



Fragless FS: File-system for IaaS

10.07.2021

Paweł Suchanecki

Linbedded Sp. z o.o.

Kaliska 19/21/30

02-316 Warszawa

Abstract

Recently¹ introduced HDD (Hard-Disk Drive) write technology called SMR (Shingle Magnetic Recording) in essence allows condensing physical tracks² thus offering more storage on a disc as compared to traditional write technologies (CMR, PMR).

Among that physical change, this technology introduces a concept of a *zone* (a set of tracks) that is new logical sub-container from HDD's perspective so each zone can be perceived as a separated storage silo.

Linbedded is creating a file-system (extension) for GNU/Linux OS called FraglessFS that supports cloud-native storage for IaaS providers by design. This file system is ideal for virtual machine ecosystems as it offers a continuously written, always defragmented backing store that is optimized for maximum performance of virtual I/O needs.

Goals

1. One Zone per Virtual Machine Instance
2. Never delete files (add. storage cost) on the physical drive, leave SMR writes intact.

Specification


- Offer functionality as additional features...AS
- extension to industry accepted filesystems (e.g. Ext4, Btrfs)... WITH
- ZBC level (Zoned Block Commands) abstracted

Example

A GNU/Linux system administrator wants to create new storage volumes to be used by his company. At the level of filesystem creation (i.e. `mkfs.ext4`) he/she specifies parameters which create partitions as dedicated storage for Virtual Machine(s). In result a set of zones (with known *ZoneIDs*) is created and each can be mounted as a separate device in the system. Each device can then be mounted with selected access rights dedicated for a given instance of the virtual machine. With that an access to a zone with given *ZoneID* is

¹ 2015+ Western Digital Technologies and others.

² Physically the write head is wider than the read head, so SMR condenses the tracks, but overwriting part of the track as next track is being written, but only to the point where the read head can still work.



separated to one and only client and the client's data are always defragmented. For resilience each zone is mirrored on a different SMR device (RAID0).

Milestones

I. In-house Edge Computing solutions ready

At this level the completed solution will be evaluated for Linbedded client installations

II. Enterprise ready

Reaching here means having established a de-facto standard for the Industry and support the major file-systems, operating systems and cloud orchestration software tools.