

VMware® View™ 的存储注意事项

最佳实践

目录

简介	3
Windows 磁盘 I/O 工作负载	3
基础映像注意事项	4
存储协议选项	6
存储技术选项	9
存储阵列决策	15
提高存储性能	16
总结	18
关于作者	18
参考资料	19

简介

VMware® View™ 是一个企业级虚拟桌面管理器，可将授权用户安全连接到集中式虚拟桌面。通过该管理器，您可以将虚拟桌面整合到数据中心服务器上，并独立管理操作系统、应用和数据，从而在提高业务敏捷性的同时使终端用户能够通过任何网络获得灵活的高性能桌面体验。

存储是 VMware View 实施的基础。本文档面向正在考虑实施 View 的架构师和系统管理员。文中列出了要了解自己环境的特征应采取的步骤，从确定您环境中的用户类型到选择存储协议和存储技术等诸多步骤。

本文档涵盖以下主题：

- Windows 磁盘 I/O 工作负载
- 基础映像注意事项
- 存储协议选项
- 存储技术选项
- 存储阵列决策
- 提高存储性能

Windows 磁盘 I/O 工作负载

要做出正确的存储大小规划和设计决策，您需要了解 Windows XP 和 Windows 7 的磁盘输入和输出 (I/O) 特征。为了便于描述特征，您可以将工作人员分为轻量级用户或重量级用户。轻量级用户在日常工作中一般使用电子邮件 (Outlook)、Excel、Word 和 Web 浏览器 (Internet Explorer 或 Firefox)。他们做的通常是数据录入或文书工作。重量级用户掌握全面的知识，他们不仅使用轻量级用户所用的各种工具 (Outlook、Excel、Word、Internet Explorer 和 Firefox)，而且还处理大型的 PowerPoint 演示文稿以及执行其他的大型文件操作。业务经理、管理层以及市场营销人员都属于此类用户。

“信息工作人员磁盘 I/O 吞吐量”改编自文档“VMware VDI 服务器调整和扩展”（请参见“资源”获取链接），它针对大量 VMware View（以前称为 VDI）桌面（每个 VMware ESX 主机上大约有 20 个桌面）对轻量级用户和重量级用户的磁盘使用情况做了比较。结果表明，平均有超过 90% 的信息工作人员的磁盘 I/O 中包含读取操作。

信息工作人员磁盘 I/O 吞吐量			
	读取磁盘 I/O 峰值	写入磁盘 I/O 峰值	总磁盘 I/O 峰值
轻量级用户	4.5 MB/秒	0.5 MB/秒	5.0 MB/秒
重量级用户	6.5 MB/秒	0.5 MB/秒	7.0 MB/秒

表 1：信息工作人员磁盘 I/O 吞吐量

为了明智地选择存储子系统，您需要将这些吞吐量的值转化为 SAN 和 NAS 存储行业所用的每秒 I/O 操作数 (IOPS) 的值。您可以利用以下公式将吞吐率转化为 IOPS：

$$\frac{\text{吞吐量 (MB/秒)} \times 1024 \text{ (KB/MB)}}{\text{数据块大小 (KB/IO)}} = \text{IOPS}$$

虽然 NTFS 文件系统分配的标准大小是 4 KB，但 Windows 可使用 64 KB 的数据块大小，Windows 7 可使用 1 MB 的数据块大小进行磁盘 I/O。

如果使用最坏情况（重量级用户）7.0 MB/秒的吞吐量以及较小的数据块大小 64 KB，则全组大约 20 个 Windows 虚拟机将产生 112 IOPS。

基础映像注意事项

大多数 View 实施都从少量的基础映像（黄金级映像）部署托管桌面。由于这些基础映像需要重复使用，因此在创建这些映像时务必要小心。您可以通过调整操作系统设置、应用程序集以及用户数据来优化映像大小。优化操作系统设置还可在提高整体系统性能方面带来更多优势。

操作系统设置

不管采用哪种底层存储技术，减少操作系统的占用空间都是 View 体系结构的最佳实践。操作系统经过虚拟化之后，传统桌面 PC 实施过程中用到的许多组件就不再需要了。您可以利用各种可用的实用程序去除不必要的功能，以此来简化基础操作系统。

尽管我们的主要目标是减小映像大小，但性能提高也是一个重要的附加效果。去除不必要的组件可以减少与虚拟桌面中不需要的流程和服务相关的开销。

应用集

在 View 部署过程中，应用的部署方式会直接影响最终桌面映像的大小。在传统的桌面环境中，应用程序可以直接安装在本地硬盘上，然后流式传输至桌面，也可以使用基于服务器的计算模型集中部署。

在 View 环境中，基本的部署方法保持不变，但 View 提供了新的应用管理机会。例如，多个 View 用户可以利用某类单个桌面映像，映像中包括基础操作系统以及必要的應用。管理员可以基于该单个黄金级映像创建桌面池。用户登录时，View 将基于基础映像从池中分配一个新桌面，桌面中包含热修补程序、应用程序更新和新增功能。

管理单个基础模板而不是多个独立的桌面映像（每个映像都有自己的应用集）可以降低应用部署的总体复杂性。您可以将 View 配置为通过各种方式克隆此基础模板（例如在虚拟磁盘级别、数据存储级别或卷级别），以使桌面能够满足各类用户的不同存储要求。

如果您使用基于存储或基于虚拟化的快照技术，则在大量虚拟机中增加本地安装的应用可能会导致存储需求有所提高，而性能有所降低。因此 VMware 建议您，仅对于基于自定义映像的单独桌面，才将应用程序直接部署到虚拟机上。

