



FEBRUARY 4-5, 2020
TEL AVIV, ISRAEL

STORAGE DEVELOPER
CONFERENCE

Shared Storage Using NVMe Over Fabrics

Murali Rajagopal (VMware)

Raj Lalsangi/Madhu Pai (NetApp)

VMware Disclaimer

This presentation may contain product features or functionality that are currently under development.

This overview of new technology represents no commitment from VMware to deliver these features in any generally available product.

Features are subject to change, and must not be included in contracts, purchase orders, or sales agreements of any kind.

Technical feasibility and market demand will affect final delivery.

Pricing and packaging for any new features/functionality/technology discussed or presented, have not been determined.

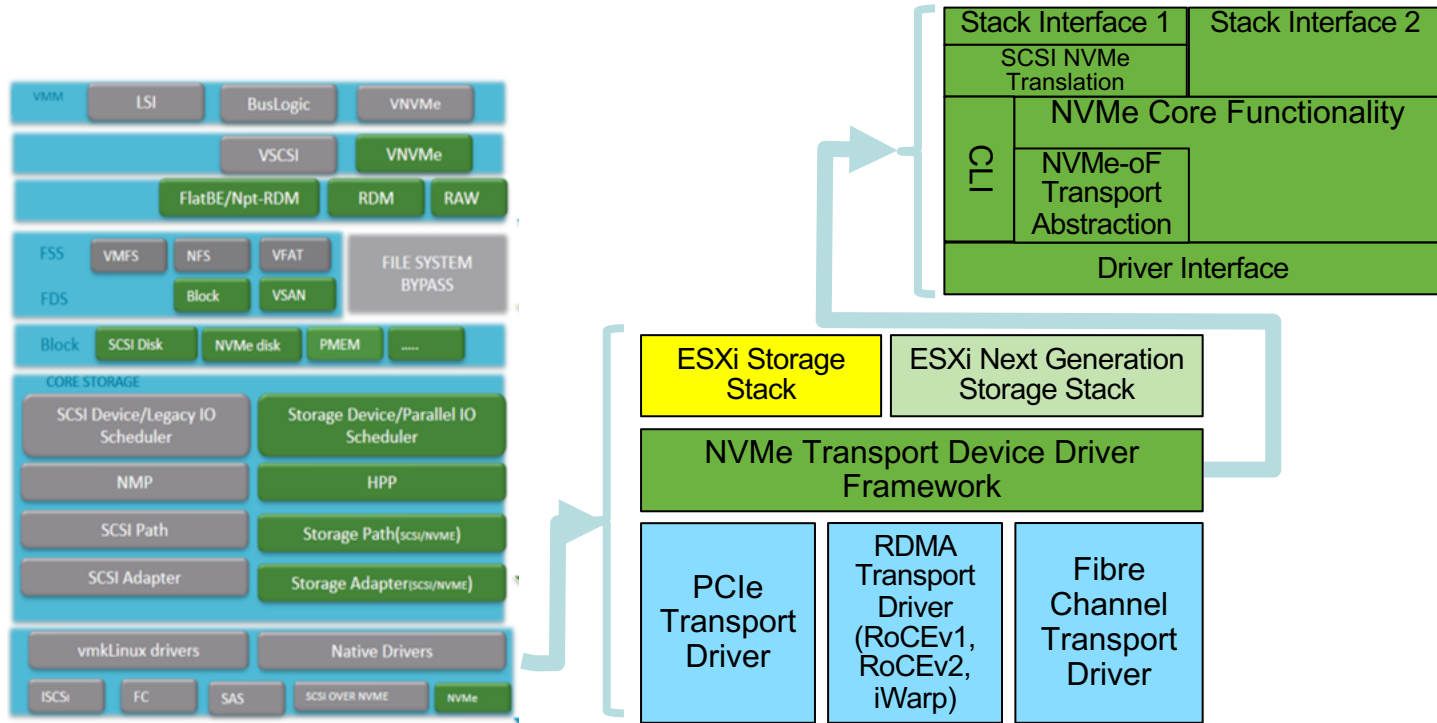
Agenda

- ❑ NVMe Support in VMWare
- ❑ NetApp ONTAP NVMe/f Architecture

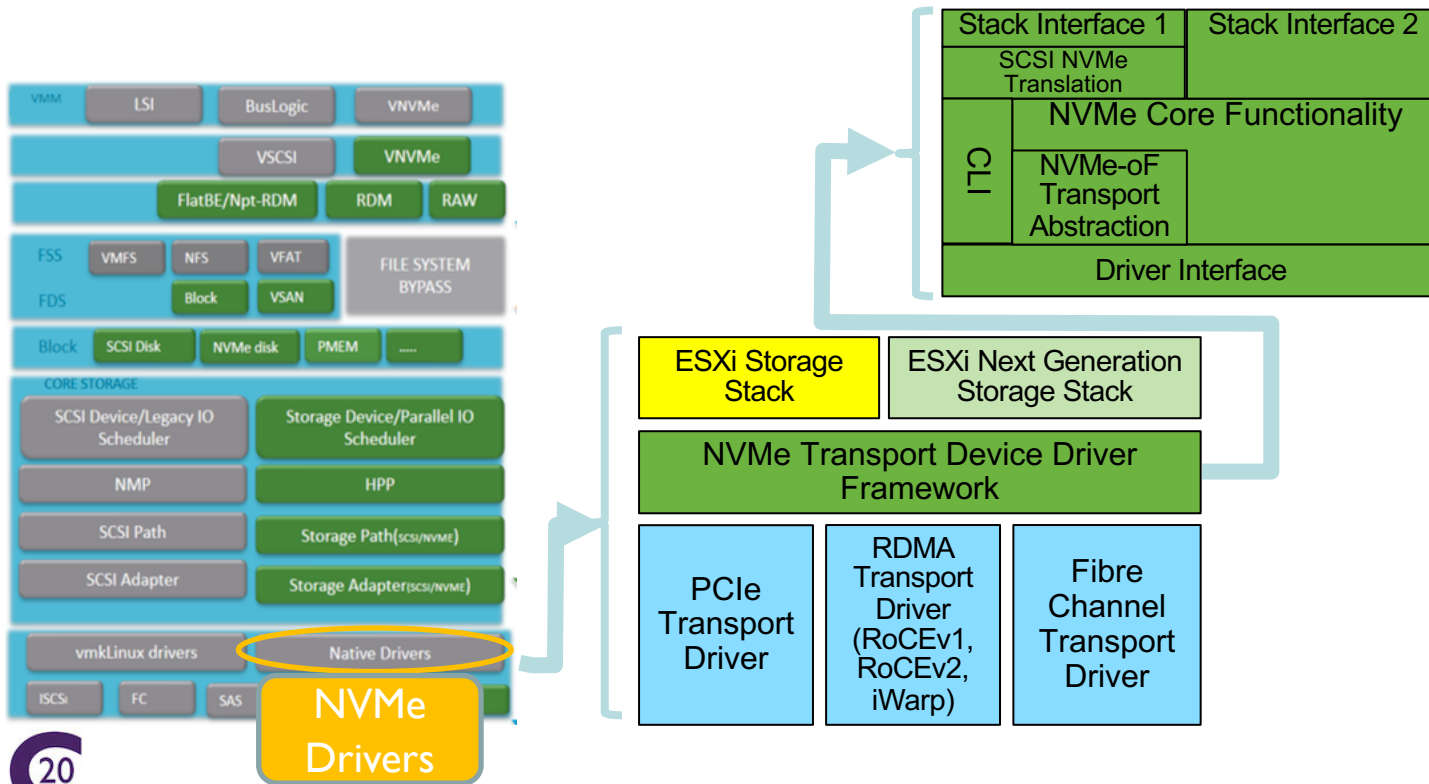
Agenda

- ❑ **NVMe Support in VMWare**
- ❑ NetApp ONTAP NVMe/f Architecture

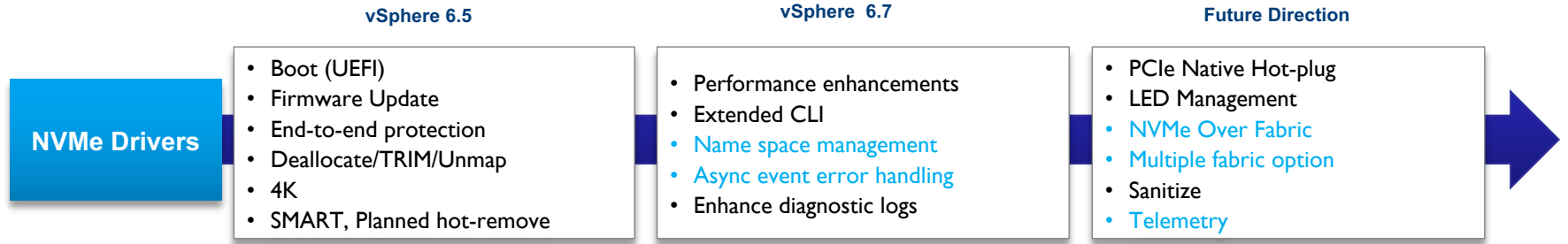
Storage Stack Architecture



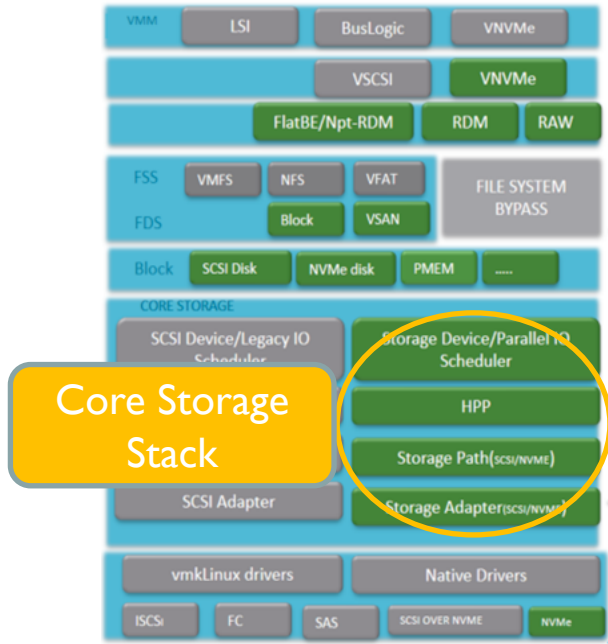
NVMe Drivers



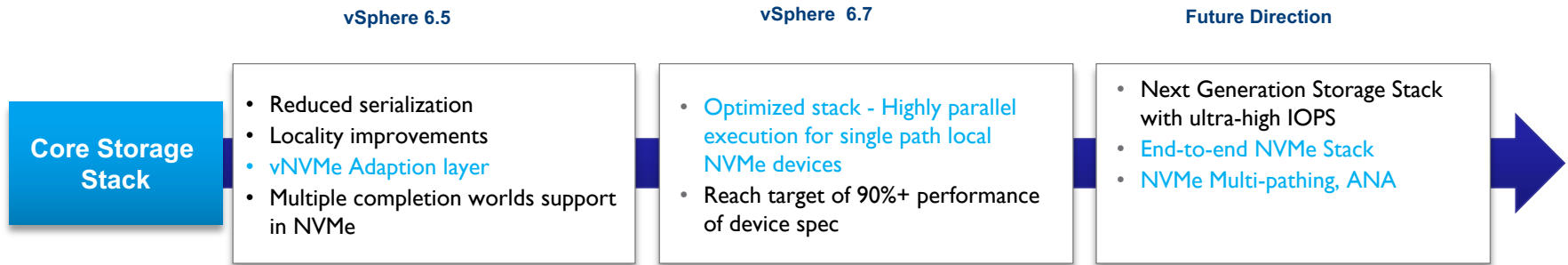
NVMe Driver Support in VMWare



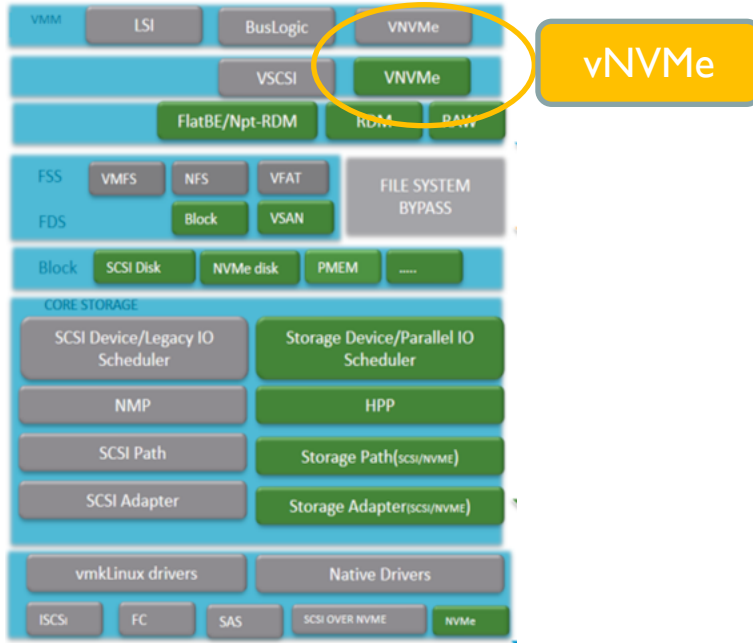
Core Storage Stack



