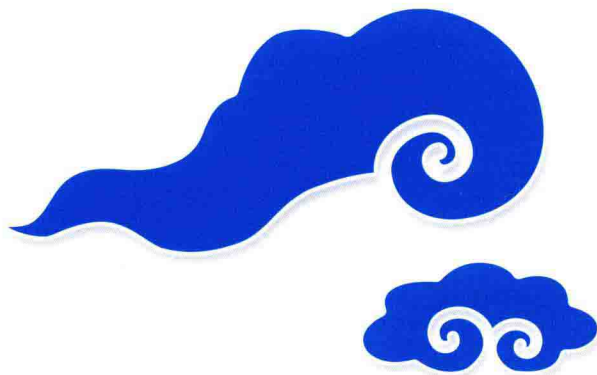




VMware Virtual SAN in Action

VMware Virtual SAN 实战

吴秋林 著

VSAN领域著名专家撰写，10余年虚拟化产品研究、实践经验结晶，名副其实的存储虚拟化领域良心之作。

源自5000篇技术文档精华，从基础概念、产品构建到原理解析，逐层解析VSAN，已帮助数千一线人员解决了实际问题。



机械工业出版社
China Machine Press

云计算与虚拟化技术丛书



VMware Virtual SAN in Action

VMware Virtual SAN 实战

吴秋林 著



机械工业出版社
China Machine Press

图书在版编目 (CIP) 数据

VMware Virtual SAN 实战 / 吴秋林著. —北京: 机械工业出版社, 2016.4
(云计算与虚拟化技术丛书)

ISBN 978-7-111-53522-5

I. V… II. 吴… III. 虚拟处理机 IV. TP338

中国版本图书馆 CIP 数据核字 (2016) 第 073002 号

VMware Virtual SAN 实战

出版发行: 机械工业出版社 (北京市西城区百万庄大街 22 号 邮政编码: 100037)

责任编辑: 孙海亮

责任校对: 殷虹

印 刷: 三河市宏图印务有限公司

版 次: 2016 年 4 月第 1 版第 1 次印刷

开 本: 186mm×240mm 1/16

印 张: 12.75

书 号: ISBN 978-7-111-53522-5

定 价: 59.00 元

凡购本书, 如有缺页、倒页、脱页, 由本社发行部调换

客服热线: (010) 88379426 88361066

投稿热线: (010) 88379604

购书热线: (010) 68326294 88379649 68995259

读者信箱: hzit@hzbook.com

版权所有·侵权必究

封底无防伪标均为盗版

本书法律顾问: 北京大成律师事务所 韩光 / 邹晓东

Preface 前言

存储从产生至今，已经经过数十个年头了，天下大势，合久必分，分久必合。在虚拟化、云计算数据中心的大背景下，在满足了企业计算高可用、网络高可用之后，企业对于存储的高可用、双活等的要求也就随之提上了议事日程。

传统的存储解决方案要想实现存储的双活，其中最大的问题就在于：对于用户来说，实现双活的费用太高了，尤其是对于用户的 TCO 与 ROI 而言，更是如此。大家可以想一下，这就相当于你买了两台 iPhone 6 Plus 手机，一台正常使用，一台放着不用。这毕竟不是豆浆，价格低廉到可以让你买一杯倒一杯。

同时，传统的共享存储有一个大问题：支持真正双活的存储设备价格很高昂，而中低端存储设备根本无法实现存储层面的双活。

可随着业务的发展，必然存在双活的要求。正是在这样的背景下，以 Nutanix、Ceph 和 Virtual SAN 之类为代表的存储虚拟化产品出现了，这类型的产品最大的卖点就在于：

- ❑ 不用独立的 Standby 存储设备；
- ❑ 不用额外的双活软件支持；
- ❑ 不用额外的兼容性；
- ❑ 不用额外的可用性；
- ❑ 不再需要考虑 LUN、Volume 等的规划；

- ❑ 价格低廉；
- ❑ 改造技术成本低廉；
- ❑ 深度整合数据中心虚拟化产品；
- ❑

在存储虚拟化的支持下，存储虚拟化产品全面转向了面向对象的服务级别。所有业务的可用性、数据可靠性，都依赖于面向对象的策略进行工作，可以个性化地针对不同业务级别定义不同的对象策略，实现不同的可用性级别和数据对于物理硬件的写入位置分割。

