



SDC 18

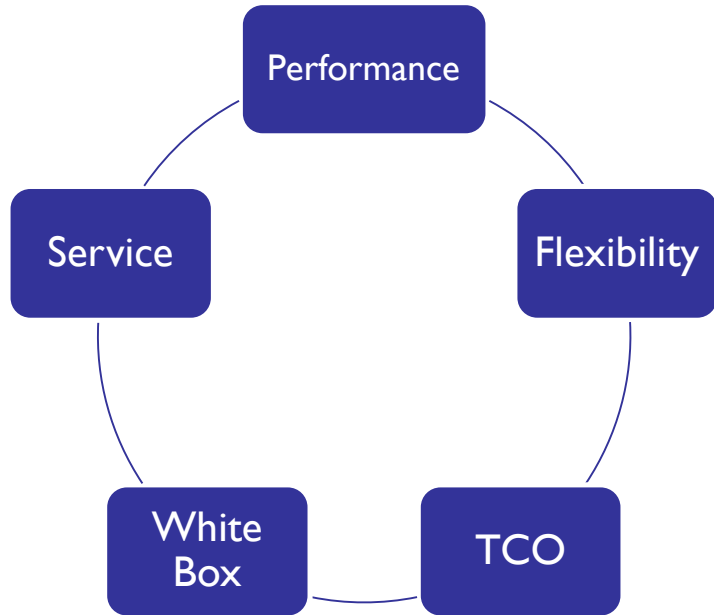
September 24-27, 2018
Santa Clara, CA

www.storagedeveloper.org

Dual-Mode SSD Architecture for Next-Generation Hyperscale Data Centers

Feng Zhu
Alibaba Group

Challenges of hyperscale data centers



- ❑ Diversified, fast-changing workloads
- ❑ Continuous pressure for TCO and supply improvement
- ❑ Demand for “white box” of I/O path - more control and determinism
- ❑ Quick response to issues and feature requests
- ❑ Demand for hardware/software co-optimization

AliFlash_Alibaba's self-developed SSDs

AliFlash V1

- High performance PCIe SSD
- Host based FTL. Control on SW/HW
- >20k devices serving in Alibaba data centers

2016

AliFlash V2

- NVMe U.2 SSD
- Customization for applications
- Cost down by 20%

2017

AliFlash V3

- Dual-mode SSD with open channel
- Deep optimization for applications
- Being productized

2018

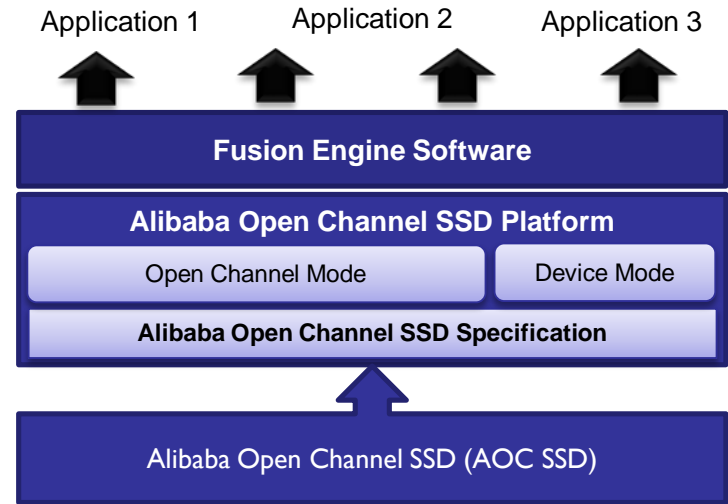
First Productionized Open Channel SSD

- ❑ Alibaba's self-developed Open Channel SSD - AliFlash V3
- ❑ Deployment ongoing in Alibaba data centers
- ❑ Major milestone since the announcement of Alibaba's Open Channel SSD Architecture in FAST'2018
- ❑ Collaborating with multiple SSD vendors to build an ecosystem



