Building a World-class Data Center Network Based on Open Standards

An IBM® Redbooks® Point-of-View publication from IBM System Networking

By Casimer DeCusatis, Ph.D.,
IBM Distinguished Engineer and CTO,
System Networking Strategic Alliances

Highlights

An open data center with an interoperable network delivers noticeable benefits:

- Scales to tens of thousands of virtual machines, reducing capital and operating expense 15 - 25 percent
- Optimizes utilization by virtualizing network resources and moving virtual machines freely within and between data centers
- Integrates pre-optimized solutions seamlessly with the rest of your data center
- Automates network provisioning, management, and troubleshooting so that it becomes more agile and less error prone
- Creates flatter, application-aware networks that need to be wired once

Transforming the data center network to create business value

Data centers are undergoing a major transition toward a smarter, more dynamic infrastructure. More flexible IT architectures are emerging to address the demands of current business applications and new areas, such as cloud computing, multitenancy, bring your own device (BYOD), big data, and analytics.

As part of the dynamic infrastructure trend, the role of data center networks is also changing. It is causing businesses to re-evaluate their current networks, which were never designed to handle modern workloads and applications. Many years ago, Ethernet was developed as a campus networking technology that interconnected “stations” (dumb terminals) through a collection of repeaters, hubs, and switches. Traditional multi-tier Ethernet networks (with access, aggregation, and core layers) were well-suited to campus networks. In these networks, the network configuration did not change much over time, and traffic patterns between users were unaffected by latency and oversubscription.

In this environment, network designs pushed Layer 2 services below the access layer. Also, protocols, such as Spanning Tree Protocol (STP), were sufficient to provide a limited set of topologies, even if they disabled links between switches in the process. This approach worked so well that the same practices were adopted for networks within the data center.

However, traditional multi-tier Ethernet networks are not well-suited to the requirements of modern data centers. Today, servers are highly virtualized, with tens to hundreds of virtual machines (VMs) on a single physical server. Because these VMs can be created, resized, moved, or destroyed at any time, the network must become more flexible and dynamic. This concept is essential to public and private cloud computing, in addition to more conventional applications that require multitenancy.

Unfortunately, the network is not virtualized to the same degree as the servers or storage. Provisioning new functions can take days or weeks for highly skilled network engineers, and often involves moving cables within the data center. The sheer number of VMs in this environment also demands that the network scales in a cost-effective and energy-effective manner.

The data center network also has many other issues. Traffic patterns have changed, with up to 75 percent of data communication occurring between servers in many data centers (so-called *east-west traffic*). The resulting changes in latency and oversubscription can significantly affect application performance. Cost pressures and shrinking margins favor converging different application-specific networks into a common fabric. Finally, organizations are seeking new ways to deploy applications quickly and reliably, with sophisticated features such as quality of service and resource pooling. Faced with this bewildering set of new applications, many organizations force themselves to spend limited resources to keep the current network running, rather than deploying new applications that can generate more revenue streams.

In light of these changes, the networking industry is experiencing a major discontinuity. Older technologies are being supplanted by a new way of doing business. Network virtualization, integration, automation, and convergence are the next big frontiers for the industry. Network virtualization technologies, such as software-defined networking (SDN) and OpenFlow, are combined with network overlays to enable feature-rich functions in management, isolation, and optimization of selected traffic flows. As networks evolve, attention will shift toward application-aware networks. This shift will involve far-reaching and fundamental changes to the entire data center infrastructure and require a roadmap to gradually implement such changes without impacting day-to-day business operations.

Just as you might expect when conquering a new frontier, you must consider many paths. The networking industry has responded to these problems with an array of new technologies that are industry standard and vendor proprietary. Regrettably, not all vendors use compatible technologies, making it difficult to evaluate and implement appropriate networking solutions without becoming locked into a single vendor. This confusing list of options makes the network difficult to manage and almost impossible to optimize.

Although it might be tempting to hand off the entire network to someone else, industry analysts have debunked this strategy. In reality, failure to implement a multivendor approach means paying 15 – 25 percent more than necessary for networking services. In fact, according to a recent Gartner Group report,[1] CIOs who do not re-evaluate long-held incumbent vendor decisions are not living up to their fiduciary responsibilities.

The IBM viewpoint is that the most cost-effective approach to building a best-of-breed data center network is based on open industry standards that are supported by multiple vendors.

