

1.5 计算机网络的性能



- 1.5.1 计算机网络的性能指标
- 1.5.2 计算机网络的非性能特征



1.5.1 计算机网络的性能指标

- 计算机网络的性能一般是指它的几个重要的性能指标，主要包括：

- 速率

- 带宽

- 吞吐率

- 时延

- 时延带宽积

- 往返时间 **RTT**

- 利用率

1.7 计算机网络的性能指标

1. 速率

- **比特 (bit)** 是计算机中数据量的单位，也是信息论中使用的信息量的单位。
- **Bit** 来源于 **binary digit**，意思是一个“**二进制数字**”，因此一个比特就是二进制数字中的一个 **1** 或 **0**。
- **速率**即**数据率(data rate)**或**比特率(bit rate)**是计算机网络中最重要的一個性能指标。速率的单位是 **b/s**，或**kb/s**, **Mb/s**, **Gb/s** 等

2.带宽

原义指某个信号具有的频带宽度。

(1) 模拟信号

带宽——表示允许信号占用的频率范围。

单位：**HZ、KHZ、MHZ**

如：话音的带宽为**3.1 KHZ(300 HZ ~ 3.4 KHZ)**

(2) 数字信号

带宽——表示在单位时间内网络中某个信道所通过的**最高数据率，单位是bit/s**

单位： 比特/秒， **bit/s**， **bps**

■ 在表示带宽时， 更常用的带宽单位是

➤ 千比特每秒， 即 **kb/s** (10^3 b/s)

➤ 兆比特每秒， 即 **Mb/s** (10^6 b/s)

➤ 吉比特每秒， 即 **Gb/s** (10^9 b/s)

➤ 太比特每秒， 即 **Tb/s** (10^{12} b/s)

■ 请注意： 在计算机上表示数据大小和磁盘容量时，

$$K = 2^{10} = 1024, \quad M = 2^{20}, \quad G = 2^{30}, \quad T = 2^{40}。$$

■ **kB/S** 和 **kbit/s** 的区别 **B** 字节 **bit** 比特

■ **1字节=8比特。**

3. 吞吐量

- **吞吐量(throughput)**表示在单位时间内通过某个网络（或信道、接口）的数据量。
- 吞吐量更经常地用于对现实世界中的网络的一种测量，以便知道**实际**上到底有多少数据量能够通过网络。
- **吞吐量受网络的带宽或网络的额定速率的限制。**

4.时延

- 时延：指数据（报文或分组，比特）从网络（或一条链路）的一端传送到另一端所需的时间。
- 时延有以下几个组成部分：发送时延、传播时延、处理时延和排队时延。

(1) .传播时延

指电磁波在信道中传播一定距离所需的时间。

$$\text{传播时延} = \frac{\text{信道长度 (米)}}{\text{电磁波在信道中的传播速率 (米/秒)}}$$

电磁波在信道中的传播速率:

太空: $3.0 \times 10^5 \text{ km/s}$ (30万km/s)

电缆: $2.3 \times 10^5 \text{ km/s}$ (23万km/s)

光纤: $2.0 \times 10^5 \text{ km/s}$ (20万km/s)

如: 1000km长的光纤, 其传播时延 $= 5 \times 10^{-3} \text{ s}$
 $= 5 \text{ ms}$

(2) 发送时延

发送数据时，数据块从结点进入到传输媒体所需要的时间。也就是从发送数据帧的第一个比特算起，到该帧的最后一个比特发送完毕所需的时间。

$$\text{发送时延} = \frac{\text{数据块长度 (bit)}}{\text{信道带宽 (bit/s或bps)}}$$

习题

- 例 如果收发两端的距离为10km,信号在媒体上传输的速率为 2.0×10^5 km/s ,数据长度为1000B,数据发送的速率为100kbit/s;

求发送时延? 传播时延?

$$\text{发送时延} = 1000 \times 8 / 100 \times 10^3 = 0.08\text{s}$$

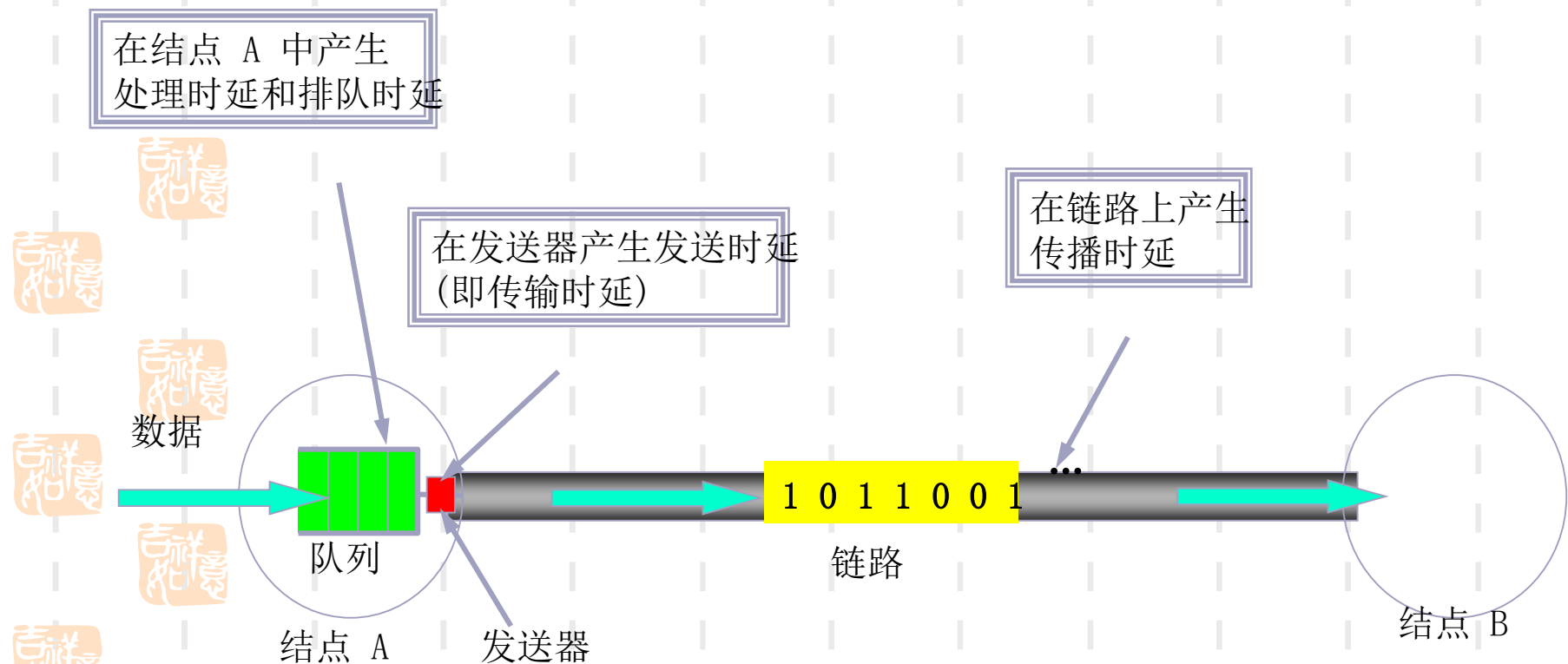
$$\text{传播时延} = 10 / 2.0 \times 10^5 = 0.00005\text{s}$$

(3) 处理时延 交换结点为存储转发而进行一些必要的处理所花费的时间。

(4) 排队时延 结点缓存队列中分组排队所经历的时延。

- 排队时延的长短往往取决于网络中当时的通信量。
- 有时可用排队时延作为处理时延。指数据在交换节点等候发送时在缓存队列中排队所经历的时间。

总时延 = 发送时延 + 传播时延 + 处理时延 + 排队时延



5. 时延带宽积(表示链路能够容纳的比特数)

时延带宽积

(传播) 时延

带宽



$$\text{时延带宽积} = \text{传播时延} \times \text{带宽}$$

链路的时延带宽积又称为以**比特**为单位的链路长度。

5.时延带宽积（以比特为单位的的链路长度）

=传播时延 \times 带宽=链路的体积

表示链路能够容纳的比特数。

6.往返时延RTT

数据从发送端开始发送，到发送端收到来自接收端的确认信息所经历的时间。

往返时延带宽积：在收到确认信息前，已经发送的比特数。

7. 利用率

- **信道利用率**指出信道有百分之几的时间是被利用的（有数据通过）。完全空闲的信道的利用率为零。
- **网络利用率**则是全网络的信道利用率的加权平均值。
- 信道利用率并非越高越好。

