DriveScale

D∠LLEMC

DriveScale-DellEMC Reference Architecture

DellEMC/DRIVESCALE

Introduction

DriveScale has pioneered the concept of Software Composable Infrastructure that is designed to radically change the way data center systems are designed, deployed, managed and consumed. We connect disaggregated compute and storage components in an intelligent manner and give IT operators the ability to manage and modify these connections over time. This results in a simpler deployment model with a fluid pool of resources that can be used for modern cloud-native workloads with significantly improved agility.

With DriveScale you can save up to 40% in up front capital costs and get 3X better utilization of your infrastructure

Smarter Scale-Out for Cloud Native Applications

Traditional Data Centers are transitioning to a Cloud Native world where instead of vertically integrated stacks of hardware and software, applications are designed to scale horizontally on several nodes. Resiliency is built into the software itself so that the hardware does not require it.

However, typical scale-out platforms have multiple limitations:

- Administrators can't respond quickly to changing application stacks and data velocity

- Resources are over-provisioned and under-utilized in order to provide service level guarantees

- Taking advantage of cost-efficient commodity platforms is not easy

- Multiple silos of hardware are created for each application workload

The opportunity cost can be massive when measured against the ability to respond to changing business conditions. This is especially true in the infrastructure design choices available for Hadoop applications. We define Hadoop as the ecosystem of applications and frameworks that rely on the HDFS filesystem.

Developed with breakthrough technology, DriveScale's rack scale architecture makes all of these problems go away. **Now administrators can flexibly deploy and manage independent pools of compute and storage resources optimized for each workload.**

Dell EMC Advantage

Dell PowerEdge servers maximize operational effectiveness and optimize flexibility at any scale. Focused on accelerated performance, enhanced automation and simplified management, the PowerEdge line-up of servers can help you experience worry-free computing through:

- Greater IT efficiency

- Superior IT agility

- Better IT reliability

PowerEdge servers are scalable, flexible, efficient solutions platforms with streamlined and intuitive management that can help you meet changing missions and drive business success.

PowerEdge rack, tower and blade servers are customer-inspired, feature-rich platforms designed to deliver the performance and versatility you need to meet all challenges in almost any setting, from small businesses to enterprise hyper-scale environments.

Dell/DriveScale Reference Architecture Designs

Entry or POC Configuration

The reference architectures shown below describes a typical single-rack configuration for deploying a Big Data infrastructure utilizing Dell EMC PowerEdge servers, Dell EMC Networking Ethernet switches and Dell EMC direct-attached storage in the form of JBOD's. This is a suggested entry configuration that might be deployed for a small workload or for proof-of-concept testing. This design highlights the value proposition of the Dell/DriveScale solution in that it demonstrates how you can achieve significantly higher levels of flexibility and agility in deploying and managing modern workloads such as Hadoop with a highly composable solution. This design also demonstrates that one can start small and add components as needed, to scale the system. You can add servers or JBOD's to the rack depending on where additional capacity is required.

DELL/DRIVESCALE REFERENCE DESIGN
ENTRY CONFIG – Rackmount

Standard 42U Rack:
12 x Dell PowerEdge R430 Servers
12 Servers in Total
2 x DriveScale Adapters with 4 cards each
2 x Dell MD3060e JBODs
2 x Dell S4048-ON Switches
1 x Dell 2848 1Gb Switch for Out of Band Management

Data Plane:
12 x 2-Socket Servers
96 x 10GbE Ethernet Switch Ports
12 x 40GbE Ethernet Switch Ports (configurable as 48 10GbE ports)
16 x 12Gb SAS Ports on DriveScale Adapters
12 x 6Gb SAS Ports on JBODs
160 Gigabits of bandwidth from Servers to DSA's
120 Max Drives

Performance:
SAS Bandwidth per port – 16 Gb/s
Peak available SAS Bandwidth – 192 Gb/s
Average SAS Bandwidth per drive – 1.6 Gb/s
Average Ethernet Bandwidth per drive – 1.33 Gb/s
Average Number of Drives per Server – 10
Average Bandwidth per Server – 13.33 Gb/s

We show two possible entry configurations, one using the Dell PowerEdge 6320 modular servers and the other using Dell PowerEdge R430 rackmount servers. Both designs achieve the same level of composabilty with similar CPU and memory configurations. The servers are joined by two Dell PowerEdge S4048-ON Ethernet Switches to two DriveScale Adapters that are then linked to two Dell MD3060e JBOD's, each with up to 60 drives.