尽管使用基于服务器的传统计算模型来部署集中式应用的成功几率有限，但您可以利用这种方法减少与向虚拟化桌面提供应用相关的开销。

此外，基于服务器的计算还能提高每个运行 View 桌面的 ESX 主机上的虚拟桌面工作负载量，因为应用处理开销被卸载到了托管这些应用的服务器上。而且，您还可以通过基于服务器的计算实现灵活的负载管理、更轻松的复杂一线业务应用更新和全局应用访问。不管您的 View 部署是如何设计的，您都可以使用基于服务器的计算作为应用分发的整体方法。

用户数据

对于 View 桌面中用户特定数据的最佳实践是，将尽可能多的用户数据重定向到基于网络的文件共享。要最大程度地发挥共享存储技术的优势，请将单独的虚拟桌面视为一次性桌面。尽管您可以设置永久虚拟桌面（通常是一对一专用桌面），但 View 的某些重要优势恰恰在于能够轻松更新基础映像，并让 View 基础架构将这些更改作为一个全新的虚拟桌面来分发。您可以通过更新黄金级映像来对 View 桌面进行更新，View 将根据需要调配新桌面。

如果您使用的是永久桌面和池，则可以将用户数据存储在本地上。不过 VMware 建议您将数据存储到集中式文件存储中。如果您将桌面映像与数据分开存储，则可以更轻松更新桌面映像。

在 Windows 基础架构中，管理员经常会遇到与漫游配置文件相关的问题。不过，漫游配置文件经过正确的设计后可以变得很稳定，您可以在 View 环境中成功使用这些文件。成功使用漫游配置文件的关键在于要保持配置文件尽可能的小。通过使用文件夹重定向，尤其是通过超越在标准组策略对象中的默认值，您可以将漫游配置文件拆分到最小。

为减小配置文件大小，您可以重定向以下关键文件夹：

- 应用数据
- 我的文档
- 我的图片
- 我的音乐
- 桌面
- 收藏夹
- Cookie
- 模板 不管您使用的是不是永久池，桌面映像中都不应包含本地数据。如果组织要求使用本地数据存储，则应对这种要求加以限制。

此外，锁定桌面也很重要，其中包括防止用户在虚拟机的根目录驱动器上创建文件夹。在设计 View 部署的安全性时，您可以充分利用基于服务器的传统计算模型中的诸多策略设置和准则。

重定向所有数据最终可以获得的好处是，您只需归档或备份用户数据和基础模板。无需备份每个单独的用户 View 桌面。

配置注意事项

构建桌面映像时，应确保虚拟机不会消耗不必要的计算资源。您可以安全地做出以下配置更改以提高性能和可扩展性：

- 关闭图形屏幕保护程序。仅使用基本的空 Windows 登录屏幕保护程序。
- 禁用脱机文件和文件夹。
- 禁用除字体平滑之外的所有 GUI 增强功能，例如主题。
- 禁用所有 COM 端口。
- 用户注销时删除本地缓存的漫游配置文件。

在您的 View 实施过程中，需要考虑托管桌面中运行的应用如何访问存储，这一点非常重要。例如，如果多个虚拟机共享一个基础映像，并且同时都在运行病毒扫描，则性能可能会明显下降，因为这些虚拟机都会试图在同一个时间使用同一个 I/O 路径。对存储资源进行过度同步访问会降低性能。

根据您的存储缩减计划，一般应将虚拟机的交换文件与虚拟机的快照文件存储在不同的位置。在有些情况下，利用快照技术节约存储时，可能会需要使用交换文件，从而增加了快照大小，以至于性能降低。举例来说，如果您在有大量磁盘活动和极高内存使用频率的环境中使用虚拟桌面，尤其是在因所有内存均已使用而导致页面文件活动增加的情况下，就可能出现这种现象。这仅适用于您使用基于阵列的快照来节约共享存储使用的情况。

存储协议选项

VMware ESX 3.0 及更高版本支持为虚拟机磁盘 (VMDK) 存储使用多种协议选项：

- 光纤通道协议 (FCP)
- iSCSI
- NFS
- 10 千兆位以太网 (10GbE)

选择协议时主要考虑以下事项：最大吞吐量、每个协议上发生的 VMDK 行为以及重用现有的存储基础架构相对于采购新存储基础架构的成本。

存储协议的吞吐量

基于存储协议的每台主机的虚拟机数显示了以下主要存储协议的最大吞吐量：

基于存储协议的每台主机的虚拟机数				
	最大传输率	理论最大吞吐量	实际最大吞吐量	单台主机上的虚拟机数量（对于每 20 个虚拟机为 7 MB/秒时）
FCP	4 Gb/秒	512 MB/秒	410 MB/秒	1171
iSCSI (硬件或软件)	1 Gb/秒	128 MB/秒	102 MB/秒	291
NFS	1 Gb/秒	128 MB/秒	102 MB/秒	291
10 GbE (光纤)	10 Gb/秒	1280 MB/秒	993 MB/秒	2837
10 GbE (铜)	10 Gb/秒	1280 MB/秒	820 MB/秒	2342

表 2：基于存储协议的每台主机的虚拟机数

最大实际吞吐量远远超出了最大 ESX 主机的需求 — ESX 4.1 主机支持 128 GB 的 RAM（每个 Windows 客户机使用 512 MB 的 RAM）。

