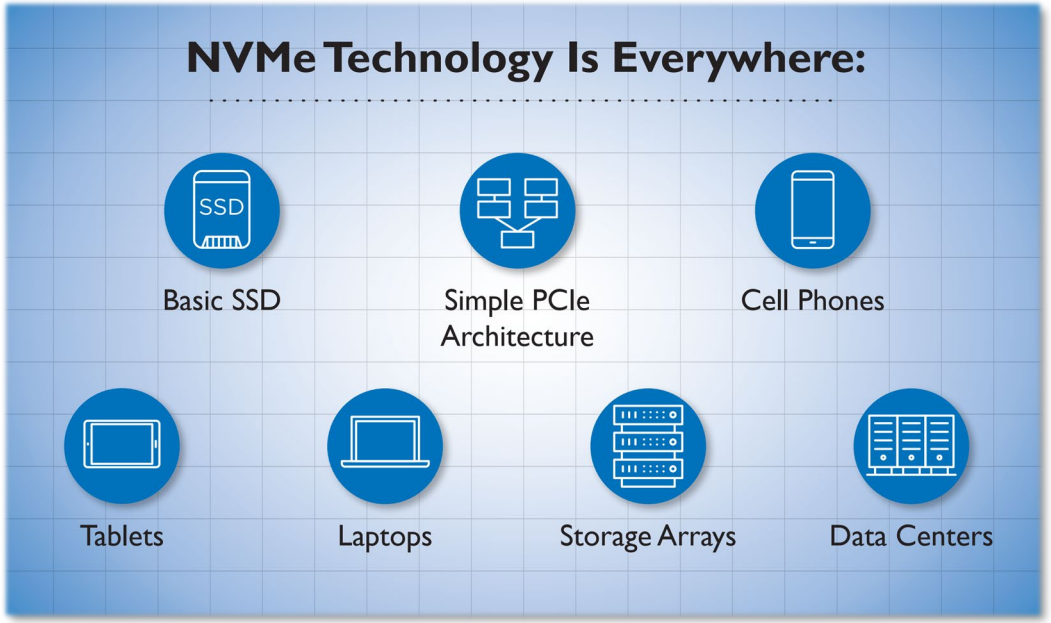
STORAGE DEVELOPER CONFERENCE
SDC 21
BY Developers FOR Developers
Virtual Conference
September 28-29, 2021