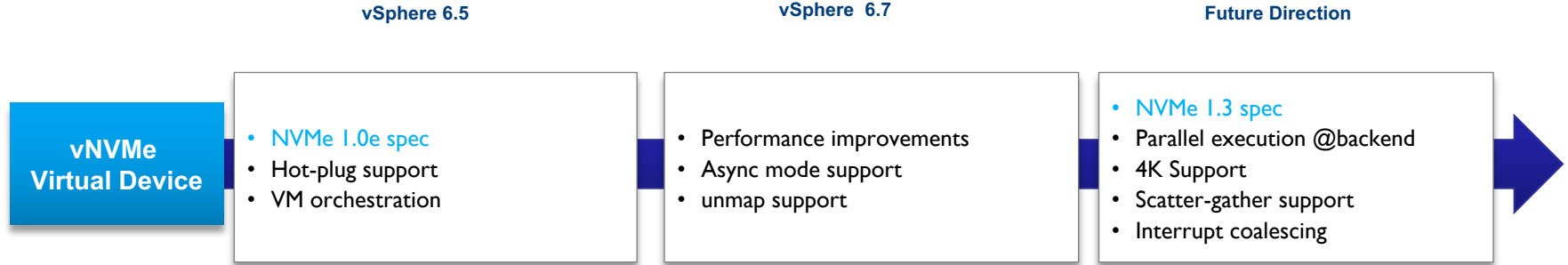
NVMe Support in Core Storage Stack



vNVMe



vNVMe Support in VMWare



Agenda

- NVMe Support in VMWare
- **NetApp ONTAP NVMe/f Architecture**

NetApp NVMe/f

- A SERIES OF FIRSTS IN THE INDUSTRY
 - 9.4 (May '18) – First NVMe/FC solution to the market with bare metal SUSE Linux.
 - 9.5 (Dec '18) – Asymmetric Namespace Access (ANA), Redhat Enterprise Linux
 - 9.6 (May '19) – Windows support
 - 9.7 (Dec '19) – ESX support.
- Supported on AFF arrays with 32G FC Target Ports
- No new hardware/fabric!!
- No licensing costs
- Seamless migration from SCSI

