VMware vCloud® Architecture Toolkit™
for Service Providers

VMware vSAN™ Two-Node Architecture VMware Cloud Provider™ Use Cases

Version 2.9

January 2018

Martin Hosken

VMware, Inc.
3401 Hillview Ave
Palo Alto, CA 94304
www.vmware.com

Contents

List of Figures

Introduction

Many VMware Cloud Providers™ deliver business compute and storage platforms which extend across on-premises and provider data centers to their customers. However, these hybrid solutions typically require high levels of CapEx and OpEx, and are therefore out of reach for many small businesses and startups.

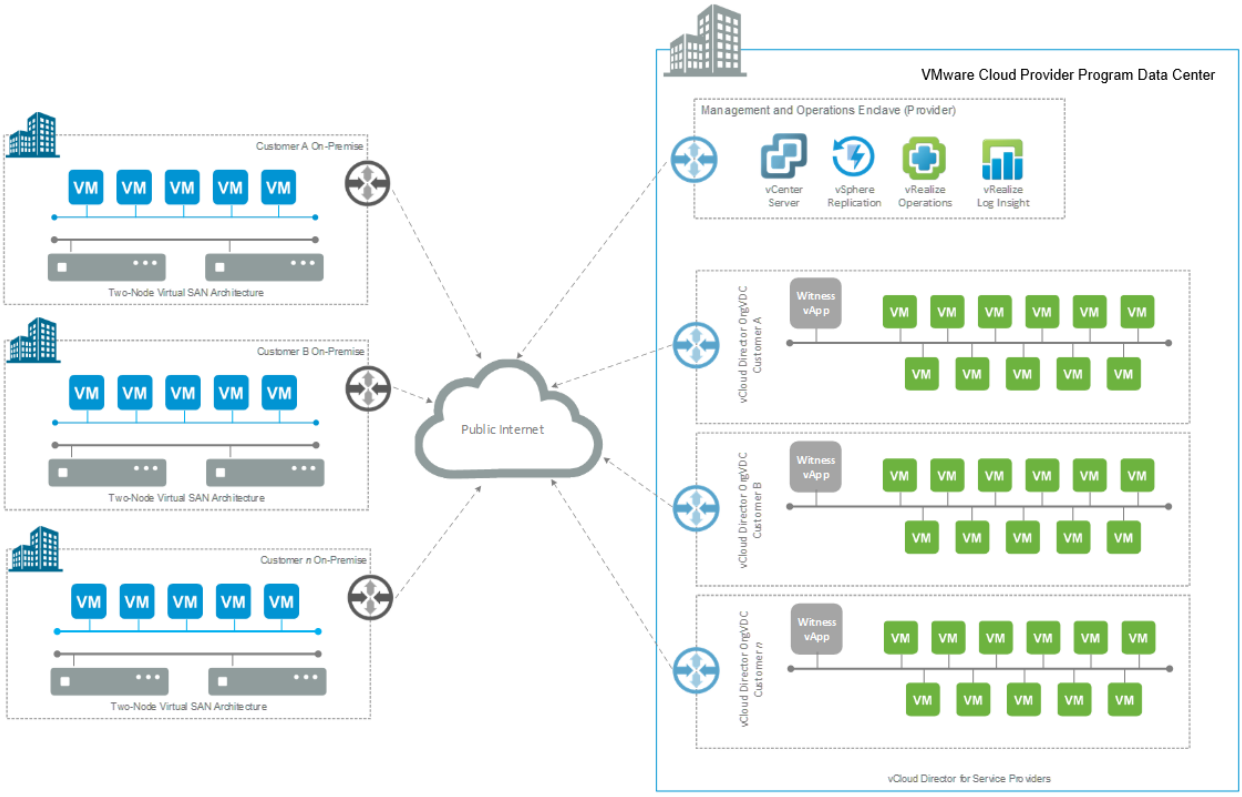
Deploying VMware vSAN™ in a two-node configuration on small sites, using low-cost standard server components or vSAN ready nodes, enables small businesses or startups with limited capital to avoid the significant up-front costs associated with storage hardware. In addition, by automating the management of virtual machines through Storage-Policy Based Management (SPBM), operational costs can also be lowered, with no need for end customers to provide administrative assistance during setup or operational management of the solution. All administration can easily be performed centrally by the service provider as part of their managed service offering, or delegated to local end-customer administrative staff through the use of VMware vCenter Server® role-based access control (RBAC) and the user-friendly VMware vSphere® Web Client.

