

2020

EDITION

SCS

sultan chand

SCS

sultan chand

Textbook for CBSE Class XII

Computer Science

— *With* —

Python

PREETI ARORA

Web Support

at

sultan-chand.com/ws/python12

- Presentations • Projects & Program Codes
- Sample Papers • Model Test Papers
- Practical File • Viva Voce

- ❖ Advanced Python Concepts
- ❖ Modules
- ❖ Recursion
- ❖ Data Structures
- ❖ Computer Networks
- ❖ MySQL
- ❖ Society, Law & Ethics
- ❖ Project Work

SCS

sultan chand

SUPPLEMENT

(as per 2020–21 Curriculum)



FROM THE AUTHOR



The CBSE always believes in global trends of education and keeps reviewing and updating the syllabus to make it more relevant to the ongoing educational transformation. The changes brought about by the CBSE in the Computer Science curriculum in the recent past will go a long way in improving the technical skills of students. Keeping in sync with the latest curriculum updates by the CBSE in Computer Science with Python for Class XII, we have brought out this Supplement incorporating all the changes effected in the syllabus for the 2020-21 academic session.

As per the new curriculum, the following additions have been made in this Supplement:

- ❖ **Addition of math and string library functions in Chapter 3 of the textbook.**
- ❖ **Addition of read and write operations on to a csv file, random access on files (`seek()` and `tell()`) and more questions on binary and csv file handling (Chapter 4).**
- ❖ **Chapter 8—“Computer Networks”—has been reframed.**
- ❖ **“More on SQL” (earlier Chapter 10) has been reframed and renamed as “Database Concepts and SQL” and is now the new Chapter 12.**
- ❖ **Major topics of the earlier Unit IV have now been included in Chapter 8 (“Computer Networks”).**

A comparative analysis of the old and the new curriculum has already been shared with teachers and is also being given on the following pages to keep the users of this Supplement up-to-date vis-a-vis the changes made in the latest syllabus. I am confident that students and teachers will benefit immensely by making the best use of this Supplement.

Happy learning!

PREETI ARORA

COMPARATIVE ANALYSIS OF CLASS XII OLD AND NEW CBSE CURRICULUM			
Computer Science (083)			
UNIT	OLD CURRICULUM (2019-20)	NEW CURRICULUM (2020-21)	CHANGES
I: Computational Thinking and Programming	30 Marks	40 Marks	Total weightage of this unit increased by 10 Marks
	Revision of the Basics of Python	Revision of the Basics of Python	No change
	Functions: scope, parameter passing, mutable/immutable properties of data objects, pass arrays to functions, return values, functions using libraries: mathematical and string functions.	Functions: scope, parameter passing, mutable/immutable properties of data objects, passing strings, lists, tuples, dictionaries to functions, default parameters, positional parameters, return values, Functions using libraries: mathematical and string functions.	math and string functions to be added
	File handling: open and close a file, read, write, and append to a file, standard input, output, and error streams, relative and absolute paths.	File handling: Need for a data file, Types of file: Text files, Binary files and CSV (Comma separated values) files. Text File: Basic operations on a text file: Open (filename – absolute or relative path, mode) / Close a text file, Reading and Manipulation of data from a text file, Appending data into a text file, standard input/output and error streams, relative and absolute paths. Binary File: Basic operations on a binary file: Open(filename – absolute or relative path, mode) / Close a binary file, pickle Module – methods load and dump; Read, Write/Create, Search, Append and Update operations in a binary file. CSV File: import csv module, functions – Open / Close a csv file, Read from a csv file and Write into a csv file using csv.reader() and csv.writer().	(1) CSV Files– open, read and write functions (2) Read from a csv file and Write into a csv file using csv.reader() and csv.writer()
	Using Python libraries: create and import Python libraries	Using Python libraries: create and import Python libraries.	No change
	Recursion: simple algorithms with recursion: factorial, Fibonacci numbers; recursion on arrays: binary search	Recursion: simple algorithms with recursion: print a message forever, sum of first n natural numbers, factorial, Fibonacci numbers; recursion on arrays: binary search	No change

	Idea of efficiency: performance defined as inversely proportional to the wall clock time, count the number of operations a piece of code is performing, and measure the time taken by a program. Example: take two different programs for the same problem, and understand how the efficient one takes less time.	Idea of efficiency: performance measurement in terms of the number of operations.	No change
	Data visualization using Pyplot: line chart, pie chart, and bar chart.		Entire Chapter removed
	Data-structures: lists, stacks, queues.	Data-structures: Lists as covered in Class XI, Stacks – Push, Pop using a list, Queues – Insert, Delete using a list.	No change
II: Computer Networks (Entire Unit reframed)	15 Marks Structure of a network: Types of networks: local area and wide area (web and internet), new technologies such as cloud and IoT, public vs. private cloud, wired and wireless networks; concept of a client and server.	10 Marks Evolution of Networking: ARPANET, Internet, Interspace Different ways of sending data across the network with reference to switching techniques (Circuit and Packet switching). Data Communication terminologies: Concept of Channel, Bandwidth (Hz, KHz, MHz) and Data transfer rate (bps, Kbps, Mbps, Gbps, Tbps). Transmission media: Twisted pair cable, coaxial cable, optical fibre, infrared, radio link, microwave link and satellite link. Network Topologies and types: Bus, Star, Tree, Types of Network: PAN, LAN, WAN, MAN.	Total weightage of this unit reduced by 5 marks (1) Evolution of Networking (2) Switching Techniques (3) Data Communication Terminologies (4) Topologies (1) Three more Network devices added

