分类号：　　密级：

UDC ：

武汉科技大学

硕士学位论文

基于FPGA的网络安全加速卡研究与设计

Research and Design of Network Security Accelerator based on FPGA

明幼林

指导教师姓名：吴谨　教授

武汉科技大学信息科学与工程学院

申请学位级别：工学硕士　专业名称：电路与系统

论文定稿日期：2010-4-27　论文答辩日期：2010-5-25

学位授予单位：武汉科技大学

学位授予日期：

答辩委员会主席：刘惠康　教授

评　阅　人：刘文予　教授

陈　静　教授

武汉科技大学

研究生学位论文创新性声明

本人郑重声明：所呈交的学位论文是本人在导师指导下，独立进行研究所取得的成果。除了文中已经注明引用的内容或属合作研究共同完成的工作外，本论文不包含任何其他个人或集体已经发表或撰写过的作品成果。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。

申请学位论文与资料若有不实之处，本人承担一切相关责任。

论文作者签名： 明幼林 **日期：** 2010.5.28

研究生学位论文版权使用授权声明

本论文的研究成果归武汉科技大学所有，其研究内容不得以其它单位的名义发表。本人完全了解武汉科技大学有关保留、使用学位论文的规定，同意学校保留并向有关部门(按照《武汉科技大学关于研究生学位论文收录工作的规定》执行)送交论文的复印件和电子版本，允许论文被查阅和借阅，同意学校将本论文的全部或部分内容编入学校认可的国家相关数据库进行检索和对外服务。

论文作者签名： 明幼林

指导教师签名： 吴谨

日　　期： 2010.5.28

摘 要

随着网络应用范围的迅速扩大和网络性能的不断提高，安全网关的吞吐量从原来的几十兆/每秒迅速增长到几百兆/每秒甚至几千兆/每秒。另外，安全网关从传统的防火墙应用扩展到UTM、在线流量分析监控、Web访问管理等领域。因此，网络管理设备需要有更加强大的CPU处理能力、更强大的数据报文分析能力和流量转发能力来支撑整个网络的高速运转。在实际应用过程中，转发工作往往消耗了过多的服务器CPU资源，使得CPU没有足够的资源对高速数据报文进行分析，影响了网络的性能，制约了网络的应用。因此，本文设计了一种网络安全加速卡来有效解决网络数据包处理的瓶颈问题。

本文首先介绍了课题的研究背景，概述了网络安全方面的威胁和对策、网络加速方面的常见技术。接着，对多种处理器的实现方式进行比较，得出了本文的主要研究内容。通过对网络隔离原理的分析，总结了网络安全隔离的要求；对数据包分类字段、TCAM的硬件查找原理和规则管理算法做深入的研究，总结出一种合适的硬件规则管理方法。在此基础上，设计一种基于FPGA的网络安全加速卡，该卡采用硬件数据包分类和硬件转发的方式来分担CPU的数据分析和转发工作量。另外，通过对匹配条目的具体配置，和对网络数据报文的采样，能够有效地拦截有威胁的报文，具备网络安全功能。

网络安全加速卡的设计包括：总体框架设计、网络接口模块设计、硬件查找模块设计、硬件转发模块设计、PCIE总线接口设计、电源时钟设计、掉电保护电路设计、FPGA设计。最后，就网络应用的主要性能，与软件加速方法做了具体的比较。

本设计在用硬件方式实现数据报文分析、规则匹配、报文转发等功能的基础上，还为未来可能的新应用保留了升级的空间。在保持现有硬件平台不变的基础上，通过软件的升级可以实现地址转换、状态跟踪和Qos等功能。本设计具有比较好的可扩展性和应用价值。

关键词： 安全网关，报文分析，流量转发，规则匹配，网络安全加速卡

Abstract

With the rapid expansion of the network applications and the continuous improvement of network performance, security gateway's throughput rapidly grows from the original scores of megabytes per second to several megahertz or multi-gigabit per second. Moreover, the security gateway is extending from the traditional firewall applications to the UTM, online data stream analysis, web access management and other fields. Therefore, the network management device needs to have a more powerful capability of CPU processing, data analysis and flow packet forwarding to support high-speed running of the entire network. In practical applications, the forward work often consumes too much server CPU resources, and makes the CPU not have enough resources to analyze high-speed data packet, limiting the network performance and application. A network security accelerator is designed to effectively solve the bottlenecks in network packet processing.

Firstly, the background of the research is introduced. An overview of threats and countermeasures about network security, and the common technology about network acceleration is given. Then, implementations based on multi processors are compared, obtaining the main contents of this paper. By analyzing the principles of network isolation, the network security requirement of the separation is summed up. Algorithms that search for hardware principles and rules management for Packet Classification, TCAM is studied deeply. An appropriate hardware rules management is summed up. On above basis, a network security accelerator card based on FPGA is designed. The accelerator shares data analysis and forwarding workload for CPU, through hardware packet classification and sharing. At last, by matching the specific configuration items, and network data packet sampling, threatening messages can be effectively blocked with network security features.

The design of network security accelerator card includes the overall framework, network interface, hardware search module, the hardware forwarding module, PCIE bus interface, power clock, power-down protection circuit, FPGA design. Finally, the main performance for network applications is compared with the software accelerated method.

In this design, space for upgrade is reserved based on data packet analysis, rule matching, packet forwarding and other functions with hardware. While maintaining the existing hardware platform unchanged, functions of the address translation, state tracking and Qos can be achieved through software upgrades. This design is of better scalability and application.

Keywords: Security gateway, Packet analysis, Stream forwarding, Rule matching, Network security accelerator

目 录

第1章 绪论

1.1 引言

目前，计算机网络发展非常迅速，各政府部门和企事业单位，都大量通过网络进行信息查询、邮件收发、数据共享等各种办公操作[2]。由于计算机网络通信具有信息量大、信息更新速度快、信息处理和利用方便等优点，使得计算机网络通信已逐渐成为各个单位日常工作不可或缺的一部份，整个社会已步入网络信息化时代。网络的飞速发展给人们带来方便的同时，也带了一系列的新问题[2]。

一方面，网络的飞速发展给企业和用户带来了便利，但同时也对网络安全管理提出了严峻的挑战。局域网内部以及局域网与互联网之间不断增长的数据通信，使网络及网络设备在负载、工作效率以及安全性方面都承受着巨大的压力，网络时断时续、网络速度慢、网络遭受攻击等故障一直制约着网络的正常运行。

另一方面，随着Internet技术和应用的飞速发展，各种新的应用不断涌现，造成网络流量不断增加。在这种情况下，网络管理设备既要有更加强大的报文分析和流量分析功能，也需要对经过本设备的流量进行高效地转发处理。因此，需要有强大的CPU处理能力来同时保证分析工作和转发工作迅速完成。而网络服务器CPU处理能力的限制往往使得转发工作挤占了分析工作所需的资源，产品应用开发人员不得不在性能和功能的平衡取舍问题上花费很多的精力，网络性能问题甚至成了制约功能进一步丰富完善的瓶颈。

因此，在保证网络安全性的同时进一步提高网络的性能，成为了一个热门的研究课题。

1.2 网络安全

网络安全是指网络系统的硬件、软件及其系统中的数据受到保护，不受偶然的因素或者恶意的攻击而遭到破坏、更改、泄露，确保系统能连续、可靠、正常地运行，网络服务不中断。通常，网络安全主要指网络上的信息安全，包括物理安全、逻辑安全、操作系统安全、网络传输安全[3]。

1.2.1 网络安全面临的威胁

计算机网络安全面临的威胁很多，包括人为的和自然的、信息安全和设备安全[3]等等。总的来说，对计算机网络造成威胁的主要原因包括以下几点：

(1)人为的无意失误：如操作员安全配置不当造成的安全漏洞，用户安全意识不强，用户口令选择不慎，用户将自己的网络账号和密码随意转借他人或与别人共享等都会对网络安全带来威胁。

(2)人为的恶意攻击：这是计算机网络所面临的最大威胁。此类攻击又可以分为以下两种：一种是主动攻击，它以各种方式有选择地破坏信息的有效性和完整性；另一类是被

动攻击，它是在不影响网络正常工作的情况下，进行截获、窃取、破译以获得重要机密信息。这两种攻击均可对计算机网络造成极大的危害，并导致机密数据的泄漏。

(3)利用网络软件的漏洞和“后门”：网络软件不可能是百分之百的无缺陷和无漏洞的，然而，这些漏洞和缺陷恰恰是黑客进行攻击的首选目标。经常发生黑客攻入内部网络的事件，这些事件大部分就是因为安全措施不完善所招致的后果。另外，软件的“后门”都是软件公司的设计编程人员为了自便而设置的，一般不为外人所知，但一旦“后门”洞开，其造成的后果将不堪设想。

1.2.2 网络安全的对策

网络容易受到攻击的原因主要由以下一些原因：网络认证薄弱、操作系统容易受到监视、攻击者采用伪装的主机地址、局域网管理存在缺陷、网络管理员的配置错误使攻击者获得访问权、用户主机自身存在安全性问题。基于以上原因，可以采取下面几种方式来应对网络安全问题[4]：

(1) 采用物理安全策略。物理安全策略有：采取一定的措施来保护用户主机、网络管理设备、网络传输设备等免受自然和人为的破坏；对网络管理设备设置用户权限，使主机、服务器、网络设备等拥有一个良好的电磁兼容工作环境；建立一定的安全管理制度，防止人为破坏、剽窃电子设备或者重要信息。

(2)建立访问控制策略。通过建立用户名和密码的方式使网络系统免受非法用户的入侵。

(3)建立网络权限控制。通过对网络用户设立不同优先级的权限，减少针对网络的非法操作。管理员用户能够对网络进行管理和资源分配，而普通用户只能根据所分配的权限进行相应的操作

(4)建立目录级的安全控制。目录权限一般有八种：管理权限、读、写、创建权限、删除权限、修改权限、查找权限、存取权限。通过设立目录级的权限管理，能有效控制用户的资源访问，提高了网络和服务器的安全性。

(5)建立属性安全控制。管理员对计算机文件设置属性，使重要的文件数据不会被普通用户的误操作所修改。

(6)网络服务器的安全控制。通过对网络服务器的设置，实现了限制用户访问时间、对非法用户进行检测等功能。

(7)网络监控与锁定控制。管理员可以对网络进行监控。网络管理服务器则能够识别非法的访问，并能产生报警。还可以通过对非法用户的 IP 或者账号进行锁定的方式来保护网络的安全。

(8)对网络端口和节点的控制。网络服务器的端口采用回呼设备、静默调制解调器等加以保护，并以加密的方式来识别节点身份。

(9)采用防火墙对局域网进行保护。这种方式可以隔离内外网络、阻挡非法入侵，使网络处于保护屏障之中。目前，防火墙有三种：包过滤防火墙、代理防火墙、双穴主机防

火墙。

(10)采用信息加密技术。常采用链路加密、端点加密、节点加密的方式来保护网络传输数据包的安全。

通常采用多种安全策略相结合的方法来保护网络数据和信息的安全。目前，主要采用安全隔离防火墙的技术来保护网络安全，使网络免受非法入侵。

1.3 网络加速技术

随着网络用户和网络服务类型的迅速增长，导致网络中的数据流量激增。再加上网络带宽有限，用户数据报文在传输过程中会产生严重延时[6]。在特别情况下，还会产生严重的数据报文丢失现象。要改善目前的这种状况，只有让网络处理数据包能力提高。这就是网络加速所要研究的问题。

1.3.1 网络加速技术的研究背景

目前，随着国家主干网带宽的增加，一定程度上解决了网络拥塞的问题。但是，由于网络数据流量增长得太快，仅仅依靠带宽的增加显然不能从根本上满足网络传输高质量的需求。另外，大量重复报文和不需要的报文在网络上传输，进一步加剧了网络的拥塞。造成网络传输性能下降的因素主要包括以下几点：

(1)传输线路带宽的大小。这种问题的解决方法是：提高网络传输带宽、使带宽的利用率达到最佳。

(2)网络上大量的重复数据报文影响了网络的性能质量[7]。当一个局域网上的所有用户都需要下载局域网服务器中的文件时，同一份文件就要向所有的用户发送。另外，当服务器的文件更新时，即使更新很小，新的文件也会在网络中多次传输。一些数据的重复传输，导致网络带宽利用率降低，也加剧了网络的进一步恶化。

(3)一些端到端的应答机制和拥塞控制等策略，也会造成数据传输的延时。

(4)应用服务类型的增多，特别是在线视频传输等应用，使数据量急剧增加。

总之，网络数据流量的快速增加，使传统的计算机网络既无法满足高质量的传输要求，也阻碍了网络新应用的实施。因此，需要一种新的机制来解决数据流量增加造成网络服务质量下降的问题。目前，网络加速技术已成为研究的热门领域。

(5)网络节点处的管理设备无法快速对大量数据包进行快速处理，是造成网络质量下降的重要原因，。

1.3.2 网络加速常用技术

目前，网络加速技术主要有以下几种：

(1)使用高速缓存技术来解决带宽瓶颈和数据包延迟问题[5]。典型的方法是采用 web 文件缓存和数据字节缓存技术。另一种方法是采用动态缓存方式，将重复性较高的应用数据包以指针的方式缓存于设备中，当用户访问同样数据时，直接从缓存中存取数据包。

(2)采用内容分发网络的形式来加速网络数据报文的传输。这种方法常见于大型网站服务器。该方法将网站的内容发布到离用户最近的网络“边缘”，用户能就近快速的访问网站。

(3)采用专用设备对传输层和应用层进行优化。这种方法主要用来解决 TCP 协议造成的延迟问题。

(4)对网络数据报文进行压缩。通过这种方式可以使网络获得更大的带宽。

(5)使用 Qos 服务质量控制。该方法主要是利用带宽分配和 Qos 工具来减轻带宽的竞争。网络管理员可以对应用业务设定服务优先级，使特定应用类型的服务质量得到保证。

(6)采用硬件包过滤防火墙来对网络流量进行优化。该种方式是在网络层对 IP 报文进行拦截，只允许符合特定规则的 IP 数据包进入局域网。这种方式有效地减少了网络中无用数据报文的传输，提高了局域网的服务质量。

传统的方法大多采用软件的方式来实现网络的加速和优化。传统的软件加速技术都是在服务器的操作系统层面对数据报文采取相应的策略[24]，达到网络优化的目的。然而，随着网络带宽的高速增长，传统的方法已显得力不从心。一方面，完全采用软件方式进行策略控制，已经无法满足越来越高的速度要求。另一方面，随着流量的逐渐加大，网络节点处服务器的 CPU 已经不堪重负，严重影响了报文的处理和传输。因此，从硬件方面来寻求解决途径成为一个热门的研究方向。

1.3.3 硬件加速技术

基于硬件的网络加速技术是采用硬件来实现字符串的匹配[4]，以达到快速数据包处理的目的。这种方式使原本由服务器 CPU 和分析软件处理的工作下放到专用的硬件设备来执行，既加快了数据报文的分类处理速度，也使 CPU 从疲于应对高速数据的流量处理中解脱。硬件加速技术的实现方法主要有以下几种：

(1)采用基于自动机的方法，把数据包的关键字集合转化为正则表达式在硬件中实现。

(2)采用内容地址存储器来对关键字进行快速匹配的方式。该方法可以在单周期内实现相关条目的匹配，并返回匹配条目的地址。由于内容地址存储器具有速度快，可配置性强等优点，与 FPGA 相结合可以快速完成对数据包的处理。

(3)采用 Bloom-filter 的方式来对网络数据报文进行快速检测。这种方法利用位数组来表示一个集合，并快速判断所流经的数据报文是否属于这个集合。

(4)采用专用指令集处理器来实现对数据报文的快速处理。

本文的研究内容，就是采用 FPGA 与内容地址存储器相结合的方式来设计一种网络安全加速卡。

1.4 实现方式比较

原来的路由器、交换机等网络设备常采用通用的处理器来设计。可是随着网络带宽的飞速增长，通用处理器已经难以满足高带宽的网络处理要求。当前，普遍采用 ASIC、NP、

FPGA 来设计网络设备[49]。这三种处理器各有自己的优势。

(1)ASIC 技术(专用集成电路，application specific integrated circuit)是当前网络设备中使用最多的技术。该方式是用硬件电路来实现路由查找，报文分析等操作。当硬件电路做到成熟之后，就可以将其固化在一颗硬件芯片内来实现，可以获得更高的运行速率，这就是 ASIC 技术。采用 ASIC 技术可以使处理速度更快、性能更好。因此，这种技术方式满足了网络带宽不断扩容的需求。特别是大批量生产时，专用集成电路的成本会变得很低。然而，这种技术也有自身的缺陷：一旦逻辑电路固化到芯片后，不能对芯片再进行升级和修改，也不能添加新的功能逻辑，这使得专用集成电路的资源重用率非常低。另一方面，ASIC 的开发周期非常长、一次流片的费用很高，导致总的开发设计成本非常昂贵、产品面世的周期很长[49]。这些缺点在某种程度上制约了 ASIC 技术的在小批量产品中的应用。ASIC 技术一般使用于功能成熟并且大量应用的场合，比如路由器等。

(2)NP(网络处理器，Network processing)是一种专门用于网络领域的可编程器件，为适应网络的应用要求，该器件采用了专门的设计和优化来对芯片进行处理。NP 包含以下功能单元：RISC 处理核、协处理器、高速总线、大容量的片内存储器，高速接口等。NP 具有很多的优势：可灵活编程、能实现不同级别的并行处理、具有良好的扩展性(这时因为 NP 由软件构架和硬件构架共同构成)。另一方面，NP 也有自身的缺点：高速应用中需要大容积的存储带宽、价格昂贵、处理性能也比 ASIC 低。

(3) FPGA (现场可编程门阵列，Field Programmable Gate Array)是通用可编程逻辑阵列。FPGA 内部逻辑资源非常丰富，内部逻辑单元之间互连灵活。FPGA 的集成度比较高、执行速度快、可灵活编程且方便重新配置，具有非常大的灵活性。目前的 FPGA 针对网络应用推出了一系列的 IP core，并且将有些 IP core 集成到芯片内部，提高了网络处理能力和控制能力。FPGA 最大的优势就是提供了另两种芯片无法比拟的二次开发能力；并且可以通过工作组的协调设计来提高开发速度。FPGA 的缺点是：处理能力比同样逻辑的 ASIC 低很多、功耗比 ASIC 大很多、大批量的价格比 ASIC 高很多[49]。

由于本设计采用 FPGA 进行系统设计，保证了系统的灵活性和可升级性。后续的新功能方便在现有系统上添加。

1.5 本文研究内容和章节安排

本文通过对网络隔离技术的分析和研究，总结出网络安全隔离技术的关键点。通过对数据包分类技术的研究总结出一种基于硬件的规则匹配方法，通过该方法可以实现数据包的快速分类。基于硬件的数据包分类方式，有效地加快了数据包分类的速度、分担了 CPU 的工作量、提高了网络的服务质量。

本设计的重点放在：设计一种基于 FPGA 的网络安全加速卡，该卡采用硬件数据包分类和硬件转发的方式来分担 CPU 的数据分析和转发工作量。另外，通过对匹配条目的具体配置，和对网络数据报文的采样，能够有效地拦截有威胁的报文，具备网络安全功能。

本文各章节的安排如下：

第一章：对课题的研究背景做了系统的介绍，通过对 ASIC，NP，FPGA 的比较，选用 FPGA 作为系统的核心芯片。最后就本文的章节内容进行了整体的安排。

第二章：对网络安全隔离技术进行系统研究，详细介绍了物理隔离技术的原理，总结了网络隔离技术的关键点所在。

第三章：对数据包分类字段、TCAM 的硬件查找原理和规则管理算法做深入的研究。总结出一种合适的硬件规则管理方法。

第四章：对网络安全加速卡进行详细电路设计。包括了各个功能单元的原理和详细的硬件电路设计过程。最后，与传统加速软件的主要性能做具体比较。

第五章：对 PCB 板卡设计过程中遇到的问题做了详细的阐述，并对主要器件的功耗做了具体分析。

第六章：全文总结与展望。对全文进行了总结，分析了本设计的网络安全加速卡的性能优势。最后，提出了网络安全加速卡的不足之处以及后续进一步完善的方向。

第2章 网络隔离技术分析

2.1 网络隔离技术的发展过程

随着互联网的发展和普及，计算机网络在给人们的生活带来巨大便利的同时也带来了更多的网络信息安全方面的问题。在信息经济高速发达的今天，任何的重要信息资料的丢失或者因网络攻击导致的网络瘫痪都会给企业或个人带来诸多的损失与不便。因此，无论是个人还是企业都迫切地需要一款设备来有效保证网络的安全和网络畅通。而网络安全隔离设备可以通过内网与外网进行隔离的方式来减少和避免网络攻击[2]。这也是网络隔离技术的发展动力所在。

