

认识网络设备

网络设备知识点：

了解网络设备的主要构成（交换机、路由器、防火墙）

了解框式设备的主要硬件模块（MPU 主控板、SFU 交换网板、LPU 接口板）

了解网络设备逻辑构架的 3 个层面（数据平面、控制管理平面和监控平面）

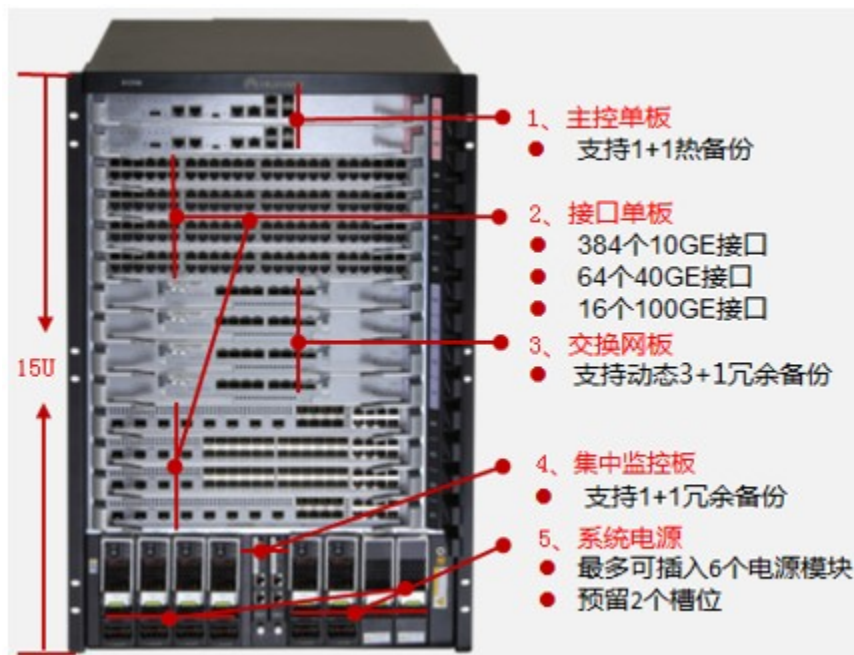
了解设备 2 种报文分类的处理

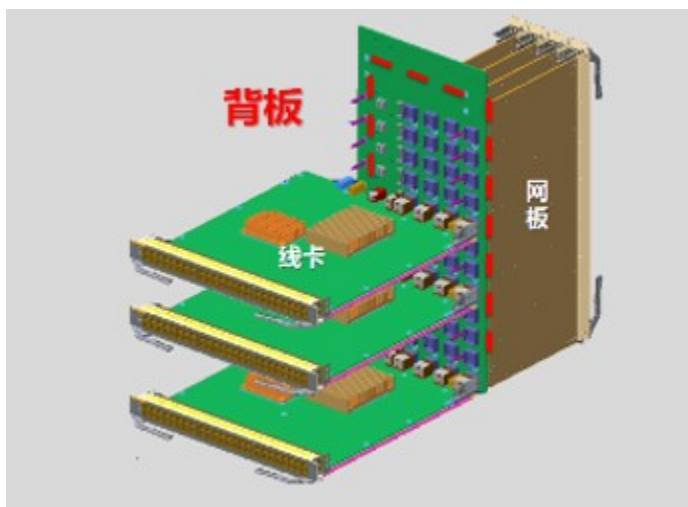
业务报文：交换网板交由下行接口板处理，不经主控板

协议报文：上送主控板由主控板的 CPU 进行处理

=====

框式设备







世界虽然很大，但是当前网络的高速发展使得我们足不出户都能了解全球最新的资讯、随时和朋友与同事交流、网络极大的丰富人们的沟通和生活。未来将会是一个万物互联的智能世界。

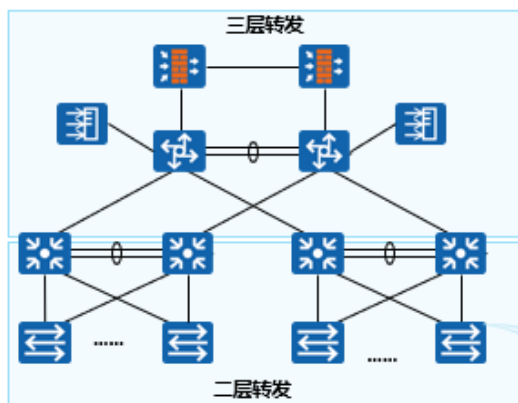
网络的搭建离不开网络设备，物理连接，以及设备之间的多种协议。其中在实现网络互通时，最常见的网络设备是路由器和交换机。

