vStart 1000v for Enterprise Virtualization using VMware vSphere: Reference Architecture

Release 1.0 for Dell PowerEdge Blade Servers, Force10 Switches, and Compellent Storage Center

Dell Virtualization Solutions Engineering

Revision: A01

September 2012
This document is for informational purposes only and may contain typographical errors and technical inaccuracies. The content is provided as is, without express or implied warranties of any kind.

© 2012 Dell Inc. All rights reserved. Dell and its affiliates cannot be responsible for errors or omissions in typography or photography. Dell, the Dell logo, OpenManage, Compellent, Force10, Kace, EqualLogic, PowerVault, PowerConnect, and PowerEdge are trademarks of Dell Inc. Intel and Xeon are registered trademarks of Intel Corporation in the U.S. and other countries. Microsoft, Windows, Hyper-V, and Windows Server are either trademarks or registered trademarks of Microsoft Corporation in the United States and/or other countries. VMware, vSphere, ESXi, vMotion, vCloud, and vCenter are registered trademarks or trademarks of VMware, Inc. in the United States and/or other jurisdictions. Linux is the registered trademark of Linus Torvalds in the U.S. and other countries. Other trademarks and trade names may be used in this document to refer to either the entities claiming the marks and names or their products. Dell disclaims proprietary interest in the marks and names of others.

September 2012
 Revision History

<table>
<thead>
<tr>
<th>Revision</th>
<th>Description</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>A00</td>
<td>Initial version</td>
<td>July 2012</td>
</tr>
<tr>
<td>A01</td>
<td>Reorganized some sections</td>
<td>September 2012</td>
</tr>
</tbody>
</table>
vStart 1000v for Enterprise Virtualization using VMware vSphere: Reference Architecture

Contents
1 Introduction.......................................................................................................................... 2
2 Audience............................................................................................................................... 2
3 Overview .............................................................................................................................. 2
4 Design Principles ................................................................................................................. 11
5 Reference Architecture ...................................................................................................... 11
6 Dell Blade Network Architecture ....................................................................................... 12
7 Network Architecture ......................................................................................................... 14
8 Storage Architecture .......................................................................................................... 16
9 Management Infrastructure ................................................................................................. 19
10 Scalability ........................................................................................................................... 20
11 Delivery Model .................................................................................................................. 21
12 Reference............................................................................................................................ 25

Figures
Figure 1: vStart 1000v Overview .......................................................................................... 3
Figure 2: vStart 1000v Network Topology (Logical View) .................................................... 12
Figure 3: I/O Connectivity for PowerEdge M620 Blade Server ........................................... 13
Figure 4: vSwitch, NPAR and I/O Module Configuration ...................................................... 15
Figure 5: Fibre Channel SAN Logical Connectivity .............................................................. 17
Figure 6: Management Components .................................................................................... 20
Figure 7: vStart 1000v Single Chassis: Rack Overview ......................................................... 23
Figure 8: vStart 1000v Two Chassis and Maximum Storage: Rack Overview ....................... 24
1 Introduction
The vStart 1000 solution is an enterprise infrastructure solution that has been designed and validated by Dell™ Engineering. It is available to be partially racked, cabled, and delivered to your site, to speed deployment. Dell Services will deploy and configure the solution tailored for business needs and ready to be integrated into your datacenter. vStart 1000 is offered in configurations with either VMware® vSphere® (vStart 1000v) or Microsoft® Windows Server® 2008 R2 with Hyper-V® role enabled (vStart 1000m) Hypervisors. This paper will define the Reference Architecture for the VMware vSphere based vStart 1000v solution.

vStart 1000v includes Dell PowerEdge™ M1000e blade chassis, Dell PowerEdge M620 blades, Dell Compellent™ Storage, Dell Force10™ network switches, Brocade Fibre Channel switches, and VMware vSphere. The solution also includes Dell PowerEdge R620 servers as management servers. VMware vCenter Server, Compellent Enterprise Manager, and OpenManage™ Essentials are included with the solution.

2 Audience
IT administrators and IT managers, who have purchased or are planning to purchase a vStart configuration, can use this document to understand the component details of the solution.

