

信号流

文档版本 01

发布日期 2014-8-25



版权所有 © 华为技术有限公司 2014。 保留一切权利。

非经本公司书面许可,任何单位和个人不得擅自摘抄、复制本文档内容的部分或全部,并不得以任何形式传播。

商标声明



HUAWEI和其他华为商标均为华为技术有限公司的商标。

本文档提及的其他所有商标或注册商标,由各自的所有人拥有。

注意

您购买的产品、服务或特性等应受华为公司商业合同和条款的约束,本文档中描述的全部或部分产品、服务或 特性可能不在您的购买或使用范围之内。除非合同另有约定,华为公司对本文档内容不做任何明示或默示的声 明或保证。

由于产品版本升级或其他原因,本文档内容会不定期进行更新。除非另有约定,本文档仅作为使用指导,本文 档中的所有陈述、信息和建议不构成任何明示或暗示的担保。

华为技术有限公司

地址: 深圳市龙岗区坂田华为总部办公楼 邮编: 518129

网址: http://www.huawei.com 客户服务邮箱: support@huawei.com

客户服务电话: 4008302118

信号流 前言

前言

本章主要介绍了信号流在设备内部的转发路径以及在转发过程中各组件单元对信号流的 处理动作,阅读本章有助于进一步了解数据报文从进入设备到转发出设备整个周期内的 处理流程。 信号流 目录

目 录

前言	ii
1通用业务流程简介	1
2控制平面信号流	
3 数据平面信号流	
4 监控平面信号流	9

信号流 1 通用业务流程简介

1 通用业务流程简介

概念类比

路由器是Internet网络中为IP数据包找到路径的设备,如同邮政系统中的邮件分发中心,只不过路由器的操作对象不是邮件而是数据包。路由器的核心功能可以归纳为四个字,即"寻址转发"。

邮政系统为了实现邮件的正确分发,首先要具备一个"邮件收集中心",将需要分发的邮件汇总收集,同时还需要具备一个"邮件查询分发中心",能够按照邮件的不同目的地查询地址后归类分发。"邮件收集中心"相当于一个数据平面,主要负责邮件的传递:"邮件查询分发中心"相当于一个控制平面,负责邮件的寻址和分发。

逻辑架构

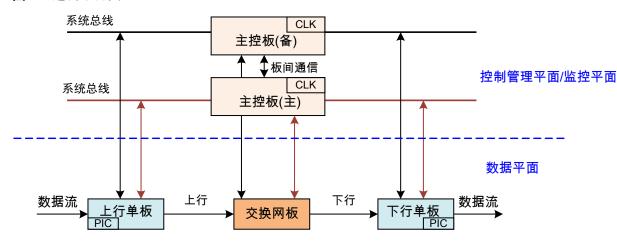
类比邮政系统,路由器要实现"寻址转发"的功能需求,也需要具备如上两个功能平面,此外,为了保证"寻址转发"的安全稳定运行,路由器还需要提供一个监控级平面对系统环境进行监控。综上,路由器在逻辑架构上有三个功能平面:数据平面、控制管理平面和监控平面。

- 数据平面:相当于"邮件收集中心",为数据包提供进出设备的接口,同时负责完成路由器的数据报文高速处理和内部无阻塞交换,包括报文的封装与解封装、以太/MPLS/IPV4/IPV6报文的转发处理、QoS与调度处理、内部高速交换以及各种统计。
- 控制管理平面:相当于"邮件查询分发中心",不仅实现"寻址",还是整个系统的中枢神经系统,通过系统总线实现对各单元的管理控制,包括路由计算、LSP的生成、生成组播、单播和MPLS转发表、协议和信令的处理、系统状态的配置与维护管理、系统状态报告与控制等。
- 监控平面:保障上述两个平面的安全稳定运行,主要完成系统的环境监控,包括电压检测、系统上下电控制、温度监测与风扇控制等,在出现单元故障的情况下及时隔离故障单元,保障系统其它部分的正常运行。

路由器的逻辑架构如图1-1所示。

信号流 1 通用业务流程简介

图 1-1 逻辑架构图



∭说明

图中箭头方向表示数据流的流向。

物理架构

在物理架构上,数据平面的转发功能主要由上下行业务接口板以及交换网板实现,控制管理平面和监控平面的监控管理功能主要由主控板以及各单板上的控制管理单元组成。物理架构与逻辑架构的对应关系可参见《VRP 硬件描述-机箱-概述-系统架构》。

