

学校编码: 10384

分类号 _____ 密级 _____

学号: 22120051302290

UDC _____

厦 门 大 学

硕 士 学 位 论 文

Linux 环境下基于 Intel 千兆网卡的高速数
据包捕获平台的研究

The Research of High-Performance Packet Capture
Platform Based on Intel Gigabit Network Card in Linux
Environment

刘 峰

指导教师姓名: 黎忠文 教授

专 业 名 称: 计算机系统结构

论文提交日期: 2008 年 5 月

论文答辩时间: 2008 年 月

学位授予日期: 2008 年 月

答辩委员会主席: _____

评 阅 人: _____

2008 年 5 月

厦门大学学位论文原创性声明

兹提交的学位论文，是本人在导师指导下独立完成的研究成果。本人在论文写作中参考的其他个人或集体的研究成果，均在文中以明确方式标明。本人依法享有和承担由此论文产生的权利和责任。

声明人（签名）：

年 月 日

厦门大学学位论文著作权使用声明

本人完全了解厦门大学有关保留、使用学位论文的规定。厦门大学有权保留并向国家主管部门或其指定机构送交论文的纸质版和电子版，有权将学位论文用于非赢利目的的少量复制并允许论文进入学校图书馆被查阅，有权将学位论文的内容编入有关数据库进行检索，有权将学位论文的标题和摘要汇编出版。保密的学位论文在解密后适用本规定。

本学位论文属于

- 1、保密（ ），在 年解密后适用本授权书。
- 2、不保密（ ）

（请在以上相应括号内打“√”）

作者签名：

日期： 年 月 日

导师签名：

日期： 年 月 日

厦门大学博硕士学位论文摘要库

摘 要

线速地实时处理网络报文流是入侵检测系统、网络协议分析、网络防火墙、高性能通信系统、高性能路由器、主机路由器，以及其他网络监视系统必须满足的首要条件。

当前大部分的入侵检测系统、防火墙等网络监视及防御系统都是基于运行在 PC 架构上的 Linux 平台。

百兆的网络环境下，普通的网卡和 Libpcap 接口就足以保证运行在 Linux 用户空间的网络报文处理程序线速地捕获网络报文。如：tcpdump、ethereal 和 snort 等程序。然而千兆网络时代的到来，使得在通用操作系统下远达不到线速捕获网络数据包。本文实现了一种高性能的网络数据包的捕获平台，省去了数据包在内核和用户空间传递过程中的拷贝，缓存，以及系统调用过程，从而使得网络报文的处理性能有很大提升。

本人主要工作如下：

1. 对 Linux 网络协议栈进行了分析，并对传统报文捕获平台 Libpcap 进行了分析研究。
2. 深入分析了 Intel 千兆网卡驱动及相关的中断机制，软中断机制及内存映射机制。
3. 在深入分析 Linux 2.6 内核的基础上，实现了一种新的基于 Intel 千兆网卡的高性能的网络数据包捕获平台 ZeroP。ZeroP 由两部分组成，一部分是位于内核空间的驱动模块，另一部分是用户空间的兼容 Libpcap 的接口。
4. 搭建了 CISCO 测试平台，对 ZeroP 平台进行了性能及兼容性测试，并对测试数据进行了分析。

关键字： NAPI；中断；E1000 网卡驱动；Mmap

厦门大学博硕士论文摘要库

Abstract

Many network monitoring systems are based on wire-speed packet capture and real-time analysis. Such as, intrusion detection systems, protocol analysis, network firewall, high-performance communications systems, high-performance router, host routers and other network monitoring systems.

At present, most of the intrusion detection system, firewall, and so on are based on PC architecture running on the Linux platform.

In fast network environment, the common interface card and Libpcap is enough to guarantee that the user space application, on the linux system, can wire-speed capture network packets. such as: tcpdump, ethereal and snort, and so on. As the time of the gigabit network, it is impossible for general operating system to capture packet or send packet at wire speed. This paper designs a high-performance zero-copy platform, it decreases the number of the copy and the buffer between kernel and user space, reduces the overhead of the system call, thus the network packet processing performance is greatly improving.

