TOMORROW starts here.

CISCO

Cisco live!

Integration of Multi-Hypervisors with Application Centric Infrastructure

BRKAPP-9005

Bradley Wong
Principal Engineer

Cisco live!

“ …The Application Centric Infrastructure (ACI) is adopting an innovative approach to addressing these challenges, through normalisation of different hypervisor encapsulations together with tight integration of the Virtual Machine Manager (VMM) of choice, providing a single point of management for both physical and virtual infrastructure as well as the applications that run on top of them. This session will address how the ACI fabric handles single and multi-hypervisor environments, and how the ACI controller provides integration into different VMMs for a single point of management...”

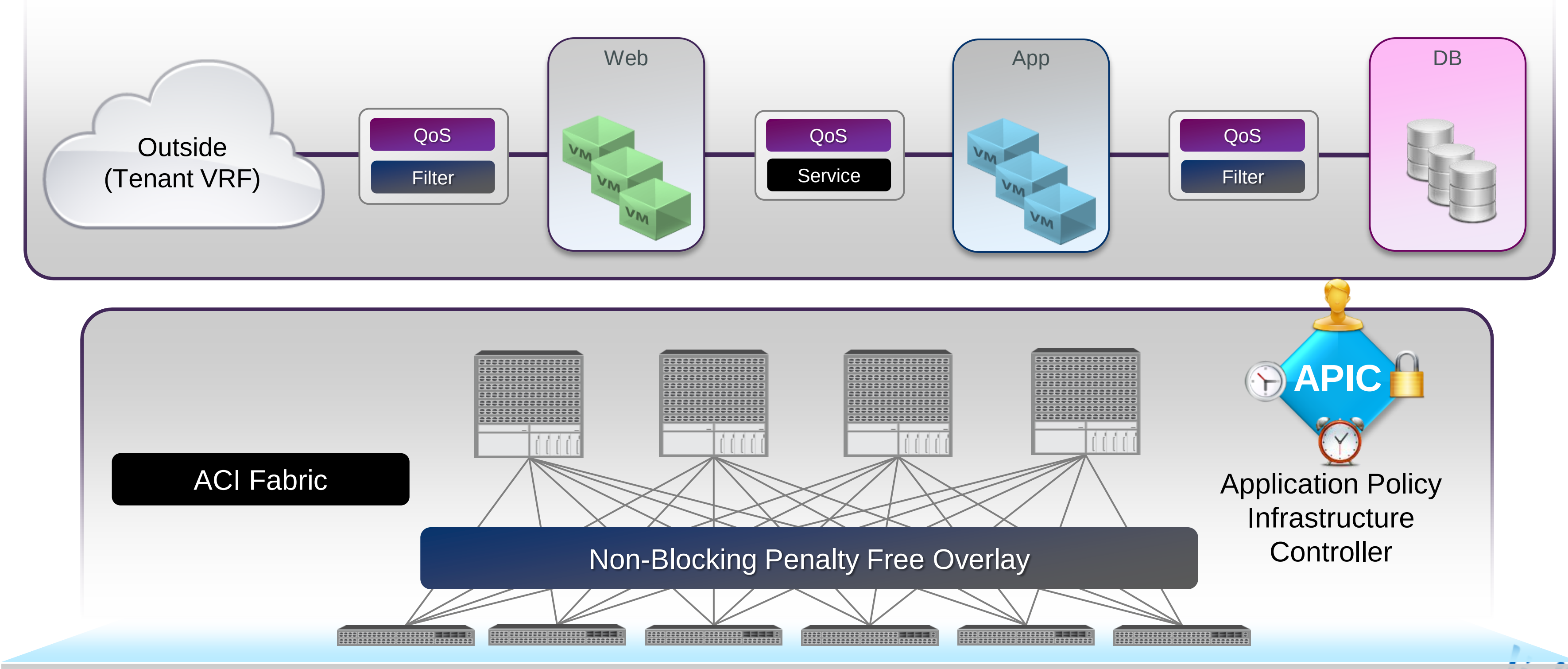
BRKAPP-9005 ABSTRACT




Introduction to ACI

Cisco ACI

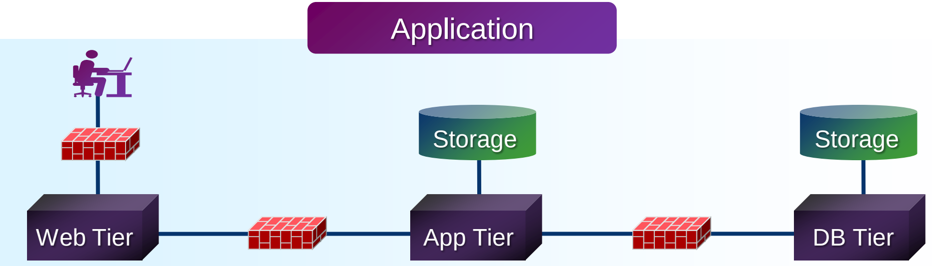
Logical Network Provisioning of Stateless Hardware



ACI Network Profile

Policy-Based Fabric Management

- Extend the principle of Cisco UCS® Manager service profiles to the entire fabric
- Network profile: stateless definition of application requirements
  - Application tiers
  - Connectivity policies
  - Layer 4 – 7 services
  - XML/JSON schema
- Fully abstracted from the infrastructure implementation
  - Removes dependencies of the infrastructure
  - Portable across different data centre fabrics



The network profile fully describes the application connectivity requirements

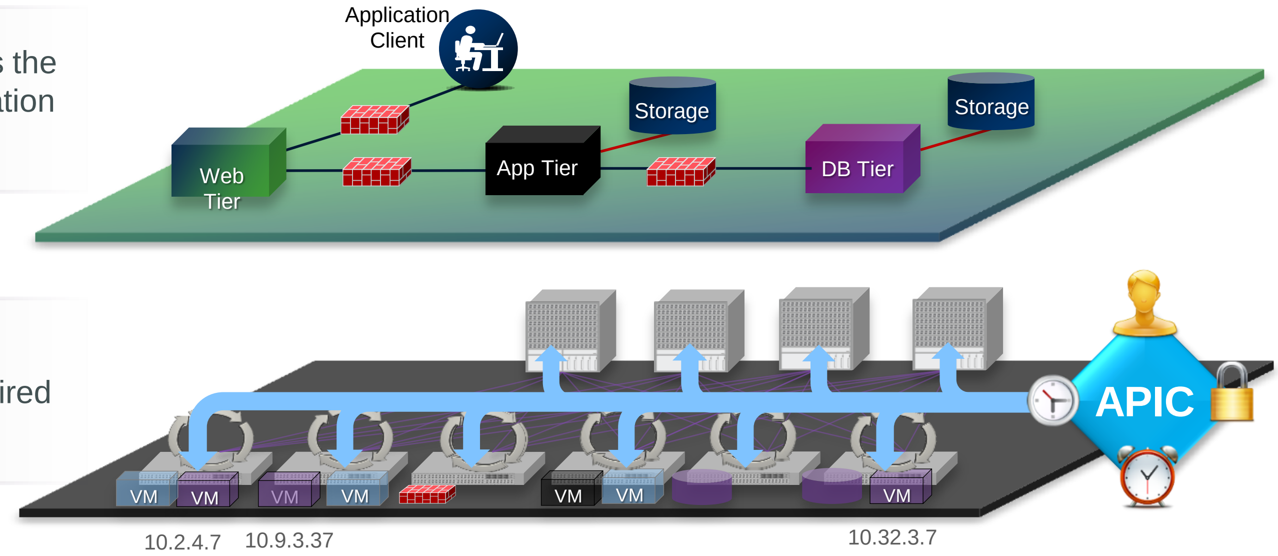
```
## Network Profile: Defines Application Level Metadata (Pseudo Code Example)

<Network-Profile = Production_Web>
  <App-Tier = Web>
    <Connected-To = Application_Client>
        <Connection-Policy = Secure_Firewall_External>
   <Connected-To = Application_Tier>
        <Connection-Policy = Secure_Firewall_Internal & High_Priority>
. . .
 <App-Tier = DataBase>
    <Connected-To = Storage>
        <Connection-Policy = NFS_TCP & High_BW_Low_Latency>
. . .
```

Application Policy Model and Instantiation

Application policy model: Defines the application requirements (application network profile)

Policy instantiation: Each device dynamically instantiates the required changes based on the policies



All forwarding in the fabric is managed through the application network profile

- IP addresses are fully portable anywhere within the fabric
- Security and forwarding are fully decoupled from any physical or virtual network attributes
- Devices autonomously update the state of the network based on configured policy requirements

Application Awareness

Application-Level Visibility

ACI Fabric provides the next generation of analytic capabilities

Per application, tenants, and infrastructure:

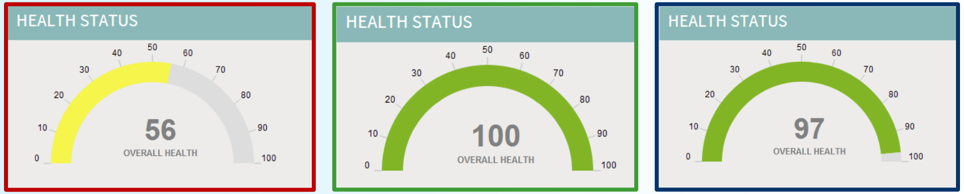
- Health scores
- Latency
- Atomic counters
- Resource consumption

Integrate with workload placement or migration

PetStore Event

Triggered Events or Queries

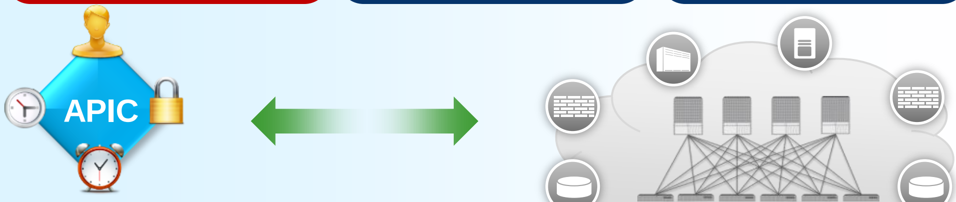
Actions:
No new hosts or VMs
Evacuate hypervisors
Re-balance clusters



PetStore Dev
- Leaf 1 and 2
- Spine 1 – 3
- Atomic counters

PetStore Prod
- Leaf 2 and 3
- Spine 1 – 2
- Atomic counters

PetStore QA
- Leaf 3 and 4
- Spine 2 – 3
- Atomic counters



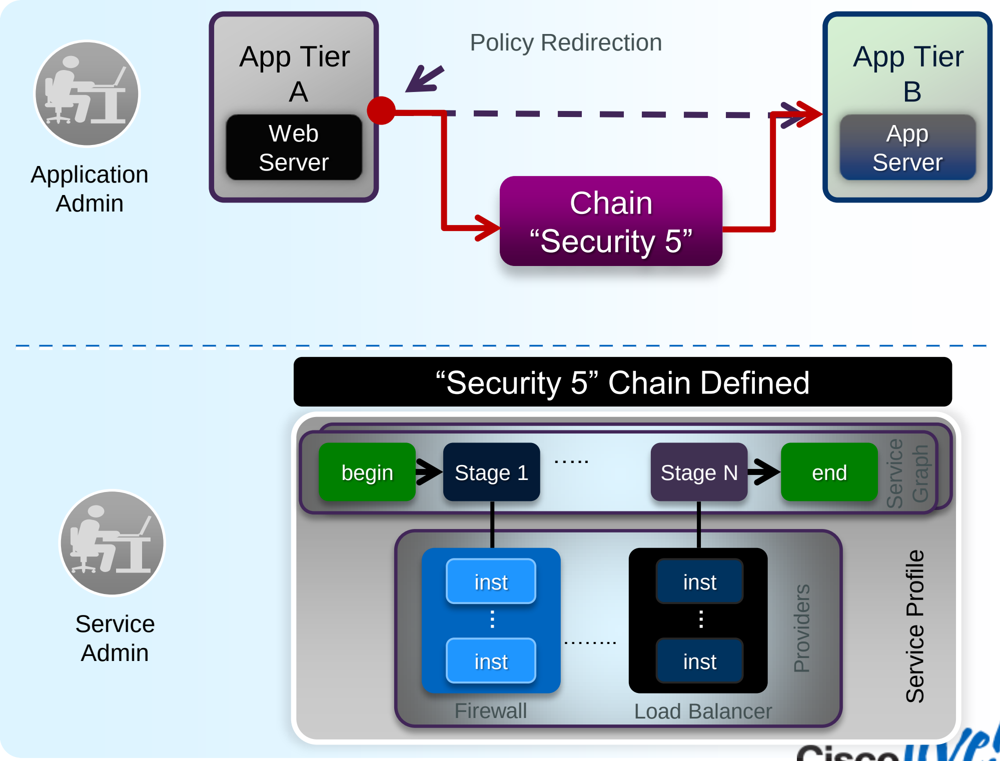
VXLAN
Per-Hop Visibility

Physical and
Virtual as One

ACI Layer 4 - 7 Service Integration

Centralised, Automated, And Supports Existing Model

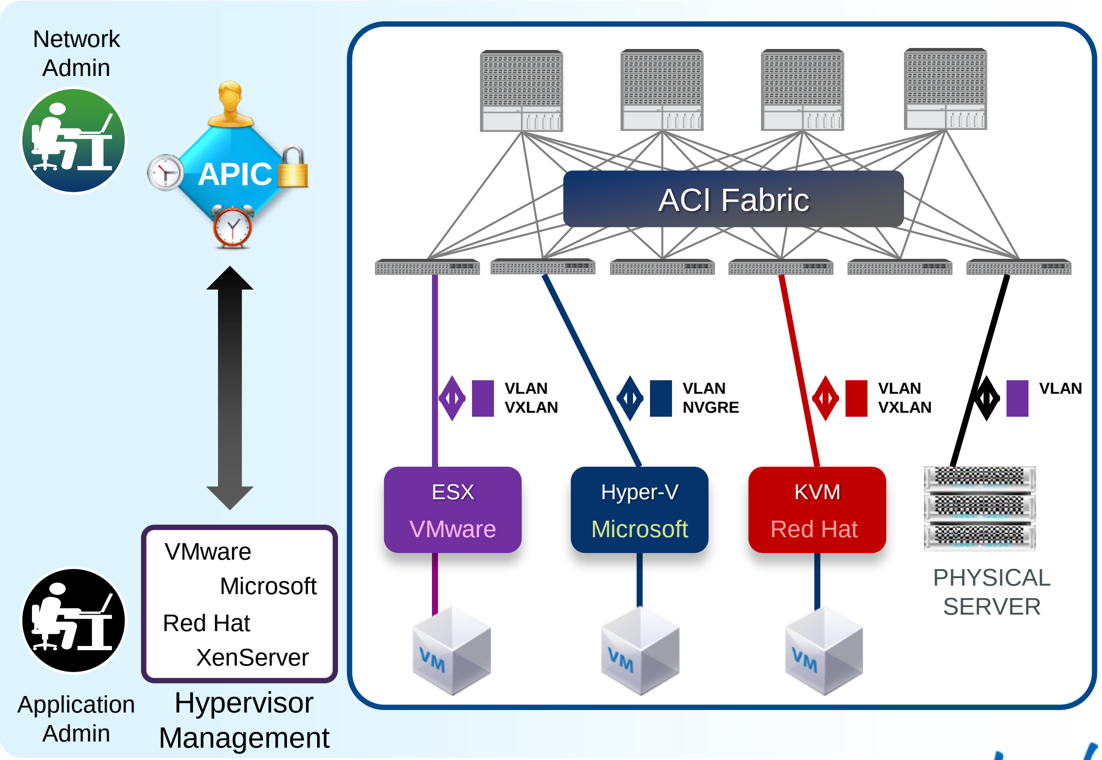
- Elastic service insertion architecture for physical and virtual services
- Helps enable administrative separation between application tier policy and service definition
- APIC as central point of network control with policy coordination
- Automation of service bring-up/tear-down through programmable interface
- Supports existing operational model when integrated with existing services
- Service enforcement guaranteed, regardless of endpoint location



Multi-Hypervisor-Ready Fabric

Virtual Integration

- Integrated gateway for VLAN, VxLAN, and NVGRE networks from virtual to physical
- Normalisation for NVGRE, VXLAN, and VLAN networks
- Customer not restricted by a choice of hypervisor
- Fabric is ready for multi-hypervisor



Open Ecosystem Framework

Full-Featured, Programmable API And Data Model

Northbound API

- Rapid integration with existing management frameworks
- OpenStack
- Tenant- and application-aware

System Management: HP, NetQoS, CA Technologies, SolarWinds, Arbor Networks, Tivoli Software, NetBrain, InfoVista

Automation Tools: Puppet Labs, Opscode, Python, CFEngine

Hypervisor Management: VMware, XenServer, Microsoft, Red Hat KVM

Orchestration Frameworks: CloudStack, OpenStack, VMware, Nebula, Eucalyptus

Object-Oriented
Centralised Automation
RESTful XML/JSON

Open Ecosystem Framework

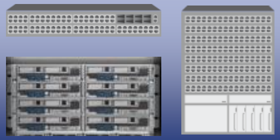
Comprehensive Programmability and System Access

