LEGAL NOTICE:

This Technical Proposal is proprietary to the NVM Express, Inc. (also referred to as “Company”) and/or its successors and assigns.

NOTICE TO USERS WHO ARE NVM EXPRESS, INC. MEMBERS: Members of NVM Express, Inc. have the right to use and implement this Technical Proposal subject, however, to the Member’s continued compliance with the Company’s Intellectual Property Policy and Bylaws and the Member’s Participation Agreement.

NOTICE TO NON-MEMBERS OF NVM EXPRESS, INC.: If you are not a Member of NVM Express, Inc. and you have obtained a copy of this document, you only have a right to review this document or make reference to or cite this document. Any such references or citations to this document must acknowledge NVM Express, Inc. copyright ownership of this document. The proper copyright citation or reference is as follows: **“© 2008 to 2023 NVM Express, Inc. ALL RIGHTS RESERVED.”** When making any such citations or references to this document you are not permitted to revise, alter, modify, make any derivatives of, or otherwise amend the referenced portion of this document in any way without the prior express written permission of NVM Express, Inc. Nothing contained in this document shall be deemed as granting you any kind of license to implement or use this document or the specification described therein, or any of its contents, either expressly or impliedly, or to any intellectual property owned or controlled by NVM Express, Inc., including, without limitation, any trademarks of NVM Express, Inc.

LEGAL DISCLAIMER:

THIS DOCUMENT AND THE INFORMATION CONTAINED HEREIN IS PROVIDED ON AN “**AS IS**” BASIS. TO THE MAXIMUM EXTENT PERMITTED BY APPLICABLE LAW, NVM EXPRESS, INC. (ALONG WITH THE CONTRIBUTORS TO THIS DOCUMENT) HEREBY DISCLAIM ALL REPRESENTATIONS, WARRANTIES AND/OR COVENANTS, EITHER EXPRESS OR IMPLIED, STATUTORY OR AT COMMON LAW, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, TITLE, VALIDITY, AND/OR NONINFRINGEMENT.

All product names, trademarks, registered trademarks, and/or servicemarks may be claimed as the property of their respective owners.

The NVM Express® design mark is a registered trademark of NVM Express, Inc.

NVM Express Workgroup
c/o VTM, Inc.
3855 SW 153rd Drive
Beaverton, OR 97003
USA
info@nvmexpress.org

NVM Express® Technical Proposal for New Feature

Technical Proposal ID	TPAR 4162 – Non Storage namespace changes
Change Date	2023-07-27
Impacted Specifications	NVM Express Base Specification 2.0c NVM Command Set 1.0c Zoned Namespace Command Set 1.1c Key Value Command Set 1.0c NVMe-MI 1.2c TP4091 Computational Programs TP4131 Subsystem Local Memory

Technical Proposal Author(s)

Name	**Company**
Bill Martin, Oscar Pinto, Michael Allison, Judy Brock	Samsung
Christoph Hellwig	WD
Kim Malone, James Harris	Intel
Jason Molgaard	Solidigm

This proposal intends to:

- Create new definitions for storage namespace and non storage namespace.
- Modify necessary uses of namespace in the base specification and determine what changes are necessary for other specifications in the NVMe family of specifications.
- Add support for indicating which command sets support namespace management (create/delete).

Revision History

Revision Date	Change Description
2023-05-04	• Initial version
2023-05-10	• Changes from TWG meeting
2023-05-11	• Additional changes from TWG meeting for phase 2 exit
2023-06-27	• Member Review comment resolution
2023-07-25	• Integrated
2023-07-27	• Reformed the number of the section so the navigation pane section numbers match the specifications being modified.

Summary of changes:

- Modify a number of usages of “namespace” throughout the NVM Base Specifications
- Change the definition of namespace in the NVM Command Set and in the KV Command Set
- Does not require any changes in the ZNS Command Set or the Management Interface Specification
- References:
  - NVM Base Specification
  - NVM Command Set Specification
  - KV Command Set Specification

Markup Conventions:

Black:	Unchanged (however, hot links are removed)
~~Red Strikethrough~~:	Deleted
Blue:	New
Blue Highlighted:	TBD values, anchors, and links to be inserted in new text
Purple	Text moved without change.
<Green Bracketed>:	Notes to editor

Modify portions of NVM Express Base Specification 2.0c as follows:

Modify Section 1 as follows:

1 Introduction

…

1.5 Definitions

…

1.5.30 Endurance Group

A portion of ~~NVM~~ non-volatile storage in the NVM subsystem whose endurance is managed as a group. Refer to section 3.2.3.