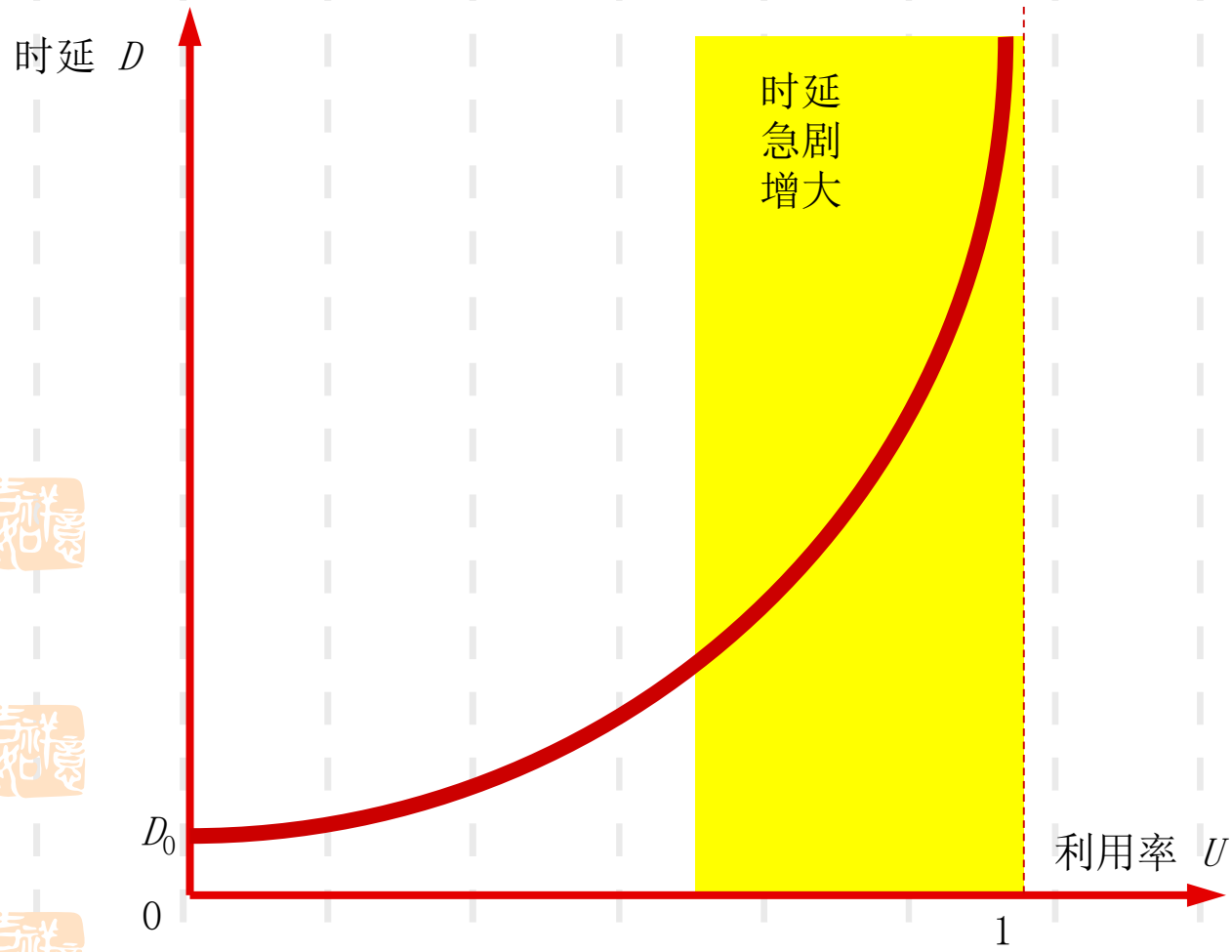
时延与网络利用率的关系

- 根据排队论的理论，当某信道的利用率增大时，该信道引起的时延也就迅速增加。
- 若令 D_0 表示网络空闲时的时延， D 表示网络当前的时延，则在适当的假定条件下，可以用下面的简单公式表示 D 和 D_0 之间的关系：

$$D = \frac{D_0}{1-U}$$

U 是网络的利用率，数值在 0 到 1 之间。

吉祥如意



1.5.2 计算机网络的非性能特征

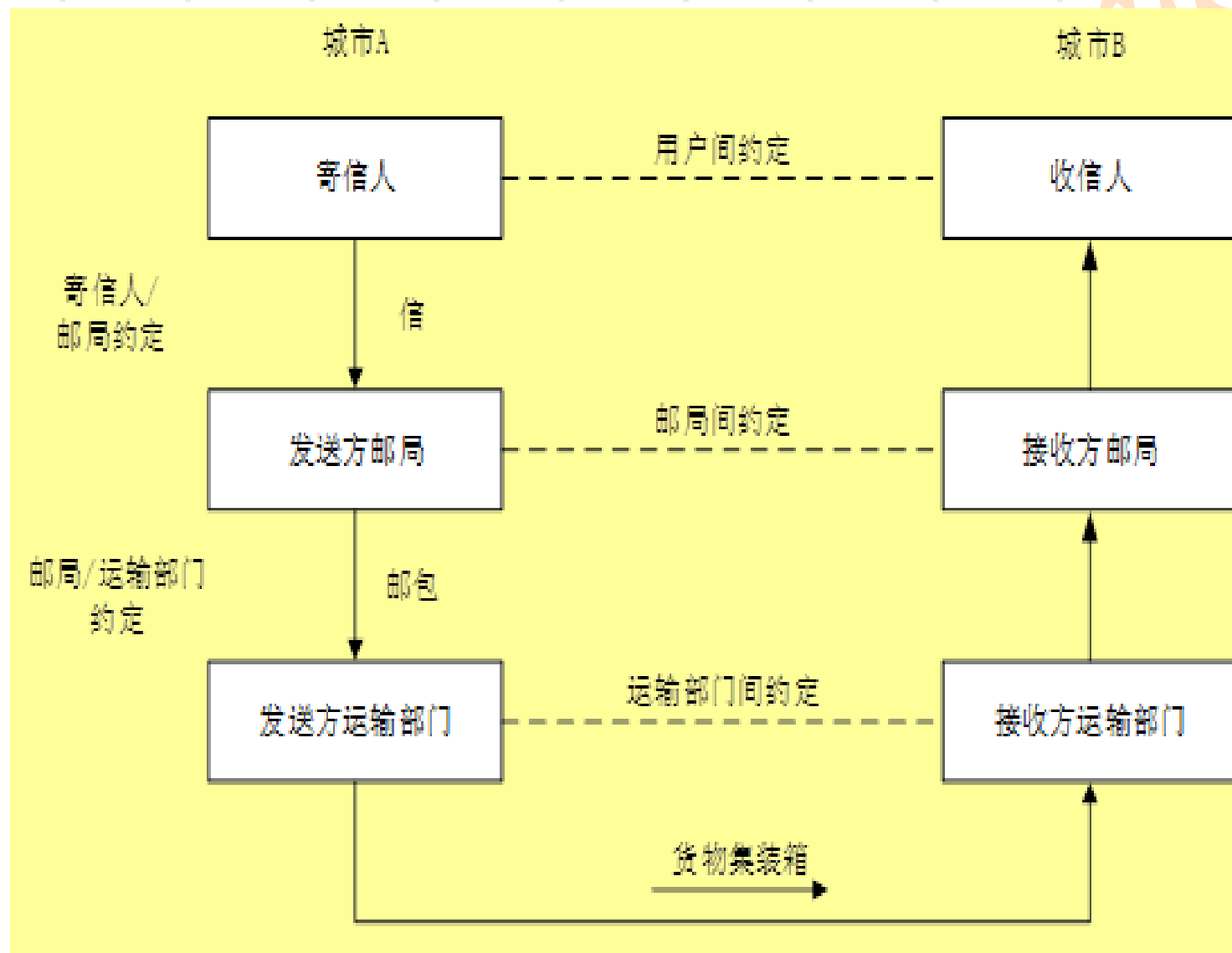
- 费用
- 质量
- 标准化
- 可靠性
- 可扩展性和可升级性
- 易于管理和维护

1.6 计算机网络的体系结构

1.6.1 计算机网络体系结构的形成

- 相互通信的两个计算机系统必须高度协调工作才行，而这种“协调”是相当复杂的。
- “分层”可将庞大而复杂的问题，转化为若干较小的局部问题，而这些较小的局部问题就比较易于研究和处理。