在保障计算可用的基础上，存储虚拟化产品同时保障了数据的可用性。在闪存设备与 HDD 设备的有机组合下，提供了远超传统共享存储的 IOPS 支持，解决了传统性能密集型业务（例如：Oracle 数据库之类的业务）无法运行在 x86 平台上的问题。

在软件定义存储产品出现之前，x86 结构下对于承载 OLTP（Online Transaction Processing）或 OLAP（Online Analytical Processing）之类的业务是有所欠缺的。因为 x86 结构下的业务模型，导致了对计算、I/O 密集型业务的承载能力的欠缺，尤其是 I/O 密集型业务，比如 Oracle RAC 之类的业务、医疗领域 PACS 系统等。传统的 SAN 存储用于虚拟化，要么成本过高，要么无法匹配 x86 虚拟化本身。正是在这个背景之下，Server SAN 出现了，其迅速迎合了这类业务的需求。

以 Virtual SAN 为例，它已经得到了国内部分三甲医院的认可，并用于 PACS 系统，也得到了部分证券型业务的认可。虽然现在其应用规模还不算太大，但无疑证明了它的可用性。

Contents 目 录

前 言

第 1 章 Virtual SAN 产品概念	1
1.1 产品介绍	1
1.2 Virtual SAN 关键概念	3
1.3 Virtual SAN 与传统存储	4
1.4 Virtual SAN 的构建方式	5
1.4.1 Virtual SAN Ready Nodes	5
1.4.2 VMware EVO: RAIL	6
1.4.3 DIY: Do It Yourself	6
1.5 Virtual SAN 产品的兼容性	8
1.6 Virtual SAN 功能限制	9
1.7 Virtual SAN 适用的业务场景	10
1.8 本章小结	11
第 2 章 软件定义数据中心与 Virtual SAN	12
2.1 VMware 软件定义数据中心产品清单	13
2.1.1 vSphere 计算虚拟化	14

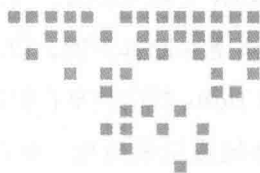
2.1.2	Virtual SAN 存储虚拟化	15
2.1.3	NSX 网络虚拟化	15
2.1.4	vRealize Automation Center 自动化云平台	17
2.1.5	vRealize Operations Manager 统一监控	19
2.1.6	Horizon 桌面虚拟化	21
2.2	软件定义数据中心之软件定义存储	23
2.2.1	构建软件定义数据中心的概念	24
2.2.2	软件定义存储选择	25
2.2.3	VVoLS 技术特征	26
2.2.4	Virtual SAN 技术特征	27
2.3	为软件定义数据中心准备 Virtual SAN	28
2.4	本章小结	29
第 3 章	Virtual SAN 配置要求	30
3.1	Virtual SAN 硬件配置要求	30
3.2	Virtual SAN 软件配置要求	31
3.3	Virtual SAN License 的要求	32
3.4	本章小结	33
第 4 章	Virtual SAN Cluster 构建准备工作	35
4.1	Virtual SAN 存储组件	36
4.1.1	Flash 缓存组件	37
4.1.2	Flash 空间组件	39
4.1.3	HDD 磁盘组件	39
4.2	Virtual SAN Cluster 配置要求	40
4.3	Virtual SAN 网络配置要求	42
4.3.1	Virtual SAN 与 vSwitch	42
4.3.2	Virtual SAN 与 Jumbo Frames 巨帧	44

