### 网络理论知识手册

**前言：**

1.蓝色字体表示命令行命令，正式执行时不要复制前面的#号，#号只是提示应该使用root权限操作

2.绿色字体表示注释，有时注释太多就不用绿色表示了

3.注意：本文档的所有操作请先在在测环境进行实践，请不要直接在真实的服务器中操作！

**版权声明**：

本文档以开源的形式发布，所有条款如下：

1. 无担保：作者不保证文档内容的准确无误，亦不承担由于使用此文档所导致的任何后果

2. 自由使用：任何人可以出于任何目的而自由地 阅读/链接/打印/转载/引用/再创作 此文档，无需任何附加条件

若您 阅读/链接/打印/转载/引用/再创作 本文档，则说明接受以上2个条款。

作者：李茂福

时间：2022-05-04

**章0、互联网Internet**

**★互联网基础结构发展的3个阶段**

1.从单个网络ARPANET向互联网发展的过程，1969～

2.建成三级结构的互联网、1985年NSFNET（主干，地区，校园网）

3.逐渐形成多层次ISP结构的互联网，1993～

|  |  |
| --- | --- |
| 1969年9月 | 美国国防部创建了第一个分组交换网 ARPANET |
| 1970年12月 | 卡恩、瑟夫参与制定出最初的通信协议NCP网络控制协议 |
| 1974年 | IBM公司宣布了SNA系统网络体系结构 |
| 1974年12月 | 卡恩、瑟夫 发表**TCP/IP协议** |
| 1977年 | 国际标准化组织ISO 成立专门机构提出**OSI/RM模型** |
| 1983年 | TCP/IP协议成为ARPANET上的标准协议（互联网诞生） |
| 1986年 | 美国NSF建立了NSFNET国家科学基金网，因特网逐步形成三级层次架构 |
| 1989年11月 | 中国第一个公用分组交换网CNPAC建成运行 |
| 1990年 | ARPANET宣布关闭，实验任务已完成 |
| 1991年 | 美国政府决定将互联网主干网转交给私人公司经营 |
| 1992年 | 成立了互联网协议ISOC（其下有一个技术组织IAB（互联网体系结构委员会），IAB下有2个部门：**IETF互联网工程部**和**IRTF互联网研究部**）所有互联网标准都是以**RFC**的形式在互联网上发布 |
| 1994年4月20日 | 中国以64kbit/s专线正式接入互联网 |

**★计算机网络**

**计算机网络** 是 （分布在不同地理位置上的具有独立功能的计算机及其外部设备 通过通信线路和通信设备连接起来，按照某种事先约定的规则进行信息交换、以实现资源共享）的系统

**★网络协议3要素**

网络协议：是指在计算机网络中为进行网络中的数据交换而建立的规则、标准或约定的集合，如交换数据的格式、编码方式、同步方式等。（在计算机网络中进行数据交换的规则、标准）

语法：数据与控制信息的结构或格式

语义：需要发出何种控制信息，完成何种动作，做出何种响应

同步：事件实现顺序的详细说明

**★IEEE**

IEEE（Institude of Electrical and Electronics Engineers）国际电气与电子工程师协会，是一个非营利性科技学会，在计算机、电信、生物医学、电力及消费性电子产品等学术领域中都是主要的权威。

IEEE的2个前身：

AIEE美国电气工程师协会，1884年成立

IRE无线电工程师协会，1912年成立。2这个协会于1963年1月1日合并为IEEE

**★OSI/RM七层模型**

Open System Interconnection / Reference Mode开放系统互联参考模型（1977年）

|  |  |
| --- | --- |
| 应用层 | 系统与用户的接口，应用程序（HTTP, FTP, DNS, POP3, TELNET） |
| 表示层 | 完成消息的加密解密，压缩解压缩，字符编码，不同系统间格式的转换 |
| 会话层 | 完成会话的进程控制，ssl |
| 传输层 | 完成端到端（进程之间）的消息的正确传输；  分组的排序，流量控制，差错控制 |
| 网络层 | 完成站到站（主机之间）的分组的正确传输  路由选择，拥塞控制，差错控制 |
| 数据链路层 | 完成数据链路间（一条传输介质的2端）帧的正确传输  封装成帧，透明传输，差错检测，链路管理，流量控制 |
| 物理层 | 完成相邻2个结点间比特流的传输，注重机械电气特征 |

每一层的实体通常定义2种不同的接口：

服务：对上层提供的服务接口，垂直的

协议：对等实体之间通信的规则，水平的

**★TCP/IP四层模型**

应用层，传输层，网络层，网络接口层（1974年）

**章n、网络工程**

**★网络工程**

**网络工程**：是研究网络系统的规划、设计与管理的工程科学。

①工程设计人员：要全面了解计算机网络的原理、技术、系统、协议、安全和系统布线的基础知识，了解计算机网络的发展现状和发展趋势

②总体设计人员：要熟练掌握网络规划与设计的步骤、要点、流程、安全、技术、设备选型以及发展方向

③工程主管人员：要懂得网络工程的组织实施过程、能把握网络工程的评审、监理和验收等环节

④工程开发人员：要掌握网络应用开发技术，网站设计和Web制作技术、信息发布技术以及安全防御技术

⑤工程竣工之后：网络管理人员使用网管工具对网络实施有效的管理维护，使网络工程发挥应有的效益。

**网络工程建设的各阶段：**

1.规划阶段

2.设计阶段

3.实施阶段

4.运行与维护阶段

**网络工程组织方式：**

1.政府机关统一实施的工程，一般指定主管领导和具体负责人

2.公司承建的具体工程，一般采用项目经理制

政府行为的网络工程一般包括3层机构：

1.领导小组

2.总体组（总承组）

3.技术开发小组

**网络工程监理：**

是指在网络建设过程中，给用户提供建设前期咨询、网络方案论证、系统集成商的确定和网络质量控制等 一系列服务；帮助用户建设一个性价比最优的网络系统。

**网络设计的目标：**

1.最大效益下最低运作成本

2.不断增强的整体性能

3.易于操作和使用

4.增强安全性

5.适应性

**网络设计应遵循的原则：**

1.最小的运行成本

2.最少的安装花费

3.最高的性能

4.最大的适应性

5.最大的安全性

6.最大的可靠性

7.最短的故障时间

**网络拓扑结构：**

网络拓扑结构是指忽略了网络通信线路的距离远近和粗细程度，忽略通信节点大小和类型后，仅用点和直线来描述的图形结构

总线型

树型

环型

星型

网状模型

**★综合布线注意事项：**

1.准备工具：压线钳、卡线器、剪刀、螺丝刀、水晶头、扎带、标签扎带、油性记号笔、寻线测线仪、设备挂耳及螺丝、M6机柜螺丝、（光纤尾纤、红光笔、光功计）等

2.办公室搬迁时，旧的网络尽量留长一点，方便下一家使用

3.新办公室要先布线，测好线序，打上编号（对应的位置及桌子号码）

4.尽量不使用配线架或只使用双面都是插水晶头的直通配线架（免打配线架）

5.所有设备先配置，后接线，配置要备份

**★域名转出注意事项：**

1.域名要实名认证且注册满60天，才可转出

2.续费后60天内不能转出

3.过户2年内不能转出

4.到期前60天内不能转出

5.在预计的域名转出日期到来之前2个月之前要算时间是否够90天，不够就续费，因为审核流程可能要花一个月

6.获取转出码后，要先确定转入后的联系人及域名持有人信息

**★互联网接入方式**

**（1）采用普通电话线拨号接入**

①使用调制解调器拨号接入，拨打ISP提供的接入号码（如16900）与ISP建立连接，提供用户名密码

速率： 上行 33.6Kbit/s 下行 56Kbit/s

②ISDN综合业务数字网（一线通）

用户在上网时可打电话，收发传真，所有信号都以数字形式传输，上下行均128Kbit/s

③ADSL非对称数字用户线路

在普通电话线上加上话音分离器和ADSL调制解调器，有效传输距离3~5km。上行 512K ~ 1Mb/s 下行 1Mb/s ~ 8Mb/s

**（2）专线接入**

Cable Modem接入，10M ~ 36M

DDN专线接入，200K ~ 2M

E1专线接入64K ~ 2M

光纤接入 FTTx

**（3）局域网接入（LAN接入）**

将已通过DDN，ADSL等方式接入了Internet的计算机/路由器设为网关或代理服务器，其他用户就可以通过网关来访问互联网

**（4）无线接入**

移动通信网： 2G, 3G, 4G, 5G

无线WLAN： wifi

**★用户接入认证**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 认证方式 | 标准 | IP地址分配 | 客户端软件 | 计费方式 |
| Web Portal | 厂家私有 | 认证前，可2次分配 | 不需要 | 一般按流量计费 |
| 802.1x | IEEE 802.1x | 认证前后都可以 | 需要 | 只支持按时长计费 |
| PPPoE | RFC2516 | 认证后 | 需要 | 时长，流量 |
| IPoE |  |  |  |  |

Web Portal 用户端先通过DHCP获取IP地址，但用户使用此ip并不能访问互联网，在认证通过前，只能访问特定的ip地址，这个地址通常是Web Portal服务器，用户打开浏览器后，Portal服务器通过URL劫持的方式强制跳转到指定的认证web页面，用户输入用户名密码后，浏览器将认证信息传给Portal Server， Portal Server再将信息传给BAS，BAS传给Radius Server或其他认证服务器去认证，认证通过后，就给客户端重新分配ip，就可上网

**★机柜，服务器高度**

1U为一个高度单位 ==> 1.75英寸 ==> 4.445厘米

19英寸规范机架布局宽度 ==> 48.3厘米

2U为8.89 cm

4U为17.78 cm

42U为186.69 cm

47U为208.915 cm

48U为213.36 cm

**章n、电磁波、通信信道**

光速c = f \* λ （频率\*波长）

f频率 = 3000/波长nm = ~(100THz) 1nm为10-9m

|  |  |  |  |
| --- | --- | --- | --- |
| 可见光 |  | 波长nm | 频率 |
| 红 | 760~622 | 760nm对应394THz |
| ... |  | （100T=10^14） |
| 紫 | 455~400 | 400nm对应750THz |

**★电磁波应用领域：**



|  |  |
| --- | --- |
| 双绞线 | 4KHz（75Km） 16MHz（18.75m） 100MHz（3m） 125MHz（2.4m）  250MHz（1.2m） |
| 调频无线电 | 76MHz（3.94m） 108MHz（2.78m） |
| 移动电话 | 850MHz（35.3cm） 900MHz（33.3cm） 1800MHz（16.66cm）  2000MHz（15cm） 2100MHz（14.2cm） 2300MHz（13.04cm） |
| 无线wifi | 2.4GHz（12.5cm） 5.0GHz（6cm） 5.8GHz（5.17cm） |
| 光纤 | 850nm（352.9THz）1310nm（229THz）1490nm（201.3THz）1550nm（193.5THz） |

**电磁波传输速度**

铜线： 2.3x10^5 km/s 230公里/ms (毫秒)

光纤： 2.0x10^5 km/s 200公里/ms

真空： 3.0x10^5 km/s 300公里/ms

**★通信信道**

信道（Channel）一般用来表示某一个方向传送信息的媒体，所以一条通信电路往往包含2条信道，发送信道和接收信道。

**3种通信方式：**

单向通信，（单工通信）

双向交替通信（半双工通信）

双向同时通信（全双工通信）

**★复用技术**

|  |  |
| --- | --- |
| FDM | 频分复用 Frequency Division Multiplexing |
| TDM | 时分复用 Time Division Multiplexing |
| STDM | 统计时分复用 Statistics TDM |
| WDM | 波分复用(光缆) Wavelength Division Multiplexing |
| DWDM | 密集波分复用 Dense WDM |
| CDM | 码分复用 Code Division Multiplexing |
| CDMA | 码分多址 |
| TDD | 时分双工 |
| FDD | 频分双工 |
| SDM | 空分复用 |

**★基本带通调制方式**

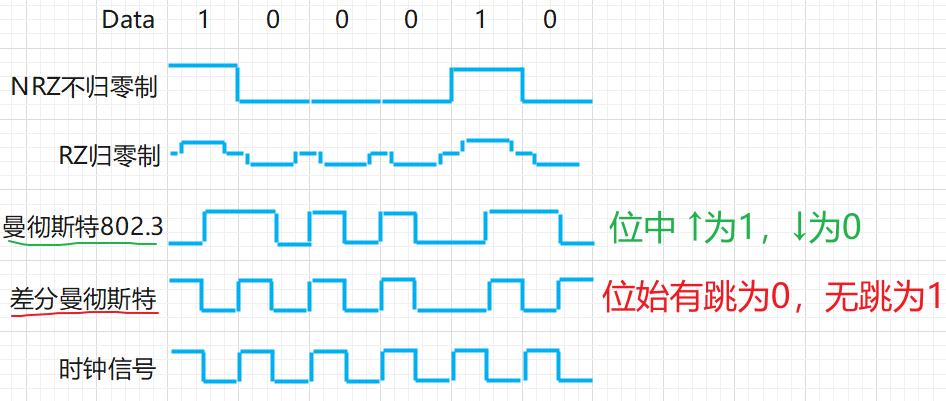
基带信号：来自信源的信号，基带信号往往带有较多的低频成分，许多信道不能传输低频，所以要对其进行调制

调制分2种：

**基带调制（编码）**：仅对基带信号波形进行变换，变换后的信号仍为基带信号

**带通调制**：使用载波进行调制，把基带信号的频率范围搬到较高的频段并转换为模拟信号

**常用的基带调制（编码）**



802.3的曼彻斯特编码 == 数据 XOR 时钟 （导或运算）

**★码元，波特率**

信道带宽W = f2 - f1 （f2为信道能通过的最高频率，f1为最低频率）

一个数据脉冲称为一个码元，

**码元速率**：单位时间内信号波形的变换次数或者 通过信道传输的码元个数

**码元速率的单位**为 波特（Baud）

所以 **码元速率**也叫作 **波特率 BaudRate**

最大码元速率等于2倍信道带宽，Baudmax = 2\*W

码元携带的信息量 由码元取的离散值个数决定，如：

码元取2个离散值，则一个码元携带 1bit信息

码元取4个离散值，则一个码元携带 2bit信息

码元取8个离散值，则一个码元携带 3bit信息

所以 数据速率 等于 波特率 \* nbit（码元携带信息bit）

但实际会受噪声干扰而达不到理想值

**香农定理**：无论用什么方式调制，只要给定了信噪比，则单位时间内最大的信息传输量就确定了

如：信道带宽为3000 Hz，信噪比30dB，则最大数据率约为 30kb/s

**★带宽 概念**

**带宽**是“**频带宽度**”的简称，**原本是**通讯和电子技术中的一个术语，指通讯线路或设备 能传送信号的范围（如4kHz ~ 1MHz），而网络中的带宽是指在规定时间内从一端传输到另一端的信号量（即数据传输率）如100Mbit/s

**吞吐量**：是指单位时间内实际传输数据的量（如一秒传了57Mbit） <= (带宽\*时间)

**章n、以太网、双绞线**

**★双绞线种类**

Bell于1876年发明了电话，也发明了双绞线

1991年7月美国电子工业协会EIA 和电信行业协会TIA联合发布了EIA/TIA-568标准（商用建筑物电信布线标准）

1995年更新至EIA/TIA-568A（定义了5个种类的UTP标准），一个扭绞周期长度叫作节距，节距越小抗干扰能力越强

UTP（Unshielded Twisted Pair）非屏蔽双绞线

STP（Shielded Twisted Pair）全屏蔽双绞线，每对线芯都有各自的屏蔽层

FTP（Foil Twisted Pair）单屏蔽，在整个电缆外有屏蔽装置

4对线要么都用顺时针绞绕，要么都用逆时针

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 种类 | 频率带宽 | 线芯绞合长度 | 数据速率 | 应用 |
| 1 | 750KHz |  |  | 电话语音（1980年代） |
| 2 | 1MHz |  | 4Mbit/s | 旧的令牌网（1980年代） |
| 3 | 16MHz | 2对4芯7.5~10cm | 10Mbit/s | 模拟电话，传统以太网（1990） |
| 4 | 20MHz | 4对8芯7.5~10cm | 16Mbit/s | 令牌局域网 |
| 5 | 100MHz | 4对8芯12.7mm以内 | 100M | 以太网（1995年） |
| 5e | 125MHz | 4对8芯 | 100M~1G | 以太网（2001年） |
| 6 | 250MHz | 4对8芯中有十字分隔条 | 1G | 以太网（2002年） |
| 6a | 500MHz | 4对8芯中有十字分隔条 | 10G | 以太网（2008年） |
| 7 | 600MHz | 屏蔽双绞线 | 10G | 以太网 |

