Communication Protocols
and Internet Architectures

Harvard University

Lecture #2

Instructor: Len Evenchik
cs40@evenchik.com or evenchik@fas.harvard.edu

ALIGHLSOD1701

Lecture Agenda

- Some Quick Logistics
- Course Goals and Objectives
- Review and Q&A
- Services versus Protocols
- Tools and Techniques for Protocol Specification
- Protocol Design Issues
- SP3: A Framework for Understanding Protocols
- What is the Internet
- One Minute Wrap-Up

Course Logistics

Course Logistics

- Information on section meetings and office hours is on the website.
- Course work load -- minimum 12 hours/week. It is very, very important that you do all of the homework assignments.
- In addition to the lectures, the course website contains announcements, course handouts, and a discussion group. The course handouts include: a weekly course information sheet, a reading assignment for each lecture, and handouts with the lecture slides
- There will be an online midterm exam and in-class, closed book final exam.
- If you live in New England, you must take the final exam at Harvard on the date it is scheduled. Distance students will need to have their exam proctored.
- **Please submit a one minute wrap-up after each lecture. Thank you!**

One Minute Wrap-Up

- Please complete this Wrap-Up at the end of each lecture.
- Please use the form on the course website.
- The form on the website is anonymous (but you can include your name if you wish.)
- Please answer three questions:
  - What is your grand "Aha" for today' s class?
  - What concept did you find most confusing in today's class?
  - What questions should I address next time?

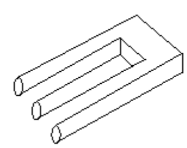
Course Goals

Course Goals

- This course will provide:
  - * Tools and a framework for understanding the constantly changing networked environment
  - * In-depth technical understanding of current network architectures and protocols
  - * Understanding of the design trade-offs that have to be made when building large networks
  - * Appreciation of the evolution of communication networks and services: from POTS to IoT and the Cloud
- What this course is not about:
  - * Programming specific network protocols
  - * Installing specific communication products

What is Wrong with the Following Statements?

- "I was at a seminar and they explained that VLANs and switches are are how networks are built today so we don't need to worry about IPv6."
- "Doing a network design will take too long, we'll just use the Internet."
- "I just heard that we ran out of IP addresses and we all know that this means that the web will stop working sometime next spring."
- "Easy, the video will look great if you just use gigabit ethernet and 4G LTE."
- "Of course using QoS (or the TLA of the moment) solves the problem."
- "Don't worry, we have an extremely reliable network since it is built using nothing but fiber optics."

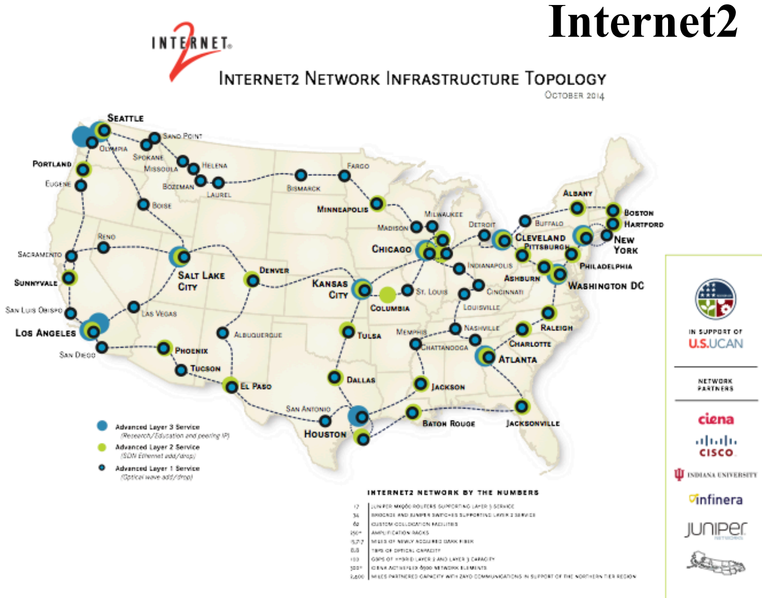


Q&A and Some Things from Previous Lectures

What is a Network and how are networks different ?

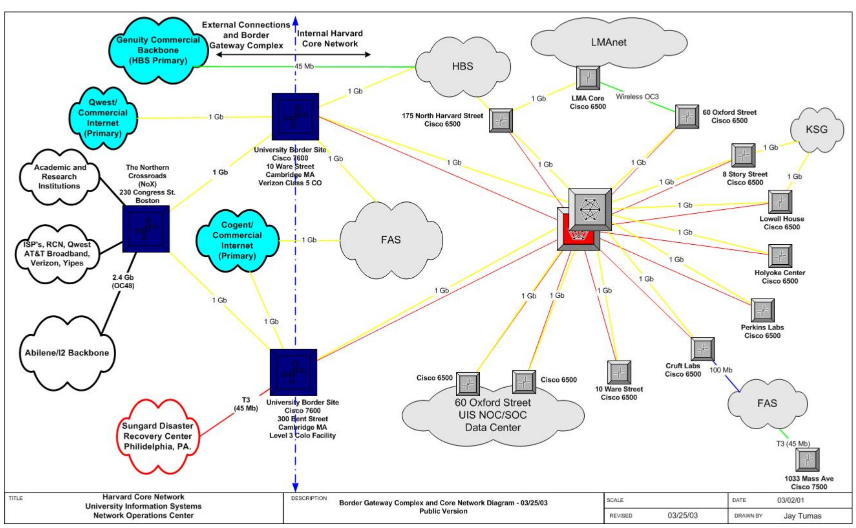
ALIGHLSOD1701

Internet2



Source and copyright is http://www.internet2.edu 2015

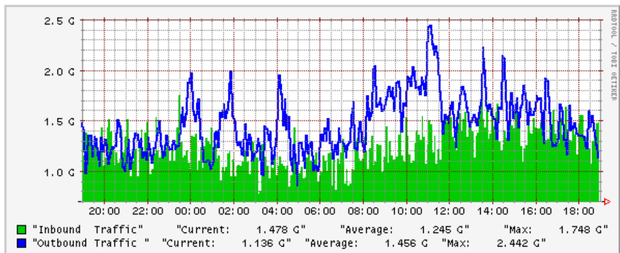
Harvard Core Network
Source and copyright is Harvard HUIT



Traffic as seen from KSCYng

The green area is a five minute average of input bits per second.

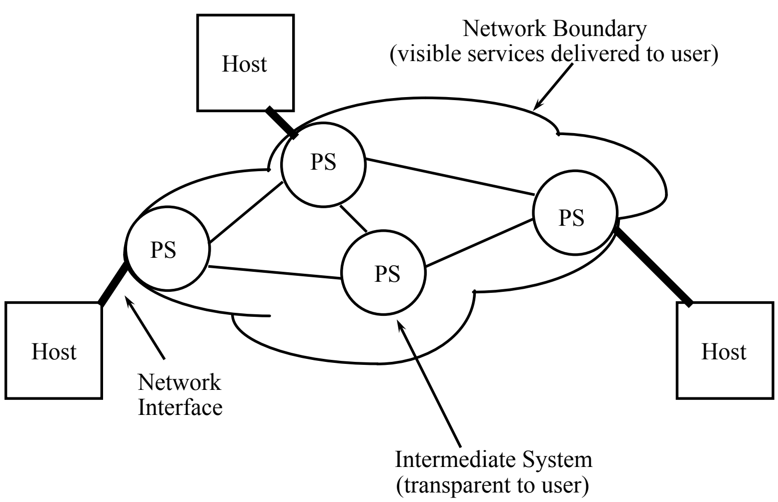
The blue line is a five minute average of output bits per second.



