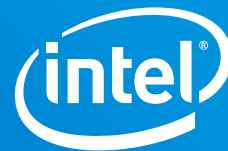


技术简介

数据中心内存和存储
英特尔® 傲腾™ 技术



为存储密集型工作负载 提供一致的低延迟

英特尔® 傲腾™ 固态硬盘通过高度可预测的访问延迟提供更可靠的性能。

Frank T. Hady 博士

英特尔院士

傲腾系统首席架构师

英特尔非易失性存储器 (NVMe)

解决方案事业部

我们都看过产品规格，这些产品规格虽然可以提供有用的信息，但无法完全告诉我们产品是否适合预期用途。现在，请打开你所用的固态硬盘 (SSD) 的产品规格，然后找到延迟参数。我钟爱的产品都可以在 intel.cn/content/www/cn/zh/products/memory-storage/solid-state-drives.html 找到。所谓访问延迟，是指固态硬盘提取系统请求的数据所需的时间。在固态硬盘规格中，你会看到一个典型延迟。典型延迟通常比较接近固态硬盘在最佳条件下可以接受写入数据或返回读取数据的最快时间。显然，我们并不能确保固态硬盘随时都能处于最佳条件。固态硬盘有时需要处理繁重的工作负载，或者处理后台工作，这都会导致延迟时间变长。本文所想探讨的正是这些延迟更长的情形。

我们把固态硬盘访问延迟的变化程度称为固态硬盘的服务质量 (QoS)。固态硬盘的 QoS 既指典型延迟，也指那些出现频率较低但时间更长的延迟。有时候，这些出现频率较低的延迟会非常长。我们可以通过一个类比来加深对 QoS 的理解。假设我们去一家咖啡店，当你走进门时，等待咖啡的延迟就开始了。如果咖啡店没有人，你可以马上点咖啡并快速拿到手——这就是咖啡店的“典型”延迟。但有时你需要排队，如果你前面有三个顾客，那么你可能就要多花三倍的时间。有时你甚至会发现已经有十多个顾客在排队为他们自己和朋友点多种饮品，而你甚至还没有进入排队区。虽然这样的情况不太可能发生，但你仍然会在意如果发生了该怎么办。如果这种情况发生不止一次，并且如果你和我一样没耐心，那么你能另找一家咖啡店。为了喝到咖啡，你需要考虑典型的等待时间、必须等待的时间以及预计等待的最长时间。你关心的是咖啡店的 QoS，而不仅仅是喝到咖啡所需的典型等待时间。

虽然固态硬盘比咖啡店要快得多，但固态硬盘的 QoS 对我来说也很重要。例如，如果你想评估特定固态硬盘的性能，那么仅仅查看平均延迟并不足以衡量这一固态硬盘进行时间关键型操作或执行需要满足服务级别协议 (SLA) 操作时所能提供的性能。即使数据请求延迟中只有一部分（甚至只有一小部分）是较长的异常值，它也可能会对您的应用性能产生重大影响并损害用户体验。你会发现你关心的是两个数字：比典型时间耗时更长的访问所占的比例，以及这些耗时更长的访问所花费的时间。这两个数字可以使用一种叫做百分位数 (percentile) 的统计工具来进行明确的描述。

内存和存储技术系列

构建您与英特尔院士
以及首席工程师的沟通桥梁

本文是关于内存和存储技术的系列文章之一，旨在帮助系统架构师、工程师和 IT 管理员理解传统内存和存储的技术局限、造成的数据中心性能和容量差距，以及英特尔® 傲腾™ 技术和英特尔® QLC 3D NAND 技术如何以全新架构填补这些差距。这一系列文章研究了影响存储性能和容量的几个因素，包括带宽、延迟、队列深度、服务质量 (QoS) 和可靠性。