生活中的层次（邮政系统模型）



在网络的层次模型中需要解决

(1) 一个网络到底应该具有哪些层次？网络中每一层的功能是什么？（功能分类与层次划分）

(2) 网络中各层之间的关系是如何设计与处理的？它们之间如何进行交互？（服务与接口）

(3) 通信双方的数据传输需要遵循哪些规则和约定？（协议）

1.6.2 划分层次的必要性

- 计算机网络中的数据交换**必须遵守事先约定好的规则**。
- 这些**规则**明确规定了所交换的数据的格式以及有关的同步问题（同步含有时序的意思）。
- **网络协议(network protocol)**，简称为**协议**，是为进行网络中的数据交换而建立的规则、标准或约定。

网络协议的组成要素

- **语义** 需要发出何种控制信息，完成何种动作以及做出何种响应。“做什么”
- **语法** 数据与控制信息的结构或格式。
“怎么做”
- **同步** 事件实现顺序的详细说明。

协议很复杂



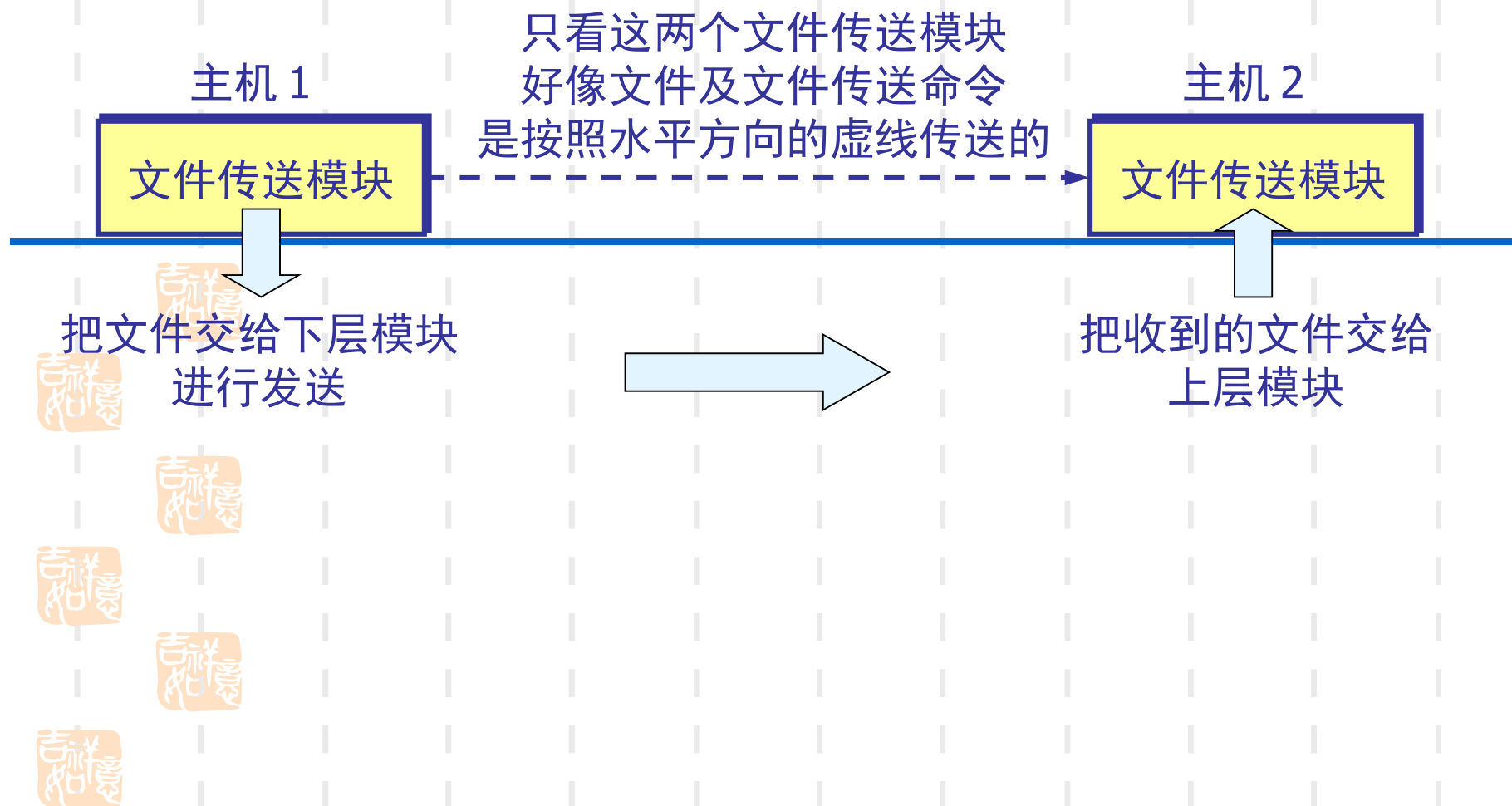
- 协议必须把所有**不利的条件**事先都估计到，而**不能假定**一切都是正常的和非常理想的。
- 看一个计算机网络协议是否正确，不能光看在正常情况下是否正确，还必须非常仔细地检查这个协议**能否应付各种异常情况**。



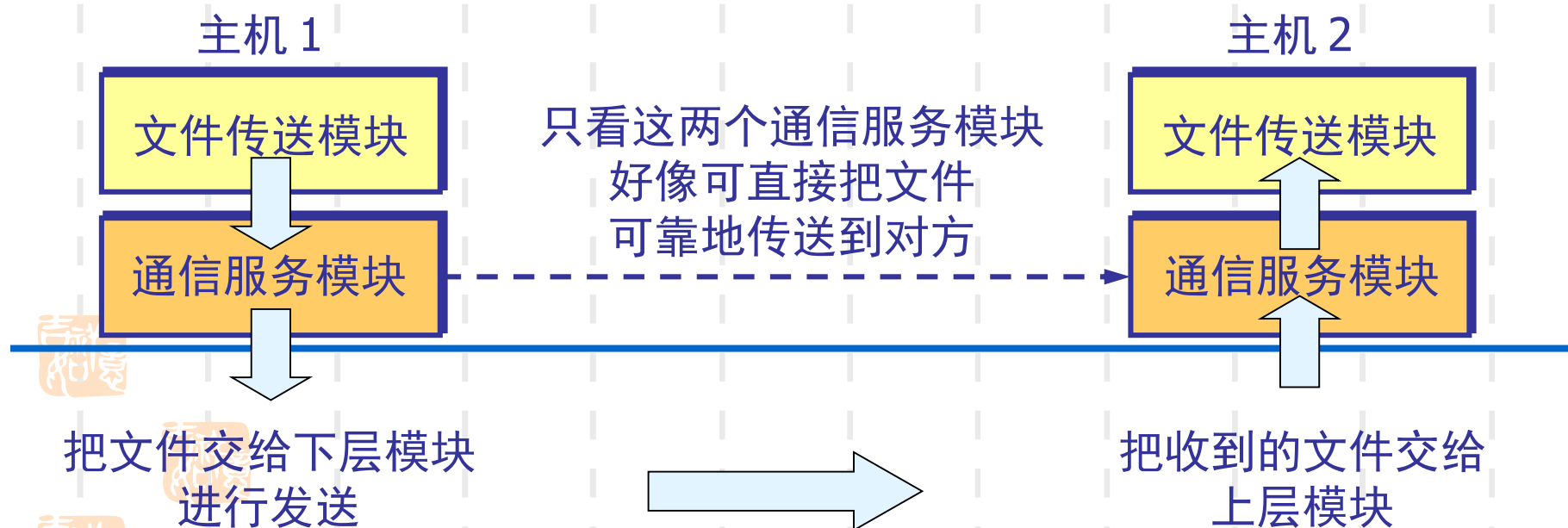
划分层次的概念举例

- 主机 1 向主机 2 通过网络发送文件。
- 可以将要做的工作进行如下的划分。
- 第一类工作与传送文件直接有关。
 - 确信对方已做好接收和存储文件的准备。
 - 双方协调好一致的文件格式。
- 两个主机将文件传送模块作为最高的一层。剩下的工作由下面的模块负责。

