



1
2
3
4

Document Identifier: DSP0235

Date: 2018-08-03

Version: 1.0.1

5
6
7

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

8 **Supersedes: 1.0.0**
9 **Document Class: Normative**
10 **Document Status: Published**
11 **Document Language: en-US**
12

13 Copyright notice

14 Copyright © 2015, 2018 Distributed Management Task Force, Inc. (DMTF). All rights reserved.

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

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30 withdrawn or modified after publication, and shall be indemnified and held harmless by any party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

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

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

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

39 This document's normative language is English. Translation into other languages is permitted.

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..... 8

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 10

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 11

67 7.7 Multiple MCTP transports 11

68 ANNEX A (informative) Notation and conventions 12

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

70

71 **Figures**

72 Figure 1 – Generic MCTP message fields 10

73

74

Foreword

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

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

80 Acknowledgments

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

82 Editor:

- 83 • Tom Slaight – Intel Corporation

84 Contributors:

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

95

96

Introduction

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

99 Document conventions

100 Typographical conventions

101 The following typographical conventions are used in this document:

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

107 ABNF usage conventions

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

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

112

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

115 1 Scope

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

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

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

126 2 Normative references

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

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

134 NVM Express, *NVMe Management Interface 1.0a*, [http://nvmexpress.org/wp-
135 content/uploads/NVM_Express_Management_Interface_1_0a_2017.04.08_-_gold.pdf](http://nvmexpress.org/wp-content/uploads/NVM_Express_Management_Interface_1_0a_2017.04.08_-_gold.pdf)

136 DMTF DSP0004, *CIM Infrastructure Specification 2.7*,
137 http://www.dmtf.org/standards/published_documents/DSP0004_2.7.pdf

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

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

142 DMTF DSP0237, *Management Component Transport Protocol (MCTP) SMBus/I2C Transport Binding
143 Specification 1.0*
144 https://www.dmtf.org/sites/default/files/standards/documents/DSP0237_1.0.pdf

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

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

150 DMTF DSP1001, *Management Profile Specification Usage Guide 1.1*,
151 http://www.dmtf.org/standards/published_documents/DSP1001_1.1.pdf

152

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

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

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

159 3 Terms and definitions

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

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

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

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

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

174 Refer to [NVMe-MI](#) for terms and definitions that are used in the NVM Express™ Management Interface
175 specification.

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

178 3.1

179 Endpoint

180 An MCTP endpoint unless otherwise specified.

181 3.2

182 NVM Express™

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

185 3.3

186 NVMe™ Management Interface

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

189 **4 Symbols and abbreviated terms**

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

192 **4.1**

193 **ACPI**

194 Advanced Configuration and Power Interface

195 **4.2**

196 **MCTP**

197 Management Component Transport Protocol

198 **4.3**

199 **MC**

200 Management Controller

201 **4.4**

202 **NVMe™**

203 NVM Express

204 **4.5**

205 **NVMe-MI**

206 NVM Express Management Interface

207

208 **5 Conventions**

209 **5.1 Reserved and unassigned values**

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

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

214 **5.2 Byte ordering**

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

217 **6 Overview**

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

221 NVM Express Management Interface Commands (NVMe Management Interface Commands) are used for
222 the accessing configuration, control, and status functions in NVMe-compatible non-volatile memory
223 devices. NVMe Management Interface Commands are defined by the NVMe Management Interface

224 specification and the members of NVM Express, Inc. Refer to www.nvmexpress.org and [NVMe-MI](#) for
225 more information.

226 This specification only defines how NVMe™ Management Interface Commands are encapsulated in
227 MCTP Messages and transferred between MCTP Endpoints over the specified transports. These are
228 referred to in this document as NVMe Management Messages over MCTP. The definitions and semantics
229 of the NVMe Management Commands themselves are outside the scope of this specification. See the
230 reference to the NVMe Management Interface specification ([NVMe-MI](#)).

