Open Ethernet Drive

--What to know?

George@DCSLab

Motivation

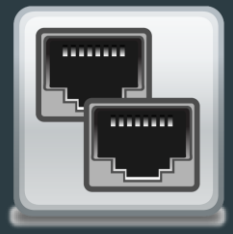
- Distribute ***resource*** to storage media as close as possible

A Drive Running Linux

Leverages the Linux ecosystem

A Drive with CPU & RAM

Run storage service directly on device



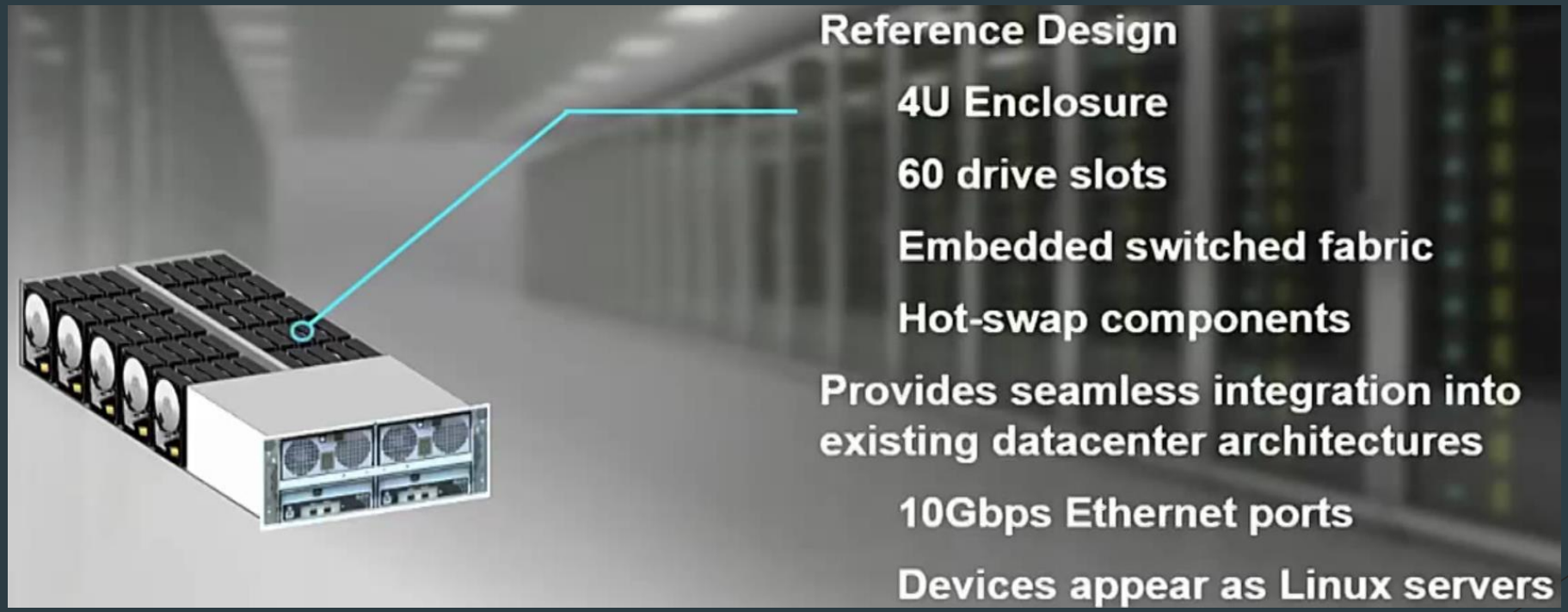
A Drive with Ethernet

Connect storage directly to data center fabric

Architecture

Hornet + drone development board

Architecture



Hive – Chassis JBOD

4U60 Storage Enclosure

Features

- 60 integrated Ultrastar® drive modules
  - — Available with 512e or 4Kn sector size (by part number)
- Host connectivity 2 x SAS QSFP+ receptacles per IO Module; optional connectivity cables are available
  - — QSFP+ to HD mini SAS cables for host
  - — QSFP+ to QSFP+ cables for daisy-chaining*
- Hot Swappable: two power supplies (with integrated fans), drive modules, and IO modules
- Fully compliant with SAS 3.0 specification for operation up to 12Gbps
- Cable management arm simplifies maintenance of hot-swappable CRUs
- SCSI Enclosure Services (SES-3)
- Microsoft Certified for Windows Server 2012 & Windows Server 2012 R2
- 5-year limited warranty on Ultrastar HDD modules, 3-year limited warranty on all other components

*Up to 4 storage enclosures supported for daisy-chaining

The HGST 4U60 Storage Enclosure is a high-density, scalable, and cost-effective design, using 60 Ultrastar® 3.5-inch drive modules in a 4U enclosure. It offers 2x2x4-lane SAS 12GB/s performance, high availability (HA), and hot-swappable components. The design is targeted towards data centers that need a dense solution with HA capabilities, while maintaining a low power profile.

