Using Open Channel SSDs to Improve QoS for Datacenter Applications

Xueshi Yang, PhD
CEO, Shannon Systems
August 2018
Open Channel SSDs for the datacenter
Transactions, the history
Transactions, today
Transactions: Volume

2007: 1.5 Billion
2017: 3.5 Billion

Increase: 2.3x
Transactions: Velocity

¥10,014,231,205

$1,510,861,063

Top-up GMV in USD.
GMV for GMV from Mobile 93%.
900,000,000 swipes a day!

Transactions: Variety
HyperScale Datacenter
Powered by NAND Flash
- Higher QoS
- Lower TCO
Open Channel SSDs for the datacenter
Applications
Middleware
Database

VFS
XFS, EXT4...
Block layer
NVMe Driver

FTL
NAND Flash

User
Kernel
Hardware
The Agent Problem

The Matrix, 1999
Open Channel SSD: Reducing Agents

- APIs tailored for applications
- Data hints sent to IO managers
- NAND status hints sent to Applications
Open Channel SSD: Eliminating Agents

Applications with IO and flash management

Physical Block Addressing: Channel/CE/Block/Page

NAND
OCS in **Kernel** Mode

- **User**
- **Kernel**
  - IO Engine (Block Device/NS/Key/Value etc) and flash manager
  - LightNVM Driver
- **Hardware**
  - Controller logic (R/W/E) + NAND
OCS in User Mode

- **User**: IO Engine (Block Device/NS/KeyValue etc) and flash manager
- **Kernel**: LightNVM Driver
- **Hardware**: Controller logic (R/W/E) + NAND
OCS Become Integrated Part of Applications
Open Channel SSDs for the datacenter
Introducing Industry’s First Commercialized OCS Platform

The “Venice”
Partnership on OCS conforming to the Alibaba Open Channel SSD Spec

In production today
OCS for the Cloud: QoS Provisioning

- **Instance 0**: High Priority Read, 50K IOPS
- **Instance 1**: High Priority Write, 100K IOPS
- **Instance 3**: Normal Priority Read, 20K IOPS
- **Instance 4**: Normal Priority Write, 100K IOPS

**OCS IO Engine**

**Venice Controller + NAND**
Prioritized IO Reduces Latency

90% reduction
OCS for the Cloud: Thin Provisioning

- Instance 0
  - Volume 0
    - 4TB

- Instance 1
  - Volume 01
    - 4TB

- Instance 99
  - Volume 99
    - 4TB

OCS IO Engine

Venice Controller + NAND (4TB)
OCS for Databases: Atomic Writes

User
- Database
- Any Filesystems

Kernel
- FTL block device with atomic write support
- LightNVM Driver

Hardware
- Venice Controller + NAND

Atomicity
Hints
OCS for Databases
Atomic Writes

Native support after Ver 10.2
OCS Atomic Writes – MariaDB® Support

- TPS increases by up to 50%

OCS for NoSQL: **Key-Value**

- KV access
- Kernel/user-space mode
- Physical based addressing

**KV Applications**

**User-space KV Engine**

**Venice Controller + NAND**
OCS for NoSQL: Key-Value
**KV-SSD: Breakthrough Performance**

Get(Key):
- LevelDB+SSD: 780%
- KV-SSD: 1,500%

Put(Key):
- LevelDB+SSD: 780%
- KV-SSD: 1,500%

Random 8KB value size access
Partnership on KV-SSD
Key Takeaways

• **OCS have arrived**

• OCS enable applications to **directly manage flash**, bypassing agents and OCS become part of applications

• OCS offer significant opportunity for **QoS improvement** for **datacenter applications**:
  – Cloud and virtualization
  – Databases
  – Key value store
Come to Visit US at **Booth #219**

Next to the food and drink stand!