

Inspur Information, Intel, and Lightbits Labs jointly provide a distributed all-flash storage solution based on NVMe/TCP

Inspur Information, Intel and Lightbits Labs jointly provide a distributed all-flash storage solution based on NVMe/TCP that promotes business innovation through technological innovation, helps users apply the most advanced architecture to achieve innovation and optimization of cloud and data center infrastructure, and helps enterprise storage technology achieve performance breakthroughs.

Abstract

With the development and large-scale application of modern technologies such as the Internet, the Internet of Things (IoT), and artificial intelligence/machine learning (AI/ML), the scale of data has increased exponentially, and the storage and use of that data faces new challenges.

Inspur Information and Intel have teamed up with Lightbits Labs® (Lightbits®) to provide an all-flash storage solution based on NVMe® over TCP (NVMe/TCP) that promotes business innovation through technological innovation, helps users apply the most advanced architecture to achieve innovation and optimization of cloud and data center infrastructure, and helps enterprise storage technology achieve performance breakthroughs.

Storage Challenges

With the rapid development of emerging technologies such as big data, cloud computing, and IoT, the world has entered the era of the digital economy. As the cornerstone and enabler of the digital era, storage technology has

undergone tremendous advancements in the past decade. New storage media represented by flash memory have an order of magnitude improvement in performance and latency compared to mechanical hard drives (HDDs), and solid-state drives (SSDs) have begun to gradually replace HDDs as mainstream storage devices. In order to give full play to the advantages of high-performance hardware, the NVMe protocol specification came into being. The NVMe standards organization defines the NVMe over Fabric (NVMe-oF™) protocol. However, NVMe-oF only supported Fibre Channel (FC) and Remote Direct Memory Access (RDMA) at the beginning, both of which require the support of dedicated network card components and switches, which increases the difficulty and complexity of deployment and cost. What was needed is a new, lower-cost way to fully leverage the speed and performance of NVMe.

Non-Volatile Memory Express over TCP (NVMe/TCP) is an industry storage transport standard developed by the NVM Express® Alliance, which includes various industry organizations including Intel and Lightbits. Lightbits is the inventor of the NVMe/TCP protocol, and has standardized the protocol with industry partners such as Intel, making it part of NVMe over Fabrics, thus providing the industry with a protocol basis for implementing high-performance storage servers. Using NVMe/TCP as a high-performance transport protocol, the commonly used standard Ethernet can be supported without any special equipment and configuration. With lower latency and higher throughput, NVMe/TCP is a better choice for existing data center storage systems.

Inspur Information and Lightbits Distributed All-Flash Storage Solution

Inspur Information, Intel and Lightbits jointly launched a distributed all-flash storage solution based on NVMe/TCP, using Inspur servers equipped with Intel® Optane™ persistent memory and Inspur NS8610 G1 solid state drives as a hardware infrastructure for efficient data processing and storage, Lightbits is an efficient data processing platform for software-

defined storage, providing higher-performance software-hardwareintegrated storage solutions for cloud and data centers.



Figure 1 Inspur and Lightbits software and hardware integrated all-flash storage solution

Figure 1 shows the Inspur and Lightbits software and hardware integrated all-flash storage solution. This solution has the following advantages:

- **Flexible configuration and elastic expansion**

The solution adopts Inspur's new-generation 2U2-socket NF5280M6 rack server hardware platform and Lightbits distributed software-defined storage software platform.

- The NF5280M6 server has a very rich configuration to meet the application requirements of various industries. It supports a maximum of 20x 3.5" drives or 39x 2.5" drives, 12x SATA/SAS/NVMe hot-pluggable drives in the front, optional rear M.2/RSSD module to meet diversified storage requirements, and up to 32 DDR4 ECC memory. The memory supports RDIMM and DCPMM, which can provide excellent speed and large capacity memory. It supports optional OCP3.0 module, and provide 10G, 25G, 40G, 100G, 200G multiple network interface options.
- Lightbits distributed storage software supports flexible hardware configuration, flexible expansion of NVMe resource pools, and dynamic migration and expansion of logical volumes. Decoupling storage and

computing improves the efficiency of infrastructure usage and thus meets the needs of cloud and data centers for flexible configurations and elastic expansion of infrastructure.