To help guide organizations through the current networking discontinuity and transform their legacy data centers, IBM proposed an end-to-end data center network reference architecture that is based on open industry standards. Known as the *Open Data Center Interoperable Network* (ODIN), this approach takes advantage of existing industry standards and best practices to create world-class data center networks with multivendor interoperability. ODIN is not a new standard, but rather uses existing standards and best practices from standards bodies, such as the IEEE, IETF, INCITS, IBTA, and ONF. IBM announced a series of five technical briefs about ODIN that are available from the System Networking website at:

http://www.ibm.com/systems/networking/solutions/odin.html

The ODIN initiative has been well received since its introduction at InterOp 2012. To date, ODIN has been publicly endorsed by many industry leaders, including Juniper Networks, Brocade, Huawei, NEC, Adva, Ciena, BigSwitch, Extreme Networks, and Marist College. IBM has developed a strong system of partners to help implement this vision of an open, interoperable network.

The next-generation data center fabric

An explosion of activity has occurred in the past few years in data networking, with dozens of new standards and architectures being proposed, on a scale unprecedented since the introduction of the Internet. It is important to understand not just what these standards can do for your business, but how they interact with each other when you deploy them in an end-to-end solution.

ODIN addresses many of the key problems faced by modern data networks, including automation, integration, and management. Compared to classical Ethernet architectures, ODIN fabrics have several distinguishing benefits.

As shown in Figure 1 on page 3, ODIN provides a flatter network that reduces latency and improves performance. Larger layer 2 domains support predominantly east-west traffic. By facilitating VM mobility, they enable load balancing across multiple servers and storage. Arbitrary topologies and east-west traffic patterns are enabled. They range from switch stacking and link aggregation techniques to meshed, Clos networks, based on IETF control plane IS-IS algorithms that replace STP with TRILL.

[1] *Debunking the Myth of the Single-Vendor Network*, Gartner RAS Core Research Note, 17 November, 2010.

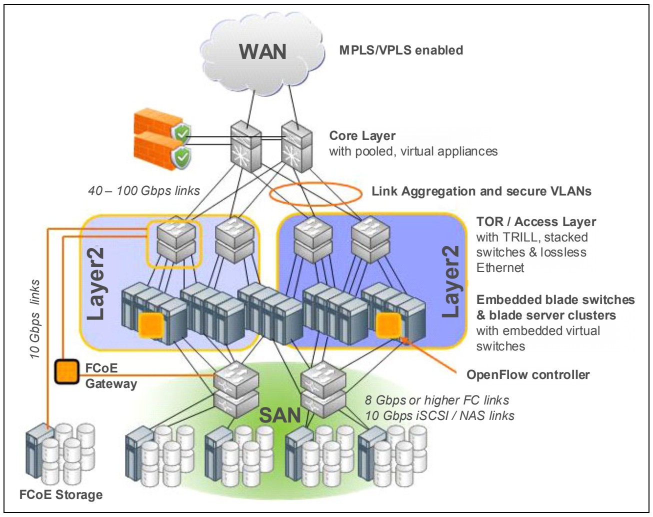

Figure 1 Key characteristics of ODIN

In addition, when using SDN, a network controller (for example, OpenFlow) discovers the switches and creates the topology. If a device or cable fails, its nearest neighbors report the event, and the controller routes around the failure with fast reconvergence times. Existing storage is supported through bridges to the SAN (Fibre Channel or iSCSI) and enables the use of lossless Ethernet and converged FCoE. Physically disjoint paths for high availability or storage access are enabled automatically, simplifying management and provisioning.

Data flow is optimized for greater efficiency. Traffic between two VMs on the same server does not have to leave the server. A virtualized network allows for multitenancy that extends back to the server hypervisor or across multiple sites that are hundreds of kilometers apart. This type of network uses a virtual overlay technology called *Distributed Overlay Virtual Ethernet* (DOVE). DOVE is useful for public and private cloud services, and can enable the consolidation of multiple data centers after a merger or acquisition.

Network functions are delivered as a service, including multitenant VM security and load balancing. Therefore, new functions can be added to the network much more quickly, because reusable software applications are on the network controller, a VM guest, or hardware. A dynamic network means that bandwidth can be managed depending on application needs. For example, it can automatically combine small amounts of available bandwidth into a large, virtual connection to handle high traffic spikes.

As you might expect from an open standard, ODIN offers a choice of blade or rack mounted servers (x86, IBM POWER®, or IBM System z®), virtual or physical switches, and operating systems or hypervisors. ODIN allows you flexibility to tailor an end-to-end solution around your application requirements. ODIN also addresses the interaction between the standards listed in Figure 1. It provides a nondisruptive migration path from your current network and an end-to-end approach that extends from the server hypervisor to the metropolitan area network (MAN) or wide area network (WAN). ODIN promotes networking innovation by enabling open application developers to automate key tasks, such as provisioning new services.