NVMe ® 2.0 Specifications: The Next Generation of NVMe Technology

Presented by Peter Onufryk

NVMe Technical Workgroup Chair

NVMe® Technology Powers the Connected Universe

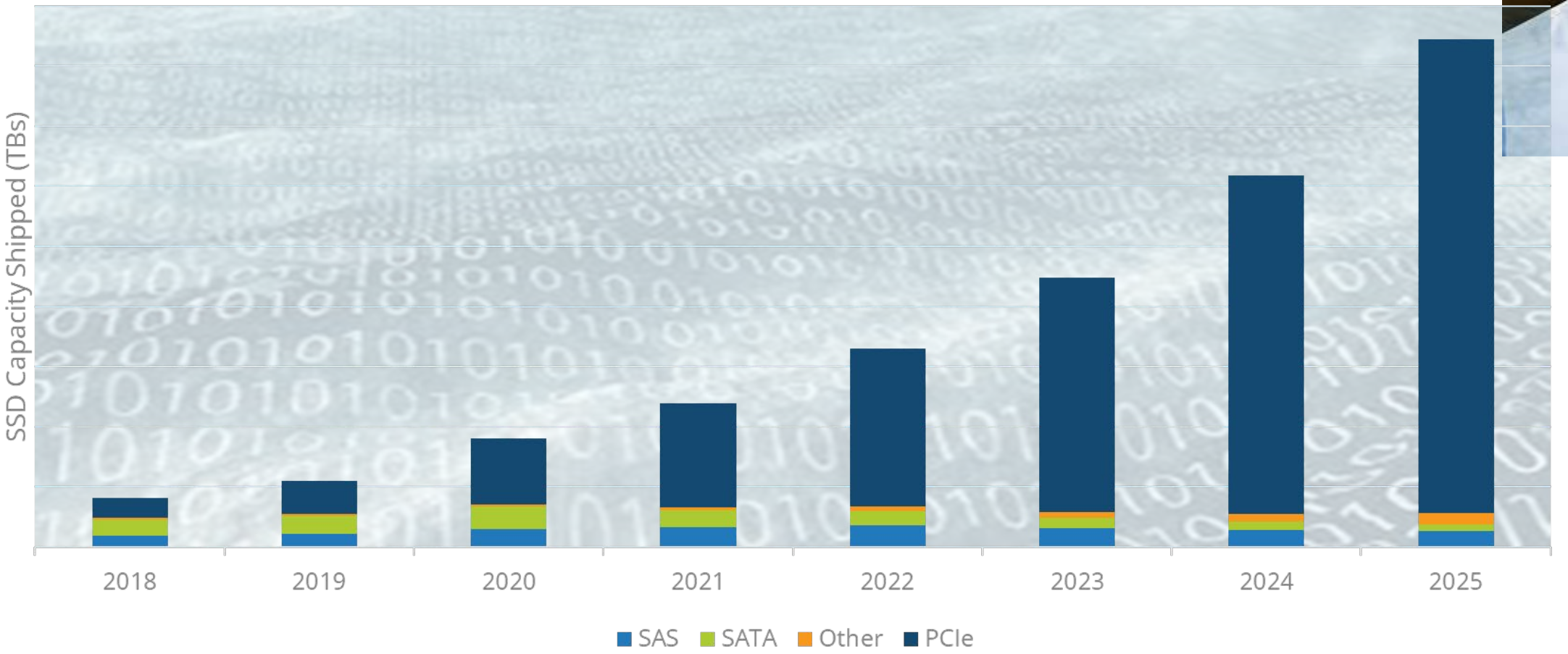


Units (Ku)	2016	2017	2018	2019	2020	2021*
Enterprise	364	749	1,069	2,045	4,910	7,290
Cloud	2,051	3,861	10,369	12,276	19,205	20,349
Client	33,128	48,951	82,587	143,236	226,221	350,253

* Data and projections provided by Forward Insights Q2'21

NVMe architecture has evolved into the new language of storage

The Leading SSD Interface



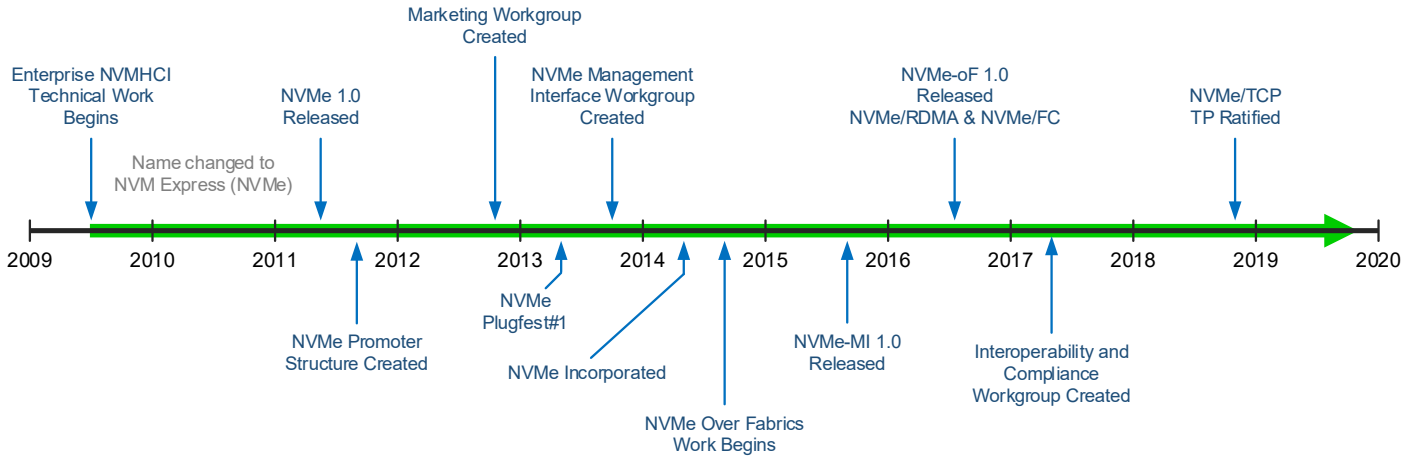
Source: Worldwide Solid State Drive 2020-2025 Forecast Update (May 2021) IDC #US46412021

Evolution of NVMe® Specifications

NVMe Management Interface Specification	
NVMe-MI 1.1	4/29/2019
NVMe-MI 1.0	11/17/2015

NVMe Base Specification	
NVMe 1.4	6/10/2019
NVMe 1.3	5/1/2017
NVMe 1.2	11/3/2014
NVMe 1.1	10/11/2012
NVMe 1.0	3/1/2011

NVMe Over Fabrics Specification	
NVMe-oF 1.1	10/22/2019
NVMe-oF 1.0	6/5/2016



The Evolution of NVMe® Technology

Focus #1

- Define NVMe architecture and NVM command set
- Unify PCIe® SSDs around a common interface
- Get an in-box driver in all major operating systems

Focus #2

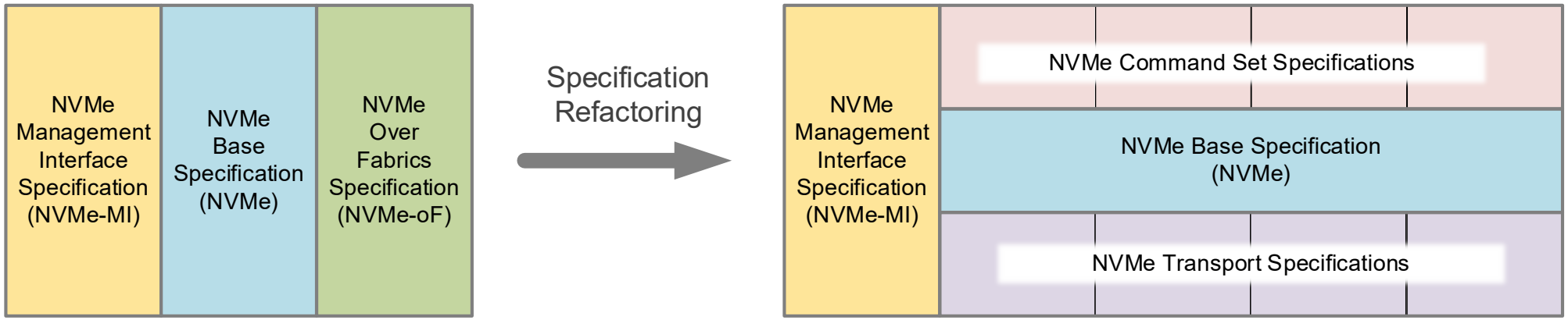
- Scale NVMe architecture and command set over arbitrary fabrics

