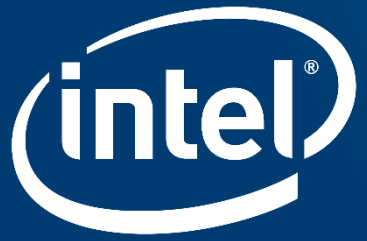


STORAGE PERFORMANCE DEVELOPMENT KIT STATE OF THE PROJECT

Jim Harris

Principal Software Engineer

Intel Data Center Group

Notices & Disclaimers

Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software or service activation. Performance varies depending on system configuration.

No computer system can be absolutely secure.

Tests document performance of components on a particular test, in specific systems. Differences in hardware, software, or configuration will affect actual performance. For more complete information about performance and benchmark results, visit http://www.intel.com/benchmarks .

Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. For more complete information visit http://www.intel.com/benchmarks .

Benchmark results were obtained prior to implementation of recent software patches and firmware updates intended to address exploits referred to as "Spectre" and "Meltdown." Implementation of these updates may make these results inapplicable to your device or system.

Intel® Advanced Vector Extensions (Intel® AVX)* provides higher throughput to certain processor operations. Due to varying processor power characteristics, utilizing AVX instructions may cause a) some parts to operate at less than the rated frequency and b) some parts with Intel® Turbo Boost Technology 2.0 to not achieve any or maximum turbo frequencies. Performance varies depending on hardware, software, and system configuration and you can learn more at http://www.intel.com/go/turbo.

Intel's compilers may or may not optimize to the same degree for non-Intel microprocessors for optimizations that are not unique to Intel microprocessors. These optimizations include SSE2, SSE3, and SSSE3 instruction sets and other optimizations. Intel does not guarantee the availability, functionality, or effectiveness of any optimization on microprocessors not manufactured by Intel. Microprocessor-dependent optimizations in this product are intended for use with Intel microprocessors. Certain optimizations not specific to Intel microarchitecture are reserved for Intel microprocessors. Please refer to the applicable product User and Reference Guides for more information regarding the specific instruction sets covered by this notice.

Cost reduction scenarios described are intended as examples of how a given Intel-based product, in the specified circumstances and configurations, may affect future costs and provide cost savings. Circumstances will vary. Intel does not guarantee any costs or cost reduction.

Intel does not control or audit third-party benchmark data or the web sites referenced in this document. You should visit the referenced web site and confirm whether referenced data are accurate.

© 2018 Intel Corporation.
Intel, the Intel logo, and Intel Xeon are trademarks of Intel Corporation in the U.S. and/or other countries.
*Other names and brands may be claimed as property of others.

- Current State of the Project
- Looking Forward
- Call to Action

- Current State of the Project
- Looking Forward
- Call to Action

One year ago…

Was difficult for people outside Intel to contribute

No transparency on work going on inside Intel

Internal automated test framework

SPDK needed to truly become an open source community – not just an Intel open source code repository!

Since then…

Since then…

Patch reviews – Gerrithub

GerritHub™

All My Projects People Documentation MAKE A DONATION GitHub

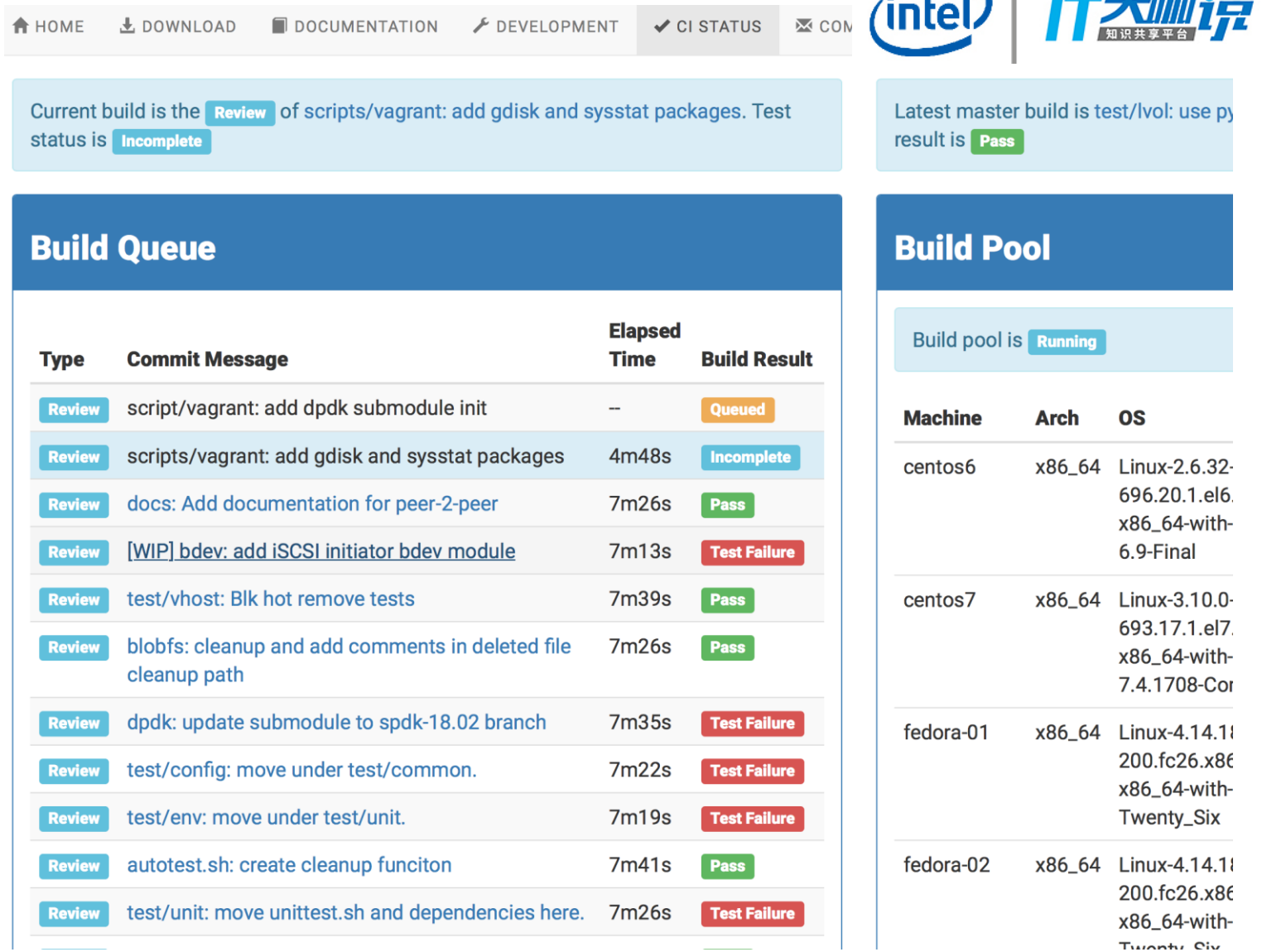
Changes SPDK Drafts Draft Comments Edits Groups Merged

