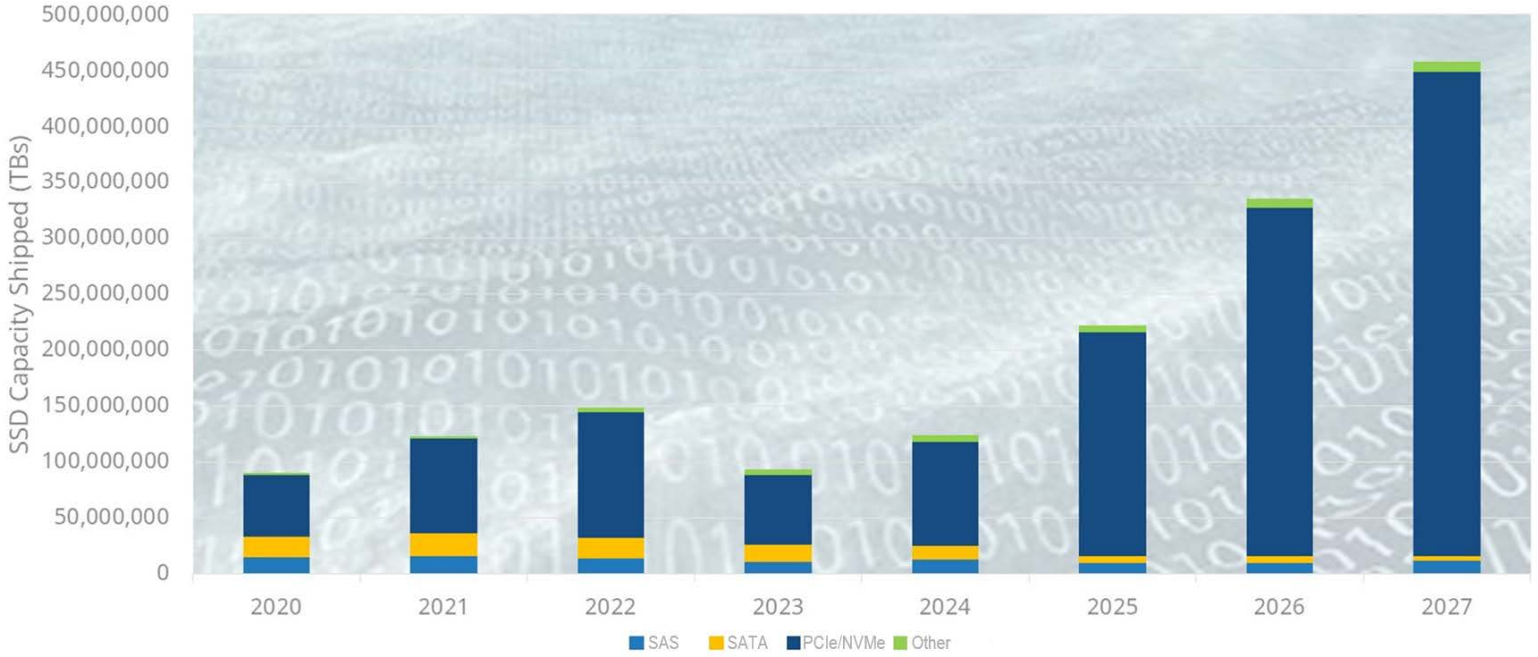
NVM Express State of the Union

Sponsored by NVM Express organization, the owner of NVMe® Specifications

Speakers

Phil Cayton

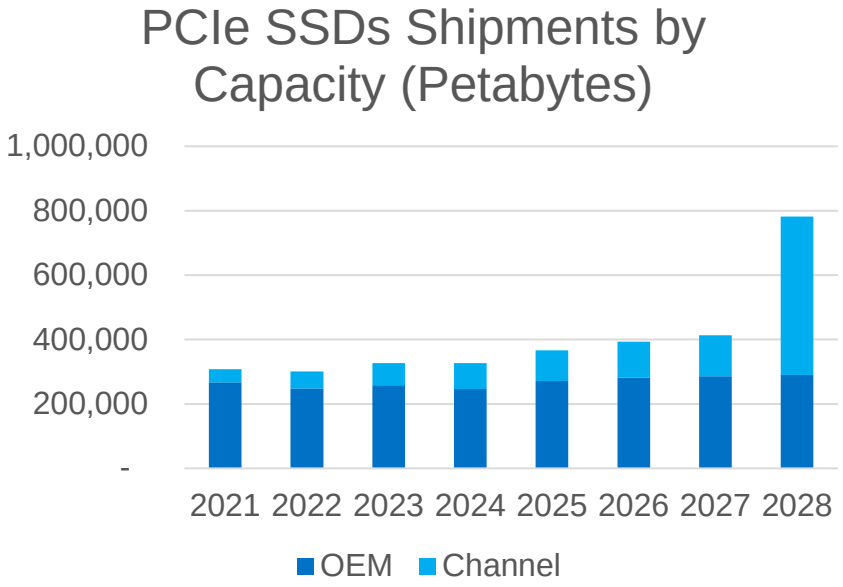
