

为何在存储的缓存层采用英特尔® 傲腾™ 固态硬盘是优于 NAND 闪存固态硬盘的选择

始终如一的高性能、低时延和出色的耐用性使英特尔® 傲腾™ 固态硬盘成为优于 NAND 闪存固态硬盘的选择

作者

Flavio Fomin

解决方案架构师, 英特尔公司

目录

- 执行概要1
- 简介..... 2
- VMware vSAN 架构组件和 I/O 处理流程 2
- 测试环境 3
 - 硬件和软件配置..... 3
 - 虚拟机 (VM) 配置..... 3
 - 工作负载及测试方法..... 3
 - 报告工具..... 4
- 测试结果 4
 - vDisk 准备用时 4
 - 32 KB 块大小下 70/30 读/写工作负载 4
 - 数据从缓存到磁盘的落盘对 VM 时延的影响 5
 - 物理层的 I/O 事务..... 6
- 在缓存中使用低耐用性硬盘的风险 7
- 固态硬盘性能和耐用性对 vSAN 集群配置预估的影响..... 7
 - 轻型 I/O 工作负载的配置预估..... 7
 - 中型 I/O 工作负载的配置预估..... 8
- 结论..... 9

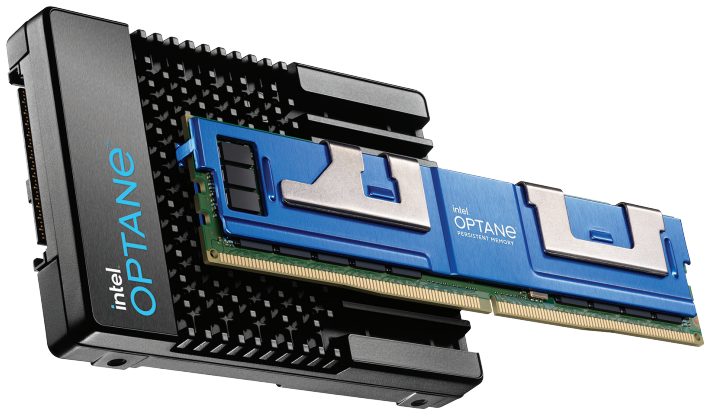
执行概要

本文探讨了 NAND 闪存和英特尔® 傲腾™ 固态硬盘的功能将如何影响 VMware vSAN 配置, 尤其是在存储的缓存层。我们考量的因素包括成本、性能、占用空间和数据丢失风险。

在选择用于缓存的硬盘时, 通常会采用通用的性能和耐用性准则。这些准则能避免硬盘配置不足, 但如果作为推荐准则, 则可能会使您错过优化成本和性能的机会。在某些情况下, 所选的硬盘可能无法为特定工作负载提供所需的耐用性, 从而增加数据丢失的风险。

在本文中, 我们提供了通用的配置预估准则中未包含的额外信息。我们介绍了当英特尔® 傲腾™ 技术部署在缓存层时, 将如何为多层分布式存储 (如 VMware vSAN) 带来额外的价值。这些价值正源于英特尔® 傲腾™ 固态硬盘与 NAND 闪存固态硬盘相比所提供的更高吞吐量和耐用性:

- 在处理 50% 读/写负载时, 英特尔® 傲腾™ 固态硬盘提供的吞吐量提升高达 2.5 倍¹。
- 英特尔® 傲腾™ 固态硬盘的使用寿命比 NAND 闪存固态硬盘长 16 倍²。



当部署在缓存层时, 英特尔® 傲腾™ 技术为可多层分布式存储带来额外的价值

简介

在配置新的 VMware vSphere 和 vSAN 集群时，大家都希望使工作负载的要求与主机配置中硬件组件的特性相匹配，从而优化成本和性能，实现更优的服务级别。

所有主机组件都会影响这一配置预估过程，但配置多少 NAND 闪存固态硬盘可能比配置其他组件更具挑战性，因为 I/O 性能 (MB/s 和 IOPS) 会随着工作负载配置文件而变化。例如，当我们增加一项并发的写入 I/O 事务时，NAND 闪存固态硬盘的读/写吞吐量和读取时延就会发生变化。在预估存储解决方案配置，尤其是在配置缓存层时，应该考虑到这些差异。

在配置用于缓存的硬盘时，耐用性是另一个不容忽视的因素。如果我们只是遵循以 TB 写入量 (TBW) 为单位的最低耐用性水平准则，而不去匹配工作负载与硬盘的耐用性，则硬盘有可能在保修期之前就已经磨损。尽管供应商可能会为您更换故障的硬盘，但故障仍可能使企业面临数据丢失风险，并对服务级别产生影响。例如，在 VMware vSAN 集群中，如果缓存硬盘出现故障，那么整个磁盘组都将变得不可用，数据丢失风险将大大增加，直到整个磁盘组被替换。

在本研究中，我们比较了 VMware vSAN 在使用 NAND 闪存固态硬盘与英特尔® 傲腾™ 固态硬盘时的性能差异。我们评估了不同读/写 I/O 组合下 NAND 闪存固态硬盘性能的变化。也就是说，数据从缓存到磁盘的落盘速率逐渐增加会影响整体性能；这就解释了为什么在处理相同的读/写混合负载时，英特尔® 傲腾™ 介质可以支持比 NAND 闪存固态硬盘更高的吞吐量³。我们还关注了不同的耐用性水平，并注意到使用过时的耐用性准则来选择硬盘，会导致数据丢失的风险升高。最后，我们对所有的发现进行了总结，并通过两种集群配置对成本、性能、占用空间和数据丢失风险进行了比较。

