第3章

计算机网络体系结构

计算机网络是由各类具有独立功能的计算机系统和终端通过通信线路连接起来的复杂系统，网络中各计算机必须遵从通信规定才能相互协调工作。为了设计这样复杂的系统，网络工作者提出了分层实现计算机网络功能的方法。

本章要点

※ 协议与网络体系结构及其相关的基本概念

※ OSI/RM 的分层结构，以及各层的功能与服务

※ TCP/IP 体系结构

3.1 网络体系结构基本概念

网络通信协议和网络体系结构是计算机网络技术中重要的内容，要掌握计算机网络的基本原理，就必须对协议和网络体系结构的概念和相关知识有比较深入的了解。

3.1.1 通信协议

通信协议是计算机网络中为进行数据通信而制定的通信双方共同遵守的规则、标准或约定的集合。在网络中通信双方之间必须遵从相互可以接受的网络协议（相同或兼容的协议）才能进行通信，如目前因特网上使用的协议是 TCP/IP。

协议本质上是一系列规则和约定的规范性描述，它不仅定义了通信时信息必须采用的格式和这些格式的意义，而且还要对事件发生的次序做出说明。所以，任何一种网络协议都应包括如下 3 个要素。

※ 语法（syntax)：规定通信双方“如何讲”，是将若干协议元素和数据组合起来表达一个更完整的内容时所应遵循的格式，即数据与控制信息的结构、编码及信号电平等。

※ 语义（semantics)：规定通信双方“讲什么”，即协议元素的含义，如控制信息、执行的动作和返回的应答等。

※ 时序（timing，又称时序或定时）：规定通信双方“讲的顺序”或“应答关系”。即对事件实现顺序的说明，解决何时进行通信的问题。

协议的三要素看起来十分抽象，拿电报来做比喻，可以对它们有一个清晰的认识。拍电报时，必须首先规定好报文的传输格式、多少位的码长、什么样的码字表示开始、什么样的码字表示结束等，这种预先定好的格式就是语法，格式中的内容如发报人的名字和地址等就是语义，而电报收发的先后次序就是时序，这些要素构成了协议。

3.1.2　网络体系结构

体系结构是研究系统中各组成部分及其关系的一门学科，这个术语后来被计算机网络工作者所采用，为了使计算机网络系统能够在同一原则和方法下进行设计、构建和使用，提出了计算机网络体系结构的概念，对构成整个计算机网络的主要部分及应具备的功能给出了一组定义。要理解网络体系结构，首先必须了解分层的设计思想。

1．体系结构的层次化

将一个复杂系统分解为若干个容易处理的子系统，然后“分而治之”，这种结构化设计方法是工程设计中常见的手段。计算机网络是由各类具有独立功能的计算机系统和终端通过通信线路连接起来的复杂系统，网络中各计算机或结点之间的数据通信，数据从发送端的处理、发送，到经中继结点的交换转发，再到接收端的接收，发送端和接收端必须相互协调工作，才能保证正常的相互通信。为了设计实现这样的复杂系统，人们提出了分层实现计算机网络功能的方法，将复杂的问题进行分解、简化，分而治之。

分层结构是指把一个复杂系统的设计分解成层次分明的局部问题，并规定每一层次所必须完成的功能。为了便于理解分层，以两个城市邮寄信件的工作过程为例来说明。

在如图 3-1 所示的分层结构中，一个寄信的过程被分成 3 个层次来完成，即寄信人、邮局和传输部门。各层只需要完成自己的功能，下层为上层提供服务，同时各层还必须遵守各层的约定，通过这种模式来完成信件的邮寄任务。类似信件投递的过程，为了便于对计算机网络的组成成分、功能及协议的描述、设计和实现，现在都采用分层的体系结构，如图 3-2 所示。

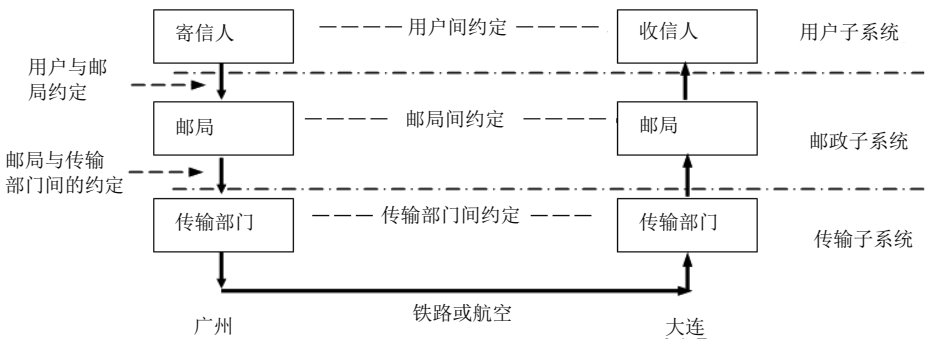


图 3-1　邮政系统分层结构

层次结构的好处在于每一层实现相对独立的功能。每一层不必知道下面一层是如何实现的，只要知道下层通过层间接口提供什么服务，以及本层应向上层提供什么服务，就能独立设计。系统经过分层后，每一层的功能相对简单且易于实现和维护。此外，若某一层需要改动或替代时，只要不改变它和上下层的服务关系，则其他层次都不会受到影响，因此具有很大的灵活性。每一层的功能和所提供的服务都有精确的说明，有助于标准化。

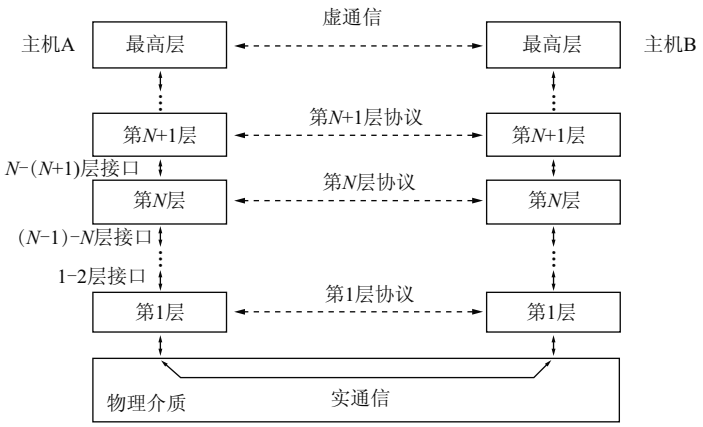


图 3-2　网络的层次结构

分层结构提供了一种按层次观察网络的方法，它描述了网络中任意两个结点间的逻辑连接和信息传输。同一系统分层结构中的各相邻层间的关系是：下层为上层提供服务，上层利用下层提供的服务完成自己的功能，同时再向更上一层提供服务。因此，上层是下层的用户，下层是上层的服务提供者。

2．网络体系结构的概念

计算机网络体系结构是从通信功能上来描述计算机网络结构的，网络体系结构为了完成计算机之间的通信，将通信的功能划分成定义明确的层次，规定了同层次进程通信的协议及相邻层之间的接口及服务。因此，将计算机网络的层次结构以及各层服务、协议的集合统称为计算机网络体系结构。

计算机网络体系结构是一个抽象的概念，是对网络通信所需要完成的功能的精确定义，只解决了“做什么”的问题，而不涉及“怎么做”。对于体系结构中所确定的功能用何种硬件或软件实现，如协议如何制定与实现，不属于网络体系结构的内容。可见体系结构是抽象的，是存在于纸上的，而实现则是具体的，是真正运行的计算机硬件和软件。

不同的网络体系结构，分层的数量、名称、功能、协议和接口可能不同，但是都遵守分层的原则，即各层完成的功能相对独立，某一层的内部变化不能影响到另一层，高层使用下层提供的服务时，下层服务的实现是不可见的。

3.1.3　分层结构中的相关概念

1．通信实体和对等实体

在网络通信中，通信实体是层功能实现的真正承担者（相应的软硬件），能发送和接收信息。例如，文件传输系统、电子邮件系统等，也可以是一块网卡、一个智能 I/O 芯片。系统中的各层次都存在一些实体，不同系统的相同层次称为对等层，对等层之间的通信称为对等层通信，而对等实体是指相互通信的两个不同系统上的同一层的通信实体。

2．服务和分层协议

网络服务是指相邻两层之间下层为上层所提供的操作功能或通信能力。由于网络分层的结构中的单向依赖关系，使下层总是向它的上层提供服务，上层可看成是下层的服务用户，

下层是上层的服务提供者。*N* 层使用 *N*-1 层及以下各层所提供的服务，向更高的 *N*+1 层提供服务。

在网络分层结构中，通信协议相应地被分为各层协议，每一层都可能有若干个协议，因此，网络中提到的协议总是指某一层的协议。*N* 层协议规定了第 *N* 层对等实体之间进行的虚通信必须遵守的规则。对等层通信所遵守的规则或约定称为同层协议。

在网络体系结构中，常提到“功能”“服务”和“协议”这几个术语，它们有着不同的含义。功能是本层内部的活动，是为了实现对外服务而从事的活动；而服务是本层提供给高一层使用的操作功能，属于外观的表象，只有那些能够被高一层看得见的功能才能称为服务；协议则相当于一种工具，对外的服务是依靠本层的协议实现的。

服务和协议的关系是：服务是“垂直的”，是下层为上层用户的需要而执行的一组操作，但并不规定这些操作是如何实现的；协议则是定义同层对等实体间信息交换的规则，所以协议是“水平的”。实体在实现其服务时必须遵守协议，但是只要不改变对用户而言可见的服务，对等实体可以选择或改变其协议。

3．面向连接的服务和无连接的服务

从通信角度看，各层所提供的服务有两种服务形式——面向连接的服务和无连接的服务。所谓“连接”是指在对等层的两个对等实体间所设定的逻辑通路。

（1）面向连接的服务。利用建立的连接进行数据传输的方式就是面向连接的服务。面向连接的服务思想来源于电话系统，即在开始通话之前，发送方和接收方必须通过电话网络建立连接线路，然后开始通话，通话结束后再拆除连接线路。面向连接的服务过程可分为 3 个部分——建立连接、传输数据和撤销连接。面向连接的服务比较适合于数据量大、实时性要求高的数据传输应用场合。

（2）无连接的服务。无连接的服务过程类似于邮政系统。通信前，无须在两个对等层之间事先建立连接，通信链路资源完全在数据传输过程中动态地进行分配，无论何时，计算机都可以发送数据。此外在通信过程中，双方并非需要同时处于“激活”（或工作）状态，如同在信件传递中收信人没必要当时位于目的地一样。因此，无连接服务的优点是灵活方便，信道的利用率高，特别适合于短报文的传输。

与面向连接的服务不同的是，由于无连接服务在通信前没有建立“连接”，因此传输的每个分组中必须包括目的地址，同时由于无连接方式不需要接收方的应答和确认，在此服务方式的数据传输中可能会出现分组的丢失、重复或乱序等错误。

4．接口和服务访问点

接口（interface）是同一系统相邻两层之间的边界，定义下层向上层提供的原语操作和服务。同一系统相邻两层实体交换信息的地方称为服务访问点 (Service Access Point，SAP）。SAP 很像常用的邮政信箱，它实际上是相邻两层实体的逻辑接口，也可以说 N 层的 SAP 就是 *N*+1 层可以访问 *N* 层的地方。SAP 有时也称为端口。任何层间服务都是在接口的 SAP 上进行的，每个 SAP 有唯一的识别地址，供服务用户之间建立连接之用，每个层间接口可以有多个 SAP。

5. 服务原语

从上面的讲述可知，当上层实体向下层实体请求服务时，服务用户与服务提供者之间通过服务访问点进行信息交互，在信息交互时所要交换的信息由服务原语来描述。

服务原语用来在形式上描述层间提供的服务，并规定通过 SAP 所必须传递的信息。上层利用服务原语来通知下层要做什么；下层利用服务原语来通知上层已做了什么。服务原语是描述服务的一种简洁形式，类似编程时的程序调用和参数传递，但不是可执行的程序语言。一个完整的服务原语由原语名、原语类型和原语参数 3 部分组成。

例如，一个网络连接请求原语的写法是：N-CONNECT.Request（目的地址，源地址）。

这里 N-CONNECT 是原语名字，Request 是原语类型，中间用圆点隔开，而括号内的内容则是原语参数。服务原语类型有以下 4 种：

※ 请求（request）。由服务用户发往服务提供者，请求它完成某些操作的服务，如建立连接、发送数据、释放连接等。

※ 指示（indication）。由服务提供者发往服务用户，指示发生了某些事件，如连接指示、释放连接指示等。

※ 响应（response）。由服务用户发往服务提供者，作为对前面指示的响应，如接受连接、接收释放连接等。

※ 证实（confirm）。由服务提供者发往服务用户，作为对前面发生请求的证实。

服务分为有证实服务和无证实服务。有证实服务服务包括请求、指示、响应和证实 4 个原语，无证实服务只有请求和证实两个原语。

3.2 OSI 参考模型

3.2.1 OSI/RM 的制定

计算机网络体系结构的出现加快了计算机网络的发展，但在计算机网络产生之初，一些大的计算机厂商开展了计算机网络的研究与产品开发，提出了各自的网络体系结构和协议，多数网络都采用分层的体系结构。如 IBM 公司于 1974 年提出的系统网络结构 SNA，DEC 公司于 1975 年提出的数字网络体系 DNA，其他计算机厂商也分别提出了各自的网络体系结构，以适应本公司的生产和商业目的，因此，不同的网络使用不同的网络体系结构和通信协议，彼此不认识各自的数据格式，使不同厂家生产的网络设备之间很难相互通信，在一定程度上阻碍了计算机网络的发展和应用。

为了解决不同网络设备之间的互联问题，国际标准化组织（ISO）在 20 世纪 80 年代初提出了著名的开放系统互连参考模型 OSI/RM（Open Systems Interconnection Reference Model）。

OSI/RM 是根据比较成熟的分层体系结构理论，结合当时比较成功的体系结构的经验制定的。制定过程中所采用的方法是分层处理法，将整个庞大而复杂的问题划分为若干个容易处

理的小问题。先根据网络的功能将网络划分成定义明确的层次，然后定义层间的接口及每层提供的功能和服务，最后定义每层必须遵守的规则，即协议。设计采用三级抽象技术，即体系结构、服务定义、协议规格说明。

第 1 级抽象：提出 OSI/RM，建立计算机网络在概念和功能上的框架，包括确定开放系统的层次结构，以及公共术语、子层功能等。

第 2 级抽象：服务定义，说明各个子层所提供的服务。

第 3 级抽象：协议规格说明，定义一组为确保子层服务的提供而应遵循的规则。

3.2.2　OSI/RM 结构及各层功能

1. OSI/RM 的结构

OSI 参考模型定义了计算机网络系统的层次结构、层次之间的相互关系及各层所包括的服务。它将网络通信功能划分为 7 个层次，规定了每个层次的具体功能。自顶向下的 7 个层分别是应用层、表示层、会话层、传输层、网络层、数据链路层和物理层，如图 3-3 所示。

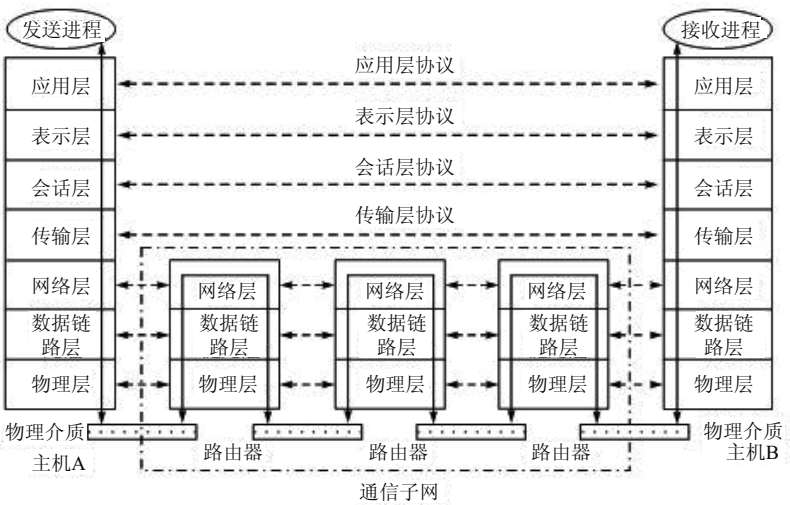


图 3-3　OSI/RM 示意图

从图 3-3 中可见，整个开放系统环境由资源子网中的主机和通信子网中的结点通过物理媒介连接构成。只有在主机中才可能需要包含所有 7 层的功能，而在通信子网一般只需要最低 3 层的功能。