几乎所有生产 View 部署都不会使用最大支持内存量，即每台 ESX 主机 128 GB RAM，这是因为受到成本上的约束，例如为主机配备 8 GB SIMM 内存而不是 2 GB 或 4 GB DIMM 的成本。ESX 主机很有可能在遭遇磁盘 I/O 瓶颈之前就用完 RAM 或 CPU 时间。但是，如果磁盘 I/O 真的成为瓶颈，则极有可能是由于磁盘布局和磁盘轴数（即没有足够的 IOPS）造成的。Windows 虚拟机的吞吐量需求通常不是存储设计的决定因素。

注意：为了呈现只使用一个物理路径的一个数据会话的最坏情况，我们没有考虑链路聚合。

各个协议上发生的 VMDK 行为

FCP 和 iSCSI 都是数据块级协议。ESX 可以直接访问磁盘数据块，并可以控制将数据块装配到文件中。数据块级协议被 ESX 主机格式化为 VMware VMFS，并使用 ESX 文件锁定机制，最多只允许 32 个 ESX 主机访问同一个 LUN。数据块级协议还使用整盘式（厚磁盘）VMDK 格式，即每个 VMDK 在创建时均为完全调配，因此，无论 VMDK 内容占用多大空间，一个 20 GB 的磁盘都可以使用 20 GB 的数据块级存储空间。

NFS 是一种文件级协议。NFS 设备控制文件锁定以及将数据块装配到文件中。文件级协议使用精简磁盘 VMDK 格式，即 VMDK 只能与其内容一样大，因此，一个包含 10 GB 数据的 20 GB 磁盘将使用 NFS 存储设备上的 10 GB 空间。ESX 支持在单个主机上使用最多 32 个 NFS 数据存储。

自 ESX 版本 2.0 开始，就一直使用格式化为 VMware VMFS 的已连接 FCP 的 LUN。数据块级协议还允许对虚拟机使用裸磁盘映射 (RDM)。但 RDM 一般不用于 Windows XP 或 Windows 7 虚拟机，因为终端用户通常没有必须要使用 RDM 的存储要求。FCP 在基于 Windows 的数据中心内用于生产的时间要远远长于 iSCSI 或 NFS。

VMware 在 ESX 3.0 中引入了对 iSCSI 和 NFS 的支持。

iSCSI 是一种数据块级协议，与 FCP 的功能相同，但 iSCSI 采用的一般是比较便宜的介质（1 Gb/秒的以太网）。iSCSI 解决方案可以使用内置的 iSCSI 软件启动器，也可以使用硬件 iSCSI HBA。使用软件启动器将增加 ESX 主机上的 CPU 负载。与光纤通道 HBA 一样，iSCSI HBA 也会将这种处理负载卸载到专用卡上。要增加 TCP/IP 传输的吞吐量，您应该将巨型帧与 iSCSI 结合使用。VMware 建议的帧大小为 9000 字节。

NFS 解决方案始终采用软件驱动。因此，存储流量会增加 ESX 主机上的 CPU 负载。

对于 iSCSI 和 NFS 而言，新型网卡的 TCP/IP 卸载功能可以减小这些协议的 CPU 负载。

如果您使用 iSCSI 或 NFS，则根据当前数据中心网络的容量和体系结构，可能需要构建一个独立的物理以太网结构，以将存储流量与常规的生产网络流量分开。FCP 始终需要独立的光纤结构，这种结构在特定的数据中心内可能已存在，也可能不存在。

使用现有的基础架构

要确定是使用现有的存储基础架构还是购买新的基础架构（结构或阵列），您应该对数据中心内所有现有设备的容量和功能进行评估，并考虑回答以下问题：

- 您是否有位于 ESX 4.1 硬件兼容性列表中的存储阵列？
- 现有的阵列是否有足够的 IOPS 容量用于预期数量的虚拟机？
- 现有的阵列是否有足够的存储容量用于虚拟机？
- 您是否有可支持预期数量的 ESX 主机的结构（以太网或光纤通道）？
- 您是否有用于虚拟化服务器的 VMware Infrastructure 环境？如果有，是否有足够的存储容量来支持新的 VMware View 环境？

最终的协议选择

在为支持 VMware View 实施的 VMware ESX 存储选择最终协议时，相对于技术而言，往往更注重财务和心理方面的因素。如果您需要购买新的结构和阵列，则总体拥有成本和投资回报率将成为决定是否购买该存储结构和阵列的主导因素。如果您可以使用现有的结构和阵列，则新的 VMware View 实施将沿用现有基础架构的技术功能。

存储技术选项

VMware View 为您提供六个主要的存储选项：

- 标准共享存储，与用于基础虚拟服务器部署的存储类似
- 存储级精简配置，位于虚拟机级别或卷级别
- 单实例，又称重复数据消除
- 仅存储用户数据并通过网络引导操作系统
- 本地存储
- 分层存储

通过正确使用上述一种或多种技术，可以降低大多数 VMware View 环境的总体存储要求。您在既定环境中采用哪种技术将取决于组织的可用性及性能要求。

标准共享存储

使用标准共享存储的 VMware View 与典型虚拟服务器的存储使用类似。从共享存储上的卷创建逻辑单元号 (LUN)，并将其提供给 ESX 主机，ESX 主机使用该 LUN 来存储经过完全调配的各个虚拟机。

相对于用于虚拟服务器的 LUN 而言，使用标准共享存储的 VMware View 工作负载所需的 LUN 更大。桌面工作负载的 I/O 需求通常比服务器工作负载低得多，但相对于其 I/O 需求而言，桌面占用的空间往往较大，尤其是在应用集安装在模板虚拟机上时。服务器 LUN 可能为 300 GB，而桌面 LUN 可能高达 500 GB。但对于 VMware View，基于阵列的快照可能需要更小的 LUN，因为使用基于阵列的快照会影响存储子系统的性能。