See full historical data

Source and copyright is http://abilene.internet2.edu/
Indiana University Abilene NOC Weathermap

The Network "Cloud"



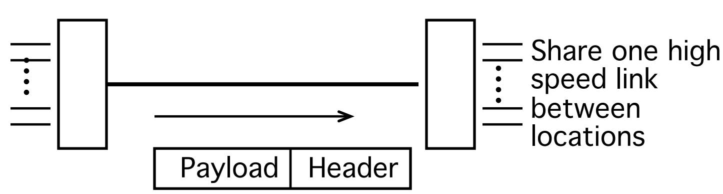
Source: A Protocol for Packet Network Interconnection,
published in 1974, IEEE Transactions on Communication

Multiplexing

- The division of a single physical channel into two or more logical channels. (Muxing is the basis of packet switching.)
- Users share a single physical channel and this provides for equipment and cost savings. The question is how to share it.
- Frequency Division Multiplexing is one option (radio, satellite)
- Time Division Multiplexing is another option
  - Slotted TDM - each user is assigned a particular slot (time interval) whether it is needed or not - this provides guaranteed delay and bandwidth. Also called fixed TDM.
  - Statistical TDM - each user requests access or enters a queue and waits for access to the shared medium. What about the queuing delay and resulting quality of service (QoS) in this case?
- Many other combinations possible

What About the 25th User?

- Now, imagine that you need to connect a 25th user between Boston and NYC.
- A couple of options exist:
  - Increase bandwidth of the backbone link
  - Decrease each users bandwidth
  - Refuse to service the 25th user
  - Use statistics and assign bandwidth as needed to individual users. The question is how can this be done?



Switching

Switching

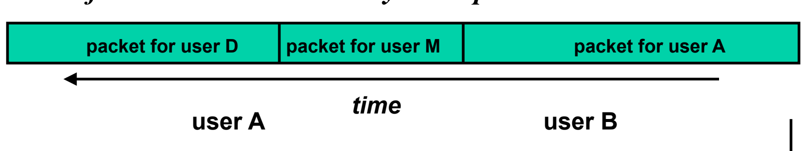
- Circuit Switching
- Message Switching (Telegraph, 1800s)
- Packet Switching
- Cell Switching (Used fixed size packets: ATM is the common example from the 1990s. Lots of development and dollars, but little deployment. Skip references in any readings.)
- Optical Switching
- .. and other options

When is switching used, and when is multiplexing used?

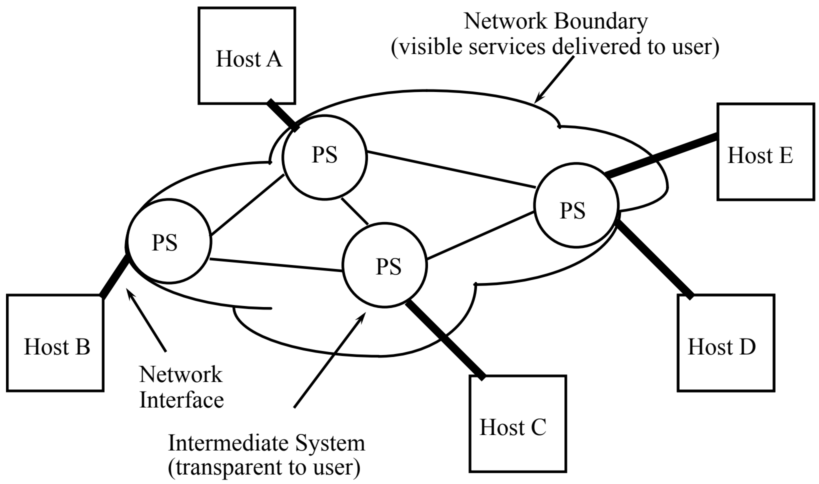
Packet Switching

- Packet Switching evolved to support computer communications
  - Separate data into packets
  - Move a packet of data from one computer to another
  - Allows for the resource sharing of the links and switches
  - Packet contains a header and a payload, and a minimum and maximum packet size (along with many other details) is set by the protocol
  - Emphasis is on efficiently moving blocks/packets of data with no errors

- ***Individual packets are directed out a particular port on the switch based on the address in the packet, and then the outbound packet utilize the entire capacity of the link, but for a very, very, very short period of time. This is not an easy concept to understand.***



The Network "Cloud"



Source: A Protocol for Packet Network Interconnection, published in 1974, IEEE Transactions on Communication

Protocol Models and Layering

Why Layered Network Architectures?

- To help manage complexity
- To facilitate the building block approach to implementation & management

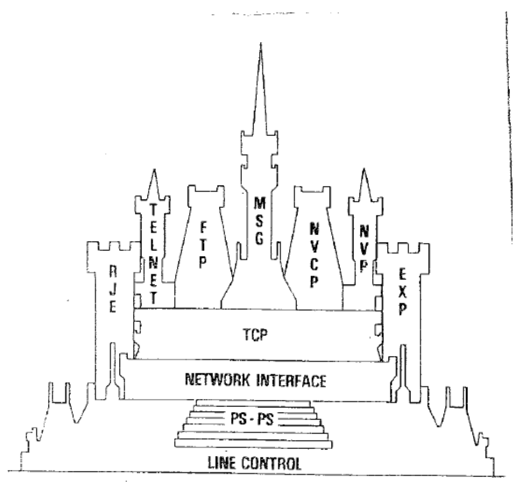
How Many Layers are Enough?

Depends on who's counting.

- OSI / ISO model “said” 7 about 30 years ago and this is still talked about today
- Internet model uses 5 and most systems are built using this approach. **This is the model we will use for this course.**

However, counting the number of layers is more art than science when you build systems in the real world.

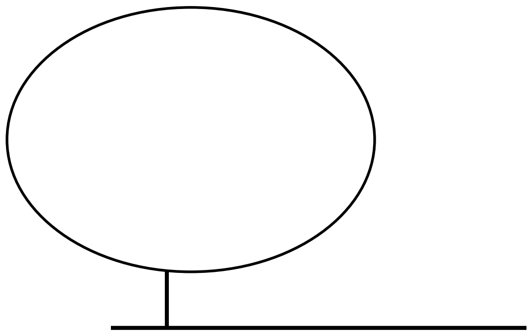
Protocol Model, circa 1974



Source: IEEE Tutorial on Communication Protocols

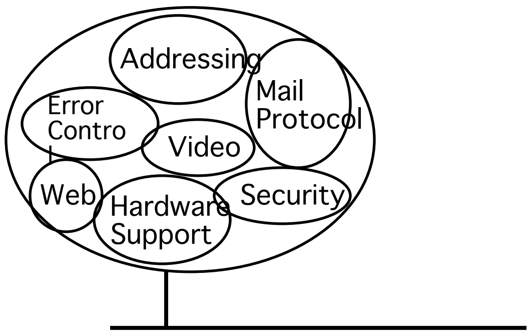
Communication Functionality versus the Model

You can divide up the functionality required in a communication system in many different ways.



Communication Functionality

You can divide up the functionality required in a communication system in many different ways.

