

DELL EMC VXRAIL™ 系统技术手册

来自 Dell EMC® 和 VMware® 的超融合基础架构系统

摘要

本技术手册从概念和体系结构方面介绍采用 VMware vSAN 和 Intel Inside® 的 Dell EMC VxRail 系统。本技术手册首先介绍超融合基础架构如何推动数字转型, 然后重点介绍作为业界卓越的超融合技术解决方案的 VxRail 系统。

2019 年 8 月

目录

| | |
|--|-----------|
| IT 面临的新挑战：引领数字转型 | 7 |
| Dell EMC 数字转型：更快获得成果。简化 IT。 | 8 |
| 利用 Dell EMC 超融合基础架构加快 IT 转型 | 9 |
| 完全转型为混合云 | 9 |
| 创新而不是集成 | 10 |
| 超融合基础架构：旨在应对现代 IT 挑战的现代化基础架构 | 11 |
| HCI 的促成技术 | 12 |
| 采用超融合基础架构的推动因素 | 12 |
| Dell EMC 超融合基础架构平台 | 14 |
| Dell EMC HCI 提供全包式客户体验 | 15 |
| Dell EMC VxRail 系统 | 16 |
| VxRail 安全性与合规性 | 18 |
| VxRail 硬件体系结构 | 20 |
| VxRail 群集 | 21 |
| VxRail 型号和规格 (基于第 14 代 Dell EMC PowerEdge 服务器) | 21 |
| VxRail 节点 | 22 |
| 英特尔® 至强® 可扩展处理器：VxRail 强大的处理能力 | 24 |
| VxRail 节点存储磁盘驱动器 | 24 |
| VxRail 硬件选项 | 25 |
| VxRail 扩展 | 26 |
| 可升级的选项 | 27 |
| VxRail 网络 | 27 |
| 1GbE 网络选项 | 29 |
| Dell EMC 开放式网络和 VxRail | 30 |
| Dell EMC SmartFabric Services (SFS) | 30 |
| VxRail 软件体系结构 | 31 |
| 系统管理 | 31 |
| VxRail HCI 系统软件 | 31 |
| VxRail Manager | 34 |

| | |
|---|-----------|
| 可由客户升级的软件..... | 36 |
| VxRail 分析咨询引擎 (ACE)..... | 37 |
| vSphere 和 vSAN 订购信息..... | 38 |
| VMware vSphere..... | 40 |
| VMware vCenter Server..... | 40 |
| vCenter 服务和接口..... | 43 |
| 增强型链接模式..... | 44 |
| VMware vSphere ESXi..... | 44 |
| vCenter Server 与 ESXi 主机之间的通信..... | 45 |
| 虚拟机..... | 46 |
| 虚拟机硬件..... | 47 |
| 虚拟机通信..... | 47 |
| 虚拟网络..... | 48 |
| 虚拟分布式交换机..... | 48 |
| vMotion 和虚拟机移动性..... | 49 |
| 增强的 vMotion 兼容性..... | 51 |
| Storage vMotion..... | 51 |
| vSphere Distributed Resource Scheduler..... | 52 |
| vSphere High Availability (HA)..... | 54 |
| vCenter Server 监视器..... | 56 |
| vSphere 加密..... | 56 |
| vSAN..... | 57 |
| 磁盘组..... | 58 |
| 混合和全闪存差异..... | 59 |
| 读缓存: 基本功能..... | 59 |
| 写缓存: 基本功能..... | 59 |
| 闪存耐用性..... | 60 |
| vSAN 对闪存耐用性的影响..... | 60 |
| 客户端缓存..... | 60 |
| 对象和组件..... | 60 |
| 见证组件..... | 61 |
| 复制副本..... | 61 |
| 基于存储策略的管理 (SPBM)..... | 62 |
| 动态策略更改..... | 62 |
| 存储策略属性..... | 63 |

| | |
|--|-----------|
| 稀疏交换 | 65 |
| I/O 路径和缓存算法 | 66 |
| 读缓存 | 66 |
| 混合读取剖析 | 67 |
| 全闪存读取剖析 | 68 |
| 写缓存 | 68 |
| 写 I/O 剖析 — 混合和全闪存 (FTM=镜像) | 69 |
| 分布式高速缓存注意事项 | 70 |
| vSAN 高可用性和故障域 | 71 |
| 故障域概述 | 71 |
| 故障域和机架级故障 | 72 |
| 部署最低群集配置时的注意事项 | 73 |
| vSAN 延伸群集 | 73 |
| 采用本地保护的延伸群集 | 74 |
| 站点局部性 | 75 |
| 网络 | 75 |
| 延伸群集心跳和站点偏向 | 75 |
| 延伸群集的 vSphere HA 设置 | 75 |
| 双节点配置 | 76 |
| 快照 | 77 |
| 使用重复数据消除和压缩的存储效率 | 77 |
| 重复数据消除和压缩开销 | 79 |
| 擦除编码 | 80 |
| 启用擦除编码 | 81 |
| 擦除编码开销 | 82 |
| vSAN 加密 | 82 |
| VxRail 集成软件 | 83 |
| 虚拟机复制 | 83 |
| VMware vSphere Replication | 83 |
| Dell EMC RecoverPoint for Virtual Machines | 84 |
| VxRail 复制使用情形 | 86 |
| 对外部网络存储的支持 | 87 |
| 光纤通道与 VxRail | 87 |
| iSCSI 与 VxRail | 87 |
| NFS 与 VxRail | 88 |

| | |
|---|------------|
| VxRail 解决方案和生态系统..... | 90 |
| VMware 验证设计与 VxRail | 90 |
| VMware Cloud Foundation on VxRail | 92 |
| Pivotal Ready Architecture (PRA)..... | 93 |
| 灵活消费选项..... | 94 |
| Flex on Demand — 类云消费选项..... | 94 |
| VDI 完整解决方案..... | 95 |
| VMware Horizon..... | 95 |
| VMware Horizon 及 VxRail..... | 96 |
| VMware vSphere Platinum..... | 97 |
| IsilonSD Edge | 98 |
| SAP HANA 认证及 VxRail | 99 |
| 适用于 Splunk 的参考体系结构 | 99 |
| | |
| 其他产品信息..... | 101 |
| Dell EMC ProSupport for Enterprise | 101 |
| 针对 VxRail 系统的 Dell EMC ProDeploy 服务 | 102 |
| 了解更多信息..... | 103 |
| 联系人..... | 103 |
| 查看更多详情..... | 103 |

前言

Dell EMC 技术手册从概念和体系结构方面介绍采用 VMware vSAN 和 Intel Inside 的 Dell EMC VxRail™ 系统。本技术手册介绍超融合基础架构如何推动数字转型，并且重点介绍作为业界卓越的超融合技术解决方案的 VxRail 系统。

受众

本技术手册面向参与设计、购置、管理或操作 VxRail 系统解决方案的 Dell EMC 现场人员、合作伙伴和客户。

相关资源和文档

请参阅以下各项以了解相关的补充文档、技术白皮书和网站。

Dell EMC VxRail 网络指南: <https://vxrail.is/networkplanning>

《VxRail Planning Guide for Virtual SAN Stretched Cluster》(VxRail 虚拟 SAN 延伸群集规划指南): <https://vxrail.is/stretchedclusterplanning>

《An overview of VMware vSAN Caching Algorithms》(VMware vSAN 缓存算法概述): <https://www.vmware.com/files/pdf/products/vsan/vmware-virtual-san-caching-whitepaper.pdf>

《VMware vSAN 6.2 Space Efficient Technologies Technical White Paper》(VMware vSAN 6.2 空间效率技术的技术白皮书): <http://www.vmware.com/files/pdf/products/vsan/vmware-vsan-62-space-efficiency-technologies.pdf>

《VxRail Stretched Cluster》(VxRail 延伸群集): <https://www.dell.com/resources/zh-cn/asset/offering-overview-documents/products/converged-infrastructure/vxrail-stretch-cluster-so.pdf>

《vSphere Virtual Machine Administration Guide》(vSphere 虚拟机管理指南): <https://www.vmware.com/support/pubs/vsphere-esxi-vcenter-server-6-pubs.html>

《vSphere Resource Management》(vSphere 资源管理): <http://pubs.vmware.com/vsphere-65/topic/com.vmware.ICbase/PDF/vsphere-esxi-vcenter-server-65-resource-management-guide.pdf>

《Dell EMC Hyperconverged infrastructure》(Dell EMC 超融合基础架构): <http://www.dell.com/zh-cn/work/shop/category/hyper-converged-infrastructure>

《VxRail vCenter Server Planning Guide》(VxRail vCenter Server 规划指南): <https://vxrail.is/vcenterplanning>

IT 面临的新挑战：引领数字转型

在数字经济中，应用是现代企业的脸面和基石。

对于数字化客户来说，用户体验胜过一切。面向客户的应用程序必须能够随时随地在任何设备上使用，并且必须提供实时更新和智能交互。对于企业来说，利用从这些交互收集的数据获得见解可以了解和推动未来发展需求。

应用程序和底层基础架构对企业具有战略意义。能够高效利用现代数据中心技术快速向客户提供创新能力的企业有望取得真正的成功。

应用程序在现代企业中的重要性为 IT 组织带来了巨大的商机。IT 不再只是后勤职能部门，它可以引领数字转型，促进企业不断取得成功。IT 可以积极推动企业发展。

传统 IT 团队面临非常复杂的应用程序构建、配置、维护和扩展工作。组织需要成功地部署和运营一个充分利用全行业创新的环境，同时避免由于拼凑各种工具和为这些拼凑式工具提供支持所产生的复杂性问题。

他们面临的挑战是如何实现这一转型。Dell EMC 调查了¹ 多个行业的 1000 多位高管，了解他们的数字转型状况。调查问题侧重于：

现代化数据中心技术利用情况：例如使用全闪存阵列、横向扩展体系结构、融合和超融合平台，以及跨网络和存储领域的软件定义的解决方案。

自动化 IT 流程：衡量标准是组织在更像公有云提供商那样运营 IT 方面取得的进展（例如，实现自助式基础架构资源调配；快速扩展能力；以及基于使用量进行跟踪和计费）。

转型后的业务与 IT 关系：是通过 IT 与业务部门利益相关者的一致沟通以及业务线 (LOB) 领导团队对 IT 成效的持续审查而实现的。

调查结果显示，他们的进展充其量是喜忧参半。一些公司几乎没有开始数字转型。许多公司采取了零敲碎打的方法。只有少数一些公司差不多完成了数字转型。为什么要花这么长时间？

归根结底是 IT 转型很难。它需要对基础架构技术和应用程序进行大量规划、评估、重组，使之实现现代化。许多因素（包括成本、技能组合、治理、创新推动力以及转型意愿）都会影响企业能否超越传统的三层数据中心结构。

¹ IT 转型成熟度如何推动实现 IT 敏捷性、创新和更丰硕的业务成果，Enterprise Strategy Group, 2017 年 4 月。
<https://China.emc.com/collateral/analyst-reports/esg-dellemc-it-transformation-maturity-report.pdf>

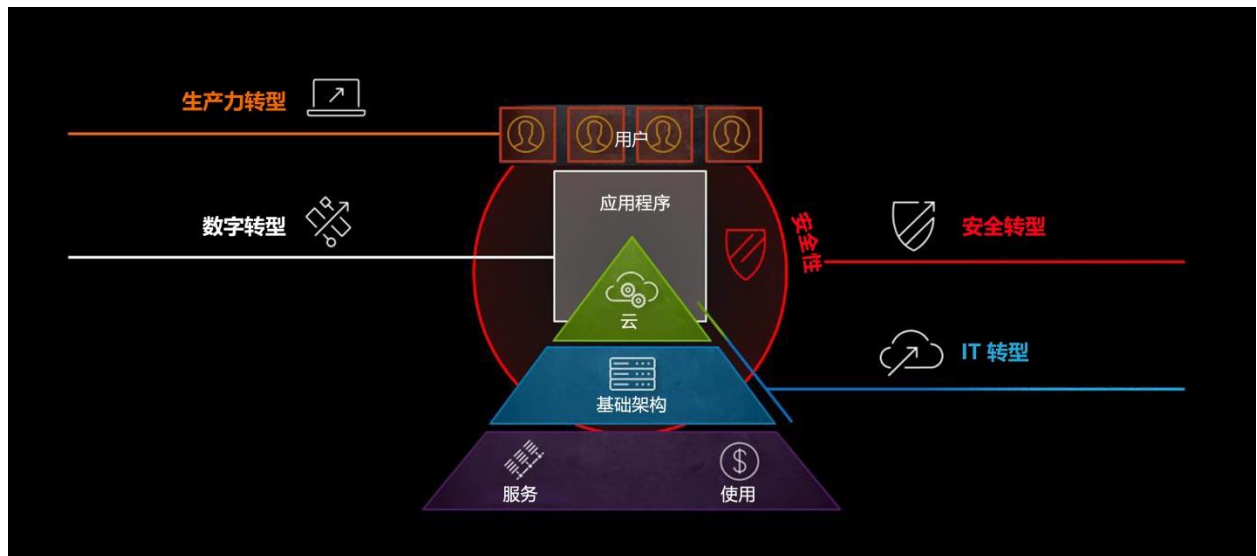
每个企业以不同的速度展开 IT 转型，并且对 IT 转型有不同的目标。并非每个企业都想要或需要采用全面的云服务交付模式。企业需要的是这样一种方法：使企业能够转型到可实现所需转型优势的程度，并以适合业务模式的速度进行转型。

Dell EMC 数字转型：更快获得成果。简化 IT。

要从战略上全面了解转型，您必须首先采取以应用程序为中心的视角。确保应用程序拥有适当的服务级别协议战略，无论是在本地还是异地部署，都需要采用混合云战略。Dell EMC 利用超融合基础架构 (HCI) 提供完全集成且受支持的 HCI 解决方案和混合云平台，促进企业数据中心和 IT 实现现代化、自动化和转型。

下图显示了用于数字转型的 Dell EMC 模型。

图 1 用于数字转型的 Dell EMC 模型



Dell EMC 提供精心设计的全包式超融合基础架构解决方案，使企业能够更快地进行创新和加快 IT 运营。

Dell EMC 融合和超融合基础架构可以提供基于卓越的硬件和软件构建而成的以应用程序为中心的解决方案，在提供真正的业务价值的同时，大大减少部署业务关键型应用程序、一般用途应用程序和云原生应用程序的风险和成本。

最后一点是，Dell EMC 提供全套服务和灵活消费模式，让您可以更快、更轻松地开始使用这些解决方案。

利用 Dell EMC 超融合基础架构加快 IT 转型

企业在其转型之旅中可以采取的一个初步措施是通过向环境中引入超融合基础架构 (HCI) 来简化基础架构部署和管理。HCI 系统实质上是将传统的三层服务器、网络和存储模型折叠起来, 以便基础架构本身更易于管理。

采用可在本机集成计算、存储、虚拟化、管理和数据服务的超融合基础架构解决方案, 可以大幅减少 IT 管理任务并为实现现代化 IT 基础架构奠定基础。无论工作负载部署和实施程度如何, HCI 解决方案都是降低基础架构成本和简化管理的理想选择。

完全转型为混合云

许多企业最终都希望通过混合云的自助服务目录实现 IT 服务交付自动化。混合云可带来以下好处:

- 为本地和异地资源提供单一控制点
- 通过自动化精简 IT 资源的交付, 可以以一致、可重复而且与业务最佳做法相一致的方式交付这些资源
- IT 团队可以利用计量功能宣传服务的价值, 同时让业务部门能够看到资源的成本和使用情况
- 自助功能使应用所有者和业务用户能够在需要时访问他们所需的资源
- 容量管理让 IT 团队可以跨整个混合云更好地管理资源
- 监控和报告可以提供对环境的容量、性能和运行状况的可见性
- 内置安全功能可以保护企业工作负载
- 服务级别选择可使各项工作负载对应于正确的服务级别和成本目标
- 能够以应用程序级别粒度满足服务级别协议要求

混合云的愿景不是新生事物。企业已尝试基于通过在大区块中部署和扩展的存储网络来访问的纵向扩展存储, 使用传统基础架构来部署混合云。尽管可以利用纵向扩展存储在传统的三层基础架构上构建云功能, 但这不是最佳解决方案。

如果企业希望通过将 IT 全面转型到云, 从而支持其应用程序环境, Dell EMC 可以用基于超融合基础架构而构建的全包式混合云平台来帮助让您的 IT 运营实现现代化、自动化和转型。

创新而不是集成

企业可以选择构建完全自定义的解决方案。将存储、网络、计算、数据保护、监视和报告集成在一起，然后确定如何让所有这些组件一起工作可能很耗时，但对于希望将指定的供应商组件作为其解决方案的一部分的组织来说，这提供了极大的灵活性。规划、设计和构建自定义解决方案是一个复杂的项目，通常需要数月甚至数年才能实现，这对需要推出解决方案以满足当前业务需求的企业来说太过久远，并且从长远来看，其维护和升级成本也很高。

IT 部门的难题是，复杂性存在于这些层中的每一层，所以构建并维护一个可正常发挥功能且有弹性的云会非常困难。许多公司发现，靠他们自己完成这件事需要占用他们 70% 的 IT 资源和预算，所以将只有很少的资源可用于推动创新和实施为业务带来真正价值的项目。

对大多数企业而言，使用 HCI 解决方案的好方法是购买它们，它们完全集成了生命周期管理和单一支持来源。与自主构建的传统三层基础架构相比，购买可缩短收效时间、简化运营，并让五年总拥有成本实现 619% 的节省。²

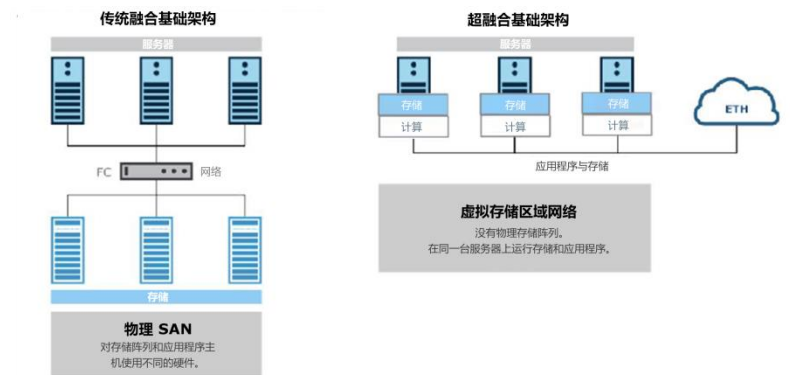
² 来源：IDC 2017 年 10 月

超融合基础架构：旨在应对现代 IT 挑战的现代化基础架构

融合基础架构平台是完全预集成的服务器、传统存储阵列和网络硬件“堆栈”。超融合基础架构平台是在行业标准服务器群集中提供计算、软件定义的存储和网络基础架构服务的解决方案。

超融合基础架构通过结合软件定义的存储 (SDS) 的虚拟化功能，扩展融合基础架构模型。超融合基础架构将传统数据中心的组件 (计算和存储) 融合到一个服务器中，从而有效地消除成本高昂和复杂的 SAN 环境。请参见下图。

图 2 CI 和 HCI



由于 HCI 是软件定义的，这意味着基础架构运营在逻辑上与物理硬件分离，组件之间的集成与 CI 相比要紧密得多。HCI 通过一个通用工具集将所有的组件作为一个系统来管理。

超融合基础架构特别有价值，因为它让您可以快速纵向扩展，而不会产生大量的额外费用。在传统环境中情况并非如此：因为预计会进行纵向扩展，所以客户必须购买超出其需求的更多资源，或者等到当前工作负载耗尽分配的资源，然后再添加基础架构。在不适当的时间购买意味着资源没有得到最佳分配，甚至会拖慢业务扩张的速度。

HCI 支持随增长付费的方法，开始时先购买当前所需的资源，然后逐步扩展，而不用提前购买大量计算和存储设备。在打算持续多年使用技术时，这种做法还可以解决经常发生的过量配置和过度购买的问题。

HCI 部署模型

Dell EMC 为 IT 转型确定了四种部署模型，这些模型根据企业所需实现的最终状态和为实现该最终状态需要达到的运营就绪情况来定义。

自主构建

企业希望获得 HCI 带来的益处，但更希望在服务器供应商和配置选择方面保持控制和灵活性。

HCI 系统

系统是一个更全包式的选项，包括服务器、软件堆栈和生命周期管理。它们完全作为单个产品进行设计和获得支持。

机架级 HCI

机架级 HCI 解决方案扩展了 HCI 系统提供的功能，此外，它还将物理和软件定义的网络以及软件定义的存储、计算、虚拟化、数据保护和管理软件紧密集成到一起，以便通过软件定义的管理和编排层为完整的基础架构配置和资源调配实现完全自动化。像 HCI 系统一样，它们完全作为单个产品进行设计和获得支持。

基于 HCI 的全包式混合云

将系统或机架级 HCI 与附加云管理、自动化与编排软件以及预先设计的服务和工作流相结合，提供一个可与异地公有云提供商集成的敏捷、弹性、自动化和自行维持的私有云。

HCI 的促成技术

下表列出了促进超融合基础架构增长和发展的技术融合。

表 1. HCI 的促成技术

| 技术 | 描述 |
|---------|---|
| 软件定义的存储 | 从底层存储基础架构提取存储智能。 将直连存储虚拟化到共享池中。 自动执行调配和负载平衡。 让企业可以通过向群集添加整个节点（例如，具有存储软件和介质的服务器）来增加可用存储资源（包括容量和处理能力）。生成的节点群集反过来可用作单个存储容量池。 |
| 虚拟化 | 抽取计算和网络功能。 实现物理资源的共享。 提高利用率、移动性和安全性。 |
| X86 服务器 | 高性能处理器、大容量内存。 闪存介质提供一致且可预测的性能。 |
| 固态存储 | 使用固态驱动器（最常见的是各种类型的闪存）来存储数据。此存储可以驻留在存储控制器或服务器中，但对于此评估而言，我们考虑的使用情形限于分层和全闪存存储阵列。 在混合阵列中，阵列中的部分驱动器是固态的，其中存放阵列中最活跃的数据。 在全闪存阵列中，阵列中的所有驱动器都是固态的。 |
| 高速网络 | 将节点连接在一起形成群集。 使 HCI 能够提供 IOPS 和更短的延迟。 将应用程序连接到用户 |

采用超融合基础架构的推动因素

已经过渡或计划过渡到 HCI 的客户表示，他们期望在实施 HCI 时获得的主要优势包括成本缩减、部署加快、扩展能力和运营效率提高以及基础架构任务减少。

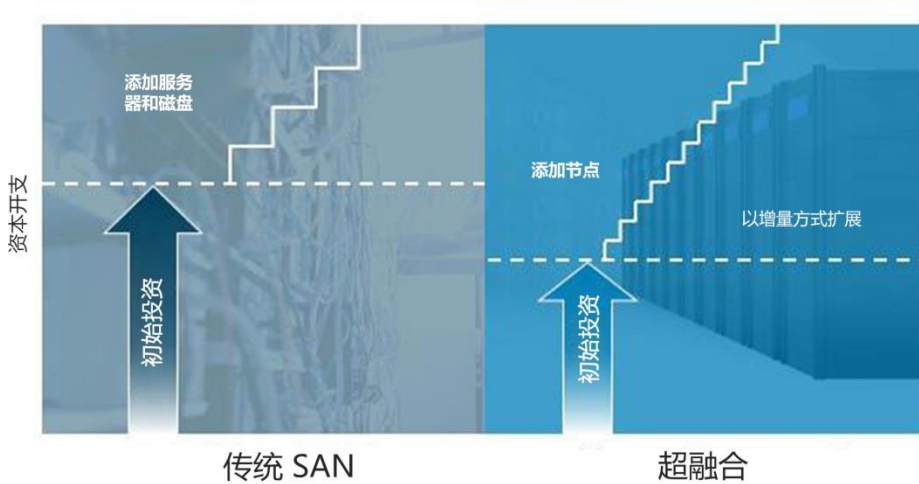
HCI 为资本支出和运营支出提供了具有说服力的案例。

与传统的三层体系结构相比，初始投资较低，运营支出也较低。成本节约包括电力、散热、日常系统管理成本的减少，以及由于免除了中断性升级和数据迁移而节约的成本。

企业可以购买针对特定工作负载的基础架构，而不是购买基于 SAN 的单体基础架构。超融合解决方案之所以能够降低总拥有成本和提高灵活性的一个主要促成因素是，它能够从更小规模开始并以增量方式扩展。

不仅初始资本支出较低，而且您还可以按增量方式扩展，以更小的幅度逐步增加计算量，甚至只根据需要扩展存储容量，如下图所示。

图 3 HCI：根据需要购买，随增长付费



此外，某些 HCI 系统在单个机箱中支持多个节点。因此，如果有可用空间，您只需几分钟便可将节点插入现有机箱以扩展基础架构。部署 HCI 可缩短时间、降低风险和复杂性。

下图显示了与传统 SAN 环境相比，HCI 如何简化 IT 管理和运营。

图 4 HCI 简化 IT 管理和运营



最令人信服的 HCI 解决方案利用的管理框架可显著改善运营任务、减轻生命周期管理负担以及提高 IT 员工响应敏捷性以满足其业务需求。理想情况下，也可以利用管理和编排堆栈中的现有投资。

Dell EMC 超融合基础架构平台

Dell EMC 使组织可以灵活地选择最适合其当前 IT 转型状态的 HCI 解决方案，同时在他们转向云实施时保持 IT 确定性、持续创新和可预测的发展。无论企业是要实现现有应用程序的现代化，还是要部署全包式的工程化解决方案，Dell EMC HCI 产品组合均可提供企业进行数字转型的下一阶段所需的强大功能、简易性和确定性。

下图显示了由 Intel® 提供支持的 Dell EMC HCI 产品组合。

图 5 Dell EMC HCI 产品组合



Dell EMC 超融合基础架构产品组合同时包括系统和机架级产品，而且与众不同之处是，它既可以提供基于 VMware 的全面集成解决方案，也允许客户通过选择虚拟机管理程序、裸机或多虚拟机管理程序选项而获得全包式解决方案。

系统通过基于全闪存、软件定义的横向扩展体系结构提供全包式解决方案而加快了客户数据中心的计算和存储层的转型。机架级产品提供了进一步的转型能力，适合于那些已准备好通过以全面集成、全包方式采用软件定义的网络连接以及计算和存储而全面实现数据中心现代化的客户。

Dell EMC 构建在现代数据中心的基础上，利用全新第 14 代 Dell EMC PowerEdge 服务器平台，率先为 HCI 提供专用超融合基础架构产品组合。该产品组合提供量身定制的性能和可靠性，足以适合任何工作负载，并与一种恰当的智能部署和运营方法结合，可以简化和加快 IT 运营。新一代 PowerEdge 服务器上的 Dell EMC HCI 是功能强大、有针对性且精心打造的超融合平台，可为实施软件定义的数据中心计划奠定坚实的基础。

随着 HCI 进入核心数据中心，必须运行更广泛的应用程序和工作负载，客户需要更高的性能和可靠性。在第 14 代 PowerEdge 服务器上运行的 Dell EMC HCI 效率更高（是 IOPS 的两倍）、更经济实惠（每个节点的 VDI 用户增加最多 3 倍）且灵活性更高（具有更高的可配置性，可满足更多使用情形），具备足够强劲的性能，适合任何工作负载。

Dell EMC PowerEdge 服务器为 HCI 而量身打造。PowerEdge 服务器内置多达 150 项客户 HCI 要求，专为依赖于服务器及存储原则的 HCI 工作负载而设计和量身打造。这样便可实现更加一致、可预测且可靠的高性能 HCI，从而满足任何使用情形的要求。借助我们全面的产品组合，Dell EMC 可以提供极其适合您客户的特定 HCI 要求（从工作负载要求到客户环境/标准化，再到部署偏好）的选项。

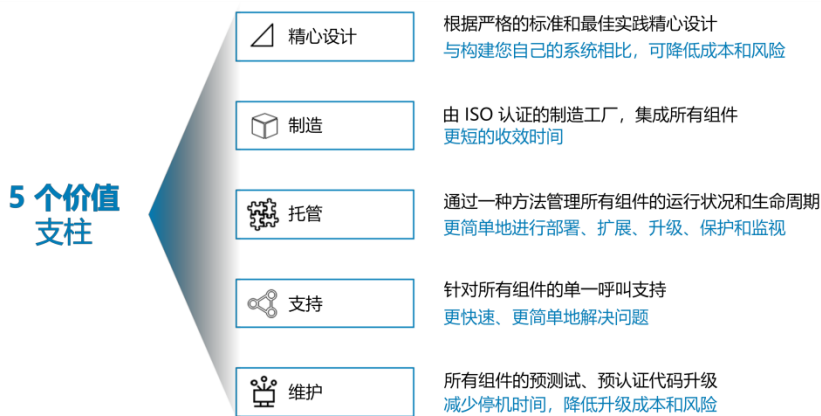
根据 IDC 统计³，Dell EMC 在超融合产品销售方面保持卓越地位，其市场份额超过 28%。越来越多的客户选择 Dell EMC HCI，而不是选择其他产品。Dell EMC PowerEdge 是全球畅销的服务器。业界卓越的 Dell EMC HCI 以出类拔萃的 PowerEdge 为基础，加上覆盖整个系统的单点支持和完整生命周期管理，已然成为一种制胜解决方案。

Dell EMC HCI 提供全包式客户体验

Dell EMC 在设计、集成和测试其超融合解决方案上投入了数千小时的工程设计时间，以确保所有组件得到强化、协同工作并实现一体化。因此，实施 Dell EMC HCI 解决方案只需几周时间即可，而不需要数月之久。

下图指出了 Dell EMC HCI 全包式体验的关键价值。

图 6 Dell EMC 全包式客户体验



Dell EMC 不断创新并提供新功能，确保其 HCI 解决方案继续发展以满足新的客户需求。我们会在解决方案发布周期中陆续推出新功能，并由 Dell EMC 对整个解决方案进行端到端的测试（包括从之前的版本进行升级）。此外，支持由 Dell EMC 全球服务部门在整个解决方案堆栈中进行重大升级，为 IT 运营团队消除互操作性管理问题。

³ 基于 IDC 融合跟踪报告，2018 年第 4 季度

Dell EMC VxRail 系统

VxRail 系统由 Dell EMC 和 VMware 联合开发, 而且是经过全面集成、预配置和测试的 HCI 系统, 采用“面向软件定义的存储”的 VMware vSAN 技术。VxRail 通过无处不在的 VMware vCenter Server 界面进行管理, 可提供熟悉的 vSphere 体验, 从而可实现精简部署并且能够扩展对现有 IT 工具和流程的使用。



VxRail 要点

全面集成、经过预先配置和测试的超融合基础架构一体机可简化并扩展 VMware 环境。

通过与现有 VMware 生态系统管理解决方案无缝集成, 精简 VMware 环境中的部署和管理。

从小配置起步, 少至三个节点。单节点扩展、存储容量扩充和 vSphere 许可证独立性可支持增长以满足业务需求。

利用包括 RecoverPoint for VMs 在内的集成数据保护选项备份分布式应用程序或工作负载。

面向硬件和软件的单一全球全天候支持点

VxRail 系统全面装载了来自 Dell EMC 和 VMware 的任务关键型集成式数据服务, 包括压缩、重复数据消除、复制和备份。VxRail 能提供弹性和集中管理功能, 可以帮助您更快、更好、更简单地管理整合工作负载、虚拟桌面、业务关键型应用程序和远程办公室基础架构。作为由 Dell EMC 和 VMware 独家提供的超融合基础架构系统, VxRail 是建立完全虚拟化 VMware 环境的极为简单快捷的途径。

VxRail 是市场上少有的将基于 Intel 的 Dell EMC PowerEdge 服务器与 VMware vSphere 和 vSAN 完全集成的 HCI 系统。VxRail 由 Dell EMC 与 VMware 联合设计, 作为单个产品提供支持, 由 Dell EMC 交付。VxRail 可与现有 (以及可选) VMware 生态系统和云管理解决方案 (包括 vRealize、NSX、Horizon、Platinum 以及属于任何庞大稳健的 vSphere 生态系统一部分的任何解决方案) 无缝集成。

VxRail 为大多数工作负载提供软件定义的数据中心 (SDDC) 的入门级选项。各种规模和类型的客户都能受益于 VxRail, 包括中小型环境、远程和分支办公室 (ROBO) 以及边缘部门, 还可为较大的数据中心提供坚实的基础架构基础。

小型 IT 部门人员可受益于系统模式的简单性, 以加快应用程序部署流程, 同时仍可利用通常仅在高端系统中可用的数据服务。

较大数据中心可受益于快速部署, 在这种情况下, 完整的 vSphere 环境可以在系统到达现场后的数小时内进行安装并准备就绪, 以部署应用程序。VxRail 使企业可以从小规模开始, 随后无中断地进行扩展。配置存储为可满足相应的应用程序容量和性能要求。

此外，节点具有不同的计算能力、内存和缓存配置，以便良好符合新的和扩展使用情形的要求。随着要求增加，系统可按精细的增量轻松地进行横向扩展和纵向扩展。最后，由于 VxRail 进行了联合设计、集成和测试，因此组织可以利用 Dell EMC 提供的单一来源支持和远程服务。

