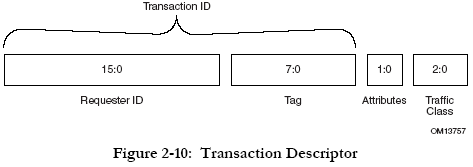
### 2.2.6 事务描述符

#### 2.2.6.1 概述

事务描述符是请求方和完成方之间传输事务信息的一种机制。事务描述符由三部分组成：

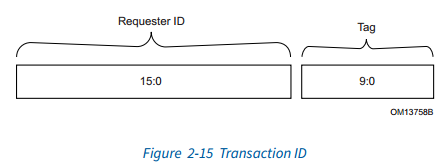


事务ID用于标识事务，属性域说明事务的特性，通信等级确定事务所需的服务。

事务ID的各部分在TLP包头中并不连续。

#### 2.2.6.2 事务描述符——事务ID

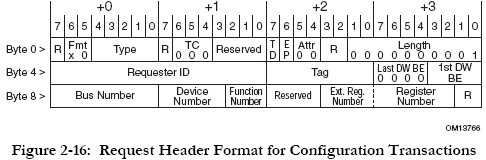
事务ID包括两部分：请求方ID和Tag，如图2-15所示。



在[PCIe-4.0]中引入的10bit Tag功能将总Tag字段大小从8位增加到10位。两个额外的Tag位，Tag[8] (T8)和Tag[9] (T9)，与TLP报头中的其他Tag[7:0]位不相邻。这两个额外的位在本规范的先前版本中是保留的。

