

应用软件运行速度提升 16 倍： Lightbits 与 Ceph 的存储性能对比

作者：

Russ Fellows

Mohammad J Rabin

2022 年 3 月



Evaluator Group



lightbits

intel®

概述

云计算和基于容器的应用是推动当今 IT 领域产生重大变革的两大趋势。与其他技术和方法相比，二者都能够以更高的灵活性和更低的成本去运行和操作应用，因此日益受到欢迎。然而，在所有新的机遇面前，要实现潜在优势也都会有巨大的挑战。

层出不穷的新技术通常会对业务运营产生渐进式的影响。但是，以云计算作为运行模式，结合管理云原生应用的 **Kubernetes**——这一做法已经改变了信息技术的交付、管理和使用方式。

此外，为支持这些环境而设计的系统需要选择计算、网络，尤其是存储系统，这些系统具备高度可扩展的容量和性能，而无需采用传统管理方式。**Kubernetes** 提供机制，帮助应用以透明的方式利用多种计算和存储资源。但是，要确保应用对持久数据的访问，需要底层存储基础架构具有弹性，同时还能满足成千上万个应用的容量和性能需求。

在此背景下，**Evaluator Group** 获邀对适合云原生应用环境的两种存储系统展开对比。

- **Ceph**: 云原生环境中常用的开源软件定义存储
- **Lightbits** 云数据平台: 专为云环境设计的软件定义存储

如下所示，在使用相同硬件和存储介质的情况下，**Lightbits** 在所有工作负载测试中，实现的性能均远远优于 **Ceph**。表 1 汇总显示了使用 QLC 介质的 NVMe 固态硬盘 (SSD) 的测试结果。该测试使用在 OpenShift Kubernetes 环境中作为容器运行的工作负载来进行。

	Ceph IOPs	Lightbits IOPs	Lightbits Advantage
4K-100% Read	1,032,428	4,068,462	3.94 X
4K-100% Write	30,728	515,697	16.78 X
8K-80% Read	90,363	1,129,335	12.50 X
16K-70% Read	35,740	372,009	10.41 X
32K-50% Read	19,797	111,852	5.65 X

表 1: 在容器环境中使用 QLC 介质时的 **Lightbits** 与 **Ceph** I/O 速率比较。 (来源: **Evaluator Group**)

本文的其余部分将更深入地探讨二者的技术和性能差异，以及使用 NVMe SSD 与 TLC 介质时的性能数据。

新兴技术

如前所述，多个趋势正在推动 IT 领域发生重大变化；其中，利用公共云等服务的趋势正在改变 IT 运行的地点和付费方式，影响了应用程序的运行方式。云原生应用程序（CNA）的推行正在改变应用程序的开发和管理方式，这在一定程度上得益于企业向服务和云端迁移的趋势。此外，可通

过 Fabric 去访问 NVMe 存储的新技术，以及新的非易失性存储技术的出现，使得那些能够利用这些功能的系统拥有显著的性能优势。

云原生应用程序

基于容器的应用程序设计旨在运行于云环境，包括本地云和公共云。云原生应用程序被设计为可独立扩展的服务，又称为微服务。借助微服务架构，CNA 能够在少量或甚至无需管理介入的情况下，按需扩展应用程序的各个部分。

此外，CNA 部署在一个位置后，可以根据可用资源甚至根据资源的价格而移动到其他位置。另外，许多此类应用程序所需的持久性存储也必须动态响应不断变化的容量和性能需求。

NVMe over Fabrics 协议

非易失性存储主机控制器接口规范（NVMe）是一个支持固态硬盘（SSD）去使用高速 PCIe 总线进行通信的接口，以实现对固态设备的低时延访问。与 SCSI 或 SATA 设备相比，NVMe 支持数十亿条命令，显著减少了延迟或时延。为了兼容与外部设备的通信，NVMe over Fabrics 扩展了 NVMe 协议，支持通过 Fabric 去访问 NVMe。

NVMe over Fabric 访问的最新选择之一是使用 TCP 传输协议，即 NVMe/TCP。实施这项协议的技术负担可显著减少，无需任何特定的硬件或交换机设置即可支持商用以太网。总体而言，与 TCP 上的其他存储协议相比，在仍然使用现有网络基础设施的情况下，NVMe over TCP 具有更低的时延和更高的吞吐量。

持久性存储

持久性存储（PMem）兼具存储特性（特别是无需能耗即可保留或“持久保存”数据的能力），以及内存总线访问特性（例如极低的时延和字节寻址能力）。持久性存储有两种主要类型，分别是 NVDIMM 和英特尔傲腾持久性存储（Optane Persistent Memory）。NVDIMM 和傲腾 PMem 之间的主要区别在于存储密度和成本，其中，与 NVDIMM 设备相比，英特尔傲腾 PMem 持久性存储在这两个方面都具有显著优势。

英特尔 PMem 既可用于“应用程序直接访问模式”，也可用于“内存模式”。前者需要修改应用程序，后者以低于 DRAM 的成本、透明地扩展系统内存容量。Lightbits 在 App Direct 模式下使用 PMem 作为非易失性写入缓冲区，将写入和元数据访问直接定向至 PMem，减少了备份介质的磨损，改善了系统写入性能。

解决方案概述

过去几十年里，使用商用硬件、以“软件或软件定义存储（SDS）”的形式运行的存储系统日益受到追捧。这种方法之所以引人入胜，是因为它能够在不可使用专用或专有硬件的环境中运行弹性的数据中心级存储。此外，其灵活性有利于在不同的环境和部署规模中使用相同的 SDS 堆栈。

