## 背景介绍

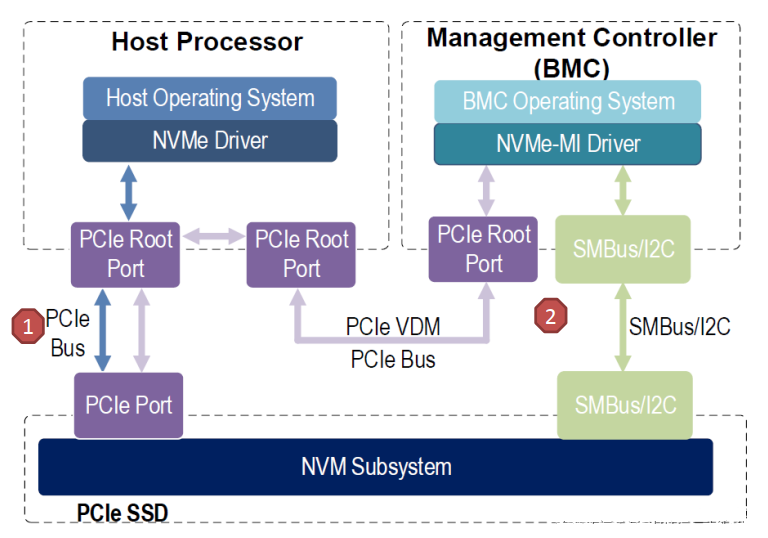
### 1.1 NVMe-MI

NVME-MI（NVMe Management Interface）规范定义了用于NVMe存储设备带外和带内管理的架构和命令集，以及用于监控和控制NVMe Enclosure元件的架构和机制。该规范建立在NVM Express规范的基础上。

### 1.2 带内管理和带外管理特性

带内管理（in-band）使用由操作系统控制的硬件资源和组件进行操作的管理，带外管理（out-of-band）是使用独立于操作系统控制的资源。

图片中通道1：NVMe驱动程序根据NVMe规范通过PCIe与NVMe控制器通信，通道2：两种OOB通信方式，PCIe VDM和SMBus/I2C。



1. 带内管理

* 支持多种操作系统（Windows、Linux、VMWare 等）
* 每种操作系统有几种不同发行版
* 为每个操作系统开发、维护、验证管理应用程序的资源成本过高
* 随着时间的推移，会发布新版本的操作系统和 NVMe 驱动程序
* 持续消耗 CPU cycle
* 安全隐患
* 管理功能因操作系统而异

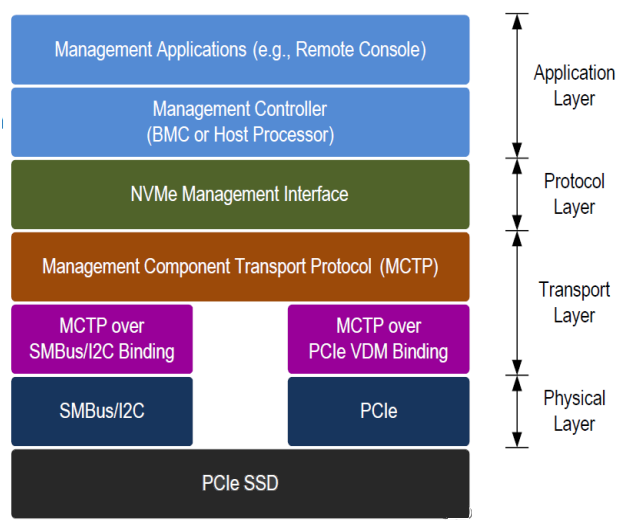
1. 带外管理

* 在一个操作环境中开发管理应用程序，在任何主机操作系统上都一样工作
* 可在没有操作系统的情况下工作(pre-boot, deployment)
* 不需要消耗主机 CPU cycle

