IPv6 Microsegmentation

Ivan Pepelnjak (ip@ipSpace.net)
Network Architect

ipSpace.net AG

Who is Ivan Pepelnjak (@ioshints)

Past

- Kernel programmer, network OS and web developer
- Sysadmin, database admin, network engineer, CCIE
- Trainer, course developer, curriculum architect
- Team lead, CTO, business owner

Present

- Network architect, consultant, blogger
- Webinar and book author

Focus

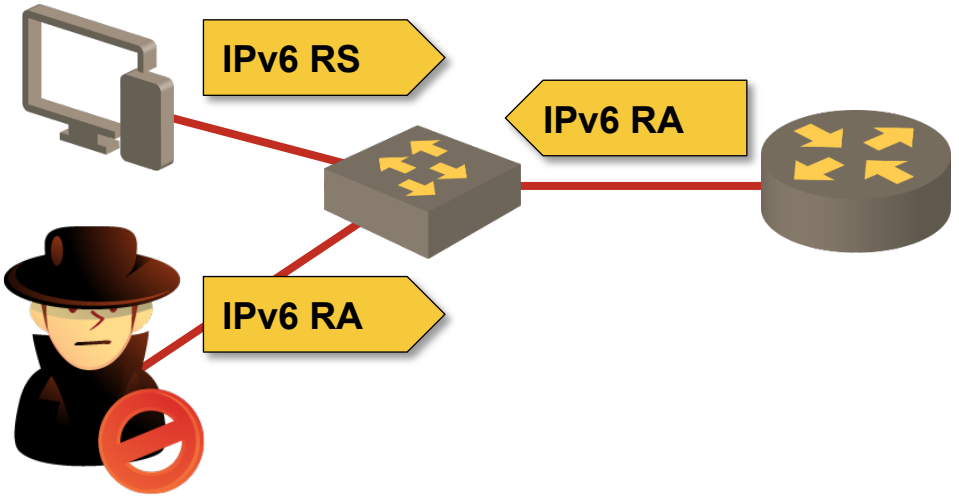
- Large-scale data centers, clouds and network virtualization
- Scalable application design
- Core IP routing/MPLS, IPv6, VPN

