



Xilinx 大学计划指定教材

零存敷取

NetFPGA 开发指南

陆佳华 杨卫 编著
周剑 张克农





Xilinx 大学计划指定教材

零存整取 NetFPGA 开发指南

陆佳华 杨卫周剑 张克农 编著

北京航空航天大学出版社

内 容 简 介

本书以通俗易懂的语言,由浅入深地向读者阐述了 NetFPGA 的安装、开发,剖析了 NetFPGA 的参考路由器的逻辑、驱动以及上层软件设计,分析了已有的各种典型应用,同时结合作者的开发实例介绍了如何在此平台上开发用户设计以及注意事项。本书在阐述 NetFPGA 开发的同时,更侧重于介绍在 FPGA 上进行以太网相关逻辑开发的方法与设计思路;同时也阐述了板卡如何与主机系统交互、如何开发驱动等整个设计流程。因此本书虽然是基于 NetFPGA 平台,但是其中的源代码与设计思路同样适用于 Xilinx 其他 FPGA 平台。

本书可作为 NetFPGA 初学者、FPGA 上网络相关硬件开发人员的参考书,亦可作为大专院校从事 FPGA 网络硬件加速研究方向的相关教师和研究生的参考书。

图书在版编目(CIP)数据

零存整取 NetFPGA 开发指南 / 陆佳华等编著. -- 北京
: 北京航空航天大学出版社, 2010. 6
ISBN 978 - 7 - 5124 - 0107 - 5

I. ①零… II. ①陆… III. ①可编程序逻辑器件一系
统设计—指南 IV. ①TP332. 1 - 62

中国版本图书馆 CIP 数据核字(2010)第 094850 号

版权所有,侵权必究。

零存整取 NetFPGA 开发指南

陆佳华 杨 卫 周 剑 张克农 编著
责任编辑 刘 星

*

北京航空航天大学出版社出版发行

北京市海淀区学院路 37 号(邮编 100191) <http://www.buaapress.com.cn>

发行部电话:(010)82317024 传真:(010)82328026

读者信箱: cmsbook@gmail.com 邮购电话:(010)82316936

北京时代华都印刷有限公司印装 各地书店经销

*

开本: 787×960 1/16 印张: 16.25 字数: 364 千字

2010 年 6 月第 1 版 2010 年 6 月第 1 次印刷 印数: 4 000 册

ISBN 978 - 7 - 5124 - 0107 - 5 定价: 32.00 元



FOREWORD

;; of which (1) is the original work by the author(s), and (2) is not covered by a copyright notice. In the event that the author(s) has rights under such a license, then the terms of such license shall take precedence over those of this license.

Designing networking hardware is hard, and very expensive. Although thousands of engineers in networking companies design ASICs, code FPGAs and build new networking products every day, our students don't get the opportunity to learn how to build networking hardware in college campuses. NetFPGA aims to change that.

The NetFPGA is the first platform in the world designed to allow students—in the classroom—to design their own networking hardware, running at line rate, and then actually deploy it in a real running network. NetFPGA was designed for teachers and students, to give the type of hands-on experience not normally possible in a university.

Several hundred students have used NetFPGA to do lab assignments where they build 4Gb/s Ethernet switches and Internet routers.

In the past couple of years, the number of NetFPGA boards has grown; at the time of writing, over 150 universities have 1,500 boards in 18 countries. And the numbers keep growing. An increasing number of researchers (graduate students, professors and researchers in industrial research labs) are using NetFPGA to prototype new ways to process packets.

The main goal of the program is to build a thriving community of "Open Source Networking Hardware" designs, contributed to by students, researchers and ASIC designers in industry. Our goal is to exchange designs and modules worldwide, and build a valuable repository of useful designs.

It is very fitting that the first book about NetFPGA is for China. As we all know, the

FOREWORD

networking industry is growing at a tremendous rate, and many college graduates are going on to design networking hardware. A large and growing fraction of the world's networking hardware is designed and built in China. How appropriate for students in China to gain hands-on experience learning networking hardware in preparation for their future job.

