



SPDK: STATE OF THE PROJECT

Jim Harris

Principal Software Engineer

Data Center Group



By Edmund Garman from Salem, Oregon, USA - Waikiki, Honolulu Panorama
CC BY 2.0, <https://commons.wikimedia.org/w/index.php?curid=39574339>

AGENDA

- Community Growth
- Feature Growth
- Looking Forward



- **Community Growth**
- Feature Growth
- Looking Forward



SPDK PATCH COMMIT STATISTICS

Version Range	Total Commits	Non-Intel Committers	Non-Intel Commit %
v16.06 .. v17.07	2221	22	2.7%
v17.07 .. v18.07	3152	37	9.0%
v18.07 .. v19.07	3524	58	16.8%

SPDK PATCH COMMIT STATISTICS

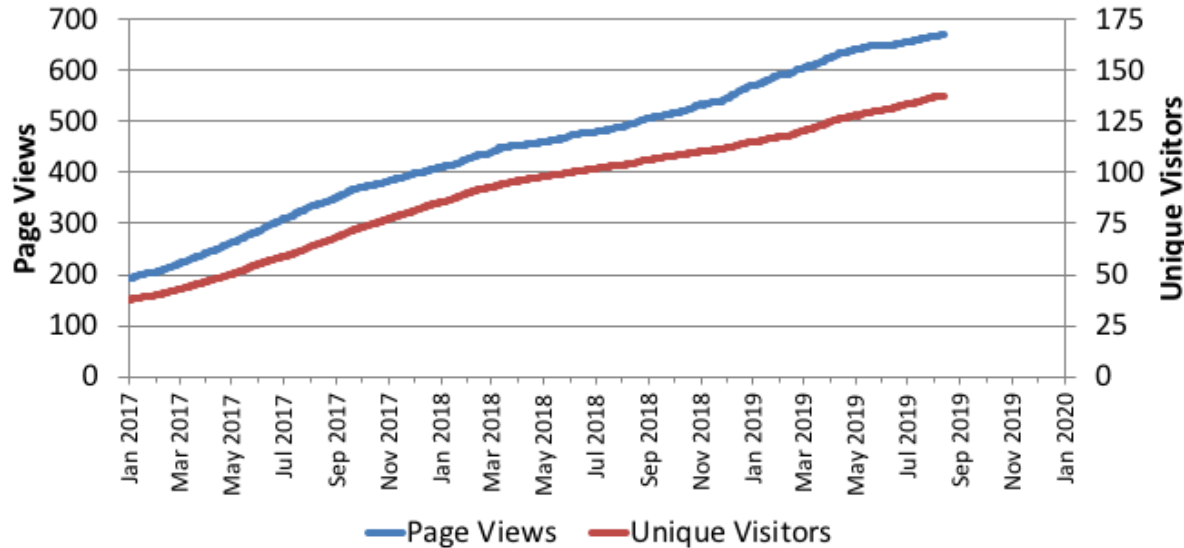
Version Range	Total Commits	Non-Intel Committers	Non-Intel Commit %
v16.06 .. v17.07	2221	22	2.7%
v17.07 .. v18.07	3152	37	9.0%
v18.07 .. v19.07	3524	58	16.8%

Last 12 months
+57% Non-Intel Committers
2x Non-Intel Commits

Last 3 Years
Over 1/3 of Non-Intel
Committers from PRC!

COMMUNITY GROWTH

SPDK Github Traffic 52 Week Running Daily Averages



+33% YoY
+247% since Jan'17

- Community Growth
- **Feature Growth**
- Looking Forward



RECENTLY ADDED FEATURES

Storage Protocols

NVMe-oF*
Target

RDMA

vhost-nvme
Target

iSCSI
Target

vhost-scsi
Target

vhost-blk
Target

Linux nbd

NVMe

SCSI

Storage Services

Block Device Abstraction (bdev)

QoS

3rd Party

Logical
Volumes

BlobFS

Blobstore

NVMe

Linux
AIO

Ceph
RBD

PMDK
blk

virtio
(scsi/blk)

iSCSI

Drivers

NVMe Devices

Fabrics
Initiator

RDMA

PCIe

virtio

virtio-
PCIe

vhost-
user

Intel® QuickData
Technology
Driver

Integration

VPP TCP/IP

RocksDB

Ceph

QEMU

Tools

fio

nvme-cli

RECENTLY ADDED FEATURES

Storage Protocols

NVMe-oF*
Target

RDMA
TCP

vhost-nvme
Target

iSCSI
Target

vhost-scsi
Target

vhost-blk
Target

Linux nbd

NVMe

SCSI

Storage Services

Block Device Abstraction (bdev)