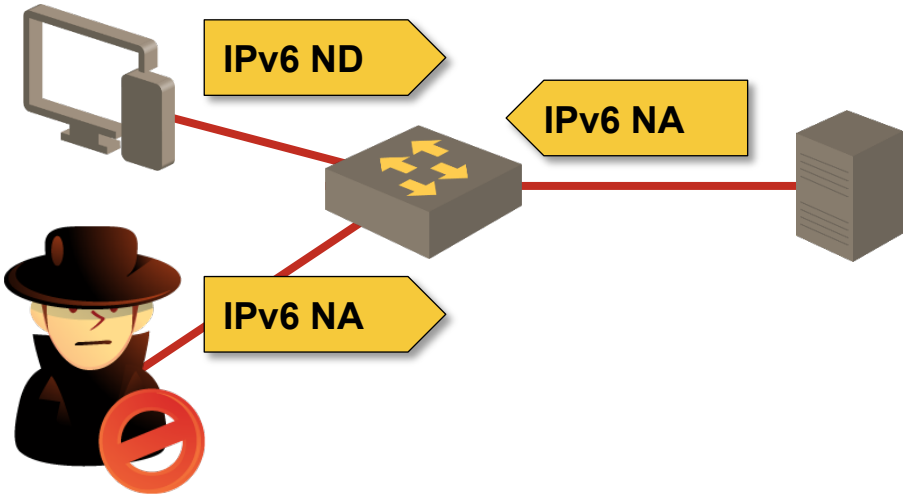
CISCO CCIE EMERITUS

vmware vEXPERT

IPv6 Layer-2 Security Challenges

The Problem



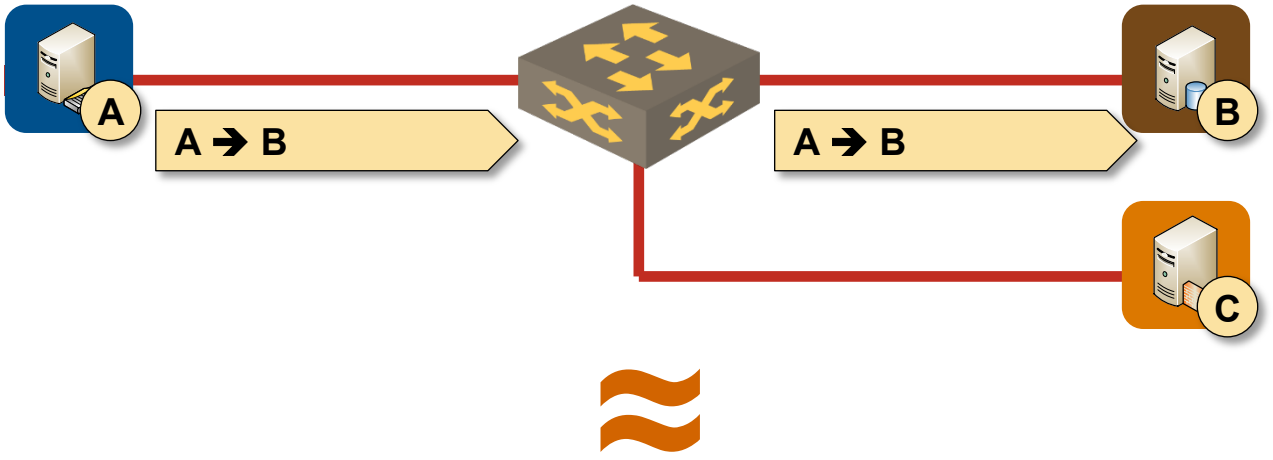


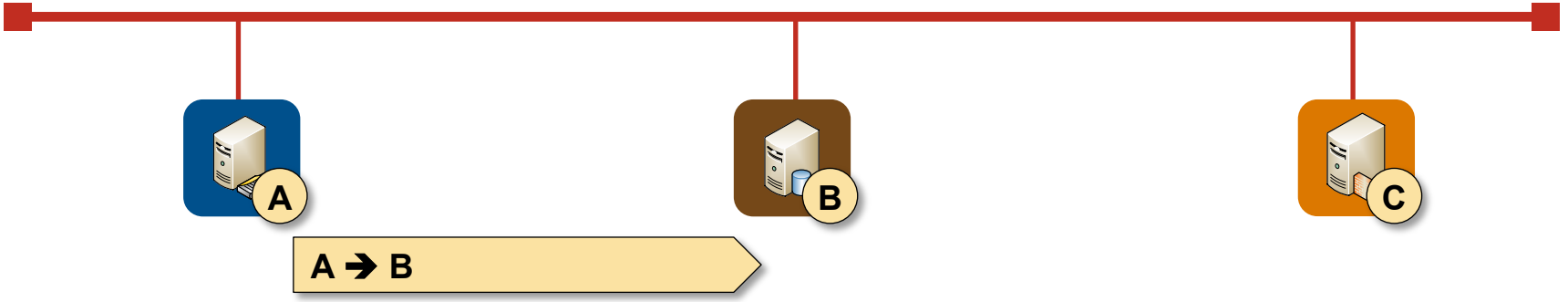
- **Assumption:** one subnet = one security zone
- **Corollary**: intra-subnet communication is not secured
- **Consequences**: multiple first-hop vulnerabilities

Sample vulnerabilities:

