

应用层

1. 网络应用模型 *

1.1. 客户/服务器模型 *

- 工作流程：
 - 服务器处于**接收请求**状态
 - 客户机发出**服务请求**，并等待接收结果
 - 服务器收到请求，分析请求，进行处理，得到结果并发送给客户机
- 特征：
 - 客户是 **服务请求方**，服务器是 **服务提供方**。一半协同工作的进程是 **两个**。
 - 整个网络中只需要**少数服务器**就可以管理客户机，使得网络管理集中方便。
 - **客户机之间不直接通信**。
 - 可拓展性**不强**，服务器支持客户机数量有限。
- 举例：Web，FTP，远程登陆，电子邮件。

1.2. P2P模式 *

- 概述：
 - 网络中每个结点都具有**上传，下载**功能
 - **不设中心服务器**，没有固定客户和服务器划分，每个结点权力和义务对等。
 - 任意一对计算机称为 **对等方**，可以 **直接互相通信**
- 优点：
 - 减轻服务器计算压力，可以将任务分配到各个节点，**提高系统效率和资源利用率**。
 - 客户机之间可以直接共享文档。
 - **可拓展性好**，不受传统服务器的响应和带宽限制。
 - 网络健壮性强，单个结点失效不影响其他部分结点。
- 类型：集中目录式/全分布式/混合式等

2. 域名系统(DNS) *

- 作用：把**具有特定含义的主机名**转换为**IP地址**，使用户可以通过主机名访问网络上的主机。
- DNS系统采用 **C/S模型**，协议运行于 **UDP** 上，使用53号端口

2.1. 层次域名空间 *

- 因特网采用 **层次树状结构** 的命名方法。任何一个连接到因特网的主机/路由器都有一个**唯一层次结构名称**，称为 **域名**。
- **域** 是名字空间中一个可被管理的部分。域可划分为 **子域**，子域可划分为更小的子域。每个域名之间用 **.** 分隔。
- 注意事项：
 - 标号英文**不区分大小写**
 - 标号中除连字符 **-** 外不能使用其他标点符号。
 - 每个标号不超过63个字符，多标号组成完整域名最长不超过255个字符。
 - 级别最低的域名写在最左边，级别最高的顶级域名写在最右边。
- 顶级域名分为三大类：

- **国家顶级域名 (nTLD)**: 国家和某些地区的域名, 如.cn, .us, .uk等。
- **通用顶级域名 (gTLD)**: 常见的有.com, .net, .org, .edu, .gov等。
- **赞助顶级域名 (sTLD)**: 有说法是例如.edu, .org等属于赞助顶级域名, 但书上把它们都归入通用顶级域名内。
- **基础结构域名 (sTLD)**, 反向域名解析, 又称**反向域名**, 仅包含一个, 地址和路由参数区, .arpa。

2.2. 域名服务器 *

- **域名服务器**: 负责把域名解析成为IP地址, 管辖范围称为 **区**。每个区**所有节点相通**。
- 因特网上所有主机的映射分布在所有DNS上。
- DNS域名服务器都把数据复制到几个域名服务器来保存, 其中的一个是**主域名服务器**, 其他的是**辅助域名服务器**。
- 当主域名服务器出故障时, 辅助域名服务器可以保证 DNS 的查询工作不会中断, 提高**域名服务器的可靠性**。

2.2.1. 根域名服务器 *

- **根域名服务器**: 负责管理顶级域名服务器, 是**最高层次**的域名服务器, 所有根域名服务器都知道所有顶级域名服务器的IP地址。
- 无法解析的域名, 首先向**根域名服务器**查询, 返回**顶级域名服务器**的IP地址。
- **IPv4根域名服务器**全球一共 **13** 个: 一个主根服务器 (美国), 12个辅助根服务器 (美国9个, 瑞典斯德哥尔摩1个, 英国伦敦1个, 日本东京1个)