- **High-bandwidth, high-IOPS, low-latency**

The Lightbits high-performance software-defined storage platform is specially designed and developed to take full advantage of the high performance of NVMe SSDs, leveraging the latest technologies, including NAND flash optimization, Intel® Optane™ persistent memory and Intel® E810 Ethernet Network Adapters with ADQ to provide a high-bandwidth, high-IOPS, low-latency high-performance storage solution.

- This solution uses Inspur NS8610 G1 enterprise-grade NVMe SSDs, and adopts the advanced firmware architecture and micro-channel cooling hardware design independently developed by Inspur Information. It provides 3.5GB/s high-speed bandwidth and 825K 4K IOPS, and has extremely low I/O jitter. Its 4K random read and write IOPS consistency is more than 99%, with 86/14μs read and write low latency, meeting the high-performance and low-latency requirements of data center and enterprise-level application storage.
- The solution is equipped with Intel® Xeon® 3rd Generation Scalable Processors, with increased multi-core density, built-in Intel® AVX-512, Intel® Speed Select Technology and other acceleration technologies, which increase the average performance by 46%¹ compared to the previous generation, ensuring the system High performance computing power.
- This scheme uses Intel® Optane™ persistent memory 200 series as write cache and metadata storage to take advantage of the high-performance, low-latency features of Optane™ persistent memory to increase system bandwidth and IOPS for consistent low-latency.

1 See Index [125] at <https://edc.intel.com/content/www/us/en/products/performance/benchmarks/3rd-generation-intel-xeon-scalable-processors/>.

- The solution uses the Intel® E810 NIC and uses the application device queue (ADQ) function of the NIC to establish a dedicated queue for high-priority application traffic, ensuring stable and consistent low latency and improving throughput.

- **High Data Reliability, Security and High Availability**

The joint solution adopts multiple protection technologies at both the software and hardware levels to enhance data reliability, security and availability.

- Inspur NS8610 G1 enterprise-class NVMe SSD, built-in innovatively designed NAND feature self-learning algorithm model, NAND Flash P/E increased by 40%, greatly extending the life of solid-state drives.
- Lightbits software has a built-in Intelligent Flash Management (IFM) feature that maximizes SSD performance and extends SSD endurance.
- The non-volatility of Intel® Optane™ persistent memory ensures that write cache data will not be lost after power failure, ensuring high system reliability.
- The elastic erasure coding mechanism is used inside the storage nodes to ensure that data will not be lost even if the SSD is damaged; a flexible replication mechanism is implemented across storage nodes in the cluster to ensure that data services are still reliable and available when storage nodes fail.
- A powerful error-tolerance and error-correction mechanism, which can verify and correct silent errors of data; full-path end-to-end data protection, and supports backing up volumes to S3 and restoring volumes from S3.

- **Rich enterprise-class storage capabilities**

The solution fully supports mainstream enterprise-level storage features, including multiple replications and erasure coding, compression and thin provisioning, snapshots, clones, QoS, load balancing, silent data error checking and recovery, SSD fault self-healing, and storage node fault self-healing, multi-tenancy support, etc.

- **Simplified deployment, intelligent operation and maintenance**

This solution is an integrated software and hardware solution, which has been fully verified to ensure compatibility and stability. The solution is based on the NVMe/TCP protocol, supports standard Ethernet deployment methods, supports mainstream operation systems such as Linux and VMware, and can be seamlessly integrated with virtual machines and container orchestration systems such as Kubernetes and OpenStack, without the need to install proprietary software or special drivers. Simple command line operations implement system installation and deployment, provide plug-and-play storage services, and greatly simplify the complexity of deployment.

BMC technology is applied to guide devices through the Web management interface and fault diagnosis LEDs to quickly locate problems, simplify operation and maintenance, and improve system availability. The hardware layer and platform layer monitoring alarms are combined to ensure the stable operation of the machine and reduce the probability of downtime.

