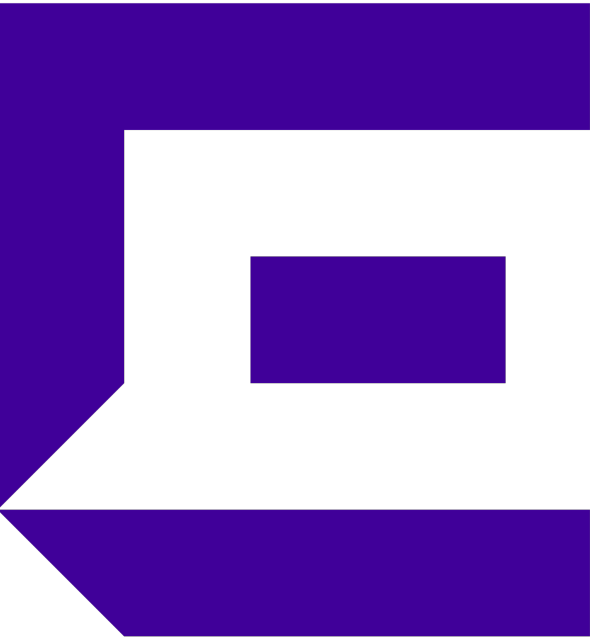


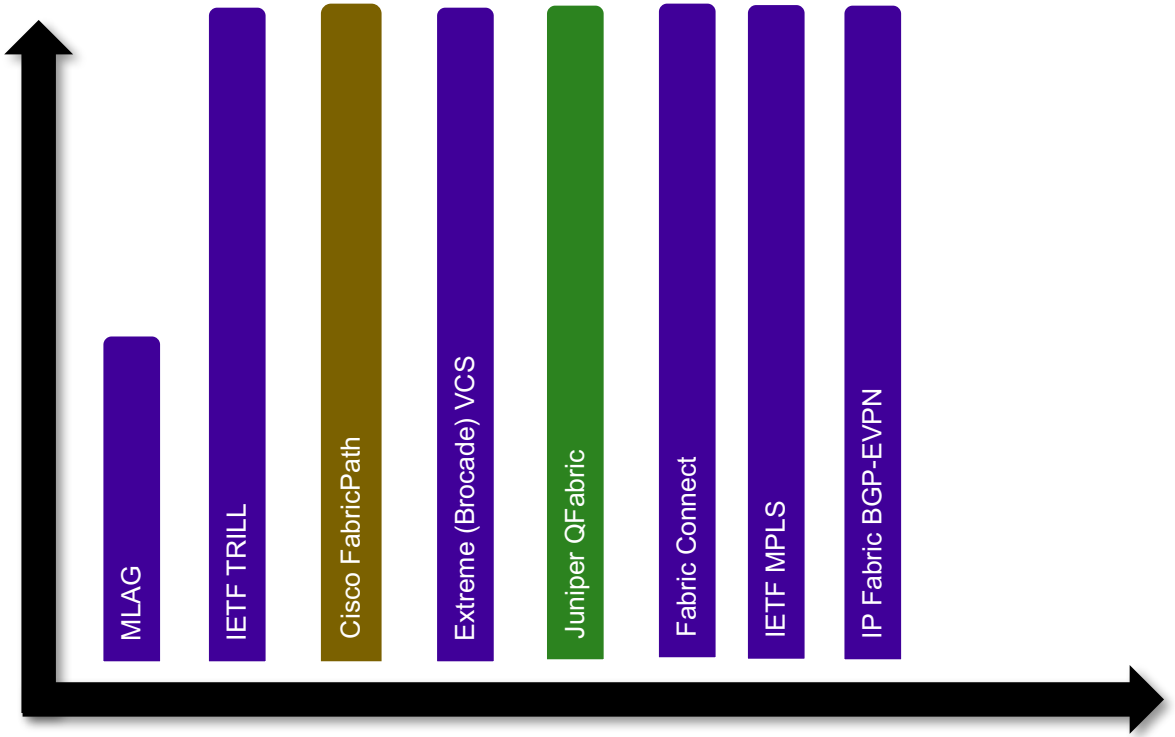
Extreme Networks

How to Build Scalable and Resilient Fabric Networks

Mikael Holmberg
Distinguished Systems Engineer

Extreme®
Connect Beyond the Network

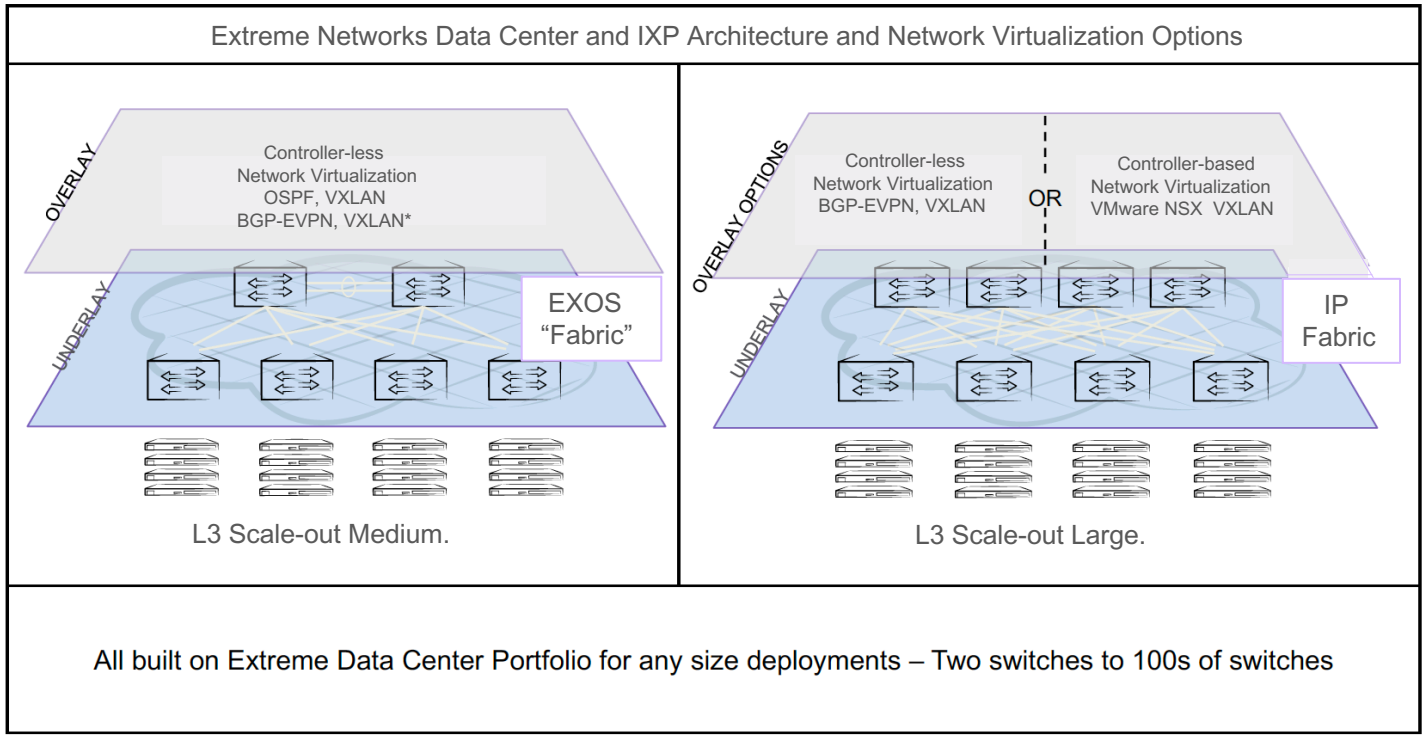
Fabrics



Next Generation Data Center Technology Stack

Orchestration and Management		Vendors Specific and Opensource
Automation	Visibility and Automation	Workflow Composer, Automation Suites & Integrations
Visibility	Visibility and Automation	Flow Optimizer, Network Packet Broker, Insight Architecture, Visibility Services
Programmability	Infrastructure	REST, NetConf, OpenFlow, gRPC and more
Network Virtualization	Infrastructure	BGP-EVPN, NSX (VXLAN), MPLS
Protocols	Infrastructure	IP Fabric, VCS Fabric
Flexible Platforms and ASICs	Infrastructure	SLX, MLX and EXOS