信号流 2 控制平面信号流

2 控制平面信号流

控制平面作为系统的神经中枢,集成了系统的控制管理功能,通过系统总线实现对各单元的管理控制。

控制平面由主控板的各管控单元及各单板上的管控单元构成。主控板上主要有3类功能单元:系统管理控制单元、系统时钟单元和系统维护单元。前两种属于控制平面功能单元,最后一种属于监控平面功能单元(该单元在后面的"监控平面信号流"章节进行介绍),主控板与单板通过系统总线与系统背板相连实现各单元之间的互通。

控制平面信号流

控制平面信号流原理框图如图2-1所示。

信号流 2 控制平面信号流

主控板(主) 上行单板 信令信号流 系统管理控制 操作维护信 单元 管理控制单元 号流 时钟单元 时钟信号流 (CLK) 系统时钟单元 (CLK) 交换网板 系 统 背 管理控制单元 板 主控板(备) 时钟单元 信令信号流 (CLK) 系统管理控制 操作维护信 单元 号流 下行单板 时钟信号流 系统时钟单元 管理控制单元 (CLK) 时钟单元 (CLK)

图 2-1 控制平面信号流原理框图

控制平面负责处理3类信号流:信令信号流、操作维护信号流和时钟信号流。其中前两种信号流由系统管理控制单元负责处理,时钟信号流由系统时钟单元负责处理。控制平面信号流为双向信号流。

- 信令信号流: 指一些携带信令信号的协议类的主机报文,如路由协议等。
 - 单板上送主控板方向:完成数据计算。由于信令类报文不能直接作用于数据平面指导转发,必须先经过系统管理控制单元的解析,数据平面才能识别,因此信令类信号必须上送到主控板计算后下发到数据平面。
 - 主控板下发单板方向:完成转发表更新。主控板完成计算后会将计算结果下发 给对应单板,从而指导数据平面的业务数据转发处理。

以路由计算为例:报文由端口进入,在单板的NP上被识别出该报文是需要上送主控板计算的信令报文,则由NP直接上送主控板进行计算,主控板计算后将目的IP和下一跳地址下发到NP的路由表中,报文根据路由表的目的IP和下一跳地址进行转发。

操作维护信号流:操作类信号流主要指用户在设备上配置的数据,包括配置、调试和查询等操作类数据;维护类信号流主要指系统本身产生的一些日志和告警等用于维护的信息。

信号流 2 控制平面信号流

- 单板上送主控板方向: 指维护类的信号流,系统管理控制单元会通过系统总线 定期收集各模块产生的日志和告警等维护类信息并存于"信息中心",为设备 维护查询时提供有力依据。

如:用户需要查询当前活动的告警,用户通过执行display alarm active命令,经控制平面解析后作用于信息中心,将信息中心已收集的当前所有活动的告警信息显示给用户。

- 主控板下发单板方向:指操作类信号流,由于控制平面本身提供了软件层面与外界通信的接口,并在内部提供了与主机软件的通信接口以及与其他单板和其他模块通信的LAN Switch接口,这些接口构成了逻辑上互通的"网络",所以用户的操作命令均配置在控制平面,并由控制平面将不同类型、不同重要程度的信息分别输出到不同的信息通道,最终通过互通的各接口下发到各单元模块。如:用户配置给1号槽位单板下电,用户执行power off slot 1命令,这些命令是基于软件层面的配置,由于控制平面与软件平面有互通接口,控制平面会解析命令信息,随后通过内部的带外通信接口,最终将命令信息作用于1号单板上进行下电。
- 时钟信号流: 指同步以太时钟信号或者1588时钟信号。
 - 单板上送主控板方向: 提取时钟信号,报文在流经上行单板时,上行单板会提取时钟频率信号并上送主控板的系统时钟单元,系统时钟单元开始进行时钟选源。
 - 主控板下发单板方向:恢复系统时钟,主控板完成选源后,会恢复出一个系统时钟,同时下发给所有单板,从而保证报文通过出接口转发到下游设备时,会携带统一的系统时钟信号。

