

文章编号: 1001-893X(2005)05-0036-06

基于 CSIX-L1 接口协议的通用交换芯片设计*

江 浩^{1,3}, 郭东辉^{1,2,3}

(1 厦门大学 物理系 EDA 实验室, 福建 厦门 361005; 2 厦门大学 电子工程系, 福建 厦门 361005;

3 厦门睿智微电子有限公司, 福建 厦门 361005)

摘要: 在原有 ATM 交换芯片设计成果的基础上, 设计了一款具有 CSIX-L1 接口的通用交换芯片。本文主要介绍该芯片的设计方案。首先简要介绍 CSIX-L1 协议的基本内容, 然后具体说明如何直接采用 ATM 交换芯片的空分交换结构来实现传递交换 CSIX-L1 协议帧内容, 并指出 CSIX-L1 接口扩展是该芯片的设计关键, 最后介绍如何采用 Bit-slice 技术来实现 CSIX-L1 接口扩展。

关键词: 网络处理器; 通用交换接口; 交换结构; 位片技术

中图分类号: TN915 **文献标识码:** A

Design of a Switch Chip Based on CSIX-L1 Interface Protocol

JIANG Hao^{1,3}, GUO Dong-hui^{1,2,3}

(1 EDA Lab, Department of Physics, Xiamen University, Xiamen 361005, China; 2 Department

of Electronic Engineering, Xiamen University, Xiamen 361005, China; 3 Xiamen Rich II

Microelectronic Technologies LTD, Xiamen 361005, China)

Abstract Based on the IP core of ATM switch fabric, a universal switch with CSIX-L1 interface is proposed and implemented in verilog codes. This paper presents the design scheme of this universal switch. It firstly introduces the CSIX-L1 protocol and then shows the architecture of the universal switch. The key of the design is the implementation of CSIX-L1 interface, which can be conveniently expanded by using bit-slice technology.

Key words Network processor; CSIX-L1; Switch fabric; Bit-slice technology

一、引言

随着网络技术的迅速发展, 网络处理器 (Network Processor, 以下简称 NP) 被认为是推动下一代网络发展的核心技术^[1]。为了保证不同 NP 设计和不同厂家生产的 NP 能够实现交换操作, 推进 NP 的发展和最大限度地保护用户利益, 网络处理器论坛^[2]于 2001 年 2 月 19 日正式成立。该论坛主要是由原先的通用交换接口协会 (CSK) 和通用编程接

口联盟 (CPIX) 的成员组建的, 其中, CSK 主要针对硬件接口标准的制定, 而 CPIX 则是针对软件接口标准的制定。

CSIX-L1^[3] 是 CSK 协会于 2000 年 8 月公布的 NP 和交换结构 (Switch Fabric, 以下简称 SFabric) 之间的通用接口协议, 它适用于如 ATM、MPLS、IP、以太网等各种数据包交换的通信设备。在 CSK 并入 NP 论坛之后, 该标准也被 NP 论坛所采纳, 因此, 基于 CSIX-L1 的交换芯片将被各大通信系统

* 收稿日期: 2005-04-06

基金项目: 国家自然科学基金资助项目; 国家人事部留学人员创业基金项目; 福建省自然科学基金资助项目 (A0410007)

开发商采用,各个芯片设计公司^[2]开始注重该通用交换芯片的设计开发。我们在原有设计 ATM 交换芯片的基础上^[6,7,8],也开始设计这类通用交换芯片。本文介绍的就是我们所设计的基于 CSK-L1 交换芯片的设计方案。

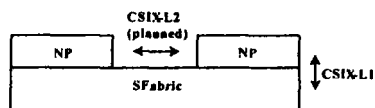


图1 网络设备中 CSK 协议的定义示意图

二、CSK-L1协议的内容简介

通用交换接口 (CSK-L1) 协议^[3,4]是一种提供在 NP 和 SFabric 之间基于信元和分组信息传送的标准接口协议,如图 1 所示,它给出了该接口在网络设备中所处的位置,其中 CSIX-L2 是 NP 与 NP 之间的标准接口协议。CSK-L1 协议的标准建立可以使系统设备厂商方便地选用不同芯片厂家的 NP 芯片和 SFabric 芯片来设计生产网络设备产品,而 CSIX-L1 协议主要是通过定义 NP 与 SFabric 之间传送的信息元帧 (CellFrame, 简称 CF 帧) 来定义的。下面具体介绍 CF 帧的格式和类型。