### 1.3 带外管理

NVMe-MI规范1.2版本定义了一种通过MCTP（Management Component Transport Protocol）传输协议管理NVMe存储设备和NVMe机箱的带外机制。带外管理路径通过SMBus/I2C上的MCTP或PCIe VDM上的MCTP从管理控制器到管理终端。与NVMe协议不同，NVMe-MI协议是通过MCTP协议进行传输的，同时底层物理层支持PCIe或者SMBus/I2C。带外管理的好处就是和NVMe协议分离，即使SSD在NVMe看来是故障的，还能有另外一条路径查看SSD状态。

带外管理的协议分层如下图所示，包括应用层(application layer)、协议层(protocol layer)、信息传输层(transport layer)、物理层(physical layer)：



* 应用层(application layer)包括NVMe子系统管理程序，以及代表管理程序跟子系统沟通的管理控制器。
* 协议层(protocol layer)包括对应每个Management Controller的NVMe-MI接口，该接口和NVMe子系统的Management Endpoint相协调来执行管理任务；
* 传输层(transport layer)包含MCTP协议，以及作为与实体层之间接口的MCTP绑定(binding)，并支持PCIe与SMBus/I2C绑定；
* 物理层(physical layer)则是信息传输层接收资料所需的输出输入口，每1个接口对应1个管理端点，SMBus/I2C接口与PCIe接口的数量分别可配置0或多个。PCIe与SMBus/I2C接口对应的管理端点都支持相同的NVMe-MI指令且提供相同的功能，不过PCIe接口支持的资料传输速度远高于SMBus/I2C接口。

### 1.4 MI命令交互过程

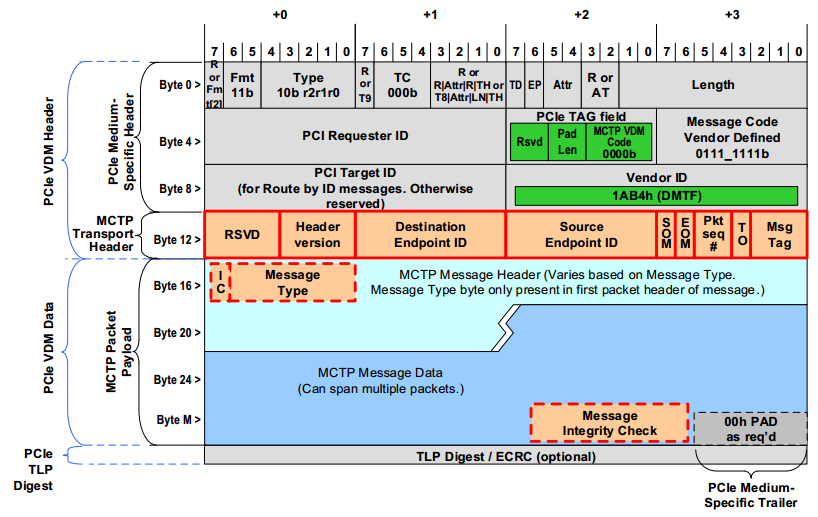
1. host将发送消息按照MI的协议规范组成MI报文；
2. 将MI报文当作MCTP的数据封装为MCTP报文；
3. 完整的MCTP报文通过I2C/PCIe链路发给SSD；
4. SSD收到MCTP报文后，解封MCTP报文得到MI报文；
5. 解析MI报文后获取发送消息，然后将要回复的数据按照MI协议规范组成MI报文；
6. 再将回复MI报文当作MCTP的数据封装为MCTP报文；
7. 再将MCTP报文通过I2C/PCIe链路发送给host；
8. host解封MCTP报文，再解析MI报文获取回复消息

## MCTP over PCIe VDM包格式

支持MCTP over PCIe VDM的PCIe设备需要至少在一个PCIe物理功能（PF）上支持MCTP over PCIe VDM通信。如果兼容MCTP over PCIe VDM的PCIe设备在多个PCIe功能上支持MCTP over PCIe VDM通信，则每个功能上的MCTP over PCIe VDM通信应独立于其他PCIe功能上的MCTP over PCIe VDM通信。

