Hewlett Packard Enterprise

HPE Cloud-First reference architecture guide

Enabling the transformation to a hybrid infrastructure

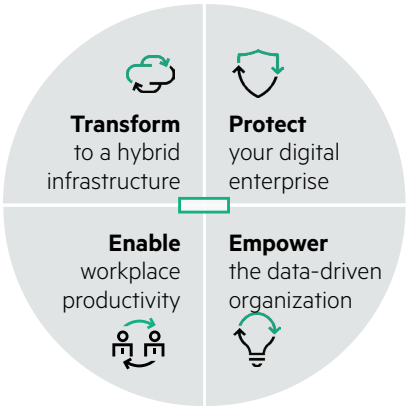
Contents

Architecture guide

Introduction

Transforming the enterprise for the Idea Economy

We are living in a digital world where everyone is connected, everywhere. It's also an Idea Economy, where the ability to turn an idea into a new product or service has never been easier. Anyone with an idea can now actually change the world.



It's an age of relentless, disruptive change for businesses and governments. Every Fortune 1000 company today is at risk of missing a market opportunity, not securing their enterprise, and being disrupted by a new idea or business model.

In the Idea Economy, no industry is immune to disruption. Whether in energy, healthcare, manufacturing or telecommunications, companies—start-ups or large enterprises—can only survive if they have both the vision and technological agility to respond to market opportunities and threats and quickly turn ideas into reality.

Today, an entrepreneur with a good idea has access to all of the infrastructure and resources that a traditional Fortune 1000 company would have…and they can pay for it all with a credit card. And yet the winners aren't always those with the best ideas. Rather, they are companies of every size that can execute on good ideas and deliver value faster and better than their competitors.

That means using the power of technology to quickly fuel the power of ideas. In the Idea Economy, IT strategy and business strategy are increasingly inseparable. And so IT's role must evolve from providing technical services to generating business value.

Businesses must change along four axes in order to survive and thrive in the Idea Economy. They must transform to a hybrid infrastructure; protect the digital enterprise; empower the data driven organization; and enable workplace productivity.

- **Transform to a hybrid infrastructure:**

  Accelerate the delivery of apps and services to your enterprise with your right mix of traditional IT, private, and public cloud. This is about having the right infrastructure optimized for each of your applications—whether in your traditional data center or in a public, private or managed cloud. And, it all has to work together seamlessly. In order to create and deliver new value instantly and continuously, businesses need infrastructure that can be composed and re-composed to meet shifting demands, infrastructure that will allow you to pivot when the inevitable disruption arrives. But, it isn't just in your data center. It isn't just in the cloud. Your infrastructure has to be everywhere, at the right cost, at the right performance with the right management, at the right scale. A hybrid infrastructure—one that combines public cloud, private cloud and traditional IT—can maximize performance allowing for continuous delivery, improved efficiency and optimized costs.

- **Protect the digital enterprise:**

  Protect your most prized digital assets whether they are on premise, in the cloud or in between. This is about security and risk management in the digital world. IT security used to be about defending the perimeter from external threats. The transformation of enterprise IT has created a matrix of widely distributed interactions between people, applications and data—on and off premise, on mobile devices and in the cloud. Security threats can be external or internal in nature and can represent malicious or unintentional actions. Enterprises lack the skills, resource and expertise required to proactively manage this threat. In parallel, enterprises must adhere to complex regulatory, compliance and data protection issues—and ensure enterprise-wide resiliency & business continuity in the face of natural and cyber disasters. All of this requires new thinking and security strategies that enable new ways to do business.

- **Empower the data-driven organization:**

  Harness 100% of your relevant data to empower people with actionable insights that drive superior business outcomes. This is about how companies harness all relevant business, human and machine data to drive innovation, growth and competitive advantage. This means empowering stakeholders to make timely and targeted decisions based on actionable, data-derived insights. A data-driven organization strives to rapidly and iteratively discover the value of its data through an optimized data-centric infrastructure. This foundation understands and engages with customers by listening and interpreting patterns within your customer data, uncovers competitive advantages and new market opportunities, and uses data to streamline operations and enable a leaner and faster organization.

- **Enable workplace productivity:**

  Deliver experiences that empower employees and customers to achieve better outcomes. The workplace is now digital, with interactions and experiences delivered to employees and customers across a multiplicity of locations, time and devices. Enterprises must deliver rich digital and mobile experiences to customers, employees and partners in order to engage employees and improve customer experience. Users expect personal, contextual and secure experiences. As enterprises look to improve productivity and drive customer loyalty, they must deliver highly engaging employee experiences and better serve customers through seamless, personal, contextual experiences.

HPE Cloud-First data center networking solution

To power your new and legacy applications and workloads, different IT environments are required based on your unique industry and application needs. Organizations need the right mix of traditional infrastructure, private cloud and public cloud with consistent management and control software to meet the demands for agility and excellent user experience.

Hewlett Packard Enterprise uniquely combines IT expertise with innovations in infrastructure, software and services to create a Cloud-First data center that is built to solve the issues that matter when transforming your business.

- **Speed:**

  HPE is the one stop shop that has the IT expertise and solutions that allow your organization to quickly execute on new opportunities better than your competition

- **Infrastructure:**

  As the leader in all major infrastructure categories including private cloud, servers, storage and networking[1], HPE is uniquely positioned to help your organization find the right combination of infrastructure to power the diversity of applications and business requirements

- **Open architecture:**

  HPE provides open architecture flexibility which can avoid vendor lock in and enable the right applications for the business

- **Ecosystem matters:**

  HPE brings the right internal expertise and a broad set of external partners to deliver best of breed infrastructure, services and solutions

- **Flexibility matters:**

  HPE can design and execute a roadmap that moves your business to your desired state in a step-by-step manner, transforming your business in the right sequence at the right time

Today's enterprise data center network architects and managers are expected to build networks that can concurrently consolidate and geographically distribute resources, which can include virtual and physical application servers, data storage, management platforms and network devices. These expectations have been fueled by the accelerating needs of businesses to be more agile, to do more with less and to increase their IT efficiency.

To reduce complexity, data center architects adopted designs that utilize network convergence, where data and storage I/O are merged onto a single network. This converged approach can eliminate physical clutter and complexity, while making more efficient use of networking resources. However, the simplification, in many cases, is only true at the surface level. Data center networks today can be more complex below the surface than ever before.

Cloud solutions present businesses with flexible solutions, but they also present challenges as you need to decide what the right mix is for you. Each enterprise must define their right mix of hybrid infrastructure based on their needs. The right mix of infrastructure varies by industry, company size, type of services and many other factors. Defining the right mix can help increase efficiency and reduce cost.

[1] **Cloud**—IDC press release, 2015 Worldwide Cloud IT Infrastructure Market Grows; idc.com/getdoc.jsp?containerId=prUS25732215
Servers—IDC Quarter x86 Server Tracker, 2015**Storage**—IDC; Worldwide total disk storage systems market, 2015
Networking—IDC WLAN Vendors, 2014

Public clouds will play a role in most organizations' transformation to a hybrid infrastructure. Public clouds are sometimes suitable for short term dev/test needs or for cloud native applications that do not require storing sensitive data or ensuring compliance.

However, many applications will continue to be locked down in your data centers on traditional systems or in private clouds. The fastest growth will actually be in private clouds because most businesses want control and assurance over the core applications and data. Some applications will be moved to a Virtual Private Cloud or hosted on a dedicated and managed private cloud. And some will be delivered from the public cloud including SaaS applications.

The Cloud-First data center is able to provide the necessary data center infrastructure solutions that can support this new Hybrid infrastructure which allows you to move traditional applications to the cloud, and develop and deploy cloud native applications.

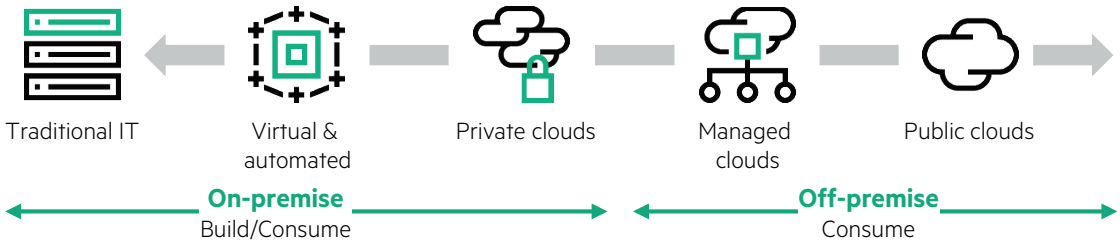


Figure 1. Balancing a hybrid infrastructure

The HPE Cloud-First data center can fulfill the following visions:

- **Cloud optimized:**

  The HPE Networking infrastructure is able to support the faster and flexible requirements of the cloud. It can enable definition of the necessary network resources, verification of resource availability, and aligns the resources with the application—allowing the network to deliver the expected SLA.

- **Software-defined networking (SDN):**

  Allows cloud providers and enterprises complete programmatic control of a dynamic network. Provides agility by orchestrating network services and automatically controlling the network according to high-level policies, rather than low-level network device configurations.

- **Strong technology partnerships:**

  HPE has a strong product and interoperability alliance with many vendors, including F5, Alcatel-Lucent, VMware® and Microsoft®. As your virtualization deployments continue to grow, these alliances will continue to add performance, flexibility and simplified management of VMs and their attached networks.

- **Enhanced L2 and L3 networking solutions:**

  HPE Networking solutions provide options for small-to-large scale deployments, which meet the demands of virtualization, convergence and massive east-west traffic patterns. Deployments can be optimized for VM migration by utilizing large-scale, flat L2 networks or by utilizing the advantages of L3 deployments in conjunction with L2 overlay technologies.

  All HPE Networking solutions are enhanced by standards-driven architecture and Intelligent Resilient Fabric (IRF) from HPE.

- **Intuitive, comprehensive management tools:**

  Orchestration is highlighted in the HPE network and server hardware in the data center through the use of HPE IMC. The HPE IMC solution is a next-generation management software, which provides the data center operations team with a comprehensive platform that integrates network technologies and provides full fault, configuration, accounting, performance and security (FCAPS) management functionality. With support for third-party devices, IMC enables network administrators to centrally manage all network elements with a variety of automated tasks including discovery, categorization, baseline configurations and software images. HPE Insight Management is a complete suite of server lifecycle management capabilities that can flexibly operate from embedded on-system utilities, your preferred CMS and now even from the cloud. Managing ProLiant servers with Insight Management delivers increased efficiency and precise control of your server infrastructure resources.

- **Proven servers:**

  Meeting the demands of virtualized and non-virtualized data center deployments. For example, HPE ProLiant BladeSystem and Moonshot System both meet the demands of the converged network infrastructure through enhancements in processing density, LAN and SAN connectivity with reduced cabling, and enhanced ability to function in flatter Layer 2 (L2) architectures to meet the demands of server virtualization.

- **High performance storage networking solutions with proven availability:**

  HPE storage solutions can meet any requirement that can be configured for budget-conscience customers or for customers that need the highest level of availability or are cloud ready.

- **Power and cooling:**

  HPE Data Center Smart Grid technology collects and communicates thousands of energy use measurements across IT systems.

HPE Cloud-First document series

This HPE Cloud-First reference architecture guide is complemented by other documents which, when referenced all together, can provide a clear view of today's data centers and HPE solutions within the data center:

- HPE Cloud-First reference architecture-Data Center trends: describes the trends and business technology drivers that are changing the shape of data centers.
- HPE Cloud-First reference architecture—100 server: provides reference architecture design examples for a 100-physical server data center.
- HPE Cloud-First reference architecture—500 server: provides reference architecture design examples for a 500-physical server data center.
- HPE Cloud-First reference architecture—2,000 server: provides reference architecture design examples for a 2,000-physical server data center.
- HPE Cloud-First reference architecture deployment guide: provides specific configuration examples, and best practices that should be followed in the data center.
- HPE Cloud-First reference architecture-building data center networks using HPE and Alcatel-Lucent: incorporates this dual vendor strategy specifically with Alcatel-Lucent.
- HPE Cloud-First reference architecture-building data center networks using HPE and F5: incorporates, this dual vendor strategy specifically with F5.

This guide is intended for technology decision-makers, solution architects and other experts tasked with improving data center networking. It can serve as a baseline guide when drafting data center architectures.

HPE Cloud-First reference architecture guide

Data centers contain the server compute, storage and networking devices needed to support businesses around the world. As such, it is critical that data centers are well designed and always available so that they are able to satisfy the demanding needs within each environment. Of course when designing a data center, networking professionals need to consider a myriad of factors including architectures (topologies + networking protocols and technologies), data center interconnections (DCI), port and table scalability, future growth, oversubscription tolerance, latency requirements, power, cooling and space constraints. A discussion about data center architectures should cover all of the above mentioned factors.

This HPE CFRA guide will simplify and provide clarification with regards to the architecture choices to consider. The data center architectures discussed in this guide can be used as a reference for data center IT professionals.

This section of the HPE CFRA guide will:

- Detail the approach to data center architectures which can be simplified as:
  - Layer 2 (L2) architectures
  - Layer 3 (L3) architectures
  - L3 with overlay solutions
  - Multiprotocol Label Switching (MPLS)/Virtual Private LAN Service (VPLS)

- List what needs to be considered when deciding on an architecture
- Provide information on the various technologies that are commonly used in modern data centers
- Provide details on HPE's approach to data center topologies

The remaining sections of this guide include:

- HPE data center interconnect solutions which can be used to connect geographically dispersed data centers
- Management of the HPE data center with HPE Intelligent Management Center (IMC)
- HPE data center networking portfolio overview which includes paths for more detailed information, including scalability, power consumption and cooling requirements

Data center architectures

Data center architectures can be simplified by breaking them down into the following four types of architectures:

L2 architecture

L2 networks were widely used in the early days of networking when networks were small. As networks grew they shifted towards L3 designs so they could scale and reduce broadcast domains. However, the rapid growth and deployment of virtualization saw the resurgence of L2 architectures, and the deployment of new large scale L2 architectures. Server virtualization solutions and cloud orchestration solutions such as HPE Helion OpenStack® are two examples of frameworks that leverage L2 and VLAN technologies for VM communications and migration. These deployments are designed so that VLANs can extend from rack to rack and even across data center boundaries, creating a large L2 environment optimized for the growing use of VM migration and disaster recovery.

L2 architecture recommendations

- Use in 1-tier or spine and leaf (2-tier) topologies:
  - 1-tier or spine and leaf topologies are not required, but recommended. If scaling to 3-tier topologies is needed, consider routing between aggregation and core layers, or consider L3 overlay architectures
- Standards based solutions allow for flexibility:
  - IEEE 802.1Q:
    - Extending necessary VLANs end to end
  - Transparent Interconnection of Lots of Links (TRILL):
    - TRILL-based designs are very suitable for large flat L2 architectures. TRILL combines L2 switching simplicity and flexibility with L3 stability and scalability. Scales to 4K VLANs.
  - Provider Bridging (QinQ), Provider Backbone Bridging (PBB, MAC-in-MAC) and Shortest Path Bridging (SPB):
    - These technologies have evolved over time to the current version of SPB. Originally developed for the Service Provider (SP) space, SPB is being used by customers to scale and provide multitenancy environments, allowing them to isolate tenants while also extending L2 networks over service provider networks and enterprise intranets. SPB scales to 16M VLANs.
- Leverage the benefits of HPE IRF:
  - Simplifies the network by creating large logical switches which can be managed as one device
  - Allows the network port density to easily scale
  - Provides device redundancy allowing a network to seamlessly recover from a device failure
  - Reduces the needs for legacy protocols like STP and VRRP while still ensuring that all links are active
  - Combine benefits of HPE IRF and standards-based L2 technologies:
    - TRILL and IRF solutions complement each other. TRILL provides for efficient multipathing resiliency, with efficient shortest path traffic flows. HPE IRF provides for port level scalability and hardware redundancy.
    - SPB and IRF solutions can also complement each other. SPB provides for multipathing support for up to 16M VLANs, multitenancy and optimal traffic flow, while HPE IRF provides for port level scalability and hardware redundancy

L3 architecture

L3 architectures that route packets at each device were widely deployed before the proliferation of virtualization. Virtualization drove the deployment of L2 architectures to satisfy the requirements of VM migration. However, within many environments L2 extension is still not a priority. These types of environments can benefit from the scaling and efficiency advantages that L3 architectures provide.

L3 architecture recommendations

- Use in spine and leaf or 3-tier topologies:
  - L3 can be deployed in 1-tier, spine and leaf (2-tier) or 3-tier topologies. However, L3 architectures are typically deployed in spine and leaf and 3-tier topologies due to the greater scale provided by those topologies. L3 is also widely used to connect data centers and service providers.
- Utilize standards based solutions which allow for easy interoperability preventing vendor lock-in:
  - Open Shortest Path First (OSPF):
    - A link-state interior gateway routing protocol (IGP), OSPF is a protocol which has knowledge of the complete topology, allowing it to be able to provide for good traffic engineering so that routes can be manipulated based on requirements
    - Allows for good scaling but as networks get larger, the size and frequency of topology updates can disrupt stability and delay route calculation while topologies converge. However, OSPF offers summarization and area isolation techniques that can help to overcome these types of issues.
    - Equal Cost Multipath (ECMP) provides deterministic load balancing across multiple paths
  - Border Gateway Protocol (BGP):
    - Widely associated as an exterior routing protocol that runs the internet, BGP can also be used as an interior protocol or both
    - Designed to exchange reachability information between separate autonomous systems. BGP is able to make routing decisions based on network policies and paths available.
    - Proven to scale very large, BGP sends updates only when a topology change has occurred and only the affected part of the table is sent
    - BGP devices contain two sets of routing tables. One is for Internal BGP (iBGP) routes within the same autonomous system, and the other table is reserved for routes between autonomous systems (eBGP)
- Continue to leverage benefits of HPE IRF:
  - HPE IRF is fully supported even in L3 architectures
  - Simplifies the network by creating large logical switches which can be managed as one device
  - Allows the network port density to easily scale
  - Provides device redundancy allowing a network to seamlessly recover from a device failure
  - Works when using any of the standards based L3 networking protocols

Overlay architecture

With the development of overlay technologies, enterprises can now enjoy the benefits of L3 architectures while also providing L2 connectivity for applications within the DC. Overlay solutions can also provide support for multitenant solutions which require client traffic separation.

Overlay networking is a solution that creates a virtual L2 network, which encapsulates and forwards traffic over the physical infrastructure. This virtual network allows VMs to remain isolated from other segments, but to also seamlessly move around the network even though the underlay network is running at L3.

An overlay network solution based on Virtual Extensible LAN (VXLAN) allows virtual networks to be easily created, enables multitenancy, scales beyond 4K VLANs and allows networks to exceed the ARP/MAC table capabilities of the underlay network.

Overlay architecture recommendations

- Use in spine and leaf or 3-tier topologies:
  - Overlays with L3 underlays will be deployed most often in spine and leaf or 3-tier topologies
- Use a L3 underlay as the foundation:
  - OSPF/BGP/ISIS etc.
  - Leverage ECMP for load balancing across multiple paths
- Enable L2 extension by using the following overlay solutions:
  - VXLAN:
    - Non controller based, static VXLAN solutions will benefit environments where the networking and server IT departments are firmly isolated from each other. Networking teams can deploy VXLAN tunnels across the underlay without engaging the server team.
    - Standards-based MP-BGP EVPN can now be used as a scalable VXLAN control plane protocol to solve the flood and learn problem associated with static VXLAN tunnels. Admins can use MP-BGP EVPN to discover VTEPs/end-host information and build VXLAN tunnels dynamically.
    - HPE Helion OpenStack, together with open source Neutron framework, enabled by HPE's Virtual Cloud Networking (VCN) SDN application will allow virtual overlay networks to be created between multiple hypervisors such as KVM and ESXi. VXLAN tunnels are created dynamically, and terminated within Open vSwitch instances, or by using either direct OVSDB or SDN-enabled termination on HPE switches such as the HPE FlexFabric 12900E/7900/5940/5930.
    - VMware direct OVSDB method allows HPE hardware VXLAN Tunnel End Points (VTEPs) to integrate directly with VMware NSX to bridge VMs on virtual networks to physical bare metal servers/physical firewalls/WAN routers etc.
    - HPE Distributed Cloud Networking (DCN) solution enables large enterprises and service providers with the ability to build distributed, scale-out, multi-cloud environment in a simple, standard and agile method using SDN and networking virtualization. L2 and L3 services can be delivered through distributed virtual switching using VXLAN tunneling and a BGP EVPN control plane to extend the service across controllers.
- Continue to leverage benefits of HPE IRF:
  - HPE IRF is also supported in Overlay solutions with L3 underlays
  - Simplifies the network by creating large logical switches which can be managed as one device
  - Allows the network port density to easily scale
  - Provides device redundancy allowing a network to seamlessly recover from a device failure

    Works when using any of the standards based L3 networking protocols

Multiprotocol Label Switching (MPLS)/Virtual Private LAN Service (VPLS)

MPLS, a standardized protocol, can simplify and increase efficiency of a network, by using "label switching" technology to transmit data across a network infrastructure.

MPLS operates at a layer that is generally considered to exist between traditional definitions of L2 and L3, and thus is often referred to as a "L2.5" protocol. In an MPLS network, unlike a L3 network, packets are assigned MPLS labels. This essentially replaces traditional IP forwarding which required complicated address matching performed at each hop in the network. The MPLS label also describes how the packet should be handled within the network and assigns the packet to a Class of Service (CoS).

By adopting an MPLS design, the need for large routing tables can be avoided in the core backbone of enterprises of service provider networks.

MPLS supports the creation of point-to-point L2 Virtual Private Networks (VPNs) and VPLS supports point to multipoint L2VPN services over an IP backbone. Because it is point to multipoint, VPLS is able to simulate an Ethernet switch among multiple switches in an MPLS network, allowing for inter-L2 port forwarding decisions based on MAC address or the combination of MAC address + VLAN ID.

MPLS/VPLS architecture recommendations:

- Use in spine and leaf or 3-tier topologies:
  - This solution will usually be deployed over spine and leaf or 3-tier topologies
  - Seen mainly in carrier networks, but also in some enterprises
- Can still leverage the benefits of HPE IRF:
  - Simplifies the network by creating large logical switches which can be managed as one device
  - Allows the network port density to easily scale
  - Provides device redundancy allowing a network to seamlessly recover from a device failure
- Advantages:
  - Available on HPE FlexFabric products without the need for a license
  - Used to extend L2 between data centers (VPLS)
  - Traffic engineering (MPLS-TE) allows traffic to be sent over non-standard paths for better path optimization
  - MPLS L2 VPNs provide for point-to-point L2 VPN connections
  - VPLS provides for multipoint-to-multipoint L2 VPN connections
- Disadvantages:
  - Implementation is not ubiquitous among carriers
  - Can be challenging to manage with large numbers of MAC addresses
  - Adds increased complexity as new locations are added or removed as each switch must be "touched"

DC solution guide

When drafting a data center solution, there are a number of questions and possible solutions that should be investigated before choosing a technology and architecture to utilize.

The below figure and table simplify the decision process and they break down the overall solution into three recommended options. Of course not all data centers may fall fully in line with one of the provided options, but these options can be used as a starting point. HPE data centers are very flexible so that the solution can be modified in a way that meets the specific needs of the environment.

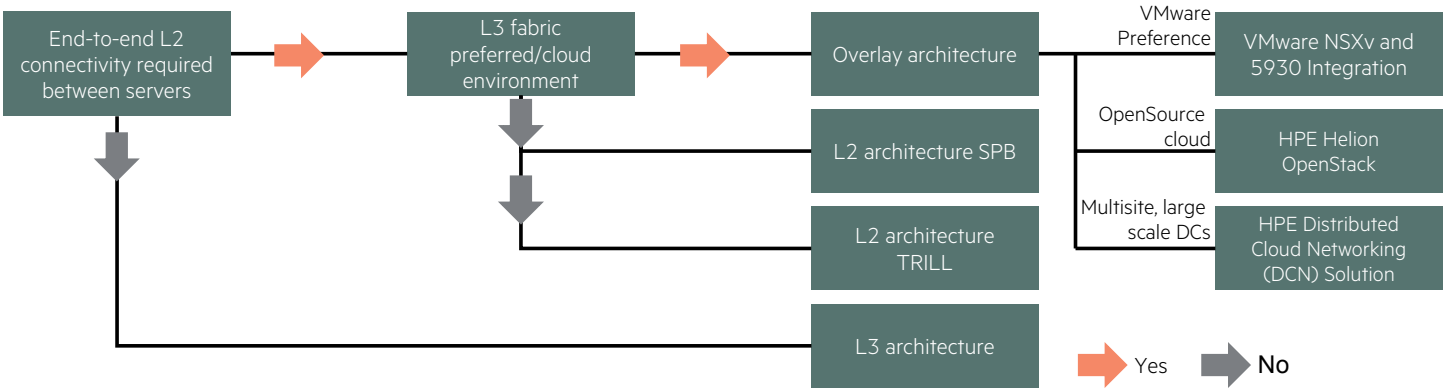


Figure 2. DC architecture guide

These DC solutions are summarized below as three cookie cutter options.

Table 1. Data Center architecture options

DC NETWORK DESIGN	OPTION 1 (RECOMMENDED)	OPTION 2	OPTION 3
Traditional 3-Tier	IRF	MSTP	PVST/PVST+
Spine & Leaf—Layer 2	IRF	MSTP	PBB/SBP (IS-IS)
Spine & Leaf—Layer 3	IRF & BGP (v4/v6)	OSPF (v4/v6)	IS-IS (v4/v6)
L3 overlay	EVPN with distributed L3 GW	EVPN with distributed L3 GW	EVPN with centralized L3 GW
LAN/SAN convergence	FC/FCoE	iSCSI	iSCSI
DC Interconnect	EVI (MACoverGRE)	EVI (MACoverGRE)	MPLS/VPLS

- Option #1:

  This option has solutions that are able to fit in many data center environments. This option leverages HPE IRF in the traditional 3-tier design, the Spine and Leaf L2 design, and the Spine and Leaf L3 design. HPE IRF is able to simplify the network while also providing high availability with device and path redundancy with ultra-fast failover support. The IRF based solutions can be deployed in pure L2 environments, eliminating the need for STP and VRRP. HPE IRF can also be deployed in pure L3 environments where it still is able to provide for a simpler network solution while still enjoying the scaling and traffic control benefits of L3.

  This option can also leverage any overlay required, however, EVPN is recommended thanks to its ability to provide distributed gateways, and dynamic and automatic configuration of the underlay and overlay.

  HPE EVI based MACoverGRE can be used to ensure the data centers have L2 connectivity.

  This option allows for a converged FCoE SAN that can take advantage of a single network with link speeds up to 10GbE, 40GbE and even 100GbE.

- Option #2:

  For those environments that IRF is not a preferred solution, the recommended data center architectures should be based on mature standards based technologies.

  Traditional 3-tier and Spine and Leaf L2 designs should utilize MSTP for L2 loop prevention. Spine and Leaf L3 designs should utilize OSPF as the dynamic routing protocol between each layer. The traditional 3-tier design should also use OSPF for L3 routing, whether performed at the aggregation layer or the core layer.

  This option can also leverage any overlay required, however, EVPN is recommended thanks to its ability to provide distributed gateways, and dynamic and automatic configuration of the underlay and overlay.

  HPE EVI based MACoverGRE can be used to ensure the data centers have L2 connectivity.

  If FCoE or FC is not preferred this option is able to provide the infrastructure needed to deploy a modern DCB based iSCSI solution. These iSCSI solutions are able to provide, in some cases, superior application performance while also leveraging the single network with link speeds up to 10GbE, 40GbE and even 100GbE.

- Option #3:

  This option highlights even more flexibility with regards to possible solutions.

  PVST/PVST+ may still be required in some mixed vendor solutions, and PBB/SPB is preferred by a portion of those admins which want a L2 loop free network.

  IS-IS is also an option for L3 solutions which do not want to leverage OSPF or BGP. This option also uses EVPN with a centralized gateway in L3 overlay solutions.

  Some environments may also still require legacy type MPLS/VPLS based DCI solutions.

  If FCoE or FC is not preferred this option is able to provide the infrastructure needed to deploy a modern DCB based iSCSI solution. These iSCSI solutions are able to provide, in some cases, superior application performance while also leveraging the single network with link speeds up to 10GbE, 40GbE and even 100GbE.

Questions to ask

Once, after careful deliberation, an architecture has been chosen—specific products that meet the needs of the deployment need to be selected. Below is a brief starting list of factors and questions that should be investigated when deciding which product to choose.

- Performance:
  - Can the networking devices and bandwidth between devices provide the performance needed?
  - Are large buffers required?
  - What is my server connectivity—1GbE, 10GbE or 40GbE—single or dual?
- Scalability:
  - What are the projected 5–10 year growth rates?
  - Can the networking devices scale to meet the port densities needed?
  - Can the chosen architecture scale?
  - Can the FIB/RIB/ACL/ARP/MAC tables scale to needed levels?
  - Does scalability tables support current and future VM requirements?
- Resiliency and redundancy:
  - Does the chosen architecture provide the resiliency and redundancy needed?
  - Server redundancy connections?
- Technologies:
  - What type of traffic?
  - VXLAN?
  - FCoE?
  - TRILL?
  - SPB?
  - HPE Multitenant Device Context (MDC)?
  - HPE Ethernet Virtual Interface (EVI)?
  - IRF? How many devices are supported?
  - Is cloud orchestration required?
  - Will overlay solutions be preferred (pure OpenStack, NSX or DCN)?
  - Should the solution use network, security or application-based SDN apps in their data center?
- If using a Blade Enclosure, which Interconnect Module is being used?
  - Can the Interconnect Module scale to meet the port densities needed?
  - Does the Interconnect Module FIB/RIB/ACL/ARP/MAC tables scale to the needed levels?

DC technologies

This section discusses the various technology solutions that are available for use in today's HPE data centers.

Intelligent Resilient Framework

IRF is a network virtualization technology from HPE that allows you to connect multiple devices through physical IRF ports to combine them into a single logical virtual device (IRF-fabric). From the point of view of an external switch, IRF-fabrics behave as a single switch in every aspect, such as a single Ethernet switch, single routing peer and single managed device (for example: single SNMP object instance).

Figure 3. HPE IRF

HPE IRF provides the following benefits:

- Simplified network configuration management and increased operational efficiency:

  Multiple linked physical devices look like one logical device and provide a single point of management. Only a single IP address and configuration file needs to be maintained.

- Scalable performance:

  IRF and Link Aggregation Control Protocol (LACP) used together can boost performance by bundling several parallel links between devices, allowing scalable "on-demand" performance and capacity to support critical business applications.

- IP address configuration is simplified:

  Each VLAN requires a single gateway IP address, eliminating the need to create identical configurations on all devices. Additionally, the resulting logical device is viewed as a single entity in the network management system, significantly simplifying network management.

