Dell PowerEdge FX2 Network Architecture

Network Flexibility versus Cisco UCS

Tolly Report #216101
Commissioned by
Dell, Inc.

January 2016

Executive Summary

While the introduction of rack server architecture has led to a revolution in the way that data centers are built, not all rack server solutions are the same. Differences in density, footprint, price and network architecture/performance can have a significant impact on the value proposition of a given solution. In particular, the goal of this project was to illustrate the practical and performance benefits of Dell's flexible networking architecture.

THE BOTTOM LINE

Dell's PowerEdge FX2 Architecture delivers:

1. Comparable or better networking performance at less than half the price of the Cisco Systems solution
2. Non-blocking Stacking FX2 modules, providing managed networking to all nodes in an enclosure
3. Intelligent, integrated networking architecture optimized for both "east/west" and "north/south" traffic
4. Smaller footprint that translates to greater density along with fewer cables
5. Completed VM migration tasks up to 28% faster than Cisco UCS with background traffic

Flexible Network Architecture

The key to Dell's flexible network architecture is its FN410S/T and FN2210S series of network interconnect modules. These modules provide eight 10GbE internal facing ports and multiple 10GbE and/or Fibre Channel external ports. The internal facing ports provide built-in connectivity among all of the eight servers that can reside in a single chassis.

Cisco's network architecture is inflexible. Cisco forces network communications between servers residing in the same chassis to leave the chassis, transit the top-of-rack (ToR) switch, re-enter the chassis and, finally, reach the target server. Not only does this waste bandwidth and increase latency (out and back in) but it increases complexity.

This project will explore the characteristics of systems built using gear from each vendor. Test configurations consisted of eight servers from each provider configured as VMware ESXi 6 hosts. Tolly engineers benchmarked virtual machine migration with and without background traffic within (intra) and between (inter) server enclosures.

Network Performance

Dell's network architecture provides for "east/west" traffic between servers to flow directly without needing to go through a ToR switch as Cisco's architecture demands. In VM migration tests within the same chassis this architectural difference resulted in Dell's performance being 19% faster (with 25% fewer cables) than the Cisco solution without background traffic and 28% better in the presence of even modest (10% load) background traffic. Dell's performance through the ToR switch was also 21% better than Cisco in the presence of background traffic. Only in a single ToR scenario with no background load did the Cisco solution run 14% faster than Dell. Both Dell and Cisco solutions may be able to deliver improved results with custom tuning.

Price

Even though the network performance of the Cisco UCS solution could not match Dell in either of the "east/west" traffic test configurations or in tests with background traffic present, the cost was significantly greater. The Cisco solution selling price of $123,000 was more than 2X the cost of Dell at $58,000. Cisco's list price exceeded $215,000.

Footprint & Density

The Dell system footprint was significantly smaller given that Dell implements quarter-size servers vs. half-size for Cisco. Thus, the Dell solution tested required only 4U where the Cisco solution required 14U (2x6U enclosures + 2x1U fabric interconnects). Ultimately, this results in Dell being able to provision 168 servers in a single rack compared to 48 for Cisco. This efficiency extends to cabling as well. Dell required four power cables where Cisco required 12 (4 per enclosure and 2 per fabric interconnect). Finally, Dell required 12 SFP+ cables (4 uplinks x 2 enclosures + 4 stacking cables) where Cisco required 16 cables (4 per enclosure + 4 uplinks.) The extra physical space in the Dell solution can also be used for additional components such as switches and SANs - an important consideration for mid-sized businesses.

Background

A network-level bottleneck can impede server communication and thus degrade the performance of applications that depend upon network bandwidth to perform optimally.

Whether the network connectivity provided in a rack of servers will enhance or degrade server communication is largely dictated by the network architecture of a given rack system implementation. A system can never perform better than its theoretical maximum. Period.

The fundamental network performance of these two platforms is a function of their individual network architectures. As Dell and Cisco Systems have chosen very different approaches to architecting server communications in a server/rack environment, it is important to understand the core elements of each architecture. The next section will introduce key elements of each architecture[1].

Traffic Flows: North/South vs East/West

Central to the understanding and implications of network architecture is the understand ing of data flows. In any server/rack environment there will be traffic that flows between servers and this is referred to as "east/west" traffic while traffic destined for the ToR switch is referred to as "north/south" traffic.

Researchers[2] estimate that over 75% of network traffic is "east/west" and that such traffic is representative of virtualized and cloud environments.

As will be seen below in the detailed discussion, the Dell network architecture is optimal in that "east/west" server traffic can be handled within the enclosure network or chassis stack without traversing a ToR switch. In contrast, the Cisco UCS architecture treats all traffic as "north/south" traffic. That is, even traffic that is between two servers in the same enclosure must transit the ToR switch. It must leave and then re-enter the same enclosure resulting in additional bandwidth usage as well as any additional latency added by the "back and forth" transit of the ToR switch.

---

[1] Dell refers to each individual server as a "node" where Cisco Systems refers to each as a "blade." For the purposes of this discussion, they are synonymous.

[2] http://bnrg.cs.berkeley.edu/~randy/Courses/CS294.S13/3.4.pdf

Cisco UCS Network Architecture

In the Cisco UCS network architecture, all server-to-server communications flow through a ToR LAN switch. Whether servers reside in different racks, different chassis in the same rack or are in adjacent slots in the same chassis, all network traffic must flow through the ToR fabric interconnect. Period.