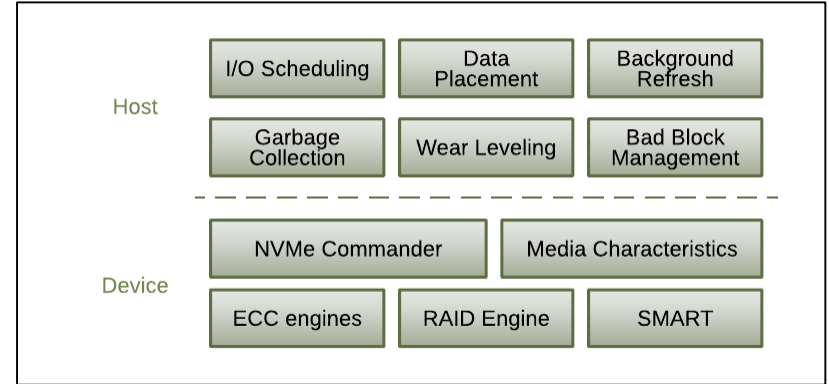
Why open channel

- ❑ New SSD architecture for next-generation data centers
- ❑ Enable hardware/software co-optimization
- ❑ Maximum flexibility for different workloads
 - Generic block device
 - Customized FTL
 - Deep integration with applications
- ❑ Reduce time & complexity of SSD qualification
- ❑ New media ready: QLC NAND, 3D X-point, SCM, ...

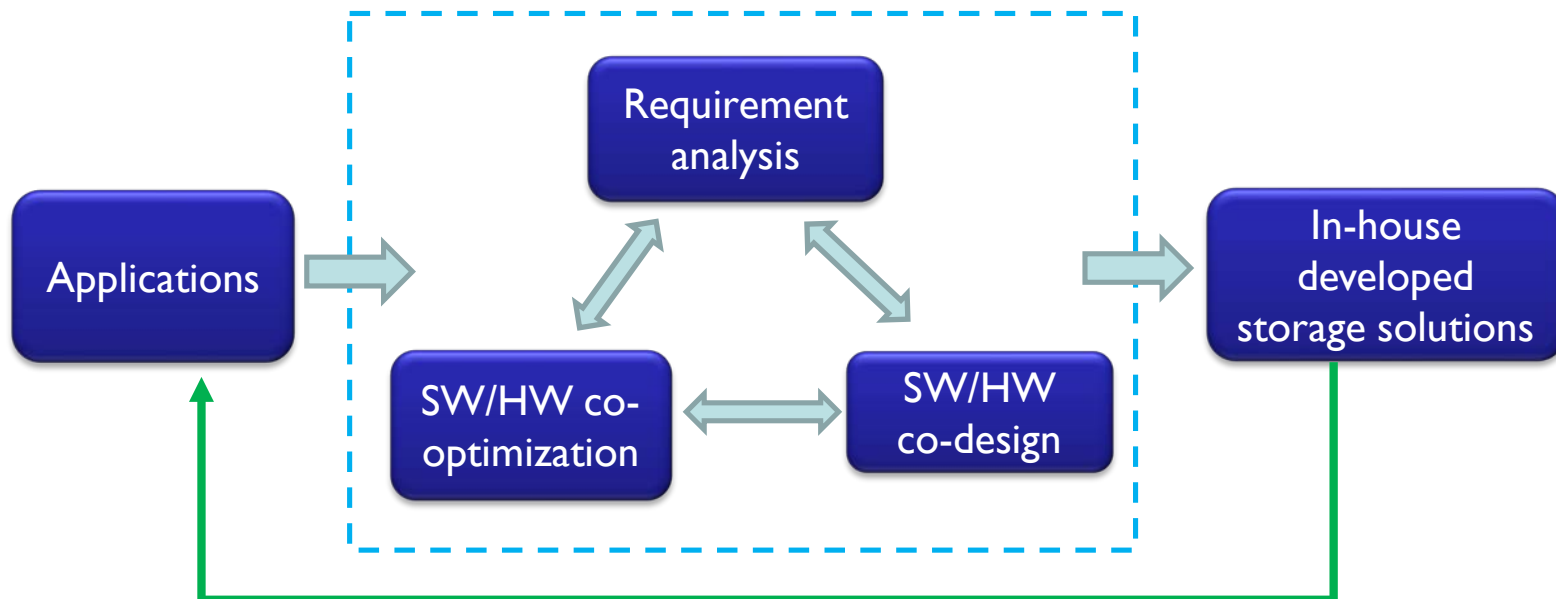


Alibaba Open Channel Architecture

- Open Channel mode
 - Direct access to the physical media, and max utilization of the media capability
 - Fully control of the data placement and I/O scheduling
 - FTL/GC customization based on application I/O pattern -> reduced WA