4.3.3 Virtual SAN 与 NIC Teaming 负载均衡策略	46
4.4 Virtual SAN Fault Domain 配置要求	47
4.5 Virtual SAN Cluster 日志问题	48
4.6 本章小结	49
第 5 章 构建一个全新的 Virtual SAN Cluster	50
5.1 软件准备	51
5.2 Virtual SAN Cluster 主机节点准备	52
5.3 安装 vSphere 主机	52
5.4 配置 vSphere 主机信息	57
5.5 部署 vCenter Server	62
5.6 将 vSphere 主机添加到 vCenter	71
5.6.1 添加 Identity Source	72
5.6.2 输入 License	73
5.6.3 添加 vSphere 主机到 vCenter	75
5.7 Virtual SAN 节点主机磁盘准备	77
5.7.1 ESXCLI 命令行方式	77
5.7.2 vSphere Web Client 图形界面	79
5.8 Virtual SAN 主机网络准备	82
5.8.1 为 vSphere Standard Switch 创建 VSAN 网络	82
5.8.2 为 vSphere Distributed Switch 创建 VSAN 网络	85
5.9 创建 Virtual SAN Cluster	98
5.9.1 创建 Manual 模式下的 Virtual SAN Cluster	99
5.9.2 查看 Virtual SAN Datastore	104
5.10 本章小结	106
第 6 章 Virtual SAN 结构性原理解析	108
6.1 Virtual SAN 的对象与组件结构	108

6.1.1	混合存储的数据读取流程	109
6.1.2	存储结构的数据写入流程	110
6.1.3	混合存储的 Cache 与 Capacity	111
6.1.4	全闪存的 Cache 和 Capacity	111
6.2	Virtual SAN I/O 数据变化	112
6.3	Virtual SAN 中的 Witness 详解	113
6.4	Virtual SAN 5.5 中的 Witness 类型	113
6.5	本章小结	115
第 7 章	维护 Virtual SAN Cluster	116
7.1	管理 Virtual SAN Storage Providers	117
7.2	管理 Virtual SAN Cluster 设备	118
7.3	扩容 Virtual SAN Cluster	122
7.4	Maintenance Mode	124
7.5	Fault Domains 管理	126
7.6	Storage Policy-Based Management 概念	130
7.6.1	默认 VM Storage Policies	130
7.6.2	自定义 VM Storage Policies	132
7.7	VM Storage Policies 深度解析	135
7.8	为对象更改 VM Storage Policies	138
7.9	为对象调整全新 VM Storage Policies	140
7.10	禁用 Virtual SAN	141
7.11	本章小结	144
第 8 章	升级 Virtual SAN Cluster	145
8.1	升级前的准备工作	145
8.2	vCenter Server 和 ESXi Hosts 升级	146
8.3	Virtual SAN Disk Format 升级	146

8.4 本章小结	148
第 9 章 简单监控 Virtual SAN Cluster	149
9.1 Virtual SAN Cluster 简单监控	149
9.2 创建针对 VSAN 的 Alarms	154
9.3 RVC 监控	155
9.3.1 在 VCSA 上使用 RVC	157
9.3.2 RVC 中关于 VSAN 参数的简单运用	157
9.4 VSAN Observer 监控	158
9.5 vRealize Operations Manager 监控 Virtual SAN	160
9.6 本章小结	161
第 10 章 Virtual SAN 常见故障	162
10.1 Absent 与 Degraded 故障	163
10.2 Compliant 与 Not Compliant 故障	164
10.3 Applicable 与 Not Applicable 故障	165
10.4 Out of Date 故障	165
10.5 None 故障	166
10.6 healthy 与 unhealthy 故障	166
10.7 inaccessible 与 orphaned 故障	167
10.8 Cache 的闪存故障	167
10.9 HDD 及 Network 相关故障	167
10.10 Storage Provider 故障	171
10.11 本章小结	172
第 11 章 Virtual SAN 简单性能	173
11.1 I/O 的类型	174
11.2 FTT 参数与性能输出	175

11.3	Stripe 参数与性能输出	175
11.4	Disk Groups 数量与性能输出	176
11.5	网络参数与性能输出	177
11.6	其他因素与性能输出	178
11.7	Virtual SAN 性能相关结论	178
11.8	本章小结	179
第 12 章	Virtual SAN 6.1 快速预览	180
12.1	Virtual SAN 6.1 的新功能	181
12.2	Virtual SAN 6.1 新功能的优势	181
12.2.1	支持 Stretched Cluster	181
12.2.2	SMP-FT 与 Virtual SAN	182
12.2.3	结合 SRM 中的 VR 实现 RPO 最小为 5	182
12.2.4	Virtual SAN for vROPS Suite Packs 组件	183
12.3	Virtual SAN 6.1 与 6.0 的差异	184
12.4	Virtual SAN 结构针对 Witness 的变化	184
12.5	Virtual SAN 6.1 高级参数	185
12.6	Virtual SAN Beta 未来版预期	188
12.7	本章小结	188
结束语		189
附录	参考资料清单	191



