COMP312-09A Communications and Systems Software

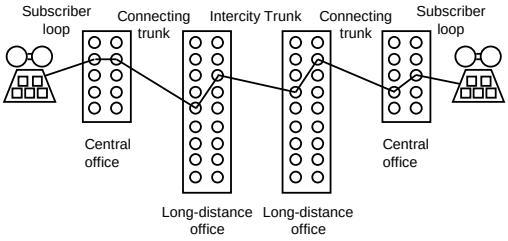
Lecture 9 – Switching and Queues II

THE UNIVERSITY OF WAIKATO

Overview

- Last lectures
  - TCP uses packet loss as an indication of congestion
  - Congestion occurs when a node is unable to forward packets at the rate they arrive at.
  - Queues tend to operate best when the presented load is no more than 70% of the server's capacity.
- This lecture
  - A bit more on circuits, including virtual circuits
  - Random early detection
  - Explicit congestion notification

Circuit switching



Circuit Switching

- Circuits are suited to
  - Continuous steady rate traffic such as voice
  - Analogue signals
- Three phases of a circuit
  - Setup
  - Information transfer
  - Tear down
- Signalling is used to setup, maintain, and tear down circuits

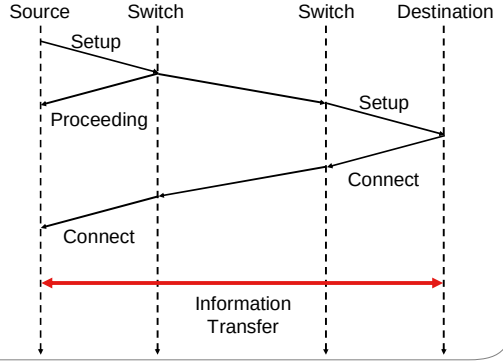
User Signalling

- User to network signalling is done:
  - When a circuit is required
  - To select destination of the circuit
  - When the destination connects to the circuit
  - To enquire about status of the circuit
  - When a circuit it to be released

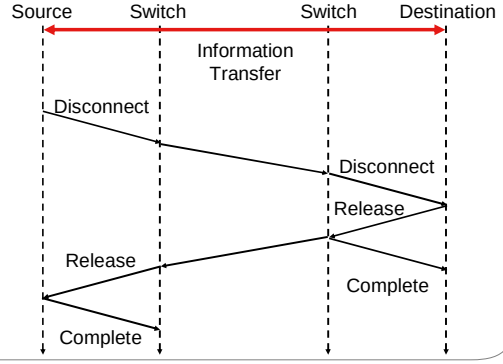
Network Signalling

- Network switch to network switch signalling is required
  - To find a route
  - To set up circuits.
  - To release circuits.
  - For network management.

Circuit switching



Circuit Switching



Issues with circuit switching

- Comparatively expensive
- If node or link in circuit fails, the circuit fails too
  - In a packet switched network, each packet is routed independently of any previous packet, so failures can often be routed around
- Switches need to keep per-connection state

Issues with packet switching

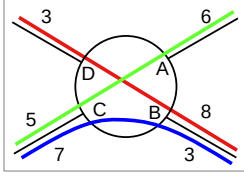
- Difficult to provide Quality of Service
- Very difficult to guarantee information delivery
- Arrival order of packets not guaranteed
- If packets take different routes, some delay variation can occur
- Switches must match destination address to topology for routing every packet at every node.

Connection-oriented packet switched networks

- Virtual circuits
  - X.25
  - Frame Relay
  - Asynchronous Transfer Mode (ATM)
  - Multi-protocol Label Switching (MPLS)
- Signalling just like a circuit switched network
- State about each circuit held by VC switches
- Packets are identified as belonging to a particular connection
  - "labelled"

Virtual Circuit Forwarding

- Labels are normally unique only on each link
  - Switch re-labels packet before forwarding it
- Each switch maintains a list of
  - <incoming port, circuit label> pairs
  - corresponding <outgoing port, circuit label>



