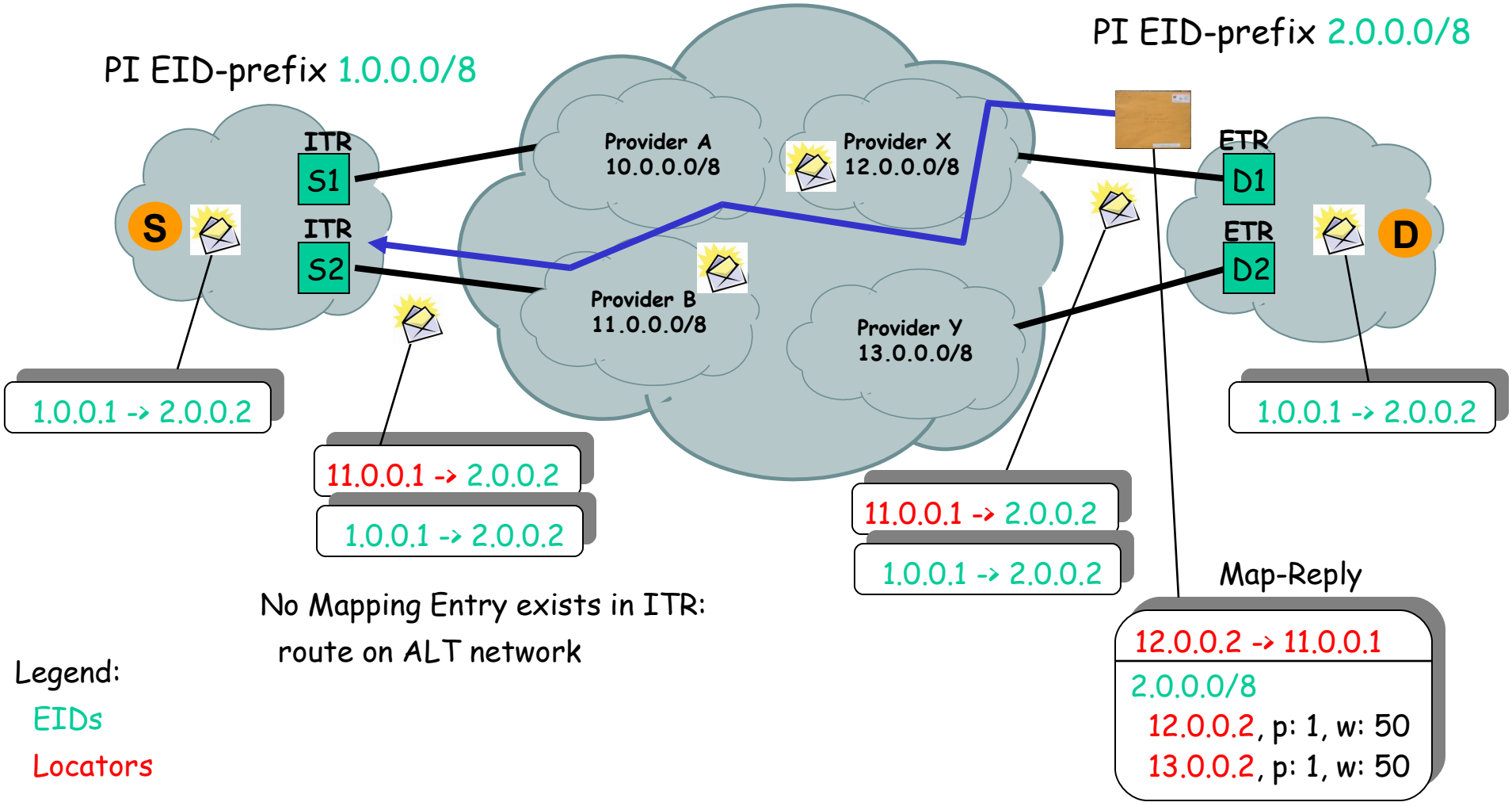
EIDs

Locators

When the ITR has no Mapping

- ITR needs to obtain from ETR
- ITR sends Map Request (or Data Probe)
- ETR returns Map Reply

Map Request & Reply



Finding the ETR for an EID

- Need a scalable EID to Locator mapping lookup mechanism
- Network based solutions
  - Have query/reply latency
  - Can have packet loss characteristics
  - Or, have a full table like BGP does
- How does one design a scalable Mapping Service?

