| **Syllabus** |
|---|
| **Application Layer:** BOOTP and DHCP, DNS, TELNET, FTP, SMTP, HTTP, WWW, VoIP, Four aspects of Network security, Privacy, Digital Signatures. |

## WWW and the HTTP

The World Wide Web (WWW) is a repository of information linked together from points all over the world. The WWW has a unique combination of flexibility, portability, and user-friendly features that distinguish it from other services provided by the Internet. The WWW project was initiated by CERN (European Laboratory for Particle Physics) to create a system to handle distributed resources necessary for scientific research.

The WWW today is a distributed client-server service, in which a client using a browser can access a service using a server. However, the service provided is distributed over many locations called sites. Each site holds one or more documents, referred to as Web pages. Each Web page can contain a link to other pages in the same site or at other sites. The pages can be retrieved and viewed by using browsers.

When a web address (or URL) is typed into a web browser, the web browser establishes a connection to the web service running on the server using the HTTP protocol. URLs (or Uniform Resource Locator) and URIs (Uniform Resource Identifier) are the names most people associate with web addresses.

The URL http://www.cisco.com/index.html is an example of a URL that refers to a specific resource - a web page named index.html on a server identified as cisco.com.

Web browsers are the client applications our computers use to connect to the World Wide Web and access resources stored on a web server. As with most server processes, the web server runs as a background service and makes different types of files available.

To better understand how the web browser and web client interact, we can examine how a web page is opened in a browser. For this example, we will use the URL: http://www.cisco.com/web-server.htm.

First, the browser interprets the three parts of the URL:

1. http (the protocol or scheme)

2. www.cisco.com (the server name)

3. web-server.htm (the specific file name requested).

The browser then checks with a name server to convert www.cisco.com into a numeric address, which it uses to connect to the server. Using the HTTP protocol requirements, the browser sends a GET request to the server and asks for the file web-server.htm. The server in turn sends the HTML code for this web page to the browser. Finally, the browser deciphers the HTML code and formats the page for the browser window.

The ***Hypertext Transfer Protocol (HTTP),*** one of the protocols in the TCP/IP suite, was originally developed to publish and retrieve HTML pages and is now used for distributed, collaborative information systems. HTTP is used across the World Wide Web for data transfer and is one of the most used application protocols.

HTTP specifies a request/response protocol. When a client, typically a web browser, sends a request message to a server, the HTTP protocol defines the message types the client uses to request the web page and also the message types the server uses to respond. The three common message types are GET, POST, and PUT.

**GET** is a client request for data. A web browser sends the GET message to request pages from a web server. As shown in the figure, once the server receives the GET request, it responds with a status line, such as HTTP/1.1 200 OK, and a message of its own, the body of which may be the requested file, an error message, or some other information.

**POST** and **PUT** are used to send messages that upload data to the web server. For example, when the user enters data into a form embedded in a web page, POST includes the data in the message sent to the server.

**PUT** uploads resources or content to the web server.

## TELNET

TELNET is an abbreviation for Terminal Network. It is the standard TCP/IP protocol for virtual terminal service as proposed by the International Organization for Standards (ISO). TELNET enables the establishment of a connection to a remote system in such a way that the local terminal appears to be a terminal at the remote system.

TELNET was designed at a time when most operating systems, such as UNIX, were operating in a timesharing environment. In such an environment, a large computer supports multiple users. The interaction between a user and the computer occurs through a terminal, which is usually a combination of keyboard, monitor, and mouse. Even a microcomputer can simulate a terminal with a terminal emulator.

In a timesharing environment, users are part of the system with some right to access resources. Each authorized user has an identification and probably, a password. The user identification defines the user as part of the system. To access the system, the user logs into the system with a user id or log-in name. The system also includes password checking to prevent an unauthorized user from accessing the resources.

When a user logs into a local timesharing system, it is called local log-in. As a user types at a terminal or at a workstation running a terminal emulator, the keystrokes are accepted by the terminal driver. The terminal driver passes the characters to the operating system. The operating system, in turn, interprets the combination of characters and invokes the desired application program or utility.