VMware vSAN 架构组件和 I/O 处理流程

为了便于解释下一节中给出的数据结果，我们对核心硬件组件上的 vSAN I/O 处理流程 (参见图 1) 进行了概括性总结。

- 写入请求:** 在容量层仅使用固态硬盘的 VMware vSAN 配置中，vSAN 将所有写入 I/O 请求缓冲至缓存固态硬盘，读取也来自仍驻留在缓存中的数据。在 vSAN 7.0U3 之前，写入缓冲限制为 600 GB。
- 数据复制:** 根据所使用的保护机制，对一台主机的写入请求必须复制到其他主机。值得注意的是，该保护机制会增大缓存和容量层硬盘的事务。例如，对每个虚拟机 (VM) 写入请求使用 RAID 1 会导致缓存层中产生两个写入 I/O 请求。
- 数据落盘:** 数据从缓存层到容量层的落盘速率是可变的 (具体速率与占用缓存的百分比有关)：
 - 0-30% 已满: 无落盘 (实际缓存命中率为 100%)。
 - 30-70% 已满: 限制速率落盘，试图调节数据落盘速率，在前端和后端 I/O 请求之间达到平衡⁴。
 - >70% 已满: 不限速率落盘，同时数据进入缓存的速率被限制，以与落盘速率达到平衡。
- 读取请求:** 读取来自容量层的硬盘，除非数据仍驻留在缓存层中。

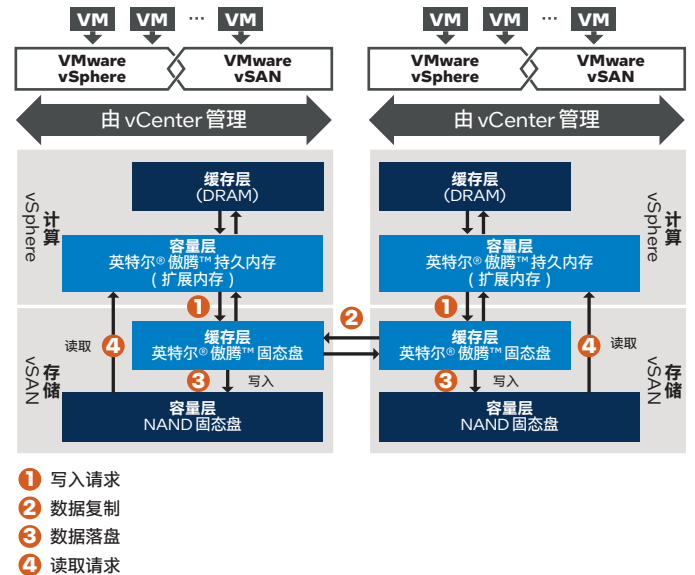


图 1. 硬件架构组件和 I/O 处理流程

测试环境

在本研究中，我们比较了使用两种硬盘型号作为缓存导致的 VMware vSAN 性能差异：

- 800 GB NAND 闪存固态硬盘，采用 SAS 接口，每日整盘写入次数 (DWPD) 为 3 次。
- 400 GB 英特尔® 傲腾™ 固态硬盘，采用 PCIe/NVMe 接口，每日整盘写入次数 (DWPD) 为 100 次。

我们根据原始设备制造商 (OEM) 的反馈选择了目前较为流行的 800 GB NAND 闪存固态硬盘。1.6 TB PCIe Gen4 NAND MU 闪存盘的性能规格优于 800 GB SAS MU 固态硬盘，但不如 400 GB 英特尔® 傲腾™ 固态硬盘 P5800X⁵。考虑到 800 GB 每日整盘可写入 3 次的固态硬盘是入门级硬盘，我们还是选择了这一款，希望在本研究中能展示两种极端情况下的性能。

硬件和软件配置

测试在一个 4 节点集群上执行，硬件和软件配置如表 1 所示。

表 1. 测试集群硬件和软件物料清单

主机组件	产品型号
CPU	2 x 英特尔® 至强® 金牌 6348 处理器 (2 x 28 核 @ 2.6 GHz)
内存	256 GB (16 x 16 GB DIMM, 3200 MT/s)
网卡	1 x 英特尔® 以太网适配器 E810 100 GbE
容量层硬盘	8 x 英特尔® 固态硬盘 D7-P5510 3.84 TB
缓存层	
主机配置 1 (NAND 闪存固态硬盘)	2 x 三星 PM1645 800 GB MU (混合用途) SAS 3 DWPD
主机配置 2	2 x 英特尔® 傲腾™ 固态硬盘 P5800X 400 GB NVMe 100 DWPD
软件	产品名称/版本
虚拟机管理程序	vSphere 7.0U2
SDS	vSAN 7.0U2
基准测试工具	HCIBench 2.5.3/VDBench 作为 I/O 驱动程序

虚拟机 (VM) 配置

测试环境由表 2 中提供的虚拟机 (VM) 配置方案组成。

表 2. VM 配置详情

测试环境元素	配置方案
VM 数量	每台主机 8 台 VM/总计 32 台 VM
vCPU/VM	8
内存/VM	8 GB
vDisk/VM	2 个 vDisk, 每个 150 GB
存储容量分配	9.6 TB 可用容量/ 19.2 TB 原始容量 (基于 RAID 1)

工作负载及测试方法

比较测试使用以下 I/O 工作负载对固态硬盘进行基准测试：

- HCIBench 磁盘准备，将测试配置中分配给 64 个 vDisk 使用的 9.6 TB 区域写零。
- 每次测试前都会清空缓存。
- 70/30 32 KB VDBench 上升曲线 (ramp-up curve)。

测试中使用了 VDBench 的“性能曲线”选项来生成上升曲线。使用此选项时，VDBench 一开始会在不限制 I/O 速率的情况下运行测试，以确定测试 I/O 配置文件可执行的最大 I/O 速率。接着 VDBench 会再运行 10 次测试，首先使用前一次达到的最大 I/O 请求的 10%，随后每次增加 10%，直到达到最大值 (参见图 2)。

重要提示： 确定最大 I/O 速率时，写入缓冲区的利用率接近于零。随着测试从 10% 逐渐上升到 100%，VM 的 I/O 请求升级，落盘事务也会逐渐增加。

70/30 32 KB VDBench 上升曲线

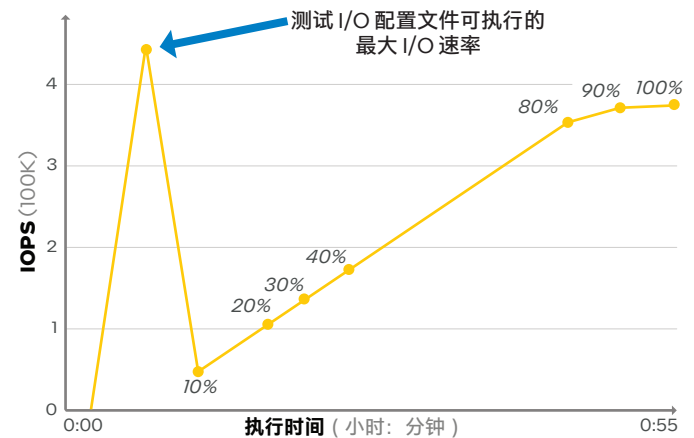


图 2. VDBench 的上升曲线选项

报告工具

除了总结测试工作负载 I/O 事务及时延的 HCI Bench 报告之外，本研究还使用了来自 vSAN 性能数据库的数据，其中包括写入缓冲区利用率、数据从缓存到磁盘的落盘速率以及缓存和容量层固态硬盘的物理事务等（参见图 3）。