- RA spoofing
- NA spoofing
- DHCPv6 spoofing
- DAD DoS attack
- ND DoS attack

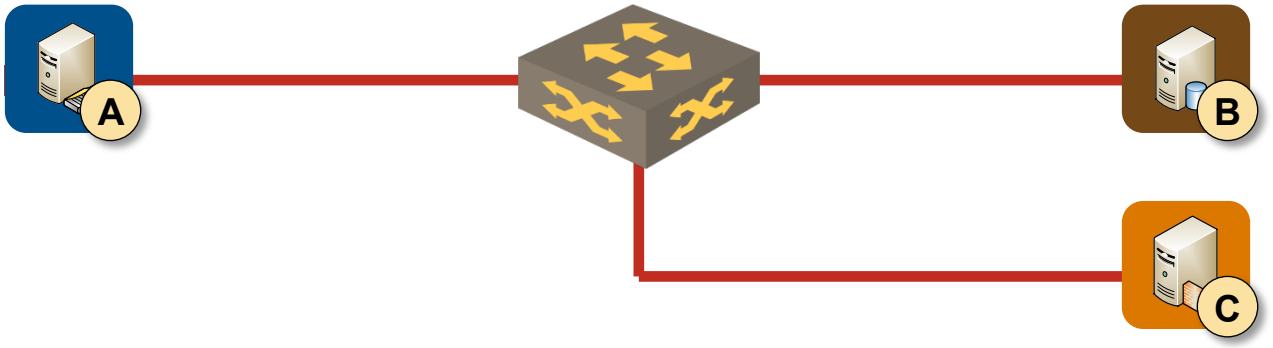
Root Cause





All LAN infrastructure we use today emulates 40 year old thick coax cable

The Traditional Fix: Add More Kludges



Typical networking industry solution

- Retain existing forwarding paradigm
- Implement layer-2 security mechanisms

Sample L2 security mechanisms

- RA guard
- DHCPv6 guard
- IPv6 ND inspection
- SAVI

Benefits

- Non-disruptive deployment (clusters and Microsoft NLB still works)
- No need to educate customers

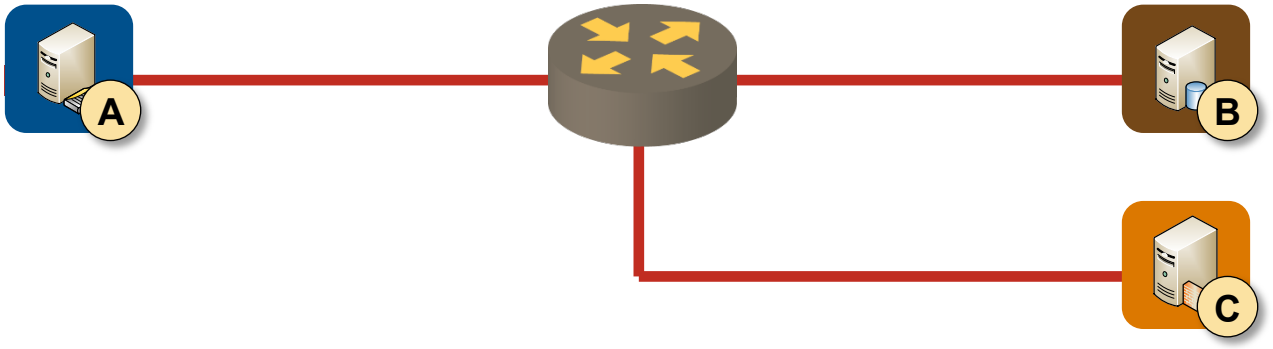
Drawbacks

- Not available on all platforms
- Expensive to implement in hardware
- Exploitable by infinite IPv6 header + fragmentation creativity

Can we do any better than that?