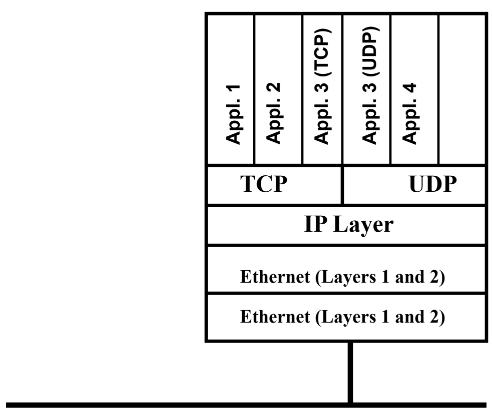


Protocol Encapsulation

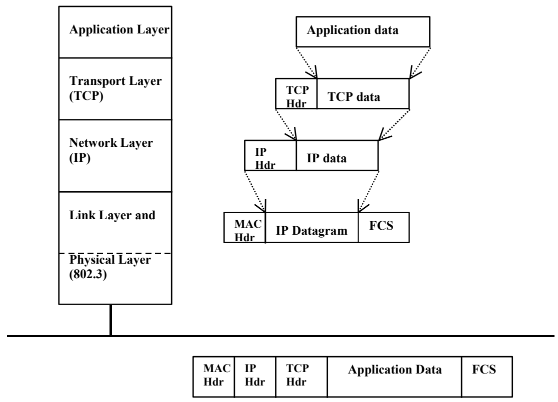
Encapsulation - 5 Layer Model

- Application layer
- Transport layer
- Internet layer (sometimes called the Network layer)
- Link layer (sometimes called Network Interface layer, or Data Link layer)
- Physical layer

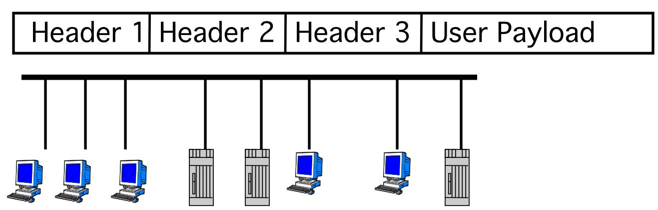
Five Layer Simplified Software Schematic



Protocol Encapsulation

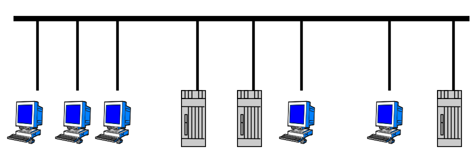


Encapsulation Results in Multiple Headers



What Does a Packet Look Like

What Does a Frame Look Like on an Ethernet?

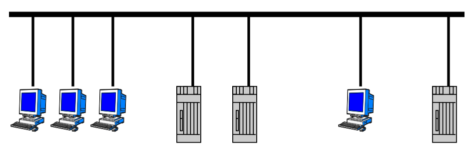


Ethernet Frame

(Sometimes incorrectly called a Packet)

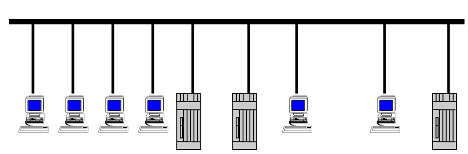
Frame starts here … . . . . and continues off the page….

```
00 00 a7 11 57 dd 08 00 09 38 24 31 08 00 45 00 00 38 b3 ….
```



Ethernet Frame

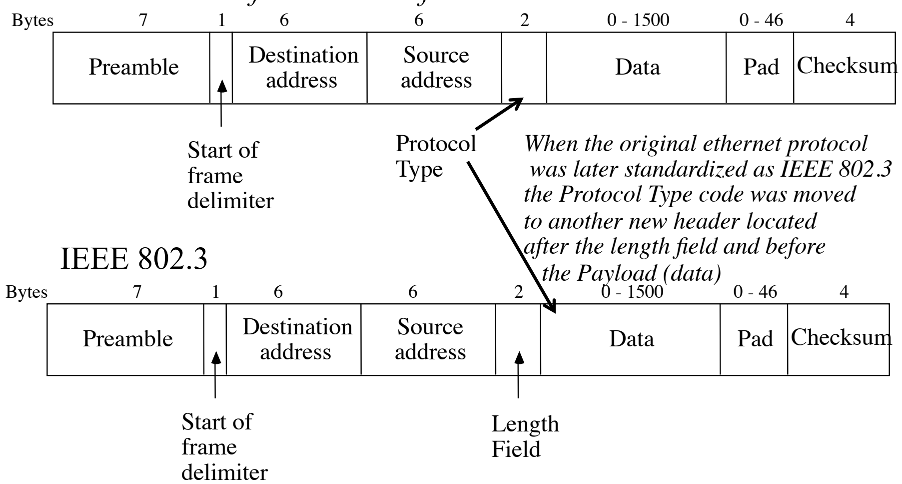
```
Byte Count                Content of Frame
00 to 15: 00 00 a7 11 57 dd 08 00 - 09 38 24 31 08 00 45 00
16 to 31: 00 38 b3 e3 00 00 1e 06 - 8b 3e c6 04 69 c9 c6 04
32 to 47: 69 cf 05 bd 17 70 3a e4 - 44 b9 2c 0a 9e ad 50 18
48 to 63: 20 00 88 53 00 00 3d 11 - 11 04 11 80 11 02 34 a3
64 to 69: 00 76 34 0c 00 13
```



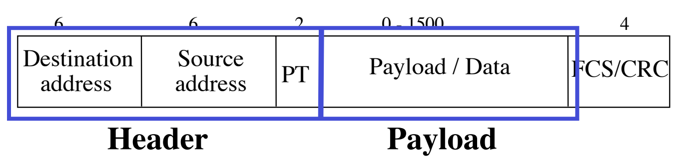
Ethernet and IEEE 802.3 Frame Formats

(There are two compatible formats used today)

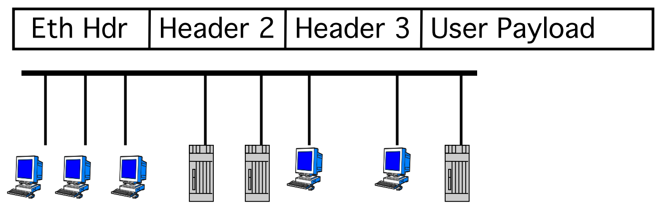
Ethernet – *We will focus on this format.*



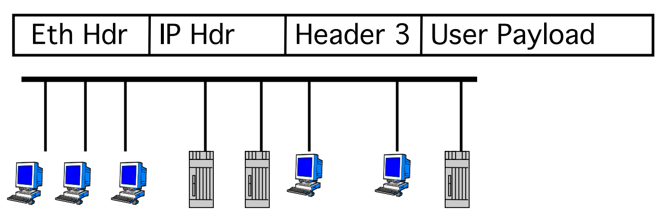
Typical Representation of the Ethernet Frame Format and the One We Will Use in Class



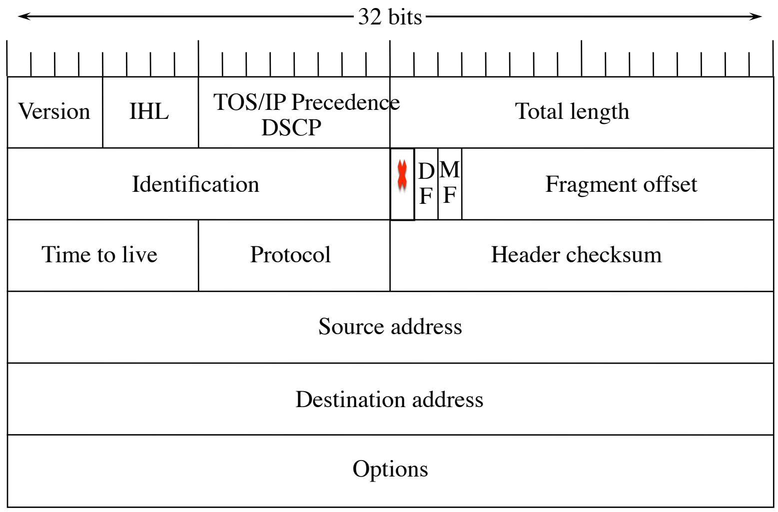
Encapsulation Results in Multiple Headers



