Big Data Is a Big Deal:
Is Your Network Ready?

Michael Harris
Kinetic Strategies

TIME WARNER CABLE Business Class®

NaviSite®

Big Data Is a Big Deal: Is Your Network Ready?

Big data is a big deal. Not only is it redefining business management, marketing and operations principles, big data is also transforming network performance requirements. Seven out of ten enterprises and more than half of small and medium businesses have already deployed or plan to deploy big data projects.[1] According to management consulting firm McKinsey & Company, "the use of big data is becoming a crucial way for leading companies to outperform their peers."[2] Indeed, research finds that a majority of companies with a strategy focused on collecting and analyzing their most valuable data best their competitors financially.[3]

Consider these big data success stories. A hotel chain now analyzes video monitoring of front desk operations to prevent long check-in lines for its preferred customers. A health insurance firm analyzes speech-to-text data from call center recordings to gain insight into customer dissatisfaction triggers and implement tactics to avoid them. A manufacturer of jet engines is eyeing opportunities to leverage sensor data to help airlines achieve reductions in fuel consumption that could deliver $30 billion in savings over 15 years. A delivery company crunches data to improve vehicle routing, eliminating millions of miles driven annually across its fleet. With big data a department store chain is able to optimize pricing for the millions of items in its sales inventory in only 1 hour, compared to 27 hours several years ago.[4]

70% of enterprises

56% of SMBs

Have already deployed or plan to deploy big data projects

Source: IDG Enterprise

Big Data Defined

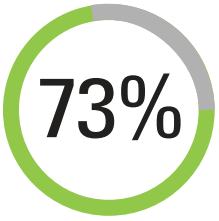
So what exactly is big data? As the name implies, big data involves collecting, storing and analyzing huge quantities of business information from a wide range of devices and applications. This includes real-time and historical data from customer relationship, sales, manufacturing, supply chain and retail sales systems, as well as social media, website and other network usage information. Data can be collected from an array of devices, including servers, PCs, smartphones, tablets, video cameras, sensors, smart meters and other connected machines. IT professionals say the average amount of data being managed within their organization is expected to increase by 76 percent within the next 12-18 months.[5]

76%

Increase in amount of data being managed over 12-18 months

Source: IDG Enterprise

By skillfully surfing this data tsunami, advanced analytics solutions can deliver powerful business intelligence. More than three out of four companies report that their current big data and analytics projects are delivering quantified benefits that either meet or exceed expectations.[6] Big data helps organizations fine-tune product and service offerings, enables more personalized interactions with customers and prospects, creates efficiencies in manufacturing, better detects fraud and improves regulatory compliance.



of companies have increased revenue with big data

Source: Wakefield Research

A recent survey finds that 84 percent of companies say having more data has helped them make better business decisions. This trend is driving the bottom line as 73 percent of organizations say leveraging data has increased revenue from existing or new sources.[7] Nearly half of IT executives say their company CEO is directly supporting or sponsoring their efforts, highlighting the visibility and strategic importance of big data initiatives.[8]

Big Data Types

Structured Data

- CRM, ERP, POS
- Database
- Data warehouse
- Spreadsheets

Unstructured Data

- Email, tweets
- WP documents
- Telephone calls
- Web, social media
- Clickstreams
- Photos, audio, video
- Sensor and device data

Managing the Three Vs

Organizations eager to create business value from big data must manage three Vs of information flows: variety, velocity and volume. On the variety front, big data analysis means crunching information from both structured and unstructured sources.

Structured data is arranged with organized fields in sources such as databases, spreadsheets, customer relationship management (CRM), enterprise resource planning (ERP) and point of sale (POS) systems. Researchers estimate structured data accounts for only 10 percent of the digital universe.[9]

The vast majority of available information for analysis is unstructured data. Included in this category are email and text messages, telephone conversations, word processing documents, web and social media content, sensor and device data, plus photos, audio and video. Given its unruly nature, crunching and analyzing unstructured data poses a greater challenge for organizations. However, given the opportunities available to gain insight into operations and manufacturing, as well as customer and prospect behavior and needs, unstructured data is high priority for big data initiatives. A recent survey of IT managers finds 84 percent are currently analyzing unstructured data. Among the minority who are not yet doing so, 44 percent expect to be analyzing unstructured data in the next 18 months.[10]

When it comes to velocity, big data analysis is handled either on a batch or real-time basis. While both types of analytics are now employed equally, that is changing. By 2015 organizations expect to perform nearly 63 percent of their analytics in real time to gain instant intelligence.[11] That means data must be available immediately across the network for analysis, rather than being staged through managed bulk transfers to minimize network traffic and processing loads.