Subject	Status	Owner	Project	Branch	Updated	Size	CR	V
Outgoing								
[WIP] bdev: add iSCSI initiator bdev module		Jim Harris	spdk/spdk	master	14 minutes ago			✗
blobfs: cleanup and add comments in deleted file cleanup path		Jim Harris	spdk/spdk	master	31 minutes ago			✓
bdev/fio_plugin: flush any remaining events after done		Jim Harris	spdk/spdk	master	2 hours ago			✓
blob: queue sync requests if one already in progress	Merge Conflict	Jim Harris	spdk/spdk	master	32 hours ago		+1	✓
examples/blobfs: add fio plugin	Merge Conflict	Jim Harris	spdk/spdk	master	8 months ago			✓
Needs My +2								
nvmf: add Namespace deletion RPC support		Changpeng Liu	spdk/spdk	master	4 hours ago		✓	✓
nvmf: allow configuration of ns UUID	Merge Conflict	Daniel Verkamp	spdk/spdk	master	32 hours ago		✓	✓
bdev/qos: add the QoS rate limiting support on bdev		GangCao	spdk/spdk	master	2 days ago		✓	✓
Waiting For Another +2								
nvme_pci: Correct bug in CMB allocation		Stephen Bates	spdk/spdk	master	28 minutes ago		✓	✓
autotest.sh: create cleanup funciton		Seth Howell	spdk/spdk	master	2 hours ago		✓	✓
lib/iscsi: add a check in spdk_iscsi_conn_sock_cb		Ziye Yang	spdk/spdk	master	4 hours ago		✓	✓
scripts/ceph: remove unused variable 'fsid'		Daniel Verkamp	spdk/spdk	master	5 hours ago		✓	✓
lib/iscsi: rename spdk_iscsi_conn_execute		Ziye Yang	spdk/spdk	master	5 hours ago		✓	✓
lib/iscsi: move location of spdk_iscsi_conn_handle_queued_datain_tasks		Ziye Yang	spdk/spdk	master	5 hours ago		✓	✓
bdev/qos: add the QoS parameters for the bdev channel		GangCao	spdk/spdk	master	2 days ago		✓	✓
blobstore: Add a blob_bs_dev that provides back_bs_dev for clones		Piotr Pelpliński	spdk/spdk	master (396387)	2 days ago		✓	✓
doc: Programmer's guide template and example for blobstore		Paul Luse	spdk/spdk	master (18.01)	2 days ago		✓	✓
Needs Review								
docs: Add documentation for peer-2-peer		Stephen Bates	spdk/spdk	master	4 minutes ago			✓
lib: Return instead of exit in event		Lance Hartmann	spdk/spdk	master	51 minutes ago		-1	✓
scripts/ceph: remove subshell and cd		Daniel Verkamp	spdk/spdk	master	4 hours ago			✓
bdev/virtio/scsi: hotremove targets on EVENTS_MISSED rescan		Dariusz Stojaczyk	spdk/spdk	master	5 hours ago			✓
bdev/rpc: add "get_bdev_dependencies" RPC call		Shuhei Matsumoto	spdk/spdk	master	5 hours ago		-1	✓
autotest: iscsi_fio_test (porting from nightly to autotest)		Chen, LongX O	spdk/spdk	master	12 hours ago		-1	✓

Since then…

Patch reviews – Gerrithub

Public Continuous Integration

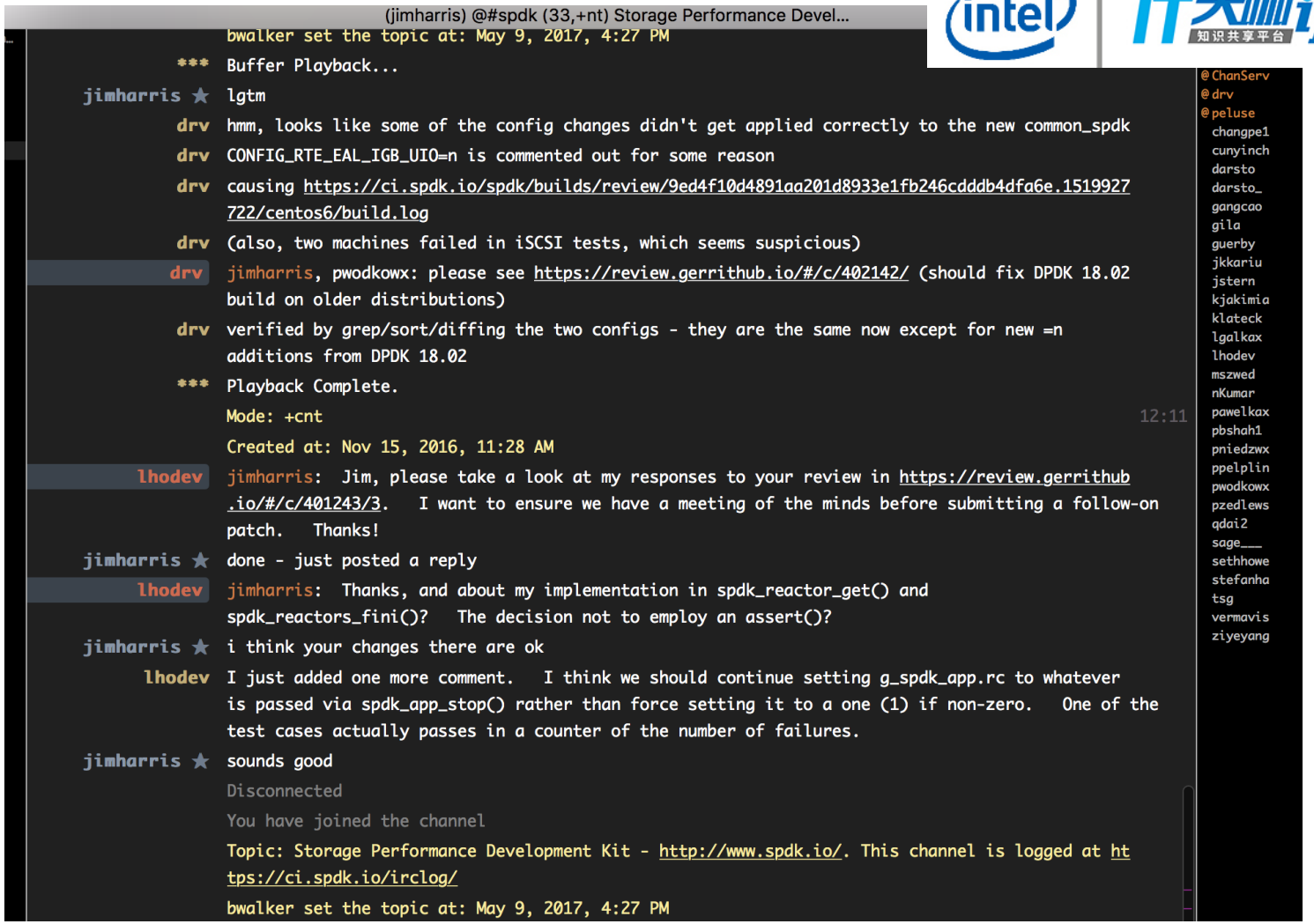


Since then…

Patch reviews – Gerrithub

Public Continuous Integration

#spdk (IRC)



