

光纤通道 SOPC 中的 DMA 研究与实现

张雷锋, 王忆文, 李 辉

(电子科技大学 电子薄膜与集成器件国家重点实验室, 四川 成都 610054)

摘 要: 针对可编程片上系统(SOPC)的光纤通道适配器设计, 分析了 XPS Central DMA 和 Scatter-Gather DMA 两种 DMA 的特点. 并在开发板上做了硬件实现, 实验结果表明 Scatter-Gather DMA 传输速率可以达到 225MB/s, 满足光纤通道传输速率要求(200MB/s).

关键词: 光纤通道适配器; DMA; Xilinx 嵌入式开发套件; 片上可编程系统

中图分类号: TN47

文献标识码: A

文章编号: 1000-7180(2013)04-0131-04

Research and Implementation of DMA in the Fibre Channel SOPC

ZHANG Lei-feng, WANG Yi-wen, LI Hui

(State Key Laboratory of Electronic Thin Films & Integrated Devices,
University of Electronic Science & Technology, Chengdu 610054, China)

Abstract: In order to design a Fibre Channel based on Programmable System on Chip (SOPC), both the XPS Central DMA and the Scatter-Gather DMA are studied. Finally some tests are done on a development board, and the experimental results show that the Scatter-Gather DMA data transmission rate can reach 225MB/s, which meets the Fibre Channel transmission rate requirements (200MB/s).

Key words: fibre channel adapter; DMA; Embedded Development Kit(EDK); SOPC

1 引言

随着计算机通信技术的不断进步, 在存储设备和服务器之间要求快速、安全地传输数据. 由于光纤通道(Fiber Channel, FC)具有优异的性能, 从最初应用于存储区域网络, 正逐步成为航空电子网络中主要的传输标准^[1-2].

直接内存访问(Direct Memory Access, DMA)是由 DMA 控制器完成的存储器与外部设备或者存储器之间的大量数据传输方式. 在 DMA 传输中, 硬件设备代替 CPU 进行数据传输, 提高数据传输效率, 并降低了处理器负担, 从而提高了系统的性能.

本设计中, FC 传输速率为 2.125Gb/s, 去除帧头、帧尾以及 8b/10b 编码所带来的冗余位, 实际有效数据传输速率大约为 200MB/s. FC 的高传输速率要求光纤通道适配器具有足够高的数据处理和搬运能力. 在光纤通道适配器中大块数据的传输主要

发生在 PCIe 与 DDR2 SDRAM 和 DDR2 SDRAM 与 FC 收发端之间, 为了减轻 PowerPC 的负担, 提高系统的性能, 这两处的数据传输由 DMA 实现.

2 系统架构

嵌入式开发套件(EDK)是 Xilinx 公司推出的 SOPC 开发工具包, 它是用于设计嵌入式可编程片上系统的全面解决方案. 本设计通过 EDK 中的硬件开发平台 XPS(Xilinx Platform Studio)完成了基于 FPGA 的硬件系统设计. 系统框图如图 1 所示.

本系统采用的是基于 CoreConnect 总线的 PowerPC440 作为处理器, 此外还包含了 DDR2 SDRAM、UART、Flash、BRAM 以及 PCIe 数据通道和光纤通道. PCIe 数据通道用于实现 PC 机与本系统之间的数据通信, 光纤通道用于实现与本节点相连的另一节点之间的数据通信. PowerPC 通过总线实现对光纤通道模块的控制和相关寄存器的读