4U enclosure
60 Ultrastar drive modules
Available in 3 capacities[1]: 480TB, 360TB and 240TB

Specs

Version 1

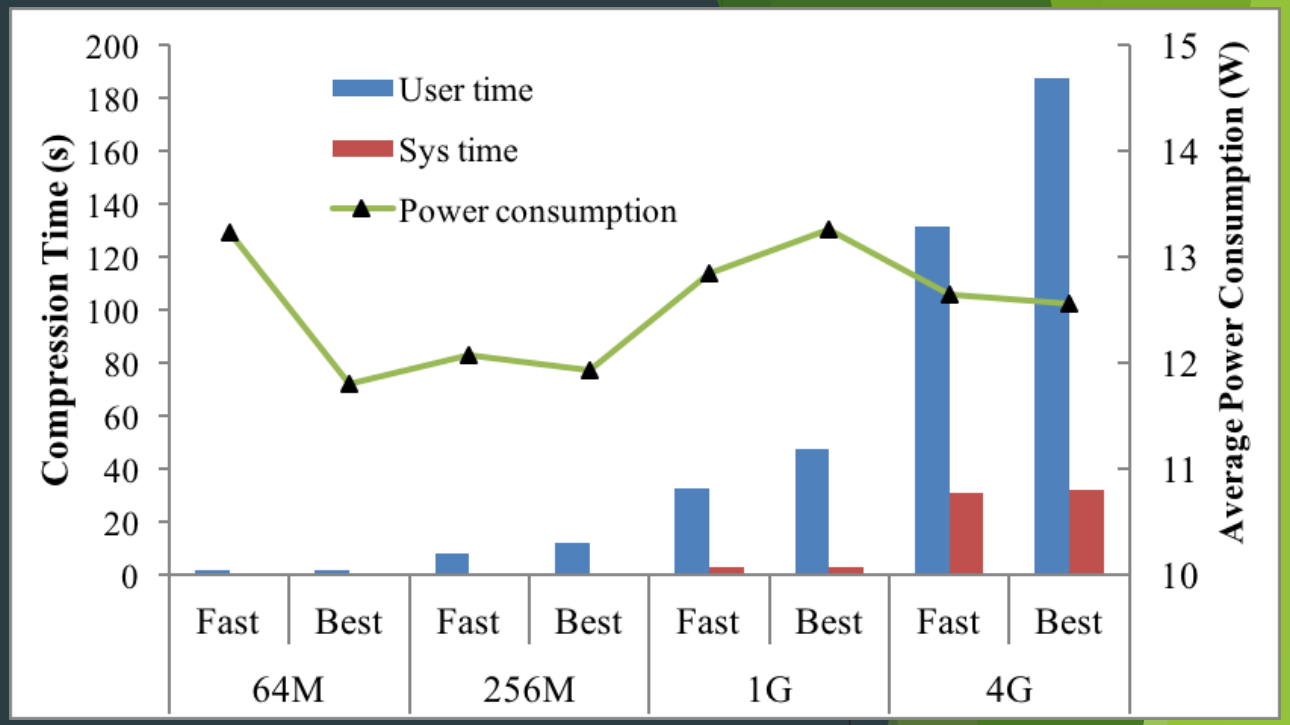
- Linux – Debian 7,4 + (Wheezy) on demo system
- CPU – single core 32-bit ARM, 512KB L2 Cache
- Memory – 2 GB DRAM, DDR-3
  - 1792 MB available to Linux
- Block storage driver – 4TB
  - Drive enumerates as a SCSI disk (/dev/sda)
- Ethernet network driver
  - Enumerates as eth0 device

Version 2

- Debian 8.4 Jessie
- Dual Core ARM
- 1 GB DRAM, DDR3
  - Dedicated to Linux Application
- 8TB