OSI/RM 作为一个框架来协调和组织各层协议的制定，是对网络内部结构最精炼的概括与描述。它的最大特点是开放性，“开放”这个词表示只要遵循 OSI 标准，一个系统就可以和位于世界上任何地方的同样遵循 OSI 标准的其他任何系统进行连接。

OSI/RM 成功之处在于清晰地分开了服务、接口和协议这 3 个概念：服务描述每一层的功能；接口定义某层提供的服务如何被高层访问；而协议是每一层功能的实现方法。通过区分这些抽象概念，OSI/RM 将功能定义与实现分开，概括性高，具有普遍的适应能力。

应该注意的是，OSI/RM 并没有提供一个可以实现的方法，也就是说，OSI/RM 并不是一个标准，而只是一个在制定标准时所使用的概念性框架。在 OSI/RM 中只有各种协议是可以

实现的，网络中的设备只有与 OSI/RM 有关协议相一致时才能互连。OSI/RM 模型有 7 层，但 OSI/RM 本身并不满足网络体系结构要求，按照定义，网络体系结构是网络的层次结构和分层协议的集合，OSI/RM 没有精确定义各层的协议，只是描述了每一层的功能与服务。然而，国际标准化组织还是对各层制定了标准，每一层都作为一个单独的国际标准来颁布，尽管这些标准不是 OSI/RM 本身的一部分。

OSI/RM 从理论上为网络的发展指明了方向，对计算机网络起到了规范和指导作用。但是，在实际应用中，完全符合 OSI/RM 的产品却很少。现在，实际使用的网络互联协议是 TCP/IP 协议集，随着以 TCP/IP 协议为基础建立的 Internet 的飞速发展，TCP/IP 协议已成为计算机网络事实上的工业标准，得到了相当广泛的实际应用。

2. OSI/RM 各层主要功能简述

※ 应用层：是 OSI/RM 的最高层，提供用户应用软件与网络之间的接口服务。

※ 表示层：主要解决用户信息的语法表示问题。它将欲交换的数据从适合于某一用户的抽象语法，转换为适合于网络系统内部使用的传送语法，即提供格式化的表示和转换数据服务。数据的压缩和解压缩、加密和解密等工作也由表示层负责。

※ 会话层：是“进程 - 进程”的层次，其主要功能是组织和同步不同的主机上各种进程间的通信（也称为对话）。不参与数据传输，但对数据传输进行管理。在会话层及以上的高层次中，数据传送的单位不再另外命名，统称为报文。

※ 传输层：是“端 - 端”层次，该层的任务是根据通信子网的特性，最佳地利用网络资源，并以可靠和经济的方式，为两个端系统（源站和目的站）的会话层之间提供建立、维护和取消传输连接的功能，负责可靠地传输数据。这一层数据传送单位是报文。

※ 网络层：是“结点 - 结点”层次，在计算机网络中进行通信的两个计算机之间可能会经过很多个数据链路，也可能还要经过很多通信子网。网络层主要负责如何使数据分组跨越通信子网从一个结点到另一个结点的正确传送，即在通信子网中进行路由选择。当分组要跨越多个通信子网才能到达目的地时，还要解决网际互联的问题。另外，为避免通信子网中出现过多的分组而造成网络拥塞，需要对流入通信子网的分组数量进行拥塞控制。这一层数据传送单位是分组。

※ 数据链路层：是相邻结点层次，主要功能是通过校验、确认和反馈重发等手段，将不可靠的物理链路改造成对网络层来说无差错的数据链路，为网络层在相邻结点间无差错的传送以帧为单位的数据。数据链路层还要协调收发双方的数据传输速率，即进行流量控制，以防止接收方因来不及处理发送方发来的高速数据而导致缓冲器溢出丢失。这一层的数据传送单位是帧。

※ 物理层：要传递数据就要利用一些物理媒体，如双绞线、同轴电缆等，但具体的物理媒体并不是物理层。物理层的任务是为它的上一层提供一个物理连接，定义了为建立、维护和拆除物理链路所需的机械的、电气的、功能的和规程的特性，其作用是确保原始的数据比特流能够在物理媒体上传输。在物理层数据的传送单位是位（bit）。

3.2.3　OSI/RM 中的数据传输

在 OSI/RM 中，每一层将上层传递过来的数据加上若干控制位后再传递给下一层，最终由物理层传递到对方物理层，再逐级上传，从而实现对等层之间的逻辑通信，如图 3-4 所示。不同主机对等层之间按相应协议进行通信，同一主机不同层之间通过接口进行通信。除了最低层的物理层是通过传输介质进行物理数据传输外，其他对等层之间的通信均为逻辑通信。图 3-4 中自上而下的实线表示的是数据实际的传送过程。

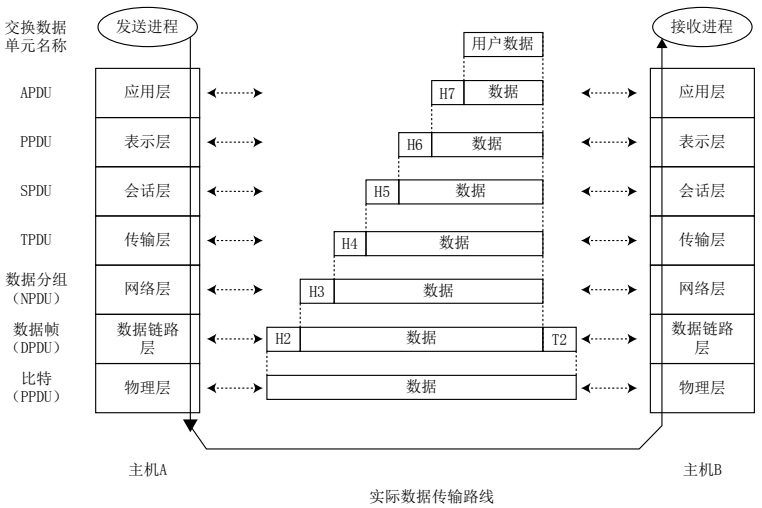


图 3-4　OSI/RM 中的数据传输

用户数据首先要经过发送方的应用层，应用层在用户数据前面加上本层的控制信息 H7，称作“头信息”。H7 加上用户数据一起传到表示层，表示层则将 H7 和原始用户数据当作一个整体数据部分对待。同样，表示层也在整体数据前面加上本层的控制信息 H6，传到会话层，并作为数据部分。这个过程一直进行到数据链路层，数据链路层除了增加头信息 H2 以外，还要增加一个尾信息 T2，然后整个作为数据传送到物理层。数据到物理层成为由 0 或 1 组成的数据比特流，然后再转换为电信号在物理媒体上传输至接收方。接收方收到数据后在向上传递时其过程正好相反，要逐层剥去发送方相应层加上的控制信息，其中数据链路层负责去掉 H2 和 T2，网络层去掉 H3，一直到应用层去掉 H7，最终把原始数据传递给了接收进程。

这个在发送方自上而下逐层增加头信息的过程称为数据的封装，而在接收方又自下而上逐层去掉头信息的过程称为数据的拆封。因接收方的某一层不会收到底下各层的控制信息，而高层的控制信息对于它来说又只是透明的数据，所以它只阅读和去除本层的控制信息，并进行相应的协议操作。发送方和接收方的对等实体看到的信息是相同的，就好像这些信息通过虚通信直接给了对方一样。

协议数据单元（Protocol Data Unit, PDU）是对等层实体之间通过协议传送的数据单元，包括应用层的协议数据单元 APDU（Application Protocol Data Unit）、表示层的协议数据单元 PPDU（Presentation Protocol Data Unit）……网络层的协议数据单元 NPDU（Network Protocol Data Unit）。通常人们把网络层的 PDU 称为分组或包（packet），数据链路层的 PDU 称为帧（frame），物理层是比特（bit）。

3.3 物理层

3.3.1 物理层功能与协议

物理层是 OSI/RM 中的最底层，是整个开放系统的基础。计算机网络中有许多物理设备和传输介质，但物理层不是指这些连接设备的具体传输介质，它是介于数据链路层和传输介质之间的一层，起着数据链路层到传输介质之间的接口作用。由于物理层的存在，使数据链路层感觉不到传输介质的差异，这样，数据链路层就可以不必考虑网络的具体传输介质，而只完成本层的服务。

1. 物理层的功能与服务

物理层的基本功能是负责实际或原始的数据“位”传送，目的是在通信设备 DTE 和 DCE 之间提供透明的比特流传输。物理层向数据链路层提供的服务是建立、维持和释放物理连接，并在物理连接上透明传输比特流。

另外，CCITT 在 X.25 建议书第 1 级（物理级）中也做了类似的定义：利用物理的、电气的、功能的和规程的特性，在 DTE 和 DCE 之间实现对物理信道的建立、保持和拆除功能。这里的 DTE(Date Terminal Equipment) 指的是数据终端设备，是对属于用户所有的连网设备或工作站的统称，它们是通信的信源或信宿，如计算机、终端等；DCE(Date Circuit Terminating Equipment 或 Date Communications Equipment) 指的是数据电路终接设备或数据通信设备，是对为用户提供入接点的网络设备的统称，如自动呼叫应答设备、调制解调器等。

2. 物理层协议

物理层协议规定了网络物理设备之间的物理接口特性及通信规则，即定义了为建立、维护和拆除物理链路所需的机械、电气、功能和规程特性。物理层协议实际上是 DTE 和 DCE 之间接口及传输比特的规则的一组约定，主要解决网络设备与物理信道如何连接的问题，其作用是确保比特流能够在物理信道上传输，DTE-DCE 接口如图 3-5 所示。例如，PC 上的 COM1 和 COM2 接口称为 RS-232 接口，使用的是典型的物理层协议 RS-232C。

图 3-5　DTE-DCE 接口

物理层协议用 4 个特性对网络设备和传输介质之间的接口进行定义。

※ 机械特性：规定物理连接器的规格尺寸、插针或插孔的数量和排列方式等。

※ 电气特性：规定传输二进制比特流有关的特性，如信号电压的高低、阻抗匹配、传输速率和距离限制等，通常包括发送器和接收器的电气特性，以及与电缆相关的规则等。

※ 功能特性：规定各信号线的功能。信号线按功能可分为数据线、控制线、定时线和接地线等。

※ 规程特性：定义 DTE 和 DCE 通过接口连接时，各信号线进行二进制位流传输的一组操作规程（动作序列），如怎样建立、维持和拆除物理连接，以及全双工或半双工操作等。

3.3.2　物理层协议举例

目前使用的计算机和调制解调器的串行接口 EIA RS-232C 就是物理层协议的一个例子。EIA RS-232C 是由美国电子工业协会（Electronic Industry Association，EIA. 在 1969 年颁布的一种串行物理接口。RS（Recommended StandarD. 的意思是“推荐标准”，232 是标识号码，而后缀 C 则表示该推荐标准已被修改过的次数。

RS-232C 的机械特性规定使用一个 25 针、接口形状为 D 型的标准连接器（DB-25），宽 47.04mm ± 0.13mm，每个插座有 25 个插头，编号为 1~25。

RS-232C 的电气特性规定逻辑 1 的电平为 −15 ~ −5V；逻辑 0 的电平为 +5 ~ +15V，即 RS-232C 采用 +15V 和 −15V 的负逻辑电平，+5V 和 −5V 之间为过渡区域，不做定义。允许的最大传输速率为 20kb/s，最长可驱动电缆 15m，如图 3-6 所示。

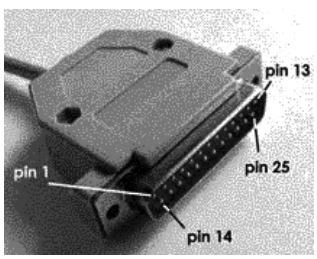


图 3-6　RS-232C 串行接口

RS-232C 的功能特性定义了 25 针标准连接器中的 20 根信号线，其中包括 2 根地线、4 根数据线、11 根控制线、3 根定时信号线，剩下的 5 根线作为备用或未定义。表 3-1 给出了其中最常用的 10 根信号线的功能特性。

表 3-1　RS-232C 功能特性

引脚号	功能说明	信号线型	连接方向
1	保护地线 (GND)	地线	
2	发送数据 (TD)	数据线	→ DCE
3	接收数据 (RD)	数据线	→ DTE
4	请求发送 (RTS)	控制线	→ DCE
5	清除发送 (CTS)	控制线	→ DTE
6	数据设备就绪 (DSR)	控制线	→ DTE
7	信号地线 (Sig.GND)	地线	
8	载波检测 (CD)	控制线	→ DTE
20	数据终端就绪 (DTR)	控制线	→ DCE
22	振铃指示 (RI)	控制线	→ DTE

RS-232C 的工作过程是在各根控制信号线有序的 ON（逻辑 0）和 OFF（逻辑 1) 状态的配合下进行的。在 DTE-DCE 连接的情况下，只有数据终端就绪 DTR 和数据设备就绪 DSR

均为 ON 的状态时，才具备操作的基本条件。此后，若 DTE 要发送数据，则须先将请求发送 RTS 置为 ON 状态，等待清除发送 CTS 应答信号为 ON 状态后，才能在发送数据 TD 上发送数据。例如，发送规程为：4 针置位，请求发送；5 针置位，准许发送；数据通过 2 针发出。

3.4 数据链路层

数据链路层是 OSI/RM 中的第 2 层，它在物理层基础上向网络层提供服务。物理层是通过传输媒体形成物理连接，但在物理媒体上传输的数据难免受到各种因素的影响而产生差错，使物理连接是有差错的、不可靠的。另外，发送端和接收端的物理设备之间可能存在发送和接收速度不匹配，导致缓冲区溢出和数据丢失等问题。为了进行有效的、可靠的数据传输，就需要对传输操作进行严格的控制和管理。设立数据链路层的主要目的是对物理层传输原始比特流的功能的加强，将物理层提供的可能出错的物理链路通过数据链路层协议改造为逻辑上无差错的数据链路，使之对网络层表现为一条无差错的传输通路。这就是数据链路层的任务，也就是数据链路层协议的任务。

3.4.1 数据链路层的功能与服务

数据链路层解决两个相邻结点间的通信问题，提供的服务是通过数据链路层协议在不太可靠的物理链路上实现可靠的数据传输，向网络层提供透明的和可靠的数据传送服务。

相邻结点的数据交换应保证帧同步和各帧顺序传送，对损坏、丢失和重复的帧应能进行处理，为网络层提供一条可靠的、无差错的数据传输通路，这种处理过程对网络层是透明的。为了实现这个目的，数据链路层必须能完成如下的主要基本功能：

※ 链路管理。当网络中的两个相邻结点要进行通信时，数据的发送方必须明确知道接收方是否已经处于准备接收的状态。为此，通信双方必须先交换一些必要的信息，建立一条数据链路。同样地，在数据传输时要维持数据链路，而在通信完毕时要释放数据链路。数据链路的建立、维持和释放就叫作链路管理。

※ 帧同步。数据链路层的数据传送单位是帧，数据一帧一帧地传送。帧同步是指接收端应当能够从收到的比特流中准确地区分出一帧的开始与结束。

※ 差错控制。处理传输中可能出现的差错。

※ 流量控制。协调传输中发送方的发送速率大于接收方的问题。

※ 将数据和控制信息区分开。由于要传输的数据和控制信息处于同一帧中，因此要有相应的措施使接收方能够将二者区分开。

※ 透明传输。无论什么样的比特串，都应当能在链路上传输。当出现实际传输数据中的比特串恰巧与某一控制信息的比特串完全一样时，必须采取适当的措施，使接收方不会将这样的比特串误解为某种控制信息，这就是透明传输。

