



LEGAL NOTICE:

© Copyright 2007 to 2020 NVM Express™, Inc. ALL RIGHTS RESERVED.

This NVM Express revision 1.4 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 1.4 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 to 2020 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	TP 6026 Boot Partition Read Access using OOB
Change Date	2020-11-15
Builds on Specification	NVMe 1.4a TP 4071a Commands and Effects Log Enhancement TP 4047a Aborting Commands During Format NVM Command
References Specification	TP 4046a Command Group Control Feature

Technical Proposal Author(s)

Name	Company
Mike Allison, Myron Loewen	Intel

This proposal enables access of Boot Partitions via a log page to allow SMBus access in the event PCIe is not accessible.

Revision History

Revision Date	Change Description
2020-08-10	Initial version
2020-08-24	Updates section numbers and made sure inserted text is blue.
2020-08-27	Made it clear that reading Boot Partition from log page can be supported if the NVMe Boot Partition registers are not supported.
2020-09-09	Changed the title as asked during Phase 2 exit ballot. Aligned the Log Specific Fields to match the changes identified by TP 4046a. Small editorial changes.
2020-09-14	Removed last sentence of section 8.13.1.
2020-10-01	Accepted for member review. Removed all commands and accepted all changes.
2020-11-15	Integrated into the NVMe Base Specification.

Description for NVMe 1.4a Changes Document

This technical proposal adds:

- New log page to return Boot Partition contents.

Description of Specification Changes

Markup Conventions:

Black: Unchanged (however, hot links are removed)

~~Red Strikethrough~~: Deleted

Blue: New

Technical input submitted to the NVM Express™ Workgroup is subject to the terms of the NVM Express™ Participant's agreement. Copyright © 2014 to 2020 NVMe™ Corporation.

Blue Highlighted: TBD values, anchors, and links to be inserted in new text.

<Green Bracketed>: Notes to editor

Modify portions of NVMe 1.4a as shown below:

Modify portions of section 5 as defined by TP 4047a shown below:

5 Admin Command Set

...

Figure TBD: Sanitize Operations and Format NVM Command – Admin Commands Allowed

Admin Command	Additional Restrictions for Format NVM command	Additional Restrictions for sanitize operations
Abort		
Asynchronous Event Request		
Create I/O Completion Queue		
Create I/O Submission Queue		
Device Self-test	Only Controller DST should be allowed	Prohibited
Delete I/O Completion Queue		
Delete I/O Submission Queue		
Get Features		
Get Log Page	The log pages allowed for both Format NVM command and sanitize operations are listed below.	
	Log Pages	Additional Restrictions for both Format NVM command and sanitize operations
	Error Information	Return zeros in the LBA field.
	SMART / Health Information	
	Changed Namespace List	
	Reservation Notification	
	Asymmetric Namespace Access	
	Sanitize Status	
	Vendor specific	Prohibited for Sanitize
	Persistent Event Log	Prohibited for Sanitize
	Boot Partition	
Identify		
Keep Alive		
NVMe-MI Receive	Allowed unless explicitly prohibited in the NVMe Express Management Interface Specification.	
NVMe-MI Send		
Sanitize		Prohibited
Set Features	Namespace Write Protection Config Feature is not allowed.	

Figure TBD: Sanitize Operations and Format NVM Command – Admin Commands Allowed

Admin Command	Additional Restrictions for Format NVM command	Additional Restrictions for sanitize operations
Opcode 7Fh	The Fabric Commands allowed are listed below. Refer to the NVMe over Fabrics specification.	
	Fabrics Commands	Additional Restrictions for both Format NVM command and sanitize operations
	Property Set	
	Connect	
	Disconnect	
	Property Get	
	Authentication Send	
	Authentication Receive	
	Vendor Specific	Commands are allowed that do not affect or retrieve user data.
Vendor Specific	Commands are allowed that do not affect or retrieve user data.	

Modify portions of Figure 191 in section 5.14.1 as shown below:

5.14.1 Log Specific Information

...