1. CF 帧格式

与其他信元交换的帧格式一样, CSIX-L1 的 CF 帧如表 1 所示,由基本头部 (Base Header)、可选的扩展头部 (Extension Header)、可选净负荷 (payload)、可选填充比特和垂直校验域 (vertical parity) 等部分组成。每个 CF 帧的基本头部为 2 字节,用来说明帧的类型和长度,其格式如表 2 所示。基本头部后面跟的是特定类型的扩展头部,不同类型的扩展头部不一样。扩展头部主要用来附加说明对帧信息的特定处理,如对单播操作的目的交换端口。Payload 部分主要为所要传递的信息内容,为了保证 CF 帧有合适的长度,采用填充比特方法来增加 Payload 的长度。而校验域为 2 字节共 16 比特来实现垂直校验功能。

表 1 CF 帧格式

CFrame Component	Header		Payload	Vertical Parity
	Base Header	Extension Header		
Length	2 bytes	0~4 bytes	Maximum allowable length is 256 bytes	2 bytes

表 2 基本头部

Byte number	Bit position															
	7	6	5	4	3	2	1	0	7	6	5	4	3	2	1	0
0/1	ready		Type			CR	P	Payload length								

可见, CF 帧的内容首先是由帧的基本头部决定的,其中基本头部的各域作用如下:

Ready 域: 用来指示发送者是否准备好接收信息,为低 (0) 时表示现在还不能接收信息,为高 (1) 时表示准备好可以接收信息。通常,低位比特 (第 6 比特) 用于控制业务,如命令、流控、和状态控制帧;高位比特 (第 7 比特) 用于数据业务,如组播、单播和广播;

Type 域: 用来指明所要传送 CF 帧的信息类型。目前定义的信息类型有空闲帧 (Idle)、单播帧 (Unicast)、组播帧 (Multicast)、广播帧 (Broadcast)、流控帧 (Flow Control Frame)、命令和状态帧 (Command

and Status), 它们对应的类型码可参见文献 [3]。除了流控帧以外,其他所有帧的净负荷都没有被 CSIX-L1 协议所指定;

CR (CSK Reserved): 预留比特;

P (Private): 私有比特,在 CSIX 中没有定义;

Payload length 表示所传送净负荷的长度,用字节数来表示,不包含填充比特部分。

2 帧的类型

由前面的介绍可知,不同类型的 CF 帧有不同的扩展头部。下面我们重点介绍各个类型 CF 帧的扩展头部。

空闲帧,是在系统启动和没有数据传输期间发送的帧,用以保持基本头部中 ready域的活动状态和维护时钟的同步。4个字节,由基本头部和垂直奇偶校验域构成,如表3所示。

表3 空闲帧的结构

Byte Number	Contents
0~ 1	Base Header
2~ 3	Vertical Parity

数据帧,是用来在 NP和 SFabric之间传送数据

信息。按不同的信息类型又分为单播帧、组播掩码帧、组播 ID 帧、组播二元拷贝帧和广播帧。由于篇幅所限,在这里我们只介绍单播帧、组播 ID 帧和广播帧(其它的类型帧可以参考文献[3])。

单播帧,它的扩展头部如表4所示,12比特的地址标明不同的 NP接口,可支持多达 4 096个交换端口。目的地址只在入口方向有效,当 CF 帧被正确传送到出端口之后,目的地址就不再有效。它还有 8比特类型域以表示 256个不同的单播信息类型。

表4 单播帧的扩展头部

Byte Number	Bit Position															
	7	6	5	4	3	2	1	0	7	6	5	4	3	2	1	0
2/3	Class								P		CSK Reserved					
4/5	CSIX Reserved				Destination Address											

组播帧,它的扩展头部如表5所示,第3字节的第7比特是专用比特,第6比特是 CSK 预留比特。22比特的 ID是一个查询标签, SFabric 根据这

个查询标签决定哪些 NP能收到当前的组播帧。组播帧的 ID在入口和出口都是有效的。类型域则表示当前组播帧归属的组播队列。

表5 组播 ID数据的扩展头部

Byte Number	Bit Position															
	7	6	5	4	3	2	1	0	7	6	5	4	3	2	1	0
2/3	Class[7 0]								P	CR	Multicast ID[2: 16]					
4/5	Multicast ID[15 0]															

