The Open Systems Interconnection (OSI) reference model can be used to describe layered networked systems. What are the seven layers of the OSI reference model, in order from lowest to highest? [4]

The lower three layers of the OSI reference model operate in a hop-by-hop manner, while the upper four layers operate end-to-end. With references to the roles of the different layers, explain what is meant by this, and why the distinction arises. [9]

Identify and briefly describe the key functions of any three layers in the OSI reference model. Name an exemplar protocol in each layer; expand any acronyms in your answer. [6]

What is encapsulation and why is it commonly carried out in data networks? [3]

Networked systems often follow a layered architecture. This is useful for explaining how a networked system works and is helpful when writing protocol specifications. Discuss whether it’s also a good way of implementing a networked system, or whether an alternative software architecture would be more suitable. Outline the trade-offs involved in developing software to implement a network protocol stack, giving examples of possible design decisions and their impact if appropriate. [8]

A common form of wired connection is an unshielded twisted pair. Briefly state 1) why the two wires that form the twisted pair are twisted together, and 2) why a typical cable includes more than one twisted pair.

When using wireless links, the signal to be transmitted is modulated onto a carrier wave, rather than using a baseband encoding. With the aid of a diagram, demonstrate the operation of amplitude, frequency, and phase modulation.

Briefly state what is the Baud rate of a link. Explain what it means to say a link can send more than one bit per Baud. Give an example of how a wireless signal can be modulated so that two bits can be sent per Baud.

Manchester encoding is a commonly used baseband encoding scheme for wired links. Describe how the binary digits 1 and 0 are encoded when using Manchester encoding. With the aid of a diagram, show how the sequence 0010 is encoded in Manchester encoding. [5]

There are several different ways in which a signal may be modulated onto a carrier wave. With the aid of a diagram, briefly explain the operation of one such modulation scheme. [3]

There are also several different encoding schemes used for baseband signals, for example NRZ, NRZI, and Manchester encodings. Draw a diagram to show how the binary signal 0010111101000010 would be transmitted using each of these encoding schemes. Two problems are avoided by using Manchester encoding instead of NRZ or NRZI encoding; describe these problems and explain why they don’t occur in Manchester encoding. [7]

A key function of the data link layer in the OSI reference model is framing: converting the raw bit-stream provided by the physical layer into a structured communication channel which can be used by the upper layers. To determine the start of a frame, many data link layer implementations use a start code comprising the binary pattern 01111110. It is essential that the start code only appear at the beginning of a data link layer frame, and not within the data. Discuss why this limitation exists and explain how the data link layer can prevent the start code occurring within the data, while still allowing arbitrary data to be transferred. [8]

Before data link layer frames can be used, it is necessary to check those frames for errors that may have occurred during transmission. This check can be done in a number of different ways, for example using a parity code, a checksum, or a cyclic redundancy check (CRC), depending on the likelihood of errors. Discuss what physical processes might cause transmission errors, and how these might affect the bit stream. [2]

Explain what a parity code is, how it can be calculated, and how it can be used to detect transmission errors. [4] (Parity code / Parity bit / check bit)

What is framing and how it is performed? [4]

Outline the advantages and disadvantages of using a parity code to detect transmission errors, compared to using a checksum or a CRC. [2]

The Internet Protocol provides a connectionless, best effort, packet delivery service at the network layer of the protocol stack. Explain what is meant by best effort in this context and discuss how this aspect of the delivery service impacts the transport layer. [4]

One of the key roles of the network layer is inter-networking between different link layer technologies, to allow local area networks to be combined to form a seamless wide area network. Each link layer technology may have a different maximum transmission unit (MTU), and the network layer is responsible for handling this mismatch. Describe how IPv4 and IPv6 handle inter-networking between links with different MTUs. Discuss why this behaviour was changed when IPv6 was designed. [10]

At the network layer, the Internet is a connectionless packet network, providing a best-effort packet delivery service. By way of contrast, the traditional telephone network provides a reliable circuit switched service. Discuss what are the advantages and disadvantages of these two approaches to network design. [6]

Many of the key design decisions of the Internet were influenced by the end-to-end argument. With the aid of an example, briefly describe the end-to-end argument as it applies to network protocol design. [4]

Since hosts on the Internet typically have both an IP address and MAC address, it may be tempting to assume the nature of both addresses must be very similar. Briefly outline the salient differences between the IP address and MAC address of a typical host on the Internet. Your answer should consider factors such as the relative address sizes and structures, how the addresses are allocated, organisations responsible for developing the relevant standards, layers in which the addresses are typically used and the way both addresses are used to switch/route data to a host. [7]

A routing algorithm is used to find the best network layer path between two hosts that are not directly connected. Two types of routing algorithm are used: intradomain routing and inter-domain routing. State what is the difference between these two types of routing and describe in what environment each type of algorithm would be used. [4]

