Cisco Virtual Wide Area Application Services: Technical Overview

What You Will Learn

Organizations are offering private and virtual private cloud-based application delivery over the WAN to their end users as they seek to reduce the cost of application delivery while increasing deployment agility. Organizations want to migrate their business applications to the cloud while reducing the negative effect on performance of cloud-based application delivery to end users.

IT managers are dealing with increasing pressure to reduce application delivery costs while helping ensure application availability. As a result, they are increasingly adopting virtualization and cloud computing. However, they face several challenges in migrating applications to the cloud, with application performance being often cited in industry surveys as one of the main concerns. The presence of a WAN or Internet link between the cloud and the end user affects business application performance. A solution to the application performace challenge is to make WAN optimization a part of the cloud infrastructure.

In this document, you will learn how the Cisco® Virtual Wide Area Application Services (vWAAS) solution solves application-delivery challenges and mitigates poor application performance over private and virtual private cloud-based networks. In addition, Cisco vWAAS enables cloud service providers to offer optimized application performance over the WAN as a value-added service in their catalogs of cloud services. It also helps their internal or external customers migrate their business applications to the cloud.

Challenges with Cloud Adoption

Cloud-based models promise to change the consumption model for technology, allowing enterprises to increase their business agility and save costs through on-demand provisioning and tear-down of infrastructure and services while being charged through a utility-based model. Public clouds were adopted first, and now private clouds and virtual private clouds are being built. Private clouds typically are deployed within an enterprise's premises; virtual private clouds are built by cloud service providers to offer enterprises a way to extend their resources through a use-based model.

The adoption of enterprisewide private cloud offerings has been cautious. A recent IDC industry study cites application performance as one of the top three inhibitors of cloud adoption, in addition to security and availability.

Main Requirements for a WAN Optimization Service in the Cloud

To offer WAN optimization as a service in a cloud environment, several requirements must be met:

- On-demand deployment model: Associate the WAN optimization service with specific application server virtual machines, giving the cloud provider the agility and capability to offer a differentiated optimized-application service in contrast to basic computing service.

- Minimal network configuration with increasing scale: As new server virtual machines are instantiated to accommodate increasing application load, the network configuration needed to optimize these new virtual machine instances should be minimal; ideally, no work should be required at all.

- Virtual machine-mobility awareness: Given the dynamic nature of cloud resource pools, the WAN optimization service provided to a set of application servers needs to be continually available regardless of the physical location of the application server virtual machines.

- Support for multi-tenant deployments: A multi-tenant deployment model must be supported, to reduce the amount of hardware required and to lower the fixed hardware costs for the cloud service provider.
- Elastic scale-out deployment: The WAN optimization service will need to scale in different ways simultaneously for different tenants as demand grows or diminishes. This process requires an elastic scale-out deployment architecture.

Any solution that attempts to meet these requirements should be virtualized. Virtual appliances can be deployed on demand, unlike hardware appliances, which need to be racked, stacked, and cabled. Virtualization reduces hardware costs in environments serving multiple organizations because a dedicated device is not needed for each tenant. As user demand increases, IT managers can easily scale up performance of virtual appliances by moving them to a more powerful platform or by allocating more resources on existing platforms and using the scalable nature of virtual computing environments. However, virtual form-factor alone is not sufficient to meet all requirements.

Cisco Virtual WAAS: The First Cloud-Ready WAN Optimization Solution

Cisco Virtual WAAS (vWAAS) is the industry's first cloud-ready WAN optimization solution. Cisco vWAAS is a virtual appliance that accelerates business applications delivered from private and virtual private cloud infrastructures, helping ensure an optimal user experience. Cisco vWAAS runs on the VMware ESXi hypervisor and Cisco Unified Computing System™ (UCS) x86 servers, providing an agile, elastic, and multi-tenant deployment (Figure 1).

Cisco vWAAS is the only WAN optimization solution that is deployed in an application-specific, virtualization-aware, and on-demand manner. Using policy-based configuration in the Cisco Nexus® 1000V Switch, Cisco vWAAS service is associated with application server virtual machines as they are instantiated or moved. This approach helps enable cloud providers to offer rapid delivery of WAN optimization services with little network configuration or disruption in cloud-based environments.

Cisco vWAAS is designed for both enterprises and service providers who plan to offer private and virtual private cloud-based application delivery services over the WAN to their internal or external customers. Cisco vWAAS can be deployed in the physical data center and in private clouds and in virtual private clouds offered by service providers.

Cisco vWAAS enables migration of business applications to the cloud, reducing the negative effect on performance of cloud-based application delivery to end-users. It enables service providers to offer an excellent application experience over the WAN as a value-added service in their catalogs of cloud services.

