

INTRODUCTION TO TELECOMMUNICATIONS NETWORKS MSc MODULE (ITN)

Packet Network Technologies

Prof. George Pavlou
Communication and Information Systems Group
Dept. of Electronic and Electrical Engineering
<http://www.ee.ucl.ac.uk/~gpavlou/>
g.pavlou@ucl.ac.uk

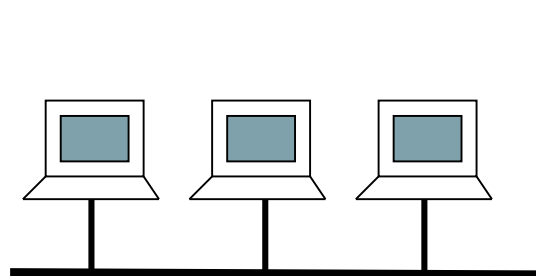
Introduction

- Packet networks connect computing devices together.
 - Also referred to as computer or data networks.
- Key services offered are:
 - Electronic mail, file transfer, file and printer sharing, remote login, electronic directory services, hypertext document access (the web), audio/video streaming.
- Present and future: higher speeds, better user quality of experience, advanced multi-services, next generation mobile networking (5/6G), mobile edge computing (MEC).

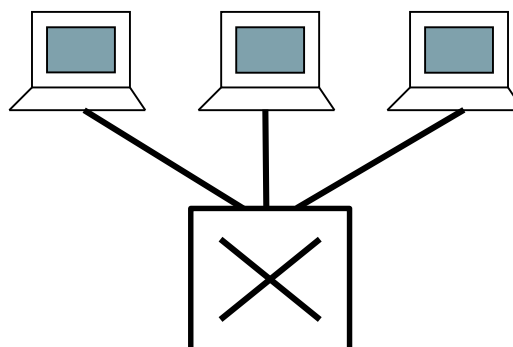
Types of Packet Networks

- **Local Area Networks (LANs)** - typically used within a building or in buildings situated close together.
 - E.g. Gigabit Ethernet , Wireless LAN
- **Metropolitan Area Networks (MANs)** - employ LAN-like technologies over distances of some kilometres e.g. in a town/city.
 - E.g. Metro Ethernet
- **Wide Area Networks (WANs)** - from distances of several kilometres upwards.
 - WANs are “mesh” type networks, with network nodes connected in a graph and with customer networks attached to “edge” nodes.

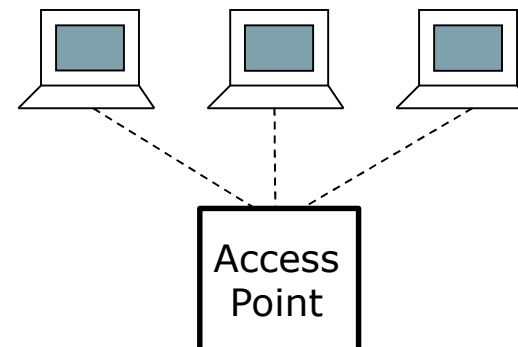
Local Area Networks



Older style broadcast Ethernet



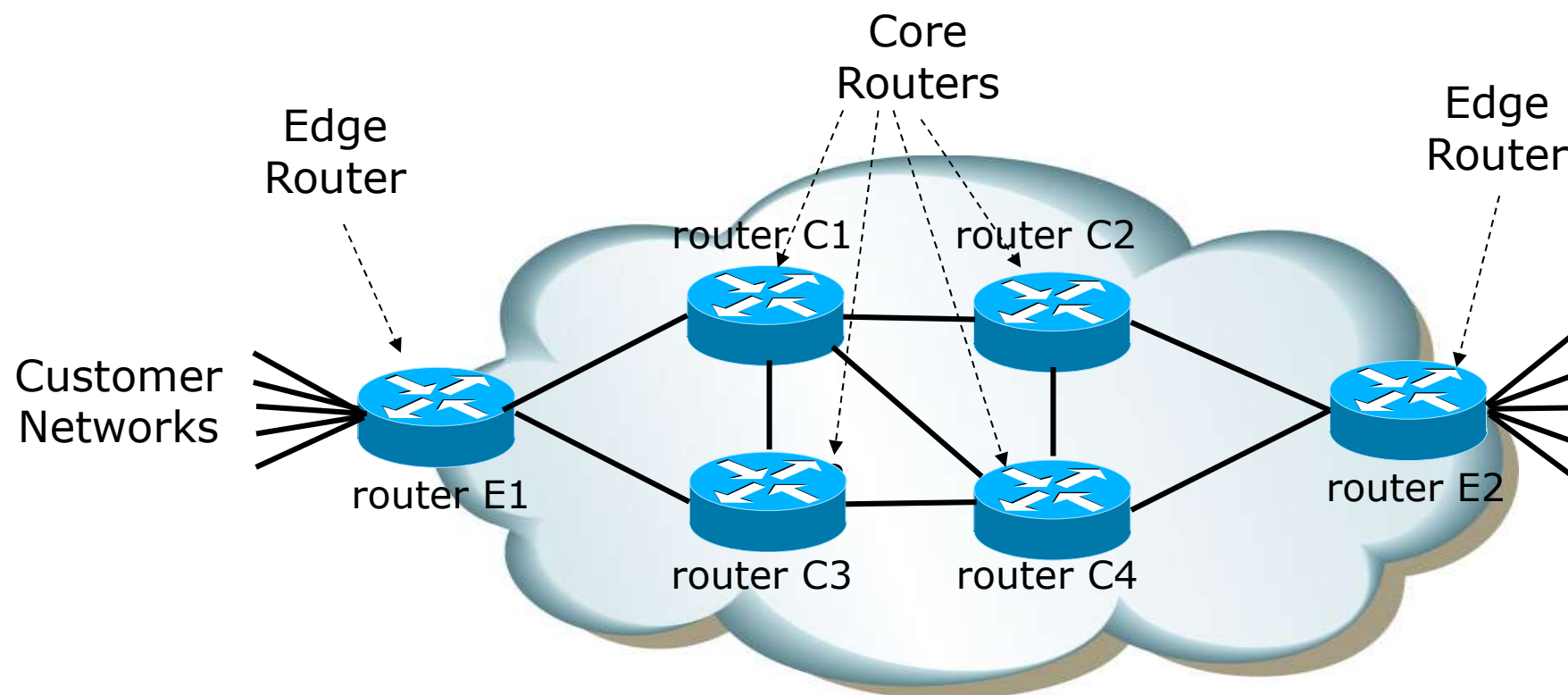
Switched Gigabit Ethernet



Wireless LAN (WLAN)

- In older broadcast Ethernet and in WLANs, all terminals listen to all other terminals; regulation of access to the shared medium is required in order to avoid packet “collisions”
- Gigabit Ethernet is “switched”: all terminals are connected to a central switch which switches a packet only to the destination terminal, hence there are no collisions
- All nodes have Layer 2 (Ethernet in this case) addresses