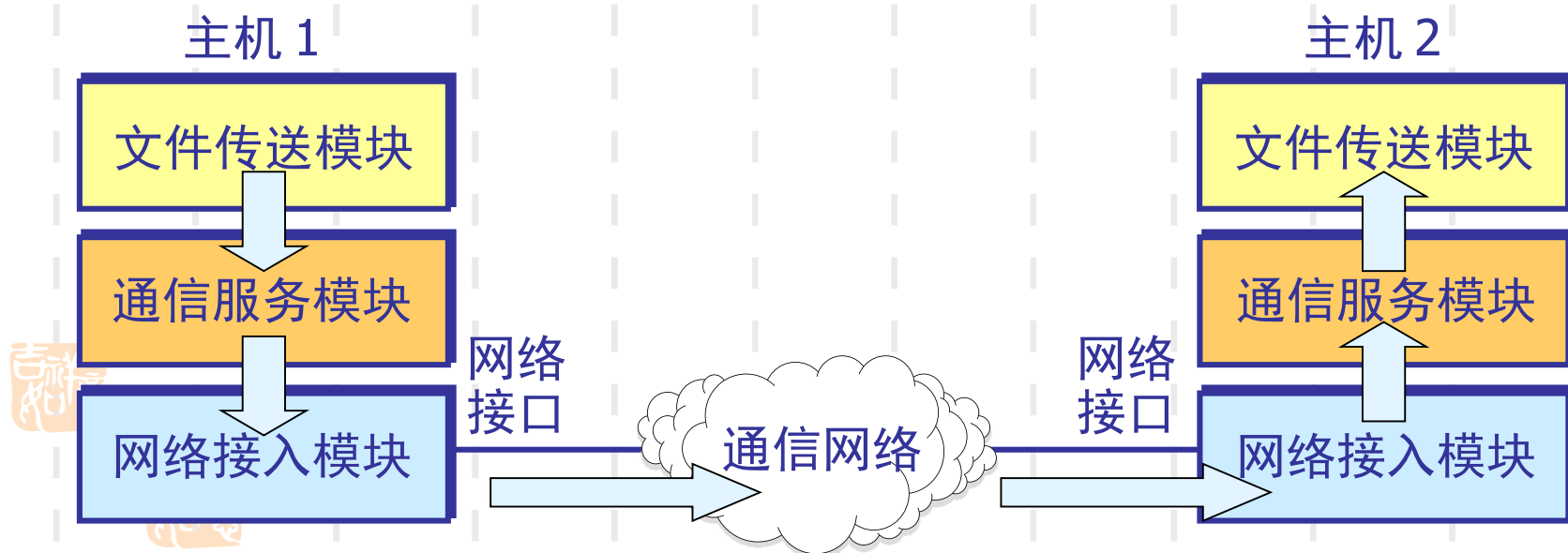
两个主机交换文件



再设计一个通信服务模块



再设计一个网络接入模块



网络接入模块负责做与网络接口细节有关的工作
例如，规定传输的帧格式，帧的最大长度等。

分层的好处

- 各层之间是独立的。
- 灵活性好。
- 结构上可分割开。
- 易于实现和维护。
- 能促进标准化工作。



计算机网络的体系结构

- 计算机网络的**体系结构(architecture)**是计算机网络的各层及其协议的集合。
- 体系结构就是这个计算机网络及其部件所应完成的功能的**精确定义**。
- **实现(implementation)**是遵循这种体系结构的前提下用何种硬件或软件完成这些功能的问题。
- 体系结构是抽象的，而实现则是具体的，是真正正在运行的计算机硬件和软件。

关于开放系统互连参考模型

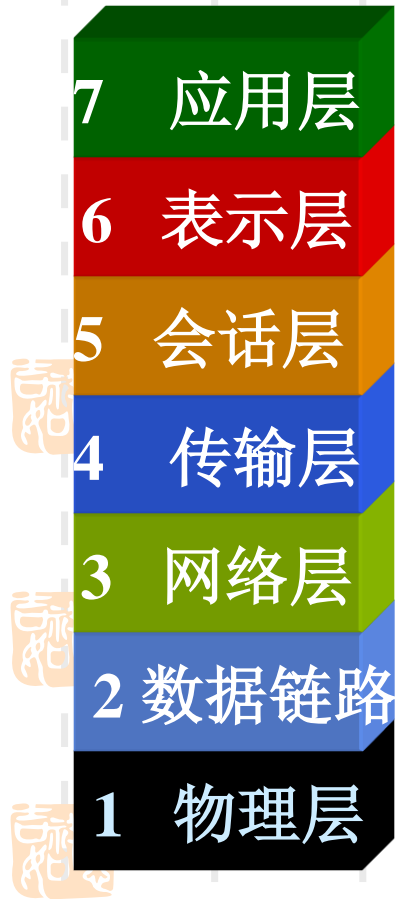
OSI/RM

- 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、也遵循这同一标准的其他任何系统进行通信。

■ 在市场化方面 **OSI** 却失败了。

- **OSI 的专家们在完成 OSI 标准时没有商业驱动力；**
- **OSI 的协议实现起来过分复杂，且运行效率很低；**
- **OSI 标准的制定周期太长，因而使得按 OSI 标准生产的设备无法及时进入市场；**

1. OSI/RM模型结构



OSI/RM参考模型的功能划分为七个层次：

- 应用层（Application Layer）
- 表示层（Presentation Layer）
- 会话层（Session Layer）
- 传输层（Transport Layer）
- 网络层（Network Layer）
- 数据链路层（Data Link Layer）
- 物理层（Physical Layer）



