

基于NP的流量测量系统

学生：陈衡

指导老师：李英壮、于广辉

摘要

随着互联网规模飞速扩大和异构化发展，其控制机制和行为特征日趋复杂。为了认识和理解现代互联网的行为特征和性能表现、保证和提高现有网络服务质量、推动互联网和信息基础结构的健康发展、有必要建立建立一套完整的网络测量和管理体系。

目前常规的网络流量测量系统需要使用能产生符合NetFlow或sFlow等工业标准的采样数据包的交换机或路由器，有成本高和可扩展性差的缺点。并且一旦对流量采集的要求发生变化，则很难对采集数据的方式和数据包的具体格式做修改。

而网络处理器的应用为下一代网络向高性能、可编程可扩展方向发展提供了极佳的选择。本文描述了一种运用IXP2400网络处理器和sFlow技术，在IXA（internet exchange architecture）架构上实现网络流量测量系统功能的设计。

内容

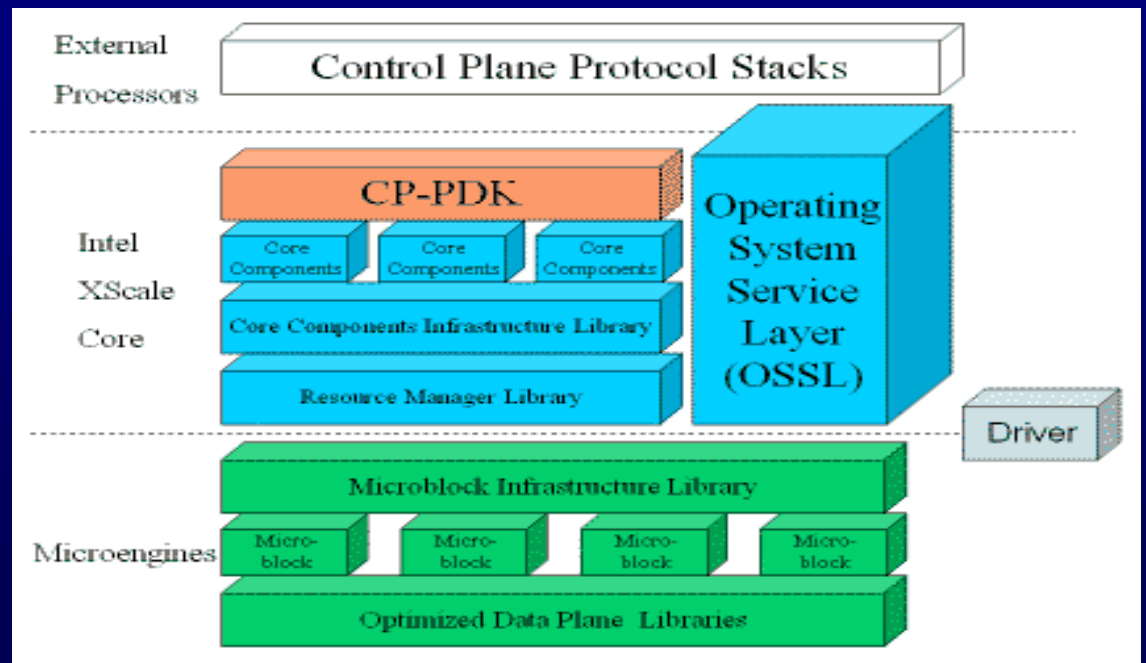
- 背景：网络处理器的发展应用和基于IXA架构流量测量系统的提出
- 网络处理器和Intel IXA架构技术的基本知识及应用研究现状
- 测量理论的介绍和几种流量采样方案比较
- IXP2400网络处理器上流量测量系统的软硬件框架设计和具体实现
- 以后要做的工作