MCTP over PCIe VDM传输绑定通过带数据的PCIe Type 1 VDM消息报文传输MCTP消息。MCTP消息使用MCTP VDM Code(0000b)，该值唯一地将MCTP消息与其他DMTF VDM区分开来。

MCTP over PCIe VDM报文通过Type对应的路由方式、MCTP VDM Code、Message Code和Target ID这四个字段来区分其他message报文。



标记为“PCIe Medium-Specific Header”和“PCIe Medium-Specific Trailer”的字段是针对使用PCIe vdm承载MCTP报文的。标记为“MCTP Transport Header”和“MCTP Packet Payload”的字段是所有MCTP数据包和消息的通用字段，并在MCTP基础协议中定义。本文档定义了这些字段在PCIe VDM中携带时的位置。PCIe规范允许PCIe VDM头的最后四个字节由供应商定义。MCTP over PCIe VDM传输绑定规范使用这些字节作为DMTF供应商ID下的MCTP传输头字段。本文档还指定了MCTP“Hdr Version”字段的特定的媒介（PCIe协议）的用法。

表1列出了MCTP over PCIe VDM通信中应使用与PCIe相关的字段和字段值。当不指定时，请根据PCIe规范设置。请注意，通过PCIe VDM数据包的MCTP中TLP前缀的存在取决于实现，并且超出了本规范的范围。

表1 特定于PCIe协议的MCTP包字段

(参考MCTP PCIe VDM Transport Binding Specification)