- Redundancy protocol and loop prevention:

  Because multiple devices are virtualized into one logical device, loop prevention, reliability and redundancy protocols—such as VRRP, STP, RSTP and MSTP—can be eliminated. This simplifies network configuration and maintenance and eliminates design complexity, while enabling significantly decreased convergence times.

- Guaranteed system reliability:

  A device failure does not impact applications that rely on network state information. Additionally, LACP allows higher performance while eliminating single points of failure in the system. L2/L3 protocols do not need to re-converge when there's a link failure within a LAG group.

- Expanding bandwidth capacity:

  A virtualized system provides an effective load balancing mechanism between member devices, thus fully utilizing available bandwidth.

- Much more than just stacking:

  IRF provides all the traditional functions of a stacking technology and more. For example, IRF is a multiprotocol (L2, L3 IPv4, L3 IPv6, MPLS, VPLS, Unicast and Multicast) technology that removes the need for technologies like VRRP, providing not only the virtual IP gateway but also allowing full active/active L3 forwarding.

IRF operational fundamentals

Think of IRF as a framework on which Ethernet operates, rather than a protocol that operates on top of Ethernet. This is not to say that you cannot still make use of the benefits of SPB, VXLAN or TRILL in a data center bridged Ethernet network, but IRF does so much more than L2 Ethernet protocols. It was designed with much more in mind than to be just a replacement of STP.

IRF technology extends network control over multiple active switches. Management of a group of IRF-enabled switches is consolidated around a single management IP address, which vastly simplifies network configuration and operations. You can combine as many as nine HPE Comware OS based switches to create an ultra-resilient virtual switching fabric comprising hundreds or even thousands of 1GbE, 10GbE, 40GbE or 100GbE switch ports.

HPE IRF provides a simplified, higher performing, more resilient and flatter network design. IRF and HPE Comware OS based switches allow enterprise networks to be designed with fewer devices and fewer networking layers. It is a big improvement over the low performance, high cost, and crippling latency of conventional multitier legacy solutions, which often rely on a variety of different operating systems and complex resiliency protocols.

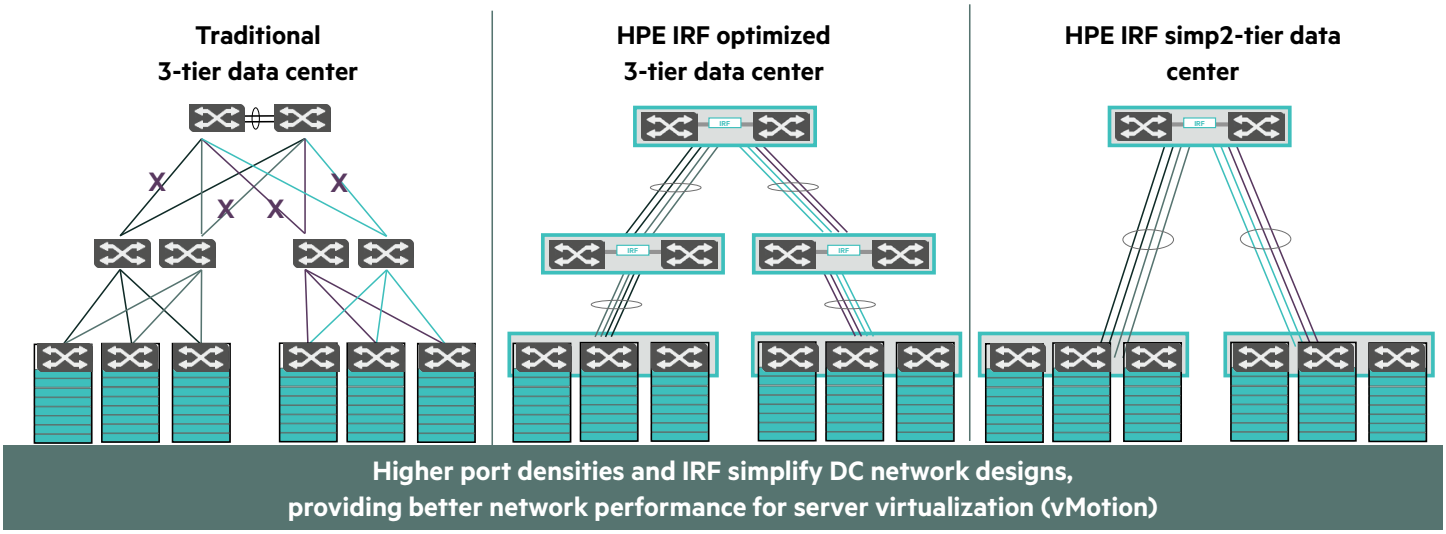


Figure 4. Flatter networks with IRF

HPE IRF solutions are providing flatter networks with fewer devices, managed in many cases as one device, but IRF doesn't restrict which technologies can be used when deploying networks. IRF is a complimentary technology allowing network admins to utilize IRF-based solutions while still deploying the network with end-to-end L2, L3, TRILL, VXLAN, multitenant, SDN, cloud or a combination.

For more information on HPE IRF, refer to HPE Intelligent Resilient Framework.

L2 or L3?

Many data centers use a combination of L2 and L3 solutions, but how much and at what layer each solution is deployed varies widely. Administrators should have a thorough understanding of why and when L2 or L3 should be used.

Virtualization

Server virtualization has been broadly deployed globally across data centers for many years. With it, the operating systems, applications and servers work in a non-dedicated or loosely coupled manner, interacting as needed to meet an enterprise's needs.

Virtualization provides several key benefits:

- Higher efficiency and lower CAPEX/OPEX:

  Allows for more efficient use of a server and its resources.

- Agility and flexibility:

  Provides the ability to do on-the-fly migration of virtual servers or applications across physical servers in the data center.

- Resiliency:

  Provides the ability to restack or shuffle applications in support of business continuity or even for routine maintenance.Widespread deployment of virtualization brought about the following architectural issues in data centers:

- East-west traffic patterns:

  Virtualization has changed the data center traffic flows from the old style client/server (north-south) to horizontal server/server (east-west) patterns.

- VM mobility:

  Virtualized data centers require large L2 domains so that VMs can move from rack to rack without having to change IP addresses.

- Higher bandwidth requirements in less space:

  With a virtualized environment, a complete enterprise information system infrastructure can be deployed in much less space, which means that the number of servers requiring bandwidth in a given rack has drastically increased.

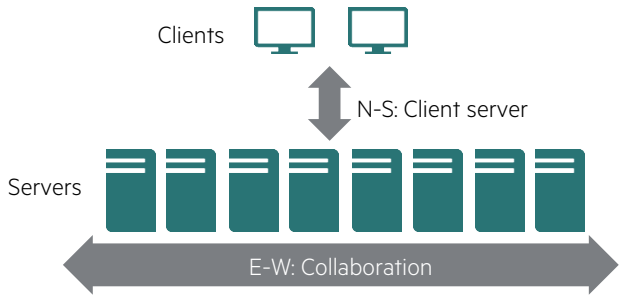


Figure 5. N-S vs E-W

These virtualization issues have helped to drive both flatter network architectures with fewer hops, and the adoption of large scale, high-performance data center access and core switches that now widely support 10GbE, 40GbE and 100GbE uplinks.

These flatter architectures were then deployed as large L2 domains, where typically L3 routing was only done in the core switches. This way, VMs within the virtualized environment could freely move whenever and wherever they need. Many enterprises have also worked to extend L2 deployments to other data centers so they can take advantage of long distance VM migration and disaster recovery protection.

These L2 environments have seen lots of success, however, not one technology solution fits all needs of all customers. As with everything, there are advantages to L2 and L3 solutions.

L2 advantages:

- Lower CAPEX:

  L2 switches typically offer lower upfront capital expenditure

- Simpler solution:

  For small-to-medium sized data centers L2 deployments can be simpler to deploy and manage with fewer staff.

- VM migration support:

  L2 deployments allow VMs to move as needed, between racks or even between data centers without having to change IP addresses. However, as discussed later in this document, newer overlay solutions such as VXLAN are now able to allow VM migration over L3 environments. Customers with large virtualized environments are not locked into L2 solutions now.

- Less latency:

  With L3 networks, the IP portion of the datagram has to be stripped off the frame and then reassembled. After reassembly, the hop count is reduced, the checksum is recalculated, a routing lookup is performed, and then the IP datagram is inserted back into the frame and transmitted to the next hop. Of course this can add latency, however, the performance of today's switches has increased to the point that this latency may only be an issues in certain types of environments.

L3 advantages:

- Reduced broadcast storms:

  L2 networks forward all traffic, especially ARP and broadcasts. Anything transmitted by one device is forwarded to all devices. When the network gets too large, the broadcast traffic can begin to create too much congestion thus decreasing network efficiency. L3 networks are able to avoid this problem by segmenting out the network into multiple smaller broadcast domains.

- More efficient forwarding:

  Suboptimal forwarding can happen in L2 environments when servers (or VMs) in different VLANs, which are connected to the same top-of-rack (ToR) switch have to travel to the core to get routed, and then hairpinned back down the same path so they can communicate. L3 solutions avoid this problem by making routing decisions closer to the server. Additionally, L3 solutions can efficiently distribute traffic across multiple paths using ECMP and per flow hashing, as opposed to per host hashing used at L2.

- Better large scale capabilities:

  Every VM in a data center is going to have its own MAC address, and when DCs grow large this can quickly exhaust the MAC address and VLAN limits on L2 networks, especially when not using DC hardware. L3 networks are better able to scale to large size more efficiently. In fact a properly configured L3 network could theoretically have infinite growth.

- More efficient use of network uplinks:

  Traditional L2 deployments used STP to resolve loops in a network. This inherently meant that there will be blocked unused links in the network. However, networks that utilize HPE IRF can eliminate the use of STP. With that said, L3 deployments don't need to use IRF to eliminate STP. L3 deployments are able to efficiently tolerate failures and utilize redundant links with technologies such as ECMP.

- VM migration support:

  With the recent adoption of overlay technologies like VXLAN, enterprises now can deploy L3 solutions from core to ToR. These L3 environments are able to leverage efficient L3 forwarding technologies to distribute traffic across the core spine switches, while still supporting vMotion migration in the overlay environment.

- Troubleshooting:

  Small-to-medium L2 solutions can be considered "simpler", but as these L2 environments grow they can get complex to troubleshoot. L3 networks have been proven to scale very large (i.e., the Internet), while still offering administrators with efficient troubleshooting capabilities.

Bottom line

Administrators of highly virtualized environments should consider L2/L3 deployments carefully. They will need to decide when and where they want VMs to migrate. An administrator could create a single large L2 environment, which works well for small-to-medium size data centers. However, as data centers grow it can be difficult to keep expanding this environment, especially as tables reach maximum scalability and broadcast domains get too large.

The larger a data center gets, L3 deployment starts to make more sense. However, to still support L2 migration these data center admins may also need to consider an overlay or POD (isolated clusters of L2 racks, which require routing between clusters) environment. This could mean the data center will require more investment for equipment and training of support staff.

Finally, administrators need to determine which advantage is important to them. For example, is having a more efficient forwarding mechanism more important than lower latency per hop? Is lower CAPEX more important than possible troubleshooting advantages L3 provides in larger environments?

HPE VXLAN overlay solutions

VXLAN is a network encapsulation overlay solution that solves the problems of L2 VM mobility, network infrastructure MAC address and VLAN scalability limitations, and supports multitenancy. VXLAN increases the speed and agility of deploying virtual networks by allowing a virtualization administrator to self-provision their required networks with no requirement to wait for the network admin. VXLAN decouples the VM from the underlying L3 network thereby allowing VMs to communicate across the same L2 segment without reconfiguring the underlying network.

VXLAN Tunnel Endpoint (VTEP)

VXLAN uses VTEP devices to map VMs or physical devices to Virtual Network Identifier (VNI) segments and to perform VXLAN encapsulation and de-encapsulation. VTEPs are available both as software vSwitches on hypervisors or as hardware switches such as the HPE FlexFabric 5930/5940, 7900 and 12900E Switch Series. Hardware VTEPs can be beneficial in scenarios where VMs need connectivity with physical devices such as servers, firewalls or routers.

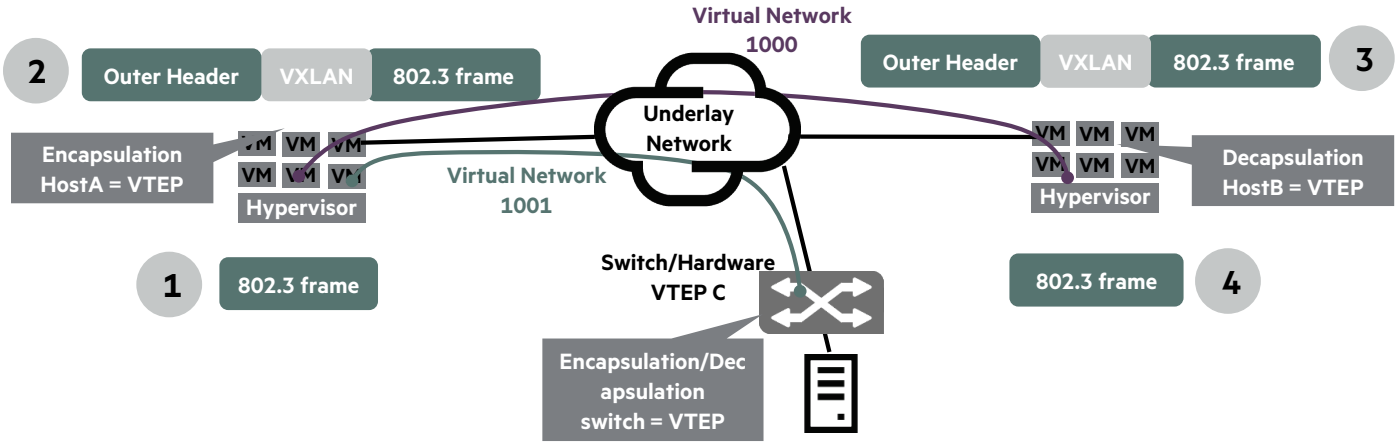


Figure 6. VTEPs and VXLAN tunnels with different VNIs across an underlay network

An example of the HPE VXLAN implementation would be a solution that utilizes the HPE FlexFabric 5930/7900/12900E Switch as a VTEP supporting different VNIs and deployed in strategic places within the DC to bridge traffic between VMs on hypervisors and physical devices such as servers/firewalls/routers.

L2 & L3 VXLAN gateways

Basic L2 VXLAN VTEPs are VXLAN capable networking devices, which provide connectivity for VMs and bare metal devices residing on the same L2 segment.

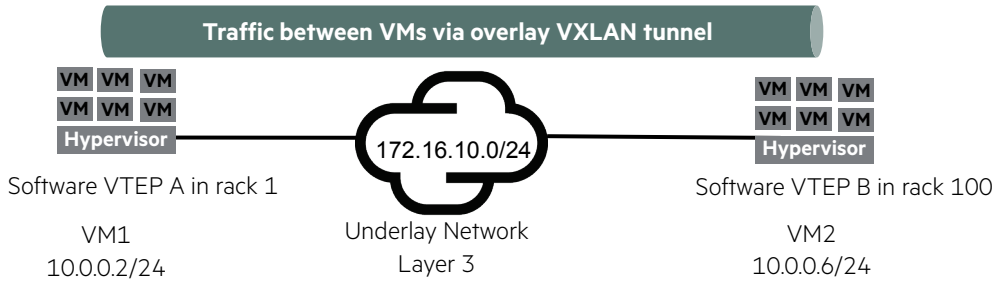


Figure 7. L2 VTEPs providing connectivity for VMs on same segment

L3 VXLAN IP gateways are networking devices which act as IP gateways to provide Layer 3 forwarding of VM/server traffic as it exits the VXLAN tunnel to a device on a different subnet. This example provides VM-to-VM connectivity for VMs on hypervisors which are not VTEP capable and reside on separate network subnets. L3 VXLAN gateways can also connect VMs to external networks by providing gateway functionality, which de-encapsulates VXLAN traffic and then routes the traffic outside of the underlay network onto a different L2 segment.

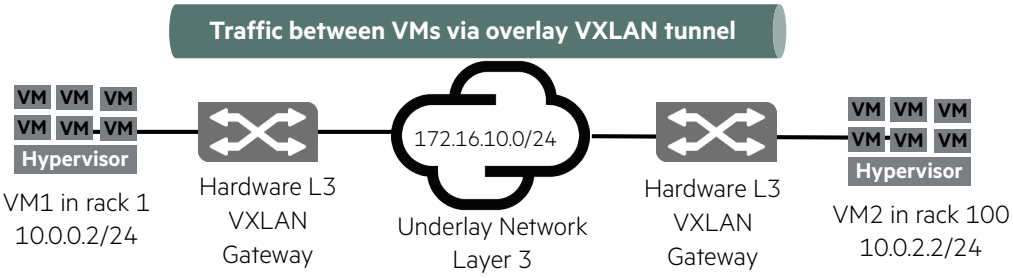


Figure 8. L3 VXLAN gateways providing connectivity for VMs on different segments (de-encapsulate, route and re-encapsulate)

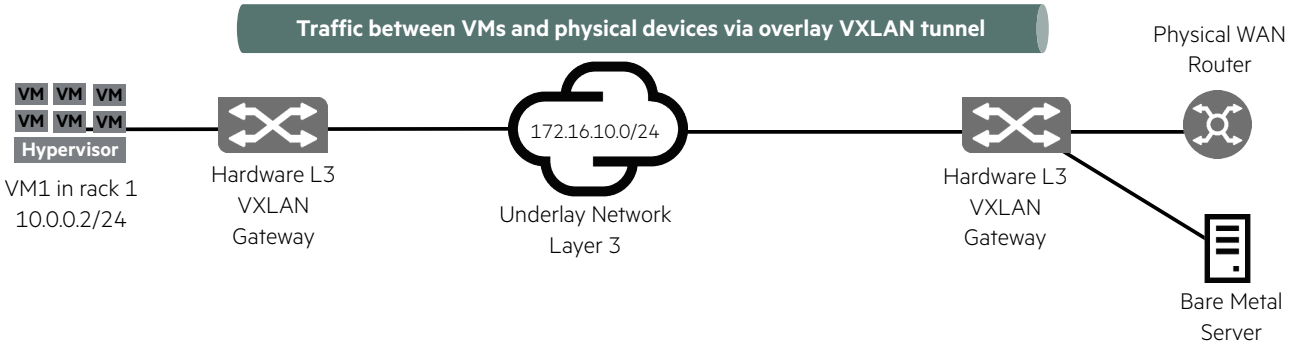


Figure 9. L3 gateway providing VM connectivity outside of underlay fabric (de-encapsulate, route)

Centralized L3 VXLAN gateway redundancy is achieved through VTEP grouping of the VXLAN L3 gateway devices. With VTEP grouping, all of the spine switches are configured as VXLAN L3 gateways and the leaf switches are configured as L2 VTEPs. The spine switches essentially provide L3 routing for VLAN-to-VLAN traffic within the VXLAN fabric or for routing outside the VXLAN fabric. All the spine switches can be configured to provide Layer 3 gateway redundancy through the VTEP group functionality. The leaf switches will establish VXLAN tunnels to all of the gateway switches in the VTEP group using one virtual IP address for the entire group and one local IP address for each gateway. Virtual tunnels are used for all normal VXLAN packet forwarding to all of the gateways, and the local tunnels are used for ARP request flooding for specific L3 gateways. When any spine switch fails, other spine switches learn the MAC addresses of the prevailing traffic through ARP using the local tunnels and they then provide L3 gateway redundancy.

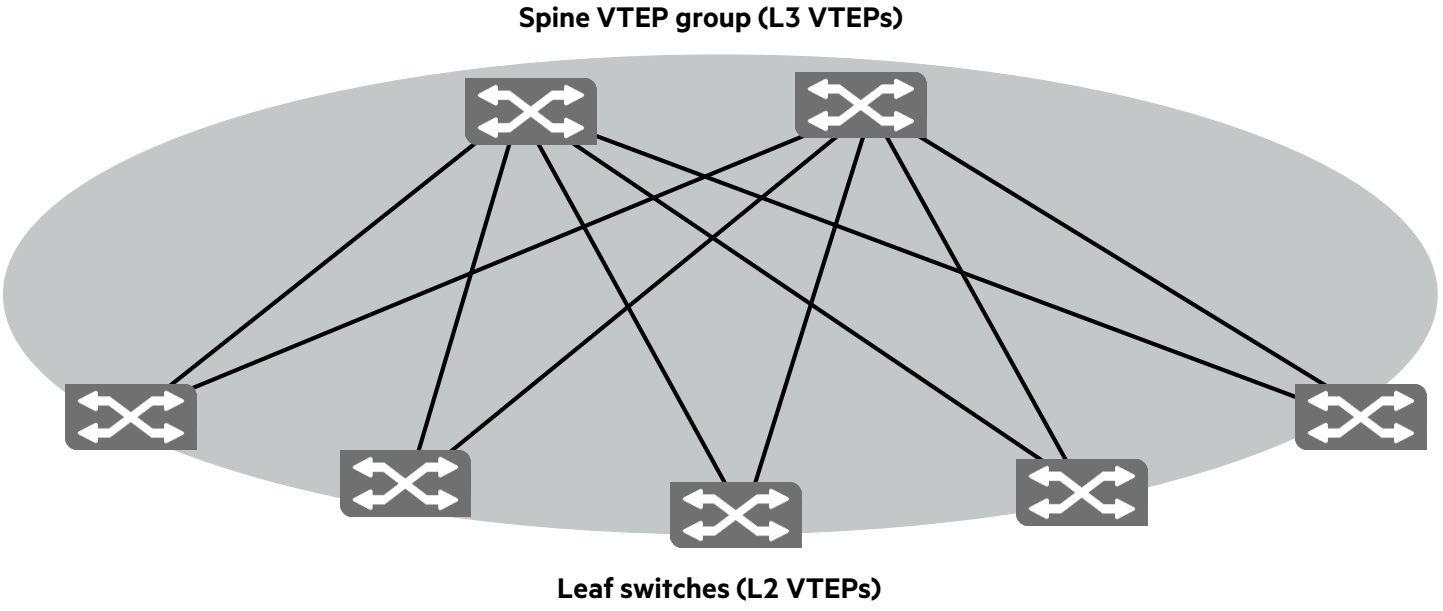


Figure 10. Centralized L3 gateway redundancy through the VTEP group functionality

Distributed L3 VXLAN gateways are another option available to provide Layer 3 forwarding capability at multiple devices for VM/server traffic beyond its own subnet. In the example below, the same default gateway is located on multiple devices to provide optimal forwarding of traffic for devices on the same switch but different subnets.

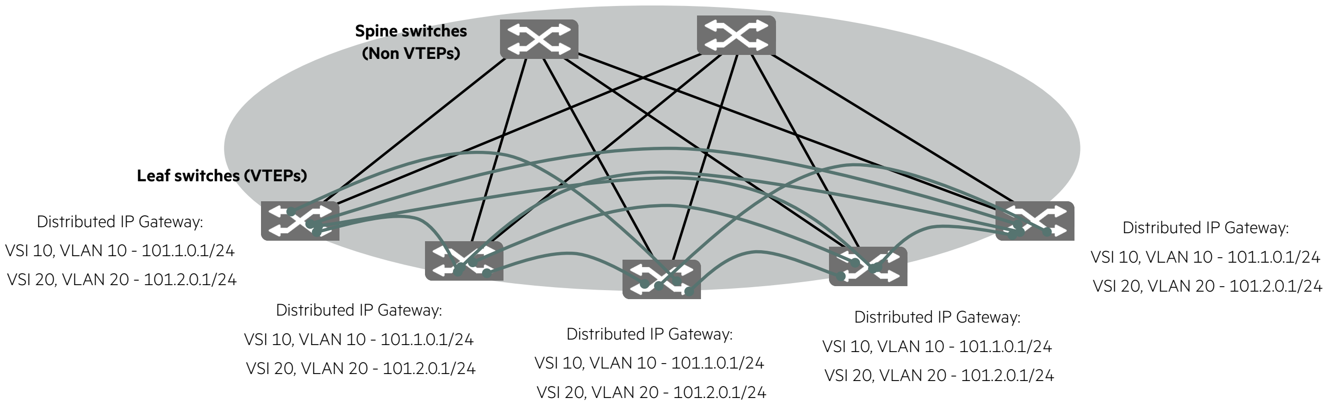


Figure 11. Distributed L3 gateways

EVPN (VXLAN at scale)

VXLAN tunnels can be manually configured without a control plane or automatically established via a control plane. Static manual configuration using the CLI is not recommended unless the scale is very small or it is used for test purposes. The recommended approach for larger scalable solutions with dynamic tunnel management is to use MP-BGP EVPN, which is able to dynamically discover the VTEPs and establish VXLAN tunnels.

MP-BGP EVPN offers the following benefits:

- The MP-BGP EVPN protocol is based on industry standards
- It enables control-plane learning of end-host Layer2 and Layer3 reachability information, enabling more robust and scalable VXLAN overlay networks
- It uses MP-BGP to support scalable multitenant VXLAN overlay networks
- It minimizes network flooding through protocol-based host MAC/IP route distribution

Multicast functionality with VXLAN fabric

To achieve the best performance for multicast applications in a VXLAN fabric, it is recommended that multicast traffic is sent into the Layer 3 underlay network directly.

This is illustrated in the topology below. IGMP PIM is configured normally on the underlay network. Multicast VLAN feature can be used to map multicast traffic to a dedicated VLAN on the Top of the Rack (ToR). With this configuration, unicast and broadcast traffic uses the overlay while multicast uses the underlay network providing efficient use of both the overlay and underlay.

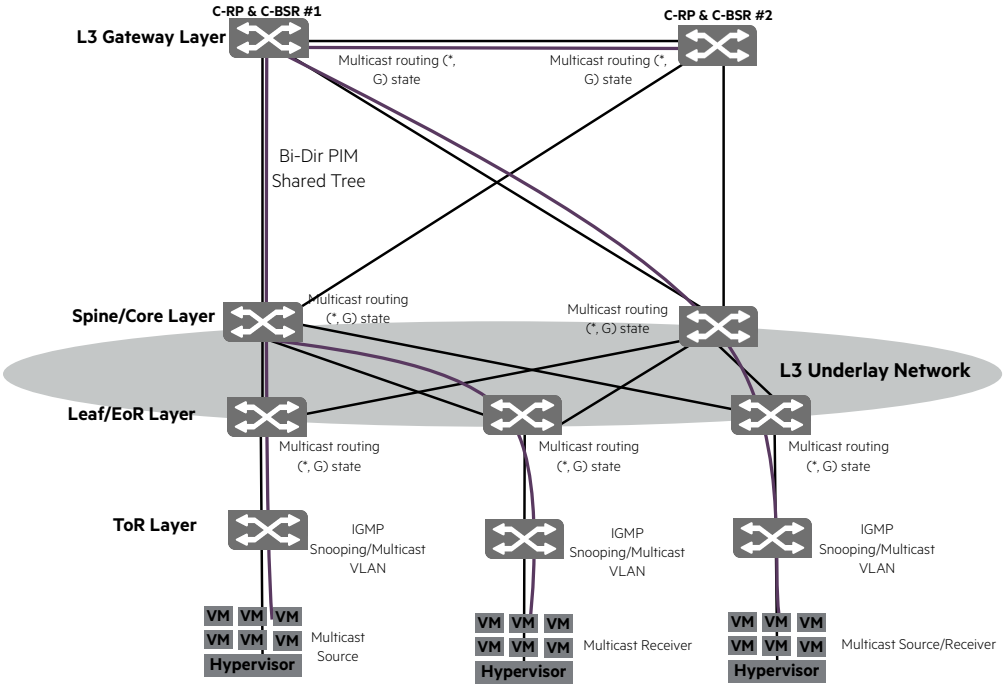


Figure 12. Multicast in VXLAN fabric

The HPE FlexFabric 5930/5940/7900/12900E switches support both VTEP and gateway functionality. VXLAN is currently utilized in three main HPE solutions as shown below.

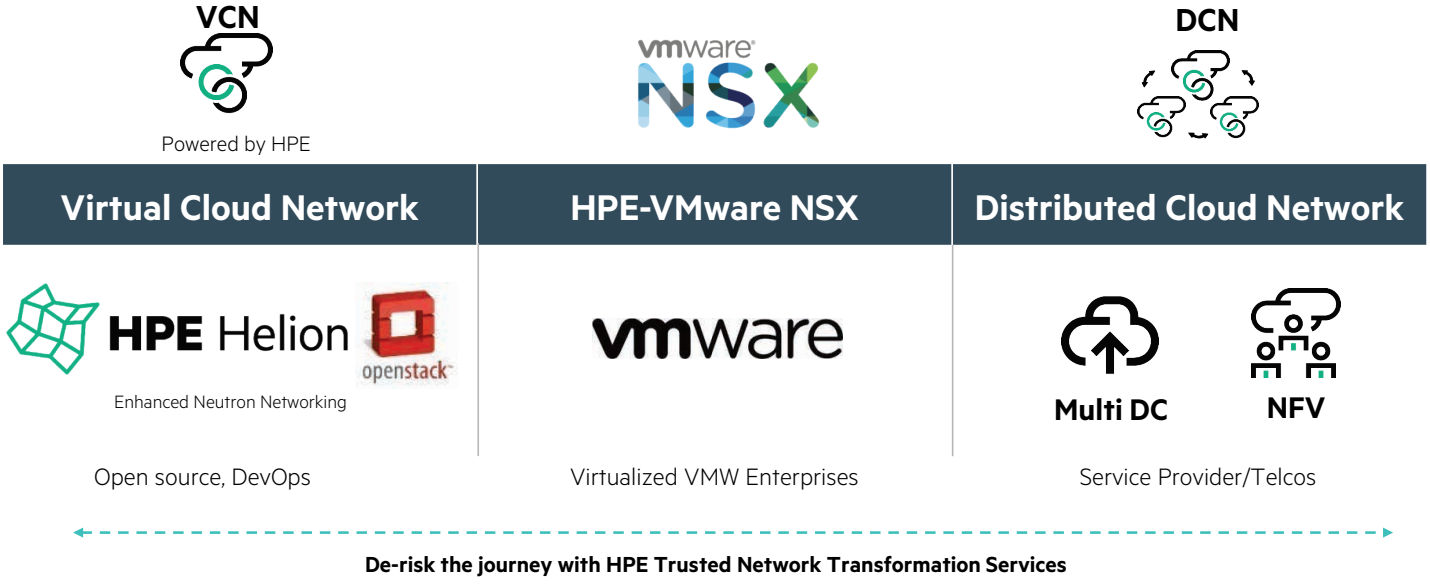


Figure 13. HPE VXLAN solutions

HPE Helion OpenStack — Virtual Cloud Networking (VCN)

HPE VCN is the enhanced Neutron networking component included as part of HPE Helion OpenStack. In the HPE Helion OpenStack solution, virtual overlay VXLAN networks are created via the HPE Helion OpenStack Horizon dashboard. VCN uses the Distributed Virtual Router (DVR) to dynamically set up and tear down VXLAN tunnels between compute nodes.

