



arm

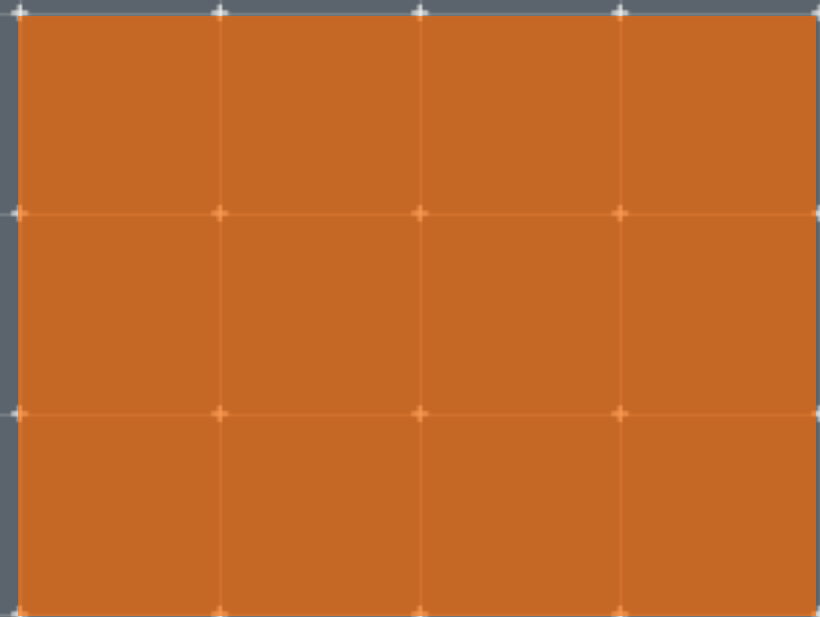
The latest storage status on arm64

Jun He, jun.he@arm.com

Agenda

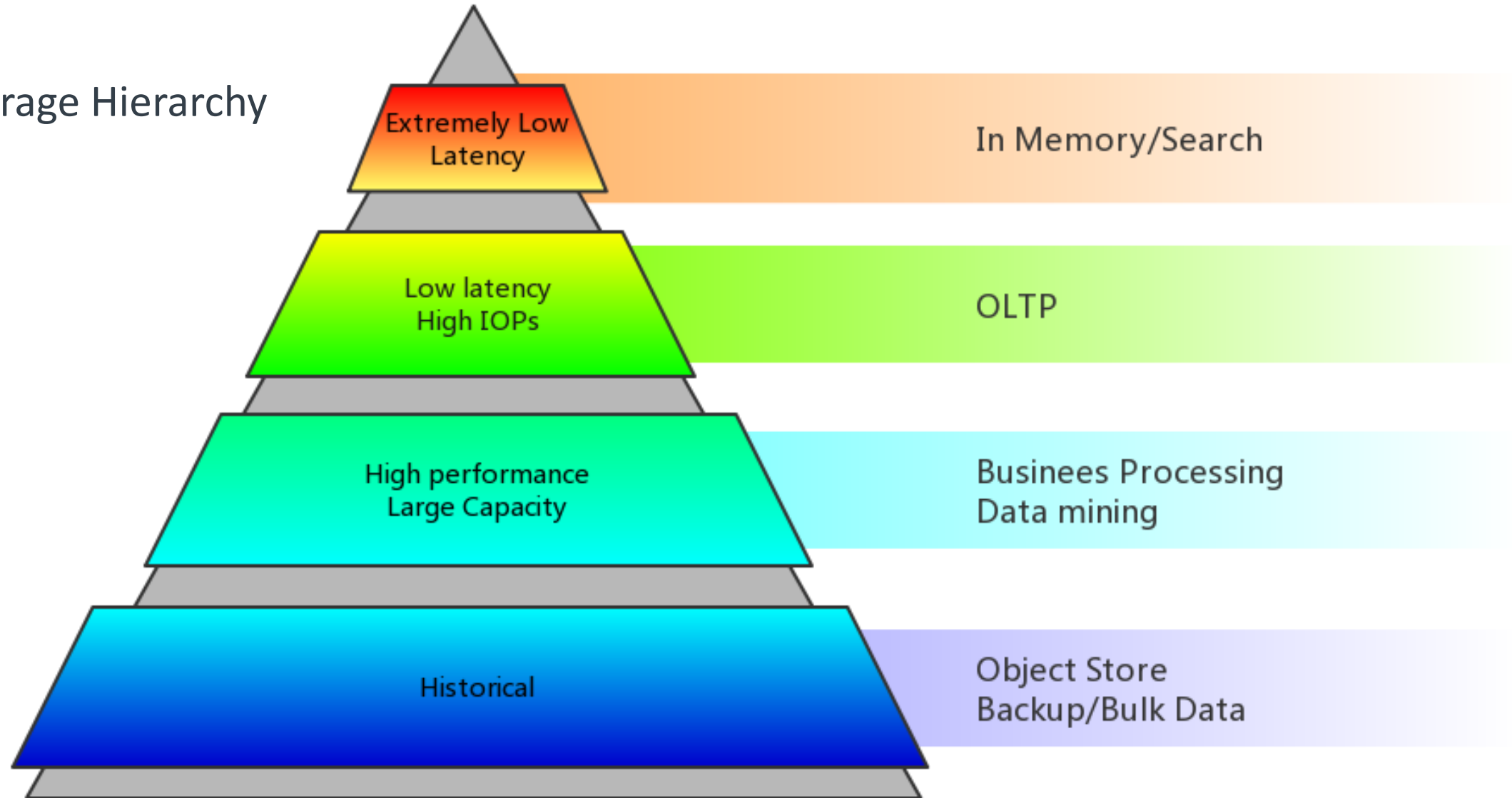
- Enterprise Storage Overview
- Arm's Fit
- Key Takeaways
- Q&A