It is therefore wonderful that Joshua Lu has written this book, to help the new NetFPGA users to get started. His thoughtful material takes us through the basics of setting up and using the NetFPGA, and some exercises to get started creating our own designs. I am confident that this new book will open up a whole new world of possibilities for networking researchers. Every NetFPGA user and developer in China should buy this book. It is an invaluable guide.

—Nick McKeown
Stanford

前 言

NetFPGA，拆分开就是 Network 与 FPGA，这两个名词对于电子工程师来说都是耳熟能详的，直译可简单称为网络 FPGA。通常普通的 FPGA 用户也知道 FPGA 上能实现哪些网络的应用，但这两个名词混合到一起又表示什么，又有哪些高明之处呢？

NetFPGA 起源于斯坦福大学由 Nick McKeown 教授领导的研究小组，作为一个软硬件皆可编程的开放平台，应用在网络设计课程中。学生只需要大约十周的时间，就可以开发出实际的网络设备，如网卡、交换机、路由器、防火墙等。NetFPGA 小组刚卸任的负责人之一 John Lockwood 教授也是一直从事着网络应用硬件化加速的工作，其最主要的开发、验证平台就是 Xilinx FPGA。

记得 2003 年在读硕士时就了解到 Lockwood 教授在开展利用 FPGA 进行网络加速的研究，而作者当时所在的课题组也在尝试做一些网络应用的硬件加速，然而那时对于 FPGA 和网络都是一知半解，一切都得从头开始、白手起家。为了搭建一个 PCI 接口的 FPGA 加速平台，从原理图、PCB、硬件调试、驱动开发、逻辑开发到上层应用整整花了一年半的时间，花费了很多精力，却也没有得到特别好的效果。等平台搭建好，也就离毕业不远了，因此反而没有太多时间研究硬件加速应用本身。

毕业后有幸来到 Xilinx 公司工作，在一次与大学计划经理谢凯年博士的聊天中得知 Lockwood 已经加入 Nick 的小组，而斯坦福大学正在和我们公司合作开展 NetFPGA 的项目，顿觉我的师弟们前途一片光明，也明白这对他们的课题来讲意味着什么。NetFPGA 不仅提供硬件平台，还提供驱动、逻辑源代码。这些源代码能够让读者更快地熟悉网络相关的逻辑设计，更好地利用搭建好的底层逻辑；同时易用的驱动程序和上位机程序也帮助硬件开发人员极大地减少了与上位机交互的工作量，这就能让开发人员更专注于应用和加速本身，而不用考虑硬件平台，降低了做网络加速应用开发与研究的门槛，这也和 Xilinx 近期推出的“目标设计平台”的概念不谋而合，NetFPGA 也算是以网络加速应用为目标的设计平台。

因为工作的便利可以第一时间把这个平台推荐给了仍在学校的师弟杨卫、王飞、周剑等，在此基础上他们也确实做了不少工作，比如流量采集、流量回放、内容匹配、数据包产生、入侵

前言

检测等。2008年的夏天,曾经和作者一起并肩作战于交大实验室、现正在美国UCSB攻读博士的好友高明回国休假,了解到有NetFPGA这一平台,于是我们和师弟王飞一起花了不到两个月的时间就开发了一个千兆环境下BT流量监测的应用模型,也算是弥补了我们毕业时的遗憾。能在这么短的周期内开发出完整的应用模型,也是得益于NetFPGA这个平台,站在巨人的肩膀上,起点自然就高了很多。

