一、计算机网络层次

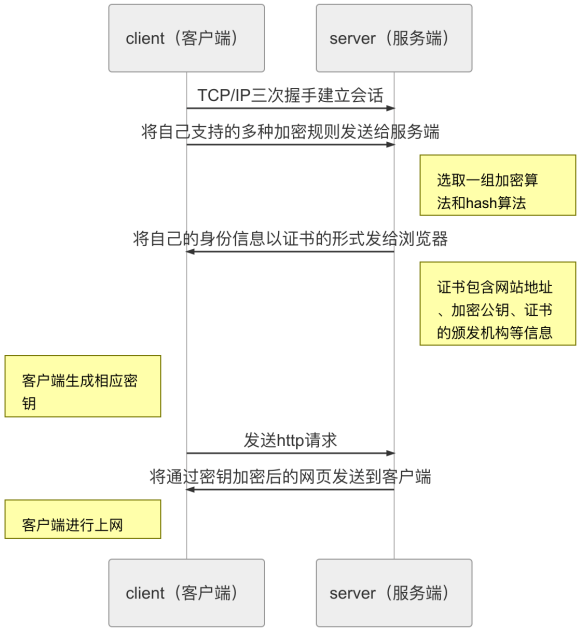
1. 因特网协议栈分为应用层、运输层、网络层、链路层、物理层；OSI 参考模型分为应用层、**表示层、会话层**、运输层、网络层、链路层、物理层。
表示层：使通信的应用程序能够解释交换数据的含义。包括数据压缩，数据加密，数据描述。
会话层：提供数据交换定界和同步功能，包括建立检查点和恢复方案的方法。

二、应用层

1. 应用程序中的两种体系结构：客户服务器结构、P2P 体系结构、混合模式。

> 即时通讯是混合模式，服务器被用于实时跟踪用户的 IP 地址

2. 识别计算机某个进程：IP+PORT，IP 标识计算机，端口标识进程。
3. TCP 是明文传输，为了提高安全性，研制出了安全套接字层(SSL)，它是位于应用层，并且提供了加密、数据完整性和端点鉴别，SSL 有自己的套接字 API。

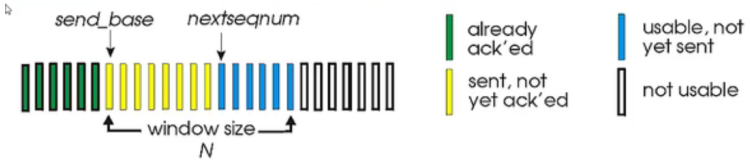


4. DNS(Domain Name System) 是一个由分层的 DNS 服务器实现的分布式数据库，通常运行在 UNIX 服务器上，DNS 协议运行在 UDP 上，端口号为 53。功能：主机别名、邮件服务器别名、负载均衡。
5. DNS 记录分布式存储在不同机器上的，没有一台 DNS 服务器拥有因特网上所有主机的映射。DNS 主要使用 UDP 数据报传送报文，不含前面的各种头部，DNS 报文要求被控制在

512 字节之内，主要考虑是这个大小几乎可以在互联网上畅通无阻，不会因为路径中某个 MTU 太小(MTU 通常总会 >= 576）而导致 IP 分片，从而预防了各种不可预期的后果。这 13 个根域名服务器，并不是只有 13 台物理的服务器。这 13 个根，只是一个逻辑上的概念，每个根 DNS，背后都有多台真正的物理服务器在工作！国内也是有根服务器镜像的。所有根和镜像都有着同样的根区文件。它们一起共享 13 个 IP，泛播技术做到的。(泛播是指某组中任意发送方对应拓朴结构中几个最接近的接收方之间的通信)。

