



## TECHNOLOGY STRATEGY BRIEF

# Integrating Physical and Virtual Networking – Virtualized Server Connect to the Data Center Fabric

---

# Integrating Physical and Virtual Networking – Virtualized Server Connect to the Data Center Fabric

## Introduction

With the promise of reduced capital investment, higher agility and lower operational expense, server virtualization has revolutionized all facets of the data center from servers to storage to the network. While virtualization has become a mainstream technology, its success depends on the integration of both the physical and the virtual network. This integration brings many considerations and challenges along with it.

For the full value of virtualization to be realized, enterprises must move beyond server consolidation with locally contained feature usage to taking full advantage of all the features that the virtualization technology offers to automate the operation of the server and the network infrastructure.

Server and storage virtualization enable rapid changes on the services layer, but the dynamic nature of virtualization places drastic requirements on the data center network. “Motion” technologies require rapid configuration changes on the network layer as servers/virtual machines (VMs) are added or moved among physical machines. In order to deliver network services in real-time within a virtualized environment, technologies are being developed to bridge the divide between VM/server and network or fabric provisioning applications. Organizational hurdles that exist between server and network operations must also be overcome or the increased operational costs would offset the positive ROI benefits that virtualization promises.

Another challenge is the Hypervisor-based virtual Switch (vSwitch) configuration which connects VMs locally to each other, as well as to the physical network. When VMs reside on the same physical server, communication might not be exposed to the external physical switch, making policy enforcement within the physical network a challenge. The vSwitch configuration needs to be aligned and orchestrated with the physical network, which further increases the complexity of a combined fabric infrastructure.

The Enterasys OneFabric architecture addresses these challenges by providing a unified network fabric that delivers centralized visibility and control over the entire network from the data center to the edge. Purpose-built for dynamic and complex IT environments leveraging virtualization technologies, OneFabric makes the once-complex task of provisioning and de-provisioning servers and network infrastructure simple: defined locally and enforced globally achieving significant scale, improved operational efficiency, and more reliable and successful application delivery.

Focusing on the integration of the physical and virtual network infrastructure that must take place for a virtualization solution to be successful, this paper provides an overview of the most viable options that are currently available. This paper will also discuss how leveraging the Enterasys OneFabric architecture can enable organizations to simply deploy virtualization solutions today, while leaving the door open to rapidly scale the network for the business solutions of tomorrow.

## OneFabric™ Benefits

### Simplicity

- The OneFabric architecture is easy to deploy and leverages existing standards and features

### Automation

- The provisioning of the virtual and physical network infrastructure is automated and dynamic

### Visibility and Control

- All traffic inside the data center fabric is enforced the same way if hairpin mode is used
- Statistics are collected for all traffic inside the data center fabric if hairpin mode is used

### OPEX Reduction

- Simplicity and automation of deployment, add/move/changes results in lower operational costs

### Investment Protection

- Emerging standards can be supported through software upgrades on CoreFlow2 based data center products – no rip and replace

## Integrated Solutions

Integrated solutions aim to orchestrate the configuration of physical (controlling) bridges and virtual edge bridges (VEBs)/switches. VEBs are implemented today in several hypervisors. The goal of this approach is to have policy enforcement and traffic forwarding taking place in concert on the physical switches. The umbrella term for the series of developing standards is called Edge Virtual Bridging (EVB). More specifically, the EVB environment encompasses VEBs, Virtual Ethernet Port Aggregator (IEEE 802.1Qbg) technologies and protocols that help automate the coordination and configuration of network resources.

### Leveraging Private VLANs

This option uses the private VLAN function that the VMware vSwitch supports (please note that a specific license is required) to enforce policies at the physical switch. This provides a means to gather traffic statistics via Netflow and include advanced security controls, such as intrusion prevention or firewalling, for any inter-VM traffic. This option works in conjunction with Enterasys Fabric Routing, the policy enforcement scheme on the data center switches and the Data Center Manager (DCM), which is part of the OneFabric Control Center. It is available today based on using existing standards and features (VMware 4.x and 5.x dvSwitch and its private VLAN implementation) to orchestrate the enforcement of policies, create Netflow traffic statistics and enable the external inspection of VM traffic. Leveraging private VLANs addresses the deficiencies of today's vSwitch/VEB implementations by allowing an external switch to control and enforce policies even for traffic that originally would have been internally forwarded by the vSwitch within the hypervisor layer.