Wide Area Networks

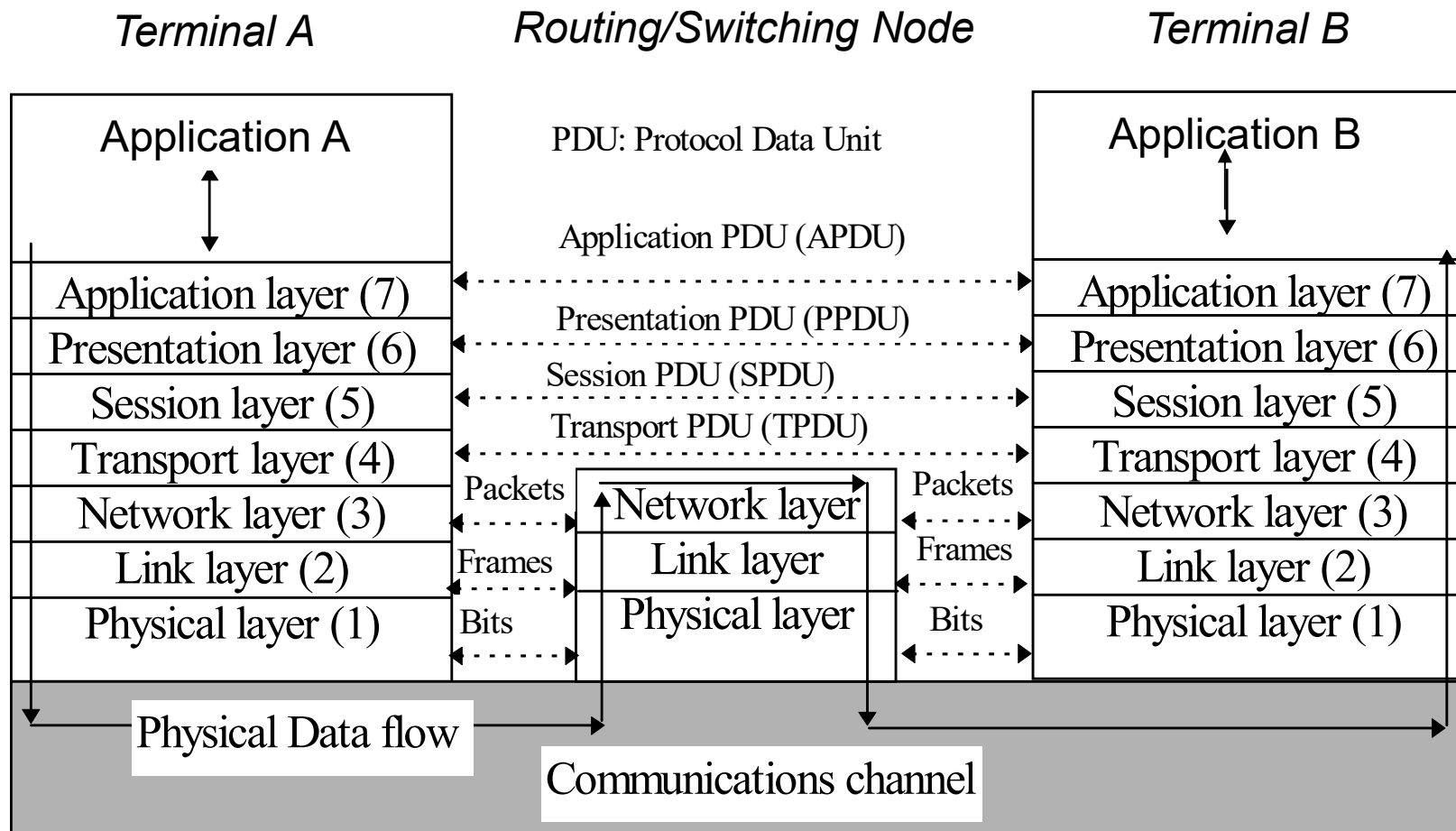


- All nodes build up routing tables through routing protocols *beforehand*, stating the next node towards a remote destination
- Packets are forwarded based on routing table information
- All nodes have Layer 3 (e.g. IP) addresses

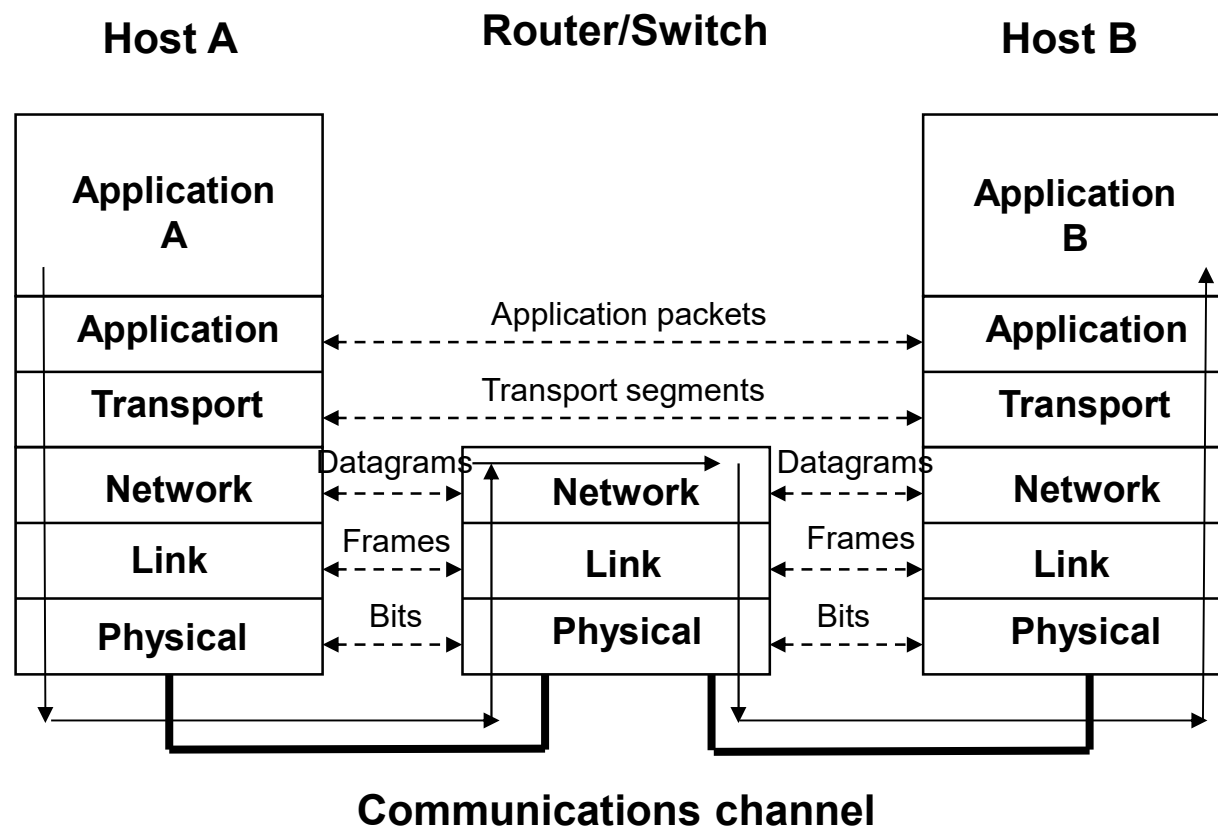
Packet Network Standardisation

- ISO/ITU-T have produced standards based on the Open Systems Interconnection (OSI) reference model
 - X25 and ATM have been relevant network technologies which are in little use today
 - Cellular/mobile communication technologies have and are also being standardised by the ITU-T
- The Internet Engineering Task Force (IETF) produces standards for the Internet
 - IP is the key network technology today, together with MPLS
- IEEE has produced standards for LANs
 - The 802 family of standards, Ethernet (802.3) and Wireless LAN (802.11) among other

The ISO/ITU-T OSI Reference Model



The Internet Reference Model



- No session and presentation layer: such functionality is provided when required directly by the application layer or by a particular application

Packet Reliability

- Reliability is crucial for data communications
 - Packets should be received exactly as transmitted, i.e. with no bits corrupted, in the same sequence sent and with no packets lost
 - Packets may be corrupted because of channel noise, which is high in wireless channels but very low in wired channels e.g. optical links
 - Packets may be lost because of congestion - router buffer overflow
- Reliability is achieved through error and flow control provided by sliding window protocols (we will discuss them later)
 - These protocols retransmit corrupted or lost packets while keeping their flow/sequence the same
- Reliability is **not good** for real-time audio & video as there is no point retransmitting, also the sliding window mechanism required for reliability reduces transmission speeds
 - Modern network technologies leave reliability to the transport and application layers

Link Layer Functionality

- The Physical layer provides a bit stream with noise / errors
- The Link layer adds the following functionality over it:
 - Medium access control (MAC) for a shared medium, e.g. WLAN; the MAC is effectively the lower part of the link layer
 - Framing: delimiting the beginning and end of a frame
 - Link connection management: the two ends of the link establish a connection and negotiate necessary parameters for it
 - Per-link error and flow control: provided in reliable service only - reliable service means frames will never be corrupted or lost
- Key link layer protocols are the following:
 - Point to Point Protocol (PPP) used in IP WANs, no error/flow control
 - Logical Link Control (LLC) in LANs; no error/flow control over Ethernet but error/flow control in WLAN as part of the 802.11 MAC protocol
 - Link Access Protocol Balanced (LAPB) in X.25 and Message Transfer Part 2 (MTP2) in SS#7 – both support error/flow control

Network Layer Functionality

- The Network layer can be organised for **either** connectionless (CL) or connection-oriented (CO) / virtual circuit (VC) operation and has the following functionality:
 - Network addressing, every node has a unique network address
 - Network routing, building routing tables through routing protocols
 - Packet forwarding, forwarding a packet based on a network address in CL operation or on a VC identifier (VCI) in the CO/VC operation
 - Virtual circuit set-up & release – only in CO/VC operation
 - End-to-end error & flow control – only for reliable service
- Key network protocols are the following:
 - IP, which provides CL unreliable service (no error/flow control)
 - ATM/MPLS, which provide CO unreliable service (no error/flow control)
 - X.25, which provides CO reliable service (with error/flow control)
 - SS7 MTP3 which provides CL reliable service (with error/flow control)