Enterprise SSD Capacity Shipment by Interface

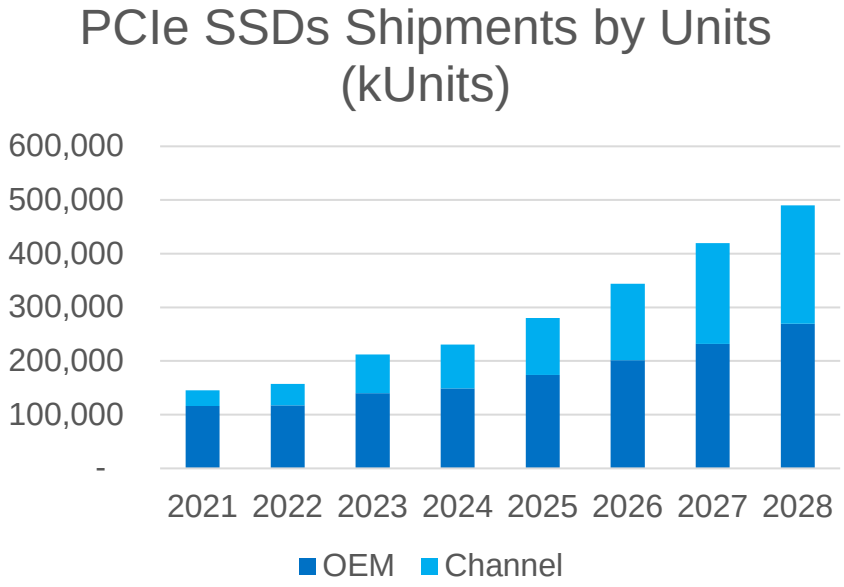


SAS
SATA
PCIe
other

Source*: IDC Worldwide Solid-State Drive Forecast, 2023-2027 Doc # US50021623 , Dec 2023*

