Lyve Rack: Mass Capacity Object Storage

Open, affordable, and durable object storage solution—ideal for high-capacity, high-availability needs.

Solution Summary

Lyve™ Rack is a high-capacity, high-availability object storage solution built for fast and easy deployment to better manage growing volumes of data. An open reference architecture, Lyve Rack can store large amounts of data—both structured and unstructured. This tightly integrated solution bundles Seagate storage, Dell and SuperMicro servers, and Mellanox Technologies networking, along with many third-party tools and applications.

Lyve CORTX™, open source object storage software supported by Seagate, is included in the Lyve Rack solution. Customers receive all the benefits of open source, as well as the ease of deployment of turnkey object storage systems.

Benefits Summary

- Store and access data for modern, leading-edge workloads
- Simpler, easier deployment
- High-capacity, high-availability object storage solution
- Flexibility and transparency with open source object storage software

Leading-edge workloads and applications require tremendous amounts of data to be effective.

Technologies like the Internet of Things (IoT), edge computing, big data, and artificial intelligence (AI) are proliferating—all while demand for consumer endpoint devices is rising fast. These are just two of many factors that are driving the explosion of data. Two concepts help us better understand data proliferation: growth and sprawl. Data growth refers to the percentage by which the datasphere increases over time. Sprawl describes the spread of this growing data throughout various configurations—from endpoint to edge to cloud. Data growth has been, and will continue to be, unprecedented in volume, whereas data sprawl represents how data is scattered.

We live in an exciting era of innovation and opportunity. At the same time, we live in an era of unprecedented challenges. Advances in AI are unlocking opportunities and solutions to critical challenges in almost all fields of human endeavor. Whether it's improving our health with custom DNA diagnoses or making roads safer with autonomous vehicles, data, more than ever, is poised to transform our world in unimaginable ways.

The Challenge

With so much data generation, organizations are struggling to leverage the full extent of their data for business insights and better decision making. This is primarily because of the difficulty associated with storing and accessing growing data. Organizations are tasked with managing the data they currently have, as well as the large volumes of data they will continue to receive. And if they are unable to tap into that data, they will miss out on opportunities to better understand customers, obtain competitive advantages, prevent fraud, and optimize supply chains. Smart management of the ever-growing troves of information requires that organizations understand how and where their data is gaining in volume.

Many of the leading-edge technologies that businesses are using require enormous amounts of data to be effective. A key driver of AI enhancement has been innovations in machine learning (ML)—the technique by which machines are trained to act autonomously—and corresponding innovations in analytics, which enable us to discover hidden insights. But the use of existing storage software infrastructure and parallel file systems delays time to insights. Object storage, on the other hand, can support many of these modern workloads by handling unstructured data and retrieving data faster—all in a more cost-effective manner.

That said, object storage software needs an ideal hardware and infrastructure platform to leverage its full potential. Currently, many object storage solutions are difficult to deploy, proprietary, and cost prohibitive.

As the cost of data storage continues to rise, bytes of data remain difficult to access. Converting mass capacity hard drives into consumable data in data centers requires a tremendous amount of overhead. In fact, costs are so high that the International Data Corporation (IDC) predicted less than 10 percent of the 175ZB of data produced in 2025 will actually be stored with today's storage technology. This untapped data represents an enormous amount of lost opportunity.

The Solution

Object storage has innovative characteristics that make it highly suitable for customers deploying modern workloads that require lots of data, such as AI, big data, high-performance compute (HPC), and edge data centers.

Customers need affordable object storage solutions that are easier to configure; ones that use denser storage, have a tighter integration of software and hardware, and are based on open standards.

Lyve Rack is an object storage, converged infrastructure solution that solves many data management challenges by storing large amounts of data—both structured and unstructured. The integrated software and hardware bundle includes storage, compute, networking, and object storage, along with many third-party tools and applications. Lyve CORTX was co-designed by Seagate and development partners who validated existing concepts and suggested new features, yielding an open, community-driven approach.

Lyve CORTX is 100% open source, with the core utilizing Apache Licensing Version 2—one of the least restrictive open source licensing models. Lyve Rack democratizes data access by making object storage software open and affordable, enabling organizations to deploy more storage so they can store and access more of their data.

Artificial Intelligence and Machine Learning

AI and ML applications need a tremendous amount of data to be effective. With larger data sets for AI/ML training models, inference engines can provide deeper levels of insight. Lyve Rack's high-density, high-capacity storage enhances this process. Further, Lyve CORTX was designed from the ground up for AI/ML—with research and scientific organizations contributing many cutting-edge use cases, as well as validating and suggesting new features to support AI/ML requirements.

Big Data

Big data is characterized by lots of unstructured data that is difficult to manage and costly to access. Object storage can store unstructured data in native formats, making it available for use in big data analytics, as well as for creating data lakes and supporting newer tools and technologies.

HPC

HPC environments are capacity intensive, with users requiring fast access to large amounts of data, high availability, and stringent data protection. This makes HPC environments ideal for object storage and Lyve Rack.