※ 寻址。在多点连接的情况下，必须保证每一帧数据都能送到正确的目的地。

物理链路与数据链路存在着概念性区别，物理链路是指一条中间没有任何交换结点的物理线段，是有线或无线的传输通路。而数据链路则具有逻辑上的控制关系，这是因为在相邻

计算机之间传输数据时，除了需要一条物理链路外，还必须有一些必要的规程或协议来控制这些数据的传输。把实现控制数据传输规程的硬件和软件加到物理链路上去，就构成了数据链路。因此，数据链路就好像一条将物理链路加以改造后的数字通道。当采用复用技术时，一条物理链路可以在逻辑上分解成多条数据链路。

3.4.2　帧的封装

数据链路层中传输的协议数据单元是帧，帧是逻辑的、结构化的数据块。上层的协议数据单元传到数据链路层后，数据链路层通过添加头部和尾部将数据封装成帧。数据链路层之所以要把比特流组合成帧进行传送，是为了出错时只重发有错的帧，而不必重发全部数据，另外，帧的头部和尾部含有数据链路层需要使用的协议信息。协议不同，帧的长短、语法、语义也有差别。对帧进行首尾定界，目的是识别出每一帧的开始与结束，保证相邻结点之间数据交换的同步，也就是所谓的“帧同步”问题。帧同步需要解决的问题是：接收方必须能够从物理层收到的比特流中准确地识别出帧的起始与终止位置。帧的封装（帧同步）有下面 4 种方法。

1．字节计数法

这种方法以一个特殊控制字符（例如 SOH）表示一帧的起始，并以一个专门的字节计数字段来标明帧内的字节数。接收方可以通过对该特殊控制字符的识别从比特流中区分出帧的起始位置，并从专门字段中获知该帧的数据字节数，从而确定出帧的终止位置。

在字节计数法中，“字节计数”字段十分重要，一旦“字节计数”字段出错，即缺失了帧边界划分的依据，将造成灾难性的后果。由于采用字段计数方法来确定帧的终止边界不会引起数据及其他信息的混淆，因而不必采用任何措施便可实现数据的透明性，即任何数据均可不受限制地传输。

2．字符填充的首尾定界法

该方法用特定的 ASCII 字符序列 DLE STX 和 DLE ETX 分别标识一帧的起始与终止。为了不使数据信息位中出现的与特定字符相同的字符被误判为帧的首尾定界符，可采用字符填充技术，即在这种数据字符前填充一个转义控制字符 DLE 以示区别，在接收端将成对的 DLE 丢掉一个，从而实现数据传输的透明性。

3．带比特填充的首尾标记法

该方法以一组特定的比特模式 (如 01111110) 来标志一帧的起始与终止。3.4.5 节介绍的 HDLC 规程即采用该法。

为了不使信息位中出现的与该特定模式相同的比特串被误判为帧的首尾标志，可以采用比特填充的方法。例如，采用特定模式 01111110，则对信息位中的任何连续出现的 5 个 1，发送方自动在其后插入一个 0，而接收方则进行该过程的逆操作，即每收到连续 5 个 1，则自动删去其后所跟的 0，以此恢复原始信息，实现数据传输的透明性。比特填充很容易由硬件来实现，性能优于字符填充方法。

4．违法编码法

该方法在物理层采用特定的比特编码方法时采用。例如，曼彻斯特编码方法，是将数据比特 1 编码成“高 - 低”电平对，将数据比特 0 编码成“低 - 高”电平对。而“高 - 高”电平对和“低 - 低”电平对在数据比特中是违法的。可以借用这些违法编码序列来定界帧的起始与终止。局域网 IEEE 802 标准中就采用了这种方法。违法编码法不需要任何填充技术，便能实现数据传输的透明性，但它只适合采用冗余编码的特殊编码环境。

以上 4 种方法有各自的优缺点，目前较普遍使用的是比特填充法和违法编码法。

3.4.3 差错控制

差错控制是数据链路层的主要功能之一，数据链路层采用检错和纠错技术，变不可靠的物理连接为可靠的数据链路，从而保证相邻结点的数据传输正确性。

检错的方法通常使用在第 2 章中介绍过的差错控制编码进行检错，由于检错码不能自动纠正所发现的错误，所以当接收方发现错误时，一般采取反馈重发机制来纠正错误，即发送方将要发送的数据帧附加一定的冗余检错码一并发送，接收方则根据检错码对数据帧进行差错检测，若发现错误，就返回请求重发的应答，发送方收到请求重发的应答后，便重新传送该数据帧。这种差错控制方法称为自动重发请求法 (Automatic Repeat reQuest)，简称 ARQ 法。

有时链路上的干扰严重，或由于其他原因，接收结点收不到发送结点的数据帧。这种情况称为数据帧丢失。此时，接收方当然不会向发送结点给出反馈应答帧。发送方因接收不到应答帧，将永远等待下去，于是就出现了死锁现象。同理，应答帧的丢失也同样会造成这种死锁现象。要解决死锁问题，需要引入计时器，可在发送完一个数据帧时，就启动超时定时器，若到了计时器所设置的重发时间而收不到来自接收方的任何应答帧，则发送方就重传前面所发送的数据帧。

然而问题并没有完全解决，当数据帧丢失时，超时重发没有问题。但当丢失的是反馈应答帧时，则超时重发将使接收方收到两个同样的数据帧，因而会产生重复帧的错误。要解决重复帧的问题，需要对每个数据帧进行编号，从而使接收方能根据数据帧的不同编号来区分是新发送的帧还是已被接收但又重新发送来的帧。

这样，数据链路层通过计时器和序号来保证每帧最终都能正确地交付给目标网络层一次。

实用的差错控制方法既要传输可靠性高，又要信道利用率高。ARQ 法仅需返回少量控制信息，便可有效地确认所发数据帧是否被正确接收。ARQ 法有两种最基本的实现方法：停止等待协议和连续重发请求协议。

1）停止等待（Stop and Wait）协议

该协议就是传好一帧再传下一帧。收发双方仅需设置一个帧的缓冲存储空间，即可有效地实现数据重发并确保接收方接收的数据不会重复。实现过程如下：

（1）发送方每次仅将当前信息帧作为待确认帧保留在缓冲存储器中。

（2）发送方发送一个信息帧后，就停止发送动作，随即启动计时器，等待反馈结果。

（3）当接收方收到无差错信息帧后，即向发送方返回一个确认帧。

（4）当接收方检测到一个含有差错的信息帧时，便舍弃该帧。

（5）若发送方在规定时间内收到确认帧，就将计时器清零，继而开始下一帧的发送。

（6）若发送方在规定时间内未收到确认帧，即计时器超时，则重发缓冲器中的待确认帧。

停止等待协议最主要的优点就是所需的缓冲存储空间最小，收发双方仅需设置一个帧的缓存空间；其缺点是发送方每帧都要停下来等待确认帧后再继续发送而造成信道浪费。因此该方法在链路端使用简单终端的环境中被广泛采用。

2）连续重发请求（Continuous RQ）

连续重发请求方案是指发送方可以连续发送一系列信息帧，即不用等前一帧被确认便可发送下一帧。这就需要在发送方设置一个较大的缓冲存储空间（称作重发表），用以存放若干待确认的信息帧。当发送方收到对某信息帧的确认帧后便可从重发表中将该信息帧删除。所以，实现连续重发请求协议的链路传输效率大大提高，但相应地需要更大的缓冲存储空间。实现过程如下：

（1）发送方连续发送信息帧而不必等待确认帧的返回。

（2）发送方在重发表中保存所发送的每个帧的备份。

（3）重发表按先进先出 (FIFO) 队列规则操作。

（4）接收方对每一个正确收到的信息帧返回一个确认帧。

（5）每一个确认帧包含一个唯一的序号，随相应的确认帧返回。

（6）接收方保存一个接收次序表，它包含最后正确收到的信息帧的序号。

（7）当发送方收到相应信息帧的确认后，从重发表中删除该信息帧的备份。

（8）当发送方检测出失序的确认帧（即第 N 号信息帧和第 N+2 号信息帧的确认帧已返回，而 N+1 号的确认帧未返回）后，便重发未被确认的信息帧。

上面连续 RQ 过程是假定在不发生传输差错的情况下描述的，如果出现差错，如何进一步处理还可以有两种策略，即回退 -N 策略和选择重发策略。

回退 -N 策略的基本原理是，当接收方检测出失序的信息帧后，要求发送方重发最后一个正确接收的信息帧之后的所有未被确认的帧；或者当发送方发送了 N 个帧后，若发现该 N 帧的前一个帧在计时器超时后仍未返回其确认信息，则该帧被判为出错或丢失，此时发送方就不得不重新发送出错帧及其后的 N 帧。这就是回退 -N 法名称的由来。因为，对接收方来说，由于这一帧出错，就不能以正常的序号向它的高层递交数据，对其后发送来的 N 帧也可能都不能接收而丢弃。回退 -N 策略操作过程如图 3-7 所示。图中假定发送完 8 号帧后，发现 2 号帧的确认帧 ACK 在计时器超时后还未收到，则发送方只能退回从 2 号帧开始重发。

回退 -N 策略可能将已正确传送到目的方的帧再重传一遍，这显然是一种浪费。另一种效率更高的策略是当接收方发现某帧出错后，其后继续送来的正确帧虽然不能立即递交给接收方的高层，但接收方仍可收下来，存放在一个缓冲区中，同时要求发送方重新传送出错的那一帧。一旦正确收到重传来的帧后，就可以和原已存于缓冲区中的其余帧一起按正确的顺序递交高层。这种方法称为选择重发（selective repeat）协议，其工作过程如图 3-8 所示。图中

2 号帧的否认返回信息 NAK2 要求发送方选择重发 2 号帧。显然，选择重发减少了浪费，但要求接收方有足够大的缓冲区空间。

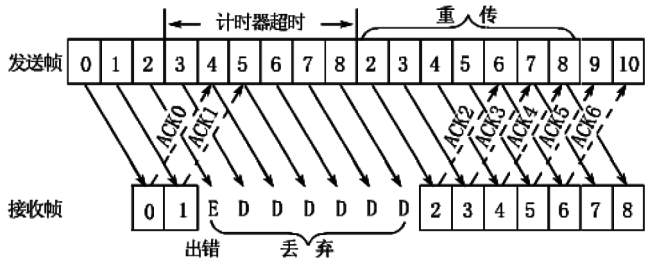


图 3-7　回退 -*N* 策略

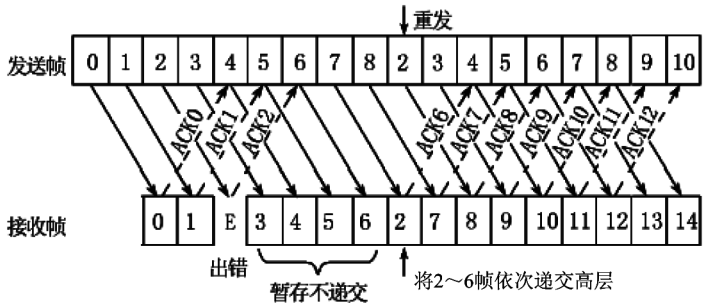


图 3-8　选择重发

3.4.4　流量控制

由于系统性能的不同，如硬件能力（CPU 速度、缓冲存储空间等）和软件功能的差异，会导致发送方与接收方处理数据的能力有所不同。当发送方以较快的发送速率发送，而接收方的接收速率较慢时，就会出现发送方发送能力大于接收方接收能力的现象，此时，接收方来不及接收的帧最终会被不断发送来的后续帧“淹没”，从而造成帧的丢失而出错。解决的办法是进行流量控制，协调发送方的发送速度或能力大于接收方的问题。流量控制实际上是控制发送方所发出的数据流量，使其发送速率不要超过接收方所能接收的速率，防止接收方被数据淹没。需要说明的是，流量控制并不是数据链路层特有的功能，许多高层协议中也提供流量控制功能，只不过流量控制的对象不同而已。例如，对于数据链路层来说，控制的是相邻两结点之间数据链路上的流量；在传输层上控制的是源端到目的端的流量。

实现流量控制的关键是需要有一种信息反馈机制，使发送方能了解接收方是否能接收到，常见的实现方法是窗口机制。

为了提高信道的有效利用率，如前所述采用了不等待确认帧返回就连续发送若干帧的方案。由于允许连续发送多个未被确认的帧，这就要求发送方有较大的发送缓冲区保留可能要求重发的未被确认的帧。但是缓冲区容量总是有限的，如果接收方不能以发送方的发送速率处理接收到的帧，则还可能用完缓冲容量而暂时过载。为此，可引入类似于停止等待的调整措施，其本质是在收到一个确认帧之前，对发送方可发送帧的最大数目加以限制。这是由发送方调整保留在重发表中的待确认帧的数目来实现的。如果接收方来不及对新到的帧进行处理，则停发确认帧，此时发送方的重发表就会增长，当达到重发表的最大限度时，发送方就不再发送新帧，直至再次收到确认信息为止。

为了实现此方案，发送方存放待确认帧的重发表中应设置待确认帧数目的最大限度，这一限度被称为发送窗口。显然，如果窗口设置为 1，即发送方缓冲能力仅为一个帧，则传输控制方案就回到了停止等待协议，此时传输效率很低。故窗口限度应选为使接收方尽量能处理或接收发送方发来的所有帧。当然选择时还必须考虑诸如帧的最大长度、可使用的缓冲存空间，以及传输速率等因素。

重发表是一个连续序号的列表，对应发送方已发送但尚未确认的那些帧。这些帧的序号有一个最大值，这个最大值即发送窗口的限度。所谓“发送窗口”就是指示发送方已发送但尚未确认的帧序号队列的界，其上、下界分别称为发送窗口的上沿、下沿，上、下沿的间距称为窗口尺寸。接收方类似也有接收窗口，它指示允许接收的帧的序号。

发送方每次发送一帧后，待确认帧的数目便增 1；每收到一个确认信息后，待确认帧的数目便减 1。当重发表长度计数值，即待确认帧的数目等于发送窗口尺寸时，便停止发送新的帧。

一般帧号只取有限位二进制数，到一定时间后就反复循环。若帧号配 3 位二进制数，则帧号在 0~7 之间循环。如果发送窗口尺寸取值为 2，如图 3-9 所示。图中发送方阴影部分表示打开的发送窗口，接收方阴影部分则表示打开的接收窗口。当传送过程进行时，打开的窗口位置一直在滑动，所以也称为滑动窗口（ sliding window ）协议。

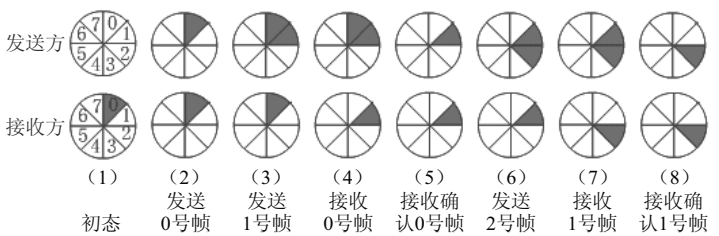


图 3-9　滑动窗口状态变化的过程

图 3-9 中的滑动窗口变化过程可叙述如下 (假设发送窗口尺寸为 2，接收窗口尺寸为 1)。

（ 1 ）初态，发送方没有帧发出，发送窗口上、下沿相重合。接收方 0 号窗口打开，表示等待接收 0 号帧。

（ 2 ）发送方已发送 0 号帧，此时发送方打开 0 号窗口，表示已发出 0 帧但尚无确认返回信息。此时接收窗口状态同前，仍等待接收 0 号帧。

（ 3 ）发送方在未收到 0 号帧的确认返回信息前，继续发送 1 号帧。此时，1 号窗口打开，表示 1 号帧也属等待确认之列。至此，发送方打开的窗口数已达规定限度，在未收到新的确认返回帧之前，发送方将暂停发送新的数据帧。接收窗口此时状态仍未变。