使用百分位数更准确地测量固态硬盘延迟

通过百分位数，我们可以确定延迟小于特定值的访问所占的百分比。假设我们针对给定的工作负载运行固态硬盘，并记录延迟。然后，我们将这些延迟按由小（快）到大（慢）的顺序排列，并从左（小）到右（大）列出。特定延迟值在这个列表中的位置能告诉你该延迟发生的可能性，也可以将其有效地指定为百分位数。通过使用百分位数，我们可以避免列出每个可能延迟（因为那样可能会有太多值）的百分位数测量值，只需选择所关注的百分位数并指定延迟即可。第 50 个百分位数（在列表的中间）值得关注，它其实就是平均数。第 90 个百分位数（从列表右侧开始数的第十个值）、第 99 个百分位数（100 个百分位数中的第 99 个）、第 99.999 个百分位数（100,000 个百分位数中的第 99,999 个）都值得关注，且通常可在固态硬盘规格中找到相关数据。应用的应用性能很大程度上取决于特定的百分位数测量值（如下文所示）。

以下是一个有关如何使用 QoS 百分位数的示例。假设你负责一个网站托管系统，其中每个页面需要从存储中实时提取众多元素，包括客户专用数据、社交媒体帖子、库存的实时状态以及多个来源的广告。当用户访问网站时，将进行成百上千次固态硬盘访问以收集页面内容。加载该页面所需的时间取决于显示的最慢项，亦即耗时最长的延迟。在这种特定情景中，你需要考虑固态硬盘的 QoS。仅仅考虑典型甚至平均延迟是不够的。如果页面由 100 个项目组成，那么第 99 个百分位数的延迟就会是一个值得关注的数字，因为它提供了用户访问 100 个项目时平均需要等待的最长延迟。这个最长延迟决定了用户在看到完整页面之前需要等待的时间。完整的固态硬盘规格包括百分位数延迟测量值，有助于你了解预期可获得的 QoS。

本文将这些百分位数测量原理应用于英特尔® 傲腾™ 固态硬盘和 3D NAND 固态硬盘，从而全面描述它们的相对性能。这些测量值表明，对于重要的应用使用点，英特尔® 傲腾™ 固态硬盘提供的 QoS 显著优于 NAND 固态硬盘。

NAND 技术的 QoS 局限性

对于所有固态硬盘来说，从底层介质的行为开始就会对 QoS 产生影响了。对于 NAND 固态硬盘，典型延迟由介质的读取时间 ($T_{\text{读取}}$) 决定，因为它往往比固态硬盘其余部分和系统中的典型延迟更长（10 微秒级）。更重要的是，要将数据写入 NAND 固态硬盘，须先擦除大块，而且擦除速度很慢。尽管没有擦除慢，但写入的速度也很慢。由于擦除是在 NAND 上的大块中进行的，因此在擦除原始块之前，必须先将块中的良好数据移动到已经完成擦除的新块中。这是一个被称为垃圾回收的过程。想象一下，你去的咖啡店只有一名员工，既要煮咖啡，又要打扫卫生。在清理（垃圾回收）期间，你需要等待。如果清理工作包括将积累的垃圾带到垃圾填埋场并去商店购买更多杯子，则你需要等待更长的时间。

读取或写入延迟增加的可能性有很多。要成功完成访问，可能需要等待保存数据的 NAND 固态硬盘先完成之前的读取操作，并且可能会因对特定芯片的多次访问而滞后，甚至可能需要等待芯片先完成写入操作。这些读写可能来自于固态硬盘内部的垃圾回收操作。读取可能需要在完成已经开始的擦除操作之后再行进行。更糟糕的是，对固态硬盘的大量写入操作可能会促进垃圾回收，从而迫使这些写入操作需要等待许多其他操作完成才能进行。此外，由于固态硬盘同时还负责返回高度准确的数据，因此有时需要进行大幅纠错，这可能会需要进行多次读取操作。