目前NetFPGA这个平台在全球也已经有了很多的用户,但是国内开展这方面研究的确实还不多。一来,NetFPGA虽然降低了做网络加速应用开发的门槛,但这个平台的入门本身就有一定的难度;二来,很多用户都觉得必须要有NetFPGA这个平台,才能做这方面的开发,但这个平台相对价格还是较高,个人购买有一定的困难。适逢全国大学生信息安全邀请赛即将举行,这次邀请赛的开发平台之一就是NetFPGA,所以在Xilinx大学计划经理谢凯年博士的促成下,决定以此为契机编写一本NetFPGA开发指南。这样,不仅可以把我们做的这些工作和学习经验介绍给国内的开发人员,解决入门问题(本书名中“零存整取”之意正来源于此);而且也是想告诉大家并不是必须有NetFPGA这个平台才可以做这方面的开发,只要有心去做工作,理解了NetFPGA理念和源代码,一样可以在其他硬件平台上进行研究开发。

2

本书第1章介绍了NetFPGA及其技术渊源;第2章介绍了NetFPGA平台的搭建;第3章仔细分析了Reference Router的硬件结构和数据流程;第4章介绍了驱动的开发;第5章介绍了其他经典应用;第6章介绍了高阶开发,利用BT流量的检测实例介绍;第7章在理解了NetFPGA理念和源码的基础上,介绍了对其移植的想法及注意事项。

本书由陆佳华列出具体内容提纲,硬件设计大部分由杨卫完成编写,软件驱动部分由周剑完成,第2章由大学计划实习生阙志强完成,最终由杨卫和陆佳华一起完成统稿与校正。师弟陈海荣、郑伟、陈灿、高西洋完成了书中插图的绘制,并通阅了全书;西安交通大学张克农副教授通阅了全书,并对全书进行了指导;斯坦福大学的James Hongyi Zeng通读了全书并提出了大量宝贵意见;Xilinx大学计划经理谢凯年博士在本书的编写过程中对本书提供了宝贵意见,在此对他们表示衷心的感谢。

本书只求能抛砖引玉,为国内从事网络加速应用开发研究人员尽一份工程师的绵薄之力。由于编写较为仓促,个人经验及能力有限,对网络应用和FPGA开发的理解可能有很多不准确之处,恳请各位专家和读者不吝赐教,以便在适当的时间再做修订补充。有兴趣的读者,可以发送邮件到:joshua.lu@xilinx.com,与作者进一步交流;也可以发送邮件到:emsbook@gmail.com,与本书策划编辑进行交流。

陆佳华

2010年2月于上海

目 录

第 1 篇 初识 NetFPGA

第 1 章 网络 FPGA	3
1.1 NetFPGA 溯源	3
1.2 核心部件	5
1.2.1 FPGA	5
1.2.2 Memory	6
1.2.3 PHY	7
1.2.4 PCI	8
1.2.5 SATA	9
第 2 章 NetFPGA 平台搭建指南	11
2.1 NetFPGA 主机清单	11
2.1.1 官方网站推荐主机清单	11
2.1.2 Xilinx 大学计划使用的主机清单	12
2.1.3 预装机购买	12
2.1.4 机器选购的一些建议	12
2.2 操作系统介绍及其安装	15
2.2.1 NetFPGA 兼容的操作系统介绍	15
2.2.2 Bios 设置	16
2.2.3 Cent OS 4.4 安装指南	16
2.3 NetFPGA 系统快速安装法	24
2.3.1 Java 环境安装	25
2.3.2 rpmforge 安装	28
2.3.3 NetFPGA 基础开发包安装	29

