IDC
Analyze the Future

IDC TECHNOLOGY SPOTLIGHT

# The New Need for Speed in the Datacenter Network

*March 2015*

Adapted from *Worldwide Datacenter Network 2014–2018 Forecast and Analysis* by Brad Casemore, Petr Jirovsky, Rohit Mehra, IDC #248757, and *Market Analysis Perspective: Datacenter Networks, 2014* by Brad Casemore, Petr Jirovsky, Rohit Mehra, IDC #253190

Sponsored by Cisco

*Today's datacenter networks must better adapt to and accommodate business-critical application workloads. Datacenters will have to increasingly adapt to virtualized workloads and to the ongoing enterprise transition to private and hybrid clouds. Pressure will mount on datacenters not only to provide increased bandwidth for 3rd Platform applications such as cloud and data analytics but also to deliver the agility and dynamism necessary to accommodate shifting traffic patterns (with more east-west traffic associated with server-to-server flows, as opposed to the traditional north-south traffic associated with client/server computing). Private cloud and legacy applications will also drive daunting bandwidth and connectivity requirements. This Technology Spotlight examines the increasing bandwidth requirements in enterprise datacenters, driven by both new and old application workloads, cloud and noncloud in nature. It also looks at how Cisco is meeting the bandwidth challenge posed by 3rd Platform application workloads in the datacenter.*

## Introduction

The 3rd Platform — comprising cloud, mobility, data analytics, and social business technologies — is providing a robust foundation for enterprise business model and business process transformation. Enterprises are increasingly leveraging the 3rd Platform to better engage with their customers, to expedite the development and delivery of products and services, to innovate continually, and to run their businesses and their IT operations more efficiently.

What is the 3rd Platform? IDC introduced the concept in 2007, framing it as a software-based, next-generation IT platform, succeeding the mainframe (1st Platform) and client/server (2nd Platform) eras. The 1st and 2nd Platforms were generational, hugely influential, and notably disruptive. The 3rd Platform will be equally, if not more, disruptive. In 2015, IDC expects the 3rd Platform to account for one-third of global information and communication technology (ICT) spending and 100% of spending growth (see Figure 1).

That said, the core elements of the 3rd Platform are driving a pronounced need for increased bandwidth in the datacenter. Virtualization, a precursor and integral component of cloud computing, has had far-reaching bandwidth implications, driving the need for greater bandwidth capacity on both leaf and spine switches in Clos architectures. This need is further heightened by the heavily east-west (within and between servers) traffic patterns of virtualization, which are markedly different from the traditional north-south traffic flows associated with client/server applications.

**Figure 1**

Technology-Fueled Business Transformation Shifts Technology Buying Centers

The 3rd Platform creates the underpinnings for business process transformation…and, increasingly, business model transformation:

- How enterprises **engage** with customers
- The **speed** at which they deliver their products and services
- How they **innovate**
- Their **resiliency**
- The **reliability** of their operations

With such high stakes, the business is increasingly driving technology initiatives.


**Inversion of business from *make-sell* to *sense-and-respond***

Source: IDC, 2015

Although a growing number of large enterprises are transitioning to hybrid cloud, using SaaS applications where appropriate and also considering, if not tapping into, IaaS and PaaS options, most enterprises will retain legacy application workloads in their own datacenters as well as virtualized private cloud applications that they will run internally, for compliance or other reasons. Perhaps overlooked amid the growing market interest in public cloud, these private cloud and legacy applications will continue to drive daunting bandwidth and connectivity requirements, demanding their own bandwidth upgrades in enterprise datacenters.

In this context, it's worth noting that some applications initially spun up in the public cloud, for time-to-revenue or other reasons, also are subject to being brought in-house as they become more important to the business and reach a point where they can be run more cost effectively in-house than in the public cloud.

From previous market studies and surveys, IDC knows that network refreshes and extensions in the datacenter are being driven by the desire to improve application performance, by the need for the network to accommodate the capacity demands of new applications (including private cloud and big data analytics), and by the ongoing need to accommodate increased use of existing legacy applications.