IXP Architecture Flexibility

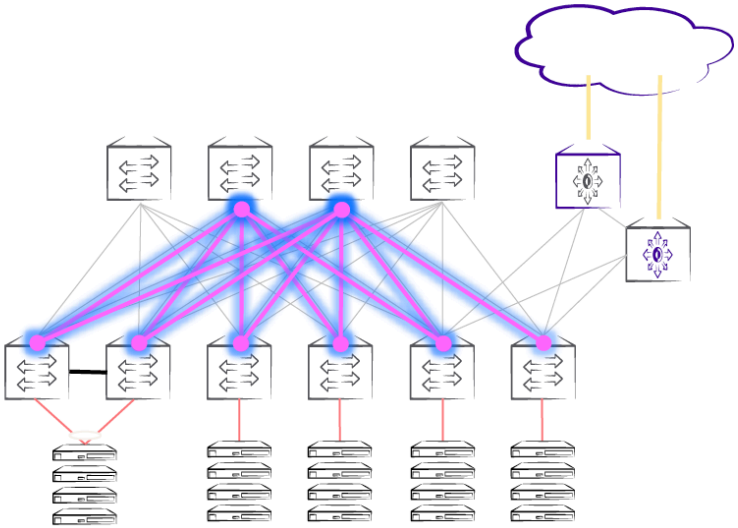


* EXOS 22.5

IXP Portfolio

Automation & Visibility	**N**etwork **V**isibility & **A**nalytics	**E**xtreme **W**orkflow **C**omposer	**E**xtreme **F**low **O**ptimizer
Place in Network	"Static IP Fabric"	IP Fabric Core Routing	Core Routing
Platforms	**EXOS**	**SLX-OS**	**NetIron**
	EXOS Ethernet Switches	SLX Ethernet Switches/Routers	MLX Ethernet Routers

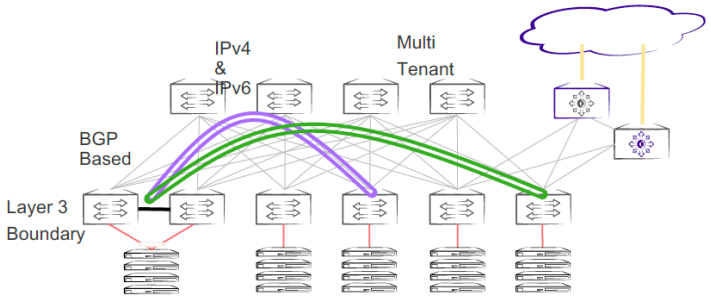
IP Fabric in IXPs



- Based on the community driven architectures and technologies of the digital era
- Simplified protocol stack; Internet proven IP and BGP
- Consumable by customers regardless of their size – from small to hyper-scale.
- Adapts to the resource constraints; skills, cost and time to value
- Available on VDX and SLX family platforms (EXOS possibly)

