ENHANCED BUSINESS CONTINUITY WITH APPLICATION MOBILITY ACROSS DATA CENTERS

P/N: H7238
Date: June 2010

Table of Contents

Executive summary

The evolution of the global economy is redefining the paradigm of business continuity. Most businesses can no longer afford to operate only in their native time zones, but are mandated to operate round the clock. This puts a very heavy burden on Information Technology (IT) groups because the operational windows set aside for planned downtime of the data center are reduced. IT groups are forced to perform most of the operational tasks within and across data centers in a very limited amount of time in order to adhere to the service level agreements (SLAs) defined for business continuity, while the procedures and processes defined for unplanned outages remain the same. This implies that IT departments need to adapt and innovate to provide the required SLAs. Data centers are being extended beyond geographic boundaries, and the IT industry is moving toward the concept of a virtualized data center. The need of the hour is to enable IT departments to reduce the planned downtime of business-critical applications without compromising the reliability of the data center operations.

This solution highlights an end-to-end architecture in which applications can be seamlessly migrated between data centers with zero application downtime and very minimal application performance degradation (≤5%). Furthermore, the architecture can be applied to connect two or more Vblock™ Infrastructure Packages between data centers, thus extending the solution's benefit of business continuity to the Vblock Infrastructure Packages environment. The reduction of planned downtime greatly enhances the overall business continuity of the operations, but also enables the IT department to be more efficient in its operations without affecting any of the performance SLAs. This white paper is based on innovations in server, network, and storage virtualization driven by VMware®, Cisco®, and EMC®, respectively. Figure 1 provides a high-level overview of this architecture.

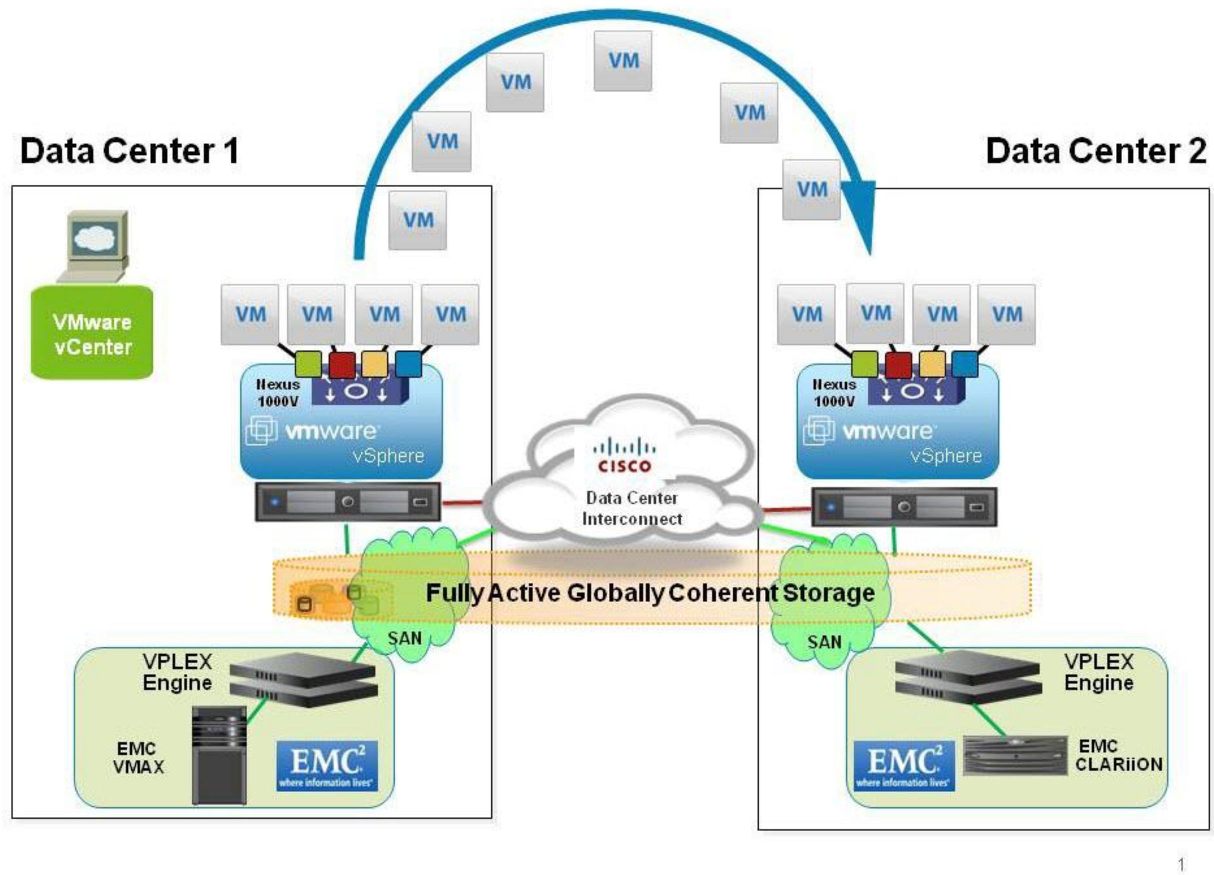

Figure 1 **VMware VMotion across data centers**

Drivers for application mobility across data centers

Data centers are being extended beyond geographic boundaries and the IT industry is moving toward the concept of a virtualized data center. This trend is driven by applications running on servers that can be virtualized using hypervisors, such as VMware vSphere™. The changing model of data center management and provisioning allows VMware VMotion to be used for several purposes without violating the application SLAs.

- **Data center maintenance without downtime**: Applications on a server or data center infrastructure requiring maintenance can be migrated offsite without downtime.
- **Disaster avoidance**: Data centers in the path of natural threats (such as hurricanes) can proactively migrate the mission-critical application environment to another data center.
- **Data center migration or consolidation**: Migrate applications from one data center to another without business downtime as part of a data center migration or consolidation effort.
- **Data center expansion**: Migrate virtual machines to a secondary data center as part of data center expansion to address power, cooling, and space constraints in the primary data center.
- **Workload balancing across multiple sites**: Migrate virtual machines between data centers to provide compute power from data centers closer to the clients ("follow the sun") or to load balance across multiple sites. Enterprises with multiple sites can also conserve power and reduce cooling costs by dynamically consolidating virtual machines in fewer data centers (automated by VMware Dynamic Power Management, or DPM), another feature enabling the green data center of the future.

The application mobility discussed in this document provides the foundation necessary to enable cloud computing—for example, cloud import and export—providing the flexibility to move virtual machines into the cloud from an enterprise data center, to move them between different clouds, and to move them back into the enterprise data center.

Note: This document does not address disaster recovery in the event of a data center outage. Cisco, EMC, and VMware provide comprehensive disaster recovery solutions, and these solutions are discussed in other documents.

Challenges and considerations with application mobility between data centers

Successful VMware VMotion migration between data centers in different physical locations poses certain challenges. VMware VMotion migration across long distances requires careful evaluation of data center network and storage design, as shown in Figure 2.