Figure 191: Get Log Page – Log Page Identifiers

Log Identifier	Scope	Log Page Name	Reference Section
0Eh	Controller	LBA Status Information	5.14.1.14
0Fh	NVM subsystem	Endurance Group Event Aggregate	5.14.1.15
...			
15h	Controller	Boot Partition	5.14.1.TBD
16h to 6Fh	Reserved		
NVMe over Fabrics Implementations			
70h	Discovery (refer to the NVMe over Fabrics specification)		
71h to 7Fh	Reserved for NVMe over Fabrics implementations		
I/O Command Set Specific			
80h to BFh	Reserved		
Vendor Specific			
C0h to FFh	Vendor specific ⁵		
KEY:			
Namespace = The log page contains information about a specific namespace.			
Controller = The log page contains information about the controller that is processing the command.			
NVM subsystem = The log page contains information about the NVM subsystem.			
NOTES:			
1. For namespace identifiers of 0h or FFFFFFFFh.			
2. For namespace identifiers other than 0h or FFFFFFFFh.			
3. Bit 0 is cleared to '0' in the DSTO field in the Identify Controller data structure (refer to Figure 247).			
4. Bit 0 is set to '1' in the DSTO field in the Identify Controller data structure.			
5. Selection of a UUID may be supported. Refer to section 8.24.			

Add section 5.14.1.TBD as shown below:

5.14.1.TBD Boot Partition (Log Identifier 15h)

The Boot Partition Log page provides read only access to the Boot Partition (refer to section 8.13) accessible by this controller through the BPRSEL register (refer to section 3.1.14).

This log consists of a header describing the Boot Partition and Boot Partition data as defined by Figure TBD_FIG. The Boot Partition Identifier bit in the Log Specific field determines the Boot Partition.

A host reading this log page has no effects on the BPINFO (refer to section 3.1.13), BPRSEL, and BPMBL (refer to section 3.1.15) registers.

Figure 201: Command Dword 10 – Log Specific Field

Bits	Description
14:09 <<TP 4046a >>	Reserved
08	Boot Partition Identifier: This bit specifies the Boot Partition identifier for the Boot Partition to return.

Figure TBD_FIG: Get Log Page – Boot Partition Log (Log Identifier 15h)

Bytes	Description								
Boot Partition Header									
00	Log Identifier: This field shall be set to 15h.								
03:01	Reserved								
07:04	Boot Partition Information: Contains defines the characteristics of Boot Partitions.								
	<table><tr><th>Bits</th><th>Description</th></tr><tr><td>31</td><td>Active Boot Partition ID (ABPID): This bit indicates the identifier of the active Boot Partition.</td></tr><tr><td>30:15</td><td>Reserved</td></tr><tr><td>14:00</td><td>Boot Partition Size (BPSZ): This field defines the size of the Boot Partition Data field in multiples of 128 KiB.</td></tr></table>	Bits	Description	31	Active Boot Partition ID (ABPID): This bit indicates the identifier of the active Boot Partition.	30:15	Reserved	14:00	Boot Partition Size (BPSZ): This field defines the size of the Boot Partition Data field in multiples of 128 KiB.
	Bits	Description							
	31	Active Boot Partition ID (ABPID): This bit indicates the identifier of the active Boot Partition.							
	30:15	Reserved							
14:00	Boot Partition Size (BPSZ): This field defines the size of the Boot Partition Data field in multiples of 128 KiB.								
15:08	Reserved								
Boot Partition Data									
BPSZ*128 KiB + 15:16	Boot Partition Data: Contains the contents of the specified Boot Partition.								

Modify a portion of Figure 422 in section 7.1.1 as shown below:

Figure 422: I/O Controller – Log Page Support

Log Page Name	Command Support Requirements ¹
...	
Endurance Group Event Aggregate	O
Boot Partition	O

Figure 422: I/O Controller – Log Page Support

Log Page Name	Command Support Requirements ¹
Notes: 1. O = Optional, M = Mandatory, P = Prohibited	

Modify a portion of Figure 430 in section 7.1.2 as shown below:

Figure 430: Administrative Controller – Log Page Support

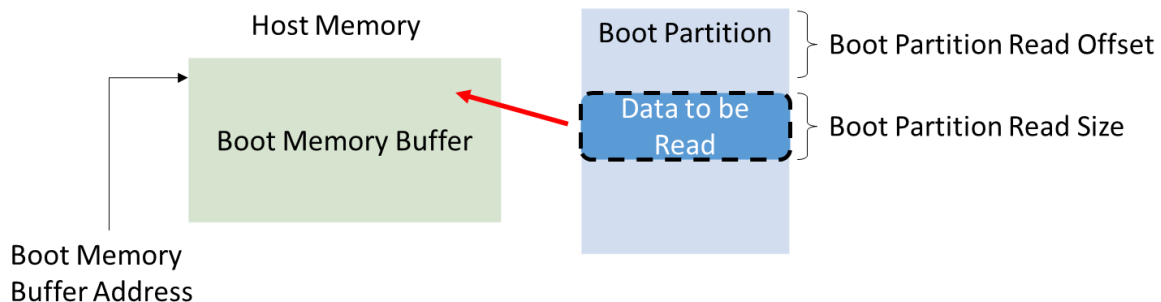
Log Page Name	Command Support Requirements ¹
...	
Endurance Group Event Aggregate	O
Boot Partition	O
Notes: 2. O = Optional, M = Mandatory, P = Prohibited	

Modify a portion of section 8.13.1 as shown below:

8.13.1 Reading from a Boot Partition

A Boot Partition is a continuous block of data as shown in Figure 478, that the host may read [via NVMe registers](#).

Figure 478: Boot Partition Overview



To read the contents of a Boot Partition [using NVMe registers](#), the host allocates a Boot Partition Memory Buffer in host memory for the controller to copy contents from a Boot Partition. The host initializes the Boot Partition Memory Buffer Base Address. The host sets the Boot Partition ID, Boot Partition Read Size, and Boot Partition Read Offset to initiate the Boot Partition read operation. The host may continue reading from the Boot Partition until the entire Boot Partition has been read.

A portion of the Boot Partition may be read by the host any time the NVM subsystem is powered (i.e., whether or not CC.EN is set to '1'). The host shall not modify the PCI Express registers (described in section 2), reset, or shutdown the controller while a Boot Partition read is in progress.

To read data from a Boot Partition, the host follows these steps:

1. Initialize the transport (e.g., PCIe link), if necessary;
2. Determine if Boot Partitions are supported by the controller (CAP.BPS);
3. Determine which Boot Partition is active (BPINFO.ABPID) and the size of the Boot Partition (BPINFO.BPSZ);

4. Allocate a physically contiguous memory buffer in the host to store the contents of a Boot Partition;
5. Initialize the address (BPMBL.BMBBA) into the memory buffer where the contents should be copied;
6. Initiate the transfer of data from a Boot Partition by writing to the Boot Partition Read Select (BPRSEL) register. This includes setting the Boot Partition identifier (BPRSEL.BPID), size of Boot Partition Read Size (BPRSEL.BPRSZ) and Boot Partition Read Offset (BPRSEL.BPROF). The controller sets the Boot Read Status (BPINFO.BRS) field while transferring the Boot Partition contents to indicate a Boot Partition read operation is in progress; and
7. Wait for the controller to completely transfer the requested portion of the Boot Partition, indicated in the status field (BPINFO.BRS). If BPINFO.BRS is set to 10b, the requested Boot Partition data has been transferred to the Boot Partition Memory Buffer. If BPINFO.BRS is set to 11b, there was an error transferring the requested Boot Partition data and the host may request the Boot Partition data again.

In constrained memory environments, the host may read the contents of a Boot Partition with a small Boot Partition Memory Buffer by reading a small portion of a Boot Partition, moving the data out of the Boot Memory Buffer to another memory location, and then reading another portion of the Boot Partition until the entire Boot Partition has been read.

If the Boot Partition log page is supported (refer to section 5.14.TBD2 <TP 4071a>), then the Boot Partition can be accessed through the Boot Partition log page (refer to section 5.14.1.TBD).

Modify portions of NVMe-MI 1.1 as shown below:

No modifications needed.