目录

[第一章 计算机网络和因特网 3](#_Toc109248630)

[1.1什么是因特网 3](#_Toc109248631)

[1.1.1具体构成描述 3](#_Toc109248632)

[1.1.2服务描述 3](#_Toc109248633)

[1.1.3 什么是协议 3](#_Toc109248634)

[1.2网络边缘 3](#_Toc109248635)

[1.2.1 接入网 4](#_Toc109248636)

[1.2.2 物理媒介 4](#_Toc109248637)

[1.3网络核心 5](#_Toc109248638)

[1.3.1分组交换 5](#_Toc109248639)

[1.3.2电路交换 5](#_Toc109248640)

[1.3.3网络的网络 5](#_Toc109248641)

[1.4分组交换网中的时延、丢包和吞吐量 6](#_Toc109248642)

[1.4.1分组交换网中的时延概述 6](#_Toc109248643)

[1.4.2排队时延和丢包 6](#_Toc109248644)

[1.4.3端到端时延 6](#_Toc109248645)

[1.4.4 计算机网络中的吞吐量 6](#_Toc109248646)

[1.5协议层次及其服务模型 6](#_Toc109248647)

[1.5.1分层的体系结构 6](#_Toc109248648)

[1.5.2封装 7](#_Toc109248649)

[1.6面对攻击的网络 7](#_Toc109248650)

[1.7计算机网络和因特网的历史 8](#_Toc109248651)

[1.7.1分组交换的发展：1961~1972 8](#_Toc109248652)

[1.7.2专用网络和网络互联：1972~1980 8](#_Toc109248653)

[1.7.3网络的激增：1980~1990 8](#_Toc109248654)

[1.7.4因特网爆炸：20世纪90年代 8](#_Toc109248655)

[1.7.5最新发展 8](#_Toc109248656)

[1.8小结 8](#_Toc109248657)

[第二章 应用层 8](#_Toc109248658)

[2.1应用层协议原理 8](#_Toc109248659)

[2.1.1网络应用程序体系结构 8](#_Toc109248660)

[2.1.2进程通信 9](#_Toc109248661)

[2.1.3可供应用程序使用的运输服务 9](#_Toc109248662)

[2.1.4因特网提供的运输服务 9](#_Toc109248663)

[2.1.5应用层协议 9](#_Toc109248664)

[2.1.6本书涉及的网络应用 9](#_Toc109248665)

[2.2 Web和HTTP 10](#_Toc109248666)

[2.2.1 HTTP概况 10](#_Toc109248667)

[2.2.2非持续连接和持续连接 10](#_Toc109248668)

[2.2.3 HTTP报文格式 10](#_Toc109248669)

[2.2.4用户与服务器的交互：cookie 10](#_Toc109248670)

[2.2.5 Web缓存 11](#_Toc109248671)

[2.2.6条件GET方法 11](#_Toc109248672)

[2.3因特网中的电子邮件 11](#_Toc109248673)

[2.3.1 SMTP 11](#_Toc109248674)

[2.3.2与HTTP的对比 11](#_Toc109248675)

[2.3.3邮件报文格式 11](#_Toc109248676)

[2.3.4邮件访问协议 12](#_Toc109248677)

[2.4 DNS：因特网的目录服务 12](#_Toc109248678)

[2.4.1 DNS提供的服务 12](#_Toc109248679)

[2.4.2 DNS工作机理概述 12](#_Toc109248680)

[2.4.3 DNS记录和报文 13](#_Toc109248681)

[2.5 P2P文件分发 13](#_Toc109248682)

[2.6视频流和内容分发网 14](#_Toc109248683)

[2.6.1因特网视频 14](#_Toc109248684)

[2.6.2 HTTP流和DASH 14](#_Toc109248685)

[2.6.3内容分发网 14](#_Toc109248686)

[2.6.4 学习案例：Netflix、YouTube和“看看” 15](#_Toc109248687)

[2.7 套接字编程：生成网络应用 15](#_Toc109248688)