Encapsulation – Ethernet and IP Headers

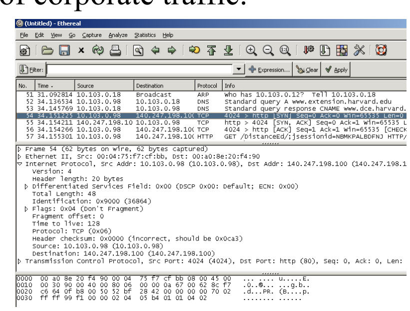


IPv4 Packet Header Format

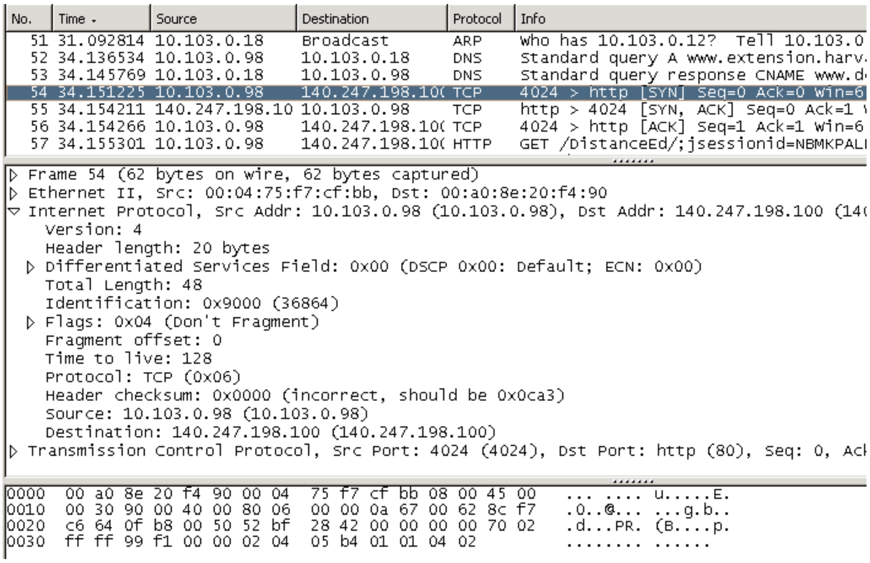


Network Packet Trace

- Network analysis tool is called Wireshark and it is available from www.wireshark.com
- Wiretapping or monitoring of any Harvard network traffic is strictly forbidden under all circumstances. See the Harvard Computer Policy & Responsibilities website for the details.
- Almost all organizations also have policies against the monitoring of corporate traffic.



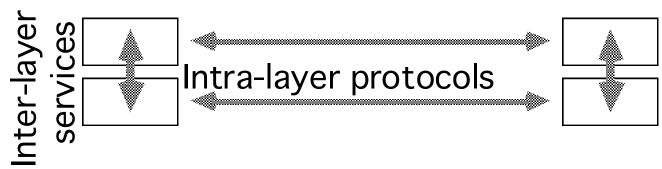
Network Packet Trace



Services versus Protocols

"Services" vs "Protocols"

- Entities in each layer offer ***services*** to entities in the layer above
- Same-layer entities implement these services by conversing across the net according to ***protocols***.
- Both services and protocols must be documented.
- The reason to do this is because it allows protocols to change while services stay the same.



Types of Service

Imagine that you wanted to transfer data between two systems. What types of services might you offer?

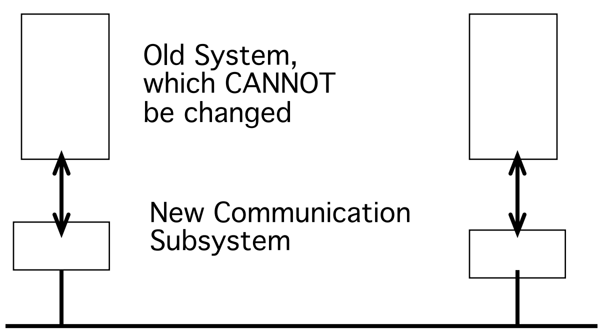
Types of Service (2)

Imagine that you wanted to transfer data between two systems. The services you could offer include:

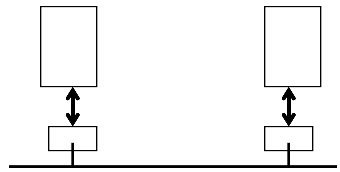
- Reliable and sequenced delivery. This is known as connection-oriented service (HDLC, TCP)
- Unreliable, not sequenced. This is known as connectionless or datagram service. (802.3, IP)
- Unreliable, but sequenced. (RTP)
- Are there others?

Why Distinguish Services from Protocols?

- Imagine you are asked to add a reliable communication service between two old system (which cannot be changed for some arcane reason.)

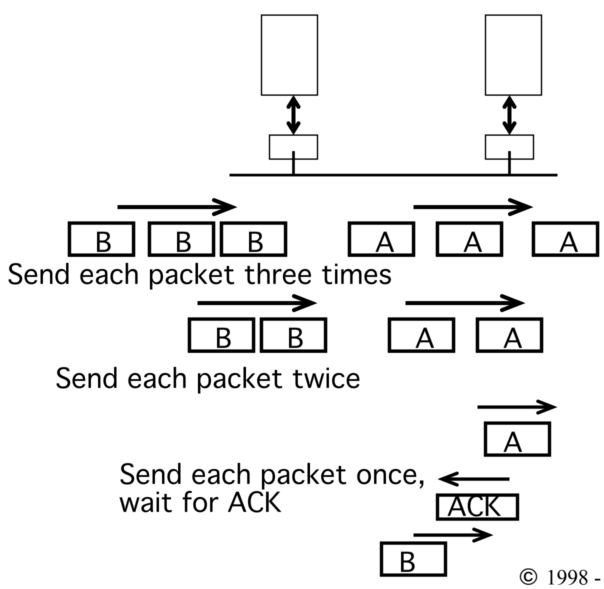


How could you implement a Reliable service between these two systems



© 1998 - 2017 L. Evenchik

Some Possible Solutions...