Protocols

Simplified Protocol Stack with Internet Reliability

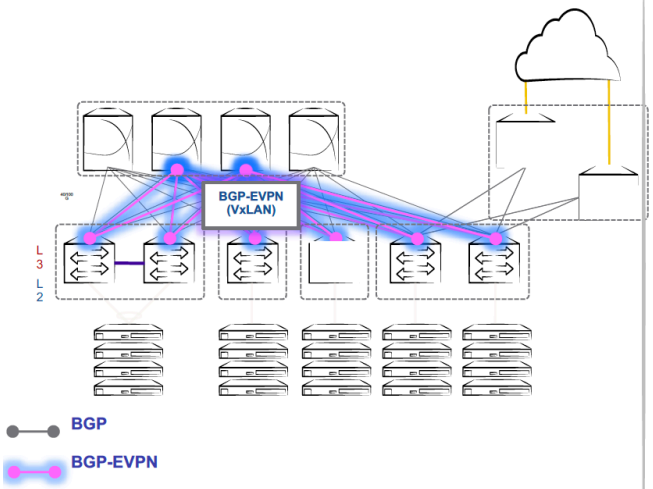


Feature	Benefit
Everything done with IP and BGP!	Easy to use, easy to understand and easy to operate
Standards based fabric	Multi-vendor support at all layers Support to thousands of standard tools

Data Center Architecture for Digital Age

Virtualization

BGP-EVPN (VXLAN)



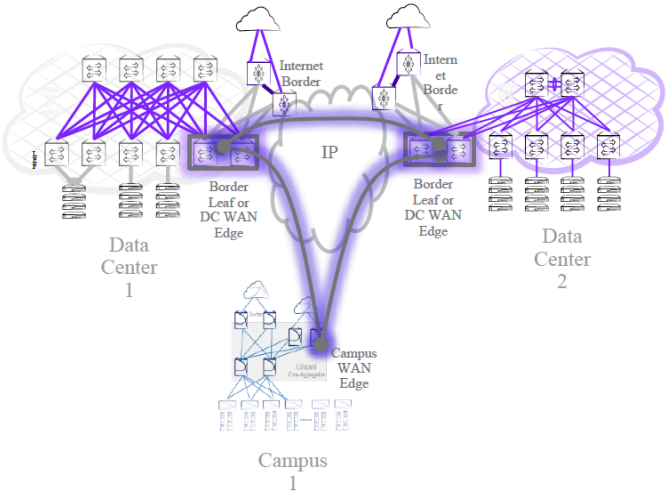
Feature	Benefit
Overlapping VLANs and IP Subnets	Ease of onboarding tenants
Layer 2 & 3 VNI support	From micro-segmentation to shared services; a simple efficient solution
Standards based BGP-EVPN	Massively scalable Distributed multi-tenant environments within and across data center
Distributed architecture	Eliminates controller costs and performance bottlenecks

Why an other VPN Technology?

- MPLS/VPLS and PBB are both proven technologies for Ethernet services, but
  - The control plane approach hasn't changed.
  - Still relies on flooding and learning to build the L2 forwarding database (FDB)
- EVPN introduces a new model for delivery of Ethernet services
  - Inherits a decade of VPLS operational experience in production networks
  - Incorporates flexibility for service delivery over L3 networks
  - Abstracts and separates the control and data planes
- Allows operators to meet emerging needs in their networks for Ethernet L2VPNs
  - Data Center Interconnect (DCI)
  - Cloud and virtualization services, and connectivity management
  - Multi-homing with all-active forwarding
  - Integrated L2 and L3 VPN services
  - Optimizing the delivery of Multi-destination frames (BUM)
  - Easier provisioning of services
  - L3VPN-like operation for scalability and control
  - Delivering L2 and L3 services over the same interface
  - Overlay technology that simplify topologies, and remove protocols from the network
  - VPLS and L3VPNs are proven technology but cannot meet all of these requirements
  - EVPN supports integrated routing and bridging VPN solutions with MAC/IP mobility over the same VLAN
  - Multiple data plane encapsulation choices

IXP Design

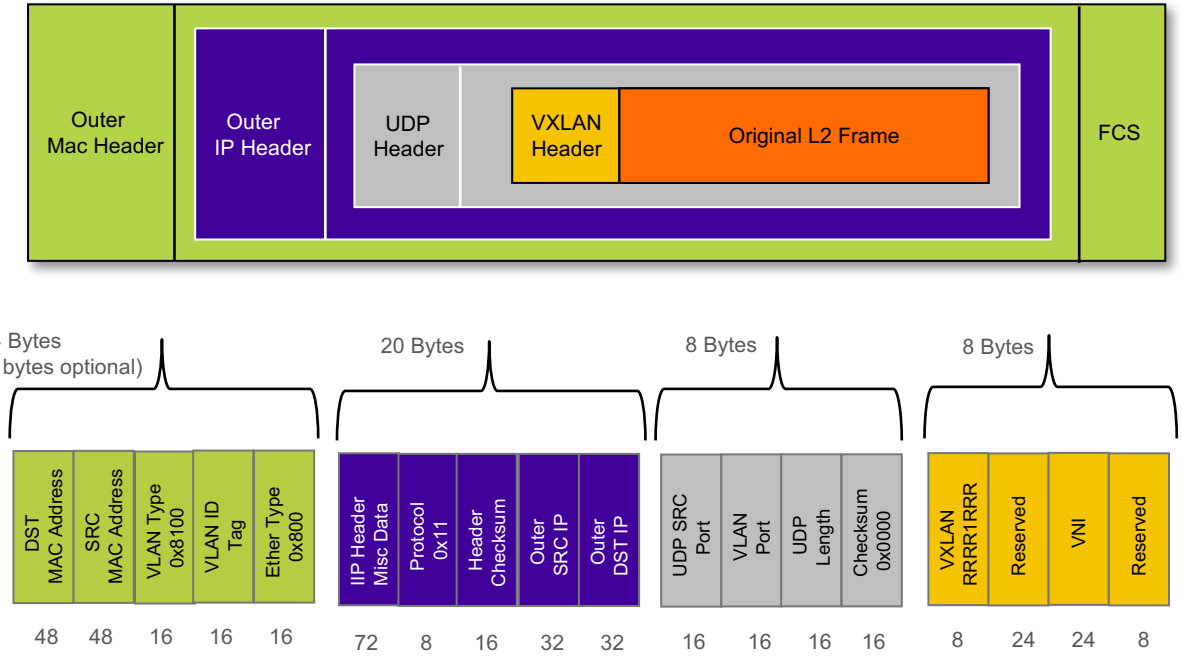
BGP EVPN-VXLAN



Feature	Benefit
Standards based BGP-EVPN	Interoperable across multiple vendor
Seamless scale for multiple sites	Ideal solution for distributed data centers and IXPs
Overlapping VLANs and IP Subnets	Ease of onboarding tenants and ISPs
Layer 2 & 3 VNI support	Single service for L2 & L3 reachability lowering deployment cost