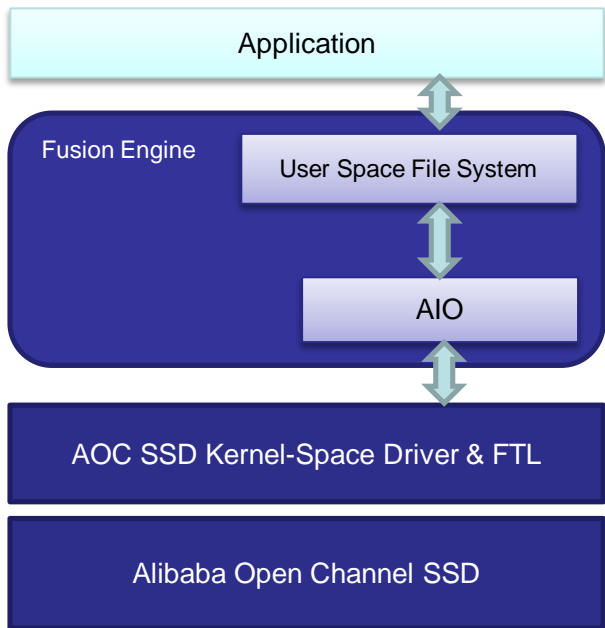
Customization for Alibaba data centers



Alibaba Open Channel SSD Software

- ❑ Host-side software stack developed by Alibaba
- ❑ Kernel-space AOC SSD Driver + Block Device FTL
- ❑ User-space AOC SSD Driver integrated with Alibaba's Fusion Engine software
- ❑ User-space customized FTL solutions
- ❑ Management, monitoring and test tools

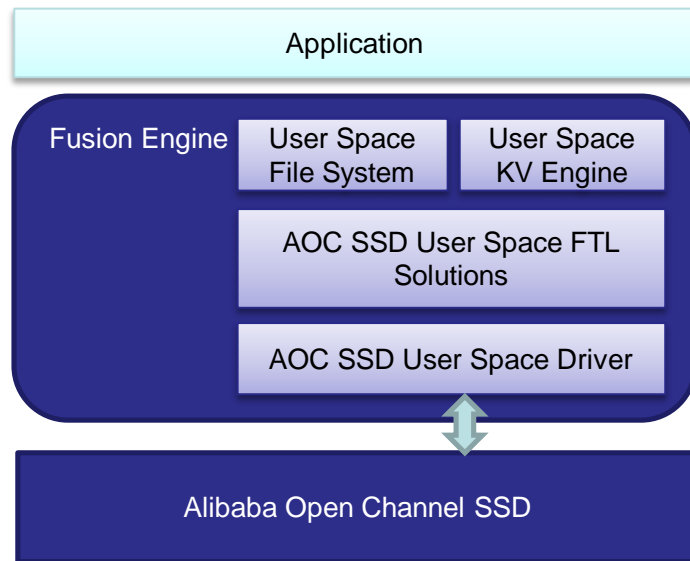
Kernel-space AOC SSD Solution



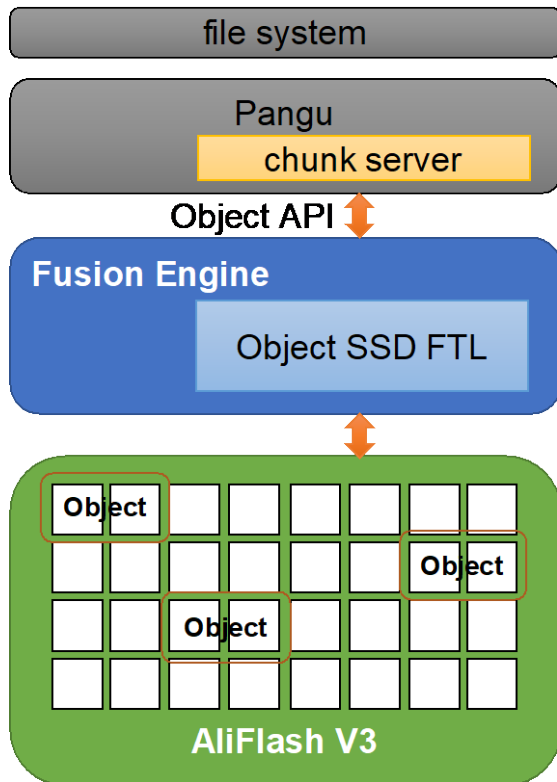
- ❑ Work like generic block device
- ❑ Covers most of the legacy use cases
- ❑ Equivalent performance and functionality as standard NVMe SSD
- ❑ Customized improvements and extensions (e.g. scheduling & QoS)

