# Open Source Data Deduplication

Nick Webb

Red Wire Services, LLC
nickw@redwireservices.com www.redwireservices.com @disasteraverted

RED WIRE SERVICES

SHELTER YOUR DATA

# Introduction

- What is Deduplication?  Different kinds?
- Why do you want it?
- How does it work?
- Advantages / Drawbacks
- Commercial Implementations
- Open Source implementations, performance, reliability, and stability of each (disclaimer)

# What is *Data Deduplication*

Wikipedia:

. . . data deduplication is a specialized data compression technique for eliminating coarse-grained redundant data, typically to improve storage utilization. In the deduplication process, duplicate data is deleted, leaving only one copy of the data to be stored, along with references to the unique copy of data. Deduplication is able to reduce the required storage capacity since only the unique data is stored.

Depending on the type of deduplication, redundant files may be reduced, or even portions of files or other data that are similar can also be removed . . .

# Why Dedupe?

- Save disk space and money (less disks)
- Less disks = less power, cooling, and space
- Improve Writes (maybe)
- Be smart

# Where does it Work Well?

- Backups/Archives (Secondary Storage)
  - Online backups with limited bandwidth
  - Save disk space – additional full backups take little space
  - Replicate disk to disk over slow links, only unique blocks are transfered
- Virtual Machines (Primary & Secondary)
- File Shares (Primary)

# Not a Fit

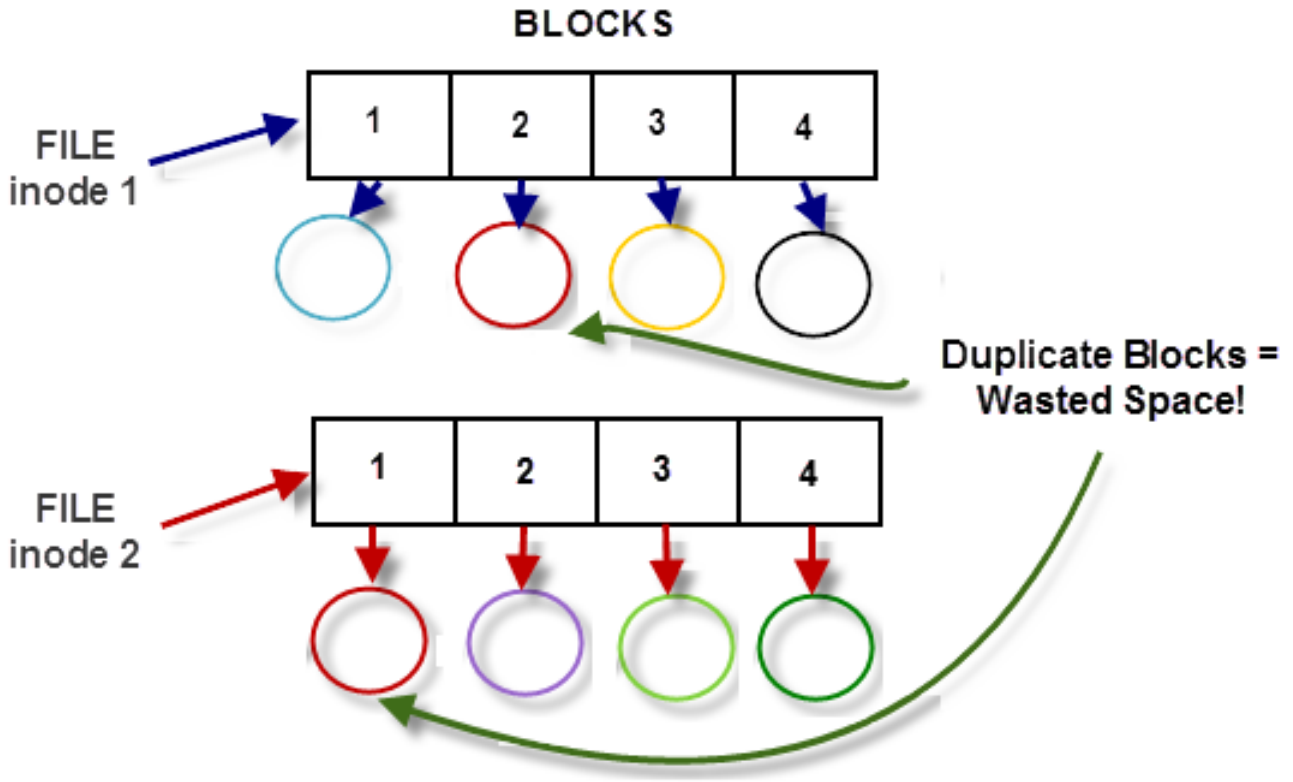
- Unique data
  - Video
  - Pictures
  - Music
  - Encrypted files (many vendors dedupe, then encrypt)

