vSAN 部署运维手册

GCR vSAN SE Team

Dec 2019

vmware[®]

dential | ©2019 VMware, Inc



| 设计 | 部署 | 运维 |
|---|--|---------------------------------|
| 两个途径 通过Ready Node 完全DIY 关键组件 | 选择较新版本并且ESXi和VC和ESXI版本匹配 确保Controller、SSD Firmware 驱动 严格符合兼容性要求 为vSAN配置专用网络并冗余,如果配 置了MTU,请确保所有节点相同 | 正确的使用习惯 确保符合兼容性 正确配置,重视告警 |
| ESXI系统引导 磁盘控制器 机械硬盘HDD 固态硬盘SSD | 如果使用SD卡部署ESXi,记得重定向 日志 确保vSAN health check 全部通过检查 符合vSAN交付检查清单 | |

Agenda

1 规划设计与最佳实践

2 部署指南

- 3 运维管理—集群操作
- 4 运维管理一故障处理



vSAN的集群要求

- vSAN 软件:
 - ESXi与vCenter版本对应,具体参考KB(2150753)
 - 生产环境建议选择update3版本
- vSAN集群需要注意:
 - 所有硬件、软件、驱动程序和固件版本满足VCG,健康检查状态全部是绿色
 - 生产环境建议至少使用4个节点,确保更高的可用性和灵活性
 - •3节点有限的恢复能力,集群无法在维护期间从节点中迁出所有数据
 - •3节点集群只能容忍单一组件故障,如果故障发生,无法重新构建组件,也不能置备新的虚拟机来容忍故障

在进行初始部署时,考虑在集群中留出容量用于未来的虚拟机部署,并留出足够的闪存缓存来容纳未来的容量增长

- 向vSAN集群添加容量时,首选方法是通过添加主机来横向扩展
- 故障域可以进一步拓展vSAN的高可用特性,在设计规模较大的集群时,建议考虑利用故障域

vSAN版本与vSphere版本对应要求

| vSAN version | ESXi version | ReleaseDate | BuildNumber | vSAN on-disk format version |
|-----------------------------|---------------------------|-------------|-------------|-----------------------------|
| vSAN 6.7 Patch 01 | ESXi 6.7 Patch 01 | 2019-12-05 | 15160138 | 10 |
| vSAN 6.7 Express Patch 13 | ESXi 6.7 Express Patch 13 | 2019-11-12 | 15018017 | 10 |
| vSAN 6.7 Update 3 | ESXi 6.7 Update 3 | 2019-08-20 | 13981272 | 10 |
| vSAN 6.7 Express Patch 10 | ESXi 6.7 Express Patch 10 | 2019-06-20 | 13981272 | 7 |
| vSAN 6.7 Express Patch 09 | ESXi 6.7 Express Patch 09 | 2019-05-14 | 13644319 | 7 |
| vSAN 6.7 Express Patch 08 | ESXi 6.7 Express Patch 08 | 2019-04-30 | 13473784 | 7 |
| vSAN 6.7 Update 2 | ESXi 6.7 Update 2 | 2019-04-11 | 13006603 | 7 |
| vSAN 6.7 Express Patch 07 | ESXi 6.7 Express Patch 07 | 2019-03-28 | 13004448 | 7 |
| vSAN 6.7 Express Patch 06 | ESXi 6.7 Express Patch 06 | 2019-01-17 | 11675023 | 7 |
| vSAN 6.7 Express Patch 05 | ESXi 6.7 Express Patch 05 | 2018-11-09 | 10764712 | 7 |
| vSAN 6.7 Update 1 | ESXi 6.7 Update 1 | 2018-10-17 | 10302608 | 7 |
| vSAN 6.7 Express Patch 4 | ESXi 6.7 Express Patch 4 | 2018-10-02 | 10176752 | 6 |
| vSAN 6.7 GA | ESXi 6.7 GA | 2018-04-17 | 8169922 | 6 |
| vSAN 6.6.1 Express Patch 18 | ESXi 6.5 Express Patch 18 | 2019-12-05 | 15177306 | 5 |
| vSAN 6.6.1 Express Patch 17 | ESXi 6.5 Express Patch 17 | 2019-11-12 | 14990892 | 5 |
| vSAN 6.6.1 Express Patch 16 | ESXi 6.5 Express Patch 16 | 2019-10-24 | 14874964 | 5 |
| vSAN 6.6.1 Express Patch 15 | ESXi 6.5 Express Patch 15 | 2019-08-20 | 14320405 | 5 |
| vSAN 6.6.1 Update 3 | ESXi 6.5 Update 3 | 2019-07-02 | 13932383 | 5 |
| vSAN 6.6.1 Express Patch 14 | ESXi 6.5 Express Patch 14 | 2019-05-14 | 13635690 | 5 |
| vSAN 6.6.1 Express Patch 13 | ESXi 6.5 Express Patch 13 | 2019-03-28 | 13004031 | 5 |
| vSAN 6.6.1 Express Patch 12 | ESXi 6.5 Express Patch 12 | 2019-01-31 | 11925212 | 5 |
| vSAN 6.6.1 Express Patch 11 | ESXi 6.5 Express Patch 11 | 2018-09-11 | 10719125 | 5 |
| vSAN 6.6.1 Express Patch 10 | ESXi 6.5 Express Patch 10 | 2018-10-23 | 10390116 | 5 |
| vSAN 6.6.1 Express Patch 9 | ESXi 6.5 Express Patch 9 | 2018-10-02 | 10175896 | 5 |
| vSAN 6.6.1 Update 2 | ESXi 6.5 U2 | 2018-05-03 | 8294253 | 5 |

新环境部署建议: vSAN 6.7 U3 旧版本升级建议: vSAN 6.5 P15

业界选择最灵活的超融合 – VMware vSAN



◆ 开始设计前,需要调研清楚业务类型特点以及对存储的容量性能需求

Confidential | ©2019 VMware, Inc.

vSAN两种部署类型



vSAN

vSAN HCL (VMware Compatibility Guide, 简称VCG)

https://www.vmware.com/resources/compatibility/search.php?deviceCategory=vSAN

| 查找的内容: | vSAN | ▼ 兼容性指南 - | 帮助 | 当前结果: 0 |
|--------|------|-----------|----|---------|
| | | | | |

需要帮助吗? 尝试使用 vSAN ReadyNode™ 配置程序。

步骤 1: 参考 "vSAN 硬件快速参考指南", 了解如何构建 vSAN ReadyNode。

步骤 2: 构建 vSAN ReadyNode:

从以下认证 vSAN ReadyNode 中选择 vSAN ReadyNode。

| vSAN ReadyNode 类型: | vSAN ReadyNode 供应商: | vSAN ReadyNode 配置文件: |
|--------------------------|----------------------------|-----------------------|
| All | All | All |
| | Cisco | HY-2 Series |
| vSAN ReadyNode 支持的版本: | DELL | HY-4 Series |
| All | Ericsson AB | HY-6 Series |
| ESXi 6.7 (vSAN 6.7) | Fujitsu | HY-8 Series |
| vSAN 6.6 (ESXi 6.5.0d) | Hewlett Packard Enterprise | AF-6 Series |
| ESXi 6.5 U1 (vSAN 6.6.1) | | |
| ESXi 6.5 (vSAN 6.5) | vSAN ReadyNode 版本: | vSAN ReadyNode 服务器类型: |
| | All | All |
| 预安装选项: | Gen1 - 6G | 1U |
| ESXi Pre-Installed | Gen2 - 12G | 20 |
| ESXi Not Pre-Installed | Gen3 - Xeon Scalable | Blade |
| | Gen3 - AMD-EPYC | Rackmount |
| | | |
| 关键字:🕡 | 发布日期范围: | 原始存储容量 (TB): |
| | 全部 | All |
| | | |
| 更新并查看结果重置 | | |

步骤 3: 要自定义 vSAN ReadyNode,请参考<u>此处</u>所述的指南

VMware 强烈建议使用认证 vSAN ReadyNode™,经验证可以为您的 vSAN 部署提供预期性能和可扩展性。如果您仍希望使用认证组件构建自己的 vSAN,请单击 基于认证组件构建自己的 vSAN。

vSAN Ready Node

步骤 2: 构建 vSAN ReadyNode:

