15-441 Computer Networks — **Lecture 1**

Mor Harchol-Balter — **August 26, 2002**

Lecture 1: Introduction, Network Infrastructure, Delays, Throughput, Bandwidth

1 Why are you here?

- **Question:** How many of you plan to go to industry after finishing your undergraduate degree? How many of you are planning on graduate school?
- Chances high that you use material from this class.
- Cisco, AT&T, IBM, Intel, Microsoft, all consider themselves to be "networking companies."
- Networking material is popular choice for interviewing questions :-)

2 What's different this year?

As I'm sure you've heard, 15-441 is a project-oriented class. For the past 10 years, 15-441 has used the same project set. This project set has involved writing a simplified version of IP and TCP on top of a simulated environment. While this project set has been instructive, we felt that it could use some updating. This summer your 4 TAs, and Srini and I worked on devising a new project set. We worked all summer, but there may still be kinks, so please be patient! The new project set differs in the following respects:

1. We ditched the simulator. Students have complaned for years that the simulated environment made the projects very difficult to debug and unrealistic in many ways. This year, all the code you write will be done directly on top of IP or UDP. This is exactly what you'll have to do when you work at a company.
2. Instead of just re-programming the existing IP and TCP protocols, you will get to define and program your own protocols as well. You will design protocols tailored to applications like sending videos and peer-to-peer networking. Again, the aim here is to give you experience doing the kind of programming and design that you'll be asked to do at a company or in graduate school.

3. We have added a measurement component. A lot of networking research today involves measuring the performance of the internet. We feel that it's important to introduce you to the tools used to make such measurements.

3 Helpful Hints

- Networking is Easy. Two skills needed:
  1. *Memory* for the thousand definitions and acronyms.
  2. Strong sense of *logic*. Solution used is always the logical one. Keep in mind that the heart of networking is all algorithms.
- Best way to get your networking questions answered is to use the Internet itself. Search the Web!

4 Beginning Networking Vocabulary

First 3 lectures are introductory. You may have heard it all before, but for completeness we're starting at the beginning.

The **big goal** in networking is to allow **applications** running on different computers to **communicate** with each other, reliably, and speedily.

The machines on which the applications are running are often called **end systems**.

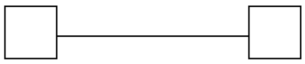
Question: What's an **end system** (a.k.a. **end host**)?

Answer: Typically a computing device: PC, Unix-based workstation, server, pager, mobile computer, toaster etc..

End systems are connected by some **physical medium**: could be copper wire, optical fiber, coaxial cable, radio spectrum, etc.. We'll discuss the various physical media in detail in a later lecture (10/28). The physical medium is abstractly referred to as a **communication link**.

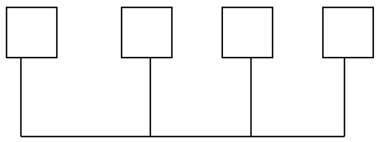
Communication links have different **transmission rates**, or **bandwidths** (bits/second that they are capable of sending.

A link is a **point-to-point** link if it connects two nodes:



Question: If you want to connect > 2 end systems to each other, what do you need?

Answer: One option is to use a single **multiple-access link**, e.g. Ethernet wire to connect them. This works well if the endsystems are close together, and utilization of the link is low.



If the end systems are far apart or attached to different physical media, we may need a **switch** or **router** (we use these terms interchangeably). A switch has several **interfaces** or **ports**. The purpose of a switch is to **forward** packets from one link to another.

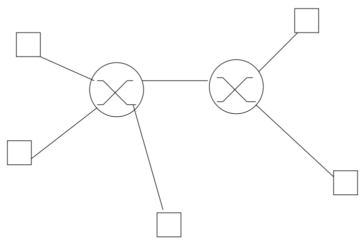


Figure 1: *Picture of switched network. Circles represent switches. Boxes represent hosts.*

When one host wants to send message to another end host, it sends message over a path or **route** through the network.

Question: What is an **internetwork**?

Answer: An internetwork is a set of independent networks that are interconnected. A **"network of networks"**.

Question: What is the Internet?

Answer: The word "internet" is short for internetwork and just means a network of networks. The word Internet (with a capital I) denotes an internet, where all hosts and switches speak the IP protocol and all have IP addresses. Anyone who obeys the IP protocol can join the Internet, including your toaster!

