



1
2
3
4

Document Identifier: DSP0235

Date: 2015-03-06

Version: 1.0.0

5
6
7

NVMe™ (NVM Express™) Management Messages over MCTP Binding Specification

8
9
10
11
12

Supersedes: None
Document Type: Specification
Document Class: Normative
Document Status: Published
Document Language: en-US

13

14 Copyright notice

15 Copyright © 2015 Distributed Management Task Force, Inc. (DMTF). All rights reserved.

16 DMTF is a not-for-profit association of industry members dedicated to promoting enterprise and systems
17 management and interoperability. Members and non-members may reproduce DMTF specifications and
18 documents for uses consistent with this purpose, provided that correct attribution is given. As DMTF
19 specifications may be revised from time to time, the particular version and release date should always be
20 noted.

21 Implementation of certain elements of this standard or proposed standard may be subject to third party
22 patent rights, including provisional patent rights (herein "patent rights"). DMTF makes no representations
23 to users of the standard as to the existence of such rights, and is not responsible to recognize, disclose,
24 or identify any or all such third party patent right, owners or claimants, nor for any incomplete or
25 inaccurate identification or disclosure of such rights, owners or claimants. DMTF shall have no liability to
26 any party, in any manner or circumstance, under any legal theory whatsoever, for failure to recognize,
27 disclose, or identify any such third party patent rights, or for such party's reliance on the standard or
28 incorporation thereof in its product, protocols or testing procedures. DMTF shall have no liability to any
29 party implementing such standard, whether such implementation is foreseeable or not, nor to any patent
30 owner or claimant, and shall have no liability or responsibility for costs or losses incurred if a standard is
31 withdrawn or modified after publication, and shall be indemnified and held harmless by any party
32 implementing the standard from any and all claims of infringement by a patent owner for such
33 implementations.

34 PCI-SIG, PCIe, and the PCI HOT PLUG design mark are registered trademarks or service marks of PCI-
35 SIG.

36 All other marks and brands are the property of their respective owners.

37 For information about patents held by third-parties which have notified the DMTF that, in their opinion,
38 such patent may relate to or impact implementations of DMTF standards, visit
39 <http://www.dmtf.org/about/policies/disclosures.php>.

40

41

CONTENTS

42 Foreword 4

43 Acknowledgments 4

44 Introduction..... 5

45 Document conventions..... 5

46 Typographical conventions 5

47 ABNF usage conventions 5

48 1 Scope 6

49 2 Normative references 6

50 3 Terms and definitions 7

51 4 Symbols and abbreviated terms..... 7

52 5 Conventions 8

53 5.1 Reserved and unassigned values..... 8

54 5.2 Byte ordering..... 8

55 6 Overview..... 8

56 7 Message Type-specific considerations 9

57 7.1 Message Type number 9

58 7.2 Supported transport bindings..... 9

59 7.3 MCTP specification versioning and version compatibility..... 9

60 7.3.1 Base specification and control protocol version compatibility 9

61 7.3.2 NVMe Management Messages over MCTP-specific version information 9

62 7.3.3 Packet header version compatibility 9

63 7.4 Timing specifications 10

64 7.5 Encapsulation 10

65 7.6 Maximum message size 10

66 7.6.1 Additional semantics for MCTP fields 10

67 7.7 Multiple MCTP transports 11

68 ANNEX A (informative) Notation and conventions 12

69 ANNEX B (informative) Change log..... 13

70 Bibliography 14

71

72 Figures

73 Figure 1 – Generic MCTP message fields 10

74

75

Foreword

76 The NVMe™ (*NVM Express™*) Management Messages over MCTP Binding Specification (DSP0235)
77 was prepared by the Platform Management Components Intercommunications (PMCI Working Group) of
78 the DMTF.

79 DMTF is a not-for-profit association of industry members dedicated to promoting enterprise and systems
80 management and interoperability.