1到7类线都是最长100米

**★AWG线径**

AWG（American wire gauge）美国线规，是一种区分导线直径的标准

22 AWG ==> 0.643 mm

23 AWG ==> 0.574 mm 或 0.573 mm

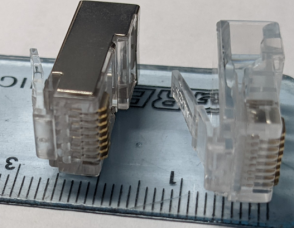
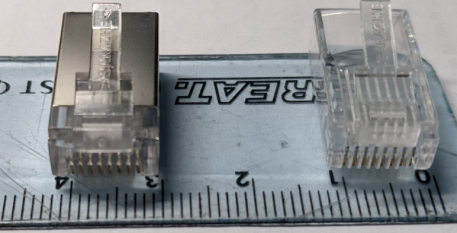
24 AWG ==> 0.511 mm

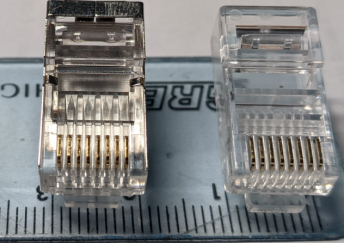
25 AWG ==> 0.44 mm

26 AWG ==> 0.404 mm

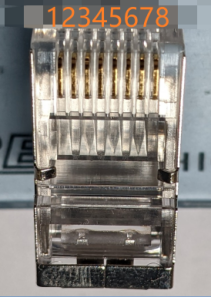
**★水晶头形状及线序**

RJ45水晶头 宽约11mm，高约6mm，插入部分长度约10mm，



 RJ45网口

水晶头金手指朝上，面向我们，从左到右线序为1到8



EIA/TIA-568B标准：白**橙**，**橙**，白**绿**，**蓝**，白**蓝**，**绿**，白**棕**，**棕**

EIA/TIA-568A标准：白**绿**，**绿**，白**橙**，**蓝**，白**蓝**，**橙**，白**棕**，**棕** （1，3交换，2，6交换）

**<1,2> Tx发送， <3,6> Rx接收 4，5，7，8双向数据**

**★RS-232串口，DB9**

1969年 RS-232-C

COM：Cluster Communication Port 串行通讯端口

DB9母头， DB9公头（电脑台式机箱后面一般为公头）

|  |  |  |
| --- | --- | --- |
| 1 | DCD | 载波检测，Data Carrier Detect |
| 2 | RxD | 接收数据，Received Data |
| 3 | TxD | 发送数据，Transmit Data |
| 4 | DTR | 数据终端准备好，Data Terminal Ready |
| 5 | GND | 信号地线 |
| 6 | DSR | 数据准备好，Data Set Ready |
| 7 | RTS | 请求发送，Request to Send |
| 8 | CTS | 清除发送，允许发送，Clear to Send |
| 9 | RI | 振铃指示，Ring Indicator |

一般只用2，3，5号线，不用流控

流控：

DTR/DSR

RTS/CTS

Xon/Xoff

TxD和RxD电压：

逻辑1 Mark -3 ~ -15V

逻辑0 Space +3 ~ +15V

**★交换机console口通用线序**

在

|  |  |
| --- | --- |
| Rj45水晶头线序说明 | 对应交换机console口的 |
| 1，RTS | 1，CTS |
| 2，DTR | 2，DSR |
| 3，TxD | 3，RxD |
| 4，GND | 4，GND |
| 5，GND | 5，GND |
| 6，RxD | 6，TxD |
| 7，DSR | 7，DTR |
| 8，CTS | 8，RTS |

**★以太网**

以太网（Ethernet）于1975年由美国Xerox施乐公司的PaloAlto研究中心研制，当时速率2.94Mbit/s，基带总线型局域网

1980年9月DEC公司、Intel公司和施乐公司 联合提出10Mbit/s以太网第一版（DIX v1）

1982年修改为DIX v2 成为世界上第一个局域网产品规约

1983年IEEE 802委员会的802.3工作组制定了IEEE 802.3（10M）以太网标准

1990年代后DIX v2已取得垄断地位，802.3的LLC逻辑链路控制子层LLC的作用已消失，很多厂商的适配器上仅装有MAC协议，而没有LLC协议

严格来说以太网是指符合DIX Ethernet v2标准的局域网，以太网发送的数据使用曼彻斯特编码，CSMA/CD载波监听 多点接入/碰撞检测

1988年IEEE 802.3ac标准定义了以太网的扩展帧，支持VLAN

1990年IEEE制定出星形以太网，802.3i标准，10BASE-T

MAC地址就是适配器地址或适配器标识符EUI-48

10Mbit/s以太网：

争用期 51.2us ==> 512bit ==> 64字节，最短帧长64字节（10M网里）

帧间最小间隔 9.6us ==> 96bit ==> 12字节间距

1995年IEEE把100BASE-T快速以太网定为正式标准，802.3u

争用期 5.12us ==> 512bit ==> 64字节，同10M

帧间最小间隔 0.96us ==> 96bit ==> 12字节间距

1997年IEEE通过了吉比特以太网标准（802.3z和802.3ab）1998年成为正式标准

1000Mbit/s以太网标准：

争用期增大为 512字节 ==> 4.096us

最短帧长仍为 64字节，载波延伸，分组突发

2002年6月完成10Gbit/s以太网标准802.3ae的前三项标准，只工作在全双工模式，不使用CSMA/CD，不存在争用期问题

**以太网标准**

|  |  |  |  |
| --- | --- | --- | --- |
| 时间 | IEEE标准 | 协议 | 说明 |
| 1990 | 802.3i | 10 BASE-T | 双绞线，100米 |
| 1995 | 802.3u | 100 BASE-T | 双绞线，100米 |
| 1997 | 802.3z | 1000 BASE-SX  1000 BASE-LX  1000 BASE-CX | 光缆，550米，多模，850nm  光缆，5000米，单模，1310/1550 nm  双绞线，25米，STP屏蔽 |
| 1997 | 802.3ab | 1000 BASE-T | 双绞线，100米，UTP，cat 5e以上双绞线 |
| 2002 | 802.3ae | 10G BASE-SR  10G BASE-LR  10G BASE-ER | 光缆，300米，多模，850nm  光缆，10km，单模，1310 nm  光缆，40km，单模，1550 nm |

**★数据包转发速率**

|  |  |
| --- | --- |
| 100M端口 最大 包转发速率 | 156.038 kpps （1024进制），最短包  148.8 kpps （1000进制），最短包 |
| 100M端口 最小 包转发速率 | 8522 pps （1024进制），最长包  8127 pps （1000进制），最长包 |

最小包/短包 为 64字节 + 8字节前导 + 12字节间隔 = 84字节

100Mbit/s ==> 12.5MB/s

12.5MB/s除以84B ==> 156038 包/秒 （1024进制），148809包/秒（1000进制）

最长包 为 1518字节 + 8字节前导 + 12字节间隔 = 1538字节

12.5MB/s除以1538B ==> 8522 pps（1024进） 8127 pps（1000进制）

**★无线网络 802.11**

1997年IEEE制定无线局域网标准802.11系列

在MAC层使用CSMA/CA载波监听，多点接入/碰撞避免

|  |  |  |  |
| --- | --- | --- | --- |
| 无线标准 | 频率 | 速率 | 时间 |
| 802.11a | 5GHz | 54M bit/s | 1999年9月 |
| 802.11b | 2.4GHz | 11M bit/s | 1999年9月 |
| 802.11g | 2.4GHz | 54M bit/s | 2003年6月 |
| 802.11n | 2.4G/5G Hz | 1 x 1 40MHz ==> 150Mb/s  4 x 4 40MHz ==> 600Mb/s | 2009年9月 |
| 802.11ac | 5GHz | 1G bit/s | 2014年1月 |
| 802.11ah | 900MHz | 18M bit/s | 2016年传输距离1Km |
| 802.11ax | 2.4G/5G Hz | 单流1201Mb/s | 2018年10月4日 |

802.1b 2.4G标准定义了14个信道，（2.4~2.485 GHz中有85MHz带宽可用）

仅当2个信道由4个及以上信道隔开才无重叠，

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 信道 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 频率GHz | 2.412 | 2.417 | 2.422 | 2.427 | 2.432 | 2.437 | 2.442 |
| 信道 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
| 频率GHz | 2.447 | 2.452 | 2.457 | 2.462 | 2.467 | 2.472 | 2.484 |

蓝牙：2400~2483.5 MHz

5G的频段分为7~196个信道（40到165信道：5.2GHz ~ 5.825GHz）

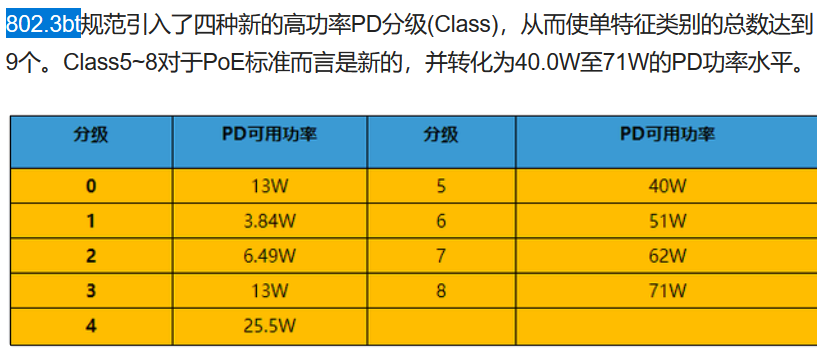
36信道：5170MHz

中国可用的：36~64, 149~165信道

**★POE供电**

Power Over Ethernet，利用以太网双绞线传输电力，无需额外的电源线就可给设备供电，一般用在AP无线接入设备上。一般AP为48V

|  |  |  |  |
| --- | --- | --- | --- |
| 标准 | 供电能力(标准，实际) | 电压 |  |
| 802.3af | 15.4W ，12.95W | 44~57 V | af为PD受电设备提供3.84~12.95W四个等级的功率请求 |
| 802.3at（PoE+） | 30W ，25.5W | 50~57V |  |
| 802.3bt（PoE++） |  |  | 7类网络及以上 |



标准的五类网线有四对双绞线，IEEE 802.3af允许两种线序供电方法：

①在4、5、7、8线对上传输电流，并且规定，4、5为正极，7、8为负极。

②在1、2、3、6线上传输电源，极性较为任意，1、2为正极，3、6为负极或是1、2为负极，3、6为正极。

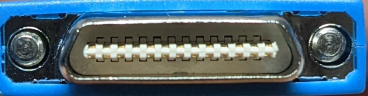
POE交换机采用第二种供电线序，无线AP或者物联网LoRa网关两种供电线序都支持

**★V.35接口**

v.35通用终端接口最初版本是ITU-T于1968年发布，在48~64 Kbps之间的宽带模拟调制解调器和DTE之间的接口，修定后版本可支持6Mbps

v.35对连接器的形状并未规定，但48~64 kbps的美国Bell调制解调器的普及，34引脚的ISO 2593被广泛采用









**章n、电话交换机及载波**

**★电话交换机发展史**

1876年3月10日 贝尔发明了电话

1877年在波士顿的第一条电话线路开通了

1891年3月10日发明步进制电话接线器，共电式电话交换机

1892年11月3日步进制自动电话交换机在美国印第安纳州的拉波特城投入使用，这是世界上第一个自动电话局

1919年发明纵横制接线器

1929年瑞典松兹瓦尔布 建成世界上第一个大型纵横制电话局

1965年美国研制和开通了第一部空分程控交换机（贝尔系统的1号电子交换机）

1970年在法国拉尼翁成功开通了世界上第一个程控数字交换机系统E10

1990年代中国研制并生产了HJD-04, ZXJ10, C&C08 交换机

PBX俗称程控交换机，程控用户交换机，电话交换机，集团电话

1995年IP-Phone技术，采用了压缩编码及基于TCP/IP的包交换技术，

G.729速率8kb/s G.723.1输出6.3kb/s或5.3kb/s

ip-phone封装在tcp中后，实际传送码流约为 24kbit/s 时延20ms（打包2个话音包）

比PCM的64kbit/s要小，利用率提高了

**★电话交换机按控制方式分类**

①布线逻辑控制WLC（Wired Logic Control）使用机电接线器，只是将控制部分更新为电子器件，也称为 布控半电子式交换机

②存储程序控制SPC（Stored Program Control）全电子型，采用程序控制方式，分2种

|  |  |
| --- | --- |
| 程控空分交换机 | 在话路中一般传送模拟话音信号 |
| 程控时分交换机 | 一般传送数字信号，采用PCM技术，不仅可接入电话业务，还可接入ISDN数字业务 |

空分：就是用户在打电话时要占用一对线，一直到打完电话才释放，以前机电式电话交换机都是空分方式（1965~1975年间绝大多数电话交换机为空分的，模拟的）

**★PCM脉冲编码调制**

PCM（Pulse Code Modulation）脉冲编码调制是话音信号的数字化方法，包括 抽样、量化、编码 三种功能单元

模拟话音经过防混叠低通滤波得到 300~3400 Hz（fm约为4kHz）的话路信号，再将其抽样变成脉冲调幅PAM信号，抽样频率fs >= 2fm（8kHz）

CCITT G.711, G.712建议每抽样值编为8位码（256个量化级），所以每路模拟话音相应的数字话音标准数码率为8kHz x 8bit = 64kbit/s，有的用了128个量化级，数码率为56kb/s

**★奈奎斯特采样定理**

1928年美国电信工程师H.Nyquist 提出采样定理，该定理说明了采样率与信号频率之间的关系，在进行模拟/数字信号的转换过程中，当采样频率fs大于信号中的最高频率fmax的2倍时，采样之后的数字信号完整地保留了原始信号中的信息，实际应用2.56~4倍的fs

|  |  |
| --- | --- |
| 人类发音频率 | 85 ~ 1100 Hz |
| 一般乐器频率 | 20 ~ 4000 Hz |
| 人类听觉 | 20 ~ 20000 Hz |
| 高保真耳机 | 50 ~ 12500 Hz |

语音编码：

G.711 G.723.1 G.729

**★T1载波 1.544Mb/s**

1960年代

T1载波是专用电话连接，时分多路数字传输设施，采用PCM和TDM技术

T1线路逻辑上是由24个单独的通道组成，每个通道支持64kbit/s的传输速度，数据率为56kb/s

当T1系统用于模拟传输时，复用24路话音信道，每条话音信道输出7bit数据和1bit控制信号（随路信令）

当T1系统完全用于数字传输时，仅23条信道用于数据传输，第24条信道用于同步模式，即D信道（共路信令）

T1载波的每一个帧均分为24段，每段为一个所谓的通道/路，每段也叫一个时隙

T1线路每秒传输8000帧，每帧用时125us，每帧长193bit

帧结构：

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 长度bit | 8 | 8 | 8 |  | 8 | 1 |
|  | 1路 | 2路 | 3路 | ....路 | 24路 | 帧同步码 |

一个帧共24路（每路7bit数据 + 1bit控制信号）24路之后加一bit帧同步码

所以T1载波的传输速率为 1.544Mbit/s（193bit x 8000/s） 开销8kbit/s（1bit x 8000/s）

每一路速率为 56kbit/s（7bit x 8000/s） 或 64kbit/s（8bit x 8000/s）

T1信号由100Ω平衡双绞线传送，使用4线电路，2条发送，2条接收，电压-135V

**★E1载波 2.048Mb/s**

E1载波原理同T1，E1将30个话音信道和2个控制/同步 信道复合在一条高速信道上

即E1的每一帧均分为32个时隙，每时隙8bit

E1线路每秒也是传输8000帧，每帧用时125us，每帧长256位

帧结构：

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 长度bit | 8 | 8 |  | 8 |  | 8 | 8 |
|  | 时隙0 | 时隙1 | ...时隙 | 时隙16 | ...时隙 | 时隙30 | 时隙31 |

E1载波的传输速率为2.048Mbit/s（256bit x 8000/s）

E1有3种成帧方式：

|  |  |
| --- | --- |
| 成帧E1 | 第0时隙传输帧同步数据（同步码），其他31个时隙传输有效数据 |
| 成复帧E1 | 第0时隙传输帧同步数据，第16时隙传信令，其余30个传有效数据 |
| 不成帧E1 | 32个时隙都可以传输有效数据 |

