Chapter 8. Internetworking and the Internet

Table of Contents

Introduction to Internetworking and the Internet

Context

This unit follows on from units 6 and 7 in continuing the treatment of the technology of networks.

Introduction

Any internet, including the Internet, is a collection of networks linked together in some way that allows them to work together. The way in which routers are used to create the Internet and are then used to route messages through that it are explained, thereby expanding the view presented in the previous unit. The form of addressing used by the Internet is examined. Essentially, the address of an Internet computer has two parts: its own (computer) address and the address of the network to which it is attached. Once a communication reaches the network containing the computer to which it is addressed, it can be delivered using just the computer address. When it is in another network, it can be directed to the network containing the destination computer using the network address. In this way, despite the fact that diverse networks may have been linked together with routers to form an internet, messages can be routed across the resulting internetwork in a uniform fashion.

The stack of layers of the communications software is completed by adding the Application Layer, which possesses the distributed applications that generate Internet traffic. The precise correspondence of the Internet's software, known as TCP/IP, and its TCP and IP layers, with the layers introduced in the previous units is explained. The applications possessed by the Application layer are briefly introduced and examined in terms of both their operation and their usage. These applications include electronic mail, file transfer and Telnet.

Objectives

At the end of this module, you should be able to:

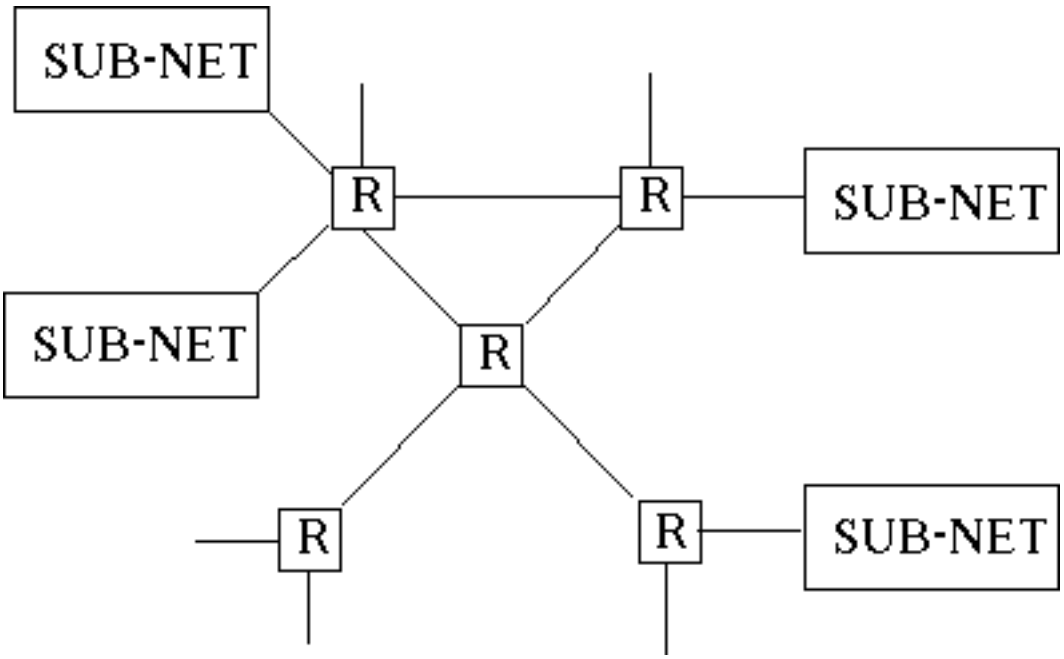
- discuss the role of routers in the creation of an internetwork, and in routing packets through an internetwork;
- explain Internet addressing, and its roles within the Internet;
- appreciate the role of the Application layer, and know some of the applications it provides;
- understand the layered structure of Internet software;
- explain how applications communicate across the Internet.

Content

In parallel with this unit, you should read the relevant parts of your textbooks.

Introduction

We can give a more detailed description of the structure of the Internet than that presented at the end of the previous unit by viewing it as a collection of networks. The Internet is created by linking existing networks (the sub-nets) with routers (R). This can be represented by:



In this way, any existing network can be linked into the Internet by means of a router. The interconnected routers, some of which have no sub-nets attached directly to them, in themselves create a network. This router network can be regarded as a backbone for the Internet.

When packets are routed across the Internet, they must negotiate both sub-nets and the back-bone network of routers. To understand how this operates, we must first understand Internet addressing.

