

在科技飞速发展的今天,网络的普及已经惠及你我,对网络相关知识的了解和掌握必不可少。本章介绍网络与信息系统的基础知识,主要包括三部分的内容:网络基础知识、Internet 概述、信息系统概述。其中网络基础知识主要介绍网络的定义与分类、网络的拓扑结构、网络的功能;Internet 概述主要介绍 IP 地址、域名、网络协议简介、URL;信息系统概述主要介绍信息系统的分类、信息系统的功能以及网络信息系统的实例。通过本章的学习,可以更好地了解 and 掌握计算机网络及信息系统的基础知识。

3.1 网络基础知识

网络是一个复杂的人或物的互联系统。人们周围时刻存在着各种网络,如电话网、交通网等;即使人体内部也是由许许多多的网络系统组成的,如神经系统、消化系统等。在网络盛行的今天,使用计算机网络是现代人必须掌握的一项基本技能。本节主要介绍什么是计算机网络、网络的结构与分类、网络的特征与功能。

3.1.1 网络的定义与分类

1. 计算机网络的定义

网络原来是指用一个巨大的虚拟画面,把所有东西连接起来。在计算机领域中,网络就是用物理链路将各个孤立的工作站或主机相连在一起,组成数据链路,从而达到资源共享和通信的目的。凡是将地理位置不同,并具有独立功能的多个计算机系统通过通信设备和线路连接起来,且以功能完善的网络软件(网络协议、信息交换方式及网络操作系统等)实现网络资源共享的系统,都可称为计算机网络。

2. 网络的分类

按照网络覆盖的地理范围大小,可以将网络分为局域网、城域网和广域网三种类型。这也是网络最常见的分类方法。

1) 局域网

局域网(Local Area Network, LAN)是将较小地理区域内的计算机或数据终端设备连接在一起的通信网络。局域网覆盖的地理范围比较小,一般在几十米到几千米之间。它常用于组建一个办公室、一栋楼、一个楼群、一个校园或一个企业的计算机网络。局域网的结构如图 3.1.1 所示。

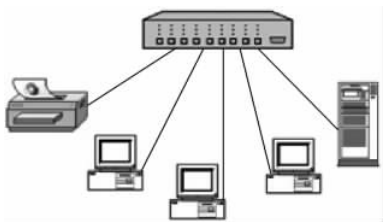


图 3.1.1 局域网的结构

2) 城域网

城域网(Metropolitan Area Network,MAN)的覆盖范围介于局域网和广域网之间,一般为几千米至几十千米。城域网的覆盖范围在一个城市内,它将位于一个城市之内不同地点的多个计算机局域网连接起来实现资源共享。城域网的结构如图 3.1.2 所示。

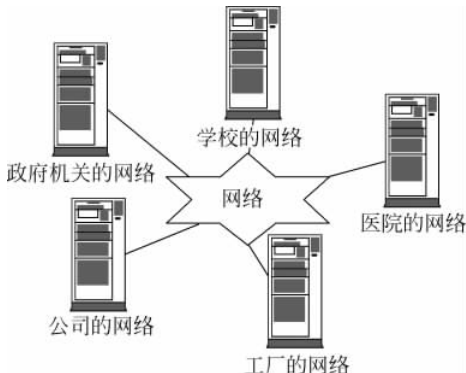


图 3.1.2 城域网的结构

3) 广域网

广域网(Wide Area Network,WAN)是在一个广阔的地理区域内进行数据、语音、图像信息传输的计算机网络。由于远距离数据传输的带宽有限,因此广域网的数据传输速率比局域网要慢得多。广域网可以覆盖一个城市、一个国家甚至于全球。广域网的结构如图 3.1.3 所示。

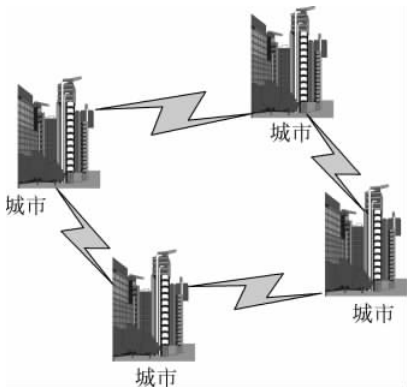


图 3.1.3 广域网的结构

网络除了按照覆盖范围来分类外,还可以有一些其他的分类方式,总结如下。

按交换方式分类:线路交换网络(Circuit Switching)、报文交换网络(Message Switching)和分组交换网络(Packet Switching)。

按网络拓扑结构分类:总线网络、环状网络、星状网络、树状网络和网状网络。

按通信介质分类:双绞线网、同轴电缆网、光纤网和卫星网等。

按传输带宽分类:基带网和宽带网。

按使用范围分类:公用网和专用网。

按速率分类:高速网、中速网和低速网。

按通信传播方式分类:广播式和点到点式等。

3.1.2 网络的拓扑结构

拓扑(Topology)是指将各种物体的位置表示成抽象位置。在网络中,拓扑用于描述网络的安排和配置,包括各种结点和结点之间的相互关系。值得注意的是,拓扑并不关心事物的细节,也不在乎事物相互之间的比例关系,而是将事物之间的相互关系通过图表示出来。网络中的计算机等设备要实现互连,就需要以一定的结构方式进行连接,这种连接方式就叫作“拓扑结构”,通俗地讲就是这些网络设备是如何连接在一起的。

计算机网络的拓扑结构主要有总线拓扑、星状拓扑、环状拓扑、树状拓扑、网状拓扑和混合型拓扑。

1. 总线拓扑结构

总线拓扑结构采用一条单根的通信线路(总线)作为公共的传输通道,所有结点都通过相应的接口直接连接到总线上,并通过总线进行数据传输。由于所有结点共享同一条公共通道,所以在任何时候只允许一个站点发送数据。当一个结点发送数据,并在总线上传播时,数据可以被总线上的其他所有结点接收。各站点在接收数据后,分析目的物理地址再决定是否接收该数据。

粗、细同轴电缆以太网就是这种结构的典型代表。总线拓扑结构如图 3.1.4 所示。

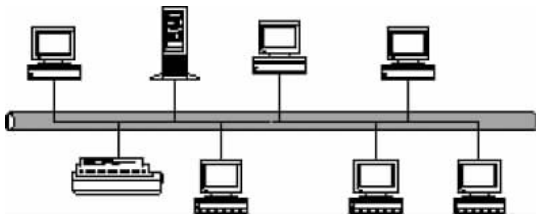


图 3.1.4 总线拓扑结构

1) 总线拓扑结构的优点

- (1) 总线拓扑结构所需要的电缆数量少,线缆长度短,易于布线和维护。
- (2) 总线拓扑结构简单,有较高的可靠性,传输速率高,可达 100Mb/s。
- (3) 总线拓扑结构组网容易,易于扩充,增加或减少用户比较方便。
- (4) 总线拓扑结构信道利用率高,多个结点共用一条传输信道。

