Labeled ARP

Kireeti Kompella
Balaji Rajagopalan
IETF 89

Acknowledgments:
Shane Amante
Thomas Morin
Luyuan Fang
The Juniper “MPLS-in-DC” team

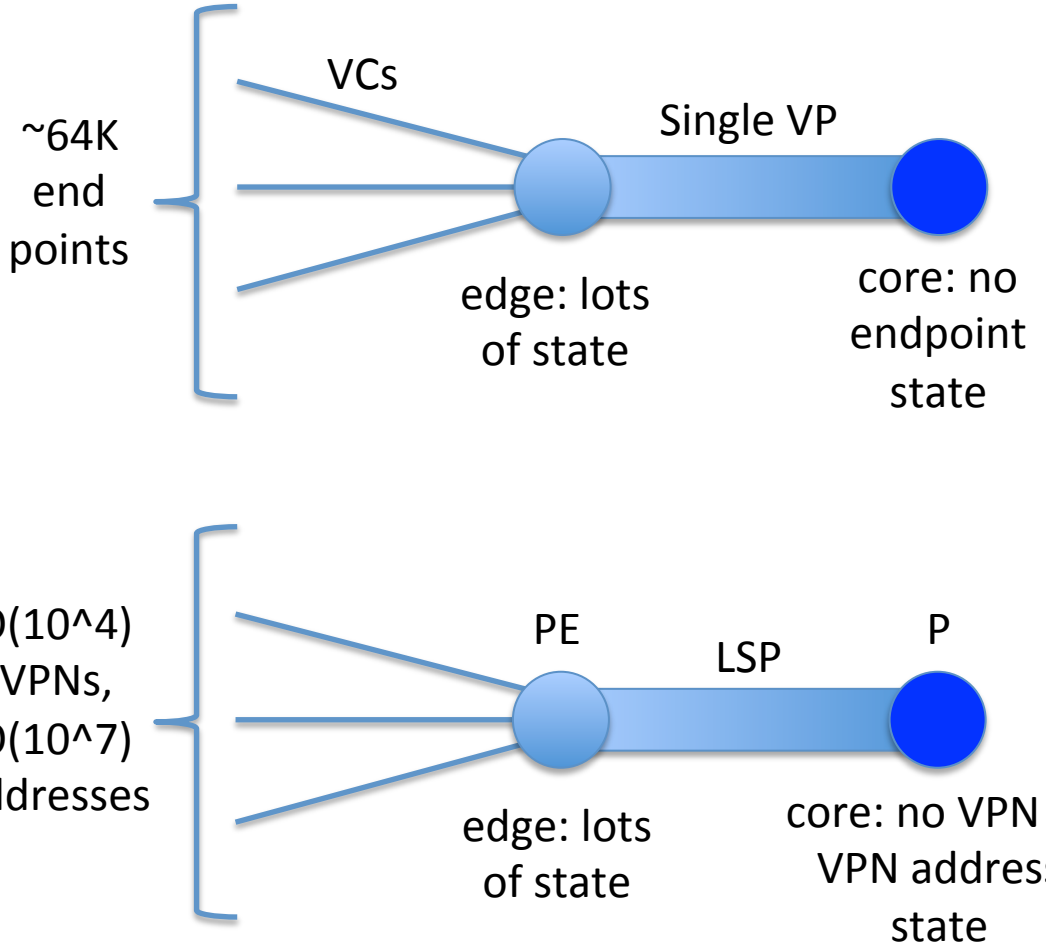
Problem Statement (DC)

- Overlays are all the rage in the data center
  - except that we've been doing overlays/underlays with MPLS pretty much since 1997
- The DC overlays start at the host (server)
  - which requires true "plug-and-play" operation
- To have an MPLS underlay network, the host must be part of the underlay
  - this draft is about making that easy and p-n-p

Problem Statement (access)

- Many have asked that MPLS start at the access node (DSLAM, OLT, cell-site gateway)
- “Seamless MPLS” has suggested the use of LDP DoD (Downstream on Demand) for this
- There haven’t been many implementations of LDP DoD from access node vendors
  - Maybe a different approach/protocol for the same functionality is needed

