



计算机网络的定义：

计算机网络是使用单一技术互相连接的、**自治计**
算机的集合。





网络建立的主要目的是实现主机**通信**和资源的**共享**。





计算机网络的性能指标:

- 带宽
- 时延 (发送时延、传播时延、时延带宽积、往返时延)
- 吞吐量
- 利用率





计算机网络的性能指标—带宽

“带宽” (bandwidth) 本来是指信号具有的频带宽度，即最高频率和最低频率之差，单位是赫兹（或千赫、兆赫、吉赫等）。

线路的带宽表示通信线路允许通过的信号频带范围。

如传统的通信线路上传送的电话信号的标准带宽是 3.1kHz(即话音的频率范围是从300Hz到3400Hz) 。



计算机网络的性能指标—带宽



在计算机网络中，**带宽**表示在单位时间内从网络中的某一点到另一点所能通过的“**最高数据率**”，单位是“**比特每秒**”，或 b/s (bit/s)。

更常用的速率单位是

千比每秒，即 kb/s (10^3 b/s)

兆比每秒，即 Mb/s (10^6 b/s)

吉比每秒，即 Gb/s (10^9 b/s)

太比每秒，即 Tb/s (10^{12} b/s)





在计算机中的数据容量用字节B (Byte) 作为度量单位。 “千” 字节用KB表示, $1\text{KB} = 2^{10}\text{B} = 1024\text{B}$; 而 $1\text{MB} = 2^{20}\text{B}$, $1\text{GB} = 2^{30}\text{B}$, $1\text{TB} = 2^{40}\text{B}$ 。

