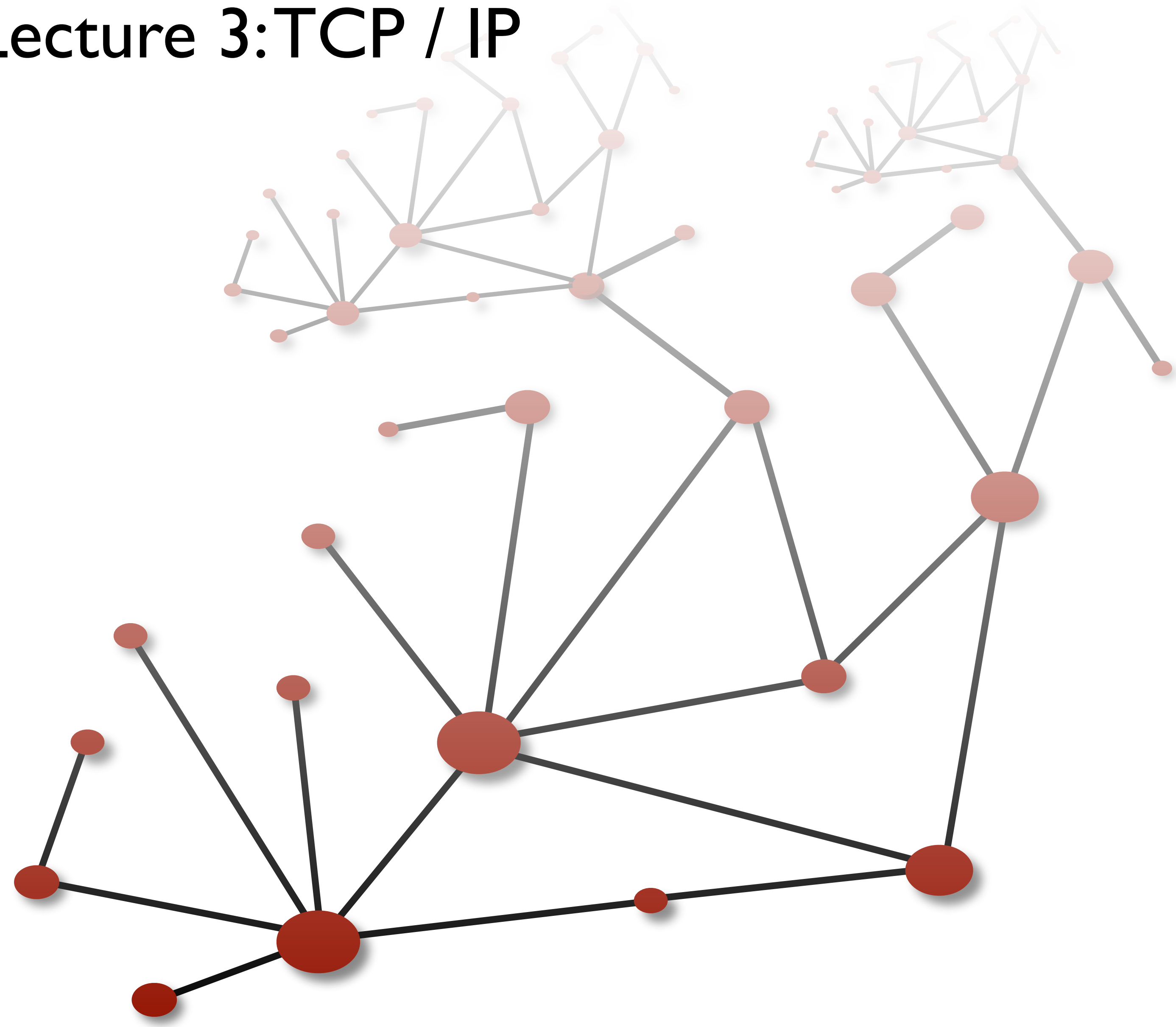
CS 598: Advanced Internet

Lecture 3: TCP / IP

Brighten Godfrey

pbg@illinois.edu

Fall 2009



Today

Announcements

A few more project ideas

Cerf and Kahn: TCP / IP

Clark: TCP / IP design philosophy

Announcements

- Project proposals due Sept 15
  - Find partner and chat with me before then
  - Submit half-page description of your plan and roughly what each group member will do
- Readings through Sept 15 on web site
- Thursday:
  - Jon Postel. Internetwork protocol approaches. IEEE Transactions on Communications, 28(4):604-611, April 1980.
  - J.H. Saltzer, D.P. Reed and D.D. Clark. End-to-End Arguments in System Design. ACM Trans. on Computer Systems, Vol. 2, No. 4, Nov 1984, pp. 277-28

Upcoming presentations

- Soliciting volunteers to present Tuesday Sept 8 (and later):
  - Van Jacobson. Congestion Avoidance and Control. Proc. SIGCOMM 1988, pp. 314-329.
  - Dah-Ming Chiu and Raj Jain. Analysis of the Increase and Decrease Algorithms for Congestion Avoidance in Computer Networks. Computer Networks and ISDN Systems, Vol. 17, No. 1, June 1989, pp. 1-14.

Paper reviews

- Preferred format: plain text pasted into email. No attachments necessary.
- No summary necessary. Just criticism.
- One paragraph is sufficient. Longer is not better. :-)

Today

Announcements

A few more project ideas

Cerf and Kahn: TCP / IP

Clark: TCP / IP design philosophy

Optimally hierarchical distributed systems

