HPE aruba networking

atmosphere'23 BELGIUM

CX in the data center

Dik van Oeveren, Consulting Systems Engineer EMEA

October 19, 2023

AGENDA

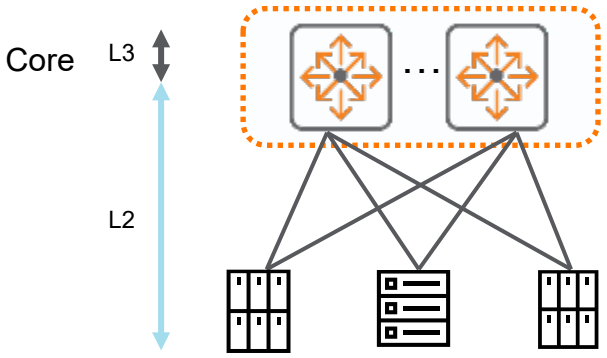
- Data Center reference architectures
- Data Center security challenges and solutions
- Flow behavior and security features
- Data Center orchestration

Data Center network reference architectures

Does Every DC Network Architecture require Spine/Leaf with VXLAN?

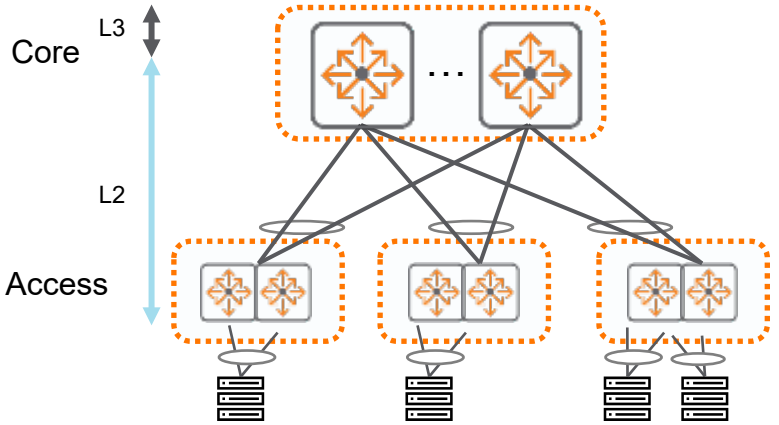
- These are still valid based on customer requirements, they all support HA and network automation

1-Tier Data Center



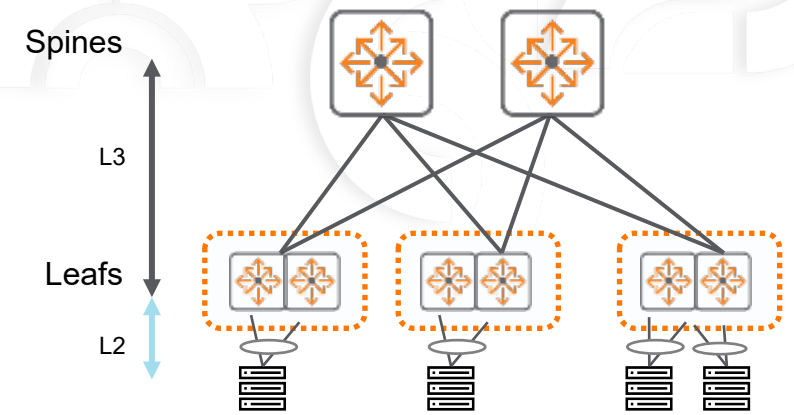
- Supports L2 (e.g. vMotion) /L3 connectivity between racks/servers
- Modular/Fixed port core switches are possible, this will determine how many servers can be connected
- Link aggregation from core to servers provides traffic load sharing and link/switch redundancy

2-Tier Data Center (Star Topology)



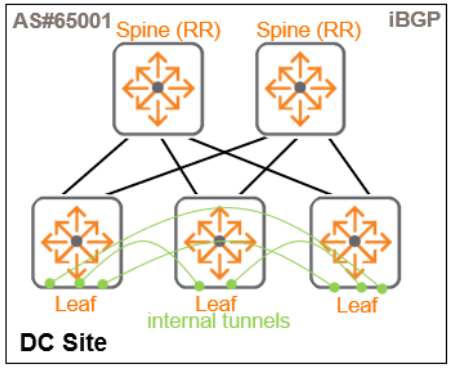
- Supports L2 (e.g. vMotion) /L3 connectivity between racks/servers
- Modular/Fixed port core switches are possible, this will determine how many access switches can be used
- Loop free topology as link aggregation is used between Access/Core for traffic load sharing and link/switch redundancy
- STP enabled as a backup mechanism to prevent loops
- Link aggregation from access to servers provides traffic load sharing and link/switch redundancy

L3 Spine/Leaf Data Center

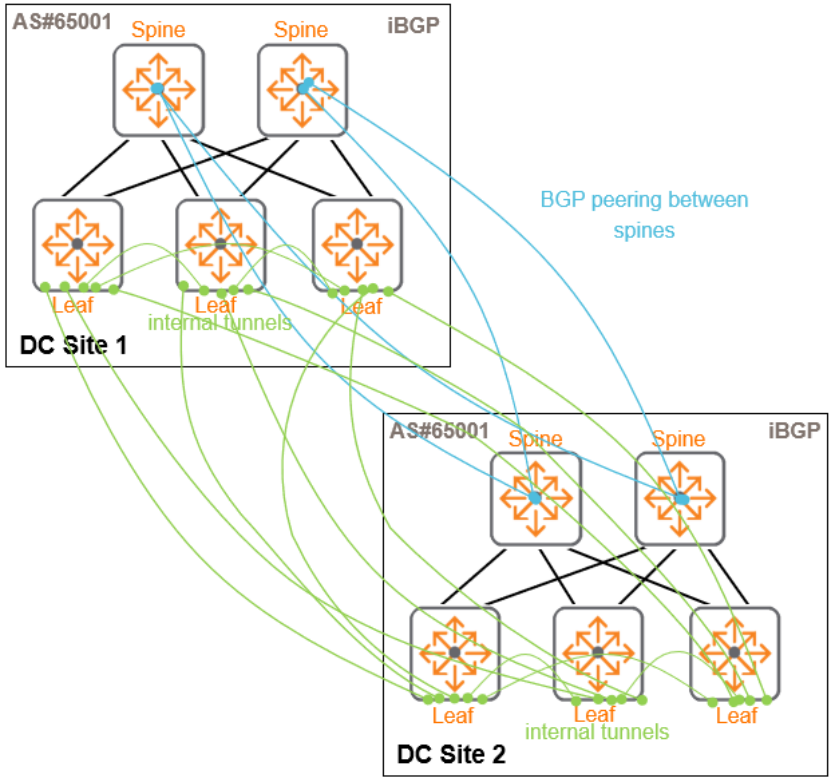


- Supports L3 connectivity between racks/leafs
- Removes STP since an L3 IP fabric is used
- Failure domain contained at L2 leafs
- Modular/Fixed port spines are possible, this will determine how large the fabric can grow
- Link aggregation from leafs to servers provides traffic load sharing and link/switch redundancy

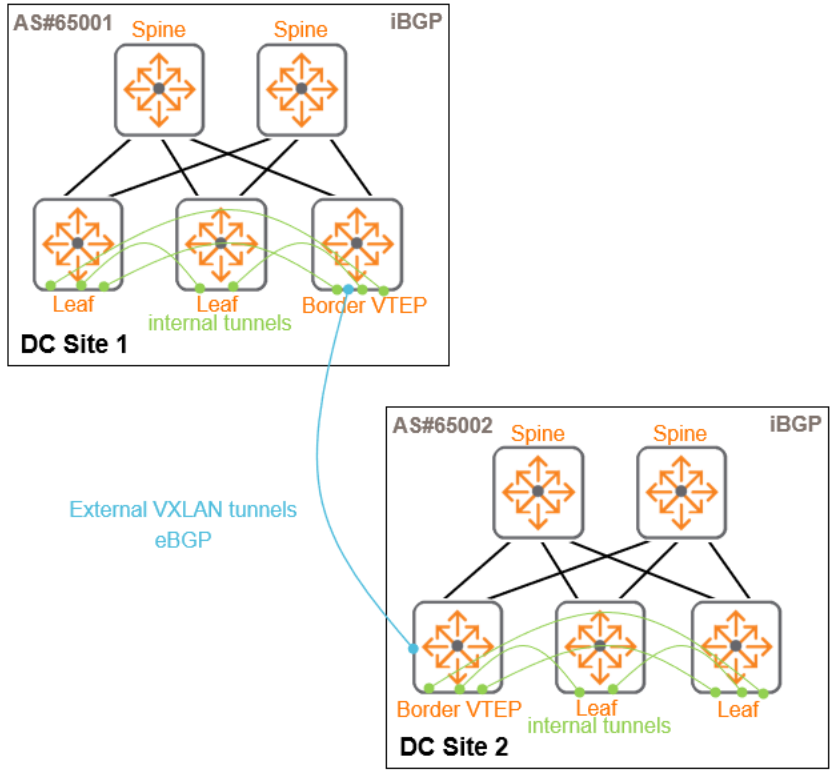
Reference architectures that scale



Single site, single fabric

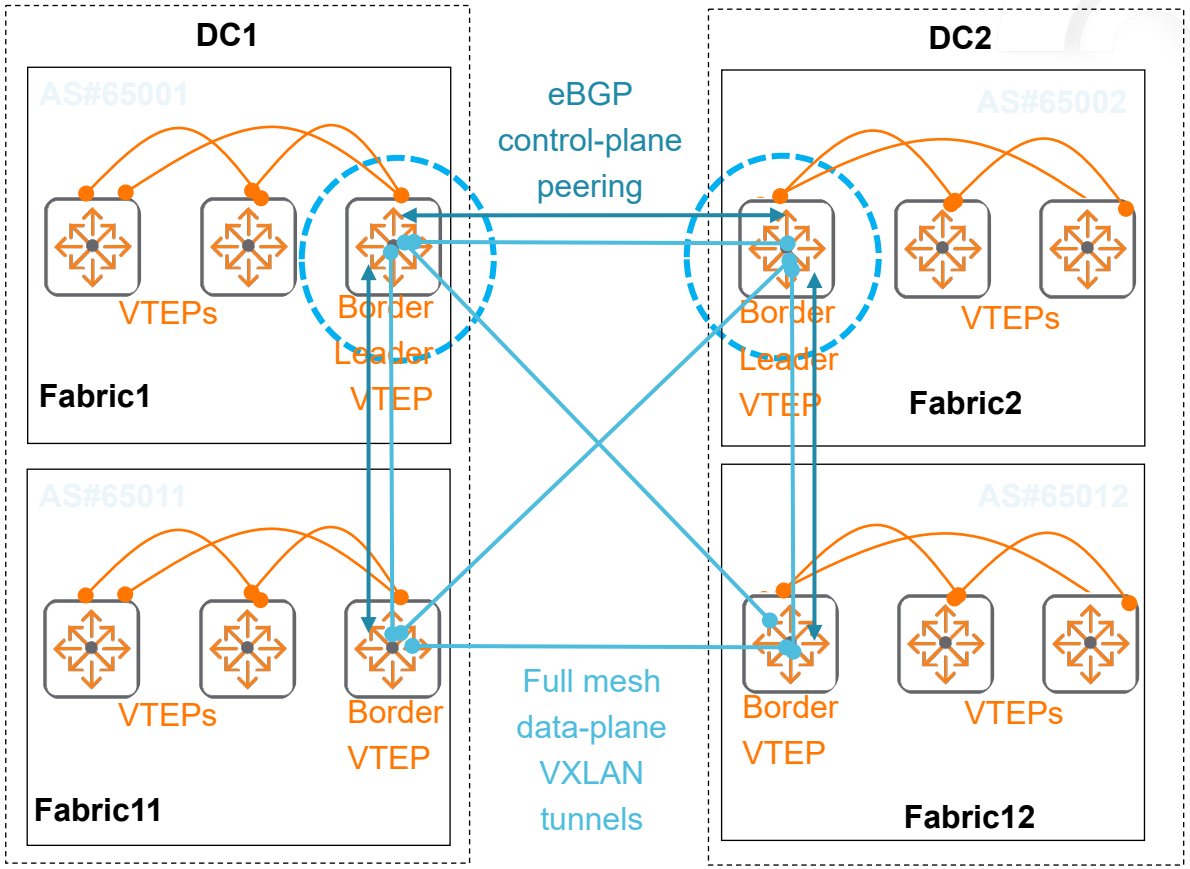


Single fabric, multi site



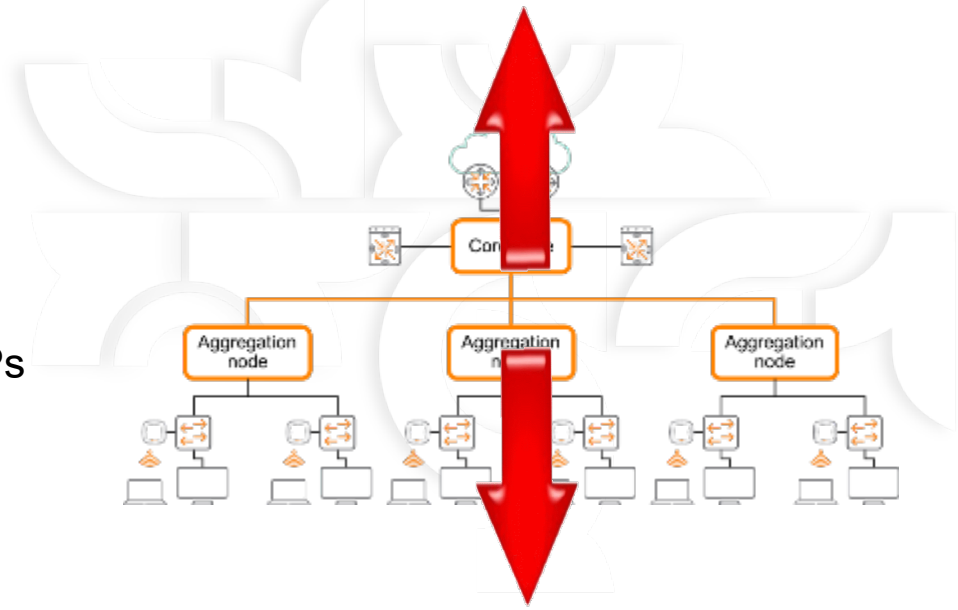
Multi site, single fabric per site

Reference architecture: Multiple Fabric, multiple sites, multi fabric per site



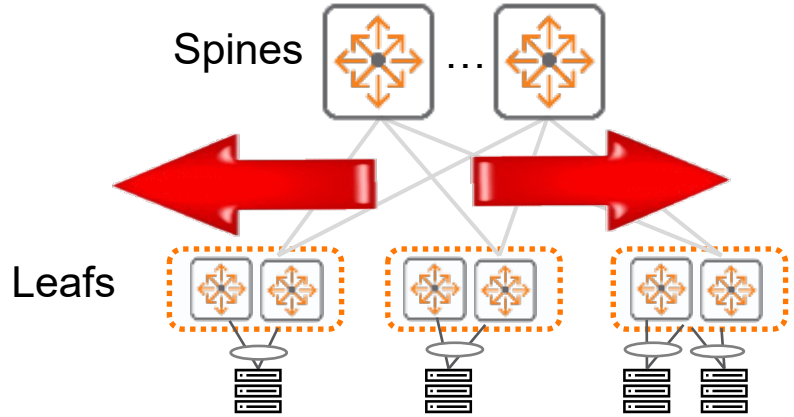
Data Center security today

What makes a Data Center Network?