Southbound API

- Publish data model
- Open source
- Enables application portability

Microsoft

XenServer



Red Hat KVM

*Only straight chains supported at FCS

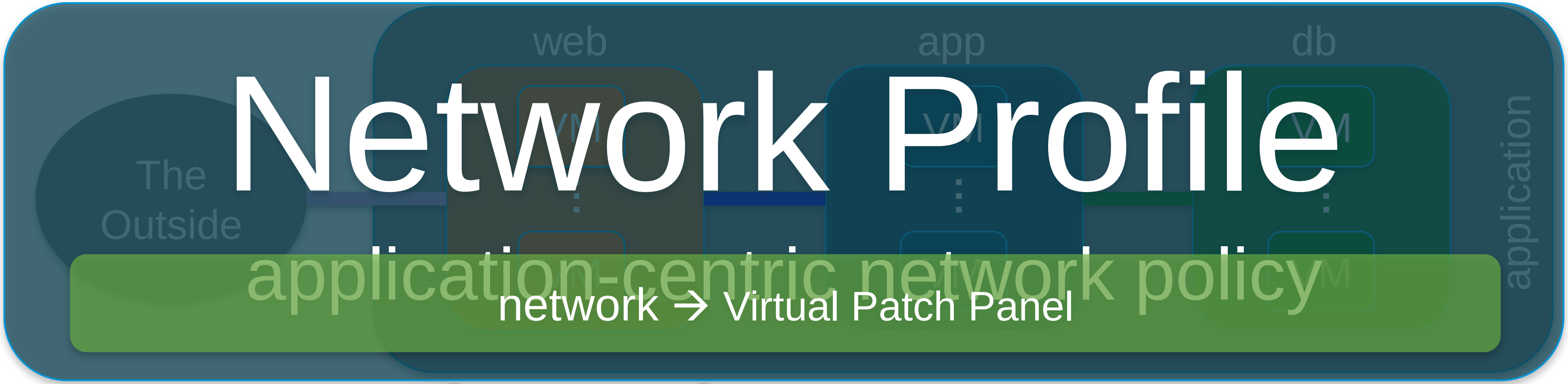



ACI Fabric Policy Constructs

rules of how application communicates to the external private or public networks

a set of network requirements specifying how application components communicate with each other

Contract
Access Control
QoS
Network Services

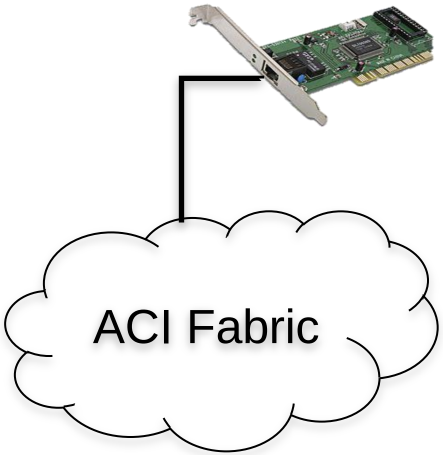


a collection of end-points connecting to the network… VMs, physical compute, …

Component
Tier
End Point Group

End-points

→ Things that connect to the fabric and use it to interface with other things
→ A compute, storage or service instance attaching to a fabric



NIC
vNIC
.
.
.

end-points [**EP**]

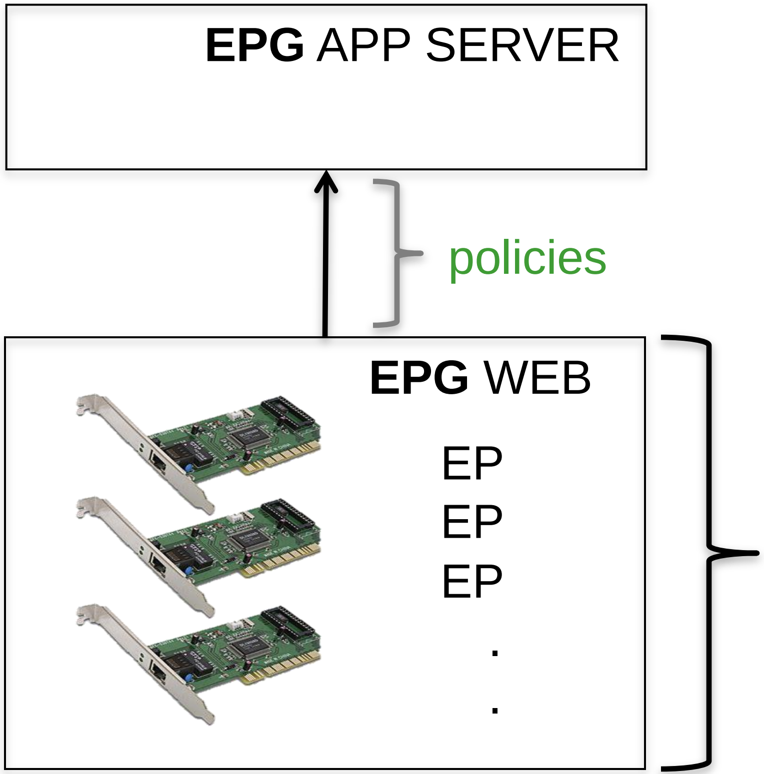
End-points

→ Things that connect to the fabric and use it to interface with other things
→ A compute, storage or service instance attaching to a fabric

EP
EP
EP
.
.
.

A collection of end-points with identical network behaviour form a … ***End Point Group (EPG)***

End-point Groups (EPGs)



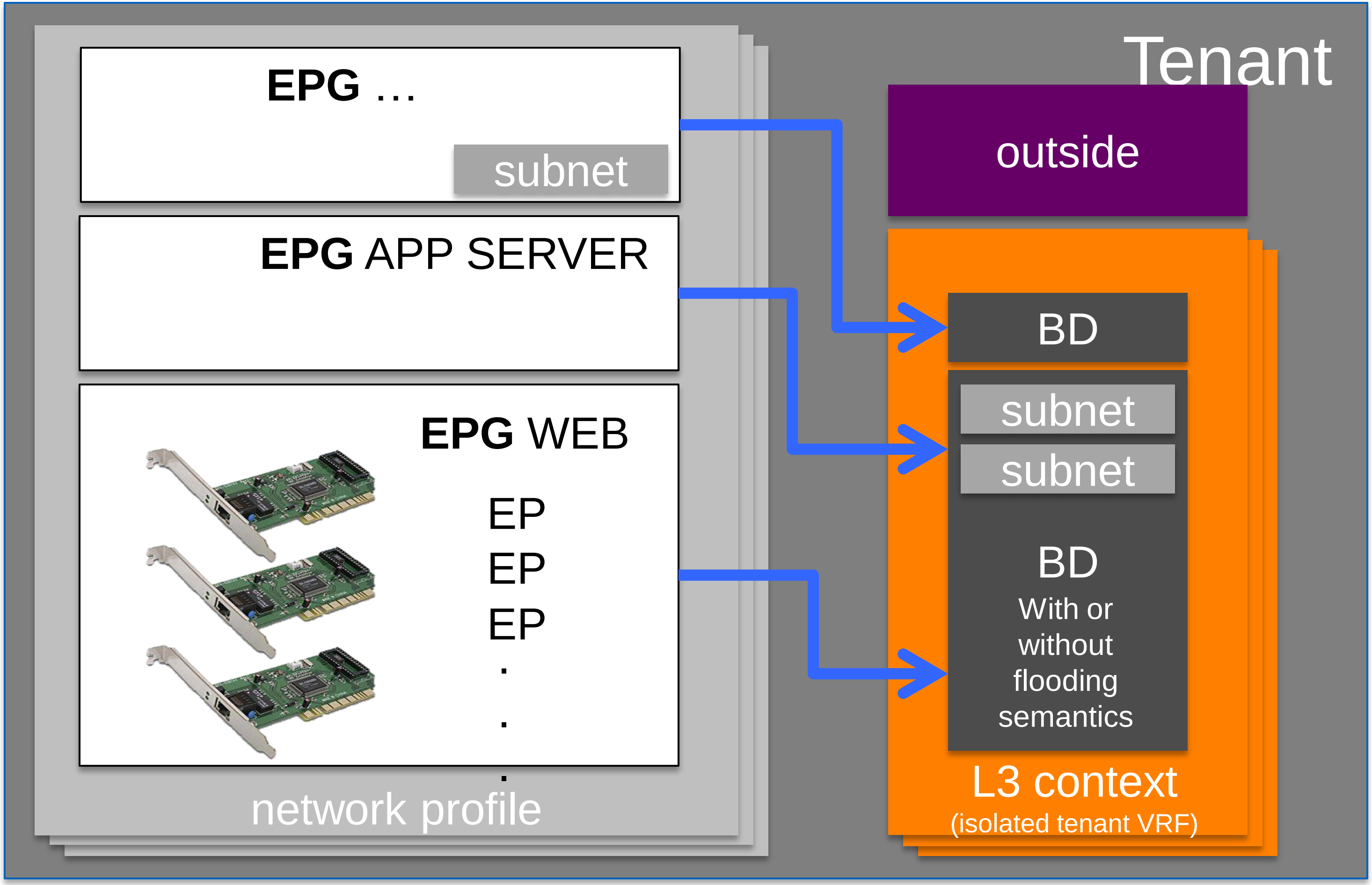
Allows to specify rules and policies on groups of physical or virtual end-points without understanding of specific identifiers and regardless of physical location.

Can flexibly map into

→application tier of multi-tier app
→segmentation construct (ala VLAN)
→a security construct
→ESX port group
→…

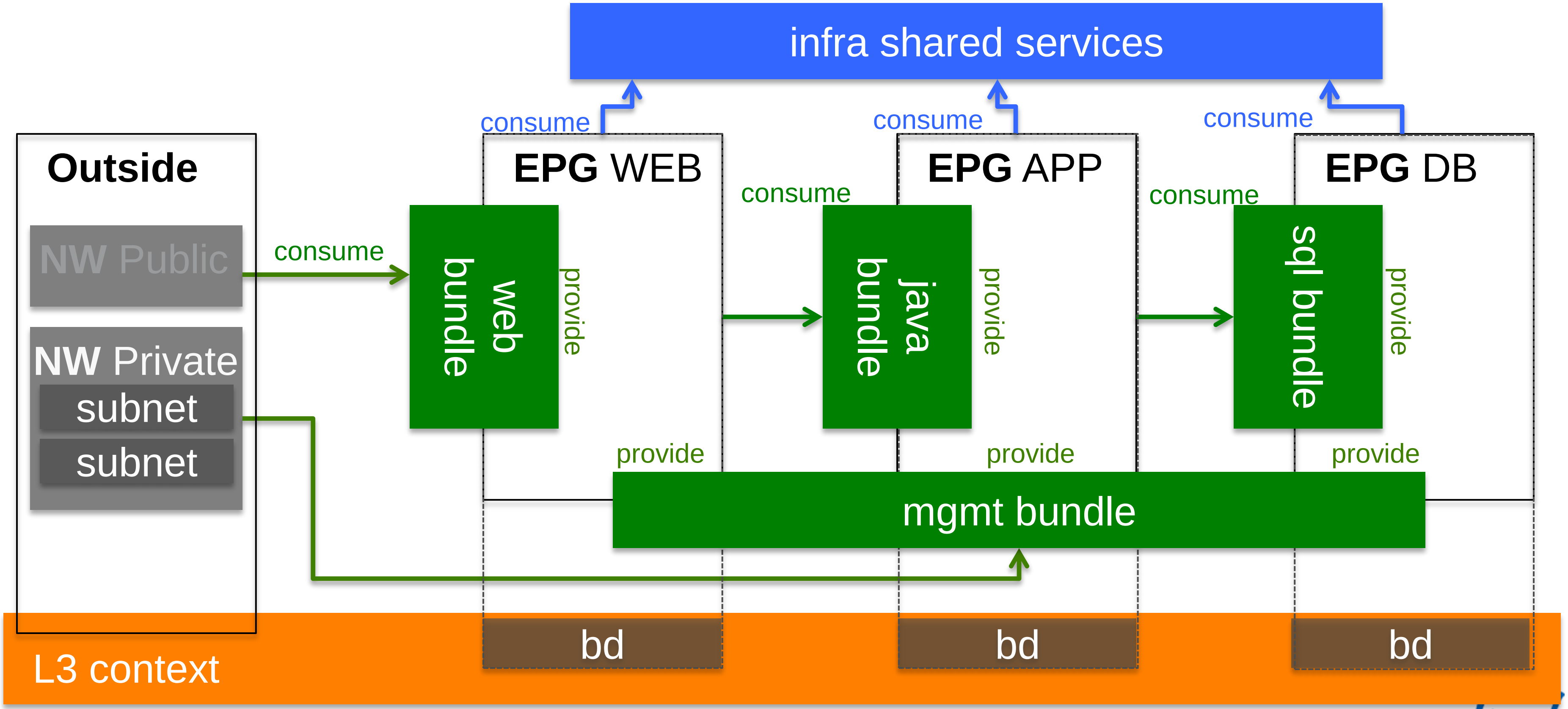
… **end-point group** [***EPG***]

Tenant L3, L2 Isolation



self-contained tenant definition representable as a recursive structured text document

EXAMPLE: Three-tier APP

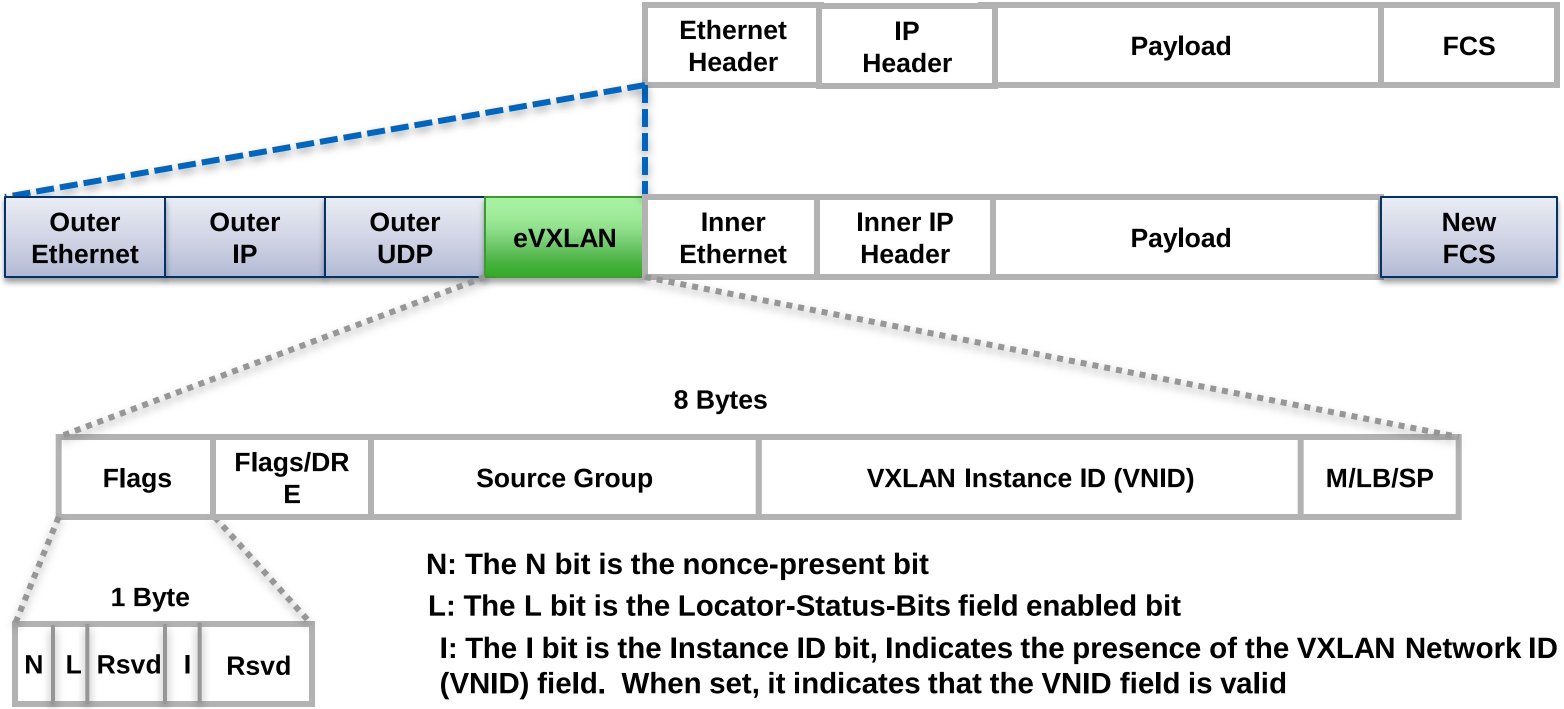





Overview of Hypervisor Integration

ACI Fabric Architecture

ACI VXLAN (eVXLAN) Header



N: The N bit is the nonce-present bit

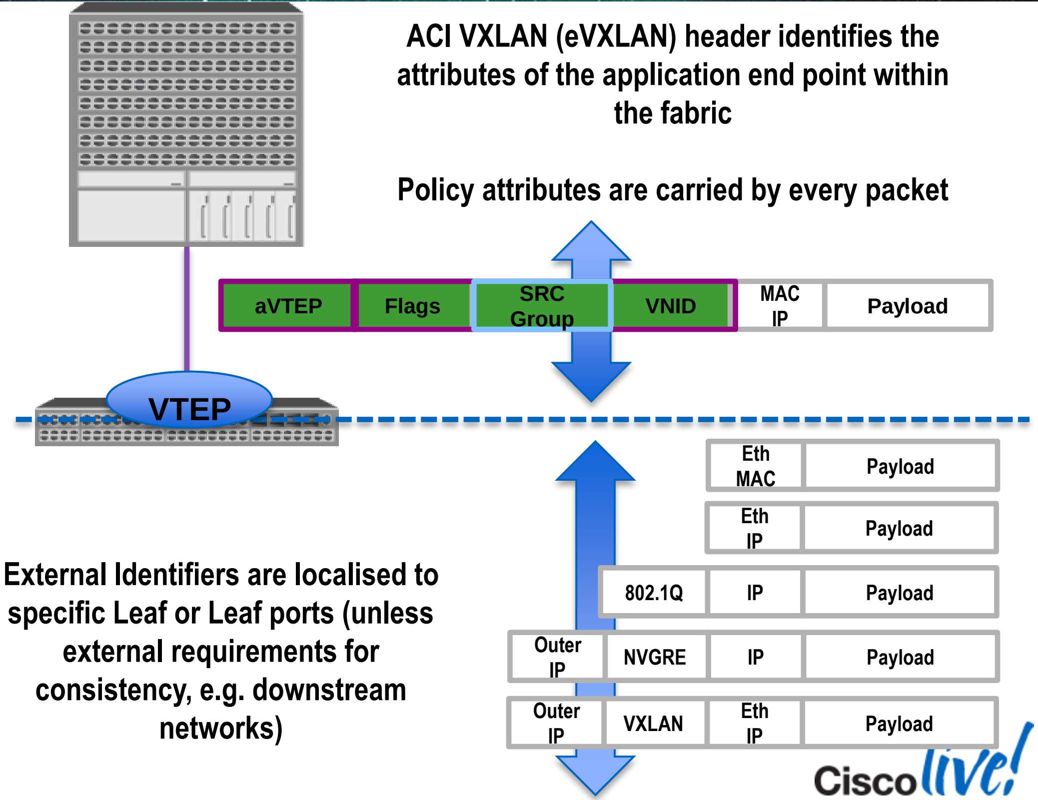
L: The L bit is the Locator-Status-Bits field enabled bit