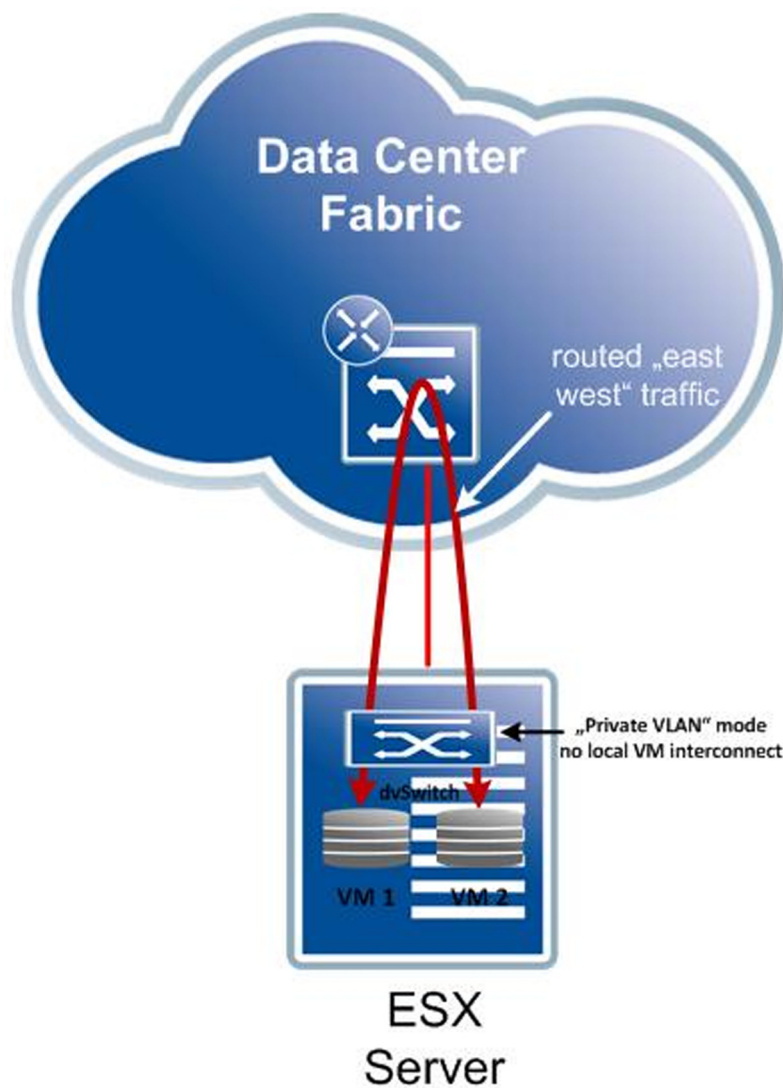


Figure 1: Hair-pin mode using private VLAN

Here's how it works: the DCM component provisions port groups (VLANs) on distributed vSwitches as private and isolated, so the VM's cannot communicate with each other. The ARP requests going out are intercepted by the Fabric Routing enabled switch, and the VRRP MAC (Default Gateway) address for that given IP subnet/VLAN is returned. The VM then starts to send packets towards the default gateway (MAC). These packets are intercepted by the Fabric Routing enabled switch and locally routed. Policies that have been defined (access control list, priority, bandwidth control and Quality of Service) are enforced for each application stream of that VM.

## VEPA Virtual Ethernet Port Aggregator

VEPA provides another alternative switching approach to VEB by sending “all” VM-generated traffic to an adjacent “controlling” bridge. VEPA moves the network demarcation back to the physical switch. When a VEPA mode is used, the external switch applies policy and forwarding rules to the VMs and can see all traffic generated by the VMs. If necessary, traffic destined for VMs on the same physical server is returned to that server via a ‘reflective relay’ process commonly known as a “hairpin turn”. For the most efficient delivery of broadcast and multicast, a VEPA replicates the traffic to each VM locally on the server. The implementation of a VEPA is ideally implemented in hardware by the Network Interface Card (NIC) of a server using direct I/O functionality so the VM can access these interfaces, bypassing the hypervisor, and thereby reducing the load on the CPU of the server.

In the VEPA scenario, the VMs are connected to the physical network as closely as possible so that the virtual network interface is directly connected to the physical network interface of the host without requiring a traditional vSwitch. VEPA offers the following advantages:

- Reduced complexity and the potential to enable higher performance by off-loading advanced network functions from the VM or hypervisor to the physical switch.
- Consistent levels of network policy enforcement by routing all network traffic through the adjacent bridge with its more complete policy-enforcement capabilities.
- Visibility of inter-VM traffic is provided to network management tools designed for physical networks.
- Reduced complexity for the network administrator, as less network configuration is required by server administrators.