[2.7.1 UDP套接字编程 15](#_Toc109248689)

[2.7.2 TCP套接字编程 15](#_Toc109248690)

[2.8 小结 15](#_Toc109248691)

[第三章 运输层 15](#_Toc109248692)

[3.1 概述和运输层服务 15](#_Toc109248693)

[3.1.1 运输层和网络层的关系 15](#_Toc109248694)

[3.1.2 因特网运输层概述 16](#_Toc109248695)

[3.2 多路复用与多路分解 16](#_Toc109248696)

[3.3 无连接运输：UDP 16](#_Toc109248697)

[3.3.1 UDP报文段结构 16](#_Toc109248698)

[3.3.2 UDP检验和 16](#_Toc109248699)

[3.4 可靠数据传输原理 16](#_Toc109248700)

[3.4.1 构造可靠数据传输协议 17](#_Toc109248701)

[3.4.2 流水线可靠数据传输协议 17](#_Toc109248702)

[3.4.3 回退N步 17](#_Toc109248703)

[3.4.4 选择重传 17](#_Toc109248704)

[3.5 面向连接的运输：TCP 18](#_Toc109248705)

[3.5.1 TCP连接 18](#_Toc109248706)

[3.5.2 TCP报文段结构 18](#_Toc109248707)

[3.5.3 往返时间的估计与超时 18](#_Toc109248708)