I: The I bit is the Instance ID bit, Indicates the presence of the VXLAN Network ID (VNID) field. When set, it indicates that the VNID field is valid

ACI Fabric – Integrated Overlay

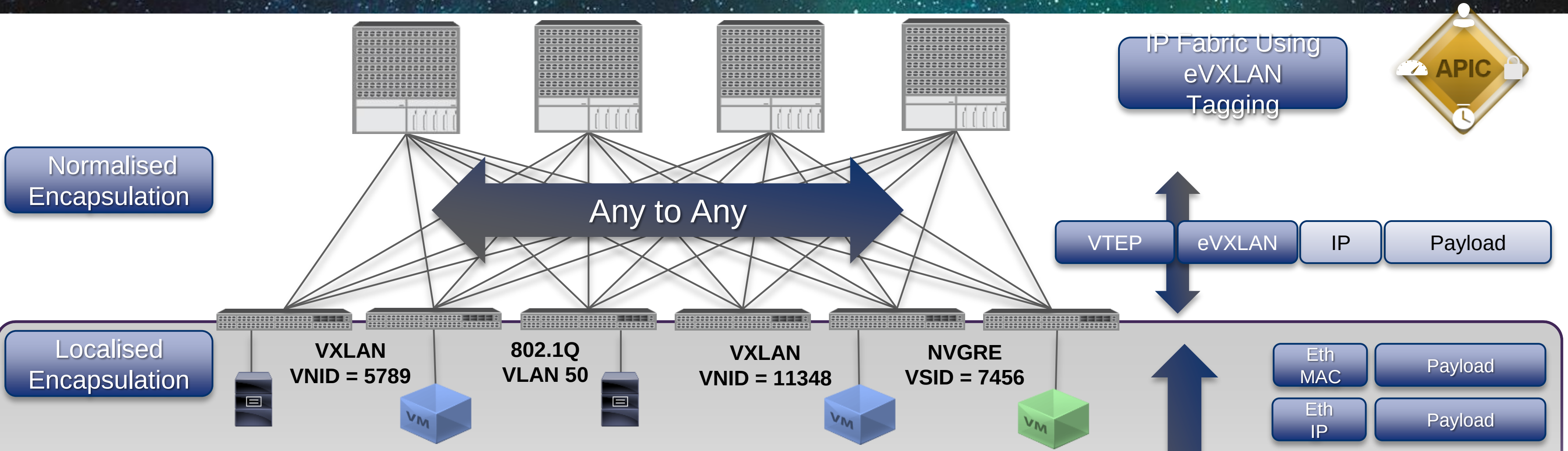
ACI VXLAN (eVXLAN) Header

- All Tenant traffic within the Fabric is tagged with an ACI VXLAN (eVXLAN) header which identifies the policy attributes of the application end point within the fabric
- At the ingress port the Fabric translates an external identifier which can be used to distinguish different application end points via the ACI eVXLAN tagging format



ACI Fabric – Integrated Overlay

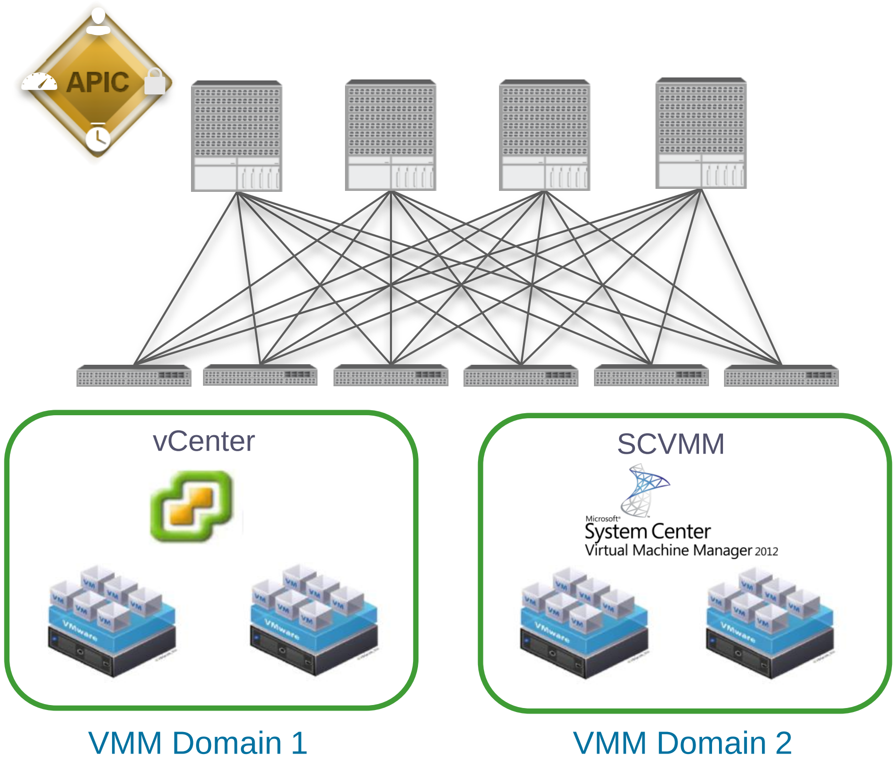
Multi-Hypervisor Encapsulation Normalisation



- All traffic within the ACI Fabric is encapsulated with an extended VXLAN (eVXLAN) header
- External VLAN, VXLAN, NVGRE tags are mapped at ingress to an internal eVXLAN tag
- Forwarding is not limited to, nor constrained within, the encapsulation type or encapsulation ‘overlay’ network
- External identifies are localised to the Leaf or Leaf port, allowing re-use and/or translation if required

Hypervisor Integration with ACI

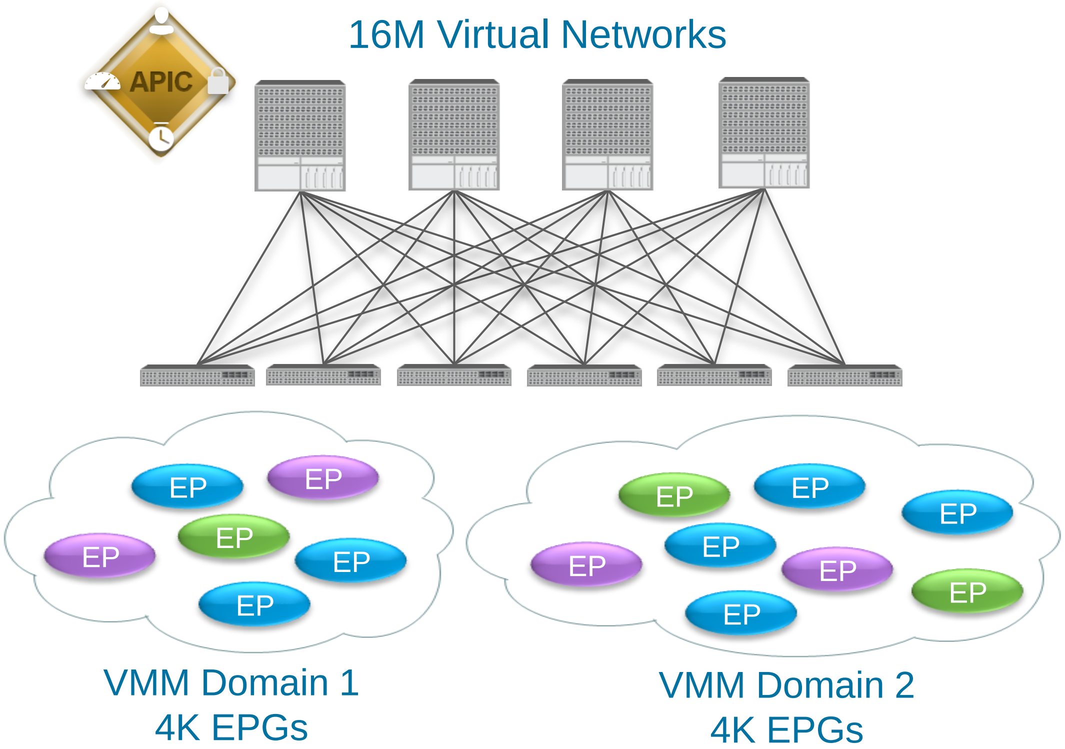
VMM Domains



- Multiple Virtual Machine Managers (VMMs) likely on a single Fabric
- Each VMM and associated Virtual hosts are grouped within APIC
- Called VMM Domain

Hypervisor Integration with ACI

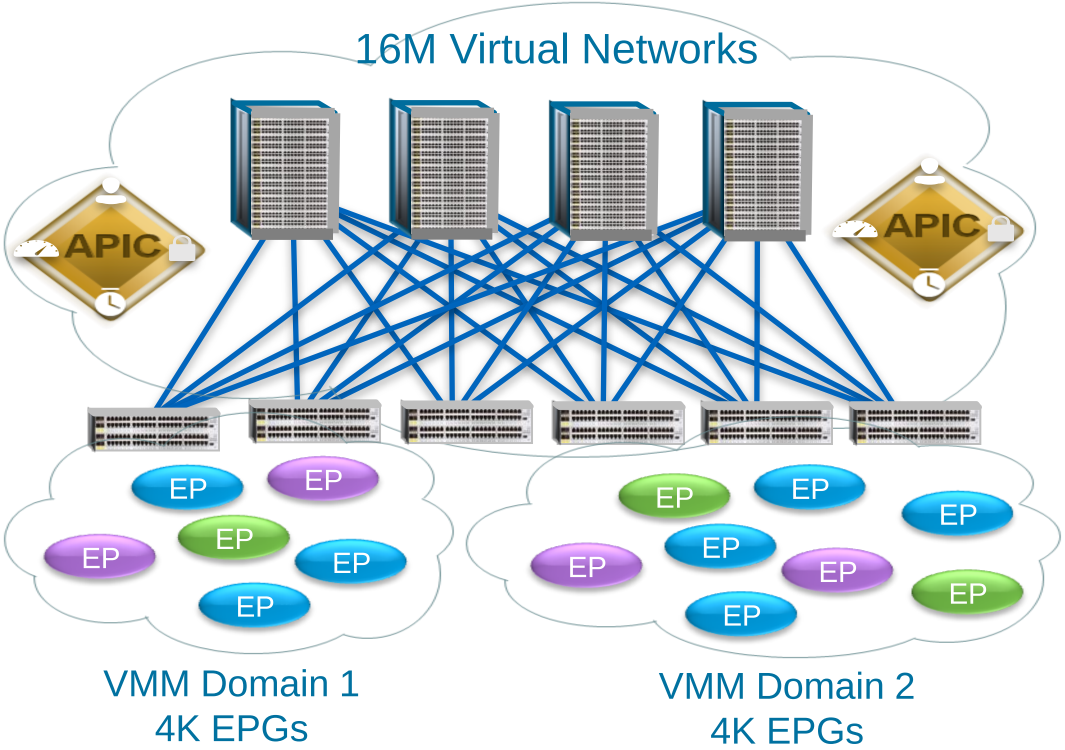
VMM Domains & VLANs



- VLAN ID only gives 4K EPGs (12 bits)
- Scale by creating "**pockets**" of 4K EPGs
- Map to scope of live migration
- Place VM anywhere
- Live migrate within VMM domain

Hypervisor Integration with ACI

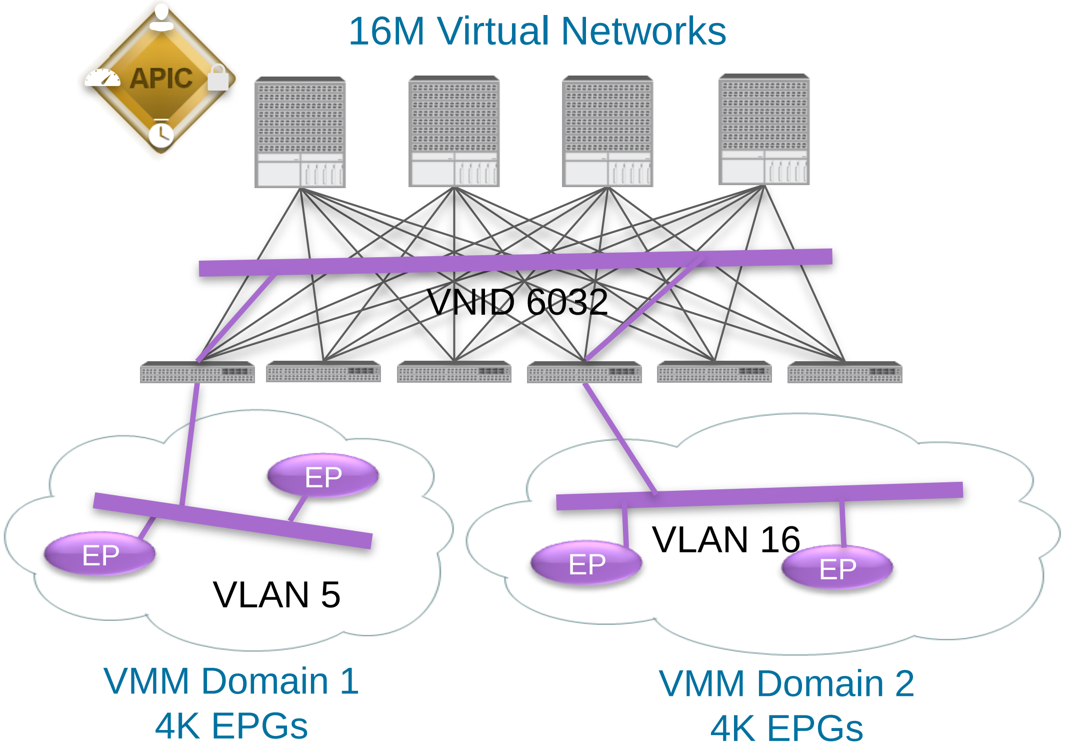
VMM Domains



- VLAN ID only gives 4k EPGs (12 bits)
- Scale by creating "**pockets**" of 4k EPGs
- Map to scope of live migration
- Place VM anywhere
- Live migrate within VMM domain

Hypervisor Integration with ACI

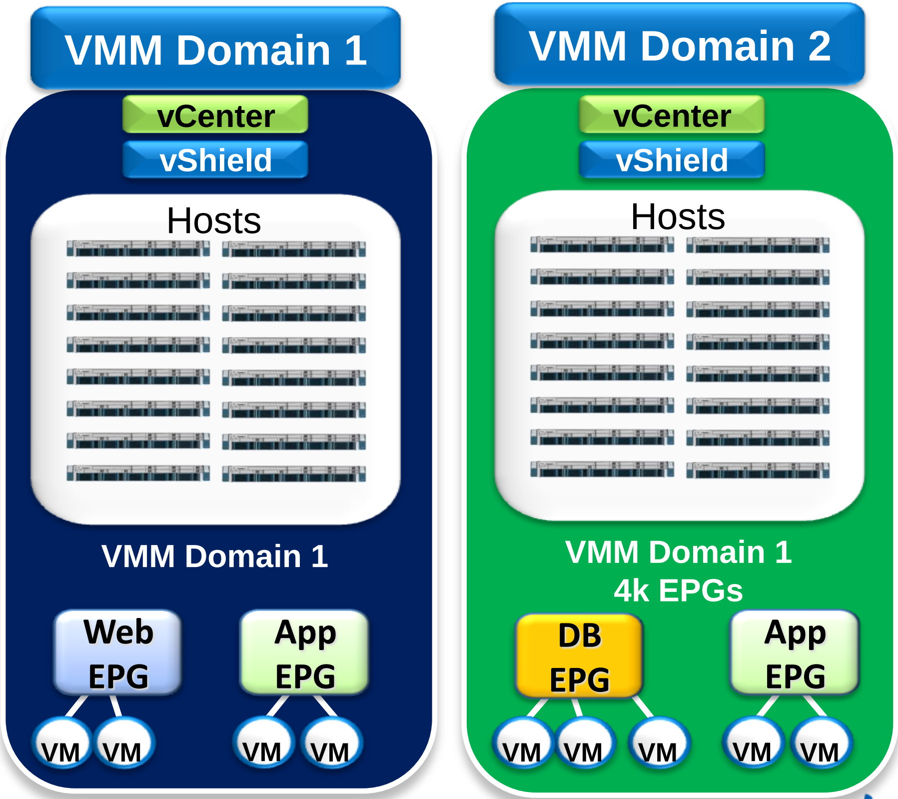
VMM Domains & VLANs



- VLAN ID only gives 4K EPGs (12 bits)
- Scale by creating "**pockets**" of 4K EPGs
- Map to scope of live migration
- Place VM anywhere
- Live migrate within VMM domain

