



LEGAL NOTICE:

© Copyright 2007 - 2018 NVM Express, Inc. ALL RIGHTS RESERVED.

This NVM Express revision 1.3 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 NVM Express revision 1.3 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: "© 2007 - 2018 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.

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

NVM Express Technical Proposal for New Feature

Technical Proposal ID	TP 4031
Change Date	08/15/2018
Builds on Specification	NVM Express 1.3c

Technical Proposal Author(s)

Name	Company
Fred Knight	NetApp
Peter Onufryk	Microsemi
Bill Martin	Samsung

This technical proposal enables write streams to be shared across multiple hosts.

Revision History

Revision Date	Change Description
01/16/2018	Initial Creation
03/04/2018	Address exclusive host resources.
04/04/2018	Switch from 2-bit field to a single bit (no "undefined case").
05/11/2018	Add 9.3.1.2 text – clear tracked changes & change bars
08/15/2018	Incorporate integration changes.
08/20/2018	Ratified

Description of Specification Changes

Make changes to section 9.3 (Streams) as shown below:

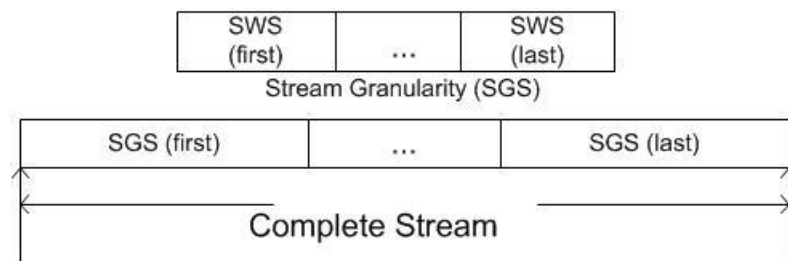
9.3 Streams (Directive Type 01h, Optional)

The Streams Directive enables the host to indicate (i.e., by using the stream identifier) to the controller that the specified logical blocks in a write command are part of one group of associated data. This information may be used by the controller to store related data in associated locations or for other performance enhancements.

The controller provides information in response to the Return Parameters operation about the configuration of the controller that indicates Stream Write Size, Stream Granularity Size, and stream resources at the NVM subsystem and namespace levels.

Data that is aligned to and in multiples of the Stream Write Size (SWS) provides optimal performance of the write commands to the controller. The Stream Granularity Size indicates the size of the media that is prepared as a unit for future allocation for write commands and is a multiple of the Stream Write Size. The controller may allocate and group together a stream in Stream Granularity Size (SGS) units. Refer to Figure .

Figure 290: Directive Streams – Stream Alignment and Granularity



If the host issues a Dataset Management command to deallocate logical blocks that are associated with a stream, it should specify a starting LBA and length that is aligned to and in multiples of the Stream Granularity Size. This provides optimal performance and endurance of the media.

Stream resources are the resources in the NVM subsystem that are necessary to track operations associated with a specified stream identifier. There are a maximum number of stream resources that are available in an NVM subsystem as indicated by the Max Stream Limit (MSL) field in the Return Parameters data structure.

Available NVM subsystem stream resources are stream resources that are not allocated for exclusive use in any namespace. Available NVM subsystem stream resources are reported in the NVM Subsystem Streams Available (NSSA) field and may be used by any host in any namespace that:

- has the Streams Directive enabled; **and**
- has not been allocated exclusive stream resources by that host **if bit 0 of the NSSC field is cleared to '0'; and**
- **has not been allocated exclusive stream resources by any host if bit 0 of the NSSC field is set to '1'.**

Each time stream resources are allocated for exclusive use in a specified namespace, the available NVM subsystem stream resources reported in the NSSA field are reduced.

For a given namespace:

- a) a host allocates stream resources to that namespace for the exclusive use of that host(s) by issuing the Allocate Resources operation;
- b) other hosts may concurrently allocate stream resources to that namespace for their exclusive use; and
- c) hosts which have not allocated stream resources to that namespace may use available NVM subsystem stream resources for access to that namespace.

The Directive operations that shall be supported if the Streams Directive is supported are listed in Figure . The Directive Specific field in a command is referred to as the stream identifier when the Directive Type field is set to the Streams Directive.