Consumer SSD Shipments by Capacity and Units



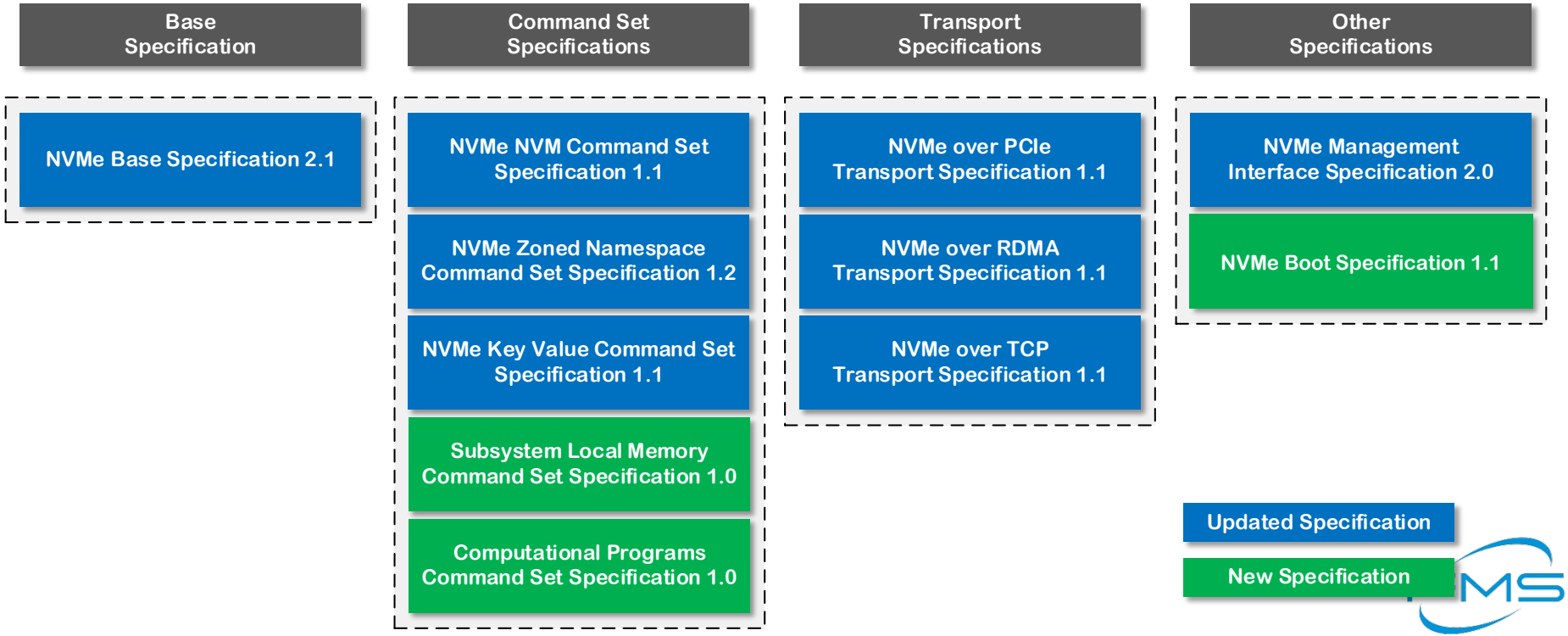


Source: Data and projections provided by Forward Insights Q2 2024

Activity Since Release of NVMe® 2.0 Specifications

New Authorized Technical Proposals	Ratified Technical Proposals	Ratified ECNs
71	82	22

NVMe® Specifications



3 New Specification and 8 Updated Specifications

New NVMe® Features and Enhancements

Fabrics Enhancements

- NVMe Boot (NVMe Boot Specification, TP8012, TP4126, TP8027, TP8029)
- Discovery Controllers Enhancements (TP8009, TP8010, TP8013)
- Fabric Zoning (TP8016)

Emerging Technologies

- Computational Storage (Computational Programs Command Set)
- Flexible Data Placement (TP4146)
- NVM Subsystem Memory (Subsystem Local Memory Command Set)

Security Enhancements

- Key per I/O (TP4055)
- Post Sanitize Media Verification (TP4152)
- Fabrics Security Enhancement including Support for TLS 1.3 (TP8019, TP8025, TP8018)

Data and VM Migration

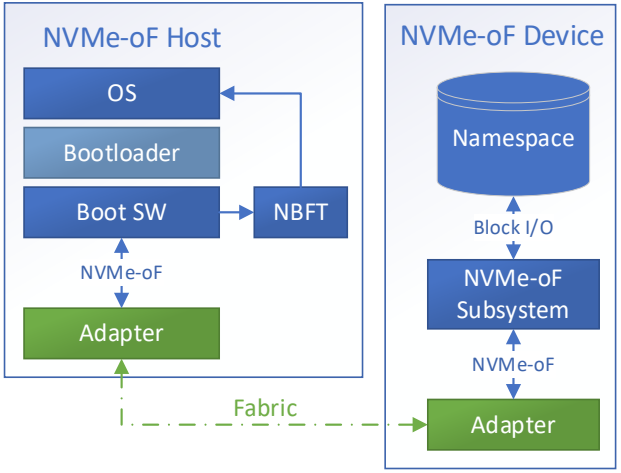
- PCIe Infrastructure for Live Migration (TP4159)
- Cross Namespace Copy (TP4130)
- Tracking LBA Allocation with Granularity (TP4165)

Management

- Scalable Resource Management (TP6011)
- High Availability Out-of-Band Management (TP6034)
- Out-of-Band Management Asynchronous Events (TP6035)
- Management Support for I3C (TP6037)