Internet addresses

Every Internet computer is assigned an Internet address (or IP address). The addresses are 32 bits long, and are used by Internet packets as source addresses and destination addresses. To make them slightly more readable than a sequence of thirty-two 0s and 1s, they are usually written as:

158.94.39.1

(When each of the numbers is converted to its 8-bit binary representation, the four 8-bit sequences give the 32-bit address.)

The 32-bit sequences are formatted to indicate three things: The type of the sub-network, the address of the sub-network, and the address of the computer on that sub-network. There are three types of sub-net: large, medium-sized and small. This gives the following three types of IP address:

Type	Network address	Computer address
0	< 7 bits >	< 24 bits >

Type	Network address	Computer address
10	< 14 bits >	< 16 bits >

Type	Network address	Computer address
110	< 21 bits >	< 8 bits >

To Do

Do Review Questions 1, 2, 3 and 4.

Packets and Routing

Packets have to be routed through the Internet whenever it is used. The elements of a packet basic to this activity are:

Source Address	Destination Address	Data	Parity

Since the addresses have two parts, a network address (Net Ad) and a computer address (Comp Ad), this can be elaborated as:

Net Ad	Comp Ad		Net Ad	Comp Ad		data	parity

A packet to be sent, for example, from computer 3 on network 22 (address 22: 3) to computer 12 on network 7 (address (7:12) takes the form:

22: 3	7: 12	information	parity

The form of the addresses shows that it is possible to determine automatically whether a packet is being sent to a computer in the same network or to a computer in another network.

To Do

Do Review Question 5.

Routing in the Internet

When a packet is routed through the Internet, there are two situations to be covered. They are:

1. When the destination computer is in the same sub-net as the source computer.

   The computer parts of the source and destination addresses will be the same. In this case, routing can proceed with the use of routing tables, as described in unit 8. Note that once it has been determined that the network parts of the two addresses are the same, only the computer part of the destination address is needed for routing.

2. When the destination computer is in a different sub-net from the source computer.

   In this case, routing is a three-stage process.

   Stage 1 Packets are routed through the sub-net of the sending computer to its router, as in case 1.

   Stage 2 Packets are routed through the backbone network to the router of the sub-net containing the destination computer. At this stage, when the network parts of the two addresses are different, only the network part of the destination address is needed for routing. It is done with routing tables that take the form:

Destination Network	**Send to Router**
101	R1
102	R2

Each router has a table of this kind, so that packets can be routed through the backbone network from router to router by the familiar store-and-forward procedure.

Stage 3 At the final router, packets have reached the destination sub-net, and they are routed through it to the destination computer as in case 1.

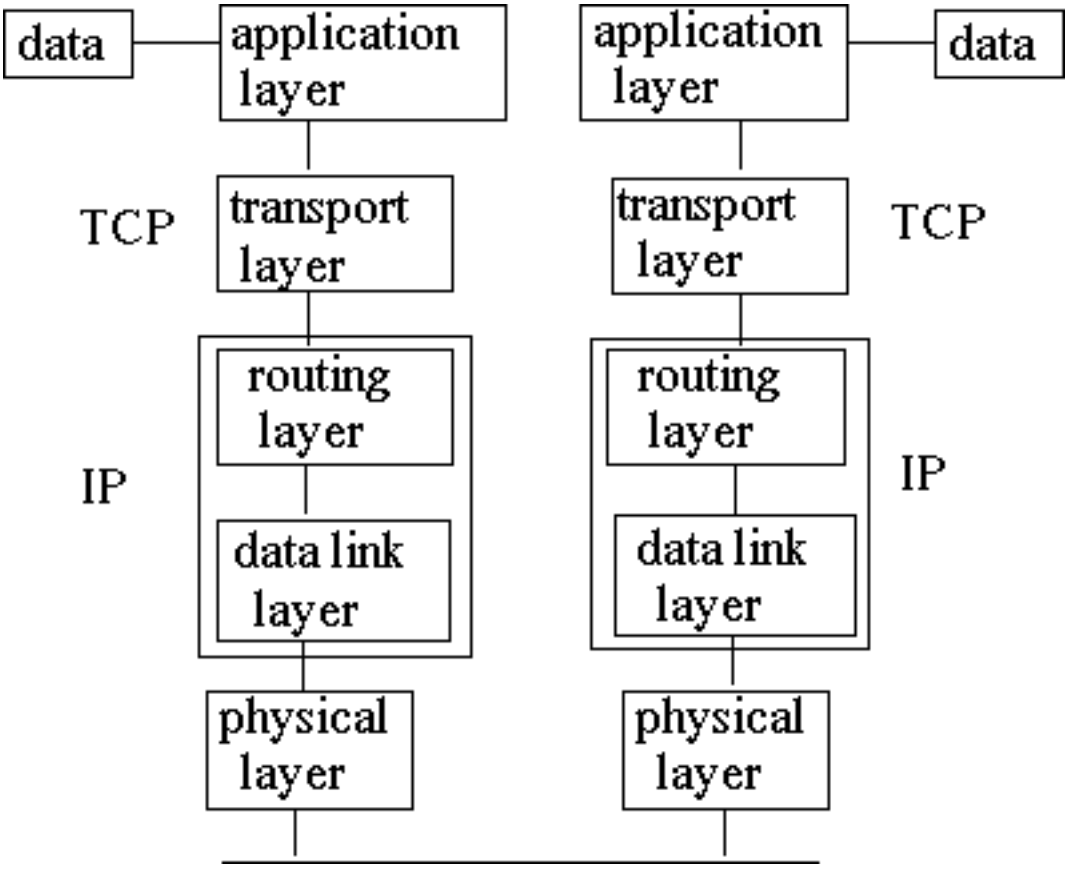
Note the underlying assumption here that the router by which a sub-network is linked to the Internet is a part of the sub-network. This level of treatment allows us to proceed without the detail need to introduce 'gateways'.