Transport Layer Functionality

- The Transport layer provides **both** connectionless (CL) unreliable and connection-oriented (CO) reliable service and has the following functionality:
 - Transport addressing, every application has a transport port
 - Higher layer packet segmentation, braking a large application packet to the required transport layer packets/segments in order to fit the underlying maximum transmission unit (MTU) – typically 1500 bytes
 - Connection set-up and release – setting up and releasing transport connections, only in CO reliable service
 - Per connection error & flow control – only in CO reliable service
- Key transport protocols are the following:
 - UDP, which provides CL unreliable service (no error and flow control); used for real-time audio/video and also for some data services in which reliability is provided by the application protocol or application
 - TCP, which provides CO reliable service (with error and flow control); used mostly for data services and also for non real-time audio/video

Upper/Application Layer Functionality

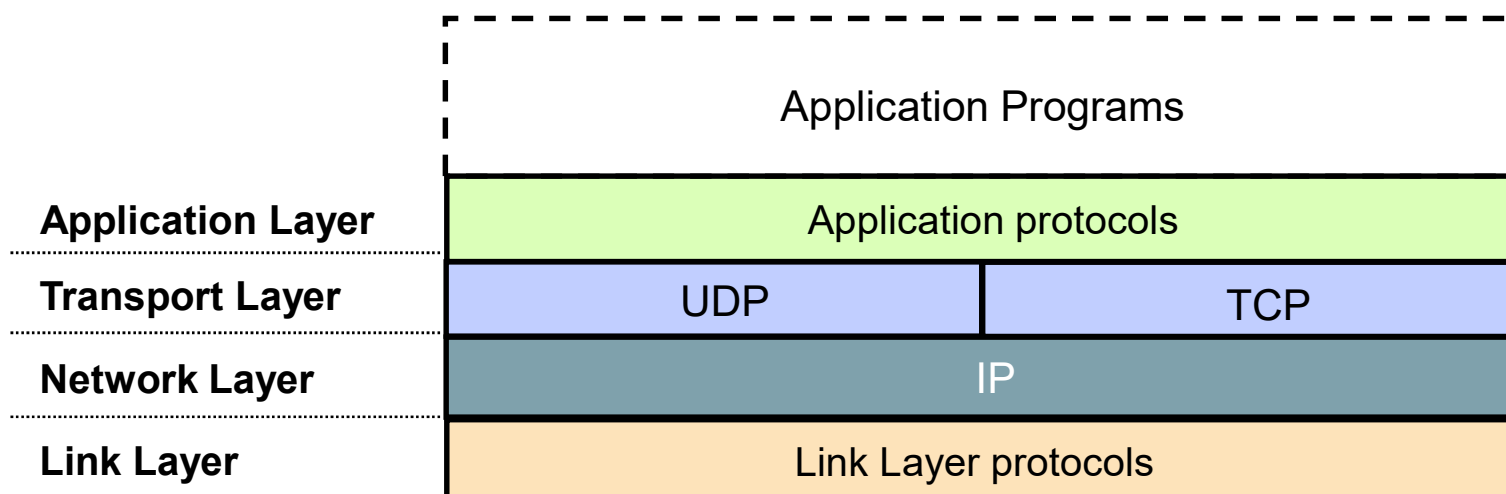
- Session Layer (5):
 - session management for applications by coordinating/merging multiple connections into a single session
 - e.g. recovering lost connectivity for a large data transfer and continuing the session when connectivity is recovered
- Presentation Layer (6):
 - data structure encoding/decoding
 - data compression
- Application Layer (7):
 - specific application-level protocols for email, file transfer, directory services, network management, web access, etc.

Link Layer Framing

- Framing: the ability to detect the beginning and end of a frame
- Frames are delimited by the flag pattern “01111110”, which signifies both the beginning and end of a frame
 - Issue if a flag byte happens to be part of the frame payload
- Byte-oriented framing: every flag byte in the payload is substituted by the flag-flag pattern and the additional flag byte is removed at the receiver – the technique is also known as “byte stuffing”
 - PPP uses byte-oriented framing
- Bit-oriented framing: every flag byte in the payload is substituted by “011111010” (a 9-bit pattern) and the additional 0 in the 7th position of this pattern is removed by the receiver – “bit stuffing”
 - If after 5 1’s there is a 0 followed by 10, it is removed by the receiver and the resulting byte is interpreted as simply a flag in the payload
 - LLC, X.25 LABP and SS7 MTP2 all use bit-oriented framing



Internet Host Layered Architecture



Packet Structure



- A packet “on-the-wire” contains a set of headers, payload/data and finally a checksum (CS) used to detect transmission errors.
- Every layer, from application to link, adds a header which contains information such as packet type, addressing, sequence number, etc.
 - A layer’s header is added as a packet goes down the stack – the link layer also adds the checksum – and it is removed as it goes up
- The payload/data is the “useful” information to transfer across.
- The length of the header and checksum reduce the effective bandwidth, so they should be kept relatively small
 - Headers/CS are 50 bytes, so best utilisation is $(1500-50)/1500 = 96.7\%$
- Some packets have no payload/data part, being used entirely for control purposes, e.g. acknowledgment packets (ACKs/NAKs).

Checksum-based Error Detection/Correction

- Bit Error Rate (BER): the probability that a single bit in a frame/packet is corrupted, e.g. 10^{-3} means 1 bit in 1000.
- The sender produces a checksum (CS) on the headers and payload which is appended to the packet/frame
- The receiver reproduces the CS and compares it with the CS on the packet: if not the same, the packet is corrupted due to BER, hence it is discarded and retransmission is requested
 - The CS is based on polynomial codes, a checksum of R bits (typically 16 or 32) is produced
 - The checksum production & check can be implemented in hardware and be very fast

ARQ Protocols

- Automatic Repeat Request (ARQ) protocols support error and flow control for connection-oriented reliable services
 - Found in the link (e.g. WLAN, X.25 LAPB, SS7 MTP2), network (e.g. X.25) and transport layer (e.g. TCP)
- **Error control** is implemented by the sender retransmitting corrupted packets (due to channel errors) or lost packets (dropped due to congestion / buffer overflow)
 - Selective-Repeat – only the corrupted/lost packet is resent, the receiver needs to re-order packets hence it is relatively complex
 - Go-Back-N – all packets sent after and including the lost one are resent - the receiver simply ignores packets after the lost one so it does not have to re-order anything, hence it is simpler
- **Flow control** is implemented through an agreed window size between the sender and receiver which signifies the maximum number of packets that can be un-acknowledged
 - The sender stops when it reaches the window size; as ACKs are received, the window “slides forward” and more packets can be sent