如今在各种级别的网络随处可见各种低、中、高端的路由器、交换机，种类繁多，这些不同种类的设备内部是如何工作、处理数据报文的呢？

本课程将会以路由器、交换机为例学习网络设备的工作机制、报文转发流程。



网络设备



- 网络基础设施由交换机、路由器、防火墙等构成，这些设备日复一日地接收、发送数据，从一个接口收到的数据如何经过设备内部转发到另外一个接口？
- 设备有哪些组件构成？这些组件如何协同工作？。

交换机内部如何把报文从一个接口转发到另外一个接口？



框式设备硬件模块



网络设备典型面板图

为方便理解网络设备内部的各个功能模块，以S12700E-8为例讲解典型网络设备的构架：

- 主控板(MPU, Main Processing Unit)：负责整个系统的控制平面。
- 交换网板(SFU, Switch Fabric Unit)：负责整个系统的数据平面。数据平面提供高速无阻塞数据通道，实现各个业务模块之间的业务交换功能。
- 接口板(LPU, Line Processing Unit)：线路处理单元是物理设备上用于提供数据转发功能的模块，提供不同速率的光口、电口。
- 交换网板、接口板上都有自己的管理芯片，与主控板共同组成整个设备的控制管理平面。



主控板



主控板

主控板提供了整个系统的控制平面和管理平面。

- 控制平面完成系统的协议处理、业务处理、路由运算、转发控制、业务调度、流量统计、系统安全等功能。
- 管理平面完成系统的运行状态监控、环境监控、日志和告警信息处理、系统加载、系统升级等功能。



交换网板



交换网板

交换网板提供整个系统的数据平面。接口板、主控板之间通过交换网板完成通信。



接口板



接口板

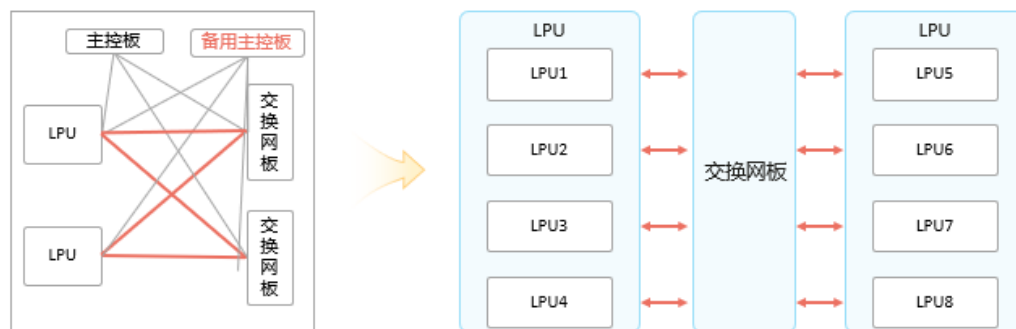
接口板提供了不同类型（光口、电口），不同速率的接入接口，通过分布式数据平面对数据进行转发。

盒式设备



不同于框式设备，盒式设备的各个业务模块并不是独立的硬件模块，而是集成在一个框内。

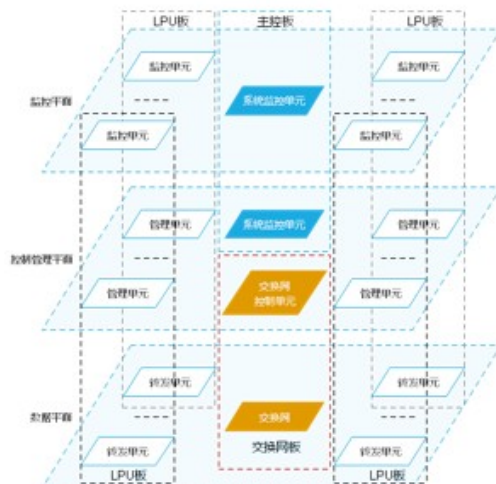
模块连接逻辑图



- 框式设备各个模块分为不同的单板，单板之间通过框式设备内部的连接进行通信。
- 盒式设备内部集成了这些模块，各个模块之间同样也是通过内部连接进行通信。
- 接口板和接口板之间通过交换网板连接了起来，接口板之间的通信统一经由交换网板进行转发。