2.2.2. 顶级域名服务器 *

- 负责管理**所有二级域名**, 收到DNS查询请求时, 就给出相应的回答。可能是最后的结果, 也可能是下一步查找的域名服务器的IP地址。

2.2.3. 权限域名服务器 *

- 每台主机都必须在**权限域名服务器**上登记。为了更加可靠地工作, 一台主机最好有两个授权域名服务器。

2.2.4. 本地域名服务器 *

- 一台主机发送DNS请求时, 查询请求白问就发送给主机的 **本地域名服务器**。

2.3. 域名解析过程 *

- **域名解析**: 把域名映射称为IP地址 (**正向解析**) 或把IP映射成域名 (**反向解析**) 的过程。
- **递归查询**: 本地域名服务器向根域名服务器查询**一次**, 其余查询递归地在其他几个域名服务器之间进行。
- 根域名服务器负载过大, 实际几乎不适用, 大多使用**递归与迭代结合**的查询方式。
- 例: 查询**www.super2021.com**
- 主机向**本地域名服务器**查询: **递归查询**, 与之前相同
- **本地域名服务器**向**根域名服务器**查询: **迭代查询**
 - 客户机向**本地域名服务器**发送DNS请求报文 (**递归查询**)。

- **本地域名服务器**查询 **本地缓存**，如果没有记录，则以**DNS客户机**的身份向**根域名服务器**发送解析请求报文 (**迭代查询**)
- **根域名服务器**收到请求后，判断域名属于.com顶级域，返回**com顶级域名服务器**的IP地址。
- **本地域名服务器**向**com顶级域名服务器**发送解析请求报文 (**迭代查询**)。
- **顶级域名服务器**收到请求后，判断域名属于**super2021.com**二级域名，返回**super2021.com授权域名服务器**的IP地址。
- **本地域名服务器**向**super2021.com授权域名服务器**发送解析请求报文 (**迭代查询**)。
- **授权域名服务器**收到请求后，判断域名属于 **www.super2021.com** 主机，返回 **www.super2021.com主机** 的IP地址。
- **本地域名服务器**将结果保存到 **本地缓存**，同时返回给**客户机**。
- **高速缓存**：提高查询效率，减少因特网上DNS查询报文数量。DNS服务器在一段时间后会丢弃高速缓存中的信息。

3. 文件传输协议(FTP)

3.1. FTP的工作原理

- FTP提供交互式的访问，允许用户指明文件的类型与格式，并允许文件具有存取权限，并屏蔽了各计算机系统的细节，适用于在异构网络中任意计算机之间传送文件
- FTP使用C/S方式实现
 - 服务器主进程打开TCP21端口，等待客户进程发出的连接请求
 - 客户可以用分配的任意一个本地端口号与服务器进程的TCP21端口进行连接
 - 客户请求到来时，服务器主进程启动从属进程来处理客户进程发来的请求
 - 服务器从属进程对客户进程的请求处理完毕后即终止，但从属进程在运行期间根据需要还可能创建其他一些子进程
 - 服务器主进程返回，继续等待接收其他客户进程发来的连接请求，服务器主进程与从属进程并行工作

3.2. 控制连接与数据连接

3.2.1. 控制连接

控制连接在整个会话期间一直保持，客户进程发出的文件传输请求通过控制连接发送给服务器控制进程（工作在TCP21端口），但控制连接不用来传输文件

3.2.2. 数据连接

服务器控制进程在接收到客户进程发送来的文件传输请求后就创建数据传输进程（工作在 TCP20端口）和数据连接

- 数据连接用来连接客户进程和服务器数据传输进程，实际完成文件的传输。服务器数据传输进程在文件传输完毕后关闭数据连接并结束运行

4. 电子邮件 *

4.1. 电子邮件系统的组成结构 *

- 电子邮件是一种**异步通信方式**，通信时不需要双方在场。
- 电子邮件系统又三个组成构建：**用户代理UA**、**邮件服务器**、**协议**（SMTP，POP3，IMAP等）。

- **用户代理UA**：用户与电子邮件系统的接口。通常情况下，用户代理就是一个运行在PC上的程序。
- **邮件服务器**：发送和接收邮件，同时向发信人报告邮件传送情况。采用 **C/S** 方式工作，但要**同时充当客户和服务器**。
- **协议**：规定了电子邮件系统的工作方式。**SMTP邮件发送协议**（推）、**POP3邮件接收协议**（拉）。
- 电子邮件发送过程：
 - 发信人写邮件，**用户代理**用 **SMTP** 把邮件发送给**发送端邮件服务器**。
 - **发送端邮件服务器**将邮件放入缓存队列，等待发送
 - **发送端SMTP客户进程**发现缓存中有待发送邮件，向**接收端SMTP服务器**发起TCP连接的建立。
 - TCP连接建立后，**发送端SMTP客户进程**开始向**接收端SMTP服务器**发送邮件，发送完毕后关闭TCP连接。
 - **接收端SMTP进程**收到邮件，把邮件放入收信用户邮箱，等待用户读取。
 - 收件人收信时，调用**用户代理**，使用 **POP3/IMAP** 协议将邮件从**接收端邮件服务器**中取回。