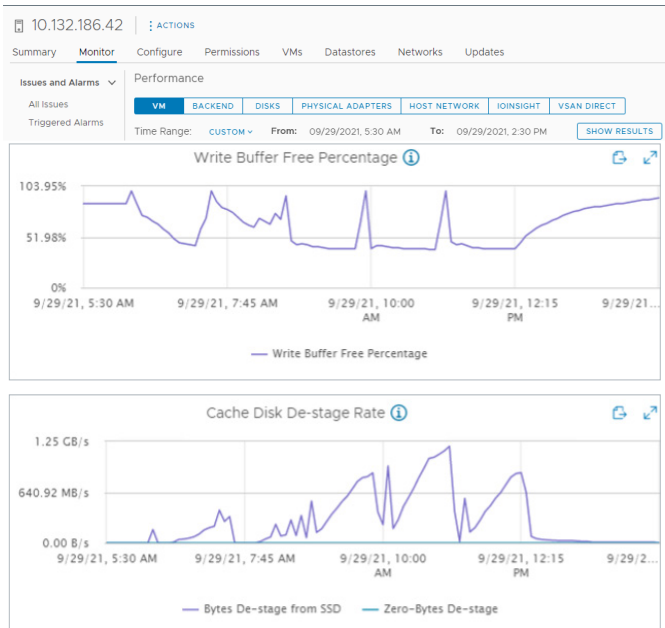


图 3. vSAN 性能数据库信息

测试结果

以下部分介绍了评估 VMware vSAN 性能的四中不同方法：

- vDisk 准备用时
- 32 KB 块大小下 70/30 读/写工作负载
- 数据从缓存到磁盘的落盘对 VM 时延的影响
- 物理层的 I/O 事务

vDisk 准备用时

作为测试环境准备的一部分，vDisk 必须被完全写入，以确保在测试之前已经完成精简配置空间分配。在此测试配置中，vDisk 准备总共写入了 9.6 TB 的数据。当使用 800 GB NAND 闪存固态硬盘作为缓存硬盘时，此操作花费了 1 小时 55 分钟。相比之下，使用 400 GB 英特尔® 傲腾™ 固态硬盘作为缓存硬盘时，同样的操作只需 30 分钟即可完成，速度快了近 4 倍（见图 4）⁶。我们可以将这些结果推广到其他写入密集型的工作负载上，例如视频编辑、备份和恢复作业以及缓冲区刷新等。请注意，如果将 400 GB 英特尔® 傲腾™ 固态硬盘 P5800X 与 1.6 TB PCIe Gen4 NAND 闪存固态硬盘进行比较，写入吞吐量差异将降低至 1.8 倍⁷。虽然差距有所降低，但 1.8 倍的速度差异仍然十分显著。

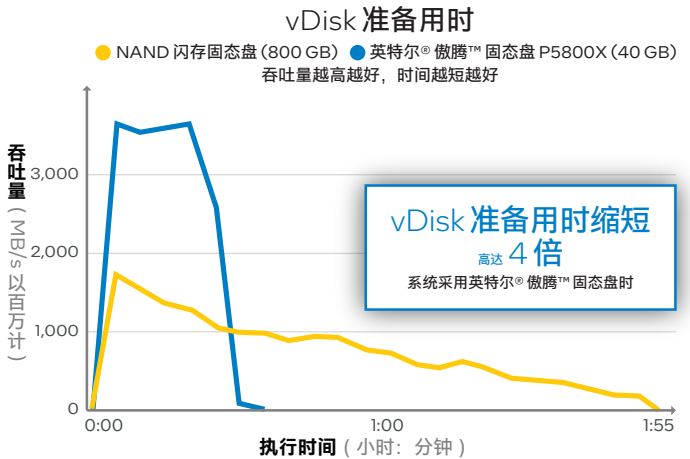


图 4. 英特尔® 傲腾™ 固态硬盘完成 vDisk 准备的速度比 NAND 闪存固态硬盘快了将近 4 倍

32 KB 块大小下 70/30 读/写工作负载

图 5 展示了随机 I/O 工作负载下（70% 读，30% 写，块大小为 32 KB），I/O 与时延关系的曲线。用作缓存的固态硬盘是测试之间唯一更改的硬件组件。我们可以观察到，使用 400 GB 英特尔® 傲腾™ 固态硬盘 P5800X 比使用 800 GB NAND 闪存固态硬盘得到的 IOPS 高出 2 倍⁸。

我们还观察到，在接近最大 I/O 速率时，800 GB NAND 闪存固态硬盘的曲线表现得非常不稳定。发生这种情况的原因是，当 vSAN 系统无法足够快地让数据从缓存硬盘中落盘时，它会开始限制 VM 的 I/O 请求，以平衡 VM 请求和落盘事务。相比之下，英特尔® 傲腾™ 固态硬盘的曲线在工作负载升级时表现出了可预测的性能。