|  |  |  |
| --- | --- | --- |
| Field Name | Size | Description |
| R 或者  Fmt[2] | 1 bit | PCIe 1.1/2.0: PCIe reserved bit (1 bit).  PCIe 2.1, 3.X, 4.X, 5.X: Fmt[2]. Set to 0b. |
| Fmt | 2 bit | Format (2 bits). 设置11b 来表示4 DW的带有数据的报头. |
| Type | 5 bit | Type and Routing (5 bits).[4:3] 设置10b表明为消息报文  [2:0] PCIe消息路由(r2r1r0)  000b : Route to Root Complex  010b : Route by ID  011b : Broadcast from Root Complex  MCTP不支持其他路由域段 |
| R 或者T9 | 1 bit | PCIe 1.1/2.0/2.1/3.X: PCIe预留位（1bit），参考PCIe协议。设置为0b。  PCIe 4.X/5.X: T9 (1bit). 参考PCIe Gen 4。设置为0b。 |
| TC | 3 bit | Traffic Class (3 bits). 对于MCTP over PCIe VDM设置为000b |
| R  或者  R | Attr | R | TH  或者  T8 | Attr | LN | TH | 4 bit | PCIe 1.1/2.0: PCIe预留位 (4 bits)。Set to 0000b。  PCIe 2.1/3.X: PCIe预留位 (1 bit), Set to 0b。  Attr[2] (1 bit)，Set to 0b。  预留位 (1bit), Set to 0b。  TH (1bit), Set to 0b。  PCIe 4.X/5.X: T8 bit (1 bit) – Set to 0b,  Attr[2] (1 bit) – Set to 0b,  LN (1bit) – Set to 0b,  TH (1bit) – Set to 0b. |
| TD | 1 bit | TLP Digest (1 bit)。  1b indicates the presence of the TLP Digest field at the end of the PCIe TLP (transaction layer packet). The TD bit should be set in accordance with the devices overall support for the TLP Digest capability, and whether that capability is enabled. See description of the TLP Digest / ECRC field, below, for additional information. Note that earlier versions of this specification erroneously required this bit to be set to 0b, which would have required devices to not support the TLP Digest capability.  1b表示在PCIe TLP(事务层报文)的末尾存在TLP Digest字段。应根据设备对TLP Digest能力的总体支持情况和是否使能TLP Digest能力来设置TD位。有关更多信息，请参阅下面TLP Digest / ECRC字段的描述。请注意，该规范的早期版本错误地要求将该位设置为0b，这将要求设备不支持TLP Digest功能。 |
| EP | 1 bit | 错误中毒(1 bit). |
| Attr[1:0] | 2 bit | Attributes (2 bits). 所有的MCTP over PCIe VDM包设置为00b or 01b |
| R或者AT | 2 bit | PCIe 1.1: PCIe预留位 (2 bits).  PCIe 2.0/2.1/3.X/4.X/5.X: 地址类型域（Address Type）. Set to 00b. |
| Length | 10 bit | Length: Length of the PCIe VDM Data in dwords. Implementations shall support the baseline transmission unit defined in the MCTP Base Specification. For example, supporting a baseline transmission unit of 64 bytes requires supporting PCIe VDM data up to 16 dwords. An implementation may optionally support larger transfer unit sizes.  长度:PCIe VDM的长度，单位为DW。实现应支持MCTP基本规范中定义的baseline传输单元。例如，如果支持64字节的baseline传输单元，则需要支持PCIe VDM数据，最长可达16DW。实现可以选择支持更大的传输单元大小。 |
| PCI Requester ID | 16 bit | Bus/device/function or bus/function number of the managed endpoint sending the message.  发送消息的且被管理的EP的Bus/device/function或者bus/function |
| Pad Len | 2 bit | Pad Length (2-bits). 1-based count (0 to 3) of the number of 0x00 pad bytes that have been added to the end of the packet to make the packet dword aligned with respect to PCIe . Because only packets with the EOM bit set to 1b are allowed to be less than the transfer unit size, packets that have the EOM bit set to 0b will already be dword aligned and will thus not require any pad bytes and will have a pad length of 00b.  Pad长度(2位)。基于1的计数(0 ~ 3)，在报文的末尾添加0x00 pad字节，使报文的DW相对于PCIe保持一致。因为只有EOM位设置为1b的数据包才允许小于传输单元大小，所以EOM位设置为0b的数据包将已经是双字对齐的，因此不需要任何填充字节，并且填充长度为00b。 |