Mapping Service

- Build a large distributed mapping database service
- Scalability paramount to solution
- How to scale:

  `(state * rate)`

- If both factors large, we have a problem
  - `state` will be $O(10^{10})$ hosts
    - Aggregate EIDs into EID-prefixes to reduce state
  - `rate` must be small
    - Damp locator reachability status and locator-set changes
    - Each mapping system design does it differently

Mapping Service Designs

- DNS – considered, many issues
- DHTs – considered, research pending
- CONS – new protocol, hybrid push+pull
  - Push EID-prefixes at top levels of hierarchy
  - Pull mappings from lower levels of hierarchy
- ALT – GRE/BGP based, **current focus**
- EMACS – like ALT, but multicast-based
- NERD – pure Push design

LISP+ALT Design Goals

- Use as much technology as reasonable
  - Use what works and no more
- Minimal memory impact on ITRs
- Provide data path to reduce latency
- Allow infrastructure players to achieve new revenue source

LISP+ALT: What and How

- Hybrid push/pull approach
  - ALT pushes aggregates, LISP pulls specifics
- Hierarchical EID prefix assignment
- Aggregation of EID prefixes
- GRE-based overlay network
- BGP used to advertise EIDs on overlay
- Option for data-triggered Map-Replies

LISP-ALT Routers and the LAT

- LISP+ALT routers form "Alternative Logical Topology" (ALT)
  - Interconnected by tunnels (GRE or …)
  - BGP used for EID prefix propagation
  - Logical hierarchy
- ITRs and ETRs connect at "edge"
- Who runs LISP+ALT routers?
  - ISPs, IXCs, RIRs, Neutral parties?

