



1
2
3
4

Document Number: DSP0236

Date: 2010-04-22

Version: 1.1.0

5 **Management Component Transport Protocol**
6 **(MCTP) Base Specification**
7 **Includes MCTP Control Command Specifications**

8 **Document Type: Specification**
9 **Document Status: DMTF Standard**
10 **Document Language: en-US**
11

12 Copyright Notice

13 Copyright © 2009–2010 Distributed Management Task Force, Inc. (DMTF). All rights reserved.

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

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28 withdrawn or modified after publication, and shall be indemnified and held harmless by any party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

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

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

CONTENTS

39	Foreword	7
40	Introduction	8
41	1 Scope	9
42	2 Normative References.....	9
43	3 Terms and Definitions.....	10
44	3.1 Requirement Term Definitions	10
45	3.2 MCTP Term Definitions.....	12
46	4 Symbols and Abbreviated Terms.....	18
47	5 Conventions	20
48	5.1 Byte Ordering.....	20
49	5.2 Reserved Fields.....	20
50	6 Management Component Relationships.....	20
51	7 MCTP Overview	21
52	8 MCTP Base Protocol.....	23
53	8.1 MCTP Packet Fields	23
54	8.2 Special Endpoint IDs.....	26
55	8.3 Packet Payload and Transmission Unit Sizes	26
56	8.4 Maximum Message Body Sizes.....	27
57	8.5 Message Assembly.....	27
58	8.6 Dropped Packets	27
59	8.7 Starting Message Assembly	28
60	8.8 Terminating Message Assembly/Dropped Messages	28
61	8.9 Dropped Messages.....	29
62	8.10 MCTP Versioning and Message Type Support	29
63	8.11 MCTP Message Types	30
64	8.12 Security	30
65	8.13 Limitations.....	30
66	8.14 MCTP Discovery and Addressing.....	31
67	8.15 Devices with Multiple Media Interfaces.....	32
68	8.16 Peer Transactions.....	32
69	8.17 Endpoint ID Assignment and Endpoint ID Pools	32
70	8.18 Handling Reassigned EIDs.....	37
71	8.19 MCTP Bridging.....	38
72	8.20 Bridge and Routing Table Examples	45
73	8.21 Endpoint ID Resolution	49
74	8.22 Bridge and Bus Owner Implementation Recommendations.....	51
75	8.23 Path and Transmission Unit Discovery.....	52
76	8.24 Path Transmission Unit Requirements for Bridges.....	55
77	9 MCTP Control Protocol	55
78	9.1 Terminology	55
79	9.2 MCTP Control Message Format.....	56
80	9.3 MCTP Control Message Fields.....	56
81	9.4 MCTP Control Message Transmission Unit Size	57
82	9.5 Tag Owner (TO), Request (Rq), and Datagram (D) Bit Usage.....	57
83	9.6 Concurrent Command Processing.....	58
84	10 MCTP Control Messages	59
85	10.1 MCTP Control Message Command Codes	59
86	10.2 MCTP Control Message Completion Codes.....	61
87	10.3 Set Endpoint ID.....	61
88	10.4 Get Endpoint ID	63
89	10.5 Get Endpoint UUID Command	64

90	10.6	Get MCTP Version Support	65
91	10.7	Get Message Type Support	67
92	10.8	Get Vendor Defined Message Support	68
93	10.9	Resolve Endpoint ID	69
94	10.10	Allocate Endpoint IDs	70
95	10.11	Routing Information Update	72
96	10.12	Get Routing Table Entries	73
97	10.13	Prepare for Endpoint Discovery	75
98	10.14	Endpoint Discovery	76
99	10.15	Discovery Notify	76
100	10.16	Query Hop	76
101	10.17	Get Network ID Command	77
102	10.18	Transport Specific	78
103	11	Vendor Defined – PCI and Vendor Defined – IANA Messages	79
104	11.1	Vendor Defined – PCI Message Format	79
105	11.2	Vendor Defined – IANA Message Format	79
106		ANNEX A (informative) Notation	80
107	A.1	Notations	80
108		ANNEX B (informative) Change Log	81
109		Bibliography	82
110			

111 Figures

112		Figure 1 – Management Component Relationships	20
113		Figure 2 – MCTP Networks	21
114		Figure 3 – MCTP Topology	23
115		Figure 4 – Generic Message Fields	23
116		Figure 5 – Topmost Bus Owners	33
117		Figure 6 – Split Bridge	34
118		Figure 7 – Acceptable Failover/Redundant Communication Topologies	38
119		Figure 8 – Routing/Bridging Restrictions	39
120		Figure 9 – EID Options for MCTP Bridges	40
121		Figure 10 – Basic Routing Table Entry Fields	43
122		Figure 11 – Routing Table Population	44
123		Figure 12 – Example 1 Routing Topology	46
124		Figure 13 – Example 2 Routing Topology	47
125		Figure 14 – Example 3 Routing Topology	48
126		Figure 15 – Endpoint ID Resolution	50
127		Figure 16 – Resolving Multiple Paths	51
128		Figure 17 – Example Path Routing Topology	53
129		Figure 18 – Path Transmission Unit Discovery Flowchart	54
130		Figure 19 – MCTP Control Message Format	56
131		Figure 20 – Structure of Vendor ID Field for Get Vendor Defined Capabilities Message	69
132		Figure 21 – EID Pools from Multiple Bus Owners	71
133			

134 **Tables**

135 Table 1 – MCTP Base Protocol Common Fields 24

136 Table 2 – Special Endpoint IDs..... 26

137 Table 3 – MCTP Message Types Used in this Specification 30

138 Table 4 – Example 1 Routing Table for D2..... 46

139 Table 5 – Example 2 Routing Table for D1..... 47

140 Table 6 – Example 3 Routing Table for D2..... 48

141 Table 7 – Additional Information Tracked by Bridges 49

142 Table 8 – MCTP Control Protocol Terminology 55

143 Table 9 – MCTP Control Message Types..... 55

144 Table 10 – MCTP Control Message Fields 56

145 Table 11 – Tag Owner (TO), Request (Rq) and Datagram (D) Bit Usage..... 58

146 Table 12 – MCTP Control Command Numbers 59

147 Table 13 – MCTP Control Message Completion Codes..... 61

148 Table 14 – Set Endpoint ID Message 62

149 Table 15 – Get Endpoint ID Message..... 63

150 Table 16 – Get Device UUID Message Format 65

151 Table 17 – Example UUID Format..... 65

152 Table 18 – Get MCTP Version Support Message 66

153 Table 19 – Get Message Type Support Message 68

154 Table 20 – Get Vendor Defined Message Support Message 68

155 Table 21 – Vendor ID Formats..... 69

156 Table 22 – Resolve Endpoint ID Message..... 70

157 Table 23 – Allocate Endpoint IDs Message 71

158 Table 24 – Routing Information Update Message 73

159 Table 25 – Routing Information Update Entry Format 73

160 Table 26 – Get Routing Table Entries Message..... 74

161 Table 27 – Routing Table Entry Format..... 74

162 Table 28 – Prepare for Endpoint Discovery Message 75

163 Table 29 – Endpoint Discovery Message 76

164 Table 30 – Discovery Notify Message 76

165 Table 31 – Query Hop Message 77

166 Table 32 – Get Network ID Message Format 78

167 Table 33 – Transport Specific Message 78

168 Table 34 – Vendor Defined – PCI Message Format..... 79

169 Table 35 – Vendor Defined – IANA Message Format 79

170

172

Foreword

173 The *Management Component Transport Protocol (MCTP) Base Specification* (DSP0236) was prepared
174 by the PMCI Working Group.

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

177

Introduction

178 The Management Component Transport Protocol (MCTP) defines a communication model intended to
179 facilitate communication between:

- 180 • Management controllers and other management controllers
- 181 • Management controllers and management devices

182 The communication model includes a message format, transport description, message exchange
183 patterns, and configuration and initialization messages.

184 MCTP is designed so that it can potentially be used on many bus types. The protocol is intended to be
185 used for intercommunication between elements of platform management subsystems used in computer
186 systems, and is suitable for use in mobile, desktop, workstation, and server platforms. Management
187 controllers such as a baseboard management controller (BMC) can use this protocol for communication
188 between one another, as well as for accessing management devices within the platform.

189 Management controllers can use this protocol to send and receive MCTP-formatted messages across the
190 different bus types that are used to access management devices and other management controllers.
191 Management devices in a system need to provide an implementation of the message format to facilitate
192 actions performed by management controllers.

193 It is intended that different types of devices in a management system may need to implement different
194 portions of the complete capabilities defined by this protocol. Where relevant, this is called out in the
195 individual requirements

196
197

Management Component Transport Protocol (MCTP) Base Specification

1 Scope

199 The *MCTP Base Specification* describes the command protocol, requirements, and use cases of a
200 transport protocol for communication between discrete management controllers on a platform, as well as
201 between management controllers and the devices they manage.

202 This document is intended to meet the following objectives:

- 203 • Describe the MCTP Base transport protocol
- 204 • Describe the MCTP control message protocol

205 The MCTP specifies a transport protocol format. This protocol is independent of the underlying physical
206 bus properties, as well as the "data-link" layer messaging used on the bus. The physical and data-link
207 layer methods for MCTP communication across a given medium are defined by companion "transport
208 binding" specifications, such as [DSP0238](#), *MCTP over PCIe® Vendor Defined Messaging*, and [DSP0237](#),
209 *MCTP over SMBus/ I²C*. This approach enables future transport bindings to be defined to support
210 [DSP0237](#) additional buses such as USB, RMIII, and others, without affecting the base MCTP
211 specification.

2 Normative References

212 The following referenced documents are indispensable for the application of this document. For dated
213 references, only the edition cited applies. For undated references, the latest edition of the referenced
214 document (including any amendments) applies.

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

218 DMTF DSP0238, *PCIe VDM Transport Binding Specification 1.0*,
219 http://www.dmtf.org/standards/published_documents/DSP0238_1.0.pdf

220 DMTF DSP0239, *Management Component Transport Protocol (MCTP) IDs and Codes 1.0*,
221 http://www.dmtf.org/standards/published_documents/DSP0239_1.0.pdf

222 Hewlett-Packard, Intel, Microsoft, Phoenix, and Toshiba, *Advanced Configuration and Power Interface*
223 *Specification v3.0*, ACPI, September 2, 2004, <http://www.acpi.info/DOWNLOADS/ACPIspec30.zip>

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

226 Intel, Hewlett-Packard, NEC, and Dell, *Intelligent Platform Management Interface Specification: Second*
227 *Generation v2.0*, IPMI, 2004, ftp://download.intel.com/design/servers/ipmi/IPMIv2_0rev1_0markup.pdf

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

230 OMG, *Unified Modeling Language (UML) from the Open Management Group (OMG)*, <http://www.uml.org>

231 PCI-SIG, *PCI Express Base Specification v1.1*, PCIeV1.1, March 28, 2005,
232 http://www.pcisig.com/members/downloads/specifications/pciexpress/PCI_Express_Base_11.pdf

233 PCI-SIG, *PCI Express Base Specification v2.0*, PCIeV2.0, December 20, 2006,
234 http://www.pcisig.com/members/downloads/specifications/pciexpress/PCI_Express_Base_2.pdf

235 Philips Semiconductors, *The I²C-Bus Specification v2.0*, I²C, December 1998

236 SMBus, *System Management Bus (SMBus) Specification v2.0*, SMBus, 2000,
237 <http://www.smbus.org/specs/smbus20.pdf>

238 **3 Terms and Definitions**

239 For the purposes of this document, the following terms and definitions apply.

240 **3.1 Requirement Term Definitions**

241 This clause defines key phrases and words that denote requirement levels in this specification.

242 **3.1.1**

243 **can**

244 used for statements of possibility and capability, whether material, physical, or causal

245 **3.1.2**

246 **cannot**

247 used for statements of possibility and capability, whether material, physical or causal

248 **3.1.3**

249 **conditional**

250 indicates requirements to be followed strictly to conform to the document when the specified conditions
251 are met

252 **3.1.4**

253 **deprecated**

254 indicates that an element or profile behavior has been outdated by newer constructs

255 **3.1.5**

256 **mandatory**

257 indicates requirements to be followed strictly to conform to the document and from which no deviation is
258 permitted

259 **3.1.6**

260 **may**

261 indicates a course of action permissible within the limits of the document

262 NOTE: An implementation that does *not* include a particular option must be prepared to interoperate with another
263 implementation that *does* include the option, although perhaps with reduced functionality. An implementation that
264 *does* include a particular option must be prepared to interoperate with another implementation that does *not* include
265 the option (except for the feature that the option provides).

266 **3.1.7**

267 **may not**

268 indicates flexibility of choice with no implied preference

- 269 **3.1.8**
270 **must**
271 indicates that the item is an absolute requirement of the specification
- 272 **3.1.9**
273 **must not**
274 indicates that the definition is an absolute prohibition of the specification
- 275 **3.1.10**
276 **need not**
277 indicates a course of action permissible within the limits of the document
- 278 **3.1.11**
279 **not recommended**
280 indicates that valid reasons may exist in particular circumstances when the particular behavior is
281 acceptable or even useful, but the full implications should be understood and carefully weighed before
282 implementing any behavior described with this label
- 283 **3.1.12**
284 **obsolete**
285 indicates that an item was defined in prior specifications but has been removed from this specification
- 286 **3.1.13**
287 **optional**
288 indicates a course of action permissible within the limits of the document
- 289 **3.1.14**
290 **recommended**
291 indicates that valid reasons may exist in particular circumstances to ignore a particular item, but the full
292 implications should be understood and carefully weighed before choosing a different course
- 293 **3.1.15**
294 **required**
295 indicates that the item is an absolute requirement of the specification
- 296 **3.1.16**
297 **shall**
298 indicates requirements to be followed strictly to conform to the document and from which no deviation is
299 permitted
- 300 **3.1.17**
301 **shall not**
302 indicates requirements to be followed strictly to conform to the document and from which no deviation is
303 permitted
- 304 **3.1.18**
305 **should**
306 indicates that among several possibilities, one is recommended as particularly suitable, without
307 mentioning or excluding others, or that a certain course of action is preferred but not necessarily required
- 308 **3.1.19**
309 **should not**
310 indicates that a certain possibility or course of action is deprecated but not prohibited

311 3.2 MCTP Term Definitions

312 For the purposes of this document, the following terms and definitions apply.

313 3.2.1

314 Address Resolution Protocol

315 ARP

316 refers to the procedure used to dynamically determine the addresses of devices on a shared
317 communication medium

318 3.2.2

319 baseline transmission unit

320 the required common denominator size of a transmission unit for packet payloads that are carried in an
321 MCTP packet. Baseline Transmission Unit-sized packets are guaranteed to be routable within an MCTP
322 network.

323 3.2.3

324 baseboard management controller

325 BMC

326 a term coined by the IPMI specifications for the main management controller in an IPMI-based platform
327 management subsystem. Also sometimes used as a generic name for a motherboard resident
328 management controller that provides motherboard-specific hardware monitoring and control functions for
329 the platform management subsystem.

330 3.2.4

331 binary-coded decimal

332 BCD

333 indicates a particular binary encoding for decimal numbers where each four bits (*nibble*) in a binary
334 number is used to represent a single decimal digit, and with the least significant four bits of the binary
335 number corresponding to the least significant decimal digit. The binary values 0000b through 1001b
336 represent decimal values 0 through 9, respectively. For example, with BCD encoding a byte can
337 represent a two-digit decimal number where the most significant nibble (bits 7:4) of the byte contains the
338 encoding for the most significant decimal digit and the least significant nibble (bits 3:0) contains the
339 encoding for the least significant decimal digit (for example, 0010_1001b in BCD encoding corresponds to
340 the decimal number 29).

341 3.2.5

342 bridge

343 generically, the circuitry and logic that connects one computer bus or interconnect to another, allowing an
344 agent on one to access the other. Within this document, the term *bridge* shall refer to MCTP bridge,
345 unless otherwise indicated.

346 3.2.6

347 bus

348 a physical addressing domain shared between one or more platform components that share a common
349 physical layer address space

350 3.2.7

351 bus owner

352 the party responsible for managing address assignments (can be logical or physical addresses) on a bus
353 (for example, in MCTP, the bus owner is the party responsible for managing EID assignments for a given
354 bus). A bus owner may also have additional media-specific responsibilities, such as assignment of
355 physical addresses.

356 **3.2.8**357 **byte**

358 an 8-bit quantity. Also referred to as an *octet*.

359 NOTE: PMCI specifications shall use the term *byte*, not *octet*.

360 **3.2.9**361 **endpoint**

362 see [MCTP endpoint](#)

363 **3.2.10**364 **endpoint ID**365 **EID**

366 see [MCTP endpoint ID](#)

367 **3.2.11**368 **Globally Unique Identifier**369 **GUID**

370 see [UUID](#)

371 **3.2.12**372 **host interface**

373 a hardware interface and associated protocols that is used by software running locally on the host
374 processors to access the hardware of a management subsystem within a managed system.

375 **3.2.13**376 **Inter-Integrated Circuit**377 **I²C**

378 a multi-master, two-wire, serial bus originally developed by Philips Semiconductor

379 **3.2.14**380 **intelligent management device**381 **IMD**

382 a management device that is typically implemented using a microcontroller and accessed through a
383 messaging protocol. Management parameter access provided by an IMD is typically accomplished using
384 an abstracted interface and data model rather than through direct "register level" accesses.

385 **3.2.15**386 **Intelligent Platform Management Bus**387 **IPMB**

388 name for the architecture, protocol, and implementation of an I²C bus that provides a communications
389 path between "management controllers" in IPMI -based systems

390 **3.2.16**391 **Intelligent Platform Management Interface**392 **IPMI**

393 a set of specifications defining interfaces and protocols originally developed for server platform
394 management by the IPMI Promoters Group: Intel, Dell, HP, and NEC

395 **3.2.17**396 **managed entity**

397 the physical or logical entity that is being managed through management parameters. Examples of
398 *physical* entities include fans, processors, power supplies, circuit cards, chassis, and so on. Examples of
399 *logical* entities include virtual processors, cooling domains, system security states, and so on.

400 **3.2.18**401 **Management Component Transport Protocol**402 **MCTP**

403 The protocol defined in this specification.

404 **3.2.19**405 **management controller**

406 a microcontroller or processor that aggregates management parameters from one or more management
407 devices and makes access to those parameters available to local or remote software, or to other
408 management controllers, through one or more management data models. Management controllers may
409 also interpret and process management-related data, and initiate management-related actions on
410 management devices. While a native data model is defined for PMCI, it is designed to be capable of
411 supporting other data models, such as CIM, IPMI, and vendor-specific data models. The microcontroller
412 or processor that serves as a management controller can also incorporate the functions of a management
413 device.

414 **3.2.20**415 **management device**

416 any physical device that provides protocol terminus for accessing one or more management parameters.
417 A management device responds to management requests, but does not initiate or aggregate
418 management operations except in conjunction with a management controller (that is, it is a *satellite*
419 device that is subsidiary to one or more management controllers). An example of a simple management
420 device would be a temperature sensor chip. A management controller that has I/O pins or built-in analog-
421 to-digital converters that monitor state and voltages for a managed entity would also be a management
422 device.

423 **3.2.21**424 **management parameter**

425 a particular datum representing a characteristic, capability, status, or control point associated with a
426 managed entity. Example management parameters include temperature, speed, volts, on/off, link state,
427 uncorrectable error count, device power state, and so on.

428 **3.2.22**429 **MCTP bridge**

430 an MCTP endpoint that can route MCTP messages not destined for itself that it receives on one
431 interconnect onto another without interpreting them. The ingress and egress media at the bridge may be
432 either homogeneous or heterogeneous. Also referred to in this document as a "bridge".

433 **3.2.23**434 **MCTP bus owner**

435 responsible for EID assignment for MCTP or translation on the buses that it is a master of. The MCTP bus
436 owner may also be responsible for physical address assignment. For example, for SMBus bus segments,
437 the MCTP bus owner is also the ARP master. This means the bus owner assigns dynamic SMBus
438 addresses to those devices requiring it.

439 **3.2.24**440 **MCTP control command**

441 commands defined under the MCTP *control* message type that are used for the initialization and
442 management of MCTP communications (for example, commands to assign EIDs, discover device MCTP
443 capabilities, and so on)

444 **3.2.25**445 **MCTP endpoint**

446 an MCTP communication terminus. An MCTP endpoint is a terminus or origin of MCTP packets or
447 messages. That is, the combined functionality within a physical device that communicates using the
448 MCTP transport protocol and handles MCTP control commands. This includes MCTP-capable
449 management controllers and management devices. Also referred to in this document as "endpoint".

450 **3.2.26**451 **MCTP endpoint ID**

452 the logical address used to route MCTP messages to a specific MCTP endpoint. A numeric handle
453 (logical address) that uniquely identifies a particular MCTP endpoint within a system for MCTP
454 communication and message routing purposes. Endpoint IDs are unique among MCTP endpoints that
455 comprise an MCTP communication network within a system. MCTP EIDs are only unique within a
456 particular MCTP network. That is, they can be duplicated or overlap from one MCTP network to the next.
457 Also referred to in this document as "endpoint ID" and abbreviated "EID".

458 **3.2.27**459 **MCTP host interface**

460 a host interface that enables host software to locally access an MCTP Network in the managed system.

461 **3.2.28**462 **MCTP management controller**

463 a management controller that is an MCTP endpoint. Unless otherwise indicated, the term "management
464 controller" refers to an "MCTP management controller" in this document.

465 **3.2.29**466 **MCTP management device**

467 a management device that is an MCTP endpoint. Unless otherwise indicated, the term "management
468 device" refers to an "MCTP management device" in this document.