Cisco's approach is sometimes referred to as a "north/south" network architecture as all traffic flows up and down to and from the ToR switch. Traffic flows from a given chassis to the ToR switch via a Cisco UCS fabric extender (FEX).

The highest-bandwidth FEX available for the Cisco UCS 5108 Chassis is the 2208XP which provides for 80 Gbps of bandwidth between a given chassis and a ToR switch.

Two of these FEX modules are supported per chassis, resulting in the maximum 160Gbps[3] bandwidth between any chassis and the proprietary fabric interconnects[4]. See Figure 1.

The ToR switch communicates with up to two FEX modules in each chassis in a rack. Any and all traffic between servers must flow through a ToR switch. See Figure 2.

Cisco UCS Network Architecture
Current Implementation Chassis-to-Switch Limitation: 160Gbps

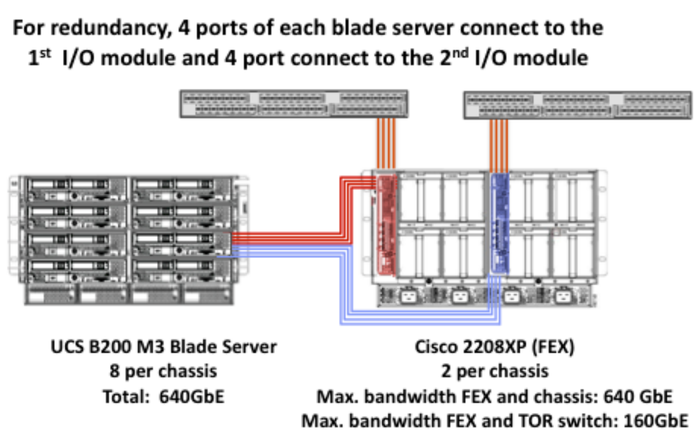


Figure 1

Cisco UCS Network Architecture
All Traffic Flows via TOR Switch

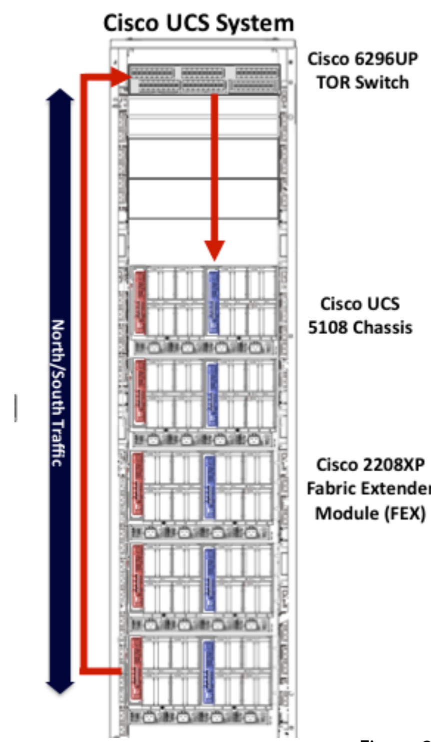


Figure 2

[3] Testing described herein was conducted with 40Gbps of chassis connectivity for each vendor. Test engineers experimented with adding bandwidth for the Cisco solutions but results did not vary.

[4] Note: simply increasing bandwidth in a UCS environment will increase the overall cost of the solution due to port licensing requirements on the UCS 6248 Fabric Interconnects.

Dell, Inc.

PowerEdge FX2 System

Network Architecture Evaluation

Tolly Certified

October 2015

Dell PowerEdge FX2 Network Architecture

TheDell FX2 System architecture offers several options for network connectivity. The most basic solution provides a pair of eight-port 10Gbps Ethernet Pass-through modules that can be connected directly to the ToR switch. For this testing, the more advanced "I/O Module" (IOM) technology was utilized. Available in three different variants, IOMs are installed within the FX2 chassis in pairs and act as integrated switches that can be interconnected to provide seamless east/west AND north/south traffic flows. Variants include: the FN410S which provides 4 SFP+ ports, the FN410T that provides 4 10Gb Base-T ports and the FN2210S that provides 2 2/4/8G Fibre Channel ports and 2 SFP+ ports (all ports can be configures for SFP+ if desired).

The Dell architecture provides for 40Gbps up to the ToR switch as well as 40Gbps as a dedicated stacking link for "east/west" traffic within an enclosure. See Figure 3. A customer wishing to maximize networking bandwidth further can supplement the IOMs by adding additional 10GbE ports using the PCIe slot in each server node and connecting these directly to the switch providing for a theoretical maximum throughput of 240Gbs – compared to 160Gbs for the Cisco System UCS.

With the Dell System, data can travel between stacked IOMs without traversing the ToR switch. IOMs can be stacked across up to 6 enclosures supporting a maximum "performance domain" of up to 48 server nodes.