E1的应用：

1，将整个2M用作一条链路，如DDN 2M，帧中继，x.25

2，将2M用作若干个64kb/s及其组合，如128k, 256k, ....称为 CE1

3，原本的用法，3种成帧方式来使用，如PRI接入方式

CE1在使用时可将一条E1的2M划分为n x 64kb/s，CE1和0时隙和15时隙不用来传输用户数据

CE1也可用E1 2M相连，只是CE1必须当成E1 2M来使用

E1使用G.703标准，平衡 120Ω 使用RJ45接口；非平衡75Ω 使用同轴电缆

**★ISDN综合业务数字网 BRI和PRI**

ISDN（Integrated Services Digital Network）是一个数字电话网络国际标准，俗称一线通

在ITU的建议中，ISDN是一种在IDN数字电话网的基础上发展起来的通信网络，ISDN能够支持多种业务，如 电话业务，非电话业务

ISDN 1980年代兴起

中国于1995年上海最早商用，1996年正式将ISDN业务命名为 一线通

ISDN是一个全数字的网络，实现了端到端的数字连接，ISDN用户可在一对双绞线上传输2个B信道和一个D信道：

B表示Bearer承载，通信通路，64kb/s传输话音/数据

D表示Data控制，控制通路，64/16 kb/s传输信令

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ISDN | BRI基本速率接口（窄带ISDN） | 2B + D | 2x64 + 16 | 144 kbps |
| PRI主速率接口（宽带ISDN） | T1（23B + D） | 23\*64 + 1\*64 | 1544 kbps |
| E1（30B + 2D） | 30\*64 + 2\*64 | 2048 kbps |

BRI和PRI的B承载都为64kbps，BRI的D为16kbps，PRI的D为64kbps

PRI，在中国大陆、欧洲使用E1载波；在北美、日本使用T1载波

PRI信令又称为PRA信令，DSS1信令

**★No.1信令和No.7信令**

程控交换机的信令（Signalling）按其作用分为用户信令和局间信令

用户信令：用户与交换机之间的用户线上传递的信令

局间信令：交换机之间/交换局之间 中继线上传递的信令

为了统一局间信令，国际电信联合会ITU-T 陆续提出并形成了1~7号信令系统（局间）

1~5号为随路信令，6~7号为共路信令

|  |  |
| --- | --- |
| 随路信令CAS | 在同一通路上传递话音信息与其相应的信令 |
| 共路信令CCS | 将一组话路所需的各种局间控制信号集中到一条与话音通路分开的公共信号通路上传递 |

**★ADSL非对称数字用户线**

1994年

ADSL（Asymmetric Digital Subscriber Line）是一种异步传输模式，以ITU-T G.992.1标准为例：

ADSL在一对双绞线上支持 上行512K ~ 1Mb/s ，下行1M ~ 8Mb/s，有效传输距离3~5公里

较成熟的ADSL标准有： G.DMT和G.lite

G.DMT全速率ADSL标准，上行1.5M，下行8M，用户端要安装POTS分离器

G.lite标准，上行512K，下行1.5M，不用POTS，成本低

DMT（Discrete Multi-Tone）离散多音调（多载波）

DSL数字用户线技术是贝尔通信研究所于1989年研发的

ADSL由ANSI制定，1994年通过，1999年ITU批准了G.lite标准

**★DDN数字数据网**

DDN（Digital Data Network）向用户提供永久性/半永久性 连接的数字数据传输信道，可使用x.25和帧中继 等协议，传输质量高，时延小

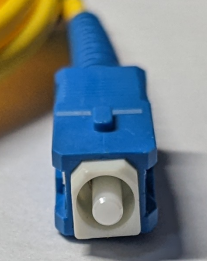
专线上网是一种利用光纤，微波，卫星等数字传输通道和数字交叉利用设备 组成的数字数据传输网

中国公用数字数据骨干网 CHINA DDN 于1994年正式开通

**章n、光纤，光网络**

**★光纤接口**

① SC 塑料外壳（光纤入户常用接口）



② LC 塑料外壳，比SC小（机房交换机端口上常用）



③ FC 金属接头



④ ST 金属接头有弹簧卡扣



单模光纤 一般为黄色外皮，9/125 um（9um纤芯，125um包层）

多模光纤 一般为橘红色外皮 50/125 um（50um纤芯，125um包层）

**★光模块类型及协议**

在

GBIC（GigaBit Interface Converter）千兆光模块，热拔插

SFP（Small Formufactor Pluggable）是GBIC的升级版本

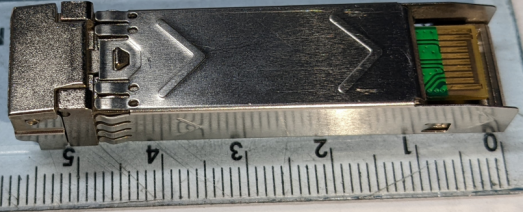
SFP+尺寸和SFP一样大，但它为10Gb/s

BiDi单纤双向（BiDirectional），利用WDM波分复用技术，通过光模块中的滤波器进行滤波同时完成1310nm和1490/1550nm光信号的发射与接收。（只有一根光纤）

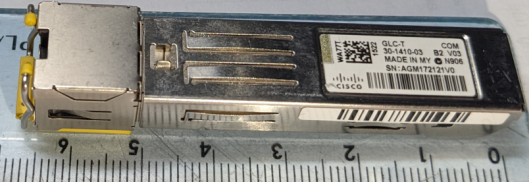
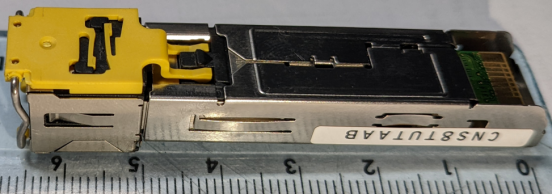
**①SFP 光模块**

宽约13mm，高约8mm，插入部分约47mm 或48mm



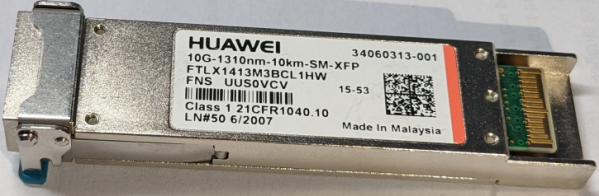
**②SFP千兆光转电模块**，插入部分同SFP，外面接口不是接光纤，而是rj45以太网口



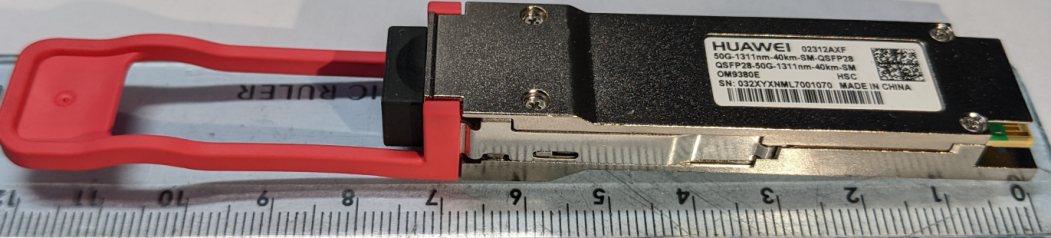
③XFP光模块

宽约18mm，高约8mm，长度约70mm

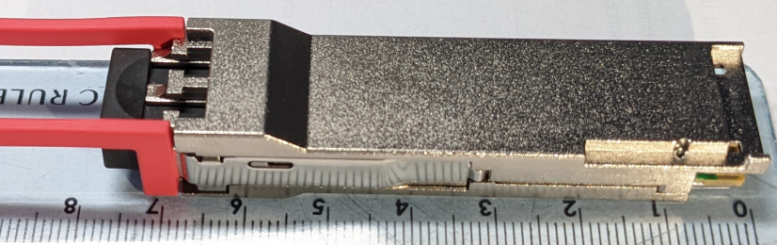
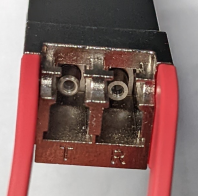


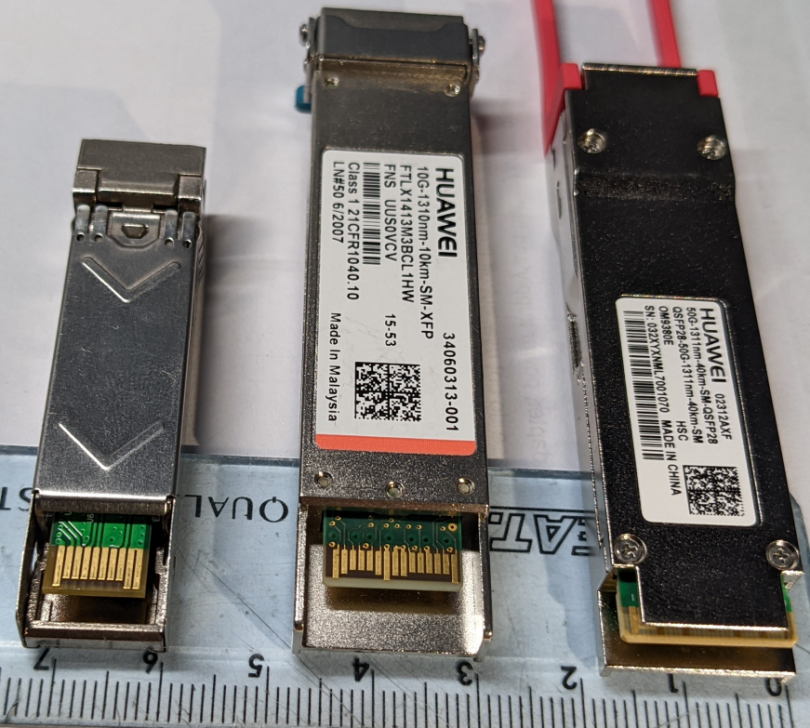


④QSFP光模块



宽约18mm，高约8mm，插入部分约52mm



左边为sfp，中间为xfp，右边为qsfp

**★光功率单位**

1mW为0dBm（分贝瓦），dBm也写作dBW

小于1mW的dBm值为负数

换算公式 dBm == 10 log（功率值/1mW）

1个基准：30dBm == 1W

2个原则：

\*加3dBm功率值乘2，减3dBm功率值除以2

\*加10dBm功率值乘10，减10dBm功率值除以10

如：

30dBm == 1W

20dBm == 0.1W

10dBm == 0.01W （10mW）

0dBm == 0.001W （1mW）

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| dBm | 功率值 | dBm | 功率值 | dBm | 功率值 |
| 0 | 1mW | -21 | 7.94328uW | -45 | 31.62nW |
| -10 | 100uW | -22 | 6.3095uW | -50 | 10nW |
| -11 | 79.4328uW | -23 | 5.0118uW | -55 | 3.16nW |
| -12 | 63.0957uW | -24 | 3.981uW | -60 | 1nW |
| -13 | 50.1187uW | -25 | 3.162uW | -65 | 0.3nW |
| -14 | 39.8107uW | -26 | 2.5118uW | -70 | 0.1nW |
| -15 | 31.6227uW | -27 | 1.99526uW |  |  |
| -16 | 25.1188uW | -28 | 1.5849uW |  |  |
| -17 | 19.9526uW | -29 | 1.2589uW |  |  |
| -18 | 15.8489uW | -30 | 1uW |  |  |
| -19 | 12.589uW | -35 | 316.23nW |  |  |
| -20 | 10uW | -40 | 100nW |  |  |

光模块 过载光功率 High Alarm 最大可接收功率 0dBm

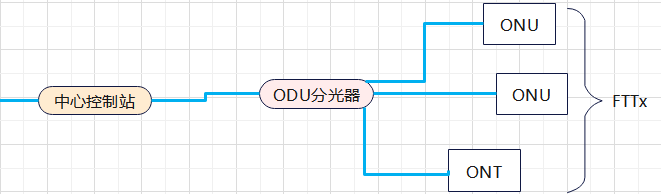
光模块 接收灵敏度 Low Alarm 最小可接收功率 -20dBm

**★PON无源光网络**

PON（Passive Optical Network）无源光网络

ODU（Optical Distrubution Unit）分光器

ODN（Optical Distribution Network）光分配网，位于OLT与ONU之间



ONU（Optical Network Unit）光网络单元，与用户之间还可以有其他网络，如以太网，Ems

ONT（Optical Network Terminal）光网络终端，直接位于客户端

ODF（Optical Distribution Frame）光纤配线架

OLT（Optical Line Terminal）光线路终端，电信局端设备

OLT广播数据包（加密了的）到所有ONU，目的ONU有选择地提取

ONU发出的包只能到达OLT，时分复用

PON有2个标准（GPON和EPON）

**①GPON** 由ITU/FSAN制定，Gigabit PON 无源光吉比特网

接收灵敏度 -27dBm，上行1310nm，1.244Gb/s

过载光功率 -8dBm，下行1490nm，2.488Gb/s

**②EPON** 由IEEE 802.3ah工作组制定，Ethernet PON 无源光以太网

2004年6月IEEE 802.3EFM工作组发布了EPON标准（802.3ah）

2005年并入IEEE 802.3-2005标准，定义了2种PON光接口：

1000 BASE-PX-10 U/D 10km

1000 BASE-PX-20 U/D 20km

接收灵敏度 -27dBm，上行1310nm，1Gb/s

过载光功率 -3dBm，下行1490nm，1Gb/s（线路速度都是1.25Gb/s）

数据链路层采用MPCP（Multi-Point Control Protocol）多点MAC控制协议

**章n、网络基础协议及报文**

**★MAC地址及以太网MAC帧**

MAC（Media Access Control）媒体访问控制地址 长度48bit，6字节，用冒分十六进制表示

也可用点分十六进制，杠分十六进制表示

mac地址前24bit由IEEE分配，表示OUI（organizationally Unique Identifier）网络硬件制造商的编号，后24位由厂商自行分配

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 第一字节 | 第2字节 | 第3字节 |  |  | 第6字节 |
| xxxx xxug | xxxx xxxx | xxxx xxxx | xxxx xxxx | xxxx xxxx | xxxx xxxx |

第一字节的低第2位（从左到右第7位）为U/L位（Universal/Local）置0表示全局地址，置1表示本地地址

第一字节的低第1位（从左到右第8位）为I/G位，置0表示单播，1为多播地址

IP组播地址对应的MAC地址：

ip组播地址28bit的后面23位对应MAC中的低23位（有32个组播ip对应相同的MAC）

前25位固定为01-00-5E（第25位为0）

VRRP的虚拟网关对应MAC： 0000.5E00.01xx xx为vrid号

HSRP的虚拟网关对应MAC： 0000.0C07.ACxx xx为standby组号

EUI-64：

把48bit的MAC地址转为64bit的EUI-64

在mac地址第3，4字节之间插入2字节（FF.FE）再把第1字节第7bit（U/L位）取反，就得到EUI-64（ipv6的接口id）

**★以太网MAC帧**

DIX Ethernet v2标准：

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 长度字节 | 8 | 6 | 6 | 2 | 46~1500 | 4 |
|  | 前同步码，帧开始 | 目的MAC | 源MAC | 类型 | 数据（46~1500字节MTU） | FCS |

7字节前同步码（1010 1010），1字节帧开始定界符（1010 1011）

IEEE 802.3与DIX v2的区别在于MAC帧的第3个字节（类型）的值，其值大于0x0600时，2者相同，当其值小于0x0600时，表示mac帧的数据长度，其数据为LLC子层

类型字段值表示的上层协议：

|  |  |  |  |
| --- | --- | --- | --- |
| 0x0800 | IPv4协议 | 0x0806 | ARP报文 |
| 0x86dd | IPv6协议 | 0x8863 | PPPoE Discovery |
| 0x8100 | VLAN帧 | 0x8864 | PPPoE Sesson |
| 0x8137 | IPX/SPX报文 |  |  |

VLAN帧：

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 字节 | 8 | 6 | 6 | 4 | 2 | 46~1500 | 4 |
|  | 前同步码，帧开始 | 目的MAC | 源MAC | VLAN标记 | 类型 | 数据（46~1500字节MTU） | FCS |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 长度bit | 16 | 3 | 1 | 12 |
|  | 0x8100 | qos | CFI | VID |