VxRail 环境配置为群集，其中最少包含两个服务器节点，每个节点都包含内部存储驱动器。VxRail 系统在交付时已加载了软件，已准备好连接到客户提供的网络。虽然大多数环境将 10Gb 以太网用于内部和外部通信，不过 25Gb 或 1Gb 以太网连接也可使用。在安装时使用一个简单向导便可按照独有的站点和网络要求来配置系统。

VxRail 系统使组织可以从小规模开始，然后随着 IT 组织转型进行横向扩展，并适应融合基础架构的管理，避免出现孤岛。通过一套丰富的数据服务（包括数据保护、分层到云以及支持双主用数据中心），VxRail 可以成为 IT 的基本基础架构。最重要的是，您只需将新系统添加到现有群集（并停用旧系统），即可提供永久的 HCI 环境，而无需再担心成本高昂的 SAN 数据迁移。随着组织继续转型到云模式，与 VMware vRealize Suite 集成可实现完整的云自动化和服务交付能力。

Dell EMC VxRail 系统提供众多的 Dell EMC PowerEdge 服务器选择，这些服务器采用全新的英特尔® 可扩展® 处理器、可变 RAM 和存储容量，使客户可按当前需求进行购买。VxRail 系统使用一种模块化、分布式系统体系结构，从最少两个节点开始，并以线性方式纵向扩展至最多 64 个节点。单节点扩展和存储容量扩展针对以后在业务和用户要求不断发展时进行纵向扩展和横向扩展提供一种“随增长而付费”的可预测方法。

VxRail 软件层将 VMware 技术用于服务器虚拟化和软件定义的存储。VxRail 节点配置为 ESXi 主机，VM 和服务使用逻辑网络的虚拟交换机进行通信。VMware vSAN 技术在 ESXi 内核级别实施，对存储资源进行池化。这一高效的 SDS 层占用极少的系统资源，从而使更多资源可用于支持用户工作负载。内核级别集成还可显著降低基础架构管理中涉及的复杂性。vSAN 向群集中的节点提供了熟悉的数据存储区，而通过基于存储策略的管理可以灵活地为每个 VM 轻松配置合适的服务级别。

VxRail HCI 系统软件 (VxRail 管理平台) 是 VxRail 的一种战略优势，可进一步降低运营复杂性。VxRail HCI 系统软件为第 0 天到第 2 天基于系统的操作任务提供立即可用的自动化和编排，从而减少管理堆栈所需的总体 IT 运营支出。无需自主构建的 HCI 解决方案可实现这一级别的生命周期管理、自动化和简化操作。

借助 VxRail HCI 系统软件，只需单击一下即可轻松进行升级。您可以坐下来放松一下，知道自己正在从一个已知良好状态转变到下一个已知良好状态，包括所有托管软件和硬件组件固件。您不再需要验证硬件兼容性列表、运行测试和开发方案、序列和试用升级等。我们已为您完成了持续发展和生命周期管理的重要任务。

在 VxRail HCI 系统软件内，VxRail Manager 插件在 vCenter Server 上提供了一个简单的集成控制面板界面，用于基础架构监视和自动执行生命周期管理任务（如软件升级）。由于 VxRail 节点用作 ESXi 主机，因此 vCenter Server 用于 VM 相关的管理、自动化、监视和安全性。

VxRail 系统采用 VMware vSAN 软件，该软件完全集成到 vSphere 的内核中并提供功能齐全且经济高效的软件定义的存储。vSAN 实施了一种高效的体系结构，它直接内置在虚拟机管理程序中。这可将 vSAN 与通常安装作为来宾 VM 在每个主机上运行的虚拟存储设备 (VSA) 的解决方案区分开来。将 vSAN 嵌入 ESXi 内核层在性能和内存要求方面具有优势。它对 CPU 利用率的影响很小（不到 10%），可基于工作负载和资源可用性进行自我平衡。它可为存储提供熟悉的数据存储区构造，并且可无缝地与其他 vSphere 功能（如 VMware vSphere vMotion）配合工作。

vSphere 是一个成熟的虚拟化平台 — 在大多数数据中心内是熟悉的可用实体。Dell EMC 将 vSphere 用于多个产品组合中的基于 ESXi 的虚拟化和虚拟机网络，并且它们支持一组通用的 VMware 和 Dell EMC 服务。这使得 VxRail 实施可以顺畅地集成到以 VMware 为中心的数据中心，并与 Dell EMC 融合、超融合和传统存储产品/服务协同工作。NSX for SDN 也可以作为一个选项添加到 VxRail 解决方案。NSX for SDN 便于使用单个管理平台维护所有数据中心资产。高效且可靠地执行监视、升级和诊断活动。

集成到 VxRail 的其他数据服务包括 RecoverPoint for VM 复制和 Dell EMC Remote Secure Services (ESRS)。

总之，VxRail 建立了 IT 确定性。它基于经企业验证的 VMware 虚拟机管理程序，采用 vSAN 技术，可提供简化的生命周期和日常运营管理，并支持动态稳健的 vSphere 生态系统。系统体系结构旨在随业务一起以可预测的方式发展。系统的扩展方式是部署预先加载的系统，以便轻松地扩展环境，而不是构建并手动集成服务器。

由 vSAN 提供支持并采用基于 Intel 的 Dell EMC PowerEdge 服务器的 VxRail 在继续发展，因此企业可以成长壮大而无需担心 IT 基础架构。VxRail 是全面集成、预先配置且经过预先测试的 VMware 超级融合基础架构系统系列。VxRail 系统基于 VMware 的 vSphere 并采用 vSAN 和 Dell EMC 软件，可为数字转型提供 IT 基础架构基础，使组织能够不断创新。

VxRail 安全性与合规性

Dell EMC VxRail 系统是一种弹性、安全和现代化的超融合基础架构系统，可直接解决当今环境中的安全性和合规性难题。

VxRail 系统按照 Dell EMC 安全开发生命周期进行设计、构建、配置和维护，在产品上市之前遵循严格的方法来保护产品开发，包括高管级别风险管理。此外，VMware vSphere 是 VxRail 超融合基础架构的一个重要部分，也使用类似的安全开发生命周期进行开发。

构成 VxRail 的一切都是安全的，可以在下图中看到。每个组件都有内置的安全功能，包括公司安全流程、独特的安全功能和供应链控制，因此您可以放心，VxRail 可以适应您的安全 IT 基础架构设计。硬件包括 Dell EMC PowerEdge 服务器和英特尔处理器。虚拟化和软件层包括集成到 vSphere 内核的 vSphere 和 vSAN。VxRail 附带的集成软件和管理组件包括来自 VMware 的软件、vRealize Log Insight 和 vCenter，以及来自 Dell EMC 的软件、RP4VM、ESRS 和 VxRail HCI 系统软件，以实现对此堆栈中的所有组件的生命周期管理。（RP4VM 不包括在 LCM 中。）所有这些都由 Dell EMC 和 VMware 联合设计，并作为单个产品 (VxRail) 由 Dell EMC 全面负责交付和支持。

图 7 VxRail 包括的组件



VxRail 满足众多标准，获得了通用标准 EAL2+ 证书、USGv6 认证，这使其适用于 IPv6，并提供 VxRail 产品安全配置指南，进一步强化 VxRail 部署。此外，客户还可以利用 VxRail STIG 合规性指南和自动化脚本来进一步强化其环境。

要详细了解 VxRail 全面的安全性设计，请下载白皮书：

<https://China.emc.com/collateral/white-papers/vxrail-comprehensive-security-design.pdf>

VxRail 硬件体系结构

Dell EMC VxRail 系列是超融合基础架构中的标准，提供极高的灵活性以精细地按需添加容量和性能并使客户可以在整个 VMware 虚拟化环境中轻松地扩展使用情形。基于系统的设计使 IT 中心可以无中断地扩展容量和性能，因此它们可以从小规模开始，并以极少的预先规划逐步扩展。VxRail 环境可以设计为支持少量虚拟机并扩展到数千个虚拟机。

VxRail 体系结构支持可预测的随增长付费方法，可与不断变化的业务目标和用户需求保持一致。Dell EMC 和 VMware 在持续进行创新，并且 VxRail 引入基于 Dell EMC PowerEdge 的新型号，可提供极高的配置灵活性。这种灵活性使客户可以根据需要为 VMware 环境选择性能、显卡和容量，并且支持更多使用情形。

Dell EMC VxRail 的系统系列提供了一系列平台：

G 系列 — 通用和计算密集的多节点外形规格，非常适合广泛部署的通用应用程序和 VDI 工作负载（不需要 GPU 卡）

E 系列 — 从数据中心核心到边缘部署，借助自身的密度、驱动器组以及在半高 1U 外形规格中的均衡资源组合，它能够在各种使用情形下进行部署。

P 系列 — 针对具有业务关键型、性能密集型应用程序和/或内存中数据库的高端使用情形进行了性能优化。

V 系列 — 针对具有图形密集型应用程序（如高端 2D/3D 可视化应用程序）的专业化使用情形进行了 VDI 优化，它是唯一支持 GPU 卡的系列。

S 系列 — 针对在服务器级别需要较高存储容量的专业化使用情形（如大数据、分析或协作应用程序）的存储密集型配置。

E、P、V 和 S 系列是基于 Dell EMC PowerEdge 服务器技术（销量最好的 X86 服务器平台）的单节点系统，具有更高存储容量、更大内存和更强大的 CPU 选项。G 系列是集于紧凑型 2U 机箱中的四节点系统，可提供高密度计算，占用空间小。

VxRail 系统使用一种分布式群集体系结构进行构建，该体系结构由可以随系统增长而线性扩展（从小至 3 个节点，到大至 64 个节点）的模块化构造块组成。节点具有不同外形规格，单节点系统适用于以下使用情形：半高系统；性能优化；VDI 优化，带 GPU；以及支持大容量 HDD 驱动器的存储优化配置。

广泛的计算、内存和存储选项旨在适应多种使用情形。客户可以从一系列新一代英特尔处理器、可变内存大小、存储和缓存容量中进行选择，从而提供计算、内存和存储的适当平衡。单节点扩展和低成本入门级选项使客户可以仅购买适量的存储和计算能力来满足当今的要求，并轻松进展以适应未来增长。系统采用全闪存存储配置，这些配置可为需要最高性能和低延迟的应用程序提供业界强大的 HCI。

VxRail 群集

VxRail 节点封装在单节点单服务器系统中，每个节点由一个或两个多核处理器以及全闪存固态硬盘 (SSD) 或闪存 SSD 和硬盘驱动器 (HDD) 混合组成。节点组成联网群集，其中最少有三个节点，最多 64 个节点。群集中的节点必须属于相同存储配置 (全混合或全闪存)。支持灵活地在群集内混用节点。前三个节点必须具有相同的计算、内存和存储配置，并且不支持混用 1GbE、10GbE 和 25GbE。从最低配置到最高配置，VxRail 群集可轻松地一次扩展一个节点。

系统型号支持 25GbE、10GbE 或 1GbE 网络。对于全闪存配置和将要扩展到八个以上节点的环境，需要 10Gb 和 25Gb 以太网网络。提供了其他端口，从而使客户能够扩展虚拟机网络流量。

VxRail 型号和规格 (基于第 14 代 Dell EMC PowerEdge 服务器)

构建在新的第 14 代 Dell EMC PowerEdge 服务器平台上的 VxRail 系统提供您的客户实现广泛工作负载所需的性能和可靠性，同时通过单点支持实现完整的生命周期管理。简言之，VxRail 是实现基础架构转型的极为快速简单的方法。设计高性能和可靠的 HCI 解决方案需要做大量工作，而且还要具备专业知识，并且在初始部署后工作也不会停止。还需要持续验证，以使其能够通过软件升级和添加节点实现平稳运行。作为全包式、预先集成、经过测试和验证的 HCI 解决方案，VxRail 可以快速部署、轻松地分发，并且应该可以提高 IT 环境的可预测性、可用性和性能。

下一代服务器上的 VxRail 系统包括多种专门构建的平台，采用照单生产配置，支持范围广泛的客户使用情形，包括图形密集型 VDI、大数据和分析、高性能计算、远程办公室等。借助更多的处理器选项、新的 SATA SSD、更多附加网络连接选项以及更多的 GPU 扩展，现在您可以将 VxRail 与工作负载的要求更密切地匹配起来。这种情况下不会出现过度调配：您可以在需要时仅购买所需的内容。

提供各种 VxRail 型号来满足一系列广泛的使用情形的要求。对于较小的工作负载，我们提供使用 1U 单节点系统的半高系统配置，可以节省空间。针对性能优化和针对 VDI 优化的配置采用全闪存和混合配置。对于需要甚至更大存储的使用情形，提供了使用较大容量 3.5 英寸驱动器的混合存储密集型配置。所有型号都具有广泛的可用内存、SSD 缓存和容量存储配置选项，可以从最少两个节点开始。

第 14 代 PowerEdge 服务器上的 VxRail 现在比上一代更强大、更具可预测性，可实现更高的灵活性，以满足 VMware 环境中的任何使用情形和更严苛的工作负载要求。

IOPS 翻倍，响应时间减半 — VxRail
 足够强大，适合任何工作负载






持续响应时间低于 1 毫秒 — VxRail
 可实现高度可预测的性能

Dell EMC 提供极易配置的 HCI
 系统 — VxRail 完全符合任何 HCI
 要求



下图展示了一系列支持多种使用情形的
 的平台。

图 8 基于第 14 代 Dell EMC PowerEdge 服务器的 VxRail

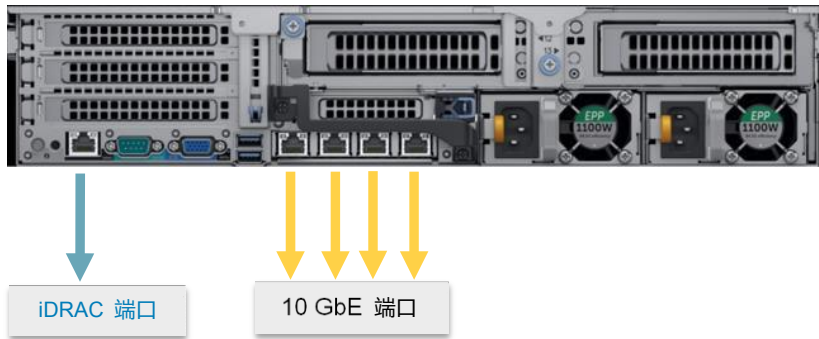
| G 系列节点 | E 系列节点 | P 系列节点 | V 系列节点 | S 系列节点 |
|--|---|---|--|---|
|  |  |  |  |  |
| 计算密集型 | 半高 | 优化了性能 | 优化 VDI | 存储密集型 |
| G560/F/N | E560/F/N | P570/F/N | V570/F | S570 |
| 支持全新一代基于英特尔® 至强® 第 1 代和第 2 代处理器可扩展系列的 Dell EMC PowerEdge 服务器 更多的内核数量、更快的时钟频率、更多的内存通道、更快的内存速度、更高的耐用性和冗余引导设备 | | | | |
| 2000W 或 2400W PSU 10 GbE 傲腾和 NVMe 缓存 混用 SAS 缓存 | 1100W PSU 10GbE 或 25GbE 傲腾和 NVMe 缓存 混用 SAS 缓存 FC HBA | 1100W 或 1600W PSU 10GbE 或 25GbE 傲腾和 NVMe 缓存 混用 SAS 缓存 FC HBA | 2000W PSU 10GbE 或 25GbE 多达 6 个 GPU 混用 SAS 缓存 FC HBA 仅双插槽 | 1100W PSU 10GbE 或 25GbE 仅混合机型 混用 SAS 缓存 FC HBA |

VxRail 节点

VxRail 系统装配有经验证的服务器节点硬件，该硬件由 Dell EMC 作为完整的解决方案进行了集成、测试和验证。当前一代 VxRail 中的所有节点都使用英特尔至强可扩展系列处理器。英特尔至强可扩展系列处理器采用多线程的多核 CPU，旨在为云服务、高性能计算和网络处理各种工作负载。核心数和内存容量因每个 VxRail 型号而异。

下图显示一台节点服务器及其处理器、内存和支持组件的物理视图。所有 VxRail 型号都具有类似组件，但可能会以不同方式进行物理布局。

图 9 VxRail P 系列节点服务器后视图



每个节点服务器包括以下技术：

一个或两个英特尔至强可扩展处理器，每个处理器具有最多 28 个核心

最多 24 个 DDR4 DIMM，对每个节点可提供从 64 GB 到 3072 GB 的内存容量，具体取决于型号

一个支持 12 GB SAS 速度的 PCIe SAS 磁盘驱动器控制器

一对镜像的 BOSS 240GB SATA M.2 卡，用于在“四端口 10GbE 网络子卡”（可自动协商到 1GbE）节点上启动 ESXi

iDRAC 端口

英特尔® 至强® 可扩展处理器：VxRail 强大的处理能力

英特尔® 至强® 可扩展平台是功能强大的基础架构，代表了敏捷性和可扩展性方面的飞跃。它采用颠覆性创新设计，在计算、存储、内存、网络 and 安全性方面为平台融合和功能树立了新的基准。英特尔® 至强® 可扩展处理器采用创新的平台设计方法，为当今的数据中心和通信网络带来可扩展性能 — 从最小型工作负载到任务关键型应用程序，全面覆盖。

凭借多达 28 个核心（每个核心实现高度增强的性能），而且内存带宽（六个内存通道）和 I/O 带宽（48 条 PCIe 通道）显著增加，对数据需求量大、对延迟敏感的应用程序（例如内存中的数据库和高性能计算）将通过更密集的计算和更快速的大型数据卷访问，实现明显的改进。而且名称中带有“M”的最新一代处理器可支持更密集的内存，每个处理器可支持多达 1536 GB 内存。

计算、内存、网络和存储性能的融合与软件生态系统优化相结合，使得英特尔® 至强® 可扩展平台非常适合可基于工作负载需求，在本地、通过网络和在公有云中动态地进行资源自助调配的完全虚拟化的软件定义的数据中心。

VxRail 节点存储磁盘驱动器

VxRail 系统的存储容量由 Dell EMC 进行了集成、测试和验证的磁盘驱动器提供。大多数 VxRail 配置使用 2.5 英寸外形规格的 SSD 和机械 HDD，而且还为密集存储需求提供使用 3.5 英寸外形规格驱动器的配置。磁盘驱动器以逻辑方式组织为磁盘组。磁盘组采用两种方式进行配置：



Intel Inside®。外部可信云。

Intel 的创新正在推动传统企业数据中心向现代化和混合云转型。

迁移到新一代基于英特尔的高性能和高能效硬件，可以优化数据中心，以便在广泛的企业工作负载范围内获得高度优化的性能，同时降低成本并提高资源利用率。

随着时间的推移，数据中心的所有关键域（计算/存储/网络）发展到软件定义的基础架构 (SDI)，这样就可提供关键的自动化、编排和遥测功能，帮助企业充分释放多云计算的全部潜力。

凭借在软件定义的基础架构上运行的现代化行业标准英特尔® 服务器和技术，您可以无缝地管理支持云原生应用程序开发和交付的环境和位于安全私有云上的任务关键型工作负载，同时还可与公有云集成，有许多公有云已经在英特尔® 体系结构上运行。

混合配置, 包含用于高速缓存的单 SSD 基于闪存的磁盘 (高速缓存层) 以及用于容量的多个 HDD 磁盘 (容量层)

全闪存配置, 包含用于高速缓存和 SAS 的所有 SAS SSD 或 NVMe 驱动器, 或用于容量的 SATA SSD。仅在双处理器配置中才支持 NVMe。

用于缓存和容量的闪存驱动器具有不同的耐用性水平。耐用性水平指在必须更换之前, 在五年内每天可以写入整个闪存磁盘的次数。耐用性较高的 SSD 用于写缓存, 经过容量优化的 SSD 用于容量。目前, 400 GB、800 GB 和 1600 GB SAS SSD 或 1600 GB NVMe 用于缓存, 而 1.92 或 3.84 TB 闪存 SSD、1.2、1.8 和 2.4 TB 10K HDD、2 TB (7.2K) HDD 和 4 TB 7.2K (3.5 英寸外形规格) 用于容量。所有 VxRail 磁盘配置都使用精心设计的缓存/容量比, 以确保一致的性能。较高耐用性 SAS 和 SATA 中均提供容量 SSD。SATA SSD 是一种成本较为低廉的选择, 每个驱动器的成本最高可减少 30%, 非常适合读取工作负载或适度密集的工作负载。下图说明了在 VxRail 系统产品中可用的广泛组件集。

VxRail 硬件选项

VxRail 节点可配置任意的处理器、内存、存储 (高速缓存和容量驱动器)、基本网络、电源和 GPU (仅在 V 系列中)。下图显示了在整个系列中提供的一组全面的选项。客户可以放心, VxRail 系列中有数百万种可能的配置组合, 我们会根据他们的工作负载要求以一种非常规范的方式来配置 VxRail。将此与众多按需扩展方法相结合, 很明显 VxRail 提供了如今的现代 IT 所要求的敏捷性。VxRail 的一些升级选项包括内存、GPU、NIC 卡、缓存 SSD 和容量驱动器。

图 10 VxRail 系列中提供的一系列代表性组件选项

适合您的工作负载的 VxRail 配置灵活性

G、E、P、S、V 系列基于最新 Dell EMC PowerEdge 服务器

处理器

可选择 32 种英特尔® 可扩展® 第 2 代处理器
可选择 40 种英特尔® 可扩展® 第 1 代处理器

每个系统配置 4-56 个内核

RAM

24 个 DIMM 插槽
16 GB RDIMM
32 GB RDIMM
64 GB LRDIMM、RDIMM
128 GB LRDIMM

电源

| | |
|----------------------|---------------|
| 1100 W | 100-240 V 交流电 |
| 1600 瓦、2000 瓦、2400 瓦 | 200-240 V 交流电 |
| 1100 W | 48V DC |



存储

缓存驱动器: 傲腾 375 GB 和 NVMe 1600 GB
SAS 400 GB、800 GB、1600 GB

容量 SSD (SAS 和 SATA): 1.92 TB、3.84 TB
硬盘驱动器: 1.2 TB、1.8 TB、2.4 TB、2.0 TB、4.0 TB

基本网络连接

SFP28、SFP+、RJ45
2 个 25 GbE
4 个 10 GbE
2 个 10 GbE
4 个 1 GbE (4 个 10 GbE, 自动协商)
可选附加 NIC、FC HBA

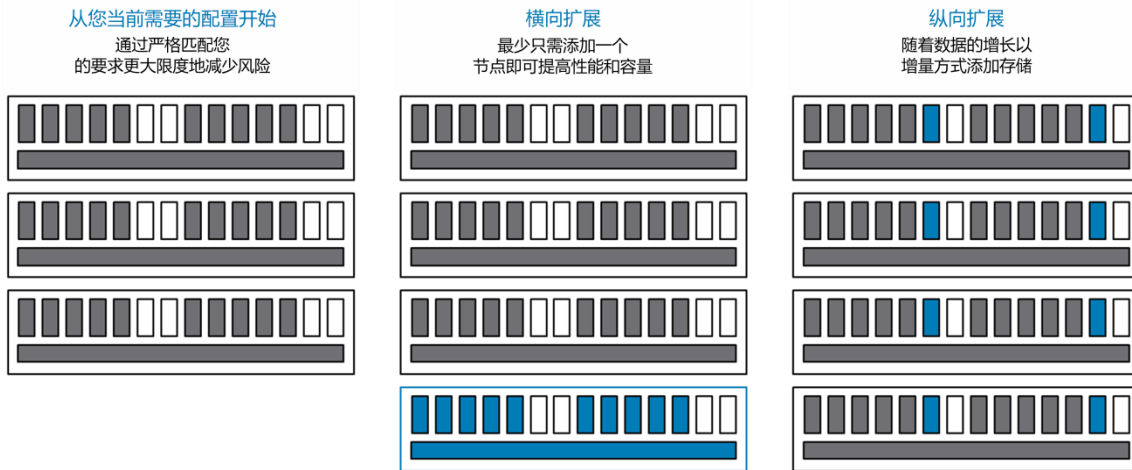
GPU

NVIDIA V100、T4、P40、M10
注意: GPU 软件和驱动程序另售

VxRail 扩展

VxRail 群集从少至三个节点开始，可以按一个节点的增量扩展到最多 64 个节点，从而提供性能和容量以满足广泛的使用情形。支持两节点 VxRail 群集，但目前无法扩展。可以无中断地添加新系统，并且可以在 VxRail 群集中混合使用不同型号的系统。灵活的存储选项还使节点可以从几个驱动器开始，随着容量要求的增加而添加驱动器，如下图所示。单节点升级和驱动器可扩展性通过使客户可以从所需配置开始来保护优化的初始投资，并通过根据需要添加节点和/或驱动器以提高性能和容量来扩展 VxRail 群集。咨询 Dell EMC 代表以获得帮助。

图 11 VxRail 按需扩展



在进行计划时需要考虑有关扩展的一些基本规则：

1. 平衡

- 群集中的前三个节点必须具有相同的处理器和驱动器配置。(不支持 2 节点 vSAN 配置。)
- 所有节点都必须运行相同版本的软件。
- 在同一群集内不可以混用混合节点和全闪存节点。
- 不能在同一群集中混用 1GbE、10GbE 和 25GbE 基本网络。
- 1GbE 必须是混合和单处理器节点类型。
- 对于 G 系列，同一机箱中的所有节点必须完全相同。

2. 灵活性

- 群集中的系统可以是不同的型号或系列，并且可以具有不同数量的节点。
- 群集可以有不同数量的驱动器、CPU、内存和型号类型。
- 一个群集可以有 3-64 个节点，但在使用 1GbE 网络时最多只能有 8 个节点。
- 对于 G 系列，机箱可以不完全填充。

可升级的选项

使用 VxRail，可以升级节点或添加内存、NIC 卡、高速缓存驱动器以及容量驱动器。在 V 系列中可以升级或添加 GPU。不能从单处理器升级到双处理器 VxRail 节点。请参阅下表，了解客户可安装（可更换）哪些组件。

图 12 VxRail 客户和现场可更换部件

| 硬件组件 | 客户可换部件 (CRU) | 现场可换部件 (FRU) |
|------------------------|--------------|--------------|
| 系统内存 | 是 | |
| 硬盘驱动器 | 是 | |
| 固态硬盘（缓存和容量） | 是 | |
| NVMe 缓存驱动器 | | 是 |
| PCIe 网络接口卡 | 是 | |
| 图形处理单元 (GPU) | 是 | |
| Micro SDHC 卡 | 是 | |
| 电源 | 是 | |
| 处理器 | | 是 |
| 系统主板 | | 是 |
| 主机总线适配器 (HBA330) | 是 | |
| BOSS 控制器卡和 M.2 SATA 磁盘 | | 是 |
| 网络子卡 (NDC) | | 是 |

*上表是一个非详尽的 FRU 清单，体现了常见顶层装配部件。

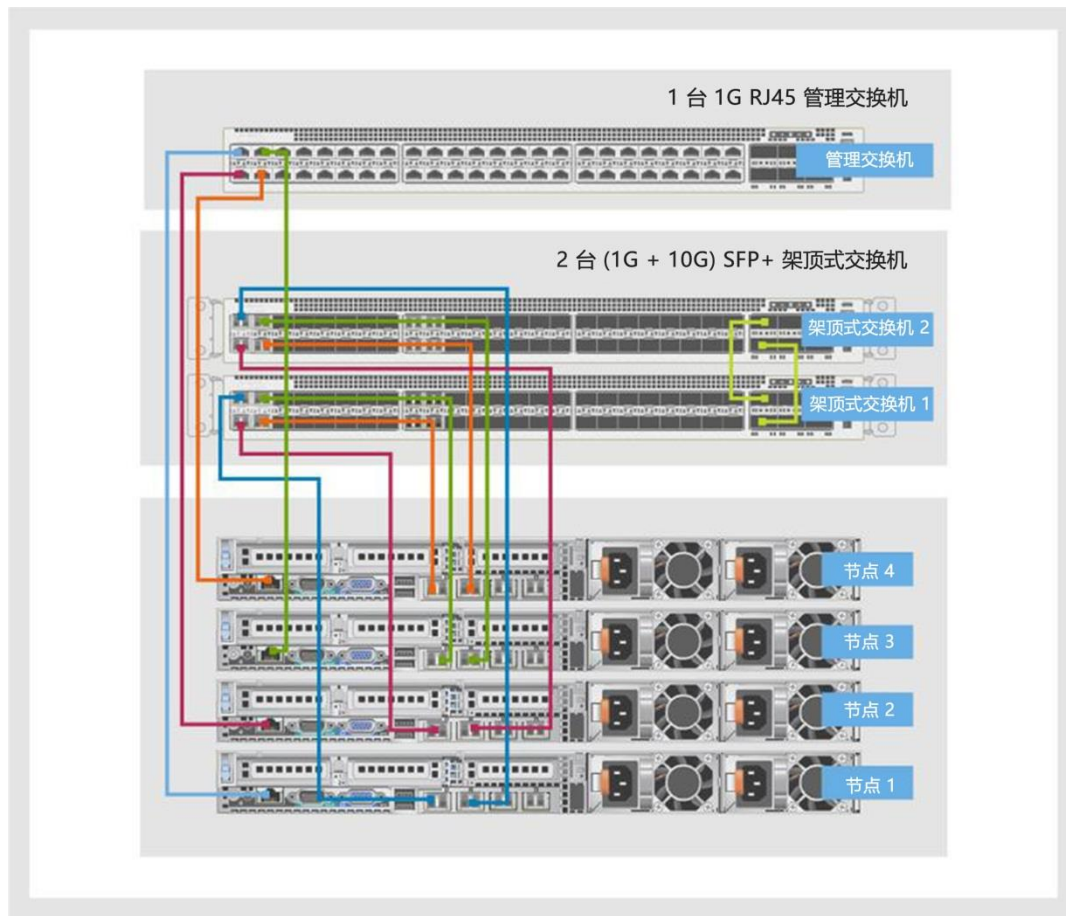
VxRail 网络

VxRail 系统是一种功能完备的环境，具有组成超融合基础架构的计算、存储、服务器虚拟化和管理服务。分布式群集体系结构使独立节点可以作为单个系统协同工作。每个节点都同时提供并消耗系统资源。节点之间的这种紧密耦合通过 IP 网络连接来实现。通过 IP 网络还可访问虚拟机以及它们所提供的服务。

尽管 VxRail 是一种功能完备的基础架构；但它不是一个独立的环境。它旨在与客户现有数据中心网络进行连接和集成。典型实施使用一个或更多客户提供的 10GbE 架顶式 (ToR) 交换机连接 VxRail 群集中的每个节点。对于较小的环境，提供了一个使用 1GbE 交换机的选项，但是这些较低带宽的网络会限制性能和规模。虽然网络交换机通常由客户提供，不过 Dell EMC 提供了可以随系统附带的 S4048 以太网交换机。

下图显示使用两个交换机实现冗余的典型网络连接。还支持单交换机实施。

图 13 用于 10GbE 配置的典型 VxRail 物理网络连接



所需以太网交换机端口数取决于 VxRail 型号以及它是配置为混合存储还是配置为全闪存。全闪存系统需要两个或四个 10GbE 端口，混合系统对每个节点使用两个 10GbE 端口，或对每个节点使用四个 1GbE 端口。对于 1GbE 网络，10GbE 端口将自动向下协商至 1GbE。还可以选择两端口 25GbE SFP28。通过添加额外的 NIC 卡可以实现额外的网络连接。额外的 PCIe NIC 不通过 VxRail 管理进行配置，但可供客户用于支持非 VxRail 流量（主要是虚拟机流量）。附加端口通过 vCenter 进行管理。

网络流量使用基于交换机的 VLAN 技术和 vSphere 网络 I/O 控制 (NIOC) 进行隔离。VxRail 群集中存在四种类型的网络流量：

管理。管理流量用于连接到 VMware vCenter Web Client、VxRail Manager 和其他管理接口，并用于在管理组件与群集中的 ESXi 节点之间进行通信。默认 VLAN 或特定管理 VLAN 用于管理流量。

vSAN。用于读取和写入活动以及用于优化和数据重建的数据访问通过 vSAN 网络执行。低网络延迟对于此流量十分关键，通过特定 VLAN 隔离此流量。

vMotion。 VMware vMotion™ 允许虚拟机在节点之间迁移。使用单独 VLAN 隔离此流量。

虚拟机。 用户访问虚拟机和 VM 网络上提供的服务。最初配置系统时会至少配置一个 VM VLAN，并且可以根据需要定义其他 VLAN。

安装前计划包括验证是否有足够物理交换机端口可用以及是否针对合适的 VLAN 配置了端口。在安装期间配置系统时会使用 VLAN 以及 IP 地址和其他网络配置信息。详细计划和配置信息包含在 [《VxRail Network Guide》 \(VxRail 网络指南\)](#) 中。

在安装期间初始化系统时，配置向导会自动按照 VxRail 标准和最佳做法配置所需上行链路。向导会要求提供 NIC 配置并接受两个选项：

4 个 1GbE。这仅对带有单个处理器和使用混合存储配置的系统有效。四 (4) 个 10GbE 端口自动向下协商至 1GbE。管理、vSAN、vMotion 和 VM 流量通过适当的网络绑定策略和 NIOC 设置与这些端口相关联。

2 个 10GbE。管理、vSAN、vMotion 和 VM 流量通过适当的网络绑定策略和 NIOC 设置与这些端口相关联。请注意，附加虚拟机流量可使用附加的两个端口。