- Distributed systems frequently hierarchical: some set of nodes are picked for greater responsibilities (e.g., content distribution systems, Skype, distributed hash tables)
- Larger set of these "superpeers" brings more capacity (good) but potentially greater overhead and worse service quality (bad!)
- How do you balance these tradeoffs optimally? (e.g., if n superpeers incur $\log(n)$ overhead factor, and you know the distribution of node capacities, what is the optimal set of superpeers?)

emailfs

- We use email a lot like a filesystem
- Admit it, and design an email system that has the best of both worlds. Compared with email,
  - avoid explicit duplication of content
  - integrated versioning of files?
  - ideas from distributed filesystems to deal with large files?

Incentive compatibility of congestion control

- What congestion control schemes are both efficient and incentive compatible?
- Intermediate problem: convergence with feedback effects
- Simulate these effects using ns2 or similar packet-level evaluation, working with Brighten and coauthors

Today

Announcements

A few more project ideas

Cerf and Kahn: TCP / IP

Clark: TCP / IP design philosophy

Interconnection challenges

- Different ways of addressing, supported packet lengths, latency, status information, routing
- Must let each network operate independently
- Solution:

IP

"unacceptable alternative"

Hosts

Protocols

Gateways and IP

- Gateways sit at interface between networks

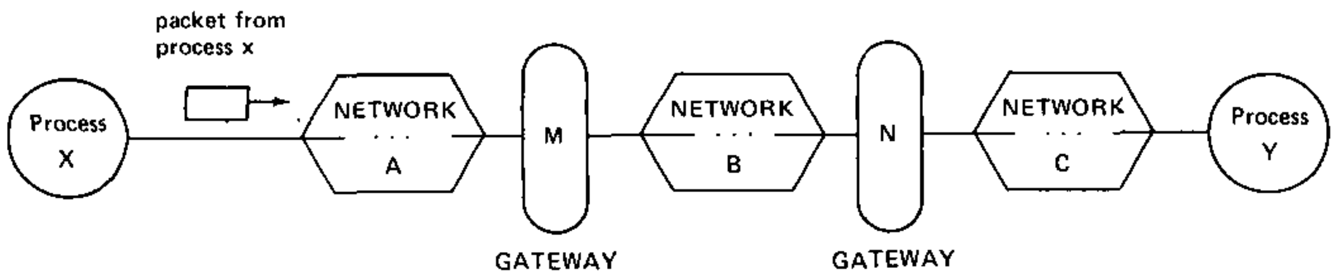


Fig. 2. Three networks interconnected by two GATEWAYS.