广播 (Broadcast)帧,它的扩展头部如表6所示,8比特的类型域可以表示 256种广播队列中的任何一种。

表6 广播数据的扩展头部

Byte Number	Bit Position															
	7	6	5	4	3	2	1	0	7	6	5	4	3	2	1	0
2/3	Class[7 0]								Private				CSK Reserved			
4/5	CSK Reserved															

流控帧,用来提供比链路层更为精确的流控,因为它可以指定专门的 NP类型。一个流控 CF 帧如表7所示,由基本头部、流控条目和两字节的垂直校验域组成。其中,流控条目是可变长的,其最多条目数由最大允许净负荷长度决定。

每个流控条目的格式如表8所示,共4个字节,其中,第3字节的第4、5比特分别表示端口通配符和类型通配符,类型域也是8比特,它对于单播操作与单播扩展头部是一致,而对于组播则是指定256种组播队列中的任何一种。流控条目类型主要有单

播、多播、广播等,其类型值可参见文献[3]。Speed域为4比特,指明NP可以以多快的速度发送CF帧到指定的目的地址。

SCell通过SFabric模块,然后,由Tx模块将SCell转换为CF帧输出给接在输出端口的NP。

表7 流控CF帧格式

Byte number	Contents
0 1	Base Header
2 3	Flow Control Entry 1a
4 5	Flow Control Entry 1b
6 7	Flow Control Entry 2a
8 9	Flow Control Entry 2b

	Last Flow Control Entry in frame, bytes & 2
	Last Flow Control Entry in frame, bytes & 2
N	(pad to interface width)
	Vertical Parity

表8 CF流控条目的格式

Byte number	Bit Position															
	6	5	4	3	2	1	0	7	6	5	4	3	2	1	0	
7																
2/3	Class						FC Entry Type		C*		P*		Speed			
4/5	P	CSX	Destination Address													



图2 通用交换芯片的整体设计

SFabric模块是采用Crossbar无阻塞的交换矩阵,如图3是一个4×4的Crossbar交换矩阵。只要SCell中有输入端口和输出端口的地址信息,Crossbar矩阵就可以容易地被控制来调度实现SCell的交换。鉴于我们原有设计的ATM交换芯片具有相同的SFabric模块^[7,8],因此,直接采用ATM交换芯片的IP核(即16个端口且每个端口32bit数据线的SFabric模块)。这样,SCell可以定义成如同ATM Cell的信元帧结构,如图4所示。该SCell具有56个字节,其中交换标签1个字节;输入输出端口地址1个字节;1个字节的类型域,用来表示不同的信息类型;还有1个字节的分类,表示不同信息类型的子类型,如对于单播信息和组播信息分别表示不同的单播类型和组播信息归属的组播队列;一个字节的预留字;剩余部分为净负荷。

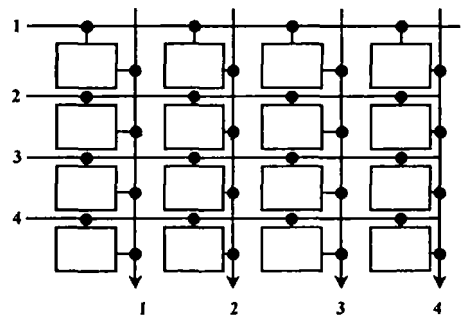


图3 4×4型的矩阵型 crossbar结构

标签标记 (FF)	输入端口, 输出端口
类型 (type)	净负荷长度
分类 (class)	预留
净负荷	

图4 内部交换信元SCell的帧结构

三、通用交换芯片的结构设计

从上面协议帧的定义上看,CSIX-L1协议支持多达4096个通用交换端口。由于它只是定义用于将NP连接到SFabric的界面标准,因此,符合该协议标准的通用交换芯片需要在SFabric的输入输出端口上加上收发处理模块。我们所设计的这个具有CSIX-L1接口的通用交换芯片是以16个收发端口且每个端口为32位数据线构成的SFabric^[7,8]为基础的,整体设计框图如图2所示,它是由CF帧接收处理(Rx)模块、SFabric模块和CF帧发送处理(Tx)模块组成的。Rx模块接收来自入口NP的CF帧并转换为交换信元(Switch Cell以下简称SCell),以