3 Overview
This section provides a high-level product overview of the VMware vSphere, Dell PowerEdge blade servers, Dell Force10 S4810, Force10 S55, Brocade 5100 Fibre Channel Switches, and Dell Compellent Storage, as illustrated in Figure 1. Readers can skip the sections of products with which they are familiar.
Figure 1: vStart 1000v Overview

**VMware vSphere 5**
- vMotion, Storage vMotion
- VMware HA and DRS

**Dell PowerEdge Blade Servers**
- Energy efficient PowerEdge M1000e enclosure
- 12th generation M620 blade server
- Flex Address
- CMC and iKVM for enclosure management

**Dell Force10 S4810 Switches**
- High-density 48-port 10 GbE switch with four 40 GbE uplinks
- Ultra-low-latency, non-blocking, cut-through switch for line-rate L2 and L3 performance
- Integrated network automation and virtualization tools via the Open Automation Framework

**Brocade 5100 Fibre Channel Switches**
- 1U 40 port high density switch with 8Gbps Fibre Channel

**Dell Compellent Storage array**
- Fluid Data Architecture with Thin Provisioning and Automated Tiered Storage
- Centralized management using Enterprise Manager
- Fast Track

**Integrated Management**
- Dell management plug-in for VMware vCenter
- Compellent vSphere plug-in
- OpenManage Essentials

**Cloud Enablement**
- VIS Creator for private cloud
- VMware vCloud Connector for Dell vCloud connectivity
Table 1 below describes the key solution components and the roles served.

Table 1: Solution Components

<table>
<thead>
<tr>
<th>Component</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypervisor Server</td>
<td>PowerEdge M1000e chassis with PowerEdge M620 Blade Servers and embedded VMware vSphere 5</td>
</tr>
</tbody>
</table>
| Management components hosted in the management infrastructure | • VMware vCenter Server  
• Compellent Enterprise Manager  
• Dell Management Plugin for VMware vCenter  
• Virtual Integrated System (VIS) Creator  
• OpenManage Essentials  
• VMware vCloud Connector  
• Compellent Plugin for VMware vCenter |
| LAN Switch                                 | Two Force10 S4810 and 10GbE Pass-through-k modules for the Chassis     |
| SAN Switch                                 | Two Brocade 5100 Fiber Switch and Dell 8 | 4 I/O modules for the chassis |
| Storage                                    | Two Compellent Series 40 controllers with SAS enclosures             |
| Management Infrastructure                  | Two PowerEdge R620 servers with embedded VMware vSphere 5 hosting management VMs. One Force10 S55 used as a 1Gb management switch |

3.1 VMware vSphere 5

VMware vSphere 5 includes the ESXi™ hypervisor as well as vCenter™ Server which is used to configure and manage VMware hosts. Key capabilities for the ESXi Enterprise Plus license level include:

- **VMware vMotion**: VMware vMotion technology provides real-time migration of running virtual machines (VM) from one host to another with no disruption or downtime.

- **VMware High Availability (HA)**: VMware HA provides high availability at the VM level. Upon host failure, VMware HA automatically re-starts VMs on other physical hosts running ESXi. VMware vSphere 5 uses Fault Domain Manager (FDM) for High Availability.

- **VMware Distributed Resource Scheduler (DRS) and VMware Distributed Power Management (DPM)**: VMware DRS technology enables vMotion to automatically achieve load balancing according to resource requirements. When VMs in a DRS cluster need fewer resources, such as
during nights and weekends, DPM consolidates workloads onto fewer hosts and powers off the rest to reduce power consumption.

- **VMware vCenter Update Manager**: VMware vCenter Update Manager automates patch management, enforcing compliance to patch standards for VMware ESXi hosts.

- **VMware Storage vMotion™**: VMware Storage vMotion enables real-time migration of running VM disks from one storage array to another with no disruption or downtime. It minimizes service disruptions due to planned storage downtime previously incurred for rebalancing or retiring storage arrays.

- **Host Profiles**: Host Profiles standardize and simplify the deployment and management of VMware ESXi host configurations. They capture and store validated configuration information, including host compliance, networking, storage, and security settings.

For more information on VMware vSphere, see [www.vmware.com/products/vsphere](http://www.vmware.com/products/vsphere).

### 3.2 VIS Creator

VIS Creator is an elastic, policy-driven platform that simplifies workload delivery and lifecycle management for IT department and end users. It empowers authorized end users to deploy their own workloads for quick response to business needs. Additionally it controls sprawl and optimizes utilization of virtual machines with automated provisioning and reclamation policies.

Dell VIS Creator enables authorized users to access a customized catalog of IT resources, some of which can be ready for deployment in minutes. With VIS Creator, authorized users can:

- Request new services from a customized catalog of resources
- Modify existing resources
- View details and consumption of existing resources

Utilizing the automated self-service delivery model of VIS Creator, users can request and purchase solutions that are the right size today, with the confidence that changes can be made easily and quickly in the future.

