

基于全新英特尔® 架构的新华三 ONEStor 分布式存储系统



“人工智能、5G 网络、自动驾驶、物联网等新技术和新应用带来了指数级的数据量爆炸，对数据存储的容量、密度、持久性与可靠性需求也飞速增加，英特尔在存储领域与新华三进行技术合作，应用了非易失性持久化内存、高性能 NVMe 闪存等新的高性能存储技术，并结合分布式存储软件解决方案端到端的优化，将高性能分布式存储产品推上了一个新的台阶。”

— 李亚东
英特尔公司数据中心集团
政企事业部中国区总经理

“新华三是全球领先的智能联接、信息安全、大数据、云计算、分布式存储解决方案提供商，也是英特尔的战略合作伙伴。在 H3C 全闪分布式存储优化项目中，英特尔和 H3C 团队密切合作，在基于英特尔® 至强® 第二代可扩展处理器的平台上，引入英特尔傲腾持久内存和 SPDK 等软硬件产品，着力于软件优化、软硬件协同，最终大幅提升该存储解决方案的性能，为 H3C 的用户提供更高性能、高性价比的全闪存储方案。未来，我们还将紧密合作，为用户持续提供与时俱进的产品。”

— 王幼康
H3C 分布式存储产品副部长

在人工智能 (AI)、视频直播、互联网交易等应用的推动下，企业需要实时存储并处理的数据出现了爆发式增长，实时读写、随机访问超大规模数据集等场景对于存储系统的性能带来极高要求。随着固态硬盘 (SSD) 单位容量成本的不断下降，高 IO 吞吐率、低延迟业务的需求越发广泛，促进了高性能、全闪存存储介质的快速发展和普及。目前，全闪存存储已经逐渐成为企业在数字化转型策略中，部署或优化存储系统的优先选择。

在此背景下，新华三技术有限公司 (以下简称：H3C) 在 H3C ONEStor 分布式存储系统中搭载了全闪存阵列。该产品可以有效满足高并发、高实时性要求的业务应用场景，并能够提供比传统分布式存储环境更高的性能和更低的延迟。此外，H3C ONEStor 分布式存储系统还具备高扩展性、高灵活性、智能化、高可用性等优势，凭借卓越的性能密度及容量密度，满足绝大多数企业级用户对关键业务环境的需求。

挑战：如何进一步提升全闪存存储的性能

凭借高性能可靠性、耐用性、可管理性等优势，全闪存已经成为企业级存储演进的重要方向。IDC 发布的《全球季度企业存储系统跟踪报告》显示，2019 年第三季度全球全闪存阵列市场实现创收 25.8 亿美元，同比增长 11.7%，显著高于当季度企业存储市场的整体增长率¹。与此同时，越来越多的企业基于全闪存阵列来部署分布式存储，

在利用分布式存储低成本、易扩展等特性的同时，满足关键业务对于实时数据访问的苛刻需求。

随着闪存性能不断提升，以及软件优化技术的持续创新，全闪存的性能在近几年已经实现了快速飞跃。但全闪存存储追求极致性能的道路依然没有止境，通过采用更高性能的 SSD、搭载创新的软件优化技术、使用专有的性能加速芯片等方式，全闪存存储的性能仍然在持续快速提升。为了提升 H3C ONEStor 分布式存储系统的性能，H3C 从元数据 (Metadata) 访问加速、降低处理器 (CPU) 开销、优化分布式存储系统等方面进行了技术探索。

首先，从云数据访问来看，由于性能高、搭载了高级功能特性等特点，全闪存系统中的元数据的读写访问往往非常频繁。而在分布式存储系统上，系统的元数据和数据缓存通常部署在固态硬盘上，固态硬盘性能的瓶颈会导致元数据读写缓慢，进而拖慢全闪存存储的整体性能。因此，H3C 希望找到比闪存性能更高、且能持久化保存的介质存储元数据。