469 **3.2.30**470 **MCTP message**

471 a unit of communication based on the message type that is relayed through the MCTP Network using one
472 or more MCTP packets

473 **3.2.31**474 **MCTP network**

475 a collection of MCTP endpoints that communicate using MCTP and share a common MCTP endpoint ID
476 space

477 **3.2.32**478 **MCTP network ID**

479 a unique identifier to distinguish each independent MCTP network within a platform

480 **3.2.33**481 **MCTP packet**

482 the unit of data transfer used for MCTP communication on a given physical medium

- 483 **3.2.34**
484 **MCTP packet payload**
485 refers to the portion of the message body of an MCTP message that is carried in a single MCTP packet
- 486 **3.2.35**
487 **message**
488 see [MCTP message](#)
- 489 **3.2.36**
490 **message assembly**
491 the process of receiving and linking together two or more MCTP packets that belong to a given MCTP
492 message to allow the entire message header and message data (payload) to be extracted
- 493 **3.2.37**
494 **message body**
495 the portion of an MCTP message that carries the message type field and any message type-specific data
496 associated with the message. An MCTP message spans multiple MCTP packets when the message body
497 needs is larger than what can fit in a single MCTP packet. Thus, the message body portion of an MCTP
498 message can span multiple MCTP packets.
- 499 **3.2.38**
500 **message disassembly**
501 the process of taking an MCTP message where the message's header and data (payload) cannot be
502 carried in a single MCTP packet and generating the sequence of two or more packets required to deliver
503 that message content within the MCTP network
- 504 **3.2.39**
505 **message originator**
506 the original transmitter (source) of a message targeted to a particular message terminus
- 507 **3.2.40**
508 **message terminus**
509 the name for a triplet of fields called the MCTP Source Endpoint ID, Tag Owner bit value, and Message
510 Tag value. Together, these fields identify the packets for an MCTP message within an MCTP network for
511 the purpose of message assembly. The message terminus itself can be thought of as identifying a set of
512 resources within the recipient endpoint that is handling the assembly of a particular message.
- 513 **3.2.41**
514 **most significant byte**
515 **MSB**
516 refers to the highest order byte in a number consisting of multiple bytes
- 517 **3.2.42**
518 **nibble**
519 the computer term for a four-bit aggregation, or half of a byte
- 520 **3.2.43**
521 **packet**
522 see [MCTP packet](#)
- 523 **3.2.44**
524 **packet payload**
525 see [MCTP packet payload](#)

- 526 **3.2.45**
527 **pass-through traffic/message/packets**
528 non-control packets passed between the external network and the management controller through the
529 network controller
- 530 **3.2.46**
531 **payload**
532 refers to the information bearing fields of a message. This is separate from those fields and elements that
533 are used to transport the message from one point to another, such as address fields, framing bits,
534 checksums, and so on. In some instances, a given field may be both a payload field and a transport field.
- 535 **3.2.47**
536 **physical transport binding**
537 refers to specifications that define how the MCTP base protocol and MCTP control commands are
538 implemented on a particular physical transport type and medium, such as SMBus/I²C, PCI Express™
539 Vendor Defined Messaging, and so on.
- 540 **3.2.48**
541 **Platform Management Component Intercommunications**
542 **PMCI**
543 name for a working group under the Distributed Management Task Force's Pre-OS Workgroup that is
544 chartered to define standardized communication protocols, low level data models, and transport
545 definitions that support communications with and between management controllers and management
546 devices that form a platform management subsystem within a managed computer system
- 547 **3.2.49**
548 **point-to-point**
549 refers to the case where only two physical communication devices are interconnected through a physical
550 communication medium. The devices may be in a master/slave relationship, or could be peers.
- 551 **3.2.50**
552 **Reduced Media Independent Interface**
553 **RMII**
554 a reduced signal count MAC to PHY interface, based on the IEEE Media Independent Interface (MII),
555 which was specified by the RMII Consortium (3Com Corporation; AMD Inc.; Bay Networks, Inc.;
556 Broadcom Corp.; National Semiconductor Corp.; and Texas Instruments Inc.)
- 557 **3.2.51**
558 **simple endpoint**
559 an MCTP endpoint that is not associated with either the functions of an MCTP bus owner or an MCTP
560 bridge
- 561 **3.2.52**
562 **Transmission Unit**
563 refers to the size of the portion of the MCTP packet payload, which is the portion of the message body
564 carried in an MCTP packet
- 565 **3.2.53**
566 **transport binding**
567 see [physical transport binding](#)

568 **3.2.54**

569 **Universally Unique Identifier**

570 **UUID**

571 refers to an identifier originally standardized by the Open Software Foundation (OSF) as part of the
572 Distributed Computing Environment (DCE). UUIDs are created using a set of algorithms that enables
573 them to be independently generated by different parties without requiring that the parties coordinate to
574 ensure that generated IDs do not overlap. In this specification, [RFC4122](#) is used as the base specification
575 describing the format and generation of UUIDs. Also sometimes referred to as a globally unique identifier
576 (GUID).

577 **4 Symbols and Abbreviated Terms**

578 The following symbols and abbreviations are used in this document.

579 **4.1**

580 **ACPI**

581 Advanced Configuration and Power Interface

582 **4.2**

583 **ARP**

584 Address Resolution Protocol

585 **4.3**

586 **BCD**

587 binary-coded decimal

588 **4.4**

589 **BMC**

590 baseboard management controller

591 **4.5**

592 **CIM**

593 Common Information Model

594 **4.6**

595 **EID**

596 endpoint identifier

597 **4.7**

598 **FIFO**

599 first-in first-out

600 **4.8**

601 **GUID**

602 Globally Unique Identifier

603 **4.9**

604 **I²C**

605 Inter-Integrated Circuit

606 **4.10**

607 **IANA**

608 Internet Assigned Numbers Authority

609	4.11
610	IMD
611	intelligent management device
612	4.12
613	IP
614	Internet Protocol
615	4.13
616	IPMB
617	Intelligent platform management bus
618	4.14
619	IPMI
620	Intelligent platform management interface
621	4.15
622	ISO/IEC
623	International Organization for Standardization/International Engineering Consortium
624	4.16
625	KCS
626	Keyboard Controller Style
627	4.17
628	MCTP
629	Management Component Transport Protocol
630	4.18
631	MSB
632	most significant byte
633	4.19
634	PCIe
635	Peripheral Component Interconnect (PCI) Express
636	4.20
637	PMCI
638	Platform Management Component Intercommunications
639	4.21
640	RMII
641	Reduced Media Independent Interface
642	4.22
643	SMBus
644	System Management Bus
645	4.23
646	TCP/IP
647	Transmission Control Protocol/Internet Protocol

648 **4.24**
 649 **USB**
 650 Universal Serial Bus
 651 **4.25**
 652 **UUID**
 653 Universally Unique Identifier
 654 **4.26**
 655 **VDM**
 656 Vendor Defined Message

657 **5 Conventions**

658 The conventions described in the following clauses apply to this specification.

659 **5.1 Byte Ordering**

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

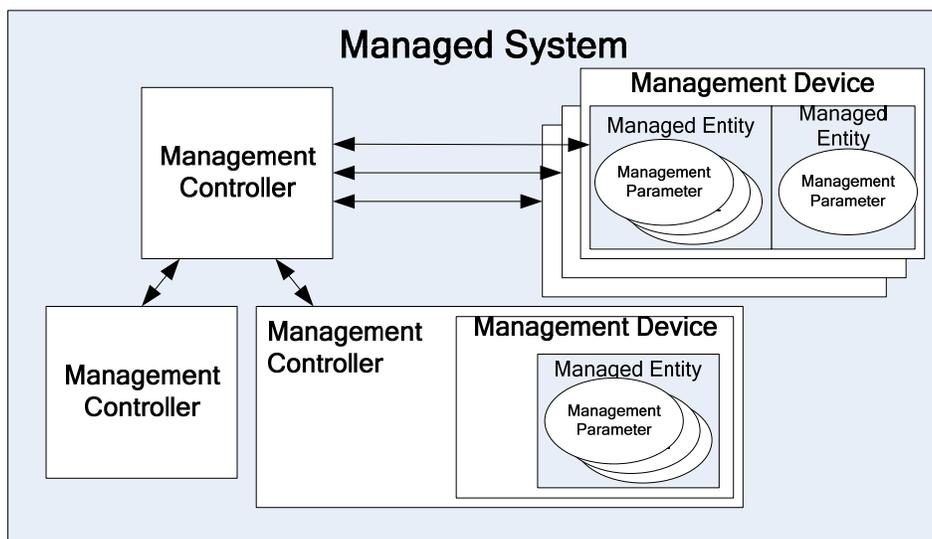
662 **5.2 Reserved Fields**

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

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

667 **6 Management Component Relationships**

668 Figure 1 illustrates the relationship between devices, management controllers, management devices, and
 669 managed entities, which are described in Clause 3.2.



670

671

Figure 1 – Management Component Relationships

672 **7 MCTP Overview**

673 This clause provides an overview of the main elements of MCTP. Additional overview information is
 674 available in the MCTP white paper, [DSP2016](#).

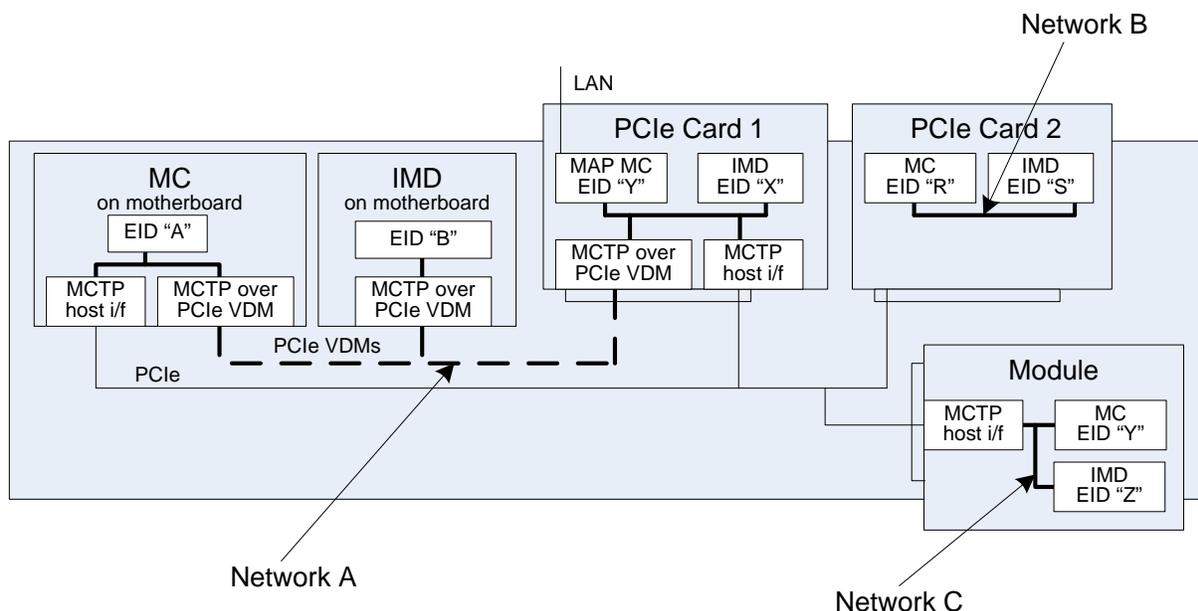
675 MCTP is a transport independent protocol that is used for intercommunication within an MCTP Network.
 676 An MCTP Network that consists of one or more physical transports that are used to transfer MCTP
 677 Packets between MCTP Endpoints. MCTP Transport Binding Specifications define how the MCTP
 678 protocol is implemented across a particular physical transport medium. For example, the DMTF has
 679 defined transport bindings for MCTP over SMBus/I²C and MCTP over PCIe using PCIe Vendor Defined
 680 Messages (VDMs), and others.

681 An MCTP Endpoint is the terminus for MCTP communication. A physical device that supports MCTP may
 682 provide one or more MCTP Endpoints. Endpoints are addressed using a logical address called the
 683 Endpoint ID, or EID. EIDs in MCTP are analogous to IP Addresses in Internet Protocol networking. EIDs
 684 can be statically or dynamically allocated.

685 A system implementation can contain multiple MCTP Networks. Each MCTP Network has its own
 686 separate EID space. There is no coordination of EIDs between MCTP Networks. EIDs can overlap
 687 between MCTP Networks.

688 An MCTP Network may provide an MCTP Network ID that can be used to differentiate different MCTP
 689 Networks when more than one MCTP Network can be accessed by an entity such as system software.
 690 The Network ID is also used when an entity has more than one point of access to the MCTP Network. In
 691 this case, the MCTP Network ID enables the entity to tell whether the access points provide access to the
 692 same MCTP Network or to different MCTP Networks.

693 MCTP also includes the definition of MCTP host interfaces. MCTP host interfaces are used by software
 694 that runs locally on the host processors of the managed system to access an MCTP Network.
 695



696

697

Figure 2 – MCTP Networks

698 Figure 2 shows the different ways MCTP Networks can exist in a system. In this example, Network A
699 connects a Management Controller (MC) and Intelligent Management Device (IMD) on a motherboard
700 with devices on PCIe Card 1 using MCTP over PCIe Vendor Defined Messages. Note that there are two
701 host interfaces (host i/f) on standard PCIe that can be used by host software to access this particular
702 network. This network thus requires an MCTP Network ID so that the host software can tell that the two
703 host interfaces connect to the same MCTP Network.

704 Network B represents a network that is solely used for interconnecting devices within PCIe Card 2. This
705 MCTP Network would typically not require an MCTP Network ID since it is not visible to host software or
706 any other entity that would need to differentiate Network B from another MCTP Network in the system.

707 Network C represents an MCTP Network on an add-in module. This network is separate from networks A
708 and B but can be accessed by host software through PCIe. Thus, this network requires a Network ID so
709 that host software can differentiate that Network C is a different network than Network A.

710 MCTP Messages are comprised of one or more MCTP Packets. MCTP defines fields that support the
711 assembly of received MCTP Packets into MCTP Messages and the disassembly of MCTP Messages into
712 packets for transmission.

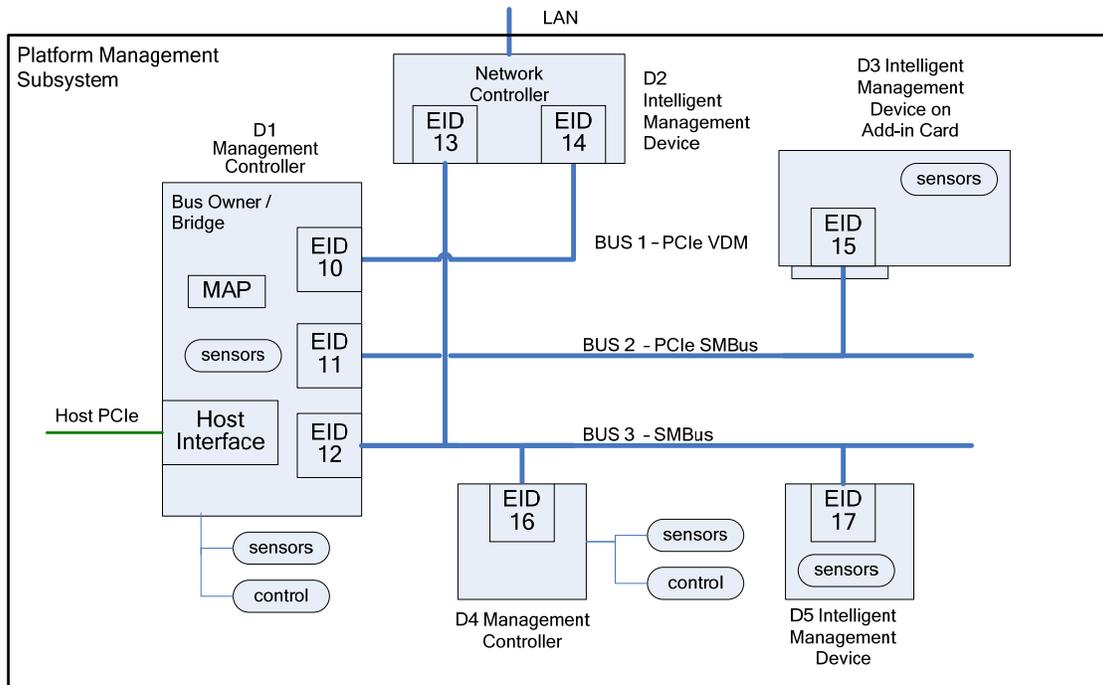
713 MCTP is designed to be able to transfer multiple Message Types in an interleaved manner using the
714 same protocol. MCTP Message Types identified using a Message Type number. The use of the message
715 type number is similar to a well-known port number in Internet Protocol. It identifies MCTP Messages that
716 are all associated with a particular specification. This specification defines a Message Type for MCTP
717 Control Messages that are used to initialize and maintain the MCTP Network. The DMTF has also defined
718 Message Types for use by the PMCI (Platform Management Communications Interconnect)
719 specifications, Vendor-specific Messaging over MCTP, and so on. Additional message types for functions
720 such as implementing NC-SI (network controller sideband interface) communications over MCTP and
721 others are expected in the future. MCTP Message Type number assignments are provided in [DSP0239](#).

722 MCTP Control Messages use a request/response protocol. It is important to note that the base transport
723 protocol defined by MCTP just defines a protocol for the transport of MCTP messages. Whether the
724 message content is a request, a response, or something else is part of the particular Message Type
725 definition.

726 In MCTP, a Bus is defined as a physical medium that shares a single physical address space. MCTP
727 includes the definition of a function called the MCTP Bus Owner. The Bus Owner provides to main
728 functions. It distributes EIDs to Endpoints when the MCTP implementation uses EIDs that are dynamically
729 allocated, and it provides the way for an Endpoint to resolve an EID into the physical address used that is
730 required to deliver a message to the target Endpoint.

731 Busses can be interconnected within an MCTP Network using MCTP Bridges to forward MCTP packets
732 between busses. Bridges also handle the task of managing the difference in moving packets from one
733 type of physical media to another, such as moving an MCTP packet between SMBus and PCIe.

734 The following example illustrates how MCTP can be used within a hypothetical platform management
735 subsystem implementation. More complex topologies, with multi-levels of bridges and greater numbers of
736 busses and devices can be readily supported by MCTP as required.



737

738

Figure 3 – MCTP Topology

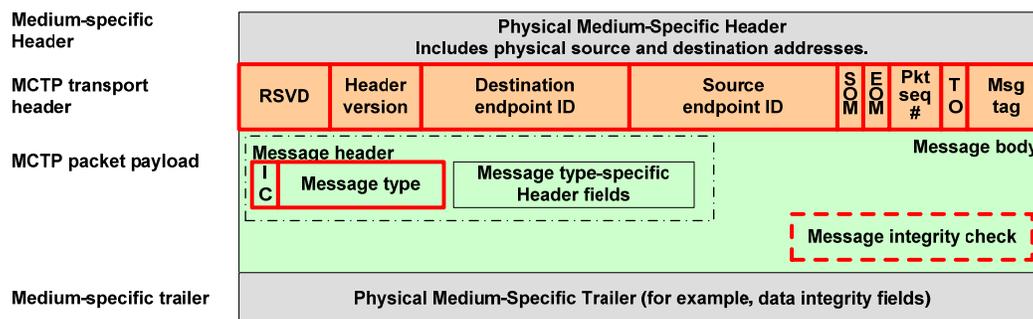
739 8 MCTP Base Protocol

740 The MCTP base protocol defines the common fields for MCTP packets and messages and their usage.

741 Though there are medium-specific packet header fields and trailer fields, the fields for the base protocol
 742 are common for all media. These common fields support the routing and transport of messages between
 743 MCTP endpoints and the assembly and disassembly of large messages from and into multiple MCTP
 744 packets, respectively. The base protocol's common fields include a message type field that identifies what
 745 particular higher layer class of message is being carried using the MCTP base protocol.

746 8.1 MCTP Packet Fields

747 Figure 4 shows the fields that constitute a generic MCTP packet.



748

749

Figure 4 – Generic Message Fields

750 Table 1 defines the base protocol common fields.

751 **Table 1 – MCTP Base Protocol Common Fields**

Field Name	Field Size	Description
Medium-specific header	see description	This field represents the physical addressing and framing information that is used for transferring MCTP packets between devices on a particular physical medium. The size and type of any sub-fields or data within this field are defined by the corresponding transport binding specification for MCTP messaging on a given medium (for example, MCTP over SMBus, MCTP over PCIe, and so on).
Medium-specific trailer	see description	This field represents any additional medium-specific trailer fields (if any) that are required for transferring MCTP packets between devices on a particular physical medium. A typical use of this field would be to hold per-packet data integrity fields (for example CRC, checksum, and so on) that would be specified for the particular medium.
MCTP transport header	32 bits	The MCTP transport header is part of each MCTP packet and provides version and addressing information for the packet as well as flags and a "Message Tag" field that, in conjunction with the source EID, is used to identify packets that constitute an MCTP message. The MCTP transport header fields are common fields that are always present regardless of the physical medium over which MCTP is being used. Note: The positioning of the sub-fields of the MCTP transport header may vary based on the physical medium binding.
RSVD	4 bits	(Reserved) Reserved for future definition by the MCTP base specification.
Hdr version	4 bits	(Header version) Identifies the format, physical framing, and data integrity mechanism used to transfer the MCTP common fields in messages on a given physical medium. The value is defined in the specifications for the particular medium. Note: The value in this field can vary between different transport bindings.
Destination endpoint ID	8 bits	The EID for the endpoint to receive the MCTP packet. A few EID values are reserved for specific routing. See Table 2 – Special Endpoint IDs.
Source endpoint ID	8 bits	The EID of the originator of the MCTP packet. See Table 2 – Special Endpoint IDs.
SOM	1 bit	(Start Of Message) Set to 1b if this packet is the first packet of a message.
EOM	1 bit	(End Of Message) Set to 1b if this packet is the last packet of a message.
Pkt Seq #	2 bits	(Packet sequence number) For messages that span multiple packets, the packet sequence number increments modulo 4 on each successive packet. This allows the receiver to detect up to three successive missing packets between the start and end of a message. Though the packet sequence number can be any value (0-3) if the SOM bit is set, it is recommended that it is an increment modulo 4 from the prior packet with an EOM bit set. After the SOM packet, the packet sequence number must increment modulo 4 for each subsequent packet belonging to a given message up through the packet containing the EOM flag.

Field Name	Field Size	Description
Msg tag	3 bits	<p>(Message tag) Field that, along with the Source Endpoint IDs and the Tag Owner (TO) field, identifies a unique message at the MCTP transport level. Whether other elements, such as portions of the MCTP Message Data field, are also used for uniquely identifying instances or tracking retries of a message is dependent on the message type.</p> <p>A source endpoint is allowed to interleave packets from multiple messages to the same destination endpoint concurrently, provided that each of the messages has a unique message tag.</p> <p>For messages that are split up into multiple packets, the Tag Owner (TO) and Message Tag bits remain the same for all packets from the SOM through the EOM.</p>
TO	1 bit	<p>The TO (Tag Owner) bit identifies whether the message tag was originated by the endpoint that is the source of the message or by the endpoint that is the destination of the message. The Message Tag field is generated and tracked independently for each value of the Tag Owner bit. MCTP message types may overlay this bit with additional meaning, for example using it to differentiate between "request" messages and "response" messages.</p> <p>Set to 1b to indicate that the source of the message originated the message tag.</p>
Message body	See description	The message body represents the payload of an MCTP message. The message body can span multiple MCTP packets.
IC	1 bit	<p>(MCTP integrity check bit) Indicates whether the MCTP message is covered by an overall MCTP message payload integrity check. This field is required to be the most significant bit of the first byte of the message body in the first packet of a message along with the message type bits.</p> <p>0b = No MCTP message integrity check 1b = MCTP message integrity check is present</p>
Message type	7 bits	Defines the type of payload contained in the message data portion of the MCTP message. This field is required to be contained in the least-significant bits of the first byte of the message body in the first packet of a message. Like the fields in the MCTP transport header, the message type field is one of the common MCTP fields that are present independent of the transport over which MCTP is being used. Unlike the MCTP transport header, however, the message type field is only required to be present in the first packet of a particular MCTP message, whereas the MCTP transport header fields are present in every MCTP packet.
Message header	0 to M bytes	Additional header information associated with a particular message type, if any. This will typically only be contained in the first packet of a message, but a given message type definition can define header fields as required for any packet.
Message data	0 to N bytes	Data associated with the particular message type. Defined according to the specifications for the message type.
MCTP packet payload	See description	The packet payload is the portion of the message body that is carried in a given MCTP packet. The packet payload is limited according to the rules governing packet payload and transfer unit sizes. See 8.3 for more information.

Field Name	Field Size	Description
Msg integrity check	Message type-specific	(MCTP message integrity check) This field represents the optional presence of a message type-specific integrity check over the contents of the message body. If present, the Message integrity check field must be carried in the last bytes of the message body. The particular message type definition will specify whether this is required, optional, or not to be used, the field size, and what algorithm is to be used to generate the field. The MCTP base protocol does not specify whether this field is required on single packet messages (potentially dependent on transmission unit size) or is only required on multiple packet messages. This is left to the message type specification.

752 8.2 Special Endpoint IDs

753 Table 2 lists EID values that are reserved or assigned to specific functions for MCTP.

754 **Table 2 – Special Endpoint IDs**

Value	Description
Destination endpoint ID 0	Null Destination EID. This value indicates that the destination EID value is to be ignored and that only physical addressing is used to route the message to the destination on the given bus. This enables communication with devices that have not been assigned an EID. Because the physical addresses between buses are not guaranteed to be unique, MCTP does not support bridging messages with a null destination EID between different buses.
Source endpoint ID 0	Null Source EID. This value indicates a message is coming from an endpoint that is using physical addressing only. This would typically be used for messages that are delivered from an endpoint that has not been assigned an EID. Because the physical addresses between buses are not guaranteed to be unique, MCTP does not support bridging messages with a null source EID between different buses.
Endpoint IDs 1 through 7	Reserved for future definition.
Endpoint ID 0xFF	Broadcast EID. Reserved for use as a broadcast EID on a given bus. MCTP network-wide broadcasts are not supported. Primarily for use by the MCTP control message type.
All other values	Available for assignment and allocation to endpoints.