VIS Creator provides administrators and end users with a powerful set of tools, including the ability to:

- Eliminate time-consuming steps by formalizing the resource-deployment process and streamlining the approval process
- Respond to changes quickly by selecting and deploying standard IT resources almost instantly
- Customize an extensive set of out-of-the-box workflows to help automate many common tasks associated with workload deployment

Giving end users the ability to select, deploy, and manage IT resources sounds as though it could wreak havoc with internal governance, but the truth is quite the opposite. With VIS Creator in place, IT administrators can:

- Define policies to limit resource consumption
- Ensure that IT requests are provisioned to standards
- Enforce process workflows, as well as user rights and access control
vStart 1000v for Enterprise Virtualization using VMware vSphere: Reference Architecture

- Define which users have access to specific resources and processes

With its robust automation of policies and access controls, VIS Creator can help you establish consistent IT governance through processes, not people. For administrators, this means the ability to:

- Organize resources, policies, processes and management access controls for each group or service tier
- Reserve dedicated computer resources for a group from a shared physical infrastructure
- Utilize workflows for building, managing, reclaiming, decommissioning, and archiving each machine
- Define group membership and user rights

For more information on Dell Compellent, see [Dell.com/VISCreator](http://Dell.com/VISCreator).

### 3.3 Dell Cloud Connectivity using VMware vCloud Connector

VMware vCloud Connector lets you view, operate on and transfer your computing resources across vSphere and vCloud Director in your private cloud environment, as well as Dell vCloud public cloud.

- Expand your view across hybrid clouds. Use a “single pane of glass” management interface that seamlessly spans your private vSphere and public Dell vCloud environment.
- Extend your datacenter. Move VMs, vApps, and templates from private vSphere to a Dell vCloud to free up your on-premise datacenter resources as needed.
- Consume cloud resources with confidence. Run Development, QA, and production workloads using Dell vCloud, a VMware technology-based public cloud.

The Dell Cloud with VMware vCloud™ Datacenter is an enterprise-class, multi-tenant infrastructure-as-a-service (IaaS) public cloud solution that is hosted in secured Dell data centers. Utilizing VMware vCloud Connector, Dell Cloud provides you with unique hybrid cloud capabilities to extend your internal data center with Dell and VMware by transitioning your VMware virtualized workloads into our vCloud data center. vCloud hosting provides you with a secure, manageable and flexible public cloud application.

For more information, see [Dell vCloud website](http://Dell.com/vCloud).

### 3.4 Dell Management Plugin for VMware vCenter

Dell Management Plug-in for VMware vCenter is included in the solution. This enables customers to:

- Get deep-level detail from Dell servers for inventory, monitoring, and alerting — all from within vCenter
- Apply BIOS and Firmware updates to Dell servers from within vCenter
- Automatically perform Dell-recommended vCenter actions based on Dell hardware alerts
- Access Dell hardware warranty information online
- Rapidly deploy new bare metal hosts using Profile features

For more information, see the web page for [Dell Management Plugin for VMware vCenter](http://Dell.com/ManagementPlugin).
3.5 OpenManage Essentials and OpenManage Server Administrator

The Dell OpenManage™ Essentials (OME) Console provides a single, easy-to-use, one-to-many interface through which to manage resources in multivendor operating system and hypervisor environments. It automates basic repetitive hardware management tasks — like discovery, inventory, and monitoring— for Dell servers, storage, and network systems. OME employs the embedded management of PowerEdge™ servers — Integrated Dell Remote Access Controller 7 (iDRAC7) with Lifecycle Controller — to enable agent-free remote management and monitoring of server hardware components like storage, networking, processors and memory.

OpenManage Essentials helps you maximize IT performance and uptime with capabilities like:

- Automated discovery, inventory and monitoring of Dell PowerEdge™ servers, EqualLogic™ and PowerVault™ storage and PowerConnect™ switches
- Server health monitoring as well as BIOS, firmware and driver updates for Dell PowerEdge servers, blade systems and internal storage
- Control of PowerEdge servers within Windows®, Linux®, VMware® and Hyper-V® environments

For more information on OpenManage Essentials, see Dell.com/openmanageessentials.

OpenManage Server Administrator (OMSA) is installed on each ESXi server to provide comprehensive one-to-one management solution.

3.6 Dell PowerEdge Blade Servers

Blade Modular Enclosure: The Dell PowerEdge M1000e is a high-density, energy-efficient blade chassis that supports up to sixteen half-height blade servers, or eight full-height blade servers, and six I/O modules. A high-speed passive mid-plane connects the server modules to the I/O modules, management, and power in the rear of the chassis. The enclosure includes a flip-out LCD screen (for local configuration), six hot-pluggable/redundant power supplies, and nine hot-pluggable N+1 redundant fan modules.

