cisco Live!

TURN IT UP

#CiscoLive

The bridge to possible

The Choice of LISP for Switching Fabrics

Victor Moreno, Distinguished Engineer
@victoratcisco

BRKENT-2102

Agenda

- **Statement of requirements**
- LISP primer
- Mobility and Scale factors
- Extensibility
- Wired + Wireless
- Operations
- Conclusion

Network Requirements and Implications

Requirements

- High Densities of End-points
- Mobility
- Segmentation
- Stack of Services: IP unicast, Multicast, Layer 2, NAT, Service Insertion, etc.

Implication

- Scale
- Performance
- Simplicity: Consolidated Stack
- Extensibility

66%
Mobility
Mobile device traffic will be 66% of total IP traffic by 2020[1]

26B
IoT
26 billion networked devices and connections will exist by 2020[2]

93%
Cloud
93% of organizations will use multiple clouds by 2019[4]

100
Security
100 days Industry average to detect a common threats[3]

How did we end up here?

- Requirements: Mobility, Segmentation, Path engineering, IP transparency, Scale, Migration
- We tried to do it all in the network: Things got very complicated
- SDN separated the requirements from the underlay network
- New functionality moved to the overlay network
  - DC Networking ➔ VM Networking ➔ Container Networking ➔ Service Meshes
  - Access/WAN Networking ➔ Software Defined Access, Software Defined WAN

Overlay Virtual Networks
Control and Data Plane Separation

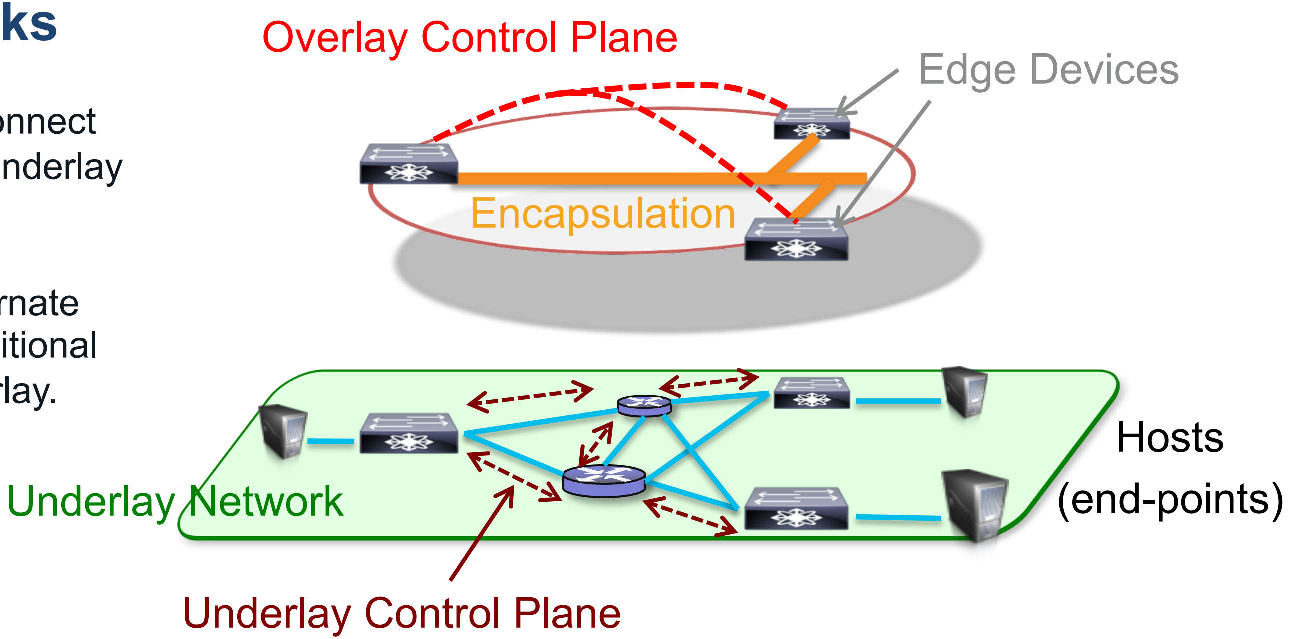
Overlay Virtual Networks

Logical topology used to virtually connect devices, built on top of a physical Underlay topology.

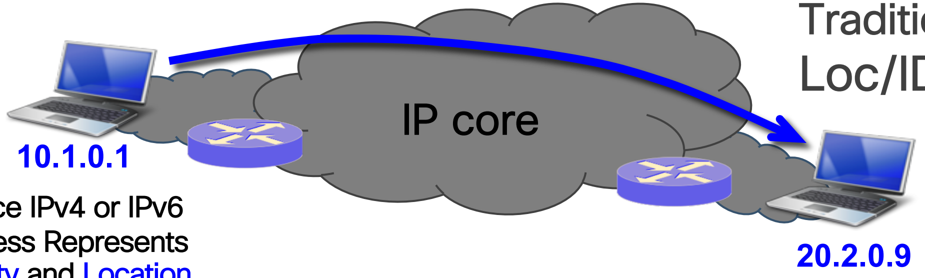
An Overlay network often uses alternate forwarding attributes to provide additional services, not provided by the Underlay.

Benefits:

Single hop networking
Multi-homing may be avoided
Extensible data plane semantics
Connection focus (vs. routing focus)



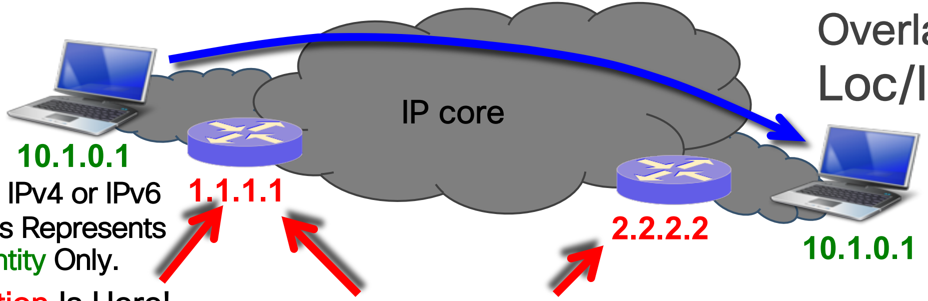
Location and Identity Separation



Device IPv4 or IPv6 Address Represents Identity and Location

Traditional Behaviour
Loc/ID "Overloaded" Semantic

When the Device Moves, It Gets a New IPv4 or IPv6 Address for Its New Identity and Location



Device IPv4 or IPv6 Address Represents Identity Only.

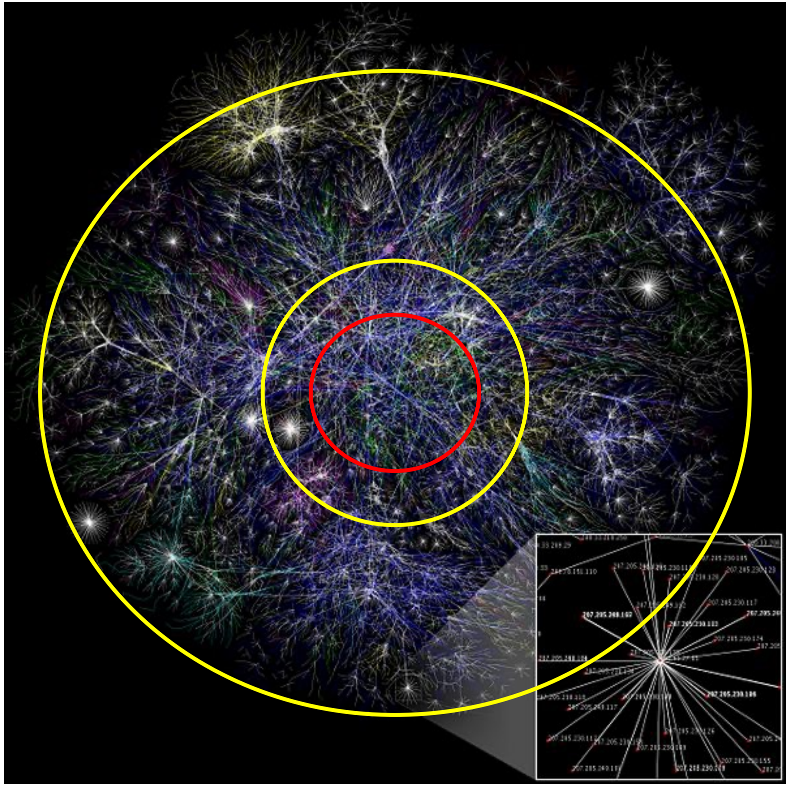
Its Location Is Here!

Only the Location Changes

Overlay Behaviour
Loc/ID "Split"

When the Device Moves, Keeps Its IPv4 or IPv6 Address.
It Has the Same Identity

Clean up the core, move churn to the Overlay



- Only locators remain in the core
  - Locators are stable & low churn
  - Multi-path routing on locators
- Move End-Point IPs to the Overlay
  - Entropy and churn is in the End-points
  - No routing, just map to locators
- **“End-point addressing can be independent of the topology”**

Function distribution in overlay networks

Function	Underlay Routing	Overlay Routing
Loop-free path computation	Yes	
Multi-homing	Yes	
Liveness & Failover	Yes, aggressive	
Dissemination of updates	Pervasive	Selective/scoped
Mobility		Yes
Segmentation		Yes
Policy		Yes
Extensibility		Yes
Programmability		Yes
Path engineering		Yes
Encryption		Yes

New requirements lead to new approaches

Does this still look like a network?

