**hEDVIg**

**Introduction**

In order to help the business create private, hybrid, or multi-cloud environments, Hedvig offers software-defined storage. Hedvig is perfect for legacy, current, and backup workloads since it is built for both main and secondary data. Block, file, and object storage are combined into a single, API-driven platform by the Hedvig Distributed Storage Platform, which keeps up with the need for data's constant expansion. Utilizing commodity servers or cloud computing, Hedvig's proprietary Universal Data Plane technology creates a distributed, scale-out cluster that serves as a flexible basis for bare metal, hypervisor, and container architecture.

Now Hedvig is known as Commvault Distributed Storage. It is a modern software solution that allows the user to tailor a flexible, data management system built on industry-standard x86 server hardware to support any application, hypervisor, container, or cloud. The solution is based on the former Hedvig product, which was acquired by Commvault in September 2019.

**Project Summary**

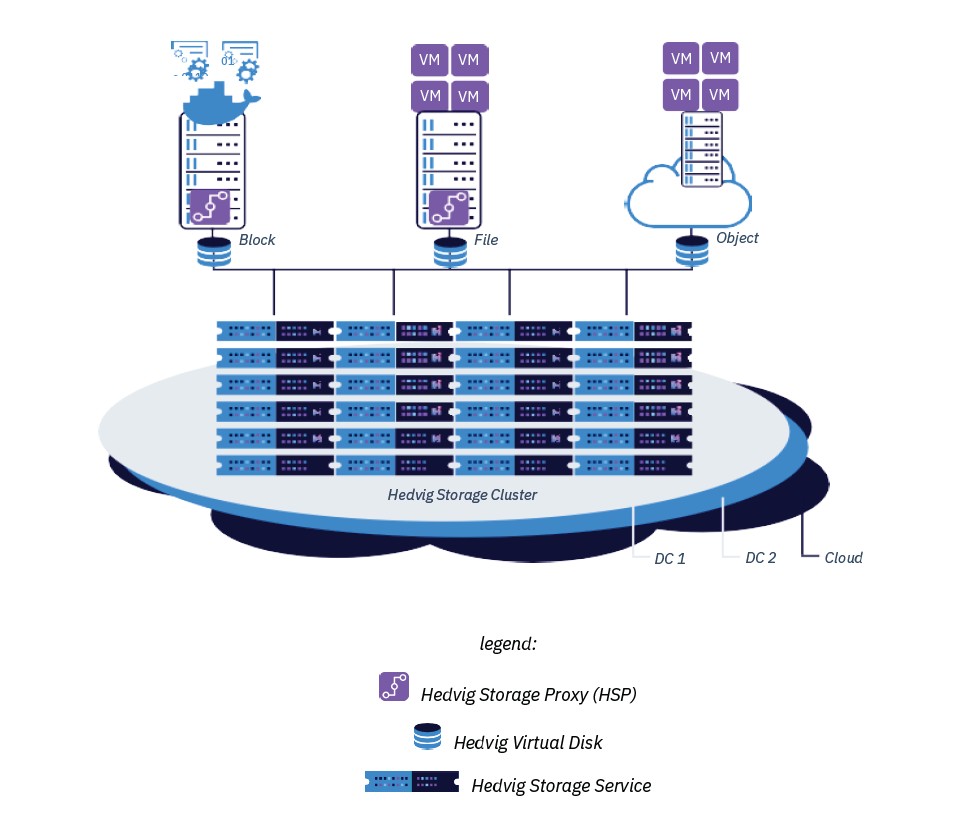
|  |  |
| --- | --- |
| Website | <https://www.commvault.com/software-defined-storage> |
| Organization/Foundation Name | Hedvig (A Commvault Venture) |
| License | Software Licenses (Requires Subscription) |
| Open/Proprietary | Private |
| Brief Description | Hedvig’s software-defined storage solution provides the flexibility needed for rapidly changing data, application, and user requirements in private, hybrid, and multi-cloud environments. |