2) 总线拓扑结构的缺点

- (1) 总线拓扑结构的传输距离有限,通信范围受到限制。
- (2) 故障诊断和隔离较困难。
- (3) 分布式协议不能保证信息的及时传送,不具有实时功能。
- (4) 站点要有媒体访问控制功能,从而增加了站点的硬件和软件开销。

2. 星状拓扑结构

星状拓扑结构的每个结点都由一条点对点链路与中心结点(公用中心交换设备,如交换机、集线器等)相连,如图 3.1.5 所示。星状拓扑结构中的一个结点如果向另一个结点发送数据,首先将数据发送到中央设备,然后由中央设备将数据转发到目标结点;信息的传输是通过中心结点的存储转发技术实现,并且只能通过中心结点与其他结点通信。星状拓扑结构是局域网中最常用的拓扑结构。

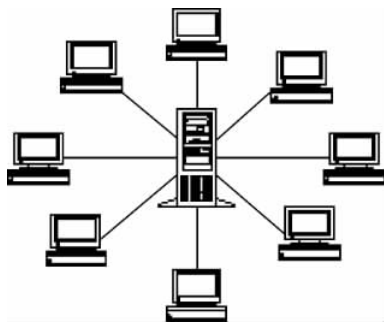


图 3.1.5 星状拓扑结构

1) 星状拓扑结构的优点

- (1) 结构简单,连接方便,管理和维护都相对容易,而且扩展性强。
- (2) 网络延迟时间较小,传输误差低。
- (3) 同一网段内支持多种传输介质,除非中央结点故障,否则网络不会轻易瘫痪。
- (4) 每个结点直接连到中央结点,结点故障容易检测和隔离。

2) 星状拓扑结构的缺点

- (1) 安装和维护的费用较高。
- (2) 共享资源的能力较差。
- (3) 一条通信线路只被该线路上中央结点和边缘结点使用,通信线路利用率不高。
- (4) 对中央结点要求相当高,一旦中央结点出现故障,则整个网络将瘫痪。

3. 环状拓扑结构

环状拓扑结构的各个网络结点通过环接口连在一条首尾相接的闭合环状通信线路中,如图 3.1.6 所示。其每个结点设备只能与它相邻的一个或两个结点设备直接通信。如果要与网络中的其他结点通信,数据需要依次经过两个通信结点之间的每个设备。环状拓扑结构既可以是单向的也可以是双向的。

环状拓扑结构有两种类型,即单环结构和双环结构。令牌环(Token Ring)是单环结构的典型代表,光纤分布式数据接口(FDDI)是双环结构的典型代表。

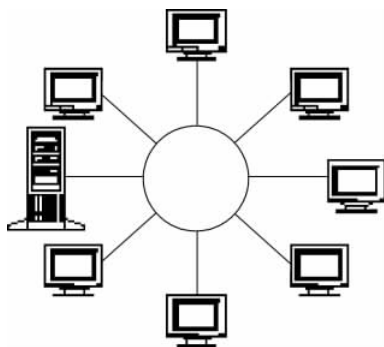


图 3.1.6 环状拓扑结构

1) 环状拓扑结构的优点

- (1) 环状拓扑网络所需的电缆长度和总线拓扑网络相似,电缆长度短。
- (2) 增加或减少工作站容易,仅需要简单的连接操作。
- (3) 可使用光纤,光纤的传输速率很高,十分适合环状拓扑结构的单方向传输。

2) 环状拓扑结构的缺点

- (1) 结点的故障会引起全网故障。这是因为环上的数据传输要通过接在环上的每一个结点,一旦环中某一结点发生故障就会引起全网的故障。
- (2) 故障检测困难。因为不是集中控制,故障检测需要在网络中各个结点进行,因此就不是很容易。
- (3) 信道利用率低。环状拓扑结构的媒体访问控制协议都采用令牌传递的方式,当网络负载很轻时,信道利用率就比较低。

4. 树状拓扑结构

树状拓扑结构可以由多级星状结构组成的,这种多级星状结构自上而下呈三角形分布,就像一棵树一样,最顶端的枝叶少些,中间的树叶多些,最下面的枝叶最多。树的最下端相当于网络中的边缘层,树的中间部分相当于网络中的汇聚层,树的顶端则相当于网络中的核心层。它采用分级的集中控制方式,其传输介质可有多条分支,但不形成闭合回路,每条通信线路都必须支持双向传输。树状拓扑结构如图 3.1.7 所示。

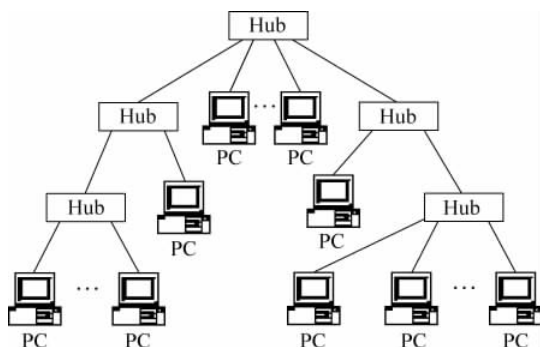


图 3.1.7 树状拓扑结构

1) 树状拓扑结构的优点

(1) 树状拓扑结构易于扩展,可以延伸出很多分支和子分支,这些新结点和新分支都能容易地加入网内。

(2) 故障隔离比较容易。如果某一分支的结点或线路发生故障,则很容易将故障分支与整个系统隔离开来。

2) 树状拓扑结构的缺点

各个结点对根的依赖性太大,如果根发生故障,则全网不能正常工作。从这一点来看,树状拓扑结构的可靠性有点类似于星状拓扑结构。

5. 网状拓扑结构

网状拓扑结构是指将各网络结点与通信线路连接成不规则的形状,每个结点至少与其他两个结点相连,或者说每个结点至少要有两条链路与其他结点相连。大型互联网一般都采用这种结构,如我国的教育科研网 CERNET、Internet 的主干网都采用网状拓扑结构。网状拓扑结构如图 3.1.8 所示。

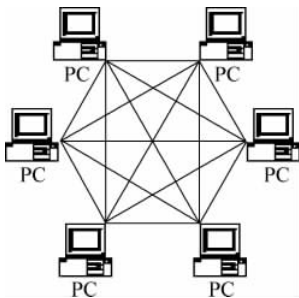


图 3.1.8 网状拓扑结构

1) 网状拓扑结构的优点

(1) 结点间路径多,碰撞和阻塞减少。

(2) 局部故障不影响整个网络,可靠性高。

2) 网状拓扑结构的缺点

(1) 网络关系复杂,建立网络比较难,不容易扩充。

(2) 网络控制机制复杂,必须采用路由算法和流量控制机制。

6. 混合型拓扑结构

混合型拓扑结构是将两种单一拓扑结构混合起来,取两者的优点构成的拓扑结构。例如,有一种混合型拓扑结构就是星状拓扑结构和总线拓扑结构混合成的“星-总”拓扑结构。用一条或多条总线把多组设备连接起来,相连的每组设备呈星状分布,采用这种拓扑结构,用户很容易配置网络设备。总线拓扑结构采用同轴电缆,星状拓扑结构可采用双绞线。混合型拓扑结构如图 3.1.9 所示。

