

● 互联网的组成

- **核心部分**：由大量网络和连接这些网络的路由器组成。这部分是为边缘部分提供服务的（提供连通性和交换）。
- **边缘部分**：由所有连接在互联网上的主机组成。这部分是用户直接使用的，用来进行通信（传送数据、音频或视频）和资源共享。
 - **端系统之间的通信方式**
 - ✓ 客户-服务器方式（C/S方式）**改进**→ B/S方式
 - ✓ 对等连接方式（P2P方式）

● 互联网的标准化工作

● 计算机网络类别

- 按覆盖范围分类：**广域网**、**城域网**、**局域网**、**个域网**
- 按网络所有权分类：**公用网**、**专用网**

● 性能指标



- **速率**指的是**数据的传送速率**，是**单位时间内**在信道中传输的**数据量**。
- **带宽**表示网络中某通道传送数据的能力，是网络中的某信道所能通过的“**最高数据率**”。
- **吞吐量**表示在**单位时间内**通过某个网络（或信道、接口）的**实际数据量**。
 - **吞吐量受网络的带宽或网络的额定速率的限制**。
 - ✓ 额定速率是绝对上限值。
 - ✓ 可能会远小于额定速率，甚至下降到零。
 - 有时可用**每秒传送的字节数或帧数**来表示。

性能指标——时延

- 时延 (delay 或 latency) 有时也称为**延迟**或**迟延**，是指数据（一个报文或分组，甚至比特）从网络（或链路）的一端传送到另一端所需的时间。
- 网络中的时延由以下几个不同的部分组成：
 - 也称为**传输时延**。
 - 发送时延：发送数据时，数据帧从节点进入到传输媒体所需要的时间。
 - 传播时延：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。
 - 处理时延
 - 排队时延

$$\text{发送时延} = \frac{\text{数据帧长度 (bit)}}{\text{发送速率 (bit/s)}}$$

性能指标——时延

- 时延 (delay 或 latency)有时也称为**延迟或迟延**，是指数据（一个报文）一端传送到另一端

电磁波传播速率：

- 自由空间的传播速率是光速 = 3.0×10^5 km/s
- 在铜线电缆中的传播速率约 = 2.3×10^5 km/s
- 在光纤中的传播速率约 = 2.0×10^5 km/s

- 网络中的时延由

- 发送时延
 - 传播时延
 - 处理时延
 - 排队时延
- 信号在信道中需要传播一定的距离而花费的时间。
 - 发送时延与传播时延有本质上的不同。
 - 信号发送速率和信号在信道上的传播速率是完全不同的概念。

$$\text{传播时延} = \frac{\text{信道长度 (米)}}{\text{信号在信道上的传播速率 (米/秒)}}$$

性能指标——时延

- 时延 (delay 或 latency)有时也称为**延迟**或**迟延**，是指数据（一个报文或分组，甚至比特）从网络（或链路）的一端传送到另一端所需的时间。
- 网络中的时延由以下几个不同的部分组成：
 - 发送时延
 - 传播时延
 - **处理时延**
 - 主机或路由器在收到分组时，为**处理分组**（例如，分析首部、提取数据、差错检验或查找路由）所**花费的时间**。
 - 排队时延