网络设备逻辑构架

网络设备从逻辑上可以分为以下三个平面：数据平面、控制管理平面和监控平面。



以交换机为例，设备可以分为以下的三个平面（参考 RFC7426 中对网络设备平面划分）：

交换机转发平面：转发平面提供高速无阻塞数据通道，实现各个业务模块之间的业务交换功能。交换机的基本任务是处理和转发交换机各不同端口上各种类型的数据。L2/L3/ACL/QoS/组播/安全防护等各种具体的数据处理转发过程，都属于交换机转发平面的任务范畴。

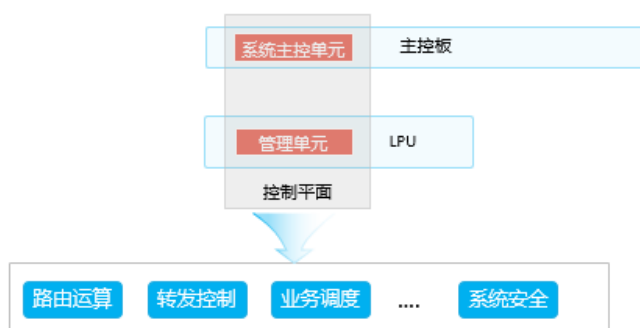
交换机控制平面：控制平面完成系统的协议处理、业务处理、路由运算、转发控制、业务调度、流量统计、系统安全等功能。交换机的控制平面用于控制和管理所有网络协议的运行。控制平面提供了数据平面数据处理转发前所必须的各种网络信息和转发查询表项。

交换机管理平面：管理平面完成系统的运行状态监控、环境监控、日志和告警信息处理、系统加载、系统升级等功能。交换机的管理平面是提供给网络管理人员使用 TELNET、WEB、SSH、SNMP、RMON 等方式来管理设备，并支持、理解和执行管理人员对于网络设备各种网络协议的设置命令。管理平面必须预先设置好控制平面中各种协议的相关参数，并支持在必

要时刻对控制平面的运行进行干预。
在华为产品具体的实现上，部分系列产品将功能组合区分为数据平面、管理平面和监控平面。

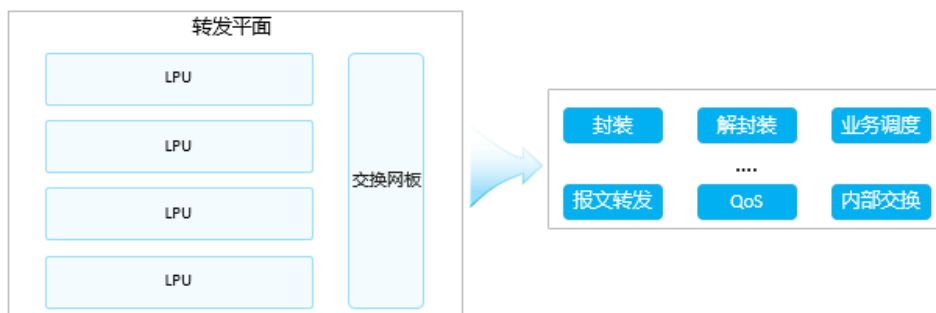
控制平面

- 设备的控制平面由主控板以及接口板的管理单元组成。
- 控制管理平面完成系统的控制管理功能，是整个系统的中枢神经系统。控制平面完成系统的协议处理、业务处理、路由运算、转发控制、业务调度、流量统计、系统安全等功能。交换机的控制平面用于控制和管理所有网络协议的运行。控制平面提供了数据平面数据处理转发前所必须的各种网络信息和转发查询表项。



转发平面

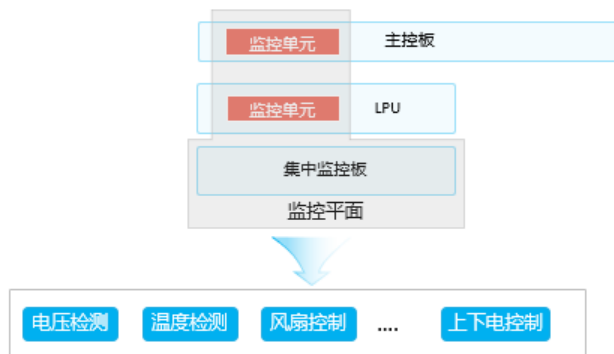
- 设备的转发平面由交换网板以及接口板组成。
- LPU上存在FPE（转发引擎），其本质也是一个交换芯片，完成本接口板端口之间的交换。
- 数据平面完成数据报文的高速处理和内部无阻塞交换。包括报文的封装与解封装、IPv4/IPv6/MPLS转发处理、QoS与调度处理、内部高速交换以及各种统计。