Layer-3-Only IPv6 Networks

Goal: Remove Layer-2 from the Network



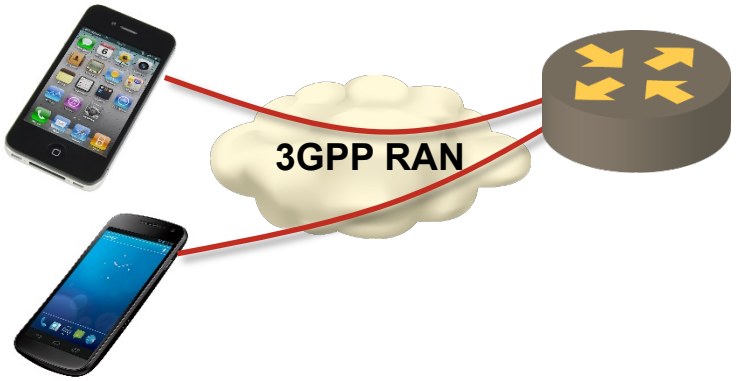
Change the forwarding paradigm

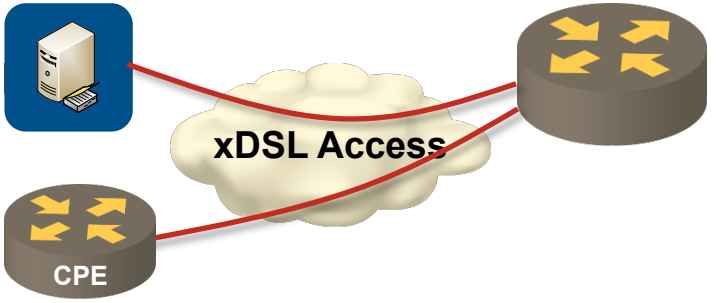
- First-hop network device is a router (layer-3 switch in marketese)
- Fake router advertisements or ND/NA messages are not propagated to other hosts

Simplistic implementation

- Every host is in a dedicated /64 subnet
- Default behavior on 3GPP and xDSL networks
- Somewhat harder to implement on Carrier Ethernet, hard on cable networks

IPv6 over 3GPP and PPPoX Networks





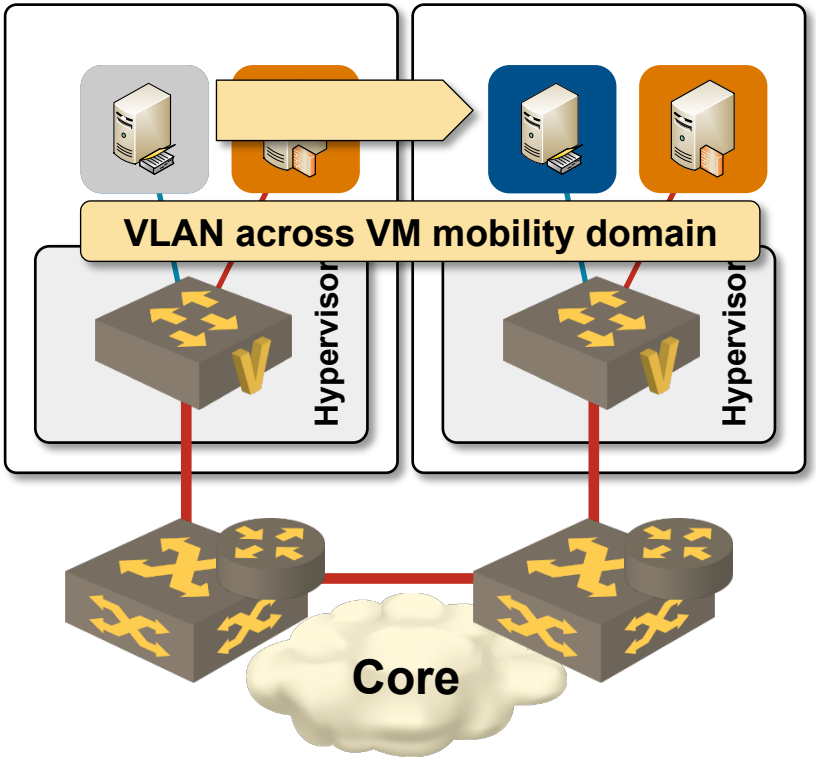
- Each device-to-network connection is a separate dial-up interface on BRAS/GGNS
- Customer device (phone, computer, CPE) interacts directly with the first-hop router
- A /64 subnet is allocated to each dial-up interface (usually from a local pool)
- Aggregate IPv6 prefix is advertised to the network core to minimize number of prefixes advertised in the core

