## 网络基础知识

|  |
| --- |
| 12《Java语言程序设计及实训》绘图  图10‑1 TCP/IP协议簇 |

### TCP/IP协议

TCP/IP协议是网络通信技术事实上的标准，Java的网络编程主要是通过TCP/IP协议实现计算机之间的通信。

TCP/IP协议是一组协议的总称，它包括了HTTP、SMTP、TCP、UDP、IP等协议，所以也将其称为TCP/IP协议簇。在TCP/IP的体系结构中，这些协议分为四个层次（见图10‑1），从下向上依次为网络接口层、互联层、传输层和应用层。

* 网络接口层：包括网络硬件设备，如双绞线、光纤、无线电波、网卡、交换机等设备，以及网络协议，如以太网、令牌环网等协议。该层提供了网络传输的基础。
* 互联层：由多个协议组成，这些协议互相合作，共同完成互联层的功能，其中最主要的协议是IP协议。
* 传输层：有多种协议，每种协议提供不同的服务。主要的协议是TCP和UDP协议。
* 应用层：有非常多种协议，几乎包括了互联网的所有应用，例如用于WWW服务的HTTP协议，用于邮件传输的SMTP协议和POP3协议，用于文件传输的FTP协议等。

### IP协议

IP协议（互联网协议，Internet Protocol）提供无连接的传输服务，它的功能是把数据发送到目标主机。IP协议传输的数据单元称为数据包，每个IP数据包都包含了数据的源IP地址和目的IP地址。

#### IP地址

IP地址是一台主机（即网络中的计算机）在网络中的唯一标识，它是全球唯一的。IP地址是一个32位的数字，一般以点分十进制的形式表示，如210.28.144.22。IP地址由IANA（Internet Assigned Numbers Authority）组织统一负责管理分配，从而保证全球所有主机的IP地址不会重复。

每台主机都必须正确配置IP地址、子网掩码、默认网关地址、DNS服务器地址等有关参数后，才能在网络上与其他主机通信。

#### 回环地址

有一个特殊的地址127.0.0.1，被称为回环地址，它永远回环到本机，也就是说，访问127.0.0.1即访问本机，是本机地址的代名词。回环地址的主要用途是测试。

#### 私有地址和公网地址

私有地址是指内部网络（局域网）上主机的IP地址，公网地址是指在因特网上主机的IP地址，公网地址才具有全球唯一的特点。TCP/IP规定将下列范围的IP地址保留用作私有地址：10.0.0.0 ~ 10.255.255.255、172.16.0.0 ~ 172.31.255.255、192.168.0.0 ~ 192.168.255.255。也就是说，这三个范围内的地址不会在因特网上被分配，它们仅在局域网内部使用。私有IP地址在全球范围内不具有唯一性，仅在局域网内具有唯一性。

### TCP和UDP协议

传输层主要的二种协议是TCP协议（传输控制协议，Transmission Control Protocol）和UDP协议（用户数据报协议，User Datagram Protocol）。

#### TCP协议

TCP协议是一种面向连接的可靠的协议。它通过序列号确认以及包重发机制，提供可靠的数据传输。用于可靠性要求较高、传输大量数据的场合。

TCP协议把数据分割成适当长度的分段，为每个分段加上序列号、端口号（见后面讨论）等信息，这样的分段称为TCP报文，然后把报文交给IP层，由IP协议通过网络将报文传输给接收端的IP层，再交给TCP协议。接收端的TCP协议负责将多个TCP报文按顺序恢复（合并）成原来的数据。

基于TCP协议的应用层协议有FTP、SSH、Telnet、SMTP、HTTP和POP3等。

#### UDP协议

UDP是一种不可靠的、无连接的服务，它直接向网络发送数据给一台主机或多台主机，而不需要确认数据是否送达。主要用于一次性传输少量数据，不仅用于点对点的通信，也用于一对多（如组播）的通信。

基于UDP协议的应用层协议有SNMP和DNS等。

#### 端口

在一台主机上可以运行多个网络应用程序（通信进程），网络通信的对象准确地说不是主机，而是主机上运行的通信进程。这时用IP地址是无法标识这些通信进程的。

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 表10‑1 常用的默认端口号   |  |  |  | | --- | --- | --- | | 传输层  服务类型 | 应用层  协议 | 端口号 | | 基于TCP | FTP | 21 | | SSH | 22 | | Telnet | 23 | | SMTP | 25 | | HTTP | 80 | | POP3 | 110 | | 基于UDP | SNMP | 161 | | DNS | 53 | |