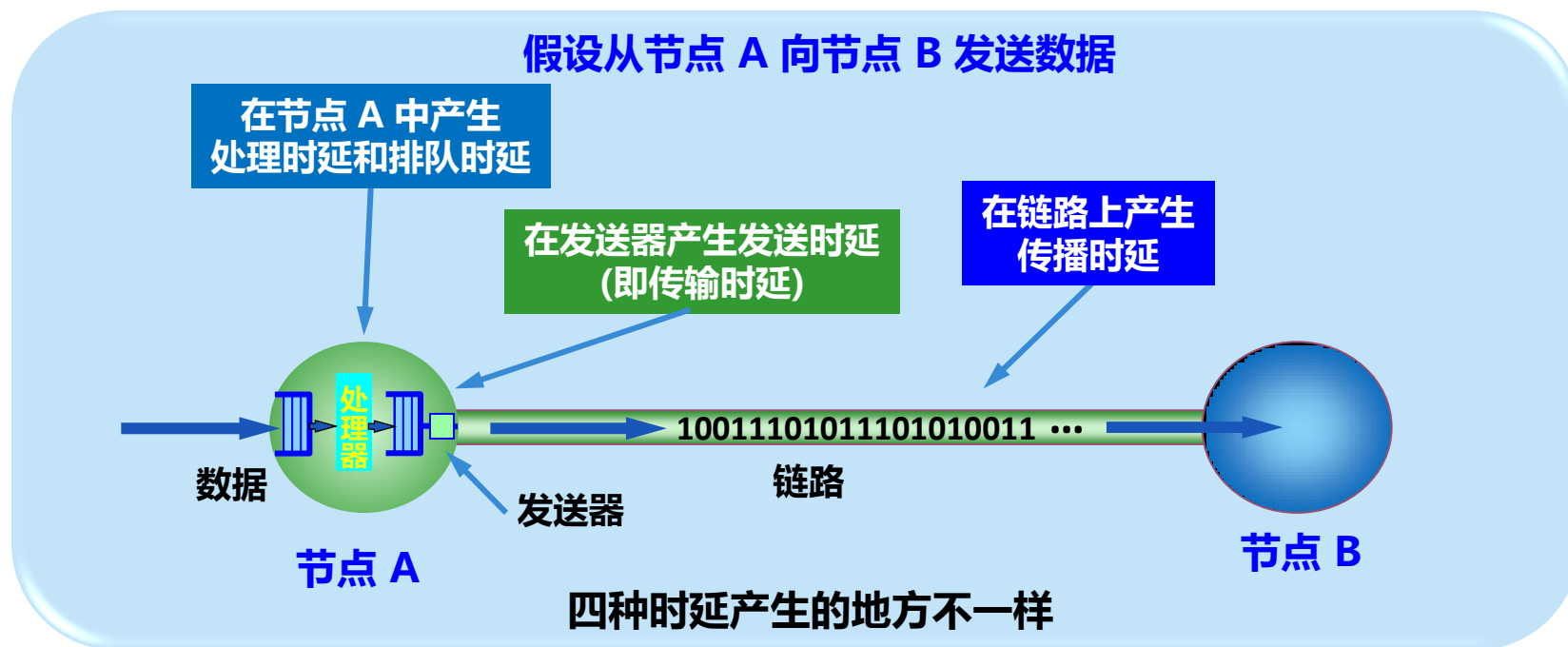
性能指标——时延

- 时延 (delay 或 latency)有时也称为**延迟或迟延**，是指数据（一个报文或分组，甚至比特）从网络（或链路）的一端传送到另一端所需的时间。
- 网络中的时延由以下几个不同的部分组成：
 - 发送时延
 - 传播时延
 - 处理时延
 - **排队时延**
 - 分组在路由器的接收缓存队列中**排队等待接收的时间**与在发送缓存队列中**排队等待发送**所经历的时延之和。
 - **排队时延的长短**往往取决于网络中**当时的通信量**。

性能指标——时延

总时延 = 发送时延 + 传播时延 + 处理时延 + 排队时延

注意：在总时延中，究竟是哪一种时延占主导地位，必须具体分析。



性能指标——时延

容易产生的错误概念：

- 对于高速网络链路，我们提高的仅仅是数据的**发送速率**而不是比特在链路上的**传播速率**。
- 提高数据的发送速率只是减小了数据的发送时延。

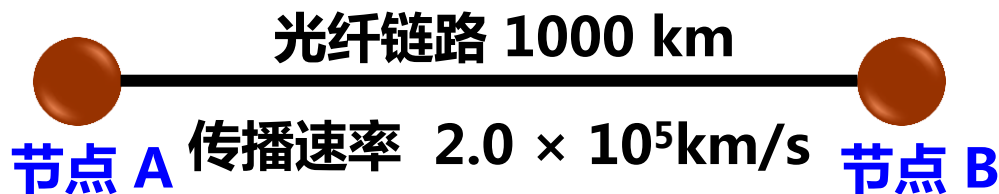


以下说法是错误的：

- “在高速链路（或高带宽链路）上，比特会传送得更快些”。



性能指标——时延



节点 A 要将一个数据块通过 1000 km 的光纤链路发送给节点 B。忽略处理时延和排队时延。请分别计算下列情况时的总时延，并验证“数据的发送速率越高，其传送的总时延就越小”的说法是否正确。

- (1) 数据块大小为 100 MB，信道带宽为 1 Mbit/s；
- (2) 数据块大小为 100 MB，信道带宽为 100 Mbit/s；
- (3) 数据块大小为 1 B，信道带宽为 1 Mbit/s；
- (4) 数据块大小为 1 B，信道带宽为 1 Gbit/s。



