Supercharged MySQL Database Using Lightbits™ LightOS® Reduces Database by over 40% and Tail Latency by 30%!

MySQL is one of the most widely deployed relational database management engines in the world.

To meet the needs of modern cloud-scale architectures running data-hungry applications system administrators have resorted to installing high-performance NVMe Flash storage directly into their MySQL compute nodes. This direct-attached storage (DAS) model introduces expensive trade-offs. The storage is now tied to the compute node preventing independent scaling of the CPU & Storage.

Today, thanks to Lightbits™ LightOS® Software Defined Storage (SDS) solution, compute and storage can be easily disaggregated for each MySQL compute node. This enables high performance without stranding capacity and wasting storage within each node.

We benchmarked a MySQL DAS configuration versus a MySQL disaggregated solution using Lightbits in a head to head comparison. Our results showed a 2x reduction in tail latency while delivering the same transactional performance.

Next, we enabled compression on LightOS, and reduced the database size by 40% —while delivering 30% lower latency than the DAS configuration with no compression! Low tail latencies are critical for enabling cloud-scale workloads.

LightOS provides this improvement to the quality of the IO while also including data services such as Data Reduction, Thin Provisioning, and Erasure Coding through the LightOS Global FTL.

The Lightbits LightOS SDS Solution

Installed on standard servers in large scale data centers, LightOS is the Lightbits SDS solution. It runs on standard Linux servers and is optimized to deliver high performance and consistently low latency for I/O intensive compute clusters, such as Cassandra, MySQL, MongoDB, FDB and time series databases.

LightOS is a patent pending multi-NVMe SSD management system and NVMe over TCP (NVMe/TCP) storage target that requires no changes to the networking infrastructure or application servers.
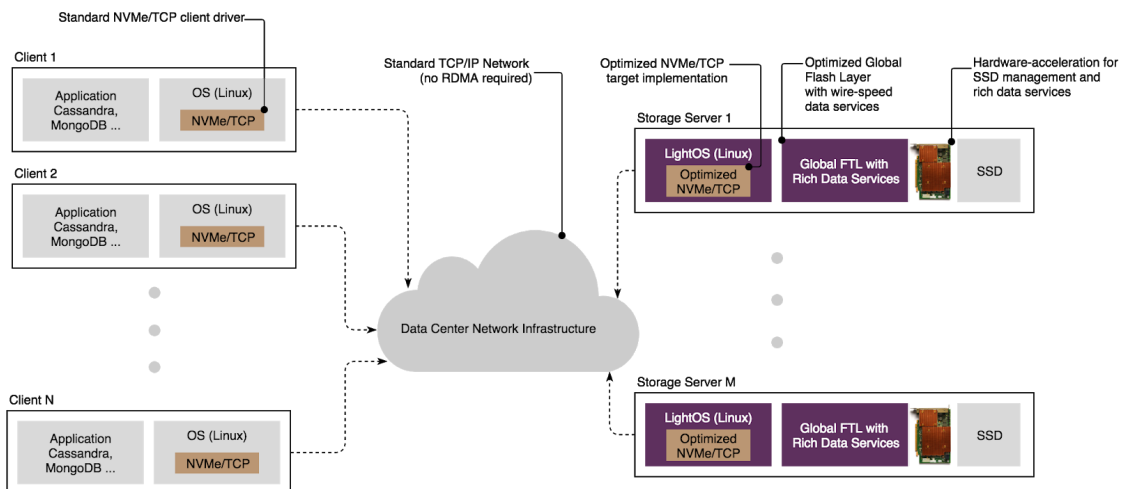
LightOS uses a pool of storage provisioned into volumes as a unique NVMe namespace. Clients connect with NVMe/TCP client-side software that is freely available from operating system software repositories.

LightOS is designed to disaggregate storage from compute. It provides to you:

- **Increased Application Performance:** When MySQL updates millions of records, the LightOS Global FTL operating on a pool of flash with inline data reduction dramatically reduces tail latency. LightOS can work in software-only mode or in concert with the optional LightField(™) hardware accelerator which reduces latency further.
- **Standard, Simple, Secure, Storage Access:** Full support for standard NVMe-oF using standard NVMe connection protocols and Access Control List functionality. Using standard NVMe-oF mechanisms, storage administrators can set Access Control Lists for each NVMe-oF volume exported by LightOS. LightOS will only let NVMe/TCP connected clients access volumes that the client has access permissions for.
- **Software-Defined Storage (SDS) Solution:** Install LightOS on your own storage servers.
- **Reduced Total Cost of Ownership:** Utilize all of your storage using Thin Provisioning and only buy more storage when you need it. Avoid stranded capacity and limits to scalability.
- **Separate Storage and Compute:** Only upgrade the storage servers or the application servers when you need to. No need to bundle and waste storage or CPU capacity.
- **Simple, Safe Data Management:** Data is protected from SSD failures using a highly optimized Erasure Coding (EC) algorithm managed in concert with the high-performance LightOS Global FTL, streamlining data management across the pool of SSDs, increasing performance and ensuring data is safe.