在接下来的两节中，我们将详细评估 vSAN 层数据和物理层数据，以更好地了解 I/O 性能的这种差异。

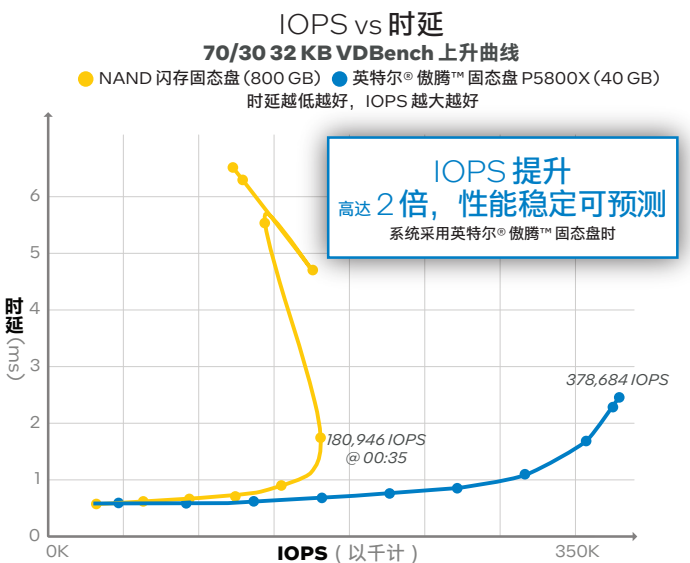


图 5. 即便 I/O 工作负载升级，英特尔® 傲腾™ 固态硬盘也能提供可预测的 I/O 性能

数据从缓存到磁盘的落盘对 VM 时延的影响

如图 6 图表所示，我们观察了在处理上一节中同一 70/30 32 KB 工作负载时，写入缓冲区利用率和数据从缓存到磁盘的落盘速率。我们可以观察到缓存事务与 VM 的 I/O 事务运行情况之间的关联。

如“工作负载及测试方法”部分所述，测试的前六分钟不对 I/O 事务运行加以限制，目的是为了找到该配置下的最大 I/O 速率。此信息随后被用于生成上升曲线。在本节中，我们将此任务称为“ratemax”。执行 ratemax 任务时，写入缓冲区几乎 100% 空闲，这表明此时没有落盘事务争用资源。

在使用 800 GB NAND 闪存固态硬盘的测试中，系统执行 ratemax 时 IOPS 达到了 305,000。因为 NAND 闪存固态硬盘上有一个 600 GB 的写入缓冲区，写入缓冲区的利用率达到了近 40%，这触发了 8 MB/s 的轻量落盘事务。随着 I/O 事务的进行，落盘速率随着 I/O 工作负载的增加而增加，落盘速率稳定在 280 MB/s 附近。此时，当测试进行到 60% ratemax 附近的时候，VM 写入 I/O 时延跃升至 15.5 毫秒，同时 IOPS 开始降低，因为 vSAN

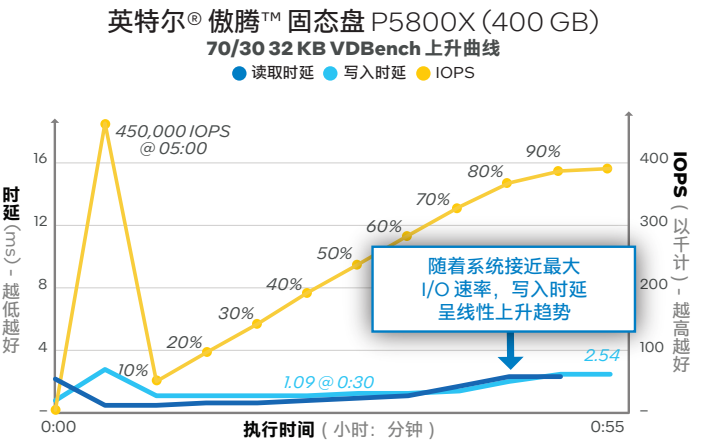
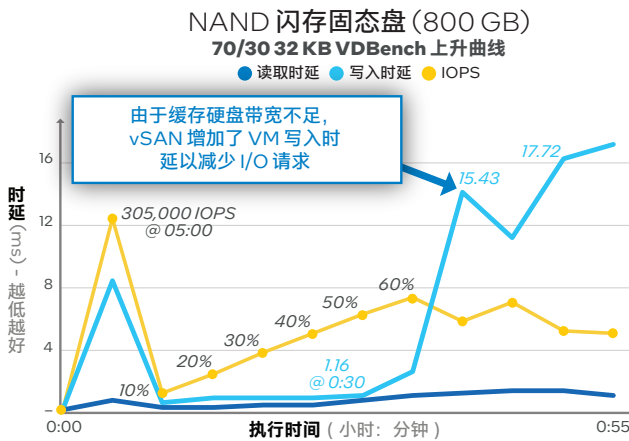
会限制 I/O 请求以避免缓存溢出。这就解释了上一节中描述的曲线不稳定的成因。

再来看英特尔® 傲腾™ 固态硬盘的图表，我们观察到在 ratemax 任务期间，I/O 速率提高了 50%，IOPS 达到了 450,000。由于写入事务更频繁，vSAN 触发了更高的初始落盘速率 (434 MB/s)，从而将缓冲区利用率保持在 60%。ratemax 任务下的缓存很快被耗尽；随着上升曲线测试开始，落盘速率降低到 146 MB/s，然后随着 VM 的 I/O 工作负载增加呈线性上升趋势。

这一次，我们比较了数据从缓存到磁盘的落盘对系统 I/O 能力的影响，发现缓存硬盘性能对整个 vSAN 的 I/O 性能至关重要。借此机会，我们也简要地讨论一下写入缓冲区大小的差异。

如“VMware vSAN 架构组件和 I/O 处理流程”部分所述，当缓存利用率在 30% 到 70% 之间时，vSAN 会采用“限制速率落盘”的策略；也就是说，为了跟上 VM I/O 事务的增加，落盘速率会逐步上升。