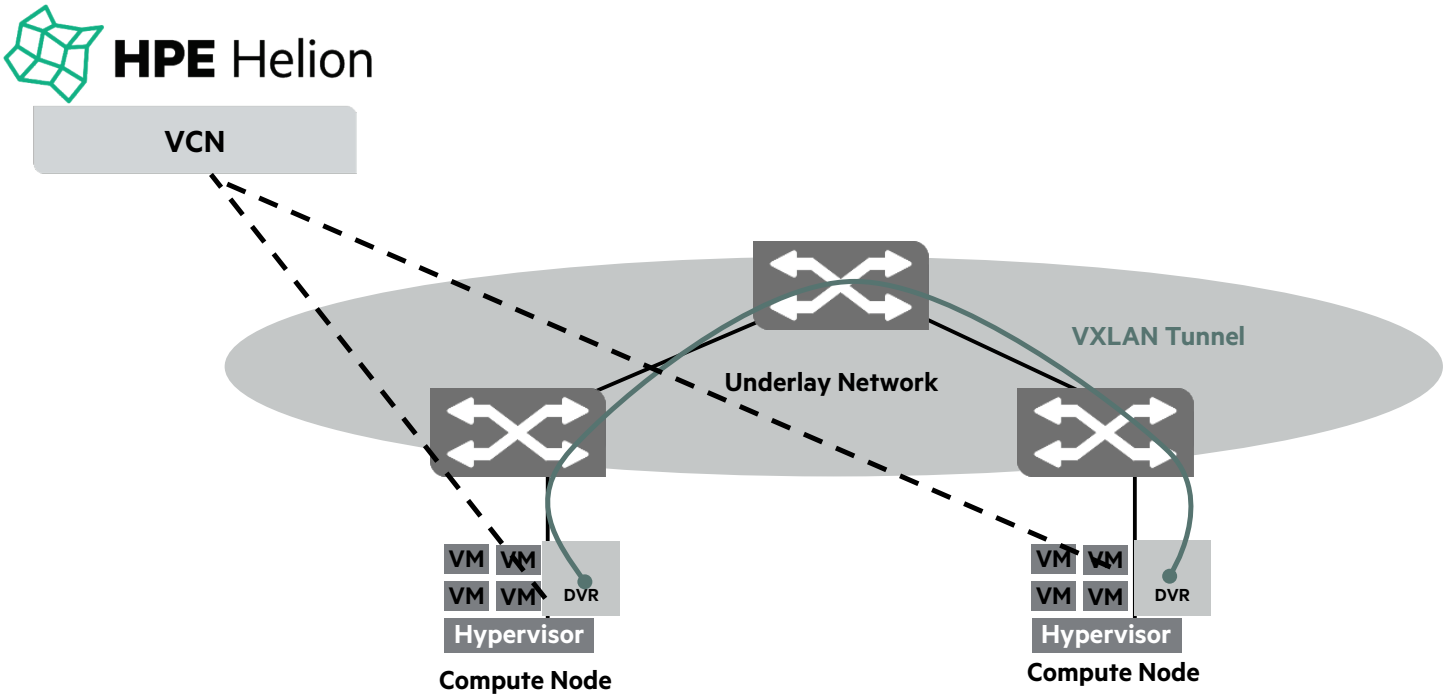


Figure 14. VCN dynamically setting up and tearing down VXLAN tunnels between software VTEPs compute nodes

Helion OpenStack 3.0 includes support for hardware L2 VTEPs to bridge VMs on compute nodes and bare metal servers on the same subnet.

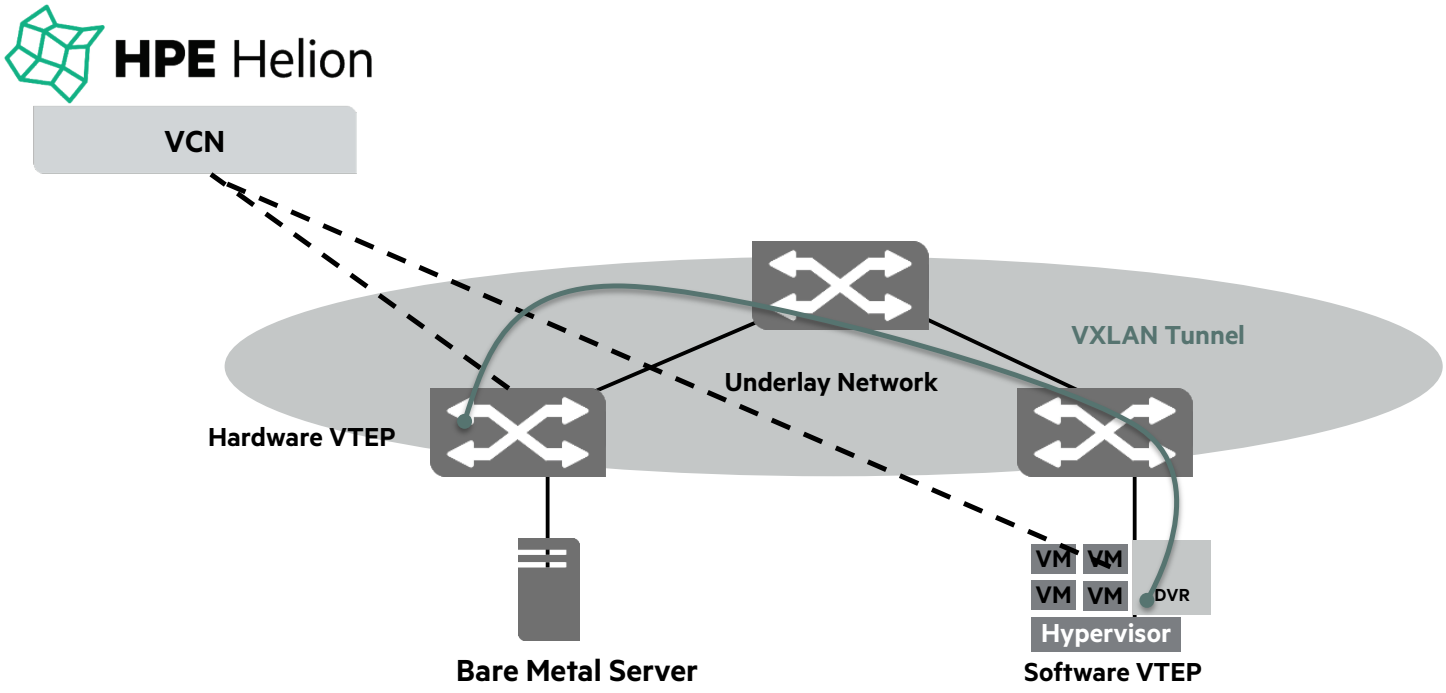


Figure 15. VCN dynamically setting up and tearing down VXLAN tunnels between hardware and software VTEPs

For more information on HPE Helion OpenStack, refer to the HPE Helion OpenStack webpage.

VMware NSXv integration

With the HPE 5930 VMware NSXv integration, HPE hardware VTEPs are able to integrate with VMware NSXv to bridge VMs on virtual networks to physical bare metal servers/physical firewalls/WAN routers etc. As shown in figure 16, virtual networks are created on the vSphere web client. This info is sent to the NSX manager and controller, which uses OVSDB to dynamically set up and tear down VXLAN tunnels between HPE hardware VTEPs and ESXi hypervisors. Since the NSXv controller knows about all the MAC addresses in the virtual networks, it shares these remote MAC addresses to VTEPs requiring this info via OVSDB. In this implementation, security certificates are required on all VTEPs to prove its identity, however, the underlay and overlay networks function separately.

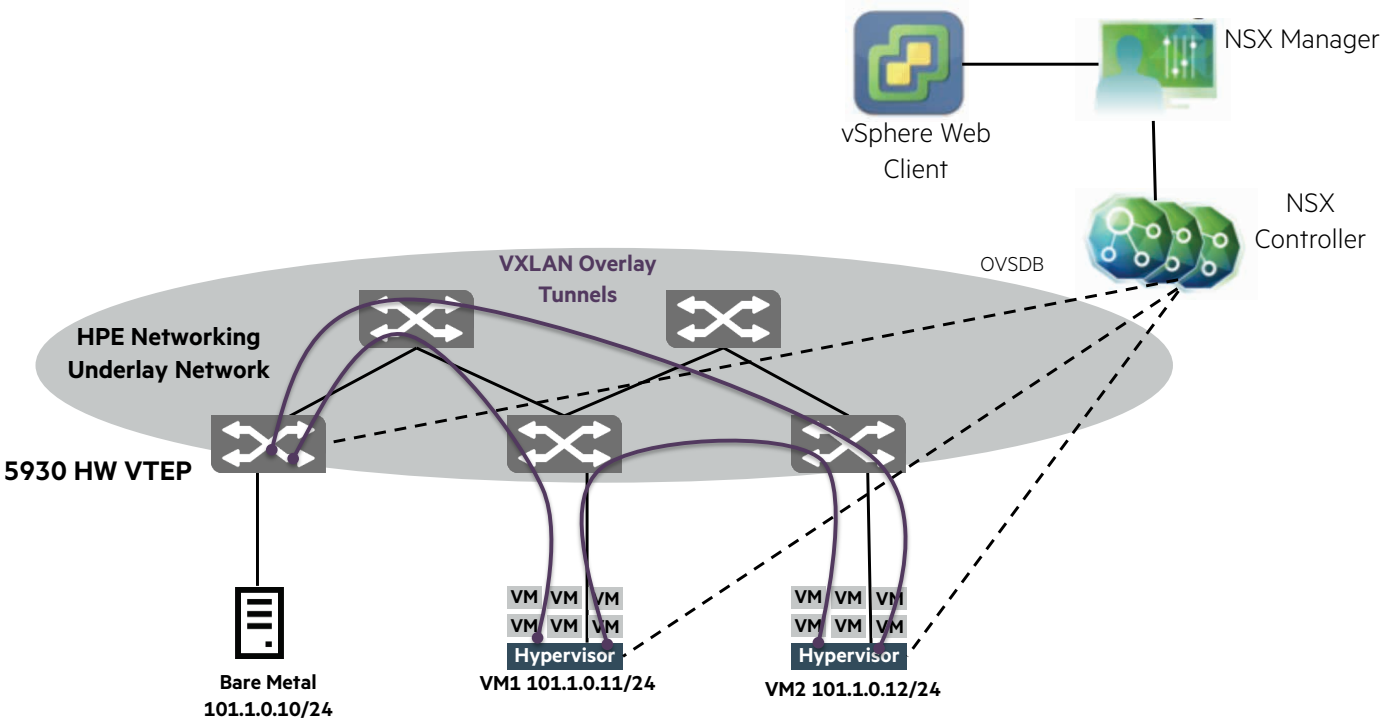


Figure 16. HPE 5930 VMware networking solution — direct OVSDB NSXv integration

For more information on HPE and NSX integration, refer to the HPE-VMware Networking Solution Technical Brief.

HPE Distributed Cloud Networking

HPE DCN provides large scale network virtualization, including federation, across multiple DCs. Virtual networks can be configured and consumed through HPE Helion OpenStack or HPE Matrix Operating Environment (MOE)/Cloud Service Automation (CSA). This acts as a broker for cloud services providing information which is also sent to HPE Virtualized Services Directory (VSD), which in turn provides a centralized orchestration and policy management framework. These policies are passed to HPE Virtualized Services Controller (VSC) for dynamic deployment seamlessly within and across datacenters. To dynamically create VXLAN tunnels, the HPE VSC uses OVSDB to HPE hardware VTEPs and OpenFlow to software VTEPs with HPE Virtual Routing and Switching (VRS) agents residing in the hypervisors. The solution supports OpenSource KVM, VMware ESXi, Xen and Hyper-V.

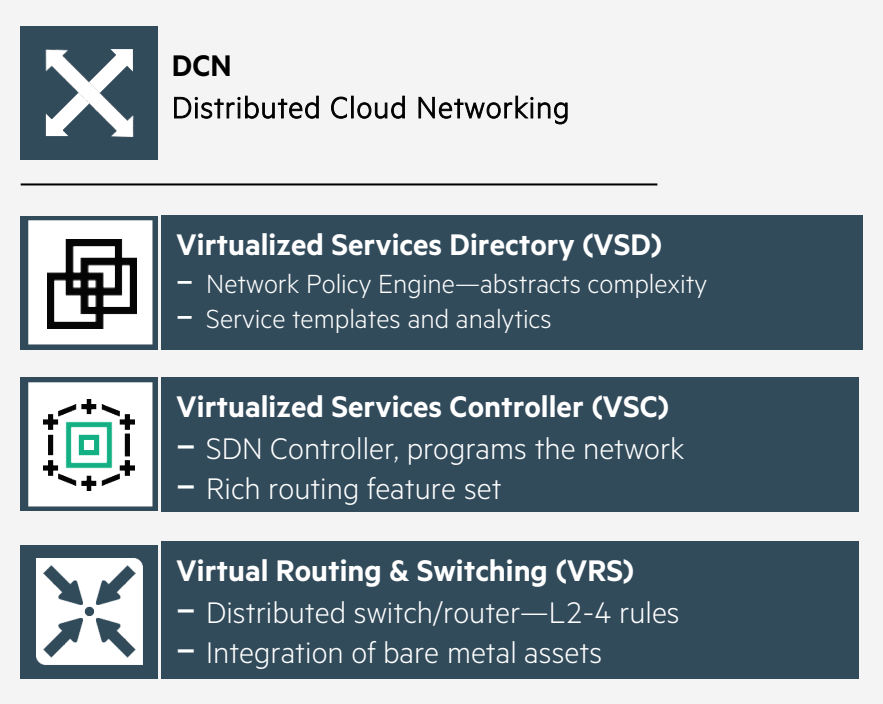


Figure 17. HPE DCN

For detailed information about these solutions, refer to VXLAN in HPE data center solutions.

Cloud

Cloud computing environments provide the flexibility businesses need so they can add compute, networking and storage on demand to respond to advancing technology and shifting business priorities. Cloud technology provides the necessary orchestration of resources to allow IT organizations to deploy applications on-demand, making IT services immediately accessible to internal customers in a standardized, automated way. By providing the same level of agility through private cloud services, security and compliance can be maintained while improving overall availability, performance and cost.

As customers started moving to the cloud, the following models of cloud delivery emerged:

- Private cloud:

  Owned and managed by the customer

- Managed cloud:

  Owned by the customer, but managed by a third party

- Public cloud:

  Cloud as a service, so the customer only pays for the resources they use

- Hybrid cloud:

  The organization provides and manages some resources in-house and others externally

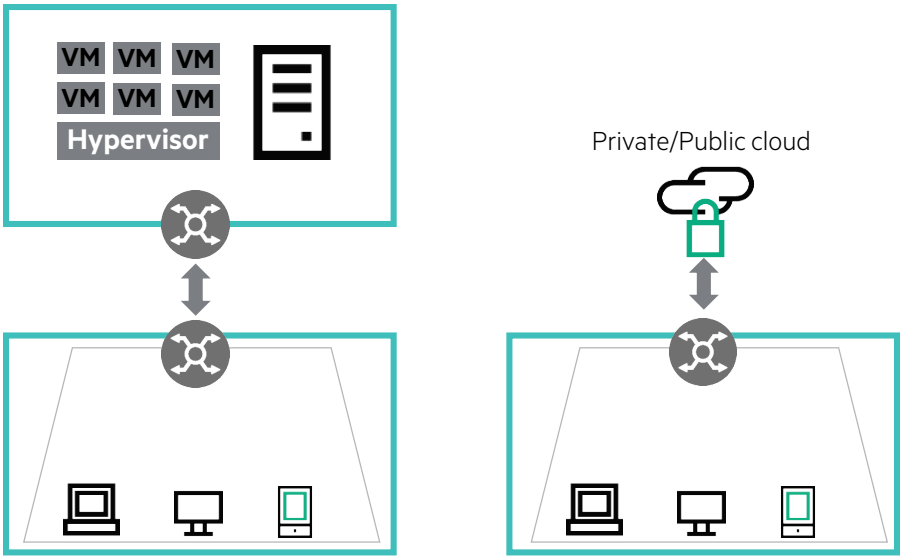


Figure 18. Centralized data center vs. cloud

HPE delivers scalable cloud fabrics with virtualization solutions in both overlay and underlay.
HPE Networking offers cloud-based network virtualization solutions such as HPE Virtual Cloud Networking and HPE Distributed Cloud Networking, discussed earlier, which allow cloud administration teams to provide support for public, private, and hybrid multi-cloud integration environments where multitenant and multi-application policies can be deployed.

TRILL

TRILL deployments combine the simplicity and flexibility of L2 switching with the stability, scalability and rapid convergence capability of L3 routing. This makes TRILL very suitable for large flat L2 domains in data centers. TRILL provides a mechanism that allows every single node to have a tree rooted at itself, allowing the optimal (shortest path) distribution of traffic as well as multipathing for failure recovery.

Why TRILL?

TRILL is an evolutionary step in Ethernet technology designed to address some of the shortcomings within Ethernet, specifically Spanning Tree and loop prevention. TRILL uses L3 multipathing and shortest path routing techniques to create large flat L2 domains, so that clients and VMs can move and migrate without having to change their IP addresses.

Benefits of TRILL include:

- Vendor neutral, non-proprietary technology
- No Spanning Tree Protocol (STP), loop-free, multipathing Ethernet fabric
- Distributed scale out Ethernet fabric with all ToR switch server ports having equal latency
- Stable underlay network for overlay SDN networks in the data center

Architecture and scalability

Each device in a TRILL architecture is called a routing bridge (RB). TRILL architectures can be as small as a few RBs meshed in a 2-tier type topology, but they can also expand so that they consist of many access layer devices that are meshed to many core layer devices.

TRILL establishes and maintains adjacencies between RBs by sharing link state protocol (LSP) messages and creating a link state database (LSDB) for all RBs in the network. Based on the LSDB, each RB uses a shortest path first (SPF) algorithm to calculate forwarding entries destined to other RBs.

The example below uses four chassis switches configured as core TRILL designated routing bridges (DRB). The DRBs are responsible for creating a distribution tree, which guides the forwarding of multi-destination frames such as multicast, broadcast and unknown unicast frames in the TRILL fabric.

A TRILL fabric using RBs that are not be able to perform L3 routing on TRILL encapsulated traffic will require a solution as shown in Figure 21. L3 routing in this example is offloaded to separate devices and can be configured as VRRP routers that act as the gateways for the access VLANs.

IRF is not specifically shown in the below image, but IRF and TRILL are not mutually exclusive. In the real world, each individual device shown could actually be a stack of devices virtualized into a single device using IRF.

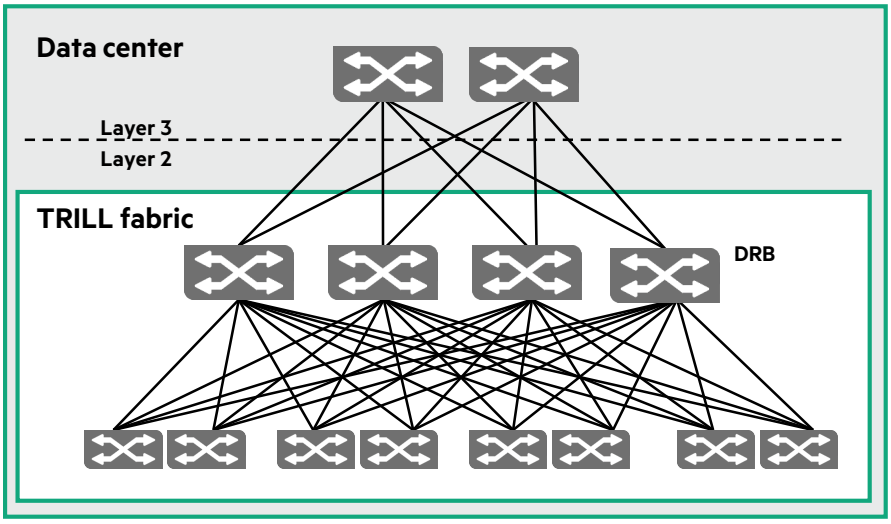


Figure 19. TRILL fabric with L3 offloaded to separate device

For detailed information on TRILL, refer to hpe.com/go/flexfabric.

Shortest Path Bridging IEEE 802.1aq

Shortest Path Bridging is an IEEE standards-based technology, ratified in 2012, which is able to present the data center network as one large L2 non-blocking network. SPB can greatly simplify how customers create and configure networks, both across the enterprise and in the cloud, by requiring provisioning at the edge of the network.

SPB is an edge provisioning solution, i.e., when new VMs are introduced at the edge, they only need to be configured at the network edge rather than throughout the entire network.

Why SPB?

SPB is able to simplify a network because it is an end-point provisioning system. The backbone devices in between the edge devices in a DC network are automatically provisioned through SPB's link state protocol and configuration is really only required at the edge of the network. SPB is able to transparently extend L2 and/or L3 domains across the core backbone with very little effort.

Benefits of SPB include:

- Satisfies VM migration requirements by extending L2 networks across the backbone
- Extends the L2 environment between data centers for long distance VM migration and disaster recovery
- Provides optimal bandwidth with multipathing support, equal cost paths, and shortest path forwarding
- Standards-based (IEEE 802.1aq), resilient protocol with sub-second failover
- Provides carrier-class multitenancy support with enterprise friendly features
- Simplicity that can offer plug-and-play deployments helping to reduce service time
- Provides loop free L2 environments eliminating previous constraints on network topologies
- Scalability up to 16 million tenants

Architecture and scalability

There are two variants of SPB—SPBV and SPBM. SPBV uses the VID and incorporates the same MAC-in-MAC technique from QinQ (802.1ad), while SPBM leverages Provider Backbone Bridging (PBB, 802.1ah). SPBM is the current version supported by HPE.

SPBM networks which can scale to support up to 16 million tenants, include backbone edge bridges (BEB) and backbone core bridges. The backbone edge bridges connect customer sites to the SPBM backbone core network. The SPB control plane utilizes IS-IS link-state routing protocol to prevent loops and calculate the shortest path from a frame.

BEBs, which are identical to PEs in an MPLS VPN network, encapsulate frames into MAC-in-MAC frames before forwarding them to the Backbone Core Bridge (BCB) network. They also decapsulate MAC-in-MAC frames before sending them to a customer site.

For customer frames to be transmitted across an SPBM network, the ingress BEB encapsulates them in MAC-in-MAC format. In the outer frame header, the source MAC address is a Backbone MAC address (B-MAC) of the ingress BEB, and the destination MAC is a B-MAC of the egress BEB. All devices in the SPBM network forward the MAC-in-MAC frames based on the destination B-MAC and B-VLAN.

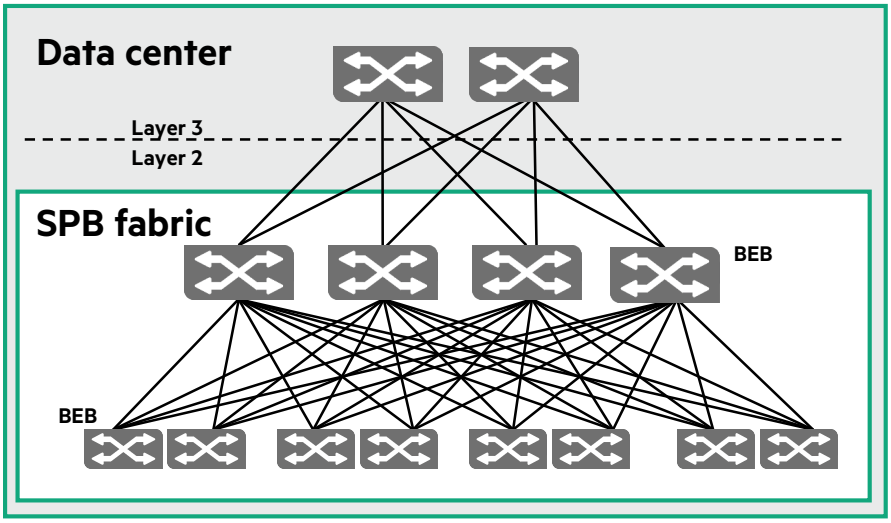


Figure 20. SPBM fabric with L3 offloaded to separate device

Provider Bridging (QinQ) IEEE 802.1ad

Provider Bridging is an IEEE standard, ratified in 2005 and sometimes referred to as 802.1QinQ, which enables multitenant environments.

Provider bridging enables an Ethernet frame to carry two VLAN tags, which allows the service provider to offer an Ethernet LAN service over their existing network where the tenant can maintain their own VLANs. In the 802.1ad frame, the outer tag is the VLAN allocated to the tenant in the provider network and as such is called the S-Tag (service tag), which contains the service VLAN ID (S-VID). The inner tag represents the VLANs belonging to the tenant and is called the C-Tag (customer tag) and contains the customer VLAN-ID (C-VID). Provider Bridging allows for up to 4096 S-Tags and 4096 C-Tags.

Forwarding in Provider Bridging networks requires switches to perform two new operations:

- Tag pop—removes the outer VLAN tag
- Tag push—adds an outer VLAN tag

While QinQ allows for each customer to be contained to a unique VLAN, it does not abstract the customer and provider network entirely. Switch forwarding decisions are still based on the S-Tag and the Destination MAC address, therefore, the provider network needs to learn the MAC addresses of all the customer devices. Furthermore QinQ does not restrict the flow of L2 control frames for protocols such as Spanning Tree. To address these shortcomings, Provider Backbone Bridging (PBB) was developed.

Provider Backbone Bridging (MAC-in-MAC) IEEE 802.1ah

PBB or MAC-in-MAC is an IEEE standard, ratified in 2008, which extends the work achieved in 802.1ad by providing a protocol-based, hierarchical multitenant network infrastructure that completely abstracts the service provider backbone from the customer network. PBB networks can be deployed to aggregate existing QinQ networks.

When a frame arrives at a Backbone Edge Bridge (BEB) it is encapsulated and the BEB adds the backbone source, destination MAC (B-SA and B-DA) and backbone VLAN. The I-SID or service identifier field is a 24-bit customer specific identifier, which supports 224 or ~16 million tenants.

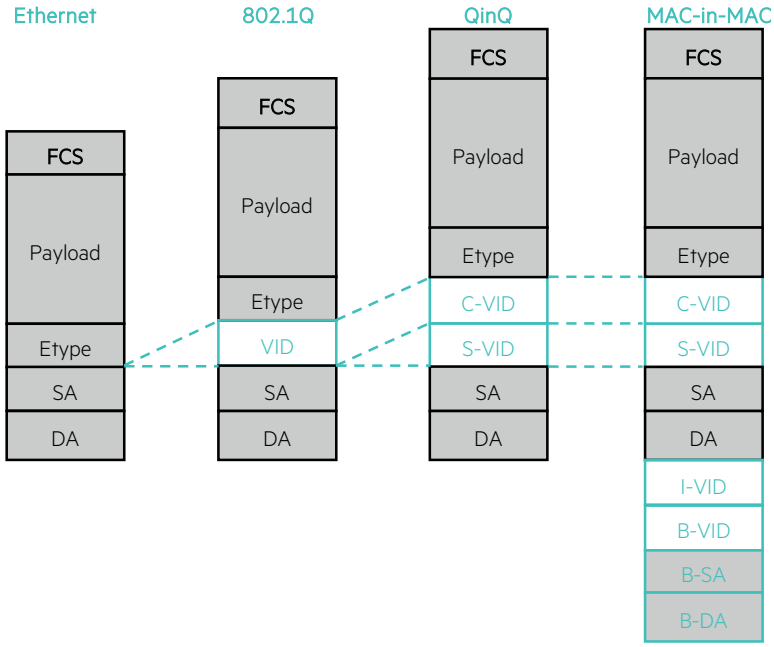


Figure 21. Evolution from Ethernet to MAC-in-MAC

HPE Multitenant Device Context

Multitenant data centers can provide segmentation between different organizations within the same company, as well as segmentation between multiple different companies.

Multitenant data centers are becoming more and more common, and they have been bringing many challenges to the networking environment. Customers or tenants, naturally want to be able to migrate their workloads from an internal enterprise network to a service provider's data center, while still keeping the same networking configurations as the internal enterprise network. The service providers are being challenged to meet these needs while still keeping their own operational flexibility, efficiency and standards.

HPE MDC is a device-level, multitenant technique that provides the ability to logically partition a single physical device into many virtual devices. This capability gives an administrator the flexibility to set up multiple customers or lines of business on the same physical hardware while ensuring dedicated and discrete management, security and network services separation and isolation from other "tenants".

MDC provides true isolation in a multitenant environment by allowing the memory of a single device to be separated into protected partitions running different instances of the switch. This allows for full separation of forwarding databases and traffic in each MDC for those different tenants. As an example, in the enterprise context, those tenants could be the following departments: finance, marketing, R&D and legal.

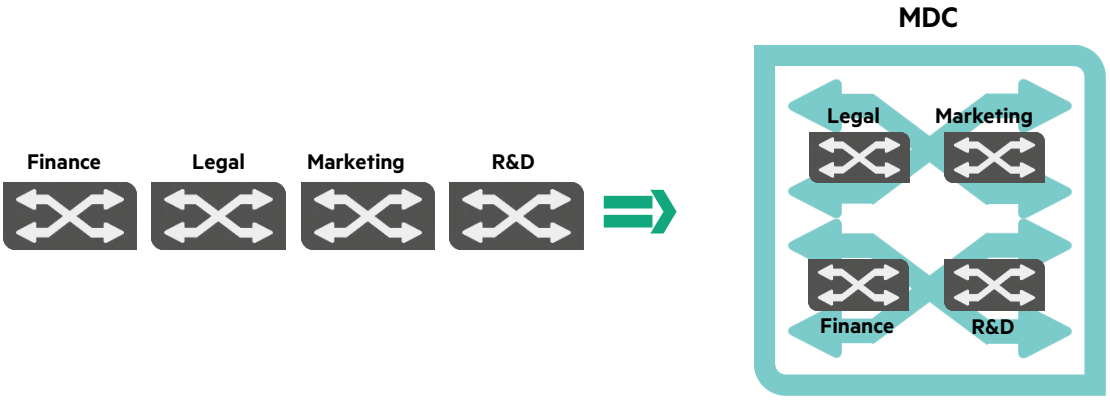


Figure 22. HPE Multitenant Device Context

MDC provides complete separation and increased resiliency, especially when used with IRF across a group of switches in the core of a network. MDC can exist in all the switches in that group. In this way the organization could have complete transparent failover in the event of a single device failure, or when a single device needs to be taken offline for service.

Serving up to nine tenants in a single device, with full separation instead of building nine separate networks, organizations can use up to 75 percent less equipment and spend up to 75 percent less money. They consume up to 75 percent less power in cooling and space, and administration.