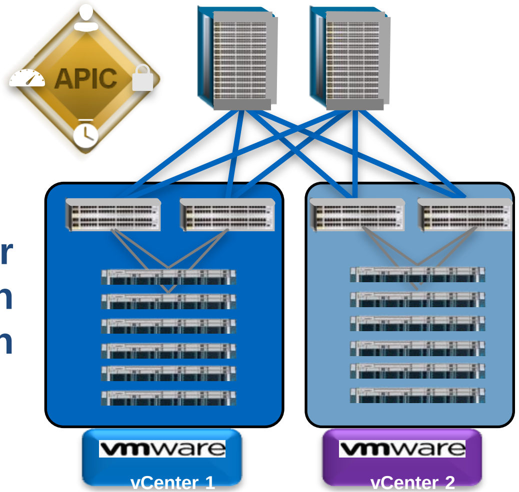
EPG Spanning Across VMM Domains

- EPGs can take different network identities across VMM Domain
- Applications can be deployed across VMM Domains
- VM Mobility is not allowed between VMM Domain due to vCenter/SCVMM limitation

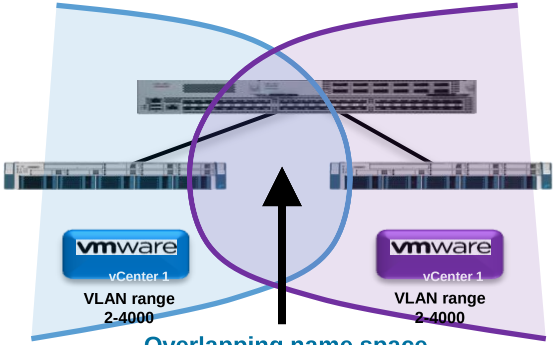


Recommended Practice for VLAN Networks

- Well separated VMM Domains
- Separate VLAN name space when VMM domains share TOR



Best Practice for VMM Domain definition

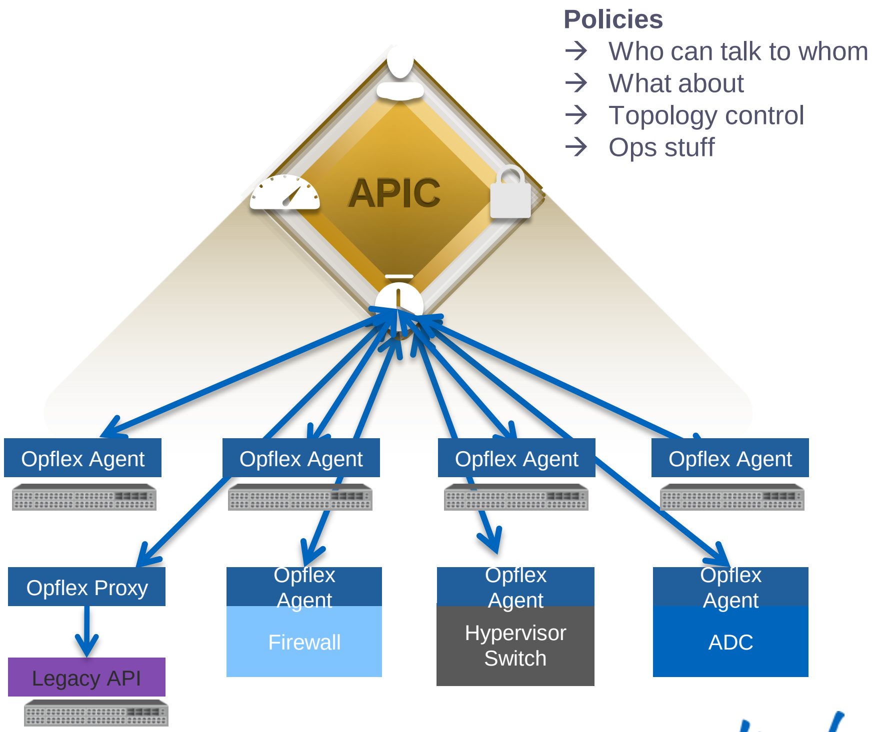


VMM Definition to avoid

OpFlex – A Flexible, Extensible Policy Protocol

OPFLEX is a new extensible policy resolution protocol designed for declarative management of any data centre infrastructure. Unlike legacy protocols such as OVSDB, OPFLEX was designed to offer:

- Declarative resolution – Push + Pull API support
- Abstract policies rather than device-specific configuration
- Flexible, extensible definition of using XML / JSON
- Support for any device – vswitch, physical switch, network services, servers, etc.

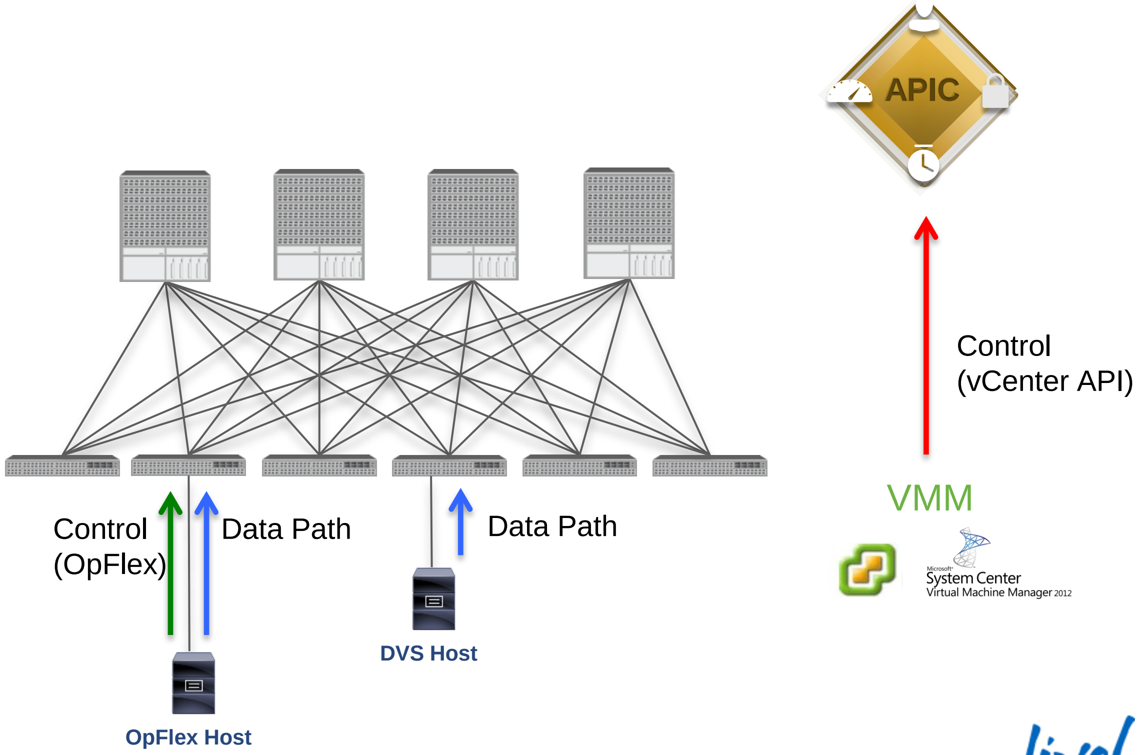


Cisco live!

Hypervisor Integration with ACI

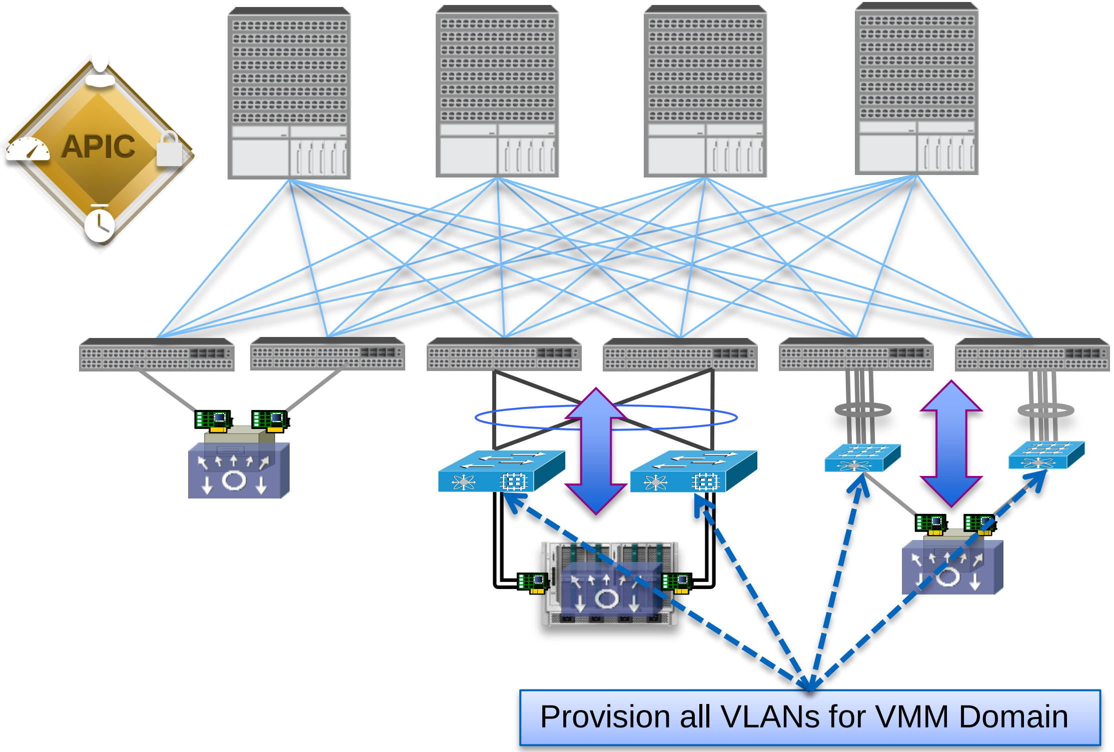
Endpoint Discovery

- Virtual Endpoints are discovered for reachability & policy purposes via 2 methods:
- Control Plane Learning:
  - Out-of-Band Handshake: vCenter APIs
  - Inband Handshake: OpFlex-enabled Host (N1KV, Windows Server 2012, etc.)
- Data Path Learning: Distributed switch learning
- LLDP used to resolve Virtual host ID to attached port on leaf node (non-OpFlex Hosts)



Design Considerations

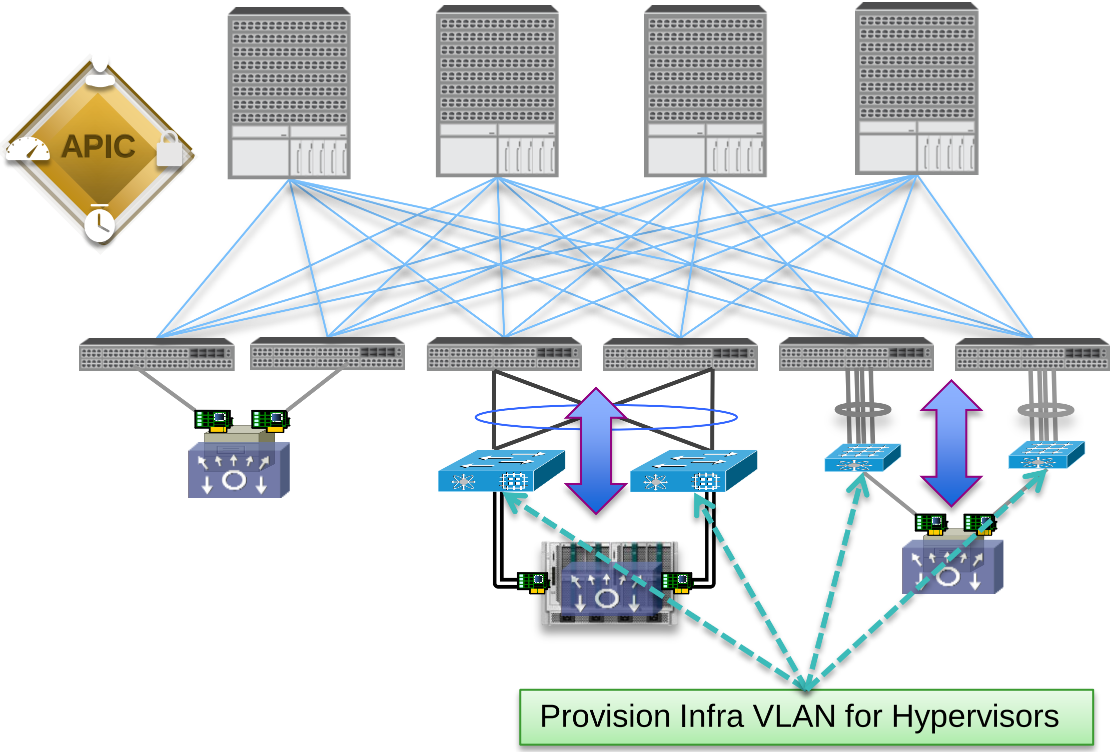
VLAN-Based Hypervisor Networks



- Hosts are assigned VLAN ID to EPG binding through VMM & APIC Integration
- Intermediate L2 nodes not managed – need to manage VLANs on these for each VMM Domain
- Endpoint location discovered through "stitching" LLDP TLVs (non OpFlex-enabled Hosts)

Design Considerations

VXLAN & NVGRE-based Hypervisor Networks

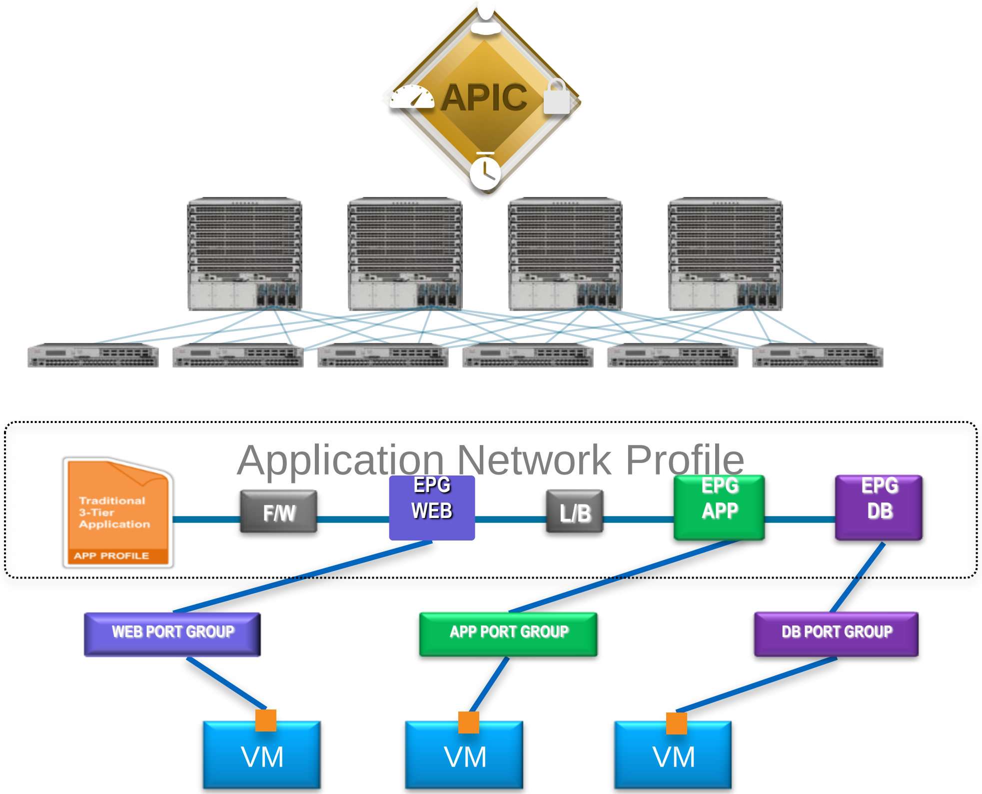


- Hosts are assigned VNID and VSID to EPG binding through VMM & APIC Integration
- Infra-VLAN is extended out to front-panel tenant ports - Infra-VLAN needs to be provisioned on intermediate L2 Nodes
- Endpoint location discovered though “stitching” LLDP TLVs (non OpFlex-enabled Hosts)



