vmware® PRESS

Essential Virtual SAN

Administrator's Guide to VMware VSAN

Cormac Hogan
Duncan Epping

Essential Virtual SAN

Administrator's Guide to VMware® Virtual SAN

VMware Press is the official publisher of VMware books and training materials, which provide guidance on the critical topics facing today's technology professionals and students. Enterprises, as well as small- and medium-sized organizations, adopt virtualization as a more agile way of scaling IT to meet business needs. VMware Press provides proven, technically accurate information that will help them meet their goals for customizing, building, and maintaining their virtual environment.

With books, certification and study guides, video training, and learning tools produced by world-class architects and IT experts, VMware Press helps IT professionals master a diverse range of topics on virtualization and cloud computing and is the official source of reference materials for preparing for the VMware Certified Professional Examination.

VMware Press is also pleased to have localization partners that can publish its products into more than forty-two languages, including, but not limited to, Chinese (Simplified), Chinese (Traditional), French, German, Greek, Hindi, Japanese, Korean, Polish, Russian, and Spanish.

For more information about VMware Press, please visit
http://www.vmwarepress.com.

This page intentionally left blank

Essential Virtual SAN

Administrator's Guide to VMware® Virtual SAN

Cormac Hogan
Duncan Epping

vmware® PRESS

Upper Saddle River, NJ • Boston • Indianapolis • San Francisco
New York • Toronto • Montreal • London • Munich • Paris • Madrid
Capetown • Sydney • Tokyo • Singapore • Mexico City

Essential Virtual SAN

Published by Pearson Education, Inc.

Publishing as VMware Press

ISBN-10: 0-13-385499-X

ISBN-13: 978-0-13-385499-2

Library of Congress Control Number: 2014942087

Printed in the United States of America

First Printing: August 2014

Warning and Disclaimer

Every effort has been made to make this book as complete and as accurate as possible, but no warranty or fitness is implied. The information provided is on an "as is" basis. The authors, VMware Press, VMware, and the publisher shall have neither liability nor responsibility to any person or entity with respect to any loss or damages arising from the information contained in this book or from the use of any digital content or programs accompanying it.

The opinions expressed in this book belong to the author and are not necessarily those of VMware.

Special Sales

For information about buying this title in bulk quantities, or for special sales opportunities (which may include electronic versions; custom cover designs; and content particular to your business, training goals, marketing focus, or branding interests), please contact our corporate sales department at corpsales@pearsoned.com or (800) 382-3419.

For government sales inquiries, please contact governmentsales@pearsoned.com.

For questions about sales outside the U.S., please contact international@pearsoned.com.

VMWARE PRESS PROGRAM MANAGER
David Nelson

ASSOCIATE PUBLISHER
David Dusthimer

ACQUISITIONS EDITOR
Joan Murray

TECHNICAL EDITORS
Christos Karamanolis
Paudie O'Riordan

SENIOR DEVELOPMENT EDITOR
Christopher Cleveland

MANAGING EDITOR
Sandra Schroeder

PROJECT EDITOR
Mandie Frank

COPY EDITOR
Keith Cline

PROOFREADER
Paula Lowell

INDEXER
Lisa Stumpf

EDITORIAL ASSISTANT
Vanessa Evans

DESIGNER
Chuti Prasertsith

COMPOSITOR
Bumpy Design

We would like to dedicate this book to the VMware VSAN engineering team.
Without their help and countless hours discussing the ins and outs of
Virtual SAN, this book would not have been possible.
—Cormac & Duncan

This page intentionally left blank

Contents

This page intentionally left blank

Foreword by Ben Fathi

When I arrived at VMware in early 2012, I was given the charter to deliver the next generation of vSphere, our flagship product. It was a humbling experience, but exhilarating at the same time. A few months into the role, I welcomed our storage group to the team, and I had the honor of working closely with a dedicated team of engineers. I saw that they were building something very unique—what I believe will be a significant turning point in the history of storage.

We set out to build a distributed fault tolerant storage system optimized for virtual environments. Our goal was to build a product that had all the qualities of shared storage (resilience, performance, scalability, and so on) but running on standard x86 servers with no specialized hardware or software to maintain. Just plug in disks and SSDs, and vSphere takes care of the rest. Add to that a policy-based management framework and you have a new operational model, one that drastically simplifies storage management.

There were problems aplenty, as is usual in all long-term software projects: long nights, low morale, competing priorities, and shifting schedules. Through it all, the team persevered. We hit a particularly painful point in June 2013. We were getting ready to ship vSphere 5.5, and I had to be the one to tell the team that they weren't ready to ship VSAN. Instead, they would have to go through a broad public beta and much more rigorous testing before we could call the product "customer-ready."

The stakes were far too high, particularly because this was VMware's first foray into software-defined storage and a key part of our software-defined data center vision.

They were disappointed, of course, not to be showcased up on stage at VMworld, but still they persevered. I think we took the right course. Six months and 12,000 beta testers later, Virtual SAN was ready: It's robust, proven, and ready for action. VSAN can scale from a modest three-node configuration in a branch office to a multi-petabyte, mega-IOPS monster capable of handling all enterprise storage needs.