This solution is in contrast to the use of local storage at similar on-premises deployments, and the limitations associated with such configurations. The use of local disk VMFS datastores, without vSAN, introduces multiple risks to uptime of local applications, and provides no redundancy for failure. For instance, with this local disk architecture, a single copy of the virtual machine's files is stored. If that disk fails, the virtual machine will fail, and the only option might be to restore from backups, which is likely to be time consuming and may result in the loss of business data, depending on the Recovery Point Objective (RPO) set out in the application's Service Level Agreement (SLA). While other availability options might exist for local storage solutions such as these, they are typically not automated, are unlikely to minimize data loss, and will likely not provide a recovery time of less than 90 seconds, which is a common RPO and Recovery Time Objective (RTO) offered by the VMware vSphere High Availability (HA) mechanism.

This article demonstrates how VMware Cloud Providers can provide low-cost hybrid solutions to small businesses and startups by benefitting from the simple two-node on-premises architecture offered through vSAN, which is typically referred to by VMware and others as the Remote Office/Branch Office (ROBO) solution.

The primary aim of this architecture is to offer small businesses and startup customers a simple, on-premises compute and shared storage solution, while minimizing up-front deployment costs and long-term operational overhead.

Figure 1. Solution Overview



Benefits to IT users within the end customer's organization include providing local, on-premises services, such as Active Directory authentication, file and other large data services, and optimized and low-cost recovery of the data center resources provided by the VMware Cloud Provider. However, in the past, server resources on site at the end customer's location could prove to be cost prohibitive, due to the requirement for shared storage to facilitate vSphere features such as vSphere HA.

This article includes several use cases where a vSAN two-node architecture offers benefits to the service provider's portfolio:

- Extended Data Center Hosted Services
- On-Premises / Off-Premises Stretched Applications
- Small and Medium Enterprise (SME) Public Cloud Broker Deployments

Before addressing each of the use cases, this article provides an overview of the vSAN two-node architecture and its implementation considerations.

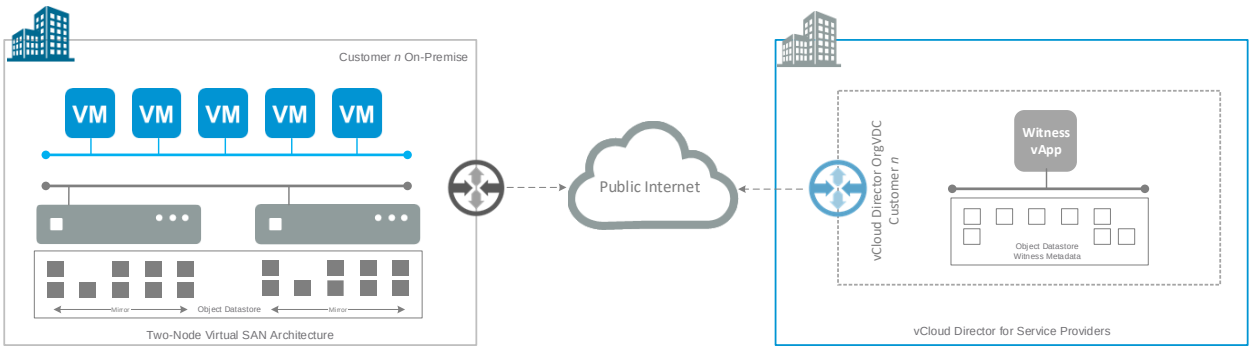
vSAN Two-Node Architecture Overview

As of vSAN 6.1, a new type of two-node solution has been supported, typically referred to as Remote Office/Branch Office (ROBO) environments. This solution architecture allows small office implementations to benefit from shared storage, while also minimizing cost. Prior to this, and without this specific architecture, three-node clusters remain the minimum supported configuration for vSAN enabled environments.