目 录

2.3.4 其他设置.....	34
2.4 NetFPGA 系统详细安装法	35
2.4.1 设置 Grub 参数	35
2.4.2 下载 NetFPGA 基础开发包	35
2.4.3 设置环境变量.....	38
2.4.4 检查是否安装了对应版本的 Linux 内核源代码	39
2.4.5 安装 perl 支持包	39
2.4.6 安装 Java	40
2.4.7 安装 NetFPGA 驱动	40
2.4.8 验证是否安装成功.....	43
2.4.9 执行 CPCI	43
2.5 安装 NetFPGA 开发工具——综合工具	44
2.5.1 ISE 版本规定	44
2.5.2 Linux 如何安装 ISE9.1.03	44
2.5.3 如何设置环境变量.....	47
2.6 安装 NetFPGA 开发工具——仿真工具及其相关设置	48
2.6.1 ModelSim 的安装	48
2.6.2 安装内存仿真模块.....	48
2.7 安装 NetFPGA 开发工具——调试工具	49
2.8 NetFPGA 的测试	49
2.8.1 selftest 版本 1	49
2.8.2 selftest 版本 2	50
2.8.3 regress test	52

第 2 篇 近观 NetFPGA

第 3 章 深入浅出 Router 硬件	63
3.1 为什么是 Router	63
3.2 纵观 Router Architecture	64
3.3 硬件设计结构的思考.....	66
3.3.1 关键技术之 Packet 和 Register Bus	68
3.3.2 关键技术之 5 级 pipelining	69
3.3.3 关键技术之统一 Packet 格式	69
3.3.4 我们需要关注什么.....	70
3.4 链路层	82
3.4.1 认识 MAC 核	82

3.4.2 Router 中的 MAC 核	84
3.4.3 链路层的辅助设计	86
3.4.4 如何使用 TEMAC 核	91
3.5 核心层面的网络层	95
3.5.1 简单的队列调度	95
3.5.2 出色的转发引擎	97
3.5.3 管理好输出缓冲	110
3.5.4 SRAM 接口设计	112
3.5.5 留给读者的电路	116
3.6 数据交互的 PCI 接口	117
3.7 HDL 源码探究	118
第 4 章 深入浅出 Router 软件	123
4.1 驱动程序的结构	123
4.1.1 驱动概述	123
4.1.2 NetFPGA 驱动简介	124
4.1.3 PCI 驱动介绍	127
4.1.4 nf2 设备探测和初始化	129
4.1.5 nf2 设备卸载	131
4.2 设备驱动的操作	132
4.2.1 打开与关闭	132
4.2.2 数据包是如何接收的	133
4.2.3 驱动如何发送数据包	135
4.2.4 这样来配置硬件板卡——ioctl	137
4.2.5 换一种方式来实现驱动程序	138
4.3 用户界面分析	140
4.3.1 为什么要有用户界面	140
4.3.2 用户界面如何操控硬件	142
第 3 篇 再会 NetFPGA	
第 5 章 经典应用剖析	147
5.1 视频流 demo	147
5.2 通用的 Packet Generator	151
5.2.1 硬 件	152
5.2.2 软 件	158

目 录

5.3 新颖的 OpenFlow	159
5.3.1 了解 OpenFlow Switch	160
5.3.2 如何在 NetFPGA 上搭建 OpenFlow	164
5.4 丰富的 Project	170
5.4.1 值得分析的 Project	170
5.4.2 更多的 Project	180
5.5 贡献你的 Project	182
第 6 章 开发实践	185
6.1 选择流量检测	185
6.2 硬件设计方法	189
6.2.1 开始前的准备	190
6.2.2 设计正确的 module	191
6.2.3 提交放心的 module	198
6.2.4 添加新的 module	201
6.3 驱动设计方法	210
6.3.1 驱动设计准备	210
6.3.2 提升数据传输速率的两种方法	212
6.3.3 怎样更加轻松地使用驱动程序	214
6.4 应用程序设计方法	219
6.4.1 功能验证利器 C 语言程序	219
6.4.2 Java 编写 GUI 让你的演示更 nice	226
6.4.3 Makefile 浅谈	230
6.5 系统调试	232
第 7 章 皆可 NetFPGA	238
7.1 高性能的 NetFPGA	238
7.1.1 目标平台	238
7.1.2 硬件移植	243
7.1.3 PCIe 驱动开发	246
7.2 轻量级的 NetFPGA	246
7.3 NetFPGA 资源共享	248
附录 NFP2.0 的改进	249
参考文献	250