81 Acknowledgments

82 The DMTF acknowledges the following individuals for their contributions to this document:

83 Editor:

- 84 • Tom Slaight – Intel Corporation

85 Contributors:

- 86 • Patrick Caporale – IBM
- 87 • John Carroll – Intel Corporation
- 88 • Philip Chidester – Dell
- 89 • Yuval Itkin – Mellanox
- 90 • Patrick Kutch – Intel Corporation
- 91 • Myron Loewen – Intel Corporation
- 92 • Eliel Louzoun – Intel Corporation
- 93 • Pat Schoeller – Hewlett-Packard Company
- 94 • Hemal Shah – Broadcom Corporation
- 95 • Bob Stevens – Dell

96

97

Introduction

98 The NVMe™ Messages over MCTP Binding Specification defines a new MCTP message type used to
99 convey NVMe™ Management Messages over MCTP to storage devices.

100 Document conventions

101 Typographical conventions

102 The following typographical conventions are used in this document:

- 103 • Document titles are marked in *italics*.
- 104 • Important terms that are used for the first time are marked in *italics*.
- 105 • Terms include a link to the term definition in the "Terms and definitions" clause, enabling easy
106 navigation to the term definition.
- 107 • ABNF rules are in `monospaced font`.

108 ABNF usage conventions

109 Format definitions in this document are specified using ABNF (see [RFC5234](#)), with the following
110 deviations:

- 111 • Literal strings are to be interpreted as case-sensitive Unicode characters, as opposed to the
112 definition in [RFC5234](#) that interprets literal strings as case-insensitive US-ASCII characters.

113

114 NVMe™ (NVM Express™) Management Messages over MCTP 115 Binding Specification

116 1 Scope

117 The NVMe™ (NVM Express™) Management Messages over MCTP Binding Specification defines the
118 bindings between NVMe Management Interface protocol elements and MCTP elements in order to
119 transport NVMe Management Messages for storage devices using MCTP. The specific NVMe
120 management message contents will be documented outside of DMTF directly by the NVMe Management
121 Interface working group.

122 Portions of this specification rely on information and definitions from other specifications, which are
123 identified in clause 2. The following references are particularly relevant:

- 124 • DMTF [DSP0236](#), *Management Component Transport Protocol (MCTP) Base Specification*,
125 defines the MCTP transport protocol over which the NVMe over MCTP messages are to be
126 conveyed.

127 2 Normative references

128 The following referenced documents are indispensable for the application of this document. For dated or
129 versioned references, only the edition cited (including any corrigenda or DMTF update versions) applies.
130 For references without a date or version, the latest published edition of the referenced document
131 (including any corrigenda or DMTF update versions) applies.

132 Unless otherwise specified, for DMTF documents this means any document version that has minor or
133 update version numbers that are later than those for the referenced document. The major version
134 numbers must match the major version number given for the referenced document.

135 Refer to the Bibliography for additional, non-normative, reference information.

136 DMTF DSP0223, *Generic Operations 1.0*,
137 http://www.dmtf.org/standards/published_documents/DSP0223_1.0.pdf

138 DMTF DSP0236, *Management Component Transport Protocol (MCTP) Base Specification 1.2*
139 http://www.dmtf.org/standards/published_documents/DSP0236_1.2.pdf

140 DMTF DSP0237, *Management Component Transport Protocol (MCTP) SMBus/I2C Transport Binding*
141 *Specification 1.0*
142 http://www.dmtf.org/standards/published_documents/DSP0237_1.0.pdf

143 DMTF DSP0238, *Management Component Transport Protocol (MCTP) PCIe VDM Transport Binding*
144 *Specification 1.0*
145 http://www.dmtf.org/standards/published_documents/DSP0238_1.0.pdf

146 DMTF DSP0239, *Management Component Transport Protocol (MCTP) IDs and Codes 1.2*
147 http://www.dmtf.org/standards/published_documents/DSP0239_1.2.pdf

148 IETF, RFC4122, *A Universally Unique Identifier (UUID) URN Namespace*, July 2005
149 <http://www.ietf.org/rfc/rfc4122.txt>

150 IETF RFC5234, *ABNF: Augmented BNF for Syntax Specifications*, January 2008,
151 <http://tools.ietf.org/html/rfc5234>

152 ISO/IEC Directives, Part 2, *Rules for the structure and drafting of International Standards*,
153 <http://isotc.iso.org/livelink/livelink.exe?func=ll&objId=4230456&objAction=browse&sort=subtype>

154 3 Terms and definitions

155 In this document, some terms have a specific meaning beyond the normal English meaning. Those terms
156 are defined in this clause.