Focus #3

- Standardize NVM enabled storage innovations
  - Spec Enhancements
  - New Command Sets
- Expand NVMe technology into new use cases (e.g., automotive, warehouse-scale storage, computational storage)

NVMe® Specification Refactoring

- Why Refactor?
  - Ease development of NVMe-based technology
  - Enable rapid innovation while minimizing impact to broadly deployed solutions
  - Create extensible spec infrastructure that enables the next phase of growth for NVMe technology



NVMe® 2.0 Family of Specifications

NVMe Base Specification

Command Set Specifications	Transport Specifications
NVMe NVM Command Set Specification	NVMe over PCIe Transport Specification
NVMe Zoned Namespace Command Set Specification	NVMe over RDMA Transport Specification
NVMe Key Value Command Set Specification	NVMe over TCP Transport Specification

NVMe Management Interface Specification

NVMe® Multiple I/O Command Set History

NVMe 1.0

CAP.CSS

Bit	Definition
37	NVM command set
38	Reserved
39	Reserved
40	Reserved

CC.CSS

Value	Definition
000b	NVM command set
001b To 111b	Reserved

Up to 4 I/O Command Sets

NVMe 1.1

CAP.CSS

Bit	Definition
37	NVM command set
38	Reserved
39	Reserved
40	Reserved
41	Reserved
42	Reserved
43	Reserved
44	Reserved

CC.CSS

Value	Definition
000b	NVM command set
001b To 111b	Reserved

Up to 8 I/O Command Sets

NVMe 1.4

CAP.CSS

Bit	Definition
37	NVM command set
38	Reserved
39	Reserved
40	Reserved
41	Reserved
42	Reserved
43	Reserved
44	No I/O Command Set

CC.CSS

Value	Definition
000b	NVM command set
001b To 110b	Reserved
111b	Admin Cmd Set Only

Up to 7 I/O Command Sets

TP 4056 Namespace Types

CAP.CSS

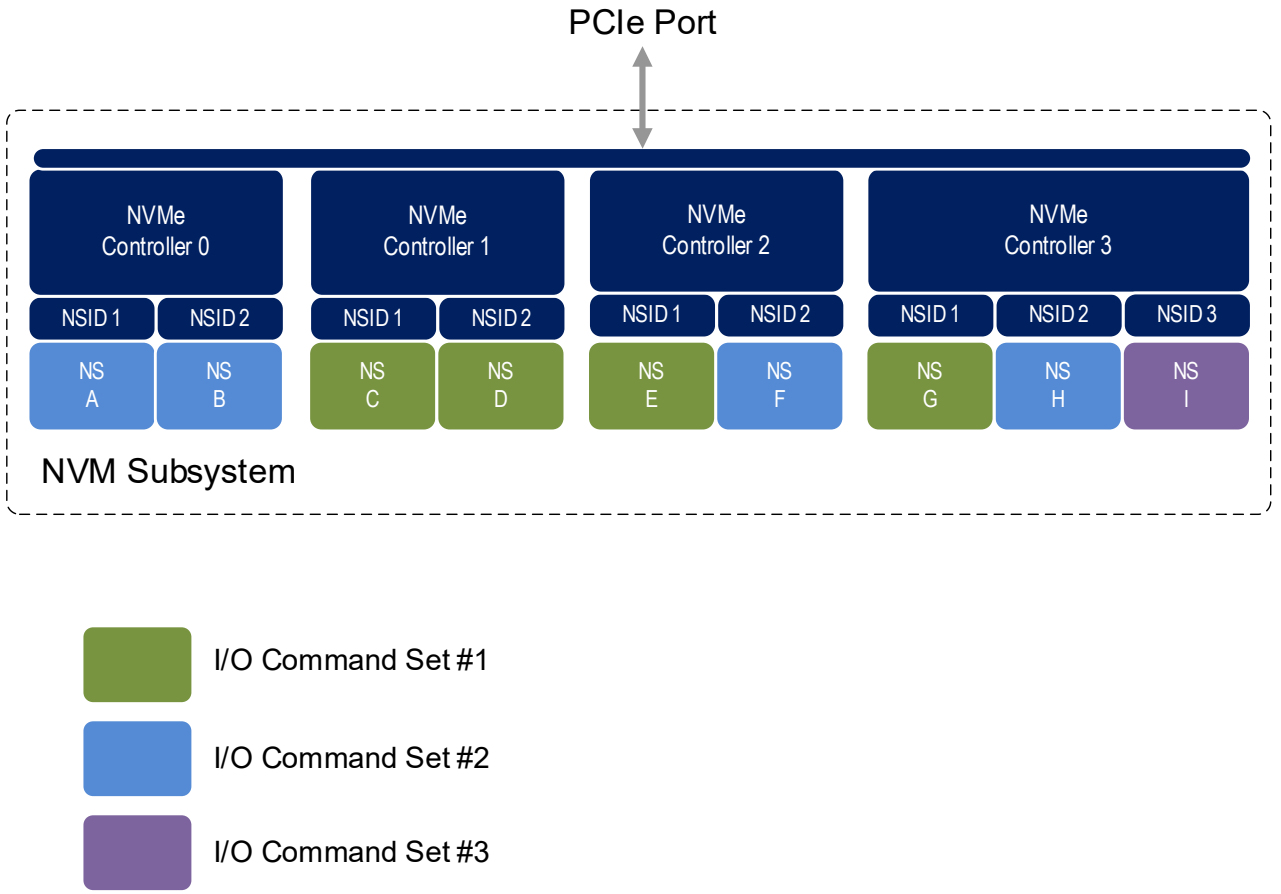
Bit	Definition
37	NVM command set
38	Reserved
39	Reserved
40	Reserved
41	Reserved
42	Reserved
43	Identify I/O Cmd Set
44	No I/O Command Set

