DNS

根域名服务器知道所有的顶级域名服务器的域名和IP地址，世界上共有13个不同IP地址的根域名服务器，每个IP地址有多台完全相同的服务器分布在世界各地，对任何一个域名解析，只要本地域名服务器无法解析，就首先求助根域名服务器，就近选择根域名服务器。

顶级域名服务器

com net org cn us uk asia biz jobs ...

二级域名服务器

com.cn, ac.cn, edu.cn

三级域名

sust.edu.cn mail.cctv.cn

主机

[www.sust.edu.cn](http://www.sust.edu.cn) xy.sust.edu.cn library.sust.edu.cn fcwr.jstv.com.cn

一个单位拥有了一个域名，可以自己决定是否要进一步划分其下属的子域，并且不必由向其上级机构批准。域名树的树叶就是单位计算机的名字。

每一级的域名都有一个相对应的域名服务器。



权限域名服务器

abc公司买下了abc.com这个二级域名，该公司把[www.abc.com](http://www.abc.com)作为公司主页域名，[www.abc.com](http://www.abc.com)

不是域名，是指abc.com域名下的www这台服务器。mail.abc.com 作为公司的邮件服务器。

market.abc.com作为销售服务器，后来市场火爆，尤其是中国市场。market.abc.com服务器受不了，公司就把market.abc.com改成域名，主机有us.market.abc.com,jp.market.abc.com，europe.market.abc.com等,由于中国市场太过火爆，把china.market.abc.com弄成域名，设置henan.china.market.abc.com, beijing.china.market.abc.com...为服务器名。注意market.abc.com和china.market.abc.com都是域名，都有独立的域名服务器，区别是market.abc.com是三级域名，china.market.abc.com是四级域名，market.abc.com域名服务器只能解析前缀是主机名如[www.market.abc.com的域名，碰到henan.china.market.abc.com会返回china.market.abc.com域名服务器的IP](http://www.market.abc.com的域名，碰到henan.china.market.abc.com会返回china.market.abc.com域名服务器的IP)地址。

以上就是权限服务器。

本地域名服务器

假设域名为[www.baidu.com的主机想知道beijing.china.market.abc.com的IP](http://www.baidu.com的主机想知道beijing.china.market.abc.com的IP)地址，过程：

1. [www.baidu.com主机向本地根域名服务器dns.baidu.com](http://www.baidu.com主机向本地根域名服务器dns.baidu.com)查询，缓存命中，就返回beijing.china.market.abc.com的IP地址，并标记成非权威服务器的应答。china.market.abc.com域名服务器或者它的父域名服务器回复的IP地址才是权威的，其他域名服务器（本地域名服务器）回复的都是非权威服务器的应答。
2. 缓存没有命中，本地域名服务器就近访问根域名服务器，根域名服务器回复dns.baidu.com .com域名服务器的IP地址，dns.baidu.com又向.com域名服务器请求，被回复abc.com的IP地址，...被回复market.abc.com，china.market.abc.com的IP地址，最后被china.market.abc.com域名服务器直接回复了beijing.china.market.abc.com的IP地址，本地服务器dns.baidu.com把IP地址反馈给主机[www.baidu.com](http://www.baidu.com)。

HTTP

URL

Uniform Resource Loacator URL 统一资源定位符

格式: 协议 – 主机域名/主机IP - 端口 – 文件路径 – 文件名

例如: http://www.sust.edu.cn/zzjg/xzbm.htm

陕西科技大学 - 组织机构 - 行政部门 的页面

协议是http,主机域名[www.sust.edu.cn,路径/zzjg/](http://www.sust.edu.cn,路径/zzjg/) ,文件名xzbm.htm,http默认端口80可省略

URL使得web上的所有资源都有了唯一的标识符,并且标识符蕴藏着资源的位置.

网页与HTML

HTML文档可以包含本地资源(直接插入文档显示),本地资源的本地链接(解析文档去链接处找到资源显示它),或者位于其他主机的资源链接(比如www.sust.edu.cn/zzjg/xzbm.htm,实现网页互连,跳转).

网页显示的是HTML文件,网页与网页互连,形成了万维网!