Dell PowerEdge FX2 Network Architecture - Stacking Dataflows
Full, Intra-Chassis, Non-Blocking Networking

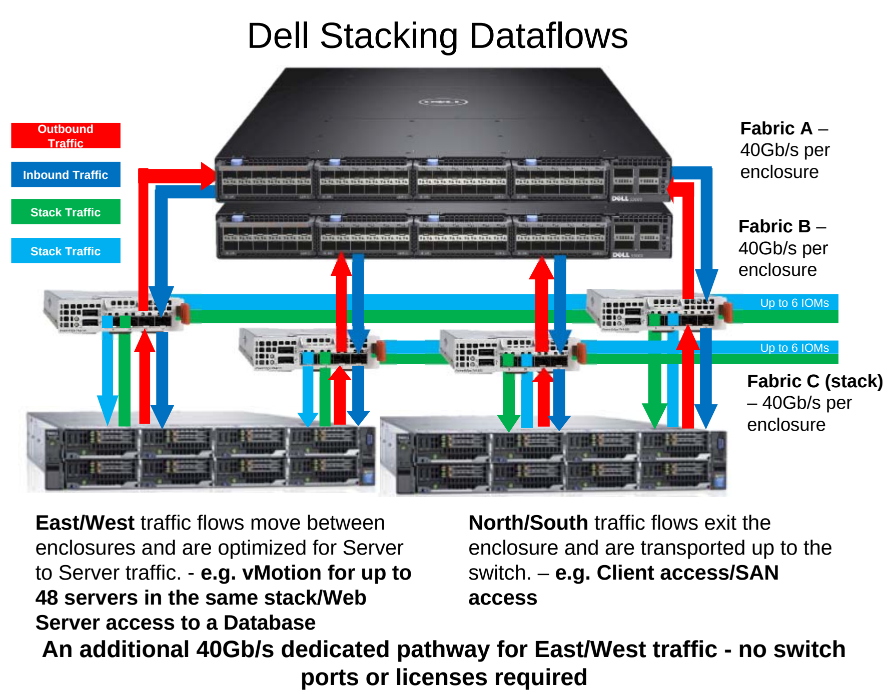


Figure 3

Dell PowerEdge FX2 System Network Architecture: Demonstrable Benefits

Architectural benefits need to be more than just academic. To benefit users, architectural attributes must translate into demonstrable benefits relevant to real-world implementations. Given the importance of virtualization and the prevalence of its use in rack server deployments, Tolly engineers benchmarked the impact of network performance on a key element of virtualized server environments - VM real-time migration.

Challenge

To optimize and load balance resources of virtualization hosts, virtualization providers like VMware migrate virtual servers, as they are running, between physical host servers. VMware refers to this live migration feature as vMotion[5]. Since the server data must be moved to the target server across the network, engineers built a series of tests that would illustrate the impact of unrelated network activity on the time required for virtual server migration.

Test 1: Live Migration of ESXi Server Node

In the first of two tests, four sets of eight related virtual servers (called a " vApp tile" by VMware) were queued for live migration from one physical host to a second empty host on another similarly-configured chassis. Thirty-two virtual servers were migrated as a whole, in an emulation of a full-node evacuation. A deliberate two-minute waiting period from the time one virtual server completed migration to when the next began in order to isolate the migration time. Two simple network load scenarios were used. The test was first run without any other network traffic. This test was repeated with an additional background load of approximately: 15Gbps on Cisco UCS, and 27Gbps on the Dell PowerEdge FX2. The background traffic was generated between one 'standby' VM, positioned in each vApp tile (on-host), and a central off-host client designated for each vApp tile in the same chassis. The traffic generated was 'north-south', terminating at a physical client.

While network bandwidth available to each host was similar for each environment, the differences in network topology and configuration enable the Dell solution to perform this migration over the I/O Module's 'stack' links instead of directing traffic through the ToR switch, unlike in the case of the Cisco UCS. (The Cisco solution passes all traffic through the dual 6248UP Fabric Interconnects in a 'north-south' traffic pattern.)

Because the Dell PowerEdge FX2 System supports up to 8 nodes per chassis in a 2RU form factor, and Cisco UCS provides 8 blades per chassis , the overall rack density is very different for each of these solutions. Due to network architecture differences overall

---

[5] https://www.vmware.com/products/vsphere/features/vmotion.html

product bandwidth differences, the background network load was significantly different as noted above.

See Test Methodology section for additional details.

Test 1: Outcome

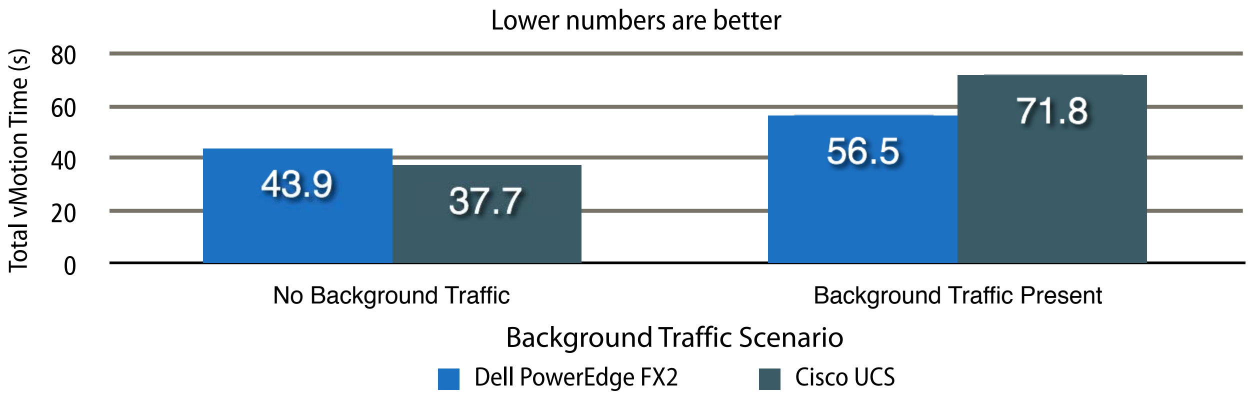
When migrations were performed without background traffic, the Cisco system was 14% faster. With background traffic, however, the Dell system was greater than 20% faster. See Figure 4.