1) 混合型拓扑结构的优点

(1) 故障诊断和隔离比较方便。一旦网络发生故障,只要诊断出哪个网络设备有故障,将该网络设备和全网隔离即可。

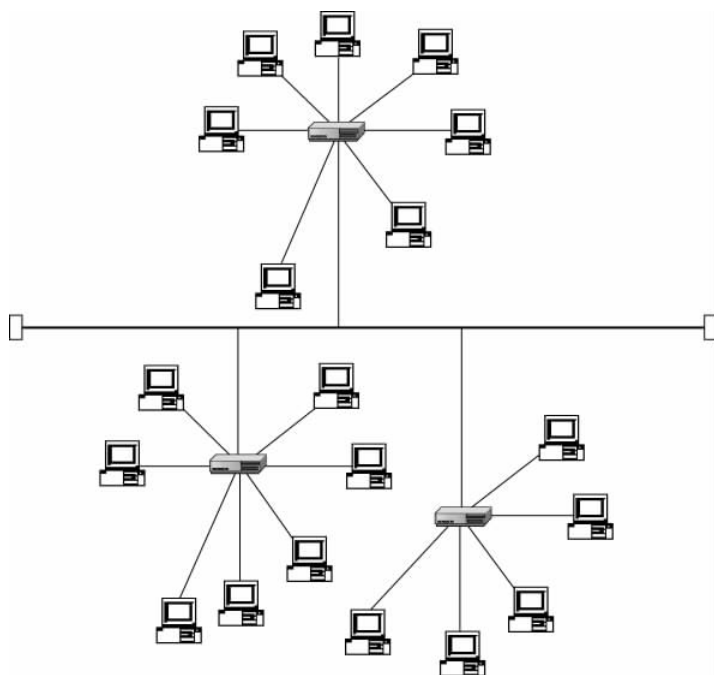


图 3.1.9 混合型拓扑结构

(2) 易于扩展。要扩展用户时,可以加入新的网络设备,也可在设计时,在每个网络设备中留出一些备用的可插入新站点的连接口。

(3) 安装方便。网络的主链路只要连通汇聚层设备,然后再通过分支链路,连通汇聚层设备和接入层设备。

2) 混合型拓扑结构的缺点

网络建设成本比较高,需要选用智能网络设备实现网络故障自动诊断和故障结点的隔离。汇聚层设备到接入层设备的线缆安装长度会增加较多。

3.1.3 网络的功能

计算机网络的功能主要包括实现资源共享,实现数据信息的快速传递(即数据通信),提供负载均衡与分布式处理能力,提高可靠性。

1. 资源共享

资源共享是计算机网络最有吸引力的功能。由于计算机的许多资源成本昂贵,如大型数据库、海量存储器、特殊外部设备等,资源共享使网上用户能部分或全部地享用这些成本昂贵的资源,使网络中各地区的资源互通有无、分工协作,从而大大提高网络资源的利用率。网络资源具体来说包括以下几种。

(1) 硬件资源: 主要包括大型主机、大容量磁盘、光盘库、打印机、网络通信设备、通信线路和服务器硬件等。

(2) 软件资源: 主要包括网络操作系统、数据库管理系统、网络管理系统、应用软件、开发工具和服务器软件等。

(3) 数据资源：主要包括数据文件、数据库和光磁盘所保存的各种数据。数据是网络中最重要资源。

2. 数据通信

通信即在计算机之间传输信息,是计算机网络最基本的功能之一。通过计算机网络使不同地区的用户可以快速和准确地相互传输信息,这些信息包括数据、文本、图形、动画、声音和视频等。很多网络用户通过收发 E-mail、VOD(视频点播)、IP 电话、即时通信等方式进行数据通信。

3. 分布式处理与负载均衡

计算机网络中,用户可以根据需要合理选择网络内的资源,以便就近处理。例如:用户在异地通过远程登录可直接进入单位内部网络;对于像人口普查、售火车票这样需要综合处理的大型作业,通过一定的算法将负载比较大的作业分解,交给多台计算机进行分布式处理,起到负载均衡的作用,这样就能提高处理速度,充分发挥利用率,提高效率。

4. 提高可靠性

提高可靠性表现在计算机网络中的多台计算机可以通过网络彼此间相互备用,一旦某台计算机出现故障,其任务可由其他计算机代其处理。这样可以避免在单机情况下,一台计算机发生故障引起整个系统瘫痪的现象,从而提高网络系统的可靠性。

3.2 Internet 概述

Internet(因特网)是一组全球信息资源的总汇。Internet 以相互交流信息资源为目的,基于一些共同的协议,并通过许多路由器和公共互联网相连接,它是一个信息资源和资源共享的集合。Internet 是当今世界上最大的、最具影响力的国际性计算机综合网络。

我国在 1994 年 4 月正式接入因特网。1994 年 5 月中国科学院计算机网络信息中心接管中国国家顶级域名(CN)。1994—1996 年先后建成中国的四大网络,即有资格设置独立国际信息出口的 Internet 服务机构,具体如下。

(1) 中国教育和科研计算机网(China Education and Research Network,CERNET):是由政府资金启动的全国范围教育与学术网络。

(2) 中国科技网:主要为中科院在全国的研究所和其他相关研究机构提供科学数据库和超级计算资源。

(3) 中国公用计算机互联网:中国电信经营和管理的中国公用 Internet 网。

(4) 中国金桥信息网:国家公用经济信息通信网,由吉通通信有限责任公司负责建设、运营和管理。

3.2.1 IP 地址

1. IP 地址的概念

IP 是英文 Internet Protocol 的缩写,意思是“网络之间互连的协议”,即为计算机网

络相互连接进行通信而设计的协议。在因特网中,它是使连接到网上的所有网络设备(如计算机、交换机)实现相互通信的一套规则,规定了网络设备在因特网上进行通信时应当遵守的规则。任何厂家生产的网络设备,只要遵守 IP 协议就可以与因特网互联互通。正是因为有了 IP 协议,因特网才得以迅速地发展成为世界上最大的、开放的计算机通信网络。因此,IP 协议也可以叫作“因特网协议”。现行的 IP 地址分配格式,也称为 IPv4 格式。

2. IPv4 简介

1) IPv4 的地址结构

IPv4 地址是一个 32 位(4 字节)的二进制数字,为了方便记忆,采用“点分十进制”法来表示。即将 4 个字节的二进制数值转换成 4 个十进制数值,每个数值的取值为 0~255,数值之间用点号“.”隔开。如 192.168.0.1,IPv4 地址的二进制表示法与点分十进制表示法的比较如图 3.2.1 所示。