- **Local Area Network (LAN) / Campus Networks**
  - Same geographical location, building, campus etc.
  - Wired and wireless network connects users, IP phones and wireless APs
  - Typical features required: POE, 802.1X etc

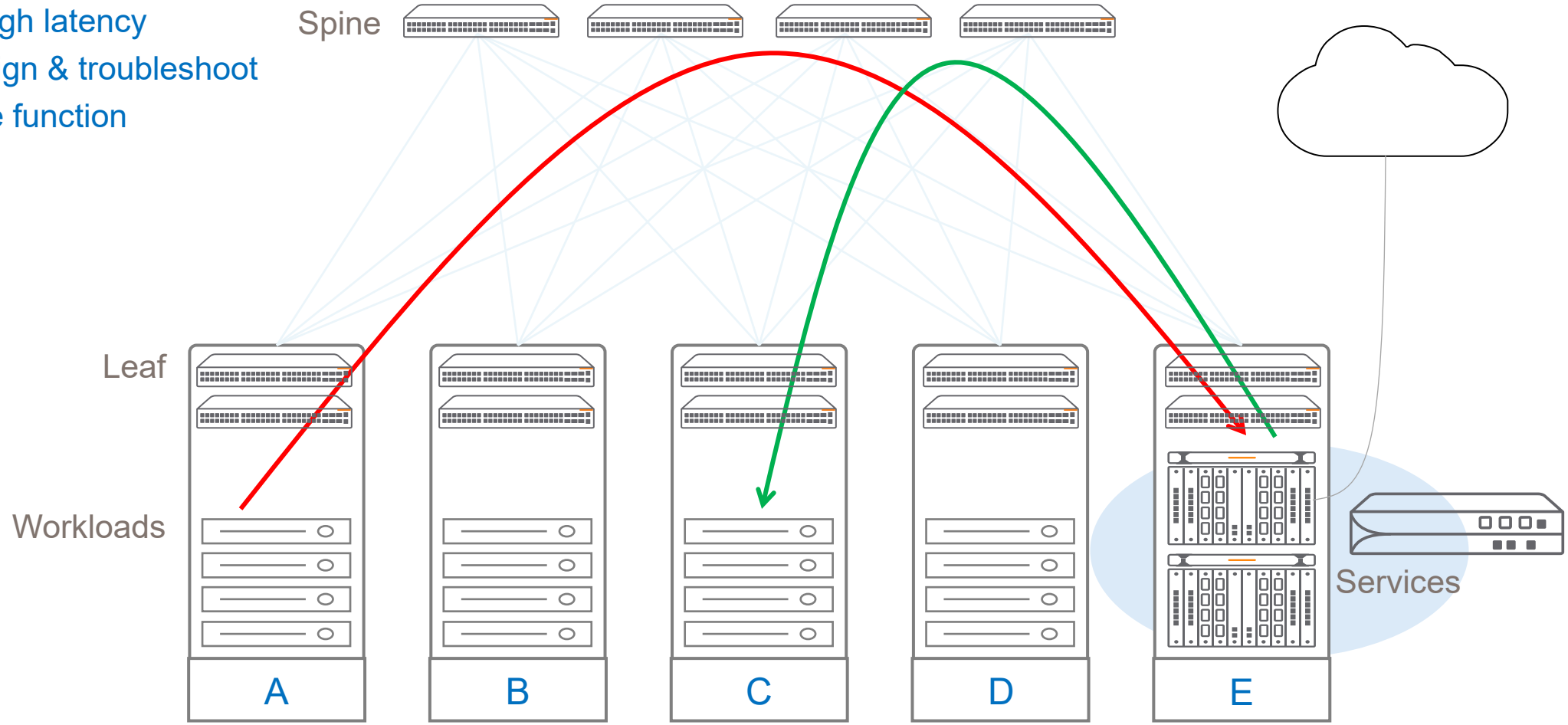
- **Data Center Networks**
  - Same geographical location (single data center)
  - Connects Servers/VMs/Containers, applications, storage, firewalls/ load balancers, etc. – wired connectivity
  - Stable, low latency fabrics with high availability / high performance and throughput / density and scale
  - Build revenue for business (E-Commerce)!
  - Typical features required: VXLAN/EVPN, BGP, OSPF, DCB, etc..
  - Focus on improving East - West traffic between racks



Security enforcement today: centralized services architecture

Centralized Services

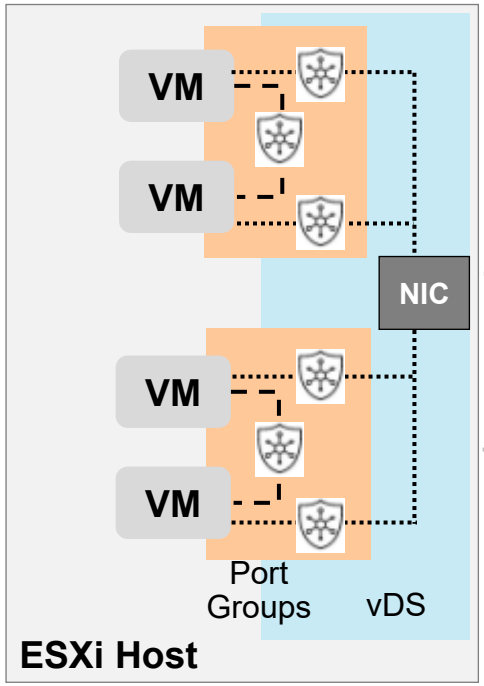
- Waste of bandwidth
- Congestion & high latency
- Complex to design & troubleshoot
- Limited to single function
- Very expensive

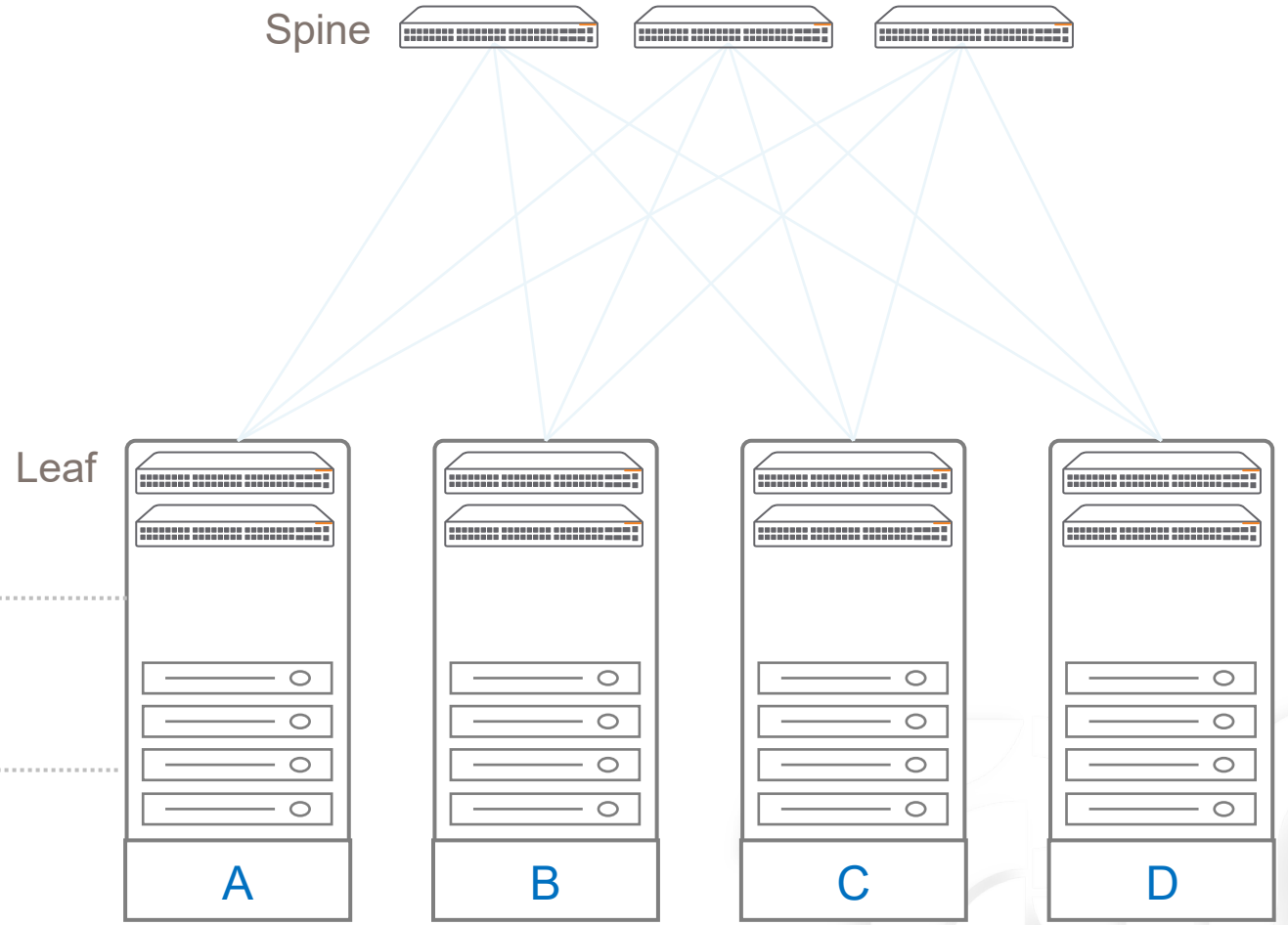


Security enforcement today: distributed Services architecture

Software based Services

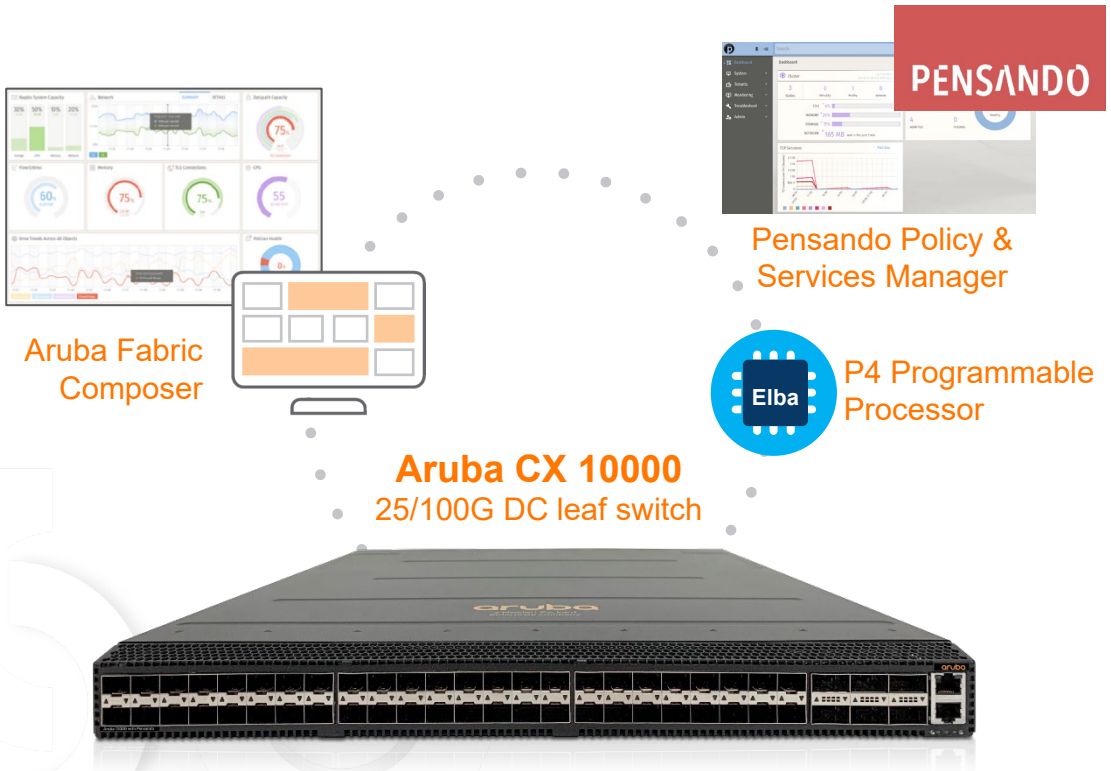
- High resource requirement on host (CPU/Memory)
- Congestion & high latency (ms)
- Complex to design & troubleshoot
- Very expensive (Licenses)





