Software-Defined Storage for Private Clouds: Lightbits vs Ceph

Lightbits delivers up to 16x better performance with lower TCO and enterprise resiliency.

March 2024

Table of Contents

1. Overview

Cloud computing as an operational model and cloud native as an application architecture are together transforming the way information technology is delivered, managed, and consumed. Both are gaining traction due to their ability to run and operate applications with greater flexibility - and often, with lower costs than traditional technologies and approaches.

To meet these evolving application requirements, enterprises in industries from e-commerce to financial services to healthcare are building private clouds for their cloud-native environments and are turning to software-defined storage for their data platform. These modern storage solutions deliver highly scalable capacity and performance with greater efficiency at a lower total cost of ownership (TCO) than storage-attached networks (SAN), directed-attached storage (DAS), or hyper-converged infrastructure.

At the same time, these modern storage solutions ensure that containerized workloads retain access to persistent data through a resilient underlying storage infrastructure, while supporting thousands of applications. For virtual machine (VM) workloads, software-defined storage seamlessly pools capacity supporting availability during infrastructure dynamics. Features like compression, replication, self-healing, and high performance uphold workload continuity by leveraging the scalable, resilient software foundation.

This white paper compares two software-defined storage solutions suited for private clouds.

- **Ceph**: Open-source and software-defined block, file, and object storage often used in cloud-native environments.
- **The Lightbits® cloud data platform from Lightbits Labs™:** Software-defined, disaggregated, and composable block storage architected from the ground up for cloud environments.

As illustrated below, Lightbits significantly outperformed Ceph for all workloads tested-using identical hardware, with testing performed using workloads running as containers in an OpenShift Kubernetes environment. The summary in Table 1 shows results using Non-Volatile Memory Express (NVMe®) SSDs. The result: Lightbits delivers up to 16x better performance with lower TCO and enterprise resiliency. We dive deeper into the architectural comparisons to explain this performance advantage in the following sections.

	Ceph IOPs	Lightbits IOPs	Lightbits Advantage
IOPS for Small Block Sizes			
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
Bandwidth for Bigger Block Sizes (MB/s)			
16K-70% Read	558	5,813	10.41 x
32K-50% Read	619	3,495	5.65 x

Table 1: Lightbits vs. Ceph I/O Performance Comparison in Containerized Environment.

The remainder of this paper explores the technology and performance differences in more detail, along with architectural differences.

2. Emerging Technologies

The move towards cloud architectures has changed not only where IT operations occur and how they're paid for, but also with how applications are developed to be cloud-native. At the same time, new technologies enable NVMe storage to be disaggregated across standard TCP/IP networks, boosting performance and reducing infrastructure and management costs.

Cloud-Native Applications

Designed to operate in both private and public cloud environments, cloud-native applications are built as services that can be independently scaled. These are known as microservices, an architecture that enables cloud-native applications to scale on-demand, with little or no administrative input.

Cloud-native applications are typically containerized, making them easily deployed in one location, then moved to another based on available resources or even based on the price of the resources. For this reason, it is imperative that the persistent storage required by many of these applications also responds dynamically to changing requirements for capacity as well as performance.

NVMe over Fabrics

NVMe is a protocol designed specifically for flash-based storage devices, offering improved performance and reduced latency compared to older protocols such as SCSI or SATA. NVMe is known for its ability to take full advantage of the high-speed capabilities of NAND flash, providing faster data transfer rates and lower latency. As a result, it has become the preferred choice for modern storage solutions.

To accommodate communications to external devices, NVMe over Fabrics (NVMe-oF) extends the NVMe protocol to support accessing NVMe devices over a network fabric.

The most flexible and scalable transport of NVMe-oF is TCP/IP, known as NVMe over TCP (NVMe/TCP) which imposes far fewer technical burdens to implement, supporting commodity Ethernet without any specific hardware or switch settings. Overall, NVMe/TCP provides significantly lower latency and higher throughput than other storage protocols over TCP, while still using existing Ethernet networking infrastructure.