With no background load on the network, the Dell FX2 System completed the vMotion virtual server migration in 43.9 seconds total time compared with 33.7 seconds for the Cisco UCS.

As background traffic was introduced, the Dell FX2 migration time increased modestly to 56.5 seconds, while the Cisco UCS vMotion run time increased over 90% to 71.87 seconds. While the background traffic was well below the maximum levels that the Cisco UCS will support, the presence of even a relatively low level of unrelated traffic clearly had a negative impact on the performance of the VMware live migration.

While the Cisco UCS5108 chassis has a theoretical maximum of 160Gbps between a single chassis and the ToR switch, in our test configuration the maximum theoretical throughout was 40Gbps. The maximum background traffic load generated by Netperf in this test was approximately 15Gbps.

Dell PowerEdge FX2 System vs. Cisco Unified Computing System (UCS)
Node to Node (Different Enclosure)
4x vApp Tile (32 VMs) Migration Time Without and With Background Network Load
(as reported by VMware vCenter 6)



Notes: Four tiles of 8VMs each migrated simultaneously. Background traffic generated with NetPerf tool. Background traffic varied by solution depending upon what the solution could carry. For Cisco, engineers measured ~15Gbps of unidirectional background trafficFor Dell PowerEdge FX2 max background traffic measured was ~27Gbps. 4 blades in Cisco UCS 5108 Chassis, 4 nodes in PowerEdge FX2Chassis.

Source: Tolly, October 2015

Figure 4

Test 2: Live Migration of Full Server Enclosure

The second test conducted was a slight variation of the first test. While the first test benchmarked the migration of four vApp tiles between ESXi nodes, the second test benchmarked sixteen 8-virtual server vApp tiles (128 virtual machines in total) going through VMware live migration simultaneously. All other conditions and background loading scenarios remained the same between these two benchmark tests.

Test 2: Outcome

In all test cases, the Dell PowerEdge FX2 system completed the migration task faster than the Cisco UCS. See Figure 5.

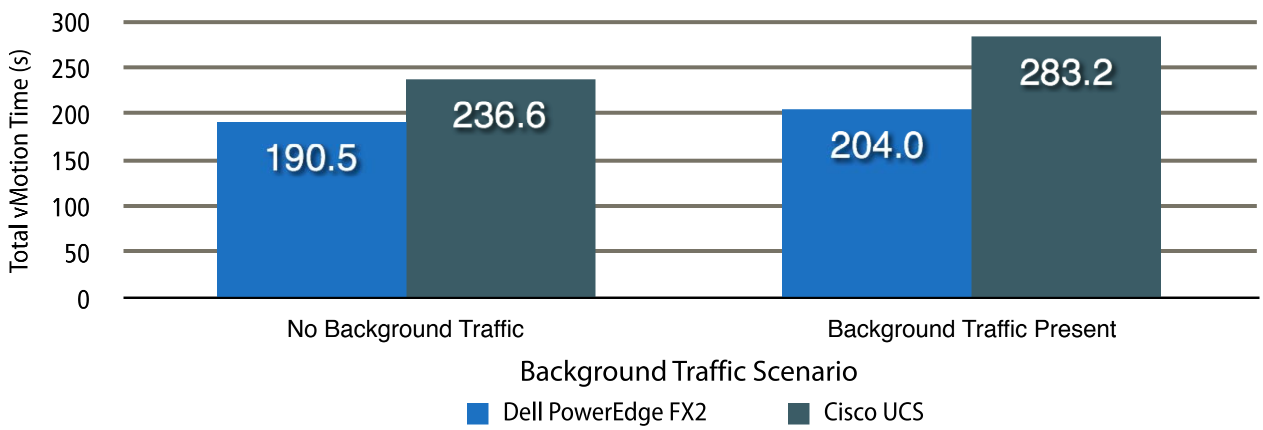
When migrations were performed without injected background traffic, Cisco performed the migration of all 128 VMs (16-tiles) in 3:56.6, whereas Dell completed the same task in 3:10.5, a time savings of nearly 25%.

When traffic was introduced, at the same levels as Test 1, the Dell migration time increased to 3:24, while the Cisco UCS time increased approximately 20% to 4:43.

Compared to Cisco, the Dell PowerEdge FX2 system required 38% less time to complete the migrations, under an increased network load. The fundamentals of this test are similar to Test 1, and the migration traffic takes the same path through each network environment.

