NEW SERVICE DELIVERY

Cisco Global Cloud Index: Forecast and Methodology

Extract from the Cisco® Global Cloud Index: Forecast and Methodology, 2011–2016 White Paper

The Cisco® Global Cloud Index is an ongoing effort to forecast the growth of global data center and cloud-based IP traffic. The forecast includes trends associated with data center virtualization and cloud computing.

Forecast Overview

Global data center traffic:

- Annual global data center IP traffic will reach 6.6 zettabytes by the end of 2016. By 2016, global data center IP traffic will reach 554 exabytes per month (up from 146 exabytes per month in 2011).
- Global data center IP traffic will nearly quadruple over the next 5 years. Overall, data center IP traffic will grow at a compound annual growth rate (CAGR) of 31 percent from 2011 to 2016.

Data center virtualization and cloud computing transition:

- The number of workloads per installed traditional server will increase from 1.5 in 2011 to 2.0 by 2016.
- The number of workloads per installed cloud server will increase from 4.2 in 2011 to 8.5 by 2016.
- By 2016, nearly two-thirds of all workloads will be processed in the cloud.

Global cloud traffic:

- Annual global cloud IP traffic will reach 4.3 zettabytes by the end of 2016. By 2016, global cloud IP traffic will reach 355 exabytes per month (up from 57 exabytes per month in 2011).
- Global cloud IP traffic will increase six-fold over the next 5 years. Overall, cloud IP traffic will grow at a CAGR of 44 percent from 2011 to 2016.
- Global cloud IP traffic will account for nearly two-thirds of total data center traffic by 2016.

Regional cloud readiness:

- North America and Western Europe led in broadband access (fixed and mobile) in 2011 and will continue to lead in this category through 2016. However, all regions will show measurable improvement in broadband access to their respective populations throughout the forecast period. Asia Pacific leads in the number of subscribers throughout the forecast period due to the region's large population (see Broadband Ubiquity section for details).
- Western Europe leads all regions with an average fixed download speed of 11.7 Mbps. North America follows with an average fixed download speed of 10.3 Mbps. Asia Pacific and Central and Eastern Europe lead all regions in average fixed upload speeds of 6.7 Mbps and 5.9 Mbps, respectively.
- Western Europe and Central and Eastern Europe lead all regions in average fixed network latency with 58 ms and 59 ms, respectively.

Evolution of Data Center Traffic

In just the past year since the first Cisco Global Cloud Index was released, the industry has seen cloud adoption evolving from an emerging technology to an established networking solution that is gaining widespread acceptance and deployment. Enterprise and government organizations are moving from test environments to placing more of their mission-critical workloads in the cloud. And for consumers, cloud services offer ubiquitous access to content and services, on multiple devices, almost anywhere network users are located.

Qualitatively, the main drivers for cloud adoption include faster delivery of services and data, increased application performance, as well as improved operational efficiencies. While security and integration with existing IT environments continue to represent concerns for some potential cloud-

www.intercomms.net InterComms 47

based applications, a growing range of consumer and business cloud services are currently available. Today's cloud services address varying customer requirements (for example, privacy, mobility, and multiple device access) and support near-term opportunities as well as long-term strategic priorities for network operators, both public and private.

Quantitatively, the impact of cloud computing on data center traffic is clear. It is important to recognize that most Internet traffic has originated or terminated in a data center since 2008. Data center traffic will continue to dominate Internet traffic for the foreseeable future, but the nature of data center traffic is undergoing a fundamental transformation brought about by cloud applications, services, and infrastructure. The importance and relevance of the global cloud evolution is highlighted by one of the top-line projections from this updated forecast — by 2016 nearly two-thirds of data center traffic will be cloud traffic.

Global Data Center IP Traffic: Nearly Four-fold Increase by 2016

Figure 1 summarizes the forecast for data center IP traffic growth from 2011 to 2016.

The Internet is forecast to reach the zettabyte era in 2016, but the data center has already entered the zettabyte era. While the amount of traffic crossing the Internet and IP WAN networks is projected to reach 1.3 zettabytes per year in 20161, the amount of data center traffic is already 1.8 zettabytes per year, and by 2016 will nearly quadruple to reach 6.6 zettabytes per year. This represents a 31 percent CAGR. The higher volume of data center traffic is due to the inclusion of traffic inside the data center (typically, definitions of Internet and WAN stop at the boundary of the data center).

The global data center traffic forecast, a major component of the Cisco Global Cloud Index, covers network data centers worldwide operated by service providers as well as private enterprises.

Data Center Traffic Destinations: Most Traffic Remains Within the Data Center

Consumer and business traffic flowing through data centers can be broadly categorized into three main areas (Figure 2):

- Traffic that remains within the data center
- Traffic that flows from data center to data center
- Traffic that flows from the data center to end users through the Internet or IP WAN

The portion of traffic residing within the data center will remain the majority throughout the forecast period, accounting for 76 percent of data center traffic in both 2011 and 2016. Factors contributing to traffic remaining in the data center include functional separation of application servers, storage, and databases, which generates replication, backup, and read/write traffic traversing the data center. Furthermore, parallel processing divides tasks and sends them to multiple servers, contributing to internal data center traffic.