According to Gartner, NVMe/TCP fast block storage solutions are gaining traction as organizations aim to support low-latency applications with scalable architectures leveraging high-speed networking. They forecast that by 2027, 25% of enterprises will deploy NVMe/TCP as a storage networking protocol, up from less than 10% in mid-2023. Gartner recognizes NVMe/TCP's ability to deliver performance exceeding iSCSI and low-end Fibre Channel while simplifying network infrastructure. Use cases where the benefits justify deployment include AI/ML, databases, transaction processing, and replacing iSCSI.

As organizations evaluate options to stay ahead of evolving business demands, purpose-built NVMe/TCP storage platforms like Lightbits are well positioned to meet the needs of new cloud-native applications. In fact, Gartner has recently included NVMe/TCP fast block storage solutions such as Lightbits in their Top Trends in Enterprise Data Storage.

3. Comparing Software-Defined Storage Architectures

Storage running as software on commodity hardware, known as Software-defined Storage (SDS), has gained traction for providing scalable data center class capabilities without requiring proprietary appliances. SDS offers flexibility to operate consistently across various on-premises and cloud-based environments, while gaining cost efficiencies by leveraging standardized server infrastructure.

These traits make SDS well-suited for modern containerized applications and microservices - as well as legacy VMs that need storage mobility across infrastructures. SDS enables resilient data services for containers and can be adapted dynamically to changing workloads. However, not all SDS options excel in integrating with container ecosystems or meeting the performance demands of cloud-native applications.

Both Ceph and the Lightbits cloud data platform are SDS solutions supporting containerized or virtualized workloads but take different architectural approaches. In the following sections of this paper, we will compare both solutions side-by-side.

While Ceph offers open-source advantages and unified storage protocols, Lightbits' purpose-built and cloud-native design delivers substantially better storage performance, efficiency, and operational simplicity at scale under demanding production workloads, such as SQL, NoSQL, real-time, and vector databases.

Lightbits Cloud Data Platform

The Lightbits cloud data platform is software-defined storage designed to utilize more modern technologies, such as NAND flash, and supports NVMe/TCP using high-speed 100GbE network interface cards. As depicted in Figure 1, Lightbits uses a disaggregated architecture, enabling it to scale CPU, memory, and NVMe devices independently as needed. This allows Lightbits to deliver scalable, enterprise-class storage with performance that surpasses local NVMe devices.

Lightbits' Intelligent Flash Management (IFM) is a set of features that maximize the performance and extend the endurance of SSDs. IFM implements Elastic RAID, a newer RAID architecture, which combines self-healing capabilities with per node erasure coding to maximize data protection from SSD failures.

Access to Lightbits storage for Kubernetes-managed containers as Persistent Volumes (PVs) is provided through the Kubernetes Container Storage Interface (CSI). The Lightbits CSI Plugin enables Kubernetes to store PVs in the Lightbits cluster. The driver is part of the official CNCF CSI drivers list, which demonstrates Lightbits' commitment to compatibility and interoperability within Kubernetes environments.

Included in the Lightbits software license are robust data services such as compression, replication, volume snapshots, clones, QoS and Role-Based Access Control (RBAC) for multi-tenant environments.

Lightbits offers enhanced functionality through its Cinder plugin, delivering high-performance block storage services specifically designed for OpenStack-based cloud infrastructure. This seamlessly integrated plugin comes pre-installed in the OpenStack Yoga version and subsequent releases, ensuring optimal compatibility and performance.