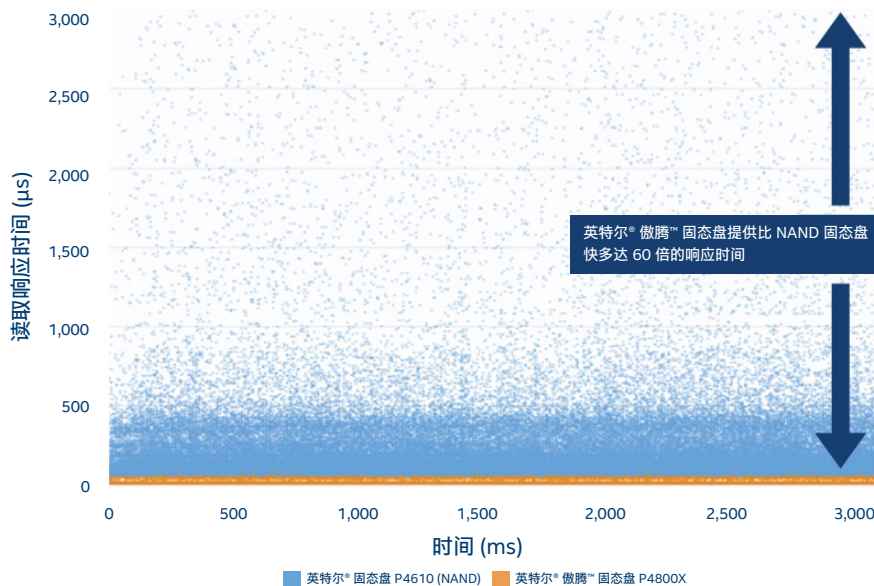


图 1. 在 QoS 的第 99 个百分位数，英特尔® 傲腾™ 固态硬盘提供比 NAND 固态硬盘快多达 60 倍的响应时间¹

NAND 是为低成本大容量存储而构建的技术，无法始终提供理想性能。看一下 NAND 固态盘的 QoS，这一点显而易见。所有 NAND 的固态硬盘都是如此。不同供应商的 NAND 固态硬盘会有不同的 QoS 特性，但所有 NAND 固态硬盘都显示了相对较宽的延迟分布范围。图 1 中的蓝点显示了这种分布的示例。此图以所示吞吐量绘制了固态硬盘进行 4 KB 70% 读取/30% 写入的响应表现，其中每个点的高度就是测得的延迟。值得注意的是，尽管不频繁，但这种高性能 NAND 固态硬盘延迟有时会高达 3 毫秒。我们将继续使用这一数据集，与英特尔® 傲腾™ 固态硬盘进行比较，由此展示描述 QoS 的更好方法。

英特尔® 傲腾™ 技术, 革新固态硬盘 QoS 的先进技术

英特尔® 傲腾™ 固态硬盘基于英特尔® 傲腾™ 内存介质这一全新内存技术而设计，具备与 NAND 固态硬盘完全不同的延迟级别和分布。这种新型内存的读/写延迟比 NAND 低很多——读/写速度要快几个数量级。它还支持无需先进行擦除，即可在适当的位置进行写入操作。此外，它支持字节寻址，这意味着仅需写入要修改的一个或多个扇区。不需要擦除大块，因此也不需要垃圾回收。因为无需进行擦除操作，因此英特尔® 傲腾™ 固态硬盘可提供高速读/写性能。所有这些优势让英特尔® 傲腾™ 固态硬盘具备比 NAND 固态硬盘优异几个数量级的 QoS。

使用英特尔® 傲腾™ 固态硬盘，就好比你去的咖啡店配备了有史以来最快的咖啡机、拥有多名员工随时准备为你服务，并且无需清理垃圾。而 NAND 固态硬盘就像是一家普通的咖啡店，你有时需要等待很长时间才能喝到咖啡。在图 1 中，橙色点表示英特尔® 傲腾™ 固态硬盘在运行上述 NAND 固态硬盘所运行的相同工作负载时测得的延迟。底部的橙色区域不是 x 轴，而是针对英特尔® 傲腾™ 固态硬盘