4.2. 电子邮件格式与MIME *

4.2.1. 电子邮件格式 *

- 电子邮件由 **信封** 和 **内容** 组成，内容又分为 **首部** 和 **主体**。
- **首部**包含**首部行**，首部行由**关键字 + : + 值**组成。最重要的关键字是 **To** 和 **Subject**。
- **To**：必须的关键字，后面填入一个或多个收件人的电子邮箱地址。格式规定为：**收件人邮箱名 @ 邮箱所在主机域名**，例如3239695085@qq.com，邮件地址在整个因特网上**唯一**。
- **Subject**：可选关键字，邮件的主题，反应邮件的主要内容。

```
From:3239695085@qq.com      }  
To:20373614@buaa.edu.cn    }-首部  
Subject:Computer Network   }  
  
Hello World!                主体
```

4.2.2. 多用途网际邮件扩充(MIME) *

- 局限：**SMTP**只能传送定长的**ASCII码**邮件，很多语言无法传送。
- MIME并未改动SMTP协议，而是在SMTP协议的基础上，增加了主体结构，定义了传送**非ASCII码**的**编码规则**。
- 内容：
 - 5个新的首部字段：**MIME版本**，**内容描述**，**内容标识**，**传送编码**和**内容类型**
 - 定义众多邮件内容格式，对多媒体电子邮件的表示方法进行标准化
 - 定义传送编码，可对任何内容格式进行转化。

4.3. SMTP和POP3 *

4.3.1. SMTP *

- **简单邮件传输协议SMTP**：使用 **C/S** 方式，因此发送邮件的SMTP就是客户，接收邮件的SMTP就是服务器。
- **连接建立**：

- 发送方使用熟知端口号25与接收方建立 **TCP** 连接。接收方发出 **220 Service ready** 服务就绪的响应，发送方发出 **HELO** 命令，并附上发送方主机名。
- **不使用中间的邮件服务器**，发送方直接与接收方建立连接。
- **邮件传送：**
 - 客户端从 **MAIL** 命令开始，后跟From首部行。
 - 服务器若准备好接收邮件，回答 **250 OK** 或 **550 No such user here**。
 - 客户端发送一个或多个 **RCPT** 命令，后跟To首部行。
 - 每发送一个RCPT命令，都有相应信息从服务器返回。
 - 获得回答后，即确认接收方系统已做好接收邮件准备，客户端发送 **DATA** 命令，告诉服务器开始发送邮件。
 - 服务器回答 **354 Start mail input; end with <CRLF>.<CRLF>**，之后客户端发送邮件内容，以 **<CRLF>.<CRLF>** 结束。
- **连接释放：**
 - 客户端发送 **QUIT** 命令，服务器回答 **221 Service closing transmission channel**，表示SMTP同意释放TCP连接，至此邮件传送过程结束。

4.3.2. POP3和IMAP *

- **邮局协议POP**：十分简单但功能有限的**邮件读取协议**。
 - POP采用 **C/S工作方式**，**传输层使用TCP**，端口号110。
 - 两种工作方式：**下载并保留**，**下载并删除**（用户是否可以再次从服务器上读取）
 - 第3个版本 **POP3采用拉**的通信方式，即用户读取邮件时，用户代理向邮件服务器发出请求，**拉取**用户邮箱中的邮件。
- **因特网报文存取协议IMAP**：相比POP复杂许多
 - 为用户提供创建文件夹，移动邮件等功能，为此IMAP维护了会话用户的状态信息。
 - IMAP允许用户代理只获取报文的某些部分，例如可以只读取一个报文的首部，或多部份MIME报文的一部分。
 -

5. 万维网(WWW) *

5.1. 万维网的概念与组成结构 *

- **万维网WWW**：是一个**分布式的、联机式**的信息存储空间，在此空间中：一样有用的事物称为**一样资源**，并由一个**全域统一资源定位符URL**标识，这些资源通过**超文本传输协议HTTP**传送给使用者，而后者通过单机链接获取资源。
- 万维网使用**链接**的方法从因特网上的一个站点访问另一个站点。**超文本标记语言HTML**能很方便地用一个超链接从页面链接到因特网上任何一个万维网页面。
- 万维网内核部分组成：
 - **统一资源定位符URL**：标识万维网上各种文档
 - **超文本传输协议HTTP**：应用层协议，使用TCP连接进行可靠传输。是万维网客户程序和服务器程序之间交互必须遵守地协议。
 - **超文本标记语言HTML**：一种文档结构的标记语言。
- URL格式：**<协议>://<主机>:<端口>/<路径>**
- 万维网以 **C/S** 方式工作，浏览器是用户主机上万维网客户程序，万维网文档驻留的主机是万维网服务器。
 - Web用户用浏览器与服务器建立连接，发送浏览请求