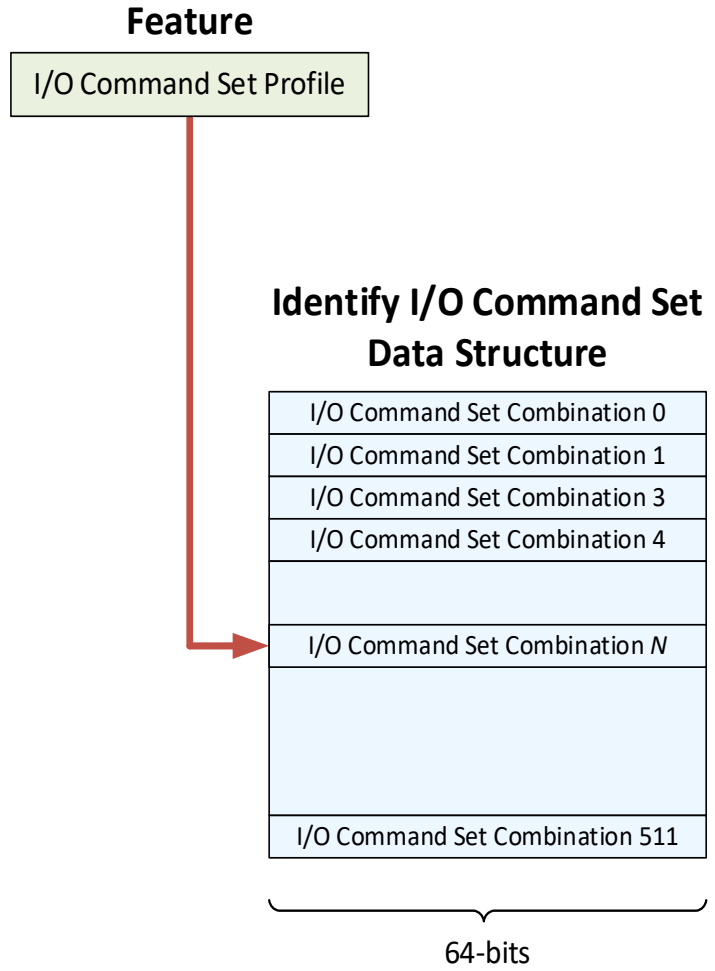
CC.CSS

Value	Definition
000b	NVM command set
001b To 101b	Reserved
110b	Identify I/O Cmd Set
111b	Admin Cmd Set Only