Data Center Considerations

Implications of Live VM Mobility

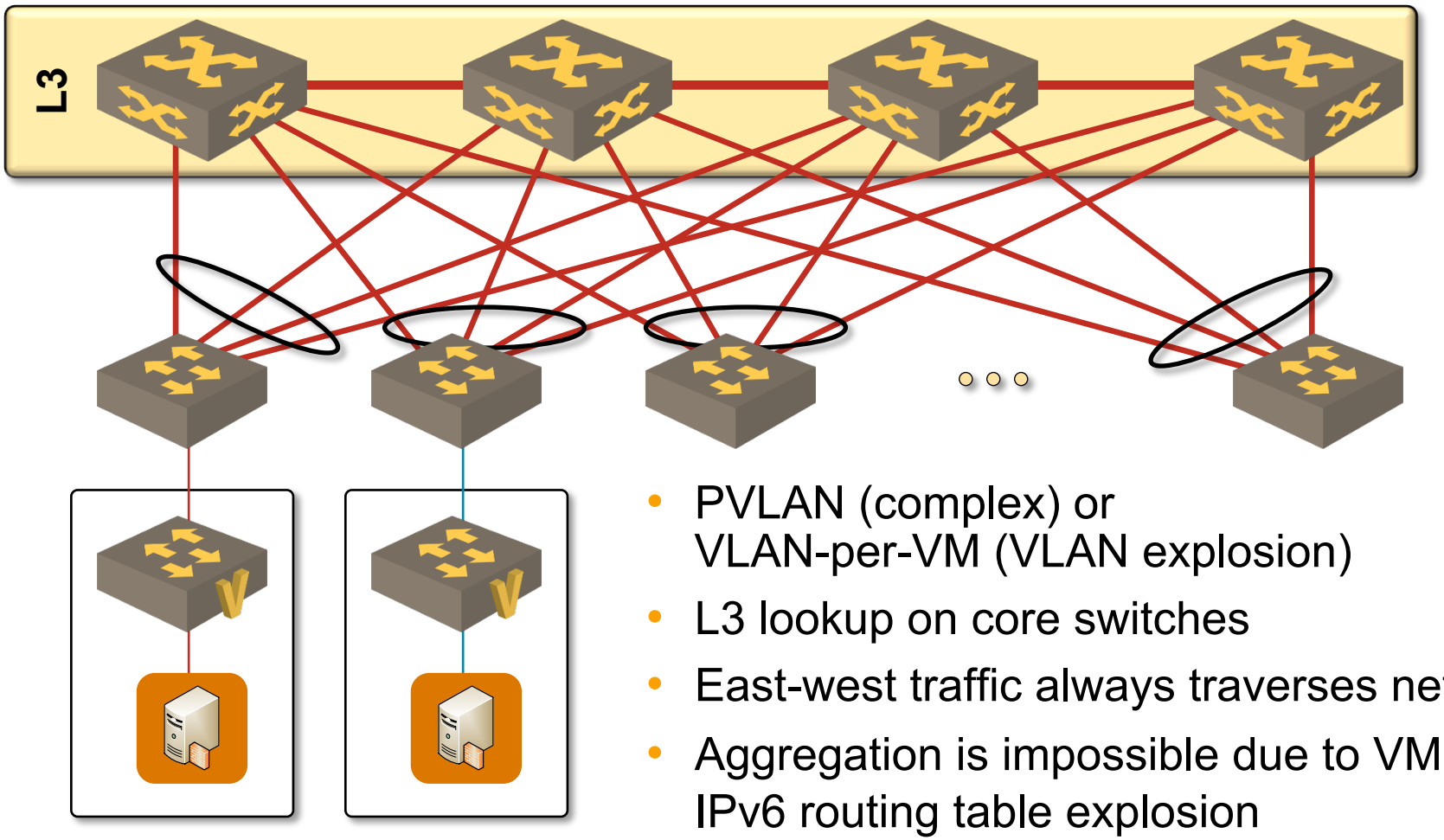
Challenges

- VM moved to another server must retain its IPv6 address and all data sessions
- Existing L3 solutions are too slow for non-disruptive VM moves
- Live VM mobility usually relies on L2 connectivity between physical servers
- Large VLANs must span the whole VM mobility domain