P4800X 记录的紧密聚集的读取延迟。图 1 中的蓝点显示的则是英特尔® 3D NAND 固态硬盘的各个读取响应时间。在处理随机写入工作负载时，英特尔® 傲腾™ 固态硬盘始终表现出低延迟和出色的 QoS，其第 99 个百分位数的读取响应时间比高耐用型 NAND 固态硬盘相同百分位数的读取响应时间快多达 60 倍。¹

图 1 清楚地显示了英特尔® 傲腾™ 固态硬盘的 QoS 优势，但是你却无法从这张图中读取特定百分位数的延迟，因此无法用它准确预测特定性能。为此，我们需要借助累积密度函数 (CDF) 来进行统计。图 2 显示了与图 1 相同的数据，但是使用了两个 CDF，每个分别对应一种固态硬盘类型。在这张图中，对于任何给定的百分位数 (x 轴)，都可以读取到预期延迟 (y 轴)。英特尔® 傲腾™ 固态硬盘的平均延迟约为 14 微秒 (μs)，而 NAND 固态硬盘约为 200 微秒。对于英特尔® 傲腾™ 固态硬盘，第 90 个百分位数的值仍为约 14 微秒，而 NAND 固态硬盘已经增至约 400 微秒。这一趋势一直延续到第 99.999 个百分位数。通过 CDF，我们可以非常明显地看出英特尔® 傲腾™ 固态硬盘巨大的 QoS 优势。

图 2 中的 CDF 很有用，但也存在限制，因为它仅显示了测试中单个每秒输入/输出 (I/O) 操作次数 (IOPS) 负载的相对 QoS。对于各类应用，需要考虑的不仅仅是不同百分位数的测量值，还有在不同负载下测量的 QoS。为了更好地解在不同负载下英特尔® 傲腾™ 内存介质相比 NAND 技术的 QoS 优势，图 3 针对内存而非存储应用情况绘制了延迟分布范围。图 3 显示了英特尔® NAND 固态硬盘和英特尔® 傲腾™ 固态硬盘的延迟与 IOPS 的变化百分位数 (例如第 99 位和第 99.9 位) 的一系列曲线。每个负载将生成不同的雪花图和不同的 CDF。使用线性标度时，英特尔® 傲腾™ 固态硬盘的曲线将聚集在 y 轴底部附近，因此为了清晰起见，图 3 使用对数标度。

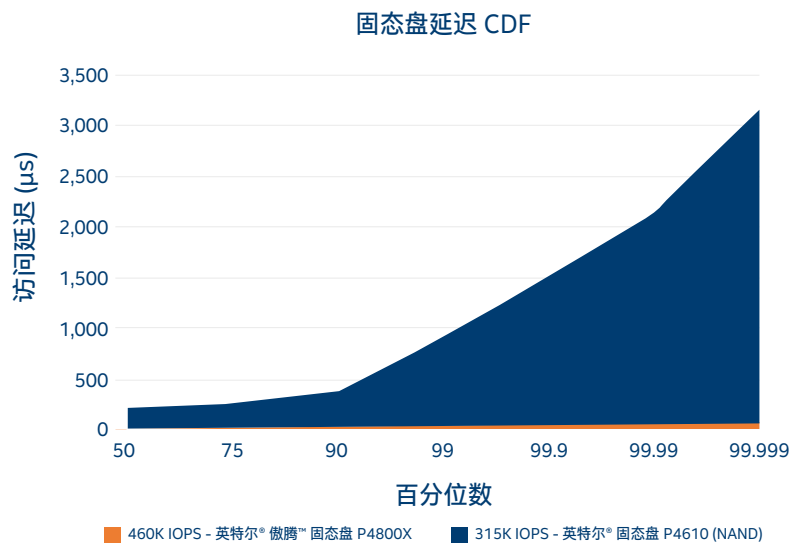


图 2. CDF 显示，英特尔® 傲腾™ 固态硬盘 P4800X 的平均延迟远低于英特尔® 固态硬盘 P4610 (NAND)