<https://www.netapp.com/us/media/tr-4684.pdf>

Core Design Principles

- Lockless
- Zero Context Switching
- Polling mode
- Small Code path length
- Dynamic threading
- Zero copy

Innovation Velocity

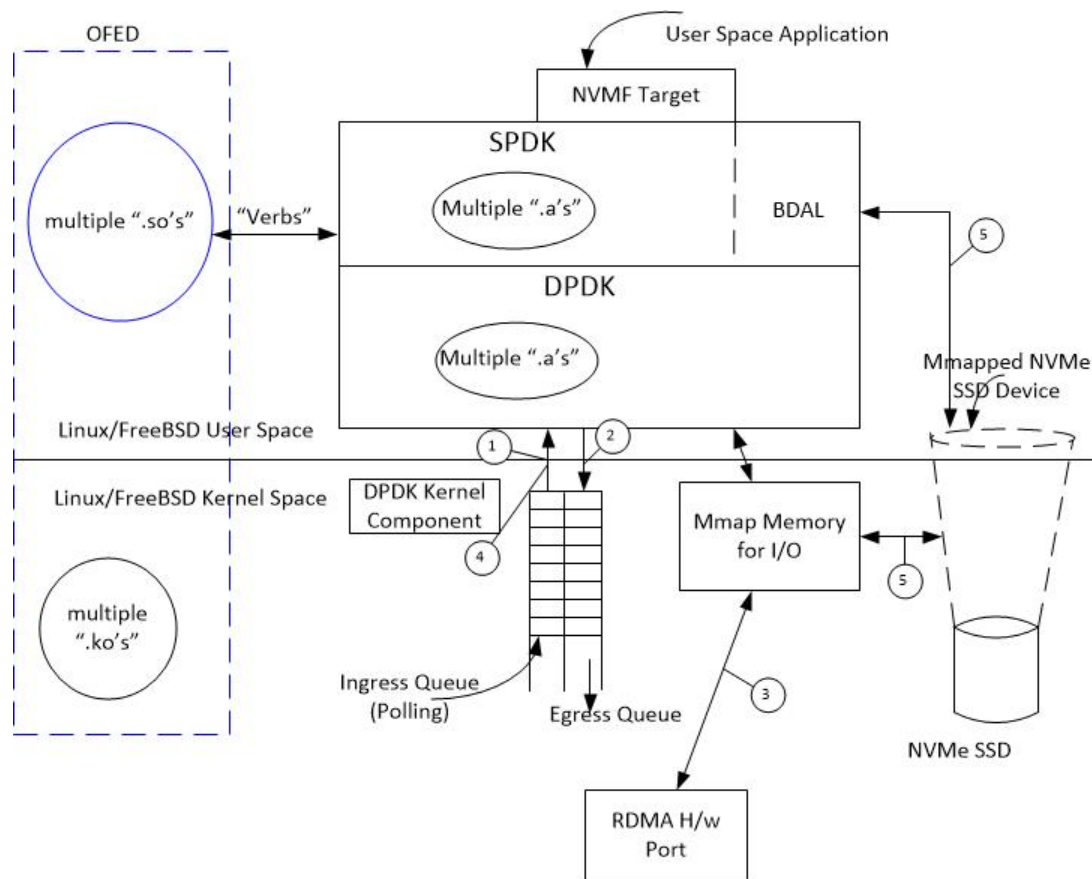
- Reuse wherever possible
 - Leverage SAN management stack
 - Leverage existing FCP driver
 - Build new flows into existing components
 - Open source components
- Partner wherever possible
 - Broadcom for FC transport layer