QoS

3rd Party

Logical
Volumes

BlobFS

Blobstore

NVMe

Linux
AIO

Ceph
RBD

PMDK
blk

virtio
(scsi/blk)

iSCSI

Drivers

NVMe Devices

Fabrics
Initiator

RDMA
TCP

PCIe

virtio

virtio-
PCIe

vhost-
user

Intel® QuickData
Technology
Driver

Integration

VPP TCP/IP

RocksDB

Ceph

QEMU

Tools

fio

nvme-cli

NVME/TCP



NVMe TP ratified November 2018

SPDK added TCP transport for

- NVMe driver
- NVMe-oF target

Supports alternative TCP stack implementations

RECENTLY ADDED FEATURES

Storage Protocols

NVMe-oF*
Target

RDMA

TCP

vhost-nvme
Target

iSCSI
Target

vhost-scsi
Target

vhost-blk
Target

Linux nbd

NVMe

SCSI

Storage Services

Block Device Abstraction (bdev)

QoS

3rd PartyLogical
Volume

Crypto

BlobFS

Blobstore

NVMe

Linux
AIOCeph
RBDPMDK
blkvirtio
(scsi/blk)

iSCSI

Drivers

NVMe Devices

Fabrics
Initiator

RDMA

TCP

PCIe

virtio

virtio-
PCIe

vhost-
user

Intel® QuickData
Technology
Driver

Integration

VPP TCP/IP

RocksDB

Ceph

QEMU

Tools

fio

nvme-cli

CRYPTO

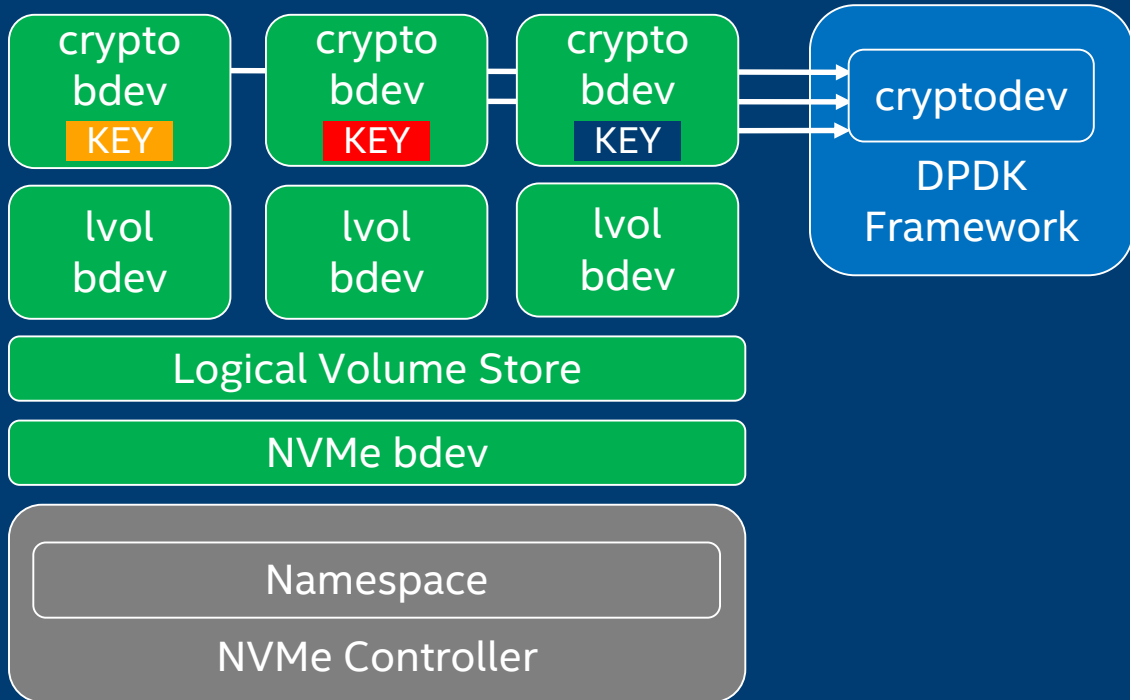


Hardware Accelerators

- Intel® QuickAssist Accelerator

Optimized Software

- ISA-L



RECENTLY ADDED FEATURES

Storage Protocols

NVMe-oF*
Target

RDMA

TCP

vhost-nvme
Target

iSCSI
Target

vhost-scsi
Target

vhost-blk
Target

Linux nbd

NVMe

SCSI

Storage Services

Block Device Abstraction (bdev)

