pFS: A partitioned filesystem targeting Virtual Machine images

Anastasios Papagiannis, Yannis Sfakianakis, Stelios Mavridis,
Manolis Marazakis and Angelos Bilas
{apapag, jsfakian, mavridis, maraz, bilas}@ics.forth.gr

Abstract

In this poster we present a new filesystem for storing virtual machine images. In the current filesystem design all I/O calls pass through a singe path inside the Linux kernel, resulting in contention on shared resources and interference along independant virtual machine images. We propose a partitioned I/O path through Linux kernel to minimize the contention and inteference. This partitioned scheme contains a filesystem and an allocator. The other parts are a partitioned DRAM cache and partition journal meachanism which are beyond the scope of this work.

Motivation

- On today's servers, data-centric applications can experience severe performance degradation when ran concurrently with other applications.
- The performance of a workload can vastly vary due to I/O, even though the load of the server seems to be relatively low.
- Our intention is to isolate the I/O path providing that way almost dedicated I/O resources for each VM, ensuring that the workloads will issue I/O undisturbed.
- Our aim is to make an individual server operate with maximum efficiency without migrating VMs.

Filesystem

We provide several optimization for files that are VM images:

- Different I/O path through Linux kernel to minimize the interference across different VMs.
- Remove the indirect and double-indirect pointers from inodes that translates from file offset to block offset. This allows the faster translation from file offset to block offset.
- Because of the limited need of resizing we can preallocate the VM image. Thus we can allocate the VM images as a big contiguous allocation in the I/O devices.
- Separate the reliabily from the filesystem. Filesystem exposes transactions from another device to ensure atomicity.