16位的0x8100表示vlan帧

3位的用户优先级User Priority，取值0~7，越大越优先，默认为0

1位规范格式指示符CFI（Canonical Format Indicator）

12位VLAN标识符，vlan id，0~4095，有效的1~4094

**★802.11 MAC帧（无线局域网）**

使用CSMA/CA载波监听，多点接入/碰撞避免

802.11无线局域网有3种mac帧：

**①数据帧**（总长 34 ~ 2346字节）

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 字节 | 2 | 2 | 6 | 6 | 6 | 2 | 6 | 0 ~ 2312 | 4 |
|  | **帧控制** | 持续期 | 地址1 | 地址2 | 地址3 | 序号控制 | 地址4 | 帧主体 | FCS |

帧控制结构：（2字节）

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| bit | 2 | 2 | 4 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
|  | 协议版本 | 类型 | 子类型 | 去往AP | 来自AP | 更多分片 | 重试 | 功率管理 | 更多数据 | WEP | 顺序 |

帧控制说明：

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | 去往AP | 来自AP | 地址1 | 地址2 | 地址3 | 地址4 |
| AP发出的报文 | 0 | 1 | 目的MAC | AP的MAC | 源MAC | 无 |
| 发给AP的报文 | 1 | 0 | AP的MAC | 源MAC | 目的MAC | 无 |

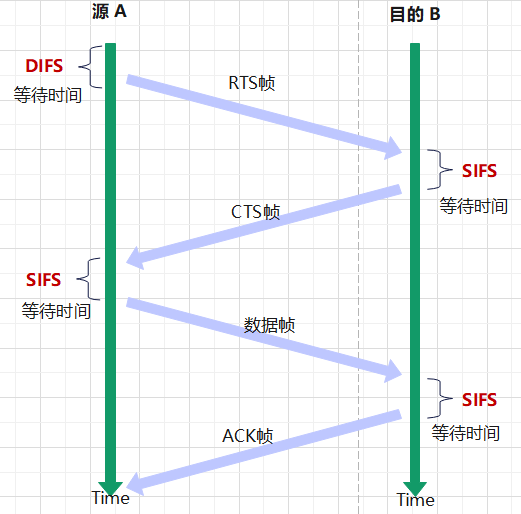
**②RTS帧**（请求发送）共20字节

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 字节 | 2 | 2 | 6 | 6 | 4 |
|  | 帧控制 | 持续期 | 接收地址 | 发送地址 | FCS |

**③CTS/ACK帧**（清除发送，确认）共14字节

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 字节 | 2 | 2 | 6 | 4 |
|  | 帧控制 | 持续期 | 接收地址 | FCS |

**\*源/目之间的 帧交互**：（A向B发送一个数据帧的过程）



DIFS > PIFS > SIFS

DIFS：分布协调功能帧间间隔

PIFS：集中协调功能帧间间隔

SIFS：短帧间间隔

**★IPv4地址**

1981年ipv4发布，分类的ip地址（InternetProtocol v4）

ipv4的5类地址：

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| A类 | 0xxx xxxx | 24位主机号 | | | 1.x.x.x | 126.x.x.x |
| B类 | 10xx xxxx | xxxxxxxx | 16位主机号 | | 128.1.x.x | 191.255.x.x |
| C类 | 110x xxxx | xxxxxxxx | xxxxxxxx | 8位主机号 | 192.0.1.x | 223.255.255.x |
| D类 | 1110 xxxx | 28位多播地址（组播） | | | 224.x.x.x | 239.x.x.x |
| E类 | 1111 xxxx | 保留 28位 | | | 240.x.x.x | 255.x.x.x |

127.x.x.x 环回地址

128.0.x.x 不指派

192.0.0.x 不指派

\*子网的划分，1985年 RFC 950

\*VLSM（Variable Length Submit Mask）可变长子网掩码，在一个划分子网的网络中可同时使用几个不同的子网掩码

\*CIDR（Classless Inter-Domain Routing）无分类编址，使用斜线记法，如 192.2.4.0/24

无分类域间路由选择，构成超网

**★私有地址段（专用地址）**

|  |  |  |  |
| --- | --- | --- | --- |
|  | 网段 | 子网掩码 | 可用地址 |
| 1个A类 | 10.0.0.0/8 | 255.0.0.0 | 10.0.0.0/8 |
| 16个B类 | 172.16.0.0/12 | 255.240.0.0 | 172.16.0.0/16 ~ 172.31.0.0/16 |
| 256个C类 | 192.168.x.0/24 | 255.255.255.0 | 192.168.0.0/24 ~ 192.168.255.0/24 |

**★运营商级NAT地址段**

2011年2月3日IANA宣告ipv4地址枯竭

2012年4月RFC 6598 IANA-Reserved IPv4 Prefix for Shared Address Space 100.64.0.0/10

NAT444/NAT44方案：采用运营商级NAT，不用传统NAT设备

LSN（Large Scale NAT）专门为满足运营商级NAT而产生的设备，LSN可支持Full Cone NAT

|  |  |  |  |
| --- | --- | --- | --- |
|  | 网段 | 子网掩码 | 可用地址 |
| A类子网 | 100.64.0.0/10 | 255.192.0.0 | 100.64.0.0/16 ~ 100.127.0.0/16 |

**★ipv4组播地址**

组播MAC地址前25位固定为 01-00-5E-0 第25位为0（剩余23位可用）

IP组播地址有28位可用，后23位对应MAC组播地址中的后23位

所以每一个mac组播地址都对应32个组播ip地址

**组播地址分类：**

|  |  |
| --- | --- |
| 组播地址段 | 用途 |
| 224.0.0.0 ~ 224.0.0.255 | 预留组播地址（永久组播地址） |
| 224.0.1.0 ~ 238.255.255.255 | 临时组播地址，用户可用的地址，全网范围内有效 |
| 239.0.0.0 ~ 239.255.255.255 | 本地管理组播地址，仅在特定的本地范围内有效 |

**常用预留组播地址：**

|  |  |
| --- | --- |
| 224.0.0.0 | 基准地址，保留 |
| 224.0.0.1 | 所有主机的地址，包括路由器 |
| 224.0.0.2 | 所有组播路由器的地址 |
| 224.0.0.5 | OSPF |
| 224.0.0.6 | OSPF DR |
| 224.0.0.9 | RIP-V2路由器 |
| 224.0.0.13 | PIM路由器组播地址 |
| 224.0.0.18 | VRRP |
| 224.0.0.22 | IGMP v3地址 |
| 224.0.0.251 | mDNS （5353端口） |
| 224.0.0.252 | LLMNR本地链路多播名称解析 |
| 239.255.255.250 | SSDP简单服务发现协议（1900端口） |

**★IPv4报文**

IPv4在IETF于1981年9月发布的RFC 791中被描述

IP协议规定在互联网中所有主机和路由器必须能够授受长度不超过576字节的IP数据报，当大于576字节时可先了解目的主机是否能接受此长度的报文，不能则分片

ipv4报文格式：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 长度bit | 0~3 | 4~7 | 8~16 | 16~18 | 19 ~ 31 |
|  | 版本 | 首部长度 | 区分服务 | 总长度 | |
|  | 标识ID | | | 标志 | 片偏移offset |
|  | 生存时间 | | 上层协议 | **首部校验和** | |
|  | 源ip地址（4字节） | | | | |
|  | 目的ip地址（4字节） | | | | |
|  | 可选字段 + 填充，4字节的倍数 | | | | |
|  | 数据部分（0~65515） | | | | |

①版本，4bit，指ip协议的版本，值为4表示ipv4，值为6表示ipv6

②首部长度，4bit，单位4字节取值5~15表示20~60字节，固定最小长度为5\*4=20字节

③区分服务DSCP，服务类型8bit ：

|  |  |  |  |
| --- | --- | --- | --- |
| bit | 3 | 4 | 1 |
|  | 优先级 | TOS（DTRC） | 保留位 |

TOS的4bit中同时只能有一位置1，保留位为0

④总长度（首部+数据）16bit，单位字节，当IP报文长度大于链路层的MTU时必须分片

⑤标识Identification，16bit，计数器，每产生一个ip报文，计数器就加1

⑥标志flag，3bit，

|  |  |  |  |
| --- | --- | --- | --- |
| bit | 1 | 1 | 1 |
|  | 0 | DF | MF |

DF置0表示允许分片，DF置1表示不允许分片

MF置0表示当前报文为最后一个分片，DF置1表示后面还有分片

⑦片偏移，13bit，单位8字节，每个分片的数据部分长度一定是8字节的整数倍

⑧生存时间TTL，8bit，TTL=0,1时，只能在同一物理网络中传递

⑨上层协议，8bit，10进制时对应的协议：

|  |  |  |  |
| --- | --- | --- | --- |
| 6 | TCP | 17 | UDP |
| 1 | ICMP | 47 | GRE |
| 89 | OSPF | 0 | IP |
| 50 | ESP | 41 | IPv6 |
| 51 | AH | 27 | RDP |

⑩首部校验和（只对首部校验，反码相加，Head Checksum

**★TCP报文及建立/关闭连接**

TCP（Transmission Control Protocol）传输控制协议，

熟知端口号： 0 ~ 1023

登记端口号： 1024 ~ 49151

短暂/用户端口号： 49152 ~ 65535

报文：

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| bit | 0 ~ 3 | 4 ~ 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 ~ 31 |
|  | 源端口 16bit | | | | | | | | 目的端口 16bit |
|  | 序列号 SEQ | | | | | | | | |
|  | 确认号 AN | | | | | | | | |
|  | 首部长度 | 保留 | URG | ACK | PSH | RST | SYN | FIN | 窗口大小 |
|  | 校验和 16bit | | | | | | | | 紧急指针 |
|  | 选项+填充 | | | | | | | | |

⑤首部长度，4bit，单位4字节（5~15）x 4 == 20 ~ 60 字节

⑥保留字段，6bit，未使用。之后6bit为控制字段

⑦控制字段：

URG == 1时，紧急指针有效；ACK == 1时，确认号有效，该报文为确认报文

PSH == 1时，接收方应尽快将这个报文交给应用层处理

RST == 1时，重建连接

SYN == 1时，发起一个连接

FIN == 1时，表示发送端完成发送任务，关闭连接

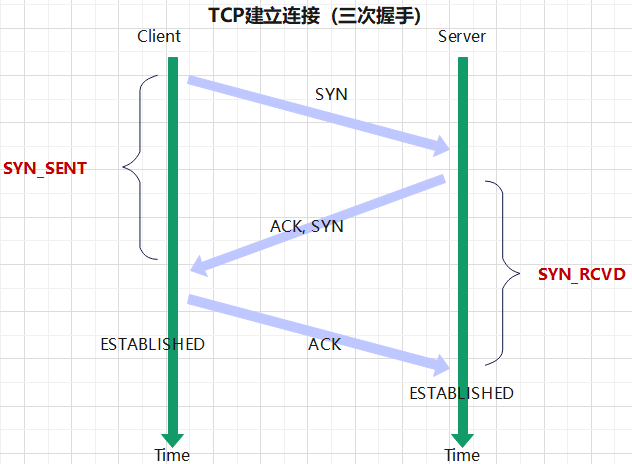
⑧窗口大小，16bit，用来配合TCP中的流量控制，单位字节，该字段值表示在确认号AN字段给出的字节数后面还可发送的字节数

为了不在IP层分片，TCP层本身可以根据MSS（Maximum Segment Size）进行分片

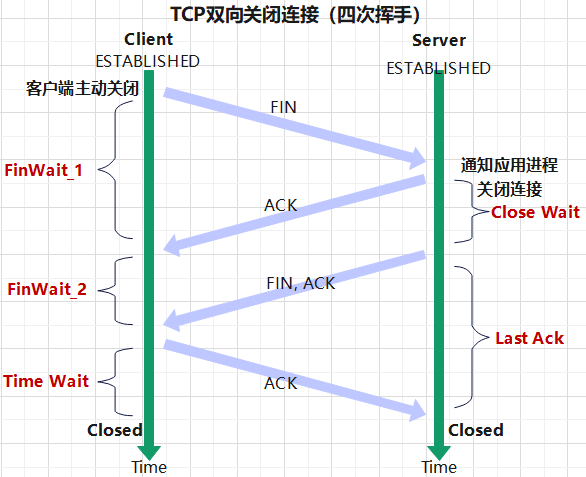
ipv4 tcp adjust-mss 1448 (1460 - 12) 建议在1024到1300之间

ipv6 tcp adjust-mss 1428 (1440 - 12 )

★三次握手建立连接



★4次挥手关闭连接



**★UDP报文**

UDP（User Datagram Protocol）用户数据报协议，无连接

报文：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 字节 | 2 | 2 | 2 | 2 | 0 ~ 65535 |
|  | 源端口 | 目的端口 | 总长度 | 校验和 | 数据 |

UDP本身不支持分片，数据长度推荐（1500 - 20 - 8）== 1472以内，局域网

公网建议（576 - 20 - 8）== 548 字节以内

**★TCP/UDP伪首部**

TCP/UDP的校验和都引入了伪首部的概念，只用于求校验和

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 字节 | 4 | 4 | 1 | 1 | 2 |  |
|  | 源IP | 目的IP | 0 | 6 tcp 17 udp | 长度 | TCP/UDP正式报文 |

**★常用端口号（tcp/udp）**

一般指报文的目的端口，或者服务端监听的端口

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 端口号 | 协议 | 传输类型 | 端口号 | 协议 | 传输类型 |
| 20 | FTP数据 | Tcp | 21 | FTP控制 | Tcp |
| 22 | SSH | Tcp | 23 | Telnet | Tcp |
| 25 | SMTP | Tcp | 42 | Wins复制 | Tcp |
| 53 | DNS | Tcp & Udp | 63 | Whois |  |
| 67 | DHCP Server | Udp | 68 | DHCP Client | Udp |
| 69 | TFTP | Udp | 80 | HTTP | Tcp |
| 88 | Kerberos | Tcp | 110 | POP3 | Tcp |
| 111 | RPC linux | Tcp | 119 | NNTP网络新闻 |  |
| 123 | NTP, SNTP | Udp | 135 | RPC windows | Tcp |
| 137 | NetBIOS名称解析 | Tcp & Udp | 138 | NetBIOS数据报服务 | Udp |
| 139 | NetBIOS会话服务 | Tcp | 143 | IMAP 4 |  |
| 161 | SNMP | Udp | 162 | SNMP trap | Udp |
| 389 | LDAP非加密 | Tcp | 443 | HTTPS | Tcp |
| 445 | SMB微软DS | Tcp & Udp | 464 | kerberos密码v6 | Tcp |
| 465/994 | SMTP ssl | Tcp | 514 | Syslog日志服务器 | Udp |
| 520 | RIP | Udp | 587 | SMTP tls | Tcp |
| 593 | RPC over HTTPS | Tcp | 993 | IMAP ssl |  |
| 994 | SMTP ssl |  | 995 | POP3 ssl |  |
| 1900 | ssdp简单服务发现 | Tcp & Udp | 2049 | nfsd | udp |
| 3389 | Mstsc | Tcp & Udp | 5060 | SIP | U |
| 5353 | mDNS | Udp | 5722 | RPC |  |
| 9389 | AD DS Web Services | Tcp | 20048 | Mountd | Tcp & Udp |
| 636 | LDAP ssl 加密 | Tcp | 1723 | PPTP vpn | Udp |
| 500,1701  5500 | L2TP vpn | Udp | 500  4500 | IPsec vpn | Udp |
| 860  3260 | iSCSI | Tcp | 1433  1434 | SQL Server | Tcp |
| 3306 | Mysql |  | 1521 | Oracle db |  |
| 5000 | DB2 |  | 27017 | Mongo DB |  |
| 6379 | Redis |  | 11211 | Memcached |  |
| 1645 | Radius认证 思科 | Udp | 1646 | Radius计费 思科 | Udp |
| 1812 | Radius认证 标准 | Udp | 1813 | Radius计费 标准 | Udp |
| 4789 | vxlan 交换机IANA | Udp | 8472 | vxlan linux | Udp |

**★ICMP报文（ping）**

ICMP（Internet Control Message Protocol）网际控制报文协议

icmp报文封装在ip数据报中，作为ip报文的数据部分（ipv4上层协议号为 1）

报文：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 字节 | 1 | 1 | 2 | 4 |  |
|  | 类型 | 代码 | 校验和 | id | icmp数据 |

|  |  |  |  |
| --- | --- | --- | --- |
| 类型 | 差错报告 | 类型 | 询问报文 |
| 3 | 终点不可达 | 8 | 回送echo请求 |
| 4 | 源点不可达 | 0 | 回送echo回答 |
| 11 | 时间超时，TTL为0 | 13 | 时间戳请求 |
| 12 | 参数问题 | 14 | 时间戳回答 |
| 5 | 改变路由Redirect |  |  |