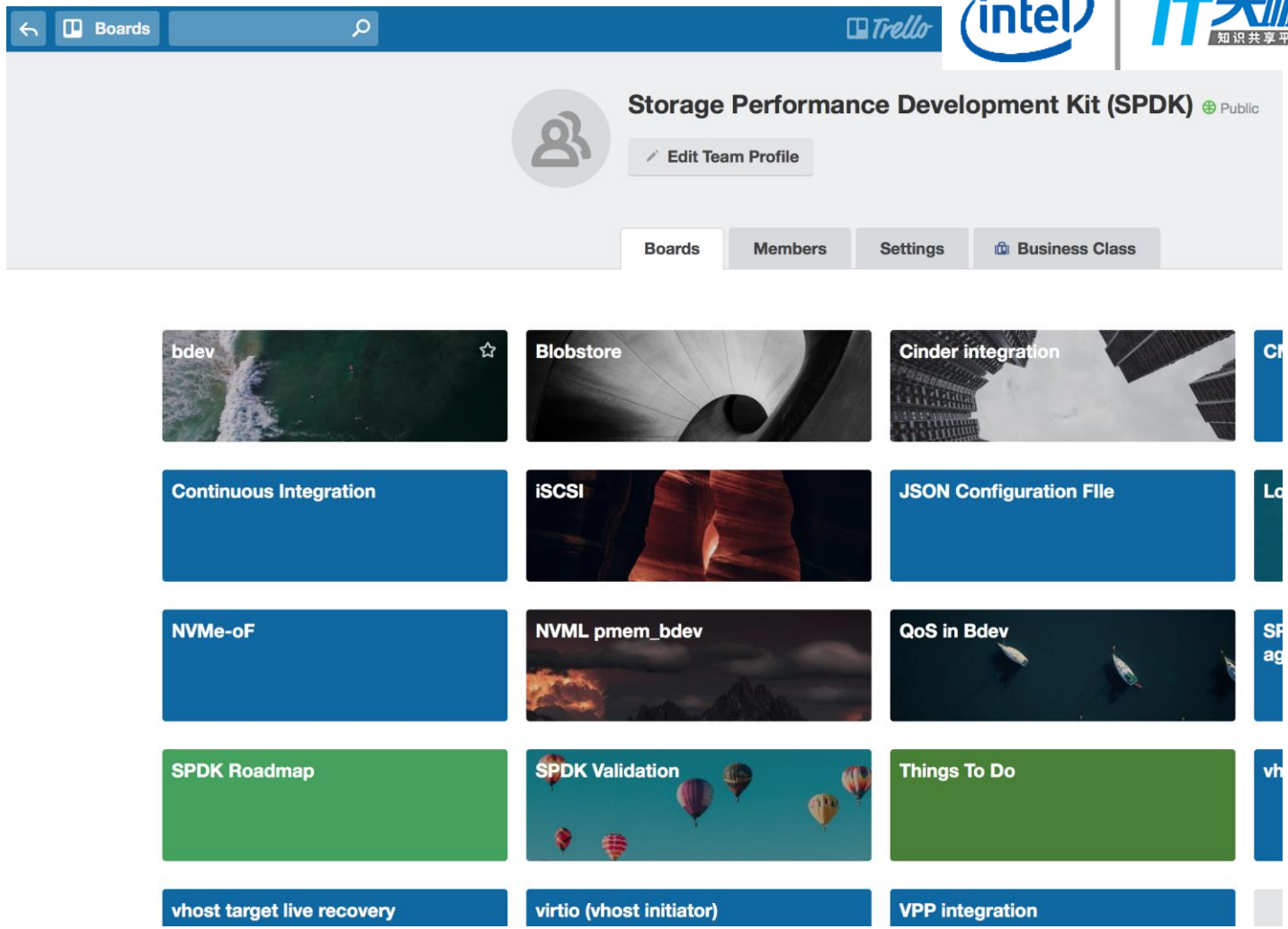
Since then…

Patch reviews – Gerrithub

Public Continuous Integration

#spdk (IRC)

Trello



Since then…

Patch reviews – Gerrithub

Public Continuous Integration

#spdk (IRC)

Trello

Meetups

- SPDK US Summit (Santa Clara, CA)
- SPDK Developer Meetup (Chandler, AZ)

Since then…

Patch reviews – Gerrithub

Public Continuous Integration

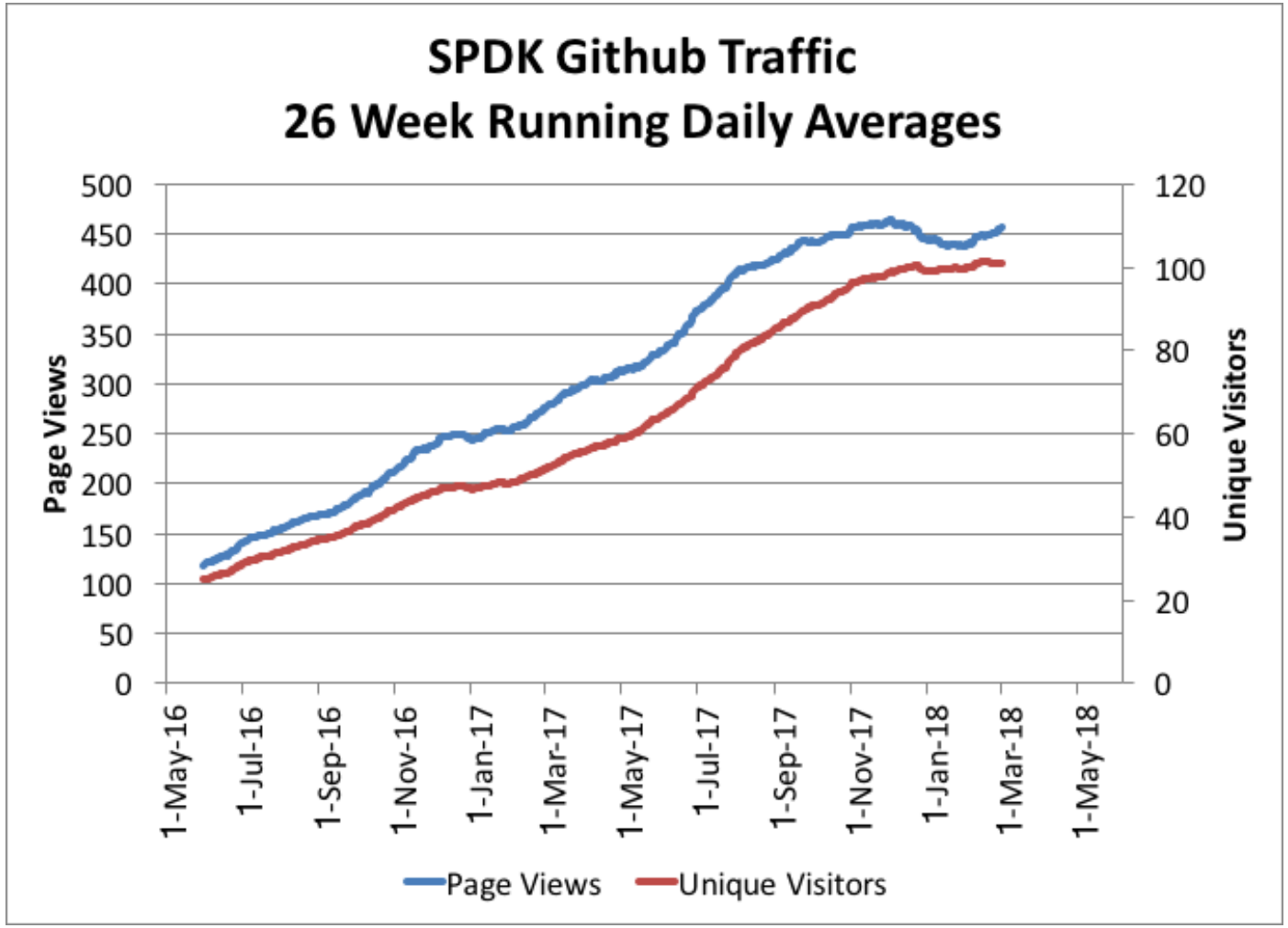
#spdk (IRC)