"By 2015, 63% of big data analytics will be performed real-time to gain instant intelligence"

Volume is also a major challenge. Big data storage is scaling from gigabytes (GB) to terabytes (1 TB = 1,000 GB) and petabytes (1 PB = 1,000,000 GB). According to a 2014 IDC survey, the average organization now stores 164 TB of data, with that number expected to skyrocket to 289 TB within 18 months. Nearly a third of the biggest enterprises expect to be working with a 1 PB dataset by the end of 2014.[12]

Data Sources IT Managers Are Asked to Analyze for Insight

Data Source	Total %
Documents	84%
Transactions in DB	82%
Email	74%
Imaging data	68%
Sensor or device data	57%
Search engine indexing	57%
Weblogs	55%
Social media	54%
Phone conversations	52%
Videos	52%
Pictures	46%
Clickstreams	42%

Source: Intel

Price and performance breakthroughs in data center storage and processing are a key factor in successfully meeting big data variety, velocity and volume requirements. For example, Apache Hadoop, an open-source software solution that enables large-scale data storage and processing using commodity server hardware, has become a high-tech superstar. While companies like Facebook have created a buzz with massive Hadoop implementations, organizations are deploying big data with a wide range of approaches, including NoSQL—short for Not Only SQL—is a distributed, non-relational database framework.

Big Data Means Big Traffic

Whichever big data architecture approach is employed, organizations must be ready to handle a corresponding surge of big traffic across the wide area network (WAN) to collect, transport, store, analyze and retrieve it.

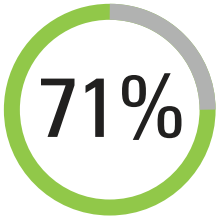
"These workloads are bringing the WAN to the breaking point," concludes research firm IDC. "Network throughput is paramount to the successful implementation of big data in transactional environments."[13]

In a survey of IT managers, Cisco Systems found that 71 percent expected their network loads to at least double over a two-year period. However, only 40 percent reported they are ready for a surge in network traffic.[14] To stay ahead of the curve, one in four companies is already increasing network bandwidth to prepare for the flood of big data.[15]

"59% of organizations store their data in three or more geographical locations"

Not only are companies managing more data, they are storing it in more places. A survey by cloud service provider NaviSite, a Time Warner Cable company, finds that 59 percent of organizations store their data in three or more geographic locations, including different data centers within their own network as well as those hosted by external cloud service providers.[16]

The explosion in distributed big data storage correlates with surging network bandwidth. Researchers estimate that 0.5 Gbps of network bandwidth is needed per petabye (PB) of big data under management by Hadoop. One network engineering consultant observed that, for many customers, Hadoop traffic alone accounted for 20 percent or more of WAN bandwidth utilization between data centers.[17]



of companies are interested in third-party cloud service providers for big data analysis

Source: Intel

IT professionals are challenged not only to address WAN bandwidth capacity but also quality. That's because the bulk of big data traffic is server-to-server. Unlike user-to-server applications like email, web surfing and file sharing, server-to-server traffic is latency sensitive. In other words, distributed big data storage and analytics applications require WAN performance that resembles the bandwidth and latency levels traditionally provisioned within a data center. Highlighting the trend, IT executives identify speed and availability, in addition to security, as their top data-related challenges.[18]

Cloud Power

Organizations are increasingly eyeing opportunities to partner with cloud providers to host big data applications. This way big data storage, analytics processing and traffic can be scaled within a cloud provider's network, rather than an organization's own WAN and data centers. A survey of IT professionals finds that 71 percent are interested in third-party cloud service providers for analysis of their data sets.[19]

Utilizing cloud solutions for big data may yield significant cost savings. Accenture compared the costs and benefits of a Hadoop cluster built and operated within a company for big data processing with a Hadoop-as-a Service offering delivered by a cloud solution provider. The results revealed that the big data cloud solution offered a better price-performance ratio.[20]

Selecting a Big Data Network Solution

Even when an organization leverages a cloud provider to assist with big data storage, processing and analytics, the data must be collected and transmitted across the WAN. Therefore, organizations are responding to the networking demands of big data by deploying highly reliable, scalable and secure WAN connections with quality of service (QoS) capabilities, such as Business Ethernet, and dedicated Internet access solutions.

"Unplanned data center outage expenses have increased to $7900 per minute, a 41% increase in 3 years"

The cost of an unplanned data center outage is now more than $7,900 per minute—a 41-percent increase over three years ago. The dramatic spike in outage costs—including expenses for business disruption, lost revenue and end-user productivity—highlights the mission-critical role of real-time data availability for today's businesses.[21]

Because WAN performance, resiliency and diversity are essential when deploying big data, partnering with the right service provider is crucial. Organizations should consider the following factors when choosing a service provider for WAN and Internet connectivity.

1. Facilities-Based Foundation: When evaluating WAN and Internet and service options, select a provider that owns its last-mile network and can offer future-proof fiber-based solutions. Facilities-based providers with broad network reach are better able to guarantee service quality. Using a variety of access technologies, they may also be able to serve a wide range of locations, including retail sites and branch offices located in residential areas.