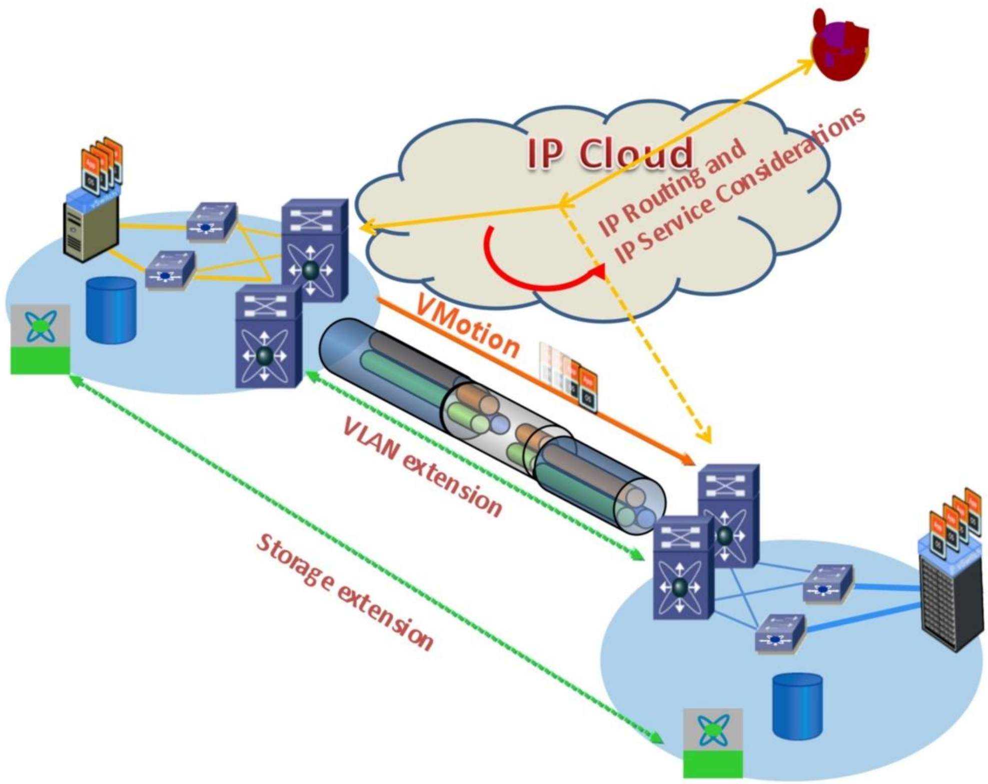

Figure 2 Infrastructure challenges with VM mobility across data centers

These challenges are:

- **VLAN extension**: Layer 2 VLANs must be extended across data centers without compromising the availability, resilience, and security that exist within a single physical location. Layer 2 domain elasticity should be possible over different connection media such as dark fiber or IP and Multiprotocol Label Switching (IP/MPLS)-based infrastructure. The obvious challenge in extending VLANs across data centers is the possibility of increased risk from expanded spanning tree domains, the possibility of loops if the spanning tree is isolated across data centers, and packet broadcast and flooding in multiple sites. However, some broadcast traffic is essential for application-level communication. Layer 2 domain elasticity should thus help ensure a loop-free topology while isolating the spanning tree across data centers, scalability to connect multiple data center sites, and optimal use of WAN bandwidth.

- **Storage extension**: The availability of identical storage devices with concurrent read- and write-access to two VMware ESX servers in physically separated data centers is critical to a successful application migration across geographies. The WAN design should intelligently and optimally manage the large data sets associated with applications. Storage network designs and storage systems should take into account these parameters to help ensure that data is not only available, but that it is secure, with I/O latencies and performance that will not affect the SLAs of the applications. Application bandwidth and latency requirements partially determine the storage architecture: For example, workflows that consist primarily of data read operations minimize the bandwidth requirements and the average I/O response time elongation for the federation of the data between the data centers.

- **IP routing and IP service considerations**: An application migrated across data centers using VMware VMotion maintains its existing IP and MAC addresses. If the traffic to the virtual machine originates in the same Layer 2 domain, the Layer 2 extension will suffice. If the traffic to the virtual machine is traversing a Layer 3 network, such as an IP cloud or the Internet, the traffic needs to be rerouted to the new data center location. Existing application sessions may continue to be routed through the existing data center due to specific or existing IP service requirements such as firewalls. Because of this behavior, the following IP routing considerations are required:

    - Routing from remote clients to application servers: Requires intelligent routing-based or Domain Name System (DNS)-based mechanisms to adapt to IP mobility.

    - Routing from application servers to remote clients: Requires forwarding of application traffic to the appropriate default gateway (preferably in the local data center pod) to achieve optimal routing as well as symmetrical routing for IP services such as firewalls.

- **WAN characteristics**: The WAN bandwidth and latency requirements for VMware VMotion are critical factors in a successful VMotion application migration across data centers.

- **VMware VMotion considerations**: VMware VMotion application mobility is based on certain infrastructure requirements:

    - A minimum bandwidth of 622 Mb/s is required between data centers and at least 1 Gb/s links within a data center. The source and destination VMware ESX servers must be on the same IP subnet and broadcast domain.

    - The maximum round-trip latency between the source and destination VMware ESX servers cannot exceed 5 milliseconds. Based on the speed of light over fiber and certain guard bands for network delays, a maximum distance of 400 km is supported today.

    - The IP subnet on which the virtual machine resides must be accessible from both the source and destination VMware ESX servers. This requirement is very important because a virtual machine retains its IP address when it moves to the destination VMware ESX server, to help ensure that its communication with the outside world (for example, with TCP clients) continues smoothly after the move.

    - The data storage location, including the boot device used by the virtual machine, must be active and accessible by both the source and destination VMware ESX servers at all times. If servers are present in two distinct locations, the sets of data must be identical.

    - Access from VMware vCenter (the vSphere management GUI) to both VMware ESX servers must be available to accomplish the migration. This implies that a single VMware vCenter server spans both data centers.

Cisco, EMC, and VMware joint solution overview

The solution, jointly engineered by the three companies, addresses a VMware VMotion migration with multiple components, one of which is a storage solution that removes physical barriers within a single data center and multiple data centers by presenting a single copy of data independent of the physical location. Applications provisioned on the VMware vSphere server can be migrated, live, across the data centers or a private cloud with no application downtime. Migration is granular at the virtual machine level. The validated architecture is shown in Figure 3 (logical topology) and Figure 4 (physical topology). This architecture uses the following components:

- **VMware VMotion technology** with a VMware vSphere server cluster enabled with VMware VMotion in each data center. A VMware vCenter server manages all virtual machine migrations.

- **Cisco Overlay Transport Virtualization (OTV) technology** on Cisco Nexus® 7000 Series switches as data center interconnect (DCI) technology, to enable data center interconnection across an IP network connecting the data centers.

- **Distributed Virtual Switching** using Cisco Nexus 1000V as a distributed virtual switch at the virtual machine level in each data center.

- **DNS-based Routing Optimization** to optimally route users to the application in the new data center using automatic interaction between vCenter, Cisco Global Site Selector (GSS), and the Cisco Application Control Engine (ACE).

- **EMC VPLEX™** family with the EMC GeoSynchrony™ operating system to remove physical barriers within a single data center and multiple virtualized data centers and to enable a single copy of data to be shared, accessed, and relocated over distance.

- **Cisco MDS-based storage area network (SAN)** to connect VPLEX clusters to storage arrays and to connect VPLEX clusters across data centers.

Figure 3 shows that, conceptually, any storage can be connected to EMC VPLEX (including Vblock Infrastructure Packages in the future) to provide application mobility and enhanced business continuity across data centers. For more detail on individual components of this architecture, refer to Appendix A.