- Only work inside JBOD Chassis

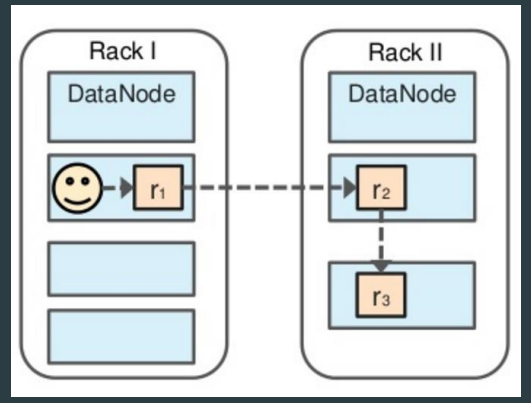
Performance

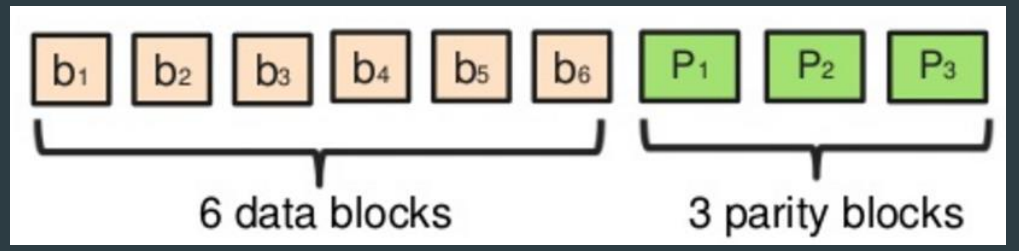


- Gzip file 64M ~ 4G
  - File: dd if=/dev/zero of=…
- Compression throughput:
  - 29.3 MB/s for fast compression – 230 : 1 ratio
  - 22.9 MB/s for best compression – 1030 : 1 ratio
- Power Consumption:
  - 12.5 Watts on average
- Same Dataset on Intel Core Due Desktop:
  - 4 x faster
  - Power Consumption: ~175w for desktop & ~60w for laptop

Application

- OED Cluster + Erasure coding as resilience storage system

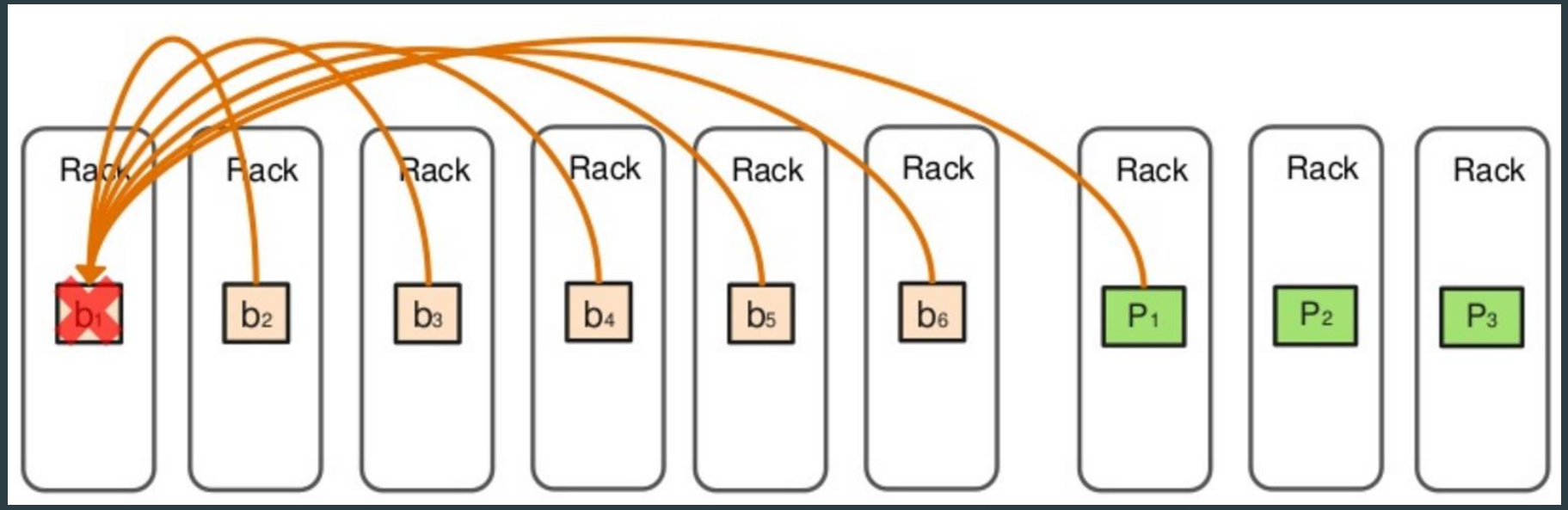




Reed-Solomon 6+3

- k + m erasure coding tolerate m failures
- Save disk spaces & IO bandwidth on write path