Figure 291: Streams – Directive Operations

Directive Command	Directive Operation Name	Directive Operation Value	Definition
Directive Receive	Return Parameters	01h	Section 9.3.1.1
	Get Status	02h	Section 9.3.1.2
	Allocate Resources	03h	Section 9.3.1.3
	Reserved	All others	
Directive Send	Release Identifier	01h	Section 9.3.2.1
	Release Resources	02h	Section Error! Reference source not found.
	Reserved	All others	

Stream identifiers are assigned by the host and may be in the range 0001h to FFFFh. The host may specify a sparse set of stream identifiers (i.e., there is no requirement for the host to use ~~Stream~~ identifiers in any particular order).

~~The host may be accessing a namespace through multiple controllers in the NVM subsystem. The controllers in an NVM subsystem distinguish if the stream identifier has the same meaning for a particular namespace by the Host Identifier. If more than one Host Identifier has the same non-zero value, then that value represents a single host that is accessing the namespace through multiple controllers and the stream identifier is used across controllers to access the same stream on the namespace. If a Host Identifier is zero or has a unique value, then that value represents a unique host that is accessing the namespace and the stream identifier does not have the same meaning for a particular namespace.~~

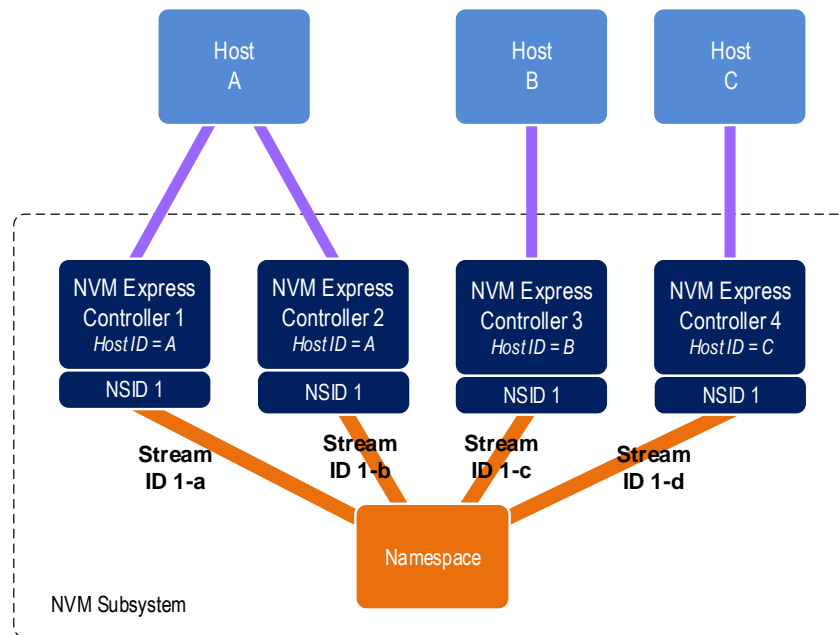
~~The host may access a namespace through multiple controllers in the NVM subsystem. The controllers in an NVM subsystem indicate in bit 0 of the NSSC field (refer to Figure 293) if a stream identifier is unique based on the Host Identifier (i.e., the same stream identifier used to access the same namespace by a host that has registered a different Host Identifier is referencing a different stream), or if a stream identifier may be used by multiple Host Identifiers (i.e., the same stream identifier used to access the same namespace by a host that has registered a different Host Identifier is referencing the same stream). All controllers in an NVM subsystem shall report the same value in the NSSC field.~~

~~If multiple controllers receive a registration of a Host Identifier (refer to section 5.21.1.19) that has the same non-zero value, then that value represents a single host that is accessing the namespace through those controllers and a stream identifier is used across those controllers to access the same stream on the namespace. If a Host Identifier has a unique non-zero value, then each value represents a unique host that is accessing the namespace and:~~