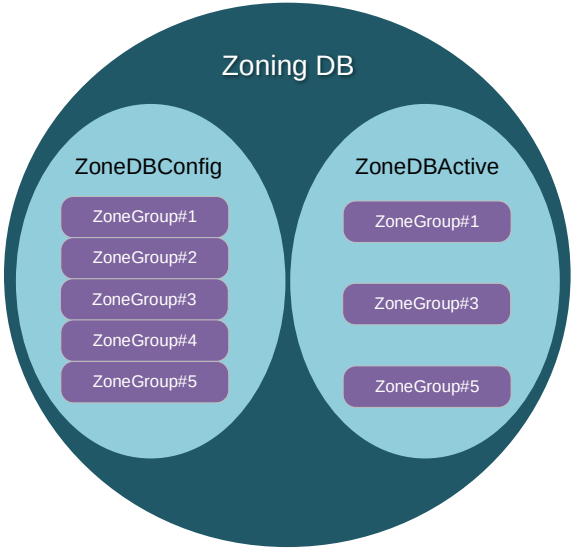
NVMe® Boot Specification

- New NVM Express® Boot Specification
  - The specification defines construct and guidelines for booting from NVMe technology
  - While the specification covers all transports, the current specification only describes mechanisms for NVMe/TCP technology
  - Recent enhancements
    - Standardized the format of UUIDs displayed in an NBFT
    - Added support for IPV4/IPV6 DHCP Identifiers
    - Improved error codes for common Subsystem Connection failures



Discovery Enhancements

- NVMe-oF™ Automated Discovery
  - Simplifies provisioning of Hosts by allowing them to locate NVMe®/TCP Discovery controllers
- Centralized Discovery Controller
  - Enable discovery information to be consolidated and retrievable from a single Discovery Service
  - **Centralized Discovery Controller (CDC)**: a Discovery controller that reports discovery information registered by Direct Discovery Controllers and hosts
  - **Direct Discovery Controller (DDC)**: a Discovery controller capable or registering discovery information with a CDC
- Fabric Zoning
  - A ZoneGroup is a set of access control rules enforced by the CDC
  - Members of the same Zone are allowed to communicate between each other
  - A Direct Discovery Controller may provide Fabric Zoning formation to a CDC using push or pull registrations



Key Per I/O

- Self encrypting drives perform encryption on LBA ranges within namespaces
- Key per I/O provides dynamic fine grain encryption control by indicating which encryption key to use per I/O
  - Assigning an encryption key to a sensitive file or host object
  - Easier support of General Data Protection Regulation (GDPR)
  - Easier support of erasure when data is spread and mixed with other data that should be preserved (e.g., RAID and erasure coding)
- Mechanisms to download and manage keys are outside the scope of the specification
  - Keys are stored in volatile memory and are lost when powered off
- Supported by both NVM Express and the Trusted Computing Group (TCG)

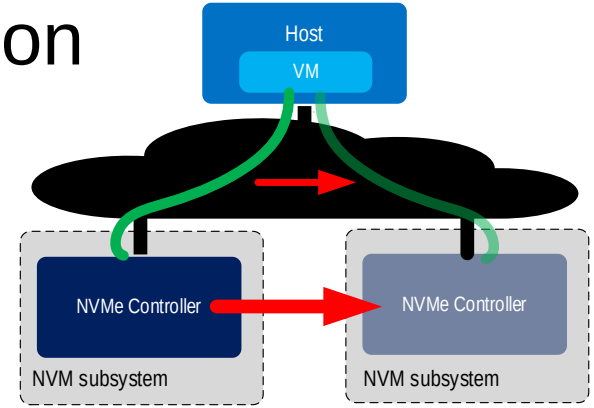
Fabrics Security Enhancements

- NVMe®/TCP Transport Layer Security (TLS) updates redesign how TLS is used
  - Improves security
  - Removes key identification ambiguities
  - Strongly discourages TLS 1.2 in favor of TLS 1.3 for improved security
- Centralized authentication verification entity for DH-HMAC-CHAP
  - Removes need for widespread deployment of pre-shared keys for authentication verification
  - Each PSK is deployed in exactly two places
    - The entity (host or NVM subsystem) that uses the PSK to authenticate
    - The centralized AVE that verifies authentications.
- Added NVMe-oF™ Security configurations for consistent behavior
  - based on the configuration individual NVMe-oF Hosts or NVM Subsystems