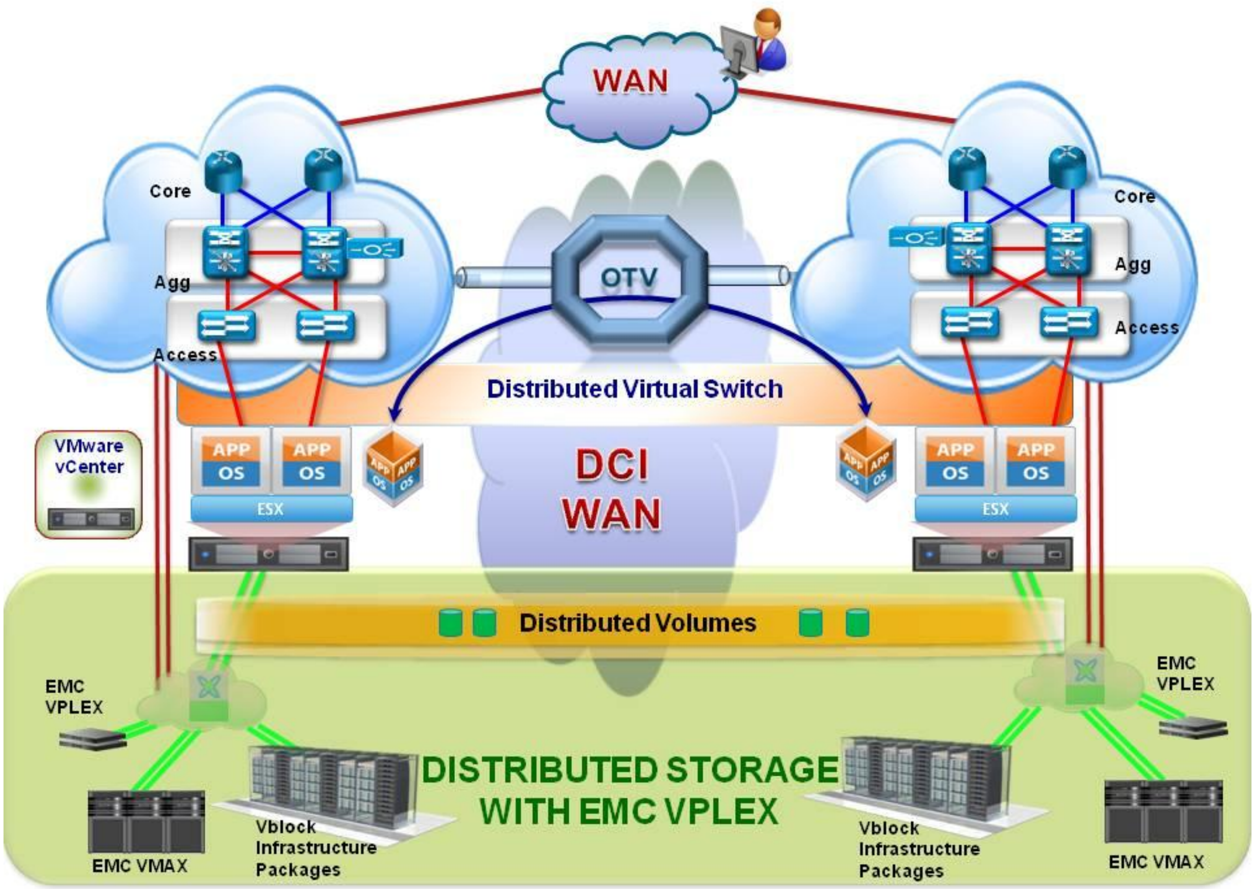

Figure 3 VMware, Cisco, and EMC validated architecture – logical topology

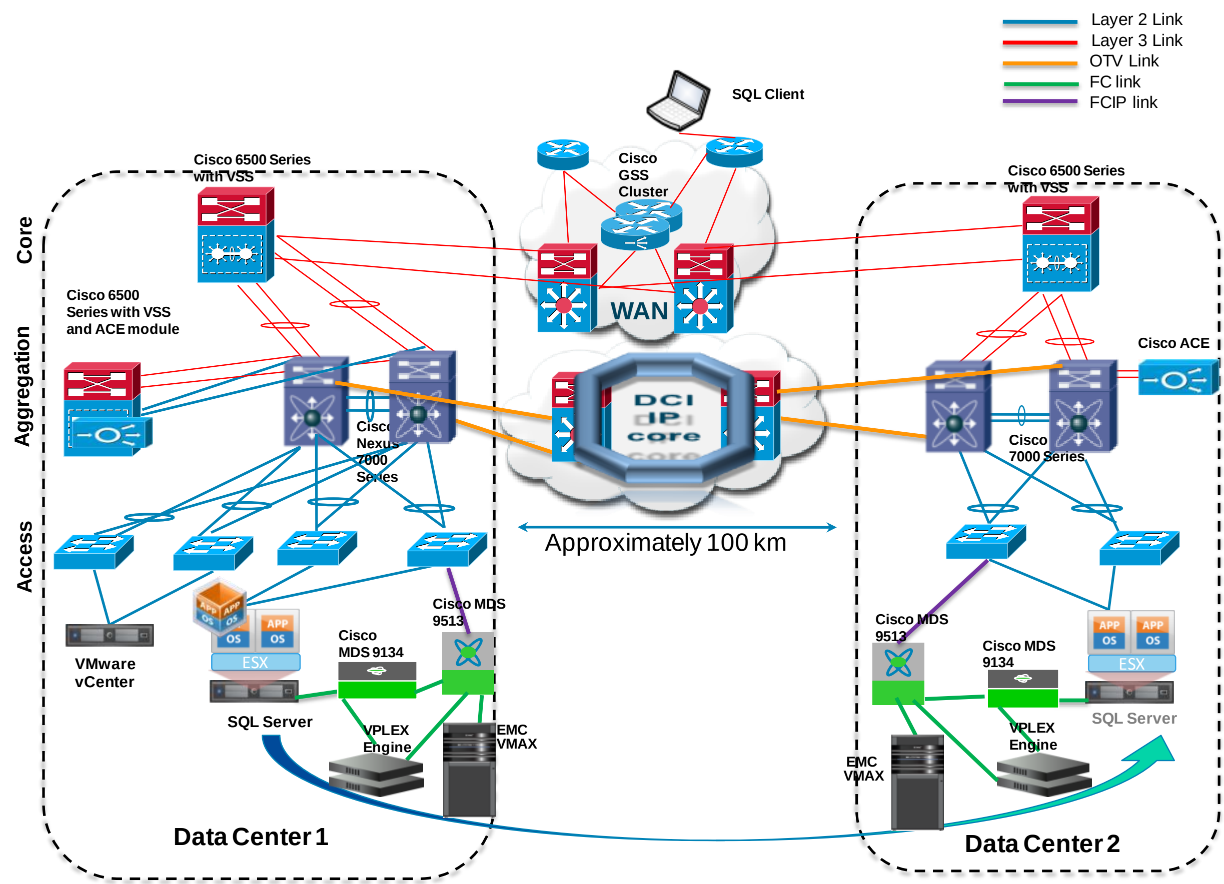

Figure 4 VMware, Cisco, and EMC validated architecture – physical topology

The network topology used in the joint solution test simulates two data centers that are about 100 km apart and are connected through an IP network. The distance between the two data centers was simulated using a WAN simulator (Empirix PacketSphere).

Both data centers use Cisco Nexus 7000 Series switches at the aggregation layer. Both data centers use OTV to extend Layer 2 across long distances. With OTV, the DCI layer has been combined into the aggregation layer. Both data centers use virtual device contexts (VDCs) on the Cisco Nexus 7000 Series switches. In Data Center 1, Cisco Catalyst® 6500 Series switches run as a VSS system with the Cisco ACE modules as a Data Center Service Node and are connected to the Cisco Nexus 7000 Series Switches at the aggregation layer. Data Center 2 has a Cisco ACE appliance connected directly to the Cisco Nexus 7000 Series switches at the aggregation layer. Cisco ACE is used in Layer 3 one-arm mode with source NAT in both data centers.

Both data centers use Cisco Catalyst 6500 Series switches at the data center core layer to connect to the Internet. VSS technology is used between pairs of Cisco Catalyst 6500 Series switches for high availability and to provide Multichassis EtherChannel (MEC) connectivity. Both data centers are built on a Cisco three-tier data center architecture.

Each data center has a VMware ESX server on which Microsoft SQL Server can be run as a virtual machine. Initially, SQL Server runs in Data Center 1. This SQL Server instance is migrated between data centers using VMware VMotion. A SQL client connects to SQL Server over the Internet, using the data center WAN core, to access e-commerce applications (refer to “Test topology and tools” for application details).

The Cisco GSS cluster in the Data Center WAN is used for DNS-based lookup, to access SQL Server. Cisco GSS points to the Cisco ACE device in the data center where SQL Server currently resides.

The storage for the solution is provisioned using EMC VPLEX Metro with a single engine at each site. The VPLEX engines are connected to EMC Symmetrix® VMAX™ storage arrays in each data center. This configuration highlights the capability of the VPLEX clusters to federate heterogeneous physical storage to create distributed volumes across multiple data centers. The VPLEX engines provide a single logical storage view to both data centers. Cisco MDS storage fabric provides SAN connectivity between the VPLEX engines and VMware ESX Server. The communication and synchronization of the data between the VPLEX engines at the two data centers are facilitated through the use of a separate VSAN on the MDS directors. The VSAN is extended between the data centers through the use of an FCIP tunnel that is routed using the OTV links between the data centers.