网络隔离技术就是把两个或者多个可路由的计算机网络(如 TCP/IP 协议栈的网络)通过不可路由的协议(如 IPX/SPX 等)进行数据交换达到网络隔离的目的。与传统的人为截断网络连接的方式不同，网络隔离技术的基本原理是采用不同的网络传输控制协议来实现网络的逻辑断开。因此，目前的网络安全隔离技术实际上是采用协议隔离技术来隔离网络之间的直接联系。

隔离技术经过了五个不同的发展阶段[4]，每一个阶段的发展都带来了性能上的实质性提升。下面对这五个阶段的发展做一个简要的介绍。

第一个阶段：采用完全隔离的方式。这种方式是将内部专用网络与外部网络完全逻辑断开。这种方式虽说可以完全避免来自外部网络的数据包攻击，但是将内部专用网络与外部互联网完全隔离，让内部网络完全断开了与外部网络的联系。当内网和外网之间要实现数据的交换时，必须采用两套以上的网络和操作系统来完成交换过程，使建设成本提高、也给网络维修和网络数据交换带来诸多不便。该方式不利于网络的互通也违背了网络互连的初衷。

第二个阶段：采用硬件卡隔离。通过在客户端增加一个硬件卡的方式，使客户端的硬盘或者存储设备，通过一张硬件隔离卡连接到主机主板上。用户在选择不同的硬盘或者存储设备时就选择隔离卡上相应的网络接口，通过该卡可以控制客户端的硬盘等存储设备连接到不同的网络。

第三个阶段：采用数据转播隔离的方式。该方式采用时分复用文件的方式来实现传播系统的有效隔离。该方式与上一代的卡片式隔离方式相比具有更多的灵活性，可以通过编程的方式实现灵活的时分复用。

第四个阶段：采用空气开关隔离的方式。该方式采用一种电气驱动的双掷开关的方式使内部专用网和外部网络分时访问缓存器来完成数据交换。这种双掷开关可以采用控制器件控制继电器 relay 的方式来实现，简单易行，延时主要是 relay 的开关延时，实时性较好。

第五个阶段：采用安全通道隔离。隔离方式采用专用硬件设备与专有的网络安全协议相结合的安全机制，有效的将内外网络进行逻辑隔离。该方式将一些转发和协议分析等工

作由控制器交给专用硬件设备来完成，显著提高了数据安全交换的效率，同时透明支持多种网络应用的实施。是新一代网络安全隔离设备的优势所在和进一步完善的方向。

2.2 网络隔离原理

隔离技术是要把带有恶意攻击代码的攻击数据包和一切不相关的数据包进行隔离，保证内部专用网络和外部不可信网络之间的数据安全交换。通常所说的隔离是指物理隔离技术，即包括网络隔离和数据隔离[58]。一般采用协议隔离的技术来实现。

一般来说，外网是安全性不高的因特网，是不可信网络；内部专用网络是安全性很高的网络，是可信网络。要保证内部专用网的绝对安全，不受到来自外网的数据包攻击，必须保证外部网络和内部专用网在逻辑上是完全断开的。内网和外网之间的隔离设备通常是由存储介质和一个控制调度电路组成。网络隔离设备充当内网和外网通信的桥梁作用。

当外网有数据要向内网发送时，外服务器就会向隔离设备的控制部分发起非 TCP/IP 协议的数据连接请求。链接建立后，网络安全隔离设备接收来自外网的数据流，并对接收到的所有数据包进行协议剥离(包括 TCP/IP 协议栈的协议剥离和应用层协议的剥离)，最终只保留原始应用数据，最后控制原始数据写入存储介质当中。在整个过程中还必须对数据流做安全性和完整性的检查，防止恶意代码等有害数据包写入存储介质当中，保证数据流的安全性。外网数据包在网络安全隔离设备的控制下写入存储介质的过程如图 2.1 所示：

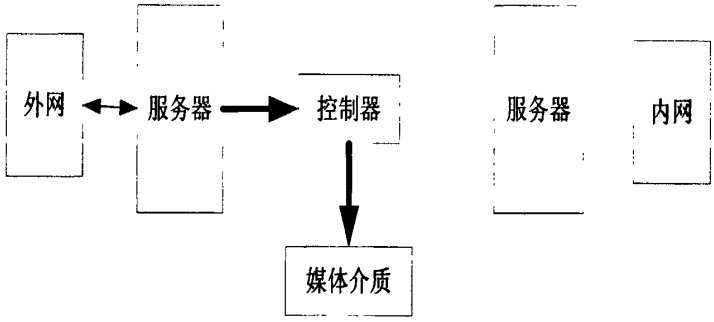


图 2.1 外网数据包写入存储介质

数据写入媒体介质以后，控制器随即断开与外网的连接，同时发起与内部专用网络的非 TCP/IP 协议的数据连接。内网与隔离设备建立连接以后，内网从媒体介质中获取所需数据。内网服务器通过 TCP/IP 协议和应用协议对获取的原始数据报文进行重新封装，封装完毕后传送到应用系统做相应的应用处理。这样内网就安全成功的接受到了来自外网的所需数据流。内网从存储介质中获取数据流并提交给应用系统的过程如图 2.2 所示：

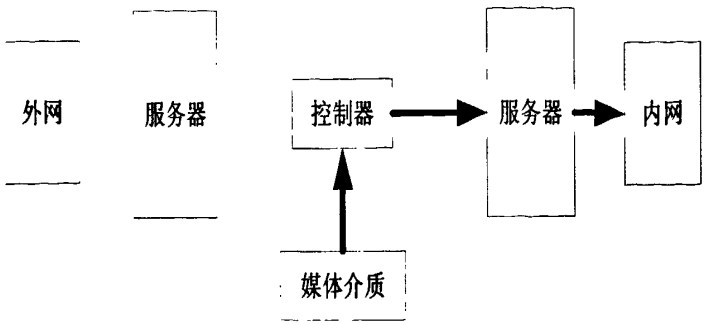


图 2.2 内网从存储介质获取数据流

内网接收到完整的数据流之后，会向控制器发送接收完毕指令。控制器收到接受完毕指令后，立即断开隔离设备与内网服务器的连接。到此，完成了一个完整的从外网接收一段完整数据流并比较给应用系统的全过程。

同样，当内网需要向外网发送数据流时(比如电子邮件应用数据报文等)，内网服务器就会向隔离设备发出连接请求。隔离设备收到连接请求指令后，响应连接请求并通过内网服务器与内部专用网络之间建立起非 TCP/IP 协议的数据连接通道。网络安全隔离设备与内网建立连接后就能收到来自内部网络的数据流。从收到内网数据流的首帧开始，网络隔离设备会对所有接收到的数据包的 TCP/IP 协议头部和应用协议头部进行剥离[58]，最终只剩下原始应用数据。这时，控制器件将得到的原始数据流写入到媒体介质。内部专用网络数据报文写入存储介质的过程如图 2.3 所示：

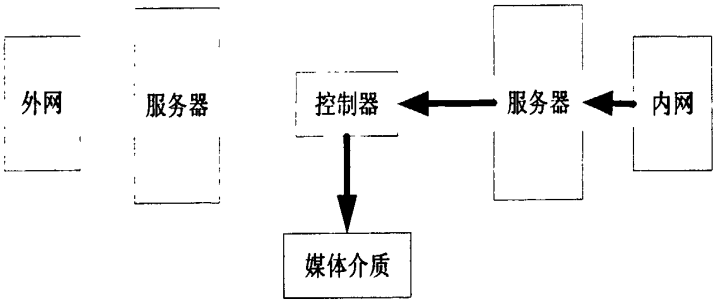


图 2.3 内网数据流写入存储介质

一旦数据向媒体存储介质写入完毕，隔离设备立即中断与内网的逻辑连接。随即向外网发起数据连接请求(该连接同样是非 TCP/IP 协议的连接)。外网服务器接到并响应网络安全隔离设备的连接指令后，隔离设备与外网的数据连接通道随即建立。这时，控制器控制媒体介质中的数据流向外网发送。外网服务器收到数据后，随即对接收到的数据报文进行 TCP/IP 协议和应用协议的封装，将封装好的帧向外网系统发送。隔离设备与外部网络建立连接并发送数据流到外网的过程如图 2.4 所示：

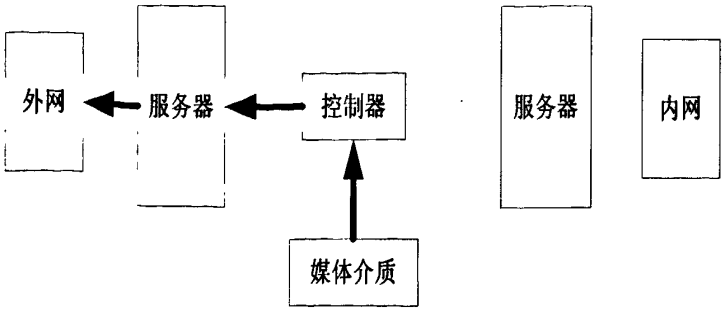


图 2.4 隔离设备向外网发送数据流

当数据发送完毕后，网络安全隔离设备立即断开与外网的连接。这时，内网向外网发起的一次数据流传送工作宣告完成，隔离设备与外网和内部专用网都不再有数据通道的连接，恢复完全隔离的初始状态。

基于 TCP/IP 协议栈的封包过程如图 2.5 所示：

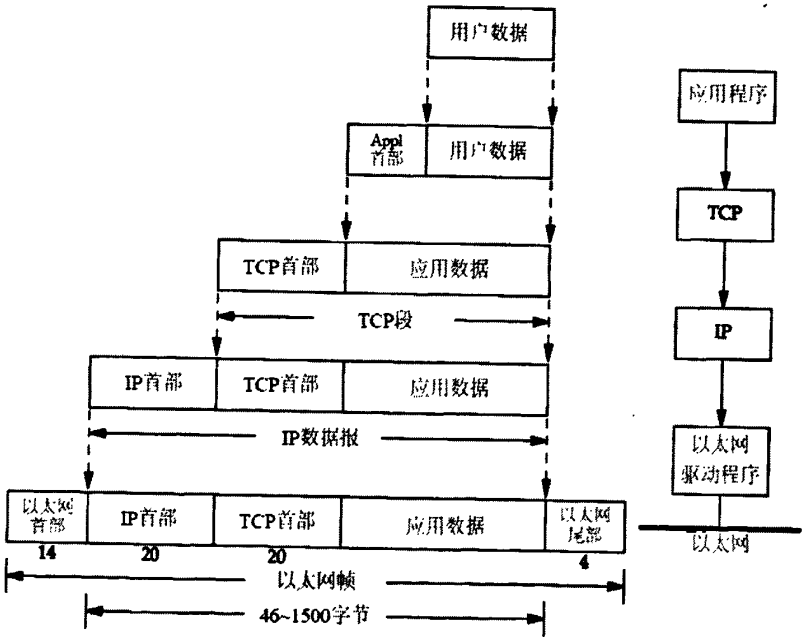


图 2.5 数据进入协议栈封包的过程

外网和内网之间的一次完整数据交换经历了接收，存储和转发这三个步骤。根据计算机体系结构原理，数据的交换过程是由服务器的内核进行处理并调度内存来完成的，整个过程的延时比较小，速度比较快。由于在数据交换的过程中，外网和内网之间并没有直接相连，所有的数据交换都是通过中间的网络安全隔离设备来完成，在很大程度上可以有效避免内部专用网受到外部网络的恶意代码等攻击数据包的入侵。

2.3 隔离技术的要求

网络隔离技术的隔离原理是将内部专用网络和外部互联网络做一个逻辑通道的隔离。要实现一个完整的隔离系统必须注意以下一些要点：

首先，网络安全隔离设备必须保证设备自身的安全性。要保证隔离设备自身的安全，除了采用操作系统加固优化和使用安全的操作系统等常用办法之外，还要在内部专用网络和外部网络的接口使用不同的操作系统，实现两个网络之间的接口分离。只有让两个不同的操作系统分别控制不同的内网接口和外网接口，并在网络安全隔离设备中采用非TCP/IP的协议进行通道连接和数据交换，才能保证设备本身的安全性。

其次，要保证内外网络之间是绝对的逻辑隔离。即任何一个网络的数据包都无法直接进入本部门的专用内部网络。只有通过剥离了相关的协议头部和应用协议头部之后，剩下的原始数据流经过再次重新封装，组成新的帧格式后才能进入对方网络。

第三，要保证网络之间交换的数据只有应用数据，不包含任何的网络层攻击数据包。可以采用命中匹配安全规则和采用内容分类技术实现协议的分析和原始应用数据的提取。网络安全隔离设备在交换原始数据时还要具有拦截网络攻击数据包的能力。一般可以通过内容解析和命中规则等方法来有效解决。

第四，网络之间的互访要有完善的控制和检查机制。要确保所交换的数据流可信可控且安全，需要采取相关的技术对数据流进行完善的控制和检查。

第五，要保证网络之间数据通信的畅通。网络交换产品必须具有极高的线速处理的能力，否则，很可能成为网络数据交换的速率瓶颈。

总之，网络隔离系统要能对网络交换数据流进行有效控制，在不可路由的协议下完成网络之间的数据流交换。同时，网络安全隔离设备必须对网络应用层的多种应用实现透明支持，实现在有安全保障和高带宽要求的情况下进行数据流的高速交换。

2.4 隔离技术的发展方向

网络隔离技术进过了长时间的发展，市场需求不再只是追求安全隔离的单一性能。新的市场需求希望网络隔离设备不仅能够解决数据包处理的速度瓶颈，并且具有先进安全设计思路。在当前的这种背景下，第五代网络隔离技术应运而生，并且不断完善。第五代网络隔离技术的出现，明显的提升了产品的安全性，显著提高了数据包处理速率，并且具有更加优秀的隔离防护手段[58]。

第五代网络隔离技术的实现原理是通过专用通信设备、专有安全协议、加密验证机制、配合应用层数据提取和鉴别认证技术，进行不同安全级别网络之间的数据交换。网络隔离设备通过与网络之间的非TCP/IP协议建立连接，同时对双方网络通信的内容及过程实施严格的身份认证和内容过滤等多重安全机制，保证了网络之间数据通信的安全可控性和可操作性，提高了网络产品的网络安全性能。这是网络安全隔离产品的重要发展方向。

本硬件加速系统在部分吸收新一代网络安全技术的基础上，采用硬件进行快速协议分析，用硬件实现快速数据包过滤等方法显著地解决了网络管理设备的带宽处理问题，同时有效地保护了内专用网络的安全性。

2.5 本章小结

本章阐述了网络隔离设备的五个发展阶段，并对主流的物理隔离原理进行了深入的分析，总结了网络隔离设备的性能要点和设计要求。传统物理隔离设备的缺陷推动了新一代网络安全隔离设备的诞生。文章最后就新一代网络安全隔离设备的性能要求和多个发展方向进行了系统的阐述。

本设计介绍的网络安全加速卡的设计将采用硬件实现快速协议分析、快速命中规则匹配、快速数据包过滤和快速数据包转发。系统设计在充分吸取了新一代网络安全隔离设备的安全机制的同时，将原本由服务器处理的一些转发和协议分析功能下放到网络安全加速卡来处理。这种方式将改变服务器既要承受超高速报文的转发和报文协议分析，又要进行应用系统的运行的格局。应用了网络安全隔离新技术的硬件加速卡可以实现高速报文协议分析和高速数据包转发，服务器运行网络应用系统和管理系统，在保证安全性的同时有效的改善了网络管理设备节点的处理速度。

第3章 数据包分类技术研究

3.1 数据包分类技术概述

用一定的分类匹配规则，对数据包进行处理的方法，称为数据包分类技术。这些规则通常是数据包协议首部的某个字段或者一些字段的组合[1]。在区分服务，虚拟专用网络，基于策略的路由技术，流量统计等网络应用方面，数据包分类技术得到了广泛应用。

数据包首先经过协议解析和字段抽取，然后根据协议类型和分类字段进行数据包的分类工作[23, 24]。本节从网络的层次结构入手，分析每层常用的分类字段，以此对数据包分类做一个系统的描述。

在数据链路层，源 MAC 地址和目的 MAC 地址可以作为分类字段。因为该地址指示了数据包从何处发出，要发往何处。另外，对于不同类型的网络，还可以在数据链路层使用不同的关键字段，如：地址解析协议等。

在网络层，源 IP 地址，目的 IP 地址可以作为分类字段。同样，运输层使用的是 TCP 协议还是 UDP 协议，可以通过对协议类型的分析来确定。因此，协议类型也可以作为分类的字段。

在运输层，源端口号和目的端口号可以作为分类字段。因为他们确定了上层应用程序的端到端的端口号。此处以 TCP 协议为例，对数据包分类做一个简要的描述。TCP 首部中的 URG(紧急)，ACK(确认)，PSH(推送)，RST(复位)，SYN(同步)，FIN(终止)这些标识字段都可以用来对数据进行分类。TCP 包首部结构如表 3.1 所示：

表 3.1 TCP 包首部结构

<table>
<tr><td colspan="9">16位源端口</td><td colspan="8">16位目的端口</td></tr>
<tr><td colspan="17">32位序号</td></tr>
<tr><td colspan="17">32位确认序号</td></tr>
<tr><td colspan="4">4位首部长度</td><td colspan="4">保留6位</td><td>URG</td><td>ACK</td><td>PSH</td><td>RST</td><td>SYN</td><td>FIN</td><td colspan="4">16位窗口大小</td></tr>
<tr><td colspan="9">16位检验和</td><td colspan="8">16位紧急指针</td></tr>
<tr><td colspan="17">选项</td></tr>
<tr><td colspan="17">数据</td></tr>
</table>

在应用层，由于使用的协议种类非常繁多，常用 URL 作为分类字段。URL 由协议类型，主机名，路径和文件名组成。这些组成因子包含了文件的类型、目录结构和服务站点等信息，可以用于分类的依据。

目前，大多采用五元组的分类组合来对数据包进行五维分类。五元组包括：源 IP 地址，目的 IP 地址，源端口号，目的端口号，协议类型。这种方式常碰到同时匹配多个规则的情况，这个时候需要一种机制能够保证数据包匹配最佳的规则条目。本章重点讨论利

用 CAM 实现最佳规则匹配的原理，并深入研究规则管理的最佳方法。

3.2 基于 TCAM 的硬件规则匹配原理

分类查找是实现数据报文分类的核心步骤。目前，分类查找已有多种实现方式。大体上来说，基于软件的分类算法有：基于查找树的数据结构算法、几何算法、启发式算法和并行软件搜索算法等。随着查找维数的增加，这些算法各有优劣。单纯从数据结构的角度来说，这些算法要么具有比较高的时间复杂度，要么具有比较高的空间复杂度，要么具有非常差的更新复杂度。而基于硬件的内容地址存储器，采用并行的方式对利用硬件存储的规则进行匹配，因此具有最快的效率。本节将就CAM硬件算法的查找原理做深入的探讨和总结。

CAM是一种专用的内容地址存储器件，它能够快速对存储在CAM中的大量数据进行并行的查找工作[25, 26]。和其他的数字存储器一样，CAM的每一个存储位存储0或者1。目前，市场上大多的新型CAM叫做三态内容地址存储器简称TCAM，TCAM的每个存储为存储0或者1或者为不定态X，这种方式会使应用更加合理。TCAM的数据存储在数据存储位，而与之对应的真实值却是由掩码位来决定的。数据存储位和掩码位的“与”逻辑，就是该位数据的实际值。当掩码为“1”时，数据存储位的值就是该位的实际存储状态；当掩码为“0”时，不管数据存储位为何值，存储位的实际状态都是不定态X。