755 8.3 Packet Payload and Transmission Unit Sizes

756 For MCTP, the size of a transmission unit is defined as the size of the packet payload that is carried in an
757 MCTP packet.

758 8.3.1 Baseline Transmission Unit

759 The following are key information points regarding baseline transmission unit:

- 760 • The baseline transmission unit (minimum transmission unit) size for MCTP is 64 bytes.
- 761 • A message terminus that supports MCTP control messages must always accept valid packets
762 that have a transmission unit equal to or less than the baseline transmission unit. The message
763 terminus is also allowed to support larger transmission units.
- 764 • The transmission unit of all packets in a given message must be the same size, except for the
765 transmission unit in the last packet (packet with EOM bit = 1b). Except for the last packet, this
766 size must be at least the baseline transmission unit size.

- 767 • The size of the transmission unit in the last packet must be less than or equal to the
768 transmission unit size used for the other packets (if any).
- 769 • If a transmission unit size larger than the baseline transmission unit is negotiated, the
770 transmission unit of all packets must be less than or equal to the negotiated transmission unit
771 size. (The negotiation mechanism for larger transmission units between endpoints is message
772 type-specific and is not addressed in this specification.)
- 773 • A given endpoint may negotiate additional restrictions on packet sizes for communication with
774 another endpoint, as long as the requirements of this section are met.
- 775 • All message types must include support for being delivered using packets that have a
776 transmission unit that is no larger than the baseline transmission unit. This is required to support
777 bridging those messages in implementations where there are MCTP bridges that only support
778 the baseline transmission unit.

779 **8.4 Maximum Message Body Sizes**

780 The Message Body can span multiple packets. Limitations on message body sizes are message type-
781 specific and are documented in the specifications for each message type.

782 **8.5 Message Assembly**

783 The following fields (and *only* these fields) are collectively used to identify the packets that belong to a
784 given message for the purpose of message assembly on a particular destination endpoint.

- 785 • Msg Tag (Message Tag)
- 786 • TO (Tag Owner)
- 787 • Source Endpoint ID

788 As described in 3.2, together these values identify the message terminus on the destination endpoint. For
789 a given message terminus, only one message assembly is allowed to be in process at a time.

790 **8.6 Dropped Packets**

791 Individual packets are dropped (silently discarded) by an endpoint under the following conditions. These
792 packets are discarded before being checked for acceptance or rejection for message assembly.
793 Therefore, these packets will *not* cause a message assembly to be started or terminated.

- 794 • **Unexpected "middle" packet or "end" packet**

795 A "middle" packet (SOM flag = 0 and EOM flag = 0) or "end" packet (SOM flag = 0 and EOM
796 flag = 1) for a multiple-packet message is received for a given message terminus without first
797 having received a corresponding "start" packet (where the "start" packet has SOM flag = 1 and
798 EOM flag = 0) for the message.

- 799 • **Bad packet data integrity or other physical layer error**

800 A packet is dropped at the physical data-link layer because a data integrity check on the packet
801 at that layer was invalid. Other possible physical layer errors may include framing errors, byte
802 alignment errors, packet sizes that do not meet the physical layer requirements, and so on.

- 803 • **Bad, unexpected, or expired message tag**

804 A message with TO bit = 0 was received, indicating that the destination endpoint was the
805 originator of the tag value, but the destination endpoint did not originate that value, or is no
806 longer expecting it. (MCTP bridges do not check message tag or TO bit values for messages)

807 that are not addressed to the bridge's EID, or to the bridge's physical address if null-source or
808 destination-EID physical addressing is used.)

809 • **Unknown destination EID**

810 A packet is received at the physical address of the device, but the destination EID does not
811 match the EID for the device or the EID is un-routable.

812 • **Un-routable EID**

813 An MCTP bridge receives an EID that the bridge is not able to route (for example, because the
814 bridge did not have a routing table entry for the given endpoint).

815 • **Bad header version**

816 The MCTP header version (Hdr Version) value is not a value that the endpoint supports.

817 • **Unsupported transmission unit**

818 The transmission unit size is not supported by the endpoint that is receiving the packet.

819 8.7 Starting Message Assembly

820 Multiple-packet message assembly begins when the endpoint corresponding to the destination EID in the
821 packet receives a valid "start" packet (packet with SOM = 1b and EOM = 0b).

822 A packet with both SOM = 1b and EOM = 1b is considered to be a single-packet message, and is not
823 assembled per se.

824 Both multiple- and single-packet messages are subject to being terminated or dropped based on
825 conditions listed in the following section.

826 8.8 Terminating Message Assembly/Dropped Messages

827 Message assembly is terminated at the destination endpoint and messages are accepted or dropped
828 under the following conditions:

829 • **Receipt of the "end" packet for the given message**

830 Receiving an "end" packet (packet with EOM = 1b) for a message that is in the process of being
831 assembled on a given message terminus will cause the message assembly to be completed
832 (provided that the message has not been terminated for any of the reasons listed below). This is
833 normal termination. The message is considered to be accepted at the MCTP base protocol
834 level.

835 • **Receipt of a new "start" packet**

836 Receiving a new "start" packet (packet with SOM = 1b) for a message to the same message
837 terminus as a message assembly already in progress will cause the message assembly in
838 process to be terminated. All data for the message assembly that was in progress is dropped.
839 The newly received start packet is not dropped, but instead it begins a new message assembly.
840 This is considered an error condition.

841 • **Timeout waiting for a packet**

842 Too much time occurred between packets of a given multiple-packet message. The timeout
843 interval is specific to a particular medium. All data for the message assembly that was in
844 progress are dropped. This is considered an error condition.

- 845 • **Out-of-sequence packet sequence number**
- 846 For packets comprising a given multiple-packet message, the packet sequence number for the
847 most recently received packet is not a mod 4 increment of the previously received packet's
848 sequence number. All data for the message assembly that was in progress is dropped. This is
849 considered an error condition.
- 850 • **Incorrect transmission unit**
- 851 An implementation may terminate message assembly if it receives a "middle" packet (SOM =
852 0b and EOM = 0b) where the MCTP packet payload size does not match the MCTP packet
853 payload size for the start packet (SOM = 1b and EOM bit = 0b). This is considered an error
854 condition.
- 855 • **Bad message integrity check**
- 856 For single- or multiple-packet messages that use a message integrity check, a mismatch with
857 the message integrity check value can cause the message assembly to be terminated and the
858 entire message to be dropped, unless it is overridden by the specification for a particular
859 message type.
- 860 NOTE: The message integrity check is considered to be at the message-type level error condition rather
861 than an error at the MCTP base protocol level.

862 **8.9 Dropped Messages**

- 863 An endpoint may drop a message if the message type is not supported by the endpoint. This can happen
864 in any one of the following ways:
- 865 • The endpoint can elect to not start message assembly upon detecting the invalid message type
866 in the first packet.
- 867 • The endpoint can elect to terminate message assembly in process.
- 868 • The endpoint can elect to drop the message after it has been assembled.

869 **8.10 MCTP Versioning and Message Type Support**

870 There are three types of versioning information that can be retrieved using MCTP control messages:

- 871 • MCTP base specification version information
- 872 • MCTP packet header version information
- 873 • Message type version information

874 The version of the MCTP base specification that is supported by a given endpoint is obtained through the
875 Get MCTP Version Support command. This command can also be used to discover whether a particular
876 message type is supported on an endpoint, and if so, what versions of that message type are supported.

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

881 **8.10.1 Compatibility with Future Versions of MCTP**

882 An Endpoint may choose to support only certain versions of MCTP. The command structure along with
883 the Get MCTP Version Support command allows endpoints to detect and restrict the versions of MCTP

884 used by other communication endpoints. To support this, all endpoints on a given medium are required to
 885 implement MCTP Version 1.0 control commands for initialization and version support discovery.

886 8.11 MCTP Message Types

887 Table 3 defines the values for the Message Type field for different message types transported through
 888 MCTP. The MCTP control message type is specified within this document. Baseline requirements for the
 889 Vendor Defined – PCI and Vendor Defined – IANA message types are also specified within this
 890 document. All other message types are specified in the [MCTP ID](#) companion document to this
 891 specification.

892 NOTE: A device that supports a given message type may not support that message type equally across all buses
 893 that connect to the device.

894 **Table 3 – MCTP Message Types Used in this Specification**

Message Type	Message Type Code	Description
MCTP control	0x00	Messages used to support initialization and configuration of MCTP communication within an MCTP network. The messages and functions for this message type are defined within this specification.
Vendor Defined – PCI	0x7E	Message type used to support VDMs where the vendor is identified using a PCI-based vendor ID. The specification of the initial message header bytes for this message type is provided within this specification. Otherwise, the message body content is specified by the vendor, company, or organization identified by the given vendor ID.
Vendor Defined – IANA	0x7F	Message type used to support VDMs where the vendor is identified using an IANA-based vendor ID. (This format uses an "enterprise number" that is assigned and maintained by the Internet Assigned Numbers Authority, www.iana.org , as the means of identifying a particular vendor, company, or organization.) The specification of the initial message header bytes for this message type is provided within this specification. Otherwise, the message body content is specified by the vendor, company, or organization identified by the given vendor ID.

895 8.12 Security

896 The basic premise of MCTP is that higher layer protocols will fulfill security requirements (for example,
 897 confidentiality and authentication) for communication of management data. This means that the data
 898 models carried by MCTP must fulfill the security requirements of a given management transaction. The
 899 MCTP protocol itself will not define any additional security mechanisms.

900 8.13 Limitations

901 MCTP has been optimized for communications that occur within a single computer system platform. It has
 902 not been designed to handle problems that can typically occur in a more generic inter-system networking
 903 environment. In particular, compared to networking protocols such as IP and TCP/IP, MCTP has the
 904 following limitations:

- 905 • MCTP has limited logical addressing. MCTP been optimized for the small number of endpoints
 906 that are expected to be utilized within the platform. The 8-bit range of EIDs is limited compared
 907 to the ranges available for IP addresses.

- 908 • MCTP assumes an MCTP network implementation that does not include loops. There is no
909 mechanism defined in MCTP to detect or reconcile implementations that have connections that
910 form routing loops.
 - 911 • MCTP assumes a network topology where all packets belonging to a given message will be
912 delivered through the same route (that is, MCTP does not generally support some packets for a
913 message arriving by one route, while other packets for the message arrive by a different route).
 - 914 • MCTP does not support out-of-order packets for message assembly.
 - 915 • The MCTP base protocol does not address flow control or congestion control. These behaviors,
916 if required, are specified at the physical transport binding level or at the message type or higher
917 level.
 - 918 • MCTP is not specified to handle duplicate packets at the base protocol message assembly
919 level. If a duplicate packet is received and passed on to MCTP message assembly, it can cause
920 the entire message assembly to be terminated.
- 921 NOTE: Transport bindings are not precluded from including mechanisms for handling duplicate packets
922 at the physical transport level.

923 8.14 MCTP Discovery and Addressing

924 This section describes how MCTP endpoints and their capabilities are discovered by one another, and
925 how MCTP endpoints are provisioned with the addresses necessary for MCTP communication.

926 MCTP discovery occurs over the course of several discrete, ordered steps:

- 927 1) Bus enumeration
- 928 2) Bus address assignment
- 929 3) MCTP capability discovery
- 930 4) Endpoint ID assignment
- 931 5) Distribution and use of routing information

932 This section gives an overview of the methods used for accomplishing each of these steps in various
933 operational scenarios. Clause 10 gives details on the messages used to implement these operations.

934 8.14.1 Bus Enumeration

935 This step represents existing bus enumeration. (The actions taken in this step are specific to a given
936 medium.) Because enumeration of devices on the physical bus is medium-specific, this information is
937 provided in the transport binding specification for the medium.

938 8.14.2 Bus Address Assignment

939 MCTP endpoints require a bus address that is unique to a given bus segment. This step deals with
940 assignment of these addresses. Some bus types (such as PCIe) have built-in mechanisms to effectively
941 deal with this. Others (such as SMBus) require some additional consideration. Because bus address
942 assignment is medium-specific, this information is provided in the transport binding specification for the
943 medium.

944 8.14.3 MCTP Capability Discovery

945 Capability discovery deals with the discovery of the characteristics of individual MCTP endpoints.
946 Capabilities that can be discovered include what message types are supported by an endpoint and what
947 message type versions are supported. See 8.10 for a description of the methods used to accomplish
948 capability discovery.

949 **8.14.4 Endpoint ID Assignment**

950 Endpoint IDs are system-wide unique IDs for identifying a specific MCTP endpoint. They can be
951 dynamically assigned at system startup or hot-plug insertion. See 8.17 for a description of the methods
952 used to accomplish EID assignment.

953 **8.14.5 Distribution and Use of Routing Information**

954 Bridging-capable MCTP endpoints need routing information to identify the next hop to forward a message
955 to its final destination. See 8.19 for a description of how routing information is conveyed between MCTP
956 endpoints.

957 **8.15 Devices with Multiple Media Interfaces**

958 MCTP fully supports management controllers or management devices that have interfaces on more than
959 one type of bus. For example, a device could have both a PCI Express (PCIe) and an SMBus interface. In
960 this scenario, the device will typically have a different EID for each interface. (Bridges can include
961 instantiations that have an endpoint shared across multiple interfaces; see 8.19.2 for more information.)

962 This concept can be useful in different operational scenarios of the managed system. For example,
963 typically a PCIe interface will be used during [ACPI](#) "S0" power states (when the system is fully powered
964 up), which will provide significantly higher bandwidths, whereas the SMBus interface could be used for
965 "S3–S5" low-power sleep states.

966 The baseline transmission unit is specified to be common across all media, enabling packets to be routed
967 between different media without requiring bridges to do intermediate assembly and disassembly
968 operations to handle differences in packet payload sizes between different media.

969 **8.16 Peer Transactions**

970 Endpoints can intercommunicate in a peer-to-peer manner using the physical addressing on a given bus.

971 A special value for the EID is used in cases when the physical address is known, but the EID is not
972 known. This capability is used primarily to support device discovery and EID assignment. A device that
973 does not yet have an EID assignment is not addressed using an EID. Rather, the device gets its EID
974 assigned using an MCTP control command, Set Endpoint ID, which uses physical addressing only.

975 Similarly, depending on the transport binding, a device can also announce its presence by sending an
976 MCTP message to a well-known physical address for the bus owner (for example, for PCIe VDM, this
977 would be the root complex; for SMBus, the host slave address, and so on).

978 It is important to note that in cases where two endpoints are on the same bus, they do not need to go
979 through a bridge to communicate with each other. Devices use the Resolve Endpoint ID command to ask
980 the bus owner what physical address should be used to route messages to a given EID. Depending on
981 the bus implementation, the bus owner can either return the physical address of the bridge that the
982 message should be delivered to, or it can return the physical address of the peer on the bus.

983 **8.17 Endpoint ID Assignment and Endpoint ID Pools**

984 MCTP EIDs are the system-wide unique IDs used by the MCTP infrastructure to address endpoints and
985 for routing messages across multiple buses in the system. There is one EID assigned to a given physical
986 address. Most intelligent management devices (IMDs) or management controllers will connect to just a
987 single bus and have a single EID. A non-bridge device that is connected to multiple different buses will
988 have one EID for each bus it is attached to.

989 Bus owners are MCTP devices that are responsible for issuing EIDs to devices on a bus segment. These
990 EIDs come from a pool of EIDs maintained by the bus owner.

991 With the exception of the topmost bus owner (see 8.17.1), a given bus owner's pool of EIDs is
 992 dynamically allocated at run-time by the bus owner of the bus above it in the hierarchy. Hot-plug devices
 993 must have their EID pools dynamically allocated.

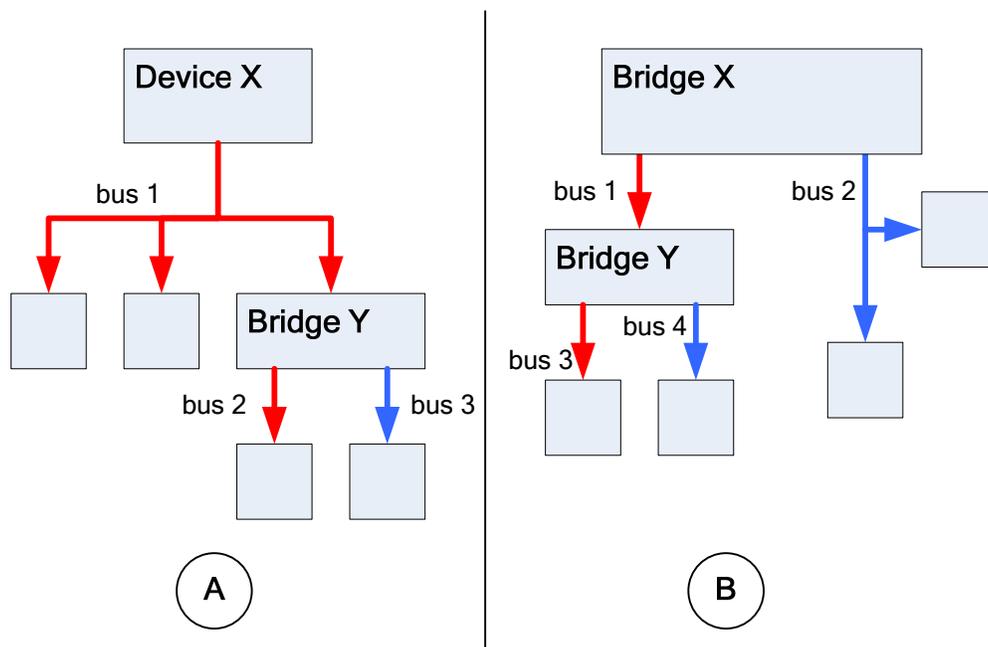
994 Once EIDs are assigned to MCTP endpoints, it is necessary for MCTP devices involved in a transaction
 995 to understand something about the route a given message will traverse. Section 8.19 describes how this
 996 routing information is shared among participants along a message's route.

997 **8.17.1 Topmost Bus Owner**

998 The topmost bus owner is the ultimate source of the EID pool from which all EIDs are drawn for a given
 999 MCTP network.

1000 This is illustrated in Figure 5, in which the arrows are used to identify the role of bus ownership. The
 1001 arrows point outward from the bus owner for the particular bus and inward to a device that is "owned" on
 1002 the bus.

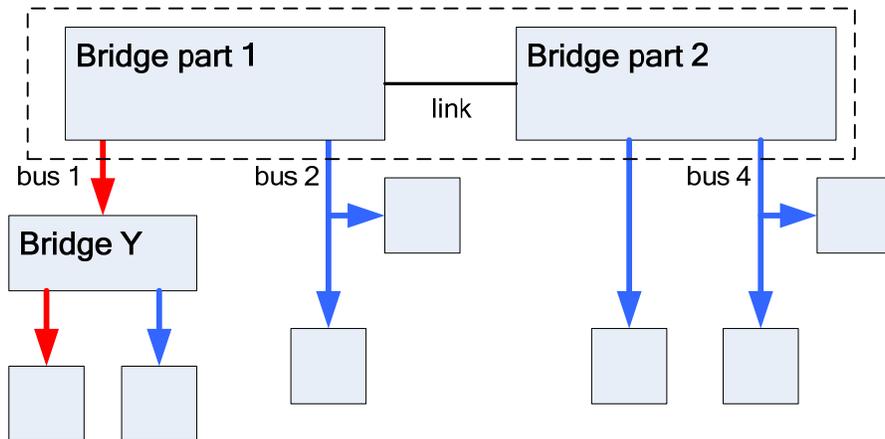
1003 In Figure 5, device X in diagram A and bridge X in diagram B are examples of topmost bus owners.
 1004 Diagram A shows a device that connects to a single bus and is the topmost bus owner for the overall
 1005 MCTP network. Diagram B shows that a bridge can simultaneously be the topmost bus owner, as well as
 1006 the bus owner for more than one bus. The different colors represent examples of different media.



1007

1008 **Figure 5 – Topmost Bus Owners**

1009 An implementation may need to split a bus owner or bridge across two physical devices. Such an
 1010 implementation must include a mechanism (for example, a link as shown in Figure 6) that enables the two
 1011 parts to share a common routing table, or have individual copies of the routing table that are kept
 1012 synchronized. The definition of this mechanism is outside the scope of this specification.



1013

1014

Figure 6 – Split Bridge

1015 8.17.2 Use of Static EIDs and Static EID Pools

1016 In general, the only device that will require a static (pre-configured) EID assignment will be the topmost
 1017 bus owner. It needs a static EID because there is no other party to assign it an EID through MCTP.
 1018 Otherwise, all other devices will have their EIDs assigned to them by a bus owner.

1019 The same principle applies if the device functions as an MCTP bridge. If the device is the highest device
 1020 in the MCTP bus hierarchy, it will require a static pool of EIDs to be assigned to it as part of the system
 1021 design. Otherwise, the device will be dynamically allocated a pool of EIDs from a higher bus owner.

1022 An MCTP network implementation is allowed to use static EIDs for devices other than the topmost bus
 1023 owner. Typically, this would only be done for very simple MCTP networks. Other key EID assignment
 1024 considerations follow:

- 1025 • Endpoints that support the option of being configured for one or more static EIDs must also
 1026 support being configured to be dynamically assigned EIDs.
- 1027 • No mechanism is defined in the MCTP base specification for a bridge or bus owner to discover
 1028 and incorporate a static EID into its routing information. Thus, a simple endpoint that is
 1029 configured with a static EID must also be used with a bus owner that is configured to support
 1030 the static EIDs for the endpoint.
- 1031 • All bus owners/bridges in the hierarchy, from the topmost bus owner to the endpoint, must have
 1032 their routing configurable to support static EID routing information.
- 1033 • Although an endpoint that uses a static EID must be used with a bus owner that supports static
 1034 EIDs, the reverse is not true. A bus owner that uses static EIDs does not need to require that
 1035 the devices on the buses it owns be configured with static EIDs.
- 1036 • How the configuration of static EIDs occurs is outside the scope of this specification.
- 1037 • No specified mechanism exists to "force" an override of a bridge's or bus owner's routing table
 1038 entries for static EIDs. That is, commands such as Allocate Endpoint IDs and Routing
 1039 Information Update only affect entries that are associated with dynamic EIDs.

- 1040 • MCTP does not define a mechanism for keeping routing tables updated if static EIDs are used
1041 with dynamic physical addresses. That is, static EIDs are not supported for use with dynamic
1042 physical addresses.
- 1043 • Bridges can have a mix of both static and dynamic EID pools. That is, the routing table can
1044 have both static and dynamic entries and can allocate from static and dynamic EID pools. Only
1045 the dynamic EID pool is given to the bridge by the bus owner using the Allocate Endpoint IDs
1046 command. There is no specification for how a static EID pool gets configured or how a bridge
1047 decides whether to give an endpoint an EID from a static or dynamically obtained EID pool.
1048 There is also no MCTP-defined mechanism to read the static EID pool setting from the bridge.
- 1049 • MCTP bridges and bus owners (except the topmost bus owner) are not required to include
1050 support for static EIDs.
- 1051 • MCTP does not define a mechanism for allocating EID pools that take static EID assignments
1052 into account. That is, a bridge cannot request a particular set of EIDs to be allocated to it.
- 1053 • MCTP bridges/bus owners may be configurable to use only static EIDs.

1054 8.17.3 Use of Static Physical Addresses

1055 In many simple topologies, it is desirable to use devices that have statically configured physical
1056 addresses. This can simplify the implementation of the device. For example, an SMBus device that is not
1057 used in a hot-plug application would not need to support the SMBus address assignment (SMBus ARP)
1058 protocol. Fixed addresses can also aid in identifying the location and use of an MCTP device in a system.
1059 For example, if a system has two otherwise identical MCTP devices, a system vendor will know that the
1060 device at address "X" is the one at the front of the motherboard, and the device at address "Y" is at the
1061 back, because that is how they assigned the addresses when the system was designed.