Dell PowerEdge FX2 System vs. Cisco Unified Computing System (UCS)
Intra-Chassis
VM Migration Time for 16 Tiles (128 VMs) Simultaneously Without and With Background Network Load (as reported by VMware vCenter 6)

Lower numbers are better



Notes: Four tiles of 8VMs each migrated simultaneously. vApp consisted of 8 individual VMs, 16 of which were migrated with two minutes between each vMotion event. Average of three to five runs. Background traffic generated with NetPerf. Under full load, engineers measured 15Gbps of bidirectional traffic, Dell PowerEdge FX2 max. throughput measured at 27Gbps. 4 blades in each Cisco UCS 5108 Chassis, 4 blades in Dell PowerEdge FX2 Chassis.

Source: Tolly, October 2015

Figure 5

While Cisco may be able to improve their results through tuning (no tuning was done to either Cisco or Dell), doing so could incur significant penalties. Tuning increases complexity and customers often indicate to Dell that they do not have the time to implement tuning. Since a Cisco UCS system can support up to 160Gbps of throughput per enclosure, adding bandwidth could improve the results however, this would significantly reduce UCS scaling, it would significantly increase cost (cables + port licenses) and it would add additional complexity (cabling) to the solution.

System Pricing

As noted above, the costs differ dramatically for each solution with Cisco's selling price being more than 2X that of Dell. The US web price for the Cisco UCS solution as tested exceeds $123,000 at list price where the Dell FX2 solution costs less than $60,000. Furthermore, Dell's approach minimizes licensing costs. For example, iDRAC remote management licensing is for the full lifecycle, there is never any renewal required even should the server be sold later on in its lifecycle. The new owner gains the benefit of the existing license thus improving the after market value of Dell. Cisco's policy on license expiration/renewal is unclear. See Tables 1 - 3.

Solution Pricing: Cisco Systems (Web Price)

Part Number	Description	Quantity	Unit Price	Extended Price
UCS-SPL-5108-AC2	UCS 5108 Blade Server AC Chassis with no Power Supply Unit, 8 fans, 1-2208	2	$6,805	$13,610
UCS-IOM-2208XP	UCS 2208XP Fabric Extender/8 external, 32 internal 10Gb ports	2	$4,227	$8,454
UCS-FI-6248UP	UCS 6248UP 48-port Fabric Interconnect (1U) with no Power Supply Unit, 2 fans, 32 x 10GbE/FCoE/FC SFP+ Unified ports, 12 port license preinstalled	2	$14,953	$29,906
UCSB-PSU-2500ACPL	2500W AC power supply unit for UCS 5108 (80 Plus Platinum Certified)	8	$454	$3,632
UCS-PSU-6248UP-AC	600W power supply unit for UCS 6248UP/100-240VAC	4	$658	$2,632
UCSB-B200-M4	UCS B200 M4 half-width Blade Server with no processor, no memory, open bay (diskless), no cards	8	$1,400	$11,200
UCS-CPU-E52630D	2.40GHz Xeon E5-2630 v3 processor (1 chip, 8 cores) with 20MB L3 cache	16	$999	$15,984
UCS-MR-1X162RU-A	16GB DDR4-2133 PC3-17000 2Rx4 1.2V dual rank low voltage RDIMMs	64	$350	$22,400
UCSB-MLOM-40G-03	VIC 1340 modular LOM adapter for blade servers	8	$642	$5,136
UCSB-MRAID12G	Cisco FlexStorage 12G SAS RAID controller with drive bays	8	$918	$7,344
A03-D300GA2	300GB 6Gb SAS 10K rpm 2.5" SFF hot plug disk drive with drive sled	8	$277	$2,216
UCSB-HS-EP-M4-F	Cisco UCS B200 M4 Server - Heat Sink (Front)	8	$49	$392
UCSB-HS-EP-M4-R	Cisco UCS B200 M4 Server - Heat Sink (Rear)	8	$49	$392
			System Total:	$123,298.00

Note: Cisco prices cited are selling prices on CompSource, their lowest prices available for new equipment.

Source: CompSource, December 2015

Table 1

Solution Pricing: Dell, Inc. (Web Price)