从以下认证 vSAN ReadyNode 中选择 vSAN ReadyNode。

| ···오슈티 모== | ··CAN Deet-Nede 供应查· | ···오사 P 지문수사. | | | |
|---|---|--|-------------------|---|---|
| | All Cisco Sci | All HY2 Series HY2 Series | | HY-8-DELLEMC-R/40X0 vSAN Rea 配置文件:) 面型文件:) | dyNode 版本: Gen3 - Xeon Scalable HY-8 Series 2号・25 2 エロ |
| vSAN ReadyNode 支持的版本: All ESXI 6.7 (vSAN 6.7) vSAN 6.6 (ISXI 6.5 0c) | DELL Ericsson AB Fujitsu Hewlett Packard Enterprise | HY-4 Series HY-6 Series AF-6 Series | | | 9. 23.2 ID |
| ESXi 6.5 U1 (vSAN 6.6.1) ESXi 6.5 (vSAN 6.5) | vSAN ReadyNode 版本:� | vSAN ReadyNode 服务器类型: | 单击此处了解 vSAN Ready | Node 型号详细信息: HY-8-DELLEMC-R740xd | |
| 预安装选项: ● ESXi Pre-Installed ESXi Not Pre-Installed | All Gen1 - 6G Gen2 - 12G Gen3 - Xeon Scalable Gen3 - AMD-EPYC | All 1U 2U Blade Rackmount | 组件 SKU | 详细信息 Contact Dell Sales Rep | 数量 |
| 关键字:✔ | 发布日期范围: 全部 | 原始存储容量 (TB): ♀ | 是否已预安装 ESXi? | Yes | |
| 更新并查看结果重置 | | | 系统 | PowerEdge R740xd | 3 |
| 单击此处 查看 重要支持信息。 | | | CPU | Intel® Xeon® Gold 6126 2.6G, 12C/24T, 10.4GT/s 2UPI, 19.25M Cache, Turb DDR4-2666 | bo,HT (125W) 6 |
| 单击"型号"了解详细信息。 | | 加为书签!打印!导出为 CSV 全部展开 全部折叠 | 内存 | 32GB DDR-4, Dual Rank, 2666MHz | 36 |
| 搜索结果: 搜索" vSAN ReadyNodes " 返回 5 统 | 条结果。 <u>回到顶部 关闭自动滚动</u> | 교가: 10 📀 | 缓存层 | 800GB SSD SAS Write Intensive 12Gb 512n 2.5in Hot-plug Drive | 9 |
| ▼ 提供以下 vSAN ReadvNode 型号的供 | | | 容量层 | 1.2TB 10K RPM SAS ISE 12Gbps 2.5in | 63 |
| | | | 控制器 | HBA330 Adapter | 3 |
| | | 北重义件和谷重 | 网卡 | Intel X710 Quad Port 10Gb DA/SFP+ Ethernet | 3 |
| DELL HY-8-D | ELLEMC-R740xd | vSAN ReadyNode 版本: Gen3 - Xeon Scalable 配置文件: HY-8 Series 原始存储容量: 25.2 TB | 引导设备 | BOSS controller card + with 2 M.2 Sticks 120G (RAID 1),FH or Internal SD 16GB SD Card | Module with 2x 3 |
| | | | 支持的版本 | ESXi 6.7 (vSAN 6.7), vSAN 6.6 (ESXi 6.5.0d), ESXi 6.5 U1 (vSAN 6.6.1), E | SXi 6.0 U3 (vSAN 6.2 Update 3) |

Confidential | ©2019 VMware, Inc.

基于vSAN Ready Node进行调整

如果调整vSAN Ready Nodes,哪些可以改,哪些不改?可参考KB52084

| Components | Modifiable | Guidance |
|---------------|--------------|---|
| CPU | \checkmark | Higher core count with similar or better CPU clock speed is supported. Switching between different generation of CPU platform is not supported e.g. Romley vs Purley. |
| Memory | \checkmark | Adding more memory than what is listed is supported. |
| Caching Tier | \checkmark | Caching tier device needs to be of same or higher endurance and performance class. Follow Caching to Capacity ratio guidance discussed in this <u>blog</u>. |
| Capacity Tier | \checkmark | Capacity tier device needs to be of same or higher endurance and performance class. |
| Controller | × | Only tested, certified and listed controller configurations for that vSAN ReadyNode is supported. The firmware and driver for the controller should be an exact match as listed on the vSAN VCG. |
| NIC | \checkmark | You can add more NICs if there are available slots in the server. NIC of similar or higher configuration allowed and supported e.g. 10G with 25G or 40G. The NIC needs to be IOVP (ESXi) certified. |
| Boot Device | \checkmark | Changing boot devices is allowed. Follow guidance in this <u>KB article</u> . |

磁盘控制器 Disk Controller

storage

| 参数 | 设计及选型规则 |
|-------------------------------|---|
| 兼容性要求 (品牌 , 型号 , 驱动程 序) | 符合vSAN对Disk Controller硬件的技术兼容性要求 Firmware 驱动版本严格对齐 |
| vSAN磁盘控制卡 | RAID模式: Pass-Through最佳(或RAID-0,每块磁盘单独做一个 RAID0,但不是最佳) 队列深度:至少256 |
| | 单独用于ESXi系统磁盘(可选) 用于vSAN磁盘,同时需要注意:控制器数量 x 支持端口数 >= 磁盘数 多控制器实现更高的性能,以及将控制器故障隔离到更小范围的磁盘组 避免使用停止兼容性更新的Raid卡 |
| ESXi系统磁盘控制卡 | RAID模式:RAID1 控制器缓存:打开 |

vSAN HCL步骤: 以RAID 卡示范

- 1. 点选Search For当中的 I/O Controller选项
- 2. 点选Brand Name(本范例为DELL)
- 3. 点选Supported Releases,代表计划安装的ESXi版本(本范例为6.5)
- 4. 选择vSAN Type (有All Flash or Hybrid两个选项,本范例为Hybrid)
- 5. 最后点选Update and View Results察看结果

| Search For: | Supported Releases: | Device Type: | vSAN Type: |
|-------------------------------|---------------------|---------------|------------|
| I/O Controller | All | All | Hybrid |
| HDD <u></u> 上 耶1 | vSAN 6.6 | SAS | 上 啷 / |
| SSD JV J/K | ESXi 6.5 | SAS-RAID | 少称4 |
| | ESXi 6.0 U3 | SAS/SATA-RAID | |
| Deve d Name | ESXi 6.0 U2 | SATA | |
| Brand Name: | ESXi 6.0 U1 | | |
| All Adaptes by PMC 上顶2 | ESXi 6.0 | | |
| Adapted by PMC 少家Z | ESXi 5.5 U3 | | |
| DELL | ESXI 5.5 U2 | | |
| Fuiltsu | ESXi 5.5 01 | | |
| HP | | | |
| Huawei Technologies Co. Ltd. | | | |
| Inspur | - | | - |
| | | | |
| Keyword: | Posted Date Range: | | |
| 步骤5 | All | | |
| | | | |
| Update and View Results Reset | | | |

RAID Card的检查结果

- 6. 从系统列出的结果中,点选配置的Raid Card (本范例为H730P)
- 7. 点选之后,即可从系统所列结果,确认H730P的Firmware与Driver

| | | | | | | ESXI 5.5 02 ESXI 5.5 U1 | | | |
|------|---------------------------|---------------------|-------------------------------|--|---|--|--|-------------------|---------------------|
| DELL | PERC H730 Mini | All Flash Hybrid | Pass-Through | Device Type: SAS-RAID VID: 1000 SVID: 1028 DID: 005D SSID: 1F49 | 891,895 | vSAN 6.6 ESXi 6.5 ESXi 6.0 U3 ESXi 6.0 U2 ESXi 6.0 U1 ESXi 6.0 ESXi 5.5 U3 | | | |
| | | | 📄 vSAN 6.6 | } | lsi_mr3 version 6.9 1vmw.650.0.0.456 | 910.18.00- 4106 | 25.5.0.0018 | inbox | All Flash Hybrid |
| DELL | PERC H730P Adapter 步骤6 | All Flash Hybrid | Feature vSAN C Footnote | Category Compatible es : To prevent known issues controller: http://kb.vmwa vSAN can only support in | Features All Flash,Hybrid,Pass-Throu with vSAN compatibility, plea re.com/kb/2109665 nternal drives and cannot be u | ugh ase refer to the following used with the external ca | KB article for optimal driver pabilities of this controller | r, firmware and c | onfiguration deta |
| | | All Elect | ESXi 6.5 | | lsi_mr3 version 6.9 1vmw.650.0.0.456 | 910.18.00- 4106 | 25.5.2.0001 | inbox | All Flash Hybrid |
| DELL | PERC H730P Mini | All Flash Hybrid | Feature vSAN C | Category Compatible | Features All Flash,Hybrid,Pass-Throu | 步骤 ugh | / | | |
| | | | Footnote | es: vSAN 6.5 can only suppo | ort internal drives and cannot | be used with the externa | l capabilities of this control | ler. | |
| | | | ESXi 6.5 | | lsi_mr3 version 6.9 1vmw.650.0.0.456 | 910.18.00- 4106 | 25.5.0.0018 | inbox | All Flash Hybrid |
| | | | Feature | Category | Features | | | | |
| | | 19 VMware Inc | VSAN C | Compatible | All Flash, Hybrid, Pass-Throu | Jgh | | | |

控制器举例

华为SR 320BC,只支持混合阵列,而且只能用RAID 0方式;

华为SR 430C,在全闪存的时候,只支持直通(Pass-Through);混合阵列的时候,支持直通(Pass-Through),或者RAID-0。

| Huawei Technologies Co. Ltd. | SR 320BC | Hybrid | RAID 0 | Device Type: SAS VID: 1000 SVID: 19e5 DID: 005b SSID: d205 | 975 | ESXi 6.0 U3 ESXi 6.0 U2 ESXi 6.0 U1 ESXi 6.0 |
|------------------------------------|-------------------|---------------------|---|---|-----|--|
| Huawei Technologies Co. Ltd. | SR 420BC | Hybrid | RAID 0 | Device Type: SAS VID: 1000 SVID: 19e5 DID: 005b SSID: d206 | 975 | ESXi 6.0 U3 ESXi 6.0 U2 ESXi 6.0 U1 ESXi 6.0 |
| Huawei Technologies Co. Ltd. | SR 430C / RU 430C | All Flash Hybrid | All Flash (Pass- Through) Hybrid (Pass- Through or RAID-0) | Device Type: SAS-RAID VID: 1000 SVID: 19e5 DID: 005d SSID: d207 | 895 | ESXi 6.5 U1 vSAN 6.6 ESXi 6.5 ESXi 6.0 U3 ESXi 6.0 U2 ESXi 6.0 U1 ESXi 6.0 |