仅对于某些具有特定要求的桌面类型，才应将标准共享存储与 VMware View 一起使用，例如用于特定用户或者信息技术人员的一次性应用集，也可能是在只需要有限数量的桌面的小型 VMware View 实施中。

精简配置

精简配置术语用于描述减少存储子系统上所用存储空间量的多种方法。主要方法包括：

- 减少虚拟机磁盘文件中的空白部分。
- 通过在相同卷内克隆虚拟磁盘文件，仅向卷上的一小部分区域执行写入操作，大量虚拟机共享一个基础映像，从而减少相同虚拟机所用的空间。
- 通过在共享存储设备内对整个卷进行精简克隆，来减少一组克隆虚拟机所用的空间。每个卷本身都有基础卷的虚拟克隆。

您可以先对某些类型的标准共享存储进行精简配置。然后克隆各个虚拟机或克隆包含这些虚拟机的整个卷。您可以在存储层创建此克隆。虚拟机的磁盘写入可能会进入某种类型的快照文件，或者用于克隆各个虚拟机或卷的存储子系统可能会跟踪数据块级写入。

您可以在许多区域使用精简配置方法，也可以对精简技术分层使用或组合使用，但应评估这些技术在分开使用或组合使用时对性能的影响。分层方法从理论上讲是不错，但由于存储子系统中使用了查询表，因此效果可能并不理想。您必须在存储要求上的降低与给定解决方案为设计增加的开销这两个方面进行权衡。

虚拟机空白部分

减少虚拟磁盘文件中的空白部分是指消除虚拟磁盘文件中未使用的空间。虚拟机所用的存储取决于磁盘文件中实际的数据量。例如，使用 ESX 时，默认情况下会对 NFS 装载上的 VMDK 文件进行精简配置，这样的话，一个具有 24 GB 数据的 40 GB VMDK 在经过精简配置后可以使用 24 GB 的存储。

对各个虚拟机进行精简配置

一段时间以来，在虚拟化管理程序级共享基础 VMDK 虚拟磁盘一直是一种在技术上可行的方法。通过这种方法，您可以利用具有多个虚拟机快照的基础 VMDK 文件，无需手动配置或自定义 .vmx 文件。由于此方法利用了虚拟机快照，因此可以大幅减少存储，无需借助任何其他存储精简技术。

您还可以在存储子系统内共享基础映像。与在虚拟化管理程序级上共享 VMFS 卷内的 VMDK 一样，某些存储设备也可以在存储层上共享基础 VMDK 映像文件。在存储层进行共享，而不在 VMFS 卷内使用快照，这可以大幅提高可扩展性。提高比例可以达到 1:20 甚至更高。

文件层精简配置提示

- 配置客户操作系统，使其可以使用 Microsoft 的 Sysprep 实用程序进行复制。
- 密切关注各个虚拟机如何写入数据。例如，请不要在各个虚拟机克隆中安装大型应用，而是返回至基础映像并对其进行更新。
- 了解 Windows 页面文件的使用情况。某些虚拟机使用情形可能会导致基于阵列的快照和性能方面的异常问题。

对整个数据存储进行精简配置

在数据存储级别利用精简配置可以提供更多的存储精简机会。您可以克隆整个数据存储，而不是各个虚拟机，并将其作为不同的数据存储提供给 ESX 集群。但是，如果在共享存储设备的存储子系统中对存储进行虚拟克隆，则该克隆实际上不会使用两倍的存储。

此方法的原理是，您对原始的黄金级基础数据存储进行虚拟克隆，因此可以对具有多个（20 或更多）虚拟机的数据存储进行多次虚拟克隆。然后，您可以分别打开并使用每个虚拟机，而对于全部 100 个或更多虚拟机，您将仅使用原始黄金级基础数据存储占用的公共存储空间。

这种类型的精简配置严重依赖于存储制造商的快照技术。有些存储供应商可提供比其他供应商更好的性能。

数据存储层精简配置提示

- 对于 VMFS 数据存储，在 vCenter 中启用 `enableresignature` 选项以使用克隆数据存储。
- 与文件层配置一样，基础数据存储中的虚拟机中的客户操作系统必须处于可复制的状态。使用 Sysprep，准备好在打开虚拟机时即对其进行修改，包括新主机名、SID、域成员身份和网络连接信息。
- 性能不会是线性的。可能需要多个基础映像数据存储。每个基础数据存储的克隆的可扩展性主要取决于存储供应商的快照方法和效率。
- 注意 VMFS 数据存储和 NFS 装载对每个卷上 LUN 的限制。

精简配置摘要

预计可以将多种精简配置方法结合使用来压缩 View 存储。将来，可以将映像级精简配置与数据存储虚拟克隆结合使用。这些解决方案的可扩展性非常有前景，尤其是在使用 NFS 时。但是，对于 View 应用的真实场景，几乎尚无准确的现场数据。

单实例

共享存储数据的单实例（即重复数据消除）的概念非常简单：系统在共享存储设备上搜索重复数据，并通过匹配相同的数据来减少所需的实际物理磁盘空间量。重复数据消除是数据库领域的一个概念，管理员为此术语赋予的含义是在合并数据库中搜索重复记录。在共享存储中，重复数据消除是用来查找并删除重复数据对象（文件、块或数据块）的算法。存储系统中的原始指针会被修改，以便系统仍可以查找该对象，但磁盘上的物理位置会与其他指针共享。如果写入数据对象，则写入会进入新的物理位置，并且不再共享指针。