HPE MDC can be complimented and works in conjunction with protocol-based, multitenant solutions like QinQ, PBB or SPB.

For more information on MDC, refer to hpe.com/go/flexfabric and the switch configuration guide.

SDN

SDN as defined by the Open Networking Foundation (ONF), is the physical separation of the network control plane from the forwarding plane and where the control plane controls several devices. SDN promises an easier, more dynamic interaction with the network through a "clean" interface obtained through this abstraction of the control plane. This reduces the complexity of managing, provisioning, and changing the network.

The HPE SDN architecture departs from legacy solutions by building networks from three abstractions or layers. First, the infrastructure layer acts as the foundation for an SDN architecture. The infrastructure consists of both physical and virtual network devices such as switches and routers. These devices implement the OpenFlow protocol as a standards-based method of implementing traffic forwarding rules.

Second, the control layer consists of a centralized control plane for the entire network. The control plane is decoupled from the underlying infrastructure to provide a single centralized view of the entire network. The HPE VAN SDN controller provides this control layer and utilizes OpenFlow southbound to communicate with the infrastructure layer.

Third, the application layer consists of network services, orchestration tools and business applications that interact with the control layer. These applications leverage open interfaces to communicate with the control layer and the network state.

In the HPE SDN architecture, the HPE IMC VAN SDN Manager module provides comprehensive management—including fault, configuration, accounting, monitoring and security for SDN environments. IMC VAN SDN Manager Software integrates with the IMC platform, providing administrators with a single interface to manage both the traditional network as well as the SDN.

The HPE SDN architecture provides an end-to-end solution to automate the network from data center to campus and branch. Expanding the innovation of SDN, HPE SDN ecosystem delivers resources to develop and create a market place for SDN applications.

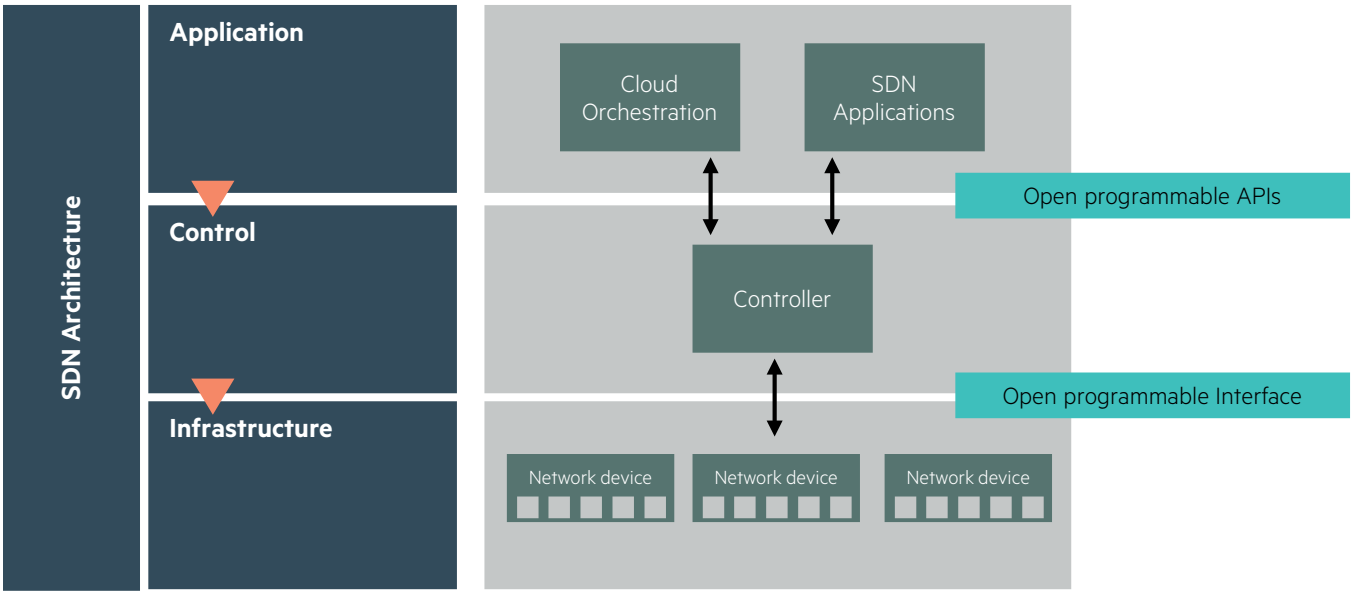


Figure 23. SDN architecture framework

HPE CloudFirst SDN is currently focused on cloud networking and network virtualization to support network automation, VMs and multitenancy at scale. This SDN approach allows customers to integrate physical and virtual environments, and unlock never-before realized capability, intelligence and visibility. The customer has a choice in how this integration is accomplished—a choice which has direct implications on whether they will merely solve their network virtualization problems or whether they'll place themselves on a path to unlock the full potential of an intelligent, SDN-enabled converged infrastructure.

At the core of HPE approach to SDN is a hybrid architecture, combining the historic benefits of standard, distributed L2 and L3 network protocols with an OpenFlow-based policy and overlay provisioning model. This ensures that existing network behaviors and management tools can be preserved, while the benefits of software programming can be used to enhance the visibility and security of the data center network to respond to dynamic changes in the operational network and application environment. All packets are optionally processed by a set of OpenFlow rules before being passed to the normal hardware processing pipeline where standard bridging or routing protocols can determine the forwarding path. This addresses scaling and maturity concerns that customers might otherwise have with this new technology.

SDN applications

SDN applications provide a true end-to-end service level for network performance, quality of service, and security, which can be tuned to an application's needs. For example, SDN applications can provision policy, inspect flows, or perform other network control functions via the HPE SDN controller.

SDN applications can be network-service oriented, but they can also accommodate virtually any business service use case. For example, for applications that require fast response, such as IP voice or high-frequency trading, an SDN application can establish policy programmatically and with a precise level of control, and the policy will be enforced across the enterprise network by the SDN Controller.

Northbound APIs utilize a REST architecture and provide easy access to applications that are co-resident or run remotely from the Controller, such as OpenStack drivers or network management tools. Native APIs, provided in Java, deliver support to Network Control applications that are integrated directly in the controller. The fundamental extensibility and open RESTful APIs with the HPE SDN Controller will allow innovative new applications to be created that make requests of the underlying network without the need to physically uproot or reconfigure the underlying infrastructure.

In fact, industry experts predict a swift migration to SDN. The overall SDN market will reach USD $6.3 billion by 2017 and nearly USD $1.1 billion will come from SDN applications.[2]

[2] Technology assessment: SDN Momentum Builds in Datacenter and Enterprise Networks, August 2014, IDC #250288.

HPE is actively building an SDN application developer ecosystem to foster the creation of SDN applications. HPE is offering an SDN developer community, as well as forums, events and other services, to help developers and go-to-market partners build and sell SDN applications. Third-party SDN applications are available through an HPE App Store. And HPE is working with partners to sell SDN applications as part of a broader effort that encompasses the application, network and services sales.

HPE and partners have already created several SDN applications, and this ecosystem will continue to grow.

These applications include:

- Network Visualizer:

  The HPE Network Visualizer application combines dynamic network traffic capture capabilities with detailed real-time monitoring, allowing fast network diagnosis and verification. This delivers a rapid incident to fix transition, as compared with traditional investigation methods which are often disruptive and time consuming.

- Network Protector:

  The result of a partnership with enterprise-security specialists TippingPoint and HPE ArcSight, this HPE application enables real-time threat detection and security policy enforcement at the edge of campus networks. Network Protector leverages the Virtual Application Networks SDN Controller and OpenFlow to program the network infrastructure with security intelligence from the TippingPoint RepDV Labs database. If it detects a threat, the user is prevented from accessing the threat site and the incident is logged by HPE ArcSight Logger.

- BlueCat DNS Director:

  Delivers network-driven enforcement of DNS policies that allow security infrastructures to gain complete visibility and control through IP address management data across all devices and applications.

- GuardiCore Defense Suite:

  Provides highly scalable, SDN-based network security for software-defined data centers, detecting and mitigating advanced persistent threats, malware propagation, and insider attacks, at an early stage.

These are just a few examples of possible SDN applications that can be developed with the HPE SDN Controller and ecosystem.

Visit hpe.com/us/en/networking/sdn.html to learn more about our SDN strategy and our growing ecosystem of SDN partners and applications.

Network security

Network security is one of the most important factors a company needs to consider. Increasing network security decreases the chance of an intentional hacker gaining access to a system as well as preventing unintentional breaches exposed by negligent employees. If the network is not secured even against topology and configuration changes, then all other layers of security built on top of the network are susceptible to attack.

Below is a list of base security features that should be deployed to secure a network:

- Physically secure:

  While there are features to protect a device that is physically accessible, it is not as secure as a device that is physically secured.

- Management access control:

  After physically securing a device, it is necessary to control access to the management interfaces of the devices. This includes strong passwords and ideally a centrally managed authentication infrastructure, such as RADIUS. The recommended management protocols: SSH, SSL, SNMPv3 and SFTP are encrypted protocols and help with privacy and authenticity.

  RADIUS is recommended so that usernames and passwords do not have to be shared as well as for accounting and auditing.

- IP authorized managers and management ACLs:

  IP authorized managers and management ACLs are used to deny access to the device except from specified IP addresses.

- Management VLAN:

  Can be used to further secure access to the device as a management VLAN is usually not routed to other VLANs and requires that a management station be on the Management VLAN.

- MAC Lockdown and MAC Lockout:

  These two features provide a type of port-based security. Both involve the specification of MAC addresses as part of their configuration. MAC Lockdown is used to ensure a particular device can access the network only through designated ports whereas MAC Lockout is used to ensure a particular device does not access the network through one or more switches.

- Port security:

  This feature enables an administrator to configure each switch port with a unique list of device MAC addresses that are authorized to access the network through that port. This enables individual ports to detect, prevent and log attempts by unauthorized devices to communicate through the switch.

- ACL traffic filters:

  Traffic can be controlled based on multiple tuples.

- Spanning Tree Protection:

  This feature set protects against rouge devices being inserted into the network, and causing topology changes and service interruptions. The specific features are: BPDU Filtering, BDPU Protection, Root Guard and TCN Guard.

- DHCP protection:

  DHCP is designed to work in the trusted internal network and does not provide authentication or access controls. Because of this lack of built-in security, a DHCP server has no way of verifying that the client requesting an address is a legitimate client on the network. Similarly, the DHCP client has no way of knowing if the DHCP server that offers it an address is a legitimate server or not. Therefore, DHCP is vulnerable to attacks from both rogue clients and servers.

- Address spoofing:

  A rogue DHCP server on the network can assign invalid IP addressing information to client devices. This includes the IP addresses of the client itself, the default gateway, DNS servers and WINS servers. Without valid IP addresses, the legitimate client devices are unable to contact other legitimate IP network devices and users are prevented from reaching the resources they need to do their jobs. A rogue DHCP server can also be used to perform a man-in-the-middle attack.

- Address exhaustion:

  An attacker can access the network and request IP addresses until the DHCP server's supply of available IP addresses is exhausted. This prevents legitimate clients from receiving IP addresses and accessing the network.

- ARP protection:

  ARP is used to resolve a device's IP address to its MAC address. ARP creates and populates a table of known IP addresses and the associated MAC addresses as it requests information for unknown MAC addresses. Most ARP devices update their tables every time they receive an ARP packet even if they did not request the information. This makes ARP vulnerable to attacks such as ARP poisoning.

  ARP poisoning occurs when an unauthorized device forges an illegitimate ARP response, and other devices use the response to change their ARP tables. By positioning itself using a traditional "man-in-the-middle" style attack, a rogue device can capture information such as usernames and passwords, email messages, and other confidential company information.

  ARP poisoning can also take the form of unsolicited ARP responses and can lead to DoS attacks. For example, a rogue device can poison other devices' ARP tables by associating the network gateway's IP address with the MAC address of some endpoint station. Because the endpoint station does not have access to outside networks, outgoing traffic is prevented from leaving the network. The endpoint station may also become easily overwhelmed by the unexpected traffic.

  ARP protection can defend against these types of attacks by validating that an ARP message comes from the legitimate owner of the MAC and IP address.

- IP Spoofing Protection:

  Many network attacks occur when an attacker injects packets with forged IP source addresses into the network. Also, some network services use the IP source address as a component in their authentication schemes. For example, the BSD "r" protocols (rlogin, rcp, rsh) rely on the IP source address for packet authentication. SNMPv1 and SNMPv2c also frequently use authorized IP address lists to limit management access. An attacker who is able to send traffic that appears to originate from an authorized IP source address may gain access to network services for which he is not authorized.

Comware switches provide a feature called IP Source Guard that can help to mitigate an IP spoofing attack. The IP Source Guard function can be enabled on user access ports of the switch to improve network security. It prevents illegal packets from traveling through the ports. When a port enabled with the IP Source Guard function receives a packet, the port looks up the key attributes (including IP address, MAC address and VLAN tag) of the packet in the binding entries of the IP source guard. If there is a match, the port forwards the packet. If there is no match, the port discards the packet. IP source guard bindings are on a per-port basis. After a binding entry is configured on a port, it is effective only on that port.

Network firewalls

Firewalls are used to deter unwanted traffic on a network or specific network segments. Firewalls have gone through a number of iterations and continue to evolve. First generation firewalls used basic ACLs as packet filters and were often utilized for NAT/PAT.

Second generation firewalls are stateful filters. A stateful firewall is able to hold significant attributes of each connection in memory, from start to finish. The advantage of stateful firewalls is that rules only need to be written for traffic in one direction and only return traffic is permitted back through the firewall.

Third generation firewalls provide application layer filters. An Application Specific Packet Filter (ASPF) implements application layer and transport specific, namely status-based packet filtering.

Firewalls come in various forms; PC software, network appliance, blade for network device, or features in router/switch OS.

There are three areas in the enterprise network architecture where firewalls are being deployed. The traditional deployment is between the enterprise LAN and the Internet. This is still required in today's networks in addition to deployments between the data center and users and finally between different types of users. For example, a firewall might be deployed between finance and R&D VLANs or between employee and guest VLANs.

Next generation firewalls

Next Generation Firewalls (NGFW) add the ability to apply policy to an application. Today many different applications use the same port, such as TCP 80 or TCP 443, and it is not possible to distinguish between applications by the port number alone. NGFWs perform a deep packet inspection to identify different applications and allow for customized policy per application. NGFWs can also include full IPS as well as traditional firewall and routing capabilities.

Intrusion prevention system

In today's network environments, the "network perimeter" is becoming blurred. This is due to traffic entering the network using a VPN or mobile users—employees and guests connecting to the network while at the customer site, particularly using wireless access points.

This drives the need to consider a "defense-in-depth" strategy; having a firewall alone is no longer considered the only security device on the network. In addition to the network border, the internal network is subdivided into separate "attack domains" (also known as "security broadcast domains"), by IPSs that contain outbreaks within a LAN.

The IPS's main component is the Threat Suppression Engine (TSE). The TSE reconstructs and inspects flow payloads at the application layer. As each new packet belonging to a flow arrives, the flow is re-evaluated for malicious content. The instant a flow is deemed malicious, the current packet and all subsequent packets pertaining to the flow are blocked. This ensures that the attack never reaches its destination.

Each flow is tracked in the "connection table" on the IPS. A flow is uniquely identified by the port on which it was received and its packet header information, referred to as the "6-tuple":

- IP protocol (ICMP, TCP, UDP, other)
- Source IP address
- Source ports (TCP or UDP)
- Destination IP address
- Destination ports (TCP or UDP)
- IPv6

Once classified, each packet is inspected by the appropriate set of protocol and application filters. The IPS filter engine combines pipelined and massively parallel processing hardware to perform simultaneous filter checks on each packet. The parallel filter processing ensures that the packet flow continues to move through the system with a bounded latency (on the order of microseconds) for the most part, independent of the number of filters that are applied.

IPS deployment

There are two options for deployment. The first is to use an NGFW, combining the capabilities of a traditional firewall and IPS plus new features into a single appliance. The second is to deploy a traditional firewall and an IPS.

The most common deployment is at the network perimeter, which in a data center are those links connecting the data center network to the Internet. It is desirable to deploy a network security solution, NGFW or traditional firewall/IPS combination, between any high value resource and a less trusted device or user. For example, it is recommended to deploy a network security solution between a data center and a campus network. It is also recommended to deploy between different types of users such as guests and employees, or finance and engineering.

An IPS may be deployed in front of the firewall, however, most customers will deploy it behind the firewall. In these deployments, the first device can reduce load on the next device by first dropping packets that don't pass its specific rules.

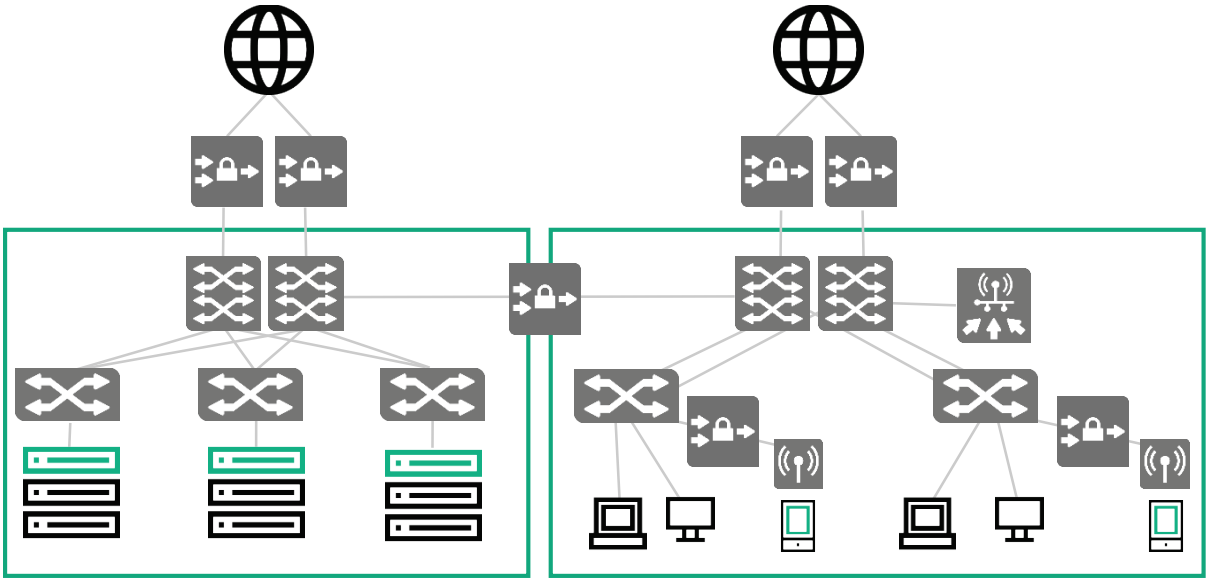


Figure 24. IPS deployment options

The IPS is placed in line between two network elements (i.e., between two routers or switches) or can be placed on a switch where it can inspect individual VLANs.

The IPS doesn't act as a network element in the sense that it does not route traffic, it simply inspects the traffic. Because the IPS is an inline device, the physical interfaces must match the segment in which it will be placed. These are individual segments and are not shared.

Data center topologies

Data center topologies have traditionally been deployed as multitier environments where servers are connected to access layer switches. These then connect to aggregation or end/row switches, which then connect to core layer switches. Although successfully deployed widely, these multitier topologies can also have disadvantages in that they can require many devices, routing occurred only in the core, and they can allow some very high subscription ratios.

These types of topologies may still have their place but modern data centers are able to take advantage of a combination of newer high performance, scalable modular core, and access layer switches with rich feature sets to flatten and simplify the data center.

Objectives

There can be many diverse objectives when architecting a data center. Some may be more or less important depending on the situation. Below is a list of typical objectives to be aware of:

- Public, private, hybrid cloud support:
  - It is critical when building out a new Hybrid Infrastructure data center that the infrastructure chosen provides developers with the on demand infrastructure they require to go from idea to revenue.
  - Enable DevOps orchestration infrastructure which can accelerate IT operations, app delivery and quality
- Create flexibility:
  - Open and standard-based systems and environments reduce risk and increase flexibility for your applications, infrastructure, and data needs.
- Flatter network with increased frame forwarding and packet forwarding:
  - Traditional legacy 3-tier data centers were not able to keep providing the performance needed in the modern virtualized data centers that now see immense amounts of east-west traffic flows. Data center networks can see great benefits from eliminating these unneeded hops needed to get from rack to rack.
  - 1-tier networks are able to truly optimize the traffic flows between racks by ensuring each rack can reach another rack with one hop.
  - 2-tier spine and leaf deployments have become one of the most common data center network architectures used by the industry today. These types or designs offer greater flexibility and scalability than 1-tier solutions while still providing for exceptional performance, predictably, efficiency and resiliency with a maximum of two hops from rack to rack.
  - HPE IRF can be used as a switch virtualization feature, helping to create larger scalable devices which provide many benefits including adding redundancy but also allowing extra layers like aggregation layers to be eliminated, helping to reduce the number of hops and increasing performance.
- L2 vMotion flexibility:
  - L2 connectivity always needs to be considered. Data centers can deploy a single L2 domain, leveraging TRILL or SPB, or by using L3 with VXLAN overlay technology, or vSphere 6.0.
- Reduced management complexity:
  - Flattening the network with large core switches and the elimination of aggregation switches combined with leveraging IRF in the various layers of the network simplifies what was typically a complex management scheme.
- Zero need for STP/RSTP:
  - HPE IRF can eliminate the need for STP, by presenting many physical switches as a single logical switch. This allows active/active connections between the layers, instead of relying on a loop prevention technology like STP, which results in better performance and faster fail-over.

Blade server 1-tier topology

When looking at flattening a network and providing substantial support for virtualization and East-West traffic flows, blade server 1-tier topologies have provided many advantages. Server blade enclosures allow for substantial compute density per rack and row, and HPE has optimized the HPE BladeSystem c-Class server portfolio to support the vision and reality of virtualization. This type of blade server 1-tier network topology optimizes and combines the reality of high performance networking with simplicity. It allows flexibility in networking by supporting IRF and a variety of interconnect options, including Comware based HPE 6127XLG switches as well as HPE Virtual Connect (VC) modules with server optimized features such as server profiles.

Network design

Typical blade server 1-tier deployments will usually consist of two to four core switches utilizing IRF, which then connect to blade-servers housed in C-Class enclosures. These types of deployments are able to extend VLANs across the entire data center and are optimized for virtualized environments where VMs may be moving from rack to rack or data center to data center.

The physical connections between the spine and leaf in these topologies will be either LAGs or routed links of 10GbE, 40GbE, and even 100GbE uplinks, providing for very high performance that can provide line rate performance between the spine and leaf.

The following devices can be used as the core switch in this type of topology:

- Jumbo data centers (~6000 physical servers)
  - HPE FlexFabric 12900E Switch Series
- Large data centers (~2000 physical servers)
  - HPE FlexFabric 12900E Switch Series
- Medium data centers (~500 physical servers)
  - HPE FlexFabric 7900 Switch Series
- Small data centers (~100 physical servers)
  - HPE FlexFabric 59xx Switch Series

Rack and logical diagrams

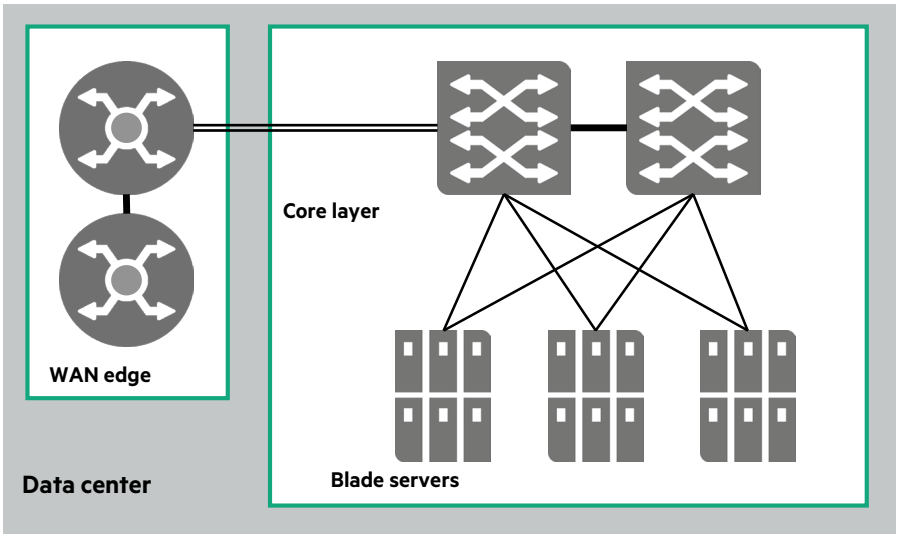


Figure 25. Blade server 1-tier topology rack view

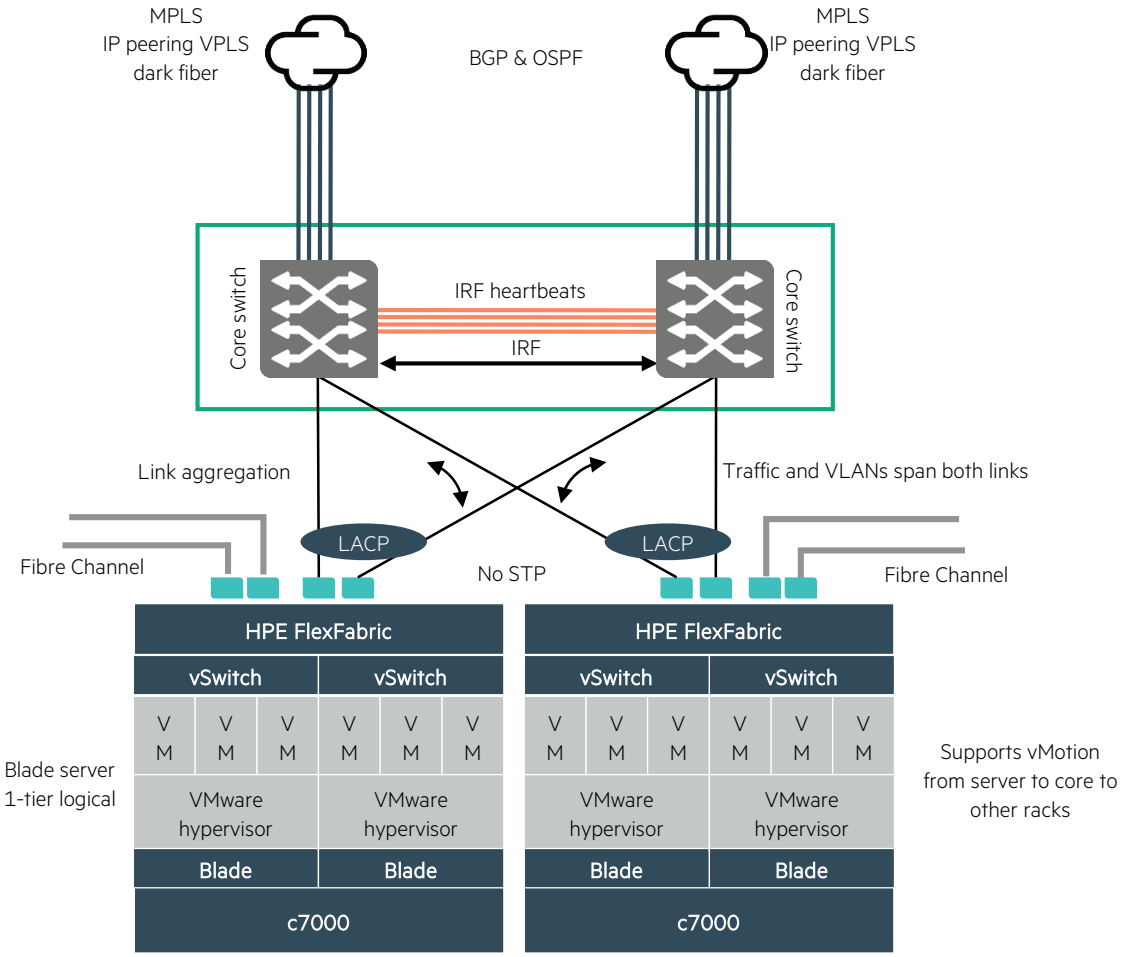


Figure 26. Blade server 1-tier logical

Spine and leaf topologies

Spine and leaf topologies, sometimes referred to as 2-tier, can provide a balanced network that can be optimized for virtualization while still providing for great scale and flexibility so the data center can adapt to changing application requirements. Spine and leaf are also used when the physical cabling plant cannot support a 1-tier design.

Although these topologies may also utilize rack servers, enhancements to the HPE BladeSystem c-Class server portfolio allow these types of topologies to combine the reality of high-performance networking with simplicity. The enclosures allow flexibility in networking, supporting a variety of interconnect options, including Comware based 6127XLG switches with IRF solutions as well as HPE VC modules with server optimized features such as server profiles.

Network design

Many spine and leaf deployments will utilize two to four spine (core) switches, which then connect to various leaf switches (ToR). These leaf switches, which may or may not be utilizing IRF, will then connect to servers usually within the same rack. Leaf switches can also be blade switches which are installed directly into HPE C-Class BladeSystem enclosures.

Large scale L2 spine and leaf topologies can be built using TRILL or SPB as the technology used to extend L2 networks across the backbone data center network. Devices running SPB or TRILL can still leverage IRF. However, large scale spine and leaf solutions usually use L3 solutions (either BGP or OSPF) coupled with overlay solutions like VXLAN to enable L2 connectivity between leafs.

Spine and leaf solutions are easily able to scale well past the common dual spine solution. In these solutions that use more than two spine switches the WAN edge is usually connected to a pair of leaf switches, known as border leafs. This allows for dual redundant connections to the WAN edge without requiring the WAN edge to be directly connected to each and every spine switch.

The physical connections between the spine and leaf in these topologies will be either LAGs or routed links of 10GbE, 40GbE and even 100GbE uplinks, providing for very high performance that can provide line rate performance between the spine and leaf.