Overlays/Underlays



ATM overlay (simple)

MPLS overlay (sophisticated)

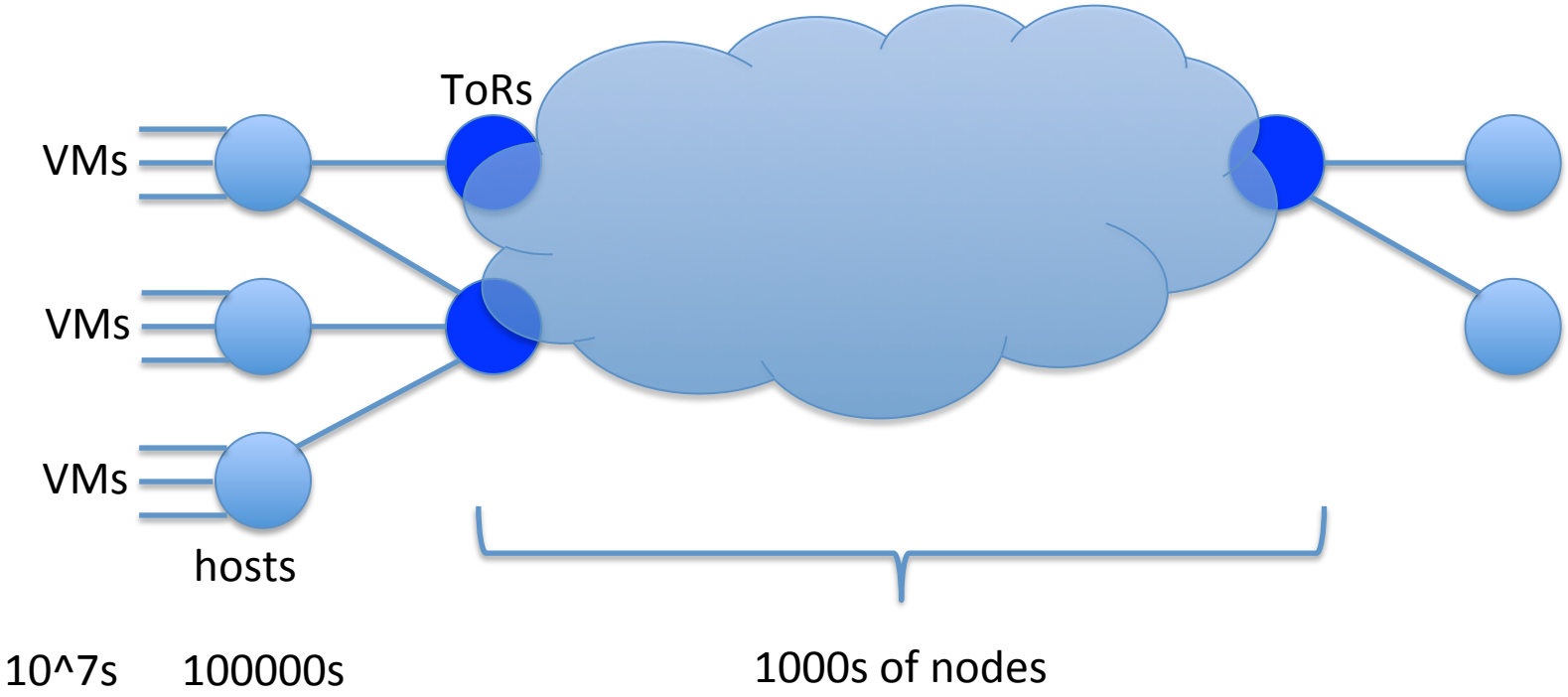
Overlay/Underlay Data Plane

- Commodity chips implemented the MPLS data plane about a decade ago
- Now, some have implemented just one of a largish crop of new overlay encapsulations
  - And, as we speak, there is yet another one