性能指标——时延

不能笼统地认为“数据的发送速率越高，其传送的总时延就越小”。

解：

$$\text{传播时延} = 1000 \text{ km} / 2.0 \times 10^5 \text{ km/s} = 5 \text{ ms}.$$

$$(1) \text{ 发送时延} = 100 \times 2^{20} \times 8 \div 10^6 = 838.9 \text{ s},$$

$$\text{总时延} = 838.9 + 0.005 \approx 838.9 \text{ s}.$$

$$(2) \text{ 发送时延} = 100 \times 2^{20} \times 8 \div 10^8 = 8.389 \text{ s}$$

$$\text{总时延} = 8.389 + 0.005 = 8.394 \text{ s}.$$
 缩小到 (1) 的近 1/100.

$$(3) \text{ 发送时延} = 1 \times 8 \div 10^6 = 8 \times 10^{-6} \text{ s} = 8 \mu\text{s},$$

$$\text{总时延} = 0.008 + 5 = 5.008 \text{ ms}.$$

$$(4) \text{ 发送时延} = 1 \times 8 \div 10^9 = 8 \times 10^{-9} \text{ s} = 0.008 \mu\text{s}$$

$$\text{总时延} = 0.000008 + 5 = 5.000008 \text{ ms}.$$
 与 (3) 相比没有明显减小.

性能指标——时延带宽积

带宽延时积(Bandwidth-Delay Product, BDP)

- 链路的时延带宽积又称为以**比特为单位的链路长度**。

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



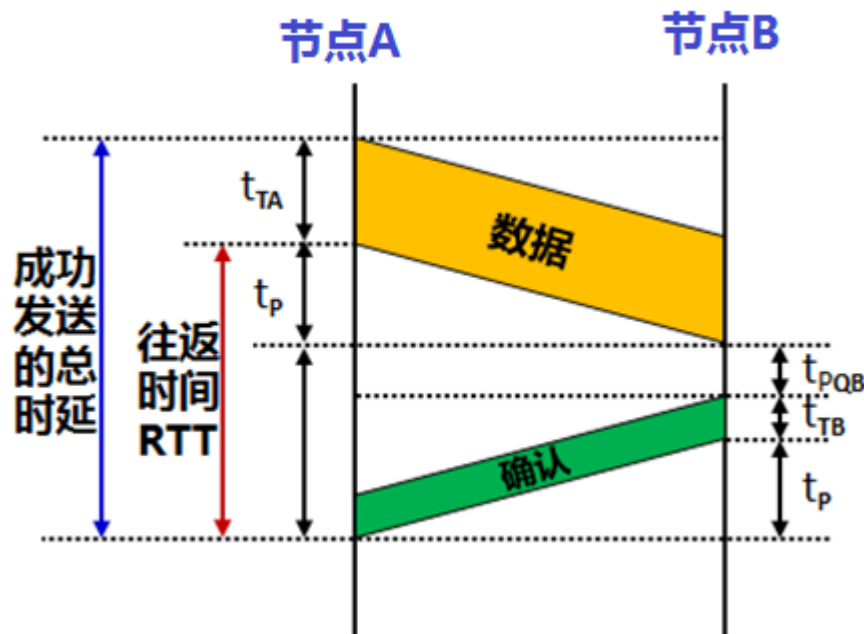
只有在代表链路的管道都充满比特时，
链路才得到了充分利用。

性能指标——往返时间 (RTT)

- **往返时间 RTT** (Round-Trip Time) 表示从发送端**发送数据开始**，到发送端**收到**来自接收端的**确认**，**总共经历的时间**。
- 在互联网中，往返时间还包括**各中间节点**的处理时延、排队时延以及转发数据时的发送时延。

往返时间 RTT

$$\begin{aligned} &= \text{节点 A 到 B 的传播时延 } t_p \\ &+ \text{节点 B 处理和排队时延 } t_{pQB} \\ &+ \text{节点 B 发送时延 } t_{TB} \\ &+ \text{节点 B 到 A 的传播时延 } t_p \\ &= 2 \times \text{传播时延 } t_p \\ &+ \text{节点 B 处理和排队时延 } t_{pQB} \\ &+ \text{节点 B 发送时延 } t_{TB} \end{aligned}$$



性能指标——往返时间 (RTT)