The following devices can be used as the spine switch:

- Large data centers (>25,000 physical servers)
  - HPE FlexFabric 12900E Switch Series
- Medium data centers (>1,000 – <25,000 physical servers)
  - HPE FlexFabric 7900 Switch Series
- Small data centers (<1,000 physical servers)
  - HPE 59xx Switch Series

The following devices can be used as the leaf switch:

- HPE FlexFabric 12900E Switch Series
- HPE FlexFabric 7900 Switch Series
- HPE 59xx Switch Series

Rack and logical diagrams

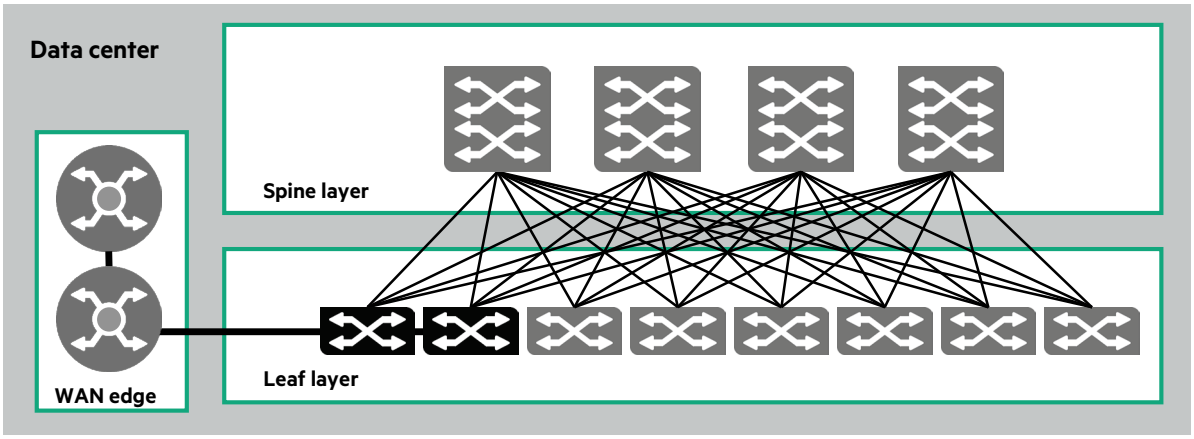


Figure 27. Spine and leaf topology rack view

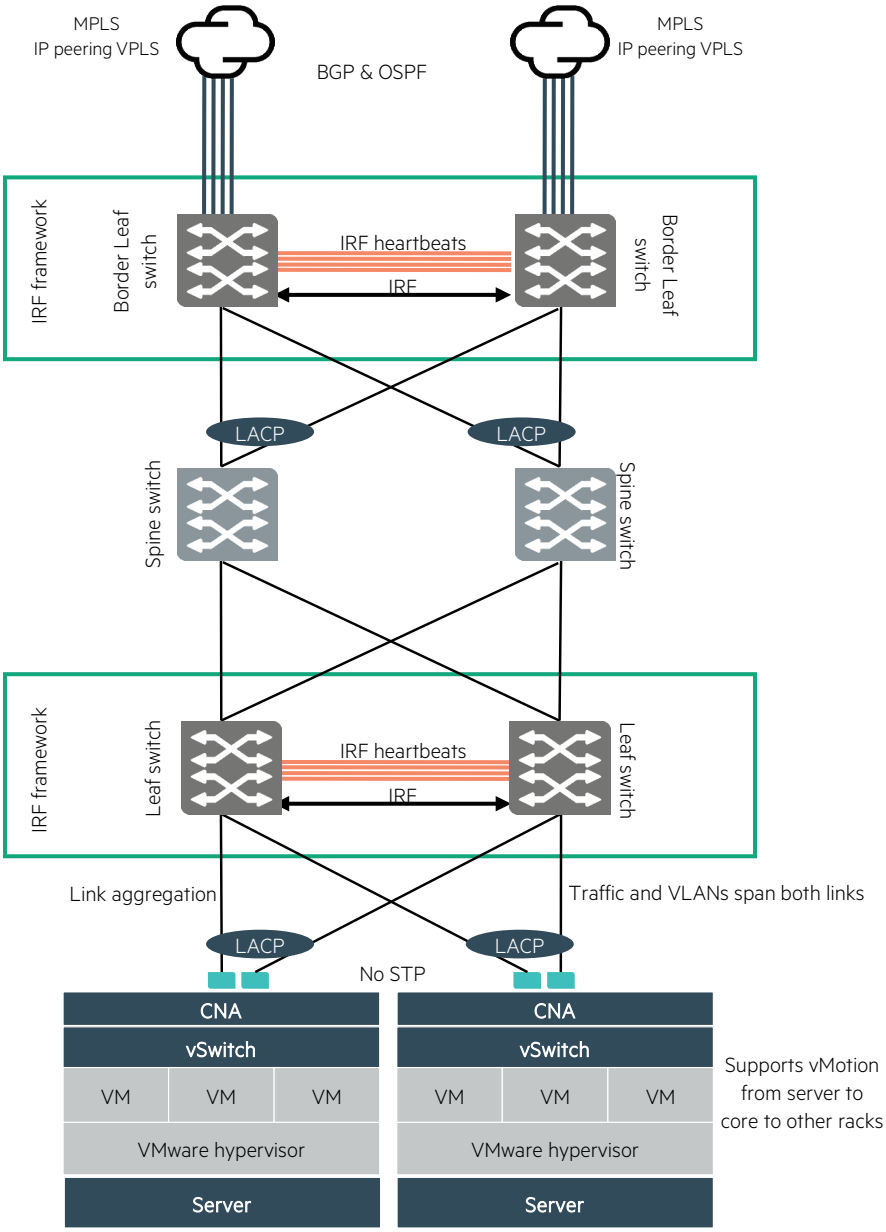


Figure 28. Spine and leaf topology logical

3-tier topology

A 3-tier topology is traditionally built using multiple layers that when drawn out look like a tree. These types of topologies usually consist of at least the following:

- Access layer:

  Usually ToR devices which connect to servers

- Aggregation layer:

  Provides aggregation connectivity between the access layer and the core layer

- Core layer:

  Usually provides the top layer routing services within and between data centers

Network design

Typical 3-tier topologies are similar to the 2-tier spine/leaf topologies, however, these solutions will have an added aggregation layer positioned in between the core and ToR layer. This aggregation layer will generally consist of chassis-based switches, and L3 routing typically occurs between the aggregation layer to the core and then the WAN. This way, L2 domains can be isolated to each aggregation layer device and the ToR switches that connect to it.

The type and number of switches deployed at each layer in these topologies will vary based on the requirements within each data center. Because these topologies aggregate large number of access devices into aggregation layer switches which then connect to core switches, the uplinks between switches can present bottleneck and oversubscription issues. It is imperative that these deployments take this into consideration and leverage large LAGs and high bandwidth uplinks between each layer.

Another common 3-tier design could be the use of PODs. In these scenarios each POD is essential, a 2-tier design with the third tier providing connectivity between the PODs as well as outside the DC. Designed properly, using this approach the oversubscription ratios can be kept lower since most of the traffic will stay within the POD.

As with the spine and leaf topology, redundancy is usually built into the network at the device level by leveraging HPE IRF and multi-chassis LAGs between the layers. However, redundancy can also be built in at the network level by designing multiple paths and leveraging L3 ECMP solutions.

3-tier topologies are able to scale to large sizes, which can provide 50,000 or more physical servers.

The following devices can be used as the core layer switches in this type of topology:

- HPE FlexFabric 12900E Switch Series

The following devices can be used as the aggregation layer switches in this type of topology:

- HPE FlexFabric 12900E Switch Series
- HPE FlexFabric 7900 Switch Series

The following devices can be used as the access layer switches in this type of topology:

- HPE FlexFabric 12900E Switch Series
- HPE FlexFabric 7900 Switch Series
- HPE 59xx Switch Series

Rack and logical diagrams

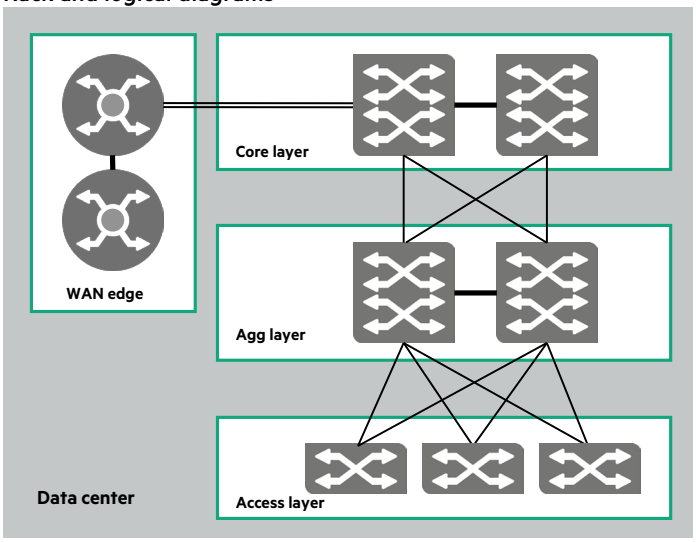


Figure 29. Legacy 3-tier topology rack view

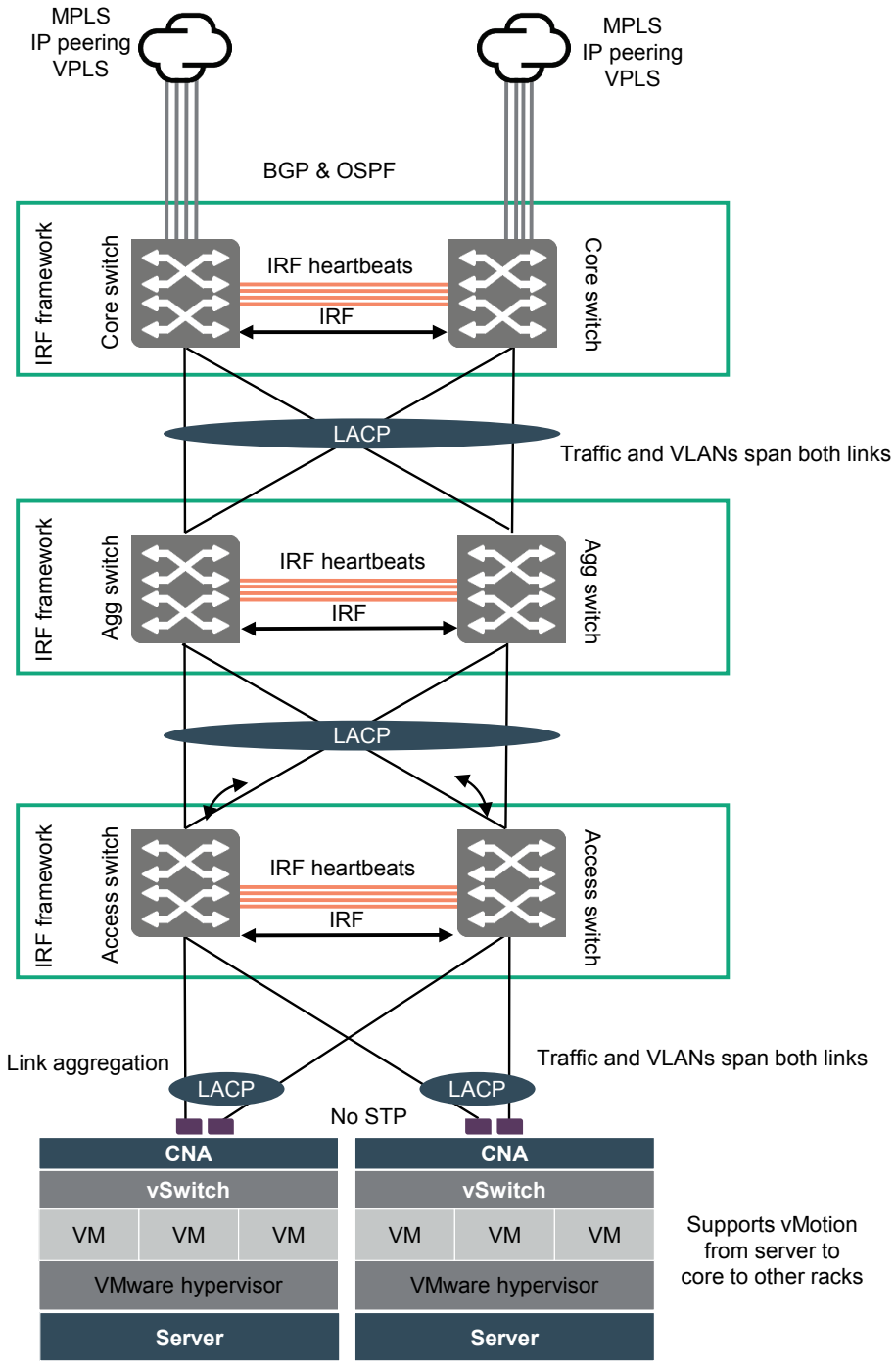


Figure 30. 3-tier logical view

Uplinks and downlinks in a data center

Today's 40GbE and 100GbE uplink speeds are providing a great benefit to data centers with regards to performance and scalability. However, when moving to 40/100GbE fibre uplinks a customer needs to rethink the cabling infrastructure. Current multi-mode optic standards for 40GbE and 100GbE use multiple 10Gbps lasers or multiple 25Gbps lasers simultaneously transmitting across multiple fiber strands to achieve the high data rates. Because of the multi-lane nature of these optics, they use a different style of fiber cabling, known as MPO or MTP cabling. For 40GbE, we can use 8-fiber or 12-fiber MPO/MTP cables. These MPO/MTP type cabling solutions can be used in existing data centers that are making incremental upgrades to the uplinks. However, upgrading an entire cabling plant can be cost prohibitive.

A preferred option for these existing data centers is to leverage 40GbE BiDi transceivers which allows the use of the same two MMF fiber strands with duplex LC connectors, currently used by 10GbE connections, to be used by the new 40GbE connections. In many cases the 40GbE BiDi optics are less costly than MPO/MTP optics and cabling combined. Using 40GbE BiDi optics in the existing data center means that migrating from 10GbE to 40GbE will be a smooth, cost-effective transition.

When building out new data centers from the ground up, however, customers will need to consider building out this new data center using single mode fibre rather than multi-mode fibre. It has been standard practice to build a data center using MMF which has, until recently, been appealing because lower cost solutions which are able to meet most distance requirements. However, new data centers are starting to see that moving to SMF will provide many benefits, including cost and performance. SMF solutions are able to:

- Leverage dual strand fibre:

  These solutions can leverage dual strand SMF fibers for 10, 40 and even 100GbE connections

- Network Taps:

  Many customers will tap optical fibers which means that there is one tap (two splitters for two fiber duplex communications) per connection. For new data centers that leverage SMF 40/100GE, this does not change. However, if this new data center was built using MMF, then for 40/100GbE a customer must plan for an entirely new cabling infrastructure because there are 8-20 fibers per connection, which means a single bidirectional, passive optical tap will require 8-20 splitters. In addition, a 40/100GbE data center will also have to ensure that all of the patch panels are MTP connectorized patch panels to terminate the associated 8-20 fiber cables, further increasing costs.

- Attenuation:

  Due to the higher bandwidth more attenuation is incurred resulting in smaller optical power budgets to accommodate splitter, splice and fiber losses. With 40/100GE loss budgets around 1.5dB, it is difficult to insert passive optics to monitor these links. If SMF is used, the link loss budget is substantially higher (distance dependent of course), so tapping these optical signals for monitoring is virtually the same as it is for 10GE today.

Access layer switching and server deployment considerations

Access layer switching

In current data center deployments, there are two access layer switching deployment models: ToR and end-of-row (EoR).

Top-of-rack

As the name implies, in a ToR placement, servers within a rack connect to an access layer switch generally placed at the top of the rack, as shown below ("ToR placement").

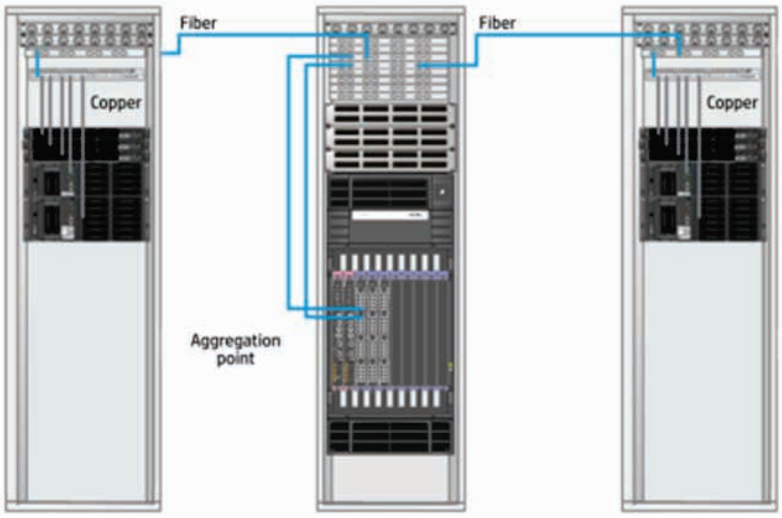


Figure 31. ToR placement

In a ToR design, servers connect to an access switch via copper or fiber within the rack, while the access switch connects to other consolidation or backbone switches within the data center via multimode fiber (MMF). Often, the data center backbone consists of high-capacity aggregation or distribution switches with L2/3 capabilities.

Even though today 10GbE copper connections are supported up to 100 m using Cat 6a/7 cables, Ethernet copper is generally restricted to relatively short runs inside the rack, while connections outside the rack can be made with smaller form factor multimode or singlemode fiber. This provides the ability to be connected to different capacity interfaces to support the bandwidth requirements of the rack.

Shorter copper Ethernet runs in the rack allow for multiple choices on cable types which can support various required speeds dictated by the systems in the rack. In many cases, server to switch connections can be up to 10GbE connections with support for integrated I/O.

Longer runs from the rack to core can utilize 10GbE, 40GbE or 100GbE MMF or SMF. HPE currently offers 40GbE interfaces on most HPE data center switches, and 100GbE interfaces for even higher bandwidth on the HPE FlexFabric 5950, 7900 and 12900E data center switch series.

Note

Current multi-mode optic standards for 40GbE and 100GbE use multiple 10Gbps lasers or multiple 25Gbps lasers simultaneously transmitting across multiple fiber strands to achieve the high data rates. Because of the multi-lane nature of these optics, they use a different style of fiber cabling, known as MPO or MTP cabling. For 40GbE, we can use 8-fiber or 12-fiber MPO/MTP cables. These MPO/MTP type cabling solutions can be used in existing data centers that are making incremental upgrades to the uplinks.

An option for existing data centers is to leverage 40GbE BiDi transceivers which allows the use of the same two MMF fiber strands with duplex LC connectors, currently used by 10GbE connections, to be used by the new 40GbE connections. In many cases the 40GbE BiDi optics are less costly than MPO/MTP optics and cabling combined. Using 40GbE BiDi optics in the existing data center means that migrating from 10GbE to 40GbE will be a smooth, cost-effective transition.

When building out new data centers from the ground up customers will need to consider building out this new data center using single mode fibre rather than multi-mode fibre. It has been standard practice to build a data center using MMF which has, until recently, been appealing because lower cost solutions which are able to meet most distance requirements. However, new data centers are starting to see that moving to SMF will provide many benefits, including cost and performance.

Advantages

- Issue isolation:

  Each rack, or group of racks using IRF, with a ToR configuration is treated as an independent module. Any issues or outages that occur with an access layer switch typically affect only the servers within that rack which are connected to that access switch.

- Traffic isolation:

  Because each access switch is a home run back to a backbone aggregation/core switch, traffic can be monitored and isolated to an identified switch port within a specific rack.

- Physical disasters:

  A potential physical disaster affecting cabling at the ToR will have adverse effects on the ToR rather than an entire row.

Disadvantages

- Number of switches:

  Each rack adds to the number of switches that need to be managed as independent entities.

- Additional rack-to-rack hops:

  ToR switches can introduce additional hops if used in conjunction with aggregation layer devices.

End-of-row

In an EoR placement, a rack containing the switching equipment is typically placed at either end of a row of cabinets or racks. Bundles of cabling provide the connectivity from each server rack to the switching equipment rack. Servers are usually connected to a patch panel inside each server rack. The copper or fiber bundles are home run to another patch panel in the rack containing the EoR switches. The switches are then connected to the patch panel via patch cables. EoR switches are typically connected to the core with a series of fiber patch cables.

EoR does not imply that the network racks have to be placed at the end of the row. There are situations where network switch racks are placed together in the middle of cabinet/rack rows. Placing the switch rack in the middle of a row limits the length of cables required to connect the furthest server racks to the nearest network rack.

Unlike the ToR model, where each rack is treated as an independent module, in the EoR placement model, each row is usually treated as an independent module.

Advantages

- Number of switches:

  There are fewer switches to manage as separate entities.

- Network preplanning:

  A data center can be preplanned and cabled without deploying a substantial number of switches into individual racks.

- Fewer rack-to-rack hops:

  Using EoR switches can reduce hops when EoR connects to core.

Disadvantages

- Issue isolation:

  A potential physical disaster affecting cabling or the EoR aggregation switch can have adverse effects on not just one rack of servers, but an entire row.

- Traffic isolation:

  Because each EoR switch is a home run back to a backbone aggregation/core switch, traffic will be for an entire row rather than just within a specific rack.

- Cabling:

  The magnitude of cabling home runs back to the EoR switch makes for substantial cabling bundles.

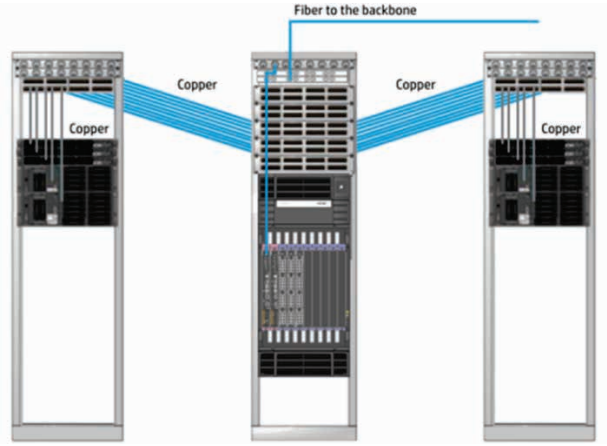


Figure 32. EoR

Server deployment considerations

Choosing the correct type of server to deploy is a critical decision IT departments regularly face. Primarily, servers come in two distinct form factors: rack servers and blade servers. There is no clear superior option between the two types. They are two distinctly different solutions that each have pros and cons.

Many factors can come into play when making these decisions. Primarily, however, these factors will focus on space, cost and power.

Rack servers

Rack servers vary in size, but are typically 1RU devices. Each containing their own power, cooling, expansion, I/O ports and processors.

Advantages

- Performance:

  A single server with its own dedicated CPU has direct uplink to network, rather than sharing uplink backplane in an enclosure.

- Simplicity:

  One physical system sharing common CPU(s), memory and storage controller.

- More expansion slots available for network and storage adapters:

  Rack servers can provide a number of slots for network interface cards (NICs) or storage controllers for load balancing (LB) and fault tolerance.

- Traditional servers have a greater internal capacity for local disk storage:

  If you're running several VMs on local disk storage, rack servers are a better choice, as they have more drive bays for internal disk storage.

- Serial, parallel and USB I/O ports:

  These ports allow you to connect external storage devices and optical drives. They can also accept hardware dongles used for licensing software.

- Lower cost:

  From a one-to-one hardware comparison, rack servers are lower cost. Rack servers can be advantageous for environments that have few total servers, with limited growth.

Disadvantages

- Green data center and long-term costs:

  These types of systems do not optimize the computing power per rack unit; they utilize more electrical power per rack unit, and add to the overall requirement for cooling systems. This added costs can, in the long run, negate the initial cheaper cost of the hardware.

- More space needed:

  When server deployments start to grow, rack servers will consume more space than blade servers. If space is limited, blade servers may be the better choice.

- Cabling:

  Cabling sprawl can become an issue for those deployments that choose to deploy multiple NICs on every rack server.

Blade servers/enclosures

Blade servers can reduce the external switch count and significantly reduce the number of cables in a rack. The blade enclosure can include server-to-network interconnect modules, which can replace the access layer switch at either the ToR or EoR placements.

Advantages

- Data center space:

  Blade servers can increase rack density. Compared with traditional servers, this means up to 50 percent more servers in the same amount of rack space.

- Green data centers consume less power:

  They are more energy efficient and require less cooling. A fully-loaded chassis will consume far less power than an equivalent amount of traditional servers. If the chassis is not full, it will still consume less power than an equal amount of rack servers.

- Integrated Ethernet switches:

  Ethernet cables plug into a chassis with a single connector, which reduces cabling costs and makes cabling neater and eliminates the clutter of cables common with rack servers.

- Flexible SAN deployments:

  Blade server/enclosures provide a variety of options with regards to both LAN/SAN convergence, as well as direct SAN solutions where embedded blades are able to replace Fibre Channel edge switches. This reduces the number of physical connections at the server edge to both the network and storage media.

- Centralized management:

  From a single user interface, administrators can monitor and manage all of the blades.

Disadvantages

- Local expansion slots:

  Relatively limited number of expansion slots for storage and/or internal blade NICs.

- Expansion cost:

  Blade server chassis have substantial capacity, but once full, the next incremental CPU blade requires another chassis and the additional integrated chassis components.

- Space:

  Although an advantage to blade servers is that they take up less rack space; that is only true when typically more than 11 blade servers are used in an enclosure. Sizing correctly and planning for future growth is a critical step when making these decisions.

- Vendor lock-in:

  When choosing to expand and add another server blade, customers will be locked into the same vendor for the life of the chassis.

SAN convergence

Virtualization of data, web-based and multimedia applications have driven the requirement for new servers, new platforms and associated network storage. This requirement has resulted in unplanned growth in server complexity. IT departments are being charged with the improvement of computing efficiencies in their data centers. The response has been server consolidation utilizing virtualization technologies implemented on both blade and rack servers.

In a traditional server environment, each server is connected to two networks: The Ethernet general purpose LAN and the SAN. The Ethernet network has the following features:

- A single set of servers or enclosures may connect to ToR switches (typically deployed in pairs for redundancy) which uplink to a redundant core or aggregation network layer.
- Several servers or enclosures may connect to EoR switches, which consolidate the network connections at the edge before going into the core or aggregation network layer.
- Flatter, simplified deployments may connect the servers directly to the core layer, eliminating the ToR altogether.

These types of networks have different interconnect implementations. Each server or enclosure needs to have one set of interconnects to the LAN and one set to the SAN. In addition, each server or enclosure has its own power, cooling and management requirements.

However, data centers are now moving away from these traditional dual LAN/SAN network deployment and towards deployments, which converge the dual networks into one single network. A converged network is able to transmit the Ethernet and SAN traffic over a single cable while leveraging the DCB set of protocols, which ensure the SAN traffic remains completely lossless. These converged networks are able to reduce costs by reducing the number of networking devices, cabling and server NICs/HBAs.

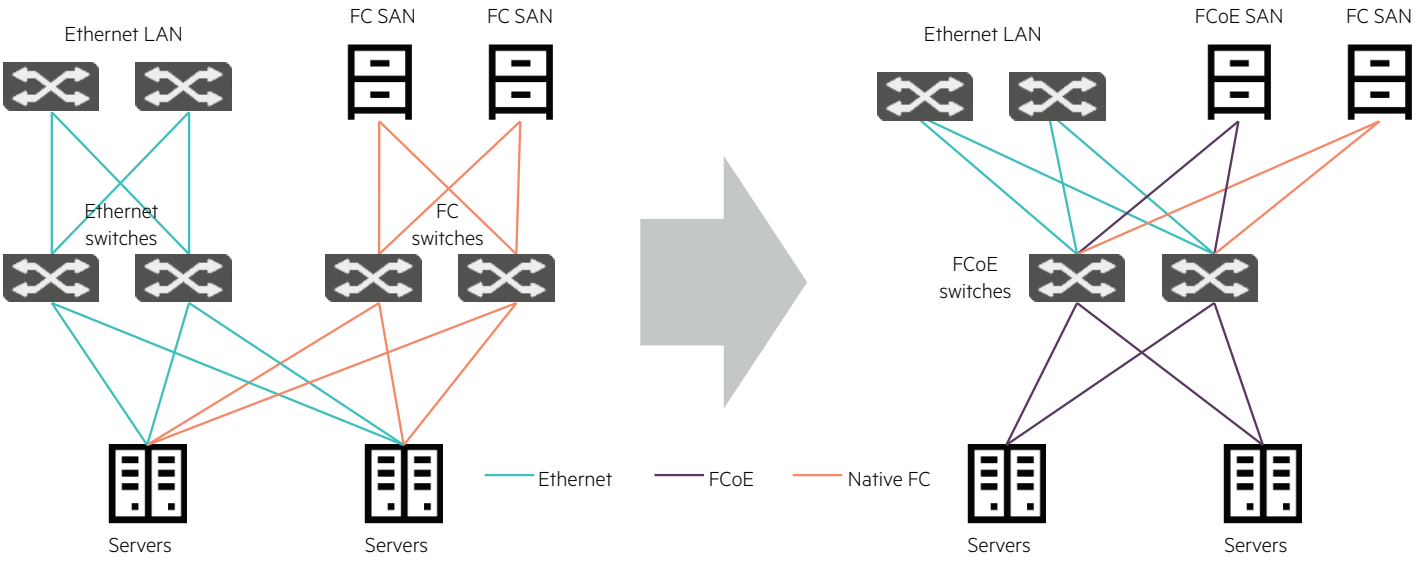


Figure 33. Converged network

HPE Networking solutions are able to converge the LAN and SAN into a single network, by leveraging DCB protocols ensuring lossless SAN traffic, as well as by leveraging higher bandwidth connections available on most HPE data center switches.

Within a blade enclosure, administrators can leverage the HPE 6125/6127XLG Switch, which provides the ability to converge LAN and SAN traffic within the enclosure while also leveraging the same Comware OS used on the HPE data center switches.

Administrators also have the option to use HPE VC modules, which also are able to converge LAN and SAN traffic. In addition, they can use Flex-10 technology to allocate one 10GbE server port into four individual Ethernet connections or functions. Each connection can be adjusted dynamically to meet the demands of the workload. The VC FlexFabric Module takes this a step further by allowing one connection per port to be used as either a Fibre Channel or iSCSI function. Additionally, HPE VC solutions are able to further simplify the solution by supporting direct connection to HPE 3PAR storage, eliminating the SAN network.

Use cases

To simplify the understanding and implementation of FC/FCoE configurations, HPE has validated a set of use case topology designs. The use cases describe the recommended ways to use the HPN products, switch modes and port types in different server-storage deployment scenarios. Some of the use cases show multiple types of connectivity within the same configuration. This is meant to show the different connection options that are available.

Below are three major use case scenarios, each describing a different implementation based on the hardware being utilized for servers and storage. Use case 1 shows a single layer or single tier of HPE FlexFabric 59xx switches connected to rack servers, while use case 2 shows blade servers using Interconnect modules connected to HPE FlexFabric 59xx switches, and use case 3 shows fabric connectivity on the storage side to existing FC/FCoE switch series fabrics via HPE FlexFabric 59xx switches running in NPV gateway mode. Each use case has A, B and C variations available as detailed in the table below. The solutions are validated with HPE 59xx, 7900 and 12900E Switch Series. If the use case requires native FC ports, then the solution would require either HPE FlexFabric 5900CP switches or the modular versions of the HPE FlexFabric 5930 or 5940 with CP modules.