To Do

Carry out Activity 1.

Communications Software

The complete organisation of the Internet's communications software can be shown, as below. It has a top layer, the application layer, that provides applications such as e-mail and FTP. These applications are the source of the messages and files that are presented to the transport layer to be divided into packets to be routed correctly to another Internet computer. The transport layer, to repeat, is known as TCP. It passes packets to IP, which determines the next stage of the route and deals with error handling. As such, the routing and data link layers more or less correspond to IP. The destination computer, of course, reverses the process to receive the message or file that was sent to it. Its IP layer is responsible for the receipt of correct packets, which are passed to TCP. The transport layer assembles the message or file, and passes it to the Application layer and so to the appropriate application.



To Do

Do Review Question 6.

Internet Applications

Having introduced the Application layer, in this section we give brief descriptions of two of the Internet's most widely used applications. They illustrate the ways in which Internet applications generate messages to be delivered by the Internet.

E-mail

An e-mail application provides facilities for:

- creating an electronic envelope,
- composing a document, and
- adding attachments.

The envelope has slots for the:

- the destination address,
- subject,
- copies to, etc.

The form of an e-mail address is:

- Mailbox@Computer, for example:
- mit@cs.uct.ac.za

When composed, the message is submitted to TCP which divides it into packets and submits the packets to a mail submission server, where the Computer part of the e-mail address is converted to an IP address before the packets are launched into the network. At the destination mail server, the packets are reassembled to recover the message that is then stored in the recipient's mail box.

To Do

Carry out Activity 2.

World Wide Web

By running a browser, a computer is able to retrieve documents held by any World Wide Web server on the Internet. The form of a request is:

http://Computer/Document,

for example:

http://www.cs.uct.ac.za/mit/index.html

When a request is submitted, the http protocol is used to contact the Computer, and, if this is successful, to relay the request for the Document. On receiving the request, the computer (a Web server) retrieves the Document and sends it to the requesting computer where it is displayed. (The document will be held and transferred as an HTML file which the browser can interpret and display.)

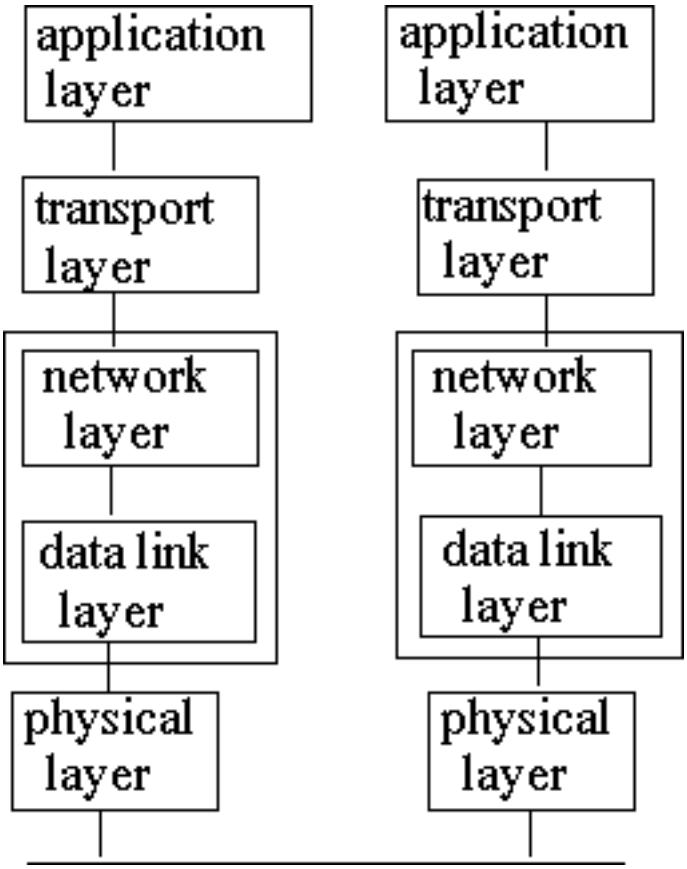
To Do

Do Review Questions 7, 8 and 9.