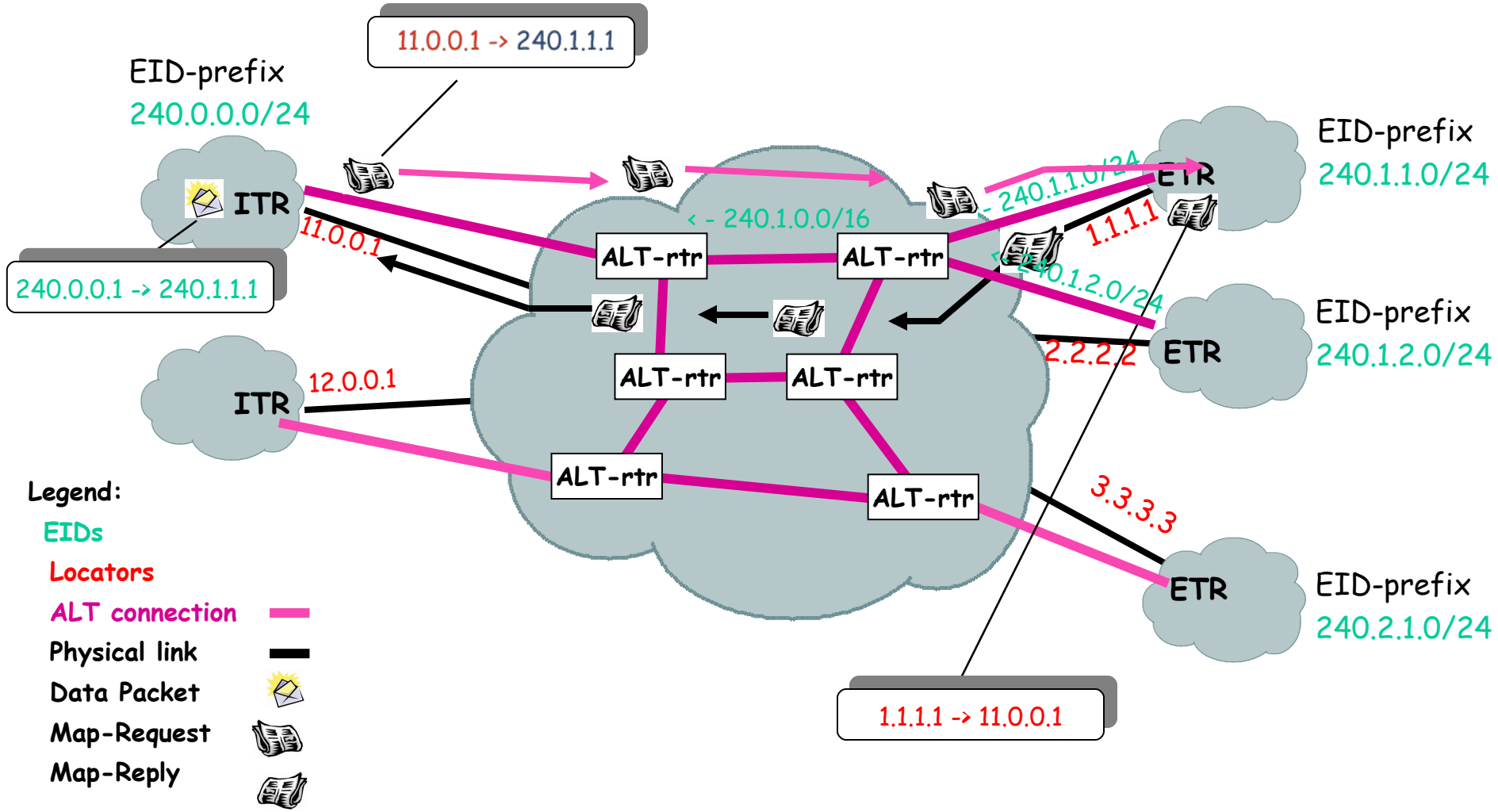
Tunnel and BGP Operation

- EID prefixes originated into BGP at edge
  - by ETRs or LISP+ALT routers on behalf of site
- ITRs learn EID prefixes via BGP from LISP+ALT routers or use "default"
  - Map-Requests are forwarded into the ALT via first-hop LISP+ALT router(s)
  - ALT forwards Map-Request to "owning" ETR for EID prefix
- LISP+ALT routers aggregate prefixes "upward" in the alternative topology

