



FLARE: The High-Performance Indexing Solution for Virtual Disk Environments

Overview of Technologies and Concepts Developed
for the File-Level Restore FLARE Project

Executive Summary

Comtrade 360 has developed a file system indexing software (FLARE) with greatly improved performance and drastically decreased catalog disk space requirements.

FLARE can index virtual disks (VMware, Hyper-V and RedHat Virtualization), raw disks, partitions and mounted filesystems.

Introduction

Modern storage is increasingly optimized for virtualized workloads; VM snapshots are done in hardware and are extremely space and CPU efficient. In these environments, backups are done frequently, by snapshotting whole VMs. However, restoring whole VM snapshots is rarely desired, thus we investigated and developed FLARE to enable file-level restore from snapshot-based VM backups.

Problem Statement

We needed to provide efficient indexing, display and search of files on VM snapshots. Each snapshot results in one or more virtual disks. For scaling, we assumed 1000 VMs, each with 100 snapshots and 1 new snapshot/sec. Snapshots are created frequently and rotated to keep their number under control. Thus, deleting a snapshot is equally performance critical.

These requirements presented two problems:

Traverse File Systems On Virtual Disks

The current industry standard for virtual disks traversal is Libguestfs, which is extremely CPU and time intensive, and requires nested virtualization.

Traditional recursive traversal is extremely slow for NTFS, and often unacceptably slow for other filesystems.

Insert To Catalog

Centralized cataloging is widely regarded as impractical by traditional backup solutions.

It requires frequent table locking, may need reference counting, etc.

Although multiple SQL databases have extensions for hierarchical data, we are dealing with **versioned hierarchical data**, akin to git.

While centralized storage saves space by saving unique file names only once, file names account for less than 30% of all file attributes – file size, file times, etc., account for the rest.

Solution

Our solution consists of a C++17 indexing daemon and a Java main daemon, which serves as an API endpoint. The C++ indexing daemon, flare-catalog, is a systemd service that can be installed on Debian-like systems. The flare-catalog can run standalone or as a daemon. Communication between the flare-catalog and Java daemon occurs via inbox/outbox directories.