| MCTP VDM Code | 4 bit | Value that uniquely differentiates MCTP messages from other DMTF VDMs. Set to 0000b for this transport mapping as defined in this specification.  该值唯一地将MCTP消息与其他DMTF VDM区分开来。为此规范中定义的传输映射设置为0000b。 |
| Message Code | 8 bit | (8 bits). 设置为0111\_1111b来表示是一个Type 1 VDM. |
| PCI Target ID | 16 bit | (16 bits). For Route By ID messages, this is the bus/device/function number or bus/function number that is the physical address of the target endpoint. This field is ignored for Broadcast and for Route to Root Complex messages.  对于Route By ID消息，这是总线/设备/功能号或总线/功能号，即目标端点的物理地址。对于广播和路由到RC消息，该字段被忽略 |
| Vendor ID | 16 bit | (16 bits). Set to 6836 (0x1AB4) for DMTF VDMs. The most significant byte is in byte 10, the least significant byte is byte 11.  DMTF VDM设置为6836 (0x1AB4)。最高位字节在第10位，最低位字节在第11位。 |
| RSVD | 4bit | MCTP reserved (4 bits). Set these bits to 0 when generating a message. Ignore them on incoming messages。  MCTP保留(4位)。在生成消息时将这些位设置为0。在收到的消息中忽略它们 |
| Hdr Version | 4bit | MCTP version (4 bits)  0001b : For MCTP devices that conform to the MCTP Base Specification and this version of the PCIe VDM transport binding.  All other settings: Reserved to support future packet header field expansion or header version.  MCTP版本(4位)0001b:  适用于符合MCTP Base Speciation和此版本的PCIe VDM传输绑定的MCTP设备。  所有其他设置:保留以支持未来的报头字段扩展或报头版本。 |
| 00h PAD | 见描述 | Pad bytes. 0 to 3 bytes of 00h as required to fill out the overall PCIe VDM data to be an integral number of dwords. Because only packets with the EOM bit set to 1b are allowed to be less than the transfer unit size, packets that have the EOM bit set to 0bwill already be dword aligned, and will thus not require any pad bytes and will have a pad length of 00b.  填充字节。0 ~ 3个字节的00h被要求填写使整个PCIe VDM数据为整数个dwords。因为只有EOM位设置为1b的数据包才允许小于传输单元大小，所以EOM位设置为0b的数据包将已经是双字对齐的，因此不需要任何填充字节，并且填充长度为00b。 |
| TLP Digest /  ECRC | 32bit | TLP Digest / ECRC (End-to-end CRC). This field is defined for all PCIe TLPs (Transaction Layer Packets). Device support for this field is optional. However, per PCIe v2.1/3.X/4.X/5.X: "If a device Function is enabled to generate ECRC, it must calculate and apply ECRC for all TLPs originated by the Function. If the device supports generating this field, it must support it for all TLPs." Additionally, per PCIe v2.1/3.X/4.X/5.X, if the ultimate PCI Express Receiver of the TLP does not support ECRC checking, the receiver must ignore the TLP Digest.  TLP摘要/ ECRC(端到端CRC)。该字段是为所有PCIe TLP (Transaction Layer packet)定义的。设备对该字段的支持是可选的。但是根据PCIe v2.1/3.X/4.X/5.X:“如果一个设备的Function使能了生成ECRC功能，它必须计算并应用该功能生成的所有tlp。”如果设备支持生成该字段，则必须支持所有tlp。”此外，根据PCIe v2.1/3.X/4.X/5.X，如果TLP的最终PCI Express Receiver不支持ECRC检查，则该Receiver必须忽略TLP Digest。 |