Trello

Meetups

- SPDK US Summit (Santa Clara, CA)
- SPDK Developer Meetup (Chandler, AZ)
- **SPDK China Summit (北京)**

Community Growth



Compare:

- 6 months ending 1-Mar-17
- 6 months ending 1-Mar-18

Page Views +65%

Unique Visitors +96%

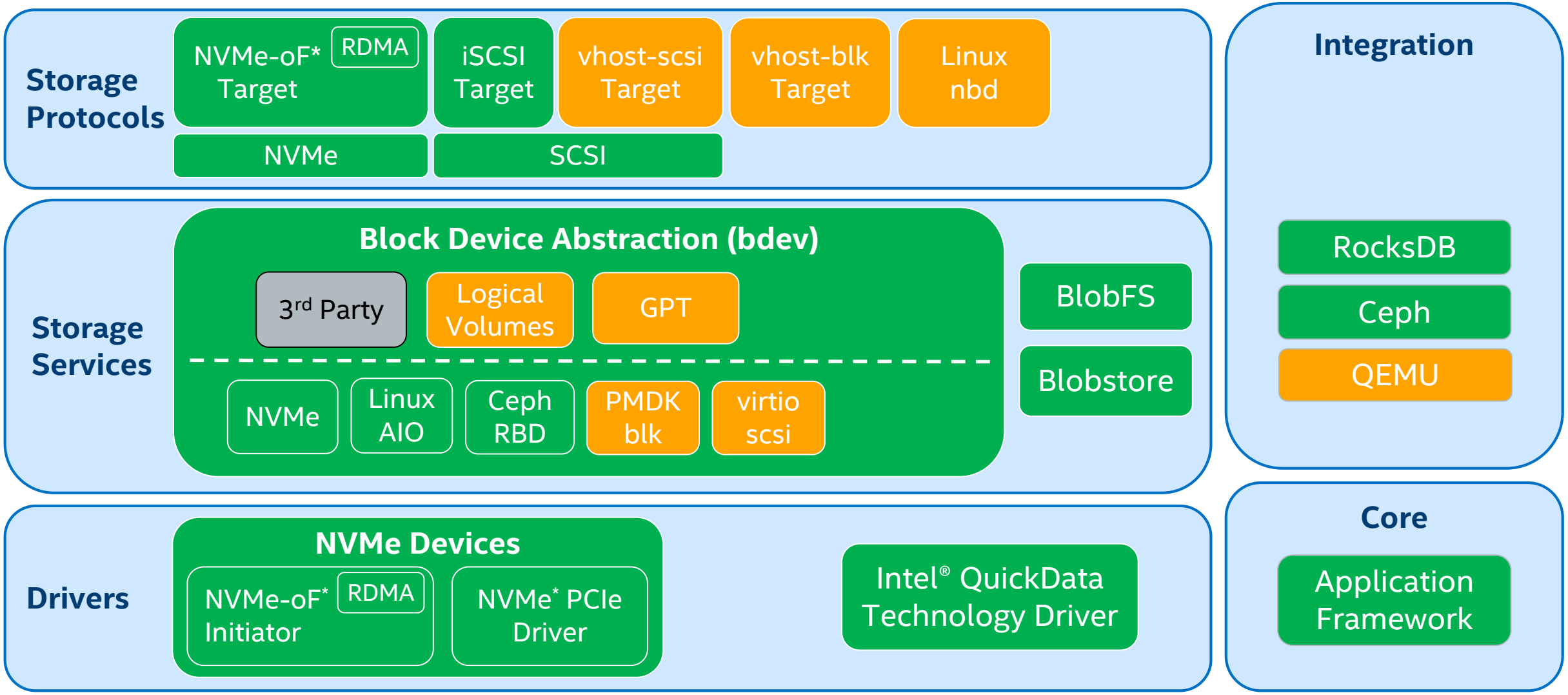
SPDK Patch Commit Statistics

SPDK Version	Total Commits	Committers outside Intel	Commits outside Intel	% Commits outside Intel
18.04 (so far)	439	7	56	12.8%
18.01	775	14	79	10.2%
17.10	753	8	17	2.2%
17.07	685	7	11	1.6%
17.03	483	8	18	3.7%