The team has delivered something truly unique in the industry: a fully distributed storage architecture that's seamlessly integrated into the hypervisor.

VSAN isn't bolted on, it's built in.

Much has already been written about VSAN and how it brings something completely new to the storage world. This book, however, is different. Duncan and Cormac have worked closely with the development team all throughout the project. Not only are they intimately familiar with the VSAN architecture, but they've also deployed and managed it at scale. You're in good hands.

Ben Fathi
CTO, VMware

This page intentionally left blank

Foreword by Charles Fan

Earlier this year, I had the pleasure of sitting in a session by Clayton Christensen. His seminal work, *Innovator's Dilemma*, was one of my favorite business readings, and it was an awesome experience to hear Clayton in person. For the whole session, I had this surreal feeling that there were no other people in the room, just Clayton and me, and we were discussing Virtual SAN (VSAN).

The topic of the discussion? Is VSAN a disruptive innovation or a sustaining innovation?

As they were defined in Clayton's book, sustaining innovations are the technological advances that make things better, faster, more powerful to answer to the increasing demand from customers. Sustaining innovations do not require any change in the business model, business process, or target customers. This is how big companies become bigger. Given their resources and customer relationships, they almost always win against smaller companies when it comes to sustaining innovation.

However, there will be times that the technology advances outpace the growth of customer demand. At this time, the innovation comes from the bottom. Those innovations will offer a different way of getting things done, which may not deliver the same level of feature and performance initially, but they are cheaper, simpler, and often introduce the technology to more and different customers and sometimes completely change the business model. This is the disruptive innovation. It is extremely difficult for incumbent leaders to deal with disruptive innovations, and this type of innovation redefines industries. And new leaders are born.

So, is VSAN a disruptive innovation or a sustaining innovation? It might seem like a dumb question. Of course, it is a disruptive innovation. It is a radically simple, software-only, hypervisor-converged distributed storage solution fully integrated with vSphere, running on commodity hardware. It redefines both the economics and the consumption models of storage. Although it lacks (so far) a list of classic storage features and goodies, it is offering orders of magnitude more simplicity than classic enterprise storage arrays, sold to a different set of users from the storage admins, at a lower cost. Thus it is a classic disruptive innovation, similar to the point-and-shoot cameras. Compared to "real" cameras, point-and-shoot cameras had fewer features initially, but they were also radically simpler and targeted a different set of users. Guess what? Quickly there were more point-and-shoots than the real ones.

Then why the question? As a storage product, yes, VSAN is without a doubt a revolutionary product that will disrupt the entire storage industry and usher in a new era. However, if we change our perspective, and look at it as the natural extension of vSphere Server virtualization platform to the software-defined data center, it is a sustaining innovation.

VSAN is being sold to the same vSphere customers and empowers them to do more. It extends server environments into converged infrastructure. It extends the vSphere abstractions and policy-based automation from compute to storage.

So, we have a rare winner on our hands, a combination of being a sustaining innovation on top of our hypervisor platform that is natural for VMware to extend the value we offer to our customers, and at the same time a disruptive innovation that will reshape the storage industry. In other words, it is a product that will do to storage what vSphere did to servers.

The VSAN product is the result of 4 years of hard work from the entire VSAN product team. This team is more than just the core architects, developers, testers, and product managers. Duncan and Cormac are two critical members of the team who brought real-world experiences and customer empathy into the program, and they have also been two of our strongest voices back out to the world. I am really glad that they are writing this timely book, and hope you will find it as useful as I did. VSAN is a unique product that will have a lasting impact on the industry, and I welcome you to join us in this exciting journey.

Charles Fan
SVP of VMware R&D, Storage, and Availability

About the Authors

Cormac Hogan is a storage architect in the Integration Engineering team at VMware. Cormac was one of the first VMware employees at the EMEA headquarters in Cork, Ireland, back in 2005, and has previously held roles in VMware's Technical Marketing and Support organizations. Cormac has written a number of storage-related white papers and has given numerous presentations on storage best practices and new features. Cormac is the owner of CormacHogan.com, a blog site dedicated to storage and virtualization.

He can be followed on twitter **@CormacJHogan**.

Duncan Epping is a principal architect working for VMware R&D. Duncan is responsible for exploring new possibilities with existing products and features, researching new business opportunities for VMware. Duncan specializes in software-defined storage, hyper-converged platforms, and availability solutions. Duncan was among the first VMware Certified Design Experts (VCDX 007). Duncan is the owner of Yellow-Bricks.com and author of various books, including the VMware vSphere Clustering Technical Deepdive series.

He can be followed on twitter **@DuncanYB**.