Overlay/Underlay Control Planes

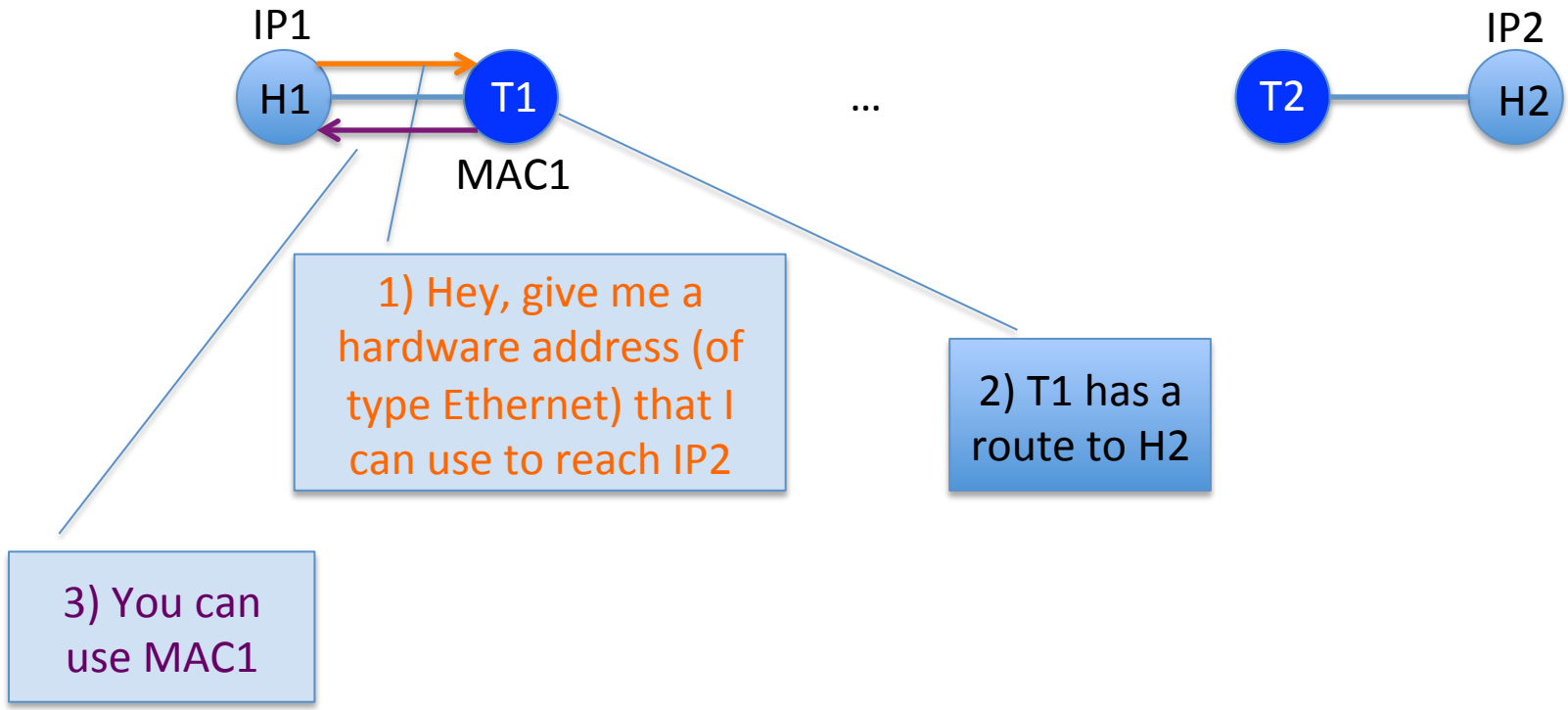
- MPLS has a very sophisticated, robust, scalable and interoperable control plane
  - Various types of hierarchy are supported
  - {BGP, T-LDP} [overlay] over {LDP, RSVP-TE, LDP/ RSVP-TE} [underlay]
- None of the new overlays encapsulations have well-specified, interoperable control planes for either the overlay or the underlay
  - BGP for an overlay (EVPN/IPVPN over VXLAN) is just being proposed

Can the MPLS Control Plane Be Too Sophisticated (in Some Cases)?