Solution Verification: Compatible, stable and excellent performance

In order to verify the application advantages of this solution, Inspur Information, Intel and Lightbits jointly tested product compatibility and performance indicators that users care about most.

The test uses the popular Benchmark tool FIO. The test setup configuration is shown in Table 1.

	Lightbits Storage Server	Client
--	--------------------------	--------

Model	NF5280M6	NF5280M6
BIOS Version	06.00.05	06.00.05
BMC Version	4.21.00	4.21.00
CPU	2*Intel® Xeon® 6338 32Cores@ 2.0GHz	2*Intel® Xeon® Gold 6342 24Cores@ 2.8GHz
Memory	16*16GB DDR4, Total 256GB	32*32GB DDR4, Total 1TB
Intel® Optane™ Persistent Memory	8*128GB , a total of 1TB	without
SSD hard disk	12*1.6TB NS8610 G1 SSD	without
network	1*Intel® E810 100Gb dual- port Ethernet card	1*Intel® E810 100Gb dual- port Ethernet card
OS	CentOS 7.9	Cent OS 7.9
Kernel	4.14.252_001730324769e3ea3 c709_rel_lb	5.12.8-1.el7.elrepo.x86_64
FIO Version		fio-3.7

Table 1 Test server configuration

Test results²

- 1) 4KB data block, 2-replica volume, 100% random read performance test²

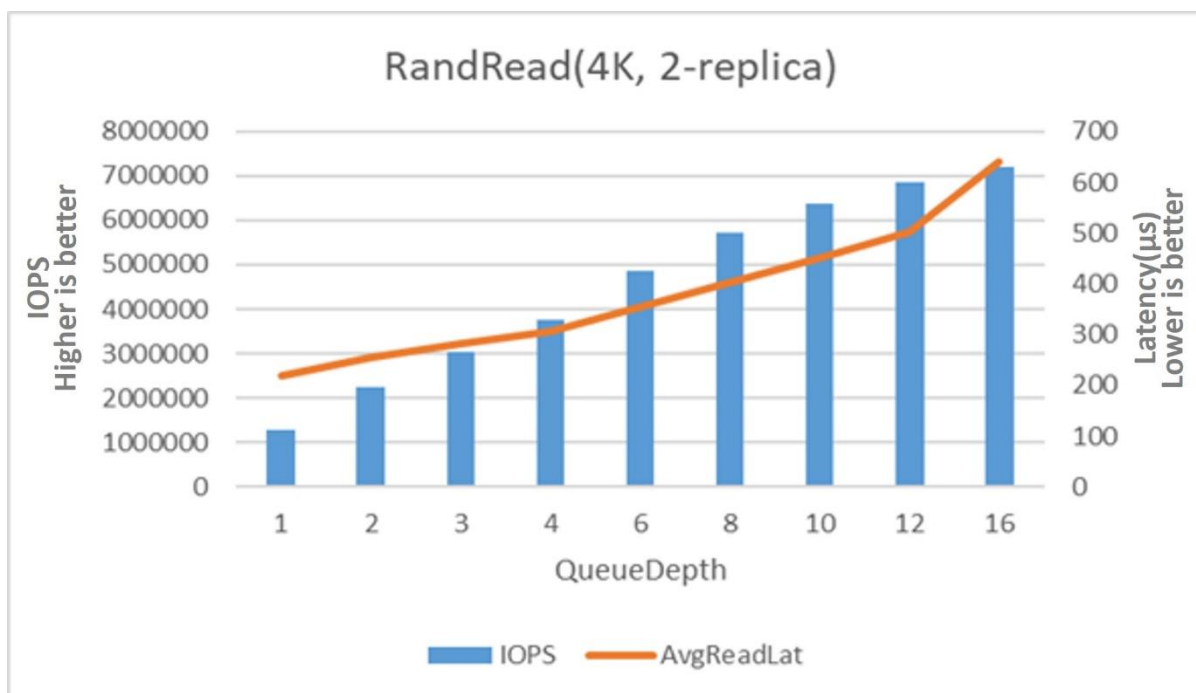


Figure 2 4KB data block, 2-replica volume, 100% random read performance test

2) 4KB data block, 2-replica volume, 100% random write performance test²

² Jointly tested by Inspur, Intel and Lightbits as of 9/6/2022 using FIO version 3.7. See Table 1 for configuration details. Results may vary.