②代码字段是为了进一步区分某种类型中的不同情况（如 引起差错的具体原因）

**★ARP报文（封装在mac帧上）**

ARP（Address Resolution Protocol）地址解析协议，根据ip地址查询相应的mac地址

RAPR（Reverse Address Resolution Protocol）反向地址解析协议，mac地址==> ip地址

网络上的主机都可以自主发送ARP应答消息，其他主机收到应答报文之后不会检测该报文的真实性就会将其记入本地ARP缓存

ARP报文封装在mac帧中，mac帧中的类型字段值为0x0806

报文：

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 字节 | 2 | 2 | 1 | 1 | 2 | 6 | 4 | 6 | 4 |
|  | 硬件类型 | 协议类型 | 硬件地址 长度 | 协议地址 长度 | OP | 发送者 硬件地址 | 发送者ip | 目标硬件 地址 | 目标ip |

①硬件类型，值为1时表示以太网

②协议类型，表明上层协议类型，0x0800为ip协议

③硬件地址长度，表示物理地址长度，单位字节，以太网mac地址值为6

④协议地址长度，上层协议为ipv4时，此值为4

⑤OP操作类型：

1为arp请求，2为arp应答，

3为rarp请求，4为rarp应答

应答包的源mac地址一般应该是请求包中的目标ip对应的mac，也可由网关代答

交互过程：

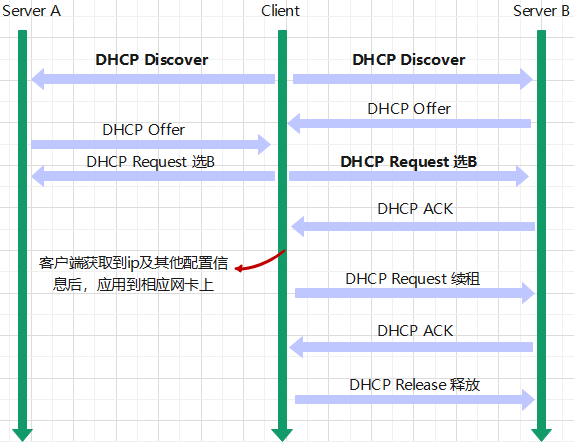
刚开始主机A不知道主机B的mac地址，于是A广播发送一个arp请求包，目标ip为B的

B收到后发现目标ip是自己，就回复A一个单播应答包，A收到后记录B的ip--mac信息，之后A再刷新时，以单播arp去查询B的mac地址

**★DHCP协议**

DHCP（Dynamic Host Configuration Protocol）动态主机配置协议，报文封装在UDP数据中，服务端监听67号端口，客户端监听68号端口（属于应用层协议）

工作过程：



**①发现阶段**：dhcp客户端以广播方式（255.255.255.255）发送dhcp Discover报文来寻找dhcp服务器

**②提供阶段**：同一广播域中每个dhcp服务器收到客户端发出的discover报文后，都会作出响应，向客户端发回一个dhcp Offer报文（含有出租的ip地址及其他配置）

**③选择阶段**：客户端只接受第一个收到的dhcp offer，然后以广播方式回答一个dhcp Request报文，目的是通知所有dhcp服务器，它（客户端）将选择某台dhcp服务器所提供的ip等配置信息的offer

**④确认阶段**：dhcp服务器收到客户端的request报文后，便向客户端发送一个包含各配置信息的dhcp Ack确认报文，客户端便将这些信息配置到网卡上（若客户端发现这些信息不可用，则回复一个dhcp Decline报文，通知服务器禁用该ip，客户端再重新发送dhcp discover）

**⑤重新登录**：客户端获取过一次配置后，之后每次登录网络时就不用再发discover报文，而是直接发一个含前一次获取的配置信息的dhcp Request报文给上一次的dhcp server，上一次的dhcp服务器收到后，就回一个ack，如果之前分配的ip不想再分配给这个客户端了，服务器就会回复一个dhcp Nack报文，客户端发现nack后，就重新发dhcp discover报文

**⑥更新租期**：dhcp服务器给客户端分配的ip都有一个租期，租期满后，服务器会回收此ip，dhcp客户端从开始使用ip时就会根据报文里的租期T设置2个计时器：

T1 == 0.5 T

T2 == 0.875 T

T1时间到后，客户端发送dhcp Request报文要求更新租用期，服务器同意则返回ack，不同意则返回nack，若客户端没有收到回复，则在T2时再发送一次Request续租

**⑦释放阶段**：客户端可以随时提前终止ip地址的使用，发送dhcp Release报文给服务器

**★DHCP报文**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 字节 | 1 | 1 | 1 | 1 |
| 4 | OP | Htype | Hlen | hops |
| 4 | Xid（Transaction ID） | | | |
| 4 | secs | | flags | |
| 4 | Ciaddr （Client ip addr） | | | |
| 4 | Yiaddr （Your Client ip） | | | |
| 4 | Siaddr （Next Server ip） | | | |
| 4 | Giaddr （Relay Agent ip） | | | |
| 16 | Chaddr (Client hardware addr） | | | |
| 64 | Sname （Server host name） | | | |
| 128 | File （Boot file name） | | | |
|  | options 可变长 | | | |

OP：报文的操作类型，1表示请求报文，client发给server的，2为应答报文，server发给client的

Htype：客户端的硬件地址类型，1表示以太网mac地址

Hlen：客户端的硬件地址长度，单位字节

hops：dhcp报文经过的dhcp中继的数目

Xid：客户端通过dhcp discover报文发起一次请求时选择的随机数，相当于请求标识，在一次请求中，所有报文的xid都一样

Secs：dhcp客户端从获取到ip地址或续约ip到现在 经过的时间，单位秒

Flags：只使用最左边的一个bit，用来标识dhcp服务器应答报文是采用单播（0）还是广播（1）的方式

Ciaddr：客户端的ip地址，仅在ack报文中显示，其他报文均为0

Yiaddr：服务器分配给客户端的ip，仅在offer和ack报文中

Siaddr：若客户端需要通过网络启动，则此字段写开机程序文件所在的服务器ip，offer,ack报文中

Giddr：客户端发出请求报文后经过的第一个dhcp中继的ip，若没有则置0

Chaddr：客户端的mac地址，一般用6个字节，之后10字节以0填充

Sname：为客户端分配ip地址的dhcp服务器名称，offer,ack中

File：服务端为client指定的启动文件名称及路径信息，稍后以tftp协议传送，仅offer报文中

Options：可选字段，可有多个可选字段，格式为TLV形式：

|  |  |  |  |
| --- | --- | --- | --- |
| 字节 | 1 | 1 | 变长 |
|  | 类型代码 | 长度（后面数据的） | 数据 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 代码 | 长度 | 数据说明 | 代码 | 长度 | 数据说明 |
| **1** | 4 | 子网掩码 | **3** | 4 x N | 默认网关，可有多个 |
| **6** | 4 x N | dns服务器ip，可有多个 | 12 | N | 客户端主机名 |
| 15 | N | 域名称 | **51** | 4 | 有效租期（秒） |
| 53 | 1 | 报文类型 | **43** | N | Vendor Specifier厂商自定义 |
| 60 | N | Class ID | 67 | N | Boot File Name |
| 82 | N | 中继代理信息 |  |  |  |

#

**★NAT（udp转换）4种类型**

Tuple表示元组，下面指ip:port

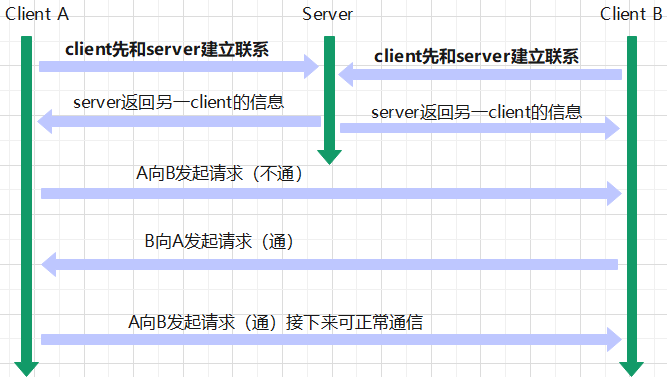
NAT表示把内部源ip:port（tuple）转为外部源tuple，再去访问目标tuple

根据UDP的tuple转换方式，可分为4种类型：

|  |  |
| --- | --- |
| Full Cone NAT | 全锥形NAT，接收任意外部tuple的回包，都能回到内部tuple |
| Restricted Cone NAT | 受限锥形NAT，只接收内部tuple访问过的目标ip的回包，对port没要求 |
| Port Restricted Cone NAT | 端口受限锥形NAT，只接收内部tuple访问过的目标ip：port的回包 |
| Symmetric NAT | 对称NAT，只有同一内部源tuple访问同一目标tuple时，才被转为同一外部源tuple |

只有前3种锥形NAT才支持建立P2P通道（俗称 打洞）

打洞过程（需要借助一个第三方Server）



**★IPv6地址**

1993年9月IETF互联网工程任务组建立了一个临时的ad-hoc下一代ip领域来专门解决下一代ip的问题

1994年7月25日IETF采纳了IPng模型

2012年6月6日国际互联网协会举行了世界IPv6启动纪念日，全球ipv6网络正式启动

IPv6在1994年7月15日制定，ipv6地址采用冒分十六进制表示，长度16字节

如： FF01:2D88:3014:2000:5500:6600:7700:8081 共8块，每块2字节

一个全0的块简写为一个0，多个全0的块化简为 ::

一个ipv6地址只能出现一个 :: ，化简最长的0，等长靠左

**ipv6分为3类地址：**

①单播地址 Unicast ：同ipv4单播概念，唯一地标识一个接口

②组播地址 Multicast ：一个组播地址对应一组接口，发往组播地址的包会被这组接口的所有接口接收。ipv6的广播功能是由组播来完成的

③任播地址 Anycast ：一个任播地址对应一组接口，发往任播地址的包 会被这组接口其中的某一个接口接收，具体由哪个接口收，由路由协议决定

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | 地址前缀 | ipv6地址前缀标识 |
| 单播地址 | 未指定地址 | 0000 0000 ...00（128个0） | ::/128 |
| 环回地址 | 0000 0000 ...01（127个0，1个1） | ::1/128 |
| 链路本地地址 | 1111 1110 10xx xxxxxxxx | FE80::/64 |
| 唯一本地地址 | 1111 110x xxxx xxxxxxxx | FC00::/7 FD00::/7 |
| 站点本地地址 | 1111 1110 11xx xxxxxxxx | FEC0::/10 （已弃用） |
| 全球单播地址 | 其他形式 | 非F,非0 开头的地址 |
| 组播地址 |  | 1111 1111 xxxx xxxxxxxx | FF00::/8 |
| 任播地址 |  | 从单播地址空间中分配 |  |

①未指定地址 ::/128 只能做源ip，不能做目的ip

②环回地址 ::1/128 等同于127.0.0.1

③全局单播地址 等同于ipv4中的公网ip，可以在ipv6 互联网上进行全局路由和访问，允许路由前缀聚合

④本地单播，相当于ipv4中的局域网专用地址：

|  |  |
| --- | --- |
| 链路本地地址 | 仅用于单个链路，不能跨vlan，不能在不同子网中路由 |
| 唯一本地地址 | 是本地全局的，用于本地通信，不通过Internet路由 |
| 站点本地地址 | 已被唯一本地代替 |

⑤任播地址：只能分配给ipv6路由器

★IPv6基础知识

**★IPv6报文**

ipv6报文：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| bit: | 0~3 | 4~11 | 12~ | 16~23 | 24~31 |
|  | 版本号 | 流量等级 | 流标签 | | |
|  | 载荷长度 0~15 | | | 下一报头 | 跳数限制 |
|  | 源地址 128 bit | | | | |
|  | 目的地址 128 bit | | | | |
|  | 扩展首部 变长 | | | | |

①版本号4位，表示ip协议版本，值为6

②流量等级，8位（Traffic Class）也叫流量类别，主要用于QoS

③流标签，20位（Flow Label）用于标识同一个流量里面的报文，用于某些对连接的服务质量有特殊要求的通信

④载荷长度，16位，单位字节，ipv6头部之后的长度（含扩展首部及数据）

⑤下一报头，8位，指明下一报头的类型或上层协议：

0x3A 表示icmp v6 ； 0x2C 表示数据分片， 0x11（十进制17）表示UDP

⑥跳数限制，8位，生存时间，TTL

\*ipv6主首部中没有错误校验，为网络设备节省了执行校验的时延

ipv6的上层协议申明与数据分片都在扩展首部中完成

**章n、网络常用协议及报文**

★FTP

★TFTP

★NTP

★DNS

★PPPoE

**★HTTP报文**

Http工作原理：

①客户端浏览器与web服务器的http端口（一般为80/tcp）建立一个tcp连接

②客户端向服务器发送一个文本的请求报文

③服务器接受请求并返回http响应

④客户端浏览器解析html内容并展示

⑤释放连接：

若connection模式为close则服务器主动关闭tcp连接

若connection模式为keepalive则该连接保持一段时间，该时间段内可连续请求

HTTP请求报文：

|  |  |
| --- | --- |
| GET /dir1/fds/dd/index.html HTTP/1.1 | 请求行 |
| Host: www.xxx.com:4600 | 请求头部 一行一个 参数 |
| User-Agent: Mozilla/5.0 Firefox/77.0 |
| Connection: keep-alive |
| Accept-Language: en-US,en;q=0.5 |
| Accept-Encoding: gzip, deflate |
| 其他参数 |
|  | **空行** |
| 请求的数据data | 请求数据 |

请求行格式： 请求方法 URI http版本

请求方法一般有 GET PUT POST

HTTP回答报文：

|  |  |
| --- | --- |
| HTTP/1.1 200 OK | 状态行 |
| Date: Sun, 03 Jul 2022 02:54:59 GMT（生成响应的时间） | 消息报文附加信息 |
| Connection: Keep-Alive |
| Keep-Alive: timeout=5, max=100 |
| Server: Apache |
| Content-Length: 204 |
| Content-Type: text/html; charset=iso-8859-1 |
|  | **空行** |
| 数据正文 <html> xxxx </html> | 数据正文 |

状态行格式： http版本 状态码 状态消息

状态码：

|  |  |
| --- | --- |
| 1xx | 指示信息 表示请求已接收，继续处理 |
| 2xx | 成功，请求已被成功接收处理 |
| 3xx | 重定向，要完成请求必须进行更进一步的操作，URL重定向 |
| 4xx | 客户端错误，请求有语法错误或请求无法实现 |
| 5xx | 服务端错误，服务器未能实现合法的请求 |

200 : 成功，表示访问成功，正常状态。

301 : 永久移动，表示本网页已经永久性的移动到一个新的地址，在客户端自动将请求地址改为服务器返回的新地址。

302 : 临时重定向，表示网页暂时性的转移到一的新的地址，客户端在以后可以继续向本地址发起请求。

303 : 表示必须临时重定向，并且必须使用GET方式请求。

304 : 重定向至浏览器本身，当浏览器多次发起同一请求，且内容未更改时，使用浏览器缓存，这样可以减少网络开销。

401 : 表示协议格式出错，可能是此IP地址被禁止访问该资源，与403类似。

403 : 表示没有权限，服务器拒绝访问请求。

404 : 这是最常见的错误，表示找不到系统资源，但是只是暂时性地。

500 : 表示服务器程序错误，一个通用的错误信息。

503 : 表示服务器繁忙，或者服务器负载，通常这只是一个临时状态

★SSL数字证书

**★URI和URL**

URI（Uniform Resource Identifier）统一资源标识符，资源

**URL**（Universal Resource Locator）统一资源定位符，资源 + 访问方法

uri定义了一个资源，如： www.xxx.com:4888/path/to/resourcexx

url定义了一个资源同时说明如何访问这个资源，如： https://xxx.com/path/to/resourcey

让uri成为url只需要增加访问机制

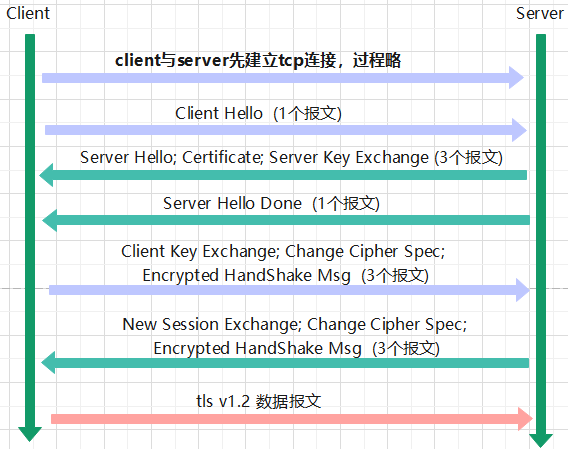
URL格式：

scheme://host:port/path/[params][?query-string][#anchor]

scheme有： http, https, ftp, ssh, ed2k 等

**★TLS交互报文**

TLS v1.2交互过程：



Client Hello / Server Hello报文：

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 字节 | 1 | 2 | 2 | 1 | 3 | 2 | 变长 |
| 字段： | 0x16 | 0x0303 | Length1 | 0x01 | Length2 | 0x0303 | Random; Session-ID; Cipher Suite |
| 说明： | CT | Ver1 |  | HT |  | Ver2 | 随机数，会话id，加密套件 |