HTTP协议

HTTP协议定义了浏览器如何向万维网服务器请求万维网文档,以及服务器如何把万维网文档传送给浏览器.

浏览器使用HTTP1.0协议访问http://www.sust.edu.cn/zzjg/xzbm.htm的过程:

1. 浏览器向域名服务器获取www.sust.edu.cn的IP地址.
2. 浏览器与服务器进程建立TCP连接.
3. 浏览器发出取文件命令:GET /zzjg/xzbm.htm.
4. 服务器把xzbm.htm.发送给浏览器.
5. 关闭TCP连接.

HTTP1.0和HTTP1.1区别

假设客户端请求的HTML文件包含文本和2个图片以及3个图片的链接:

(1)HTTP1.0非持久性连接,传送一个对象就断开TCP.

* 建立TCP连接 RTT
* 接收HTML文档 RTT
* 断开TCP连接 2RTT
* 解析HTML文档,显示文本和2个图片,再次建立TCP连接, 根据图片链接传输图片,操作3次

耗时(RTT+RTT+2RTT) \* 4 = 16RTT

(2)HTTP1.1持久性连接,保持TCP连接,后续的HTTP消息可通过同一个TCP传输.

* 建立TCP连接 RTT
* 接收HTML文档 RTT
* 解析HTML文档,流水机制(客户端未对上一对象确认收到就可请求下一对象)请求3个图片RTT
* 关闭TCP连接2RTT

耗时RTT+RTT+RTT+2RTT = 5RTT

HTTP消息格式

HTTP请求消息Request

GET /zzjg/xzbm.htm HTTP/1.1 请求行 GET POST PUT HEAD... 文件路径文件名 版本

Host: [www.sust.edu.cn](http://www.sust.edu.cn) 头部行 主机名,使用代理服务器/缓存时用到

User-agent: Opera/5.0 头部行 浏览器版本,服务器找到网页的相应的最佳版本

Accept-language: fr 头部行 网页显示语言

消息体...... (请求消息一般没有,需要时有)

请求消息包含给服务器的信息

GET 方法: 信息比较少时,将信息放在请求行文件名后面, GET/zzjg/xzbm.htm/animal?monkeys HTTP/1.1.

HEAD:请sever不要把请求的对象放入消息体,测试时用.

POST:网页经常填写表格,将信息放在消息体

PUT:将消息体中的文件上传到URL指定的路径

DELETE:删除URL指定的文件

HTTP响应消息response

HTTP/1.1 200 OK 状态行

Date:Thu,06 Aug 1998 12:00:15 GMT 响应消息生成的时间

Server:Apache/1.3.0(Unix) 服务器软件和系统

Last-Modified: Mon,12 Aug 1997 15:20:47 GMT 网页最后修改的时间

Content-Type:text/html 文本类型

data ... data ... data ..........

HTTP相应状态代码

200 OK

304 Not Modified

400 Bad Request

404 Not Found

文件传输协议FTP

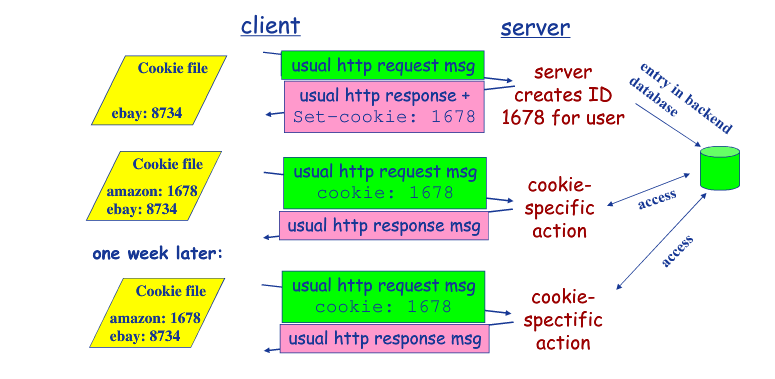
Email应用

Cookie技术

某些网站为了辨别用户身份而存在用户本地终端上的数据.

Cookie组件:HTTP请求消息的Cookie头部行,HTTP响应消息的Cookie头部行

保存在用户本地的Cookie文件,由浏览器管理.Web服务器的后台数据库.