The two-node vSAN architecture builds on the concept of Fault Domains, first introduced in vSAN 6.0. Each of the two VMware ESXi™ hosts, located on the tenant's premises, represent a single Fault Domain. In vSAN architecture, the objects that make up a virtual machine are typically stored in a redundant mirror across two Fault Domains, assuming the Number of Failures to Tolerate is equal to 1. As a result of this, in a scenario where one of the hosts goes offline, the virtual machines can continue to run, or be restarted, on the alternate node. To achieve this, a Witness is required to act as a tie-breaker, to achieve a quorum, and enable the surviving nodes in the cluster to restart the affected virtual machines.

However, unlike a traditional vSAN enabled cluster, where the witness objects are local to the configured cluster hosts, in a two-node architecture, the witness objects are located externally at a second site on a dedicated virtual appliance specifically configured to store metadata, and to provide the required quorum services for a host failure. In the use cases that follow, this Witness Appliance is located in the VMware Cloud Provider's data center.

Figure 2. Witness Object Metadata



By employing a dedicated virtual appliance to provide Witness services, this architecture eliminates the need to deploy a third vSphere host on the end customer's site, which reduces over all costs without sacrificing the availability benefits of shared storage. The Witness Appliance is actually a specially modified nested ESXi host, specifically designed to only store witness objects and cluster metadata. The Witness Appliance does not contribute to the compute and storage capacity of the solution, and cannot be employed to host virtual machines. The use of a Witness Appliance in a vSAN enabled configuration is only supported by VMware for this type of two-node architecture, and with vSAN Stretched Cluster designs.

In the following use cases, the Witness Appliance is deployed as an OVA into the service provider's data center. As most deployments of this kind only host a small number of virtual machines at the end customer's site, the "Tiny" configuration of two vCPUs with 8 GB of assigned memory is typically more than sufficient, allowing support for up to 750 components.

The nested ESXi vSAN Witness Appliance is automatically deployed with both flash and mechanical disks embedded, where one of the appliance's VMDKs is tagged as a flash device during provisioning. No manual configuration is required by the service provider's vSAN administrator. In addition, there is no requirement for a physical flash device in the vSphere host that is hosting the Witness. All of the appliance's virtual disks can be thin provisioned, if required.

To store the required metadata, the Witness requires 16 MB of storage capacity for each Witness component stored, with one Witness component per object. Object types include:

- Virtual disk objects
- Snapshot delta VMDK objects
- Virtual machine namespace object
- Virtual machine swap object

With this in mind, the number of components stored on the Witness Appliance directly reflects the number of objects associated with the virtual machines running on the on-premises hosts. For instance, with each virtual machine requiring at least one virtual disk (VMDK), one namespace, and one swap file, the result is a minimum of three objects per virtual machine, with each snapshot also adding one object per VMDK.

Because the Witness Appliance does not host virtual machines, and therefore does not have to service virtual machine read and write requests, the network connectivity requirements between the end customer site and the service provider's data center are minimal. Typically, a WAN interconnect with 1.5 Mb/s of available bandwidth and latency as high as 500 ms Round-Trip Time (RTT) is sufficient to provide network communication between the two-node cluster at the customer's offices and the Witness Appliance located at the service provider's data center. However, like a traditional vSAN deployment, multicast must be enabled for communication between hosts in the two-node on-premises cluster, although there is no requirement for multicast to be enabled for WAN communication with the Witness Appliance.

The vSAN Storage-Policy Based Management (SPBM) employs a capability referred to as the Number of Failures to Tolerate (FTT), which is pertinent to this architecture. This capability provides the RAID 1 standard mirrored configuration that provides the *n*+1 redundancy for a virtual machine. In a vSAN two-node architecture, the FTT=1 policy is explicitly required, because there are exactly three configured Fault Domains. With this configuration applied, a mirrored copy of each virtual machine is created and automatically maintained on separate physical nodes. It is this mechanism which allows one host within this two-node architecture to fail, with users either retaining full access to the application continuously, or the workload being restarted by the vSphere HA process. The continuous availability of the application depends on whether or not it was hosted by the affected node, or configured at the application layer for high availability, and deployed and load balanced across both nodes in the cluster. Another option that provides application availability across the two-node cluster, is VMware vSphere Fault Tolerance, which is also compatible with this vSAN architecture, and can provide continuous availability to workloads with up to four virtual CPUs, in the event of a host failure.

