写在前面

TCP/IP模型是计算机网络的事实标准,它在历史上出现得也比国际标准OSI模型要早.因为在日本大学院的专攻就是网络工学,再加上过去几年担当塾讲师,讲过多轮计算机网络的课,对各层的功能,主要协议和设备等都有一些基本的了解.

阅读笔记

TCPIP详解一共有3卷，我手上有英文原版，也有中文翻译版，阅读以译本为参考，然后每卷单列一个子目录。

1. 卷1-协议
2. 卷2-实现
3. 卷3-TCP事务协议□HTTP,NNTP和UNIX域协议

a1-概述

> 协议的定义: 国家事务或外交场合的正式程序或规则系统.

不同厂商,不同型号的计算机之间也需要进行数据交换,它们也需要一系列通用的协议来完成指定任务.所以就需要TCP/IP参考模型.

a2-internet地址结构

表示IP地址

大多数用户所熟悉的IP地址,是IPv4地址,一般采用点分十进制。 与之相对的,是在IPv6中,地址长度从32位扩展到128位,它使用的是8段四位十六进制数,每个字段由冒号: 分隔.并且它也有一些简化的书写方法.

1. 一个字段前导的0可以不写,比如5f05:2000:80ad:5800:0058:0800:2023:1d71可以写为5f05:2000:80ad:5800:58:800:2023:1d71
2. 全零的字段可以省略,并用符号::代替,但为了避免歧义,只能使用一次::,比如0:0:0:0:0:0:0:1可以简写为::1.而2001:0db8:0:0:0:0:0:2可以简写为2001:db8::2.
3. 如果在IPv6中嵌入IPv4地址要, 可以混合点分十进制和十六进制的写法,比如::ffff:10.0.0.1表示IPv4地址10.0.0.1,即被**IPv4映射的IPv6地址**.
4. IPv6的低32位通常采用点分十进制写法,比如::0102:f001相当于地址::1.2.240.1,它被称为**IPv4兼容的IPv6地址**,但这种写法只用于IPv4和IPv6的过渡计划.

基本的IP地址结构 根据IPv4地址长度,可以计算它的地址空间大小为$2^{32}=4,294,967,296$,而IPv6的地址空间为$2^{128}=340,282,366,920,938,463,463,374,607,431,768,211,456$. 所以势必要进行地址分块,才能方便使用.

分类地址

为了在茫茫互联网世界中定位到某一台电脑,需要先找到它所在的网络,(因为互联网本身就是由类型不同,规模不同的网络组成的),所以IPv4地址也反映了这一要素,它分为**网络号**和**主机号**.

早期分类

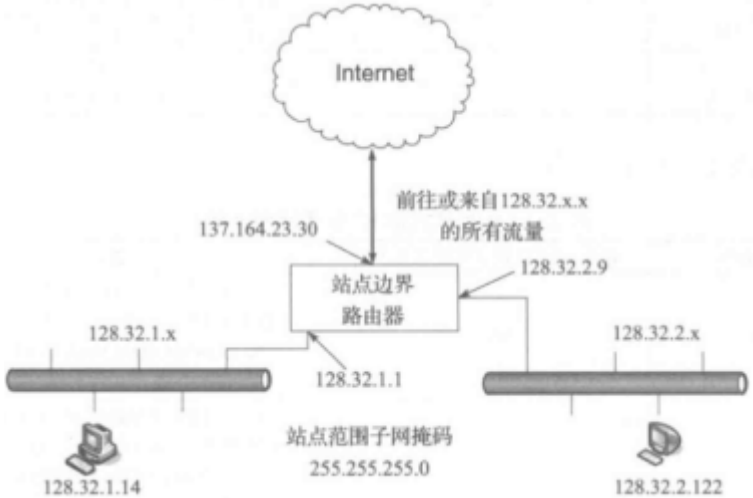
早期将IPv4地址分为ABCDE这5类,列表如下:

类别	网络号	主机号	特征	网络数	主机数
A	8位	24位	网络号以0开头	128	16,777,216
B	16位	16位	网络号以10开头	16384	65536
C	24位	8位	网络号以110开头	2097152	256
D			网络号以1110开头,用于组播	NA	NA
E			网络号以1111开头,保留地址	NA	NA

补充说明: 有些地址通常不作为单播地址使用,比如地址块中的第一个和最后一个地址,前者用于表示该网络,后者表示网络广播地址.

子网寻址

但这种分类方法慢慢出现了新问题,A类和B类网络号规模太大,浪费了许多主机号(太多用不完),而C类规模较小,不能提供足够的主机号. 为了解决这一问题,在保持ABC类网络号集中分配的前提下,各站点的管理员获得权利进一步划分子网络号.但同时也带来了新问题,之前的ABCDE分类方法只要知道网络号,就知道主机号位数了,现在管理员是如何划分子网络的需要反映到路由器和主机中.即只有被划分子网的网络中的主机和路由器知道子网结构.