Test topology and tools

The solution was validated using real-life application servers migrating across data centers while clients accessed the applications. The applications used were Microsoft SQL Server and Microsoft Exchange Server. Table 1 lists the configurations and the test tools used.

Table 1 Test tool configuration

Application	Server Configuration	Stress-Generation Tool	Application Performance Metrics	Description
Microsoft SQL Server 2005 (64-bit)	CPU: 4 virtual CPUs (vCPUs) Memory: 8 GB Storage: EMC VPLEX, EMC VMAX, EMC CLARiiON® OS: Microsoft 2008 64-bit server	Dell DVD Store open source benchmark	Orders per minute (OPMs)	The DVD Store benchmark is an online transaction processing (OLTP) benchmark that simulates the operation of a DVD store. Performance is measured in OPMs, indicating the number of orders successfully inserted into the database per minute.

Test methodology

Microsoft SQL Server test

- Reinitialize Microsoft SQL Server by rebooting the VMware ESX server on which it resides and the target VMware ESX server to reset the statistics data.
- Start the Dell DVD Store client on a virtual machine that has IP connectivity to both VMware ESX servers.
- Run the Dell DVD client and wait for 30 minutes for the client to attain a steady state; note the operations per minute (OPM) on that VMware ESX server.
- Migrate the system to the corresponding target.
- Wait 30 minutes for the client to attain steady state; note the OPMs for that VMware ESX server.
- Perform 18 more migrations with a 10-minute wait between each migration.
- Collect test statistics to evaluate the total elapsed time.

Test results

The goal of the joint testing was to measure the completion time taken for the overall VMware VMotion migration and the impact to application performance in terms of operations per minute (OPMs) due to the migration of the workload between the data centers.

The overall migration time is an important measure and it becomes critical when multiple VMware VMotion migrations are being performed. The duration of a VMware VMotion migration largely depends on the distance between the source and destination VMware ESX servers, the amount of memory configured on the virtual machine, and the amount of bandwidth available between the data centers.

The application used to validate the solution is an e-commerce suite, which was hosted on Microsoft SQL Server 2005. Dell DVD Store Version 2 (DS2) is a complete online e-commerce test application with a back-end database component, a Web application layer, and driver programs. The virtual machine hosting the back-end Microsoft SQL Server database was migrated across the data centers and the performance of the application in OPMs was captured.

The test was performed on a pair of EMC VMAX arrays, one in each data center, providing the raw storage. A pair of EMC VPLEX systems provided distributed volumes for the database across the data centers.

Figure 5 shows application performance, in OPMs, when SQL Server and its back-end database were moved to a remote data center up to 100 kilometers away. The two graphs shown in the figure highlight the performance of SQL Server with VPLEX providing the logical storage and VMAX providing the physical storage.

The figure contrasts the performance of SQL Server with VMAX storage within a single data center with the performance of SQL Server accessing the VMAX data across a physical distance. The OPM numbers dropped by up to 10% when the storage for the database was 100 kilometers from the server. However, the use of the VPLEX/VMAX combination for the storage minimizes the effects of latency between data centers by allowing read I/Os to be serviced at the local data center. This results in a performance drop that is less than 2%. The fact that a typical Online Transaction Processing (OLTP) e-commerce application has a 65%-to-35% read-to-write ratio indicates that it is well-suited to take advantage of the performance benefits offered by VPLEX.

The SQL client performing the benchmark maintained all sessions. A momentary drop in performance was observed during the migration for ~90 seconds before the performance returned to steady-state values during the migration.

To sum up, the use of EMC VPLEX storage ensures that the performance of the application is almost the same at both data centers. VPLEX storage improves the overall performance of the application—in fact the performance within the same data center improved by ~9%. Performance degradation between data centers was less than 2%; thus, application performance across data centers exceeded the single data center performance. In addition, an Active/Active data center proves to be an operational reality. Finally, migration to a remote data center is feasible from not only a technical perspective (application mobility is possible) but also from a business standpoint (application performance is not adversely affected).

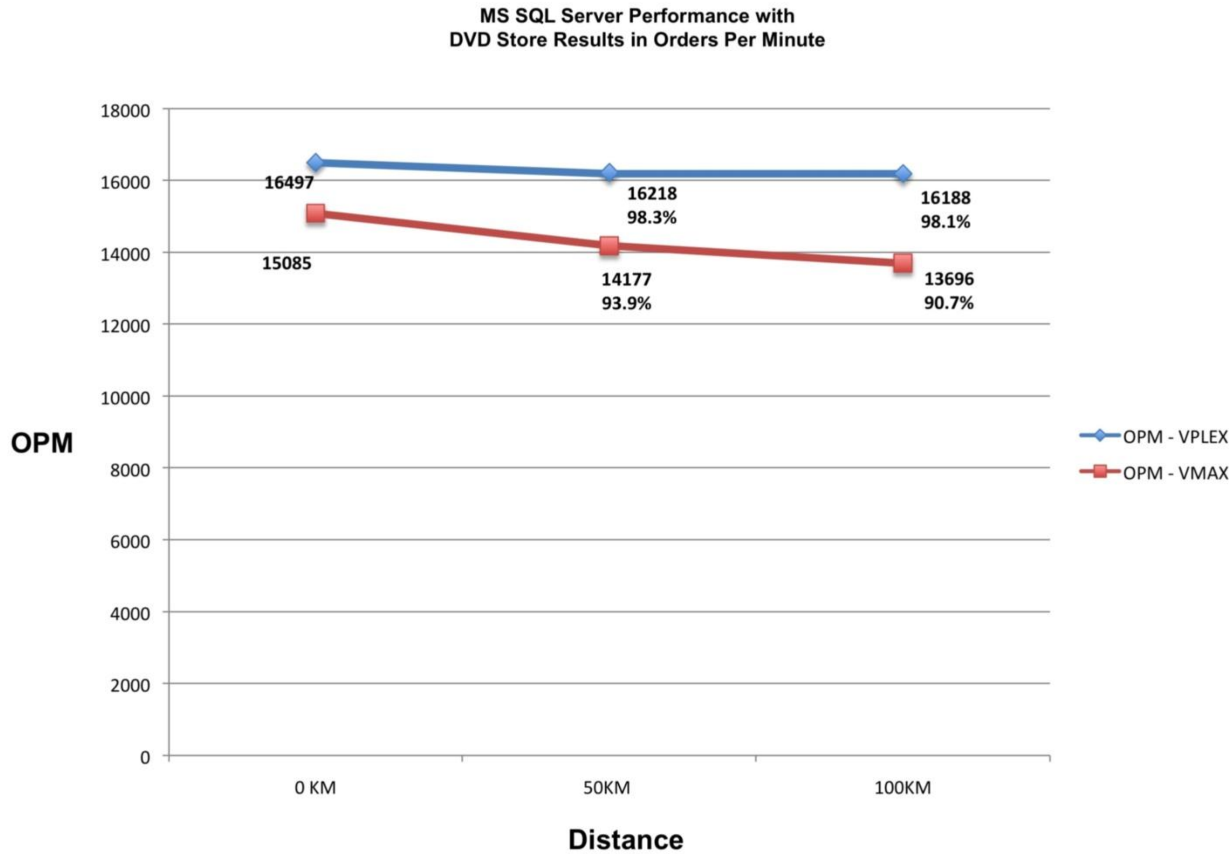

Figure 5 **Microsoft SQL Server and Dell DVD Store performance**

Recommended operational procedure

The recommended procedure for implementing the joint Cisco, EMC, and VMware solution is to set up the VMware vSphere high-availability clusters so that they are independent of each other in the two data centers. VMware VMotion migration across data centers should be a manually initiated task to keep VMware Dynamic Resource Scheduling (DRS) from automatically moving virtual machines across data centers.