端口（Port）被用来标识运行在一台主机上的多个通信进程。端口号是一个16位的正整数，其中0~1023已被系统使用，分配给常用的网络服务程序，称为默认端口号（表10‑1）。用户能够使用的端口号是1024~65535，网络上的每一种应用程序都需要一个端口号。

客户进程的端口一般由所在主机的操作系统动态分配，当客户进程要求与一个服务器进程进行TCP或UDP 连接，操作系统为客户进程随机分配一个还未被占用的端口，当客户进程与服务器进程断开连接，这个端口就被释放。

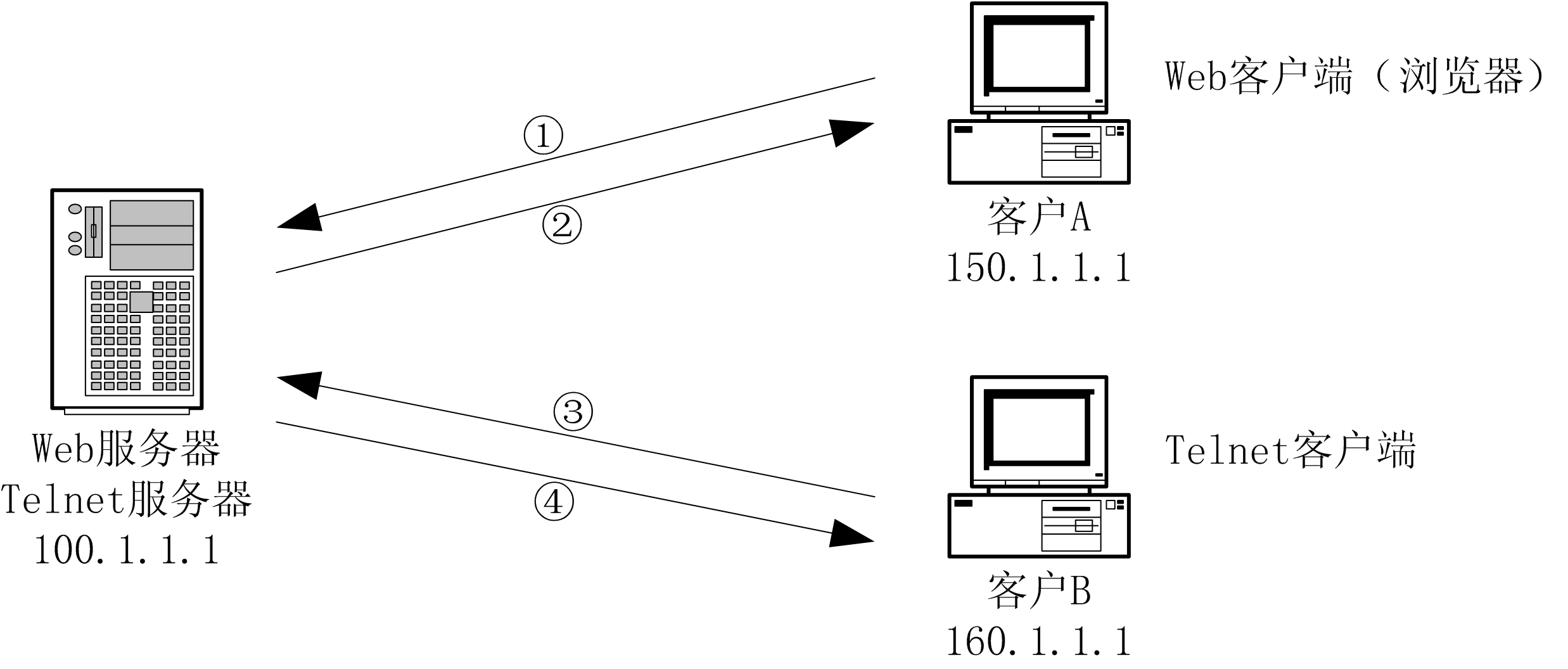


图10‑2 客户机与服务器的通信

为说明服务器端和客户端的端口号，以图10‑2为例进行说明。图中IP地址为100.1.1.1的主机是一台服务器，安装了二个网络程序：Web服务器和Telnet服务器，这两种服务的默认端口号是80和23。客户A和客户B都只有一个网络程序，将分别访问服务器的这二个服务，这时在网络上传输的TCP报文有四种，见表10‑2。