<p>Network stack: amplitude and frequency modulation, collision in wireless networks, error checking, and the notion of a MAC address, main idea of routing. IP addresses: (v4 and v6), routing table, router, DNS, and web URLs, TCP: basic idea of retransmission, and rate modulation when there is congestion (analogy to a road network),</p>	<p>Network Protocol: TCP/IP, File Transfer Protocol (FTP), PPP, HTTP, SMTP, POP3, Remote Login (Telnet) and Internet, Wireless / Mobile Communication protocol such as GSM, GPRS and WLL.</p>	<p>Topics removed: (1) AM & FM (2) CSMA/CA & CD (3) Error Checking (4) Idea of Routing (5) Network Congestion</p>
<p>Protocols: 2G, 3G, 4G, Wi-Fi. What makes a protocol have a higher bandwidth?</p>		
<p>Application layer: HTTP (basic idea), working of email, secure communication: encryption and certificates (HTTPS), network applications: remote desktop, remote login, HTTP, FTP, SCP, SSH, POP/IMAP, SMTP, VoIP, NFC.</p>	<p>Mobile Telecommunication Technologies: 1G, 2G, 3G, 4G and 5G; Mobile processors; Electronic mail protocols such as SMTP, POP3, Protocols for Chat and Video Conferencing: VoIP, Wireless technologies such as Wi-Fi and WiMax</p>	<p>No major changes</p>
<p>Basic network tools: traceroute, ping, ipconfig, nslookup, whois, speed-test.</p>		<p>Entire Topic removed</p>
	<p>Network Security Concepts: Threats and prevention from Viruses, Worms, Trojan horse, Spams. Use of Cookies, Protection using Firewall, https; India IT Act, Cyber Law, Cyber Crimes, IPR issues, hacking.</p>	<p>Mentioned Topics have been added to this unit; taken from Unit-4 (Society Law and Ethics)—old curriculum</p>
	<p>Introduction to Web services: WWW, Hyper Text Markup Language (HTML), Extensible Markup Language (XML); Hyper Text Transfer Protocol (HTTP); Domain Names; URL; Website, Web browser, Web Servers; Web Hosting, Web Scripting</p>	<p>Newly-added topics</p>

	sultan chand	E-commerce payment transactions using online banking, mobile banking, payment apps and services.	Newly-added topics
	15 Marks	20 Marks	Total weightage of this unit increased by 5 marks
	Write a minimal Django-based web application that parses a GET and POST request, and writes the fields to a file—flat file and CSV file.		Complete Chapter removed
	Interface Python with an SQL database	Interface of Python with an SQL database Connecting SQL with Python Creating Database connectivity Applications Performing Insert, Update, Delete queries Display data by using fetchone(), fetchall(), rowcount()	No change
III: Data Management	SQL commands: aggregation functions – having, group by, order by.	Database Concepts: Introduction to database concepts and its need. Relational data model: Concept of domain, relation, tuple, attribute, degree, cardinality, key, primary key, candidate key, alternate key and foreign key Structured Query Language: General Concepts: Advantages of using SQL, Data Definition Language and Data Manipulation Language; Data Types: number / decimal, character / varchar / varchar2, date; SQL commands: CREATE TABLE, DROP TABLE, ALTER TABLE, UPDATESET , INSERT, DELETE; SELECT, DISTINCT, FROM, WHERE, IN, BETWEEN, LIKE, NULL / IS NULL, ORDER BY, GROUP BY, HAVING; SQL Aggregate functions: SUM (), AVG (), COUNT (), MAX () and MIN (); Joins: equi-join and natural join	The mentioned topics have been added from Class XI previous curriculum

	SCS sultan 10 Marks	SCS 0 Marks	No weightage
IV: Society, Law and Ethics	<p>Intellectual property rights, plagiarism, digital rights management, and licensing (Creative Commons, GPL and Apache), open source, open data, privacy.</p> <p>Privacy laws, fraud; cyber-crime—phishing, illegal downloads, child pornography, scams; cyber forensics, IT Act, 2000.</p> <p>Technology and society: understanding of societal issues and cultural changes induced by technology. E-waste management: proper disposal of used electronic gadgets.</p> <p>Identity theft, unique ids, and biometrics.</p> <p>Gender and disability issues while teaching and using computers.</p>		<p>Major Topics of this Unit have been added to Computer Networks.</p> <p>This Unit has been completely removed.</p>

RED - Removed Topics
GREEN – Detailed added topics
PURPLE – No Change
BLUE - Added Topics

CONTENTS

CHAPTER 3	Python Libraries (Additions)	3.1 – 3.8
CHAPTER 4	Data File Handling (Additions)	4.1 – 4.25
CHAPTER 8	Computer Networks	8.1 – 8.76
CHAPTER 12	Relational Database and SQL	12.1 – 12.71

3

Python Libraries (Additions)

3.11 USING PYTHON STANDARD LIBRARY'S FUNCTIONS AND MODULES

Python's standard library is very extensive, offering a wide range of modules and functions. The library contains built-in modules (written in C) that provide access to system functionality such as file I/O, that would otherwise be inaccessible to Python programmers, as well as modules written in Python that provide standardized solutions for many problems that occur in everyday programming. Some of these important modules are explicitly designed to encourage and enhance the portability of Python programs such as **math** module for mathematical functions; **random** module for pseudo-random number generation; and **string** module for string/text-related functions.

We have already discussed important math and string modules and their functions.

Let us talk about some more built-in mathematical functions.

3.11.1 Python's built-in Mathematical Functions

We have already discussed major functions using **math** module. Python provides many other mathematical built-in functions as well. These include trigonometric functions, representation functions, logarithmic functions, angle conversion functions, etc. In addition, two mathematical constants are also defined in this module.

Pie (π): Pie (π) is a well-known mathematical constant, which is defined as the ratio of the circumference to the diameter of a circle and its value is 3.141592653589793. In order to import this single object (pie) from the **math** module through any of the two methods given below:

```
>>> import math  
>>> math.pi  
3.141592653589793
```

Alternatively, it can be done without prefixing the module name and writing the name of the object directly after keyword **import** as shown below:

```
>>> from math import pi  
>>> print(pi)  
>>> 3.141592653589793
```

Similarly, another well-known mathematical constant defined in the math module is **e**. It is called **Euler's number** and it is a base of the natural logarithm. Its value is 2.718281828459045.

```
>>> math.e
2.718281828459045
```

Hence, these constants can be directly called and used from math module.

It should be remembered not to use module name with imported object if imported through **from <module> import** command because now the imported object is part of your program's environment.

The math module contains functions for calculating various trigonometric ratios for a given angle. The functions (sin, cos, tan, etc.), which we have already discussed, need the angle in radians as an argument. We, on the other hand, are used to express the angle in degrees.