9 Months to deliver a new Blocks protocol in ONTAP!!

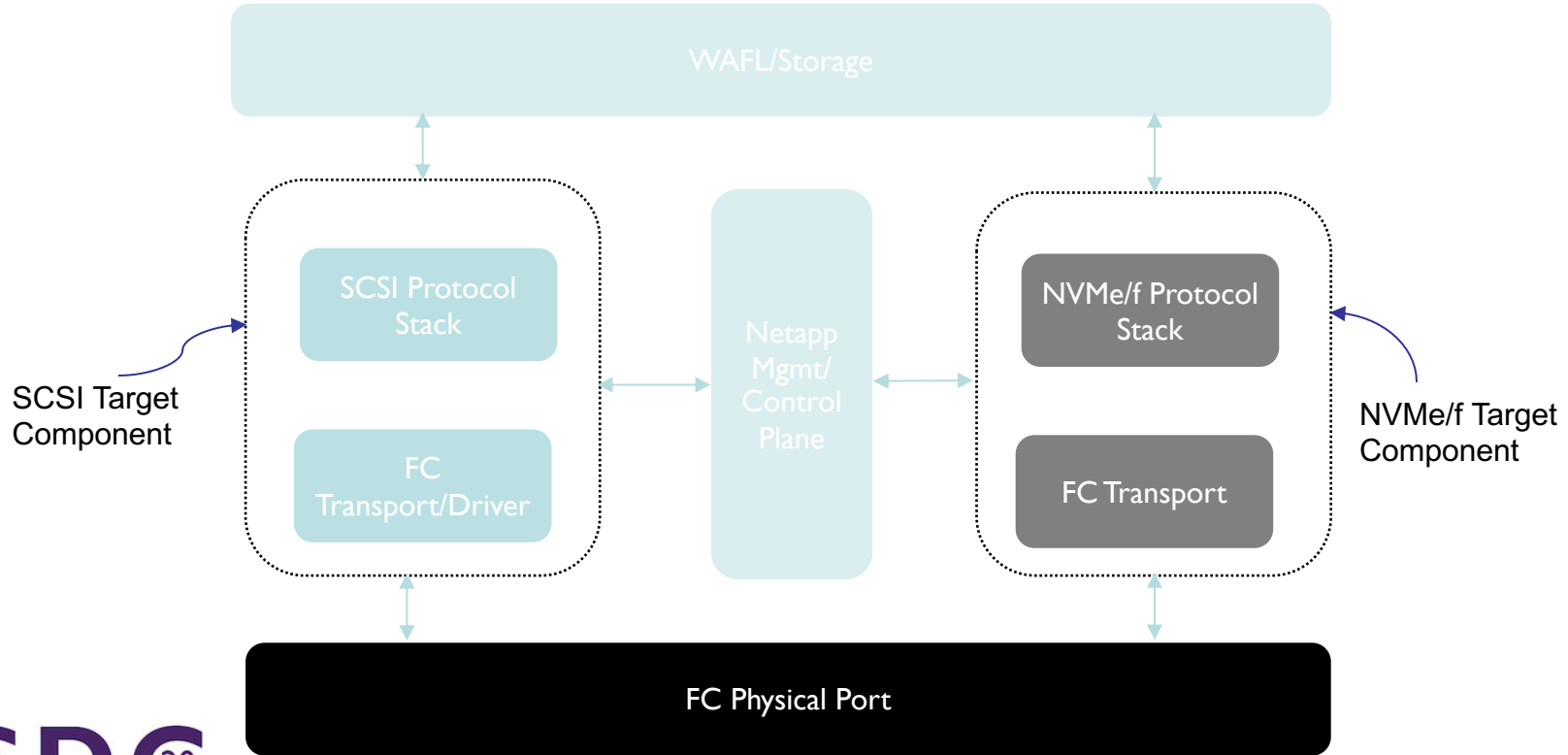
Open Source Leverage

• SPDK

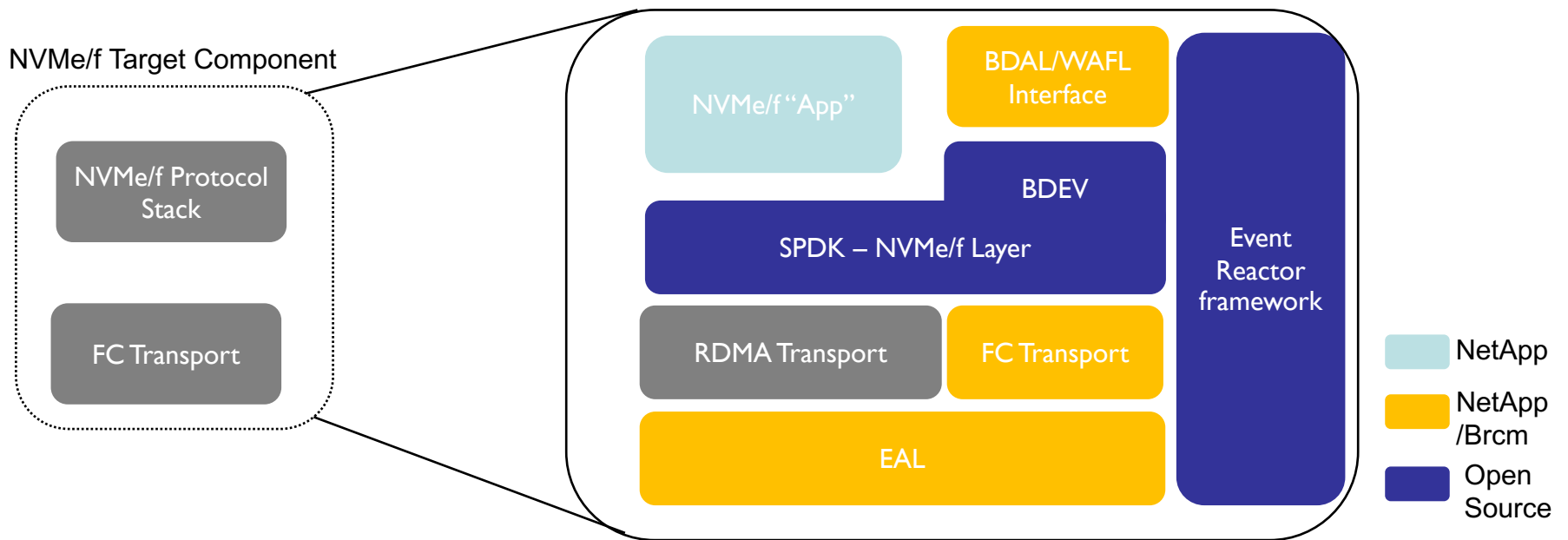
- Storage Performance Development Kit
- FreeBSD licensed software
- Intel powered
- <http://www.spdk.io/>
- Implements the NVMeoF Protocol stack
- Implements Core Design Principles
 - Poll mode driver
 - Lean and optimized code path length
 - Modular with multiple libraries
 - Static number of threads
 - RDMA transport
 - Built for user space OFED verbs
 - Used DPDK – another Intel framework