解释上图:

客户端本地Cookie文件已有eBay交互ID8734,客户端尚未与Amazon交互过.

客户端发送HTTP消息,服务器检测到没有Cookie头部,便为用户创建ID存入自己的数据库,并把ID返回给用户写入到自己的本地Cookie文件,以后每次交互,客户端都发送带Cookie头部的HTTP消息.

SMTP协议

邮件系统由邮件客户端和邮件服务器组成,所有的用户都是客户端,邮件服务器是核心.

Simple Mail Transfer Protocol 简单邮件传输协议,是邮件服务器之间传递消息使用的协议.

Tom要给Jack发邮件?

1. Tom申请邮箱tom@qq.com,qq.com邮件服务器开辟空间[received mail],[need send mail]

前者用于存储别人发给自己的邮件,后者用于暂存自己要发给别人的文件.

1. Jack申请邮箱jack@163.com,同上.
2. Tom发送邮件存到@qq.com服务器上自己的[need send mail]空间,然后由@qq.com和@163.com交互,最终转发到@163.com邮箱服务器的Jack的[received mail]内.当某个时间点Jack打开自己的电脑邮件客户端,163邮件服务器检测到Jack的[received mail]内有邮件,就会提醒Jack,Jack从服务器接收邮件到自己电脑的邮件客户端.

SMTP邮件服务器交互过程:

qq.com : HELO qq.com

163.com : 250 hello,qq.com,Nice to meet you.//也许会163公司会屏蔽qq公司的邮箱用户

qq.com : MAIL FROM : <tom.qq.com>

163.com : 250 tom.qq.com ...... Sender is OK. //屏蔽垃圾邮件

qq.com : RCPT TO : <jack.163.com>

163.com : 250 jack.163.com.....Recipient is OK.//可能邮件服务器没有此用户,没注册.

qq.com : DATA //告诉163我开始发数据了

163.com : 354 Enter Email,end with "." on a line by itself.//以"."作为数据传完标志

qq.com : Dear Jack.

qq.com : I Send Some AV From Janpan,

qq.com : Wish You like.

qq.com : .

163.com : 3 Message accepted for delivery.// 3条消息已经全部接收

POP3是Post Office Protocol 3的简称，即邮局协议的第3个版本，是TCP/IP协议族中的一员（默认端口是110）。本协议主要用于支持使用客户端远程管理在服务器上的电子邮件。POP3协议允许电子邮件客户端下载服务器上的邮件，但是在客户端的操作（如移动邮件、标记已读等），不会反馈到服务器上，比如通过客户端收取了邮箱中的3封邮件并移动到其他文件夹，邮箱服务器上的这些邮件是没有同时被移动的 。

IMAPIMAP全称是Internet Mail Access Protocol，即交互式邮件访问协议，是一个应用层协议（端口是143）。用来从本地邮件客户端（Outlook Express、Foxmail、Mozilla Thunderbird等）访问远程服务器上的邮件。 IMAP提供Webmail 与电子邮件客户端之间的双向通信，客户端收取的邮件仍然保留在服务器上，同时在客户端上的操作都会反馈到服务器上（如：删除邮件，标记已读等，服务器上的邮件也会做相应的动作。所以无论从浏览器登录邮箱或者客户端软件登录邮箱，看到的邮件以及状态都是一致的。）。IMAP更好地支持了从多个不同设备中随时访问新邮件。所有通过IMAP传输的数据都会被加密，从而保证通信的安全性。

FTP（File Transfer Protocol，文件传输协议） 是 TCP/IP 协议组中的协议之一。FTP协议包括两个组成部分，其一为FTP服务器，其二为FTP客户端。其中FTP服务器用来存储文件，用户可以使用FTP客户端通过FTP协议访问位于FTP服务器上的资源。在开发网站的时候，通常利用FTP协议把网页或程序传到Web服务器上。此外，由于FTP传输效率非常高，在网络上传输大的文件时，一般也采用该协议。默认情况下FTP协议使用TCP端口中的 20和21这两个端口，其中20用于传输数据，21用于传输控制信息。

TCP对应的协议：FTP，21 Telnet，23，远程登录协议，用户可以以自己的身份远程连接到计算机上，可提供DOS环境下的通信服务。

SMTP，25 POP3，110 HTTP，80

UDP对应的协议：DNS，53，域名解析协议 TFTP，69，简单文件传输协议。

网络层协议：IP，ICMP，ARP,RARP

运输层协议：TCP，UDP

应用层协议：HTTP,Telent，FTP，TFTP,SMTP,POP3,DNS

代理服务器缓存技术

Web缓存代理服务器技术:在不访问服务器的前提下满足客户端的HTTP请求.