三、运输层

1. 运输层刚好位于网络层之上。网络层主要提供了**主机**之间的逻辑通信，而运输层为运行在不同主机上的**进程**之间提供了逻辑通信。
2. 将主机间家父扩展到进程间交付被称为运输层的多路复用和多路分解。
3. 运输层提供了两种协议 UDP(用户数据报协议)和 TCP(传输控制协议)，UDP 是一种不可考的、无连接的服务，TCP 是一种可靠的、面向连接的服务。但是两个协议都保证传送的数据是正确的，因为有校验和。
4. rdt2.0：产生位错误，接收方显示告知发送方 ACK 或者 NAK，若 NAK 则发送方重传（停等协议，发一个停下等 ACK）。
rdt2.1：ACK/NAK 也可能会产生错误，方法：对 ACK 加一个序列号，ACK0/NAK0 或者 ACK1/NAK1。
rdt2.2：取消 NAK，只存在 ACK0，ACK1 的协议，接收方只确认最后一个被接受的消息。如果校验和失败，则返回上一个接受的 ACK。
rdt3.0：数据包可能发生错误，也可能丢失。当分组丢失时，接收方和发送方就无限等待了。所以增加一个超时重传。发送方等待一段时间，没有 ACK 则重传。任何发送的一方发送时都要启动定时器（接收方发送 ACK 也要启动）。
流水线：几个分组一起发，接收方一起收。改进：更大的序列号范围，需要更大的存储空间缓存分组。
滑动窗口协议：GBN，SR。
5. GBN(回退 N 步)。



ACK(n): 确认收到序列号 n (包含 n)的分组均已被正确接受。为空中的分组设定计时器 Timer，超时事件：只有一个定时器，如果超时则重传所有序列号大于等于 n，但还没收到 ACK 的**所有分组**。当 ACK 都被确认，则 base 相应的增加（if base == nextseqnum），base 增加，停掉计时器。GBN 的接收方没有缓存分组，只需要维护一个期望 expect 收到的序列号，当来到的分组与期望收到的相等，则向上传递并且 expect 增加。对于乱序到达的分组，直接丢弃，重新确认最大的收到的分组。
6.SR(选择重复)。

GBN 缺点：重传时需要重传所有大于等于 base 小于 nextseqnum 分组，这样会产生大量重复。

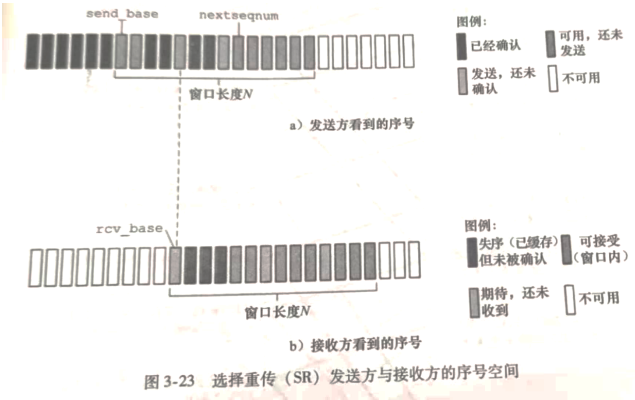


图 3-23　选择重传（SR）发送方与接收方的序号空间

接收方对每个分组单独进行确认。接收方设置缓存机制，缓存乱序到达的分组。发送方只重传那些没收到 ACK 的分组，为每一个分组单独设置定时器。发送方窗口：N 个连续的序列号，限制已发送且未确认的分组。接收方发送的是分组编号的 ACK，而不是 Base 的 ACK。

7.**TCP 连接的三次握手**：客户首先发送一个特殊的 TCP 报文段(SYN 置为 1)，服务器用另一个特殊的 TCP 报文段来响应(SYN/ACK)，最后客户再用第三个特殊报文段作为响应(SYN/ACK)。前两个报文段不承载内容，第三个报文段可以承载内容。

TCP 为什么要建立三次握手而不建立两次？因为服务器第一次的响应可能会丢失，这样客户端超时会开启新的连接，而不是使用旧有的，所以三次及以上的握手可以避免重复连接。