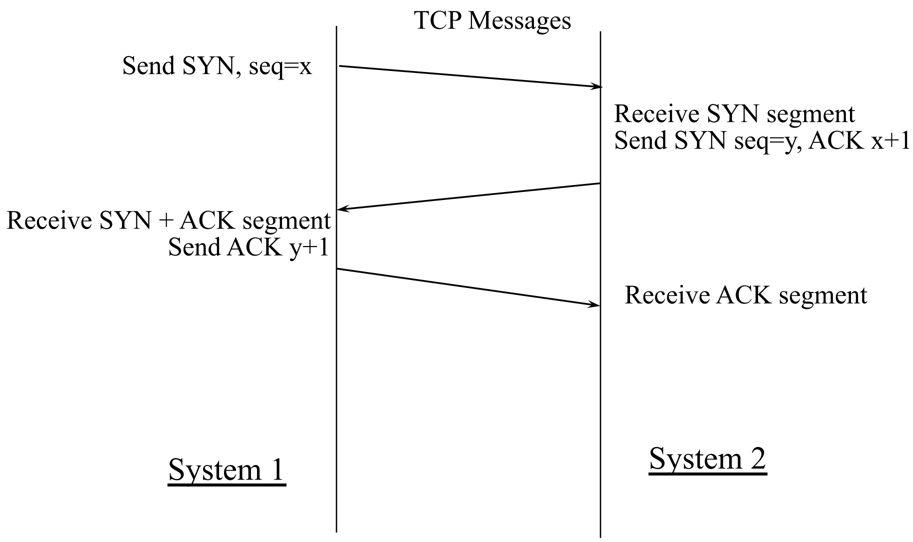


Common Tools and Techniques for Protocol Specification

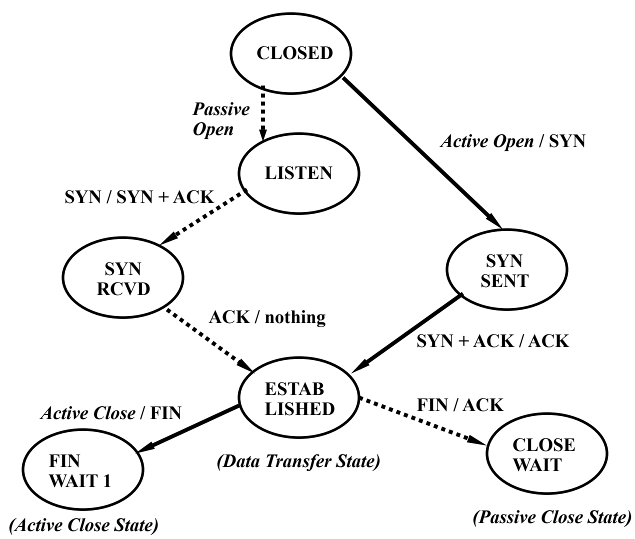
Protocol Specification Tools and Descriptive Tools

- Time Sequence Diagram, also called a Ladder Diagram
- State Transition Diagram (or sometimes called State Diagram)
- Packet traces, Wireshark
- Plus many other formal and ad-hoc tools and methods

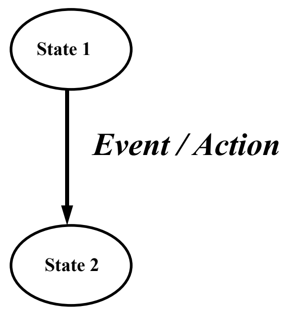
Time Sequence Diagram, aka Ladder Diagram



TCP State Transition Diagram, Client/Server - abridged

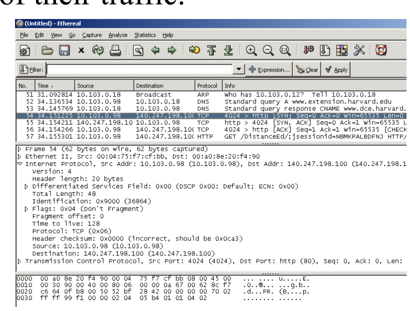


State Transition Diagram



Network Packet Trace

- Network analysis tool is called Wireshark and it is available from www.wireshark.com
- Wiretapping or monitoring of any Harvard network traffic is strictly forbidden under all circumstances. See the Harvard Computer Policy & Responsibilities website for the details.
- Almost all organizations also have policies against the monitoring of their traffic.



Protocol Design 101

Classic Network Topologies

- Point to Point (pt-to-pt)
- Star
- Bus
- Ring
- Mesh

Notes:

- This list describes the topology, not the type of media or the protocol that is used to communicate
- A separate question is whether the topology and media support broadcast (or multicast)

Protocol Design Issues for a Protocol Used in a Very Simple Pt-2-Pt Network

- Does the protocol need addressing?
- Should the protocol be reliable?
  - What does reliable mean?
- What else should be considered

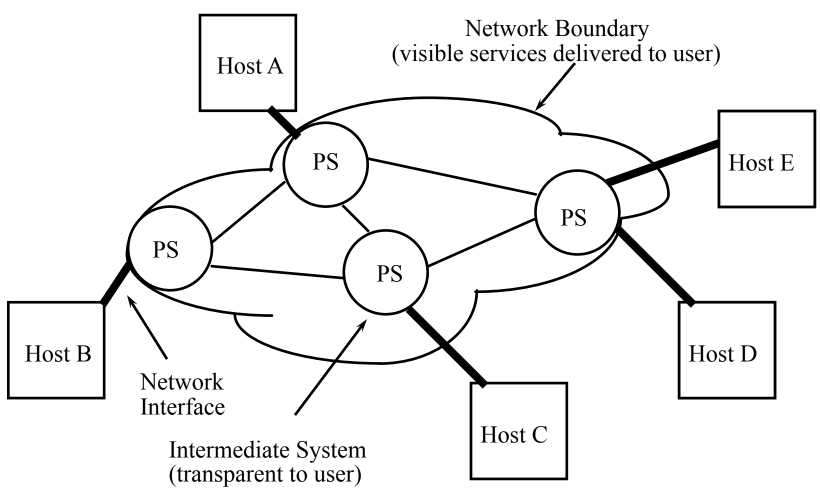
Protocol Design Issues for a Protocol Used in a Network with a Bus Topolgy

- Does the protocol need addressing?
- Can each station hear every other station? How should the protocol support this?
- Should the protocol be reliable?
- What else should be considered

Protocol Design Issues for a Protocol Used in a Mesh Network

- What is the difference between the host and the switch?
- Does the protocol need addressing?
- Can each station hear every other station?
- Should the protocol be reliable?
- What else should be considered

The Network "Cloud"

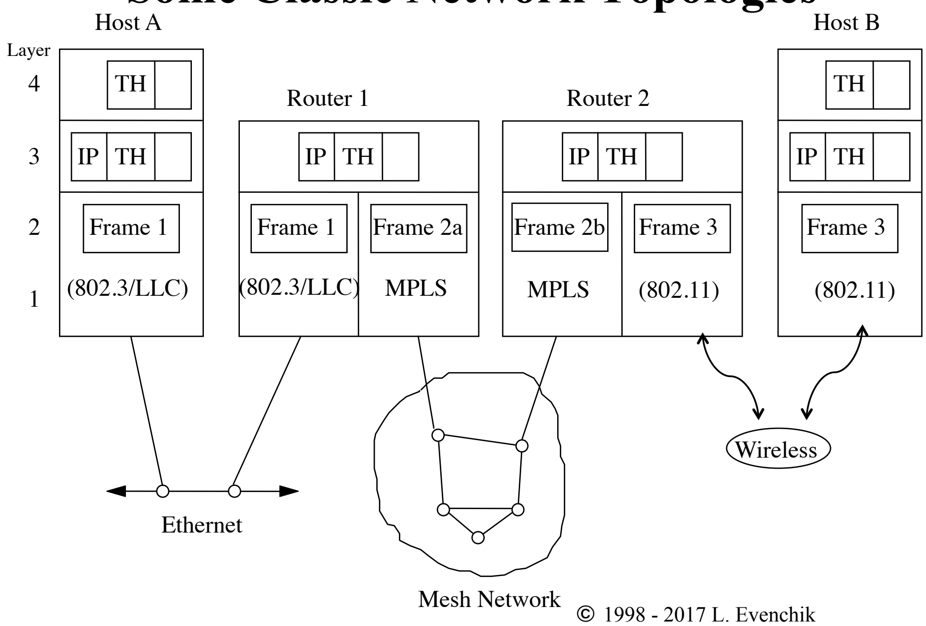


Source: A Protocol for Packet Network Interconnection, published in 1974, IEEE Transactions on Communication

SP3: Service, Purpose, Packet and Procedure

SP3 is a Framework for Understanding and Describing Protocols

Some Classic Network Topologies



SP3 Protocol Framework

- Service
  - The Service is a description of what the protocol does, not how it is done. This should be a few sentences long.
- Purpose
  - The Purpose describes the specific functionality that the protocol provides and how it is accomplished. Examples are flow control, error detection, error correction, etc.
- Packets
  - The Packet layout determines how the various bits and fields within the packet are defined, assembled and used.
- Procedures
  - The Procedures describe the various packet exchanges and the reason for each exchange.

SP3 - Service

What type of Service does the protocol provide?