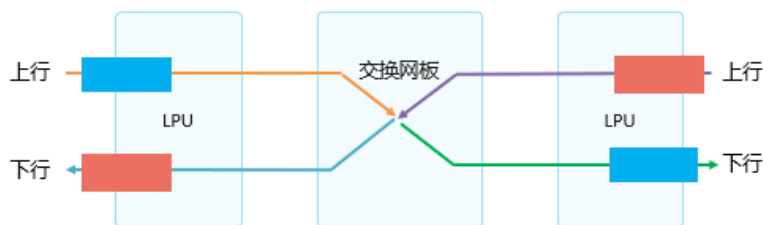
监控平面

- 监控平面由主控板、接口板的监控单元构成，部分框式设备还会存在单独的集中监控板(CMU)。
- 监控平面独立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，以保证系统的安全稳定运行，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。



报文转发上行、下行

以交换网板为中心，可将报文在设备的行程一分为二，前半程称为“上行”，后半程称为“下行”。

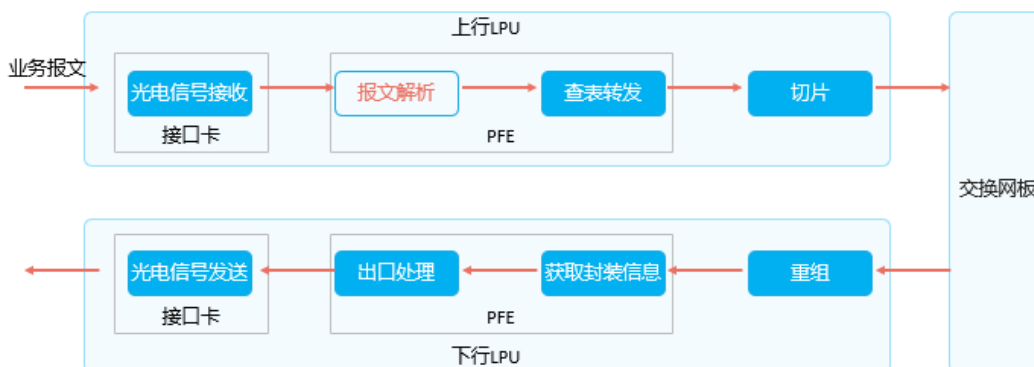


- 设备处理报文分类：一种是业务报文、一种是协议报文。
- 对于业务报文设备只会进行转发，从一个接口进入之后依据转发表项从另外一个接口发送出去。
- 协议报文（如 ARP、OSPF、BGP 等协议的报文）设备在收到之后会交由控制层面进行处理，如 ARP 报文交由控制层面处理、判断之后决定是否回应，是否学习 ARP 报文中的源 MAC、源 IP。



业务报文转发处理流程

业务报文从接口进入上行接口板处理之后，通过框式交换机内部总线交由交换网板，交换网板交由下行接口板处理之后从接口发出去。

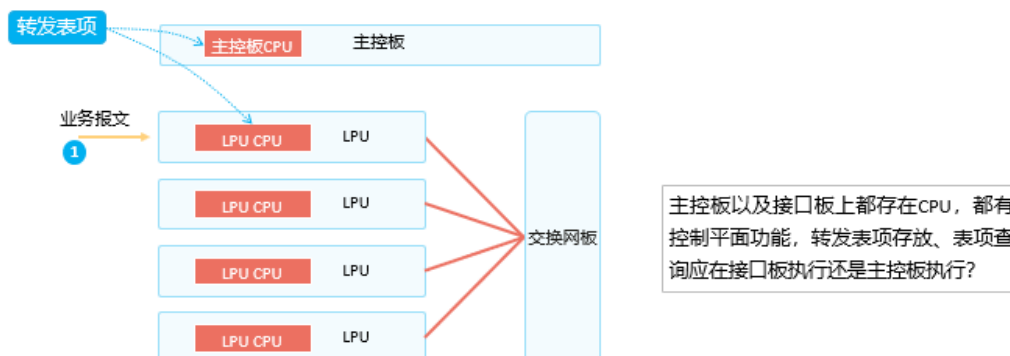


- PFE (Packet Forwarding Engine)：包转发引擎。
- 业务报文：服务、应用在交互过程中涉及的报文。
- 切片：把报文送往交换网板之前，进行切片处理，也就是把报文按一定粒度进行切片，切成固定长度。
- 重组：将交换网板发送过来的已经切片的报文进行重新组合。