MySQL Benchmarking Environment

First, MySQL benchmarks were run using a typical cloud-based workload where each MySQL node contained NVMe SSDs - the DAS configuration. A dedicated Stress Client machine ran the benchmark workload on each MySQL compute node. Then, the same Stress Clients ran the same benchmark workload on the same MySQL nodes, this time with the MySQL nodes using disaggregated storage residing in a Lightbits LightOS storage server and accessed over NVMe/TCP.

Stress Client and MySQL Node Hardware Details

Component	Stress Client	MySQL Node
Description	A compute node with Yahoo! Cloud Serving Benchmark (YCSB) installed.	A server with the MySQL application running on it and with direct attached storage (in DAS benchmark runs)
Memory	64GB	64GB
CPU	Intel ® Xeon ® E5-2603 v4 @ 1.70GHz	Intel® Xeon® Gold 5120 CPU @ 2.20GHz
Network Card	Mellanox ® ConnectX-4™ LX 25GbE	Mellanox® ConnectX®-5
Storage	N/A	2x Intel 4510 2TB NVMe SSDs (4 PCIE gen 3 lanes / SSD)

LightOS Storage Server

| Component	Description
Memory	512GB
CPU(x2)	Xeon(R) Gold 5120 CPU @ 2.20GHz (56HT) (SkyLake)
Network Card(x2)	Mellanox® ConnectX®-5
Storage	16x Intel 4510 2TB NVMe SSD (4 PCIE gen 3 lanes / SSD)

Benchmarking Methodology

MySQL v5.7 was installed on bare metal servers running Linux (Ubuntu 16.04). MySQL DAS nodes contained two NVMe SSDs configured using standard Linux based Software RAID-0 to utilize the full capacity and IOPS available on both devices.

The Yahoo! Cloud Serving Benchmark (YCSB) tool was executed on the Stress Clients, connected to the MySQL nodes over a standard TCP/IP network infrastructure.