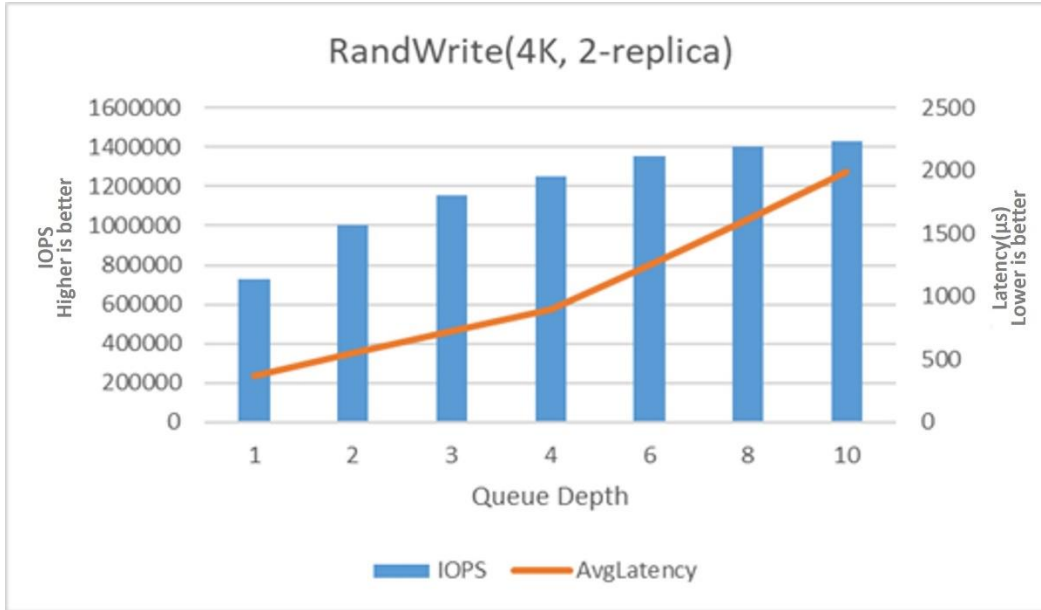


Figure 3 4KB data block, 2-replica volume, 100% random write performance test

3) 4KB data block, 2-replica volume, 70/30 random read write mix performance test²

