

# Storage performance development kit: USING DPDK TO Accelerate Storage services

Jim Harris

Principal Software Engineer

Intel Data Center Group

# Legal Disclaimer

INFORMATION IN THIS DOCUMENT IS PROVIDED IN CONNECTION WITH INTEL PRODUCTS. NO LICENSE, EXPRESS OR IMPLIED, BY ESTOPPEL OR OTHERWISE, TO ANY INTELLECTUAL PROPERTY RIGHTS IS GRANTED BY THIS DOCUMENT. EXCEPT AS PROVIDED IN INTEL'S TERMS AND CONDITIONS OF SALE FOR SUCH PRODUCTS, INTEL ASSUMES NO LIABILITY WHATSOEVER AND INTEL DISCLAIMS ANY EXPRESS OR IMPLIED WARRANTY, RELATING TO SALE AND/OR USE OF INTEL PRODUCTS INCLUDING LIABILITY OR WARRANTIES RELATING TO FITNESS FOR A PARTICULAR PURPOSE, MERCHANTABILITY, OR INFRINGEMENT OF ANY PATENT, COPYRIGHT OR OTHER INTELLECTUAL PROPERTY RIGHT.

UNLESS OTHERWISE AGREED IN WRITING BY INTEL, THE INTEL PRODUCTS ARE NOT DESIGNED NOR INTENDED FOR ANY APPLICATION IN WHICH THE FAILURE OF THE INTEL PRODUCT COULD CREATE A SITUATION WHERE PERSONAL INJURY OR DEATH MAY OCCUR.

Intel may make changes to specifications and product descriptions at any time, without notice. Designers must not rely on the absence or characteristics of any features or instructions marked "reserved" or "undefined." Intel reserves these for future definition and shall have no responsibility whatsoever for conflicts or incompatibilities arising from future changes to them. The information here is subject to change without notice. Do not finalize a design with this information.

The products described in this document may contain design defects or errors known as errata which may cause the product to deviate from published specifications. Current characterized errata are available on request.

Contact your local Intel sales office or your distributor to obtain the latest specifications and before placing your product order.

All products, computer systems, dates, and figures specified are preliminary based on current expectations, and are subject to change without notice.

This document contains information on products in the design phase of development.

Applies only to halogenated flame retardants and PVC in components. Pursuant to JEP-709, halogens are below 1,000ppm bromine and 1,000ppm chlorine. The replacement of halogenated flame retardants and/or PVC may not be better for the environment.

Intel and the Intel logo are trademarks of Intel Corporation in the U.S. and other countries.

\*Other names and brands may be claimed as the property of others.

Copyright © 2017 Intel Corporation. All rights reserved.

# The Challenge: Media Latency

Latency

10,000

200

175

150

125

100

75

50

25

0

kernel driver Overhead <0.01%

kernel driver overhead 1-8%

kernel driver overhead 30%-50%

HDD +SAS/SATA  
SSD NAND +SAS/SATA  
SSD NAND +NVM eM  
SSD optane™ +NVMe™

Drive Latency Controller Latency Driver Latency

Technology claims are based on comparisons of latency, density and write cycling metrics amongst memory technologies recorded on published specifications of in-market memory products against internal Intel specifications.