重复数据消除的方法各种各样，而且有关重复数据消除可为一线存储带来哪些好处的可用信息也可能是误导性的。

不管是哪种方法，重复数据消除都基于两个元素：哈希值和索引。 **哈希值** 哈希值是提供给每个对象的唯一数字指纹。哈希值通过一个公式产生，在该公式中，同一个哈希值不可能使用两次。但有一点需要注意，那就是两个对象可以具有相同的哈希值。某些线内系统仅使用基础哈希值，这种方法可能会导致数据损坏。缺乏对重复哈希值的辅助检查的任何系统都可能给您的 VMware View 部署带来风险。

索引 索引使用哈希目录或查询表。存储子系统使用哈希目录确定重复记录，即使在它们当前未执行磁盘读取和写入操作时也是如此。此方法使

索引可以独立工作，而不管本机磁盘的使用情况。

存储子系统使用查询表将哈希目录扩展到不支持多个数据块引用的文件系统。查询表可以在系统 I/O 和本机文件系统之间使用。不过此方法的劣势是，查询表可能会导致失败。

重复数据消除的应用

在备份存储市场中，重复数据消除具有明显的优势。重复数据的量可能是巨大的。20:1 甚至更高的比例都是很常见的。备份中的相同数据甚至更多，相对于一线虚拟机工作负载而言，在备份中执行重复数据消除时的速度往往慢得多。

两种重复数据消除方法是线内（实时）重复数据消除和写入后重复数据消除。选择哪种方法主要取决于您使用的存储类型。写入后重复数据消除是您应在 View 体系结构中考虑的主要方法。线内重复数据消除的速度尚未快到能够满足大型基础架构的速度要求。它主要是一种备用方法。

线内重复数据消除现在主要部署用于备份和存档，它可以实际跟踪对象记录，并仅将不重复的数据块写入（通常是专有的）后端文件系统。这种方法需要两个文件系统，一个在前端，一个在后端。它必须在向磁盘写入数据之前搜出重复的对象，而且必须非常快速地完成此操作。随着处理器计算能力的提高以及固态磁盘成本的降低，线内重复数据消除将成为一线存储越来越可行的一种选择。

供应商现在也在此领域做着巨大的努力。但线内重复数据消除仍缺乏 View 存储所需的成熟度。

使用写入后重复数据消除时，将在数据写入磁盘后再执行重复数据消除。这样就可以在系统资源空闲时执行重复数据消除过程，不会干扰一线存储的速度。但是这种方法也有劣势，那就是它需要获得足够的空间先写入所有数据，然后再整合数据，因此您必须在您的存储体系结构中规划出该临时使用的容量。

网络引导的操作系统

通过网络将操作系统直接引导至虚拟或物理设备是一个相对比较新的概念。仅在过去的几年，该技术已经成为标准方法（即使用已安装的操作系统）的替代方法。

这种技术允许设备通过网络从一对一的映像或共享映像引导操作系统。操作系统的磁盘驱动器是远程服务器上的一个映像文件。虚拟机通常会使用 PXE 引导至映像服务器，然后，映像服务器会通过网络为引导虚拟机提供正确的虚拟磁盘。

此技术所用的协议与专用于网络的 iSCSI（原始本地 SCSI 封装在 TCP/IP 数据包中）非常相似。当映像服务器提供虚拟磁盘作为只读映像（每个虚拟机实例有一个私有写入磁盘）时，流式传输增加了另一个注意事项。此写入磁盘中存在有自引导计算机以来所做的全部更改。您可以将缓存存储在各个位置，包括虚拟机的 RAM、本地硬盘或者网络文件服务器上的文件。虚拟机关闭时，所做的更改将消失，除非映像被配置为将写入映像保留在文件服务器上。

您还可以设置私有一对一映射。您仍然必须要为各个磁盘文件提供存储，但在这种情况下，存储位于映像服务器上。

通过网络引导的优势

- 节约 SAN 存储空间 — 每个虚拟机只需一个很小的 VMDK，因此似乎只需本地磁盘即可正常运行。
- 速度 — 几乎瞬间即可部署新虚拟机。
- 灵活性 — 您可以非常轻松地更改虚拟机正在引导和查看的内容，而且可以非常轻松地更新将许多单独的虚拟机部署在池中的一个基础映像。

通过网络引导的劣势

- 目前无法脱机工作。
- 默认情况下仅支持非永久池。
- 网络流量繁重。
- 提高了服务器硬件要求。
- 在大型部署中每个映像服务器都存在扩展问题。
- 由于只在一个位置上存放私有映像，因此引发访问问题。如果映像服务器出现故障，则这些桌面就无法访问，这样您可能就无法满足桌面虚拟机的 SLA。

小心：如果从单个映像主机流式传输私有唯一映像，则缺少该主机可能会导致服务中断。

本地存储

您可以使用本地存储（服务器中的内部备盘磁盘驱动器），但 VMware 不建议使用此方法。在决定是否要使用本地存储时，应考虑以下几点：

- 无法与商用存储阵列性能相匹配。本地磁盘驱动器可能具有所需的容量，但却没有存储阵列的吞吐量。拥有的磁盘轴越多，吞吐量越好。
- 使用本地存储时，无法使用 VMware VMotion™ 来管理卷。
- 本地存储不允许您跨资源池使用负载平衡。
- 本地存储不允许您使用 VMware High Availability。
- 您必须管理和维护本地存储和数据。

• 克隆模板以及从模板克隆虚拟机比较困难。 分层存储 分层存储已经问世多年，最近因其节约成本也得到关注。如果要问什么是分层存储，您可能会获得从“一种

数据分类方法”到“实施新存储的技术”的不同答案。分层存储确实是硬件、软件和流程（允许公司基于数据的价值或者对公司的重要性更好地对数据进行分类和管理）的组合。在任何公司实施分层存储环境的计划中，数据分类都是相当重要的一部分。如果不进行这种分类，几乎不可