如前所述，我们的评估中就使用两种不同的 SDS 选项，为容器应用提供持久存储。随后将概述它们的功能。采用英特尔技术的 Lightbits 解决方案如图 1 所示。

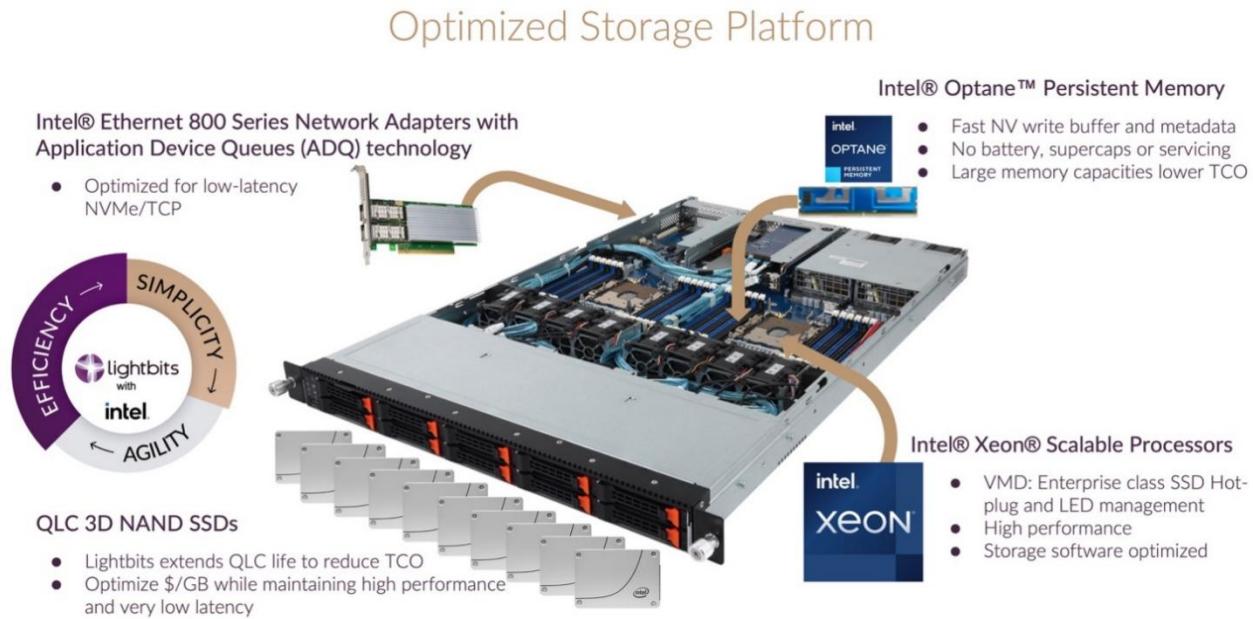


图 1：用于 Lightbits 软件定义存储的英特尔平台（来源：英特尔）

英特尔

英特尔为当今许多领先的服务器和存储解决方案提供基础平台组件。这些组件包括英特尔至强处理器、英特尔网络适配器和英特尔傲腾™技术等，其中，英特尔傲腾™可作为 NVMe 固态硬盘，或内存形式的傲腾持久性存储（PMem）设备提供。此外，具有应用设备队列（ADQ）的英特尔 Ethernet 800 系列提供了高性能、低时延的 NVMe/TCP 网络接口卡。

Lightbits 云数据平台

Lightbits 云数据平台是一种软件定义存储，旨在运用最新的技术，包括 NAND 闪存和英特尔傲腾持久性存储（PMem）；它使用高速 100GbE 网络接口卡，支持 NVMe over TCP。Lightbits 使用分布式架构，能够按需独立扩展 CPU、内存、PMem 或 NVMe 设备。因此，Lightbits 能够提供可扩展的企业级存储，性能优于本地 NVMe 设备。

Lightbits 智能闪存管理解决方案 Intelligent Flash Management™ (IFM) 支持最大限度地提高 SSD 性能和延长其耐用性。其主要特性包括写条带化、无 IOP 元数据访问、智能垃圾收集、追加写策略和并行读/写通道。

通过将英特尔的高性能硬件平台与 Lightbits 方案进行耦合，可提供可扩展的解决方案，降低总体拥有成本 (TCO)。Lightbits 解决方案特别采用了以下技术：

- 英特尔®至强®处理器系统
 - 高性能 CPU、内存、I/O 和 PCIe Gen4
 - 使用高性能 CPU 架构的高效存储软件
 - 企业级 SSD 热插拔和 LED 支持 NVMe 介质
- 英特尔傲腾持久性存储 (PMem)
 - 低时延、非易失性写缓冲区和元数据存储
 - 无需电池或电容器即可维持持久性
 - 大容量持久性存储，总拥有成本低于其他替代方案
- 英特尔 Ethernet 800 系列网络适配器
 - 提供应用设备队列 (ADQ)
 - 高性能应用设备队列
 - 低时延 NVMe/TCP，无需使用定制的驱动程序

通过 Kubernetes 容器存储接口 (CSI)，Kubernetes 管理的容器可以作为持久卷 (PV) 访问 Lightbits 存储。Lightbits CSI 插件使 Kubernetes 能够使用其“Elastic RAID”在 Lightbits 集群中存储持久卷 (PV)，该“Elastic RAID”提供纠删码保护，防止每个节点的数据丢失。其他 Lightbits 服务包括压缩、复制、卷快照、克隆和用于多租户环境中的基于角色的访问控制 (RBAC)。

