# 一、ACS介绍

PCIe协议允许P2P传输，这就意味着同一个PCIe Switch下连接不同End Point可以在不经过RC的情况下相互通信。除了End Point和End Point之间（End Point和End Point都是挂在同一个PCIE Switch下）可以P2P传输，同一个Root Complex下的Root Port之间也可以支持P2P传输。而对于多功能的PCIe dev，dev的不同功能之间也可以进行P2P传输，只是在functions之间需要通过transparent embedded switch进行互相连接（interconnected）（Conceptually, the Functions are interconnected through a transparent embedded switch）。若使用过程中不希望P2P直接通信又不采取相关措施，则该漏洞很有可能被无意或有意触发，使得某些EP收到无效、非法甚至恶意的访问请求，从而引发一系列潜在问题。

ACS（Access Control Services）协议通过设置相关的Control Bit能够决定一个TLP可以被正常路由,阻塞或者重定向。ACS可以被应用于RCs、Switches、Multi-Function Device。如果单功能设备支持SR-IOV功能，则ACS可以将该设备当成多功能设备而使能某些ACS功能。通过在PCIe Switch上开启ACS服务，可以禁止P2P发送，强迫PCIe Switch将所有地址的访问请求送到RC，从而避开P2P访问的风险。

ACS服务是默认关闭的，想要使用ACS访问控制服务需要通过软件配置ACS控制寄存器来打开对应能力。在PCIe系统中可以选择关闭ACS，也可选择全部开启或部分开启ACS访问控制能力。不同PCIe组件对ACS访问控制功能的需求不同，不同的ACS访问控制功能也并非适用于所有的PCIe组件。至于要开启或关闭哪些特定的ACS访问控制能力，则取决于具体的Function类型及其是否为多Function设备或者单功能设备支持SR-IOV的能力。ACS并不适用于PCIE to PCI Bridge dev和Root Complex Event Collector Function。

# 二、访问类型（DN Ports对ACS控制选项的需求）

## 1、ACS Source Validation来源验证（V已实现）

当ACS Source Validation Control Bit被使能时，Downstream Port将会检查该Port接收到来自下游设备发上来的Upstream Request的bus number来确定该Upstream Request是否真正来自Switch Downstream Port的下游dev发出的。如果接收到的Upstream Request的Requester ID中的bus number不在Switch Downstream Port的Secondary Bus Number Register和Subordinate Bus Number Register的范围内，则该Switch Downstream Port将会产生一笔ACS Violation Error。

该部分做在了配置空间的校验模块（pcie5\_cfgchk），报文头校验后得到校验结果信号（tlp\_chk\_tovc）送到接收模块（pcie5\_rxvc）进行分发，分发模块识别ACS ERR后将其分发（rx\_toconf\_n）进而配置空间通路r2c，但由于conf模块识别到该携带acs\_error所以并不写入配置空间，而是在配置空间读写模块（pcie5\_cfgrw）产生c2t数据产生CPL应答报文直通发送。

路由模块目前均校验目标并不校验来源，且目标不路由下行端口的数据均会路由至上行端口。该错误类型的报文不会传输至路由模块已经被干掉且处理完成。

## 2、ACS Translation Blocking转换阻塞（B已实现）

当Downstream Port的ACS Extend Capability的ACS Translation Blocking的Enable Bit被使能时，Downstream Port将会检查该Port接收到的Upstream Memory Request的Address Type（AT），如果AT域不是default值(00)该Switch Downstream Port将会报告一个ACS Violation Error。因为ACS Translation Blocking检查而导致Switch Downstream Port报告ACS Violation Error的优先级顺序高于ACS Upstream Forwarding（第4条）和其他ACS P2P控制机制报告的ACS Violation Error。(This error must take precedence over ACS Upstream Forwarding and any applicable ACS P2P control mechanisms.)

AT与PCIe总线的地址转换相关，在一些PCIe设备中设置了ATC（Address Translation Cache）部件，该部件的主要功能是进行存储器域和PCI域地址转换，但主要目的是方便多个虚拟主机共享同一个PCIe设备；

1、AT字段为0xb00：TLP的Address字段没有通过ATC进行转换，存放的PCI总线域物理地址；如果PCIe设备进行DMA操作，那么该地址将被RC转化为存储器域物理地址，然后对存储器域进行读写操作；

2、AT字段为0xb01：表示当前TLP报文为“Translation Request”报文，该报文由PCIe设备通过存储器读写请求TLP发出，其目的地为TA；而当TA收到该报文后将根据I/O页表的设置，将合适的地址转换关系，通过存储器读完成TLP，发送给PCIe设备；