为什么使用这项技术?

1. 缩短客户请求的响应时间
2. 减少组织机构的流量
3. 翻墙

代理服务器工作细节?

1. 客户端设置浏览器通过缓存进行Web访问
2. 客户端浏览器向代理服务器发送所有的HTTP请求.
3. 请求的对象在缓存中命中,代理服务器把对象发给客户端,否则,代理服务器向原始服务器请求对象,然后把接收的对象返回给客户端,并自己保存一份.

代理服务器一般由ISP架设.

工作原理?

某机构的局域网内部网速是10Mbps,路由器连接因特网的网速是1.5Mbps,从机构路由器到互联网的往返迟延是2s.

该机构向原始服务器平均每秒请求15个对象,每个对象平均大小是100,000bit,请分析网络性能.

局域网的利用率 = 1.5 / 10 = 15% 时延几微秒

接入互联网的链路利用率 = 1.5 / 1.5 = 100% 时延几分钟,甚至不确定

总时延 = 2s + 几分钟 + 几微秒

改善方案一:提升接入互联网的带宽到10Mbps, (成本太高).总时延2s+ 几微秒 + 几微秒

改善方案二:安装web缓存,假设命中率是0.4(一般介于0.2 – 0.7)

总时延 = 0.6 \* 2.01s + 0.4 \* 几微秒 < 1.5s

低成本却达到更佳的访问速度.

代理服务器引来的新问题:缓存的网页不是最新版本怎么办?

客户端发送到代理服务器的每个HTTP消息,代理服务器都会从中取出要访问的文件,如果命中,把该文件的版本时间发给原始服务器,如果原始服务器在这个时间后都没更改内容,就回复响应消息,状态码304 Not Modified,消息体中并没有请求的文件;如果更改了,就重传一份新的.

TCP的三次握手

TCP三次握手过程

服务器处于LISTEN状态,等待客户端连接.

客户端发送SYN请求连接包(SYN = 1,seq = m,进入SYN-SENT状态)

服务器收到SYN包,同意连接.回复ACK-SYN包(SYN = 1,ACK = 1,ack = m + 1,seq = n,进入SYN-RCVD状态,第一次握手完成).

客户端收到ACK包,检查其ack,若ack = m + 1,表示服务器同意连接,回复ACK包(ACK = 1,seq = m + 1,ack = n + 1,进入ESTABLISHED状态,第二次握手完成).

服务器收到ACK包,检查其ack = n + 1, 进入ESTABLISHED状态,第三次握手完成.

TCP为什么握手三次,而不是两次或四次？

假设一种情况,客户端发送了第一个SYN包,但是由于第一个SYN包网络延时,没按时到达服务器,客户端没及时收到确认,会再发出第二个SYN包,如果采用了两次握手机制,服务器一旦收到SYN包便进入ESTABLISHED状态,服务器收到第二个SYN包,成功建立了一个TCP连接,此时,延迟的SYN包到达了服务器,服务器又建立了一个连接并回复了一个无效的确认包,客户端收到无效的确认包直接丢弃.那么,延迟的SYN包使得仅有服务器进入ESTABLISHED状态,但永远不会有数据传输,造成了资源浪费.

TCP的四次分手

A发送FIN包(FIN = 1,seq = u = 上次传输的最大的字节序号+1,进入FIN-WAIT-1状态)

B收到FIN包,回复ACK包(ACK = 1,ack = u + 1,进入CLOSE-WAIT状态)

A收到ACK包,进入FIN-WAIT-2状态

此时TCP连接处于半关闭状态,B若还有数据发给A,A仍可接收.