Enterprise Storage Overview



Enterprise Storage Overview

- Storage Hierarchy



Enterprise Storage Overview

From Core-Tech to Solutions

Core-Tech

- Algorithms
 - CRC, Hash
 - RAID, Erasure Coding
 - Encryptions: AES, SMx
 - Compression
 - Bloom Filters
 - ...
- Resource Access
 - NVMe
 - Fibre Channel
 - RDMA

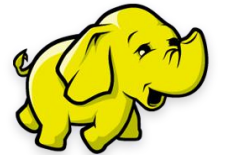
File Systems and Accelerators

- File System
 - EXT4
 - XFS
 - ZFS/OpenZFS
 - BTRFS
- Accelerations
 - SPDK
 - DPDK
 - NVMe-oF

Solutions



ceph



MINIO



ROOK



CINDER

an OpenStack Community Project



SWIFT

an OpenStack Community Project



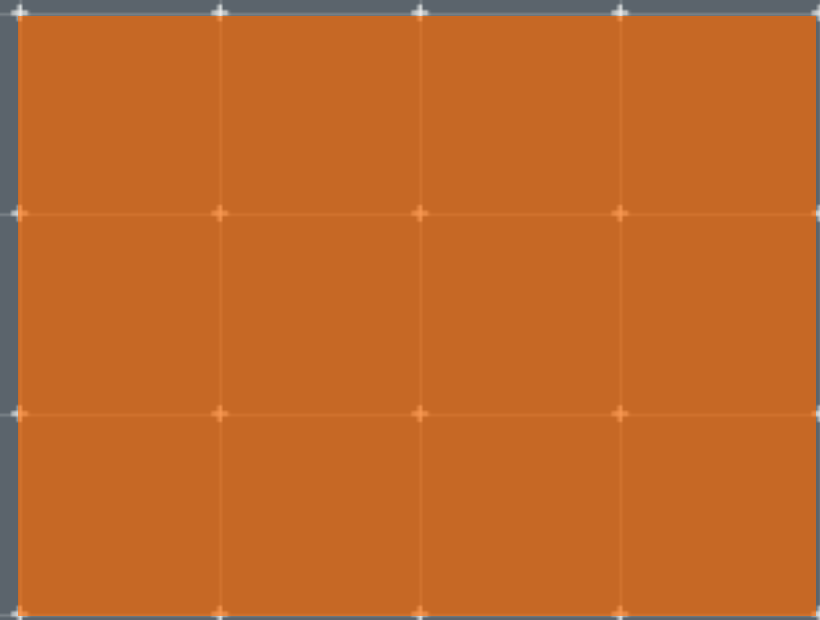
arm

Enterprise Storage Overview

Trend

- With SSDs being everywhere NVMe is becoming more popular
- Compute resources are separated from storage resources to get right sizing and independent scaling
- Software defined storage is important for system deployment, particularly for general databases and backup

Arm's Fit



Core Tech

Algorithms

- Fundamental algorithms are optimized with Neon and specific extensions
- CRC/SHA/AES optimizations have been contributed to various popular open source projects
- A complete reference implementations set will be contributed to ISA-L
 - RAID has been done and merged
 - CRC and multi-buffer Hash are in progress
 - AES will be next

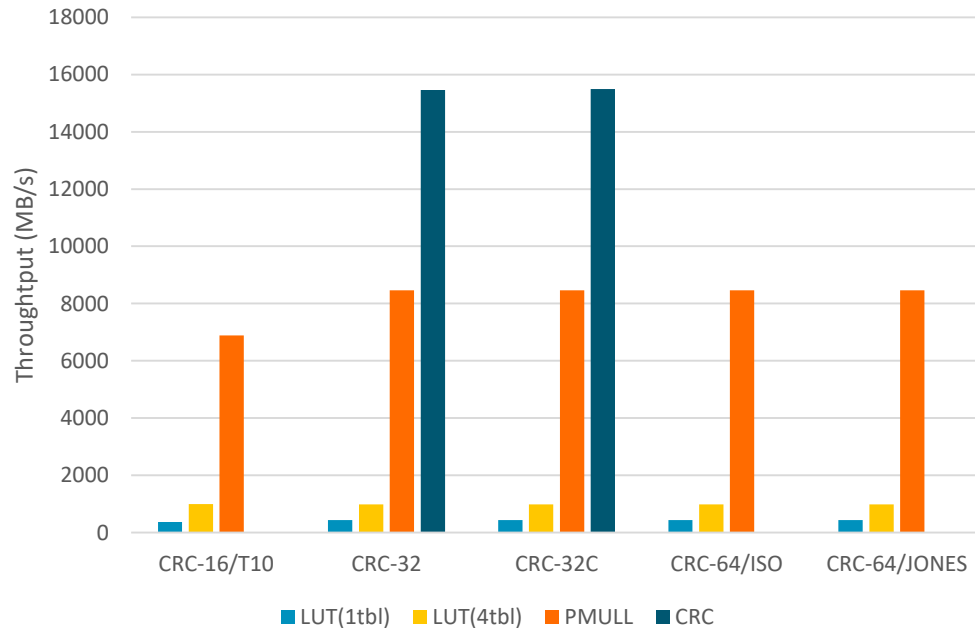
Resource Access

- NVMe
 - Quite a few NVMe SSDs from different vendors have been validated on Arm
- RDMA
 - Validated Mellanox ConnectX series on Arm
 - 4KB kernel page size
 - 64KB kernel page size