Part Number	Description	Quantity	Unit Price	Extended Price
321-BBPI	PowerEdge FX2 Enclosure for up to 8 Quarter-Width Server Nodes, no nodes, no power supplies	2	$6,339.00	$12,678
321-BBGH	PowerEdge FX2S Chassis Configuration with 8 x PCI Express 3.0 Slots	2	Included in price above	
FX2-U3OS-100	PowerEdge FX2 - Minimum Warranty, 10x5 NBD Onsite, 3 Year	2		
540-BBGF	PowerEdge FN410S 10GbE SFP+ I/O Module (4 x 10GbE SFP+ ports)	4		
634-0286	** NO COST ** CMC Express	2		
634-0287	CMC Enterprise	2		
450-ADHQ	2 x 1600W Power Supplies, Hot-swap	4		
770-BBER	ReadyRails Sliding Rails	2		
210-ADYI	PowerEdge FC430 Server Node (no TPM), no processors, no memory, diskless	8	$5,713.46	$45,707.68
FC430-U3OS-100	PowerEdge FC430 - Minimum Warranty, 10x5 NBD Onsite, 3 Year	8	Included in price above	
329-BCLP	** NO COST ** PowerEdge FC430 Motherboard with On-Board 1GbE Network Adapter	8		
329-BCLQ	PowerEdge FC430 Motherboard with On-Board 10GbE Network Adapter	8		
338-BFFU	2.4GHz Xeon E5-2630 v3 processor (1 chip, 8 cores) with 20MB L3 cache	8		
374-BBHD	Additional 2.4GHz Xeon E5-2630 v3 processor (1 chip, 8 cores) with 20MB L3 cache	8		
370-ABUG	16GB Memory - 1 x 16GB dual ranked 2133MHz registered DIMM	64		
330-BBEV	** NO COST ** PCIe Riser for FX2S Configuration	8		
385-BBHN	** NO COST ** iDRAC8 Express	8		
385-BBHO	iDRAC8 Enterprise	8		
406-BBHN	** NO COST ** 1.8" Storage Backplane with up to 2 x Hard Drives	8		
406-BBHM	1.8" Storage Backplane with up to 1 x Hard Drive and 1 x InfiniBand Port	8		
400-AEIF	200GB Solid State Drive uSATA Mix Use Slim MLC 6Gbps 1.8" Hot-swap Drive	8		
			System Total:	$58,385.68

Note: Dell pricing as provided by Dell website.

Source:Dell.com, December 2015

Table 2

Solutions Under Test

	Dell PowerEdge System		Cisco Unified Computing System	
	Model	Notes	Model	Notes
Chassis	Dell PowerEdge FX2s Enclosure	2RU | 8 Slot	Cisco UCS 5108 Blade Server Chassis	6RU | 8 slot
Chassis Networking	PowerEdge FN410S I/O Module	2 per chassis	Cisco UCS 2208XP Fabric Extender	2 per chassis
Top of Rack Switch	Dell S5000 Switch	48 x 10GbE 4x 40GbE	2x Nexus 6248UP	5.2(3) N2(2.23g)
vMotion Blades/Nodes	Dell PowerEdge FC430 Nodes	4 per chassis	Cisco UCS B200 M3	4 per chassis
Blade/Node CPU	2x Intel Xeon E5-2630v3	8 core, 2.40GHz	2x Intel Xeon E5-2630v3	8 core, 2.40GHz
Blade/Node RAM	128 GB	1600MHz	128 GB	1600MHz
Blade/Node Networking	FN410S 2x 40GbE NDC		Cisco UCS VIC 1340 (mLOM)	2.2(3c)

Note: Both systems accessed a share storage environment consisting of:2x Dell Compellent SC8000 controllers, 4x Compellent SC220, each 24x 300GB 15k SAS, 2x2x 10Gb connections to network. VMware systems were ESXi 6.0.0. The two Cisco Nexus 6248UP were connected to a Dell S5000 TOR switch.

Source: Tolly, October 2015 Table 3

Test Setup And Methodology

Physical Environment

Tolly configured one Cisco UCS environment, consisting of 2x Cisco UCS 5108 Chassis, each equipped with two Cisco UCS 2208XP Fabric Extenders in an Active/Active configuration. Each chassis was outfitted with 4x Cisco UCS B200 M4 Blades, each of which contained 2x Intel Xeon E5-2630v3 series processors, providing 16 hardware cores per blade. Each blade was configured with 128GB RAM, and ESXi 6.0.0 installed on one local 300GB 10K hard drive (used only for booting the operating system).

In the configuration tested, each Cisco UCS 6248UP Fabric Interconnect was connected to each blade chassis with 2x 10GbE links and each Fabric Interconnect was independently connected via 4x 10GbE to a Dell S5000 switch, which provided connectivity to the shared storage environment. In this configuration, the UCS system would have a maximum scalability of 128 blades.

The Dell FX2 System consisted of a Dell FX2s Enclosure, equipped with two Dell PowerEdge FN410S I/O Modules and a total of 8x Dell PowerEdge FC430 FX2 based , quarter-width servers. As with the Cisco environment, each node was equipped with 2x Intel Xeon E5-2630v3 series processors, with 128GB RAM, installed with ESXi 6.0.0. Each node was equipped with two 10GbE ports and one 200GB SSD hard drive (used only for booting the operating system).

Two links from each of the FN410S connected to the Dell S5000 Top-of-rack switch, while the two remaining links provided stack networking between the enclosures.

Storage Networking

The shared Dell S5000 switch was connected to two Dell Compellent SC8000 controllers, each attached with 2x 10GbE links. The controllers served data from 4x Compellent SC220 disk enclosures, each equipped with 24x 300GB 15k SAS drives. All test VMs were hosted in a RAID 10 volume on this disk array, served via iSCSI.

Blade/Node Networking

The Cisco UCS B200 M3 Blades were equipped with the Cisco UCS VIC-1340, providing up to 80Gbps of fabric throughput per server. Each Virtual Interface Card exposed two vNICs to the system, which map to 4x 10GbE traces on the Chassis backplane, such that one vNIC from each card was pinned to Fabric A, and the other to Fabric B. Between the two NICs, a total of 8 interfaces were configured, each on a separate subnet .

An MTU of 9000 was set across both environments and vMotion network QoS was set to silver.

