CXL™ Fabric Manager API over MCTP Binding Specification

Supersedes: None

Document Class: Normative

Document Status: Published

Document Language: en-US
Copyright Notice

Copyright © 2021 DMTF. All rights reserved.

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

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patent owner or claimant, and shall have no liability or responsibility for costs or losses incurred if a standard is withdrawn or modified after publication, and shall be indemnified and held harmless by any party implementing the standard from any and all claims of infringement by a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

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

This document’s normative language is English. Translation into other languages is permitted.
CONTENTS

42 Foreword .......................................................................................................................... 5
43 Introduction....................................................................................................................... 6
44 Document conventions..................................................................................................... 6
45 1 Scope .............................................................................................................................. 7
46 2 Normative references.................................................................................................... 7
47 3 Terms and definitions .................................................................................................. 8
48 4 Symbols and abbreviated terms ................................................................................... 8
49 5 Conventions .................................................................................................................. 9
50 5.1 Reserved and unassigned values .............................................................................. 9
51 5.2 Byte ordering ............................................................................................................ 9
52 6 Overview ....................................................................................................................... 9
53 7 Message Type-specific considerations ......................................................................... 10
54 7.1 Message Type number ............................................................................................. 10
55 7.2 CXL FM API over MCTP specification version information ..................................... 10
56 7.3 Timing specifications.................................................................................................. 10
57 7.4 CXL FM API over MCTP message format ................................................................. 10
58 7.4.1 Integrity Check (IC), Tag Owner (TO) and Message Tag (Msg Tag) usage .......... 11
59 7.4.2 Message assembly ............................................................................................... 11
60 7.5 Maximum message size .......................................................................................... 11
61 7.6 Multiple MCTP transports ....................................................................................... 12
62 ANNEX A (informative) Change log ............................................................................ 13
63
Figures

Figure 1: MCTP Message fields ........................................................................................................ 10

Tables

Table 1: CXL FM API over MCTP Message field descriptions .................................................... 11
Foreword

The CXL™ (Compute Express Link™) Fabric Manager API over MCTP Binding Specification (DSP0234) was prepared by the Platform Management Components Intercommunications (PMCI Working Group) of the DMTF.

DMTF is a not-for-profit association of industry members dedicated to promoting enterprise and systems management and interoperability. For information about the DMTF, see http://www.dmtf.org.

The CXL Consortium is an open industry, standards group formed to develop technical specifications that facilitate breakthrough performance for emerging usage models while supporting an open ecosystem for data center accelerators and other high-speed enhancements. For information about the CXL consortium, see https://www.computeexpresslink.org.

Acknowledgments

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

Editors:
- Balaji Natrajan – Microchip Technology Inc
- Mahesh Natu – Intel Corporation

DMTF Contributors:
- Patrick Caporale – Lenovo
- Yuval Itkin – NVIDIA Corporation
- Eliel Louzoun – Intel Corporation
- Hemal Shah – Broadcom Inc.
- Bob Stevens – Dell Technologies

CXL Consortium Contributors:
- Vincent Hache – Microchip Technology Inc.
- Ariel Sibley – Microchip Technology Inc.
The CXL™ Fabric Manager API over MCTP Binding Specification defines a new MCTP message type used to convey CXL™ Fabric Manager API Messages over MCTP to devices.

Document conventions

Typographical conventions

This document uses the following typographical conventions:

- Document titles are marked in *italics*.
- Important terms that are used for the first time are marked in italics.
- Terms include a link to the term definition in the “Terms and definitions” clause, enabling easy navigation to the term definition.
- ABNF rules are in monospaced font.

ABNF usage conventions

Format definitions in this document are specified using ABNF (see RFC5234), with the following deviations:

- Literal strings are to be interpreted as case-sensitive Unicode characters, as opposed to the definition in RFC5234 that interprets literal strings as case-insensitive US-ASCII characters.
CXL™ Fabric Manager API over MCTP Binding Specification

1 Scope

The CXL™ Fabric Manager API over MCTP Binding Specification defines the bindings between CXL Fabric Manager API protocol elements and MCTP elements in order to transport Fabric Manager API Messages for CXL devices using MCTP. The specific Fabric Manager API message contents will be documented outside of DMTF directly by the CXL consortium.