This page intentionally left blank

About the Technical Reviewers

Christos Karamanolis is the Chief Architect and a Principal Engineer in the Storage and Availability Engineering Organization at VMware. He has more than 20 years of research and development experience in the fields of distributed systems, fault tolerance, storage, and storage management. He is the architect of Virtual SAN (VSAN), a new distributed storage system, and vSphere's policy-based storage management stack (S-PBM, VASA). Previously, he worked on the ESX storage stack (NFS client and vSCSI filters) and Disaster Recovery (vSphere Replication). Prior to joining VMware in 2005, Christos spent several years at HP Labs as a researcher working on new-generation storage products. He started his career as an Assistant Professor at Imperial College, while he also worked as an independent IT consultant. He has coauthored more than 20 research papers in peer-reviewed journals and conferences and has 24 granted patents. He holds a Ph.D. in Distributed Computing from Imperial College, University of London, U.K.

Paudie O'Riordan is a Staff Integration Liaison at VMware R&D. Formerly he worked in EMC Corporation (1996-2007) as an IT Admin, EMC Global Technical Support, Corporate Systems Engineer, and Principle Software Engineer at EMC R&D. Previously at VMware, he had a role in VMware Global Services as Senior Staff Technical Support Engineer. He holds the VCP certification (VCP4).

This page intentionally left blank

Acknowledgments

The authors of this book both work for VMware. The opinions expressed in the book are the authors' personal opinions and experience with the product. Statements made throughout the book do not necessarily reflect the views and opinions of VMware.

We would like to thank Christos Karamanolis and Paudie O'Riordan for keeping us honest as our technical editors. Of course, we want to thank the Virtual SAN engineering team. In particular, we want to call out two individuals of the engineering team, Christian Dickmann and once again Christos Karamanolis, whose deep knowledge and understanding of VSAN was leveraged throughout this book. We also want to acknowledge William Lam, Wade Holmes, Rawlinson Rivera, Simon Todd, Alan Renouf, and Jad El-Zein for their help and contributions to the book.

Lastly, we want to thank our VMware management team (Phil Weiss, Adam Zimman, and Mornay van der Walt) for supporting us on this and other projects.

Go VSAN!

Cormac Hogan and Duncan Epping

This page intentionally left blank

We Want to Hear from You!

As the reader of this book, *you* are our most important critic and commentator. We value your opinion and want to know what we're doing right, what we could do better, what areas you'd like to see us publish in, and any other words of wisdom you're willing to pass our way.

We welcome your comments. You can email or write us directly to let us know what you did or didn't like about this book—as well as what we can do to make our books better.

Please note that we cannot help you with technical problems related to the topic of this book.

When you write, please be sure to include this book's title and author as well as your name, email address, and phone number. We will carefully review your comments and share them with the author and editors who worked on the book.

Email: VMwarePress@vmware.com

Mail: VMware Press
ATTN: Reader Feedback
800 East 96th Street
Indianapolis, IN 46240 USA

Reader Services

Visit our website at www.informit.com/title/9780133854992 and register this book for convenient access to any updates, downloads, or errata that might be available for this book.

This page intentionally left blank

Introduction

When talking about virtualization and the underlying infrastructure that it runs on, one component that always comes up in conversation is storage. The reason for this is fairly simple: In many environments, storage is a pain point. Although the storage landscape has changed with the introduction of flash technologies that mitigate many of the traditional storage issues, many organizations have not yet adopted these new architectures and are still running into the same challenges.

Storage challenges range from operational effort or complexity to performance problems or even availability constraints. The majority of these problems stem from the same fundamental problem: legacy architecture. The reason is that most storage platform architectures were developed long before virtualization existed, and virtualization changed the way these shared storage platforms were used.

In a way, you could say that virtualization forced the storage industry to look for new ways of building storage systems. Instead of having a single server connect to a single storage device (also known as a logical unit or LUN for short), virtualization typically entails having one (or many) physical server(s) running many virtual machines connecting to one or multiple storage devices. This did not only increase the load on these storage systems, it also changed the workload patterns and increased the total capacity required.

As you can imagine, for most storage administrators, this required a major shift in thinking. What should the size of my LUN be? What are my performance requirements, and how many spindles will that result in? What kind of data services are required on these LUNs, and where will virtual machines be stored? Not only did it require a major shift in thinking, but it also required working in tandem with other IT teams. Whereas in the past server admins and network and storage admins could all live in their own isolated worlds, they now needed to communicate and work together to ensure availability of the platform they were building. Whereas in the past a mistake, such as a misconfiguration or underprovisioning, would only impact a single server, it could now impact many virtual machines.

There was a fundamental shift in how we collectively thought about how to operate and architect IT infrastructures when virtualization was introduced. Now another collective shift is happening all over again. This time it is due to the introduction of software-defined networking and software-defined storage. But let's not let history repeat itself, and let's avoid the mistakes we all made when virtualization first arrived. Let's all have frank and open discussions with our fellow datacenter administrators as we all aim to revolutionize datacenter architecture and operations!

Motivation for Writing This Book

During the early stages of the product development cycle, both of us got involved with Virtual SAN. We instantly knew that this was going to be a product that everyone would be talking about and a product that people would want to know more about. During the various great water-cooler type of conversations we were having, we realized that none of the information was being captured anywhere. Considering both of us are fanatic bloggers we decided to, each independently, start writing articles. We quickly had so much material that it became impossible to release all of it as blog posts and decided to join forces and publish it in book form. After some initial research, we found VMware Press was willing to release it.