Top level Architecture

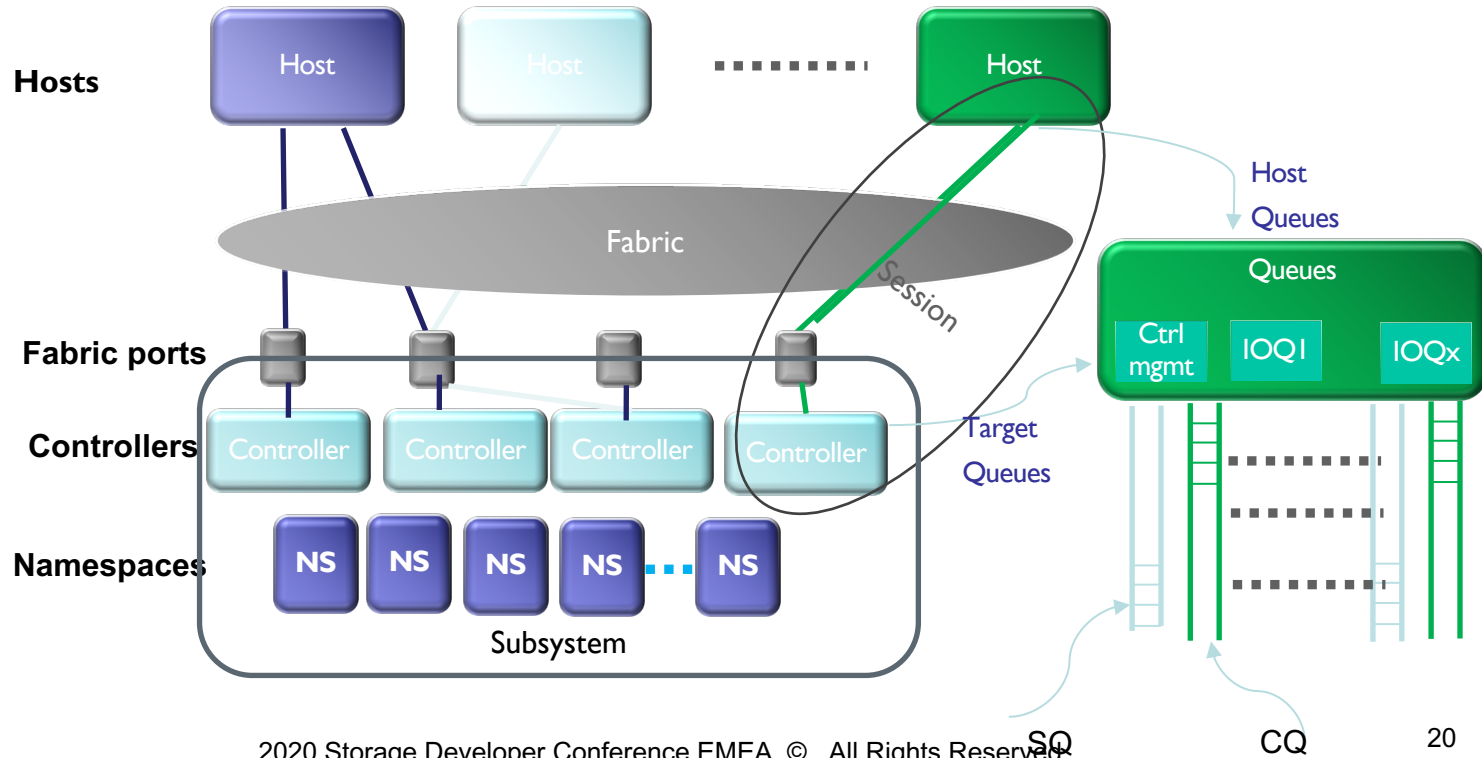


NVMe/f Target Component



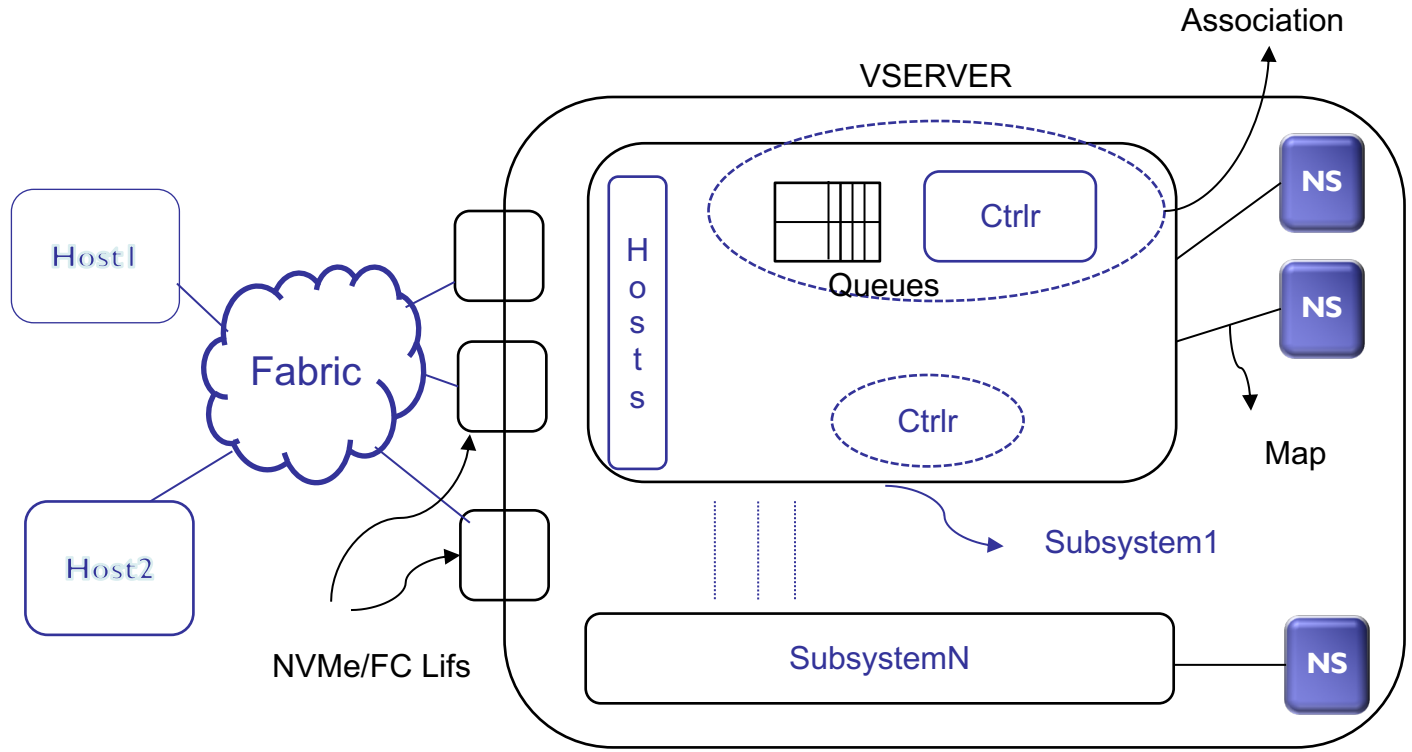
Protocol Objects

- Host
- Fabric
- Subsystem
- Namespace
- Session
- Controller
- Queues
 - SQ
 - CQ

