# PCIe的DMA通道设计需求及概要方案

## 1、双通道DMA方案

SmartNIC网卡在应用中，主要面向为服务器提供一种低时延、高带宽、智能化的网卡解决方案，基于FPGA实现的smartNIC，又赋予网卡可编程的特性，本文考虑在此应用环境下，PCIe的DMA模块实现需求及方案，并着重优化网卡的时延特性。

SmartNIC的网络接口可能存在以下配置：

100G MAC的AXI4-ST接口 512bit@322MHz

40G MAC的AXI4-ST接口 [256bit@312.5MHz](mailto:256bit@312.5MHz)

10G MAC的AXI4-ST接口 [32bit@312.5MHz](mailto:32bit@312.5MHz)

SmartNIC的系统侧接口配置为：

PCIe硬核的物理层接口 Gen3 x16 （8GT/s x 16 lane），内部逻辑接口AXI4-ST 512bit@250MHz，相关特性遵循Xilinx的PG213文档。

DMA功能的需求特性：

1. 使用PCIe硬核的RC/RQ通道，实现memory write操作和memory Read操作
2. DMA模块面向用户逻辑的接口为AXI4-ST接口，用于用户数据的收发
3. DMA模块的描述符由逻辑电路控制，可每帧配置帧长、收发地址等信息
4. DMA模块面向用户最大提供4个AXI4-ST接口，每接口独立配置（可选）
5. DMA模块按照优先级或轮询调度的方案给每个接口分配C2H的带宽
6. 支持32/64位的读写地址
7. 每个描述符支持的最大传送长度为256MB（待定）
8. 只支持1个PF
9. 在满足传输带宽的前提下，尽量减少经过DMA模块的传输时延
10. 在RC/RQ通道采用DWORD对齐方式下，可选支持帧拼接模式（Straddle）

下图为双向DMA的实现框图，为了实现低时延的数据传输，考虑从两个方面优化设计方案。首先，DMA控制数据收发的描述符，原来是由Host主机创建，通知DMA读取并使用，DMA完成相关操作后，再由中断通知Host主机。描述符的传递会消耗一定的时间，并部分叠加在数据的传输时延上。如果将DMA的需要使用的描述符交给FPGA进行创建和维护，可以节省描述符传递消耗的时间。同时可以给CPU传输一份描述符的镜像，用以同步CPU和DMA模块的工作状态。其次，压缩数据经过DMA处理的需要节拍数，减少逻辑的处理时延。

C2H方向，AXI4-ST接口向数据写入缓存buf后，根据调度模块的仲裁，选择一路数据使用空闲的描述向RQ通道生成TLP帧，如果帧长大于TLP的最大playload，则进行分段处理。

H2C方向，描述符生成模块将FPGA维护的描述符链表发送给CPU，并读取CPU使用状态，一旦发现CPU有数据写入空闲的描述符空间，则启动读取操作。生成READ TLP帧从RQ接口发出，从RC接口收到的TLP帧进行解析，将数据写入缓存buf中，根据目的地址（或端口号）写入到相应的AXI4-ST接口中。



图1 双向DMA实现框图

## 2、写通道DMA方案

在面向智慧金融的应用中，只需要C2H方向的DMA操作，重点强调低时延的性能，相应的需求规格调整为：

1. 上行的用户侧接口为数据位宽大于等于512比特的读FIFO接口
2. 上行的用户侧接口数量为可选配置项（测试环境最大支持到4）
3. 上行的用户侧接口可独立配置数据位宽（需同时满足需求1）
4. 上行的系统侧接口为PCIe硬核的RQ通道，固定配置为AXI4-ST 512bit@250MHz
5. 用户侧接口支持按帧进行轮询调度的方式使用RQ通道带宽
6. C2H发送数据的描述符由逻辑电路控制，可每帧配置帧长、收发地址等信息
7. 支持32/64位的读写地址
8. 每个描述符支持的最大传送长度为？字节（待定）
9. PCIe硬核的RQ通道的最大理论带宽为128Gb/s，在使用最大可用带宽的情况下，尽量减少经过DMA模块的传输时延
10. 在RC/RQ通道采用DWORD对齐方式下，可选支持帧拼接模式（Straddle）

下图为C2H方向的DMA实现框图。RR调度模块在帧接口模块收到数据Ready的指示后，根据调度模块的算法，从可用的接口中选择一路读取数据，并写入到缓存buf中，一旦缓存buffer中的数据满足发送条件，则TLP的发送控制模块即开始封装TLP帧，然后进行发送。TLP的分段控制模块识别出发送帧长大于TLP的最大playload，则将用户帧分割成多个TLP帧进行发送。RQ通道的tready信号反压RR调度模块从用户接口读取数据帧的操作。帧接口模块在每帧数据到来时，将该帧的帧长信息传送给描述符生成模块，该模块生成TLP帧使用的目的地址，并支持配置多个HOST缓存区。



图2 C2H方向的DMA实现框图