- **往返时间 RTT** (Round-Trip Time) 表示从发送端**发送数据开始**，到发送端**收到**来自接收端的**确认**，**总共经历的时间**。
- 在互联网中，往返时间还包括**各中间节点**的处理时延、排队时延以及转发数据时的发送时延。
- **当使用卫星通信时，往返时间 RTT 相对较长，是很重要的一个性能指标。**

$$\text{有效数据率} = \frac{\text{数据长度}}{\text{发送时延} + \text{RTT}}$$

性能指标——利用率

信道利用率

- 某信道有百分之几的时间是被利用的，即有数据通过。
- 完全空闲的信道的利用率是零。

$$\text{信道利用率} = \frac{\text{有数据通过时间}}{(\text{有} + \text{无}) \text{数据通过时间}}$$

网络利用率

- 全网络的信道利用率的**加权平均值**。

问题：信道利用率越高越好吗？

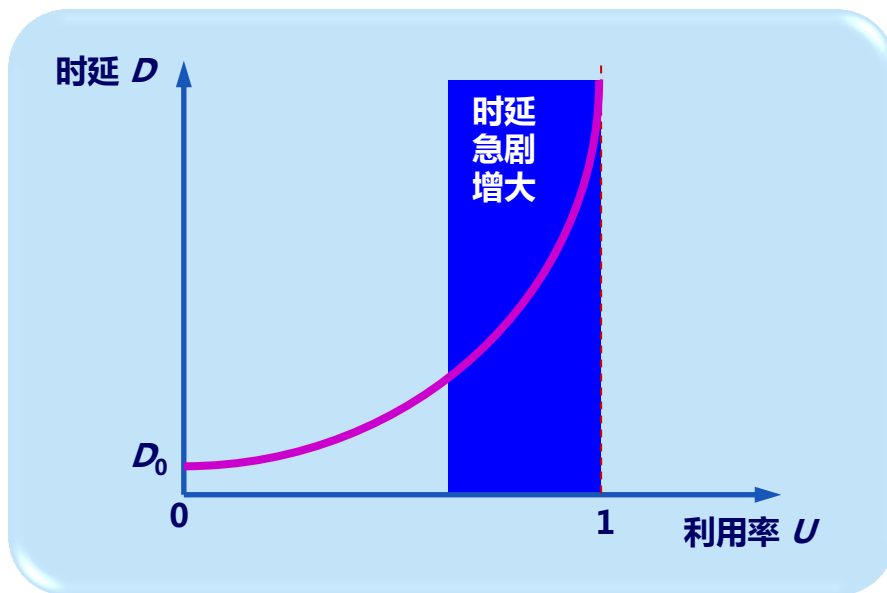
性能指标——利用率

- 信道利用率**并非越高越好**。
- 根据排队论，当某信道的利用率增大时，时延会迅速增加。

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

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

其中， U 是网络的利用率，数值在
0 到 1 之间。



性能指标

小 结

速率	数据传输速率
带宽	有两种含义：频率成分之差（频域）、数据传输速率（时域）
吞吐量	单位时间内通过网络、信道或接口的数据量
时延	发送时延，处理时延，排队时延，传播时延
往返时延	在链路上发送数据开始到收到确认的时间
时延带宽积	传播时延 \times 带宽
利用率	信道利用率，网络利用率

非性能指标

计算机网络的一些非性能指标与前面介绍的性能指标有很大的关系。



基本概念

- 计算机网络体系结构是指计算机网络的**层次结构模型**与其**各层协议的集合**。也就是说，计算机网络体系结构是关于计算机网络应该设置哪些层次，每个层次又应该提供哪些功能的精确定义。
- 计算机网络**体系结构**是**抽象**的，而实现网络协议的技术是具体的，是指一些能够运行的软/硬件。
- 计算机网络体系结构采用层次模型的优点
 - 各层次之间相互独立
 - 灵活性好
 - 易于实现和维护
 - 有利于促进标准化

基本概念

- **实体** (Entity) 表示任何可发送或接收信息的**软件进程** (如某一特定的软件模块) 或实现该层协议的**硬件单元** (如网卡、智能I/O芯片)。
- **网络协议** (network protocol), 简称为**协议**, 是控制两个**对等实体** (或多个实体) 进行通信的规则的组合。

➤ 协议三要素

- **语法** : 如何讲, 用来规定信息格式, 涉及数据与控制信息的格式、编码和信号等级 (电平的高低) 等。
- **语义** : 讲什么, 用来说明通信双方应该怎么做, 涉及数据的内容、含义及用于协调和差错处理的控制信息。
- **定时 (时序)** : 详细说明事件的先后顺序, 涉及速率匹配和排序等。

➤ 协议的制定和实现采用**分层**结构