8.TCP 四次挥手：客户端向服务器发送 FIN 控制 segment，服务器收到 FIN 返回 ACK 并关闭连接，服务器继续发送一个 FIN，客户端接受到后返回一个 ACK。原因：客户端第一次发送 FIN，服务器接收到那么服务器不会再发送数据但是会继续接收数据，服务器会发送一个应答告诉客户端收到了这个 FIN。服务器继续发送 FIN 则告诉客户端服务器不会再接受消息了，客户端接受到这个 FIN 后知道服务器不会接受消息则准备停止，并返回 ACK，服务器知道客户端收到消息，也准备停止。

服务器发送 FIN 客户端接受到则表示客户端不会再发送数据，客户端发送 FINACK 表示客户端不会再接受数据，服务器收到 FINACK 表示服务器不会再接受数据。

9.TCP 内容。TCP 序号是建立在传送的字节流上而不是报文段的序列号上。报文段的确认号是主机 A 期望从主机 B 收到的**下一字节**的序号。当超时的时候，TCP 只确认该流中至第一个丢失字节为止的字节，所以 TCP 被称为累积确认(类似 GBN)。TCP 接收方会对数据进行缓存(类似 SR)。

TCP 快速重传机制：如果收到对于一个特定报文段的 3 个冗余 ACK，则在超时事件发生前就会对该报文段进行重传，这大大节约了时间。为什么要三次冗余 ACK？因为可能和乱序混淆，三次冗余 ACK 基本可以断定就是包丢失了。

TCP 更像一个 GBN，但是还是有区别，当第 n 个确认包丢失，GBN 会重传所有 n, n+1, … N 个分组，而 TCP 只重传一个分组 n，此外如果 n+1 的 ACK 在第 n 之前到，则 TCP 甚至不会重传报文段 n。

TCP 流量控制：接收方维护 Buffer 缓存发送方的数据，通过在 Segment 的头部将 RcvWindow 告诉 Sender，Sender 限制自己已经发送但未被收到 ACK 的数据不超过 RcvWindow 尺寸。

RcvWindow == 0，还是会发送一个很小的段，防止死锁。
10.拥塞控制。流量控制是因为接收方的接受能力不行，拥塞控制是因为网络的传递能力不行。两种拥塞控制方法：端到端的拥塞控制、网络辅助拥塞控制(路由器及其他的硬件)。
11.TCP 拥塞控制：使用端到端的方法。类似于流量控制，拥塞控制中 Sender 设置一个变量 CongWin，并且 LastByteSent – LastByteAcked <= CongWin。网络拥塞事件：timeout 或 3 个重复 ACK，此时发送方会降低发送速率。
加性增-乘性减：逐渐增加发送速率，谨慎探测可用带宽。当发生 loss，则肯定已经拥塞应当迅速降低发送速率。每个 RTT 将 CongWin 增大一个 MSS (最大分段大小)，发生 loss 时则将 CongWin 减半。
慢启动：TCP 建立时，CongWin = 1 MSS。当连接开始时，指数型增长。
Threshold：当指数增长达到 threshold 时，则增长转变为线性增长。Reno 拥塞时 CongWin 变为原来一般，然后继续线性增长；Tahoe 则 loss 后 CongWin 变为 1，然后继续指数增长、线性增长。目前：三个重复 ACK 时，CongWin 切到一半然后线性增加；timeout 则直接设为 1 个，然后指数增长，达到 threshold 后线性增长。

四、网络层

1.转发：从一个输入链路接口转移到适当输出链路接口。路由选择：网络范围内决定从源到目的地所采用的端到端路径。
2.转发表：每台路由器都有一个将目的地址映射到链路接口的转发表，当分组到达路由器时，路由器使用该分组的目的地址在转发表中查找合适的输出链路接口。
3.路由器用分组的目的地址的前缀与该表中的表项进行匹配，并且遵循最长匹配规则找到前缀最长的表项。
4.网络层三个组件：IP 协议，路由选择，差错检验(ICMP 协议)。网络层是主机与主机之间的互相连接，所以报文头只有源 IP、目的 IP、寿命、首部长度、运输层协议等。运输层协议表明需要将数据交给上层的哪个运输层协议 UDP 或者 TCP。
5.IP 分片与重组：因为不同链路层的最大传送单元 MTU 不一样，所以 IP 层会选择将其分片成链路层帧然后向下发送。帧重组的过程放在端系统而不是网络路由器中。
6.地址标法是点分十进制，每一个接口都代表了一个子网，接口/xx 表示子网掩码，表示该网络前 24 位是一个子网。