- 服务器把URL转换为文件路径并返回给Web浏览器。
- 通信完成，关闭连接。

5.2. 超文本传输协议(HTTP) *

- HTTP是**面向事务的应用层协议**，规定了浏览器和服务器的请求和相应的格式和规则，是万维网上能够可靠地交换文件的重要基础。

5.2.1. HTTP的操作过程 *

- 浏览器和服务器的请求和响应的交互，必须遵循HTTP，故HTTP有两类报文：**请求报文**和**响应报文**。
- 例：访问**www.super2021.com**
 - 浏览器分析链接指向页面 **URL** (www.super2021.com)
 - 浏览器向DNS请求解析 **www.super2021.com** 的IP地址
 - 域名系统DNS解析出IP地址
 - 浏览器与该服务器建立TCP连接（默认端口80）
 - 浏览器发出HTTP请求：**GET /index.htm**
 - 服务器通过响应把文件**index.htm**发送给浏览器
 - 释放TCP连接
 - 浏览器释放文件index.htm，并将Web页显示给用户

5.2.2. HTTP的特点 *

- HTTP使用 **TCP**作为传输层协议，保证了数据的可靠性，但是HTTP本身是**无连接**的。
- HTTP是**无状态**的，同一个客户第二次访问同一个服务器上的页面时，服务器响应与第一次被访问是一样的。
- HTTP的无状态特性**简化了服务器的设计**，使服务器更容易支持大量并发的HTTP请求。实际应用中，通常使用 **Cookie加数据库** 的方式跟踪用户的活动。
- **Cookie** 工作原理：
 - 当某个用户浏览某个使用Cookie的网站时，该网站的服务器就为用户产生一个**唯一识别码**，并在用户的**响应报文**中添加一个**Set-cookie**的首部行。
 - 用户继续浏览这个网站时，会取出这个网站的识别码，并放入**请求报文**的首部行。
 - 服务器根据请求报文中的 **Cookie识别码** 就能从数据库中查询到该用户的活动记录，执行一些个性化操作。
- HTTP既可以使用**非持久性连接**，也可以使用**持久连接**。
- 对于**非持久性连接**，每个网页元素对象的传输都要**单独建立TCP连接**。每个对象的引入都要导致 **2×RTT** 的开销，一次用于建立TCP连接，另一次用于请求和接收文档。这使得万维网服务器的负担很重。
- 对于**持久连接**，指的是服务器在发送响应后**仍然保持这条连接**，使同一个客户和服务器可以继续在这条连接上传送后续的HTTP请求和响应报文。
-

5.2.3. HTTP的报文结构 *

- HTTP是 **面向文本** 的，因此报文中每个字段都是一些ASCII码串，且每个字段长度都是不确定的。分为 **请求报文**和 **响应报文**，两者都由三部分组成，区别在于开始行不同。
 - **开始行**：用于区别报文是**请求报文**还是**响应报文**。请求报文中称为**请求行**，响应报文中称为**状态行**。
 - 开始行三个字段之间以空格分隔，最后的 **CR** 和 **LF** 分别表示 **回车** 和 **换行**。
 - 请求报文的请求行包含三个字段：
 - **方法**：表示请求的类型，如 **GET**、**POST**、**HEAD** 等。
 - **URL**：表示请求的资源的位置。
 - **HTTP版本**：表示请求报文的HTTP版本。
 - 响应报文的**状态行**包含三个字段：
 - **HTTP版本**：表示响应报文的HTTP版本。
 - **状态码**：表示服务器对请求的处理结果。
 - **状态码的原因短语**：表示状态码的含义。
 - **首部行**：用来说明浏览器，服务器或报文主体的一些信息。每个首部行都有**首部字段名**和它的**值**。每一行结束都要有**回车和换行（CRLF）**。首部行结束时，还要有一个空行把首部行和后面的实体主体分开。
 - **实体主体**：请求报文一般不用，响应报文很少用。