能将正确的数据放在正确的存储层中。例如，您希望将最重要的数据保存在快速的高 I/O 存储上，即固态硬盘驱动器 (SSDD) 或光纤通道中，而将不太重要或者不常访问的数据保存在相对比较便宜的驱动器上，例如 SAS 或 SATA。通过做出这种更改，您可能会体验到更快的访问速度，并获得最佳的存储性能，因为您已经从那些存储设备中有效转移了非重要数据及其访问用户。此外，网络速度的加快可以提高访问速度，进而提高响应能力。实施分层存储环境时，并不是只有一种“正确”方法。根据公司的业务需求、公司的数据分类计划

以及对存

储环境的硬件、软件和日常维护的预算，实施过程也有所不同。

将 View Composer 副本和链接克隆存储在单独的数据存储上

VMware View 4.5 及更高版本提供分层存储选项。您可以将 View Composer 副本存储在 SSDD 上，而将链接克隆存储在相对比较便宜的驱动器上，例如 SATA。借助这一新的分层存储选项，可以加快密集型操作（例如同时调配许多链接克隆）的速度。

举例来说，您可以将副本虚拟机存储在基于固态硬盘的数据存储上。固态硬盘具有较低的存储容量和较高的读取性能，通常支持每秒 20,000 次 I/O (IOPS)。固态硬盘也是比较昂贵的。但是，View Composer 为每个 ESX 集群上的每个 View Composer 基础映像快照只创建一个副本，因此副本不需要很多存储空间。当在许多链接克隆上同步执行某项任务时，固态硬盘可以提高 ESX 读取副本的操作系统磁盘的速度。

链接克隆可以存储在比较传统的基于旋转介质的数据存储上。这些磁盘提供的性能较低，通常支持 200 IOPS。它们比较便宜，可提供高存储容量，因而适合于将许多链接克隆存储在大型的池中。ESX 不需要在链接克隆上执行密集型的同步读取操作。

通过这种方式配置副本和链接克隆可以降低当同时创建许多链接克隆时出现的 I/O 风暴所造成的影响。

注意：此功能专用于高性能磁盘解决方案供应商提供的特定存储配置。如果您的存储硬件不支持高读取性能，请不要将副本存储在单独的数据存储上。此外，您在为链接克隆池选择单独

的数据存储时，必须要遵循特定的要求：

- 如果共享副本数据存储，则该数据存储必须可从集群中的所有 ESX 主机访问。
- 如果共享链接克隆数据存储，则必须共享副本数据存储。仅当您在同一台 ESX 主机上的本地数据存储上配置了所有链接克隆时，副本才能位于本地数据存储上。

注意：仅在 vSphere 模式下才支持此功能。链接克隆必须部署在运行 ESX 4 或更高版本的主机或集群上。

将副本存储在单独的数据存储或共享数据存储上的可用性注意事项

池的可用性是需要考虑的另一个事项。将副本与链接克隆存储在相同的数据存储上时，View Composer 会在每个数据存储上创建一个单独的副本，以提高可用性。如果某个数据存储变得不可用了，则只有该数据存储上的链接克隆会受影响。其他数据存储上的链接克隆可以继续运行。

将副本存储在单独的数据存储上时，池中的所有链接克隆都将绑定到该数据存储的副本上。如果该数据存储变得不可用了，那么整个池都将不可用。

要提高链接克隆桌面的可用性，您可以为存储副本的数据存储配置高可用性的解决方案。

存储阵列决策

当您针对 VMware View 实施选择存储阵列时，您需要选择要使用的磁盘类型和 RAID 类型。此外，您可能要考虑快照、克隆、复制和可管理性。

当前的磁盘选项包括：

- 光纤通道 (FC)
- 串行 ATA (SATA)
- 串行连接的 SCSI (SAS)

这些选项指的是存储阵列所用的协议，而不是阵列与 ESX 主机之间的协议。您还必须选择磁盘速度，通常介于 7,200 RPM 与 15,000 RPM 之间。您通常可以基于预算对磁盘协议和速度做出选择，只要容量（IOPS 以及所需空间量）可以支持 View 部署的预计大小即可。

当前的 RAID 级别选项包括：

- RAID 4
- RAID 5
- RAID 1+0
- RAID 6
- RAID-DP（仅 NetApp）
- MetaRAID（仅 EMC Clariion）
- vRAID（仅 HP EVA）

您选择的 RAID 类型主要取决于所购买的存储阵列。通常情况下，您做出购买决策不取决于特定中端存储解决方案中支持的 RAID 类型。所有中端存储阵列均可通过多种方式以高冗余获得高性能。与选择各个磁盘一样，只要选择的 RAID 可以提供所需的 IOPS 速度、容量 (GB) 和磁盘故障冗余，即可使用上述任何一种 RAID。

注意：最多可以有 32 台 ESX 主机访问同一组 256 个分配的数据块级 LUN 或同一组 32 个 NFS 数据存储。

提高存储性能

以下部分介绍了用于提高存储性能的各种技术。用于提高存储性能的一种方式是通过缩短满足给定的 IO 请求所需的响应时间（又称延迟）。闪存技术的高性能和低成本产生了一些创建其他缓存进而提高存储性能的创新方式。此外，闪存技术还可通过减少与旋转磁盘相关的占地空间、能耗和散热来帮助减少对环境的影响。