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

234 7 Message Type-specific considerations

235 7.1 Message Type number

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

238 7.2 Supported transport bindings

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

- 240 • MCTP SMBus Binding Specification ([DSP0237](#))
- 241 • MCTP PCIe Binding Specification ([DSP0238](#))

242 7.3 MCTP specification versioning and version compatibility

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

- 245 • MCTP base specification version information
- 246 • MCTP control protocol version information
- 247 • NVMe Management Messages over MCTP-specific version information

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

- 249 • MCTP packet header version information

250 7.3.1 Base specification and control protocol version compatibility

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

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

257 7.3.2 NVMe Management Messages over MCTP-specific version information

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

262 **7.3.3 Packet header version compatibility**

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

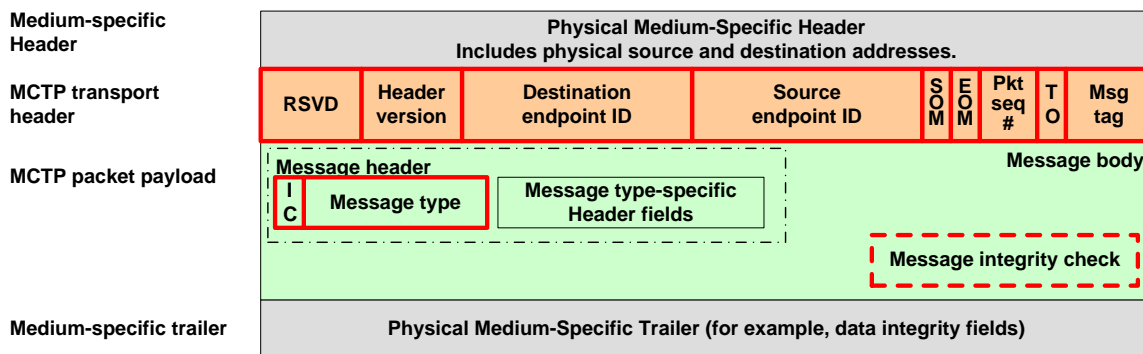
267 Unless otherwise specified herein, NVMe Management Messages over MCTP shall meet the
 268 requirements that are associated with the header version value that is used with the NVMe Management
 269 Messages over MCTP, as specified by the corresponding MCTP transport binding specification. This
 270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

272 **7.4 Timing specifications**

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

280 **7.5 Encapsulation**

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



283

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

285 **7.6 Maximum message size**

286 The MCTP message body (including IC bit, Message Type, Message type-specific header fields,
 287 message payload and message integrity check if present) for NVMe Management Messages over MCTP
 288 shall be less than or equal to 4224 (4K+128) bytes.

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

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

294 **7.6.1 Additional semantics for MCTP fields**

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

300 **7.7 Multiple MCTP transports**

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

305
306
307

ANNEX A (informative) Notation and conventions

308 A.1 Notations

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

- 310 • 2:N In field descriptions, this will typically be used to represent a range of byte offsets
311 starting from byte two and continuing to and including byte N. The lowest offset is on
312 the left; the highest is on the right.
- 313 • (6) Parentheses around a single number can be used in message field descriptions to
314 indicate a byte field that may be present or absent.
- 315 • (3:6) Parentheses around a field consisting of a range of bytes indicates the entire range
316 may be present or absent. The lowest offset is on the left; the highest is on the right.
- 317 • [PCIe](#) Underlined, blue text is typically used to indicate a reference to a document or
318 specification called out in the "Normative references" clause or to items hyperlinked
319 within the document.
- 320 • rsvd This case-insensitive abbreviation is for "reserved."
- 321 • [4] Square brackets around a number are typically used to indicate a bit offset. Bit offsets
322 are given as zero-based values (that is, the least significant bit [LSb] offset = 0).
- 323 • [7:5] This notation indicates a range of bit offsets. The most significant bit is on the left; the
324 least significant bit is on the right.
- 325 • 1b The lowercase "b" following a number consisting of 0s and 1s is used to indicate the
326 number is being given in binary format.
- 327 • 0x12A A leading "0x" is used to indicate a number given in hexadecimal format.

328

329
330
331
332

ANNEX B (informative)

Change log

Version	Date	Description
1.0.0	2015-03-06	
1.0.1	2018-08-03	Corrected the maximum message size text to state message payload not packet payload. Updated references to NVMe-MI.

333
334