NVME-based MidPlane FTL

Bharadwaj Pudipeddi

CTO, BiTMICRO
Abstract: Cloud-Scale PCIe SSDs

- NVMe and SFF-8639 create new challenges: Thermal, Cost, Scale
- Problems of localized FTL (in the SSDs) and fully centralized FTL (in the Xeons)
- The solution: a de-localized shared FTL built for cloud-scale rack-optimized systems
Server-Class PCIe SSDs: Cost Trendline

Flash vs Non-flash Cost in a 2TB PCIe SSD

- Flash Cost
- Non-Flash Cost
- %Overhead

Cost of Goods


Flash Memory Summit 2015
Santa Clara, CA
Common Solutions to Flash Cost Efficiency

- Run FTL in Xeon (even integrate into filesystem)
- Increase density of SSDs (amortize overhead $$)

FTL in CPUs

FAT SSDs
Problems of Current Approaches

- **FTL running on CPUs –**
  - Unpredictable: varies with Xeon workload
  - Does NOT benefit from custom h/w accelerations
  - ECC does not scale

- **FAT SSDs (4TB, 8TB, 16TB etc.)**
  - Harder to design with limited board space
  - Too localized: limited FTL efficiency, costly scale-out
  - Huge failure disruption: losing 8TB is worse than 2TB
  - Wasted PCIe Energy: Performance bottleneck in CPU’s I/O path
First: Anatomy of a Server PCIe SSD

PCIe G3

ECC DRAM

NVMe and FTL: Power ~5W

DMA+Ac

cel

SRAM

Pn

P0

NVME

Flash Controller & Flash Memory: Power: 0.8W/TB

 Toggle ONFI

FSM

LDPC

Custom

Power: 0.8W/TB

Power: ~5W

Flash Memory Summit 2015
Santa Clara, CA
The Midplane Solution:

- De-localizing FTL: Separate NVMe+FTL Card
  - Larger Pool of Flash: Higher Data Efficiency
  - Consistent FTL performance
  - Scope for Custom H/W features
- Separate Flash Memory Modules: Low-power
  - Scaling Error Correction with Capacity
  - Higher Capacity with more Board Space
NVMe-based Midplane FTL

NFC: NVMe/FTL Controller
FMC: Flash Memory Controller

Flash Memory Summit 2015
Santa Clara, CA
Advantages of Midplane-FTL Server

- Overall: Much better power efficiency
- Thermal problem localized to NVMe FTL Controller (NFC)
- Easier Hot Plug / Add Solutions for FMCs
- Higher Density

Expect to use 50% less power, 30% more density, and much lower costs
**Flash8x2TB = 1 NFC + 8 FMC**

**Flash6x2TB = 1 NFC + 6 FMC**

**Topology: 1-level**

**Power Savings:** 80% (Flash8x2TB); 50% (Flash6x2TB)
Challenges for Midplane Solution

- NVMe FTL Controller (NFC) – ability to parallelize and efficiently run FTL across larger pools
- High B/W Backplane “cable-friendly” Interface
- NFC Failure FRS (Dual-NFC mode)
- NFC Metadata Power-loss Protection

Using smart hardware customization, and spanning FTL algorithms – these challenges can be met!
Questions?
Bharadwaj Pudipeddi (short form: Bharad)
- CTO and VP of Engg, BiTMICRO Networks, Fremont, USA
- BiTMICRO which makes military-grade SSDs has recently announced SSD and All Flash Arrays for Cloud/Enterprise
- Previously worked as Architect at Violin Memory, 3Leaf Systems, and Intel Corp on key products including AFAs and Xeon processors
- Please contact me at bharadwaj@bitmicro.com
- Visit us at the BiTMICRO stall in the FMS exhibit hall