4 个 10GbE。管理、vSAN、vMotion 和 VM 流量通过适当的网络绑定策略和 NIOC 设置与这些端口相关联。

2 个 25GbE。管理、vSAN、vMotion 和 VM 流量通过适当的网络绑定策略和 NIOC 设置与这些端口相关联。

如果节点具有其他物理 NIC 端口，则它们可以在安装之后使用标准 vSphere 过程进行配置。

1GbE 网络选项

以太网网络不仅提供与 VM 和服务之间的连接，它还为超融合基础架构中的节点提供底板以聚合和共享系统资源。因此，网络带宽对于系统规模和性能十分关键。当前，大多数数据中心采用 10Gb 以太网连接进行构建，但在许多环境中仍存在 1GbE。为了支持这些环境，Dell EMC 还为要求较低的较小工作负载提供 1GbE VxRail 型号。以下是使用 1GbE 连接选项的注意事项：

混合存储配置仅支持 1GbE，因为它未提供所需带宽来实现全闪存系统所需的超高性能。

支持的最大节点计数是每个群集八个节点，因为 vSAN 流量会随节点数而增加。

仅支持具有单插槽 CPU 的节点。

每个节点需要至少四个端口。这会增加所需的交换机端口总数。

Dell EMC 开放式网络和 VxRail

随着超融合群集的扩展，网络结构成为成功部署的关键部分。Dell EMC 结构提供：

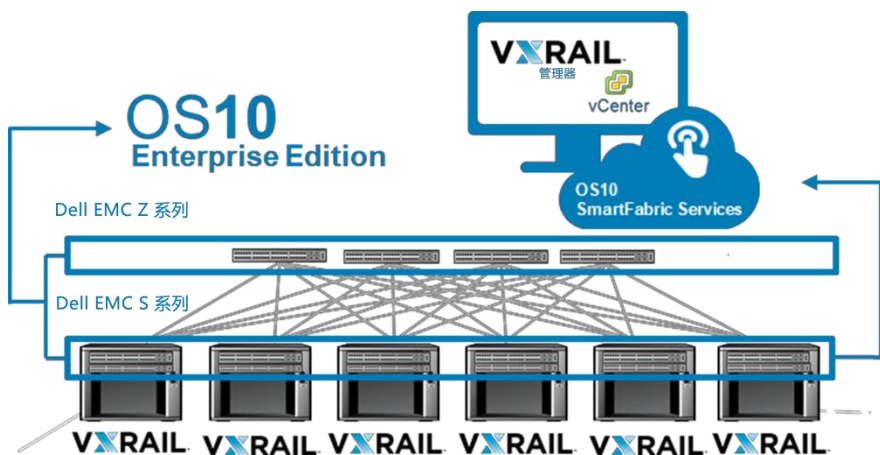
- 按需可扩展性：为了保持竞争力，现代数据中心需要能够根据业务需求动态增长和缩减。Dell EMC Networking 交换结构和 Dell EMC VxRail 共同提供了一种智能且功能强大的体系结构，可按需扩展并提高数据中心的效率。
- 提高大规模可用性：强大而冗余的结构和存储是当今数据中心所必需的。单个故障不应导致服务完全中断。

Dell EMC 结构解决方案提供了经验证、完全集成和适应性强的基础架构。我们的解决方案利用超融合体系结构的组件和专用网络基础架构元素，提供久经考验的完全集成且适应性强的基础架构，专门用于满足超融合和高效数据消耗的要求。

Dell EMC SmartFabric Services (SFS)

Dell EMC Networking 和 VxRail 通过为软件定义的企业提供精心设计的解决方案来推动产品和解决方案的发展，这些产品和解决方案可提供以前无法提供的运营和基础架构效率。Dell EMC 是一家拥有全套硬件和软件产品的供应商，能够为数字转型提供必要的工具。随着 VxRail 4.7 和 Dell EMC Networking OS10 Enterprise Edition SmartFabric Services 的推出，可以最终实现围绕自动化动态基础架构部署的对话。Dell EMC 使用 VxRail 与 SmartFabric 搭配，是在设置、群集扩展和日常管理期间提供完全自动化的网络感知和配置，以帮助促成 IT 确定性的少有的解决方案。

图 14 SmartFabric Services



SmartFabric Services 是 Dell EMC OS10 企业版旗舰网络操作系统的独特功能部件。它的引入在结构和超融合群集基础架构 (如 VxRail) 之间创建了一个完全集成的解决方案。借助 SmartFabric Services，客户可以快速轻松地部署和自动化数据中心网络结构。这样可以缩短任何规模的超融合和私有云环境投入生产前的准备时间，同时可与现有数据中心基础架构完全互操作。

VxRail 软件体系结构

前面的部分介绍了基于经验证的服务器技术的灵活 VxRail 硬件体系结构。虽然 VxRail 硬件可帮助区分 VxRail 与其他 HCI 解决方案，不过 VxRail 是完整的系统，其中包括可实现软件定义的数据中心的软件。有关软件体系结构的各部分提供所有 VxRail 软件组件及其关系和相互依赖性的全面探讨。

VxRail 系统在构建时配备了用于系统管理、虚拟化和 VM 管理的软件堆栈。堆栈已经预安装好，只需在现场运行配置向导便可将系统集成到现有网络环境中。包括了 VxRail HCI 系统软件，用于系统管理、操作和自动化。VMware 虚拟化和虚拟基础架构管理软件包括：

- VMware vCenter Server
- vSphere ESXi
- vSAN (软件定义的存储)
- VMware vRealize™ Log Insight™

其他 Dell EMC 软件包括：

RecoverPoint for VM — 每个节点 5 个虚拟机许可证 (对于单节点系统)，每个机箱 15 个虚拟机 (对于 G 系列)

VxRail 可为 VMware 环境提供紧密集成的独有体系结构。VxRail 深入集成了 VMware 虚拟化软件。具体而言，VMware vSAN 在内核级别进行了集成，可使用 VMware vSphere 进行管理，VMware vSphere 可为 VxRail 系统实现更高性能以及自动扩展和基于向导的升级。

系统管理

本技术手册的简介部分讨论了软件定义的数据中心的复杂性以及管理和维护 SDDC 环境的挑战，而 VxRail Manager 可直接应对此挑战。

VxRail HCI 系统软件

VxRail HCI 系统软件 (VxRail 管理平台) 是 VxRail 群集的系统硬件生命周期管理和可维护性接口。它是 VxRail 的一种战略优势，可进一步降低运营复杂性。VxRail Manager 为第 0 天到第 2 天基于系统的操作任务提供立即可用的自动化和编排，从而减少了管理堆栈所需的总体 IT 运营支出。无需自主构建的 HCI 解决方案可实现这一级别的生命周期管理、自动化和简化操作。

所有虚拟化管理都使用 vCenter 执行。
VxRail HCI 系统软件不执行任何虚拟化管理。

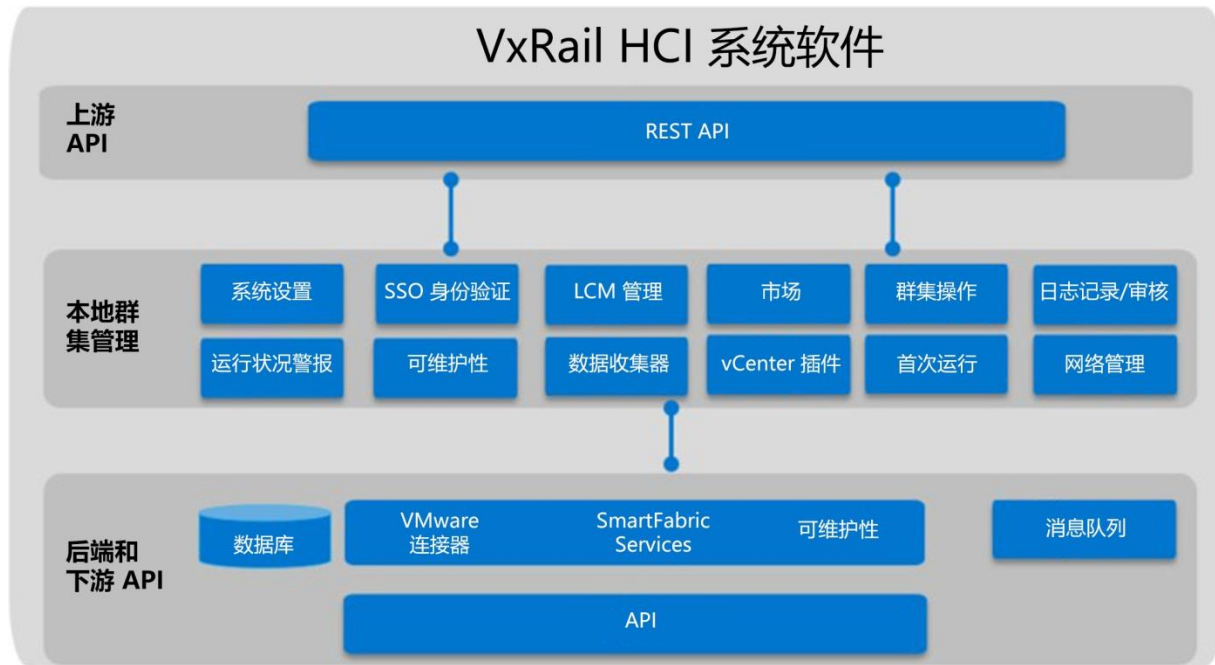
图 15 VxRail HCI 系统软件



为了向客户提供这一级别的独特价值，VxRail HCI 系统软件依赖一组后端服务来收集信息，并使用包括 vSAN 群集在内的底层基础架构组件（即 ESXi、vSAN、vCenter Server）、Dell PowerEdge 服务器和可维护性平台（即 ESRS 和 eServices）来提供事件协调。通过此集成，VxRail 可以将基础架构流程自动化并编排到 HCI 客户高度重视的关键服务（即生命周期管理、单程序包硬件/软件升级、自动化部署功能）中。用户可以通过 VxRail Manager 插件或通过一组 RESTful API 从其 vCenter Server 访问这些群集管理服务。如果客户（例如，服务提供商）希望大规模部署和管理 VxRail 群集，则通过提供公共 API，可以将 VxRail for HCI 的价值扩展到 SDDC 解决方案，例如 VMware Cloud Foundation、自定义云解决方案（例如 Puppet、Ansible）或脚本化解决方案。

对 VxRail 管理平台不熟悉是客户实现 IT 转型之旅的另一个关键因素。虽然运营简单性对运营支出的影响很大，但基础架构机器学习也迅速成为促进 IT 转型的重要因素。VxRail 分析咨询引擎 (ACE) 是新推出的分析平台，它利用从 VxRail 群集收集的数据和最佳做法为客户提供有关其 HCI 环境的运营情报。

图 16 VxRail HCI 系统软件的体系结构



上面的体系结构示意图描绘了 VxRail HCI 系统软件。自下而上，后端服务包括将下游连接到基础架构层中各成员的 API。VMware 连接器提供与构成 vSAN 群集的组件之间的通信。SmartFabric 服务为 VxRail 管理平台提供自动化网络资源调配功能。有关 SmartFabric 服务的更多详细信息，请参阅[此部分](#)。要在 VxRail 中提供可维护性功能，需要建立与 ESRS 和 eServices 的通信。对于 VxRail 收集的所有不同类型的信息，使用数据库来存储指标，并使用消息队列来帮助实现流程的自动化和编排所需的一系列事件/事务。

本地群集管理层是一组专有服务，旨在为客户提供运营简单性和情报带来的好处，以便更有效地管理其 VxRail 群集。

- 系统设置 — 硬件存储模块和组件的状态与信息传播到 VxRail
- SSO 身份验证 — 与 vCenter Server 单点登录服务集成
- 生命周期管理 — VxRail 的硬件和软件组件的端到端升级
- Marketplace — 从一个统一区域访问以下载 VxRail 生态系统软件
- 群集操作 — 群集扩展或节点删除服务
- 日志记录/审核 — 能够将数据发送到 vRealize Log Insight 的系统日志记录和审核服务
- 运行状况警报 — 组件运行状况的通知
- 可维护性 — 访问 ESRS 以获得客户支持，以及访问 eServices 以获得产品知识

- 数据收集器 — 收集群集指标并将其发送到 VxRail ACE 以进行基础架构机器学习
- vCenter 插件 — 用于本地群集管理的 VxRail Manager 界面
- 首次运行 — 专业服务的自动部署向导，以便在第 0 天提供 VxRail
- 网络管理 — 由 SmartFabric 服务启用的自动网络资源调配服务

希望利用这些服务在本地管理其 VxRail 群集的客户可以在其 vCenter Server 上使用 VxRail Manager 插件。还有一些上游 API 提供这些功能的子集（如生命周期管理和群集操作服务），以供希望为云服务交付解决方案编排 HCI 资源调配的云解决方案开发人员使用。对于希望编写自己的解决方案以便大规模地管理 VxRail 群集的客户来说，这些 API 也很有益。

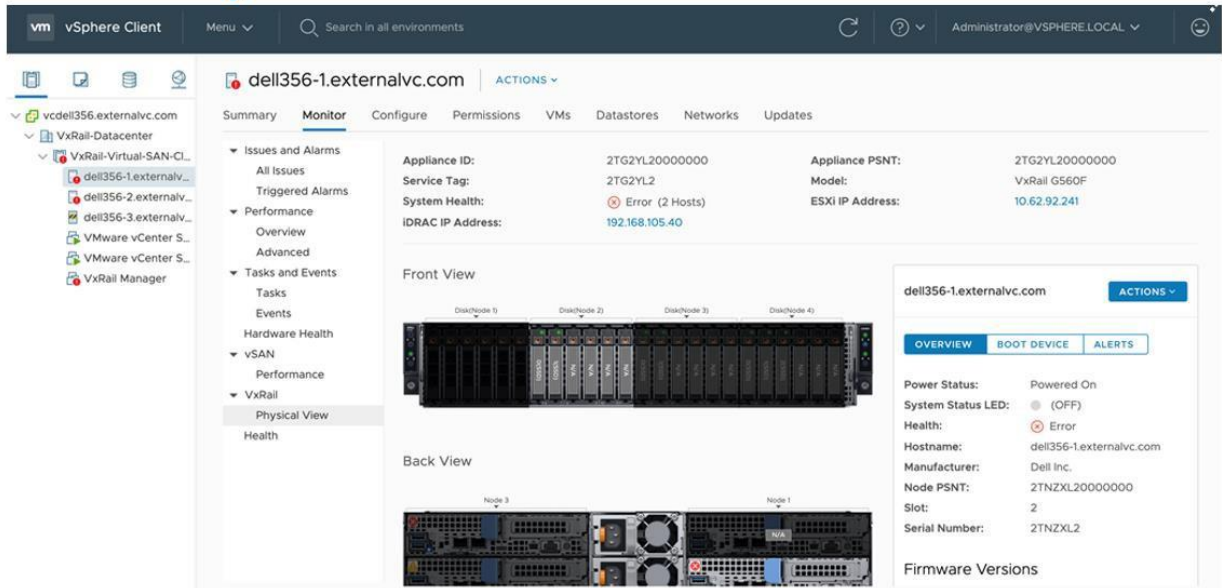
VxRail Manager

VxRail Manager 采用方便用户使用的工作流，以便自动执行 VxRail 部署和配置以及监视整个群集中的各个系统和各个节点的运行状况。它还具备硬件可维护性和系统平台生命周期管理的功能。例如，它可指导系统管理员完成将新系统添加到现有群集的过程，以及在新系统上线时自动检测到它们。VxRail Manager 还用于在不中断可用性的情况下更换出现故障的磁盘驱动器、生成和下载诊断日志包以及在 VxRail 节点间无中断地应用 VMware 更新或软件修补程序。

在极少数情况下，如果需要重建 VxRail 虚拟机，则 VxRail HCI 系统软件基于文件的备份可帮助确保业务连续性。

VxRail HCI 系统软件在 VxRail 系统上预装为单个虚拟机。利用适合 vCenter Server 的 VxRail Manager 插件，所有 VxRail Manager 功能都与 vCenter Server 集成并可从中访问，以便用户可在熟悉的管理界面上使用这些有价值的功能。使用 VxRail Manager 插件时，vCenter Server 可以管理 VxRail 群集的物理硬件。

图 17 vCenter Server 的 VxRail Manager 插件



VxRail 还利用 VMware vRealize Log Insight 监视系统事件，以及提供有关虚拟环境和系统硬件状态的持续全面通知。它通过日志监视、智能分组和分析来为 VxRail 系统提供实时自动日志管理，从而跨 VxRail 物理、虚拟和云环境提供更好的大规模故障排除。而且，通过提供可以无中断或无宕机地自动安装的修补程序软件和更新通知，VxRail HCI 系统软件可简化系统平台生命周期管理。也可从 VxRail Manager 插件或 REST API 内访问 Dell EMC Software Remote Services (SRS)，提供企业级支持和服务。SRS 包括在线聊天支持和 Dell EMC 现场服务协助（如下图所示）。

图 18 VxRail Manager ESRS 详细信息

- 为 VxRail 提供安全、高速并且全天候运行的主动式双向远程连接。
- 信号检测功能，可确保持续监视、通知和故障处理，从而将可用性提高 15%。
- 在潜在问题对业务造成影响之前，快速对其进行远程诊断和修复，从而将服务事件解决速度提高 5 倍。
- 以无缝方式自动接收支持更新。

客户环境

Dell EMC 全球支持

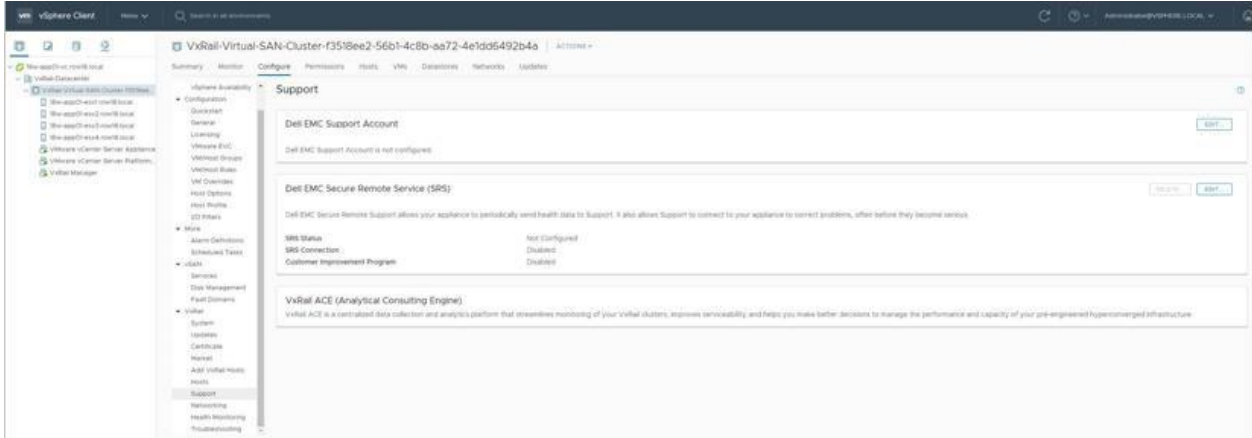




ESRS

除了 ESRS 特定支持，vCenter Server 上的 VxRail 支持页面链接到 VxRail 社区页面，以获取 Dell EMC 知识库文章，以及链接到用户论坛，以获取常见问题解答信息和 VxRail 最佳做法。下图是支持视图的一个示例。

图 19 VxRail Manager 的 “Support” 选项卡



通过 VxRail Manager 插件可访问数字市场以查找和下载经鉴定的软件包，如 VMware Horizon Cloud、Data Domain Virtual Edition、RecoverPoint for VM、vSphere Data Protection 及 VxRail 系统的其他软件选项。

可由客户升级的软件

VxRail HCI 系统软件可由客户通过完全自动化的经验证的过程进行升级。从 VxRail Manager 插件或 REST API 发起一键式软件升级，并且它会自动下载准备好进行更新的所有软件（包括 VxRail HCI 系统软件、vCenter Server 和 PSC、ESXi 主机以及 SRS）。自动化过程由四个步骤组成，包括 VxRail 软件下载、就绪性检查、软件的实际更新，最后是验证和升级后检查。最后的验证步骤可确保升级成功，并且 VxRail 系统可采用升级后的新软件版本完全正常运行。

下图显示客户执行的 VxRail HCI 系统软件升级的四个自动化步骤。在此工作流中，客户有两个接触点，第一个是在选择获取更新程序包的位置时，第二个是执行群集升级（硬件固件和软件更新一起执行）

图 20 可由客户执行的 VxRail HCI 系统软件升级的自动化过程步骤



一次对一个节点执行步骤 3, 其中 ESXi 主机置于维护模式, 并且 VM 使用 vMotion 移动到其他节点, 从而使升级过程无中断。在全新 VxRail 软件版本中, 在升级当前节点的过程中, 升级过程会在升级序列中的下一个节点上预置更新捆绑包。此改进缩短了升级节点的时间, 最终缩短了完成群集升级的总时间。

VxRail 分析咨询引擎 (ACE)

VxRail ACE 是一个集中式数据收集和分析平台, 可以简化对客户的多 VxRail 群集的监视, 提高可维护性, 并帮助客户作出更好的决策, 更明智地管理其 HCI 的性能和容量。它是一个基于云的分析平台, 利用从 VxRail 群集收集的高级遥测功能实现其基础架构机器学习, 以便提供报告和切实可行的见解。VxRail ACE 的基础架构机器学习利用内置的 Dell EMC 最佳做法和超过 700 个常见问题的知识。此产品为整个 HCI 堆栈提供运行状况评分, 使客户能够快速确定要进行故障排除和处理的方面, 以便根据预计的 IT 资源增长进行有效地扩展。

VxRail ACE 可用, 不需要为 VxRail 群集提供额外的硬件或软件。它依赖于在 VxRail 节点上运行的 VxRail HCI 系统软件提供的数据收集器服务, 聚合来自 vSAN 群集以及 VxRail 系统的指标。正式名称为 Adaptive Data Collector, 该服务将此数据捆绑包传输到 VxRail ACE 平台, 为 Dial Home 服务使用相同的 Secure Remote Services (SRS) 管道。由于它使用 SRS, 因此需要在线支持帐户, 还需要 SRS 的配置和启用, 以便将数据传输到 VxRail ACE 数据湖。此存储库驻留在 Dell EMC。通过将 Pivotal Cloud Foundry 作为其基于云的服务平台, VxRail ACE 整合了基础架构机器学习, 以便生成报告和洞察, 使客户能够改进可维护性和提高运营效率。VxRail ACE 功能完全通过基于 SaaS 的 Web 门户来使用, 该门户为客户的 VxRail 环境提供一个全局视图。

vSphere 和 vSAN 订购信息

VxRail 系统允许客户将任何符合条件的现有 vSphere 许可证用于其 VxRail，也可以随 VxRail 一起购买许可证。此 VxRail vSphere 许可证独立模式（也称为“自带”或 BYO vSphere 许可证模式）使客户可以利用他们可能已购买的各种 vSphere 许可证。

VxRail 支持多个 vSphere 许可证版本，包括 Enterprise+、Standard 和 ROBO 版本（vSphere Enterprise 也受支持，但不再可从 VMware 获取）。当系统专用于 VDI 时，还支持来自 Horizon 捆绑包或附加项的 vSphere 许可证。

如果需要购买 vSphere 许可证，则应通过 Dell EMC（客户的首选 VMware 渠道合作伙伴）或直接从 VMware 订购它们。通过 VMware ELA、VMware 合作伙伴或 Dell EMC 获取的许可证会从 Dell EMC 收到一次呼叫支持。

在确定与 VxRail 系统配合使用的最佳 vSphere 许可证时，一个关键考虑因素是 VxRail 功能的效果。DRS（本技术手册前面部分介绍的一个重要 vSphere 功能）本技术手册向 VxRail 群集提供最大程度的功能差异。客户应考虑 DRS 提供的自动化程度，以确定所需的 vSphere 许可证是否包含此功能。

VxRail 支持灵活的 vSAN 许可选项，并且要求 vSAN 随 VxRail 一起订购，或者通过 VMware 的 vSAN ELA 来申请。VxRail 支持 vSAN 的所有许可证版本，包括标准版、高级版和企业版。有关 vSAN 版本之间差异的详细信息，请参阅此处的《VMware vSAN 比较指南》：

<https://www.vmware.com/content/dam/digitalmarketing/vmware/en/pdf/products/vsan/vmware-vs-an-67-licensing-guide.pdf>

下图列出了 VxRail 随附的软件。

图 21 VxRail 附带的软件



注意：请咨询您的 Dell EMC 代表，验证支持的最新软件版本级别。

注意：随 VxRail 附带的 vCenter 许可证适用于部署在 VxRail 上的 VxRail，不能转移到其他外部服务器。与 VxRail 集成的所有其他软件也是不可转移的。

使用下图中的信息作为指导。

图 22 一般 vSphere 选项



|  vSphere Enterprise Plus |  vSphere Standard |
|---|--|
| <p>大大提升管理员的工作效率</p> <ul style="list-style-type: none"> • 自动化工作负载重新平衡和相关性规则 • 自动维护模式 • 简化的驱动器更换 • 一键式软件更新 • VxRail V 系列的 vGPU 支持 | <p>更低的前期成本；手动管理</p> <ul style="list-style-type: none"> • 手动工作负载平衡 • 手动维护模式 • 多步骤驱动器更换 • 一键式软件更新 • 不支持 vGPU |

图 23 vSphere 版本

| vSphere 版本和产品包 | vSphere 版本功能级别 | vSphere 许可证密钥输入设备 |
|---|----------------------|-------------------|
| vSphere Enterprise Plus | ENT+ | CPU |
| vSphere Enterprise | ENT+ | CPU |
| vSphere Standard | STD | CPU |
| vSphere ROBO Advanced | STD* (其他功能, 无 DRS) | VM (数量: 25) |
| vSphere ROBO Standard | STD | VM (数量: 25) |
| vSphere Desktop | ENT+ , 仅限虚拟桌面 | 用户 (数量: 100) |
| vSphere Platinum | ENT+ | CPU + 订阅 |
| vSphere with Operations Management Enterprise Plus (VSOM) | ENT+ | CPU |
| Horizon Enterprise | ENT+ , 仅限虚拟桌面 | 用户 (数量: 10 或 100) |
| Horizon Advanced | ENT+ , 仅限虚拟桌面 | 用户 (数量: 10 或 100) |
| Horizon Standard | ENT+ , 仅限虚拟桌面 | 用户 (数量: 10 或 100) |
| vCloud Suite | ENT+ | CPU |

VMware vSphere

VMware vSphere 软件套件可提供业界卓越的虚拟化平台，从而采用高度可用、有弹性、高效的按需基础架构提供应用程序虚拟化，这使它成为 VxRail 系统的理想软件基础。ESXi 和 vCenter 是 vSphere 软件套件的组件。ESXi 是直接安装在物理 VxRail 服务器节点上的虚拟机管理程序，使之分区为多个逻辑服务器或虚拟机。虚拟机在 ESXi Server 上进行配置。VMware vCenter Server 是用于管理 ESXi 主机和虚拟机的集中化管理应用程序。

以下各部分提供了对 VxRail 软件体系结构中实施的 VMware 软件组件的深入探讨。

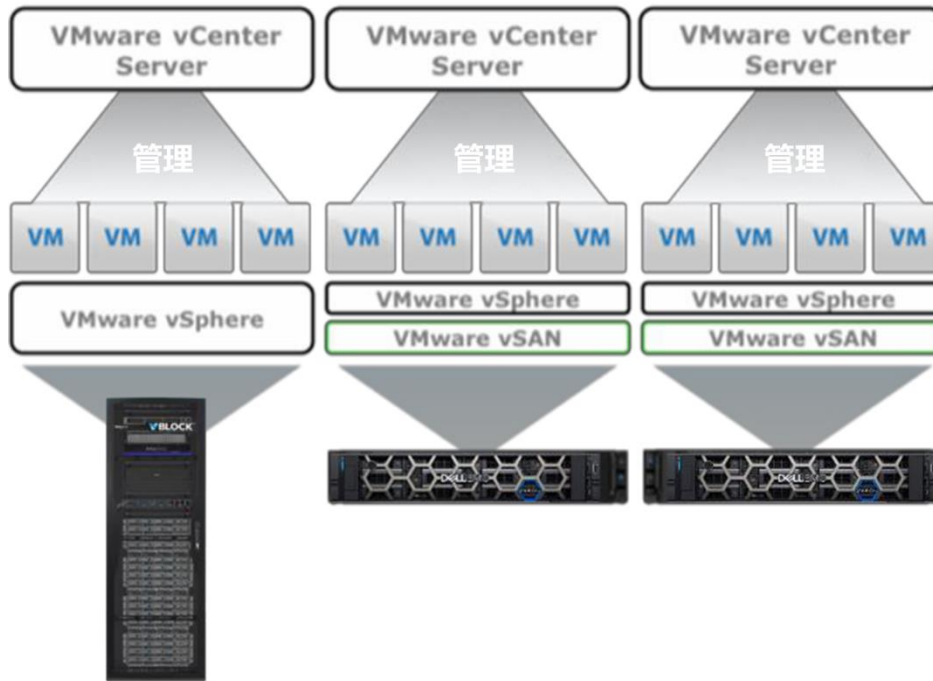
VMware vCenter Server

vCenter Server 是用于管理 VMware 环境的集中化平台。VxRail 包含用于运行 VxRail 上托管的 vCenter 的许可证（许可证不可转让来在外部服务器上运行 vCenter）。它是服务器虚拟化和 vSAN 的主要管理点，并且是 vMotion、Distributed Resource Scheduler (DRS) 和高可用性 (HA) 等高级功能的实现技术。vCenter 可扩展为企业级别，其中单个 vCenter 可以支持多达 1000 个主机 (VxRail 节点) 和 10000 个虚拟机。vCenter 支持数据中心、群集和主机的逻辑层次结构，这使资源可以按使用情形或业务线进行隔离，并使资源可以根据需要动态移动。所有这些都通过单个界面完成。

作为 VxRail 部署的一部分，可以选择在 VxRail 上配置独立 vCenter 实例，为虚拟机和 vSAN 环境提供主要管理点。vCenter 实例会使用外部 Platform Services Controller (PSC) 进行配置，其中 vCenter Server 和 PSC 会配置为单独的虚拟机。

下图是三个单独 vCenter 环境的示例。

图 24 三个单独的 vCenter 环境

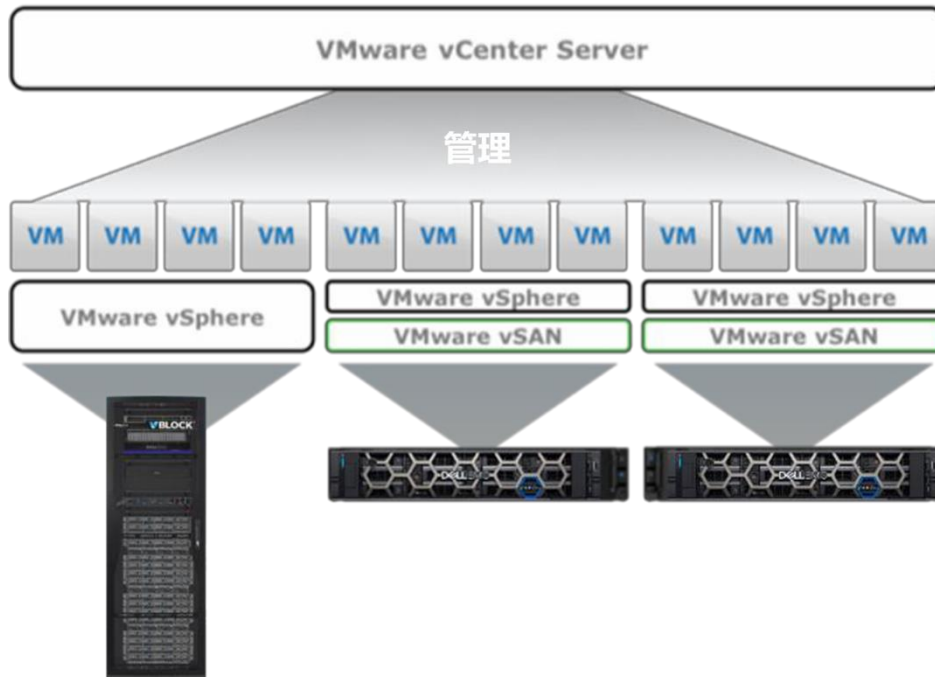


VxBlock 和两个 VxRail 环境作为单独实体进行管理。这是一个简单环境，在某些情形下计划和部署可能更为容易，因为不存在交互。

在初始配置期间，可以选择将 VxRail系统加入现有的外部托管的 vCenter 服务器环境（如果版本与 VxRail 软件兼容）。这样便可以使中央 vCenter Server 通过单一控制台管理多个 VxRail 系统。每个 VxRail 群集都会出现在 vCenter 中。