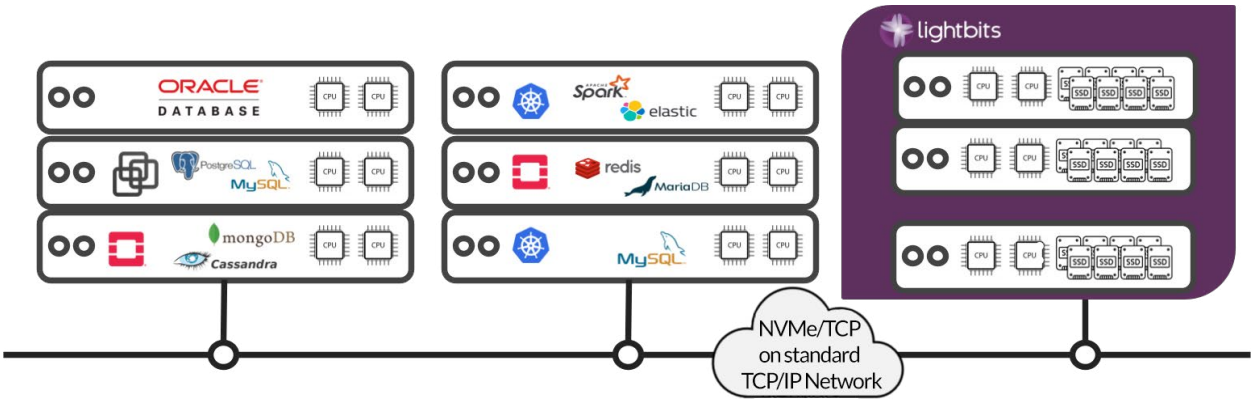

Figure 1: Lightbits cloud data platform implementation: high-performance direct NVMe/TCP connectivity from clients to storage.

Ceph

Ceph is an open-source, software-defined storage platform employing a distributed object storage architecture designed for scalability, flexibility, and fault tolerance on commodity infrastructure.

Originally intended for spinning hard disk drives (HDDs), Ceph has incrementally evolved its architecture over the past decade to take advantage of flash storage advances. Recent versions of the software utilize solid-state drives (SSDs) for metadata operations, improving performance. However, many core design elements optimizing HDD behavior remain unchanged.

Ceph aims to provide unified storage, with object, block, and file interfaces in a single platform. The block storage service (RBD) presents virtual block devices to applications. A key benefit of Ceph's architecture is leveraging commodity hardware to scale capacity across thousands of nodes. However, scaling capacity can substantially increase demands on the supporting cloud network fabric and has been shown to introduce latency with larger-scale private cloud environments.

NVMe/TCP Support for Ceph

Ceph recently unveiled a technology preview for NVMe/TCP connectivity. However, this implementation involves the integration of Ceph's NVMe-oF gateway. Figure 2 illustrates how the gateway exports RADOS Block Devices (RBD) to clients over NVMe/TCP. This model introduces additional architectural complexity, leading to bottlenecks and increased storage networking latency caused by the extra hop and protocol translation.

While Ceph strives to support modern high-speed protocols such as NVMe/TCP, the current approach involves the use of protocol gateways and translation layers atop the existing Ceph architecture. This model may improve Ceph's interoperability, but it deviates from the originally intended design of NVMe/TCP fabric architectures. This design is implemented by Lightbits, which is meticulously engineered to offer direct and high-performance host connectivity.

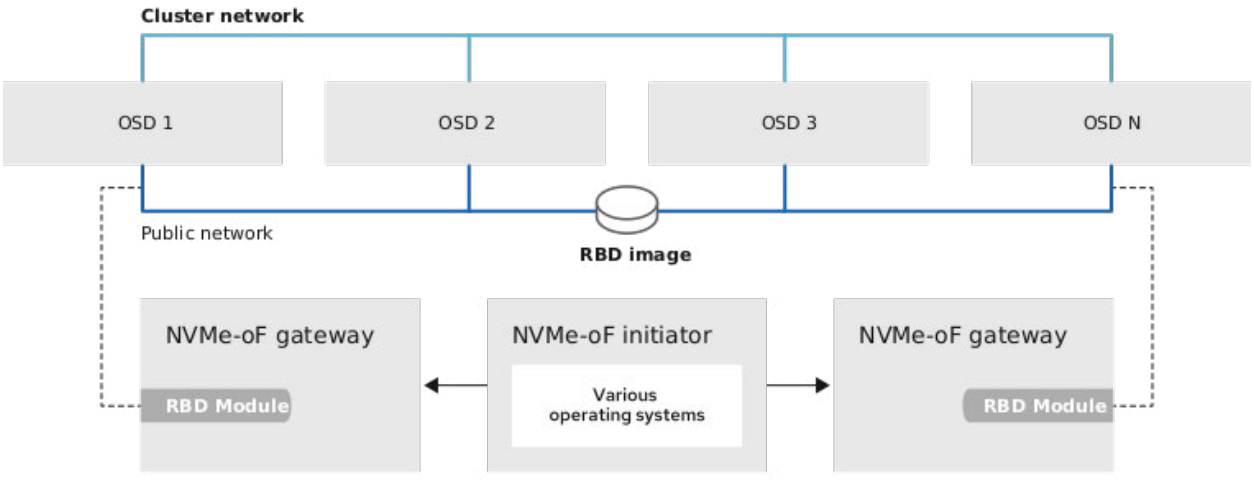

Figure 2: Ceph NVMe-oF gateway from IBM Storage Ceph product documentation, "Ceph NVMe-oF gateway (Technology Preview)".