Blade Servers: The Dell PowerEdge M620 blade server is the 12th generation half height blade server offering:

- New high-efficiency Intel® Xeon® E5-2600 family processors for more advanced processing performance, memory, and I/O bandwidth.
- Greater memory density than any previous PowerEdge server. Each PowerEdge M620 can deploy up to 24 x 16GB DIMMs, or 768GB RAM per blade - 12TB or RAM in a single PowerEdge M1000e chassis.
- 'Agent Free’ management with the new iDRAC7 with Lifecycle Controller allows customers to deploy, update, maintain, and monitor their systems throughout the system lifecycle without a software management agent, regardless of the operating system.
- The PowerEdge Select Network Adapter on the PowerEdge M620 offers three modular choices for embedded fabric capability. With 10Gb CNA offerings from Broadcom, QLogic & Intel, our customers can choose the networking vendor & technology that’s right for them and their applications, and even change in the future as those needs evolve over time. The Broadcom &
QLogic offerings offer Switch Independent Partitioning technology, developed in partnership with Dell, which allows for virtual partitioning of the 10Gb ports.

I/O Modules: The enclosure provides three redundant fabrics using six I/O modules. The modules can be populated with Ethernet switches, Fibre Channel (FC), and pass-through modules. InfiniBand™ switch modules are also supported.

Chassis Management: The Dell PowerEdge M1000e has integrated management through a redundant Chassis Management Controller (CMC) module for enclosure management and integrated Keyboard, Video, and Mouse (iKVM) modules. Through the CMC, the enclosure supports FlexAddress Plus technology, which enables the blade enclosure to lock the World Wide Names (WWN) of the FC controllers and Media Access Control (MAC) addresses of the Ethernet controllers to specific blade slots. This enables seamless swapping or upgrading of blade servers without affecting the LAN or SAN configuration.

Embedded Management with Dell’s Lifecycle Controller: The Lifecycle Controller is the engine for advanced embedded management and is delivered as part of iDRAC Enterprise in Dell PowerEdge 12th generation servers. It includes 1GB of managed and persistent storage that embeds systems management features directly on the server, thus eliminating the media-based delivery of system management tools and utilities previously needed for systems management. Embedded management includes:

- Unified Server Configurator (USC) aims at local 1-to-1 deployment via a graphical user interface (GUI) for operating system install, updates, configuration, and for performing diagnostics on single, local servers. This eliminates the need for multiple option ROMs for hardware configuration.
- Remote Services are standards-based interfaces that enable consoles to integrate, for example, bare-metal provisioning and one-to-many OS deployments, for servers located remotely. Dell’s Lifecycle Controller takes advantage of the capabilities of both USC and Remote Services to deliver significant advancement and simplification of server deployment.
- Lifecycle Controller Serviceability aims at simplifying server re-provisioning and/or replacing failed parts and thus reduces maintenance downtime.

For more information on Dell Lifecycle Controllers and blade servers, see http://content.dell.com/us/en/enterprise/dcsm-embedded-management and Dell.com/blades.

3.7 Dell Force10 S4810 Switches

The Force10 S-Series S4810 is an ultra-low-latency 10/40 GbE Top-of-Rack (ToR) switch purpose-built for applications in high-performance data center and computing environments. Leveraging a non-blocking, cut-through switching architecture, the S4810 delivers line-rate L2 and L3 forwarding capacity with ultra-low latency to maximize network performance. The compact S4810 design provides industry leading density of 48 dual-speed 1/10 GbE (SFP+) ports as well as four 40GbE QSFP+ uplinks to conserve valuable rack space and simplify the migration to 40Gbps in the data center core. (Each 40GbE QSFP+ uplink can support four 10GbE ports with a breakout cable).

Powerful QoS features coupled with Data Center Bridging (DCB) support via a future software enhancement, make the S4810 ideally suited for iSCSI storage environments. In addition, the S4810 incorporates multiple architectural features that optimize data center network flexibility, efficiency,
and availability, including Force10's stacking technology, reversible front-to-back or back-to-front airflow for hot/cold aisle environments, and redundant, hot-swappable power supplies and fans.

For more information on Force10 switches, see [Dell.com/force10](https://www.dell.com/force10).

### 3.8 Dell Force10 S55

The Dell Force10 S-Series S55 1/10 GbE ToR switch is designed for high-performance data center applications. The S55 leverages a non-blocking architecture that delivers line-rate, low-latency L2 and L3 switching to eliminate network bottlenecks. The high-density S55 design provides 48GbE access ports with up to four modular 10GbE uplinks in 1-RU to conserve valuable rack space. The S55 incorporates multiple architectural features that optimize data center network efficiency and reliability, including reversible front-to-back or back-to-front airflow for hot/cold aisle environments and redundant, hot-swappable power supplies and fans.

For more information on Force10 switches, see [Dell.com/force10](https://www.dell.com/force10).