网络处理器基本概念和结构

- ④ 网络处理器(Network Processor, 简称NP)是一种专门用于执行网络数据处理和转发的高速可编程处理器, 被经过专门设计和高度优化来完成各种网络功能。它兼具ASIC (application specific integrated circuit) 的高性能和RISC (reduced instruction set computing) 通用CPU的可编程灵活性, 并有低成本和低风险的优势。网络处理器大多通过采用多个微处理器并行工作来提高系统的吞吐量。由于其可对数据包进行2至7层的处理, 并能外接通用CPU开发网络管理的附加功能, 它正使得网络服务设备的设计和实现越来越规范, 成为一种网络系统设计的技术理念。

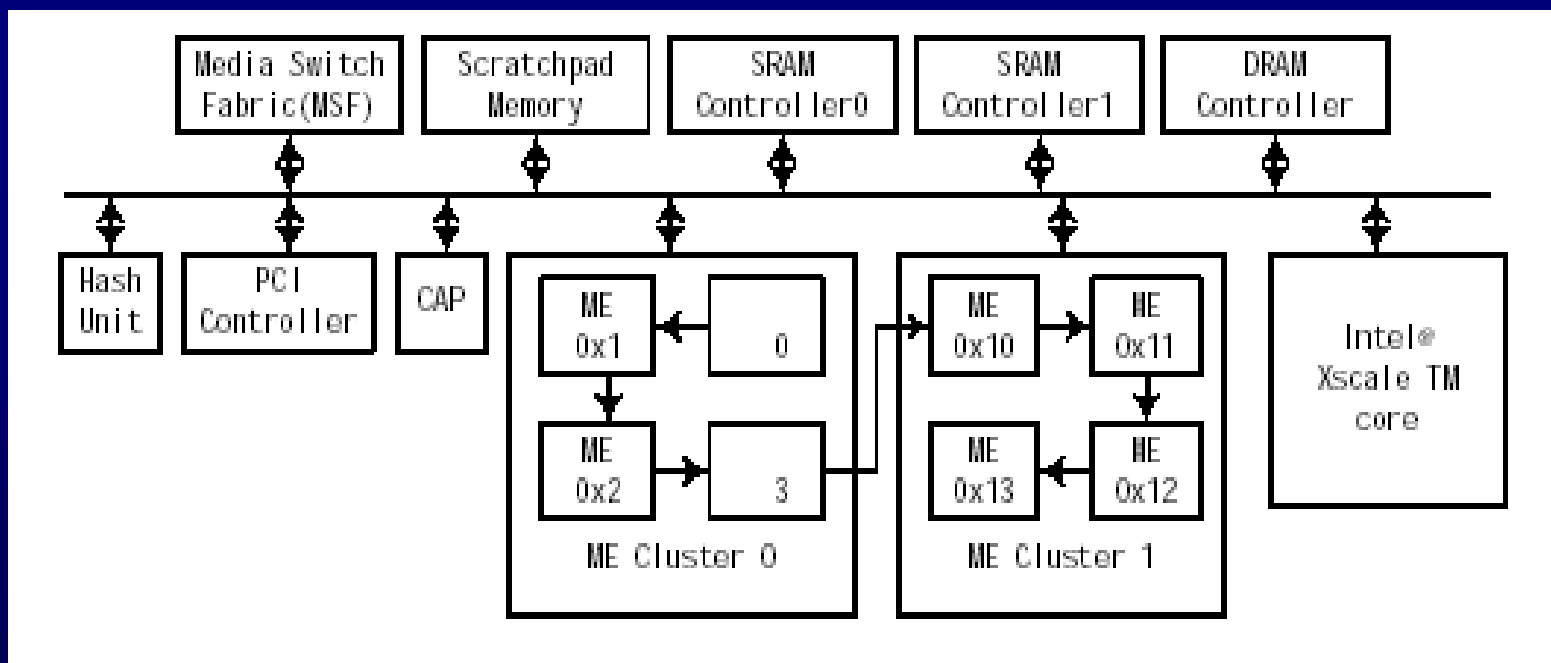
Intel 的 IXA 架构技术

- ① Intel IXA 即 Intel 互联网交换架构 (Intel Internet Exchange Architecture), 是一个网络处理架构
- ② 微引擎技术
- ③ Intel Xscale 技术
- ④ Intel IXA 移动架构