QoS

3rd Party

Logical
Volumes

Crypto

BlobFS

Blobstore

NVMe

Linux
AIO

Ceph
RBD

PMDK
blk

virtio
(scsi/blk)

iSCSI

Cinder

RocksDB

Ceph

QEMU

Tools

fio

nvme-cli

Drivers

NVMe Devices

Fabrics
Initiator

RDMA

TCP

PCIe

virtio

virtio-
PCIe

vhost-
user

Intel® QuickData
Technology
Driver

CINDER

Provision ephemeral storage using SPDK

- NVMe-oF target
- Logical volumes
- Dynamic configuration via JSON-RPC

Kubernetes integration in progress

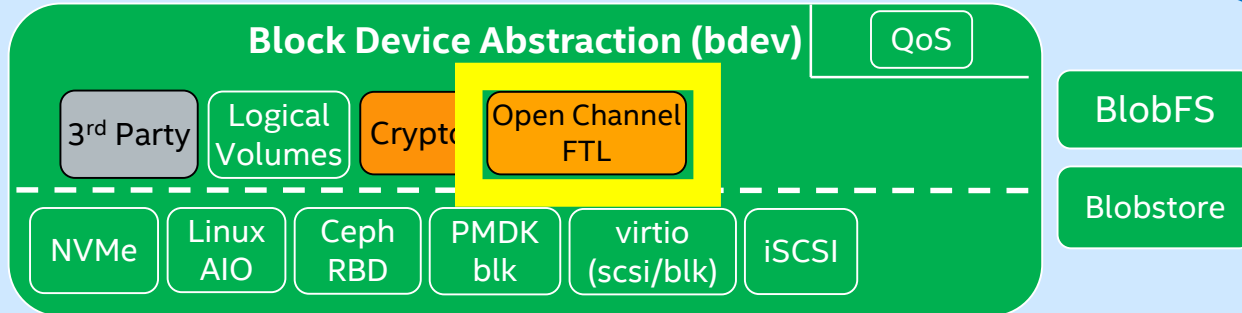


RECENTLY ADDED FEATURES

Storage Protocols



Storage Services



Drivers



Integration

Cinder

VPP TCP/IP

RocksDB

Ceph

QEMU

Tools

fio

nvme-cli

OPEN CHANNEL BLOCK FTL

Host FTL enabling smart data placement

- Based on OC2.0 specification

Block FTL support added to bdev nvme module

RECENTLY ADDED FEATURES

Storage Protocols

NVMe-oF*
Target

RDMA
TCP

vhost-nvme
Target

iSCSI
Target

vhost-scsi
Target

vhost-blk
Target

Linux nbd

NVMe

SCSI

Storage Services

Block Device Abstraction (bdev)

QoS

3rd Party

Logical
Volumes

Crypto

Open Channel
FTL

OCF
Cache

BlobFS

Blobstore

NVMe

Linux
AIO

Ceph
RBD

PMDK
blk

virtio
(scsi/blk)

iSCSI

Integration

Cinder

VPP TCP/IP

RocksDB

Ceph

QEMU

OCF

Tools

fio

nvme-cli

Drivers

NVMe Devices

Fabrics
Initiator

RDMA
TCP

PCIe

virtio

virtio-
PCIe

vhost-
user

Intel® QuickData
Technology
Driver

OPEN CAS FRAMEWORK

Block Caching Library

Open Sourced in Q4'18

SPDK integration added v19.01

RECENTLY ADDED FEATURES

Storage Protocols

NVMe-oF*
Target

RDMA
TCP

vhost-nvme
Target

iSCSI
Target

vhost-scsi
Target

vhost-blk
Target

Linux nbd

NVMe

SCSI

Storage Services

Block Device Abstraction (bdev)