- Maintain a directory of end-points
  - And their attributes – Location is one of them, but tags, geo-coordinates and other information is common.
- Enforce different types of policies at lookup time
- Make the directory scalable and accessible
- Optimize lookups and information distribution

- Does DNS come to mind? What about a service mesh?
- Note that resolving topologies and calculating best paths is not a requirement

Overlay Control Plane: Push vs. Pull

Push

- Routing Protocols (e.g. BGP)
- Distribute/push updates to all routers
- Run optimal path computation on all updates received

- Ideal for the underlay
  - Relatively static (infrequent changes)
  - Responsible for multi-path routing decisions
  - High computation requirement, distributes and calculates routes with reliable failover

Pull

- Mapping protocols (e.g. DNS, LISP)
- Updates are pulled only where needed
- No need to run path computation algorithms

- Ideal for overlay
  - Very dynamic (high rates of change)
  - Responsible for end-point attachment & services
  - Computationally nimble, updates and services queries to a database.

Agenda

- Statement of requirements
- **LISP primer**
- Mobility and Scale factors
- Wired + Wireless
- Operations
- Extensibility
- Conclusion

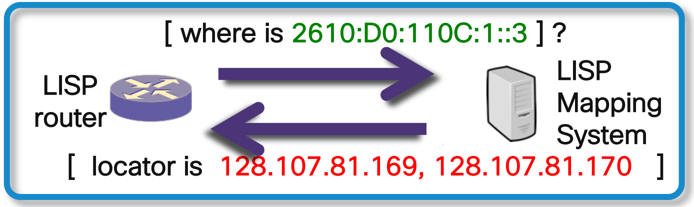
LISP Operations

LISP :: Mapping Resolution “Level of Indirection”

- LISP “Level of Indirection” is analogous to a DNS lookup
  - DNS resolves IP addresses for URL   Answering the “WHO IS” question

DNS
Name-to-IP
URL Resolution

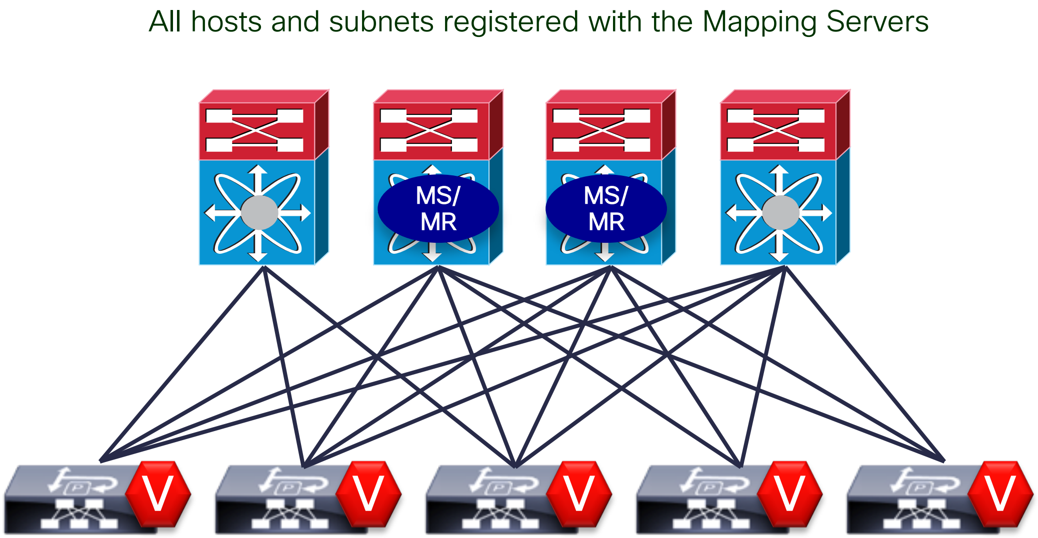
  - LISP resolves locators for queried identities   Answering the “WHERE IS” question



LISP
Identity-to-locator
Mapping Resolution

LISP Control Plane for VXLAN

Host and Subnet Route Registration



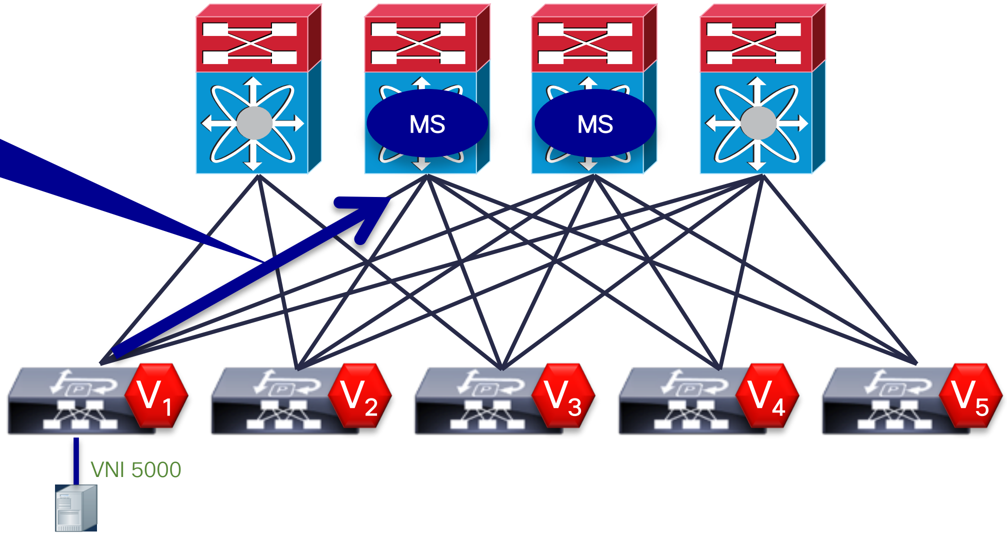
- Host Route Registration decoupled from the Underlay protocol
- Use LISP on the leaf nodes to resolve internal host/subnet routes and external reachability information

LISP Control Plane

Host Registration

	Endpoint ID (EID)	RLOC (Routing Locator)
IP	VNI	Next-Hop
1	5000	IP V1

Map Register
EID = IP1, VNI 5000
RLOC = xTR IP V1



V1 LISP Tunnel Router (xTR) & VTEP*

MS LISP Mapping System

* VTEP = VXLAN Tunnel End-Point

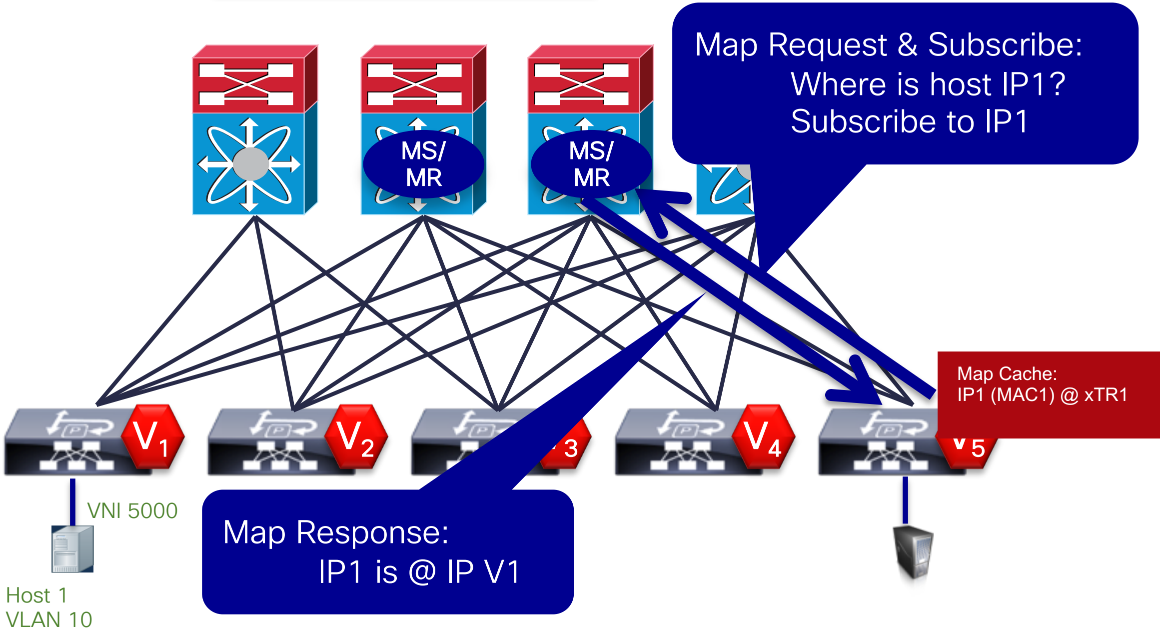
1. Host Attaches
2. Attachment xTR registers host's IP (+MAC) in LISP