读取 vs 时延 vs IOPS 比较结果



空闲写入缓冲区 vs 数据落盘传输比较结果

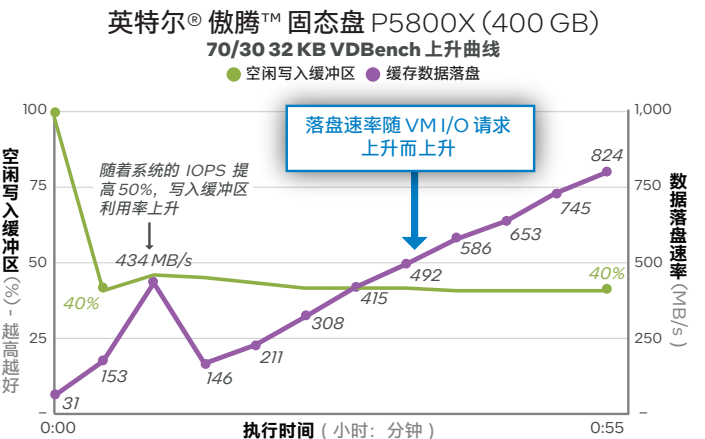
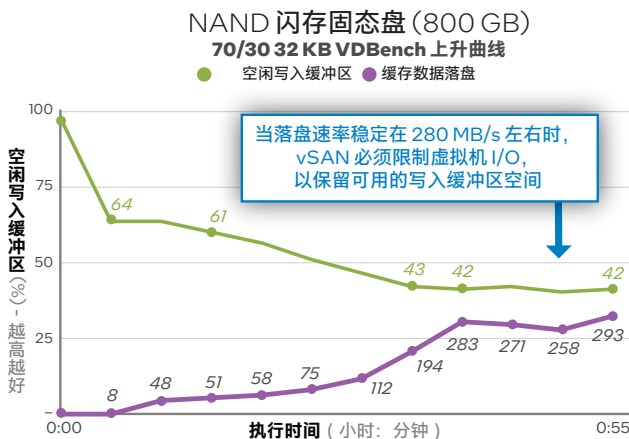


图 6. 英特尔® 傲腾™ 固态硬盘提供出色的性能，包括更高的 IOPS、更低的时延和更高的可预测性

使用 800 GB 固态硬盘时，写入缓冲区配置为最大的 600 GB，而使用英特尔® 傲腾™ 固态硬盘时，写入缓冲区为 400 Gb。后者的上升曲线测试以高出前者 50% 的 I/O 速率开始，这导致从测试一开始，后者的落盘速率就更高。但我们也可以观察到，在采用英特尔® 傲腾™ 固态硬盘的测试中，即使落盘速率明显更高，系统 VM 的 I/O 事务还是更活跃，且时延更低。

这样的比较也回答了一个常见的问题：缓冲区小于 600 GB 是否会对 vSAN 的 I/O 性能产生负面影响？我们发现，**核心差异点在于维持更高读/写带宽的能力，而非缓冲区大小。数据落盘是 vSAN I/O 可扩展性的重要限制性因素，而落盘速率则直接受硬盘性能和 vSAN 内部限制的影响。**如图 6 所示，英特尔® 傲腾™ 固态硬盘的落盘速率达到了约 820 MB/s，而 NAND 闪存固态硬盘的落盘速率则被限制在约 280 MB/s。

物理层的 I/O 事务

英特尔® 傲腾™ 固态硬盘可达到更高写入吞吐量的核心原因之一就是其“就地写入”的技术。

将数据写入 NAND 闪存固态硬盘时，需要进行编程-擦除 (P/E) 循环。P/E 循环就是将内存单元从编程状态转为擦除状态，然后再返回编程状态。P/E 循环非常耗时，直接影响硬盘性能。英特尔® 傲腾™ 介质的“就地写入”技术在将数据写入介质时不需要触发 P/E 循环，从而显著提高了写入的 I/O 密度。

在图 7 的左侧图表中，我们可以观察到，当工作负载几乎完全为写入时，NAND 闪存固态硬盘的写入吞吐率达到了 750 MB/s 左右。但是，随着落盘事务增加，硬盘上的事务接近读/写各

50% 的混合负载时，写入吞吐率在 439 MB/s 达到峰值，此时读取吞吐率为 368 MB/s。

在图 7 的右侧图表中，英特尔® 傲腾™ 固态硬盘在处理大部分均为写入的负载时达到了接近 1,100 MB/s 的写入吞吐率，并且在处理读/写各 50% 的混合负载时能维持几乎相同水平的写入吞吐率，写入和读取吞吐率分别达到了 1,019 MB/s 和 965 MB/s。
注：英特尔® 傲腾™ 固态硬盘的写入速度可远超 1 GB/s，但在此测试中，我们此时已达到 vSAN 的内部限制。

请注意，时延的增加与硬盘 I/O 限制相关。在 NAND 闪存固态硬盘上进行 ratemax 任务时，测试时延上升到 8 毫秒，此时写入缓冲区仍然处于清空状态，这表明硬盘的运行性能已饱和。在测试进行到近 35 分钟时，由于必须限制 VM 的 I/O 请求以在读写之间达成平衡，时延再次增加。但在英特尔® 傲腾™ 固态硬盘上，时延保持不变，直到测试接近 100% 最大 IOPS 时到达了 vSAN 的内部限制。

这一比较为上一节中有关数据从缓存到磁盘的落盘讨论提供了新的视角，进一步说明了英特尔® 傲腾™ 技术固有介质差异的重要性。相较于 800 GB NAND 闪存固态硬盘，400 GB 英特尔® 傲腾™ 固态硬盘吞吐率提升高达 2.5 倍 (1,984 MB/s vs 772 MB/s) 如最佳实践指南《部署经济高效的高性能 vSAN 集群》中所述，英特尔® 傲腾™ 固态硬盘的这种性能优势也可以扩展到其他比较中：与采用 NVMe/PCIe Gen 4 接口的 1.6 TB NAND 闪存固态硬盘相比，400 GB 英特尔® 傲腾™ 固态硬盘的每 GB 写入吞吐量提升高达近 8 倍⁹。

