Deploying Reliable and High-Performance Storage with LightOS

The first Software-Defined NVMe/TCP Storage Solution with Clustering Support

Table of Contents

The Cloud-Native Block Storage Revolution

When deploying cloud-native applications at any scale, access to reliable, performant and easily deployable infrastructure is key to business agility and application experience, regardless of whether you deploy five nodes, five hundred nodes, or millions of nodes.

It is well known that hyperscale Cloud Services Providers (CSP's) such as Facebook, Google, and Amazon Web Services (AWS) cracked the code of infrastructure efficiency because it was critical to their business. What's less well known is how they achieved that level of hyper-efficiency without compromising on any of the following:

- Performance, low latency, and scalability
- Flexibility and efficiency of deployed storage
- Easily adaptable to operate at cloud scale
- Compatibility with existing data center infrastructure without having to rip out and replace

The above four criteria are core requirements of any cloud-native storage solution. The challenge is that building such a storage solution is a complex and daunting task due to the engineering talent required.

The good news is that now this barrier was removed with Lightbits Labs LightOS software- defined storage with clustering support. LightOS sets a new standard for cloud-native block storage deployments offering the next step for your infrastructure deployment.

Achieving Cloud Scale Reliability

Reliability is a key component of a cloud-native block storage solution, which must also be reliable despite being built out of unreliable components such as standard COTS servers, NICs, and NVMe SSDs. A cluster of SSD servers must replicate data internally and keep it fully consistent and available in the presence of failures.

From the perspective of clients accessing the data, data replication has to be transparent, and server failover must be seamless. Furthermore, any cloud-native block storage solution should reduce application latencies, boost application performance, increase scalability, and provide the reliability of traditional storage area network (SAN) with existing data center infrastructures.

Big cloud providers incredible scaling and reliability is not achieved by building one large system that grows indefinitely but rather by constructing larger systems out of many loosely aggregated smaller systems. Similarly, any cloud-native storage solution performs best when it is made out of many smaller systems, loosely aggregated and managed as one.

The problem for cloud providers is that everything fails. Hardware fails. Software fails. Everything fails.

In the old days, we used to build storage area networks (SANs) out of highly-available and reliable components and by carefully writing bug-free software, to the extent possible.

In the cloud, you achieve reliability by accepting that everything will fail and constructing reliable systems out of unreliable components.

The LightOS® Solution

LightOS®, the Lightbits Labs™ software-defined storage solution, offers a revolution in disaggregated storage for cloud and data center infrastructures with a software-only or software with hardware acceleration solution.

LightOS unlocks the potential of a disaggregated storage, using high-performance NVMe solution to maximize resource utilization and flexibility. It enables data centers to move from an inefficient direct-attached SSD model to a shared model in which compute and storage can be scaled independently cost effectively. But it does so without compromising reliability by offering clustering capabilities.

Installed on standard servers in large scale data centers, LightOS is optimized for I/O intensive compute clusters, such as Cassandra, MySQL, FDB, MongoDB, and Time Series Databases. It delivers high performance and consistently low latency based on patent-pending multi NVMe SSD management and the innovative NVMe over TCP (NVMe/TCP) storage communication protocol. LightOS NVMe/TCP does not require any changes to networking infrastructure or application server software. LightOS can deliver millions of IOPS with latencies that can be as low as tens of microseconds.

LightOS is ideal for containerized environments such as Kubernetes that require large scale clusters with persistent storage for rapid node migration, workload rebalancing, or recovery from failures without copying data over the network.

LightOS clustering maximizes resource utilization and flexibility and provides high availability and durability with a configurable number of replicas per volume. A LightOS cluster can be composed of servers residing in different failure domains and located anywhere within the data center.

In addition, LightOS provides rich storage services such as data reduction, thin provisioning for best utilization of SSD capacity, and replication and Erasure Coding (EC) for protection from SSD and server failures to ensure service and data availability under failures.