Figure 1. Cisco vWAAS Deployment Architecture

Cisco WAAS Central Manager, a common Cisco WAAS management tool, can be used to manage physical Cisco WAAS Appliances, Cisco vWAAS instances, Cisco Intergrated Services Routers (ISRs) integrated with Cisco Services-Ready Engine (SRE) Modules, and Cisco WAAS Express in the network, providing the same level of high availability and scalability across the system. In addition, Cisco vWAAS Central Manager vCM) can be deployed in a virtual appliance form factor, providing central management for both physical and virtual Cisco WAAS instances.

Cisco vWAAS is deployed as a virtual machine based on VMware vSphere 4.0 infrastructure, providing integration with VMware vCenter management for Cisco vWAAS virtual machine orchestration and provisioning.

Cisco vWAAS provides the same level of comprehensive WAN optimization and application acceleration as Cisco WAAS. For more information about the Cisco WAAS solution, please see the Cisco WAAS solution technical overview.

Traffic to Cisco vWAAS cab be redirected for optimization using various interception options. Cisco vWAAS supports the innovative new Cisco Nexus 1000V vPATH interception technology deployed deep in the data center along with server virtual machines for cloud-ready WAN optimization. In additional, Cisco vWAAS supports traditional out-of-path interception technologies such as Web Cache Control Protocol (WCCP) and Policy-Based Routing (PBR). Both interception options are available in data center and private clouds and in virtual private clouds. Regardless of the interception technique used, multiple Cisco vWAAS virtual machines can coexist together with other server virtual machines in the same VMware ESX/ESXi host in both branch offices and data center and private clouds.

Cisco vWAAS Deployment Options

Cisco vWAAS supports two deployment options:

1. **Traditional WAN edge deployment with out-of-path Interception:** This deployment is applicable in branch offices and data center and private clouds. The WAN edge is the typical place in the network for interception of the WAN-facing traffic to be optimized. Out-of-path interception using technologies such as WCCP is the preferred method for intercepting traffic at this layer. Cisco vWAAS provides full support for WCCP Version 2 (WCCPv2), allowing up to 32 Cisco vWAAS devices to be deployed in a single device group with load balancing, failover, and nondisruptive Cisco vWAAS virtual machine insertion and removal (Figure 2).

Figure 2.　　Cisco vWAAS Traffic Interception
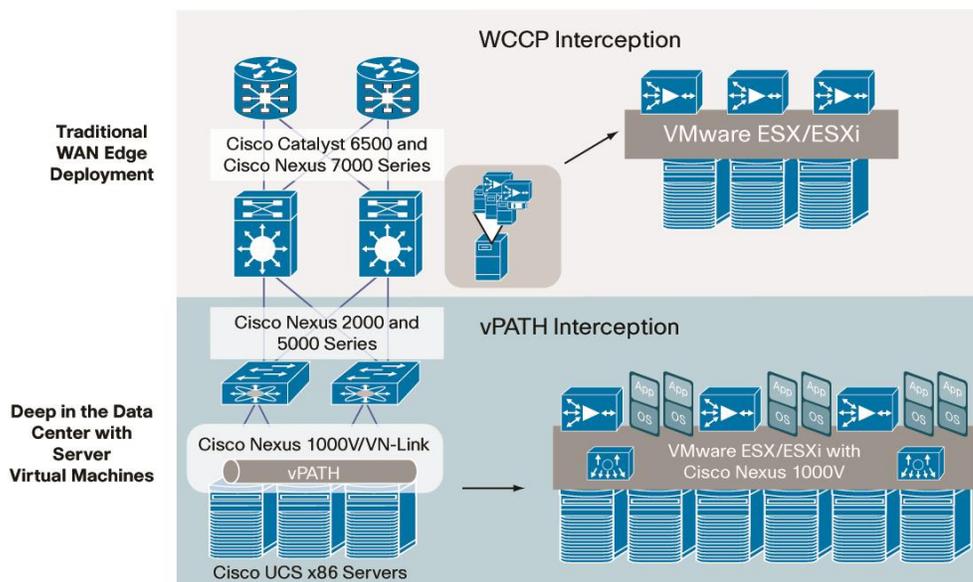
In this deployment model, VMware ESXi hosts carrying Cisco vWAAS virtual machines are connected to WCCP-enabled switches or routers near the WAN edge as shown in Figure 2. Multiple Cisco vWAAS virtual machines can co-exist together in a single WCCP cluster, optimizing all traffic intercepted by the WCCP router. These virtual machines can be spread across single or multiple physical servers. Physical and virtual appliances can be mixed in the same WCCP cluster. All standard WCCP configurations and best practices are applicable to Cisco vWAAS.