下图显示多个 VxRail 群集属于现有 vCenter 环境的示例。每个 VxRail 都是 vCenter 中唯一且独立的 vSAN 数据存储区。

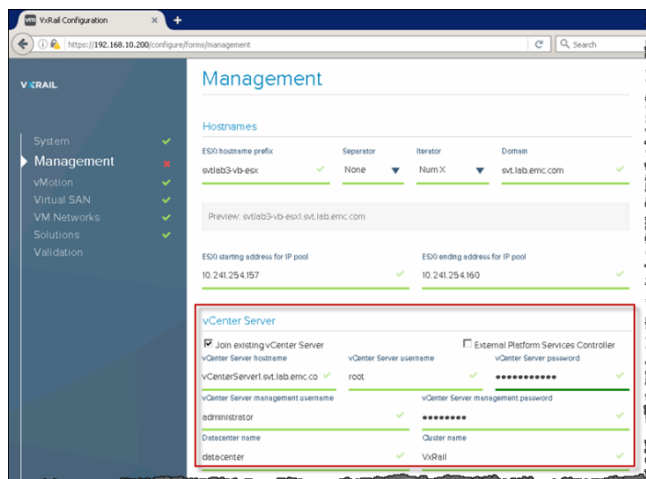
图 25 现有 vCenter 环境中的 VxRail



外部 vCenter Server 可以是物理服务器或虚拟服务器（作为 VCSA 运行或在 Windows 虚拟机中运行）。Platform Services Controller (PSC) 可以是嵌入式或非嵌入式。作为初始配置的一部分，VxRail 向导的管理页面提供用于加入现有 vCenter 的选项。如果选择该选项，则指定 vCenter Server 的主机名和管理员密码，标识数据中心以添加 VxRail 环境，以及提供群集的名称。

下图显示用于指定加入现有 vCenter Server 的 VxRail 初始化向导对话框示例。

图 26 初始化向导



数据中心必须已在 vCenter 中存在，并且群集会在安装过程中进行创建。有关配置选项的详细信息，请参阅 [VxRail vCenter Server Planning Guide](#) (VxRail vCenter Server 规划指南)。

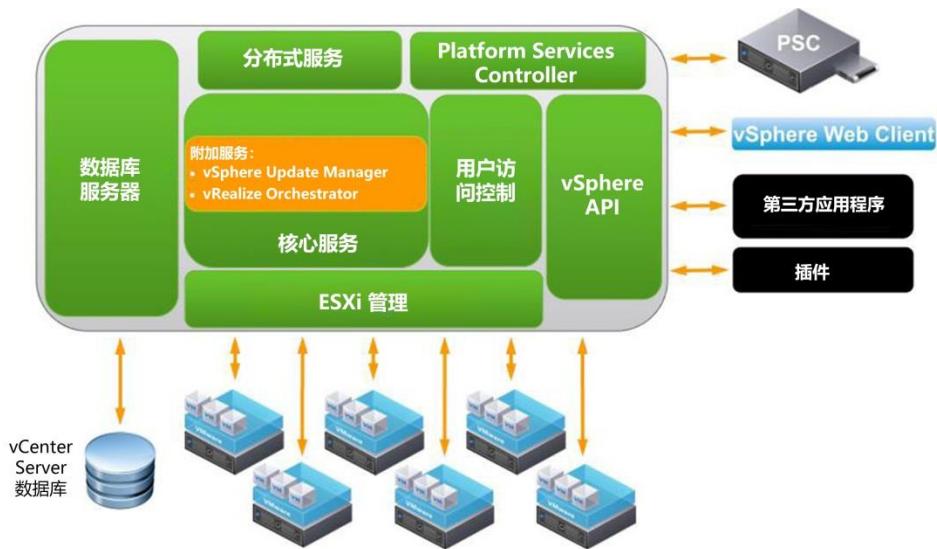
vCenter 服务和接口

vCenter 提供一些服务和接口，包括：

- 核心 VM 及其资源服务，如清单服务、任务计划、统计信息日志记录、警报和事件管理以及 VM 调配和配置
- 分布式服务，如 vSphere vMotion、vSphere DRS 和 vSphere HA
- vCenter Server 数据库接口

下图阐明了 vCenter 服务在 vSphere 环境中的组织结构。

图 27 vCenter 服务



PSC 部署选项

Platform Services Controller (PSC) 可以部署为嵌入式（在外部托管的 vCenter 中）或外部，如下所示。当 VxRail 托管 vCenter 时，PSC 在外部部署为独立的虚拟机。请参见下图。

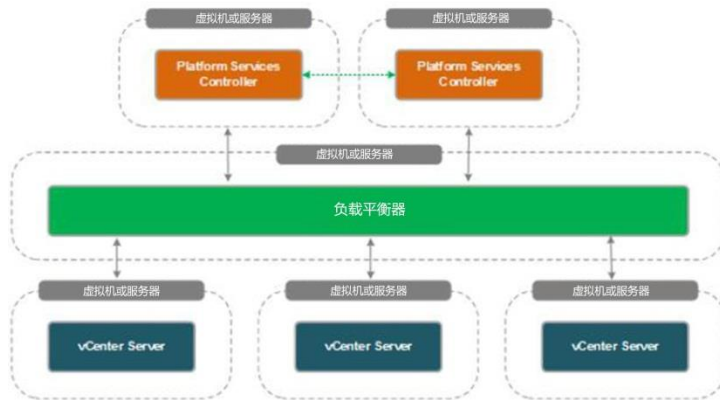
图 28 嵌入式和外部 PCS 部署



嵌入式 PSC 非常适合于小型环境, 或非常适合于简单性和降低资源利用率是针对环境的关键因素的情况。vCenter Server 与嵌入式 PSC 捆绑在一起, 所有 PSC 服务都驻留在与 vCenter Server 相同的主机上。

外部 PSC (参见下图) 非常适合于存在多个 vCenter Server 但您希望对站点使用单一控制台的较大环境。与 PSC 和 vCenter Server 捆绑在一起的服务部署在不同的虚拟机上, 甚至是部署在不同的物理服务器上。利用外部托管 vCenter 时, 支持外部 PSC、增强型链接模式 (ELM)。

图 29 针对高可用性配置的外部 PSC



增强型链接模式

多个 vCenter Server 可以连接到相同的外部 Platform Service Controller。增强型链接模式可对配置为使用 Platform Services Controller 域的多个 vCenter Server 实现整合管理视图。这包括一个通用清单, 管理员可以在其中跨 vCenter Server 搜索对象。系统间的角色、权限、许可证和其他关键数据也会在 vCenter 实例间进行复制。

此单一控制台视图可提供企业级扩展, 特别适用于大型多 VxRail 群集环境或 VxRail 系统加入现有大型 vSphere 环境的情况。此外, 增强型链接模式还可实现跨 vCenter vMotion (可以在 vCenter 之间移动 VM) 等功能。

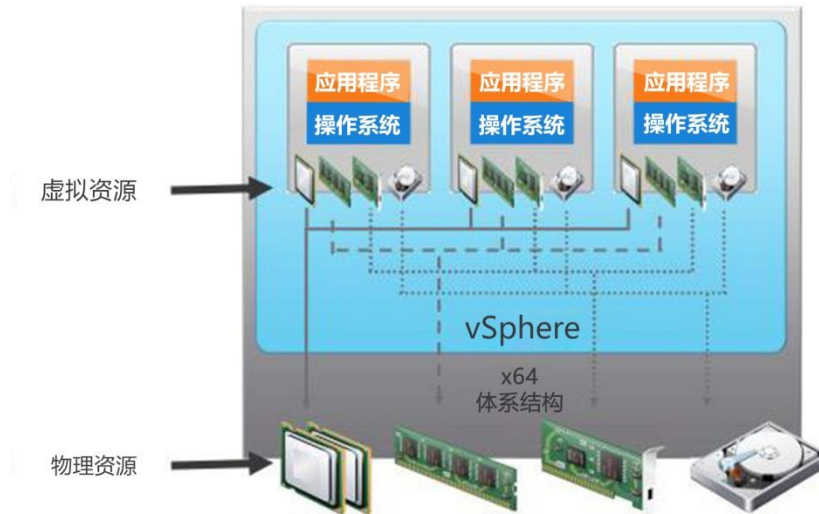
运行采用 VxRail 的外部托管 vCenter 以及利用外部 PSC 时, VxRail 支持增强型链接模式。

VMware vSphere ESXi

vSphere 是 VxRail 系统中的核心操作软件。vSphere 聚合了一组全面的功能, 这些功能可高效地对 ESXi 主机下可用的资源进行池化和管理工作。请记住, 本技术书籍重点介绍了 vSphere 技术, 特别是介绍了它与 VxRail 系统配合使用的情况。其他 vSphere 实施中包含的功能可能并不适用于 VxRail, 而 VxRail 中包含的功能可能并不适用于其他实施。

VMware ESXi 是对虚拟机进行部署并提供服务的 enterprise 级虚拟机管理程序。下图展示了其基本体系结构。

图 30 鸟瞰视图: vSphere ESXi 体系结构



ESXi 将一个物理服务器分区成可以在相同物理服务器并行运行的多个安全且可移植的 VM。每个 VM 都表示一个完整系统 (具有处理器、内存、网络、存储和 BIOS)，因此任何操作系统 (来宾操作系统) 和软件应用程序都可以在虚拟机中安装并运行而无需任何修改。

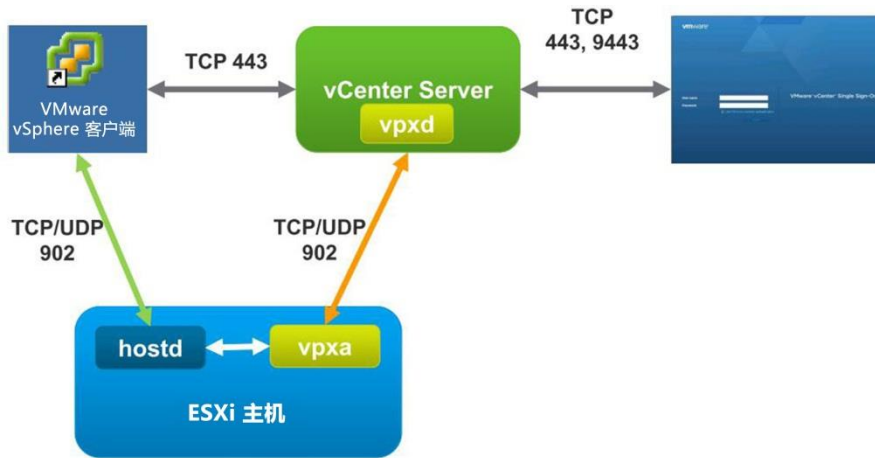
虚拟机管理程序根据需要向虚拟机 (VM) 动态提供物理硬件资源，以支持 VM 操作。虚拟机管理程序使虚拟机可以在一定程度上独立于底层物理硬件进行操作。例如，虚拟机可以从一个物理主机移动到另一个。此外，VM 的虚拟磁盘可以从一种类型的存储移动到另一种类型，而不会影响虚拟机的正常功能。

ESXi 还使 VM 相互隔离，因此当一个 VM 中运行的来宾操作系统出现故障时，相同物理主机上的其他 VM 不受影响，可继续运行。虚拟机共享对 CPU 的访问，由虚拟机管理程序负责 CPU 调度。此外，ESXi 会向 VM 分配可用内存区域，并提供对与物理主机关联的物理网卡和磁盘控制器的共享访问。不同虚拟机可以在相同物理计算机上运行不同操作系统和应用程序。

vCenter Server 与 ESXi 主机之间的通信

vCenter Server 通过 vCenter Server 代理 (也称为 vpxa 或 vmware-vpxa 服务，它会在添加到 vCenter Server 清单时在 ESXi 主机上启动) 与 ESXi 主机进行通信。请参见下图。

图 31 vCenter 与 ESXi 主机之间的通信



具体而言，vCenter vpxd 守护程序通过 vpxa 服务与 ESXi 主机守护程序（称为 hostd 进程）进行通信。vpxa 进程充当 vCenter Server 上运行的 vpxd 进程与 ESXi 主机上运行的 hostd 进程之间的中间方，从而中继要在主机上执行的任务。hostd 进程直接在 ESXi 主机上运行，负责管理 ESXi 主机上的大多数操作，包括创建 VM、迁移 VM 和打开 VM。

虚拟机

虚拟机由以下的一组核心相关文件（或一组对象）组成，如下图所示。

图 32 虚拟机文件

| Name | Type | Size |
|------------------------|--------------------------|-----------------|
| .dvsData | Folder | |
| .sdd.sf | Folder | |
| TestVM01-572b3eef.v... | File | 1,060,449.20 KB |
| TestVM01.nvram | Non-volatile Memory File | 8.48 KB |
| TestVM01.vmdk | Virtual Disk | 6,711,296.00 KB |
| TestVM01.vmsd | File | 0.00 KB |
| TestVM01.vmx | Virtual Machine | 3.44 KB |
| vmware-1.log | VM Log File | 374.14 KB |
| vmware-2.log | VM Log File | 221.98 KB |
| vmware-3.log | VM Log File | 219.72 KB |
| vmware.log | VM Log File | 218.44 KB |

除了日志文件，每个文件的名称都以虚拟机名称 (VM_name) 开头。这些文件包括：

一个配置文件 (.vmx) 和/或一个虚拟机模板配置文件 (.vmtx)

一个或多个虚拟磁盘文件 (.vmdk)

一个包含虚拟机 BIOS 设置的文件 (.nvram)

一个虚拟机当前日志文件 (.log), 以及一组用于归档旧日志条目的文件 (-#.log)

用于在争用期间回收内存的交换文件 (.vswp)

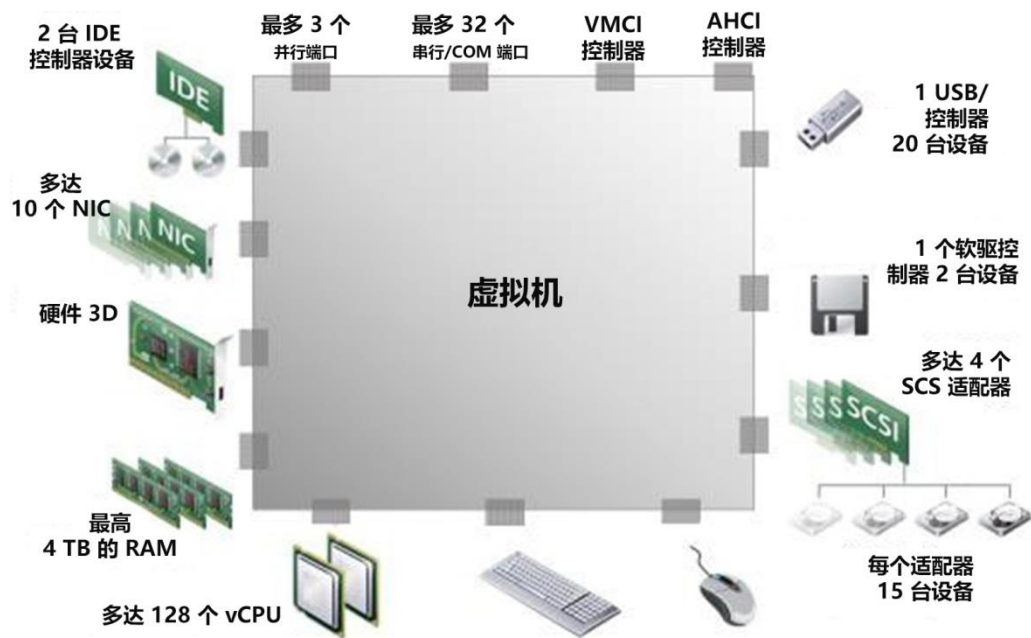
一个快照描述文件 (.vmsd), 该文件在虚拟机没有快照时为空

虚拟机硬件

虚拟机使用虚拟硬件。每个来宾操作系统都可看到普通硬件设备, 但是不了解这些设备是虚拟设备。

下图中显示了硬件资源。

图 33 VM 的硬件资源



所有虚拟机都具有统一硬件 (除了系统管理员可以应用的几个变体)。统一硬件使虚拟机可跨 VMware 虚拟化平台进行移植。vSphere 支持许多最新 CPU 功能, 包括虚拟 CPU 性能计数器。可以添加虚拟硬盘和 NIC, 以及配置虚拟硬件 (如 CD/DVD 驱动器、软盘驱动器、SCSI 设备、USB 设备和最多 16 个 PCI vSphere DirectPath I/O 设备)。

虚拟机通信

虚拟机通信接口 (VMCI) 在虚拟机与虚拟机管理程序之间提供了高速通信通道。不能添加或删除 VMCI 设备。SATA 控制器提供对虚拟磁盘和 DVD/CD ROM 设备的访问。SATA 虚拟控制器向虚拟机显示为 AHCI SATA 控制器。如果没有 VMCI, 虚拟机会使用网络层与主机进行通信, 这会增加通信开销。借助 VMCI, 通信开销极低, 并且需要这种通信的任务可以进行优化。内部网络使用 VMXNET3 时可以达到的平均传输速率略高于 2Gbps。借助 12 个 8k 大小队列对, VMCI 可以最高达到几乎 10Gbps 速度。

VMCI 提供套接字 API，它们非常类似于已用于 TCP/UDP 应用程序的 API。

有关虚拟硬件的更多信息，请参阅《*vSphere Virtual Machine Administration Guide*》（*vSphere 虚拟机管理指南*）：<https://www.vmware.com/support/pubs/vsphere-esxi-vcenter-server-6-pubs.html>。

虚拟网络

VMware vSphere 提供了一套丰富的网络功能，这些功能可很好地与复杂企业网络集成。这些网络功能由 ESXi Server 提供，由 vCenter 管理。通过虚拟网络可以采用物理机联网方式相同的方式对虚拟机进行联网。可以在单个 ESXi Server 主机中或是跨多个 ESXi Server 主机构建虚拟网络。VxRail ESXi 主机使用虚拟交换机在 VxRail 群集中的虚拟机间进行通信（使用的协议与物理交换机上使用的协议相同），而无需其他网络硬件。虚拟交换机还支持与来自交换机供应商的标准 VLAN 实施兼容的 VLAN。与物理以太网交换机一样，虚拟交换机会在数据链路层上转发帧。

虚拟以太网适配器是虚拟网络的关键 vSphere 组件。虚拟机可以使用一个或多个虚拟以太网适配器进行配置，其中每个适配器都具有自己的 IP 地址和 MAC 地址。因此，从网络角度来看，虚拟机具有与物理机相同的属性。此外，虚拟网络可实现当前使用物理网络无法实现的功能。虚拟以太网适配器由各个虚拟机以及将 VM 相互连接并将虚拟机和 ESX Server Service Console 连接到外部网络的虚拟交换机使用。

虚拟交换机通过出站以太网适配器（称为 vmnic）链接到外部网络，并且虚拟交换机可以将多个 vmnic 绑定在一起（很像传统服务器上的 NIC 绑定），从而向它所服务的虚拟机扩展可用性和带宽。

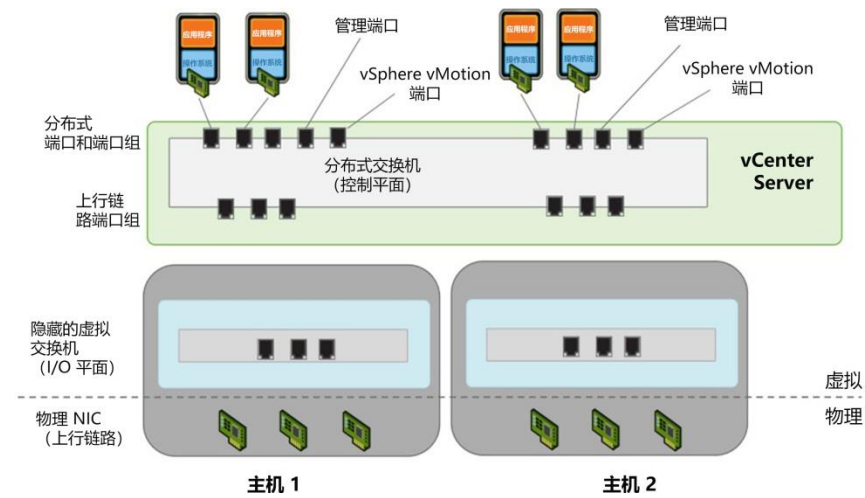
虚拟交换机类似于其物理交换机对应项。与物理网络设备一样，每个虚拟交换机都进行隔离以实现安全性，并且具有自己的转发表。一个表中的条目不能指向另一个虚拟交换机上的另一个端口。交换机仅查找与帧的来源虚拟交换机上的端口匹配的目标。此功能可阻止潜在黑客破坏虚拟交换机隔离。虚拟交换机还支持端口级别的 VLAN 分段，因此每个端口都可以配置为针对单个 VLAN 的访问端口或是针对多个 VLAN 的中继端口。

虚拟分布式交换机

VxRail 群集使用 VMware 虚拟分布式交换机 (VDS)，后者充当在相同群集中跨越多个节点的单个交换机。此交换机使虚拟机可以在多个主机之间迁移时保持一致的网络配置。分布式交换机在数据中心级别上采用 vCenter Server 进行配置，使配置在所有主机间保持一致。vCenter Server 将分布式端口的状态存储在 vCenter Server 数据库中。当虚拟机在主机之间移动时，网络统计信息和策略会随虚拟机一起迁移。如接下来各部分中所讨论的那样，vSAN 依赖于 VDS 来实现其存储虚拟化功能，而 VxRail 系统将 VDS 用于系统流量。

下图概括介绍了 VDS。

图 34 虚拟分布式交换机



vMotion 和虚拟机移动性

通过 VMware vMotion 可以将正在运行的虚拟机从一个物理服务器实时迁移到另一个物理服务器，同时保证零宕机时间、连续的服务可用性和完全的事务完整性。vMotion 是创建动态、自动化和自我优化数据中心的一项关键支持技术。vMotion 可连续自动地在资源池中分配虚拟机。它还通过在不中断业务运营的情况下执行维护，来提高可用性。

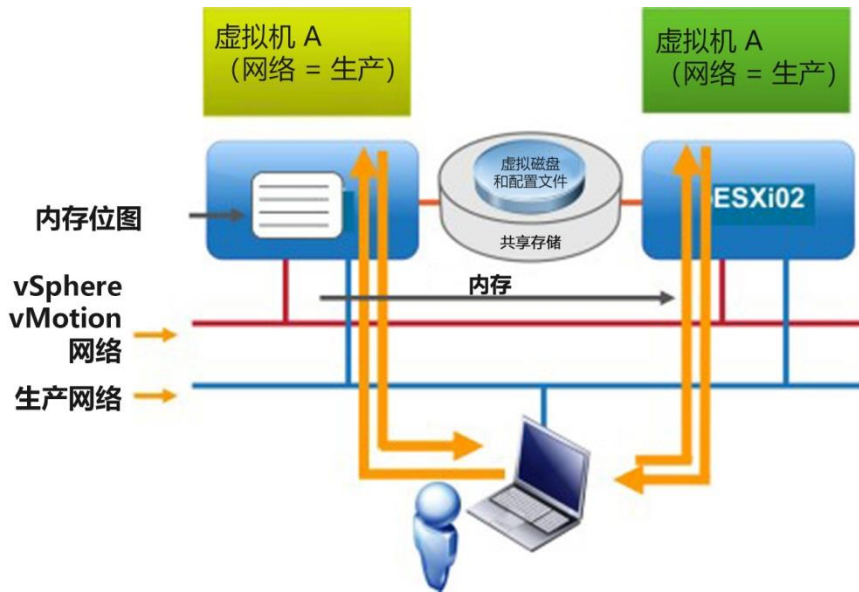
用于无中断地迁移工作负载的高级功能是区分 VxRail 解决方案与其他 HCI 选项的功能之一。在 vSphere 虚拟基础架构中，迁移指将虚拟机从一个主机、数据存储区或 vCenter Server 系统移动到另一个主机、数据存储区或 vCenter Server 系统。存在不同类型的迁移，其中包括：

冷迁移，这是将关闭的 VM 迁移到新主机或数据存储区

挂起的迁移，这是将挂起的 VM 迁移到新主机或数据存储区

实时迁移, 这是使用 vSphere vMotion 将“实时”的已启动 VM 迁移到新主机和/或使用 vSphere Storage vMotion 将实时的已启动 VM 的文件迁移到新数据存储区 vMotion 支持在 ESXi 主机之间无中断或无宕机地实时迁移虚拟机。下图总结了该过程。

图 35 vMotion 迁移



使用 vMotion 时, 虽然会迁移虚拟机的整个状态, 不过数据仍保留在相同数据存储区中。状态信息包括当前的内存内容, 以及用于定义和标识虚拟机的所有信息。内存内容包括内存中的事务数据以及任意位数的操作系统和应用程序。状态中存储的定义和标识信息包括映射到虚拟机硬件元素的所有数据 (包括 BIOS、设备、CPU、以太网卡的 MAC 地址)。

vMotion 迁移包含以下步骤:

1. VM 内存状态通过 vMotion 网络从源主机拷贝到目标主机。用户继续访问 VM, 并且可能更新内存中的页面。内存中已修改的页面的列表会保留在源主机上的内存位图中。
2. 大部分 VM 内存从源主机拷贝到目标主机之后, VM 会停顿。VM 上不会发生其他任何活动。停顿期间, vMotion 会将 VM 设备状态和内存位图转移到目标主机。
3. VM 在源主机上停顿之后, VM 会立即在目标主机上进行初始化并开始运行。免费地址解析协议 (GARP) 请求会向子网通知, VM 的 MAC 地址现在处于新交换机端口上。
4. 用户访问目标主机 (而不是源主机) 上的 VM。源主机上的 VM 所使用的内存页面会标记为可用。

增强的 vMotion 兼容性

增强的 vMotion 兼容性 (EVC) 是一种防止 vMotion 迁移由于 CPU 不兼容而失败的群集功能。EVC 可确保群集中的所有主机向虚拟机呈现相同的 CPU 功能集，即使主机中的实际 CPU 不同时也是如此。它可防止 CPU 不兼容所造成的迁移失败。此功能默认情况下在 VxRail 系统中处于开启状态。

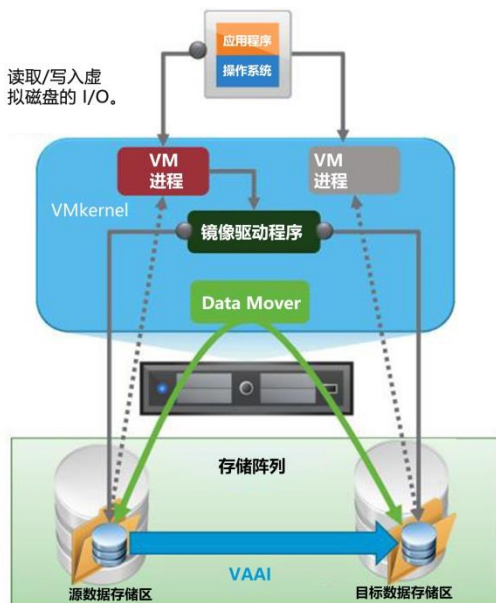
Storage vMotion

Storage vMotion 使用 I/O 镜像体系结构在源与目标之间拷贝磁盘数据块：

1. 启动存储迁移。
2. 使用 VMkernel Data Mover 并提供 vSphere Storage APIs for Array Integration (VAAI) 以拷贝数据。
3. 启动新 VM 进程。
4. 将 I/O 调用映射到已拷贝到目标数据存储区上的虚拟磁盘的文件数据块。切换到目标 VM 进程以开始访问虚拟磁盘拷贝。

下图说明了该过程

图 36 Storage vMotion



存储迁移过程只拷贝磁盘一次，镜像驱动程序会同步源和目标数据块而无需递归传递。换句话说，如果源数据块在迁移之后发生更改，则镜像驱动程序会同时写入两个磁盘，这可保持事务完整性。与更传统的存储迁移选项相比，Storage vMotion 的镜像体系结构会产生可预测性更强的结果、更短的迁移时间和更少的 I/O 操作。其速度快到使终端用户察觉不到。即使在使用较慢的目标磁盘时也能保证迁移成功。

vSphere 支持以下 Storage vMotion 迁移：

- 群集之间
- 数据存储区之间（包括非 vSAN 到 vSAN，以及相反方向）
- 网络之间
- vCenter Server 实例之间，对于配置为增强型链接模式的 vCenter Server 与使用时间同步的主机
- 长距离（最多 150 毫秒往返时间）

请注意，对于 VxRail 群集，Storage vMotion 只能用于迁入或迁出 vSAN 数据存储区。

vSphere Distributed Resource Scheduler

VMware Distributed Resource Scheduler (DRS) 是一种随 vSphere Enterprise Plus 和 vSphere with Operations Management Enterprise Plus 附带的功能。DRS 在聚合到逻辑池中的 VxRail 服务器资源集合中平衡计算容量。它可在 VM 间连续平衡并优化计算资源分配。

当 VM 遇到工作负载增加时，DRS 会按照用户定义的资源分配规则和策略评估 VM 优先级。如果合理，则 DRS 会分配其他资源。它还可以配置为将一致的资源专用于特定业务部门应用程序的 VM，以满足 SLA 和业务要求。

DRS 向 VM 分配资源的方法是将 VM 迁移到具有更多可用资源的其他服务器，或是将其他 VM 迁移出服务器源，从而在相同服务器上使更多“资源”可用于 VM。在 VxRail 系统中，所有 ESXi 主机都是 vMotion 网络的一部分。通过 vMotion 将 VM 实时迁移到不同的节点服务器对于终端用户是完全透明的（请参见下面的图）。DRS 可自动布置虚拟机，并确保一致且可预测的应用程序工作负载性能，从而向 VxRail 群集增加了巨大价值。

图 37 VM 跨节点服务器进行的 DRS 移动

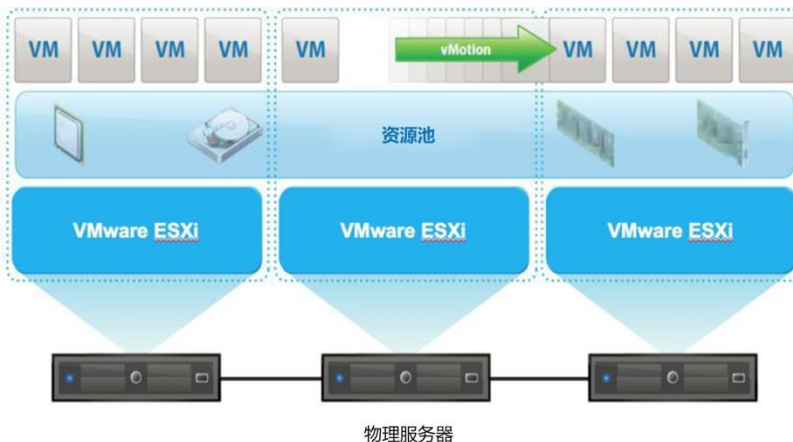
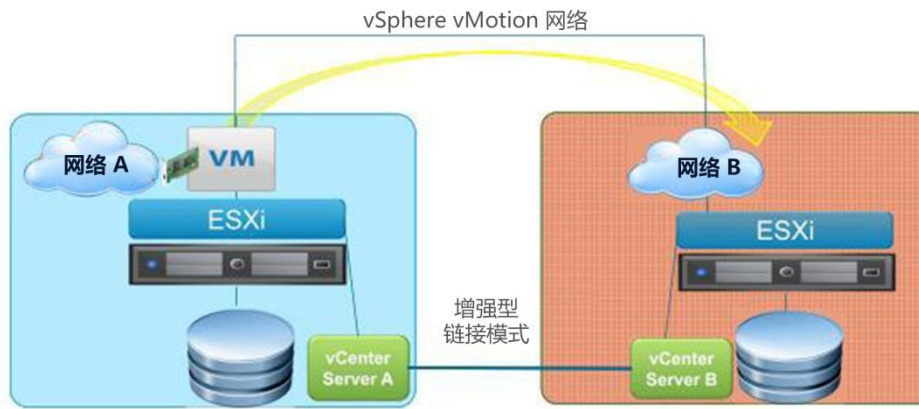


图 38 跨 vMotion 网络进行的 VM 迁移 需要外部 vCenter。

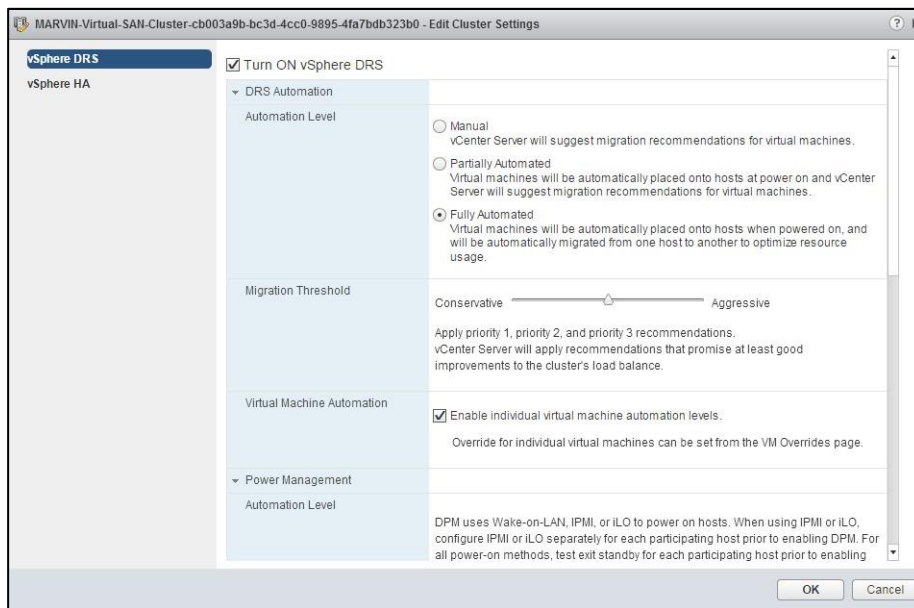


DRS 在维护期间向 VxRail 用户提供了显著优势，因为它可自动执行升级或修复操作期间手动移动实时计算机时通常涉及到的任务。DRS 通过将所有 VM 动态迁移到其他物理服务器，来促进维护自动化，从而提供透明的连续操作。这样，服务器可以参与进行维护，或是新节点服务器可以添加到资源池，在执行所有这些操作的同时，DRS 会随着物理资源的更改，在可用服务器间自动重新分布 VM。