More details in *VMware Networking* and *Cloud Networking* webinars

Live VM Mobility with IPv6 Microsegmentation

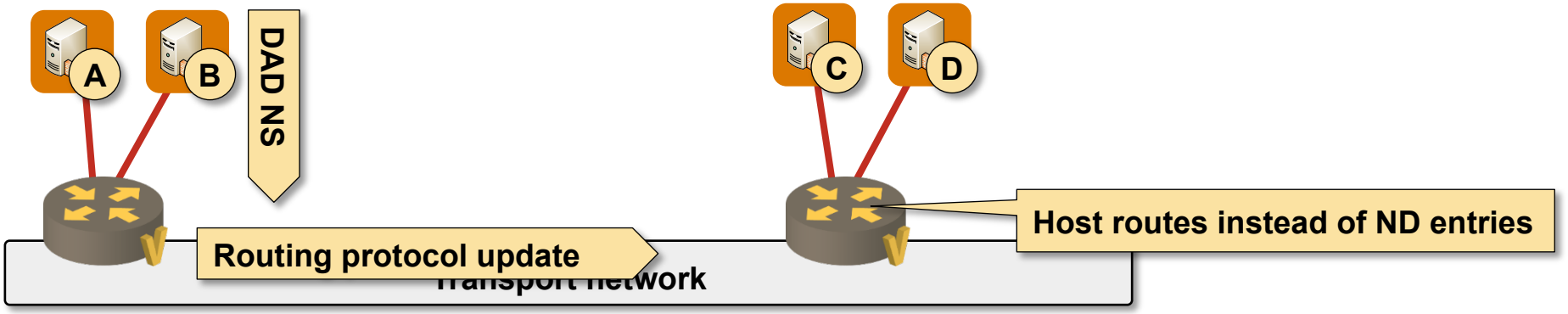


- PVLAN (complex) or
  VLAN-per-VM (VLAN explosion)
- L3 lookup on core switches
- East-west traffic always traverses network core
- Aggregation is impossible due to VM mobility ➔
  IPv6 routing table explosion

We need something better in data centers

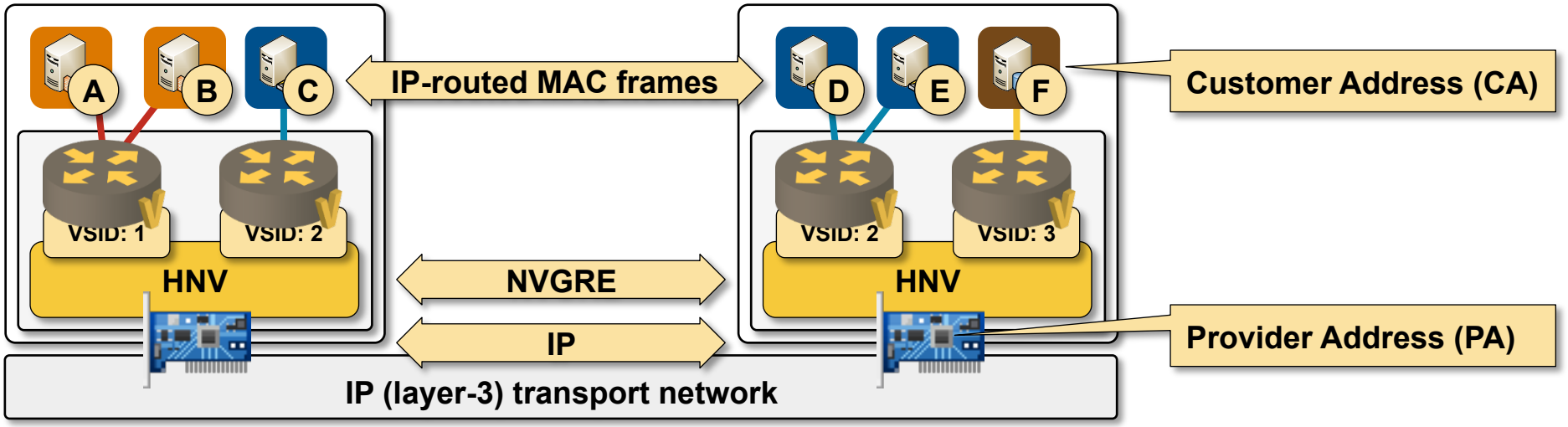
Thinking Outside of the Box

Intra-Subnet (Host Route) Layer-3 Forwarding



- Hosts are connected to layer-3 switches (routers)
- Numerous hosts share a /64 subnet
  ➔ a /64 subnet spans multiple routers
- First-hop router creates a host route on DAD, ND or DHCPv6 transaction
- IPv6 host routes are propagated throughout the local routing domain
- Host-side IPv6 addressing and subnet semantics are retained
- IPv6 ND entries are used instead of IPv6 routing table entries

