

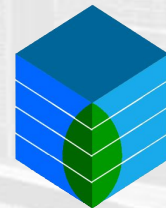


**Barcelona
Supercomputing
Center**
Centro Nacional de Supercomputación



HetFS: A Heterogeneous File System for Everyone

Georgios Koloventzos



BigStorage

The research leading to these results has received funding from the European Community under the BIGStorage ETN (Project 642963 of the H2020-MSCA-ITN-2014), by the Spanish Ministry of Economy and Competitiveness under the TIN2015-65316 grant and by the Catalan Government under the 2014-SGR-1051 grant.

HetFS: A Heterogeneous File System for Everyone

- Background
- Our approach
- Results
- Next steps

HetFS: A Heterogeneous File System for Everyone

- Background
- Our approach
- Results
- Next steps

HetFS: A Heterogeneous File System for Everyone

More and more storage media classes
Each one with its benefits and its drawbacks

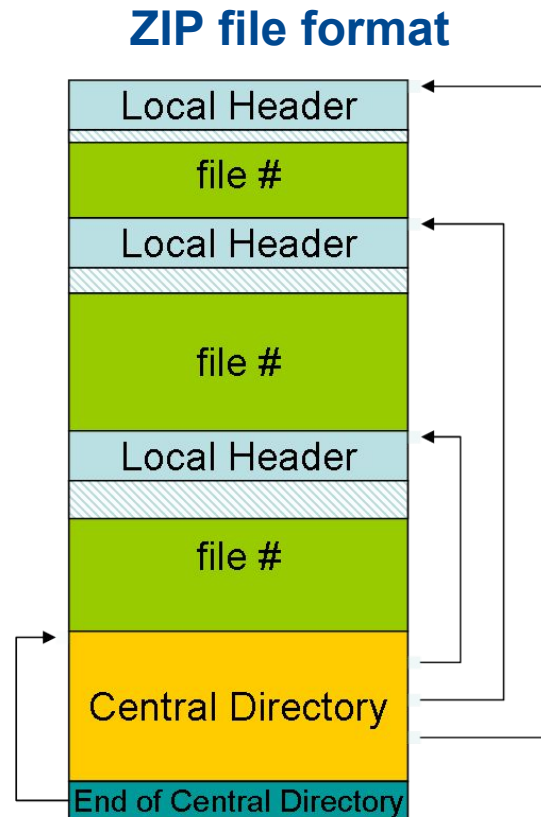
- HDD
- SSD - NAND
- NVRAM
- Helium-filled HDD
- Shingled magnetic recording (SMR)
- Combination of HHDD -SMR
- Network Storage
-
-
-
-



HetFS: A Heterogeneous File System for Everyone

More and more file formats (complex structures)