- ...and speak an Internetworking protocol

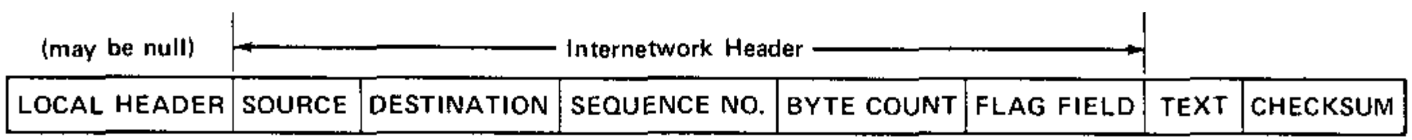


Fig. 3. Internetwork packet format (fields not shown to scale).

IP packet fragmentation

- Allow maximum packet size to evolve
- Protocol mechanisms to split packets in-transit (byte-level sequence numbers)
- Reassemble at end-hosts
  - Why not gateways?

Unreliable datagrams

- No need for reliability support from underlying network
- Greatly simplifies design
  - Exception handling always adds complexity
  - Any problem? Just drop the packet.
- What's not a stated reason for datagrams?
  - Statistical multiplexing.

Addressing & routing

- Unspecified—but not unconstrained!
  - Picks address format
  - Hierarchical (network, host)
  - Route computed within network
  - 8 bits for network. *"This size seems sufficient for the forseeable future."* Later: 32 bits in three size classes (A,B,C), and then CIDR.

Fig. 4. TCP address.

- Just about every new routing algorithm we'll see needs to change the packet's address format.

Ports

- Associated with a process on a host
- Identify endpoints of a connection ("association")
- Rejected design: connection at host level; packet may include bytes for multiple processes

What we now call TCP

- Window-based scheme
- Provides reliability, ordering, flow control
  - And now, congestion control too
  - Note you might want only some of these
- Not in this version: 3-way handshaking, congestion control (next section of course!)

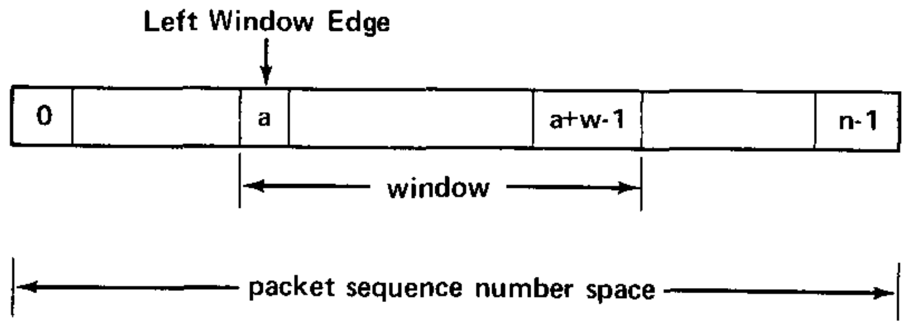


Fig. 10. The window concept.

Today

Announcements

A few more project ideas

Cerf and Kahn: TCP / IP

Clark: TCP / IP design philosophy

Goals of the architecture

0. Interconnect existing networks
1. Survivability
2. Multiple communication services
3. Variety of networks
4. Distributed management
5. Cost effective
6. Easy host attachment
7. Resource usage accountability

0. Interconnect networks

- Assumption: One common architecture
- Technique: packet switching
  - Met target application needs
  - Already used in ARPANET, ARPA packet radio network
- Interconnect with layer of gateways (packet switches)

1. Survivability

- Definition: even with failures, endpoints can continue communicating without resetting high-level end-to-end conversation
  - Except when?
- Did this work?

1. Survivability

Key question for survivability:
Where is connection state stored?