The math module presents two-angle conversion functions -- **degrees()** and **radians()** -- to convert the angle from degrees to radians and vice versa. *For example*, the following statements convert the angle of 30 degrees to radians and back.

degrees() converts a value in radians to degrees using formula as $\text{degrees}(x) = \frac{\pi}{180}x$

On the contrary, **radians()** converts a value in degrees to radians: $\text{radians}(x) = \frac{180}{\pi}x$

POINT TO REMEMBER

π radian is equivalent to 180 degrees.

For example,

To convert an imputed angle of 50 degree to radians and vice versa.

```
prog_rad_degree1.py - C:/Users/preeti/AppData/Local/Programs/Python/Python37-32/prog_r... - □ x
File Edit Format Run Options Window Help
import math
print("Conversion of degree to radians",math.radians(50))
print("Conversion of radians to degree",math.degrees(math.pi/6))

Ln: 7 Col: 0
```

```
>>>
RESTART: C:/Users/preeti/AppData/Local/Programs/Python/
ree1.py
Conversion of degree to radians 0.8726646259971648
Conversion of radians to degree 29.999999999999996
```

Practical Implementation-1

Write a Python program to calculate the area of a regular polygon.

For solving this problem, we have to understand the formula to be used first, which is given as:

Area of a Regular Polygon



n = number of sides

s = length of a side

$$\text{Area} = \frac{ns}{4 \tan\left(\frac{\pi}{n}\right)}$$

```
prog_area_polygon.py - C:/Users/preeti/AppData/Local/Programs/Python/Python37-32/prog_area... - □ x
File Edit Format Run Options Window Help
#Python program to calculate the area of a regular polygon

from math import tan, pi
no_sides = int(input("Input number of sides: "))
side_length = float(input("Input the length of a side: "))
poly_area = no_sides * (side_length ** 2) / (4 * tan(pi / no_sides))
print("The area of the polygon is: ", poly_area)

Ln: 16 Col: 0
```

Practical Implementation-2

Write a Python program to find the roots of a quadratic function.

A quadratic equation is represented as: $ax^2 + bx + c = 0$

Here, a, b and c are numbers, where $a \neq 0$; and

x is to be calculated.

The **roots** of any **quadratic equation** are given by the formula:

$$x = [-b \pm \sqrt{-b^2 - 4ac}] / 2a$$

```
>>>
RESTART: C:/Users/preeti/AppData/Local/Programs/Python/lygon.py
Input number of sides: 6
Input the length of a side: 10
The area of the polygon is:  259.8076211353316
```

```
prog_quad_root.py - C:/Users/preeti/AppData/Local/Programs/Python/Python37-32/prog_quad_ro
File Edit Format Run Options Window Help
#Program to find the roots of a quadratic function
from math import sqrt

print("Quadratic equation is : (a * x^2) + b*x + c")
a = float(input("a: "))
b = float(input("b: "))
c = float(input("c: "))

r = b**2 - 4*a*c

if r > 0:
    num_roots = 2
    x1 = ((-b) + sqrt(r))/(2*a)
    x2 = ((-b) - sqrt(r))/(2*a)
    print("The two roots are: ", x1)
    print("and", x2)
elif r == 0:
    num_roots = 1
    x = (-b) / 2*a
    print("There is one root: ", x)
else:
    num_roots = 0
    print("No roots, discriminant < 0.")
```

```

>>>
RESTART: C:/Users/preeti/AppData/Local/Programs/Python
ot.py
Quadratic equation is : (a * x^2) + b*x + c
a: 6
b: 25
c: 15
The two roots are: -0.7267649503250245
and -3.4399017163416423

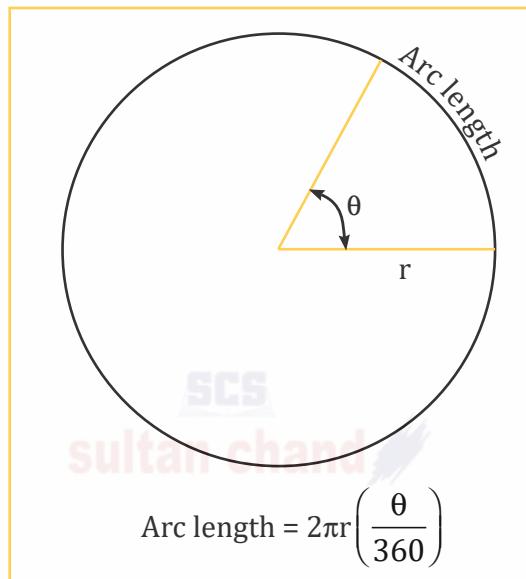
```



Practical Implementation-3

Write a Python program to calculate arc length of an angle.

In a planar geometry, an angle is a figure formed by two rays called the sides of the angle, sharing a common endpoint known as the vertex of the angle. Angles formed by two rays lie in a plane, but this plane does not have to be a Euclidean plane.



```

prog_arc_angle1.py - C:/Users/preeti/AppData/Local/Programs/Python/...
File Edit Format Run Options Window Help
#Python program to calculate arc length of an angle

import math
def arclength():    #user-defined function
    diameter = float(input('Diameter of circle: '))
    angle = float(input('angle measure: '))
    if angle >= 360:
        print("Angle is not possible")
        return
    arc_length = (math.pi*diameter) * (angle/360)
    print("Arc Length is: ", arc_length)

arclength()

```

RESTART: C:/Users/preeti/AppData/Local/Programs/Python
ot1.py
Diameter of circle: 20
angle measure: 45
Arc Length is: 7.853981633974483