Example: Hyper-V Network Virtualization



Full layer-3 switch in the hypervisor (distributed routing functionality)

- L3-only switching for intra-hypervisor and inter-hypervisor traffic
- IPv4 and IPv6 support in customer (virtual) and provider (transport) network
- ARP and ND proxies ➔ no ARP or unknown unicast flooding
- Source node flooding or Customer ➔ Provider IP multicast mapping

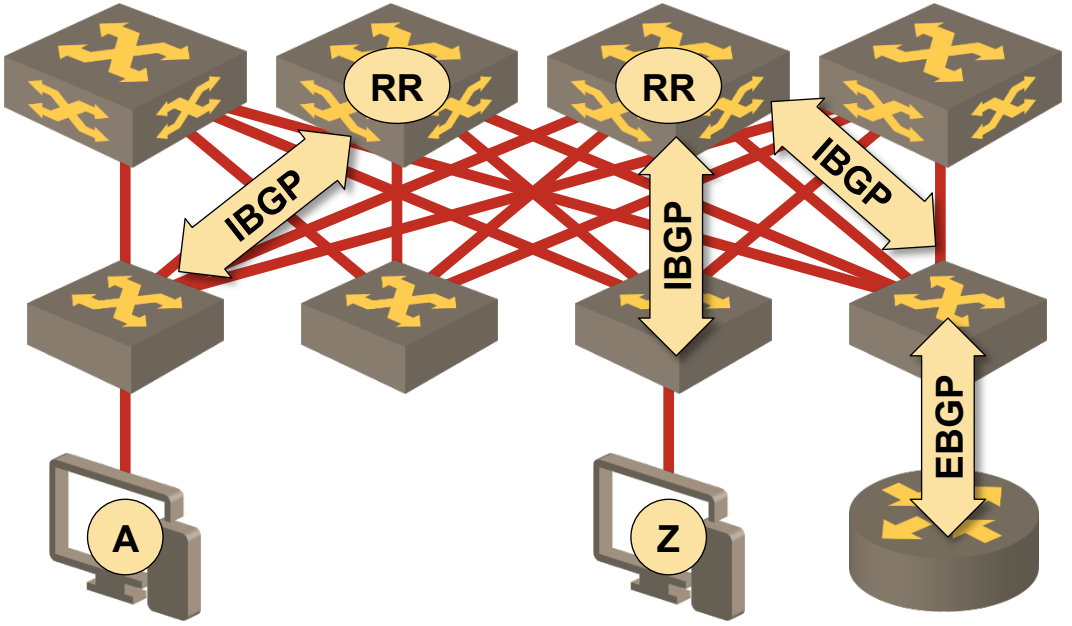
IP Routing in Cisco Dynamic Fabric Automation (DFA)

IP routing information distribution

- Host routes generated from ARP/ND/DHCP information or based on VDP messages (Nexus 1000v only)
- Subnet routes generated from configuration information
- External routes learned through routing protocols
- All IP routes inserted into MP-BGP and distributed across fabric

Each fabric node knows

- All intra-fabric host routes
- All intra-fabric subnets
- All external routes



Summary

IPv6 Microsegmentation Solutions

Why?

- Removes first-hop (L2) IPv6 security challenges

How?

- Dedicated dynamic interface per host (mobile, PPPoX)
- Dedicated VLAN per host (Carrier Ethernet, campus, data center)
- Host routing

Implementations of Host Route-Based Forwarding

IPv6 and IPv4

- Hyper-V Network Virtualization
- Juniper Contrail
- Cisco Dynamic Fabric Automation (DFA)