157 The terms "shall" ("required"), "shall not," "should" ("recommended"), "should not" ("not recommended"),
158 "may," "need not" ("not required"), "can" and "cannot" in this document are to be interpreted as described
159 in [ISO/IEC Directives, Part 2](#), Annex H. The terms in parenthesis are alternatives for the preceding term,
160 for use in exceptional cases when the preceding term cannot be used for linguistic reasons. Note that
161 [ISO/IEC Directives, Part 2](#), Annex H specifies additional alternatives. Occurrences of such additional
162 alternatives shall be interpreted in their normal English meaning.

163 The terms "clause," "subclause," "paragraph," and "annex" in this document are to be interpreted as
164 described in [ISO/IEC Directives, Part 2](#), Clause 5.

165 The terms "normative" and "informative" in this document are to be interpreted as described in [ISO/IEC](#)
166 [Directives, Part 2](#), Clause 3. In this document, clauses, subclauses, or annexes labeled "(informative)" do
167 not contain normative content. Notes and examples are always informative elements.

168 Refer to [DSP0236](#) for terms and definitions that are used across the MCTP specifications.

169 Refer to [NVMeMI](#) (see Bibliography) for terms and definitions that are used in the NVMe Express™
170 Management Interface specification.

171 The terms defined in [DSP0223](#), and [DSP1001](#) apply to this document. The following additional terms are
172 used in this document.

173 3.1

174 Endpoint

175 An MCTP endpoint unless otherwise specified.

176 3.2

177 NVM Express™

178 NVM Express is an optimized register interface, command set, and feature set for PCI Express based
179 storage. The NVMe specifications are maintained by NVM Express, Inc.

180 3.3

181 NVMe™ Management Interface

182 The NVMe Management Interface allows management entities to communicate with an NVMe non-
183 volatile memory subsystem over one or more external interfaces.

184 4 Symbols and abbreviated terms

185 The abbreviations defined in [DSP0004](#), [DSP0223](#), and [DSP1001](#) apply to this document. The following
186 additional abbreviations are used in this document.

187 4.1

188 ACPI

189 Advanced Configuration and Power Interface

190 **4.2**
191 **MCTP**
192 Management Component Transport Protocol

193 **4.3**
194 **MC**
195 Management Controller

196 **4.4**
197 **NVMe™**
198 NVM Express

199 **5 Conventions**

200 **5.1 Reserved and unassigned values**

201 Unless otherwise specified, any reserved, unspecified, or unassigned values in enumerations or other
202 numeric ranges are reserved for future definition by the DMTF.

203 Unless otherwise specified, numeric or bit fields that are designated as reserved shall be written as 0
204 (zero) and ignored when read.

205 **5.2 Byte ordering**

206 Unless otherwise specified, byte ordering of multibyte numeric fields or bit fields is "Big Endian" (that is,
207 the lower byte offset holds the most significant byte, and higher offsets hold lesser significant bytes).

208 **6 Overview**

209 Non-Volatile Memory Express (NVMe) is an optimized register interface, command set, and feature set
210 for PCI Express based storage. The NVMe Management Interface protocol may also be used for other
211 types of non-volatile memory devices.

212 NVM Express Management Interface Commands (NVMe Management Interface Commands) are used for
213 the accessing configuration, control, and status functions in NVMe-compatible non-volatile memory
214 devices. NVMe Management Interface Commands are defined by the NVMe Management Interface
215 specification and the members of NVM Express, Inc. Refer to www.nvmexpress.org and [NVMeMI](#) in the
216 Bibliography for more information.