Ceph

Ceph 是一款开源的软件定义存储产品，带有多个不同的接口，可作为块设备、文件系统和目标存储进行访问。设计 Ceph 时，采用旋转硬盘驱动器作为主要的存储介质，基于 NAND 的 SSD 在大多数主流存储系统中不发挥作用。

自 Ceph 设计以来的 15 年间，各种外形规格的固态闪存已成为高性能存储系统的主要介质。但 Ceph 初始设计中旨在提高旋转物理介质性能的许多元素仍得以保留。最近，更新后的 Ceph 已经可以使用固态的“BlueStore”和“RocksDB”进行元数据操作。迄今为止，Ceph 还没有针对持久性存储进行优化，也不支持 NVMe over fabrics 或 NVMe over TCP。

我们在测试中使用 Ceph 的开源版本，详见附录。

存储效率

与一些软件定义存储系统相比，Lightbits 的弹性设计拥有更多优势。Lightbits 针对每个节点使用纠删码保护，确保在丢失存储设备时的数据可用性；此外，节点之间可进行复制，因此，Lightbits 能够在减少开销和存储容量浪费的同时，提供更大的弹性。在测试期间，Lightbits 的配置是：在每个节点上使用纠删码来保护数据，同时在 Lightbits 集群中的另一个节点上额外制作一份数据拷贝。

与此相反, **Ceph** 的配置是: 创建三个数据副本来进行数据保护, 此时, 如果一个节点的驱动器和另一个节点出现故障断电, 则无法确保数据可用性。尽管这两种保护方法的存储效率相似, 但 **Lightbits** 方法的可用性更高。

不同容器工作负载下的性能

Evaluator Group 需对 **Kubernetes** 应用环境中作为容器运行的两种不同的存储产品进行分析。我们对 **Lightbits** 软件定义存储与面向容器的开源软件定义存储产品 **Ceph** 进行了比较。测试环境由 12 个 **Kubernetes** 节点组成, 在专用的 3 节点 **Lightbits** 集群或 3 节点 **Ceph** 集群中运行容器工作负载。容器应用测试图表和其他细节、硬件、软件和其他方面的详细信息, 请参见附录描述。

测量存储性能时, 使用众所周知的“**vdbench**”工具创建工作负载, 每个节点使用 8 个运行 **vdbench** 的容器实例, 总共 96 个 **vdbench** 实例。

测试包括性能敏感型应用中常见的 5 种不同的访问模式和块大小:

- 4KB, 100% 读取, 100% 随机访问
- 4KB, 100% 写入, 100% 随机访问
- 8KB, 80% 读取/20% 写入, 80% 随机访问
- 16KB, 70% 读取/30% 写入, 80% 随机访问
- 32KB, 50% 读取/50% 写入, 80% 随机访问

这些工作负载用于比较 **Lightbits** 与 **Ceph** 的性能, 使用 QLC 固态介质作为主要存储介质, 并根据每个存储系统使用适当的高速持久介质。在所有情况下, 分别多次运行 5 个工作负载, 获得可用于比较的平均结果。有关 **Lightbits** 和 **Ceph** 配置的具体细节, 请参见附录。

Evaluator Group 评论: 测试结果显示, 使用 **Lightbits** 具有巨大的性能优势, 其中, 4KB 写入工作负载的性能是 **Ceph** 的 16 倍。考虑到在所有比较中, 底层介质的数量和类型是相同的, 这种优势尤为明显。

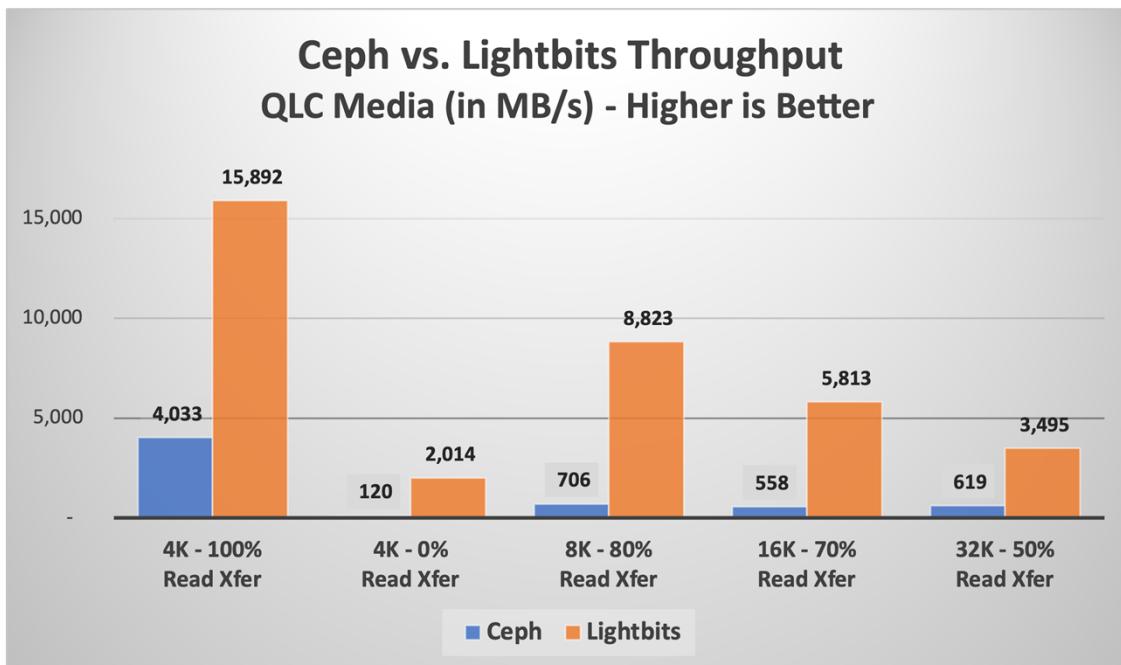


图 2: Ceph 与 Lightbits 吞吐量比较 (来源: Evaluator Group)