# Drawbacks

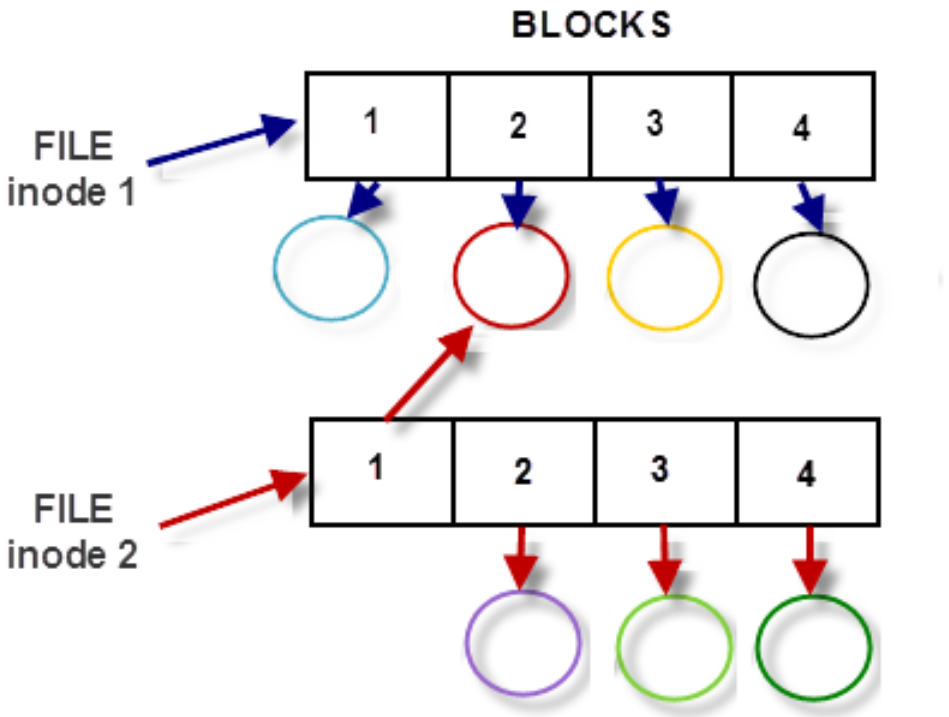
- Slow writes, slower reads
- High CPU/memory utilization (use a dedicated server)
- Increases data loss risk / corruption

# How Does it Work?

# Without Dedupe



# With Dedupe



# Block Reclamation

- In general, blocks are not removed/freed when a file is removed
- We must periodically check blocks for references, a block with no reference can be deleted, freeing allocated space
- Process can be expensive, scheduled during off-peak

# Commercial Implementations

- Just about every backup vendor
  - Symantec
  - CommVault
  - Cloud:  Asigra, Baracuda, Dropbox, Mozy, etc.
- NAS/SAN/Backup Targets
  - NEC HydraStor
  - DataDomain/EMC Avamar
  - Quantum
  - NetApp

# Commercial Implementations

- Dedupe is a competitive advantage on the commercial side
- Most do it, but many different flavors
- Fixed block versus sliding block size
- Global Dedupe

# Open Source Implementations

- Lessfs
- SDFS (Open Dedupe)
- ZFS
- btrfs (soonish)
- Limited (file based / SIS)
  - BackupPC (reliable!)
  - Rdiff-backup
- Others?