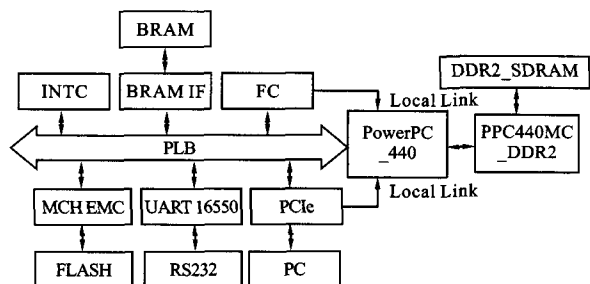


图1 光纤通道适配器系统构架

写^[2]。

3 XPS 中的 DMA

在 XPS 中有两类 DMA: XPS Central DMA 和 SGDMA (Scatter - Gather DMA)^[3]。XPS Central DMA 原理与应用比较简单, 直接实现数据从源地址向目标地址的搬运; 而 SGDMA 则类似于以太网, 由描述符(Buffer Descriptor, BD)控制 DMA 工作。SGDMA 又包括软核 DMA (Soft DMA, SDMA) 和硬核 DMA (Hard DMA, HDMA), 这两种 DMA 原理相同, 收发通道都是通过 Local Link 接口连接, 不同的是 SDMA 控制器由多端口内存控制器(MPMC)提供, 在速度上不如嵌在 PowerPC 处理器模块中的 HDMA。所以, 在此处只对 XPS Central DMA 和 HDMA 做出分析。

3.1 XPS Central DMA

XPS Central DMA 是基于 PLB (Processor Local Bus) 总线的 DMA, 可以完成基于总线的外设和内存之间的数据传输。XPS Central DMA 由从模块、主模块和 FIFO 三个主要模块组成^[4], 如图 2 所示。

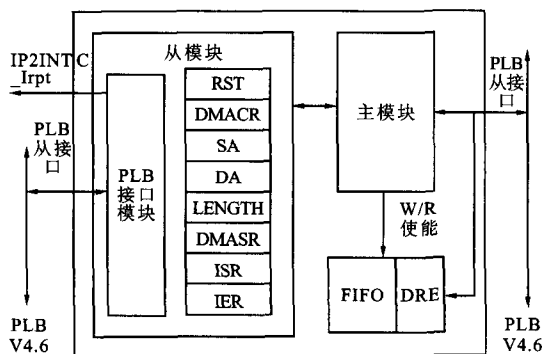


图2 XPS Central DMA 结构图

从模块通过 PLB 从接口模块和 PLB V4.6 相连, 它主要响应 CPU 对 DMA 寄存器的读写, 例如源地址、目标地址、传输长度、DMA 状态和中断寄存器等。主模块通过 PLB 主接口可以连接在 SPLB

或 MPLB 上, 作为读写的发起者, 除了进行数据的搬运, 主模块在数据的传输过程中还要负责检测错误和更新源地址、目标地址、传输长度以及状态寄存器。FIFO 位宽为 32bit, 主要用作数据缓存。

利用 XPS Central DMA 进行数据传输的流程比较简单:

- (1) 配置源地址、目标地址和传输长度等寄存器;
- (2) DMA 将源地址的数据读入缓冲 FIFO, 然后再写入目标地址;
- (3) 重复(2)直到数据传输完毕;
- (4) 更新 DMA 状态寄存器。

XPS Central DMA 的优点是原理简单, 软件编程不复杂; 不足之处是数据的传输需要通过 PLB 总线, 延时较大, 而且在数据传输过程中, 要求源物理地址和目标物理地址必须是连续的, 传输完一块物理地址上的连续数据后引起一次中断, 需要再由主机进行下一块物理地址上的传输^[5]。

3.2 HDMA

在 PPC440 处理器块中有四个独立的可以全双工工作的 HDMA 模块, 每个 HDMA 都包含一个本地链路通道 (Tx 和 Rx), 用户外设通过 Local Link 接口和 DMA 进行数据交互, DMA 模块则通过另外的接口和处理器相接^[6]。结构示意图如图 3 所示。

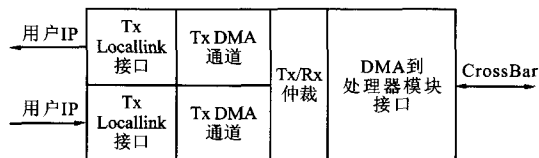


图3 HDMA 结构图

3.2.1 描述符

HDMA 是基于描述符进行数据传输控制的^[6]。DMA 描述符定义了下一个描述符指针、当前缓冲区数据地址、待发送的数据长度、状态控制寄存器以及 4 个应用参数, 其数据结构如图 4 所示。