QoS

3rd Party

Logical
Volumes

Crypto

Open Channel
FTL

OCF
Cache

BlobFS

Blobstore

NVMe

Linux
AIO

Ceph
RBD

PMDK
blk

virtio
(scsi/blk)

iSCSI

Drivers

NVMe Devices

Fabrics
Initiator

RDMA
TCP

PCIe

virtio

virtio-
PCIe

vhost-
user

Intel® QuickData
Technology
Driver

Integration

Cinder

VPP TCP/IP

RocksDB

Ceph

QEMU

OCF

Tools

fio

spdk-cli

spdk-cli

SPDK-CLI

```
0- / ..... [..]
  0- bdevs ..... [..]
    | 0- aio ..... [Bdevs: 0]
    | 0- error ..... [Bdevs: 0]
    | 0- iscsi ..... [Bdevs: 0]
    | 0- logical_volume ..... [Bdevs: 0]
    | 0- malloc ..... [Bdevs: 0]
    | 0- null ..... [Bdevs: 0]
    | 0- nvme ..... [Bdevs: 0]
    | 0- pmemblk ..... [Bdevs: 0]
    | 0- rbd ..... [Bdevs: 0]
    | 0- split_disk ..... [Bdevs: 0]
    | 0- virtioblk_disk ..... [Bdevs: 0]
    | 0- virtioscsi_disk ..... [Bdevs: 0]
  0- lvol_stores ..... [Lvol stores: 0]
  0- nvme
    0- subsystem ..... [Subsystems: 1]
      | 0- nqn.2014-08.org.nvmeexpress.discovery .... [st=Discovery, Allow any host]
      |   0- hosts ..... [Hosts: 0]
      |   0- listen_addresses ..... [Addresses: 0]
    0- transport ..... [Transports: 0]
/>
```