（ 4 ）接收方已收到 0 号帧，0 号窗口关闭，1 号窗口打开，表示准备接收 1 号帧。此时发送窗口状态不变。

（ 5 ）发送方收到接收方发来的 0 号帧确认返回信息，关闭 0 号窗口，表示从重发表中删除 0 号帧。此时接收窗口状态仍不变。

（ 6 ）发送方继续发送 2 号帧，2 号窗口打开，表示 2 号帧也纳入待确认之列。至此，发送方打开的窗口又已达规定限度，在未收到新的确认返回帧之前，发送方将暂停发送新的数据帧，此时接收窗口状态仍不变。

（7）接收方已收到 1 号帧，1 号窗口关闭，2 号窗口打开，表示准备接收 2 号帧，此时发送窗口状态不变。

（8）发送方收到接收方发来的 1 号帧收毕的确认信息，关闭 1 号窗口，表示从重发表中删除 1 号帧。此时接收窗口状态仍不变。

一般来说，凡是在一定范围内到达的帧，即使它们不按顺序到达，接收方也要接收下来。若把这个范围看成接收窗口，接收窗口的大小也应该是大于 1 的。实际上，停止等待协议是发送窗口等于 1 的滑动窗口协议的特例。

3.4.5 数据链路层协议举例

当相邻两个结点传递数据时，除了需要一条物理线路和设备外，还需要一些必要的通信规则来控制数据的传输，这些控制相邻结点间数据传输的通信规则就是数据链路层协议。

数据链路控制协议也称链路通信规程，也就是 OSI/RM 中的数据链路层协议。链路控制协议可分为异步协议和同步协议两大类。

异步协议（异步传输）规定以字符为独立的信息传输单位，在每个字符的起始处对字符内的位实现同步，但字符与字符之间的间隔时间是不固定的，即字符之间是异步的。由于每个传输字符都要添加诸如起始位、停止位等冗余位，故信道利用率很低，一般用于数据速率较低的场合。

同步协议（同步传输）规定以许多字符或许多比特组织成的数据块——帧为传输单位，在帧的起始处同步，帧内维持固定的时钟。由于采用帧为传输单位，所以同步协议能更有效地利用信道，也便于实现差错控制、流量控制等功能。同步协议又可分为面向字节计数的同步协议、面向字符的同步协议和面向比特的同步协议。其中，面向比特的同步协议的典型代表是高级数据链路通信规程（HDLC.。

HDLC 为英文 High Level Data Link Control 的缩写，称为高级数据链路控制协议，由 ISO 颁布，前身为 IBM 公司开发的 SDLC（Synchronous Data Link Control）。HDLC 协议的特点是面向比特，不依赖于任何一种字符编码集；实现透明传输的“0 比特插入 / 删除法”易于通过硬件实现；全双工通信，不必等待确认便可连续发送数据，有较高的数据链路传输效率；所有帧均采用 CRC 校验；对信息帧进行顺序编号，可防止漏收或重发，传输可靠性高等。

1. HDLC 帧格式

在 HDLC 中，数据和控制报文均以帧的标准格式传送。完整的 HDLC 帧由如图 3-10 所示的字段组成。其中：

※ 标志字段 F：也称为帧间隔符，用特殊比特串 01111110 标志帧的起始和终止。

※ 地址字段 A：通信方的地址，内容取决于所采用的操作方式。命令帧中的地址字段携带的是相邻结点的地址，而响应帧中的地址字段携带的是本结点地址。

※ 控制字段 C：表示帧的类型，8 位不同的编码构成各种命令和响应，以便对链路进行监视和控制。该字段是 HDLC 协议的关键部分。

※ 信息字段 I：表示传送的实际数据，下限可以为 0（无信息字段），上限未做严格限定，但实际上要受 FCS 字段或站点缓冲器容量的限制，一般是 1000 ～ 2000b。

※ 帧校验序列字段 FCS：可以使用 16 位或 32 位的 CRC，用于差错检测，对两个标志字段之间的整个帧的内容进行校验。

起始标志	地址	控制	信息	帧校验序列	结束标志
F 01111110	A 8位	C 8位	I *N*位	FCS 16位	F 01111110

图 3-10　HDLC 帧的标准格式

为了保证帧间隔符 01111110 的唯一性和帧内数据的透明性，保证 A（地址字段）、C（控制字段）、I（信息字段）、FCS（帧校验序列）中不出现 01111110 的位模式，HDLC 采用了 0 比特插入法。发送端发送 01111110 后，开始数据发送，并在数据发送过程中检查发送的位流，一旦发现连续的 5 个 1，则自动在其后插上 1 个 0，并继续传输后继的位流；数据发送结束后，追加帧间隔符 01111110。接收端执行相反的动作：一旦识别出帧间隔符 01111110 之后的位流不是 01111110，则启动接收过程；若识别出连续 5 个 1 和 1 个 0，则自动丢弃该 0，以恢复原来的位流；若识别出连续的 6 个 1，表示数据结束，该数据帧接收完成。

2. HDLC 帧的 3 种类型

HDLC 有信息帧（I 帧）、监控帧（S 帧）和无编号帧（U 帧）3 种不同类型的帧，各类帧中控制字段的格式及比特定义如图 3-11 所示。

控制字段位	1	2	3	4	5	6	7	8
I 帧	0	N(S)			P	N(R)		
S 帧	1	0	S1	S2	P/F	N(R)		
U 帧	1	1	M1	M2	P/F	M3	M4	M5

图 3-11　HDLC 的控制字段

控制字段中的第 1 位或第 1、第 2 位表示传送帧的类型。第 5 位是 P/F 位，即轮询 / 终止 (Poll/Final) 位。当 P/F 位用于命令帧（由主站发出）时，起轮询的作用，即当该位为 1 时，要求被轮询的从站给出响应，所以此时 P/F 位可称为轮询位（或 P 位）；当 P/F 位用于响应帧 (由从站发出) 时，称为终止位（或 F 位），当其为 1 时，表示接收方确认的结束。为了进行连续传输，需要对帧进行编号，所以控制字段中也包括了帧的编号。

（1）信息帧（I 帧）：控制字段第 1 位为 0，标志该帧是 I 帧。信息帧用于传送有效信息或数据，简称 I 帧。I 帧以控制字段中的第 2 ~ 4 位 N(S) 用于存放发送帧序号，以使发送方不必等待确认而连续发送多帧。N(R) 用于存放接收方下一个预期要接收的帧的序号，如 N(R)=5，即表示接收方下一帧要接收 5 号帧，换言之，5 号帧前的各帧接收方都已正确接收到。N(S) 和 N(R) 均为 3 位二进制编码，可取值为 0~7。

（2）监控帧（S 帧）：控制字段第 1、2 位为 10，标志该帧是 S 帧。监控帧用于差错控制和流量控制，简称 S 帧。S 帧不带信息字段，帧长只有 6B，即 48b。S 帧的控制字段的第 3、4 位为 S 帧类型编码，共有 4 种不同组合，含义分别如下：

00——接收就绪（RR），由主站可以使用 RR 型 S 帧来轮询从站，即希望从站传输编号为 N(R) 的 I 帧，若存在这样的帧，便进行传输；从站也可用 RR 型 S 帧来响应，表示从站期

望接收的下一帧的编号是 N(S)。

01——拒绝（REJ），由主站或从站发送，用以要求发送方对从编号为 N(R) 开始的帧及其以后所有的帧进行重发，这也暗示 N(R) 以前的 I 帧已被正确接收。

10——接收未就绪（RNR），表示编号小于 N(R) 的 I 帧已被收到，但目前正处于“忙”状态，尚未准备好接收编号为 N(R) 的 I 帧，这可用来对链路流量进行控制。

11——选择拒绝（SREJ），它要求发送方发送编号为 N(R) 的单个 I 帧，并暗示其他编号的 I 帧已经全部确认。

可以看出，接收就绪 RR 型 S 帧和接收未就绪 RNR 型 S 帧有两个主要功能：首先，这两种类型的 S 帧用来表示从站已准备好或未准备好接收信息；其次，确认编号小于 N(R) 的所有接收到的 I 帧。拒绝 REJ 和选择拒绝 SREJ 型 S 帧用于向对方站指出发生了差错。REJ 帧对应回退 -*N* 策略，用以请求重发 N(R) 起始的所有帧，而 N(R) 以前的帧已被确认，当收到一个 N(S) 等于 REJ 型 S 帧的 N(R) 的 I 帧后，REJ 状态即可清除。SREJ 帧对应选择重发策略，当收到一个 N(S) 等于 SREJ 帧的 N(R) 的 I 帧时，SREJ 状态即应消除。

（3）无编号帧（U 帧）：控制字段第 1、2 位为 11，标志该帧是 U 帧。无编号帧因其控制字段中不包含编号 N(S) 和 N(R) 而得名，简称 U 帧。U 帧用于提供对链路的建立、拆除以及多种控制功能，这些控制功能用于 5 个 M 位 (M1~M5，也称修正位) 来定义，可以定义 32 种附加的命令或应答功能。

3. HDLC 的工作过程

数据从发送到被接收的完整数据传输过程中，发送方和接收方要对传输操作进行一系列控制，主要包括请求与响应。这些控制是根据帧中各字段中位的数值变化来实现的。图 3-12 给出了 HDLC 用于有确认面向连接的服务的工作过程。

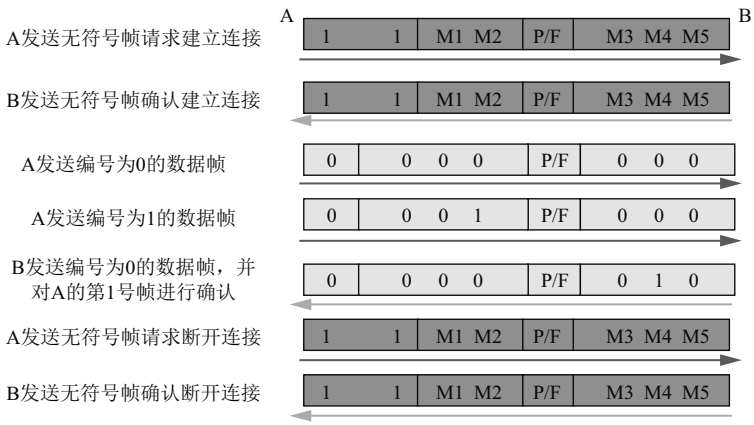


图 3-12　HDLC 的工作过程示意图

3.5　网络层

在 OSI/RM 中，网络层作为通信子网的最高层，关系到通信子网的运行控制，是通信子网中最为复杂、关键的一层。网络层介于数据链路层和传输层之间，以数据链路层提供的无

差错传输为基础，把高层发来的数据组织成分组，从源结点经过若干个中间结点传送到目的结点。传送过程要解决的关键问题是选择路径，在选择路径时还要考虑解决流量控制问题，防止网路中出现局部的拥挤或全面的阻塞。设置网络层的目的就是要为报文分组提供最佳路径，通过通信子网到达目的主机，实现两个端系统之间的数据的透明传送，主要功能包括路径选择、拥塞控制和网际互联等。

3.5.1　网络层提供的服务

两个端点之间的通信是依靠通信子网中的结点间的通信来实现的，既然网络层是通信子网中网络结点的最高层，所以网络层将体现通信子网向端系统所提供的网络服务。在 OSI/RM 中规定网络层提供面向连接的服务和无连接的服务两种类型的网络服务。在分组交换网络中，这两种网络服务的具体实现分别称为虚电路服务和数据报服务。

1．虚电路服务方式

所谓虚电路（virtual circuit），顾名思义，就是非实在性的电路。采用虚电路方式传输时，每个分组除了包含数据之外，还包含一个虚电路号，在预先建好的路径上的每个结点都知道把这些分组引导到哪里去，不再需要路由选择判定。

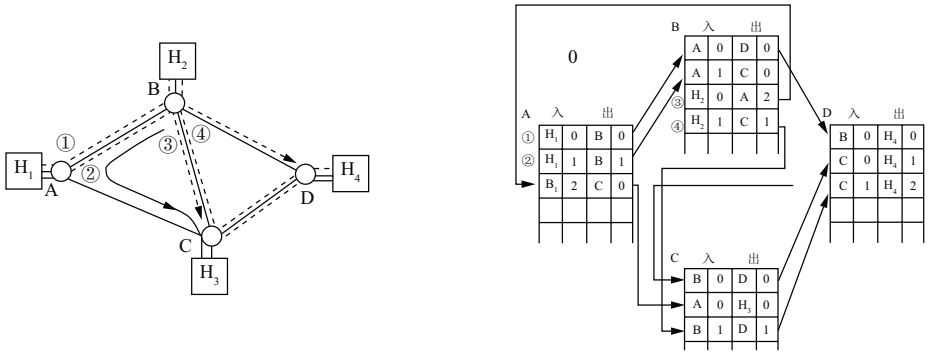
工作过程类似于电路交换，分虚电路建立、数据传送和释放 3 个阶段。不同之处在于电路交换自始至终固定占用一条物理链路，而虚电路是断续地占用一段一段的链路，此时的这条逻辑通路不是专用的，在每个结点上仍然采用“存储 - 转发”的方式处理分组，所以称之为虚电路。虚电路服务是以可靠的、面向连接的数据传送方式，向传输层提供的一种使所有分组按顺序到达目的结点的数据传输服务。面向连接是指在数据交换之前，必须先建立连接，当数据交换结束后，释放这个连接。

（1）虚电路建立。发送方发送含有地址信息的特定格式的呼叫分组，该分组除了包含源、目的地址外，还包含源端系统所选取的不用的最小虚电路号。该呼叫分组途经的每个中间结点根据当前的逻辑信道使用状况，分配虚电路号，并建立虚电路输入和输出映射表，即虚电路表。所谓主机之间建立虚电路，实际上就是在途经的各结点上填写虚电路表。

例如，图 3-13（a）所示的网络，其各结点的虚电路表如图 3-12（b）所示。当主机 H_1 的网络层收到传输层请求与主机 H_4 建立连接时，主机 H_1 的网络层发一个呼叫请求分组。呼叫请求分组在它通过通信子网到达目的主机的过程中，在所经过结点的虚电路表上登记入口和出口信息。入口信息的输入线就是前方结点名，虚电路号即是前方结点输出线的虚电路号。而出口信息的输出线可根据目标地址查输出线选择得到，其虚电路号则取该输出线当前尚未使用的最小虚电路号。

假设主机 H_1 发的呼叫请求分组在通信子网中沿路径① A → B → D 从 H_1 到达 H_4，沿途登记的入口出口信息如图 3-12（b）所示。由于这条路径是最先建立的，因此各结点虚电路表中的虚电路号均为 0。同理，按序可建立② A → B → C → D、③ B → A → C、④ B → C → D 三条虚电路，这对单工通信是可行的。但对双工通信，还必须保证两个相邻结点之间正、反两个方向的两条虚电路不能混淆。为此，不仅要考虑与下一结点之间的虚电路号不相同，还要考虑与下一结点在另一条虚电路上作为上结点时所选取的虚电路号相区别。因此，第③条虚电路结点 B 的出口虚电路号取 2，而不取 0 或 1。这样做后，不同虚电路的分组虽然从同一

条输入线进入了同一个结点，且可能从同一输出线输出，但它们却不会有相同的输出虚电路号。



(a)网络　　(b)虚电路表

图 3-13　虚电路的建立

当呼叫请求分组到达目的主机（如 H_4），若目的主机同意通信，就发同意通信的应答给源主机，则虚电路建立阶段结束。

（2）数据传送。虚电路一旦建立以后，在传输中，当一个分组到达结点时，结点根据其携带的虚电路号查找虚电路表，以确定该分组应发往下一段信道上所占用的虚电路号，用该虚电路号替换分组中原先的虚电路号后，再将该分组发往下一个结点。这样，分组上面就不需要目的主机的网络地址，只要带虚电路号即可。

（3）虚电路的释放。各结点的虚电路表空间和虚电路号都是网络资源，当虚电路拆除时必须回收。这可通过某端系统发出一个拆链请求分组，告知虚电路中各结点删除虚电路表中有关表项来实现。

虚电路有永久性和交换型的虚电路两种。永久性虚电路（PVC）是一种提前定义好的，基本上不需要任何建立时间的端点站点间的连接；交换型虚电路（SVC）是端点站点之间的一种临时性连接。虚电路服务方式主要的特点如下：