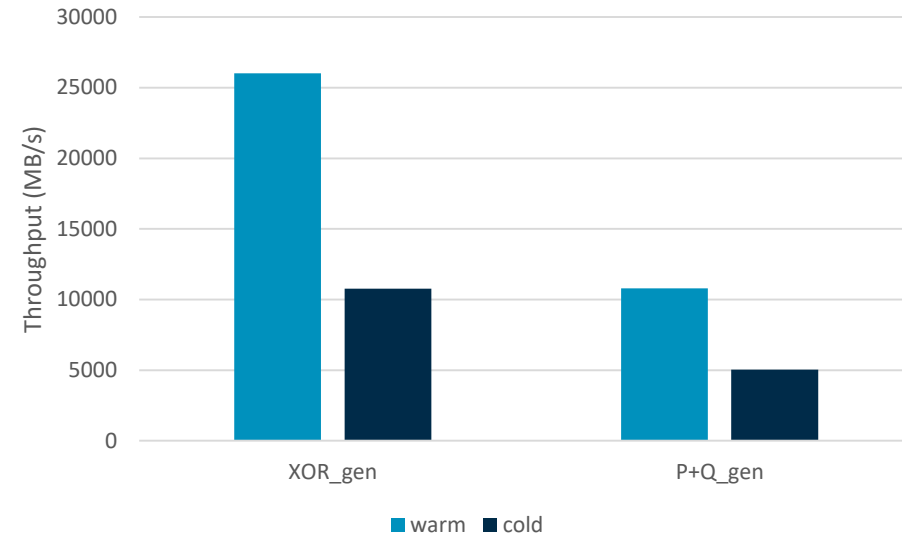
Core Tech

Algorithms

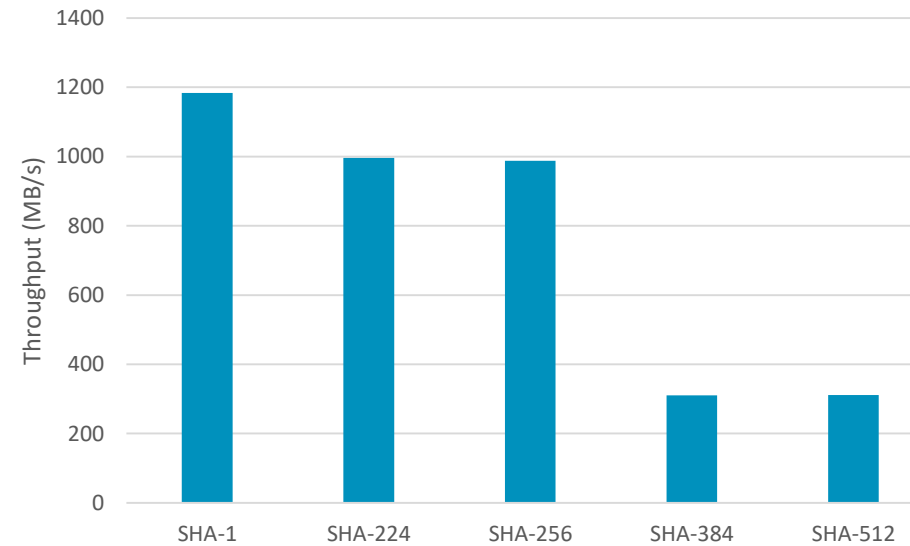
- Benchmarks on Arm64



CRC-16/32/64 Benchmark



RAID-5/6 Benchmark



SHA Benchmark

File Systems and Accelerators

File Systems

- EXT4
- XFS
- ZFS
- Btrfs

Accelerations

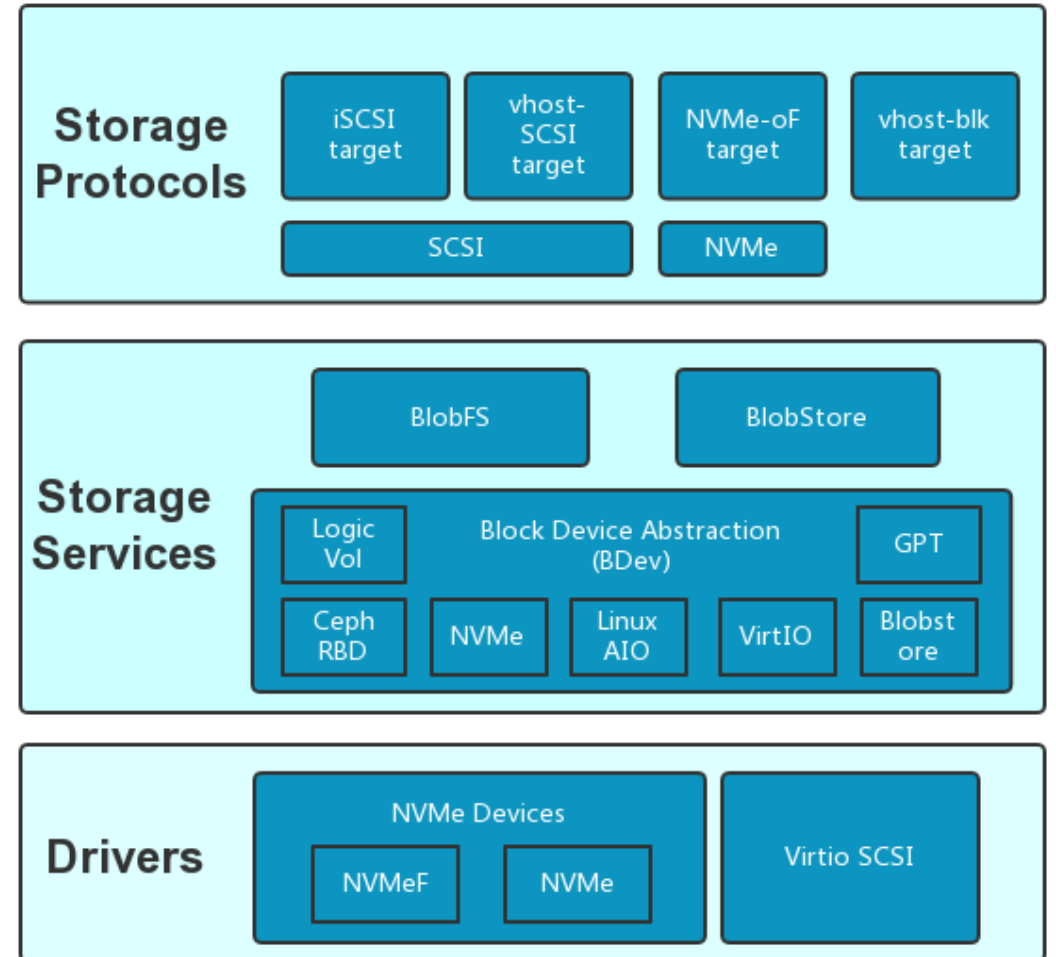
- SPDK
 - Enabled and tested on Arm.
 - Memory barrier
 - VA address space