* Tag[9:0]是由每个请求者生成的一个10位字段，对于需要Completion的请求者，其所有未完成的请求都必须是唯一的。不支持10bit Tag Requester Capability的请求者必须将Tag[9:8]设置为00b。
  + 支持16.0 GT/s或更高数据速率的Function(包括Switch中的Function)必须支持10bit Tag Completer capability。如果一个Function支持10bit Tag Completer capability，它可以选择支持10bit Tag Requester capability。参见本节后面的7.5.3.15节和“实现10bit Tag能力的注意事项”的实现说明。
  + 声明支持10bit Tag Completer capability的RCs必须通过所有支持该能力并且作为PCIe请求者的目标的寄存器和内存区域正确处理10bit Tag请求；例如，DMA请求或RCiEP中的MMIO区域所针对的主机内存。
    - 每个表示支持的RP必须处理由其入口端口接收的此类请求。
    - 每个表示支持的RCiEP都必须处理来自支持的内部路径的此类请求，包括通过RP的请求。
  + 如果RC包含表示支持10bit Tag Requester capability的RCiEP，RC必须正确处理来自这些RCiEP的所有寄存器和内存区域作为这些RCiEP的目标所支持的10bit Tag Requests；例如，DMA请求或RCiEP中的MMIO区域所针对的主机内存。
  + Receivers/Completers必须正确处理8bit的Tag值，而不管它们的Extended Tag Field Enable位的设置(参见章节7.5.3.4)。有关Extended Tag的桥处理的详细信息，请参阅PCI Express到PCI/PCI- X桥规范。
  + 支持10bit Tag Completer capability的Receivers/Completers必须正确处理10bit Tag值，无论其10bit Tag Requester Enable位设置如何。参见7.5.3.16节。
  + 10bit Tag capability不是为PCI Express到PCI/PCI- X桥而设计的，它们不能指示10bit Tag Requester capability或10bit Tag Completer capability。
  + 如果10bit Tag Requester Enable位为Clear，Extended Tag Field Enable位为Clear，则每个功能的最大未处理请求数应限制为32，并且仅使用Tag字段的低5位，其余的高5位要求为0 0000b。
  + 如果10bit Tag Requester Enable位为Clear，并且设置了Extended Tag Field Enable位，则最大值增加到256，并且只使用Tag字段的低8位，其余的高2位需要为00b。
  + 如果设置了10bit Tag Requester Enable位，则单个Completer的最大目标增加到768。Requester被允许在向它认为合适的Completer发送10bit Tag Requests时使用Tag字段的所有10位，尽管Requester仍然被允许向其他Completer发送较小的Tag请求。以下内容适用于设置了“10bit Tag Requester Enable”位的10bit Tag capable Requesters。
    - 如果一个Endpoint支持向其他Endpoint发送请求(而不是host memory)，除非特定实现的机制确定 Endpoint 支持 10-Bit Tag Completer capability，Endpoint 不得向另一个给定 Endpoint 发送 10-Bit Tag Requests。对于某些实现，根本不向其他端点发送10bit Tag Requests可能是可以接受的。更复杂的机制超出了本规范的范围。
    - 如果PIO请求者具有10bit Tag Requester capability,，那么请求者如何确定何时使用10位Tag与更小的Tag不在本规范的范围之内。
    - 对于10位Tag，有效的Tag[9:8]值为01b、10b或11b。Tag[9:8]等于00b的10位Tag值无效，请求者不得生成。这使请求者能够确定其接收到的应具有10位Tag的completion是否包含无效标记，这通常是由于完成者不支持10bit Tag Completer capability造成的。
    - 如果请求者向缺乏（不支持）10bit Tag Completer capability的完成者发送10bit Tag Request，则返回Completion的Tag中Tag[9:8]等于00b。由于请求者被禁止为10bit Tag生成这些Tag值，因此此类Completion将被处理为意外完成，默认情况下为咨询性非致命错误（Advisory Non-Fatal Errors）。请求者必须遵循标准PCI Express错误处理要求。
    - 当请求者将带有无效10bit Tag的Completion处理为意外完成（Unexpected Completion）时，原始请求可能会引发完成超时（Completion timeout）。如果请求者为了避免数据损坏而以某种特定于设备的方式处理完成超时条件，则允许请求者根据需要，通过标准PCI Express错误处理机制抑制（suppress）完成超时的处理。
    - 如果请求者支持向某些完成者发送10bit Tag Requests，并同时支持向其他完成者发送较小的Tag Requests，则请求者必须遵守较小Tag Requests的Extended Tag Field Enable位设置。也就是说，如果该比特是Clear，则只有Tag字段的低5比特可以是非零的；如果该位是Set，则只有Tag字段的较低8位可以是非零的。
    - 如果请求者支持向某些完成者发送10bit Tag Requests，并同时支持向其他完成者发送较小的Tag Requests，则请求者必须确保，如果任何10bit Tag Requests由缺乏（不支持）10bit Tag Completer capability的完成者完成，则没有未完成的10bit Tag可以别名为未完成的较小Tag。参见本节后面的“同时使用10bit Tag和更小的Tag”实现说明。
  + Extended Tag Field Enable位的缺省值与具体实现有关。10bit Tag Requester Enable位的默认值为0b。
  + 如果发送的多个未完成的请求具有非唯一的Tag值，那么Receiver/Completer的行为在规范中是未定义的。
  + 若使用幻象功能号（Phantom Function Numbers）来扩展最大请求数，请求方必须保证在其生成的所有未完成且需要返回完成包的请求中，每一个请求包的幻象功能号和Tag域的组合具有唯一性。
* 对于不需要返回完成包的请求（Posted Requests），Tag[7:0]的值没有定义，可以是任意值。
* 对于设置TH位的Posted请求，Tag[7:0]字段被重新用于ST[7:0]字段(详细信息请参阅章节2.2.7.1)。对于TH位为Clear的Posted请求，Tag[7:0]字段是未定义的，可以包含任何值。(对于某些Vendor\_Defined消息，此规则的例外情况请参见表F-1。)
  + 对于TH位为Clear的Posted请求，Tag[7:0]字段中的值不能影响接收方对请求的处理。
  + 对于设置TH位的Posted请求，ST[7:0]字段中的值可能会影响请求的完成处理(详见2.2.7.1)。
* 请求方ID和Tag为系统中的每个事务组成了一个全局性的标识符，即事务ID。
* 请求包和完成包中都包含2 事务ID。
* 请求方ID是一个16位的值，并且系统中的每个PCI Express功能模块都有一个唯一的请求ID。
* 功能必须捕获由功能完成的所有类型0配置写入请求提供的总线和设备编号，并在设备/功能发起的所有请求的请求者ID的总线和装置编号字段中提供这些编号。建议仅为成功完成的请求捕获数字。
  + 例外：Root Complex上逻辑设备的总线号和设备号的设置方式以及Switch下游端口的设备号设置方式由具体实现确定。
  + 注意，设备的总线号和设备号可能在工作时发生变化，所以在收到每个配置写请求时，都有必要重新获取总线号和设备号。
  + 建议对未实现功能的配置写请求不影响捕获的总线号和设备号。

当Switch代表自己生成请求（例如，用于错误报告）时，Switch必须使用与导致生成请求的端口（见第7.1节）逻辑关联的网桥主侧相关的请求者ID。



* 在对设备进行初始化配置写之前，该设备不允许发送Non-Posted请求（需要有效的请求者ID才能正确路由生成的完成）。
  + 例外：Root Complex上的逻辑设备可以在软件初始化配置之前发起请求，访问系统引导设备（system boot devices）。

这条规则和例外，与现有的系统初始化和配置的PCI模型一致。

* 逻辑设备的每个功能模块必须有一个唯一的功能模块号，只能响应对该逻辑设备对应功能号的配置请求。注：每个非ARI设备最多可包含八个功能。每个ARI设备最多可包含256个功能。
* Switch在传递请求时不允许修改事务ID。
* 在某些情况下，PCI Express-PCI桥需要为来自PCI/PCI-X总线的请求生成事务ID。
* **实现时注意——使用Phantom Functions增加未完成请求数**