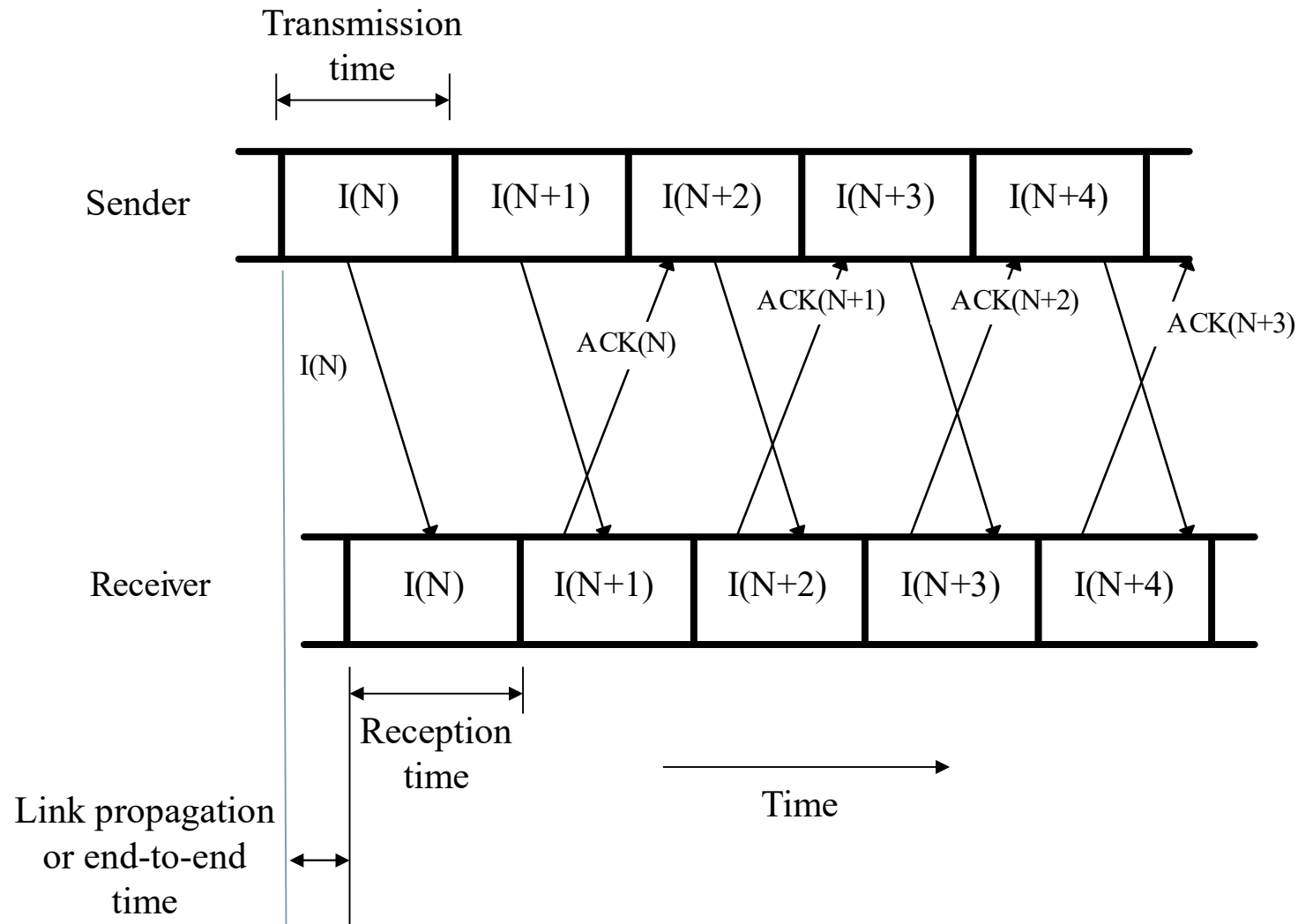
ARQ Protocol Operation

- Initially the two communicating parties exchange packets to establish a connection, in which key parameters such as the window size are agreed
- Information packets should be acknowledged by ACK packets or negative ACK (NAK) packets if not received correctly or lost
 - Information packets sent are kept by the sender until ACKed
- **Sequence numbers** are used to mark both information packets and ACK/NAK packets
- When an information or a NAK packet is sent, a timer is set and if it expires without response, the packet is re-sent
 - The timer should be bigger than the Round Trip Time (RTT)
 - RTT is easy to compute over a single link for a link-level ARQ protocol but difficult for an end-to-end transport ARQ protocol such as TCP, as it may fluctuate widely due to congestion and needs constant updating

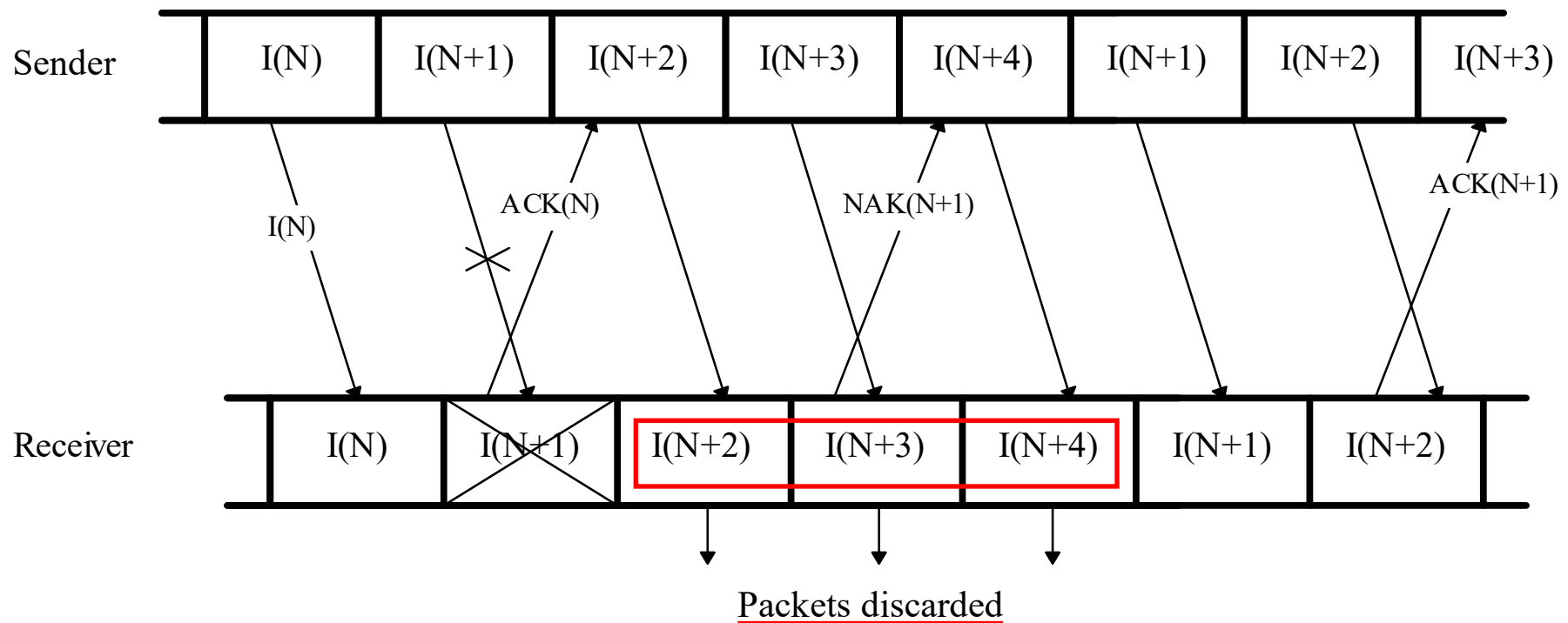
ARQ Protocol Operation (2)

- We will show next the operation of both Go-Back-N and Selective-Repeat ARQ protocols
- Some simplifying assumptions have been made in the following diagrams which though have no impact on the protocol operation
 - We assume that the window is large enough, so a continuous flow of packets is shown
 - We assume that the transmission time of ACK/NAK packets is too small in comparison to information packets, so it is omitted
- The key time depicted is packet transmission and reception
 - **For example**, for a 2Mbit/s link, the transmission/reception time for a max size IP 1500 byte packet is $1500 / ((2/8) * 10^6) = 0.006\text{sec} = 6\text{msec}$ - for a 50 byte ACK/NAK packet is $(50/1500) * 6 = 0.2\text{msec}$
 - Propagation time over the link is also depicted but tends to be very small for most links, apart from trans-oceanic and satellite ones
 - For end-to-end ARQ protocols, e.g. TCP, the overall propagation and queuing end-to-end time is significant, typically $> 100\text{msec}$