When a user wants to access an application program or utility located on a remote machine, she performs remote log-in. Here the TELNET client and server programs come into use. The user sends the keystrokes to the terminal driver, where the local operating system accepts the characters but does not interpret them. The characters are sent to the TELNET client, which transforms the characters to a universal character set called network virtual terminal (NVT) characters and delivers them to the local TCP/IP protocol stack.

## File Transfer Protocol – FTP

File Transfer Protocol (FTP) is the standard mechanism provided by TCP/IP for copying a file from one host to another. Although transferring files from one system to another seems simple and straightforward, some problems must be dealt with first. For example, two systems may use different file name conventions. Two systems may have different ways to represent text and data. Two systems may have different directory structures. All these problems have been solved by FTP in a very simple and elegant approach.

FTP differs from other client/server applications in that it establishes two connections between the hosts. One connection is used for data transfer, the other for control information (commands and responses). Separation of commands and data transfer makes FTP more efficient. The control connection uses very simple rules of communication. We need to transfer only a line of command or a line of response at a time. The data connection, on the other hand, needs more complex rules due to the variety of data types transferred. However, the difference in complexity is at the FTP level, not TCP. For TCP, both connections are treated the same. FTP uses two well-known TCP ports: Port 21 is used for the control connection, and port 20 is used for the data connection.

## Simple Mail Transfer Protocol – SMTP

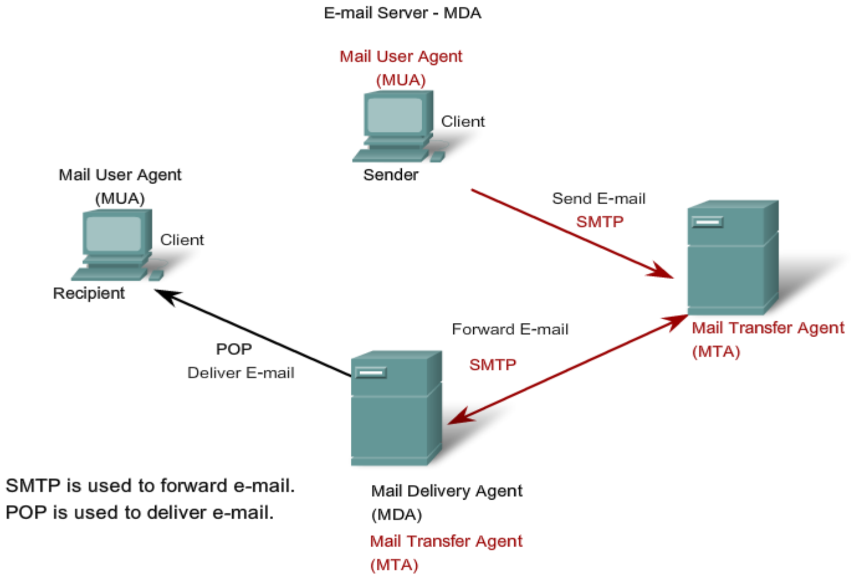
When people compose e-mail messages, they typically use an application called a Mail User Agent (MUA), or e-mail client. The MUA allows messages to be sent and places received messages into the client's mailbox, both of which are distinct processes.

In order to receive e-mail messages from an e-mail server, the e-mail client can use POP. Sending e-mail from either a client or a server uses message formats and command strings defined by the SMTP protocol. Usually an e-mail client provides the functionality of both protocols within one application.

E-mail can use the protocols, POP and SMTP are inbound mail delivery protocols and are typical client/server protocols. They deliver e-mail from the e-mail server to the client (MUA). The MDA listens for when a client connects to a server. Once a connection is established, the server can deliver the e-mail to the client.

The Simple Mail Transfer Protocol (SMTP), on the other hand, governs the transfer of outbound e-mail from the sending client to the e-mail server (MDA), as well as the transport of e-mail between e-mail servers (MTA). SMTP enables e-mail to be transported across data networks between different types of server and client software and makes e-mail exchange over the Internet possible.

The SMTP protocol message format uses a rigid set of commands and replies. These commands support the procedures used in SMTP, such as session initiation, mail transaction, forwarding mail, verifying mailbox names, expanding mailing lists, and the opening and closing exchanges.



Some of the commands specified in the SMTP protocol are:

HELO - identifies the SMTP client process to the SMTP server process.