 - 4KB + 64KB kernel page size support
 - Fixed several UT failures
 - Optimized CRC-32/32C using CRC extension. Significant performance improvement is observed in NVMe-oF/TCP
- DPDK
 - Added 64KB kernel page size support to pci_vfio
 - Updated IOMMU configuration setup for Arm64
- NVMe-oF

File Systems and Accelerators

SPDK

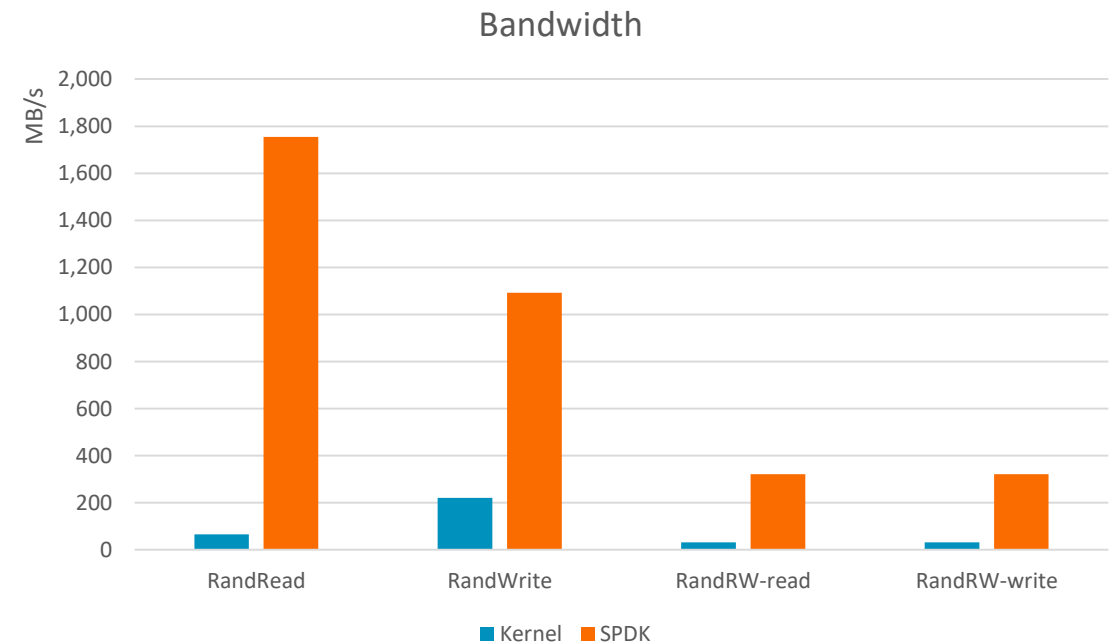
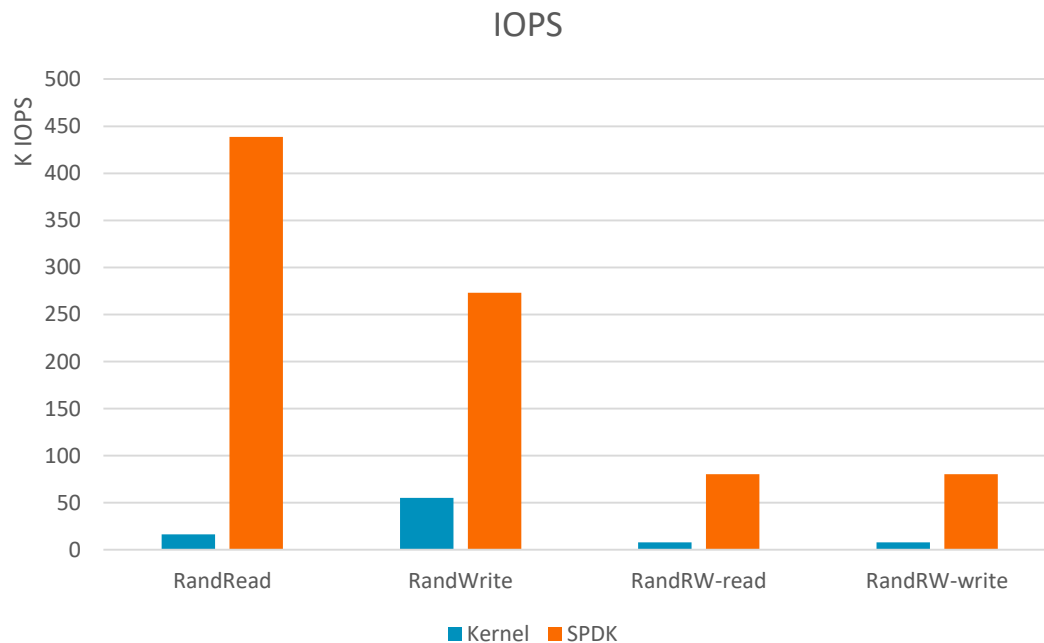
What is it