SPDK ARCHITECTURE

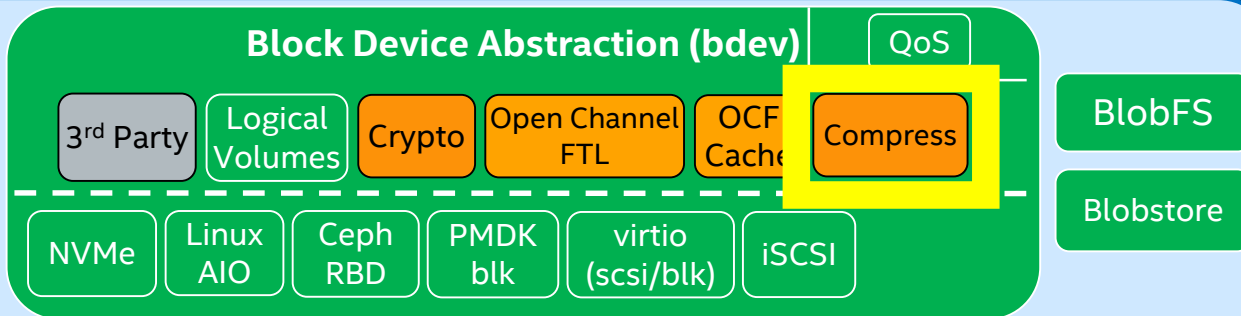
SPDK 18.07

Added since 18.07

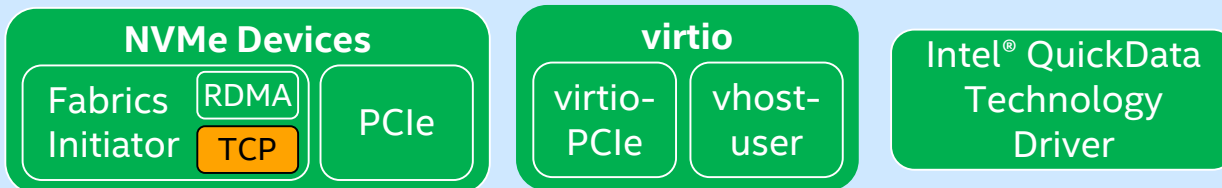
Storage Protocols



Storage Services



Drivers



Integration

Cinder

VPP TCP/IP

RocksDB

Ceph

QEMU

OCF

Tools

fio

nvme-cli

spdk-cli

BLOCK COMPRESSION



DPDK Framework for compression algorithms