The ratio of traffic exiting the data center to traffic remaining within the data center could be expected to increase over time, because video files are bandwidth-heavy and do not require database or processing traffic commensurate with their file size. However, the increasing use of applications such as desktop virtualization is likely to offset this trend. In addition, the virtualization of storage and other data center capabilities increases traffic within the data center because functions may no longer be local to a rack or server. Table 1 provides details for global data center traffic growth rates.

Figure 1. Global Data Center IP Traffic Growth

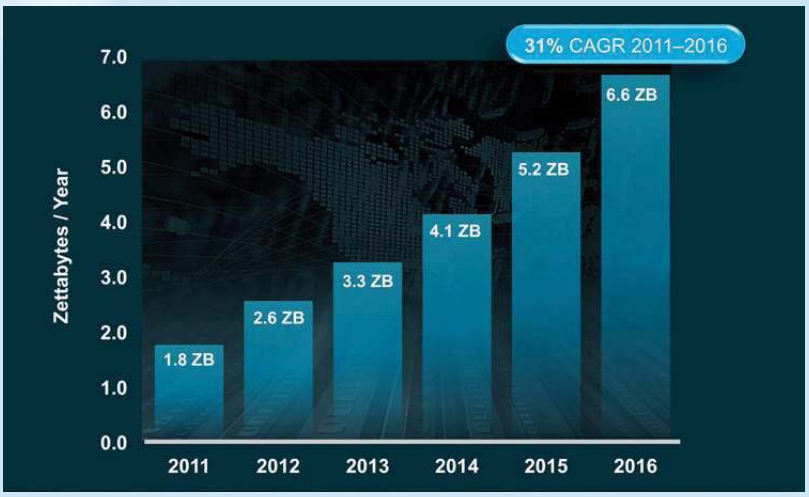

Figure 2. Global Data Center Traffic by Destination

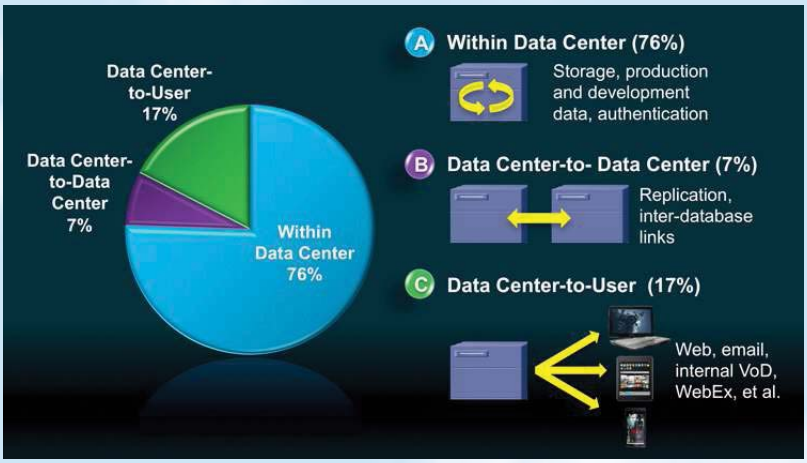

Table 1. Global Data Center Traffic, 2011–2016

Data Center IP Traffic, 2011–2016							
	2011	2012	2013	2014	2015	2016	CAGR 2011–2016
By Type (EB per Year) Data center to user	299	438	561	714	912	160	1,31%
Data center to data center	118	173	222	284	365	468	32%
Within data center	1,338	1,940	2,468	3,126	3,969	5,021	30%
By Segment (EB per Year) Consumer	1,404	2,107	2,700	3,439	4,418	5,672	32%
Business	351	444	551	685	828	977	23%
By Type (EB per Year) Cloud data center	683	1,181	1,694	2,324	3,166	4,255	44%
Traditional data center	1,072	1,370	1,557	1,800	2,080	2,394	17%
Total (EB per Year) Total data center traffic	1,755	2,551	3,251	4,124	5,246	6,649	31%

Definitions:

- **Data center to user:** Traffic that flows from the data center to end users through the Internet or IP WAN
- **Data center to data center:** Traffic that flows from data center to data center
- **Within data center:** Traffic that remains within the data center
- **Consumer:** Traffic originating with or destined for consumer end users
- **Business:** Traffic originating with or destined for business end users
- **Cloud data center:** Traffic associated with cloud consumer and business applications
- **Traditional data center:** Traffic associated with noncloud consumer and business applications
- **Data Center Workloads:** Nearly Two-Thirds Will Be Cloud-Based by 2016

A workload can be defined as the amount of processing a server undertakes to run an application and support a number of users interacting with the application. The Cisco Global Cloud Index forecasts the continued transition of workloads from traditional data centers to cloud data centers. By 2016, nearly two-thirds of all workloads will be processed in cloud data centers (Figure 3). For regional distributions of workloads, see Appendix E.

Growth of workloads in cloud data centers will be five and a half times that of the growth in traditional workloads between 2011 and 2016. Traditionally, one server carried one workload. However, with increasing server computing capacity and virtualization, multiple workloads per physical server are common in cloud architectures. Cloud economics, including server cost, resiliency, scalability, and product lifespan, are promoting migration of workloads across servers, both inside the data center and across data centers (even centers in different geographic areas).

To read/download this White Paper in full, please see:
http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns1175/Cloud_Index_White_Paper.pdf