VXLAN Overview

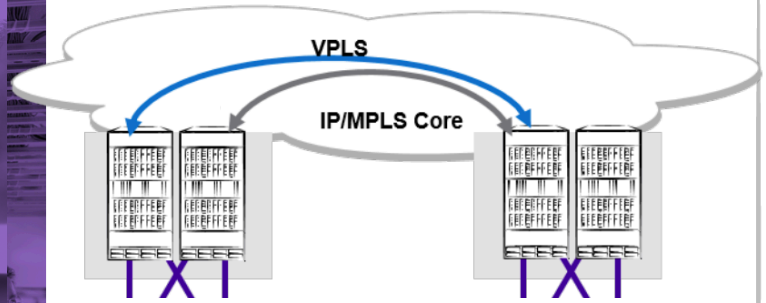
- Although the network overlay concept is not new, network overlays have gained interest as they address some of the scaling challenges. They have also gained interest with the introduction of new encapsulation frame formats purpose-built for the Data Center, like Virtual Extensible LAN (VXLAN).



MPLS / VPLS

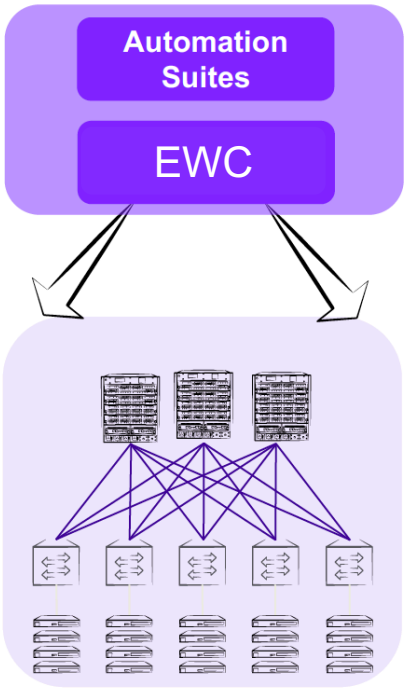
- Switch traffic based on a MPLS label
  - Originally for “faster lookup” for IP switching
  - Now mostly for creating a tunnel through an IP network
- Labels distributed via new/extended protocols
  - Label Distribution Protocol (LDP), BGP
  - Traffic Engineering extensions for RSVP (RSVP-TE
- With Layer 2 VPN’s, VPLS allows carriers to provide Ethernet (Transparent LAN Service – TLS) services
  - Using a very robust, scalable, Layer 3 infrastructure
  - No need for Service Providers to participate in customer IP addressing plans
  - Overcomes scaling issues, such as learning customer MACs on P nodes, scales beyond 4095 VLAN tags
- Either “Full Mesh” or “Hub-and-Spoke” topologies can also be supported
- Backup RSVP-TE LSP’s can be used in conjunction with L2 VPN’s to ensure fast failover

IXP Design
MPLS / VPLS



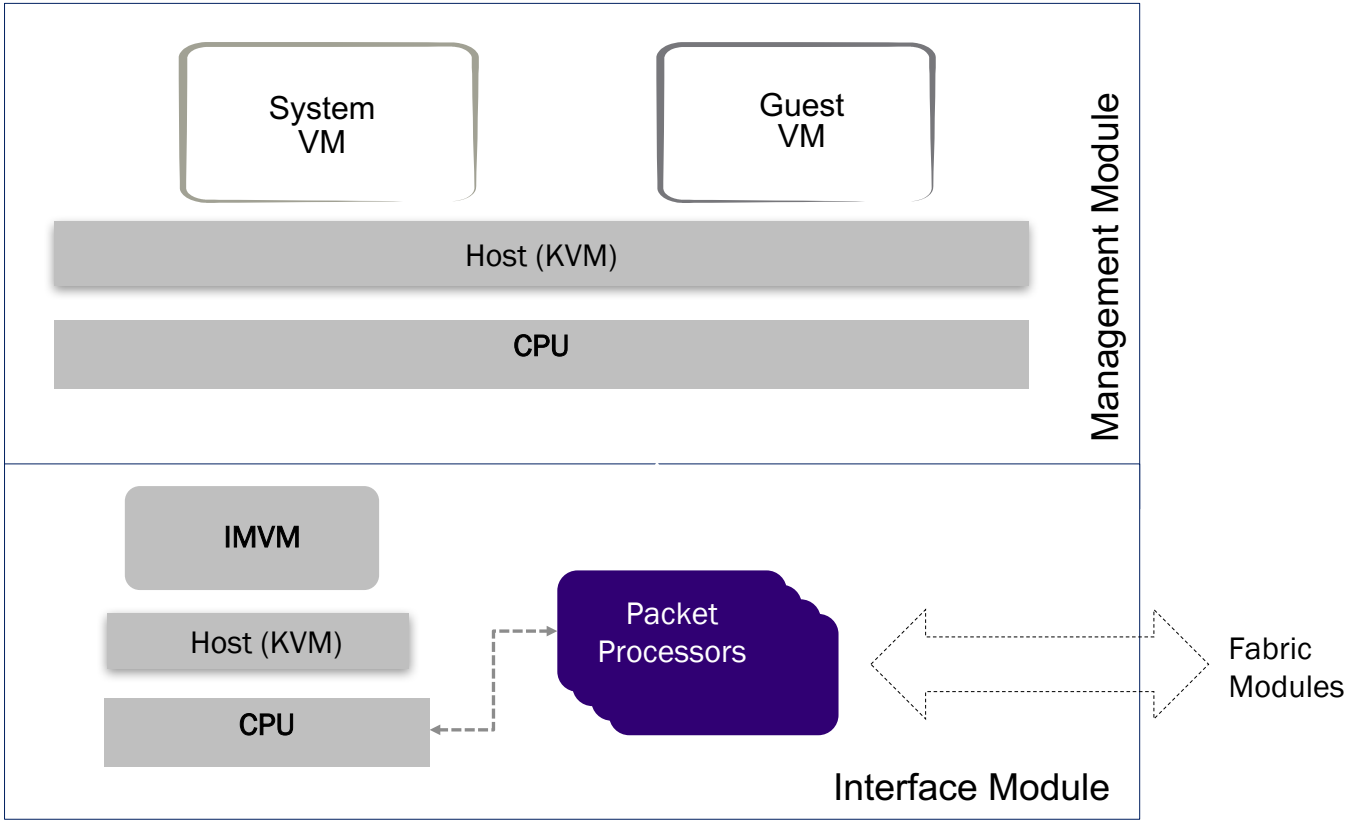
Feature	Benefit
Redundant VPLS psuedowires with Fast Re-Route (FRR)	Enables seamless disaster recovery
MPLS TE	Optimizes traffic flow between PEs
Advanced L2 resiliency and load balancing	Highly available and performance DCI for mission critical applications