LISP Control Plane

Host Resolution

IP	VNI	Locator
1	5000	IP V1

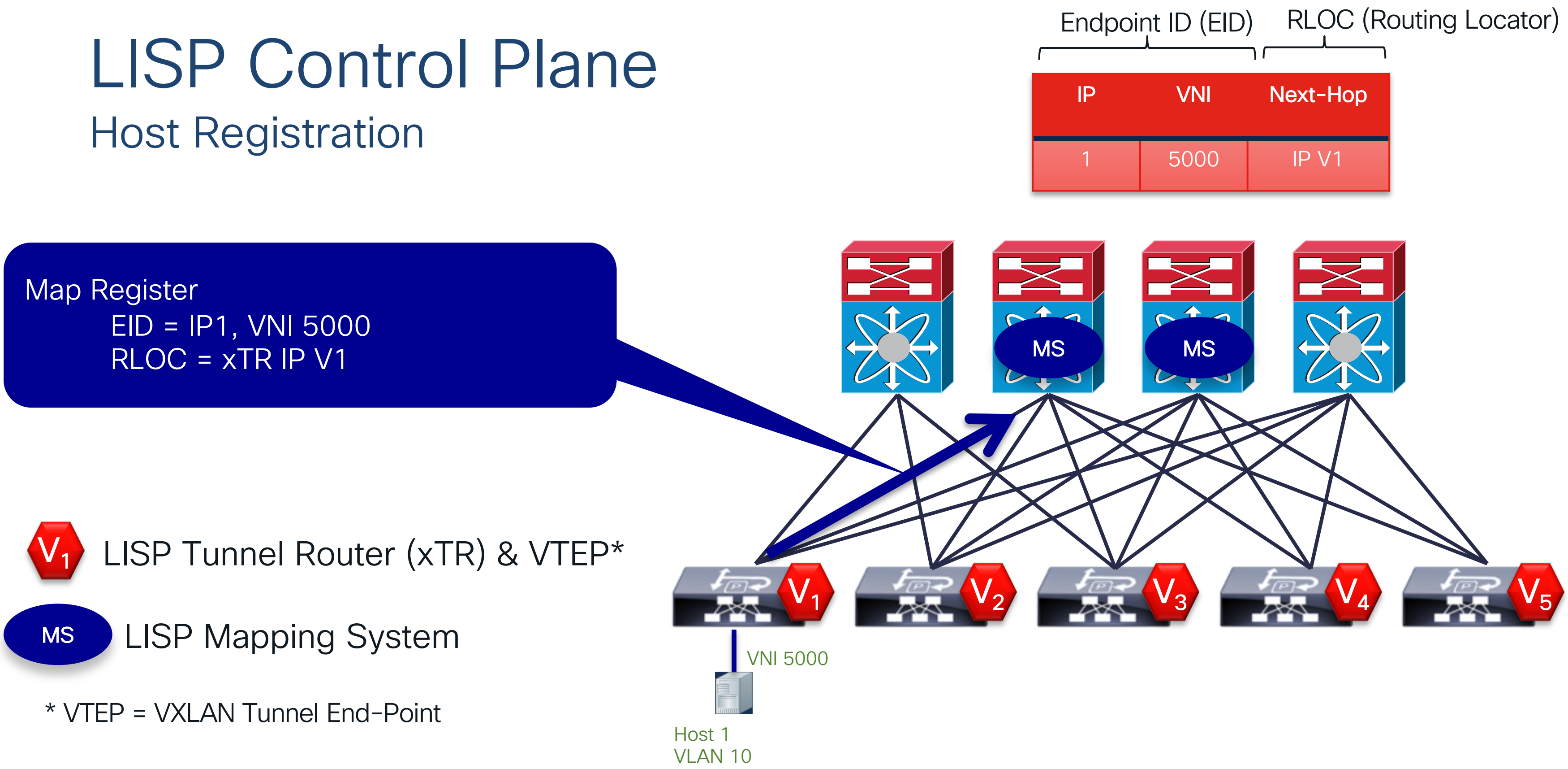
Publication	Subscriber
IP 1	V5



1. Host 2 wants to talk to host 1, the xTR (V5) issues a map-request
2. The Mapping System responds and starts a subscription to host 1 for v5
3. The response is cached at the requesting xTR (V5): LISP map-cache

LISP Control Plane

Host Registration



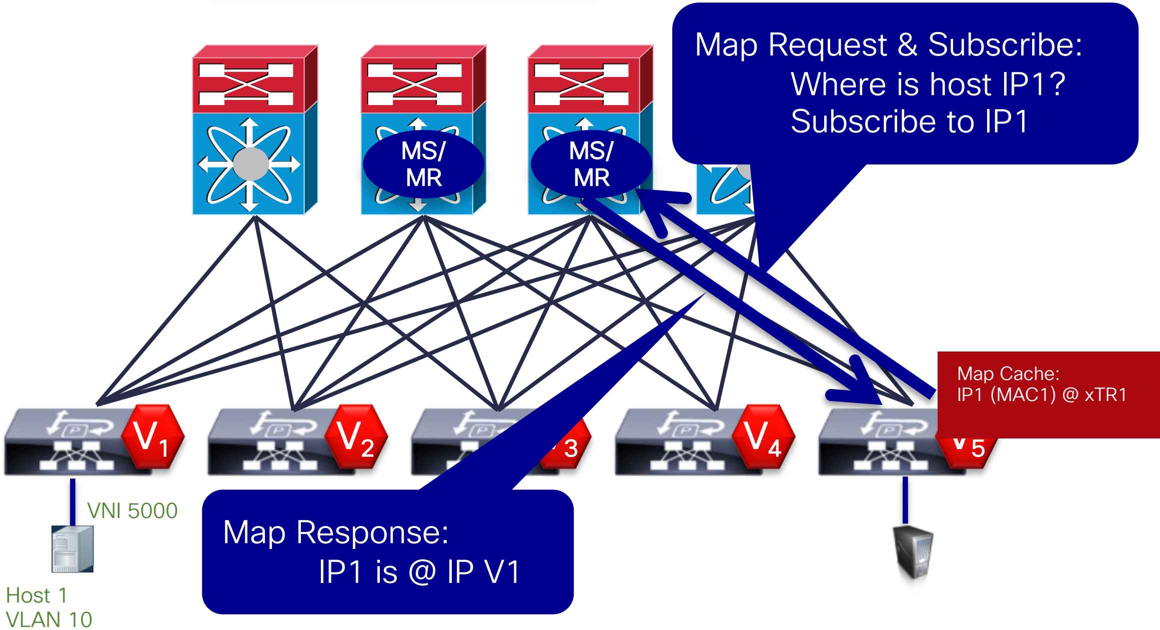
1. Host Attaches
2. Attachment xTR registers host's IP (+MAC) in LISP

LISP Control Plane

Host Resolution

IP	VNI	Locator
1	5000	IP V1

Publication	Subscriber
IP 1	V5

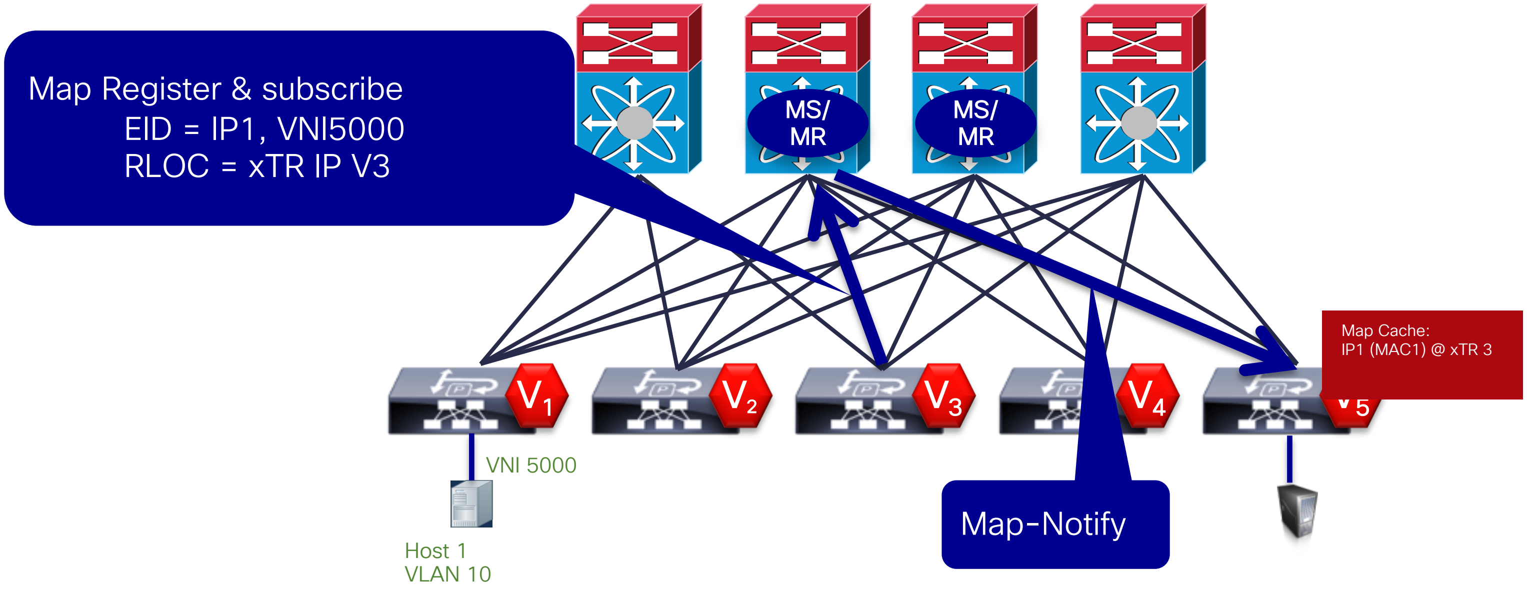


1. Host 2 wants to talk to host 1, the xTR (V5) issues a map-request
2. The Mapping System responds and starts a subscription to host 1 for v5
3. The response is cached at the requesting xTR (V5): LISP map-cache

LISP Control Plane

Host Moves

IP	VNI	Next-Hop
1	5000	IP V3



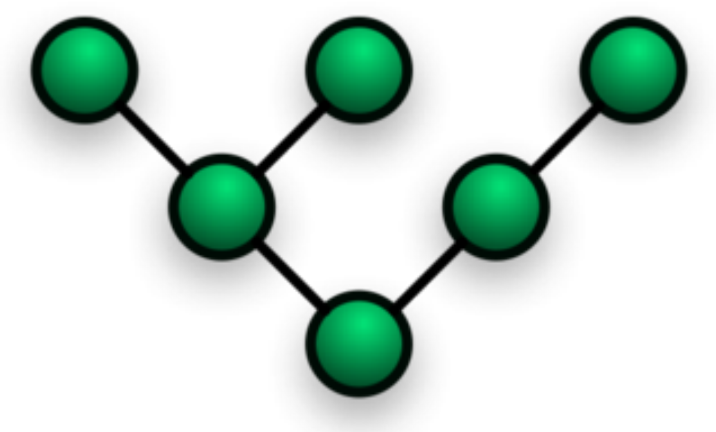
1. Host Moves to V3
2. V3 detects Host1 and registers H1
3. V5 is notified of the move and updates its map-cache