Storage Efficiency

Both Lightbits and Ceph use a clustered architecture where nodes run on commodity server hardware and communicate with each other to store, retrieve, and replicate data. Nodes can be added to the cluster to linearly increase scalability, performance, and availability. Since adding nodes increases hardware costs as well as management complexity and cost, maximizing performance on a per-node basis reduces overall TCO.

For example, the performance benchmarking results later in this paper demonstrate that with 8 KB block sizes and 80% reads, Lightbits delivers 1.1 million IOPS per server, while Ceph delivers only 90,000. To reach a target of four million IOPS, the Ceph cluster would require 45 servers while the Lightbits cluster requires only four servers. That's a difference of more than a full rack of servers versus less than a 1/10th of a rack. This reduction in servers to achieve the 12.5x performance increase over Ceph has a direct impact on lowering infrastructure costs. Over ten times more servers, and more storage media, space, cooling, and cabling are required for Ceph— thus inflating IT budgets.

Availability and Resiliency

The resiliency design of Lightbits provides additional benefits. The erasure coding on each node provides data availability in the event of the loss of a storage device. When combined with replication between nodes, Lightbits provides higher resiliency with less overhead and wasted storage capacity.

For the purposes of this whitepaper, Lightbits testing was configured to use erasure coding for data protection on each node, along with making one additional copy of data on another node within the Lightbits cluster. In contrast, Ceph was configured to create three data copies for data protection, which does not ensure data availability if a drive in one node along with an additional node experiences an outage.

Although the storage efficiency of these two protection methods is similar, the Lightbits approach provides higher availability.

4. Performance of Container Workloads

Both Lightbits and Ceph storage offerings were analyzed running as containers in a Kubernetes applications environment.

Both the performance of Lightbits and Ceph were evaluated using identical hardware configurations consisting of 12 Kubernetes nodes running container workloads against either a dedicated three-node Lightbits cluster, or a three-node Ceph cluster. These were equipped with:

- Dual socket 3rd Gen Intel Xeon Gold 6338
- 256 GB of DRAM
- Intel E810 100Gbe NIC
- 8 x 15.36TB NVMe SSDs (D5-P5316)

The Kubernetes nodes were running vdbench, a commonly used benchmarking tool to simulate workloads and assess storage system performance. Each node hosted eight instances of vdbench, totaling 96 instances across all 12 nodes. Each node was equipped with:

- Dual socket 1st Gen Intel Xeon Platinum 8173M
- 96 GB of DRAM
- Intel XXV710 2x25 Gbe NIC
- 8 x 15.36TB NVMe SSDs (D5-P5316)

Storage performance was measured using the well-known ‘vdbench’ tool to create workloads, using eight container instances running vdbench per node, for a total of 96 instances of vdbench.

Testing consisted of five different access patterns and block sizes often found in performance-sensitive applications:

- 4KB, 100% read, with 100% random access
- 4KB, 100% write, with 100% random access
- 8KB, 80% read / 20% write with 80% random access
- 16KB, 70% read / 30% write with 80% random access
- 32KB, 50% read / 50% write with 80% random access

These workloads were used to compare the performance of Lightbits to Ceph, with higher-speed persistent media where appropriate for each storage system.

Benchmark Results

We conducted a comparative analysis between Lightbits and Ceph across specific workload profiles. Our evaluation included testing for maximum IOPS using random 4KB with 100% reads. Recognizing the diverse needs of many databases and applications that rely on writes and use larger block sizes, we extended our examination to use 8KB block size with an 80% read and 20% write mix. As expected, the introduction of writes resulted in a decline in IOPS due to replication overhead.