1062 Therefore, MCTP transport bindings, such as for SMBus, are allowed to support devices being at static
1063 physical addresses without requiring the binding to define a mechanism that enables the bus owner to
1064 discover MCTP devices that are using static addresses.

1065 In this case, the bridge or bus owner must have a-priori knowledge of the addresses of those devices to
1066 be able to assign EIDs to those devices and to support routing services for those devices. To support this
1067 requirement, the following requirements and recommendations are given to device vendors:

- 1068 • Devices that act as bus owners or bridges and are intended to support MCTP devices that use
1069 static physical addresses should provide a non-volatile configuration option that enables the
1070 system integrator to configure which device addresses are being used for devices on each bus
1071 that is owned by the bridge/bus owner.
- 1072 • The mechanism by which this non-volatile configuration occurs is specific to the device vendor.
1073 In many cases, the physical address information will be kept in some type of non-volatile
1074 storage that is associated with the device and gets loaded when the device is manufactured or
1075 when the device is integrated into a system. In other cases, this information may be coded into
1076 a firmware build for the device.

1077 8.17.4 Endpoint ID Assignment Process for Bus Owners/Bridges

1078 The bus owner/bridge must get its own EID assignment, and a pool of EIDs, as follows. These steps only
1079 apply to bus owner/bridge devices that are not the topmost bus owner.

- 1080 • Bus owners/bridges must be pre-configured with non-volatile information that identifies which
1081 buses they own. (How this configuration is accomplished is device/vendor specific and is
1082 outside the scope of this specification.)

- 1083 • The bus owner/bridge announces its presence on any buses *that it does not own* to get an EID
1084 assignment for that bus. The mechanism by which this announcement occurs is dependent on
1085 the particular physical transport binding and is defined as part of the binding specification.
- 1086 • The bus owner/bridge waits until it gets its own EID assignment for one of those buses through
1087 the Set Endpoint ID command.
- 1088 • The bus owner/bridge indicates the size of the EID pool it requires by returning that information
1089 in the response to the Set Endpoint ID command.
- 1090 • For each bus where the bus owner/bridge is itself an "owned" device, the bus owner/bridge will
1091 be offered a pool of EIDs by being sent an Allocate Endpoint IDs command from the bus owner.
- 1092 • The bus owner/bridge accepts allocations only from the bus of the "first" bus owner that gives it
1093 the allocation, as described in the Allocate Endpoint IDs command description in 8.10. If it gets
1094 allocations from other buses, they are rejected.

1095 The bus owner can now begin to build a routing table for each of the buses that it owns, and accept
1096 routing information update information. Refer to 8.19 for more information.

1097 **8.17.5 Reclaiming EIDs from Hot-Plug Devices**

1098 Bridges will typically have a limited pool of EIDs from which to assign and allocate to devices. (This also
1099 applies when a single bus owner supports hot-plug devices.) It is important for bridges to reclaim EIDs so
1100 that when a device is removed, the EID can later be re-assigned when a device is plugged in. Otherwise,
1101 the EID pool could become depleted as devices are successively removed and added.

1102 EIDs for endpoints that use static addresses are not reclaimed.

1103 No mechanism is specified in the MCTP base protocol for detecting device removal when it occurs.
1104 Therefore, the general approach to detecting whether a device has been removed is to re-enumerate the
1105 bus when a new device is added and an EID or EID pool is being assigned to that device.

1106 The following approach can be used to detect removed hot-plug devices: The bus owner/bridge can
1107 detect a removed device or devices by validating the EIDs that are presently allocated to endpoints that
1108 are directly on the bus and identifying which EIDs are missing. It can do this by attempting to access each
1109 endpoint that the bridge has listed in its routing table as being a device that is directly on the particular
1110 bus. Attempting to access each endpoint can be accomplished by issuing the Get Endpoint ID command
1111 to the physical address of each device and comparing the returned result to the existing entry in the
1112 routing table. If there is no response to the command, or if there is a mismatch with the existing routing
1113 information, the entry should be cleared and the corresponding EID or EID range should be returned to
1114 the "pool" for re-assignment. The bus owner/bridge can then go through the normal steps for EID
1115 assignment.

1116 This approach should work for all physical transport bindings, because it keeps the "removed EID"
1117 detection processing separated from the address assignment process for the bus.

1118 In some cases, a hot-plug endpoint may temporarily go into a state where it does not respond to MCTP
1119 control messages. Depending on the medium, it is possible that when the endpoint comes back on line, it
1120 does not request a new EID assignment but instead continues using the EID it had originally assigned. If
1121 this occurs while the bus owner is validating EIDs to see if any endpoints are no longer accessible, it is
1122 possible that the bus owner will assume that the endpoint was removed and reassign its EID to a newly
1123 inserted endpoint, unless other steps are taken:

- 1124 • The bus owner must wait at least T_{RECLAIM} seconds before reassigning a given EID (where
1125 T_{RECLAIM} is specified in the physical transport binding specification for the medium used to
1126 access the endpoint).

- 1127 • Reclaimed EIDs must only be reassigned after all unused EIDs in the EID pool have been
1128 assigned to endpoints. Optionally, additional robustness can be achieved if the bus owner
1129 maintains a short FIFO list of reclaimed EIDs (and their associated physical addresses) and
1130 allocates the older EIDs first.
- 1131 • A bus owner shall confirm that an endpoint has been removed by attempting to access it after
1132 $T_{RECLAIM}$ has expired. It can do this by issuing a Get Endpoint ID command to the endpoint to
1133 verify that the endpoint is still non-responsive. It is recommended that this be done at least three
1134 times, with a delay of at least $1/2 * T_{RECLAIM}$ between tries if possible. If the endpoint continues
1135 to be non-responsive, it can be assumed that it is safe to return its EID to the pool of EIDs
1136 available for assignment.

1137 8.17.6 Additional Requirements for Hot-Plug Endpoints

1138 Devices that are hot-plug must support the Get Endpoint UUID command. The purpose of this
1139 requirement is to provide a common mechanism for identifying when devices have been changed.

1140 Endpoints that go into states where they temporarily do not respond to MCTP control messages shall
1141 re-announce themselves and request a new EID assignment if they are "off line" for more than $T_{RECLAIM}$
1142 seconds, where $T_{RECLAIM}$ is specified in the physical transport binding specification for the medium used
1143 to access the endpoint.

1144 8.17.7 Additional Requirements for Devices with Multiple Endpoints

1145 A separate EID is utilized for each MCTP bus that a non-bridge device connects to. In many cases, it is
1146 desirable to be able to identify that the same device is accessible through multiple EIDs.

1147 If an endpoint has multiple physical interfaces (ports), the interfaces can be correlated to the device by
1148 using the MCTP Get Endpoint UUID command (see 10.5) to retrieve the unique system-wide identifier.

1149 Devices connected to multiple buses must support the Get Endpoint UUID command for each endpoint
1150 and return a common UUID value across all the endpoints. This is to enable identifying EIDs as belonging
1151 to the same physical device.

1152 8.18 Handling Reassigned EIDs

1153 Though unlikely, it is still possible that during the course of operation of an MCTP network, a particular
1154 EID could get reassigned from one endpoint to another. For example, this could occur if a newly hot-swap
1155 inserted endpoint device gets assigned an EID that was previously assigned to a device that was
1156 subsequently removed.

1157 Under this condition, it is possible that the endpoint could receive a message that was intended for the
1158 previously installed device. This is not considered an issue for MCTP control messages because the
1159 control messages are typically just used by bus owners and bridges for initializing and maintaining the
1160 MCTP network. The bus owners and bridges are aware of the EIDs they have assigned to endpoints and
1161 are thus intrinsically aware of any EID reassignment.

1162 Other endpoints, however, are not explicitly notified of the reassignment of EIDs. Therefore,
1163 communication that occurs directly from one endpoint to another is subject to the possibility that the EID
1164 could become assigned to a different device in the middle of communication. This must be protected
1165 against by protocols specific to the message type being used for the communication.

1166 In general, the approach to protecting against this will be that other message types will require some kind
1167 of "session" to be established between the intercommunicating endpoints. By default, devices would not
1168 start up with an active session. Thus, if a new device is added and it gets a reassigned EID, it will not
1169 have an active session with the other device and the other device will detect this when it tries to
1170 communicate.

1171 The act of having a new EID assigned to an existing device should have the same effect. That is, if a
 1172 device gets a new EID assignment, it would "close" any active sessions for other message types.

1173 The mechanism by which other message types would establish and track communication sessions
 1174 between devices is not specified in this document. It is up to the specification of the particular message
 1175 type.

1176 **8.19 MCTP Bridging**

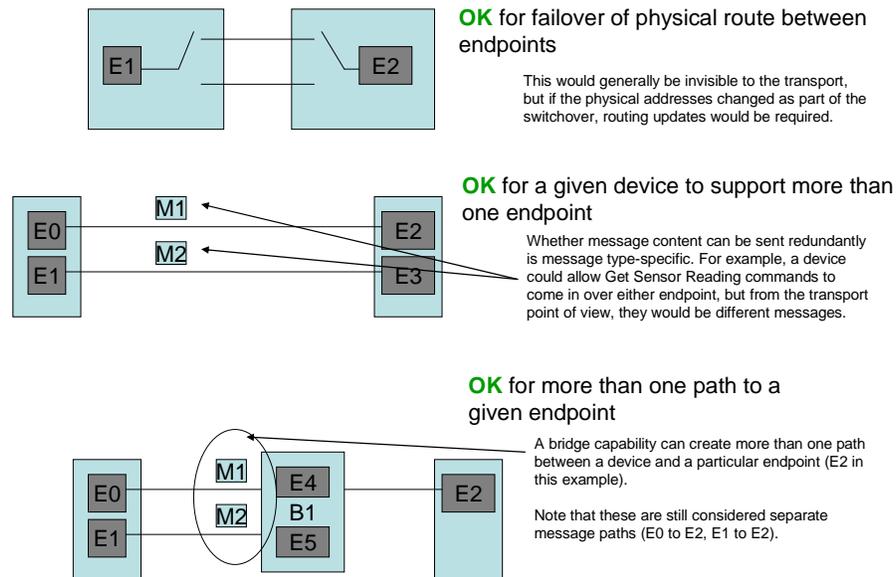
1177 One key capability provided by MCTP is its ability to route messages between multiple buses and
 1178 between buses of different types. This section describes how routing information is created, maintained,
 1179 and used by MCTP bridges and MCTP endpoints. Keep the following key points in mind about MCTP
 1180 bridges:

- 1181 • An MCTP bridge is responsible for routing MCTP packets between at least two buses.
- 1182 • An MCTP bridge is typically the bus owner for at least one of those buses.

1183 **8.19.1 Routing/Bridging Restrictions**

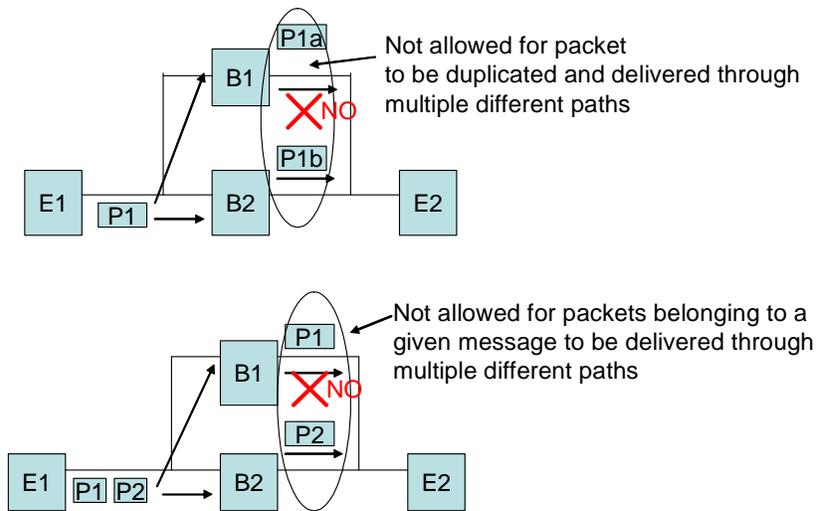
1184 Figure 7 and Figure 8 illustrate some of the supported and unsupported bridging topologies. As shown, it
 1185 is acceptable for a given topology to have more than one path to get to a given EID. This can occur either
 1186 because different media are used or because a redundant or failover communication path is desired in an
 1187 implementation.

1188 A bridge shall not route or forward packets with a broadcast destination ID.



1189

1190 **Figure 7 – Acceptable Failover/Redundant Communication Topologies**



1191

1192

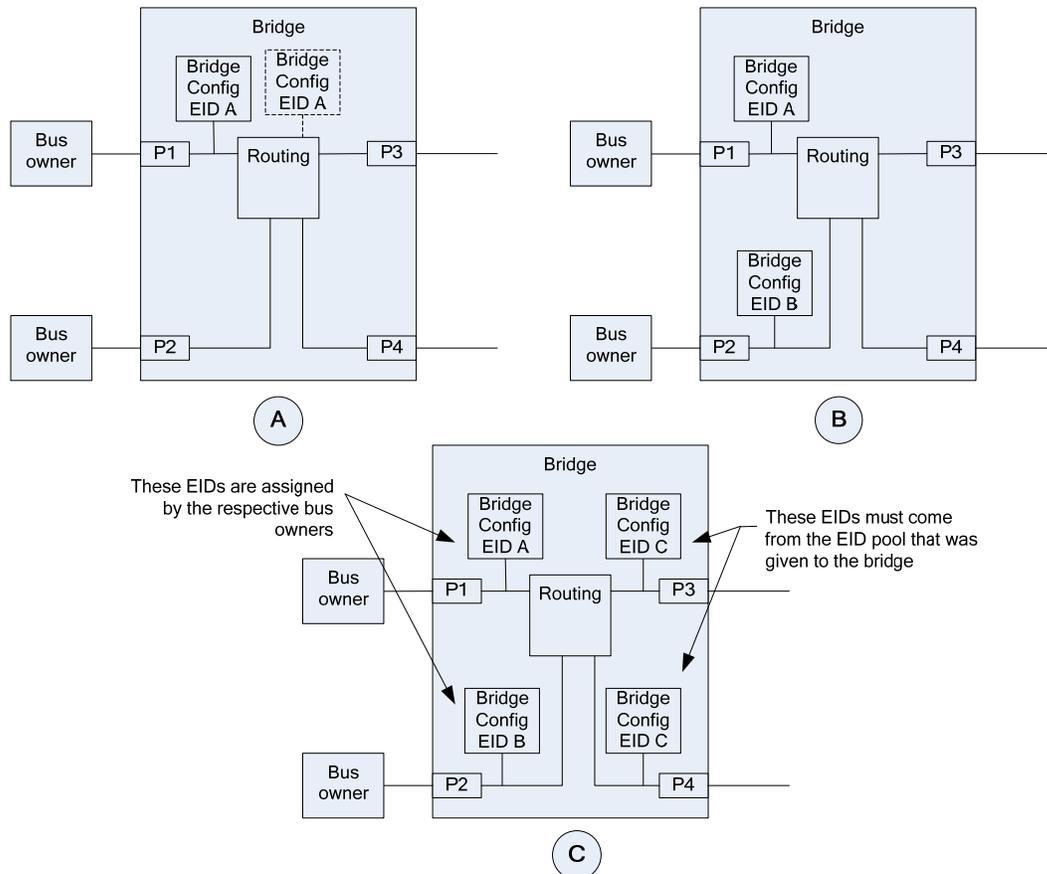
Figure 8 – Routing/Bridging Restrictions

1193 **8.19.2 EID Options for MCTP Bridges**

1194 An MCTP bridge that connects to multiple buses can have a single EID or multiple EIDs through which
 1195 the bridge's routing configuration and endpoint functionality can be accessed through MCTP control
 1196 commands. There are three general options:

- 1197 • The bridge uses a single MCTP endpoint
- 1198 • The bridge uses an MCTP endpoint for each bus that connects to a bus owner
- 1199 • The bridge uses an MCTP endpoint for every bus to which it connects

1200 Examples of these different options are shown in Figure 9, and more detailed information on the options
 1201 is provided following the figure.



1202

1203

Figure 9 – EID Options for MCTP Bridges

1204 A bridge has only one EID pool. To prevent issues with getting an EID pool allocation from multiple bus
 1205 owners, a bridge that is accessible through multiple EIDs will only accept EID pool allocation from the first
 1206 bus that allocation is received from using the Allocate Endpoint IDs command. This behavior is described
 1207 in more detail in the specification of the Allocate Endpoint IDs command.

1208 If necessary, the Get Endpoint UUID command can be used to correlate that EIDs belong to the same
 1209 MCTP bridge device. (This correlation is not required for normal initialization and operation of the MCTP
 1210 network, but it may be useful when debugging.)

1211 The following is a more detailed description of the different EID options for bridges:

1212 • **Single endpoint**

1213 A single endpoint is used to access the bridge's routing configuration and endpoint functionality.
 1214 Referring to diagram (A) in Figure 9, an implementation may elect to either have the endpoint
 1215 functionality be directly associated with a particular bus/port (for example, P1) or the
 1216 functionality can be located on a "virtual bus" that is behind the routing function. In either case,
 1217 the routing functionality ensures that the EID can be accessed through any of the buses to
 1218 which the bridge connects.

1219 Although there is a single endpoint, the bridge shall report the need for EID assignment for that
 1220 endpoint on each bus that is connected to a bus owner (for example, P1, P2). The multiple
 1221 announcements provide a level of failover capability in the EID assignment process in case a
 1222 particular bus owner becomes unavailable. The multiple announcements also help support a

1223 consistent EID assignment process across bus owners. To prevent issues with getting
 1224 conflicting EID assignments from multiple bus owners, the bridge will only accept EID pool
 1225 allocation from the first bus that an allocation is received from using the Set Endpoint ID
 1226 command. This behavior is described in more detail in the specification of the Set Endpoint ID
 1227 command. The bridge shall not report the need for EID assignment on any buses that the bridge
 1228 itself owns.

1229 • **Endpoint for each bus connection to a bus owner**

1230 The bridge has one endpoint for each bus connected to a bus owner. This is shown as diagram
 1231 (B) in Figure 9. There are no explicit endpoints associated with buses that are not connected to
 1232 a bus owner (for example, the buses connected to ports P3 and P4, respectively.) Because of
 1233 the way packet routing works, EID A and EID B can be accessed from any of the ports
 1234 connected to the bridge. Thus, the bridge's configuration functionality may be accessed through
 1235 multiple EIDs. Because a separate endpoint communication terminus is associated with each
 1236 port (P1, P2), the bridge can accept an EID assignment for each bus independently.

1237 The bridge shall only report the need for EID assignment on buses that connect to a bus owner,
 1238 and only for the particular MCTP control interface that is associated with the particular bus. For
 1239 example, the bridge would announce the need for EID assignment for the interface associated
 1240 with EID A only through P1, and the need for EID assignment for the interface associated with
 1241 EID B only through P2. The bridge shall not report the need for EID assignment on any buses
 1242 that the bridge itself owns.

1243 • **Endpoint for every bus connection**

1244 The bridge has one endpoint for each bus connected to it, as shown as diagram (C) in Figure 6.
 1245 This includes buses that connect to bus owners (for example, P1, P2) and buses for which the
 1246 bridge is the bus owner (for example, P3, P4). Because of the way packet routing works, any of
 1247 these EIDs can be accessed from any of the ports connected to the bridge.

1248 Because a separate endpoint communication terminus is associated with each owned port (P1,
 1249 P2), the bridge can accept an EID assignment for the bus owners of each bus independently.
 1250 The EIDs associated with the buses that the bridge itself owns (for example, P3, P4) must be
 1251 taken out of the EID pool that is allocated to the bridge.

1252 The bridge shall only report the need for EID assignment on buses that connect to a bus owner,
 1253 and only for the particular MCTP control interface that is associated with the particular bus. For
 1254 example, the bridge would announce the need for EID assignment for the interface associated
 1255 with EID A only through P1, and the need for EID assignment for the interface associated with
 1256 EID B only through P2. The bridge shall not report the need for EID assignment on any buses
 1257 that the bridge itself owns.

1258 **8.19.3 Routing Table**

1259 An MCTP bridge maintains a routing table where each entry in the table associates either a single EID or
 1260 a range of EIDs with a single physical address and bus ID for devices that are on buses that are directly
 1261 connected to the bridge.

1262 If the device is a bridge, there will typically be a range of EIDs that are associated with the physical
 1263 address of the bridge. There may also be an entry with a single EID for the bridge itself.

1264 **8.19.4 Bridging Process Overview**

1265 When a bridge receives an MCTP packet, the following process occurs:

- 1266 1) The bridge checks to see whether the destination EID in the packet matches or falls within the
 1267 range of EIDs in the table.
- 1268 2) If the EID is for the bridge itself, the bridge internally consumes the packet.

- 1269 3) If there is a match with an entry in the routing table, the following steps happen:
- 1270 • The bridge changes the physical addresses in the packet and reformats the medium-
- 1271 specific header and trailer fields as needed for the destination bus.
- 1272 • The destination physical address from the source bus is replaced with the destination
- 1273 physical address for the destination bus obtained from the entry in the routing table.
- 1274 • The bridge replaces the source physical address in the packet it received with the bridge's
- 1275 own physical address on the target bus. This is necessary to enable messages to be
- 1276 routed back to the originator.
- 1277 • Packet-specific transport header and data integrity fields are updated as required by the
- 1278 particular transport binding.
- 1279 4) If there is no match, packets with EID values that are not in the routing table are silently
- 1280 discarded.

1281 **8.19.5 Endpoint Operation with Bridging**

1282 A bridge does not track the packet transmissions between endpoints. It simply takes packets that it

1283 receives and routes them on a per-packet basis based on the destination EID in the packet. It does not

1284 pay attention to message assembly or disassembly or message type-specific semantics, such as

1285 request/response semantics, for packets that it routes to other endpoints.

1286 Most simple MCTP endpoints will never need to know about bridges. Typically, another endpoint will

1287 initiate communication with them. The endpoint can then simply take the physical address and source

1288 EID information from the message and use that to send messages back to the message originator.

1289 An endpoint that needs to originate a "connection" to another MCTP endpoint does need to know what

1290 physical address should be used for messages to be delivered to that endpoint. To get this information, it

1291 needs to query the bus owner for it. An endpoint knows the physical address of the bus owner because it

1292 saved that information when it got its EID assignment.

1293 The Resolve Endpoint ID command requests a bus owner to return the physical address that is to be

1294 used to route packets to a given EID. (This is essentially the MCTP equivalent of the ARP protocol that is

1295 used to translate IP addresses to physical addresses.) The address that is returned in the Resolve

1296 Endpoint ID command response will either be the actual physical address for the device implementing the

1297 endpoint, or it will be the physical address for the bridge to be used to route packets to the desired

1298 endpoint.

1299 Because the physical address format is media-specific, the format of the physical address parameter is

1300 documented in the specifications for the particular media-specific physical transport binding for MCTP (for

1301 example, MCTP over SMBus, MCTP over PCIe, and so on).

1302 If endpoint A has received a message from another endpoint B, it does not need to issue a Resolve

1303 Endpoint ID command. Instead, it can extract the source EID and source physical address from the

1304 earlier message from endpoint B, and then use that as the destination EID and destination physical

1305 address for the message to Endpoint B.

1306 **8.19.6 Routing Table Entries**

1307 Each MCTP device that does bridging must maintain a logical routing table. A bus owner must also

1308 typically maintain a routing table if more than one MCTP device is connected to the bus that it owns. The

1309 routing table is required because the bus owner is also the party responsible for resolving EIDs to

1310 physical addresses.

1311 The internal format that a device uses for organizing the routing table is implementation dependent. From
 1312 a logical point of view, each entry in a routing table will be comprised of at least three elements: An EID
 1313 range, a bus identifier, and a bus address. This is illustrated in Figure 10.