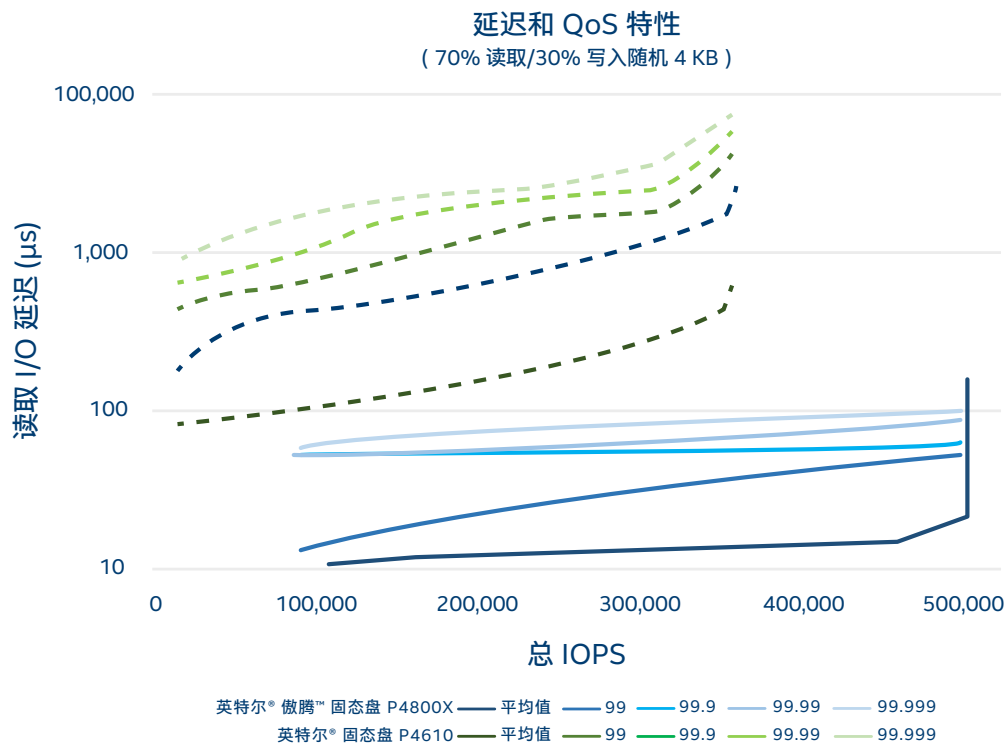


图 3. 英特尔® 傲腾™ 固态硬盘与英特尔® NAND 固态硬盘的延迟分布对比¹

图 3 对了解和比较固态硬盘的 QoS 非常有用。对于每个固态硬盘 (英特尔® NAND 固态硬盘为绿色, 英特尔® 傲腾™ 固态硬盘为蓝色), 你可以看到与不同百分位数或 QoS 水平相对应的一系列曲线。你现在可以根据数据中心特定应用的需求, 查看代表所需保证延迟级别的曲线。例如, 如果你正在编写一个需要支持 300,000 IOPS 的应用, 并且你关心第 99.9 位百分位数 (这一数值代表了 1,000 次访问中最慢的一次), 那么在曲线中可以看到 NAND 固态硬盘的延迟大约是 1,500 微秒 (1.5 毫秒)。这与英特尔® 固态硬盘 P4610 数据表中标明的 77 微秒平均读取延迟相差甚远。²