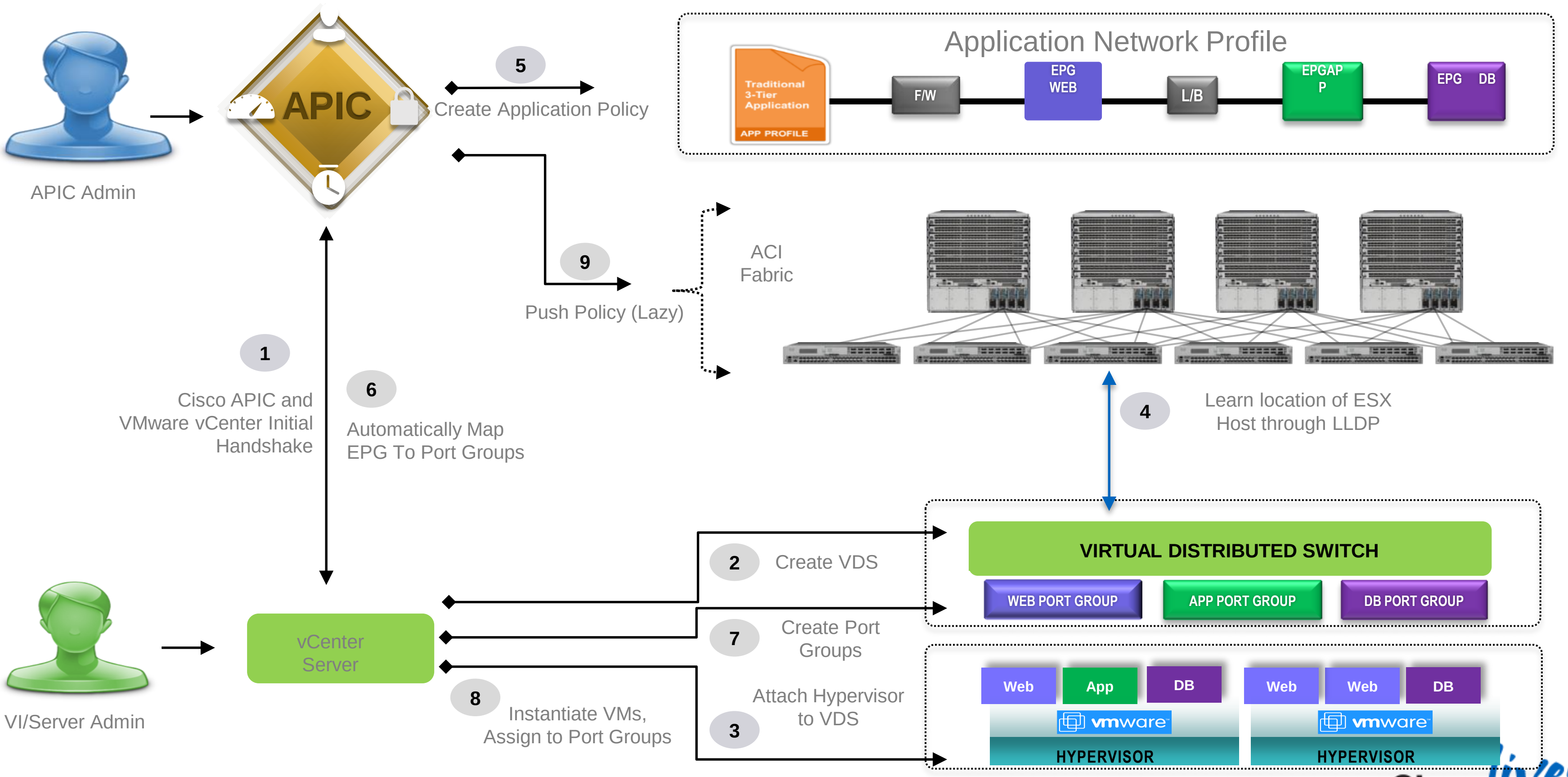

Integration with VMware DVS

ACI Fabric and VMware DVS Integration



- How does ACI Fabric implement policy?
  - Assigning EPs to EPGs
- What are EPs in virtual environment?
  - VM vNICs
- How does VMware apply network configuration?
  - Port Groups
- How are EPGs exposed to VMware?
  - Map EPGs to Port Groups

Cisco ACI Hypervisor Integration – VMware DVS

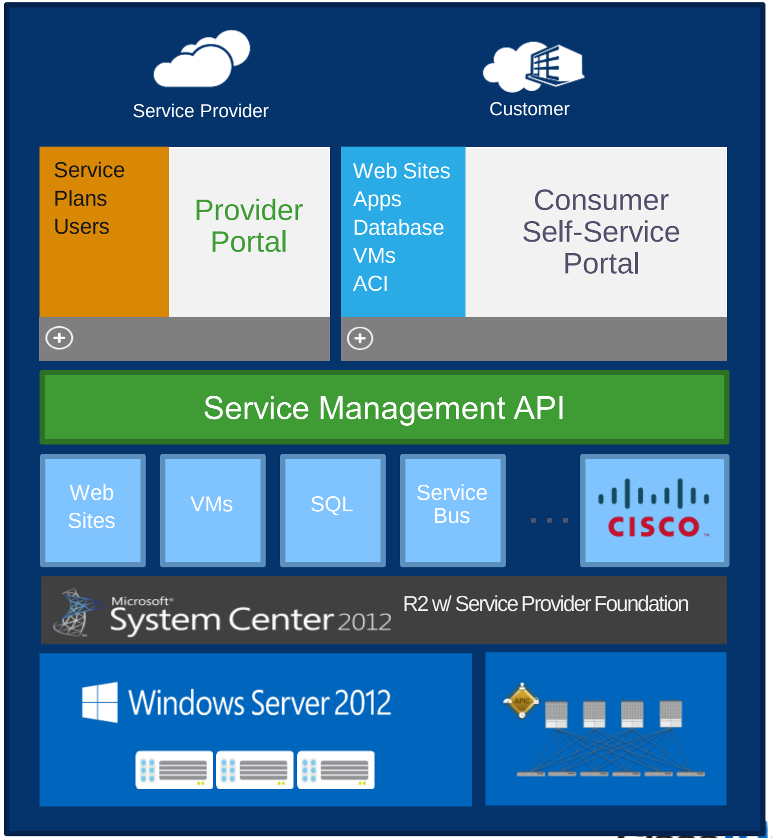





Integration with Microsoft

Microsoft Azure Pack Integration

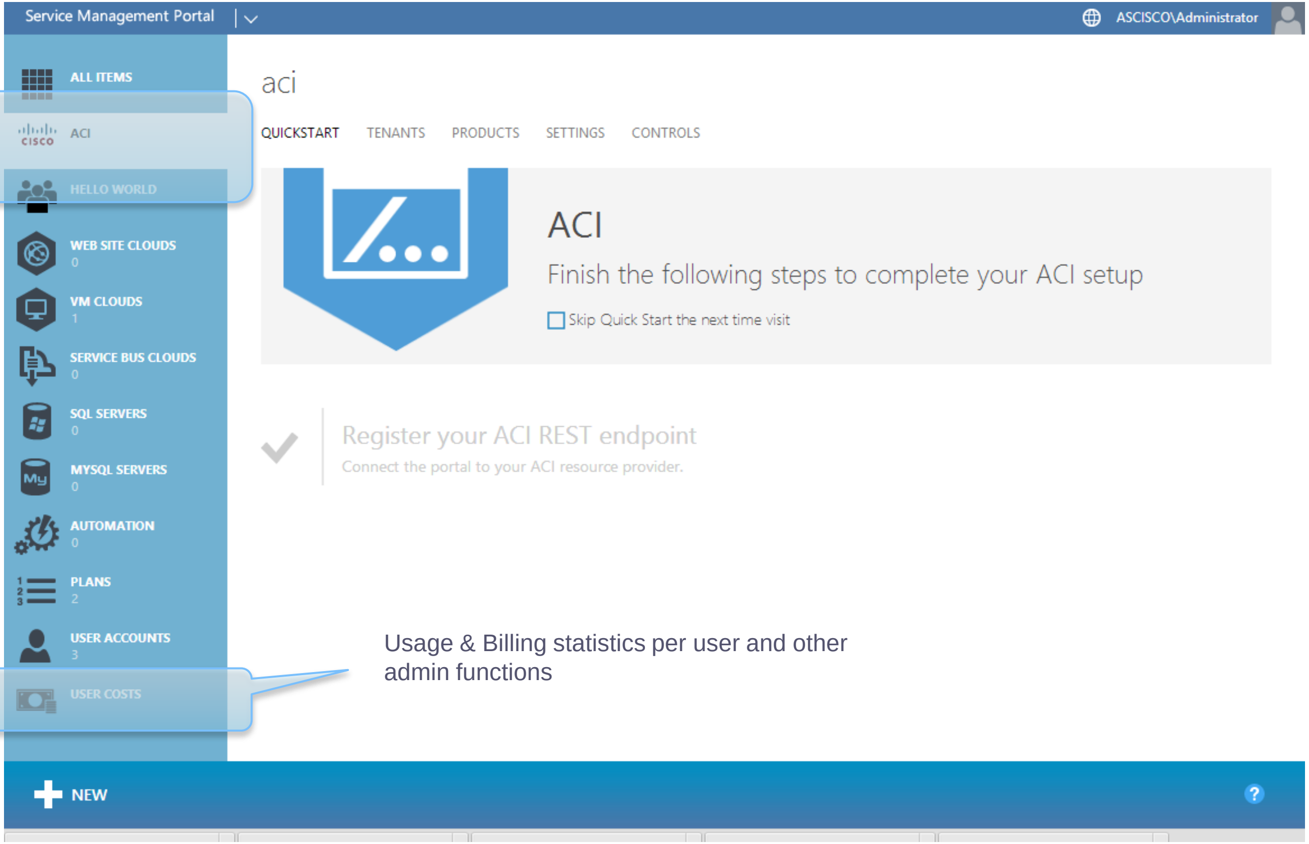
- Integration with Microsoft requires:
  - Windows Server 2012
  - Systems Centre 2012 R2 with SPF
  - Windows Azure Pack
- Azure Pack provides single pane of glass for Definition, creation, management of their cloud service
- Divided into Provider (Admin) portal and Consumer Self-Service (Tenant) portal
- Cisco ACI Service Plugin enables management of Network Infrastructure through APIC REST API



Microsoft Azure Pack Integration

Admin Experience

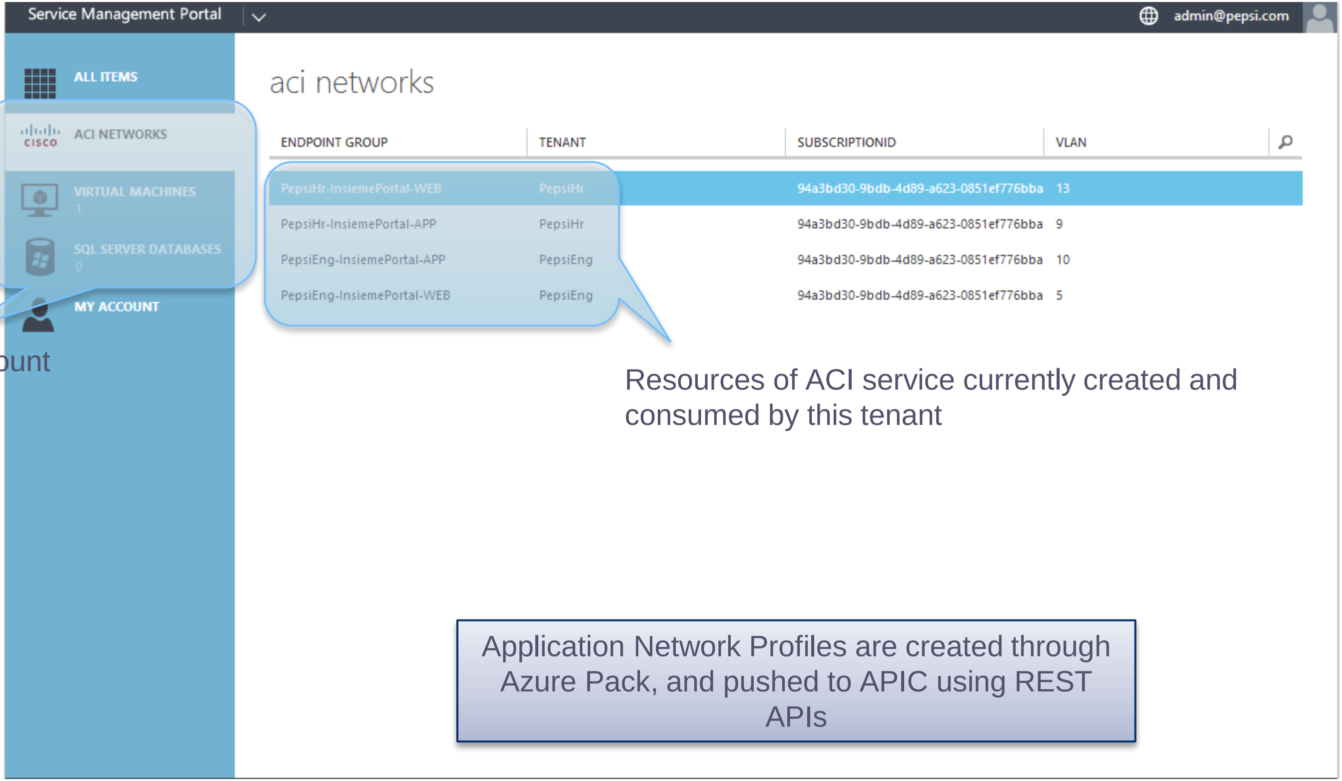
Add & Configure service providers for this deployment (APIC IP Address, Login Credentials, etc.)



Usage & Billing statistics per user and other admin functions

Microsoft Azure Pack Integration

Tenant Experience

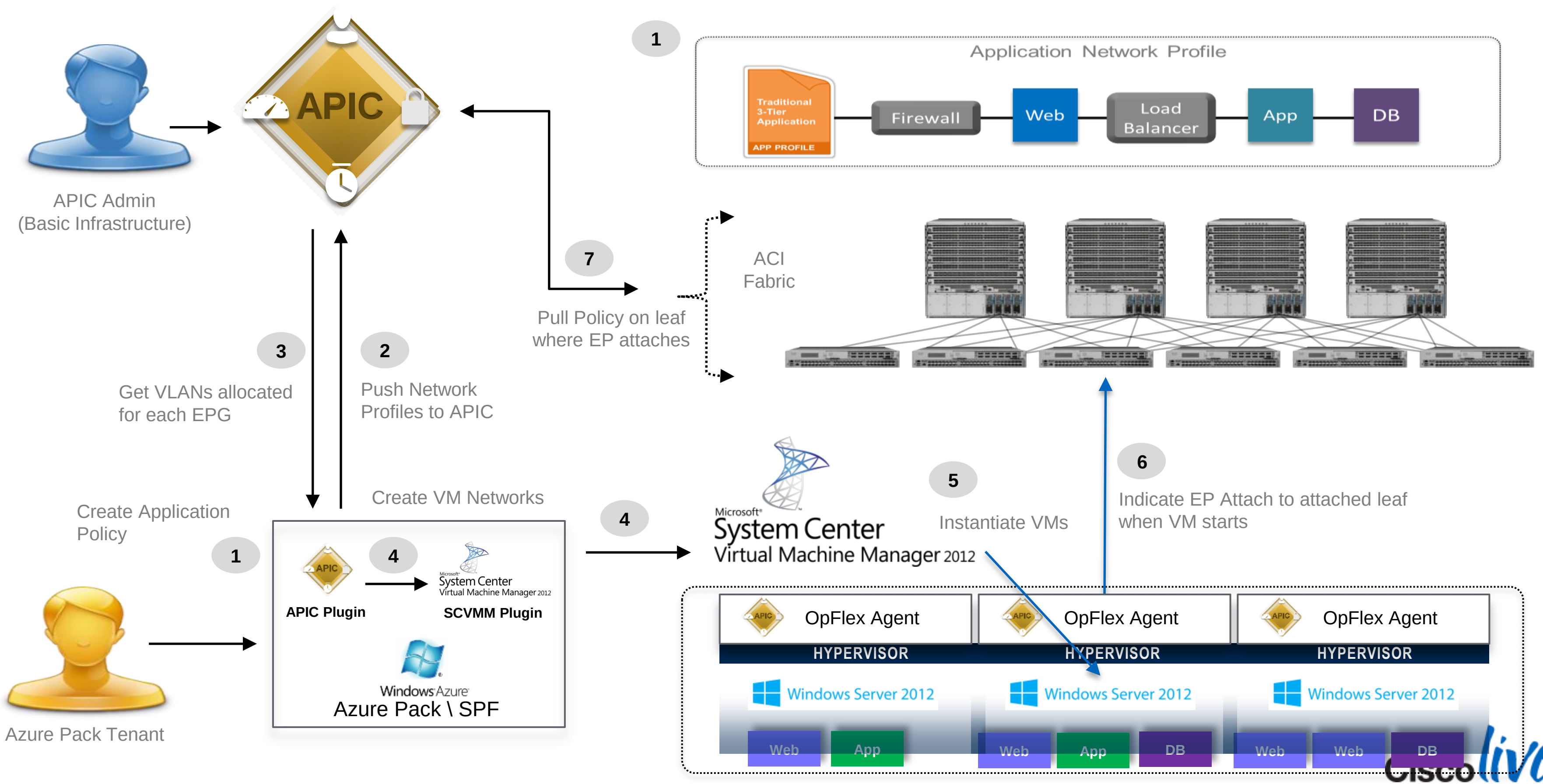


Services this account has access to

Resources of ACI service currently created and consumed by this tenant

Application Network Profiles are created through Azure Pack, and pushed to APIC using REST APIs