其次，随着闪存介质性能不断提升，全闪存存储的主要性能瓶颈已经从闪存转移到了 CPU。在数据中心应用场景中，多

块固态硬盘的同时读写常常导致 CPU 的性能开销达到 90% 以上。如果无法破除 CPU 的性能瓶颈，就只能牺牲全闪存的扩展性，降低固态硬盘的搭载数量，这就使得全闪存的性能、容量都受到极大的限制。

再次，H3C ONEStor 分布式存储系统是基于开源 Ceph 开发优化的，但 Ceph 较早版本却并未针对全闪存存储进行优化。要进一步提升 H3C ONEStor 的性能，就需要识别这些优化点，并进行相应代码移植再开发。

解决方案：基于英特尔架构的 H3C ONEStor 分布式存储系统

H3C ONEStor 分布式存储系统采用了分布式的架构，为实现存储系统的可靠性、可用性、自动运维、高性能等特性提供了有力保证。该系统具有诸多优势，包括领先的分布式融合架构、线性扩展、灵活的分区存储、丰富的企业级功能特性和自动化运维能力等，可以应用于块存储、文件存储、对象存储等多种应用场景，满足不同行业用户的存储需求。

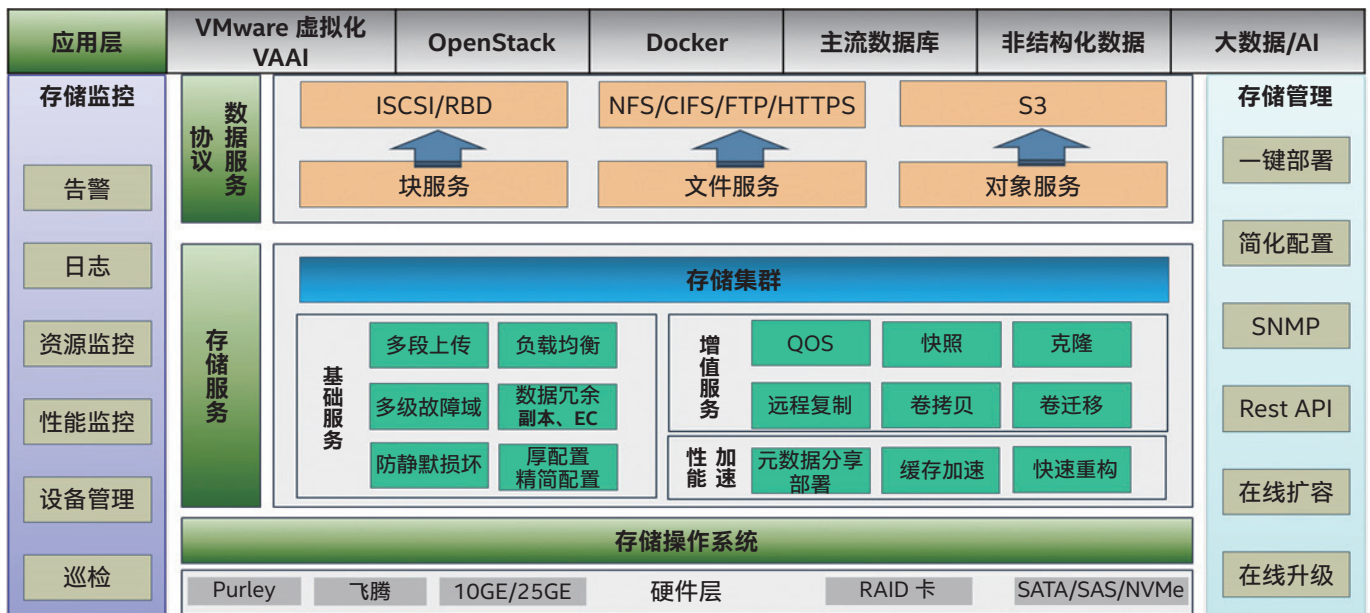


图 1. H3C ONEStor 分布式存储系统架构图

第二代英特尔® 至强® 可扩展处理器

第二代英特尔至强可扩展处理器：

- 为企业数据中心提供高可扩展性
- 与上一代处理器相比，为虚拟基础设施提供更高性能
- 实现卓越的资源利用效率和敏捷性
- 针对数据中心解决方案，提高数据和工作负载的完整性，确保监管合规性