可靠性

主控板和时钟均采用1:1冗余备份工作方式,主备主控板之间通过板间通信相互进行状态监视,一旦主用主控板出现故障,备用主控板自动升级为主用,从而大大提高了设备的可靠性。

信号流 3 数据平面信号流

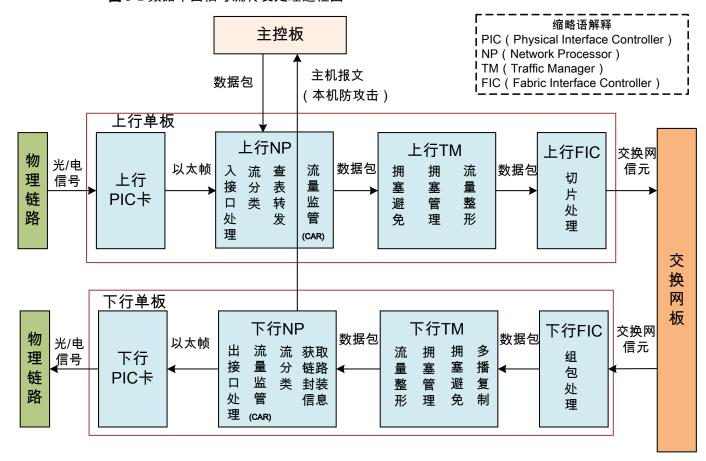
3 数据平面信号流

数据平面主要承载了数据报文的高速处理和转发,主要处理业务数据信号流。

数据平面信号流流向可以概括为如下几个阶段:物理链路—>上行单板处理—>交换网板交换—>下行单板转发—>物理链路。

数据平面信号流的转发处理过程如图3-1所示。

图 3-1 数据平面信号流转发处理过程图



信号流 3数据平面信号流

上行信号流处理过程:

1. 上行PIC卡为信号流提供入接口:物理链路的光/电信号由上行单板的PIC卡进入设备,在PIC卡上主要完成物理层时钟频率的提取、链路层协议的协商并将数据报文打包成以太帧,之后流入上行NP。

□ 说服

PIC卡接收的业务数据可能是高速业务如POS、以太业务等,也可能是低速业务如ATM、TDM、PPP业务等,PIC卡针对不同的业务会进行不同的处理,将数据报文打包成不区分高低速业务的以太帧。对于打包前的不同业务介绍这里不做赘述。

- 2. 上行NP查表为信号流找到目的出接口:
 - a. 入接口处理:信号流流经NP时,首先经过入接口处理,对数据报文携带的链路层协议进行解析,识别出报文类型。
 - b. 流分类:根据入接口的配置划分出数据报文的优先级并进行流分类和流标记。
 - c. 查表转发:根据数据包信息(如目的MAC地址、目的IP地址、MPLS标签等), 查对应转发表项,从而获得出接口或下一跳信息。

□ 说明

通常数据报文分为3类:基于MAC地址的2层报文、基于MPLS标签的2.5层报文和基于IP地址的3层报文。入接口处理模块识别出报文类型后,查表转发模块则会根据不同的报文类型查找对应的转发表项(2层报文查找MAC转发表,2.5层报文查找MPLS转发表,3层报文查找路由转发表),从而使得不同类型的报文都可以找到对应的出接口或下一跳信息。

- d. 流量监管:上行NP根据入接口配置的CAR,或者是上行流分类中配置的CAR 进行上行限速。对于转发行为为丢弃的报文,在上行NP被丢弃,不经过CAR 统计。
- 3. 上行TM对信号流进行流队列处理:
 - 拥塞管理和拥塞避免:上行TM监督网络资源的使用情况,当发现拥塞有加剧的 趋势时采取主动丢弃报文的策略并通过调整流量来解除网络的过载。
 - 流量整形: 在网络发生拥塞时,将报文放入队列中缓存,并采取某种调度算法安排报文的转发次序,进行流量整形。

经过OoS的相关处理后,报文将有序的发送给上行FIC。

4. 上行FIC对信号流进行分片:由于交换网板是基于固定信元长度进行交换的,类似于ATM信元,所以数据包送入交换网板前要对其进行切片。上行FIC将数据包进行切片,并封装成交换网信元后发送给交换网板。