Conclusion

Cisco, EMC, and VMware jointly tested and verified an application mobility architecture that allows customers to move virtual machine-based application workloads between data centers without affecting application uptime. The described solution uses innovative technologies from the three companies to solve common problems that customers face when trying to extend Layer 2 domains between data centers.

VMware VMotion and SRM are the building blocks for achieving enhanced business continuity by reducing both planned and unplanned downtime, respectively. Cisco OTV technology provides a powerful mechanism for easily and flexibly extending the LAN across any type of transport network without requiring a network redesign. The integration of VMware vCenter with Cisco GSS and Cisco ACE delivers crucial route optimization functions. The EMC VPLEX Active-Active storage solution offers an elegant and efficient way to deliver storage access with nearly no performance degradation between data centers.

The joint Cisco, EMC, and VMware solution gives IT departments a powerful tool for better provisioning, using, and maintaining a virtualized data center with resources spread across multiple physical locations.

For more information

- VMware VMotion:
  - http://www.vmware.com/products/vi/vc/vmotion.html
  - http://www.vmware.com/products/vmotion/
- VMware SRM: http://www.vmware.com/products/site-recovery-manager/
- Data Center Interconnect (DCI): Layer 2 Extension Between Remote Data Centers:
  - http://www.cisco.com/en/US/prod/collateral/switches/ps5718/ps708/white_paper_c11_493718.html
  - http://www.cisco.com/en/US/netsol/ns975/index.html
- Cisco Catalyst 6500 Series Switches: http://www.cisco.com/go/6500
- Cisco Nexus 7000 Series Switches: http://www.cisco.com/go/nexus7000
- Cisco MDS 9000 Family:
  - http://www.cisco.com/go/mds
  - http://www.cisco.com/en/US/products/hw/ps4159/ps4358/prod_white_papers_list.html
  - http://cisco.com/en/US/products/ps10497/index.html
- Cisco Overlay Transport Virtualization (OTV): http://www.cisco.com/en/US/prod/switches/ps9441/nexus7000_promo.html
- Cisco GSS and ACE products: http://www.cisco.com/go/ace
- Cisco Nexus 1000V Switch:
  - http://www.cisco.com/go/1000veval
  - http://www.cisco.com/en/US/products/ps9902/index.html
- EMC VPLEX Storage
  - http://www.emc.com/products/family/vplex.htm
  - http://www.emc.com/products/detail/hardware/vplex-metro.htm
  - http://www.emc.com/products/detail/hardware/vplex-local.htm

Appendix A

Virtual Computing Environment (VCE) overview

IT is undergoing a transformation. The current “accidental architecture” in the data center increases procurement, management costs, and complexity while making it difficult to meet customer SLAs. This makes it difficult for IT to respond to the needs of the business in a timely manner and creates the perception that IT is a cost center. The data center is now moving toward a “private cloud” model, which is a new model for delivering IT as a service, whether that service is provided internally (IT today), externally (service provider), or in combination. This new model requires a new way of thinking about both the underlying technology and the delivery model for customer success.

While the need for a new model has never been clearer, navigating the path to that model has never been more complicated. The benefits of private clouds are capturing the collective imagination of businesses and organizations of all sizes around the world. The realities of outdated technologies, rampant incremental approaches, and the absence of a compelling end-state architecture are impeding adoption by customers.

By harnessing the power of virtualization, private clouds place considerable business benefits within reach, including:

- **Business enablement**—Increased business agility and responsiveness to changing priorities; speed of deployment and the ability to address the scale of global operations with business innovation.
- **Service-based business models**—Ability to operate IT as a service.
- **Facilities optimization**—Lower energy usage and better (less) use of data center real estate.
- **IT budget savings**—Efficient use of resources through consolidation and simplification.
- **Reduction in complexity**—Moving away from fragmented, “accidental architectures” to integrated, optimized technology that lowers risk, increases speed, and produces predictable outcomes.
- **Flexibility**—Ability of IT to gain responsiveness and scalability through federation to cloud service providers while maintaining enterprise-required policy and control. The VCE coalition provides a range of infrastructure platforms for deploying virtualized and non-virtualized applications. Collectively, these platforms are known as Vblock Infrastructure Packages.

The VCE coalition, formed jointly by Cisco, EMC, and VMware, represents an unprecedented level of collaboration in development, services, and partner enablement that reduces risk for companies looking to transform their infrastructures for the private cloud. Enabled by Vblock Infrastructure Packages, the coalition delivers the industry’s first completely integrated IT offering that combines best-of-breed virtualization, networking, computing, storage, security, and management technologies with end-to-end vendor accountability. Rather than assembling individual components, customers can now purchase them as Vblock Infrastructure Packages. The three companies have invested in an industry-leading seamless support experience, featuring innovative collaboration that starts with a single point of contact for customer support.

Cisco is the industry leader in networking technologies. The Cisco switching and routing portfolio enables a robust foundation for intelligent network connectivity within and across data centers. Cisco offers a range of data center products and technologies to address the need for cutting-edge data center designs.

EMC Corporation, an industry leader in networked storage, develops, delivers, and supports information infrastructure and virtual infrastructure technologies and solutions. The company’s “Information Storage” segment offers networked information storage systems and software, which are deployed in storage area network (SAN), networked-attached storage (NAS), unified storage combining NAS and SAN, content-addressed storage, and direct-attached storage environments.

VMware has been the industry leader in virtualization technologies for the past decade and has brought to the data center critical technologies such as VMware vSphere, VMotion for virtual machine mobility, and Site Recovery Manager (SRM) for disaster recovery.

Cisco, EMC, and VMware joint solution details

VMware VMotion technology

VMware VMotion enables the live migration of running virtual machines from one physical server to another with zero downtime, continuous service availability, and complete transaction integrity.

VMware VMotion migration is achieved when the active memory and precise execution state of a virtual machine is rapidly transmitted over a high-speed network from one physical server to another and access to the virtual machine's disk storage is instantly switched to the new physical host. Since the network is also virtualized by VMware vSphere, the virtual machine retains its network identity and connections, helping ensure a transparent migration process. VMware VMotion is a crucial enabling technology for the creation of a single highly available virtual data center spanning geographically disparate data centers.

Cisco Overlay Transport Virtualization LAN extension technology

Cisco OTV technology provides an operationally optimized solution for the extension of Layer 2 connectivity across any transport. OTV is therefore critical to the effective deployment of distributed data centers to support application availability and flexible workload mobility. OTV is a "MAC in IP" technique. By using the principles of MAC address routing, OTV provides an overlay that enables Layer 2 connectivity between separate Layer 2 domains while keeping these domains independent and preserving the fault-isolation, resiliency, and load-balancing benefits of an IP-based interconnection.

OTV uses a control protocol to map MAC address destinations to IP next hops that are reachable through a routed network core. OTV can be thought of as MAC address routing, in which the destination is a MAC address, the next hop is an IP address, and traffic is encapsulated in IP so it can simply be carried to its MAC address routing next hop over the core IP network. Thus, a flow between source and destination host MAC addresses is translated in the overlay into an IP flow between the source and destination IP addresses of the relevant OTV edge devices. This process is referred to as encapsulation rather than tunneling because the encapsulation is imposed dynamically and tunnels are not maintained. Since traffic is IP forwarded, OTV is as efficient as the core IP network and will deliver optimal traffic load balancing, multicast traffic replication, and fast failover just like the core would. Figure 6 illustrates this dynamic encapsulation mechanism.