Programmability



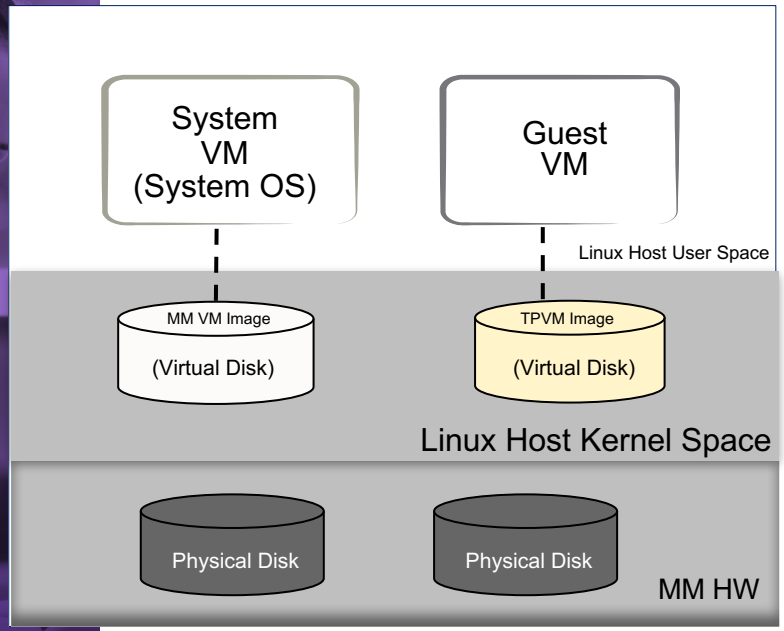
Feature	Benefit
APIs (netconf, rest, restconf)	Model based configuration
Streaming and pub/sub message buses	Ease of integration into third party tools and applications
Workflow Integration with dedicated automation suite	Flexibility of DevOps automation, ease of embedded automation
Guest VM on switch*	Run readily available tools and containers directly on the switch

SLX-OS Architecture



Birds-Eye View

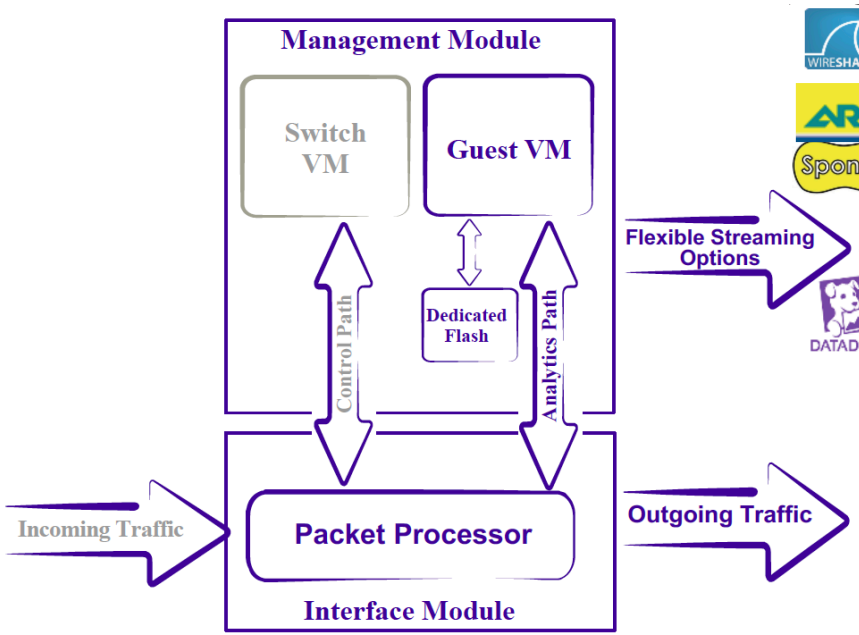
SLX-OS on MM



- Two VMs are deployed on MM – System VM to run SLX-OS ‘System OS’ software and Guest VM for third-party application hosting.
- The two VMs sit on the user space of the MM
- “MMVM Image” and “TPVM Image” are the virtual disks for the MM and Guest VMs.
- Virtual disks are files in the host Linux file system.
- Virtual disks appear as real hard disk to the VMs.