### 3.9 Brocade 5100

The Brocade 5100 switch is a high density FC switch providing 40 ports in a 1U form factor. The 5100 includes redundant power supplies and fans making it well suited to the high availability needs of virtualization infrastructures. It also includes the Ports-on-Demand capabilities for cost reduction when installing in smaller environments.

For more information on Brocade 5100 Fibre Channel Switches, see [Dell.com/brocade](https://www.dell.com/brocade).

### 3.10 Dell 8/4 Gbps FC SAN Module

The Dell 8/4 Gbps FC SAN Module is a 24-port FC module with eight external ports and 16 internal ports that installs in a Dell PowerEdge M1000e Blade Enclosure. Built on industry-standard N_Port ID Virtualization (NPIV) technology, the module eliminates the traditional challenges of heterogeneous switch-to-switch interoperability and can non-disruptively connect Dell blades to NPIV-enabled FC SANs, including Brocade, Cisco, McData, and others. The Dell 8/4 Gbps FC SAN Module eliminates incremental switch management and configuration by presenting FC connections as a logical device (rather than switch domains) to the SAN fabric. The module enables the benefits of port aggregation, failover, and redundancy without the complexities of additional SAN switches or additional switch domains.


### 3.11 Dell Compellent Series 40 Storage Center

Dell Compellent Storage is available in a dual-controller configuration with 4|8 Gb FC interconnects and 4 Gb FC or 36 Gb SAS backend interconnects. The Series 40 controllers connect to servers without server-side agents and automatically fail over to another controller when clustered. The Series 40 features six PCI-e expansion slots, battery-less cache, redundant hot-swappable power supplies, and cooling fans to help ensure non-stop operation and high availability. Virtual ports increase port capacity, disk bandwidth, I/O connectivity, and port failover.
Features of the Dell Compellent Series 40 Storage Array include:

- **Fluid Data Architecture** - Storage is managed at the most granular level with built-in system intelligence to enable the dynamic flow of enterprise data.

- **Storage Virtualization** - Storage is virtualized at the disk level to create a flexible pool of storage resources shared by all servers all the time.

- **Thin Provisioning** - Allocation is completely separated from utilization so any size volume can be created at any time, yet capacity is only consumed when data is written.

- **Automated Tiered Storage** - Data dynamically cascades from tier to tier according to actual usage, freeing up high-performance drives for mission-critical applications.

- **Space-efficient Replays** - Continuous snapshots only capture changes in data for real-time protection with instant recovery to any point in time.

- **Thin Replication** - Data is replicated between local and remote sites using space-efficient snapshots and native IP or FC connectivity, eliminating the need for high-speed data links or identical system configurations.

- **Unified Storage Resource Management** - All storage resources are managed through a single point-and-click interface, providing a complete view of the entire storage environment.

- **Open, Agile Hardware Platform** - Storage is designed for persistence, not obsolescence, leveraging a single modular hardware platform coupled with technology independence.

**Compellent Enterprise Manager**: Compellent Enterprise Manager is included in the solution. It simplifies network storage management by providing a single, centralized console for the administration of multiple local and remote Compellent systems. Users can configure and verify remote replication processes, monitor storage capacity and disk utilization in real time, and generate comprehensive enterprise storage usage and performance reports.

For more information on Dell Compellent, see [Dell.com/Compellent](http://Dell.com/Compellent). Contact Dell sales representative for more information on Compellent storage configurations and sizing guidelines.

### 3.12 PowerEdge R620 Management Server

The Dell PowerEdge R620 uses Intel Xeon E5-2600 series processors and Intel chipset architecture in a 1U rack mount form factor. These servers support up to ten 2.5” drives and provide the option for an LCD located in the front of the server for system health monitoring, alerting, and basic management configuration. An AC power meter and ambient temperature thermometer are built into the server, both of which can be monitored on this display without any software tools. The server features two CPU sockets and 24 memory DIMM slots.

For more information, see the PowerEdge R620 guides at [Dell.com/PowerEdge](http://Dell.com/PowerEdge).
4 Design Principles

The following principles are central to the design and architecture of the vStart 1000v Solution.

1. **Redundancy with no single point-of-failure**: Redundancy is incorporated in the critical aspects\(^1\) of the solution, including server high availability features, networking, and storage.

2. **Management**: Provide integrated management using VMware vCenter, Dell Management Plug-in VMware vCenter, Dell OpenManage Essentials, and Compellent plug-in for VMware vCenter.

3. **Cloud Enabled**: The solution includes VIS Creator, which enables customers to manage their virtualization infrastructure as a private cloud. The private cloud can in turn be connected to Dell vCloud using VMware vCloud Connector.

4. **Integration into an existing data center**: This architecture assumes that there is an existing 10 Gb Ethernet infrastructure with which to integrate.

