## 2.4具体研究开发内容

### 2.4.1零拷贝网卡读写功能实现；

考虑到评析系统在多个地方需要面对大规模的网络数据报文，经过测试发现，传统的操作系统在数据报文协议栈处理上消耗了大量的CPU时间，并且单通道网卡中断的SMP系统中，这些消耗的CPU无法均衡到各个CPU上，从而使得系统的总体吞吐量有限，为了提升系统吞吐量，我们采用一种绕过系统内核处理流程的报文处理机制，通过修改网卡驱动，在驱动中加入报文直接读写的接口，该接口可以在操作系统内核读取到数据报文之前提供一个报文读写的机会给应用程序，并且通过标准使该接口满足NETMAP函数库的调用规则，由此用户的应用程序可以直接使用NETMAP的标准库函数进行编程，既可以避免应用开发的复杂度，又能提供足够的报文吞吐性能。

### 2.4.2数据报文捕获功能实现

完成零拷贝网卡驱动修正后，便可以直接使用NETMAP的库函数。考虑到系统运营时将面对庞大的数据流冲击，还需要提供良好的接口给报文透传模块。由此我们设计报文捕捉模块使用NETMAP零拷贝操作的模块来绕开操作系统内核的报文协议栈，将网卡的数据报文直接映射到报文捕捉空间，由此解决大数据流量冲击问题。在数据报文透传接口方面，采用消息队列的通讯机制，并且建立简单的负载分发方式，让上层应用可以是根据自己的运算需求和SMP系统的相关参数来定制报文捕获模块的接口，由此充分的挖掘系统资源的使用率，简化应用构建和运营的难度。

### 2.4.3快速报文透传功能实现

由于报文绕过了系统协议栈直接到达报文捕捉系统，这些报文当中含有业务报文（ＤＮＳ报文）、非业务报文，业务报文需要提交给上层处理，而对于非业务报文，如果直接抛弃就会影响非业务流程（网管，http通讯等）的正常进行。为此我想设计了一种方法让报文透传系统能把非业务报文正常提交给系统协议栈。首先修改网卡驱动，在报文透传系统与系统协议栈之间建立一条通道，并提供必要的读写接口；然后报文透传系统在过滤出非业务数据后，把数据通过这个通道重新提交到系统协议栈，这样就重建了原有的通讯流程，而且避免了重新实现系统协议栈的重复工作，而上层应用感觉不到有任何不同。

### 2.4.4路由收敛功能实现

由于报文绕过了系统协议栈直接到达报文捕捉系统，业务报文发送需要根据实现环境发向不同的目标，如果使用系统协议栈，性能受到严重影响，为此需要实现一种路由收敛策略，让报文能像系统协议栈那样把数据路由到特定的目标。首先需要实时同步系统路由信息，实时监听路由配置的改变，然后把路由以降序方式排列，这样可以使用二分法快速查找到合适的路由。

## 技术关键点及难点攻关

### 2.7.1技术关键点及亮点

|  |  |
| --- | --- |
| 功能点 | 实现方式 |
| 高速率大容量网络数据的捕捉 | 系统采用直接硬件读写的方式来捕获流经网络数据接口的数据报文，这种方式有着极好的数据处理效率，结合单臂流量的重组功能，基本可以在服务器上处理10GE接口链路上的数据流。 |
| 快速报文重组 | 使用最短协议栈的处理规则，支持是用单臂流量来预测整个连接的交互过程，单服务器数据处理上可以与零拷贝网卡读写的速率相匹配。 |
| 路由收敛 |  |

### 2.7.2技术难点攻关

|  |  |
| --- | --- |
| 功能点 | 实现方式 |
| 高速的网络数据报文捕获和处理 | 很多的时候评析系统的报文处理前端要单独面对整个评析对象的全部数据流量（在一些不方便做负载的场合，或者小预算场景），为了使得评析系统能够尽量适应各种环境，需要开发一套完整的超高速的报文处理系统，这种处理速率有时甚至需要和纯硬件设计的交换机和路由器进行速率匹配。 |
| 高频率大剂量数据存储和分析 | 评析系统的报文数据处理部分会产生大量日志数据，这样的大数据需要采用合理有效的方式进行存储和管理。同时，对这样的海量数据进行采集分析，也需要运用时下领先的数据挖掘技术，数据仓库技术，精确挖掘用户行为，详尽的数据报表，为网络运营商提供运营决策依据。 |