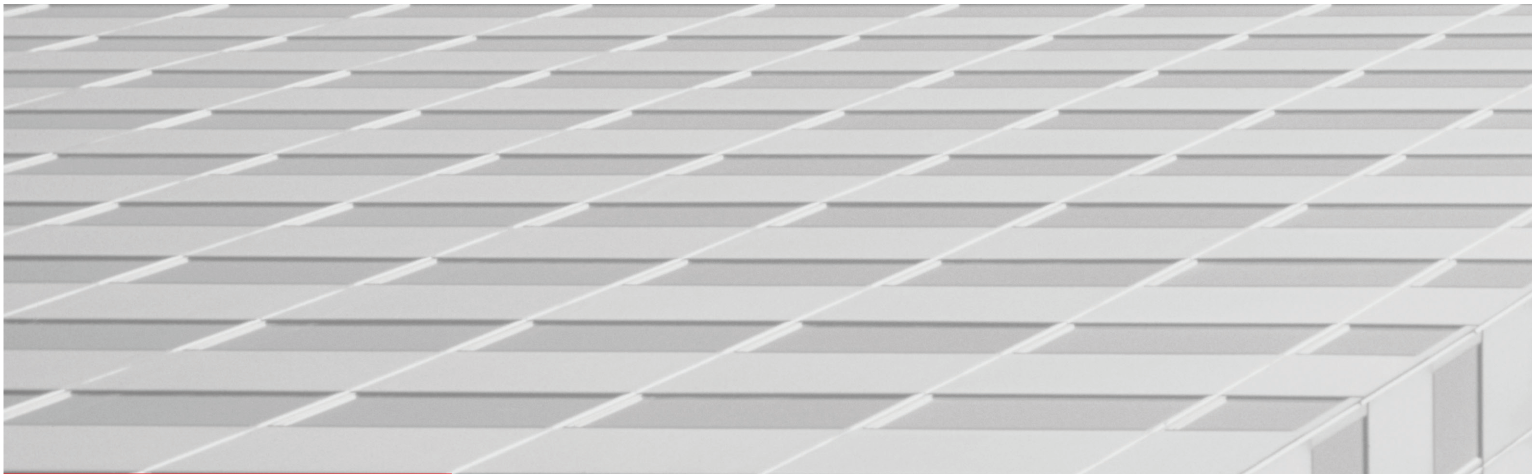

DATA CENTER

Scale-Out Storage, Scale-Out Compute, and the Network

Brocade VDX switches with Brocade VCS Fabric technology provide an automated, efficient, scale-out network architecture optimized for today's scale-out storage and compute environments.

BROCADE

The volume of unstructured data—commonly generated by video, audio, social media feeds, and Virtual Machine (VM) images—is exploding and placing new demands on scale-out storage and scale-out compute architectures. The requirement for petabytes, or even exabytes, of storage to keep the data available and the need to manage costs have led to a rapid evolution of distributed file systems, distributed storage, and NoSQL distributed databases to house the data. Additionally, the need to continuously mine this data on-demand to extract business value has spawned new analytics engines and ecosystems. Combined, these trends are driving new distributed, scale-out infrastructures for storage and compute. Yet all of these new scale-out solutions have a significant impact on the network. Brocade® VDX® switches with Brocade VCS® Fabric technology provide an automated, scale-out architecture with the next-generation network capabilities needed to help ensure optimal performance and efficiency in these new storage and compute environments.

THE IMPACT OF UNSTRUCTURED DATA ON STORAGE, COMPUTE, AND THE NETWORK

The ratio of unstructured to structured data is changing rapidly in today's data center environments, driven by mobile devices, social media, video, industrial telemetry, machine learning, and video surveillance. IDC estimates that in 2017, unstructured data will account for 79.2 percent of capacity shipped, growing at a Compound Annual Growth Rate (CAGR) of 42.5 percent.[1] This plethora of unstructured data, fueled by new data sources, is greatly affecting the infrastructure that supports it, including the storage, the analytics engines that mine the value of the data, and the underlying network infrastructure. To understand why these new data types are having a profound impact on storage, compute, and the network, it is important to understand the difference between structured, unstructured, and semi-structured data sets.