- Storage Performance Development Kit
- A set of tools and libraries to create high performance, scalable, user mode storage applications
- Designed for new storage HW devices (NVMe). Can achieve millions of IOPS per core. Better tail latency.



File System and Accelerators

- SPDK Benchmark



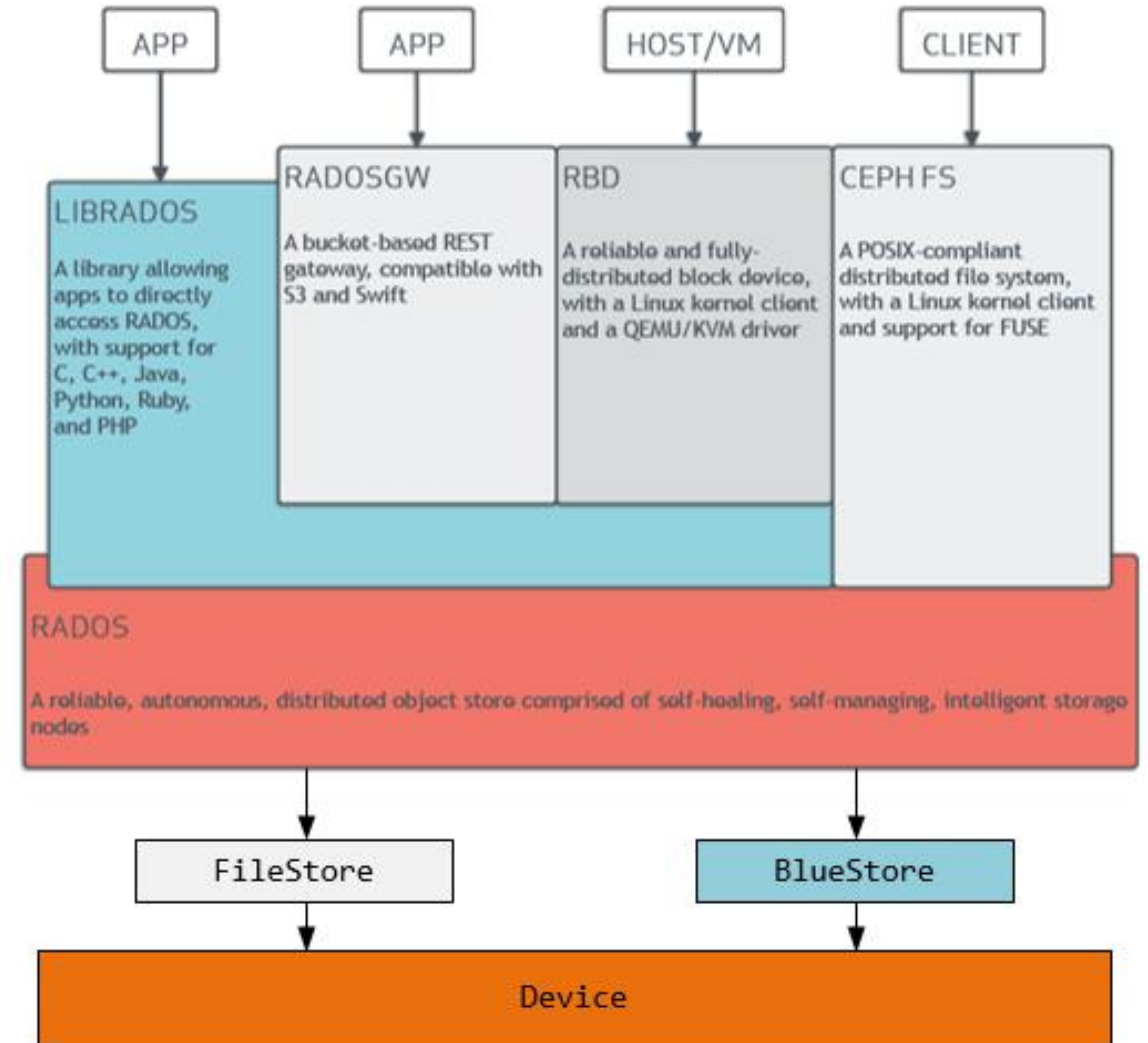
System configuration: 2.5GHz AArch64 multi-core, 96GB DDR4 Memory, 1NVMe
FIO configuration: direct=1, bs=4096, rwmixread=50, iodepth=32, ramp=30s, run_time=180s, jobs=1