You, the Reader

This book is targeted at IT professionals who are involved in the care and feeding of a VMware vSphere environment. Ideally, you have been working with VMware vSphere for some time and perhaps you have attended an authorized course in vSphere, such as the "Install, Configure, and Manage" class. This book is not a starters guide, but there should be enough in the book for administrators and architects of all levels.

How to Use This Book

This book is split into ten chapters, as described here:

- **Chapter 1, "Introduction to VSAN":** This chapter provides a high-level introduction to software-defined storage and VSAN.
- **Chapter 2, "VSAN Prerequisites and Requirements for Deployment":** This chapter describes the requirements from a physical and virtual perspective to safely implement VSAN.
- **Chapter 3, "VSAN Installation and Configuration":** This chapter goes over the steps needed to install and configure VSAN.
- **Chapter 4, "VM Storage Policies on VSAN":** This chapter explains the concept of storage policy-based management.
- **Chapter 5, "Architectural Details":** This chapter provides in-depth architectural details of VSAN.
- **Chapter 6, "VM Storage Policies and Virtual Machine Provisioning":** This chapter describes how VM storage policies can be used to simplify VM deployment.

- **Chapter 7, "Management and Maintenance":** This chapter describes the steps for most common management and maintenance tasks.
- **Chapter 8, "Interoperability":** This chapter covers interoperability of Virtual SAN with other VMware features and products.
- **Chapter 9, "Designing a VSAN Cluster":** This chapter provides various examples around designing a VSAN cluster, including sizing exercises.
- **Chapter 10, "Troubleshooting, Monitoring, and Performance":** This chapter covers the various (command line) tools available to troubleshoot and monitor VSAN.

This page intentionally left blank

Chapter 1

Introduction to VSAN

This chapter introduces you to the world of the software-defined datacenter, but with a focus on the storage aspect. The chapter covers the basic premise of the software-defined datacenter and then delves deeper to cover the concept of software-defined storage and associated solutions such as the server storage-area network (Server SAN).

Software-Defined Datacenter

VMworld, the VMware annual conferencing event, introduced VMware's vision for the software-defined datacenter (SDDC) in 2012. The SDDC is VMware's architecture for the public and private clouds where all pillars of the datacenter—compute, storage, and networking (and the associated services)—are virtualized. Virtualizing datacenter components enables the IT team to be more flexible. If you lower the operational complexity and cost while increasing availability and agility, you will ultimately lower the time to market for new services.

To achieve all of that, virtualization of components by itself is not sufficient. The platform used must be capable of being installed and configured in a fully automated fashion. More importantly, the platform should enable you to manage and monitor your infrastructure in a smart and less operationally intense manner. That is what the SDDC is all about! Raghu Raghuram (VMware senior vice president) captured it in a single sentence: The essence of the software-defined datacenter is "abstract, pool, and automate."

Abstraction, pooling, and automation are all achieved by introducing an additional layer on top of the physical resources. This layer is usually referred to as a *virtualization layer*. Everyone reading this book is probably familiar with the leading product for compute virtualization, VMware vSphere. Fewer people are probably familiar with network virtualization, sometimes referred to as software-defined network (SDN) solutions. VMware offers a solution named NSX that is based on the solution built by the acquired company Nicira. NSX does for networking what vSphere does for compute. These layers do not just virtualize the physical resources but also allow you to pool them and provide you with an application programming interface (API) that enables you to automate all operational aspects.

Automation is not just about scripting, however. A significant part of the automation of virtual machine (VM) provisioning (and its associated resources) is achieved through policy-based management. Predefined policies allow you to provision VMs in a quick, easy, consistent, and repeatable manner. The resource characteristics specified on a resource pool or a vApp container exemplify a compute policy. These characteristics enable you to quantify resource policies for compute in terms of reservation, limit, and priority. Network policies can range from security to quality of service (QoS). Unfortunately, storage has thus far been limited to the characteristics provided by the physical storage device, which in many cases did not meet the expectations and requirements of many of our customers.

This book examines the storage component of VMware's SDDC. More specifically, the book covers how a new product called Virtual SAN (VSAN), releasing with VMware vSphere 5.5 Update 1, fits into this vision. You will learn how it has been implemented and integrated within the current platform and how you can leverage its capabilities and expand on some of the lower-level implementation details. Before going further, though, you want to have a generic understanding of where VSAN fits in to the bigger software-defined storage picture.

Software-Defined Storage

Software-defined storage is a term that has been used and abused by many vendors. Because software-defined storage is currently defined in so many different ways, consider the following quote from VMware:

> Software Defined Storage is the automation and pooling of storage through a software control plane, and the ability to provide storage from industry standard servers. This offers a significant simplification to the way storage is provisioned and managed, and also paves the way for storage on industry standard servers at a fraction of the cost. (Source: http://cto.vmware.com/vmwares-strategy-for-software-defined-storage/)

A software-defined storage product is a solution that abstracts the hardware and allows you to easily pool all resources and provide them to the consumer using a user-friendly user interface (UI) or API. A software-defined storage solution allows you to both scale up and scale out, without increasing the operational effort.