Structured data adheres to predefined schemas conventionally provided by applications, such as relational databases, in which periodic extractions are done for SQL-based business analytics queries. This model scales up well and continues to be the dominant way that transactional applications and tools interact with data.

Unstructured data, on the other hand, has no predefined format and must be stored and retrieved differently. It is stored as files in file systems, or objects in object-based storage systems. Unstructured data types include audio, video, social media feeds, and VM images.

Semistructured data is an emerging data class driven largely by the Internet of Things (IoT), in which a network of machines and sensors is generating data. In this case, data is stored in a NoSQL database rather than a relational database, where the data must adhere to a predefined schema. NoSQL databases are flexible and adapt to the data.

The explosive growth of this unstructured data (see Figure 1), the requirement for petabytes or even exabytes of storage to keep it available, and the need to manage costs has led to a rapid evolution of distributed file systems, distributed storage, and NoSQL distributed databases to house the data. Additionally, the need to continuously mine this data on-demand to extract business value has spawned new analytics engines and ecosystems. Combined, these trends are driving new distributed, scale-out infrastructures for storage and compute. Often overlooked, however, is the impact these new scale-out solutions have on the network. This white paper explores specific examples of scale-out compute and storage and highlights the requirements these solutions are imposing on the underlying network infrastructure.

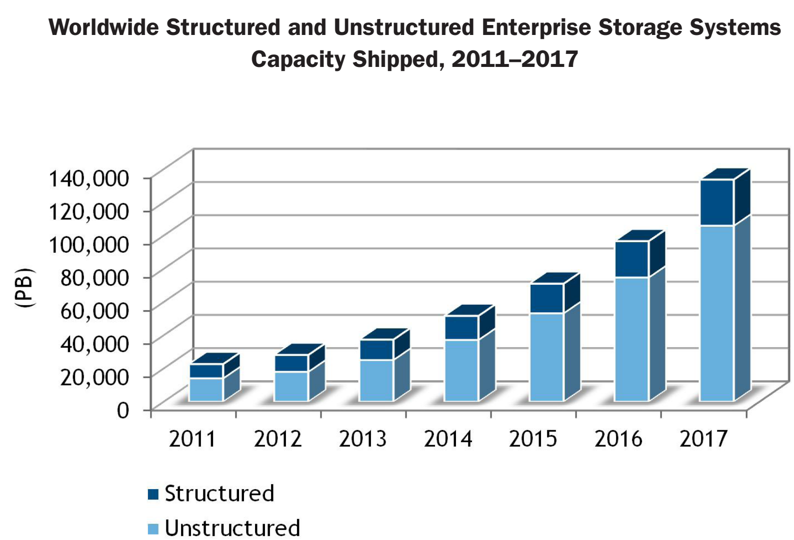

Figure 1.
The explosive growth of unstructured data is fueled by new data sources.

[1] Source: IDC, *Structured Versus Unstructured Data: The Balance of Power Continues to Shift*, March 2014.

BROCADE VDX SWITCHES: DELIVERING OPTIMAL THROUGHPUT

Brocade VDX switches, powered by Brocade VCS Fabric technology, are optimized for scale-out architectures. Brocade VDX switches provide optimal throughput by delivering frame-by-frame multipathing on Inter-Switch Links (ISLs) and flow-based multipathing on all links. This enables significantly higher trunk utilization—up to 30 to 40 percent higher throughput—than competing products, depending on flow diversity. In addition, network capacity can be added nondisruptively, as existing switch configurations do not need to be modified when adding new switches.