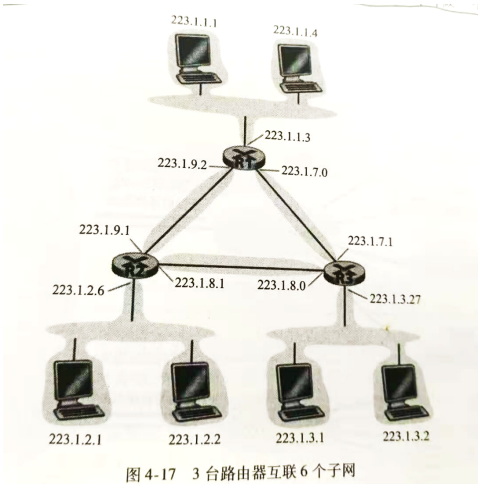


图 4-17　3 台路由器互联 6 个子网

7.DHCP：动态主机配置协议，主机被分配一个临时的 IP 地址，并且还能够知道它的子网掩码、第一跳路由器地址（默认网关）、本地 DNS 服务器地址。
8.NAT：网络地址转换。NAT 路由器对外界的行为反过来就如同一个具有单一 IP 地址的单一设备。NAT 通过 NAT 转换表包含了子网地址+端口和公网地址+端口。
9.路由算法：链路状态算法(Dijkstra, OSPF 路由选择协议)，距离向量算法(Bellman-Ford 算法, RIP, BGP 协议)。距离向量公式：d(x -> y) = min_v{c(x, v) + d(v -> y)}，初始时只更新 y 的一跳距离，循环|V|-1 次的松弛操作，每次操作里遍历所有边集合进行松弛。
10.层次路由选择：将路由器组织为自治系统 AS，在 AS 内路由器全部运行相同的路由选择算法，且彼此拥有相互的信息。在 AS 内的一台或多台路由器负责将本 AS 之外的目的地转发分组，这个路由器被称为网关路由器。AS 间的路由选择协议也必须是同一个，被称为 BGP4.
因特网自治系统内部的路由选择：RIP 和 OSPF。RIP 是距离向量协议，使用条数作为费用测度，最大路径被限制在 15 跳以内，邻居这件使用一种 RIP 响应报文来交换信息，每 30 秒交换一次。OSPF 实际上就是一个使用洪泛链路状态信息的协议+Dijkstra，OSPF 向自治系统内的所有路由器广播路由选择信息，而不仅仅是其相邻的路由器广播。
11.自治系统间的路由选择协议：BGP。在 AS 间的会话称为 eBGP，在 AS 内的两个路由器会话称为 iBGP。BGP 使得每个 AS 知道其相邻 AS 可达哪些目的地。在任何 AS 内的网关路由接受到 eBGP 学习到的前缀后，它会通过 iBGP 向 AS 内的路由器发布这些前缀。
12.广播算法使用最小生成树，每个节点对路径上的邻居进行广播。