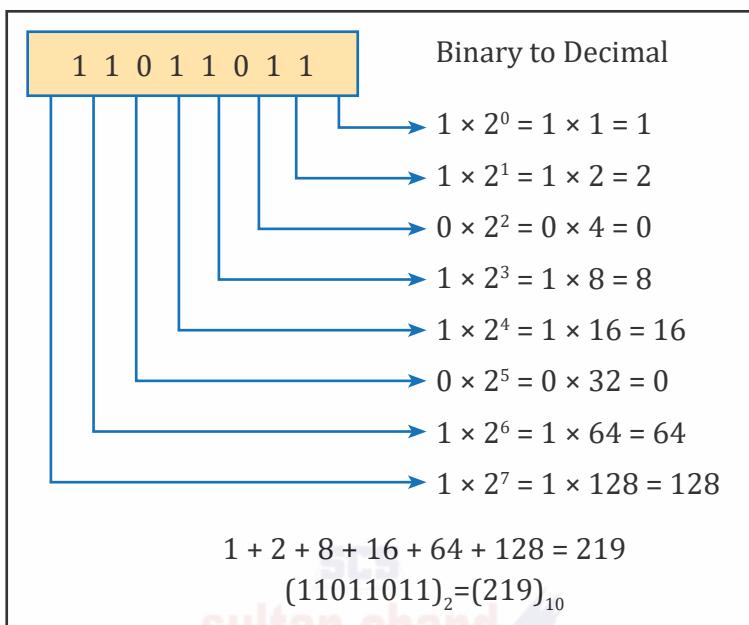
Practical Implementation-4

Write a Python program to convert a binary number to decimal number.

A binary number is a number which is represented in the form of 0 and 1 (binary digits) only and decimal number system constitutes numbers from 0 to 9, the contemporary number system used mathematically.

Also, a binary number system has 2 as its base while the decimal system has 10 as its base.

The method of conversion of a binary number to its decimal equivalent is:



```
prog_bin_decimal.py - C:/Users/preeti/AppData/Local/Programs/Python/Python37-32... - □ x
File Edit Format Run Options Window Help
#Python program to convert a binary number to decimal number
import math
b_num = list(input("Input a binary number: "))
value = 0

for i in range(len(b_num)):
    digit = b_num.pop()
    if digit == '1':
        value = value + math.pow(2, i)
print("The decimal value of the number is", value)

Ln: 13 Col: 0
```

```
>>>
RESTART: C:/Users/preeti/AppData/Local/Programs/Python/Python37-32/prog_bin_decimal.py
Input a binary number: 11011011
The decimal value of the number is 219
```

In the above program, the input taken from the user is converted to a list using the function `list()` and gets stored inside variable `b_num`. The loop gets executed up to the length of the list, *i.e.*, the number of list elements. With every iteration, each digit is extracted or popped from the list using function `pop()`. This extracted digit is checked for character as 1 and is raised as the power of 2. Finally, the value gets printed which is in decimal format. In case the digit is 0, no execution takes place since the product of any value with 0 results in 0 itself. Hence, the output is obtained as 219.

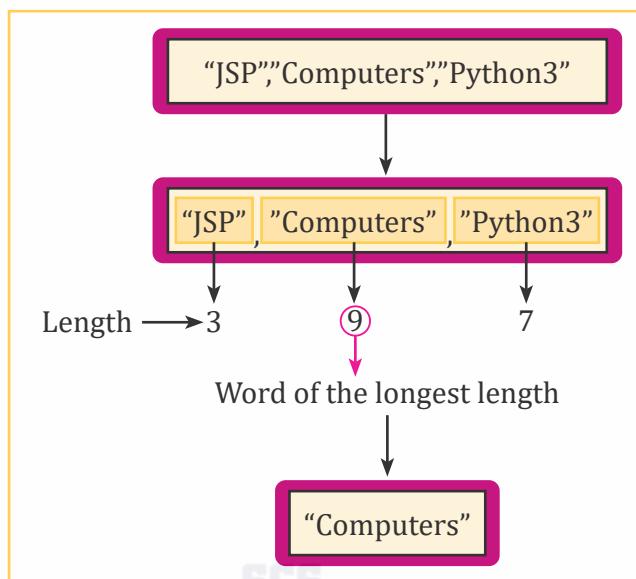
3.11.2 Python's built-in String Functions

We have already discussed the built-in String functions in the previous chapters.

Let us now see some more practical implementations of most commonly used inbuilt string functions.

Practical Implementation-5

Write a Python function that takes a list of words and returns the length of the longest one.



```
File Edit Format Run Options Window Help
#Python function that takes a list of words and returns
#the length of the longest one

def find_longest_word(words_list):
    word_len = []
    for n in words_list:
        word_len.append((len(n), n))
    word_len.sort()
    return word_len[-1][1]

print(find_longest_word(["JSP", "Computers", "Python3"]))
```

```
>>>
RESTART: C:/Users/preeti/AppData/
_word.py
Computers
```

In the above program, two in-built string functions, `append()` and `sort()`, have been used for adding elements to an empty list, `word_len`, and the list elements are sorted in ascending order using `sort()` method.

Practical Implementation-6

Write a Python program to capitalize first and last letters of each word of a given string.



```
prog_capitalize2.py - C:/Users/preeti/AppData/Local/Programs/Python/Python37-32/prog_capitalize2.p... - □ x
File Edit Format Run Options Window Help
#Python program to capitalize first and last letters of
#each word of a given string
def capitalize_first_last_letters(str1):
    str1 = result = str1.title()
    result = ""
    for word in str1.split():
        result += word[:-1] + word[-1].upper() + " "
    return result[:-1]

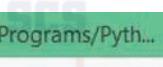
print(capitalize_first_last_letters("python string library functions"))
print(capitalize_first_last_letters("implementation"))

Ln: 14 Col: 0
```

```
>>>
RESTART: C:/Users/preeti/AppData/Local
ize2.py
PythoN StrinG LibrarY Functions
Implementation
```

Practical Implementation-7

Write a Python program to swap each character of a given string from lowercase to uppercase and vice versa.



```
prog_swap_case.py - C:/Users/preeti/AppData/Local/Programs/Pyth... - □ x
File Edit Format Run Options Window Help
#Program to swap each character of a given string
#from lowercase to uppercase and vice-versa
def swap_case_string(str1):
    result_str = ""
    for item in str1:
        if item.isupper():
            result_str += item.lower()
        else:
            result_str += item.upper()
    return result_str
print(swap_case_string("String Module in"))
print(swap_case_string("Python"))
print(swap_case_string("Matplotlib Library"))

Ln: 18 Col: 0
```

```
>>>
RESTART: C:/Users/preeti/AppData
se.py
STRING MODULE IN
PYTHON
mATPLOTLIB LIBRARY
```