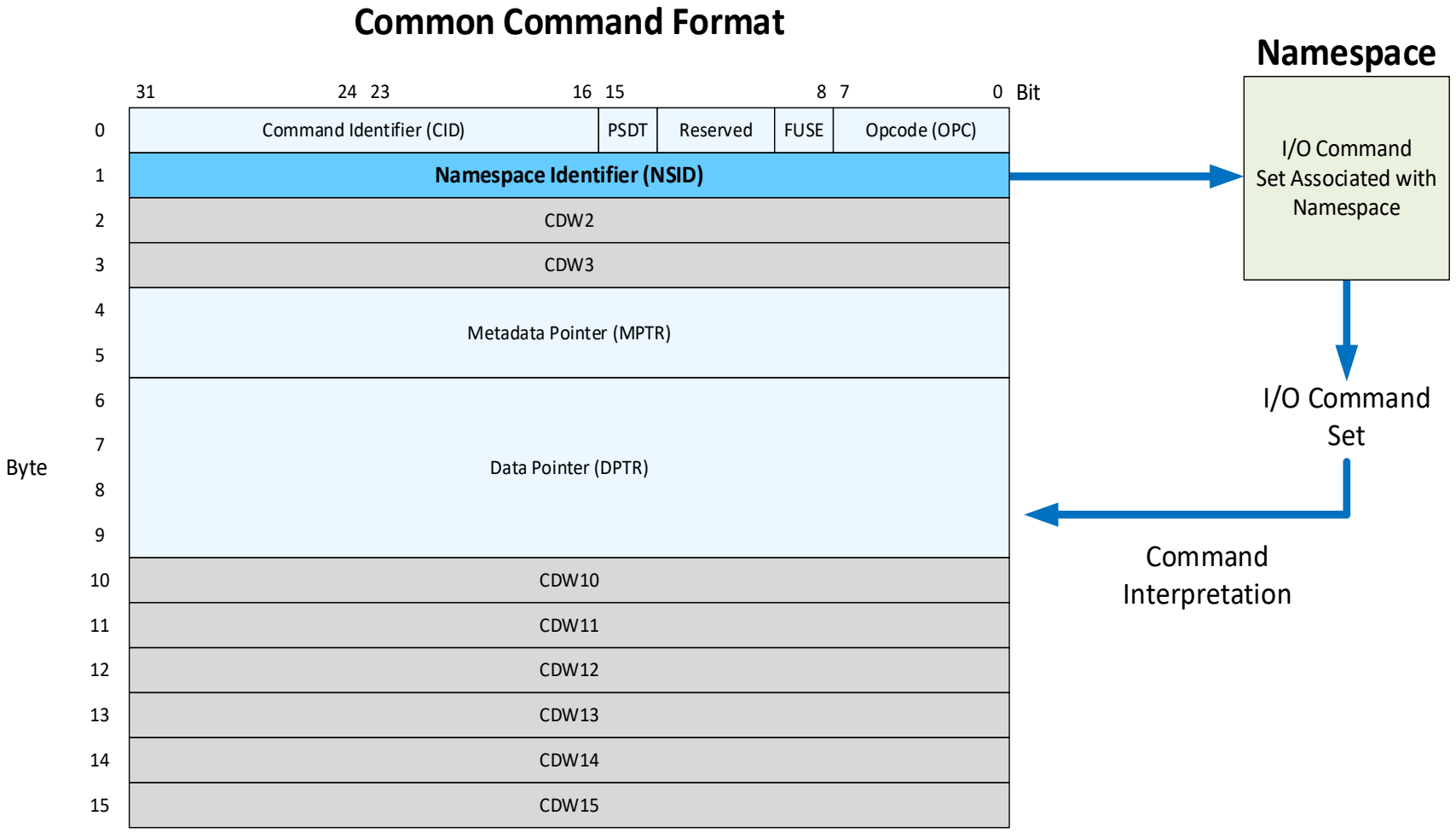
Up to 64 I/O Command Sets

Enabling Multiple Command Sets

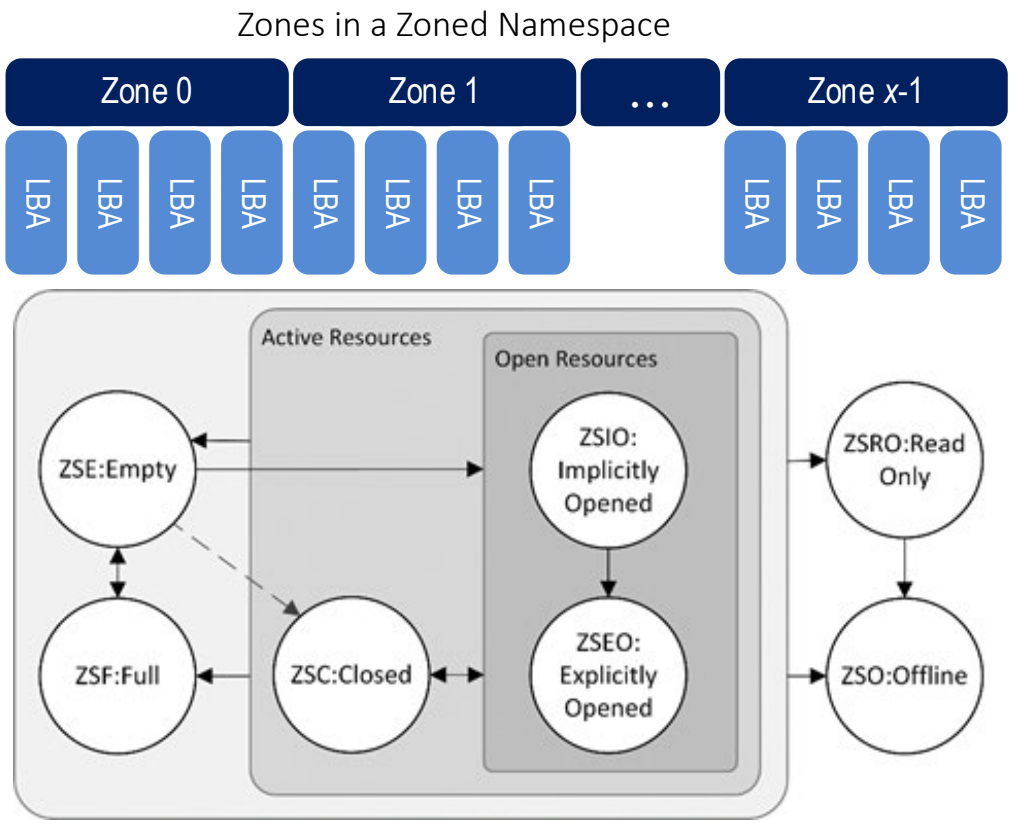




I/O Command Interpretation

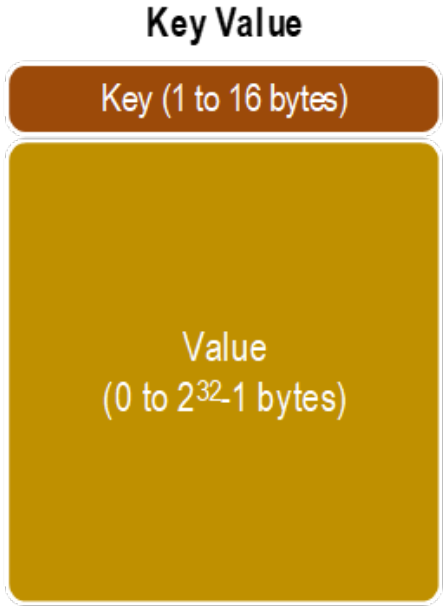


Zoned Namespaces Command Set



- Logical blocks are grouped into zones
  - Logical blocks are written sequentially within a zone
- State machine associated with each zone
  - Controls operational characteristics of each zone
  - State transitions may be explicitly controlled by the host or implicitly by host actions
- Benefits
  - Reduced write amplification
  - Reduced overprovisioning
  - Reduced memory on Storage Device (DRAM)