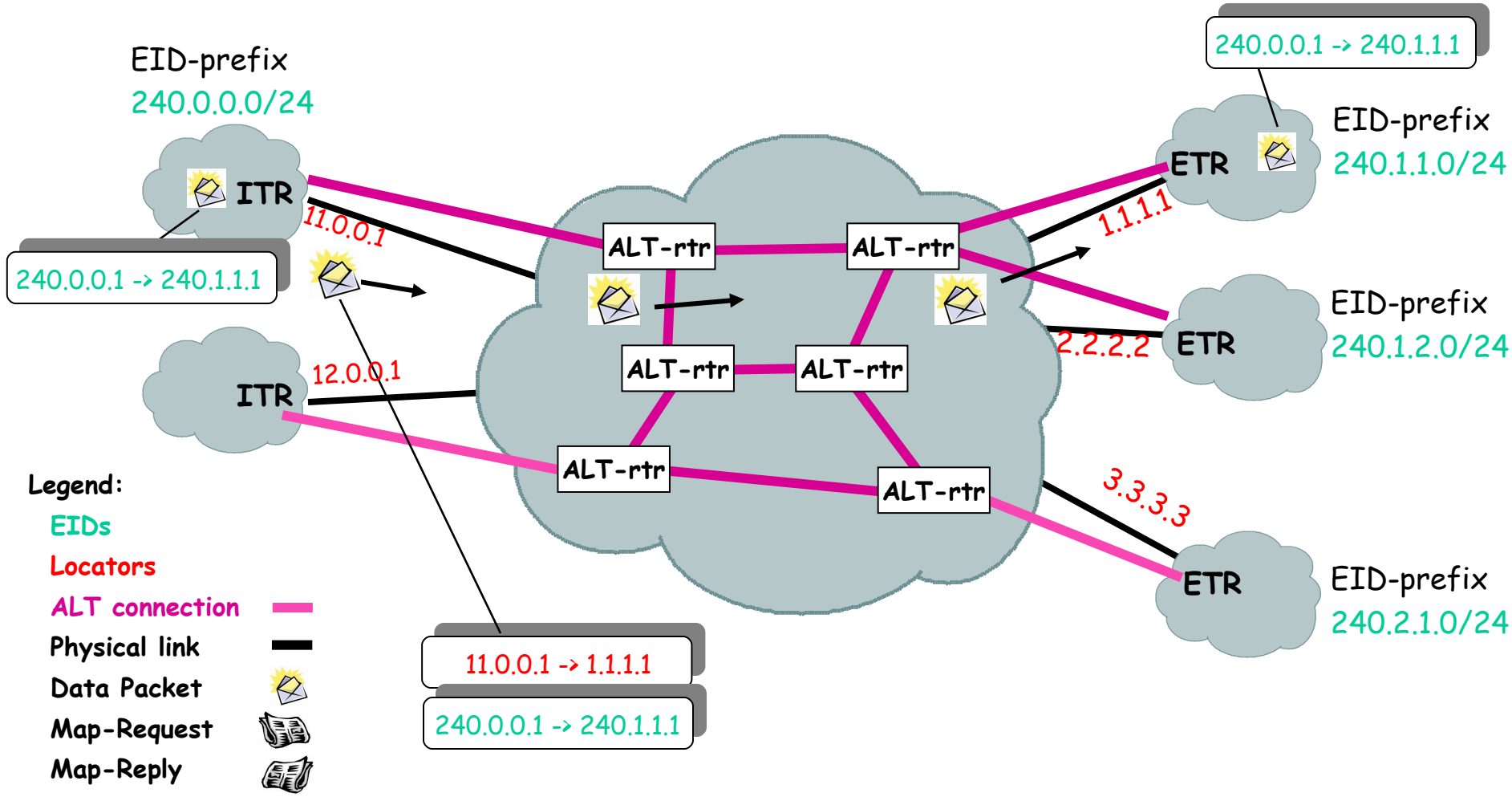
Data-Triggered Mappings

- ITRs have the option of forwarding data for "un-mapped" EIDs into ALT
- Data forwarded across ALT to ETR that originates the EID prefix
- LISP Map-Reply "triggered" from ETR to ITR, installed in ITR cache
- Following traffic uses cached RLOCs
- Scaling/performance issues