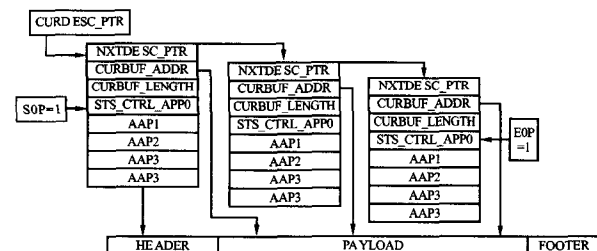


图4 描述符和链表示意图

将多个描述符链接起来可以构成 BD 链表, 该链表对所需传输的数据进行控制, 图 4 的链接结构

表示一段数据,该数据由内存中非连续的三部分组成,可以用三个描述符链接来表示。

一个描述符的生命周期有如下四部分:

- (1)在内存中创建并初始化;
- (2)由软件程序预处理,准备好数据和控制值并连接到描述符链上;
- (3)提交给 DMA 控制器进行处理;
- (4)DMA 控制器处理完毕,用户程序可以通过描述符信息获得 DMA 处理结果。

3.2.2 中断合并机制

HDMA 在每传输完一个数据包后都会提出中断申请,如果 CPU 响应每个中断,无疑会增加 CPU 负担,使中断处理延时增大,降低传输速度。为解决这一问题,在 DMA 传输过程中使用了中断合并机制:预先设定中断合并计数器 TxIrqCountReg(0~255)的值,即 TxIrqCountReg 个中断发生后才产生一个给 CPU 的中断。

以发送通道为例,原理如图 5 所示。复位后,8 bit 中断合并计数器载入中断阈值 TxIrqCountReg。每当 eop 或 irq-on-end(interrupt on end)到来时,计数器进行减一运算,直至计数器的值为 0。此时,4bit 的计数器开始加法计数,只要该计数器的值不等于 0,就会向 CPU 提出中断请求。中断请求得到响应后,DCR ack 置 1,4 bit 计数器进行减法计数,8 bit 计数器重新载入中断阈值,进行下一次运算。

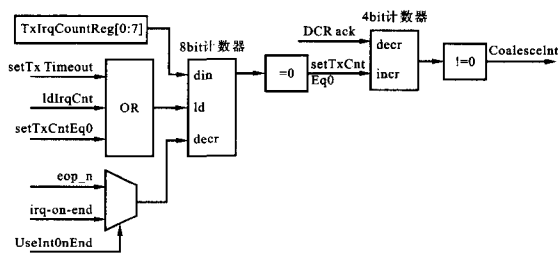


图5 中断合并机制

HDMA 数据传输方式利用专门的 DMA 通道,延时小,在 BD 的控制下可以传输不连续地址上的数据。不足之处是软件程序设计复杂,中断处理时延时较大^[5](在实际应用中,用中断合并可以改善这一情况)。

4 性能测试

本设计中所用的硬件平台是基于 5VFX70T FPGA 的定制开发板,除了基础外设电路,该开发板还配置了光纤适配器所必需的 PCIe 接口和光模

块。开发工具是 Xilinx ISE 11.1 以及 EDK 11.1,片上系统内的 PowerPC 时钟频率为 400Mhz,板上 DDR2 SDRAM 时钟频率为 200Mhz,其余部分如 Flash、HDMA、XPS Central DMA 等均采用总线时钟频率,为 100Mhz。软件部分采用了 xilinx 公司的 xapp1121 项目作为参考^[7],主要实现以下功能:

- (1)配置 DMA 控制寄存器,控制数据发送和接收;
- (2)设置发送的数据量,通过检测开始和结束标志位,将传输所用时钟周期写入相应寄存器;
- (3)计算传输速率,并在超级终端上打印出来。

为了测试 HDMA 的 Tx 和 Rx 通道传输速率,需要添加一个用户 IP,该 IP 实现简单的回环功能,即从 HDMA Tx 通道接收数据,存入缓存 FIFO 中,然后再发送给 HDMA 的 Rx 通道。用户 IP 包括两个接口:进行数据传输的 Local Link 接口和与 PLB 总线连接的状态寄存器配置接口,IP 核结构如图 6 所示。