ACI Azure Pack Integration

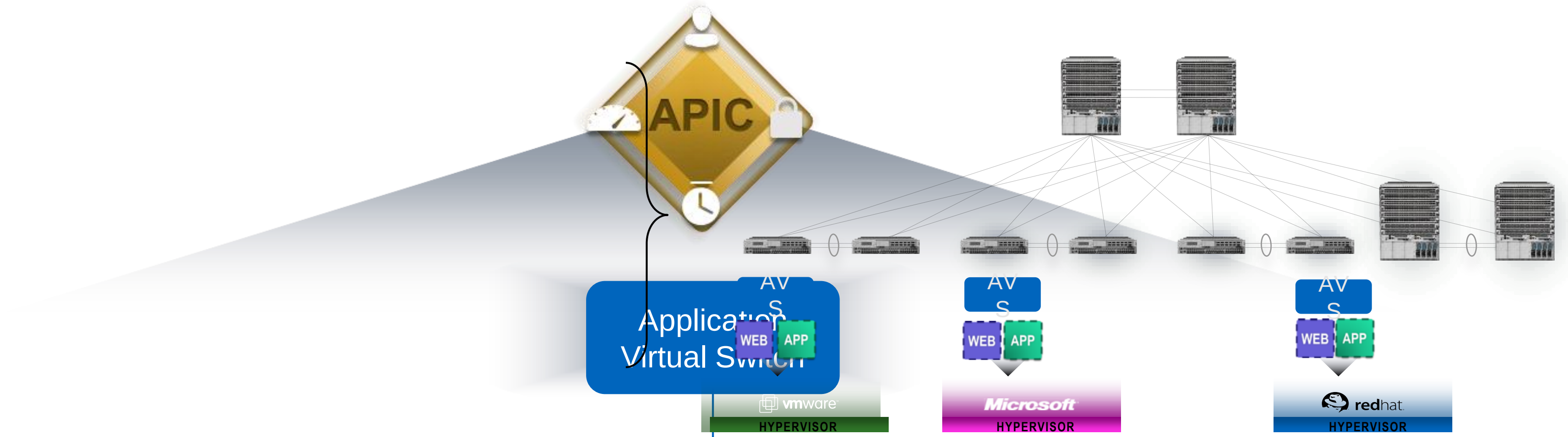





Nexus 1000V – ACI Edition

Cisco ACI - Application Virtual Switch (AVS)

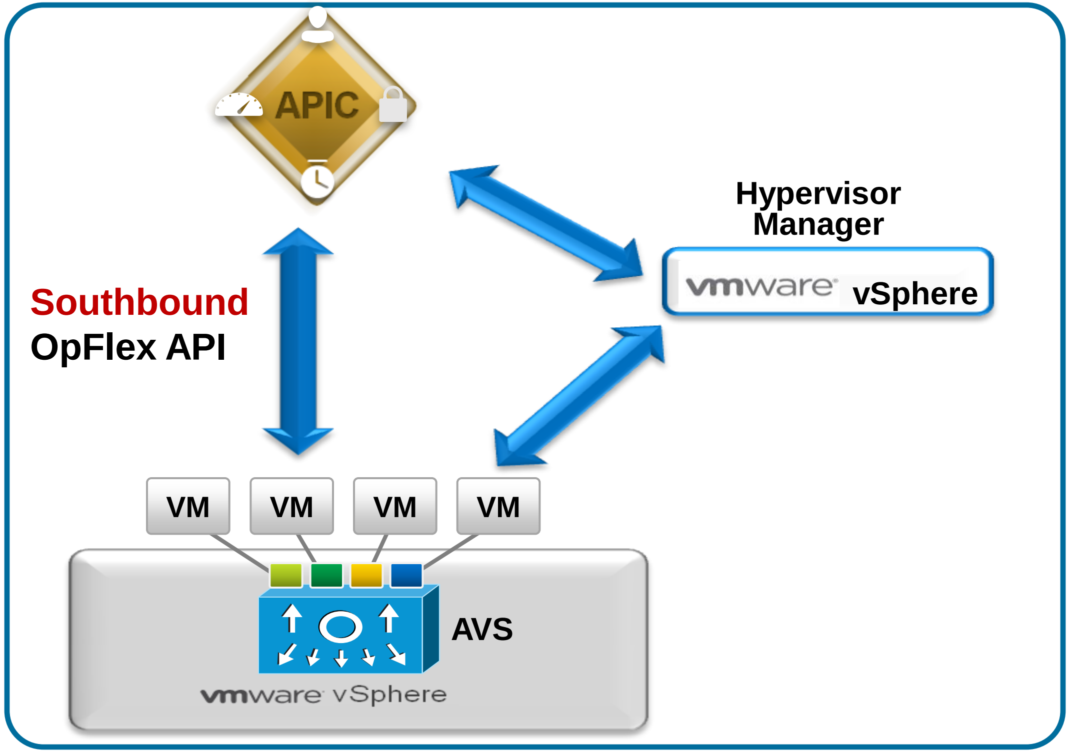
NETWORK VIRTUALISATION SUPPORT



PURPOSE BUILT VIRTUAL MEMBER OF ACI	OPTIMAL TRAFFIC STEERING	INTEGRATED VISIBLITY THROUGH APIC (PHYSICAL AND VIRTUAL)	COMMON MANAGEMENT MODEL THROUGH APIC	MULTI-HYPERVISOR SUPPORT	OPEN APIS

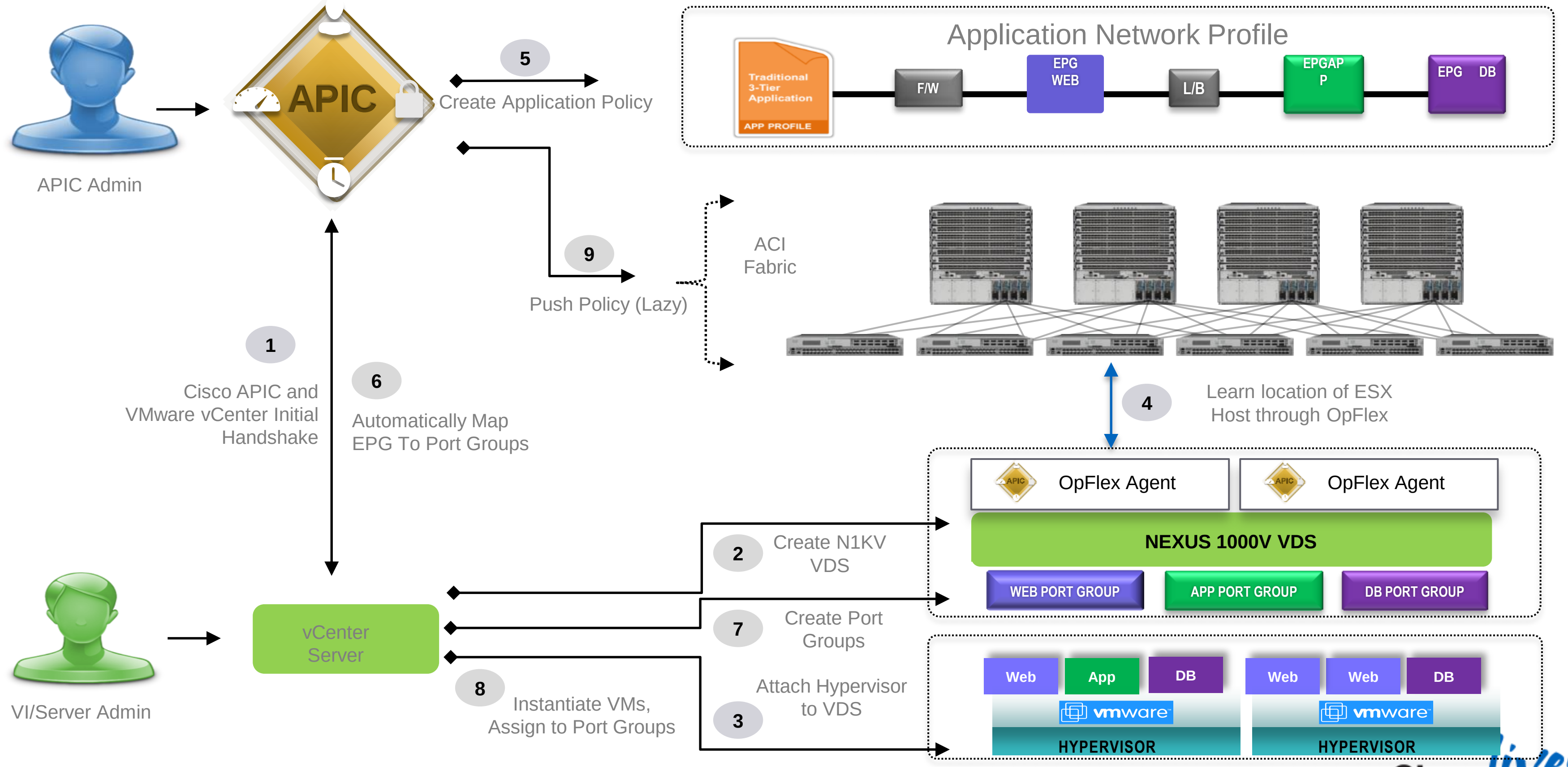
Nexus 1000V Integration Overview

- **OpFlex Control protocol**
  - Control channel
  - VM attach/detach, link state notifications
- VEM extension to the fabric
- vSphere 5.0 and above
- BPDU Filter/BPDU Guard
- SPAN/ERSPAN
- Port level stats collection



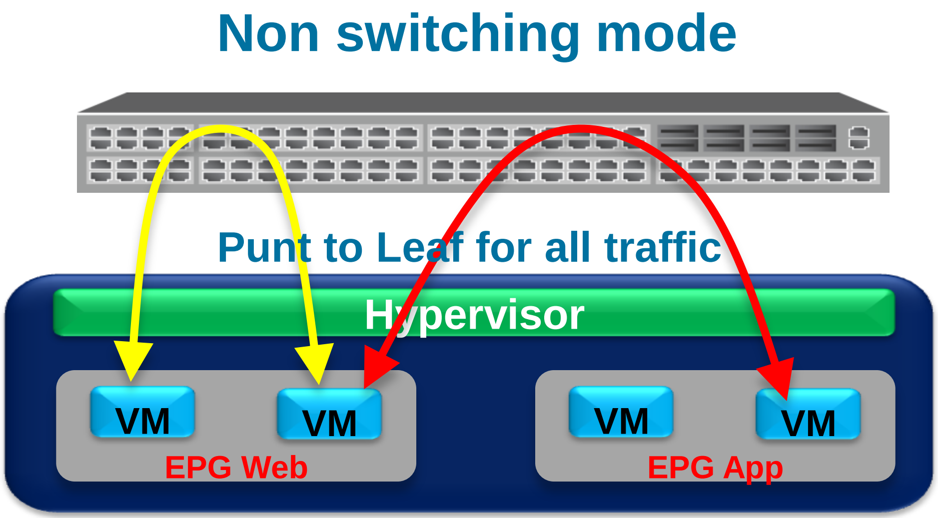
Cisco ACI Hypervisor Integration

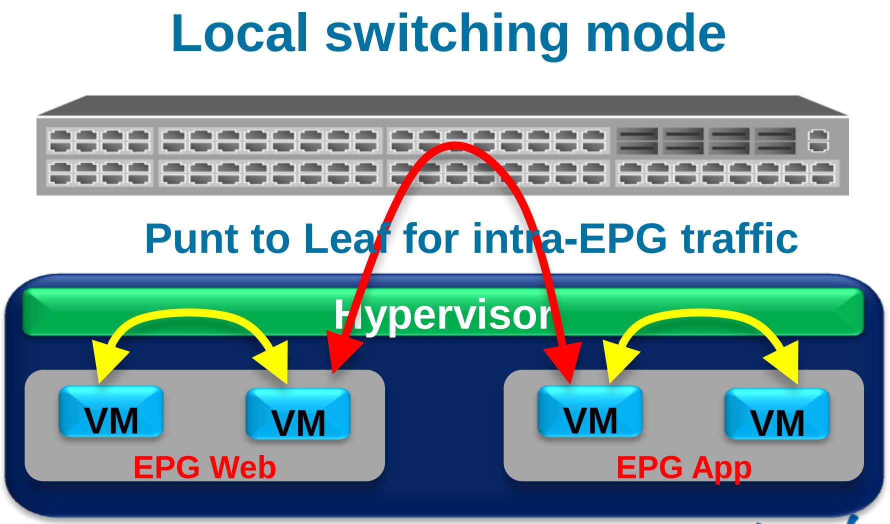
VMware N1KV VEM



Nexus 1000V Switching Modes

- NS – Non switching mode, FEX mode
- LS – Local switching within EPGs on the same host, similar behaviour as ESX VDS





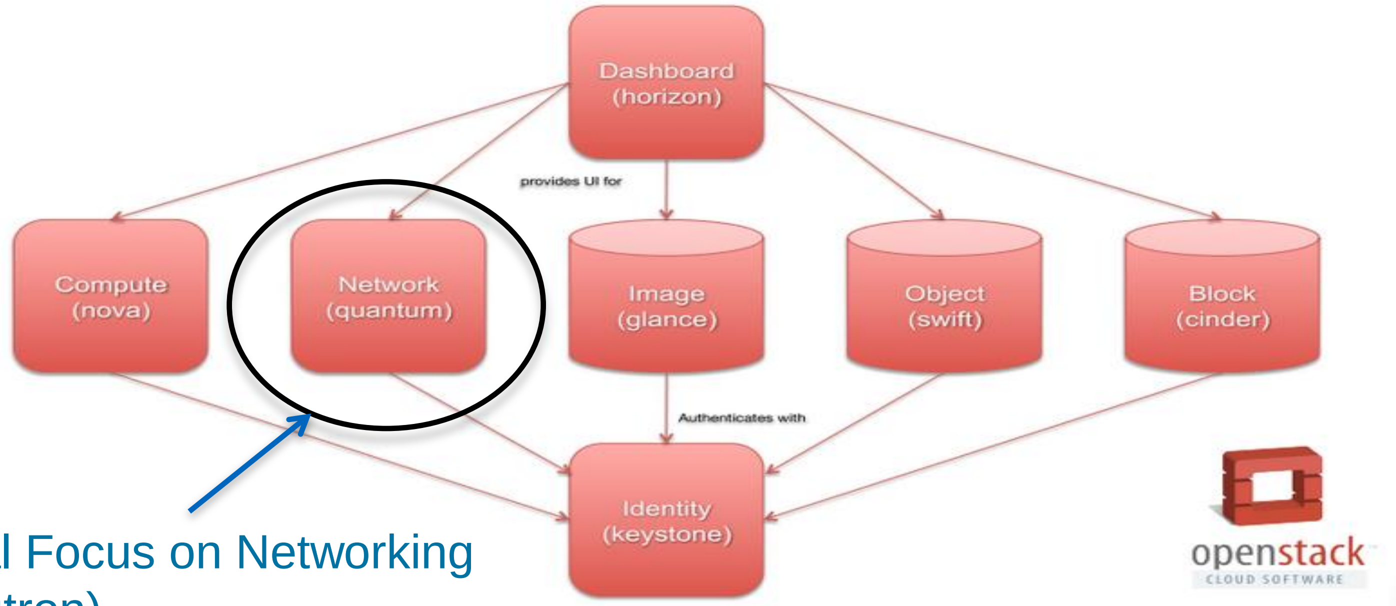
Cisco AVS Differentiation with ACI

Hypervisor Networking	VDS/OVS	AVS
No Switching	Yes	Yes
Local Switching	Yes	Yes
Full Switching(routing etc)	No	Yes
Optimal Traffic Steering	No	Yes
Local (on-host) Policy Enforcement	No	Yes
Single Point of Management with APIC	No	Yes, Robust
Atomic Counters	No	Yes
End-to-End Visibility	Yes	Enhanced
Consistency Across Hypervisors	No	Yes
Enhanced NX-OS	No	Yes
Ease of Install/Upgrade	Separate	Integrated