# How Good is it?

- Many see 10-30x deduplicaiton meaning 10-30 times more logical object storage than physical
- Especially true in backup or virtual environments

# SDFS / OpenDedupe

www.opendedup.org

- Java 7 Based / platform agnostic
- Uses fuse
- S3 storage support
- Snapshots
- Inline or batch mode deduplication
- Supposedly fast (290MBps+ on great H/W)
- Support for global/clustered dedupe
- Probably most mature OSS Dedupe (IMHO)

# SDFS

- Pro
  - Works when configured properly
  - Appears to be multithreaded
- Con
  - Slow / resource intensive (CPU/Memory)
  - Fragile, easy to mess up options, leading to crashes, little user feedback
  - Standard POSIX utilities do not show accurate data (e.g. df, must use getfattr -d <mount point>, and calculate bytes → GB/TB and % free yourself)
  - Slow with 4k blocks, recommended for VMs

# LessFS

www.lessfs.com

- Written in C = Less CPU Overhead
- Have to build yourself (configure && make && make install)
- Has replication
- Uses fuse

# LessFS

- Pro
  - Does inline compression by default as well
  - Reasonable VM compression with 128k blocks
- Con
  - Fragile (i.e. not robust)
  - Single Threaded?
  - Stats/FS info hard to see:

    ```
    sudo cat /fuse/.lessfs/lessfs_stats|less
    ```
  - Per file accounting, no totals

# Other OSS

- ZFS?
  - Tried it, and empirically it was a drag, but I have no hard data (got like 3x dedupe with IDENTICAL full backups of Vms)

# Kick the Tires

- Test data set; ~330GB of data
  - 22GB of documents, pictures, music
  - Virtual Machines
    - 220GB Windows 2003 Server with SQL Data
    - 2003 AD DC ~60GB
    - 2003 Server ~8GB
    - Two OpenSolaris VMs, 1.5 & 2.7GB
    - 3GB Windows 2000 VM
    - 15GB XP Pro VM