基于全新英特尔架构的新华三 ONEStor 分布式存储解决方案推荐采用第二代英特尔至强可扩展处理器。



英特尔® 傲腾™ 持久内存

英特尔傲腾持久内存可以提供：

- 更低的总体拥有成本 (TCO)：要组建性能或容量相当的内存池，将仅需使用更低的成本，这也有助于企业降低基础设施的 TCO，获得更大的竞争优势。
- 敏捷的应用模式：企业用户可以根据自己的应用需求，在不对硬件进行调整的前提下，自由地使用英特尔傲腾持久内存的内存模式或是应用直接访问模式。
- 强大的性能，远超传统固态硬盘。

基于全新英特尔架构的 H3C CAS 虚拟化平台快速克隆解决方案推荐采用英特尔傲腾持久内存。



为了提升 H3C ONEStor 分布式存储系统的性能，H3C 与英特尔进行了合作，结合英特尔® 傲腾™ 持久内存、第二代英特尔® 至强® 可扩展处理器、英特尔® SPDK 性能工具包等产品，对全闪存分布式存储进行了针对性的性能优化。存储服务器推荐使用第二代英特尔至强可扩展处理器，以满足对于性能、敏捷性、扩展性等方面的更高要求。

元数据访问加速

为了确保元数据非易失的前提下，提升元数据读写的性能，H3C ONEStor 分布式存储系统采用了英特尔傲腾持久内存作为介质来存储元数据。英特尔傲腾持久内存是一种颠覆传统的持久内存，集大容量、经济性和持久性于一身，读写带宽比 NVMe 设备高数倍，并能够大幅降低延时问题。英特尔傲腾持久内存分为内存模式 (Memory Mode) 和应用直接访问模式 (App Direct Mode)。在应用直接访问模式下，英特尔傲腾持久内存可以在确保数据长期存储的同时加速数据访问。

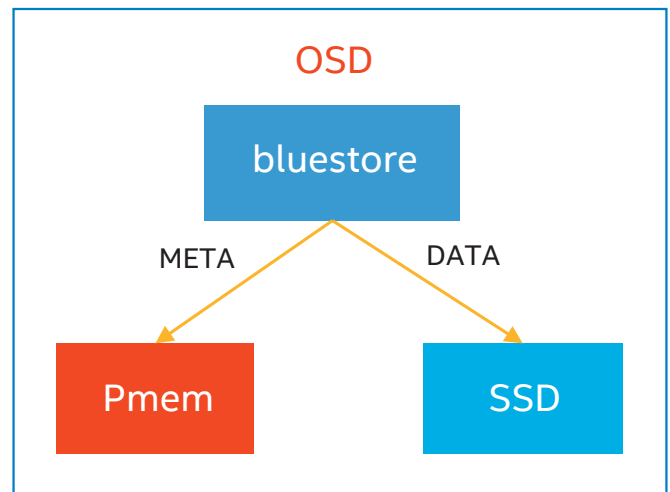


图 2. H3C ONEStor 分布式存储系统元数据访问加速示意图

在优化实践中，H3C 在应用直接访问模式下使用英特尔傲腾持久内存，来承载 H3C ONEStor 分布式存储系统的元数据，

并通过持久内存开发工具包 (PMDK) 访问英特尔傲腾持久内存, 以加速元数据读写。

SPDK 驱动降低系统开销

H3C ONEStor 分布式存储系统搭载了高性能的英特尔至强可扩展处理器, 该处理器具备强大的性能表现, 有利于为关键工作负载提供高吞吐量与 IOPS、以及低时延的存储性能。除了搭载英特尔至强可扩展处理器, H3C 还重点采用英特尔 SPDK 进行了优化。英特尔 SPDK 是一套基于 NVMe 的开发套件, 目标是通过使用英特尔的网络、计算和存储技术, 尽可能地发挥闪存介质出色的性能。