…

1.5.36 namespace

~~A formatted quantity of non-volatile memory that may be directly accessed by a host.~~

A set of resources that may be directly accessed by a host (e.g., formatted non-volatile storage).

Modify section 2 as follows:

2 Theory of Operation

…

A namespace is a ~~formatted quantity of non-volatile memory~~ set of resources (e.g., formatted non-volatile storage) that may be accessed by a host. A namespace has an associated namespace identifier that a host uses to access that namespace. The set of resources may consist of non-volatile storage and/or other resources.

Associated with each namespace is an I/O Command Set that operates on that namespace. An NVM Express controller may support multiple namespaces ~~that are referenced using a namespace ID~~. Namespaces may be created and deleted using the Namespace Management command and Capacity Management command. The Identify Namespace data structures (refer to section 1.5.29) indicate capabilities and settings that are specific to a particular namespace.

…

Modify section 3 as follows:

3 NVM Express Architecture

3.1 NVM Controller Architecture

…

3.1.2 Controller Types

…

3.1.2.2 Administrative Controller

…

Figure 26 shows an NVM subsystem with one Administrative controller and two I/O controllers within an NVM subsystem that contains a non-volatile storage medium and namespaces. I/O controller one has two attached namespaces, private namespace A and shared namespace B. I/O controller two also has two attached namespaces, private namespace C and shared namespace B. ~~Since an Administrative controller does not provide access to user data stored on an NVM subsystem's non-volatile storage medium~~, ~~the~~ An Administrative controller has no attached namespaces. The Administrative controller in this example may be used for tasks such as NVM subsystem namespace management and efficiently polling NVM subsystem health status via NVMe-MI. While this example shows a single Administrative controller, an NVM subsystem may support zero or more Administrative controllers.

…

3.2 NVM Subsystem Entities

…

3.2.2 NVM Sets

…

An NVM Set is a collection of NVM that is separate (logically and potentially physically) from NVM in other NVM Sets. One or more namespaces that contain formatted storage may be created within an NVM Set and those namespaces inherit the attributes of the NVM Set. A namespace that contains formatted storage is wholly contained within a single NVM Set and shall not span more than one NVM Set.

Figure 72 shows an example of three NVM Sets. NVM Set A contains three namespaces (NS A1, NS A2, and NS A3). NVM Set B contains two namespaces (NS B1 and NS B2). NVM Set C contains one namespace (NS C1). Each NVM Set shown also contains 'Unallocated' regions that consist of NVM that is not yet allocated to a namespace.

…

The NVM Set with which a namespace that contains formatted storage is associated is reported in the Identify Namespace data structure (refer to the applicable NVMe I/O Command Set specification). When a host creates a namespace that contains formatted storage using the Namespace Management command, the host specifies the NVM Set Identifier of the NVM Set that the namespace is to be created in. The namespace that is created inherits attributes from the NVM Set (e.g., the optimal write size to the NVM).

If NVM Sets are supported, then all controllers in the NVM subsystem shall:

- Indicate support for NVM Sets in the Controller Attributes field in the Identify Controller data structure;
- Support the NVM Set Identifier in all commands that use the NVM Set Identifier;
- Support the NVM Set List for the Identify command;
- Indicate the NVM Set Identifier with which ~~the~~ any namespace that contains formatted storage is associated in the Identify Namespace data structure for that namespace;
- Support Endurance Groups; and
- For each NVM Set, indicate the associated Endurance Group as an attribute.

…

3.8 NVM Capacity Model

3.8.1 Overview

NVM subsystems may report capacity-related information for multiple entities within the NVM subsystem. This capacity reporting model includes capacity reporting for the NVM subsystem, the domain (refer to section 3.2.4), the Endurance Group (refer to section 3.2.3), the NVM Set (refer to section 3.2.2), ~~the~~

namespaces that contain formatted storage (refer to section 3.2.1), and the Media Unit (refer to section 1.5.34). Some, all, or none of this reporting may be supported by an NVM subsystem.