Distributed services in the Data Center with CX 10000

Aruba CX 10000 Distributed Services Switch - Powered by AMD Pensando



70% of all breaches caused by end point security vulnerabilities, followed by Lateral Movement

–1RU Fixed Switch Form Factor:

- T3 Switching ASIC - 3.2 Tbps, 32MB Buffer (shared)
  - **Used for forwarding/routing/other features**
- 2 x Pensando Elba (7nm) Programmable Processor
  - **Used for smart stateful services (all forwarding performed by T3)**
- 2 x Redundant Power Supplies (N+1)
- AOS-CX Network OS, full protocol stack support

–Port Configuration:

- 48 x 1/10G/25G SFP28, 6 x 100G QSFP
- 1 x 1G RJ45 management, 1 x RJ45 console port, 1 x USB

–Services/Use-Cases:

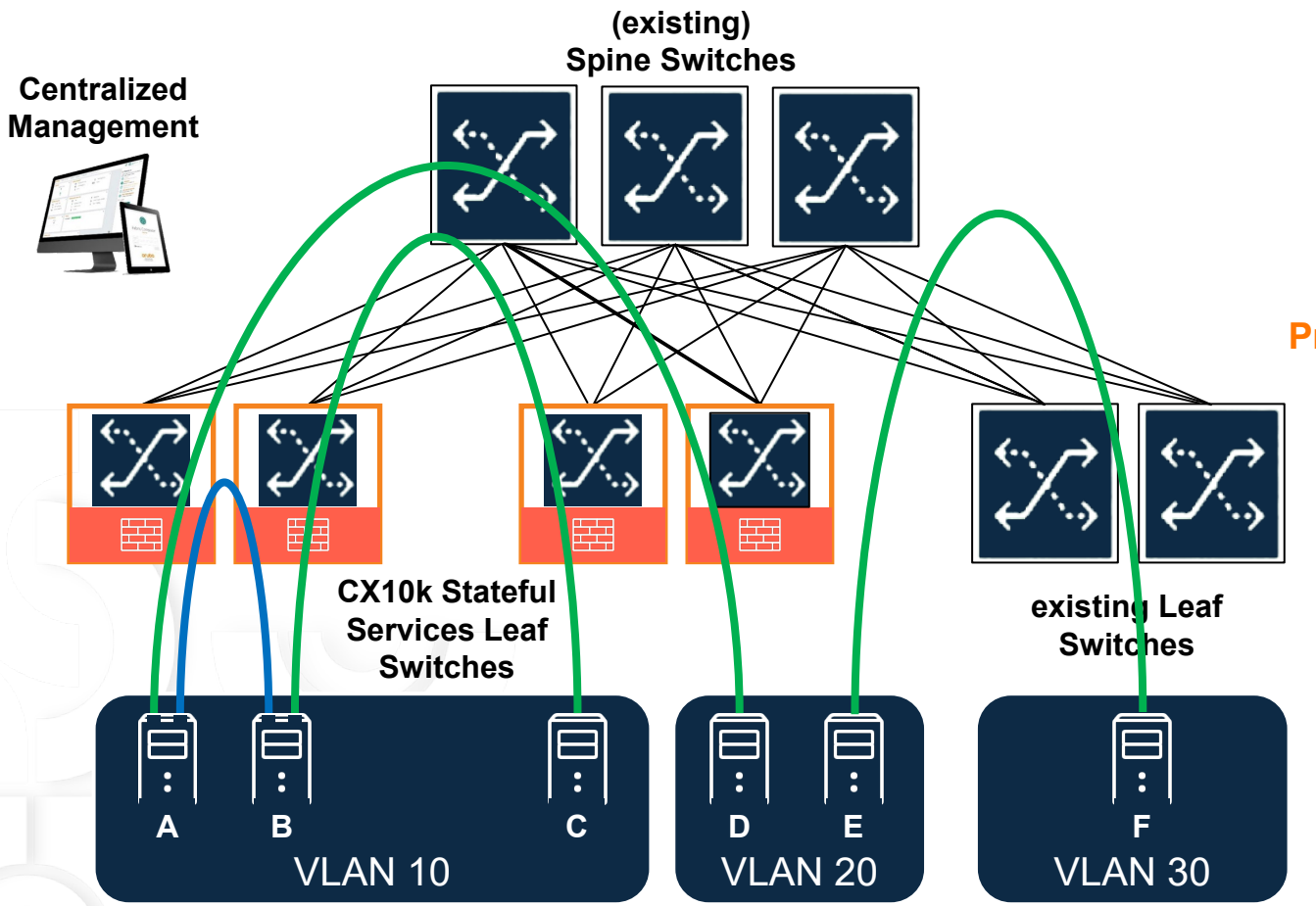
- East-West DC Segmentation (Distributed Firewall & DDoS)
- Micro segmentation
- Complete visibility (Packet Capture, Flow Logging/Statistics)
- IPSec
- NAT

–Platform Management Options:

- Aruba AFC & AMD Pensando PSM
- PSM & DevOps Tools (Terraform/Ansible), REST API

Stateful network firewall

Protect the services inside your Data Center



Secure traffic between two servers through Stateful Firewall:

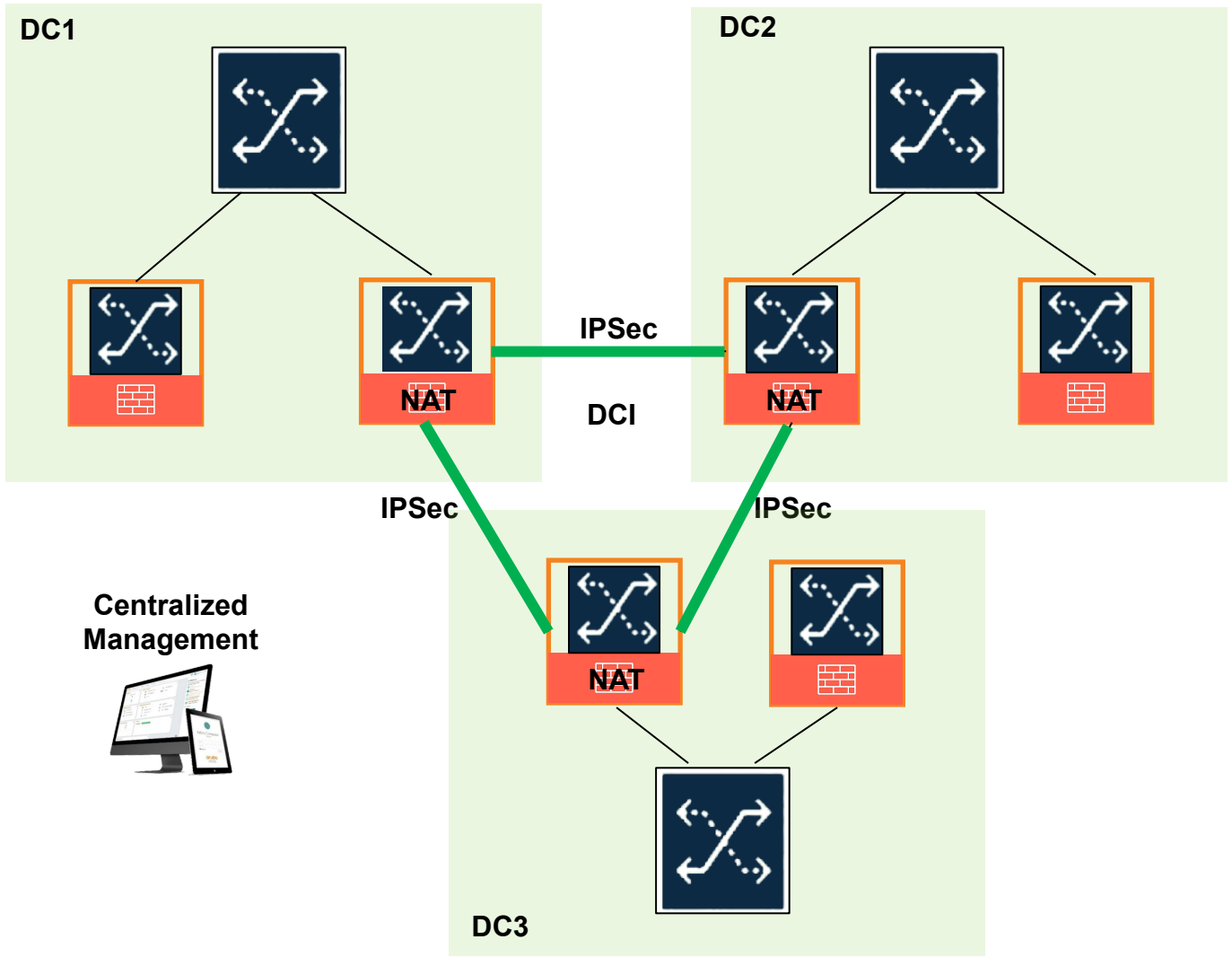
- In the same VLAN
- In different VLANs
- Both connected any leaf Distributed Services Switch
- Where one server is connected to an existing leaf
- High performance (800Gbps)
- Low latency (4us)

Protect the Unprotected:

- Hypervisors (management, storage)
- Backup Servers
- IP Storage Appliances
- Shared Services
- Bare Metal Servers

IPSec and NAT

Protect the services between data centers



Secure traffic between data centers:

- IPSec
- Network Address Translation
- 400Gbps full duplex encryption/decryption
- Includes stateful layer 4 firewalling

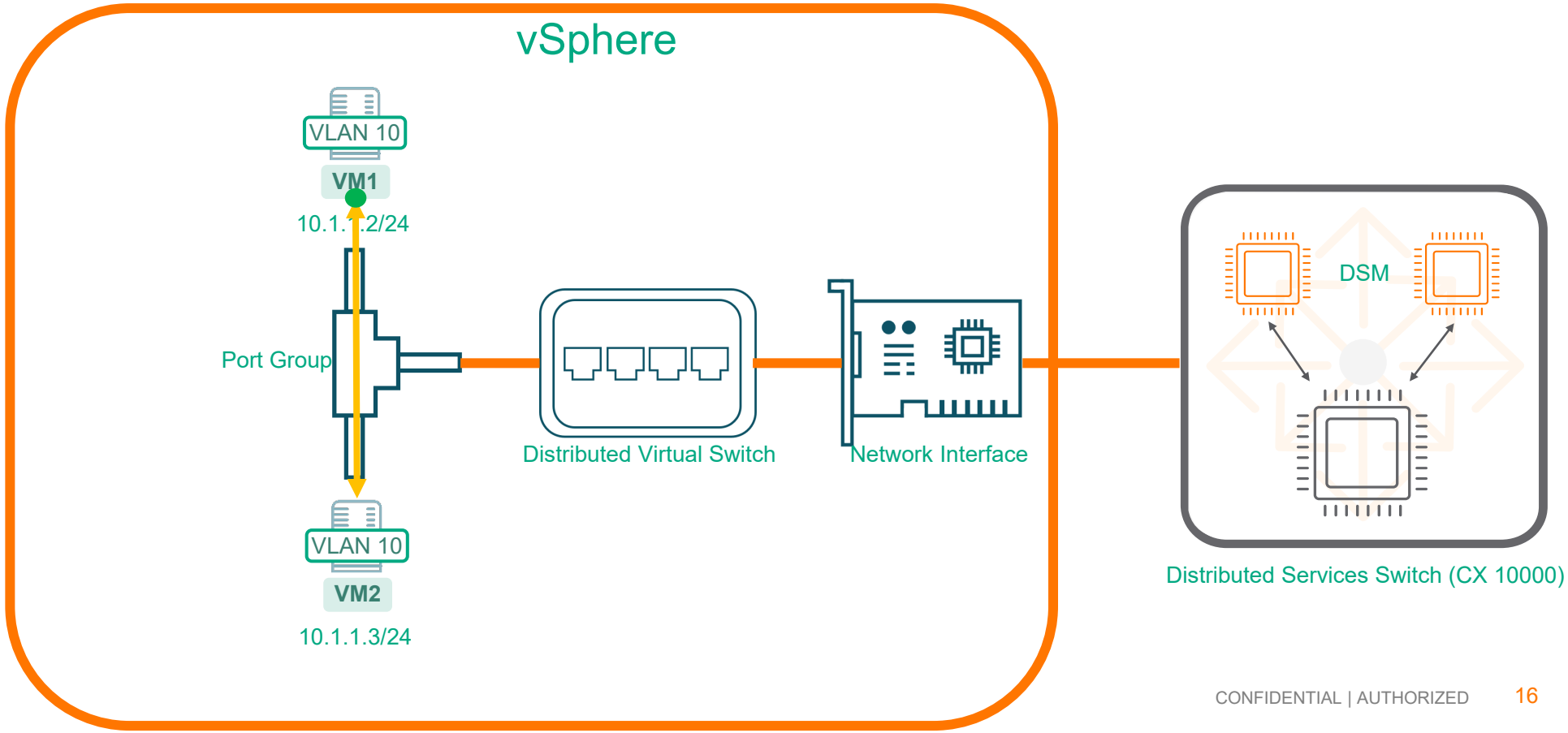
Micro Segmentation

How does it work: micro segmentation

Traffic inspection for workloads that are on the same network

By default, on a vSphere port group traffic within the port group is allowed

How can we create micro segmentation that allows stateful firewalling between workloads that are on the same subnet/VLAN?