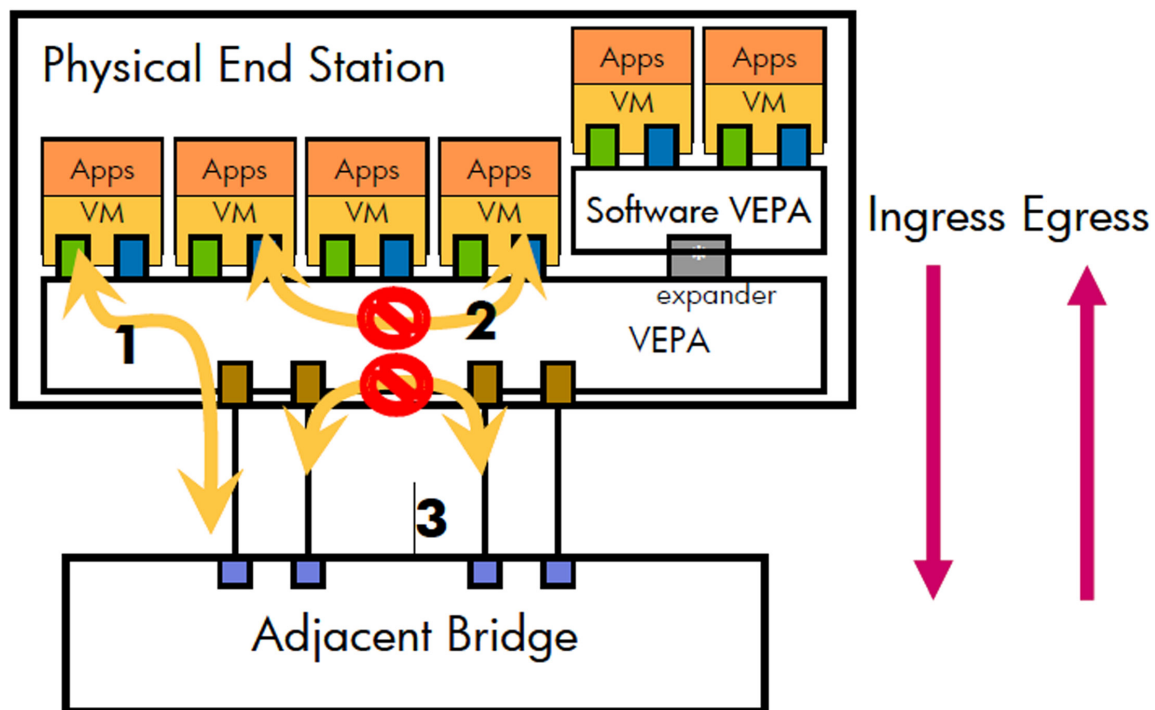


Figure 2: VEPA overview (source: IEEE proposal Apr 2010)

VEPA is on a standards track inside the IEEE but not yet widely deployed in pre-standard implementations or committed to by all major hypervisor vendors. CoreFlow2 based data center switches from Enterasys can support VEPA through a software upgrade and will also support this standard once it has been ratified and accepted in the markets that are most relevant for Enterasys.

## Proprietary Options

Today one can find a series of fabric extension solutions based on blade center and/or other top of rack switches in the market. The implementation is typically tied to a specific hypervisor and tries to achieve the same results as the standards based approaches – sometimes with dedicated NIC cards in the server as well. Often I/O virtualization and consolidation for storage (Fibre Channel) is put into the mix. This approach locks customers into a very specific server, network and hypervisor solution that is leveraging proprietary protocols and results in higher lifetime costs and less flexibility. For these reasons, this approach is generally not recommended.

# Overlay Solutions

To address the operational and organizational challenges which can slow down the adoption of virtualization technologies in today's immature data center network infrastructures, a new series of standards have been proposed. These standards also address the challenges faced by service providers who are required to rapidly scale to meet the needs of these new multi-tenancy data centers.

In general, the proposed solutions are using tunneling or overlay techniques between the servers/hypervisor to become network infrastructure “agnostic”. There is also the option to implement these standards inside the physical network for large scale deployments, but in enterprises, this is expected to be used as an overlay model.