浪潮 Inspur 2308IR,只支持混合阵列,不支持RAID方式,只支持到ESXi 6.0 U3版本,目前不支持vSAN 6.5和6.6

| 品牌名称 | 型号 | 类型 | 功能 | 产品描述 | 队列深度 | 支持的版本 |
|--------|---------------|--------|--------------|--|------|--|
| Inspur | Inspur 2308IR | Hybrid | Pass-Through | Device Type: SAS VID: 1000 SVID: 1BD4 DID: 0087 SSID: 0009 | 600 | ESXi 6.0 U3 ESXi 6.0 U2 ESXi 6.0 U1 ESXi 6.0 ESXi 5.5 U3 ESXi 5.5 U2 ESXi 5.5 U1 |

推荐方式:采用HBA卡,无需RAID功能 常用型号如下

| 厂商 | HBA卡型号(注意与主机型号的适配) | 直通模式 | RAID支持程度 |
|---------|--|------|----------|
| DELLEMC | HBA330/HBA330+/HBA330 mini | 支持 | No |
| HPE | H240/H240ar(须禁用 HP SSD 智能路径功能,KB2092190) | 支持 | 0/1/10/5 |
| Lenovo | ThinkSystem 430-8i/430-8i Dense/430-16i | 支持 | No |
| Huawei | SR 150-M/SP 150-M/SR 760-M | 支持 | 0/1/10 |

设计vSAN集群的最佳性能-硬件

硬件是vSAN性能最重要的因素



硬件必须在vSAN VMware兼容性指南(vSAN VCG) 上

使用经认证最新的固件和驱动程序

建议使用SAS接口SSD设备

- SATA SSD较慢的总线速度
- SATA SSD具有较小的队列深度

建议选择PCIe、NVMe闪存设备,无需控制器

使用HBA代替RAID控制器,更经济、更简单 •HBA无需禁用高级功能

推荐最少每台主机配置2个磁盘组,一般2-4个磁盘 组

将BIOS配置文件设置为OS Controllered, High performance带来更好的整体性能



| 工作负载定义 队列深度: 16 个或更少 传输长度: 4 KB 操作: 写入 模式: 100% 随机 延迟: 不到 5 毫秒 |
|--|
| 耐久性 • SanDisk的800 GB, LT0800RO – Class C >=3650 TBW • HGST 的 HUSMH8080BSS200 – Class D >=7300 TBW |
| |

缓存容量至少为可用存储容量的 10% 10%是参考值,容量决定于热点数据的多少,通过查看RC Hit Rate可判断缓存是否足够 如果需要可调整为15%

WWARE CONFIDENTIAL I CONFIDENTIAL

17 17

检查磁盘参数和适用性示例如下

型号详细信息 1.92TB Solid State Drive SATA Read 合作伙伴名称: DELL 型든: Intensive 2.5in, PM863a 设备类型: SATA 供应商 ID: ATA 耐用性: 3504 TBW 生能等级: Class D: 20,000-30,000 writes per second 耐用等级: N/A 闪存技术: TLC 部件号: 09W12R 接口速度: 6 Gbps 外形规格: 2.5" 容量: 1920 GB 系列: PM863a 产品 ID: MZ7LM1T9HMJP0D3 注: rss feed 🔝 版本详细信息 全部展开 | 全部折叠 层 版本 最低固件版本 vSAN Hybrid Caching Tier ESXi 6.7 U2 (vSAN 6.7 Update 2) GC55 vSAN All Flash Capacity Tier 功能类别 功能 Mandatory Drive Performance, Drive Reliability, Queue Depth, Surprise Power Removal Protection, Write Cache, Write Failure Notification vSAN Hybrid Caching Tier GC55 ESXi 6.7 U1 (vSAN 6.7 Update 1) vSAN All Flash Capacity Tier 功能类别 功能 Mandatory Drive Performance, Drive Reliability, Queue Depth, Surprise Power Removal Protection, Write Cache, Write Failure Notification ESXi 6.7 (vSAN 6.7) vSAN Hybrid Caching Tier GC55 vSAN All Flash Capacity Tier 功能类别 功能 Mandatory Drive Performance, Drive Reliability, Queue Depth, Surprise Power Removal Protection, Write Cache, Write Failure Notification vSAN 6.6 (ESXi 6.5.0d) GC55 vSAN Hybrid Caching Tier vSAN All Flash Capacity Tier

Mware[®]

Confidential | ©2019 VMware, Inc.

功能类别 功能

机械硬盘HDD

| 参数 | 设计及选型规则 |
|--------------|---|
| 兼容性要求 | 符合vSAN对HDD硬件的技术兼容性要求 |
| (品牌,型号,驱动程序) | • 要求Partner Number和Product ID都符合兼容列表 |
| | • HDD 微码不低于兼容版本,SSD微码与兼容性要求一致 |
| 接口类型 | • 使用 SAS 和 NL-SAS 可以获得最佳结果 |
| | • 大容量磁盘应当只在以容量为中心且不优先考虑性能的环境里使用 |
| | • vSAN 6.0开始,不支持SATA接口的HDD |
| | • vSAN 6.5 版本之后,开始支持 512e 大容量盘, vSAN 6.7支持4kn |
| | |
| HDD数量/服务器 | • 做磁盘组容量使用:最多35个HDD(每磁盘组最多7个HDD) |

• 做系统盘使用: 通常2个

vSAN引导设备

- 支持USB、SD卡、M.2、SATADOM、SSD/HDD作为ESXi引导设备,
- 对于内存占用很小的小规模部署,可采用USB/SD,需要额外配置远程服务器重定向log、traces和coredump(使用syslog服务器和net dump Collector),请勿将日志信息存储在vSAN数据存储上
- SSD/HDD缺失了一个磁盘槽位,还需要将引导设备磁盘放在单独的控制器后面(LSI3108系列不能有本地datastore)
- SATADOM基于SLC技术,提高启动设备的成本,且如果设备失败,需要立即替换它(单一模式)
- M.2可以是一个或两个设备,双设备的优点:镜像(硬件RAID),为引导设备提供了额外的冗余,这是SATADOM设备所缺乏的。它们通常连接到板载AHCI控制器,这意味着在vSAN数据存储和引导设备之间不需要额外的控制器

| 设备 | 价格 | 大小 | 存放log | 存放vSAN traces | 存放coredump | 备注 |
|-----------------|--------|------------------|-------|------------------------|-----------------------|--|
| USB/SD | \$ | 最小: 4G 推荐: 8G | Ν | Y 系统重启保存到SD卡的分 区 | Y,ESXi主机内存 <=512GB | 持久度过低,日志和trace 的写入量大易烧毁 设置syslog server&Network dump collector重定向 |
| SSD | \$\$ | 最小: 30GB | Y | Y | Y | 占用磁盘槽位 |
| HDD | \$\$ | 最小: 30GB | Y | Y | Y | 占用磁盘槽位 |
| SATA-DOM | \$\$\$ | 最小: 16GB | Y | Y | Y | 成本高 |
| M.2 SSD (推荐) | \$\$ | 最小: 30GB | Y | Y | Y | 1. 镜像M.2提供冗余性 2. vSANDatastore和引导设备 控制器分离 |

vSAN 网络适配器选用

vSAN 太网适配器:

- 对于全闪配置,建议专用带宽为10 Gbps
- 对于混合配置,专用或共享带宽为 10 Gbps
- 强烈推荐为vSAN配置两个万兆网络冗余
- 如果与其他流量共享,使用 Network I/O Control 保证vSAN流量份额

网络模式:

- vSAN 6.6以前需要物理交机和路由器上启用多播

- 纯vSAN6.6 环境则无需启用多播

巨型帧(MTU 9000):

- 在大型部署环境中开启巨型帧可以减少CPU的开销

- 针对全新部署启用

Confidential | ©2019 VMware, Inc.

vSAN网络最佳实践

Since vSAN is a distributed system, network can greatly impact performance



链路聚合(LACP)应该会有所帮助,特别是在跨网络的负载平衡方面,但是它会带来复杂性 对于共享单个上行链路的服务,使用网络IO控 制(NIOC)

