

6-01 互联网的域名结构是怎样的?它与目前的电话网的号码结构有何异同之处?

互联网的域名结构是由标号序列组成的,各标号之间用点隔开,例如三级域名.二级域名.顶级域名,各标号分别代表不同级别的域名。

电话网的号码结构分为国家号、区号和本机号,与中国电话网的国家号结构(中国+86)相同。

不同之处在于电话号码中包含的是数字,而域名系统中包含的是字母和数字的组合。此外,域名系统中各标号之间用点隔开,而电话号码中各部分之间用短横线隔开。

6-02 域名系统的主要功能是什么?域名系统中的本地域名服务器、根域名服务器、顶级域名服务器以及权限域名服务器有何区别?

域名系统的主要功能是将域名转换为 IP 地址。本地域名服务器将本地主机的域名转换为 IP 地址;根域名服务器是最高级别的域名服务器,负责管理根域;顶级域名服务器负责管理顶级域;权限域名服务器能够将主机名转换为 IP 地址。

6-03 举例说明域名转换的过程。域名服务器中的高速缓存的作用是什么?

域名转换的过程称为域名解析。例如,当用户在浏览器地址栏中输入 `www.example.com` 时,域名解析过程如下:

- (1) 浏览器在本地的 DNS 缓存中查找是否有 `www.example.com` 的记录,如果有,则返回对应的 IP 地址;
- (2) 如果缓存中没有,则浏览器会向本地 DNS 服务器(即本地域名服务器)发送查询请求;
- (3) 如果本地 DNS 服务器也没有 `www.example.com` 的记录,则本地 DNS 服务器会向根域名服务器发送查询请求;
- (4) 根域名服务器返回一个 TLD(顶级域名)服务器的 IP 地址,本地 DNS 服务器将这个 IP 地址存入缓存;
- (5) 本地 DNS 服务器向 TLD 服务器发送查询请求,TLD 服务器返回负责管理 `example.com` 域的权限域名服务器的 IP 地址,本地 DNS 服务器将这个 IP 地址存入缓存;
- (6) 本地 DNS 服务器向权限域名服务器发送查询请求,权限域名服务器返回 `www.example.com` 的 IP 地址,本地 DNS 服务器将这个 IP 地址存入缓存;
- (7) 本地 DNS 服务器将 IP 地址返回给浏览器,浏览器将该 IP 地址解析为对应的网页内容并显示。

域名服务器中的高速缓存的作用是减少 DNS 查询的次数,提高查询速度。通过将已解析过的域名和对应的 IP 地址存储在缓存中,下次再查询相同域名时就可以直接从缓存中获取 IP 地址,而不需要再次进行繁琐的查询过程。

6-05 文件传送协议 FTP 的主要工作过程是怎样的?为什么说 FTP 是带外传送控制信息?主进程和从属进程各起什么作用?

文件传送协议 FTP 的主要工作过程包括以下三个阶段:

- (1) 建立连接:客户端发送连接请求到 FTP 服务器,服务器接受连接并建立连接。
- (2) 登录和身份验证:客户端发送用户名和密码进行登录和身份验证,服务器进行验证并返回结果。

(3) 文件传输：客户端发送文件传输请求，服务器接受请求并开始文件传输。

FTP 是带外传送控制信息，意味着控制信息和数据信息使用不同的通道进行传输。在 FTP 中，控制信息通过控制连接进行传输，而数据信息则通过数据连接进行传输。控制连接负责传输命令和响应，而数据连接则负责实际的数据传输。这种带外传送控制信息的方式可以避免数据连接出现的问题影响控制连接。

在 FTP 中，主进程负责处理客户端的请求，包括建立连接、发送和接收命令等。从属进程则负责处理数据传输，包括接收和发送数据等。主进程和从属进程协同工作，实现了 FTP 的文件传输功能。

6-08 解释以下名词。各英文缩写词的原文是什么？

WWW, URL, HTTP, HTML, CGI, 浏览器, 超文本, 超媒体, 超链, 页面, 活动文档, 搜索引擎。

- **WWW (World Wide Web)**

WWW 是 World Wide Web 的缩写，中文称为“万维网”。它是一个由许多互相链接的超文本组成的网络，这些超文本可以通过互联网进行访问。

- **URL (Uniform Resource Locator)**

URL 是 Uniform Resource Locator 的缩写，中文称为“统一资源定位符”。它是一种用于标识和定位互联网上资源的地址，例如网页、图片、视频等。

- **HTTP (Hypertext Transfer Protocol)**

HTTP 是 Hypertext Transfer Protocol 的缩写，中文称为“超文本传输协议”。它是一种用于在互联网上传输超文本的协议，例如 HTML 文档。

- **HTML (Hypertext Markup Language)**

HTML 是 Hypertext Markup Language 的缩写，中文称为“超文本标记语言”。它是一种用于编写和格式化网页的语言，可以将文本、图片、链接等元素组合在一起。

- **CGI (Common Gateway Interface)**

CGI 是 Common Gateway Interface 的缩写，中文称为“通用网关接口”。它是一种用于将网页服务器与外部应用程序或脚本连接起来的标准接口。通过 CGI，外部应用程序或脚本可以接收来自用户的请求，并将结果返回给网页服务器。

- **浏览器 (Browser)**

浏览器是一种用于访问和浏览互联网上的网页的工具。它可以将用户输入的 URL 转换为对应的网页，并显示在屏幕上。常见的浏览器有 Chrome、Firefox、Safari 和 Edge 等。