Agenda

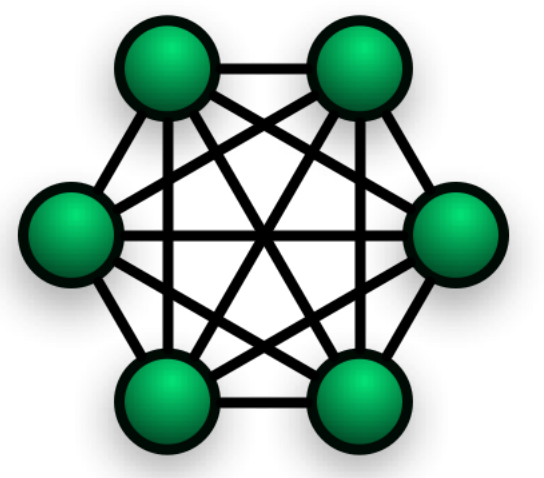
- Statement of requirements
- LISP primer
- **Mobility and Scale factors**
- Wired + Wireless
- Operations
- Extensibility
- Conclusion

Static Tree vs. Dynamic Full Mesh

Different protocols for different purposes

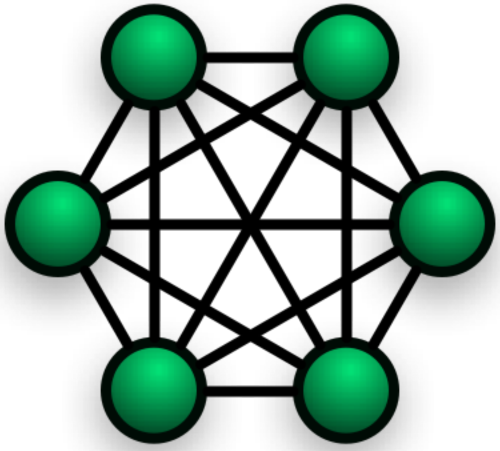


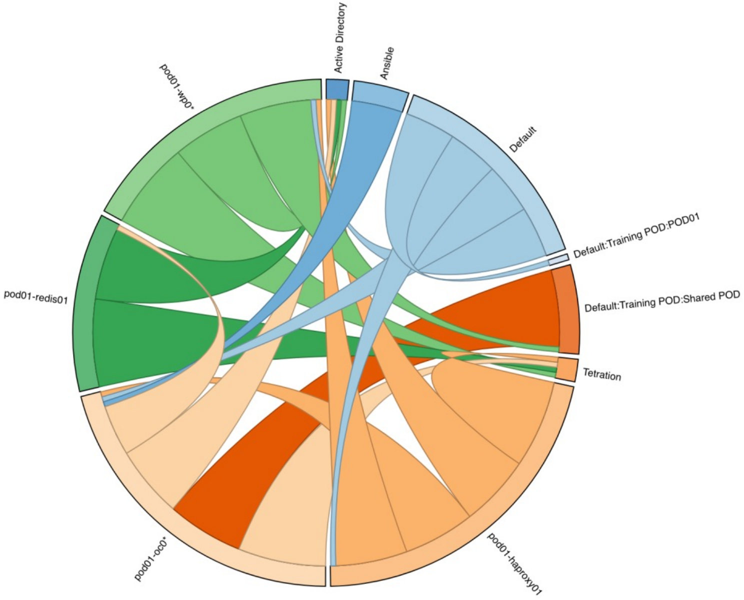
The Internet
(Static / Low Churn)



Network Overlay
(Dynamic / High Churn)

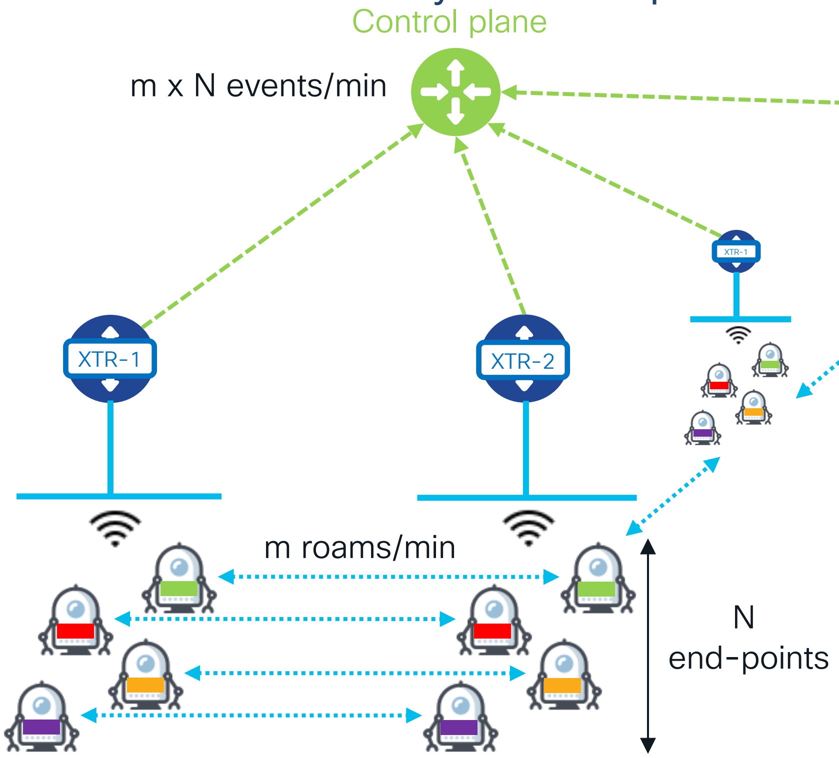
Overlay Topology vs. effective connectivity



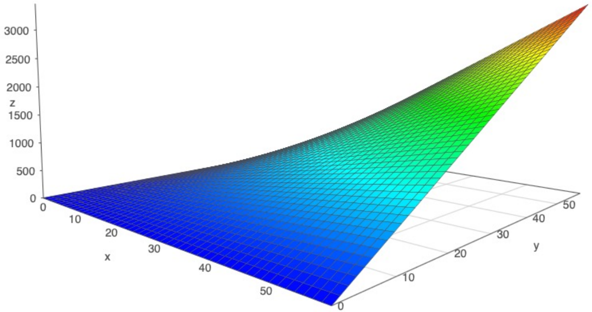


Mobility in the Access

Rates of Mobility are compounded by end-point density

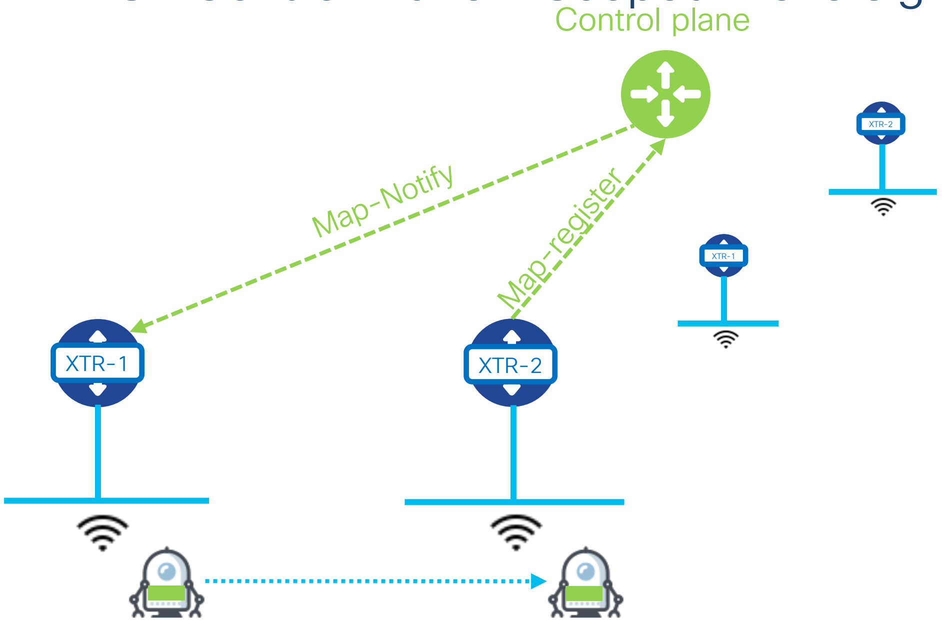


- Rate (@CP) = m x N
- IoT example:
  - 16K Robots, each moves 20 times/minute
  - 20 r/min x 16K = 320K r/min = 5,333 roams/s
  - Sub 70 ms convergence
  - 64K Robots (21K r/s) 100K Robots (31K r/s)
- Stadium 64K@1r/min ➔ ~ 1K r/s