217 This specification only defines how NVMe™ Management Interface Commands are encapsulated in
218 MCTP Messages and transferred between MCTP Endpoints over the specified transports. These are
219 referred to in this document as NVMe Management Messages over MCTP. The definitions and semantics
220 of the NVMe Management Commands themselves are outside the scope of this specification. See the
221 Bibliography for reference to the NVMe Management Interface specification ([NVMeMI](#)).

222 The MCTP Transport Bindings that are used for NVMe Management Messages over MCTP are defined in
223 other companion specifications such as *MCTP SMBus Binding Specification* ([DSP0237](#)) and *MCTP PCIe*
224 *Binding Specification* ([DSP0238](#)).

225 7 Message Type-specific considerations

226 7.1 Message Type number

227 The Message Type number for NVMe Management Messages over MCTP messages is defined in the
228 MCTP IDs and Codes Specification ([DSP0239](#)).

229 7.2 Supported transport bindings

230 As of this writing, use of the specified Message Type is defined for the following transport bindings:

- 231 • MCTP SMBus Binding Specification ([DSP0237](#))
- 232 • MCTP PCIe Binding Specification ([DSP0238](#))

233 7.3 MCTP specification versioning and version compatibility

234 Per [DSP0236](#), the following types of versioning information that can be retrieved using the Get MCTP
235 Version Support command:

- 236 • MCTP base specification version information
- 237 • MCTP control protocol version information
- 238 • NVMe Management Messages over MCTP-specific version information

239 Additionally, the MCTP packet carries the following versioning information:

- 240 • MCTP packet header version information

241 7.3.1 Base specification and control protocol version compatibility

242 Unless otherwise specified herein, NVMe Management Messages over MCTP shall meet the
243 requirements of the base specification and control protocol that are identified by the MCTP base
244 specification and control protocol version information, respectively, that are obtained from the endpoint
245 using the Get MCTP Version Support command.

246 Endpoints that implement NVMe Management Messages over MCTP must also meet the requirements
247 for MCTP Control Messages that are defined by the base specification.

248 7.3.2 NVMe Management Messages over MCTP-specific version information

249 The complete semantics of the differences between versions of NVMe Management Messages is left to
250 the NVM Express Management Interface working group, and is outside the scope of this specification.
251 However, the versioning approach should follow the major/minor/update/alpha convention as defined in
252 the Get MCTP Version Support command in [DSP0236](#).

253 7.3.3 Packet header version compatibility

254 The Header Version field in MCTP packets identifies the media-specific formatting used for MCTP
255 packets. It can also indicate a level of current and backward compatibility with versions of the base
256 specification, as specified by the header version definition in each medium-specific transport binding
257 specification.

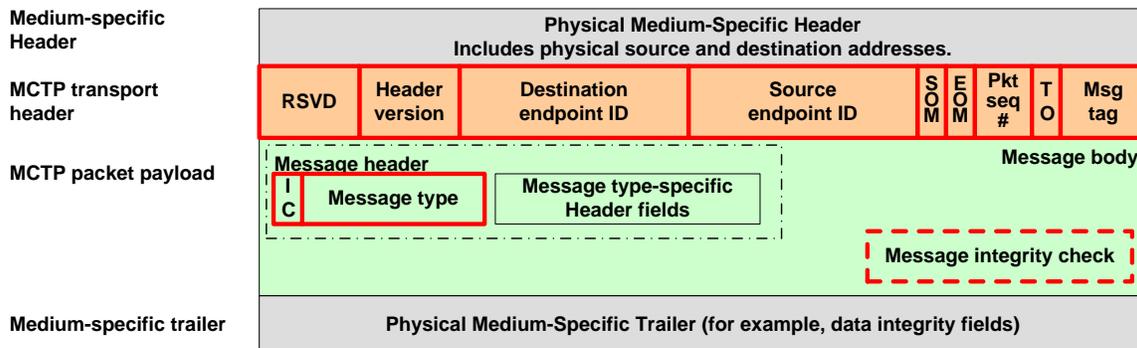
258 Unless otherwise specified herein, NVMe Management Messages over MCTP shall meet the
259 requirements that are associated with the header version value that is used with the NVMe Management
260 Messages over MCTP, as specified by the corresponding MCTP transport binding specification. This
261 includes meeting requirements for any transport-binding-specific MCTP Control Messages that are called
262 out by the particular transport binding specification.

