Hammerspace Erasure Coding

TECHNICAL BRIEF

Hammerspace is a Data Platform that unifies, automates, and provides fast access to unstructured data across any site, any cloud, any storage. The recent addition of high-performance erasure coding services to the product expands the range of supported storage options, providing customers with a new way to construct highly resilient storage environments with commodity off-the-shelf hardware.

Built using the Mojette Transform, Hammerspace erasure coding performs up to 2x faster than traditional erasure coding schemes, while providing data storage efficiency, data durability, and flexible scalability.

This document explains how erasure coding fits into the Hammerspace architecture, describes the technology, shows how it is implemented via Erasure-Coding Groups (EC-Groups), and includes examples of hardware configurations that may be used to implement EC-Groups.

Hammerspace Architecture and Storage Options

For a comprehensive review of Hammerspace architecture and functionality refer to the Hammerspace Technology White Paper.

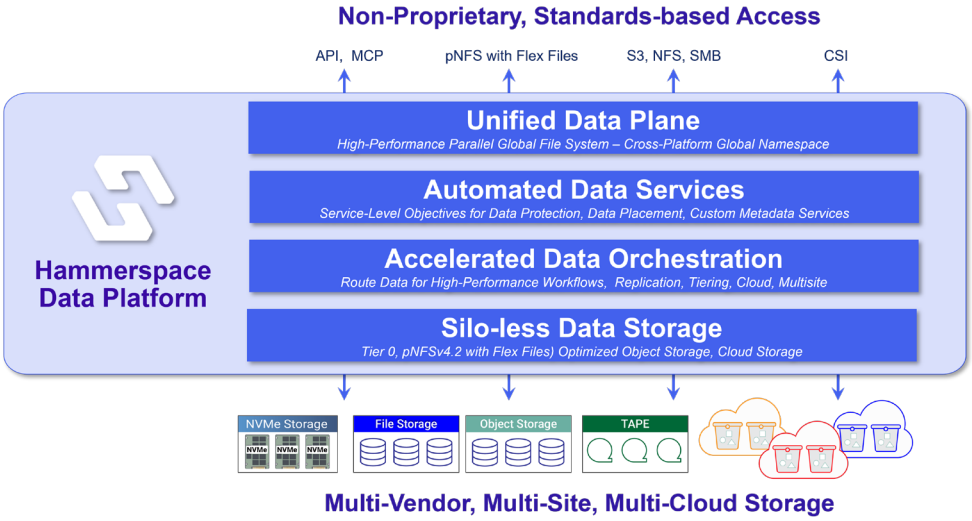


Figure 1 - Hammerspace Architecture

Erasure Coding Benefits

Efficient Data Protection

- Only 33% capacity overhead
- No single points of failure
- Failures don't affect efficiency

Self-Healing

- Automatic failure detection & recovery
- Protects against block, drive, & node failure
- CRC protection against data corruption

High Performance

- 2x the speed of traditional approaches
- Small block size for low latency
- Failed drives don't reduce performance

High Security

- Optional AES-256 data-at-rest encryption
- Directory-level granularity

A Hammerspace cluster consists of two types of nodes.Redundant Anvil nodes provide all metadata services and direct data orchestration activities. Data Services Extension (DSX) nodes host storage and execute data orchestration tasks. There are multiple options for data storage within a Hammerspace environment:

- Standards-based parallel high-performance file systems using pNFSv4.2
- Third-party file, block, object, cloud, and tape storage
- EC-Groups, erasure code-based storage clusters on commodity hardware