User-space AOC SSD Solution

- ❑ Entire I/O path in user space - low software overhead
- ❑ Deep integration with application - full benefits from hardware/software co-optimization
- ❑ Optimize data placement, coordinated background tasks (garbage collection)
- ❑ Customized FTL solutions (e.g. Object SSD)

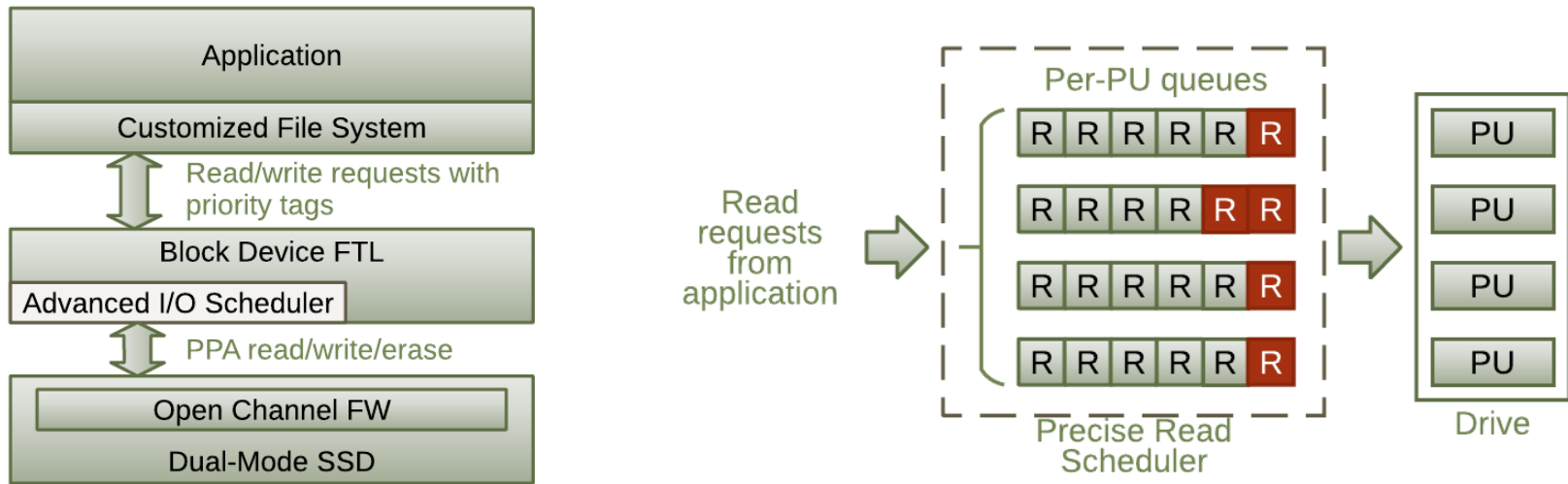


Object SSD



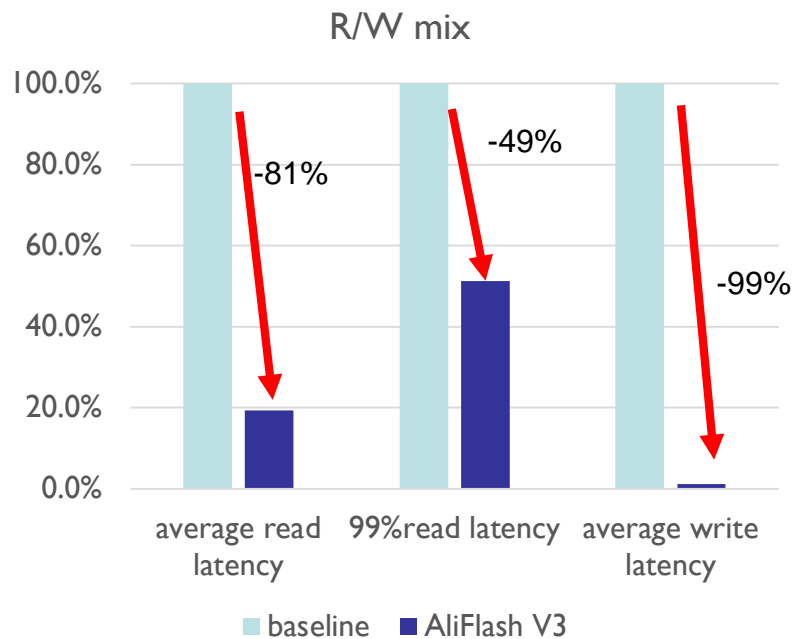
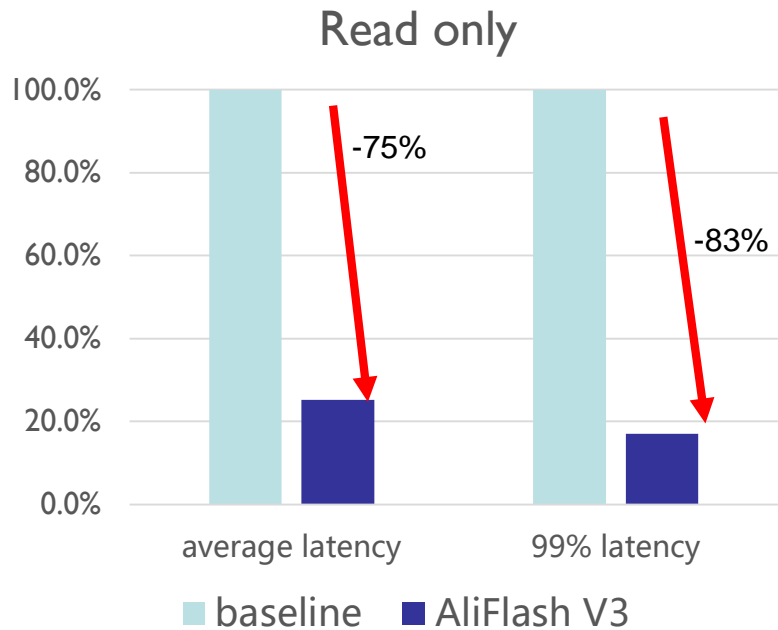
- ❑ Designed for append-only write scenario from Pangu chunk server
- ❑ Reduce WA significantly and enhance SSD lifetime
- ❑ Improve performance uniformity and QoS ~5X

QoS and I/O scheduling



- ❑ Open channel opens the door for precise I/O scheduling from host side

QoS improvements



Current Status of Alibaba Open Channel SSD

- ❑ Alibaba Open Channel SSD (AliFlash V3) is already being deployed in datacenter
- ❑ Massive deployment of AOC SSDs in 1H'2019

Alibaba Open Channel Ecosystem

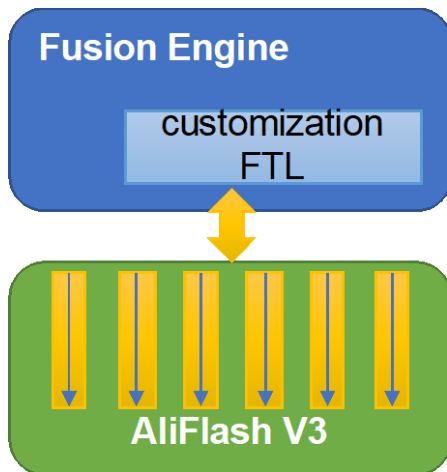


As Alibaba's **strategic partner** on Open Channel SSDs, Intel has worked with Alibaba extensively since 2017 to co-develop and co-validate this innovative solution. Alibaba's strength as a leading cloud service provider combined with Intel's strength as the leading memory and storage innovator puts us in a position to deliver the industry's 1st Open Channel SSD product.



- Alibaba is collaborating with major vendors in industry to build an ecosystem for Open Channel SSD
- Share development & debug resources
- Reduce time & complexity for SSD qualification in Alibaba
- Massive deployment in 2019

Future work



- ❑ Multi-streams
- ❑ NVMe set
- ❑ Further I/O scheduling and QoS optimization
- ❑ Low latency/high endurance solution
- ❑ QLC
- ❑ 3D X-point
- ❑ ...

Perspective

user mode

customization

HW/SW co-
design

- ❑ Alibaba is focusing on the innovation of storage solution for next generation hyperscale data center applications
- ❑ The collaboration are always welcome!

Conclusion Remarks

- ❑ Hyperscale data centers face unique challenges
- ❑ Hardware/software integrated solution is the key
- ❑ Customized designs driven by applications
- ❑ Alibaba is open to industry collaboration