Total Solution

Lyve Rack solves many challenges around deploying object storage software and simplifies the build out of mass capacity storage solutions using fully tested, pre-configured architectures with a single-pane-of-glass management. A highly reliable and efficient enterprise-class solution, Lyve Rack's high-density drives enable customers to store more data at a lower cost.

Lyve CORTX, supported by Seagate, is installed and tested with Lyve Rack. This object storage software ingests and stores the large amount of structured and unstructured data being generated by enterprises, governments, and other businesses. CORTX is 100% open source, so customers benefit from transparency, open source licensing, and the flexibility to adapt the code. CORTX

provides an open, community-based framework for object storage. Lyve CORTX is the certified, supported, and indemnified version of CORTX community.

Lyve Rack can be bought via traditional or subscription licensing. CORTX utilizes Apache License, Version 2.0 for its core and AGPLv3 for peripheral components.
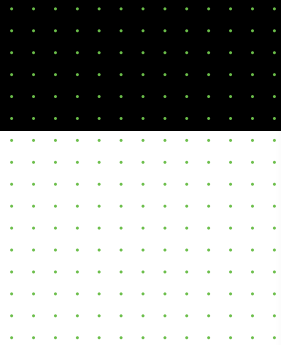
Features and Benefits

- Simpler and faster deployment with validated, pre-configured reference design and Amazon Simple Storage Service (S3) cloud native protocols

- 100% open source object storage software to optimize HDD capacity, performance, and efficiency with no vendor lock-in

- High availability: ADAPT erasure coding for faster rebuilds; redundant I/O paths

- Increased data durability: self-healing and bit-rot detection and repair

- Secure: internal audit logs for easy tracking, data provenance, and fingerprinting

- Bucket/object tagging for tighter control of AWS resources (in later release)

- Handling of mixed object sizes (in later release)
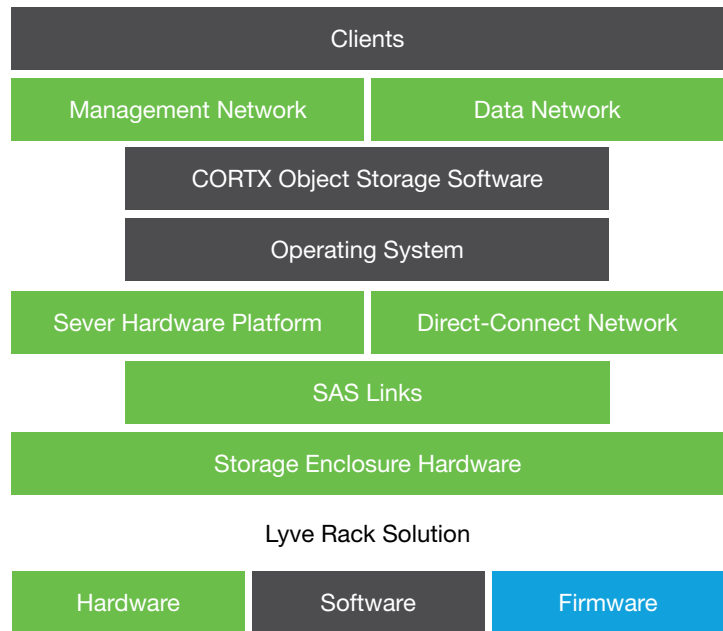
- Load balancing (in later release)
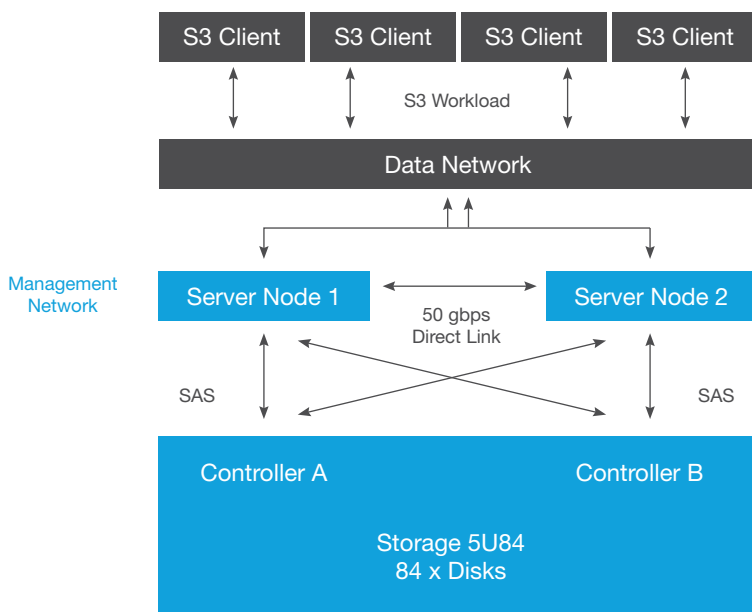
Reference Architecture

Lyve Rack converged infrastructure tests various components together to deliver an enterprise-class object storage solution. The base reference architecture is shown in Figure 1.