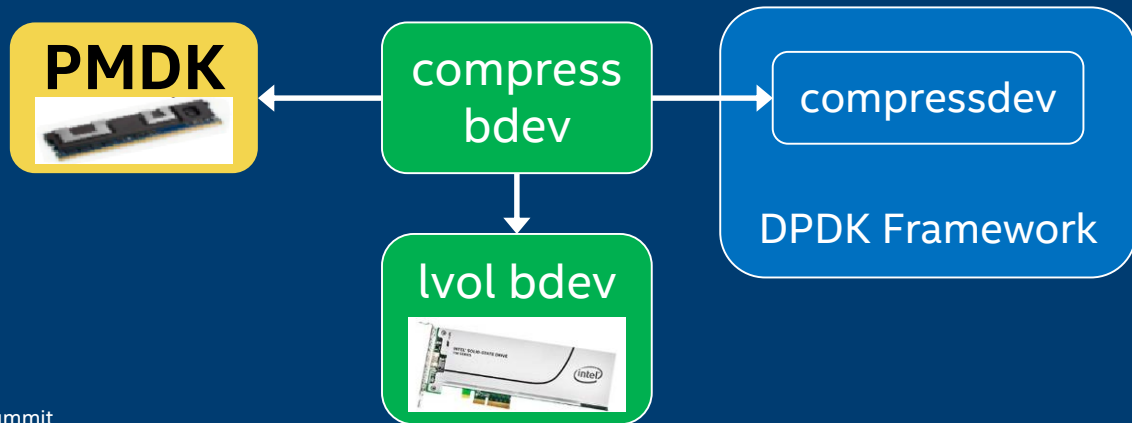
- Accelerators or Optimized SW

PMDK for persistent memory access

- For compression metadata

SPDK Logical Volumes for storage

- For realizing capacity savings



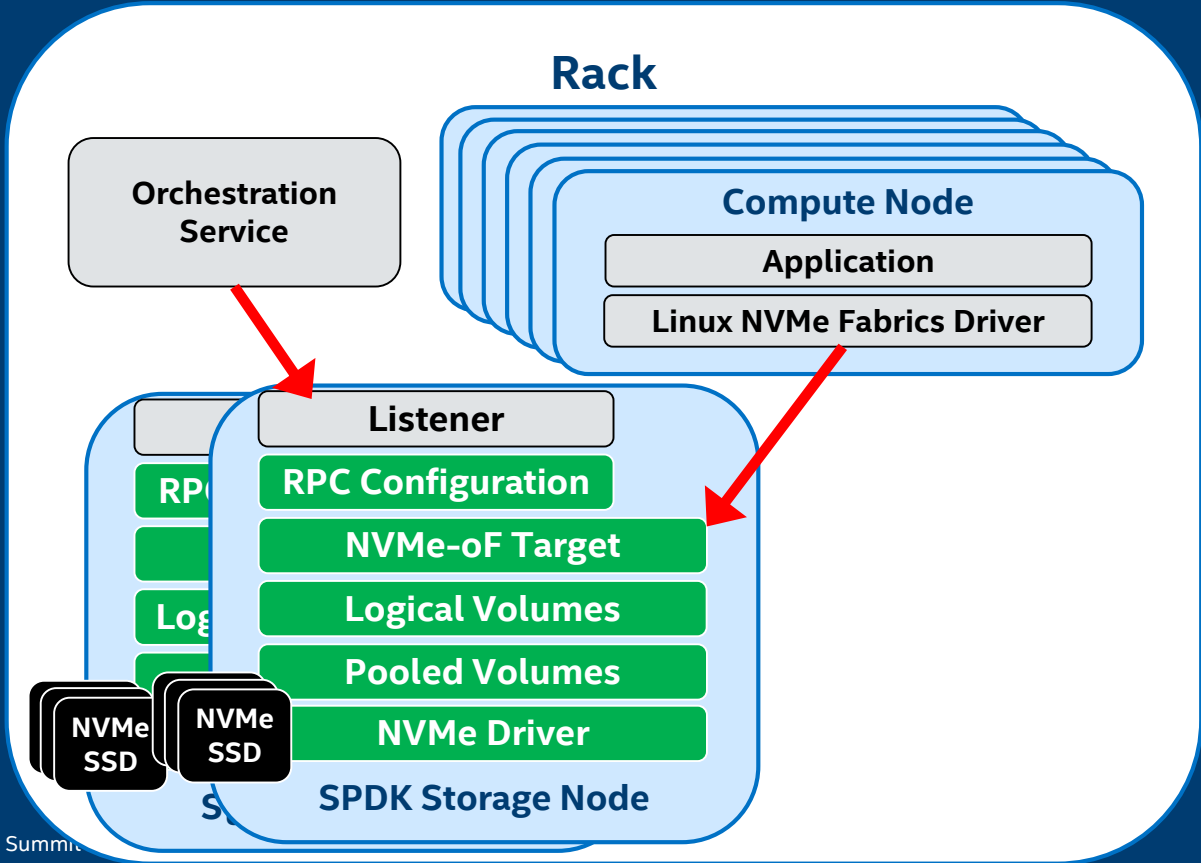
- Community Growth
- Feature Growth
- **Looking Forward**