TCAM中实际存储的数据称为一个条目，一个条目由多位数据存储位和多位掩码位组成。每一个条目对应自己的存储地址和标志位，标志位指定了该条目的有效与否，便于对条目进行管理。只要是有效的条目，查找关键字都会与之进行比对，并且对所有条目的比对过程是并行的。当关键字的数据位与一个条目中的所有掩码位为“1”的数据存储位都相同时，才表示该关键字与该条目相匹配。匹配之后，TCAM随即返回该条目的地址。如果有任何一个数据位不相同都会返回不匹配的结果。匹配了一个条目，就相当于命中了一条规则。如果查找关键字与多个条目都能匹配，则返回所有地址中最低的一个地址结果。这是因为在CAM存储器中，条目的优先级是从低地址到高地址逐渐降低。

TCAM器件采用了流水线结构，这样使多个关键字在匹配过程中采用流水操作来保证每个时钟周期内都能进行一次匹配。同时，TCAM采用了增多引脚的方式使一个周期能输入一个关键字，并采用DDR技术使查找速度进一步加快。基于这些特性，器件在一个时钟的上升沿或者一个时钟的下降沿就能锁存一个关键字并执行一次匹配。

目前，IDT等公司都有非常丰富的TCAM产品，内存一般有：2Mbit、9Mbit、18Mbit、36Mbit等规格。查找的带宽有：36bit，72bit，144bit，288bit，576bit等。表3.2对TCAM常用的操作了详细的描述。

通过对CAM的深入分析，CAM有如下优良性能：用硬件电路的方式并行查找；支持DDR操作和Pipelining操作极大地提高了每周期的操作速度；硬件算法优势明显等。同时，CAM也有自身的缺陷：集成度比较低，价格昂贵等[39]。由于本设计的应用场合所需的规则

数有限，因此采用CAM来进行查找匹配不仅能满足应用要求，还显著了提高了执行速率。

表3.2 TCAM常用操作表

操作	描述
Read	读一个条目的数据位、掩码位，内部寄存器或外部SRAM
Write	写一个条目的数据位、掩码位，内部寄存器或外部SRAM
Lookup	查找有效条目，根据匹配情况返回条目地址或者不匹配标志
Learn	寻找空闲条目，并将新条目写入空闲条目中
Move	将一个条目从源地址移动到目的地址，目的地址条目的有效状态与源地址条目一致。移动结束后，源条目变成无效
Copy	将一个条目从源地址移动到目的地址，目的地址条目的有效状态与源地址条目一致
NFA Lookup	寻找下一条空闲条目，该条目的地址保存在一个NFA(Next Free Address)寄存器中

3.3 硬件的最佳规则匹配与管理

用TCAM进行数据分类时，需要不断的将新的规则写入空闲规则中，以满足新的规则匹配要求。如果一个条目写入TCAM之后，不能再对该条目进行更新或者其他操作，则TCAM必然会因耗尽空闲条目导致新的条目无法写入TCAM规则表中。另一方面，所有的规则都应该具有一个生命周期。当一条网络服务结束时，有些规则可能不会再使用。因此，需要一种机制把不用的规则条目删除，从而增加新的规则条目就有写入的地址空间，以保证新的规则匹配能够继续运行。基于上述原因，需要一种机制对TCAM中的匹配规则进行灵活的管理。通常采用以下几种方式对匹配规则进行管理[27, 28, 29]：

(1)通过简单的单向队列操作的管理规则。

(2)通过集中管理的方式的管理规则。

(3)采用分布式的并行方式的管理规则。

第一种规则管理方式是一种最简单的管理方式。当系统上电，FPGA内的数据报文分类器就会将TCAM中的所有空闲条目地址推入一个单向队列。当有新的规则需要写入时，就会将队列首部的空闲条目地址推出队列，将新的规则条目写入该空闲条目地址。当有不会再用的规则条目需要删除时，将需要删除的规则条目地址推入单向队列，该地址能继续给新的规则条目使用，这样就实现了一条规则条目的删除。这种队列式的规则管理方法，添加和删除操作的时间复杂度都是“1”，操作方便，速度快。

如果本设计采用基于队列的规则管理方式，则会经历以下一些过程。设定数据包分类采用协议类型(PT)、源端口号(SP)、目的端口号(DP)、源IP地址(SIP)、目的IP地址(DIP)进行五维分类。由于所有规则的生命周期各不相同，所以在匹配的过程中，不用的规则将会被不断的删除，空闲的条目也会被不断地回收并重新分配。假设TCAM在现时刻拥有如下六条规则，如表3.3所示。从上到下表示条目的地址不断增大。

表 3.3　TCAM 现时刻的存储规则

0	PT				
1			DIP		
2		SIP		SP	
3	PT		DIP		DP
4		SIP	DIP		
5	PT	SIP	DIP	SP	DP

由上面的六条规则很容易看出：第 0 条目的范围大于第 3 条目的范围，第 3 条目的范围大于第 5 条目的范围；第 2 条目的范围与第 4 条目的范围不具备可比性，这两条目的范围都大于第 5 条目的范围。即有些条目是另外一些条目的一个子集。基于 CIDR 的路由匹配规则是将最长的字符串匹配路由作为最佳路径选择，因为最长字符串匹配的路由最具体。因此，当查找的关键字匹配了最具体的条目、最小子集的条目、或者具有最高优先级别的条目时，即完成了最佳规则匹配。

三态内容地址存储器中的条目存储的地址越低，其对应的优先级就越高[25, 26]。因此，当有多个条目与该关键字同时匹配时，返回的结果是所有匹配条目中地址最低的地址值。例如当关键字和条目 2 和 5 同时匹配时，则返回条目 2 的地址。但是条目 5 显然比条目 2 具有更具体的匹配度，却不能得到应有的服务操作。这种简单的采用单向队列的规则管理方式具有非常严重的弊端，只能适用非常简单的匹配系统(多为单维匹配)。

本设计的数据包分类使用的是多维分类方式。只有根据匹配的条目的具体程度、重要性等对条目分别设定不同的优先级，才能实现最佳规则匹配。将分类好的优先级根据优先级的高低从低地址向高地址依次存储。低地址存储的都是高优先级的条目，高地址则存储的是低优先级的条目，TCAM 中的条目才能和优先级一一对应。当有关键字再次匹配多个条目时，TCAM 返回的最低地址就是最高优先级的条目地址。这样才能保证匹配的最佳性。

通过集中的方式对规则进行管理则是：所有分类的有效规则集中写入 TCAM 的低地址，空闲规则全部处于高地址端。传统的规则管理多采用集中的规则管理方式[46] [47]。假设下表 3.4 表示现时刻一个集中式管理过程中的规则存储表。表中用有效位为“1”和“0”分别表示该条目是有效条目还是空闲条目，用“1”的个数表示规则内容的具体程度。

表 3.4 中，存放有多条规则，有效的规则都存在表的顶端(低地址)，无效的空闲条目都存在表的底端。假设系统在当前状态下有一条优先级为 8 的规则需要写入 TCAM，则需要将条目 3、条目 4、条目 5 依次向下移动一个条目，空出地址 2 让新的条目写入。只有这样才能使所有的规则都按照优先级排列，保证匹配的结果是最佳匹配。在这种方式下，只要有新的条目要写入 TCAM，就要将写入地址之后的所有有效规则依次下移一个条目。反之，只要有旧的条目需要删除，则需要将该条目之后的所有条目逐一上移一个条目。

表 3.4 现时刻规则存储表

顺序	地址	规则内容	有效位	优先级
1	0	111111	1	优先级 9
2	1	111111	1	
3	2	111	1	优先级 6
4	3	11	1	优先级 3
5	4	11	1	
6	5		0	空闲条目
7	6		0	
…	…		0	

由于相同优先级的规则存储顺序的先后不会有影响，依次可以对上述的集中算法进一步改进。假设现时刻一条优先级为 10 的条目要写入 TCM。由于优先级为 10 的条目比现时刻任何一条目的优先级都要高，因此必须写入表的顶端。此时，第 4 条目移动到地址 5，第 3 条目移动到地址 3,，第 1 条目移动到地址 2，将新的条目写入到地址 0。通过这种方法，不需要移动所有的条目，减小了算法的复杂度。

而采用分布式的并行规则管理方式则在此基础上做了进一步的改进。这种方法首先将所有条目按照不同的优先顺序分组，空闲条目被“分布”到各规则组中。系统初始化时，用户就可以随意的将空闲条目分配到各个组中。每个规则组的数据结构包括：组的硬件基地址、有效的条目数和空闲条目数。数据结构如下：

```
Struct ruleclass {
  int BaseAdder;      //组的基地址
  int ValidRule;      //有效的规则数
  int FreeEntry;      //空闲的条目数
};
```

表 3.5 是采用分布式进行规则管理的条目存储图。图中将各条有效位为“0”的空闲规则分布存储在不同的规则组中。

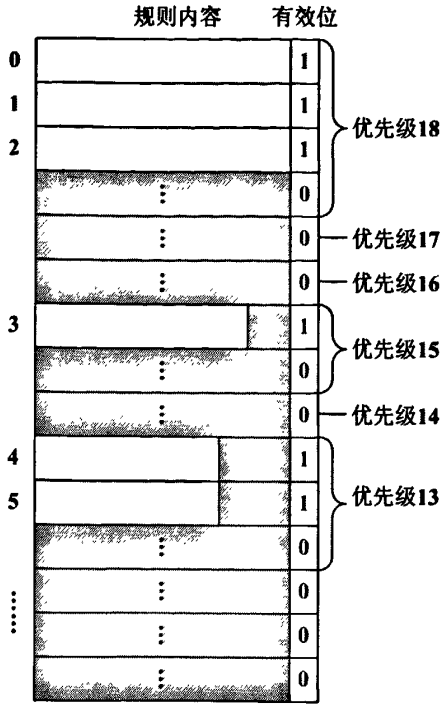
如果系统要添加一条新的优先级为 17 的规则 A，会按如下的过程来进行：

(1)检查优先级对应的规则组 17。如果该规则组的空闲条目数不是 0，则将规则 A 写入第一条空闲条目中。如果该规则组中的空闲条目数为 0，则执行第 2 个步骤。

(2)检查规则组 16 的空闲条目数是否为 0。如果不为 0，则将该规则组的第一条规则执行 move 操作写入到该组的第一条空闲条目中，并将新规则存储到规则组 16 的第一个条目中，并将优先级改为 17。如果检查结果为 0，则会继续检查下一个优先级的规则组。依次类推。如果向下遍历了所有的规则组都未发现空闲条目，则执行步骤 3。

(3)向上递归借用空闲条目(如规则组 18，规则组 19 等依次类推)。

表 3.5　分布式的规则存储表



采用分布式的并行规则管理方式，既避免了简单的单向队列管理方式不支持最佳规则匹配的缺陷，又比集中方式的规则管理执行速度更快。这种方式能够在保证比较快的执行速度的同时，使所有的应用都能得到相应的服务。

3.4　本章小结

本章从计算机网络结构入手，对从数据链路层到应用层的各个层面相对应的分类字段进行介绍，对数据包分类技术做了系统描述。接着对硬件查找芯片 CAM 及其匹配原理做了详细的分析。同时将 CIDR 路由最佳匹配方法推广到 TCAM，总结了最佳规则匹配的要点：匹配最具体条目、匹配最小子集条目和最高优先级条目。最后就命中规则管理的三种算法做了详细的介绍，总结了基于分布式规则管理方式的优点。

第4章 系统整体设计

网络安全加速卡硬件平台的设计，是在充分吸收了网络安全隔离新技术和硬件规则匹配技术等基础上完成。通过专用通信设备、自定义协议格式、并配合应用层数据提取，满足了新一代网络隔离技术的需求。采用硬件规则匹配技术加快了数据处理速度。本章对网络安全加速卡的整体设计和各个功能模块的硬件电路设计做详细的介绍。

4.1 总体设计框架

该网络安全加速卡实现的基本功能如下：

(1)支持 4-8 个 GE 网口，包括 RJ45 接口和光纤接口。

(2)本系统的 PCIE 接口可以通过金手指直接与主机或者服务器的 CPU 主板相连。CPU 根据首包的信息、安全策略和路由信息等通过 PCIE 总线向 FPGA 核心下达处理表，FPGA 根据处理表的匹配规则来执行转发等功能。FPGA 根据匹配的结果可以通过 DMA(直接存储器访问)的方式经 PCIE 总线将所需数据直接送至服务器，实现线速采样。

(3)本设计具有最大支持小包 4.2Gb/S，大包 6.2Gb/S 的报文分析处理能力

(4)本设计使用了三态内容地址存储器（TCAM）与 FPGA 相配合，通过硬件查找的方式实现了高性能的内容匹配处理工作，极大的提高了报文规则匹配速度。

(5)本设计通过五元组(SIP，DIP，SP，DP，PT)精确匹配的数据流转发方式来处理网络数据报文。

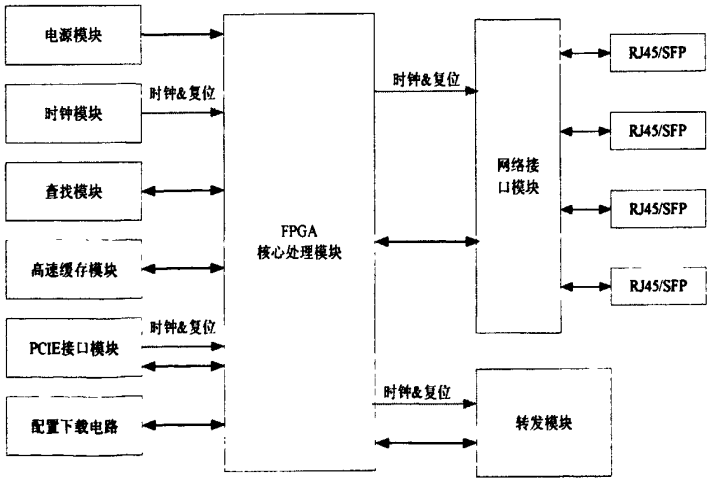


图 4.1 总体设计框架

根据总体功能需求，本系统硬件设计包括如下模块：网络接口模块、硬件查找模块、数据包转发模块、PCIE 总线接口模块、电源网络模块、时钟分配模块、FPGA 核心处理模块和掉电保护电路。系统设计涉及到物理层、数据链路层、网络层和运输层等多个层面，

配合软件程序可以实现对数据包的层层解析。以 4 个网络接口为例，系统的总体设计框图如图 4.1 所示。

后续将对各个功能部分的设计做详细地分析和介绍。

4.2 千兆网络接口模块设计

网络接口模块是外部网络和内部核心处理单元的桥梁。它的设计包括 PHY 设计和 MAC 接口设计[17]。

网络安全加速卡支持 RJ45 接口和 SFP 接口，单个接口速率达到 1000Mbps，其中 RJ45 接口还能向下兼容 10Mbps 和 100Mbps 网络速率。根据网络接口模块的性能需求，PHY 芯片采用 Marvell 公司的 88E1145 芯片。

88E1145 具有完全适合网络安全加速卡的工作性能：

(1)使用 RJ45 接口设计时，支持双绞线接入的速率从 10M 到 1000M。

(2)支持的 MAC 接口包括：GMII/MII，TBI，RGMII，RTBI，SGMII 等灵活的 MAC 接口方式.

(3) PHY 内部设计了 1.25GHZ 的 SERDES 串并转换高速收发器，支持 1000M 光纤接入设计[63]。

(4) PHY 内部专门设计了 PECL 电平接口，能满足 100M 光纤的设计方案。可以兼容 100M 和 1000M 的光模块接入口。

88E1145 集成了 4 端口的以太网收发器，支持从 10M 到 1000M 的速率要求。每个端口都具有相同的功能却能独立工作。他们内部设计了独立的 MDC/MDIO 接口。通过该接口可以对 PHY 进行灵活的逻辑地址设置和工作方式的设置。可以灵活配置 4 个端口是并行工作或由一个 MDC 串行控制。图 4.2 以其中的一个端口为例，展示了端口内部的逻辑功能框图。

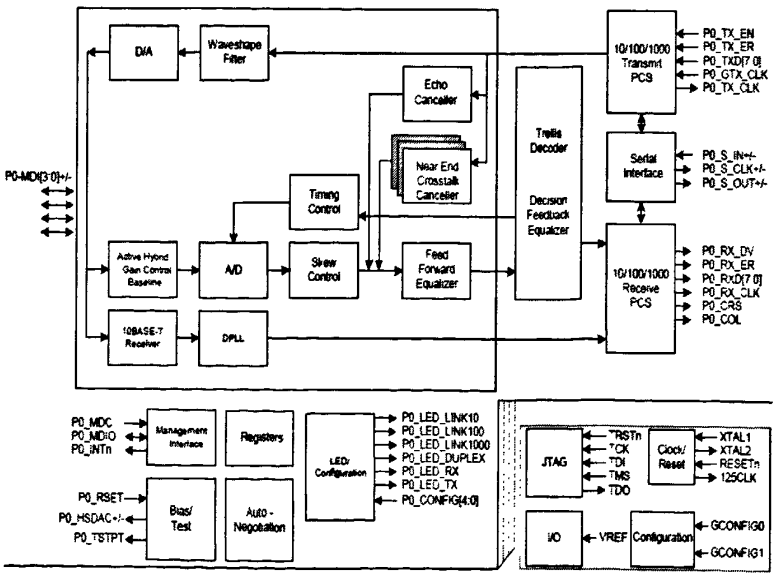


图 4.2　端口功能示意图

芯片同时支持光纤接入和 RJ45 铜接口接入。在两种接入方式下系统所能支持的网络速率等级和对芯片的设计都有所不同。当用 RJ45 铜接口设计接入时，支持的速率包括十兆、百兆、千兆兼容，支持的 MAC 接口包括：GMII/MII、TBI、RGMII、RTBI、SGMII 和串行接口。RJ45 铜口接入方式如图 4.3 所示：

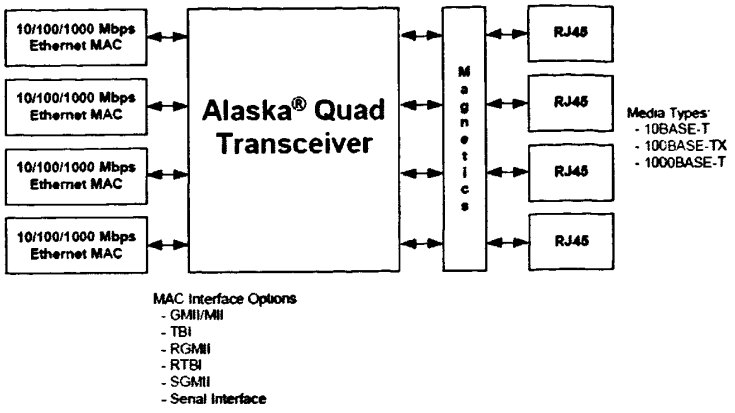


图 4.3 RJ45 接入方式设计图

在 RJ45 铜接口设计模式下，使用铜缆或者双绞线接入。设计时将 P[3:0]和 MDI[3:0]连接到物理媒介，其中 MDI 引脚需要端接 100 欧姆的差分电阻通过网络变压器连接 RJ45(本文不对芯片的具体引脚作描述)。本设计的网络变压器采用 S558-5999-Q2F 设计。网络变压器的电路设计如图 4.4 所示，以 P0 口为例的端接电阻的电路设计如图 4.5 所示。其中端接电阻的设计很好地抑制了高速信号的反射，经测试达到了很好的耦合标准，增强了系统的整体精度和性能。