Main tasks:

1. Analysed the Linux network protocol stack, analyzed the traditional packet capture platform Libpcap.
2. Analysed the Intel Gigabit Ethernet driver deeply, the interrupt mechanism, the soft interrupt mechanism and the memory mapping mechanism.
3. We implemented a new high-performance network packet capture platform ZeroP based on Intel-based Gigabit Ethernet in the Linux 2.6 kernel. The ZeroP platform contains two components: a device driver module in kernel-space and a user-space library compatible with Libpcap.
4. Built CISCO test platform, tested the performance and the compatibility, analysed the test data.

Key Words: NAPI; Interrupt; E1000 driver; Mmap

厦门大学博硕士学位论文摘要库

目 录

摘 要.....	i
Abstract.....	iii
目 录.....	v
Contents.....	ix
第一章 绪论.....	1
1.1 课题背景	1
1.2 数据包捕获技术国内外研究现状	1
1.2.1 国外现状	1
1.2.2 国内现状	3
1.3 本文组织安排	6
第二章 数据包捕获技术研究.....	7
2.1 Linux 标准协议栈的数据包捕获过程分析.....	7
2.2 基于 Libpcap 的数据包捕获技术介绍.....	8
2.3 数据包捕获过程的影响因素	9
2.3.1 系统调用	10
2.3.2 数据包拷贝和数据校验	10
2.3.3 硬件中断	10
2.3.4 协议处理	11
2.4 零拷贝思想介绍	11
2.4.1 零拷贝思想引出	11
2.4.2 零拷贝技术实现中的关键问题	12
第三章 Intel 千兆网卡相关技术研究	13
3.1 Intel 千兆网卡接收描述符介绍	13
3.1.1 接收描述符结构	13
3.1.2 接收描述符队列环结构	13
3.1.3 接收数据结构整体框架图	14

3.2 Intel 千兆网卡中断节制机制介绍	15
3.2.1 基本的中断处理流程	15
3.2.2 带中断节制机制的中断处理流程	15
3.2.3 中断节制机制的折中之道	16
3.2.4 中断节制机制的特性	17
3.3 NAPI 技术简单介绍	21
3.4 Linux 内存映射机制分析	22
3.4.1 物理内存管理	22
3.4.2 物理地址和虚拟地址之间的映射	22
3.4.3 Linux 内核内存使用	24
3.4.4 I/O 访问	27
3.5 红黑树原理介绍	28
3.5.1 红黑树插入操作	30
3.5.2 红黑树删除操作	31
3.6 Libpcap 接口分析	33
3.6.1 Libpcap 应用程序框架	33
3.6.2 打开网络设备	35
3.6.3 用户应用程序接口	37
第四章 基于 Intel 千兆网卡的零拷贝网络数据包捕获平台的研究与实现	41
4.1 设计思想	41
4.1.1 操作系统的选择	41
4.1.2 内核空间和用户空间的数据拷贝问题	41
4.1.3 缓冲环管理问题及同步问题	43
4.1.4 网络接口的选择	44
4.2 整体框架	44
4.3 基于 E1000 网卡的 NAPI 系统框架及工作原理	45
4.3.1 核心数据结构	45
4.3.2 基于 E1000 网卡的 NAPI 总体逻辑流程	46
4.4 ZeroP 平台的系统框架及工作原理	49
4.4.1 系统初始化流程	49
4.4.2 网络数据包接收处理流程	49
4.5 具体实现	51

4.5.1 E1000 网卡标准驱动的修改.....	51
4.5.2 上层接口函数	52
第五章 基于 Intel 千兆网卡的零拷贝网络数据包捕获平台的性能测试与数据分析.....	53
5.1 测试参数	53
5.2 测试环境	53
5.3 测试网络拓扑图	54
5.4 测试项目	55
5.5 测试结果	55
5.5.1 一个发包机下的接收性能	55
5.5.2 两个发包机下的接收性能	57
5.5.3 三个发包机下的接收性能	59
5.5.4 兼容性测试	61
5.6 测试结果分析	62
第六章 总结与前景展望.....	65
参考文献.....	67
攻读硕士期间发表学术论文.....	71
致谢	72