表10‑2客户机与服务器的TCP报文

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| TCP报文 | 源IP地址 | 源端口号 | 目的IP地址 | 目的端口号 |
| 客户A发往服务器（见图10‑2①） | 150.1.1.1 | 2035 | 100.1.1.1 | 80 |
| 服务器发往客户A（见图10‑2②） | 100.1.1.1 | 80 | 150.1.1.1 | 2035 |
| 客户B发往服务器（见图10‑2③） | 160.1.1.1 | 4178 | 100.1.1.1 | 23 |
| 服务器发往客户B（见图10‑2④） | 100.1.1.1 | 23 | 160.1.1.1 | 4178 |

通信是由客户端发起的，表10‑2中的二个端口号80和23是服务器上二种服务的默认端口号，其他二个端口号2035和4178分别是客户A和客户B的网络程序在发起通信时随机分配的。服务器接收到端口号为80的报文，将交由Web服务程序进行处理，接收到端口号为23的报文则由Telnet服务程序进行处理。这二种服务的响应报文发往对应客户机的相应端口。

TCP 和UDP 都用端口来标识进程，二者之间是独立的，因此允许存在取值相同的TCP端口与UDP端口。

#### 套接字

套接字（Socket）由IP地址和端口号组成，它唯一标识了网络上运行的通信进程。因此定义了网络上运行的两个不同主机的进程间进行双向通信的端点，用于建立两个不同应用程序之间通过网络进行通信的信道。

根据传输层协议的不同，Socket分为二种：TCP Socket和UDP Socket。

### 应用层协议

应用层协议的种类有非常多，常见的如表10‑1所示。

本章编写的Java网络程序就是应用层的应用程序，只是没有被标准化，不能称之为应用层协议。

### 域名

#### 域名

由于数字形式的IP地址难于记忆，因此，主机地址常用域名表示，例如本书作者博客的地址是www.ngweb.org。域名与IP地址的转换由域名系统（DNS）来完成，只要正确地设置了DNS服务器的地址，主机就能使用域名，而不需使用难记的IP地址来访问网络。

域名必须注册后才能使用，从而保证域名的全球唯一性。每台主机都预置了一个名字localhost，它指向127.0.0.1，也就是主机本身。与回环地址127.0.0.1的作用一样，localhost常常用于测试。

#### URL

统一资源定位符（(Uniform Resource Location，URL）表示网络上资源的位置，通常所说的网址就是一种URL，由协议名称、IP地址或域名等组成，形式如下：

协议名称://主机地址:端口号/文件名[#引用]

* 协议名称：这是一个应用层协议的名称，如HTTP、FTP等。
* 主机地址：可以是主机的IP地址，也可以是域名。
* 端口号[可选]：如果省略，则使用协议的默认端口号，如HTTP协议使用80。
* 文件名[可选]：资源的文件名，由路径和查询字符串二部分组成。
* 引用[可选]：即锚点。

例如http://www.ngweb.org/wp?id=4表示该资源使用HTTP协议，主机地址是www.ngweb.org，端口号是80（HTTP的默认端口号，省略），文件名是wp?id=4，缺少引用一项。

### C/S模式与B/S模式

#### C/S模式

C/S（客户机/服务器，Client/Server）是一种网络应用模式。客户进程向服务器进程发出要求某种服务的请求，服务器进程响应该请求。

在这种模式下，需要分别编写客户机程序和服务器程序，每台客户机都需要安装客户端程序。本章讨论的程序属于C/S模式。

#### B/S模式

B/S（浏览器/服务器，Browser/Server）是另一种网络应用模式。客户利用浏览器向服务器进程发出要求某种服务的请求，服务器进程响应该请求。

在这种模式下，只需编写服务器程序，任何客户机都不需安装客户端程序，客户机通过浏览器访问服务器。B/S模式的应用程序开发属于Java EE的范畴，如JSP、Servlet等。