FIGURE 1

Clients

Management Network	Data Network

CORTX Object Storage Software

Operating System

Sever Hardware Platform	Direct-Connect Network

SAS Links

Storage Enclosure Hardware

Lyve Rack Solution

Hardware	Software	Firmware

FIGURE 2

S3 Client	S3 Client	S3 Client	S3 Client

S3 Workload

Data Network

Management Network

Server Node 1 ⟷ Server Node 2

50 gbps Direct Link

SAS | SAS

Controller A   Controller B

Storage 5U84
84 x Disks

Lyve Rack supports only S3 in its first release, so S3 clients are the initial users of the Lyve Rack. Two server nodes include operating system, I/O software, and management software. The diagram shows how CORTX benefits from a tighter integration with storage. CORTX enables the data to communicate directly with the storage drives without an intervening file system. This direct-to-drive architectural feature improves performance and reliability while granting better control over storage drives.

Figure 2 gives a little more detail on how the hardware components of Lyve Rack are connected.
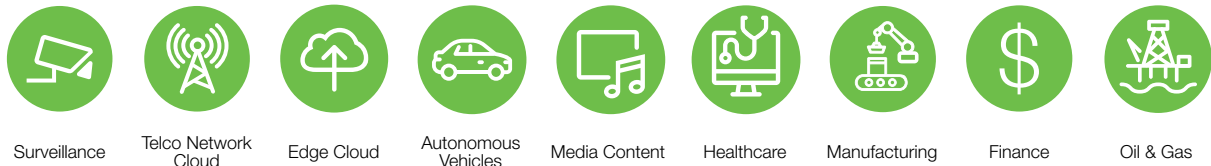
Based on the customer or industry use case, different user applications can be deployed on top of CORTX—all accessing data in various ways, but still working on the same base set of objects without data copies or physical segmentation of the persistent data store. CORTX provides a common metadata format to allow multiple protocols using CORTX to access the same data/objects. Developing on the CORTX platform is done via the standard Amazon S3 and file protocols. Special support for key-value storage is also available using the native libCORTX API.
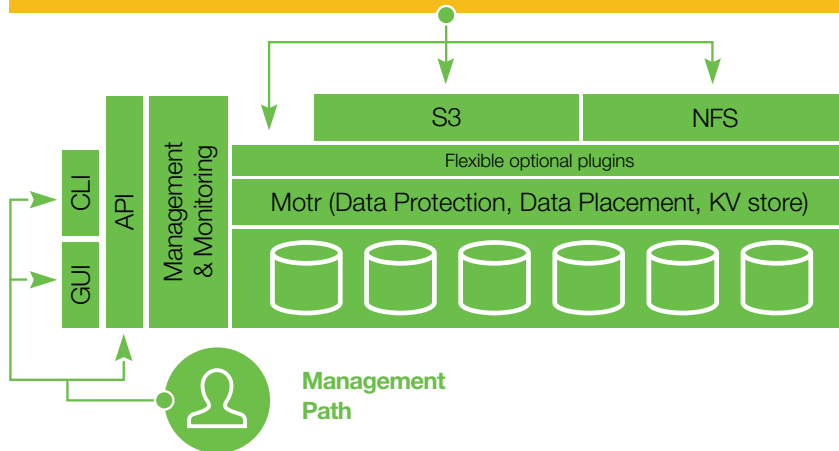

FIGURE 3 — **The role of CORTX in the IT 4.0 ecosystem**

Surveillance · Telco Network Cloud · Edge Cloud · Autonomous Vehicles · Media Content · Healthcare · Manufacturing · Finance · Oil & Gas

Applications: AI/ML, Big Data, Hybrid Cloud, IT 4.0, HPC

Frameworks: Spark, pyTorch, TensorFlow, Hadoop, MPI, Sidecar, Arrow, NoSQL, Splunk

S3 · NFS

CLI · API · Management & Monitoring · GUI

Flexible optional plugins

Motr (Data Protection, Data Placement, KV store)

Management Path

Bandwidth workloads, mass-capacity, metadata search

CORTX
Mass-Capacity Storage Software Platform

Reference Architecture Main Components:

- Storage: Seagate 5U84
- Software: Lyve CORTX object storage software
- Servers: SuperMicro, Dell
- Switches: Mellanox
- Operating System: RedHat Enterprise Linux
- Open Source Components: Elastic Search, Pacemaker, Consul, Cluster IP, and Rabbit MQ, among others

Conclusion

Lyve Rack delivers a simple and easy-to-deploy, high-capacity object storage solution that customers need to manage increasing data growth. The validated, pre-configured reference architecture and Lyve CORTX open source software lowers both the cost and risk of adopting Lyve Rack. Further, customers can choose different partner products for their infrastructure while also having a choice about the native cloud protocols they use to easily manage and move data across storage solutions.

Designed as an open and tightly integrated storage architecture, Lyve Rack is a building block solution architecture for customers in need of mass capacity storage. Strong community backing supports the development of new features and innovations to meet the changing needs of today's businesses.

Visit **www.seagate.com/products/storage/object-storage-solutions/lyve-rack/**

Download CORTX at GitHub **github.com/Seagate/cortx**

Contact Sales or Seagate Partner

seagate.com

SEAGATE