The IBM viewpoint is that data center networks are going through a discontinuity, or period of highly disruptive change, in which network virtualization, automation, integration, and convergence will be the next big frontier.

A smarter data center infrastructure by using ODIN

The ODIN reference designs address three key issues in the transformation of data center networks, as illustrated in Figure 2.

Integrated	**Simple, consolidated** management **Network service** agility **Software-defined networking** platform
Automated	**Application-aware** networking **Dynamic provisioning** **Wire-once** fabric
Optimized	**Converged** (FCoE, iSCSI, NAS, RDMA) **Single, flat** (one managed switch cluster) **Secure, grow as you need** architecture

Figure 2 Components of a smarter data center infrastructure

First, existing networks are built up from a collection of discrete switches, routers, adapters, and other devices, each with its own element manager. Such configurations result in excessively complex network designs that do not scale well. They have multiple management domains and suboptimal point services that are delivered by each network device. ODIN leads to an integrated solution that combines servers, networking, and storage with their respective management domains.

Second, traditional networks are static configurations that require extensive manual intervention for provisioning, maintenance, and troubleshooting. Manual intervention makes network security and service quality susceptible to human error. Further, legacy networks were never intended to support multitenant cloud-computing applications. ODIN promotes an automated, dynamic, virtualized network based on industry standard SDN. An ODIN fabric does not require cables to be reconfigured when new services are brought online. In fact, this automated network is application-aware, meaning that it can adapt to changes in workload, security threats, and other factors in an agile business environment.

Third, an organization has too many conventional networks, each one over-provisioned with too many switches, tiers, and appliances. This approach wastes capital equipment, takes up too much space, uses too much energy, and fails to deliver low latency and efficient performance. ODIN is optimized for a flatter fabric that handles traffic patterns in a modern data center. Multiple fabric types can be converged and consolidated over time, enabling an efficient “pay as you grow” approach for the network. By providing an integrated, automated, and optimized fabric, ODIN delivers a more elastic, scalable data infrastructure that improves end-to-end application performance and reduces capital and operational expenses.

Which platform to choose

IBM and our partners offer a range of ODIN-compliant solutions for your data center network. You do not have to transform your data center overnight. Various convenient starting points are available depending on your immediate interests. For example, you might want to lower capital expenditure and operating expense through server integration, back up your important data over the WAN, or just experiment with a proof of concept based on your particular needs.

The IBM PureSystems™ product line is a good example of how network integration can add value to your business. By providing factory tested and preconfigured servers, storage, and networking, IBM can reduce deployment time from weeks or months to one day or less than a day. PureSystems makes an ideal entry-level cloud-computing platform, hosting up to 54 percent more VMs per compute node than previous options.

As you might expect from an open data center, PureSystems is all about flexibility. You can select the type of server, operating systems, hypervisor, or other attributes to suit your needs, rather than being locked into a single solution. The PureSystems cluster or *integrated system* includes virtual networking (IBM Distributed Virtual Switch 5000V), blade chassis embedded switches, and top-of-rack switches (IBM RackSwitch™ G8264). All your physical and virtual resources can be managed from one interface with Flex System Manager, and the networking within PureSystems supports ODIN. The PureSystems internal network is pre-integrated to provide the lowest possible latency and maximize VM mobility. It is also interoperability tested with a growing number of IBM networking partners for easy integration with your existing data center resources.

As another example, storage backup and VM mobility over distance were pioneered many years ago by the IBM SAN Volume Controller. VM mobility not only improves application availability, it provides a more efficient way to use limited storage resources. VM mobility is useful for business continuity or disaster avoidance or recovery, including planned events such as migrating one data center to another or eliminating downtime due to scheduled maintenance.

Considering an increasingly global work force, this approach also provides load balancing and enhanced user performance across multiple time zones (the so-called “follow the sun” approach). Furthermore, by moving workloads over distance, you can optimize the cost of power consumption because the lowest cost electricity is usually available at night, known as “follow the moon”. Today, SAN Volume Controller stretch clusters support ODIN features including lossless Ethernet and industry standard storage (Fibre Channel and iSCSI). Distance extensions up to 300 km are available by using dark fiber Wavelength Division Multiplexing (WDM), FC-IP, or Multi-Protocol Label Switching (MPLS) or Virtual Private LAN Service (VPLS) options. Also, multiple vendor partners are tested and qualified by IBM.