# 计算机网络和因特网

## 1.1什么是因特网

### 1.1.1具体构成描述

主机host(端系统end system):与因特网相连的设备

通信链路communication link:传输数据的物理媒介

分组交换器packet switch

传输速率transmission rate:单位bit/s或bps

分组packet：发送端将数据分段并为每段加上首部字节所形成的信息包

路由器router:分组交换器的一种，常用于网络核心

链路层交换器link-layer switch:分组交换器的一种，常用于接入网

路径route/path:分组所经历的通信链路和分组交换器

因特网服务提供商Internet Service Provider, ISP

协议protocol

传输控制协议Transmission Control Protocol, TCP：运输层协议，向应用程序提供面向连接的服务，包括确保传递、流量控制、拥塞控制

网络协议Internet Protocol, IP：定义在数据报中的各个字段以及端系统和路由器如何作用于这些字段；决定路由的路由选择协议

因特网标准Internet standard

因特网工程任务组Internet Engineering Task Force, IETF

请求评论Request For Comment, RFC：IETF的标准文档

### 1.1.2服务描述

分布式应用程序distributed application：涉及多个相互交换数据的端系统的应用程序

套接字接口socket interface:发送程序必须遵循的规则集合

### 1.1.3 什么是协议

协议定义了在两个或多个通信实体之间交换的报文的格式和顺序，以及报文发送和/或接收一条报文或其他事件所采取的动作

## 1.2网络边缘

客户client

服务器server

数据中心data center

物联网Internet of Things, IoF

### 1.2.1 接入网

接入网:将端系统物理连接到其边缘路由器的网络

边缘路由器edge router:端系统到其他任何远程端系统的路径上的第一台路由器

数字用户线Digital Subscriber Line, DSL:高速下行信道50kHz~1MHz；中速上行信道4kHz~50kHz；双向电话信道0~4kHz

中心局Center Office, CO

数字用户线接入复用器DSLAM:将模拟信号转换回数字信号

电缆因特网接入cable Internet access

混合光纤同轴Hybrid Fiber Coax, HFC

电缆调制解调器cable modem

电缆调制解调器端接系统Cable Modem Termination System, CMTS:将HFC网络划分为下行和上行信道；将模拟信号转换回数字信号

光纤到户Fiber To The Home, FTTH

有源光纤网络Active Optical Network, AON:即交换以太网

无源光纤网络Passive Optical Network, PON:所有从OLT发送到分配器的分组在分配器处复制

光纤网络端接器Optical Network Terminator, ONT

分配器splitter

光纤线路端接器Optical Line Termination, OLT:提供光信号和电信号之间的转换

局域网LAN

WiFi:基于IEEE 802.11技术的无线LAN接入

长时演进Long-Term Evolution

### 1.2.2 物理媒介

物理媒介physical medium

导引型媒体guided media

非导引型媒体unguided media

无屏蔽双绞线Unshielded Twisted Pair, UTP

共享媒体shared medium

光载波Optical Carrier, OC：OC-n表示链路速率为

同步卫星geostationary satellite

近地轨道Low-Earth Orbiting, LEO

## 1.3网络核心

### 1.3.1分组交换

报文message：位于应用层的信息分组

存储转发传输store-and-forward transmission:在交换机能够开始向输出链路传输该分组的第一个比特之前必须接收到整个分组

输出缓存(输出队列)output buffer(output queue):分组交换器中用于储存准备发送的分组的缓存

排队时延queuing delay:链路正在传输其他分组时，该分组在输出缓存等待的时间

分组丢失(丢包)packet loss:由于输出缓存已满，又有新分组到达，到达的分组或正在排队的分组之一被丢弃

转发表forwarding table:路由器通过转发表将分组首部包含的目的IP地址映射为输出链路

路由选择协议routing protocol：用于配置转发表

### 1.3.2电路交换

电路交换circuit switching

电路circuit:电路交换网络中发送方与接受方的连接

端到端连接end-to-end connection

频分复用Frequency-Division Multiplexing, FDM:将链路跨越的频谱划分为频段，不同的电路使用不同的频段

时分复用Time-Division Multiplexing, TDM:时间被划分为帧，帧再划分为时隙，不同的电路使用每个帧中的指定时隙

带宽bandwidth:频段的宽度

静默期silent period

### 1.3.3网络的网络

客户customer

提供商provider

区域ISP regional ISP

第一层ISP tier-1 ISP

存在点Point of Presence, PoP:提供商ISP的路由器群组，客户ISP可以通过第三方电信提供商租用高速链路将自己的路由器与连接到PoP中的一个路由器中

