

Storage Caching

Tim Fletcher
Brighter Connections



Who are we?

15+ year Linux experience

13 years work in education

RHCE since 1999 (RedHat 6.1)

Joined Brighter Connections in October

IT Practise in Huddersfield

Specialise in innovative solutions

Why your SAN is slow

Centralised storage, good for cost bad for performance

Lots of different competing workloads mean disks "thrash"

Making your SAN faster

Hardware

Spindles

RAM

Software

Copy on Write

Quality of Service

Hard Drives

Technology basically the same for over 30 years, known as "Spinning Rust".

Good for cheap bulk storage

Good at linear access

Bad at random access

Power hungry

Fragile



Memory

Very fast access

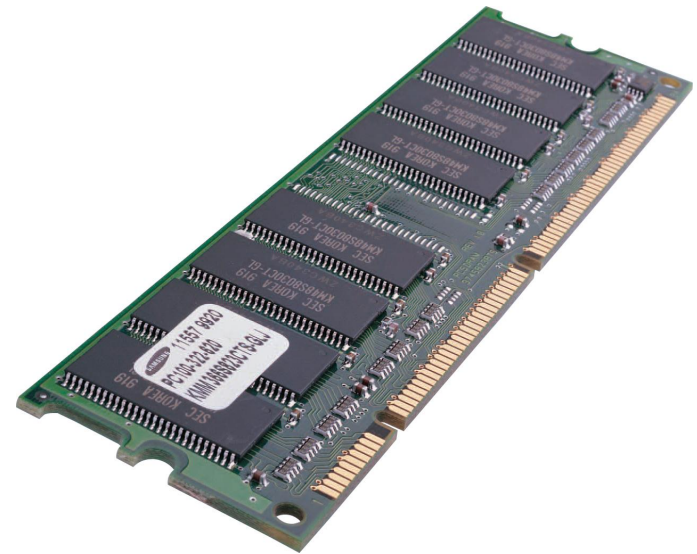
Very good random access

Directly connected to the CPU

Expensive

Volatile

Power Hungry



Flash

Very fast, especially at random access

Can be connected directly to the PCI-e bus

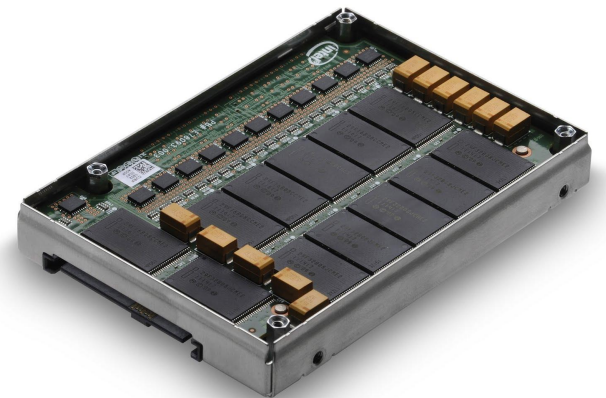
Non volatile

Low power

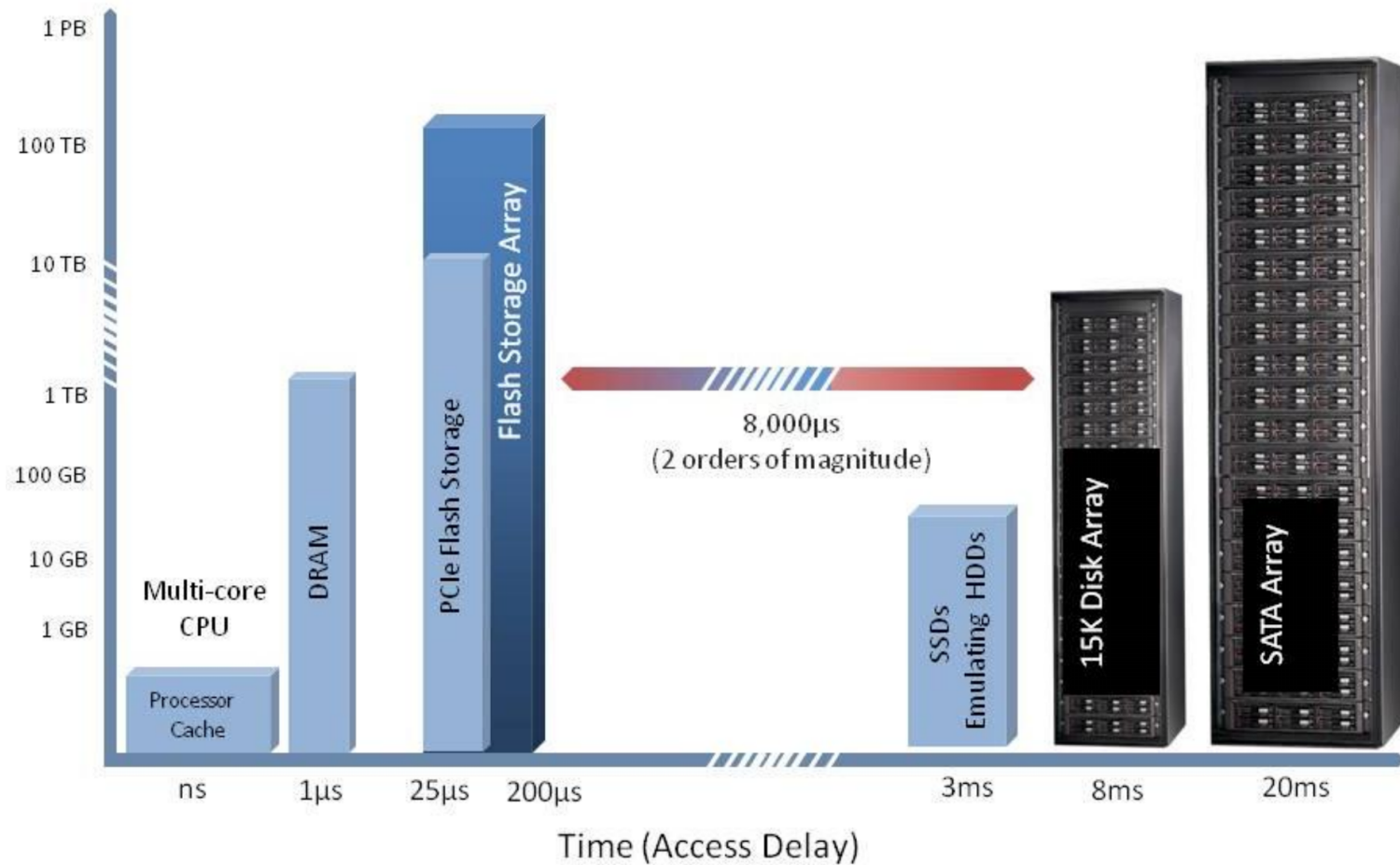
Massive drops in price, but still expensive

Cells can wear out

Includes firmware



Storage Speeds



PCIe vs Hard disk emulation

Hard disk emulation

- Easy to deploy

- Well understood

- Cheap

- Many layers of translation

PCIe

- Expensive, or is it?