固态硬盘

什么是固态硬盘？

固态硬盘 (SSD) 又称闪存驱动器，是一种使用闪存存储永久数据的存储设备。SSD 与硬盘驱动器 (HDD) 的不同之处在于 SSD 没有旋转磁盘或可移动读 / 写磁头。由于没有移动零件，因此 SSD 噪音更小，而且不易受到物理冲击的影响。SSD 的接口与 HDD 的接口相同，具有更短的访问时间和延迟，从而使其非常适于快速读写数据。

闪存是一种特定类型的 EEPROM（电可擦可编程只读存储器）。它无需电源即可维护存储的信息。闪存与典型 EEPROM 的不同之处在于，闪存能够整块擦除其中的数据，这对于需要频繁更新大量数据的应用而言将大有裨益。

戴尔的企业级硬盘和超额配置

戴尔的多种产品都有 SSD 驱动器。如同在其有关 SSD 驱动器的产品概述中提到的那样，“通过了解企业级固态硬盘的基础知识，客户在比较解决方案时可以进行对等比较。”戴尔称，“SLC 闪存（单级单元）每个单元可以存储一位信息，而 MLC 闪存（多级单元）每个单元可以存储多位信息。”他们还称，“SLC 闪存的写容忍容量大约是 MLC 闪存的写容忍容量的 10 倍。”

超额配置用于提高 SSD 的写性能。戴尔称，“要重写已写入数据的 SSD 中的某个区域，必须先擦除这些数据，然后再写入。”为了最大程度地降低此操作的影响，戴尔的 SSD 驱动器还采用了一种名为超额配置的技术。“这种做法使本机闪存容量超出指定的使用容量，并利用超出的容量用作源数据的暂存区，以将应用写入数据快速放入已处于已擦除状态的闪存区域。SSD 会在通常不会影响应用性能的时段对此超额配置闪存空间执行清除功能。”

戴尔的多种产品线都利用了 SLC 闪存和超额配置。戴尔的解决方案包括：Dell PowerEdge、EqualLogic 和 Dell-EMC SSD 解决方案。

EMC 的 FAST Cache

什么是 FAST Cache？

FAST Cache 与 EMC 的 FAST（全自动存储分层）配合使用。FAST 可以通过连续监控和识别数据的活动级别，将活动数据自动移至高性能的存储层，而将不活动数据移至低成本 / 高容量的存储层。FAST Cache 是 EMC 开发的一项功能，此功能可通过扩展 DRAM 缓存的功能（通过将经常访问的数据映射至闪存驱动器）优化其存储系统的性能，这样可以为每个驱动器提供很大的容量（相对于 DRAM 容量而言）。FAST Cache 不需要其他硬件，只需现有的企业级闪存驱动器即可扩展缓存。

如何实施 FAST Cache？

FAST Cache 使用容量范围为 73 GB 至 2 TB 的现有闪存驱动器，如同《EMC CLARiiON 和 Celerra Unified FAST Cache 详细介绍》(EMC's CLARiiON and Celerra Unified FAST Cache – A Detailed Review) 白皮书中所述。FAST Cache 是在受 RAID 保护的读 / 写模式下创建的，容量选项取决于存储系统型号以及所安装的闪存驱动器的数量和类型。

FAST Cache 是否可以配置？

如前所述，FAST 可以连续监控数据并识别数据的活动级别，然后基于某种策略自动将活动数据和不活动的数据移至相应的存储层。通过 FAST 中的这些策略设置，管理员可以控制和管理自动化活动。

基本上说，FAST Cache 会使最常访问的数据尽可能长时间地保存在闪存驱动器上。如同《EMC CLARiiON 和 Celerra Unified FAST Cache》(EMC's CLARiiON and Celerra Unified FAST Cache) 白皮书中所述，当新数据需要复制到 FAST Cache 中时，“最近最少使用” (LRU) 算法可以确定应将哪些数据从 FAST Cache 中移出。这样可以确保常访问的数据尽可能长时间地位于 FAST Cache 中。

配置 FAST Cache 是一个无中断的联机过程。如同《EMC CLARiiON 和 Celerra Unified FAST Cache》(EMC's CLARiiON and Celerra Unified FAST Cache) 白皮书中所述，它使用现有的内存分配接口并且不占用主机(服务器)周期。您可以创建 FAST Cache，在存储卷上启用它，并通过 Unisphere 管理它。无需用户干预即可使应用程序体验到 FAST Cache 的性能优势。

哪些产品支持 FAST Cache？

在 CLARiiON 和 Celerra 统一存储系统中，对于 FLARE® 30 及更高版本而言，您可以将闪存驱动器用作 FAST Cache。

FAST Cache 的容量范围是 73 GB 至 2 TB。

NetApp 虚拟存储分层

什么是 Flash Cache (PAM II)？

NetApp Flash Cache 有助于提高随机读取密集型虚拟化工作负载的性能，而无需添加高性能的昂贵 SSD 驱动器。这种智能化的读取缓存可以提高数据的访问速度，从而将延迟缩短为旋转磁盘的十分之一甚至更短。Flash Cache 还可提供堪比 SSD 的性能，而且消除了必须将数据移至另一个存储层的复杂性。活动数据可以自动流入 Flash Cache，因为存储控制器后的每个卷和 LUN 均可使用该缓存。这就是虚拟存储分层的概念。NetApp 虚拟存储分层的工作原理如下：当 ESX 服务器向存储控制器发出读取请求时，将会检查主内存和 Flash Cache 以了解数据块是否已存在。如果该数据块在任一缓存中都不存在，则从磁盘中读取数据，并在缓存中为此数据保存一份副本。此独特技术的原理很简单：数据在磁盘上进行数据块级重复数据消除，因此缓存中的数据也是已进行重复数据消除后的数据。因此，当出现引导风暴或登录风暴等事件时，这些事件的影响将会由于缓存和重复数据消除而得以抵消。