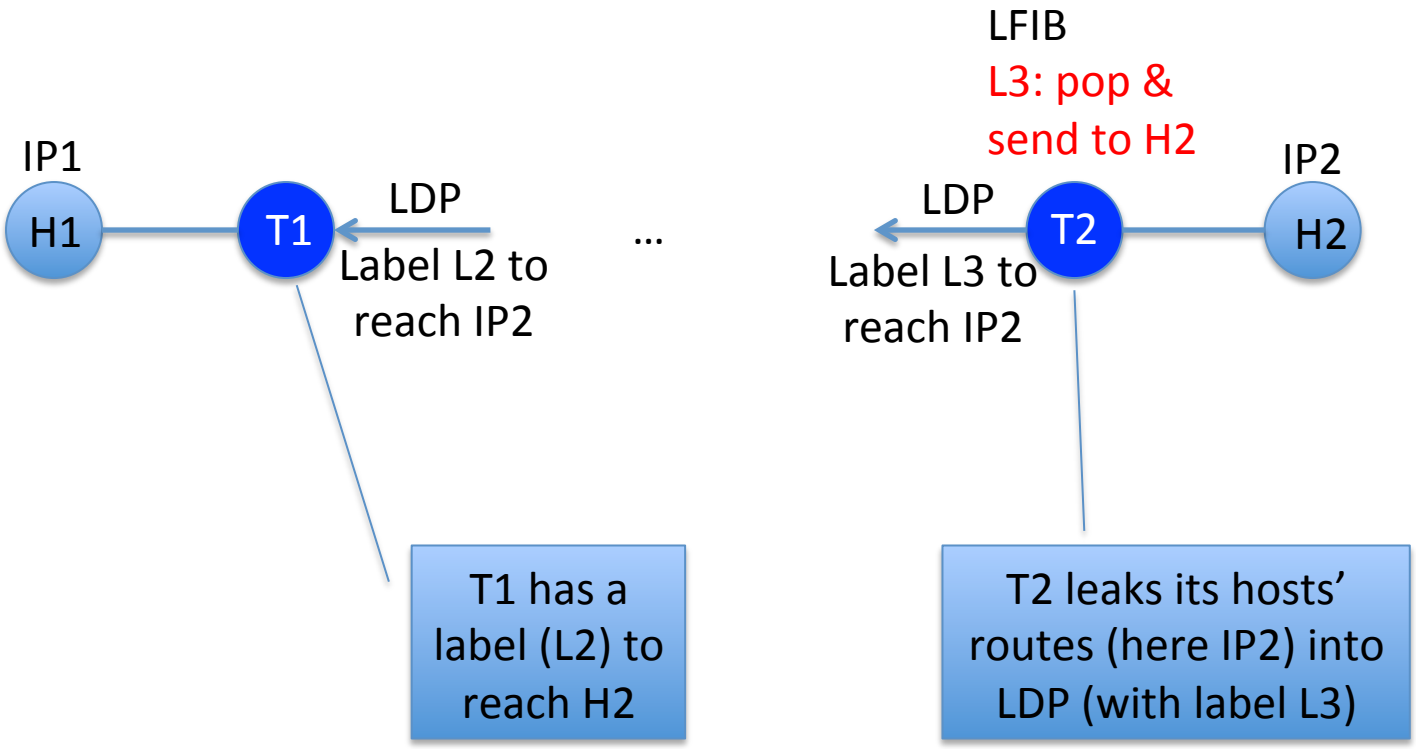


Can't have a flat IGP with so many hosts
LDP DoD with static routing is a possibility, but not ideal
Absolutely has to be plug-and-play – new hosts are added at a high rate

Proxy ARP recap



Labeled ARP (1)



Labeled ARP (2)

1) Hey, give me a hardware address (**of type MPLSoEthernet**) that I can use to reach IP2