①CT（Content Type）值为0x16时表示 HandShake

②Version，值为0x0303时表示tls 1.2 ；值为0x0301时表示tls 1.0

③Length

④HT（HadnShake Type）值为0x01时表示Client Hello

⑤Length

⑥Version，值为0x0303时表示tls 1.2

⑦随机数（含时间）；Session-id；加密套件（2字节代表一种），或者是其他参数

Certificate报文：

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 字节 | 1 | 2 | 2 | 1 | 3 | 3 | 变长 |
| 字段： | 0x16 | 0x0303 | Length1 | 0x0B | Length2 | len | Certificate data |
| 说明： | CT | Ver1 |  | HT |  | 证书长度 | 证书内容 |

④HT（HadnShake Type）值为0x0B 时表示Certificate

Server Key Exchange报文：

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 字节 | 1 | 2 | 2 | 1 | 3 | 变长 |
| 字段： | 0x16 | 0x0303 | Length1 | 0x0C | Length2 | EC-DH-参数 |
| 说明： | CT | Ver1 |  | HT |  | 加密套件 |

④HT（HadnShake Type）值为0x0C 时表示Server Key Exchange

Server Hello Done报文：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 字节 | 1 | 2 | 2 | 1 | 3 |
| 字段： | 0x16 | 0x0303 | Length1 | 0x0E | 0x00000 |
| 说明： | CT | Ver1 |  | HT | Length2 |

④HT（HadnShake Type）值为0x0E 时表示Server Hello Done

Client Key Exchange报文：

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 字节 | 1 | 2 | 2 | 1 | 3 | 变长 |
| 字段： | 0x16 | 0x0303 | Length1 | 0x10 | Length2 | EC-DH-参数 |
| 说明： | CT | Ver1 |  | HT |  | 加密套件 |

④HT（HadnShake Type）值为0x10时表示Client Key Exchange

New Session Ticket报文：

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 字节 | 1 | 2 | 2 | 1 | 3 | 变长 |
| 字段： | 0x16 | 0x0303 | Length1 | 0x04 | Length2 | TLS Session Ticket |
| 说明： | CT | Ver1 |  | HT |  | 含Lifetime，tk长度及内容 |

④HT（HadnShake Type）值为0x04时表示New Session Ticket

数据报文：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 字节 | 1 | 2 | 2 | 变长 |
| 字段： | 0x17 | 0x0303 | Length1 | data |
| 说明： | CT | Ver1 |  | 加密的数据 |

①CT（Content Type）值为0x17时表示数据报文

**章n、路由与交换**

在

**★网络RFC标准**

引导地址： https://www.rfc-editor.org/rfc-index.html

FTP 文件传输协议 RFC 959

HTTP 超文本传输协议 RFC 1945,2616

IMAP4 因特网信息访问协议第四版 RFC 1730

ISAKMP Interne安全连接和密钥管理协议 RFC 2048

DNS 域名系统 RFC 4343

DHCP 动态主机配置协议 RFC 2131

BOOTP 引导协议 RFC 951

NTP 网络时间协议 RFC 958

NNTP 网络新闻传输协议 RFC 977

POP3 邮局协议第三版 RFC 1939

Radius 远程用户拨号认证服务协议 RFC 2138

RLOGIN 远程登陆协议 RFC 1258,1282

SMTP 简单邮件传输协议 RFC 821,2821

SNMP 简单网络管理协议 RFC 1157

TACACS 终端访问控制器访问控制系统协议 RFC 1492

TELNET 终端仿真协议 RFC 854

TFTP 简单文件传输协议 RFC 1350

X-Window RFC 1198

NBSSN NetBIOS会话服务协议 RFC 1001

TLS 传输层安全协议 RFC 2246

LDAP 轻量级目录访问协议 RFC 1777

RPC 远程过程调用协议 RFC 1050,1057,1831

TCP 传输控制协议 RFC 793

UDP 用户数据报协议 RFC 768

EGP 外部网关协议 RFC 827

OSPF 开放最短路径优先协议 RFC 2178,2328

ICMP Internet控制信息协议 RFC 792

ICMPv6 Internet控制信息协议第6版 RFC 1885,2463

IGMP Internet组管理协议 RFC 1112, 2236,3376

IP ipv4互联网协议 RFC 791

IPv6 互联网协议第6版 RFC 1883,2460

PIM-SM 稀疏模式独立组播协议 RFC 2362

PIM-DM 密集模式独立组播协议 RFC 3973

RIP2 路由信息协议第2版 RFC 2453

RIPng IPv6路由信息协议 RFC 2080

VRRP 虚拟路由器冗余协议 RFC 2338,3768

AH 认证头协议 RFC 2402

ESP 安全封装有效载荷协议 RFC 2406

ARP 地址解析协议 RFC 826

RARP 逆向地址解析协议 RFC 903

MPLS 多协议标签交换协议 RFC 3031,3032

L2TP 第二层隧道协议 RFC 2661

PPTP 点对点隧道协议 RFC 2637

**★网络互连设备**

物理上的互连能力：指所支持的物理接口，能连接的物理介质类型

协议上的互连能力：指工作在不同协议类型的网络之间，实现不同协议数据包的转换

**①中继器 Repeater**

主要功能是对接收到的信号进行再生放大，以延伸网络的传输距离，提供物理层的互连。

（1）过滤电磁干扰(MEI)和射频干扰(RFI)引起的信号干扰和噪声

（2）放大和修整进入的信号，使重新传输更精确

（3）对信号重定时

（4）在所有网段上复制信号

5个网段，4个中继器，3个可挂计算机

**②信线器 Hub**

Hub是双绞线以太网对网络进行集中管理的最小单元，信线器是一个共享设备，其实质是一个多端口的中继器，是一个信号放大和中转的设备，不具备自动转发和自动寻址能力

所有传输到Hub的数据均被广播到与之相连的各个端口，遵循CSMA/CD控制方式，因此所有Hub的端口同属于一个冲突域，连接到Hub的每台工作站共享Hub的实际带宽

Hub主要用于共享式以太网的组建

**③网桥 Bridge**

网桥也叫桥接器，一般有2个端口，是连接2个局域网的存储转发设备，工作在数据链路层，网桥可以截获所有的网络信息，并读取每一帧的目标mac地址，以确定此帧是否应该转发到某一网段。

网桥在发送数据帧前，通常要接收到完整的数据帧并执行帧检测序列FCS后，才开始转发该数据帧。

**④交换机 Switch**

交换机在转发数据时，根据事先存储的mac地址表选择目标主机转发数据，所以交换机各端口的带宽是独立的。交换机初始工作时都有一个获得所连接主机的mac地址的过程，这一过程都是由交换机自动学习完成的

交换技术：

（1）直通方式 Cut Through

不需要对数据帧做差错校验和附加处理，因此时延最小 转发速度最快，不能过滤出错的帧

1，端口在接收帧的14个字节后，交换模块就取出帧的目的地址并送交端口查询模块

2，端口查询模块从地址映射表中查出帧所要转发的正确𦫢号，并通知交换模块，交换模块将帧发送到正确的端口线路上

（2）存储转发方式 Store adn Forward

需要对帧进行差错校验以及其他附加服务，如速率匹配，协议转换等，因此必须设置缓冲器

1，端口将整个数据帧完整地接收下来并存在共享缓冲器中

2，对帧进行差错校验，校验正确后将帧交给交换模块

3，交换模块取出帧头部交给端口查询模块进行地址转换

4，端口查询模块查出帧所需转发的正确端口号，并通知交换模块

5，交换模块将处理过的帧送还共享缓冲器，并发送到正确的端口线路上

（3）无碎片直通方式 Fragment Free Cut Through

先保存帧的前64字节，如果是不健全的帧或有冲突的帧就立即舍弃，如正确就继续发送

**④路由器 Router**

基本任务是在广域网上提供最短、最优、最高带宽路径 的查询和包转发功能

还提供包过滤，组播，服务质量QoS，数据加密，流量控制，拥塞控制

**⑤网关 Gateway**

网关是指一种能使2种不同类型的网络系统进行通信的接口

网关(Gateway)又称网间连接器、协议转换器。网关在网络层以上实现网络互连，是复杂的网络互连设备，仅用于两个高层协议不同的网络互连。网关既可以用于广域网互连，也可以用于局域网互连。

**★级联和堆叠**

**①级联**：是指2台及以上的交换机通过一定的方式相互连接，如总线型，树型，星型等，进行级联时应尽量保证交换机中继链路具有足够的带宽，可采用链路聚合技术提升带宽

级联的交换机可以相距较远（在媒体允许的范围内）

级联一般采用普通端口，可以在 不同厂商的交换机间级联

**②堆叠**：是指将2台及以上的交换机组合起来共同工作，以便在有限的空间内提供尽可能多的端口，多台交换机经过堆叠形成一个堆叠单元，“最大可堆叠参数”指一个堆叠单元中所能堆叠的最大交换机数量，一个堆叠单元内的交换机之间距离很近。堆叠一般采用专用的堆叠模块和堆叠电缆，只能用同一厂商的交换机进行堆叠

**③设备虚拟化**：级联和堆叠是实现设备虚拟化的前提，将多台互相连接的交换机虚拟为一台逻辑设备进行管理，整个虚拟集群中一般只有一台交换机做管理功能，称为命令交换机，可在命令交换机上统一管理。

不同厂商对设备虚拟化有不同的实现方案

\*级联和堆叠是基于硬件实现的，设备虚拟化是软件实现的

**★链路层协议**

在

|  |  |  |
| --- | --- | --- |
| SDLC | 1974 |  |
| HDLC |  | 高级数据链路控制协议 |
| X.25 | 1976 | 分组以太网 |
| Ethernet | 1980 | 以太网 |
| PPP | 1992 | 点到点协议 |
| Frame Relay | 1992 | 帧中继 |
| ATM |  | 异步传输网 |

**★VLAN及vlan的划分**

VLAN（Virtual Local Area Network）虚拟局域网，由IEEE 802.1Q标准支持，1998年

vlan是由一些局域网网段构成的与物理位置无关的逻辑组，而这些网段具有某些共同的需求

**vlan的划分方法：**

①基于端口划分vlan，将局域网交换机中的几个端口指定成一个vlan

②基于mac地址划分vlan，根据局域网内工作站的mac地址划分vlan

③基于网络地址（ip地址）划分vlan

④基于用户定义规则划分vlan

**★STP生成树协议**

**★生成树协议类型**

Spanning Tree Protocol生成树协议 是一个数据链路层的管理协议，通过有选择性地阻塞网络冗余链路来达到消除网络环路的目的，同时 具备链路冗余备份的功能。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 协议类型 | 全称 | 标准 | 实例数 | 发布时间 |
| STP | Spanning Tree Protocol | IEEE 802.1D | 单 | 1998 |
| RSTP | Rapid STP | IEEE 802.1W | 单 | 2001 |
| MSTP | Multiple STP | IEEE 802.1S | 多 | 2002 |
| PVST | Per-Vlan Spanning Tree | 思科专有 | 多 |  |
| PVST**+** | PVST Plus 增强型PVST | 思科专有 | 多 |  |
| Rapid-PVST+ | 快速 增强型pvst | 思科专有 | 多 |  |

STP和RSTP都是单生成树实例，把整个交换机组成的网络当成一个生成树，

PVST、PVST+、Rapid-PVST+是多生成树实例，为每个vlan分配并维护一个生成树实例

MSTP为每个实例维护一个生成树，一个实例可映射多个vlan

**★STP工作流程**

**①选择根桥**（默认选择桥ID最小的设备为根桥，每个STP网络中只有一个根桥）

**②选择根端口**（仅在每台 **非根桥上** 选择到根桥路径开销最小的端口，自己设备上的所有端口选一个，一台非根桥设备上只有一个根端口）

**③选择指定端口**（2台设备之间，每一条链路之间选一个端口为指定端口，有根端口的，其对端一定为指定端口，没有根端口的链路按最小开销来选）

**④**根端口与指定端口设为**转发状态**，其他端口阻塞

一旦根桥、根端口和指定端口选举成功，整个树形拓扑就建立完毕了。

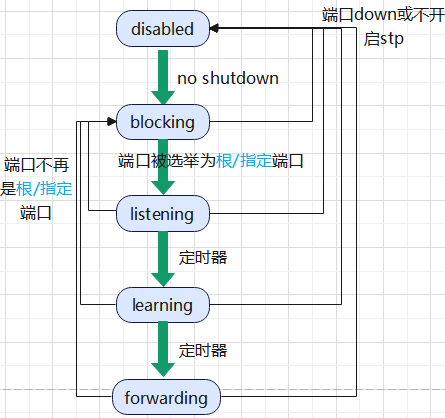
\*桥ID == 优先级 + MAC地址 #越小越优先

\*根端口 本交换机 到根桥的路径开销最小的端口（一台非根桥设备上**只有一个**根端口）

\*为一对交换机间选指定端口，一条链路（2台交换机之间的）的2个端口，谁到根桥开销小，谁就是指定端口（一台设备**可有多个**指定端口）

**★交换机stp 5种端口状态**

|  |  |  |
| --- | --- | --- |
| 状态 | 时间(秒) | 说明 |
| disabled |  | **禁用**，端口状态Down，不处理BPDU，也不转发数据流量 |
| blocking | 20 | **阻塞**，接收BPDU，不学习mac，不转发数据 |
| listening | 15 | **学习**，接收BPDU，不学习mac，不转发数据，开始STP计算，参与选举 根/指定 端口 |
| learning | 15 | **监听**，接收BPDU，学习mac，不转发数据 |
| forwarding |  | 正常**转发**，只有根端口及指定端口才能进入forwarding状态 |



blocking：如果一个端口处于阻塞状态，并在 一个最大老化时间(默认20秒)内 没有接收到新的BPDU报文，端口也会从阻塞状态转换为监听状态。

listening/learning：定时器默认15秒（转发延迟）

forwarding：正常转发状态，生成树稳定后，BPDU仍然会定时(默认每隔2秒)从各个交换机的指定端口发出，以维护链路的状态。如果网络拓扑发生变化，生成树就会重新计算，端口状态也会随之改变。

\*发生变化的交换机会在它的根端口上每隔 hello time 时间就发送TCN BPDU（拓扑变化通知BPDU）直到 生成树上游的指定网桥确认了该TCN BPDU为止。

\*当**根桥**收到TCN BPDU后，会发送TC（Topology Change,拓扑改变）置位的BPDU，通知整个生成树拓扑结构发生了变化。

\*新选出的根端口和指定端口要经过2倍的Forward Delay延时后才能进入转发状态，这个延时保证了新的配置消息传遍整个网络，从而防止了临时环路的产生。

\* Forward Delay Timer指一个端口处于Listening和Learning状态的各自持续时间，默认是15秒。即Listening状态持续15秒，随后Learning状态再持续15秒。这两个状态下的端口均不转发数据流量，这正是STP用于避免临时环路的关键。

• TCN BPDU报文主要用来向上游设备乃至根桥通知拓扑变化。

• TCA标记置位的配置BPDU报文主要是上游设备用来告知下游设备已经知道拓扑变化，通知下游设备停止发送TCN BPDU报文。

• TC标记置位的配置BPDU报文主要是上游设备用来告知下游设备拓扑发生变化，请下游设备直接删除桥MAC地址表项，从而达到快速收敛的目的。

**★生成树协议最短路径的选择过程:**

①首先比较本交换机到达根网桥的路径开销，选择开销最小的路径

②如果路径开销相同，则比较发送BPDU交换机的网桥id，越小越优先

③如果网桥id相同（去往根桥中间经过同一台交换机）则比较发送者交换机的端口id，最小优先

④若不同链路发送者的网桥id一致，则比较接收者的端口id

**★STP路径开销表**（一条链路上的2个端口只算一次，进的算，出的不算）

|  |  |  |  |
| --- | --- | --- | --- |
|  | 旧的 cost | 新修订 802.1D/1998 | 最新 802.1T/2001 cost |
| 10M | 100 | 100 | 2 000 000 |
| 100M | 10 | 19 | 200 000 |
| 1000M | 1 | 4 | 20 000 |
| 10G | 1 | 2 | 2 000 |