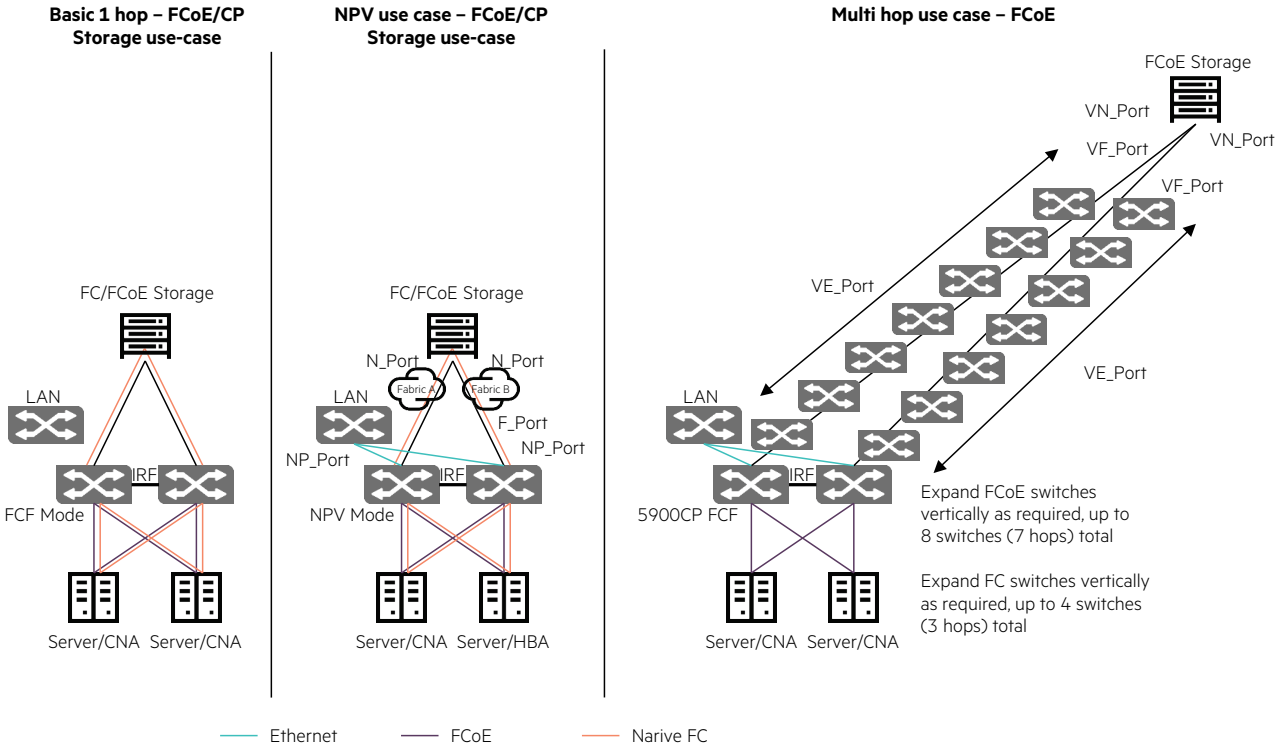


Figure 34. Major HPE Networking FC/FCoE use cases

Table 2. HPE FlexFabric converged use-cases

MAJOR SAN FABRIC USE CASE	VARIANT	SERVER CONNECT	STORAGE CONNECT	SWITCH/SWITCH MODE
1: Single-tier fabric, rack servers	1A: Rack server>FCoE storage	Rack/CNA	Native FCoE	FCF—HPE FlexFabric 5900CP/5900AF/Modular 5930 or 5940/Fixed 5930/5940/7900/12900E
	1B: Rack server>FC storage	Rack/CNA/HBA	Native FC	FCF—HPE FlexFabric 5900CP/Modular 5930 or 5940
2: Single-tier fabric, BladeSystem	2A: BladeSystem > 59xx > FCoE storage	Blade/VC/6125 or 6127XLG	Native FCoE	FCF—HPE FlexFabric 5900CP/5900AF/Modular 5930 or 5940/Fixed 5930/5940/7900/12900E
	2B: BladeSystem > 59xx >FC storage		Native FC	FCF—HPE FlexFabric 5900CP/Modular 5930 or 5940
3: Single-tier fabric, NPV gateway	3A: Rack server>FC NPV gateway>FC storage	Rack/CNA/HBA	FC via B/C/H FC switch	NPV—HPE FlexFabric 5900CP/Modular 5930 or 5940
	3B: BladeSystem>FC NPV gateway>FC storage	Rack/VC		
	3C: Rack/BladeSystem>FCoE NPV gateway	Rack/CNA/HBA Blade/VC/6125XLG/FC	FC/FCoE via existing SAN	NPV—HPE FlexFabric 5900CP/5900AF/Modular 5930 or 5940/Fixed 5930/5940/7900/12900E
4: Multihop Fabric, Rack Servers	4A: Rack server>FCoE storage	Rack/CNA	Native FCoE	FCF—HPE FlexFabric 5900CP/5900AF/Modular 5930 or 5940/Fixed 5930/5940/7900/12900E
	4B: Rack server>FC storage	Rack/CNA/HBA	Native FC	FCF—HPE FlexFabric 5900CP/Modular 5930 or 5940

MAJOR SAN FABRIC USE CASE	VARIANT	SERVER CONNECT	STORAGE CONNECT	SWITCH/SWITCH MODE
5: Multihop Fabric, Blade Servers	5A: BladeSystem > 5900 > FCoE storage	Rack/Blade/VC/6125XLG	Native FCoE—7 hops	FCF—HPE FlexFabric 5900CP/5900AF/Modular 5930 or 5940/Fixed 5930/5940/7900/12900E
	5B: BladeSystem > 5900 >FC storage	Rack/Blade/VC/6125XLG	Native FC—3 hops	FCF—HPE FlexFabric 5900CP/Modular 5930 or 5940
6: Leaf-Spine fabric	Leaf-Spine using 79xx or 129xxE with 5700/5900CP/5930/5940 Modular switches	Rack/CNA/HBA Blade/VC/FC	Native FCoE/FC	FCF—HPE FlexFabric 5900CP/5900AF/Modular 5930 or 5940/Fixed 5930/5940/7900/12900E
#7: Edge-Core-Edge fabric	Edge-Core-Edge using 79xx or 129xx with 5700/5900CP/5930 Modular switches	Rack/CNA/HBA Blade/VC/FC	Native FCoE/FC	FCF—HPE FlexFabric 5900CP/5900AF/Modular 5930 or 5940/Fixed 5930/5940/7900/12900E
#8 Multi-hop fabric, NPV (EoR-Spine) gateway	For all FlexFabric 59xx NPV mode configurations, HPE supports a maximum of seven hops between any two devices in FCoE fabric and the legacy FC fabric connected via NPV/NPIV. FC solutions have been qualified for three hops with four switches between server and storage.	Rack/CNA/HBA Blade/VC/FC	Native FCoE/FC	NPV—HPE FlexFabric 5900CP/5900AF/Modular 5930 or 5940/Fixed 5930/5940/7900/12900E
#9 Dual-mode configurations	9: Dual-mode configurations allow for the setting of a different switch mode for individual VSANs (FCF or NPV). This applies to the 5900CP, 5930 and 5940 switch models	Rack/CNA/HBA Blade/VC/FC	Native FCoE/FC	FCF and NPV

HPE has validated multihop FC solutions which have been qualified for three hops with four switches between server and storage. FCoE solutions have been qualified for seven hops with eight switches between server and storage.

HPE Networking has also validated support for 10 Km 10/40GbE FCoE as well as 10 Km FC solutions. HPE 3PAR Remote Copy testing has been used to validate these solutions and ensure reliability.

Figure 35. Remote copy validation

Many environments are happily using iSCSI. In fact, iSCSI serves the same purpose as Fibre Channel in building SANs, but iSCSI avoids the cost, complexity and compatibility issues associated with Fibre Channel SANs. Unlike native FC, iSCSI can be implemented over 1GbE, 10GbE and 40GbE TCP/IP networks. Thanks to adoption of the DCB protocol suite lossless iSCSI is also gaining in popularity.

There are a set of minimum switch capabilities that make building a high performance fault-tolerant storage network a relatively easy and cost-effective task. As a rule of thumb, any Enterprise-class managed switch typically has the necessary capabilities most iSCSI SAN customers require. HPE recommends the HPE FlexFabric 59xx for ToR convergence, and the HPE FlexFabric 12900E/7900 Switch Series for core converged switches.

For more details on HPE FC/FCoE/iSCSI solutions, refer to the resources available at hpe.com/go/flexfabric as well as hpe.com/go/storage.

HPE Data Center Interconnect—connecting geographically dispersed data centers

As IT technology has been maturing, many organizations have been reducing costs and increasing efficiency by following the trend to consolidate their data centers. This consolidation has been driving organizations to heavily leverage virtualization in the data center as well as private clouds and multitenant environments, which support different business units. This allows organizations to increase their availability and help them to move and/or deploy new workloads on resources that can best serve them.

These advancements mean that connecting geographically dispersed data centers is now more important than ever. Geographic data center interconnections allow the IT designer to put in place disaster avoidance and disaster-recovery mechanisms that increase the availability of the applications. Geographic dispersion also enables shared resource utilization, optimization of application response, and allows the flexible mobility of workloads and services.

Unfortunately, most interconnect methods suffer from limitations including transport dependency, complexity and lack of resiliency. HPE DCI solutions are designed to address these limitations by delivering responsive, efficient, and resilient data center interconnect solutions.

HPE DCI features and benefits include:

- **High availability:**

  The DCI solution establishes active/active data centers. Along with HPE FlexFabric product portfolio, DCI establishes reliable links with link aggregation and failover capabilities. This gives users uninterrupted access to applications in the event of disruption of service at one data center.

- **Disaster Recovery and Data Replication:**

  DCI enables disaster recovery where the application workloads can be migrated in between different data centers. Data can also be replicated across data centers using DCI. In the event of failure, DCI provides the data recovery path from remote data centers and facilitates business continuity.

- **Migration path for DC consolidation as well as path towards the cloud:**

  DCI solution enables the customer to connect existing Data Center to new construction site for consolidation so that customer can migrate application toward the destination site, either in new data center sites or in the cloud environment.

- **Mobile workloads and long distance workload migration:**

  DCI establishes reliable connections between data centers that serve as a platform for fast and reliable vMotion between distant data centers.

- **Multitenant enabled L2 extension solution:**

  The HPE EVI solution can leverage HPE MDC technology to partition a single device into multiple logical devices, which provides up to 75 percent reduction of the number of physical platforms leading to CAPEX and OPEX reductions.

- **Bursty and seasonal traffic patterns:**

  Organizations could load balance heavy or bursty workloads among geographically dispersed data centers making more effective use of data center resources.

- **Investment protection:**

  DCI solutions can be customized to fit the customer's exact environment, from IP to MPLS to DWDM. DCI solutions can work with customer's existing IP networking infrastructure without requiring changes. This protects networking investments and allows easy and seamless deployment.

There are many different technologies that can be used for Data Center Interconnect. The network resources of a user between data centers determines the solution to be used as follows:

- **HPE Ethernet Virtual Interconnect:**

  This IP-based solution option is extremely useful in simplifying DCI. Transmission between data centers over DWDM or MPLS can be complex to manage and is often highly dependent upon costly, dedicated and rigid service provider's infrastructures. In contrast, EVI runs over any IP infrastructure so it can be deployed without requiring changes to an existing infrastructure. This characteristic simplifies deployment by allowing Layer 2 connectivity across the network without having to deal with Layer 3 networking dependencies.

- **VXLAN:**

  VXLAN has also become widely supported as a multitenant solution allowing 4K VLANs for up to 4K tenants. Since VXLAN interconnects all the VETPs over an IP infrastructure, each VXLAN gateway can be enabled as a VTEP connecting all the servers or virtual workloads over Layer 2 network prospective. VXLAN VTEPs can be very efficient in enabling workload migration within the data center with overlay network such as EVPN, similar network device VXLAN VTEPs can be also used for the interconnection between data centers to connected selected networks or VLANs for DR, HA or application migration purpose.

- **Generic Ethernet LAN extension:**

  This option extends Ethernet natively over a dark fiber or DWDM optical transport. As such, this solution mostly applies to point-to-point deployments, where the sites are connected via dedicated dark fiber links or DWDM optical circuits.

- **MPLS point-to-point or multipoint using MPLS or VPLS:**

  This option uses MPLS technologies to provide L2 connectivity services over a L3 network service. Depending on the nature of the transport infrastructure between data center sites and the number of data center sites to be interconnected, different technologies can address the connectivity requirements.

- **PBB/SPB:**

  PBB/SPB is another alternative way to implement extended Layer 2 networking over existing Layer 2 infrastructure. The routing based control plane can remove any potential loop without requiring legacy STP to be involved across the entire network across different data centers.

Key considerations for DCI design

Transport Agnostic

When interconnecting geographically dispersed Data Centers, the main issue to take into consideration is how to interconnect those sites. Typically those data center sites are separated by a few kilometers or miles to thousands, and there are a variety of technologies or service offerings that can provide the wide area connections, such as DWDM, dark fiber, metro Ethernet to service provider offering of leased line. The ideal way to implement DCI is to leverage IP based technology where the Layer 3 protocol can operate on top of any existing infrastructure.

High availability, STP isolation and loop management

In the perspective of VM resource scheduling and remote cluster access, multiple interconnected data centers can be considered as a logical large-size data center. The links interconnecting data centers can be considered as the DCI backbone links of the large-size data center. More importantly, the backbone links interconnecting data centers transmit control signaling, in addition to vMotion or data packets. Therefore, once the links interconnecting the data centers fail, the large-size data center fails to work properly, and causes service interruption for users.

Therefore, a key consideration of L2 DCI is to improve availability. The best way of improving HA is to design multiple DCI links and IRF (DCI devices). To improve HA and increase the interconnecting bandwidth at the same time, you can design load-sharing interconnection links, so that you can increase the bandwidth and enable the services to rapidly converge when the system encounters errors, thus improving HA. New generation HPE FlexFabric switches such as the HPE FlexFabric 12900E/7900 can also support automatic EVI link load-balancing without requiring extra configuration steps.

For example, whenever possible, HPE recommends using IRF technology on the DCI devices. Aggregating two or more links not only simplifies the DCI topology but provides HA and achieves load sharing. Use LACP to aggregate two or more links between the dedicated DCI devices into a logical link, so that the DCI topology is greatly simplified. At the same time, the bandwidth of the two HA links is optimized, and load sharing is achieved.

When you deploy STP in an L2 DCI network, HPE recommends that you isolate STP domains to each data center. On each port connected to the DCI link, enable BPDU drop (Edge port) and disable STP.

Additionally, end-to-end loop management needs to be addressed to manage the loops across multiple interconnected L2 domains. By leveraging control plane based Layer 2 routing such as EVI, STP domains will be automatically isolated at each data center ensuring no loops exist. Of course this could be done by utilizing only one single link between data centers, and manually disabling STP from each site, but a better solution would be to use IRF and LACP with automatic loop prevention.

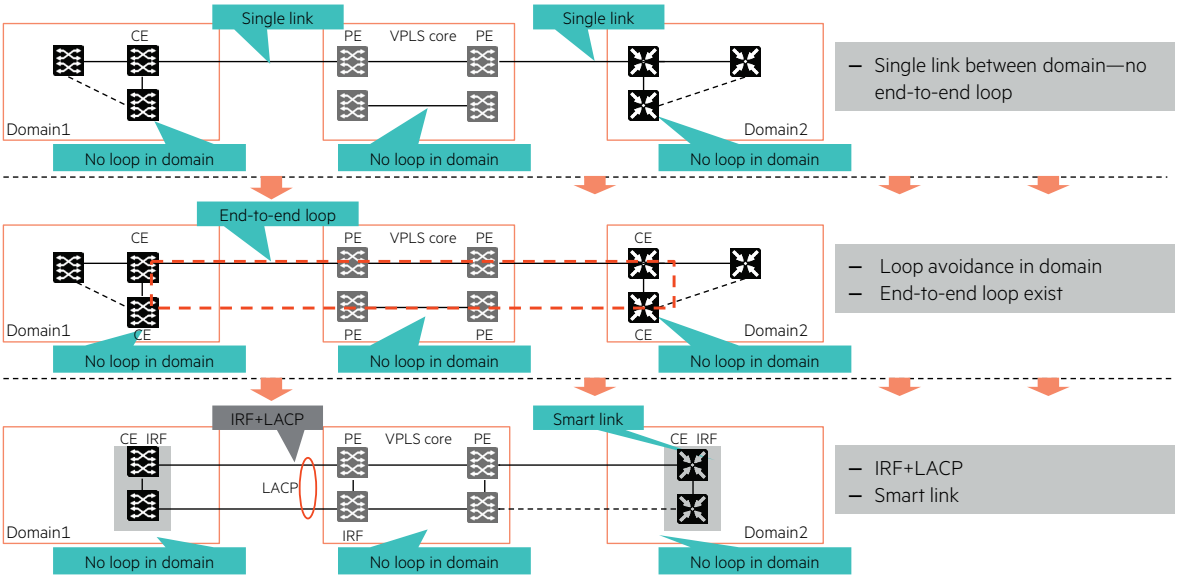


Figure 36. End-to-end loop management design

Path optimization for DCI

Data centers are becoming highly virtualized with East to West virtual machines communications reaching up to 75 percent of the overall data center traffic. Layer 3 gateways have become a critical aggregation point of data center traffic between different VLANs, but having a simple Layer 2 network extend to other data centers will aggregate all inter-VLAN traffic to a specific data center which results in inefficient trombone routing.

Distributed Layer 3 gateways can address this issue by providing a local Layer 3 gateway in each data center with optimization for virtual machine migration using VRRP. Having the East-West traffic across different VLANs routed locally can significantly remove trombone routing between the Layer 3 gateways in between different data centers, especially for an active/active data center design.

Imagine when a large volume of data traffic passes through the DCI core network links, the data traffic consumes heavy bandwidth and degrades the quality of transmitted control data. This causes a loss in control data packets and affects the migration or disaster recovery process. The key to data path optimization is drawing the response traffic of servers away from the original DCI core network path and towards the data center local gateway.

To remove trombone routing, you can configure an identical VRRP group on each data center gateway, as shown in figure 37. EVI based DCI solutions can significantly simplify and automate the path optimization by allowing each data center to have active, independent and identical VRRP gateways providing Layer 3 routing for local workloads as well as for workloads that are being migrated from other data centers.

Inbound path optimization can also be achieved with different technologies where the focus is to route the client traffic to the exact data center based on the application location information. Since the workload can be migrated between different data centers, application location information can be very dynamic. It is essential for the client side to keep polling the central data base for such location information. To be able to provide a central location data base, a global DNS can be leveraged to provide IP address to domain name resolution where the IP address is used to represent the different data centers. When the application is migrated to a different data center the local traffic load-balancer IP address can be used to identify the application as well as the location. F5 is one of the HPE Alliance partners that offer global DNS service and it has been validated with HPE Networking DCI solutions for active/active data center implementation.

Alternatively Locator/ID Separation Protocol (LISP) is another way to provide a central location data base (a MAP server). IP addresses are designed not only to be routable, but also they also can be embedded with location semantics of the network segment. LISP implementation can leverage a separate set of IP addresses to provide the location information between the client site and the data center. LISP is currently under development and being implemented in the Comware based router platforms to enable client path optimization.

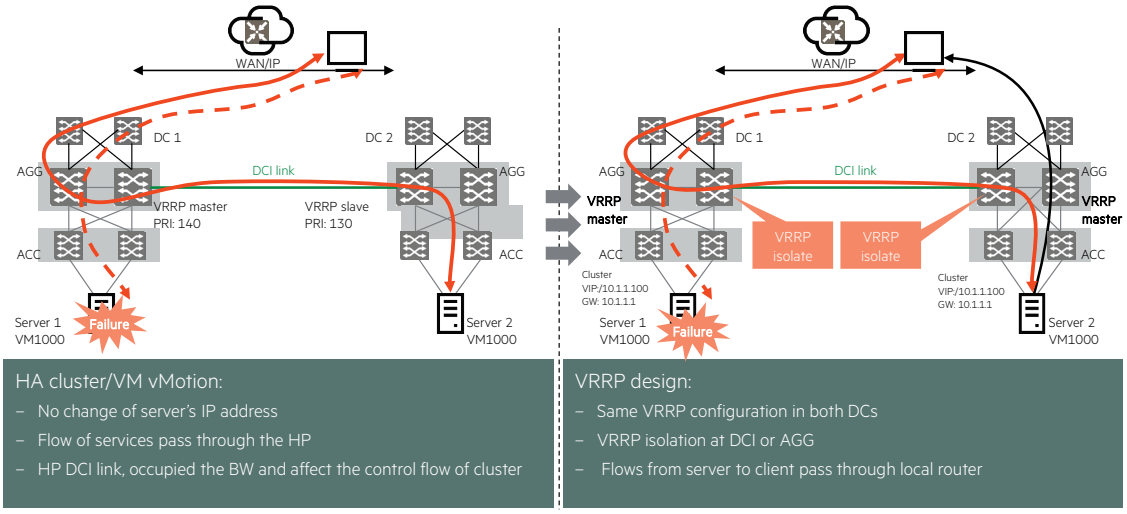


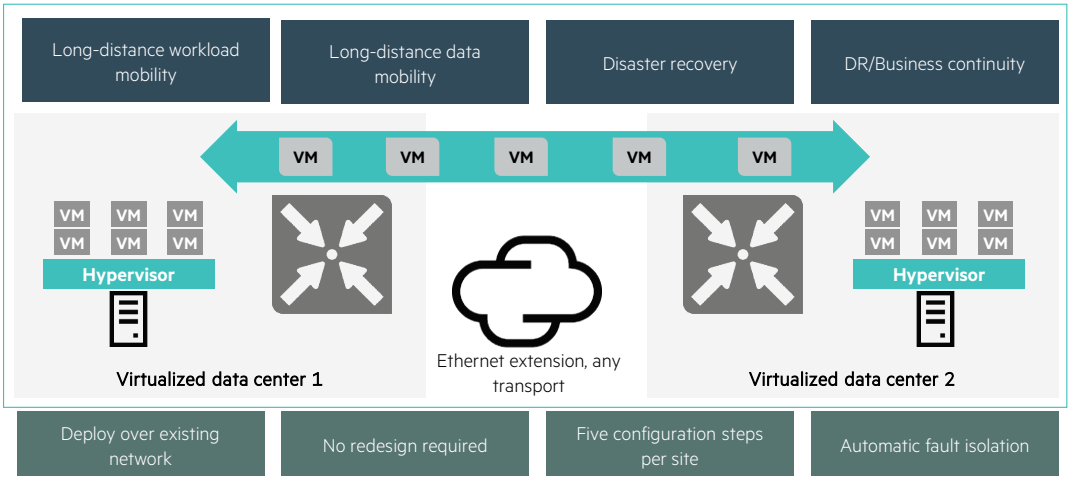
Figure 37. Data path optimization

HPE Ethernet Virtual Interconnect

HPE EVI runs over any Internet Protocol transport and extends L2 domains across a WAN network, typically between data centers. By virtualizing and automating the link layer domain across data centers, EVI delivers the elements necessary to enable an SDN data center infrastructure. It enables several data centers to work as one that is more responsive, with higher efficiency and solid high availability for business resiliency.

With HPE EVI, enterprises are able to:

- Accelerate the delivery workload mobility with remote vMotion
- Increase application performance with multipathing and load balancing
- Allow organizations to scale up to 16 geographically dispersed data centers without requiring them to change the underlying network
- Simplify the L2 connection by encapsulating traffic over GRE and automatically isolate Spanning Tree Protocol
- Achieve optimum degrees of high availability and disaster recovery for valuable data
- Allows clients to have a simple set of L2 routing extensions that can provide data interconnectivity in minutes rather than the months of legacy approaches like VPLS



When used along with HPE IRF switch virtualization technology, HPE EVI delivers greatly enhanced reliability, resilience, and faster remote vMotion capabilities. The combination of EVI and MDC brings multitenancy to cloud-ready and remotely connected data centers.

HPE EVI is a L2 routing technology that uses EVI Links and GRE tunnels to extend VLANs across up to 16 to 32 locations. Each EVI network has a unique network ID, extends a unique list of VLANs, and has separate control and forwarding planes. When used in conjunction with MDC, each MDC can support 32 EVI networks with each MDC running up to 4K VLANs.

EVI features

The following features are used to optimize the control plane traffic thus increasing efficiency.

- **EVI VLAN mapping:**

  When a data center is initially deployed, not all the vlan numbers are planned consistently. EVI support mapping the vlan number of one Data Center to another vlan number in a different Data Center so that customer may connect different VLAN from different DC for the application backup or HA or even for the clustering.

- **Access port support:**

  Normally EVI is designed to interconnect different VLAN across data center, but EVI can also be used to interconnect un-tagged access VLAN on 12900E FX series LPU or 7900 FX.

- **Selective MAC routing:**

  This feature stops unknown unicast and multicast frames from flooding the EVI links. The internal interfaces are capable of flooding to the internal interfaces while the EVI-tunnel interfaces will drop these frames. Similar to selective flooding, it is possible to permit or deny a MAC route.

- **Automatic loop prevention:**

  EVI has the following loop prevention mechanisms built in and enabled by default on both the data and control planes:

  – EVI-Split Horizon:

    This is enabled on the data plane to prevent frames received from EVI tunnels from being forwarded to the transport layer (EVI Links). Its primary function is to prevent loops among EDs.

  – STP Domain Boundary:

    This disables STP on the EVI Links. STP domains and BPDUs are not extended across sites keeping topology changes local. BPDUs are blocked from one location to another so that STP changes are contained within a site. This also allows each site to run different versions of STP. The following versions of STP are supported 802.1d, 802.1s and 802.1w.

- **Selective Flooding:**

  By default, unknown unicast and multicast are dropped at the EVI Links. If an application uses a special MAC address for traffic identification, it would break. Selective flooding enables the ED to flood frames with a certain unknown destination MAC to an EVI tunnel interface.

  An example of this would be Microsoft Network Load Balancer (NLB) that uses a special MAC address (cluster MAC) to identify a cluster. If cluster members are located in multiple sites, selective flooding would be used to propagate the cluster MAC address on the EVI-Tunnel interface(s).

- **Automatic VRRP Isolation:**

  To provide routing path optimization, EVI enables VRRP isolation by default. This allows each DC to have an active Layer 3 gateway that leverages the VRRP protocol to hide the specific HW details (IP and MAC addresses). All data centers will run separate sets of VRRP instances. VRRP isolation stops the population of VRRP keep-a-lives of each VLAN over EVI Links, so that each data center always has active/active gateways.

- **ARP Flooding Suppression:**

  This feature will reduce the number of broadcasts that traverse the EVI network. When an ARP request is made and initially flooded across the EVI network, the EVI ED listens to ARP responses on the EVI Link. These ARPs are cached for the remote MACs so that subsequent ARP requests can be handled directly by the ED.

EVI with IRF

At each location, HA is achieved by configuring IRF on the switches to simplify the network topology and increase uptime. IRF can also be configured in conjunction with MDC.

EVI and MDC

EVI and MDC work together to provide a customer DCI solution where EVI is designed to establish connectionless tunnels between data centers.

MDCs can be configured with IRF in an EVI environment to create completely secure isolation between tenants.

Up to 32 EVI networks can be deployed in each MDC and each MDC also has a completely functional L2 and L3 environment on HPE FlexFabric 12500 based platforms.

EVI networks should be deployed on admin MDC or chassis when using HPE FlexFabric 7900/12900E based platforms.

When active/active VRRP gateway is deployed across different data center, the VRRP gateway needs to be created on a separate MDC than where EVI is configured and deployed. This is required for 7900 and 12900E based implementation.

Refer to the hpe.com/networking/dci for more details.

EVI general product positioning guideline

EVI is currently supported on the HPE Comware 7 based HPE FlexFabric 12500/12900/7900 Switch Series, and also on HPE HSR, MSR, and VSR routers.

Design considerations should be taken when positioning these switches and/or routers in the DCI solution.

HPE FlexFabric 12500 based EVI implementation can form direct EVI connections with all 12500 LPUs as well as all Comware v7 based router implementations.

HPE FlexFabric 7900/12900E based EVI implementation can form direct EVI connections with all 7900/12900E LPUs as well as all Comware v7 based router implementations.

Comware v7 based HPE MSR/HSR/VSR routers support EVI with both HPE FlexFabric 12500 and 7900/12900E based platforms. In DCI solutions between HPE FlexFabric 12500 switches and HPE FlexFabric 7900/12900E switches, a router can be positioned at either site as the EVI device.

Comware v7 based HPE MSR/HSR/VSR routers can be positioned as lower cost high performance EVI devices for DCI solutions, they also provide secured DCI solutions with IPsec encryption. IPsec can provide automatic packet fragmentation and defragmentation for the DCI solution to free up customers from service provider limitations.

EVI is currently not supported on 1RU or 2RU top of rack switches.

VPLS

VPLS controls packet forwarding by using two layers of labels and can implement point-to-multipoint DCI connections. With VPLS, you can simulate an Ethernet switch among multiple data centers in the MPLS network to make inter-L2 port forwarding decisions based on MAC address or the combination of MAC address + VLAN ID. Data center aggregation layer switches directly communicate with the other aggregation layer switches associated with the VPLS instance.

Such a design is relevant in these two scenarios:

- Enterprise customer owns or manages their own Layer 1 and VPLS network
- Enterprise customers acquire a L1 or L2 type of service from a provider, and VPLS is run between the enterprise PE devices at the edge of the provider's cloud

Interconnecting multiple data centers using VPLS + smart link

For those VPLS scenarios where IRF is not supported, you can deploy VPLS + smart link solutions to improve HA performance.

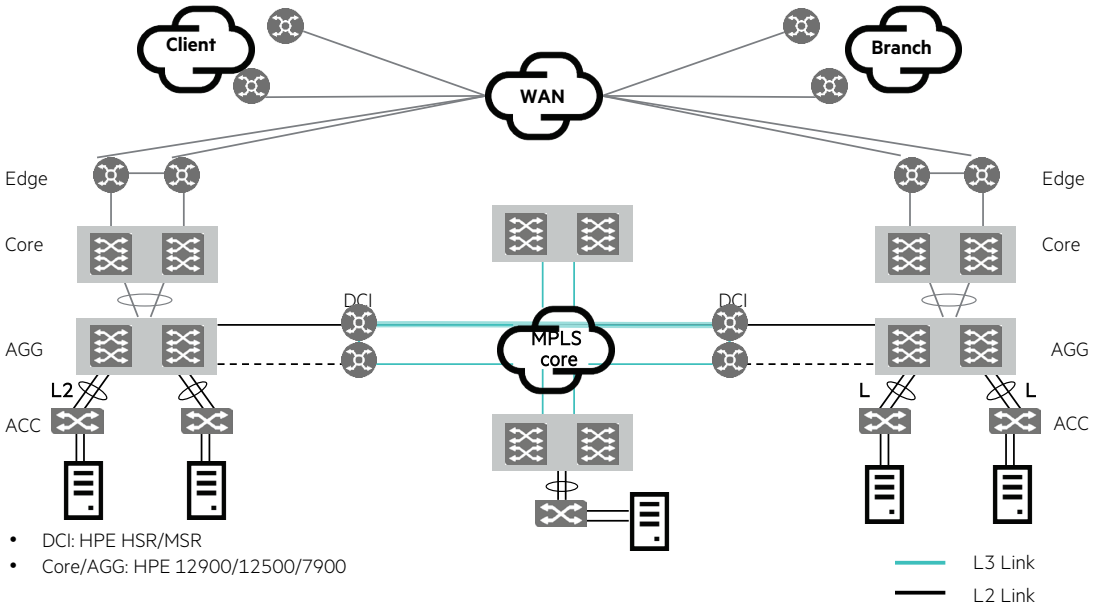


Figure 38. VPLS + smart link solution

Interconnecting multiple data centers using VPLS + IRF + LACP

Suppose a VPLS network provides the following conditions:

- The MPLS network supports Ethernet interfaces. Therefore, you can use dedicated IRF-enabled switches (could be viewed as customer-owned PE devices) for connecting to the MPLS network.
- The users need dual-homing aggregation layer devices to improve the HA performance.