相对于传统 IO 方式, 英特尔 SPDK 运用了两项关键技术: UIO 和 pooling。首先, 英特尔 SPDK 可以将设备驱动代码运行在用户态 (User Mode), 避免内核上下文切换和中断, 这有助于节省大量的处理开销, 允许更多的时钟周期被用来做实际的数据存储。在 H3C ONEStor 分布式存储系统中, H3C 针对 NVMe 固态硬盘将原来的内核驱动替换为 SPDK 的用户态驱动, 以提升性能。

开源 Ceph 优化

H3C 与英特尔展开代码级紧密合作, 引入开源社区中的针对 BlueStore 和 RBD 关于闪存存储介质的优化实现, 并对原生

Ceph 的部分配置参数进行了优化, 实现了全闪场景下的最优性能。

效果: 基于英特尔架构的 H3C ONEStor 分布式存储系统加速数据变革

经过上述各项优化, H3C ONEStor 分布式存储系统的性能得到了很大的提升, SPDK 驱动单项优化提高 36%, 最新开源社区的相关模块集成也带来 34.6% 的性能, 综合性能提高 30% 以上。

通过使用基于英特尔傲腾持久内存、英特尔 SPDK 进行优化的全闪存阵列, H3C ONEStor 分布式存储系统进一步增强了性能优势, 可以有效满足政府、教育、企业、电力能源、金融、运营商等行业的关键业务对于实时数据处理的需求。除了高性能之外, 该解决方案还具备以下优势:

- **更优的成本效益:** 通过使用英特尔傲腾持久内存、英特尔 SPDK 来优化全闪存性能, H3C ONEStor 可以降低用户对于分布式存储系统节点数量的需求, 节省相应的采购、能耗、空间、运维等支出, 降低系统的总体拥有成本 (TCO)。
- **更强的扩展性与敏捷性:** 英特尔傲腾持久内存兼具内存和存储属性, 在存储系统中有着非常广泛而灵活的应用场景, 有助于存储系统实现性能、成本等方面的平衡。

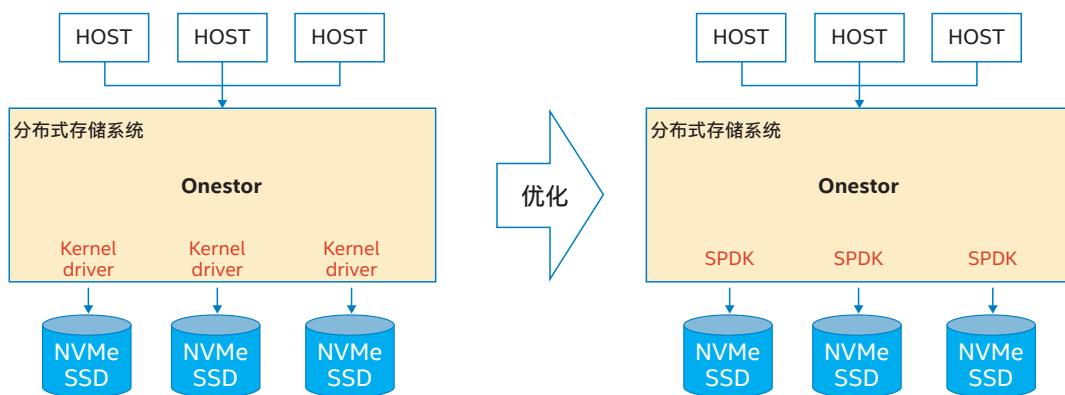


图 3. H3C ONEStor 分布式存储系统 CPU 开销优化示意图

总体而言，基于英特尔架构进行全闪存优化的 H3C ONEStor 分布式存储系统能够帮助企业与行业用户有效应对数据爆炸式增长带来的挑战，加速数据引发的颠覆式变革。