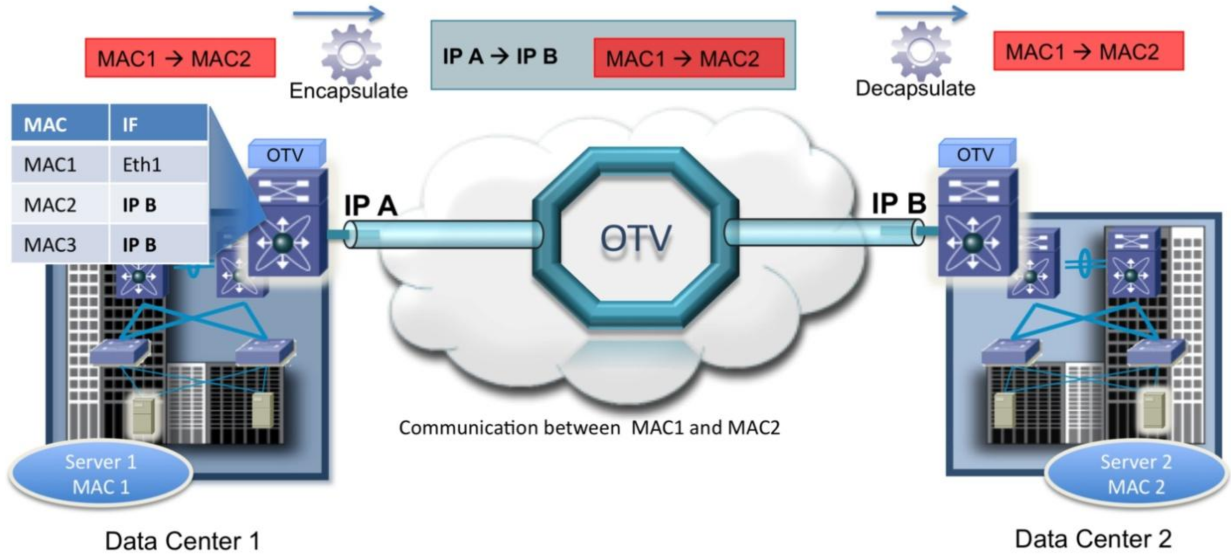

Figure 6 Cisco OTV operations

OTV provides the following benefits:

- **Transport agnostic**: OTV is IP encapsulated and can therefore use any core capable of forwarding IP traffic. OTV, therefore, does not pose any requirements for the core transport.

- **High availability**: OTV preserves the failure boundary and site independence: OTV does not rely on traffic flooding to propagate reachability information for MAC addresses. Instead, a control protocol is used to distribute such information. Thus, flooding of unknown traffic is suppressed on the OTV overlay, Address Resolution Protocol (ARP) traffic is forwarded only in a controlled manner, and broadcasts can be forwarded based on specific policies. Spanning tree Bridge Protocol Data Units (BPDUs) are not forwarded at all on the overlay. The result is failure containment comparable to that achieved using a Layer 3 boundary at the Layer 2 domain edge. Sites remain independent of each other and failures do not propagate beyond the OTV edge device. The loop prevention mechanisms in OTV prevent loops from forming on the overlay and also prevent loops from being induced by sites when these are multihomed to the overlay.

- **Full WAN bandwidth utilization and optimal multicast replication**: When sites are multihomed, OTV provides the capability to actively use multiple paths over multiple edge devices. This capability is crucial to keeping all edge devices active and thus optimizes the use of available bandwidth. OTV uses the IP-multicast capabilities of the core to provide optimal multicast traffic replication to multiple sites and avoid head-end replication that leads to suboptimal bandwidth utilization.

- **Transparent to the sites**: OTV extensions do not affect the design or protocols of the Layer 2 sites they interconnect. Interconnection is as transparent as connection of a router to the Layer 2 domain and therefore does not affect the local spanning tree or topology.

Distributed virtual switching

VMware VMotion moves virtual machines running applications across physical ports and data centers; however, the network provisioning, security, and management policies as well as visibility at the virtual machine level must be available. With current hardware switches, it is impossible to view or apply network policy to locally switched traffic between virtual machines at the virtual network interface card (vNIC) level. For example, it is not possible to correlate traffic on the same physical link from multiple virtual machines.

VMware and Cisco jointly developed the concept of a distributed virtual switch (DVS), which essentially decouples the control and data planes of the embedded switch and allows multiple, independent virtual switches (data planes), called Virtual Ethernet Modules (VEMs), to be managed by a centralized management system (control plane), called a virtual switch module (VSM). VEM enables advanced switching capability on the hypervisor and provides each VM with dedicated switch ports. Figure 7 shows a Cisco Nexus 1000V in a VMware vSphere environment.

The Cisco Nexus 1000V virtual switch (vSwitch) has the following strengths:

- It is the industry's first DVS for VMware.
- It overcomes network challenges and accelerates server virtualization.
- It is compatible with all switching platforms. The Cisco Nexus 1000V maintains the existing VMware vCenter provisioning model for server administration while allowing network administration of the virtual network using the Cisco NX-OS Software command-line interface (CLI).
- It allows server teams to offload vSwitch responsibility to the network teams, helping ensure proper network connectivity and security. Network teams get virtual machine-level visibility, NetFlow, Encapsulated Remote Switched Port Analyzer (ERSPAN), and port statistics that continue through VMware VMotion migration.

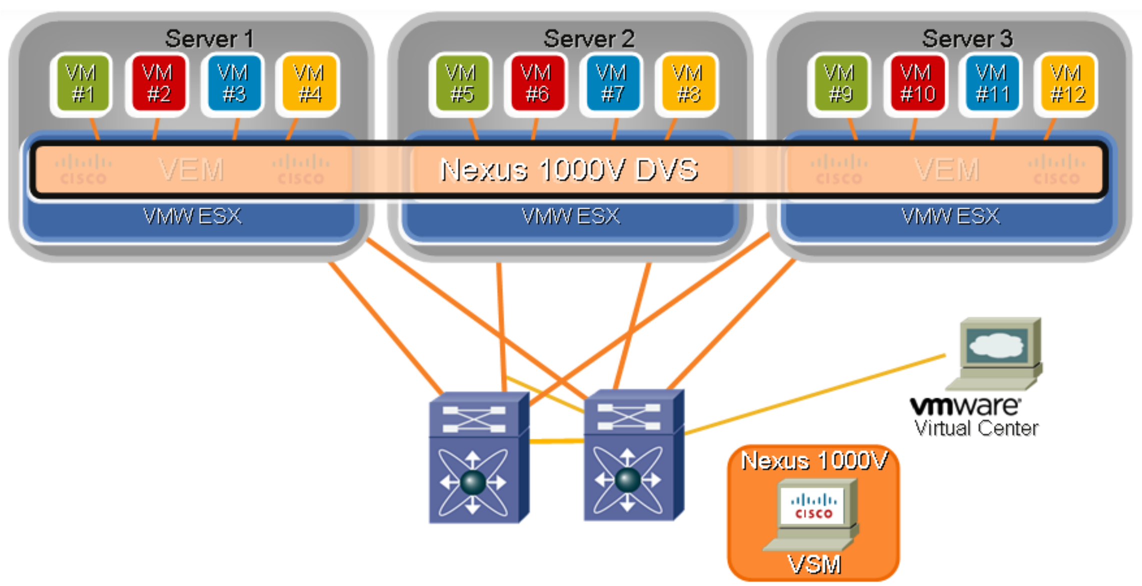

Figure 7 Cisco Nexus 1000 as a distributed virtual switch

DNS-based routing optimization

Routing optimally to a virtual machine migrated by VMware VMotion is critical because the virtual machine maintains its existing IP and MAC addresses as it is moved across data centers. The Layer 3 and Layer 2 reachability design of the data center should accommodate this behavior.

If the traffic to the virtual machine originates in the same Layer 2 domain, the Layer 2 extensions will suffice for connectivity across pods or data centers. OTV facilitates this design. In the example in Figure 4 earlier in this document, in Data Center 1, Layer 2 traffic requiring reachability from the Cisco 6500 Catalyst Series pod to the Cisco Nexus 7000 Series pod is switched using Layer 2. No additional configuration is needed.

If the traffic to the virtual machine is traversing a Layer 3 network, such as an IP cloud or the Internet, the traffic needs to be rerouted to the new data center location. Existing application sessions may continue to be routed through the existing data center due to specific or existing IP service requirements, such as firewalls.

The following IP routing considerations are required:

- Routing from remote clients to application servers: This consideration can be addressed using the Cisco Global Site Selector (GSS) and the Cisco Application Control Engine (ACE). Cisco GSS acts as the authoritative DNS entity to guide clients coming from the internet or a cloud to the application server in the correct data center, where the virtual machine resides. Cisco GSS knows about the location of the virtual machines (and the SQL server running on it) because of the workflow integration of Cisco GSS and ACE and VMware vCenter.