换句话说，在添加新服务器时或是在现有服务器完成其维护周期时，DRS 会在有其他资源可用之后立即动态平衡 VM。DRS 会为 VM 分配独有的 CPU 和内存资源，并将 vSAN 用于共享存储。

下图显示了配置 DRS 的设置。

图 39 配置 DRS 设置



与其他条件和业务运营相比，某些条件和业务运营可保证更积极的 DRS 迁移策略。可调整群集参数建立触发 DRS 迁移的阈值（如上面的屏幕截图所示）。例如，级别 2 阈值仅应用指定迁移建议来显著影响群集的负载平衡，而级别 5 阈值会应用所有建议，甚至只是为了稍微改进群集的负载平衡。

DRS 仅适用于 VxRail 虚拟机。（vSAN 使用单个数据存储区并在内部处理布置和平衡。vSAN 当前不支持 Storage DRS 或存储 I/O 控制。）

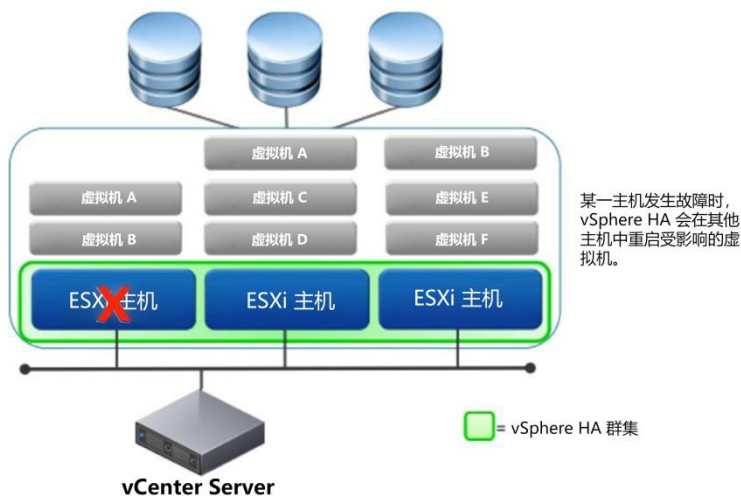
vSphere High Availability (HA)

vSphere 提供了多种解决方案，确保在计划内和计划外宕机情况下实现高可用性级别。vSphere 依靠以下技术来确保在环境中运行的虚拟机保持可用：

- 虚拟机迁移
- 多个 I/O 适配器路径
- 虚拟机负载平衡
- 故障承受能力
- 灾难恢复

vSphere HA 与 vSAN 相结合，产生了适用于 VxRail 虚拟机工作负载的有弹性、高度可用的解决方案。vSphere HA 通过在出现主机故障时重新启动虚拟机来保护它们。（参见下图。）它利用 ESXi 群集配置来确保从宕机中快速恢复，从而为虚拟机中运行的应用程序提供经济高效的高可用性。主机加入群集时，其资源会成为群集资源的一部分。群集可管理其中所有主机的资源。在 vSphere 环境中，ESXi 群集负责 vSphere HA、DRS 以及提供 VxRail 软件定义的存储功能的 vSAN 技术。请参见下图。

图 40 vSphere HA



vSphere HA 为应用程序提供多个保护点：

它通过在群集中的其他主机上重新启动虚拟机来规避任何服务器故障。

它连续监视虚拟机并重置任何检测到的 VM 故障。

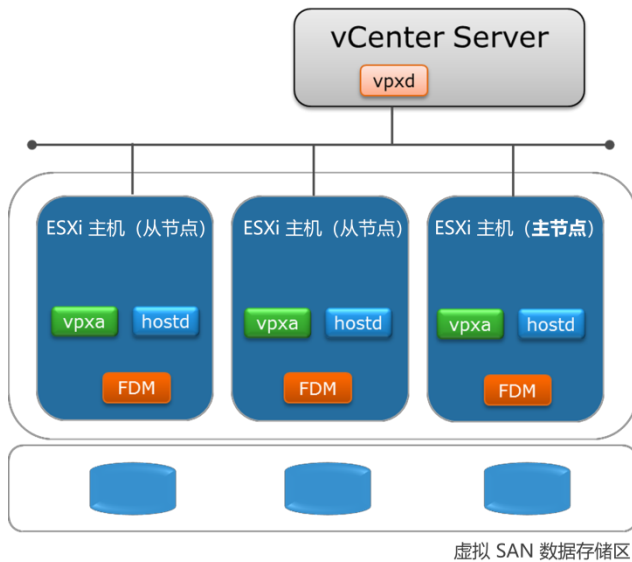
它可防范数据存储区可访问性故障，并为受影响的虚拟机提供自动恢复。借助虚拟机组件保护 (VMCP)，受影响的虚拟机会在仍然有权访问数据存储区的其他主机上重新启动。

它通过在虚拟机的主机在管理或 VMware vSAN 网络上成为隔离状态时重新启动虚拟机，来针对网络隔离保护虚拟机。即使在网络进行了分区时，也可提供此保护。

VSphere HA 进行配置之后，所有工作负载都会受到保护。无需任何操作来保护新虚拟机，并且无需特殊软件存在于应用程序或虚拟机中。

vSphere HA 中的故障切换功能包含一个名为 Fault Domain Manager (FDM) 的服务，它在成员主机上运行（如下图所示）。FDM 代理启动之后，群集主机会成为故障域的一部分，在一个故障域中，一次只能存在一个主机。如果主机处于维护模式、待机模式或与 vCenter Server 断开连接，则这些主机无法加入故障域中。

图 41 Fault Domain Manager



vpxa: hostd 与 vpxd 之间的中介
vpxd: 与 hostd 通信的 vCenter 代理
hostd: 负责管理 ESXi 操作
FDM: Fault Domain Manager

FDM 使用主-从操作模式（参见上图）。自动指定的主节点主机会管理故障域，而其余主机是从节点主机。从节点主机上的 FDM 代理使用安全 TCP 连接与主节点主机上的 FDM 服务进行通信。在 VxRail 群集中，仅当配置了 vSAN 群集之后，vSphere HA 才会启用。一旦启动了 vSphere HA，vCenter 服务器便会联系主节点主机代理，向它发送群集成员主机的列表以及群集配置。该信息会保存到主节点主机上的本地存储，然后向外推送到群集中的从节点主机。如果在正常操作期间将其他主机添加到群集，则主代理会将更新发送到群集中的所有主机。

主节点主机向 vCenter Server 提供接口以用于查询和报告故障域和虚拟机可用性的状态。

vCenter Server 治理 vSphere HA 代理, 从而标识要保护的虚拟机并维护 VM 到主机兼容性列表。

代理通过 hostd 了解状态更改, 而 vCenter Server 通过 vpxa 了解这些更改。主节点主机监视从节点的运行状况, 并对在出现故障的从节点主机上运行的虚拟机负责。同时, 从节点主机监视其本地虚拟机的运行状况, 并将状态更改发送到主节点主机。从节点主机还监视主节点主机的运行状况。

vSphere HA 通过 vCenter Server 进行配置、管理和监视。群集配置数据由 vCenter Server vpxd 进程进行维护。如果 vpxd 向主代理报告任何群集配置更改, 则主节点会播发群集配置信息的新拷贝, 随后每个从节点会获取更新的拷贝, 并将新信息写入本地存储。每个数据存储区都包含受保护虚拟机的列表。在 vCenter Server 通知任何用户启动的打开 (受保护) 或关闭 (不受保护) 操作之后, 该列表会更新。

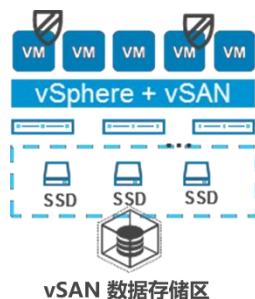
vCenter Server 监视器

提供 vCenter Server 可用性的一种方法是在 vSphere HA 群集中使用监视器功能。监视器可监视并保护 vCenter Server 服务。如果有任何服务失败, 则监视器会尝试重新启动它们。如果它由于主机故障而无法重新启动服务, 则 vSphere HA 会在新主机上重新启动运行服务的虚拟机。监视器可以使用 vCenter Server 进程 (PID 监视器) 或 vCenter Server API (API 监视器) 提供更好的可用性。

vSphere 加密

vSphere 加密使客户能够在每虚拟机级别加密数据。这个加密级别非常适合那些担心冒失的管理员将虚拟机及其所有数据发送到非安全位置的客户。应该加密哪些虚拟机由虚拟化管理团队决定, 并且可以按虚拟机进行选择 (如下图所示)。需要符合 KMIP 规范的密钥管理服务器 (如 CloudLink 或 Hytrust)。

图 42 每虚拟机级别加密 (使用 vSphere 加密)



vSAN

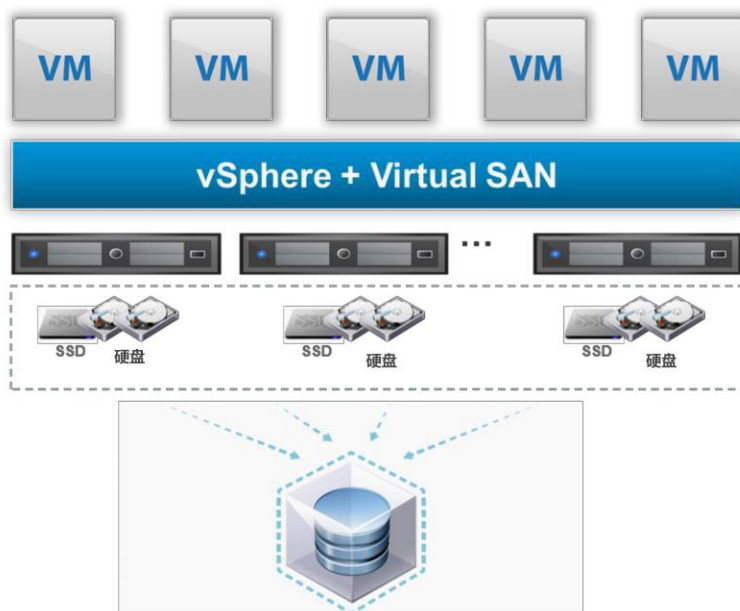
VxRail 系统利用 VMware 的 vSAN 软件，该软件与 vSphere 完全集成并提供功能齐全且经济高效的软件定义的存储。vSAN 实施了一种显著高效的体系结构，它直接内置在虚拟机管理程序中。这可将 vSAN 与通常安装作为来宾 VM 在每个主机上运行的虚拟存储设备 (VSA) 的解决方案区分开来。将 vSAN 嵌入 ESXi 内核层在性能和内存要求方面具有明显优势。它对 CPU 利用率的影响非常小（不到 10%），可基于工作负载和资源可用性进行自我平衡。它可为存储提供熟悉的数据存储区构造，并且可无缝地与其他 vSphere 功能（如 vMotion）配合工作。

vSAN 可聚合 vSphere 群集中的主机的本地连接磁盘，以创建分布式共享存储池。容量可通过向群集中的主机添加附加磁盘来轻松地纵向扩展，以及通过添加附加 ESXi 主机来横向扩展。这样便可以灵活地从非常小的环境开始，并随着时间推移而进行扩展。存储特征使用基于存储策略的管理 (SPBM) 进行配置，这样可以动态地设置和修改对象级策略，以控制存储资源调配和存储服务级别协议 (SLA) 的日常管理。vSphere 和 vSAN 为 VxRail 性能扩展提供了基础，并为托管虚拟机实现了企业级功能（包括使用 vMotion 的移动性、高可用性和 Dynamic Resource Scheduler (DRS)）。

vSAN 在 VxRail 系统首次初始化时进行预配置，并通过 vCenter 进行管理。VxRail 系统初始化过程会从群集中的每个 ESXi 节点查找本地连接的存储磁盘，以创建分布式共享存储数据存储区。vSAN 数据存储区中的存储量是群集中所有容量驱动器的总和。

下图显示了一个混合配置示例，其中每个节点都向 vSAN 数据存储区贡献了存储容量。SSD 驱动器可提供缓存来优化性能，而硬盘驱动器 (HDD) 可提供容量。

图 43 vSAN 数据存储区



vSAN 支持在 VM 创建和部署操作过程中，在 vCenter 中进行快速存储调配。vSAN 是策略驱动型，旨在简化存储调配和管理。它基于 VM 级别存储策略，自动且动态地将要求与底层存储资源匹配。VxRail 提供了两种不同的 vSAN 节点存储配置选项：利用闪存 SSD 和机械 HDD 的混合配置，以及全闪存 SSD 配置。混合配置在缓存层使用闪存 SSD，而使用机械 HDD 提供容量和持久数据存储。全闪存配置对缓存层和容量层都使用闪存 SSD。

磁盘组

VxRail 主机中的磁盘驱动器按磁盘组进行组织，而磁盘组向 vSAN 群集提供存储。在 VxRail 系统中，磁盘组包含最多一个闪存缓存驱动器和六个容量设备。根据型号，每个 VxRail 节点可以使用最多四个磁盘组进行配置。下图显示对于每种 VxRail 型号，每个节点的磁盘组数和每个磁盘组的驱动器数。

图 44 基于第 14 代服务器的 VxRail 的磁盘组配置

| 型号 | 磁盘插槽数 | 固定缓存磁盘插槽数 | 最大磁盘组 (DG) 数 | 每个 DG 的最大容量磁盘数 |
|-----------|--------------------------|----------------|--------------|----------------|
| E 系列 | 10 (2.5 英寸) | 8、9 | 2 | 4 |
| P 系列、V 系列 | 24 (2.5 英寸) | 20、21、22、23 | 4 | 5 |
| S 系列 | 12 (3.5 英寸) + 2 (2.5 英寸) | 12、13 (2.5 英寸) | 2 | 6 |

对系统进行设计和调整规模时，会确定基于容量、性能和可用性的磁盘组布局。下面是设计磁盘组布局时的注意事项：

SSD 缓存驱动器数。每个磁盘组都需要一个且只需要一个高耐用性 SSD 闪存缓存驱动器。更多磁盘组需要更多 SSD 闪存缓存驱动器。当前，VxRail 系统配有 400 GB、800 GB 和 1600 GB 缓存闪存设备。对于 VxRail，SSD 缓存驱动器必须安装在指定插槽中。请参见特定插槽位置的产品文档。

可用容量驱动器总数。每个磁盘组至少要有容量驱动器。仅当确定可用于 vSAN 群集的总容量时，才会对容量驱动器进行计数。

缓存要求。缓存可提高混合存储配置中的读取和写入性能，以及全闪存配置中的写入性能和耐用性。最佳缓存量取决于活动工作集大小。具有较大工作集大小的工作负载可能需要更多缓存，这可以通过使用较大 SSD 缓存驱动器或配置多个磁盘组（其中每个磁盘组都具有一个缓存驱动器）来进行配置。

注意：全闪存配置支持每个磁盘组最多 600 GB 的写缓存。可考虑使用多个磁盘组对每个节点实现较大写缓存。

性能。对于相同的总容量，较多数量的较小驱动器提供的 IOPS 要多于较少数量的较大驱动器。根据型号，配置更多驱动器可能需要多个磁盘组。

故障域和可恢复性。如果 SSD 缓存驱动器出现故障，则整个磁盘组进入离线状态，必须恢复。较大磁盘组需要较长时间来恢复。具有单个磁盘组的节点不会在恢复过程中为群集贡献任何 IOPS，这可能会降低群集性能。

混合和全闪存差异

在混合和全闪存配置中以不同方式使用缓存。在混合磁盘组配置（将机械 HDD 用于容量，将闪存 SSD 设备用于缓存）中，缓存算法会尝试最大程度提高读取和写入性能。闪存 SSD 设备有两种用途：读缓存和写缓冲区。70% 的可用缓存分配用于存储经常读取的磁盘数据块，从而最大程度减少对速度较慢的机械磁盘的访问。其余 30% 的可用缓存分配进行写入。多个写入会进行合并尽可能按顺序写入，从而再次最大程度提高机械 HDD 性能。

在全闪存配置中，一个指定 SSD 闪存设备用于缓存，而其他 SSD 闪存设备用于容量。在全闪存磁盘组配置中，有两种类型的 SSD：充当写缓存的非常快速且持久的闪存设备，以及充当容量的更经济高效的 SSD 设备。其中缓存层 SSD 是 100% 分配用于写入。没有闪存缓存用于读取；容量层闪存 SSD 提供的读取性能可绰绰有余地实现高性能。全闪存配置中的缓存 SSD 可以容纳更多写入，并且写入仅在需要时才写入到容量中，这延长了容量层 SSD 的使用寿命。

尽管这两种配置都可显著提高 vSAN 上运行的 VM 的性能，不过全闪存配置可提供可预测性更强且统一的性能（与工作负载无关）。

读缓存：基本功能

读缓存（仅存在于混合配置中）保留最近读取的磁盘数据块的集合。这可在发生缓存命中时减少 I/O 读取延迟，即可以从缓存（而不是机械磁盘）获取磁盘数据块。对于给定 VM 数据块，vSAN 始终从相同的复制副本/镜像进行读取。当存在多个复制副本（以便承受故障）时，vSAN 会在复制副本拷贝之间均匀划分数据块的缓存。

如果从第一个复制副本读取的数据块不在缓存中，则会引用目录服务以发现数据块是否存在于群集中其他镜像的缓存中（在其他主机上）。如果在其中找到数据块，则检索数据。如果数据块不在其他主机上的缓存中，则发生读缓存未命中。在这种情况下，会直接从机械 HDD 检索数据。

写缓存：基本功能

写缓存（在混合和全闪存配置中都存在）用作非易失性写缓冲区。这可极大提高混合和全闪存配置中的性能，还可延长全闪存配置中的闪存容量设备的使用寿命。写入数据写入到缓存中时，vSAN 会确保在群集中的其他位置写入数据的拷贝。使用 vSAN 部署的所有 VM 都使用默认可用性策略进行设置，该策略可确保至少有 VM 数据的一个额外拷贝可用。这包括确保写入在群集中的多个写缓存中结束。

来宾操作系统中运行的应用程序发起一个写入之后，该写入便会复制到包含存储对象复制副本的主机上的写缓存。这意味着在发生主机故障时，数据的拷贝处于缓存中，不会发生数据丢失。VM 只需使用缓存数据的复制拷贝即可。

闪存耐用性

闪存耐用性与高速缓存层闪存 SSD 在其可用周期内可以承受的写入/擦除周期数相关。根据预期的写入活动级别，必须为每个应用程序或使用情形选择合适的耐用性级别，这很重要。

对于 vSAN 和 VxRail 配置，耐用性规格使用每天写入的 TB 数 (TBW) 作为确定耐用性的指标。对于全闪存 vSAN 部署，高速缓存设备的规格是每天 4 TBW，这是写入密集型工作负载的适当耐用性。

vSAN 对闪存耐用性的影响

有两种常用方法可提高 NAND 闪存耐用性：改进磨损均衡以及更大程度减少写入活动。侧重于在虚拟机所驻留的相同节点上本地化数据的分布式存储实施可防止跨群集中的所有驱动器分布写入。此本地化不可避免地会增加驱动器使用量，从而导致较早地更换驱动器。

与之相反，vSAN 会跨 VxRail 群集中的所有磁盘组分布 VM 的对象和组件。此分布通过推迟写入来显著改进磨损均衡并减少写入活动。vSAN 还使用数据减少技术（如重复数据消除和压缩）来减少写入。

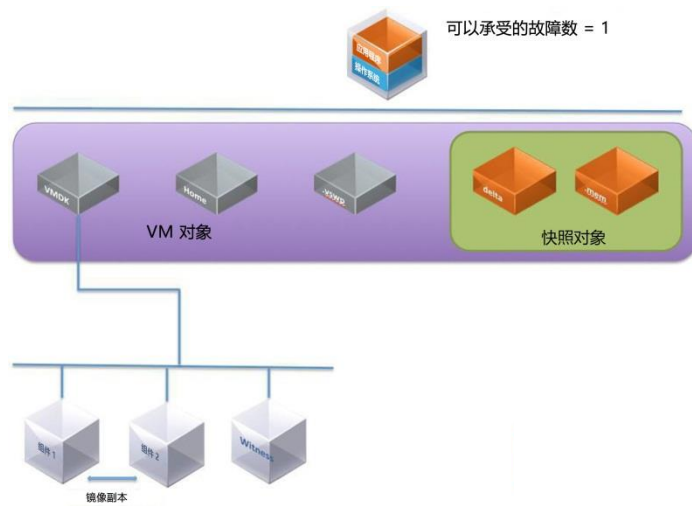
客户端缓存

混合和全闪存配置中都会使用客户端缓存。它利用节点中的本地 DRAM 服务器内存 (客户端缓存) 加速读取性能。分配的内存量是每个主机 0.4% -1 GB。vSAN 会首先尝试从本地客户端缓存履行读取请求，因此 VM 可以避免跨网络完成读取。如果数据在客户端缓存中不可用，则查询缓存层 SSD 以履行读取请求。客户端缓存有益于读缓存友好型工作负载。

对象和组件

VxRail 虚拟机由一组对象组成。例如，VMDK 是对象，快照是对象，VM 交换空间是对象，而 VM 主命名空间（存储 .vmx 文件、日志文件等的位置）也是对象。请参见下图。

图 45 vSAN 对象和组件



基于在应用于 VM 对象的存储策略中定义的性能和可用性要求将虚拟机对象拆分为多个组件。例如，如果 VM 使用可承受单个故障的策略进行部署，则对象具有两个复制副本组件。分布式存储使用磁盘分条过程在多个设备间分布数据块。条带本身指的是数据存储片；条带化设备是保存条带的单个驱动器。如果策略包含条带宽度，则对象会在容量层中的多个设备间进行条带化，每个条带都是一个对象组件。

每个 vSAN 主机最多具有 9,000 个组件。最大组件大小是 255 GB。对于大于 255 GB 的对象，vSAN 会自动将它们划分为多个组件。例如，62 TB 的 VMDK 会生成超过 500 个的 255 GB 组件。上图展示如何基于可以承受的故障数策略，在节点上的驱动器间分布组成 VM 对象的组件。

见证组件

在 vSAN 中，在对象配置为至少可承受一个故障并且使用镜像作为故障容错方法 (FTM) 时，见证组件是每个存储对象不可或缺的组件。见证组件是不包含任何数据，只包含元数据的组件。其用途是在制定可用性决策以满足可以承受的故障数 (FTT) 策略设置时充当决定因素，在确定群集中是否存在组件仲裁时使用。

在 vSAN 中，存储组件可以采用使它们可保证可用性而无需依赖见证组件的方式进行分布。在这种情况下，每个组件都具有一些投票数 — 至少一个或更多。仲裁基于需要“超过 50% 的投票数”的规则进行计算。

复制副本

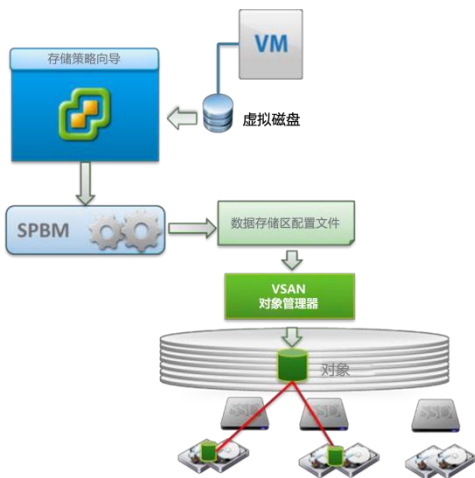
复制副本组成虚拟机的存储对象。为虚拟机指定可用性策略 (FTT) 时，复制副本会进行实例化。可用性策略决定创建的复制副本数，使虚拟机可以完整补充数据的情况下继续运行（即使是在群集中出现主机、网络或磁盘故障时）。

基于存储策略的管理 (SPBM)

vSAN 策略针对性能和可用性定义虚拟机存储要求。这些策略确定如何在数据存储区中调配和分配存储对象，以保证所需服务级别。

vSAN 实施基于存储策略的管理，vSAN 数据存储区中部署的每个虚拟机都分配有至少一个策略。创建 VM 并向它分配存储策略时，会将策略要求推送到 vSAN 层。请参见下图。

图 46 SPBM



存储策略是一组规则，可手动分配给虚拟机，也可以自动分配默认策略。一个系统可能有多个存储策略。例如，在其名称或资源组中包含 PROD-SQL 的所有虚拟机都可以设置为 RAID 1 和 5% 读缓存保留，而 TEST-WEB 会自动设置为 RAID 0。

动态策略更改

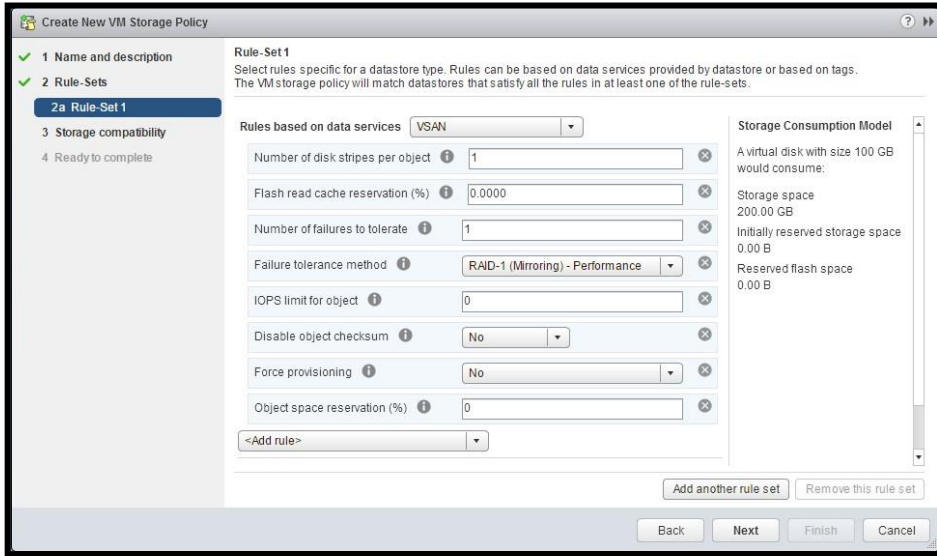
管理员可以动态更改 VM 存储策略。更改属性（如可以承受的故障数）时，vSAN 会尝试查找具有新配置的复制副本新位置。在某些情况下，可以重复使用当前配置的现有部分，随后只需更新或扩展配置。例如，如果某个对象当前使用“可以承受的故障数=1”，而用户要求“可以承受的故障数=2”，则 vSAN 可以只添加另一个镜像（和见证组件）。

在其他情况下（如将条带宽度从一更改为二），vSAN 不能重复使用现有复制副本，它会创建全新的复制副本，而不会影响现有对象。如果属性与策略匹配，则存储对象的状态为“合规”。在线更改策略时，对象可能在 vSAN 重新配置对象期间显示为“不合规”。

存储策略属性

下图显示存储策略的规则集示例。

图 47 vSAN 策略属性



Number of disk stripes per object

此规则建立用于对每个虚拟机复制副本进行条带化的最小容量设备数。大于 1 的值可能会产生更好的性能，但是也会导致更高的资源消耗。默认值最小值为 1，最大值为 12。条带大小为 1 MB。

vSAN 可能会在没有任何条带宽度策略要求的情况下，决定某个对象需要在多个磁盘间进行条带化。这样做的原因可能会有所不同，不过通常发生在 VMDK 太大，无法放置在单个物理驱动器上时。但是，在需要特定条带宽度时，它不应超过对群集可用的磁盘数。

Flash read cache reservation

闪存缓存保留指保留用作虚拟机对象的读缓存的闪存容量，仅适用于混合配置。默认情况下，vSAN 会按需向存储对象动态分配读缓存。因此，通常无需更改此参数的默认值 0。

在非常特殊的情况下，当少量增加单个 VM 的读缓存可以使性能发生显著变化时，可选择这样做。应谨慎地使用它以避免浪费资源或从其他 VM 抢占资源。

默认值为 0%。最大值为 100%。

Failures to tolerate

此 FTT 选项通常定义虚拟机对象可以承受的主机和设备故障数。如果承受 n 个故障，则将创建 $n+1$ 个 VM 对象拷贝，并且需要 $2n+1$ 个具有存储的主机。

默认值为 1。最大值为 3。

当为群集启用了擦除编码（通过设置“FTM=Capacity”）时，如果可以承受的故障数设置为 1，则应用 RAID 5；如果可以承受的故障数设置为 2，则应用 RAID 6。请注意，一个 vSAN 群集对于 RAID 5 至少需要四个节点，对于 RAID 6 至少需要六个节点。

Failure tolerance method

“Failure tolerance method”（FTM，故障容错方法）指定是否针对性能或容量优化数据复制方法。RAID 1 故障容错方法可提供更好性能并占用更少内存和网络资源，但会使用更多磁盘空间。RAID 5/6 擦除编码可提供更多可用容量，但会占用更多 CPU 和网络资源。（后面有关擦除编码的部分会提供更多信息。）

IOPS limit for object (QoS)

此属性通过对 VM/VMDK 可以执行 IOPS 数量定义上限，来为对象建立服务质量 (QoS)。默认行为是所有对象都不受限制，低优先级应用程序占用资源的方式可能会影响更重要的工作负载。此规则允许对不同应用程序设置不同限制，可以用于防止工作负载相互影响（吵闹的邻居问题），或是为不同工作负载提供服务级别协议 (SLA) 并为第 1 层应用程序维护性能。

基于此 VM/VMDK 上的所有操作（读/写，包括快照）计算 IOPS。

关于对象的 IOPS 限制的一些说明：

计算 IOPS 时，读取和写入被认为是等效的，但请记住，不考虑缓存命中率和顺序性。

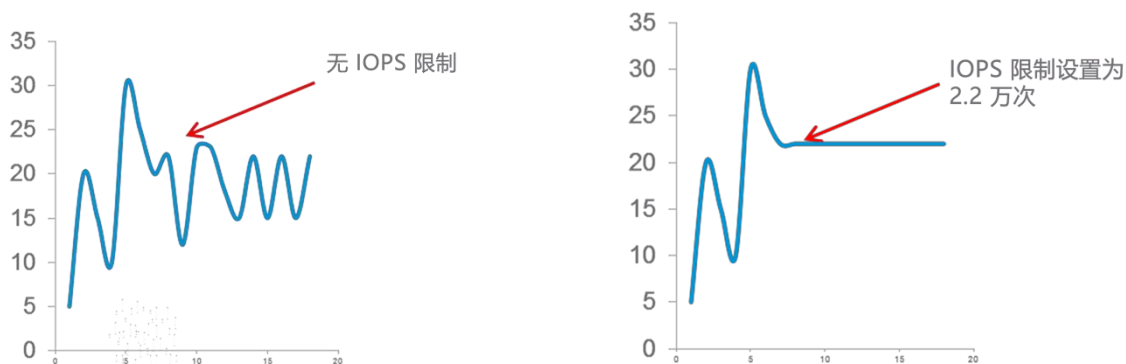
对象超过其磁盘 IOPS 限制时，I/O 操作会受到限制。

如果对象的 IOPS 限制设置为 0，则不实施 IOPS 限制。

vSAN 允许对象在操作的第一秒或是在一段时间的非活动状态之后使 IOPS 限制率加倍。

下图展示 IO 如何在设置了 IOPS 限制策略时受到限制。

图 48 IOPS 限制确立 QoS



Disable object checksum

vSAN 通过确认对象的 4K 数据区块的每个拷贝是否完全相同，来使用端到端校验并确保数据完整性。校验和的长度为五个字节，随数据一起保留。写入数据时，会验证校验和以确保在数据通过网络时没有发生损坏。读取数据时，校验和会验证读取的数据是否与写入的数据匹配。如果检测到错误，则 vSAN 使用复制副本恢复数据并记录错误。校验和计算和纠错对用户是透明的。

群集中所有对象的默认设置为“On”，这意味着校验和处于启用状态。最佳做法是不更改此设置，因为检测数据损坏是 vSAN 的一个关键且有价值的功能，不应禁用。

Force provisioning

如果此选项设置为“**Yes**”，则即使在数据存储区无法满足存储策略中指定的规则时，也会调配对象。

当资源受限且无法满足正常资源调配策略时，有时会在中断期间使用此参数。

默认值为“**No**”，如果无法满足策略规则，则不允许创建对象。这适用于大多数生产环境。如果设置为“**Yes**”，则即使没有足够资源可用于满足策略规则，也可以创建对象。在这种情况下，对象会显示为“**不合规**”。