物理缓存磁盘事务测试结果

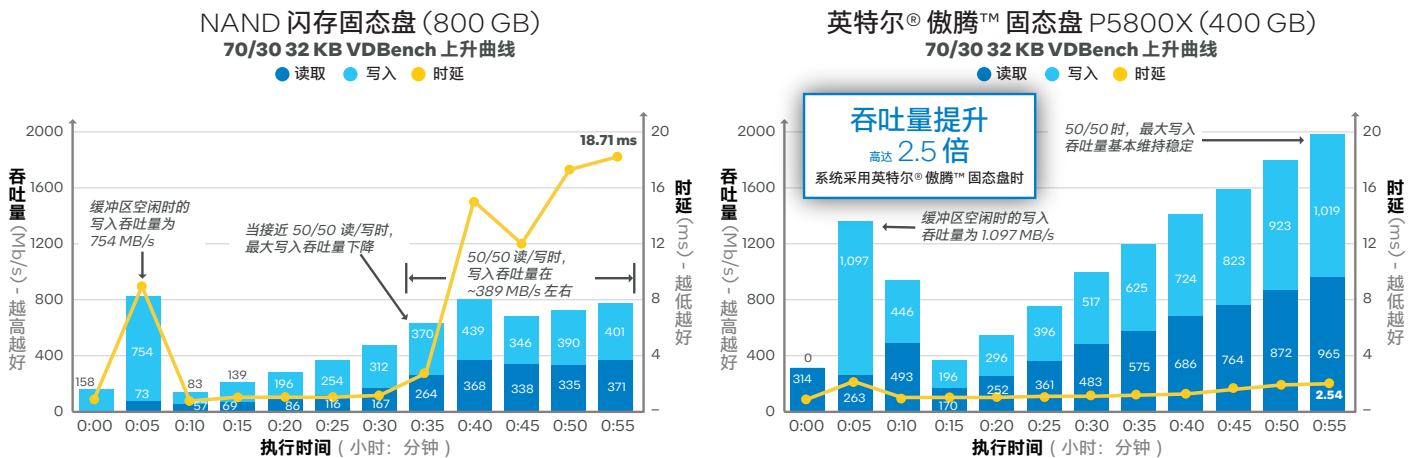


图 7. 本次比较显示，400 GB 英特尔® 傲腾™ 固态硬盘的吞吐率比 800 GB NAND 闪存固态硬盘高 2.5 倍

在缓存中使用低耐用性硬盘的风险

硬盘的使用寿命在商业上用 DWPD（每日整盘写入次数）来描述，即硬盘在五年内每天可写入的次数。在相同的时间段内，它也可以表示为 TBW（即写入的总 TB 数）。

使用耐用性准则（如 VMware 的 [面向 vSAN 的通用固态硬盘耐用性准则](#)）来确定在缓存层中使用什么固态硬盘是很常见的。在此 vSAN 准则中，全闪存配置有两种耐用性等级：C 类固态硬盘要求 TBW ≥ 3650，D 类固态硬盘要求 TBW ≥ 7300。

$$5 \text{ 年的 TBW} = \text{固态硬盘大小} \times \text{DWPD} \times 365 \text{ 天} \times 5 \text{ 年}$$

这些准则具有一定参考价值，因为它们可以避免缓存固态硬盘配置不足；但如果我们把这些准则当作金科玉律，所选的固态硬盘就有可能无法满足目标工作负载需求。

自大约 2016 年这些 TBW 准则提出以来，NAND 闪存固态硬盘行业已经推出了几种更大且更便宜的硬盘，以满足成本和容量的需求。在本文撰写时，我们观察到 800 GB SAS 3 DWPD (4300 TBW) 的固态硬盘已成为缓存固态硬盘的流行选择。这个现象值得进一步商榷。它表明，多年前提出的入门级硬盘耐用性建议在如今配置一个 VM 数量和 I/O 密度均有显著提升的系统（由于引入了新一代 CPU，且当今的混合云 I/O 工作负载更加密集）时仍然有效。

有时候，我们会认为这是一个小问题，因为如果硬盘在 5 年保修期内磨损，供应商会负责更换。但我们不应忽视这些故障（以及后续替换）可能给大规模部署带来的风险及其对服务级别的影响。例如，如果缓存硬盘磨损，那么整个 vSAN 磁盘组都将不可用。这种情况会增加数据丢失的风险，还会对运营产生影响，因为此磁盘组中的所有硬盘都必须取下，随后重新创建。在大规模部署中，运营成本是很重要的。

因此，我们建议在评估缓存硬盘时，一定要考虑其耐用性所能支持的最大写入 I/O 工作负载。例如，800 GB 3 DWPD 的硬盘在达到 5 年使用寿命时可支持 28.4 MB/s 的写入负载。

$$(3 \text{ DWPD} \times 800 \text{ GB} \times 1024) \div (3600 \text{ s} \times 24 \text{ h}) = 28.44 \text{ MB/s}$$

在评估硬盘时还必须考虑写入 I/O 增大的问题。在 800 GB 3 DWPD 硬盘的例子中，当我们考虑 RAID 1 保护机制的 I/O 增大时，28.4 MB/s 的值将变为 14.2 MB/s。

我们进而使用这个参考值评估了不同的 I/O 配置文件。例如，70% 读、32 KB I/O 大小、1,500 IOPS 的工作负载会产生 14 MB/s 的写入负载。

在表 3 中，我们将 800 GB 和 1600 GB NAND 闪存固态硬盘和 400 GB 英特尔® 傲腾™ 固态硬盘的耐用性与其在五年内可支持的 I/O 事务进行了比较。

表 3. 耐用性及其与 I/O 事务的关系

硬盘大小	耐用性 DWPD/ TBW	VM 写入 MB/s	70/30 32 KB IOPS
800 GB NAND 固态硬盘	3/4,300	28	1,500
1600 GB NAND 固态硬盘	3/8,600	56	3,000
400 GB 英特尔® 傲腾™ 固态硬盘	100/73,000	470	25,000

固态硬盘性能和耐用性对 vSAN 集群配置预估的影响

在最后一节中，我们将考虑缓存硬盘的特性，评估成本、性能、占用空间和数据丢失风险方面的差异，由此探讨如何进行集群配置预估。我们进行了两项比较：一项采用轻型 I/O 工作负载，另一项则采用中型 I/O 工作负载。

轻型 I/O 工作负载的配置预估

在第一种情况中，我们配置了一个可满足 150 台 VM 需求的集群。VM 的特性如表 4 所示。此工作负载达到了我们在使用 800 GB NAND 闪存固态硬盘进行的上升曲线测试中所观察到的有效最大 I/O 请求的 30%。满足表 5 所述特征的 4 节点集群可以满足该工作负载要求。

表 4. 轻型 I/O 工作负载 VM 特性

工作负载要求	每 VM	总计
#VM	N/A	150
总 IOPS-MB/s	360 I/O 请求 (70/30 32 KB)	54,000/1,912
可用存储	250 GB	38.8 TB
CPU 内核/GHz	4 个 vCPU;	154 个内核/325 GHz
内存	10 GB	1,500 GB

基于 NAND 闪存方案的考量因素

MU SAS 硬盘的带宽足以满足这一轻型 I/O 工作负载。然而，对于该 800 GB NAND 闪存固态硬盘而言，3 DWPD 的耐用性可能不足以满足这一工作负载，因为它每天需要在 8 块缓存硬盘上写入 13.5 次。