交换机收到TCN BPDU时，将MAC表老化时间缩为15秒？

**★RSTP**

RSTP仍然有 根端口/指定端口，阻塞端口改为 备份端口（Backup Port）或者 替代端口（Alternate Port）

备份端口是向树叶下游的口（指定端口的备份），替代端口是向上游的（根端口的备份）

RSTP将端口状态和端口角色区分开来

**★RSTP 4种端口角色**

|  |  |
| --- | --- |
| 根端口↑ | Root Port， |
| 指定端口↓ | Designed Port， |
| 备份端口↓ | Backup，由学习到自己发送的BPDU而阻塞，作为指定端口的备份，提供了另一条从根桥到相应网段的备份通路，向树叶下游的端口 |
| 替代端口↑ | Alternate，由学习到其他网桥发送的BPDU而阻塞的端口，提供了从指定桥到 根桥的另一条可切换路径，作为根端口的备份，向树上游的端口 |

**★RSTP 3种端口状态**

|  |  |  |  |
| --- | --- | --- | --- |
| 状态 | 时间(秒) | 说明 | RSTP端口角色 |
| discarding |  | **丢弃**，类似stp中 监听/阻塞/禁用 的集合，不处理BPDU，也不转发数据流量 | **根/指/Alter/Backup** |
| learning |  | **学习**，接收BPDU，学习mac，不转发数据 | **根/指** |
| forwarding |  | 正常**转发**，只有根端口及指定端口才能进入forwarding状态 | **根/指** |

在一个运行RSTP的网络中，检测拓扑是否发生变化只有一个标准：一个非边缘端口迁移到Forwarding状态。

**★RSTP所做的改进有：**

①新增2种端口角色，删除了3种端口状态，并将端口状态和端口角色解耦。而且在配置BPDU的格式中，充分利用Flag字段，明确了端口角色。

②配置BPDU的处理方式发生了变化。

（2.1）拓扑稳定后，对于非根桥设备，无论是否收到根桥传来的配置BPDU报文，都会自主地按照Hello Timer规定的时间间隔发送配置BPDU。

（2.2）如果一个端口在超时时间（超时时间＝Hello Time × 3 × Timer Factor）（2s x 3 x 3）内没有收到上游设备发送过来的配置BPDU，那么该设备认为与此邻居之间的协商失败。而不像STP那样需要先等待一个Max Age（20s）

（2.3）当一个端口收到上游的指定桥发来的RST BPDU报文时，该端口会将其与自身存储的RST BPDU进行比较。如果该端口存储的RST BPDU的优先级较高，则直接丢弃收到的RST BPDU，并立即向上游设备回应自身存储的RST BPDU。当上游设备收到回应的RST BPDU后，会根据其中相应的字段立即更新自己存储的RST BPDU。由此，RSTP处理次等BPDU报文不再依赖于任何定时器通过超时解决拓扑收敛，从而加快了拓扑收敛。

③引入快速收敛机制，包括Proposal/Agreement机制、根端口快速切换机制、边缘端口。

④引入多种保护功能，包括BPDU保护、根保护、环路保护、防TC-BPDU攻击。

★Proposal/Agreement机制

**★根端口快速切换机制**

如果RSTP网络中一个根端口失效，那么网络中最优的Alternate端口将成为根端口并直接进入Forwarding状态。因为通过这个Alternate端口连接的网段上必然有个指定端口可以通往根桥。

**★边缘端口**

RSTP增加 边缘端口（Edge Port）的概念，可直接进入转发状态，直接连接服务器的，不连接交换机，不会有BPDU入站，如果有BPDU入站则此边缘端口转为普通的RSTP端口

对于开启了生成树协议的交换机，每当有终端设备接入时会导致生成树重新计算收敛，这会导致终端设备获取IP地址的时间比较长。此时，可以关闭交换机上连接终端接口的生成树协议或将交换机上连接终端的端口配置成边缘端口

**★BPDU保护机制**

在RSTP网络中，正常情况下，边缘端口不会收到RST BPDU，启用了BPDU保护功能后，如果有人伪造RST BPDU恶意攻击交换机，边缘端口将被Error-Down。

未启用BPDU保护时，当边缘端口接收到RST BPDU时，交换机会自动将边缘端口设置为非边缘端口，并重新进行生成树计算，可能会导致业务流量中断。

**★根保护**

在指定端口上，配置根保护。对于启用Root保护功能的指定端口，其端口角色**只能保持为**指定端口。一旦启用Root保护功能的指定端口收到优先级更高的RST BPDU时，端口状态将进入Discarding状态，不再转发报文。在经过一段时间（通常为两倍的Forward Delay），如果端口一直没有再收到优先级较高的RST BPDU，端口会自动恢复到正常的Forwarding状态。

（防止在指定端口上接入一台优先级更高的交换机，导致抢了原来的根桥地位引起网络拓扑的变化）

Root保护功能**只能在指定端口上**配置生效。#开启根保护的端口不可以再配置环路保护

**★环路保护**

在运行RSTP协议的网络中，根端口和其他阻塞端口状态是依靠不断接收来自上游交换设备的RST BPDU维持。当由于链路拥塞或者单向链路故障导致这些端口收不到来自上游交换设备的RST BPDU时，此时交换设备会重新选择根端口。原先的根端口会转变为指定端口，而原先的阻塞端口会迁移到转发状态，从而造成交换网络中可能产生环路。

在启动了环路保护功能后，如果根端口或Alternate端口长时间收不到来自上游设备的BPDU报文时，则向网管发出通知信息（此时根端口会进入Discarding状态，角色切换为指定端口），而Alternate端口则会一直保持在阻塞状态（角色也会切换为指定端口），不转发报文，从而不会在网络中形成环路。直到链路不再拥塞或单向链路故障恢复，端口重新收到BPDU报文进行协商，并恢复到链路拥塞或者单向链路故障前的角色和状态。

环路保护功能**只能在根端口或Alternate端口上**配置生效。

#开启环路保护的端口不可再配置根保护

**★防TC-BPDU攻击**

交换设备在接收到TC BPDU报文后，会执行MAC地址表项和ARP表项的删除操作。如果有人伪造TC BPDU报文恶意攻击交换设备时，交换设备短时间内会收到很多TC BPDU报文，频繁的删除操作会给设备造成很大的负担，给网络的稳定带来很大隐患。

启用防TC-BPDU报文攻击功能后，在单位时间内，交换设备处理TC BPDU报文的次数可配置。如果在单位时间内，交换设备在收到TC BPDU报文数量大于配置的阈值，那么设备只会处理阈值指定的次数。对于其他超出阈值的TC BPDU报文，定时器到期后设备只对其统一处理一次。这样可以避免频繁的删除MAC地址表项和ARP表项，从而达到保护设备的目的。

**★RSTP和STP兼容**

RSTP可以和STP互相兼容，但是此时会丧失快速收敛等RSTP优势。

当一个网段里既有运行STP的交换设备又有运行RSTP的交换设备，STP交换设备会忽略RSTP BPDU；运行RSTP的交换设备在某端口上接收到运行STP的交换设备发出的配置BPDU，在两个Hello Time时间之后，便把自己的端口转换到STP工作模式，发送配置BPDU，从而实现了互相兼容。

**★术语缩写：**

AP（Alternate Port）替代端口

DP（Designed Port）指定端口

RP（Root Port）根端口

BID（Bridge ID）桥id

PID（Port ID）端口id

RPC（Root Path Cost）根路径开销

EP（Edge Port）边缘端口

**★MSTP**

为了弥补STP和RSTP的缺陷，IEEE于2002年发布的802.1s标准定义了MSTP。MSTP兼容STP和RSTP，既可以快速收敛，又提供了数据转发的多个冗余路径，在数据转发过程中实现VLAN数据的负载均衡。

MSTP把一个交换网络划分成多个域，每个域内形成多棵生成树，生成树之间彼此独立。每棵生成树叫做一个多生成树实例MSTI（Multiple Spanning Tree Instance），每个域叫做一个MST域（MST Region：Multiple Spanning Tree Region）。

MSTP网络中包含1个或多个MST域（MST Region），每个MST Region中包含一个或多个MSTI。组成MSTI的是运行STP/RSTP/MSTP的交换设备

MSTP区域（Region）只有三个属性都一致的交换机才认为是在同一个MST域内：

域名 Region-Name （32字节内）

修正号（修订级别） Revision-Number （2字节，int）

vlan到实例的映射 （1~64个实例）

公共生成树CST（Common Spanning Tree）是连接交换网络内所有MST域的一棵生成树。

如果把每个MST域看作是一个节点，CST就是这些节点通过STP或RSTP协议计算生成的一棵生成树。

公共和内部生成树CIST（Common and Internal Spanning Tree）是通过STP或RSTP协议计算生成的，连接一个交换网络内所有交换设备的单生成树。

内部生成树IST（Internal Spanning Tree）是各MST域内的一棵生成树。

IST是一个特殊的MSTI，MSTI的ID为0，通常称为MSTI0。

IST是CIST在MST域中的一个片段。

总根是CIST（Common and Internal Spanning Tree）的根桥。

域根（Regional Root）分为IST域根和MSTI域根。

IST域根如图1所示，在MST域中IST生成树中距离总根（CIST Root）最近的交换设备是IST域根。

一个MST域内可以生成多棵生成树，每棵生成树都称为一个MSTI。MSTI域根是每个多生成树实例的树根。如图2所示，域中不同的MSTI有各自的域根。

主桥（Master Bridge）也就是IST Master，它是域内距离总根最近的交换设备。如图1中的黄色交换机。

如果总根在MST域中，则总根为该域的主桥。

★根端口、指定端口、Alternate端口、Backup端口和边缘端口的作用同RSTP

**★ARQ**

自动重传请求（Automatic Repeat-reQuest，ARQ）是OSI模型中数据链路层的错误纠正协议之一。它包括停止等待ARQ协议和连续ARQ协议，错误侦测（Error Detection）、正面确认（Positive Acknowledgment）、逾时重传（Retransmission after Timeout）与负面确认继以重传（Negative Acknowledgment and Retransmission）等机制。

**★BPDU报文**

为了计算生成树，交换机之间需要交换相关的信息和参数，这些信息和参数被封装在BPDU（Bridge Protocol Data Unit）中。

BPDU报文被封装在802.3 MAC帧中（类型字段为LLC数据长度，含LLC子层及bpdu）

**目的MAC**地址是组播MAC： 01-80-C2-00-00-00

源MAC地址是交换机设备mac-address

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 字节 | 6 | 6 | 2 | 3 | 35 |
|  | 目的MAC | 源MAC | 类型 | LLC子层 | STP BPDU |
| 例值： | 0180C2000000 | xxxxxx | 0x0026 | 0x424203 | xxxxxx |

802.3和以太网 MAC帧区别在于类型字段，其值小于0x0600时，为802.3帧，类型字段值表示mac帧的数据长度，（数据为LLC子层及STP BPDU）

**★STP BPDU格式**

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 字节 | 2 | 1 | 1 | 1 | 8 | 4 | 8 | 2 | 2 | 2 | 2 | 2 |
|  | PID | PVI | BPDU Type | flags | Root ID | RPC | 桥ID | 端口ID | Message Age | Max Age | Hello Time | 转发延迟 |

①PID（Protocol ID）协议ID，2字节，（0x0000 IEEE 802.1D）

②PVI（Protocol Version ID）协议版本，1字节，（0x00 STP，0x02 RSTP，0x03 MSTP）

③BPDU Type，网桥协议数据单元类型，1字节，在标准STP（IEEE 802.1D）中有两种类型：

0x00为 配置BPDU，负责建立，维护STP拓扑;

0x80为 TCN BPDU，传达拓扑变更;

（1）CBPDU（Configuration BPDU）Type=0x00，表示配置BPDU，含Bridge ID、路径开销、Port ID等参数。STP协议通过在交换机之间传递配置BPDU来选举根桥，以及确定每个交换机端口的角色和状态。在初始化过程中，每个桥都主动发送CBPDU。在网络拓扑稳定以后，只有根桥主动发送CBPDU，其他交换机在收到上游传来的CBPDU后，才会发送自己的CBPDU。

（2）TCN BPDU，Type=0x80：下游交换机感知到拓扑发生变化时，向上游发送的拓扑变化通知（这是控制信息，从根端口转发），此时使用该类型报文。

④Flags标志位，1字节，在STP中仅使用最高位（TCA位Topology Change Acknowledgement）与最低位（TC位Topology Change）



★在CBPDU中，包含足够的信息来保证设备完成生成树计算，其中包含的重要信息如下：

⑤Root ID，8字节，由Root Bridge的优先级和MAC地址组成，每个STP网络中有且仅有一个根桥

⑥RPC（Root Path Cost）根路径开销，4字节，到根桥的最短路径开销

⑦Bridge ID，指定桥ID，8字节，由指定桥的优先级和MAC地址组成。指定桥：指定端口所在的桥（即BPDU的来源）

⑧Port ID，指定端口ID，2字节，由指定端口的优先级和端口号组成

⑨Message Age，2字节，CBPDU在网络中传播的生存期，每经过一个交换机 值+1，当达到20时，其他交换机不会再使用

⑩Max Age，2字节，CBPDU在设备中能够保存的最大时间（老化时间默认20s），在Max Age之内，如果未收到根桥的BPUD帧，则认为网络出现问题，进行重新选举根

⑪Hello Time，2字节，CBPDU发送的周期，默认每2s发送一次；

⑫Forward Delay，转发延迟，2字节，端口状态迁移的延时，默认15s

STP协议的核心是BPDU报文

BPDU报文：Root BridgeID | Cost | BridgeID| PortID

其初始值为：BridgeID | 0 | BridgeID | PortID

Root BridgeID：Root Bridge 的 BridgeID

Cost：本Bridge 到Root Bridge的距离，实际上，其是由逐步更新逐渐生成的。

BridgeID ：网桥ID，由Bridge Priority 和 Bridge MacAddress组成，其数值可以作为一个有效的定序尺度。

BridgeID = Bridge Priority<<48 + MacAddress

PortID：端口ID，其由Port Priority 和 Port Index组成，其数值可以作为一个有效的定序尺度。

BPDU报文可以比较大小，比较方法按元组（Root BridgeID ，Cost ，BridgeID，PortID ）的比较方式。

如果一个Bridge 有 N 个 Port, 那么其在同一时刻自己有N种BPDU报文，一个 Port 一种。

生成树步骤：

1.分别从自己的Port 出发，分享自己的这个Port的BPDU 到相连的LAN上

2.接收来自LAN上的BPDU，如果这个Port收到的BPDU比这个Port的BPDU好（值更小），那么更新本Port的BPDU。

3.循环1，2

输出：

找到本Bridge中使BDPU.Cost最小的那个Port，然后把这个Port打上标记 [Root Port ， R]

for 本Bridge的 each Port：

如果发现这个Port收到的BPDU中的Cost都比自己的这个Port的BPDU大，

那么把这个Port打上标记[Designated Port ,D]

TCN BPDU的结构与配置BPDU基本相同，但Payload部分只有Protocol Identifier（协议号）、Protocol Version Identifier（协议版本）和BPDU Type（BPDU类型）。类型字段是固定值0x80，长度只有4个字节。

TCN BPDU是指在下游拓扑发生变化时向上游发送拓扑变化通知，直到根节点。TCN BPDU在如下两种情况下会产生：

•端口状态变为Forwarding状态。

•指定端口收到TCN BPDU，复制TCN BPDU并发往根桥。

**★RSTP BPDU格式**

在

**★RIP路由协议**

RIP（Routing Information Protocol）是内部网关协议的一种，基于距离向量的路由选择协议

RIP中从一个路由器到非直连网络的距离定义为：所经过的路由器数加1

RIP协议的距离也称为跳数（hop count），一条路径最多只允许跳数为15，16则认为不可达

1988年6月 发布RIP v1（RFC1058）

1998年11月发布RIP v2（RFC 2453）

RIP协议数据封装在UDP报文中，目的端口号520，

RIP v2最大报文长度（4+20x25）=504字节

RIP v2报文：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 字节 |  |  |  |  |
| 4 | 命令 | 版本 | 0填充 | RIP首部 |
| 4 | 地址族标识符 | | 路由标记 | 路由信息 20字节 一个rip报文可带25个路由信息 |
| 4 | 网络地址 | | |
| 4 | 子网掩码 | | |
| 4 | 下一跳路由器地址 | | |
| 4 | 距离（1-16） | | |