2. **Deep in the data center with innovative vPATH interception (new):** This deployment is applicable in private clouds and virtual private clouds and helps meet several requirements to help enable WAN optimization as a cloud-ready service, as explained earlier. In this deployment model, Cisco vWAAS virtual machines can be placed next to server virtual machines in the same VMware ESXi host. Using policy-based orchestration of virtual network services provided by the Cisco Nexus 1000V architecture, Cisco vWAAS can be deployed on a per-application or per-server-group basis. vPATH interception in the Cisco Nexus 1000V provides a mechanism for intercepting all traffic to and from these servers and forwarding it to the Cisco vWAAS virtual machine for optimization.

Cisco vWAAS Supports Cisco Nexus 1000V Virtual Network Services

Network services in private clouds and virtual private clouds should coexist and associate with server virtual machines to provide agility, elasticity, and multi-tenancy. The Cisco Unified Network Services framework allows virtual network services (like Cisco vWAAS) to coexist with application services. This architecture provides virtualized network services with minimal service interruption and high availability even as application server nodes dynamically move from one location to another.

This approach provides investment protection for customers planning next-generation data center architectures, enabling them to build private and virtual private cloud offerings intended for a rapid service-creation IT delivery model for internal or external customers. Only Cisco offers a clear vision for the ongoing integration of WAN optimization into the network infrastructure as customers migrate to the cloud.

Cisco Nexus vPATH Interception

The Cisco Nexus 1000V provides virtualization-aware network services to all application server virtual machines. Central to this capability are port profiles, which are a collection of interface configuration commands that can be dynamically applied at either physical or virtual interfaces. Any changes to a given port profile are propagated immediately to all ports that have been associated with it. Port profiles are visible as VMware port groups in the VMware vCenter management console. For more information about the Cisco Nexus 1000V, see http://www.cisco.com/go/nexus1000v.

The Cisco Nexus 1000V provides a mechanism for attaching Cisco vWAAS to the port profiles of servers that need to be optimized. All traffic to and from these servers will be intercepted by vPATH and forwarded to the Cisco vWAAS virtual machine for optimization. vPATH interception uses Cisco Nexus 1000V port-profile attributes (vn-service) to redirect traffic to Cisco vWAAS. Administrators needs to identify the port profiles of servers to be optimized by Cisco vWAAS. After the port profile is identified, Cisco vWAAS needs to attach to one or multiple port profiles to optimize the traffic. Cisco WAAS autodiscovery helps ensure that a particular TCP connection will be optimized only by the end-point devices (Cisco Wide Area Application Engine [WAE] or Cisco vWAAS).

Cisco vPATH interception, based on the Cisco Nexus 1000V port profiles, provides the following advantages:

- Network attributes move together with virtual machines: This feature provides virtual machine mobility awareness to WAN optimization traffic. Even if server virtual machines move from one VMware ESXi host to another, traffic to and from these servers will continue to be optimized by the same Cisco vWAAS.

- A port profile provides inheritance of network attributes to by any new server virtual machines created with that port profile: Cisco vWAAS starts optimizing traffic immediately for these server virtual machines. This

approach provides an on-demand deployment model with minimal network configuration and with elastic scalability.

- Separate port profile can be configured for each tenant in a multi-tenant environment: These port profiles can be attached to separate Cisco vWAAS instances to provide full multi-tenancy.

In Figure 3, virtual machines with port profiles shown in green are the servers to be optimized. All traffic to and from these server virtual machines will be intercepted by the Cisco Nexus 1000V vPATH module and redirected to Cisco vWAAS for optimization. Server virtual machines with red port profiles are unoptimized virtual machines; vPATH will not intercept traffic to and from these virtual machines.