第 1 篇

初识 NetFPGA

- 网络 FPGA
- NetFPGA 平台搭建指南

第 1 章

网络 FPGA

只用一样东西，不明白它的道理，实在不高明。

——林语堂

认识 NetFPGA 开发平台是过程、是手段，最终目的是良好运用 NetFPGA。作者希望读者不仅要清楚地了解 NetFPGA 的历史渊源，掌握开发平台上的硬件资源；更重要的是体会 NetFPGA 板卡的设计特点，寻找采用目前这种设计的原因和优势。

1.1 NetFPGA 溯源

NetFPGA，从字面上看可以称为网络 FPGA，这个平台也确实与网络息息相关；从技术上看，除了作为教学、研究下一代网络结构的应用外，其在网络安全领域也有不小的斩获，比如 UC Berkeley 的 Nicholas 发表的文章《The Shunt: An FPGA Based Accelerator for Network Intrusion Prevention》就是在 NetFPGA 平台上实现的，这篇文章主要阐述了如何用 FPGA 实现对高速网络数据流量的实时入侵检测及防护。因为 NetFPGA 提供了易用的软硬件接口和驱动程序，所以用户基本不用花太多的时间去处理 PCI 驱动、板卡与主机的交互、MAC 与 PHY 的交互等，可以更集中精力实现其核心的内容匹配、过滤等算法，并快速添加到 NetFPGA 平台上进行测试。

其实刚卸任斯坦福 NetFPGA 小组负责人的 John Lockwood 教授的主要研究方向之一就是，如何利用 FPGA 来解决高速网络中 IDS/IPS 的性能瓶颈（早期主要是研究一些模式匹配算法在 FPGA 上的实现）以及如何在 FPGA 上构建高速的网络应用的架构。在加入斯坦福 NetFPGA 小组前，一个比较典型的例子就是其在 2003 年发表的一篇文章《An Extensible, System - On - Programmable - Chip Content - Aware Internet Firewall》，在这篇文章中可以看到如图 1.1 所示的结构框图。

在图 1.1 中，大家可以看到以 Xilinx Virtex2 FPGA 为平台，主要实现了网络数据包的解包、负载（Payload）的扫描过滤等，这个部分的架构和思想也可以算作是 NetFPGA 的硬件逻辑雏形了。在后期大家可以越来越多地看到 NetFPGA 与网络安全、网络应用的硬件加速有着千丝万缕的联系。