Object space reservation

“Object space reservation”指定为总对象大小的百分比。它反映部署虚拟机所需的密集调配保留空间。

默认值为 0%。最大值为 100%。

值应设置为 0% 或 100%（将 RAID-5/6 与重复数据消除和压缩结合使用时）。

稀疏交换

稀疏交换是一种由于其空间效率而值得强调的 vSAN 功能，可用于全闪存和混合环境。默认情况下，vSAN 交换文件会进行密集调配 — 使用 100% 空间预留进行创建。（例如，如果 VM 具有 4 GB 的 RAM，则交换文件也具有 4 GB 的 RAM。）虽然这可为 VM 保证充足的容量，不过可能会占用太多内存 — 尤其是在具有大量虚拟机的大型群集中。稀疏交换作为高级主机设置启用时，VM 为交换对象保留的内存空间会少于 100%。

稀疏交换通过启用 *Swap thick provision disabled* 设置来建立。稀疏交换文件的 FTT 设置为 1，FTM 设置为 RAID1（镜像）。

在尚未过量使用内存的环境中，特别是对于交换文件表示所需总数据存储区容量的相当大一部分的全闪存 VxRail 环境，利用稀疏交换是一种减少内存开销的有效方法。

I/O 路径和缓存算法⁴

此部分详细说明了迄今为止随有关 vSAN 缓存算法的其他常规信息介绍的一些 vSAN 概念。后续各段简要介绍 vSAN 如何利用闪存、内存和旋转磁盘。它们还阐述了来宾操作系统与永久存储器区域之间的 I/O 路径。

读缓存

每个磁盘组都包含一个用作缓存层的 SSD 驱动器。在混合系统上，默认情况下将 70% 的容量用于读缓存 (RC)。最活跃的数据保留在缓存层中，通过最大程度减小从机械磁盘进行读取所带来的延迟影响来提高性能。

RC 按照缓存行进行组织。它们表示 RC 中的数据管理单元，当前大小为 1 MB。数据会提取到 RC 中，按缓存行粒度进行回收。除了 SSD 读缓存，vSAN 还维护小型内存中 (RAM) 读缓存，其中保存从 RC 最近访问的缓存行。内存中缓存会基于系统中的可用内存来动态调整规模。

vSAN 会维护内存中元数据，这些元数据用于跟踪 RC (SSD 和内存中) 的状态，包括缓存行的逻辑地址、每个缓存行中的有效和无效区域、期龄信息等。这些数据结构旨在进行压缩以提高效率，从而在不对常规操作造成巨大 CPU 开销的情况下使用内存空间。无需将 RC 元数据换入或换出永久存储器。(这是 VMware 在其中保存重要 IP 的区域)。

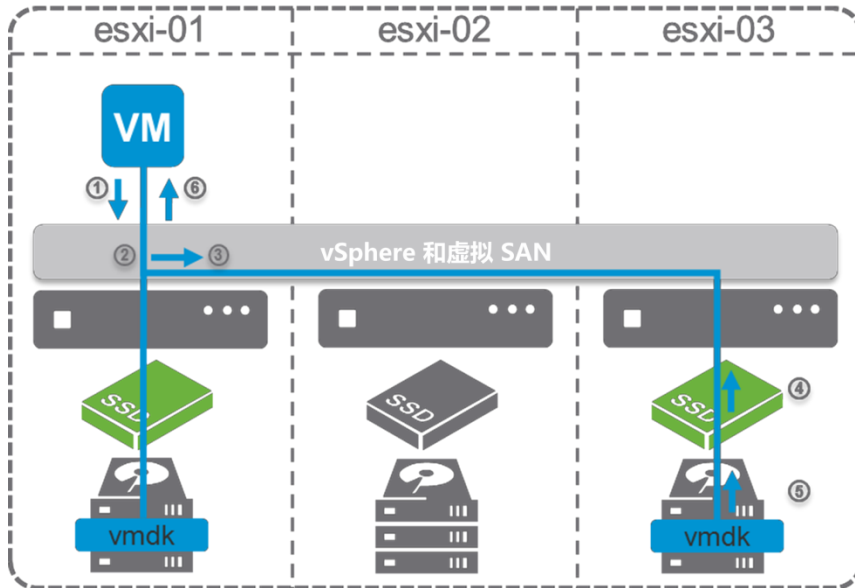
不在主机的电源周期操作间根据读缓存内容。如果电源中断，然后恢复，则 RC 会从头开始重新填充 (预热)。因此，RC 实质上用作快速存储层，无需在电源周期间保持。这种方法的基本原理是避免每次修改 RC (如缓存行提取和回收，或是在写入操作使缓存行的子区域无效时) 都保持 RC 元数据时所需要的常见数据路径上的任何开销。

⁴ 此特定部分中的很多内容提取自现有技术白皮书：《An overview of VMware VSAN Caching Algorithms》(VMware VSAN 缓存算法概述)。

混合读取剖析

读取操作遵循一个已定义过程。为进行阐明，下面示例中 VMDK 在 esxi1 和 esxi3 上具有两个复制副本。请参见下图。

图 49 混合读取



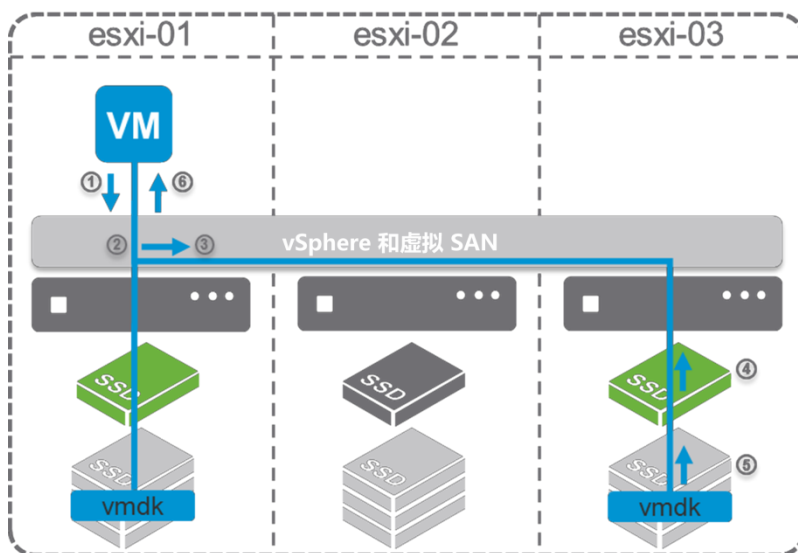
1. 来宾操作系统在虚拟磁盘上发出读取操作
2. 所有者选择从中读取数据的复制副本
 - 在复制副本之间平衡负载
 - 不一定是本地复制副本（如果有一个）
 - 始终从同一个复制副本读取数据块
3. 在选择的复制副本 (esxi-03) 上：从闪存写缓冲区（如果存在）读取数据
4. 在选择的复制副本 (esxi-03) 上：从闪存读缓存（如果存在）读取数据
5. 否则，从 HDD 读取数据并将数据放置在闪存读缓存中
 - 为缺少的缓存行分配 1 MB 缓冲区并替换“最冷”数据（清除最冷的数据，为新读取腾出空间）
 - 以 64 KB 区块倍数的形式从 HDD 读取每个缺少的行（从包含引用的数据的区块开始）
6. 将数据返回给所有者
7. 完成读取并将数据返回 VM
8. 1 MB 缓存行添加到内联读缓存后，其填充会继续以异步方式进行。这是为了探索引用的空间和时间局部性，不断增加后续读取会在读缓存中找到的更改。

全闪存读取剖析

1. 来宾操作系统在虚拟磁盘上发出读取操作
2. 所有者选择从中读取数据的复制副本
 - 在复制副本之间平衡负载
 - 不一定是本地复制副本（如果有一个）
3. 在选择的复制副本 (esxi-03) 上：从闪存写缓冲区（如果存在）读取数据
4. 否则，从容量闪存设备中读取
5. 将数据返回给所有者
6. 完成读取并将数据返回 VM

下图显示了此操作。

图 50 全闪存读取



主要区别在于读缓存未命中不会导致严重的性能下降。从闪存容量设备进行读取的速度应几乎与从缓存 SSD 进行读取的速度一样快。另一个重要区别是无需如混合配置一样，将数据块从容量层移动到缓存层。

写缓存

在混合配置中，回写缓存完全是为了性能。虚拟化基础架构中的聚合存储工作负载几乎总是随机的，这是因为对共享基础架构的许多 VM 和应用程序进行统计多路复用。

HDD 只能执行少量随机 I/O，与 SSD 相比具有高延迟。因此，将工作负载的随机写入部分直接发送到旋转磁盘可能会导致性能下降。另一方面，磁盘对于顺序工作负载表现出了出色的性能。现代 HDD 即使在工作负载不是完全按顺序时，也可能会表现出类似于顺序的行为和性能。“临近 I/O”可足够使用。

在混合磁盘组中，vSAN 使用 SSD 的写缓冲区 (WB) 部分 (默认情况下为设备容量的 30%) 作为暂存所有写入操作的回写缓冲区。关键目标是采用为构成容量层的 HDD 创建无害的近似顺序 (临近) 写入工作负载的方式转储写入的数据 (不是各个写入操作)。

在全闪存磁盘组中，vSAN 完全利用第 1 层 SSD 作为回写缓冲区 (设备容量的 100% — 最大 600 GB)。在这种情况下，WB 的用途大不相同。它可在高耐用性设备中承受最高写入操作速率，只允许将少量数据写入容量闪存。此方法允许为容量使用低耐用性的较大容量 SSD。

然而，容量 SSD 能够处理非常大量的读取 IOPS。因此，不会出现读缓存，除非读取操作引用的最新数据仍驻留在 WB 中。

在混合和全闪存中，每个写入操作都通过事务流程进行处理：操作的记录保留在 SSD 中的事务日志中。

操作的数据 (有效负载) 保留在 WB 中。

更新的内存中表反映新数据及其逻辑地址空间 (用于跟踪) 以及它在容量层中的物理位置。

写入操作在事务成功提交之后完成上游部分。

在典型稳定状态工作负载下，多个写入操作的日志记录会在保留在日志中之前进行合并。这可为 SSD 减少元数据写入操作量。根据定义，日志是循环缓冲区，以顺序方式进行写入和释放。这样可以避免写入放大 (可很好地提高设备耐用性)。WB 区域以循环方式分配数据块 (考虑到磨损均衡)。

即使在写入操作覆盖现有 WB 数据时，vSAN 也绝不会重写相应位置处的现有 SSD 页面。相反，它会分配新数据块，并更新元数据以反映旧数据块无效。vSAN 会先填充整个 SSD 页面，然后再移动到下一个页面。最终，在所有数据都无效时，会释放整个页面。(重新缓冲数据以便可以释放 SSD 页面的情况非常少见。)

此外，因为设备固件无法看到失效数据，所以它看不到页面中的“孔”。事实上，内部写入均衡 (通过四处移动数据以填充页面中的孔) 几乎已消除。这可延长设备的整体耐用性。vSAN 经过用心设计，尽可能减少不必要的写入操作，以便更大限度地提高高速缓存 SSD 耐用性。因此，在 vSAN 中实施的 SSD 预期使用寿命可能会超过制造商的规格 (这些规格是在考虑更加一般的工作负载的情况下制定的)。

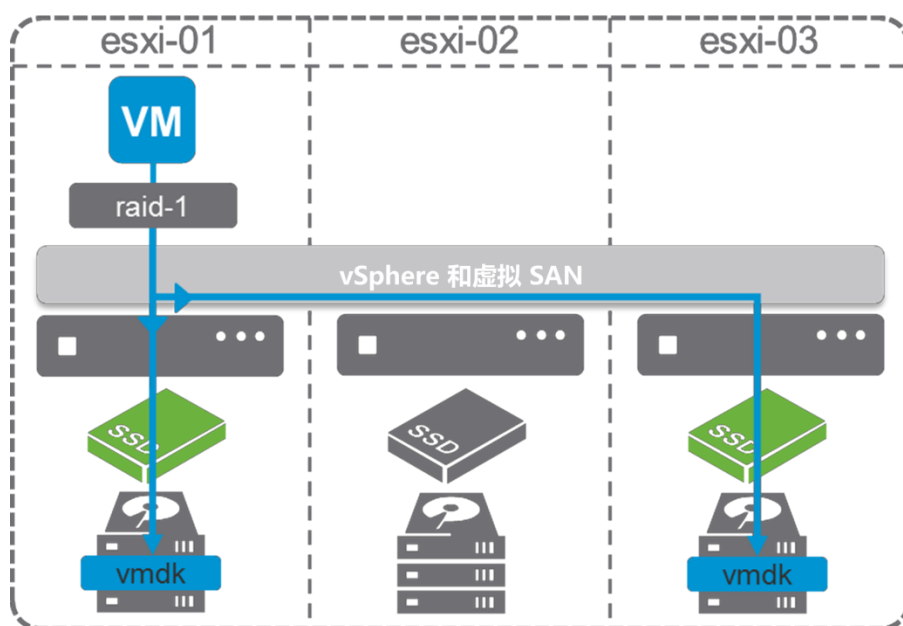
写 I/O 剖析 — 混合和全闪存 (FTM=镜像)

1. VM 运行在主机 esxi-01 上。
2. esxi-01 是虚拟磁盘对象的所有者。
 - 可以承受的故障数 = 1
3. 对象在 esxi-01 和 esxi-03 上具有两 (2) 个复制副本。
4. 来宾操作系统向虚拟磁盘发送写入操作。

5. 所有者克隆写入操作。
 - 并行：将写入操作发送到 esxi-01（本地）和 esxi-03
6. esxi-01、esxi-03 将操作保留在闪存（日志）中。
7. esxi-01、esxi-03 向所有者确认写入操作。
8. 所有者等待两个写入操作的确认并完成 I/O。
9. 之后，后端主机提交写入批处理。

下图说明了该操作。

图 51 混合和闪存写入 I/O



分布式高速缓存注意事项

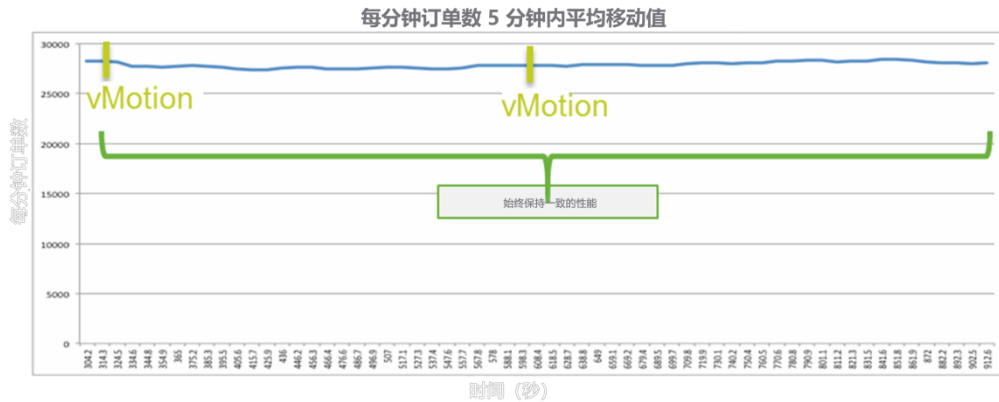
vSAN 缓存算法和数据局部性技术反映了一些与分布式缓存相关的目标和观察结果：

vSAN 利用时间和空间局部性进行缓存。

vSAN 在整个群集中的闪存上实施分布式持久缓存。缓存在磁盘（数据复制副本存在于其中）前面（而不是在客户端）进行。分布式缓存机制会产生更好的整体闪存缓存利用率。

分布式缓存的另一个优势体现在 VM 迁移过程中（在一些数据中心内，这种迁移可能一天进行十次以上）。借助 DRS 和 vMotion, VM 可以在群集中从主机移动到主机。如果不使用分布式缓存，则在每次进行 VM 迁移时，迁移都必须移动大量数据并重新预热缓存。如下图所示，vSAN 可防止在 VM 迁移之后出现任何性能下降。

图 52 vSAN 防止在 VM 迁移之后出现性能下降



访问另一个主机上的数据时，网络会引入少量延迟。10 GbE 网络中的典型延迟的范围为 5 – 50 毫秒。通过 SCSI 层访问的闪存驱动器的典型延迟对于小型 (4 K) I/O 数据块是几乎 1 毫秒。因此，对于在系统中执行的大部分 I/O，网络影响会使延迟几乎增加 0.1%。

几乎没有工作负载实际上是缓存友好型，这意味着它们不会受益于缓存大小的少量增加可以显著提高 I/O 速率的方式。这些工作负载可以受益于本地缓存，启用客户端缓存会是正确的方法。

vSAN 使用查看加速器 (进行了重复数据消除的内存中读缓存)，这对于 VDI 使用情形尤其有效。

vSAN 采用客户端缓存，这种缓存利用虚拟机的本地 DRAM 内存来加速读取性能。在任何位置，分配的内存量对于每个主机都是 0.4% 到 1 GB。

vSAN 高可用性和故障域

vSAN 策略属性可建立用于防范节点故障的参数，不过它们可能不是针对机架故障这类事件建立承受能力的最有效或最高效的方式。此部分回顾 VxRail 系统上的 vSAN 群集可用性功能。它首先讨论节点少于四个的小型 VxRail 部署中的可用性意义。

故障域概述

vSAN 和 VxRail 系统使用故障域作为针对机架和站点故障配置承受能力的方式。默认情况下，节点被视为故障域。vSAN 会在故障域间分布组件，因此默认情况下，vSAN 会在节点间分布组件。例如，考虑一个具有四 (4) 个四节点 VxRail 系统 (各自放置在不同机架中) 的群集。通过将每个四节点系统显式定义为单独的故障域，vSAN 会在不同机架间分布冗余组件。

在实施方面，不属于其他故障域一部分的任何主机都被视为其自己的单主机故障域。VxRail 需要至少三个故障域，并且各自具有至少一个主机。故障域定义可识别表示域本身的物理硬件构造。启用域之后，vSAN 便将活动虚拟机存储策略应用于整个域，而不是只应用于单个主机。群集中的故障域数基于 FTT 属性进行计算： $(故障域数) = 2 * (可以承受的故障数) + 1$ 。管理员可以从 vSphere Web Client 管理故障域 (如下图所示)。

图 53 管理故障域



故障域和机架级故障

故障域机制可以检测配置易受攻击的情况。考虑一个包含四个服务器机架（各自具有两个节点）的群集。如果 FTT 设置为 1，并且未启用故障域，则 vSAN 可能会使用相同机架中的主机存储对象的两个复制副本。如果是这种情况，则应用程序会面临潜在的机架级故障。如果启用了故障域，则 vSAN 可确保每个保护组件（复制副本和见证组件）都放置在单独的故障域中，确保节点不能同时发生故障。

下图展示一个四机架设置，每个机架都具有两个 ESXi 节点（VxRail 系统中的可用主机的子集）。定义了四个故障域：

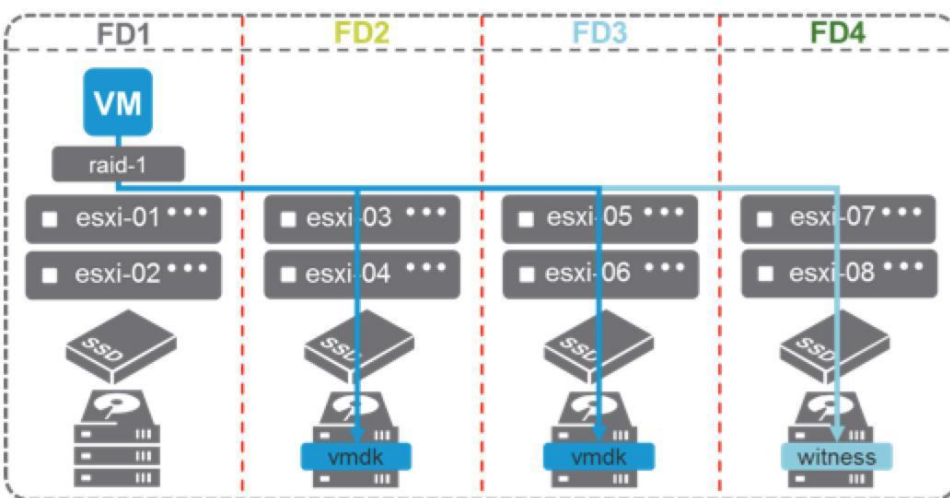
FD1 = esxi-01, esxi-02

FD2 = esxi-03, esxi-04

FD3 = esxi-05, esxi-06

FD4 = esxi-07, esxi-08

图 54 四机架 VxRail 配置的故障域



此配置可保证对象的复制副本存储在不同机架存储模块的主机中，从而确保出现机架级故障时的可用性和数据保护。

部署最低群集配置时的注意事项

部署仅满足最低要求的群集时，务必要了解高可用性的含义。为 RAID-1 保护选择 3 节点最低配置，或为 RAID-5 保护选择 4 节点最低配置，这意味着，如果一台主机出现故障，群集将无法通过在另一台主机上重建数据而进行自我修复。当主机处于维护模式（如节点升级）时，如果再发生故障，则数据会面临潜在故障或不可访问。

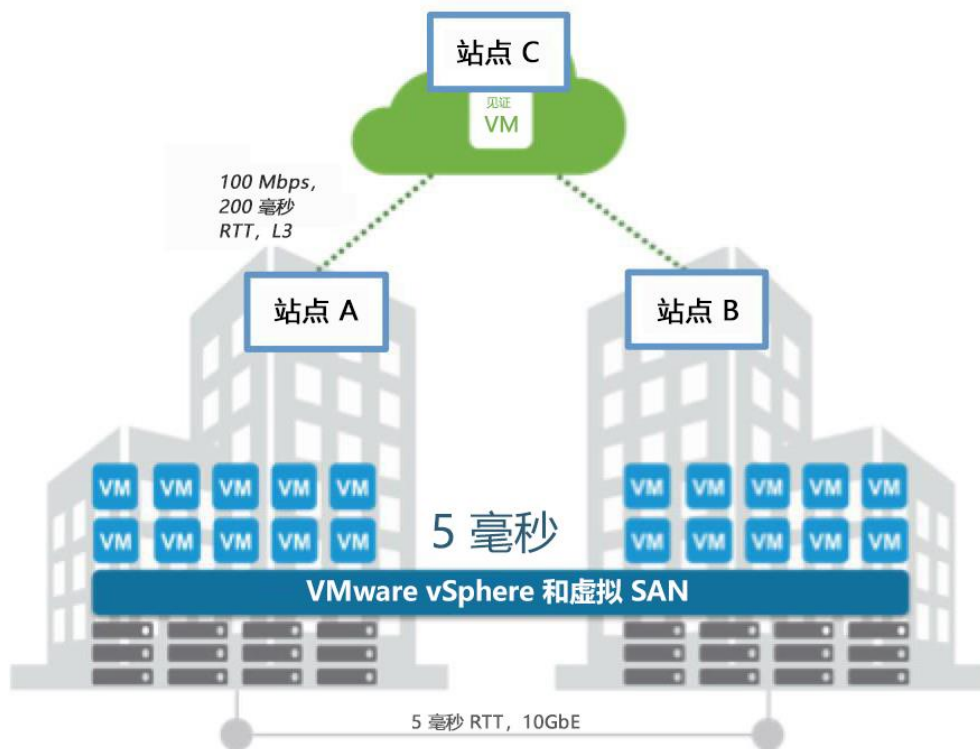
vSAN 延伸群集

延伸群集的概念是 vSAN 与 vSphere 的本机集成的一个好示例。借助 VxRail，延伸群集可扩展大型企业数据中心的可用性。延伸群集是在绝对要求避免数据中心级宕机的环境中实施的特定配置。与故障域针对机架故障实现“机架感知”的方式类似；延伸群集提供了“数据中心感知”，从而维持虚拟机可用性（无论特定数据中心故障情形是怎样的）。

在 VxRail 环境中，具有见证主机的延伸群集是指这样一种部署：一个 vSAN 群集包含两个主动/主动站点，它们具有相同数量的 ESXi 主机，这些主机在它们之间均匀分布。站点通过高带宽/低延迟网络进行连接。

在下图中，每个站点都配置为 vSAN 故障域。用于描述延伸群集配置的命名法是 X+Y+Z，其中 X 是站点 A 上的 ESXi 主机数，Y 是站点 B 上的 ESXi 主机数，而 Z 是站点 C 上的见证主机数。

图 55 延伸 VxRail 群集



延伸群集上部署的虚拟机在站点 A 上具有其数据的一个拷贝，在站点 B 上具有另一个拷贝，并且见证组件放置在站点 C 的主机上。

它是单一配置，通过故障域、主机和 VM 组以及相关性规则来实现。出现整个站点故障时，另一个站点仍具有虚拟机数据的完整拷贝和至少一半的可用资源组件。这意味着所有 VM 都在 vSAN 数据存储区上保持活动状态并可用。恢复点目标 (RPO) 为零，数据恢复时间目标 (RTO) 为零。应用程序 RTO 取决于应用程序可恢复性。

VxRail 支持的最低配置是 3+3+1 (7 个节点)；最高配置是 15+15+1 (31 个节点)。混合和全闪存 VxRail 配置都支持延伸群集。新版 VxRail 软件中运行的延伸群集支持客户驱动的升级。对于运行旧版 VxRail 软件的群集，客户将需要与支持人员联系以协助升级。

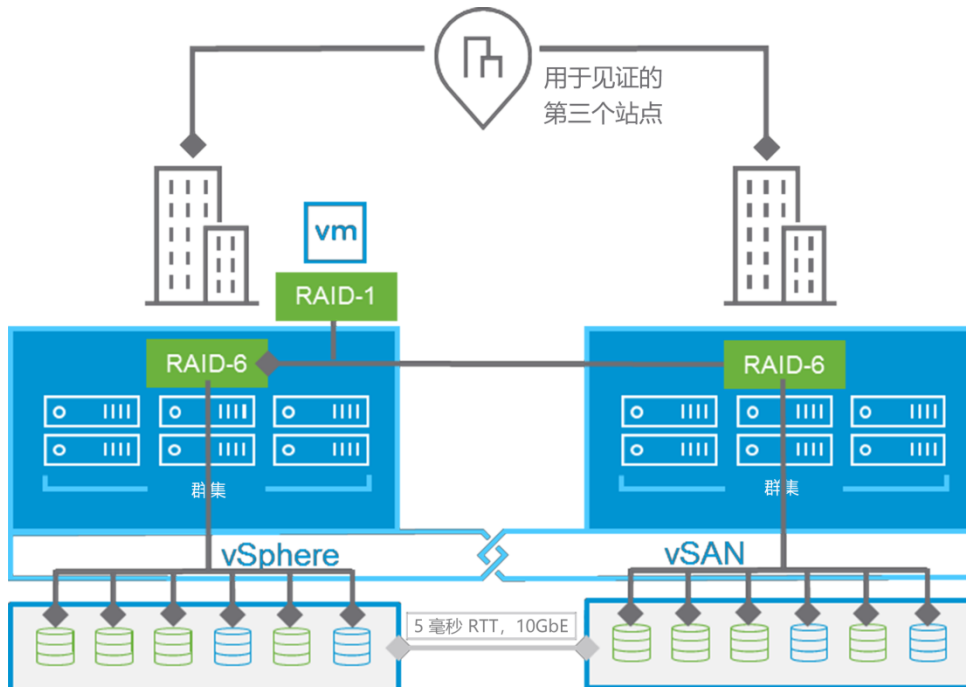
有关详细信息，请参阅《VxRail vSAN Stretched Clusters Planning Guide》(VxRail vSAN 延伸群集规划指南)：<https://vxrail.is/stretchedclusterplanning>

采用本地保护的延伸群集

VxRail 软件版本 4.5 和 vSAN 6.6 或更高版本支持采用本地保护的延伸群集。此功能会在站点之间镜像数据，每个站点应用本地数据保护以增强保护能力。使用两个参数来指定保护：可以承受主要故障 (PFTT) 和可以承受次要故障 (SFTT)。PFTT 是指始终为 RAID1 镜像的站点之间的保护。SFTT 是在每个站点应用的本地保护。混合配置支持采用 RAID1 (镜像) 故障容错方法 (FTM) 的 SFTT 0、1、2 或 3。全闪存配置支持采用 RAID1 FTM 的 SFTT 0、1、2 或 3，或支持采用擦除编码 FTM 的 SFTT 1 或 2。

下图显示了全闪存延伸群集配置的本地保护。

图 56 采用本地保护的延伸群集



站点局部性

在传统存储群集配置中，读取在复制副本间进行分布。在延伸群集配置中，vSAN 分布式对象管理器 (DOM) 还会考虑对象的故障域，并且仅从相同域中的复制副本进行读取。这样，它便可避免与使用站点间网络执行读取关联的任何延迟时间。

网络

第 2 层（相同子网）和第 3 层（路由）配置都用于延伸群集部署。第 2 层或第 3 层连接在数据站点之间配置，而第 3 层连接在见证主机与数据站点之间配置。

数据站点之间的带宽取决于工作负载，但 Dell EMC 要求 VxRail 系统在延伸群集配置中至少可处理 10 Gbps 工作负载。见证主机支持的延迟为最多 200 毫秒 RTT，并且对每 1,000 个 vSAN 对象使用 2 Mbps 的带宽。另请记住，数据站点之间的延迟不应超过 5 毫秒（在 500 千米或大约 310 英里距离上通常估计是此延迟）。

延伸群集心跳和站点偏向

延伸群集配置实际上具有三个故障域。第一个故障域充当首选数据站点，第二个是辅助数据站点，第三个只是见证主机站点。

vSAN 主节点位于首选站点上，vSAN 备份节点位于辅助站点上。只要节点（ESXi 主机）在首选站点上可用，则始终从此站点上的一个节点选择主节点 — 对于辅助站点也是类似情况（只要节点在辅助站点上可用）。

主节点和备份节点会每秒发送心跳。如果心跳通信丢失时间达到五个连续心跳（五秒），则见证主机被视为出现故障。如果见证主机遭受永久故障，则可以配置新的见证主机并将它添加到群集。首选站点会在进行分区时获得所有权。

出现完全故障之后，主节点和备份节点最终都会是唯一剩余的实时站点。出现故障的站点恢复之后，它会继续承担其指定角色（首选或辅助），主节点和辅助节点会迁移到其相应位置。

对于与见证主机之间的通信，如果心跳暂停时间达到五个连续心跳，则主节点会假定见证主机出现故障。如果这是永久故障，则需要配置新的见证主机并将它添加到群集，首选站点会在进行分区时获得所有权。

延伸群集的 vSphere HA 设置

延伸群集需要以下 vSphere HA 设置：

默认情况下，主机监视在所有 VxRail 部署（当然包括延伸群集配置）中都处于启用状态。此功能也使用网络心跳来确定参与群集的主机的状态。它会指示可能需要进行补救，如在其他群集节点上重新启动虚拟机。

配置准入控制可确保 vSphere HA 具有足够的可用资源，以便在出现故障之后重新启动虚拟机。这在延伸群集中甚至可能比单站点群集中更加显著，因为它可使整个多站点基础架构具有弹性。

工作负载可用性可能是大多数延伸群集实施的主要动机。

部署需要足够容量来应对整个站点故障。由于延伸群集在站点之间均匀划分 ESXi 主机数，因此 Dell EMC 建议为 CPU 和内存将准入控制策略配置为 50%，以确保所有工作负载都可以由 vSphere HA 重新启动。

双节点配置

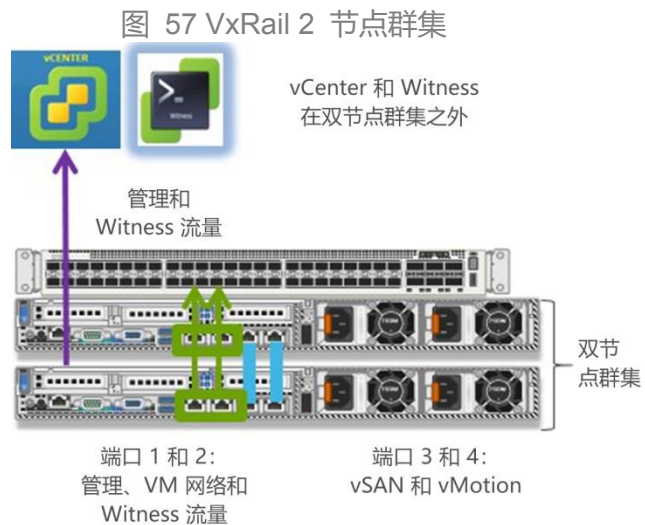
VxRail 软件版本 4.7.100 引入了对固定的 2 节点配置的支持，该配置使用两个节点之间的直接连接进行群集操作。它的占用空间小，可为空间和工作负载要求有限的位置提供合理、经济高效的解决方案。该支持要求配置为全新的部署，这意味着现有群集无法利用节点删除来转换为 2 节点配置。不支持群集扩展，因此该解决方案应仅针对特定的使用情形。2 节点群集中仅支持 VxRail E560 和 E560F。用户仍然可以受益于自动生命周期管理。

该配置包含两个直接连接的节点和一个用于为群集提供仲裁的见证组件。直接连接用于 vSAN 和 vMotion 流量。见证组件是在 ESXi 主机上安装的虚拟设备，必须驻留在 2 节点群集的外部，即在另一个数据中心或同一机架/位置中的另一台物理主机上。见证组件与两个节点有单独的连接，这需要 VLAN 将见证管理流量与 vSAN 流量分隔开来。该配置仅支持镜像 (FTT=1)。见证主机用作 tiebreaker。每个节点和见证组件是一个单独的故障域，群集中总共有三个故障域。