As previously outlined, vSphere HA is a critical component of this architecture and these use cases, because shared storage is tightly integrated with the process of restarting virtual machines after a host outage. With vSphere HA enabled, if a host fails, the virtual machines that are impacted by the outage reboot on other hosts in the cluster, minimizing downtime. However, in a two-node architecture to make sure that enough CPU and memory resources are available to restart all impacted virtual machines—effectively running 100 percent of the workload on a single host—the vSphere HA Admission Control Policy must be configured to reserve 50 percent of both memory and CPU, irrespective of the amount of available storage. As a result, in this configuration only 50 percent of the compute resources in the two-node architecture is available to run workloads.

Service Provider Sample Use Cases

This two-node vSAN architecture provides VMware Cloud Providers with a range of service offering opportunities that can be added to their portfolio. The following section of this article provides three examples of use cases, and illustrates the high-level architecture of each. These solutions are targeted at single office and small-size or medium-size enterprises, working in collaboration with one VMware Cloud Provider's data center and delivering witness and other services, such as disaster recovery and continuous availability to its customers.

3.1 Use Case 1– Extended Data Center Hosted Services

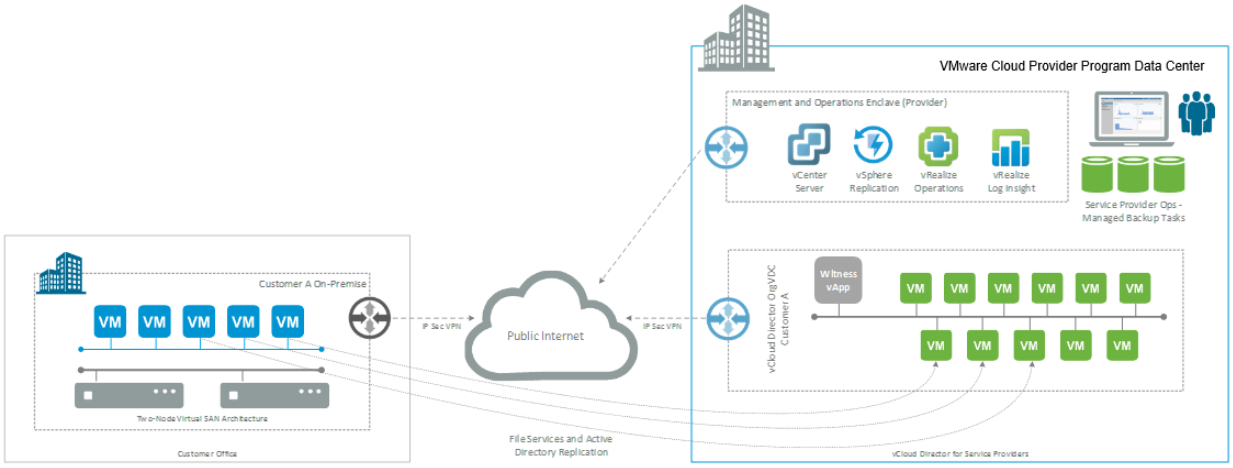
In this first sample use case, a vSAN two-node architecture is employed to provide a common small business solution where Microsoft Active Directory, File Services, other applications and site management tools are provided on-premises. In this example, the service provider furnishes a fully managed service offering. Microsoft Active Directory and File Services are provided locally to the end users to facilitate quick authentication and fast access to large media files from within their small office, which is only enabled with a limited 10Mb/s connection to the Internet.

All File Server data is being continuously replicated back to the service provider's data center through Microsoft Windows built-in Distributed File System Replication (DFSR) service. In addition, domain controllers at both locations are synchronized continuously through their service's internal replication mechanism.

Data backups are performed only at the service provider's data center. This solution also provides end users with the ability to work at their office, even in the event of a communication failure between their office location and the service provider's data center. Another service might also be added that provides the ability to allow remote access to the data center's replica data for remote and home users.

The following figure illustrates what the high-level architecture might look like for a service offering such as that outlined in this first use case. The figure shows a small office with highly available access to data, service provider backups, and local access to large media files.