An Internet Example

An example may help to show how all this fits together. This section sketches what happens when a file is transferred over the Internet. Much the same thing happens whether this is done using FTP (the File Transfer Protocol) or by using the Web to retrieve a file from a Web server.

The file transfer capability is possessed by the Application layer. When the application is run, the file and the computer holding it can be identified from the user's input. The application then generates a request addressed to that computer for that file. The request is passed to the Transport layer (TCP) which adds its own header to it identifying the application as a file transfer and, in turn, passes it to IP which (assuming the request is nothing like long enough to fill even one packet to capacity) makes it the content of a routable packet. The bits of this packet are passed to the Physical layer which signals them on the first stage of their journey. Intermediate computers store and forward the packet until it reaches the computer holding the file. When the packet has been delivered and the request it contains passed to the receiving application, the requested file is located and passed to the Transport layer. TCP splits the file into segments, adds its own header to each, and passes the result to the next layer, IP. This layer formats each chunk passed to it as an Internet packet, and passes each packet to the Physical layer to be signalled over the Internet.



The packets are routed across the Internet to their destination using the two-stage routing procedure described above.

When a packet is received at the destination computer, the computer at which the file was requested, it is passed to IP, which strips off its header, and passes what is left to TCP. TCP strips off its header, leaving just a segment of the file. When all the packets have been received, TCP reconstructs the file and passes it to the file transfer application so that the file is now available to the user.

Although this explanation is simplified, it catches the essence of the communication that occurs when a distributed application runs. Other distributed applications, including electronic mail, generate traffic on the network in much the same way.

The explanation shows how all the elements of the architecture of the network come into play. The hardware is the substrate for the activity, and is static. The protocol, by exercising the rules that govern communication and providing a format for the items communicated, allows dynamic activity to occur on the static support.

To Do

Do Review Question 10.

Carry out Activitities 3 and 4.

The Internet as a Data Delivery Network

In the light of the descriptions given above, we can see the Internet as a network capable of delivering data in a standard fashion regardless of which of its applications generated the data in the first place. The network, in moving the data between computers has no knowledge of (and no need to know) what the data represent. E-mails, image files, Web pages and digitised music are all treated in exactly the same fashion, that is, as data.

Activities>

Activity 1 - Routing in the Internet.

Using the diagram at the beginning of this unit showing the structure of the Internet, replace two of the sub-net blocks by a small network of computers. Assign addresses to the networks, the computers in them, and the routers between them. Select a source computer in one sub-net and a destination computer in the other. Select a route between them. Devise routing tables for the computers and routers along the route. Finally, trace the way in which a packet is routed between the computers and through the Internet.

You can find a discussion of this activity at the end of the chapter.

Activity 2 - Domains

Find out what is meant by a 'domain'. Identify the domains in an e-mail address such as : mit@cs.uct.ac.za.

Sketch the tree-structure of e-mail addresses, starting with a country domain, which branches to academic, company, organisation domains, and so on, each of which then branches further. Starting at any 'leaf' of this tree, and reading to the 'root' will give a complete e-mail address.

You can find a discussion of this activity at the end of the chapter.

Activity 3 - Electronic Mail

Using your textbooks and other references, find out how it is that all Internet computers, regardless of their type, the character codes they use, the type of e-mail program they run and so on, can exchange e-mail with each other.

Activity 4 - TCP and IP Details

Using your textbooks and other references, find out what the purpose and use of the following fields that occur in either TCP or IP frames:

1. Piggyback acknowledgement,

2. Time to live,

3. Don't fragment, and
4. Urgent pointer.

Review Questions

Review Question 1

To what type of network does the computer with IP address 158.94.39.1 belong?

What is the network address, and what is the computer address?

You can find an answer/comment for this review question at the end of the chapter.

Review Question 2

How many large, medium and small sub-networks can there be on the Internet?

You can find an answer/comment for this review question at the end of the chapter.

Review Question 3

How big are large, medium and small networks?

You can find an answer/comment for this review question at the end of the chapter.