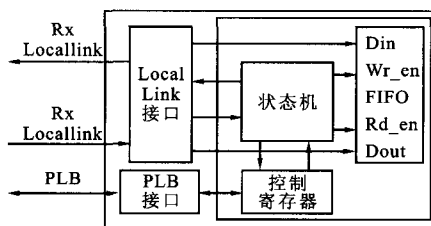


图6 用户 IP 结构示意图

在 HDMA 软件测试程序中,定义数据包大小为 2KB,128 个描述符,中断合并阈值设为 32。用户逻辑中设置标志位检测机制,该机制用于检测第一个帧的开始位 SOF(Start of Frame)和最后一个帧的结束位 EOF(End of Frame),计算 SOF 和 EOF 之间的时钟周期,并存入用户寄存器 slv_reg0 中,供软件计算速度时调用。

与 HDMA 测试类似,为了测试 XPS Central DMA 的传输速率需要添加一个用户 IP,该 IP 功能较为简单:检测传输开始标志位 PLB_pavalid 和完成标志位 Done,计算 PLB_pavalid 和 Done 信号之间的时钟周期,存入寄存器中 slv_reg0,供软件调用。在软件测试程序中对源地址、目标地址、传输数据长度以及相关寄存器完成配置后 DMA 即可工作。在此次测试中传输数据量设为 4MB。

HDMA 和 XPS Central DMA 测试结果如图 7 所示。由输出结果可以看出,HDMA 收发通道传输速率均可以达到 225MB/s,符合光纤通道适配器的速度要求(200MB/s)。

```

Starting XPS Central DMA Transfer(s):
XPS Central DMA 83.220846 MB/s
Starting HDMA Transfer(s) to DDR2:
HDMA Tx 225.518182 MB/s
HDMA Rx 225.479983 MB/s
Test Completed!

```

图 7 XPS Central DMA 和 HDMA 数据传输速率

5 结束语

本文针对基于 Xilinx PowerPC 的光纤通道 SOPC 设计,研究了 XPS Central DMA 和 HDMA 两种 DMA 的特点,对比了它们的优缺点,并在开发板上做了硬件实现.实验结果表明 HDMA 收发通道传输速率均可以达到 225MB/s,满足光纤通道的传输速率(200MB/s)要求.

参考文献:

- [1] Fibre Channel Framing and Signaling-3 (FC-FS-3), Rev0.94[S]. 2010.
 - [2] 陈俊,基于 FPGA 的光纤通道系统设计与实现[D].成都:电子科技大学,2012.
 - [3] PLB v3.4, OPB to PLB v4.6 System and Core Migration User Guide, REV1.4[S]. 2008.
 - [4] XILINX XPS Central DMA Controller, REV2.01a[S]. 2009.
 - [5] 张宇,冯丹.针对 Xilinx 可编程片上系统的硬件加速方案的研究[J].小型微型计算机系统,2010,31(6):1147-1151.
 - [6] XILINX Embedded Processor Block in Virtex-5 FPGAs, Rev1.8[S]. 2010.
 - [7] James Lucero, XILINX Reference System: Optimizing Performance in PowerPC 440 Systems, REV1.0[S]. 2008.
- 作者简介:**
张雷锋 男,(1987—),硕士研究生.研究方向为光纤通道、片上可编程系统.
王忆文 男,(1966—),教授.研究方向为 SOC/SOPC 的设计与测试、光纤通道和无线网络.
-
- (上接第 130 页)
- [7] 杨继萍,王颖.一种基于节点交换的 DHT 优化方法[J].计算机工程,2011:34-38.
 - [8] 郝结.基于 DHT 的结构化 p2p 路由协议 chord 的研究与改进[D].北京:北京邮电大学,2010.
 - [9] A Gupta, B Liskov and, Rodrigues. One hop lookups for peer-to-peer overlays[C]//Proceedings of the 9th Conference on Hot Topics in Operating Systems, 2007: 34-41.
 - [10] S Saroiu, P Gummadi, S Gribble. A measurement study of peer-to-peer file sharing systems[J]. Of SPIE/ACM MMCN, 2008: 35-40
- 作者简介:**
杜经纬 男,(1978—),硕士,讲师.研究方向为计算机网络与信息安全.