A data center built upon the Enterasys OneFabric architecture already addresses these challenges for enterprise customers. While there is no immediate need to support these standards, Enterasys CoreFlow2 based products can provide value in these environments as highlighted in the following sections.

## VXLAN

To highlight the goals of this proposed standard, here is an excerpt from the experimental IETF draft of VXLAN from August 2011:

*“Server virtualization has placed increased demands on the physical network infrastructure. At a minimum, there is a need for more MAC address table entries throughout the switched Ethernet network due to potential attachment of hundreds of thousands of Virtual Machines (VMs), each with its own MAC address. Second, the VMs may be grouped according to their Virtual LAN (VLAN). In a data center one might need thousands of VLANs to partition the traffic according to the specific group that the VM may belong to. The current VLAN limit of 4094 is inadequate in such situations. A related requirement for virtualized environments is having the Layer 2 network scale across the entire data center or even between data centers for efficient allocation of compute, network and storage resources. Using traditional approaches like Spanning Tree Protocol (STP) for a loop free topology can result in a large number of disabled links in such environments. Another type of demand that is being placed on data centers is the need to host multiple tenants, each with their own isolated network domain. This is not economical to realize with dedicated infrastructure, so network administrators opt to implement this over a shared network. A concomitant problem is that each tenant may independently assign MAC addresses and VLAN IDs leading to potential duplication of these on the physical network. The last scenario is the case where the network operator prefers to use IP for interconnection of the physical infrastructure (e.g. to achieve multipath scalability through Equal Cost Multipath [ECMP]) while still preserving the Layer 2 model for inter-VM communication.”*

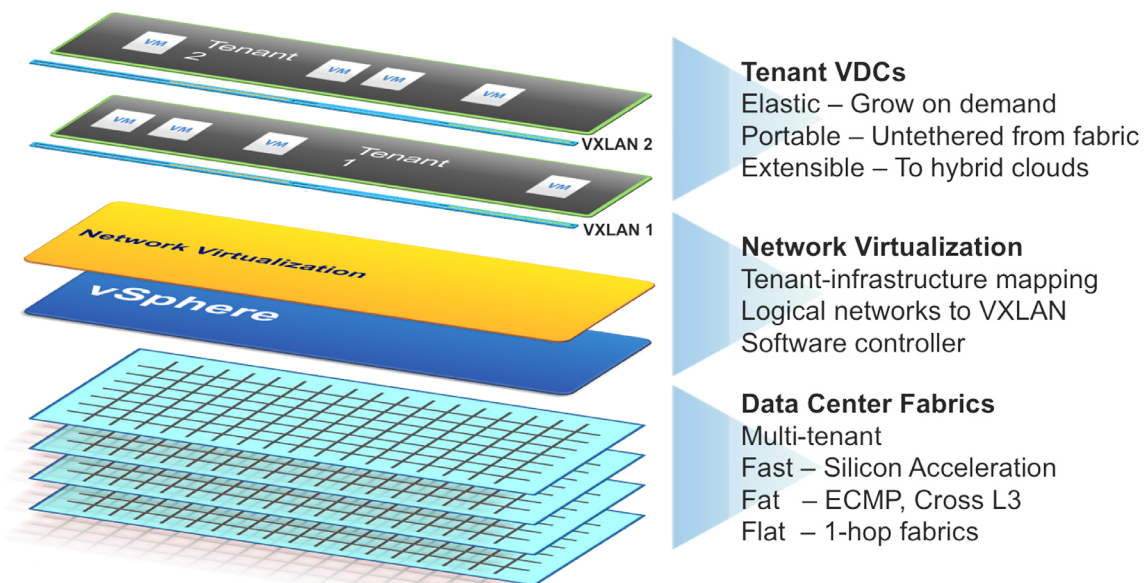


Figure 3 : VXLAN in the VMWare Cloud Stack (source: VMWare)

It becomes obvious that this standard is largely targeted towards a different deployment than an enterprise data center or private cloud deployment. But let's analyze each of the drivers for the proposed standard separately:

- Need for more MAC addresses – currently typical data center switches support between 16k and 64k addresses; this is typically more than sufficient for an enterprise deployment.
- Grouping of VMs in VLANs – the scale of 4k VLANs is sufficient for separation even in large scale enterprise data centers which require multi-tenancy for different business units and segmentation.
- Layer 2 network scale – this is a requirement today in order to leverage advanced features like vMotion, Dynamic Resource Scheduling (DRS) and others. Solutions like virtual switching are available today to address the need to build larger scale layer 2 networks without STP – where MSTP is still an option in some cases. The OneFabric Data Center architecture will also support IEEE SPB Shortest Path Bridging 802.1aq to address this requirement. For transit of Layer 3 core or service provider networks, other options such as GRE/L2 exist today and they will also be supported as part of the OneFabric architecture.
- Multi-tenancy – this is also addressed today with VLANs, VRF in a typical enterprise data center, and is part of the OneFabric Data Center architecture.
- IP interconnect – this can be still achieved with existing tunneling techniques.