Preventing Data Loss with LightOS Clustering Support

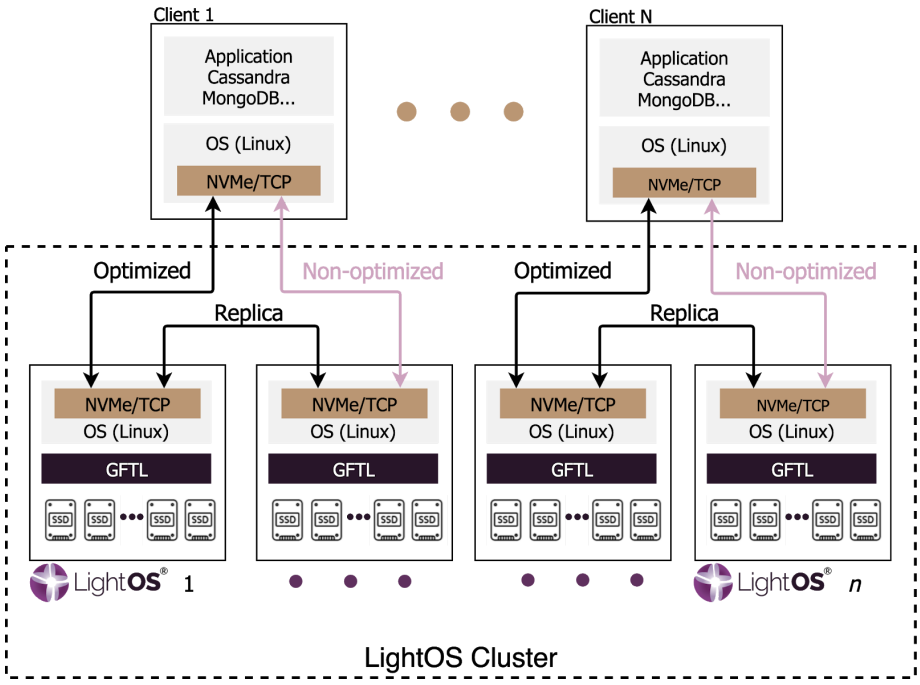
Typically to prevent data loss, software-defined storage solutions replicate data across different storage servers, guaranteeing service, and data availability whenever storage servers experience transient or permanent failures. Remember, at cloud scale, everything fails.

A cluster of LightOS servers replicates data internally and keeps it fully consistent and available, thereby guaranteeing service reliability and data availability even if one or several storage servers fail sporadically (for a limited time) or permanently (forever).

From the perspective of clients accessing the data, LightOS clustering architecture is transparent, and server failover in the event of failure is seamless. Data remains consistent and available. Clients do not require any special drivers or agents beyond the standard NVMe/TCP drivers now available in all major Linux distributions. With LightOS clustering support, client CPU cycles are used for applications and not for storage processing.

A LightOS cluster can be composed of servers residing in different failure domains and located anywhere within the data center. The architecture maximizes resource utilization and flexibility and provides high availability and durability with a configurable number of replicas per volume.

Figure 1: LightOS Clustering Multipath Replication Design

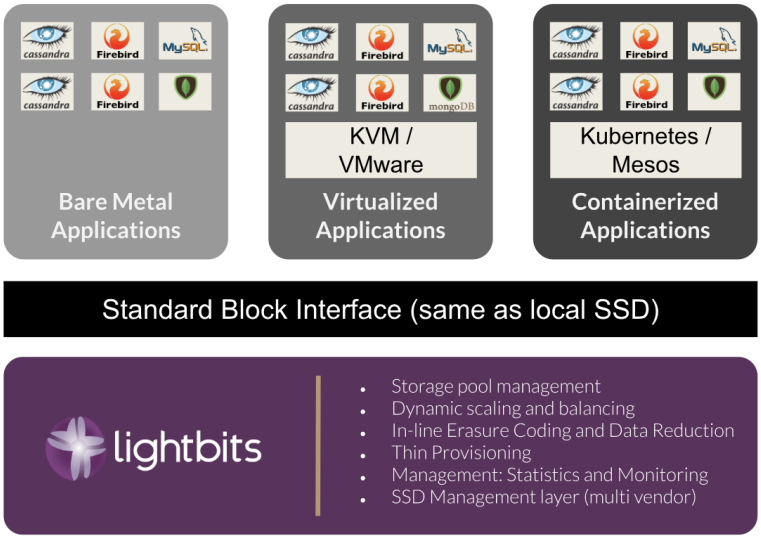


In addition to its dependability, LightOS delivers for cloud data centers the same input/output operations per second (IOPS) as direct-attached NVMe SSDs and up to a 50% reduction in tail latency compared to direct-attached storage (DAS). It also enables scaling storage and compute independently, which drives up infrastructure utilization and efficiency, reduces total cost of ownership (TCO), and provides unparalleled flexibility.