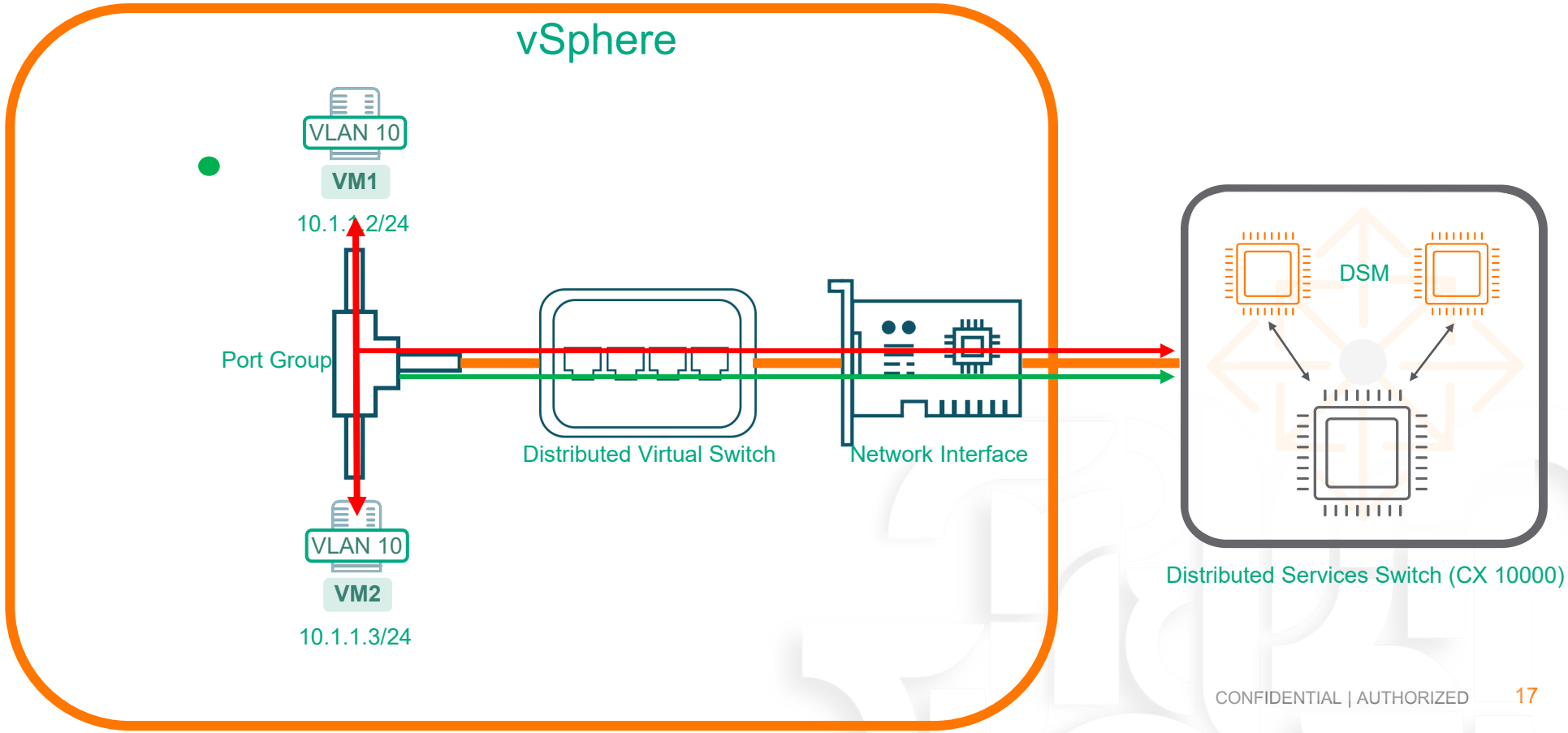


How does it work: micro segmentation

Traffic inspection for workloads that are on the same network

Micro segmentation can be achieved through Private VLAN functionality in vSphere and on Aruba CX switches

The primary VLAN (VLAN 20) is used for egress traffic into the CX 10000. VLAN 10 traffic is also egressed, there is still isolation between hosts

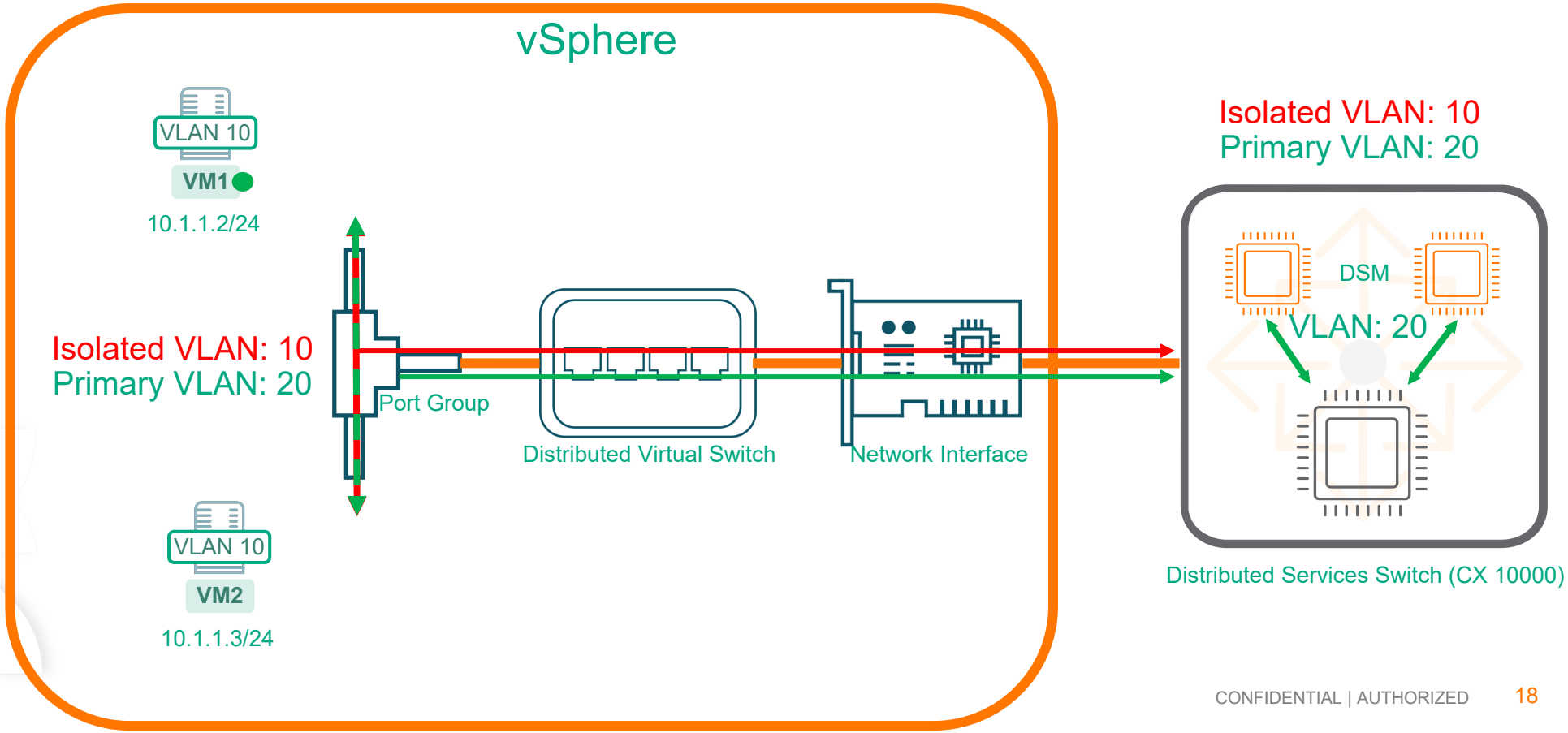


How does it work: micro segmentation

Traffic inspection for workloads that are on the same network

The CX 10000 is also configured for Private VLAN where VLAN 10 is the isolated VLAN and VLAN 20 the primary (promiscuous) VLAN

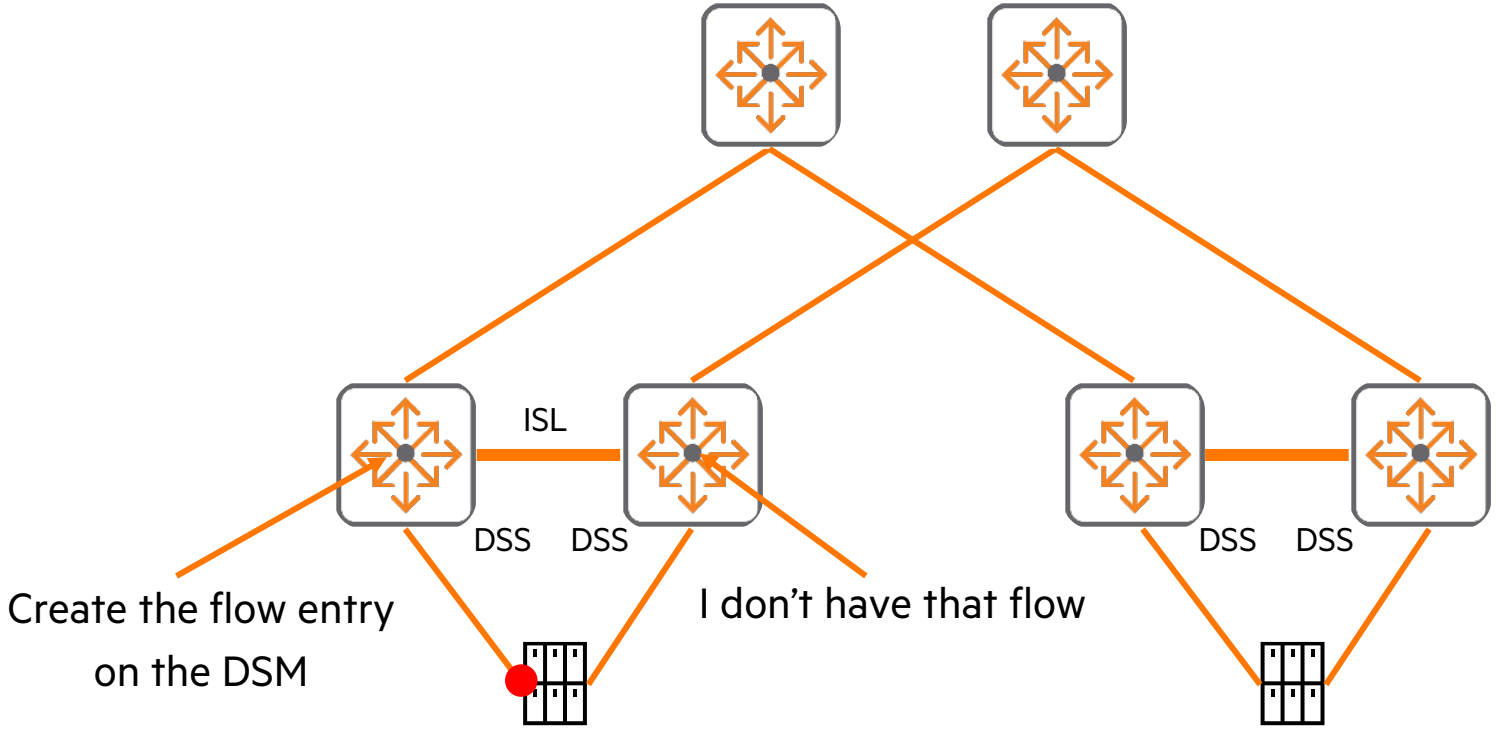
When a VLAN (Network) exists on the DSM for the primary VLAN (20), traffic is redirected to the DSM for stateful inspection



VSX flow synchronization

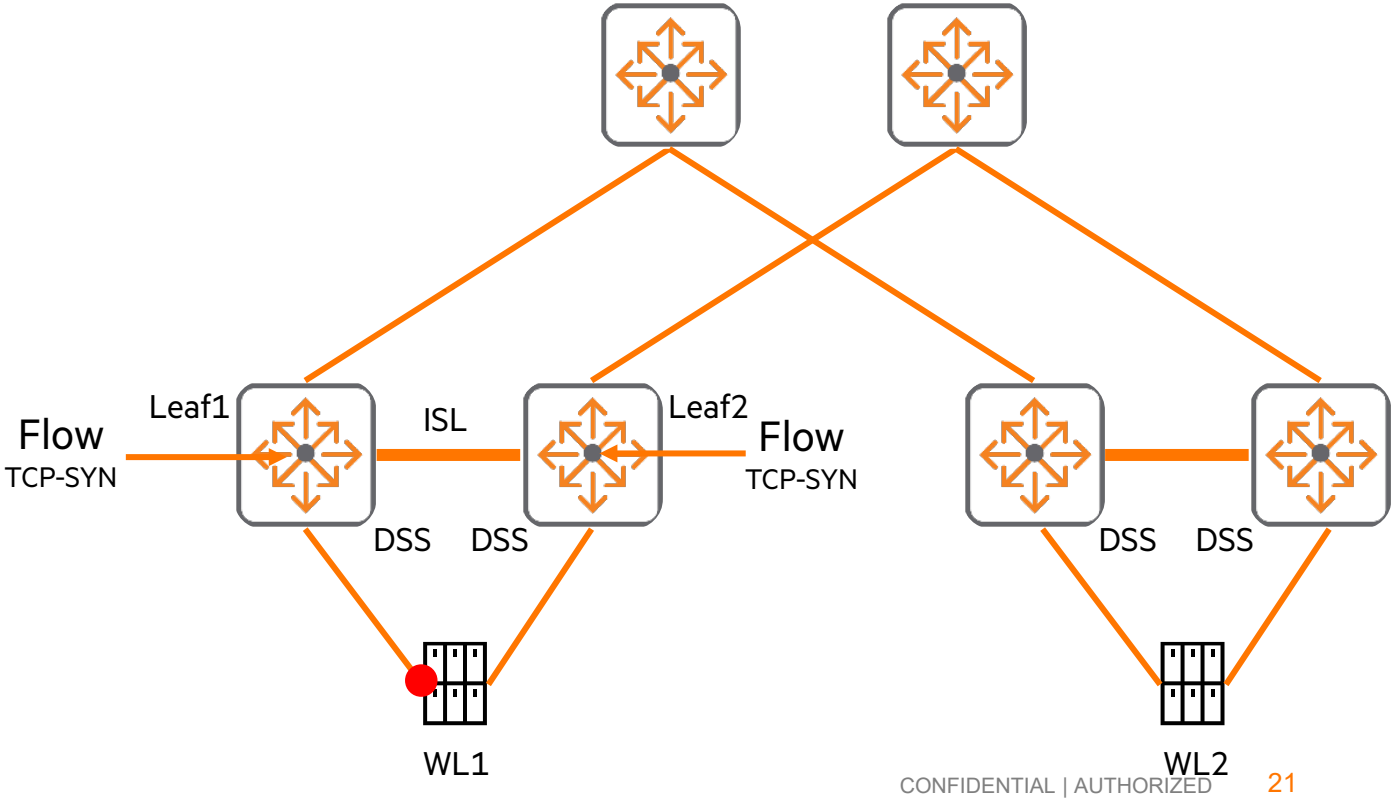
Flow behavior with VSX

- VSX is a dual control plane construct, this also applies to the Distributed Services Modules (AMD Pensando)
- Asymmetric traffic can result in forward and reverse packets arriving at different VSX nodes
  - Non asymmetric traffic can shift from one VSX node to another in case of ISL or host link failures
  - Traffic from spine layer can shift due to routing changes
- With VSX, the connection state flow table on the DSM's are kept in complete sync