In network

So, must replicate

- Complicated
- Does not protect against all failures

On end hosts

Shared fate

- Simpler
- If state lost, then it doesn't matter

Conclusion: stateless network — datagram packet switching

2. Multiple types of service

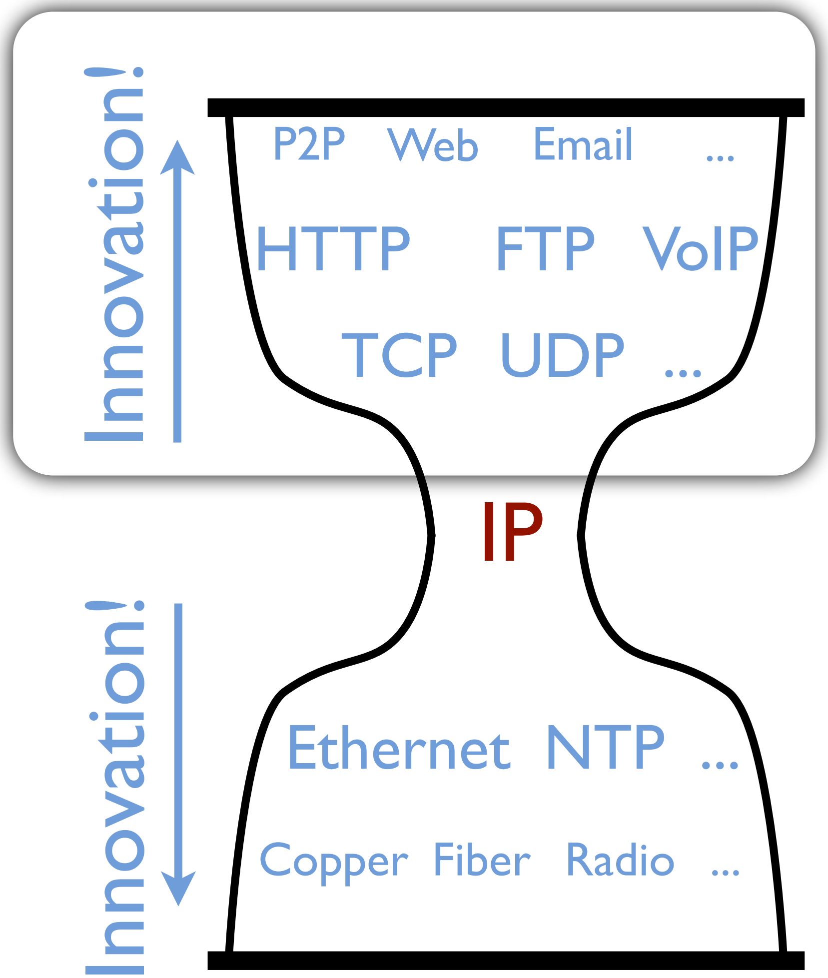
- Initially, just TCP

- But some apps reliability!

  (VoIP, XNET)