ARQ Continuous Flow With No Errors

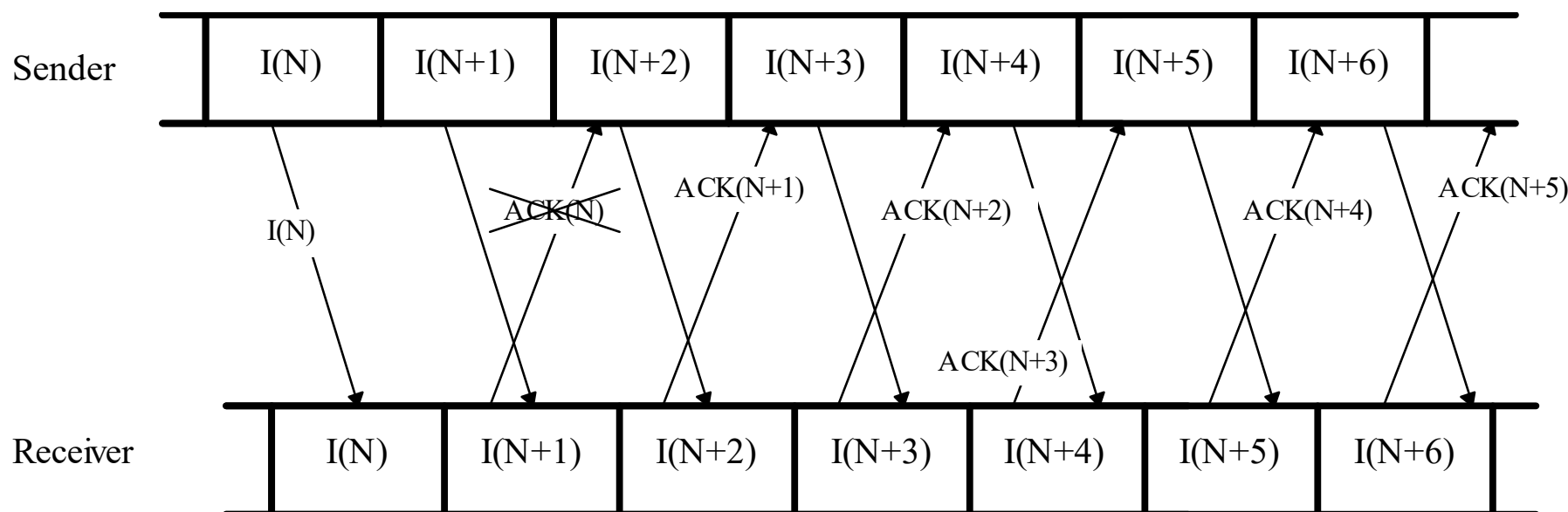


Go-Back-N: Corrupted/Lost Information Packet



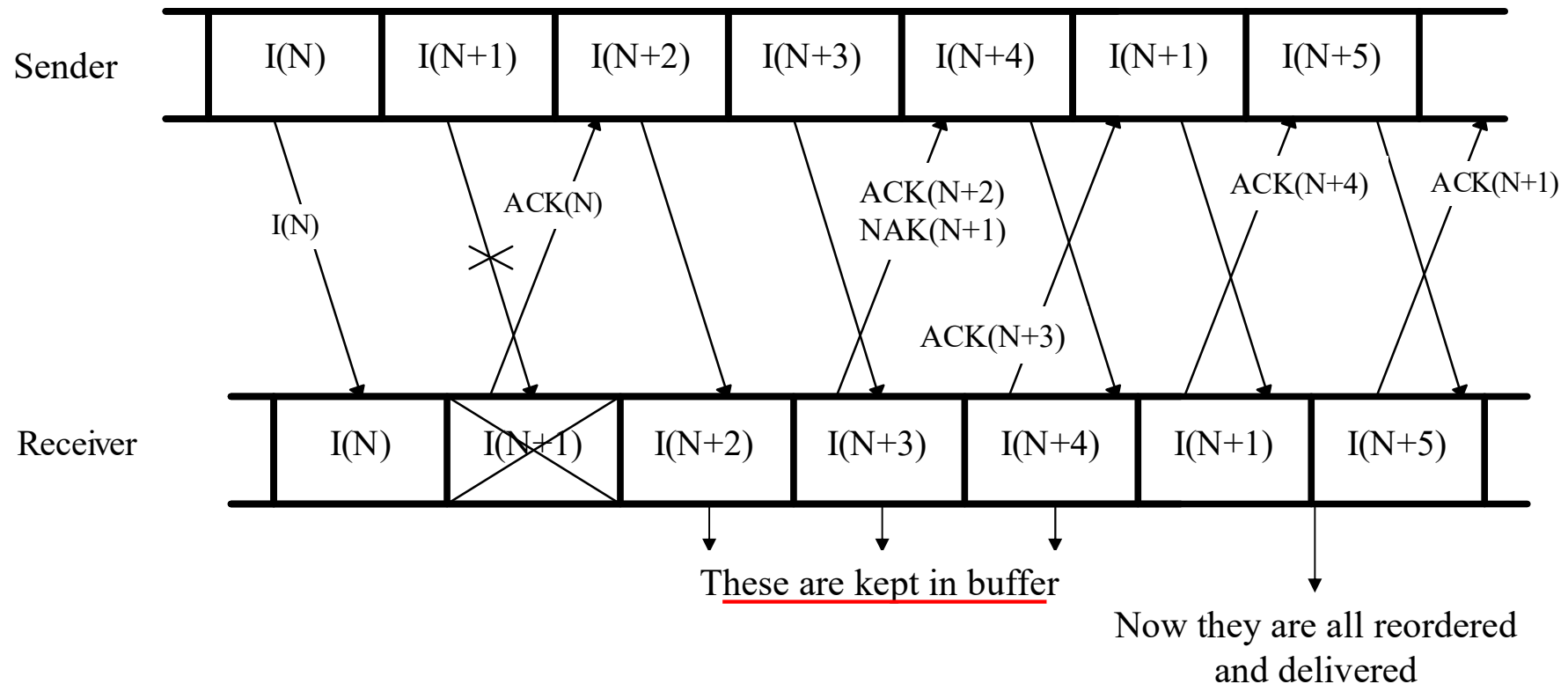
- Go-Back-N: wastes bandwidth with more retransmissions but the receiver is kept simple: just passes up or discards packets

Go-Back-N: Corrupted ACK Packet



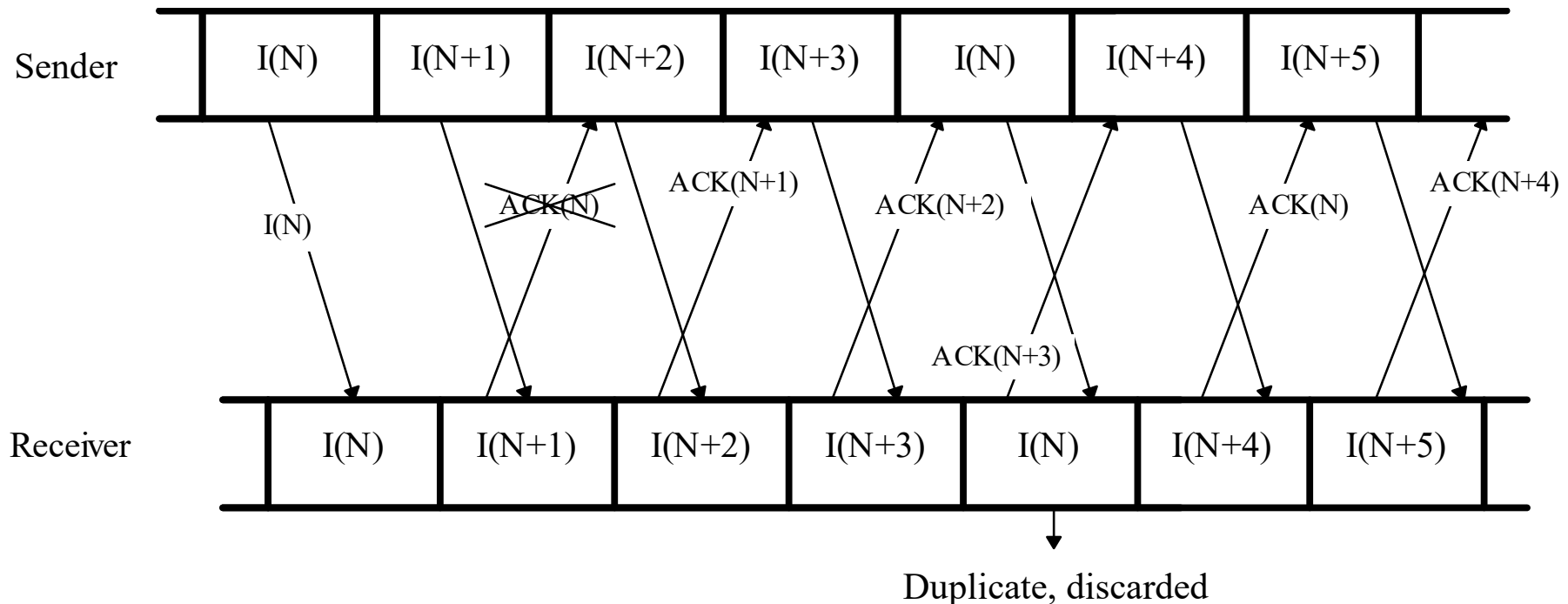
- As packets are acknowledged sequentially, the sender sees $ACK(N+1)$ without $ACK(N)$ and deduces that $ACK(N)$ has been lost so it takes no action
 - If there was no subsequent flow of frames, it would resend frame N after the relevant timeout – which would then be discarded by the receiver