Virtual SAN 产品概念

章节概要

本章的核心内容如下：

- ❑ Virtual SAN 产品介绍；
- ❑ Virtual SAN 关键概念；
- ❑ Virtual SAN 与传统存储的异同；
- ❑ Virtual SAN 的多种构建方式；
- ❑ Virtual SAN 的兼容性；
- ❑ Virtual SAN 的优缺点；
- ❑ Virtual SAN 与其他 VMware 公司产品的结合度。

1.1 产品介绍

VMware Virtual SAN 是由 Hypervisor-Converged 软件定义的分布式存储基础平台，它内置于 VMware 公司的 vSphere 服务器虚拟化平台，最早出现在 vSphere 5.5 Update 1 版本中，截止至 2015 年 3 月，最新版本为 vSphere 6.0 的 Virtual SAN 6.0，它是一个存储虚拟化产品。

Virtual SAN 又称 VSAN，当前版本是 6.1，也是第 3 个生产发布版本。它利用物理服务器本地硬件设备资源，以类似分布式存储技术的方式，为 vSphere Cluster 中的所有 ESXi Hosts 提供共享存储服务。Virtual SAN 创造性地利用策略驱动技术来帮助 vSphere 环境提供最简化、最高效和最快速的存储部署与管理支持。利用 VM 级别的存储策略，Virtual SAN 支持自动化、动态的按需匹配 VMs 存储资源分配。有了 Virtual SAN，针对存储的管理、维护、扩容等，都将变得异常简单。同时，存储带来的额外硬件成本开销也变得更加低廉。相对于传统存储而言，Virtual SAN 的优势主要体现在下列几个方面：

- ❑ 低成本；
- ❑ 低管理成本；
- ❑ 高可用性；
- ❑ 管理简单；
- ❑ 高弹性；
- ❑ 高性能；
- ❑ 支持融合式能力；
- ❑ 同城双活。

利用位于物理服务器本地的 SSD 与 HDD 磁盘资源的组合，构成了 Virtual SAN 节点的基础构成单元。Virtual SAN 5.5 作为第一个 VMware 公司发布的存储虚拟化产品，仅支持 SSD 与 HDD 的混合存储组合的底层物理硬件构成。最新发布的 Virtual SAN 6.0 和 VSAN 6.1 则支持全闪存与混合存储两种结构形态。

在混合存储结构下，SSD 设备被用作 Cache 来优化读写性能，HDD 则用作对象存放目标，负责为存储对象提供存储空间和永久性数据存放点。

在全闪存结构存储下，无论是 Cache，还是存储空间，都由 SSD 设备负责提供。表 1-1 所示是混合存储在 Virtual SAN 5.5、Virtual SAN 6.0 中与全闪存存储的功能比较。

表 1-1

	Virtual SAN 5.5	Virtual SAN 6.0 混合	Virtual SAN 6.0 全闪存
每个集群的主机数	32	64	64
每个主机的虚拟机数	100	200	200
每个主机的 IOPS	20K	40K	90K
快照深度按虚拟机数量	2	32	32
虚拟磁盘大小	2 TB	62 TB	62 TB

1.2 Virtual SAN 关键概念

Virtual SAN 作为区别于传统集中存储的类分布式存储技术，在基础概念上，与传统存储的概念有着巨大的差异。Virtual SAN 的关键基础概念如表 1-2 所示。