(1) 应用层



为用户的应用程序提供网络通信服务



识别目的通信方的可用性；



应用程序之间进行同步；



判断是否为通信过程申请了足够的资源；





2. 表示层

数据表示问题：即信息的语法和语义。

如：数据加密、解密；压缩、解压缩等

3. 会话层

建立、管理、中止不同机器上的应用程序间的会话。

“会话”——完成一项任务而进行的一系列相关的信息交换。





(4) 传输层

保证多端口多进程通信；
端到端可靠传输；

流量控制；

拥塞控制；

(3) 网络层

路由选择、网络互联



OSI各层功能

(2) 数据链路层

差错控制

封装成帧

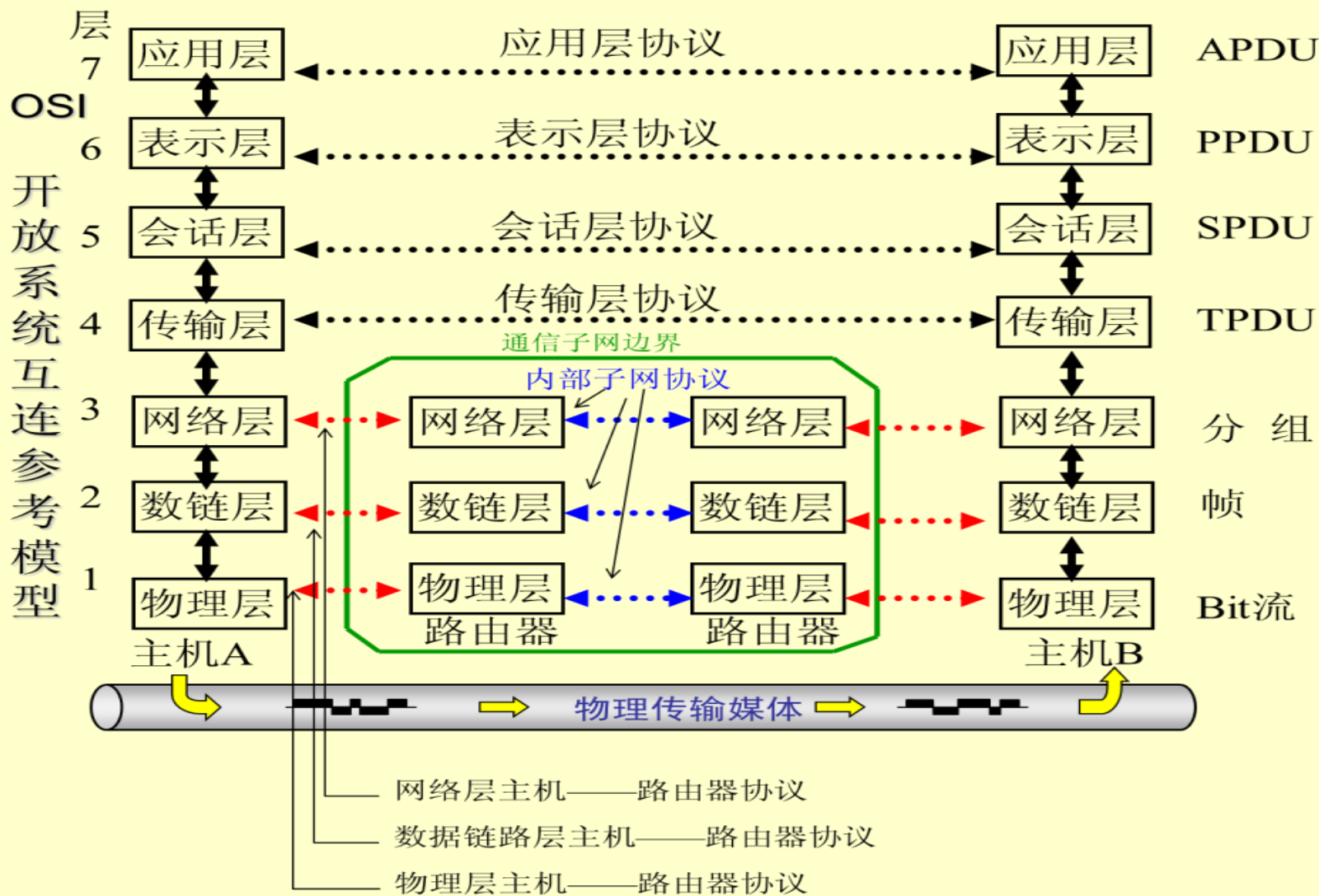


OSI各层的功能

(1) 物理层

任务是透明地传送比特流。





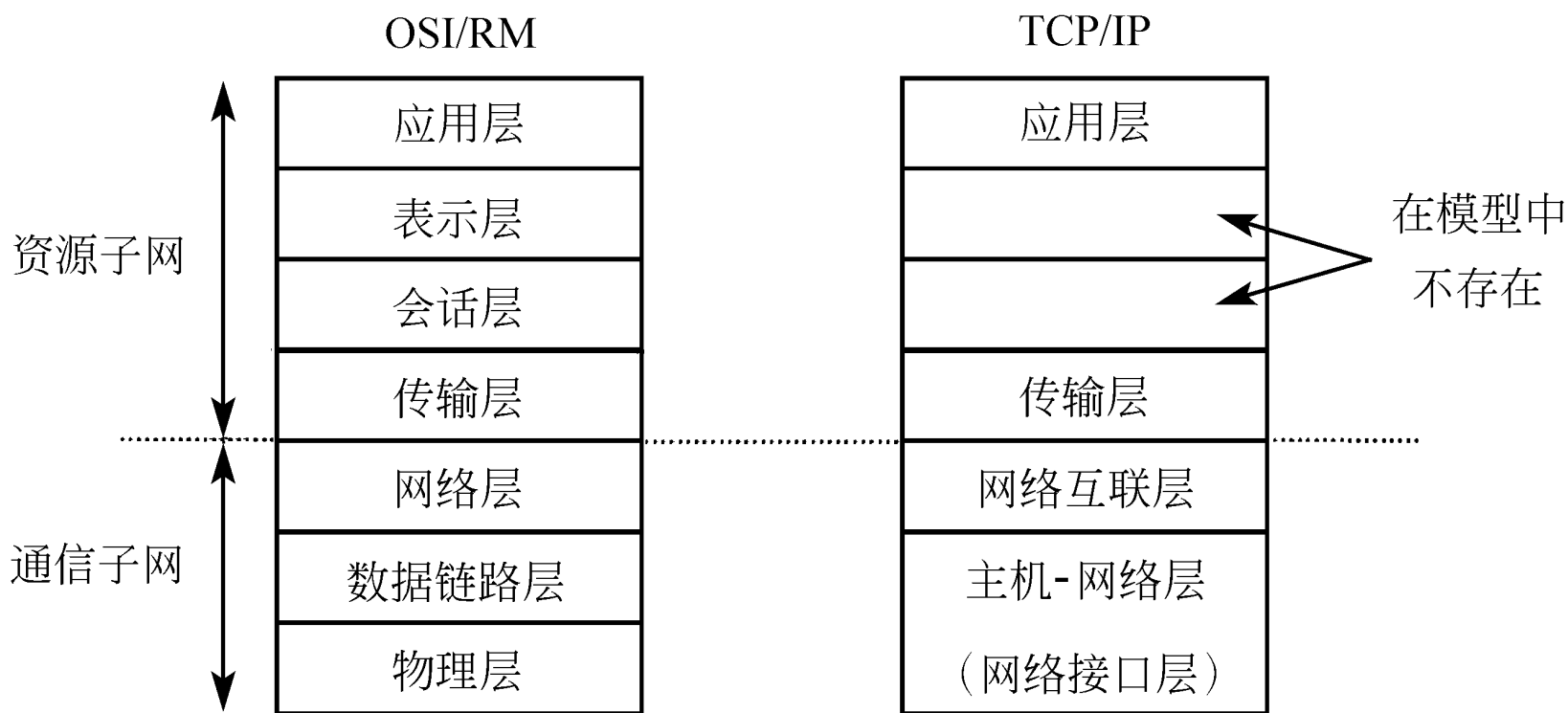
计算机网络的体系结构

■ TCP/IP体系结构（Internet 的标准）

应用层
运输层
网际层
网络接口

应用层	Dns , Http , Ftp , Telnet , Nntp , SMTP , SNMP,			
传输层	TCP		UDP	
网际层	IP	IGMP	ICMP	
			ARP	RARP
网络接口层	网络接口（物理网络）			

OSI模型与TCP/IP协议比较



1.6.3 具有五层协议的体系结构

- **TCP/IP** 是四层的体系结构：应用层、运输层、网际层和网络接口层。
- 但最下面的网络接口层并没有具体内容。
- 因此往往采取折中的办法，即综合 **OSI** 和 **TCP/IP** 的优点，采用一种只有五层协议的体系结构。（书上采用的结构）

五层协议的体系结构

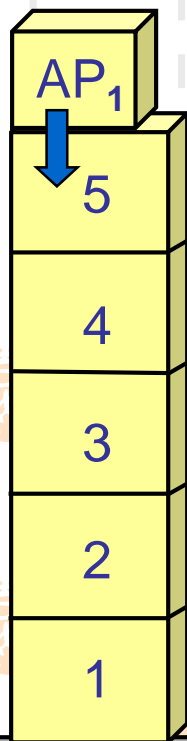


- 应用层(application layer)
- 运输层(transport layer)
- 网络层(network layer)
- 数据链路层(data link layer)
- 物理层(physical layer)