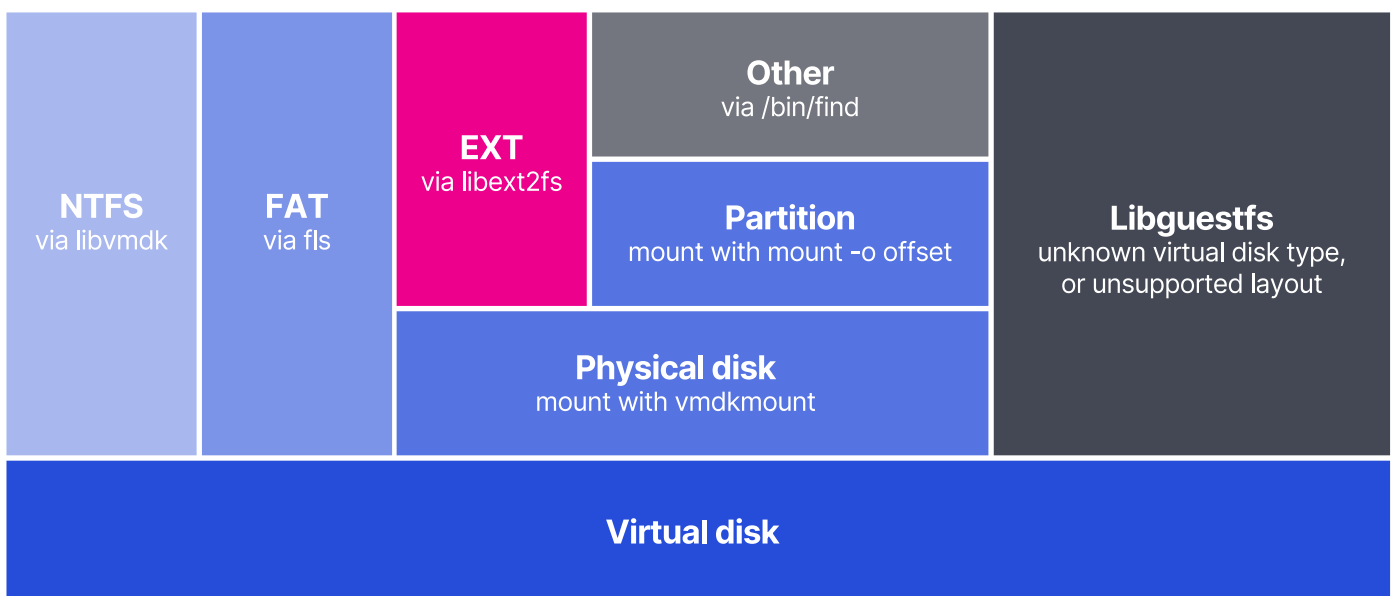
Traverse File Systems On Virtual Disks

We combined several approaches:

- Read filesystem metadata directly, without traditional recursive traversal.
- We have implemented this approach for NTFS and EXT2/3/4 filesystems.
- There is extensive research in this field. We used "File System Digital Forensics" as a basis.
- Use Sleuth Kit open source kit for digital forensics.
- Traditional mount + find.
- Libguestfs.

User can configure which approach to use for which filesystem via configuration file.

The defaults are:



Compared to the Libguestfs, we are 50x faster while using a fraction of the CPU:

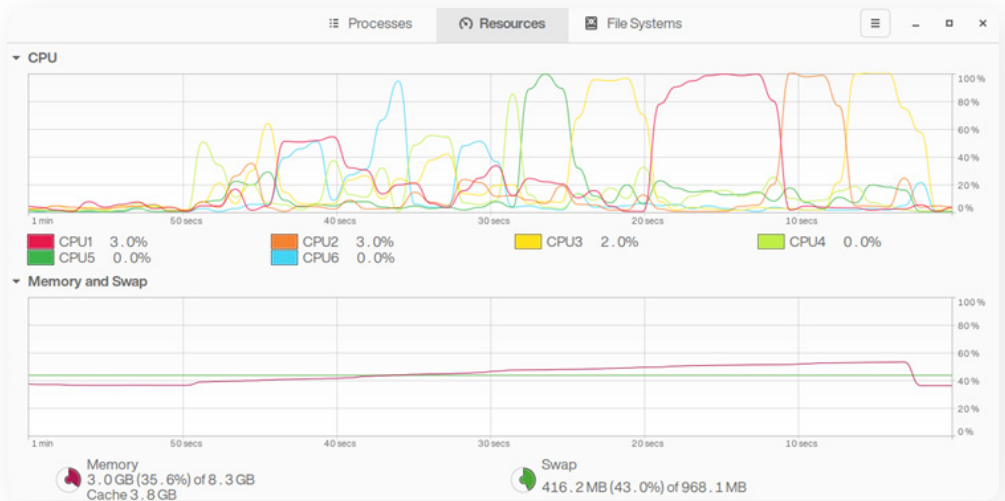


Figure 1 Resource use during VMDK traversal with Libguestfs

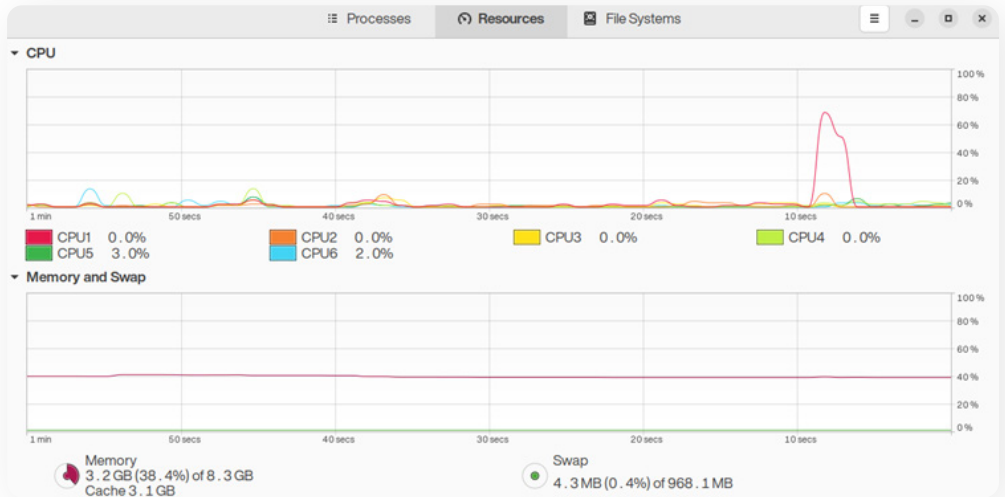


Figure 2 Resource use during same VMDK traversal with FLARE

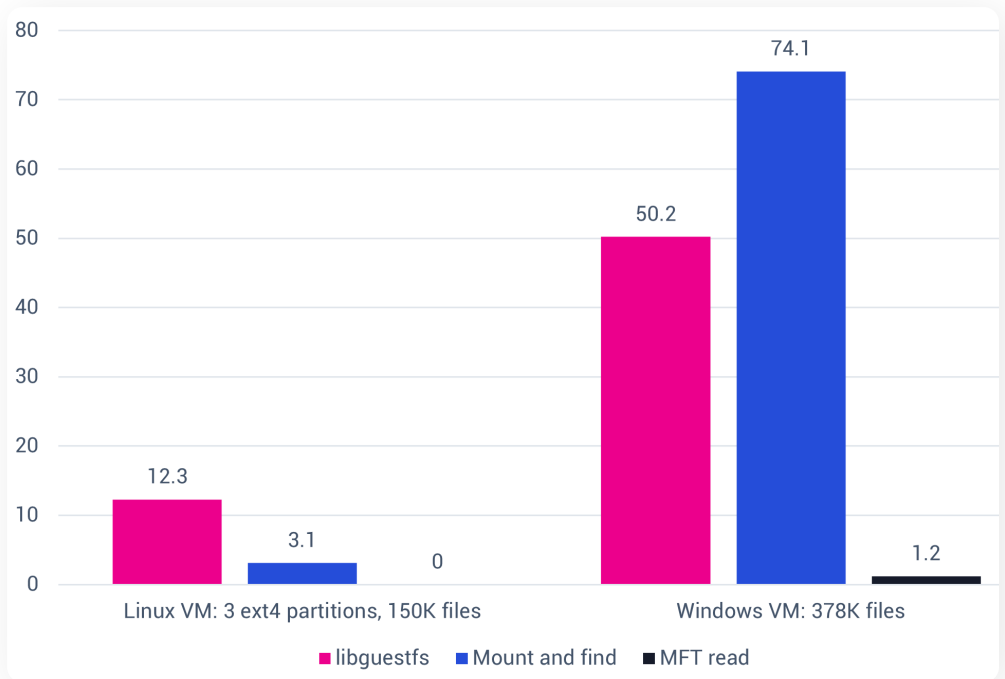


Figure 3 Comparison of filesystem traversal times (in seconds)

Insert To Catalog

Each virtual disk is indexed to its own SQLite database. The database uses approximately 100 bytes per file/directory. We achieved 1 million records/second on a commodity VM. The initial index of a virtual disk contains all file names and attributes. Subsequent indexes of the same virtual disk contain only changes since the full index – we call this a **differential catalog**. Once differential catalogs grow large enough, we repeat the full catalog. With this schema, we require ~80% less disk space compared to creating a full catalog every time.

Summary

Our solution contains several technologies that can be used, directly or with minor changes, to areas not specific to VMs, snapshots or virtual disks:

- Fast traversal of file systems on or out of virtual disks for 3rd party backup solutions
- Fast Windows search for NTFS filesystem, akin to WizTree
- Space efficient storage of versioned hierarchical data, usable for backup solutions and auditing software

In summary, FLARE offers a versatile and efficient solution for indexing and restoring files from snapshot-based VM backups. Its innovative approaches to file system traversal and catalog management provide faster performance, reduced CPU usage, and considerable disk space savings. Additionally, FLARE's technologies can be adapted for use in other areas, such as third-party backup solutions, fast Windows search for NTFS filesystems, and space-efficient storage of versioned hierarchical data for backup and auditing applications.



info@comtrade360.com
+38681605200
comtrade360.com