- Pointers
- Bitmaps
- Tables
- Data
- Header
-
-
-



 extra data in Local Header

NetCDF file format



Records grow in the UNLIMITED dimension

HetFS: A Heterogeneous File System for Everyone

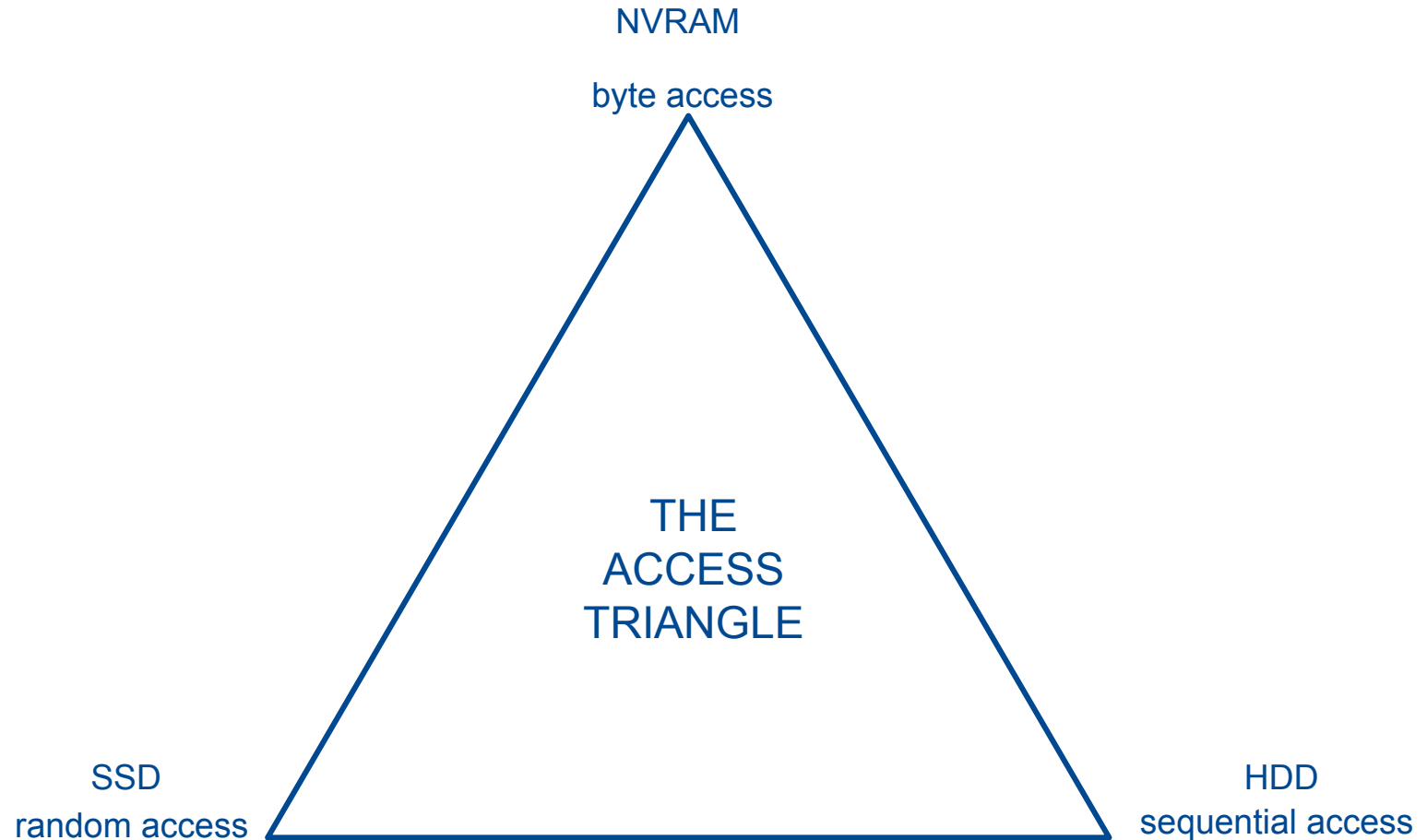
Usage of a file

- Files are mainly accessed in 3 ways (random, sequential, byte accessed)
- Are they accessed at all ? Accessed in special ways (e.g. stripes)
- Do we need the whole file or some blocks?

HetFS: A Heterogeneous File System for Everyone

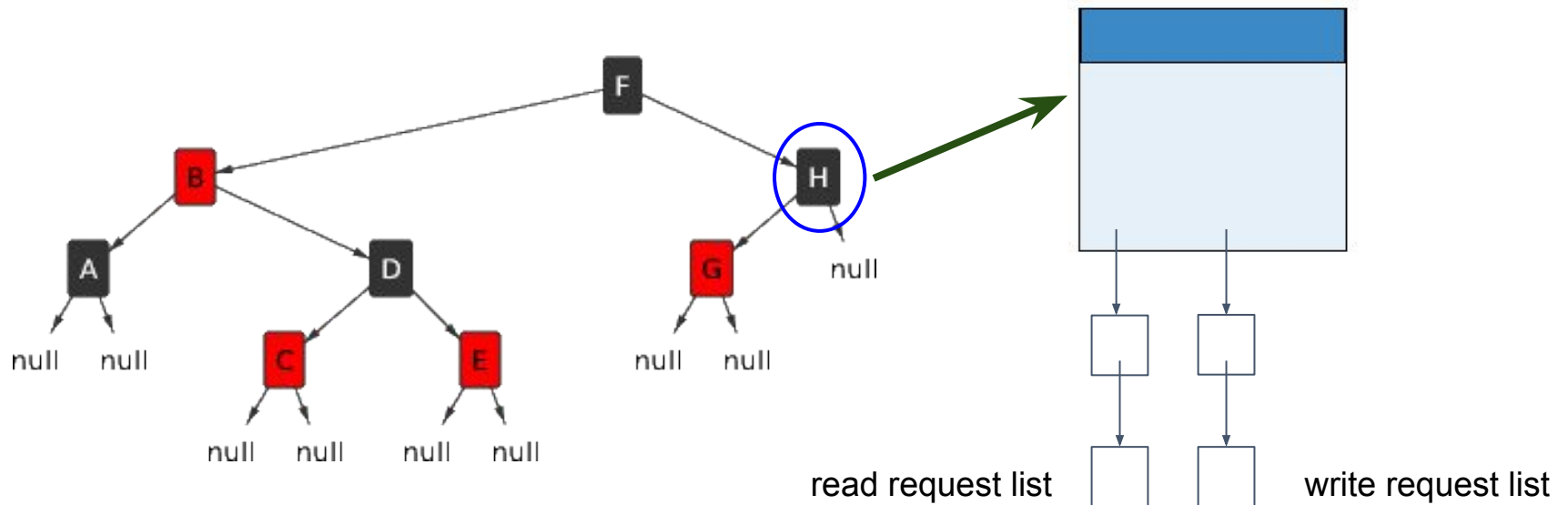
- Background
- Our approach
- Results
- Next steps