Figure 3 illustrates the outcomes, showcasing that Lightbits significantly outperforms Ceph in both scenarios. With a 4KB block size, Lightbits demonstrated a remarkable 12.5x advantage, while in the 8KB workload, it surpassed Ceph by over 3.94 times.

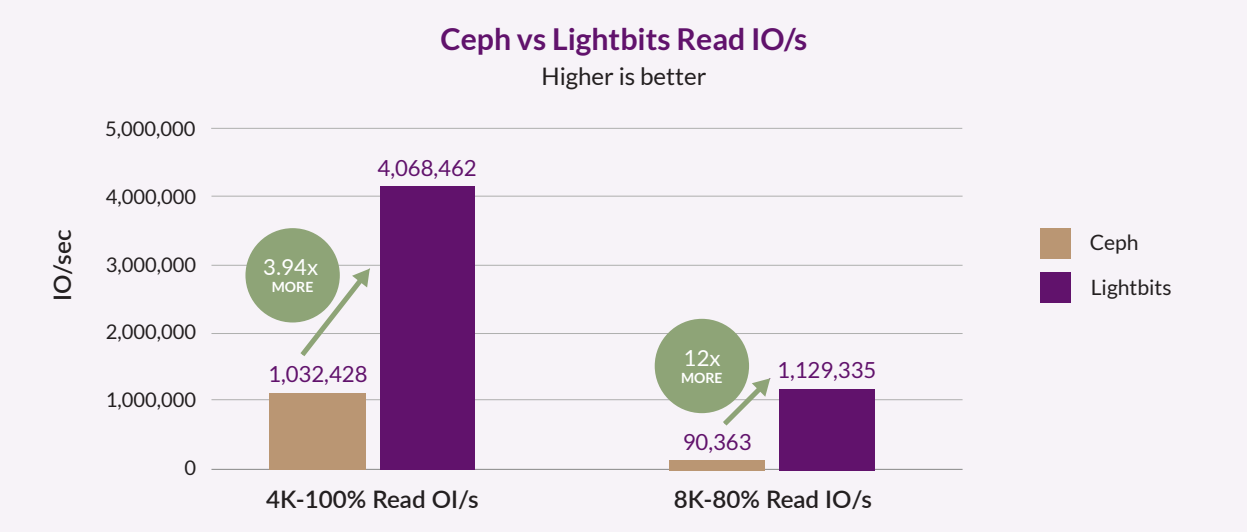

Figure 3: Lightbits IOPS outperforms Ceph for 4KB and 8KB block sizes.

While 4KB read operations can occur, 4KB write operations are quite often used for database logging operations, with Lightbits providing nearly 17X better performance than Ceph, as shown below in Figure 4.