Many hold that software-defined storage is about moving functionality from the traditional storage devices to the host. This is a trend that was started by virtualized versions of storage devices such as HP's StoreVirtual VSA and evolved into solutions that were built to run on many different hardware platforms. One example of such a solution is Nexenta. These solutions were the start of a new era.

Hyper-Convergence/Server SAN Solutions

In today's world, the hyper-converged/server SAN solutions come in two flavors:

- Hyper-converged appliances
- Software-only solutions

A hyper-converged solution is an appliance type of solution where a single box provides a platform for VMs. This box typically contains multiple commodity x86 servers on which a hypervisor is installed. Local storage is aggregated into a large shared pool by leveraging a virtual storage appliance or a kernel-based storage stack. Typical examples of hyper-converged appliances that are out there today include Nutanix, Scale Computing, SimpliVity, and Pivot3. Figure 1-1 shows what these appliances usually look like: a 2U form factor with four hosts.

Figure 1-1 Commonly used hardware by hyper-converged storage vendors

You might ask, "If these are generic x86 servers with hypervisors installed and a virtual storage appliance, what are the benefits over a traditional storage system?" The benefits of a hyper-converged platform are as follows:

- Time to market is short, less than 4 hours to install and deploy
- Ease of management and integration

- Able to scale out, both capacity and performance-wise
- Lower total costs of acquisition compared to traditional environments

These solutions are sold as a single stock keeping unit (SKU), and typically a single point of contact for support is provided. This can make support discussions much easier. However, a hurdle for many companies is the fact that these solutions are tied to hardware and specific configurations. The hardware used by hyper-converged vendors is often not the same as from the preferred hardware supplier you may already have. This can lead to operational challenges when it comes to updating/patching or even cabling and racking. In addition, a trust issue exists. Some people swear by server Vendor X and would never want to touch any other brand, whereas others won't come close to server Vendor X. This is where the software-based storage solutions come in to play.

Software-only storage solutions come in two flavors. The most common solution today is the virtual storage appliance (VSA). VSA solutions are deployed as a VM on top of a hypervisor installed on physical hardware. VSAs allow you to pool underlying physical resources into a shared storage device. Examples of VSAs include VMware vSphere Storage Appliance, Maxta, HP's StoreVirtual VSA, and EMC Scale IO. The big advantage of software-only solutions is that you can usually leverage existing hardware as long as it is on the hardware compatibility list (HCL). In the majority of cases, the HCL is similar to what the used hypervisor supports, except for key components like disk controllers and flash devices.

VSAN is also a software-only solution, but VSAN differs significantly from the VSAs listed. VSAN sits in a different layer and is not a VSA-based solution.

Introducing Virtual SAN

VMware's plan for software-defined storage is to focus on a set of VMware initiatives related to local storage, shared storage, and storage/data services. In essence, VMware wants to make vSphere a platform for storage services.

Historically, storage was something that was configured and deployed at the start of a project, and was not changed during its life cycle. If there was a need to change some characteristics or features of the logical unit number (LUN) or volume that were being leveraged by VMs, in many cases the original LUN or volume was deleted and a new volume with the required features or characteristics was created. This was a very intrusive, risky, and time-consuming operation due to the requirement to migrate workloads between LUNs or volumes, which may have taken weeks to coordinate.

With software-defined storage, VM storage requirements can be dynamically instantiated. There is no need to repurpose LUNs or volumes. VM workloads and requirements may change over time, and the underlying storage can be adapted to the workload at any time. VSAN aims to provide storage services and service level agreement *automation* through a software layer on the hosts that *integrates* with, *abstracts*, and *pools* the underlying hardware.

A key factor for software-defined storage is storage policy–based management (SPBM). This is also a key feature in the vSphere 5.5 release. SPBM can be thought of as the next generation of VMware's storage profile features that was introduced with vSphere 5.0. Where the initial focus of storage profiles was more about ensuring VMs were provisioned to the correct storage device, in vSphere 5.5. SPBM is a critical component to how VMware is implementing software-defined storage.

Using SPBM and vSphere APIs, the underlying storage technology surfaces an abstracted pool of storage space with various capabilities that is presented to vSphere administrators for VM provisioning. The capabilities can relate to performance, availability, or storage services such as thin provisioning, compression, replication, and more. A vSphere administrator can then create a *VM storage policy* (or profile) using a subset of the capabilities that are required by the application running in the VM. At deployment time, the vSphere administrator selects a VM storage policy. SPBM pushes the VM storage policy down to the storage layer and datastores that understand that the requirements placed in the VM storage policy will be made available for selection. This means that the VM is always instantiated on the appropriate underlying storage based on the requirements placed in the VM storage policy.

Should the VM's workload or I/O pattern change over time, it is simply a matter of applying a new VM storage policy with requirements and characteristics that reflect the new workload to that specific VM, or even virtual disk, after which the policy will be seamlessly applied without any manual intervention from the administrator (in contrast to many legacy storage systems, where a manual migration of VMs or virtual disks to a different datastore would be required). VSAN has been developed to seamlessly integrate with vSphere and the SPBM functionality it offers.