Mobility in the Access

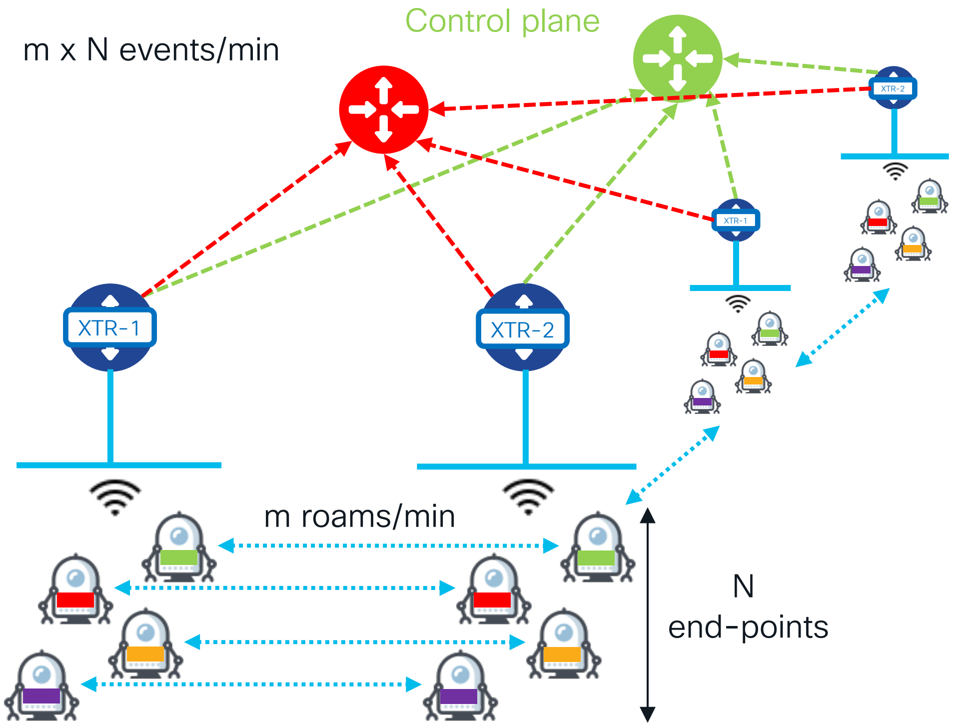
LISP Control Plane - Scoped move signaling



- Sparse signaling: Only the xTRs involved in the move
- Light processing: Only mapping updates, no path calculation
- The rate of events at the xTRs is a fraction of the total rate of events
- Rate@xTR = (m x N) / number of xTRs
- IoT example
  - In a network with 100 access routers
  - Rate @xTR = 5,333 r/s /100 = ~54 r/s

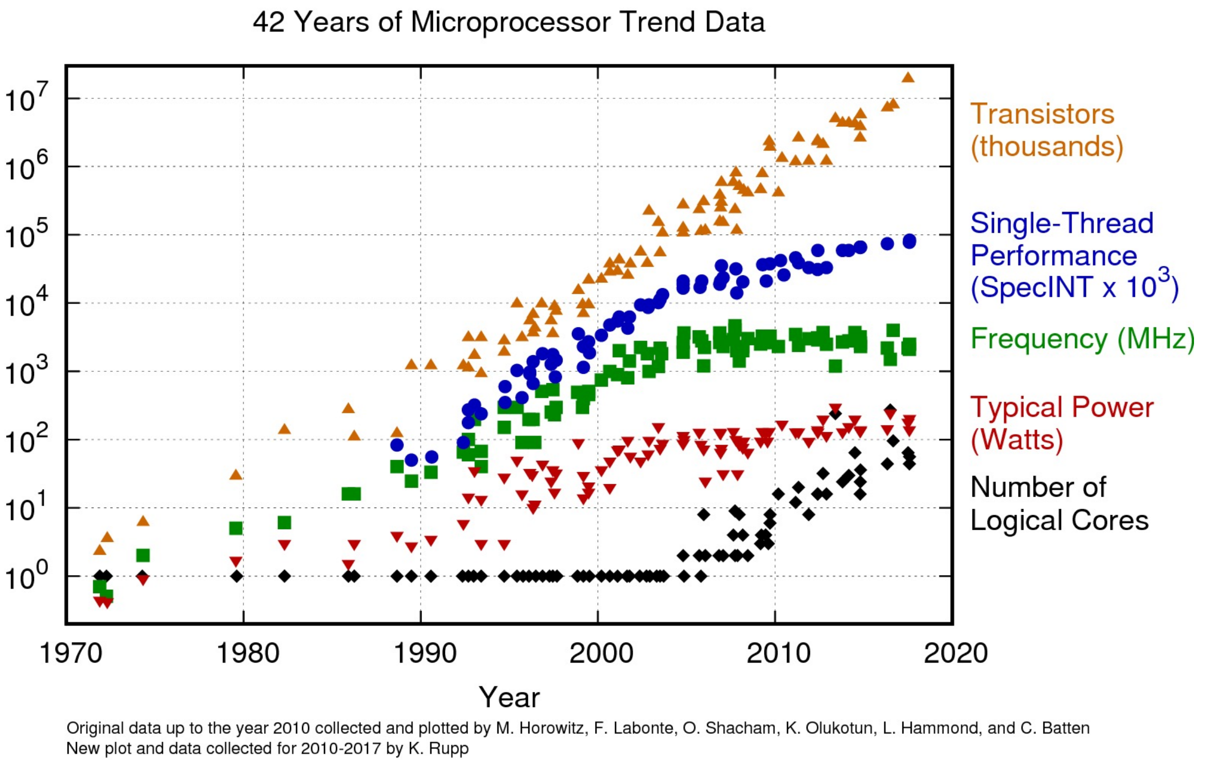
Mobility in the Access

LISP Control Plane – Horizontal Scale of the Control Plane



- Prefixes can be scoped to different control plane nodes
  - Horizontally scale by adding nodes
- Tested performance for one control plane node:
  - Up to 800 r/s while converging faster than 70 ms
- IoT example:
  - 5,333 r/s / 800 r/s = 7 CP nodes

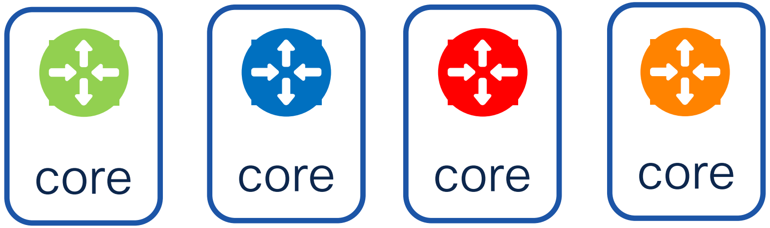
Moore’s law in perspective



As clock frequencies flatten out …

Performance improvements come from multi-core parallel processing

Horizontal scaling ➔ Micro-services

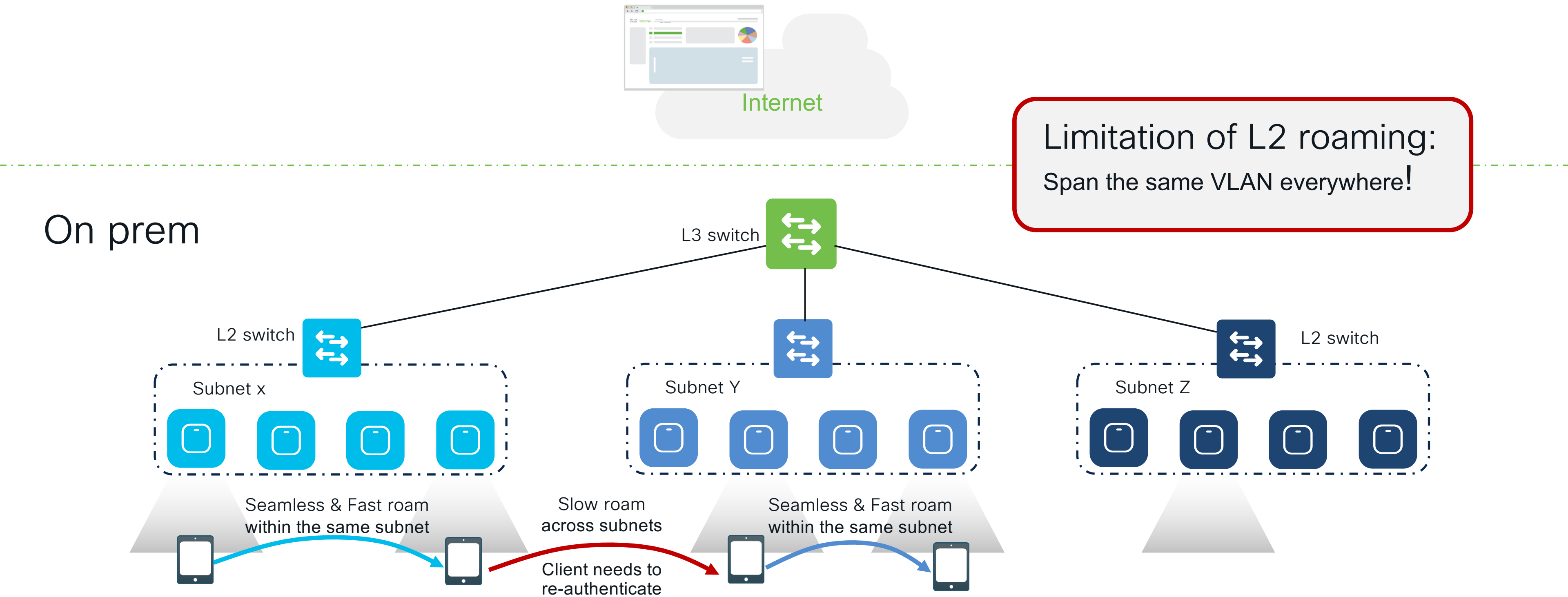


Small process footprint ➔ handle more load with less CPU/memory
Independent processes per end-point group ➔ Map to multiple cores