Extreme Insight Architecture

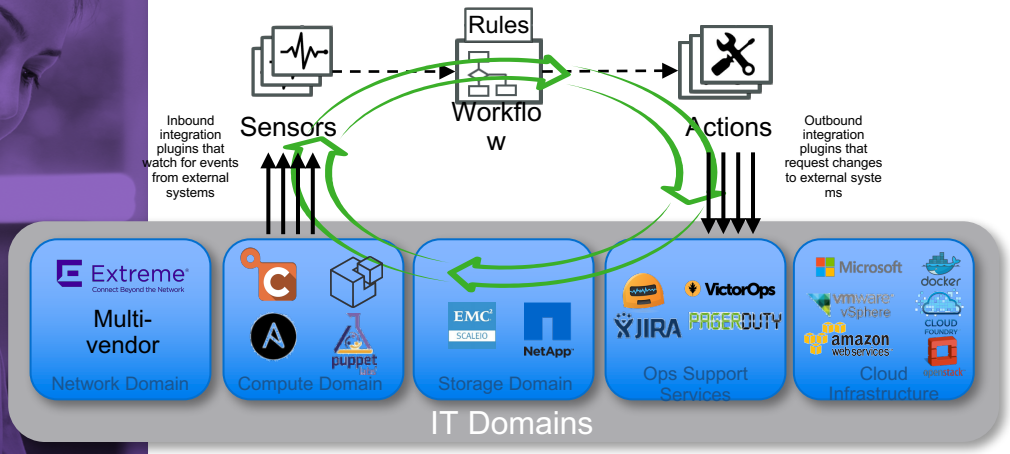
Enhancing Operational Efficiency



- **Guest VM** runs in KVM environment on every router enabling visibility for network-wide insight
- **Dedicated 10GE analytics path** ensures high performance without disruption to network traffic
- **Flexible streaming** options for in-band and out-of-band delivery of captured data to analytics applications
- **Dedicated analytics storage** provides real-time capture for easy and fast access

Extreme Workflow Composer

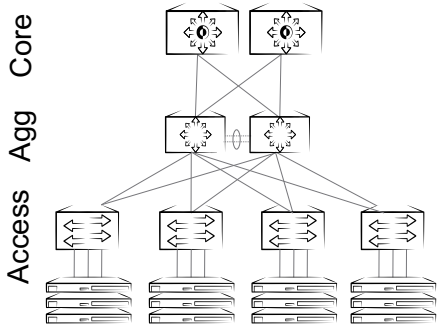
Powered By StackStorm



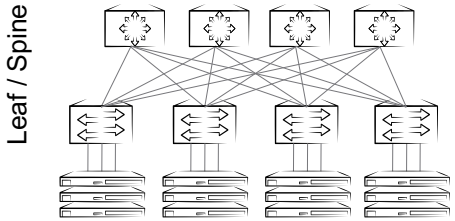
Extreme Workflow Composer
(Simple View)

- **Workflow-centric**
- **Integrations** into cross domain tool chains enabled through Actions and Sensors
- **Open** at all Layers
- **Extensible** through micro-services architecture for rapid innovation and development
- Enables **event-driven**, closed-loop automation
- Use of **community preferred** technologies