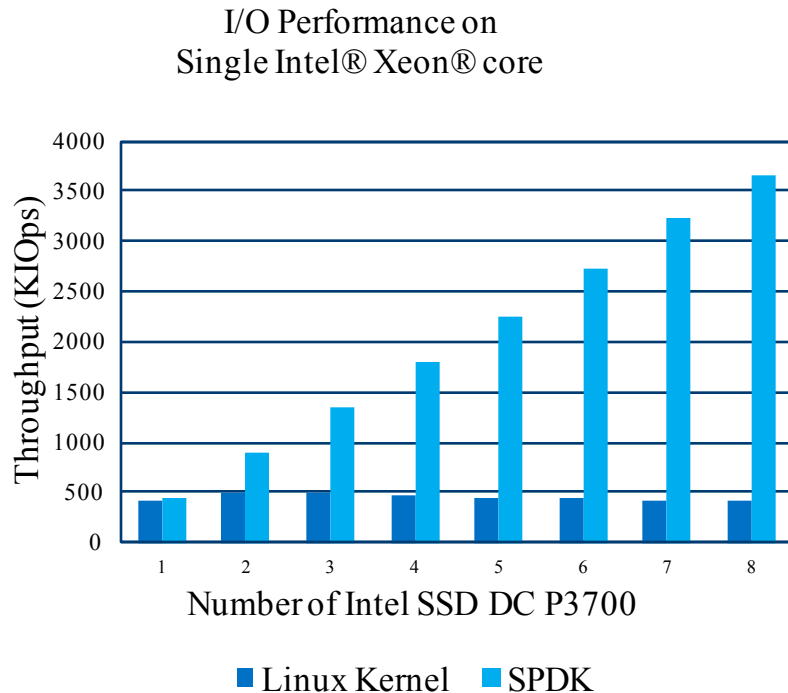


# Storage Performance Development Kit (SPDK)

## What is SPDK?

- Userspace polled-mode drivers, libraries and applications for storage, storage networking and storage virtualization
- Leverages DPDK
- Started in 2013, open sourced in 2015
- BSD licensed
- <http://SPDK.io>

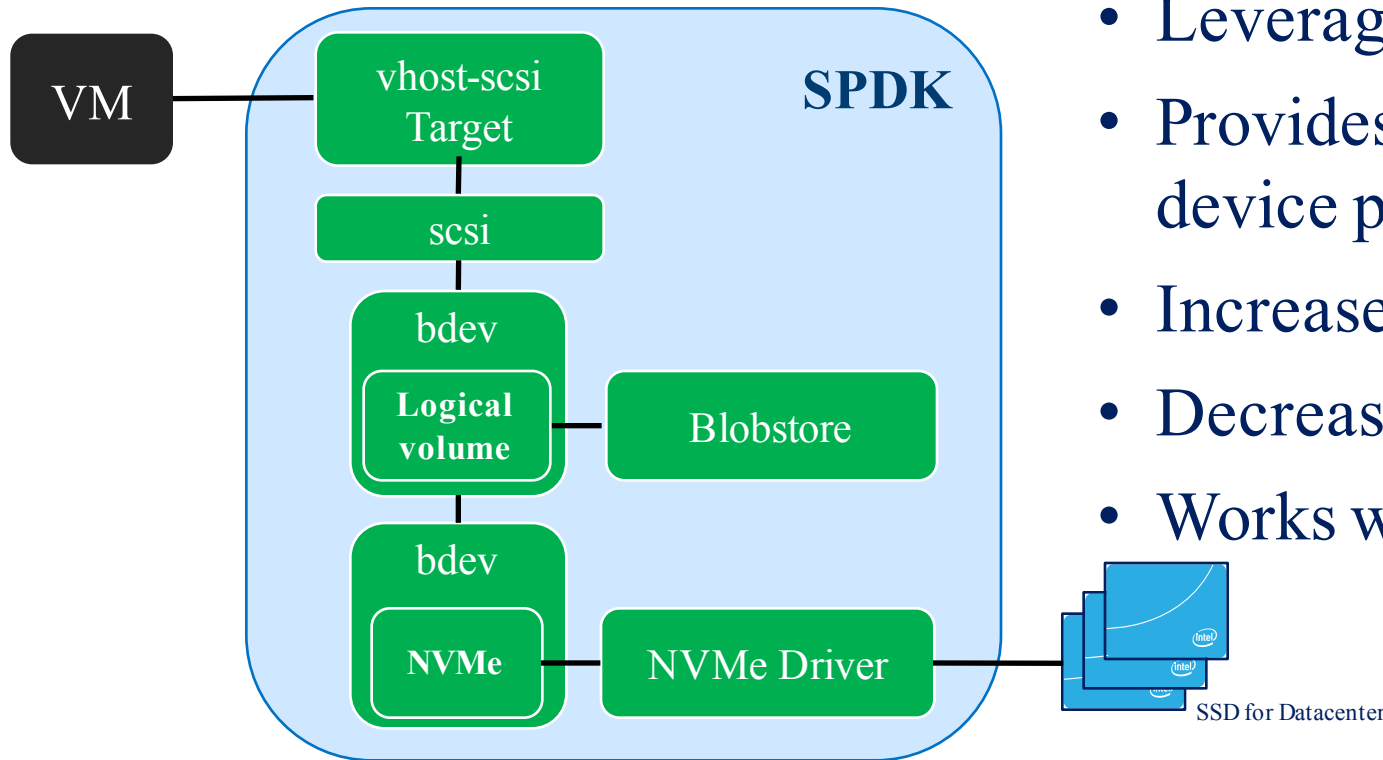
# NVM Express\* Driver Throughput Scalability