Multiple MySQL servers with Direct Attached Storage

When benchmarking MySQL using LightOS, there were two LightOS nodes configured on the Lightbits storage server. This configuration enabled testing a scaled up MySQL configuration.
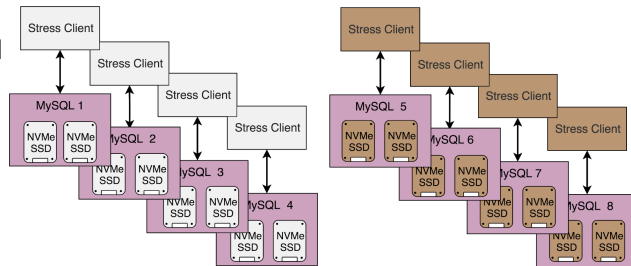
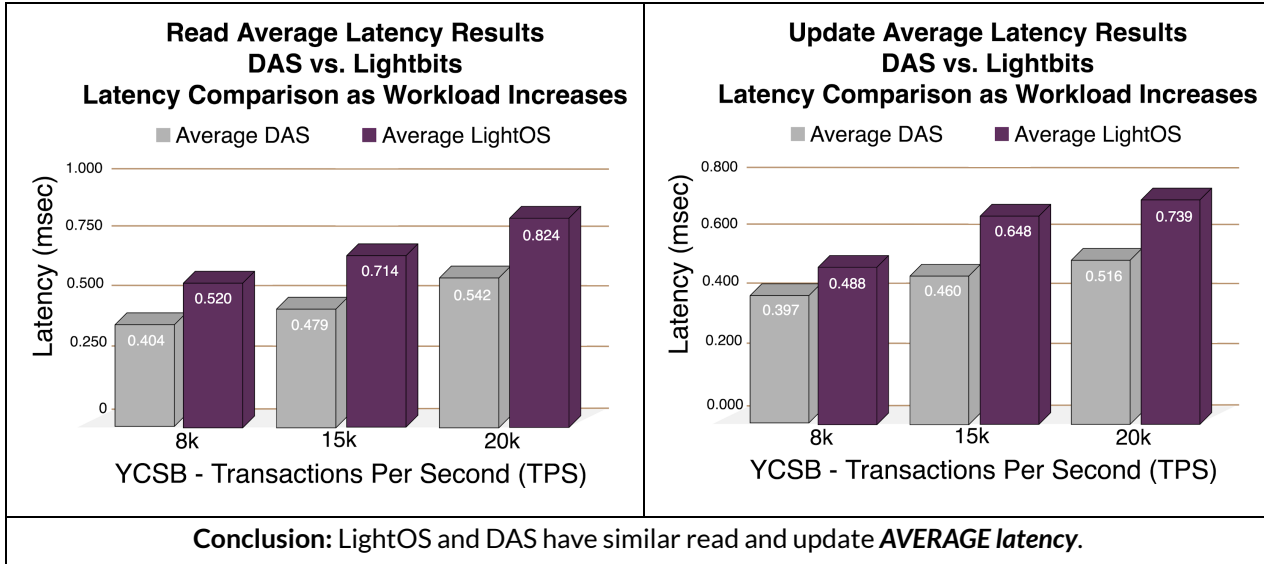**Multiple MySQL servers disaggregated to LightOS SDS Platform**
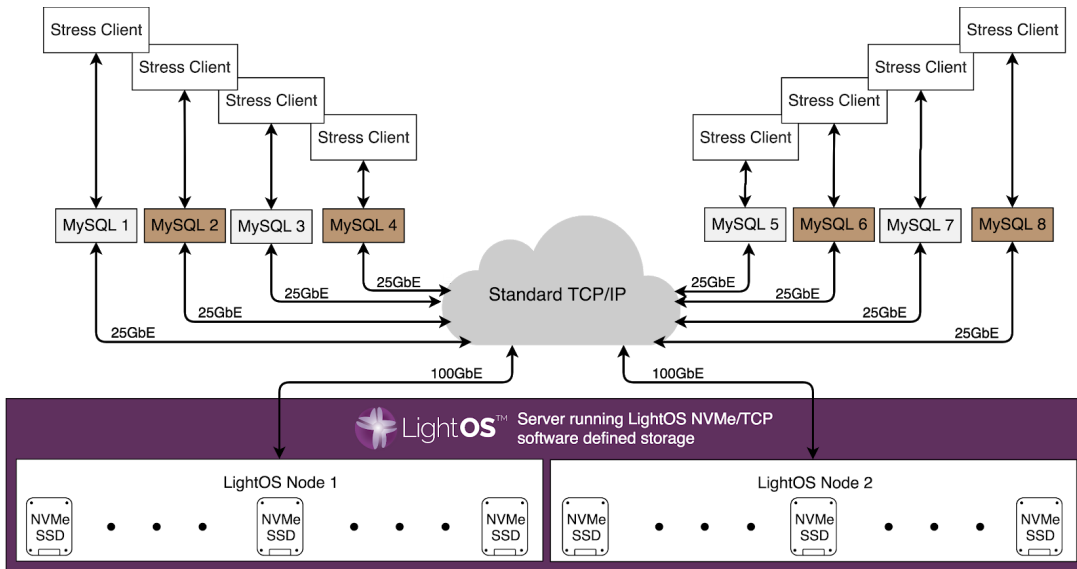
In both DAS and LightOS configurations, the YCSB benchmarking software was configured to run on a MySQL database where the database size was calculated to consume ~60% of the usable capacity. The YCSB workload was a mix of 50% reads and 50% updates, with 4K record sizes using a uniform distribution.