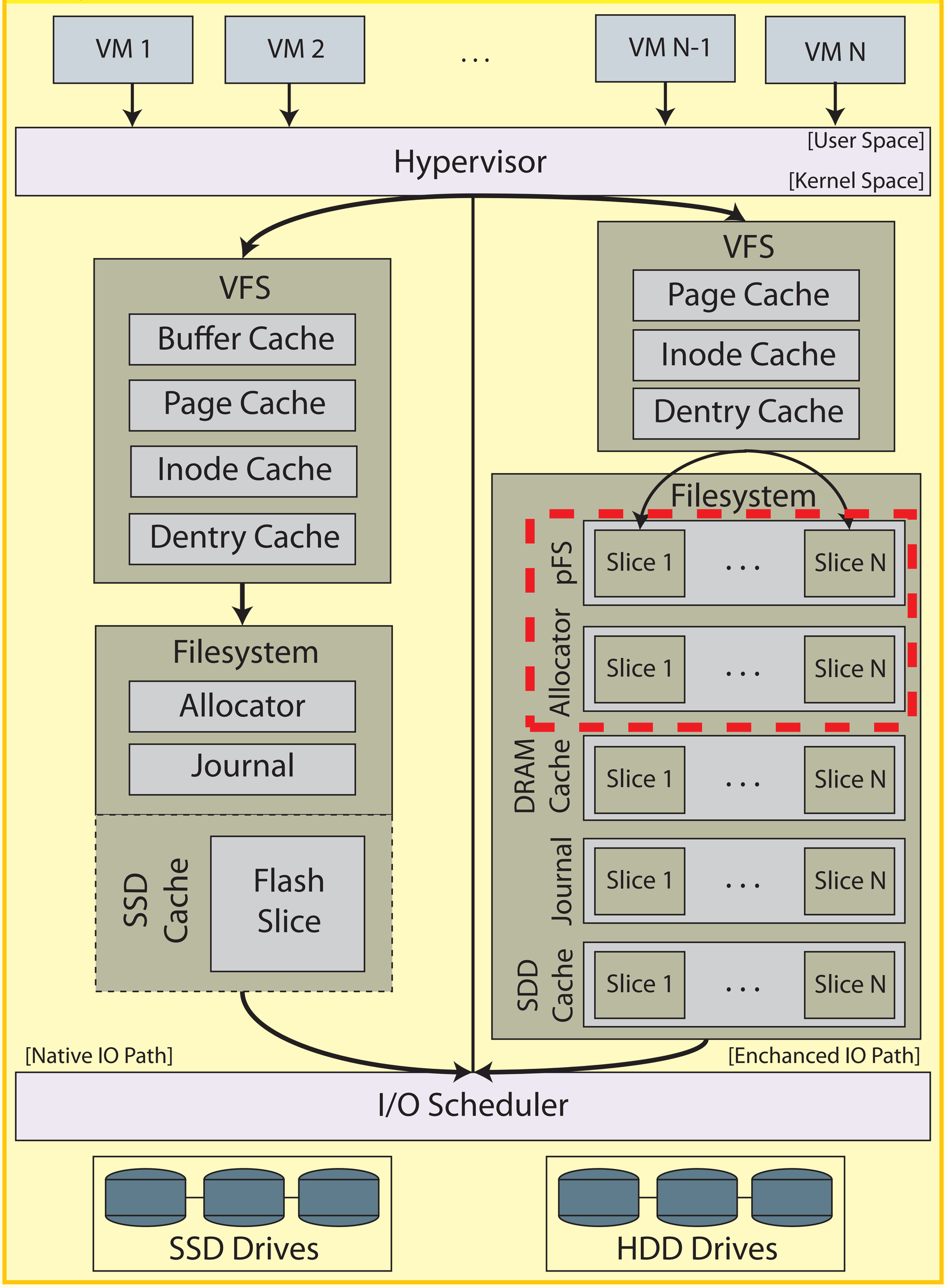
References

[1] Yannis Klonatos, Manolis Marazakis, and Angelos Bilas. A Scaling Analysis of Linux I/O Performance. Poster in *Eurosys 2011*

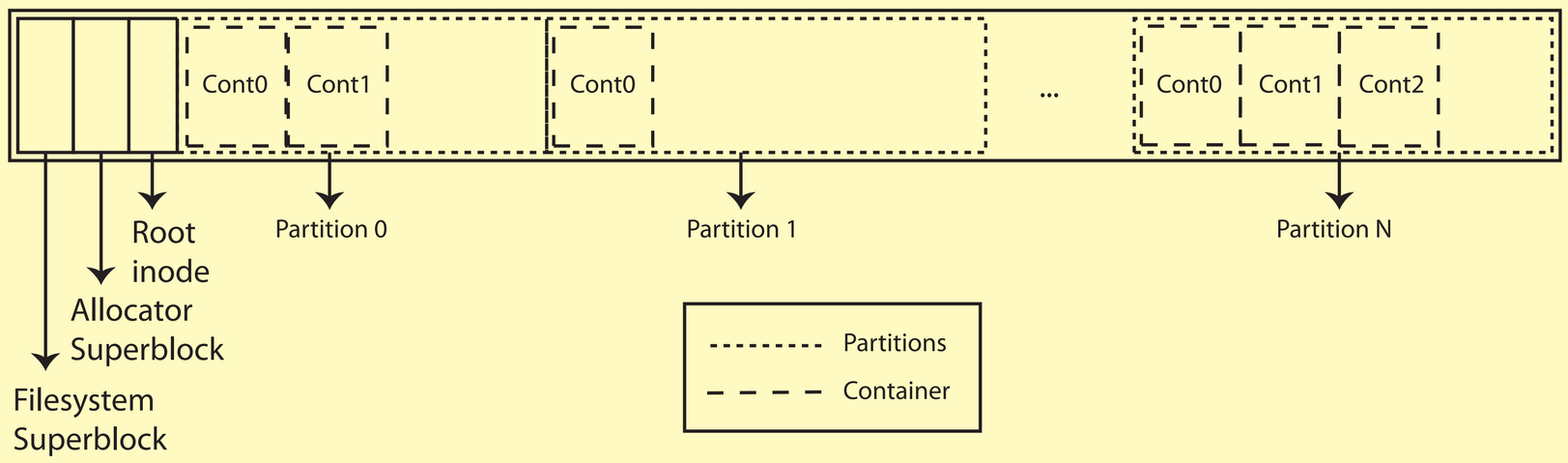
[2] Melanie Kambadur, Tipp Moseley, Rick Hank, and Martha A. Kim. Measuring interference between live datacenter applications. In *Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis, SC '12*

[3] Y. Koh, R. Knauerhase, P. Brett, M. Bowman, Z. Wen and C. Pu. An analysis of performance interference effects in virtual enviroments. In *IEEE International Symposium on Performance Analysis of Systems Software, 2007. ISPASS 2007.*

Design



Allocator



Acknowledgments

We thankfully acknowledge the support of the European Commission under the 7th Framework Programs through the IOLANES (FP7-ICT-248615) and HiPEAC3 (FET-ICT-287759) projects.