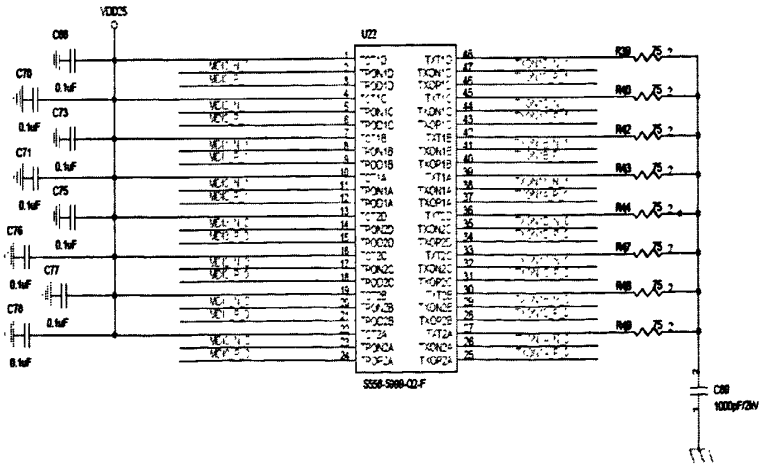


图 4.4 网络变压器电路设计

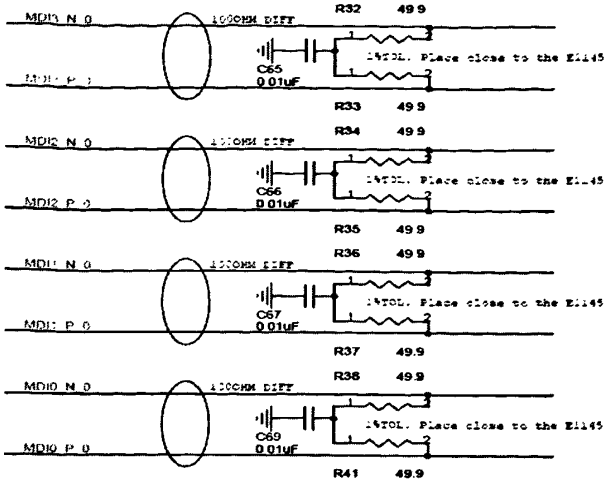


图 4.5 网络接口端接电路设计

在 SFP 光接口设计模式下，使用光纤接入。SFP 光接口设计模式下的接入方式如图 4.6 所示：

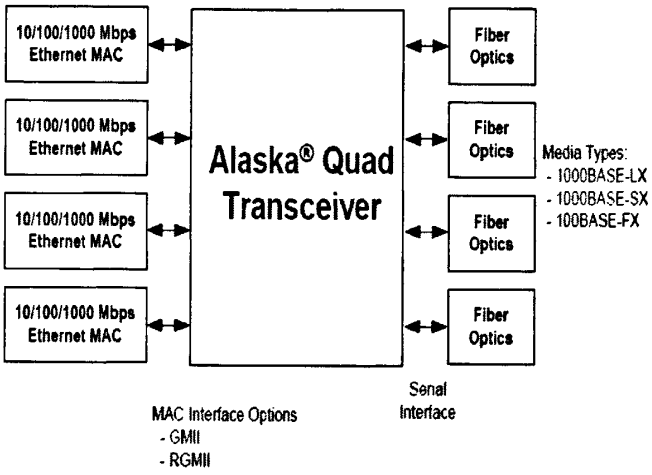


图 4.6 SFP 光接口接入方式设计图

在光纤接入时，不再需要网络变压器件，因为光模块已经对信号进行了光电转换。光模块通过串行接口(PHY 中的 S_IN,S_OUT，SD 引脚)与 PHY 芯片相连。PHY 与 FPGA 之间的 MAC 接口为 GMII 接口或者 RGMII 接口。数据的快速收发则通过内部的 SERDES 电路来完成。该芯片还对 SERDES 电路内置了端接设计，这样简化了外部电路的设计而且为板卡的设计降低了难度。内部还设计了 CML 电路来作为 PHY 的 I/O 缓冲。

经过以上的分析，要同时在一个芯片中支持 RJ45 的铜接口输入和 SFP 的光接口输入模式，则必须将 MAC 接口设计成 RGMII 接口。本网络安全加速卡的 MAC 接口采用 RGMII 的设计模式。RGMII 接口将 PHY 和 MAC 之间的引脚数减少到 12 个，还有外部连线少的优点。这种 MAC 接口在 RJ45 接入情况下，支持 10M/100M/ 100M 的模式[53]。当接入千兆网络时，启动 125M 时钟上下沿采样，四路数据线(收发各四路，分别是 TXD[3:0]和 RXD[3:0])即可实现千兆速率的收发。同样的，当接入 100M 网络时，芯片时钟自动减小

到25M四路数据线刚好实现百兆的收发带宽。接入10M时，芯片时钟则自动减小到2.5M。其中TXC信号由MAC产生，RXC信号由PHY产生，这两个时钟信号分别作为发送和接收数据的时钟信号。图4.7所示为RGMII接口模式下的MAC外部电路原理图：

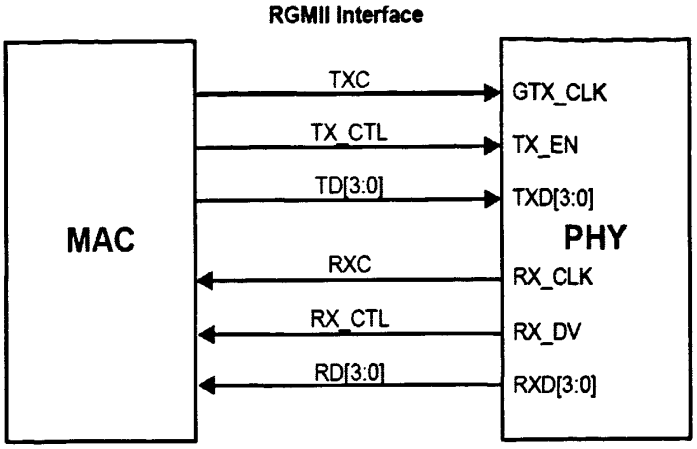


图4.7 RGMII接口图

在确定MAC接口方式之后，还必须对相应的硬件寄存器进行设置来实现每个接口的性能需求。对寄存器的每一位的配置不同代表了不同的接口性能，如：设置HWCFG MODE[3:0]=0011就表示该MAC接口采用RGMII的光纤接入方式，若HWCFG MODE[3:0]= 1011则表示该MAC接口采用RGMII的RJ45铜口接入方式。PHY芯片的所有硬件寄存器的映射表如表4.1所示。根据相应的应用必须对相应的硬件寄存器位进行设置。

表4.1 PHY芯片硬件寄存器映射表

Pin To Configuration Register Mapping					Config
Pin	Bit [3]	Bit [2]	Bit[1]	Bit[0]	Bit[3..0]
Config0	PHYaddr3	PHYaddr2	PHYaddr1	PHYaddr0	0000
Config 1	Mode 3	Mode 2	Mode 1	Mode0	1011
Config 2	ANEG3	ANEG2	ANEG1	Aneg0	1111
Config 3	PHYADR4	ENA_XC	DIS_FC	Dis_Sleep	0110
Config 4	Reserved=0	Reserved=0	SEL_TWSI	Ena_pause	0000
G config 0	Dis_dte	75/50 ohm	1/4 MDIO	Dis_125	1011/1001
G config 1	LED_tx Blink	Power Down	Sig_Det	Int_Pol	1110

表中最后一列是网络安全加速卡的PHY芯片配置数据。

通过对相应的寄存器的设置可以完成以下功能：物理接口地址配置、硬件配置模式、MAC接口模式、自协商、检测信号、内部端接电阻、MDIO管理接口和时钟等的配置。相应的配置最后由电路原理图来实现。图4.8所示为铜口模式下的RGMII接口配置方式的电路连接图。

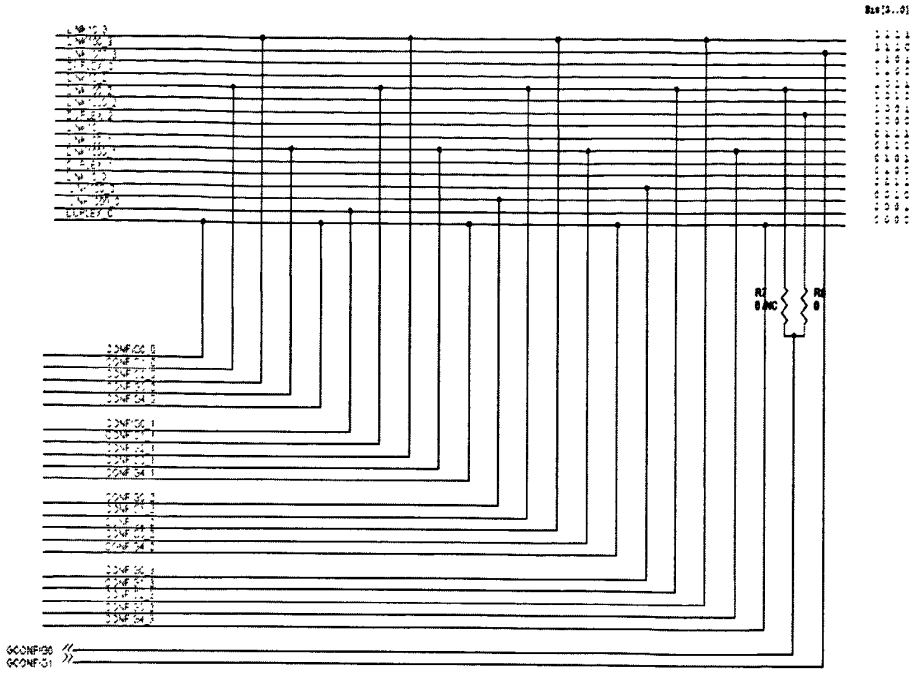


图 4.8 PHY 芯片配置电路

4.3 硬件查找模块设计

第三章对基于 TCAM 的规则命中原理做了详细的分析。此处，将结合具体的 TCAM 芯片的应用方式，来设计具体的硬件电路。

网络数据流量进入 FPGA 后，FPGA 首先会对高速数据流进行数据分析，提取分类关键字。FPGA 将提取出来的关键字与内容地址存储器中的所有条目进行比对，最后返回一个匹配条目的地址[34]。该条目的信息，就是系统对数据所要进行的操作信息。使用三态内容地址存储器用硬件的方式实现匹配条目存储和比对，具有很高的匹配速度。

本文采用 IDT75P42100 芯片作为网络数据包搜索引擎。

IDT75P42100 是一个 2Mbit 的三态内容地址存储器。它能对数据带宽为 36bits，72bits，144bits，288bits 和 576bits 的数据进行查找和比较。它有 8 个可灵活配置的有效字段，可以很好地适应不同的应用系统。每一个字段的配置方式有如下几种：4K*72,2K*144,1K*288,512*576。根据应用，本系统采用 144bits，16KEntry 的查找模式。

CAM 具有两种应用方式。本系统设计所用的 IDT75P42100 也不例外。这两种应用方式各有自身的应用场景。第一种应用方式如图 4.9 所示。

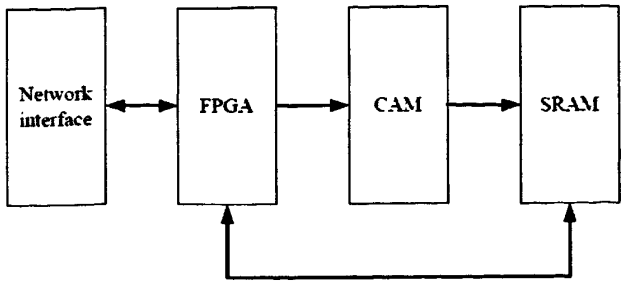


图 4.9 CAM 存储器的应用方式一

该方式下的应用过程有如下步骤：

(1) 将过滤规则或者采样规则下载到 FPGA，FPGA 再将规则下载到 CAM 和 SRAM 中。其中 CAM 的数据寄存器中存放有规则的关键字，SRAM 中存放着每个规则对应的操作，如：抓取或者丢弃等。存放有规则关键字的 CAM 芯片的数据寄存器地址与存放操作的 SRAM 的地址是一一对应的。

(2) 网络数据通过网络接口传输到 FPGA，FPGA 按规则提取数据包中的关键字，然后将关键字送到 CAM，通知 CAM 执行 Lookup 操作。

(3) CAM 执行完 Lookup 操作之后，若找到相匹配项则通过相应总线选中 SRAM 中的相应地址。

(4) FPGA 从 SRAM 中读取相应操作码，执行相关操作。

第二种应用方式如图 4.10 所示：

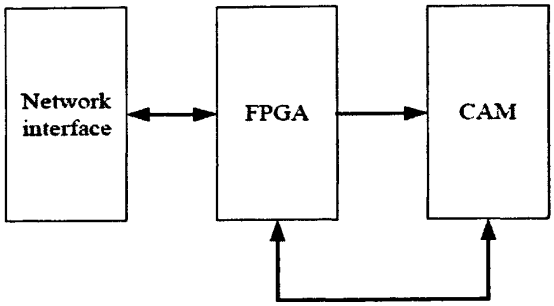


图 4.10 CAM 存储器应用方式二

(1)首先将过滤规则或者采样规则下载到 FPGA 中，FPGA 将规则下载到 CAM 中，类似于场景一，CAM 的数据寄存器中存放有规则的关键字。

(2)网络数据通过网络接口传输给 FPGA，FPGA 按规则提取数据包中的关键字并传送到 CAM 中，通知 CAM 执行 Lookup 操作。

(3) CAM 将执行结果反馈给 FPGA。

(4) FPGA 通过 CAM 提供的索引找到相应的操作码，执行相应操作。此时，FPGA 用内部逻辑配置 RAM，取代外部 SRAM 的功能，从而使操作更快。

由于第二种应用方式具有更快的操作速度，更加符合网络安全加速卡的应用环境。本设计采用第二应用方式设计网络安全加速卡，系统采用 144bits，16KEntry 的查找模式，此硬件电路设计如图 4.13 所示(电路图所示为时钟和配置部分电路)。

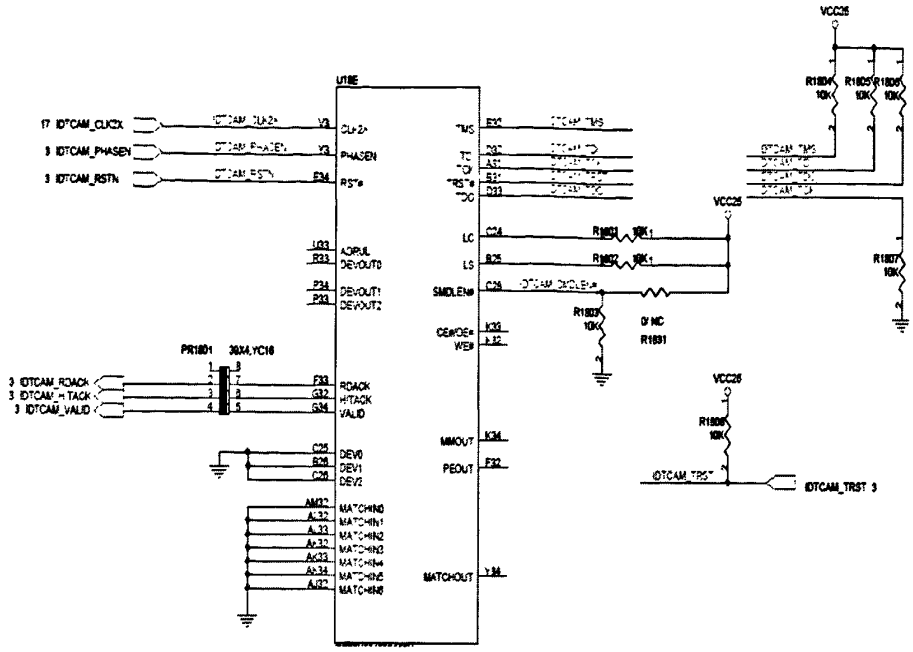


图 4.11　CAM 时钟与配置电路

数据总线扩展与地址总线扩展的电路设计如图 4.12 所示：

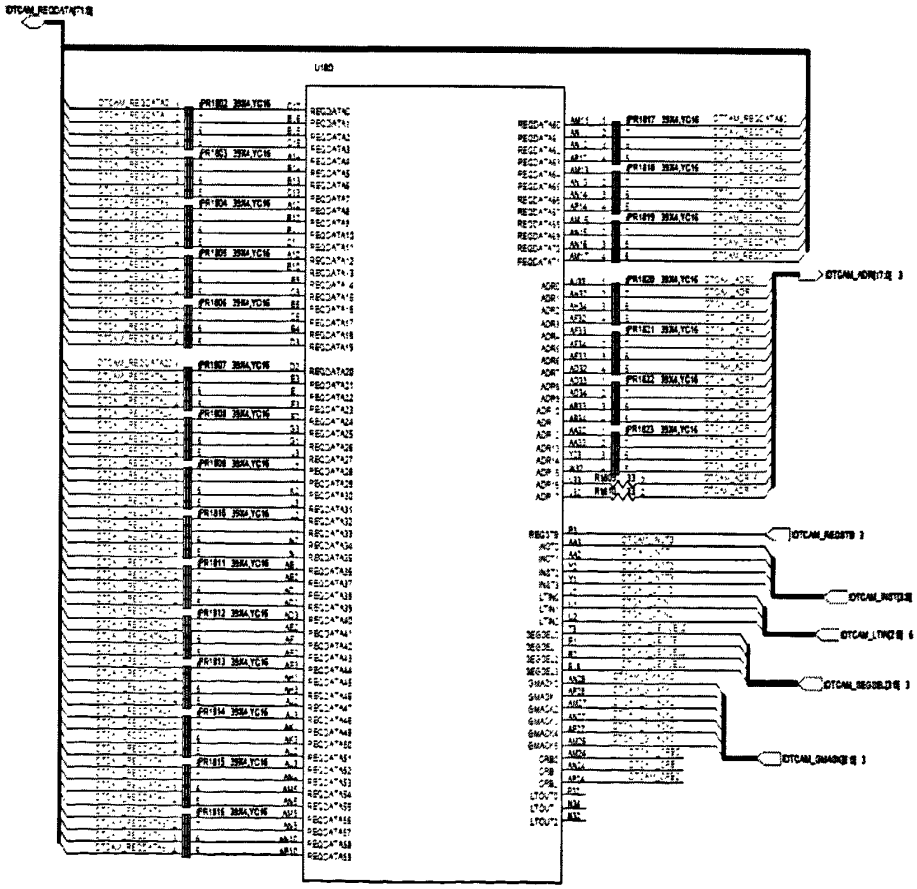


图 4.12　CAM 总线扩展电路

4.4 转发模块设计

进入系统的数据包经过FPGA数据包分析和CAM命中规则匹配之后，如果命中的规则是转发，则要求系统将所需转发的数据流直接向外网转发。由于系统数据包流量非常大，因此需要一种机制来实现数据包的硬件线速转发，否则无法满足系统的高速应用要求。

本设计通过对硬件存储器的读写来完成数据包的高速存储转发。在半导体存储器中，只有随机存储器才能够实现数据的随时写入和读出操作。随机存储器目前常用的有：SRAM静态随机存取存储器(包括了SSRAM同步静态随机存取存储器)和SDRAM同步动态随机存取存储器(DRAM的一种)。SSRAM的集成度低，但是速度很快，常用于系统缓存。而SDRAM是同步动态随机存取，需要电路对行列的不断刷新来维持数据的存储，掉电后数据随机丢失。虽然SDRAM的存储速度不如SSRAM那样快，但是目前的DDR2-SDRAM和DDR3-SDRAM都提高了时钟速率，并且采用上下沿采样数据，存储器的读写速度得到了极大改善。更重要的是SDRAM具有比SSRAM大得多的内存容量。SDRAM的这些优势与本设计的高速大容量存储转发应用非常吻合。

与传统的SDRAM相比，DDR2-SDRAM由于在时钟上下沿都进行采样，因此具有双倍于SDRAM的读写速度，常作为系统内存使用。本设计出于需要存储和转发的流量非常巨大，因此选择了由多个DDR2-SDRAM内存颗粒做成的内存模组的并行工作来实现报文地迅速读写，与核心FPGA控制芯片配合完成存储转发功能。

本系统采用DDR2-SDRAM模组来实现数据的高速存储和读出。设计采用Micron公司的MT16HTF12864H(I)-1GB模组可以很好的满足系统的要求。其各种速度等级的模组对应的时序参数如表4.2所示：

表4.2 MT16HTF12864H(I)时序参数

Speed Grade	Industry Nomenclature	Data Rate (MT/s)				^{t}RCD (ns)	^{t}RP (ns)	^{t}RC (ns)
		CL = 6	CL = 5	CL = 4	CL = 3			
-80E	PC2-6400	-	800	533	-	12.5	12.5	55
-800	PC2-6400	800	667	533	-	15	15	55
-667	PC2-5300	-	667	533	400	15	15	55
-53E	PC2-4200	-		533	400	15	15	55
-40E	PC2-3200	-		400	400	15	15	55