What Is Virtual SAN?

VSAN is a new storage solution from VMware, released as a beta in 2013 and made generally available to the public in March 2014. VSAN is fully integrated with vSphere. It is an object-based storage system and a platform for VM storage policies that aims to simplify VM storage placement decisions for vSphere administrators. It fully supports and is integrated with core vSphere features such as vSphere High Availability (HA), vSphere Distributed Resource Scheduler (DRS), and vMotion, as illustrated in Figure 1-2.

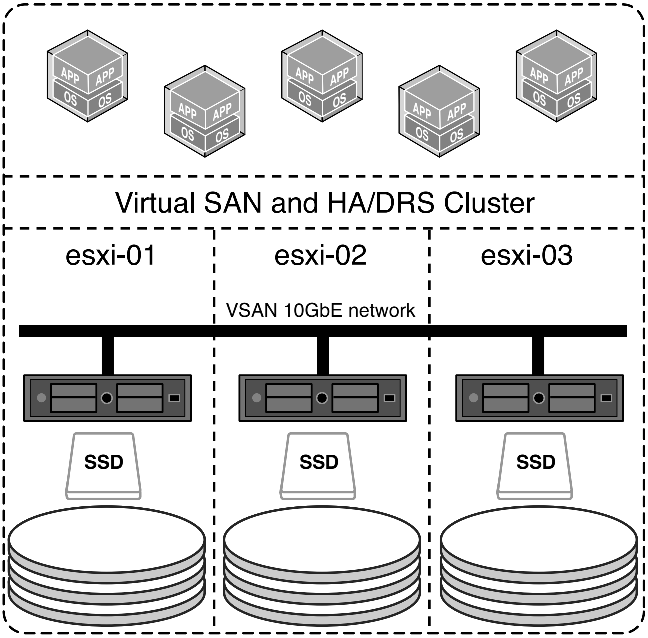

Figure 1-2 Simple overview of a VSAN cluster

VSAN's goal is to provide both resiliency and scale-out storage functionality. It can also be thought of in the context of QoS in so far as VM storage policies can be created that define the level of performance and availability required on a per-VM, or even virtual disk, basis.

VSAN is a software-based distributed storage solution that is built directly in the hypervisor. Although not a virtual appliance like many of the other solutions out there, a VSAN can best be thought of as a kernel-based solution that is included with the hypervisor. Technically, however, this is not completely accurate because components critical for performance and responsiveness such as the data path and clustering are in the kernel, while other components that collectively can be considered part of the "control plane" are implemented as native user-space agents. Nevertheless, with VSAN there is no need to install anything other than the software you are already familiar with: VMware vSphere.

VSAN is about simplicity, and when we say *simplicity*, we do mean simplicity. Want to try out VSAN? It is truly as simple as creating a VMkernel network interface card (NIC) for VSAN traffic and enabling it on a cluster level, as shown in Figure 1-3. Of course, there are certain recommendations and requirements to optimize your experience, as described in further detail in Chapter 2, "VSAN Prerequisites and Requirements for Deployment."

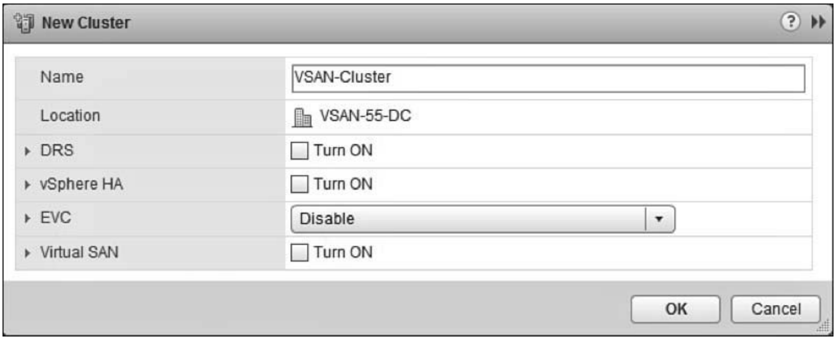

Figure 1-3 Two-click enablement

Now that you know it is easy to use and simple to configure, what are the benefits of a solution like VSAN? What are the key selling points?

- **Software defined**: Use industry standard hardware
- **Flexible**: Scale as needed and when needed, both scale up and scale out
- **Simple**: Ridiculously easy to manage and operate
- **Automated**: Per-VM and disk policy-based management
- **Converged**: Enables you to create dense/building-block-style solutions

That sounds compelling, doesn't it? Of course, there is a time and place for everything; Virtual SAN 1.0 has specific use cases. For version 1.0, these use cases are as follows:

- **Virtual desktops**: Scale-out model using predictive and repeatable infrastructure blocks lowers costs and simplifies operations
- **Test and dev**: Avoids acquisition of expensive storage (lowers total cost of ownership [TCO]), fast time to provision
- **Management or DMZ infrastructure**: Fully isolated resulting in increased security and no dependencies on the resources it is potentially managing.
- **Disaster recovery target**: Inexpensive disaster recovery solution, enabled through a feature like vSphere Replication that allows you to replicate to any storage platform

Now that you know what VSAN is, it's time to see what it looks like from an administrator's point of view.