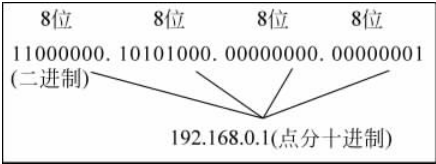


图 3.2.1 IPv4 地址的二进制表示法与点分十进制表示法的比较

IPv4 地址是由网络号(ID)和主机号(ID)两部分组成的。网络号(ID)标识某个逻辑网络,类似电话号码中的区号;主机号(ID)标识该网络上的某台主机,类似电话号码中的本机号码。这种分级的标识方式可以方便主机地址的分配和管理。如将 IPv4 地址 192.168.0.1 分成两部分,其中网络号(ID)占 24 位,主机号(ID)占 8 位。

网络号(ID)用于标识 IP 地址是否在同一个网段,如果网络号(ID)不同,则主机间不能直接通信,如果通信就需要路由器连接;如果网络号(ID)相同,则主机间可直接通信,不需要路由。在同一个网段内的计算机的网络号(ID)相同,而主机号(ID)不同。

2) IPv4 的地址分类

Internet 委员会定义了 5 种 IPv4 地址类型以适合不同容量的网络,即 A~E 类。其中 A、B、C 三类由 Internet NIC 在全球范围内统一分配,D、E 类为特殊地址。A、B、C 三类 IPv4 地址如表 3.2.1 所示。

表 3.2.1 A、B、C 三类 IPv4 地址

类别	最大可用网络数	IP 地址范围	最大主机数	私有 IP 地址范围
A	$126(2^7-2)$	0.0.0.0~127.255.255.255	16 777 214	10.0.0.0~10.255.255.255
B	$16\,384(2^{14})$	128.0.0.0~191.255.255.255	65 534	172.16.0.0~172.31.255.255
C	$2\,097\,152(2^{21})$	192.0.0.0~223.255.255.255	254	192.168.0.0~192.168.255.255

D 类 IPv4 地址在历史上被叫作多播地址(Multicast Address),即组播地址。在以以太网中,多播地址命名为一组在这个网络中接收一个分组的站点。多播地址的最高位必须

是 1110, 范围是 224. 0. 0. 0~239. 255. 255. 255。

E 类 IPv4 地址是以 1111 开始, 为将来使用保留。它的第一个字节的范围是 240~255, 主要用于 Internet 试验和开发。

除了上述的五类 IPv4 地址外, 还有一些特殊的 IPv4 地址。比如, 每一个字节都为 0 的地址 (0. 0. 0. 0) 对应于当前主机; IPv4 地址中的每一个字节都为 1 的 IPv4 地址 (255. 255. 255. 255) 是当前子网的广播地址; IPv4 地址不能以十进制数 127 作为开头, 该类地址 127. 0. 0. 1~127. 255. 255. 255 用于回路测试, 如 127. 0. 0. 1 可以代表本机 IPv4 地址, 用 [http://127. 0. 0. 1](http://127.0.0.1) 就可以测试本机中配置的 Web 服务器。网络号 (ID) 的第一个 8 位组也不能全部设置为 0, 全 0 表示本地网络。

3. IPv6 简介

IPv4 具有相当强盛的生命力, 易于实现且互操作性良好, 经受住了从早期小规模互联网络扩展到如今全球范围 Internet 应用的考验。但是随着 Internet 呈指数级飞速发展, IPv4 地址空间几近耗竭。随着连网设备的急剧增加, IPv4 公共地址总有一天会完全耗尽。除此之外, 新技术的出现也对 IP 协议提出了更多的需求。在此基础之上, 科研人员研发出了 IPv6。

1) IPv6 的定义

IPv6 是 Internet Protocol Version 6 的缩写, 其中 Internet Protocol 译为“互联网协议”。IPv6 是 IETF 设计的用于替代现行 IPv4 的下一代 IP 协议, 号称可以为全世界的每一粒沙子都编上一个网址。

IPv4 最大的问题在于网络地址资源有限, 严重制约了互联网的应用和发展。IPv6 的使用, 不仅能解决网络地址资源数量的问题, 而且也解决了多种接入设备连入互联网的障碍。

2) IPv6 的优点

IPv6 采用 128 位地址结构, 提供充足的地址空间。128 位地址可以包含约 43 亿 $\times$ 43 亿 $\times$ 43 亿个地址结点。

(1) 层次化的网络结构, 提高了路由效率。IPv6 的地址分配一开始就遵循聚类 (Aggregation) 原则, 这使得路由器能在路由表中用一条记录表示一片子网, 大大减小了路由器中路由表的长度, 提高了路由器转发数据包的速度。

(2) 报文头简洁、灵活, 效率更高, 易于扩展。IPv6 使用新的头部格式, 其选项与基本头部分开, 如果需要, 可将选项插入到基本头部与上层数据之间。这就简化并加速了路由选择过程, 因为大多数的选项不需要由路由选择。

(3) 支持自动配置, 即插即用。它可以通过地址自动配置方式使主机发现网络并获取 IPv6 地址。

(4) IPv6 具有更高的安全性。在使用 IPv6 的网络中, 用户可以对网络层的数据进行加密, 并对 IP 报文进行校验, IPv6 中的加密与鉴别选项提供了分组的保密性与完整性。

(5) 新增流标签功能, 更利于支持 QoS。IPv6 报文头中新增了流标签域, 这使得网络上的多媒体应用为服务质量 (Quality of Service, QoS) 控制提供了良好的网络平台。

3) IPv6 地址的结构

IPv6 的地址长度为 128 位,是 IPv4 地址长度的 4 倍。IPv6 采用十六进制表示,共有三种表示方法。

(1) 冒分十六进制表示法。

格式为 X:X:X:X:X:X:X:X,其中每个 X 表示地址中的 16 位,以十六进制表示。例如一个用 IPv6 表示的地址可以写为 ABCD:EF01:2345:6789:ABCD:EF01:2345:6789。在实际表示中,每个 X 的前导 0 是可以省略的。例如 2001:0DB8:0000:0023:0008:0800:200C:417A 可以表示成 2001:DB8:0:23:8:800:200C:417A。

(2) 0 位压缩表示法。

在某些情况下,一个 IPv6 地址中间可能包含很长的一段 0,可以把连续的一段 0 压缩为“::”。但为保证地址解析的唯一性,地址中“::”只能出现一次。例如 FF01:0:0:0:0:0:0:1101 表示为 FF01::1101; 0:0:0:0:0:0:0:1 表示为 ::1; 0:0:0:0:0:0:0:0 表示为 ::。

(3) 内嵌 IPv4 地址表示法。

为了更好地实现 IPv4 和 IPv6 的互通,可以将 IPv4 地址嵌入到 IPv6 的地址中,此时地址常表示为 X:X:X:X:X:X:d.d.d.d,前 96 位采用冒分十六进制表示,最后的 32 位地址则使用 IPv4 的点分十进制表示。例如 ::192.168.0.1 与 ::FFFF:192.168.0.1 就是两个典型的例子,注意在前 96 位中,0 位压缩的表示方法依旧适用。