主机 1 向主机 2 发送数据

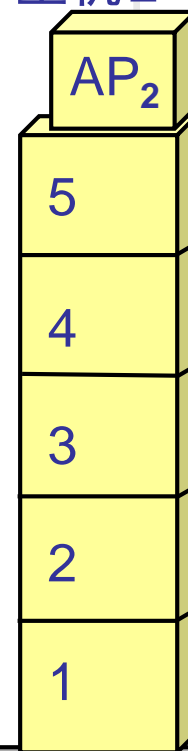
主机 1



应用进程数据先传送到应用层

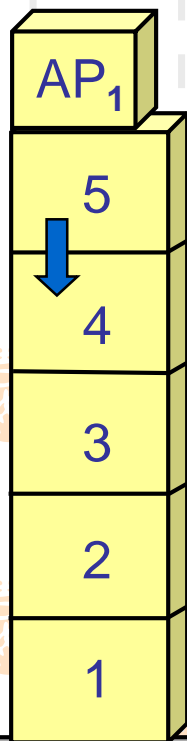
加上应用层首部，成为应用层 PDU

主机 2



主机 1 向主机 2 发送数据

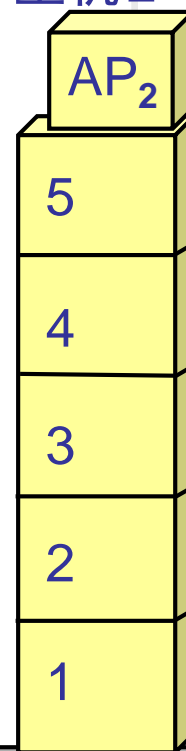
主机 1



应用层 PDU 再传送到运输层

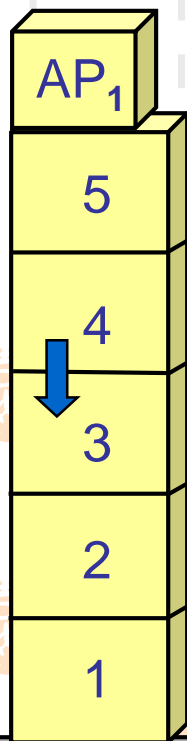
加上运输层首部，成为运输层报文

主机 2



主机 1 向主机 2 发送数据

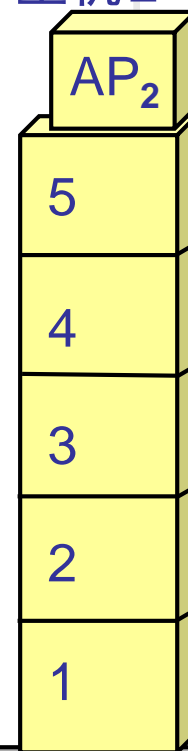
主机 1



运输层报文再传送到网络层

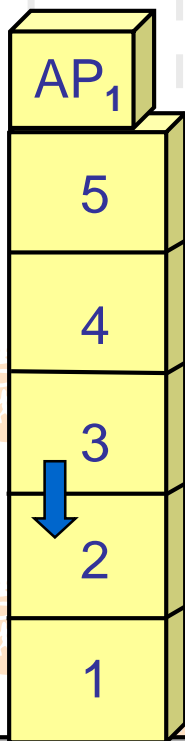
加上网络层首部，成为 IP 数据报（或分组）

主机 2



主机 1 向主机 2 发送数据

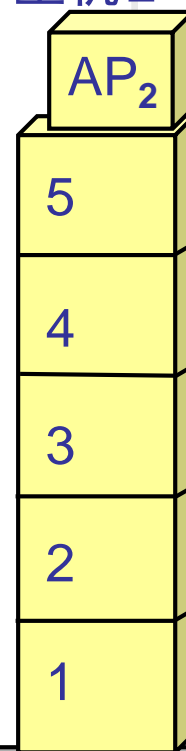
主机 1



IP 数据报再传送到数据链路层

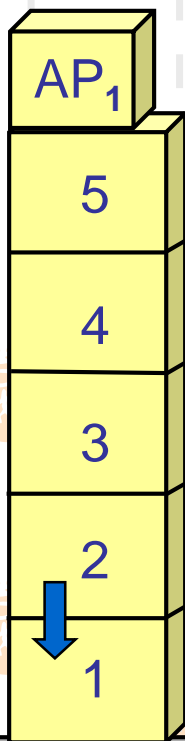
加上链路层首部和尾部，成为数据链路层帧

主机 2



主机 1 向主机 2 发送数据

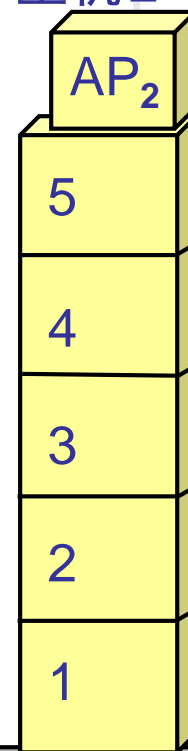
主机 1



数据链路层帧再传送到物理层

最下面的物理层把比特流传送到物理媒体

主机 2

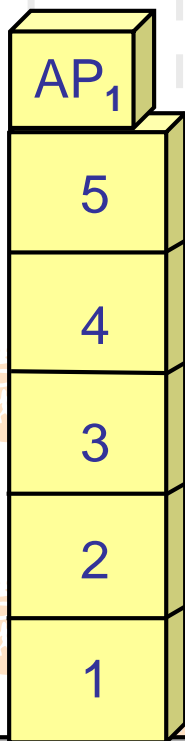


主机 1 向主机 2 发送数据

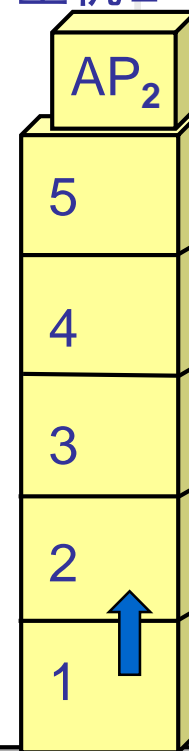


主机 1 向主机 2 发送数据

主机 1



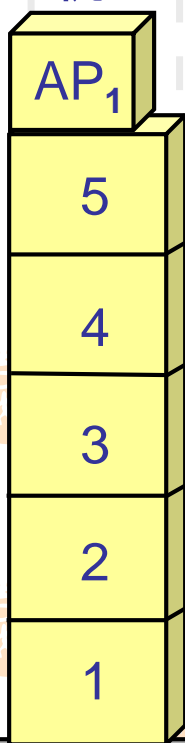
主机 2



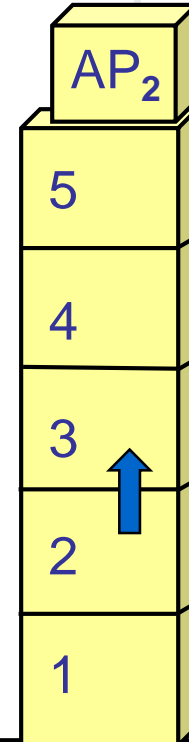
物理层接收到比特流，上交给数据链路层

主机 1 向主机 2 发送数据

主机 1



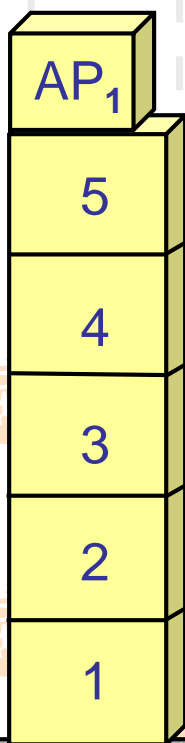
主机 2



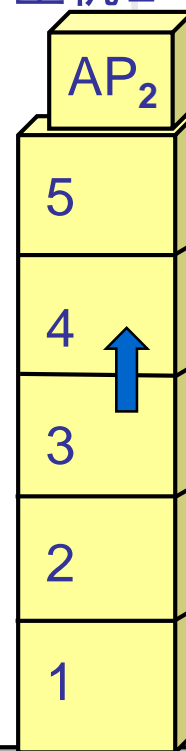
数据链路层剥去帧首部和帧尾部
取出数据部分，上交给网络层

主机 1 向主机 2 发送数据

主机 1