这里给出了一个示例,左右两个以太网都是子网络,按ABCDE分类法,它们属于B类网络号,即128.32.x.x. 外网的流量先经过边界路由器(这里路由器对外的IP地址为137.164.23.30),因为下面连着2个子网络,所以路由器要区分流向2个子网络的流量. 它要找到子网ID.

子网掩码

子网掩码是一台主机或路由器所使用的分配位,为了确定子网ID.除了用二进制表示外,一般更常用的方式是/xx(范围是1~32) 上面的例子里,网络管理员选择的**子网掩码**是255.255.255.0,即/24.每个子网可以连接256-2=254台主机. 子网掩码中某位的1表示IP地址的对应位与一个地址的**网络/子网络**部分对应, 子网掩码

中某位的0表示IP地址的对应位与一个地址的**主机号**部分对应.

a3-链路层

3.1 引言

链路层的目的是为IP模块发送和接收IP数据报(即PDU),在L2链路层可以携带一些IP辅助协议,比如ARP. TCP/IP支持多种不同的链路层,即该层所处的局域网,它于网络硬件类型有关,常见的有以下几种:

- 有线局域网,如以太网(Ethernet)
- 城域网(MAN),如ISP提供的有线电视和DSL连接
- 有线语音网络,比如支持调解解调器的电话线
- 无线网络,如Wi-Fi
- 基于蜂窝技术的各种无线数据服务,比如HSPA,EV-DO,LTE和WiMAX(即电话运营商提供的)

这几种以第1,4和5种最为常见. 在L2层,我们使用帧(Frame)来与其他层的PDU进行区分.帧格式支持可变帧长度,范围是64Bytes到1518Bytes,上限称之为MTU(最大传输单元). 关于帧最小为64Bytes的解释,有兴趣可以点开看一下:

折叠部分

帧大小 最早的以太网是10Mb/s,为了能让发数据的站知道哪个帧发生了冲突,将一个以太网的最大长度限制在2500m (此时使用4个中继器Repeater连接5个500m的电缆段). 这里就需要加点物理知识了,电子在铜线中的传播速度约为0.77c(c表示光速,学过高中物理的同学应该都知道),则64Bytes采用10Mb/s线缆传输时,用时为 $64\times 8bit/ 10 \times 10^6 =51.2 \mu s$ 最小长度的帧能在电缆中传输约11000m $51.2 \times 10^{-6} \times 0.77\times 300000000 m/s=11000 m$ 如果电缆最大长度限定为2500m,则一次往返距离为5000m 一个输出帧的最位位(比特)在所需时间后仍处于传输过程中,这个时间是信号到达位于最大距离的接收器并返回的时间,如果这时检测到冲突,传输中的站能知道哪个帧发生冲突(即当前正在传输的帧),这里该站会发送一个干扰信号来提醒其他站,然后启动一个随机的二进制指数退避过程(Back-to-N)

3.2 IEEE802 LAN/MAN标准

以太网标准是1980年首次发布,并在1982年加以修订,第一个常见格式的以太网是10Mb/s以太网,被IEEE采纳被经过轻微修改成为802.3标准. 它的结构是,一个或多个站(即主机)组成的共享一个电缆段的区域.因为线路共享,为了减少冲突,采用了CSMA/CD机制,可以协调哪些计算机可以访问共享介质,不需要其他特殊协议或同步. CSMA/CD的处理流程如下:

待补充

a4-地址解析协议

a5-internet协议

a6-系统配置_dhcp和自动配置

a7-防火墙和网络地址转换

a8-icmpv4和icmpv6_internet控制报文协议

a9-广播和本地组播_igmp和mld

a10-用户数据报协议和ip分片

a11-名称解析和域名系统

a12-tcp_传输控制协议_初步

a13-tcp连接管理

a14-tcp超时与重传

a15-tcp数据流与窗口管理

a16-tcp拥塞控制

a17-tcp保活机制

a18-安全_可扩展身份认证协议_ip安全协议_传输层安全_dns安全_域名密钥识别邮件

接下来是第2卷，内容是TCP/IP的具体实现,它有32章之多.

b1-概述

b2-mbuff_存储器缓存

b3-接口层

b4-接口_以太网

b5-接口_SLIP和环回

b6-IP编址

b7-域和协议

b8-IP网际协议

b9-IP选项处理

b10-IP的分片与重装

b11-ICMP_Internet控制报文协议

b12-IP多播

b13-IGMP_Internet组管理协议

b14-IP多播选路

b15-插口层

b16-插口I/O

b17-插口选项

b18-Radix树路由表

b19-选路请求和选路消息

b20-选路插口

b21-ARP_地址解析协议

b22-协议控制块

b23-UDP_用户数据报协议

b24-TCP_传输控制协议

b25-TCP的定时器

b26-TCP输出

b27-TCP的函数

b28-TCP的输入

b29-TCP的输入(续)

b30-TCP的用户需求

b31-BPF_BD分组过滤程序

b32-原始IP