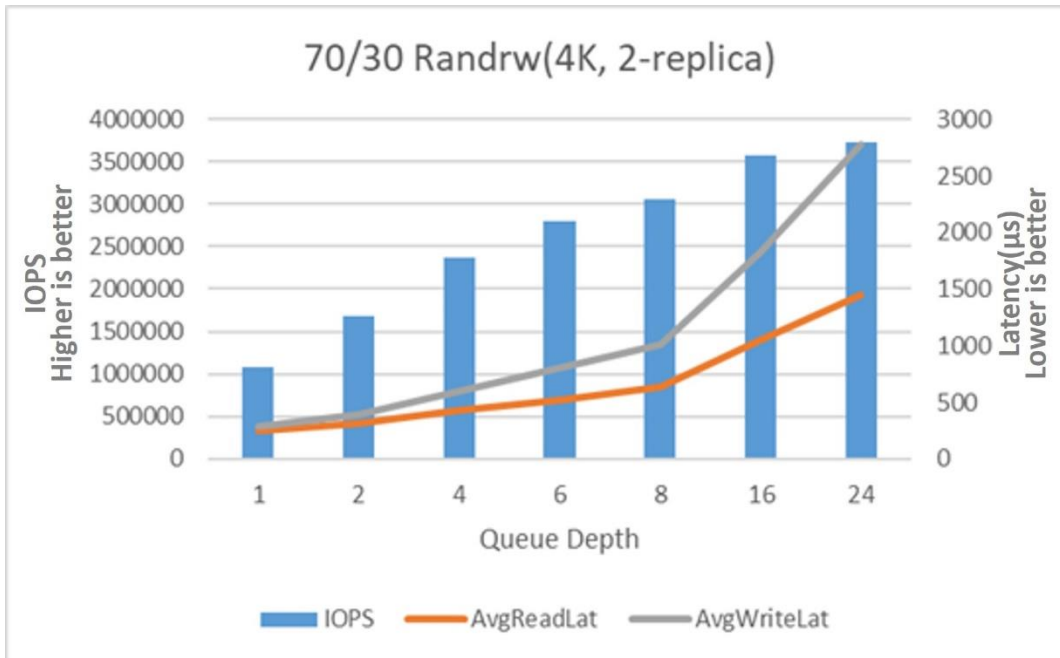


Figure 4 4KB data block, 2-replica volume, 70/30 random read write mix performance test

2 Jointly tested by Inspur, Intel and Lightbits as of 9/6/2022 using FIO version 3.7. See Table 1 for configuration details. Results may vary.

4) 4KB data block, 2-replica volume, 50/50 random read write mix performance test²

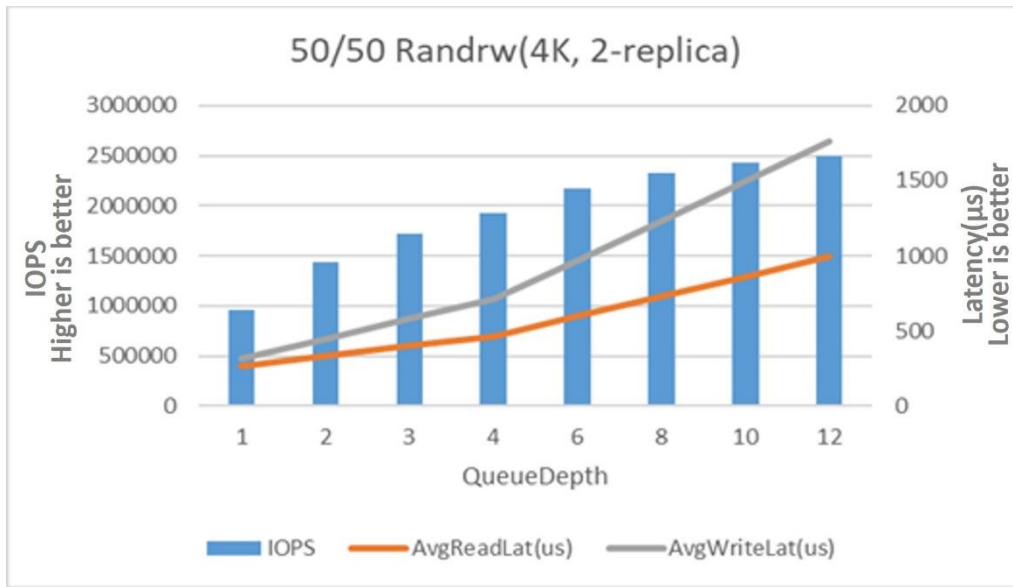


Figure 5 4KB data block, 2-replica volume, 50/50 random read write mix performance test

Detailed test data refer to Table 2²:

Test items	100% randread	100% randwrite	70/30 randrw	50/50 randrw
4Krandom IOPS	7202K	1430K	3726K	2551K
4K Typical random I/O delay	6866K (at QD=12)	1001K (at QD=2)	2800K (at QD=6)	1924K (at QD=4)
4K Average delay (us)	502	551	read: 525 write: 808	read: 463 write: 712
4K tail latency: 99% (us)	798	1155	read: 1105 write: 1499	read: 985 write: 1444

Table 2: 4K data block random read/write mix performance test results

² Jointly tested by Inspur, Intel and Lightbits as of 9/6/2022 using FIO version 3.7. See Table 1 for configuration details. Results may vary.