263 **7.4 Timing specifications**

264 NVMe Management Messages over MCTP are made up of one or more MCTP packets. Each MCTP
 265 packet shall comply with the timing, arbitration, and fairness requirements of the transport binding
 266 specifications for the media through which it passes. The MCTP endpoint may choose to negotiate longer
 267 packet payload lengths than the 64-byte baseline (when the longer packets are not blocked by bridges)
 268 provided they do not prevent other devices on the MCTP network from also meeting their transport
 269 binding specification requirements, nor prevent them from meeting the message timing specifications for
 270 their supported message types.

271 **7.5 Encapsulation**

272 Referring to Figure 1, the NVMe Management Messages over MCTP are carried via the MCTP packet
 273 payload of one or more MCTP packets.



274

275 **Figure 1 – Generic MCTP message fields**

276 **7.6 Maximum message size**

277 The MCTP packet payload for NVMe Management Messages over MCTP shall be less than or equal to
 278 4224 (4K+128) bytes.

279 This corresponds to a transfer of 66 MCTP packets using a baseline transmission unit of 64 bytes for the
 280 MCTP packet payload.

281 The maximum message size includes the IC bit and Message Type fields plus any additional Message
 282 Type-specific header fields and Message Integrity check fields, as required by [NVMeMI](#). Refer to
 283 [NVMeMI](#) for any additional restrictions on message sizes.

284 **7.6.1 Additional semantics for MCTP fields**

285 NVMe Management Messages over MCTP shall meet the requirements for the MCTP Message Fields
 286 per [DSP0236](#). Additional semantics, for example whether the Tag Owner bit or Msg Tag field are to be
 287 used to identify particular message streams, or to identify request/response messages, and so on, may
 288 be specified by [NVMeMI](#) as long as such semantics do not conflict with [DSP0236](#) or the transport binding
 289 specifications.

290 **7.7 Multiple MCTP transports**

291 In order to facilitate identification of devices that are accessible via multiple transports, the endpoints in
292 the device **must** support the Get Endpoint UUID MCTP command. Otherwise, this specification does not
293 define any additional behaviors related to communicating with NVM Express™ devices that may be
294 accessed through more than one type of MCTP transport on a given MCTP network.

295
296
297

ANNEX A (informative) Notation and conventions

298 A.1 Notations

299 Examples of notations used in this document are as follows:

- 300 • 2:N In field descriptions, this will typically be used to represent a range of byte offsets
301 starting from byte two and continuing to and including byte N. The lowest offset is on
302 the left; the highest is on the right.
- 303 • (6) Parentheses around a single number can be used in message field descriptions to
304 indicate a byte field that may be present or absent.
- 305 • (3:6) Parentheses around a field consisting of a range of bytes indicates the entire range
306 may be present or absent. The lowest offset is on the left; the highest is on the right.
- 307 • [PCIe](#) Underlined, blue text is typically used to indicate a reference to a document or
308 specification called out in the "Normative references" clause or to items hyperlinked
309 within the document.
- 310 • rsvd This case-insensitive abbreviation is for "reserved."
- 311 • [4] Square brackets around a number are typically used to indicate a bit offset. Bit offsets
312 are given as zero-based values (that is, the least significant bit [LSb] offset = 0).
- 313 • [7:5] This notation indicates a range of bit offsets. The most significant bit is on the left; the
314 least significant bit is on the right.
- 315 • 1b The lowercase "b" following a number consisting of 0s and 1s is used to indicate the
316 number is being given in binary format.
- 317 • 0x12A A leading "0x" is used to indicate a number given in hexadecimal format.

318

319
320
321
322

ANNEX B (informative)

Change log

Version	Date	Description
1.0.0	2015-03-06	

323
324

325

Bibliography

326 NVMeMI NVM Express Inc., NVM Express™ Management Interface Specification 1.0, December
327 2014
328 www.nvmexpress.org

329