Figure 12 shows the hierarchical relationships of the entities within an NVM subsystem which are used to manage NVM capacity.

The capacity in an NVM Set is able to be allocated to one or more namespaces, and each namespace that contains formatted storage exists entirely in that NVM Set (refer to section 3.2.2). Not all of the capacity in the NVM Set is required to be allocated to a namespace.

…

A host uses the Namespace Management create operation (refer to section 8.11) to allocate capacity to namespaces that contain formatted storage.

…

3.8.3 Capacity Reporting

For an NVM subsystem that does not support multiple domains, the capacity information reported in the Identify Controller data structure (i.e., the TNVMCAP field and the UNVMCAP field in Figure 275) describes the capacity for the NVM subsystem. If the MEGCAP field is non-zero, that field indicates the largest entity (e.g., Endurance Group, NVM Set, namespace that contains formatted storage) that may be created in the NVM subsystem.

For an NVM subsystem that supports multiple domains, the capacity information reported in the Identify Controller data structure describes the capacity accessible by the controller processing the Identify command. The host may use the Identify command to access the Domain List data structure (refer to section 5.17.2.17) to determine the domains that are accessible by the controller and the capacity information for each of those domains. If the Max Endurance Group Domain Capacity field is non-zero, then the field describes the largest entity (e.g., Endurance Group, namespace that contains formatted storage) that may be created by this controller in the domain described by that Domain Attributes Entry.

…

For the management of Endurance Groups, NVM Sets, and namespaces that contain formatted storage, Figure 108 describes the effects of the support of NVM Sets, Endurance Groups, and domains on which capacity information is used for each management operation.

…

<change “Namespace” to “Namespace that contains formatted storage” in Figure 108 column 1>

…

Modify section 5 as follows:

5 Admin Command Set

…

5.14 Format NVM command

The Format NVM command is used to low level format the NVM media. This command is used by the host to change the attributes of the NVM media (e.g., the LBA data size and/or metadata size for the NVM Command Set). A low level format may destroy all data and metadata associated with all namespaces that contain formatted storage or only the specific namespace associated with the command (refer to the Format NVM Attributes field in the Identify Controller data structure, Figure 275). After the Format NVM command successfully completes, the controller shall not return any user data that was previously contained in an affected namespace.

…

5.16 Get Log Page command

…

5.16.1 Log Specific Information

…

5.16.1.6 Commands Supported and Effects (Log Identifier 05h)

…

Figure 211: Commands Supported and Effects Data Structure

Bits	Description
…	**…**
15:05	Reserved
04	**Controller Capability Change (CCC):** If this bit is set to '1', then this command may change controller capabilities. If this bit is cleared to '0', then this command does not modify controller capabilities. Controller capability changes include a firmware update that changes the capabilities reported in the CAP property.
03	**Namespace Inventory Change (NIC):** If this bit is set to '1', then this command may change the number of namespaces or capabilities for multiple namespaces. If this bit is cleared to '0', then this command does not modify the number of namespaces or capabilities for multiple namespaces. Namespace inventory changes include adding or removing namespaces.
02	**Namespace Capability Change (NCC):** If this bit is set to '1', then this command may change the capabilities of a single namespace. If this bit is cleared to '0', then this command does not modify any namespace capabilities for the specified namespace. Namespace capability changes include a logical format change.
01	**Logical Block Content Change (LBCC):** If this bit is set to '1', then this command may modify user data content in one or more namespaces that contain formatted storage. If this bit is cleared to '0', then this command does not modify user data content in any namespace. User data content changes include a write to user data. NOTE: This field applies to all user data content changes. The original name has been retained for historical continuity.
00	**Command Supported (CSUPP):** If this bit is set to '1', then this command is supported by the controller. If this bit is cleared to '0', then this command is not supported by the controller and all other fields in this structure shall be cleared to 0h.

…

5.16.1.14 Persistent Event (Log Identifier 0Dh)

…

5.16.1.14.1 Persistent Event Log Events

…

5.16.1.14.1.6 Change Namespace Event (Event Type 06h)

The Changed Namespace Event (refer to Figure 236) persists the host parameters used for successful Namespace Management commands. This event type is specific to namespaces that are associated with I/O Command Sets that specify logical blocks (e.g., the NVM Command Set and the Zoned Namespace Command Set). The event contains a Persistent Event Log Event header and the Change Namespace Event data.