The Border Gateway Protocol (BGP) is used for inter-domain routing in the Internet. A BGP router builds its routing tables based on exchange of Autonomous System (AS)-path vectors giving routes to destination IP address prefixes. The routing information exchanged is often filtered to enforce policy, with the GaoRexford filtering rules being widely used. Describe the Gao-Rexford rules, and explain why they are desirable. [5]

A distance vector algorithm is sometimes used for intra-domain routing. Describe what information is stored in the routing tables at each host when using distance vector routing. [4]

Consider the network graph pictured below. Nodes represent devices on the network, edges represent links and the numbers by the edges represent the cost of forwarding a message across that link (note: in the lecture example, these weights were assumed to all be 1). Assume that the network is using a Distance Vector protocol, that all message exchanges happen at the beginning of every round of the protocol, and that hosts update their state after they’ve received all messages destined to them in the current round. Show the initial routing state of all nodes in the network, and their routing state after every round of the protocol. In your answer please show the routing table entries (distance/next hop) of all nodes in the network (i.e., a 5 x 5 matrix), for each iteration of the algorithm. Your answer should include 1 such matrix per iteration of the algorithm, including the one for the original state (i.e., nodes only know of themselves and their 1-hop neighbours). [12]

Assuming all of the information from the network graph above has been gathered at one host, show the steps of Dijkstra’s algorithm for this host; make sure to outline the currently selected node and the changes, if any, to the best route to other hosts. [10]

The link-state routing algorithm is a popular alternative to the distance vector algorithm. Considering the information stored in the routing tables, the complexity of the routing algorithm, and the convergence time, outline the advantages and disadvantages of link-state routing when compared to distance vector routing. [6]

A major problem with distance vector routing is the count to infinity problem. Discuss what is this problem, and how it can be mitigated. [6]

Many Internet service providers use the Open Shortest Path First (OSPF) routing protocol for intra-domain routing. OSPF is an example of the class of link-state routing protocols. At a conceptual level, describe how link state routing works. [4]

The alternative to link-state intra-domain routing is distance vector routing. Briefly outline the advantages and disadvantages of link-state routing when compared with distance vector routing. [3]

The two most widely used transport protocols in the Internet are TCP and UDP. Describe the service models provided by these two protocols, being sure to highlight their main differences. [8]

Networked applications often use the Berkeley Sockets API. This API provides two different sets of functions that can be used to send/receive data. TCP-based applications call the send() and recv() functions, while UDP-based applications typically use the sendto() and recvfrom() calls. Discuss why these two protocols use different calls to send and receive data. State what extra information is passed across the API when using UDP, in both sending and receiving cases. [6]

Applications that use UDP often include some form of sequence number as part of the application data they send. With reference to the behaviour of UDP, and the underlying IP layer, explain why this is needed. [6]

A commonly used application that runs over UDP is the Domain Name System (DNS). Explain why DNS was written to use UDP, and how it ensures reliability. Discuss whether you think UDP was an appropriate choice of transport. [8]

TCP senders implement congestion control by using an additive increase multiplicative decrease (AIMD) algorithm to vary the sliding window size during the congestion avoidance phase of a connection. With reference to the AIMD parameters a and b, describe how the AIMD algorithm works. Explain why an AIMD algorithm with these parameters was chosen for TCP congestion control. [5]

You are implementing an interactive video conferencing application, to run over an IP-based network. You can implement this application using either TCP or UDP at the transport layer. Which of these two transport layer protocols would you choose? Discuss the advantages and disadvantages of the two approaches and outline any cases where the other transport layer protocol would be more appropriate. [12]

When making a TCP connection, it is necessary to specify both an IP address and a TCP port number. Describe the purpose of these two parameters and explain why both are needed. [3]

A TCP connection offers a byte stream abstraction that delivers data in order but doesn’t preserve record boundaries. This complicates protocol implementations, since they must explicitly scan the data read from a TCP socket to find the end of the protocol data unit (PDU) that they are processing (e.g., when reading an HTTP request, an implementation must call read() in a loop, accumulating data into a buffer, until it sees the blank line signifying the end of request). With reference to the packet-level dynamics of a TCP connection, explain why record boundaries are not preserved. [6]

TCP congestion control relies on “ACK clocking” to determine when the sender can transmit new packets. Describe what information is contained in TCP ACKs. [2]

Explain what is ACK clocking, and how it helps prevent network congestion.[4]

The Transmission Control Protocol (TCP) is designed to provide end-to-end reliable delivery of data across an unreliable Internet Protocol (IP) network. To achieve this, each transmitted data packet contains a sequence number, and the receiver sends cumulative positive acknowledgements for packets as they are received. The sender uses these acknowledgements to decide which packets have been lost and should be retransmitted. Using diagrams to show transmission and retransmission of data packets and acknowledgements, illustrate how TCP behaves when:

1. a single packet is lost in transit;
2. when two consecutive packets are reordered in transit
3. Describe how TCP distinguishes these events, based on the acknowledgements. [7]

You have written a basic server program in the first assessed exercise for this course, using the Python programming language, and the Berkeley Sockets API. Outline the design of your server, explaining the high-level structure of your code when serving a “get” request, and highlighting how it uses the functions in the Berkeley Sockets API. [12]

When transmitting data from one Internet host to another over a TCP/IP connection, that data is generally protected by checksums and error correcting codes at the TCP, IP, and data-link layers. Explain why the checksum at the TCP layer is necessary, given that data is also protected by checksums at the IP- and link-layers. [2]

Many network protocols send binary data directly, rather than encoding it in some textual format. This binary data is usually sent in network byte order. Explain what network byte-ordering is and discuss why it is needed. [4]

Other network protocols are structured using a textual encoding, and indeed many older protocols are defined using the US ASCII character set. Such protocols cannot directly transfer binary data, forcing the binary content to be encoded using, for example, base-64 before it can be transmitted. Explain why these protocols do not support binary data, then describe (at a high level) how base-64 encoding works and solves this problem. [9]

Many applications use the Domain Name System (DNS) to map between host names and IP addresses. Imagine some catastrophic failure happened, so that the DNS root name servers all failed simultaneously and stopped answering queries. Discuss how the effects of such a total DNS failure would manifest themselves, and how quickly they would become visible. [5]

Explain what effects this failure would have on applications using the Internet. [3]

Traditionally, top-level DNS names have been limited to choose ending in two level country code domains (e.g., “.uk” or “.sg”) or a small number of generic top-level domains (e.g., “.com”, “.org”). More recently, many more domains have been introduced, including a mixture of internationalised domain names (i.e., using nonASCII characters), and large numbers of new generic top-level domains (e.g., “.clothing”, “.coffee”). Discuss whether you think this expansion is a good idea from a social and business perspective, justifying your answer. [6]

Consider the technical impact of the changes outlined in the above question and describe how they will affect the operation of the DNS, and the load on various servers that make up the DNS infrastructure. [6]

The DNS has been described by some an essential component of the Internet, that is required for its correct operation. Others claim that it’s just one application out of the many that run on the Internet. Which viewpoint is correct? Discuss. [5]

Domain names are organised in a hierarchy, derived from a single root zone. Explain why this hierarchy is necessary to effectively manage and scale the DNS. Why is there only a single DNS root zone? What problems would occur if there were multiple root zones? [8]

Network Address Translation (NAT) devices are widely used in the Internet. Describe the purpose of a NAT and give four reasons why NAT devices are used in the Internet. [6]

Home networks often connect to the wider Internet using a network address translation (NAT) router, rather than using a standard IPv4 router. Explain how addressing and routing for a sub-network were intended to work in the Internet architecture, and how they work in a NAT-based network. [10]

Describe the benefits that come from using network address translation in an IPv4 based network. [3]

Describe the problems that are caused by the use of network address translation in an IPv4-based network. [5]

Could it make sense to use network address translation in an IPv6 network? Justify your answer. [2]

To ensure the confidentiality of data sent across a network, it is important to encrypt that data before transmission. Encryption can be performed using either symmetric cryptography or public-key cryptography. Outline the differences in the number of keys, and how the keys are used, between these two approaches [4]

Explain what important problem is solved by public-key cryptography that is not solved by symmetric cryptography. Briefly explain why public-key cryptography solves this problem. [3]

The Transport Layer Security (TLS) protocol, which is used to secure HTTP connections, uses a mixture of both symmetric and public-key cryptography. Explain why this is done, and how it ensures both security and good performance.[5]

To subvert encrypted communication, an eavesdropper might perform a man-in-the-middle attack on a connection, to try to fool the sender into thinking they are the legitimate receiver. Man-in-the-middle attacks can be detected by sending a digital signature along with the data, provided the receiver checks the signature. Explain the basic steps that are involved in generating and checking a digital signature, and how the receiver can use the digital signature to detect a man-in-the-middle attack. [8]

A frequently reported security problem with the Internet is the distributed denial of service (DDoS) attack, where many attackers conspire to send unwanted traffic to the victim at the same time. This massively overloads the victim’s network connection, making it difficult, if not impossible, for them to conduct their usual business. In the long-term, such attacks are resolved by the actions of the law enforcement officials. Discuss how, in the short-term, the victim and its Internet Service Provider (ISP) can cooperate to reduce the effects of a DDoS attack? [5]