- Systems with multiple NVM Express\* (NVMe) SSDs capable of millions of I/O per second
- Results in many cores of software overhead with kernel-based interrupt-driven driver model
- SPDK enables:
  - more CPU cycles for storage services
  - lower I/O latency

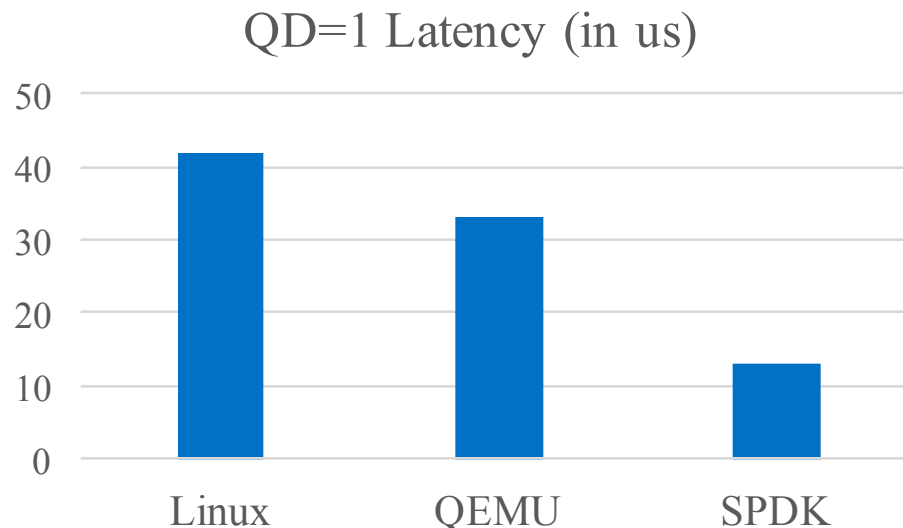
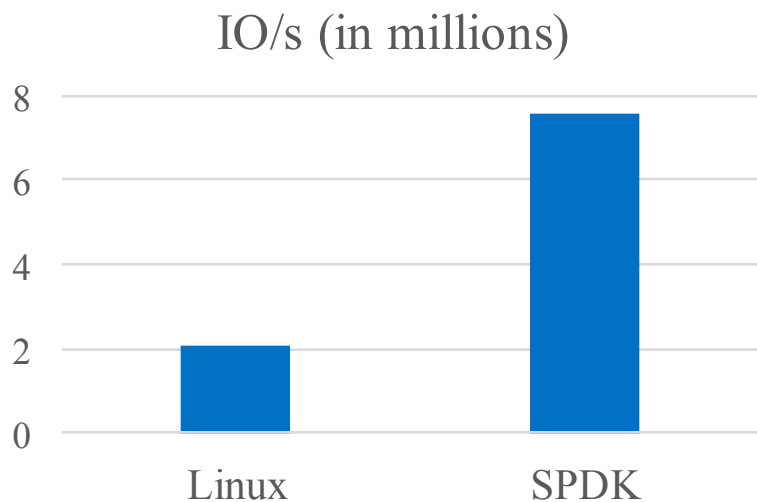
**SPDK saturates 8 NVMe SSDs with a single CPU core!**

# VM Storage Acceleration



- Leverages DPDK vhost
- Provides dynamic block device provisioning
- Increase VM Density
- Decrease Guest Latency
- Works with KVM/QEMU