LISP+ALT in action



LISP+ALT in action



Interworking Deployability

- 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

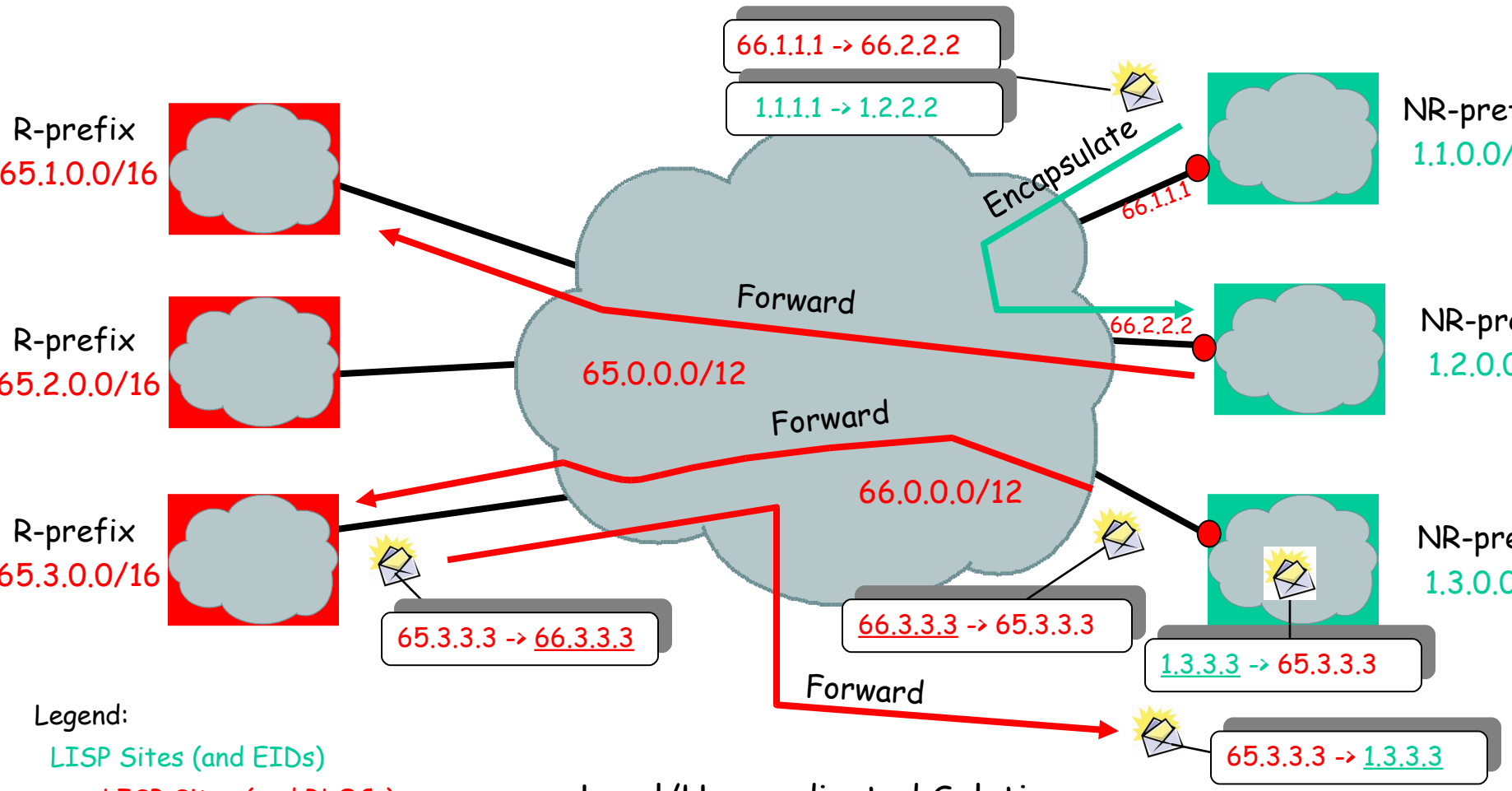
Interworking Deployability

- LISP-R site to non-LISP site
  - ITR at LISP site detects non-LISP site when no mapping exists
    - Does not encapsulate packets
  - Return packets to LISP site come back natively since EIDs are routable
  - Same behavior as the non-LISP to non-LISP case