如果你在图 3 中查看英特尔® 傲腾™ 固态硬盘在相同 300,000 IOPS 负载下的第 99.9 位百分位数, 则会发现它的延迟少于 80 微秒。因此, 对于 300,000 读/写 IOPS, 我们可以发现英特尔® 傲腾™ 固态硬盘拥有几乎高达 20 倍的 QoS 优势。得益于独特的内存技术和设计, 即使 IOPS 增加, 英特尔® 傲腾™ 技术仍可保持一致的低延迟和更高的 QoS。这一优势对于现代企业至关重要, 因为它能够为大型数据集和延迟敏感型工作负载提供一致且出色的性能。

从图中可以明显看出, 英特尔® 傲腾™ 固态硬盘可为任何工作负载提供巨大的 QoS 优势; 并且工作负载越密集以及百分位数越高, 优势也越大。同样明显的是, 与 NAND 固态硬盘相比, 英特尔® 傲腾™ 固态硬盘提供更加一致的性能: 随着负载的增加, 延迟的变化率更小。这种更为理想的响应能力让程序员可以更轻松地使用英特尔® 傲腾™ 固态硬盘, 从而让他们无需花费太多精力进行负载限制亦可

英特尔院士 Frank Hady

Frank Hady 是英特尔院士兼英特尔非易失性存储器解决方案事业部 (NSG) 傲腾系统首席架构师。Frank 负责带领团队研究和定义英特尔® 傲腾™ 技术产品及其与计算系统的集成。他曾担任英特尔首席平台 I/O 架构师, 为英特尔® QuickAssist 技术奠定了研究基础, 推动了平台性能的显著提升。他撰写或联合撰写了 30 多篇关于网络、存储和 I/O 创新主题的论文, 并经常发表关于内存和存储的文章, 拥有 30 多项美国专利。Frank 获得了弗吉尼亚大学的电子工程学士和硕士学位, 并拥有马里兰大学的电子工程博士学位。

将延迟水平维持在可接受水平。将 NAND 固态硬盘用于需满足 SLA 要求的工作负载时, 这种场景是真实存在的。无论你的应用需要考虑哪个 QoS 百分位数, 英特尔® 傲腾™ 固态硬盘始终能够提供比 NAND 固态硬盘低得多的延迟, 即使在高负载下也是如此。更低的延迟可以直接在业务中转化为更快、更加可预测的应用性能, 并帮助你更好地满足 SLA 要求。

英特尔® 傲腾™ 固态硬盘助益实际用例

在与客户合作的过程中, 我们发现英特尔® 傲腾™ 固态硬盘为许多不同的应用提供了差异化的解决方案级优势。英特尔® 傲腾™ 固态硬盘一致且强大的 QoS 和高耐用性使这些优势得以实现。

得益于出色的 QoS，英特尔® 傲腾™ 固态硬盘可以高效地提供数据缓存或分层存储。将访问频率较高的数据存储在英特尔® 傲腾™ 固态硬盘中，将访问频率较低的数据存储在容量更大的 NAND 固态硬盘中。通常，将 10% 的数据存储在高速英特尔® 傲腾™ 固态硬盘层中，该层可响应 90% 的总数据访问。其余 10% 的访问将会引用存储在 NAND 层中的另 90% 的较冷数据。此外，英特尔® 傲腾™ 固态硬盘能够以更高的速率提供 IOPS，并且能够始终保持较低的延迟。这样的低延迟意味着，使用仅占 10% 容量的英特尔® 傲腾™ 固态硬盘即可保持 90% 的强劲吞吐率和 QoS 质量。

文件系统、数据库和高可靠性分类存储都采用双重写入策略，以确保不会产生写入丢失。第一次写入的日志文件要比实际数据库小得多，第二次写入才会写入数据库本身。通过使用这种方法，万一断电只会中断一次写入操作，从而确保可在电力恢复后能正确地重构数据库事务。这意味着数据库写入操作需要完成两次操作才能完成，而其中一次写入操作（日志写入操作）将会存储到一个较小的存储库中。这就要求存储日志的固态硬盘具有更高的 IOPS 速率。英特尔® 傲腾™ 固态硬盘在高 IOPS 负载下具有出色的 QoS，因而是日志数据存储的理想选择。使用