EHLO - Is a newer version of HELO, which includes services extensions.

MAIL FROM - Identifies the sender.

RCPT TO - Identifies the recipient.

DATA - Identifies the body of the message.

## Voice Over IP – VOIP

The idea is to use the Internet as a telephone network with some additional capabilities. Instead of communicating over a circuit-switched network, this application allows communication between two parties over the packet-switched Internet. Two protocols have been designed to handle this type of communication: SIP and H.323.

### SIP

The Session Initiation Protocol (SIP) was designed by IETE It is an application layer protocol that establishes, manages, and terminates a multimedia session (call). It can be used to create two-party, multiparty, or multicast sessions. SIP is designed to be independent of the underlying transport layer; it can run on UDP, TCP, or SCTP.

### H.323

H.323 is a standard designed by lTV to allow telephones on the public telephone network to talk to computers (called terminals in H.323) connected to the Internet. A gateway connects the Internet to the telephone network. In general, a gateway is a five-layer device that can translate a message from one protocol stack to another. The gateway here does exactly the same thing. It transforms a telephone network message to an Internet message. The gatekeeper server on the local area network plays the role of the registrar server.

## Dynamic Host Configuration Protocol - DHCP

The Dynamic Host Configuration Protocol (DHCP) service enables devices on a network to obtain IP addresses and other information from a DHCP server. This service automates the assignment of IP addresses, subnet masks, gateway and other IP networking parameters.

DHCP allows a host to obtain an IP address dynamically when it connects to the network. The DHCP server is contacted and an address requested. The DHCP server chooses an address from a configured range of addresses called a pool and assigns ("leases") it to the host for a set period.

On larger local networks, or where the user population changes frequently, DHCP is preferred. New users may arrive with laptops and need a connection. Others have new workstations that need to be connected. Rather than have the network administrator assign IP addresses for each workstation, it is more efficient to have IP addresses assigned automatically using DHCP.

DHCP distributed addresses are not permanently assigned to hosts but are only leased for a period of time. If the host is powered down or taken off the network, the address is returned to the pool for reuse. This is especially helpful with mobile users that come and go on a network. Users can freely move from location to location and re-establish network connections. The host can obtain an IP address once the hardware connection is made, either via a wired or wireless LAN.

DHCP makes it possible for you to access the Internet using wireless hotspots at airports or coffee shops. As you enter the area, your laptop DHCP client contacts the local DHCP server via a wireless connection. The DHCP server assigns an IP address to your laptop.

## Domain Name System – DNS

In data networks, devices are labeled with numeric IP addresses, so that they can participate in sending and receiving messages over the network. However, most people have a hard time remembering this numeric address. Hence, domain names were created to convert the numeric address into a simple, recognizable name.

On the Internet these domain names, such as www.cisco.com, are much easier for people to remember than 198.133.219.25, which is the actual numeric address for this server. Also, if Cisco decides to change the numeric address, it is transparent to the user, since the domain name will remain www.cisco.com. The new address will simply be linked to the existing domain name and connectivity is maintained. When networks were small, it was a simple task to maintain the mapping between domain names and the addresses they represented. However, as networks began to grow and the number of devices increased, this manual system became unworkable.

The Domain Name System (DNS) was created for domain name to address resolution for these networks. DNS uses a distributed set of servers to resolve the names associated with these numbered addresses.

The DNS protocol defines an automated service that matches resource names with the required numeric network address. It includes the format for queries, responses, and data formats. DNS protocol communications use a single format called a message. This message format is used for all types of client queries and server responses, error messages, and the transfer of resource record information between servers.

DNS is a client/server service; however, it differs from the other client/server services that we are examining. While other services use a client that is an application (such as web browser, e-mail client), the DNS client runs as a service itself. The DNS client, sometimes called the DNS resolver, supports name resolution for our other network applications and other services that need it.

When configuring a network device, we generally provide one or more DNS Server addresses that the DNS client can use for name resolution. Usually the Internet service provider provides the addresses to use for the DNS servers. When a user's application requests to connect to a remote device by name, the requesting DNS client queries one of these name servers to resolve the name to a numeric address.