Question: So what's the **IP protocol**?

Answer: IP protocol defines a specific communication format for routers and hosts to use when forwarding messages. We'll get there soon – in fact we will spend a third of the class on this.

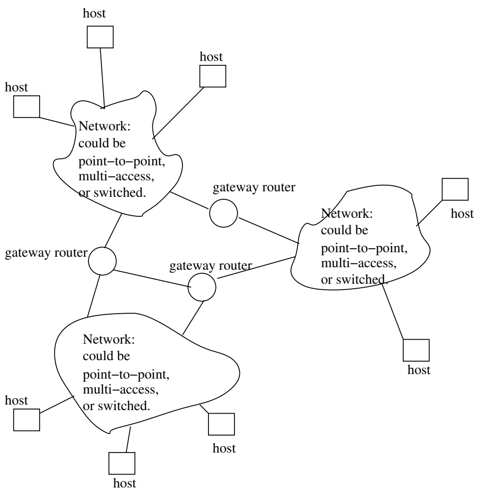


Figure 2: *Picture of internetwork.*

5 Pricing on the Internet

A **single network**, a.k.a. an **autonomous system**, or **administrative domain** is a network which is administratively fairly homogeneous. Same set of protocols/policies are used, although the physical components may be different. All managed by the same organization. Typically no charges for "internal" communication within the network.

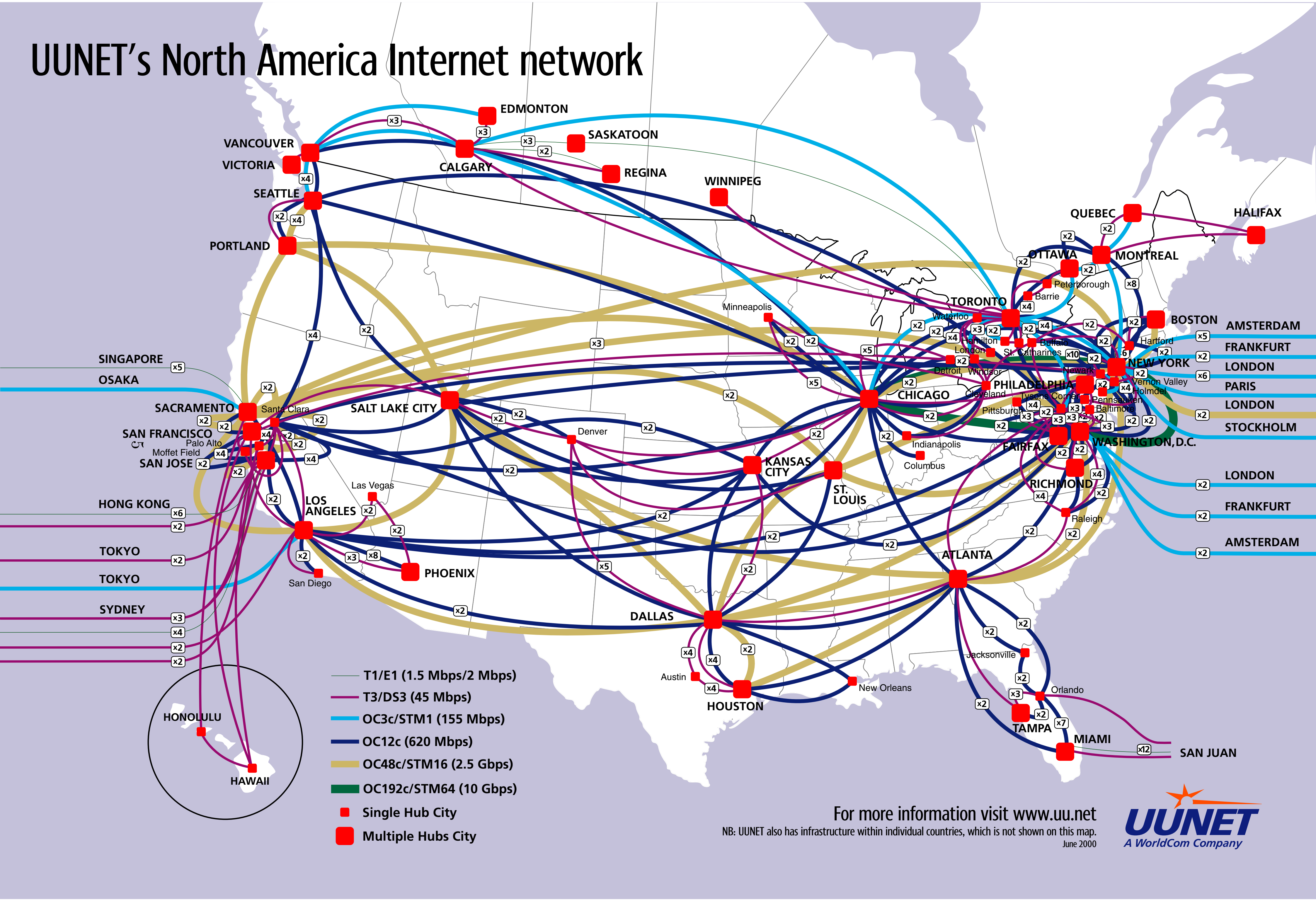
We saw that the Internet is a network of autonomous systems.