Rx模块和Tx模块的原理框图分别如图5和图6所示。Rx模块将接收到的数据(RxData[31:0])缓存(Rx Buffer),在数据拆分单元(Data Destructor Unit)中将信息转换成SCell然后输出SFabric模块,

在校验产生 (Parity Gen) 单元中生成奇偶校验码。这些过程的完成都是由 CSK 状态机 (CSIX state machine) 来控制完成的。Tx 模块接收来自 SFabric 模块输出的 SCell 数据 (input_bus[15:0]), 在数据重组单元 (Data Constructor Unit) 中将数据转换成 CF 帧, 在校验生成单元中产生奇偶校验码, CF 帧和输出校验经过输出缓存 (Tx Buffer) 输出给 NP。同样, 在 Tx 模块中, 这些过程的完成也是由 CSK-L1 的状态机来控制完成的。

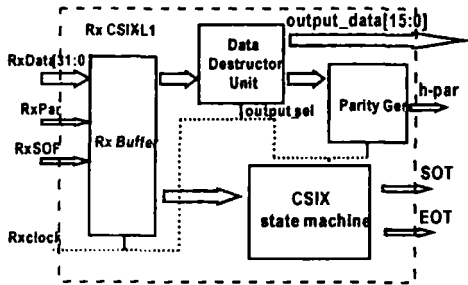


图 5 Rx 模块原理框图

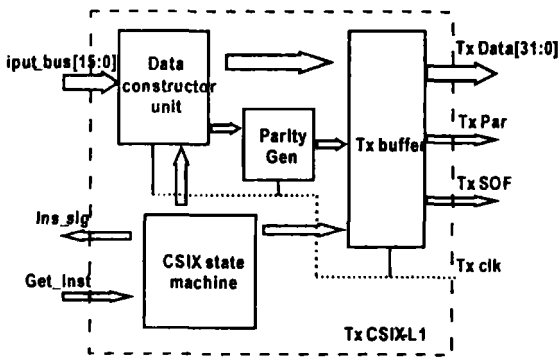


图 6 Tx 模块原理框图

四、CSK-L1 接口的扩展兑现

在 CSK-L1 协议中, 发送与接收间的数据总线宽度可以是 $n \times 32$ 比特 ($n=1, 2, 3, 4$), 即其数据总线分别是 32、64、96、128 比特。如在 100 MHz 下, 32 比特数据线就可以支持 2.5 Gbps OC-48 的速率; 在 200 MHz 下, 64 比特数据线可支持 10 Gbps OC-192 的速率。而前面所述的交换数据总线是 32 比特, 为了符合 CSK-L1 接口标准的协议, 即实现 64、96、128 比特的数据总线的可扩展, 我们可以采用 bit-slice(位片)技术^[10, 11, 12]。

表 9 32-bit 的 CSK-L1 接口信号功能

信号	方向	功能
RxDat[31: 0]	从信息管理单元到交换结构	接收数据
RxPar[0]	从信息管理单元到交换结构	接收数据奇校验
RxSOF[0]	从信息管理单元到交换结构	开始接收 CF 帧
RxCk[0]	从信息管理单元到交换结构	接收时钟
TxDat[31: 0]	从交换结构到信息管理单元	发送数据
TxPar[0]	从交换结构到信息管理单元	发送数据奇校验
TxSOF[0]	从交换结构到信息管理单元	开始发送 CF 帧
TxCk[0]	从交换结构到信息管理单元	发送时钟

当接口的数据总线为 32 比特时, 我们所设计的接口信号线图 7 所示, 各个信号功能见表 9。对于 64 比特数据宽度, 奇偶校验位 (TxPar RxPar)、时钟位 (Txclk Rxclock) 和开始发送 CF 帧位 (TxSOF、RxSOF) 各为 2 比特的宽度, 每个比特对应 32 比特的数据, 并且高位对应数据的高 32 位, 低位对应数据的低 32 位。要实现 64 比特的数据宽度, 应用 bit-slice 技术只要将 32 比特做为一个 slice, 将 2 个 32 比特的 slice 并联就可以构成一个 64 比特数据宽度的接口。同理很容易就可以扩展成 96、128 比特的接口。