The Service is a short description of what the protocol does, not how it is done. For example, a link layer protocol could provide any of the following services

- Reliable service, including sequenced delivery. This is commonly known as connection-oriented service.
- Reliable service, but not with sequenced delivery.
- Unreliable service. This is known as connectionless or datagram service. (IP, UDP)
- Unreliable service, but with the sequenced delivery of messages. (RTP is an example of this.)
- Are there more?

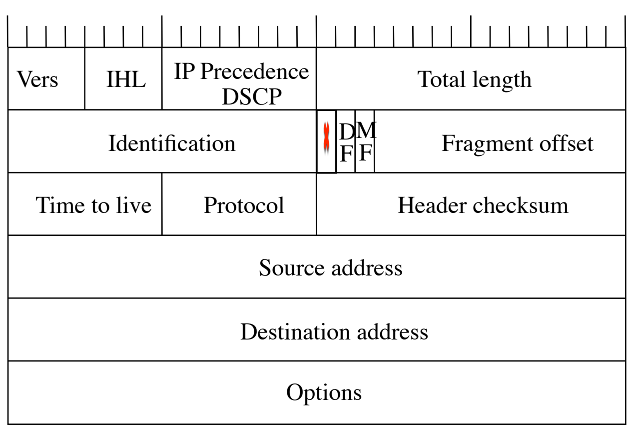
SP3 – Purpose

What specific functionality does the protocol provide and how does it do it?

- Addressing
- Multiplexing
- Sequencing
- Error control - two parts to this, detection and correction
- Flow control
- Option negotiation
- Encryption
- Fragmentation and reassembly
- *plus many others...*

SP3 Packet

What do the fields in this IPv4 packet header do?



SP3 – Procedures

What Procedures does a protocol use to implement the following functionality:

- Connection Establishment or Initialization
- Data Transfer
- Connection Release or Disconnect
- Error Handling (of many different types)

Applying SP3 to Understand Link Layer Protocols

Our Example for a Link Layer protocol is based on a simplified version of HDLC

HDLC (High-Level Data Link Control) is a simple but reliable layer 2 protocol. It was an ISO standard developed over 40 years ago, and it has provided the framework for protocols that are still used today.

We are using parts of this old protocol as a simple example of a reliable link layer protocol. You are not responsible for knowing HDLC, but you need to understand how to design a reliable layer 2 protocol.

SP3 – Service
Definition of a Reliable Data Link Layer Service

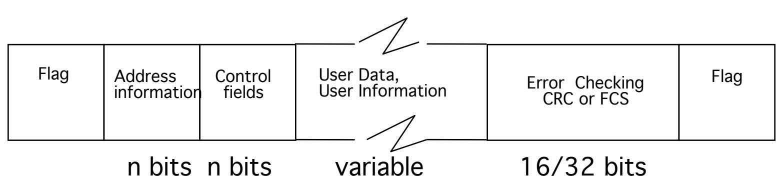
Our example protocol provides a reliable data link layer (Layer 2) service to Layer 3. In other words, it gives the physical layer the appearance, as seen by layer 3, of being an error-free link, even though it is not.

SP3 - Purpose

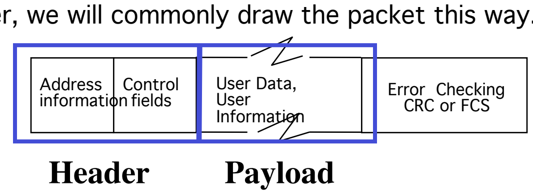
To deliver the promised level of service, a reliable protocol (such as HDLC or TCP) must handle the following problems:

- Synchronization and framing – where does the packet start and end?
- Data transparency
- Data transfer
- Addressing
- Flow control
- Error detection
- Error correction
- plus others

Example Link Layer Packet Format



Remember, we will commonly draw the packet this way.



Note that the bundle of bits at layer 2 is properly called a frame, but it is common today to call it a packet.

Procedures for a Reliable Protocol

A reliable protocol (such as HDLC or TCP) implements the following procedures:

- Connection Establishment or Initialization
- Data Transfer
- Connection Release or Disconnect
- Error Handling (of many different types)

Synchronization, Framing and Data Transparency

- Synchronization and Framing: describes the order of bit transmission, determines when the packet starts and ends versus the line being idle. information vs. noise
- Transparency: the protocol must prevent random bit patterns in the data sent by the user from being mistaken for protocol control characters or control packets.

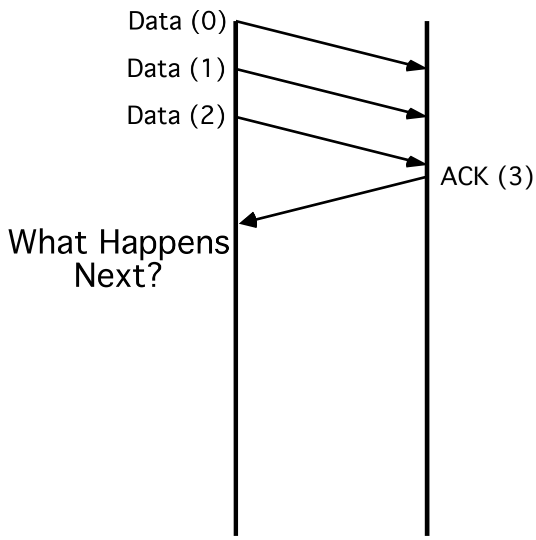
Addressing

- Should addresses have local or global significance?
- How should addresses be structured?
- How many addresses are needed in a packet or frame given that they can have local or global significance?
- How are addresses assigned?
- Can addresses be private?
- What is the difference between an address, a name, and a route?

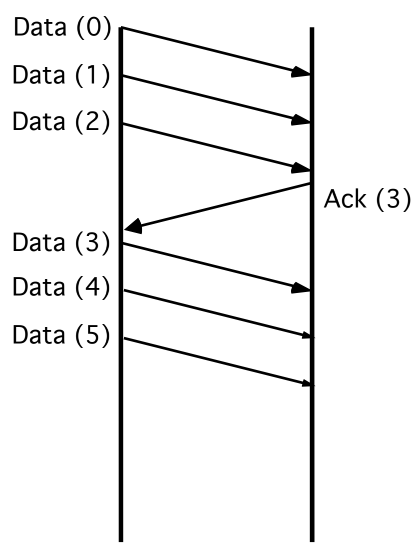
Flow Control

- Need to keep the sender from overwhelming the receiver with more data than it can handle at the moment.
- A variety of simple and complex methods to choose from, for example:
  - Stop/Go: Xon/Xoff, RTS/CTS
  - Prior buffer registration mechanisms
  - Sliding Window or Credits (used in TCP)
- Functionally, the receiver can stop the sender from sending more data for a specific reason (such as buffer space) or for no specified reason other than "not now, please".