SOLVED QUESTIONS

1. Write a program that performs the following operations on a string:
- Prompt the user to input a string.
 - Extract all the digits from the string.
 - If there are digits in the inputted string:
 - Calculate and display the sum of digit.
 - Also display:
 - The original string
 - The digits
 - The sum of the digits
 - If there are no digits:
 - Display the original string along with an appropriate message as "No Digits are present"

```
Ans. str1 = input("Enter the string: ")
sum = 0
num = 0
if str1.isalpha() == False:
    for i in str1:
        if i.isdigit() == True:
            num = num*10 + int(i)
            sum += int(i)
    print("Original String: ",str1)
    print("Digits: ",num)
    print("Sum of digits is: ",sum)
else:
    print("Original String: ", str1, "has no digit")
```

2. Write a program with a user-defined function with string as a parameter which replaces all vowels in the string with '*'.

```
Ans. #Function to replace all vowels in the string with '*'
def replaceVowel(st):
    #Create an empty string
    newstr = ''
    for character in st:
        #Check if next character is a vowel
        if character in 'aeiouAEIOU':
            #Replace vowel with *
            newstr += '*'
        else:
            newstr += character
    return newstr
    #End of function
st = input("Enter a String: ")
st1 = replaceVowel(st)
print("The original String is:", st)
print("The modified String is:", st1)
```

UNSOLVED QUESTIONS

- Write a user-defined function to calculate the area of a triangle.
- Write a user-defined function to convert a string with more than one word into title case string where string is passed as parameter. (Title case means that the first letter of each word is capitalized.)
- Write a function deleteChar() which takes two parameters -- one is a string and the other is a character. The function should create a new string after deleting all occurrences of the character from the string and return the new string.
- Write a program to print the number of occurrences of a substring into a line using built-in string function find().

4

Data File Handling (Additions)

4.11 BINARY FILES OPERATIONS

Most of the files that we see in our computer system are called binary files.

Example:

- **Document files:** .pdf, .doc, .xls etc.
- **Image files:** .png, .jpg, .gif, .bmp etc.
- **Video files:** .mp4, .3gp, .mkv, .avi etc.
- **Audio files:** .mp3, .wav, .mka, .aac etc.
- **Database files:** .mdb, .accde, .frm, .sqlite etc.
- **Archive files:** .zip, .rar, .iso, .7z etc.
- **Executable files:** .exe, .dll, .class etc.



All binary files follow a specific format. We can open some binary files in the normal text editor but we can't read the content present inside the file. That's because all the binary files will be encoded in the binary format, which can be understood only by a computer or a machine.

In binary files, there is no delimiter to end a line. Since they are directly in the form of binary, hence there is no need to translate them. That's why these files are easy and fast in working.

We have already discussed about pickle module and handling (reading and writing) binary file operations using load() and dump() methods of this (pickle) module. We will now discuss the four major operations performed using a binary file such as—

1. Inserting/Appending record in a binary file.
2. Read records from a binary file.
3. Search a record in a binary file.
4. Update a record in a binary file

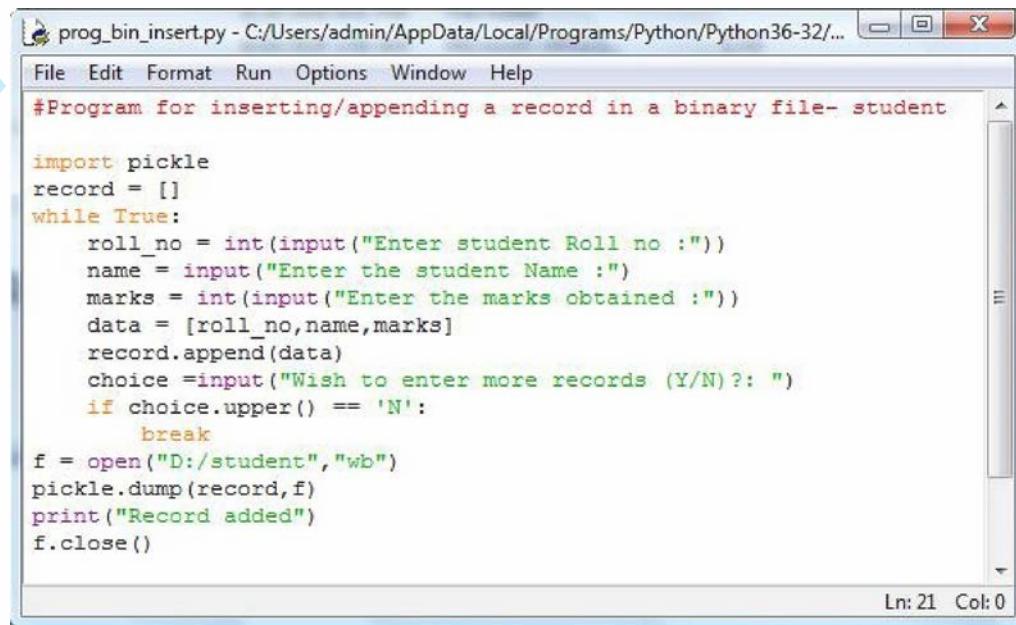
We will now discuss each of them. The file in consideration is “student” file with fields as Roll_no, name and marks.

Insert/Append a Record in Binary File

Inserting or adding (appending) a record into a binary file requires importing pickle module into your program followed by dump() method to write onto the file.