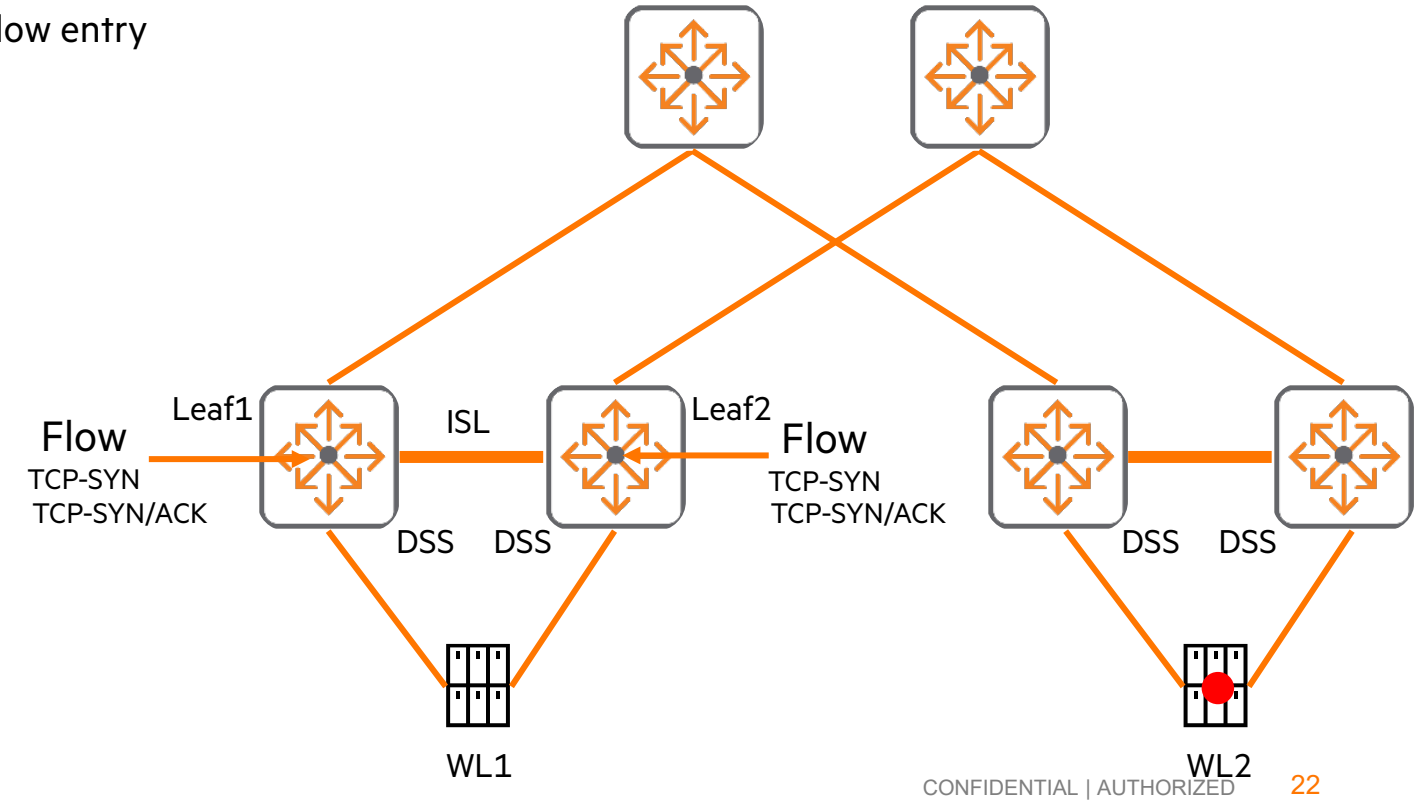
Flow behavior with VSX, how does this work

- WL1 sends a TCP SYN packet to the destination
- There is no flow entry on Leaf1, a copy of the packet is sent to Leaf2 through the ISL
- Leaf2 evaluates the flow, installs the flow entry (on Leaf2)
- Leaf2 sends back the original packet to Leaf1
- Leaf1 evaluates the flow and installs the flow entry with the SYN (5 tuple)
- Leaf1 sends the packet to the destination (WL2)



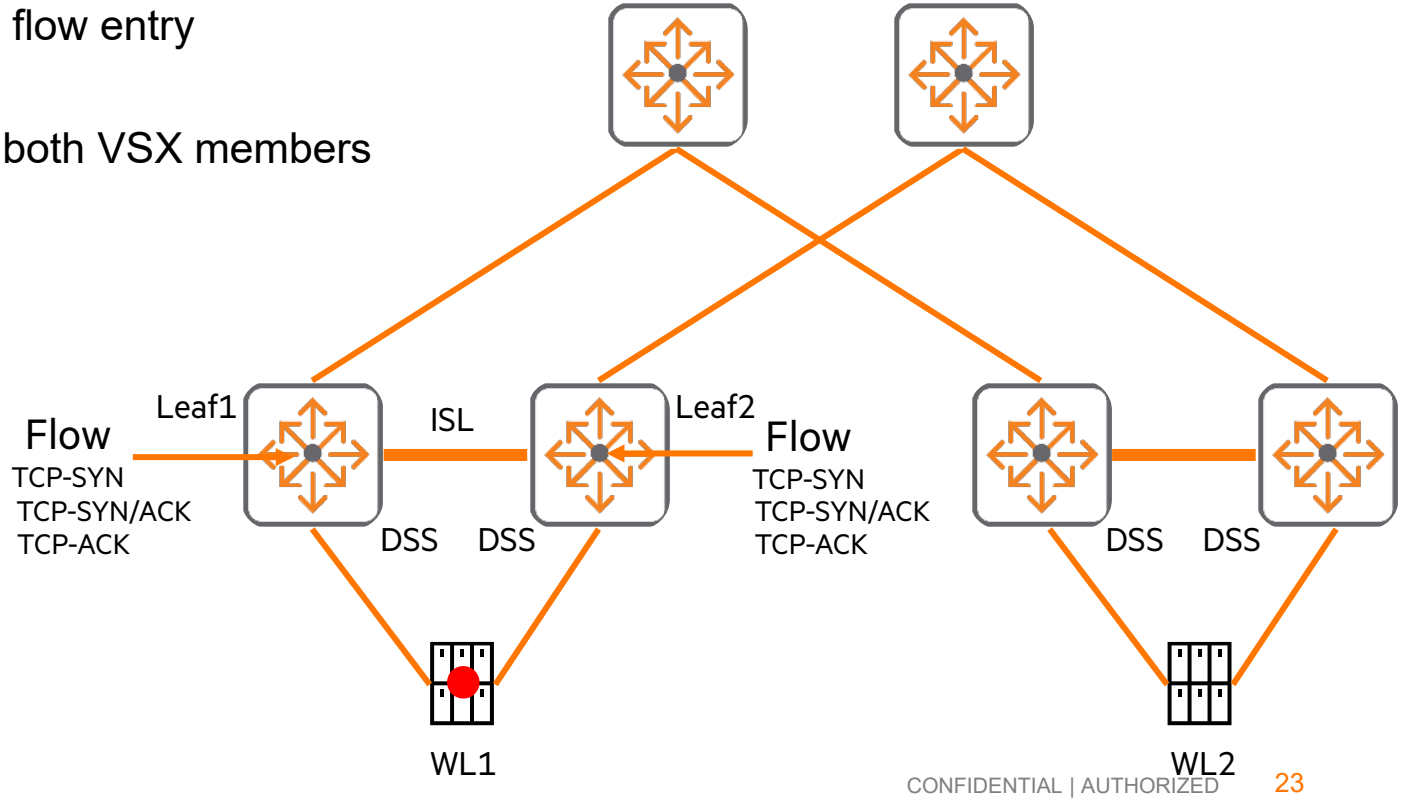
Flow behavior with VSX, how does this work

- WL2 returns the SYN-ACK on a different path
- Leaf2 has the flow programmed, but it is not completed, it requires SYN / SYN-ACK / ACK for a complete flow entry
- A copy of the packet is sent to Leaf1 through the ISL
- Leaf1 adds the SYN-ACK to the flow table entry
- Leaf1 sends back the original packet to Leaf2
- Leaf2 evaluates the flow and adds the SYN-ACK to the flow entry
- Leaf2 sends the packet back to the source(WL1)



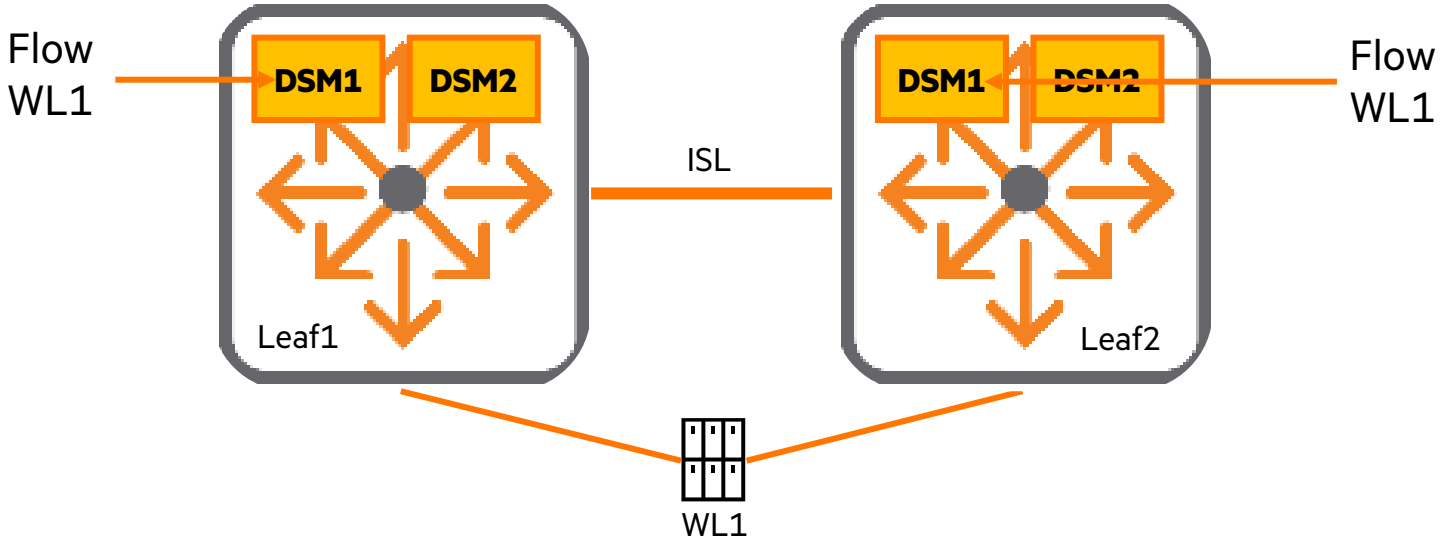
Flow behavior with VSX, how does this work

- WL1 sends the ACK to the destination
- Leaf1 has flow programmed, but it's not completed, it requires SYN / SYN-ACK / ACK for a complete flow entry)
- A copy of the packet is sent to Leaf2 through the ISL
- Leaf2 adds the ACK to the flow table entry
- Leaf2 sends back the original packet to Leaf1
- Leaf1 evaluates the flow and adds the ACK to the flow entry
- Leaf1 sends the packet to the destination (WL2)
- The flow entry is completed and synchronized on both VSX members