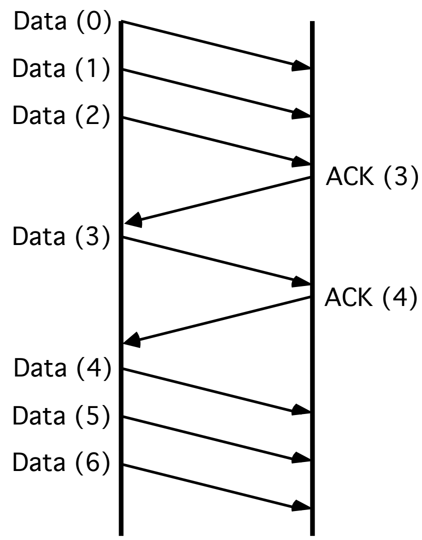
The Sliding Window - Flow Control



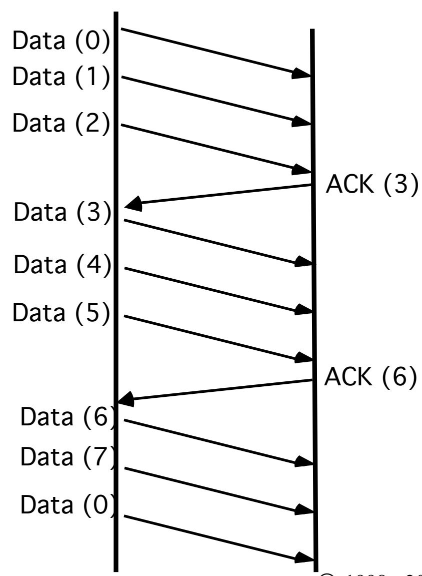
The Sliding Window - Flow Control



The Sliding Window - Flow Control



The Sliding Window - Wrapping



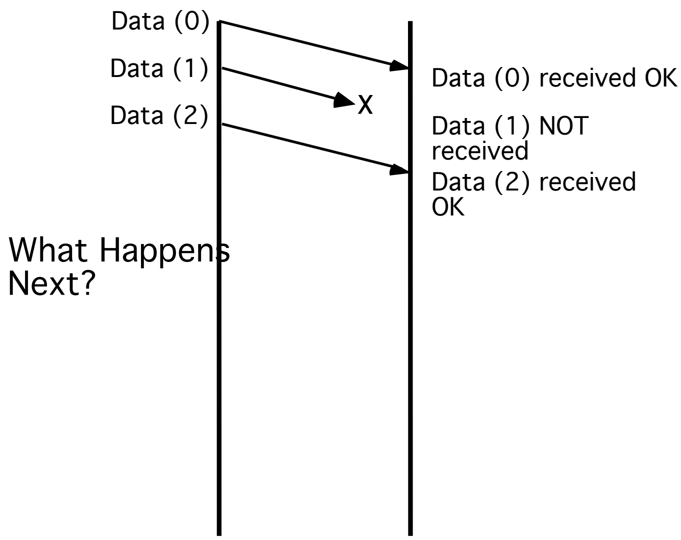
Error Control

- Error control is needed for both data transmission errors and procedural errors
- Error control consists of error detection and in some cases, error correction. When is correction not needed?
- Extra information is transmitted with the data so that the receiver can detect the existence of errors - examples are parity, checksum or CRC. If required, errors are then corrected by retransmission.
- Where retransmission would be difficult, enough extra information can be sent to allow the receiver to detect and regenerate damaged data. This is known as Forward Error Correction or FEC. We will not study FEC in class.
- Procedural errors are different that data transmission errors, and they include protocol events which don't make sense or control packets that are sent at the wrong time.

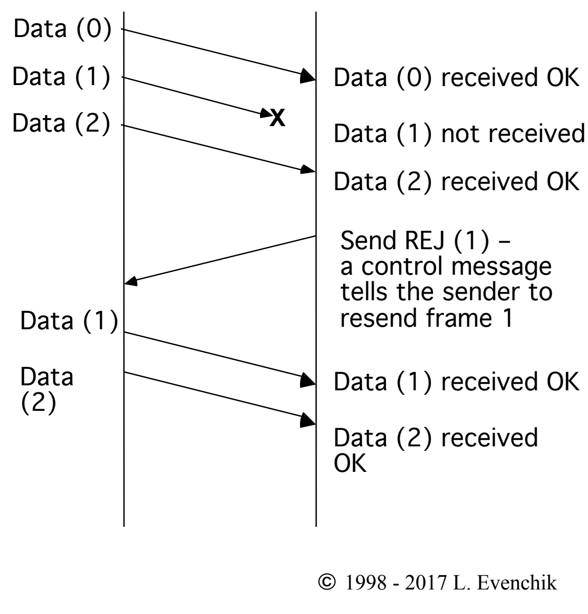
Data Error Recovery Mechanisms

- The simplest form is to wait for a Positive Response from the receiver on every frame that is sent; this means that there is a single frame in flight at any point in time.
- Using the functionality provided by sliding windows and sequence numbers, multiple frames can be sent at one time, up to some "limited number"
  - The limited number can be administrative defined, or determined on the fly.
  - "go back *n* " -- retransmit rejected frame (which was dropped for some reason) and all the frames that followed it
  - use explicit retransmission of only the frame which was rejected
- Plus others…

The Sliding Window - Error Control



The Sliding Window - Error Control

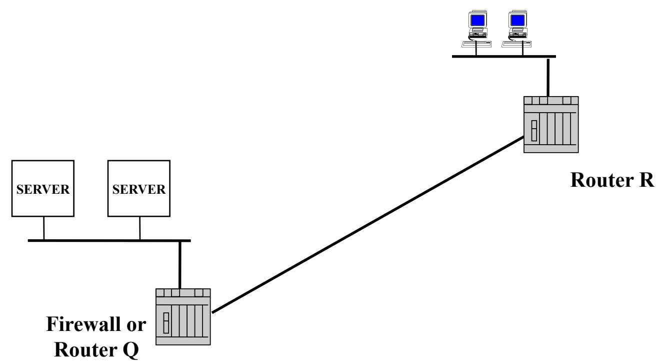


Error Control

- Transmission errors:
  - Sequence numbers are used to allow detection of lost frames (a noise burst may damage or destroy a frame)
  - Numbered frames must be acknowledged within some timed interval.
  - Sliding window is needed to prevent wraparound of the sequence number before acknowledgment
- Procedural errors:
  - Commands or responses which can't be interpreted or are sent at the wrong time are rejected by the receiver. An error message of some type is sent back to say there was an error, force a restart or reset the connection.
  - The loss of sequence numbering will also cause the connection to be reset, or close the connection

The SP3 Framework Applied to PPP

Network with pt-to-pt Line



PPP: Point-to-Point Protocol

- PPP provides for encapsulation of IP datagrams (or other network layer protocols) for use on a serial link, typically a WAN link
- The Link Control Protocol (LCP) is used for establishing, configuring and testing the data link connection.
- There are a family of Network Control Protocols (NCP) for establishing and configuring the network layer protocols.
- The protocol is defined in RFC 1661 (plus others).
- PPP frames look like HDLC frames.
- So the questions include: How do you handle data transparency? What is the PPP frame overhead? What is PPPOE?

PPP Frame Formats
(source is RFC 1661)

PPP Encapsulation (section 2)

```
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Protocol      |   Information       |     Padding     |
|  8 / 16 bits   |  (variable length)  |                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

LCP Packet Formats (section 5)

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Code          | Identifier    |     Length                    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Data ...
+-+-+-+-+-+-+-+-+
```


PPP Link Procedures/Operation

- Link Dead Phase
- Link Establishment
- Authentication Phase (not mandatory)
- Network Layer Protocol Phase
- Link Termination Phase

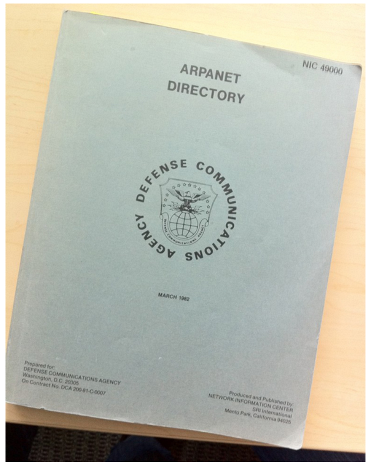
PPP Phase Diagram
(source is RFC 1661)

```
3.2.  Phase Diagram

   In the process of configuring, maintaining and terminating the
   point-to-point link, the PPP link goes through several distinct
   phases which are specified in the following simplified state diagram:

   +------+        +-----------+           +--------------+
   |      | UP     |           | OPENED    |              | SUCCESS/NONE
   | Dead |------->| Establish |---------->| Authenticate |--+
   |      |        |           |           |              |  |
   +------+        +-----------+           +--------------+  |
      ^               |                        |             |
      |          FAIL |                   FAIL |             |
      +<--------------+             +----------+             |
      |                             |                        |
      |            +-----------+    |           +---------+  |
      |       DOWN |           |    |   CLOSING |         |  |
      +------------| Terminate |<---+<----------| Network |<-+
                   |           |                |         |
                   +-----------+                +---------+

   Not all transitions are specified in this diagram.  The following
   semantics MUST be followed.
```