4) IPv6 地址的分类

IPv6 将地址分为三种类型:单播地址(Unicast Address)、组播地址(Multicast Address)和任播地址(Anycast Address)。它与 IPv4 地址相比,新增了任播地址类型,取消了原来 IPv4 地址中的广播地址。

(1) 单播地址:用来唯一标识一个接口,类似于 IPv4 中的单播地址。发送到单播地址的数据报文将被传送给此地址所标识的一个接口。

(2) 组播地址:用来标识一组接口(通常这组接口属于不同的结点),类似于 IPv4 中的组播地址。发送到组播地址的数据报文被传送给此地址所标识的所有接口。

(3) 任播地址:用来标识一组接口(通常这组接口属于不同的结点)。发送到任播地址的数据报文被传送给此地址所标识的一组接口中距离源结点最近(根据使用的路由协议进行度量)的一个接口。

3.2.2 域名

1. 域名的概念

为了方便用户使用和记忆,将每个 IP 地址映射为一个由字符串组成的主机名,使 IP 地址从无意义的数字变为有意义的主机名,这个与网络上的数字型 IP 地址相对应的字符型地址称为域名(Domain Name)。通俗地说,域名就是企业、政府、非政府组织等机构或者个人在互联网上注册的名称,是互联网上企业或机构间相互联络的网络地址。例如,我们熟悉的百度域名就是 www.baidu.com。

2. 域名的构成

一个域名一般由英文字母和阿拉伯数字以及连字符“-”组成,最长可达 67(包括后缀)个字符,并且字母的大小写没有区别,每个层次最长不能超过 22 个字母。这些符号构成了域名的前缀、主体和后缀等几个部分,组合在一起构成一个完整的域名,称为完全合格域名。

3. 域名的层次结构

域名采用层次结构,如图 3. 2. 2 所示。每一层构成一个子域名,子域名之间用圆点“.”隔开,自上至下分别为根域、顶级域、二级域、子域及主机名,例如 www. tsinghua. edu. cn,其中 www 就是主机名,tsinghua 是子域名,edu 是二级域名,cn 则是顶级域名。

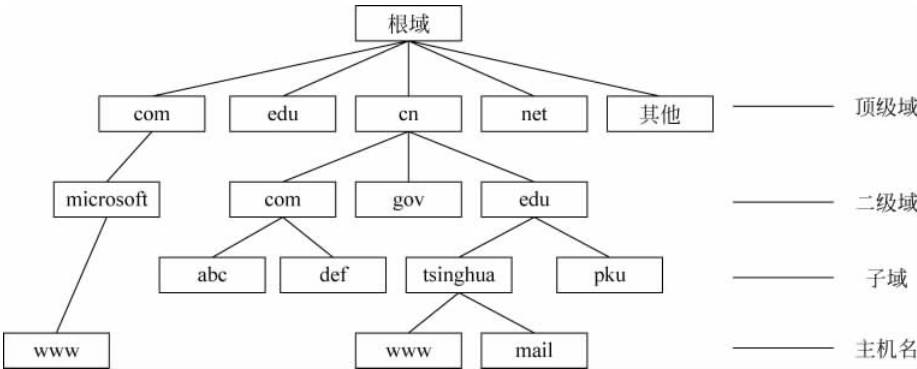


图 3. 2. 2 域名的层次结构

(1) 顶级域名。顶级域名又可以分为两大类：一是国家顶级域名；二是国际顶级域名。顶级域名如表 3. 2. 2 所示。

表 3. 2. 2 顶级域名

顶级域名		机构类型
组织机构	com	工、商、金融等企业
	edu	教育机构
	gov	政府部门
	net	互联网络、接入网络的信息中心
	org	各种非营利性组织
	int	国际组织
国家(地区)代码	cn	中国
	uk	英国
	us	美国
	jp	日本

(2) 二级域名。二级域名是指顶级域名之下的域名。在国际顶级域名下,是指域名注册人的网上名称,例如 ibm、yahoo、microsoft 等;在国家顶级域名下,表示注册企业类别的符号,例如 com、edu、gov、net 等。在顶级域名下,我国的二级域名又分为类别域名和行政区域域名。二级域名如表 3. 2. 3 所示。

表 3.2.3 二级域名

二级域名		机构类型
类别域名	ac	科研机构
	com	工、商、金融等企业
	edu	教育机构
	gov	政府部门
	net	互联网信息中心和运行中心
	org	各种非营利性组织
行政区域域名	共 34 个	对应我国各省、自治区和直辖市

除了上述介绍的英文域名外,现在也可以使用一些中文域名。

中文域名,就是以中文表现的域名。由于互联网起源于美国,英文成为互联网上资源的主要描述性文字。这一方面促使互联网技术和应用的国际化,另一方面,随着互联网在非英文国家和地区的普及,英文又成为非英语文化地区人们融入互联网世界的障碍。

中文域名是含有中文的新一代域名,同英文域名一样,是互联网上的门牌号码。中文域名属于互联网上的基础服务,注册后可以对外提供 WWW、E-mail、FTP 等应用服务。根据 2018 年《中国互联网域名体系》所述,中文域名分为国家顶级中文域名和通用顶级中文域名,国家顶级中文域名为“. 中国”,通用顶级中文域名有. 公司、. 网站、. 政务、. 公益、. 手机、. 网址等。例如,http://清华大学. 中国、http://新浪. 公司和 http://中央电视台. 中国。

3.2.3 网络协议简介

在计算机网络中,两个相互通信的实体处在不同的地理位置,当实体中的两个进程相互通信时,需要通过交换信息来协调它们的动作并达到同步,而信息的交换必须按照预先共同约定好的过程进行。这种预先约定好的过程就是网络协议。网络协议是为计算机网络中进行数据交换而建立的规则、标准或约定的集合。目前常见的协议有 TCP/IP、IPX/SPX、NetBEUI 协议等。

网络协议主要由语义、语法和交换规则三部分组成,即协议三要素。

(1) 语义:规定通信双方彼此“讲什么”,即确定协议元素的类型,如规定通信双方要发出什么控制信息、执行的动作和返回的应答。

(2) 语法:规定通信双方彼此“如何讲”,即确定协议元素的格式,如数据和控制信息的格式。

(3) 交换规则:规定信息交流的次序。

1. OSI 参考模型

为了使不同计算机厂家生产的计算机能够相互通信,以便在更大的范围内建立计算机网络,国际标准化组织(ISO)在 1978 年提出了“开放系统互连参考模型”,即著名的 OSI 参考模型。它将计算机网络体系结构的通信协议划分为 7 层,自下而上依次为物理层、数据链路层、网络层、传输层、会话层、表示层、应用层,如图 3.2.3 所示。

低3层(第1~3层)是依赖网络的,涉及将两台通信计算机连接在一起所使用的数据通信网的相关协议,实现通信子网功能。高3层(第5~7层)是面向应用的,涉及允许两个终端用户应用进程交互作用的协议,通常是由本地操作系统提供的一套服务,实现资源子网功能。中间的传输层为面向应用的上3层遮蔽了跟网络有关的下3层的详细操作。从实质上讲,传输层建立在由下3层提供服务的基础上,为面向应用的高层提供与网络无关的信息交换服务。