For more information refer to https://nvmexpress.org/wp-content/uploads/NVMe-oF-Security-Updates-Webinar-Slides.pptx

PCIe® Infrastructure for Live Migration

- Provides building blocks that allow a system to be constructed that enables a host to migrate an NVMe® controller from one NVM subsystem to another



Pre-Copy Phase Host Actions

- Requests the controller track LBA changes (dirty LBAs) of attached namespaces
- Migrate the allocated LBAs of attached namespaces
- Migrate the dirty LBAs

Stop-and-Copy Phase Host Actions

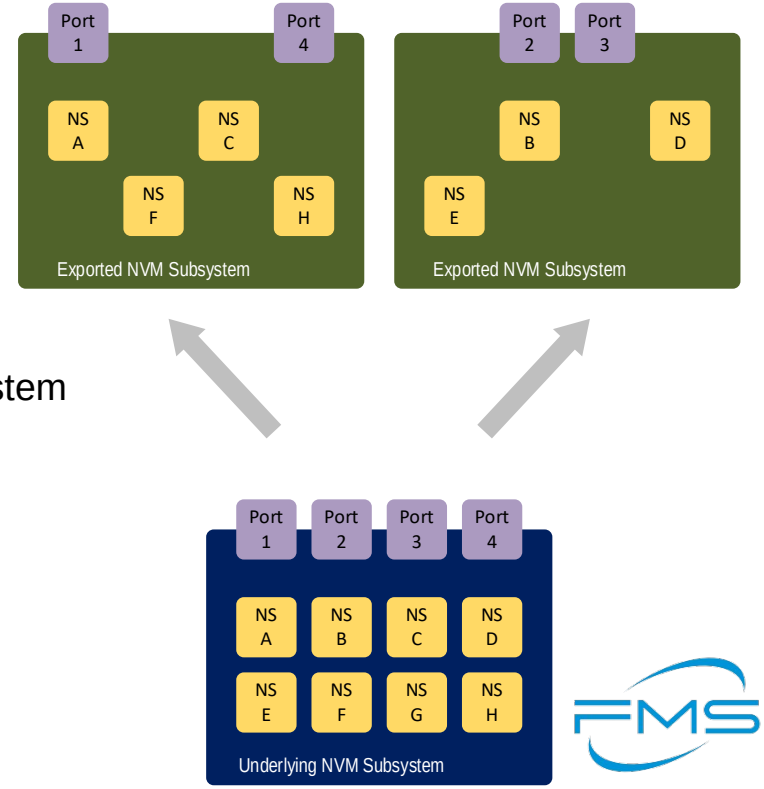
- Requests the controller pause causing all commands to be completed
- Migrate any remaining dirty LBAs

Post-Copy Phase Host Actions

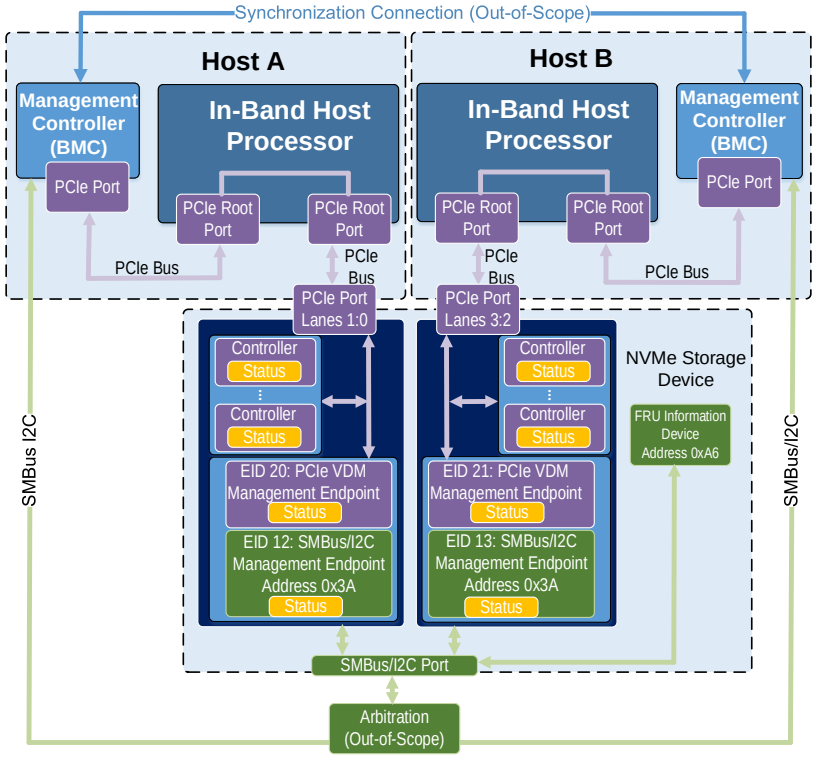
- Migrate controller state
- Resume the migrated controller