SPDK Architecture

intel | IT大咖说 知识共享平台

1H'18



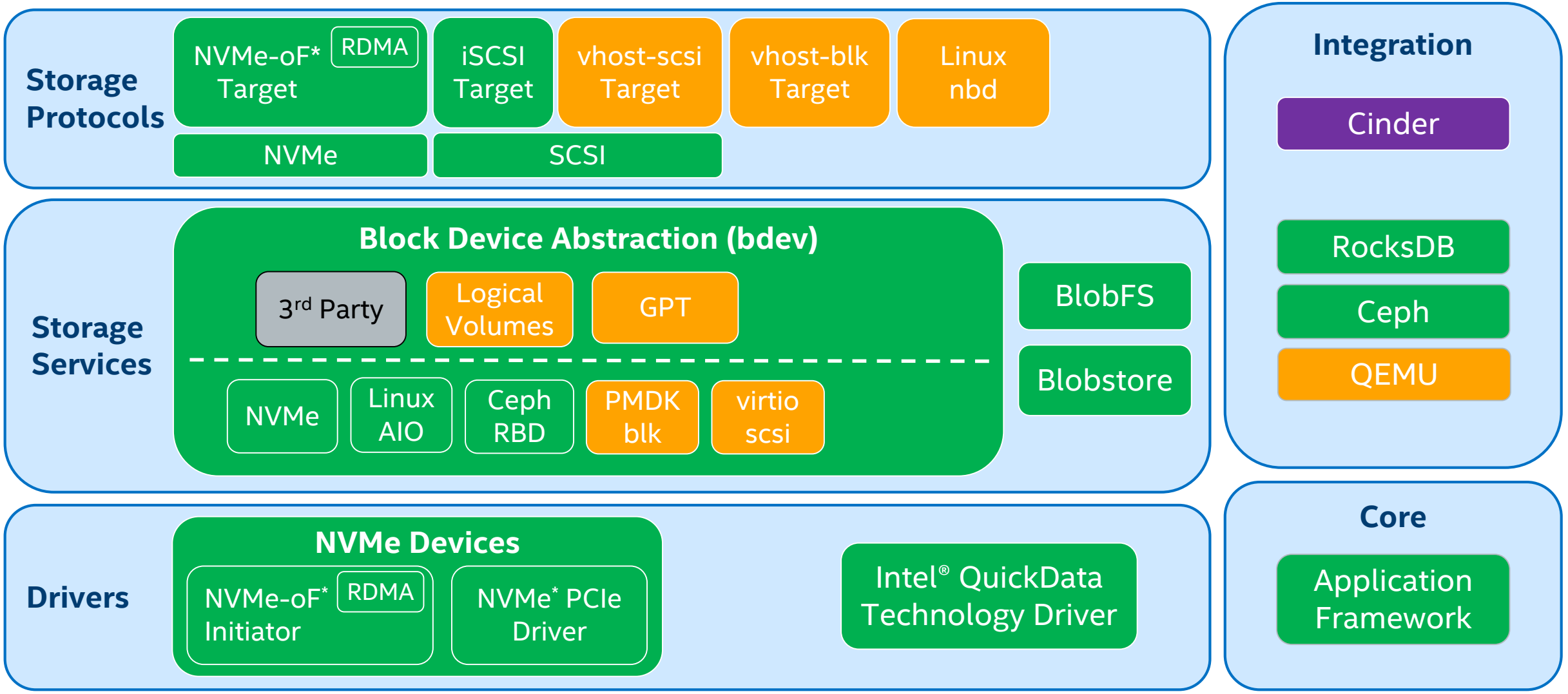
intel

- Current State of the Project
- Looking Forward
- Call to Action

SPDK Architecture

intel IT大咖说 知识共享平台

1H'18



intel

Orchestration

Cinder plugin for SPDK

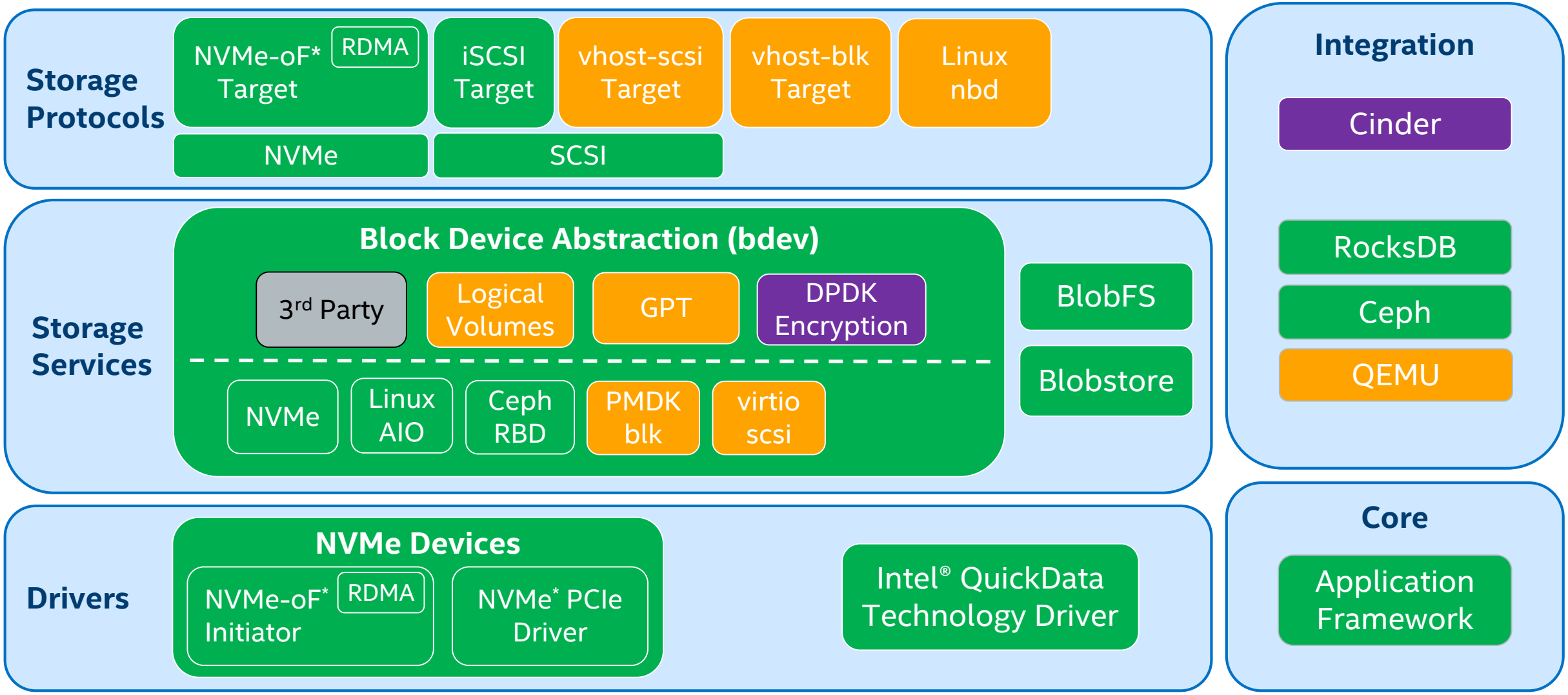
Provision ephemeral storage using SPDK:

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

Enables SPDK ephemeral storage provisioning for OpenStack!