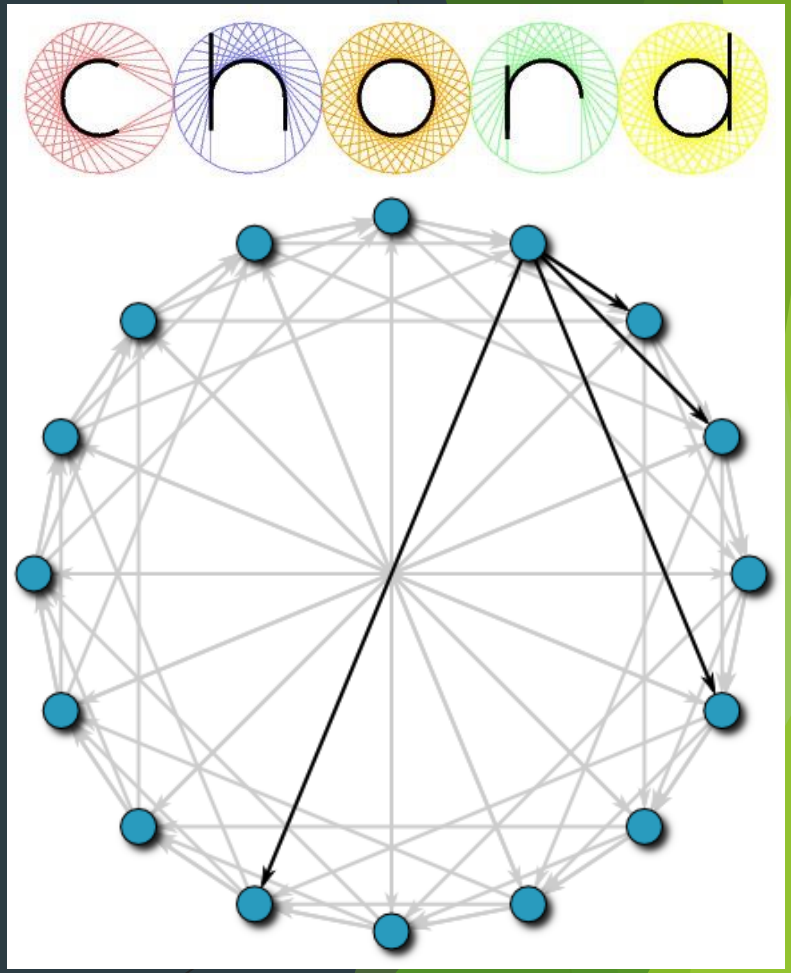
Application



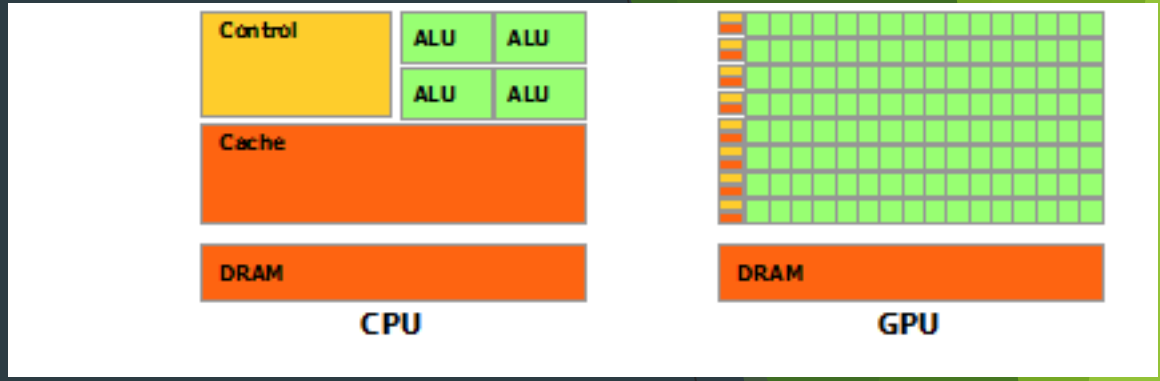
Block Reconstruction

Application

- P2P network using Distributed Hash Table (DHT)
- Near Data Computing



OED vs GPU



- Less computation power each core but overall more capability
- Hardware parallelism
- Through-put oriented
- Power efficiency