厦门大学博硕士论文摘要库

Contents

Chinese Abstract	i
English Abstract	iii
Chinese Contents	v
English Contents	ix
Chapter1 Introduction	1
1.1 Research Background	1
1.2 Packet Capture Technology at Domestic and Foreign on the Status quo	1
1.2.1 Foreign Status.....	1
1.2.2 Domestic Status quo.....	3
1.3 Structure of Thesis	6
Chapter2 Research of Packet Capture Technology	7
2.1 Process Analysis of Packet Capture Based on Linux Protocol Stack	7
2.2 Introduction of Packet Capture Technology Based on Libpcap	8
2.3 Impacting Factors of Packet Capture	9
2.3.1 System Call	10
2.3.2 Packet Copy and Packet Checksuming.....	10
2.3.3 Hardware Interrupt.....	10
2.3.4 Protocol Processing	11
2.4 Presentation of Zero-copy Technology	11
2.4.1 Introduction of Zero-copy Technology.....	11
2.4.2 The Key Issues in Zero-copy Technology.....	12
Chapter3 Research of Intel Gigabit Ethernet Technology	

.....	13
3.1 Receive Descriptor of Intel Gigabit Ethernet.....	13
3.1.1 Structure of Receive Descriptor.....	13
3.1.2 Structure of Receive Descriptor Ring	13
3.1.3 Framework Figure of Receiving Data Structure	14
3.2 Introduction of Interrupt Moderation of Intel Gigabit Ethernet	15
3.2.1 Basic Interrupt Handling Processing	15
3.2.2 Interrupt Handling Processing with Interrupt Moderation	15
3.2.3 The Compromise of Interrupt Moderation	16
3.2.4 Features of Interrupt Moderation.....	17
3.3 The Simple Introduction of NAPI Technology.....	21
3.4 Analysis of Linux Memory Mapping Mechanism.....	22
3.4.1 Physical Memory Management	22
3.4.2 Mapping between Physical Address and Virtual Address.....	22
3.4.3 Use of Linux kernel Memory	24
3.4.4 Access of I/O Memory.....	27
3.5 Intruduction of RED-BLACK TREE Principle	28
3.5.1 Insert Operation of RED-BLACK TREE.....	30
3.5.2 Delete Operation of RED-BLACK TREE	31
3.6 The Analysis of Libpcap Interface	33
3.6.1 Framework of Libpcap Application	33
3.6.2 Open Network Device	35
3.6.3 Application Interface	37
Chapter4 Research and Implementation of Zero-copy Packet Capture Platform Based on Intel Gigabit Ethernet	41
4.1 General Idea of the Design.....	41
4.1.1 Choice of Operating System.....	41
4.1.2 Data Copy Issue between Kernel Space and User Space	41
4.1.3 Buffer Ring Management Issue and the Synchronization Issue.....	43
4.1.4 Choice of Network Interface	44
4.2Overall Framework	44

4.3 System Framework and Principle of NAPI Based on E1000	
Card.....	45
4.3.1 Core Data Structure.....	45
4.3.2 General Logic Processing of NAPI Based on E1000 Card.....	46
4.4 System Framework and Principle of The ZeroP Platform	49
4.4.1 System Initialization Processing	49
4.4.2 Packet Receiving Processing	49
4.5 Implementation	51
4.5.1 Modification of E1000 Card Driver	51
4.5.2 Upper Interface Functions	52
Chapter5 Performance Testing and Data Analysis of Zero-copy	
Packet Capture Platform Based on Intel Gigabit Ethernet	53
5.1 Test Parameters.....	53
5.2 Test Environment.....	53
5.3 Test Network Topology	54
5.4 Test Items	55
5.5 Test Results	55
5.5.1 Receiving Performance Using One Sender.....	55
5.5.2 Receiving Performance Using Two Senders	57
5.5.3 Receiving Performance Using Three Senders	59
5.5.4 Compatibility Test.....	61
5.6 The Analysis of Test Results.....	62
Chapter 6 Summary and Future Works.....	65
References	67
Publications.....	71
Acknowledgements.....	72

厦门大学博硕士学位论文摘要库

Degree papers are in the "[Xiamen University Electronic Theses and Dissertations Database](#)". Full texts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to etd@xmu.edu.cn for delivery details.

厦门大学博硕士学位论文摘要库