表 1-2

基础概念	功能
JBOD	JBOD (Just a Bunch Of Disks) 主要针对刀片服务器环境。中文简称为简单磁盘捆绑，简单地说就是一个非阵列类型的磁盘设备
Virsto	Virsto 是 VMware 公司收购的一家虚拟化存储设备管理公司，其产品现内置于 vSphere 中，为 Virtual SAN 提供磁盘文件格式支持。这是一个 log-based 的文件系统，针对虚拟机快照和克隆部分提供了巨大的改善空间。在它的帮助下，Virtual SAN 6.0 可以支持 32 个快照，而 Virtual SAN 5.5 仅支持 2 个快照
Fault Domains	Fault Domains 是 Virtual SAN 6.0 推出的新功能，利用它可以保障跨刀片中心或跨机架的 Virtual SAN Cluster 故障冗余。它极大地帮助规避了各种计划外停机保护能力
Storage Policies	Virtual SAN 作为一款面向对象的存储虚拟化产品，它最大的特点就是提供了 Virtual Machine 级别的存储策略。利用基于 Virtual Machine 的策略级别，可以实现不同的业务可用性支撑和不同的数据构成分布等功能。由于选用了 VM 级别作为对象参照物，存储策略区别于传统的集中存储产品，不再适用全局性的存储相关配置，取而代之的是个性化的存储策略匹配，极大地增强了混合业务模型下的不同性能、容灾等级需求
Disk Group	Disk Group (磁盘组) 是 Virtual SAN 中的新概念，也是 Virtual SAN 存储的核心概念之一。它由 SSD 与 HDD 或者全闪存构成，它是每台 Virtual SAN Cluster 节点的基础存储单元。利用 Disk Group，可以灵活针对 Virtual SAN 构成的 Datastore 进行扩容或维护

作为一款面向对象的存储虚拟化产品，Virtual SAN 中最重要的基础概念就是对象。在 Virtual SAN 中，对象涵盖的内容如表 1-3 所示。

除了对象之外，由于 Virtual SAN 是一款 Storage Policies Based 的存储，它以虚拟机对象为颗粒度参考，因此可为不同的虚拟机以及其对应的对象提供不同程度的可

用性级别，而为了满足可用性级别的工作需求，在 Virtual SAN 有一个基础概念叫作 Witness。Witness 是一个仅包含 Metadata，不包含应用数据的组件。它唯一的用途就是在故障发生后，Datastore 需要做可用性调整时，充当可用性调整时的见证用途。在 Virtual SAN 1.0 文件系统中，它大约会消耗 Metadata 2MB 的空间，Virtual SAN 2.0 中，它大约会消耗 Metadata 4MB 的空间。

表 1-3

基础概念	内容描述
VM Home Namespace	存放虚拟机配置文件，如 *.vmx, *.log, *.vmdk 等内容的根目录
VMDK	虚拟机根目录下的磁盘相关文件，包含原始磁盘和快照磁盘等
VM Swap Object	虚拟机开机后产生的交换文件
Snapshot Delta VMDK	为虚拟机创建快照时产生的相关文件
Memory Object	运行状态下，虚拟机挂起后或内存快照后产生的相关文件

1.3 Virtual SAN 与传统存储

从概念上讲，Virtual SAN 与传统存储存在很多差异，实现形式、对象处理、扩展能力等各方面，都存在着差异。

同时，Virtual SAN 的优点也很多，例如：

- ❑ Virtual SAN 不再要求外部存储，也就是说不再需要 FC、iSCSI、FCoE 之类的外部存储设备。
- ❑ 利用本地存储的特征，Virtual SAN 不再强调存储设备厂家的差异。
- ❑ Virtual SAN 是一款面向对象的存储，它不再有传统存储的 LUN、Volumes 之类的概念。
- ❑ 传统的存储协议，如 iSCSI、FCP、NFS 之类的都不再适用于 Virtual SAN。
- ❑ Virtual SAN 的部署只需要通过 vSphere Web Client 即可完成，不再需要额外的存储管理软件。
- ❑ Virtual SAN 帮助企业实现了一体化的管理员交付模式。
- ❑ Virtual SAN 利用 Storage Policies，帮助 VMs 在生成时，自动匹配自己的策略，灵活地实现了不同可用性级别与条带化级别的交付。

Virtual SAN 很好，但是当前它最大的局限就在于它只支持 vSphere 产品，这一点是它相对于其他存储产品最大的优势，也是最大的劣势。所以 Virtual SAN 不是万能的，它只能用作 Virtual SAN Cluster 节点的存储用途，无法用作其他用途。不过，有消息称，在后续的版本中会加入对第三方服务器虚拟化产品的支持及推出类似 NAS 的功能，以便让它也具备传统存储的文件共享功能。