Selective Repeat: Corrupted Information Packet



- Selective Repeat: saves bandwidth but the receiver is more complicated: has to buffer out-of-sequence packets and reorder

Selective Repeat: Corrupted ACK Packet

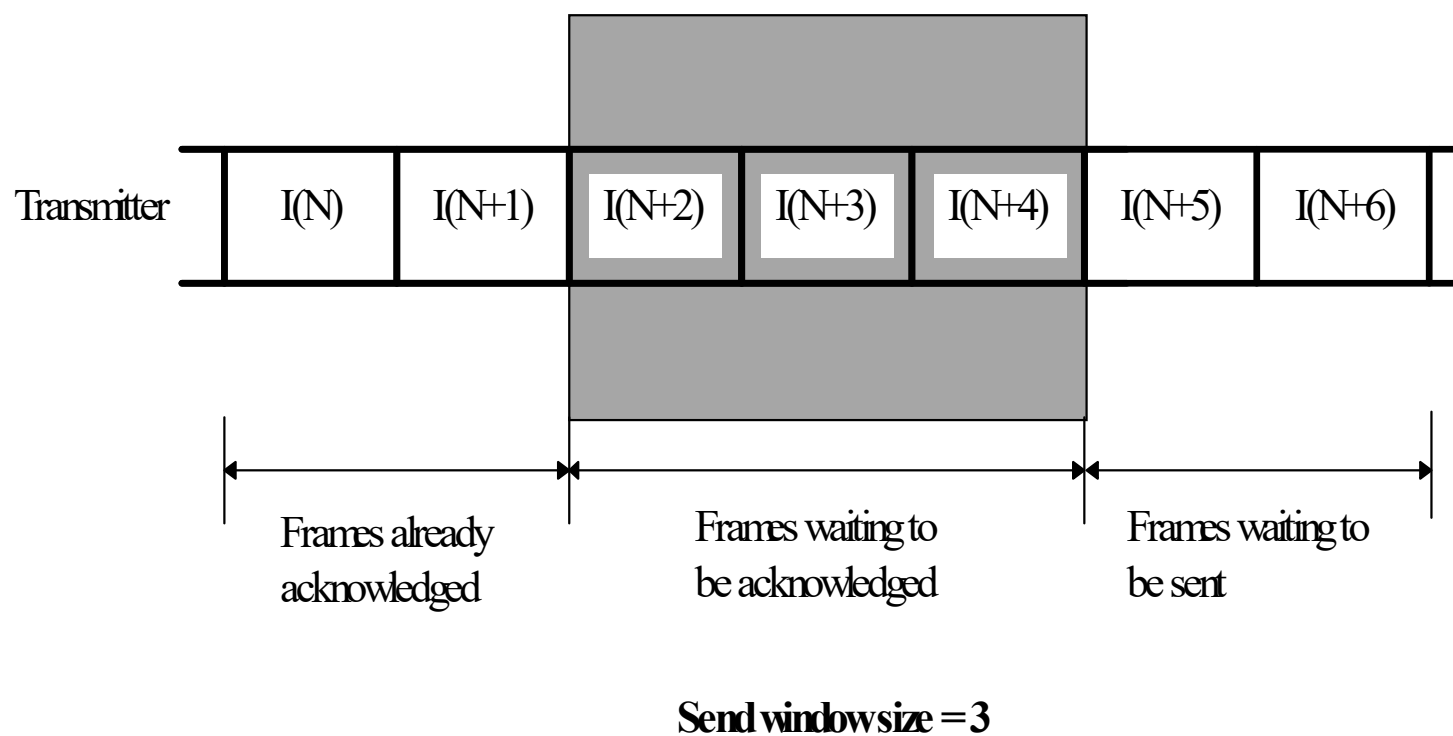


- As packets can be acknowledged in any order, when the sender sees $ACK(N+1)$ without $ACK(N)$ before it, it assumes that the frame was lost and resends it – to be simply discarded by the receiver

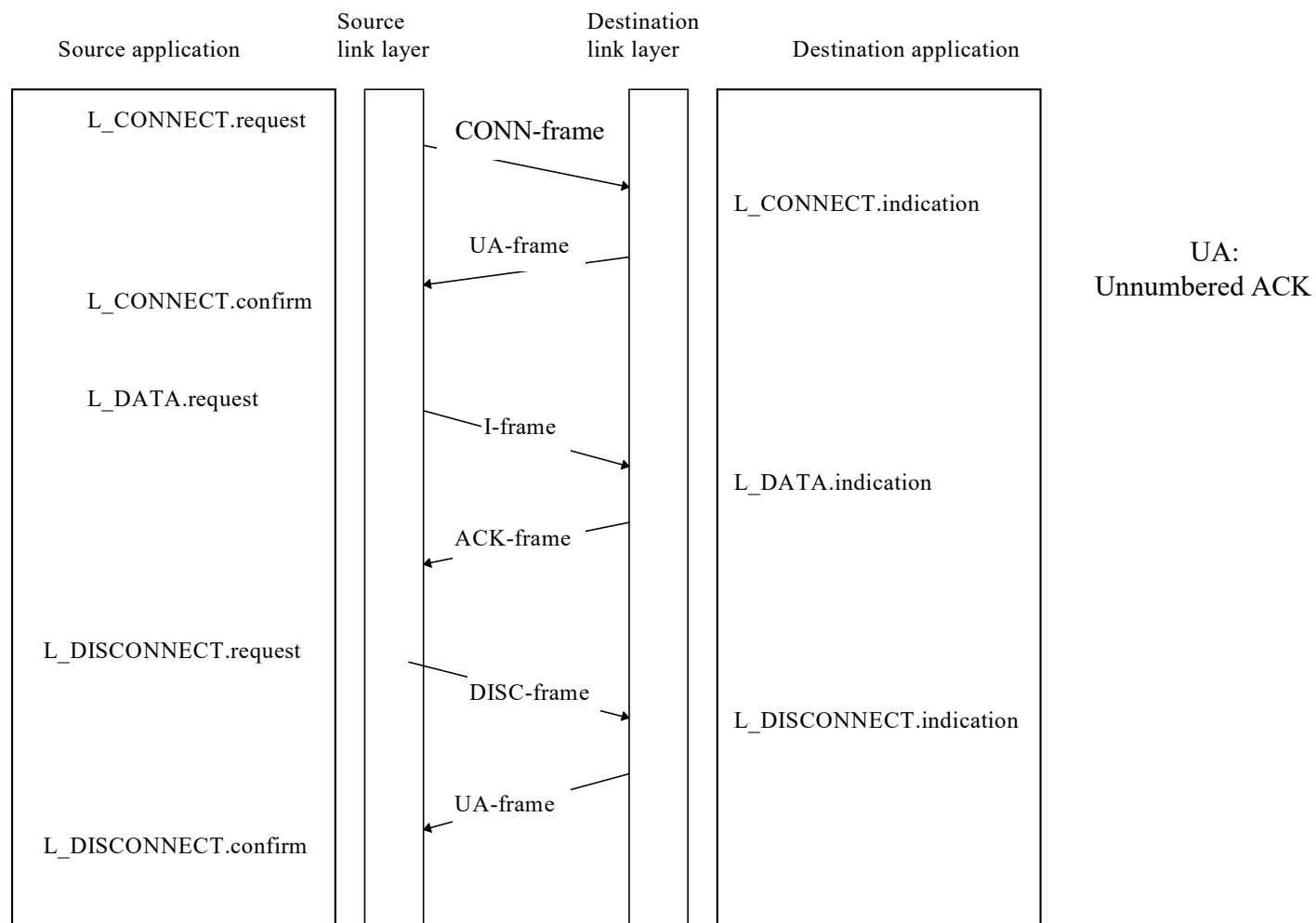
Sliding Window: Flow Control

- Flow control is implemented by stopping sending more packets when the window fills up and continuing when ACKs are received, with the window “sliding forward”.
 - The (max) window size is agreed at connection establishment time but can be also adaptive to network congestion and shrink e.g. in TCP
 - Allows senders and receivers of different capabilities, avoiding the possibility of a fast sender “overwhelming” a slow receiver; also avoids “overwhelming” the network with packets sent continuously
- Flow control has the following drawbacks:
 - Introduces gaps between successive packet bursts, i.e. introduces delay variation or jitter – no good for real-time audio/video, also packet reliability/retransmissions meaningless for those
 - Reduces throughput because of the start-stop approach in the sender, i.e. send packets up to the window size and then stops - hence wastes link bandwidth and cannot support high speeds
 - Ok for data but no good for real-time audio and video, which is why it is not included in ATM/MPLS and IP