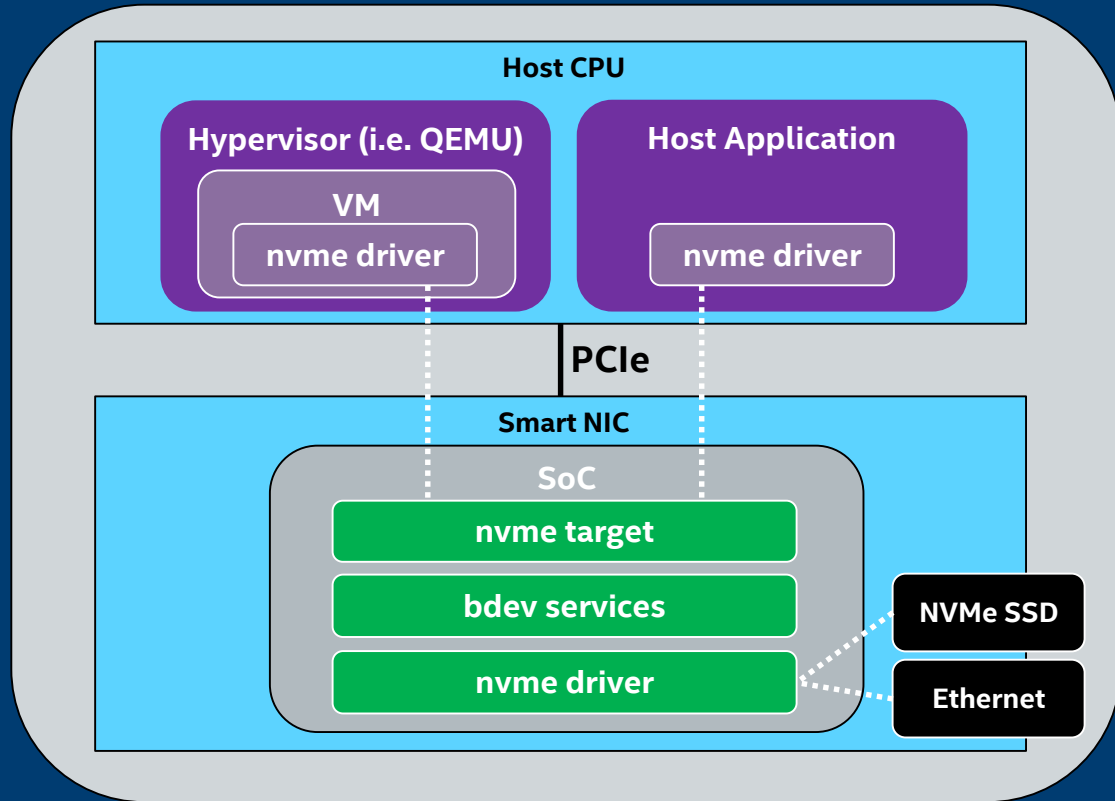
RACK SCALE ARCHITECTURE

SPDK Software

3rd Party Software



SMART NICs

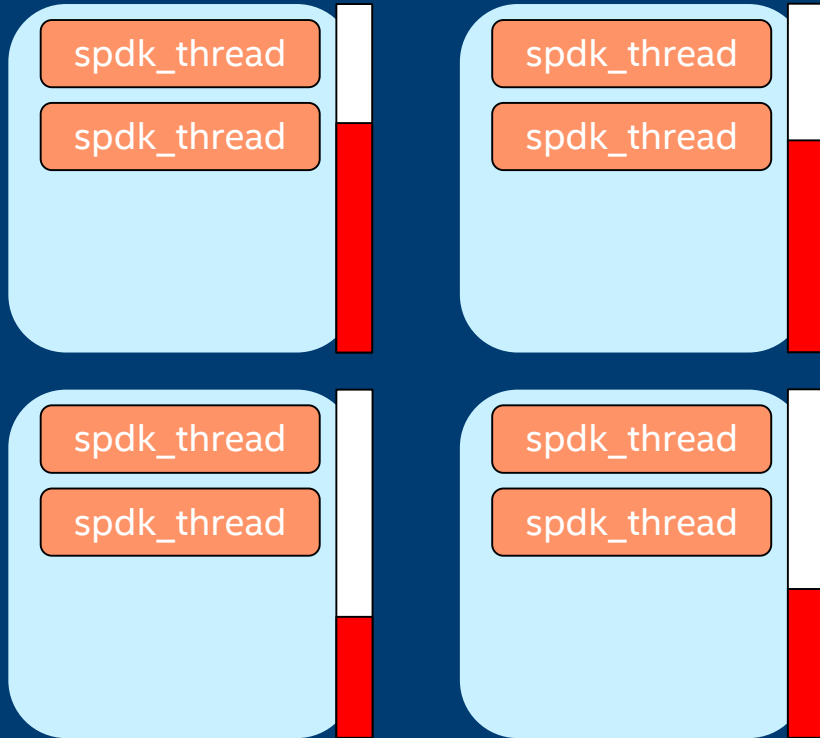


Present NVMe Interface to Host

Use SPDK to emulate NVMe

- Block services (crypto, logical volumes)
- Forward across network
- Or “local” SSD

DYNAMIC SCHEDULING

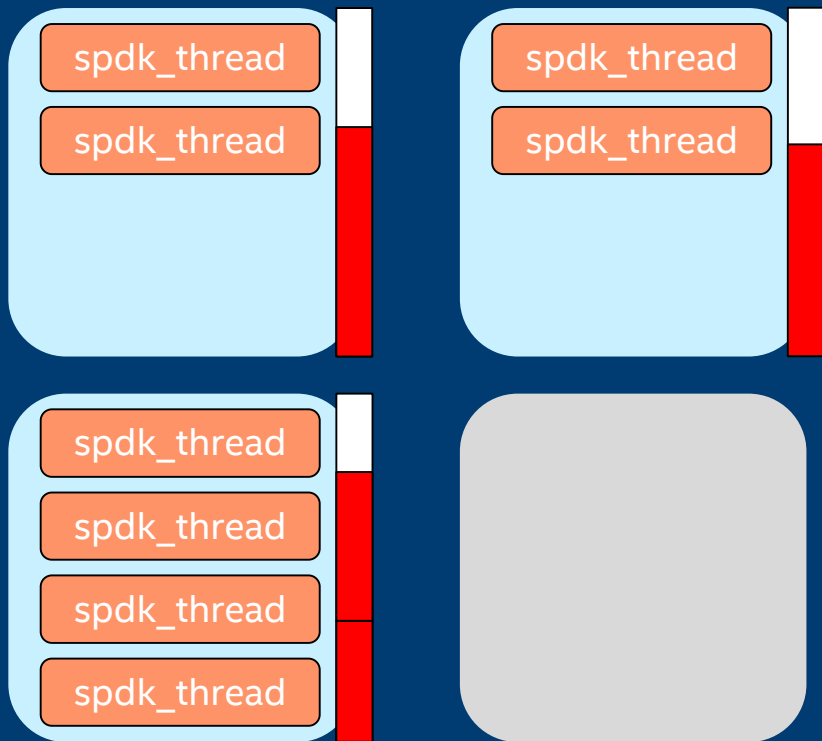


Polled mode = cores always 100%!

Differentiation now between “core” and “thread”