Figure 3. Cisco vPATH Interception
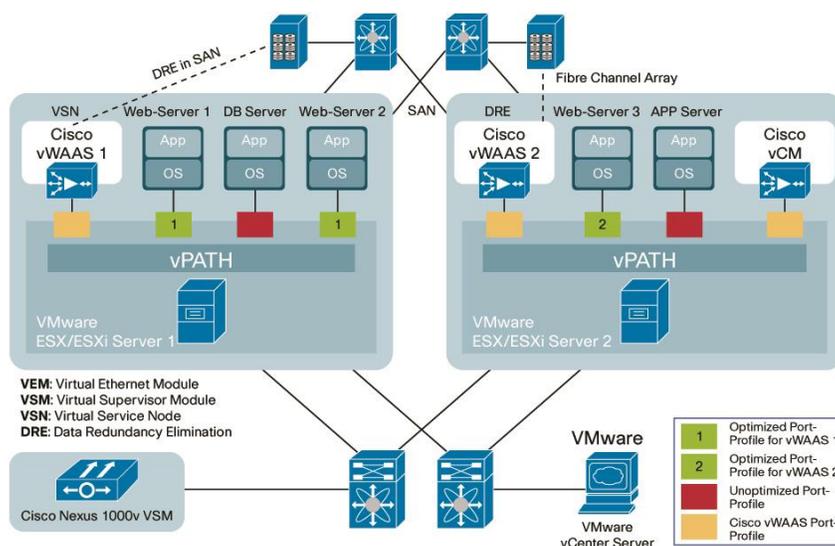
Multiple server port profiles with virtual machines in different physical hosts can be redirected to one Cisco vWAAS instance. In Figure 3, Web-Server 3 can be optimized by Cisco vWAAS 1 or 2 based on the configuration of the port profiles. vPATH helps ensure nondisruptive operation in case of VMware vMotion processing or virtual machine mobility. If Web-Server 2 moves from VMware ESX Server 1 to Server 2, it will continue to be optimized by Cisco vWAAS 1 on VMware ESX Server 1 without any configuration change required.

Cisco vWAAS SAN-Based Disk Storage (New)

Cisco WAAS Appliances and Cisco ISR network modules have hard-disk storage space to store the Cisco WAAS data redundancy elimination (DRE) byte cache and Common Internet File System (CIFS) cache. Cisco vWAAS supports the following options for disk storage:

- Direct-attached storage (DAS), as with physical appliances
- SAN with Fibre Channel (both Small Computer System Interface [SCSI] and SCSI over IP [iSCSI])

With SAN storage, Cisco vWAAS can use the shared storage infrastructure that VMware ESX and ESXi require to enable advanced features such as VMware High Availability, vMotion, and Storage vMotion. For more information, please see the VMware SAN configuration guide.

Customers are encouraged to use SAN-based storage (both SCSI and iSCSI) with Cisco vWAAS; however, SAN-based storage and DAS provide similar levels of performance for WAN optimization.

Main Features and Benefits

Cisco vWAAS is the industry's first cloud-ready WAN optimization solution to deliver the main requirements for a cloud, using policy-based operations based on the Cisco Nexus 1000V architecture:

- On-demand orchestration: Cisco vWAAS provides the industry's first on-demand orchestrated insertion of WAN optimization in the cloud, using policy-based operations. Using policy-based configuration in the Cisco Nexus 1000V, Cisco vWAAS is associated with application server virtual machines even as they are instantiated or moved.

- High availability: If one Cisco vWAAS virtual machine fails, another Cisco vWAAS virtual machine will become active on the same or a different host using VMware High Availability (HA). Cisco vWAAS also supports VMware Distributed Resources Scheduler (DRS) for optimizing and managing resources.

- Fault-tolerant persistent performance: Cisco vWAAS SAN-based storage helps ensure cache preservation and high persisitent performance in the event of failure. If a Cisco vWAAS virtual machine fails, a new virtual machine will become active using VMware HA. The new virtual machine will use the same cache storage as the failed virtual machine used, providing the compression benefits of WAN optimization without interruption.

Conclusion

Cloud-based application delivery to the branch office has been inhibited by poor performance. WAN optimization is a strategic technology that addresses this problem. Cloud service providers can differentiate their services by including a cloud-ready WAN optimization service and mitigate the eventual commoditization of computing, networking, and storage as a service.

Requirements for a cloud-ready WAN optimization solution include on-demand orchestration in response to the creation or movement of application server virtual machines, little need for network configuration in a dynamic environment, elastic scale-out, and multi-tenancy support. A virtual form-factor for WAN optimization is necessary but not sufficient.

Cisco vWAAS is the industry's first cloud-ready WAN optimization solution to meet these requirements, using policy-based operations based on Cisco Nexus 1000V Series architecture. Cisco vWAAS is fully interoperable with existing Cisco WAAS appliances and router-integrated modules, which can all be managed by a common central manager, available in virtual form as well. Cisco vWAAS offers advanced features such as SAN storage for its DRE cache, allowing faster recovery after failure through separation of computing and storage resources. Cisco vWAAS provides a path for private and virtual private cloud architects to deliver an excellent application experience over the WAN as a value-added service in their catalogs of cloud services.

For More Information

- Cisco WAAS: http://www.cisco.com/go/waas
- Cisco Unified Network Services: http://www.cisco.com/go/uns