DDR2-SDRAM模组各参数的解释如表4.3所示。

表 4.3　DDR2-SDRAM 模组参数表

参数名称	参数描述
Speed Grade	速度等级
Industry nomenclature	通用名称
CL	CAS 潜伏期
CAS	列地址选通脉冲
tRCD	RAS 到 CAS 延迟。在发送列读写命令时必须要与行有效命令有一个间隔，这个间隔就是 tRCD
tRP	预充电有效周期。在发出预充电命令之后，要经过一段时间才能允许发送 RAS 行有效命令打开新的工作行，这个间隔就是 tRP。预充电有效周期。在发出预充电命令之后，要经过一段时间才能允许发送 RAS 行有效命令打开新的工作行，这个间隔就是 tRP。
tRC	行活动周期。

DDR2-SDRAM 模组的另外一个突出优点是可以根据自己的特点内建合适的终端电阻(ODT)。在第一代的 DDR 中为了防止数据线终端反射信号就在主板上做了大量的终端电阻。但是不同的内存模组需要不同的端接电路。端接电阻的大小与数据线的信号反射密切相关：端接电阻小则信号反射小，信噪比也小；端接电阻大则信号反射大，信噪比也高。因此主板上的端接并不能很好的解决信号发射问题。本设计根据本 DDR2 模组的特点，除了结合内建合适终端电阻和外接合适大小的电阻进行终端匹配，此外还通过外部的 fly-by 端接设计，很好的保证了信号的最佳波形。DDR2 模组的内部端接设计如图 4.13 所示，DDR2 模组的 fly-by 端接设计电路如图 4.14 所示。

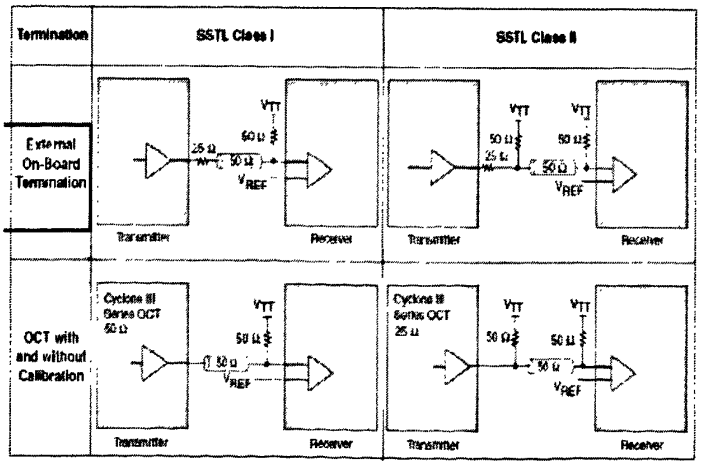


图 4.13　DDR2 模组内部端接电阻示意图

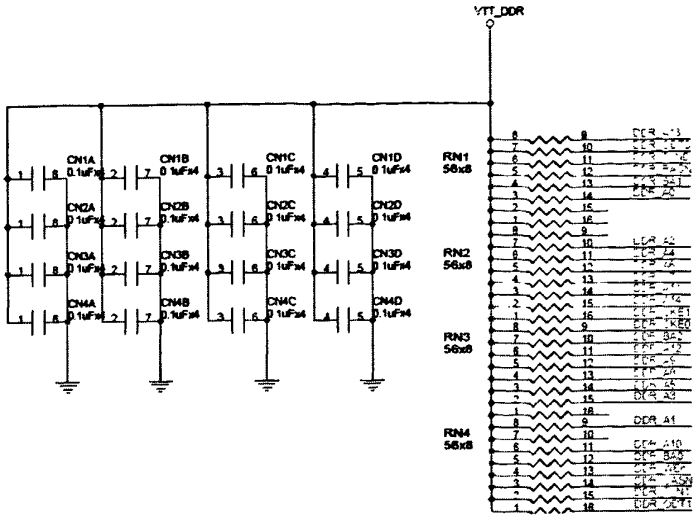


图 4.14 DDR2 模组 fly-by 端接示意图

4.5 PCIE 总线接口设计

PCIE 接口是新一代的高速总线接口。该总线接口向设备提供专用串行连接，采用点对点的通信方式，无需总线请求带宽，因此具有相当高的数据传输率。与 PCI 总线的单向传输方式不同，PCIE 总线采用双工连接方式，提高了信号通信的速率和质量。第一代的 PCIE 规范中，单通道提供 2.5G/S 的传输速率；在第二代 PCIE 规范中，单通道的传输速率达到了惊人的 5G/S。与之前的系统总线 PCI 和 PCIX 相比，PCIE 总线具有明显的优势：

(1) 在两个设备之间点对点串行互联（两个芯片之间采用接口连线互联；设备之间使用数据电缆互联，而 PCI Express 接口的扩展卡之间使用连接插槽进行连接）：与 PCI 所有设备共享同一条总线资源不同，PCI Express 总线采用点对点技术[64]，能够为每一块设备分配独享通道带宽，不需要在设备之间共享资源，这样充分保障了各设备的宽带资源，提高了数据传输速率。

(2) 双通道，高带宽，传输速度快：在数据传输模式上，PCI Express 总线采用独特的双通道传输模式，类似于全双工模式，大大提高了数据传输速度。在传输速度上，1.0 版本的 PCI Express 将从每个信道单方向 2.5Gbps 的传输速率起步，而它在物理层上提供的 1～32 个可选信道带宽特性更使其可以轻松实现近乎"无限"的扩展传输能力[64]。 2.0 版本的 PCI Express 则具有每个信道单方向 5Gbps 的传输速率。

(3) 具有灵活扩展性能：与 PCI 不同，PCI Express 总线能够延伸到系统之外，采用专用线缆可将各种外设直接与系统内的 PCI Express 总线连接在一起。这样可以允许开发商生产出能够与主系统脱离的高性能的存储控制器，不必再担心出于改用 FireWire 或 USB 等其它接口技术而使存储系统的性能受到影响。

(4) 低电源消耗，并有电源管理功能：新一代的 PCI Express 总线采用比 PCI 总线少得多的物理结构[64]，如单 x1 带宽模式只需 4 线即可实现调整数据传输，实际上是每个通道只需 4 根线，发送和接收数据的信号线各一根，另外各一根独立的地线。当然实际上在单通道 PCI Express 总线接口插槽中并不是 4 针引脚，而是 18 针，其余的 14 针是通过 4 根芯线相互组合得到的。由于减少了数据传输芯线数量，所以它的电源消耗也就大大降低了。

(5) 支持设备热拔插和热交换：PCI Express 总线接口插槽中含有"热拔插检测信号"，所以可以像 USB、IEEE 1394 总线那样进行热拔插和热交换。

(6) 支持 QoS 链接配置和公证策略

(7) 支持同步数据传输：PCI Express 总线设备可以通过主机桥接器芯片进行基于主机的传输，也可以通过交换器进行点对点传输。

(8) 具有数据包和层协议架构：采用类似于网络通信中的 OSI 分层模式，各层使用专门的协议架构，所以可以很方便地在其它领域得到广泛应用。

(9) 每个物理链接含有多点虚拟通道：PCI Express 总线技术在每一个物理通道中也支持多点虚拟通道[64]，理论上来讲，每一个单物理通道中可以允许有 8 条虚拟通道进行独立通信控制，而且每个通信的数据包都定义不同的 QoS。正因如此，它与外设之间的连接就可以达到非常高的数据传输速率。

(10) 可保持端对端和链接级数据完整性：由于 PCI Express 总线的分层架构，非常有利于保持端到端和链路级联数据的完整性。

(11) 具有错误处理和先进的错误报告功能：

由于 PCI Express 总线的分层架构中软件层的主要功能就是进行错误处理和提供错误报告，具有先进的错误处理与报告功能。

(12) 使用小型连接，节约空间，减少串扰：

PCI Express 技术与 PCI 相比，总线的导线数量减少了将近 75%[64]（PCI Express 总线也会有好几种版本的），速度会加快而且数据不需要同步。由于接口导线和主板上导线的减少，从而可以使通过增加走线数量提升总线宽度的方法就更容易实现，同时各走线之间的间隔就可以更宽，减少了相互之间的串扰。

(13) 在软件层保持与 PCI 兼容：

跨平台兼容是 PCI Express 总线非常重要的一个特点。目前被广泛采用的 PCI 2.0 设备可以在这一新标准提供的低带宽模式下运行，不会出现类似 PCI 插卡无法在 ISA 或者 VLB 插槽上使用的问题，从而为广大用户提供了一个平滑的升级平台。同时由 IBM 创导的 PCI-X 接口标准在 PCI Express 标准中也得到了兼容，但要注意的是它不兼容目前的 AGP 接口。

PCIE 时钟频率采用差分的 100M 时钟[64]。规范规定在发送和接受之间的延迟不能超过 10ns。根据传输线理论，为了有效地防止信号反射，PCIE 的端接电阻接法如图 4.15 所示：

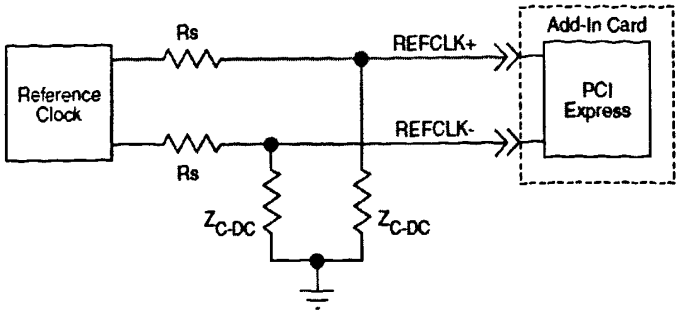


图 4.15 PCIE 端接电阻设计

PCIE 参考时钟点接法如图 4.16：

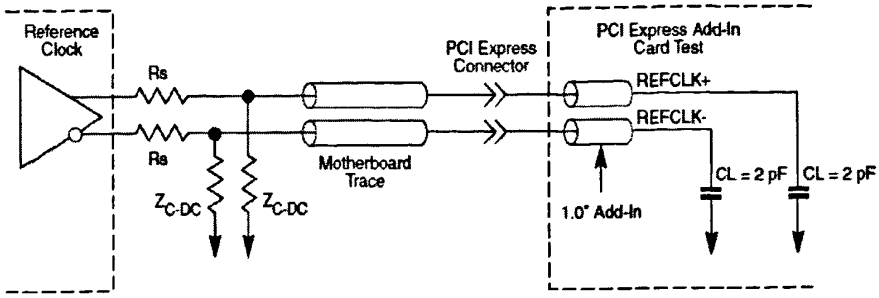


图 4.16 PCIE 参考时钟设计

PCIE 接口电路设计连接图如图 4.17 所示：

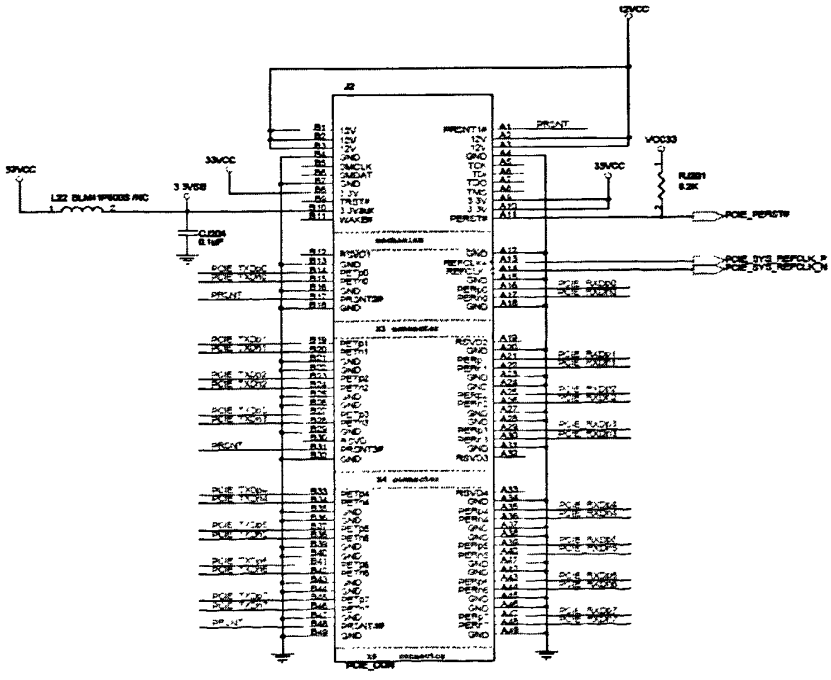


图 4.17 PCIE 接口设计电路图

4.6 电源系统设计

电源模块是一个系统正常工作的基础，一个良好的电源系统必须满足系统的功耗并保证各个功能模块地正常运行。本系统需要的电源类型有：3.3V、2.5V、1.8V、1.2V、硬件转发模块的参考电压和端接电压。而 PCIE 插槽上能提供的电源只有 12V 和 3.3V，因此需要在充分考虑功耗的情况下将现有的两种电压转化为系统所需的多种电压。本系统采用 TI 公司的 DC-DC 降压模块分别将 12V 电压和 3.3V 电压转换出系统所需的所有类型的电压值。本系统采用 TI 公司的 PTH04T 230W 和 PTH04T-240W 变压模块对 PCIE 插槽的 3.3V 电源进行转换；采用 TI 公司的 TPS40200 变压模块对 PCIE 插槽的 12V 电源进行转换。

PTH04T-230W 模块能将 PIE 插槽的 3.3V 电源电压转换为 0.69V 到 3.6V 的任何一种输出电压，支持最大电流 6A。该模块采用了 Turbo Trans 技术使输出误差极小，并且减小了模块面积。PTH04T-240W 电源模块和 PTH04T-230W 电源模块功能相似，但是该模块最大能支持高达 10A 的电源电流，能提供更大的功耗。

TPS40200 变压模块则能把 12V 的电压转换为 0.7V 到 1.8V 的任何一种电压，只需改变芯片外部的电路连接即可得到相应的电压值。

此外，系统还设计了上电检测电路。每个电源模块都能通过一个 Track 引脚来控制不同模块的上电顺序或者让所有的模块同时上电运行。上电检测电路能检测 PCIE 插槽过来的电源信号，若上电不稳定或者小于检测电路给定的阀值，检测电路就自动发出一个复位信号给所有电源模块，达到禁止电源输出的目的。直到检测到 PCIE 电源稳定时，检测电路才会失能 Track 引脚，使电源模块产生稳定的输出电压。这种设计方式保证了系统电源的稳定性，为各功能模块的正常运行打下了良好的基础。图 4.18 为系统电源网络分配框图。

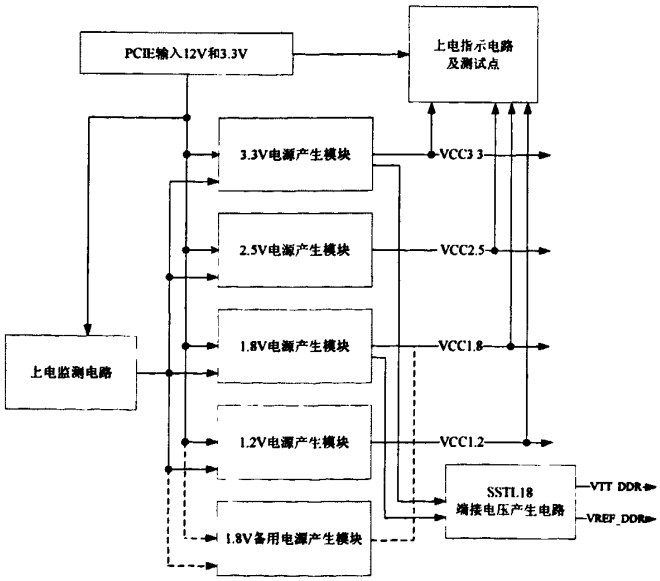


图 4.18 系统电源网络设计框图

4.7 FPGA核心处理模块设计

本设计的核心FPGA采用EP2S60器件。该器件是Altera公司Stratix II系列的一款高性能FPGA，采用1.2V核心电压、90nm芯片技术、9层金属走线和全铜SRAM的制造工艺。主要特性有：内嵌RAM块，DSP块，锁相环(PLL)和外部的存储器接口等。与上一代的Stratix系列器件相比具有一些新的改进：采用全新的自适应逻辑模块(ALM)，增加了对新的存储器的接口支持(如DDR2 -SDRAM等)，增加了源同步通道动态相位对准电路(DPA)以及增加了128位密钥对配置文件进行加密。

EP2S60拥有丰富的内部资源，完全可以实现系统所需的性能：拥有M512 RAM块329个，M4K RAM块255个，M-RAM块2个，DSP块36个，逻辑阵列块有62列51行。图4.19分别表示了该器件的平面布局框架和内部逻辑阵列结构。

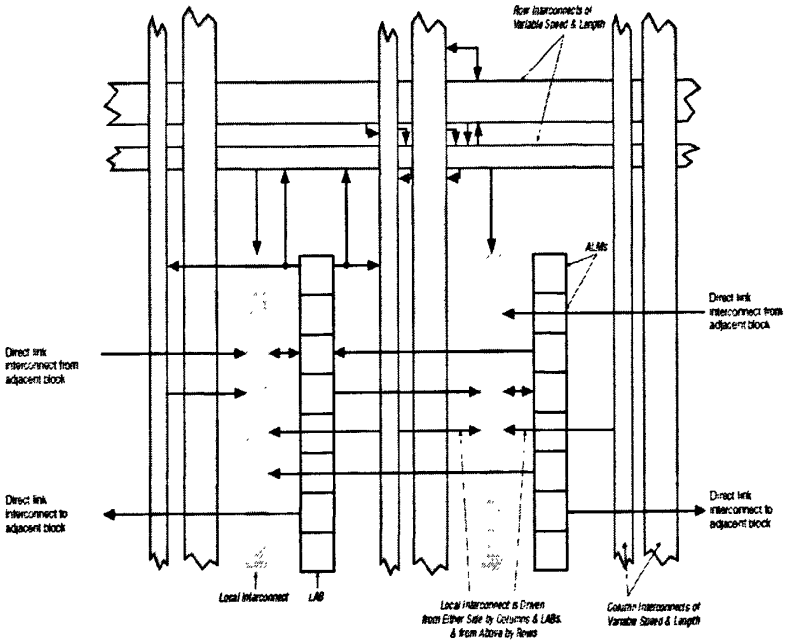


图4.19 EP2S60内部逻辑阵列结构

相比于以前的FPGA，该器件最具革命性的一点是采用ALM结构。ALM内部有两个3输入的加法器，与传统的2输入加法器相比提高了计算性能并且大大地减少了加法电路的级数。更具灵活性的是ALM中的组合逻辑模块可以根据用户的需求自动配置成所需的模式。

ALM可以灵活配置成以下的多种模式：4输入和4输入的LUT；5输入和5输入的LUT；5输入和3输入的LUT；6输入的LUT；5输入和4输入的LUT；6输入和6输入的LUT。查找表的配置方式如图4.20所示。

根据本系统的应用，FPGA需要完成：RGMII接口功能、多路数据复用与拆分功能、DDR2高速转发接口功能、数据包管理流控功能、PCIE接口功能、DMA功能、SSRAM高速缓存接口功能、CAM管理接口功能、数据包分类器功能和数据包分析器功能。其中图4.21展示了DDR2模块的DMA设计逻辑图。