There is *another way* to think about the Internet: as a *money machine* consisting of a hierarchy of **ISP**s - Internet Service Providers. At the top of the hierarchy there are national or international ISPs, called national backbone providers **NBP**s. Each NBP is a backbone network spanning North America and possibly the world. NBPs include UUNet Technologies, SprintLink, internetMCI, and others. The fact that there is more than one NBP implies there must be lots of money to be made in this business. Check out this web site for a list:

http://www.nthelp.com/maps.htm

On this web site, you can see the network which each NBP spans.

NBPs have high-bandwidth transmission links (as high as 10 Gbps). The goal of the NBPs is to make money. They do this by charging out their bandwidth to regional ISPs (Internet Service Providers). Note: running an NBP is not cheap, so the money made is not all profit. The **regional ISPs** connect to the NBPs at certain hubs. Regional ISPs pay the NBPs a fee proportional to the amount of bandwidth they need from the NBP. The regional ISPs



UUNET's North America Internet network
EDMONTON
SASKATOON
VANCOUVER
VICTORIA
CALGARY
REGINA
WINNIPEG
SEATTLE
PORTLAND
QUEBEC
HALIFAX
OTTAWA
MONTREAL
Peterborough
Barrie
TORONTO
Minneapolis
Waterloo
Hamilton
London
St. Catharines
Buffalo
Detroit
Windsor
BOSTON
Hartford
NEW YORK
Newark
Vernon Valley
Holmdel
Pennsauken
Baltimore
PHILADELPHIA
Tysons Corner
Cleveland
Pittsburgh
CHICAGO
WASHINGTON, D.C.
FAIRFAX
Indianapolis
Columbus
RICHMOND
Raleigh
AMSTERDAM
FRANKFURT
LONDON
PARIS
LONDON
STOCKHOLM
LONDON
FRANKFURT
AMSTERDAM
SINGAPORE
OSAKA
SACRAMENTO
Santa Clara
SALT LAKE CITY
SAN FRANCISCO
Palo Alto
Moffet Field
SAN JOSE
Denver
KANSAS CITY
ST. LOUIS
HONG KONG
Las Vegas
LOS ANGELES
TOKYO
TOKYO
San Diego
PHOENIX
ATLANTA
SYDNEY
DALLAS
Austin
HOUSTON
New Orleans
Jacksonville
Orlando
TAMPA
MIAMI
SAN JUAN
HONOLULU
HAWAII
T1/E1 (1.5 Mbps/2 Mbps)
T3/DS3 (45 Mbps)
OC3c/STM1 (155 Mbps)
OC12c (620 Mbps)
OC48c/STM16 (2.5 Gbps)
OC192c/STM64 (10 Gbps)
Single Hub City
Multiple Hubs City
For more information visit www.uu.net
NB: UUNET also has infrastructure within individual countries, which is not shown on this map.
June 2000
UUNET
A WorldCom Company

are also networks. The Regional ISPs in turn sell their bandwidth to the Local ISPs. Your machine is connected to a **local ISP**, which is responsible for providing you Internet access at some bandwidth. The network connecting your machine to the local ISP is called an **access network**. Access network might be telephone line + modem, or might be cable-based access, etc.. You will discover your local ISP in one of your first homeworks.