巨型帧可能(也可能不)提供2-15%的改进

不要忽视驱动程序/固件

确保高质量的连接

更高的带宽通常提供更低的延迟

配件自组总结: 兼容条件应严格遵守,以下"例外"也应仔细审核: 以vSAN ReadyNode为基准,有条件地选择组件

- CPU: 可以使用核数更多和频率更高的,但不能跨微架构(Grantley -> Purley)
- Memory: 可选用更大容量的内存, 内存频率没有强制限制
- SSD/HDD:
 - 可选用性能更好的SSD;对HDD来说意味着接口速度(6Gbps/12Gbps)可以更好,转速可以更高,但不能切换接口协议种类(SAS换成SATA,或者反过来)。
 - 可以使用同系列不同容量的盘(更大的Cache盘,更大或者更小的Capacity盘), 但盘需要被vSAN HCL认证过
 - 可以互换512e 和512n的磁盘
- 控制器:
 - 可以添加更多的同款的控制器,以机箱规格为限
 - SAS/SATA 扩展器(Expander): 不能在vSAN ReadyNode范围以外添加或者删除服务器的Expander
- Boot 设备(也即ESXi启动盘)
 - vSAN不控制Boot设备,可以使用任何满足KB的设备
 - 仅用于安装OS的硬盘及控制器,不需要符合HCL或ReadyNode
- NIC(也即网卡): vSAN不控制NIC,可以使用任何vSphere认证过的NIC

Confidential | ©2019 VMware, Inc.

| 系统 | 型号: ThinkAgile VX 2U Node | 系统类型: Rackmount | 1 |
|------|--|---|--|
| CPU | 可以使用核数更多和频率更高的 125W 2.6GHz (Skylake-SP) | CPU,但不能跨微架构 6126 12C | 2 |
| 内存 | 可选用更大容量的内存,内存频 | 率没有强制限制 | 12 |
| 缓存层 | # 可选用HCL中性能更好的SSD 可以更好,转速可以更高,但不 来)。 ^b Hot Swap SSD 设备类型: SAS 性能等级: Class F: 100,000+ writes per second # 可以使用HCL中同系列不同容 | ,对HDD来说意味着接口速度(能切换接口协议种类(SAS换成S 容量: 400 GB TBW 耐用等级: Endurance Class D >=7300 TBW 量的盘(更大的Cache盘,更大算 | 6Gbps/12Gbps) SATA,或者反过 2 或者更小的 |
| 容量层 | C <u>apacity盘)</u> 2.5" 1.2TB 10K SAS 12Gb # 可以互换512e 和512n的磁盘 设备类型: SAS | 合作伙伴名称: Lenovo 容量: 1200 GB | 12 |
| 控制器 | # 可以添加更多的" <u>同款</u> "的控制 # SAS/SATA 扩展器(Expand 或者删除服务器的Expander | 器,以机箱规格为限 er): 不能在vSAN ReadyNod | e范围以外添加 |
| 网卡 | vSAN不控制NIC,可以使用任何 | 可vSphere认证过的NIC ^{SFP+} | 1 |
| 引导设备 | # vSAN不控制Boot设备,可以 # 仅用于安装OS的硬盘及控制器 | 使用任何满足KB的设备 器,不需要符合vSAN HCL或Rea | dyNode |

部署实施vSAN的最佳实践要点 微信关注Vmware中国>VMware微站>资料下载>产品技术支持

检查点:

在需要高I/O性能的场合,将主机电源管理设置为服务器BIOS中的"OS Controlled"。通过在vSphere Client中检查主机的电源管理,验证 设置是否已生效:应该显示为APCI P-状态和C-状态,并且主动策略应该显示"高性能"。

确认禁用RAID卡上的写缓存

确认用于vSAN的硬盘和不用于vSAN的硬盘使用相同的控制器模式,建议都为"直通"模式

确认用于启动的硬盘不要与vSAN硬盘共用RAID卡;如果一定要共用,请参照上一条,设为直通或RAID-0

确认主机系统日志没有存放在vsanDatastore

确认vSAN traces没有存放在vsanDatastore

确认没有VM运行在本地Datastore

确认vSAN HealthCheck是否有遗留告警

vCenter不建议放在vSAN节点上

确认所有网络接口MTU是否一致,建议为9000

在vSAN集群上部署虚拟机的建议

- 1. 如果有可能,请使用多个VMDK。许多应用程序(如Microsoft SQL, Exchange, SharePoint)都受益于多个存储设备的使用。这可以改善VM的性能。
 - 多个虚拟SCSI适配器。当使用多个VMDK的VM时,请使用多个虚拟SCSI适配器,VMDK最多分散在4个虚拟适配器 上。客户操作系统通过磁盘控制器执行基于OS的排队。更多控制器将减少争用并提高性能。"。
 - 使用准虚拟SCSI适配器。配置VMDK以使用VMware准虚拟SCSI(PVSCSI)适配器而不是传统的模拟虚拟SCSI适配器可以提供更好的性能以及更低的CPU利用率。传统虚拟SCSI适配器的队列深度相对较小,为32,这可能会妨碍客户操作系统将数据传递到存储堆栈中的下一层的能力。PVSCSI适配器默认为队列深度为64。有关详细信息,请参阅知 识库文章1010398。
 - 如果需要,增加PVSCSI控制器的队列深度。某些应用程序和工作负载(如OLTP数据库活动)会生成高度过渡的序列化I/O。一个或多个PVSCSI控制器的默认队列深度可能不足,可以通过增加队列深度来改进。有关更多详细信息,请参阅知识库文章2053145。
- 2. 重新审视分配给VM的RAM。某些应用程序(如SQL Server)可以从分配的更大内存大小中获益。这可以潜在地减少 生成的I/O量,因为这些应用程序可以在VM内使用I/O的缓存,缓冲和合并来提高I/O处理效率。
- 3. 重新审视VM的虚拟CPU分配。虚拟CPU分配给VM的太低或太高,两者都可能会对性能产生负面影响。

为虚拟磁盘采用PVSCSI KB1010398

| DC02 - Edit Settings | | | | | | |
|--------------------------------|---------------------------|--------|--|--|--|--|
| Virtual Hardware VM Option | SDRS Rules VApp Options | | | | | |
| F CPU | 1 🔹 🖬 | | | | | |
| ▶ 🌆 Memory | 4096 * MB * | | | | | |
| ▶ 🛄 Hard disk 1 | 20 GB V | | | | | |
| ▶ 🛄 Hard disk 2 | 40 GB V | | | | | |
| ▶ 🛃 SCSI controller 0 | LSI Logic SAS | | | | | |
| Metwork adapter 1 | VM Network 🗸 🗸 Connected | | | | | |
| ▶ 🝥 CD/DVD drive 1 | Datastore ISO File | | | | | |
| ▶ 📃 Video card | Specify custom settings | | | | | |
| ▶ 🎲 VMCI device | | | | | | |
| ▶ Other Devices | | | | | | |
| ✓ ④ New SCSI controller | VMware Paravirtual | | | | | |
| 2 SCSI Bus Sharing | None | | | | | |
| → Change Type | VMware Paravirtual 🗸 | | | | | |
| | | | | | | |
| New device: | SCSI Controller - Add | | | | | |
| Compatibility: ESXi 5.5 and la | ter (VM version 10) OK | Cancel | | | | |

PVSCSI 适配器是高性能的存储适配器,可增加吞吐量,同时降低 CPU 使用率。

PVSCSI 适配器最适合硬件或应用程序将使 I/O 吞 吐量达到非常高的环境,尤其是 SAN 环境。

查询KB确认支持的版本和OS

vmware[®]



| | 最小值 | 最大值 | 默认值 | 说明 |
|--------------------|-----|------|-----|---|
| 每个对象的条带宽度 | 1 | 12 | 1 | vSAN可以决定某个对象可能需要跨多个磁盘条带化,通常管理员请求创建的 VMDK过大,无法在单个物理驱动器中容纳。vSAN上最大组件的大小为 255GB。对于大于255GB的对象,vSAN自动将它们划分为多个组件。因此, 如果管理员部署2TB的VMDK,可以在构成该VMDK对象的相同RAID-O条带 配置中看到8个或更多的组件 |
| 允许的故障数FTT/SFTT | 0 | 3 | 1 | 定义存储对象能容忍的主机、 磁盘或网络故障的数量。若要容许n个故障,则 集群中需要有2n+1个主机。如果故障域是设计的一部分,则集群中需要 有2n+1个故障域才能容纳vSAN集群中的n个故障。 容许3个故障是RAID-1镜像容错方法支持的最大数量。如果使用RAID-5/6纠 删码,则支持容许的最大故障数为2。 在延伸集群中,这将是站点内使用的保护措施。 |
| 站点之间的允许故障数 PFTT | 0 | 1 | 0 | 确定是否在vSAN延伸集群中在两个站点之间对数据进行镜像处理。 |
| 闪存读缓存预留 | 0% | 100% | 0% | 仅适用于混合配置,配置了在SSD上为存储对象读取缓存预留的闪存容量的大小 |
| 对象空间预留 | 0% | 100% | 0% | 调配虚拟机时要预留(厚配置)的存储对象的逻辑大小的百分比,其余存储对象精简配置。 |

Agenda

1 规划设计与最佳实践

2 部署指南

- 3 运维管理—集群操作
- 4 运维管理—故障处理



简化Day 0,1和2运维





配置



性能监控



Day O – 智能、自动化vSAN简易安装



Confidential ©2019 VMware, Inc.