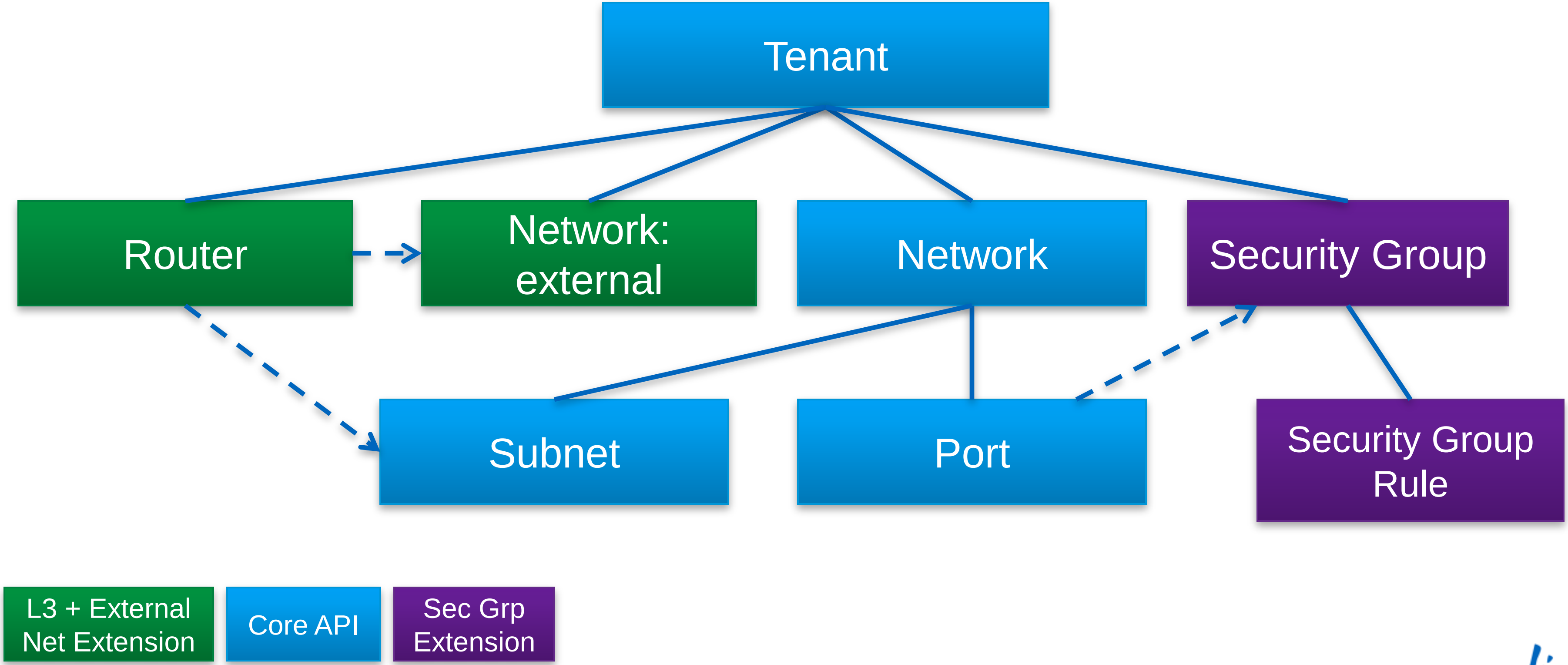
Integration with OpenStack

OpenStack Components

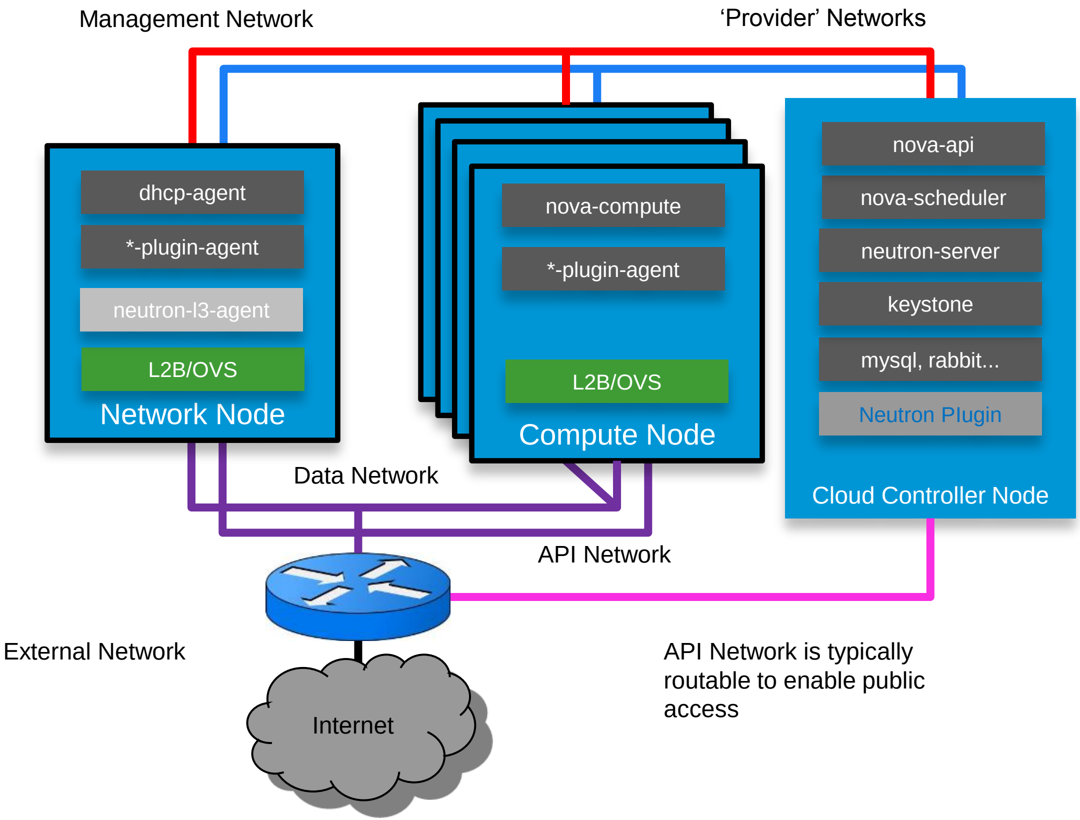


Initial Focus on Networking (Neutron)

OpenStack Neutron Networking Model

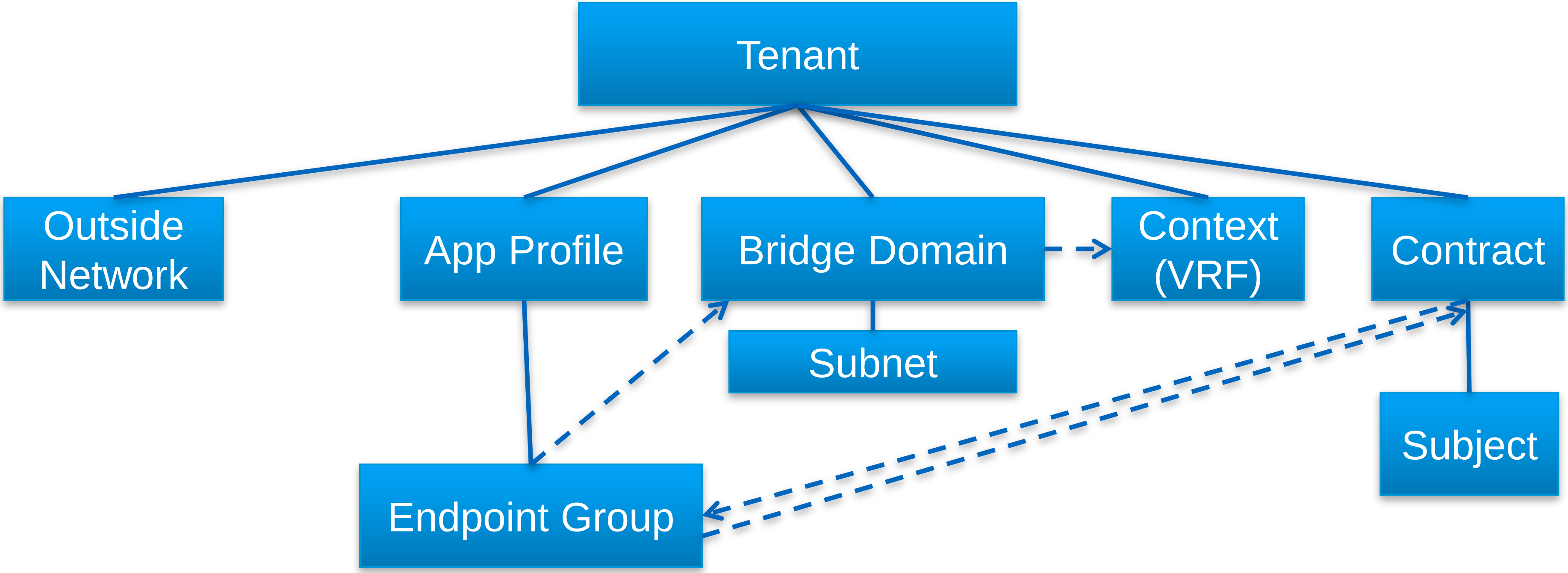


OpenStack Deployment



API Network is typically routable to enable public access

Cisco ACI Model



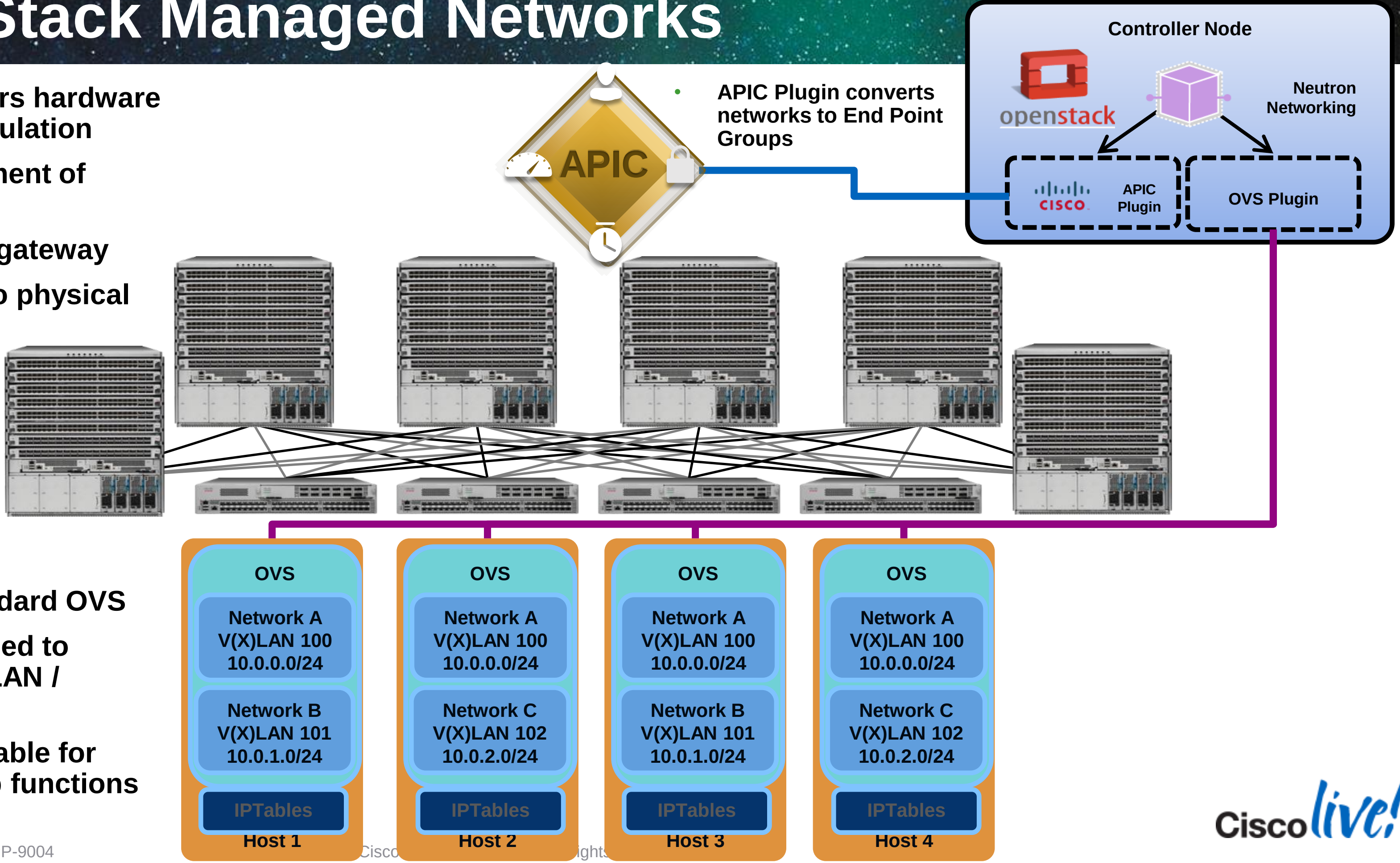
Cisco ACI Model

Neutron API Mapping

OpenStack	ACI
Tenant	Tenant
No equivalent	Application Profile
Network	EPG
Subnet	Subnet
Security Group	Handled by Host
Security Group Rule	Handled by Host
Router	Context
Network:external	Outside

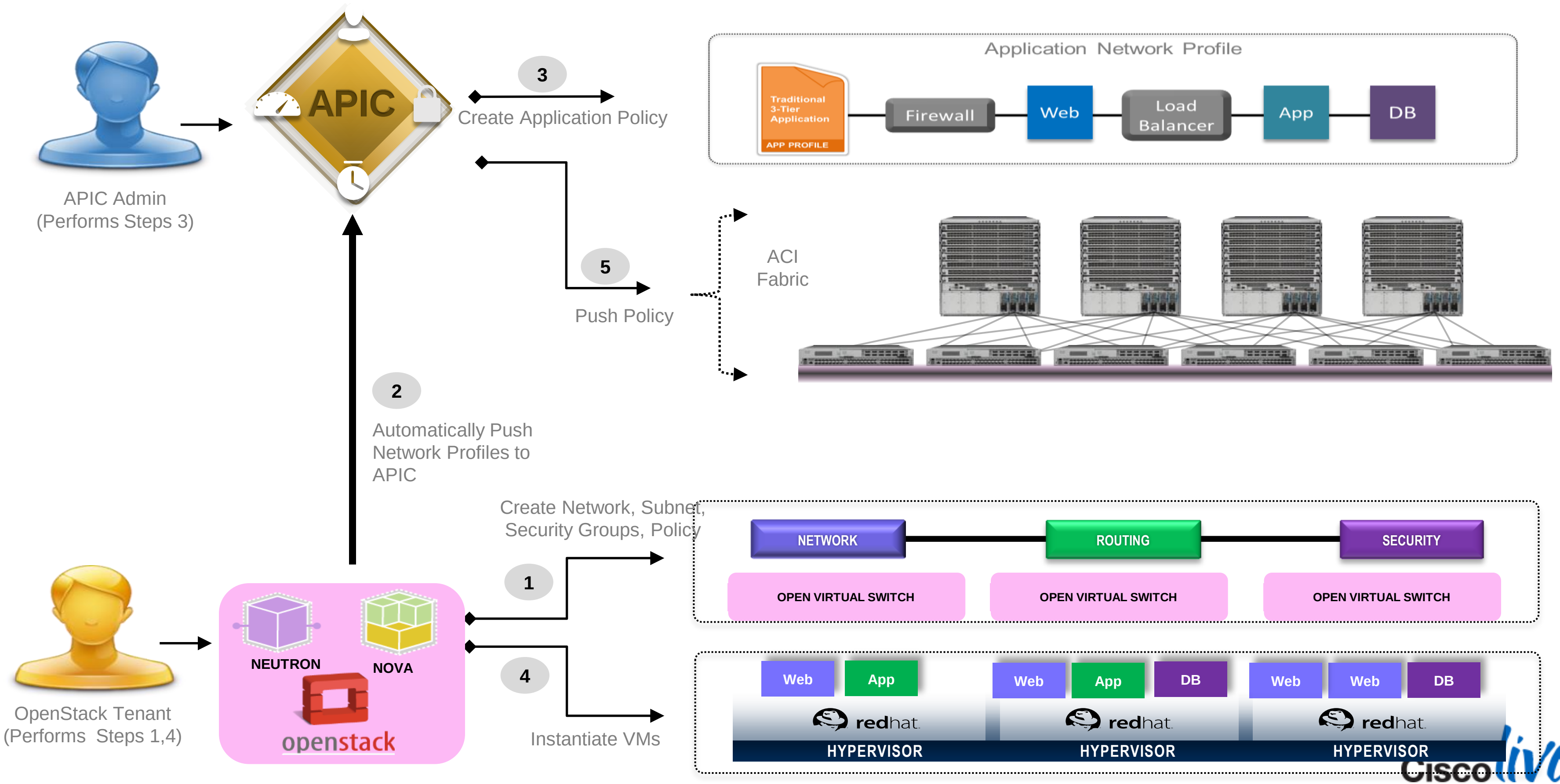
OpenStack Managed Networks

- ACI Fabric offers hardware VXLAN encapsulation
- Flexible placement of endpoints
- Distributed L3 gateway
- L2 Extension to physical servers, etc.

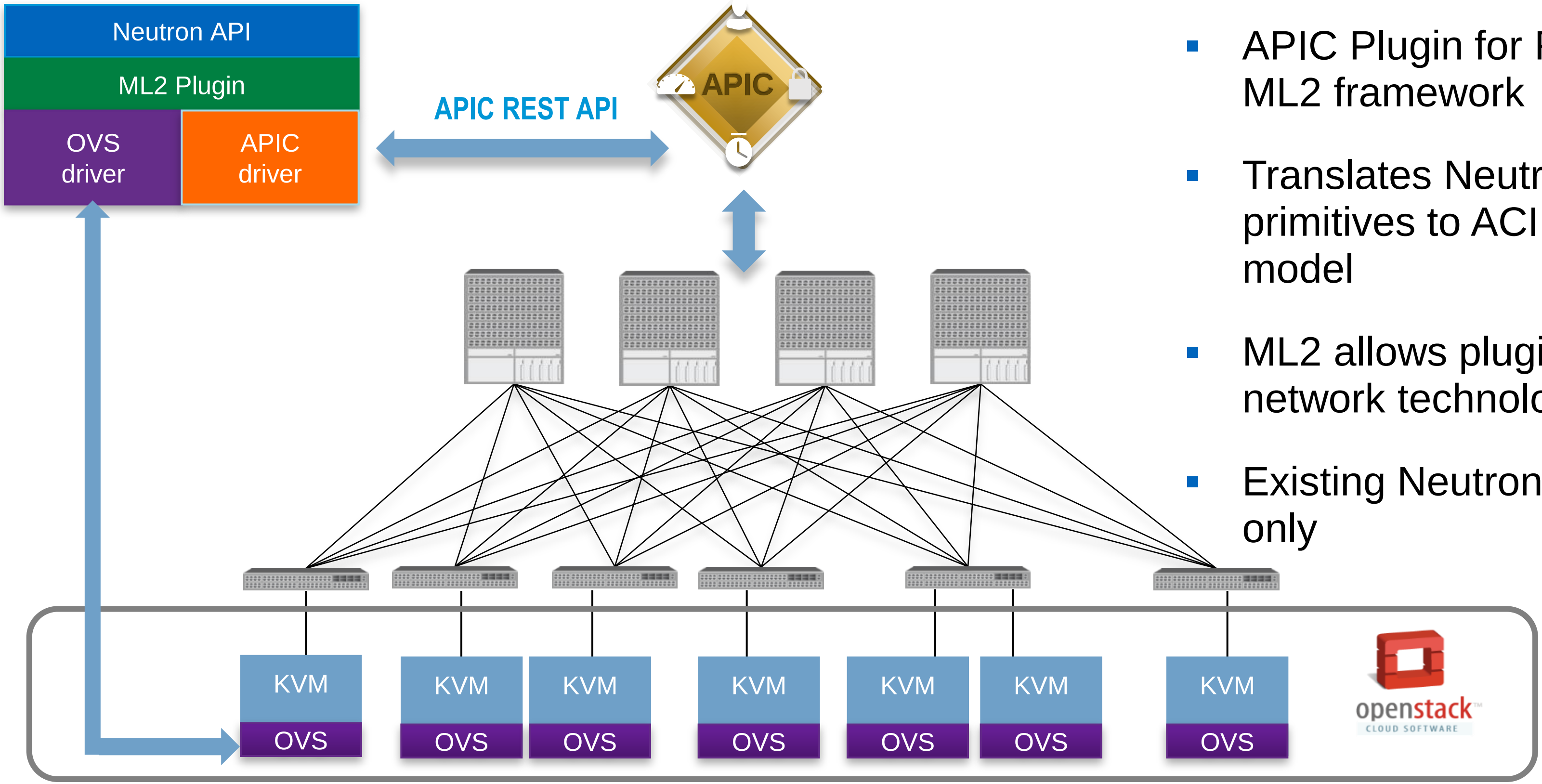


- Leverage standard OVS
- Network mapped to segment ID VLAN / VXLAN
- IPTables available for security group functions