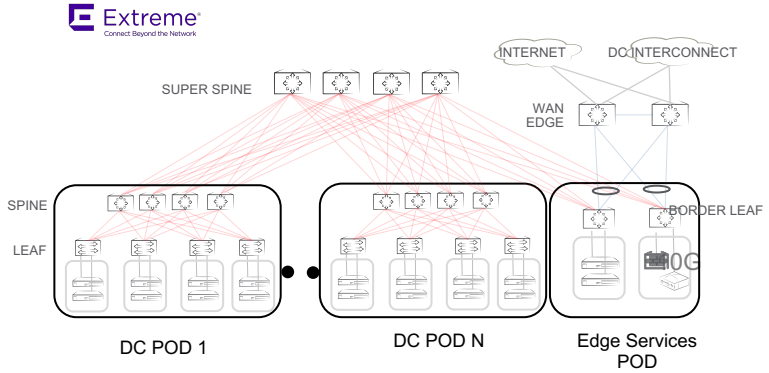
Evolution of Network Architectures



Phase 1:
3-tier Architectures



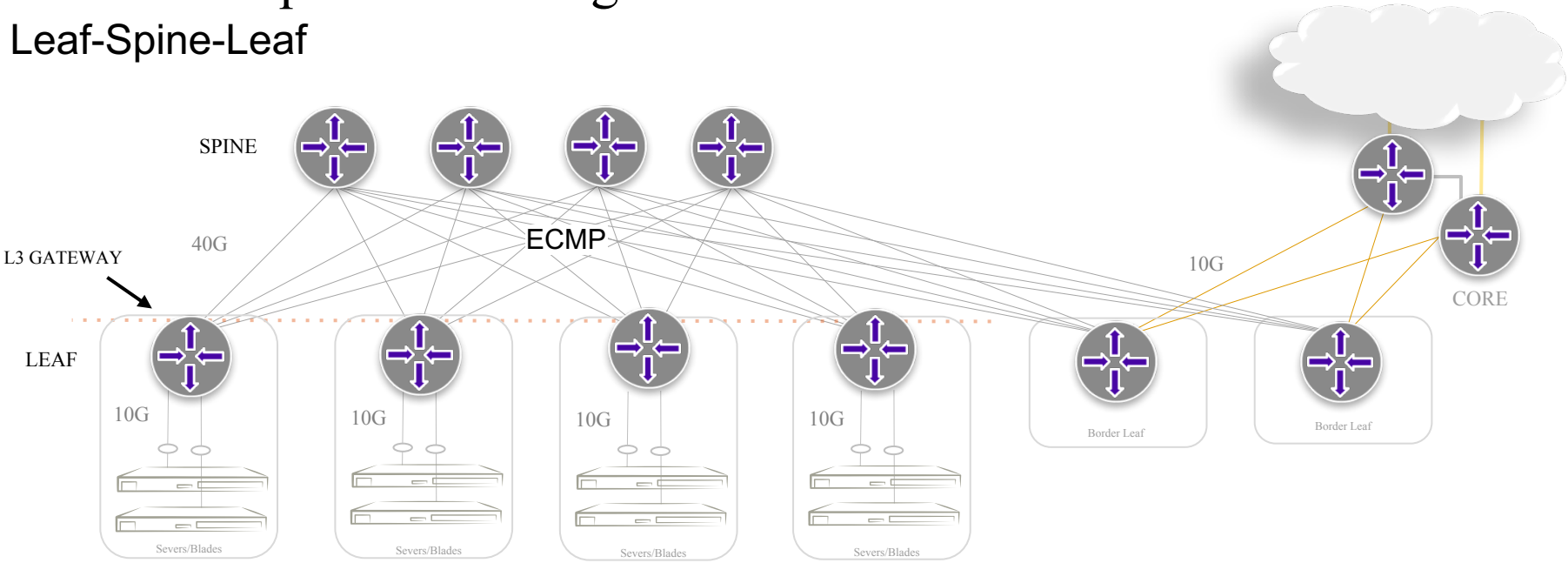
Phase 2:
Scale-out Architectures



Phase 3:
Advanced Scale-out Architectures

IP Fabric: Optimized 3 Stage Folded Clos

Leaf-Spine-Leaf



Rack
1 TOR per Rack
40 Servers/TOR
10G TOR to Server
LAG:N x 10G

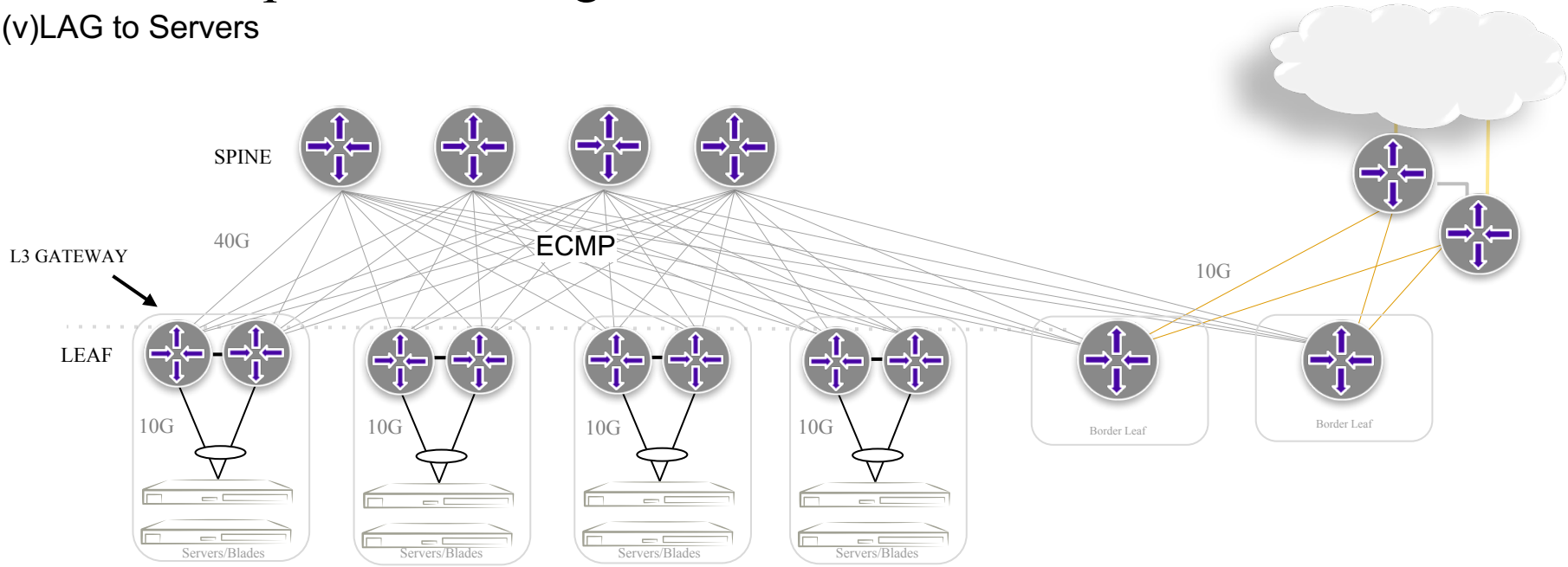
TOR
4 x 40G Leaf to Spine
# of Spine α 40G density
Oversubscription 3:1

SPINE
Scale α 40G port density
Non Blocking
Scale 36 → 216 Racks
Not interconnections between Spine Sw.

Edge
Border Leaf to Core
Oversubscription based on WAN Bandwidth

IP Fabric: Optimized 3 Stage Folded Clos

(v)LAG to Servers



Rack
2 TOR per Rack
10G TOR to Server
Dual Homed
vLAG to TOR (VCS Config)