EID Range	Bus ID	Bus Address
-----------	--------	-------------

1314 **Figure 10 – Basic Routing Table Entry Fields**

1315 The *EID range* specifies the set of EIDs that can be reached through a particular bus address on a given
 1316 bus. Because the bus ID and bus address may correspond to a particular "port" on a bridge, it is possible
 1317 that there can be multiple non-contiguous ranges (multiple routing table entries) that have the same bus
 1318 ID/bus address pair route. EIDs and EID ranges can be categorized into three types: downstream,
 1319 upstream, and local. "Downstream" refers to EIDs that are associated with routing table entries that are
 1320 for buses that are owned by the bridge that is maintaining the routing table. "Upstream" refers to EIDs that
 1321 are associated with routing table entries that route to buses that are not owned by the bridge that is
 1322 maintaining the routing table.

1323 "Local" refers to the EIDs for routing table entries for endpoints that are on buses that are directly
 1324 connected to the bridge that is maintaining the routing table. A particular characteristic of entries for local
 1325 EIDs is that the Resolve Endpoint ID command is issued from the same bus that the endpoint is on. The
 1326 bridge/bus owner delivers the physical address for that endpoint rather than the physical address
 1327 associated with a routing function. This facilitates allowing endpoints on the same the bus to
 1328 communicate without having to go through an MCTP routing function.

1329 A routing table entry may not be "local" even if two endpoints are located on the same bus. An
 1330 implementation may require that different endpoints go through the routing function to intercommunicate
 1331 even if the endpoints are part of the same bus.

1332 The *bus ID* is an internal identifier that allows the MCTP device to identify the bus that correlates to this
 1333 route. MCTP does not require particular values to be used for identifying a given physical bus connection
 1334 on a device. However, this value will typically be a 0-based numeric value.

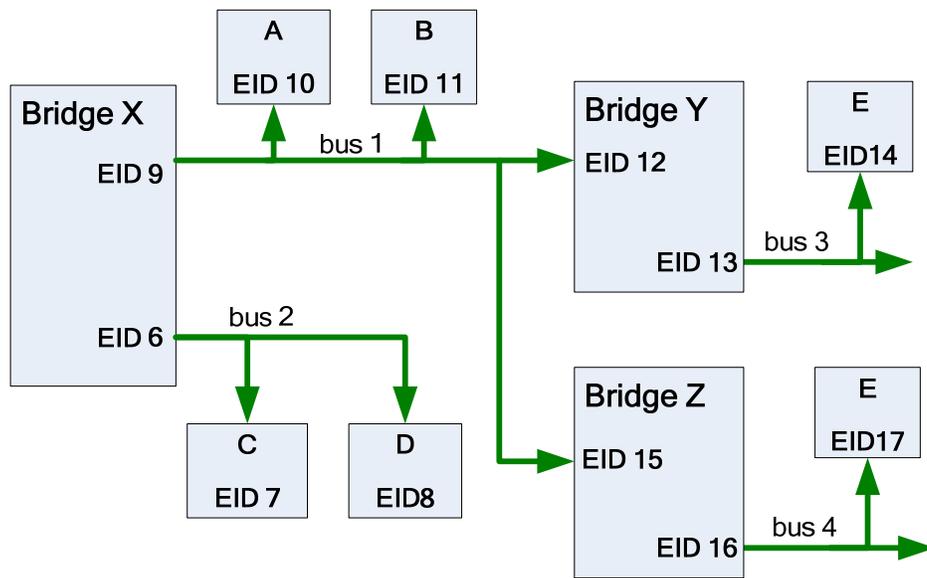
1335 EXAMPLE: A device that had three buses would typically identify them as buses "0", "1", and "2".

1336 The *bus address* is the physical address of a specific device on the bus through which the EIDs specified
 1337 in the *EID range* can be reached. This can either be the physical address corresponding to the
 1338 destination endpoint, or it can be the physical address of the next bridge in the path to the device. The
 1339 format of this address is specific to the particular physical medium and is defined by the physical medium
 1340 transport binding.

1341 **8.19.7 Routing Table Creation**

1342 This section illustrates the types of routing information that a bridge requires, and where the information
 1343 comes from. This section also describes the steps that a bus owner must use to convey that information
 1344 for a given bus.

1345 Figure 11 helps illustrate the steps that are required to completely establish the routing information
 1346 required by a bridge (bridge Y). The arrows in Figure 11 point outward from the bus owner and inward to
 1347 "owned" endpoints on the bus.



1348

1349

Figure 11 – Routing Table Population

1350 8.19.7.1 Routing Table Population Example

1351 With reference to Figure 11, the following items describe the information that bridge Y will need for routing
1352 messages in the example topology shown:

- 1353 • It needs a set of EIDs allocated to it to use for itself and to allocate to other devices (for
1354 example, EIDs 12:14). These are allocated to it by the bus owner (bridge X).
- 1355 • It needs a routing table that has an entry that maps EID 14 to the physical address for device E
1356 on bus 3.
- 1357 • It needs routing table entries for the local devices on bus 1, which are: bridge X (EID 9), device
1358 A (EID 10), device B (EID 11), and bridge Z (EID 15), assuming that devices A and B are to be
1359 reached by bridge Y without having to go through bridge X. This information must be given to it
1360 by the bus owner (bridge X).
- 1361 • It needs to know that EIDs 6:8 are accessed through bus owner/bridge X. Therefore, it needs a
1362 routing table entry that maps the EID range 6:8 to the physical address for bridge X on bus 1.
1363 This information must also be given to it by the bus owner (bridge X).
- 1364 • It needs to know that EIDs 16:17 are accessed through bridge Z. Therefore, it needs a routing
1365 table entry that maps the EID range 16:17 to the physical address for bridge Z on bus 1.
1366 Because the bus owner (bridge X) allocated that range of EIDs to bridge Z in the first place, this
1367 information is also given to bridge Y by the bus owner (bridge X).

1368 8.19.7.2 Bus Initialization Example

1369 Starting with the description of what bridge Y requires, the following task list shows the steps that bridge
1370 X must take to provide routing information for bus 1. Bridge X must:

- 1371 1) Assign EIDs to devices A, B, C, D, bridge Y, and bridge Z. This is done using the Set Endpoint
1372 ID command. The response of the Set Endpoint ID command also indicates whether a device
1373 wants an additional pool of EIDs.

- 1374 2) Allocate EID pools to bridge Y and bridge Z. This is done using the Allocate Endpoint IDs
1375 command.
- 1376 3) Tell bridge Y the physical addresses and EIDs for devices A and B, bridge X (itself), and bridge
1377 Z on bus 1. This is done using the Routing Information Update command.
- 1378 4) Tell bridge Y that EIDs 16:17 are accessed through the physical address for bridge Z on bus 1.
1379 This is also done using the Routing Information Update command. (Steps 3 and 4 can be
1380 combined and covered with one instance of the command.)
- 1381 5) Tell bridge Z the physical addresses and EIDs for devices A and B, bridge X (itself), and bridge
1382 Y on bus 1. This is also done using the Routing Information Update command.
- 1383 6) Tell bridge Z that EIDs 13:14 are accessed through the physical address for bridge Y on bus 1.
1384 This is also done using the Routing Information Update command. (Steps 5 and 6 can be
1385 combined and covered with one instance of the command.)
- 1386 7) Tell bridge Y and bridge Z that EIDs 6:8 are accessed through bridge X on bus 1. This is also
1387 done using the Routing Information Update command. This step could also be combined with
1388 steps 3 and 4 for bridge Y and steps 5 and 6 for bridge Z.

1389 **8.19.8 Routing Table Updates Responsibility for Bus Owners**

1390 After it is initialized for all bridges, routing table information does not typically require updating during
1391 operation. However, updating may be required if a bridge is added as a hot-plug device. In this case,
1392 when the bridge is added to the system, it will trigger the need for the bus owner to assign it an EID,
1393 which will subsequently cause the request for EID pool allocations, and so on. At this time, the bus owner
1394 can simply elect to re-run the steps for bus initialization as described in 8.19.7.2.

1395 **8.19.9 Consolidating Routing Table Entries**

1396 MCTP requires that when an EID pool is allocated to a device, the range of EIDs is contiguous and
1397 follows the EID for the bridge itself. Thus, a bridge can elect to consolidate routing table information into
1398 one entry when it recognizes that it has received an EID or EID range that is contiguous with an existing
1399 entry for the same physical address and bus. (The reason that EID allocation and routing information
1400 updates are not done as one range using the same command is because of the possibility that a device
1401 may have already received an allocation from a different bus owner.)

1402 **8.20 Bridge and Routing Table Examples**

1403 The following examples illustrate different bridge and MCTP network configurations and the
1404 corresponding information that must be retained by the bridge for MCTP packet routing and to support
1405 commands such as Resolve Endpoint ID and Query Hop.

1406 The following clauses (including Table 4 through Table 6) illustrate possible topologies and ways to
1407 organize the information that the bridge retains. Implementations may elect to organize and store the
1408 same information in different ways. The important aspect of the examples is to show what information is
1409 kept for each EID, to show what actions cause an entry to be created, and to show how an EID or EID
1410 range can in some cases map to more than one physical address.

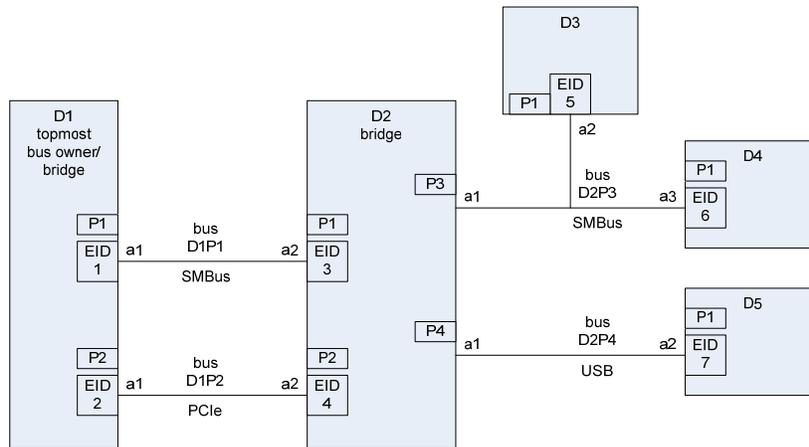
1411 The examples show a possible time order in which the entries of the table are created. Note that a given
1412 implementation of the same example topology could have the entries populated in a different order. For
1413 example, if there are two bus owners connected to a bridge, there is no fixed order that the bus owners
1414 would be required to initialize a downstream bridge. Additionally, there is no requirement that bus owners
1415 perform EID assignment or EID pool allocation in a particular order. One implementation may elect to
1416 allocate EID pools to individual bridges right after it has assigned the bridge its EID. Another
1417 implementation may elect to assign all the EIDs to devices first, and then allocate the EID pools to
1418 bridges.

1419 **8.20.1 Example 1: Bridge D2 with an EID per "Owned" Port**

1420 Figure 12 shows the routing table in a bridge (D2), where D2 has an EID associated with each bus
 1421 connected to a bus owner. In this example, D1 is not implementing any internal bridging between its P1
 1422 and P2. Consequently, EID2 cannot be reached by bridging through EID1 and vice versa (see Table 4).

1423 NOTE: If there was internal bridging, D1 would need to provide routing information that indicated that EID2 was
 1424 reachable by going through EID1 and vice versa. In this case, D1 would provide routing information that EID range
 1425 (EID1...EID2) would be accessed through D1P1a1 on SMBus and D1P1a2 on PCIe.

1426 **Key: D = device, P = port, a = physical address**



1427

1428 **Figure 12 – Example 1 Routing Topology**

1429

Table 4 – Example 1 Routing Table for D2

Time	EID	EID Access Port	Medium Type	Access Physical Address	Device/Entry Type	Entry Was Created and Populated By
↓	EID 3	P1	SMBus	D1P1a2	Bridge, Self	Self when EID was assigned by D1
	EID 4	P2	PCIe	D1P2a2	Bridge, Self	Self when EID was assigned by D1
	EID 5	P3	SMBus	D2P3a2	Endpoint	Self after D1 assigned EID pool (typically the entry will not be created until after the bridge D2 assigns EID 5 to D3)
	EID 6	P3	SMBus	D2P3a3	Endpoint	Self after D1 assigned EID pool (typically the entry will not be created until after the bridge D2 assigns EID 6 to D4)
	EID 7	P4	USB	D2P4a2	Endpoint	Self after D1 assigned EID pool (typically the entry will not be created until after the bridge D2 assigns EID 7 to D5)
	EID 1	P1	SMBus	D1P1a1	Bridge	D1 through Routing Information Update command
	EID 2	P2	PCIe	D1P2a1	Bridge	D1 through Routing Information Update command

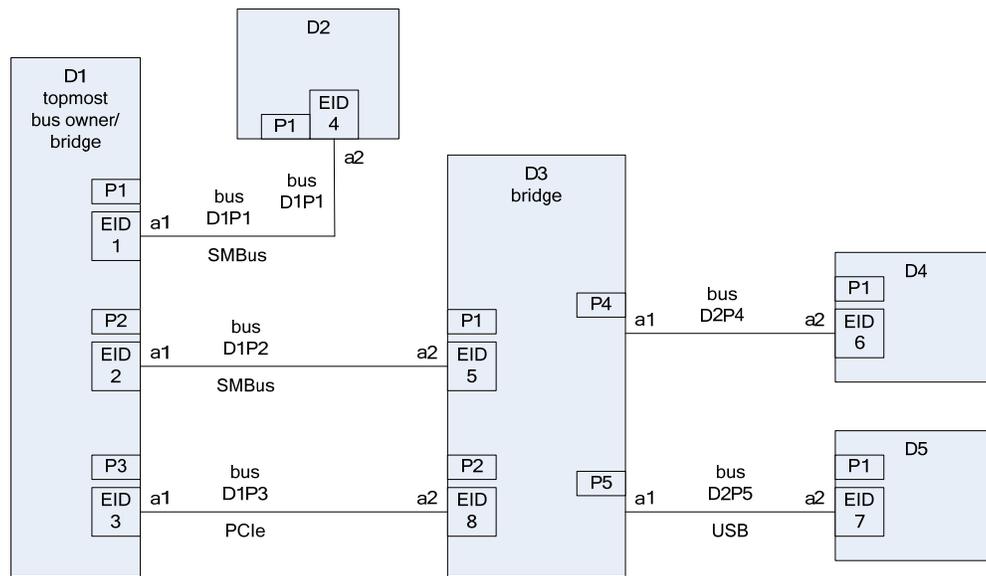
1430 **8.20.2 Example 2: Topmost Bus Owner D1**

1431 Figure 13 assumes the following:

- 1432 • D1 assigns its internal EIDs first.
- 1433 • The buses are handled in the order D1P1, D1P2, D1P3.
- 1434 • D1 allocates the EID pool to bridges right after it has assigned the EID to the device.

1435 Similar to Example 1, this example assumes that there is no internal bridging within D1 between P1, P2,
 1436 and P3. This scenario is reflected in Table 5.

1437 **Key: D = device, P = port, a = physical address**



1438

1439 **Figure 13 – Example 2 Routing Topology**

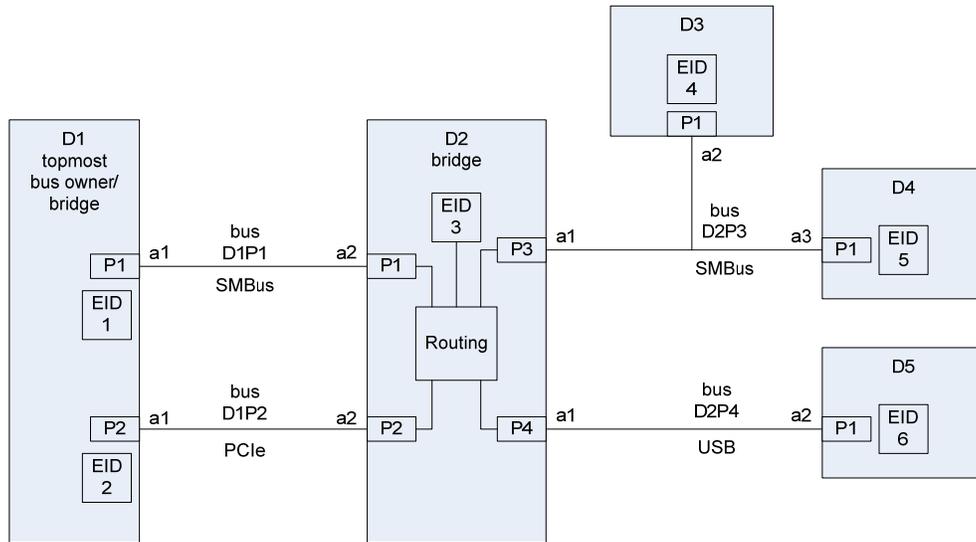
1440 **Table 5 – Example 2 Routing Table for D1**

EID	EID Access Port	Medium Type	Access Physical Address	Device/Entry Type	Entry Was Created and Populated By
EID 1	P1	SMBus	D1P1a1	Bridge, self	Self
EID 2	P2	SMBus	D1P2a1	Bridge, self	Self
EID 3	P3	PCIe	D1P3a1	Bridge, self	Self
EID 4	P1	SMBus	D1P1a2	Endpoint	Self upon assigning EID to device D2
EID 5	P2	SMBus	D1P2a2	Bridge	Self upon assigning EID 5 to bridge D3
EID 6:7	P2	SMBus	D1P2a2	Bridge pool	Self upon assigning EID pool to bridge D3
EID 8	P3	PCIe	D1P3a2	Bridge	Self upon assigning EID 8 to bridge D3
EID 6:7	P3	PCIe	D1P3a2	Bridge pool	Self upon issuing an Allocate Endpoint IDs command and finding that bridge D3 already has an assigned pool, D1 creates this entry by extracting the EIDs for this entry from the response to the Allocate Endpoint IDs command

1441 **8.20.3 Example 3: Bridge D2 with Single EID**

1442 Figure 14 assumes that bridge D2 has a single EID and gets its EID assignment and EID allocation
 1443 through bus D1P1 first, and that bus D1P2 later gets initialized. This scenario is reflected in Table 6.

1444 **Key: D = device, P = port, a = physical address**



1445

1446

Figure 14 – Example 3 Routing Topology

1447

Table 6 – Example 3 Routing Table for D2

Target EID	Target Endpoint Access Port	Target EID Access Physical Address	Device/Entry Type	Entry Was Created and Populated By
EID 3	P1	D1P1a2	Bridge, self	All four entries created by self (bridge) upon receiving initial EID assignment from D1 through P1
EID 3	P2	D1P2a2	Bridge, self	
EID 3	P3	D2P3a1	Bridge, self	
EID 3	P4	D2P4a1	Bridge, self	
EID 4	P3	D2P3a2	Endpoint	Self after D1 allocated EID pool (typically the entry will not be created until after the bridge D2 assigns EID 4 to D3)
EID 5	P3	D2P3a3	Endpoint	Self after D1 allocated EID pool (typically the entry will not be created until after the bridge D2 assigns EID 5 to D4)
EID 6	P3	D2P4a2	Endpoint	Self after D1 allocated EID pool (typically the entry will not be created until after the bridge D2 assigns EID 6 to D5)
EID 1:2	P1	D1P1a1	Bridge	D1 through Routing Information Update command
EID 1:2	P2	D1P2a1	Bridge	D1 through Routing Information Update command

1448 **8.20.4 Additional Information Tracked by Bridges**

1449 In addition to the information required to route messages between different ports, a bridge has to track
 1450 information to handle MCTP control commands related to the configuration and operation of bridging
 1451 (shown in Table 7).

1452 **Table 7 – Additional Information Tracked by Bridges**

What	Why
Which buses are connected to a bus owner	This information tells the bridge from which buses it should request EID assignment. This will typically be accomplished as a non-volatile configuration or hardware-strapping option for the bridge.
Which bus the bridge received its EID assignment through the Set Endpoint ID command	If the bridge uses a single EID that is shared across multiple "owned" buses, this information is used to track which bus the request came in on, so that the bridge can reject EID assignment requests from other buses.
Which bus it received the Routing Information Update command from for creating a particular routing table entry	This information is required so that if a future Routing Information Update command is received, the bridge will update only the entries corresponding to that bus.
Which bus it received its EID pool allocation from through the Allocate Endpoint IDs command	This information is used to track which bus the request came in on so that the bridge can reject EID pool allocations from other buses.
The physical medium and physical addressing format used for each port	This information is used to provide the correctly formatted response to commands such as Resolve Endpoint ID and for bridging MCTP packets between the different buses that the bridge supports. Because this is related to the physical ports and hardware of the bridge, this information will typically be "hard coded" into the bridge.

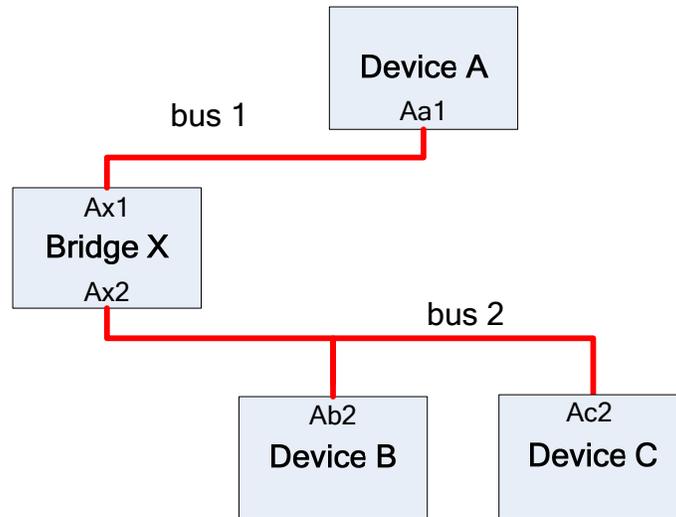
1453 **8.21 Endpoint ID Resolution**

1454 When a device uses the Resolve Endpoint ID command to request the resolution of a given endpoint to a
 1455 physical address, the bridge must respond based on which bus the request came in on.

1456 For example, consider Figure 15. If device A wishes to get the physical address needed to send a
 1457 message to device C, it sends a Resolve Endpoint ID command to bus owner bridge X through address
 1458 Ax1. Because device A must go through bridge X to get to device C, bridge X responds with its physical
 1459 address Ax1.

1460 When device B wishes to know the address to use to communicate with device C, it sends a Resolve
 1461 Endpoint ID request to bridge X through address Ax2. In this case, bridge X can respond by giving device
 1462 B the direct physical address of device C on bus 2, Ac2.

1463 Thus, the Resolve Endpoint ID command can return a different response based on the bus from which
 1464 the Resolve Endpoint ID command was received.



notation:

Ab2 = physical Address of device b on bus 2.

1465

1466

Figure 15 – Endpoint ID Resolution

1467

8.21.1 Resolving Multiple Paths

1468

Cases can occur where there can be more than one possible path to a given EID. A likely scenario is shown in Figure 16. In Figure 16, assume that the system topology supports cards that connect to either SMBus, PCIe, or both. Bridge X is the bus owner for both buses.

1469

1470

1471

NOTE: This is a logical representation of MCTP buses. Physically, the buses may be formed of multiple physical segments, as would be the case if one of the MCTP buses was built using PCIe.

1472

1473

As shown, card C contains a bridge that connects to both buses. Thus, the device with EID 100 can be reached either from bus 1 or bus 2.

1474

1475

If device D wishes to send a message to EID 100, bridge X can choose to route that message either through bus 1 or bus 2. MCTP does not have a requirement on how this is accomplished. The general recommendation is that the bridge preferentially selects the faster available medium. In this example, that would be PCIe.