Figure 3. Use Case 1 – Extended Data Center Hosted Services



3.2 Use Case 2 – On-Premises / Off-Premises Stretched Applications

In this second use case, a vSAN two-node architecture is employed to provide redundancy, continuous availability, and disaster avoidance to Microsoft Tier 1 applications distributed across the on-premises environment and into the service provider's data center.

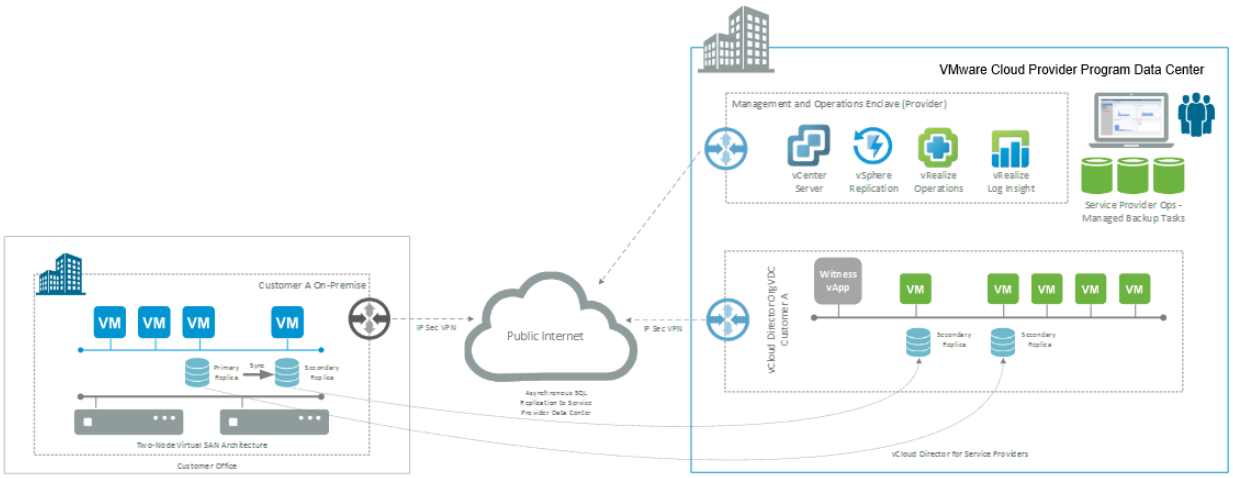
Applications such as Microsoft Exchange and Microsoft SQL Server provide redundancy and availability at the application layer. For instance, Microsoft SQL Server 2016 provides AlwaysOn technology that can be leveraged for high availability, disaster recovery, and fast application failover for maximum availability and data protection of business-critical database systems.

The AlwaysOn Availability Group feature was developed for applications that require high uptime, protection against failures within a single site, and adequate redundancy against failures across locations. This solution architecture helps protect application databases from planned and unplanned downtime, without the need for physical Raw Device Mapping (RDM) shared storage between nodes.

With this use case, the end customer has local access to the business SQL databases to provide responsive results to queries across multiple users within their office. In addition, this managed service offering provides a backup solution for the customer's data from the replica databases located in the service provider's facility, offloading the burden of these tasks, as well as providing a secure, offsite location for its business critical data.

The following figure illustrates the high-level architecture this use case. Supported are stretched Tier 1 applications with local access to databases, and managed offsite backups.

Figure 4. Use Case 2 – On-Premises / Off-Premises Stretched Applications



For more information about configuring cross-site solutions with Microsoft Exchange and SQL, see the following *VMware vCloud Architecture Toolkit for Service Providers* (vCAT-SP) publications:

- Hybrid Database Strategy with Microsoft SQL Server:

  http://www.vmware.com/files/pdf/vcloud-air/Hybrid-Database-Strategy-with-Microsoft-SQL-Server.pdf

- Hybrid Messaging Strategy with Microsoft Exchange 2013:

  http://www.vmware.com/files/pdf/vcloud-air/Hybrid-Messaging-Strategy-with-Microsoft-Exchange-2013.pdf