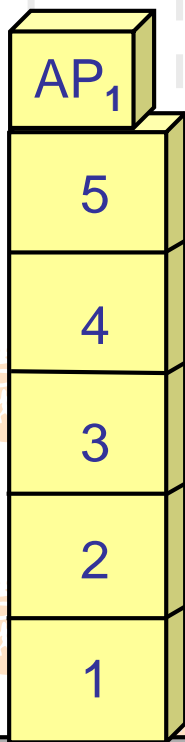
主机 2



网络层剥去首部，取出数据部分
上交给运输层

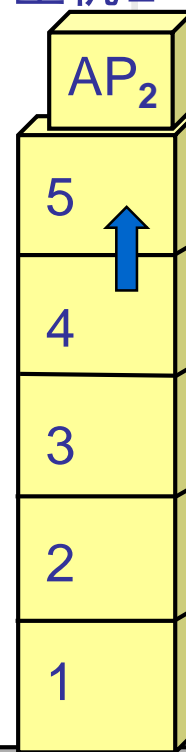
主机 1 向主机 2 发送数据

主机 1



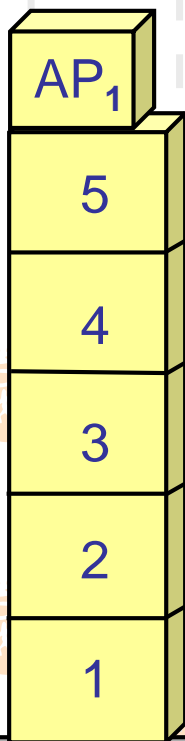
运输层剥去首部，取出数据部分
上交给应用层

主机 2



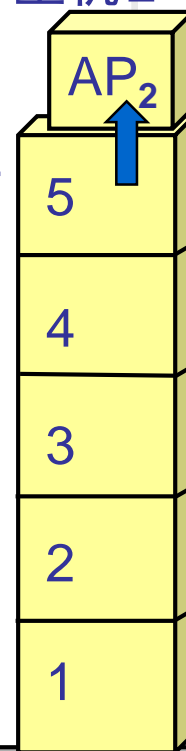
主机 1 向主机 2 发送数据

主机 1



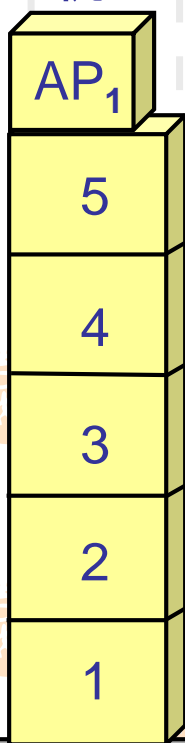
应用层剥去首部，取出应用程序数据
上交给应用进程

主机 2



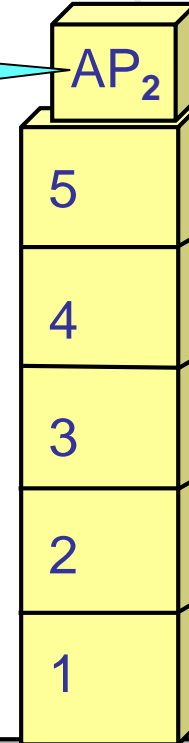
主机 1 向主机 2 发送数据

主机 1

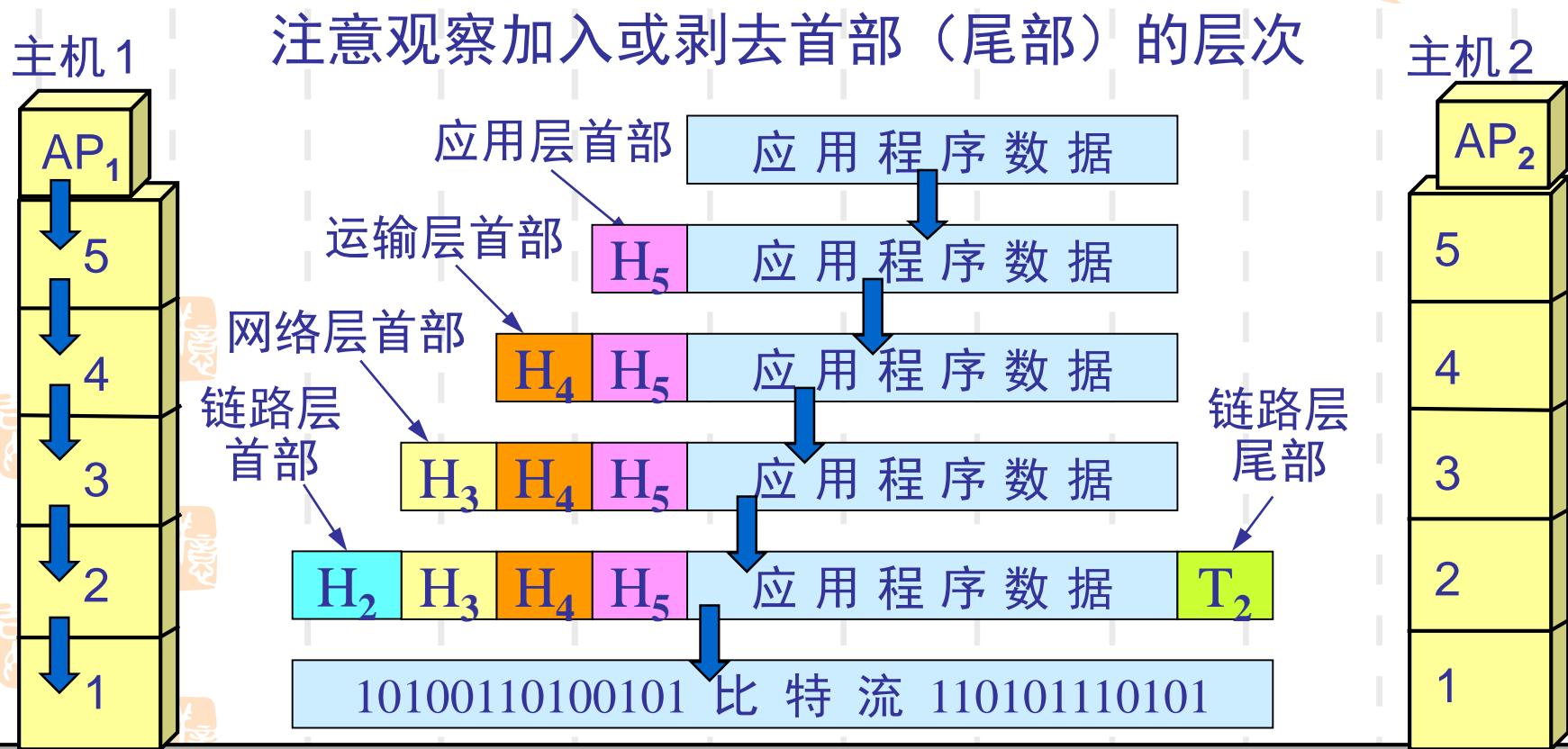


我收到了 AP_1 发来的
应用程序数据！

主机 2

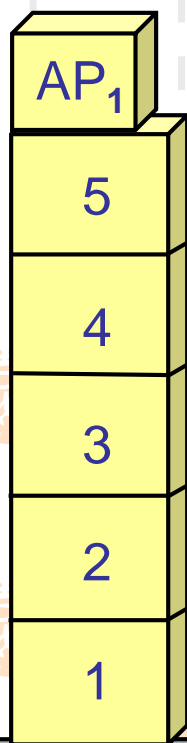


主机 1 向主机 2 发送数据

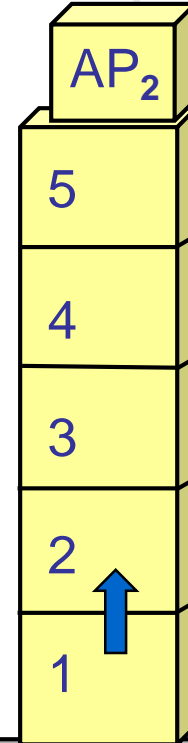


主机 1 向主机 2 发送数据

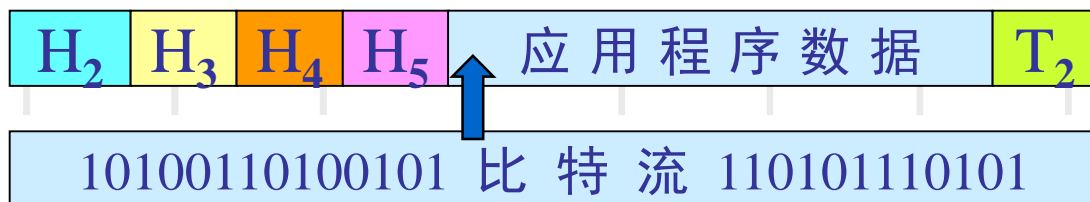
主机 1



主机 2

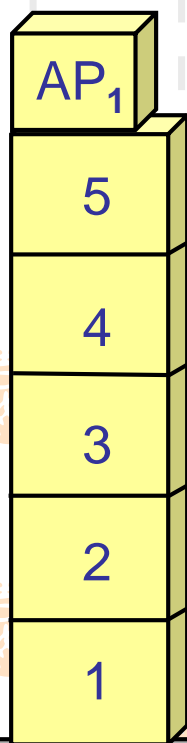


计算机 2 的物理层收到比特流后
交给数据链路层

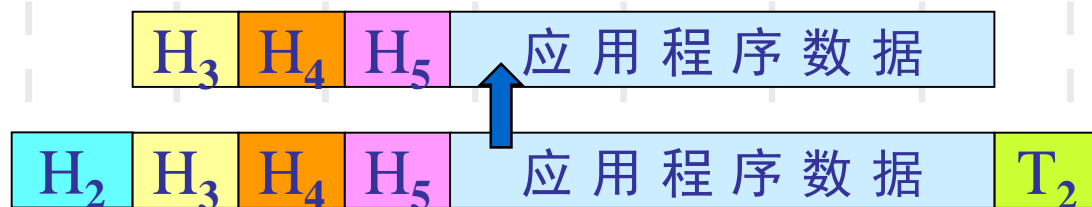


主机 1 向主机 2 发送数据

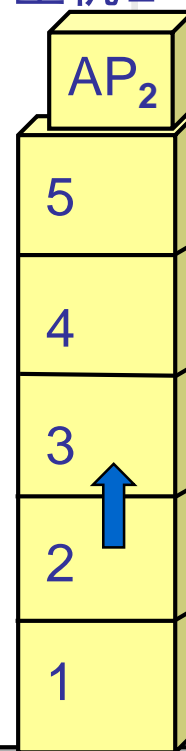
主机 1



数据链路层剥去帧首部和帧尾部后
把帧的数据部分交给网络层

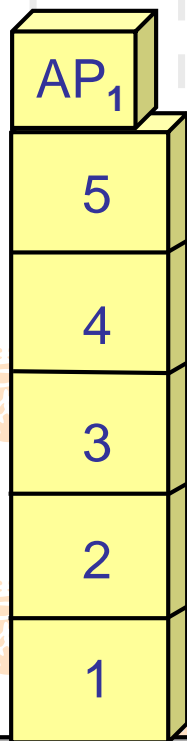


主机 2

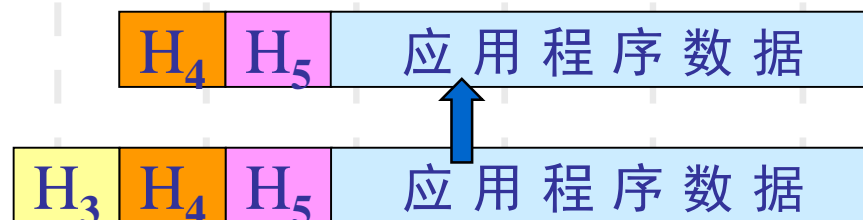


主机 1 向主机 2 发送数据

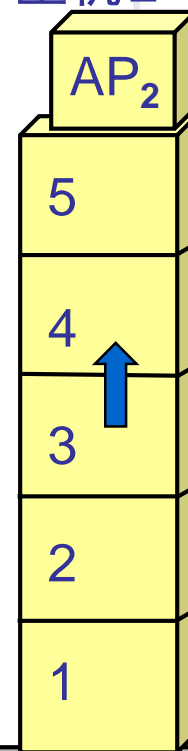
主机 1



网络层剥去分组首部后
把分组的数据部分交给运输层

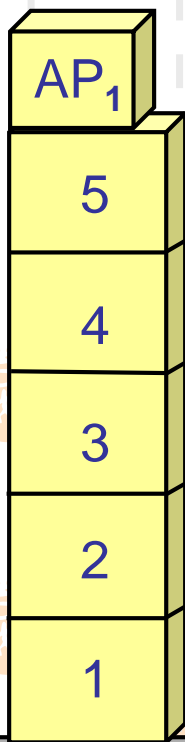


主机 2

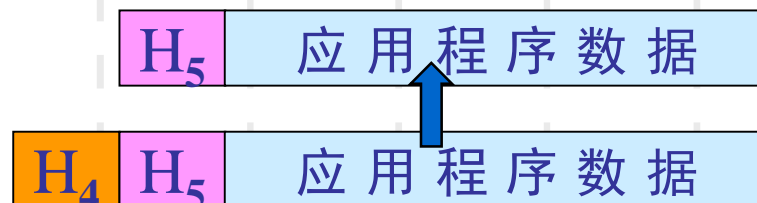