IXP2400内部功能模块结构图

- ① 核心处理器Xscale，8个微引擎microengines(ME)，ShaC功能单元，MSF(media switch fabric)，PCI、SRAM、DRAM控制器以及其他一些提升性能的协处理单元



流量测量

流量工程包括流量测量、流量控制和流量管理等内容。主要研究对操作性网络中流量处理过程的优化。其优化的中心是当网络的总体容量仍有富余时，要减少个别容量的超载。使网络中的链路、路由器和交换机等成分既不会过度使用，也不会未充分使用。

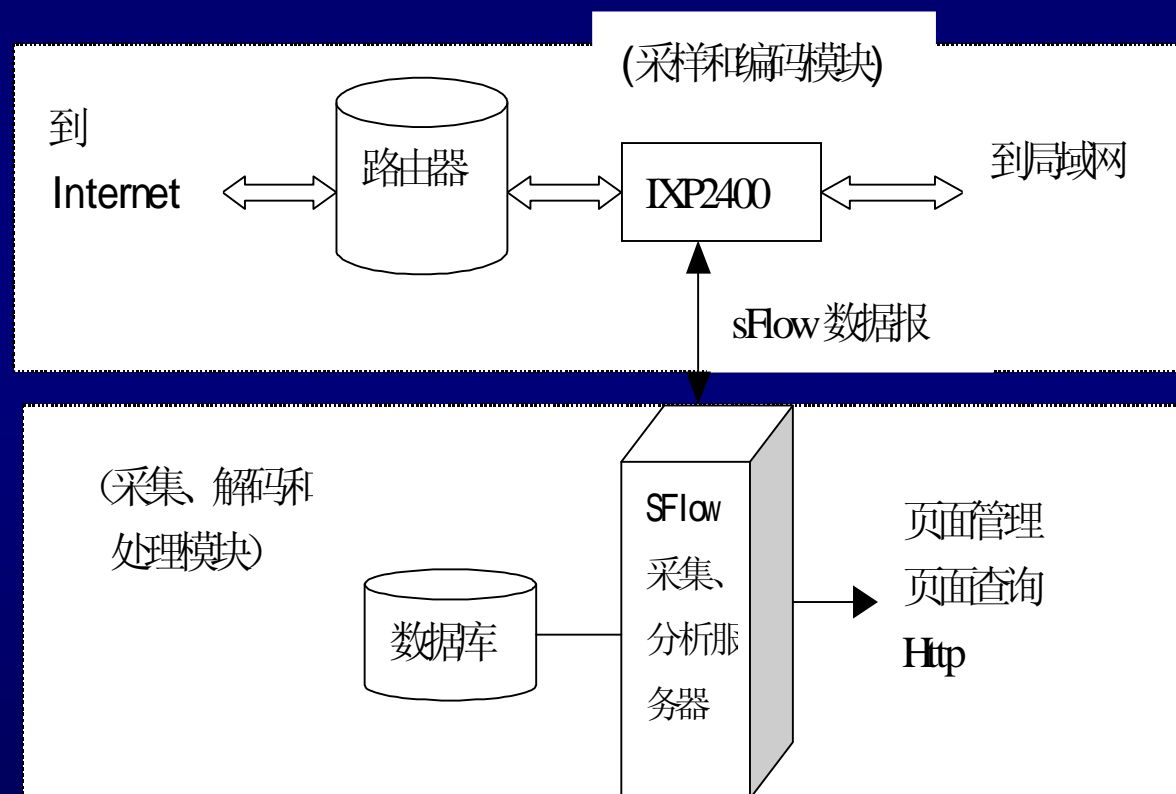
流量测量是流量工程的基础，也是网络监测和管理研究的重要组成部分。它为各种流量工程的性能分析应用提供原始数据。