3.3 Use Case 3 – Small Medium Enterprise (SME) Public Cloud Broker Deployments

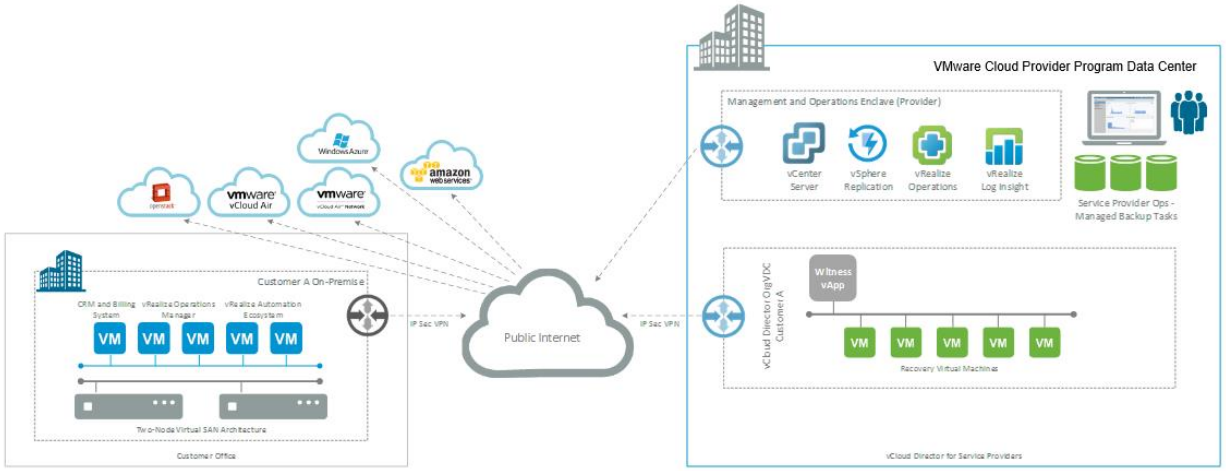
A cloud broker is a third-party individual or business that provides an intermediary service between the purchaser of a cloud computing service and the sellers of that service, often with the broker providing additional cloud management services as part of, or alongside, their broker business offerings.

In this use case, the end consumer wants to employ VMware vRealize® Automation™ as a cloud service broker, allowing the customer to host workloads across a wide range of cloud providers, such as Amazon Web Services, VMware Cloud Providers, Microsoft Azure, and OpenStack offerings from various providers.

In this use case, the cloud provisioning and management components are hosted on-premises, along with VMware vRealize Operations Manager™, for the operational monitoring of the various public cloud platforms and several internal applications that the consumer wants to keep local to their office. All locally hosted virtual machines are replicated to the VMware Cloud Provider's data center, through VMware vSphere Replication™, to provide a redundant solution, and to provide the managed service provider with the data to perform daily backups. Add-on services, such as automated recovery through VMware Site Recovery Manager™, might also be included.

The following diagram illustrates the high-level architecture for this final use case.

Figure 5. Use Case 3 – Small Medium Enterprise (SME) Public Cloud Broker Deployments



For more information about licensing, see the flexible VMware Cloud Provider Program licensing model announcement at https://www.vmware.com/content/dam/digitalmarketing/vmware/en/pdf/partners/vspp/vmw-vcloud-air-network-program-guide.pdf.

See also VMware Cloud Provider Program details at http://www.vmware.com/partners/service-provider.

Summary

The two-node vSAN architecture is a cost-effective platform for running a variety of virtual machine workloads in an on-premises environment in conjunction with a VMware Cloud Provider. This solution allows for business critical applications and services to benefit from shared storage, without the cost and complexity of dedicated storage devices being deployed in small offices.

In addition, a two-node deployment enables a simple scale-up or scale-out strategy at the end customer's site without incurring application downtime, and maintenance can be scheduled by evacuating a single host during a soft maintenance window. Features such as vSphere HA and vSphere Replication can also be enabled to facilitate quick recovery from unplanned downtime.

vSAN health monitoring can be implemented by the provider through the use of the vRealize Operations Management Pack for Storage Devices (MPSD). This management pack enables a number of vSAN focused dashboards to provide deep insight into the vSAN enabled clusters running across multiple locations. In addition, this pack provides proactive alerting and health monitoring to streamline operational management for the service provider and to simplify troubleshooting activities.