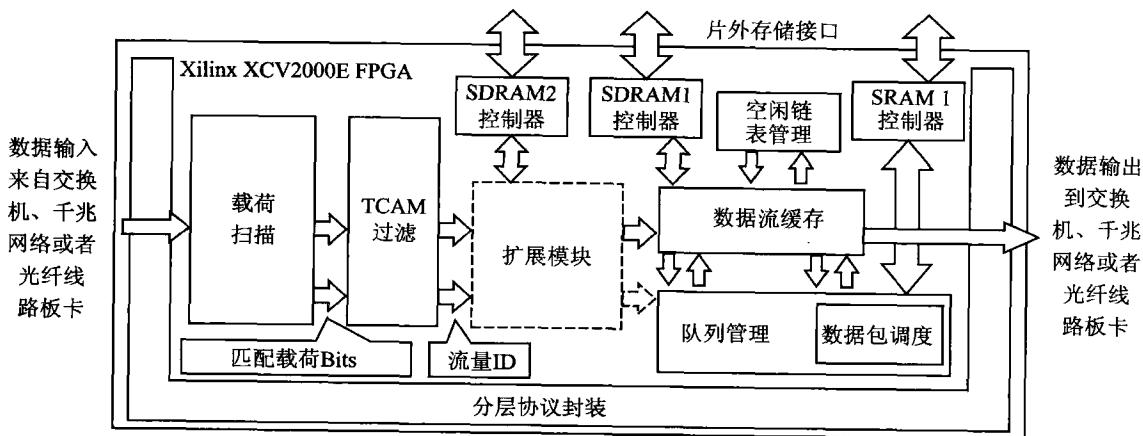


图 1.1 NetFPGA 结构框图

作为一个软硬件皆可编程的开放性平台,NetFPGA 至今经历了两种设计版本。

4

2001 年的 NetFPGA - v1 包括 3 个 Altera 的 EP20K400 APEX 芯片、一个 8 端口的以太网控制器、3 个 1 MB 的 SRAM 和时钟电路;将其应用在网络课程设计中,学生只需要大约十周的时间,就可以开发出实际的网络设备,如网卡、交换机和路由器等。但是早期版本的开发平台存在一定的缺陷:

- 只提供了 8 个 10 Mb/s 的以太端口,其传输速率在课程设计中是足够了,但是很难满足实际研究中的需要;
- 缺少板级的 CPU,无法开发复杂的控制软件,也不利于系统功能的扩展。

时隔 3 年推出的 NetFPGA - v2 拥有 1 个 32 位、33 MHz 的 PCI 总线接口,1 片 Xilinx 的 Spartan 芯片用来实现 PCI 总线接口控制,核心器件 V2P30 用来实现用户逻辑,2 个 SATA 接口实现平台间的通信,还包括 2 片 512K×36 位的 SRAM。V2P30 通过标准的 GMII 接口连接在一个 Marvell 多端口 10/100/1 000 Mb/s 以太网 PHY 上,整个系统的时钟频率为 62.5 MHz。与先前版本相比,一个比较大的改进是增加了 PCI 总线接口,方便了板级硬件和 CPU 软件之间的通信,也可以通过 PCI 总线来配置 FPGA。

作者将要介绍的是 2007 年新推出的 NetFPGA 2.1 开发平台,它在 v2 的基础上对主芯片的逻辑规模进行了扩展,增加了更快速的存储器,其硬件资源主要有:

- Virtex - II Pro XC2VP50 FPGA;
- Spartan - II XC2S200 FPGA;
- ZBT SRAM;
- DDR2 SDRAM;
- BCM5464 PHY 芯片;