※ 在每次分组传输前，都需要在源结点和目的结点之间建立一条逻辑连接。

※ 一次通信的所有分组都通过虚电路顺序传送，因此分组不必自带目的地址、源地址等信息。分组到达结点时不会出现丢失、重复与乱序的现象。

※ 分组通过虚电路上的每个结点时，中间结点只需要进行差错检测，而不需要进行路由选择。

※ 通信子网中每个结点可以与任何结点建立多条虚电路连接。

虚电路服务方式的优点是：端到端的差错控制由通信子网负责，可靠性高，网络层保证分组按顺序交付，不丢失，不重复；其缺点是：如有故障，则经过故障点的数据全部丢失。这种服务方式适用于数据量大、可靠性要求较高的场合。

2．数据报服务方式

数据报服务是无连接的数据传送方式，免去了虚电路方式的虚电路建立阶段，工作过程类似于报文交换。通信双方在开始通信之前不需要先建立连接，因此被称为无连接。

采用数据报方式传输时，被传输的分组称为数据报。当端系统从传输层接收到要发送的一个报文时，网络层将报文拆成若干个带有序号和地址信息的数据报，依次发给网络结点。网络结点接收到一个数据报后，根据数据报中地址信息和结点存储的路由信息，找出一条合适的出路，把数据报原封不动地传送到下一个结点，依此类推，直至目的结点。

从数据报服务方式中可以看出，在整个数据报传送过程中，不需要建立连接，但网络结点要为每个数据报进行路由选择。目标结点收到数据后也不需发送确认，因而是一种开销较小的通信方式。数据报服务方式的主要特点如下：

※ 同一报文的不同分组可以经过不同的传输路径通过通信子网。

※ 每个分组在传输过程中都必须带有目的地址与源地址。

※ 同一报文的不同分组到达目的结点时可能出现乱序、重复与丢失现象。需要在目的结点开辟缓冲区，缓存所有收到的分组，然后重新排序后按发送顺序交付给主机。

※ 由主机承担端到端的差错控制。

※ 传输过程延迟大，适用于突发性通信，不适用于长报文、会话式通信。

数据报具有健壮性和灵活性的优点，在传输途中，若某个结点或链路发生故障，数据报服务可以绕过故障把分组传送到目的结点。

3. 虚电路服务和数据报服务的比较

※ 在传输方式上：虚电路服务需要连接建立和释放的过程；而数据报服务，网络层从传输层接收报文分组，附加上源、目的地址等信息后独立传送，不需建立和释放连接。

※ 网络地址：虚电路服务仅在源主机发出呼叫分组中需要填上源和目的主机的网络地址，在数据传输阶段只需填上虚电路号，不需要网络地址。而数据报服务，由于每个数据报都单独传送，因此，在每个数据报中都必须具有源和目的主机的网络地址。

※ 路由选择：在数据报方式中，每个网络结点都要为每个分组路由做出选择；而在虚电路方式中，只需在连接建立时确定路由。

※ 分组顺序：虚电路服务的所有分组都是通过事先建立好的一条虚电路进行传输，所以能保证分组按发送顺序到达目的主机。但是，当把一份长报文分成若干个短的数据报时，由于它们被独立传送，可能各自通过不同的路径到达目的主机，因而数据报服务不能保证这些数据报按顺序到达目的主机。

※ 可靠性和适应性：虚电路服务在通信之前双方已进行过连接，而且每发完一定数量的分组后，对方也都给予确认，故虚电路服务比数据报服务的可靠性高。但是，当传输途中的某个结点或链路发生故障时，数据报服务可以绕开这些故障地区，另选其他路径，把数据传至目的地；而虚电路服务则必须重新建立虚电路才能进行通信。因此，数据报服务的适应性比虚电路服务强。

综上所述，虚电路适合于大批量的数据传输、交互式通信，不仅及时、传输较为可靠，而且网络开销小。数据报方式更适合于站点之间少量数据的传输。

3.5.2 路由选择

通信子网中源结点和目的结点之间存在多条传输路径的可能性。网络结点在收到一个分组后，根据一定的原则和算法确定向下一个结点传送的最佳路径，这就是路由选择。

路径的选择需要相应的路由选择算法来实现，路由选择算法就是网络结点用于决定达到目的网络的最佳路径的计算方法。网络上的路由器通过路由选择算法形成路由表，以确定发送分组的传输路径。路由算法应具备的特性有正确性、简单性、健壮性、稳定性、公平性和最优性，即一个理想的路由选择算法应该是正确、简单且易实现的，不增加额外开销，能适应通信量和网络拓扑结构的变化，公平地保证每个结点都有平等的机会传送数据。一个实际的路由选择算法应尽可能接近于理想的算法。根据路由算法能否依靠网络当前的流量和拓扑结构来调整它们的路由决策，路由选择算法分为静态路由选择算法和动态路由选择算法。

1. 静态路由选择算法

静态路由选择算法不用测量也无须利用网络信息，按某种固定规则进行路由选择，也称作非自适应算法。非自适应算法可分为以下几种。

※ 泛射路由选择。也称泛洪法，是一种最简单的路由选择算法。一个网络结点从某条线路收到一个分组后，再向除该条线路外的所有线路重复发送收到的分组。显然，最先到达目的结点的一个或若干个分组肯定经过了最短路径。实际应用中很少采用洪泛法，这是因为采用这种方法后，网络中的分组数目会迅速增长，会导致网络出现拥塞现象。这种方法可用于诸如军事网络等健壮性要求很高的场合，即使有的网络结点遭到破坏，只要源、目标间有一条信道存在，则泛射路由选择仍能保证数据的可靠传送。另外，这种方法也可用于将一个分组从数据源传送到所有其他结点的广播式数据交换中，它还可用来进行网络的最短传输延迟的测试。

※ 固定路由选择。这种方法是在每个网络结点存储一张预先确定好的路由表，该表格记录从本结点到某个目的结点的最短路径。当一个分组到达某结点时，该结点可根据分组中的目的地址，从路由表中查出其输出线。固定路由选择的路由表是由网络管理人员在网络运行前确定并建立的，路由表中的每一条信息都是手工配置的。路由表一旦建立，在运行中一般不会改变，当网络拓扑结构发生变化或路由器出现故障时，它都不能自动更新路由表，除非网络管理员重新配置它。固定路由选择法的优点是简便易行，负载稳定，适合在拓扑结构变化不大的小型网络中应用；它的缺点是灵活性差，无法应付网络中发生的拥塞和故障。

※ 随机路由选择。在这种方法中，收到分组的结点在所有与之相邻的结点中为分组随机选择一个出路结点。该方法虽然简单，也较可靠，但实际路由可能不是最佳路由，增加了不必要的负担，而且分组传输延迟也不可预测，故应用不广。

2. 动态路由选择算法

在实际网络中网络结点众多，随时都有结点开始和停止工作，网络的拓扑结构随时都有可能变化，各结点的通信请求不可预知，网络上的负载是变化的。静态路由算法不能根据网络流量和拓扑结构的变化来调整自身的路由表。动态路由选择算法是指结点根据当前网络的

状态信息，自动计算最佳路径，建立路由表，而且能够自动适应网络的故障、拓扑结构的变化，动态地更新路由表。这种策略能较好地适应网络流量、拓扑结构的变化，有利于改善网络的性能。但由于算法复杂，会增加网络的负担。现代计算机网络中普遍使用的是动态路由选择算法，动态路由选择算法是动态路由协议的依据。

为了能够动态地适应网络拓扑结构等网络状态的变化，结点间必须交换网络状态信息。每个结点获得网络状态信息有 3 种来源：本地、相邻结点和所有结点。相应地可以把动态策略分为 3 种：孤立路由、分布路由、集中路由。其中，分布路由是目前普遍使用的一种方法，采用分布路由的网络，网络中每个结点根据来自相邻结点的信息，通过最短花费算法计算出到每个目的结点的路径，更新自己的路由表。实现分布路由选择的动态路由选择算法有距离向量路由算法和链路状态路由选择算法。

1）距离向量路由算法（distance vector routing）

距离向量路由算法的基本要素是距离和向量。距离是最优距离度量值，度量值可以是从源结点到目的结点的路径上经过的路由器的个数，即跳数（hop），也可以用多种权值来综合度量。向量是指从源结点到达目的网络的路径，即它的下一跳路由器，也就是在转发数据分组时首先要传给的那个相邻路由器。

距离向量路由算法的基本思想是以某一结点到目的结点的距离作为算法的度量，每个结点（路由器）均存储一张路由选择表，表中记录本结点到达每个已知目标结点的最优距离度量值和路径。结点接收到数据分组后，根据分组头部的目的地址来查找路由表，并将其转发到下一跳所指定的结点。结点根据路由选择协议，通过与邻居之间相互交换路由表，对本身的路由表的距离度量值进行检查，如距离度量值有变化，就重新计算更新路由表。

距离向量路由算法的优点是简单、易于实现。但在实现时有明显缺陷，那就是要求所有结点都参与路由表信息的交换，而且不管网络是否发生了变化，都要定期地向相邻结点发送整个路由表。由于每次路由表的更新都是仅在相邻结点之间进行的，然后再由相邻结点向它的相邻结点传送，这样一级一级地传播下去，最终完成互联网所有结点的更新，这需要一定的时间。因此，距离向量算法交换信息量大，更新过程长，收敛速度慢，并且在刷新的过程中容易发生远近路由器路径不一致的问题。所谓“收敛”是指直接或间接交换路由信息的一组结点在网络的路由信息方面达成一致。另外，为了避免无限记数问题，距离向量算法对经过路由器的跳数有限制。采用距离向量算法的 RIP（路由信息协议）规定全程最多不能超过 15 跳，超过该值就被认为路径不可到达，这也是距离向量算法不适合在大型互联网的环境中应用的主要原因。因此，出现了另一种全新的算法——链路状态路由选择算法。在 1979 年以前，ARPAnet 一直使用的是距离向量路由算法，而在此后，则被替换为链路状态路由选择算法。

2）链路状态路由选择算法（link-state routing）

链路状态路由选择算法也称为最短路径优先算法（Shortest Path First，SPF），它克服了距离向量算法交换信息量大、收敛速度慢等不足，是一种更适合大型网络环境应用的路由选择算法。

这种算法需要每个结点都保存一份最新的关于整个网络链路状态信息的网络拓扑结构数据库，也称作链路状态数据库。数据库记录网络中每个结点的链路状态。链路状态指的是结

点间的邻接关系和链路代价，包括相邻结点的名称、状态，以及到达这个结点的延迟时间等。这里的结点就是连接网络的路由器。每个结点都会产生一些关于自己、本地直接连接链路的状态和所有直连邻居结点的信息。这些信息从一个结点传送到另一个结点，每个结点都做一份信息副本，但是决不改动这些信息，最终每个结点都有一个相同的网络信息的链路状态数据库。通过这个数据库，每个结点可以独立地计算各自的最优路径并产生路由表。如果把距离向量路由选择比作是由路标提供的信息，那么链路状态路由选择就是一张交通线路图，因为它有一张完整的网络图，所以不容易被欺骗而做出错误的路由决策。

在链路状态路由选择算法中，最重要的是保证网络上所有的路由器能够得到必要的链路状态信息，以保证路由表的及时更新。为此，每个路由器必须完成以下工作：

※ 发现邻居。通过网络发送链路状态询问报文 Hello 分组来实现相邻结点探查，相邻结点则返回一个应答，告知它是谁，同时返回它的链路状态信息。

※ 发送链路状态通告。网络中各结点在链路状态改变时，通过泛洪法的方式广播发送链路状态通告（LSA. 到网络中的所有结点。

※ 计算新的路由。一旦结点收到所有的链路状态通告，每个结点将构造区域中的网络拓扑结构图。然后，路由器根据结构图在本地运行最短路径算法（Dijkstra 算法），计算出到达所有可能目的结点的最短路径，完成路由表的更新。

链路状态路由选择算法对于每个结点发送路由信息到网络上所有的结点时，仅发送它的路由表中描述了其自身链路状态的那一部分。而距离向量路由算法则要求每个结点发送其路由表全部或部分信息，但仅发送到邻近结点上。从本质上来说，链路状态路由选择算法将少量的更新信息发送至网络各处，而距离向量路由算法发送大量的更新信息至相邻结点，再由相邻结点传给其相邻结点。因此链路状态路由选择算法收敛更快，在一定程度上比距离向量路由算法更不易产生路由循环。但另一方面，链路状态路由选择算法要求比距离向量路由算法有更强的 CPU 计算能力和更多的内存空间，因此链路状态路由选择算法将会在实现时显得更昂贵一些。

链路状态路由选择算法在实际网络中得到广泛的应用，在因特网中使用的 OSPF（开放路径优先）协议使用的就是链路状态路由选择算法。

3.5.3 拥塞控制

网络中多个层次都存在流量控制问题，网络层的流量控制是对进入通信子网的通信量加以控制，以防止因通信量过大造成通信子网性能下降。

1. 拥塞和死锁

拥塞现象是指到达通信子网中某一部分的分组数量过多，使得该部分网络来不及处理，以致引起这部分乃至整个网络性能下降的现象。当主机发送到通信子网中的分组数量在其传输容量之内时，能被送达目的端。但当通信量增加太快时，使得路由器不能及时处理，开始丢弃分组，由于丢弃分组而带来大量的重发分组，导致情况进一步恶化，严重时甚至会导致网络通信陷入停顿，即出现死锁现象。

通信子网吞吐量和通信子网负荷之间一般有如图 3-14 所示的关系。当通信子网负荷，即

通信子网正在传输的分组数比较少时，网络的吞吐量（单位为分组数 /s）随网络负荷的增加而线性增加。当网络负荷增加到某一值后，网络吞吐量反而下降，则表征网络中出现了拥塞现象。在一个出现拥塞现象的网络中，到达一个结点的分组将会遇到无缓冲区可用的情况，从而使这些分组不得不由前一结点重传，或者需要由源结点或源端系统重传。当拥塞比较严重时，通信子网中相当多的传输能力和结点缓冲器都用于这种无谓的重传，从而使通信子网的有效吞吐量下降，由此导致恶性循环，使通信子网的局部甚至全部处于死锁状态，网络有效吞吐量接近于零。

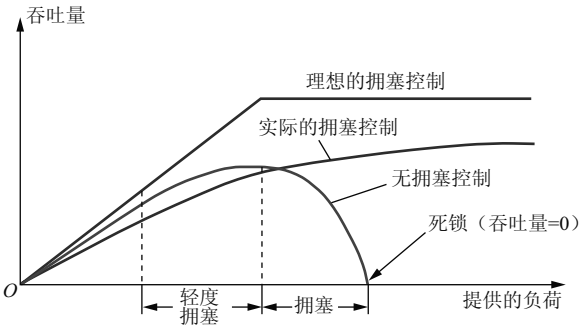


图 3-14　拥塞引起的性能下降情况

OSI/RM 中多个层次都存在流量控制问题，网络层的流量控制则对进入分组交换网的通信量加以一定的控制，以防止因通信量过大造成通信子网性能下降。拥塞控制用于确保通信子网能运送所有待传数据，是全局性的问题，涉及所有主机、路由器，并与路由器的存储转发能力和其他影响通信子网负荷的因素有关。数据链路层流量控制只涉及发送者和接收者之间的点对点通信的局部问题，其任务是确保快速的发送者不要以高于接收者所能承受的速率发送数据。