图 2 显示了使用 QLC 介质时, Lightbits 和 Ceph 在 5 种不同访问模式下的吞吐量结果。吞吐量结果是直接可比较的, 因为它们考虑了数据块大小的差异, 因此将其全部显示在一起。

性能详情

I/O 速率 (以每秒 I/O 传输次数或 IOPs 衡量) 是小数据块工作负载的常用指标, 但不利于对比不同的数据块大小。如下方图 3 所示, 对典型应用大小——8KB、16KB 和 32 KB 进行比较, 绘制 I/O 速率, 以每秒 I/O 传输次数衡量。在 32KB 时, Lightbits 的优势是 5.7 倍, 在 16KB 时, 优势是 10.4 倍, 在 8K 工作负载时, Lightbits 的性能超过 Ceph 12 倍。

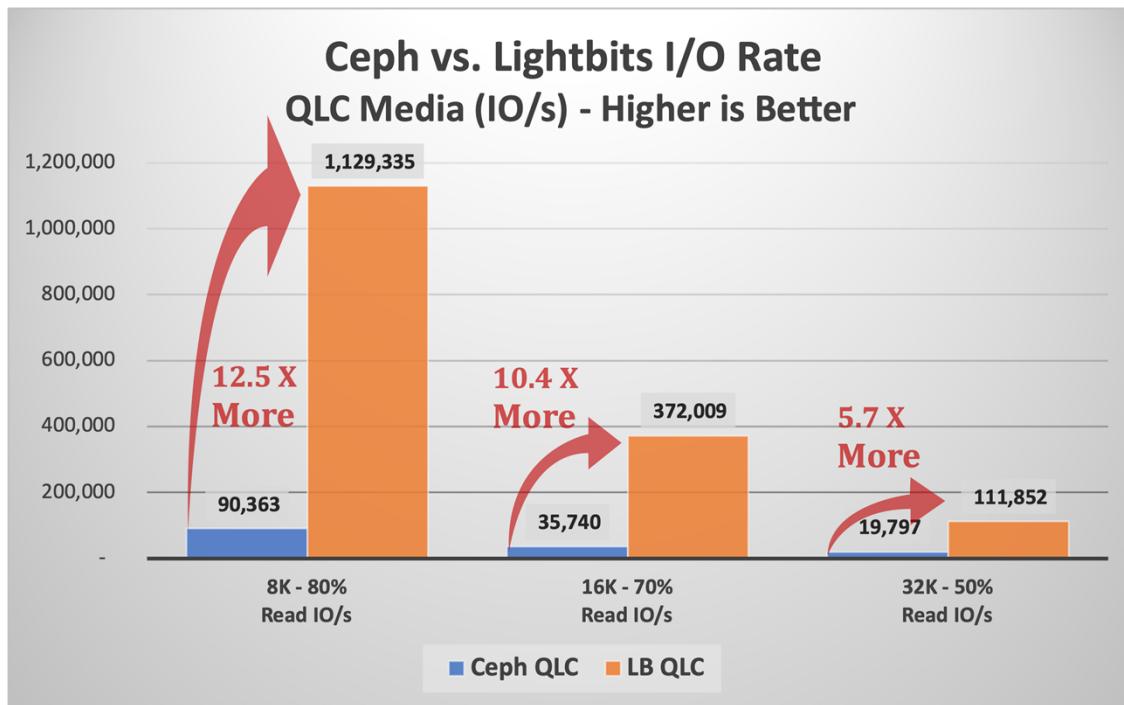


图 3: 使用 QLC 介质时 Lightbits 与 Ceph 的性能对比 - 不同数据块大小的吞吐量 (来源: Evaluator Group)

上文显示的三种工作负载是数据库或其他事务应用程序中的常见负载。尽管许多数据库尝试执行 16K 或 32K 大小的 I/O，但在许多情况下，如果更改较小或事务率较高，它们倾向于使用较小数据块。数据库将执行 8K 甚至 4K 事务，而非等待合并几个较小的 I/O。