Taken together, these needs will require enterprise IT to plan for increasing bandwidth requirements in their datacenters, driven by both new and old application workloads, cloud and noncloud in nature.

## Workloads Driving Bandwidth Requirements in Enterprise Datacenter

There's no question that virtualization continues to expand and drive the need for greater datacenter bandwidth, first at the top-of-rack (ToR) switch and subsequently in the datacenter core.

According to IDC, the typical virtualized server supported about 10 virtual machines (VMs) in 2014 and will support in excess of 12 VMs by 2017. In many organizations, the majority of production workloads are already virtualized and almost all new workloads are deployed on virtualized infrastructure, placing inexorable stress on server connectivity.

On the horizon is increased use of Docker and other container engines. The bandwidth implications of containers could be considerably greater than those associated with virtualization. Containers are inherently lightweight, allowing for greater density per server than hypervisors, as well as faster application start-ups and shutdowns. A well-tuned container system can accommodate as many as four to six times the number of server application instances that could be supported by a hypervisor running on the same system. Given the densities per server of containers — which transcend those of hypervisor-based VMs — potentially significant strain will be placed on the vSwitch at the network edge. If the datacenter network is to avoid becoming a bottleneck to IT innovation and business agility, physical leaf-spine architectures must account for burgeoning volumes of application traffic.

Big data and increased use of data analytics will contribute to the bandwidth deluge. Big data, as exemplified by Hadoop, requires high-performance compute, network, and storage infrastructure. While many early big data deployments were networked at 1Gbps, most of those deployments graduated quickly to 10GbE connections to the server and are straining those limitations and pushing toward a need for greater bandwidth, whether 25GbE (when commercially available) or 40GbE.

Mobility, too, is adding significantly to the data deluge. IDC has forecast that smart connected devices will number more than 2 billion by 2016 and expects continued growth through 2020 and beyond. IDC also anticipates that each user of a mobile device will generate about 5.2 terabytes of data, feeding the big data flood to the datacenter and putting further strain on the underlying network.

Just beginning to make an impression is the relatively nascent market represented by the Internet of Things (IoT), which will spawn over 28 billion connected devices by 2020. IDC forecasts that about 50% of IT networks will be bandwidth constrained as a result of the flood of data from the rapid proliferation of IoT devices. Moreover, IDC believes that about 10% of sites will be overwhelmed by an IoT data deluge. These organizations will have to include IoT ramifications in their planning and preparation for datacenter network upgrades.

Then there are traditional datacenter workloads, legacy applications that will remain within the enterprise datacenter for the foreseeable future. These applications will continue to be used and will necessitate incrementally more bandwidth in the years to come. Collaborative applications, such as bandwidth-intensive videoconferencing and unified communications, clearly will have a need for ever-growing datacenter pipes. In a recent IDC survey, about 86% of respondents indicated that they are using or plan to use videoconferencing solutions, while a similar percentage indicated that they are using or plan to use unified communications.

In addition, enterprise applications, such as business-related data processing (ERP, CRM, OLTP, batch), decision support (data warehousing and data mining), collaboration (email and workgroup applications), and in-house application development, will continue to grow and require incremental additional bandwidth.

All of the aforementioned application workloads will place increasing bandwidth demands on the datacenter network, with virtualization (and containerization) expected to be the biggest driver for bandwidth upgrades. The need for greater bandwidth will be felt first at the access layer, where 1GbE server connections will be supplanted by 10GbE and 25GbE connectivity. In turn, the higher bandwidths

at the access layer will drive the need for greater upstream bandwidth, with 40GbE and 100GbE connectivity meeting the challenge. In 2014, 10GbE and 40GbE accounted for shipments of more than 20 million ports, generating revenue of about $7 billion. By 2018, 10–40GbE, including the emergent 25GbE option, will account for more than 90% of spending on Ethernet switches in the datacenter.