What is the Internet?
An Extremely and Overly Simple Approach to Understanding its Complexity

Internet History
The Internet Directory: March 1982

(400 pages, 20 names per page)



What is the Internet

- A network of networks with a couple billion users
- With large national and international ISPs (Internet Service Providers) as the core networks
- With regional ISPs connected to national ISPs
- With local ISPs connected to national/regional ISPs
- All ISPs exchanging inbound and outbound traffic with other ISPs across public and/or private peering points.
- With billions of users connected to customer networks at the outer edge, that then connect to one or more of the ISPs.
- All using the TCP/IP suite of protocols

Changes in Internet Topology

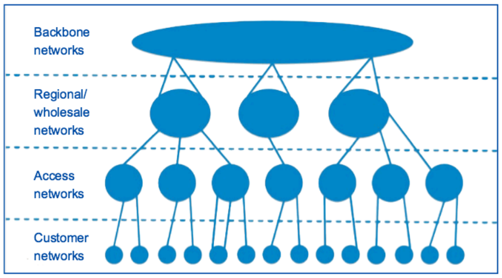


Figure 2 Traditional Internet hierarchical model

Straightforward Hierarchical Topology Early '90s

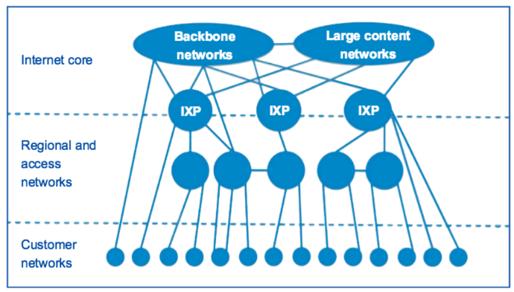
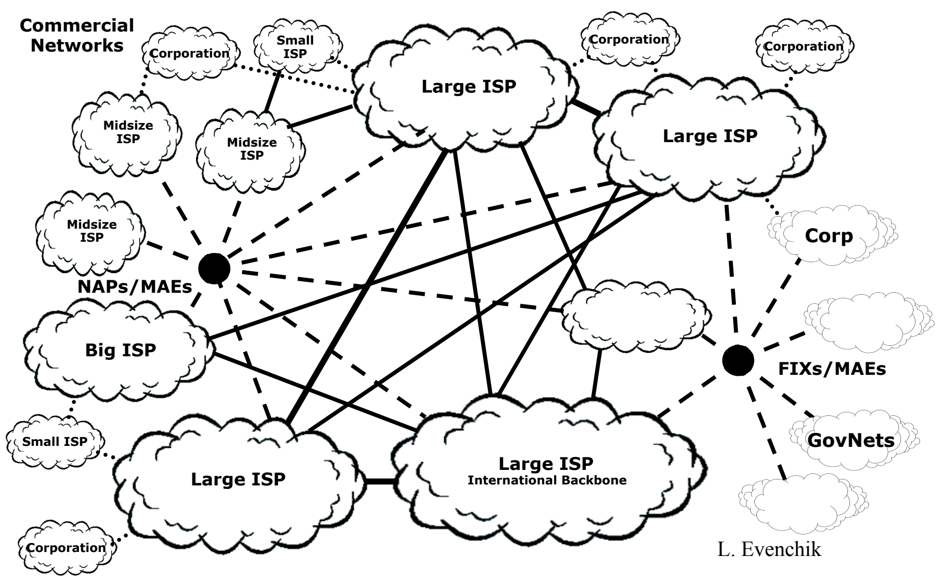


Figure 3 Flatter, more connected model

Mesh or Highly Connected Topology Today, Both ISPs and Others Provide Connectivity as well as Content

Source: ISOC Reports

The Current Internet (Very, Very, Very, Very... Simplified Topology)



Internet Trace from Harvard to MIT

```
fas% traceroute www.mit.edu
traceroute to DANDELION-PATCH.MIT.edu (18.181.0.31),40 byte packets
 1  scmr-gw.fas.harvard.edu (140.247.30.1)  1 ms  1 ms  1 ms
 2  sc-gw.fas.harvard.edu (140.247.6.2)  1 ms  1 ms  0 ms
 3  camgw1-fas.harvard.edu (140.247.20.1)  1 ms  2 ms  1 ms
 4  192.5.66.18 (192.5.66.18)  2 ms  1 ms  1 ms
 5  192.5.66.50 (192.5.66.50)  1 ms  1 ms  1 ms
 6  192.5.66.41 (192.5.66.41)  1 ms  2 ms  1 ms
 7  192.5.66.34 (192.5.66.34)  1 ms  2 ms  1 ms
 8  MIT-MEDIAONE.MIT.EDU (18.95.0.1)  30 ms  2 ms  2 ms
 9  W20-RTR-FDDI.MIT.EDU (18.168.0.8)  3 ms  3 ms  3 ms
10  DANDELION-PATCH.MIT.EDU (18.181.0.31)  2 ms *  4 ms
fas%
```


Internet Trace to Oxford University

```
fas% traceroute www.oxford.edu
traceroute to www.OXFORD.edu (163.1.0.45), 30 hops max, 40 byte packets
 1  scmr-gw.fas.harvard.edu (140.247.30.1)  1 ms  1 ms  1 ms
 2  sc-gw.fas.harvard.edu (140.247.6.2)  1 ms  1 ms  0 ms
 3  camgw1-fas.harvard.edu (140.247.20.1)  0 ms  0 ms  1 ms
 4  192.5.66.18 (192.5.66.18)  2 ms  1 ms  1 ms
 5  192.5.66.9 (192.5.66.9)  2 ms  2 ms  2 ms
 6  12.127.80.125 (12.127.80.125)  3 ms  3 ms  3 ms
 7  br2-a3110s1.cb1ma.ip.att.net (12.127.5.10)  3 ms  3 ms  3 ms
 8  br3-h20.wswdc.ip.att.net (12.127.15.177)  12 ms  13 ms  11 ms
 9  gr1-a3100s1.wswdc.ip.att.net (192.205.31.185)  13 ms  13 ms  13 ms
10  - 15 …. multiple hops in ALTER.NET, only a few shown in this slide
16  - 21 …. multiple hops in.Teleglobe.net, only a few shown in this slide
22  external-gw.ja.net (128.86.1.40)  145 ms  145 ms  143 ms
23  london-core.ja.net (146.97.251.58)  152 ms  142 ms  145 ms
24  146.97.251.82 (146.97.251.82)  150 ms  148 ms  149 ms
25  noucs2.backbone.ox.ac.uk (192.76.35.2)  152 ms  155 ms  150 ms
26  wwwtest.ox.ac.uk (163.1.0.45)  152 ms  150 ms  152 ms
fas%
```

One Minute Wrap-Up

- Please do this Wrap-Up at the end of each lecture.
- Please fill out the form on the website.
- The form is anonymous (but you can include your name if you want.)
- Please answer three questions:
  - What is your grand "Aha" for today's class?
  - What concept did you find most confusing in today's class?
  - What questions should I address next time
- **Thank you!**

Thank You!

ALIGHLSOD1701