CEPH: A DISTRIBUTED OBJECT STORAGE PLATFORM FOR UNSTRUCTURED DATA

Ceph is an open source project that provides a distributed object storage platform. When deploying Ceph, it is considered a best practice to interconnect the compute and storage nodes via a dedicated 10 Gigabit Ethernet (GbE) network, represented as the "Cluster Network" in Figure 2. The network is essentially the Ceph backplane. The optimal network architecture possesses the following attributes:

- The cluster network provides an efficient and high throughput switching infrastructure critical to Ceph, as poor network performance results in much longer replication times. In large clusters, adverse performance can translate to hours or days for replication to complete, as opposed to seconds or minutes.
- The network backplane also provides resiliency and load balancing. When a Ceph node fails, the network functions as the backplane used to reconstitute replicas of the data lost on the failed node and to propagate it to other nodes. A similar situation occurs when a new Ceph node is added and the load is balanced across the cluster. Cross-rack inter-switch throughput is optimized to ensure the health and performance of the entire distributed system. As the system scales out across racks, the inter-switch links do not become a bottleneck.

Many solutions share challenges and requirements similar to Ceph, including distributed databases such as Basho/Riak, MongoDB, and Apache CouchDB.

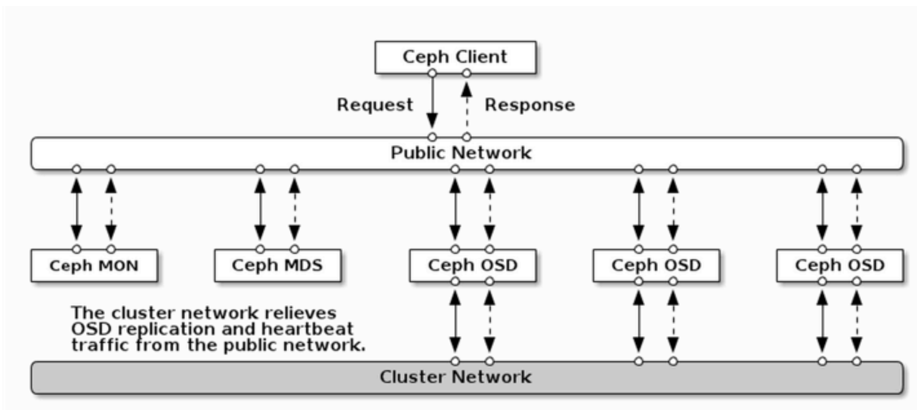

Figure 2.
A Ceph network configuration.

HADOOP: NEW ANALYTICS ENGINES DRIVE SCALE-OUT ARCHITECTURES

The explosion in unstructured data is also fueling an enormous investment in new analytics engines to mine the business value of this Big Data. These analytics engines are driving scale-out compute, storage, and network deployments. In addition, growing demand for real-time or near-real-time analytics is changing programming paradigms and driving the need for faster hardware running at higher Input/Output per Second (IOPS) and for lower latency. This has spawned an ecosystem of developers and vendors with a common goal of addressing these new requirements. A notable example of this is the Hadoop ecosystem.

The current Hadoop 2.0 model includes a cluster resource management layer called YARN, which allows programs other than the batch-oriented MapReduce to run natively in Hadoop (see Figure 3). In-memory analytics can be achieved with Apache Spark, which can run up to 100 times faster than MapReduce. Real-time analytics can be achieved with Apache Storm, which enables a new range of tasks to be performed that could not be done in batch-oriented MapReduce.

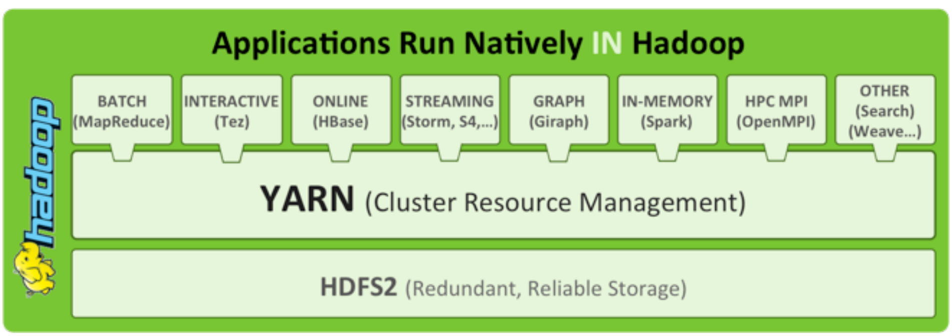

Figure 3.
Hadoop 2.0 (courtesy of Hortonworks).