主机 1 向主机 2 发送数据

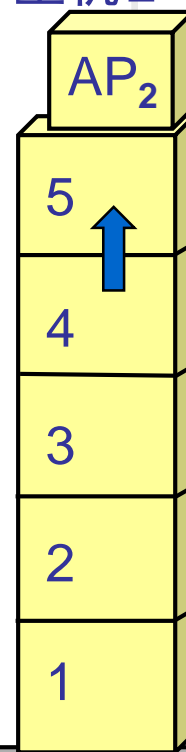
主机 1



运输层剥去报文首部后
把报文的数据部分交给应用层

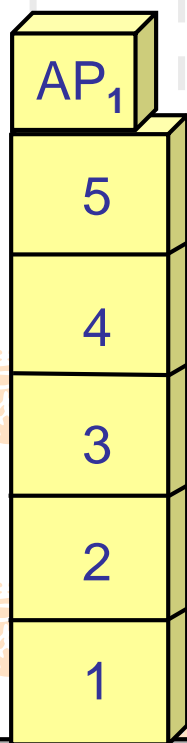


主机 2



主机 1 向主机 2 发送数据

主机 1

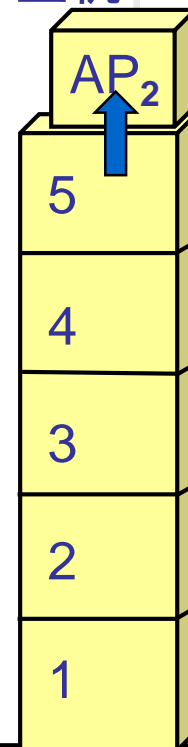


应用程序数据

H_5 应用程序数据

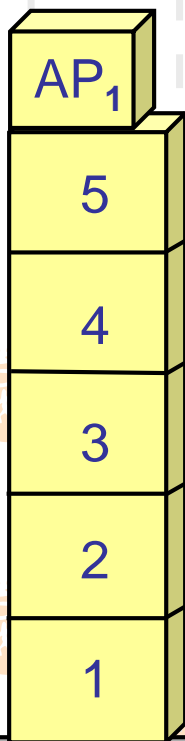
应用层剥去应用层 PDU 首部后
把应用程序数据交给应用进程

主机 2



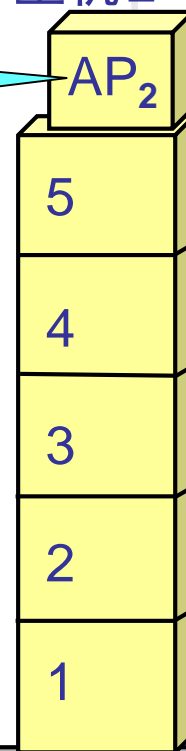
主机 1 向主机 2 发送数据

主机 1



我收到了 AP_1 发来的
应用程序数据！

主机 2



1.6.4 实体、协议、服务

- **实体 (entity)** 表示任何可发送或接收信息的硬件或软件进程。
- **协议** 是控制**两个对等实体**进行通信的规则
的集合。
- 在协议的控制下，两个对等实体间的通信
使得本层能够**向上一层提供服务**。
- 要实现本层协议，还需要**使用下层所提供的
服务**。

协议和服务在概念上是不一样的

- 协议的实现保证了能够向上一层提供服务。
- 本层的服务用户 **只能看见服务** 而无法看见下面的协议。即下面的协议对上面的服务用户是**透明**的。
- 协议是“**水平的**”，即协议是控制对等实体之间通信的规则。
- 服务是“**垂直的**”，即服务是由下层向上层通过层间接口提供的。
- 上层使用**服务原语**获得下层所提供的服务。

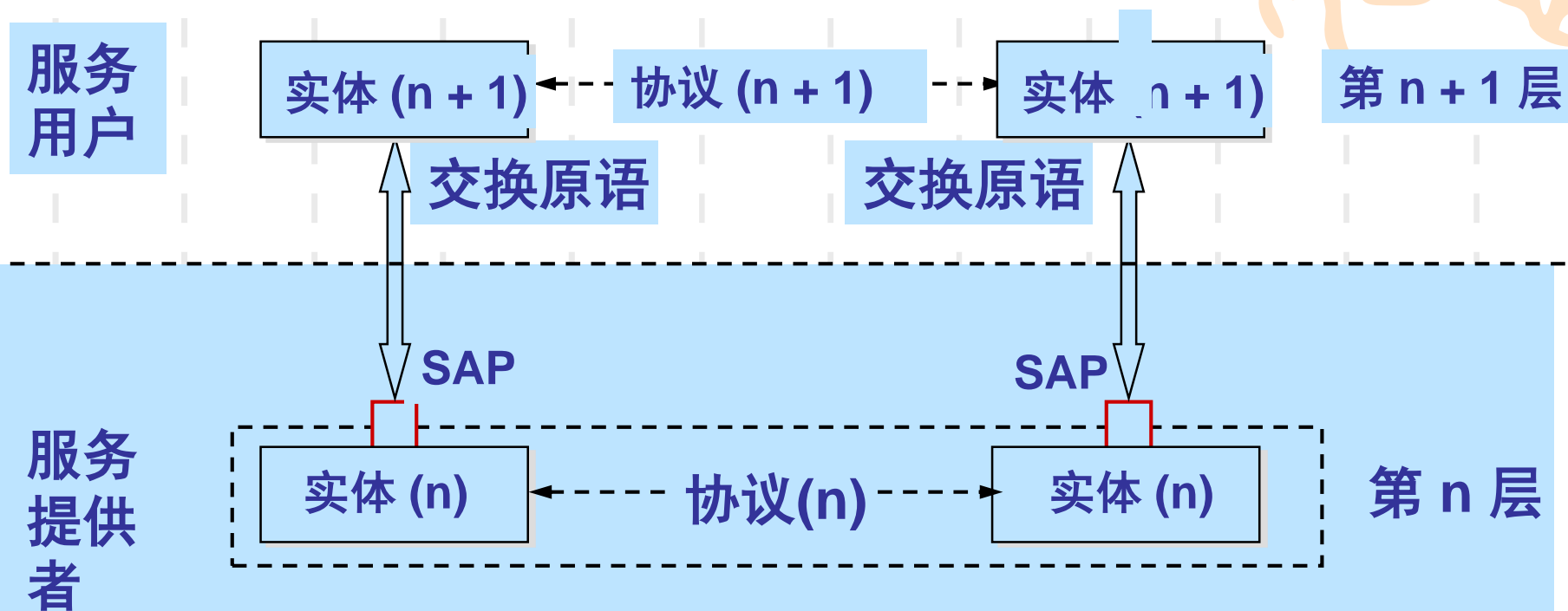
服务访问点



- 同一系统相邻两层的实体进行交互的地方，称为**服务访问点 SAP (Service Access Point)**。
- 服务访问点**SAP**是一个抽象的概念，它实际上就是一个逻辑接口。
- 传输层的服务访问点是 端口
- 网络层的服务访问点是 **IP** 地址。



1.7.4 实体、协议、服务和访问点



相邻两层之间的关系

小结



- 本章主要介绍了计算机网络的定义、组成和分类，互联网的组成和工作方式，计算机网络的主要性能指标以及计算机网络的体系结构。
- **OSI 参考模型和 TCP/IP 协议栈。**



吉祥如意

■ *The end*

