



**SWISS  
VAULT**

# **VaultFS Overview**

**(A big-data storage solution)**

---

**Introduction to VaultFS and its Advanced Features**

---

## Table of Contents:

1. Who are we?
  2. Current Problems faced by Data managing/collecting organizations
  3. Case Study : Solving Data Challenges with VaultFS
  4. Get in touch with us.
- 

## Who are we?

At Swiss Vault, we understand the critical nature of data management and archiving for business operations. Our software solution, VaultFS, optimizes data transfer speeds, minimizes the risk of data loss, maximizes robust storage with minimal overhead and supports scalable data needs. This ensures faster transfer rate with reduced wait times, safety from data loss and improved system efficiency.

With VaultFS, we employ advanced erasure coding and support remote administration of data and parity chunks, protecting against bit-rot and disk failure. VaultFS supports parallel read and write operations, reducing hardware overhead and minimizing the total cost of ownership (TCO).

Engineered for high data availability over decades, VaultFS uses erasure coding and a peer-to-peer data management strategy to deliver robust and unparalleled value to efficiently meet user's business requirements.

---

## Current Problems faced by Data-collecting/managing organizations

There are a broad range of physical and technological challenges that put the integrity of long term data storage at risk and result in partial or total loss of data. These problems could cause problems resulting in overall disruption of business operations and hence require attention.

- **Bit-Rot:** Gradual corruption of data due to environmental factors and physical wear.
  - **Disk Failure:** Sudden hardware issues leading to data loss.
  - **Limited Scalability:** Difficulty in expanding or contracting storage capacity.
  - **Wasted Disk Space:** Inefficiency in storage due to redundancy mechanisms like RAID.
  - **High Total Cost of Ownership:** Significant costs due to additional storage and expensive hardware.
-

## CASE STUDY: Solving Data Challenges with VaultFS

**Scenario:** A genomics data collection and research organization, let's say GenXYZ, is dealing with extensive and critical datasets, faces frequent data corruption, high storage costs, and scalability issues. They need a reliable, readily available, cost-effective, and scalable data management solution.

### Challenges:

1. **Bit-Rot and Disk Failure:** GenXYZ experiences gradual data corruption (bit-rot) and sudden disk failures, leading to data loss and operational disruptions. The RAID architecture usually used by data-centers offers limited data insurance, but no flexibility to manage (ZFS offers).
2. **Limited Scalability:** Their existing storage system cannot efficiently expand or contract, hindering their ability to manage growing data needs. Apart from that, they have a set configuration with RAID structure, and number of disks assigned to parity structure for them, making it hard to manage disk scalability and configurations remotely.
3. **Wasted Disk Space:** Redundancy mechanisms like RAID increase storage requirements, leading to inefficient use of disk space. Due to this, with growing data, everytime their storage needs increase by X percent, their total disk space needs to be 2X, or YX, where Y represents the number of redundancy required for data.
4. **High Total Cost of Ownership:** The need for additional storage and expensive hardware escalates costs, impacting their budget. Their overall cost gets increased along with the need to pay extra attention to ensure changing the disk, everytime there is a failure, given the low threshold for managing disk failure.

### Solutions Implemented by VaultFS

#### 1. Erasure Coding

- VaultFS uses advanced erasure coding to protect data and recover lost or corrupted data chunks. For example, a configuration of D=20 and P=5 means that 20 data chunks are protected by 5 parity chunks, allowing recovery even if five chunks are lost.
- **Benefit:** This method reduces storage overhead compared to replication, enhancing reliability and cost-efficiency.

#### 2. Peer-to-Peer Architecture:

- VaultFS employs a peer-to-peer data management strategy, eliminating single points of failure and ensuring continuous data availability.

- **Benefit:** This architecture supports incremental scalability, allowing DataXYZ to add nodes without overhauling the existing infrastructure.
- 3. **Dynamic Configuration:**
  - Administrators at GenXYZ can adjust the number of data and parity chunks dynamically, balancing storage efficiency and data resiliency as per their needs.
  - **Benefit:** This flexibility optimizes storage resources and reduces hardware costs.
- 4. **Parallel Read and Write Operations:**
  - VaultFS supports parallel I/O operations, improving data access speeds and overall system performance.
  - **Benefit:** Enhanced performance reduces reliance on high-end storage devices, further lowering costs.
- 5. **Compatibility and Ease of Integration:**
  - VaultFS is POSIX-compliant, allowing seamless integration with GenXYZ's existing applications and workflows.
  - **Benefit:** This ensures a smooth transition with minimal disruption to ongoing operations.
- 6. **Monitoring and Management:**
  - With tools like VMAN and Grafana, GenXYZ can monitor system performance, manage resources proactively, and adjust system parameters without downtime.
  - **Benefit:** These capabilities ensure optimal performance and efficient hardware utilization.

## Implementation Steps:

1. **Initial Setup:**
  - Install VaultFS using provided scripts and guides.
  - Configure system parameters to match GenXYZ's storage and resiliency requirements.
2. **Data Migration:**
  - Use standard POSIX commands (e.g., cp, rsync) to transfer data from the existing storage solution to VaultFS.
  - Monitor the migration process with VMON and Grafana to ensure data integrity and performance.
3. **Validation and Testing:**
  - Conduct performance tests to validate system capabilities.
  - Test data recovery processes to verify erasure coding and redundancy configurations.

#### 4. Full Deployment:

- Transition to using VaultFS as the primary storage solution.
- Continue monitoring and optimizing the system as needed.

**Outcome:** By implementing VaultFS, GenXYZ significantly reduces storage overhead, enhances data reliability, and lowers total cost of ownership. The advanced features of VaultFS, including erasure coding, peer-to-peer architecture, and dynamic configuration, address their data management challenges effectively, ensuring high performance and scalability.

### Software Services Overview

To effectively utilize VaultFS, below mentioned services are added as part of the package:

1. **VaultFS:** The core filesystem service ensuring data availability and resiliency.
2. **VMON:** The monitoring service for overseeing the performance and health of the VaultFS deployment. Collects the data related to files, metadata information, transfer rates etc., and provides it to grafana for processing and visualizations.
3. **VMAN:** The management service for configuring and managing the VaultFS Disk & Parity Settings. Helps the users scale the system requirements, along with control of Monitoring data.
4. **Grafana:** A dashboard service focused on offering visual and tabular information about performance of disks.

For a brief introduction, please watch the [introduction video](#).

For further information regarding Setup, Management and Monitoring included with Feature Testing and Benchmarking of VaultFS, please refer to [VaultFS.io](#).

**The above discussed case-study is applicable to multiple industries.**

**Healthcare & Scientific Research:** The exponential growth in healthcare, genomics and research data demands efficient access to data, where credibility hinges on the integrity of the data.

**Government & Enterprise :** The consistent collection of public and client records, and the need of quick and secure access to them highlights the importance of long-term data storage options.

---

*Robust structure of VaultFS ensures **data safety** by **reducing scenarios where data loss** could occur and offer **smooth operation for reliable data transfer and access**, placing itself as a **long-term solution** for all.*

---

### Get in touch with us:

VaultFS values its customers and has dedicated resources to ensure the best quality experience when demoing the product. Businesses can discover how VaultFS can serve their needs. Connect at VaultFS.io or reach out to the support team at [Swiss Vault Global Support](#)

*For more information, visit us on the web at*

<https://www.swissvault.global/>