Practical Implementation-1

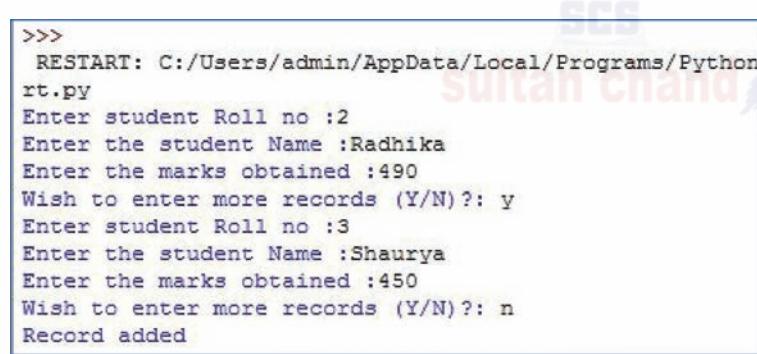
Program to insert/append a record in the binary file "student.dat"



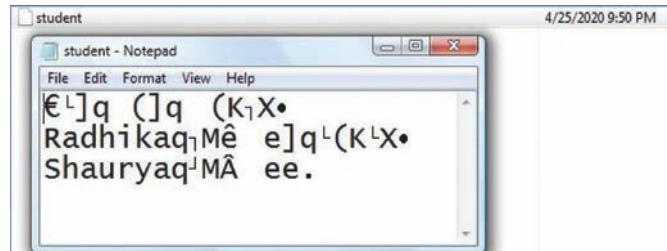
```
#Program for inserting/appending a record in a binary file- student

import pickle
record = []
while True:
    roll_no = int(input("Enter student Roll no :"))
    name = input("Enter the student Name :")
    marks = int(input("Enter the marks obtained :"))
    data = [roll_no, name, marks]
    record.append(data)
    choice = input("Wish to enter more records (Y/N) ?: ")
    if choice.upper() == 'N':
        break
f = open("D:/student", "wb")
pickle.dump(record, f)
print("Record added")
f.close()

Ln: 21 Col: 0
```

```
>>>
RESTART: C:/Users/admin/AppData/Local/Programs/Python
rt.py
Enter student Roll no :2
Enter the student Name :Radhika
Enter the marks obtained :490
Wish to enter more records (Y/N)?: y
Enter student Roll no :3
Enter the student Name :Shaurya
Enter the marks obtained :450
Wish to enter more records (Y/N)?: n
Record added
```

As shown in the above program, inserting and adding record in student file begins with importing pickle module. Then, iterating for the number of records to be added is done using while loop. Input is accepted from the user for roll number, name and marks. These fetched values are saved as a list to be appended. Once the record is created, it is appended to the binary file "student" using the dump() method, which writes the object into the opened file.

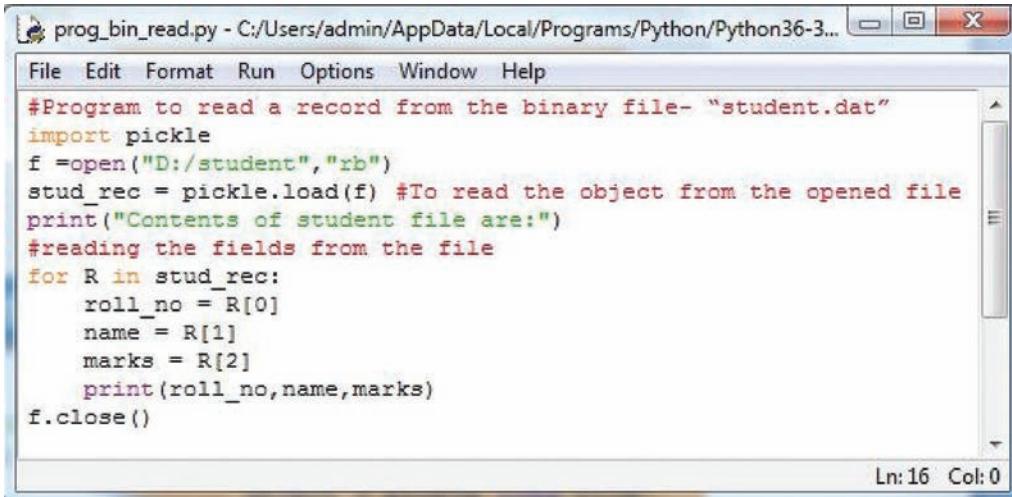
Finally the file is closed explicitly and the record is added into the student.dat file as shown in the output. Since binary file contains unreadable characters, it becomes difficult to read and understand, but directly used and understood by the computer.

Reading a record from a binary file

The following practical implementation illustrates how a record is read from a binary file.

Practical Implementation-2

Program to read a record from the binary file "student.dat"



```
#Program to read a record from the binary file- "student.dat"
import pickle
f =open("D:/student","rb")
stud_rec = pickle.load(f) #To read the object from the opened file
print("Contents of student file are:")
#reading the fields from the file
for R in stud_rec:
    roll_no = R[0]
    name = R[1]
    marks = R[2]
    print(roll_no,name,marks)
f.close()

>>>
RESTART: C:/Users/admin/AppData/Local/
.py
Contents of student file are:
2 Radhika 490
3 Shaurya 450
```

The above program deals with reading the contents from binary file student using load() method of pickle module. It is used to read the object from the opened file. The syntax for this is given by the statement—

```
object = pickle.load(file)
```

Once the contents are read, it gets stored inside the object, stud_rec. All the data of the respective fields from this object is held in the respective variables, roll_no, name and marks and are finally displayed as the output.

Searching a record in a binary file

Searching the binary file "student" is carried out on the basis of the roll number entered by the user. The file is opened in the read-binary mode and gets stored in the file object, f. load() method is used to read the object from the opened file. A variable 'found' is used which will tell the status of the search operation being successful or unsuccessful. Each record from the file is read and the contents of the field, roll no, is compared with the roll number to be searched. Upon the search being successful, appropriate message is displayed to the user as shown in the practical implementation that follows.