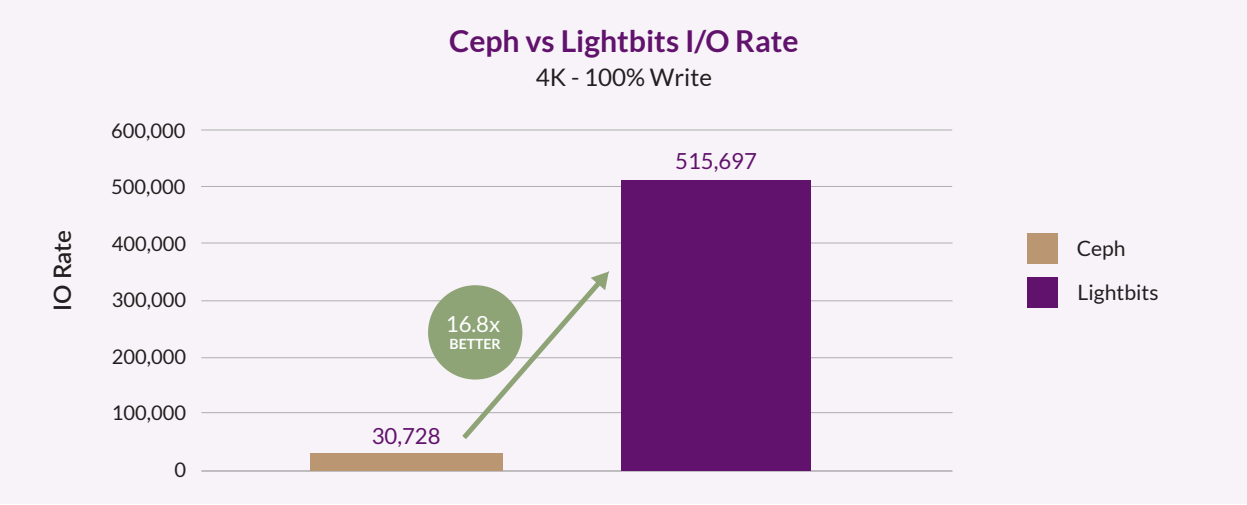

Figure 4: Lightbits outperforms Ceph for 4KB Write I/O.

We tested random read throughput for block sizes of 16KB and 32KB on Lightbits and Ceph, with a 70% read / 30% write and 50% read / 50% write mix, respectively. For the 16KB 70/30 workload, Lightbits delivered 10.41x higher MB/s performance than Ceph. At 32KB with a more write-intensive 50/50 mix, Lightbits still significantly outperformed Ceph by 5.65x, as shown in Figure 5 below.

These I/O profiles are commonly seen in applications that demand an even distribution between data retrieval and storage updates. This setup is well-suited for scenarios where both reading and writing operations play an equally crucial role, making it applicable to a range of database and transactional workloads. The higher throughput numbers illustrate Lightbits' architectural advantages in driving higher bandwidth across some application workload profiles.

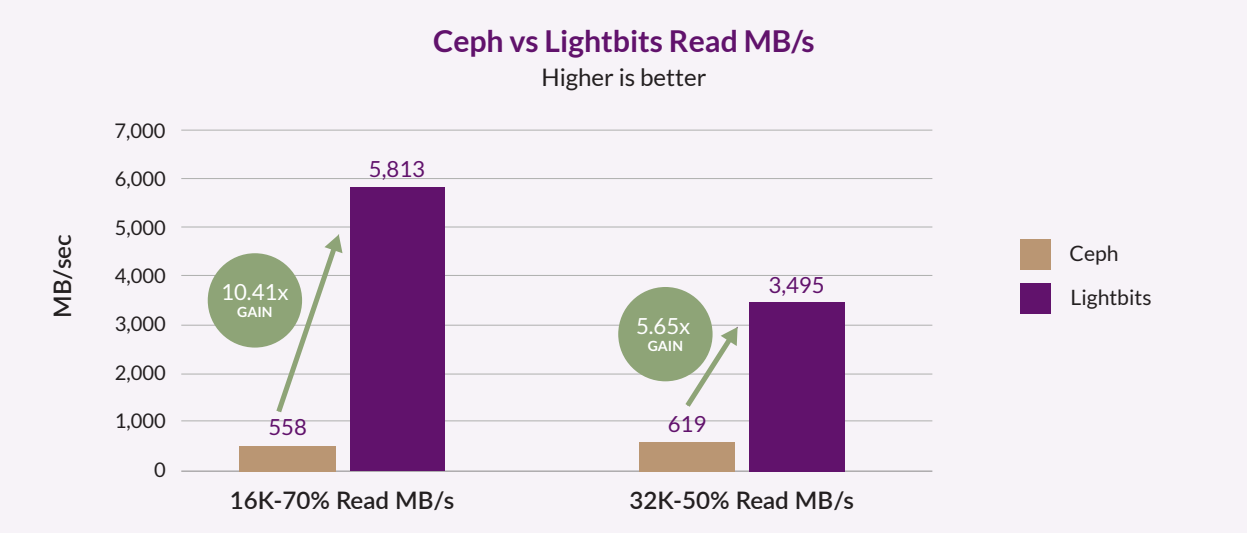

Figure 5: Lightbits throughput outperforms Ceph for 16K-30% and 32K-50% writes.

The workloads shown above are very common with databases or other transactional applications. Although small block sizes for reads are important for databases, writes are typically in bigger block sizes, ranging between 16KB and 32KB.

In addition to higher throughput, Lightbits delivers far lower read latency - which is critical for performance-sensitive applications. We measured random read latency using 100% read 4KB blocks and 80% read 8KB blocks across the two solutions. For 4KB reads, Lightbits latency was 3.95x faster than Ceph. At the 8KB mixed workload, Lightbits displayed 16x lower latency, as shown in Figure 6 below.