…

5.17 Identify command

…

5.17.2 Identify Data Structures

…

5.17.2.8 I/O Command Set Independent Identify Namespace data structure (CNS 08h)

…

Figure 280: Identify – I/O Command Set Independent Identify Namespace Data Structure

Bytes	O/M [1]	Description
…	…	**…**
09		Reserved
11:10	O	**NVM Set Identifier (NVMSETID):** For NSID other than FFFFFFFFh, this field indicates the NVM Set with which this namespace is associated. If NVM Sets are not supported by the controller, then this field shall be cleared to 0h. For namespaces that do not contain formatted storage this field shall be cleared to 0h.
13:12	O	**Endurance Group Identifier (ENDGID):** For NSID other than FFFFFFFFh, this field indicates the Endurance Group with which this namespace is associated. If Endurance Groups are not supported by the controller, then this field shall be cleared to 0h. For namespaces that do not contain formatted storage this field shall be cleared to 0h.
…	…	**…**
4095:15		Reserved

Notes:
1. O/M definition: O = Optional, M = Mandatory.

Modify section 8.3.3 Variable Capacity Management as follows:

Modify section 8 as follows:

8 Extended Capabilities

…

8.3 Capacity Management

…

8.3.3 Variable Capacity Management

…

A typical sequence of operations for allocating capacity is:

1) determine the available capacities in each domain (refer to section 3.2.4);
2) create Endurance Group with desired capacity (refer to section 5.3);
3) create NVM Set with desired capacity in the Endurance Group (refer to section 5.3); and

4) create namespace with desired capacity in the NVM Set (refer to section 5.23).

A typical sequence of operations for deallocating capacity is:

1) delete namespace that contains formatted storage, if any, (refer to section 5.23);
2) delete NVM Set, if any, (refer to section 5.3); and
3) delete Endurance Group (refer to section 5.3).

…

8.11 Namespace Management

…

The size of a namespace that contains formatted storage is based on the size requested in a create operation, the format of the namespace, and any characteristics (e.g., endurance). The controller determines the NVM capacity allocated for that namespace. Namespaces may be created with different usage characteristics (e.g., endurance) that utilize differing amounts of NVM capacity. Namespace characteristics and the mapping of these characteristics to NVM capacity usage are outside the scope of this specification.

Reporting of capacity information for the NVM subsystem, Domain, Endurance Group, and NVM Set are described in section 3.8. For each namespace that contains formatted storage, the NVM Set and the Endurance Group that contain the namespace are reported in the Identify Namespace data structure. The NVM Set to be used for a namespace that contains formatted storage is based on the value in the NVM Set Identifier field in a create operation. If the NVM Set Identifier field is cleared to 0h in a create operation for a namespace that contains formatted storage, then the controller shall choose the NVM Set from which to allocate capacity to create ~~the~~ that namespace.

If the NVM Set Identifier field and the Endurance Group Identifier field are both cleared to 0h in a create operation for a namespace that contains formatted storage, then the controller shall choose the Endurance Group and the NVM Set from which to allocate capacity to create ~~the~~ that namespace.

If the NVM Set Identifier field is cleared to 0h and the Endurance Group Identifier field is set to a non-zero value in a create operation for a namespace that contains formatted storage, then the controller shall choose the NVM Set in the specified Endurance Group from which to allocate capacity to create ~~the~~ that namespace.

If the NVM Set Identifier field is set to a non-zero value and the Endurance Group Identifier field is cleared to 0h in a create operation for a namespace that contains formatted storage, then the controller shall abort the command with a status code of Invalid Field in Command.

If the NVM Set Identifier field and the Endurance Group Identifier field are both set to non-zero values in a create operation for a namespace that contains formatted storage and the specified NVM Set exists in the specified Endurance Group, then the controller shall allocate capacity for ~~the~~ that created namespace from the specified NVM Set.

If the NVM Set Identifier field and the Endurance Group Identifier field are both set to non-zero values in a create operation for a namespace that contains formatted storage and the specified NVM Set does not exist in the specified Endurance Group, then the controller shall abort the command with a status code of Invalid Field in Command.