The Dell FC430 nodes were provisioned using NPAR which logically split each of the embedded 2x 10GbE Network Daughter Cards (NDC) into 4x10Gb/s virtual NIC's and networks were defined from the FX2 I/O Modules. A total of eight vNICs were exposed to the system, which were each configured on a separate subnet.

On both the Cisco and Dell solutions, a total of two 10GbE vNICs were used for the vMotion network.

IO Module Technology

FN-IOMs, behave as traditional network switches that are housed in the rear of the FX2 chassis. FN IOMs offer network services for both the Ethernet and storage fabrics in a data center. The FN IO Modules are capable of running in a variety of modes depending on environmental requirements. Standalone mode is the default mode in which the FN IO Modules ship. This no-touch mode allows simple and seamless network integration and allows the FN IO Modules to function out of the box. This mode moves the demarcation between server and network zones back to the top of the rack and away from the servers. This allows the FX2 to have the benefits of blade density while behaving more like a traditional rack server.[6] For our tests, each IOM was set up as part of a stack which requires minimal switch configuration.

Networking

The focus of this testing was to determine the impact of streamlining "East/West" (server to server) traffic flows since this is the type of data flow utilized by vMotion. In many cases, overall system performance can be improved in a design by segmenting "East/West" traffic from "North/South" (Server to Core/client network) traffic since doing so allows more of the available bandwidth from the ToR switches to be dedicated to core/client traffic flows. It is also important to note that in addition to vMotion traffic, many solutions benefit from "East/West" traffic flows such as: database to web server and application environments like SAP where multiple servers require access to data that resides on separate servers.

In both, Dell and Cisco test cases, a single 10GbE based Dell S5000 ToR switch was used to provide the unified fabric for data and storage. In addition to traditional Ethernet, this switch also provides connections for both iSCSI and FCoE storage fabrics without the need for additional switches.

Using the Dell IOMs in "stacked mode" in our testing required a total of 12 network cables. Four uplinks per chassis (8 total to the ToR switch) were configured into 2 port-channels and 4 additional SFP+ cables were connected between the stacked pairs. This configuration provided a total of 40GbE of "North/South) bandwidth for the 2 enclosures plus an additional 40GbE of dedicated "East/West" bandwidth between the stacks for a grand total of 120GbE of throughput. Quality of Service was implemented at the Server level and provided dynamic bandwidth allocation for: "management," "data," "storage," and "vmotion".

The Cisco solution required 4 uplinks per UCS 6248UP (8 total) configured into two port channels for "North/South" traffic and an additional 4 uplinks per enclosure (8 total) for a grand total of 16 network cables. The UCS architecture does not support "East/West" traffic flows and requires all server to server transactions to move through the "North/South" pathways. As such, the maximum bandwidth available with 8 uplinks is 40GbE. Quality of service was also implemented through the UCS interface to provide dynamic bandwidth allocation for: "management," "data," "storage," and "vmotion".

Power and Space Requirements

In our testing, each Dell FX2 Chassis could support up to 8 nodes and required 2U of rack space (4U total for 2 systems). Each Cisco chassis required 6U of rack space (12U total) plus 2x Cisco UCS 6248UP Fabric Interconnects at 1U each for a total space requirement of 14U.

---

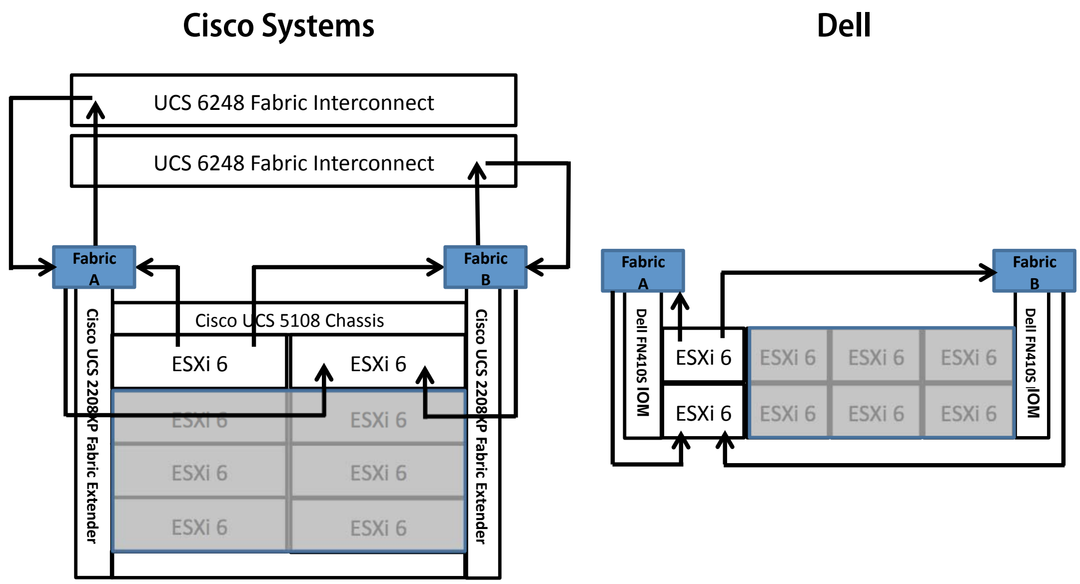
[6] http://en.community.dell.com/techcenter/networking/m/networking_files/20440793/download

Each Dell FX2 required 2 power cables (4 total) while the Cisco solution required 4 cables per chassis (8 total) plus 2 additional power cables per UCS 6248UP (4 total) for a total of 12 power cables required.

vMotion Traffic Path

The optimal traffic path available was used for each solution under test. For the Cisco UCS, this meant that all traffic between servers in the same chassis traversed the TOR switch. For the Dell FX2 System, all network traffic remained in the chassis under test when performing node-to-node migrations, and load-balanced over the cluster network when migrating workloads between chassis. See Figure 6.