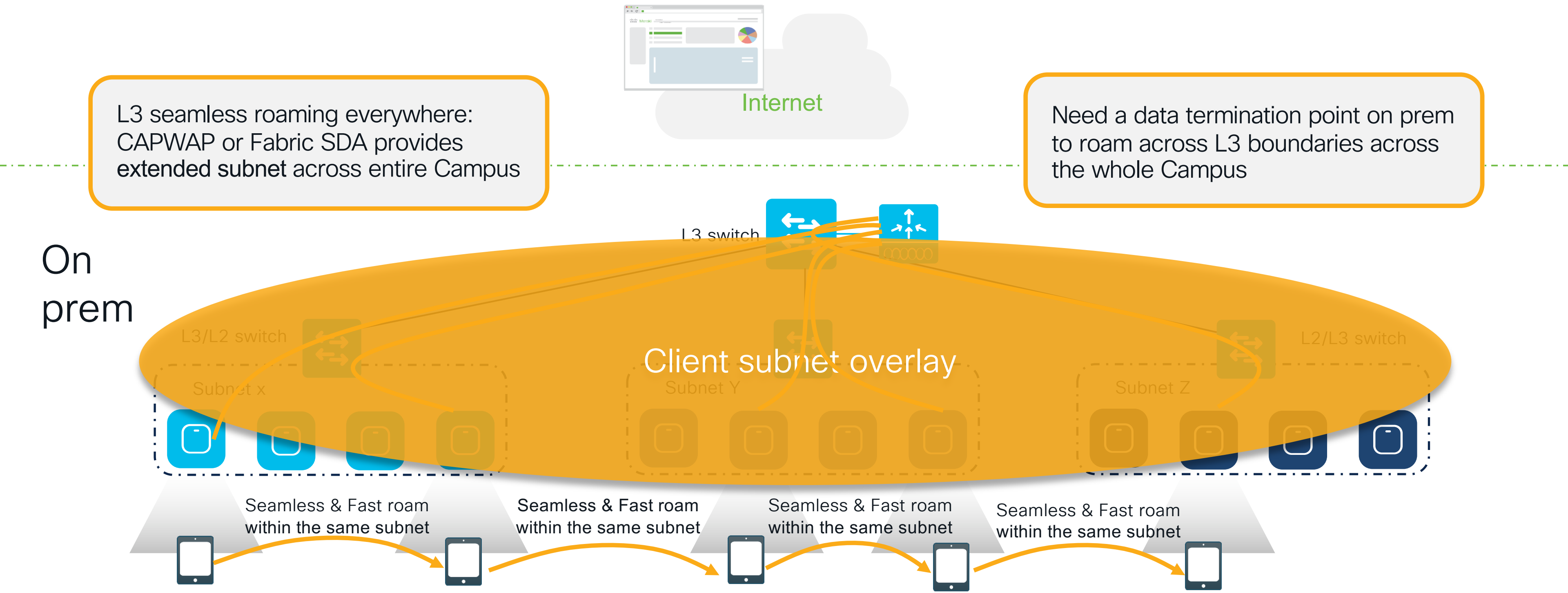
Agenda

- Statement of requirements
- LISP primer
- Mobility and Scale factors
- **Wired + Wireless**
- Operations
- Extensibility
- Conclusion

Layer 2 seamless roaming



Layer 3 seamless roaming provided by the Fabric



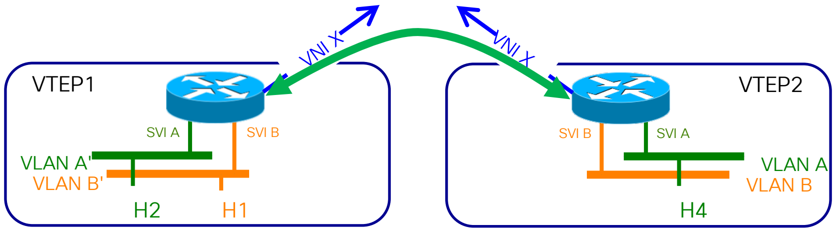
Two modes of Operation

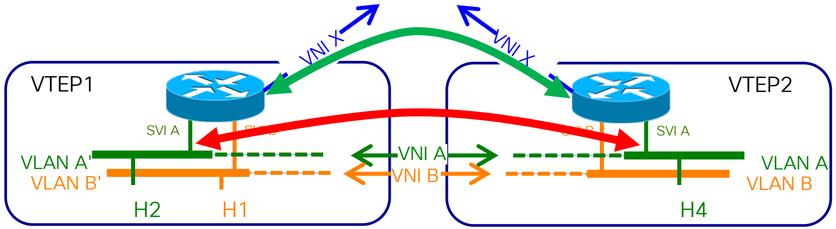
Enhanced: Strictly Routed

- Intra-subnet traffic is routed
- Optimized IP communications
- Flood suppression (BUM)

Traditional: Routing + Bridging

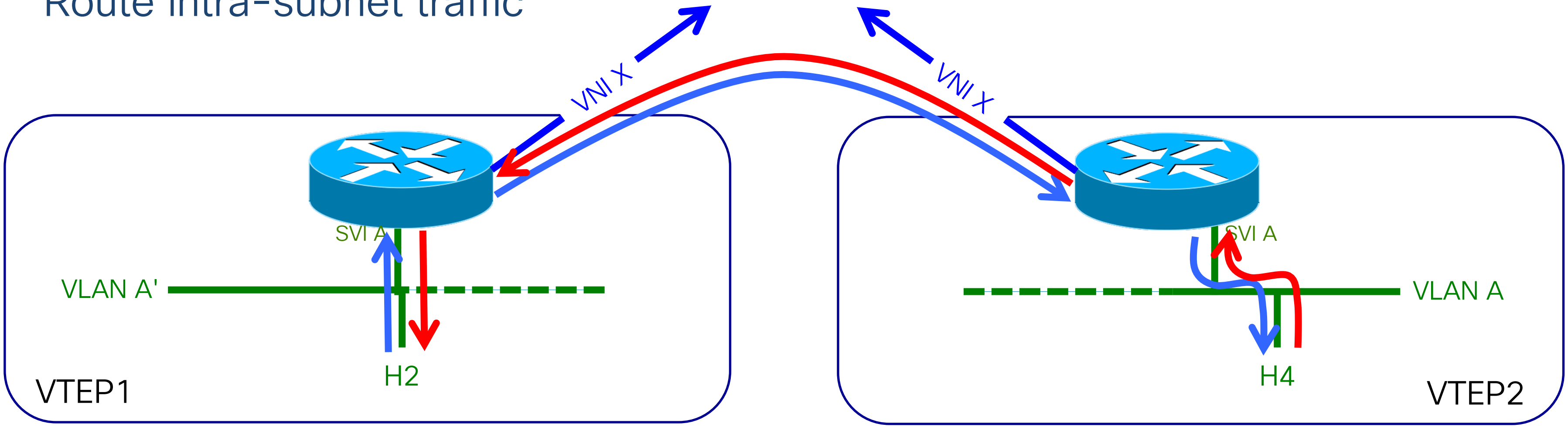
- Intra-subnet traffic is bridged
- Support non-IP communications
- Flood L2 protocols (BUM), ARP/ND





Dispersed subnets in L3 Overlay Fabrics

Route intra-subnet traffic

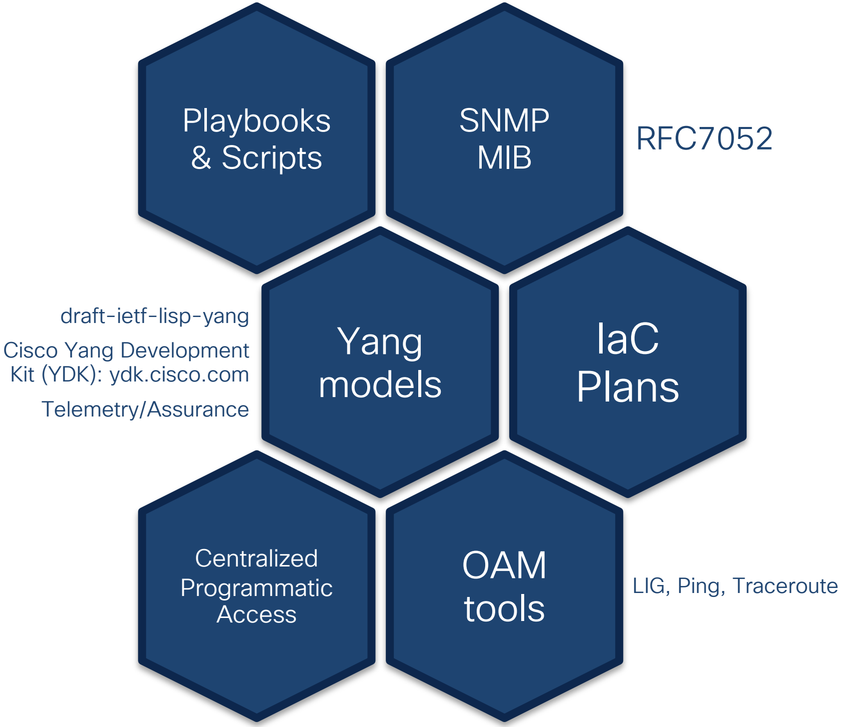


- Local SVI replies to ARPs for remote subnet members with its own MAC
  - Traffic to remote subnet members is sent to local XTR/SVI and routed
- Intra-subnet traffic between VTEPs will be encapsulated in VNI X
- Standard longest prefix match routing takes place:
  - Host routing for all known remote hosts ➔ Forward over VNI X
  - Local hosts are covered by directly connected prefix, a host route will not be present

Agenda

- Statement of requirements
- LISP primer
- Mobility and Scale factors
- Wired + Wireless
- **Operations**
- Extensibility
- Conclusion

Management and Operations Toolkit



Separation of Concerns

Underlay

- Prescriptive topologies
- Minimal entropy
- Well understood baseline
- Agile Analytics: Easy to detect anomalies and deviations

Overlay

- Single hop topology
- Rich Semantics
- Centralized database
- Simplified mechanics enable focus on data

Agenda

- Statement of requirements
- LISP primer
- Mobility and Scale factors
- Wired + Wireless
- Operations
- **Extensibility**
- Conclusion

Network Protocol Consolidation

From multiple topics to one theme with variations

Networks have evolved organically:

- Some L2, Some L3
  - L2 protocols: STP, MLAG
  - First Hop Resiliency (FHRP): VRRP, HSRP
- Multicast: PIM ASM, SSM
- L2 extensions: VPLS, EVPN
- Traffic Engineering: RSVP, MPLS
- NAT

- An L3 Access removes the need for some of these (FHRP, L2)
- The remaining services can be provided by a single protocol stack
  - With a unified operational model: Registration and Resolution are unchanged and are the same for all services
  - It is all about mapping Identifiers to Locations. Not more, not less.
  - Where necessary additional semantics can be added to the mappings
  - Services are expressed as policies governing the responses provided.
  - Pulling allows the evaluation of flow context as part of the policy, it also allows us to scale policies

Feature set Extensibility

Policy modulated

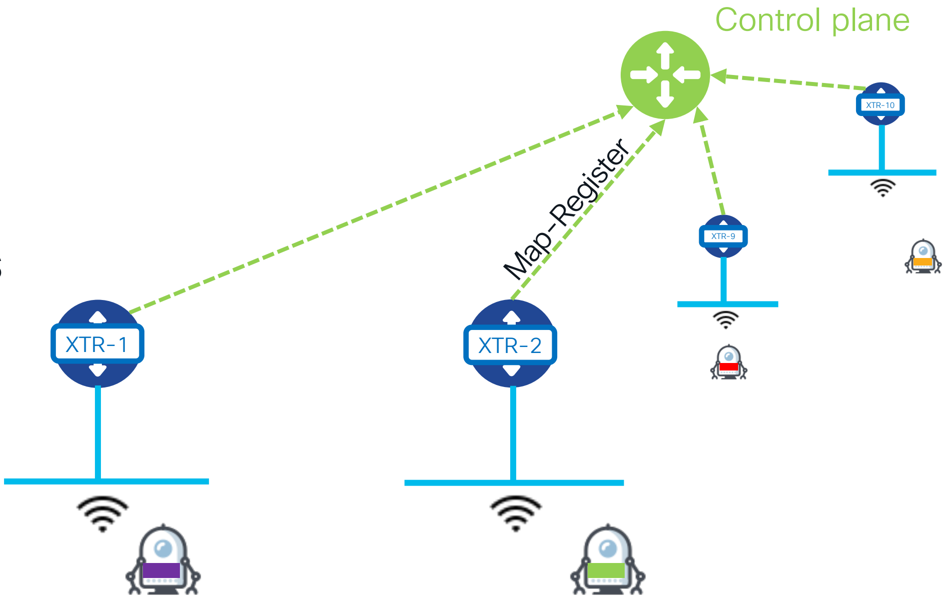
↓

Extensible Semantics + On-demand Lookups = Extensible Feature Set

Extensible Semantics

EID				RLOC	
Instance	IP	MAC	Security Group	IP	Geo-coordinates
Green	10.0.0.1	FABB.BEEF.0234.d2cF	VIP	169.8.2.1	34° 04' 21.00" N, -118° 25' 27.98" W

- Both EID and RLOC can be extended with metadata
- The LISP Canonical Address Family (LCAF) enables the encoding of the metadata in control plane messages



LISP Canonical Address Family (LCAF)

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|           AFI = 16387         |     Rsvd1     |     Flags     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    Type       |     Rsvd2     |            Length             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                             . . .                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

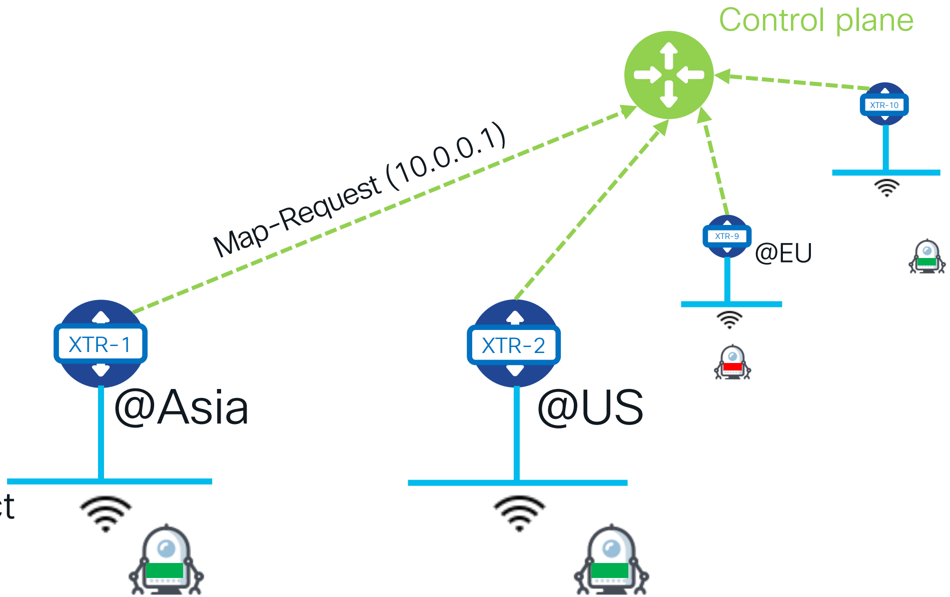
- Type 0: Null Body
- Type 1: AFI List
- Type 2: Instance ID
- Type 3: AS Number
- Type 4: Application Data
- Type 5: Geo-Coordinates
- Type 6: Opaque Key
- Type 7: NAT-Traversal
- Type 8: Nonce Locator
- Type 9: Multicast Info
- Type 10: Explicit Locator Path
- Type 11: Security Key
- Type 12: Source/Dest Key
- Type 13: Replication List Entry
- Type 14: JSON Data Model
- Type 15: Key/Value Address Pair
- Type 16: Encapsulation Format

In-context lookups

EID				RLOC	
Instance	IP	MAC	Security Group	IP	Geo-coordinates
Green	10.0.0.1	FABB.BEEF.0234.d2cF	VIP	169.8.2.1	34° 04' 21.00" N, -118° 25' 27.98" W
Green	10.0.0.1	FABB.BEEF.0234.d2cF	Overseas	IPS-IP	39° 54' 16" N 116° 23' 29" E

- Lookups are completed on demand
- This allows the inclusion of the context of the requestor in the lookup
- Different responses for different requestors

- e.g. Attributes of the requestor determine the mapping obtained
  - Requestor in Asia and Green VPN (instance) ➔ send to IPS
  - Requestor in Red VPN, deny response
  - Requestor in Green and @US ➔ Connect



Policies

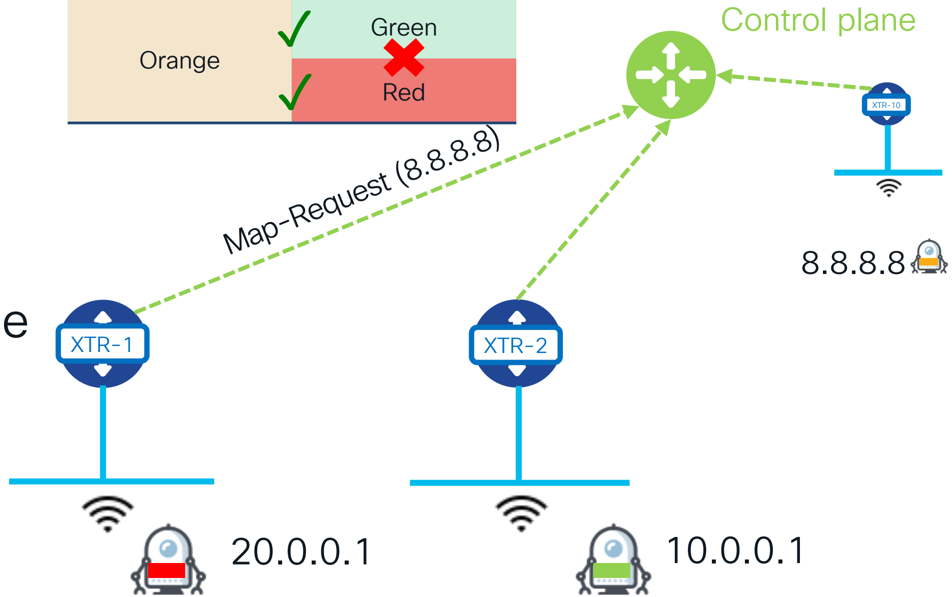
- Conditions can be imposed on the lookup to render different policies
  - Lookup rules
  - Policy Table

- A policy table may be consulted in order to modulate the lookup

- The metadata required to impose the lookup rules or consult the policy table is in the Mapping Database as well as the control plane messages

EID				RLOC
Instance	IP	MAC	Security Group	IP
Green	10.0.0.1	FABB.BEEF.0234.d2cF	VIP	169.8.2.1
Red	20.0.0.1	FABB.BEEF.0234.d2cF	Overseas	IPS-IP
Orange	8.8.8.8		Shared	100.64.10.1