Mware[®]

- 提供在单一主机上引导新的vCenter Server(VCSA)
 安装vSphere和配置vSAN数据存储的简易方式
- vSAN配置步骤内嵌于VCSA安装向导
- 简化构建新环境的过程

▶ vSAN 版本检查

Raid卡检查

缓存层SSD检查

容量层磁盘检查

| vSAN version | ESXi version | ReleaseDate | BuildNumber | vSAN on-disk format version |
|-----------------------------|---------------------------|-------------|-------------|-----------------------------|
| vSAN 6.7 Patch 01 | ESXi 6.7 Patch 01 | 2019-12-05 | 15160138 | 10 |
| vSAN 6.7 Express Patch 13 | ESXi 6.7 Express Patch 13 | 2019-11-12 | 15018017 | 10 |
| vSAN 6.7 Update 3 | ESXi 6.7 Update 3 | 2019-08-20 | 13981272 | 10 |
| vSAN 6.7 Express Patch 10 | ESXi 6.7 Express Patch 10 | 2019-06-20 | 13981272 | 7 |
| vSAN 6.7 Express Patch 09 | ESXi 6.7 Express Patch 09 | 2019-05-14 | 13644319 | 7 |
| vSAN 6.7 Express Patch 08 | ESXi 6.7 Express Patch 08 | 2019-04-30 | 13473784 | 7 |
| vSAN 6.7 Update 2 | ESXi 6.7 Update 2 | 2019-04-11 | 13006603 | 7 |
| vSAN 6.7 Express Patch 07 | ESXi 6.7 Express Patch 07 | 2019-03-28 | 13004448 | 7 |
| vSAN 6.7 Express Patch 06 | ESXi 6.7 Express Patch 06 | 2019-01-17 | 11675023 | 7 |
| vSAN 6.7 Express Patch 05 | ESXi 6.7 Express Patch 05 | 2018-11-09 | 10764712 | 7 |
| vSAN 6.7 Update 1 | ESXi 6.7 Update 1 | 2018-10-17 | 10302608 | 7 |
| vSAN 6.7 Express Patch 4 | ESXi 6.7 Express Patch 4 | 2018-10-02 | 10176752 | 6 |
| vSAN 6.7 GA | ESXi 6.7 GA | 2018-04-17 | 8169922 | 6 |
| vSAN 6.6.1 Express Patch 18 | ESXi 6.5 Express Patch 18 | 2019-12-05 | 15177306 | 5 |
| vSAN 6.6.1 Express Patch 17 | ESXi 6.5 Express Patch 17 | 2019-11-12 | 14990892 | 5 |
| vSAN 6.6.1 Express Patch 16 | ESXi 6.5 Express Patch 16 | 2019-10-24 | 14874964 | 5 |
| vSAN 6.6.1 Express Patch 15 | ESXi 6.5 Express Patch 15 | 2019-08-20 | 14320405 | 5 |
| vSAN 6.6.1 Update 3 | ESXi 6.5 Update 3 | 2019-07-02 | 13932383 | 5 |
| vSAN 6.6.1 Express Patch 14 | ESXi 6.5 Express Patch 14 | 2019-05-14 | 13635690 | 5 |
| vSAN 6.6.1 Express Patch 13 | ESXi 6.5 Express Patch 13 | 2019-03-28 | 13004031 | 5 |
| vSAN 6.6.1 Express Patch 12 | ESXi 6.5 Express Patch 12 | 2019-01-31 | 11925212 | 5 |
| vSAN 6.6.1 Express Patch 11 | ESXi 6.5 Express Patch 11 | 2018-09-11 | 10719125 | 5 |
| vSAN 6.6.1 Express Patch 10 | ESXi 6.5 Express Patch 10 | 2018-10-23 | 10390116 | 5 |
| vSAN 6.6.1 Express Patch 9 | ESXi 6.5 Express Patch 9 | 2018-10-02 | 10175896 | 5 |
| vSAN 6.6.1 Update 2 | ESXi 6.5 U2 | 2018-05-03 | 8294253 | 5 |

新环境部署建议: vSAN 6.7 U3 旧版本升级建议: vSAN 6.5 P15

Confidential | ©2019 VMware, Inc.

| ▶ vSAN 版ā | 本检查 | 软硬件配置检查 vSphere兼容性检查 Raid卡检查 缓存层SSD检查 容量层磁盘检查 ESXi日志配置检 |
|-----------|-----|---|
| | | 详细配置(Raid卡,缓存层SSD,容量层磁盘需要提供具体的Product ID和 结果 备注 Module ID) |
| | 1 | 服务器配置 |
| | 2 | 交换机配置 |
| | 3 | vSphere 版本: |
| | 4 | vCenter版本: |
| | 5 | 集群多少台主机 |
| | 6 | 是否为全闪环境 |
| | 7 | 是否为ROBO集群 / 延伸集群 |
| | 8 | 是否为Ready Node |
| | 9 | 每台服务器几个Raid卡 |
| | 10 | 每台主机几个磁盘组 |
| | 11 | 磁盘组配置 |
| | 12 | 缓存层SSD信息 |

| vSAN 版本检查 | 软硬件配置检查 | vSphere兼容性检查 | Raid卡检查 | 缓存层SSD检查 | 容量层磁盘检查 | ESXi日志配置检查 |
|-----------|---------|--------------|---------|----------|---------|------------|
|-----------|---------|--------------|---------|----------|---------|------------|

| | 服务器满足 | LvSphere兼容性 | 结果 | 备注 |
|---|-------|--------------------------------|----|----|
| 1 | 检查方法: | 前往VMware <u>vSphere兼容网</u> 站检查 | | |
| 2 | 检查结果: | 服务器兼容截图 | | |
| 3 | 检查结果: | RAID卡兼容截图 | | |
| 4 | 检查结果: | 网卡兼容截图 | | |

容量层磁盘检查

| | 要求Raid卡符合vSAN兼容性,固件版本和驱动严格对齐 | 结果 | 备注 |
|---|--------------------------------|----|----|
| 1 | 到 <u>vSAN兼容性</u> 网站检查当前Raid卡型号 | | |
| 2 | 检查结果: | | |
| 3 | 检查结果:当前vSAN配置(混合 全闪) | | |
| 4 | 检查结果: 当前Raid型号 | | |
| 5 | Raid卡的firmware和驱动版本要求对齐 | | |
| 6 | Raid 卡的firmware请前往服务器BIOS查找 | | |
| 7 | 检查结果: vSAN兼容性网站截图: | | |
| 8 | 检查结果:当前的Raid卡固件版本,截图 | | |
| 9 | 检查结果:当前的驱动版本,截图: | | |

vSAN 版本检查 软硬件配置检查 vSphere兼容性检查 Raid卡检查 缓存层SSD检查 ESXi日志配置检查 容量层磁盘检查 SSD 符合vSAN兼容性要求,性能满足Class C及以上, 结果 备注 如果普通SSD, Firmware 满足兼容性要或高于都可以 如果为NVMe SSD, 要求驱动和firmware严格对齐, 不同的firmware 对应不同 的驱动 1 到vSAN兼容性网站检查当前SSD型号和firmware SSD firmware 请到BIOS查看 2 3 驱动检查 vmkload mod –s/esxcli system module get -m driver name firmware 版本截图: 4 5 驱动版本截图: (针对需要驱动程序的SSD,如NVMe / PCI-E接口的SSD) Raid 卡的firmware请前往服务器BIOS查找 6 7 检查结果: vSAN兼容性网站截图:

Confidential | ©2019 VMware, Inc.


| | 要求容量层磁盘符合vSAN兼容性要求,firmware不低于兼容性网站要求 | 结果 | 备注 |
|---|---------------------------------------|----|----|
| 1 | 到 <u>vSAN兼容性</u> 网站检查容量层磁盘型号和firmware | | |
| 2 | 检查结果: vSAN兼容性网站截图: | | |

| vSAN 版本检查 | 软硬件配置检查 | vSphere兼容性检查 | Raid卡检查 | 缓存层SSD检查 | 容量层磁盘检查 | ESXi日志配置检查 |
|-----------|---------|--------------|---------|----------|---------|------------|
|-----------|---------|--------------|---------|----------|---------|------------|

| | 检查项目 | 结果 | 检查方法 |
|---|------------------------------------|----|--------------------------------|
| 1 | 配置ESXi主机日志重定向(SD/USB作为启动设备时需要) | | vSAN集群和主机摘要无黄色有关主 机日志位置告警 |
| 2 | 配置vSAN trace日志重定向(SD/USB作为启动设备时需要) | | vSAN集群和主机摘要无黄色有关 vSAN日志位置告警 |

部署最佳实践: 使用可靠的硬件



Confidential | ©2019 VMware, Inc.

部署最佳实践: 使用高效的软件





部署最佳实践: 使用高效的软件





部署最佳实践:进行规范的操作



Confidential | ©2019 VMware, Inc.