2. 拥塞控制方法

为防止出现拥塞和死锁现象，可采用预先分配缓冲区资源、准许结点在必要时丢弃分组、限制进入通信子网的分组数等方法。

1）缓冲区预分配法

这种方法用于虚电路分组交换网中。在建立虚电路时，让呼叫请求分组的途径结点为虚电路预先分配一个或多个数据缓冲区。若某个结点缓冲区已被占满，则呼叫请求分组另择路由，或者返回一个“忙”信号给呼叫者。这样，通过途经的各个结点为每条虚电路开设的永久性缓冲区（直到虚电路拆除），就总能有空间来接纳并转送经过的分组。此时的分组交换与电路交换很相似。当结点收到一个分组并将它转发出去之后，该结点向发送结点返回一个确认信息。该确认一方面表示接收结点已正确收到分组，另一方面告诉发送结点，该结点已空出缓冲区以备接收下一个分组。上面是“停止 - 等待”协议下的情况，若结点之间的协议允许多个未处理的分组存在，则为了完全消除拥塞，每个结点要为每条虚电路保留等价于窗口大小数量的缓冲区。这种方法不管有没有通信量，都有可观的资源（线路容量或存储空间）被某个连接占有，因此网络资源的有效利用率不高。这种控制方法主要用于要求高带宽和低延迟的场合，例如传送数字化语音信息的虚电路。

2）分组丢弃法

这种方法不必预先保留缓冲区，当缓冲区占满时，将到来的分组丢弃。若通信子网提供的是数据报服务，则用分组丢弃法来防止拥塞，拥塞发生时也不会引起大的影响。但若通信子网提供的是虚电路服务，则必须在某处保存被丢弃分组的备份，以便拥塞解决后能重新传送。有两种解决被丢弃分组重发的方法，一种是让发送被丢弃分组的结点超时，并重新发送分组直至分组被收到；另一种是让发送被丢弃分组的结点在一定次数后放弃发送，并迫使数据源结点超时而重新开始发送。但是不加分辨地随意丢弃分组也不妥，因为一个包含确认信息的分组可以释放结点的缓冲区，若因结点无空余缓冲区来接收含确认信息的分组，这便使结点缓冲区失去了一次释放的机会。解决这个问题的方法是：可以为每条输入链路永久地保留一块缓冲区，以用于接纳并检测所有进入的分组，对于捎带确认信息的分组，在利用了所捎带的确认信息释放缓冲区后再将该分组丢弃，或将该捎带确认消息的分组保存在刚空出的缓冲区中。

3）定额控制法

这种方法在通信子网中设置适当数量的被称为许可证的特殊信息，一部分许可证在通信子网开始工作前预先以某种策略分配给各个源结点，另一部分则在子网开始工作后在网中四处环游。当源结点要发送来自源端系统的分组时，它必须首先拥有许可证，并且每发送一个分组便注销一张许可证。目的结点方则每收到一个分组并将其递交给目的端系统后，便生成一张许可证。这样便可确保子网中分组数不会超过许可证的数量，从而防止拥塞的发生。

3.5.4 网络互联

网际互联的目的是使一个网络上的某一主机能够与另一网络上的主机进行通信，即让一个网络上的用户能访问其他网络上的资源，可使不同网络上的用户互相通信和交换信息。若互联的网络都具有相同的协议结构，则互联的实现比较容易。OSI/RM 正是为达到这一境界而提出的。但是在实际应用中存在着大量采用不同体系结构和协议的异构网。对于异构网（如各种类型的局域网）来说，在分组长度、寻址方式、超时控制、差错恢复方法等多方面存在着很大差异，因此，当被传送的分组需要跨越一个网络边界时，网络层应该对不同网络中的这些差异进行转换，消除网络间的差异，使异构网之间能够互联。有关网络互联的知识将在以后的章节中详细介绍。

3.5.5 网络层协议举例

典型的网络层协议是 X.25 协议，X.25 是国际电报电话咨询委员会（CCITT）提出的对于分组交换网（Packet-Switched Network，PSN）的标准访问协议，描述了主机（DTE）与分组交换网之间的接口标准，使主机不必关心网络内部的操作就能方便地实现对网络的访问。X.25 实际上是 DTE 与 PSN 之间的一组接口协议，它包括物理层、数据链路层和分组层 3 个层次，其中分组层相当于 OSI/RM 中的网络层。X.25 协议最主要的功能是向主机提供多信道的虚电路服务。

1. X.25 分组层的功能

X.25 分组层的主要功能是将链路层所提供的连接 DTE-DCE 的一条或多条物理链路复用

成数条逻辑信道，并且对每一条逻辑信道所建立的虚电路执行与数据链路层单链路协议类似的链路建立、数据传输、流量控制、顺序和差错检测、链路拆除等操作。利用 X.25 分组层协议，可向网络层的用户提供多个虚电路连接，使用户可以同时与公用数据网中若干个其他 X.25 数据终端用户（DTE）通信。X.25 提供虚呼叫和永久虚电路两种虚电路服务。

2．X.25 分组层分组格式

在分组层上，所有的信息都以分组为基本单位进行传输和处理，无论是 DIE 之间所要传输的数据还是交换网所用的控制信息都以分组形式来表示，并按照链路协议穿越 DTE-DCE 接口进行传输。因此在链路层上传输时，分组应嵌入到信息帧（I 帧）的信息字段中。每个分组均由分组头和数据信息两部分组成，其一般格式如图 3-15 所示。

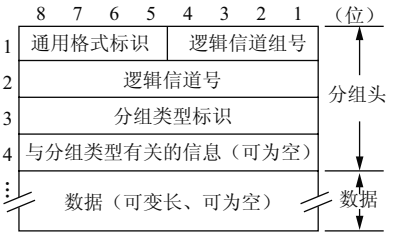


图 3-15　X.25 分组层分组格式

分组格式的数据部分通常被递交给高层协议或用户程序去处理，所以分组协议中不对它做进一步规定。分组头用于网络控制，主要包括 DTE/DCE 的局部控制信息，其长度随分组类型不同，但至少要包含前 3 个字节作为通用格式标识、逻辑信道标识和分组类型标识，它们的含义如下。

※ 通用格式标识（GFI）。由分组中第 1 个字节的前 4 位组成，用于指出分组头中其余部分的格式。第 1 位（b8）称作 q 位或限定位，只用于数据分组中。这是为了对分组中的数据进行特殊处理而设置的，可用于区分数据是正常数据还是控制信息。对于其他类型的分组，该设置为 0；第 2 位（b7）称作 d 位或传送确认位，设置该位的目的是用来指出 DTE 是否希望用分组接收序号 P(R) 来对它所接收数据做端到端确认。在呼叫建立时，DTE 之间可通过 d 位来商定虚电路呼叫期间是否使用 d 位规程；第 3、第 4 位（b6、b5）用以指示数据分组的序号是用 3 位即模 8（b6 置 1）还是 7 位即模 128（b5 置 1），这两位或者取 10，或者取 01，一旦选定，相应的分组格式也有所变化。

※ 逻辑信道标识。由第 1 个字节中的剩余 4 位（b4、b3、b2、b1）所作的逻辑信道组号（LCGN）和第 2 个字节所作的逻辑信道号（LCN）两部分组成，用以标识逻辑信道。每条虚电路都要赋予一个虚电路号，X.25 中的虚电路号由逻辑信道组号（0 ～ 15）和逻辑信道号（0 ～ 255）组成。用于虚呼叫的虚电路号范围和永久虚电路的虚电路号应在签订服务协议时与管理部门协商确定并分配。

※ 分组类型标识。由第 3 个字节组成，用于区分分组的类型和功能。若该字节的最后一位（b1）是 0，则表示分组为数据分组；若该位是 1，则表示分组为控制分组，其中可包括呼叫请求或指示分组及释放请求或指示分组。若该字节末 3 位（b3、b2、b1）均为 1，则表示该分组是某个确认或接受分组。第 4 个字节及其后各字节将依据分组类型的不同而有不同的定义。

X.25 分组层协议还规定了其他多种类型的分组，包括释放请求 / 指示、复位请求 / 指示、重启动请求 / 指示，在这里就不进行详述了。

3.6 传输层

1. 传输层的地位和作用

传输层也称为运输层，只存在于通信子网以外的主机中。它是整个协议层次结构的核心，是唯一负责源端到目的端对数据传输和控制的一层。

传输层介于低 3 层通信子网和高 3 层之间，它接收来自高 3 层的用户信息并交给网络层进行传输。传输层之上的会话层、表示层及应用层不包含任何数据传输的功能，传输层之下的网络层代表的是通信子网。有一个既存事实，即世界上各种通信子网在性能上存在着很大差异。例如电话交换网、分组交换网、公用数据交换网、局域网等通信子网都可互联，但它们提供的吞吐量、传输速率、数据延迟等通信费用各不相同。对于会话层来说，却要求有一个性能恒定的界面。因此在网络层之上加一层即传输层来承担这一功能，它采用分流 / 合流、复用 / 介复用技术来屏蔽通信子网在这些方面的细节与差异，使会话层感受不到通信子网的差异。有了传输层，用户在通信时就不必知道通信子网的构成及线路质量等，也不必考虑子网是局域网还是分组交换网，在传输数据时也不必关心数据传送方法的细节。传输层的存在使高层用户看见的就好像是在两个传输实体间已有一条已经建立好的端到端的可靠的通信通路，但传输层不对所传送的数据进行处理。

计算机网络通信是主机中应用进程间的数据通信，数据到达指定的主机后，还必须交给相应的应用程序。主机中可能会同时运行多个应用程序（如 Web 服务、FTP 服务等），因此数据中必须有相应的标识（端口号），以保证正确传送到对应的应用程序。所以，传输层完成进程间的数据传送，实际上是按端口号找到进程进行通信。网络层只是根据网络地址将源主机发出的分组传送到目的主机，而传输层负责将数据可靠地送到相应的端口。在生活中，邮包送到邮局类似于网络层的功能，而邮递员将信件交到收件人手中类似于传输层的功能。

设置传输层的两个主要作用是：第一，负责可靠的端到端通信，所谓“端到端”就是进程到进程；第二，向会话层提供独立于网络的传输服务。

2. 传输层的功能

传输层提供了主机应用进程之间的端到端服务，其主要功能是：为一个进行的会话或连接提供可靠的传输服务，完成端到端的通信链路的建立、维护和管理；在单一连接上提供端到端的端口号、流量控制以及差错恢复等服务。针对用户端的需求，采用一定的手段，屏蔽不同网络的性能差异，使用户无须了解网络传输的细节，获得相对可靠的数据传输服务。其基本功能如下：

※ 分割与重组数据。在发送方，传输层将从会话层来的数据分割成较小的数据单元，并在数据单元的头部加上一些控制信息后，形成报文。报文头部包含源端口号和目的端口号。在接收方，数据经过通信子网到达传输层时，将各报文中的头部控制信息去掉，然后按正常的顺序重组，还原为原来的数据，送交给会话层。

※ 按端口号寻址。端点是与网络地址对应的，但同一端点上可能有多个应用进程，它们在同一时间内进行通信。传输层则通过端口号寻址到端点上的不同进程，并使用多路复用技术处理多端口同时通信的问题。

※ 连接管理。面向连接的传输服务需要建立、维持和释放连接。

※ 差错控制和流量控制。传输层要向会话层提供通信服务的可靠性，避免报文的出错、丢失、延迟时间紊乱、重复、乱序等差错。因此要提供端到端的差错控制和流量控制，传输层的数据将由目标端点进行确认，如果源端点在指定的时间内未收到确认信息，将重发数据。传输层还具有流量控制的作用，使用窗口技术控制发送端口的速率，使其不要超过接收端口所能承受的范围。

3．传输层的服务类型

传输服务有两大类，即面向连接的服务和面向无连接的服务。面向连接的服务提供传输服务用户之间逻辑连接的建立、维持和拆除，是可靠的服务，可提供流量控制、差错控制和序号控制；而无连接的服务只能提供不可靠的服务。

传输层利用网络层提供给它的服务开发本层的功能，实现本层对会话层的服务。通过对网络层的学习我们已经知道，网络层提供面向连接的和无连接的服务两种形式，传输层也提供类似的面向连接和无连接的传输服务。也就是说，传输服务和网络服务十分相似。既然两种服务如此类似，为什么不直接利用网络层服务完成所有的功能，还需要传输层服务呢？这是因为网络层代表的是通信子网，提供的是数据报和虚电路服务，而这两种服务提供的服务质量是有差异的、不可靠的。对于数据报，网络层无法保证分组无丢失、无重复，无法保证分组按顺序从发送端到接收端。对于虚电路服务，虽然可以保证分组无差错、无丢失和无重复，以及按顺序到达，但在这种情况下，也不能保证网络服务能达到 100% 的可靠。对于这种情况，用户将束手无策，因为用户不能对通信子网加以控制，无法采用更优的通信处理机来解决网络服务质量的问题，更不能通过改进数据链路层纠错能力来改善它。解决这个问题的唯一可行办法就是在网络层上增加一层传输层。传输层的存在使传输服务比网络服务更可靠，分组丢失、残缺甚至网络的复位均可被传输层检测出来，并采取相应的补救措施。

4．服务质量

服务质量（Quality of Service，QoS）是传输层性能的度量，它衡量传输层的总体性能，反映传输质量及服务的可用性。在传输层中，要求服务质量达到一定的高度，从另一个角度看，传输层服务质量可以看作是网络层服务的增强。如果网络层服务质量比较完备，则传输层可以少做一些工作，实现比较简单；相反，如果网络层服务质量比较差，那么就要求传输层实现比较复杂的功能才能达到传输层服务质量的要求。

服务质量可用一些参数来描述，如连接建立延迟、连接建立失败概率、吞吐率、传输延迟、残留差错率、连接拆除延迟、连接拆除失败概率、传输失败率，等等。

（1）连接建立延迟。指从传输服务用户发出连接请求到连接建立成功之间的时间。这个延迟时间越短，服务质量就越高。

（2）连接建立失败概率。指最大延迟时间内连接未能建立的可能性。连接未能建立可能由于各种因素，如网络拥塞、缓冲区不足等。这个概率当然越小越好。

（3）吞吐率。指在一个时间段内，在一条传输连接上传输的数据字节数。

（4）传输延迟。指从源端开始传输数据到数据被目的端收到为止的时间。

（5）残留差错率。指传输连接上错误的报文数占全部传输的报文数的比例。

传输层的服务质量参数值通常由传输服务用户在请求建立连接时指定，一般需要指出期望值和最低可接受的值。传输实体在收到这个连接请求时会有下面两种情况。

一种是传输实体马上就能判断这个 QoS 是不可能达到的。此时，传输实体可能甚至不去与目的传输实体连接就马上给传输服务用户发回连接请求失败的信息，并且指明因为哪种服务质量不能达到标准而造成了连接失败；另一种是如果传输实体不能到达期望的服务质量，但是可以到达服务质量高于最低可接受的服务质量。此时，传输实体向目的传输实体发出连接请求，同时传递相应的 QoS 参数值，例如吞吐率参数的期望值为 500Mb/s，最低可接受值为 160Mb/s。如果目的传输实体只可以实现的吞吐率参数值为 300Mb/s，则目的实体以它可实现的 QoS 参数值 300Mb/s 来响应这个连接请求。这个过程称为用户与传输服务提供者之间的协商服务质量。主呼叫用户请求的服务质量可能被传输服务者、提供者降低，也可能被被呼用户降低。一旦这些参数被双方确认，在整个连接存在期间将保持不变，即协商过的服务质量适用于整个传输连接的生存期。

5. 传输层协议等级

传输服务是通过建立连接的两个传输实体之间所用的传输协议来实现的。由于传输服务是在网络服务的基础上实现的，因此传输层协议的等级与网络服务质量密切相关。根据差错性质，网络服务按质量可分为以下 3 种类型。

※ A 型网络服务。具有可接受的残留差错率和故障通知率，即可靠的网络服务。

※ B 型网络服务。具有可接受的残留差错率和不可接受的故障通知率。如有故障发生时，网络层则通过网络服务报告该故障的发生，如 X.25 即为此类服务质量的网络。

※ C 型网络服务。具有不可接受的残留差错率，可能丢失分组，提供完全不可靠的网络服务，IP 网络即为此类服务质量的网络。

这 3 种类型的网络服务中，A 型服务质量最高，B 型服务质量次之，C 型服务质量最差。

传输层的功能是要弥补从网络层获得的服务和向传输服务用户提供的服务之间的差距，它所关心的是提高服务质量。为了能够在各种不同服务类型的网络上进行数据传送，OSI 定义了 5 种协议级别，它们都是面向连接的。