Sliding Window: Flow Control (2)



Connection Management



A Complete Example: Web Access

- The protocols used and the functions exercised are the following
- The browser on a WLAN laptop uses the following protocols:
 - HTTP, which first asks to establish a TCP connection to the remote server at port 80 and then sends over it an HTTP GET request for the web page
 - TCP, which sends the HTTP GET packet over IP, and expects an ACK back from the remote TCP at the server (error control)
 - LLC over WLAN 802.11 to the local Access Point (AP) – 802.11 expects an ACK from the AP for the single frame (error control)
 - Later when receiving a stream of response packets from the local AP, the AP 802.11 performs error/flow control to the laptop so frames are not lost
- The server on an Ethernet computer uses the following protocols:
 - HTTP, which retrieves and passes down the page as a single HTTP packet
 - TCP, which segments the HTTP packet to as many 1500 byte packets (Ethernet size) required, then sends them over IP using ARQ error and flow control, expecting ACKs and resending selectively if packets are lost
 - LLC over Ethernet 802.3 to the local router – the fixed Ethernet 802.3 does not provide error and flow control

Datagram and Virtual Circuit Network Layer Organisation

- There exist two major different approaches for organising the network layer of a packet switched network.
- Datagram or Connectionless (CL) approach:
 - Each packet entering the network is treated separately, with no relationship to previous packets; packets must contain the destination address and may be routed over different paths in different path
 - IP is the key datagram CL protocol
- Virtual Circuit (VC) or Connection-Oriented (CO) Approach:
 - A virtual circuit (i.e. a fixed path) is first established between source and destination over which all subsequent packets will be sent and received; each packet only contains a path identifier in the same path
 - X.25, ATM and MPLS are all virtual-circuit CO protocols

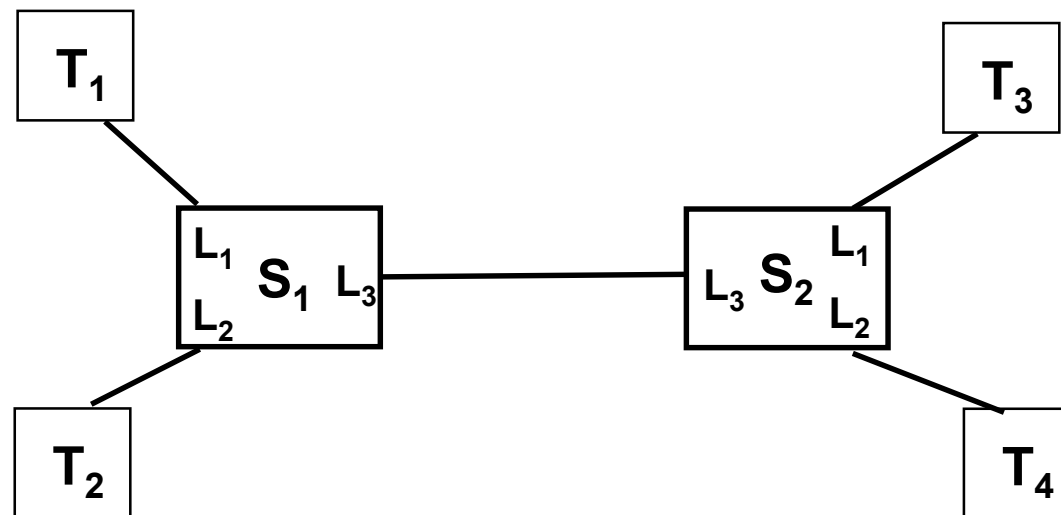
Datagram-based Operation CL

- No VC setup is required before communicating.
- Every packet is treated individually and a routing decision is made according to its destination address and the routing tables
- Packets may follow different routes to the destination if routing information changes while two applications are communicating.
- If a router or a forwarding interface in the communication path fails, packet transfer will simply suffer for a bit, e.g. 1-2 secs, until routes “re-converge”, and packets will then follow the new route.
- Advantages: simplicity, robustness against failures
- Disadvantages: high processing for address lookup => reduced forwarding speeds, packets may be delivered out of sequence

VC-based Operation CO

- Routing decisions are made only in the VC setup phase
- As a VC is setup through special signalling packets, an entry is created for every switch the VC crosses; these entries form the switching table which is different from the routing table switching table include routing table and status
- Switching tables mean additional state information => complexity
- Data packets carry only a VC identifier (and not an address)
- All packets follow exactly the VC established route, so they are guaranteed to be received in sequence
- If a switch, or simply a forwarding interface, fails, all the VCs which cross them will become non-operational; in this case, alternative VCs which follow a different route need to be setup
- Advantages: high speeds, packets delivered in sequence
- Disadvantages: switch complexity / additional state information, more difficult to deal with failures, VC establishment delay

Switching Example: the Routing Tables



S1 Routing Table

T1 -> L1
T2 -> L2
T3 -> L3
T4 -> L3

S2 Routing Table

T3 -> L1
T4 -> L2
T1 -> L3
T2 -> L3

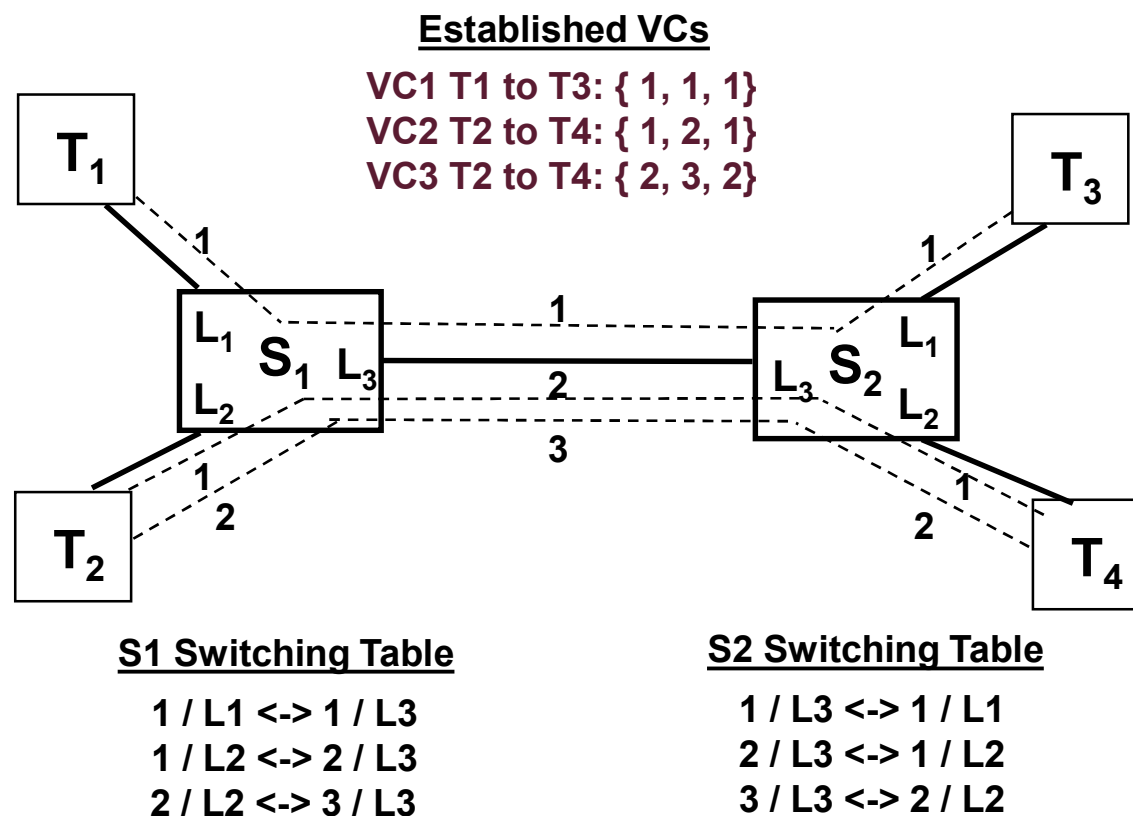
S: Switch
T: Terminal
L: Link

Note: we show routing table entries per terminal for simplicity,
in the real world entries are kept per network for scalability