Key Value Command Set

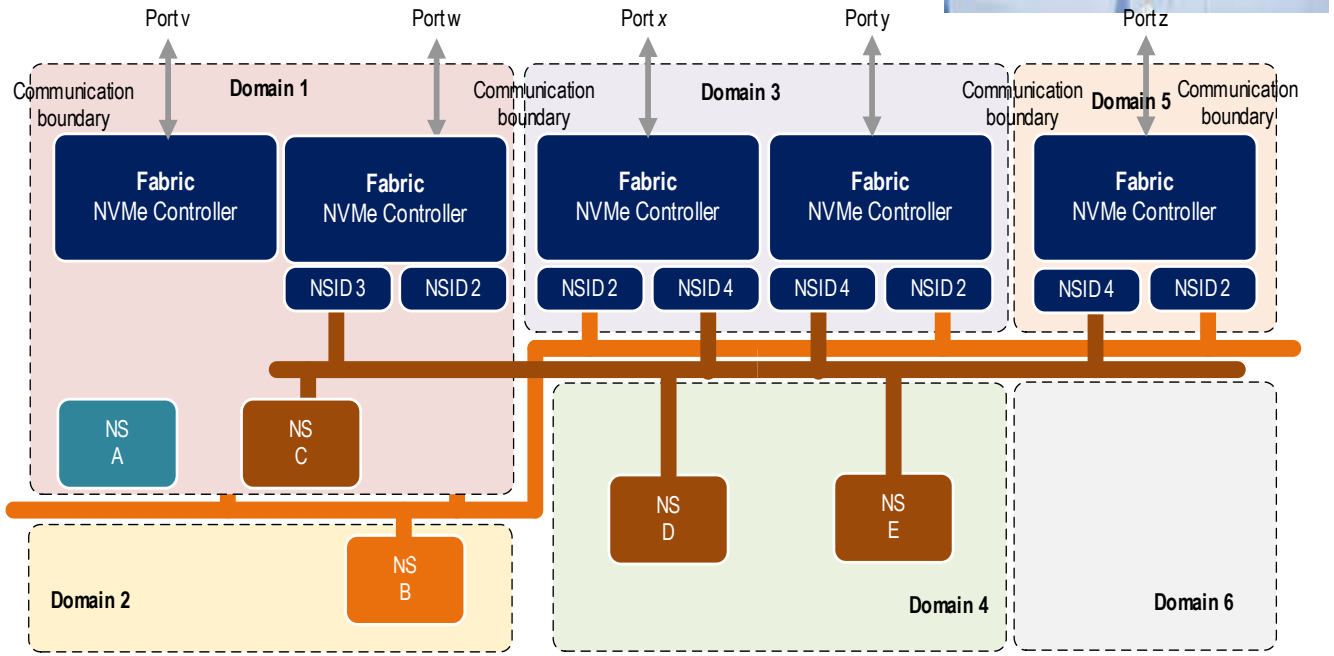


Command	Description
Delete	Delete Key and Value associated with a specified Key
List	Lists Keys that exist in a Key Value Namespace starting at a specified Key
Retrieve	Retrieve Value associated with a specified Key
Exist	Returns status indicating whether a Key Value exists for a specified Key
Store	Stores a Key Value to a Key Value Namespace

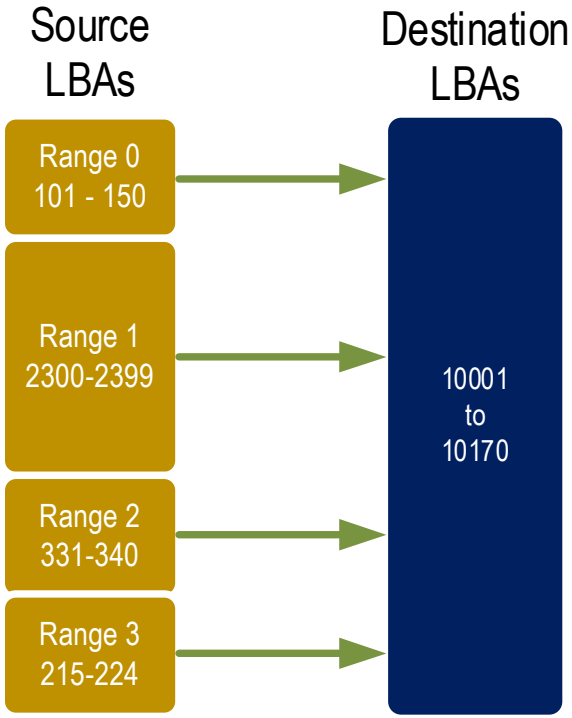
- Command set optimized for unstructured data