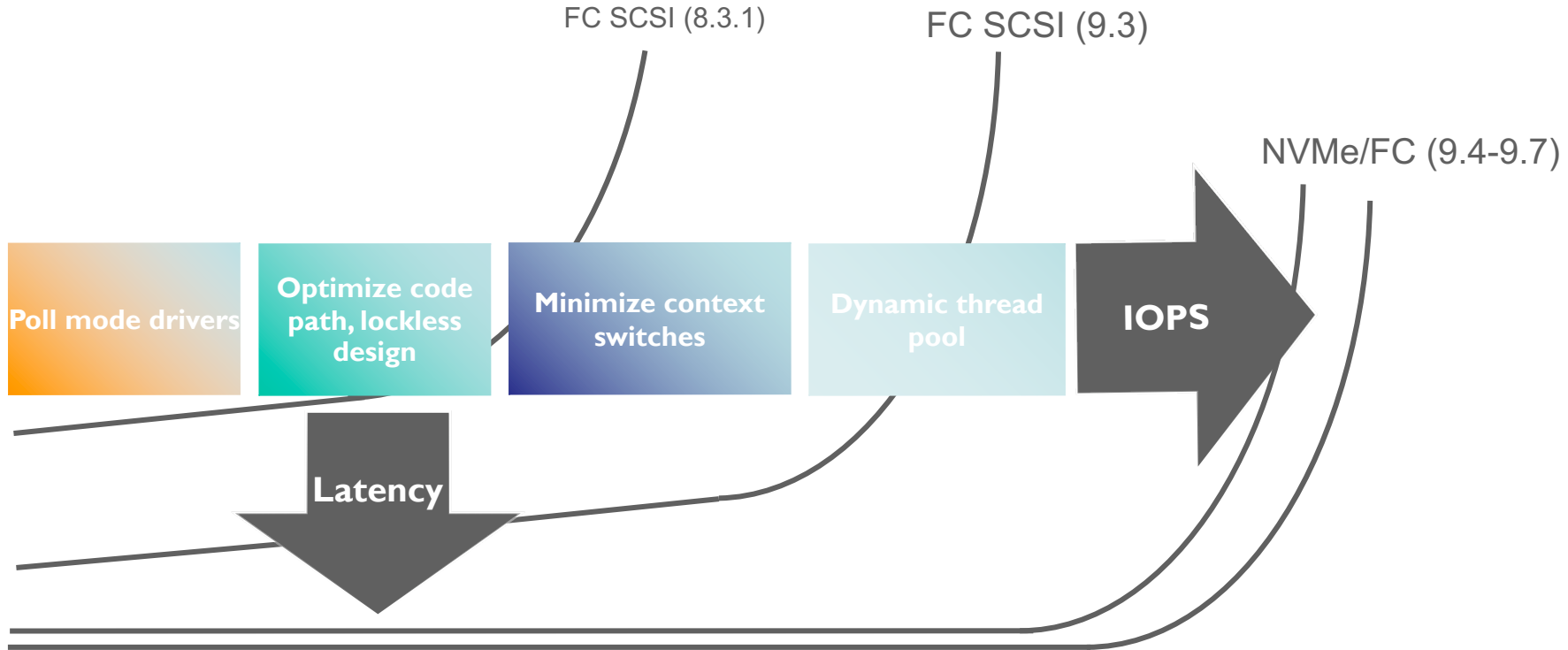


ONTAP Objects

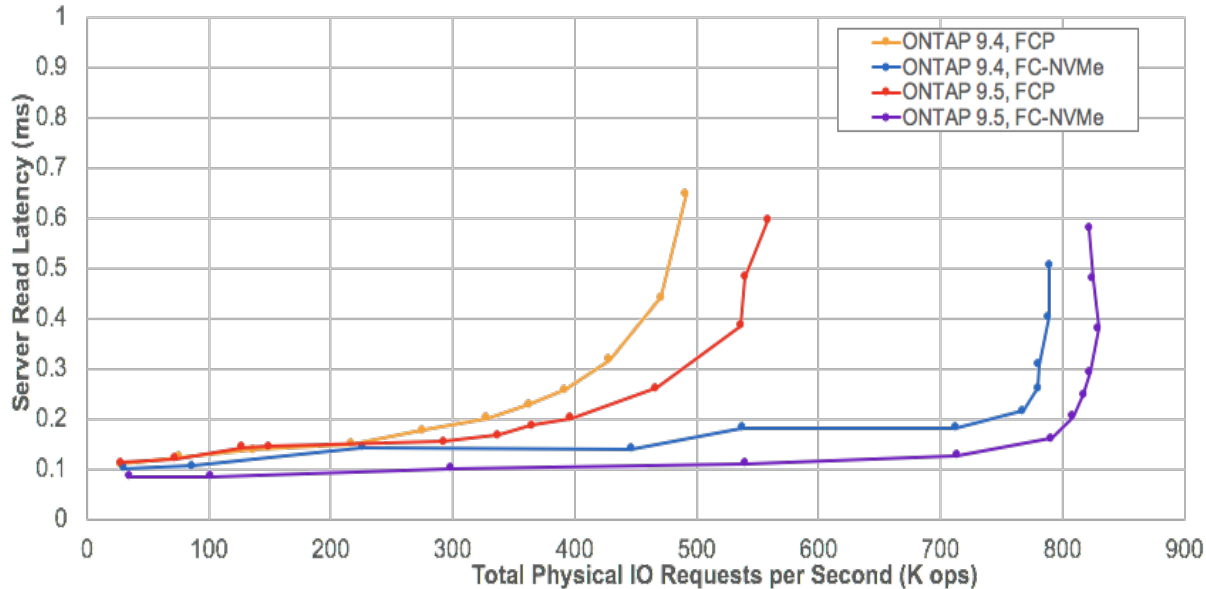
- Vserver
- Subsystems
- Namespaces
- Create
- Map
- NVMe/FC Lifs
- Hosts
- Controller



NVMe/FC Performance



Industry Leading Performance



NVMe / FC IOPS:

- Towards 1 million / controller @.5ms
- ~100us latency up to 700K IOPS

Note: Single Node A800 with 10-node SLES 12.3 Oracle RAC 75% read/ 25% write, FCP Vs. NVMe/FC tests

ONTAP NVMe/f Next Steps – Disclaimer

- ❑ Dates are not included
- ❑ Plans and Priorities are always subject to change

ONTAP NVMe/f Next Steps

Plans and priorities are always subject to change

❑ Protocol Currency

- ❑ Support for relevant TPs

❑ Support for NVMe Reservations

❑ NVMe/RoCE support

❑ NVMe/TCP support

- Support for TP-4078
 - Allows the NVMe controller to restrict or limit the number of attached namespaces.
- Support for TP-4027
 - Support Identify CNS 17h
- Support for Namespace Attach/Detach
 - Support for Identify CNS 10h (Allocated Namespace ID list)
- Support for TP-8001
 - Graceful Disconnect
- Support for TP-8005
 - Fabric SQ Flow Control
-

Future Directions

- ❑ VMware will continue to support relevant NVMe specifications in future vSphere releases
- ❑ VMware is also involved in SNIA's Swordfish specifications on NVMe-oF Management

NVMe Ecosystem @VMware

- Available as part of base ESXi image from vSphere 6.0 onwards
 - ❑ Faster innovation with async release of VMware NVMe driver
- VMware led vSphere NVMe Open Source Driver project to encourage ecosystem to innovate
 - ❑ <https://github.com/vmware/nvme>
- VMware NVMe Driver Ecosystem
<https://www.vmware.com/resources/compatibility/search.php?deviceCategory=io>
 - ❑ Close to 300 third party NVMe devices certified on VMware NVMe driver