如何实施 Flash Cache？

Flash Cache 具有硬件和软件组件，即 Flash Cache 模块和 FlexScale 软件许可证。Flash Cache 模块是一个 PCI Express 设备，每个设备包含 256 GB 或 512 GB 的 SLC 内存，存储控制器可以包含最多 8 TB 的附加缓存。

Flash Cache 模块是否可以配置？

Flash Cache 解决方案的软件组件 FlexScale 包含缓存策略，应用这些策略可以优化对小数据块的随机读取访问。缓存策略可以根据特定需求进行更改。Flash Cache 有三种工作模式：仅元数据模式、正常用户数据和低优先级数据。FlexScale 默认即时可用的配置是正常用户数据，此模式将元数据与正常用户数据模式组合在一起。这是虚拟化工作负载的优化设置。

如何优化缓存中的数据？

FlexShare 解决方案可为 Data ONTAP 提供有关如何管理卷的虚拟存储分层的提示。

缓存策略可以是以下值之一：

- 保持 — 此值告知 Data ONTAP，在重用缓存缓冲区之前，应尽可能长时间地保持缓存缓冲区。此值可以提高频繁访问（频繁对同一个缓存缓冲区进行多次访问）的卷的性能。此值会在部署链接克隆等技术时使用。“保持”缓存策略会在包含主映像或模板的卷上设置，从而将数据固定到缓存中。

- 重用 — 此值告知 Data ONTAP 应快速重用此卷的缓冲区。您可以使用此值。此值告知 Data ONTAP 应快速重用此卷的缓冲区。对于经常写入但很少读取的卷（例如数据库日志卷），或者对于即使保持缓存缓冲区也可能无法提高命中率的大数据集的卷，您可以使用此值。
- 默认值 — 此值告知 Data ONTAP 此卷应使用默认系统缓存缓冲区策略。

哪些产品支持 Flash Cache ？

NetApp 的多种 FAS/V 系列存储系统上都支持 Flash Cache 模块。Flash Cache 模块有 256 GB 和 512 GB 两种型号。根据型号不同，每个控制器支持的最大 Flash Cache PAM 模块数可以为 2 至 16 之间的任意值。这样可使缓存增大到 512 GB 至 8 TB，具体取决于您的存储系统。

NetApp 开发了一个软件包，可帮助您确定您的存储系统是否可从通过 Flash Cache 模块增加的缓存中受益。预测缓存统计 (PCS) 软件当前在 Data ONTAP 7.3 及更高版本中提供，您可以通过该软件预测添加相当于两倍、四倍或八倍系统内存的缓存所产生的影响。

总结

当您在评估您的 VMware View 实施的存储需求时，一定要考虑以下几点：

- 在为生产 View 实施做出设计选择时，应了解虚拟机的磁盘需求。Windows XP 或 Windows 7 客户端与提供服务器功能的虚拟机具有完全不同的需求。客户端的磁盘 I/O 超过 90% 是读取，速度相当慢（7 MB/秒或每 20 个虚拟机 112 IOPS）。此外，除了操作系统和应用安装外，需要很少的磁盘空间，因为所有终端用户数据都应存储在现有的网络集中存储、文件服务器或 NAS 设备上。
- 在您了解了给定 View 部署的磁盘大小、吞吐量和 IOPS 需求之后，即可获得选择存储协议、阵列类型、磁盘类型和 RAID 类型所需的信息。精简配置、重复数据消除和克隆可以大大降低所需的磁盘空间。
- 在制定存储决策时，最重要的考虑事项往往是财务，而不是技术。考虑以下问题：现有的数据中心资源是否可以重用？购买全新存储环境的价值主张和投资回报率是什么？

关于作者

Fred Schimscheimer 是 VMware 的高级技术营销工程师。作为技术营销工程师，他不仅是产品营销团队的一员，并且是虚拟桌面解决方案的存储和工作负载方面的专家。

参考资料

戴尔的企业级硬盘驱动器概述

<http://www.dell.com/content/topics/topic.aspx/global/products/pvaul/topics/en/hard-drives?c=us&cs=555&l=en&s=biz&~tab=2>

《EMC CLARiion 和 Celerra Unified FAST Cache 详细介绍》(EMC CLARiion and Celerra Unified FAST Cache – A Detailed Review)

<http://www.emc.com/collateral/software/white-papers/h8046-clariion-celerra-unified-fast-cache-wp.pdf>

EMC 术语 — 全自动存储分层 (FAST) 缓存

<http://www.emc.com/about/glossary/fast.htm> <http://www.emc.com/about/glossary/fast-cache.htm>

EMC 演示

<http://www.emc.com/collateral/demos/microsites/mediaplayer-video/demo-fast.htm>

NetApp Tech OnTap — 在不添加磁盘驱动器的情况下提高性能

<http://www.netapp.com/us/communities/tech-ontap/pam.html>

NetApp Flash Cache (PAM II)

<http://www.netapp.com/us/products/storage-systems/flash-cache/>

NetApp Flash Cache (PAM II) — 在不添加磁盘驱动器的情况下优化存储系统的性能

<http://media.netapp.com/documents/ds-2811-flash-cache-pam-II.PDF>

NetApp Flash Cache (PAM II) 技术规范

<http://www.netapp.com/us/products/storage-systems/flash-cache/flash-cache-tech-specs.html>