- **超文本 (Hypertext)**

超文本是一种将文本与其他元素（如图片、链接等）结合在一起的技术。通过超文本，用户可以通过点击链接从一个页面跳转到另一个页面，或者从一张图片跳转到另一张图片。

- **超媒体 (Hypermedia)**

超媒体是一种将文本、图片、音频、视频等多种媒体元素结合在一起的技术。它扩展了超文本的概念，允许用户在多个媒体元素之间进行导航和交互。

- **超链 (Hyperlink)**

超链是一种链接，它允许用户从一个页面跳转到另一个页面，或者从一个元素跳转到另一个元素。超链通常由链接文本或图像组成，用户可以通过点击它们来访问目标页面或元素。

- **页面 (Page)**

页面是指互联网上的一个单独的文档或界面。它通常由 HTML 文档组成，并包含文本、图片、链接等多种元素。用户可以在浏览器中打开和浏览页面。

- **活动文档 (Active Document)**

活动文档是指一种动态生成的文档，它可以响应用户的操作或事件。通过活动文档，用户可以与页面进行交互，例如填写表单、切换选项卡等。活动文档通常由脚本语言编写，例如 JavaScript。

- **搜索引擎 (Search Engine)**

搜索引擎是一种用于搜索互联网上的信息的工具。它通过爬取网页并建立索引来存储网页的信息，然后根据用户的搜索查询返回相关的结果。常见的搜索引擎有 Google、Baidu、Bing 等。

6-10 假定要从已知的 URL 获得一个万维网文档。若该万维网服务器的 IP 地址开始时并不知道。试问:除 HTTP 外，还需要什么应用层协议和运输层协议？

除了 HTTP 协议外，如果要从已知的 URL 获取一个万维网文档，还需要使用以下应用层协议和运输层协议：

应用层协议：**DNS (Domain Name System)**：用于将 URL 转换为对应的 IP 地址。通过 DNS 查询，可以将 URL 中的域名解析为相应的服务器 IP 地址。

运输层协议：**TCP (Transmission Control Protocol)**：一种可靠的传输协议，用于在网络上传输数据。TCP 提供了一种可靠的数据传输服务，通过确认、超时和重传等机制，确保数据的完整性和可靠性。

因此，为了从已知的 URL 获取一个万维网文档，需要使用 HTTP、DNS 和 TCP 协议。首先，通过 DNS 将 URL 解析为相应的服务器 IP 地址。然后，使用 TCP 协议建立与该服务器的连接，并使用 HTTP 协议发送请求并获取文档。

6-13 浏览器同时打开多个 TCP 连接进行浏览的优缺点如何?请说明理由。

- **优点：**

提高页面加载速度：浏览器使用多个 TCP 连接时，每个连接可以同时传输数据，加快了页面加载速度。这是因为每个 TCP 连接都可以独立地传输数据，减少了等待时间，从而提高了加载速度。

多任务处理：浏览器可以在多个 TCP 连接上同时处理多个请求，有助于提高浏览器性能和增加用户体验。这意味着用户可以在多个网页之间进行切换，同时进行多项操作，提高了浏览器的响应速度和效率。

提高可靠性：如果一个 TCP 连接中断或失败，其他连接可以仍然正常工作，在保持网站可靠性方面具有优势。由于多个 TCP 连接可以同时工作，即使其中一个连接出现问题，浏览器仍然可以通过其他连接获取所需的数据。

- **缺点：**

服务器负载增加：多个 TCP 连接意味着服务器需要处理更多的请求，这可能会增加服务器的负载。过多的请求可能会导致服务器过载，影响到网站的性能和响应速度。

建立 TCP 连接时间增加：TCP 连接建立需要一定时间，多个连接会导致更多的连接建立请求，从而增加了页面加载时间。这是因为每个 TCP 连接的建立都需要进行握手等过程，过多的连接会延长这个时间。

消耗更多计算和网络资源：打开过多的 TCP 连接也会消耗更多的计算资源

和网络资源。每个连接都需要浏览器和服务器进行数据交换和处理，因此过多的连接会加重浏览器的负担，消耗更多的 CPU 和内存资源。

6-18 一个万维网网点有 1000 万个页面，平均每个页面有 10 个超链。读取一个页面平均要 100 ms。问要检索整个网点所需的最少时间。

假设每个页面有 n 个超链，读取一个页面的时间为 t ms。

根据题目，我们可以建立以下模型：

每个页面有 10 个超链，所以总超链数量是 $10 \times 1000 \text{ 万} = 1 \text{ 亿}$ 个超链。

读取一个页面的时间是 100 ms。

但是，我们不需要读取所有的页面。我们只需要读取那些包含超链的页面。

所以，最少需要读取的页面数量是总超链数量除以每个页面的超链数量，即 $1 \text{ 亿} / 10 = 1 \text{ 千万}$ 个页面。

用数学模型，我们可以表示为：

最少需要读取的页面数量 = 总超链数量 / 每个页面的超链数量

最少需要的时间 = 最少需要读取的页面数量 \times 读取一个页面的时间

现在我们要来计算最少需要的时间。

计算结果为：最少需要读取的页面数量是 10000000 个，最少需要的时间是 1000000000 ms。

所以，检索整个网点所需的最少时间是：1000000000 ms。