※ 级别 0（简单级）。它建立一个简单的端到端的传输连接，可将长报文分段传送，没有差错恢复功能和将多条传输复用到一条网络连接的能力，主要面向 A 型网络服务。

※ 级别 1（基本差错恢复级）。只增加了在网络断开、连接失败等基本差错时的差错恢复功能，主要面向 B 型网络服务。

※ 级别 2（多路复用级）。具有将多条传输复用到一条网络连接功能和流量控制功能，主要面向 A 型网络服务。

※ 级别 3（差错恢复和多路复用级）。是级别 1 和 2 的综合，既有差错恢复功能又有多路复用功能，主要面向 B 型网络服务。

※ 级别 4（差错检测和恢复级）。在级别 3 的基础上增加了差错检测功能，是最复杂、功能最全的协议级别，主要面向 C 型网络服务。

服务质量较高的网络仅需要较简单的协议级别；反之，服务质量较低的网络则需要较复杂的协议级别。

3.7　高层简介

传输层之上的会话层、表示层及应用层是面向信息处理的高层。这 3 层的功能是为应用程序提供服务的，不包含任何数据传输的功能，即组织和同步进程间的通信，对数据的语法表示进行变换，以及为网络的最终用户提供服务。

3.7.1　会话层

1. 基于传输层的问题

通过前面的学习我们已经知道，在 OSI/RM 的层次体系结构中，物理层协议可以实现物理线路的连接，数据链路层协议可以实现相邻结点之间连接并无差错地传输数据，网络层协议实现源结点和目标结点的连接，传输层协议实现端到端之间连接的建立和维持。

传输层的功能使用户所需要的通信环境十分完善，可以保证用户数据从源 DTE 发出后，按照要求经过通信子网到达目的端 DTE。会话层是建立在传输层之上的，由于利用了传输层提供的服务，使两个会话实体不用考虑它们之间相隔多远、使用了什么样的通信子网等网络通信细节，即可进行透明的、可靠的数据传输。但当两个应用进程进行相互通信时，在如何控制信息的交互，网络应当提供什么样的功能来协助用户管理和控制用户之间的信息交换的问题上，希望有一个作为第三者的服务能组织它们的对话，协调它们之间的数据流，以便使应用进程专注于信息交互，设立会话层就是为了达到这个目。会话层虽然不参与具体的数据传输，但它却对数据传输进行控制和管理。

2. 会话层功能

会话层在两个互相通信的应用进程之间建立、组织和协调双方的交互活动，并使会话获得同步。会话层担负应用进程的服务要求，弥补传输层不能完成的剩余部分工作。对数据的传送提供控制和管理，协调会话过程，为表示层实体提供更好的服务。其主要的功能是会话用户之间的对话管理、数据流同步和重新同步。

会话（session）是指在两个用户进程之间为完成一次完整的通信而建立的连接。会话可以使一个远程终端登录到远地的计算机，进行文件传输或进行其他的应用。由于会话往往是由一系列交互对话组成的，所以对话的次序、对话的进展情况必须加以控制和管理。OSI/RM 之所以设立会话层，就是为了有效地组织和同步进行通信的用户之间的对话，并对它们之间的数据交换进行管理。

（1）对话管理。从原理上说，所有 OSI 的连接都是全双工的。但在许多情况下，高层软件为方便起见往往设计成半双工交互式通信。例如，远程终端访问一个数据库管理系统，需要发出一个查询，然后等待回答，要么轮到用户发送，要么轮到数据库发送，保持这种轮换并强制实行的过程就称为“对话管理”。

会话层通过令牌来进行对话的交互控制，令牌是会话连接的一个属性，表示使用会话的独占使用权。只有拥有令牌的一方才可以发送数据，令牌可在某一时刻动态地分配给一个会话服务用户，该用户用完后又可重新分配。

（2）同步与重新同步。同步与重新同步就是使会话服务用户对会话的进展情况有一致的了解，在会话被中断后可以从中断处继续下去，而不必从头恢复会话。同步与重新同步是通过设置同步点来获得的，即在数据中插入同步点。每次网络出现故障后，仅仅重传最后一个同步点以后的数据（其实就是断点下载的原理）。会话层允许会话用户在传输的数据中自由设置同步点，并对每个同步点赋予同步序号，用以识别和管理同步点。这些同步点是插在用户数据流中一起传输给对方的，当接收方通知发送方它收到一个同步点时，发送方即可确信接收方已将此同步点之前发送的数据全部收妥。

会话层中定义了两类同步点。主同步点用于在连续的数据流中划分出对话单元，一个主同步点是一个对话单元的结束和下一个对话单元的开始，所谓对话单元就是一个活动中数据的基本交换单元，通常代表逻辑上重要的工作部分；次同步点用于在一个对话单元内部实现数据结构化。主同步点与次同步点有一些不同，在重新同步时，只可能回到最近的主同步点，每一个插入数据流中的主同步点都被明确地确认，而次同步点不被确认。

举例来说，某个用户登录到一个远程系统，并与之交换信息。会话层管理这一进程，控制哪一方有权发送信息，哪一方必须接收信息，这其实是一种同步机制。若一个用户正在网络上发送一个大文件的内容，而网络忽然坏了。当网络重新工作时，用户是否必须从该文件的起始处开始重传呢？回答是否定的，因为会话层允许用户在一个长的信息流中插入同步点，如果网络崩溃了，只要将最后一个主同步点以后丢失的数据重传即可。

3.7.2 表示层

1. 基于会话层的问题

会话层向用户提供了信息交互的控制和管理的手段，完成了端到端的数据传送，并且是可靠、无差错的有序传送。但数据传送只是手段而不是目的，最终是要实现对数据的使用。然而，由于不同的计算机系统可能采用了不同的信息编码，例如，PC 采用的是 ASCII 码，而 IBM 主机采用的是 EBCDIC 码。对于同样一个整数，有些机器可能采用 2B 表示，而有些计算机系统则可能采用 4B 表示，如果不加以处理，不同的信息描述将导致通信的计算机系统之间无法正确地识别信息。在这种情况下，表示层就担负起消除这种障碍的任务。设置表示层的目的是屏蔽不同计算机在信息表示方面的差异。

2. 表示层功能

表示层是 OSI/RM 的第 6 层，主要处理不同系统被传送数据的表示问题，解释所交换数据的意义，进行数据压缩，即各种变换（如代码、格式转换等），使采用不同数据表示方法

的开放系统能够相互通信。此外，利用密码对数据进行加密和解密也是表示层的重要功能。

由于各种计算机都可能有各自的数据描述方法，所以不同类型计算机之间交换的数据一般需经过格式转换才能保证其意义不变。表示层要解决的问题是如何描述数据结构并使之与具体的设备无关，其作用是对原站内部的数据结构进行编码，使之形成适合于传输的比特流，到了目的站再进行解码，转换成用户所要求的格式。

对于用户数据来说，可以从两个方面来分析：一个是数据的含义，称作语义；另一个是数据的表现形式，称作语法，例如文字、图形、声音、数据加密和压缩都属于语法范畴。

为使各个系统间交换的信息具有相同的语义，应用层采用的是对数据一般结构描述的抽象语法。表示层为抽象语法制定一种编码规则，便构成一种传输语法。传输语法是同等表示实体之间通信时对用户信息的描述，是对抽象语法比特流进行编码得到的。在表示层中，可用这种方法定义多种传输语法，抽象语法与传输语法之间是多对多的对应关系。每个应用层协议中的抽象语法与一个能对其编码的传输语法的组合，即抽象语法与传输语法之间的对应关系，就构成了一个表示上下文。表示上下文可以在连接建立时协商确定，也可以在通信过程中重新定义。表示层主要处理通信双方之间的数据表示问题，主要功能如下：

※ 语法转换。将抽象语法转换成传输语法，并在对方实现相反的转换。通过这种转换来统一表示被传送的用户数据，使通信双方使用的计算机都可以识别。涉及的内容有代码转换、字符转换、数据格式的修改，以及对数据结构操作的适应、数据压缩、加密等。国际标准化组织定义了一种抽象语法称作抽象语法标记 1（ANS.1）及相应编码规则，包括 3 类 15 种功能单元，其中表示上下文管理功能单元允许用户选择语法和转换，沟通用户之间的数据编码规则，以便有一致的数据形式，能够相互认识。

※ 语法选择。根据应用层的要求协商选用合适的上下文，即选择传输语法传送数据。

※ 连接管理。利用会话层服务建立表示连接，管理在这个连接之上的数据传输和同步控制，以及正常或异常地释放这个连接。

3.7.3　应用层

应用层是直接面向用户的一层，为网络应用提供一个访问网络的接口，使应用程序能够使用网络服务。它采用各种不同的应用协议直接为应用进程提供服务。

应用层也称应用实体（Application Entity，AE），应用实体是被简化的应用进程，它是应用进程中与进程间交互行为有关的那部分，即与 OSI/RM 有关的那部分。而对应用进程中与 OSI/RM 无关的那部分仍称为应用进程。一个应用实体由若干个元素（element）构成，在这些元素中包括一个用户元素（User Element，UE）和若干个应用服务元素（Application Service Element，ASE）。用户元素实际上是应用进程中非标准化模块的化身，用户元素即是应用者。

在应用层中最复杂的就是各种应用要求，并且保证这些不同类型的应用所采用的低层通信协议是一样的。因此 ISO 把一系列业务处理所需的服务按其向应用程序提供的特性分成组，称为应用服务元素（Application Service Element，ASE）。ASE 是 OSI/RM 在应用层中定义的标准化模块，它是应用实体的一部分，通过应用服务元素为用户元素提供标准化服务。有

些服务元素可由多种应用程序共同使用，称为公用应用服务元素（CASE）；有些服务素则只为特定的一种应用服务程序使用，称为特定应用服务元素（SASE）。

需要说明的是，在 1998 年公布的 ISO 8650 中规定，ASE 不再分为 CASE 和 SASE，统称 ASE，只是又根据不同的用途相应地定义了各种 ASE，例如联系控制服务元素（ACSE）、可靠传输元素（RTSE）和远程操作服务元素（ROSE）等，这些以前称为公用应用服务元素。又如文件传输和管理（FTAM）、报文处理等，这类与特定应用有关的 ASE 以前也称为公用应用服务元素。

由于用户要求不同，应用层中提供多种支持不同应用的协议，典型协议如下：

※ FTAM（File Transfer, Access and Management）：提供文件传输、存取和管理。

※ MHS（Message Handling System）：报文处理系统，有关电子邮件服务系统的功能模型，源于 CCITT 的 X.400 规范。

※ VTP（Virtual Terminal Protocol）：虚拟终端协议，将不同类型的终端具有的功能一般化、标准化，以标准的虚拟终端出现。

※ DS（Directory Service）：目录服务，提供全球分布式管理的目录服务。

※ CMIP（Common Management Information Protocol），通用管理信息服务，提供网络管理功能。

总而言之，OSI/RM 的低 3 层属于通信子网，涉及为用户提供透明数据传输连接，操作主要以每条链路为基础，在结点间的各条数据链路上进行通信，由网络层控制各条链路，但要依赖于其他结点的协调操作。高 3 层属于资源子网，主要涉及保证数据以正确、可理解的形式传送。传输层是高 3 层和低 3 层之间的接口，保证透明的端到端连接，满足用户服务质量的要求。

3.8 TCP/IP 模型

OSI/RM 是一种理论上比较完整的网络概念模型，但在实际应用中，完全符合 OSI/RM 的成熟产品却很少；而以 TCP/IP 协议为基础建立的 Internet 的发展却非常迅猛，众多的网络产品都支持 TCP/IP。经过多年的发展，TCP/IP 已成为计算机网络体系结构事实上的工业标准，得到了广泛的实际应用。所以，尽管 OSI/RM 国际标准对计算机网络起到了规范和指导作用，但实际使用的标准仍然是 TCP/IP。

TCP/IP（Transmission Control Protocol/Internet Protocol，传输控制协议 / 因特网协议）是一组用于实现网络互联的通信协议，是 Internet 最基本的协议和互联网络的基础。

3.8.1 TCP/IP 模型结构

TCP/IP 模型从更实际的角度出发，形成了具有更高效率的 4 层结构，即网络接口层、网络互联层（IP 层）、传输层（TCP 层）和应用层。虽然它与 OSI/RM 各有自己的分层结构，但大体上两者仍能相互对照，如图 3-16 所示。

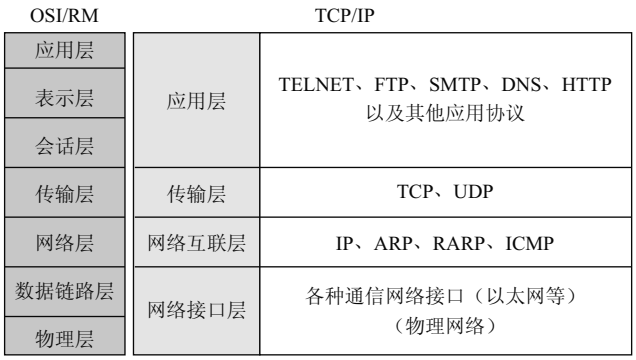


图 3-16　TCP/IP 与 OSI/RM 对比

1．网络接口层

网络接口层负责接收 IP 数据报并将其封装成适合在物理网络上传输的帧格式进行传输，或将从物理网络接收到的帧解封，取出 IP 数据报交给上层的网络互联层。

网络接口层与 OSI/RM 中的物理层和数据链路层相对应。事实上，TCP/IP 本身并未定义该层的协议，主要是为了保证通过 TCP/IP 模型可将不同的物理网络互联起来。参与互联的各网络使用自己的物理层和数据链路层协议，然后与 TCP/IP 的网络接口层进行连接。如局域网的 Ethernet、令牌网、分组交换网的 X.25、帧中继、ATM 协议等。当一种物理网络被用作传送 IP 数据包的通道时，就可以认为是这一层的内容。这充分体现出 TCP/IP 协议的兼容性与适应性，它也为 TCP/IP 的成功奠定了基础。

2．网络互联层

网络互联层（IP 层）是 TCP/IP 模型的关键部分。它负责将数据报文独立地从源主机送到目的主机，以及建立互联网络。网间的数据报可根据它携带的目的 IP 地址，通过路由器由一个网络传送到另一个网络。它主要解决数据封装、寻址、数据报的分段和路由选择等问题，对应于 OSI/RM 的网络层。这一层有 4 个主要协议：IP、ARP、RARP 和 ICMP，其中最重要的是 IP 协议。

※　网际协议（IP）：提供 IP 地址寻址、路由选择以及信息包的分段和重组功能。

※　地址解析协议（ARP）：负责将 IP 地址转换为物理地址。

※　反向地址解析协议（RARP）：负责将物理地址转换为 IP 地址。

※　互联网控制报文协议（ICMP）：用于传送控制报文和差错报告报文。

3．传输层

传输层负责在源主机和目的主机的应用程序之间提供端到端的数据传输服务，与 OSI/RM 中的传输层相似。该层处理网络互联层没有处理的通信问题，保证通信连接的可靠性。该层主要有传输控制协议 TCP 和用户数据报协议 UDP。

TCP 协议提供面向连接的、可靠的数据传输服务。双方通信之前，先建立一条连接，然后双方就可以在其上发送数据流。TCP 协议具有数据报的顺序控制、差错检测、校验以及重发控制等功能。

UDP 协议提供无连接的、不可靠的数据传输服务。UDP 是依靠 IP 协议来传送报文的，因而它的服务和 IP 一样是不可靠的，UDP 协议将可靠性问题交给应用程序解决。这种服务不用确认，也不进行流量控制。UDP 报文可能会出现丢失、重复、失序等现象。

4．应用层

TCP/IP 的应用层与 OSI/RM 的应用层有较大区别，它不仅包含了会话层以上 3 层的所有功能，而且还包括了应用进程本身。因此，TCP/IP 模型的简洁性和实用性就体现在它不仅把网络层以下的部分留给了实际网络，而且将高层部分和应用进程结合在一起，形成了统一的应用层。TCP/IP 的应用层包含所有的高层协议，为用户提供所需的各种服务。其中 FTP、TELNET、SMTP、DNS、HTTP 是几个在各种不同机型上广泛实现的协议，TCP/IP 中还定义了许多别的高层协议。随着计算机网络技术的不断发展，还会有新的协议不断加入。