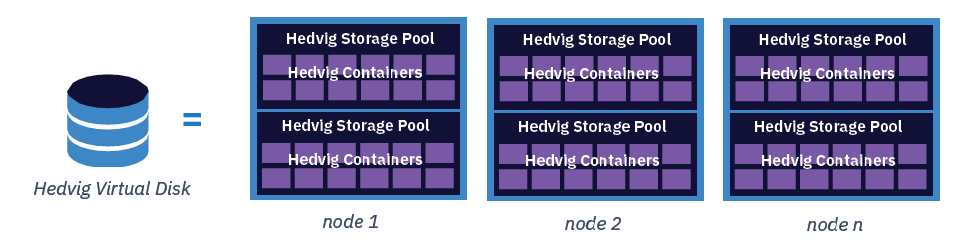
**Project Details**

**Key Features**

* **Extensive hardware independence:** Software-defined storage (SDS) allows organizations to leverage existing hardware without having to purchase new hardware. Hedvig's design not only supports commodity X86 hardware but also provides the capability to support lower-cost ARM-based infrastructure for an added level of flexibility and affordability.
* **The ability to go from hyper-converged to hyper-scale:** Traditional storage arrays are architected with a specific deployment sweet spot in mind. As an end-user company expands, capacity or performance demands may force migration from one system to another. Hedvig's solution can be deployed across just a couple of systems or it can scale to potentially thousands of storage nodes. The technology can be also deployed as a virtual instance for a hyper-converged solution.
* **On- and off-premises cloud flexibility:** Hedvig's storage technology can be deployed for either an on-premises or public cloud environment, providing a hybrid cloud storage solution. A single solution that can reside either on- or off-premises can reduce users' learning curves, thereby easing the transition to public cloud resources.
* **Uncompromising storage functionality:** Hedvig's first release includes a solid line-up of both "check box" and advanced storage capabilities, as well as a range of plug-and-play storage solutions. High availability via replication. Data protection with snapshots and clones. Unified storage protocol support with iSCSI, NFS, S3, and Swift. Auto-tiering and balancing across scale-out storage nodes. The ability to scale performance and capacity independently.
* **Storage for virtualization, cloud, and big data:** The company's flexible design and multi-protocol support will likely benefit from big data and big data deployments. Hedvig's flexibility to be deployed as an instance or virtual instance can significantly improve its affordability over traditional storage solutions.

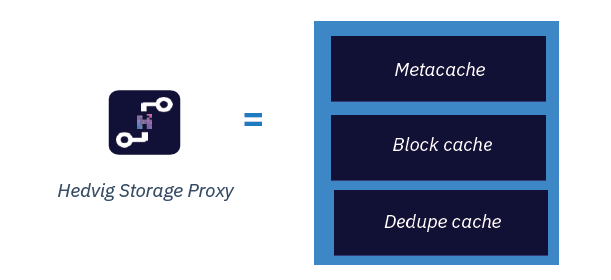
**Architecture**



1. **Hedvig Storage Service:** The foundation of the platform is the proprietary distributed systems engine created by Hedvig. It installs on standard x86 or ARM servers and converts current server and storage resources, such as SSD/flash media and hard discs, into a fully functional elastic storage cluster. To build a single storage cluster that is implicitly hybrid, the software deploys to hosted or public clouds as well as on-premise infrastructure.
2. **Hedvig Virtual Disk:** The Hedvig Distributed Storage Platform's core abstraction is the Hedvig Virtual Disk. Organizations can rapidly create any number of thinly supplied virtual drives. When deploying a virtual disc, many user-configurable properties can be changed.

**Virtual Disks, Containers, and Storage Pools** play a critical role in high availability by creating protection groups for data protection needs. A Hedvig Virtual Disk is partitioned into fixed-size virtual chunks, each of which is called a Hedvig Container. Since replica assignment occurs at the container level - not simply at a virtual disk level - the data for a virtual disk is spread across the Hedvig Cluster, thus eliminating hotspots and allowing increased parallelism during client I/Os or disk rebuilds.

Replicas are chosen according to replication factor and replication policy settings to support the application`s data protection needs. Within a replica, containers are assigned to a Hedvig Storage Pool. Hedvig Storage Pools are logical groupings of disks/drives in the Hedvig Storage Cluster Nodes and are configured as the protection group for disk/drive failures and rebuilds.

1. **Hedvig Storage Proxy:** The Hedvig Storage Proxy is a small-footprint software component that may be deployed at the application layer as a virtual machine or a Docker container. By intelligently routing I/O requests to the appropriate backend storage nodes, the storage proxy gives access to the hyper-scale storage nodes. It converts between several protocols and offers edge caching using local flash devices to boost storage performance right on the application hosts.

Three caching mechanisms – **meta cache, block cache, and dedupe cache** – are featured components of each Hedvig Storage Proxy. These caches enable distributed data operations, performing storage system operations locally at the application compute host to accelerate read performance and efficiency.