Solutions

Ceph

Quick Primer

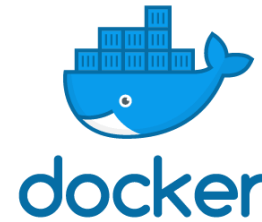
- Open source, object-based distributed storage system
- Offers three kinds of services
 - Object storage
 - Block storage
 - File system
- Highly durable, available
- Popular in Cloud, HPC and BigData domains
 - Dominate in Cinder drivers



Solutions

Ceph

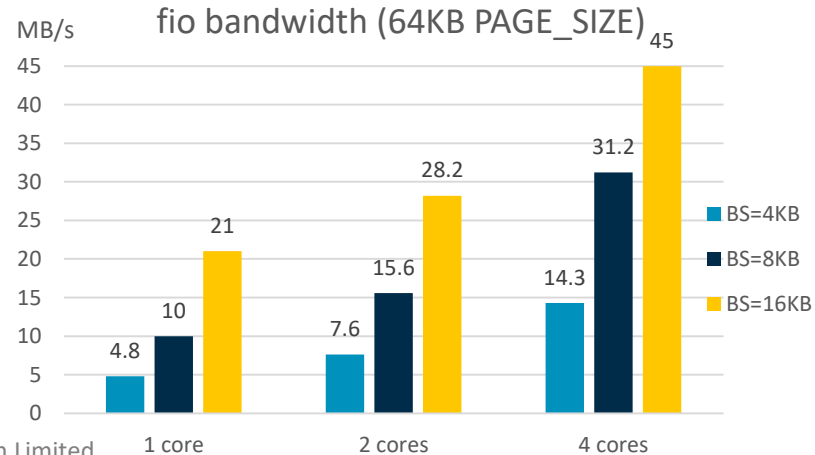
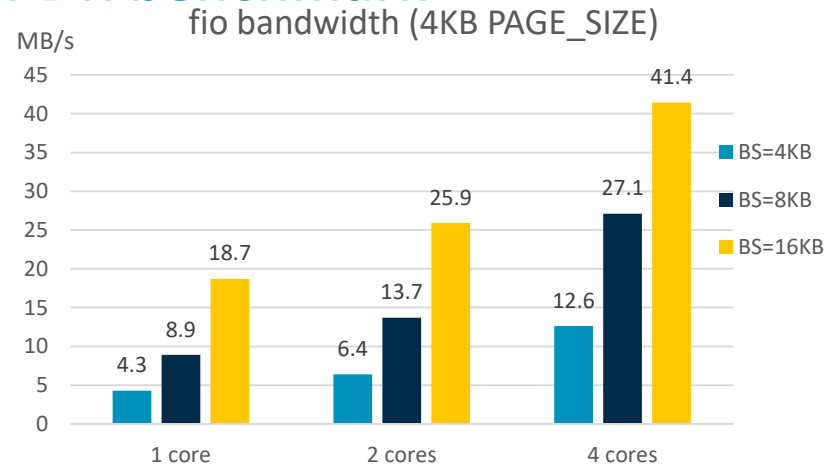
- Arm64 packages are available in main distros
- Supported in container world
- Bugfixes, features and improvements
 - CRC32 optimizations with Arm's extension
 - 64KB kernel page size support in NVMeDevices
 - NVMeDevice crash
 - NVMeManager thread hang
 - Tested Ceph + SPDK with 4KB and 64KB kernel page size
 - Validated Ceph + RDMA with 4KB and 64KB kernel page size, full coverage test is in progress



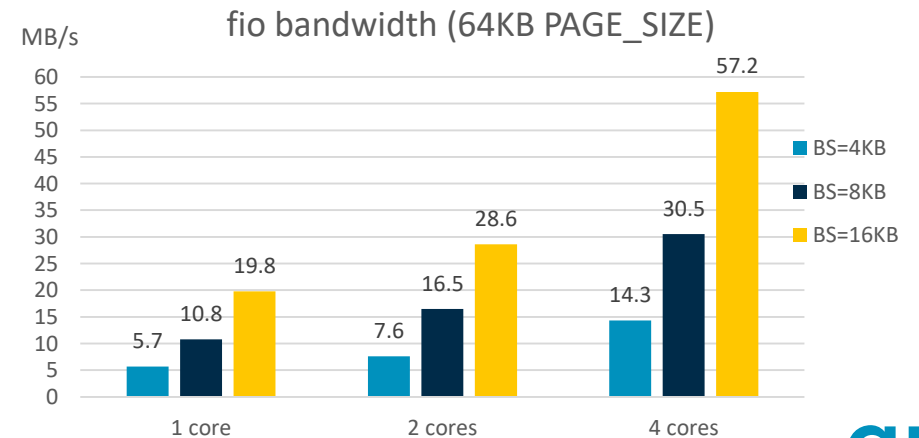
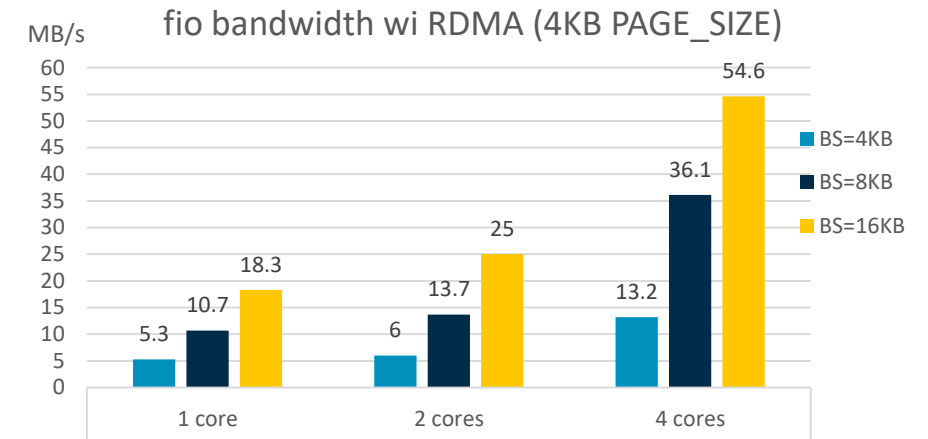
Solutions