RIP v2具有简单鉴别功能，使用鉴别功能时，将第一个路由信息块用作鉴别，地址族标识符置为全1（0xFFFF）路由标记写入鉴别类型，剩余16字节为鉴别数据，然后最多可有24个路由信息块

**★OSPF路由协议**

OSPF（Open Shortest Path First）开放最短路径优先，是一个内部网关协议，用在单一自治系统内决策路由，是对链路状态路由协议的一种实现，使用Dijkstra算法来计算最短路径树

OSPF v2用在ipv4网络，RFC 2328

OSPF v3用在ipv6网络，RFC 5340

1989年发布OSPF（RFC 2178）

1998年发布OSPF v2（RFC 2328）

OSPF中的链路代价cost/metric可取值1~65535

OSPF支持负载均衡（到同一目的网络有多条相同代价的路径时）

★OSPF网络类型

|  |  |  |  |
| --- | --- | --- | --- |
| 类型 |  | Hello时间 | 应用 |
| BMA | 广播多路访问型网络 | 10 s | 以太网，令牌环网，FDDI |
| NBMA | 非广播多路访问型网 | 30 s | 帧中继，X.25，SMDS |
| P2P | 点到点型 | 10 s | HDLC，PPP |
| P2MP | 点到多点型 | 30 s | 自动发现邻居，不选举DR/BDR |
| P2MP Nb | 点到多点非广播  p2mp Non-broadcast | 30 s | 手动配置邻居，不选举DR/BDR |

ospf报文在ip层之上，ip上层协议为89

★OSPF报文类型

|  |  |  |
| --- | --- | --- |
| Hello报文 | Hello |  |
| DD | Data Description |  |
| LSR | Link State Request | 链路状态请求报文 |
| LSU | Link State Update | 链路状态更新报文 |
| LSA | Link State Acknowledgment | 链路状态确认报文 |

LSA（Link-State Advertisement）链路状态公告报文，不是链路状态确认报文！！这个LSA以LSU更新报文形式发送

★LSA Type

|  |  |  |  |
| --- | --- | --- | --- |
| Type 1 | router lsa | O | 用于选举DR/BDR，所有路由器都发出，只在本区域内广播 |
| Type 2 | network lsa | O | 由本区域的DR/BDR发出，通告本区域汇总后的路由 |
| Type 3 | summary lsa | OIA | ABR发出，用于交换 区域与区域之间的路由信息 |
| Type 4 | asbr lsa | OIA | ABR汇总后发出，用于告诉区域内部路由器谁是ASBR |
| Type 5 | external lsa | OE1  OE2 | ASBR发出，把外部的路由信息引致OSPF自治系统内部 |
| Type 6 | MOSF | MOSF | 多播ospf |
| Type 7 | nssa lsa | ON1  ON2 | ASBR产生的关于NSSA的信息，NSSA中的ASBR产生的  5类不能有，于是产生了7类 |
| Type 8 |  |  | external attributes |
| Type 9 |  |  | Opaque LSA（Link-local scope） |
| Type 10 |  |  | Opaque LSA（Area-local scope） |
| Type 11 |  |  | Opaque LSA（AS scope） |

**★OSPF区域**

当网络规格大到一定程序时，LSA形成一定规模，给OSPF计算带来一定压力，为了缓解计算压力，OSPF采用分区域计算

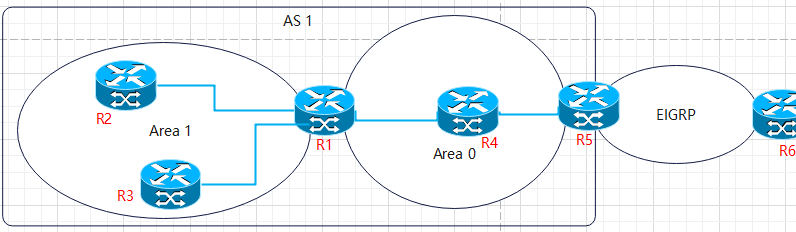
区域的命名可采用整数，如 0，1，2，3

区域的命名也可采用ip地址的形式，如 0.0.0.1，0.0.0.2

区域划分采用无环的Hub-Spoke拓扑架构，必须有一个核心区域 area 0

area 0称为BackBone骨干区域，其他为常规区域。所有常规区域应该和骨干区域相连，常规区域只能和骨干区域交换LSA，常规区域之间不能互换LSA，要通过area 0转发。

OSPF区域是基于路由器的接口划分的，一台路由器可以属于多个区域，也可只属于单一区域，不直接与area 0相连的区域要配置虚链路。根据区域的不同，路由器也可划分为不同的角色：



|  |  |  |
| --- | --- | --- |
| IR | （Internal Router）域内路由器，该类路由器所有接口都属于同一个区域 | R4 |
| ABR | （Area Borde Router）区域边界路由器，该类路由器的接口属于不同的区域，ABR用来连接骨干区域和非骨干区域，它与骨干区域之间既可以是物理连接，也可以是逻辑上的连接 | R1 |
| ASBR | （Autonomous System Boundary Router）自治系统边界路由器，位于OSPF自治系统和非OSPF网络之间。ASBR可以运行OSPF和另一路由协议（如RIP），把OSPF上的路由发布到其他路由协议上。ASBR必须处于非存根OSPF区域（NSSA）中 | R5 |

特殊区域：

|  |  |  |
| --- | --- | --- |
| Stub | 存根区域，末梢区域，不需要外部网络的信息，位于Stub边界的ABR将宣告一条默认路由到所有Stub区域内的路由器，不接受自治系统外部信息，也拒绝4类LSA | 不存在4，5，7  类LSA |
| Totally Stub | 完全存根区域，思科专用区域，不接受自治系统外部路由及自治系统内其他区域的路由汇总 | 不存在4，5，7 |
| NSSA | 非完全存根区域，可包含ASBR，不接受AS外部信息 | 不存在4，7 |
| Totally NSSA | 完全次末梢区域，不接受AS外路由，不接受AS内其他汇总路由 | 不存在4，7 |

\*骨干区域不能配为Stub区域，虚链接不能穿过Stub/NSSA 区域

\*OSPF路由在路由表中也以多种形式存在：

|  |  |  |
| --- | --- | --- |
| Intra Area Route | 同区域的路由 | O |
| Inter Area Route | 不同区域的路由(Summary Route) | OIA |
| External Route | 被重分布到OSPF中 | OE1 OE2 |

路由优先级： O > OIA > OE1 > OE2

一台路由器可以运行多个OSPF进程，不同进程的OSPF可以视为没有任何关系，要获得相互的路由信息，需要重分布（引入），路由器的链路状态数据库是分进程和分区域存放的

OSPF只有邻接状态才会交换LSA，OSPF邻居靠发送Hello报文来建立和维护，Hello报文会在启动了OSPF的接口上周期性发送，超过4倍Hello时间后，还没收到邻居的Hello报文则断开邻居关系

链路是路由器接口的另一种说法，OSPF也称为接口状态路由协议

OSPF通过路由器之间通告网络接口的状态来建立链路状态数据库，生成最短路径树，每个OSPF路由器使用这些最短路径构造路由表。

OSPF用于同一个路由域内，路由域指一个自治系统AS，它是指一组通过统一的路由政策或路由协议互相交换路由信息的网络，在这个AS中，所有的OSPF路由器都维护一个相同的描述这个AS结构的数据库，该数据库中存放的是路由域中相应链路的状态信息

OSPF将链路状态通过组播数据LSA传送给在某一区域内的所有路由器，LSA是封装在LSU报文中

OSPF使用接口带宽来计算cost/metric，只能为整数，带宽参考值默认为100M

cost等于 带宽参考值 除以 接口带宽

如：100Mbit/s接口的cost为1， 10M接口的cost为10，

cost值最小为1，所以1000M接口计算值为0.1，实际为1

到达目标相同的cost值的路径，可以做负载均衡

**★Router-ID**

路由器的身份id，不能有重名，一般为路由器上的一个ip地址，确定Router-ID的方法：

①手工指定rid

②路由器上Loopback接口中ip最大的ip

③路由器上物理接口的最大的ip（UP状态的）

如果一台路由器收到一条LSA，如果它无法到达lsa里的router-id的ip地址，则无法到达lsa里的目标网络

**★邻居**

形成OSPF邻居的2个路由器必须满足以下4个条件

①Area-ID区域号要相同

②Hello时间和Dead时间要分别一致

③认证相同，配置相同的认证密码

④Stub Area Flag末梢标签必须一致（处于相同的Stub区域内）

OSPF只能使用接口的primary地址建立邻居，邻居只发送Hello报文

**★邻接Adjacency**

2台OSPF路由器能形成邻居，但不一定能交换LSA，

只要能交换LSA则称为邻接（Adjacency）

**★DR/BDR**

Area区域中选出的一个核心路由器称为DR（Designated Router），备份的则为Backup DR

其他Other路由器只和DR/BDR互换LSA，

DR的选举：

优先级最高的（0~255，默认为1，0表示不参与选举）；Router-id最大的

Other与Otehr路由器之间可能是邻居关系，但不互换LSA

所有路由器与DR/BDR是邻接关系，互换LSA

Other使用源ip 224.0.0.6发送LSA给DR/BDR（的目的ip 224.0.0.5）

DR/BDR使用源ip 224.0.0.5发送LSA给Other（的目的ip 224.0.0.6）

**★Hello协议的目的**

①用于发现邻居

②在成为邻居前，必须对Hello报文里的一些参数进行协商

③Hello报文在邻居之间扮演着keepalive的角色

④允许邻居之间的双向通信

⑤用于在NBMA、BMA广播网络中选举DR/BDR

**★Hello报文包含的信息**

源路由器的Router-ID，Area-ID，接口掩码，接口认证类型和认证信息

源路由器的Hello报文发送的时间间隔，接口的无效时间间隔

优先级，DR/BDR接口ip地址，五个标记位

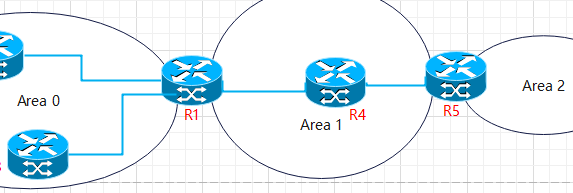
源路由器的所有邻居的Router-ID

★OSPF虚链路

虚链路是指在2台ABR之间，穿过一个非骨干区域（转换区域Transit Area）而建立的一条逻辑上的连接通道，可理解为2台ABR之间存在一个点对点的连接

这条逻辑通道只传送Type 3的LSA给骨干区域

虚连接的路由信息是作为域内路由来看待的，虚连接属于主干区域，通过虚链路连接的2台ABR可以交换路由信息



如上图，Area 2和Area 0之间隔着Area 1，要让Area 2和Area 0交换信息，则需要在Area 1上创建一条虚链路，在Area 2和Area 0分别与Area 1相连的ABR上配置Virtual-Link

因为普通区域之间不能直接交换LSA，只能通过Area 0，所以当Area 2不能与Area 0骨干区域直连时，需要通过一个中间区域Area 1去转换

**★BGP路由协议**

1989年发布BGP（Border Gateway Protocol）边界网关协议

各自治系统AS内运行不同的内部路由选择协议，且各AS指明的路径度量可能各不相同

如果互联网也用ospf则当一条路径通过几个不同AS时，计算出来的metric没有可比性

BGP只是寻找一条能够到达目的网络且比较好的路由，不一定是最佳路由。BGP使用路径向量（path vector）路由选择协议（as号的个数）报文封装在TCP中，目的端口号179

BGP v4 （RFC 4271）2006年1月发布

**★IGMP组播**

IGMP（Internet Group Management Protocol）互联网组管理协议，是TCP/IP协议族中负责IP组播成员管理的协议，用来在IP主机和与其直接相邻的组播路由器之间 建立、维护组播成员关系

IP组播要关注的问题：

|  |  |
| --- | --- |
| 组播寻址机制 | 组播源将组播信息传输到哪里 |
| 主机注册 | 网络中有哪些接收者 |
| 组播源发现 | 这些接收者需要从哪个组播源接收信息 |
| 组播路由 | 组播信息如何传输 |

IGMP有3个版本，所有的版本都支持ASM（任意信源组播）模型，IGMP v3可直接应用于SSM指定信源组播模型

**组播模型：**

ASM（Ayn Source Multicast）任意信源组播

SFM（Source Filtered Multicast）信源过滤组播

SSM（Source Specific multicast）指定信源组播

**★IGMP v1**

主要基于查询和响应机制来完成对组播组成员的管理

当一个网段内有多台组播路由器时，需要有一个查询器Querier选举机制，来确定由中始路由器作为IGMP查询器。

对于IGMP v1来说由组播路由协议（如PIM）选出的DR作为IGMP查询器

①主机会主动向其要加入的组播组发送IGMP成员关系报告报文以声明加入，而不必等IGMP查询器发来的IGMP查询报文

②IGMP查询器周期性地以组播方式（224.0.0.1）向本网段内所有主机与路由器发送IGMP查询报文

③主机们收到查询报文后，有一个抢先发了报告报文，宣告其属于G1，其他也想加入这个组播的主机看到有人发了报告报文后，自己就不发了

④重复1，2，3步骤后，IGMP路由器了解到本地网段中有G1,G2 的成员，于是由组播路由协议（PIM）生成（\*,G1）和（\*,G2）组播转发项作为组播数据的转发依据

⑤由组播源发往G1,G2的组播数据，经过组播路由到达IGMP员單，由于IGMP路由器上存在（\*,G1）和（\*,G2）转发项，于是将该组播数据转发到本地网段，主机们就能收到该组播数据了

\*IGMP v1没有定义离开组播组的报文，所以IGMP路由器在一段时间内没有收到发往某组播组的报告报文时，便删除该组播组所对应的组播转发项

**★IGMP v2**

与IGMP v1相比，IGMP v2增加了独立的查询选举机制和离开组机制

**★IGMP v3**

在v1 v2的基础上增强了主机的控制能力，增加查询报告报文的功能

主机在加入某个组播时，能明确要求接收或拒绝来自某特定组播源S的组播信息

IGMP v3报告报文目的地址为 224.0.0.22

v3不仅支持v1的普遍组查询和v2的特定组查询，还增加了 特定源组查询

|  |  |
| --- | --- |
| 普遍组查询报文 | 不携带组地址，不携带源地址 |
| 特定组查询报文 | 携带组地址，不携带源地址 |
| 特定源组查询报文 | 携带组地址，携带一个或多个源地址 |

★IGMP Snooping

交换机对流经的IGMP报文进行分析，为端口和组播mac建立映射关系，从而减少组播资源的消耗（如果交换机不支持IGMP Snooping，则对所有组播报文进行广播发送，开启IGMP Snooping后只发送到有组播成员的端口）

当交换机收到DR查询器发来的查询报文时，向同一vlan的其他端口发送，当收到某个/些 端口回复的membership report报文后，就把映射关系加入mac地址表中

当交换机收到leave group报文或动态端口的定时器超时后，删除相应的mac地址表项

**★PIM组播**

PIM（Protocol Independent Multicast）独立组播协议，协议无关组播

**章n、周边知识**

**★PCIE接口**

PCI-Express（peripheral component interconnect express）外围组件接口互联

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| PCIE | 传输速率 | 吞吐量 | | | |
| x1 | x4 | x8 | x16 |
| 1.0 | 2.5GT/s | 250MB/s | 1GB/s | 2GB/s | 4GB/s |
| 2.0 | 5GT/s | 500MB/s | 2GB/s | 4GB/s | 8GB/s |
| 3.0 | 8GT/s | 984.6MB/s | 3.938GB/s | 7.877GB/s | 15.754GB/s |
| 4.0 | 16GT/s | 1.969GB/s | 7.877GB/s | 15.754GB/s | 31.508GB/s |
| 5.0 | 32GT/s |  |  |  |  |

GT表示giga Transmission每秒传多少次，一次多少数据是由通道个数决定的

8B10B编码，250MB x 10/8 == 2.5 GT/s x 1通道

**★SATA接口**

|  |  |  |
| --- | --- | --- |
| SATA 1.0 | 1.5Gbps | 150MB/s |
| SATA 2.0 | 3Gbps | 300MB/s |
| SATA 3.0 | 6Gbps | 600MB/s |

★M.2硬盘