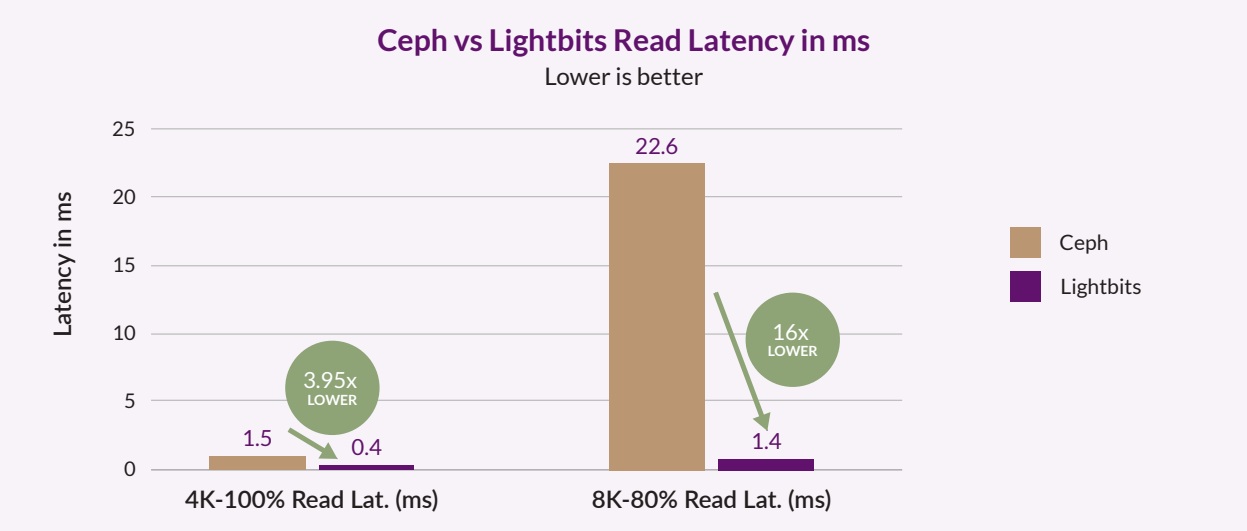

Figure 6: Lightbits outperforms Ceph with lower latencies for 4K-100% reads and 8K-20% writes.

Conclusion

While performance is always an important consideration in IT infrastructure design, its importance rises to new levels when operating cloud environments or hosting cloud-native applications. Although not every application or microservice demands high I/O rates, providing resilient, scalable storage with high IOPS, bandwidth, and low latency is critical to running modern cloud infrastructure with efficiency.

Software-defined storage has emerged as the preferred architecture to meet these data platform needs, with open source Ceph increasingly considered for its flexible, cost-effective block storage capabilities. However, Ceph has significant limitations in delivering performance at scale, and many enterprises are finding that it does not meet the needs of the growing demands of their cloud-native applications.

As highlighted in the test results, Lightbits outperformed Ceph by a significant margin while offering far greater resiliency using the same hardware configuration. With Lightbits, enterprises building private clouds can leverage the power and flexibility of software-defined storage to give their Kubernetes and other cloud-based applications the performance, availability, and ease of operations for their needs today and in the future.

About Lightbits Labs™

Lightbits Labs (Lightbits), offers a complete data platform that enables organizations to run performance-sensitive workloads on the public cloud or to build a high-performance, cost-efficient on-premises private cloud. The software-defined, NVMe/TCP and clustered architecture, coupled with essential data services solve the common cloud storage challenges of performance, efficiency, and cost eliminating the barriers to cloud adoption. Lightbits is backed by enterprise technology leaders [Cisco Investments, Dell Technologies Capital, Intel Capital, Lenovo, and Micron] and is on a mission to deliver a robust cloud storage platform with unmatched performance, efficiency, agility, and flexibility.
LBWP06/2024/03

www.lightbitslabs.com

info@lightbitslabs.com

US Office
1830 The Alameda,
San Jose, CA 95126, USA

Israel Office
17 Atir Yeda Street,
Kfar Saba, Israel 4464313