Ceph

Ceph+SPDK benchmark



Ceph+RDMA benchmark



Solutions

Ceph

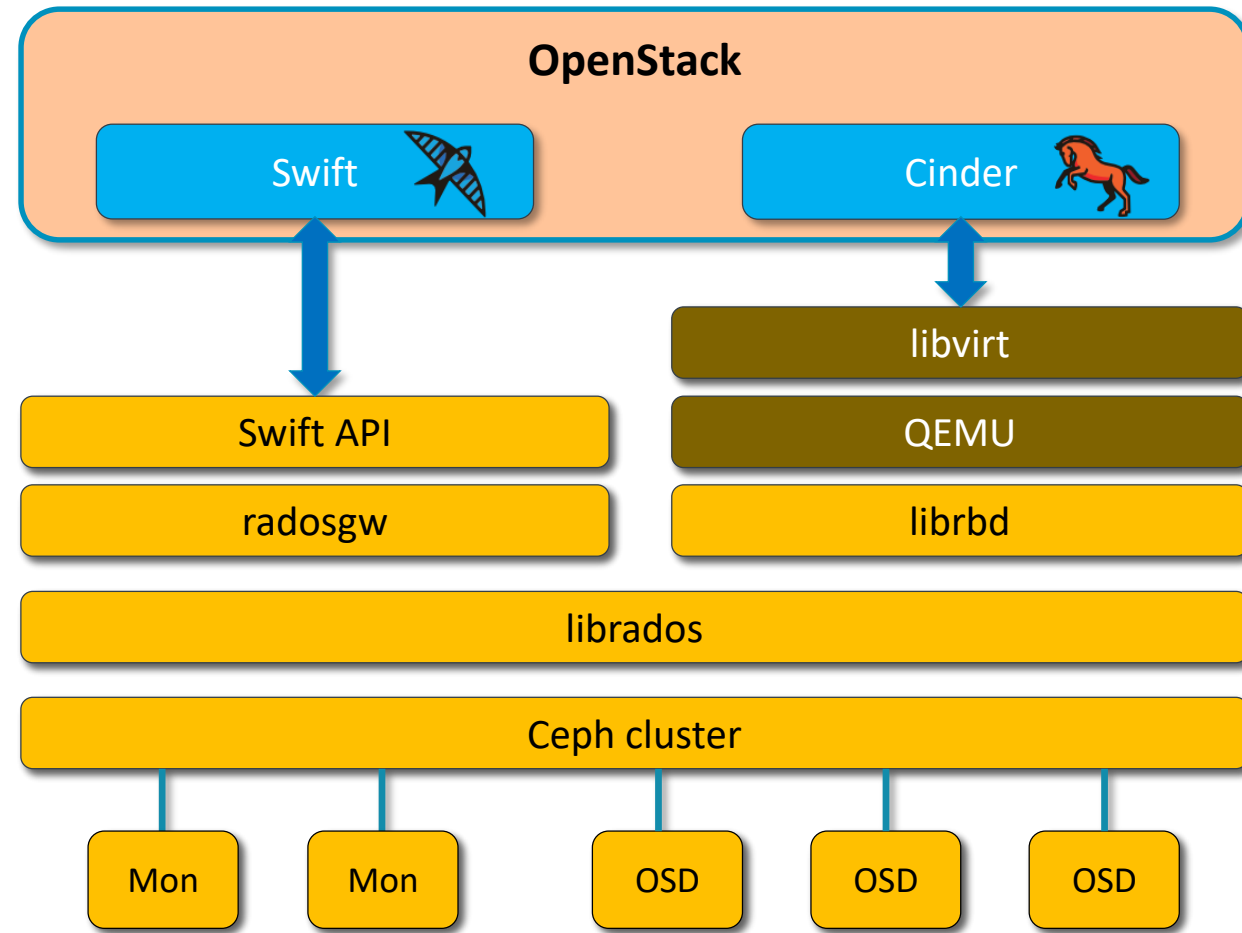
Take the Next Step

- Ceph + RDMA performance optimization
- Ceph + NVMe-oF
 - With RDMA underlying
 - Enablement with 4KB and 64KB kernel page size support
 - Performance profiling and optimization
- Ceph OSD migration to Seastar follow up
(<https://github.com/ceph/ceph/tree/master/src/crimson>)