5. **Hardware configuration for virtualization**: This solution is designed for virtualization for most general cases. Each blade server is configured with appropriate processor, memory, host bus, and network adapters as required for virtualization.

6. **Racked, Cabled, and Ready to be deployed**: vStart is available partially racked, cabled, and delivered to the customer site, ready for deployment. Components are configured and racked to optimize airflow and thermals. Based on customer needs, different rack sizes and configurations are available to support various datacenter requirements.

7. **Power, Cooling, and Weight Considerations**: vStart 1000v solution is configured with Power Distribution Units (PDUs) to meet the power requirements of the components as well as regional constraints. Power consumed, cooling required, and information regarding rack weight are provided to enable customers to plan for the solution.

8. **Flexible configurations**: vStart 1000v is pre-configured to suit most customer needs for a virtualized infrastructure. The solution also supports additional options, such as configuring racks, server processors, server memory, and storage, based on customer needs.

5 Reference Architecture

This solution consists of a PowerEdge M1000e chassis populated with PowerEdge M620 blade servers running VMware ESXi. Figure 2 provides high-level reference architecture for the solution. The figure shows high-level logical connectivity between various components. Subsequent sections provide more detailed connectivity information.

---

\(^1\) Out of band management is not considered critical to user workload and does not have redundancy.
6 Dell Blade Network Architecture

The Dell blade chassis has three separate fabrics referred to as A, B, and C. Each fabric has two I/O modules, for a total of six I/O modules slots in the chassis. The I/O modules are A1, A2, B1, B2, C1, and C2. Each I/O module can be an Ethernet physical switch, an Ethernet pass-through module, FC switch, or FC pass-through module. Each half-height blade server has a dual-port network daughter card (NDC)
and two optional dual-port mezzanine I/O cards. The NDC connects to Fabric A. One mezzanine I/O card attaches to Fabric B, with the remaining mezzanine I/O card attached to Fabric C.

In this solution, the Chassis Fabric A contains 10 GbE Pass-Through-k modules and is used for LAN. Fabric B contains Dell 8|4 Gbps SAN modules and is used for SAN. The Fabric C is unused.

PowerEdge M620 blade servers use a Broadcom 57810-k Dual port 10GbE KR bNDC (blade Network Daughter Card) to connect to the fabric A. Pass-Through-K modules uplink to Dell Force10 S4810 network switches providing LAN connectivity. QLogic QME2572 8 Gbps Fibre Channel I/O mezzanine cards are used to connect to Dell 8|4 Gbps SAN modules. The uplinks of Dell 8|4 Gbps SAN modules connect to Brocade 5100 switches providing SAN connectivity.

Figure 3 below illustrates how the fabrics are populated in a Dell blade server chassis and how the I/O modules are utilized.

**Network Interface Card Partition (NPAR):** NPAR allows splitting the 10GbE pipe on the NDC with no specific configuration requirements in the switches. With NPAR, administrators can split each 10GbE port of an NDC into four separate partitions, or physical functions and allocate the desired bandwidth and resources as needed. Each of these partitions is enumerated as a PCI Express function that appears as a separate physical NIC in the server, operating systems, and hypervisor. The vStart 1000v solution...
takes advantage of NPAR. Partitions are created for various traffic types and bandwidth is allocated, as described in the following section.

7 Network Architecture

LAN traffic in the vStart 1000v solution is categorized into four traffic types: VM traffic, management traffic, vMotion traffic, and Out-of-Band (OOB) management traffic. VM traffic, management traffic, and vMotion traffic are associated with the blade servers. OOB management traffic is associated with CMC, iDRAC, Brocade management, and Compellent management traffic. This section provides the network best practices for implementing VMware vSphere 5 on Dell blade servers and Force10 S4810 switches.

On the PowerEdge M620 blade servers, each of the 10GbE ports is partitioned into four network interfaces using NPAR resulting in a total of eight NICs for each server. A virtual switch is created for the three traffic types: VM traffic, management traffic, and vMotion traffic. Two partitions, one from each physical network port, are connected as uplinks to the virtual switch. This creates a team of two network ports, enabling NIC failover and load balancing for each vSwitch. The resultant design and concept is illustrated in Figure 4.
Traffic Isolation using VLANs: LAN traffic is separated into four unique VLANs; one VLAN each for management, out-of-band management, vMotion, and VM traffic. Network traffic is tagged with the respective VLAN ID for each traffic type in the virtual switch. Routing between the management and out-of-band management VLANs is required to be configured in the core or the S4810 switches. Additionally, the Force10 S4810 switch ports that connect to the blade servers need to be configured in VLAN trunk mode to pass traffic with different VLANs on a given physical port.