In this design, the compute elements are diskless or disk lite. Drives are hosted in JBOD's and the DriveScale Adapter provides the connectivity from the disks to the top-of-rack switches, and thereby to the servers. Utilizing DriveScale's system management software (DMS), administrators can flexibly allocate any number of drives to each server, on demand. Hadoop clusters can be deployed on the fly with minimal steps. Several racks can be managed as one and clusters can span multiple racks.

Full Rack Configuration

The design shown in this section builds on the first reference design shown above, to show what a full rack might look like. In this architecture, the aim is to achieve the highest density for servers by using the Dell EMC PowerEdge modular servers, and combine them with high bandwidth for networking and storage throughput. Finally, the design takes into consideration the optimal ratio of drives to servers and the bandwidth to each drive, to build out a rack-level configuration that is optimized for the typical big data workloads. Scaling is achieved by simply replicating this rack-level configuration.

DELL/DRIVESCALE REFERENCE DESIGN
STANDARD CONFIG
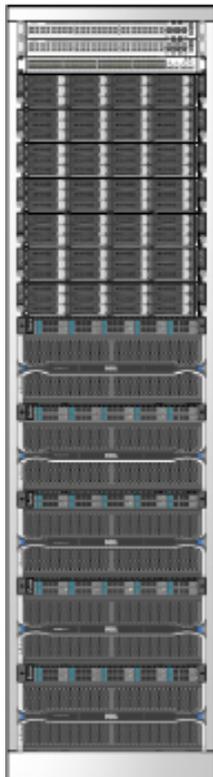
Standard 42U Rack:
7 x Dell PowerEdge C6320 Chassis- Each with 4 Server Blades
28 Servers in Total
5 x DriveScale Adapters with 4 cards each
5 x Dell MD3060e JBODs
2 x Dell S4048-ON Switches
1 x Dell 2848 1Gb Switch for Out of Band Management

Data Plane:
28 x 2-Socket Servers
96 x 10GbE Ethernet Switch Ports
12 x 40GbE Ethernet Switch Ports (configurable as 48 10GbE ports)
40 x 12Gb SAS Ports on DriveScale Adapters
30 x 6Gb SAS Ports on JBODs
400 Gigabits of bandwidth from Servers to DSA's
300 Max Drives

Performance:
SAS Bandwidth per port – 16 Gb/s
Peak available SAS Bandwidth – 480 Gb/s
Average SAS Bandwidth per drive – 1.6 Gb/s
Average Ethernet Bandwidth per drive – 1.33 Gb/s
Average Number of Drives per Server – 10.71
Average Bandwidth per Server – 14.29 Gb/s

As seen in the drawing above, Dell EMC PowerEdge C6320 servers provide a high density bladed design for compute. 28 blade servers are married with 300 drives housed in the Dell EMC MD3060e JBOD's giving the ability to create a Hadoop environment with over 10 disk drives per server on average. There are 5 DriveScale Adapters in the rack, each connected to the Dell EMC S4048-ON top of the rack switches via 10 Gigabit Ethernet ports, and to the JBOD's via SAS ports. This topology provides for an average SAS bandwidth of over 1 Gigabit per second per drive in the enclosures, which is more than sufficient for Hadoop workloads.

Compute Bound Configuration

In some cases, customers require fewer drives per server on average as the workloads are more 'compute bound'. To address such a requirement, we would recommend the following reference design.

DELL/DRIVESCALE REFERENCE DESIGN
COMPUTE-BOUND CONFIG

Standard 42U Rack:
9 x Dell PowerEdge C6320 Chassis- Each with 4 Server Blades
36 Servers in Total
4 x DriveScale Adapters with 4 cards each
4 x Dell MD3060e JBODs
2 x Dell S4048-ON Switches
1 x Dell 2848 1Gb Switch for Out of Band Management

Data Plane:
36 x 2-Socket Servers
96 x 10GbE Ethernet Switch Ports
12 x 40GbE Ethernet Switch Ports (configurable as 48 10GbE ports)
32 x 12Gb SAS Ports on DriveScale Adapters
24 x 6Gb SAS Ports on JBODs
320 Gigabits of bandwidth from Servers to DSA's
240 Max Drives

Performance:
SAS Bandwidth per port – 16 Gb/s
Peak available SAS Bandwidth – 384 Gb/s
SAS Bandwidth per drive – 1.6 Gb/s
Average Ethernet Bandwidth per drive – 1.33 Gb/s
Average Number of Drives per Server – 6.67
Average Bandwidth per Server – 8.89 Gb/s

In the above configuration, additional Dell EMC PowerEdge C6320 chassis and server blades are introduced, with a corresponding reduction in the number of DriveScale Adapters and JBODs, resulting in a slightly lower average number of disk drives per server as compared to the previous

design. Here, each server can be provisioned with 6.67 disks on average. This configuration supports Hadoop workloads that are compute intensive with smaller amounts of data while retaining the same bandwidth to each disk drive as in the previous configuration.