首次运行体验中的特殊工作流用于部署 2 节点群集。该工作流包括见证设备的设置和见证组件流量分离。该配置必须使用客户提供的 vCenter 进行管理。群集中只有两个数据节点，用户需要了解群集负载，以便防止在节点发生故障时以及单个节点为整个群集工作负载提供服务时数据不可用。

2 节点群集支持按插槽计数的 vSAN 标准、高级和企业许可证版本。请参阅 [《vSAN Licensing Guide》](#) (vSAN 许可指南)，了解更多详细信息。

有关此配置的更多详细信息，请参阅 [VxRail 技术演示文稿](#)。



快照

快照已存在有一段时间了，它们作为在特定时间点 (PIT) 捕获数据对象状态的方式，以便可以在出现逻辑或物理故障之后根据需要回滚到该状态。对于 VxRail 解决方案，管理员可以在 vSphere Web Client 中使用 Snapshot Manager 创建、回滚或删除 VM。每个 VM 都支持一个由最多 32 个快照组成的快照链。

虚拟机快照通常包括设置 (.nvram 和 .vmx) 以及所有 VM 关联磁盘的电源状态和 (可选) 内存状态。具体而言，每个快照都包含：

增量磁盘：保留创建快照时的虚拟磁盘状态。这种情况发生时，来宾操作系统无法向其 .vmdk 文件写入。相反，会在名为 VM_name-delta.Vmdk 的备用文件中捕获更改。

内存状态文件：VM_name-Snapshot#.Vlms，其中 # 是序列中的下一个编号 (从 1 开始)。此文件保存自创建快照起的内存状态。如果捕获内存，则此文件的大小是虚拟机最大内存的大小。如果未捕获内存，则此文件要小得多。

磁盘描述符文件：VM_name-00000#.vmdk，包含有关快照的信息的小型文本文件。

快照增量文件：VM_name-00000#-delta.Vmdk，其中包含在创建快照时对虚拟磁盘数据进行的更改。

VM_name.Vmsd：此快照列表文件在部署虚拟机本身时创建。它维护在 vSphere Web Client 中进入快照列表的 VM 快照信息。此信息包括快照 .Vlms 文件的名称和虚拟磁盘文件的名称。

快照状态使用 .Vlms 扩展名，存储创建快照时的必需 VM 信息。每个新的 VM 快照都会生成新的 .vmsn 文件。此文件的大小会因创建过程中选择的选项而异。例如，包括虚拟机的内存状态会增加 .vmsn 文件的大小。它通常包含每个快照的 VMDK 名称、显示名称和描述以及标识符。

可能还存在其他文件。例如，已开启的虚拟机的快照具有关联的 snapshot_name_number.vmem 文件，其中包含来宾操作系统的主内存 (保存为快照的一部分)。

停顿选项可用于为已开启的 VM 维护一致的时间点拷贝。VMware 工具可能会使用自己的同步驱动程序或使用 Microsoft 的 Volume Shadow Copy Service (VSS)，不仅停顿来宾操作系统文件系统，而且停顿理解 VSS 指令的任何 Microsoft 应用程序。

使用重复数据消除和压缩的存储效率

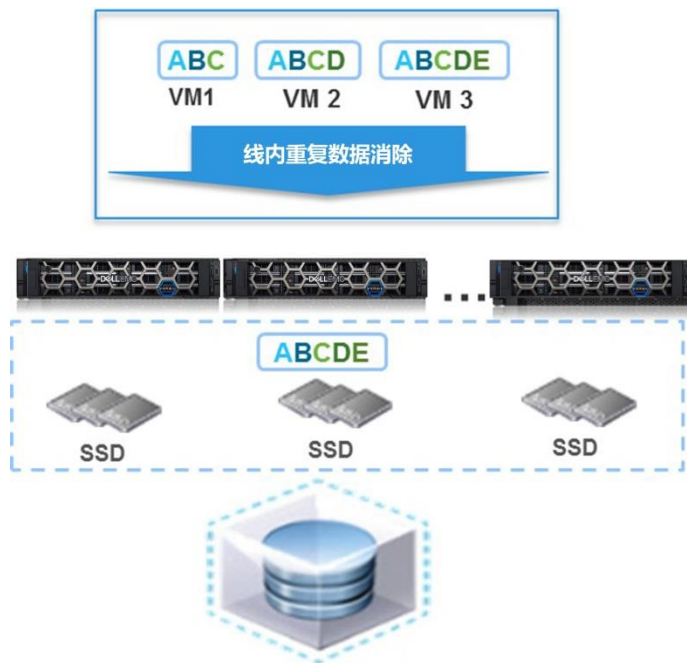
存储容量要求持续呈指数级增长，IT 组织正在寻找方法来提高存储效率，从而以最低成本满足其不断增长的容量要求。实现此目标的一种方法是使用重复数据消除和压缩，这可以按更低成本提供更多容量。许多环境可以实现的有效容量是原始容量的两倍。在全闪存 VxRail 群集中可以利用这些数据减少功能。

压缩和重复数据消除技术已实施多年，但由于实施所需的开销和系统资源而尚未得到广泛采用。目前，每个处理器具有许多核心和大量内存的 VxRail 全闪存型号是主力产品！连同 vSAN 的体系结构效率一起，空间节省不仅仅是抵销轻微的开销。与混合 HDD 解决方案相比，VxRail 全闪存配置通常以更低成本提供更多有效容量。

借助 vSAN，当数据从缓存转储到容量驱动器时，重复数据消除和压缩以内联方式进行。首先通过删除包含完全相同数据的数据块的冗余拷贝，对数据进行重复数据消除。这在 4K 数据块级进行。

下图显示一个典型虚拟机环境。

图 58 线内重复数据消除



虽然所有 VM 都是唯一的，不过它们会共享一定量的通用数据。相同数据块在 SSD 保存一次（而不是保存相同数据的多个拷贝），并使用在容量层中维护的元数据跟踪对唯一数据块的引用。

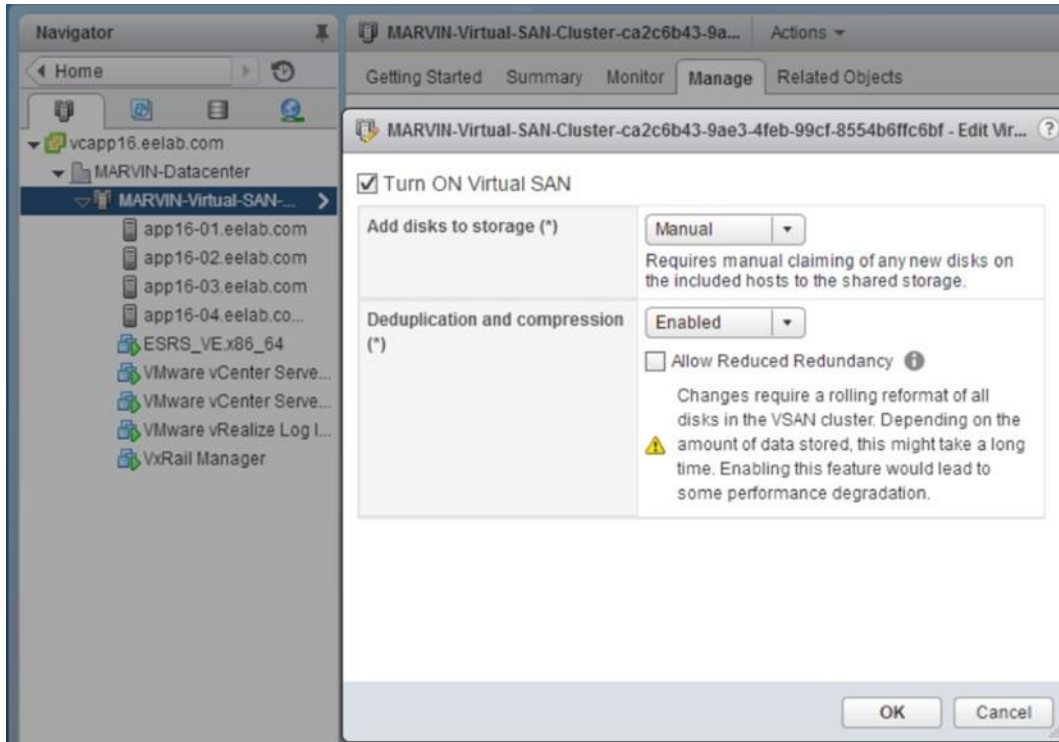
重复数据消除算法在磁盘组级别进行应用，对每个磁盘组会生成每个唯一 4K 数据块的单个拷贝。虽然进行重复数据消除的数据块可能跨多个磁盘组存在，不过通过将重复数据消除域限制到一个磁盘组，可无需全局查找表，这样可更大程度地降低网络开销和 CPU 利用率。

在数据块进行了重复数据消除之后以及在写入 SSD 之前，应用 LZ4 压缩。如果压缩产生 2 KB 或更小的数据块大小，则数据块的压缩版本会永久保存在 SSD 上。如果压缩未产生小于 2 KB 的数据块大小，则完整的 4K 数据块会写入 SSD。

几乎所有工作负载都可在一定程度上受益于重复数据消除。但是具有高度冗余数据的典型虚拟服务器工作负载（如完整克隆虚拟桌面或同构服务器操作系统）获益最大。压缩提供了进一步的数据减少。文本、位图和程序文件具有很高的可压缩性，通常可以达到 2:1。已压缩的其他数据类型（例某些图形格式和视频文件或加密文件）则可能略微减少或不减少。

重复数据消除和压缩在默认情况下处于禁用状态，并且在群集级别同时启用。（参见下图。）虽然可以随时启用，不过建议在系统进行初始设置时启用，以避免不得通过后处理（而不是以线内方式进行）对现有数据进行重复数据消除和压缩所产生的开销和潜在性能影响。

图 59 重复数据消除和压缩已启用



重复数据消除和压缩开销

重复数据消除算法会将数据文件分解为连续的数据段，或计算指纹（用于标识重复数据段并减少数据占用空间）。这是一个基本的重复数据消除概念。特定方法因系统供应商而异，但是任何重复数据消除方法都会占用 CPU 以计算数据段指纹或哈希密钥，并在对数据段索引表执行查找时执行 I/O 操作。

仅当数据从缓存转储到容量层时，vSAN 才会计算指纹并查找重复数据段。这意味着在正常操作下，对缓存 SSD 中的写缓冲区进行的 VM 写入不应产生任何延迟影响。

当数据从缓存转储到容量层时，会产生重复数据消除成本。它会占用为 vSAN 保留的一部分 CPU 能力，而索引查找生成的磁盘操作会占用一部分后端 I/O 能力。

由于资源消耗因 I/O 模式、数据类型等而异，因此请在确定对您的应用程序是否推荐使用重复数据消除之前咨询 Dell EMC 或 VMware 专家。

在《*Technical Whitepaper VMware VSAN 6.2 Space Efficiency Technologies*》（*VMware VSAN 6.2 空间效率技术的技术白皮书*）中可以找到更多信息，地址为

<http://www.vmware.com/files/pdf/products/vsan/vmware-vsan-62-space-efficiency-technologies.pdf>。

擦除编码

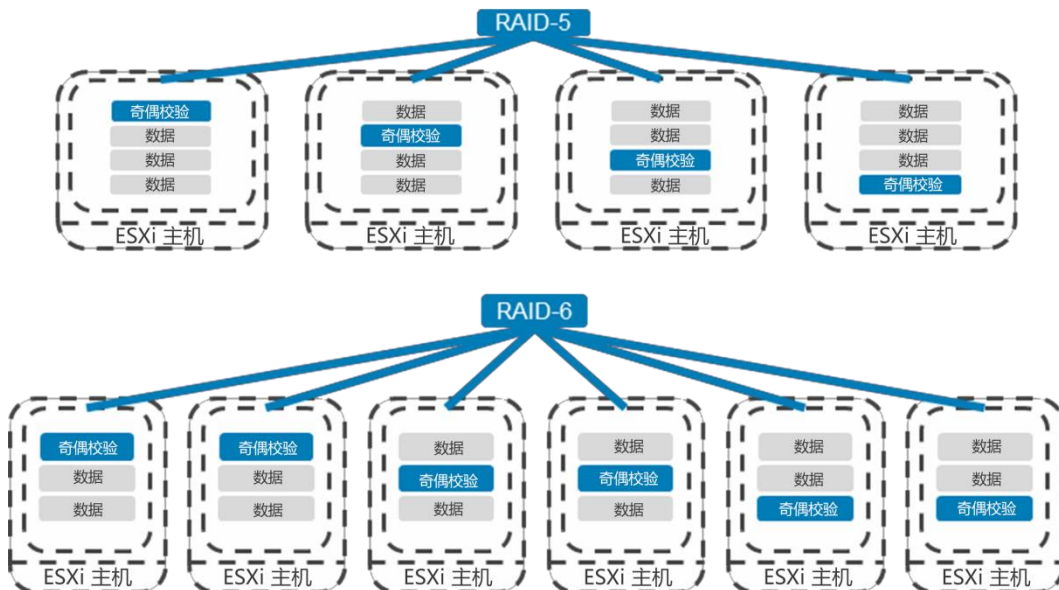
擦除编码是在全闪存 VxRail 配置中用于故障容错方法和数据保护的另一种高效容量解决方案。作为 RAID-1 镜像提供的数据复制的替代故障容错方法，擦除代码可以提供的可用容量比单纯的传统 RAID-1 镜像（这会耗尽存储空间）最高多 50%。

擦除编码将数据分解为片段并在系统中分布冗余数据区块。它使用数据块和分条来引入冗余。为从根本上进行说明，数据块按 n 个数据块进行分组，对于每组 n 个数据块，存在一组 p 个奇偶校验块。这些包含 $(n + p)$ 个块的组共同组成一个条带。关键在于， $(n + p)$ 条带中的任何 n 个块都足以恢复条带中的完整数据。

在 VxRail 群集中，属于单个条带的数据和奇偶校验块位于群集中的不同 ESXi 主机上，从而为每个条带提供了故障容错层。条带不遵循一对一分布模式。并非是以下情况：包含 n 个数据块的组位于一个主机上，而奇偶校验组位于另一个主机上。相反，算法会在 ESXi 主机间分布奇偶校验组中的各个块。

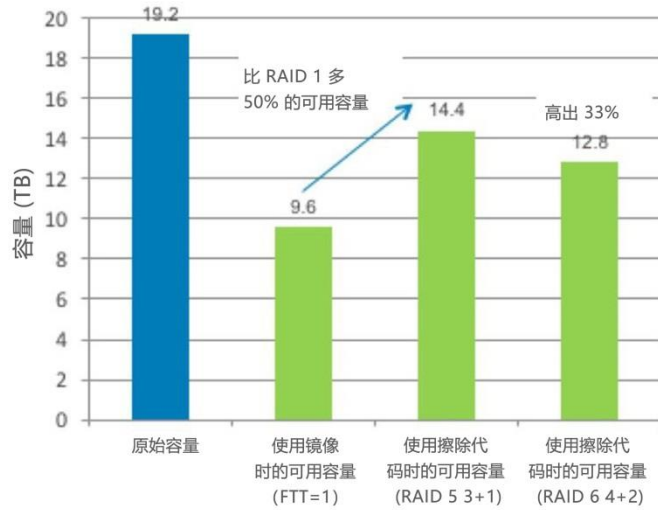
擦除编码提供可以承受一个故障 (FTT=1) 的单奇偶校验数据保护 (RAID-5)，以及可以承受两个故障 (FTT=2) 的双奇偶校验数据保护 (RAID-6)。下图展示了实施。单奇偶校验条带使用三个数据块和一个奇偶校验块 (3+1)，它需要至少四个主机或四个故障域，才能确保一个主机或磁盘出现故障时的可用性 (如下所示)。这表示可比 RAID-1 镜像节省 30% 的存储。双奇偶校验可比 RAID-1 节省多达 50% 的容量。它使用四个数据块及两个奇偶校验块 (4+2)，并且需要至少六个节点。请参见下图。

图 60 RAID-5 (FTT=1) 至少需要四个节点，RAID-6 (FTT=2) 有 4+2 个节点



看看下图中的可用容量比较。与镜像相比，擦除代码保护方法可将可用容量最多增加 50%。

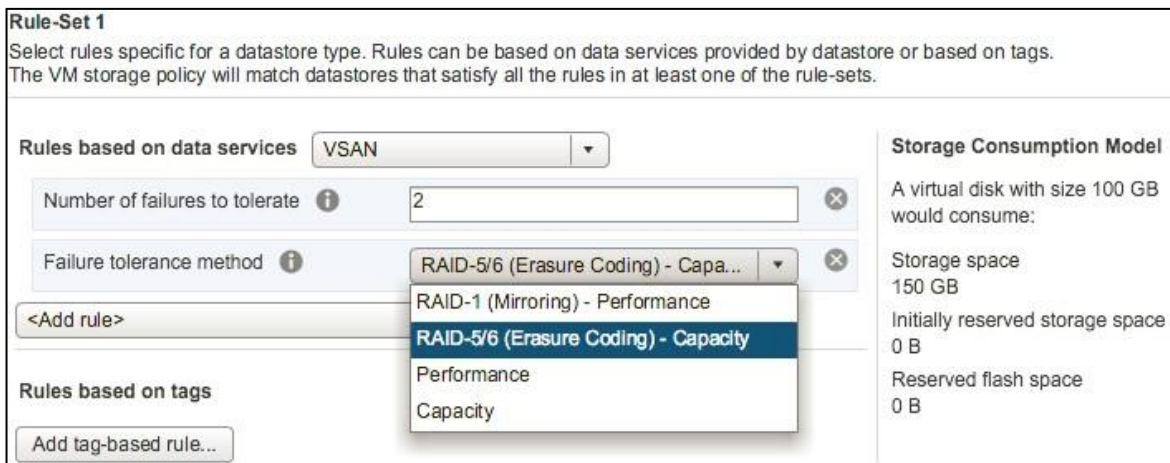
图 61 擦除编码将可用容量最多增加 50%



启用擦除编码

SPBM 策略故障容错方法 (FTM) 使管理员可以在 RAID-1 (镜像) 与 RAID-5/6 (擦除编码) 之间进行选择。FTT 策略确定擦除代码写入的奇偶校验块数。请参见下图。

图 62 FTT 策略确定擦除代码写入的奇偶校验块数



VxRail 以非常精细的级别实施擦除编码，可以应用于 VMDK，从而有助于实现精细方法。用于具有写入密集型工作负载（例如数据库日志）的配置可以包含镜像策略，而数据组件可以包含擦除编码。

擦除编码开销

擦除编码可节省空间，但会增加后端开销。计算奇偶校验块会占用 CPU 周期并向网络 and 磁盘增加开销，在多个主机间分布数据存储片也是如此。这种额外活动可能会影响延迟和整体 IOPS 吞吐量。

重建操作也增加了开销。一般情况下，重建操作会大大增加用于复制的读取和网络传输数。此处还提供了公式。如果 n 是指条带中的块数，则重建操作的成本是普通复制的 n 倍。对于 3+1 条带，这意味着对于每次传统数据复制，都会进行三次磁盘读取和三次网络传输。还可以调用重建操作以处理针对当前可用数据的读取请求。

此额外 I/O 是只有全闪存 VxRail 配置才使用擦除编码的主要原因。此处的基本原理是针对额外 I/O 的闪存磁盘补偿。

vSAN 加密

在 vSAN 6.6 中，加密是一个新的数据存储区级别设置。vSAN 加密是应用于群集中所有 VM 组件的数据存储区或群集范围级别设置。此功能可消除介质被盗的顾虑。群集范围设置的一大优势是在加密之前应用重复数据消除和压缩。与 vSphere 6.5 加密选项相比，这提供了节省空间的好处。与 vSphere 加密一样，符合 KMIP 规范的密钥管理服务器（如 CloudLink 或 Hytrust）必须与 vSAN 加密配合使用。vSAN 加密符合 FIPS 140-2 级别 1，AES 256 规范。

VxRail 集成软件

VxRail 按两种方式集成软件,这两种方法均由 Dell EMC 作为单一产品/服务进行全面设计、测试、验证、制造和支持。

VSphere 本机产品,包括 vSphere Replication

Dell EMC 软件产品,包括 RecoverPoint for Virtual Machines

虚拟机复制

可使用多个选项来复制虚拟机,以便在 VxRail 群集中实现数据保护和灾难恢复。其中包括集成到 VMware 软件堆栈的解决方案: VMware vSphere Replication (VR) 和 RecoverPoint for Virtual Machines (RP4VM), 基于经验证的企业 Dell EMC RecoverPoint 技术而构建。

选择合适的解决方案时,恢复点目标 (RPO) 是重要的考虑因素。RPO 针对由于数据丢失问题而从复制的拷贝恢复的数据,定义这类数据的可接受最长期限。例如,如果删除某个虚拟机并且该虚拟机的 RPO 是 24 小时,则虚拟机的恢复的拷贝应包含除了过去 24 小时内发生的任何更改以外的所有数据。根据技术,可以将 RPO 配置为低至几秒或几分钟。

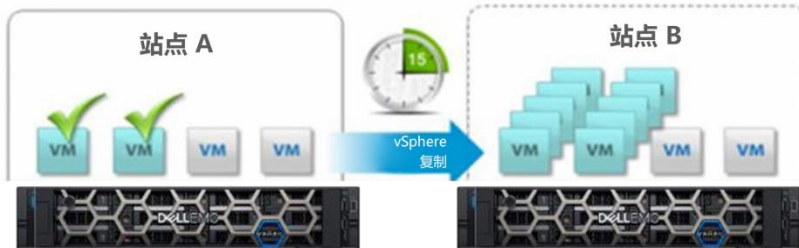
另一个考虑因素是要保留的恢复点数。为虚拟机配置复制时,管理员可以选择启用多个恢复点(时间点实例)的保留。当问题在发生后的几小时甚至是几天后才发现时,这可能会十分有用。例如,具有四小时 RPO 的复制的虚拟机感染了病毒,但是直到感染后的六个小时后才发现病毒。因此,病毒已复制到目标位置。使用多个恢复点时,可以恢复虚拟机,然后恢复到病毒问题发生之前保留的恢复点。

VMware vSphere Replication

VMware vSphere Replication 是与 VMware vCenter Server 完全集成的基于虚拟机管理程序的异步复制解决方案。它使用由 VMware 开发的专有复制引擎,包含在 VMware vSphere Essentials Plus 工具包和更高版本的 VMware vSphere 中。虽然 VR 十分适用于 VxRail 超融合环境中的 vSAN 存储,不过它完全独立于底层存储,可实现异构存储类型之间的复制。当 VxRail 系统是包括 SAN 或其他存储类型的较大虚拟化环境的一部分时,这非常有用。VR 在相同 vCenter 环境以本地方式提供数据保护以及灾难恢复并避免恢复到另一个 vCenter 站点。它还支持复制到服务提供商云(如 vCloud Air)。

下图显示两个 VxRail 站点之间的 VR 复制。

图 63 VMware vSphere Replication 及 VxRail



在此示例中，多个 VM 复制到一个 RPO 为 15 分钟的远程站点。该远程站点维护多个 VM 映像，从而允许回滚到不同时间点。

传输复制的数据的 VR 组件内置在 vSphere 中，并使用安全套接字层 (SSL) 连接。最佳做法是隔离网络流量以提高性能和安全性。VR 还包括一个或多个预先构建的基于 Linux 的 vSphere Replication vApp。第一个虚拟设备称为 vSphere Replication 管理服务器。它接收复制的数据、管理身份验证并维护源虚拟机与目标位置上的复制副本之间的映射。每个设备都需要 18 GB 的 VMDK 存储、4 GB 的内存以及两个或四个虚拟 CPU。

VR 配置包括指定处于 15 分钟到 24 小时范围内的恢复点目标 (RPO) 以及要维护的时间点时间范围数。所有这些操作都在 vSphere Web Client 中使用简单向导执行。为虚拟机配置了复制之后，vSphere Replication 会开始源虚拟机到目标位置的初始完全同步。完成此初始同步所需的时间可能会因必须复制的数据量和可用网络带宽量而异。进行初始完全同步之后，在每个复制周期中只传输更改的数据，从而最大限度降低网络带宽要求。

复制的数据首先写入重做日志。收到当前复制周期内的所有更改并写入重做日志之后，重做日志中的数据会合并到基本映像中。此过程可确保始终有一致且可恢复的 VM 映像可用。

vSphere Replication 可提供灵活、可靠且经济高效的复制来实现数据保护和灾难恢复。它在 VMware 产品堆栈中无缝集成，以进行简单的部署、配置和管理。

Dell EMC RecoverPoint for Virtual Machines

RecoverPoint for Virtual Machines (RP4VM) 基于 RecoverPoint Continuous Data Protection 技术，该技术十多年来已在企业环境中经过了验证，运行时间超过 3.5 亿小时，并保护整整 1 EB 的数据。与 VR 一样，RecoverPoint for VMs 使客户可以简单地复制虚拟机，并且可使用 RecoverPoint 插件从 vSphere Web Client 内部进行配置和管理。

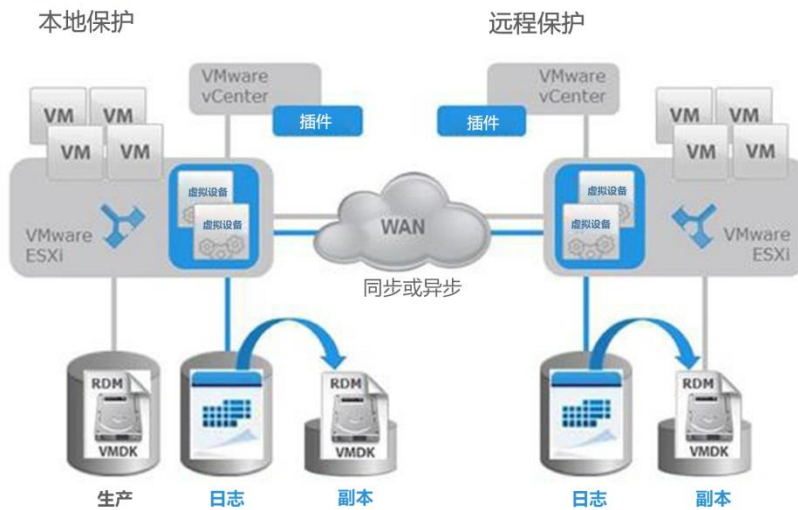
虽然在概念上与 VR 类似，不过 RP4VM 具有多种独特功能，包括能够将 RPO 为低至零秒以及并发本地和远程复制与多站点 2:1 和 1:2 配置。

RecoverPoint for VM 包含用于管理数据复制的所有方面的 RecoverPoint 虚拟设备 (vRPA)。二到八个 vRPA 分组为一个群集。对于本地复制，会配置单个群集。对于远程复制，会在本地和远程站点上配置群集。还会在每台 ESXi 服务器上安装 RecoverPoint 拆分器。拆分器会获取每个写入并将它同时发送到 VMDK 和 vRPA。RP4VM 的管理全都使用 RecoverPoint for VMs 从 vCenter 内部执行。

复制的源是虚拟机和关联应用程序数据，称为生产拷贝。RP4VM 会执行写入到生产拷贝的连续数据复制。每个写入都会进行拆分并同时发送到 VMDK 和日志。日志可提供类似于 DVR 的回滚，用于创建反映任何时间点的本地拷贝。本地拷贝可以用于从逻辑错误和数据损坏中恢复。（可选）可以将 RP4VM 配置为同步或异步复制到远程 vRPA 群集。通过 WAN 连接进行的本地和远程 vRPA 群集通信以及压缩和重复数据消除可提供优化，并减少带宽占用。

下图显示一个同时具有本地和远程数据保护的环境，以及写入如何进行拆分、日志记录并用于创建时间点映像。

图 64 连续本地和远程保护



其他 RP4VM 功能包括灾难恢复工作流的自动发现、调配和编排（包括使用一致性组对单个 VM 或多个 VM 执行测试、故障切换和回切）。

每个 VxRail 系统附带一个许可证入门产品包。可以根据需要购买其他许可证。

VxRail 复制使用情形

VR 和 RPVM 都提供数据保护和灾难恢复功能。两者都是纯软件解决方案，嵌入在虚拟机管理程序中，使用 vApp 进行控制并通过 vCenter 进行管理。使用哪种解决方案取决于使用情形。

可以配置到其他系统的本地复制以及到其他 VxRail 群集或非 VxRail vSphere 群集的远程复制。远程复制可以用于向 ROBO 站点提供灾难恢复。其他使用情形包括从其他 vCenter 环境到 VxRail 环境中以及 VxRail 环境之间的 VM 迁移。

对于使用 VMware vCenter 的完全集成且易于管理的基本复制，VR 可提供这些功能而无需额外成本。如果 RPO 少于五分钟，则只有 RP4VM 才能通过类似于 DVR 的播放方式提供连续数据复制。RPVM 还支持更灵活的部署选项，包括并发本地和远程复制以及采用 2:1 或 1:2 配置的多个站点。

无论选择何种复制，解决方案调整都应包括所需额外存储容量和复制 vApp 的开销。

对外部网络存储的支持

vSAN 向 VxRail 群集中的所有节点提供强大、安全且高效的共享数据存储区。外部 SAN 存储通常不是 VxRail 环境的一部分。但是，通常需要访问外部存储，以便将虚拟机和数据移动到 VxRail 环境或在环境之间移动数据。支持光纤通道 SAN 连接，还支持基于 IP 的存储。一个重要区别是光纤通道、iSCSI 或 NFS 数据存储区中的数据是自包含的，不会分发到 VxRail 群集中的磁盘组。

光纤通道与 VxRail

客户可以随其 VxRail 一起订购光纤通道 (FC) 主机总线适配器 (HBA)，用于外部存储。光纤通道存储可配置为对本地 VxRail 群集存储进行补充。外部存储的常见使用情形是：客户希望继续使用其现有的存储阵列作为 VxRail 的辅助存储，或者客户正在寻找一种可将数据从其 FC 存储迁移到 VxRail vSAN 数据存储区的方法。VxRail 不会为 FC HBA 提供生命周期管理。客户将需要通过 vCenter Server 管理它。

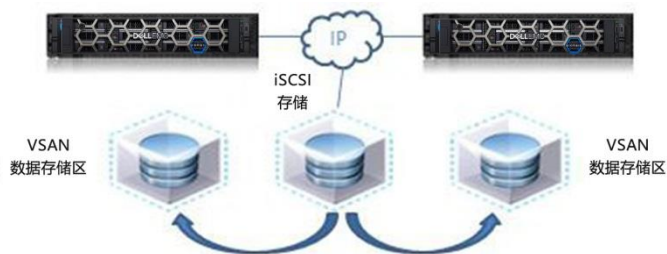
使用 FC HBA，客户可以连接到 HBA 卡支持的存储阵列，并通过 VMware 进行验证。但是，Dell EMC 将仅支持通过 HBA 连接到经 eLab 认证的 Dell EMC 存储阵列（即 Dell Compellent、Clariion/VNX/Unity、Symmetrix DMX、Symmetrix VMAX/PowerMAX 和 XtremIO）。

通过 FC HBA 配置外部存储时，允许客户安装 VM/VIB/驱动程序，以便根据需要使用外部存储开始运行。客户负责维护和更新它。如果 PCIe 总线上有插槽可用，则客户可安装多个 HBA。

iSCSI 与 VxRail

iSCSI 可以用于将 VM 和关联数据移动到 VxRail 环境以及在这些环境之间移动。下图显示包含 iSCSI 存储以及 vSAN 数据存储区的 VxRail 环境。

图 65 到 VxRail 环境中以及这些环境之间的数据移动性



iSCSI 存储上的数据可方便地移动到 VxRail vSAN 环境中或在 VxRail 环境之间移动。

现有 iSCSI 存储还可以用于向 VxRail 环境提供更多容量。但是借助 VxRail 纵向扩展和横向扩展配置灵活性，通常不使用外部存储来满足容量要求。

iSCSI 通过 IP 网络使用 SCSI 协议提供数据块级存储。SCSI 使用客户端-服务器、启动器-目标模型，其中启动器向目标设备发出读/写操作，而目标返回请求的读取数据或持久保存写入数据。VMware 环境中的 iSCSI 是标准功能。在 ESXi 主机上使用 NIC 的软件适配器配置为启动器，而外部存储系统上的目标向启动器提供 LUN。外部 LUN 可以由 ESXi 用作原始设备映射 (RDM) 设备，但是对于 VxRail，使用情形通常是将它们配置为 VMFS 数据存储区。有关详细信息，请参阅 vSphere 文档：《*Using ESXi with iSCSI SAN*》(将 ESXi 与 iSCSI SAN 结合使用)。