    - LISP site acts as a non-LISP site

Interworking Deployability

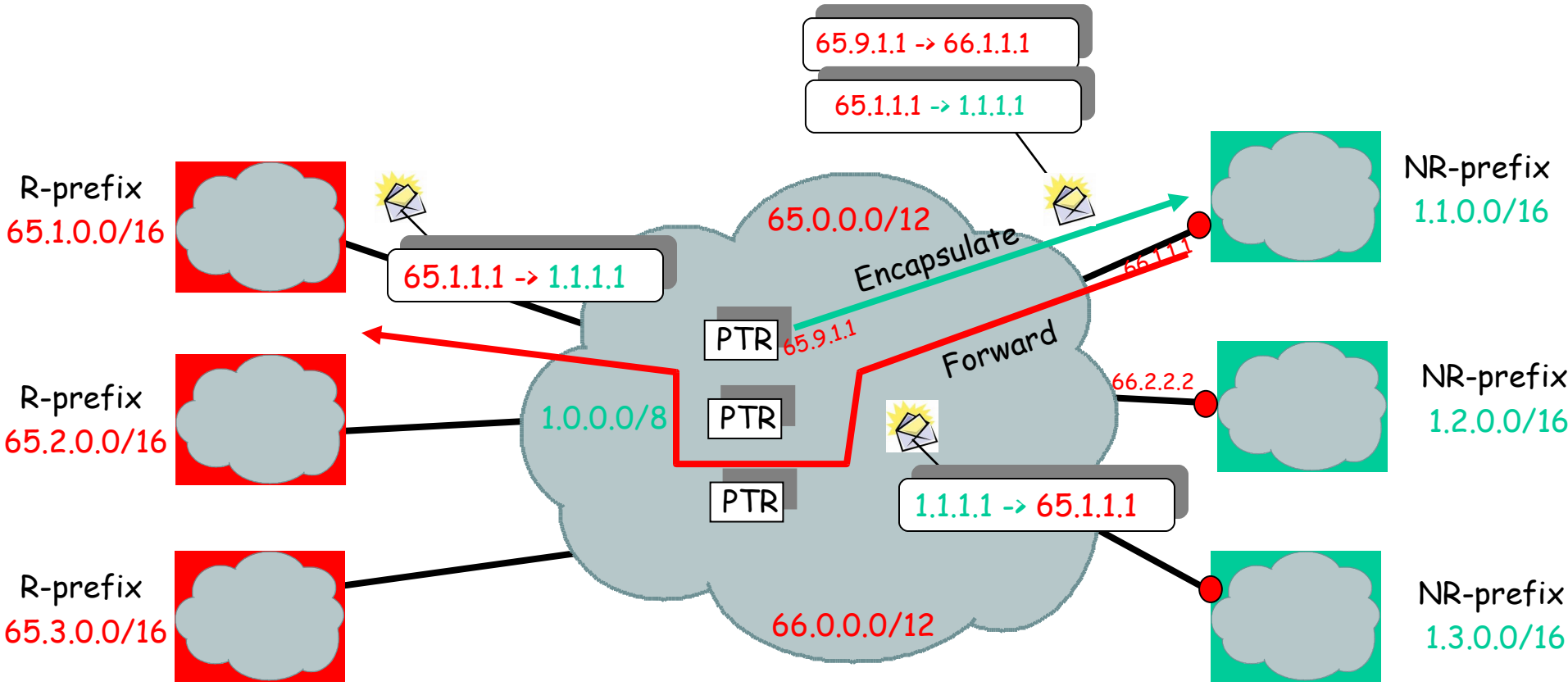
- LISP-NR site to a non-LISP site
  - ITR at LISP site detects non-LISP site when no mapping exists
    - Does not encapsulate packets
  - For return packets to LISP site
    - ITR translates to a source routable address so packets symmetrically sent natively
    - PTR advertises NR prefixes close to non-LISP sites so return packets are encapsulated to ETR at LISP site