Performance Configuration

Some workloads are I/O intensive and require a greater number of disk drives per server, with fewer overall servers.

DELL/DRIVESCALE REFERENCE DESIGN
PERFORMANCE CONFIG
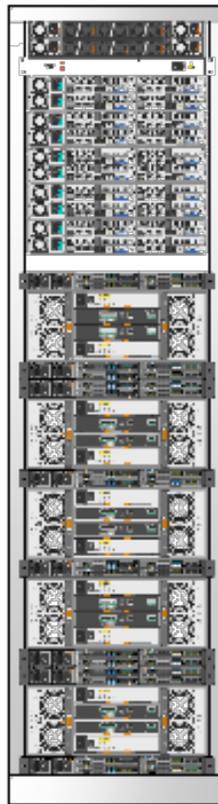
Standard 42U Rack:
5 x Dell PowerEdge C6320 Chassis- Each with 4 Server Blades
20 Servers in Total
8 x DriveScale Adapters with 4 cards each
5 x Dell MD3060e JBODs
2 x Dell S4048-ON Switches
1 x Dell 2848 1Gb Switch for Out of Band Management

Data Plane:
20 x 2 Socket Servers
96 x 10GbE Ethernet Switch Ports
12 x 40GbE Ethernet Switch Ports (configurable as 48 10GbE ports)
64 x 12Gb SAS Ports on DriveScale Adapters
80 x 6Gb SAS Ports on JBODs
640 Gigabits of bandwidth from Servers to DSA's
300 Max Drives

Performance:
SAS Bandwidth per port – 16 Gb/s
Peak available SAS Bandwidth – 480 Gb/s
SAS Bandwidth per drive – 1.60 Gb/s
Average Ethernet Bandwidth per drive 2.13 Gb/s
Average Number of Drives per Server   15
Average Bandwidth per Server – 20 Gb/s

In this reference design, the number of DriveScale Adapters has been increased to support greater connectivity and bandwidth to the JBODs and thus the disk drives. The average number of drives per server is 15 and the average bandwidth to each disk drive is 1.6 Gigabits per second to support workloads that are reading or writing more frequently from and to disk.

Summary

In all the above reference designs, servers, DriveScale Adapters and JBOD's are co-located in the same rack. This ensures that drives are always only one Ethernet switch hop away from the servers that they are attached to, ensuring optimal performance. DriveScale's solution is designed to maintain the data locality that Hadoop requires and provide performance to disk that is equivalent to direct-attached drives.

Therefore, so long as we design the systems with no more latency in the network than that produced from a single Ethernet switch, the systems will perform optimally. There are however, additional elements to consider.

Designing a data center solution utilizing Dell EMC products and DriveScale for optimal performance and cost depends on a good understanding of the type of workloads that need to be supported. These considerations start with an understanding of the average number of disk drives and the capacity of each drive that is needed for each server in the pool. That is followed by deciding the server type and configurations including CPU and memory requirement, as well as form factor. Using DriveScale, customers can deploy diskless or disk-lite server and achieve significantly higher density than conventional 2U servers used in Hadoop deployments. However, given the composable nature of the Dell/DriveScale solution, one can scale as needed on either the compute elements or the storage elements very easily, thereby adding capacity as required. This gives the solution a very flexible architecture, not just in terms of composability but also in terms of capacity scaling.

Conclusion

The Dell EMC / DriveScale reference architecture described in this document will run any standard Hadoop deployment including Cloudera, Hortonworks and Apache Hadoop. Our solution architects are available to help customers design their infrastructure for optimal cost and performance

With DriveScale, you have the ability to operate your datacenter with the agility of a cloud environment while spending less on infrastructure and operations. The architecture integrates quickly and easily into your existing environment with no changes required to the application stack.

DriveScale, Inc
1230 Midas Way, Suite 210
Sunnyvale, CA 94085

Main: +1(408) 849-4651
www.drivescale.com

ra.20171218.003.Rev001