部署最佳实践:进行规范的操作





Agenda

- 1 规划设计与最佳实践
- 2 部署指南
- 3 运维管理-集群操作
- 4 运维管理—故障诊断



确保健康检查为绿色

| 📆 VSAN-Cluster 📲 📴 😤 🗁 🎯 Actions 🗸 📃 | | | | | | | | | | |
|---|--|---|--------------|-------------------------|--|--|--|--|--|--|
| Getting Started Summary M | onitor Configure | Permissions Hosts VMs | Datastores 1 | Networks Update Manager | | | | | | |
| Issues Performance Tasks & 44 Health Capacity Resyncing Components Virtual Objects Physical Disks | K Events Profile Con vSAN Health (Last Test Result Warning Warning Passed Passed Passed | Appliance Resource Reservation Test Name > Hardware compatibility > Performance service > Network > Physical disk | VSAN Utiliz | ation | t with Online health Retest | | | | | |
| iSC SI Targets | Passed | ▶ Data | | Virtual SAN Health (I | ast chocked: Today at 8:25 M | | | | | |
| Proactive Tests | | Ciuster | _ | virtual SAN lieatui (t | Last checked. Today at 0.25 A | | | | | |
| | | | | Test Result | Test Name | | | | | |
| | | | | 😣 Failed | Cluster health | | | | | |
| | | | Select a tes | 😣 Failed | Limits health | | | | | |
| | | | ocieve a tes | 🔥 Warning | Virtual SAN HCL health | | | | | |
| | | | | Passed | ▶ Data health | | | | | |
| | | | | Passed | Network health | | | | | |
| | | | | Passed | Physical disk health | | | | | |
| | | | | | | | | | | |

vSAN健康检查工具包括:

- 硬件兼容性
 - 控制器Firmware
- 网络
- 物理磁盘
- 数据/对象
- 集群状态
- 限制
- 性能服务

检测到问题,立即在UI上显示警告

监控vSAN运行状况

| vm vSphere Client | Menu 🗸 🛛 📿 Search | | C | ?~ | Administrator@VSPHERE.LOCAL 🗸 | ٢ |
|---|---|---|------|----------------|-------------------------------|---|
| Sec-rdops-vm08-dhcp-243-1 Data Center VSAN cluster 10.160.225.54 10.160.244.18 10.160.245.105 10.160.254.39 | Summary Monitor Summary Monitor Issues and Alarms All Issues Triggered Alarms Insues Triggered Alarms Advanced Performance Overview Advanced Advanced Tasks and Events Tasks Events Scheduled Tasks Scheduled Tasks Resource Allocation CPU Memory Storage Utilization Storage Overview Security VSAN Health Virtual Objects Physical Disks Resyncing Compone. Proactive Tests Proactive Tests Capacity | ACTIONS ~ onfigure Permissions Hosts VMs Datastores Health (Last checked: Jan 29, 2018 2:52:43 PM) • A Hardware compatibility • vSAN HCL DB up-to-date • vSAN HCL DB Auto Update • vSAN HCL DB Auto Update • vSAN HCL DB Auto Update • SCSI controller is VMware certified • Controller is VMware certified • Controller driver • VMWare certified • Controller driver • Performance service • VSAN Build Recommendation • Conline health (Last check: 1 hour(s) ago) | Netw | orks RETEST | WITH ONLINE HEALTH RETEST | |



| vm vSphere Client | Menu 🗸 🛛 📿 Search | | C 🤇 🗸 Administrator@VSPHERE.LOCAL 🗸 |
|---|---|---|--|
| vm vSphere Client □ □ □ ○ > □ □ ○ ○ > □ □ □ ○ ○ > □ | Menu V Q Search VSAN Cluster Summary Monitor Issues and Alarms All Issues Triggered Alarms Performance Overview Advanced Tasks and Events Tasks Events Resource Allocation CPU Memory Storage Utilization | ACTIONS ~ Configure Permissions Hosts VMs Datastores CAPACITY USAGE CAPACITY HISTORY Capacity Overview Capacity Overview Used - Physically written 7.02 GB Used - VM overreserved 3.50 GB Used - Total 10.52 GB Used - Total 10.52 GB Free 522.39 GB Used Capacity Breakdown | Networks Updates Deduplication and Compression Overview Deduplication and compression are disabled |
| | Storage Overview Security VSAN Health Virtual Objects Physical Disks Resyncing Objects Proactive Tests Capacity Performance Performance diagno | Breakdown of the used capacity. Group by: Object types Object types Object | ************************************** |
| | | Checksum overhead | 68.00 MB (1%) |

容量概览显示已用空间和可用空间 情况

重复数据消除和压缩的节省空间

按照对象类型分类

- 文件系统开销
- 重复数据消除和压缩开销
- 校验和开销
- 虚拟磁盘
- 交换对象
- 虚拟机主目录对象





存储性能指标集成到vCenter

一键启动,默认情况下性能服务处于启动状态

健康检查监控状态

分布式架构

简化指标

- 前端指标 (VM)
- 后端指标(vSAN)

历史数据

故障场景处理

vSAN识别两种不同类型的硬件设备事件来定义故障场景

- Absent不可访问,但是无明确错误代码
 - vSAN认为数据短暂不可用
 - 默认60分钟后, vSAN启动组件重新恢复操作
- Degraded不可访问,有错误代码
 - vSAN认为数据副本永久丢失
 - 立即进行组件重新恢复

以下情况检测到IO错误, 被认定为Degraded

- 机械硬盘故障
- 闪存设备故障
- 存储控制器故障

以下情况检测到IO错误, 被认定为Absent

- 网络故障
- 网卡故障
- 主机故障

拔出磁盘引起Absent事件,若磁盘在60分钟内插回vSAN集群,则不会进行组件重构 MWAIC[®] confidential | ©2019 VMware, Inc.

组件重建



- 首先恢复IO
 - 检测到故障并重新计算对象的可访问性
 - 如果相关对象为健康(所有可用投票数 >50%),则IO恢复,整个检查过程大概为 5-7秒

●重建

- 重建的目的是使得对象重新获得由SPBM 策略定义的合规性保护
- Absent: 等待60分钟组件重建, 重建完成后删除Absent组件
- 若60分钟内恢复正常,可以避免不必要的 重建,仅同步变更数据
- Degraded: 立即开始重建
- 重建完成后
 - Stale状态,并不是直接变成可用状态,
 因为它上面的数据不是最新的,需要从其他副本同步

有效的修复实现智能化重建



● 60分钟后缺失组件重构有两种修复方式

• 构建全新副本

• 更新重新上线的现有副本

- 评估组件重新上线的两种修复方式哪种更高效
- 选择最高效的方式恢复冗余,并取消另一个操作
- 明显提高对象修复操作的速度和效率,降低风险并 最大限度减少资源使用量



- 修复过程更加灵活
- 即使资源不足,无法符合存 储合规性策略,也会尽可能 修复多的组件
- 当有足够的资源可用时会继<
 续修复其余组件

●适用于vSAN集群和延伸集群



- 修复过程更加灵活
- 即使资源不足,无法符合存 储合规性策略,也会尽可能 修复多的组件
- 当有足够的资源可用时会继<
 续修复其余组件

●适用于vSAN集群和延伸集群





- 将大型组件拆分为较小组件来实现更好的平衡,从而提高效率
- 高效的重建增强功能提高对象修复操作的速度和效率
- 在健康服务的重平衡状态可视化增强
- 引入调整重新同步操作吞吐量的选项
 - 新的调整机制可以使得管理员轻松控制重新同步流量带宽
 - 通过GUI在vSAN集群端设置限制所有重新同步操作的吞吐量
 - 在业务高峰期更好的处理主机故障或轮滚维护场景

计划内停机——维护模式

vSAN节点重启/关机

• 一定要进入维护模式,即使主机上已经没有运行的虚拟机

3种维护模式选项:

- ●确保可以访问:仅从主机迁移对运行虚拟机相关的组件(确保大于50%的组件可用)
 - 处于此模式时,如果遇到故障,虚拟机的可用性将受影响
 - 选择确保数据可访问性模式后,在故障期间不会重新保护数据,可能会遇到意外丢失数据的情况
 确保数据可访问性模式比撤出全部数据模式更快

●迁移所有数据

- 处于此模式时, 主机中的所有组件都会迁移, 并且根据主机上的数据量, 迁移可能需要较长时间
- 使用撤出全部数据模式时,即使在计划维护期间,虚拟机也可允许故障
- 若无足够资源空间,无法选择迁移所有数据选项
- ●无数据迁移(不推荐)

扩展vSAN集群

横向扩展

- ●UI界面上选择'add host'将新主机添加到集群中
- ●完成添加新主机后
 - 手动模式, 手动在新添加的主机上创建磁盘组
 - 自动模式,根据主机上磁盘设备的实际情况,自动创建磁盘组

纵向扩展

- 增加容量: 在集群中的主机上添加额外的或更大的容量盘
- ●改变缓存性能:在集群中的主机上添加额外的或更大容量的SSD
- ●同时改善缓存性能并增加容量:在集群中添加额外的磁盘组
 - vSAN集群下"配置"->"vSAN"下,单击磁盘管理,选择磁盘组,然后单击向选定的磁盘组添加磁盘图标
 - 选择添加的设备并单击确定
 - 若添加包含残留数据或分区信息的赢设备,则必须先清理设备

如何开启/关闭vSAN节点

vCenter不在vSAN中

关机顺序:

- 1. 关闭所有存储在vSAN Datastore上的所有VM
- 2. 所有主机置于维护模式
- 3. 确保没有勾选"move powered off and suspend virtual machines to other hosts in the cluster"勾选"No Data Migration",不需要迁移任何数据
- 4. 关闭所有主机

开机顺序:

- 1. 打开所有主机
- 2. 把所有主机退出维护模式
- 3. 开启存放在vSAN Datastore的所有VM

Confirm Maintenance Mode

A host in maintenance mode does not perform any activities on virtual machines, including virtual machine provisioning. The host configuration is still enabled. The Enter Maintenance Mode task does not complete until the above state is completed. You might need to either power off or migrate the virtual machines from the host manually. You can cancel the Enter Maintenance Mode task at any time.