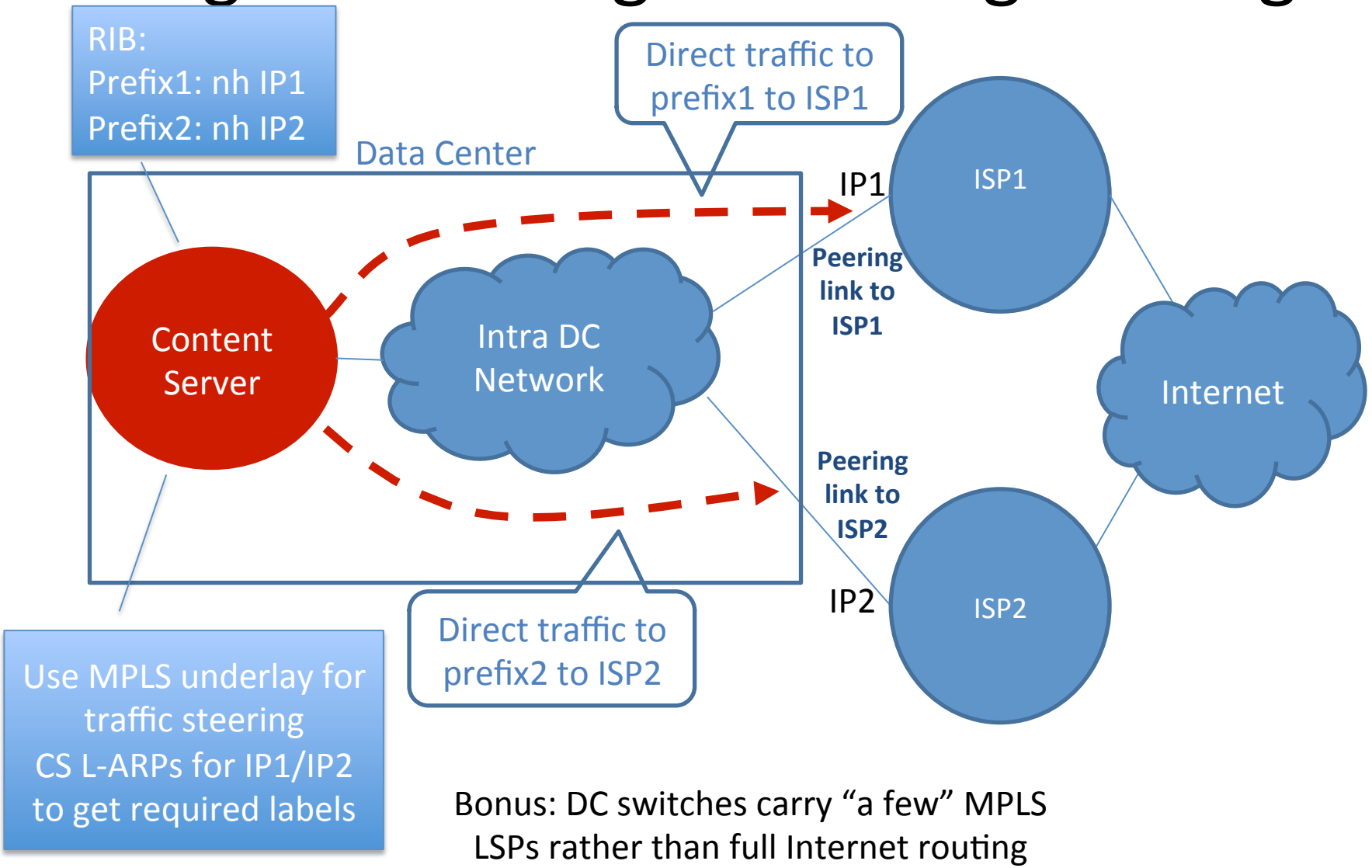
2) T1 allocates L1 for H1 to reach H2, adds an LFIB entry to swap L1 with L2