3、AT字段为0xb10：表示当前的TLP的Address字段已经通过ATC进行地址转换，当PCIe设备进行DMA操作时，RC不需要进行地址转换，而直接将数据发送给存储器。

该部分做在了配置空间的校验模块（pcie5\_cfgchk），报文头校验后得到校验结果信号（tlp\_chk\_tovc）送到接收模块（pcie5\_rxvc）进行分发，分发模块识别ACS ERR后将其分发（rx\_toconf\_n）进而配置空间通路r2c，但由于conf模块识别到该携带acs\_error所以并不写入配置空间，而是在配置空间读写模块（pcie5\_cfgrw）产生c2t数据产生CPL应答报文直通发送。

路由模块目前均校验目标并不校验来源，且目标不路由下行端口的数据均会路由至上行端口。该错误类型的报文不会传输至路由模块已经被干掉且处理完成。

## 3、ACS P2P Request Redirect请求事务重定向（R已实现）

如果Root Port与Root Port之间支持P2P传输，则该Root Ports被要求必须实现ACS P2P Request，而对于Switch Downstream Ports来讲，ACS P2P Request Redirect被要求是必须的。

ACS P2P Request Redirect访问控制权限是否有效还可能受ACS P2P Egress Control（第6条）和ACS Direct Translated P2P（第7条）访问控制权限影响。当Switch Downstream Port的ACS Extended Capability中的ACS P2P Request Redirect被使能，该Switch Downstream Port接收到P2P Requests必须被向上Redirect到RC。当Root Port的ACS Extended Capability中的ACS P2P Request Redirect被使能，P2P Request必须发送到RC的Redirected Request Validation逻辑单元中，由RC的Redirected Request Validation决定是否将该P2P Request Redirect到原始的Target中，还是被RC block住并产生一笔ACS Violation Error。

|  |  |  |  |
| --- | --- | --- | --- |
| E  P2P Egress  Control  出口控制 | R  P2P Request  Redirect  请求事务重定向 | 出口控制  向量比特 acs\_egress  \_ctrl\_vec | 对P2P请求  的处理方式 |
| 0 | 0 | x | 直接路由至P2P目的地 |
| 0 | 1 | x | 上行重定向 |
| 1 | 0 | 1 | 视为ACS违例 |
| 1 | 0 | 0 | 直接路由至P2P目的地 |
| 1 | 1 | 1 | 上行重定向 |
| 1 | 1 | 0 | 直接路由至P2P目的地 |

路由模块路由模块针对类型的报文，现有代码将E和R按照表格实现ACS路由方式。

## 4、ACS P2P Completion Redirect完成事务重定向（C已实现）

如果Root Port实现ACS P2P Request Redirect访问权限控制功能则该Root Port必须实现ACS P2P Completion Redirect。而对于Switch Downstream Port是必须的。

当Switch Downstream Port的ACS Extended Capability中的ACS P2P Completion Redirect被使能，P2P的Completions TLP包中的Relaxed Ordering Attribute Bit不为1时，该P2P的Completions TLP必须被向上redirect到RC。否则P2P Completion TLP包将会按照原始目标进行传输。

当Root Port的ACS Extended Capability的ACS P2P Completion Redirect被使能，Relaxed Ordering Bit为0的P2P Completion TLP包必须向上redirect到RC的Redirected Request Validation Logic中，由RC的Redirected Request Validation决定是否将该P2P Completion TLP发送给原始的P2P Target dev中。

路由模块针对CPLD/CPL类型的报文，如果C标记配置使能且当前报文数据的relax标记为0，则将该CPLD/CPL发送到UP口。

## 5、ACS Upstream Forwarding上行转发（U已实现）

如果RC实现ACS Redirected Request Validation，则该RC所在的Root Ports必须实现ACS Upstream Forwarding。Switch Downstream Port必须实现ACS Upstream Forwarding访问控制权限。

当Switch Downstream Port的ACS Extended Capability中的ACS Upstream Forwarding被使能，Switch Downstream Port的Ingress Port接收到来自下游设备发送的Upstream Request或者Completion TLP（其目标地址是Switch当前的Egress Port，自己发给自己的情况，例如同一个设备的multi\_function互访，或是虚拟地址与物理地址重叠），Switch Downstream Port将接收到的TLP包向上发送到RC。