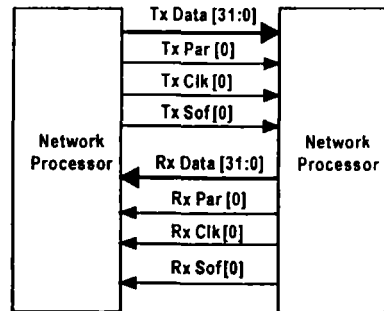


图 7 32-bit 的 CSK-L1 接口

采用 ModelSim 软件工具仿真我们用 Verilog 语言兑现方案, 结果表明 CF 帧和 SCell 之间可以正确转换。其中, 图 8 为 Tx(发送)模块的仿真结果, 第二个时钟开始在 input bus 输入 SCell 信号, 输出 CF

帧在延迟了 2 个时钟周期后出现在 output bus 图 9 为 Rx(接收)模块的仿真结果, RxData 上为接收到的 CF 帧, C41801000002 为该帧的头部, 转换后的 Scell 延迟一个时钟周期后出现在 output to system 上。这里只给出单播类型的 CF 和 Scell 之间的转换。

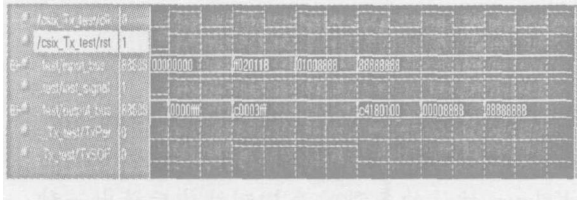


图 8 Tx 模块仿真结果

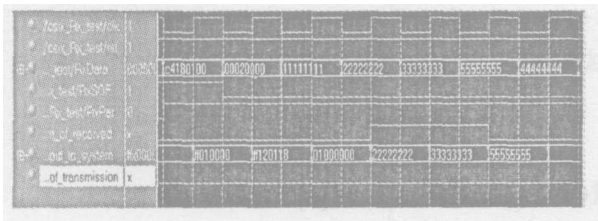


图 9 Rx 模块仿真结果

五、结 论

本文所述的是符合 CSIX-L1 协议的通用交换芯片的设计, 我们采用的是自顶向下的设计方法, 先用 Verilog 语言进行功能描述, 然后综合到 FPGA 来验证。仿真结果表明, 该设计方案是可行的, 适用于 bit-slice 技术进行接口扩展。

参考文献

[1] 周文江. 蓄势待发的网络处理器[M]. 微电脑世界, 2002 (1).

[2] www.np6nm.org
 [3] CSIX-L1, Common Switch Interface Specification-L1 Version 1.0[S].
 [4] 石晶林, 等. 网络处理器原理、设计与应用[M]. 北京: 清华大学出版社, 2003.
 [5] Ship P E, et al A building block chip for a scalable ATM architecture[J]. Circuits and system, 1994.
 [6] 李立峰. 宽带通讯交换的信元智能调度及性能分析研究[D]. 厦门: 厦门大学, 2000.
 [7] 蔡水成. 高速交换系统的研究及其专用集成电路的前端设计[D]. 厦门: 厦门大学, 2002.
 [8] 杨永红. 信元交换系统的单片集成设计及性能分析[D]. 厦门: 厦门大学, 2003.
 [9] Markus Adhivijayo. common switch interface CSIX-L1 reference design[EB/OL]. www.xilinx.com, 2002-04.
 [10] Donnamie EW hite. Bit-Slice Design: Controllers and ALUs[M]. Garland Publishing Inc, 1981.
 [11] Michael Valdez et al the design of computers using bit-slice technology[D]. Florida Department of Electrical and Computer Engineering Florida Institute of Technology Melbourne 1987.
 [12] Shambhu Upadhyaya. CSE341: Computer Organization Lecture Cost and Design[EB/OL]. http://www.cse.buffa.edu/faculty/shambhu 2002-09.

作者简介:



江浩(1974-), 女, 福建人, 厦门大学物理系硕士研究生;

郭东辉(1967-), 男, 福建莆田人, 厦门大学教授、博士生导师。

《电讯技术》专题资料

《通信系统中的无源非线性技术研究》题要(五)

HPM条件下电子器件中的PM问题研究

随着通信卫星技术和高功率微波武器的发展, 电子系统中无源器件的互调干扰成为日益严重的问题。本文简要地总结了从 20 世纪 70 年代至今国际和国内无源互调问题的研究概况, 并从无源互调产生的机理、无源互调产物阶次和类型的预测、无源互调产物电平的预测等多个方面进行分析, 提出了预测无源互调产物的新方法。