HetFS: A Heterogeneous File System for Everyone



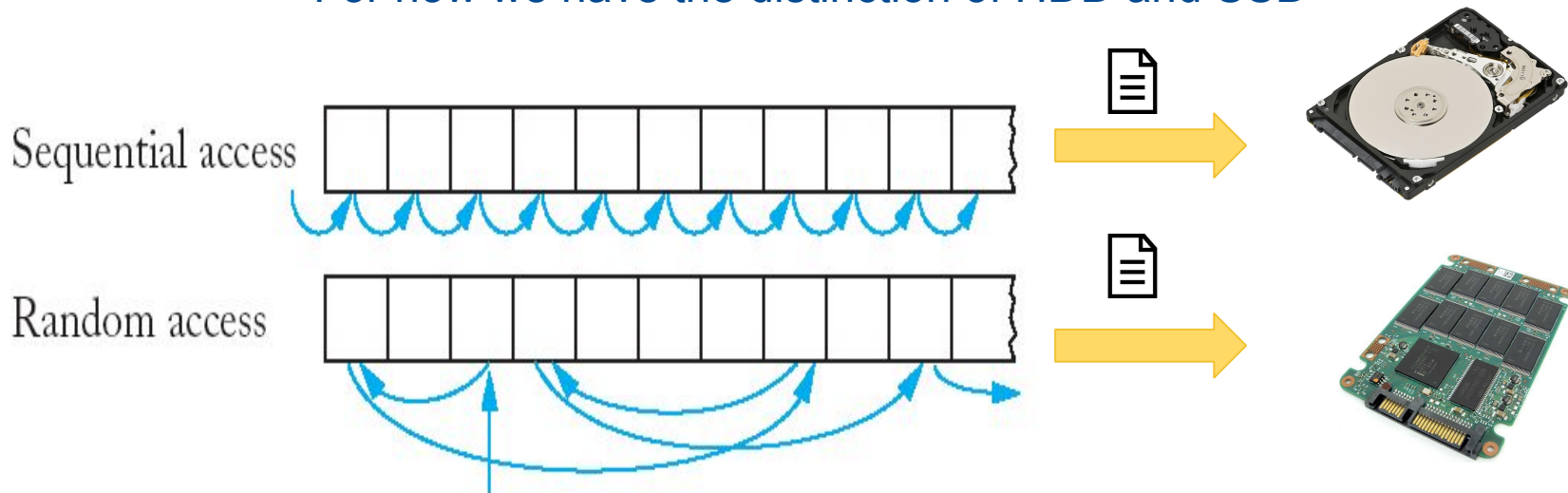
HetFS: A Heterogeneous File System for Everyone

- Log all I/O in a RB-tree for further analysis
 - Each file as a node
 - Each node has 2 lists, for read and write requests respectively
 - If 2 requests have same end - start and do not have huge time difference merge them as one