多宿multi-home:下层ISP同时与多个上层ISP连接

对等peer:相同等级的一对ISP相互连接

因特网交换点Internet Exchange Point, IXP:供多个ISP在此处对等

内容提供商网络content provider network

## 1.4分组交换网中的时延、丢包和吞吐量

## 1.4.1分组交换网中的时延概述

节点处理时延nodal processing delay

排队时延queuing delay,：分组在链路上等待传输的时间

传输时延transmission delay,：将分组的所有比特推入链路所需的时间

传播时延propagation delay,：分组从链路的起点到链路的终点所需的时间

节点总时延total nodal delay：

处理时延：检查分组首部和决定分组的出链路等所需的时间

## 1.4.2排队时延和丢包

流量强度traffic intensity

其中为分组长度，单位bit；为分组到达队列的平均速率，单位分组/秒pkt/s；为传输速率

丢弃drop

### 1.4.3端到端时延

### 1.4.4 计算机网络中的吞吐量

瞬时吞吐量instantaneous throughput

平均吞吐量average throughput

瓶颈链路bottleneck link

## 1.5协议层次及其服务模型

### 1.5.1分层的体系结构

分层layer

服务service

服务模型service model

协议栈protocol stack

自顶向下方法top-down approach

应用层：网络应用程序及应用层程序存留的地方

超文本传输协议HyperText Transfer Protocol, HTTP：提供Web文档的请求和传送

简单邮件传输协议SMTP：提供电子邮件报文的传输，使用TCP协议

FTP：提供两个端系统之间的文件传送

DNS：提供端系统名字与32为比特的网络地址的转换

运输层：在应用程序端点之间传送应用层报文。

UDP：运输层协议，不提供不必要服务

报文段segment：运输层的分组

数据报datagram：网络层分组

网络层：负责将数据报从一台主机移动到另一台主机

帧frame：链路层分组

链路层：提供的服务取决于应用于该链路的特定链路层协议

DOCSIS：链路层协议，应用于以太网、WiFi和电缆接入网

PPP：链路层协议

物理层：将帧中的每个比特从一个节点移动到下一个节点，其协议与链路层和实际传输媒体相关

国际标准化组织ISO

开放系统互联模型OSI：包含应用层、表示层、会话层、运输层、网络层、数据链路层、物理层

表示层：使通信的应用程序能解释交换数据的含义，提供数据压缩、数据加密、数据描述(使应用程序不必担心在不同计算机中表示/存储的内部格式不同的问题)等服务

会话层：提供数据交换的定界和同步功能，包括建立检查点和恢复方案的方法

### 1.5.2封装

封装encapsulation

应用层报文application-layer message

运输层报文段transport-layer segment

网络层数据报network-layer datagram

链路层帧link-layer frame

有效载荷字段payload field

## 1.6面对攻击的网络

恶意软件malware

僵尸网络botnet

自我复制self-replicating

病毒virus

蠕虫worm

拒绝服务攻击Denial-of-Serivce(DoS) attack

弱点攻击：在目标主机上运行易受攻击的应用程序或向操作系统发送特制的报文使服务器停止运行甚至崩溃

带宽洪泛：向目标主机发送大量分组，使目标接入链路拥塞，使合法分组无法到达

连接洪泛：在目标主机中创建大量半开或全开的TCP连接，使主机停止接受合法连接

分布式DoS Distributed DoS, DDoS

分组嗅探器packet sniffer

IP哄骗IP spoofing：将具有虚假源地址的分组注入因特网的能力，应对方法：端点鉴别

## 1.7计算机网络和因特网的历史

### 1.7.1分组交换的发展：1961~1972

### 1.7.2专用网络和网络互联：1972~1980

网络互联internetting

### 1.7.3网络的激增：1980~1990

### 1.7.4因特网爆炸：20世纪90年代

### 1.7.5最新发展

## 1.8小结

# 第二章 应用层

## 2.1应用层协议原理

### 2.1.1网络应用程序体系结构

应用程序体系结构application architecture

客户-服务器体系结构client-server architecture：有一台总是打开的主机称为服务器，服务于其他被称为客户的主机的请求。客户相互之间不直接通信，服务器具有固定且周知的地址，称为IP地址

P2P体系结构 P2P architecture：对位于数据中心的专用服务器有最小(或者没有)的依赖。应用程序在间断连接的主机对之间使用直接通信，这些主机对被称为对等方

自扩展性self-scalability

### 2.1.2进程通信

进程process

在一对进程之间的通信会话场景中，发起通信的进程被标识为客户，在会话开始时等待联系的进程是服务器

套接字socket

应用程序编程接口Application Programming Interface, API

端口号port number

端口号80用来标识Web服务器

端口号25用来标识使用SMTP协议的邮件服务器进程

周知端口号列表查询网址：[www.iani.org](http://www.iani.org)

### 2.1.3可供应用程序使用的运输服务

运输层协议能提供的服务可分为可靠数据传输、吞吐量、定时、安全性

可靠数据传输reliable data transfer

容忍丢失的应用loss-tolerant application

带宽敏感的应用bandwidth-sensitive application

弹性应用elastic application

### 2.1.4因特网提供的运输服务

TCP服务模型包括面向连接服务和可靠数据传输传输服务，具有拥塞控制机制

面向连接服务指交换报文前，TCP会让客户和服务器交换运输层控制信息(握手过程)，随后建立TCP连接，这条连接是全双工的(双方可以同时在连接上收发报文)，且结束时必须拆除该连接

可靠数据传输服务是指通信进程能依靠TCP无差错，按适当顺序交付所有数据而没有丢失或冗余

安全套接字层Secure Sockets Layer, SSL：一种对TCP的加强，于应用层上实现，提供加密、数据完整性、端点鉴别

UDP服务：无连接(没有握手过程)，提供不可靠数据传输服务(不保证数据是否到达和到达的顺序)，没有拥塞控制机制

### 2.1.5应用层协议

应用层协议application-layer protocol：定义交换的报文类型(如请求报文和响应报文)、报文语法、字段语义、报文发生时间和发送方式、报文响应规则