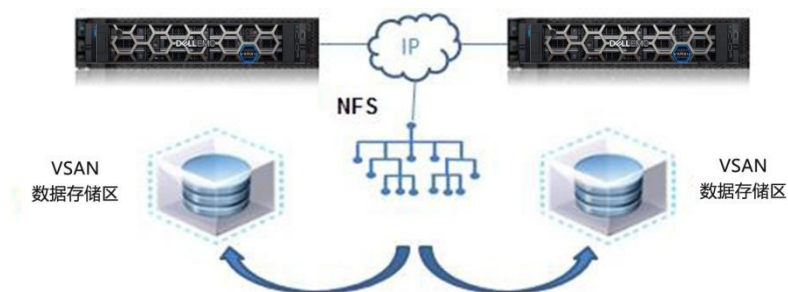
iSCSI 配置使用 vSphere Web Client 来执行。步骤涉及在 VDS 上创建端口组、创建 VMkernel 网络适配器并将它与端口组相关联以及分配 IP 地址。然后在“vCenter Manage Storage Adapters”视图中，使用“Add iSCSI Software Adapter”对话框创建软件适配器。最后一步是将 iSCSI 软件适配器与 VMkernel 适配器绑定。此操作完成之后，可以发现 iSCSI 目标和 LUN 并使用它们创建新的数据存储区以及将它们映射到群集中的主机。(有关详细信息，请参阅 VMware 文档。)

iSCSI 最适合于可提供一致且可预测的性能的网络环境，并且通常实施单独的 VLAN。针对 VxRail 环境计划网络要求时应考虑 iSCSI 网络要求，以确保存在与外部 iSCSI 存储系统之间的连接，并且额外网络流量不会影响其他应用程序。

NFS 与 VxRail

NFS 是一种网络文件系统，可通过 IP 网络使用 NFS 协议提供文件级存储。它可以处理类似于 iSCSI 的使用情形区别在于 NFS 设备呈现为文件系统，而不是数据块设备。下图显示一个从网络连接服务器导出并由 ESXi 节点装载在 VxRail 环境中的 NFS 文件系统。

图 66 网络连接文件系统及 VxRail



这可实现到 VxRail 环境中以及这些环境之间的数据移动性，以及实现额外存储容量。

外部 NFS 服务器可以是开放系统主机（通常是 Unix 或 Linux）或是专门构建的系统。NFS 服务器层采用物理存储并创建文件系统。文件系统会导出，而客户端系统（在此示例中为 VxRail 系统中的 ESXi 主机）会装载文件系统并通过 IP 网络访问它。

与 iSCSI 类似，NFS 是标准 vSphere 功能，使用 vCenter Web Client 进行配置。这在“Hosts and Clusters”视图中的“Related Objects and the New Datastore”对话框下进行。选择 NFS 作为数据存储区类型，并选择 NFS 版本、数据存储区的名称、导出文件系统的 NFS 服务器的 IP 地址或主机名以及会装载它的主机。NFS 文件系统会类似于 vSAN 数据存储区一样显示。VM、模板、OVA 文件和其他存储对象可以使用 vMotion 轻松地在 NFS 文件系统与 vSAN 数据存储区之间移动。

与 iSCSI 一样，NFS 最适合于可提供一致且可预测的性能的网络环境。在最初针对 VxRail 环境计划网络要求时，应考虑 NFS 的网络要求。

VxRail 解决方案和生态系统

Dell EMC 提供了一系列灵活消费模式,使企业能够更快、更轻松地使用 VxRail 来推动数字转型。这些使用模式包括技术本身以及企业如何为此技术付费。

VMware 验证设计与 VxRail

VMware 验证设计 (VVD) 是适合数据中心设计的一系列解决方案,跨越计算、存储、网络和管理,用作软件定义的数据中心 (SDDC) 实施的蓝图。VVD 是有关如何部署、操作和维护 VMware SDDC 的参考体系结构。VxRail 支持从 VVD 4.2 开始。

VMware VVD 在 VxRail 之上为完整的 NSX 和 vRealize 功能提供一个框架。它需要通过互操作性和可扩展性测试,对硬件和软件进行端到端验证。此外,它还提供有关如何监视、备份、恢复和故障切换管理组件的 Day 2 指导。因此,它创建一个值得信赖的实施设计,这种设计可消除部署风险、简化运营并进一步推动 IT 敏捷性,让客户能够创建私有云并加快其向多云 VMware 环境的转型。下图显示了现代化 vSphere 和 VMware 环境的一些不同部署模型,其中包括使用 VxRail 的 VVD。

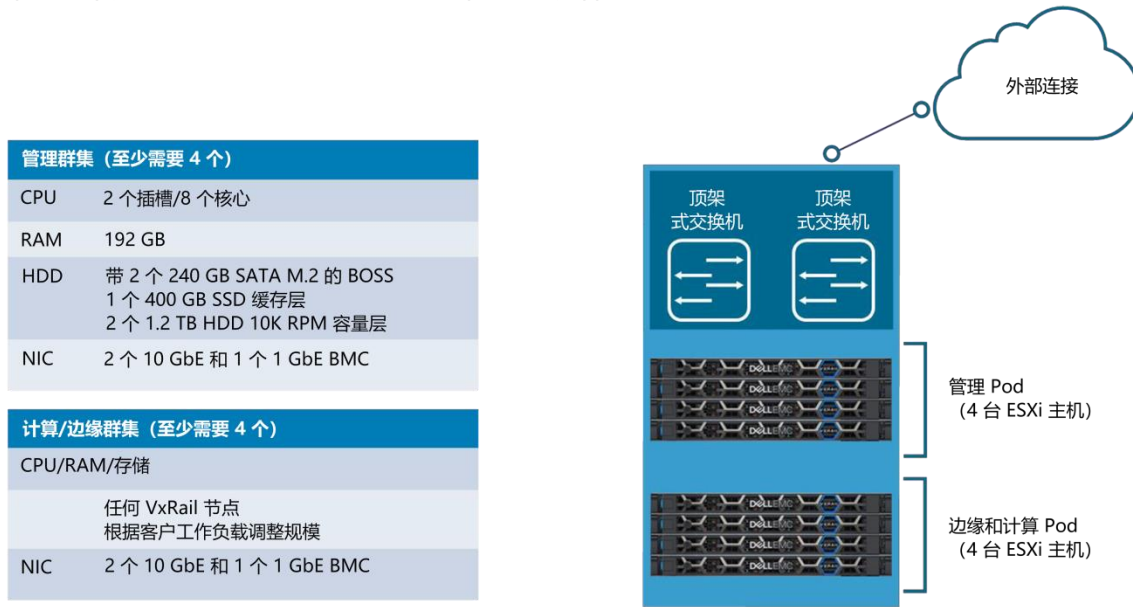
图 67 VMware SDDC 和多云 IT 模式的路径

加快 IT 向 VMware 多云的转型



VVD 需要一个包括四个节点的管理群集和一个至少有 4 个节点的计算/边缘群集。计算/边缘群集可以是任何 VxRail 节点,只要遵循本白皮书的 [VxRail 扩展部分](#)所述的群集设计规则。管理群集有一组规定的最小资源。下图详述了这两个群集的要求。

图 68 采用 VxRail 的 VVD 的配置和高层次体系结构



利用满足甚至超出这些要求和客户需求的 VxRail 型号产品

VVD 提供有关 VMware SDDC 和软件物料清单的指导，如下图所示。VMware 建议为 VxRail 客户提供 VMware Cloud Foundation (VCF) 升级许可证，这样不对已经付费的权限和软件进行重复收费。以捆绑包形式购买 VCF 许可证可降低总体许可成本，并且还将提供未来在 VxRail 上更广泛地利用 VCF 和 SDDC Manager 自动化的路径。（注意：目前在 VxRail 上不支持 SDDC Manager。用于 VxRail 的 VCF 附加组件中包括的 NSX 和 vRealize 许可证必须作为独立产品运行，而不是作为 SDDC Manager 的一部分。）

图 69 VVD 4.2 中的示例软件物料清单

| 云组件 | 产品项目 | 版本 |
|--------|---------------------------------------|---------|
| 虚拟基础架构 | ESXi | 6.5 u1 |
| | vCenter Server 一体机 (VIMISO) | 6.5 u1e |
| | NSX for vSphere | 6.3.4 |
| 云管理 | vRealize Automation 一体机 | 7.3.0 |
| | vRealize Orchestrator | 7.3.0 |
| | 适用于 NSX 的 vRealize Orchestrator 插件 | 1.0.4 |
| | vRealize Business | 7.3.1 |
| 服务管理 | vRealize Operations Manager 一体机 | 6.6.1 |
| | 适用于 vSphere 的 NSX 管理包 | 3.5.1 |
| | 适用于 vRealize Log Insight 的管理包 | 6.0 |
| | 适用于 vRealize Automation 的 Manager 管理包 | 3.0 |
| | 适用于存储设备的管理包 | 6.0.5 |
| | vRealize Log Insight | 4.5.0 |
| 基础架构 | Windows | 2012 R2 |
| | SQL Server | 2012 R2 |

有关更多信息，请访问：

VMware VVD 站点：

<https://www.vmware.com/solutions/software-defined-datacenter/validated-designs.html>

Dell EMC 社区网络，VxRail 上的 VMware 验证设计 4.2 部署指南：

<https://community.emc.com/docs/DOC-66332>

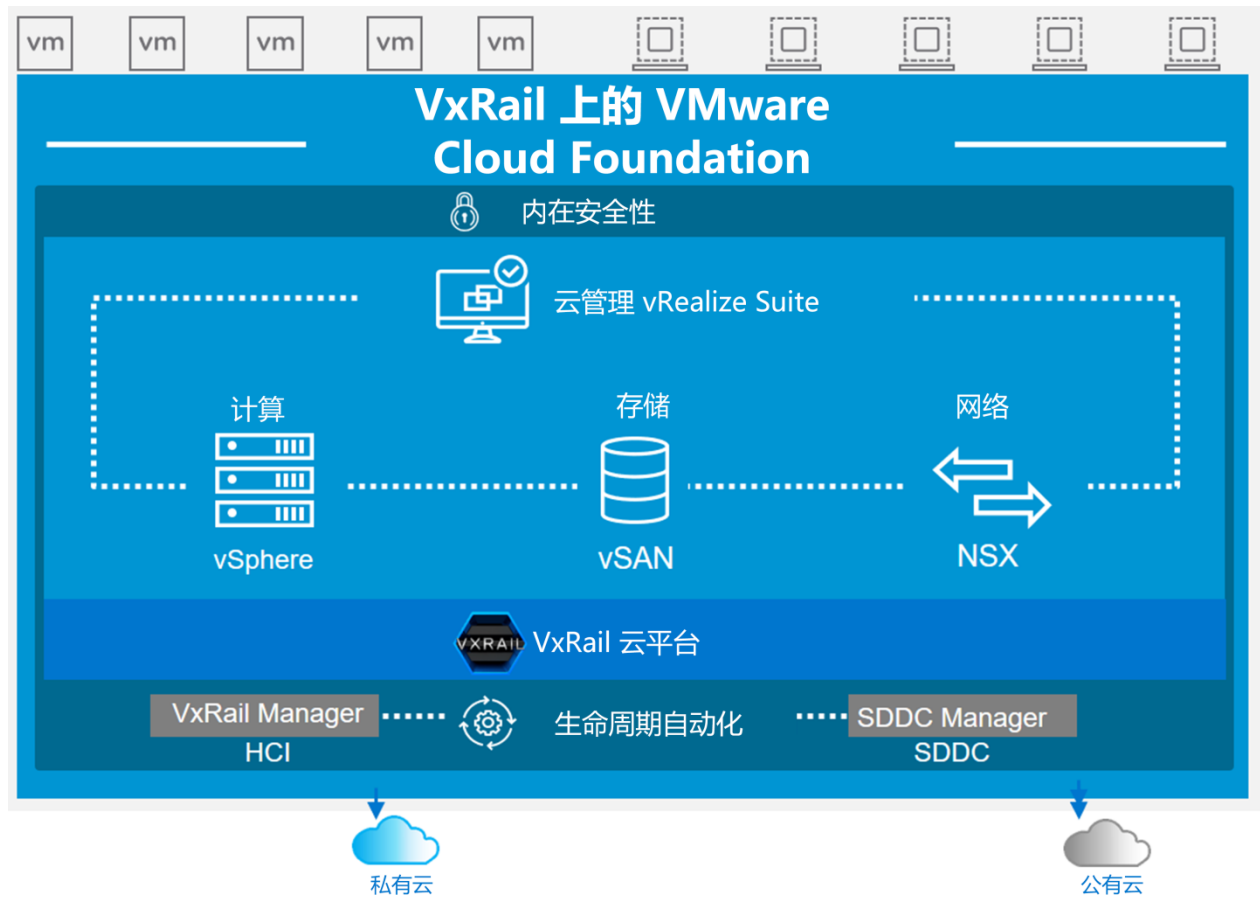
VMware Cloud Foundation on VxRail

VMware Cloud Foundation on VxRail 是 Dell EMC 和 VMware 联合设计的集成式解决方案，具有从第 0 天到第 2 天简化、精简和自动化整个 SDDC 运营的功能。新平台在私有云和公有云环境中为计算（包括 vSphere 和 vCenter）、存储（vSAN）、网络和安全（NSX）以及云管理（vRealize Suite）提供了一组软件定义的服务，使其成为混合云的运营中心。

VMware Cloud Foundation on VxRail 通过完全集成的混合云平台提供到混合云的简单路径，该平台利用本机 VxRail 硬件和软件功能以及其他 VxRail 独特集成（例如 vCenter 插件和 Dell EMC 网络）协同工作，提供全新的全包式混合云用户体验和全栈集成。全栈集成意味着您可以通过完全自动化的生命周期全包式体验同时获得 HCI 基础架构层和云软件堆栈。

该产品/服务的一个重要方面是引入了用于部署这些 SDDC 组件的标准化体系结构，另外还引入了 Cloud Foundation（基于 VVD 的集成式云软件平台）。我们承诺，通过将标准化设计纳入平台，确保这些组件已相互认证并得到 Dell Technologies 的支持。然后，您可以确保有一个自动化和经过验证的路径，以便在端到端堆栈中从一个已知良好状态转换到下一个已知良好状态。

图 70 VMware Cloud Foundation on VxRail 的体系结构



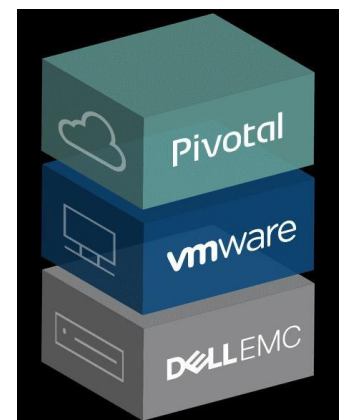
Pivotal Ready Architecture (PRA)

Pivotal Ready Architecture 是一个经过测试和验证的参考体系结构，用于在 VxRail 上部署 Pivotal Cloud Foundry。Pivotal Ready Architecture 包括适用于高可用性、全面产品支持和对象存储选项的配置，是在本地部署 Pivotal Cloud Foundry 的理想方式。

云原生模式是一种现代化的应用程序体系结构、开发和交付方法，作为对业务需求和基础架构功能变化的自然响应而出现。这种新模式直接提高了 IT 组织的应用程序交付速度和灵活性，并证明了它对初创企业和成熟企业都大有裨益。Pivotal Ready Architecture 是在数据中心启动和运行 Pivotal Cloud Foundry 的捷径。利用“轻松运行”体验加快您的转型。PRA 支持 Pivotal Application Service (PAS) 和 Pivotal Container Service (PKS)。

从 PRA 获得的业务优势包括：

- 可靠的部署。PRA 是一个经验证的硬件和软件解决方案。



- 就绪基础架构。PRA 基于市场上少有的全面集成、预先配置且经过预先测试的 VMware 超融合基础架构系统系列。
- 弹性体系结构。PRA 提供多站点、多基础和多个可用性分区配置选项，可提供更长的正常运行时间、更大的地理覆盖范围和更高限度的弹性。

PRA 提供一种经过测试、验证的参考体系结构，可基于它构建高可用企业级开发人员平台。PRA 基于超融合 VxRail 构建而成，可提供基础架构的自动化生命周期管理，这是加速向数字化业务转型的关键要素。

- VxRail 上的 Pivotal Application Service (PAS) 和 Pivotal Container Service (PKS) 参考体系结构
- 完全由软件定义的基础架构
- “永不停机”高可用配置
- 中央管理控制台
- 可随您的增长而扩展的模块化设计
- 集成式备份和灾难恢复选项

有关详细信息，请访问：<https://pivotal.io/pivotal-ready-architecture>

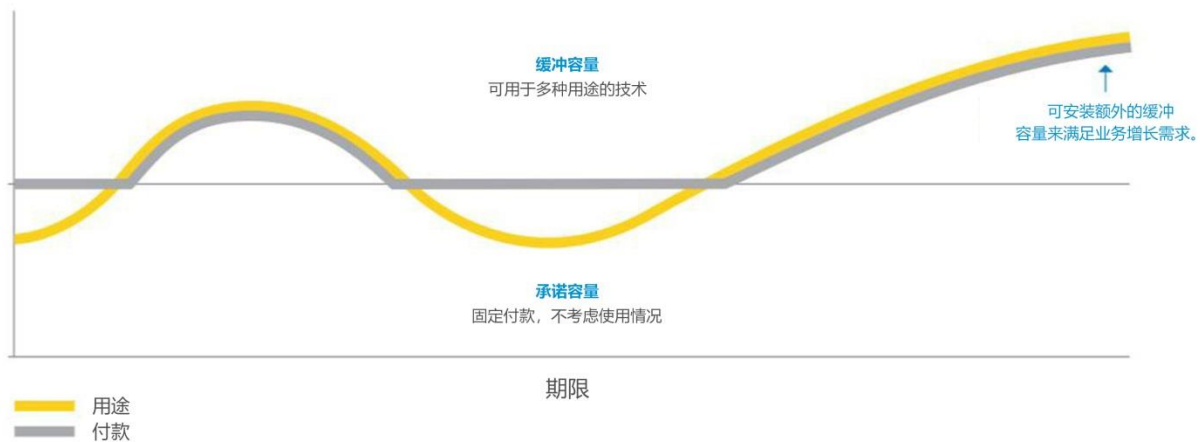
灵活消费选项

Flex on Demand — 类云消费选项

借助由 Dell Financial Services (DFS) 提供的 Flex On Demand，您可以获得所需的技术以支持不断变化的业务需求，而付款额依据实际使用情况而定。此模式可帮助您根据使用情况调整成本，避免为未使用的缓冲容量付费。在业务运营高峰期间根据使用情况提供即时容量部署，从而提高敏捷性。通过提供更好的运营经济性，它可以提高预算敏捷性和能力。

- DFS 与您一起确定您当前所需的“承诺容量”以及未来将需要的“缓冲容量”。
- 使用随您的设备安装的自动化工具衡量缓冲容量。每次付款包括两部分：固定的承诺容量金额和可变的缓冲容量金额。
- 如果您的使用量持续消耗大部分缓冲容量，则您可以选择获得额外的缓冲容量。安装后，您的已提交容量级别和相关付款将会增加。

图 71 技术使用与 Flex on Demand 付款之间的关系



有关更多信息，请参阅 <https://www.dellemc.com/zh-cn/flexibleconsumption/cloud-flex-for-hci.htm>。

VDI 完整解决方案

VDI 完整解决方案是一系列端到端桌面和应用程序虚拟化解决方案，具备卓越的解决方案堆栈和非常优化的总拥有成本。这些解决方案基于 Dell EMC VxRail 系统，并利用 VMware Horizon 虚拟桌面和应用程序。

VDI 完整解决方案是 Dell EMC 提供的端到端桌面和应用程序虚拟化解决方案，其中包括您开始工作所需的一切组件：基础架构系统、软件、存储和端点。

VDI 完整解决方案使用 Dell Technologies 的卓越技术构建而成。它利用 Dell 和 Dell EMC 经过验证且值得信赖的基础架构系统和端点。此外，它还充分利用了 VMware Horizon（行业引领者的台式机 and 应用程序虚拟化）。VDI 完整解决方案经过全面验证和测试，可降低风险和复杂性。这是购买和完整解决方案支持的单一来源。

有关更多信息，请参阅 [https://downloads.dell.com/solutions/general-solution-resources/White%20Papers/DellEMC.VxRail\(14G\).VMware.Horizon.RA.pdf](https://downloads.dell.com/solutions/general-solution-resources/White%20Papers/DellEMC.VxRail(14G).VMware.Horizon.RA.pdf)。

VMware Horizon

VMware Horizon 是 VMware 的 VDI 和桌面管理环境。Horizon 使用灵活且安全的交付模式来调配用户桌面。借助数据中心的安全性和弹性，用户可从几乎任何设备（包括移动设备）访问桌面环境。由于应用程序数据和软件组件驻留在数据中心内，因此可以应用传统的安全、备份和灾难恢复方法。



95%

初始化第 1 台 VxRail 一体机的速度加快 95%



47%

调配 10 台桌面虚拟机的速度提高 47%



96%

添加新一一体机的速度加快 96%

如果用户的设备丢失或硬件出现故障，则恢复简单明了。用户只需使用另一个设备登录即可恢复环境。在没有数据保存在用户设备上的情况下，如果设备丢失或被盗，则可能检索和泄露关键数据的可能性要小得多。

下图显示了 Horizon View 如何将操作系统、应用程序、配置文件和用户数据封装到隔离的层中，以及如何动态地按需汇集桌面以便向用户提供其各自环境的个性化视图。

图 72 高度可用且安全的桌面



可用性和安全性以及易于进行管理和支持是从传统物理台式机和笔记本电脑转到 VDI 的令人信服的原因。

VMware Horizon 是在 vSphere 环境中运行的全面桌面管理环境。环境管理通过 vCenter 集中化管理进行管理，可以利用各种高级功能（包括快照、vMotion、DRS 和 vSAN 存储）。

用户的桌面环境作为 View 桌面 VM 在 ESXi 服务器上运行，并通过使用远程桌面协议 (RDP) 或 PC over IP (PCoIP) 协议的 View 客户端进行访问。View 客户端可以是使用 View 门户在物理台式机、笔记本电脑、移动设备或 Web 浏览器中运行的应用程序。用户的桌面环境可以是专用 VM 或浮动 VM（在用户登录时从池中分配的 VM）。通过使用可选的 View Composer（而不是完整映像），链接克隆可以减少所需磁盘空间。Horizon View 包括用于管理连接、调配环境、对用户、其他应用程序和服务进行身份验证的附加组件。

VMware Horizon 及 VxRail

VxRail 系统是一种包括计算、存储和 vSphere 虚拟化和管理的功能完备的环境，非常适合于 VMware Horizon。VxRail 可加快 Horizon 基础架构部署，环境可以在数小时内（而不是数天）建立好进行运行。

VxRail 超融合基础架构采用支持数百到数千个虚拟桌面的配置。支持的桌面数取决于用户工作负载特征。

Dell EMC 开发了相应工具，通过它们能够对 VDI 环境数和预期工作负载特征进行建模，以确定可满足即时和长期要求的合适配置。随着需求的增加，VxRail 可通过添加其他系统和节点无中断地纵向扩展，同时向用户提供预期性能和一致的用户体验。

在 VxRail 系统上部署 Horizon 时，有两种常规方法：将 VxRail 环境专用于 VDI 或将 VDI 与其他工作负载混合使用。专门为与 VxRail 结合使用提供了 Horizon 版本或 Horizon Add-on 版本。VMware 或 Dell EMC 销售代表可以提供有关最佳客户特定选项的详细信息。

总之，组织通过 VxRail 与 VMware Horizon 可以快速实施桌面即服务 (DaaS) 并克服传统的桌面虚拟化资本支出 (CAPEX) 障碍。环境可以从小规模起步，并根据需要轻松地纵向扩展。这降低了初始启动投资。VxRail 超融合基础架构不仅可快速设置，而且其集成计算、存储、虚拟化和单一供应商支持模式可消除传统基础架构的复杂性。

VMware vSphere Platinum

VMware vSphere® Platinum 是一个专门构建的安全解决方案，用于保护企业应用程序、基础架构、数据和访问。它结合了两个经验证的产品：vSphere (业界卓越、高效、安全的混合云平台，适用于所有工作负载) 和 VMware AppDefense™ (由机器学习和嵌入威胁检测提供支持的数据中心端点安全，在虚拟化层中响应，以降低安全风险)。vSphere Platinum 不仅操作简单，还可确保应用程序和虚拟机在其已知良好状态下运行，并且开销和性能的影响极小。有关 VMware vSphere Platinum 的详细信息，请参阅 <https://www.vmware.com/content/dam/digitalmarketing/vmware/en/pdf/vsphere/vmw-vsphere-platinum-solution-brief.pdf>

VMware vSphere Platinum 可在 VxRail 上运行，从而在安全的现代化基础架构上构建总体应用程序防御。使用 Dell EMC VxRail 作为 VMware vSphere Platinum 的平台，确保将经过优化的网络弹性和安全性融入每层。

下图总结了在 VxRail 上运行 VMware vSphere Platinum 的好处。

图 73 VxRail 上的 VMware Platinum



IsilonSD Edge

EMC IsilonSD 产品系列结合了 Isilon 横向扩展 NAS 的强大功能与软件定义的存储的经济性。IsilonSD Edge 是专门构建的软件，可满足企业边缘位置的需求，存储不断增加的非结构化数据。借助 IsilonSD Edge，您可在 VMware 环境中快速部署简单高效的横向扩展 NAS 解决方案。通过以经济高效的方式在虚拟化基础架构中进行容量较小的部署，它还将数据湖范围从核心数据中心扩展到边缘位置。通过数据湖，您可以提高存储利用率、消除存储孤岛并降低 TCO。

在 VxRail 上运行 IsilonSD Edge 提供了一个简单、敏捷、经济高效的平台，可从 VxRail 群集内提供文件服务。它可以使用标准的 VMware 工具轻松地进行管理。此外，客户还可以使用它从远程位置整合数据以及将数据分发到远程位置。最重要的是，只需几分钟即可完成部署。IsilonSD Edge 包括所有 Isilon 的 OneFS 数据服务和协议，包括 SMB、NFS、HDFS 和 OpenStack Swift。IsilonSD Edge 与 VMware 环境紧密集成，并在 VMware ESXi 6.0（采用 VxRail 4.0 软件）或 VMware ESXi 6.5（采用 VxRail 4.5 软件）上运行。它利用 vCenter® 和作为虚拟映像运行的管理服务器，可用于安装许可证，或者在群集中添加和删除节点。IsilonSD Edge 管理服务器还会将一个插件安装到可用于配置群集的 vCenter。VxRail 上的服务器和存储资源不必是 IsilonSD Edge 专用的。如果满足了性能和容量需求，则可以在 VxRail 群集中并发运行其他应用程序。IsilonSD Edge 与 VxRail 上的 vSAN 是经过验证和测试的解决方案，它使用 VMFS 或 vSAN 数据存储区。

下图总结了在 VxRail 上运行 IsilonSD Edge 的好处。

图 74 在 VxRail 上运行 IsilonSD Edge 的好处

为 VxRail 提供企业级文件服务



| | |
|------------------|---|
| 扩展至 36 TB | ● |
| 利用 VxRail 实现快速部署 | ● |
| 消除数据中心存储孤岛 | ● |
| 加强数据保护 | ● |
| 简化管理 | ● |

VxRail 群集上的每个 IsilonSD Edge 实例可扩展至 36 TB，这足以处理许多远程和分支办公室的需求。您不必将 VxRail 环境专用于 IsilonSD Edge 群集 — 您可以将其与 VxRail 群集中的其他工作负载一起运行。

SAP HANA 认证及 VxRail

VxRail 是第一批 HCI 平台之一，并且是率先基于 VMware 且已获得认证可以运行 SAP HANA (SAP 的内存中数据库管理系统) 的 HCI。SAP 将利用 VxRail 的持久内存来支持应用程序及其使用情形。

客户将受益于在 VxRail 上运行 SAP HANA，因为系统可自动让实施更快地启动和运行，能够灵活地提供组件的适当组合，以便从第一天起就支持应用程序，并可扩展以确保满足未来的需求。使用 VxRail P 系列节点，通过自动化和完整的生命周期管理快速启动，从而快速有效地支持您的 HANA 实施。VxRail 作为 Dell EMC 就绪型解决方案的一部分经过全面认证，适用于 SAP 1.5 版本。

因为 VxRail 快速、灵活、强大且可扩展，所以它最适合 SAP HANA：

- 快速** — 自动化，易于部署/管理，确保您快速实现正常运行
- 灵活** — 借助 PowerEdge 上按订单构建的 VxRail，配置系统以满足特定需求
- 强大** — 丰富的组件相互搭配，为事务流程和分析提供性能、密度和电源效率
- 可扩展** — 无需“拆除更换”系统升级即可提高电源效率和性能

有关详细信息和解决方案指南，请访问：

<https://www.dell.com/zh-cn/solutions/business-applications/sap/hana/index.htm>

适用于 Splunk 的参考体系结构

Splunk Enterprise 是业界卓越的平台，旨在通过分析计算机生成的数据来获得宝贵的业务见解。Splunk Enterprise 使用其强大的 Splunk 搜索处理语言 (SPL™) 从计算机数据中提取有意义的信息。通过分析计算机数据而生成的见解称为运营智能，它有许多使用情形，包括：

- IT 运营** — 利用率、容量增长
- 安全** — 欺诈检测、实时检测威胁、取证
- 物联网 (IoT)** — 传感器数据、计算机到计算机、人工交互。

Dell EMC 和 Splunk 合作提供联合验证的参考体系结构，该体系结构经过优化，可实现极高的可扩展性和性能。在 Dell EMC 融合基础架构上运行的 Splunk 软件提供了推动组织的数字转型所需的运营智能。在结合使用时，Dell EMC 和 Splunk 将 Splunk 生态系统提供的运营智能与 Dell EMC 的经济高效、可扩展且灵活的基础架构结合起来。

Dell EMC 为 Splunk Enterprise 环境提供的主要优势包括：

- 经过优化的存储数据分层 — 使存储符合热/温、冷和冻结数据的要求并具有高保留期和性能。
- 经济高效和灵活的横向扩展 — 独立或作为单个融合平台提供横向扩展的容量和计算。
- 强大的数据服务 — 包括安全加密、压缩和重复数据消除以及用于保护的快速、高效的快照。

使用 Dell EMC VxRail 系统的参考体系结构与适用于虚拟化 Splunk Enterprise 环境的 Isilon™ 已由 Splunk 和 Dell EMC 联合测试和验证，以满足或超出在 Splunk 记录的参考硬件上运行的 Splunk Enterprise 的性能。VxRail 提供满足中小型企业 Splunk 部署的基础架构要求所需的性能和容量。

有关详细信息，请参阅

<https://China.emc.com/collateral/service-overviews/h15699-splunk-vxrail-sg.pdf>。

其他产品信息

有关文档、发行说明、软件更新或有关 Dell EMC 产品、许可和服务的信息，请访问 Dell EMC 在线支持网站（需要注册）：<https://support.emc.com>。

Dell EMC ProSupport for Enterprise

企业需要坚定不移的硬件和软件支持，以及管理数据中心多厂商环境的明智方法。Dell EMC 提供的统一资源具备可帮助您支持业务的专业技能、专门知识和能力。

ProSupport 在全球范围内全天候提供接受过完善培训的专业人员，帮助您解决 IT 需求、将干扰降至最低，同时维持高水平生产力。Dell EMC 在全球 165 个国家/地区拥有 55,000 多位以 55 种语言提供服务的 Dell EMC 和合作伙伴专业人员，您可以放心，在 Dell EMC 的帮助下，您将能够：

1. 利用 Dell EMC 的规模和技能，更大程度地提高生产效率
2. 训练有素的专家提供全天候服务以更大限度减少中断
3. 通过单一来源满足您的所有支持需求以提高效率

通过电话、聊天或即时消息为 VxRail 系统硬件和软件提供单一来源 24X7 全球支持。支持还包括对在线支持工具和文档的访问、快速现场部件交付和更换、对新软件版本的访问、与操作环境更新有关的帮助以及使用 Dell EMC Secure Remote Services (ESRS) 进行的远程监视、诊断和修复。

我们的 12 个卓越中心和联合解决方案中心利用 Dell EMC 与优秀的应用程序提供商（如 Oracle 和 Microsoft）的联盟，提供内部协作和行业出类拔萃的支持水平。我们的 87 个技术支持站点包括总共 71 个 Dell 技术支持站点和总计 16 个 EMC 客户服务中心。

Dell EMC 支持备受赞誉，获得 94% 的客户满意度评级，并已获得诸多奖项，包括 Temkin Group CE Excellence、TSIA STAR 奖项、Microsoft 年度部署合作伙伴等等。

Dell EMC 的不同之处是显而易见的，当涉及到您的 IT 战略时，我们让您可以无所畏惧地采用新技术，自由地重点关注您的业务。在整个基础架构中使用相同的 Dell EMC 企业级支持使您可实现这种自由度。

针对 VxRail 系统的 Dell EMC ProDeploy 服务

Dell EMC 提供 ProDeploy 安装和实施服务，以确保顺利、快速地将 VxRail 系统集成到客户网络中。针对单个系统而优化的标准服务可提供现场专家来与数据中心团队一起核对安装前清单、确认网络和架顶式 (TOR) 交换机设置、执行站点验证、机架和缆线、配置以及初始化系统。最后，现场 Dell EMC 服务技术人员会配置 EMC Secure Remote Services (ESRS)，并对基本 VxRail 系统管理任务进行简要功能概述。为较大规模的 VxRail 系统部署（包括具有多个系统或群集环境的部署）提供了此安装和实施服务的自定义版本。还提供 VxRail 系统扩展服务，该服务进行远程交付，并提供专家级服务技术人员以快速实施 VxRail 系统预加载数据服务（RecoverPoint for Virtual Machines 和 vSphere Data Protection）。