Domains and Partitions

- An NVM subsystem may represent a warehouse-scale storage system
- A warehouse-scale storage system may be constructed from multiple Domains
  - Capacity, controllers, and ports, may be partitioned among Domains
  - Domains may be added, removed, reconfigured, partitioned, or fail
- NVMe® technology now defines Domains as an architectural element



Copy Command



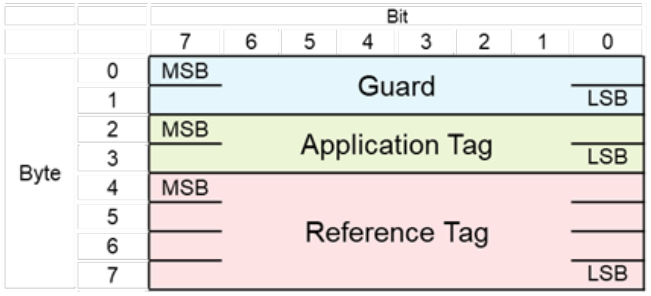
- New NVM I/O command that copies logical blocks from one or more logical block ranges to a single contiguous destination logical block range
  - Source logical block ranges described by Source Range Entries transferred from host
  - Supports protection information

Command Group Control

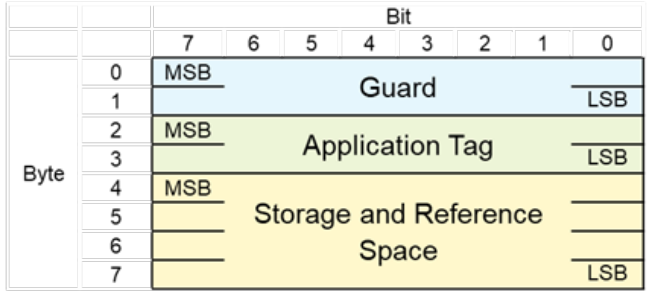
- Defines new Lockdown admin command
  - May be used to prohibit execution of a command or modification of a feature in an NVM subsystem
    - Admin command
    - Set Feature for a specified Feature Identifier
    - Management Interface Command Set command
    - PCIe Command Set command
  - Provides interface level granularity
    - Ability to lockdown in-band, out-of-band, or both

- Once a command or feature is locked down, then it remains locked down until re-enabled by the Lockdown command or NVM subsystem power cycle

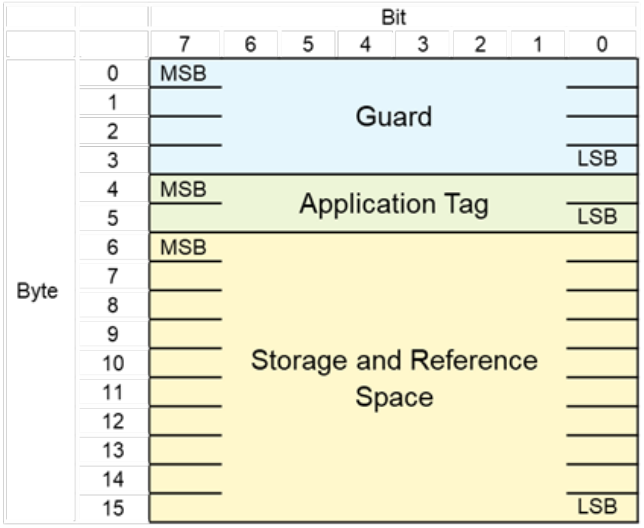
Protection Information Enhancement



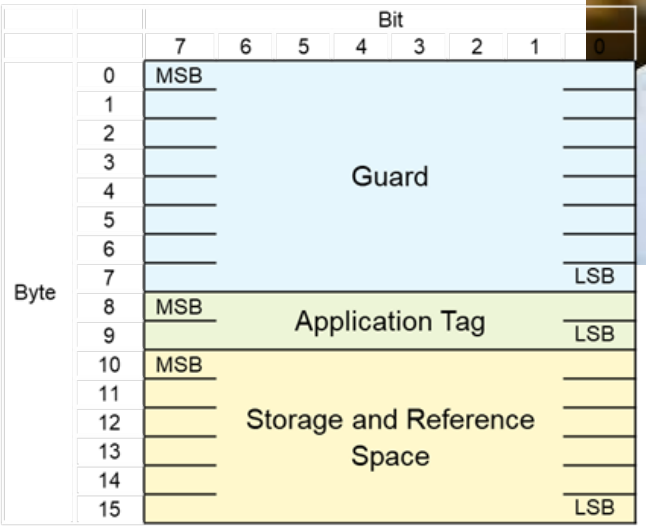
16-bit Guard (Legacy)



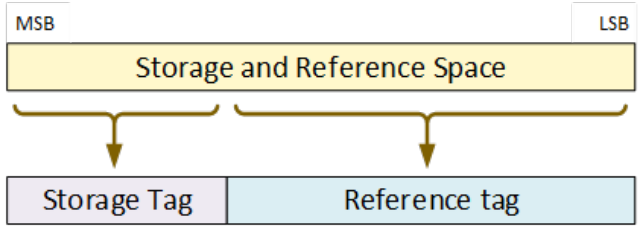
16-bit Guard (New)



32-bit Guard (New)