几种常见的流量测量技术

- ① RMON :提供基于接口、链路的测量信息
 - ① NetFlow : 提供基于流的统计信息
 - ① sFlow: 同样是对数据包流的采样信息，与NetFlow主要采样包格式的区别
- 此外还有juniper的cflow、华为的netscream等

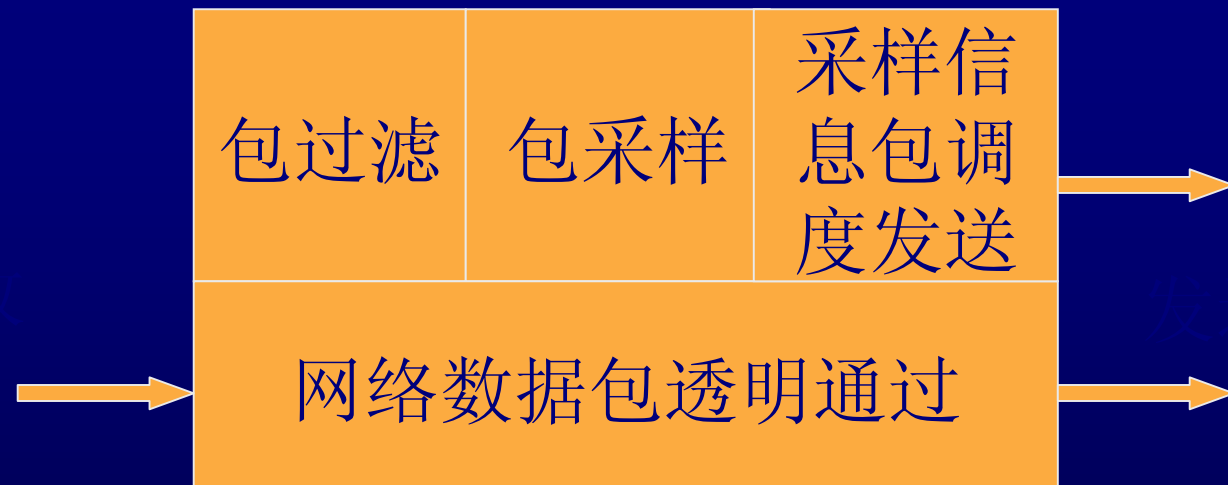
系统的整体硬件框架

包括流量信息的采样编码模块和采集解码模块两部分



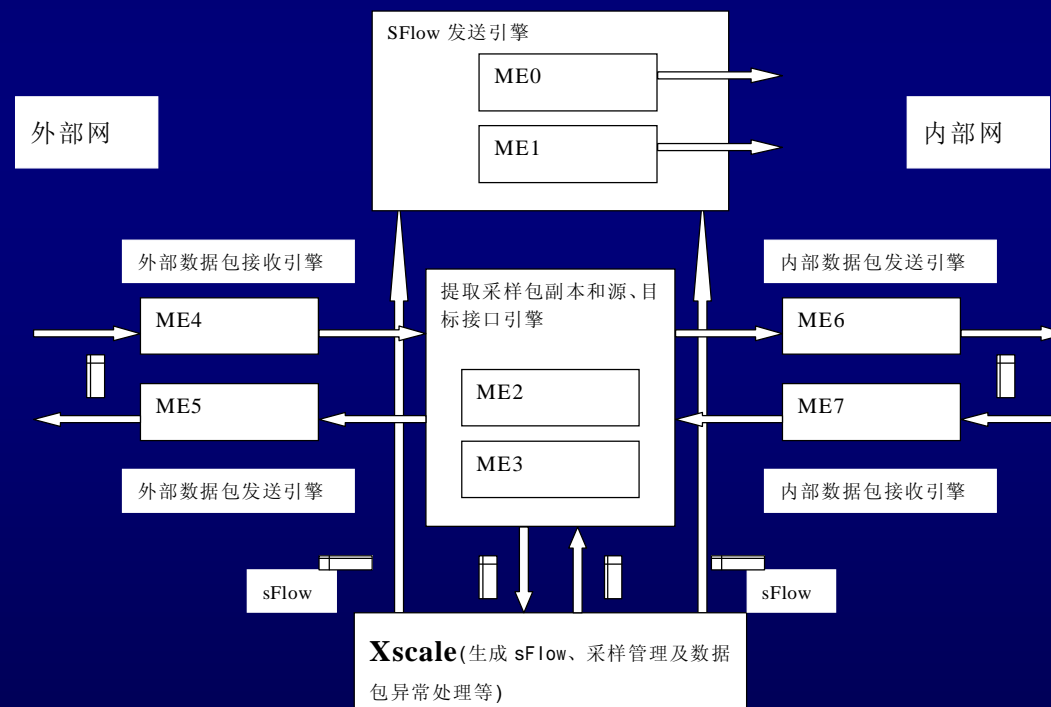
采样功能的子模块

在IXP2400上要实现的数据的采样和编码模块为设计实现和研究的重点。其中又包括实现四个子功能：基本的网络数据通路、包过滤、包采样和采样包统计发送的调度。



IXP2400上数据采样子系统的软件任务分配

- ④ 其中用于内外网数据包收发四个微引擎，两个微引擎发送数据包副本和接口号给Xscale，两个微引擎负责发送Xscale生成的sFlow包至sFlow采集服务器。



IXP2400上数据包流动过程

MSF元件上RBUF收到MAC包 (mpacket) ->接收包微引擎接收重组该包 ->重组成的数据包被接收引擎放到DRAM上的接收处理环上 ->采样处理微引擎处理该数据包 ->处理过的数据包被处理引擎放到DRAM上的处理发送环上 ->发送微引擎将数据包拆分成MAC包 ->MAC包被放到MSF元件上的TBUF上并被发送出去。



IXP2400上数据包的接收和过滤

IXP2400上接收的数据包由多个MAC包（mpacket）组成，每个MAC包被硬件标示为起始包start-of-packet (SOP)、终止包end-of-packet (EOP)、都是，或都不是。接收线程读取接收状态字重组该包。

目前的过滤主要是为采样功能服务的，检查包的二层帧头信息，丢弃有错误和不符合标准的包。

数据包采样机制

当数据包到达IXP2400的接口，首先过滤器判断其是否应被丢弃。若不被丢弃，则判断器将包指定给IXP2400的另一个接口，同时决定是否采样该包。一个决定采样间隔的计数器当有数据包通过IXP2400就减1，一旦该计数器变成零则执行一次采样并恢复计数器原值。采样间隔数应被设为一个随机整数。另有一个计数器Total_Packets统计所有可以被采样的数据包。

执行采样涉及到复制数据包头，及提取包的特性。每执行一次采样，计数器Total_Samples加1，Total_Samples统计采样执行的次数。采样功能把样本结果发给sFlow Agent的功能函数处理，样本包括数据包信息、Total_Packets值和Total_Samples值。



采样数据信息格式

```
typedef __declspec(packed) struct
{
    unsigned int agent_address,
    sequence_number,
    uptime;
    union{
        unsigned int sequence_number,
        source_id,
        sampling_rate,
        sample_pool,
        drop,
        input,
        output;
        struct{
            unsigned int length,
            protocol,
            src_ip,
            dst_ip,
            src_port,
            dst_port,
            tcp_flags,
            tos;
        }packet_data;
        struct{
            char* src_user,
            dst_user;
        }extend_data;
    }sample;
} sample_data_t;
```