IBM has already designed and deployed next-generation networking solutions for clients in multiple industries, including healthcare, financial services, cloud providers, and telecommunications. For example, IBM is actively contributing to SDN standards and protocols. As an inaugural member of the Open Networking Forum standards body, IBM is committed to the adoption of open, industry standard SDN without vendor proprietary extensions.

IBM’s SDN leadership is exemplified by being first to market with solutions, such as the first commercially available OpenFlow enabled switch, the 10/40 Gbe IBM RackSwitch G8264. IBM is working with clients across many industries, including leading financial institutions on Wall Street, to enable SDN and ODIN solutions. IBM is also helping global telecommunications companies redefine their networks to be cloud-centric by using this technology. Business solutions in many areas have been qualified in partnership with other vendors who endorse ODIN.

Next steps: How IBM can help

If your organization wants to learn more about how a multivendor network, based on open standards, can help you create a flexible, application-aware network, as a first step, consider your biggest pain points. Then pick a starting point for your network transformation.

For example, integrated offerings, such as IBM PureSystems, provide a simple way to introduce these concepts. IBM Global Services can help integrate an IBM PureFlex™ or IBM PureApplication™ solution with your existing network. At the other extreme, you can design your own network from building blocks, such as by using IBM 5000V virtual hypervisor switches, blade chassis-embedded switches, and top-of-rack switches. IBM System Networking can provide best practices and propose network architectures that use these products. For more examples of IBM solutions in your industry, see the resources listed in the next section, or contact your local IBM representative.

Resources for more information

For more information about the concepts that are highlighted in this paper, see the following resources:

- Data Center Networking blog

  https://www.ibm.com/connections/blogs/DCN

- Debra Curtis and Mark Fabbi, *Debunking the Myth of the Single-Vendor Network*, Gartner RAS Core Research Note, 17 November, 2010

  http://www.dell.com/downloads/global/products/pwcnt/en/Gartner-Debunking-the-Myth-of-the-Single-Vendor-Network-20101117-published.pdf

- IBM PureSystems, the power of patterns

  http://www.ibm.com/ibm/puresystems/us/en/index.html

- IBM SVC KnowledgeBase for VMware interoperability

  http://kb.vmware.com/selfservice/microsites/search.do?language=en_US&cmd=displayKC&externalId=2032346

- Juniper Networks interoperability guides for IBM networking solutions

  http://kb.juniper.net/InfoCenter/index?page=answers&type=search&searchid=1343918845874&question_box=Qfabric+storage+interoperability+guide

- Networking Solutions: Open Data Center Interoperable Network (ODIN)

  http://www.ibm.com/systems/networking/solutions/odin.html

- OpenFlow research sponsored by IBM and Marist College:

  http://openflow.marist.edu

- T. Bundy, M. Haley, F. Street, C. DeCusatis, *The impact of data center convergence, virtualization, and cloud on DWDM optical networks both today and into the future*, Proc. Pacific Telecommunications Council 2012 Annual Meeting, Honolulu, Hawaii, (January 2012)

- C. DeCusatis, *Enterprise networks for low latency, high frequency financial trading*, Proc. Enterprise Computing Community conference, June 12-14, 2011, Marist College, Poughkeepsie, NY (2011)

  http://ecc.marist.edu/conf2011

- C. DeCusatis, *Towards an open data center with an interoperable network: enterprise networking using open industry standards*, Proc. National Science Foundation Enterprise Computing Conference, Marist College (June 11-13, 2012)

- C. DeCusatis and B. Larson, *VM mobility over distance, leveraging IBM SVC split cluster with VMWare/VSphere*, Proc. IBM Storage Edge conference, Orlando, FL (June 4-7, 2012)

Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:
IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785 U.S.A.

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs.

This document, REDP-4933-00, was created or updated on October 3, 2012.

IBM®

Redbooks®

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. These and other IBM trademarked terms are marked on their first occurrence in this information with the appropriate symbol (or), indicating US registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at `http://www.ibm.com/legal/copytrade.shtml`

The following terms are trademarks of the International Business Machines Corporation in the United States, other countries, or both:
IBM®
POWER®
PureApplication™
PureFlex™
PureSystems™
RackSwitch™
Redbooks®
Redbooks (logo) ®
System z®

The following terms are trademarks of other companies:

Other company, product, or service names may be trademarks or service marks of others.