In this case, HPE recommends that you use the VPLS + IRF+ LACP solution.

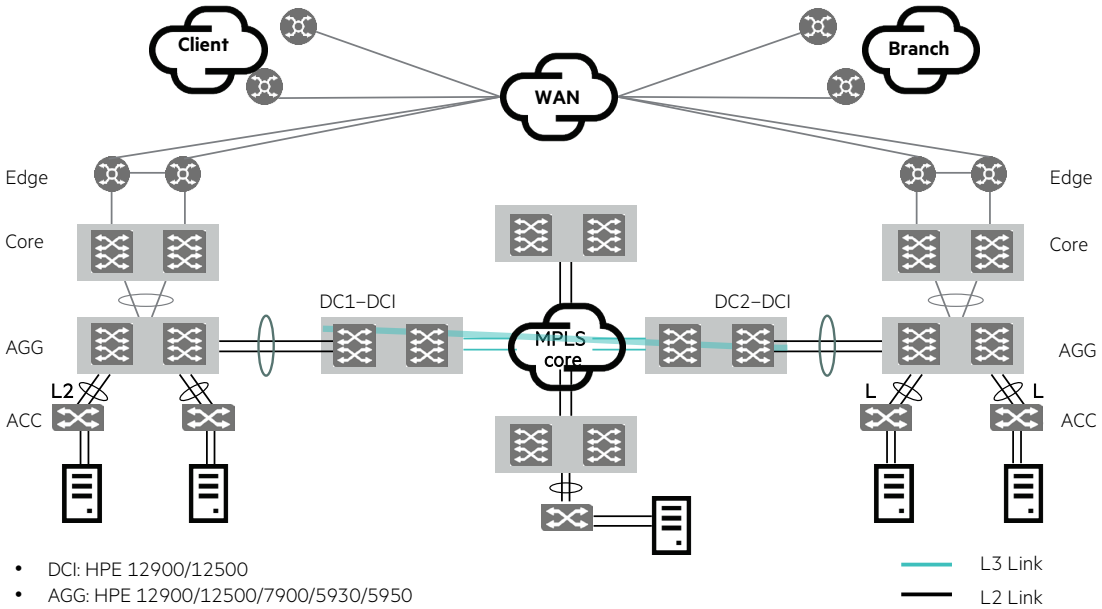


Figure 39. VPLS + IRF + LACP

The VPLS + IRF solution delivers the following benefits:

- Per-layer IRF design, which has a simple architecture and is easy to maintain.
- Per-layer IRF design, which provides high HA performance and implements millisecond-level convergence for the failure of any node or any link.
- Implementing IRF on the dedicated DCI devices decreases the number of nodes in the MPLS network and the number of broadcast packets.
- Per-layer IRF design implements load sharing for all links, improves the bandwidth utilization, and improves the performance of the whole network.

Dark fiber or DWDM

Dark fiber or DWDM DCI: Point-to-point for two data centers

When two data centers are directly connected through dark fiber or DWDM you can use pairs of dedicated DCI devices at each data center utilizing IRF and LACP to create the DCI connection. To facilitate expansion to more than two data centers you can use alternate solutions listed later in this document.

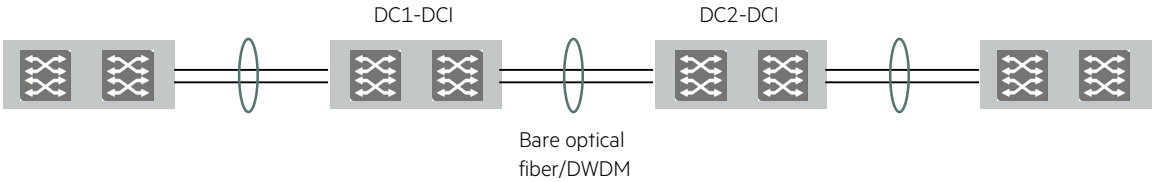


Figure 40. Point-to-point dark fiber/DWDM for two data centers

Dark fiber or DWDM DCI: RRPP solution for multiple data centers

RRPP is a convenient solution that can be used to interconnect up to four data centers. You can use RRPP to connect the aggregation layers of multiple data centers to form a DCI core network.

The major purpose of RRPP is to connect multiple DCs and to avoid a core node failure, which can split and isolate the data centers. In an RRPP network, all nodes (the aggregation layer nodes of data centers) are equal. No matter which node fails, the network connectivity can be implemented through backup links. All backup nodes and links are deployed in distributed mode.

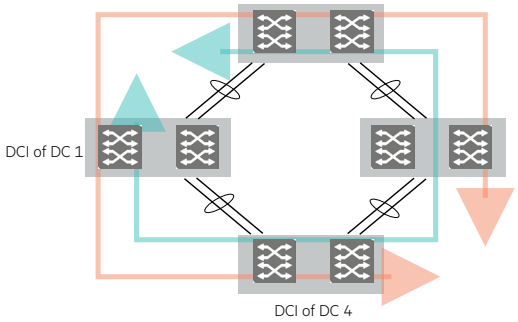


Figure 41. RRPP solution

To further improve the HA performance for the RRPP network, you can deploy IRF + LACP in the RRPP ring network. More specifically, you can deploy IRF on the aggregation layer of each data center and configure LACP to aggregate links for interconnecting aggregation layer nodes. This solution improves the HA performance for each node and links in the RRPP ring network.

The benefits of RRPP solution are:

- More robust network architecture and reasonable risk factor distribution
- Very high HA
- Easy, simple cabling

The RRPP solution has the following disadvantages:

- Supports a maximum of four data centers within the same city. More than four data centers spread across too large a distance will cause the forwarding path in the ring network to be too long. This will increase the network delay and decrease the quality of service.

Dark fiber or DWDM DCI: Hub-and-spoke solution for multiple data centers

You can use the dark fiber or DWDM and RRPP solution to interconnect up to four data centers nodes. To facilitate expanding to more data center nodes, you must use the hub-and-spoke model, where a core node is connected to the aggregation layers of multiple data centers. Logically, the multiple data centers and the core node form a hub-and-spoke star topology, where the core node is the hub, and the aggregation layer of each data center is a spoke.

RRPP + hub-and-spoke solution

The RRPP solution and the hub-and-spoke solution have their own disadvantages: The RRPP solution can cover only a small scope but provides high HA performance; and the hub-and-spoke solution provides a simple architecture, covers a large scope, delivers high scalability, but provides relatively low HA performance. However, you may find scenarios where you need to combine the two solutions to form an RRPP + hub-and-spoke solution. More specifically, you can use RRPP to connect multiple data centers in the same MAN and use the hub-and-spoke solution to interconnect the RRPP ring networks.

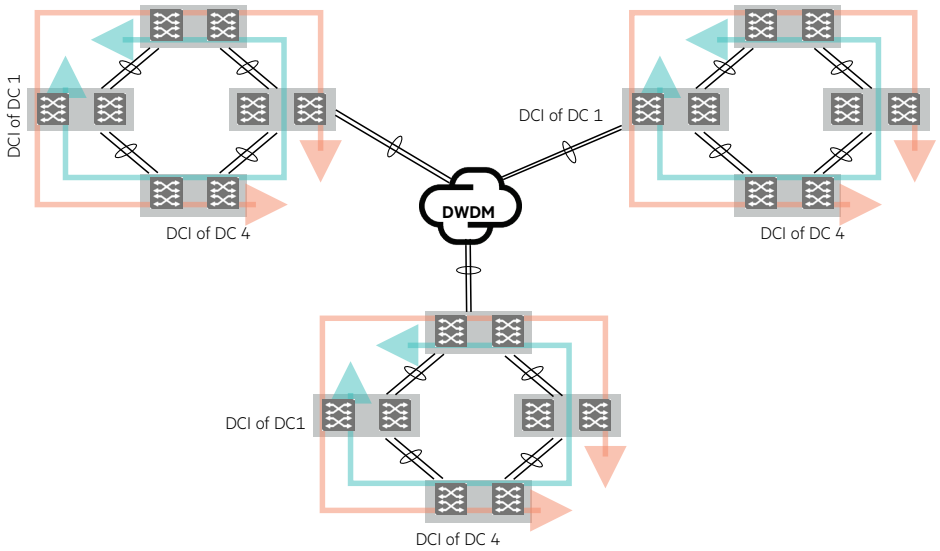


Figure 42. RRPP + hub-and-spoke solution

IP-based solution

VPLSoGRE overview

When only IP networks are used for interconnecting data centers, you can use HPE EVI or VPLS over GRE (VPLSoGRE) to implement L2 DCI. This VPLSoGRE design solution is only relevant in those scenarios, where an enterprise acquires an IP type of service from a provider and VPLSoGRE is run between the enterprise PE devices at the edge of the provider's cloud.

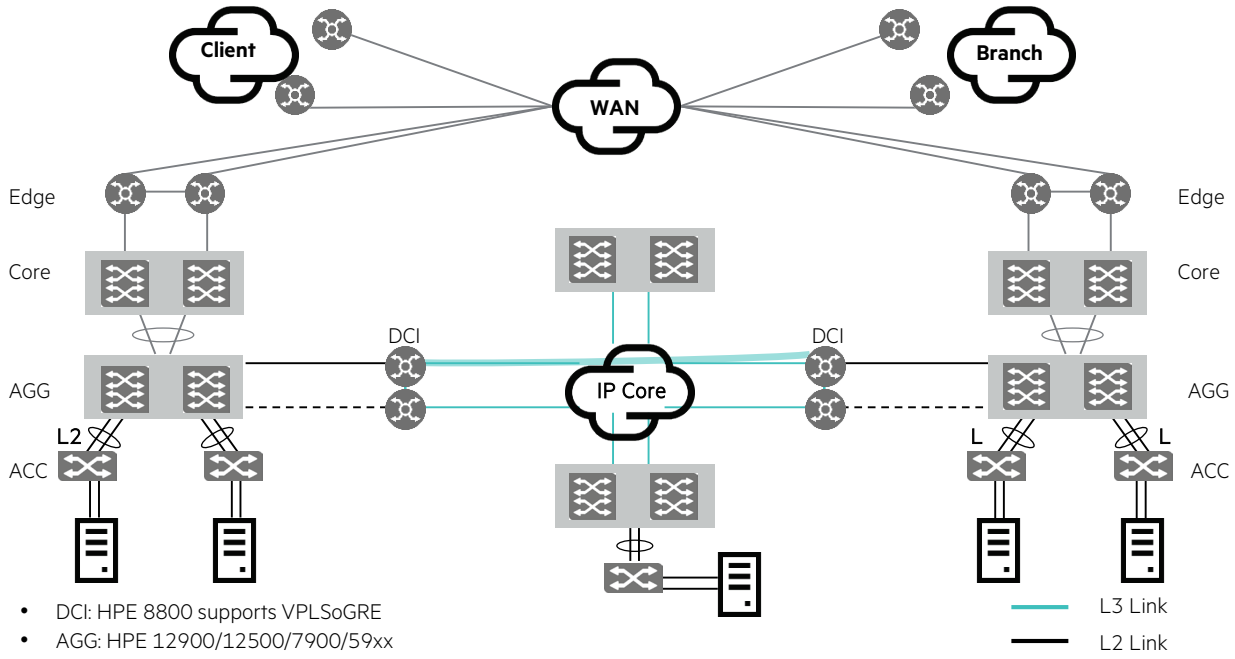


Figure 43. VPLSoGRE

The VPLSoGRE solution delivers the following benefits:

- High standardization level, powerful compatibility with all IP networks.
- Fast switchover of smart link, which improves the HA performance of the system.

The VPLSoGRE solution has the following disadvantages:

- The dedicated DCI devices do not support IRF. As a result, there are many dedicated DCI devices and the PW configuration is complicated.

The IP network provides low-service quality and you must configure network-wide end-to-end QoS, which can be very difficult.

Summary

EVI is the preferred DCI method from HPE, however, when utilizing DCI solutions other than EVI, select the proper solution by considering the performance, HA, L2 management and security factors.

Table 3. Traditional DCI performance, HA, L2 management and security factors

NETWORK RESOURCE	SOLUTION	PERFORMANCE	HA	L2 MGMT	BROADCAST CONTROL	SECURITY
Dark fiber or DWDM	RRPP	High	High	High	High	Medium
	Hub-and-spoke	High	Medium	High	High	Medium
MPLS core network	VPLS + IRF	Medium	Medium	Medium	Medium	Low
	VPLS + smart Link	Medium	Medium	Medium	Low	Low
IP core network	VPLSoGRE	Medium	Low	Medium	Low	Low

Data center management

Nowadays enterprise businesses require a high level of agility from their IT systems. Digitalization of business is a key competitive differentiator, data center network infrastructure is the spine of your company, and it has to be robust and flexible at the same time.

Every enterprise class network should include management, accounting and notification as an integral part of the complete design.

HPE Aruba IMC is a next-generation management software, which provides the data center operations team with a comprehensive platform that integrates network technologies and provides full fault, configuration, accounting, performance and security (FCAPS) management functionality.

Built from the ground up to support the Information Technology Infrastructure Library (ITIL®) operational center of excellence IT practices model, IMC's single pane-of-glass management paradigm enables efficient end-to-end business management to address the stringent demands of today's mission-critical enterprise IT operations.

IMC offers the following benefits:

- Lower operating expenses and improved TCO, because of automated features, default alerts and a consolidation of tools and correlated information
- Improved network availability and reliability that result in fewer trouble tickets, thanks to automated configuration management and comprehensive auditing
- Quicker problem recognition and troubleshooting
- Improved endpoint defense, control and visibility
- Integrated management between wired and wireless networks, traditional and SDN and even physical and virtual networks
- Excellent flexibility and scalability for networks of all sizes
- Multivendor support
- Zero Touch auto deployment plans
- Supports 6,000 devices from more than 220 manufacturers

HPE Aruba IMC base platform

HPE IMC is a comprehensive solution for the management of advanced enterprise networks ideal for large enterprise IT and data center environments. It uses a service-oriented architecture (SOA) model to deliver full and extensible device, service and user management functionality. IMC also ensures performance and scalability through distributed and hierarchical deployment models and through variable options for operating system and database support. IMC's modular design enables IMC to integrate traditionally separate management tools into a single unified platform.

IMC as a whole consists of a base platform for delivering network resource management capabilities and optional service modules for extending IMC's functionality. The base platform provides administrators and operators with the basic and advanced functionality needed to manage IMC and the devices, users and services managed by IMC. The base platform incorporates the essential functional areas of network management. The optional service modules enable administrators to extend and integrate the management of voice, wireless and MPLS VPN networks as well as end user access and endpoint defense management into IMC for a unified element management platform. The IMC base platform provides the following:

- Resource management including network device management from the SNMP, Telnet, SOAP, NETCONF, WMI, PowerShell and SSH configurations on a device to Spanning Tree configurations and PoE energy management and more.
- Baselining, configuration, auditing, and change management for device configurations and system software files for devices managed by IMC. Includes storing, backing up, comparing and deploying configuration and software files using baselines.
- Every device can have one configuration file defined as its baseline and it serves as a foundation upon which the device's configuration is evaluated and audited.
- Automation in the form of automatic discovery and mapping of all the devices in the network, automatic discovery of VMs and virtual switches, and their relationships with the physical network.

- Physical device deployment can be automated with HPE IMC in a secure dynamic fashion, providing for error free deployments.
- Applications can be quickly deployed by automating virtual resource connections and monitoring the network performance. IMC supports automatic reconfiguration when those virtual workloads are moved within and across data centers.
- NetFlow, sFlow® and NetStream flow collector support for accurate and deep analysis of traffic.
- Real-time management of events and the translation of events into faults and alarms in IMC. This includes creating, managing and maintaining alarm lists, trap and syslog filters and definitions, and configurations for notifications of alarms.
- Monitoring, reporting and alarming on the performance of network resources. This includes managing global and device specific monitors and thresholds as well as creating views and reports for displaying performance information.
- Managing Access Control List (ACL) resources including creating and maintaining ACL templates, resources and rule sets and deploying ACL rule sets to devices managed by IMC. It also includes monitoring and leveraging ACLs that exist on devices for deployment to other network devices.
- Monitoring and managing security attacks and the alarms they generate.
- Global management of VLANs for all devices managed by IMC that support VLANs.
- Administrative controls for managing IMC and access to it through operator and operator group management. System-wide management of device data collection and information is shared by all IMC modules including the creation and maintenance of device, user and service groups, and device vendor, series and model information. It also includes SNMP MIB management and other system-wide settings and functions.

HPE Aruba IMC modules

In addition to the features supported by the base IMC platform, IMC supports a variety of powerful add-on modules. Details for the various supported modules can be found at. hpe.com/us/en/networking/management.html.

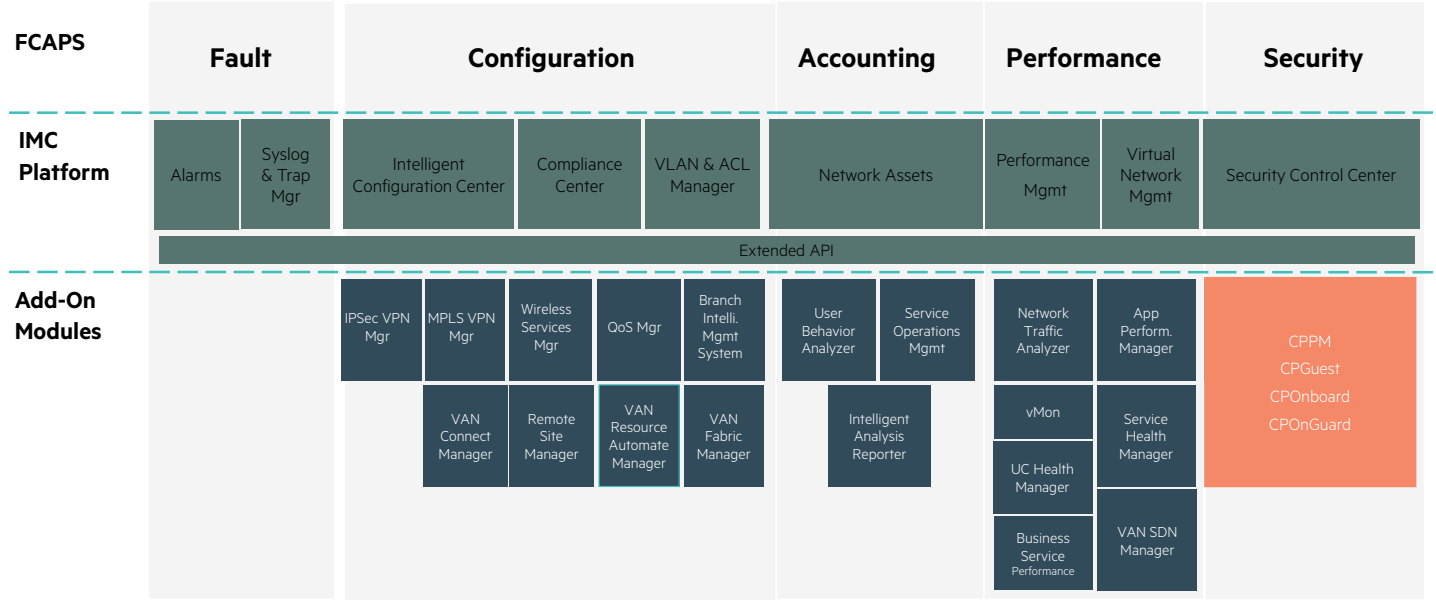


Figure 44. HPE Aruba IMC modules

Below is a brief overview of HPE IMC modules that should be considered in a DC environment.

VAN Connection Manager module

HPE IMC VAN Connection Manager is an IMC module that delivers a template-based approach for managing network configuration policies, yielding consistency and reliability across the network infrastructure. Administrators define policies in a template, which is then applied as a configuration policy to the edge switch associated with the virtual machine of interest.

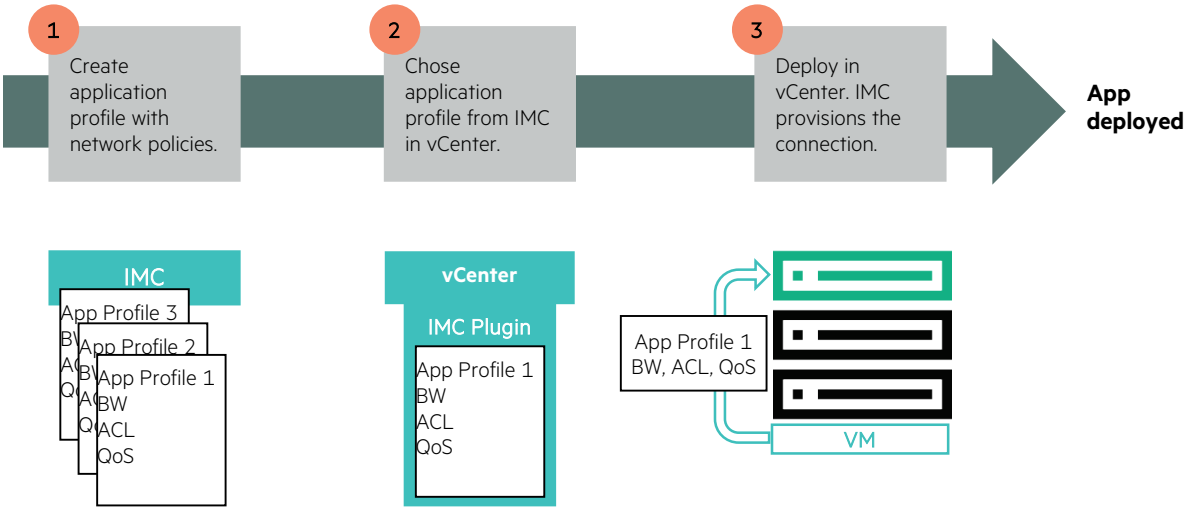


Figure 45. App deployment

Virtual machine network connectivity is automated and orchestrated by IMC VAN Connection Manager, which helps accelerate application deployment and service rollout, and greatly reduces risk during virtual machine migration. The module includes a plug-in into the VMware hypervisor manager, which enables the connection policies defined in IMC software to be applied to the virtual machine, no matter where the VM moves.

HPE Aruba IMC also integrates with the VMware vCenter and allows for centralized management and visibility of hundreds of VMware ESXi/ESX hosts and thousands of VMs, delivering operational automation, resource optimization, and HA to IT environments. Using a single management client for all tasks, administrators can provision, configure, start, stop, delete, relocate and remotely access VMs.

Application Performance Manager (APM)

HPE Aruba IMC APM module offers a view to application availability, allowing administrators to visualize and measure the health of critical business applications and their impact on network performance. It is easy to determine which business process is affected and which application issues to prioritize—all leading to quick and effective troubleshooting. It provides one user interface offering fault management, and performance monitoring of applications, servers and databases. It works with IMC Service Health Manager to provide a global assessment of network health.

Network Traffic Analyzer

As the enterprise network infrastructure expands to support different types of traffic and users, traffic management becomes critical, and complete visibility into a network's behavior becomes more important and more challenging. What is or is not happening throughout the network grid—including application performance, bandwidth utilization, network congestion, and appropriate prioritization of user and application traffic—are questions that often go unanswered.

In today's connected business environments, straightforward and effective traffic management from the network core to the network edge is essential. Enterprises need a network infrastructure that scales to meet new business needs and manages added complexity in a cost-effective manner. In addition, the data center operations team is expected to control the network in such a way that it is transparent to users. Essential information assets need to be instantly available around the clock. However, this is impossible to achieve without the right tools to make smart, informed decisions.

Most network administrators do not have simple, affordable tools that can quickly answer the following questions, regardless of the size of the network:

- Is network performance slowing down or becoming congested?
- Is NIC chattering effectively clogging the network?
- What is the current network usage, and what has it been in the past hour?
- Which network routers are most active or over-utilized?
- Why is a server slow or inaccessible?
- Which users and applications are driving network traffic?
- Which users and applications are starving for bandwidth?
- How much bandwidth do I need for new applications?

HPE Aruba IMC Network Traffic Analyzer (NTA) is a graphical network monitoring tool that utilizes industry-supported flow standards to provide real-time information about the top users and applications consuming network bandwidth.

Figure 46. Applications consuming network bandwidth

A reliable solution for data center network traffic analysis, HPE Aruba IMC NTA statistics help network administrators better understand how network bandwidth and resources are being used, as well as which source hosts carry the heaviest traffic. This information is invaluable in network planning, monitoring, optimizing and troubleshooting—IMC NTA identifies network bottlenecks and applies corrective measures to help ensure efficient throughput.

Service Health Manager (SHM) module

The HPE IMC SHM module is an IMC providing end-to-end service monitoring and service through visualization of infrastructure or network variances in the service path. It leverages data derived from other IMC components to yield critical performance metrics and aggregates key performance indicators to generate key quality indicator metrics.

Key quality indicators provide a visual representation for network administrators on their defined services and take proactive measures to maintain service level agreements.

VAN Fabric Manager

HPE Aruba IMC VAN Fabric Manager Software simplifies the management of data center and FC SAN fabrics. The software provides a unified view of network and storage devices in the data center fabric alongside the fabric health to enable quick troubleshooting and proactive management.

It reduces manual provisioning and allows for configuration of EVI, SPB or TRILL to speed fabric deployment. Additionally, it manages advanced DC technologies such as DCI/EVI—Layer 2 DC connectivity deployment and monitoring.

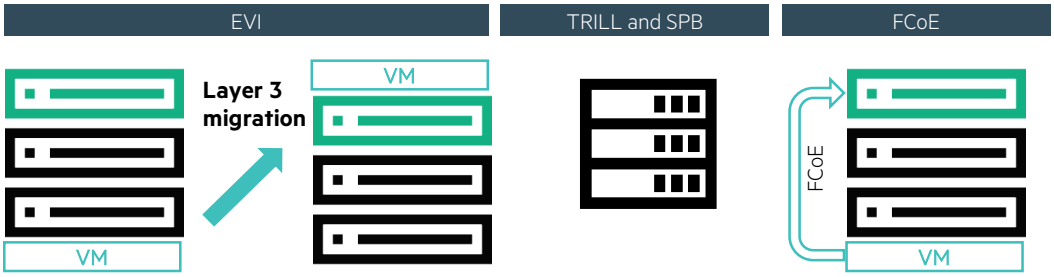


Figure 47. VAN Fabric Manager

IMC VAN Resource Automation Manager

HPE IMC VAN Resource Automation Manager Software is an IMC module providing a network fabric orchestration tool for service application delivery, optimizing the utilization of network resources for specific cloud-based or virtualized applications or tenants. This tool accelerates the deployment of applications while tuning the network to provide the best experience to users without overprovisioning valuable network resources. Converged infrastructure and cloud management becomes more robust with the end-to-end infrastructure provisioning and monitoring over the physical and virtual network.

VAN Resource Automation Manager Software has a simple-to-use service model design tool leveraging a drag-and-drop UI of HPE or third-party network resources. You can associate a specific application or tenant, desired network resources and characteristic for each service model, and provision virtual service paths through the software's orchestration capabilities. The service models allow for repeatable and consistent experience throughout the network since it can be cloned and provisioned to other parts of the network.

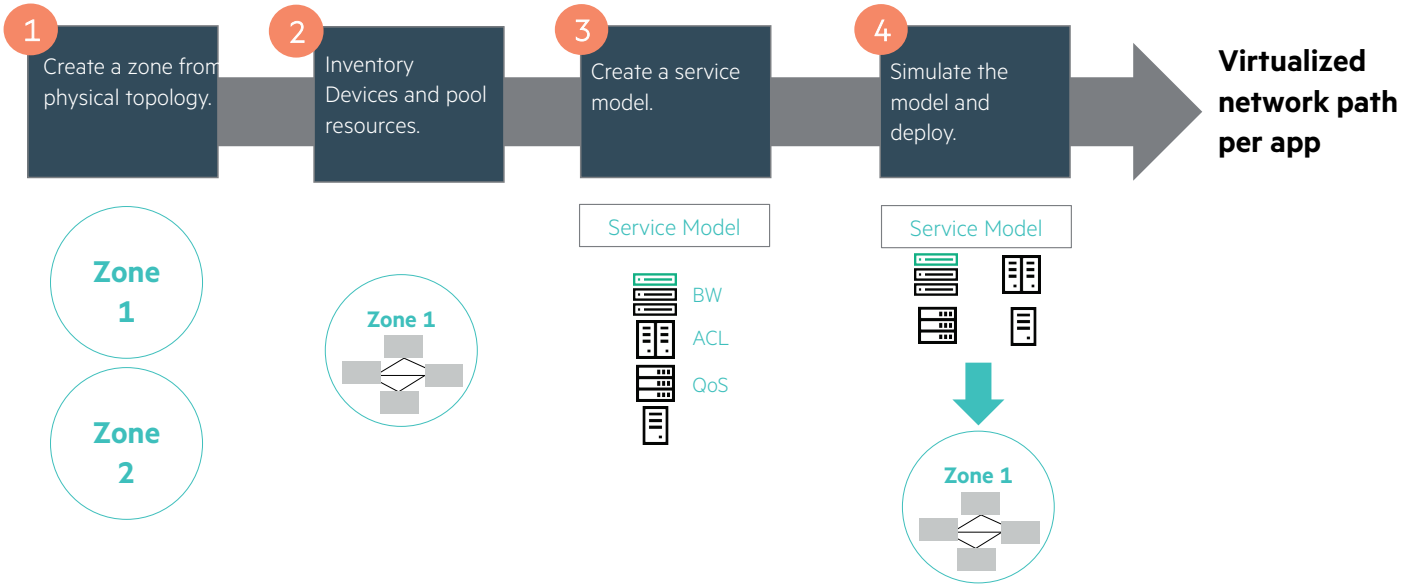


Figure 48. VAN Resource Automation Manager

HPE Data Center Networking Portfolio

HPE FlexFabric 12900E Switch Series

The HPE FlexFabric 12900E is HPE's major core data center switching platform for next generation software defined data centers. The HPE 12900E delivers unprecedented levels of performance, buffering, scale and availability with high density 10GbE, 40GbE and 100GbE interfaces. The switch series includes a 4-, 8-, and 16-slot chassis, and it supports the full Layer 2 and 3 features and advanced data center features to build resilient scalable fabric and achieve convergence.