Benchmark Results

We compared the performance of a MySQL DAS based compute node to a MySQL compute node using storage on Lightbits LightOS by pushing the MySQL nodes until they achieved the maximum Transactions Per Second (TPS) rate per node. In other words, we found the performance saturation point for a MySQL compute node using DAS and using Lightbits.
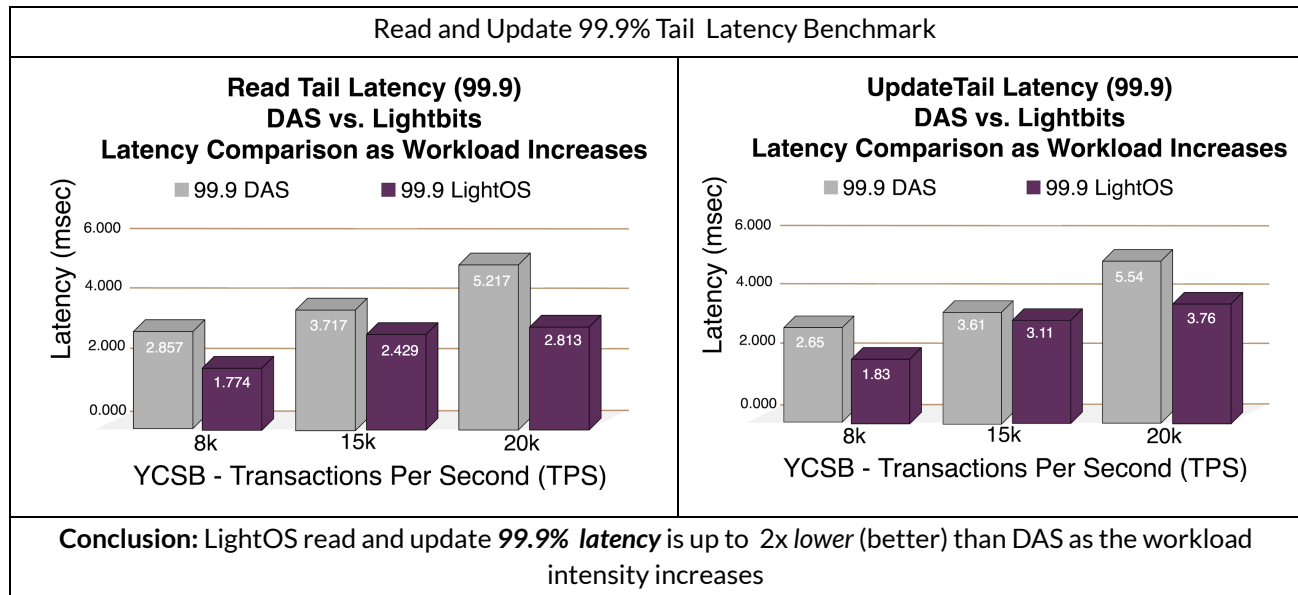
As seen in the benchmark results, when looking at the average latency, LightOS performs similar to DAS, within a few 100 microseconds of DAS, despite completing all I/O operations over a TCP/IP network.

Read and Update Average Latency Benchmark

3

**Read Average Latency Results
DAS vs. Lightbits
Latency Comparison as Workload Increases**

**Update Average Latency Results
DAS vs. Lightbits
Latency Comparison as Workload Increases**

Conclusion: LightOS and DAS have similar read and update *AVERAGE latency*.

An important metric for large scale cloud deployments is the tail latency, defined as the time it took to complete the I/O requests that required the most time to finish. When looking at tail latency, as the workload intensity increases, LightOS provides lower (better) latencies.
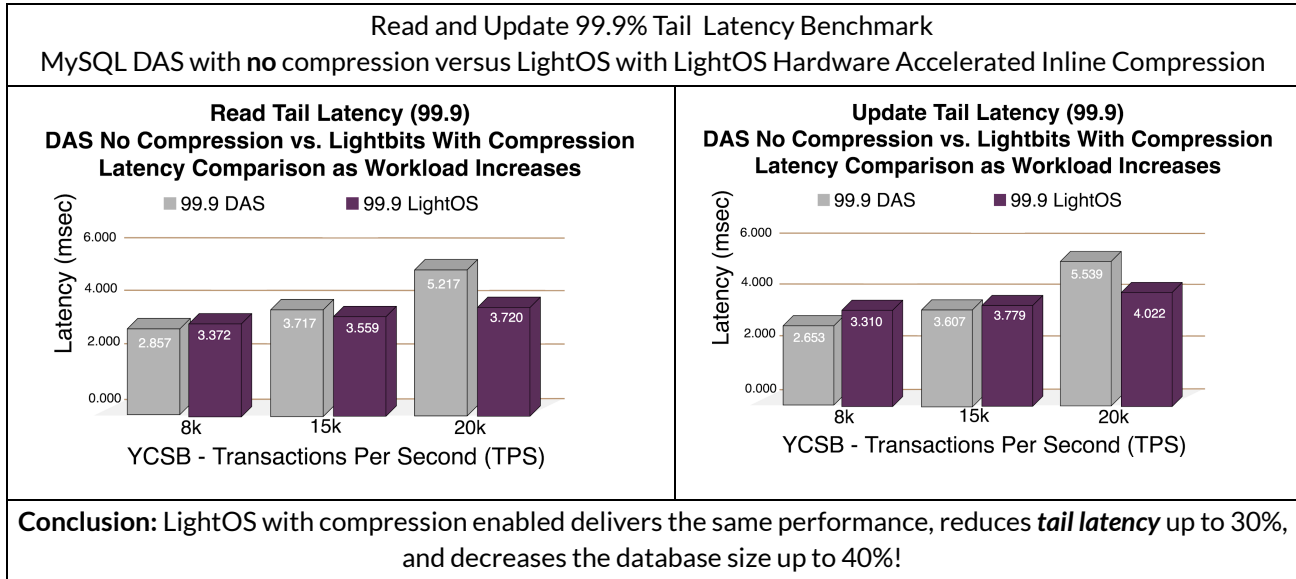
The benchmark results show that DAS needs thousands of microseconds longer to complete IOs compared to LightOS.