确定报文出口 (1)

- 当报文从接口板进入时，设备需要依据转发表项（IP路由表、MAC地址表等）确定报文的出接口（对于框式设备需要确定下行接口板）。
- 报文到达交换网板时已经明确了出接口、下行接口板，因此表项查询需要在上行接口板的处理过程中完成。

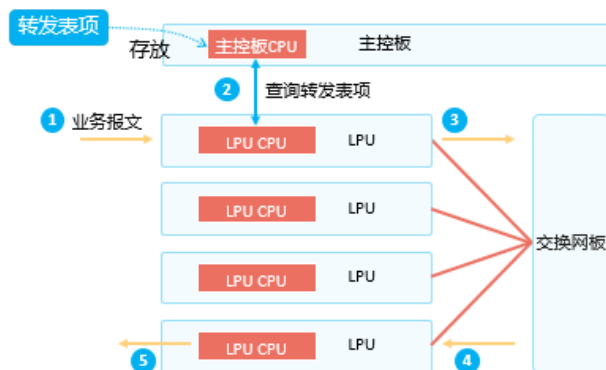


- 二层转发查询 MAC 地址表，三层转发查询三层路由表。



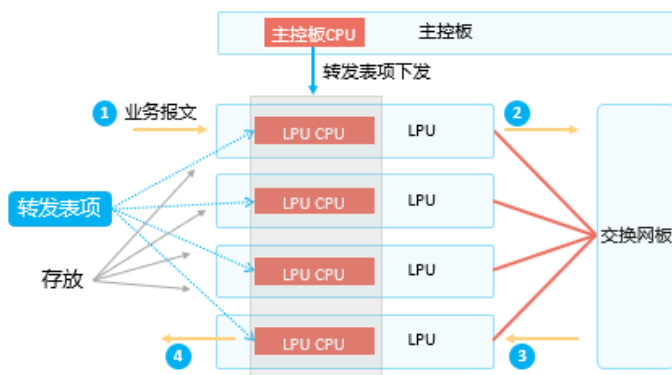
确定报文出口 (2)

- 转发表项存放在主控板上，报文进入接口板之后，接口板从主控板处查询表项。
- 每次转发都需要与主控板进行通信，转发效率低，报文时延增加，对高速率接口板而言转发速率严重下跌。



确定报文出口 (3)

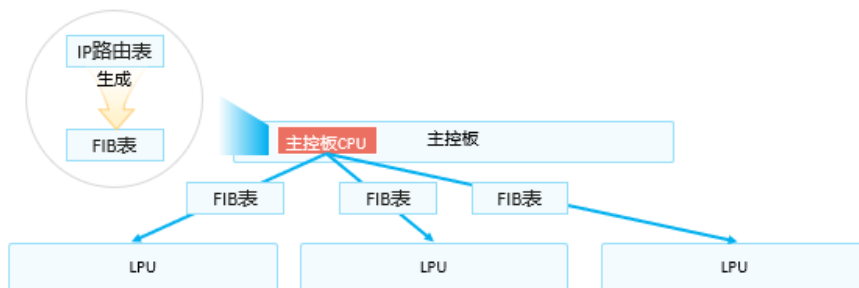
- 转发表项存放在接口板上，报文进入接口板之后直接在接口板完成报文查询，报文转发效率高。
- 所有接口板上都要存储转发表项，控制平面资源占用率高。





转发信息

- 高端设备业务报文不经过主控板CPU处理，由接口板提供转发信息查询。
- 接口板上存在的转发信息并非存在于主控板上的转发表项（IP路由表、MAC地址表...）。主控板生成转发表项之后，生成对应的转发信息下发在接口板。



以IP路由表为例，路由表生成之后，主控板根据路由表生成FIB表项（Forwarding Information Base）并下发到接口板，接口板根据FIB表进行转发。



硬件转发

- 接口板执行转发的部件为包转发引擎PFE（Packet Forwarding Engine），通常为NP或者ASIC芯片，报文直接由接口板独立完成转发，无需主控板参与。
- 高端框式设备把转发层面和控制层面分配在不同的组件，控制层面组件（主控板）负责运行转发相关协议、维护转发表项，转发平面组件（接口板）依据控制层面下发的转发信息能够独立完成转发工作，互不影响，控制层面组件高负载时并不会影响转发平面的正常工作，这种工作机制被称为转控分离。