若B没有数据要发给A,发送FIN包(FIN = 1,seq = v = 上次传输的最大的字节序号+1,进入LAST-ACK状态)

A发送ACK包(ACK = 1,ack = v + 1,进入TIME-WAIT状态)

B收到ACK包后关闭连接

A等待两个最长报文段寿命时间后,关闭

发现:A先提出的关闭,但是最终B先关闭的.第三次握手后,A没直接关闭,而是等待两个最长报文段寿命才关闭,原因如下:(1)如果第三次握手发送的ACK包失效,B无法关闭.(2)两个最长报文段寿命内,能使当前即将关闭的连接运输的报文段全部从网络中消失,不会造成本次的报文段出现在A的下一次连接中.

TCP实现可靠传输的机制

实现可靠传输需要保证三个条件:  
(1)无比特差错传输

(2)字节流不丢不重不乱序

(3)接收方的处理能力大于发送速率

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

TCP的首部检验和会检查首部和数据段,保证报文段无比特差错,满足条件(1).

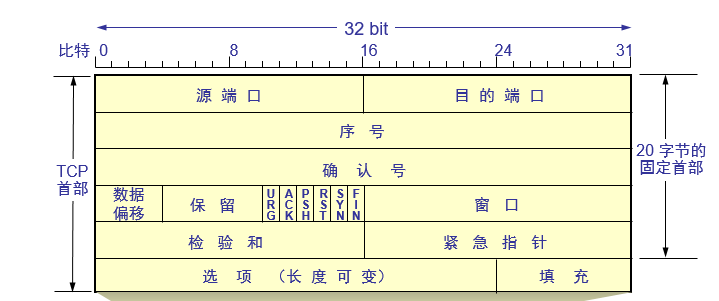
应用程序把数据缓存到发送缓存,接收方TCP把数据放到接收缓存.发送方根据接收方反馈回来的窗口设置发送窗口,窗口是缓存的前片段,缓存内窗口外的字节不允许发送,实现了流量控制,满足条件(3).

窗口内的字节分成两部分:已发送还未确认 & 可发送还未发送,接收缓存的存储的是:按序到达还未提取,未按序到达的(4没到,5到了),接收方按序发送确认包,比如4的确认包没发送,就不会发送5的确认包.接收方收到确认包,就丢弃窗口中对应的字节,窗口往前滑动,新增字节.确认机制满足条件(3).

所以TCP能实现无差错传输.

|  |  |  |  |
| --- | --- | --- | --- |
| 发送缓存 | | 接收缓存 | |
| 已发送未确认 | 可发送未发送 | 未提取的正确字节 | 未按序到达的字节 |

TCP首部



**SourcePort(源端口)和DestinationPort（目的端口）**：各占2个字节，端口是运输层与应用层的服务接口,运输层的复用和分用功能都要通过端口才能实现。

   注：TCP的包是没有IP地址的，那是IP层上的事。但是有源端口和目标端口。  
**SequenceNumber（序号）**：占4个字节，TCP连接中传送的数据流中的每一个字节都编上一个序号。序号字段的值则指的是本报文段所发送的数据的第一个字节的序号。  
**AcknowledgmentNumber（确认号）**：占4字节，是期望收到对方的下一个报文段的数据的第一个字节的序号。  
**DataOffset（数据偏移（即首部长度））**：占4位，它指出TCP报文段的数据起始处距离TCP报文段的起始处有多远。“数据偏移”的单位是32位字，即以4字节为计算单位。   
**Reserved(保留字段)**:占6位，保留为今后使用，但目前应置为0。  
**TCPFlag(标志位)**:占6位，分别表示6个标志：**紧急URG**，**确认ACK**，**推送PSH(PuSH)**，**复位RST(ReSeT)**，**同步SYN**，**终止FIN(FINis)**：  
URG—— 当URG= 1 时，表明紧急指针字段有效。它告诉系统此报文段中有紧急数据，应尽快传送(相当于高优先级的数据)。  
**ACK**—— 当ACK= 1 时确认号字段有效,当ACK= 0 时，确认号无效。  
**PSH(PuSH)** ——接收TCP收到PSH= 1 的报文段，就尽快地交付接收应用进程，而不再等到整个缓存都填满了后再向上交付。  
**RST(ReSeT)**—— 当RST= 1 时，表明TCP连接需要释放连接，然后再重新建立运输连接。  
**SYN**—— 同步SYN= 1 表示这是一个连接请求或连接接受报文。  
**FIN(FINis)**——用来释放一个连接。FIN= 1 表明此报文段的发送端的数据已发送完毕，并要求释放运输连接。