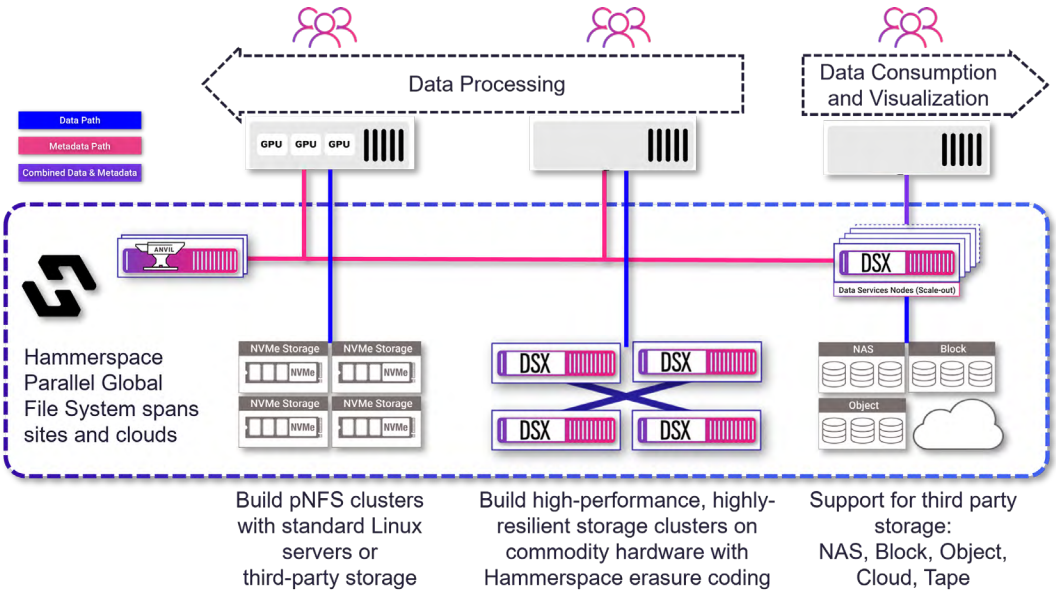


Figure 2 - Hammerspace Storage Options

The Mojette Transform

Most data storage systems that leverage erasure coding use Reed-Solomon codes, which date back to 1960. Reed-Solomon is not the only option, however. Many other types of error-correcting codes exist. Hammerspace erasure coding uses the Mojette Transform as an erasure code.

The Mojette Transform supports higher performance encoding and decoding compared to Reed-Solomon because it is computationally simpler. While Reed-Solomon requires relatively complex polynomial computations, Mojette uses only addition and subtraction, using fewer CPU cycles and resulting in two times the performance of Reed-Solomon, or more.

Because the Mojette transform is fast, Hammerspace erasure coding can use a smaller block size than other systems (4KB vs. tens of MB or more), providing increased efficiency for small files.

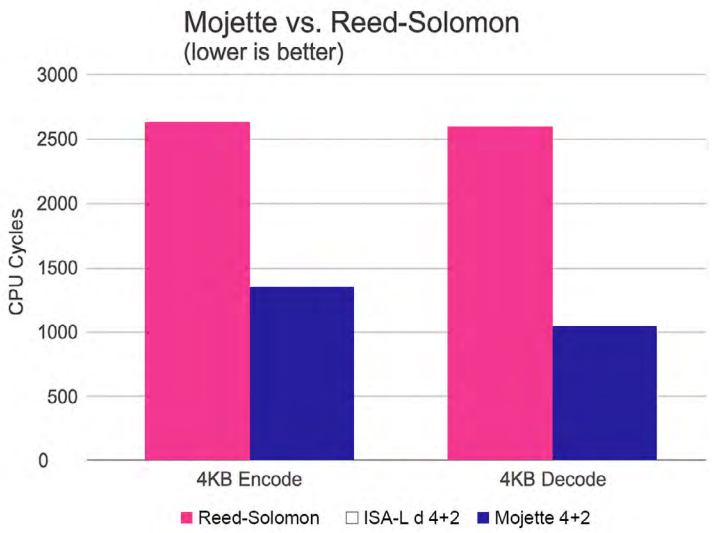


Figure 3 - Computational Effort Comparison

Data Redundancy

When using the Mojette transform for erasure coding, ‘n’ projections are created from each data block. Of these projections, only ‘m’ are required to recreate the original data, where m is a subset of n. The relationship between n and m is defined by the equation n/m = 1.5. Numerous combinations of n and m are possible, but generally n = 6 and m = 4. This allows two drives that hold projections for a specific block to fail without affecting data availability.

To ensure that writes are always protected even when some drives have failed, drives are grouped across nodes in sets larger than n. In the usual case with n = 6, sets of 8 drives are used.

In practice, it works like this:

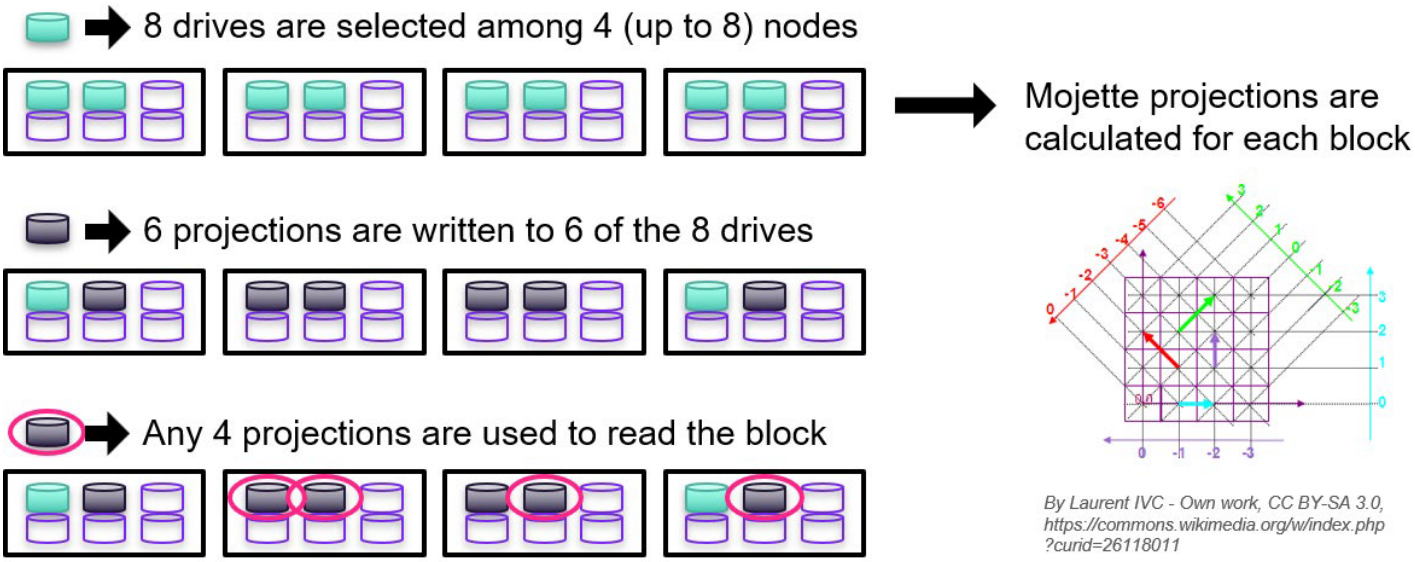


Figure 4 - Mojette Transform EC Implementation

In this example, one full node and one additional disk may be unavailable, and both write and read activity will continue, unaffected.

Data Security

EC-Groups is compatible with hardware-based encryption schemes, such as those provided by self-encrypting drives (SEDs), as well as array-based software encryption. Native encryption of data at rest is also an option. EC-Groups encryption uses the well known and trusted AES-256 algorithm to encrypt data blocks prior to erasure coding them. Encryption may be enabled on an individual directory basis if required.

EC-Groups Configurations and Scaling

Hammerspace has implemented this technology using what we call EC-Groups. An EC-Group refers to a group of servers joined together in a cluster, using Hammerspace erasure coding for durability and protection. The wide variety of server and storage hardware available and the flexibility of the software means that many different EC-Group configurations are possible. The Hammerspace team will assist in the design process to ensure systems will meet customer requirements.

Base Configurations

Figure 5 below shows two basic configurations, one based on high-capacity hard drives, and one based on high-performance NVMe flash. Most customer requirements can be met by some variation of one or both base systems.

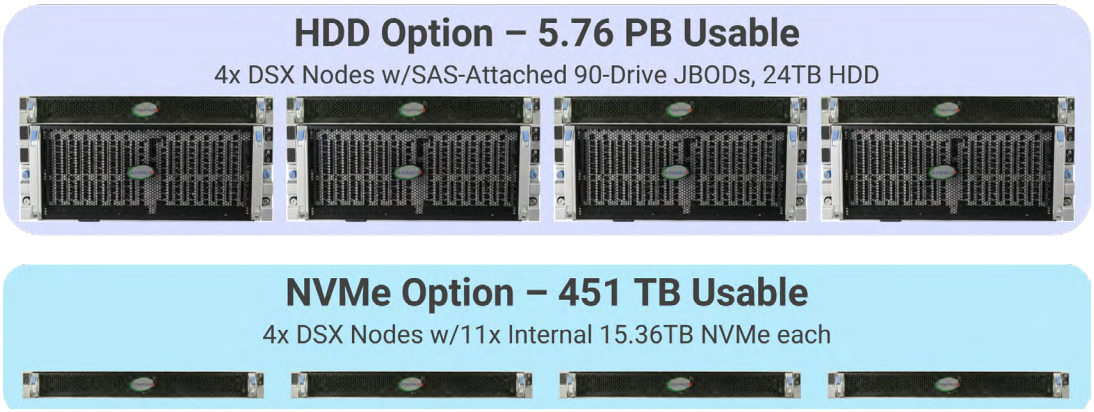


Figure 5 - Base Configurations

Scaling

EC-Groups configurations can flexibly scale capacity and performance, either by adding more drives to existing storage nodes, or adding more storage nodes, as few as one at a time. No downtime is required to scale.

Summary

Erasure coding provides a valuable new option for Hammerspace customers to build efficient, resilient, high-performance and secure storage systems on the hardware of their choice. The Mojette Transform confers unique performance benefits compared to Reed-Solomon-based approaches and enables flexible paths to scale as data grows and performance requirements increase.