2. TCP/IP

TCP/IP 是 Transmission Control Protocol/Internet Protocol 的简写,是“传输控制协议/互联网络协议”。TCP/IP 是一种网络通信协议,它规范了网络上的所有通信设备,尤其是一个主机与另一个主机之间的数据往来格式以及传送方式。TCP/IP 是由一系列协议组成的协议簇。它是 Internet 的基础协议,也是一种计算机数据打包和寻址的标准方法。在数据传送中,可以形象地理解为有两个信封,TCP 和 IP 就像是信封,要传递的信息被划分成若干段,每一段塞入一个 TCP 信封,并在该信封上记录有分段号的信息,再将 TCP 信封塞入 IP 大信封,发送网上。在接收端,一个 TCP 软件包收集信封,抽出数据,按发送前的顺序还原,并加以校验,若发现差错,TCP 将会要求重发。对普通用户来说,并不需要了解网络协议的整个结构,仅需了解 IP 的地址格式,即可与世界各地的网络设备进行通信。

近年来还有学者提出了 TCP/IP 参考模型,它将计算机网络体系结构的通信协议划分为4层,自下而上依次为网络接口层、网际互连层、传输层、应用层。TCP/IP 网络协议就分布在 TCP/IP 的各层结构中,其对应关系如图 3.2.4 所示。



图 3.2.3 OSI 参考模型

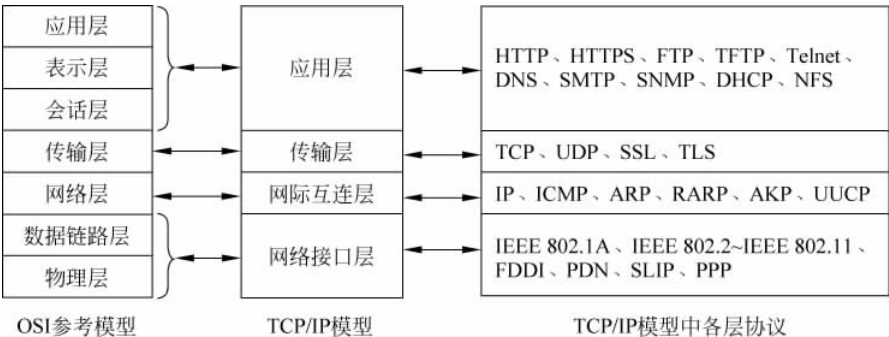


图 3.2.4 参考模型与网络协议对应关系

3. HTTP

HTTP 的全称是 HyperText Transfer Protocol,即超文本传输协议,从 1990 年开始就在 WWW 上广泛应用,是现今在 WWW 上应用最多的协议。HTTP 是应用层协议,并且是基于 TCP 连接的。当用户上网浏览网页的时候,客户端在浏览器与 Web 服务器之间就会通过 HTTP 在 Internet 上进行数据的发送和接收。HTTP 是一个基于请求/响应模式的、无状态的协议,即通常所说的 Request/Response。HTTP 请求/响应模型如图 3.2.5 所示。

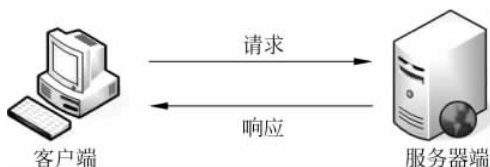


图 3.2.5 HTTP 请求/响应模型

HTTP 的功能如下。

(1) HTTP 可以使客户端的浏览器更加高效,使网络传输减少。它不仅保证计算机正确快速地传输超文本文档,还确定传输文档中的哪一部分,以及哪部分内容首先显示(如文本先于图形)等。

(2) HTTP 是客户端与 Web 服务器之间的应用层通信协议。在 Internet 上的 Web 服务器上存放了超文本信息,客户端需要通过 HTTP 传输所要访问的超文本信息。HTTP 包含命令和传输信息,不仅可以用于 Web 访问,也可以用于其他因特网/内联网应用系统之间的通信,从而实现各类应用资源超媒体访问的集成。

(3) 用户在浏览器的地址框中输入一个 URL 或者单击一个超链接时,URL 或超链接确定 Web 服务器的地址。客户端的浏览器通过 HTTP,将 Web 服务器上站点的网页代码提取出来,并翻译成漂亮的网页。

4. HTTPS

HTTP 被用于在客户端浏览器和 Web 服务器之间传递信息,HTTP 以明文方式发送内容,不提供任何方式的数据加密,如果攻击者截取客户端浏览器和 Web 服务器之间的传输报文,就可以直接还原其中的信息,因此 HTTP 不适合传输安全性比较高的敏感信息,如银行卡号、支付密码等。HTTPS(Hypertext Transfer Protocol Secure)解决了 HTTP 的这一缺陷,其建立以安全为目标的 HTTP 通道,简单讲就是 HTTP 的安全版。

HTTPS 的功能如下。

(1) 认证用户和服务器,保障网站本身以及访问网站的用户有更好的安全性。

(2) 加密数据以防止数据中途被窃取,保证数据的机密性,可以避免第三方窃听或阻断流量。

(3) 保证所传输数据的完整性和不可抵赖性,保护用户的隐私和安全。

3.2.4 URL 简介

在互联网上,每一个信息资源都有统一的且是唯一的地址,该地址就叫 URL(Uniform Resource Locator,统一资源定位符),它是 WWW 的统一资源定位标志,就是网络地址。

以如下 URL——`http://www.aspxfans.com:8080/news/index.asp?boardID=5&.ID=24618&.page=1#name` 为例,说明 URL 的基本格式。

- 协议的部分:“http”,在 Internet 中可以使用多种协议,如 HTTP、HTTPS、FTP 等,协议后的“//”为分隔符。

- 域名部分：“www.aspxfans.com”。一个 URL 中，也可以使用 IP 地址代替域名部分。
- 端口部分：“: 8080”，域名和端口之间使用“:”作为分隔符。端口不是一个 URL 必需的部分，如果省略端口部分，则将采用默认端口 80。
- 虚拟目录部分：“/news”，从域名后的第一个“/”开始到最后一个“/”为止。虚拟目录也不是必需的部分。
- 文件名部分：“/index.asp”，从最后一个“/”开始到“?”或“#”为止的部分都是文件名部分，如果没有“?”或“#”，那么就从最后一个“/”到结束都是文件名部分。文件名也不是必需的部分，如果省略则使用默认文件名。
- 参数部分：“? boardID=5&ID=24618&page=1”，从“?”开始到“#”为止的部分为参数部分，又称搜索部分、查询部分。可以允许有多个参数，参数与参数之间用“&”作为分隔符。
- 锚部分：“#name”，从“#”开始到最后都是锚部分，可以理解为定位。锚部分也不是 URL 必需的部分。