Practical Implementation-3

Program to search a record from the binary file "student.dat" on the basis of roll number.

```
# Program to search a record from the binary file- "student.dat"
import pickle
f =open("D:/student","rb")
stud_rec = pickle.load(f) #To read the object from the opened file
found = 0
rno =int(input("Enter the roll number to search :"))

for R in stud_rec:
    if R[0] == rno:
        print("Successful Search",R[1],"Found !")
        found = 1
        break
if found == 0:
    print("Sorry, record not found")
f.close()
```

```
>>>
RESTART: C:/Users/admin/AppData/Local/Programs/
ch.py
Enter the roll number to search :2
Successful Search Radhika Found !
```

Updating a record in a binary file

Updating a record in the file requires roll number to be fetched from the user whose name is to be updated.

```
# Program to update name of the student from binary file
import pickle
f =open("D:/student","rb+")
stud_rec = pickle.load(f) #To read the object from the opened file
found = 0
rollno =int(input("Enter the roll number to search :"))
for R in stud_rec:
    rno = R[0]
    if rno == rollno:
        print("Current Name is:",R[1])
        R[1]=input("New Name :");found = 1
        break

if found == 1:
    f.seek(0) #taking the file pointer to the beginning of the file
    pickle.dump(stud_rec,f)
    print("Name Updated!!!")
f.close()
```

```
>>>
RESTART: C:/Users/admin/AppData/Local/Programs/
te.py
Enter the roll number to search :2
Current Name is: Radhika
New Name :Sonia
Name Updated!!!
```

Once the record is found, the file pointer is moved to the beginning of the file using seek(0) statement, and then the changed name is written to the file and the record is updated. seek() method is used for random access to the file.

We will be learning more about it in the next section.

Practical Implementation-4

Write a menu-driven program to perform all the basic operations using dictionary on student binary file such as inserting, reading, updating, searching and deleting a record.

```
File Edit Format Run Options Window Help
import os
import pickle

# Accepting data for Dictionary
def insertRec():
    rollno = int(input("Enter roll number:"))
    name = input("Enter Name:")
    marks = int(input("Enter Marks:"))
    # Creating the Dictionary
    rec = {"Rollno": rollno, "Name": name, "Marks": marks}
    # Writing the Dictionary
    f = open("student.dat", "ab")
    pickle.dump(rec, f)
    f.close()

# Reading the records
def readRec():
    f = open("student.dat", "rb")
    while True:
        try:
            rec = pickle.load(f)
            print("Roll Num:", rec['Rollno'])
            print("Name:", rec["Name"])
            print("Marks:", rec["Marks"])
        except EOFError:
            break
    f.close()

# Searching a record based on Rollno
def searchRollNo(r):
    f = open("student.dat", "rb")
    flag = False
    while True:
        try:
            rec = pickle.load(f)
            if rec['Rollno'] == r:
                print("Roll Num:", rec['Rollno'])
                print("Name:", rec["Name"])
                print("Marks:", rec["Marks"])
                flag = True
        except EOFError:
            break
    if flag == False:
        print("No Record Found")
    f.close()

# Marks Modification for a RollNo
def updateMarks(r, m):
    f = open("student.dat", "rb")
    reclst = []
    while True:
        try:
            rec = pickle.load(f)
            reclst.append(rec)
        except EOFError:
            break
    f.close()
    for i in range(len(reclst)):
        if reclst[i]['Rollno'] == r:
            reclst[i]['Marks'] = m
    f = open("student.dat", "wb")
    for x in reclst:
        pickle.dump(x, f)
    f.close()
```



```

# Deleting a record based on RollNo
def deleteRec(r):
    f = open("student.dat", "rb")
    reclst = []
    while True:
        try:
            rec = pickle.load(f)
            reclst.append(rec)
        except EOFError:
            break
    f.close()
    f = open("student.dat", 'wb')
    for x in reclst:
        if x['Rollno'] == r:
            continue
        pickle.dump(x, f)
    f.close()

while True:
    print('Type 1 to insert rec.')
    print('Type 2 to display rec.')
    print('Type 3 to search rec.')
    print('Type 4 to update rec.')
    print('Type 5 to delete rec.')
    print('Enter your choice 0 to exit')
    choice = int(input("Enter your choice:"))
    if choice == 0:
        break
    elif choice == 1:
        insertRec()
    elif choice == 2:
        readRec()
    elif choice == 3:
        r = int(input("Enter a rollno to search:"))
        searchRollNo(r)
    elif choice == 4:
        r = int(input("Enter a rollno:"))
        m = int(input("Enter new Marks:"))
        updateMarks(r, m)
    elif choice == 5:
        r = int(input("Enter a rollno:"))
        deleteRec(r)

```

Ln: 53 Col: 32

```

>>>
RESTART: C:/Users/admin/AppData/Local/Programs/Python/
pertns.py
Type 1 to insert rec.
Type 2 to display rec.
Type 3 to Search RollNo.
Type 4 to update marks.
Type 5 to delete a Record.
Enter your choice 0 to exit
Enter you choice:1
Enter roll number:20
Enter Name:Ankur
Enter Marks400
Type 1 to insert rec.
Type 2 to display rec.
Type 3 to Search RollNo.
Type 4 to update marks.
Type 5 to delete a Record.
Enter your choice 0 to exit
Enter you choice:2
Roll Num: 20
Name: Ankur
Marks: 400
Type 1 to insert rec.
Type 2 to display rec.
Type 3 to Search RollNo.
Type 4 to update marks.
Type 5 to delete a Record.
Enter your choice 0 to exit
Enter you choice:3
Enter a rollno to search:20
Roll Num: 20
Name: Ankur
Marks: 400
Type 1 to insert rec.
Type 2 to display rec.
Type 3 to Search RollNo.

```

4.12 RANDOM ACCESS IN FILES USING TELL() AND SEEK()

Till now in all our programs we laid stress on the sequential processing of data in a text and binary file. But files in Python allow random access of the data as well using built-in methods seek() and tell().

seek()—seek() function is used to change the position of the file handle (file pointer) to a given specific position. File pointer is like a cursor, which defines from where the data has to be read or written in the file.

Python file method seek() sets the file's current position at the offset. This argument is optional and defaults to 0, which means absolute file positioning. Other values are: 1, which signifies seek is relative (may change) to the current position and 2 means seek is relative to the end of file. There is no return value.

The reference point is defined by the "from_what" argument. It can have any of the three values:

0: sets the reference point at the beginning of the file, which is by default.

1: sets the reference point at the current file position.

2: sets the reference point at the end of the file.

seek() can be done in two ways:

- Absolute Positioning
- Relative Positioning

Absolute referencing using seek() gives the file number on which the file pointer has to position itself. The syntax for seek() is—

```
f.seek(file_location)      #where f is the file pointer
```

For example, f.seek(20) will give the position or file number where the file pointer has been placed. This statement shall move the file pointer to 20th byte in the file no matter where you are.