There are hosts in a Virtual SAN cluster. Once the hosts are removed from the cluster, they will not have access to the Virtual SAN datastore and the state of any virtual machines on that datastore.

Move powered-off and suspended virtual machines to other hosts in the cluster

Virtual SAN data might reside on the hosts in a Virtual SAN cluster. Select an option to set the migration mechanism for the Virtual SAN data that will be enforced before the hosts enter maintenance mode.

Virtual SAN data migration:



Put the selected hosts in maintenance mode?



如何开启/关闭vSAN节点

vCenter在vSAN中

关机顺序:

- 1. 关闭除vCenter之外的其他所有VM
- 2. 为了方便启动,将vCenter迁移到第一台主机
- 3. 关闭vCenter, Web Client会断连
- 4. 所有主机置于维护模式,必须命令行操作
 - esxcli system maintenanceMode set –e true –m noAction
- 5. 关闭所有主机

注:先关闭vCenter,后主机进入维护模式,否则vCenter的数据变成不可访问状态

开机顺序:

- 1. 打开所有主机
- 2. 把所有主机退出维护模式: esxcli system maintenanceMode set -e false
- 3. 连接vCenter所在的主机,通过Client开启vCenter
- 4. 登录vCenter开启其他VM

Confidential | ©2019 VMware, Inc.

从vSAN移除磁盘组或设备

从磁盘组中移除一个闪存缓存设备或所有容量设备的同时会移除整个磁盘组

删除设备或磁盘组时,可以通过选择撤出全部数据选项或选择确保数据可访问性,将 Virtual SAN 主机置于维护模式

- 1. 在vsphere Web Client 中导航到集群,选择"配置"
- 2. 在vSAN下选择磁盘管理,选择故障的设备/磁盘组
- 3. 选择闪存设备移除,选择数据撤出模式

| X Pre-check evacuation - Local VMware Disk (mpx.vmhba1:C0:T0:L0) | | | | | | 🚯 RegionA01-COMP01 👔 🔒 | 🔁 🚼 😂 🎯 Actions 👻 | - | | |
|---|--------------------------|-------------|---------------------------------|--|-------------------|---------------------------|--|--------------------|------------------|---|
| Consumption of the data data and the sector of the sector | | | Getting Started Summary Monitor | Configure Permissions Hosts VMs Datastores Network | | | | | | |
| Evacuate all data to other nosts Sufficient canacity on other hosts 3.04. GB will be moved | | | | | | " | Disk Groups | | Disk Groups Disk | Disk "Local VMware Disk (mpx.vmhba1:C0:T0:L0)" is about to be removed from the disk group "Disk group (000000000766d686261313a303a30)". Removing |
| 🔘 Ensure da | ata accessibility from o | ther hosts | 3 | | | ✓ Services | a 🗙 🙀 ! 🛀 '- '= C' | Q Filte | ۱۲ - | the flash disk will remove the entire disk group. Unless the data on the disk(s) is |
| 🛕 No data | a will be moved. 5 obj | ects will b | ecome non-compliant wit | th storage policy. | | vSphere DRS | Disk Group | Disks in Use State | | evacuated first, removing the disk(s) might disrupt working VMs. |
| 🔵 No data ev | vacuation | | | | | | esx-04a.corp.local | 2 of 2 | Connected | Data on disk group: 6.43 GB |
| 🛕 5 objec | cts will become non-co | mpliant | vith storage policy. | | | General | esx-01a.corp.local | 2 2 of 2 | Connected | Select data migration mode: |
| | | | | | | Disk Management | 📰 Disk group (000000000766d686261313a303a30) | 2 | Mounted | Evacuate all data to other hosts |
| | | | | Q Fil | ter 👻 | Fault Domains & Stretched | ✓ ■ esx-03a.corp.local | 2 of 2 | Connected | Sufficient capacity on other hosts. 6.43 GB will be moved. |
| Name | Object 1 | ype | Result | Storage Policy | Compliance Status | Health and Performance | Disk group (000000000766d686261313a303a30) | 2 2 of 2 | Mounted | Ensure data accessibility from other hosts |
| | | | This list is empty. | | | iSCSI Targets | Disk group (000000000766d686261313a303a30) | 2 | Mounted | No data will be moved. 7 objects will become non-compliant with storage notice. See full pre-check evecuation results. |
| | | | | | | iSCSI Initiator Groups | 4 | | | No data evacuation |
| | | | | | | Configuration Assist | 86 | 8 items 🔒 | Export 👻 🎦 Copy | 7 objects will become non-compliant with storage policy. See full pre-check |
| | | | | | | Updates | Disk group (00000000766d686261313a303a30); Disks | | | evacuation results. |
| | | | | | | | 🖴 🔏 🙀 🥥 💿 🄯 All Actions 🗸 | Show: In us | e (2) | |
| | | | | | | Licensing | Name | Drive Type | Disk Tier Capa | er Remove disk? |
| | | | | | | VMware EVC | Local VMware Disk (mpx/mhba1:C0:T0:L0) | Flash | Cache 5.00 | Yes No |
| | | | R | | | VM/Host Groups | E Local VMware Disk (mpx.vmhba1:C0:T1:L0) | Flash | Capacity 40.0 | |
| | VIIIVVc | al G | Confidential | ©2019 V | Mware, Inc. | VM/Host Rules | | | | 60 |
| | | | | | | VM Overrides | | | | |

移除vSAN 节点

- 迁移该节点所有虚拟机到其他主机
- 进入维护模式,选择"迁移全部数据"
- 等待数据迁移结束, 主机进入维护模式, 移除磁盘组
- 从vSAN集群移除主机



vmware[®]

Confidential | ©2019 VMware, Inc.

vSAN运维中的不当操作

- vSAN使用率100%
- 使用老版本
- 同时/短时间连续重启多台主机
- 数据同步的过程中重启主机
- 启动过程中重启主机
- 未进入维护模式直接重启主机
- 直接移除硬件,例如:磁盘



容量使用率最佳实践

推荐SSD比例至少为HDD容量的10%

- ●如果SSD比例过小则容易出现磁盘拥堵的故障
- ●举例:磁盘组HDD裸容量是4TB,推荐SSD容量为400GB

vSAN尝试在磁盘空间使用率达到80%时对其执行平衡操作

- ●如果一个磁盘达到80%,则vSAN将自动修复此情况,如果所有物理磁盘都使用超过80%的容量,则 vSAN将尝试使所用容量的大小保持平衡,此时需要考虑向集群中引入额外的容量
- ●建议保留30%的空闲空间



从非vSAN环境将VM迁移到vSAN环境

常用迁移方式

- Storage vMotion
- Deploy OVF

使用 Web Client进行Storage vMotion或部署OVF,在执行过程中选择vSAN Datastore以及存储策略

| 1 Select the migration type 2 Select storage 3 Ready to complete | Select the migration type Change the virtual machines' compute resource, storage, or both. Change compute resource only Migrate the virtual machines to another host or cluster. Image: Change storage only Migrate the virtual machines' storage to a compatible datastore or datastore cluster. | 1 Select the migration type 2 Select storage 3 Ready to complete | Select storage Select the destination storage for the virtual machine migration. Select virtual disk format: As defined in the VM storage policy VM storage policy: Virtual SAN Default Storage Policy The following datastores are accessible from the destination resource that you selected. Select the destination datastor virtual machine configuration files and all of the virtual disks. | | | | | | |
|--|---|--|---|-----------|-------------|-----------|--------|--|--|
| | Change both compute resource and storage Migrate the virtual machines to a specific host or cluster and their storage to a specific datastore or datastore cluster | r | Name Compatible | Capacity | Provisioned | Free | Туре | | |
| | Select compute resource first | | cn-tse-vsanDatastore | 1.64 TB | 1.21 TB | 716.85 GB | vsan | | |
| | O Select storage inst | | Incompatible | | | | | | |
| | | | cn-tse-d07-datastore-3-SSD | 1.09 TB | 1.17 TB | 973.47 GB | VMFS 5 | | |
| | | | cn-tse-d07-datastore-4-SSD | 1.09 TB | 1.10 TB | 824.40 GB | VMFS 5 | | |
| | | | cn-tse-d07-datastore-1 | 272.00 GB | 23.16 GB | 248.84 GB | VMFS 5 | | |

运维需要注意

- vSAN数据存储推荐预留30%空闲空间,变更存储策略推荐25~30%空间用于组件重建
- 当设置FTT=O时,避免维护模式选择'无数据迁移'
- 维护模式选择'迁移所有数据'时有足够数量的故障域
- 启动去重压缩功能后,需要轮滚升级,磁盘可以添加,但不能移除
- 当集群中出现故障,检查对象状态是否健康
- 在vSAN健康状态和磁盘管理中观察是否网络分区
- 利用vSphere Update Manager自动化升级
- 主动监测磁盘使用容量以及性能,避免集群不平衡
- 主机重启时远程访问ESXi DCUI
- 升级顺序:vCenter Server (升级后重启) —>ESXi主机 (轮滚升级) —>磁盘格式
- 为vCenter Server、每个主机、虚拟设备配置相同DNS、NTP 服务器