交换网板的信号流处理过程:交换网板主要负责各LPU之间的数据交换。交换网板接收已分片的信元,经过缓存和调度后,进入SFU(Switch Fabric Unit)上的交换单元。SFU的交换单元采用了多平面负载分担,Cell单元能够被均匀的分配到每个交换平面上进行交换,这样不仅实现交换平面的负载分担,更有利于系统的容错处理。交换单元根据数据报文的目的端口将报文调度到目的出口,交换网板即完成数据交换并将数据报文发送到下行单板。

□ 说明

报文是否经过交换网处理与设备的硬件架构有关,当前VRP-X8/X16属于交换网架构设备,报文均走交换网处理,VRP-X1/X2/X3属于无交换网的架构,数据报文通过单板上的SM(Switch Module)确定链路连接关系,完成转发。

下行信号流处理过程:

1. 下行FIC将交换网信元进行组包:交换网信元经过交换网板后,发送给下行FIC,在下行FIC上进行重新组装成数据包。

信号流 3数据平面信号流

2. 下行TM对信号流进行流队列处理和调度:下行TM根据出接口配置的流队列参数对数据报文进行流队列调度,包括流量的整形、监管,拥塞避免等处理,之后将数据包发送给下行NP。

∭说明

对于多播报文,对应的下行出接口是多个,因此需要对下行的数据报文进行多播复制,多播复制的动作也是在下行TM上实现的。

- 3. 下行NP查表获取链路层封装信息:下行NP主要也承担三个功能,获取报文链路封装信息,流处理模块和出接口处理模块。
 - a. 下行NP查表获取链路封装信息并做报文封装处理。对于不同的报文,在下行NP上的处理方式也各不相同,如对于2层报文存在QinQ场景时,需要在下行NP上封装Vlan Tag;对于2.5层MPLS报文,在下行NP封装MPLS标签;对于3层IP报文,在下行NP查ARP表项,封装MAC头等。
 - b. 流处理模块根据出接口的配置划分报文优先级进行流分类和流标记。
 - c. 下行NP根据出接口下配置的CAR,或者下行流分类中配置的CAR进行下行限速。
 - d. 对于需要上送主控板的主机报文,在此直接上送主控板,不需要上送主控板的 纯转发报文,直接传递给出接口模块处理。

□□说明

由于主控板CPU的处理能力有限,为了提高设备安全性,避免网络中的恶意攻击报文或缺陷报文上送CPU导致CPU占用率过高而影响正常业务,需要对上送CPU的报文进行限制。在报文上送CPU之前进行检查,比如只上送需要CPU处理的报文,不上送有缺陷的报文,限制报文上送CPU的速率,从而实现本机防攻击,保证CPU高效处理正常。

- e. 出接口处理模块为数据报文封装新的二层头,转发给下行PIC卡。
- 4. 下行PIC卡为信号流提供出接口: PIC卡最终将数据包转换成光/电信号由指定出接口发送到物理链路。

信号流 4 监控平面信号流

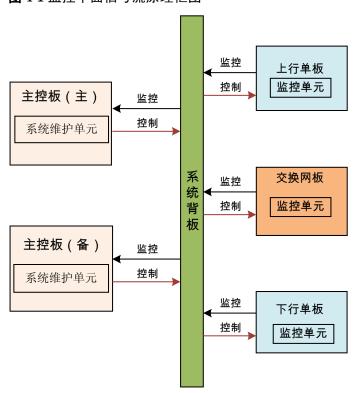
4 监控平面信号流

监控平面完成系统的环境监控,在出现单元故障的情况下及时隔离故障,保障系统其它部分的正常运行。控制平面和监控平面的有机配合,为系统的运行维护及安全稳定性能提供了有力的系统保障。

与控制平面类似,监控平面也是由主控板的"系统维护单元"和以及单板和交换网板的"监控单元"构成,通过系统背板进行板间通信。

监控平面信号流原理框图如图4-1所示。

图 4-1 监控平面信号流原理框图



监控平面信号流也是双向信号流。

信号流 4 监控平面信号流

● 单板上送主控板方向:实现监控,主控板的系统维护单元会通过监控总线 (Monitorbus)定期收集系统各单元运行数据,并根据各单元运行状态产生控制信息。

● 主控板下发单板方向:实现控制,主控板将产生的控制信息下发到各模块,如检测各单板在位情况、风扇调速等等,并能通过系统加载总线,实现从远端或近端对系统各单元进行测试或在线升级。