1476

1477

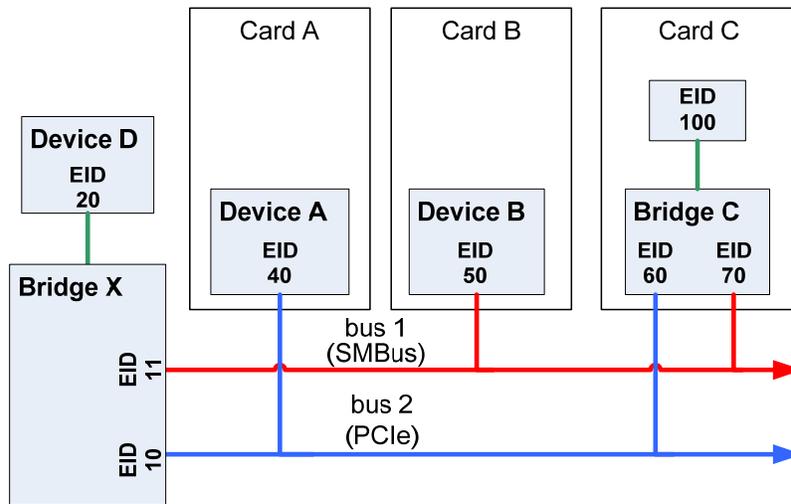
1478

1479

NOTE: There are possible topologies where that simple rule may not yield the preferred path to a device. However, in most common implementations in PC systems, this approach should be effective. A vendor making a bridge device may consider providing configuration options to enable alternative policies.

1480

1481



1482

1483

Figure 16 – Resolving Multiple Paths

1484

8.22 Bridge and Bus Owner Implementation Recommendations

1485

This clause provides recommendations on EID pool and routing table sizes for devices that implement bridge and bus owner functionality.

1486

1487

8.22.1 Endpoint ID Pool Recommendations

1488

The system design should seek to minimize the number of devices that need to allocate EID pools to hot-plug devices or add-in cards. If feasible, the system design should have all busses that support hot-plug devices/add-in cards owned by a single device.

1489

1490

1491

If only one device handles the hot-plug devices and add-in cards, it will be simpler for the system integrator to configure devices and allocate EID pools. Because any other bridges in the system that do not handle hot-plug devices only need to handle a fixed number of MCTP devices, it will be known at design time how large an EID pool will be required. The remaining number of EIDs can then simply be allocated to the single device that handles the hot-plug devices and add-in cards.

1492

1493

1494

1495

1496

To support this, it is recommended that devices that operate as bridges include a non-volatile configuration option that enables the system integrator to configure the size of the EID pool they request.

1497

1498

8.22.2 Routing Table Size Recommendations

1499

This section provides some initial recommendations and approaches on how to determine what target routing table entry support to provide in a device.

1500

1501

- **PCIe slots**

1502

To provide entries to support devices that plug into PCIe slots, assume that each slot may support both PCIe and SMBus endpoints and provide support for at least two endpoints per bus type.

1503

1504

1505

This means providing support for at least four directly connected endpoints per card. (Other endpoints may be behind bridges on the card, but this does not affect the routing table size for the bus owner.) This implies at least four routing table entries per PCIe slot. Thus, a device that was designed to support system implementations with eight PCIe slots should have support for 32 routing table entries.

1506

1507

1508

1509

- 1510 • **Planar PCIe devices**
- 1511 In most PC systems, PCIe would be typically implemented as a single MCTP bus owned by a
1512 single device as the bus owner. Thus, the number of static devices should be proportional to the
1513 number of PCIe devices that are built into the motherboard.
- 1514 Typically, this is fewer than eight devices. Thus it is recommended to support at least eight
1515 entries for static PCIe devices.
- 1516 • **Static SMBus MCTP devices**
- 1517 The routing table should also be sized to support an additional number of "static" devices on
1518 owned buses. At this time, it is considered unlikely that more than a few MCTP devices would
1519 be used on a given SMBus. Most devices would be non-intelligent sensor and I/O devices
1520 instead. Conservatively, it is recommended that at least four entries be provided for each
1521 SMBus that the device owns.
- 1522 Example 1: "client" capable device
- | | | |
|------|-----------------------------------|----------------------------|
| 1523 | Four PCIe slots | → 16 routing table entries |
| 1524 | Two owned SMBus | → +8 entries |
| 1525 | <u>Static PCIe device support</u> | <u>→ +8 entries</u> |
| 1526 | | ~32 entries or more |
- 1527 Example 2: volume server capable
- | | | |
|------|-----------------------------------|----------------------------|
| 1528 | Eight PCIe slots | → 32 routing table entries |
| 1529 | Four owned SMBus | → +16 entries |
| 1530 | <u>Static PCIe device support</u> | <u>→ +8 entries</u> |
| 1531 | | ~56 entries or more |

1532 8.23 Path and Transmission Unit Discovery

1533 The transmission unit is defined as the size of the MCTP packet payload that is supported for use in
1534 MCTP message assembly for a given message. The supported transmission unit sizes are allowed to
1535 vary on a per-message type basis.

1536 Intermediate bridges and physical media can limit the transmission unit sizes between endpoints.
1537 Therefore, the MCTP control protocol specifies a mechanism for discovering the transmission unit support
1538 for the path between endpoints when one or more bridges exist in the path between the endpoints.

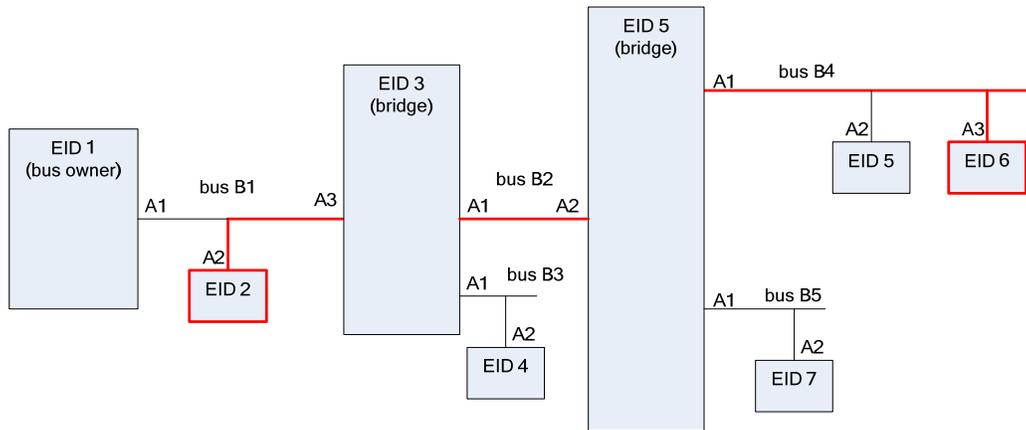
1539 The mechanism for path transmission unit discovery also enables the discovery of the bridges and
1540 number of "hops" that are used to route an MCTP packet from one endpoint to another.

1541 8.23.1 Path Transmission Unit Negotiation

1542 The MCTP control protocol only specifies how to discover what the path transmission unit size is for the
1543 path between endpoints. The MCTP control protocol does not specify a generic mechanism for
1544 discovering what transmission unit sizes a particular endpoint supports for a given message type.
1545 Discovery and negotiation of transmission unit sizes for endpoints, if supported, is specified by the
1546 definition of the particular message type.

1547 **8.23.2 Path Transmission Unit Discovery Process Overview**

1548 This section describes the process used for path transmission unit discovery. The discovery process
 1549 described here is designed to enable one endpoint to discover the path and transmission unit support for
 1550 accessing a particular "target" endpoint. It does not define a general mechanism for enabling an endpoint
 1551 to discover the path between any two arbitrary endpoints. For example, referring to
 1552 Figure 17, the process defines a way for the endpoint at EID 2 to discover the path/transmission unit
 1553 support on the route to endpoint at EID 6, but this process does not define a process for EID 2 to discover
 1554 the path/transmission unit support between EID 4 and EID 6.



1555

1556 **Figure 17 – Example Path Routing Topology**

1557 The following example provides an overview of the path/transmission unit discovery process. The
 1558 example presumes that the MCTP network has already been initialized. Referring to
 1559 Figure 17, the endpoint with EID 2 wishes to discover the path used to access the endpoint with EID 6.
 1560 This discovery is accomplished using just two commands, Resolve Endpoint ID and Query Hop, as
 1561 follows:

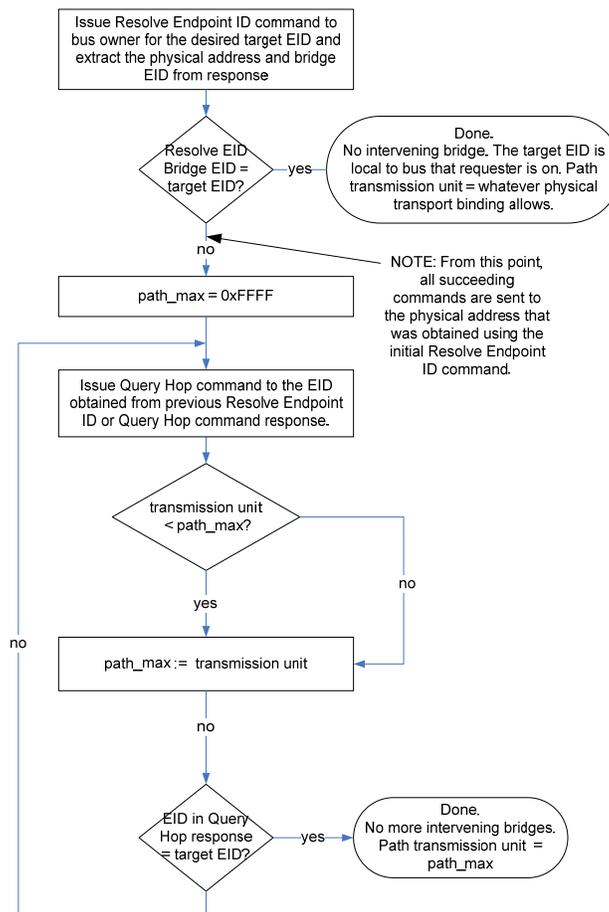
- 1562 1) EID 2 first issues a Resolve Endpoint ID command to the bus owner, EID 1, with EID 6 as the
 1563 EID to resolve.
- 1564 2) EID 1 returns the physical address and EID of the bridge, EID 3 in the Resolve Endpoint ID
 1565 command response.
- 1566 3) EID 2 queries the bridge, EID 3, using a Query Hop command with EID 6 (the "target" EID) as
 1567 the request parameter. Note that EID 2 does not need to do another Resolve Endpoint ID
 1568 command because it already received the physical address of EID 3 from the original Resolve
 1569 Endpoint ID command.
- 1570 4) Bridge EID 3 responds to the Query Hop command by returning EID 5, which is the EID of the
 1571 next bridge required to access EID 6. The bridge EID 3 also returns the transmission unit
 1572 support that it offers for routing to the target EID.
- 1573 5) EID 2 then sends a Query Hop command to the bridge at EID 5. Note that EID 2 does not need
 1574 to do another Resolve Endpoint ID command because it already received the physical address
 1575 of EID 5 from the original Resolve Endpoint ID command.
- 1576 6) Bridge EID 5 responds to the Query Hop command by returning EID 6, which, because it is the
 1577 EID of the target endpoint, tells EID2 that bridge EID 5 was the last "hop" in the path to EID 6.
 1578 The bridge EID 5 also returns the transmission unit support that it offers for routing to the target
 1579 EID.

1580 7) At this point, the bridges in the path to EID 6 have subsequently been discovered and their
 1581 respective transmission unit support returned. The effective transmission unit support for the
 1582 path to EID 6 will be the lesser of the transmission unit support values returned by the two
 1583 bridges.

1584 8.23.3 Path Transmission Unit Discovery Process Flowchart

1585 Figure 18 shows a generic algorithm for discovering the bridges in the path from one endpoint to a given
 1586 target endpoint and the path transmission unit support. The flowchart has been intentionally simplified.
 1587 Note that while the Query Hop command actually supports returning separate transmission unit sizes for
 1588 the transmit and receive paths, the flowchart is simplified for illustration purposes and just refers to a
 1589 single transmission unit for both transmit and receive.

1590 Additionally, Figure 18 does not show any explicit steps for error handling nor the process of handling
 1591 command retries. In general, errors are most likely due to either an invalid EID being sent to the bridge
 1592 (perhaps due to a programming error at the requester) or the EID not being present in the bridge's routing
 1593 table. The latter condition could occur under normal operation if the requester did not realize that a
 1594 routing table update had occurred because of a hot-plug update, for example. This error condition would
 1595 be indicated by the bridge responding with an `ERROR_INVALID_DATA` completion code.



1596

1597

Figure 18 – Path Transmission Unit Discovery Flowchart

1598 **8.24 Path Transmission Unit Requirements for Bridges**

1599 An MCTP bridge routes packets between different buses, but it does not typically interpret the packet
 1600 payload contents nor does it do assembly of those packets. Exceptions to this are when the bridge is
 1601 handling packets addressed to its own EID, receives a Broadcast EID, and if the bridge supports different
 1602 transmission units based on message type. See Table 31 for more information.

1603 **9 MCTP Control Protocol**

1604 MCTP control messages are used for the setup and initialization of MCTP communications within an
 1605 MCTP network. This section defines the protocol and formatting used for MCTP control messages over
 1606 MCTP.

1607 **9.1 Terminology**

1608 The terms shown in Table 8 are used when describing the MCTP control protocol.

1609 **Table 8 – MCTP Control Protocol Terminology**

Term	Description
Requester	The term "requester" is used to refer to the endpoint that originates an MCTP control Request message.
Responder	The term "responder" is used to refer to the endpoint that originates an MCTP control response message (that is, an endpoint that returns the response to an MCTP control Request message).
Originator or Source	The term "originator" or "source" is used to refer to the endpoint that originates any MCTP control message: Request, Response, or Datagram.
Target or Destination	The term "target" or "destination" is used to refer to the endpoint that is the intended recipient of any MCTP control message: Request, Response, or Datagram.
Asynchronous Notification	The term "asynchronous notification" is used to refer to the condition when an MCTP endpoint issues an un-requested Datagram to another MCTP endpoint.
Broadcast	The term "broadcast" is used when an MCTP control Datagram is sent out onto the bus using the broadcast EID.

1610 **9.1.1 Control Message Classes**

1611 The different types of messages shown in Table 9 are used under the MCTP control message type.

1612 **Table 9 – MCTP Control Message Types**

Type	Description
Request	This class of control message requests that an endpoint perform a specific MCTP control operation. All MCTP control Request messages are acknowledged with a corresponding Response message. (Within this specification, the term "command" and "request" are used interchangeably as shorthand to refer to MCTP control Request messages.)
Response	This class of MCTP control message is sent in response to an MCTP control Request message. The message includes a "Completion Code" field that indicates whether the response completed normally. The response can also return additional data dependent on the particular MCTP control Request that was issued.

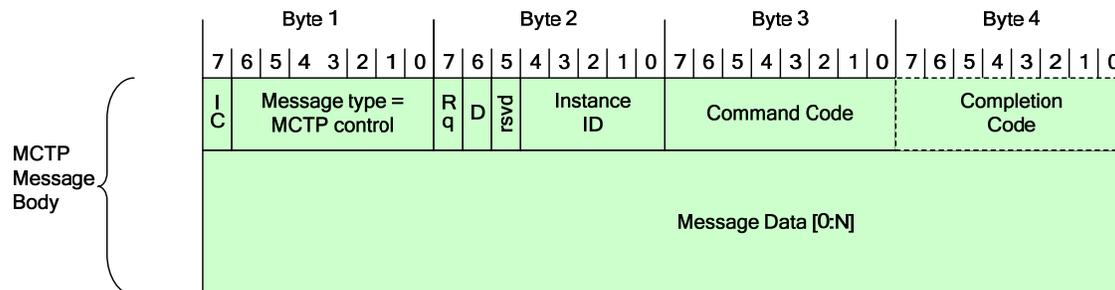
Type	Description
Datagram	Datagrams are "unacknowledged" messages (that is, Datagrams do not have corresponding Response messages). This class of MCTP control message is used to transfer messages when an MCTP control Response message is neither required nor desirable.
Broadcast Request	A broadcast message is a special type of Request that is targeted to all endpoints on a given bus. All endpoints that receive the message are expected to interpret the Request.
Broadcast Datagram	A Datagram that is broadcast to all endpoints on the bus. Broadcast Datagrams are "unacknowledged" messages (that is, broadcast Datagrams do not have corresponding Response messages).

1613 **9.2 MCTP Control Message Format**

1614 MCTP control messages use the MCTP control message type (see Table 3). Any message sent with this
 1615 message type will correspond to the definitions set forth in this section. The basic format of an MCTP
 1616 control message is shown in Figure 19. Note that the byte offsets shown in Figure 19 are relative to the
 1617 start of the MCTP message body rather than the start of the physical packet.

1618 **9.2.1 Use of Message Integrity Check**

1619 MCTP control messages do not use a Message Integrity Check field. Therefore, the IC bit in MCTP
 1620 control messages shall always be 0b.



1621

1622

Figure 19 – MCTP Control Message Format

1623 **9.3 MCTP Control Message Fields**

1624 Table 10 lists the common fields for MCTP control messages.

1625

Table 10 – MCTP Control Message Fields

Field Name	Description
IC*	Message Integrity Check bit = 0b. MCTP control messages do not include an overall Message Integrity check field.
Message Type*	MCTP control = 0x00 (000_0000b). This field identifies the MCTP message as being an MCTP control message.
Rq bit	Request bit. This bit is used to help differentiate between MCTP control Request messages and other message classes. Refer to 9.5.

Field Name	Description
D-bit	Datagram bit. This bit is used to indicate whether the Instance ID field is being used for tracking and matching requests and responses, or is just being used to identify a retransmitted message. Refer to 9.5.
Instance ID	The Instance ID field is used to identify new instances of an MCTP control Request or Datagram to differentiate new requests or datagrams that are sent to a given message terminus from retried messages that are sent to the same message terminus. The Instance ID field is also used to match up a particular instance of an MCTP Response message with the corresponding instance of an MCTP Request message.
Command Code	For Request messages, this field is a command code indicating the type of MCTP operation the packet is requesting. Command code values are defined in Table 12. The format and definition of request and response parameters for the commands is given in Clause 10. The Command Code that is sent in a Request must be returned in the corresponding Response.
Completion Code	This field is only present in Response messages. This field contains a value that indicates whether the response completed normally. If the command did not complete normally, the value can provide additional information regarding the error condition. The values for completion codes are specified in Table 13.
Message Data	Zero or more bytes of parameter data that is specific to the particular Command Code and whether the message is a Request or Datagram, or a Response.
* These fields are MCTP base protocol fields.	

1626 **9.4 MCTP Control Message Transmission Unit Size**

1627 All MCTP control messages are required to have a packet payload that is no larger than the baseline
 1628 transmission unit size of 64 bytes.

1629 MCTP control messages are carried in a single MCTP packet. Multiple messages are used if an operation
 1630 requires more data to be transferred than can be carried in a single message.

1631 **9.5 Tag Owner (TO), Request (Rq), and Datagram (D) Bit Usage**

1632 For MCTP control messages, the Rq bit is set to 1b if the message is a "command" or Request message
 1633 and 0b if the message is a Response message. For Datagram and Broadcast messages, the Rq bit is
 1634 always 1b.

1635 For the present specification, Requests and Datagrams are required to be issued from tag owners (TO
 1636 bit = 1b). Provision has been left for the definition of possible future Datagrams that are not issued from
 1637 tag owners (see Table 11).

1638

Table 11 – Tag Owner (TO), Request (Rq) and Datagram (D) Bit Usage

MCTP Control Message Class	Destination EID Value	Tag Owner (TO) bit	Request (Rq) bit	Datagram (D) bit
Command/Request Responses are expected and tracked by Instance ID at the requester.	Target EID	1b	1b	0b
Response	Target EID	0b	0b	0b
Broadcast Request Responses are expected and tracked by Instance ID at the requester.	Broadcast EID	1b	1b	0b
Datagram Unacknowledged Request – Responses are neither expected nor tracked by Instance ID at the requester. Duplicate packets are handled the same as retried Command/Request packets.	Target EID	1b	1b	1b
Broadcast Datagram (unacknowledged control command that is broadcast.)	Broadcast EID	1b	1b	1b
Reserved for future definition	all other			

1639 9.6 Concurrent Command Processing

1640 This section describes the specifications and requirements for handling concurrent overlapping MCTP
1641 control requests by endpoints.

1642 9.6.1 Requirements for Responders

1643 An endpoint is not required to process more than one request at a time (that is, it can be "single threaded"
1644 and does not have to accept and act on new requests until it has finished responding to any previous
1645 request).

1646 A responder that is not ready to accept a new request can either silently discard the request, or it can
1647 respond with an `ERROR_NOT_READY` message completion code.

1648 9.6.2 Requirements for Requesters

1649 An endpoint that issues MCTP control Requests to another endpoint must wait until it gets the response
1650 to the particular request, or times out waiting for the response, before issuing a new request, Datagram,
1651 or Broadcast Datagram.

1652 An endpoint that issues MCTP control Requests is allowed to have multiple requests outstanding
1653 simultaneously to *different* responder endpoints.

1654 An endpoint that issues MCTP control Requests should be prepared to handle responses that may not
1655 match the request (that is, it should not automatically assume that a response that it receives is for a
1656 particular request). It should check to see that the command code and source EID values in the response
1657 match up with a corresponding outstanding command before acting on any parameters returned in the
1658 response.

1659 **9.6.3 Additional Requirements for Bridges**

1660 The packets that are routed *through* a bridge's routing functionality are not interpreted by the bridge and
 1661 therefore are not considered to constitute concurrent requests.

1662 A bridge must support at least one outstanding MCTP control request for each bus connection (port)
 1663 through which MCTP control messages can be used to access the bridge's configuration and control
 1664 functionality.

1665 Bridges must retain temporal ordering of packets forwarded from one message terminus to another.

1666 **10 MCTP Control Messages**

1667 This section contains detailed descriptions for each MCTP control message. The byte offsets for the
 1668 Request and Response parameter information given in the tables for the commands indicates the byte
 1669 offset for the message data starting with the byte following the Command field.

1670 **10.1 MCTP Control Message Command Codes**

1671 Table 12 lists the MCTP control messages and their corresponding command code values. The
 1672 commands and their associated parameters are specified later in this section. For bridges, the
 1673 requirements apply equally to all endpoints within the bridge device that are used to configure and control
 1674 the bridges routing functionality.