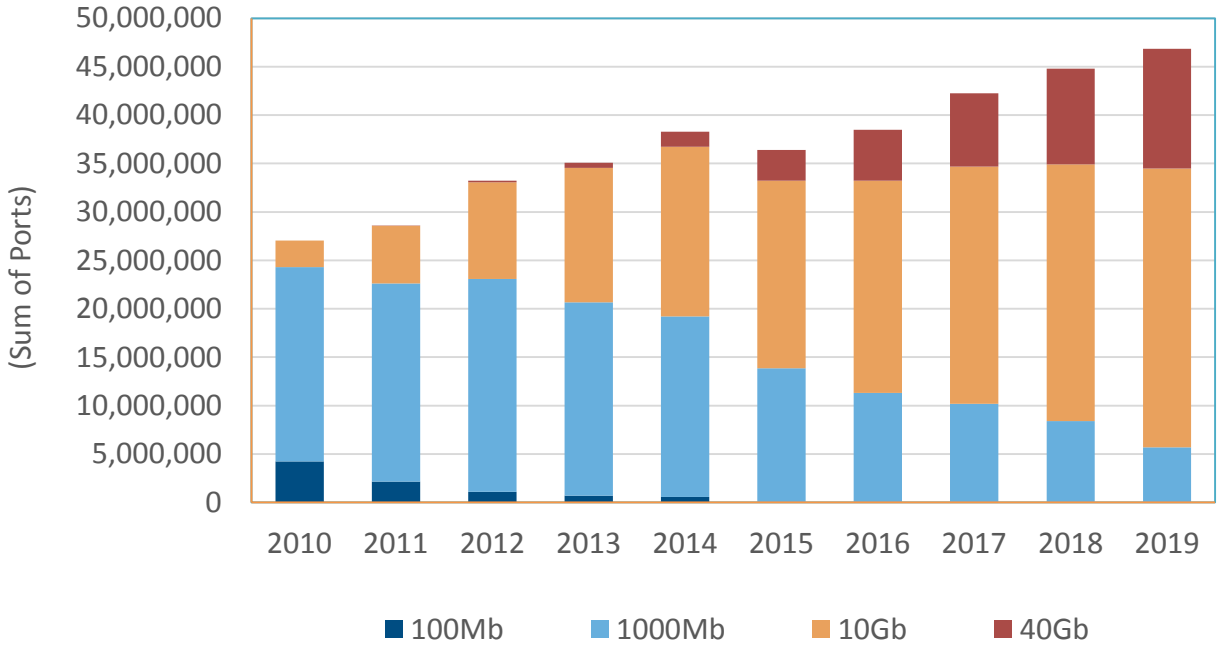
## Considering Bandwidth Progressions and Alternatives

When it comes to meeting mounting bandwidth requirements, enterprise datacenters have an increasing number of options. For enterprises using 10GbE ToR switches today, the upgrade path leads to 40GbE, with 100GbE spine uplinks.

However, other alternatives are on the near-term horizon. Major cloud service providers, as well as many large enterprises, envision a datacenter bandwidth migration path that could be served by rate progressions that include 10GbE, 25GbE, 40GbE, 50GbE, 100GbE, and — perhaps not so finally — 400GbE. Given the needs of cloud service providers, the prospects for 100GbE seem bright.

**Figure 2**

Strong Port Growth for 10GbE and 40 GbE




The relatively recent specifications for 25Gbps and 50Gbps Ethernet address the need not only for greater bandwidth but also for scalable, cost-effective connectivity infrastructure, translating into lower cost per unit of bandwidth. Essentially, major datacenter operators in the cloud are seeking more bandwidth from fewer cables. It helps that the 25 Gigabit Ethernet Consortium's specifications leverage technologies developed for 100Gbps Ethernet, with the 25Gbps specification requiring a single lane and the 50Gbps specification requiring two lanes of the four-lane architecture constituting 100Gbps Ethernet.

## Cisco's Approach to Growing Bandwidth Requirements in the Datacenter

Cisco's contributions to the IEEE standards body have helped form numerous Ethernet link speed standards, ranging from Fast Ethernet (100Mbps) to 100Gbps Ethernet. These standards have provided an incremental transition path to meet the increasing bandwidth demands as technology has evolved.