## MCTP控制消息的物理地址格式

|  |  |  |  |
| --- | --- | --- | --- |
| Format Size | Address Type | Layout and Description | |
| 2 bytes  (BDF ID) | Bus Device Function (BDF) | byte 1 | [7:0] – Bus number |
| byte 2 | [7:3] – Device number  [2:0] – Function number |
| 2 Bytes  (ARI ID) | Alternate Routing Identifier (ARI) | byte 1 | [7:0] – Bus number |
| byte 2 | [7:0] – Function number |

## 消息路由

PCIe总线内的数据包路由使用PCIe规范定义的路由。PCIe物理路由/桥接与MCTP桥接不同。PCIe物理路由/桥接通常对MCTP是透明的。没有mctp定义的用于配置或控制PCIe总线设置的功能。

### 4.1 Route by ID

All MCTP over PCIe VDM packets between endpoints that are not the bus owner shall use Route by ID for message routing.

The MCTP bus owner shall use Route by ID for messages to individual MCTP endpoints.

MCTP endpoints are required to capture the PCIe requester ID and the MCTP source EID when receiving an EID assignment MCTP control request message. This is because this request can only be issued by the MCTP bus owner

非总线所有者端点之间的所有MCTP over PCIe VDM数据包都应使用Route by ID进行消息路由。

MCTP总线所有者发送到各7个MCTP端点的消息应使用Route by ID。

当接收到EID分配MCTP控制请求消息时，需要MCTP端点捕获PCIe请求者ID和MCTP源EID。这是因为该请求只能由MCTP总线所有者发出。

### 4.2 Route to root complex

MCTP endpoints shall use this routing for the Discovery Notify request message to the MCTP bus owner as part of the MCTP over PCIe VDM discovery process.

The MCTP endpoints shall use this routing for responding to the MCTP control request messages that were sent using Broadcast from Root Complex.

Communication of MCTP PCIe VDM packets that are destined to MCTP bus owner using routed to root complex is implementation specific and is outside the scope of this specification.