# Kick the Tires

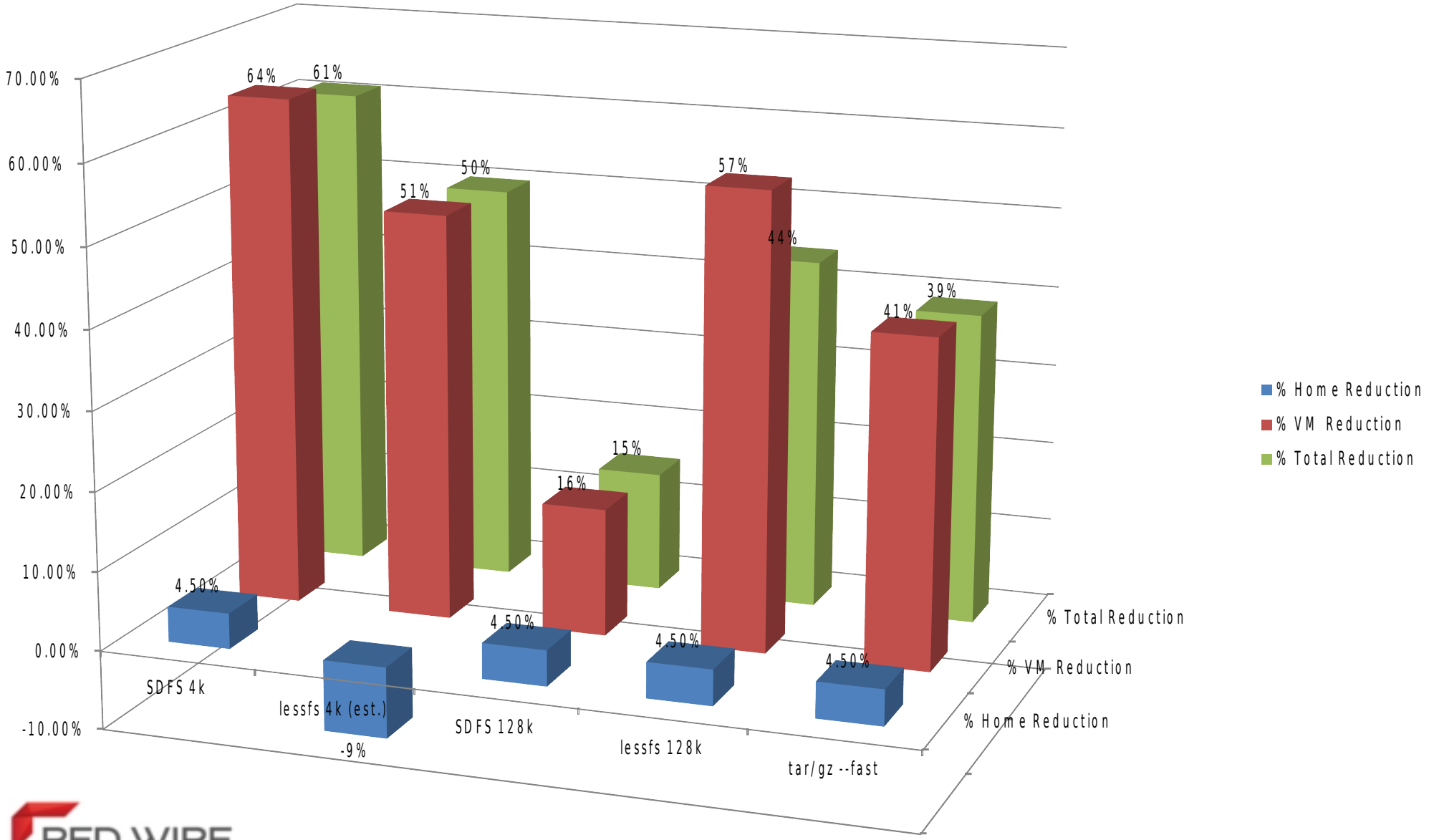
- Test Environment
  - AWS High CPU Extra Large Instance
  - ~7GB of RAM
  - ~Eight Cores ~2.5GHz each
  - ext4

# Compression Performance

- First round (all “unique” data)
- If another copy was put in (like another full), we should expect 100% reduction for that non-unique data (1x dedupe per run)

| FS | Home Data | % Home Reduction | VM Data | % VM Reduction | Combined | % Total Reduction | MBps |
|---|---|---|---|---|---|---|---|
| SDFS 4k | 21GB | 4.50% | 109 GB | 64% | 128GB | 61% | 16 |
| lessfs 4k (est.) | 24GB | -9% | N/A | 51% | N/A | 50% | 4 |
| SDFS 128k | 21GB | 4.50% | 255 GB | 16% | 276GB | 15% | 40 |
| lessfs 128k | 21GB | 4.50% | 130 GB | 57% | 183GB | 44% | 24 |
| tar/gz --fast | 21GB | 4.50% | 178 GB | 41% | 199GB | 39% | 35 |