Load Balancing and Failover: This solution uses Route based on the originating virtual switch port ID configuration at the vSwitch for load balancing the LAN traffic. Any given virtual network adapter will use only one physical adapter port at any given time. In other words, if a VM has only one virtual NIC, it will use only one physical adapter port at any given time. The reason for choosing this option is that it is easy to configure and provides good load balancing across VMs, especially in the case of a large number of VMs.

Inter Switch Links for S4810s: The two Force10's S4810 switches are connected using Inter Switch Links (ISLs) using two 40 Gbps QSFP+ links. Link Aggregation Groups (LAGs) are then created between the two 40 Gbps QSFP+ ports, providing a path for communication across the switches.

Note that the two switches can also be stacked together. However, this is not recommended, as this configuration will incur downtime during firmware updates of the switch or failure of stack links.
Uplinks: There are several options to uplink the Force10 switches to the core network. Selecting the uplink option depends on the customer core network and customer requirements. One simple option is to create multiple uplinks on each switch and connect them to the core network switches. Uplink LAGs can then be created from the S4810 to the core network.

8  Storage Architecture
In this solution, Compellent Storage Center is connected to the Dell blade servers using Brocade 5100 FC switches.

8.1  Fibre Channel Fabric Architecture
The solution is configured with two FC fabrics as shown in Figure 5. The two fabric design ensures that changes to one fabric do not impact the other fabric.
8.2 Storage Connectivity

Connectivity between the Dell FC SAN Module and Brocade 5100: Each blade is populated with a QLogic QME2572 8 Gbps Fibre Channel I/O mezzanine card, which is used to connect to Fabric B. Fabric B is populated with Dell 8|4Gbps SAN modules. The Dell FC SAN Module is configured to operate as a port aggregator for aggregating 16 internal ports to four external ports. Dell FC SAN module port aggregator operates in access gateway mode for providing N_Port ID virtualization (NPIV) functionality. The following are the FC ports that the FC SAN Module uses:

- F_Port - internal fabric port that connects a blade server (HBA).
- N_Port - external node port that connects to a switch.
Using Brocade Port Trunking, multiple external N_Ports are combined to form a single logical port. The FC SAN Module uses internal F-port to external N-port mappings as configured to direct traffic from the blade server (HBAs) to the fabric.

In this solution, a trunk group is configured on each Brocade 5100 switch with four FC ports. Each Dell SAN module automatically forms a trunk with the corresponding external N-ports connected to the FC switch ports.

**Connectivity between Brocade 5100 and Compellent Storage Controller:** In the solution, each Compellent Series 40 storage controller is configured with two quad ports FC HBAs. Two ports per HBA are used to connect each storage controller to the two Brocade 5100 switches.

**Compellent Storage Connectivity:** In this solution, Compellent series 40 storage controllers are configured with two quad port 6Gb/s SAS HBAs. Using the two quad port SAS HBAs, multiple Compellent 24 bay 2.5” SAS enclosures are connected. The enclosures are grouped into two redundant daisy chained connections for optimal performance. Each daisy chain loop can have a maximum of 96 drives.

### 8.3 Performance

Dell Compellent Series 40, with the dual-controller configuration, 8 Gb Fibre Channel interconnects provides high bandwidth for data flows. This bandwidth is complemented with a large variety of drives in multiple speeds and sizes. The Series 40 also uses virtual port IQNs and WWNs, thereby enabling higher throughput and fault tolerance.

### 8.4 Drive Types and Automated Tiered Storage

In the vStart 1000v solution, the number of storage enclosures and the drives in the enclosures can be customized based on customer requirements. Administrators can mix SSD and SAS drives in the same system, as well as SAS drives with the same form factor (but different speeds and capacities) in the same storage enclosure. A maximum of 16 enclosures is supported in vStart 1000v.

High speed drives are assigned to higher tiers and low speed drives to lower tiers. Compellent Storage Center automatically configures RAID levels for these tiers and automatically moves the data between the tiers based on access patterns. Compellent Fluid Data storage dynamically moves data to the optimal tier based on actual usage. The most active blocks reside on high-performance SSD, FC, or SAS drives, while infrequently accessed data migrates to lower-cost, high-capacity SAS or SATA drives. For more details, refer to the Automated Tiered Storage web page. Automated Tiered Storage requires Data Progression™ licenses for Compellent.

### 8.5 RAID Array Design

Dell Compellent Series 40 supports RAID 5, 6 and 10. The Compellent Storage Center will dynamically set up RAID based upon the demands of applications accessing data on the storage tier(s).
8.6 Multipath Configuration
In the solution, VMware Native Multipath Plug-In (NMP) is used to provide multi-pathing. Path Selection Plug-Ins (PSPs) run with the VMware NMP and are responsible for choosing a physical path for I/O requests. Round Robin (VMW_PSP_RR) path selection algorithm is the recommended configuration for Compellent Storage Center. Round Robin uses a path selection algorithm that rotates through all available active optimal paths enabling load balancing across the paths. This ensures all the paths are used to provide the maximum bandwidth and balance I/O across the paths/fabrics.