以下显示 Ceph 和 Lightbits 使用两种不同类型的存储介质，即 QLC 和 TLC 介质时，4KB 读取操作的 I/O 速率比较。

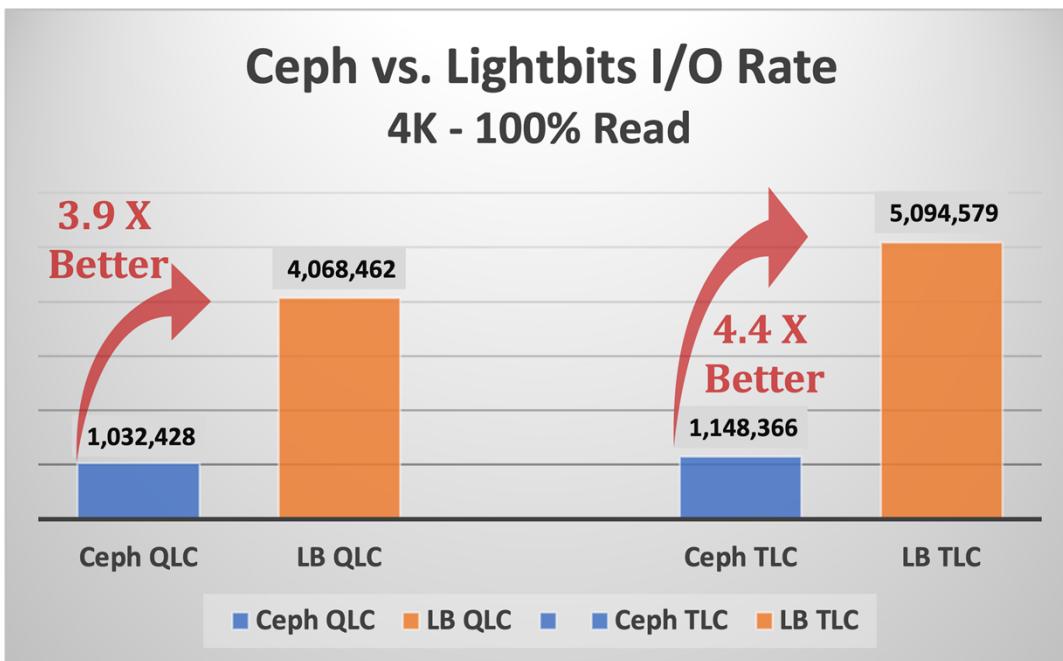


图 4: 使用 QLC 和 TLC 介质时, Ceph 与 Lightbits 的 4KB 读取 I/O 速率比较 (来源: Evaluator Group)

虽然会发生 4KB 读取操作, 但 4KB 写入操作更经常用于数据库日志记录操作。使用 QLC 介质时, Lightbits 的性能比 Ceph 高出近 17 倍, 使用 TLC 介质时, 性能比 Ceph 高出近 7 倍, 如下图 5 所示。

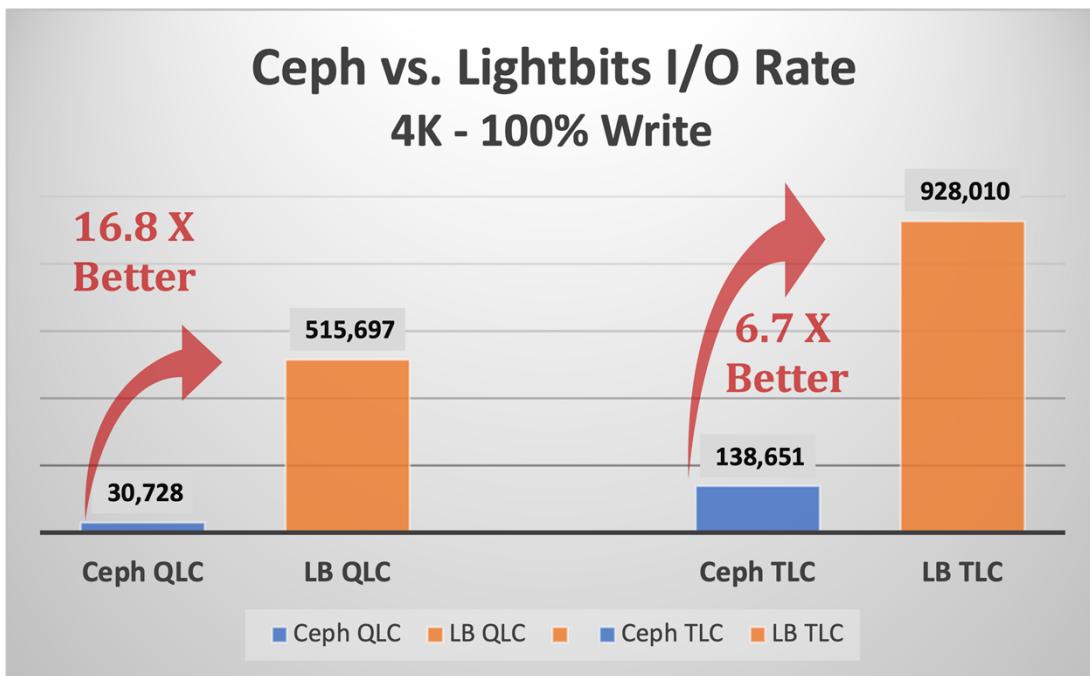


图 5：使用 QLC 和 TLC 介质时，Ceph 与 Lightbits 的 4KB 写入 I/O 速率比较（来源：**Evaluator Group**）

Evaluator Group 评论： 数据库和其他事务应用程序依赖于低时延存储，尤其是事务日志，后者必须在事务完成前写入持久介质。日志大多通过 4K I/O 操作写入，在此工作负载下，Lightbits 显示出比 Ceph 更大的优势：使用 TLC 介质时，优势高出近 7 倍，使用 QLC 时，优势高出近 17 倍。

其他性能数据

下文为性能结果概述，其中指出三个主要指标，包括吞吐量、I/O 速率以及时延。Lightbits 相对于 Ceph 的性能优势极其显著，因此用“x”倍数来体现这种优势。使用简单的除法（即 $515,697 / 30,728 = 16.78$ x）来显示 Lightbits 性能优势倍数。要点如下：

- 在所有五个工作负载下的测试结果表明，Lightbits 具有 4 到 16 倍的优势
- 数据库经常使用对时延高度敏感的小块写入日志或日记
- 在 4KB 写入下，Lightbits 的性能优势为 16.78 倍，优势显著