As a result, analytics pods and object stores are quickly making their way into a wide spectrum of businesses, research organizations, and cloud service providers. They are driving new requirements, beyond throughput, for the network infrastructure that interconnects them. For example, these systems are built to have multiple hosts working independently on tasks, which, when completed, can result in bursts of traffic converging on a single node in the network. This is known as Transmission Control Protocol (TCP) Incast. This many-to-one communication pattern can overrun the buffers of that network node, a condition known as a microburst, causing the individual hosts to back off and slow down, going into TCP slow-start, which dramatically impacts application performance.

BROCADE VDX SWITCHES: OPTIMIZED TO SUPPORT TCP INCAST

Brocade VDX data center switches are highly optimized for analytics environments where TCP Incast is common. These switches deliver industry-leading packet buffering and predictable latency to ensure analytics clusters operate smoothly without interruption. All Brocade VDX switches provide line-rate performance at very low latencies, and performance is highly deterministic with minimal jitter.

NETWORK REQUIREMENTS FOR SCALE-OUT ARCHITECTURES

Organizations are increasingly adopting scale-out architectures in order to address the unprecedented growth of unstructured data. While scale-out systems may be more cost-effective than scale-up systems, there is a trade-off. As additional capacity and performance are needed, new devices must be added to the scale-out architecture—as opposed to simply adding capacity to a system with scale-up. While the equipment may be less expensive than traditional scale-up equipment, there will be much more of it, increasing operational complexity. As a result, organizations will seek solutions that offer greater automation and programmatic control of their data center resources.

Within the network, the ability to add capacity without having to reconfigure existing switches is extremely important. To that end, new switches should automatically configure themselves in order to minimize manual intervention. Additionally, organizations should have the ability to use programmatic tools to manage the infrastructure. The network should expose industry-standard Application Programming Interfaces (APIs) for configuration and operational state retrieval in an open, programmatic way to allow the use of modern DevOps tools. All of these capabilities are provided by Brocade VDX switches.

Brocade Network OS implements the IETF Yang data model, which allows object-oriented programmatic configuration and operational state retrieval. This is accomplished using IETF NETCONF and REST APIs. Puppet and Python scripting is supported to address DevOps environments. OpenStack and CloudStack orchestration is also supported using Brocade plugins. Software Defined Networking (SDN) is supported with OpenFlow 1.3 for the Brocade Vyatta® OpenDaylight (ODL) controller as well as third-party ODL controllers.

Multiple switches in a Brocade VCS fabric can be managed as one logical chassis with configuration accessed via a single management IP address. This significantly reduces operational complexity and expenses. In addition, Brocade Network OS offers a rich set of multitenancy capabilities for deployments in which the infrastructure is shared.