Flow behavior with VSX, how does this work

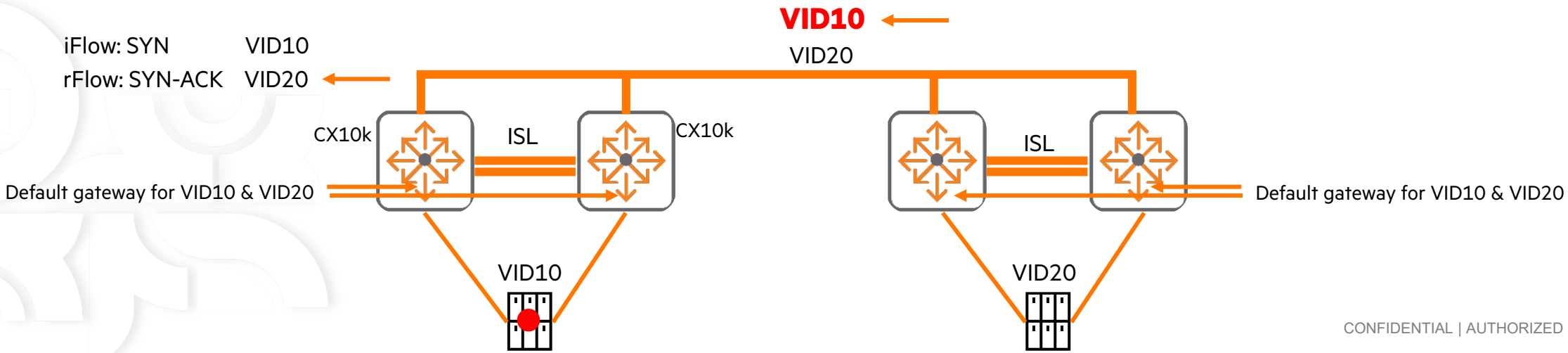
- If a VSX member reloads or ISL failure, there is no traffic interruption because the flow state is synchronized
- When ISL is restored, a full flow sync will take place
- Support for Live upgrades (ISSU), there is a graceful switch over during upgrade
- A flow is only stored on one ASIC (DSM) on each VSX member
  - When a flow is stored on DSM1 on Leaf1, it is also stored on DSM1 on Leaf2
  - Hashing mechanisms within the individual switch ensures redirection to the same DSM



Collapsed backbone with CX 10000

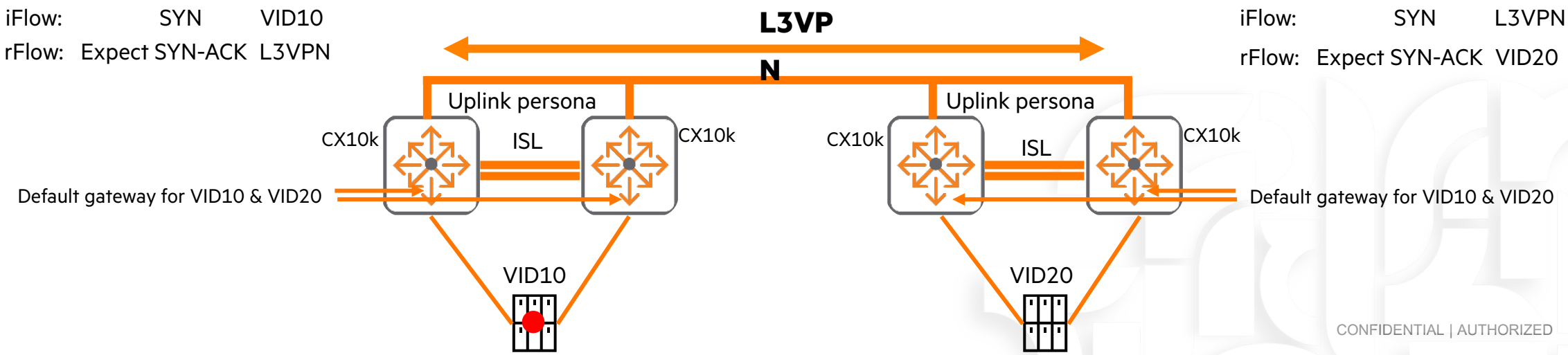
Collapsed backbone with CX 10000

- A very common use case for DCN is a single layer collapsed backbone, directly connected servers and campus backbone
- Redirection of traffic to the DSM (stateful inspection) leads to asymmetric routing and flow miss table
- Workload sends SYN packet from VLAN 10 to destination in VLAN 20
- Packet is inspected and stored in the flow table mapped to VLAN 10 on the destination switch, marked as iFlow
- Packet returned to T3, routes the packet to VLAN 20 and informs DSM that it expects SYN-ACK on VLAN 20 on return flow (rFlow)
- Packet is forwarded to the destination on VLAN 20
- Destination responds with a SYN-ACK packet, packet is received by the destination switch on VLAN 20 and is routed locally to VLAN 10
- The SYN-ACK packet is sent to the source switch over VLAN 10, the source switch however expects the SYN ACK on VLAN 20
- There is a flow miss on the returning packet, the flow entry is incomplete, and the session is dropped



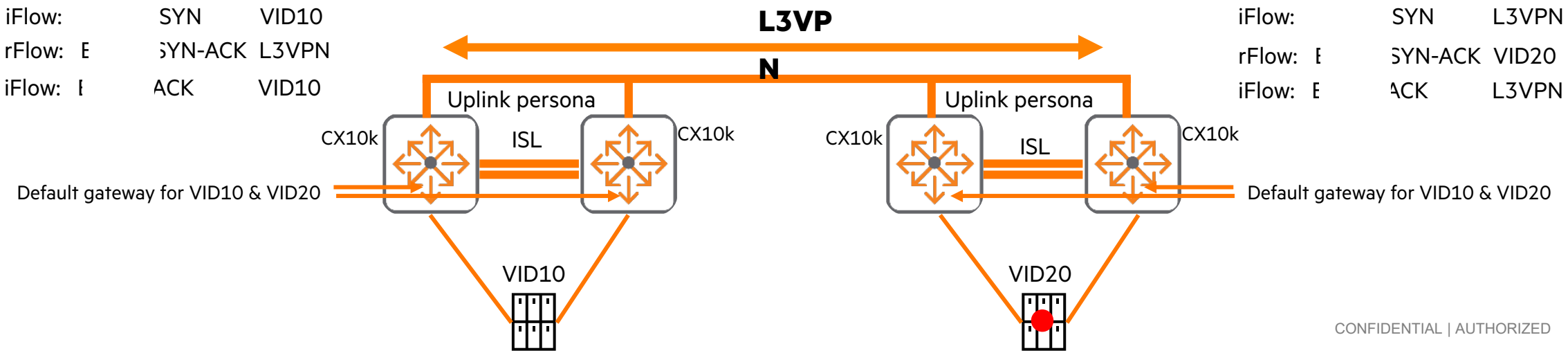
Collapsed backbone with CX 10000: Solution

- EVPN VXLAN provides symmetric routing between the VSX pairs where each VSX pair performs local routing
- This is accomplished by a L3VPN that performs routing in the overlay
- The links between the VSX pairs are configured as uplink persona, changes redirection behavior for ingress traffic
- SYN packet is sent from source, received by the source switch and redirected for inspection, flow entry is created for VLAN 10
- SYN packet is returned to T3, routes the packet into the L3VPN and notifies the DSM that it expects SYN-ACK from the L3VPN
- Source switch knows the destination IP prefix through EVPN (RT-5), after inspection, the SYN packet is sent over the L3VPN
- Destination receives the SYN packet; it is redirected for inspection (persona behavior) and flow is added
- SYN packet is returned to T3, routes the packet to VLAN 20 and notifies the DSM that it expects SYN-ACK from VLAN 20
- SYN packet is forwarded to the destination



Collapsed backbone with CX 10000: Solution

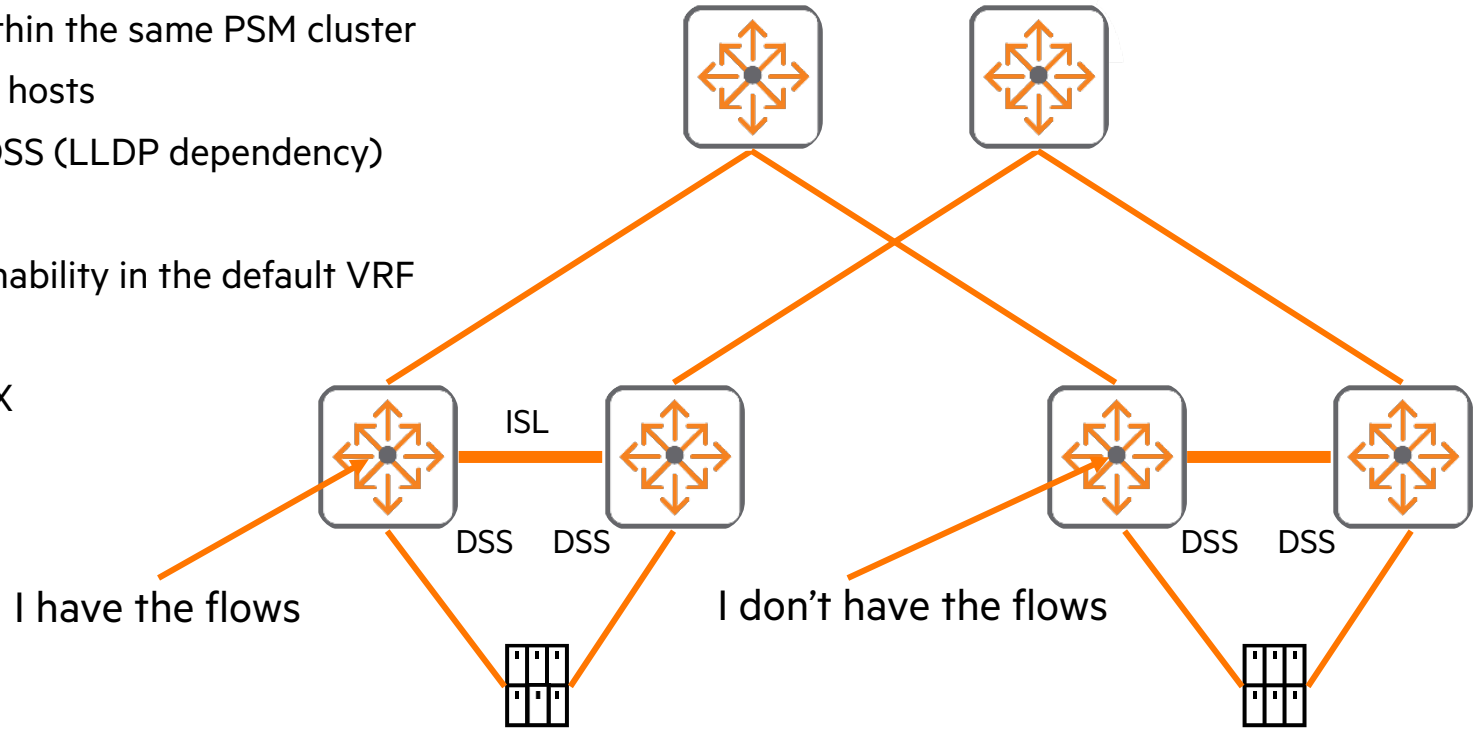
- Destination sends SYN-ACK back to the source, destination switch receives SYN-ACK on VLAN 20, packet is redirected for inspection
- SYN ACK is redirected for inspection; SYN-ACK is acknowledged (it is expected on VLAN 20) by the DSM and added to the flow
- SYN ACK packet is returned to T3, routes the packet into the L3VPN and notifies the DSM that it expects ACK from the L3VPN
- Destination switch knows the destination IP prefix through EVPN (RT-5), after inspection, the SYN-ACK packet is sent over the L3VPN
- Source switch receives the SYN-ACK packet; it is redirected for inspection (persona behavior) and flow is added (Expected on L3VPN)
- SYN ACK packet is returned to T3, routes the packet into the L3VPN and notifies the DSM that it expects ACK from VLAN 10
- SYN ACK packet is forwarded to the destination
- The same process is repeated for the ACK packet



Flow behavior with vMotion

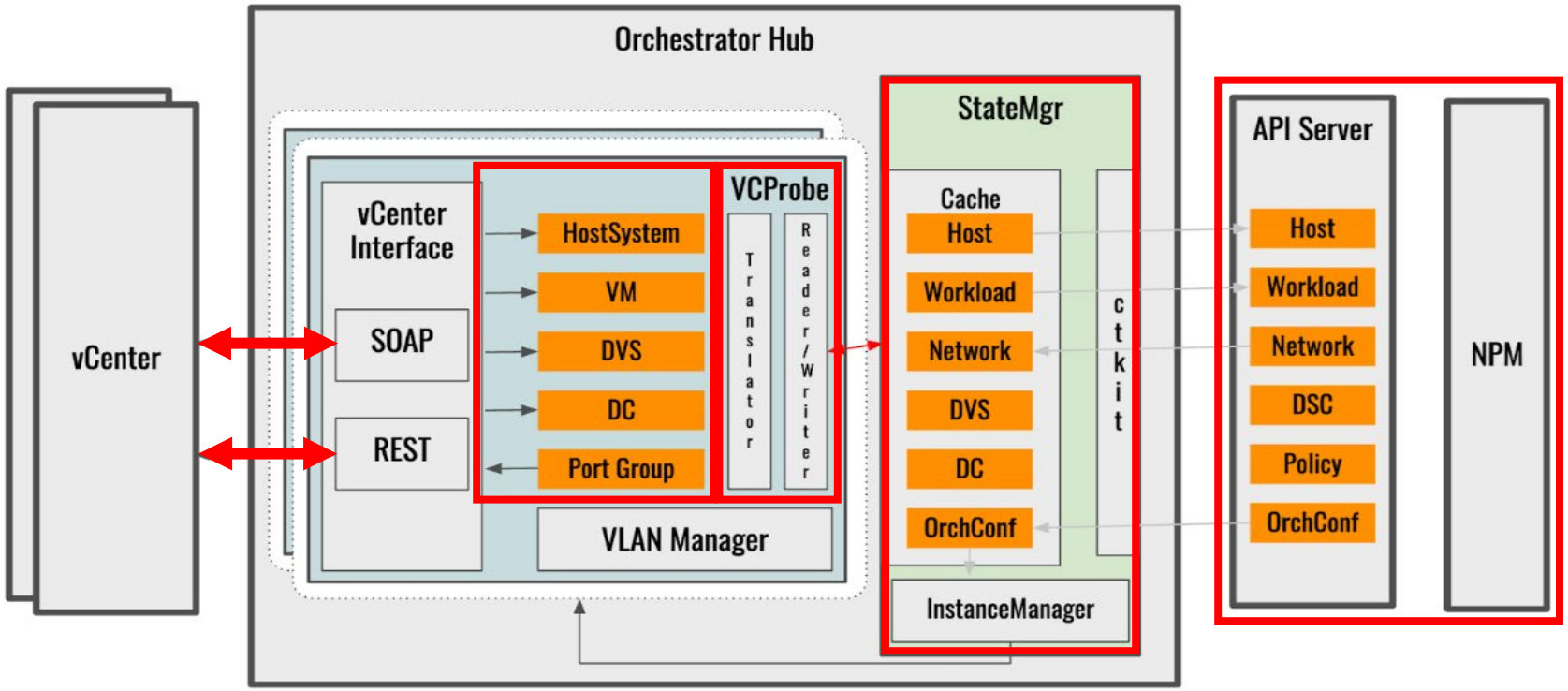
Flow behavior with vMotion

- Workload migration on ESXi hosts that are connected to different DSS switches is supported
  - Within the same PSM fabric across racks
  - No traffic disruption
  - Accomplished my migrating associated flows from the source to the destination DSS
- Requirements
  - Source and destination DSS only supported within the same PSM cluster
  - LLDP enabled on DSS ports connected to ESXi hosts
  - ESXi hosts must be directly connected to the DSS (LLDP dependency)
  - PSM must be integrated with vSphere
  - All DSS in the PSM cluster require inband reachability in the default VRF
    - Required for the flow migration sessions
  - Neighbor resolution must be enabled on the CX