当Root Port的ACS Extended Capability的ACS Upstream Forwarding被使能，Root Port的Ingress Port接收到其目标地址是Root Port 的Egress Port Upstream Request或者Completion TLP，对于接收Upstream Request的Root Port来讲，Root Port必须将该Upstream Request当作Root Port的ACS P2P Request Redirect被使能的情况处理（即先redirect到RC的Redirected Request Validation逻辑单元中，由RC的Redirected Request Validation决定是否将该P2P Request Redirect到原始的Target中，还是被RC block住并产生一笔ACS Violation Error）。对于接收到Completion TLP的Root Port来讲，Root Port必须将该Completion TLP当作Root Port的ACS P2P Completion Redirect被使能的情况处理（即Completion TLP包必须向上redirect到RC的Redirected Request Validation Logic中，由RC的Redirected Request Validation决定是否将该Completion TLP发送给原始的Target dev中）。

路由模块针对类型的报文，当收到自己发给自己的报文，将其转发给RC。

## 6、ACS P2P Egress Control出口控制（E已实现）

对于Root Port和Switch Downstream Port是可选的。ACS P2P Egress Control会受到ACS P2P Request Redirect和ACS Direct Translated P2P访问控制权限的影响。

当Switch支持ACS P2P Egress Control时，Switch可以通过选择性配置禁止或者允许某些Switch Downstream Port之间进行P2P传输。系统软件可以通过配置ACS Extended Capability中的Egress Control Vector寄存器来选择让哪些Switch Downstream Ports下的devs之间进行P2P传输。

当RC支持ACS P2P Egress Control时，RC可以通过选择性配置禁止或者允许某些Root Port之间进行P2P传输。系统软件可以通过配置ACS Extended Capability中的Egress Control Vector寄存器来选择让哪些Root Port之间进行P2P传输。

当Downstream Ports实现ACS P2P Egress Control时，Downstream Ports中的ACS P2P Egress Control决定了该Port是否应该block或者传输P2P Request。

|  |  |  |  |
| --- | --- | --- | --- |
| E  P2P Egress  Control  出口控制 | R  P2P Request  Redirect  请求事务重定向 | 出口控制  向量比特 acs\_egress  \_ctrl\_vec | 对P2P请求  的处理方式 |
| 0 | 0 | x | 直接路由至P2P目的地 |
| 0 | 1 | x | 上行重定向 |
| 1 | 0 | 1 | 视为ACS违例 |
| 1 | 0 | 0 | 直接路由至P2P目的地 |
| 1 | 1 | 1 | 上行重定向 |
| 1 | 1 | 0 | 直接路由至P2P目的地 |

路由模块路由模块针对类型的报文，现有代码将E和R按照表格实现ACS路由方式。

## 7、ACS Direct Translated P2P定向转换（T已实现）

如果Root Port支持ATS（Address Translation Services）和允许与其他Root Port之间进行P2P传输则该Root Port必须支持ACS Direct Translated P2P。而对于Switch Downstream Port来讲，ACS Direct Translated P2P是必须支持的。

当Switch Downstream Port的ACS Direct Translated P2P被使能，AT（Address Type）域的值表示为Translated Address的P2P Memory Requests必须直接被路由到peer Egress Port，无需关注ACS P2P Request Redirect（第3项）和ACS P2P Egress Control（第6项）的使能情况。而对于其他类型的P2P Requests（AT域表示为非Translated Address类型）需要考虑ACS P2P Request Redirect和ACS P2P Egress Control配置。

AT与PCIe总线的地址转换相关，在一些PCIe设备中设置了ATC（Address Translation Cache）部件，该部件的主要功能是进行存储器域和PCI域地址转换，但主要目的是方便多个虚拟主机共享同一个PCIe设备；

1、AT字段为0xb00：TLP的Address字段没有通过ATC进行转换，存放的PCI总线域物理地址；如果PCIe设备进行DMA操作，那么该地址将被RC转化为存储器域物理地址，然后对存储器域进行读写操作；

2、AT字段为0xb01：表示当前TLP报文为“Translation Request”报文，该报文由PCIe设备通过存储器读写请求TLP发出，其目的地为TA；而当TA收到该报文后将根据I/O页表的设置，将合适的地址转换关系，通过存储器读完成TLP，发送给PCIe设备；

3、AT字段为0xb10：表示当前的TLP的Address字段已经通过ATC进行地址转换，当PCIe设备进行DMA操作时，RC不需要进行地址转换，而直接将数据发送给存储器。

路由模块中针对类型的报文，判断定向转换是否使能且AT字段是否为0x10，如果满足以上条件则P2P转发，并屏蔽violation操作使得①屏蔽E/R操作问题；②屏蔽“强制”（报文本身不是路由UP口）的上行转发操作；