在出于性能考量预估缓存硬盘配置时，我们会使用峰值 I/O 请求指标。但是在考量耐用性时，我们应该使用平均写入事务。如果平均写入事务比用作输入的值低 4.5 倍（在本例中，即每 MB/s 的总 IOPS 为每 VM 360 次 I/O 请求），则 800 GB NAND 闪存固态硬盘足以满足此工作负载。否则，这就会增加此配置的数据丢失风险。

基于英特尔® 傲腾™ 固态硬盘方案的考量因素

英特尔® 傲腾™ 固态硬盘比 NAND 闪存固态硬盘更贵。在这种情况下，使用英特尔® 傲腾™ 固态硬盘会使主机成本增加 4.5%¹⁰。但是，英特尔® 傲腾™ 固态硬盘可提供高达 2 倍的 I/O 吞吐量提升¹¹。这种性能上的差异为集群提供了未来的保障，因此它可以消化比初始计划更大的 I/O 工作负载。

英特尔® 傲腾™ 固态硬盘的耐用性比 NAND 闪存固态硬盘高 16 倍¹²；这远远超过了这一轻型 I/O 工作负载在 5 年保修期内的 I/O 属性，最大限度地降低了因硬盘磨损故障而导致的数据丢失风险。

表 5. 轻型 I/O 工作负载配置方案

主机组件	基于 NAND 闪存固态硬盘的方案	基于英特尔® 傲腾™ 固态硬盘的方案
CPU	2 x 英特尔® 至强® 金牌 6330N 处理器 (28 核, 2.2 GHz)	
内存		
- DRAM	256 GB (16 x 16 GB, 3200 Mt/s)	
- 英特尔® 傲腾™ 持久内存 200 系列	1024 GB (8 x 128 GB)	
存储		
- 容量层硬盘	4 x 读取密集型 SAS 硬盘, 7.68 TB	
- 缓存层硬盘	2 x MU SAS 硬盘, 800 GB, 3 DWPD	2 x 英特尔® 傲腾™ 固态硬盘 P5800X, 400 GB, 100 DWPD ¹³
其他		
网络	1 x 25 GbE 网卡	
主机数	4	4

中型 I/O 工作负载的配置预估

在第二种情况中，我们继续配置一个可满足 150 台 VM 需求的集群。但在这个用例中，工作负载达到了我们在使用英特尔® 傲腾™ 固态硬盘进行的上升曲线测试中所观察到的有效最大 I/O 请求的 40%。VM 的特性如表 6 所示。

在这种情况下，使用 NAND 闪存固态硬盘作为缓存的成本优化型配置需要一个 7 节点集群，而使用英特尔® 傲腾™ 固态硬盘作为缓存的配置仍只需一个 4 节点集群（参见表 7）。

表 6. 中型 I/O 工作负载 VM 特性

工作负载要求	每 VM	总计
#VM	N/A	150
总 IOPS-MB/s	960 I/O 请求 (70/30 32 KB)	144,000/4,500
可用存储	100 GB	16.6 TB
CPU 内核/GHz	4 个 vCPU;	154 个内核/325 GHz
内存	8 GB	1,200 GB

表 7. 中型 I/O 工作负载配置方案

主机组件	基于 NAND 闪存固态硬盘的方案	基于英特尔® 傲腾™ 固态硬盘的方案
CPU	1 x 英特尔® 至强® 金牌 6330N 处理器 (28 核 @ 2.2 GHz)	2 x 英特尔® 至强® 金牌 6330N 处理器 (28 核 @ 2.2 GHz)
内存		
- DRAM	384 GB (8 x 16 GB + 8 x 32 GB, 3200 MT/s)	256 GB (16 x 16 GB, 3200 MT/s)
- 英特尔® 傲腾™ 持久内存 200 系列	N/A	1024 GB (8 x 128 GB)
存储		
- 容量层硬盘	6 x 读取密集型 SAS 3.84 TB	4 x 读取密集型 SAS 7.68 TB
- 缓存层硬盘	2 x 800 GB MU SAS 硬盘, 3 DWPD	2 x 英特尔® 傲腾™ 固态硬盘 P5800X, 400 GB, 100 DWPD
其他		
网络	1 x 25 GbE 网卡	
主机数	7	4

基于 NAND 闪存方案的考量因素

此 I/O 工作负载需要配置七台主机，因此需要 14 块 800 GB MU SAS 硬盘来维持足够的性能，因为较理想的 I/O 性能需每台主机使用两个磁盘组。由于主机数量增多，因此使用单路配置来优化成本。主机数量的增加会对占用空间和功耗/散热产生负面影响。

在该用例中，14 块 3 DWPD 的 800 GB NAND 闪存固态硬盘每块预计每天需要写入 20 次，所以这个配置的耐用性是远远不够的。于是，在 5 年保修期满之前更换硬盘的相关风险会大幅增加。

基于英特尔® 傲腾™ 固态硬盘方案的考量因素

由于英特尔® 傲腾™ 固态硬盘在每台主机上可执行的 I/O 请求比 NAND 闪存固态硬盘多 2 倍，因此可以整合主机使得占用空间减少 40%。

如果我们在基于英特尔® 傲腾™ 固态硬盘的方案中仅使用 DRAM（每台主机 1 TB DRAM），则该方案的成本将降低 6%¹⁴，并且仍然能够支持两倍的 I/O 工作负载要求。但是，由于主机整合增加了每台主机的内存，因此在该方案中采用英特尔® 傲腾™ 持久内存更具吸引力，因为这样的话总成本可降低 13%¹⁵。

具有 100 DWPD 的英特尔® 傲腾™ 固态硬盘可以支持该工作负载长达五年以上，从而最大限度地降低因硬盘磨损故障而导致的数据丢失风险和运营影响。

结论

这项研究表明，与测试中的 NAND 闪存固态硬盘相比，当使用英特尔® 傲腾™ 固态硬盘作为缓存时，vSAN 可实现高达 2 倍的性能提升，这是因为英特尔® 傲腾™ 固态硬盘在处理混合读/写 I/O 负载时出色的带宽能力。更高的带宽为主机整合创造了机会，从而节省了成本和占用空间。此外，它还建立了一个可预测性更高的性能环境，这在运行具有严格服务级别协议的关键任务应用时非常重要。