3.3 信息系统概述

信息系统(Information System)作为现代社会的一个组成部分,是指由一组计算机硬件、网络、通信设备、计算机软件、信息资源、信息用户和规章制度组成的以处理信息流为目的的人机一体化系统。

信息系统是新兴学科,其主要任务是最大限度地利用现代计算机及网络通信技术加强企业的信息管理,通过对企业的人力、物力、财力、设备、技术等资源的调查了解,整理准确的数据,加工处理并编制成各种信息资料,及时提供给管理人员,以便进行正确的决策,不断提高企业的管理水平和经济效益。企业的信息系统已成为企业进行技术改造及提高企业管理水平的重要手段。

3.3.1 信息系统的分类

以信息系统的发展和特点划分,可分为数据处理系统(Data Processing System, DPS)、管理信息系统(Management Information System, MIS)、决策支持系统(Decision Sustainment System, DSS)、专家系统(人工智能(AI)的一个子集)和虚拟办公室(Virtual Office, VO)5 种类型。

3.3.2 信息系统的功能

信息系统的 5 个基本功能: 输入功能、存储功能、处理功能、输出功能和控制功能。

输入功能: 实现各种信息的输入,它决定于系统所要达到的目的及系统的能力和息环境的许可。

存储功能: 信息系统存储各种信息资料和数据的能力。

处理功能: 基于数据仓库技术、数据挖掘技术等联机分析和数据处理。

输出功能：实现各种信息的输出，保证实现最佳的输出。

控制功能：通过各种程序对整个信息加工、处理、传输、输出等环节进行控制。

3.3.3 网络信息系统实例

网络信息系统的应用在今天已经越来越普及，本小节以 12306 网站为例，分析网络信息系统的应用过程。

1. 12306 网站系统概述

中国铁路客户服务中心即 12306 网站(网址为 <http://www.12306.cn>)，是中国国家铁路集团有限公司下属的信息服务网站，该网站基于中国铁道科学研究所设计的“铁路客票发售及预订系统”创建。用户通过登录网站，可以查询旅客列车时刻表、票价、列车正晚点、车票余票、售票代售点、货物运价、车辆技术参数以及有关客货运规章。12306 网站的首页如图 3.3.1 所示。



图 3.3.1 12306 网站首页

在高度信息化的今天，铁路客票发售及预订系统完成大多数人主要的订票服务，越来越多的人选择网络订票，而飞速增长的用户数量也使得 12306 网站成为世界上最繁忙的网站之一，12306 网站的访问量远远超过了淘宝、京东等国内知名的电商平台，这在某种意义上意味着 12306 网站要面临更大的挑战。首先整个售票系统是一个非常庞大而复杂的系统，是一个高负荷、高并发的云平台，其规模比淘宝大 2~3 倍，而且对于数据的实时性要求非常高。据有关统计，12306 网站系统的日访问量达到 15 亿次，如果加上各个代售点和车站售票系统，则高峰时段数据访问层的并发量在千万级别。如此大的访问和并发量，必然要求系统具有非常高的稳定性和健壮性。

2. 系统架构分析

铁路客票发售及预订系统是在铁道部原有的联网售票系统基础上开发的,所以其原有的数据架构直接影响到整个系统的扩展性和稳定性。首先 12306 网站系统是一个云平台的典型应用,系统按云平台的思想分层设计,从上而下分为 3 层,即应用层、数据访问层、数据层。每一层之间松散耦合,松散耦合使得每一层都具有很强的扩展性和伸缩性。每一层内部都基于集群技术,分组部署,每一组处理单元都即插即用,可根据计算压力动态扩充,其总体结构如图 3.3.2 所示。

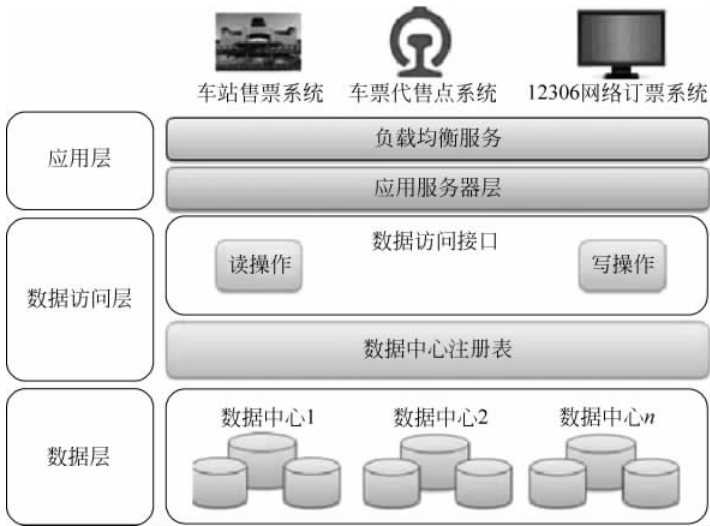


图 3.3.2 12306 网站的总体结构图

应用层：主要是指各种售票和订票系统,主要包括车站售票系统、车票代售点系统及 12306 网络订票系统 3 种。其中前两个是 C/S 结构的应用模式,后一个是 B/S 结构的应用模式。其客户端与应用服务器之间增加一个负载均衡服务,这有利于系统的并发,可以有效地根据当前用户量和访问情况自动地分配相对压力较小的服务器。

数据访问层：主要是将业务应用与底层数据库之间的操作接口专门独立出来,业务应用访问数据不是直接访问数据库,而是通过数据访问层进行间接的访问和操作。这样的好处是可以解决数据访问的并发瓶颈,可以根据系统的压力情况动态地调整和部署访问层。

数据层：根据车次和地域将车次的余票信息分别存储在很多个数据中心上,每一个数据中心是一组服务器。这样将众多的并发用户根据查询车次分散到多个数据中心上去,从而降低单点压力,提高整体的并发性能。如果数据访问是一个大瓶颈,则可增加数据中心的结点而减小数据中心的粒度(也就是每个数据中心减少车次数量),还可提高数据访问的速度。

3. 详细架构

12306 网站系统整体按分层架构处理,每一层都是可注册、可插拔的体系。这种架构的好处是每一层都可以分层优化,而互不影响。根据实际运行的情况对并发和访问量过

大的实体层进行动态扩容,很容易提高系统的并发能力和稳定性。其详细架构如图 3.3.3 所示。此架构很好地解决了应用服务器和数据访问的瓶颈问题,如果应用服务器压力大则可以通过注册表对应用服务器扩容,通过负载均衡地访问各个应用服务器。如果数据访问是一个瓶颈,则可通过增加数据中心的方式来解决数据访问拥挤情况。