### 2.1.6本书涉及的网络应用

本书讨论Web、电子邮件、目录服务、文件传输、流式视频五种网络应用

## 2.2 Web和HTTP

### 2.2.1 HTTP概况

Web页面(文档)Web page

对象object

HTML基本文件base HTML file

Web浏览器Web browser

Web服务器Web server

HTTP是无状态协议(不保存关于客户的信息)，以TCP作为支撑运输协议

无状态协议stateless protocol

### 2.2.2非持续连接和持续连接

非持续连接non-persisent connection：服务器对客户的每个请求/响应对经单独的TCP连接发送

持续连接persistent connection：服务器对客户的所有请求/响应对经相同的TCP连接发送

往返时间Round-Trip Time, RTT：一个短分组从客户到服务器再返回客户所花的时间

非持续连接的HTTP的流程：客户发起TCP连接，服务器响应连接并回应；客户发出HTTP请求，服务器发送响应；服务器通知TCP断开连接；TCP确认客户接收到响应后关闭

持续连接的HTTP服务器会在一定无请求间隔后才将未使用的连接关闭

### 2.2.3 HTTP报文格式

请求行request line：包括方法字段、URL字段、HTTP版本字段

首部行header line：仅包括值字段

实体体entity body

状态行status line：包括协议版本字段、状态码、状态信息

部分常见状态码及其相关信息：

200 OK(请求成功)

301 Moved Permanently(请求对象已被永久转移，新的URL在响应报文的首部行“Location:”中)

400 Bad Request(通用差错代码，请求不能被服务器理解)

404 Not Found(被请求的文档不在服务器上)

505 HTTP Version Not Supported(服务器不支持请求报文使用的HTTP协议版本)

### 2.2.4用户与服务器的交互：cookie

Cookie技术包含4个组件：响应报文的首部行、请求报文的首部行、用户端系统中由浏览器管理的cookie文件、Web站点的后端服务器

### 2.2.5 Web缓存

Web缓存器Web cache(代理服务器proy server)：能够代表初始Web服务器满足HTTP请求的网络实体

内容分发网络Content Distribution Network, CDN

### 2.2.6条件GET方法

条件GET conditional GET

条件GET方法能确保代理服务器中储存的对象是最新的。当浏览器请求一个代理服务器已存储的对象时，代理服务器向Web服务器发送一个包含If-modified-since首部行的请求报文，表示仅当请求对象的更新日期新与此才发送对象，否则仅发送包含304 Not Modified的状态码

## 2.3因特网中的电子邮件

因特网电子邮件系统的组成部分：用户代理user agent、邮件服务器mail server、简单邮件传输协议Simple Mail Transfer Protocol, SMTP

邮箱mailbox

报文队列message queue

### 2.3.1 SMTP

SMTP限制邮件报文的体部分使用7比特ASCII码，即传输邮件前需要将二进制多媒体数据编码为ASCII，到达接受方的邮件服务器后再解码还原，而HTTP不需要

SMTP不使用中间服务器转发，如果发送方邮件服务器无法连接接受方邮件服务器，则会等待再次尝试，多次失败后从报文队列移除并通知发送方

SMTP使用TCP的持续连接

### 2.3.2与HTTP的对比

拉协议pull protocol：由接收方发起连接

推协议push protocol：由发送方发起连接

### 2.3.3邮件报文格式

首部必须包含“From:“首部行和”To:”首部行，首部和报文体用空行分隔

### 2.3.4邮件访问协议

由于接受方访问邮件服务器获取邮件是一个拉操作，而SMTP是推协议，因此使用邮件访问协议解决这一问题，常用的有：POP3、IMAP和HTTP

授权authorization

第三版邮局协议Post Office Protocol-Version 3, POP3：极为简单的邮件访问协议。客户通过邮件服务器的110端口进行TCP连接，连接建立后POP3开始工作。其工作包括授权(用户发送用户名和口令)、事务处理(用户取回报文，同时可标记删除等操作)、更新(客户发出”quit”命令后断开连接，服务器更新操作)。POP3工作时保留但不显示会话状态信息