3.8.2 OSI/RM 与 TCP/IP 的比较

OSI/RM 和 TCP/IP 模型都采用了层次结构的概念，但前者是 7 层模型，后者是 4 层结构。不论在层次的划分还是协议的使用上都有明显的差别。它们的主要不同点如下：

※ TCP/IP 与 OSI/RM 相比，简化了高层的协议，简化了会话层和表示层，将其融合到了应用层，使通信的层次减少，提高了通信的效率。

※ 在模型和协议的关系上，OSI/RM 抽象能力高，适合于描述各种网络，它采取自上向下的设计方式，先定义参考模型，后逐步定义各层的协议，通用性强，但实现困难。TCP/IP 则正好相反，先有协议之后，人们为了对它进行分析研究，才制定了 TCP/IP 模型，实用性强，但通用性不足，不适合描述其他非 TCP/IP 网络。

※ OSI/RM 的概念清晰，明确定义了服务、接口和协议的概念及它们之间的关系；而 TCP/IP 参考模型没有明确区分这 3 个概念，功能描述和实现细节混在一起。

※ OSI/RM 的网络层既提供面向连接的服务，又提供无连接服务，但传输层仅提供面向连接的服务；TCP/IP 模型的网络互联层仅提供无连接服务，而传输层提供面向连接的服务（TCP）和无连接服务（UDP）。

总之，OSI/RM 虽然一直被人们看好，但由于没有把握时机，迟迟没有一个成熟的产品推出，大大影响了它的发展；相反，TCP/IP 参考模型虽然有不尽如人意的地方，但经实践证明它还是比较成功的，特别是近年来国际互联网的飞速发展，也使它获得了巨大的支持，TCP/IP 协议不仅应用在广域网上，在局域网上也逐渐成为被普遍采用的协议。

本章小结

本章主要介绍了通信协议和网络体系结构的概念，并介绍了开放系统互联参考模型（OSI/RM），物理层功能与协议、数据链路层功能与协议、网络层功能与协议、OSI/RM 高层功能以及 TCP/IP 体系结构。

1．通信协议与网络体系结构的概念

通信协议是计算机网络中为进行数据通信而制定的通信双方共同遵守的规则、标准或约定的集合。协议本质上是一系列规则和约定的规范性描述。它不仅定义了通信时信息必须采用的格式和这些格式的意义，而且还要对事件发生的次序做出说明。所以，任何一种网络协议都应包括如下三要素：语法、语义和时序。

计算机网络通信的功能的层次结构以及各层服务、协议的集合统称为网络体系结构。网络体系结构中主要概念有协议、服务、接口、服务原语、服务访问点和面向连接和无连接的服务等。

2．开放系统互联参考模型（OSI/RM）

OSI/RM 定义了计算机网络系统的层次结构、层次之间的相互关系及各层所包括的服务。它将网络通信功能划分为 7 个层次，规定了每个层次的具体功能。自顶向下的 7 个层分别是应用层、表示层、会话层、传输层、网络层、数据链路层和物理层。OSI/RM 成功之处在于清晰地分开了服务、接口和协议这 3 个概念，服务描述了每一层的功能，接口定义了某层提供的服务如何被高层访问，而协议是每一层功能的实现方法。

3．物理层功能与协议

物理层的基本功能是负责实际或原始的数据“位”传送，目的是在通信设备 DTE 和 DCE 之间提供透明的比特流传输。物理层向数据链路层提供的服务是建立、维持和释放物理连接，并在物理连接上透明传输比特流。物理层协议规定了网络物理设备之间的物理接口特性及通信规则。物理层协议用 4 个特性对网络设备和传输介质之间的接口进行定义：机械特性、电气特性、功能特性和规程特性

4．数据链路层功能与协议

数据链路层通过数据链路层协议，在不太可靠的物理链路上实现可靠的数据传输，向网络层提供透明的和可靠的数据传送服务。数据链路层必须完成的基本功能有链路管理、帧的封装、差错控制、将数据和控制信息区分开、透明传输和流量控制。

当相邻两个结点传递数据时，除了需要一条物理线路和设备外，还需要一些必要的通信规则来控制数据的传输，这些控制相邻结点间数据传输的通信规则就是数据链路层协议。HDLC 是一种数据链路层协议，称为高级数据链路控制协议。

5．网络层功能与协议

在 OSI/RM 中，网络层以数据链路层提供的无差错传输为基础，把高层发来的数据组织成分组，从源结点经过若干个中间结点传送到目的结点。传送过程要解决的关键问题是选择路径，在选择路径时还要考虑解决流量控制问题，防止网路中出现局部的拥挤或全面的阻塞。主要功能包括路径选择、拥塞控制和网际互联等。

通信子网中源结点和目的结点之间存在多条传输路径的可能性。网络结点在收到一个分组后，根据一定的原则和算法确定向下一个结点传送的最佳路径，这就是路由选择。根据路由算法能否依靠网络当前的流量和拓扑结构来调整它们的路由决策，路由选择算法有静态路由选择算法和动态路由选择算法。

拥塞现象是指到达通信子网中某一部分的分组数量过多，使得该部分网络来不及处理，以致引起这部分乃至整个网络性能下降的现象。

典型的网络层协议是 X.25 协议，X.25 协议实际上是主机与分组交换网之间的一组接口标准，它包括物理层、数据链路层和分组层 3 个层次，其中分组层相当于 OSI/RM 中的网络层。X.25 协议最主要的功能是向主机提供多信道的虚电路服务。

6．网络层功能与协议

传输层提供了主机应用进程之间的端到端的服务，主要功能是：为一个进行的会话或连接提供可靠的传输服务，完成端到端的通信链路的建立、维护和管理；在单一连接上提供端到端的端口号与流量控制及差错恢复等服务。

服务质量（Quality of Service，QoS）是传输层性能的度量，衡量了传输层的总体性能，反映了传输质量及服务的可用性。服务质量可用一些参数来描述，如连接建立延迟、连接建立失败的概率、吞吐率、传输延迟、残留差错率、连接拆除延迟、连接拆除失败概率、传输失败率等等。

7．OSI/RM 高层功能

会话层在两个互相通信的应用进程之间建立、组织和协调双方的交互活动，并使会话获得同步。会话层担负应用进程的服务要求，弥补传输层不能完成的剩余部分工作。它对数据的传送提供控制和管理，协调会话过程，为表示层实体提供更好的服务。其主要的功能是会话用户之间的对话管理，数据流同步和重新同步。

表示层是 OSI/RM 的第六层，主要处理不同系统传送数据的表示问题，解释所交换数据的意义，进行数据压缩即各种变换（如格式转换等），使采用不同数据表示方法的开放系统能够相互通信。此外，利用密码对数据进行加密和解密也是表示层的重要功能。

应用层是直接面向用户的一层，为网络应用提供一个访问网络的接口，使应用程序能够使用网络服务。它采用各种不同的应用协议直接为应用进程提供服务。

8．TCP/IP 体系结构

TCP/IP 体系结构具有更高效率的 4 层结构，即网络接口层、网际互联层（IP 层）、传输层（TCP 层）和应用层。

习题

一、概念解释

网络体系结构，协议，服务，接口，服务访问点，服务原语，面向连接和无连接，DTE 和 DCE，帧，数据链路，拥塞现象，路由选择算法，服务质量，同步点。

二、单选题

1．设计网络功能时采用分层实现，在下列分层原则中不正确的是（　）。

A. 每层的功能应明确

B. 层数应适中

C. 层间的接口须清晰，跨接口的信息量应尽可能少

D. 同一功能可以由多个层共同实现

2. 在组成协议的三要素中，（　）规定数据的结构和格式。

A. 语法　　B. 语义　　C. 时序　　D. 都不是

3. 通信子网不包括 OSI/RM 的层次是（　）。

A. 物理层　　B. 数据链路层　　C. 网络层　　D. 传输层

4. 通信子网一般由 OSI/RM 的（　）。

A. 高三层组成　B. 中间三层组成　C. 低三层组成　D. 以上都不对

5. 当数据在网络层时，称之为（　）。

A. 报文　　B. 分组　　C. 帧　　D. 位

6. 在 OSI/RM 中，当数据分组从较低的层移动到较高的层时，其首部会被（　）。

A. 加上　　B. 去除　　C. 重新安排　　D. 修改

7. 在 OSI/RM 中，当数据分组从设备 A 传到 B 时，在 A 的第 3 层加上的首部会在 B 的第（　）层被读出。

A.2　　B.3　　C.4　　D.5

8. 以下设备中属于 DCE 的是（　）。

A. 显示器　　B. 键盘　　C. 打印机　　D. 调制解调器

9. RS−232C 是（　）接口规范。

A. 物理层　　B. 数据链路层　　C. 网络层　　D. 运输层

10. OSI/RM 的物理层协议定义了 4 个特性，其中定义接口形状、尺寸的内容属于（　）。

A. 机械特性　　B. 电气特性　　C. 功能特性　　D. 规程特性

11. 采用 RS−232C 标准连接 PC 和调制解调器，其请求发送信号（RTS）的连接方向为（　）。

A.DCE → DTE　B.DCE → DCE　C.DTE → DTE　D.DTE → DCE

12. 差错控制和流量控制是在 OSI/RM 的（　）完成的。

A. 物理层　　B. 数据链路层　　C. 网络层　　D. 传输层

13. 数据被封装成帧是在 OSI/RM 的（ ）完成的。

A. 物理层　　B. 数据链路层　　C. 网络层　　D. 传输层

14. 流量控制实际上是对（ ）。

A. 发送方数据流量的控制　　B. 接收方数据流量的控制

C. 发送方和接收方数据流量的控制　D. 以上都不对

15. 若 HDLC 帧数据段中出现比特串 01011111110，采用比特填充技术，填充后的输出为（ ）。

A.010111111010　B.010111111100　C.010111110110　D.010111101110

16. HDLC 规程中采用的帧同步是（ ）。

A. 字节计数法　B. 字符填充法　C. 比特填充法　D. 违法编码法

17. 在 OSI/RM 的层次中，（ ）的数据传送单位是分组。

A. 物理层　　B. 数据链路层　　C. 网络层　　D. 传输层

18. 决定使用哪条路径通过通信子网是在 OSI/RM 中的（ ）完成的。

A. 物理层　　B. 数据链路层　　C. 网络层　　D. 传输层

19. X.25 协议的分组层最主要的功能是（ ）。

A. 多路复用物理链路　　B. 实现 DTE-DCE 连接

C. 链路访问控制　　D. 差错控制

20. 在采用点对点通信线路的网络中，由于连接多台计算机之间的线路采用网状结构，因此确定分组从源结点通过通信子网到达目的结点的适当传输路径需要使用（ ）。

A. 差错控制算法　　B. 路由选择算法

C. 拥塞控制算法　　D. 协议变换算法

21. 在拓扑结构变化不大的网络中，可得到较好运行效果的路由算法是（ ）。

A. 固定路由选择　　B. 随机路由选择

C. 泛射路由选择　　D. 独立路由选择

22. 为用户提供端到端的透明数据运输服务是在 OSI/RM 中的（ ）完成的。

A. 物理层　　B. 数据链路层　　C. 网络层　　D. 传输层

23. OSI/RM 的传输层中差错恢复和多路复用技术属于传输层协议等级的级别（ ）。

A.1　　B.5　　C.3　　D.4

24. 语法转换、数据加密和压缩是（　）应完成的功能。

A. 物理层　　B. 传输层　　C. 会话层　　D. 表示层

25. TCP/IP 的 IP 层的功能对应 OSI/RM 的（　）应完成的功能。

A. 应用层　　B. 传输层　　C. 网络层　　D. 会话层

三、填空题

1. 网络协议主要由 3 个要素组成：_____、_____、_____。

2. 设计网络协议采用分层实现，每层都有协议控制本层的对等实体进行通信，除了在物理介质上进行的是 _____ 通信，在其他各层间进行的都是 _____ 通信。

3. 计算机网络体系结构是一种 _____ 结构，OSI/RM 的传输层（第 4 层）处于 _____ 层提供的服务之上，给 _____ 层提供服务。

4. 数据链路层的服务用户是 _____，同时它又使用 _____ 层提供的服务。

5. 在网络体系结构中，不同系统对等层上数据传送的基本单位是 _____。

6. 所有服务都是由 _____ 来描述的，它规定了通过 _____ 所必须传递的信息。

7. 物理层协议描述了建立、维护和维持数据链路在 _____、_____、_____ 和 _____ 4 个方面的特征。

8. 物理层接口协议实际上是 DTE 和 DCE 或其他通信设备之间接口的一组规定。其中，DTE 是指 _____ 设备，属于用户所有的设备，具有数据 _____ 能力，典型设备是 _____；DCE 指 _____ 设备，为用户设备提供入网连接点，提供 _____ 功能，典型设备是 _____。

9. RS-232C 机械特性规定了使用一个 _____ 芯标准连接器，电气特性规定逻辑 1 的电平为 _____ 至 _____，逻辑 0 的电平为 _____ 至 _____。

10. 停止等待协议是 OSI/RM 中 _____ 层的一种协议，它可以实现对数据的 _____ 控制和 _____ 控制功能。

11. 通信子网由传输介质和通信设备构成，它包含了 OSI/RM 中 3 个层次的功能，即 _____ 层、_____ 层和 _____ 层。

12. HDLC 协议中有 3 种类型的帧：_____、_____ 和无编号帧，其中用于差错控制和流量控制的是 _____ 帧。

13. 采用 HDLC 协议，假设被传输的数据为比特串 101111101111110010，经 0 比特插入后的结果是 _____。

14. 在 OSI/RM 中规定，网络层提供两种类型的网络服务方式，即 _____ 和 _____。

15. 虚电路服务是 OSI/RM_____ 层向传输层提供的一种可靠的数据传送服务，它确保所有分组按发送 _____ 到达目的地端系统。

16. 到达通信子网中某一部分的分组数量过多，使该部分乃至整个网络性能下降，称为_____现象。

17. 当接收方的接收能力小于发送方的发送能力时，必须进行_____控制。

18. OSI/RM 中将服务质量（QoS）分为 3 类，_____类服务质量最高，其传输层采用的协议级别比较_____。

19. OSI/RM 环境中负责处理语义的是_____层，负责处理语法的是_____层，下面各层负责信息从源到目的地的有序传送。

20. 网络体系结构参考模型是为了规范和设计网络提出的抽象模型，具有代表性的参考模型有两个：_____和_____。

21. TCP/IP 有两个著名的协议，分别是_____和_____。

四、简答题

1. 什么是网络体系结构？网络体系结构中基本的原理是什么？

2. 网络协议的组成要素是什么？试举例说明协议及对应的要素。

3. 简要说明物理层要解决什么问题，物理层的接口有哪些特性？

4. 简要画出数据终端设备（DTE）之间通过公用电话网进行远程数据传输至少需要增加哪些设备，并画出连接示意图。

5. 说明服务和协议的关系。

6. 比较数据报与虚电路两种服务各自的优缺点及适用场合。

7. OSI/RM 中的第几层分别处理下面的问题？

① 噪声使传输的数据 0 变为 1，接收端发现错误并纠错；② 接收端检测出收到序号错误的帧；③ 决定分组使用哪条路径到达目的端；④ 分组交换网交付给接收端的分组序号错误；⑤ 在两端用户之间传送文件的过程中，连接中断后能够从中断的地方开始继续传输；⑥ 两端用户之间传送文件。

8. HDLC 的帧格式是怎样的？如何保证传送信息的透明性？

9. 路由选择的作用是什么？常用的方法有哪些？

10. 为何要引入传输层？其作用具体表现在哪几个方面？

11. 在传输层中端口的作用是什么？

12. 抽象语法与传输语法的关系如何？为什么要使用两种不同的语法？

13. 应用层实体和应用程序是否是相同的概念？说明应用层的功能。

14. TCP/IP 模型的网络接口层并没有具体的协议，为什么要这样设计？

15. 说明 TCP/IP 模型与 OSI/RM 相比有何优点和不足。