By providing a highly available and durable storage layer, application teams can focus their efforts on developing new services while LightOS guarantees availability and high-performance access to data.

LightOS clustering support enables non-disruptive maintenance routines that temporarily disable access to storage servers (e.g., TOR firmware upgrades). In such cases, clients continue working transparently with a subset of the storage servers in degraded mode until the maintenance routine completes, and all storage servers become accessible again.

Figure 2: LightOS Block Diagram



LightOS Deployment

At its core, LightOS is a software-defined disaggregated storage solution that is a natural evolution of your existing data center infrastructure.

LightOS uses the NVMe/TCP block storage access protocol for the communication links between application server-to-storage server, and storage server-to-storage server. NVMe/TCP runs on any TCP/IP network and is included, or will soon be included in every operating system and hypervisor.

The software only needs to use the standard NVMe/TCP drivers, leading to simple and easy deployment on existing application servers and with existing TCP/IP networks. Also, multiple LightOS clusters can exist in the same cloud data center, which can be easily aggregated and managed as one large cloud-scale block storage solution.

LightOS runs on commodity hardware built with off-the-shelf components and products. It offers a durable solution implementation without a single point of failure on a data and control paths to help cloud vendors keep their service up at all times.

Additional Advantages

In addition to its reliability, LightOS with clustering support provides several advantages for building cloud-native block storage infrastructure:

Performance and Latency

The prevalent storage architecture in today's clouds uses direct-attached hard disk drives and SSDs. In particular, NVMe SSDs that provide high performance and low latencies at reasonable costs.

Any cloud-native block storage solution cannot reduce performance or increase latencies beyond the level that applications currently experience.

LightOS provides applications with the same performance levels and comparable latencies as the direct-attached architecture.

Flexibility and Efficiency

Current direct-attached storage deployments often suffer from low utilization and lack of flexibility. Low usage means that purchased infrastructure is sitting idle. Lack of flexibility means that the operational expenses are higher than they should be because it's hard to adapt to changing workload conditions and increase either storage or compute capacities as needed—with direct-attached, you must always scale both together.

Low efficiency and rigid infrastructure are both contributors to a high TCO.

LightOS uses a disaggregated storage model that provides flexibility and efficiency by separating compute and storage. This increase in efficiency results in lower TCO for your infrastructure.

Cloud Scale

Clouds achieve their incredible scaling not through building one large system that grows indefinitely but rather from constructing larger systems out of many smaller systems that are loosely aggregated.

LightOS clustering support builds storage clusters out of many smaller storage servers, loosely aggregated and managed as one.

Conclusion

For enterprises and cloud providers, the time has come to build cloud-native block storage clusters.

The LightOS clustering architecture provides your infrastructure with increased flexibility, efficiency, and reliability while offering better latency and lower overall TCO. LightOS clusters help you manage one or more server failures without any loss of service or data. LightOS clustering support will be available for production deployments in late 2019.

About Lightbits Labs™

Cloud-scale data centers require disaggregation of storage and compute, as evidenced by the top cloud giants' transition from inefficient Direct-Attached SSD architecture to low-latency shared NVMe flash architecture.

The Lightbits team, who were key contributors to the NVMe standard and among the originators of NVMe over Fabrics (NVMe-oF), now bring you their latest innovation: the Lightbits LightOS® NVMe/TCP-based software defined disaggregated storage solution.

In stark contrast to other NVMe-oF solutions, Lightbits LightOS separates storage and compute without touching the network infrastructure or data center clients. With NVMe/TCP, LightOS delivers the same IOPS as direct-attached NVMe SSDs and up to a 50% reduction in tail latency.

The transition from direct-attached SSDs to NVMe/TCP with LightOS is so smooth your applications teams won't even notice the change. Once they switch, they are freed to scale storage and compute independently, with consistently better user experience, and with lower overall costs. As the trailblazers in the disaggregated storage space, Lightbits LightOS is already successfully deployed in industry-leading cloud data centers.

Finally, you too can separate storage from compute without drama.

www.lightbitslabs.com info@lightbitslabs.com

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

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

3 Habankim Street,
Haifa, Israel 3326115