展望：H3C 与英特尔协力推动全闪存系统创新

通过与英特尔的紧密合作，并使用英特尔产品与技术，H3C 全面优化了全闪存的性能表现，为 H3C ONEStor 分布式存储

系统在行业关键业务中的应用奠定了坚实的高性能基础。双方还将持续在网络相关技术（RDMA）、最新 Ceph 版本和傲腾持久内存（AD 模式）等方面深入合作，以进一步促进包括全闪存在内的存储系统创新，推动全闪存在更多存储系统与应用场景的落地，最大化地发挥全闪存对于性能的提升作用，支撑更多关键数字化业务，助力行业的数字化转型升级。

附录 1：基于全新英特尔® 架构的新华三 ONEStor 分布式存储解决方案推荐配置列表

存储服务器：UIS 3000G3 (2U)	
处理器	2x 英特尔至强金牌 5218 处理器 @ 2.30GHz
内存	8 x 32 GB DDR4-2666
存储网络	管理网：1xGE 业务网：2x25GE (bond4) 存储网：2x25GE (bond4)
硬盘	4 x 1.8TB NVMe 固态硬盘
RAID	P460
客户机：H3C R4900 G2	
处理器	2x 英特尔至强处理器 E5-2690 @ 2.60GHz
Memory	16 x 32GB 2400Ghz
网络	管理网：1xGE 业务网：2x25GE (Bond4)

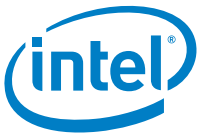
关于 H3C

紫光旗下新华三集团作为数字化解决方案领导者，致力于成为客户业务创新、数字化转型最可信赖的合作伙伴。新华三拥有计算、存储、网络、安全等全方位的数字化基础设施整体能力，提供云计算、大数据、智能联接、信息安全、新安防、物联网、边缘计算、人工智能、5G 等在内的一站式数字化解决方案，以及端到端的技术服务。同时，新华三也是 HPE® 服务器、存储和技术服务的中国独家提供商。

关于英特尔

英特尔致力于创造最不可思议的体验

您可能知道英特尔的处理器。但我们的业务不仅限于处理器。英特尔致力于提供技术创新，为企业和社会以及地球上的每个人提供卓越的体验。借助功能强大的云计算、无处不在的物联网、最新的内存和可编程解决方案技术、丰富的人工智能技术组合以及不间断的连接，英特尔正在颠覆医疗和生命科学行业，帮助他们应对最棘手的挑战。



¹【2019 年 Q3 企业外部存储系统以亚太地区收入增长最为强劲】，<http://www.elecfans.com/consume/1140031.html>

在特定系统中对组件性能进行特定测试。硬件、软件或配置的任何差异都可能影响实际性能。考虑购买时，请查阅其他信息来源以评估性能。如欲了解有关性能及性能指标评测结果的更完整信息，请访问：<http://www.intel.cn/content/www/cn/zh/benchmarks/benchmark.html>

在性能测试过程中使用的软件及工作负载可能仅针对英特尔微处理器进行了性能优化。SYSmark* 和 MobileMark* 等性能测试使用特定的计算机系统、组件、软件、操作和功能进行测量。上述任何要素的变动都有可能导致测试结果的变化。您应当参考其它信息和性能测试以帮助完整评估您的采购决策，包括该产品与其它产品一同使用时的性能。更多信息敬请登陆：<http://www.intel.cn/content/www/cn/zh/benchmarks/benchmark.html>

描述的成本降低方案旨在作为举例，说明指定的英特尔架构产品在特定环境和配置下，可能如何影响未来的成本和提供成本节省。环境各不相同。英特尔不保证任何成本和成本的节约。

英特尔技术的特性和优势取决于系统配置，可能需要支持的硬件、软件或服务激活。实际性能可能因系统配置的不同而有所差异。任何计算机系统都无法提供绝对的安全性。请咨询您的系统制造商或零售商，也可登录 [intel.cn](http://www.intel.cn) 获取更多信息。

英特尔未做出任何明示和默示的保证，包括但不限于关于适用性、适合特定目的及不侵权的默示保证，及履约过程、交易过程或贸易惯例引起的任何保证。

英特尔、英特尔标识、Optane、傲腾、Xeon 和至强是英特尔公司在美国和/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。

英特尔公司 © 2020 年版权所有。所有权保留。