- ~~• if bit 0 of the NSSC field is cleared to '0', then the same stream identifier on controllers with different non-zero Host Identifiers does not have the same meaning for a particular namespace (i.e., the stream identifier is not used across controllers with different non-zero Host Identifiers to access the same stream on the namespace); and~~
- ~~• if bit 0 of the NSSC field is set to '1', then the same stream identifier on any controller with a non-zero Host Identifier has the same meaning for a particular namespace (i.e., the stream identifier is used across controllers to access the same stream on the namespace).~~

~~If a Host Identifier is set to zero, then a unique host is accessing the namespace and the stream identifier does not have the same meaning for a particular namespace.~~

Figure 291.a: Example Multi-Stream and NSSC



In the example shown in Figure 291.a, if NSSC bit 0 is cleared to '0', then there are three streams as follows:

- Stream ID 1-a and Stream ID 1-b have the same meaning;
- Stream ID 1-c has a different meaning; and
- Stream ID 1-d has a different meaning.

In the example shown in Figure 291.a, if NSSC bit 0 is set to '1', then there is one stream as follows:

- Stream ID 1-a, Stream ID 1-b, Stream ID 1-c, and Stream ID 1-d have the same meaning.

The controller(s) recognized by the NVM subsystem as being associated with a specific host **or** hosts and attached to a specific namespace either:

- utilizes a number of stream resources allocated for exclusive use of that namespace as returned in response to an Allocate Resources operation; or
- utilizes resources from the NVM subsystem stream resources.

The value of Namespace Streams Allocated (NSA) indicates how many resources for individual stream identifiers have been allocated for exclusive use for the specified namespace by the associated controllers. This indicates the maximum number of stream identifiers that may be open at any given time in the specified namespace by the associated controllers. To request a different number of resources than are currently allocated for exclusive use by the associated controllers for a specific namespace, all currently allocated resources are first required to be released using the Release Resources operation. There is no mechanism to incrementally increase or decrease the number of allocated resources for a given namespace.

Streams are opened by the controller when the host issues a Write command that specifies a stream identifier that is not currently open. While a stream is open the controller maintains context for that stream (e.g., buffers for associated data). The host may determine the streams that are open using the Get Status operation.

For a namespace that has a non-zero value of Namespace Streams Allocated (NSA), if the host submits a Write command specifying a stream identifier not currently in use and stream resources are exhausted, then an arbitrary stream identifier for that namespace is released by the controller to free the stream resources associated with that stream identifier for the new stream. The host may ensure the number of open streams does not exceed the allocated stream resources for the namespace by explicitly releasing stream identifiers as necessary using the Release Identifier operation.

For a namespace that has zero namespace stream resources allocated, if the host submits a Write command specifying a stream identifier not currently in use and:

- NVM subsystem streams available are exhausted, then an arbitrary stream identifier for an arbitrary namespace that is using NVM subsystem stream resources is released by the NVM subsystem to free the stream resources associated with that stream identifier for the new stream; or
- all NVM subsystem stream resources have been allocated for exclusive use for specific namespaces, then the Write command is treated as a normal Write command that does not specify a stream identifier.

The host determines parameters associated with stream resources using the Return Parameters operation. The host may get a list of open stream identifiers using the Get Status operation.

If the Streams Directive becomes disabled for a host in a namespace, then all stream resources and stream identifiers are released for the host in that namespace. If the host issues a Format NVM command, or deletes a namespace, then all stream identifiers for all open streams for affected namespaces are released.

Streams Directive defines the command specific status values specified in Figure .

Figure 292: Streams Directive – Command Specific Status Values

Value	Description
7Fh	Stream Resource Allocation Failed: The controller was not able to allocate stream resources for exclusive use of the specified namespace and no NVM subsystem stream resources are available.

9.3.1 Directive Receive

This section defines operations used with the Directive Receive command for the Streams Directive.

9.3.1.1 Return Parameters (Directive Operation 01h)

The Return Parameter operation returns a data structure that specifies the features and capabilities supported by the Streams Directive, including namespace specific values. The DSPEC field in command Dword 11 is not used for this operation. The data structure returned is defined in Figure . If an NSID value of FFFFFFFFh is specified, then the controller returns the NVM subsystem specific values, may return any namespace specific values that are the same for all namespaces (e.g., SWS), and clears all other namespace specific fields to zero.