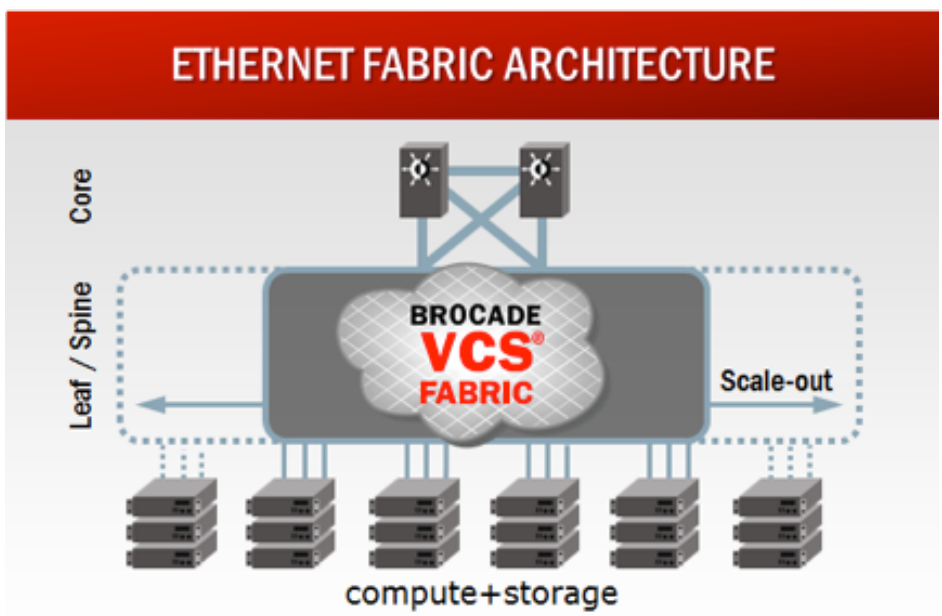

Figure 4.
A Brocade VCS fabric built on Brocade VDX switches.

SUMMARY

Brocade VDX switches provide key network capabilities that help ensure optimal scale-out performance in distributed storage and distributed compute environments:

- **Highly automated and simple to deploy:** Scale-out systems exist in a flat space. Connecting a traditional three-tier network hierarchy to a scale-out system results in a very complex, inefficient deployment model. Brocade VDX switches with VCS Fabric technology offer a flat fabric, automatic trunk formation, automatic healing, and logical chassis operations—significant advantages in a scale-out storage environment.
- **Predictable performance:** Brocade ISL Trunking, Layer 2 (L2) Equal-Cost Multipathing (ECMP), and Layer 3 (L3) Virtual Router Redundancy Protocol Extended (VRRP-E) multipathing combine to offer the industry's best and most predictable fabric utilization. Other solutions have an unequal distribution of load across links, which can be detrimental to deterministic behavior. In addition, Brocade hardware architectures minimize jitter and provide deep buffering. A lossless option provides improved predictability for TCP flows, which is important for Long-Fat Network (LFN) flows, or "elephant" flows.
- **Efficient:** Nonblocking multipathing, Brocade ISL Trunking, ECMP, and VRRP-E all provide optimal use of the fabric links for best-in-class efficiency.
- **Programmability:** Brocade VDX switches provide OpenStack Neutron ML2 support and fabric-level, programmable REST APIs with a YANG data model to enable integration with third-party and in-house network automation and cloud management tools. Support for Puppet and Python scripting offers choice and more effective configuration management. VCS fabrics provide support for OpenFlow 1.3, an industry-standard SDN communications protocol, allowing operators to address complex network behavior, optimize performance, and leverage a richer set of capabilities.

For more information about Brocade solutions, visit www.brocade.com.

ABOUT BROCADE

Brocade networking solutions help organizations achieve their critical business initiatives as they transition to a world where applications and information reside anywhere. Today, Brocade is extending its proven data center expertise across the entire network with open, virtual, and efficient solutions built for consolidation, virtualization, and cloud computing.

Learn more at www.brocade.com.

Corporate Headquarters
San Jose, CA USA
T: +1-408-333-8000
info@brocade.com

European Headquarters
Geneva, Switzerland
T: +41-22-799-56-40
emea-info@brocade.com

Asia Pacific Headquarters
Singapore
T: +65-6538-4700
apac-info@brocade.com

© 2015 Brocade Communications Systems, Inc. All Rights Reserved. 02/15 GA-WP-1850-01

ADX, Brocade, Brocade Assurance, the B-wing symbol, DCX, Fabric OS, HyperEdge, ICX, MLX, MyBrocade, OpenScript, The Effortless Network, VCS, VDX, Vplane, and Vyatta are registered trademarks, and Fabric Vision and vADX are trademarks of Brocade Communications Systems, Inc., in the United States and/or in other countries. Other brands, products, or service names mentioned may be trademarks of others.

Notice: This document is for informational purposes only and does not set forth any warranty, expressed or implied, concerning any equipment, equipment feature, or service offered or to be offered by Brocade. Brocade reserves the right to make changes to this document at any time, without notice, and assumes no responsibility for its use. This informational document describes features that may not be currently available. Contact a Brocade sales office for information on feature and product availability. Export of technical data contained in this document may require an export license from the United States government.

BROCADE