- Higher power draw

- Directly attached to the CPU bus

Caching - Why does it matter

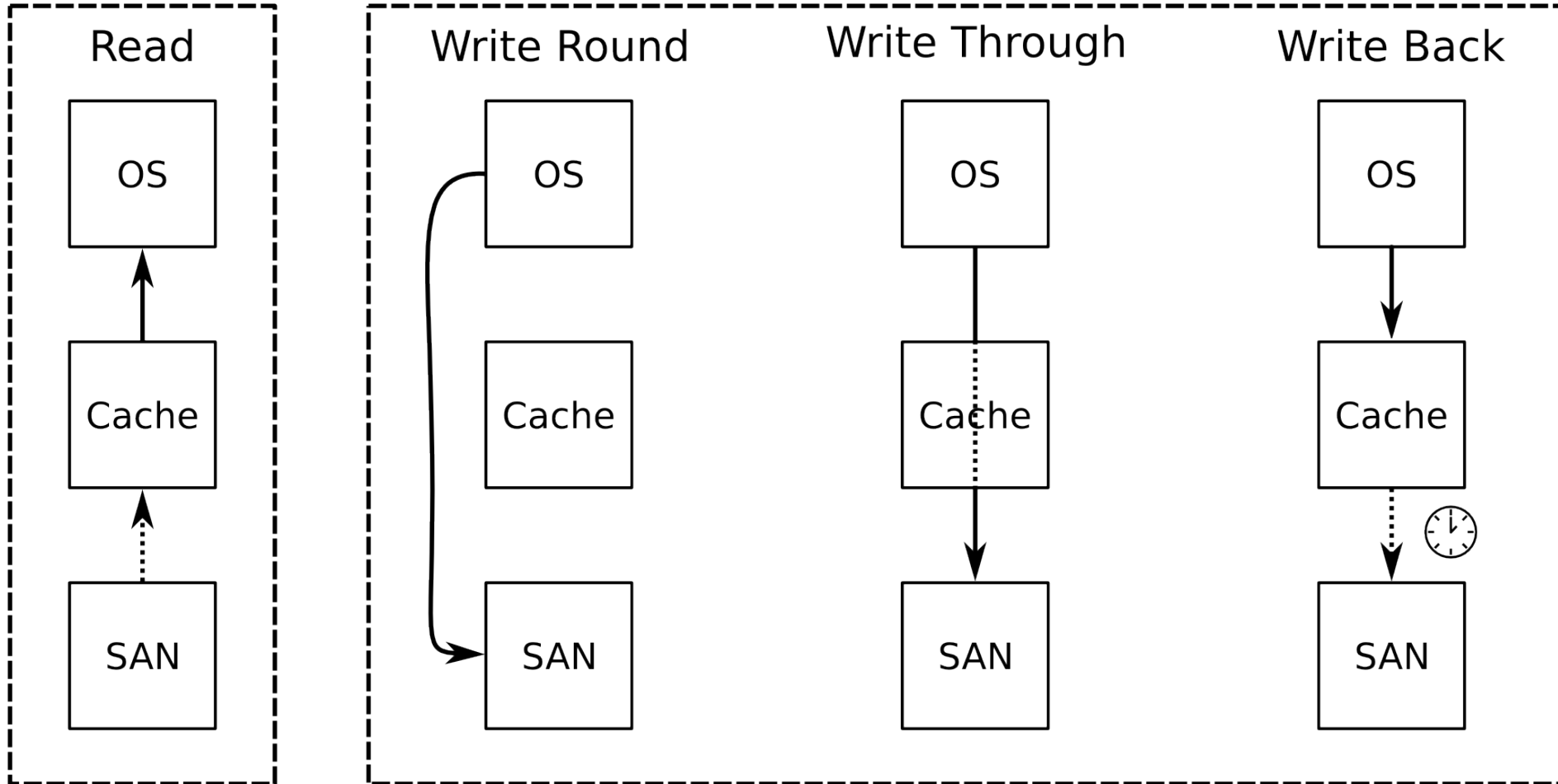
Flash is expensive and wears out

The closer the data is to the CPU the better

Keep your SAN array

Leverages the best of both worlds

Caching - Where do your blocks go?



Flashcache

Fully open source, released by Facebook
Kernel module, based in the device mapper
layer

Starting to appear in distros

Error resistant

Cache can be added and removed from
storage

Used by Facebook in production

Bcache

Fully Open Source

Designed from the ground up to cache

Designed with an understanding of flash

Higher performance

Requires devices to be rebuilt as "bcache"
devices

Requires full kernel rebuild

ZFS

Not just a filesystem

Includes device management, RAID and filesystem management.

Open Source but license incompatible with GPL

Written by Sun, available in FreeBSD, Solaris derivatives, and Linux

Can use SSDs as L2ARC and as SLOG

Cache devices can be live added and removed

Cache does not persist over reboots

Closed Source Options

Hook in to VMware

FusionIO - ioTurbine

SanDisk - FlashSoft

Questions?

tim@brighter-connections.com
www.brighter-connections.com
@timjdfletcher