为了增加需要Completion的未完成请求的最大可能数量，而不仅仅是使用标签位，如果幻象功能号使能位（Phantom Function Number Enable bit，参见7.5.3.4节）被置为有效，设备可以使用未被分配给功能模块的功能号来逻辑地扩展Tag标识符。对于单功能设备，这种方法可以增加到8倍最大请求数。

未声明的功能号称为幻象功能号（Phantom Function Numbers，PFN）。

幻象功能有很多架构上的限制，包括ARI设备、虚拟功能（VFs）以及物理功能（PFs）在启用VFs时缺乏支持。另外，地址翻译服务（ATS）和基于ID的排序（IDO）不理解幻象功能。因此，对于许多实现来说，使用10bit Tag是增加未提交的Non-Posted Requests数量的最好方法。

* **实现时注意——实现10位标记功能的注意事项**

10位标签的使用使请求者能够将其未完成的Non-Posted Requests（NPR）的数量从256个增加到768个，对于非常高速率的NPR，这可以避免标签可用性成为瓶颈。以下公式给出了有效载荷带宽、未完成的NPR数量和其他因素之间的基本关系：

BW=S\*N/RTT，其中：

BW = 有效载荷带宽；

S = 事务有效负载大小；

N = 未完成的NPR数量；

RTT = 事务往返时间。

通常，只有使用相对较小事务的高速链路上的高速请求方才能从将未完成NPR数量增加到256个以上中受益，尽管这也有助于在事务往返时间较高的配置中保持性能。

在具有10位标记请求者能力的请求者需要针对多个完成者的配置中，需要确保请求者只向具有10位标记完成者能力的完成者发送10bit Tag请求。如果所有完成器都具有此功能，则可以大大简化此操作。

为了在整个行业启用10bit Tag，强烈建议所有功能支持10bit Tag Completer Capability。通过新的实施，不需要同时操作更多数量的NPR的完成方本身通常可以在内部跟踪10位标签，并以适度的增量投资在完成时返回它们

实际同时处理更高数量NPR的完成方可能需要大量额外的硬件资源，但除非完成方实际同时处理更多数量的NPR，否则通常无法实现10位标签的全部性能优势

对于RC支持10bit Tag Completer Capability的平台，强烈建议配置PCIe层次结构的平台固件或操作软件在具有10bit Tag Requester Capability的EP中自动设置10位标签请求器启用位。这启用了一类重要的支持10位标记的适配器，这些适配器仅向主机内存发送内存读取请求。

对于RCiEP以外的端点，可以通过检查其相关RP中支持的10位标签完成器位来确定RC是否支持每个端点的10位标记完成器能力。RCiEP没有相关RP，因此，除非RC支持其10位标签完整器能力，否则不允许设置其10位标记请求器支持的位。因此，软件不需要对RCiEP执行单独的检查。

缺乏10位标签完成器能力的交换机仍然能够正确地转发携带10位标签的NPR和完成，因为两个新的标签位在以前保留的TLP报头位中，并且交换机需要在不修改的情况下转发保留的TLP-报头位。然而，如果这样的交换机检测到带有10比特标签的NPR的错误，并且该交换机通过充当NPR的完成器来处理该错误，则所得到的完成将具有无效的10比特标签。因此，强烈建议使用10位标签的任何组件之间的交换机支持10位标签完成器功能。请注意，支持16.0 GT/s或更高数据速率的交换机必须支持10位标记完成器功能。

对于具有10位标签请求者能力的请求者以完成者为目标的配置，其中一些完成者具有10位标记完成者能力，而另一些不具有10位标志完成者能力。请求者如何确定哪些NPR包括10位标签不在本规范范围内。

* **实现时注意——同时使用10位标记和较小的标记**

如本节前面所述，如果请求方支持向某些完成方发送10位标记请求，并支持同时向其他完成方发送较小的标记请求，则如果缺少10位标记完成方能力（10bit Tag Completer Capability）的完成方完成了任何10位标记申请，则请求方必须确保没有未完成的10位标记可以别名为未完成的较小标记。

一种实现方法是让请求方将其8位标签空间划分为两个区域：一个区域仅用于较小的标签（8位或5位标签），另一个区域只用于10位标签的较低8位。注意，这将强制在可用于10位标签的标签空间和较小标签之间进行权衡。

例如，如果请求方将其8位标签空间划分为仅将最低的4位用于较小的标签，则最多支持16个未完成的较小标签，并将10位标签空间减少3\*16个值，支持768-48=720个未完成10位标签。许多其他分区选项也是可能的，所有这些选项都可以减少未处理请求的总数。通常，为较小的标签保留N个值会将10位标签空间减少3\*N个值，并且较小的标签加上10位标签的总数最终为768-2\*N。