SPDK Architecture

1H'18



Accelerators

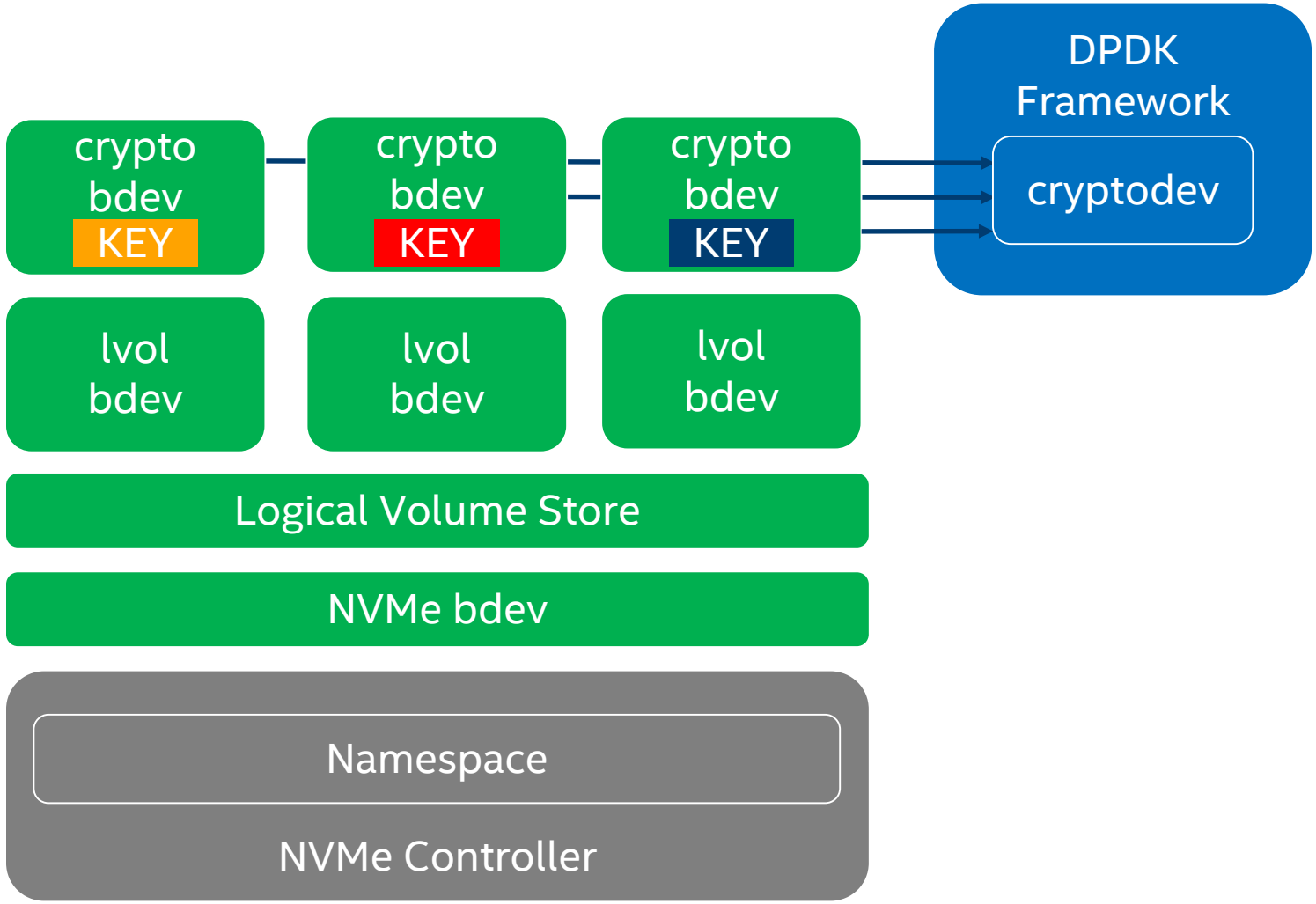
DPDK Framework

- Hardware Accelerators
- Optimized Software

Key per bdev

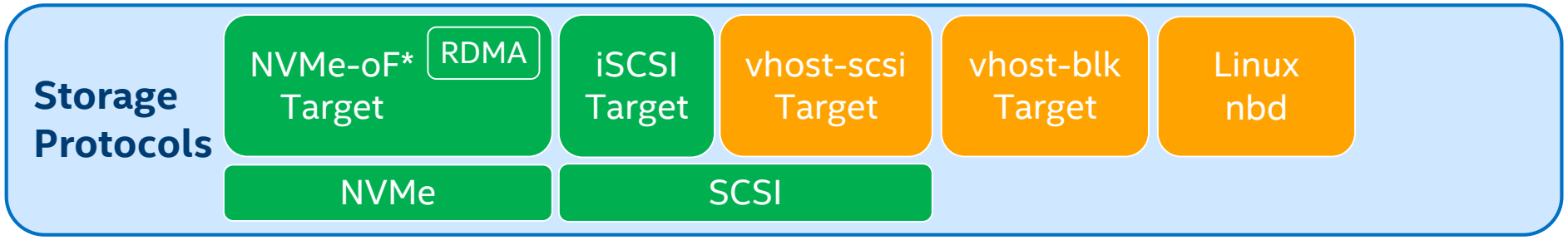
No SPDK key management

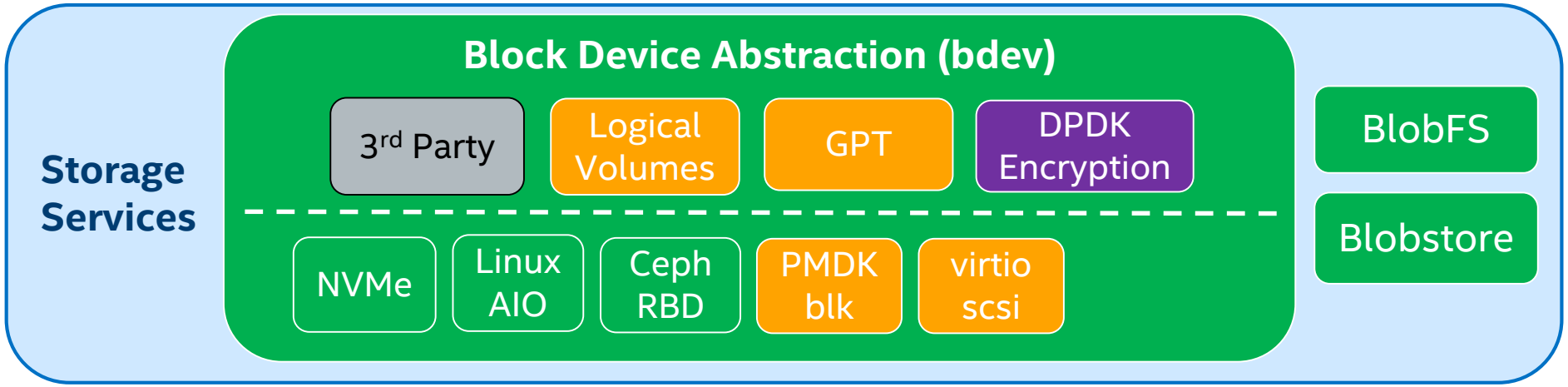
- Keys passed to SPDK over privileged JSON-RPC channel