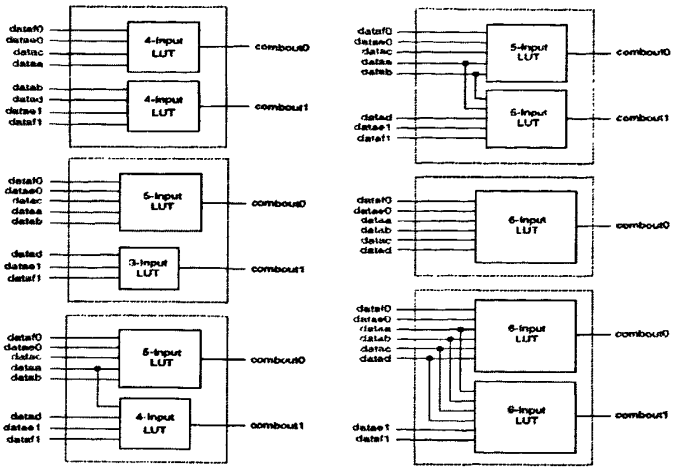


图 4.20 ALM 多种配置方式

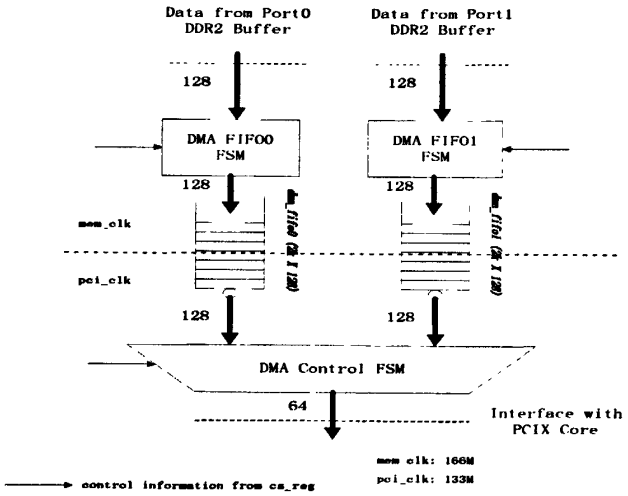


图 4.21 DDR2 接口 DMA 逻辑设计

FPGA 内部逻辑功能流图如图 4.22 所示：

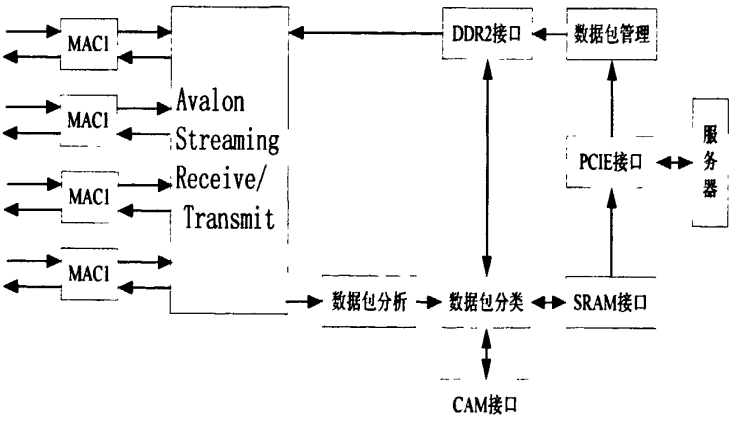


图 4.22 FPGA 内部逻辑功能设计图

多路数据包从外部 RJ45/SFP 接口进来后进入 PHY 芯片，FPGA 和 PHY 之间 MAC 接口采用 RGMII 接口方式。经由 PHY 的多路数据流在 FPGA 的内部逻辑(RGMII)通过 Avalon 总线传送到接收数据逻辑部分，多路数据在接收数据逻辑部分实现复用，合成一

路高速数据流，数据包分析单元对高速数据流进行字段抽取，提取关键字。通过关键字与硬件查找芯片 CAM 比对，返回一个地址。根据返回的地址读取内部的硬件匹配规则。同时将数据送存 DDR2 存储器和 SRAM 存储器。存入的数据加上了自己定义的信息头部。SRAM 中的数据通过 PCIE 接口送往服务器处理，服务器实现对数据流的采样和处理决定。对需要转发的数据段给一个转发指令。转发指令则通过数据包管理单元对 DDR2 的数据进行有选择的转发。由于有些数据在流经加速卡的过程中被丢弃，只有符合内网规则的数据才被转发，因此大大减少了下级网络的垃圾数据量。由于局域网一般为以太网，采用 CSMA/CD 协议侦听，由此限制一些不必要的数据流过本局域网，能非常有效的提高局域网的速度，极大地减少了拥塞。由于用硬件实现了数据的分析和规则的匹配，在实现防火墙功能的同时，极大的提高了数据的处理速度(和软件相比)。本设计采用分布式的方法对应经匹配规则进行管理。所以系统能够灵活支持新规则的写入，并能保证较快的执行速度。

4.8 时钟管理设计

由于系统所使用的模块较多，而且某些模块采用相同的工作时钟频率。在这种情况下，本设计采用有源晶振输出时钟，经过零延迟时钟 buffer 输出多路相同时钟给使用相同时钟频率的模块。不同工作时钟的模块则通过 PCIE 的 100M 系统时钟和晶振的时钟分两路送到 FPGA，在给 FPGA 提供工作时钟的同时，通过 FPGA 内部的 PLL 倍频后输出合适的时钟频率给外部的转发模块和网络接口模块，以提供合适的时钟源。时钟网络分布如图 4.23 所示：

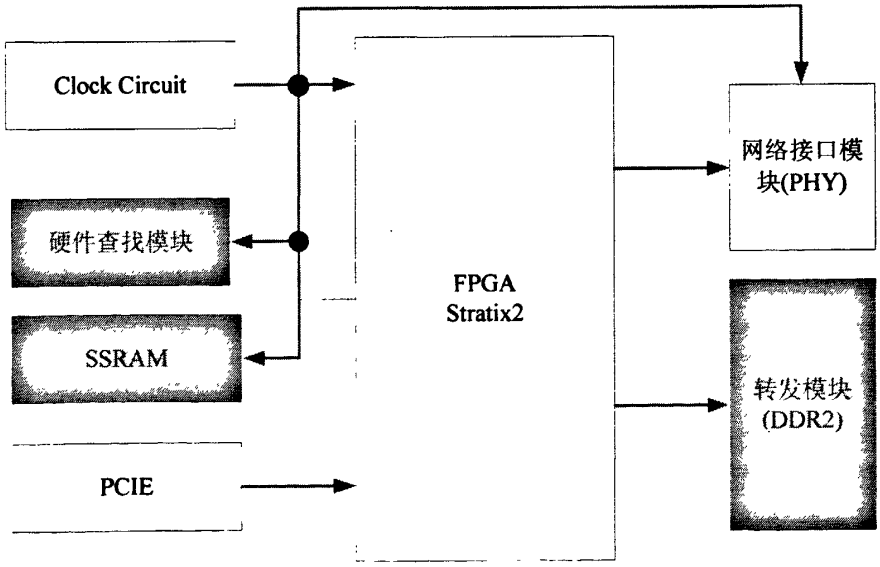


图 4.23 系统时钟分布图

网络安全加速卡设计了三种形式的复位：软件复位、按键复位、PCIE 接口自带复位。软件复位通过上位机程序给 FPGA 产生一个软复位信号，软复位信号通过 FPGA 内部的软复位逻辑产生各种软复位信号，分别给外部的功能模块接口。按键复位和来自 PCIE 接

口的复位，就是一个系统硬件复位，复位信号是一个长时间的低电平触发信号。触发信号到达 FPGA 内部的硬件复位逻辑块就会产生多路硬件复位信号，分别给各功能模块接口，同时作为 FPGA 系统的自身复位信号。至此，达到了整个系统同步复位的目的。图 4.24 表示了复位路径的分布。

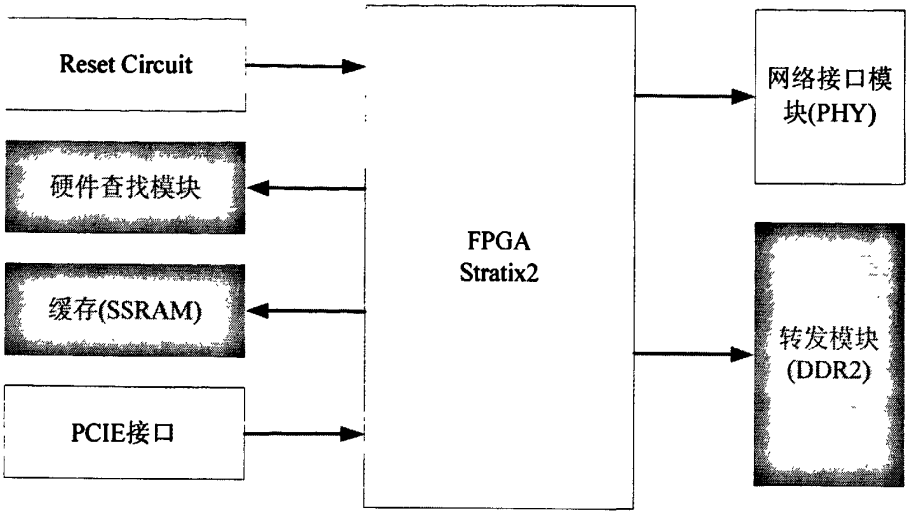


图 4.24　复位路径分布图

4.9 bypass 电路设计

由于网络安全加速卡是用在网络节点的管理设备当中，所以必须保证加速卡在掉电的情况下不影响其他网络地正常运行。本文设计了一种掉电情况下的网络 bypass 电路，在断电等特殊情况下能经过本卡的网络自动组成环路，不会影响其他网络地运行(一般为各个局域网)。本设计以 RJ45 的铜接口 bypass 为例详细介绍本文的 bypass 设计流程和功能。图 4.25 为 bypass 设计框图。

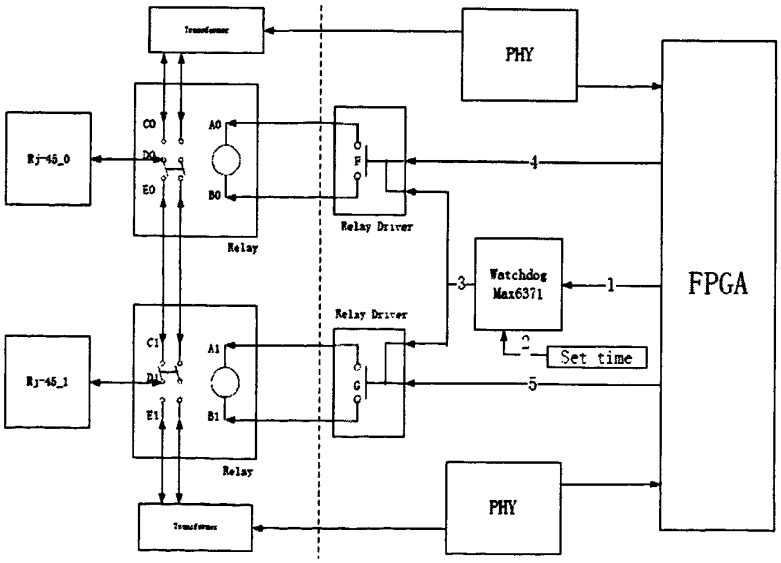


图 4.25　bypass 电路设计

Bypass电路的详细工作流程如下：

如果电路板突然掉电，此时Relay也断电，驱动开关A0、B0/A1、B1随即断开。由于驱动部分的掉电和驱动信号的消失，使Relay动触点D0/D1恢复到初始连接状态(D0连接E0，D1连接C1)。此时Rj-45_0接口与Rj-45_1接口互通，成为处于Bypass状态。

当电路板处于正常供电状态时，待FPGA及软件正常运行后，软件通过串行接口4/5对Relay Driver写入相应开关闭合的信息，开关F/G随即闭合。Relay Driver驱动Relay中的A0,B0/A1.B1的闭合，这时Relay启动并将动触点D0/D1分别打向C0/E1。此时，网络接口从Bypass状态进入正常工作的状态之中。

在FPGA逻辑与软件运行过程之中，Watchdog芯片起到监视作用。如果FPGA逻辑与软件运行正常，则会通过线路1对Watchdog中的Timer进行定时清除操作。如果FPGA逻辑或软件中有任何错误，都会导致线路1无法清除Timer的计数数据，看门狗计数器随即溢出并给Relay Driver一个复位信号，Relay Driver复位后随即断开开关F/G， Relay不再工作，动触点D0/D1将自动处于初始连接状态，系统进入Bypass状态。

Watchdog的Timer定时时间由SET引脚(通过线路2)来进行设置，可选时间有1ms，3ms，10ms，100ms，300ms，3s，60s和Disabed。Relay开关闭合或断开操作的时间最大为3ms。

Relay Driver是一个8路带开关的驱动器，它可以通过串行接口对内部的8路开关进行配置，使其闭合或是断开，从而驱动相应的Relay，同时它可以由单一的引脚SET或是Reset对8路开关进行同时闭合或是断开的操作，便于调试中使用。

在光纤接口的情况下，使用光开关来实现Bypass功能，实现的原理与RJ45的铜接口的实现方式一致。

本设计通过相应的电气开关构成了一个能快速反应的掉电保护电路，很好地保证了网络在特殊情况下的畅通。在本管理节点完全不能的情况下，也能实现外网的互通，不会造成网络数据包的丢失。

4.10 与加速软件的性能对比

本网络安全加速卡比传统的加速软件具有更高的性能。网络安全加速卡采用流转发的方式来处理报文，如有如下的优点：

(1)数据流的首包采用传统的CPU处理方式，在CPU获得必要的路由信息、状态跟踪信息后，把针对流的安全策略、路由信息和处理决策等内容下发到网络安全加速卡中。在FPGA中，每一个数据流都有一条记录这些信息的条目，称为SSN entry。

(2)数据流中的后续报文可以根据SSN entry来自行处理，不再需要CPU的介入。

(3)为了给CPU足够精细的控制粒度，CPU可以随时修改SSN entry，来决定网络安全加速卡是直接对数据流进行转发还是交给CPU来处理。

(4)网络安全加速卡把首包的流分类结果附在上行报文中一起上交给CPU是被和处理，使CPU获得了时间复杂度为O(1)的流分类处理能力，加速了CPU对首包的处理。

表4.4对基于网络应用的主要性能做了详细比较。

表4.4 网络安全加速卡与传统加速软件的性能比较

性能	传统加速软件	网络安全加速卡
总线	操作系统采用PCI/PCI-X/PCI-E进行调度。	只有首数据包上交CPU处理，后续报文处理都在网络加速卡上完成，不需要再上交到CPU。
中断	频繁的软件中断，易导致CPU产生中断锁。通常采用Polling机制来解决中断锁的问题，但这种机制增大了报文处理的时延	可以在网络安全加速卡中专为收发报文设计多个线程(128个)，而且只有首数据包才会和CPU进行交互通信。因此，不会产生中断锁的问题。
路由表查询和状态表查询	(1) 传统软件逐包查询路由表,典型的Radix trie算法的时间复杂度是O(log n)。 (2) 即使经过优化的路由查询和状态表查询算法也会大量消耗CPU大量的处理时间。 (3) 最新的多核CPU通过多线程查表的方式可以提升性能	(1)路由表和状态表统一查询。 (2)芯片中128个线程进行查表运算，相当于128 核的专用处理器。 (3)采用流水线作业，这128个线程可以对报文并行处理。
报文处理	报文处理和报文修改都需要消耗CPU的大量处理时间。常用的软件算法是checksum 计算方法。	
流分类	(1)最经典的算法是顺序查找，时间代价O(N)，在规则增多时几乎不可用。 (2)优化的流分类算法可以提升查找效率，但也极大的提升了软件实现的复杂度。	通过TCAM查询流分类策略表， 查询的时间复杂度为O(1)。
流量采样	流量监管都需要采用细粒度,时间粒度越小，消耗 CPU处理能力越多。	网络安全加速卡可以在毫秒级对报文速率进行采样而不需要CPU参与。

4.11 本章小结

本章对整个系统的硬件设计过程做了详细的介绍。首先提出了系统的整体性能设计指标并根据指标确定了系统构架，然后分别对各个部分的功能设计进行了具体的描述，包括：网络接口设计，硬件查找模块设计，硬件转发模块设计，PCIE 接口设计，电源网络设计，时钟网络设计，FPGA 核心处理模块设计和 bypass 功能设计等。各个模块的相互配合能够完成系统的预定性能指标。

首先，网络数据包从网络接口模块计入系统，数据通过 RGMII 接口进入到 FPGA 中。FPGA 对各路数据包进行复用，形成一路高速数据流。高速数据流经 FPGA 处理后将相应的数据直接存入 SSRAM 缓存和 DDR2-SDRAM 中。DDR2-SDRAM 中的数据直接转发到相应的网络接口，而 SSRAM 中的数据则需要给服务器软件做一个数据流的分析，实现采样的功能。

由于 DDR2-SDRAM 不通过 CPU 的仲裁，直接由硬件向外网转发数据包，因此在实现了数据流的高速转发功能的同时还显著地减轻了 CPU 的负荷。这种方式起到了加快数据包转发和加快网络节点 CPU 运行效率的双重作用。

采样的过程是通过一系列的仲裁过程来实现的。存入 SSRAM 中的数据报文就是需要服务器系统采用的数据报文。FPGA 通过与 SSRAM 的接口，控制 SSRAM 中的数据报文传送到服务器的 CPU 中。为了加快数据的采样速率，采用 DMA(直接存储器访问)的方式通过高速 PCIE 系统总线向服务器传送报文。经过服务器采样和处理过的报文则通过 PCIE 高速总线接口通过 DMA 的方式将要转发的数据存储到加速卡的 DDR2 中。FPGA 则会控制 DDR2 将数据转发到外网的相应接口，实现了数据报文的有效快速转发。至此，完成了一次数据包的高速分析采样和转发过程。

至于进入系统的数据包是需要采样、转发还是直接丢弃，由 FPGA 核心处理模块中的数据包分析逻辑块做裁决。数据包分析逻辑会对已经由多路合成的高速数据流做数据包分析，解析出报文的五元组(源地址，目的地址，源端口，目的端口，协议类型)。裁决的依据则是硬件查找模块返回的规则匹配标志。

用硬件的方式来实现数据的分析、匹配、采样和转发，比传统的软件实现方式要快得多。通过软件的升级，还可以在不影响报文传输的情况下实现快速防火墙的功能。系统有效的解决了网络管理节点的数据包处理速率问题，使主干网更加畅通。

第5章 PCB板卡及信号完整性设计

在大规模数字系统 PCB 的设计过程中，需要特别注意以下几个方面的问题：信号完整性设计、高速系统设计、电源完整性设计和热设计[53]。本章结合网络安全加速卡的 PCB 设计来对信号传输问题和如何在布线中避免这些问题进行详细的阐述。第一节对信号完整性问题的起因和解决作一个详细的分析。第二节对高速系统设计的 PCB 设计做详细的探讨。第三节对网络安全加速卡主要部分的功耗做了详细的分析。

5.1 信号完整性设计

若要使信号在信号线上的传输过程中能正确地采样和识别，必须使信号具有比较小的失真以达到电平数值和跳边沿的特性，这就是信号完整性所要考虑的问题。其中最核心的问题就是要防止信号因传输过程中的严重变形导致接收端无法对信号进行正确采样。引起这些问题的主要原因有：信号反射，信号振铃，电子元器件的开关噪声，地弹，信号传输衰减，信号间串扰和容性负载等[53]。

所有的信号完整性现象，深究其原因可以从以下四点来考虑：单一网络的信号传输质量，电源和地噪声，不同信号线之间的串扰和系统的电磁干扰 。

传输线效应是影响信号完整性的一个重要原因之一。在高速设计系统中，如果信号在条边沿的变化频率太快，则非常容易导致信号产生严重的波动和反射现象，这就是所说的传输线效应[53,61]。

每一条信号线都有电压和电流通过，这就导致信号线看上去有一定的阻值，这个阻值称为特性阻抗。在特性阻值不连续的几点处往往会发生信号反射现象。因此，要防止信号反射现象的发生率就要保证传输线特性阻抗的连续性，具体到 PCB 制版中就是保证信号线宽的一致。信号线宽的不一致，走线与叠层之间的间距变化，高速信号走线换层或者过孔，没有构成完整的回路，使用连接器件，走线分叉等等都易引起严重的信号反射现象。因为这些现象都导致了信号线在传输过程中的阻抗变化。图 5.1、图 5.2 分别表示了线宽导致特性阻抗的变化情况，层间距变化导致特性阻抗的变化情况。