# SPDK vhost Performance



**SPDK up to 3x better efficiency and latency**

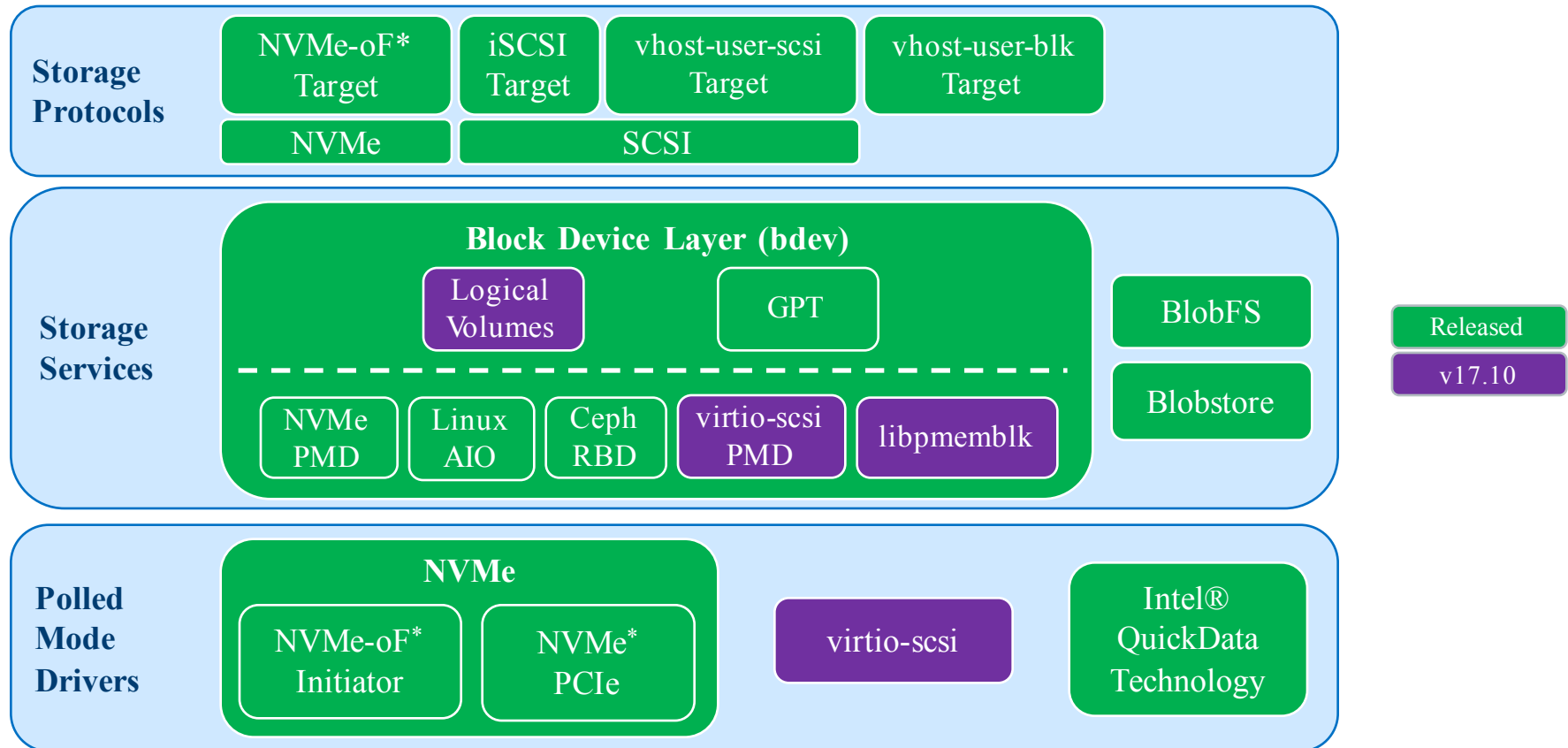
System Configuration: 2S Intel® Xeon® Platinum 8180: 28C, E5-2699v3: 18C, 2.5GHz (HT off), Intel® Turbo Boost Technology enabled, 12x16GB DDR4 2133 MT/s, 1 DIMM per channel, Ubuntu® Server 16.04.2 LTS, 4.11 kernel, 23x Intel® P4800x Optane SSD – 375GB, 1 SPDK lvolstore or LVM lvggroup per SSD, SPDK commit ID c5d8b108f22ab, 46 VMs (CentOS 3.10, 1vCPU, 2GB DRAM, 100GB logical volume), vhost dedicated to 10 cores

As measured by: fio 2.10.1 – Direct=Yes, 4KB random read I/O, Ramp Time=30s, Run Time=180s, Norandommap=1, I/O Engine = libaio, Numjobs=1

Legend: Linux: Kernel vhost-scsi QEMU: virtio-blk dataplane SPDK: Userspace vhost-scsi



# Storage Performance Development Kit (SPDK)





# DPDK Key Features for SPDK

- Threads
- PCIe Device Management
- Memory Management
- Rings, Mempools
- Multi-Process
- vhost

# Storage v. Packet Processing

- PCIe Device Hotplug
- Runtime v2phys Translation
- vhost VM Boot
- Storage is Endpoint Focused

# SPDK Community

**Home Page** : <http://www.SPDK.io/>

**Github** : <https://github.com/spdk/spdk>

**Trello** : <https://trello.com/spdk>

**GerritHub** : <https://review.gerrithub.io/#/q/project:spdk/spdk+status:open>

**IRC** : <https://freenode.net/> we're on #spdk