Agenda

- 1 规划设计与最佳实践
- 2 部署指南
- 3 运维管理—集群操作

4 运维管理--故障诊断

vSAN故障诊断工具

健康检查

ESXCLI

Ruby vSphere Console (RVC)

vSAN Observer

- ●物理磁盘层的统计信息
- ●深入探索物理磁盘组的细节
- ●CPU 使用量统计信息
- Virtual SAN 内存池消耗情况
- Virtual SAN 群集中的物理分布和内存对象分布

VDQ

ESXTOP

Health Cluster List

| [root@sa-esxi-03:/dev/vsan] esxcli vsan health clust | ter list | | | | | |
|--|--------------------------|--|--|--|--|--|
| Health Test Name | Status | | | | | |
| | | | | | | |
| Overall health | red (vSAN Object health) | | | | | |
| Data | red | | | | | |
| vSAN object health | red | | | | | |
| Cluster | yellow | | | | | |
| ESXi vSAN Health service installation | green | | | | | |
| vSAN Health Service up-to-date | green | | | | | |
| Advanced vSAN configuration in sync | green | | | | | |
| vSAN CLOMD liveness | green | | | | | |
| vSAN Disk Balance | yellow | | | | | |
| Resync operations throttling | green | | | | | |
| Software version compatibility | green | | | | | |
| Disk format version | green | | | | | |
| Network | green | | | | | |
| Hosts disconnected from VC | green | | | | | |
| Hosts with connectivity issues | green | | | | | |
| vSAN cluster partition | green | | | | | |
| All hosts have a vSAN vmknic configured | green | | | | | |
| All hosts have matching subnets | green | | | | | |

vsan.check_state

集群的状态?

快速了解集群是否有严重的问题

/sa-vcsa-01.vclass.local/SA-Datacenter/computers> vsan.check state 0 2017-07-07 14:54:52 +0000: Step 1: Check for inaccessible vSAN objects

Detected 2 objects to be inaccessible

Detected ec263959-a242-3303-fd92-00505602f99a on sa-esxi-02.vclass.local to be inaccessible Detected 26283959-7ce3-849b-f37b-00505602f99e on sa-esxi-02.vclass.local to be inaccessible

2017-07-07 14:54:52 +0000: Step 2: Check for invalid/inaccessible VMs

2017-07-07 14:54:52 +0000: Step 3: Check for VMs for which VC/hostd/vmx are out of sync Did not find VMs for which VC/hostd/vmx are out of sync

vsan.disks_stats 0

| /sa-vcsa-01.vclass.loca | al/SA-Datacenter/computers | > vsan.d | isks sta | ats 0 | | | | L | I | | | | ll |
|--|---|-------------------------|-----------------------|-------------------------------------|----------------------------------|---------------------------------|------------------------|--------------------------|------------------------|-----------------------|-----------------------|-----------------------|---------------------------------------|
| DisplayName | Host | isSSD | Num Comp | Capacity Total | Used | Reserved | Physical Capacity | Physical Used | Physical Reserved | Logical Capacity | Logical Used | Logical Reserved | Status Health |
| mpx.vmhba2:C0:T1:L0 mpx.vmhba3:C0:T1:L0 | sa-esxi-01.vclass.local sa-esxi-01.vclass.local | SSD MD | 0 12 | 5.00 GB 19.99 GB | 0.00 % 59.83 % | 0.00 % 0.64 % | N/A N/A | N/A N/A + | N/A N/A | N/A N/A | N/A N/A | N/A N/A | OK (v5) OK (v5) |
| mpx.vmhba2:C0:T0:L0 mpx.vmhba3:C0:T0:L0 | / sa-esxi-01.vclass.local sa-esxi-01.vclass.local | ' SSD MD | 0 14 | 5.00 GB 19.99 GB | 0.00 % 29.50 % | 0.00 % 7.01 % | N/A N/A | ' N/A N/A + | N/A N/A | N/A N/A | N/A N/A | N/A N/A | OK (v5) OK (v5) |
| <pre> mpx.vmhba2:C0:T0:L0 mpx.vmhba3:C0:T1:L0 mpx.vmhba3:C0:T0:L0</pre> | sa-esxi-02.vclass.local sa-esxi-02.vclass.local sa-esxi-02.vclass.local | SSD MD MD | 0 14 18 | 5.00 GB 19.99 GB 19.99 GB | 0.00 % 28.27 % 54.53 % | 0.00 % 0.72 % 27.37 % | N/A N/A N/A | N/A N/A N/A | N/A N/A N/A | N/A N/A N/A | N/A N/A N/A | N/A N/A N/A | OK (V5) OK (V5) OK (V5) |
| <pre> mpx.vmhba2:C0:T0:L0 mpx.vmhba3:C0:T0:L0</pre> | sa-esxi-03.vclass.local sa-esxi-03.vclass.local sa-esxi-03.vclass.local | + SSD MD + | + 0 16 + | 5.00 GB 19.99 GB | 0.00 % 60.98 % | 0.00 % | N/A N/A N/A | + N/A N/A + | N/A N/A N/A | N/A N/A | N/A N/A N/A | N/A N/A N/A | OK (v5) OK (v5) OK (v5) |
| mpx.vmhba2:C0:T0:L0 mpx.vmhba3:C0:T0:L0 + | sa-esxi-04.vclass.local sa-esxi-04.vclass.local + | SSD MD + | 0 17 + | 5.00 GB 19.99 GB + | 0.00 % 45.86 % | 0.00 % 21.04 % | N/A N/A + | N/A N/A + | N/A N/A + | N/A N/A | N/A N/A + | N/A N/A + | OK (v5) OK (v5) ++ |

KB 2145267 –vSAN on-disk format

升级

升级顺序:

- 升级vCenter Server到6.7
- 升级主机到ESXi 6.7
- •升级vSAN磁盘格式到version 6.0
- 注: 若磁盘测试从version 5.0升级,磁盘格式化时不需要执行数据撤出

升级路线:

- 支持vSphere 6.0升级到vSphere 6.7
- 支持从vSphere 6.5-6.5 Update1升级到vSphere 6.7
- 目前不支持从vSphere 6.5 Update2 升级到vSphere 6.7
收集vSAN日志方法

| | 主机日志 | RVC日志 | vCenter日志 | 主机性能日志 | vSAN性能数据 | Hostd dump | 截图(包含名称时间点) |
|------------------|---------------|---------|-----------|--------|----------|------------|-------------|
| VM无法启动/数据丢失 | \star | \star | 0 | 0 | 0 | 0 | • |
| vSAN使用/管理/界面相关问题 | \star | \star | \star | 0 | 0 | 0 | • |
| 健康检查报错 | \star | \star | \star | 0 | 0 | 0 | • |
| 主机无响应/断开连接 | \star | Ο | 0 | 0 | 0 | • | • |
| 磁盘/磁盘组/硬件相关问题 | \star | \star | 0 | 0 | 0 | 0 | • |
| PSOD | \star | 0 | 0 | 0 | 0 | 0 | • |
| 无法创建VM | \star | • | \star | 0 | 0 | 0 | • |
| Storage Provider | \rightarrow | 0 | \star | 0 | 0 | 0 | • |
| 元数据状态异常 | \star | \star | • | 0 | 0 | 0 | • |
| vSAN性能问题 | \star | \star | 0 | * | * | 0 | • |
| 主机性能问题 | \star | \star | 0 | * | 0 | 0 | • |
| 延伸集群/ROBO集群 | \star | 0 | \star | 0 | 0 | 0 | • |
| | | | | | | | |
| | | | | | | | |
| | | * | 需要提供 | | | | |
| | | • | 建议提供 | | | | |
| | | 0 | 不需要提供 | | | | |

VMware vSAN运维最佳实践

下载地址: http://vmpartner.incloudexpo.com/ftpdata/最佳实践/Index.htm

| VMw 第 | are vSAN 售后最佳实践 5一部:交付检查清单 | VMware vSAN 售后最佳实践 第二部:运维检查清单 | VMware vSAN 售后最佳实践 第三部:常用操作汇总 | | |
|-----------------------------|--|--|-------------------------------------|--|--|
| | | WMWARE VSAN HYPER-CONVERGED ST. 2013 | A REF CONTRACTOR | | |
| | 项目交付 | 运维检查 | 日常操作 | | |
| VM 第 | lware vSAN 售后最佳实践 驷部:日志收集方法汇总 | VMware vSAN 售后最佳实践 第五部:vSAN 升级手册 | VMware vSAN 售后最佳实践 第六部:vSAN 扩容手册 | | |
| | SINCE USAN SINCE USAN SINCE USAN | Powered By vmware [™] vSAN [™] | PROJECT | | |
| | 日志收集 | 产品升级 | 硬体动容 | | |
| vm ware [®] | Confidential 📔 ©2019 VMware. Inc. | | | | |

Thank you!



Confidential | ©2019 VMware, Inc.