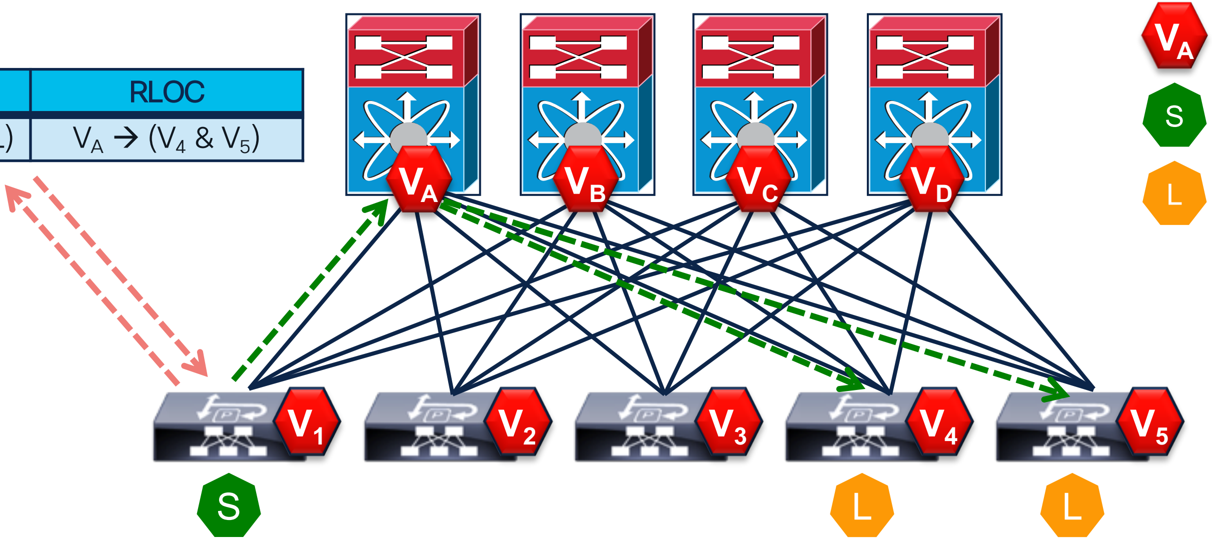
Extranet	Subscribers
Orange	✓ Green
	✓ ~~Red~~

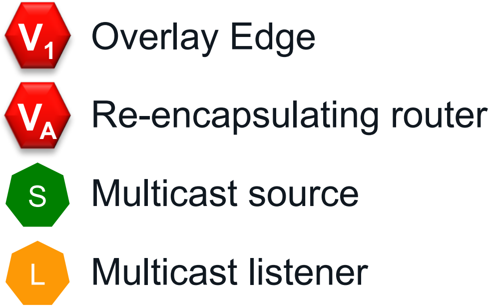


Signal Free Multicast

Normalize Unicast and Multicast behavior

EID	RLOC
Group (S,L)	$V_A \rightarrow (V_4$ & $V_5)$





Overlay Control Plane creates multicast replication lists as Listener sites register

Multicast group simply mapped to the replication list (or underlay group)

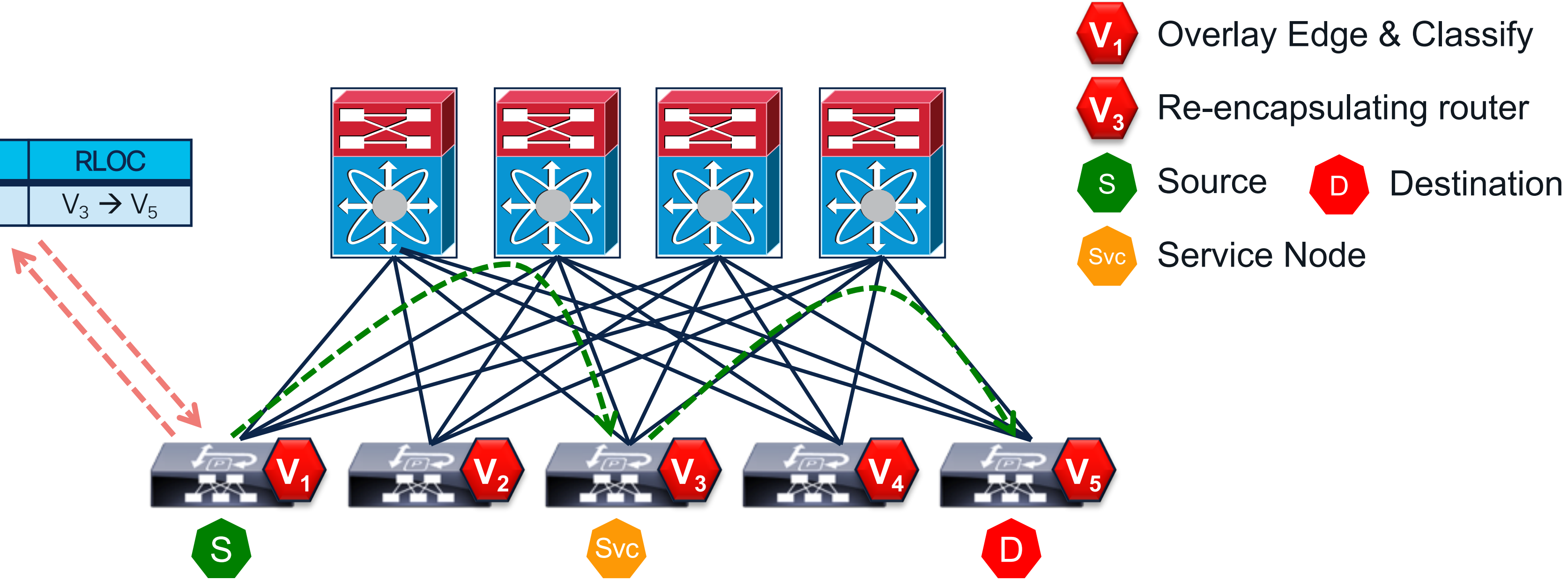
Over-the-top replication for unicast underlay: intermediate re-encapsulation points

Native support for extranet and seamless source mobility

Traffic Engineering and Service Insertion

Controller path provisioning

EID	RLOC
D	$V_3 \rightarrow V_5$



The controller distributes the desired Service or TE path to the overlay edges

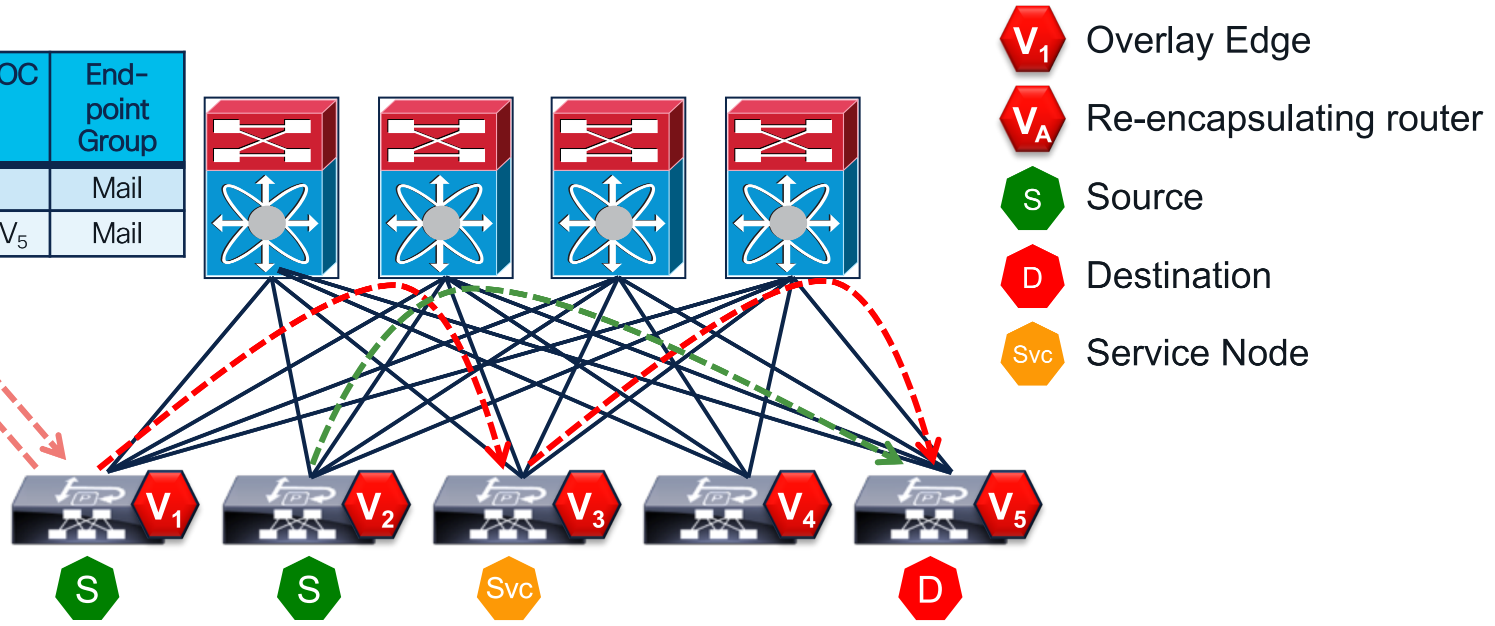
The path may be expressed as a string of RLOCs

The path may be delivered one RLOC at a time at each requesting RTR

Integrated Policy

Context aware mappings

Src-RLOC	EID	D-RLOC	End-point Group
V_2	D	V_5	Mail
V_1	D	$V_3 \rightarrow V_5$	Mail



Map Replies may include policy information such as the Group for the destination

Map Requests may be serviced differently based on the Locator originating the request

The demand protocol advantage

- No routing processing ➔ No “churn”
  - This is an underlay function
- Scoped activity (only involve the relevant network devices) = Horizontal Scale
  - Reduced processing requirements (less CPU required)
  - A lighter footprint in memory/state (smaller tables on network elements)
  - An architectural basis for the disaggregation of the control plane into Micro-service threads
  - Fast convergence at massive scale
- Conditional and contextual resolution enables policy and functionality in the control plane
  - Extend the feature set without altering operations or adding more machinery
- Simplicity of operations
  - Routing is made prescriptive and immutable in the underlay
  - The overlay is focused on simple database lookups on a flat “topology”

Summary

- Scale
  - Minimize and horizontally distribute state
  - Scoped updates to enable a wider network radius
- Performance
  - Focused update processing
  - Fast convergence decoupled from network scale
- Simplicity
  - Features consolidated into a single stack and model
- Extensibility
  - Functionality as policy

Thank you

cisco Live!

TURN IT UP

#CiscoLive