Computer operating systems also have a utility called nslookup that allows the user to manually query the name servers to resolve a given host name. This utility can also be used to troubleshoot name resolution issues and to verify the current status of the name servers.

The Domain Name System uses a hierarchical system to create a name database to provide name resolution. The hierarchy looks like an inverted tree with the root at the top and branches below.

At the top of the hierarchy, the root servers maintain records about how to reach the top-level domain servers, which in turn have records that point to the secondary level domain servers and so on.

The different top-level domains represent either the type of organization or the country of origin. Examples of top-level domains are:

.au - Australia
.co - Colombia

.com - a business or industry
.jp - Japan
.org - a non-profit organization

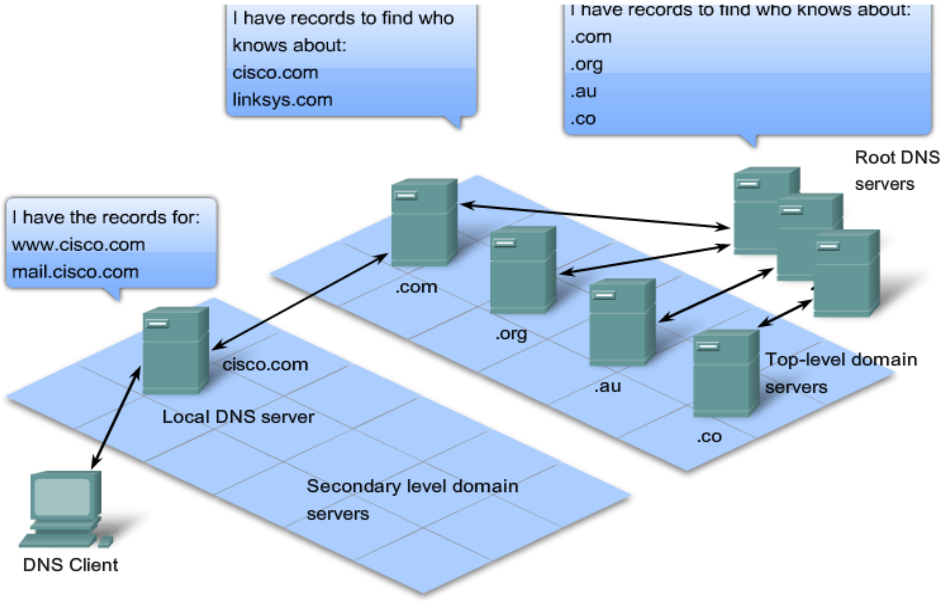
After top-level domains are second-level domain names, and below them are other lower level domains.

Each domain name is a path down this inverted tree starting from the root.

For example, as shown in the figure, the root DNS server may not know exactly where the e-mail server mail.cisco.com is located, but it maintains a record for the "com" domain within the top-level domain. Likewise, the servers within the "com" domain may not have a record for mail.cisco.com, but they do have a record for the "cisco.com" domain. The servers within the cisco.com domain have a record (a MX record to be precise) for mail.cisco.com.

The Domain Name System relies on this hierarchy of decentralized servers to store and maintain these resource records. The resource records list domain names that the server can resolve and alternative servers that can also process requests. If a given server has resource records that correspond to its level in the domain hierarchy, it is said to be authoritative for those records.

For example, a name server in the cisco.netacad.net domain would not be authoritative for the mail.cisco.com record because that record is held at a higher domain level server, specifically the name server in the cisco.com domain.



A hierarchy of DNS servers contains the resource records that match names with addresses.

## Network Security

Secured communication requires the following four basic services:

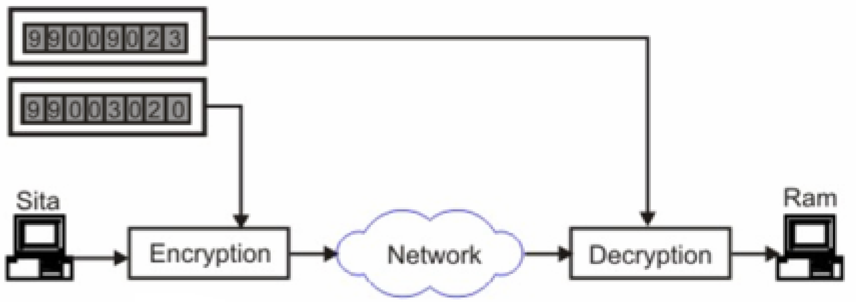
- **Privacy:** A person (say Sita) should be able to send a message to another person (say Ram) privately. It implies that to all others the message should be unintelligible.
- **Authentication:** After the message is received by Ram, he should be sure that the message has been sent by nobody else but by Sita.
- **Integrity:** Ram should be sure that the message has not been tampered on transit.
- **Nonrepudiation:** Ram should be able to prove at a later stage that the message was indeed received from Sita.

## Digital Signature

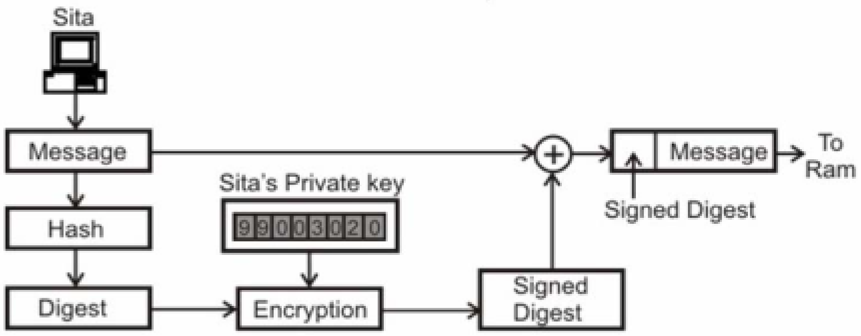
There are two alternatives for Digital Signature:

- Signing the entire document.
- Signing the digest

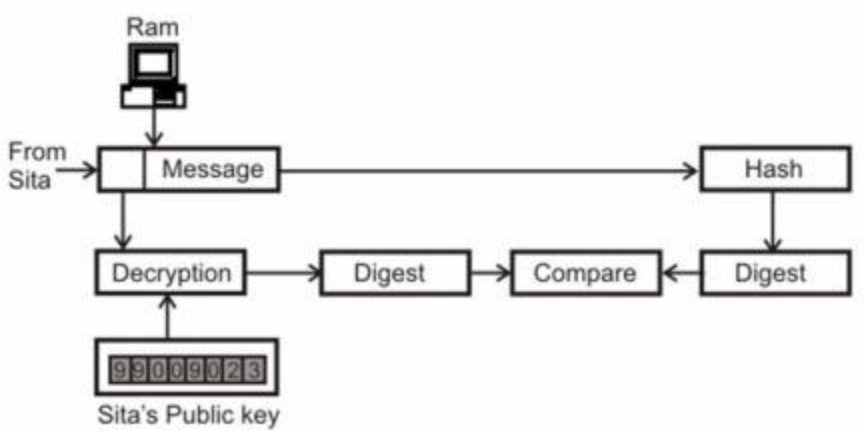
In the first case the entire document is encrypted using private key of the sender and at the receiving end it is decrypted using the public key of the sender. For a large message this approach is very inefficient. In the second case a miniature version of the message, known as *digest*, is encrypted using the private key of the sender and then the signed digest along with the message is sent to the receiver. The receiver decrypts the signed digest using the public key of the sender and the digest created using the received message is compared with the decrypted digest. If the two are identical, it is assumed that the sender is authenticated. This is somewhat similar to error detection using parity bit.



Authentication by signing the whole document.



Sender site for authentication by signed digest



Receiver site for authentication by signed digest

Some key features of this approach are mentioned below:

- Digital signature does not provide privacy.
- Hash function is used to create a message digest.
- It creates a fixed-length digest from a variable-length message.
- Most common Hash functions:
    - MD5 (Message Digest 5): 120-bit.
    - SHA-1 (Secure Hash algorithm 1): 160-bit.
- Important properties:
    - One-to-One.
    - One-way.

## References:-

1. A.S.Tanenbaum, ”Computer Networks”, Pearson Education, Fourth Edition.

2. Data Communication And Networking 4th Edition by B.A. Forouzan Tata McGrawhill Publication.

3. CISCO NETACADEMY.