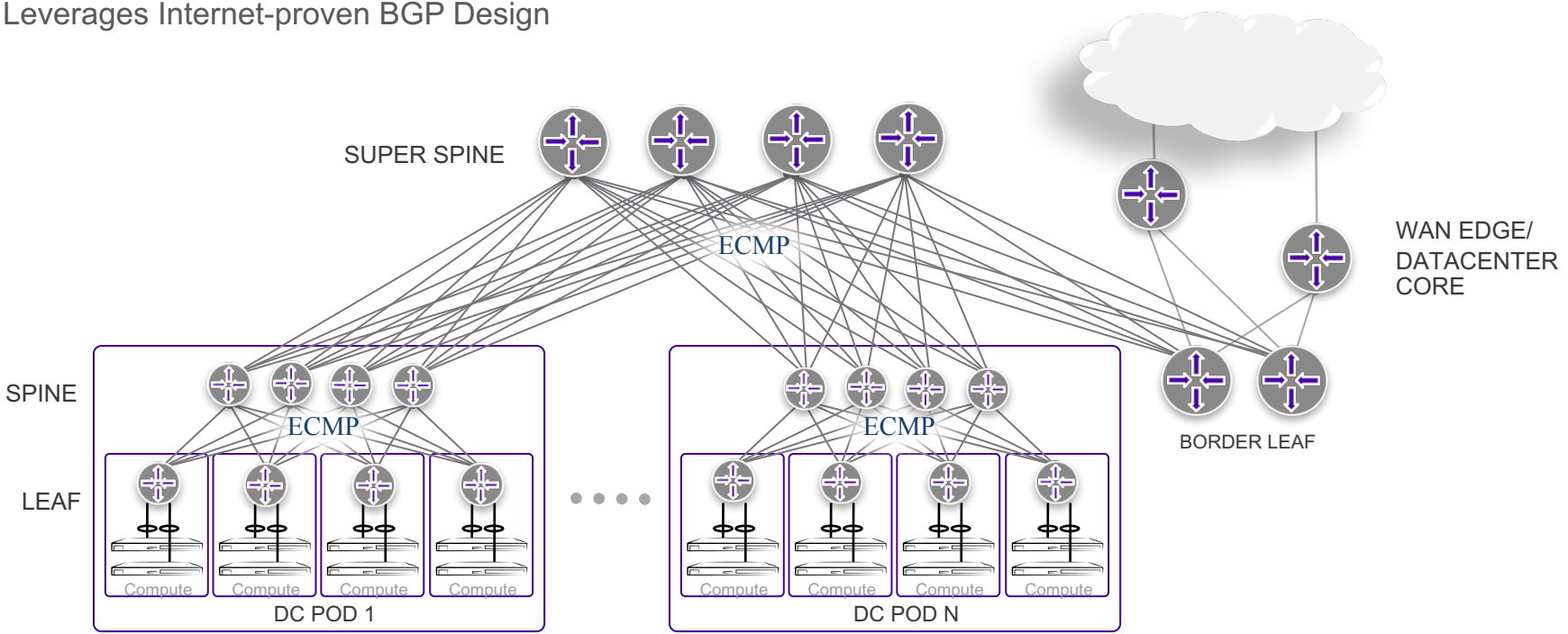
TOR
8 x 40G per Rack
Oversubscription 3:1

SPINE
Scale α 40G port density
Scale 18 → 108 Racks
Non Blocking
Not interconnected

Edge
Border Leaf to Core
Oversubscription based on WAN Bandwidth

IP Fabric: Optimized 5 Stage Folded Clos

Leverages Internet-proven BGP Design



Spine

40G from Spine to Super Spine

eBGP Peering across Spine and Super Spine

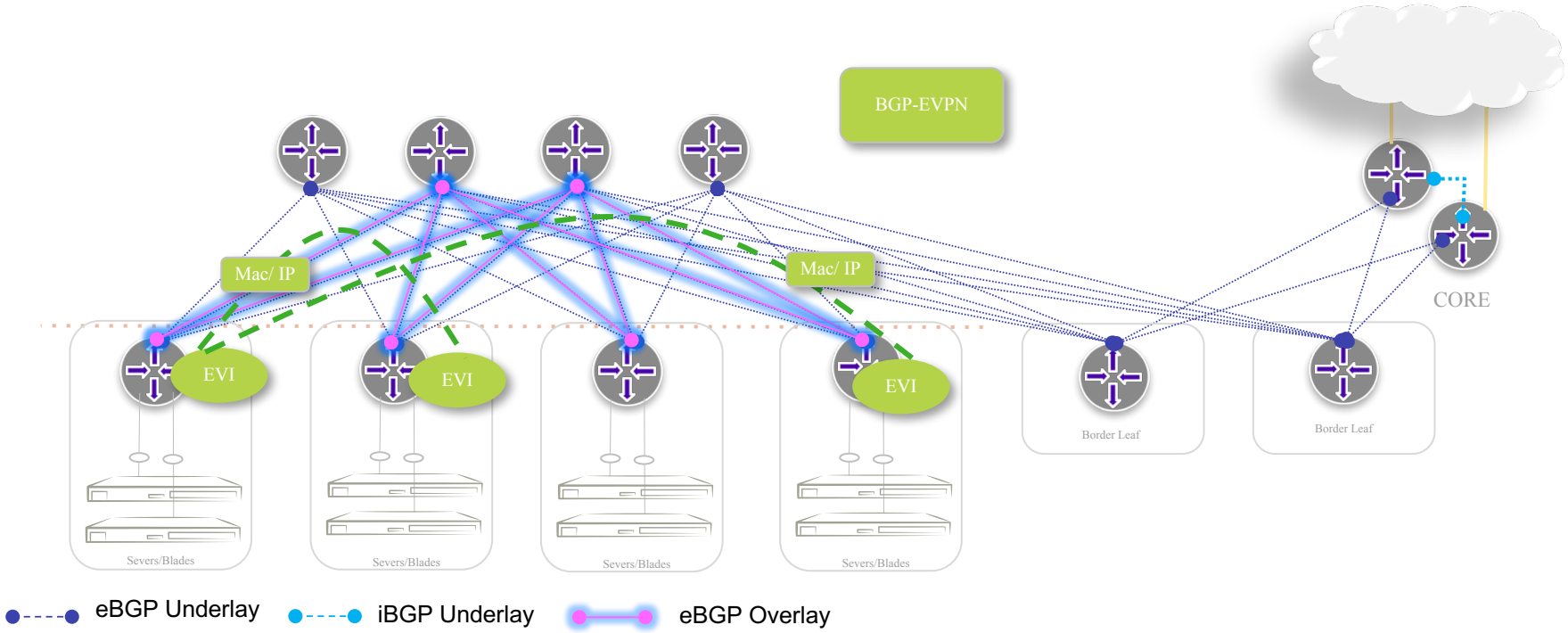
Super Spine

CORE directly connected to Super Spine

Reference: Facebook Data Center Architecture

Standards Based Network Virtualization

- Controller-less Overlay
- BGP-EVPN control plane and VXLAN data plane.



Thank You

WWW.EXTREMENETWORKS.COM