What Does VSAN Look Like to an Administrator?

When VSAN is enabled, a single shared datastore is presented to all hosts that are part of the VSAN-enabled cluster. This is the strength of VSAN; it is presented as a datastore. Just like any other storage solution out there, this datastore can be used as a destination for VMs and all associated components, such as virtual disks, swap files, and VM configuration files. When you deploy a new VM, you will see the familiar interface and a list of available datastores, including your VSAN-based datastore, as shown in Figure 1-4.

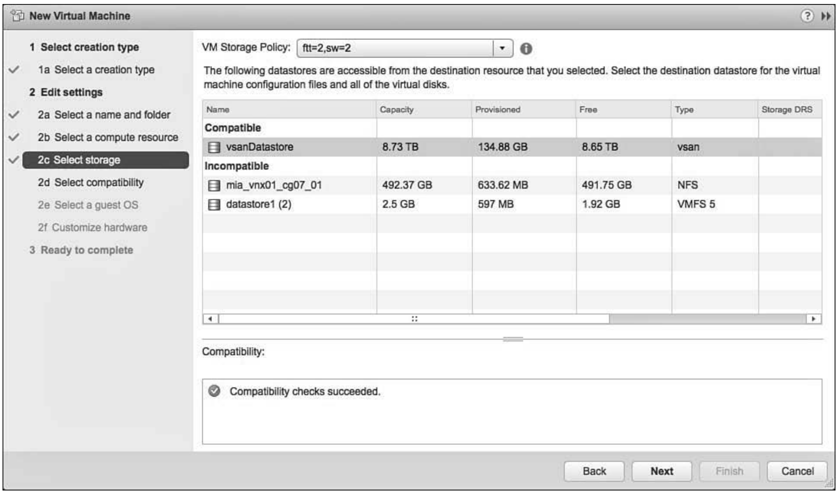

Figure 1-4 Just a normal datastore

This VSAN datastore is formed out of host local storage resources. Typically, all hosts within a VSAN-enabled cluster will contribute performance (flash) and capacity (magnetic disks) to this shared datastore. This means that when your cluster grows, your datastore will grow with it. VSAN is what is called a scale-out storage system (adding hosts to a cluster), but also allows scaling up (adding resources to a host).

Each host that wants to contribute storage capacity to the VSAN cluster will require at least one flash device and one magnetic disk. At a minimum, VSAN requires three hosts in your cluster to contribute storage; other hosts in your cluster could leverage these storage resources without contributing storage resources to the cluster itself. Figure 1-5 shows a cluster that has four hosts, of which three (esxi-01, esxi-02, and esxi-03) contribute storage and a fourth does not contribute but only consumes storage resources. Although it is technically possible to have a nonuniform cluster and have a host not contributing storage, we do highly recommend creating a uniform cluster and having all hosts contributing storage for overall better utilization, performance, and availability.

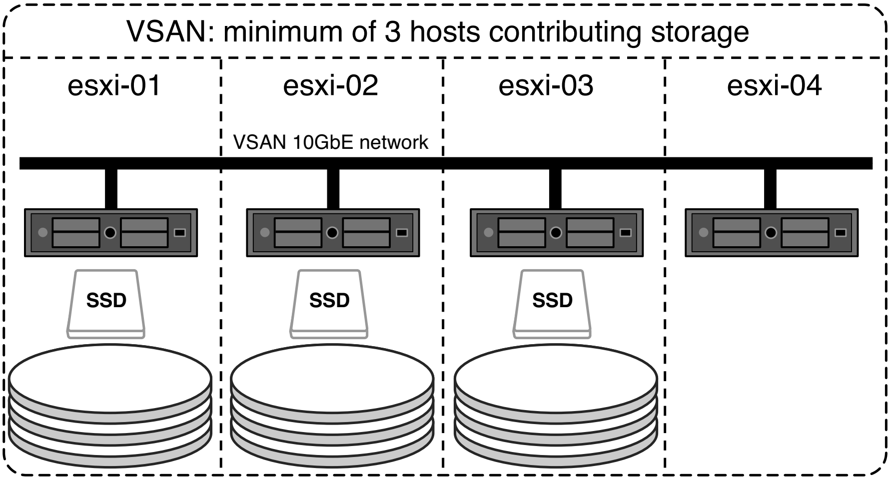

Figure 1-5 Nonuniform VSAN cluster example

Today's boundary for VSAN in terms of both size and connectivity is a vSphere cluster. This means that at most 32 hosts can be connected to a VSAN datastore. Each host can run a supported maximum of 100 VMs, allowing for a total combined of 3,200 VMs within a 32-host VSAN cluster, of which 2,048 VMs can be protected by vSphere HA.