Interworking 1



Local/Uncoordinated Solution

Interworking: 2



Legend:

LISP Sites (and EIDs)

non-LISP Sites (and RLOCs)

xTR

Infrastructure Solution

Implementation Status

- cisco has a LISP prototype implementation
  - Started the week of IETF Prague (March 2007)
- OS platform is DC-OS
  - Linux underlying OS
- Hardware platform is Titanium
  - 1 RU dual-core off-the-shelf PC with 7 GEs
- Based on `draft-farinacci-lisp-07.txt`
- Software switching only
- Supports both IPv4 and ipv6

Implementation Status

- IOS 12.4T prototype is in the works
- OpenLISP implementation

  draft-iannone-openlisp-implementation-00.txt
- Would really like to see more

Prototype Functionality

- Supports ITR encap and ETR decap
  - Load-balancing among locators
  - Respects priority & weight per mapping
- Multiple EID-prefixes per site
- Support for locator reachability
- Multi-VRF support for BGP-over-GRE

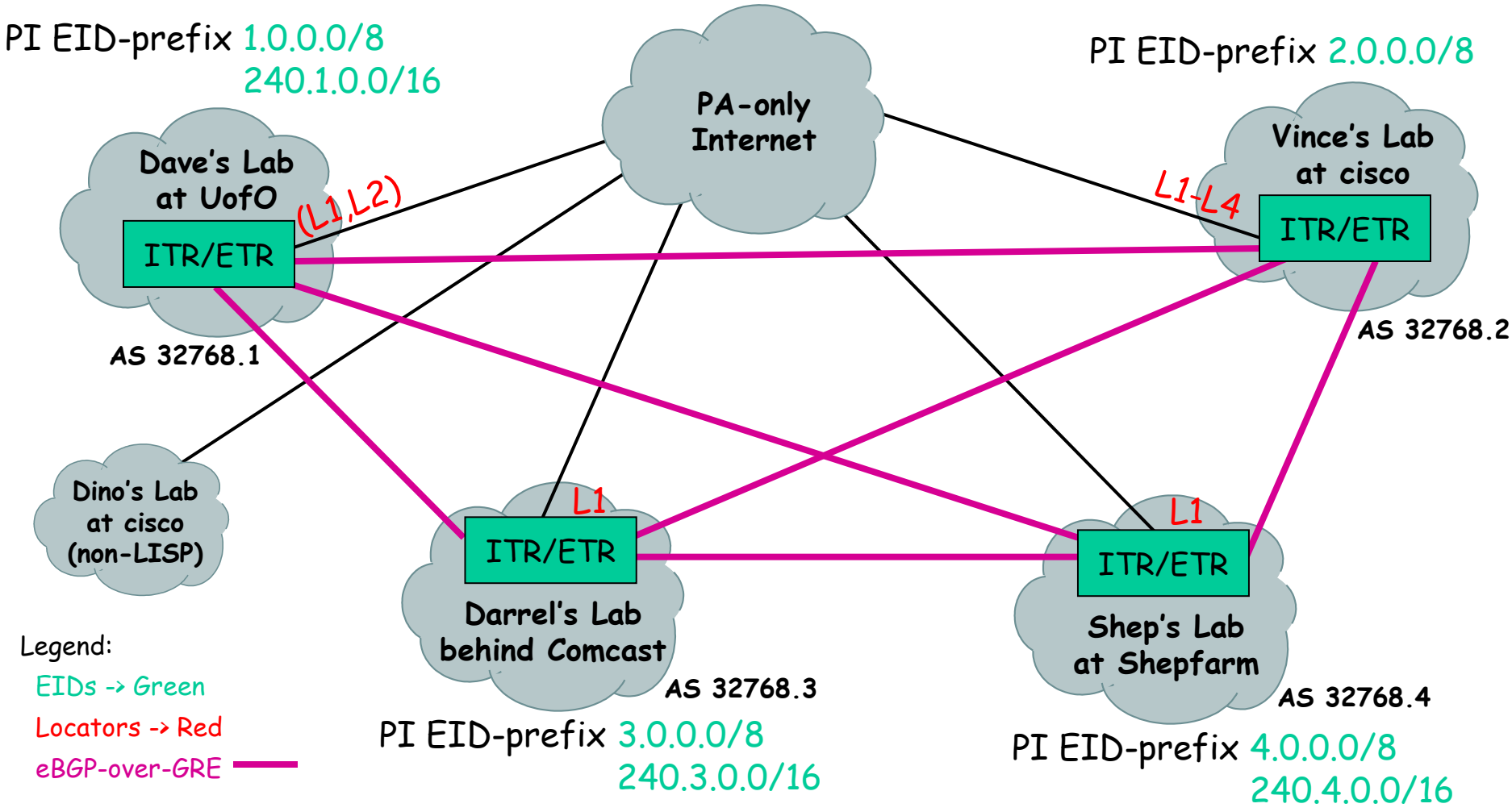
Prototype Functionality

- 240/4 support
  - To use as EIDs
- 'glean-mapping' support
  - And route-returnability check for verifying when an EID has moved to a new ITR
- LISP+ALT support
  - BGP advertises EID-prefixes over GRE tunnels
  - Data Probes sent over GRE topology
  - Map-Replies returned over GRE topology
- Interoperability – PTR and NAT functionality