The analysis above shows that in an enterprise data center VXLAN is typically not applicable but might still be considered by the server administration teams because of its overlay nature. Another problem that is created by these overlay techniques is the fact that the physical network infrastructure becomes “blind” to the traffic patterns and flows. If congestion occurs and traffic management is required, these solutions fall short. With Enterasys, the CoreFlow2 based products stand ready to provide visibility into VXLAN tunnels in the event that overlay model acceptance and adoption should ever arise. Enterasys data center switches, based on CoreFlow2 technology, are the only switches on the market today that can claim hardware support for VXLAN – allowing enterprises to protect and maximize their existing investments.

## NVGRE – Network Virtualization using Generic Routing Encapsulation

This proposed standard from IETF Informational Status highlights similar challenges as seen with VXLAN and seeks to address them with a tunneled proposal. Here is an excerpt from the September 2011 proposal:

*“Conventional data center network designs cater to largely static workloads and cause fragmentation of network and server capacity...Layer-2 networks use Rapid Spanning Tree Protocol (RSTP) which is designed to eliminate loops by blocking redundant paths. These eliminated paths translate to wasted capacity and a highly oversubscribed network. There are alternative approaches...network utilization inefficiencies are exacerbated by network fragmentation due to the use of VLANs for broadcast isolation...The current VLAN limits theoretically allow for 4K such subnets to be created...The 4K VLAN limit is no longer sufficient in a shared infrastructure servicing multiple tenants...In order to achieve efficiency it should be possible to assign workloads that operate in a single Layer-2 network to any server in any rack in the network. It should also be possible to migrate workloads to any server anywhere in the network while retaining the workload's addresses. This can be achieved today by stretching VLANs however when workloads migrate the network needs to be reconfigured which is typically error prone.”*

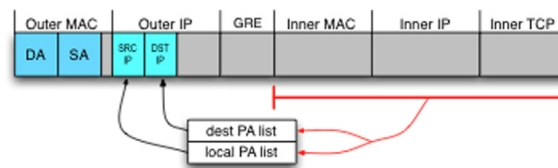


Figure 4: NVGRE header

Highlighted challenges, such as the reconfiguration issues and the potential errors associated with it, are addressed by today's OneFabric Data Center architecture. OneFabric eliminates the need for manual reconfiguration, and the errors that go along with it, by leveraging the integration of the OneFabric Control Center with all leading hypervisor management solutions to automate the process of provisioning and re-provisioning access for virtual machines on the network. A data center network built on the OneFabric architecture is virtualization aware and automatically adapts to the changes that occurs at its edge.

As in the VXLAN environment, the challenge of visibility and control inside the network with a tunneling technique remains and can also be addressed with CoreFlow2 based products in the future.

---

## Summary

Driven by server virtualization and its requirement of the integration of the physical and virtual network infrastructure, a series of proposed solutions are provided by the industry. While a diverse set of options have emerged, most of them are either not ready for deployment, are highly proprietary or have an unknown and questionable lifetime. By leveraging the intelligence embedded in Enterasys solutions, and integrating with your virtual infrastructure, OneFabric delivers an innovative architecture that provides simplicity, scalability and control for complex IT environments. OneFabric allows enterprises to deploy viable virtualization solutions today while being open to new standards-based solutions in the future.

For more information on Enterasys OneFabric please visit [www.onefabric.net](http://www.onefabric.net).

## Contact Us

For more information, call Enterasys Networks toll free at **1-877-801-7082**, or +1-978-684-1000 and visit us on the Web at [enterasys.com](http://enterasys.com)



© 2011 Enterasys Networks, Inc. All rights reserved. Enterasys Networks reserves the right to change specifications without notice. Please contact your representative to confirm current specifications. Please visit <http://www.enterasys.com/company/trademarks.aspx> for trademark information.