- Routing from application servers to remote clients: Application traffic must be forwarded to the appropriate default gateway preferably in the local data center pod to achieve optimal routing as well as symmetrical routing for IP services such as firewalls. The Hot Standby Router Protocol (HSRP) default gateway pointing from Microsoft SQL Server to the aggregation layer switch can be the local switch. This design is achieved by having an identical HSRP default gateway address in both data centers. HSRP localization techniques can be used to filter HSRP keep-alives across the Layer 2 VLAN extension between data centers. The HSRP default gateway forwards traffic to the appropriate Cisco ACE device for processing of source Network Address Translation (NAT) or firewall services. This capability can be achieved by deploying the Cisco ACE in Layer 3 one-arm mode with source NAT. This design maintains symmetrical routing from an IP services perspective.

As shown in Figure 8, the process works as follows in a stepwise manner:

1. A remote client, client1 (across the Internet or cloud), wants to connect to the application in the data center. Client1 will perform a DNS-based lookup for the application server (Microsoft SQL Server “sql-server.jsmp.cisco.com”) on the Cisco GSS in the IP WAN. The Cisco GSS will respond to the DNS lookup with the virtual IP address of the Cisco ACE in the data center containing the SQL server (ACE1 for data center 1).

2. The client will send and receive traffic for this session to/from the IP address of ACE1. ACE1 will do a NAT translation and send the traffic to the aggregation switch. The aggregation switch is the Layer3-Layer2 boundary in the data center and will forward the traffic to the SQL server. The SQL server will send the return traffic to the local HSRP gateway, which in turn will send the return traffic to ACE1.The return traffic exits the data center at the same point it entered. This ensures state maintenance with a firewall or address-translation service.

3. The SQL server is migrated (using VMotion) from data center 1 to data center 2 using vCenter. vCenter notifies GSS that the SQL server VM has moved and makes configuration changes on GSS, so that the GSS points to the updated location of the VM. The design can also allow changes on the ACE via vCenter and ACE can notify the GSS that it does not have the VM locally. In either case, changes on GSS or ACE will only affect new client sessions, which will do a DNS lookup.

4. Existing client1 sessions will use the original data center (data center 1) and traverse the DCI cloud using OTV to reach the SQL server. The return traffic is forwarded to the local HSRP gateway in data center 2, which forwards the return traffic to ACE1 in data center 1. ACE 1 will resolve the NAT translation and forward return traffic to the client.

5. A new client, client2, wishing to connect to the SQL server will do a DNS lookup with the Cisco GSS in the IP WAN. The Cisco GSS will respond to the DNS lookup with the virtual IP address of the Cisco ACE

in the data center containing Microsoft SQL Server (ACE2 for data center 2). Client2 sessions will use data center 2 to reach the SQL server. The client will send and receive traffic for this session to/from the IP address of ACE2. ACE2 will do a NAT translation and send the traffic to the aggregation switch in data center 2. The aggregation switch is the Layer3-Layer2 boundary in the data center and will forward the traffic to the SQL server. The SQL server will send the return traffic to the local HSRP gateway, which in turn will send the return traffic to ACE2. The return traffic exits the data center at the same point it entered. This ensures state maintenance with a firewall or address-translation service.

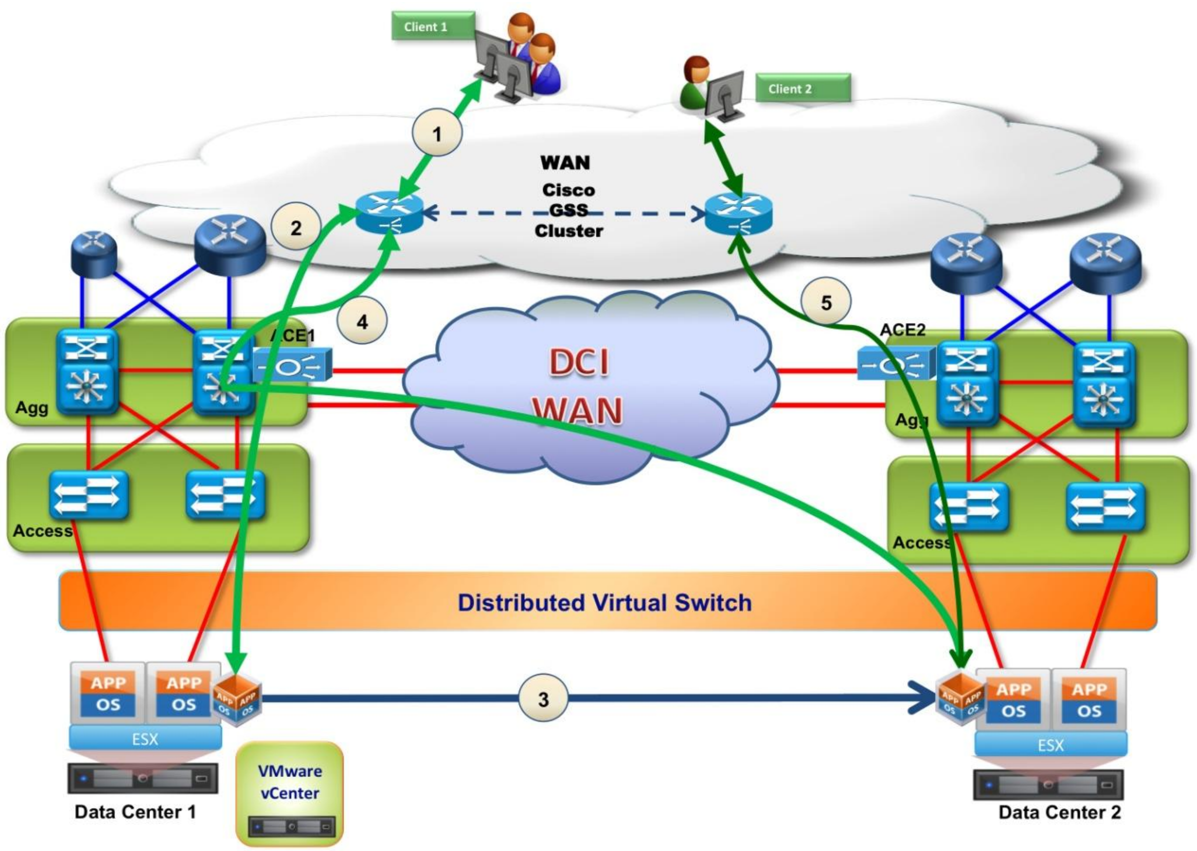

Figure 8 DNS-based routing optimization

Cisco facilitates simplified provisioning of application delivery services by integrating Cisco ACE with VMware vCenter through the implementation of a VMware vCenter plug-in that securely communicates with Cisco Application Networking Manager (ANM) 3.1. From within VMware vCenter, using the functions integrated by the plug-in, the user can:

- Deploy virtual machines as real servers into an existing server farm.
- Monitor application traffic flow for virtual machines through the Cisco ACE.
- Securely activate and suspend application traffic flows through the Cisco ACE for the associated real servers.

This single-pane provisioning, application traffic monitoring, and operations management streamlines the deployment of services and the maintenance operations for applications and virtual machines. Organizations do not need to undertake a separate integration or management application development project to gain these functions; they are all part of the Cisco ANM 3.1 offering.

EMC VPLEX for VMotion with globally unique namespace storage

EMC VPLEX with the EMC GeoSynchrony operating system breaks physical barriers of data centers and allows users to access data for read and write operations at different geographical locations concurrently. This is achieved by synchronously replicating data between the data centers while depending on the hosts accessing the storage devices to manage the consistency through the use of intelligent distributed lock management. This capability in a VMware context enables functionality that was not available earlier. Specifically, the ability to concurrently access the same set of devices independent of the physical location enables geographical VMotion based on the VMware virtualization platform. This allows for transparent load sharing between multiple sites while providing the flexibility of migrating workloads between sites in anticipation of planned events, such as hardware maintenance. Furthermore, in case of an unplanned event that causes disruption of services at one of the data centers, the failed services can be quickly and easily restarted at the surviving site with minimal effort.