* **Metacache:** It is a mandatory cache of the metadata that is stored locally on the Hedvig Storage Proxy, preferably on SSDs. This cache prevents the need to traverse the network for metadata lookups, leading to significant read acceleration.
* **Block Cache:** For virtual disks enabled with the Client-side Caching option during provisioning, the block cache stores a working set of disk blocks to local SSD/PCIe drives to accelerate reads. By returning blocks directly from local flash media, read operations avoid network hops when accessing recently used data.
* **Dedupe Cache:** For virtual disks enabled with the Enable Deduplication option during provisioning, the dedupe cache resides on local SSD media and stores fingerprint information for virtual disks that use the deduplication policy. This cache allows the storage proxy to determine whether blocks have been previously written and if so, to bypass the need to write the data over the network to the storage cluster.

Usage

A software-defined storage (SDS) system called Hedvig may be used to a number of different data storage use cases. Hedvig can be used in a variety of situations, including:

* Organizations may utilise Hedvig to offer a scalable and highly available file storage solution. It supports well-known protocols like NFS and SMB, which makes it simple to integrate with currently used file-based applications.
* Hedvig's block storage capabilities may be utilised to offer a highly-available and scalable storage solution for programmes like databases and virtualization that demand quick access to data.
* For unstructured data, such as pictures, movies, and backups, Hedvig can offer a highly available and scalable object storage solution.
* By duplicating data over different sites, Hedvig may be used to offer disaster recovery solutions for businesses, guaranteeing that data is always accessible in the event of a catastrophe.
* Hedvig serves as a link between on-premises storage and cloud storage and may be used to move data to the cloud.
* Hedvig may be used as a platform for software-defined storage (SDS), offering a solitary, integrated storage option that is simple to expand and maintain.
* Hedvig is also applicable to a wide range of other use cases, including big data, artificial intelligence, and the internet of things.

Overall, Hedvig is a flexible and highly available solution that can be used to store and manage data across a wide range of use cases, providing scalability and high availability. Customers like BNP Paribas, Scania, GE, and Pittsburg State rely on Hedvig as a fundamental enabler of digital business.

**Technical Details**

An elastic storage solution that scales from a few terabytes to tens of petabytes is the Hedvig Distributed Storage Platform. The addition of backend storage nodes allows the platform to grow performance and capacity. A cluster may scale out smoothly and automatically. By adding Hedvig Storage Proxies, the compute tier may also be scaled out (independent of the backend). Reliability will scale linearly if SSD assets are made available to the storage proxies.

For constant performance, it is generally advised that you retain storage nodes in homogenous groupings. Hedvig, however, permits heterogeneous nodes across the cluster with various hardware requirements. As the system expands, you may add groups of nodes to benefit from faster CPUs and greater capacity/performance discs as they become available. It should be noted that the Hedvig Distributed Storage Platform can handle sub-millisecond latency and expand to tens of thousands of IOPS and hundreds of MB/sec per virtual disc, while system performance is dependent on the workload and the exact system configuration.

Hedvig's Encrypt360 function offers software-based encryption. This makes it possible to encrypt data at the moment of intake (on the storage proxy server). When data is encrypted in this manner, it is safe while it is in use at the storage proxy, traveling between storage nodes (or sites) as part of replication, traveling between storage nodes (or sites), and at rest. KMIP-compliant KMS and 256-bit AES encryption are used to assure compliance with regulations like PCI, HIPAA, and Gramm-Leach-Bliley. To allay key management worries, any external key management solution may be plugged in.

**Other Information**

The first software-defined storage solution with real distributed system DNA is the Hedvig Distributed Storage Platform. This special platform can connect to any OS, hypervisor, container, or cloud and may be deployed in hyper-scale or hyper-converged modes. The Hedvig Distributed Storage Platform offers a comprehensive set of corporate storage capabilities that can be customized at the virtual disc (application) level, scaling easily and linearly from a few nodes to thousands of nodes. The implicit hybrid architecture safeguards each application with a unique disaster recovery plan across several data centers or clouds.

The Hedvig Distributed Storage Platform, which is elastic, straightforward, and versatile, is the go-to option for businesses looking to update data centers and create private and hybrid clouds. The Hedvig Distributed Storage Platform is ideal for scenarios where the cost-per-terabyte to store data and the operational burden of managing silos of dissimilar storage systems are having a negative impact on the bottom line. The Hedvig Distributed Storage Platform enables IT to lower expenses and boost business responsiveness, to put it simply.

**Reference**

[commvault.com](https://www.commvault.com/resources/commvault-distributed-storage-platform-technical-and-architectural-overview-whitepaper) : Official website of Commvault

[linkedin.com](https://www.linkedin.com/company/hedvig-inc/) : LinkedIn account of Hedvig Inc.