IPv4 only

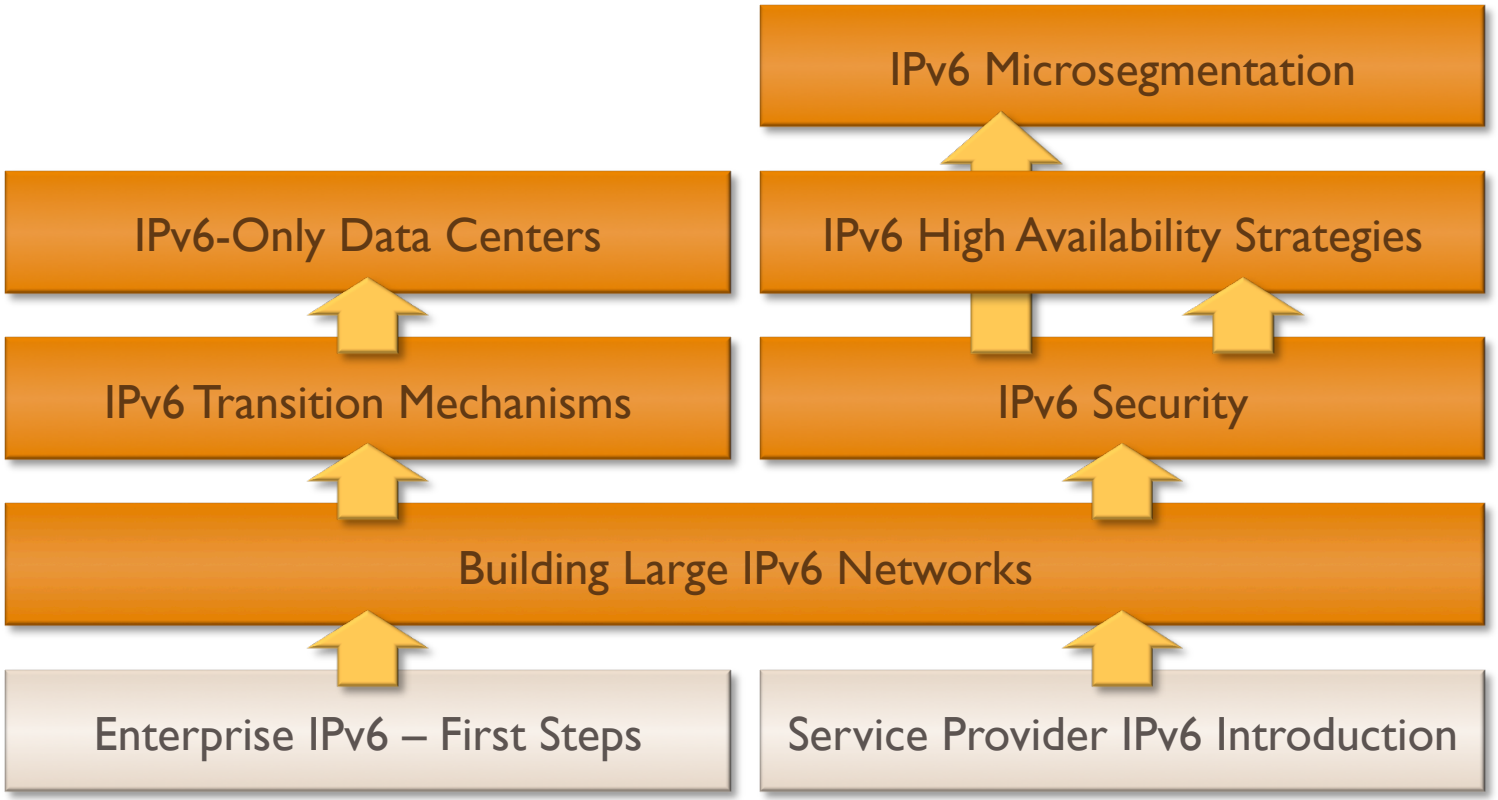
- Nuage Virtual Services Platform (VSP)
- Cisco Application Centric Infrastructure (ACI)

Unrelated honorable mention

- IPv6 RA guard and ND inspection implemented on VMware NSX

Hint: vote with your wallet!

More Information: IPv6 Webinars on ipSpace.net



Availability

- Live sessions
- Recordings of individual webinars
- Yearly subscription

Other options

- Customized webinars
- ExpertExpress
- On-site workshops

Stay in Touch

Web:	ipSpace.net
Blog:	blog.ipSpace.net
Email:	ip@ipSpace.net
Twitter:	@ioshints

IPv6:	ipSpace.net/IPv6
Webinars:	ipSpace.net/Webinars
Consulting:	ipSpace.net/Consulting