As you can imagine, with just regular magnetic disks it would be difficult to provide a good user experience when it comes to performance. To provide optimal user experience, VSAN relies on flash. Flash resources are used for read caching and write buffering. Every write I/O will go to flash first, and eventually will be destaged to magnetic disks. For read I/O it will depend, although in a perfect world all read I/O will come from flash. Chapter 5, "Architectural Details," describes the caching and buffering mechanisms in much greater detail.

To ensure VMs can be deployed with certain characteristics, VSAN enables you to set policies on a per-virtual disk or a per-VM basis. These policies help you meet the defined service level objectives (SLOs) for your workload. These can be performance-related characteristics such as read caching or disk striping, but can also be availability-related characteristics that ensure strategic replica placement of your VM's disks (and other important files).

If you have worked with VM storage policies in the past, you might now wonder whether all VMs stored on the same VSAN datastore will need to have the same VM storage policy assigned. The answer is no. VSAN allows you to have different policies for VMs provisioned to the same datastore and even different policies for disks from the same VM.

As stated earlier, by leveraging policies, the level of resiliency can be configured on a per-virtual disk granular level. How many hosts and disks a mirror copy will reside on depends on the selected policy. Because VSAN uses mirror copies defined by policy to provide resiliency, it does not require a local RAID set. In other words, hosts contributing to VSAN storage capacity should simply provide a set of disks to VSAN.

Whether you have defined a policy to tolerate a single host failure or, for instance, a policy that will tolerate up to three hosts failing, VSAN will ensure that enough replicas of your objects are created. The following example illustrates how this is an important aspect of VSAN and one of the major differentiators between VSAN and most other virtual storage solutions out there.

EXAMPLE: We have configured a policy that can tolerate one failure and created a new virtual disk. This means that VSAN will create two identical storage objects and a witness. The witness is a component tied to the VM that allows VSAN to determine who should win ownership in the case of a failure. If you are familiar with clustering technologies, think of the witness as a quorum object that will arbitrate ownership in the event of a failure. Figure 1-6 may help clarify these sometimes-difficult-to-understand concepts. This figure illustrates what it would look like on a high level for a VM with a virtual disk that can tolerate one failure. This can be the failure of a host, NICs, disk, or flash device, for instance.

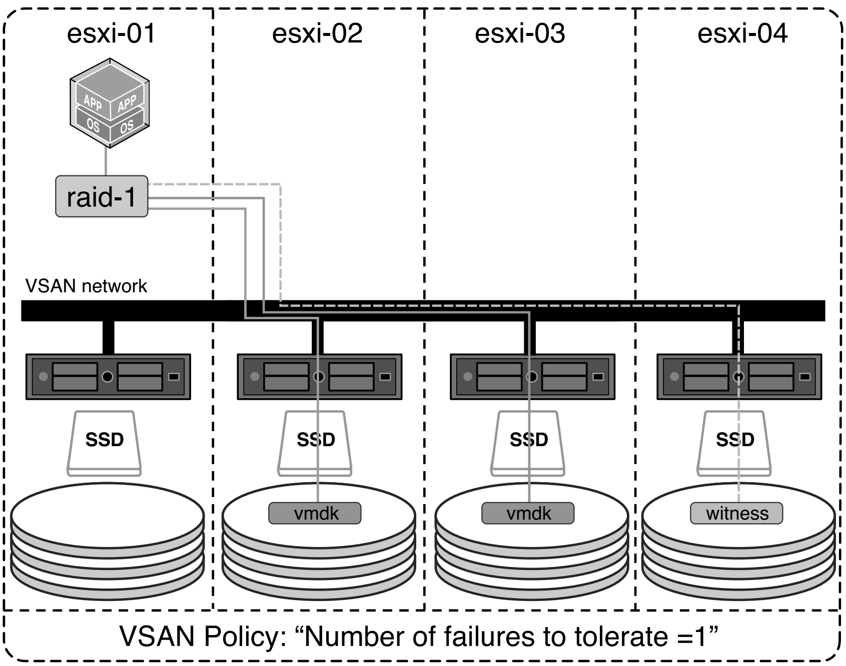

Figure 1-6 VSAN failures to tolerate

In Figure 1-6, the VM's compute resides on the first host (esxi-01) and its virtual disks reside on the other hosts (esxi-02 and esxi-03) in the cluster. In this scenario, the VSAN network is used for storage I/O, allowing for the VM to freely move around the cluster without the need for storage components to be migrated with the compute. This does, however, result in the first requirement to implement VSAN. VSAN requires at a minimum one dedicated 1Gbps NIC port, but VMware recommends a 10GbE for the VSAN network.

Yes, this might still sound complex, but in all fairness, VSAN masks away all the complexity, as you will learn as you progress through the various chapters in this book.

Summary

To conclude, vSphere Virtual SAN (VSAN) is a brand-new, hypervisor-based distributed storage platform that enables convergence of compute and storage resources. It enables you to define VM-level granular SLOs through policy-based management. It allows you to control availability and performance in a way never seen before, simply and efficiently.

This chapter just scratched the surface. Now it's time to take it to the next level. Chapter 2 describes the requirements for installing and configuring VSAN.

This page intentionally left blank

Index

A

B

C

D

E

F

S

T

W-Z