1675 **Table 12 – MCTP Control Command Numbers**

Command Code	Command Name	General Description	OMC		Section
			E	B	
0x00	Reserved	Reserved	–	–	–
0x01	Set Endpoint ID	Assigns an EID to the endpoint at the given physical address	Ma Ng	Ca ¹ Mg	10.3
0x02	Get Endpoint ID	Returns the EID presently assigned to an endpoint. Also returns information about what type the endpoint is and its level of use of static EIDs.	Ma Og	Ma Og	10.4
0x03	Get Endpoint UUID	Retrieves a per-device unique UUID associated with the endpoint	Ca ² Og	Ca ² Og	10.5
0x04	Get MCTP Version Support	Lists which versions of the MCTP control protocol are supported on an endpoint	Ma Og	Ma Og ⁵	10.6
0x05	Get Message Type Support	Lists the message types that an endpoint supports	Ma Og	Ma Og	10.7
0x06	Get Vendor Defined Message Support	Used to discover an MCTP endpoint's vendor-specific MCTP extensions and capabilities	Oa Og	Oa Og	10.8
0x07	Resolve Endpoint ID	Used to get the physical address associated with a given EID	Na Og	Ma Og	10.9
0x08	Allocate Endpoint IDs	Used by the bus owner to allocate a pool of EIDs to an MCTP bridge	Na Ng	Ma ⁶ Mg ⁶	10.10
0x09	Routing Information Update	Used by the bus owner to extend or update the routing information that is maintained by an MCTP bridge	Na Ng	Ma ⁴ Mg ⁴	10.11
0x0A	Get Routing Table Entries	Used to request an MCTP bridge to return data corresponding to its present routing table entries	Na Og	Ma Og	10.12
0x0B	Prepare for Endpoint Discovery	Used to direct endpoints to clear their "discovered" flags to enable them to respond to the Endpoint Discovery command	Ca ³ Ng	Ca ³ Cg ³	10.13

Command Code	Command Name	General Description	OMC		Section
			E	B	
0x0C	Endpoint Discovery	Used to discover MCTP-capable devices on a bus, provided that another discovery mechanism is not defined for the particular physical medium	Ca ³ Cg ³	Ca ³ Cg ³	10.14
0x0D	Discovery Notify	Used to notify the bus owner that an MCTP device has become available on the bus	Na Cg ³	Ca ³ Cg ³	10.15
0x0E	Get Network ID	Used to get the MCTP network ID	Ca ⁷	Ca ⁷	10.15
0x0F	Query Hop	Used to discover what bridges, if any, are in the path to a given target endpoint and what transmission unit sizes the bridges will pass for a given message type when routing to the target endpoint	Na Og	Ma Og	10.16
0xF0 – 0xFF	Transport Specific	This range of control command numbers is reserved for definition by individual MCTP Transport binding specifications. Transport-specific commands are intended to be used as needed for setup and configuration of MCTP on a given media. A particular transport-specific command number may have different definitions depending on the binding specification. Transport-specific commands shall only be addressed to endpoints on the same medium. A bridge is allowed to block transport-specific commands from being bridged to different media. The general format of transport-specific messages is specified in 10.18.	-	-	10.17
<p>Key for OMC (optional / mandatory / conditional) column:</p> <p>E = non-bridge, non-bus owner endpoint (simple endpoint)</p> <p>B = bridge / bus-owner endpoint</p> <p>Ma = mandatory (required) to accept. The request must be accepted by the endpoint and a response generated per the following command descriptions.</p> <p>Mg = mandatory to generate. The endpoint must generate this request as part of its responsibilities for MCTP operation.</p> <p>Oa = optional to accept</p> <p>Og = optional to generate</p> <p>Ca = conditional to accept (see notes)</p> <p>Cg = conditional to generate (see notes)</p> <p>Na = not applicable to accept. This command is not applicable to the device type and must not be accepted</p> <p>Ng = not applicable to generate. This command is used for MCTP configuration and initialization and should not be generated.</p>					
<ol style="list-style-type: none"> 1. The topmost bus owner is not required to support the Set Endpoint ID command. 2. Hot-plug and add-in devices are required to support the Get Endpoint UUID command. 3. Mandatory on a per-bus basis to support endpoint discovery if required by the physical transport binding used for the particular bus type. Refer to the appropriate MCTP physical transport binding specification. 4. The topmost bus owner is not required to accept this command. The command is required to be generated when downstream bridges require dynamic routing information from bus owners that they are connected to. Some implementations may be configured where all routing information has been statically configured into the bridge and no dynamically provided information is required, In this case, it is not required to support the command while the endpoints are configured in that manner. 5. Bridges should use this command to verify that they are initializing devices that are compatible with their MCTP control protocol version. 6. The endpoint is required to accept this command if it indicated support for a dynamic EID pool. The command must be generated by the endpoint if the configuration requires the endpoint to support allocating EID pools to downstream bridges. 7. See Section 8.19 MCTP Network IDs for information for implementation requirements of this command. 					

1676 **10.2 MCTP Control Message Completion Codes**

1677 The command/result code field is used to return management operation results for response messages. If
 1678 a `SUCCESS` completion code is returned then the specified response parameters (if any) must also be
 1679 returned in the response. If an error completion code is received (not `SUCCESS`) an MCTP endpoint
 1680 receives an error completion code, unless otherwise specified, the responder shall not return any
 1681 additional parametric data and the requester shall ignore any additional parameter data provided in the
 1682 response (if any). See Table 13 for the completion codes.

1683 **Table 13 – MCTP Control Message Completion Codes**

Value	Name	Description
0x00	SUCCESS	The Request was accepted and completed normally.
0x01	ERROR	This is a generic failure message. (It should not be used when a more specific result code applies.)
0x02	ERROR_INVALID_DATA	The packet payload contained invalid data or an illegal parameter value.
0x03	ERROR_INVALID_LENGTH	The message length was invalid. (The Message body was larger or smaller than expected for the particular request.)
0x04	ERROR_NOT_READY	The Receiver is in a transient state where it is not ready to receive the corresponding message.
0x05	ERROR_UNSUPPORTED_CMD	The command field in the control type of the received message is unspecified or not supported on this endpoint. This completion code must be returned for any unsupported command values received in MCTP control Request messages.
0x80–0xFF	COMMAND_SPECIFIC	This range of completion code values is reserved for values that are specific to a particular MCTP control message. The particular values (if any) and their definition is provided in the specification for the particular command.
all other	Reserved	Reserved

1684 **10.3 Set Endpoint ID**

1685 The Set Endpoint ID command assigns an EID to an endpoint. This command should only be issued by a
 1686 bus owner to assign an EID to an endpoint at a particular physical address. Since it is assumed the
 1687 Endpoint does not already have an EID assigned to it, or because the EID is unknown, the destination
 1688 EID in the message will typically be set to the special null destination EID value.

1689 The Set Endpoint ID command is also used to provide the Physical Address and EID of the Bus Owner to
 1690 an Endpoint. An Endpoint that needs to communicate with the Bus Owner may capture the physical
 1691 address and EID that was used to deliver the Set Endpoint ID message.

1692 NOTE: Endpoints that are not the Bus Owner should not issue the Set Endpoint ID command because it can cause
 1693 the receiver of the message to capture incorrect information for the Bus Owner's address.

1694 An MCTP bridge may elect to have a single EID for its functionality, rather than using an EID for each port
 1695 (bus connection) that is connected to a different bus owner. See 8.19.2 for more information. In this case,
 1696 the bridge will accept its EID assignment from the "first" bus to deliver the Set Endpoint ID request to the
 1697 bridge.

1698 It is recognized that different internal processing delays within a bridge can cause the temporal ordering
 1699 of requests to be switched if overlapping requests are received over more than one bus. Therefore, which
 1700 request is accepted by an implementation is not necessarily tied to the request that is first received at the
 1701 bridge, but instead will be based on which request is the first to be processed by the bridge.

1702 If an EID has already been assigned and the Set Endpoint ID command is issued from a different bus
 1703 without forcing an EID assignment, the command shall return a `SUCCESSFUL` completion code, but the
 1704 response parameters shall return an EID assignment status of "EID rejected".

1705 The Set Endpoint ID command functions in the same manner regardless of whether the endpoint uses a
 1706 static EID. The only difference is that if an endpoint has a static EID, it uses that EID as its initial "default"
 1707 EID value. The endpoint does not treat this initial EID as if it were assigned to it by a different bus owner.
 1708 That is, the endpoint shall accept the EID assignment from the first bus that the command is received
 1709 from, and shall track that bus as the originating bus for the EID for subsequent instances of Set Endpoint
 1710 ID command. See 8.17.2 for more information. The request and response parameters are specified in
 1711 Table 14.

1712

Table 14 – Set Endpoint ID Message

	Byte	Description
Request data	1	<p>Operation</p> <p>[7:3] – reserved</p> <p>[1:0] – Operation:</p> <p>00b Set EID. Submit an EID for assignment. The given EID will be accepted conditional upon which bus the device received the EID from (see preceding text). A device where the endpoint is only reached through one bus shall always accept this operation (provided the EID value is legal).</p> <p>01b Force EID. Force EID assignment. The given EID will be accepted regardless of whether the EID was already assigned through another bus. Note that if the endpoint is forcing, the EID assignment changes which bus is being tracked as the originator of the Set Endpoint ID command. A device where the endpoint is only reached through one bus shall always accept this operation (provided the EID value is legal), in which case the Set EID and Force EID operations are equivalent.</p> <p>10b Reset EID (optional). This option only applies to endpoints that support static EIDs. If static EIDs are supported, the endpoint shall restore the EID to the statically configured EID value. The EID value in byte 2 shall be ignored. An <code>ERROR_INVALID_DATA</code> completion code shall be returned if this operation is not supported.</p> <p>11b Set Discovered Flag. Set Discovered flag to the "discovered" state only. Do not change present EID setting. The EID value in byte 2 shall be ignored. Note that Discovered flag is only used for some physical transport bindings. An <code>ERROR_INVALID_DATA</code> completion code shall be returned if this operation is selected and the particular transport binding does not support a Discovered flag.</p>
	2	<p>Endpoint ID.</p> <p>0xFF, 0x00 = illegal. Endpoints are not allowed to be assigned the broadcast or null EIDs. It is recommended that the endpoint return an <code>ERROR_INVALID_DATA</code></p>

	Byte	Description
		completion code if it receives either of these values.
Response data	1	Completion code
	2	[7:6] – reserved [5:4] – EID assignment status: 00b = EID assignment accepted. 01b = EID assignment rejected. EID has already been assigned by another bus owner and assignment was not forced. 10b = reserved. 11b = reserved. [3:2] – reserved. [1:0] – Endpoint ID allocation status (see 10.10 for additional information): 00b = Device does not use an EID pool. 01b = Endpoint requires EID pool allocation. 10b = Endpoint uses an EID pool and has already received an allocation for that pool. 11b = reserved
	3	EID Setting. If the EID setting was accepted, this value will match the EID passed in the request. Otherwise, this value returns the present EID setting.
	4	EID Pool Size. This is the size of the dynamic EID pool that the bridge can use to assign EIDs or EID pools to other endpoints or bridges. It does not include the count of any additional static EIDs that the bridge may maintain. See 8.17.2 for more information. Note that a bridge always returns its pool size regardless of whether it has already received an allocation. 0x00 = no dynamic EID pool.

1713 **10.4 Get Endpoint ID**

1714 The Get Endpoint ID command returns the EID for an endpoint. This command is typically issued only by
 1715 a bus owner to retrieve the EID that was assigned to a particular physical address. Thus, the destination
 1716 EID in the message will typically be set to the special Physical Addressing Only EID value. The request
 1717 and response parameters are specified in Table 15.

1718 **Table 15 – Get Endpoint ID Message**

	Byte	Description
Request data	–	–
Response data	1	Completion Code.
	2	Endpoint ID. 0x00 = EID not yet assigned.
	3	Endpoint Type. [7:6] = reserved [5:4] = Endpoint Type:

	Byte	Description
		<p>00b = simple endpoint</p> <p>01b = bus owner/bridge</p> <p>10b = reserved</p> <p>11b = reserved</p> <p>[2:0] = reserved</p> <p>[1:0] = Endpoint ID Type:</p> <p>00b = dynamic EID. The endpoint uses a dynamic EID only.</p> <p>01b = static EID supported. The endpoint was configured with a static EID. The EID returned by this command reflects the present setting and may or may not match the static EID value.</p> <p>The following two status return values are optional. If provided, they must be supported as a pair in place of the static EID support status return. It is recommended that this be implemented if the Reset EID option in the Set Endpoint ID command is supported.</p> <p>10b = static EID supported. Present EID matches static EID. The endpoint has been configured with a static EID. The present value is the same as the static value.</p> <p>11b = static EID supported. Present EID does not match static EID. Endpoint has been configured with a static EID. The present value is different than the static value.</p> <p>See 8.17.2 for more information.</p>
	4	<p>Medium-Specific Information.</p> <p>This byte can hold additional information about optional configuration of the endpoint on the given medium, such as whether certain types of timing or arbitration are supported. This should only be used to report static information.</p> <p>This byte shall be returned as 0x00 unless otherwise specified by the transport binding.</p>

1719 10.5 Get Endpoint UUID Command

1720 The Get Endpoint UUID command returns a universally unique identifier (UUID), also referred to as a
 1721 globally unique ID (GUID), for the management controller or management device. The command can be
 1722 used to correlate a device with one or more EIDs. The format of the ID follows the byte (octet) format
 1723 specified in [RFC4122](#). [RFC4122](#) specifies four different versions of UUID formats and generation
 1724 algorithms suitable for use for a device UUID in IPMI. These are version 1 (0001b) "time based", and
 1725 three "name-based" versions: version 3 (0011b) "MD5 hash", version 4 (0100b) "Pseudo-random", and
 1726 version 5 "SHA1 hash". The version 1 format is recommended. However, versions 3, 4, or 5 formats are
 1727 also allowed. A device UUID should never change over the lifetime of the device. The request and
 1728 response parameters are specified in Table 16.

1729

Table 16 – Get Device UUID Message Format

	Byte	Description
Request data	–	–
Response data	1	Completion Code
	2:17	UUID bytes 1:16, respectively (see Table 17)

1730 The individual fields within the UUID are stored most-significant byte (MSB) first per the convention
 1731 described in [RFC4122](#). See Table 17 for an example format.

1732

Table 17 – Example UUID Format

Field	UUID Byte	MSB
time low	1	MSB
	2	
	3	
	4	
time mid	5	MSB
	6	
time high and version	7	MSB
	8	
clock seq and reserved	9	MSB
	10	
node	11	MSB
	12	
	13	
	14	
	15	
	16	

1733 **10.6 Get MCTP Version Support**

1734 This command can be used to retrieve the MCTP base specification versions that the endpoint supports,
 1735 and also the message type specification versions supported for each message type. The format of the
 1736 request and response parameters for this message is given in Table 18.

1737 More than one version number can be returned for a given message type by the Get MCTP Version
 1738 Support command. This enables the command to be used for reporting different levels of compatibility
 1739 and backward compatibility with different specification versions. The individual specifications for the given
 1740 message type define the requirements for which versions number values should be used for that
 1741 message type. Those documents define which earlier version numbers, if any, must also be listed.

1742 The command returns a completion code that indicates whether the message type number passed in the
 1743 request is supported or not. This enables the command to also be used to query the endpoint for whether
 1744 it supports a given message type.

1745 NOTE: Version numbers are listed from oldest to newest. This specification is backward compatible with version
 1746 1.0.x.x. Therefore, the response data shall return two Version Number entries: 1.0.x.x for Version Number entry 1
 1747 and 1.1.x.x for Version Number entry 2.

1748

Table 18 – Get MCTP Version Support Message

	Byte	Description
Request data	1	<p>Message Type Number</p> <p>The Message Type Number to retrieve version information for:</p> <p>0xFF = return MCTP base specification version information.</p> <p>0x00 = return MCTP control protocol message version information.</p> <p>other = return version information for a given message type. See MCTP ID and Table 3 for message type numbers.</p>
Response data	1	<p>Completion Code</p> <p>0x80 = message type number not supported</p>
	2	<p>Version Number Entry count</p> <p>One-based count of 32-bit version numbers being returned in this response. Numerically lower version numbers are returned first.</p>
	3:6	<p>Version Number entry 1: The following descriptions are informational. Refer to DSP4004 for the normative definition of version numbering of DMTF specifications.</p> <p>[31:24] = major version number. This field is used to identify a version of the specification that includes changes that make it incompatible with one or more functions that were defined in versions of the specification that have an older (smaller) major version number.</p> <p>[23:16] = minor version number. This field is used to identify functional additions to the specification that are backward compatible with older (smaller) minor version numbers that share the same major version number.</p> <p>[15:8] = update version number. This field is used for editorial updates to the specification that do not define new functionality nor change existing functionality over the given major.minor release. This field is informational and should be ignored when checking versions for interoperability.</p> <p>[7:0] = "alpha" byte. This value is used for pre-release (work-in-progress) versions of the specification. Pre-release versions of the specification are backward compatible with specification versions that have an older (smaller) minor version numbers that share the same major version number. However, since the alpha value represents a version of the specification that is presently under development, versions that share the same major and minor version numbers, but have different 'alpha' versions may not be fully interoperable.</p> <p>The encoding of the version number and alpha fields is provided in 10.6.1.</p>
	(7:X)	<p>Version Number Entries 2 through N.</p> <p>Additional 32-bit major/minor version numbers, if any.</p>

1749 10.6.1 Version Field Encoding

1750 The version field is comprised of four bytes referred to as the "major", "minor", "update", and "alpha"
1751 bytes. These bytes shall be encoded as follows:

1752 The "major", "minor", and "update" bytes are BCD-encoded, and each byte holds two BCD digits. The
1753 "alpha" byte holds an optional alphanumeric character extension that is encoded using the
1754 ISO/IEC 8859-1 Character Set. The semantics of these fields follows that specified in [DSP4004](#).

1755 The value 0x00 in the alpha field means that the alpha field is not used. Software or utilities that display
1756 the version number should not display any characters for this field.

1757 The value 0xF in the most-significant nibble of a BCD-encoded value indicates that the most-significant
1758 nibble should be ignored and the overall field treated as a single-digit value. Software or utilities that
1759 display the number should only display a single digit and should not put in a leading "0" when displaying
1760 the number.

1761 A value of 0xFF in the "update" field indicates that the field to be ignored. Software or utilities that display
1762 the version number should not display any characters for the field. 0xFF is not allowed as a value for the
1763 "major" or "minor" fields.

1764 EXAMPLES:

1765 Version 1.1.0 → 0xF1F1F000

1766 Version 3.1 → 0xF3F1FF00

1767 Version 1.0a → 0xF1F0FF61

1768 Version 3.7.10a → 0xF3F71061

1769 Version 10.11.7 → 0x1011F700

1770 10.6.2 MCTP Base Specification Version Number

1771 The Version Number Entry 1 field shall be used to indicate backward compatibility with Version 1.0 of the
1772 base specification as:

1773 **1.0** [Major version 1, minor version 0, no update version, no alpha)]

1774 This is reported using the encoding as: 0xF1F0FF00

1775 The version of the MCTP base specification for this specification shall be reported in Version Number
1776 Entry 2 as:

1777 **1.1.0** [Major version 1, minor version 1, update version 0, no alpha)]

1778 This is reported using the encoding as: 0xF1F1F000

1779 10.7 Get Message Type Support

1780 The Get Message Type Support command enables management controllers to discover the MCTP
1781 control protocol capabilities supported by other MCTP endpoints, and get a list of the MCTP message
1782 types that are supported by the endpoint. The request and response parameters for this message are
1783 listed in Table 19.

1784 The response to this command may be specific according to which bus the request was received over
1785 (that is, a device that supports a given message type may not support that message type equally across
1786 all buses that connect to the device).

1787

Table 19 – Get Message Type Support Message

	Byte	Description
Request data	–	–
Response data	1	Completion Code.
	2	MCTP Message Type Count. One-based. Number of message types in addition to the MCTP control message type that is supported by this endpoint
	(3:N)	List of Message Type numbers. One byte per number. See Table 3 and MCTP ID .

1788 10.8 Get Vendor Defined Message Support

1789 The Get Vendor Defined Message Support operation enables management controllers to discover
 1790 whether the endpoint supports vendor-defined messages, and, if so, the vendors or organizations that
 1791 defined those messages. The format and definition of the request and response parameters for this
 1792 message is given in Table 20.

1793

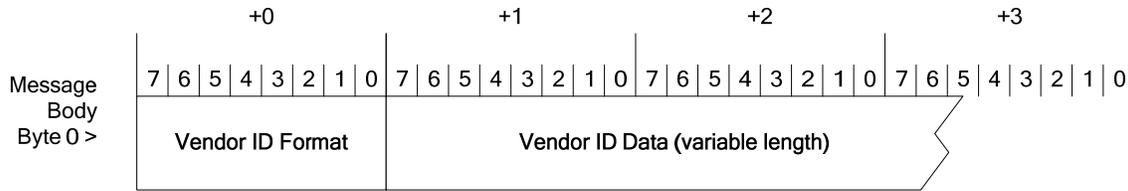
Table 20 – Get Vendor Defined Message Support Message

	Byte	Description
Request data	1	Vendor ID Set Selector Indicates the specific capability set requested. Indices start at 0x00 and increase monotonically by 1. If the responding endpoint has one or more capability sets with indices greater than the requested index, it increments the requested index by 1 and returns the resulting value in the response message. The requesting endpoint uses the returned value to request the next capability set.
	Response data	
Response data	1	Completion Code
	2	Vendor ID Set Selector 0xFF = no more capability sets.
	var	Vendor ID A structured field of variable length that identifies the vendor ID format (presently PCI or IANA) and the ID of the vendor that defined the capability set. The structure of this field is specified in Figure 20.
	2 bytes	16-bit numeric value or bit field, as specified by the vendor or organization identified by the vendor ID. This value is typically used to identify a particular command set type or major version under the given vendor ID.

1794 10.8.1 Vendor ID Formats

1795 Figure 20 shows the general structure of Vendor ID fields used in this specification. The first byte of the
 1796 field contains the Vendor ID Format, a numeric value that indicates the definition space and format of the
 1797 ID. The remainder of the field holds the Vendor ID Data with content and format as specified in Table 21.

1798 The MCTP management controller or management device can pick which format is best suited for the
 1799 device. In general, if the device does not already have an existing vendor ID that matches one of the
 1800 specified formats, it is recommended that the IANA enterprise number format be used.



1801

1802

Figure 20 – Structure of Vendor ID Field for Get Vendor Defined Capabilities Message

1803

Table 21 – Vendor ID Formats

Vendor ID Format Name	Vendor ID Format	Vendor ID Data Length	Description
PCI Vendor ID	0x00	2	16-bit Unsigned Integer. The PCI 2.3 specifications state the following about the PCI vendor ID: "This field identifies the manufacturer of the device. Valid vendor identifiers are allocated by the PCI SIG to ensure uniqueness. 0xFFFF is an invalid value for the Vendor ID." However, for MCTP this value may be used for identifying aspects other than the manufacturer of the device, such as its use in the Vendor Defined - PCI message type, where it identifies the vendor or organization that defined a particular set of vendor-defined messages. Thus, in some uses, the ID may or may not correspond to the PCI ID for the manufacturer of the device.
IANA Enterprise Number	0x01	4	32-bit Unsigned Integer. The IANA enterprise number for the organization or vendor expressed as a 32-bit unsigned binary number. For example, the enterprise ID for the DMTF is 412 (decimal) or 0x0000_019C expressed as a 32-bit hexadecimal number. The enterprise number is assigned and maintained by the Internet Assigned Numbers Authority, www.iana.org , as a means of identifying a particular vendor, company, or organization.

1804 **10.9 Resolve Endpoint ID**

1805 This command is sent to the bus owner to resolve an EID into the physical address that must be used to
 1806 deliver MCTP messages to the target endpoint. The command takes an EID as an input parameter in the
 1807 request and returns the EID and the physical address for routing to that EID (if any) in the response. The
 1808 response data will also indicate if no mapping was available.

1809 An endpoint knows the physical address of the bus owner by keeping track of which physical address
 1810 was used when the endpoint received its EID assignment through the Set Endpoint ID command. The
 1811 endpoint can send this command to the bus owner using the null destination EID value. This eliminates
 1812 the need for the endpoint to also keep track of the EID of the bus owner. The request and response
 1813 parameters are specified in Table 22.

1814

Table 22 – Resolve Endpoint ID Message

	Byte	Description
Request data	1	Target Endpoint ID This is the EID that the bus owner is being asked to resolve.
Response data	1	Completion Code
	2	Bridge Endpoint ID This is the EID for the endpoint that is providing the bridging server (if any) that is required to access the target endpoint. If the EID being returned matches the same value as the target EID, it indicates that there is no bridging function that is required to access the target endpoint (that is, the target EID is local to the bus that the Resolve Endpoint ID request was issued over).
	3:N	Physical Address. The size of this field is dependent on the particular MCTP physical transport binding used for the bus that this data is being provided for. The size and format of this field is defined as part of the corresponding physical transport binding specification.

1815 10.10 Allocate Endpoint IDs

1816 Bus owners are responsible for allocating pools of EIDs to MCTP bridges that are lower in the bus
1817 hierarchy. This is done using the Allocate Endpoint IDs command. The EID for the bridge itself is
1818 assigned separately and is *not* part of the pool given with this command.