Figure 49. HPE FlexFabric 12900E Switch Series

The HPE FlexFabric 12900E Switch Series provides:

- Non-blocking, lossless Clos architecture with VOQs and large buffers with the flexibility and scalability for future growth.
- Distributed architecture delivers enhanced fault tolerance and facilitates continuous operation and zero service disruption during planned or unplanned control-plane events.
- Up to 46 Tbps switching capacity and 28.8 Bpps, providing you non-blocking wire speed performance.
- High-density 1GbE, 10GbE, 40GbE and 100GbE interface connectivity.
- Ability to build Layer 2 fabrics which are flexible, resilient, and scalable with VxLAN, TRILL and/or Hewlett Packard Enterprise IRF.
- Multitenant Device Context (MDC) for multi-tenancy giving you the ability to virtualize a physical switch into multiple logical devices; each logical switch has its own tenants.
- Network and storage convergence with support for Fiber Channel over Ethernet (FCoE) and Data Center Bridging (DCB) protocols include IEEE 802.1Qaz Data Center Bridging Exchange (DCBX), Enhanced Transmission Selection (ETS) and IEEE 802.1Qbb Priority Flow Control (PFC).

HPE FlexFabric 7900 Switch Series

The HPE FlexFabric 7900 Modular Core Switch is a compact modular data center core switch supporting virtualized data centers and evolutionary needs of private and public clouds deployments. This switch series uses a fraction of the footprint used by traditional chassis while still delivering very high levels of performance, buffering, scale and availability. This switch series features 10GbE, 40GbE and 100GbE interfaces with support for Layer 2 and 3 features, including advanced features such as VXLAN and HPE IRF, which enables scale-out, two-tier leaf-spine architecture.

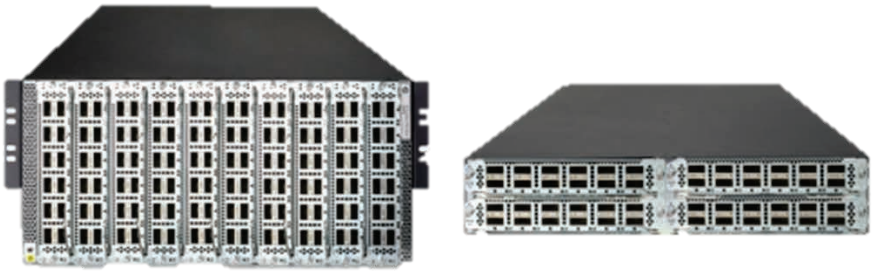


Figure 50. HPE FlexFabric 7900 Switch Series

The HPE FlexFabric 7900 Switch Series provides support for:

- Virtualized data centers and public and private clouds deployments.
- Scale-out, two-tier leaf-spine architectures with greater reliability.
- 10Gb, 40GbE and 100GbE interfaces in a smaller footprint than a traditional modular chassis.
- Rich HPE Comware capabilities and simplified zero cost/complexity licensing.
- Simplified, Automated Software-Defined Networking (SDN) Fabric
- Large Layer 2 scaling with TRILL and HPE IRF, VBXLAN, and OpenFlow 1.3.
- Federated HPE FlexFabric infrastructure with VMware NSX virtual overlay.

HPE FlexFabric 5950 Switch Series

The HPE FlexFabric 5950 is a 25/50/100GbE networking switch series that provides customers with a high density and ultra-low latency solution enabling them to deploy spine/leaf networking configurations in the data center for business critical applications.

Consisting of a 1U 32-port 100GbE QSFP28 Switch, the 5950 brings high density to a small footprint. The 100GbE ports may be split into four 25GbE ports and can also support 40GbE which can be split into four by 10GbE for a total of 128 25/10GbE ports.

Figure 51. HPE FlexFabric 5950-32QSFP28

The HPE FlexFabric 5950 Switch Series provides support for:

- 3.2 Tbps switching capacity for the most demanding applications.
- 2976 MPPS throughput for data-intensive environments.
- Under a 1µs 100GbE latency, gives your business agility.
- VXLAN support for network virtualization and overlay solutions.
  - OVSDB for dynamic VXLAN tunnel management
- IRF <50 msec convergence time enabling faster application response time.

- IRF-based In Service Software Update (ISSU) enables high availability with updates accomplished without a reboot or power cycle, in the background.
- No extra hidden cost with one license per switch for all OS features.
- All switch ports are active and ready to use without need for activation licenses.
- Automate tedious tasks with a Software-defined Network (SDN) and reclaim wasted resources.

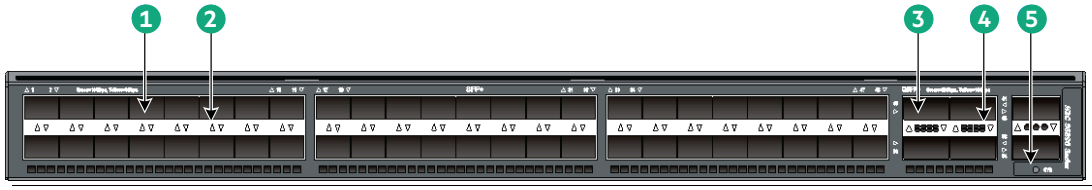
HPE FlexFabric 5940 Switch Series

The HPE FlexFabric 5940 is HPE's latest 1RU/2RU 10/40/100GbE networking ToR switch series (available summer 2016) that is specifically designed for data centers that require rich data centers feature sets. The HPE 5940 series supports L2/L3 VXLAN VTEP, large MAC and ARP tables, full Ipv4/Ipv6 routing suite, IPv4 and IPv6 dual-stack, dual power supplies, dual fans and reversible airflow.

The HPE 5940 series also supports IRF technology, allowing users to connect multiple switches to form a single logical entity, thus building high reliability, extensibility and ease of management of the new intelligence network.

The initial release of the HPE FlexFabric 5940 will include models that feature:

- 48x 10GbE SFP+ ports + 6x 100GbE QSFP28 ports
- 48x 10GbE SFP+ ports + 6x 40GbE QSFP+ ports
- 48x 10GbE BaseT ports + 6x 100GbE QSFP28 ports
- 48x 10GbE BaseT ports + 6x 40GbE QSFP+ ports



(1) SFP+ port	(2) SFP+ LED
(3) QSFP+ port	(4) QSFP+ LED
(5) System status LED (SYS)	

Figure 52. HPE FlexFabric 5940

HPE FlexFabric 5930 Switch Series

The HPE FlexFabric 5930 Switch Series consists of two modular switches (2-slot and 4-slot) and a fixed switch (32-port 40GbE). This series is a family of high-density, small footprint, low-latency, ToR switches supporting a rich data center feature set which includes VXLAN VTEP, OpenFlow/SDN, ISSU, Native FC, DCB, TRILL and MAC-sec.

Ideally suited for deployment at the aggregation or server access layer of large enterprise data centers, the HPE FlexFabric Modular 5930/Fixed 5930 Switch Series is also powerful enough for deployment at the data center core layer of medium-sized enterprises.

The HPE FlexFabric 5930 Switch Series provides support for:

- VXLAN and OVSDB support for network virtualization and overlay solutions.
- IRF support of up to nine switches simplifies management by up to 88%.
- OpenFlow and SDN automate manual tasks and speed service delivery.
- 1RU 32-port 40GbE QSFP+, 2RU 2-slot with 2 40GbE QSFP+ and 2RU 4-slot form factors.
- Modular port options include MACsec, 10GbE SFP+, 10GBASE-T and converged ports supporting 1/10GbE and 4/8Gbps Fiber Channel.
- Up to 2.56 Tbps switching capacity for the most demanding applications.

- Up to 1492 Mpps throughput for data-intensive environments.
- In Service Software Update (ISSU) which enable high availability.
- All switch ports are active and ready to use without need for activation licenses.
- Automate tedious tasks with a Software-defined Network (SDN) and reclaim wasted resources.

Figure 53. HPE FlexFabric 5930 Switch Series

HPE FlexFabric 5900AF/5900CP and 5920AF Switch Series

This is a family of high-density 10GbE and ultra-low latency ToR switches, which are ideally suited for deployment at the server access layer of large enterprise data centers. These switches support TRILL, Virtual Ethernet Port Aggregator (VEPA), FCoE and Fibre Channel for virtualized networks and data center convergence.

The HPE FlexFabric 5900/5920 Switch Series provides support for:

- Large data buffers for loss sensitive environments (HPE 5920).
- Full L2/L3 features, IPv4/IPv6 dual stack, OpenFlow and TRILL for high scalability and Software-defined Networking (SDN) support.
- IRF for simpler management, faster re-convergence and business agility.
- Low latency, under a 1.5 µs 10GbE latency, provides increased performance.
- In Service Software Update (ISSU) enables high availability with updates accomplished without a reboot or power cycle, in the background.
- All switch ports and features are active and ready to use without need for activation licenses.
- Automate tedious tasks with SDN and reclaim wasted resources.

Figure 54. HPE FlexFabric 5900 Switch Series

HPE FlexFabric 5700 Switch Series

The HPE FlexFabric 5700 Switch Series provides an open door for expanding your enterprise network by adding capacity with local switching and L2/Light L3 support. Take advantage of IRF enhancements for spine/leaf configurations, to simplify network management and expand server connectivity. While IRF reduces management complexities by up to 88%, it also delivers <50msec convergence time. You can rely on the FlexFabric 5700 to lower TCO with up to a 25% lower cost than competing devices.

The HPE FlexFabric 5700 Switch Series provides support for:

- Choices that fit your budget and environment by offering 1/10GbE ports supporting SFP and BASE-T with 10/40GbE uplinks.
- The HPE FlexFabric 5700 Switch Series delivers up to 960Gbps switching capacity for the most demanding applications.
- Low latency, under 1.5us 10GbE latency, provides improved throughput and fewer lost packets.
- <50msec convergence time enabling faster application response time.
- In Service Software Update (ISSU) enables high availability

- No hidden software licensing costs, all OS features included with switch purchase.
- Automate tedious tasks with software defined network (SDN) and reclaim wasted resources.

Figure 55. HPE FlexFabric 5700 Switch Series

HPE HSR6800

The HPE HSR6800 Router Series is a family of high-performance, multiservice routers designed for data center interconnection, enterprise WAN core, campus WAN edge, and high-speed WAN aggregation services. It runs the Comware operating system and features an advanced multi-core, distributed service processing hardware architecture that scales up to 420 Mpps forwarding and up to 2 Tbps switching capacity.

The router delivers robust routing, multicast, MPLS, IPv6, security, quality of service, carrier-level high-availability features, and high-density 10GbE and 1GbE interface options.

The HPE HSR6800 Router Series provides support for:

- Advanced hardware architecture with multi-core CPUs and fully distributed routing and service engines for increased performance.
- Scale up to 420 Mpps forwarding performance, 120 Gbps IPsec performance, 2 Tbps switching capacity and delivers up to 64 10GbE ports, 384GbE ports, 32 channelized OC-3 POS interfaces, 1512 E1 interfaces and 2016 T1 interfaces.
- Separate control and service planes to avoid interference and to facilitate service continuity during an active/standby switchover.
- Distributed design allows packet forwarding and complicated services such as NAT, GRE, NetStream, QoS and IPsec to be processed on each line card independently, thus tremendously enhancing the service processing performance of the overall system as line cards are added.
- All IPv4 and IPv6 routing protocols including RIP/RIPng, OSPF/OSPFv3, IS-IS/IS-ISv6, BGP/BGP4+, PIM/PIM6, MSDP, MBGP and policy based routing delivering increased flexibility.
- Comprehensive MPLS features, including LDP, MPLS TE, L3 VPN, L2 VPN, VPLS, Multicast VPN, 6PE and 6vPE which provides more flexibility at cost-effective price-points.
- All-round security protection features, including packet filtering, stateful firewall, ACL, attack detection and prevention, control plane rate limiting, uRPF, AAA, IPSec VPN and Auto-discovery VPN.
- Advanced QoS features, including queue scheduling, congestion avoidance, congestion management, traffic policing, traffic shaping, priority marking and hierarchical QoS.
- IRF to implement system virtualization. IRF enables two HSR6800 Routers to form and work as a single virtual device.
- Dual-MPUs, redundant power modules and advanced distributed service architecture.
- High availability features that reduce service disruption including non-stop routing, graceful restart, stateful failover, IGP fast convergence, in-service software upgrades (ISSUs), fast reroute and innovative multi-chassis resiliency.

Figure 56. HPE HSR6800 Router Series

HPE Virtual Connect

VC modules provide a better way for IT to work together and offer benefits to all parties involved, without the traditional compromises. It is simply the best way to connect blade servers to network LANs and SANs for the lowest costs and least amount of power. HPE continues to expand this technology and its capabilities across ProLiant, Integrity and Storage product lines. VC can simplify and converge your server edge connections, integrate into any standards-based networking infrastructure and reduce complexity with industry-leading flexibility.

HPE 6127XLG Blade Switch family

Designed for the HPE BladeSystem c-Class enclosure, the HPE 6127XLG Blade Switches provide 16 1GbE, 10GbE or 20GbE server downlinks and 10/40GbE options for uplinks, along with 10GbE cross-connects. A robust set of industry-standard L2/3 switching functions, IRF, Comware v7, VXLAN VTEP support, IMC management, QoS metering, security and high-availability features round out this extremely capable blade switch family.

With a variety of connection interfaces, the 6127 Blade Switch family offers excellent investment protection, flexibility and scalability as well as ease of deployment and reduced operational expense.

HPE IMC

A comprehensive platform that integrates management of physical and virtual network resources and provides full FCAPS management functionality for IT infrastructures.

Data center optimized workload solutions

HPE Helion Rack

HPE Helion Rack is a pre-configured, pre-tuned, and pre-tested private cloud solution that provides fast deployment of an OpenStack-based private cloud with lower TCO and end-to-end support by HPE. It enables rapid infrastructure provisioning for cloud-native application development and production workloads with the flexibility to meet dynamic, low-latency, high-transaction and compute-intensive workload requirements—so your cloud is ready to support the cloud-native applications your developers create from Big Data to mobile applications, and also those with high-performance database requirements.

HPE Helion Rack includes the HPE Helion Platform comprised of HPE Helion OpenStack (IaaS layer) and the HPE Helion Development Platform (PaaS layer). And it's powered by the industry leading HPE ProLiant Gen9 server platform optimized for OpenStack to give you the complete, open private cloud solution you need

HPE Helion CloudSystem

HPE Helion CloudSystem is the most complete, integrated and open cloud solution on the market. Powered by OpenStack technology and developed with an emphasis on automation and ease-of-use, HPE Helion CloudSystem redefines how you build and manage cloud services. Created with enterprises and service providers in mind, our end-to-end service lifecycle management solution gives you the ability to effortlessly provision applications and their corresponding infrastructure resources. With an Open Source approach and support for hybrid service delivery, you can run on-premises services in your private cloud while simultaneously managing off-premises services or bursting to external public cloud providers.

Whether you are transitioning existing environments to cloud or starting your journey from scratch, HPE Helion CloudSystem offers a rapid, reliable and cost-effective path with built-in extensibility for future growth and development.

HPE ConvergedSystem

HPE ConvergedSystem is a smart way to build an infrastructure. It provides for lower cost of ownership and greater flexibility to meet changing business demands. With HPE CS, deploy pre-validated, factory-tested configurations in weeks instead of months. Set up IT services in minute, not hours. Integrated workload-optimized systems eliminate infrastructure silos for improved efficiency, simplicity and speed.

HPE Technology Services—Mobility and Networking

HPE Technology Services offers a comprehensive set of services to help advice, transform, integrate, support and evolve your next-generation connectivity and communications environment. HPE Trusted Network Transformation approach can help you manage the risk in the transformation journey to an SDN-enabled business aligned data center network.

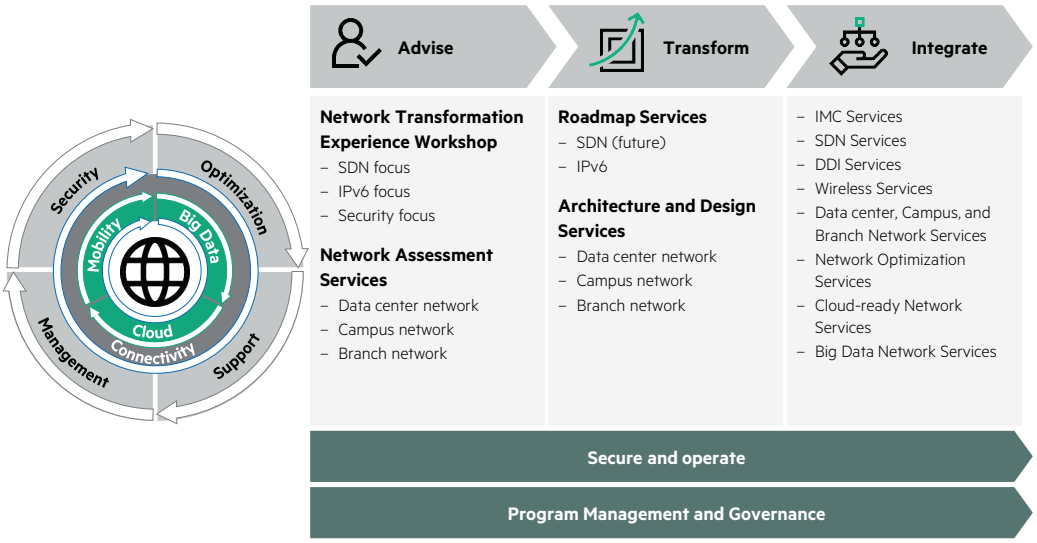


Figure 57. HPE TS

With HPE, customers can:

- Realize the business outcomes of a transformation
  - Increased business agility, higher availability and performance
  - Lower TCO, improved demand management, governance, capital and budget planning, lifecycle and capability maturity management
- Lower the risk in transformation
  - Guided steps that start with business and IT strategy alignment and addresses the change to people, process and technology
  - Use our IT and financial expertise and project experience to guide the customer data center networking project
  - Maintain business continuity with a safe and proven approach to deliver on transformation
  - Over 40 years of multivendor experience across networking, servers, storage and finance

Contact your HPE account manager or reseller to find out how HPE Technology Services can help you manage the risk in implementing an automated, agile, high-performance data center network.

Support, services and partners

The core foundation to the solution (HPE and its technology partners):

- HPE is #19 on the list of Fortune 500 companies
- HPE has a global reach, so your data center solution will be supported internationally
- HPE has industry-leading support and professional services to support your network
- HPE products can be integrated with other vendor solutions through the use of standards-based technologies
- HPE can provide a complete end-to-end virtual computing solution encompassing racks, servers, storage, networking technologies, security and management

More information regarding the HPE Services can be found at the HPE Business Services product page.

Glossary

ACL: A network Access Control List (ACL) is an optional layer of security for your VPC that acts as a firewall for controlling traffic in and out of one or more subnets.

ARP: The Address Resolution Protocol (ARP) is a protocol used by the Internet Protocol (IP) [RFC826], specifically IPv4, to map IP network addresses to the hardware addresses used by a data link protocol. The protocol operates below the network layer as a part of the interface between the OSI network and OSI link layer.

BGP: Border Gateway Protocol (BGP) is a standardized exterior gateway protocol designed to exchange routing and reachability information among autonomous systems (AS) on the Internet. The protocol is often classified as a path vector protocol but is sometimes also classed as a distance-vector routing protocol.

CEE: Converged Enhanced Ethernet (CEE) is an enhanced Ethernet that enables the convergence of various applications in data centers (LAN, SAN and HPC) onto a single interconnect technology.

CIFS: The Common Internet File System (Microsoft CIFS), an enhanced version of Microsoft Server Message Block (SMB), is the standard way that computer users share files across intranets and the Internet.

CoS: Class of Service (CoS) is one type of the techniques or methods used to deliver Quality of Service (QoS) in a network.

DoS: A Denial of Service (DoS) attack is an occasion in which a legitimate user or a group of users is prevented from accessing the services and information of network resources they would normally receive.

DCB: Data Center Bridging (DCB) is a series of enhancements to the IEEE 802.1 standard to provide extensions to Ethernet for support for converged technologies such as Fiber Channel over Ethernet (FCoE).

EoR: End-of-row topologies, which rely on larger switches placed on the end of each row for server connectivity.

ETS: Enhanced Transmission Selection (ETS) is to allocate bandwidth based on the different priority settings of the converged traffic.

EVB: Edge Virtual Bridging (EVB) is an IEEE standard that involves the interaction between virtual switching environments in a hypervisor and the first layer of the physical switching infrastructure.

EVI: Ethernet Virtual Interface (EVI) runs over Internet Protocol (IP) transport and extends Layer 2 domains across a WAN network, typically between data centers. By virtualizing and automating the link-layer domain across data centers, EVI delivers the elements necessary to enable Software-defined Networking (SDN) data center infrastructure. It enables several data centers to work as one that is more responsive, with higher efficiency and solid high availability for business resiliency.

FCAPS: An extension of the popular network management conceptual frameworks called telecommunication management network (TMN), FCAPS describes network management in four layers. Each TMN layer needs to perform some or all FCAPS functions in certain ways.

FCIP: Fiber Channel over TCP/IP (FCIP) describes mechanisms that allow the interconnection of islands of Fiber Channel storage area networks over IP-based networks to form a unified storage area network in a single Fiber Channel fabric.

FCoE: Fiber Channel over Ethernet (FCoE) is an encapsulation of Fiber Channel frames over Ethernet networks. This allows Fiber Channel to use 10 Gigabit Ethernet networks (or higher speeds) while preserving the Fiber Channel protocol.

FC: Fiber Channel, a Gigabit-speed network technology primarily used for storage networking.

HPE IMC: HPE Intelligent Management Center (IMC) delivers next-generation, integrated and modular network management capabilities that efficiently meet the end-to-end management needs of advanced, heterogeneous enterprise networks.

IDS: An intrusion detection system is a device or software application that monitors network and/or system activities for malicious activities or policy violations and produces reports to a management station.

IRF: Intelligent Resilient Framework (IRF) is a software virtualization technology developed by H3C (3COM). Its core idea is to connect multiple devices through physical IRF ports and perform necessary configurations, and then these devices are virtualized into a distributed device.

iSCSI: The Internet Small Computer System Interface (iSCSI) is a TCP/IP-based protocol for establishing and managing connections between IP-based storage devices, hosts and clients, called the Storage Area Network (SAN).

Jumbo frames: Jumbo frames often mean 9,216 bytes for Gigabit Ethernet, but can refer to anything over 1,500 bytes.

LACP: Link Aggregation Control Protocol (LACP) is part of the IEEE specification 802.3ad that allows you to bundle several physical ports to form a single logical channel.

LAG: Link Aggregation (LAG) is used to describe various methods for using multiple parallel network connections to increase throughput beyond the limit that one link (one connection) can achieve.

MAC: A Media Access Control address (MAC address), also called physical address, is a unique identifier assigned to network interfaces for communications on the physical network segment. MAC addresses are used as a network address for most IEEE 802 network technologies, including Ethernet and WiFi.

MDC: Multitenant Device Context (MDC) is an HPE feature for multi-tenancy which gives you the ability to virtualize a physical switch into multiple logical devices; each logical switch has its own tenants.

MMF: Multi-mode optical Fiber is a type of optical fiber mostly used for communication over short distances, such as within a building or on a campus.

MPLS: Multiprotocol Label Switching (MPLS) is a type of data-carrying technique for high-performance telecommunications networks that directs data from one network node to the next based on short path labels rather than long network addresses, avoiding complex lookups in a routing table.

MPO/MTO: MPO (Multi-fiber Push On) is a connector for ribbon cables with four to twenty-four fibers. MTP is a brand name for a version of the MPO connector with improved specifications.

MSTP: The Multiple Spanning Tree (MST) protocol carries the concept of the IEEE 802.1w rapid spanning tree protocol (RSTP) a leap forward by allowing the user to group and associate VLANs to multiple spanning tree instances (forwarding paths) over link aggregation groups (LAGs).

NAT-PT: Network Address Translation with Port Translation (NAT-PT) is a service which can be used to translate data sent between IP-heterogeneous nodes.

NIC: Network Interface Cards (NIC) are adapters attached to a computer (or other network device such as a printer) to provide the connection between the computer and the network.

NMS: The Network Management System (NMS) is a combination of hardware and software used to monitor and administer a network.

OSPF: Open Shortest Path First (OSPF) is a routing protocol for Internet Protocol (IP) networks. It uses a link state routing algorithm and falls into the group of interior routing protocols, operating within a single autonomous system (AS).

OVSDB: The Open vSwitch Database Management Protocol (OVSDB) is an OpenFlow configuration protocol that is designed to manage Open vSwitch implementations.

PFC: Priority Flow Control (PFC) is defined by a one-byte bitmap. Each bit position stands for a user priority. If a bit is set, the flow control is enabled in both directions (Rx and Tx).

Port mirroring: Port mirroring is used on a network switch to send a copy of network packets seen on one switch port (or an entire VLAN) to a network monitoring connection on another switch port.

PXE: Pre-execution Environment (PXE) is an environment to boot computers using a network interface independently of data storage devices (like hard disks) or installed operating systems.

QCN: Quantized Congestion Notification (QCN) is a form of end-to-end congestion management defined in IEEE 802.1Qau. The purpose of end-to-end congestion management is to ensure that congestion is controlled from the sending device to the receiving device in a dynamic fashion that can deal with changing bottlenecks.

RAID: Redundant Array of Inexpensive Disks (RAID) is a technology that provides increased storage functions and reliability through redundancy. Combining multiple disk drive components into a logical unit, where data is distributed across the drives in one of several ways called "RAID levels".

RIB: In computer networking a routing table, or Routing Information Base (RIB), is a data table stored in a router or a networked computer that lists the routes to particular network destinations, and in some cases, metrics (distances) associated with those routes.

RSTP: The Rapid Spanning Tree Protocol (RSTP IEEE 802.1w) can be seen as an evolution of the IEEE 802.1d standard more than as a revolution. IEEE 802.1w is also capable of reverting back to IEEE 802.1d in order to interoperate with legacy bridges (thus dropping the benefits it introduces) on a per-port basis.

SAN: A Storage Area Network (SAN) is a high-speed special-purpose network (or subnetwork) that interconnects different kinds of data storage devices with associated data servers on behalf of a larger network of users.

SCSI: The Small Computer System Interface (SCSI), an ANSI standard, is a parallel interface standard used by Apple Macintosh computers, PCs, and many UNIX® systems for attaching peripheral devices to computers.

SFP: The Small form-factor Pluggable (SFP) is a compact, hot-pluggable transceiver used for both telecommunication and data communications applications. The form factor and electrical interface are specified by a multi-source agreement (MSA).

SMF: Single Mode Fiber optic cable has a small diametral core that allows only one mode of light to propagate. This application is typically used in long distance, higher bandwidth runs.

SMB: The Server Message Block (SMB) protocol is an IBM protocol for sharing files, printers, serial ports, etc. between computers.

SNMP: The Simple Network Management Protocol (SNMP) is used by network management systems to communicate with network elements.

SPB: Shortest Path Bridging (SPB) provides logical Ethernet networks on native Ethernet infrastructures using a link state protocol to advertise both topology and logical network membership.

STP: The Spanning Tree Protocol (STP) is an L2 protocol designed to run on bridges and switches. The main purpose of the spanning tree is to prevent loops from forming in a bridged network.

TCP Windows® sizing: With window sizing, TCP dynamically determines the number of frames to send at once without an acknowledgement.

TOE: TCP Offload Engine (TOE) technology aims to take the server CPU out of I/O processing by shifting TCP/IP processing tasks to the network adapter or storage device. This leaves the CPU free to run its applications, so users get their data faster.

ToR: Top-of-rack utilizes a switch at the top of each rack (or close to it).

QSFP: The Quad Small Form-factor Pluggable (QSFP) is a compact, hot-pluggable transceiver used for data communications applications. It interfaces networking hardware to a fiber optic cable or active or passive electrical copper connection.

ToR: Top-of-rack utilizes a switch at the top of each rack (or close to it).

TRILL: TRILL (Transparent Interconnection of Lots of Links) is an IETF standard implemented by devices called RBridges (routing bridges) or TRILL Switches. TRILL combines techniques from bridging and routing and is the application of link state routing to the VLAN-aware customer-bridging problem.

VEPA: A standard being led by HPE for providing consistent network control and monitoring for virtual machines (of any type).

VM: A Virtual Machine is a system that enables multiple operating systems to concurrently run on a single physical server, providing much more effective utilization of the underlying hardware.

VLANs: Virtual LANs (VLANs) provide the capability to overlay the physical network with multiple virtual networks. VLANs allow you to isolate network traffic between virtual networks and reduce the size of administrative and broadcast domains.

VTEP: A VXLAN (Virtual Extensible Local Area Network) Tunnel End Point (VTEP) is a host interface which forwards Ethernet frames from a virtual network via VXLAN or vice-versa. All hosts with the same VNI configured must be able to retrieve and synchronize data (ARP and MAC tables for example).

VXLAN: Virtual Extensible LAN (VXLAN) is a network virtualization technology that attempts to improve the scalability problems associated with large cloud computing deployments. It uses a VLAN-like encapsulation technique to encapsulate MAC-based OSI Layer 2 Ethernet frames within Layer 4 UDP packets.

WWN: A World Wide Name or World Wide Identifier (WWID) is a unique identifier which identifies a particular Fiber Channel, Advanced Technology Attachment (ATA) or Serial Attached SCSI (SAS) target.

Resources, or additional links

HPE FlexFabric Reference Architecture: Data Center Trends

HPE FlexFabric Reference Architecture Guide-100 Server

HPE FlexFabric Reference Architecture guide-500 Server

HPE FlexFabric Reference Architecture guide-2000 Server

HPE FlexFabric Reference Architecture-Building data center networks using HPE and F5

Building data center networks using HPE and Alcatel-Lucent

VXLAN in HPE data center solutions

HPE Helion OpenStack

HPE Helion Rack

HPE Helion CloudSystem

HPE ConvergedSystem

HPE Ethernet Virtual Interconnect

HPE Converged Infrastructure white papers and videos

HPE Intelligent Resilient Framework

HPE Networking 6125XLG Blade Switch with HPE 5900CP ToR switch FC/FCoE solution using HPE Blade Servers and HPE Storage

HPE Networking single-tier FC/FCoE solution using HPE Rack Servers and HPE Storage

HPE Virtual Connect technology information

HPE IMC data sheets and product details

Sign up for updates

4AA5-6480ENW, July 2016, Rev. 2