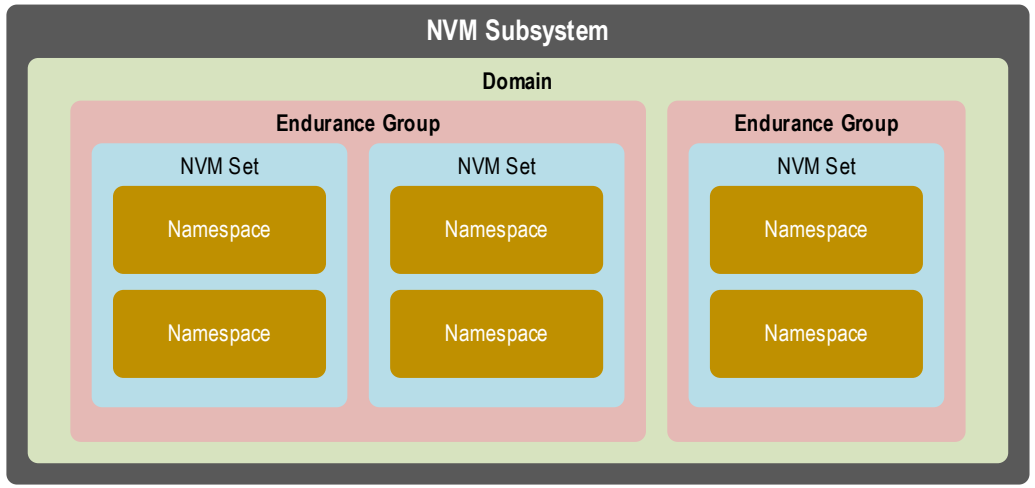


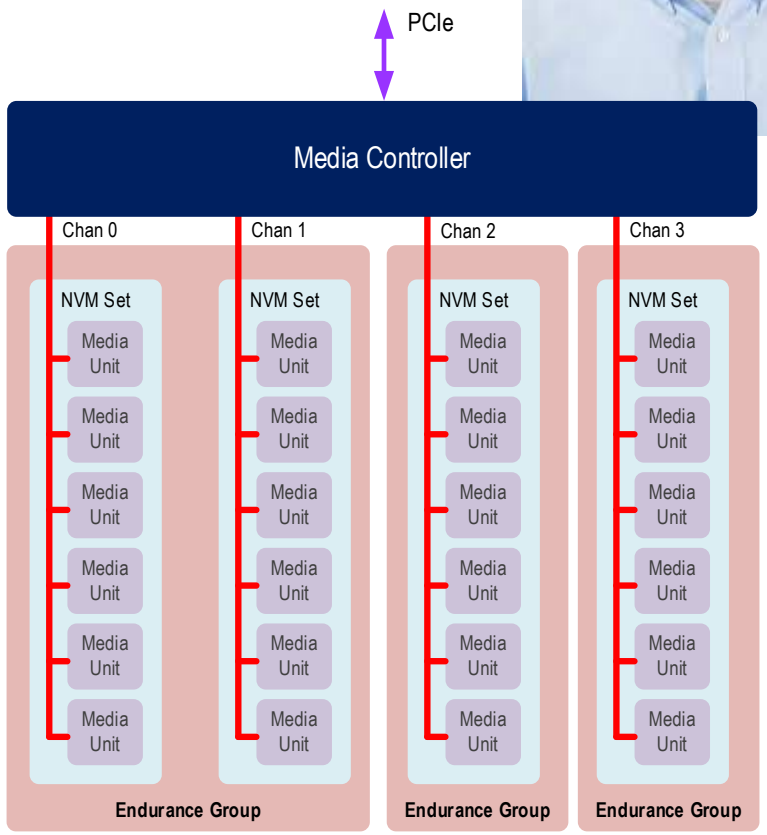
64-bit Guard (New)



Endurance Group Management

- Defines new Capacity Management admin command
  - Creation/deletion of NVM Sets
  - Creation/deletion of Endurance Groups
  - Allocation of Media Units to Endurance Groups
  - Allocation of Media Units to NVM Sets





NVMe® over Fabrics Security Features

- NVMe/TCP technology implementations that support TLS are now required to support TLS 1.3
  - TLS 1.2 is still supported, but support provisions have not been carried forward to the NVMe 2.0 family of specifications
- New mutual host and NVM subsystem in-band authentication protocol based on Diffie-Hellman HMAC-CHAP (DH-HMAC-CHAP)

Rotational Media Support

- NVMe® specifications add support for rotational media (e.g., HDDs)
- Enhancements
  - Indicator that Namespace stores data on rotational media
  - Log page that provides rotational media information
    - Number of actuators
    - Nominal rotational speed
    - Spinup count & failed spinup count
    - Load count & failed load count
  - Spinup control

Summary

- NVMe® technology has become the new language of storage
  - NVMe technology has unified client, cloud, and enterprise storage around a common command set and architecture
  - NVMe specifications now support every major storage interconnect
  - NVMe technology has moved into a third focus of development – storage innovations and new use cases
- The NVMe 2.0 family of specifications introduce new spec refactoring and features
- The NVMe technical community is accelerating technical development
  - Maintaining existing specifications
  - Enhancing existing NVMe features and capabilities
  - Delivering new innovations

nvm EXPRESS®

Architected for Performance

Please take a moment to rate this session.

Your feedback is important to us.