Read and Update 99.9% Tail Latency Benchmark

**Read Tail Latency (99.9)
DAS vs. Lightbits
Latency Comparison as Workload Increases**

**UpdateTail Latency (99.9)
DAS vs. Lightbits
Latency Comparison as Workload Increases**

Conclusion: LightOS read and update *99.9% latency* is up to 2x *lower* (better) than DAS as the workload intensity increases

The LightOS Global FTL was designed to deliver consistently low latencies. As the workload intensity increases, the DAS-based MySQL node cannot smooth out the tail of the latency.

The GFTL further improves tail latency and reduces wear on the SSDs when the data is compressed by LightOS using the LightField hardware accelerator. Our benchmarks show that MySQL data and the metadata used by MySQL is highly compressible. We first measured performance using MySQL's built-in software option, but the software based MySQL algorithm drove down the performance on each node dramatically. With MySQL software compression enabled, each node was only able to deliver 3000 TPS.

4

LightOS performance was unaffected using the LightField hardware acceleration option. The total database size was reduced by 40% using LightOS + LightField. Average latency for LightOS was the same as that measured in our previous run without LightOS compression enabled. However, once again the tail latency was lower than DAS as the workload intensity increased—even with compression enabled on LightOS.

Read and Update 99.9% Tail Latency Benchmark
MySQL DAS with no compression versus LightOS with LightOS Hardware Accelerated Inline Compression

Read Tail Latency (99.9)
DAS No Compression vs. Lightbits With Compression
Latency Comparison as Workload Increases

■ 99.9 DAS   ■ 99.9 LightOS

YCSB - Transactions Per Second (TPS)	99.9 DAS	99.9 LightOS
8k	2.857	3.372
15k	3.717	3.559
20k	5.217	3.720

Latency (msec)

Update Tail Latency (99.9)
DAS No Compression vs. Lightbits With Compression
Latency Comparison as Workload Increases

■ 99.9 DAS   ■ 99.9 LightOS

YCSB - Transactions Per Second (TPS)	99.9 DAS	99.9 LightOS
8k	2.653	3.310
15k	3.607	3.779
20k	5.539	4.022

Latency (msec)

Conclusion: LightOS with compression enabled delivers the same performance, reduces *tail latency* up to 30%, and decreases the database size up to 40%!

Again, as the workload intensity increases LightOS delivers lower tail latency than DAS. The tail latency improves when you compare MySQL running on DAS, with no software compression, versus MySQL running on LightOS, with LightOS compression enabled. Further, MySQL running on LightOS delivers the same performance as DAS and reduces the database size by 40%!

Conclusion

The Lightbits LightOS SDS solution enables disaggregation of storage and compute for MySQL workloads. It can supercharge your cloud database applications by providing:

- Increased performance and return on investment with a lower total cost of ownership
- Improved service levels and a better user experience with dramatically reduced tail latency
- Reduced database size without compromising performance and enjoying lower tail latency
- Standard, simple, secure storage access to your application servers
- Full utilization of your existing TCP/IP network with no changes to application servers

No more managing a growing number of database servers with their own underutilized, underperforming storage inhibiting simple and efficient cloud-scale architecture.

Finally, you can separate storage from compute without all the drama.

For more information about the Lightbits solution, contact info@lightbitslabs.com.

5
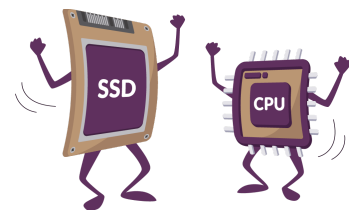
About Lightbits Labs™

The Lightbits NVMe/TCP solution is the latest innovation from the Lightbits team, who were key contributors to the NVMe standard and among the originators of NVMe over Fabrics (NVMe-oF). Unlike other NVMe-oF approaches, the Lightbits NVMe/TCP solution separates storage and compute without touching the network infrastructure or data center clients. With NVMe/TCP, Lightbits delivers the same IOPS as direct-attached NVMe SSDs and up to a 50% reduction in tail latency. The transition is so smooth your applications teams won't even notice the change. Once they switch to NVMe/TCP, they can scale storage and compute independently and with consistently better user experience.

www.lightbitslabs.com   info@lightbitslabs.com

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

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

Haifa
  3 Habankim Street,
  Haifa, Israel 3326115