# 三、增强访问能力和类型

## 1、ACS I/O Request Blocking请求阻塞（未实现）

如果Root Port和Switch Downstream Port支持ACS Enhanced Capability，Root Port和Switch Downstream Port必须支持ACS I/O Request Blocking。

当Switch Downstream Port或者Root Port使能ACS I/O Request Blocking时，该Port的Ingress接口必须将收到Upstream I/O Request当作ACS Violation处理。

## 2、ACS DSP Memory Target Access存储器目标访问（未实现）

如果Root Port和Switch Downstream Port支持ACS Enhanced Capability，Root Port和Switch Downstream Port必须支持ACS DSP Memory Target Access。

当Switch Downstream Port的ACS Control Register中的ACS DSP Memory Target Access域写入不同的值时，Switch Downstream Port对接收到其访问Memory地址是Downstream Port的Memory Bar Space的Upstream Request处理情况不同。当ACS DSP Memory Target Access的值为00（表示Direct Request Access Enabled）时，表示Downstream Port允许该Memory Request访问Downstream Port的Memory Bar Space。当ACS DSP Memory Target Access的值为01（Request Blocking Enabled），表示Downstream Port将block住该Memory Request访问Downstream Port的Memory Bar Space。当ACS DSP Memory Target Access的值为10（表示Request Redirect Enabled），表示Downstream Port将把该Memory Request Redirect到RC。

## 3、ACS USP Memory Target Access存储器目标访问（未实现）

如果Switch Downstream Port支持ACS Enhanced Capability，Switch Downstream Port必须支持ACS USP Memory Target Access。而Root Port不支持ACS USP Memory Target Access访问控制权限。

ACS USP Memory Target Access访问控制权限决定了当Switch Downstream Port接收到了其访问Memory地址是Switch Upstream Port的Memory Bar Space时，系统将根据Switch Downstream Port的ACS Control Register中的ACS USP Memory Target Access域的值不同，由Switch Downstream Port选择blocked或者redirect或者发送给Target。

## 4、ACS Unclaimed Request Redirect未声明请求重定向（未实现）

如果Switch Downstream Port支持ACS Enhanced Capability，Switch Downstream Port必须支持ACS Unclaimed Request Redirect。而Root Port不支持ACS Unclaimed Request Redirect访问控制权限。

当ACS Unclaimed Request Redirect被使能，由Switch Downstream接收到其目标访问地址是Switch Upstream Port的Memory Bar Space范围内,但并不在Switch的任何Downstream Port的Memory BAR Space的范围内的Upstream Request会被redirect到Switch的上游。

如果ACS Unclaimed Request Redirect未被使能，以上描述的Upstream Request会被Switch Downstream Port当作Unsupported Req处理。

# 四、Multi-Function Devices和支持SR-IOV的Functions Devices对ACS的要求（未实现）

我们之前介绍了Switch Downstream Port和Root Port对ACS访问控制的需求，接下来我们看看多function device和支持SR-IOV的functions device对ACS访问控制的需求。单function device(不支持SR-IOV功能)不应该支持ACS。

1、ACS Source Validation:禁止支持

2、ACS Translation Blocking:禁止支持

3、ACS P2P Request Redirect:如果该function与其他function之间P2P传输则该function dev必须支持ACS P2P Request Redirect。这也包括支持SR-IOV的虚拟functions。

如果Multi-Function device使能ACS P2P Request Redirect，在同一个dev的多function之间进行P2P Requests必须被Redirect到RC。

4、ACS P2P Completion Redirect: 如果该Multi-Function dev支持ACS P2P Request Dedirect，则该Multi-Function dev也必须支持ACS P2P Completion Redirect。

引出ACS P2P Completion Redirect的目的就是为了当P2P Request被Redirect后，相应的P2P Request Completion可能会超过Redirect后的P2P Request从而违反PCIe的Ordering规则。

当Multi-Function Device使能ACS P2P Completion Redirect，Relaxed Ordering bit为0的P2P Completion TLP必须被Redirect到RC。

5、ACS Upstream Forwarding: 禁止支持

6、ACS P2P Egress Control: 可以选择性支持。对于多功能设备，可以通过配置多功能设备的每个function dev的ACS P2P Egress Control的enable bit来禁止该function dev与其他function dev之间进行P2P传输。

7、ACS Direct Translated P2P: 如果Multi-Function Device的functions支持Address Translation Services(ATS)和function之间进行P2P传输，则该多function dev必须支持ACS Direct Translated P2P。