采样信息数据报发送调度



SRAM中的hash表

是否满足表中发送条件，若是则发送采样信息数据报

SDRAM中分配采样数据报队列空间

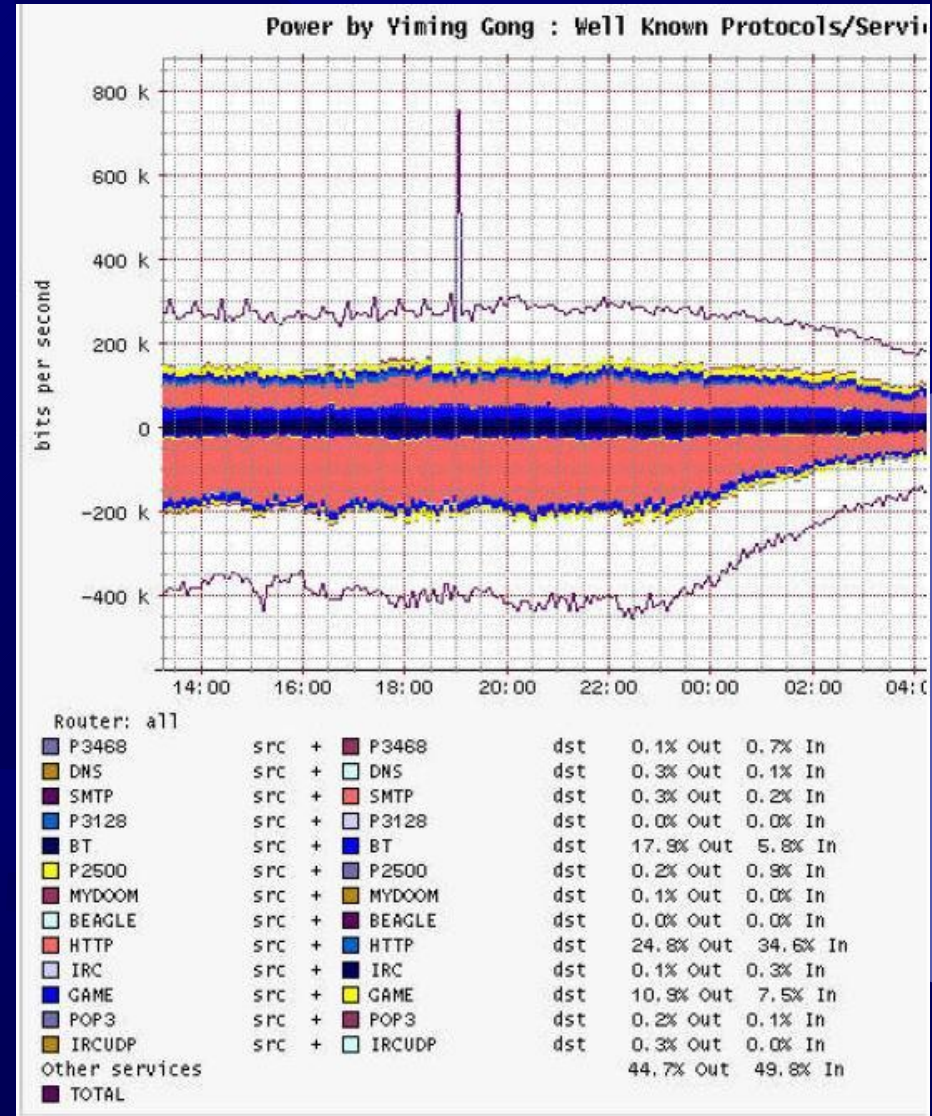
MSF单元

微引擎微块程序

采样信息数据报的采集、存储和解码分析

采样产生的数据报格式采用RFC3176规范的sFlow。因此，可以兼容基于sFlow的软硬件分析设备。

使用现在已经很成熟的sFlow分析技术工具(如InMon Traffic Server、sFlowTrend、netscout、CommonTraffic、Gentek Flow@tch等)即可实现数据存储、分析和分类整理功能，其它许多厂家也提供类似的采集软件。



以后要做的工作

- ① 加强完善 IXP2400 中对通过流量的过滤等功能，可以在流量测量的基础上实现流量控制，完成流量工程
- ② 可增加多个 IXP2400 协同处理，扩展系统性能
- ③ 开发图形化交互界面，以便进行更方便的配置和管理