因特网邮件访问协议Internet Mail Access Protocol, IMAP：解决POP3中邮件在不同端间断访问困难的问题。IMAP为用户接收到的邮件在服务器创建一个INBOX文件夹，用户可以将邮件移动至新创建的文件夹进行阅读或删除操作。IMAP提供创建、移动、查询等命令。IMAP服务器维护会话的用户状态信息，如文件夹的名字，文件夹与邮件的关系。IMAP支持仅阅读邮件的一部分

## 2.4 DNS：因特网的目录服务

主机名hostname

IP地址IP address：由4个字节组成，每个字节用”.”分隔

### 2.4.1 DNS提供的服务

域名系统Domain Name System, DNS：由分层的DNS服务器实现的分布式数据库和使主机能查询分布式数据库的应用层协议组成

DNS服务器 DNS server：通常是运行BIND(Berkeley Internet Name Domain)软件的UNIX机器

DNS协议运行在UDP上，使用53号端口、使用客户-服务器模式

主机别名host aliasing

规范主机名canonical hostname

邮件服务器别名mail server aliasing

负载分配load distribution：一个主机名可能对应一个IP地址集合，当客户发出DNS请求时，服务器会用整个集合进行响应，但客户只会向排在最前面的IP地址发送请求报文，因此服务器会调整集合里的顺序，使不同IP的负载均衡

### 2.4.2 DNS工作机理概述

DNS工作概述：用户程序需要将主机名转换成IP地址，程序调用本地的DNS客户端，客户端向网络发送DNS查询报文，得到DNS回答报文后客户端传递给程序

单点故障a single point of failure

通信容量traffic volume

远距离集中式数据库distant centralized database

维护maintenance

根DNS服务器

顶级域DNS服务器Top-Level- Domain, TLD：如”.com”、”.org”、”.edu”

权威DNS服务器

本地DNS服务器locol DNS server

递归查询recursive query：主机依次查询根服务器、TLD、权威服务器

迭代查询iterative query：主机查询根服务器，根服务器查询TLD

DNS缓存DNS caching：DNS服务器将请求链的回答存储在本地，当再次遇到相同的请求能够直接返回回答而不必传递请求。由于映射会发生变动，因此存储不是永久的，一段时间后服务器会丢弃缓存信息

### 2.4.3 DNS记录和报文

资源记录Resource Record, RR：提供主机名到IP地址的映射，由四个字段组成”(Name, Value, Type, TTL)”，每个DNS回答报文包含若干资源记录

若Type为A，则Name为主机名，Value为IP地址

若Type为NS，则Name为域，如foo.com，Value为能得到该域的IP地址的权威DNS服务器主机名

若Type为CNAME，则Name为主机名，Value为其规范主机名

若Type为MX，则Name为主机名，Value为其邮件服务器的规范主机名

TTL为存活时间，决定何时将其从缓存中删去

DNS报文只有查询报文和回答报文，且两种报文格式相同，由首部、问题、回答、权威、附加组成

首部：共12字节，分为标识符(16bit，用于匹配查询和回答)，标志(含有若干标记，如报文类型，查询0，回答1；是否来自权威DNS服务器，是1；在DNS服务器无记录时客户是否希望递归查询，希望1；回答报文中标记DNS服务器是否支持递归查询)，问题数，回答RR数，权威RR数，附加RR数

问题：包含正在进行的查询信息，包括名字字段(正在查询的主机名)，类型字段(问题类型)