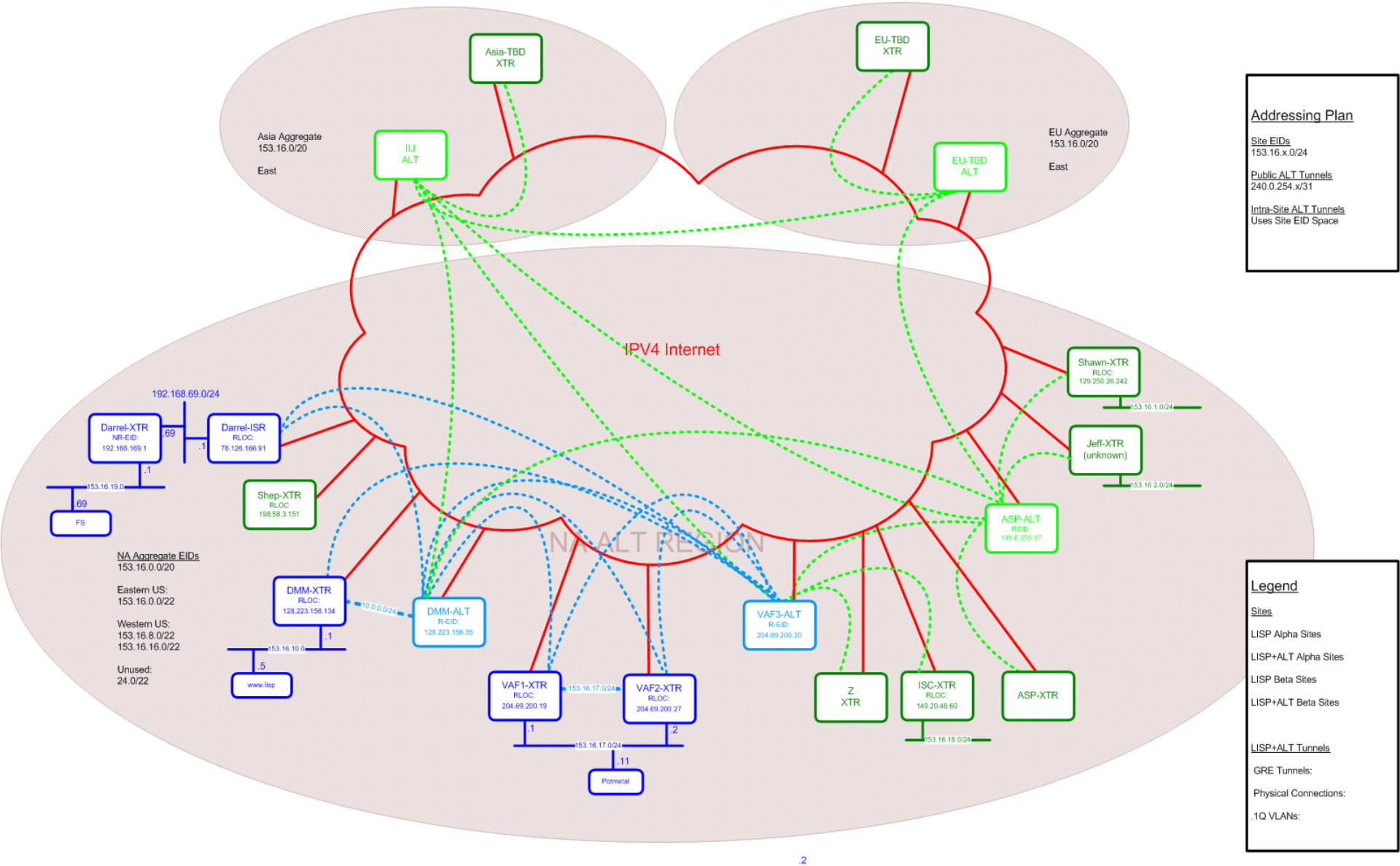
Prototype Testing

- Detailed Test Plan written and being executed against
- Multiple EID-prefix testing completed
- Multiple locator testing completed
- LISP+ALT testing underway
- Surprise: found a few bugs ☺

LISP Alpha topology



LISP and LISP+ALT Test Topology



Addressing Plan

Site EIDs
153.16.x.0/24

Public ALT Tunnels
240.0.254.x/31

Intra-Site ALT Tunnels
Uses Site EID Space

Legend

Sites

LISP Alpha Sites

LISP+ALT Alpha Sites

LISP Beta Sites

LISP+ALT Beta Sites

LISP+ALT Tunnels

GRE Tunnels:

Physical Connections:

.1Q VLANs:

.2

What's Next for Prototype and Testing

- Experiment with re-encapsulating and recursive ITRs
- Analysis of ALT data-probe latency
  - Reason for inter-continental test topology
- More testing on map entry changing
- Think more about security mechanisms

What's Next for Prototype and Testing

- Think more and experiment with movement
- Think more about aggregation and anti-entropy models
- Implement Address-Family crossover support
  - ipv6 EIDs over IPv4 Locators

Wanna play with us?

- Looking for more external test sites
  - Particular need in European region
  - Must be able to dedicate minimum of 1 day a week
- Goals:
  - Test multiple implementations
  - Experience with operational practices
  - Learn about revenue making opportunities

Wanna Play With Us?

- It's been > one year since the IAB RAWS
  - Some of us committed to working in the IETF and IRTF in an open environment
- This is not a Cisco only effort
  - We have approached and recruited others
  - There are no patents (cisco has no IPR on this)
  - All documents are Internet Drafts
- We need designers
- We need implementers
- We need testers
- We need research analysis
- We want this to be an open effort!
- Contact us: lispers@cisco.com
- See also: lisp-interest@lists.civil-tongue.net

Internet Drafts

draft-farinacci-lisp-07.txt

draft-fuller-lisp-alt-02.txt

draft-lewis-lisp-interworking-00.txt

draft-farinacci-lisp-multicast-00.txt


LISP
RULES