另外，软件定义类型的存储对于空间的开销是蛮大的，所以，如果要用它，需要有这个思想准备，也就是 50% 以上的空间浪费比例。

1.4 Virtual SAN 的构建方式

针对 Virtual SAN 的构建，基础设备准备部分支持以下 3 种模式：

- ❑ Virtual SAN Ready Nodes；
- ❑ VMware EVO：RAIL；
- ❑ DIY。

1.4.1 Virtual SAN Ready Nodes

Virtual SAN Ready Nodes 这种 Virtual SAN 准备模式表示通过了 VMware 公司对应 Virtual SAN 版本兼容性测试的现存物理硬件设备，其中包含了几家国际性厂家的成型产品，包括 CISCO、DELL、IBM、HITACHI、SUPERMICRO、HUAWEI、HEWLETT PACKARD 等几家国际知名硬件设备提供商，详情请查阅：VMware Virtual SAN Ready Nodes。利用 Virtual SAN Ready Nodes，可以直接选购 List 清单中的物理服务器设备，然后按照 Virtual SAN 的部署策略标准化部署即可。

这种部署方式是官方推荐的方式，因为 Virtual San Ready Nodes 经 VMware 公司官方反复测试过，例如驱动的版本、设备型号、设备固件版本、CPU 型号等，都经过反复测试。在稳定性和可靠性上是毋庸置疑的，而且还有一个至关重要的因素，就是 VMware 公司的官方服务。选购 Virtual San Ready Nodes 如果出了问题，VMware 公司负责解决，如果不选购 Virtual San Ready Nodes，VMware 公司是不会负责的。在市场

上曾出现过这样的情况：由于驱动版本不在兼容性列表中，所以 VMware 公司判定不在服务范畴。

因此，如果条件允许，强烈建议选择 Virtual SAN Ready Nodes 来构建存储虚拟化环境。

VSAN Ready Nodes 的查询页面如图 1-1 所示。

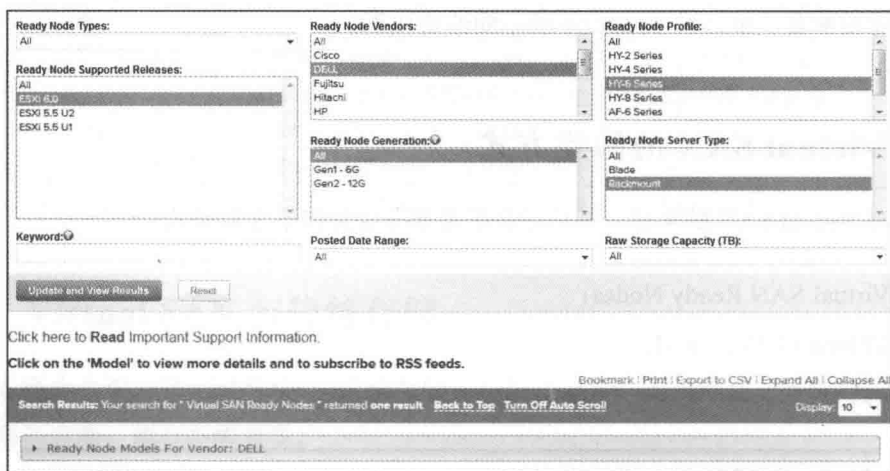


图 1-1

1.4.2 VMware EVO: RAIL

VMware EVO: RAIL 是由 VMware 公司出品的一款组合了计算虚拟化、网络虚拟化、存储虚拟化（Virtual SAN）的一体化设备，对于最终用户而言，如果选择 EVO: RAIL，只需要将该设备购买回来，即可上架使用。更多 VMware EVO: RAIL 的相关信息，请查阅地址 <http://www.vmware.com/products/evorail>。目前，全球范围之内，EVO: RAIL 一体化设备得到了包括 CISCO、DELL 等公司的支持，最初 HP 公司也支持，但截止到 2015 年 08 月 10 日，HP 已经放弃支持这个产品了。

1.4.3 DIY: Do It Yourself

DIY（Do It By Yourself）顾名思义，就是一种很草根的处理方式，同时也是多