	QLC Media Comparison			TLC Media Comparison		
	Ceph QLC	Lightbits QLC	LB- X better	Ceph TLC	Lightbits TLC	LB- X better
4K-100% Read MB/s	4,033	15,892	3.94	4,486	19,901	4.44
4K-0% Read MB/s	120	2,014	16.78	542	3,625	6.69
8K-80% Read MB/s	706	8,823	12.50	3,542	16,219	4.58
16K-70% Read MB/s	558	5,813	10.41	4,959	15,449	3.12
32K-50% Read MB/s	619	3,495	5.65	4,802	9,837	2.05
4K-100% Read IO/s	1,032,428	4,068,462	3.94	1,148,366	5,094,579	4.44
4K-0% Read IO/s	30,728	515,697	16.78	138,651	928,010	6.69
8K-80% Read IO/s	90,363	1,129,335	12.50	453,419	2,076,046	4.58
16K-70% Read IO/s	35,740	372,009	10.41	317,367	988,724	3.12
32K-50% Read IO/s	19,797	111,852	5.65	153,677	314,774	2.05
4K-100% Read Lat. (ms)	1.5	0.4	3.95	1.3	0.3	4.46
4K-0% Read Lat. (ms)	19.8	1.1	17.27	4.1	0.6	6.72
8K-80% Read Lat. (ms)	22.6	1.4	16.34	3.4	0.7	4.59
16K-70% Read Lat. (ms)	22.3	2.2	10.27	2.4	0.8	3.12
32K-50% Read Lat. (ms)	11.5	1.9	6.19	1.2	0.6	2.06

表 3: Lightbits 与 Ceph 性能比较 (来源: Evaluator Group)

关于测试，还有一点值得留意，即在使用 QLC 介质时，Ceph 的时延高于大多数工作负载可接受的水平。这是因为，我们在所有测试中使用相同的队列深度设置，确保比较结果的公平公正。如果工作负载针对 Ceph QLC 进行了优化，将使用较低的队列深度设置，从而降低 I/O 速率，同时减小时延。如果以这种方式进行测试，Ceph 的时延水平更易于接受，但性能会显著降低，从而放大 Lightbits 的优势。

最终结论

性能始终是 IT 基础设施的重要考虑因素，尤其是在运行云环境或托管云原生应用程序时。尽管并非每个应用或微服务都需要高 I/O 速率，但如要运行现代云基础设施，提供具有高 I/O 速率和低时延的弹性、可扩展存储至关重要。

Evaluator Group 评论：在使用相同硬件配置的情况下，**Lightbits** 的性能显著优于 **Ceph**，弹性也更高。**Lightbits** 针对英特尔的高性能技术进行优化，为支持云的 **Kubernetes** 环境提供了强大的存储平台。

近期技术进展层出不穷，助力 IT 用户去大规模、经济高效地运营基础设施。当与英特尔傲腾持久性存储结合使用时，QLC 等新型存储介质可以为要求不高的工作负载支持海量存储容量，并提供良好的性能。包括新处理器和端到端 NVMe 在内的其他技术也可以提供性能和成本优势。

英特尔的专有技术，如 Xeon CPU、支持 NVMe over TCP 的英特尔网卡，以及傲腾持久性存储和高密度 NVMe SSD，为软件定义产品提供了强大的硬件基础。由于高效利用了英特尔 CPU、网络接口和持久存储介质，**Lightbits** 存储可提供本地 NVMe 级别的性能，从而持续确保低时延和数据服务。

本文中所示的性能测试使用了 IT 环境和应用中常见的工具和工作负载。附录提供了进一步的详细信息，如希望进一步了解测试，或希望使用自己的硬件配置重复这些测试，可参考附录。

Evaluator Group 评论：借助存储专业人员熟知的测试工具和工作负载，我们认为，**Lightbits** 可以提升现代应用程序的性能。**Lightbits** 可以通过 QLC 介质和傲腾 PMem，支持时延敏感型应用，而 **Ceph** 在明显具有更高延迟的情况下只能提供很低的 I/O 速率。

利用持久性存储和端到端 NVMe 等新技术，**Lightbits** 即可提供比 **Ceph** 更优异的性能，且弹性更高。显而易见，**Lightbits** 的性能是 **Ceph** 的 16 倍，时延更低，是适合现代应用和云环境的更佳选择。

附录

测试环境详情

测试环境利用以下硬件、软件和应用工作负载项目。

测试持续 4 个月，从 2021 年 9 月起至 12 月止。

硬件和基础设施

- 关于测试配置的概述，请参见下方图 6
 - 在裸机上运行的应用集群，由 12 个 Kubernetes “worker” 节点组成
 - 3 节点管理集群/作为虚拟机运行的 Kubernetes “master” 节点
 - 运行软件定义存储堆栈的 3 节点存储集群
- 应用环境和编排使用 OpenShift 4.6 版