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

- DMTF DSP0236, Management Component Transport Protocol (MCTP) Base Specification 1.3, defines the MCTP transport protocol over which the CXL Fabric Manager API over MCTP messages are to be conveyed.
- CXL Consortium, Compute Express Link™ (CXL™) Specification Revision 2.0, defines the CXL Fabric Manager API and message formats.

2 Normative references

The following referenced documents are indispensable for the application of this document. For dated or versioned references, only the edition cited (including any corrigenda or DMTF update versions) applies. For references without a date or version, the latest published edition of the referenced document (including any corrigenda or DMTF update versions) applies. Earlier versions may not provide sufficient support for this specification.

3 Terms and definitions

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

The terms "shall" ("required"), "shall not", "should" ("recommended"), "should not" ("not recommended"), "may", "need not" ("not required"), "can" and "cannot" in this document are to be interpreted as described in ISO/IEC Directives, Part 2, Clause 7. The terms in parentheses are alternatives for the preceding term, for use in exceptional cases when the preceding term cannot be used for linguistic reasons. Note that ISO/IEC Directives, Part 2, Clause 7 specifies additional alternatives. Occurrences of such additional alternatives shall be interpreted in their normal English meaning.

The terms "clause", "subclause", "paragraph", and "annex" in this document are to be interpreted as described in ISO/IEC Directives, Part 2, Clause 6.

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

Refer to DSP0236 for terms and definitions that are used across the MCTP specifications.

Refer to the CXL Specification for terms and definitions that are used in the Compute Express Link™ Fabric Manager API specification. For the purposes of this document, the following additional terms and definitions apply.

3.1 Compute Express Link™
A low-latency, high-bandwidth link that supports dynamic protocol muxing of coherent accesses, memory access, and IO protocols, thus enabling attachment of coherent accelerators or memory devices.

3.2 CXL™ Fabric Manager
The Fabric Manager controls aspects of a CXL system related to binding and management of pooled ports and devices.

3.3 CXL™ Fabric Manager API
Command set defined by the CXL consortium to manage devices in a CXL system.

3.4 Endpoint
An MCTP endpoint unless otherwise specified.

4 Symbols and abbreviated terms

Refer to DSP0236 for terms and definitions that are used across the MCTP specifications. Refer to the CXL Specification for terms and definitions that are used in the Compute Express Link™ Fabric Manager API specification. For the purposes of this document, the following additional symbols and abbreviated terms apply.

4.1 CXL™
Compute Express Link
5 Conventions

5.1 Reserved and unassigned values

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

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

5.2 Byte ordering

Unless otherwise specified, the byte ordering of multibyte numeric fields or multibyte bit fields in this specification shall be "Big Endian": The lowest byte offset holds the most significant byte and higher offsets hold lesser significant bytes.

6 Overview

Compute Express Link™ (CXL) is a dynamic multi-protocol technology designed to support accelerators and memory devices. CXL provides a rich set of protocols that include I/O semantics similar to PCIe (i.e., CXL.io), caching protocol semantics (i.e., CXL.cache), and memory access semantics (i.e., CXL.mem) over a discrete or on-package link.

CXL devices can be configured statically or dynamically via a Fabric Manager (FM), an external logical process that queries and configures the system’s operational state by using the Fabric Manager Application Programming Interface (FM API) commands. FM API commands are defined by the members of CXL Consortium in the Compute Express Link specification. Refer to www.computeexpresslink.org and the CXL specification for more information.

This specification only defines how FM API Commands are encapsulated in MCTP Messages and transferred between MCTP Endpoints over transports that have a corresponding MCTP transport binding specification. These are referred to in this document as FM API Messages over MCTP. The definitions and semantics of the FM API Commands themselves are outside the scope of this specification and are defined in the CXL specification.

The MCTP Transport Bindings that are used for CXL FM API over MCTP are defined in other companion specifications including but not limited to the MCTP SMBus/I2C Transport Binding Specification (DSP0237) and the MCTP PCIe VDM Transport Binding Specification (DSP0238).
7 Message Type-specific considerations

7.1 Message Type number

The Message Type number for CXL FM API over MCTP messages is defined in the MCTP IDs and Codes Specification (DSP0239) and the number assigned is 0x07.

7.2 CXL FM API over MCTP specification version information

Implementations that follow this specification shall return the following version information in the response to the GET MCTP Version Support command when the Message Type parameter in the request is set to 0x07 (return CXL FM API over MCTP specification version information).