For each namespace that contains formatted storage, the NVM capacity used for that namespace is reported in the Identify Namespace data structure (refer to the applicable I/O Command Set specification). The controller may allocate NVM capacity in units such that the requested size for a namespace that contains formatted storage may be rounded up to the next unit boundary. The units in which NVM capacity is allocated are reported in the Namespace Granularity List (refer to the NVM Command Set Specification), if supported. For example, when using the NVM Command Set, if host software requests a namespace of 32 logical blocks with a logical block size of 4 KiB for a total size of 128 KiB and the allocation unit for the implementation is 1 MiB, then the NVM capacity consumed may be rounded up to 1 MiB. The NVM capacity fields may not correspond to the logical block size multiplied by the total number of logical blocks.

The method of allocating ANA Group identifiers is outside the scope of this specification. If the ANA Group Identifier (refer to Figure 280 and the Identify Namespace data structure in the NVM Command Set Specification) is cleared to 0h, then the controller shall determine the ANAGRPID that is assigned to that namespace.

To create a namespace, host software performs the following actions:

1. Host software requests the Identify Namespace data structure that specifies common namespace capabilities (i.e., using an Identify command with the NSID field set to FFFFFFFFh and the CNS field cleared to 0h);
2. If the controller supports reporting of I/O Command Set specific Namespace Management content (refer to the Namespace Management section in the applicable I/O Command Set specification), host software optionally requests that information (e.g. Namespace Granularity).
3. Host software determines available ~~capacity~~ resources (e.g., capacity for a namespace that contains formatted storage) (refer to section 3.8);
4. Host software creates the data structure defined in Figure 300 (e.g., taking into account the common namespace capabilities, available capacity);
5. Host software issues the Namespace Management command specifying the Create operation and the data structure. On successful completion of the command, the Namespace Identifier of the new namespace is returned in Dword 0 of the completion queue entry. At this point, the new namespace is not attached to any controller; and
6. Host software requests the Identify Namespace data structures for the new namespace to determine all attributes of the namespace.

…

8.15 Power Management

…

8.15.4 Runtime D3 Transitions

…

The RTD3 Resume Latency is the expected elapsed time from the time power is applied until the controller is able to:

a) process and complete I/O commands; and
b) access the resources (e.g., formatted storage~~NVM~~) associated with attached namespace(s), if any, as part of I/O command processing.

…

Modify portions of NVM Express NVM Command Set Specification 2.c as follows:

Modify section 2 as follows:

2 NVM Command Set Model

…

2.1 Theory of operation

…

2.1.1 Namespaces

…

A namespace is a set of resources that may be accessed by a host ~~a collection of NVM~~ and is as defined in the NVM Express Base Specification. A namespace has an associated namespace identifier that a host uses to access that namespace.

The Identify Namespace data structure (refer to Figure 97), for a namespace associated with this command set, contains related fields reporting the Namespace Size, Namespace Capacity, and Namespace Utilization:

- The Namespace Size (NSZE) field defines the total size of the namespace in logical blocks (LBA 0 through n-1).
- The Namespace Capacity (NCAP) field defines the maximum number of logical blocks that may be allocated at any point in time.
- The Namespace Utilization (NUSE) field defines the number of logical blocks currently allocated in the namespace.

…

Modify portions of NVM Express Zoned Namespace Command Set Specification 1.1c as follows:

Modify section 2 as follows:

2 Zoned Namespace Command Set Model

…

2.1 Theory of operation

…

2.1.1 Namespaces

…

A namespace is a set of resources that may be accessed by a host ~~a collection of NVM~~ and is as defined in the NVM Express Base Specification and in the NVM Command Set Specification, as modified by this specification. A namespace has an associated namespace identifier that a host uses to access that namespace.

…

Modify portions of the NVM Express Key Value Command Set Specification 2.c as follows:

Modify section 2 as follows:

2 Key Value Command Set Model

…

2.1 Theory of operation

…

2.1.1 Namespaces

…

A namespace is a set of resources that may be accessed by a host ~~a collection of NVM~~ and is as defined in the NVM Express Base Specification. A namespace has an associated namespace identifier that a host uses to access that namespace.

…