**Window（窗口字段）**：占2字节，用来让对方设置发送窗口的依据，单位为字节。表示接收缓冲区的空闲空间，16位，用来告诉TCP连接对端自己能够接收的最大数据长度。  
**Checksum（检验和）**：占2字节。检验和字段检验的范围包括首部和数据这两部分。在计算检验和时，要在TCP报文段的前面加上12字节的伪首部。  
**UrgentPointers（紧急指针字段）**：占16位，指出在本报文段中紧急数据共有多少个字节（紧急数据放在本报文段数据的最前面）。  
**options（选项字段）**：长度可变。TCP最初只规定了一种选项，即最大报文段长度MSS。MSS告诉对方TCP：“我的缓存所能接收的报文段的数据字段的最大长度是MSS个字节。”在TCP的MSS选项中，MSS值是一个16位的字段，最大值为65535.

MSS是TCP报文段数据部分的最长长度,不是总长度.MSS太小,首部20字节,有效数据1个字节,网络效率太低.MSS太大,IP要分层,浪费时间.

TCP中的计时器

TCP中的计时器?

(1)重传计时器

TCP发送完一个报文段,就设置一个专属于此报文段的计时器,规定时间内收到此报文段的确认,撤销计时器,时间走完还没收到确认包,重传此报文段并重置计时器.

(2)持续计时器

客户端收到的确认包窗口是0,便停止发送数据了.过了一会,接收端缓过来劲了,继续发送一个更高序号的字节的确认包,它的窗口大于0,客户端如果收到此确认包,检测到窗口大于0,就会重新发送数据.但是如果此”激活”确认包万一丢失,双方都会永久静默下去(TCP不会重传ACK确认包).所以为每个TCP连接设置一个计时器,每隔一段时间,探测一下是不是丢包了.

(3)保活计时器

数据传输中途,客户端崩溃,防止服务器傻傻等待而浪费资源,会每隔一段时间,检查一下客户端是不是崩溃了.

1. 时间等待计时器,TCP四次分手的时候使用的.

比较TCP和UDP

UDP首部

UDP首部固定8个字节   源端口,目的端口,数据报总长度,检验和,各占2个字节.

TCP和UDP区别

|  |  |
| --- | --- |
| TCP | UDP |
| 面向连接 | 无连接 |
| 可靠的全双工通信 | 尽力而为 |
| 拥塞控制,流量控制,确认机制 | 有就发,发出去死活不管 |
| 首部最少20字节 | 首部固定8字节 |
| 面向字节流 | 面向报文 |
| 一对一 | 任意量对任意量 |

IP地址

32位，网络号+主机号

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 网络类别 | 特征 | 最大网络数 | 第一个可指派的网络号 | 每个网络中的最大主机数 |
| A类 | 0，前一个字节 | 2^7 - 2 | 1.0.0.0 | 2^24 - 2 |
| B类 | 10，前二个字节 | 2^14 - 1 | 128.1.0.0 | 2^16 - 2 |
| C类 | 110，前三个字节 | 2^21 - 1 | 192.0.1.0 | 2^8 - 2 |

最大网络数-1：除特定位，剩余位全0的网络号保留。A类地址-2原因：0111\_1111（127）是环回地址：本主机的进程间通信之用，主机发送一个目的地址是环回地址的IP数据报，不会ARP解析，而是直接处理。

每个网络中的主机数-2的原因：主机号全0，表示主机所在的单个网络地址，主机号全1表示该网络上的所有主机。

ARP地址解析协议

把网络层的IP地址解析成数据链路层的硬件地址。

每一个主机都有一个ARP缓存，里面有本局域网上所有主机和路由器的IP地址-硬件地址映射表。当主机A要向本局域网上的主机B发送IP数据报时，就先在自己的ARP缓存中，查找主机B的IP地址对应的硬件地址，若找到，发送数据帧，若没有找到，主机A向本局域网上的所有主机广播发送：我的IP地址是…我的硬件地址是…我要找IP地址是…的主机的硬件地址。目的主机收到广播消息，恢复自己的硬件地址，其他主机不理睬广播消息。