The Version Number Entry 1 field shall be used to indicate compatibility with Version 1.0.0 of the CXL FM API over MCTP message type as:

1.0 [Major version 1, minor version 0, any update version, no alpha]

This is reported using the encoding as: 0xF1F0FF00

7.3 Timing specifications

CXL FM API messages over MCTP are made up of one or more MCTP packets. Each MCTP packet shall comply with the timing, arbitration, and fairness requirements of the transport binding specifications for the media through which it passes. For examples, refer to the MCTP SMBus/I2C Transport Binding Specification (DSP0237) and the MCTP PCIe VDM Transport Binding Specification (DSP0238) for specific packet and message timing requirements.

7.4 CXL FM API over MCTP message format

Referring to Figure 1, the CXL FM API Messages over MCTP are carried via the MCTP packet payload of one or more MCTP packets.

![Figure 1: MCTP Message fields](image)
7.4.1 Integrity Check (IC), Tag Owner (TO) and Message Tag (Msg Tag) usage

CXL FM API over MCTP messages do not include an overall message integrity check field and shall set the Integrity Check (IC) bit to 0b.

CXL FM API request and event notification messages when transported over MCTP shall have the TO bit set (TO bit = 1b). CXL FM API response messages over MCTP shall have the TO bit cleared (TO bit = 0b).

When request/response message exchange is used and the Tag Owner (TO) bit is set to 1b in the request, a responder shall return the same Message Tag with the Tag Owner bit cleared to 0b in the corresponding response message.

<table>
<thead>
<tr>
<th>Field Name</th>
<th>Field Size</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>TO</td>
<td>1 bit</td>
<td>1b - CXL FM API Request and Event Notification messages</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0b – CXL FM API Response messages</td>
</tr>
<tr>
<td>IC</td>
<td>1 bit</td>
<td>Message Integrity Check bit = 0b</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CXL FM API over MCTP messages do not include an overall Message Integrity check field.</td>
</tr>
<tr>
<td>Message type</td>
<td>7 bits</td>
<td>CXL FM API = 0x07</td>
</tr>
<tr>
<td></td>
<td></td>
<td>This field identifies the MCTP message as carrying a CXL FM API message.</td>
</tr>
<tr>
<td>CXL FM API Message Body</td>
<td>Variable</td>
<td>The CXL FM API message fields are defined in the CXL Specification</td>
</tr>
</tbody>
</table>

Table 1: CXL FM API over MCTP Message field descriptions

For a definition of CXL FM API request, response and event notification messages, refer to the CXL Specification.

7.4.2 Message assembly

CXL FM API messages over MCTP may be split into one or more MCTP packets thus requiring segmentation and assembly. All multi-packet CXL FM API over MCTP messages shall comply with the message packetization and assembly rules of the MCTP base specification (DSP236). Specifically, clauses in the MCTP base specification related to Message assembly, Dropped packets, Starting message assembly, Terminating message assembly/dropped messages, and Dropped messages shall be complied with strictly. CXL FM API messages when transported over MCTP shall not require any changes to the MCTP base specification.

7.5 Maximum message size

The CXL FM API message body over MCTP shall be less than or equal to 1088 bytes. All MCTP endpoint shall support this maximum message body size of 1088 bytes, which includes a maximum of 1024 bytes for the CXL FM API message payload and a maximum of 64 bytes for the CXL FM API header. This corresponds to a transfer of 17 MCTP packets using a baseline transmission unit of 64 bytes for the MCTP packet payload. See the CXL Specification for a definition of the CXL FM API message payload and headers.

The maximum message size includes the IC bit and Message Type fields plus any additional Message Type-specific header fields, as required by the CXL FM API. Refer to the CXL Specification for any additional restrictions on message sizes.
7.6 Multiple MCTP physical transports

In order to facilitate identification of devices that are accessible via multiple physical transports, the endpoints in the device shall support the Get Endpoint UUID MCTP command.

An MCTP endpoint is not required to support more than one outstanding command over a single physical transport. A requestor shall not have multiple requests outstanding simultaneously across multiple physical transports to an endpoint. Otherwise, this specification does not define any additional behaviors related to communicating with CXL™ devices over MCTP that may be accessed through more than one type of MCTP physical transport on a given MCTP network.
ANNEX A
(informative)

Change log

<table>
<thead>
<tr>
<th>Version</th>
<th>Date</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0.0</td>
<td>2021-05-25</td>
<td></td>
</tr>
</tbody>
</table>