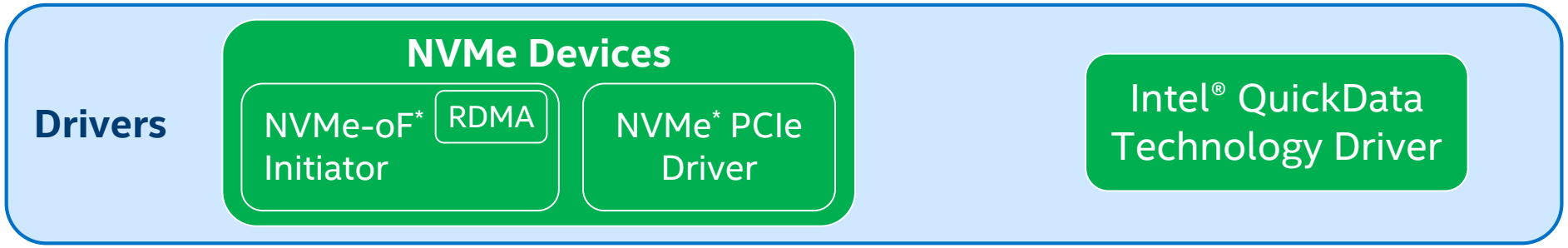


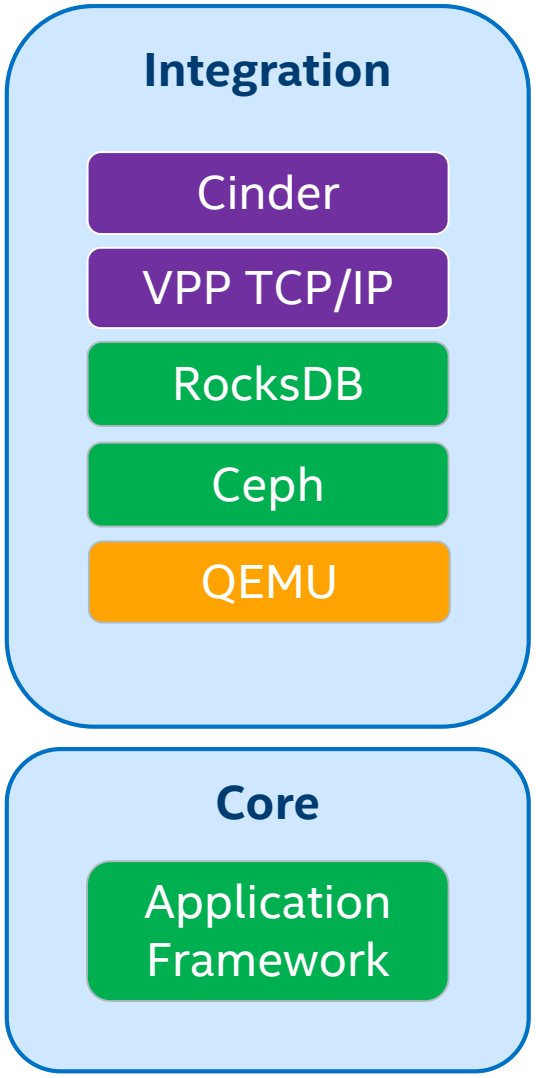
SPDK Architecture

1H'18









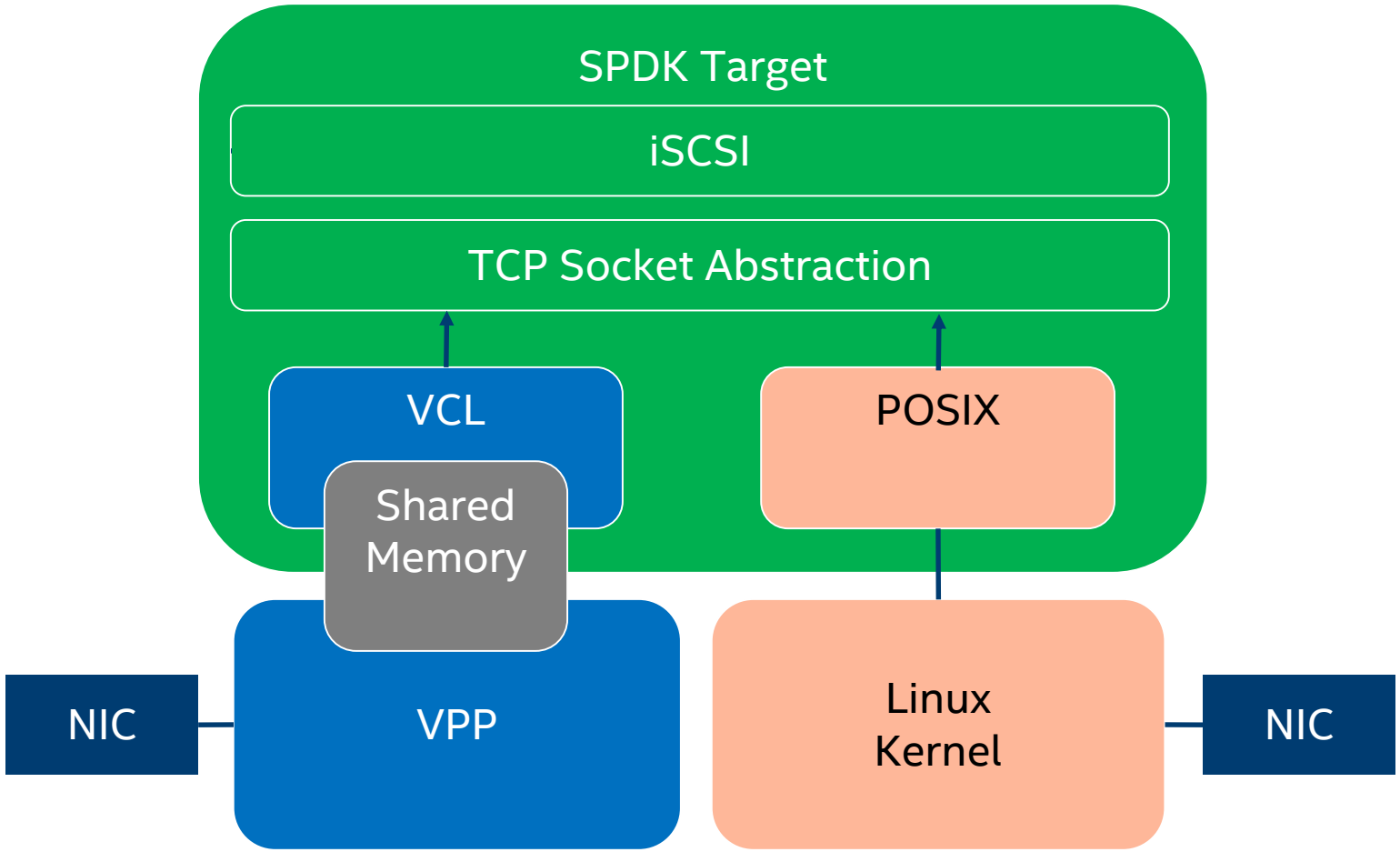
Userspace TCP/IP

FD.io – The Fast Data Project

- VPP – Vector Packet Processing

Utilize VCL (VPP Communications Library) for accelerated TCP processing

Integration with SPDK iSCSI targeted for v18.04 release



SPDK Architecture

Storage Protocols

- NVMe-oF* Target (RDMA, TCP)
- iSCSI Target
- vhost-scsi Target
- vhost-blk Target
- Linux nbd
- NVMe
- SCSI

Storage Services

Block Device Abstraction (bdev)

- 3rd Party
- Logical Volumes
- GPT
- DPDK Encryption
- NVMe
- Linux AIO
- Ceph RBD
- PMDK blk
- virtio scsi