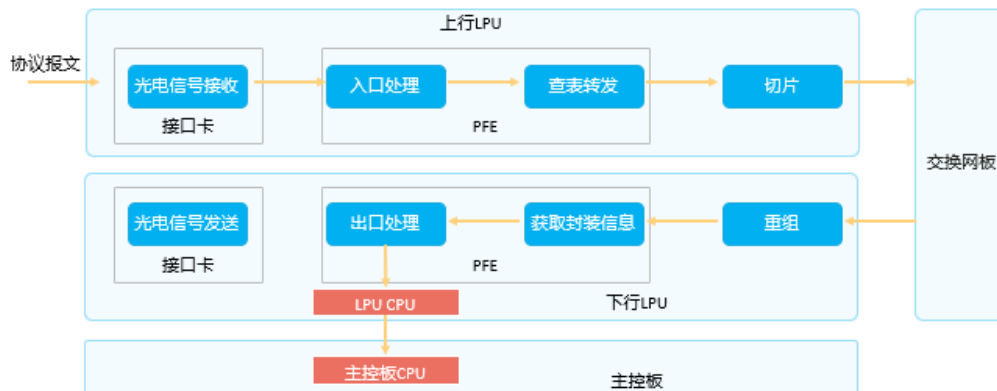


业务报文转发由接口板、交换网板独立完成，无主控板参与。



协议报文转发处理流程

设备收到的协议报文，如路由协议：OSPF、IS-IS、BGP报文，ARP报文，STP报文，对设备的ICMP请求报文等，需要交由设备的控制平面处理，即上送主控板由主控板的CPU进行处理。

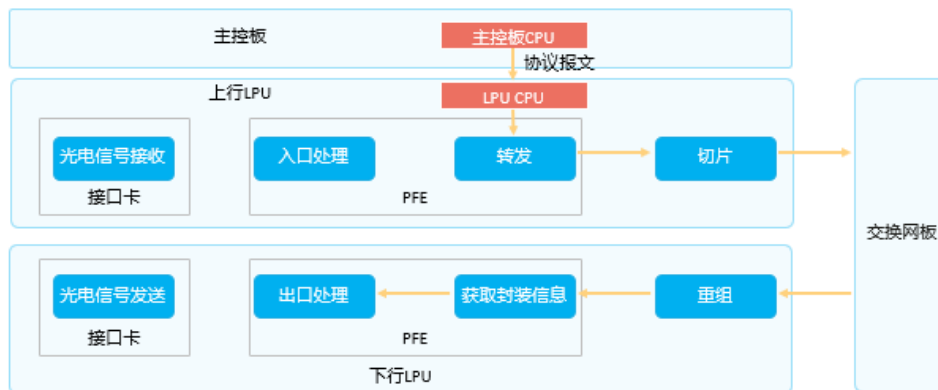


- 主控板 CPU 收到协议报文之后进行相应的处理之后如果需要回应报文，则主控板会构造协议报文进行回应，如收到发往自身的 ARP Request、ICMP Echo Request 报文，主控板处理之后构造 ARP Reply、ICMP Echo Reply 进行回应。
- 主控板 CPU 处理能力有限，如果过多的协议报文上送主控板 CPU 处理，会造成其繁忙，无法及时对协议报文进行响应，为此设备默认限制了各种类型的协议报文上送主控板 CPU 的速率。



设备自身发送协议报文处理流程

设备自身发送的协议报文，如路由协议报文：OSPF、IS-IS、BGP报文，ARP报文，STP报文，ICMP报文等，由主控板CPU构造之后交由接口板对外发送。



本章总结：

- 每台网络设备存在独立的控制平面、转发平面、监控平面，控制平面负责系统的协议处理、路由运算、业务调度等，而转发平面负责转发数据，实现各个业务模块之间的报文交互。监控平面负责系统的环境监控，以保证系统的安全稳定运行。
- 高端框式设备各个平面以不同的板卡形式存在，接口板、交换网板实现转发平面的功能，主控板实现控制平面的功能，监控板实现监控平面的功能。
- 高端框式设备采用“硬转发”，报文的转发由接口板直接完成，无需控制平面参与，报文转发效率极高。

思考题：

- （单选题）框式设备中负责运行路由协议，生成、维护路由表的模块是？
 - 接口板
 - 交换网板
 - 主控板
 - SPU
- （简答题）高端框式设备转发业务报文时是否会向主控板查询转发表项？

答案：

- C
- 不需要，高端框式设备将转发信息下发到接口板，接口板直接转发报文，无需向主控板查询转发表项。