Solutions

OpenStack

- Block storage: Ceph RBD as Cinder backend
- Object storage: Swift compatible RADOS gateway
- 100% pass rate on interoperability tests (2018.02 guidelines)
- Moving to Kolla
 - Added Ceph bluestore OSD in Kolla
 - Blueprints, improvements and CI jobs

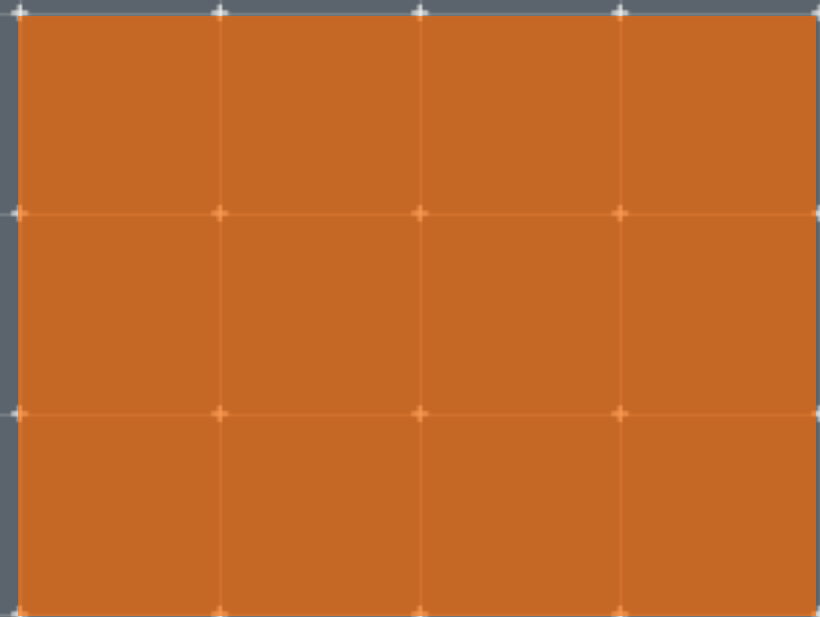


Solutions

Misc

Lustre	GlusterFS	HDFS	MiniO	ROOK
<ul style="list-style-type: none">Validated with ZFS backend and LDISKFS backendAuster Tests	<ul style="list-style-type: none">Built, deployed and unit tested on Arm64Benchmarked with gbench, fio and iozone	<ul style="list-style-type: none">Validated with BigData software stack on Arm64	<ul style="list-style-type: none">Built and benchmarked on Arm64 with both BareMetal and docker containers.While it has highway hash and SHA-256 are optimized for Arm64, CRC32 is not there	<ul style="list-style-type: none">Enabled native build on Arm64Ceph features and improvements

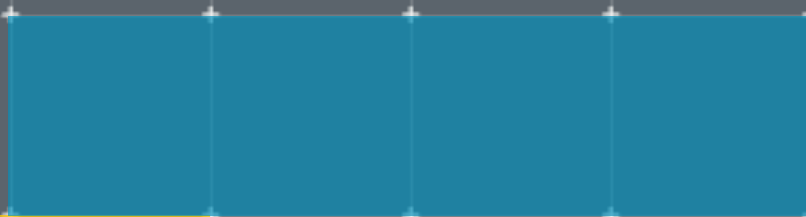
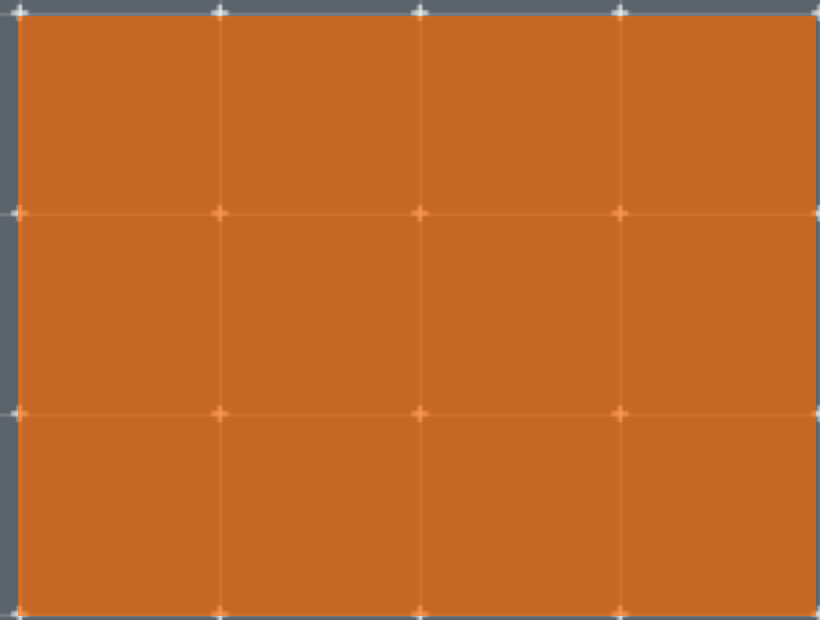
Key Takeaways



Key Takeaways

- Arm support has been widely adopted in various perspective of storage tiers.
- Optimizations for core tech are important for storage performance
- New technologies and use cases bring new requirements for storage where Arm can be a good fit

Q&A



Thank You

Danke

Merci

谢谢

ありがとう

Gracias

Kiitos

감사합니다

धन्यवाद

תודה

arm