The test results show that:

- In the randread test using 4KB I/O size, a cluster with three storage servers can achieve more than 7 million randread IOPS, approaching the limit of 100Gb network bandwidth, while maintaining very low latency. Higher performance can be achieved by increasing the number of clients in a production environment. According to the Lightbits reference configuration, if each server is configured with two 100Gb network cards, then the three server clusters can provide a total random read performance of about 14 million IOPS.
- The 100% random write performance and mixed read and write performance of the cluster consisting of three storage servers were excellent too, while maintaining stable and consistent low latency.
- After sufficient functional and stability tests, it shows that Inspur NF5280M6 server is well compatible with Lightbits software-defined storage system, and all functions are normal and stable. The reliability and high availability of the entire environment are outstanding under long-term operation.

According to the comparison test result between Lightbits and Ceph which is delivered by the third-party testing agency Evaluator Group, in the same hardware environment, the performance of Lightbits is 2-16 times higher than that of Ceph. For details, please refer to the ESG report <https://www.lightbitslabs.com/resources/run-apps-up-to-16x-faster/>

User Value

Inspur Information and Intel, together with Lightbits Labs, leverage

their advantages and cooperate strongly to build NVMe-based Inspur servers equipped with Intel® Optane™ persistent memory /TCP all-flash solution. Inspur Information and Lightbits distributed all-flash storage solution utilizes persistent storage and end-to-end NVMe/TCP and other new technologies to unleash the excellent performance of flash memory and achieve better flexible expansion capabilities. It is a better choice for high-performance storage demanding applications, such as database, data analysis, cloud infrastructure, etc.

The joint solution is seamlessly compatible with various orchestration platforms such as OpenStack, Kubernetes, and VMware. It provides a fully standardized client driver, and truly realizes the plug-and-play solution. With the Inspur Information and Lightbits distributed all-flash storage solution, cloud-native applications can achieve ultra-high performance, and cloud data centers can reduce TCO and help enterprises achieve digital transformation and upgrading.

References

1. Inspur NF5280M6 server
<https://www.inspur.com/lcjtww/2509917/2509945/2552345/index.html>
2. Inspur NS8610 G1 SSD hard drive
<https://www.inspur.com/lcjtww/2526546/2526538/2526549/2563157/index.html>
3. Lightbits Labs: Official website <https://www.lightbitlabs.com/>
4. ESG Lab Report: Run Apps up to 16X Faster: Storage Performance Comparison Lightbits vs. Ceph Storage
<https://www.lightbitlabs.com/resources/run-apps-up-to-16x-faster/>
5. Xeon® Scalable Processors
<https://www.intel.cn/content/www/cn/en/products/details/processors/xeon/scalable.html>

6. Intel® Optane™ Technology

<https://www.intel.cn/content/www/cn/en/architecture-and-technology/optane-technology/optane-for-data-centers.html>Intel®

7. Ethernet Network Adapter E810

<https://www.intel.cn/content/www/cn/en/products/details/ethernet/800-network-adapters.html>

Notices & Disclaimers

Performance varies by use, configuration and other factors. Learn more on the [Performance Index site](#).

Performance results are based on testing as of dates shown in configurations and may not reflect all publicly available updates. See backup for configuration details. No product or component can be absolutely secure.

Your costs and results may vary.

Intel technologies may require enabled hardware, software or service activation.

© Intel Corporation. Intel, the Intel logo, and other Intel marks are trademarks of Intel Corporation or its subsidiaries. Other names and brands may be claimed as the property of others.

Lightbits and Lightbits Labs are registered trademarks of Lightbits Labs, Ltd.

The NVMe®, and NVMe®/TCP word marks are registered or unregistered service marks of the NVM Express organization in the United States and other countries. All rights reserved. Unauthorized use strictly prohibited.

All trademarks and copyrights are the property of their respective owners.