Applications Need More Than IOPS and Latency

Aug. 2018
Applications Need More Than Just IOPS and Latency

- Performance consistence
- Reliability
- Endurance
- Data Recovery
- Fail Cost
- Offloading
- Multi-vendors
- Application Optimization

IOPS
Latency
Performance Consistency

Write IOPS

Consistency

Interface Limit

~3.6GB/s

PCIe 3.0 X 4

1T -> 16T

Bandwidth no upgrade

IOPS / GB ↓20X

Write IOPS

GC WAF

Consistency

R/W/E congestion

Latency

Bandwidth

~3.5GB/s

~3.2GB/s

1T   ->  16T  Bandwidth no upgrade

IOPS / GB         20X

Performance Consistency

50%

90%

99.9%

99%

99.99%
Reliability: Power Loss Protection and Data Recovery

Power loss protection challenge

SLC

QLC

Big capacity makes data recovery more difficult

SSD

SSD

SSD
Fail drive parts cost higher

Endurance reach critical point

PE Cycle /K

SLC eMLC MLC TLC QLC

Endurance and Cost

≈

64TB

BMW3
How to choose a RAID stripe size?

Erase block

SSD0
16 MB

SSD1
27 MB

SSD2
80 MB

SSD3
36 MB

Differences in Erase Block Size between different manufacturers/generational Flash
Offload or Just Pull Up to The Host

- Mapping Table
- Snap Shot/Copy
- Compress/Decompress
- EC/RAID
- Cache
- Dedup
- Power
- Mirror
- AES/DES
- Scrubbing X25
- Wear Leveling X25
- Bad Block / Page X25

SSD0 | SSD1 | SSD2 | ... | SSD23 | SSD24
The Dialog GAP Between Applications and SSD Drives

- Database
- SDS
- Cloud

Transaction throughput
Transaction time

GAP!

Bandwidth
IOPS Latency
There is no perfect solution to solve everything.

**HOST Manage SSD**

- Multi-Vendor Adoption
- Offloading

**Standard SSD**

- Performance Consistency
- Reliability
- Endurance
- Cost
What We are Doing to Make SSD Even Better?

- Standard SSD Optimization
- PLOG SSD
- Large Scale Design
Smart Multilevel Scheduler: Provides Great Latency

Smart multilevel scheduler

e.g. Priority: user read > system read > user write > system write

64 Client 99% Read Latency
Smart Workload Processing: Enables Applications Experience Better

MySQL®

Aware typical workload, make better schedule management

Typical Read Performance

DB TPMC

V3

V5

10%

27%

Smart streams: Makes More Difference

- Standard
- With Smart-Streams

22.8% 20.9%

Endurance Performance

Standard With Smart-Streams
Dorado Achieves No.1 Performance By Smart-streams

- Smart-streams
- Smart-workload processing
- Smart-scheduler

**Dorado**
Mission-Critical All-Flash Storage
Lightning Fast
7 Million IOPS
Unique FLASHLINK™ Tech

Flash-oriented design, 0.5 ms consistent latency
PLOG SSD: Bring Lots of Fun

- Medialess
- Flex PLOG Size
- Byte Addressable
- Bandwidth Self-define
Storage Performance And WAF Benefits Evaluation

- **WAF/Lower is better**
  - Standard SSD
  - PLOG SSD

- **Cliff increasing**
- **Almost same WAF**

- **Latency**
  - Standard SSD: 52%
  - PLOG SSD: 44%

- **OPS**
  - Standard SSD
  - PLOG SSD
Large Scale Design For Cloud Storage Scenario Brings More Interesting Things

Cloud Storage Architecture

- Compute node 0
- Compute node 1
- Compute node n
- NIC
- NIC
- NIC
- NIC
- Storage Node 0
- Storage Node 1
- Storage Node m
- Switch

Large Scale Design

PLOG SSD

Standard SSD Optimization
Solving the Network Latency To Make Application E2E System Better

Latency

ALARM

RoCEv1/v2

QP

1 8 16 32 64 128 256 512 1K 2K 4K 8K
Reference Case Results

**E2E Latency (ms)**
- Write: 17.25%
- Read: 33.68%

*Lower Is Better*

**E2E IOPS**
- Write: 10.16%
- Read: 21.55%

*Higher Is Better*
Hardware Innovation Enables Better Application Experience

Innovations
Co-design
Multi-solutions
THANK YOU