Figure 293: Streams Directive– Return Parameters Data Structure

Bytes	Description
	NVM Subsystem Specific Fields
01:00	Max Streams Limit (MSL): This field indicates the maximum number of concurrently open streams that the NVM subsystem supports. This field returns the same value independent of specified namespace.
03:02	NVM Subsystem Streams Available (NSSA): This field indicates the number of NVM subsystem stream resources available. These are the stream resources that are not allocated for the exclusive use by a host in any specific namespace. This field returns the same value independent of specified namespace.
05:04	NVM Subsystem Streams Open (NSSO): This field indicates the number of open streams in the NVM subsystem that are not associated with a namespace for which resources were allocated using an Allocate Resources operation. This field returns the same value independent of specified namespace.
06	<p>NVM Subsystem Stream Capability (NSSC): This field indicate the stream capabilities of the NVM subsystem.</p> <p>Bits 7:1 are reserved.</p> <p>Bit 0 indicates whether stream identifiers may be shared by multiple Host Identifiers, or if a stream identifier is associated with a single Host Identifier. Bit 0 if cleared to '0', then the stream identifier is associated with a single non-zero Host Identifier. Bit 0 if set to '1', then the stream identifier may be associated with multiple non-zero Host Identifiers.</p>

Bytes	Description
15:076	Reserved
Namespace Specific Fields	
19:16	Stream Write Size (SWS): This field indicates the alignment and size of the optimal stream write as a number of logical blocks for the specified namespace. The size indicated should be less than or equal to Maximum Data Transfer Size (MDTS) that is specified in units of minimum memory page size. SWS may change if the namespace is reformatted with a different LBA format. If the NSID value is set to FFFFFFFFh, then this field may be cleared to 0h if a single logical block size cannot be indicated.
21:20	Stream Granularity Size (SGS): This field indicates the stream granularity size for the specified namespace in Stream Write Size (SWS) units. If the NSID value is set to FFFFFFFFh, then this field may be cleared to 0h.
Namespace and Host Identifier Specific Fields	
23:22	<p>Namespace Streams Allocated (NSA): This field indicates the number of stream resources allocated for exclusive use of the specified namespace.</p> <p>If bit 0 of the NSSC field is cleared to '0', then those exclusive stream resources are shared by the controller processing the Return Parameters operation and all other controllers which that share the same non-zero Host Identifier, and which are attached to the specified namespace.</p> <p>If bit 0 of the NSSC field is set to '1', then those exclusive stream resources are shared by all controllers that are associated with any non-zero Host Identifier and are attached to this namespace.</p> <p>If this value is non-zero, then the namespace may have up to NSA number of concurrently open streams. If this field is cleared to zero, then no stream resources are currently allocated to this namespace and the namespace may have up to NSSA number of concurrently open streams.</p>
25:24	<p>Namespace Streams Open (NSO): This field indicates the number of open streams in the specified namespace.</p> <p>If bit 0 of the NSSC field is cleared to '0', then this field indicates the number of streams that were opened by the controller processing the Return Parameters operation and by all other controllers which that share the same non-zero Host Identifier, and which are attached to this namespace.</p> <p>If bit 0 of the NSSC field is set to '1', then this field indicates the number of streams that were opened by the controller processing the Return Parameters operation and all other controllers that are associated with any non-zero Host Identifier and are attached to this namespace.</p> <p>NOTE: It is not possible for a host to retrieve the number of open streams using resources allocated to the specified namespace by other hosts.</p>
31:26	Reserved

9.3.1.2 Get Status (Directive Operation 02h)

The Get Status operation returns information about the status of currently open streams for the specified namespace and the host issuing the Get Status operation. The DSPEC field in command Dword 11 is not used for this operation.

If NSSC bit 0 is cleared to '0', then the information returned describes only those resources for the specified namespace that are associated with hosts that are registered with the same non-zero Host Identifier value as the host issuing the Get Status operation. If NSSC bit 0 is set to '1', then the information returned describes the resources for the specified namespace that are associated with hosts that are registered with any non-zero Host Identifier.

If an NSID value of FFFFFFFFh is specified, then the controller shall return information about the status of currently open streams in the NVM subsystem that use resources which are not allocated for the exclusive use of any particular namespace. If a stream identifier value being returned is in use by different namespaces, then that stream identifier shall be returned only once.