3) You can use MAC1:L1

Note: new h/w type means that this code is separate from "normal" ARP code

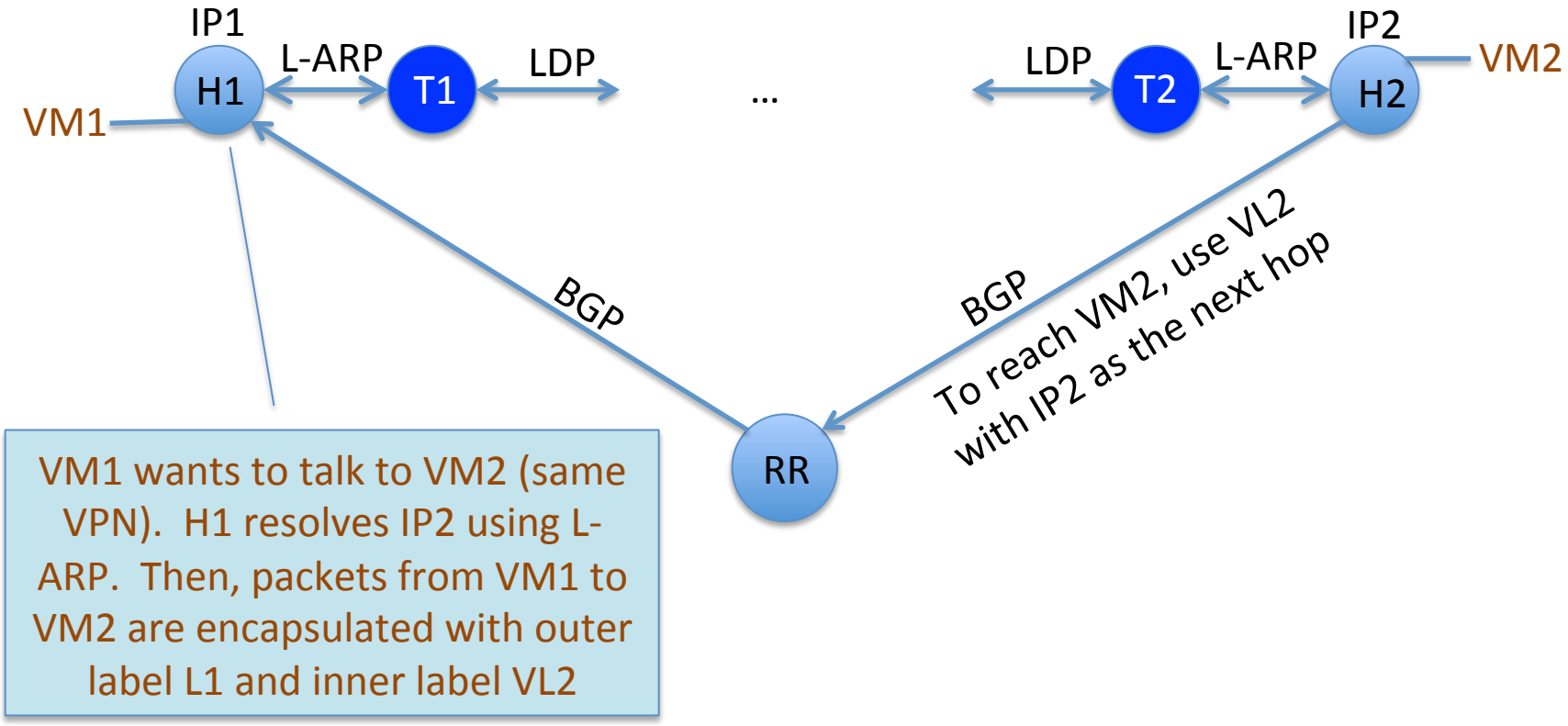
Functionality is very much like LDP DoD.
However, ARP code is plug-and-play and ubiquitous.

Use Case 1:
Egress Peering Traffic Engineering



Bonus: DC switches carry “a few” MPLS LSPs rather than full Internet routing

Use Case 2: MPLS Underlay for DCs (with VRFs/E-VPNs for overlay)



Next Steps

- There are a few problems to resolve
  - Section 4: “For Future Study”
  - We have some of the answers, but not all
  - Philosophy: keep L-ARP client simple
  - Will republish the draft with proper filename
- We will publish a use cases draft, if desired
- We have running code (for Linux hosts)
  - User space daemon, independent of Ethernet ARP
  - Now, all we need is rough consensus :)