英特尔® 傲腾™ 固态硬盘后，写入日志文件不会降低数据库的提交速度。

如上面的例子所示，当在系统中的正确位置使用英特尔® 傲腾™ 固态硬盘后，就可以在高负载下提供高 QoS，从而在重要的实际用例中带来巨大的收益。

最大延迟值同样是有用的数据

如果你想要评估应用的实际性能，平均延迟只是一个起点。读取响应时间的偶发性峰值可能关系到性能是否令人满意。

NAND 固态硬盘具有低成本、高容量的优点，与机械硬盘 (HDD) 相比具有显著优势。它能够以高性价比的容量存储访问频率不高的数据，因此在数据中心中仍发挥着关键作用。但是对于需要出色 QoS 的时间敏感型操作，英特尔® 傲腾™ 固态硬盘提供了更优的解决方案。

英特尔® 傲腾™ 技术提供了一种创新型方法，通过为密集型固态硬盘工作负载提供严格控制且一致的低延迟，从而在数据中心中实现一致的 QoS。

了解更多信息

如需进一步了解英特尔® 傲腾™ 技术如何颠覆数据中心内存和存储层次，请查看内存和存储技术系列的其他文章。

如需了解英特尔® 傲腾™ 持久内存，请访问：intel.cn/content/www/cn/zh/architecture-and-technology/optane-dc-persistent-memory.html

如需了解英特尔® 傲腾™ 固态硬盘，请访问：intel.cn/content/www/cn/zh/products/memory-storage/solid-state-drives/data-center-ssds/optane-dc-ssd-series.html



¹ 基于英特尔 2018 年 11 月 15 日进行的测试：响应时间是指在 4K 随机写入工作负载期间，在队列深度 1 使用 FIO 3.1 所测得的平均读取延迟。配置：低队列深度 (QD) 下的 4K 70/30 读/写性能。通过 FIO 3.1 测得。通用配置：英特尔® 2U 服务器系统，CentOS 7.5，内核 4.17.6-1.el7.x86_64，2 个英特尔® 至强® 金牌 6154 处理器 (3.0 GHz，18 个内核)，256 GB DDR4 RAM (2,666 MHz)。配置：375 GB 英特尔® 傲腾™ 固态硬盘 P4800X 与 1.6 TB 英特尔® 固态硬盘 P4610 对比。英特尔® 微代码：0x2000043；系统 BIOS：00.01.0013；英特尔® 管理引擎 (英特尔® ME) 固件：04.00.04.294；基板管理控制器 (BMC) 固件：1.43.91f76955；FRUSDR：1.43。由于要进行额外的测试，以上基准测试结果可能需要修改。

² 英特尔公司，“英特尔® 固态硬盘 P4610 系列”，<https://ark.intel.com/content/www/cn/zh/ark/products/140103/intel-ssd-dc-p4610-series-1-6tb-2-5in-pcie-3-1-x4-3d2-tlc.html>。

性能测试结果基于配置中所示日期进行的测试，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。

性能测试中使用的软件和工作负载可能仅在英特尔微处理器上进行了性能优化。诸如 SYSmark 和 MobileMark 等测试均系基于特定计算机系统、硬件、软件、操作系统及功能。上述任何要素的变动都有可能测试结果的变化。请参考其他信息及性能测试 (包括结合其他产品使用时的运行性能) 以对目标产品进行全面评估。更多信息，详见 intel.cn/benchmarks。

英特尔技术可能需要支持的硬件、软件或服务得以激活。请从原始设备制造商或零售商处获得更多信息。

© 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在美国和/或其他国家的商标。其他的名称和品牌可能是其他所有者的资产。