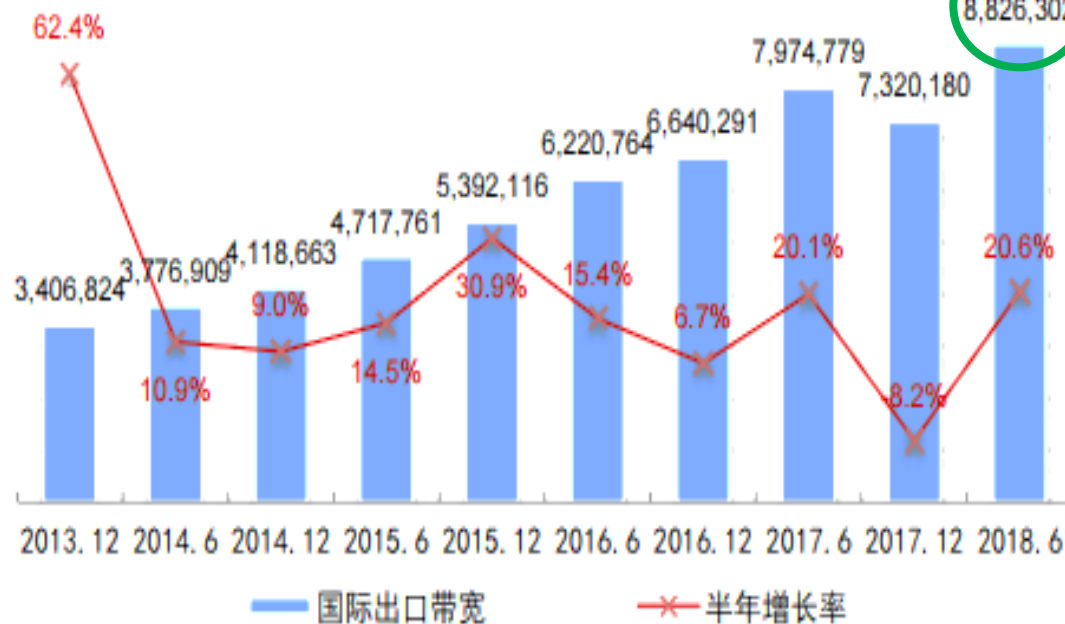


计算机网络的性能指标—带宽



中国国际出口带宽及增长率

单位: Mbps



8826Gb
ps

8.826Tb
ps

来源: CNIC 中国互联网络发展状况统计调查

2018.6

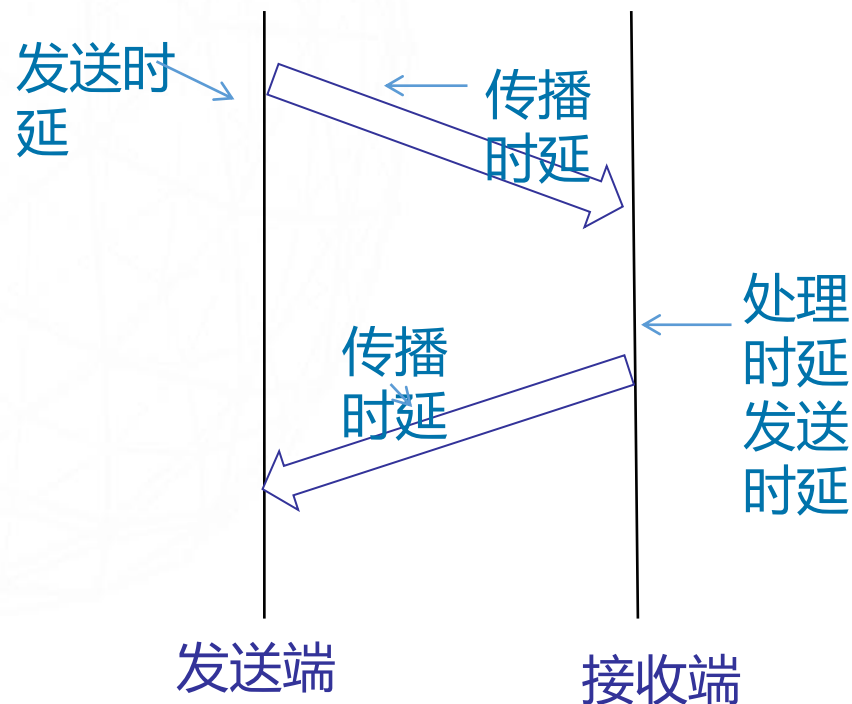


计算机网络的性能指标—时延

时延是计算机网络最重要的两个性能指标之一。

时延是指数据（一个报文或分组，比特）从网络（或链路）的一端传送到另一端所需的时间。网络时延包含以下几种时延：

- 发送时延
- 传播时延
- 处理时延
- 排队时延





计算机网络的性能指标—时延

- **发送时延**（传输时延）。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。

$$\text{发送时延} = \frac{\text{数据块长度（比特）}}{\text{信道带宽（比特/秒）}}$$





计算机网络的性能指标—时延

- **传播时延** 电磁波在信道中需要传播一定的距离而花费的时间。

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

电磁波在自由空间中的传播速率是光速，即 $3.0 \times 10^8 \text{ km/s}$ 。电磁波在网络传输媒体中的传播速率比自由空间中略低一些：在铜线电缆中的传播速率约为 $2.3 \times 10^8 \text{ km/s}$ ，在光纤中的传播速率约为 $2.0 \times 10^8 \text{ km/s}$ 。

如：1000km长的光纤线路产生的传播时延 $= \frac{1000 \text{ km}}{2.0 \times 10^8 \text{ km/s}} = 0.005 \text{ s} = 5 \text{ ms}$





计算机网络的性能指标—时延

- **处理时延** 主机或路由器在收到分组进行一些必要的处理所花费的时间。
- **排队时延** 路由器中分组排队所经历的时延是处理时延中的重要组成部分。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。





计算机网络的性能指标—时延

数据在网络中经历的**总时延**就是发送时延、传播时延与处理时延之**和**：

$$\text{总时延} = \text{发送时延} + \text{传播时延} + \text{处理时延}$$





三种时延所产生的地方

在队列中产生
处理时延

从结点 A 向结点 B 发送数据

在发送器产生发送时延
(即传输时延)

在链路上产生
传播时延



发送时延发生在发送器中，与传输信道长度无关。

传播时延发生在传输信道媒体上，与信道的带宽无关。



思考：数据在网络经历的总时延中，哪一种时延占主导地位？





计算举例：

(1) 数据块长度为100MB，信道带宽为1Mb/s，将此数据块用光纤传送到1000km远的计算机，忽略处理时延和排队时延，分别求发送时延和传播时延？

计算：发送时延：
$$\frac{100 \times 2^{20} \text{B} \times 8 \text{bit}}{1 \times 10^6 \text{b/s}} = 838.9 \text{s}$$

传播时延：
$$\frac{1000 \text{km}}{2 \times 10^5 \text{km/s}} = 5 \times 10^{-3} \text{s} = 5 \text{ms}$$

发送时延占主导。





计算举例：

(2) 数据块长度为1B，信道带宽为1Mb/s，将此数据块用光纤传送到1000km远的计算机，忽略处理时延和排队时延，分别求发送时延和传播时延？

计算：发送时延：
$$\frac{1 \times 8 \text{ bit}}{1 \times 10^6 \text{ bit/s}} = 8 \times 10^{-6} \text{ s}$$

传播时延：
$$\frac{1000 \text{ km}}{2 \times 10^5 \text{ km/s}} = 5 \times 10^{-3} \text{ s}$$

传播时延占主导。





- 对于高速网络链路，我们提高的仅仅是数据的**发送速率**（即带宽）而不是比特在链路上的**传播速率**。
- 提高链路**带宽**减小了数据的**发送时延**。





计算机网络的性能指标—时延带宽积

时延带宽积 链路的**时延带宽积**又称为以**比特**为单位的链路长度。

管道中的比特数表示从发送端已经发出但尚未到达接收端的比特。

时延带宽积

(传播) 时延

带宽

链路

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





计算机网络的性能指标—往返时延

- **往返时延 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（接收端收到数据后立即发送确认），总共经历的时延。





计算机网络的性能指标—往返时延

互联网上的信息往往是双向交互的，例如A向B发送数据，如果数据长度是100MB，发送速率是100Mbps，那么

$$\text{发送时间} = \frac{\text{数据长度}}{\text{发送速率}} = \frac{100 \times 2^{20} \times 8}{100 \times 10^6} \approx 8.39\text{s}$$

若B正确收到100MB的数据后，就立即向A发送一个确认；假定 A只有正确收到B的确认信息后，才能继续向B发送数据。显然A需要等待一个往返时延RTT。





在互联网中，往返时延还包括各中间结点的处理时延、排队时延以及转发数据时的发送时延。





计算机网络的性能指标—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。

- 吞吐量常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速率的限制。





计算机网络的性能指标—利用率

信道利用率指出某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是零。

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。



信道或网络的利用率过高会产生非常大的时延

- 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。

时延
 D