Question: But if local ISP A is connected to NBP A and local ISP B is connected to NBP B, then how do local ISP A and local ISP B communicate?

Answer: The different NBPs must have points where they connect to each other. These are called **NAPs**, or network access points.

Example: Connecting your Web server

Suppose you have a web server (really just a file server with lots of files that people want). You want to connect your web server to the Internet, so various clients can retrieve files. You do this by **leasing** an **access link** to the local ISP. The amount of money you pay monthly will depend on the bandwidth of your access link. In my research I've found that for web servers servicing static requests (of the form "get me a file"), the bottleneck resource for the web server is not its CPU or its memory, but rather the bandwidth on the access link. Why is this? For static requests, CPU is not heavily used. Even so, if you don't have enough CPU, you get a second server and connect your servers in a server farm configuration. A server farm often has 32 servers. If you don't have enough memory, you can buy more cheaply. Now CPU load is low and no memory problems. But still have to lease that access link. In fact need big fat access link (lots of bandwidth) to get all the stuff from 32 servers across the link. Costs lots of money. Companies like Akamai which service static requests always complain that the bandwidth on their access link is their performance bottleneck.

6 Delays in networks

Motivating Question 1: Suppose we're sending a single packet on 10Mbps link across the U.S.. At the same time we send a single packet on a 10Gbps link across the U.S.. How much sooner does the packet on the 10Gbps link arrive?

Answer: Both arrive at the same time.

Motivating Question 2: Two graduate students came to me. They were playing with Napster, downloading music files and other stuff over UDP. They noticed that sometimes the same file that they wanted was available at two different sites. When they pinged the sites, they saw that one of the sites was 10ms away and the other was 100ms away. They tried requesting the files from both sites and were surprised to see that the file from the far away site arrived sooner. Why?

Answer: We'll come back to both these questions in a minute.

Initial Setup : Packet must move from Router A to *adjacent* router B. What kind of delays does it incur? (assuming no loss, for now)

Processing Delay – Examining packet header at router A to determine that packet is next going to router B.

Queueing Delay – Waiting at router A in the queue of "packets destined for router B." This queue can be long, depending on how many other flows there are which need to get from router A to router B. This delay depends on the average rate of arrival (bits/sec) of data which needs to go on the link from A to B, but also on the burstiness of these arrivals.

Transmission delay – Time to get whole packet onto wire.

Question: Suppose L bits in packet. Bandwidth is R bits/sec. What is the transmission delay?

Answer: L/R seconds.

Propagation Delay – Time from when bit gets onto wire until it gets to next router.

$$\begin{aligned} d &= \text{distance} \\ s &= \text{speed of light} = 3 \cdot 10^8 meters/sec \\ \frac{d}{s} &= \text{propagation delay} \end{aligned}$$

Total delay in getting packet from router A to *adjacent router* B is often referred to as the **nodal delay**.

$$d_{nodal} = d_{proc} + d_{queue} + d_{trans} + d_{prop}$$

Question: Why are we adding d_{trans} and d_{prop}?

Answer: Time until whole packet moves to router B equals time until get last bit of packet onto wire (d_{trans}) *plus* time from when last bit of packet gets onto wire until it gets to router B (d_{prop}).

Example Question: Suppose you have a 100 bit file and a link with 10ms propagation delay over which you are sending your file. The bandwidth on your link is 1Kbps. Suppose yours is the only file on the network, so there is no queueing delay. Ignoring loss and processing delay, how long does it take to get your file across the link?

Answer: The time until the last bit gets on is

$$100bits/1Kbps = .1sec = 100ms$$

The time from when last bit gets on link until it gets to other side is $10ms$. Thus the total time to transfer the file over the link is:

$$100ms + 10ms = 110ms$$

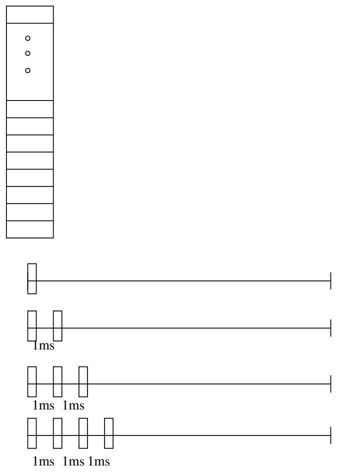


Figure 4: *Illustration of spacing of bits.*

Note proportion of contribution of each of these terms to d_{nodal} can vary wildly.