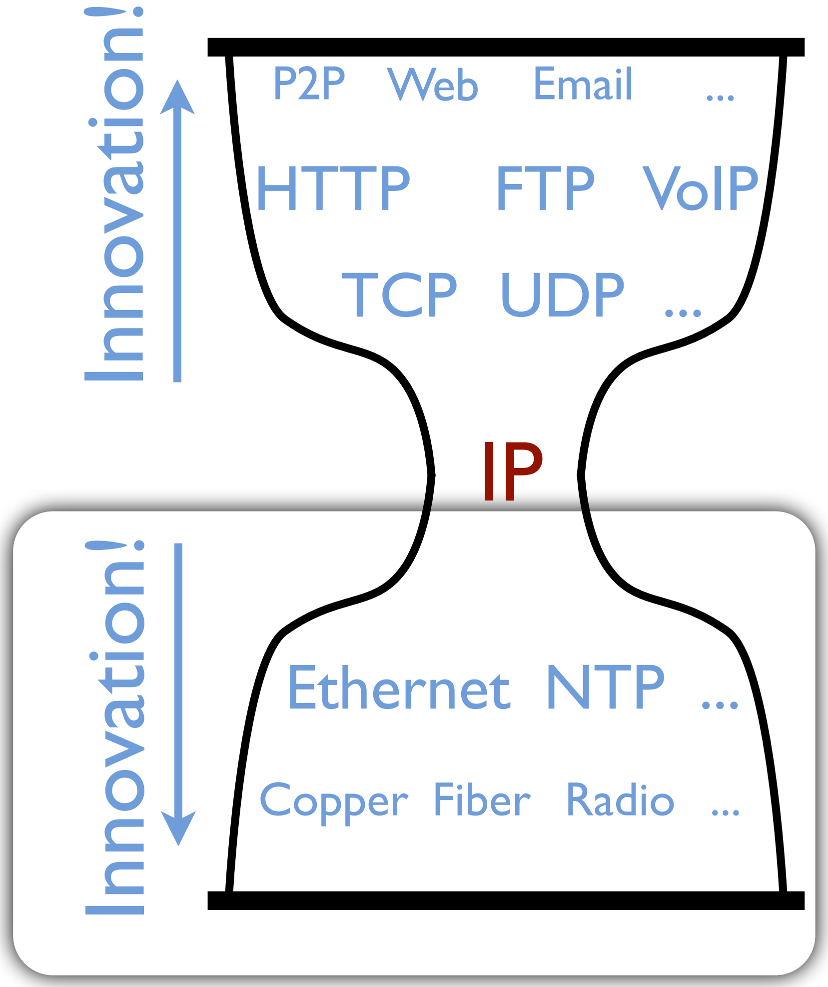
2. Multiple types of service

- So, TCP/IP split: datagram basic building block for many types of service
- Still difficult to support low latency across all networks
  - Hard to remove reliability if underlying network provides it
- For what services is IP not sufficient?



3. Variety of networks

- Datagram is simple building block
- Few requirements from underlying network technology



4. Distributed management

> “*... some of the most significant problems with the Internet today relate to lack of sufficient tools for distributed management, especially in the area of routing.*”

— David Clark, 1988

Still a problem 20 years later!

Later in this course: new easier-to-manage architectures (Ethane, OpenFlow, SEATTLE)

5. Cost effective

- Inefficiencies:
  - 40 byte header
  - retransmission of lost packets

- How much does it matter now?

6. Easy host attachment

- End-hosts must implement net services
  - once caused concern to some people
  - problems if host misbehaves!

7. Accountability

- Difficult to account for who uses what resources
  - How is this done today? Why is it only an approximation?
  - Both an economic and security issue
  - Later in this course: Anderson et al., Accountable Internet Protocol, SIGCOMM 2008.

What it doesn't do

- *"The architecture tried very hard not to constrain the range of service which the Internet could be engineered to provide."* Extremely successful! But:
    - Hard for network to report that it failed ("potential for slower and less specific error detection")
    - Resource management (coming soon: fair queueing)
    - Multipath
    - Full illusion of reliability during failures
- Security: Clark discusses host misbehavior (briefly) and accountability, but other aspects missing

Clark’s new terms

fate-sharing

flow

soft state

What if the Internet were commercial?

- Different priorities: accountability first, survivability last
- Example: Videotex networks
- e.g., France Telecom’s Minitel
  - Centralized
  - Reliable
  - Banking, news, stock transactions, ...