VMware vMotion Intra-Chassis Migration Traffic Path



Notes: Cisco UCS migration path caused traffic to exit and re-enter chassis, while Dell system traffic remained within the chassis network.

Source: Tolly, October 2015

Figure 6

Test Configuration

A "Tile" of 8 VMs were used to represent the HA application. This consisted of four web servers, two database servers, an SMTP server, and an idle VM. A ninth VM acts as a client for all workloads, and was placed on a separate physical host for both environments. This server was equipped with two Intel E5-2699 v3 (18 core) CPUs and 128GB of RAM.

There were a total of sixteen "Tiles" spread across all ESXi hosts for a given solution, which would run an independent copy of the workload.

Netperf v2.6.0 was used to generate background traffic load and was controlled by a separate server. A script was configured which would initiate a bi-directional netperf test between the 'idle' server of each tile and the client VM, generating north-south traffic flows.

Test Runtime

Each Dell enclosure was configured with 4 server nodes (8 nodes total) and each Cisco enclosure was configured with 4 server nodes (8 nodes total). 4 tiles (8 VM's per tile) were then created for each server and 4 of the nodes were activated for a total of 128 virtual machines. CPU utilization information from VMware indicated that each system was working at over 75% CPU utilization in this configuration. Testing consisted of starting the VMmark benchmark test and once it was running, using vMotion to move workloads between systems while measuring the time to completion.

A script was configured from a separate host, which would connect to a VMware vCenter 6 instance and perform the migrations on the set of VMs. VMs were migrated either as a set of four vApps (32 VMs) or as a set of sixteen vApps (128 VMs) with a minimum of two minutes wait time between each migration. The vApp workload was also configured to run for a period of three hours. The script used for 16vApps was specifically designed to initiate a parallel vmotion for one server and when finished, would instantly start the process on the next server until all 4 servers were vacated. Further improvements in vmotion performance may be possible by moving all 16vApps simultaneously but for consistency, the same script was used for both Dell and Cisco tests.

For the test runtime, Engineers first began the Netperf traffic generator. Next, ESXtop was invoked on each of the ESXi hosts, and a separate command created a dump of the vkernel logs in order to get the vMotion time from the host perspective.

The vApp workload was started, and after initialization, the vMotion script was executed to begin the events.

After each test was completed, each log was collected and analyzed. A script assisted engineers to access and strip relevant vMotion events from the vCenter database.

Testing focused on comparing the migration time due to differences in system architecture, therefore application performance statistics (i.e., VMmark results) were not captured.

VMmark

This test was based on the VMmark 2.5 benchmark suite. VMmark was designed to fill the need for a multi-server virtualization platform benchmark by incorporating a variety of common platform-level workloads. In addition to being widely recognized as the defacto standard for virtualization testing in a VMware environment, VMmark also provided some important elements to this test including:

1) Utilizes typical customer workloads like email (Microsoft Exchange), Web services, database services and video streaming services.

2) VMmark provides "real-world" data flows with "clients" located on external servers that then drive typical "north/south" traffic flows while at the same time, utilizing common backend architectures that require "east/west" traffic flows like Database to Web Server.

3) VMmark provides the configuration flexibility to allow "infrastructure tests" like vMotion, Storage Motion and deployment activities to be turned set to manual so that they could be activated on demand by the test team and performance accurately measured.

Ensuring Test Consistency

The initial measurements were taken on the Cisco systems and once completed, all virtual machines were migrated to a single enclosure. The empty enclosure was then turned off and the CPU's and memory were moved to the Dell systems. Next, one of the port channels (40GbE) was removed from the Cisco UCS 6248UP connected to a Dell FX2 system.

After powering on the Dell systems, vMotion was used to move all of the workloads to the Dell environment. Finally, the second Cisco enclosure was turned off, the memory and CPU's were moved to the Dell FX2 nodes in the second enclosure, the port-channel was disconnected from the second Cisco UCS 6248UP and reconnected to the Dell FX2 system.

About Tolly…

The Tolly Group companies have been delivering world-class IT services for over 25 years. Tolly is a leading global provider of third-party validation services for vendors of IT products, components and services.

You can reach the company by email at sales@tolly.com, or by telephone at +1 561.391.5610.

Visit Tolly on the Internet at:
http://www.tolly.com

Interaction with Competitors

In accordance with Tolly's Fair Testing Charter, Tolly personnel invited representatives from Cisco Systems, Inc. to participate in the testing, providing the test plan and configuration details. Cisco representatives did not respond.

For more information on the Tolly Fair Testing Charter, visit:
http://www.tolly.com/FTC.aspx

216101--de-l-wt-2016-01-12-Ver0