In		Out	
Label	Port	Label	Port
6	A	5	C
7	C	3	B
3	D	8	B

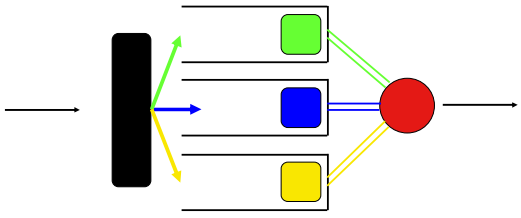
Connection-oriented packet switched networks

- Advantages
  - Minimum packet overhead – just a label per packet
  - Less per-packet processing at switches
  - Straightforward to provide Quality of Service
    - Reliability and Delay can be controlled
  - Order of packet arrival same as transmission
- Disadvantages
  - Signalling protocol is complex
  - Failure of switch causes failure of all VCs through that switch
  - Switches keep per-connection state
  - Can't provide VCs across the Internet to an arbitrary destination

Back to Queuing

- Queue types range from the obvious
  - First in, First out (FIFO)
- to more complicated
  - Priority Queuing
  - Class-based Queuing
  - Leaky bucket

Priority Queuing



(Strict) Priority Queuing

- Queues have different priorities
- Queues are serviced in order of priority
  - Serviced until empty
  - Highest priority queue experiences lowest delay, lowest drop rate
  - Other queues fare less well
    - Can receive no service for a long time if there is sufficient high priority traffic

Class Based Queuing

- Queues receive a service allocation
  - E.g. queue A gets 50% of total service, B and C 25% each
- Queues are serviced round-robin according to their allocation
- Can control whether unused allocations are shared among other classes (or not)
- Allocation size is a trade off between fairness and delay variation

Queues and Congestion Control

- Congestion occurs when a router is unable to forward packets as fast as they are arriving
  - Queues fill up
  - Packets that arrive when the queue is full are discarded
- TCP reacts to loss by at least halving the rate of transmission

Queues and Congestion Control

- When a queue fills, all TCP connections that arrive at the queue experience congestion at about the same time
  - Therefore, all currently active TCP connections will halve their rate of transmission
  - The queue in question will quickly empty, and throughput (work done) will reduce
  - The TCP connections in question will then slowly increase their rate of transmission until the queue fills
  - Repeat from top
- Known as **TCP global synchronisation**

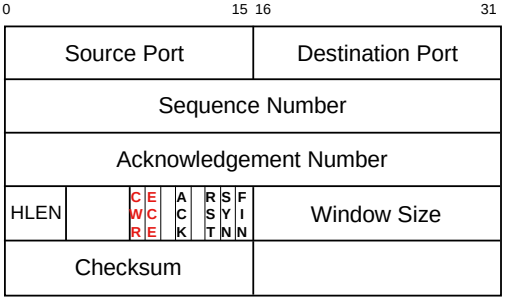
Random Early Detection

- Random Early Detection / Discard / Drop
- Router will randomly drop packets as a function of its queue size
  - If queue is empty, or below a defined threshold, an arriving packet will not be discarded
  - If queue is full, all arriving packets will be discarded
  - If queue is above threshold, an arriving packet could be discarded
- Because packets are discarded early, connections will experience congestion at different times
  - i.e. each will back off at different times, avoiding synchronisation
- The faster a TCP is transmitting, the more chance it has of experiencing an early drop
  - i.e. RED is reasonably fair

Explicit Congestion Notification (ECN)

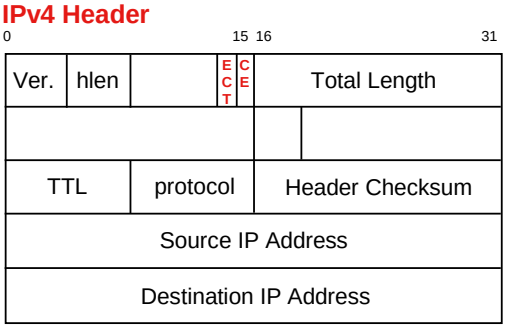
- Instead of discarding packets when the queue is not yet full (before the situation is dire) advise the sender to reduce their rate of transmission
- On the Internet, this can be done with ECN
  - in reality, ECN is not well deployed
  - however, this is a university.
- Transport protocols have to be ECN capable.
- History
  - First proposed in 1994.
  - RFC 2481, 1999 *experimental* proposal to add ECN
  - RFC 3168, 2001 proposed standard
  - 2009: not terribly well supported