回答：包含对最初请求的RR，可能包含多条RR

权威：包含其他权威服务器的RR

附加：包含其他RR，如对于一个MX请求的回答报文，回答区域包含多条RR，附加区域可能包含A记录

注册登记机构registrar：商业实体，验证域名的唯一性，将域名输入DNS数据库

因特网名字与地址分配机构Internet Corporation for Assigned Names and Numbers, ICANN

当客户向注册登记机构注册域名时，需要提供权威DNS服务器的名字和IP，机构确保顶级域DNS服务器含有一个A记录和NS记录，而客户确保权威DNS服务器含有一个Web服务器的A记录和邮件服务器的MX记录

针对DNS的攻击：DDoS、中间人攻击(截获主机的请求并返回伪造的回答)、毒害攻击(向DNS服务器发送伪造的回答，使服务器缓存这个回答)

## 2.5 P2P文件分发

分发时间distribution time

客户-服务器体系结构分发时间下界

其中

为客户-服务器体系结构分发时间

为客户数

为文件大小

为服务器上传速度

为客户最小下载速度

P2P体系结构分发时间下界

其中为P2P体系结构分发时间

为第i个客户的下载速度

洪流torrent：参与一个特定文件分发的对等方集合

块chunk

追踪器tracker

邻近对等方：成功创建TCP连接的对等方

最稀缺优先：从自己没有的块中请求邻近对等方拥有最少的块，有利于均衡每个块的副本在洪流中的数量，使最稀缺的块更为迅速的分发

疏通unchoked：向自己发送数据最快的数个对等方

## 2.6视频流和内容分发网

### 2.6.1因特网视频

### 2.6.2 HTTP流和DASH

在HTTP流中，视频被视为一个文件，所有客户请求的始终是同一个文件

经HTTP的动态适应流Dynamic Adaptive Streaming over HTTP,DASH

告示文件manifest file

在DASH中，视频编码为比特率不同的几个版本，客户首先请求服务器的告示文件得知不同的版本，然后每次根据自身带宽请求其中一个版本的一小段数据块

### 2.6.3内容分发网

内容分发网Content Distribution Network, CDN

专用CDN, private CDN

第三方CDN, third-praty CDN

CDN的服务器安置原则：深入(将服务器集群部署在接入ISP中)、邀请做客(将服务器集群部署在IXP中)

CDN操作：当用户浏览器指令检索一个特定的视频(由URL标识)时，CDN必须截获该请求以便确定合适的CDN集群和将用户请求重定向到该集群的某台服务器上

使用DNS截获和重定向请求的步骤：用户首先访问目标网页；用户点击视频链接时，用户主机发送对于该视频的DNS请求；用户的本地DNS服务器(LDNS)将DNS请求中继到其权威DNS服务器，权威DNS服务器识别到DNS请求中目标主机名的特定字符串，返回第三方CDN的主机名；LDNS向第三方CDN权威DNS服务器发送DNS请求，服务器返回CDN节点IP；LDNS向用户转发该IP；用户通过IP与第三方CDN创建连接，获得视频

集群选择策略cluster selection strategy：将用户定向到某个CDN服务器集群或数据中心的机制

地理最近geographically closest

实时测量real-time measurement

### 2.6.4 学习案例：Netflix、YouTube和“看看”

## 2.7 套接字编程：生成网络应用

### 2.7.1 UDP套接字编程

服务器进程在客户发起连接前已启动

分组的目的地址包括目的主机的IP和目的地套接字的端口号

### 2.7.2 TCP套接字编程

连接套接字connectionSocket ：在客户连接欢迎套接字后服务器专门生成供连接的新套接字

## 2.8 小结

# 第三章 运输层

## 3.1 概述和运输层服务

逻辑通信logic communication

### 3.1.1 运输层和网络层的关系

网络层提供主机间的逻辑通信

运输层提供不同主机上的进程间的逻辑通信

运输层提供的服务受制于网络层(如时延和带宽保证)，但也能提供网络层没有的服务(如可靠性和机密性)

