

1.4 分组交换网中的时延、丢包和吞吐量

回想在 1.1 节中我们讲过，因特网能够看成是一种为运行在端系统上的分布式应用提供服务的基础设施。在理想情况下，我们希望因特网服务能够在任意两个端系统之间瞬间移动我想要的大量数据而没有任何数据丢失。然而，这是一个极高的目标，实践中难以达到。与之相反，计算机网络必定要限制在端系统之间的吞吐量（每秒能够传送的数据量），在端系统之间引入时延，而且实际上能够丢失分组。一方面，现实世界的物理定律引入的时延、丢包并限制吞吐量是不幸的。而另一方面，因为计算机网络存在这些问题，围绕如何去处理这些问题有许多令人着迷的话题，多得足以开设一门有关计算机网络方面的课程，可以做上千篇博士论文！在本节中，我们将开始研究和量化计算机网络中的时延、丢包和吞吐量等问题。

1.4.1 分组交换网中的时延概述

前面讲过，分组从一台主机（源）出发，通过一系列路由器传输，在另一台主机（目的地）中结束它的历程。当分组从一个结点（主机或路由器）沿着这条路径到后继结点（主机或路由器），该分组在沿途的每个结点经受了几种不同类型的时延。这些时延最为重要的是**结点处理时延**（nodal processing delay）、**排队时延**（queuing delay）、**传输时延**（transmission delay）和**传播时延**（propagation delay），这些时延总体累加起来是**结点总时延**（total nodal delay）。许多因特网应用，如搜索、Web 浏览、电子邮件、地图、即时讯息和 IP 语音，它们的性能受网络时延的影响都很大。为了深入理解分组交换和计算机网络，我们必须理解这些时延的性质和重要性。

时延的类型

我们来探讨一下图 1-16 环境中的这些时延。作为源和目的地之间的端到端路径的一部分，一个分组从上游结点通过路由器 A 向路由器 B 发送。我们的目标是在路由器 A 刻画出结点时延。值得注意的是，路由器 A 具有通往路由器 B 的出链路。该链路前面有一个队列（也称为缓存）。当该分组从上游结点到达路由器 A 时，路由器 A 检查该分组的首部以决定该分组的适当出链路，并将该分组导向该链路。在这个例子中，对该分组的出链路是通向路由器 B 的那条链路。仅当在该链路没有其他分组正在传输并且没有其他分组排在该队列前面时，才能在这条链路上传输该分组；如果该链路当前正繁忙或有其他分组已经在该链路上排队，则新到达的分组则将参与排队。

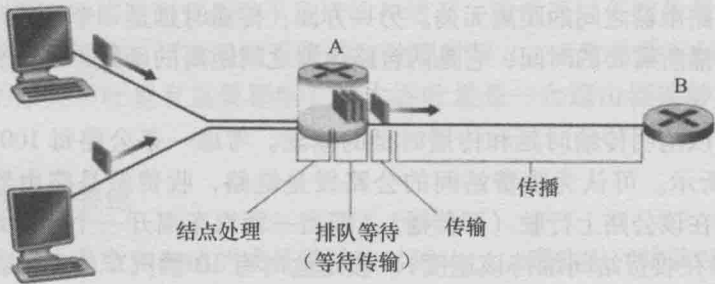


图 1-16 路由器 A 的结点时延