- BlobFS
- Blobstore

Drivers

NVMe Devices

- NVMe-oF* Initiator (RDMA, TCP)
- NVMe* PCIe Driver

- Intel® QuickData Technology Driver

Integration

- Cinder
- VPP TCP/IP
- RocksDB
- Ceph
- QEMU

Core

- Application Framework

TCP Transport for NVMe over Fabrics

NVM Express adding TCP as new NVMe over Fabrics transport

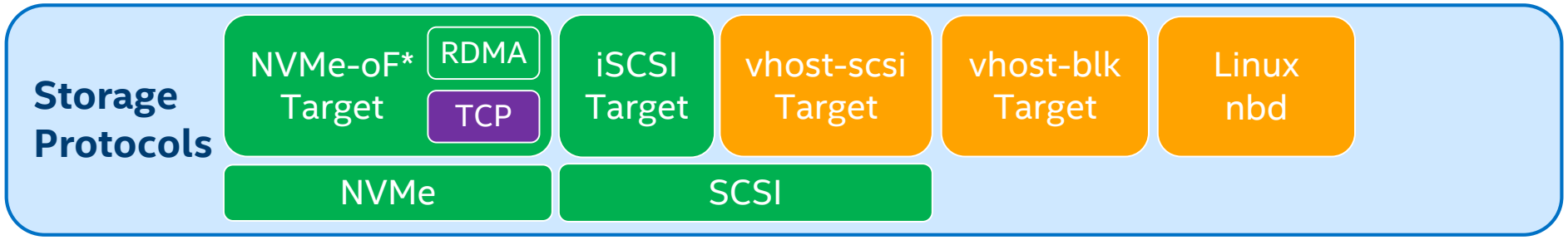
- Spec release expected later this year

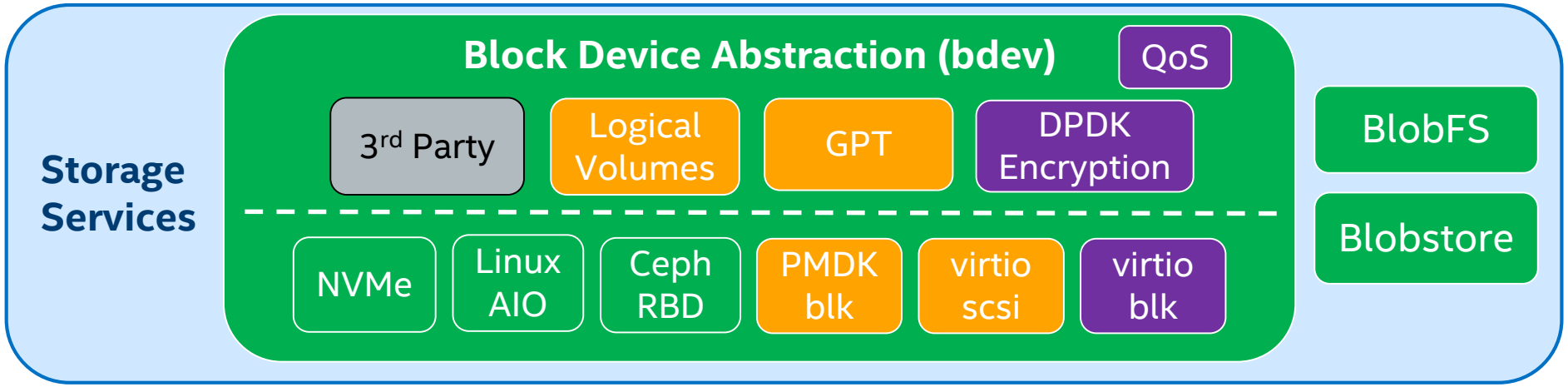
SPDK adding NVMe-oF TCP transport

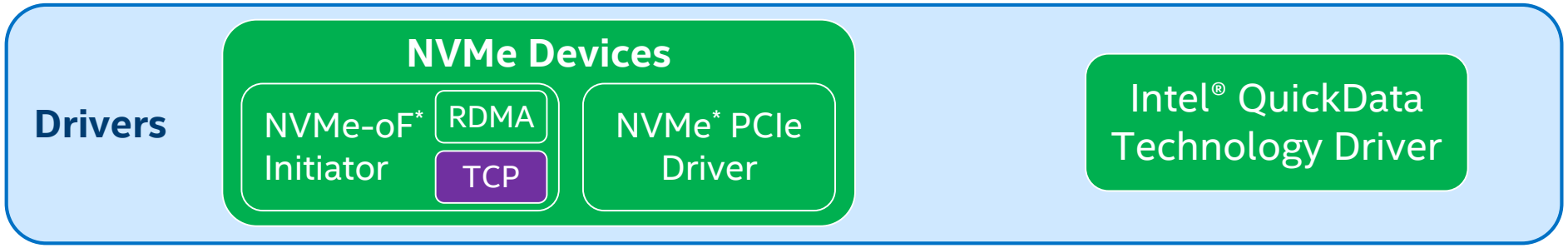
- Aligned with spec release
- Kernel TCP stack first
- VPP integration possible in the future

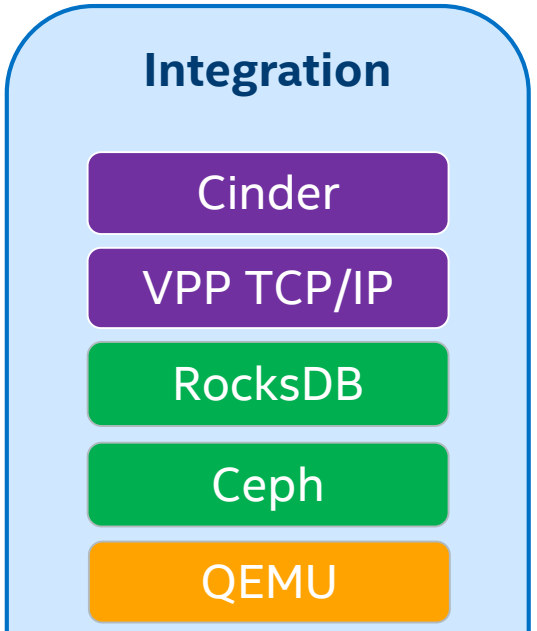
SPDK Architecture

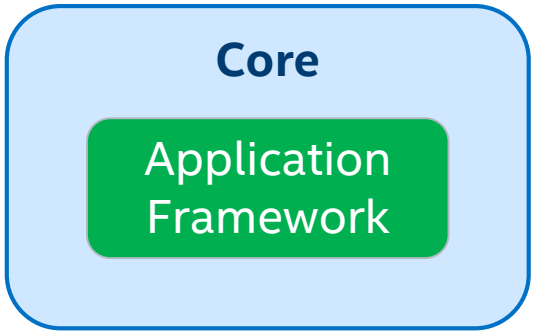
1H'18











and more...

Block Device Quality of Service

- Rate limiting per bdev

virtio-blk Polled Mode Driver

- Enable SPDK applications on virtio-blk storage in QEMU-based VMs

Packaging

- Working towards easier deployment

JSON Configuration

Future Areas of Development

Block Device Aggregation

- striping (RAID-0)
- concatenation

Mirroring/Replication

- RAID-1, remote replication

Fibre Channel

Data compression

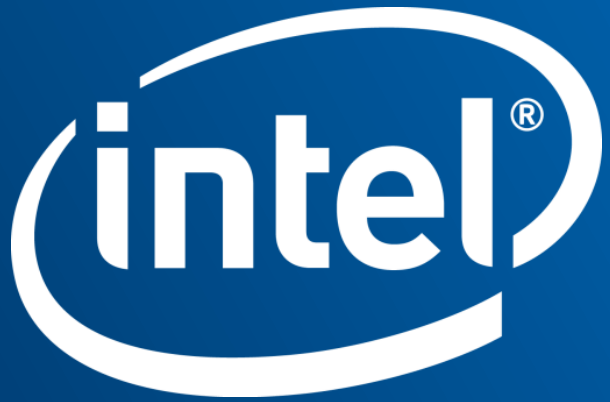
Interactive Configuration Tools

- Current State of the Project
- Looking Forward
- Call to Action

Call to Action

Learn today from industry leaders and SPDK experts

Join the community!


intel