四、链路层

1.链路层功能：成帧、链路接入(MAC 规定链帧在链路上传输的规则)、可靠交付、差错检验和纠正。
2.链路层主体部分是网卡，核心是链路层控制器，包括上面 1 的功能，其许多功能是硬件实现的。
3.差错检验纠正方法：奇偶校验
循环冗余校验(CRC)：也称为多项式编码，它将发送的比特串看成系数是 0 和 1 的多项式，对比特串的操作被解释为多项式算数。发送方和接收方协商一个 r+1 比特的多项式，叫生成多项式，称为 G。给定一个数据 D，发送方要选择 r 个附加比特 R，将其附加在 D 上，得到 d+r 长度的比特，它用模 2 算数恰好能被 G 整除。如果余数非 0，则说明传输过程中出现了差错。模 2 除法中加法不进位，减法不借位。D % G = R, T = D + R;　　T % G == 0?
另一种是海明校验，类比于二进制位，3 位校验位能检查八比特数据。
4.多路访问协议：信道划分协议、随机接入协议、轮流协议
信道划分协议：时分多路复用和频分多路复用、码分多址。
随机接入协议：碰撞后就等待一个随机时间再重发，时隙 ALOHA，ALOHA：一旦一个帧首次到达，节点就立即将该帧完整地传输进广播信道。载波监听多路访问：说话之前先监听，如果有数据正在传输则等待一小段时间没有其他数据传输后再开始传输。
轮流协议：轮询协议要求某个结点选作教材主结点，主结点以循环方式询问每个节点是否传输数据。令牌传递协议，没有主结点，一个称为令牌的特殊帧在结点之间以某种固定的次序进行交换。当结点需要有数据发送是才持有令牌，否则就立刻将令牌传递给下一位。
电缆使用 FDM 将下行和上行网络段划分为多个频率信道，上行最大带宽 6.4MHz，下行带宽 6MHz。
5.MAC 地址：是网络接口的硬件地址，当某适配器要向某些目的适配器发送一个帧时，发送适配器的目的适配器的 MAC 地址插入到该帧中，并将该帧发送到局域网上。一块适配器可

以接受并非向它寻址的帧。这样适配器收到一个帧时会检查目的 MAC 地址是否与自己的匹配，如果匹配就会向协议栈上层传递，否则就丢弃。FF-…-FF 地址则是广播帧，交换机都会接受并处理，它是特殊的帧。
6.ARP 协议：IP 地址和 MAC 地址转换的协议。ARP 表即插即用，自动建立的。当发送方的 ARP 表中存目的 IP 地址和 MAC 地址映射时，那么它直接就讲数据传送给对应的 MAC 接口即可；当不知道时，它需要先广播请求，查询这个 IP 的 MAC 地址，然后再转向发送对应的接口。
7.以太网使用 CSMA/CD 协议。

Web 页面请求过程

DNS 部分：首先向本地 DNS 服务器查询 IP 地址->向根服务器查询->向顶级域名服务器查询->权威服务器，期间再重新缓存起来。
发送请求：构建 HTTP 请求报文（如果是 HTTPS 部分，验证证书，交换密钥，发送数据）-> 如果 HTTP 请求而且有代理缓存，则缓存代理先验证是否已经存在数据，如果存在且没过期则直接返回，不去请求。如果没有缓存，则代理服务器去请求，并适当修改请求头。反向代理过程类似。代理请求完毕后返回给客户端。如果没有代理则直接请求。
HTTP 状态：如果是保持连接，则 TCP 不会被关闭，否则发送完请求后 TCP 连接被关闭。当发送完毕后，网卡产生一个中断告知 CPU，CPU 将网卡缓存的数据先拷贝到内核态，然后再从内核态拷贝到用户态，拷贝完毕后，用户使用 IO 轮询或者 EVENT 知道数据已经传送完毕，浏览器开始解析。
里面省略了 TCP 连接、可能存在的 NAT 转换、RIP/OSPF 的内部 IP 寻址、BGP 的主干道路寻址等内容。