- spdk_thread is a lightweight thread

DYNAMIC SCHEDULING



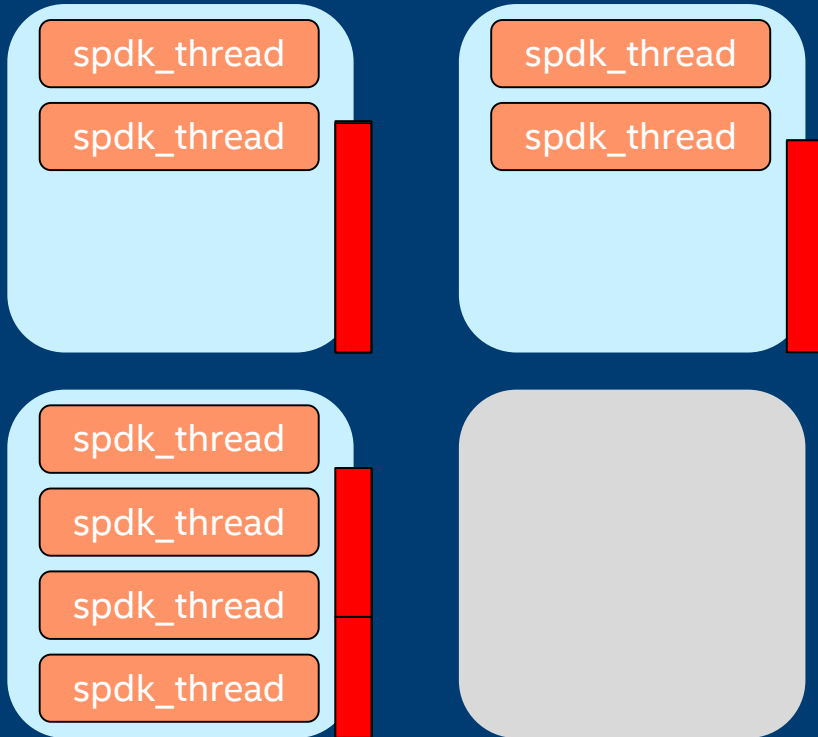
Polled mode = cores always 100%!

Differentiation now between “core” and “thread”

- spdk_thread is a lightweight thread

Dynamically move spdk_threads to enable putting cores to sleep

DYNAMIC SCHEDULING



Polled mode = cores always 100%!

Differentiation now between “core” and “thread”

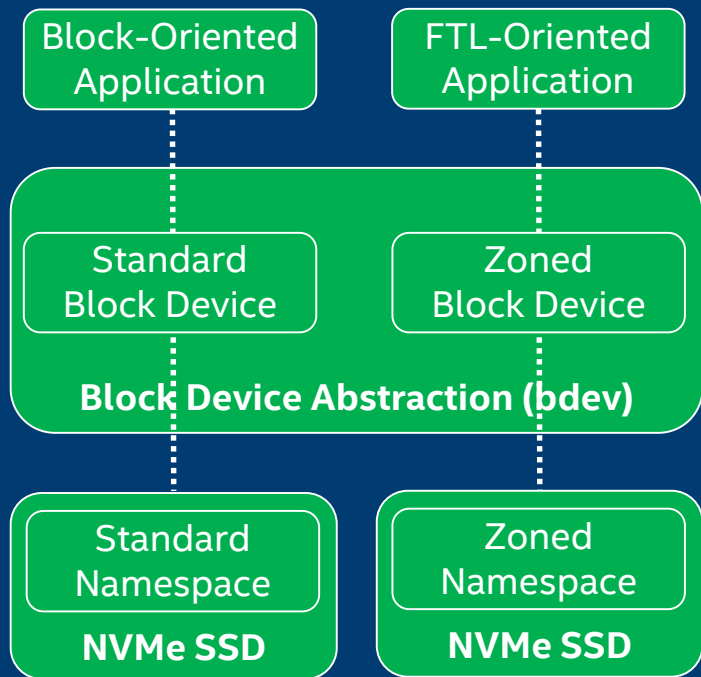
- spdk_thread is a lightweight thread

Dynamically move spdk_threads to enable putting cores to sleep

Scale CPU frequency to reduce power

- i.e. Intel Speed Select Technology

ZONED STORAGE



Provides additional capabilities and responsibilities on host to enable:

- Smart data placement on flash
- Better isolation
- Better quality of service

NVMe standardization in progress

Open Channel via adapters

FUTURE WORK

Block Indirection

Replication and Parity

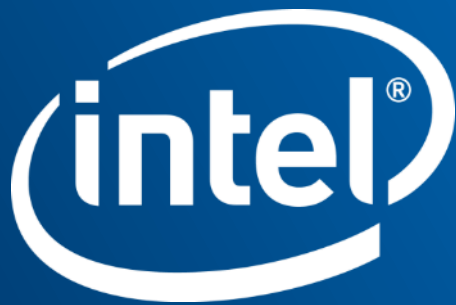
Even better performance!

CALL TO ACTION

Take advantage of talks and demos to learn more about SPDK
and PMDK and VTune™ Amplifier

Join the SPDK Community

- Submit patches! (<https://spdk.io/development/>)
- Submit issues! (<https://github.com/spdk/spdk/issues>)
- Ask a question! (<https://spdk.io/community/>)
 - E-mail, Slack, community meetings



BACKUP