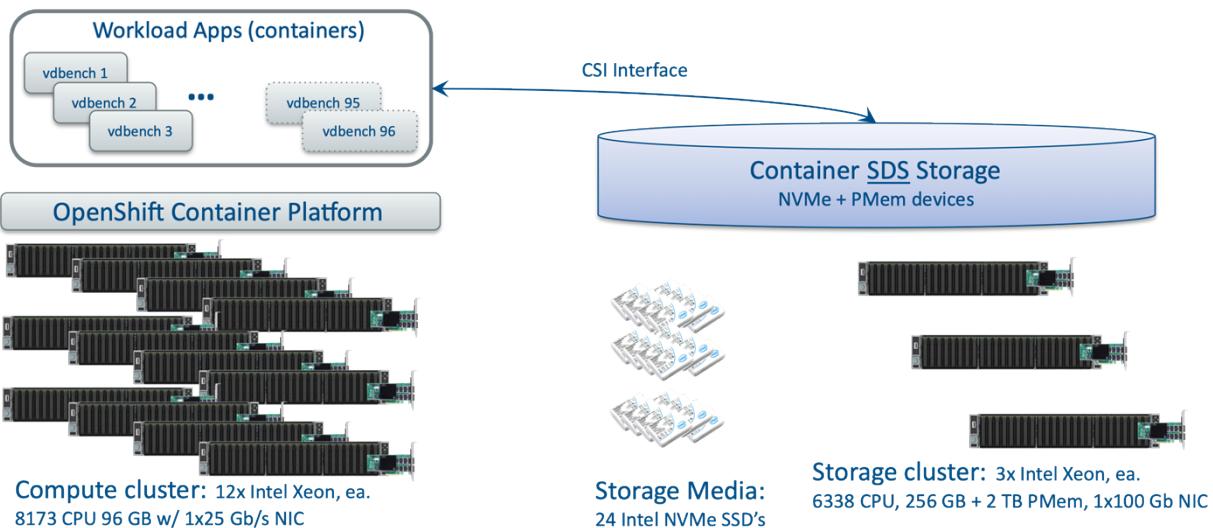


图 6: OpenShift 访问存储节点的工作负载测试设置 (来源: Evaluator Group)

软件环境

- 使用 Red Hat OpenShift 4.6 版作为 3 个管理节点和 12 个应用节点
 - 12 个应用 (OpenShift “worker”) 节点直接运行于硬件上，因此称为“裸机”
 - 3 个管理 (OpenShift “master”) 节点作为虚拟机运行，但作为“裸机”安装
- CentOS 8.4 版是 3 个 Ceph 节点使用的基础操作系统
- Ceph SDS 堆栈使用开源 Ceph 16.2.6 版 (又名“Pacific”)
- VMware 7.0; 3 台虚拟机使用 ESXi 及 vCenter

SUT 配置

- 一个 12 节点应用集群，运行 Red Hat OpenShift
 - 应用/工作节点配置:
 - 双插槽第一代英特尔至强可扩展系统 (Intel Xeon Platinum 8173M)
 - 96 GB 内存 (12 个 8 GB DIMM)
 - 英特尔 XXV710-DA2 网卡，每台主机带 2 个 25 Gb/s 连接
- 两个存储系统连接到 OpenShift 应用集群
 - SDS #1, Ceph 配置:

- 双插槽第三代英特尔至强可扩展系统 (Intel Xeon Gold 6338)
- 256 GB 内存 (16 个 16 GB DIMM)
- 2 TB 的英特尔傲腾 PMem (配置为 WAL 和 Cache 设备)
- 英特尔 E810-CQDA2 网卡, 每台主机带 1 个 100 Gb/s 连接
- 8 个英特尔 NVMe 存储介质, 可选择
 - QLC SSD 介质: 英特尔 SSD D5-P5316 @ 15.36 TB
 - TLC SSD 介质: 英特尔 SSD D7-P5510 @ 3.84 TB
- SDS #2, Lightbits 配置:
 - 双插槽第三代英特尔至强可扩展系统 (Intel Xeon Gold 6338)
 - 256b GB 内存 (16 个 16 GB DIMM)
 - 2 TB 英特尔傲腾持久性存储 Persistent Memory 200 (配置为 Lightbits 写入缓冲区)
 - 英特尔 E810-CQDA2 网卡, 每台主机带 1 个 100 Gb/s 连接
 - 8 台英特尔 NVMe 存储设备, 可选择:
 - QLC SSD 介质: 英特尔 SSD D5-P5316 @ 15.36 TB
 - TLC SSD 介质: 英特尔 SSD D7-P5510 @ 3.84 TB

客户端配置 - 基础设施

- 使用运行 VMware ESXi 的单一英特尔系统来支持基础设施
 - 3 台虚拟机用作 OpenShift “master” 节点
 - 运行 Microsoft Server 2019 的 “jump” 主机, 用于远程访问基础设施

SDS 存储配置

Ceph 配置

备注: Ceph 没有“默认”配置, 由于硬件的可用性和其他配置参数, 每次安装都可能不同。有关配置, 请参见下文:

- 使用“CentOS 8.4 版”作为所有三个节点的主机操作系统
- 按照上述定义配置硬件, 每个节点配置 8 个 NVMe 设备
- 下载后, 在全部 3 个节点上配置 Ceph “Pacific”
 - 将 2 TB 的 PMem 分成两个分区, 一个用于 WAL, 另一个用于 RocksDB
 - 每个 NVMe SSD 配置为 6 个逻辑 OSD (每台设备设置 6 个 OSD)
- Ceph 设备配置
 - 每台主机有 8 个 SSD, 分别被划分为 6 个区域
 - 在 8 个物理 SSD 上使用 6 个分区, 每个系统总共有 $6 * 8 = 48$ 个 OSD
 - Ceph 配置了驻留在 Optane PMem 上的 WAL
- 默认资源设置:
 - CPU 设置为每个 OSD 1 个 CPU, 内存设置为每个设备 4 GB

Lightbits 配置

备注: Lightbits 没有“默认”配置。Lightbits 和英特尔工程部门为 Lightbits 执行了所有的设置操作。有关配置, 请参见下文:

- 使用“Lightbits”作为全部三个节点的主机操作系统
- 按照上述定义配置硬件, 每个节点配置 8 个 NVMe 设备
- 在全部 3 个节点上配置 Lightbits
 - 指定 2 TB 的 PMem 作为缓存设备
 - 将每个 NVMe SSD 用作逻辑设备

应用工作负载

- 使用“vdbench”工具, 生成具有不同数据块大小和读写比率的合成工作负载
 - 在开始其他测试之前, 运行“写入-填充”工作负载, 完全写入被测的整个存储容量
 - 每个测试运行 15 分钟进行“预热”, 然后运行 30 分钟进行“测量”
 - 测试按顺序进行, 每完成 5 个测试称为一“组”
 - 每个测试组重复 9 次, 用 9 轮运行结果的平均值进行比较
- 针对存储配置测试以下 5 种工作负载
 - 4KB, 100% 读取, 100% 随机访问
 - 4KB, 100% 写入, 100% 随机访问
 - 8KB, 80% 读取/20% 写入, 100% 随机访问
 - 16KB, 70% 读取/30% 写入, 80% 随机访问
 - 32KB, 50% 读取/50% 写入, 50% 随机访问

测试过程概述

- 设置用于应用的服务器和网络
- 设置系统, 安装附加测试基础设施使用的 VMware
- 在 15 个系统上安装 OpenShift (12 个“worker”裸机硬件和 3 个“master”虚拟机节点)
- 安装和配置软件定义存储目标集群
 - 将 Lightbits 安装到 3 个节点上, 用于 Lightbits SDS 存储
 - 安装 CentOS, 并在用于 Ceph SDS 存储的 3 个节点上安装 Ceph Pacific
- 创建包含 96 个 vdbench 客户端容器实例和 1 个 vdbench 控制器实例的容器工作负载环境
- 在控制器容器上运行 vdbench 工作负载, 将工作负载分配给所有 96 个 vdbench 实例
- 为每个工作负载收集 9 组结果, 得到平均 I/O 速率、吞吐量和时延

使用 TLC 介质进行比较

使用 QLC 介质测试完毕后, 重新将两个存储系统配置为使用 TLC 介质设备, 而非 QLC 设备。众所周知, TLC 介质的写入 I/O 速率更优, 时延更低。同样, 测试两个存储系统时, 使用相同的工

作负载，在相同的服务器硬件、网络上运行，最重要的是，使用相同数量和类型的存储介质。基于 TLC 的 NVMe 固态设备作为主要的存储介质，并根据最佳实践，为每个存储堆栈使用英特尔傲腾 PMem。

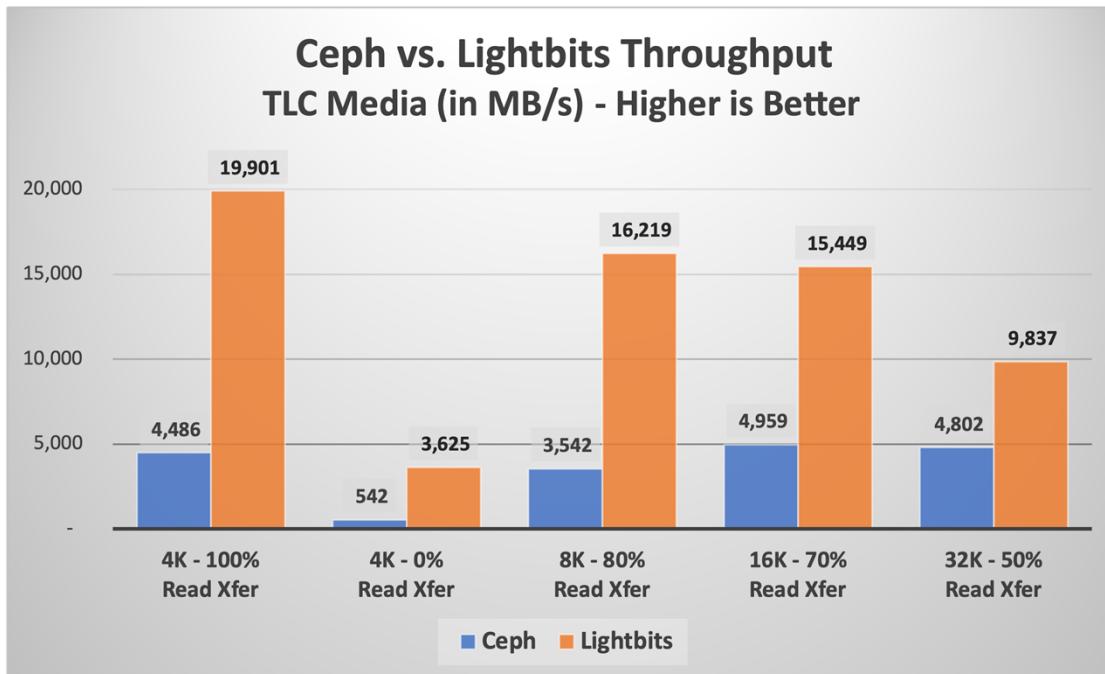


图 7: TLC 多工作负载下的 Lightbits 与 Ceph 吞吐量比较 (来源: Evaluator Group)

图 7 显示了使用 TLC SSD 介质时，五次相同测试的吞吐量结果。显示吞吐量结果旨在提供可直接比较的结果，不过 I/O 速率是更常用于显示单个小数据块工作负载的指标。