Figure 9 schematically shows the configuration of EMC VPLEX Metro that enables live migration of VMware virtual machines between two sites separated by distance. It can be seen from the figure that each site has a VPLEX cluster with access to physical storage. The cluster at each site communicates with each other through Fibre Channel protocol. The FC extension between the VPLEX clusters can be done with either dark fiber extending between the VPLEX clusters or with an FC over IP (FCIP) tunnel on the IP WAN between data centers.

Figure 9 also shows that the federation capability of VPLEX Metro allows the creation of a distributed volume that has the same SCSI identification independent of the location from which the device is accessed. Therefore, the two VMware ESX hosts shown in the figure consider the distributed volume as the same device and enable capabilities such as VMware VMotion that were traditionally available only in a single data center.

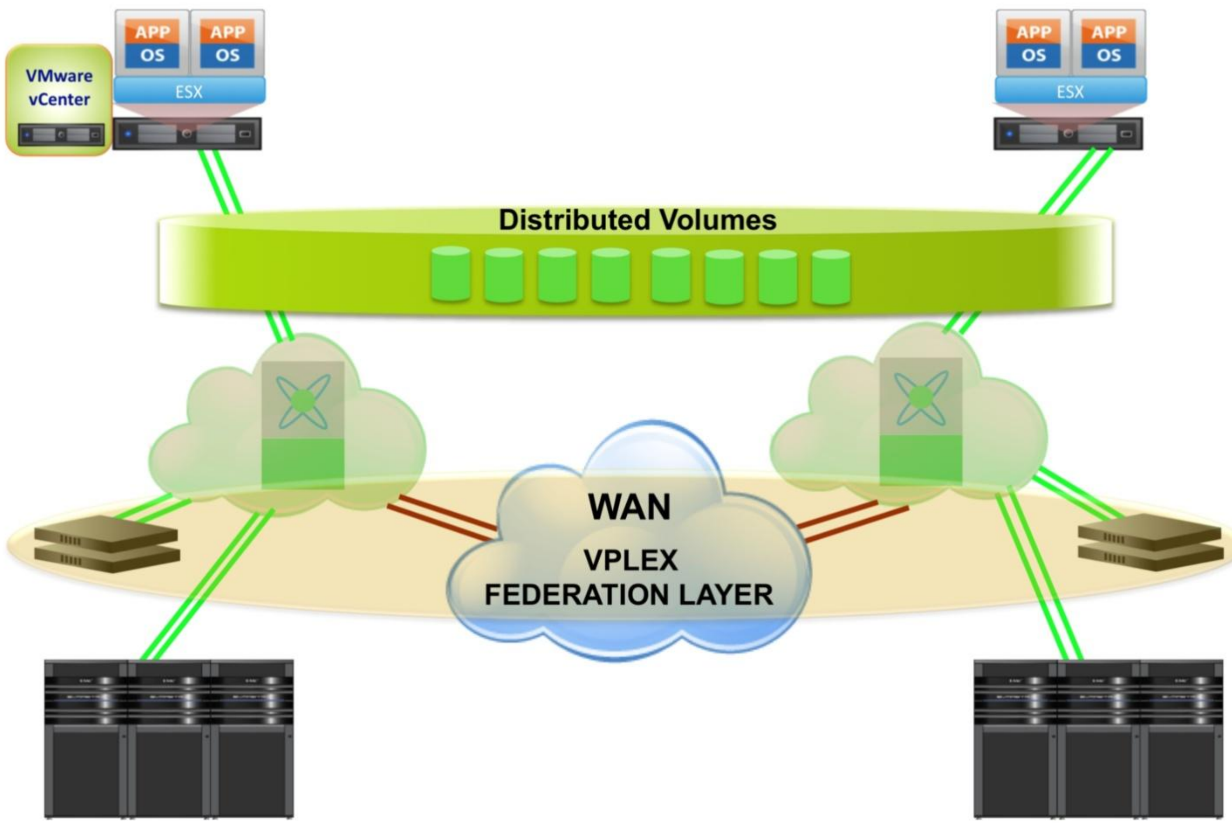

Figure 9 EMC VPLEX Metro for enabling VMware VMotion of applications across geography

Cisco Storage Area Networking (SAN) technology

The availability, scalability, security, and performance of the storage subsystem are of utmost importance to any enterprise. The task of ensuring that all these factors are addressed in a single data center is a daunting task for any storage administrator. Managing all of these factors across data centers is an even greater challenge, requiring implementation of storage best practices. These factors directly affect application performance, in turn affecting the SLAs of business-critical applications. The Cisco MDS 9000 Family of SAN switches is especially suited to these SAN topologies. Table 2 summarizes the features that can be used to address the requirements for storage across data centers.

Table 2 Cisco SAN extension solutions

Feature	Requirements	Functions
Virtual SAN	Isolation and security	The VSAN technology provides secure hardware-based network segmentation, similar to the VLAN technology that is widely deployed in LANs. Fabric services such as zoning and routing are independent per VSAN. In this validated solution, the nodes in each VMware ESX cluster are placed in a dedicated VSAN, to use a consolidated physical infrastructure and to be isolated with respect to security threat and fabric-wide errors.
	Management and access control	Cisco MDS 9000 NX-OS software management offers several levels of role-based access control (RBAC). This feature allows an administrator to be in charge of a specific VSAN without having any visibility into other VSANs. The administrator can map the roles defined in the VMware vCenter; for instance, an administrator may be able to access a specific VSAN and the corresponding VMware ESX cluster and nothing else.
Inter-VSAN Routing (IVR)	Isolation and security	In a DCI solution, each data center can implement independent VSANs, preserving the fabric services segmentation, data isolation, and administration independence. IVR allows selected devices from different VSANs, even across different data centers, to communicate without any fabric merging. In this validated solution IVR provides connectivity between the VMware ESX servers located in the secondary data center and the storage located in the primary data center (shared storage). IVR can also provide connectivity to execute VMware Storage VMotion across data centers and to perform primary-array-to-secondary-array storage replication.
SAN extension with dark fiber	Integrated solution	The capability to plug long-wave and Coarse Wavelength Division Multiplexing (CWDM) optics into the Cisco MDS 9000 Series Switches simplifies SAN extension over dark fiber. The performances are guaranteed by the extended buffer-to-buffer credits available with the Cisco MDS 9000 Series.
	Security	Cisco MDS 9000 Series switches provide Cisco TrustSec Fibre Channel Link Encryption to secure SAN extension data across native Fibre Channel links.

SAN extension with FCIP	**Integrated solution**	**Cisco MDS 9000 Series switches provide Gigabit Ethernet interfaces and support the FCIP protocol, to transparently extend the SAN over an IP network.**
	Security	The Cisco MDS 9000 Series provides native IP Security (IPsec) encryption to secure FCIP links.
Port channeling	Availability	Cisco MDS 9000 Series PortChannels are the aggregation of multiple physical Fibre Channel or FCIP links into one logical link, to provide higher aggregated bandwidth, load balancing, and link redundancy.
I/O acceleration (IOA)	Application performances	IOA is an intelligent distributed fabric service built into Cisco MDS 9000 Series Switches. IOA accelerates I/O performance across distances. This feature helps the overall application performance remain relatively the same, even when the application server and the storage are separated by considerable distance. In this validated solution, I/O performance has been enhanced over the FCIP link.

Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134 USA

Tel: 408-526-4000 or 800-553-6387 (NETS)
Fax: 408-527-0883

www.cisco.com

EMC Corporation
176 South Street
Hopkinton, MA 01748 USA

Tel: 508-435-1000

www.emc.com

VMware, Inc.
3401 Hillview Ave
Palo Alto, CA 94304 USA

Tel: 650-427-5000 or 877-486-9273
Fax: 650-427-5001

www.vmware.com