photo: wikimedia

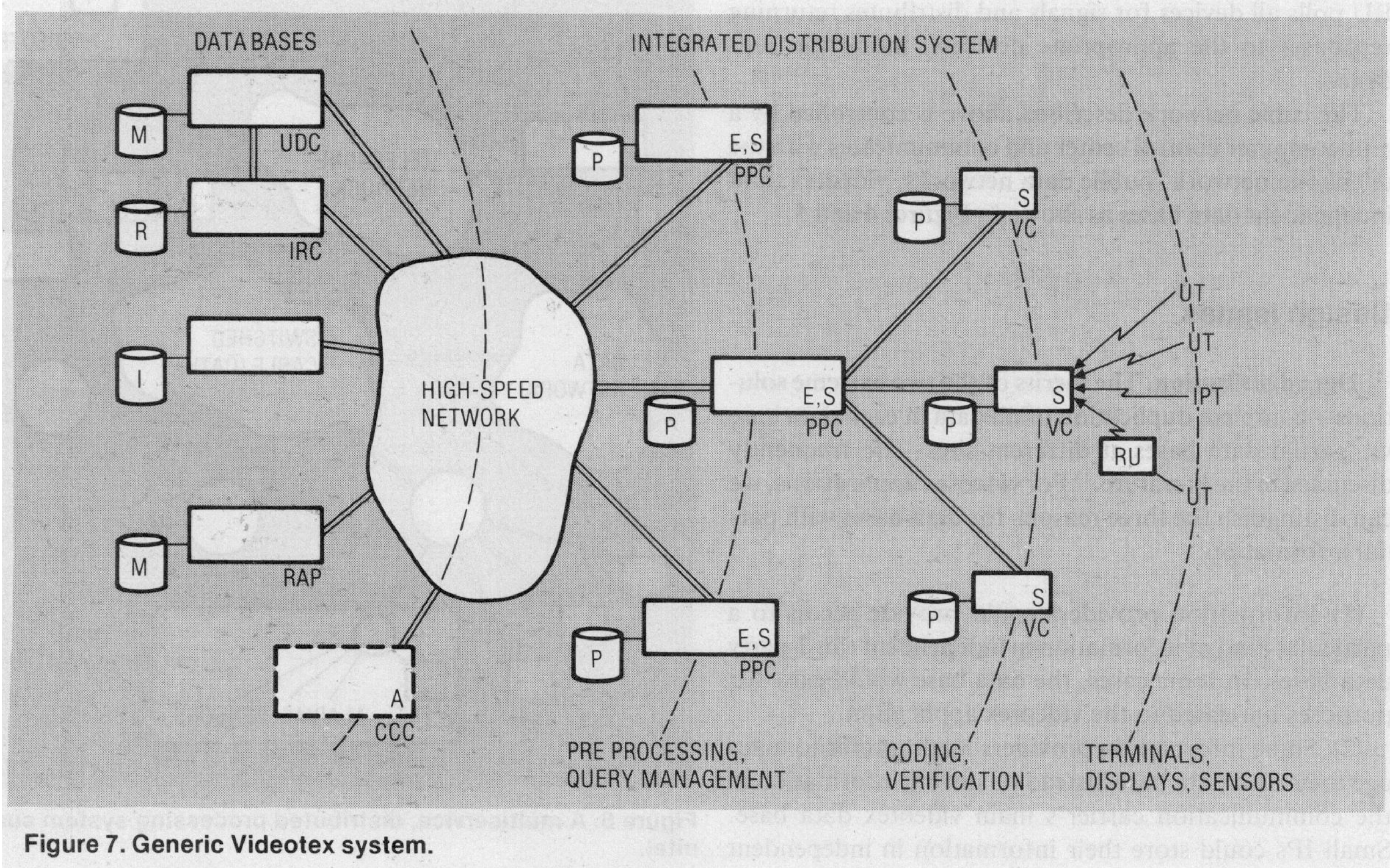


Figure 7. Generic Videotex system.

[A. J. S. Ball, G.V. Bochmann, and J. Gecsei. Videotex networks. IEEE Computer Magazine, 13(12):8–14, December 1980]

What's next

- Thursday:
  - Jon Postel. Internetwork protocol approaches. IEEE Transactions on Communications, April 1980.
  - Saltzer, Reed and Clark, "End-to-End Arguments in System Design," ACM Trans. on Computer Systems, November 1984.
- Tuesday: congestion control begins

Volunteers?

- Soliciting volunteers to present Tuesday Sept 8 (and later):
  - Van Jacobson. Congestion Avoidance and Control. Proc. SIGCOMM 1988, pp. 314-329.
  - Dah-Ming Chiu and Raj Jain. Analysis of the Increase and Decrease Algorithms for Congestion Avoidance in Computer Networks. Computer Networks and ISDN Systems, Vol. 17, No. 1, June 1989, pp. 1-14.