DOES BACKUP NEED A FILE SYSTEM OF ITS OWN?

A Deep Dive Into VembuHIVE™ File System

Jayashree Subramanian
Head of Vembu Labs

Copyright Information

Information in this document is subject to change without any notice. The entire risk of the use or the results of the use of this document remains with the user. No part of this document may be reproduced or transmitted in any form or by any means (electronic, mechanical, photocopying, recording, or otherwise), or for any purpose, without the express written permission of Vembu Technologies. A formal Non Disclosure Agreement may additionally govern usage of the information in this document.

Vembu Technologies may have patents, patent applications, trademarks, copyrights, or other intellectual property rights covering the subject matter in this document. Except as expressly provided in any written license agreement from Vembu Technologies, the furnishing of this document does not give you any license to these patents, trademarks, copyrights, or other intellectual property.

Executive Summary

SMART FILE SYSTEM

Two interesting trends in the backup industry have created the need for a smart file system. First, there is a rising demand for a more sophisticated use of backup data, than just traditional recovery. Think how interesting it would be to share and collaborate the files residing inside in your disk image backup without having to mount or boot the image. This is possible only if the file system is able to read the image file, bit-by-bit and understand what files are stored inside the image file.

Second, the demand for online backup has led the service providers to host their backup on cloud infrastructure. This means that backup applications could potentially harness the cluster file system and computing possibilities of cloud. For example, it is possible to dramatically improve the read/write speed of backup data by storing it in SAN/NAS and distributing the operations to a large cluster of servers.

This opens up new avenues for intelligent use cases for backup data such as big data analytics. The traditional file systems (NTFS, EXT, FAT, etc.) and the modern cloud file systems were not designed for backup applications. The file formats (VHD, VMDK) used by backup products do not exploit the power of cloud. A single file system cannot be a panacea for all applications which is why Vembu Technologies developed its own cloud file system called VembuHIVE™.

WHAT IS A FILE SYSTEM?

In computing, a file system (or file systems) is used to control how data is stored and retrieved. Without a file system, information placed in a storage area would be one large body of data with no way to tell where one piece of information stops and the next begins.

-Wikipedia

NEED FOR APPLICATION SPECIFIC FILE SYSTEM

Backup is just not about storage. It's the intelligence on top of storage.

Typically when businesses think of backup, they see it as a simple data copy from one location to another. Traditional file systems would suffice if the need were to just copy the data. But backup is the intelligence applied on top of storage where data can be put to actual use.

Multiple use cases for backup data

Backup data is no longer a cold storage. Businesses are planning to use the data for various purposes. Imagine the ability to use backup data for staging, testing, development and pre-production deployment. Traditional file systems are not designed to meet such complex requirements.

Businesses need the restore available NOW!

With the advent of information technology, more and more organizations are relying on IT for running their businesses. They cannot afford to have downtime on their critical applications and need instant access to data in the event of disaster. Hence, a new type of file system is necessary to satisfy this need.

What is VembuHIVE™?

EFFICIENT CLOUD FILE SYSTEM

VembuHIVE™ is an efficient cloud file system designed for large-scale backup and disaster recovery (BDR) application with support for advanced use cases. VembuHIVE™ can be thought of as a *File System of File Systems* with in-built version control, encryption, and in-built error correction. During the backup, the data present in the backup files or an image is separated from all the bookkeeping associated with it, i.e., its metadata and stored as objects.

VembuHIVE™ manages the metadata smartly through its patent-pending technology, in a way that is agnostic to the file system of the backup, which is why we call VembuHIVE™, a file system of file systems. This helps the backup application to instantly associate the data in VembuHIVE™ to any file system metadata, thereby allowing on-demand file or image restores in many possible file formats. The data and metadata storage, harness cluster file system, computing and storage.

This is really a powerful concept that will address some very interesting use cases not just in the backup and recovery domain but also in other domains, such as big data analytics.

The key to the design of VembuHIVE™ is its novel mechanism to capture and generate appropriate metadata and store it intelligently in a cloud infrastructure. The increment data (the changes with respect to a previous version of the same backup) are treated like versions in a version control system (CVS, GIT). This revolutionary way of data capture and metadata generation provides seamless support to a wide range of complex restore use cases.

BENEFITS OF VEMBUHIVE™ FILE SYSTEM

- ✅ **Built-In Version Controls**
- ✅ **Point-In-Time Restores**
- ✅ **Storage Reduction**
- ✅ **WAN Acceleration**
- ✅ **Built-In Error Correction**
- ✅ **Ultra-Reliable Data Integrity**
- ✅ **Granular Recovery**
- ✅ **Highly Scalable**
- ✅ **Instant VM Restores**
- ✅ **Block Level Storage**
- ✅ **Faster Data Processing**
- ✅ **Universal Data Format**
- ✅ **Bootable Incrementals**
- ✅ **Complex Restore Use Cases**

Use Cases of VembuHIVE™

Built-In Version Control and Point-In-Time Restores

During an incremental backup, VembuHIVE™ stores only the changed blocks since the latest backup, similar to versions in a version control system. Due to this and the smart metadata management that is flexible enough to expose the underlying data in multiple ways, VembuHIVE™ exposes every incremental as a virtual full backup. i.e. a restoration of a backup with any time stamp, will not require merging of all the changes to a previous full backup. A point-in-time full is available for every timestamp during which an incremental backup was done. These backup versions can be instantly booted or mounted without any tedious merges.

Deduplication for Storage Reduction

Besides storage capacity, the more data there is to manage, the greater is the impact and costs associated with provisioned servers, network bandwidth and even human resources to manage the infrastructure. In the face of high volume data growth, backup & restore products still need to meet recovery time and recovery point objectives (RTOs and RPOs). Vembu's innovative, global, variable-length, block level, client & server-based deduplication technology provides for dramatic storage cost and bandwidth savings.

Built-In Error Correction Techniques for Reliability

A parity file (additional redundancy) is added to each data chunk in the VembuHIVE™ using advanced error correction techniques. In the event of a data corruption, the information in the parity file is used in fixing errors in the VembuHIVE™ file storage. VembuHIVE™ also maintains such parity information at the backup file or disk image-level, chunk- level, repository-level, and client or backup-level. These capabilities are not provided in the existing file system.

Mail/Document/File Level Restores

VembuHIVE™ is intelligent enough to understand the way content is organized inside the backup data and interpret in multiple ways, irrespective of where it came from (BDR, file backup, or virtual machine backup) and thus can perform on-demand granular restores.

ABOUT VEMBU

Vembu ventured into the Backup & Disaster Recovery market in 2002 and has been striving to provide affordable software with enterprise-level features for all to achieve complete data protection.

Our flagship offering - Vembu BDR Suite, is a comprehensive Backup & Disaster recovery solution that protects data across diverse IT environments which includes virtual (VMware, Hyper-V), physical (Windows, Linux, Mac), and cloud (Office 365 & G Suite). It is ideal for all businesses regardless of type, size or budget.

CONTACT US

US & Canada:+1-512-256-8699

UK: +44-203-793-8668

Sales: vembu-sales@vembu.com | Support: vembu-support@vembu.com

DOWNLOAD FOR FREE TRIAL