MCTP端点应使用此路由将发现通知请求消息（Discovery Notify request message）发送给MCTP总线所有者，作为通过PCIe VDM发现过程的一部分。

MCTP端点应使用此路由来响应使用RC广播发送的MCTP控制请求消息。

使用路由到根复合体的MCTP PCIe VDM数据包的通信是特定于实现的，不在本规范的范围之内。

### 4.3 Broadcast from root complex

The MCTP bus owner should use this routing for the Prepare for Endpoint Discovery and Endpoint Discovery messages as part of the MCTP over PCIe VDM discovery process.

MCTP总线所有者应该将此路由用于准备端点发现和端点发现消息，作为PCIe VDM上MCTP发现过程的一部分。

### 4.4 第一代225的路由兼容性

第一代225的message类报文路由转发通过decoder模块完成，decoder模块分为decoder\_up和decoder\_dn两个模块。

在decoder\_up模块中有参数FMT\_TYPE\_MESD\_ID和FMT\_TYPE\_MESD\_ BROADCAST两种，分别对应message报文的ID路由和来自RC的广播路由两种路由方式。当报文中解析出的FMT和TYPE是对应的两个参数时，便将TLP报文按照协议规范的路由方式进行转发。

在decoder\_dn模块中有参数FMT\_TYPE\_MESD\_ID和FMT\_TYPE\_MESD\_GRRC两种，分别对应message报文的ID路由和路由到RC两种路由方式。当报文中解析当报文中解析出的FMT和TYPE是对应的两个参数时，便将TLP报文按照协议规范的路由方式进行转发。

而MCTP over PCIe VDM报文有三种路由方式，分别为ID路由、路由到RC以及来自RC的广播路由，因此第一代225支持MCTP over PCIe VDM报文的三种路由方式。

## 在PCIe总线上路由对等事务

Because the PCIe specification does not require peer-to-peer routing support in PCIe root complexes, MCTP over PCIe VDM messages are not required to be routed to peer devices directly. When peer-to-peer routing is not supported by a PCIe root complex, all MCTP over PCIe VDM messages between two MCTP endpoints shall be routed to or through the MCTP bus owner as an MCTP bridge. If the PCIe root complex, as the MCTP bus owner, supports peer-to-peer routing, it shall use direct physical addressing to support routing between two MCTP endpoints on the PCIe bus.

由于PCIe规范不需要RC支持p2p路由，因此不需要将MCTP over PCIe VDM消息直接路由到对等设备。当RC不支持p2p路由时，两个MCTP端点之间的所有MCTP over PCIe VDM消息应作为MCTP桥路由到MCTP总线所有者或通过MCTP总线拥有者。如果PCIe根复合体作为MCTP总线所有者支持对等路由，则应使用直接物理寻址来支持PCIe总线上两个MCTP端点之间的路由。

## 在PCIe和其他总线之间路由消息

All MCTP messages that span between PCIe and other buses shall be sent through the MCTP bus owner. The MCTP bus owner has the destination EID routing tables necessary to route messages between the two bus segments. If an endpoint is aware of multiple routes to a destination over multiple bus types, a higher level algorithm/protocol above MCTP shall be used to determine which bus/route to use. Typically this decision can be based on things like power state and MCTP discovery state.

所有跨越PCIe和其他总线的MCTP消息都应通过MCTP总线所有者发送。MCTP总线所有者拥有在两个总线段之间路由消息所必需的目的地EID路由表。如果一个端点知道通过多种总线类型到达目的地的多条路由，则在MCTP应使用之上的更高级别的算法/协议，以确定使用哪条总线/路由。通常这个决策可以基于诸如电源状态和MCTP发现状态之类的东西。