9 Management Infrastructure
Two PowerEdge R620 servers and one Force10 S55 1Gb Ethernet switch are used for management infrastructure. The PowerEdge R620 servers are connected to the Force10 S4810 switches using Broadcom 57810 Dual Port 10Gb Network Adapter. The servers are connected to the Compellent storage through the Brocade 5100 switches using a QLogic QLA2562 8Gbps Fibre Channel Card.

Note that the Compellent storage is shared between management cluster and compute cluster. The Compellent storage must be sized so that sufficient bandwidth is allocated for both the management VMs and compute VMs.

The PowerEdge R620 servers run VMware ESXi 5.0 hypervisor and are a part of the unique vSphere Cluster. VMware High Availability is enabled in that cluster to provide HA for virtual machines. Admission control is disabled in the VMware HA Cluster. If admission control is enabled, VMware HA would prevent putting one of the management servers in maintenance mode, since this would violate HA policy of having more than one active server in the cluster.

The following management components are installed as virtual machines in the management infrastructure as illustrated in Figure 6:

- VMware vCenter Server
- Dell OpenManage Plugin for vCenter
- Dell Compellent Enterprise Manager
- VIS Creator
- Dell OpenManage Essentials
- VMware vCloud Connector Server
- VMware vCloud Connector Node

Compellent Plugin for vCenter is installed along with VMware vCenter client in VMware vCenter Server VM. OpenManage Server Administrator Web Server is installed in the OME VM to enable one-to-one management of ESXi servers.
**Required Software Components:** The following components are required for the management components:

- **Active Directory® (AD)** - Required for the following scenarios:
  - Required if VIS Creator is part of the VMware vSphere solution. Connectivity is required between the AD Server and VIS Creator to be deployed in the management servers.
  - VMware vCenter can also be configured to use Active Directory, but is not required.

- **Domain Name Server (DNS)** - DNS must be available on the management network.

- **Network Time Protocol (NTP) Server** - NTP must be available on the management network.

- **SQL Requirements** - Required for vCenter, VIS Creator and Compellent Enterprise Manager.

- **SMTP**
  - Required for VIS Creator.
  - Optionally utilized by other components of the management stack, including OME & vCenter, for notifications.

**10 Scalability**

As workloads increase, the solution can be scaled to provide additional compute and storage resources independently.

**Scaling Compute and Network Resources:** This solution is configured with two Force10 S4810 network switches. Up to two PowerEdge M1000e chassis can be added to the two Force10 switches. In order to scale the compute nodes beyond two chassis, new Force10 S4810 switches need to be added. Additional switches can either be stacked together and/or connected to this distribution switch based on customer needs.
Scaling Storage Resources: Compellent storage can be scaled seamlessly and independent of the compute and network architectures. Additional drives and enclosures can be added to the existing controllers. New volumes can be created or existing volumes can be expanded to utilize the capacity in the added enclosures. The vStart 1000v solution can scale up to maximum of 16 array enclosures. To scale beyond this, additional racks or controllers can be added. Compellent Series 40 controller can scale up to a maximum of 960 drives.

11 Delivery Model

This Reference Architecture can be purchased as a complete solution, the vStart 1000v. This solution is available to be partially racked, cabled, and delivered to the customer site, to speed deployment. Dell Services will deploy and configure the solution tailored to the business needs of the customer and based on the architecture developed and validated by Dell Engineering. For more details or questions about the delivery model, please consult with your Dell Sales representative.
Figure 7 below shows the vStart 1000v solution with a single chassis. Figure 8 shows vStart 1000v with two chassis and maximum storage enclosures. Note that switches shown in figures are shown mounted forward for representation. In actual use, ports face the back of the rack. PDUs shown are for illustration and will vary by region or customer power requirements. Additional PDUs are utilized within the rack.
Figure 8: vStart 1000v Two Chassis and Maximum Storage: Rack Overview
12 Reference

- Dell TechCenter vStart Wiki
- Dell vStart: Dell.com/vStart
- Dell Force10 Switch Details
- Brocade 5100 Product Details
- Dell 8/4 Gbps Fibre Channel SAN module
- Dell Compellent Storage
  - Automated Tiered Storage
  - Fast Track
  - Dell Compellent Plug-In for VMware vCenter
- VMware vSphere links:
  - Virtual Networking Concepts
- Dell Management Plug-in for VMware vCenter Documentation
- NPAR: Enhancing scalability through network interface card partitioning