我们还讨论了耐用性，并说明了在选择缓存硬盘时考量一般耐用性准则之外因素的重要性。硬盘可支持的最大写入 I/O 工作负载是一项重要的考量因素，可以最大限度地减少因磨损故障导致的数据丢失风险和运营影响。

了解更多信息

以下资源可能对您有所帮助：

- [最佳实践指南：部署经济高效的高性能 vSAN 集群](#)
- [英特尔® 傲腾™ 固态硬盘](#)
- [使用英特尔® 傲腾™ 固态硬盘 P5800X 增强 VMware vSAN 集群的性能博客](#)

找到适合您的组织或机构的解决方案，请联系您的英特尔代表或访问[英特尔® 傲腾™ DC 固态硬盘](#)。



¹ 截至 2021 年 9 月 28 日英特尔所做测试。

通用配置：4 节点，2x 英特尔® 至强® 金牌 6348 处理器（28 核，2.6 GHz），总 DRAM 256 GB（16 插槽/16 GB/3200 MT/s），存储（容量）= 8x 英特尔® 固态硬盘 D7-P5510 3.84 TB，1x 英特尔® 以太网网络适配器 E810-CQDA2（100 GbE），英特尔® 超线程技术 = 启用，英特尔® 睿频加速技术 = 启用，采用 RAID1 的 VMware vSAN 7.0U2。
NAND 闪存固态硬盘缓存配置：2x 三星 PM1645 800 GB MU（混合用途）SAS。
英特尔® 傲腾™ 固态硬盘缓存配置：2x 英特尔® 傲腾™ 固态硬盘 P5800X 400 GB NVMe。

² 仅当持续写入工作负载低于每天 2,400 GB 或 28.4 MB/s 写入时，具有 3 DWPD 耐用性等级的基于 NAND 的 800 GB 硬盘寿命才能持续五年。换句话说，假设使用 RAID1 作为数据保护机制，即使是轻型的持续 I/O 负载，例如 450 32-KB 写入 IOPS 事务，也会每天填满硬盘三次。相比之下，400 GB 英特尔® 傲腾™ 固态硬盘 P5800X 的耐用性等级为 100 DWPD。它可以连续 5 年维持每天写入 40,000 GB 或 470 MB/s 的每日写入。这相当于耐用性增加了 16 倍。

³ 见尾注 1。

⁴ 前端请求来自 VM。后端请求代表 VMware vSAN 执行的所有物理 I/O 事务。

⁵ **PCI Gen4 1.6 TB NAND MU 闪存硬盘规格：**<https://www.samsung.com/semiconductor/cn/>

SAS 800 GB MU 硬盘规格：<https://www.dell.com/cn-zh/shop/dell-800-gb-solid-state-drive-serial-attached-scsi-sas-mixed-use-12gbps-512e-25-inch-hot-plug-drive-pm1645-3-dwpc-4380-tbw-ck/apd/400-azii/storage-drives-media>

400 GB 英特尔® 傲腾™ 固态硬盘 P5800X 规格：<https://www.intel.cn/content/www/cn/zh/products/sku/201861/intel-optane-ssd-dc-p5800x-series-400gb-2-5in-pcie-x4-3d-xpoint/specifications.html>

⁶ 见尾注 1。

⁷ 英特尔，“Deploying a Cost-Efficient, High-Performance vSAN Cluster,”（部署经济高效的高性能 vSAN 集群），<https://www.intel.cn/content/www/cn/zh/architecture-and-technology/deploying-vsan-cluster-best-practices-guide.html>

⁸ 见尾注 1。

⁹ 截至 2021 年 5 月 10 日英特尔所做测试。

英特尔® 傲腾™ 固态硬盘配置：4 节点，2x 英特尔® 至强® 金牌 6348 处理器（28 核，2.6 GHz），总内存 = 256 GB（16 插槽/16 GB/3200 MT/s），英特尔® 超线程技术 = 启用，英特尔® 睿频加速技术 = 启用，2x 英特尔® 傲腾™ 固态硬盘 P5800X（缓存）400 GB 和 8x 英特尔® 固态硬盘 D7-P5510 3.84 TB（容量），1x 英特尔® 以太网网络适配器 E810C 100 GbE，BIOS = 2.1 (ucode = 05003003)，VMware vSphere 7.0U2，HCI Bench 2.5.3。

英特尔® 固态硬盘 D7-5600 配置：4 节点，2x 英特尔® 至强® 金牌 6348 处理器（28 核，2.6 GHz），总内存 = 256 GB（16 插槽/16 GB/3200 MT/s），英特尔® 超线程技术 = 启用，英特尔® 睿频加速技术 = 启用，2x 英特尔® 固态硬盘 D7-P5600（缓存）1.6 TB 和 8x 英特尔® 固态硬盘 D7-P5510 3.84 TB（容量），1x 英特尔® 以太网网络适配器 E810C 100 GbE，BIOS = 2.1 (ucode = 0x8d055260)，VMware vSphere 7.0U2，HCI Bench 2.5.3。

¹⁰ 2021 年 11 月 1 日的当日标价。价格经常变化。具体成本和结果可能不同。

价格来源为 https://www.dell.com/cn-zh/work/shop/cty/pdp/spd/poweredge-750/pe_r750_14794_vi_vp?configurationid=e0d51b8e-5c01-40fd-9275-c20905029621。

¹¹ 见尾注 1。

¹² 见尾注 2。

¹³ 见尾注 2。

¹⁴ 见尾注 11。

¹⁵ 见尾注 11。

实际性能受使用情况、配置和其他因素的差异影响。更多信息详见 [intel.cn/performanceindex](https://www.intel.cn/performanceindex)。性能测试结果基于配置信息中显示的日期进行的测试，且可能并未反映所有公开可用的安全更新。详情请参阅配置信息披露。没有任何产品或组件是绝对安全的。具体成本和结果可能不同。英特尔技术可能需要启用硬件、软件或激活服务。© 英特尔公司版权所有。英特尔、英特尔标识以及其他英特尔商标是英特尔公司或其子公司的商标。其他的名称和品牌可能是其他所有者的资产。1221/GMCK/KC/PDF