2. Service Level Agreements: Because WAN and Internet reliability is essential to big data success, select a provider that offers service level agreements (SLAs). Such agreements set performance benchmarks for service reliability and, should an unplanned outage occur, responsiveness for repair and restoration.

"Ethernet services deliver up to 81% lower costs than TDM alternatives"

3. Ethernet Options: Ethernet is emerging as a preferred WAN solution to cost-effectively deliver the performance required for big data and other essential network applications. Ethernet services seamlessly scale from 5Mbps to 10 Gbps and offer important built-in quality and resiliency features. Furthermore, an analysis by ACG Research found that Ethernet services deliver 44 percent to 81 percent lower costs than legacy TDM-based WAN alternatives. Be sure a prospective WAN provider supports Ethernet services and has earned the latest Metro Ethernet Forum (MEF) certifications such as CE 2.0.

4. Cloud Capabilities: Some facilities-based providers are able to offer a complete bundle of WAN, Internet and cloud services. Purchasing these services together may yield cost savings and simplify management with a single bill to process and one number to call for support. Indeed, a recent survey finds that 73 percent of small and mid-sized businesses prefer to purchase cloud services from their current Internet service provider.[22]

NaviSite, a Time Warner Cable Company, provides essential building blocks for cloud based IT services that support the storage and management of increasing volumes of enterprise data. Services include NaviCloud® Remote Storage, NaviSite Managed Cloud Services, NaviCloud Director™, and many others. In a recent performance analysis, NaviSite's solutions exceeded those of its key competitors in the cloud services category.[23]

Big data presents big opportunities and big challenges. While big data analytics can help businesses drive revenue, operating efficiencies and enterprise value, IT professionals are tasked with enabling the mission-critical infrastructure and applications that make it possible. Partnering with a service provider that can deliver high-performance and cost-effective WAN connectivity and cloud storage is an essential ingredient for big data success.

About the Author

Michael Harris is principal consultant at Phoenix, Arizona-based Kinetic Strategies, Inc. Applying more than 15 years of experience as a strategist, research analyst and journalist, Michael consults with select clients in the networking, Internet and telecommunications industries.

About Time Warner Cable Business Services

Time Warner Cable Business Services, a division of Time Warner Cable, offers a full complement of business communications tools to small, medium and enterprise-sized companies under its Time Warner Cable Business Class brand. Its Internet, voice, television, network and cloud services are enhanced by award-winning customer service and local support teams. Through its NaviSite subsidiary, Time Warner Cable Business Services also offers scalable managed services, including application services, enterprise hosting, and managed cloud services primarily in the U.S. and U.K. Time Warner Cable Business Services, founded in 1998, serves approximately 625,000 business customers throughout Time Warner Cable's service areas. For more information, visit http://business.twc.com.

[1] IDG, Enterprise Big Data Study, 2014.
[2] Ivey Business Journal, "Why Big Data Is the New Competitive Advantage," July/August 2012.
[3] SAS, "Big Data: Harnessing a Game-Changing Asset," Economist Intelligence Unit Limited, 2011.
[4] SAS Institute, "Big Data in Big Companies," Thomas H. Davenport and Jill Dyché, May 2013.
[5] IDG, Enterprise Big Data Study, 2014.
[6] IDC, "Building a Datacenter Infrastructure to Support Your Big Data Plans," January 2014.
[7] Wakefield Research, Global Survey: Is Big Data Producing Big Returns?
[8] IDG, Enterprise Big Data Study, 2014.
[9] IDC, "Extracting Value from Chaos," June 2011.
[10] Intel, "Big Data Analytics: Intel's IT Manager Survey on How Organizations Are Using Big Data," August 2012.
[11] Intel, "Big Data Analytics: Intel's IT Manager Survey on How Organizations Are Using Big Data," August 2012.
[12] IDG, Enterprise Big Data Study, 2014.
[13] IDC, "The Intelligent WAN: IT Transformation, Content, and Big Data Pose Big Traffic Challenges," February 2012.
[14] Cisco Systems, "Connected World Technology Report," 2012.
[15] IDG, Enterprise Big Data Study, 2014.
[16] Navisite Storage Survey, 2013.
[17] "Big Traffic Networking: How Big Data Impacts WAN," Ashish Shah, August 16, 2012 .
[18] Navisite Storage Survey, 2013.
[19] Intel IT Center, Peer Research, Big Data Analytics, 2012.
[20] Accenture, "Hadoop Deployment Comparison Study: Price-performance comparison between a bare-metal Hadoop cluster and Hadoop-as-a-Service," 2013.
[21] Ponemon Institute, "2013 Cost of Data Center Outages," December 2013.
[22] Cisco Systems, "What Do SMBs Want from Commercial-Services Providers?" 2013.
[23] Rick Blaisdell, "Choosing the Right Cloud Provider for Your Business," January 2014.