While 10Gbps Ethernet increasingly is providing server access in the datacenter, Cisco sees that ever-greater demands are being placed on server connectivity as server performance improves and as the number of workloads per virtualized server also increases. This is especially true in highly virtualized farms and cloud environments running scale-out elastic applications through containers, where ToR connectivity is pushing the limits of 10GbE bandwidth. Currently, 40GbE is the next successive link speed, but, as mentioned previously, many large-scale customers with tens and thousands of servers are looking for an intermediary Ethernet speed that could offer a lower price per gigabit for server connectivity, particularly in cases where 40GbE is not required.

Cisco expects broad industry backing for 25Gbps Ethernet to result in fast tracking of the 25GbE standard, with products likely to ship as early as the second half of 2015. Cisco plans to be at the forefront of 25Gbps commercialization and product rollouts.

Where 40GbE is required, Cisco's QSFP 40G BiDi (bidirectional) transceiver helps significantly reduce the cost of migrating from 10GbE to 40GbE. It does so by allowing the reuse of current 10GbE cabling for 40GbE connectivity, thus circumventing the need for a costly upgrade to cabling infrastructure, which — in addition to being cost prohibitive — can be disruptive to datacenter operations.

### *Challenges and Opportunities*

As the market leader in datacenter networking with revenue of nearly $5.6 billion (or more than 50% of an overall revenue pie worth $11.07 billion), Cisco has a large installed base of customers spanning a wide range of organizational sizes and vertical markets. This customer base is obviously an asset, but it presents a challenge, too.

With the inherent heterogeneity of Cisco's customer base, not all customers will have an immediate need to upgrade their datacenter switches for greater bandwidth. Of course, need will be driven by application workloads and traffic volumes, and large customers, as well as those furthest along on their cloud journeys, will be the first to experience an acute need for greater bandwidth at the access layer and the aggregation/core of their networks. Conversely, customers that have yet to aggressively adopt 3rd Platform workloads will not feel a pressing requirement to pursue greater bandwidths.

Moreover, many customers adopting new applications and private cloud might not have an accurate or full appreciation of the bandwidth implications associated with emerging workloads. Cisco and its channel partners must function as trusted partners to articulate the bandwidth ramifications of virtualization, mobility, cloud, data analytics, video-based collaboration, and social business, candidly explaining the consequences of not providing for sufficient bandwidth.

Despite those challenges, Cisco's vast installed base of customers will have a need for greater bandwidth, and Cisco is well placed to seize the market opportunity and meet customer needs. It has an extensive portfolio of datacenter switches for the leaf and spine, and it has a keen understanding of customer requirements. With those requirements in mind, Cisco has sought to provide investment protection to customers by enabling its 10GbE switches to support 40GbE or 100GbE in the future.

## Conclusion

Networks exist to support the application workloads that are critical to business outcomes for enterprises and cloud service providers. As the workload mix transitions from the relatively

predictable north-south traffic patterns and volumes associated with client/server applications to the more demanding east-west traffic patterns and greater bandwidth requirements of 3rd Platform applications such as cloud and data analytics, the network must respond accordingly. This requires the network to accommodate greater bandwidth at the access layer — where the need will grow from 1GbE to 10GbE or 25GbE — and then in the core, where there will be a need for connectivity at 40GbE or 100GbE.

As enterprises and service providers deploy new applications and extend existing applications, they will confront the bandwidth limitations of their current datacenter network infrastructure. Fortunately, specifications and standards have been proposed and are in place to address these requirements, regardless of company size or where customers might find themselves in their adoption of 3rd Platform workloads. For many companies, the need for greater bandwidth is already acute, and for others, it is just a matter of time.

Good planning is always a virtue, and enterprises and service providers can navigate a prudent course by ensuring that their critical applications and services, including increasingly virtualized and even containerized workloads, are accommodated by sufficient bandwidth in their datacenter networks.

---