Flow behavior with vMotion, how does it work

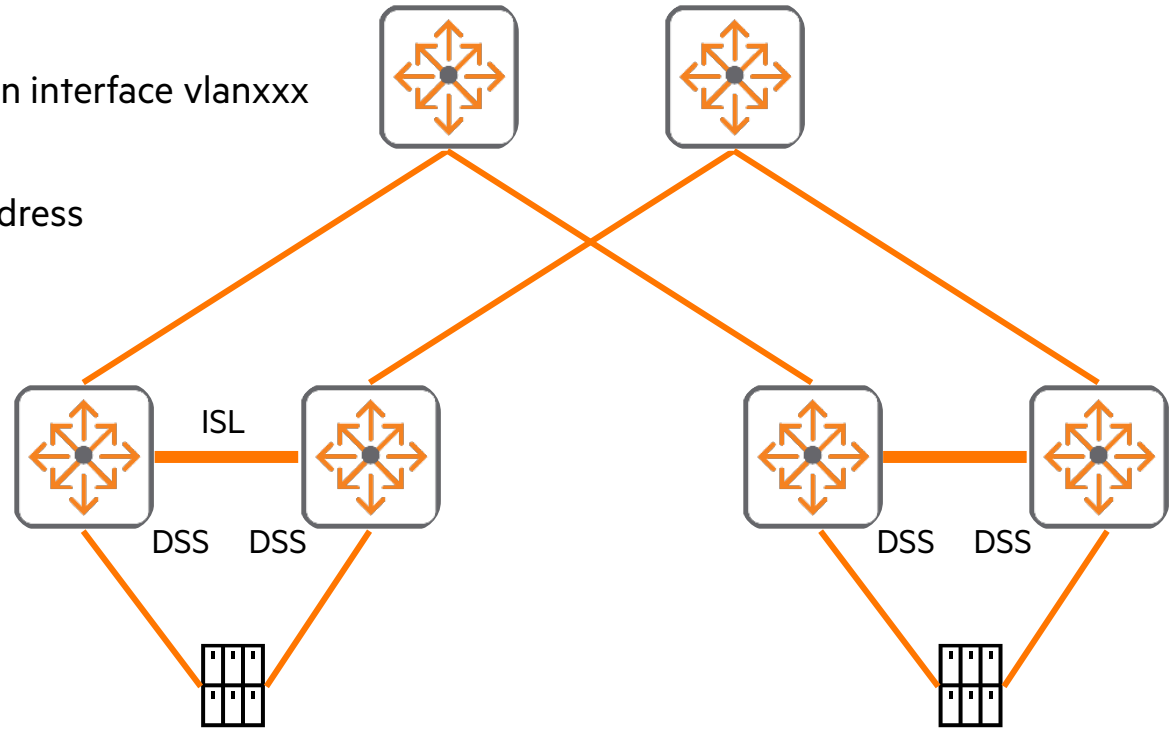
- Orchestrator Hub (service running on PSM) connects to vCenter via user provided credentials
- Orchestrator Hub reads the Host, Workloads and LLDP information from the vCenter via VIM (API) interface (LLDP has to be enabled on ESXi)
- Orchestrator Hub adds watcher for VM migration notifications
- Information received from vCenter is used to create the ESXi Host to local DSS mapping (on which switch is the host connected)
- ESXi locality database is maintained in NPM (Network Policy Manager) by correlating the Host object information and LLDP information



Flow behavior with vMotion, how does it work

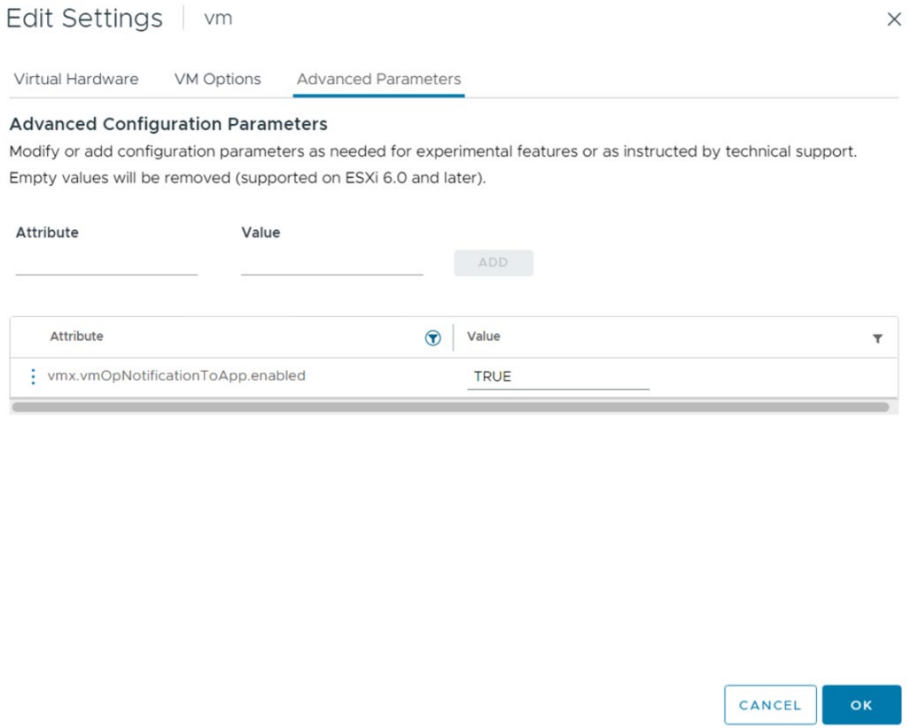
- On the DSS (CX 10000)
- Source and destination DSM cannot communicate directly
  - NAT between DSS and DSM (internal to the switch, DSMs are using 169.254.x.x)
  - Source and Destination DSS are used for translation
  - Allows migrating associated flows from the source to the destination DSS
    - Command to achieve reachability (in the default VRF):
      - Cx10k(config)#ip source-interface workload_migration interface vlanxxx

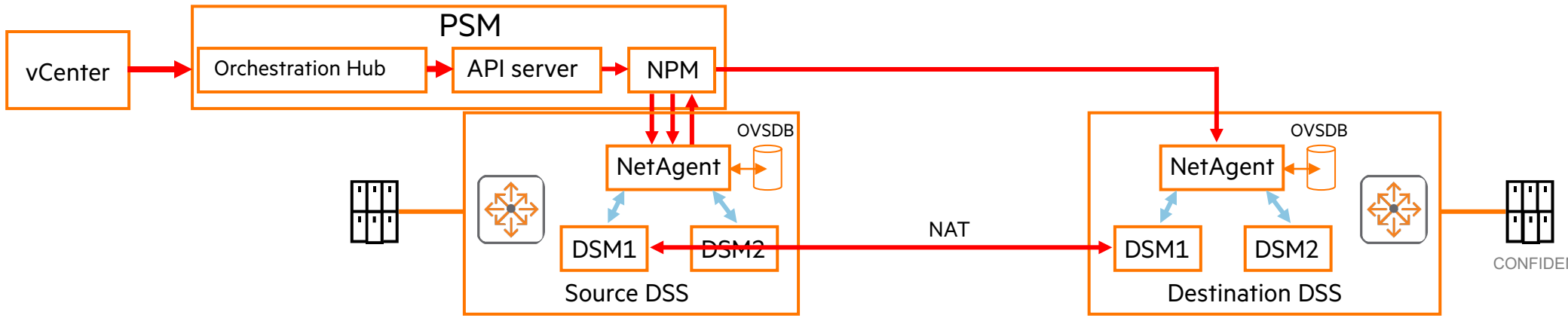
- During vMotion, the vCenter notification message contains the MAC address of the moving workload
- DSM needs to know the IP of the VM being moved to derive the associated flows
- MAC-IP bindings on VLAN ports are created by snooping the ARP request packets including the transit ARP on all the ports
  - Command to trigger ARP snooping on DSS:
    - cx10k(config)# dsm workload-migration



Flow behavior with vMotion, how does it work

- vCenter sends a migration notification to the PSM Orchestrator Hub
  - Includes the IP address of the source, destination host and MAC address of th
- The Orchestrator hub forwards the notification internally to the Network Policy Mar
- Network Policy Manager finds the source and destination DSS and triggers the flow
- Network Policy Manager notifies the source DSS
  - Endpoint migration has started, this is the source MAC address
  - Give me the IP address and DSM of this MAC address (MAC-IP binding) → d
- Source DSS returns the IP address and DSM reachability to the Network Policy Mana
- Network Policy Manager sends a migration message to the destination DSS
  - Contains the moving endpoint MAC-IP/VLAN ID and source DSM reachability
- Destination DSM pulls the flow information of the migrating VM
  - From the Source DSM through the NAT'ed connection → ip source-interface interface vlanxxx





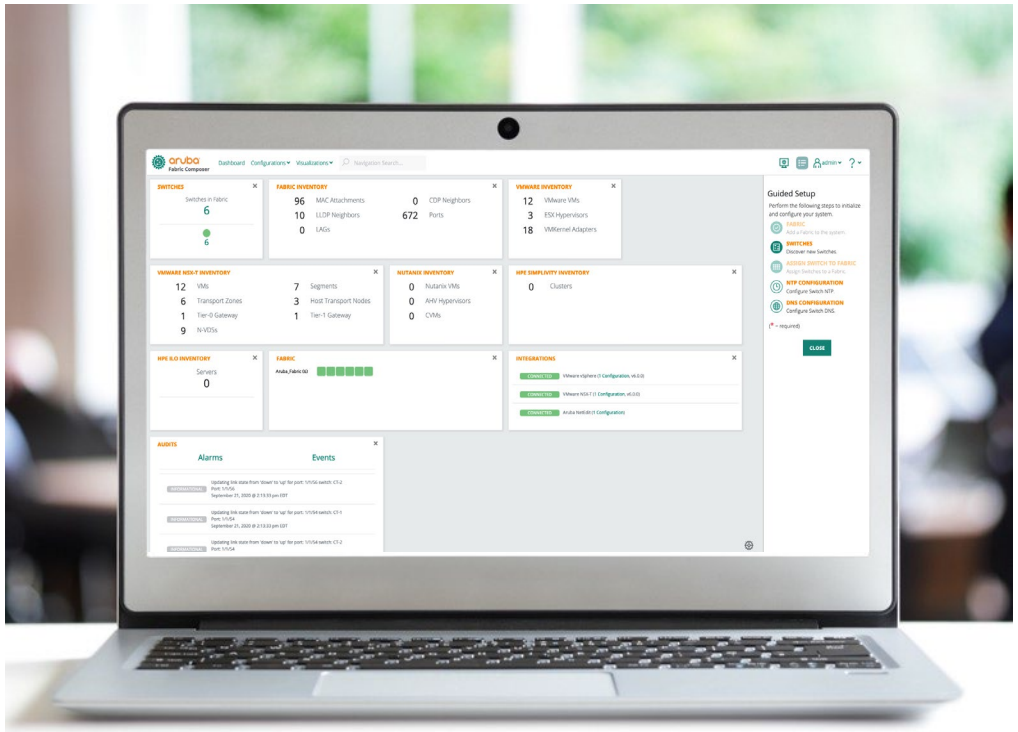
Flow behavior with vMotion, the process

- In some situations, vMotion occurs within a host, when this happens, the DSM is not aware of the flows
  - Session persistence cannot be guaranteed
- Solution for this:
  - Micro segmentation with Private VLANs
  - This redirects the workload traffic to the DSS and can then be inspected by the DSM
  - The process for moving the flow information when a vMotion occurs will take place because they are visible now