## MCTP格式

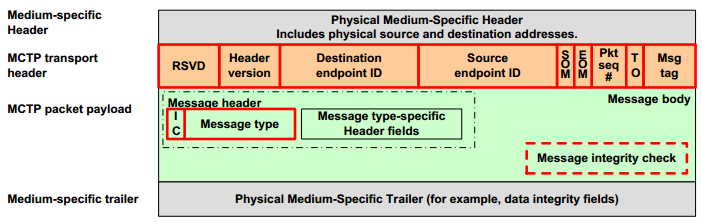


表2 MCTP基础协议公共字段

(参考MCTP Base Specification)

|  |  |  |
| --- | --- | --- |
| Field Name | Size | Description |
| MCTP Transport  Header | 32bit | 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.  MCTP传输报头是每个MCTP数据包的一部分，它提供了数据包的版本和地址信息，以及标志和“消息标签”字段，消息标签与源EID一起用于识别构成MCTP消息的数据包。MCTP传输报头字段是通用字段，无论使用MCTP的物理介质是什么，这些字段都始终存在。  注意:MCTP传输报头的子字段的位置可能根据物理介质绑定而变化。 |
| RSVD | 4bit | (预留) 未来的MCTP基本规范的定义所预留。 |
| Hdr version | 4bit | (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.  (报头版本)标识用于在给定物理介质上传输消息中MCTP公共字段的格式、物理帧和数据完整性机制。该值在特定介质的规范中定义。  注意:该字段的值在不同的传输绑定之间可能不同。 |
| Destination  endpoint ID | 8bit | 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.  终端接收MCTP报文的EID。  对特定的路由来说，一些EID的值是预留位。  参见表2 –Special endpoint ID。 |
| Source endpoint ID | 8bit | The EID of the originator of the MCTP packet. See Table 2 – Special endpoint IDs.  MCTP报文发起者的EID。参见表2 –Special endpoint ID。 |
| SOM | 1bit | (Start of Message)Set to 1b if this packet is the first packet of a message.  如果此数据包是消息的第一个数据包，则设置为1b。 |
| EOM | 1bit | (End Of Message) Set to 1b if this packet is the last packet of a message.  如果此报文是消息的最后一个报文，则设置为1b。 |
| Pkt Seq # | 2bit | (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 shall increment modulo 4 for each subsequent packet belonging to a given message up through the packet containing the EOM flag.  对于跨越多个包的消息，数据包序列号在每个连续的数据包上以模4递增。  这允许接收方在消息的开始和结束之间检测最多三个连续的丢失数据包。虽然数据包序列号可以是任意值(0-3)，但如果设置了SOM位，建议它是与设置了EOM位的前一个数据包的增量模4。在SOM包之后，包序列号应该对每个属于给定消息的后续包以模4递增，直到包含EOM标志的包 |
| TO | 1bit | 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. Set to 1b to indicate that the source of the message originated the message tag.  TO (Tag Owner)位用于标识message tag是由作为消息源头的端点还是由作为消息目的地的端点发起的。对于Tag Owner位的每个值，Message Tag字段是独立生成和跟踪的。MCTP消息类型可以用附加的含义覆盖这个位，例如用它来区分“请求”消息和“响应”消息。设置为1b表示消息的来源源自消息标记。 |
| Msg tag | 3bit | (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.  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.  When request/response message exchange is used and the Tag Owner (TO) bit is set to 1 in the request, a responder should return the same Message Tag with the Message Tag Owner bit cleared to 0 in the corresponding response Message.  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.  该字段与Source endpoint id和标签所有者(TO)字段一起用于标识MCTP传输级别的唯一消息。其他元素(例如MCTP Message Data字段的部分)是否也用于唯一标识实例或跟踪消息的重试取决于Message type。  如果每个消息都具有唯一的message tag，则允许Source endpoint同时将来自多个消息的数据包穿插到相同的目标端点。  当使用请求/响应消息交换并且在请求中设置Tag Owner (TO)位为1时，响应方应返回相同的message tag，但在相应的响应消息中将Tag Owner位清除为0。  对于拆分为多个数据包的消息，从SOM到EOM的所有数据包的Tag Owner (TO)和Message Tag位保持相同。 |
| Message body | 见描述 | The message body represents the payload of an MCTP message. The message body can span multiple MCTP packets.  消息体表示MCTP消息的有效负载。消息体可以跨越多个MCTP包。 |
| IC | 1bit | (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.  0b = No MCTP message integrity check  1b = MCTP message integrity check is present  (MCTP完整性校验位)表示该MCTP消息是否被整体MCTP消息负载完整性校验所覆盖。该字段以及message type位必须是消息的第一个数据包中消息体的第一个字节，并且该字段为消息体中第一个字节的最高有效位。  0b =没有MCTP消息完整性检查。  1b =有MCTP消息完整性检查 |
| Message type | 7bit | 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 overwhich 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. See DSP0239 and Table 3 for information on message type values.  定义包含在MCTP消息的消息数据部分中的有效负载类型。该字段必须包含在消息的第一个数据包的消息体的第一个字节的最低有效位中。  与MCTP传输报头中的字段一样，消息类型字段是与使用MCTP的传输无关的常见MCTP字段之一。然而，与MCTP传输报头不同的是，消息类型字段只需要出现在特定MCTP消息的第一个数据包中，而MCTP传输报头字段出现在每个MCTP数据包中。有关消息类型值的信息，请参见DSP0239和表3。 |
| 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.  与特定message type相关联的附加头信息(如果有的话)。这通常只包含在消息的第一个数据包中，但是给定的消息类型定义可以根据任何数据包的需要来定义报头字段 |
| Message data | 0 to N bytes | Data associated with the particular message type. Defined according to the specifications for the message type.  与特定message type相关联的数据。根据message type的规范定义。 |
| MCTP packet  payload | 见描述 | 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, Packet payload and transmission unit sizes, for more information.  数据包有效负载是在给定的MCTP数据包中携带的消息体的一部分。根据控制数据包有效载荷和传输单元大小的规则，数据包有效载荷受到限制。有关更多信息请参阅8.3，数据包有效载荷和传输单元大小。 |
| Msg integrity  chec | 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 shall 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 also 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. Use of the Msg integrity check field is specific to the particular message type specification.  (MCTP消息完整性检查) 此字段表示在消息体的内容上是否存在特定于消息类型的完整性检查。如果存在，消息完整性检查字段应在消息体的最后字节中携带。特定的消息类型定义将指定这是必需的、可选的还是不使用的、字段大小以及用于生成字段的算法。MCTP基本协议也没有指定该字段是在单个分组消息中需要(可能取决于传输单元大小)还是仅在多个分组消息中需要。Msg完整性检查字段的使用特定于特定的消息类型规范。 |

## MCTP包有效载荷和传输单元大小

对于MCTP，传输单元的大小定义为MCTP包中携带的数据包有效载荷的大小。以下是baseline传输单元的关键信息要点：

* MCTP的基线传输单元（最小传输单元）大小为64字节。
* 支持MCTP control message的消息终端应始终接受传输单元等于或小于基线传输单元的有效数据包。消息终端还允许支持更大的传输单元。
* 除了最后一个包(EOM位= 1b的包)的传输单元外，同一报文中所有包的传输单元大小必须相同。除了最后一个包之外，这个大小至少应该是基线传输单元大小。
* 最后一个数据包中传输单元的大小应小于或等于用于其他数据包(如果有的话)的传输单元大小。
* 如果协商的传输单元大小大于基线传输单元大小，则所有报文的传输单元应小于或等于协商的传输单元大小。(端点之间较大传输单元的协商机制是特定于消息类型的，在本规范（MCTP Base Speciation）中没有涉及。)
* 此条款的要求，给定端点可以协商对数据包大小的额外限制，以便与另一个端点进行通信。
* 所有消息类型都应支持使用传输单元不大于基线传输单元的数据包进行传递。这是在实现中支持桥接那些消息所必需的，其中存在仅支持基线传输单元的MCTP桥接。