Back to Motivating Question 1: Suppose we're sending single packet on a 10Mbps link across the U.S.. At same time we send a single packet on a 10Gbps link across the U.S.. How much sooner does the packet on the 10Gbps link arrive?

Answer: Both cases about 100 ms. Bandwidth effects only transmission time, but transmission time is negligible in comparison to propagation time in this example.

Back to Motivating Question 2: Two graduate students came to me. They were playing with Napster, downloading music files and other stuff over UDP. They noticed that sometimes the same file that they wanted was available at two different sites. When they pinged the sites, they saw that one of the sites was 10ms away and the other was 100ms away. They tried requesting the files from both sites and were surprised to see that the file from the far away site arrived sooner. Why?

Answer: If the file is large, transmission delay, not propagation delay, is the dominating factor. For example it is perfectly reasonable that you might be trying to download a 10 MB image and the available bandwidth might be 10 Mbps. In this case:

$$d_{trans} \approx \frac{10 \cdot 10^6 \cdot 8bits}{10 \cdot 10^6 bps} = 8seconds$$

Given that d_{prop} is under 100 ms in both scenarios, it's clear that d_{prop} is irrelevant.

Question: What is the end-to-end delay from source host to destination host, given $Q-1$ routers in between?

Answer: Depends on whether packet can span several routers at once, so can get extreme pipelining advantage. Almost all networks use **store-and-forward**: router must receive full packet before it can begin to transfer first bit of packet to next router. Thus,

$$\text{End-to-end delay} = Q \cdot d_{nodal}$$

7 Throughput versus Bandwidth

In talking about network performance, the terms network **throughput** and network **bandwidth** are sometimes used interchangeably. But these are **not the same**.

Throughput refers to bytes per second of data that Application A on machine 1 received from Application B on machine 2. Bandwidth refers to the bytes per second that the link connecting machine 1 and machine 2 can possibly transmit.

Question: There are many reasons why throughput $\neq$ bandwidth. Can you think of any? (To make things simple, let's stick to a single link).

Answer 1: Loss. Some of the packets may get lost and need to be retransmitted. This means that the actual throughput is less than the available bandwidth.

Question: How does loss arise?

Answer: Queue at router can get too long. Buffer has finite capacity, so lose packets.

Answer 2: Headers. Some of the bytes transmitted are header info not data. Again this means that throughput $>$ bandwidth.

Question: Suppose I assure you that there is zero loss and there are no headers. How might throughput be WAY less than bandwidth?

Answer 3: Imagine a situation where $t_{prop} >> t_{trans}$. For example suppose our application sends a 1MB file across a 1Gbps network with a RTT (round-trip-time) of 100 ms.

$$\text{Time to send file } = 100ms + \frac{1 * 10^6 * 8bits}{10^9 bps} = 108ms$$

Thus throughput is only $\frac{1MB}{108ms} = 74Mbps$. On the other hand, available bandwidth is $1Gbps$. Such a situation where *throughput* << *bandwidth* occurs quite often in practice when one application needs confirmation from another before it can send more data.

8 Delay × Bandwidth Product

Question: What do you think the term **delay-bandwidth product** refers to?

Answer: Total number of bits on a link simultaneously.

For example, in the Figure 4 above, the delay-bandwidth product is $1Kbps \cdot (10 \cdot 10^{-3} sec) = 10bits$. Thus the link in Figure 4 can hold 10 bits simultaneously.

The reason that delay-bandwidth product is important is that twice the delay-bandwidth product represents the total amount of data a sender might send before hearing an acknowledgement from the receiver. This comes up a lot in Ethernet examples, where several machines are trying to send data. It is important to calculate how long it takes to find out that the signal was corrupted (because of simultaneous transmissions) and how much data will need to be retransmitted. We'll come back to this at length when discussing Ethernet.

9 Class Administration

Everything you need to know if on this web page:

http://www.cs.cmu.edu/~srini/15-441/F02/ Check this page regularly for announcements!