Data Center orchestration and management

Aruba Fabric Composer

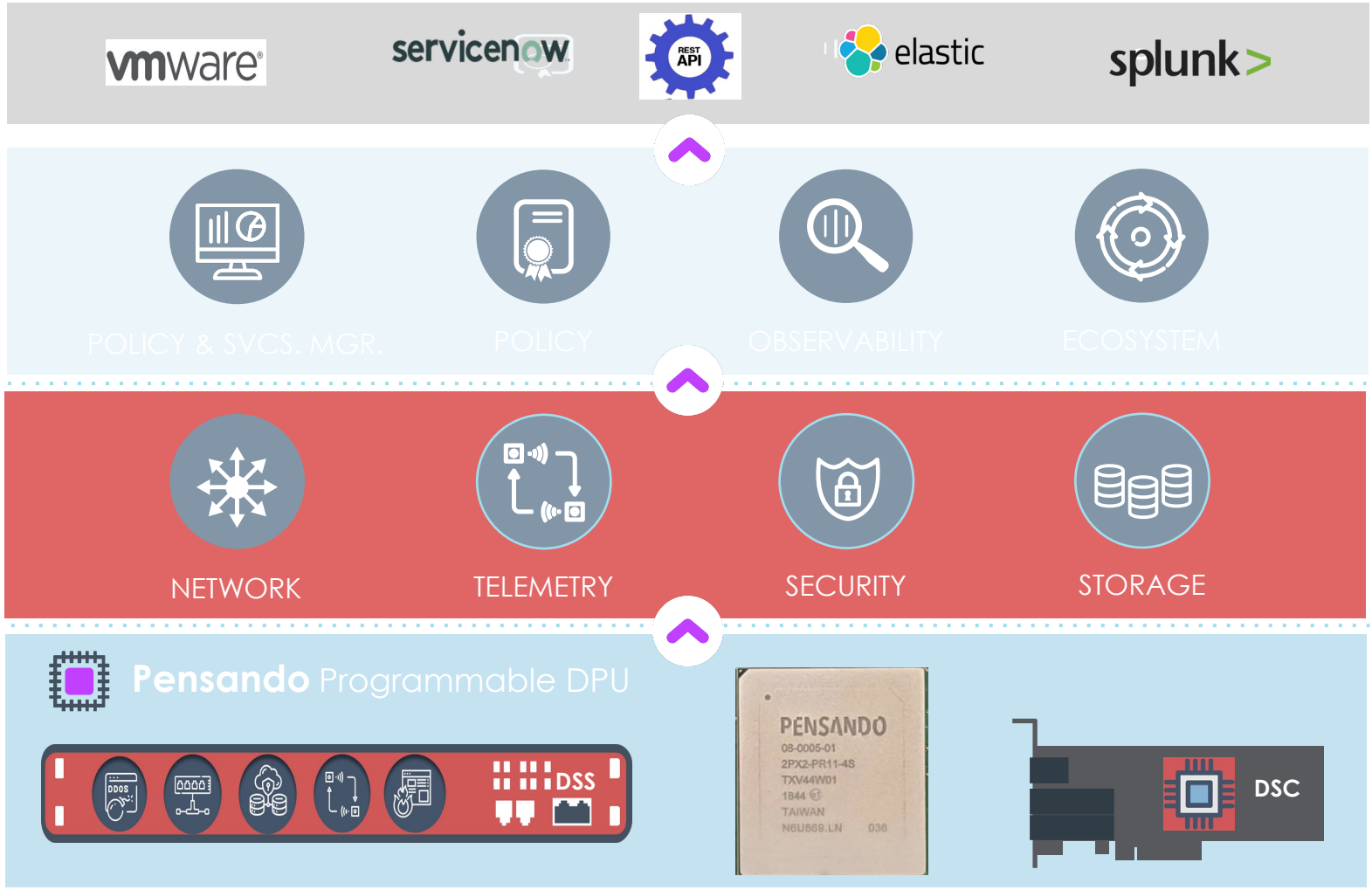
The on-site data fabric orchestration system



Key Features & Benefits

- Simplified provisioning & orchestration
- Complex workflow automation
- Manage and monitor global network configuration
- Integrate with 3rd party data center orchestration systems
- Integration with HPE Infrastructure hardware and software
- Visualize data center infrastructure
- Automate lifecycle events in the data center
- Holistic troubleshooting of end-to-end network connectivity

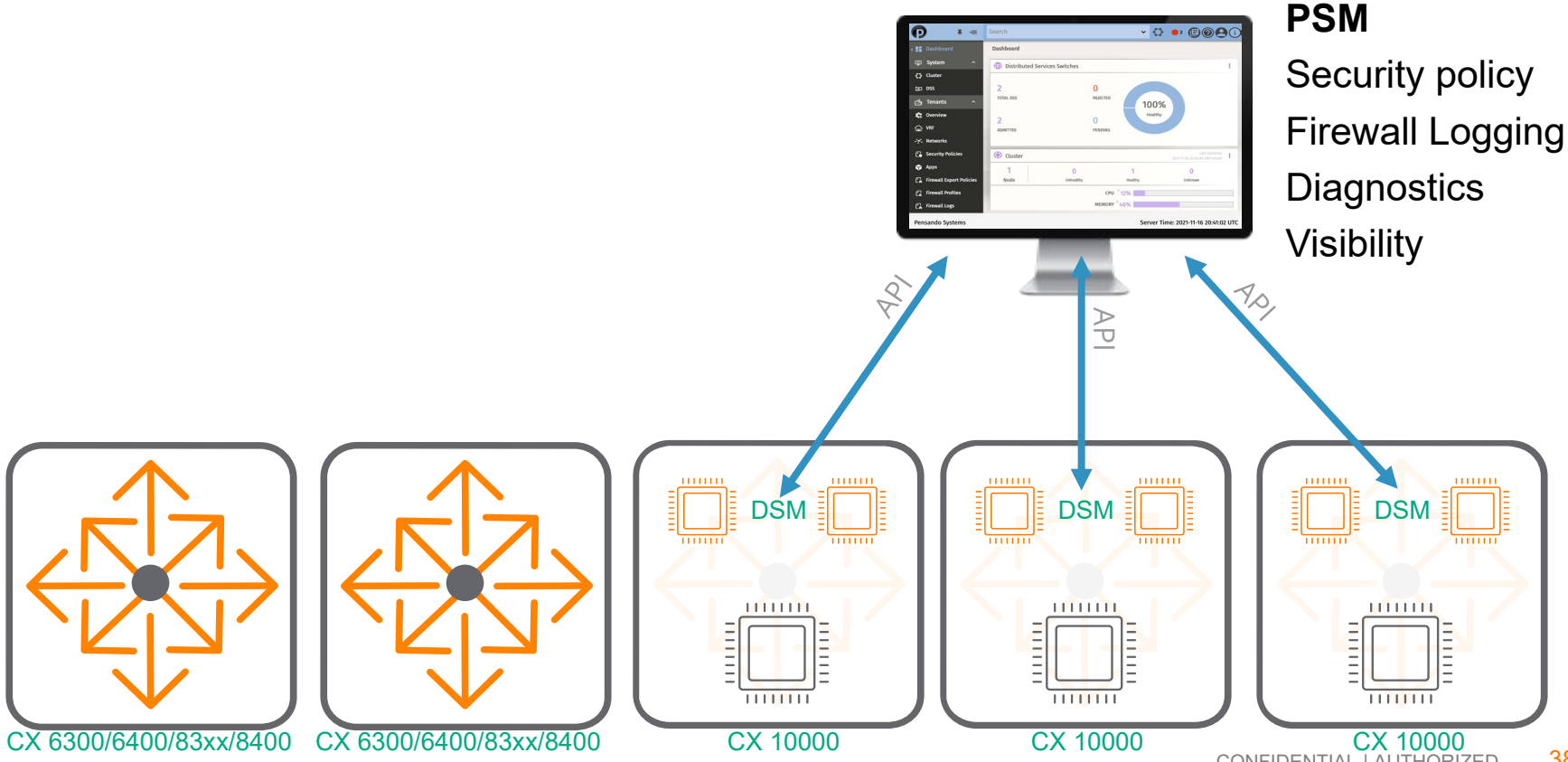
The Pensando Distributed Services Platform



– **Policy & Services Manager**
  – Centralized Lifecycle Management
  – Ensures Full-stack Enterprise-grade Security & Policy Compliance
  – REST-API integration with existing apps
– **Distributed Services**
  – Software-defined Services
  – Inline All-the-time at Wire-speed
– **Programmable ASIC**
  – Form-Factor Agnostic
  – Designed for Security
  – Low Power/Latency/Jitter
  – High Bandwidth & Scale

Orchestration and management

Pensando Services Manager provides policy enforcement, firewall logging, diagnostics and visibility for the DSM's
Pensando Services Manager does not provide fabric and switch orchestration and management

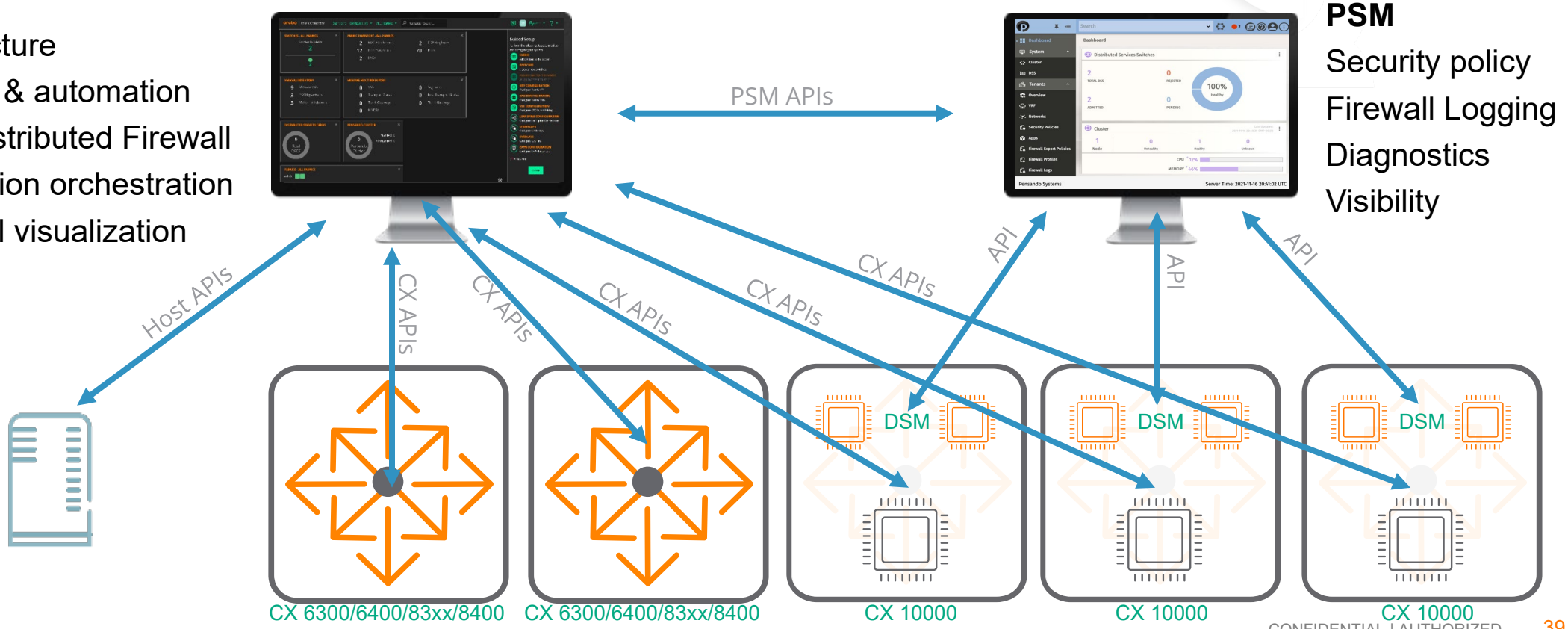


Orchestration and management

Aruba Fabric Composer provides datacenter orchestration, configuration and management for CX switches
Aruba Fabric Composer allows for security policy management by means of PSM API exchange between AFC and PSM
Aruba Fabric Composer has tight integrations with many third-party solutions (vSphere, Nutanix, Simplivity, iLO, etc)

AFC

Unified infrastructure
Fabric discovery & automation
Policy – ACL, Distributed Firewall
Micro segmentation orchestration
Physical & virtual visualization



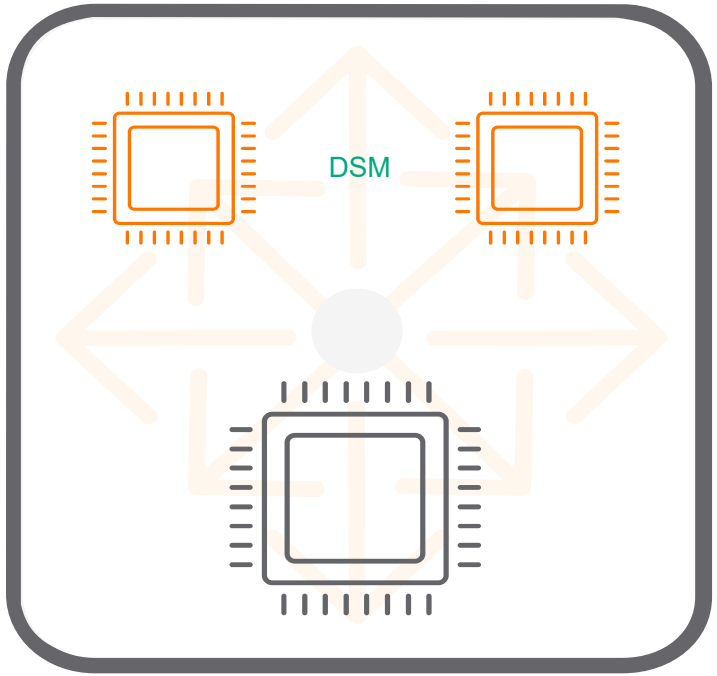
PSM

Security policy
Firewall Logging
Diagnostics
Visibility

Data Center flow visibility

Visibility

From Operational Visibility to Observability



NON-SAMPLED

SYSLOG
IPFIX

XDR

FORTINET paloalto NETWORKS exabeam splunk> CROWDSTRIKE

SIEM

splunk> LogRhythm Azure Sentinel ALIEN VAULT

DATA LAKE

elasticsearch + logstash + kibana

Chronicle

Thank you

Dik van Oeveren, Consulting Systems Engineer
dik.van.oeveren@hpe.com