# Compression (“Unique” Data)

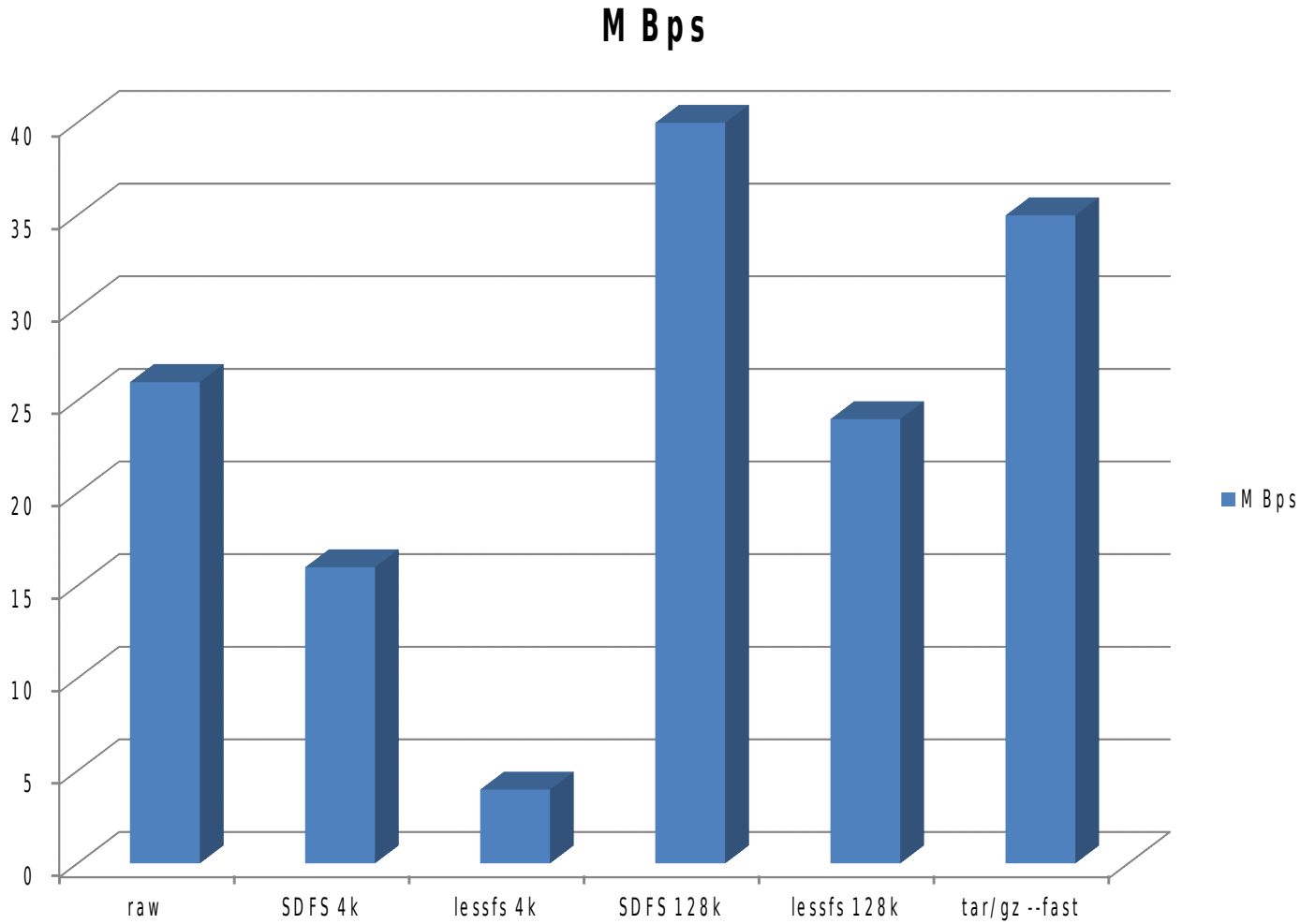



RED WIRE SERVICES

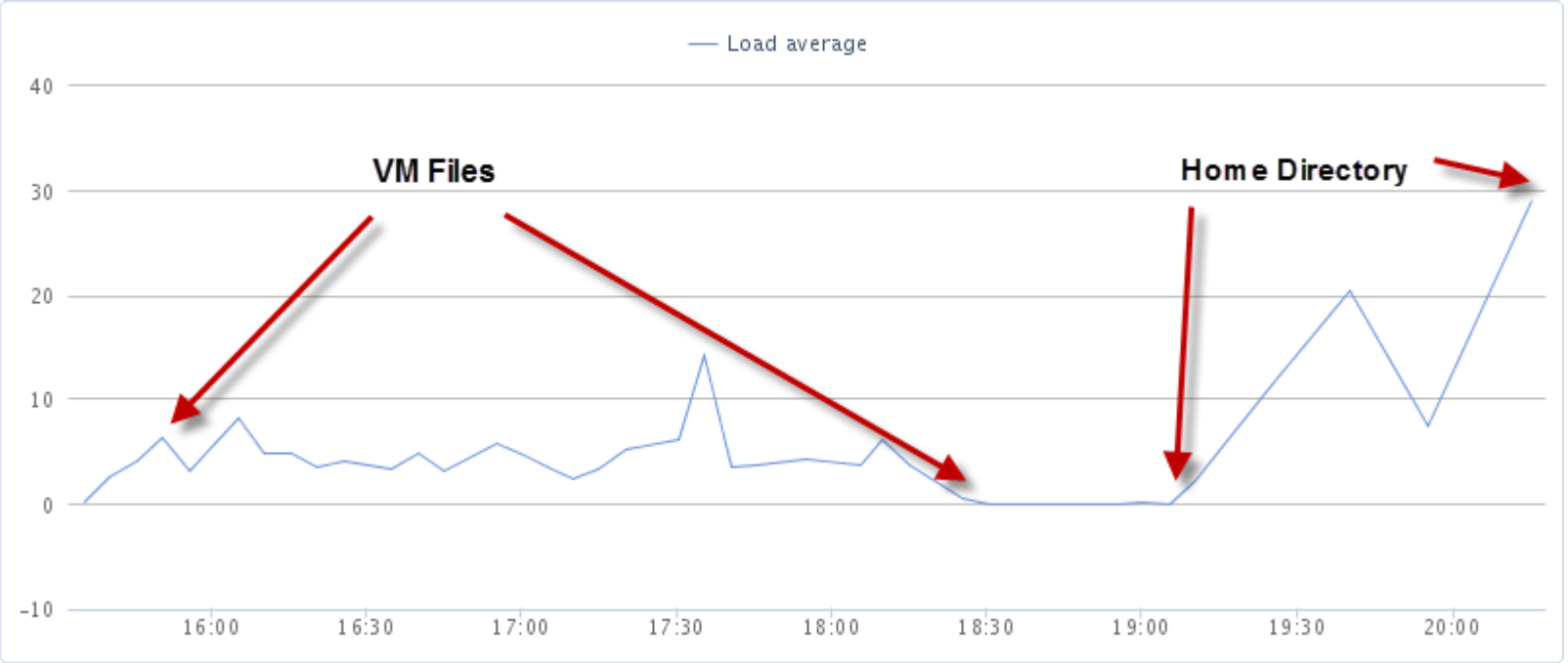
SHELTER YOUR DATA

# Write Performance

(don't trust this)



# Load
(SDFS 128k)



# Open Source Dedupe

- Pro
  - Free
  - Playtime
- Con
  - Unstable, PIA, not in repos yet
  - Efforts behind them seem very limited
  - No/Poor documentation

# The Future

- Right now all kinds of backup and storage businesses are pushing dedupe, but eventually it will become zip, and everyone will expect it
- Until then, everyone is fighting over what is better, and the OSS community seems to be waiting...
- brtfs
  - Dedupe planned (off-line only)

# Conclusion

- Dedupe is great, if it works and it meets your performance and storage requirements
- Currently, after 20+ hours using OSS Dedupe, I feel it is not ready
- Especially considering commercial versions; sometimes it's cheaper just to by 10x storage.
- If you have to do it now, in production, go commercial (sorry) – unless . . .

# Questions?

# Open Source Data Deduplication

Nick Webb

nickw@redwireservices.com
@disasteraverted

Linuxfest Northwest 2011

RED WIRE SERVICES

SHELTER YOUR DATA