多媒体网络

1.流式视频：吞吐量最重要；会话式 IP：时延高敏感、容忍丢包；
2.流式视频系统可分为 UDP 流，HTTP 流和适应性 HTTP 流，后两者最常用。
3.内容分发网(CDN)：CDN 管理分布在多个地理位置上的服务器，在它的服务器中存储视频或图片、文档）的副本，并且所有试图将每个用户请求定向到一个将提供最好的用户体验的 CDN 位置。如果客户向一个未存储该视频的集群请求某视频，该集群向其他集群/中心仓库检索该视频，然后发送给用户并且同时在本地存储一个副本，类似于因特网缓存；当某集群存储器变满时，它删除不经常请求的视频。
4.CDN 操作：首先截获用户请求，然后将客户的请求重定向到该集群的某台服务器。截**获请求**：大多数 CDN 利用 DNS 来截获和重定向请求，请求到用户本地 DNS 后，中继该 DNS 到权威服务器 NetCinema，NetCinema 并不返回 IP 地址，而是向本地 DNS 返回一个 KingCDN 域的主机名，如 a1105.kingcdn.com；本地 DNS 发送第二个请求解析该 url，KingCDN 的 DNS 系统指定流 CDN 的服务器 IP 地址，客户将能够从这台服务器接受到它的内容。集群选择策略：地理上最近、流量实时测量、IP 任播（让因特网中的路由器将客户的分组路由到最近的集群，也就是 CDN 为不同集群指派相同的 IP 地址，当 BGP 接受到对这个相同 IP 地址的多个路由通告时，它对待不同的通告就像对相同的物理位置提供了不同的路径，那么 BGP 会根据其本地路由，对该 IP 地址选择最近的路由）。

网络安全

1.对称密码加密：正常数+k、table 加密、密码块连接（每个块都用同一个 table 加密，但是加密后和一个随机向量 XOR，再加密一边，只用给定初值随机向量即可，后面块的随机向量通过前面的计算出来）。
2.不对称加密：RSA，接收方提供公钥，保留私钥，发送方根据公钥进行加密，接收方用私钥解密。一般情况下不对称加密用来交换对称加密的密钥，不做块数据加密。
3.报文鉴别：A 和 B 先共享报文鉴别码 s，然后 A 发送的数据 m + s 使用散列算法得到一个 H(m+s)，然后 A 发送(m, H(m+s))后，B 因为已经知道 s 了，然后
3.数字签名：数字签名技术是将摘要信息用发送者的**私钥加密**，与原文一起传送给接收者。接收者用**自己的公钥解密**被加密的摘要信息，然后用 HASH 函数对收到的原文产生一个摘要信息，**与解密的摘要信息对比**。如果相同，则说明收到的信息是完整的，在传输过程中没有被修改，否则说明信息被修改过，因此数字签名能够验证信息的完整性。公钥必须给一个可信的机构才可以。
4.SSL：安全套接字层，其稍加修改的版本称为运输层安全性(Transport Layer Security)。Alice 创建一条 TCP；验证 Alice 是真实的 Alice；发送给 Alice 一个主密钥，Bob 和 Alice 持用该主密钥生成 SSL 会话所需的所有对称密钥。
* 一旦创建连接，Bob 就会向 Alice 发送一个 hello 报文，Alice 就会将她的证书响应，证书中包括了她的公钥，并且 CA 证书已被某机构证实过，公钥确实属于 Alice；Bob 产生一个主密钥 MS 用于这个 SSL 会话，并 Alice 的公钥加密该 MS，生成 EMS 发送给 Alice，Alice 用私钥解密该 EMS 从而得到主密钥 MS，此时可以对称加密通信了。
* 实际上生成四个密钥 Eb,Mb,Ea,Ma，其中两个加密密钥用来加密数据（双通道不同的密钥），后两个 MAC 密钥用来验证数据完整性。
* SSL 将数据流分割成记录，然后对每个记录加一个 MAC 用于完整性检查，然后加密该“记录+MAC”。为了产生这个 MAC，发送方将数据联通密钥 Mb 放入一个散列函数。
* SSL 连接关闭：不能直接关闭 TCP，因为可能产生截断攻击，也就是说 A 关闭了，但是中间恶意者 T 捕获该报文，然后向 A 同意关闭，但是悄悄和 B 继续交谈。解决方法是给每一个报文的报文类型里增加一个该报文是否用于终止该 SSL 会话，这样 B 收到的报文就能确保是否是 A 发送的关闭报文，A 也不会被 T 欺骗（T 发送的报文无效）。
HTTPS：构建于 SSL 或者 TLS 的 HTTP 请求，默认端口是 443，CA 全程是 Certificate Authority，认证机构。CA 证书：认证机构颁发的证书。