ACI OpenStack Integration - FCS



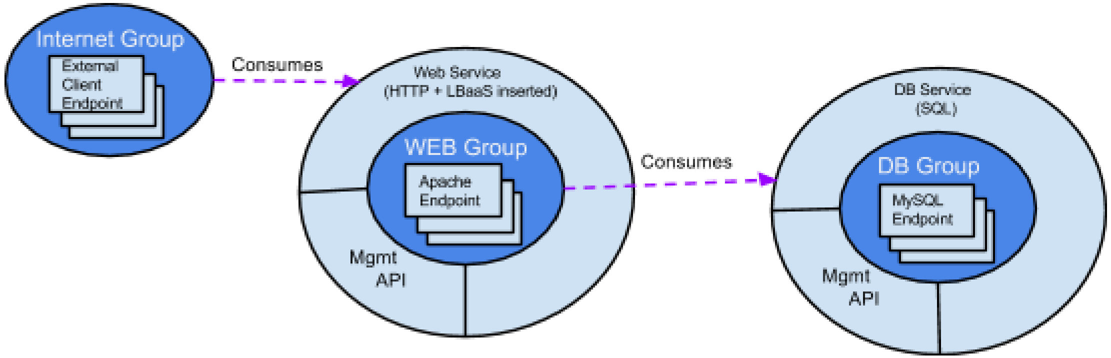
ACI Neutron Plugin



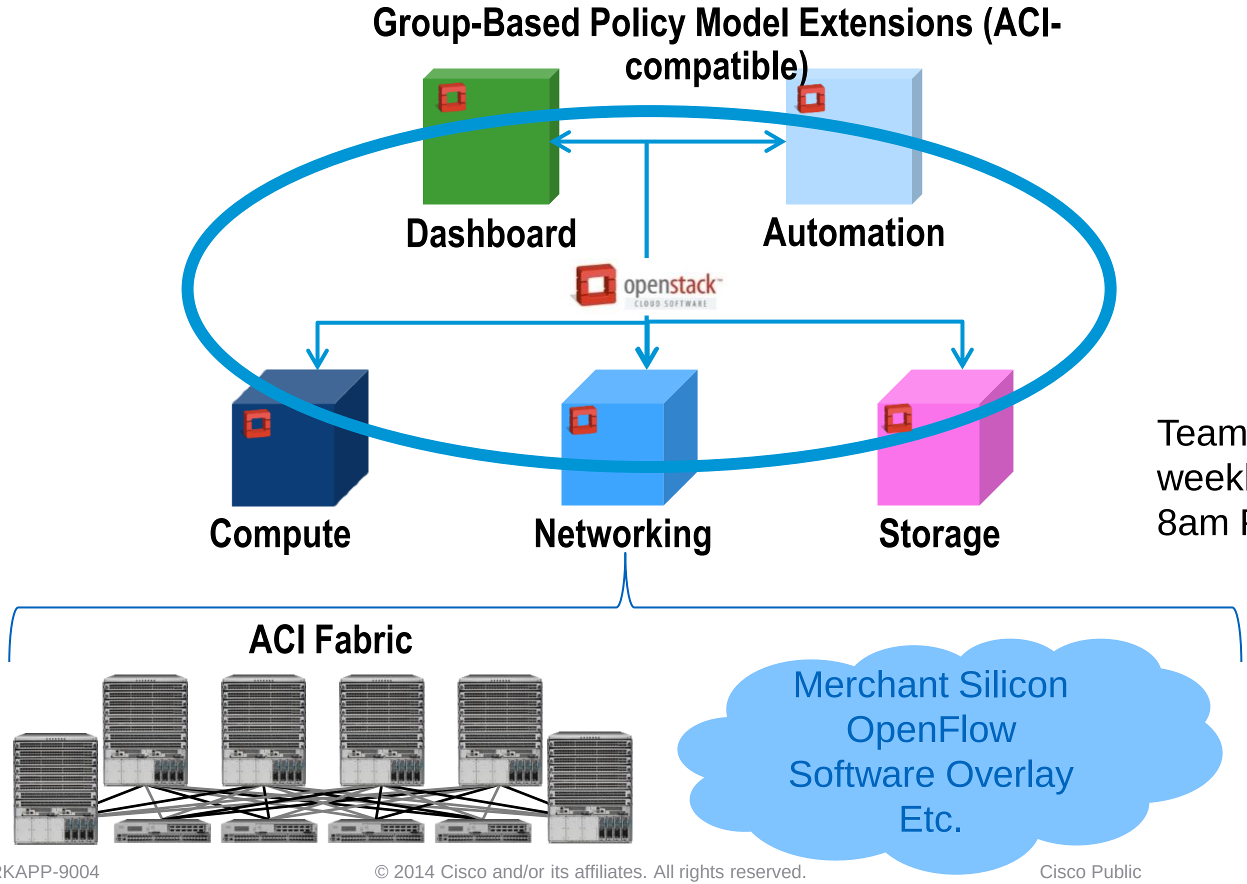
- APIC Plugin for Fabric using ML2 framework
- Translates Neutron primitives to ACI policy model
- ML2 allows plugin to select network technology
- Existing Neutron functions only

Group-based Policy in OpenStack

- Messy mapping ACI to current OpenStack components
  - Endpoint Groups (Ports + Security Groups)
  - Contracts (Security Groups + Security Group Rules)
- Goal : Introduce ACI model into OpenStack
- Starting with Groups and Group based Policies

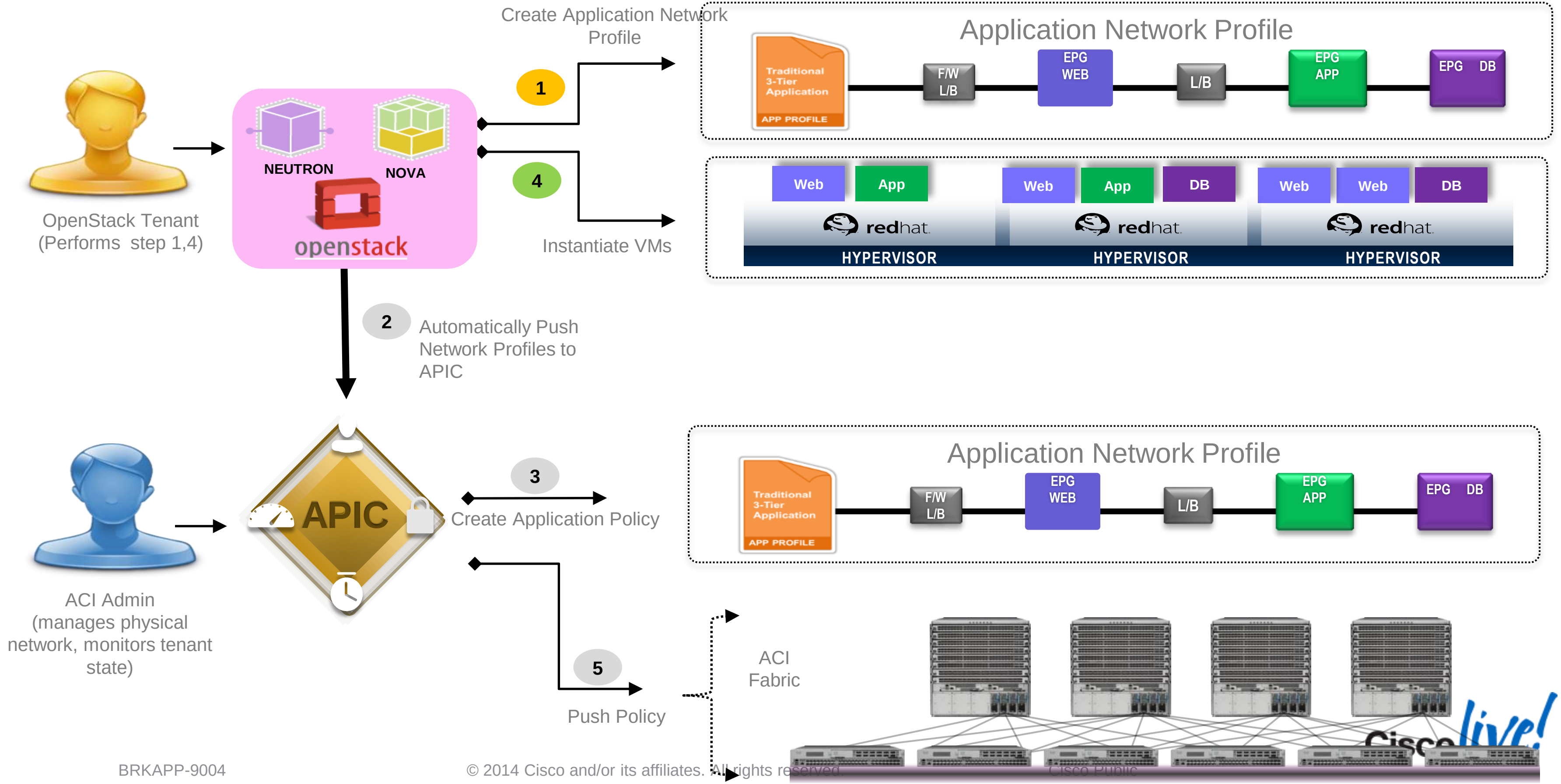


Group-based Policy in OpenStack



Team meets weekly at 8am PT

ACI OpenStack Integration – Post-FCS

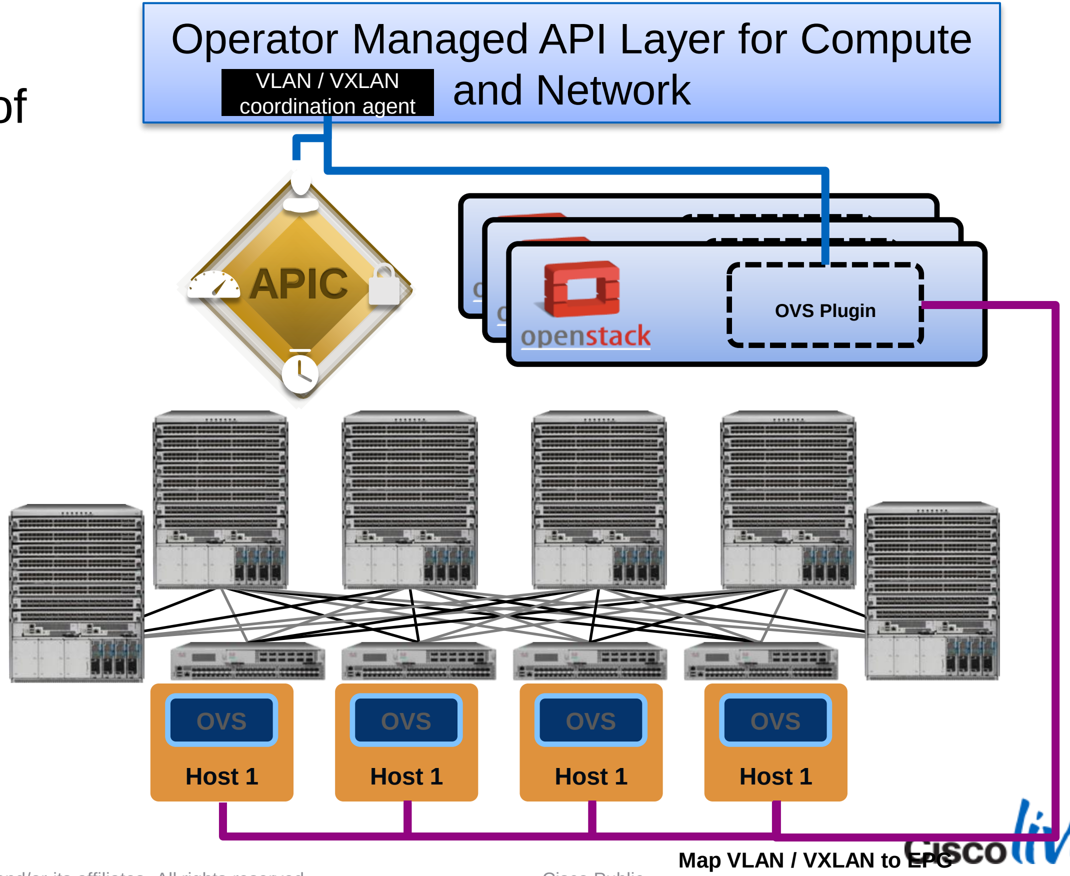


APIC Managed Networking

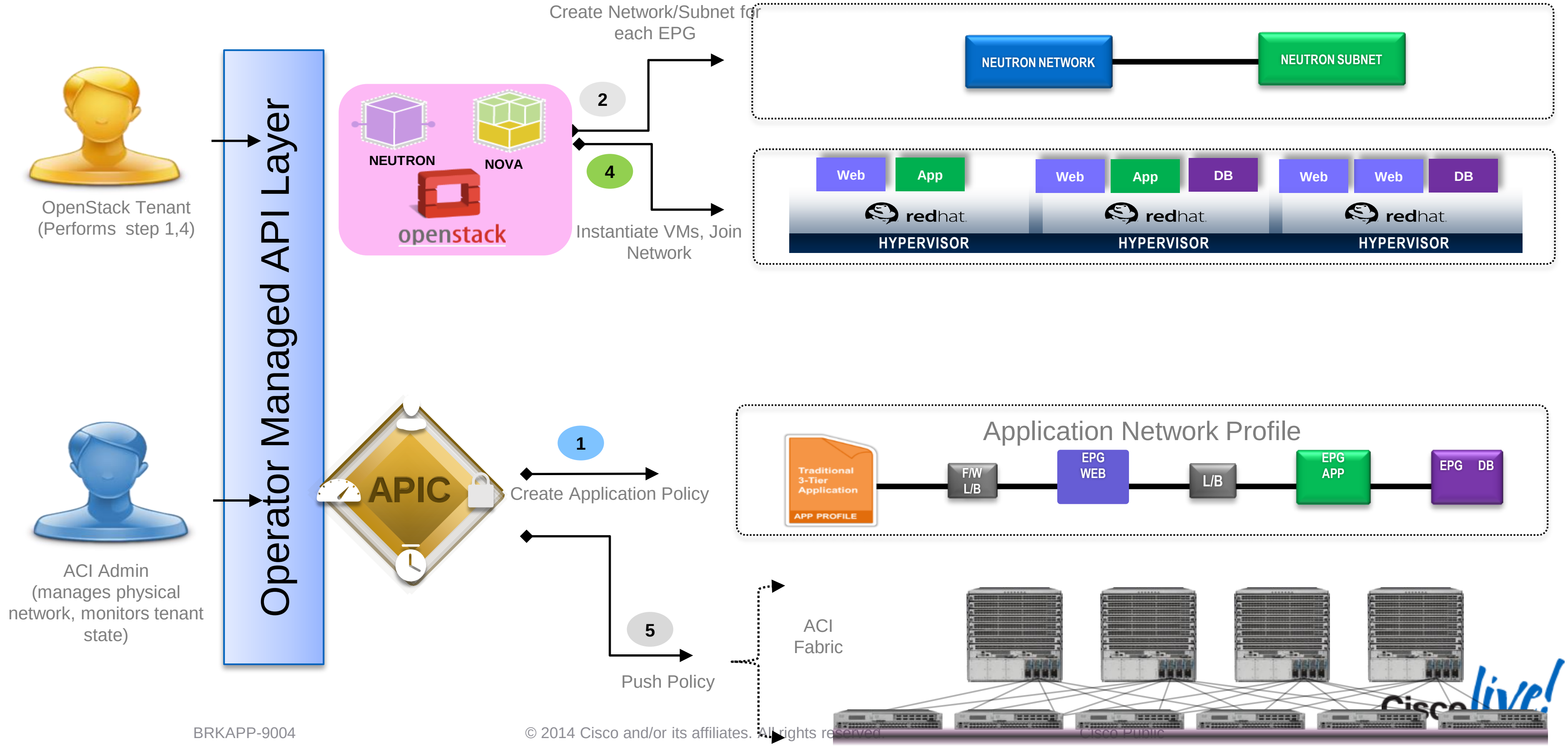
ACI Fabric also supports OpenStack through the addition of an Operator Managed API Layer

Highlights:

- No OpenStack changes required! - OpenStack running OVS Plugin or even nova network
- Network Policy defined by APIC in terms of EPGs, Contracts, etc.
- Requires Operator Managed API Layer



APIC Managed Networking






Q & A

Complete Your Online Session Evaluation

Give us your feedback and receive a Cisco Live 2014 Polo Shirt!

Complete your Overall Event Survey and 5 Session Evaluations.

- Directly from your mobile device on the Cisco Live Mobile App
- By visiting the Cisco Live Mobile Site www.ciscoliveaustralia.com/mobile
- Visit any Cisco Live Internet Station located throughout the venue

Polo Shirts can be collected in the World of Solutions on Friday 21 March 12:00pm - 2:00pm

Learn online with Cisco Live!

Visit us online after the conference for full access to session videos and presentations. www.CiscoLiveAPAC.com

Cisco live!

CISCO ™