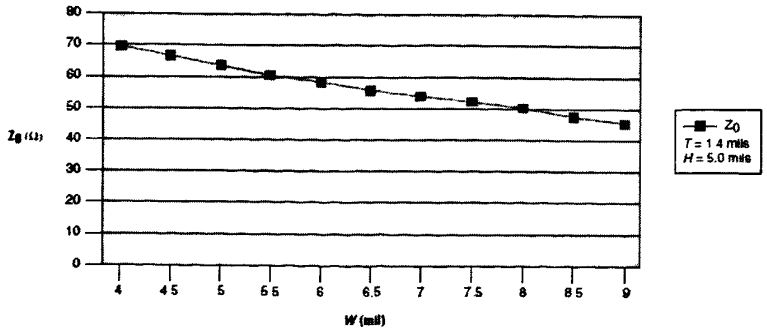


图 5.1　线宽与特性阻抗的关系图

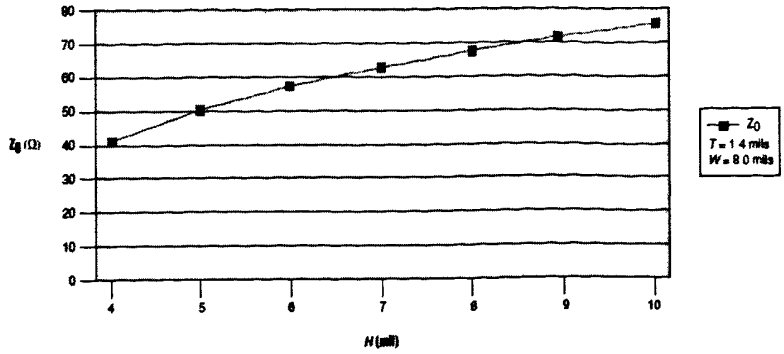


图 5.2 两参考层的间距与特性阻抗的关系图

另外，还通过端接匹配的方式可以使传输线阻抗连续，从而减小反射和振铃。常见的端接方式有：并行匹配方式、串行匹配方式、戴维南匹配方式、交流匹配方式和 FLY-BY 端接方式。

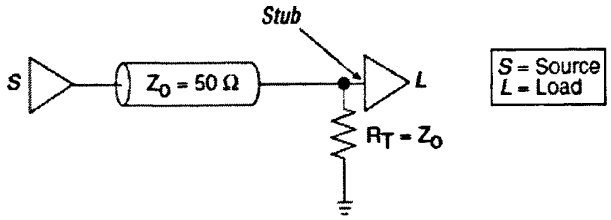


图 5.3 并行匹配方式

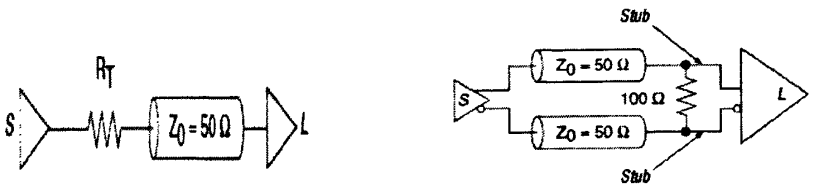


图 5.4 串行匹配方式

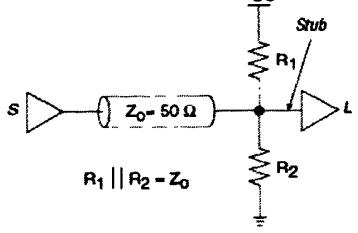


图 5.5 戴维南匹配方式

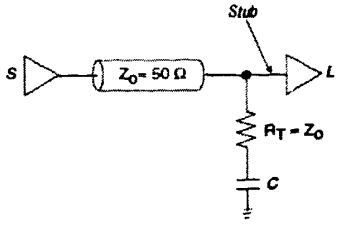


图 5.6 交流匹配方式

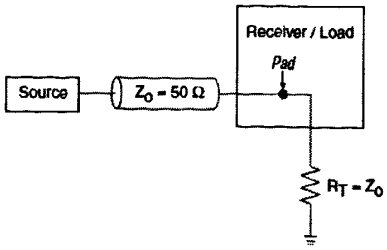


图 5.7　FLY-BY 匹配方式

第二个重要的因素是电源和地噪声。在实际的应用中，电源和地噪声总会受到来自外界和内部的一些干扰使电源产生一定的波动，对于高速板来说，这些波动会严重的影响系统的可靠性。一般外界的干扰多为低频的干扰信号。而内部的噪声干扰多为数字芯片的开关噪声，这些同步开关噪声会产生一些高频干扰信号。这些噪声让电源产生波动的原因是因为在电源分配系统中电源和地之间存在着交流阻抗，要减小干扰就必须减小电源分配系统的阻抗。对于高速 PCB 板，要让电源层和地层的间距尽量小，采用尽量薄的介质，另外就是采用一定数量的低电感的去耦电容。

第三个重要的因素就是串扰。串扰就是一条走线的电流发生变化时会在相邻的走线中耦合出一定的噪声电流。因此，要减小串扰就必须减小信号间的耦合效应。在高速 PCB 制版中要注意以下几点来减少串扰：同一层走线尽量远离，相邻层走线不要平行以及走线和参考平面间距尽量小。

最后一个因素就是在电子系统设计过程中存在无所不在的电磁干扰问题。尽量让信号线有一个很好的参考回流平面可以非常有效地减小电磁干扰的问题。这时散布到空间中的电磁干扰就小很多，对其他的电路和设备的干扰也就少很多。信号的合理分层也可以有效的减小电磁干扰，让所有的信号线都有参考平面。下图 5.8 为本设计中的信号分层情况：

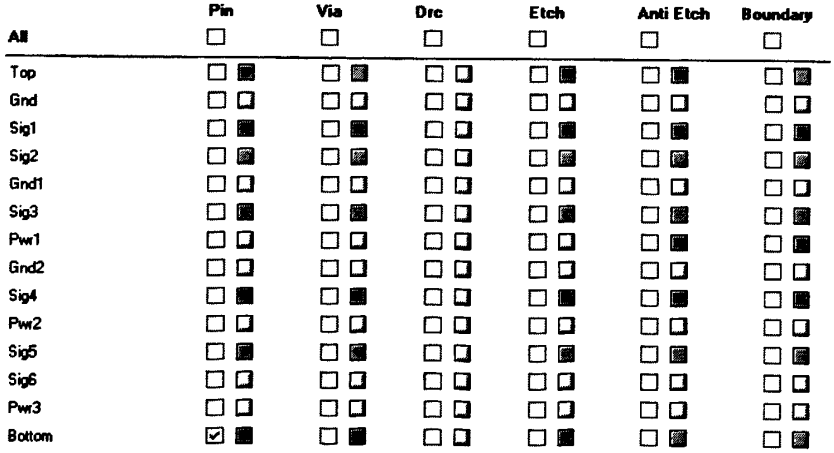


图 5.8　信号分层示意图

在 PCB 设计过程中要特别注意信号网络的信号完整性问题。

5.2 SERDES 及高速接口设计

FPGA 的高速接口包括 DPA，LVDS，SERDES 等[53]。本章首先从 SERDES 电路开始阐述 FPGA 高速通道设计的要领。

SERDES 电路是一种高速串并收发器。发送端是串行发送单元 serializer，通过高速时钟来对数据流进行调制，形成高速的编码数据流。接收端是串行接收单元 deserializer，该单元可以从高速数据流中恢复出时钟信号，并解调出原来的并行数据。接受单元的解调通过一个时钟数据恢复器(CDR)来完成解调。整个的发送过程就是在发送端将多路的并行数据流通过高速时钟调制成串行数据流进行串行传输，在接收端又从高速的串行数据流中恢复出原始的数据和时钟。SERDES 电路的结构图如图 5.9 所示：

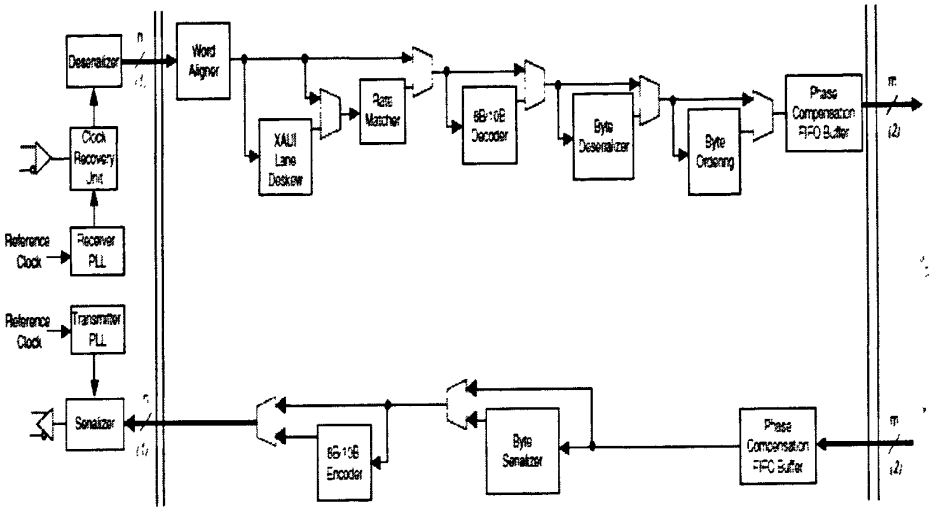


图 5.9 SERDES 结构图

接收端主要由以下功能模块组成：输入缓冲电路，收端环回缓冲电路，CDR 时钟数据恢复器，接收端 PLL，串并转换单元，码型检测器，通道对齐和数率匹配，8B/10B 编码，接收端到逻辑的接口。

发送端主要由以下功能模块组成：逻辑资源到发送端的接口电路，8B/10B 编码，发送端 PLL 和输出缓冲电路。

SERDES 电路和 LVDS 等高速通道的 PCB 设计应该注意以下事项：

(1)高速差分对的布线:根据情况使用边沿耦合，边沿耦合带状线，broadside 耦合等方式

(2)旁路电容要靠近 VCC，走线要尽可能短。采用低阻抗的小电容来作为旁路电容可以有效的滤除高速变换信号中的高频干扰。

(3)连接旁路电容时应该注意：为减小容抗，应该使用大尺寸的过孔连接电容的焊盘；使用短而宽的导线连接过孔和电容的焊盘。

(4)布高速时钟线时应注意：时钟线尽可能走直线；时钟线尽可能在单一信号处布线；为减小噪声和串扰尽量让时钟线靠近地平面;尽量少的使用过孔来减少时钟线的信号反射和阻抗不匹配问题。

(5)滤波设计：在每个电源接入端加滤波电路；在电源进入 PCB 的接入点附近用 100uf 的电解电容滤波；VCC 和 GND 处必须布置去耦电容

(6)高速翻转信号输出端必须外加 buffer，以避免过高的驱动能力要求。

(7)未使用的 IO 的输出设置为低电平，并且直接连接到地平面。

(8)电源和地分开布置

以本设计为例列举了几种常见的 PCB 走线错误如图 5.10-5.15 所示：

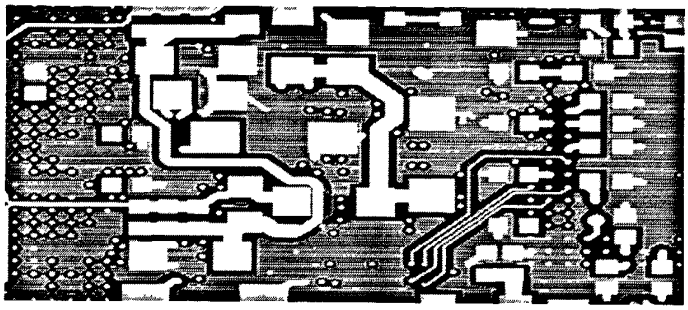


图 5.10 PCB 设计错误一

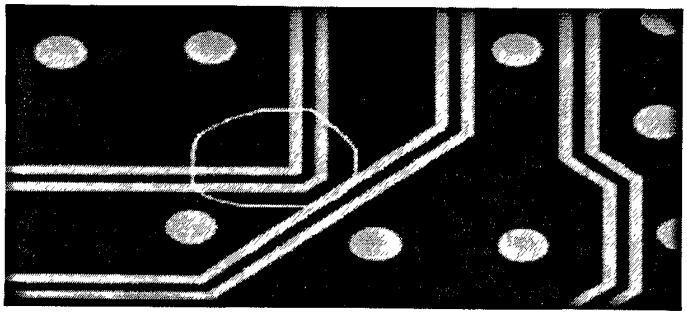


图 5.11 PCB 设计错误二

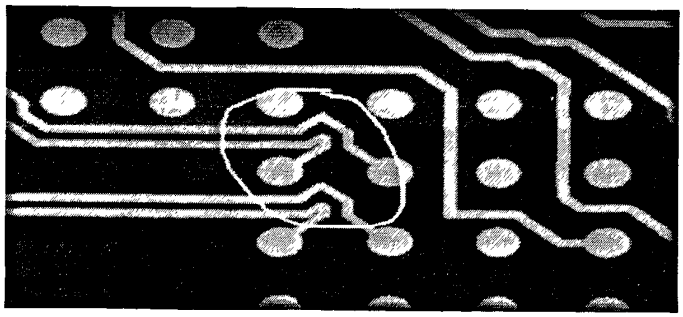


图 5.12 PCB 设计错误三

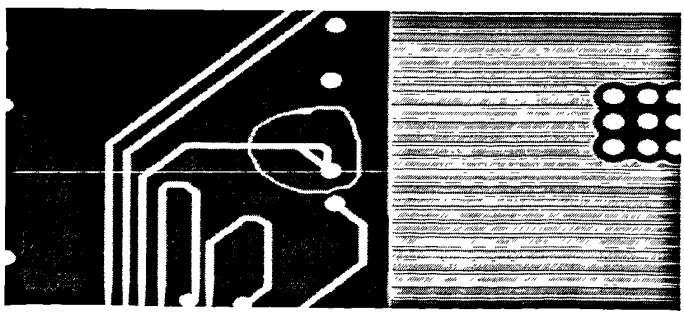


图 5.13 PCB 设计错误四

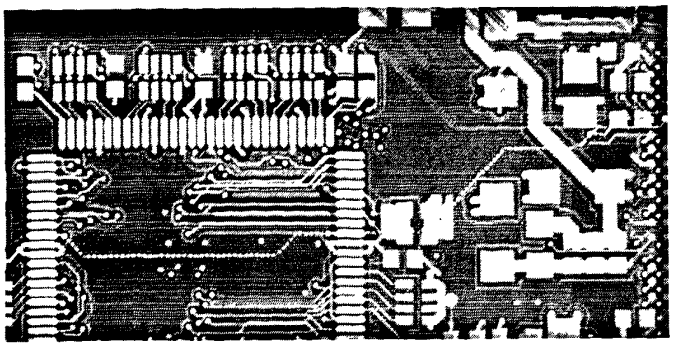


图 5.14 PCB 设计错误五

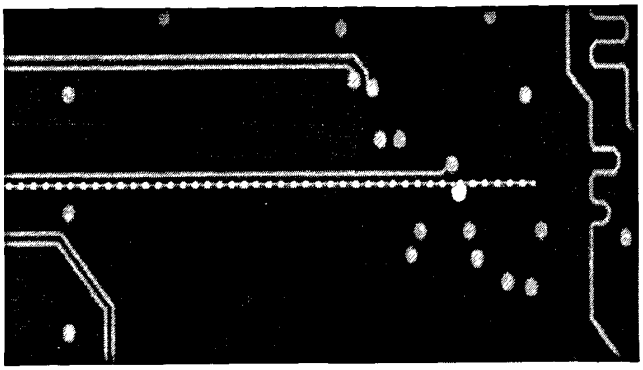


图 5.15 PCB 设计错误六

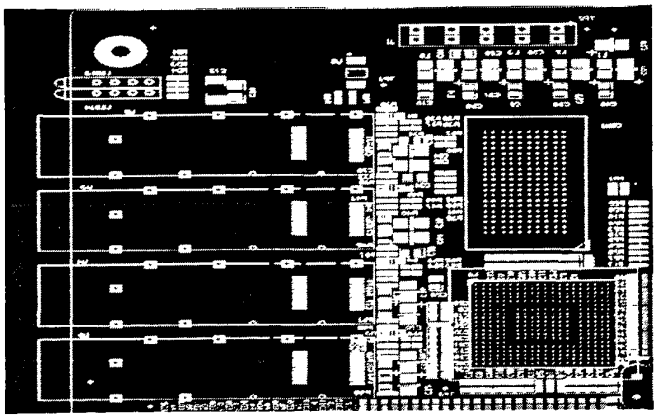


图 5.16 网络接口部分 PCB 布局图

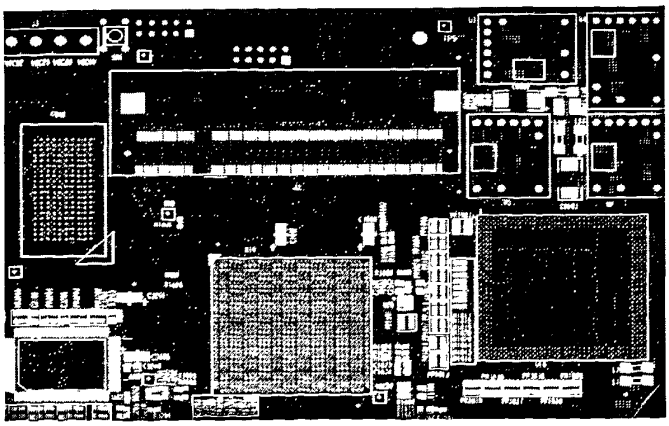


图 5.17 核心板顶层布局图

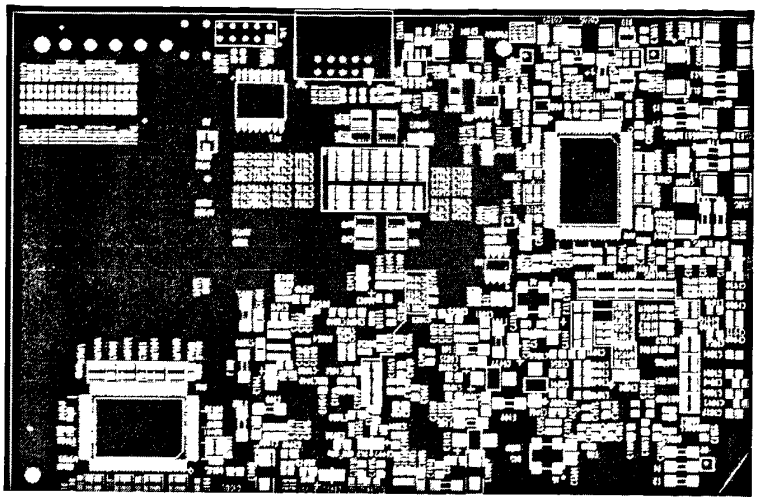


图 5.18　核心板底层布局图

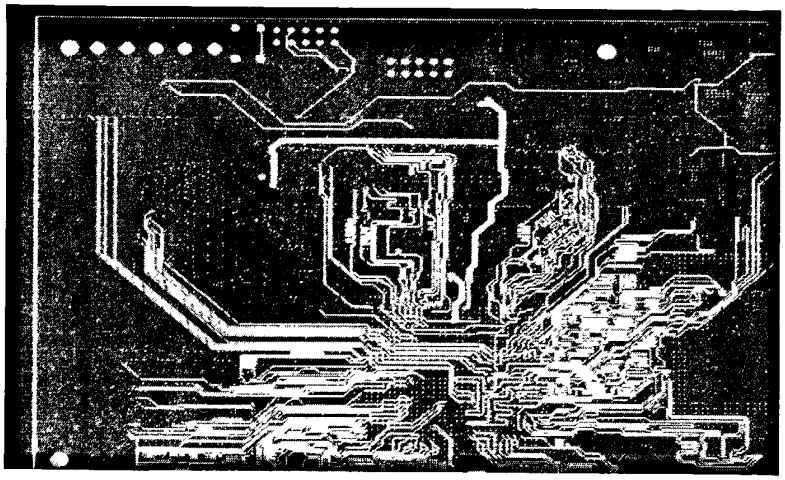


图 5.19　中间第一层信号层走线图

5.3　功耗分析设计

功耗的分析设计是系统硬件设计的关键所在。只有对系统的功耗需求分析具体到每一个模块，才能根据实际所需设计出合理的电源系统。本节主要对系统的几个核心模块做具体的功耗分析。表 5.1 列出了网络安全加速卡所用到的主要芯片。

表 5.1　网络安全加速卡主要芯片列表

编号	名称	描述
1	核心处理单元	EP2S60 芯片
2	硬件查找单元	75P42100 芯片
3	硬件转发单元	SO-DIMM 接口的 DDR2-SDRAM 模组
4	光模块	SFP
5	网络接口模块	Marvell 88E1145
6	高速缓存	SSRAM