HetFS: A Heterogeneous File System for Everyone

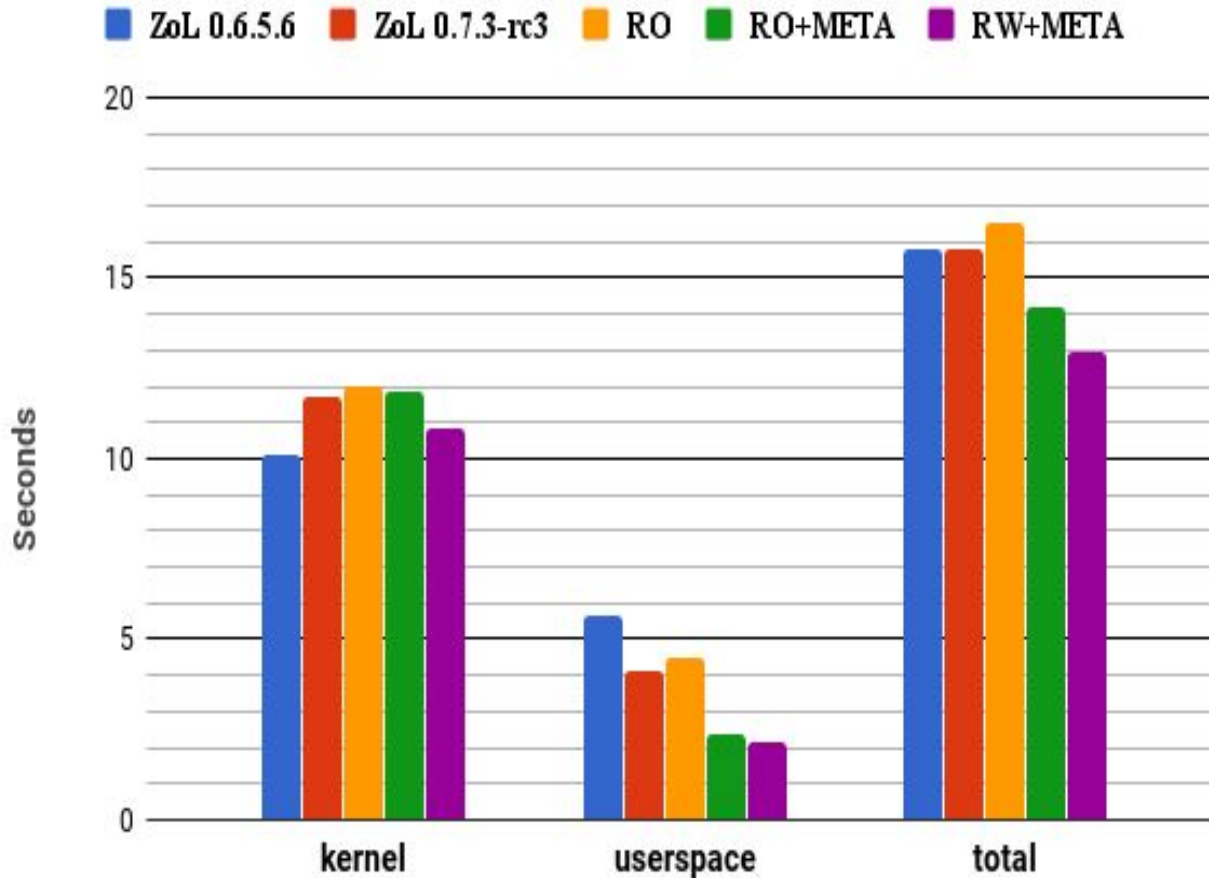
- Analyse the I/O patterns of files
 - Which parts of files are accessed more.
 - Is more than 50% of file accessed on a burst of requests
- The analysis distinguish where a part of a file should be placed
 - more than 50% accessed regularly -> sequential
 - random bursts all over -> random
 - byte - or really small parts -> byte access
 - For now we have the distinction of HDD and SSD