Review Question 4

What happens when the Internet has 128 large networks?

You can find an answer/comment for this review question at the end of the chapter.

Review Question 5

How is it possible to determine automatically whether a packet is being sent to a computer in the same network or to a computer in another network?

You can find an answer/comment for this review question at the end of the chapter.

Review Question 6

On the diagram at the beginning of the previous unit, showing the idea of packet switching, indicate the protocol layers associated with the message, packet and network elements.

You can find an answer/comment for this review question at the end of the chapter.

Review Question 7

Would it be difficult to convert a World Wide Web query (http://www.jargon.etc) to a comprehensible query?

You can find an answer/comment for this review question at the end of the chapter.

Review Question 8

Which applications, other than those dealt with above, reside in the Internet's application layer?

You can find an answer/comment for this review question at the end of the chapter.

Review Question 9

Can music, perhaps in the form of data read from a CD, be streamed across the Internet?

You can find an answer/comment for this review question at the end of the chapter.

Review Question 10

Describe the form and purpose of each of the following:

1. a routing table,
2. a datagram,
3. a file transfer capability, and
4. an error correcting code.

Which TCP/IP layer would each item associated with, and why?

You can find an answer/comment for this review question at the end of the chapter.

Discussion Topics

1. Discuss ways in which the limitations imposed by fixed-format IP addresses may be overcome.
2. The layers of the Internet software, in contrast to those of the OSI stack, have not been assigned clear and distinct responsibilities. The OSI stack, for example, includes a layer responsible for routing and another responsible for error handling, whereas in the Internet software the TCP and IP layers each play some part in both routing and handling errors.

   In the context of layered software, discuss the merits of assigning to each layer a clear and distinct responsibility.
3. It has been suggested that layering is not always the best way of organising software. Discuss the advantages and disadvantages of layering, and propose alternative forms of organising communications software that may be superior to layering.

Answers and Comments

Activity 1

In your diagram, you should assign addresses to the computers distinguished in the source and target sub-nets, and then trace a route from the source computer to the destination computer via distinguished computers in these sub-nets and any intervening sub-nets. Then you should devise routing tables for the distinguished computers along your route in a way that ensures that traffic will be directed along the route. After this, you will be able to simulate the action of an internetwork in automatically routing packets from their source to their destination.

Activity 2

For the purposes of naming, the Internet is divided into a set of domains which, in turn, are divided into sub-domains, and so on. Rather confusingly, sub-domains are commonly referred to as domains.

The domains in mit@cs.uct.ac.za: za, ac.za, uct.ac.za and cs.uct.ac.za.

If you are not familiar with the tree-structure, consult your textbooks.

Review Question 1

The equivalent of 158 is 10011110. The leading digits, 10, make the network is a medium-sized one.

The next 14 bits give the network address. They are: 011110 followed by the eight bit representation for 94, which is: 01011110. In binary, this is: 01111001011110; in decimal, it is 30*256+94 = 7774. The last 16 bits give the computer address, which is 39*256+1 = 9985.

Review Question 2

2^7 = 128, 2^14 = 16 000 and 2^21 = 2 000 000, respectively.

Review Question 3

They can have up to 2^24 = 16 000 000, 2^16 = 64 000, and 2^8 = 256, computers, respectively.

Review Question 4

There is no address available for another one! At this point, the design has to be changed. One of the reasons that IP version 4, which is what the Internet currently uses, is being upgraded to IP version 6 is to provide more addresses for networks.

Review Question 5

By examining the source and destination addresses to see if their network parts are the same or not.

Review Question 6

A file or a message is generated by an application at the Application layer and passed to TCP which divides it into packets and passes them to IP. The IP layer decides how packets will be routed correctly across the network.

Review Question 7

Since it means: "Can I have this file from this server, please?", the answer must be: "It wouldn't."

Review Question 8

FTP and Telnet.

Review Question 9

Yes. After all, the network has to do essentially the same thing as it does when transferring a long file of delivering a long e-mail, both of which it can do.

Review Question 10

1. i) A routing table is a data structure the purpose of which is to record the next part of the route for any packet in transit. As an enabler for routing, it is associated with the routing layer (or IP).

2. A datagram is a data string formatted for transmission. Its purpose is to provide a self-contained packet that can be routed across a network. It is associated with the Transport layer, which constructs datagrams.

3. A file transfer capability is distributed software the purpose of which is to transfer files. As an application, it is associated with the Application layer.

4. An error-correcting code is a way of adding redundant digits to a digit string in such a way that it is possible to correct any changes that may occur to the resulting combined string. Its purpose is to correct errors, and so it is associated with the error-handling layer (or IP).