ICMP网际控制报文协议

当IP数据报传送异常时，主机或路由器向源点报告异常情况。

ICMP差错报告报文 = 把异常IP数据报的首部和前8个字节提取出来作为ICMP报文的数据字段 + 异常原因。

最后把ICMP差错报告报文+IP首部封装成IP数据报发送给源点。

异常情况有：（1）终点不可达（目的主机崩溃）（2）源点抑制（网络拥塞）（3）时间超过（规定时间内一个大数据报未收全）等。

DHCP协议

动态主机配置协议。用途：给内部网络自动分配IP地址和配置参数，基于UDP。

在浏览器输入[www.baidu.com](http://www.baidu.com)后执行的全部过程

1. 浏览器得到DNS服务器解析到的IP地址
2. 浏览器与服务器建立TCP连接，默认端口是80
3. 浏览器发出HTTP请求，请求百度首页
4. 服务器通过HTTP响应把首页文件发送给浏览器
5. 释放TCP连接
6. 浏览器显示首页HTML文件

涉及到的协议：

应用层：HTTP DNS

传输层:TCP (HTTP) UDP (DNS)

网络层：IP，ICMP，ARP

运输层习题

运输层为应用程序提供端到端的逻辑通信,网络层负责主机寻址和分组交换.

应用进程之间的通信是可靠还是尽力而为,由运输层决定并加载到网络层.

什么是复用和分用?

应用层的数据通过不同端口加载到同一个TCP/IP连接传输,叫复用.

运输层把从IP层拿到的数据,通过正确的交给不同的端口,叫分用.

UDP/TCP和实时音频

实时场景,宜采用UDP,UDP产生数据就直接发送,不考虑网络拥塞,接收端流量,错误丢弃不重传,所以数据传输速度快,实时性好.虽然有可能数据报错误,但是人眼耳有一定的承受能力,如果采用TCP可靠传输,虽然没有数据报错误,但是重传,流量控制拥塞控制,会造成时延,综合考虑,实时性重于可靠,故选用UDP.

如果采用TCP观看收听音频,可以回看.

为什么说UDP是面向数据报的,TCP是面向字节流的?

UDP对应用层数据不拆不并,只保留边界交个IP层.TCP视应用层数据是无结构的字节流,根据需要拆分合并,但是不打乱字节流的顺序.

什么是端口?

端口16bit,来标记应用进程.各种服务器使用的端口号都是保留端口号,以便客户端能够找到服务器.

什么是套接字?

套接字 = {IP地址 : 端口号},套接字由应用程序程序产生,并指明它将由客户端还是服务器使用.

是否TCP和UDP都需要计算往返时间RTT?

TCP发送方的报文需要确认,所以要计算RTT,来设置合适的重传时间.UDP发出去是死是活就不管了,不需要计算RTT.

TCP在建立新的连接时所选择的初始序号,一定要和前面的一些连接所使用的序号不同,防止后效作用.

为什么ACK包不消耗序号,SYN包消耗序号?

TCP给传输的每个字节编号,目的是为了保证每一个字节能正确到达对方.接收方按字节的编号对收到的数据进行核对,用发送确认报文段的方法保证能够收到发送方的每一个字节.

确认报丢失不会重传,因为没必要重传,假设对1000字节的确认报丢失,可能会造成发送方多发送一个900-1000的字节段,但是对于多收的字节,接收方会丢弃.等对1100字节的确认包到达发送方,那有没有对1000字节的确认包无所谓了,因为收到对1100字节的确认包,已经意味着1100字节前的所有字节都已正确收到.

所以对ACK包没必要编号,ACK包不消耗序号,但是ACK包如果携带数据,会消耗序号.

SYN包只有首部,数据部分是空的,但是SYN包非常重要,没有它无法建立连接,且SYN包的序列号用来协商初始序列号,所以有必要也必须对SYN包编号.

为什么传输大文件用TCP而不是TCP?

因为一个大文件稍有错误,就无法打开,那不是白搭功夫了.