VC Switching Principles

- The switching tables are populated as VCs are setup; in every link they VC crosses, a Virtual Circuit Identifier (VCI) is assigned to it (see example on the next slide)
- The VCI value chosen for every VC crossing a particular link has to be different for all VCs over that link
 - A typical approach is to start from 1 and increase monotonically for every new VC until the highest number is reached, then wrap around
 - In the example we assume that the three VCs are the very first ones in the network, hence the values {1, 1, 1}, {1, 2, 1} and {2, 3, 2}
- The VCI/link pair of the incoming packet is used as the “key” to the table to locate the VC cross-connection entry: the switch simply changes the VCI and forwards the packet to the XC outgoing link

Established VCs and the Switching Tables



Example packet from T2 to T4 on VC3:
 2/L2 maps to 3/L3 in S1, then 3/L3 maps to 2/L2 in S2
 And exactly the reverse for the response(s) from T4 to T2

X.25

- X.25 was the first ISO/ITU-T network technology, it was developed in the 1970s when links were analog with relatively high BER
- Link layer: uses the X.25 Link Access Protocol Balanced (LAPB)
 - ARQ protocol with error/flow control per link
 - The SS7 MTP2 is very similar to LAPB
- Network layer: uses the X.25 Packet Layer Protocol (PLP)
 - Includes VC setup, data transfer and tear-down
 - ARQ protocol that provides error and flow control end-to-end per VC,
- Maximum packet size is 128 octets/bytes
- X.25 cannot achieve high speeds due to ARQ functionality, running both per every link and also end-to-end per every VC
 - Also not suitable for real-time audio and video
- Now legacy technology, but found good use in the past (and is still being used) in banking, credit card verification and similar services

X.25 Reliable Datagram Service Emulation

8 bytes head

- Given that data \leq 120 bytes (payload) may be passed in the VC connect request packet, the called party may also respond immediately with a disconnect request, passing back data \leq 120 bytes.
 - can be used to emulate a reliable datagram service, useful for services such as credit card verification.
- X.25 CONNECT(data-request)
 - e.g. data can be card number & transaction amount
- X.25 DISCONNECT(data-response)
 - e.g. transaction can be accepted or rejected

Signalling System 7

- Signalling System 7 (SS7) – a common channel signalling system that constitutes the “nervous system” of the PSTN.
 - Also referred to as Common Channel Signalling 7 (CCS7).
- Developed from SS6 which was introduced by AT&T in the 1970's, SS7 started being deployed in the early 1980's and soon became the common signalling standard used today.
- SS7 supports a messaging architecture for call establishment, billing, information exchange and the Intelligent Network (IN).
- It is a packet network whose protocols conform to a layered architecture influenced from the OSI-RM, which was also being developed around the same time (late 70's).

SS7 Message Transfer Parts 2 and 3

- MTP 2 (link layer) is very similar to the X.25 LAPB.
 - The difference is that when there is no traffic, special control frames are transmitted so that a faulty link or node is immediately detected.
 - Also a count of frames received in error is kept and when a threshold on the error rate is reached, the receiving end puts the link out-of-service.
 - These differences relate to the SS7 requirements for very high availability, reliability and performance.
- MTP 3 (network layer) provides a reliable datagram service, a unique combination among all protocol architectures
 - Signalling packets are acknowledged and retransmitted if lost (rare)
 - Routing uses multiple S-D routes for load balancing
 - The PSTN signalling network is separate from the voice network and is carefully provisioned so there is almost never congestion / packets lost

Asynchronous Transfer Mode (ATM)

- Asynchronous Transfer Mode (ATM) was developed in the 1990s and was the proposed answer by the ITU-T to the needs of future multiservice networks.
- The key targets were to support integrated voice/video and data (replacing/unifying the PSTN and X.25), support very high speeds and provide Quality of Service (QoS) guarantees
- The salient features of ATM are the following:
 - Fixed small size packets (called cells) of 53 bytes are used; this allows to interleave traffic in a highly flexible fashion to meet QoS requirements.
 - No error or flow control are provided to achieve high speeds and also support the transmission of real time audio and video.
 - Virtual Connections (VCs) are setup by applications through signalling with specific QoS characteristics (bandwidth, delay and packet loss).

ATM (cont'd)

- ATM was ambitious, but also complex and expensive technology, with its deployment limited to core telecom networks and never reaching the enterprise and home socket as originally envisaged.
 - This is also partly due to the dominance of IP after the mid-1990s.
- But ATM has had a lasting influence on today's technologies:
 - Digital Subscriber Loop (DSL) technologies were designed based on underlying ATM principles
 - IP Integrated Services (IntServ) and Differentiated Services (DiffServ) were influenced by ATM, aiming to support QoS in IP
 - MPLS (see next) is based on ATM switching and QoS principles but switches IP packets instead of ATM cells; as such, it can be thought as “ATM in IP clothes”.

Multi-Protocol Label Switching (MPLS)

- MPLS was developed in the late 1990s when it was clear that IP had become the dominant packet networking technology and it attempts to bring some of the ATM functionality in IP
- It brings connection-oriented forwarding “just underneath IP” and because of this it is sometimes referred to as a “layer 2.5” protocol.
 - It does not fit well the layered reference model
- Its key features are the following:
 - Equivalent VCs are known as Label Switched Paths (LSPs) as each packet carries a “Label” which is equivalent to the VCI
 - There is no error and flow control in order to achieve high speeds and support real-time audio and video
 - LSPs can be setup with specific QoS characteristics (bandwidth, delay, packet loss), in a similar manner to ATM VCs
- MPLS achieves fast forwarding speeds because of “label switching” as opposed to IP address-based routing and no error/flow control

MPLS (cont'd)

- A key difference with ATM is that LSPs are not established dynamically by applications but are pre-established by ISPs edge-to-edge i.e. from ingress to egress network nodes
 - Multiple LSPs are established for every ingress-to-egress pair in order to balance traffic (“traffic engineering”) and achieve a good grade of service i.e. good network performance
 - This circumvents the limitations of IP shortest path routing, in which all traffic ingress-to-egress follows a single shortest path
- Most big ISPs today (Tier-1 and Tier-2) use MPLS in their networks for high speeds and for traffic engineering; all ingress-to-egress traffic is switched via MPLS instead of being routed via IP
 - This means that IP over MPLS operates effectively in a CO manner within big ISP domains
- MPLS and ISP network traffic engineering are covered in detail in the Network and Services Management (NSM) module

Internet Protocol (IP)

- IP is a connectionless unreliable datagram protocol that was designed from the beginning to operate over disparate network types; simple, easy to implement and deploy
- Variable packet size depending on link layer technology, maximum packet size is 1500 bytes dictated by Ethernet LANs
- Error and flow control is left to the transport/application layers
 - Network is kept simple, “the end-to-end design principle” according to which complexity should be confined to the end systems
- There have been extensions to support QoS (Differentiated Services), security (IPSec/TLS), mobility (Mobile IP), etc.

IP (cont'd)

- In WANs, IP is used either directly over PPP or over MPLS over PPP, hence no error and flow control
- In Ethernet LANs, IP is used over LLC over Ethernet 802.3, again without error and flow control
- In Wireless LANs, IP is used over LLC over WLAN 802.11, the latter provides error and flow control
- The Transmission Control Protocol (TCP) provides a CO reliable service over IP with error and flow control
- The User Datagram Protocol (UDP) provides a CL unreliable service over IP
- Note that most big ISPs (Tier-1 and Tier-2) use IP over MPLS
 - This means connection-oriented operation in those domains

Link and Network Layer Combinations in the Presented Protocol Architectures

	X.25	SS7	MPLS	IP
Network Layer	VC reliable (X.25 PLP)	datagram reliable (MTP3)	VC unreliable	datagram unreliable
Link Layer	reliable (LAPB)	reliable (MTP2)	unreliable (PPP)	assumed unreliable (mostly PPP)

- Older data-oriented architectures (X.25, SS7) include reliability low down in the protocol stack i.e. in the link and network layer
 - Lower speeds, not suitable for real-time streaming
- Newer media-oriented architectures (MPLS, IP) leave packet reliability to the transport and application layers
 - Can achieve high speeds, especially VC-based ones such as MPLS

Summary

- Discussed fundamental principles of packet networks
 - The layered reference model and layering principles
 - Packet reliability through ARQ / sliding window protocols
 - Network layer organisation through VC and datagram approaches
- Examined older data-oriented technologies such as X.25 and SS7 which include reliability lower down in the protocol stack
 - i.e. in the link and network layer
- Examined subsequent technologies such as ATM, MPLS, IP, which leave reliability to the end systems
 - i.e. to the transport and application layer
- Gained a good understanding of the fundamental operating principles of packet network technologies