下面，将对各个芯片的所需的供电电压、电流及其功耗进行分析，以便对网络安全加速卡的功耗有个整体把握，这些模块的功耗分析，就是上一章中电源网络分配的依据所在。

FPGA 芯片采用的是 Altera 公司的 EP2S60F672C4。该芯片是网络安全加速卡的核心。此芯片能否稳定工作将会影响到整个设备的正常运行。表 5.2 显示了对该 FPGA 的各种电源类型的分析。

表 5.2 EP2S60F672 电源类型分析表

引脚名	描述	输入电压范围
VCCINT	内部逻辑阵列供电引脚 也给输入缓冲提供电源	1.2V.可遵从 LVDS，LVPECL， HSTL，SSTL 以及差分 HSTL 和差 分 SSTL
VCCIO[1..8]	Bank1 到 Bank8 的所有 IO 脚供电 端，对所有的 IO 端的输出缓冲器 供电，也可对输入缓冲器供电	LVTTL，LVCOMS 1.5V，1.8V，2.5V，3.3V-PCI 3.3V-PCI-X
VCCPD[1..8]	对 IO 预驱动以及 3.3V 配置缓冲器 的引脚和 JTAG 引脚进行供电	3.3V
VREFB[1..8]	每一个 Bank 的参考电压输入端	不用的时候，需要将它们接 VCC 或 是接 GND
VCCA_PLL[1..6]	PLLs[1..6]的模拟电源	1.2V
VCCD_PLL[1..6]	PLLs[1..6]的数字电源	1.2V

另一个重要的模块就是硬件匹配查找模块。该模块采用 IDT 公司生产的 75P42100 芯片。该芯片是一款强大的硬件匹配搜索引擎。表 5.3 则详细列出了该模块的电流需求分析。表 5.4 详细列出了其电压需求分析。

表 5.3 CAM 的供电电流情况

类型	描述	工作时的电流量
IDD1	工作时的核电流量	100MHz 查找时，IDD1 为 3A(5A) 50MHz 查找时，IDD1 为 2.28A (3.8A)
IDD2 (VDDQ=2.5V)	工作时的 IO 电流量	100MHz 查找时，IDD2 为 0.36A(0.6A) 50MHz 查找时，IDD2 为 0.36A(0.6A)
IDD2 (VDDQ=1.8V)	工作时的 IO 电流量	100MHz 查找时，IDD2 为 0.3A(0.5A) 50MHz 查找时，IDD2 为 0.3A(0.5A)
IBAIS	工作时的基准电压 电流量	100MHz 查找时，IBAS 为 60mA(0.1A) 50MHz 查找时，IBAS 为 60mA(0.1A)

表 5.4 CAM 的供电电压情况

引脚名	描述	典型电压值	最大电压值
VDD	核电压输入	1.8V	1.89V
VDDQ	IO 引脚电压输入	1.8V 或 2.5V	1.89V 或 2.625V
VBIAS	器件内部信号基准电压	2.5V	2.625V

由表 5.3 和表 5.4 的分析可知，该模块在工作过程中可能出现的峰值功耗为 10W。连续正常工作下的功耗为：4.794W(VDDQ 为 1.8V 时)和 5.154W(VDDQ 为 2.5V 时)

网络安全加速卡的硬件转发模块采用 DDR2 的 SO-DIMM 模组来实现。该模组的功耗分析如表 5.5 所示：

表 5.5 DDR2 模组的功耗分析

类型	描述	典型电压值	最大电压值
VDD	供电电压	1.8V	1.9V
VDDL	DLL 供电压	1.8V	1.9V
VDDQ	Output 端供电电压	1.8V	1.9V
VREF	输入参考端电压	0.5*VDDQ	0.51*VDDQ
VTT	终端电压	VREF	VREF+0.04V
VDDSPD	SPD 芯片核电压	1.7V-3.6V	3.6V

该模组读取 Bank 时的最大功耗电流为 IDD7＝1.48A（工作频率为 200MHz，CL=3 时）。电源设计过程中一定要能提供最大的工作电流情况。

PHY 芯片的供电电压分布如表 5.6 所列，供电电流如表 5.7 所列。

表 5.6 PHY 的电压分布分析表

类型	描述	典型电压值	最大电压值
AVDD	模拟电压	2.5V	2.63V
VDDO	MAC 接口引脚电压	For all non-HSTL 电平,则用 2.5V,如果在 HSTL 电平下,则支持 1.4V 到 1.9V	2.63V
VDDOH	XTAL1,2 LED 以及 SEL_CLK 供电电压	2.5V	2.63V
VDDOX	MDIO INTn 125CLK RESETn JTAG 引脚供电电压	1.4V 到 2.5V	2.63V
DVDD	数字电压	1.0V 或 1.2V	1.32V

表5.7 PHY的电流分析表

类型	最大典型电流值
IAVDD	681mA
IVDDO,IVDDOH,IVDDOX	156mA
IDVDD	833mA

由以上两表可以得出PHY在最坏工作情况下的最大功耗为4.12W，正常工作时的正常功耗为3.6W。

最后，对于配置芯片EPCS64。其功耗分析比较简单：正常的输入电压为3.3V，最大输入电压为4V，最大输入电流为15mA，因此需求的功耗为60mW。

5.4 本章小结

本章系统探讨了PCB设计中的信号完整性问题以及设计中应该注意的技巧。同时对本系统中的高速收发器件电路的设计进行了详细的阐述，总结了高速接口设计的技巧和方法。最后对本系统的主要模块的硬件功耗做了详细的数据分析。这些数据分析是整个系统的电源设计的依据。

第6章 全文总结

6.1 总结

网络数据包分析处理是计算机网络设备研究领域的热门技术，并处于逐渐完善成熟的过程。由于网络数据流量的持续增大，导致诸如企业机房等网络节点的服务器和普通网管类设备面对如此庞大而高速的数据流往往不堪重负。如果网络节点无法及时对数据流作出分析和处理，则会造成严重的数据包丢失现象。因此，如何在网络节点处保证局域网数据安全、加快数据流的分析处理速度成为了一个研究的热点问题。

本文首先研究了网络隔离技术的基本原理和数据包分类技术原理，接着在对数据包分类技术中的硬件快速规则命中进行了深入的研究，并在充分吸收网络安全思想和快速数据包分析技术的基础上设计了一种可行的网络加速设备。本网络安全加速卡的设计优势为：

(1) 系统设计中充分吸取了网络安全方面的思想，通过对规则的命中可以直接丢弃非可信数据包，充分保证了进入局域网的数据安全可靠性。

(2) 采用 FPGA 硬件对数据包进行数据报文分析，相比软件的解析方式明显加快了数据报文的解析速度。

(3) 采用硬件进行规则命中和数据包转发。数据包的决策和转发速度相对于传统的服务器 CPU 仲裁的方式提高了一个数量级，明显地加快了网络节点处的数据包传输速率，避免了高速数据流因处理速率不够而丢包。通过网络安全加速卡来执行数据报文的分析和转发，让服务器从繁重的数据报文的分析转发工作中解脱出来，从而有足够的资源来执行网络管理中的其他事务，改善了网络传输的质量。

(4) 采用了最新的 PCIE 接口设计，保证了设备与目前 X86 体系的任何计算机系统相兼容。

(5) 保留了良好的升级空间，通过软件的升级可以在保持本硬件平台不变的基础上进一步开发一些诸如地址转换，状态跟踪，在线流量审计等应用。

(6) 系统设计了安全可靠的 bypass 功能电路，使系统在掉电等特殊情况下不影响网络的正常数据传输。

6.2 展望

由于互联网的发展日新月异，虽然国内参与互联网安全与网络优化方面研究的单位很多，但是在快速包处理技术和控制策略以及解决网络拥塞方面还有很多的不完善之处。作者关于网络加速设备的研究也有很多待完善之处：

(1) 本网络安全加速卡的数据解析指令不够灵活。虽然本文设计的硬件系统能够方便地对数据链路层到运输层做一些协议解析，但是对种类繁多且格式更加复杂的应用层协议显得力不从心。因此对支持灵活解析指令结构的研究当成为网络安全方面的研究重点。

(2) 本网络安全加速卡只能对一些固定的规则进行匹配命中，缺乏足够的灵活性。在实现更加灵活的模糊匹配规则方面，作者未作研究。支持更加灵活的匹配规则是下一步需要继续研究的方向。

(3) 本网络安全加速卡设计的最大吞吐量是8Gb/s。每秒的数据解析能力则比8Gb/S的吞吐量低4Gb/s到2Gb/s左右，这主要取决于大小包所占的比重。随着网络吞吐量的迅速增大，如何达到网络加速设备的更高吞吐量应该是今后工作的重点。

参考文献

[1] P. Gupta, N. Mc Keown. Algorithms for Packet Classification [J]. *IEEE Network*, 2001, 15(2): 24~32

[2] 李正茂.网络隔离理论与关键技术研究[D]. 上海: 同济大学.2006, 3

[3] 严蔚敏, 吴伟民. 数据结构(C 语言版). 北京: 清华大学出版社, 1997: 80~84

[4] 郑连清, 崔捷, 马哲元. 安全网络概论[M]. 北京: 清华大学出版社, 2004

[5] 龚俭, 陆晟, 王倩. 计算机网络安全导论[M]. 南京: 东南大学出版社, 2000. 8

[6] 高峰, 许南山. 防火墙包过滤规则问题的研究[J]. 计算机应用, 2003. 6

[7] 蔡淑珍, 陆阳, 陈蕾. 基于分布的嵌入式防火墙的设计与实现[J]. 计算机工程与应用, 2003(11): 162~164

[8] 谢树新. 浅谈防火墙系统的研究现状与发展趋势[J]. IT 技术, 2007. (6), 12-13, 8

[9] S. Deering, R. Hinden. Internet Protocol, Version 6 (IPv6) Specification [S]. *IETF, RFC 2460*, 1998

[10]S.Blake, D.Black, M.Carlson et al. An Architecture for Differentiated Services [S]. *IETF, RFC 2475*, 1998

[11]R. Braden, D. Clark, S.Shenker. Integrated Services in the Internet Architecture: an Overview [S]. *IETF, RFC 1633*, June 1994

[12]谢希仁. 计算机网络(第四版) [M]. 北京: 电子工业出版社, 2003, 232~240

[13]A.Conta, S.Deering.Generic. Packet Tunneling in IPv6 Specification [S]. *IETF, RFC 2473*, 1998

[14]B. Carpenter, C. Jung. Transmission of IPv6 over IPv4 Domains without Explicit Tunnels [S]. *IETF, RFC 2529,* 1999

[15]R. Gilligan, E. Nordmark. Transition Mechanisms for IPv6 Hosts and Routers [S]. *IETF, RFC 2893*, August 2000

[16]B. Carpenter, K. Moore. Connection of IPv6 Domains via IPv4 Clouds[S]. *IETF, RFC 3056*, 2001

[17]M. Roesch. Snort-Lightweight Intrusion Detection for Networks [M]. *In: Proceedings of the 13th Conference on Systems Administration.* Washington: USENIX, 1999.

[18]G. Apostolopoulos, D. Aubespin, V. Peris, et al. Design, Implementation and Performance of a Content-Based Switch [J]. *In: Proceedings of IEEE Infocom 2000*, March 2000.

[19]A. Cohen, S. Rangarajan, H.Slye. On the Performance of TCP Splicing for URL-aware Redirection [J].*In: Proceedings of the 2nd USENIX Symposium on Internet Technologies and Systems*, October 1999: 117~125

[20]T. Y. C. Woo. A Modular Approach to Packet Classification: Algorithms and Results [J]. *In: Gruein R ed. Proceedings of IEEE Infocomm 2000. San Francisco, CA: IEEE Computer Society Press*, 2000. 1210~1217

[21]K. Thompson, G. J. Miller, R. Wilder. Wide-area Internet Traffic Patterns and Characteristics [J]. IEEE Network, 1997, 11(6): 10~27

[22]T.Kijkanjanarat.Fast Routing Lookup and Packet Classification for Next- generation Router: [PhD Dissertation] [J]. *Polytechnic University*, 2002

[23]喻中超, 吴建平, 徐恪. IP 分类技术研究[J]. 电子学报. 2001, 29(2): 260~262

[24]冯东雷, 张勇, 白英彩. 线速数据包输入处理技术[J]. 计算机研究与发展, 2002, 39(1):

41~48
[25]李胜磊, 张德运, 刘刚. 基于 CAM 的高速 IP 数据包分类技术[J]. 计算机工程, 2004, 30(4): 81~82, 101
[26]付歌, 杨明福, 陈骏.基于 TCAM 的快速更新算法[J]. 计算机工程, 2003, 29(9): 19~21
[27]田立勤, 林闯. 报文分类技术的研究及其应用[J]. 计算机研究与发展, 2003, 40(06): 765～775
[28]单征, 赵荣彩, 张铮. 报文分类算法研究[J]. 计算机工程与应用, 2005, 41(7): 149~152
[29]朱秋香, 陶军. 流分类算法研究综述[J]. 小型微型计算机系统, 2004, 25(10): 1802~1810
[30]王燕. 实现 IPv6 数据包分类的算法研究[J]. 计算机应用, 2005, 25(11):2052~2054, 2520
[31]高蕾, 谭明峰, 龚正虎. IP 报文分类算法综述与评价[J]. 计算机工程与科学, 2006, 28(3): 70~73, 105
[32]赵国锋, 谷上宇, 温智宇. 一种多域数据包快速分类算法[J]. 通信学报, 2002, 23(12): 378~384
[33]陈骏, 杨明福. 基于 Trie 结构的并行多维数据包分类[J]. 计算机应用与软件, 2003, 20(11): 61~63
[34]V. Srinivasan, S. Suri, G.Varghese, M. Waldvogel. Fast and Scalable Layer four Switching [J]. *In: Proceedings of ACM Sigcomm'98,* 1998: 191~202
[35]M. M. Buddhikot, S Suri, M Waldvogel. Space Decomposition Techniques for Fast Layer-4 Switching [J]. *In: Proceedings of Conference on Protocols for High Speed Networks. Salem, MA, USA: Kluwer Academic Publishers*, 1999: 25~41
[36]邵华钢, 杨明福. 基于空间分解技术的多维数据包分类[J]. 计算机工程, 2003, 29(12): 123~124,172
[37]付歌, 杨明福, 王兴军. 基于空间分解的数据包分类技术[J]. 计算机工程与应用, 2004, 40(8): 63~65, 139
[38]A. Feldman and S.Muthukrishnan. Tradeoffs for Packet Classification [J]. *In: Proceedings of IEEE Infocom,* 2000: 193-202
[39]F. Baboescu, G.Varghese. Scalable Packet Classification [J].*In: Proceedings of ACM Sigcomm,*2001. 199~210
[40]韩晓非, 王学光, 杨明福. 位并行数据包分类算法研究[J]. 华东理工大学学报, 2003, 29(5): 504~508
[41]唐兴艳, 汪纪锋. 一种高效的多维数据包分类算法[J]. 重庆邮电学院学报(自然科学版), 2005, 17(6): 733~735, 758
[42]尚凤军, 潘英俊. 基于 XOR Hash 的快速 IP 数据包分类算法研究[J]. 计算机工程与应用, 2005, 41(8): 1~3, 77
[43]杨勇, 瞿中, 何江平. 基于散列查找的数据包分流算法研究[J]. 计算机工程与设计, 2005, 26(4): 927~929
[44]尚凤军,王海霞.基于完全无冲突哈希的IP数据包分类算法研究[J]. 计算机工程与应用, 2004, 40(34): 173~175
[45]P. Gupta, N. McKeown. Packet Classification on Multiple Fields [J]. *In: Proceedings of ACM Sigcomm'99*, 1999: 147~160
[46]P. Gupta, N. McKeown. Packet Classification Using Hierarchical Intelligent Cuttings [J]. IEEE Micro, 2000, 20(1): 34~41
[47]V. Srinivasan, S. Suri, G.Varghese. Packet Classification Using Tuple Space Search [J].*The*

ACM Sigcomm'99, 1999. 135~146

[48]J. van Lunteren, APJ Engbersen. Multi-field Packet Classification Using Ternary CAM[J]. *IEE Electronics Letters*, 2002, 38(1): 21~23

[49]T. V. Lakshman, D. Stidialis. High-Speed Policy-based Packet Forwarding Using Efficient Multi-dimensional Range Matching [J].*In: Proceedings of ACM Sigcomm'98*, 1998. 191~202

[50]Douglas E. Comer. 网络处理器与网络系统设计(英文版) [M]. 北京: 电子工业出版社, 2004. 158

[51]S. Iyer, R. Rao Kompella, A Shelat. ClassiPI: An Architecture for Fast and Flexible Packet Classification [J]. *IEEE Network*, 2001, 15(2): 33~41

[52]D. Shah and P. Gupta. Fast Updating Algorithms for TCAMs [J]. *IEEE Micro*, 2001, 21(1): 36~47

[53]R. Panigrahy, S. Sharma. Sorting and Searching Using Ternary CAMs [J]. *IEEE Micro*, 2003, 23(1): 44~53

[54]T. Lecroq. Experimental Results on String Matching Algorithms [J].*Software- Practice & Expierence*, 1995, 25(7): 727~765

[55]I. Sourdis. Efficient and High-Speed FPGA-based String Matching for Packet Inspection: [Master's Thesis] [J]. *ECE Dept, Technical University of Crete (TUC), Chania, Greece*, July 2004

[56]王诚, 吴继华, 范丽珍, 薛宁, 薛小刚. Altera FPGA/CPLD 设计(高级篇)[M].北京: 人民邮电出版社, 2005

[57]B. Bloom. Space/Time Trade-Offs in Hash Coding with Allowable Errors[J]. *Commmunication of the ACM*, 1970, 13(7): 422~426

[58]A. Aho, MJ. Corasick. Efficient String Matching: An aid to Bibliographic Search [J]. *Communications of the ACM*, 1975, 18(6): 333~340

[59]Long Bu,John A. Chandy. FPGA Based Network Intrusion Detection using Content Addressable Memories [J].In: Proceedings of the 12th Annual IEEE Symposium on Field-Programmable Custom Computing Machines (FCCM'04), 2004. 316~317

[60]Xu Ke, Wu Jian-ping, Yu Zhong-chao et al. A Non-Collision Hash Trie-Tree Based Fast IP Classification Algorithm[J]. *Computer Science and Technology*, 2002, 17(2): 219~226

[61]彭元杰. 高速电路信号完整性分析[D]. 长沙: 湖南大学, 2007.5

[62]张华. 高速互联系统的信号完整性研究[D]. 南京: 东南大学, 2005.1

[63]王诚, 吴继华, 范丽珍, 薛宁, 薛小刚. Altera FPGA/CPLD 设计(初级篇)[M].北京: 人民邮电出版社. 2005

[64]布达科(美), 译者:田玉敏等. PCI Express 系统体系结构[M]. 2005.11

致 谢

值此论文结束之际，回顾三年来的硕士学习过程，我要衷心的感谢所有帮助过我的老师、同学、朋友和亲人。

首先我要衷心的感谢我的导师吴谨教授在三年的学习生涯中给我的悉心指导和严格要求。吴老师在理论上的高深造诣和严谨务实的治学态度让我受益匪浅。吴老师不仅在学习和研究中给我提供了优良的环境，还在方法上给了悉心的指导和帮助。更重要的是吴老师在做人做事方面留给了我深刻的印象。她的无私奉献精神和优秀的品质深刻影响了我。在此，谨向吴老师致以诚挚的谢意！

感谢实验室的肖浩华硕士、陈宇硕士、姚欣硕士、文志科硕士、郑辉硕士、陈奕奕硕士、胡婷婷硕士，感谢他们在研究生期间给我指导和帮助。还要感谢邓艾、袁金娟、余潜玉、杨林，他们的理解、包容、信任、支持与帮助，伴我度过了人生中最重要的研究生时光。

特别感谢杨莘老师，江风先生，陈振家先生，魏毅先生，感谢他们在项目研发和论文设计过程中给我的点拨和指导。

最后感谢我妈妈，没有她的默默支持我不可能如此顺利的完成漫长的求学之路。她的言传身教让我受益匪浅，在此对她表示最衷心的感谢和祝福！

二零一零年四月于

武汉科技大学

研究生期间发表的论文

[1] 明幼林, 吴谨. 千兆网络数据包分析过滤采集系统设计. 信息技术. 2010.7

基于FPGA的网络安全加速卡研究与设计

作者：明幼林
学位授予单位：武汉科技大学