Scalable Resource Management

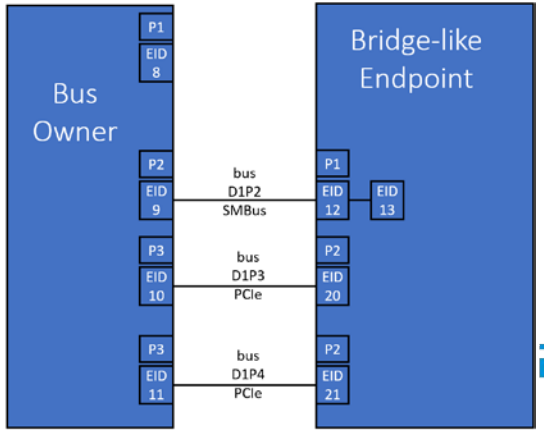
- Defines a standard framework to dynamically construct, configure, and provision “Exported” NVM Subsystems from underlying physical resources in “Underlying” NVM subsystems
- New Admin Commands that enable
  - Creation and management of an Exported NVM Subsystem
  - Manage Exported Namespaces
  - Manage Exported Ports
- Ability to manage host access to an Exported NVM Subsystem using an “Allowed Host List”



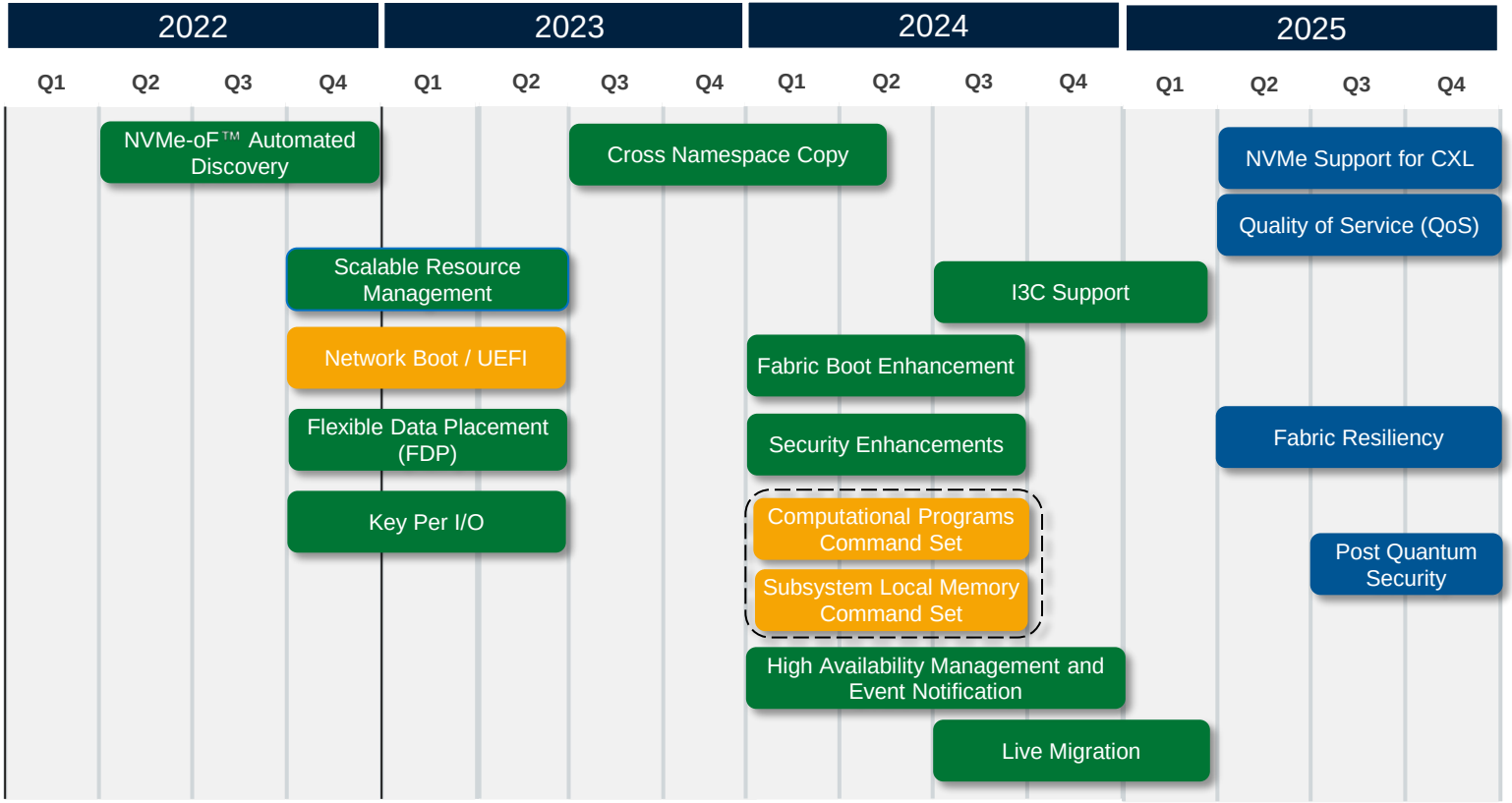
NVMe-MI™ High Availability Out of Band Management



- Management Component Transfer Protocol (MCTP) bridging on endpoints to allows multiple MCTP endpoints per physical address
- Unique instance of status/state per MCTP endpoint



NVMe® Architecture Planned Feature Roadmap



Ratified Feature
(left edge indicates ratification quarter)

Planned Feature
(left edge indicates planned ratification quarter)

Ratified New Specification
(left edge indicates ratification quarter)

Summary

- NVMe® technology adoption continues to grow and has succeeded in unifying client, cloud, AI, and enterprise storage around a common architecture
- Systems are being built using NVMe architecture as the native underlying storage technology:
  - cell phones, tablets, laptops, desktops, storage arrays, data centers, automotives, …