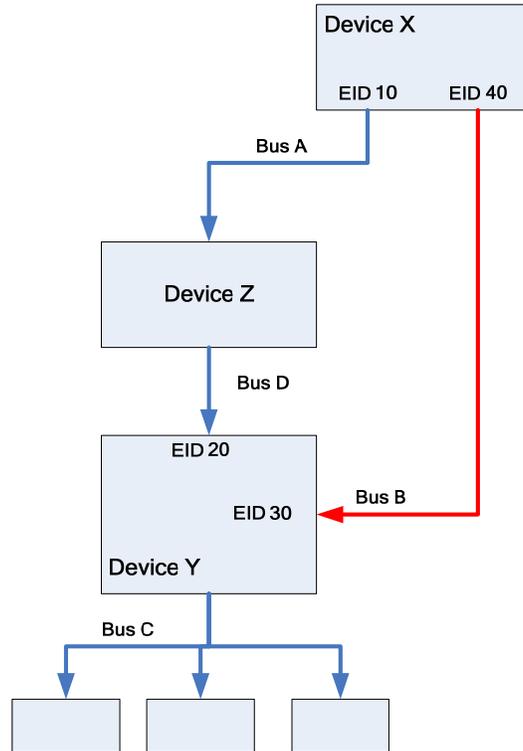
1819 The bus owner will typically use this command as part of the EID assignment process for a bus. When a
1820 device has been assigned an EID using the Set Endpoint ID command, the response to that command
1821 indicates whether the endpoint supports an EID pool. If the device indicates that it supports an EID pool,
1822 the bus owner can then issue the Allocate Endpoint IDs command to supply the pool of EIDs to the
1823 device.

1824 NOTE: The Allocate Endpoint IDs command can also cause a bridge to rebuild its routing table. See 10.11.2 for
1825 more information.

1826 When an EID or EID pool that was previously allocated becomes unused (for example, due to a hot-swap
1827 removal), the bus owner must reclaim the endpoint's EID or EID pool allocation. See 8.17 for additional
1828 details.

1829 Referring to Figure 21, there is a potential race condition with handling EID allocation. In the scenario
1830 shown in this figure, it is possible that device X and device Z might both be assigning EIDs to device Y at
1831 the same time. This also means that, unless steps are taken, device Z could allocate endpoints to device
1832 Y only to have this overwritten by a set of endpoints assigned by device X.

1833 To prevent this, the Allocate Endpoint IDs command is only accepted from the "first" bus that provides the
1834 EID pool to the device. If another bus owner attempts to deliver an EID pool through another bus, the
1835 request will be rejected unless an intentional over-ride is done.



1836

1837

Figure 21 – EID Pools from Multiple Bus Owners

1838 The Allocate Endpoint IDs message fields are described in Table 23.

1839

Table 23 – Allocate Endpoint IDs Message

Request data	Byte	Description
	1	<p>Operation Flags:</p> <p>[7:1] – reserved.</p> <p>[1:0] – Operation:</p> <p>00b = allocate EIDs. Submit an EID pool allocation. Do not force allocation. This enables the allocation to be rejected if the bridge has already received its EID pool from another bus. (See additional information in the following sections.)</p> <p>01b = Force allocation. Force bridge to accept this EID pool regardless of whether it has already received its EID pool from another bus. This shall also cause a bridge to rebuild its routing tables, See 10.11.2 for more information.</p> <p>10b = Get allocation information Return the response parameters without changing the present allocation. This can be used to query information on the dynamic pool of EIDs presently allocated to the Endpoint, if any. If this operation is selected, the Number of Endpoint IDs and Starting Endpoint ID parameters in the request shall be ignored.</p> <p>11b = Reserved</p>

1840

	Byte	Description
	2	Number of Endpoint IDs Specifies the number of EIDs in the pool being made available to this Endpoint
	3	Starting Endpoint ID Specifies the starting EID for the range of EIDs being allocated in the pool. When multiple EIDs are provided, the IDs are sequential starting with this value as the first EID in the range.
Response data	1	Completion Code An error completion code shall be returned if the number of EIDs assigned does not match the pool size. (This error condition does not apply to when the number of endpoint IDs passed in the request is 0xFF or 0x00).
	2	[7:2] – reserved [1:0] – 00b = Allocation was accepted 01b = Allocation was rejected. The Allocate Endpoint IDs command is accepted only from the "first" bus that provides the EID pool to the device. If another bus owner attempts to deliver an EID pool through another bus, the request will be rejected unless an intentional over-ride is done. (The rationale for this behavior is explained in the text of this section.) 10b, 11b = reserved
	3	Endpoint ID Pool Size This value is the size of the EID pool used by this endpoint. This is the size of the dynamic EID pool that the bridge can use to assign EIDs or EID pools to other endpoints or bridges. It does not include the count of any additional static EIDs that the bridge may maintain. See 8.17.2 for more information.
	4	First Endpoint ID This field specifies the first EID assigned to the pool for this endpoint. The value is 0x00 if there are no EIDs assigned to the pool.

1841 10.11 Routing Information Update

1842 The Routing Information Update message is used by a bus owner to give routing information to a bridge
1843 for the bus on which the message is being received.

1844 Because the physical address format is based on the bus over which the request is delivered, the bus
1845 owner must use the medium-specific physical address format for the addresses sent using this command.

1846 An MCTP bridge may be sent more than one instance of this command to transfer the update information.
1847 An integral number of routing information update entries must be provided in the command (that is,
1848 routing information update entries cannot be split across instances of the command).

1849 10.11.1 Adding and Replacing Entries

1850 The recipient of this command must check to see whether the information in the request corresponds to
1851 the EID for an existing entry for the bus over which the command was received. If so, it must replace that
1852 entry with the new information. If an entry for a given EID or EID range does not already exist, it must
1853 create new entries for the given EIDs. In some cases this may require the bridge to split existing entries
1854 into multiple entries.

1855 NOTE: A bus owner is only allowed to update entries that correspond to its bus. For each routing table entry that
1856 was created or updated through the Routing Information Update message, the bridge must keep track of which bus it
1857 received the Routing Information Update from. This is necessary so that when a Routing Information Update is

1858 received from a particular bus, the bridge only updates entries that correspond to entries that were originally given to
 1859 it from that bus.

1860 **10.11.2 Rebuilding Routing Tables**

1861 A bridge that receives and accepts the Allocate Endpoint IDs command with the "Force Allocation" bit set
 1862 (1b) must clear out and rebuild its routing table information. The bridge shall issue commands to reassign
 1863 EIDs and re-allocate EID pools to all downstream devices. The request and response parameters are
 1864 specified in Table 24, and format information is provided in Table 25.

1865 **Table 24 – Routing Information Update Message**

	Byte	Description
Request data	1	Count of update entries (1-based)
	see text	One or more update entries, based on the given count, as illustrated in Table 25
Response data	1	Completion Code 0x80 = Insufficient space to add requested entries to internal routing table

1866 **Table 25 – Routing Information Update Entry Format**

Byte	Description
1	[7:4] - reserved [3:0] - Entry Type: 00b = entry corresponds to a single endpoint that is not serving as an MCTP bridge 01b = entry reflects an EID range for a bridge where the starting EID is the EID of the bridge itself and additional EIDs in the range are routed by the bridge 10b = entry is for a single endpoint that is serving as an MCTP bridge 11b = entry is an EID range for a bridge, but does not include the EID of the bridge itself
2	[7:0] Size of EID Range. The count of EIDs in the range.
3	First EID in EID Range. The EID Range is sequential (for example, if the size of the EID Range is 3 and the First EID value given in this parameter is 21, the Entry covers EIDs 21, 22, and 23).
4:N	Physical Address. The size and format of this field is defined as part of the corresponding physical transport binding specification for the bus that this data is being provided for.

1867 **10.12 Get Routing Table Entries**

1868 This command can be used to request an MCTP bridge to return data corresponding to its present routing
 1869 table entries. This data is used to enable troubleshooting the configuration of routing tables and to enable
 1870 software to draw a logical picture of the MCTP network. More than one instance of this command will
 1871 typically need to be issued to transfer the entire routing table content.

1872 An integral number of routing table entries must be provided in the response to this command (that is,
 1873 routing table entries cannot be split across instances of the command). The request and response
 1874 parameters are specified in Table 26, and format information is provided in Table 27.

1875 **Table 26 – Get Routing Table Entries Message**

	Byte	Description
Request data	1	Entry Handle (0x00 to access first entries in table)
Response data	1	Completion Code
	2	Next Entry Handle (Use this value to request the next set of entries, if any.) If the routing table data exceeds what can be carried in a single MCTP control response. 0xFF = No more entries
	3	Number of routing table entries being returned in this response
	4:N	One or more routing table entries, formatted per Table 27. This field will be absent if the number of routing table entries is 0x00.

1876 **Table 27 – Routing Table Entry Format**

Byte	Description
1	Size of EID range associated with this entry
2	Starting EID
3	<p>Entry Type/Port Number</p> <p>[7:6] – Entry Type:</p> <ul style="list-style-type: none"> 00b = entry corresponds to a single endpoint that does not operate as an MCTP bridge 01b = entry reflects an EID range for a bridge where the starting EID is the EID of the bridge itself and additional EIDs in the range are routed by the bridge 10b = entry is for a single endpoint that serves as an MCTP bridge 11b = entry is an EID range for a bridge, but does not include the EID of the bridge itself <p>[5] – Dynamic/Static Entry.</p> <p>Indicates whether the entry was dynamically created or statically configured. Note that statically configured routing information shall not be merged with dynamic information when reporting entry information using this command. While an implementation may internally organize its data that way, dynamic and statically configured routing must be reported as separate entries. Dynamically created entries include entries that were generated from the Routing Information Update command as well as entries that were created as a result of the bridge doing EID assignment and EID pool allocation as a bus owner.</p> <ul style="list-style-type: none"> 0b = Entry was dynamically created 1b = Entry was statically configured <p>[4:0] – Port number</p> <p>This value is chosen by the bridge device vendor and is used to identify a particular bus connection that the physical address for the entry is defined under. In some cases, this number may correspond to an internal "logical" bus that is not directly connected to an external physical bus. Port numbers are required to be static.</p> <p>It is recommended, but not required, that the ports (bus connections) on the bridge be numbered sequentially starting from 0x00. This specification does not define any requirements or recommendations on how port numbers are assigned to corresponding physical connections on a device.</p>

Byte	Description
4	Physical Transport Binding Type Identifier, according to MCTP ID
5	Physical Media Type Identifier, according to MCTP ID . This value is used to indicate what format the following physical address data is given in.
6	Physical Address Size. The size in bytes of the following Physical Address field The size is defined as part of the corresponding physical transport binding specification identified by the physical media type identifier.
7 :N	Physical Address. The size and format of this field is defined as part of the corresponding physical transport binding specification. The information given in this field is given MSB first. Any unused bits should be set to 0b.

1877 **10.13 Prepare for Endpoint Discovery**

1878 The Endpoint Discovery message is used to determine if devices on a bus communicate MCTP (see
1879 Table 28). Whether this message is required depends on the particular medium. Currently, this message
1880 may be required only by a particular transport binding, such as PCI Express (PCIe) VDM, because other
1881 bindings such as SMBus may use other mechanisms for determining this information.

1882 Each endpoint (except the bus owner) on the bus maintains an internal flag called the "Discovered" flag.

1883 The Prepare for Endpoint Discovery command is issued as a broadcast Request message on a given bus
1884 that causes each endpoint on the bus to set their respective Discovered flag to the "undiscovered" state.
1885 The flag is subsequently set to the "discovered" state when the Set Endpoint ID command is received by
1886 the endpoint.

1887 An endpoint also sets the flag to the "undiscovered" state at the following times:

- 1888 • Whenever the physical address associated with the endpoint changes or is assigned
- 1889 • Whenever an endpoint first appears on the bus and requires an EID assignment
- 1890 • During operation if an endpoint enters a state that requires its EID to be reassigned
- 1891 • For hot-plug endpoints: After exiting any temporary state where the hot-plug endpoint was
1892 unable to respond to MCTP control requests for more than $T_{RECLAIM}$ seconds (where $T_{RECLAIM}$ is
1893 specified in the physical transport binding specification for the medium used to access the
1894 endpoint). See 8.17.5 for additional information.

1895 Only endpoints that have their Discovered flag set to "undiscovered" will respond to the Endpoint
1896 Discovery message. Endpoints that have the flag set to "discovered" will not respond.

1897 The destination EID for the Prepare for Endpoint Discovery message is set to the Broadcast EID value
1898 (see Table 2) in the request message to indicate that this is a broadcast message. The response
1899 message sets the destination EID to be the ID of the source of the request message, which is typically the
1900 EID of the bus owner. The request and response parameters are specified in Table 28.

1901 **Table 28 – Prepare for Endpoint Discovery Message**

	Byte	Description
Request data	–	–
Response data	1	Completion Code

1902 10.14 Endpoint Discovery

1903 This command is used to discover endpoints that have their Discovered flag set to "undiscovered". Only
 1904 endpoints that have their Discovered flag set to "undiscovered" will respond to this message. Endpoints
 1905 that have the flag set to "discovered" will not respond.

1906 This message is typically sent as a Broadcast Request message by the bus owner using the Broadcast
 1907 EID as the destination EID, though for testing purposes endpoints must also accept and handle this
 1908 command as a non-broadcast Request. Additionally, the request may be sent as a datagram, depending
 1909 on the transport binding requirements. The request and response (if any) parameters are specified in
 1910 Table 29.

1911 **Table 29 – Endpoint Discovery Message**

	Byte	Description
Request data	–	–
Response data	1	Completion Code

1912 10.15 Discovery Notify

1913 This message is available for use as a common message for enabling an endpoint to announce its
 1914 presence to the bus owner. This will typically be used as part of the endpoint discovery process when an
 1915 MCTP device is hot-plugged onto or becomes powered up on an MCTP bus.

1916 Whether and how this message is used for endpoint discovery depends on the particular physical
 1917 transport binding specification. For example, the SMBus transport binding does not use this message for
 1918 an endpoint to announce itself because it takes advantage of mechanisms that are already defined for
 1919 SMBus.

1920 This message should only be sent from endpoints to the bus owner for the bus that the endpoint is on so
 1921 it can notify the bus owner that the endpoint has come online and may require an EID assignment or
 1922 update. Additionally, the request may be sent as a datagram, depending on the transport binding
 1923 requirements. The request and response (if any) parameters are specified in Table 30.

1924 **Table 30 – Discovery Notify Message**

	Byte	Description
Request data	–	–
Response data	1	Completion Code

1925 10.16 Query Hop

1926 This command can be used to query a bridge to find out whether a given EID must be accessed by going
 1927 through that bridge, and if so, whether yet another bridge must be passed through in the path to the
 1928 endpoint, or if the endpoint is on a bus that is directly connected to the bridge.

1929 The command also returns the information about the transmission unit information that the bridge
 1930 supports in routing to the given target endpoint from the bus that the request was received over. See
 1931 section 8.23 for more information.

1932 NOTE: The physical transport binding for MCTP may place additional requirements on the physical packet sizes
 1933 that can be used to transfer MCTP packet payloads, such as requiring that physical packet sizes be in 32-byte or 64-
 1934 byte increments, or particular power of 2 increments (for example, 128, 256, 512, and so on).

1935 The request and response parameters are specified in Table 31.

1936 **Table 31 – Query Hop Message**

	Byte	Description
Request data	1	Target Endpoint ID 0x00, 0xFF = reserved. (An ERROR_INVALID_DATA completion code shall be returned.)
	2	Message type for which transmission unit information is being requested. Use the MCTP control message type number unless another message type is of interest.
Response data	1	Completion Code An ERROR_INVALID_DATA completion code must be returned if the target EID is not covered by any entry in the bridge's routing table.
	2	EID of the next bridge that is used to access the target endpoint, if any NOTE: This response depends on which bus port the Query Hop request is received over. If this EID is 00h: The EID is covered by the bridge's routing table, but the target EID does not require access by <i>going through</i> this bridge from the port the request was received over. This response will be returned if the target EID is already local to the bus over which the request is being received. This response is also returned when the target EID is an EID for the bridge itself. If this EID is non-zero <i>and</i> is different than the target EID passed in request: The EID being provided is the EID of the "next bridge" in the path to the target EID. If this EID is equal to the target EID passed in request: The target EID is accessed by going through this bridge and no additional bridges must be gone through to reach the target.
	3	Message Type. This value either returns the message type that was given in the request, or it returns 0xFF to indicate that the information is applicable to all message types that are supported by the bridge.
	4:5	Maximum supported incoming transmission unit size in increments of 16 bytes, starting from the baseline transmission unit size (0x0000 = 64 bytes, 0x0001 = 80 bytes, and so on).
	5:6	Maximum supported outgoing transmission unit size in increments of 16 bytes, starting from the baseline transmission unit (0x0000 = 64 bytes, 0x0001 = 80 bytes, and so on). The responder will return whether this transmission unit size is supported for MCTP packets that it transmits for the given message type.

1937 **10.17 Get Network ID Command**

1938 The Get Network ID command returns a universally unique identifier (UUID), also referred to as a globally
 1939 unique ID (GUID), for a given MCTP network. Typically this command is sent to the topmost MCTP bus-
 1940 owner since the topmost bus-owner has this knowledge. A Network ID is required for add-in MCTP
 1941 networks (For example, an MCTP Network on an add-in card or module). A Network ID is not required for
 1942 a fixed (not add-in) MCTP network provided there is only one network in the system implementation. A

1943 Network ID is required for fixed MCTP networks when more than one fixed network exists in the system
 1944 implementation and is simultaneously accessible by a common entity such as system software.

1945 The format of the ID follows the byte (octet) format specified in [RFC4122](#). [RFC4122](#) specifies four
 1946 different versions of UUID formats and generation algorithms suitable for use for a device UUID in IPMI.
 1947 These are version 1 (0001b) "time based", and three "name-based" versions: version 3 (0011b) "MD5
 1948 hash", version 4 (0100b) "Pseudo-random", and version 5 "SHA1 hash". The version 1 format is
 1949 recommended. However, versions 3, 4, or 5 formats are also allowed. A device UUID should never
 1950 change over the lifetime of the device. The request and response parameters are specified in Table 16.

1951

Table 32 – Get Network ID Message Format

	Byte	Description
Request data	–	–
Response data	1	Completion Code
	2:17	Network ID bytes 1:16, respectively (see Table 17)

1952 The individual fields within the UUID are stored most-significant byte (MSB) first per the convention
 1953 described in [RFC4122](#). See Table 17 for an example format.

1954 10.18 Transport Specific

1955 Transport Specific commands are a range of commands that are available for use by transport binding
 1956 specifications in order to perform additional MCTP Control functions that are defined by a particular
 1957 transport binding. Transport specific commands shall only be addressed to endpoints on the same
 1958 medium. A bridge is allowed to block transport specific commands from being bridged to different media.

1959 The request and response parameters are specified in Table 33.

1960

Table 33 – Transport Specific Message

	Byte	Description
Request data	1	MCTP Physical Transport Binding Identifier The ID of the Physical Transport specification that defines the transport specific message. This ID is defined in the MCTP ID companion document to this specification.
	2	MCTP Physical Media Identifier The ID of the physical medium that the message is targeted for. This ID is defined in the MCTP ID companion document to this specification.
	3:N	Transport specific command data. Defined by the transport binding specification identified by the MCTP Physical Transport Binding Identifier given in byte 1. If the Physical Transport Binding Identifier = Vendor Defined: The first four bytes of data shall be the IANA Enterprise ID for the Vendor. MSB first. See 10.8.1 for the information on the IANA Enterprise ID as used in this specification.
Response data	1	Completion Code

1961 **11 Vendor Defined – PCI and Vendor Defined – IANA Messages**

1962 The Vendor Defined – PCI and Vendor Defined – IANA message types provide a mechanism for
 1963 providing an MCTP message namespace for vendor-specific messages over MCTP.

1964 The PCI and IANA designations refer to the mechanism that is used to identify the vendor or organization
 1965 this is specifying the message's functionality and any parametric data or other fields provided in the
 1966 message body.

1967 Note that this specification only defines the initial bytes in the message body of these messages, and sets
 1968 the requirement that these messages must follow the requirements set by the MCTP base protocol and
 1969 any additional requirements necessary to meet the transport of these messages over a particular
 1970 medium, such as path transmission unit limitations.

1971 Otherwise, any other field definitions and higher level message behavior such as retries, error/completion
 1972 codes, and so on, is message type-specific and thus is vendor-specific.

1973 **11.1 Vendor Defined – PCI Message Format**

1974 For these messages, the MCTP message type is set to the value for "Vendor Defined – PCI" as defined in
 1975 Table 3. The request and response parameters are specified in Table 34.

1976 **Table 34 – Vendor Defined – PCI Message Format**

	Byte	Description
Request data	1:2	PCI/PCIe Vendor ID. Refer to PCIe . MSB first. This value is formatted per the Vendor Data Field for the PCI Express vendor ID format. See 10.8.1". NOTE: Because the vendor ID format is implied by the command, the Vendor ID Format bytes are not part of this field.
	(3:N)	Vendor-Defined Message Body. 0 to N bytes.
Response data	1:2	PCI/PCIe Vendor ID. Refer to PCIe . MSB first.
	(3:M)	Vendor-Defined Message Body. 0 to M bytes.

1977 **11.2 Vendor Defined – IANA Message Format**

1978 For these messages, the MCTP message type is set to the value for "Vendor Defined – IANA" as defined
 1979 in Table 3. The request and response parameters are specified in Table 35.

1980 **Table 35 – Vendor Defined – IANA Message Format**

	Byte	Description
Request data	1:4	IANA Enterprise ID for Vendor. MSB first. This value is formatted per the Vendor Data Field for the IANA enterprise vendor ID format. See 10.8.1. NOTE: Because the vendor ID format is implied by the command, the Vendor ID Format bytes are not part of this field.
	(5:N)	Vendor-Defined Message Body. 0 to N bytes.
Response data	1:4	IANA Enterprise ID for the Vendor. MSB first.
	(5:M)	Vendor-Defined Message Body. 0 to M bytes.

1981

ANNEX A (informative)

Notation

1982
1983
1984
1985

1986 A.1 Notations

1987 Examples of notations used in this document are as follows:

- 1988 • 2:N In field descriptions, this will typically be used to represent a range of byte offsets
1989 starting from byte two and continuing to and including byte N. The lowest offset is on
1990 the left, and the highest is on the right.
- 1991 • (6) Parentheses around a single number can be used in message field descriptions to
1992 indicate a byte field that may be present or absent.
- 1993 • (3:6) Parentheses around a field consisting of a range of bytes indicates the entire range
1994 may be present or absent. The lowest offset is on the left, and the highest is on the
1995 right.
- 1996 • PCIe Underlined, blue text is typically used to indicate a reference to a document or
1997 specification called out in 2, "Normative References", or to items hyperlinked within
1998 the document.
- 1999 • rsvd Abbreviation for Reserved. Case insensitive.
- 2000 • [4] Square brackets around a number are typically used to indicate a bit offset. Bit offsets
2001 are given as zero-based values (that is, the least significant bit [LSb] offset = 0).
- 2002 • [7:5] A range of bit offsets. The most-significant is on the left, and the least-significant is on
2003 the right.
- 2004 • 1b The lower case "b" following a number consisting of 0s and 1s is used to indicate the
2005 number is being given in binary format.
- 2006 • 0x12A A leading "0x" is used to indicate a number given in hexadecimal format.

2007

ANNEX B
(informative)

Change Log

2008
2009
2010
2011
2012

Version	Date	Description
1.0.0	2009-07-28	Released as DMTF Standard
1.1.0	2010-04-22	Released as DMTF Standard, with the following changes: <ul style="list-style-type: none"> • Updated the glossary and the overview section, including additions for MCTP host interfaces and descriptions of MCTP networks. • Added support for MCTP network IDs and the Get Network ID command. • Addressed Mantis issue: 0000417. • Added text to Clause 1 (Scope) referencing DSP0238 and DSP0239 per WG ballot comments.

2013

Bibliography

2014

2015 DMTF DSP2016, *Management Component Transport Protocol (MCTP) Overview White Paper*,
2016 http://www.dmtf.org/standards/published_documents/DSP2016.pdf

2017 DMTF DSP4004, *DMTF Release Process 2.2*,
2018 http://www.dmtf.org/standards/published_documents/DSP4004_2.2.pdf

2019