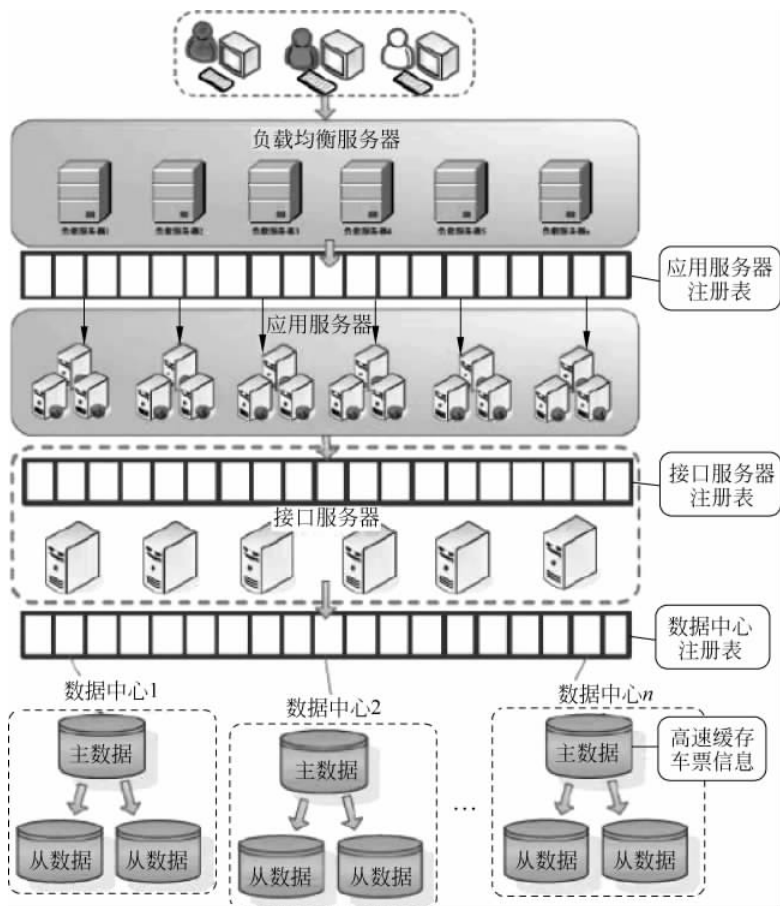


图 3.3.3 12306 网站详细架构

对于数据层系统按车次对所有的车票信息进行分组,每一组是一个数据中心,数据中心的大小可随时调整。这样可以把用户对数据的访问分散到多个结点上去,从而降低数据中心的压力。每一个数据中心由若干台服务器组成:一台主数据服务器和若干台从数据服务器。主数据服务器用于给用户出票,每一个接口调用都需要加锁,保证票数数据修改的准确性,其所管理的车次和车票数据在内存中高速缓存,同时每隔一定的时间周期同步到从数据服务器上;从数据服务器上的数据用来提供查询的数据副本,它把大量的查询操作分散到从数据服务器上。

4. 网站常见问题及分析

现行 12306 网站系统的常见问题总结如下。

- (1) 高峰时段无法登录,提示在线用户过多。
- (2) 订单提交成功之后,支付环节出问题,浏览器意外退出,再登录却发现登录不上,无法在规定时间内完成支付,购票失败。
- (3) 订单提交反馈时间长,热门线路需要等待 20min 甚至更长时间才能得到反馈。
- (4) 验证码输入总是错误,无法完成验证环节,无法登录。
- (5) 逢用户高峰,网站反应速度较慢。

针对以上问题,分析如下:无法登录的问题,原因是前端用于处理 Web 连接服务器太少或网络带宽不足,为了不让更多的用户一起连接服务器导致服务器较慢,只好拒绝一些用户的登录请求,最好的处理办法是使同时在线人数保持在一个上限以内;验证码输入错误的问题,原因也是用于处理 Web 连接服务器太少所致,为了防止一些客户端使用恶意软件不断自动登录,验证码需要由客户端向服务器提交一个验证请求,由于服务器响应实在太慢,有时响应速度会超过验证码的有效时间;对于订单提交反馈时间长、网站反应速度较慢等问题,大多是由于 Web 服务器与逻辑处理服务器在同一台机器上,导致服务器 CPU 分配过多的时间与资源处理用户请求,在执行逻辑过程时缓慢。

课后习题

一、选择题

1. 在 Internet 中,按()地址进行寻址。
 - A. 邮件地址
 - B. IP 地址
 - C. MAC 地址
 - D. 网线接口地址
2. 在网络拓扑结构中,将各结点通过一条首尾相连的通信线路连接起来而形成一个封闭环,并且数据只能沿单方向传输的结构为()。
 - A. 总线拓扑结构
 - B. 星状拓扑结构
 - C. 环状拓扑结构
 - D. 树状拓扑结构
3. 下列四项中表示域名的是()。
 - A. www.cctv.com
 - B. hk@zj.school.com
 - C. zjwww@china.com
 - D. 202.96.68.1234
4. 网址 www.pku.edu.cn 中的 cn 表示()。
 - A. 英国
 - B. 美国
 - C. 日本
 - D. 中国
5. 下列 IP 地址中书写正确的是()。
 - A. $168 * 192 * 0 * 1$
 - B. 325.255.231.0
 - C. 192.168.1
 - D. 255.255.255.0
6. 以下关于 IPv4 和 IPv6 的说法,错误的是()。
 - A. IPv4 存在局限性,所以人们开发了 IPv6
 - B. IPv6 依然沿用了 IPv4 的核心设计思想
 - C. IPv6 在协议格式、地址表示法方面和 IPv4 是相同的
 - D. IPv6 采用了 128 位地址长度

7. 计算机网络的功能有()。
- A. 用户管理 B. 资源共享 C. 病毒管理 D. 站点管理
8. 通常按网络覆盖的地理范围分类,可以分为局域网、()和广域网三种。
- A. 星状网络 B. 有线网 C. 城域网 D. 无线网
9. ()是将地理位置上相距较远的多个计算机系统,通过通信线路按照网络协议连接起来,实现计算机之间相互通信的计算机系统的集合。
- A. 广域网 B. 局域网 C. 互联网 D. 电信网
10. 下列不属于信息系统功能的是()。
- A. 存储功能 B. 输入输出功能
- C. 控制功能 D. 网络功能

二、判断题

1. 网络域名地址一般都通俗易懂,大多采用英文名称的缩写来命名。
2. 目前使用的广域网基本都采用星状拓扑结构。
3. 计算机网络按拓扑结构可分为星状网络、总线网络、树状网络、环状网络和网状网络 5 种。
4. 局域网的英文缩写为 LAM。
5. 环状结构网络中各计算机的地位相等。
6. 广域网的各连接设备比城域网的各连接设备之间的距离远,数据传输率低,错误率低。
7. IP 地址具有层次特点,将号码分割成网络号和主机号两部分,这样便能唯一地指定每一台主机。
8. IPv6 是 IPv4 的下一代协议,其 IP 协议地址长度为 256 位。
9. 计算机网络的功能不包括资源共享。
10. 信息系统的开发用不到网络技术。

三、思考题

1. 简述计算机网络的概念。
2. 简述计算机网络的功能。
3. 简述网络拓扑结构的类型。
4. 简述网络域名的作用。
5. 简述参考模型与网络协议的对应关系。
6. 简述信息系统的功能。