### 3.1.2 因特网运输层概述

IP提供尽力而为交付服务，是不可靠服务

多路复用multiplexing

多路分解demultiplexing

拥塞控制congestion control

## 3.2 多路复用与多路分解

多路分解：将运输层报文段中的数据交付给正确的套接字的工作

多路复用：从源主机的不同套接字中收集数据块，并为这些数据块封装上首部信息生成报文段，然后将报文段传递至网络层

源端口号字段source port number field

目的端口号字段destination port number field

端口号：套接字的唯一标识，16比特，0~1023为周知端口号

UDP套接字由二元组(包含目的IP地址和目的端口号)全面标识

TCP套接字由四元组(包含源IP地址，源端口号，目的IP地址，目的端口号)全面标识

## 3.3 无连接运输：UDP

### 3.3.1 UDP报文段结构

UDP报文段由首部和数据组成，首部包含源端口号，目的端口号，长度(首部和数据)，检验和四个字段组成，各2字节

### 3.3.2 UDP检验和

发送方对报文段每16比特字求和，溢出回卷，反码运算的结果放在检验和处，接受方将包括检验和的每16比特求和，若无差错，结果应全为1

端到端原则end-end principle

## 3.4 可靠数据传输原理

可靠数据传输协议reliable data transfer protocol

单向数据传输unidirectional data transfer：数据从发送端到接收端

双向数据传输(全双工数据传输)bidirectional data transfer：允许数据同时在两个方向上传输

### 3.4.1 构造可靠数据传输协议

经完全可靠信道的可靠数据传输：rdt1.0

有限状态机 Finite-State Machine, FSM

经具有比特差错信道的可靠数据传输：rdt2.0

肯定确认positive acknowledgment, ACK

否定确认negative acknowledgment, NAK

自动重传请求Automatic Repeat request,ARQ

ARQ协议包含：差错检测，接受方反馈，重传三个功能

停等stop-and-wait

冗余分组duplicate packet

序号sequence number

经具有比特差错的丢包信道的可靠数据传输 rdt3.0

冗余数据分组duplicate data packet

倒计数定时器countdown timer

比特交替协议alternating-bit protocol

### 3.4.2 流水线可靠数据传输协议

信道利用率utilization：发送方将比特送入信道的时间与发送时间之比

流水线pipelining

回退N步Go-Back-N, GBN

选择重传Selective Repeat, SR

### 3.4.3 回退N步

窗口长度window size

滑动窗口协议sliding-window protocol

GBN发送方必须响应三种类型的事件：上层调用，接收ACK，超时

累计确认cumulative acknowledgment：对序号n的分组的确认表示接收方已正确接收到序号为n及之前的分组

基于事件编程event-based programming

### 3.4.4 选择重传

对于SR协议，窗口长度必须小于序号空间的一半

## 3.5 面向连接的运输：TCP

### 3.5.1 TCP连接

面向连接的connection-oriented

全双工服务full-duplex service

点对点point-to-point

多播：在一次发送操作中，从一个发送方将数据传送给多个接收方

三次握手three-way handshake

发送缓存send buffer

最大报文段长度Maximum Segment Size, MSS：不包括TCP报文段的首部

最大传输单元Maximum Transmission Unit, MTU：最大链路层帧长度

TCP报文段TCP segment

### 3.5.2 TCP报文段结构

TCP报文段由首部字段和数据字段组成，首部字段一般20字节，包括

源端口号，目的端口号各16bit

检验和字段checksum field, 16bit

序号sequence number field，确认号字段acknowledgment number field各32bit

接收窗口字段receive field, 16bit，用于指示接收方愿意接受的字节数量

首部长度字段header length field, 4bit，指示以32bit为1单位的TCP首部长度

选项字段options field, 不定长

标志字段flag field, 6bit，包括ACK, RST, SYN, FIN, PSH, URG

紧急数据指针字段urgent data pointer field, 16bit

TCP将数据看作无结构、有序的字节流，序号表示发送方报文段首字节的在整个字节流编号，确认号表示接收方希望获得的字节编号，也即表示之前的字节已确认接收。初始序号不必是0，但即使报文段没有数据也要填入序号

回显echo back

捎带piggybacked

### 3.5.3 往返时间的估计与超时