HetFS: A Heterogeneous File System for Everyone

- Background
- Our approach
- **Results**
- Next steps

HetFS: A Heterogeneous File System for Everyone



Boot files on SSD

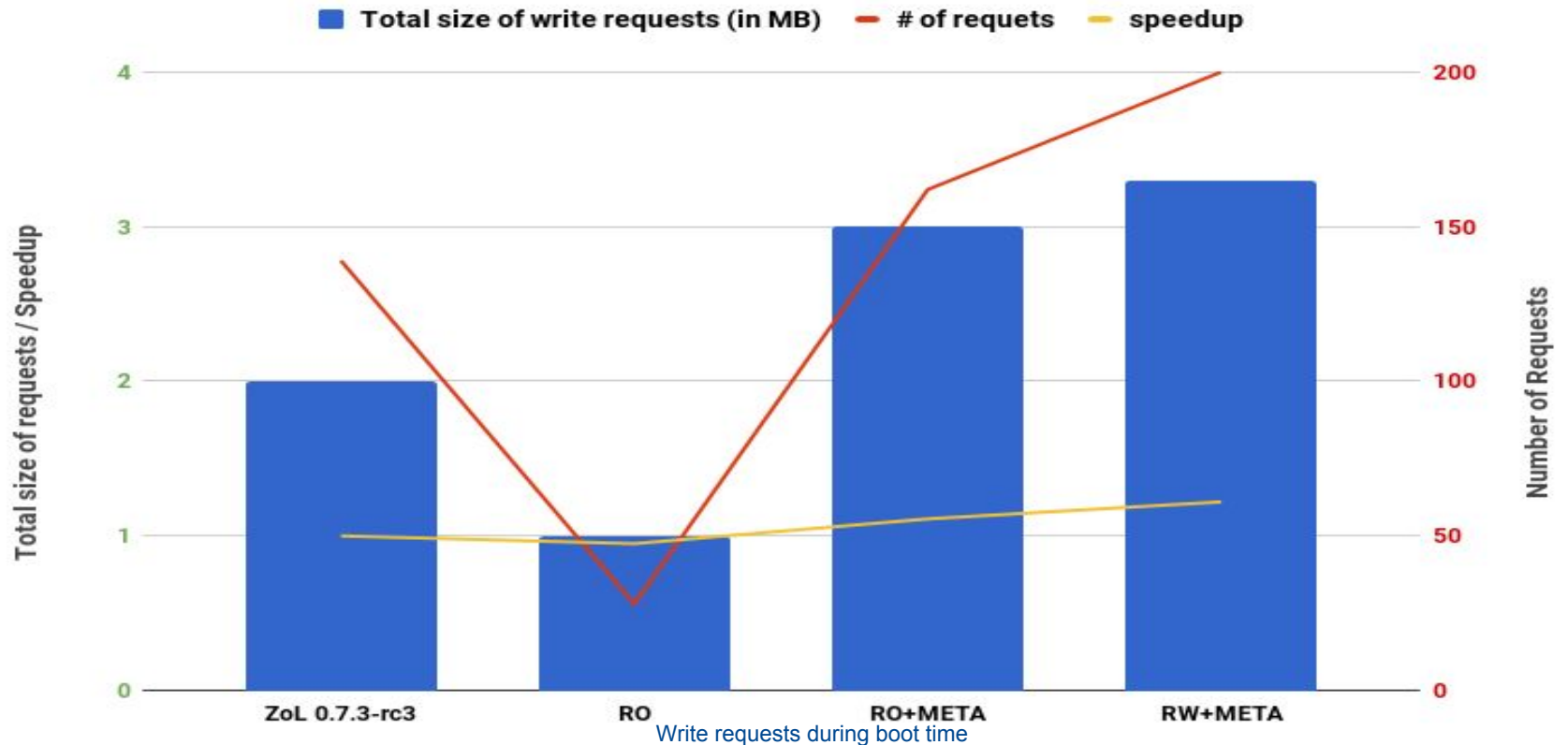
RO - Read Only files
RW - Read and Write files
META - Metadata of files

Our approach of logging the I/Os and place read only files to SSD leverage less than 1% degradation.

Write files are mostly the kernel logs.

Having all files to SSD has a 22% faster boot time. But this will decrease the life expectancy of the SSD.

HetFS: A Heterogeneous File System for Everyone



- Using SSD only for read files reduces 50% the size of the requests
 - The 1 MB are from the prime data structures of ZFS which are out of our scope
- If a Sys Admin chooses to boost the boot time now has an informed opinion how much this will impact the SSD

HetFS: A Heterogeneous File System for Everyone

- Background
- Our approach
- Results
- Next steps

HetFS: A Heterogeneous File System for Everyone

- Make analysis and file movement automatic
- Move part of files in specific medium
 - e.g In boot sequence, only the first blocks need to be read.
 - Add more media in our test bed
 - Examples from popular applications
- Sys Admins will eventually pick from roles such as “web server, file server, video stream, ...”
 - Such roles will have predefined files/directories that should be in a faster medium
- Everything done without specific hardware or complex configurations
 - Our solution could be implemented from HPC to home stations



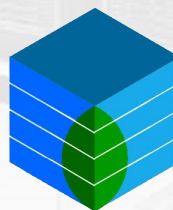
**Barcelona
Supercomputing
Center**
Centro Nacional de Supercomputación



HetFS: A Heterogeneous File System for Everyone

Georgios Koloventzos
georgios.koloventzos@bsc.es

THANK YOU!



BigStorage