拥塞控制:对资源的需求>可用资源,硬件设施不行.

解决方法是控制端口注入网络的数据或升级网络硬件设施.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 5-13 | 5-14 | 5-22 | 5-23 |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |

IP数据报的格式

|  |  |  |  |
| --- | --- | --- | --- |
| 版本 | 首部长度 | 总长度 | 可选字段 |
| 标识 | 标志 | 片偏移 | 源地址 |
| 生存时间 | 协议 | 首部检验和 | 目的地址 |

版本:IPV4 IPV6

首部长度:IPV最少20字节,IPV6的首部长度是固定的.

总长度:IP数据报首部和数据之和的长度. 16个bit,所以IP数据报的最大长度是2^16个字节.IP数据报会在数据链路层封装成数据帧,数据帧的数据部分是整个IP数据报,有长度限制,即最大传送单元(MTU),<=1500个字节.所以,太长的数据报下放到数据链路层时,会被分片.

标识,标志,片偏移:一个原始数据报被分成若干子数据报.每个子数据报有自己的首部,每个子数据报的数据部分是原始数据报数据的一个片段.每个子数据报的标识位与原始数据报的标识位相同,可辨认出哪些子数据报来自同一个大数据报,片偏移是子数据报在原始数据报的片段的起始字节(从0开始),以便有序组装成原始数据报.

标志MF=1,不是一个子数据报,MF=0是最后一个子数据报.

首部检验和:只检验IP数据报的首部无比特差错.不检验数据部分,所以IP层是不可靠传输.

生存时间:占8bit,指明此数据报最多经过几个路由器,每经过一个路由器,值减1,减至0就从网络中丢弃此IP数据报.把生存时间设置成1,表示此数据报只能在本地局域网里传输.

协议:运输层协议

源地址,目的地址:两台通信主机的IP地址.

分类地址,向一个组织分配IP地址,就只能以8,16,24为单位来分配,很不灵活.

某ISP承包4096个IP地址,也就是16个C类地址,如果采用分类地址分配,要分配给该ISP 16个C类地址,ISP的路由表项有16个,但是采用无分类地址(CIDR),只需分配给该ISP 206.0.176.0/12的地址块,只需一个路由表项就能找到该ISP,节省了路由表的存储空间.某大学向该ISP申请800个IP地址,ISP可以把206.0.184.0/10的地址块分配给该大学,共1024个IP地址.该大学还可以给各个学院分配合适的地址块.这就是地址聚合的概念.

假定大学的4系,希望转发给自己的信息不经过大学的路由器,而是直接从ISP的路由器到自己的路由器,可以在ISP的路由器表项中增添自己的地址块.当一个IP地址到达ISP的路由器,会匹配成功两个地址块(学校的网址和4系的网址),选择较长的地址块转发.

如何在1万个不重复整数中查找某个整数.

(1)把1万个整数表示成二进制

|  |  |  |  |
| --- | --- | --- | --- |
| 208 | 128 | 15 | 240 |
| 1101 0000 | 0100 0000 | 0000 1111 | 1111 0000 |

(2)逐个扫描所有整数的第一个bit,第二个bit.... 建立二叉树,详见谢希仁145页.

(3)从树根的左孩子或右孩子开始往叶子结点遍历,相当于二叉查找树.

(4)查找,增加,删除一个数都只需要比较8次(char),32次(long).

ICMP网际控制报文协议

把收到的需要进行差错报告的IP数据报的首部和数据字段的前8个字节(源端口,目的端口,确认号)提取出来,作为ICMP报文的数据字段,再加上相应的ICMP差错报告报文的前8个字节(类型1,代码1,检验和2,阐述类型4)构成了ICMP差错报文报告.

ICMP差错报告报文共有5种:

1. 终点不可达
2. 源点抑制,当路由器或主机由于拥塞而丢弃数据报时,向源点发送抑制报文,使源点知道应当把数据报发送的速率放慢.

(3)时间超过.当路由器收到生存时间为0的数据报时,除丢弃该数据报外,还要向源点发送时间超过报文.当终点在预定时间内不能收到一个数据报的全部数据片,就把已收到的数据片都丢弃,并向源点发送时间超过报文.