- 时钟资源；
- 4个千兆以太网端口；
- 33 MHz PCI 总线接口；
- 2个 Serial ATA 接口。

NetFPGA 2.1 开发平台如图 1.2 所示。

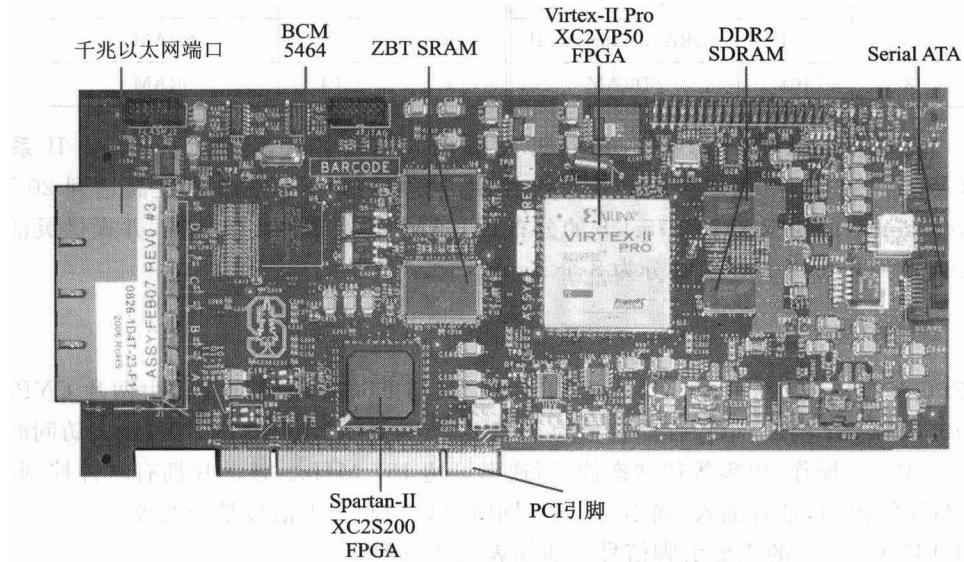


图 1.2 NetFPGA 2.1 开发平台

1.2 核心部件

使用一种以 FPGA 为 核心部件的硬件开发平台,需要关注平台上的哪些硬件资源呢? 作者以为需要充分熟悉开发平台上核心器件的功能以及器件引脚间的连接关系。

1.2.1 FPGA

开发平台上包含了两块可编程的芯片,首先来看看其中比较重要的一个。UFPGA(User FPGA, 实现用户逻辑的可编程芯片)是 Virtex - II Pro 系列芯片,型号是 XC2VP50, 封装形式为 FF1152,速度等级为 7。该芯片包括 2 个 PowerPC 405 处理器、16 个 MGT、8 个 DCM、23 616 个 slices(基本单元)和 4 176 Kb 的块 RAM。PowerPC 405 处理器支持 300 MHz 时钟频率;每个 MGT 能够支持 3.125 Gb/s 的传输速率;852 个用户可选 I/O 支持多种高速 I/O 标准,如 LVDS、SSTL2 和 PCI 等;DCM 支持 24~420 MHz 的时钟频率。

第1章 网络 FPGA

2VP50 芯片同其他芯片的链接如表 1.1 所列。

表 1.1 2VP50 接口表

Bank	可用 I/O	接 口	Bank	可用 I/O	接 口
0	69	Spartan - II	4	69	SATA
1	69	Spartan - II	5	69	BCM5464
2	104	SDRAM, Spartan - II	6	104	SRAM
3	104	SDRAM	7	104	SRAM

另外一块 CFPGA (Control FPGA) 是开发平台的辅助芯片, 选择 Spartan - II 系列的 XC2S200 - FG456C 芯片, 用来实现 PCI 总线控制, 速度等级为 6, 芯片密度高达到 20 万门。它内部具有 56 Kb 的块 RAM 和 75 Kb 的分布式 RAM, 支持多种接口标准, 并提供灵活的时钟处理, 284 个用户可用 I/O 引脚分为 8 个 Bank。

1.2.2 Memory

6

NetFPGA 2.1 上包括了 36 Mb 的 ZBT SRAM(与 FPGA 时钟同步), 由两片 CYPRESS 的 CY7C1370C - 167 构成, 单片存储容量为 512K×36 位, 工作电压为 3.3 V, 最大访问时间为 3.4 ns, 支持 Burst 操作, 以零等待状态读/写速率可达 167 MHz。该芯片拥有时钟控制信号, 支持字节写使能控制, 带有输入/输出寄存器, 同时可以控制输出信号是否有效。

CY7C1370C - 167 的主要引脚信号描述如表 1.2 所列。

表 1.2 CY7C1370C - 167 主要引脚信号描述

引 脚	I/O	描 述
A0, A1, A	I	地址输入, 与 CLK 同步
BWa, BWb, BWc, BWd	I	写字节选择, 低电平有效
WE	I	写使能控制, 低电平有效
CLK	I	时钟输入, 同步芯片的输入信号
CE1, CE3	I	芯片输入使能, 低电平有效
CE2	I	芯片输入使能, 高电平有效
OE	I	芯片输出使能, 低电平有效
CEN	I	始终使能, 低电平有效
DQa, DQb, DQc, DQd	I/O	双向数据 I/O
DQP _a , DQP _b , DQP _c , DQP _d	I/O	双向数据校验 I/O, 这些信号是 DQ[31:0] 的标记位, 在写时分别由 BW 信号来控制