TCP/IP with ECN



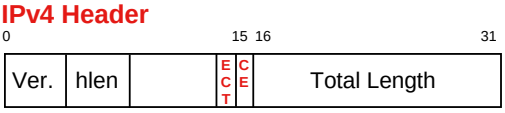
ECE: ECN Echo
CWR: Congestion Window Reduced

IPv4 Header



ECT: ECN Capable Transport
CE: Congestion Experienced

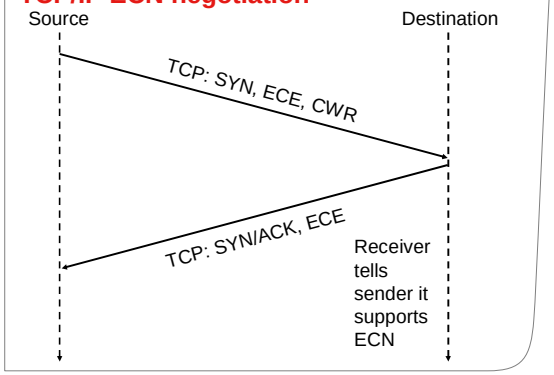
IPv4 Header



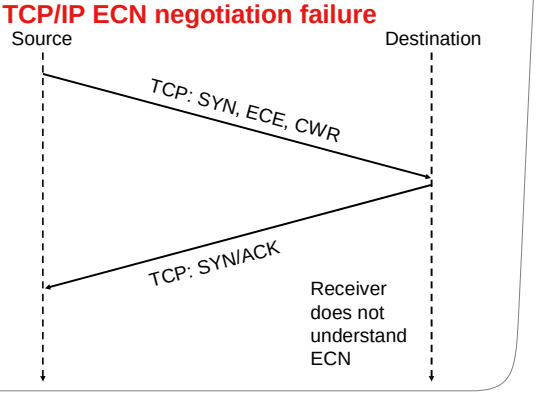
In practice, the naming of these two bits is obsolete

ECT	CE	
0	0	Not-ECT
0	1	ECT(1)
1	0	ECT(0)
1	1	CE

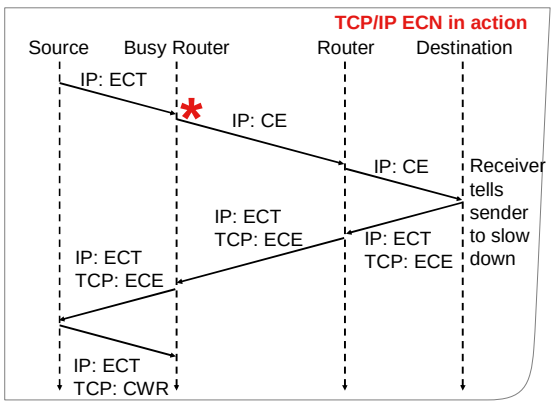
TCP/IP ECN negotiation



TCP/IP ECN negotiation failure



TCP/IP ECN in action



TCP/IP ECN

1. Sender transmits packet with ECT set in IP header
2. Router in forward direction is close to experiencing congestion, decides to mark CE in IP header
3. Destination notices CE in IP header, sets ECE in TCP acknowledgement reply packet
   - and all TCP ACK packets from then on, until it receives CWR
4. Sender receives ACK, notices ECE bit, reduces congestion window (rate at which it is sending) and sends CWR to the destination

Alternative to ECN

- ICMP Source Quench
  - Defined in 1981, RFC 792
  - Router sends a message to sender, asking it to slow down.
- Not ever used these days
  - Creating an additional packet when congested in theory just adds to the load of the network

ECN support

- Most operating systems support use of ECN in TCP
  - Linux has for a long time
  - FreeBSD 8.0 (upcoming release)
  - MacOS 10.5
  - Windows Vista
- By default, it is turned off
  - Issues with devices in the network that do not understand ECN bits in IP and TCP headers, which then discard the packets outright

Summary

- Virtual circuits provide reserved capacity in a packet-switched network
- Through smarter queues, it is possible to improve performance of TCP
  - RED – widely deployed
  - ECN – not widely deployed
- Further reading
  - Pages 344 - 350
  - Pages 391 - 395