- A dedicated NVMe technical community maintains existing specifications while developing new innovations
  - 75 new Technical Proposals authorized
  - 80 Technical Proposals ratified
  - 22 ratified ECNs
- NVMe began as short and simple PCIe® SSD specification and has grown into nearly a dozen specifications supporting all major transports, multiple command sets, and standardizing many aspects of storage
  - Technical Proposals are made public when ratified
  - Specification updates simply aggregate Technical Proposals ratified since the last specification update


nvm
EXPRESS®
Architected for Performance

Consumer SSD Shipments by Capacity and Units

Petabytes*								
	2021	2022	2023	2024	2025	2026	2027	2028
OEM	267,126	248,099	257,815	246,481	271,113	281,928	286,180	289,727
Channel	40,789	52,646	68,685	79,988	95,892	111,249	127,292	132,097
Total	**307,915**	**300,745**	**326,500**	**326,469**	**367,005**	**393,177**	**413,472**	**421,824**

Units (kU)*								
	2021	2022	2023	2024	2025	2026	2027	2028
OEM	116,331	117,701	140,302	149,164	174,321	201,632	231,607	269,642
Channel	29,279	39,707	72,106	81,476	106,089	142,586	187,839	220,431
Total	**145,610**	**157,408**	**212,409**	**230,640**	**280,410**	**344,219**	**419,446**	**490,072**

**Data and projections provided by Forward Insights Q1 2024*