Stream Identifier 1 (i.e., returned at offset 03:02) contains the value of the open stream of lowest numerical value. Each subsequent field contains the value of the next numerically greater stream identifier of an open stream.

The data structure returned is defined in Figure . All fields are specific to the namespace specified **if the NSID value was not set to FFFFFFFFh**.

Figure 294: Streams Directive – Get Status Data Structure

Bytes	Description
01:00	Open Stream Count: This field specifies the number of streams that are currently open.
03:02	Stream Identifier 1: This field specifies the stream identifier of the first (numerically lowest) open stream.
05:04	Stream Identifier 2: This field specifies the stream identifier of the second open stream.
...	...
131071: 131070	Stream Identifier 65535: This field specifies the stream identifier of the 65535 th open stream.

9.3.1.3 Allocate Resources (Directive Operation 03h)

The Allocate Resources operation indicates the number of streams that the host requests for the exclusive use ~~by the host~~ for the specified namespace. **If bit 0 of the NSSC field is cleared to '0', then those resources are for the exclusive use of hosts that are registered with the same Host Identifier as the host that made the request. If bit 0 of the NSSC field is set to '1', then those resources are for the exclusive use of any host that is registered with any non-zero Host ID.** The DSPEC field in command Dword 11 is not used for this operation. The operation returns the number of streams allocated in Dword 0 of the completion queue entry. The value allocated may be less than or equal to the number requested. The allocated resources shall be reflected in the Namespace Streams Allocated field of the Return Parameters data structure.

If the controller is unable to allocate any stream resources for the exclusive use ~~by the host~~ for the specified namespace, then the controller shall:

- return a status value of Stream Resource Allocation Failed; or
- if NVM subsystem stream resources are available, then set NSA to 0000h in the completion queue entry to indicate that the host may use stream resources from the NVM subsystem for this namespace.

If the specified namespace already has stream resources allocated for the exclusive use of the host issuing the Allocate Resources operation, then the controller shall return a status value of Invalid Field in Command. To allocate additional streams resources, the host should release resources and ~~then~~ request a complete set of resources.

No data transfer occurs.

Figure 295: Allocate Resources – Command Dword 12

Bit	Description
31:16	Reserved
15:00	Namespace Streams Requested (NSR): This field specifies the number of stream resources the host is requesting be allocated for exclusive use by the namespace specified.

Figure 296: Allocate Resources – Dword 0 of command completion queue entry

Bit	Description
31:16	Reserved
15:00	Namespace Streams Allocated (NSA): This field indicates the number of streams resources that have been allocated for exclusive use by the namespace specified. The allocated resources are available to all controllers associated with that host.

9.3.2 Directive Send

This section defines operations used with the Directive Send command for the Streams Directive.

9.3.2.1 Release Identifier (Directive Operation 01h)

The Release Identifier operation specifies that the stream identifier specified in the DSPEC field in command Dword 11 is no longer in use by the host. Specifically, if the host uses the stream identifier in a future operation, then it is referring to a different stream. If the specified identifier does not correspond to an open stream for the specified namespace, then the command completes successfully. If there are stream resources allocated for the specified namespace, then the stream resources remain allocated for this namespace, and may be re-used in a subsequent write command. If there are no stream resources allocated for the specified namespace, then the stream resources are returned to the NVM subsystem stream resources for future use by a namespace without allocated stream resources. If an NSID value of FFFFFFFFh is specified, then the controller shall abort the command with a status of Invalid Field in Command.

No data transfer occurs.

9.3.2.2 Release Resources (Directive Operation 02h)

The Release Resources operation is used to release all streams resources allocated for the namespace attached to all controllers:

- associated with the same non-zero Host Identifier of the controller that processed the operation **if bit 0 of the NSSC field is cleared to '0'; and**
- **associated with any non-zero Host Identifier if bit 0 of the NSSC field is set to '1'.**

On successful completion of this command, the **number of** allocated stream resources **are is** cleared to 0h for the specified namespace. If this command is issued when no streams resources are allocated for the namespace, the command shall complete successfully.

No data transfer occurs.