Relative referencing/positioning has two arguments, offset and the position from which it has to traverse. The syntax for relative referencing is:

```
f.seek(offset, from_what)  #where f is file pointer
```

For example,

f.seek(-10,1) from current position, move 10 bytes backward

f.seek(10,1) from current position, move 10 bytes forward

f.seek(-20,1) from current position, move 20 bytes backward

f.seek(10,0) from beginning of file, move 10 bytes forward

POINT TO REMEMBER

This is to be kept in mind that Python 3.x only supports text file seeks from the beginning of the file. seek() with negative offset only works when the file is opened in binary mode.

tell()—tell() returns the current position of the file read/write pointer within the file. Its syntax is:

```
f.tell()      #where f is file pointer
```

- When you open a file in reading/writing mode, the file pointer rests at 0th byte.
- When you open a file in append mode, the file pointer rests at last byte.

This is illustrated in the practical implementation that follows:




```
prog_seek.py - C:/Users/admin/AppData/Local/Programs/Python/3.8/python.exe prog_seek.py
File Edit Format Run Options Window Help
#Illustrating seek() and tell()
#Reading byte by byte
f = open("D:/test.txt")
print("Before reading: ",f.tell())
s = f.read()
print("After reading: ",f.tell())
f.seek(0)
#Brings file pointer to the 0th byte so data can
#be written/read from next byte
print("From beginning again:",f.tell())
s = f.read(4)
print("First 4 bytes are:",s)
print(f.tell())
s = f.read(3)
print("next 3 bytes",s)
print(f.tell())
f.close()

Ln: 20 Col: 0
```

Contents of “test.txt”

```
>>>
RESTART: C:/Users/admin/AppData/Local/Programs/Python/3.8/python.exe k.py
Before reading: 0
After reading: 48
From beginning again: 0
First 4 bytes are: This
4
next 3 bytes is
7
```



4.13 INTRODUCTION TO CSV

In every span of today's organizational working environment, data sharing is one of the major tasks to be carried out, largely through spreadsheets or databases.

A basic approach to share data is through the comma separated values (CSV) file.

CSV is a simple flat file in a human readable format which is extensively used to store tabular data, in a spreadsheet or database. A CSV file stores tabular data (numbers and text) in plain text.



Files in the CSV format can be imported to and exported from programs that store data in tables, such as Microsoft Excel or OpenOffice Calc.

Already defined, CSV stands for “**comma separated values**”. Thus, we can say that a comma-separated file is a delimited text file that uses a comma to separate values.

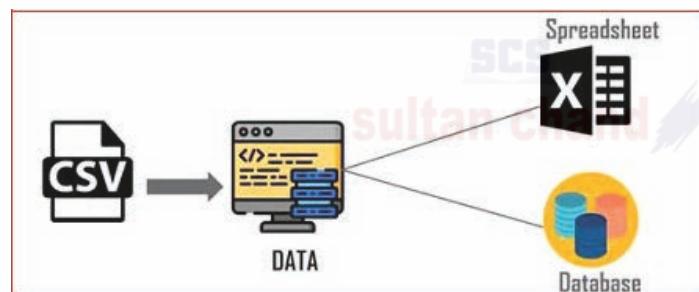


Fig. 4.1: CSV Storage Format



Fig. 4.2: Features of a CSV file

Each line in a file is known as **data/record**. Each record consists of one or more fields, separated by commas (also known as delimiters), i.e., each of the records is also a part of this file. Tabular data is stored as **text** in a CSV file. The use of comma as a field separator is the source of the name for this file format. It stores our data into a spreadsheet or a database.

CTM: The most commonly used delimiter in a CSV file is usually a **comma**.

4.14 WHY USE CSV

With the use of social networking sites and its various associated applications being extensively used requires the handling of huge data. The problem arises as to how to handle and organize this large unstructured data as shown in the Fig. 4.3(a) given below.

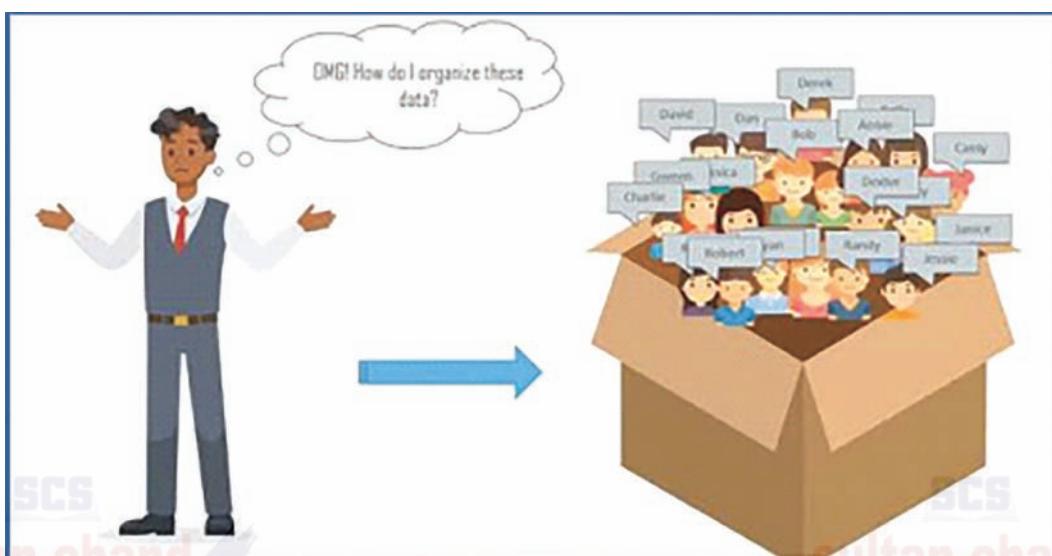


Fig. 4.3(a): Problem of Huge Data

The solution to the above problem is CSV (Fig. 4.3(b)). Thus, the CSV organizes data into a structured form and, hence, the proper and systematic organization of this large amount of data is done by CSV. Since CSV files formats are of plain text format, it makes it very easy for website developers to create applications that implement CSV.

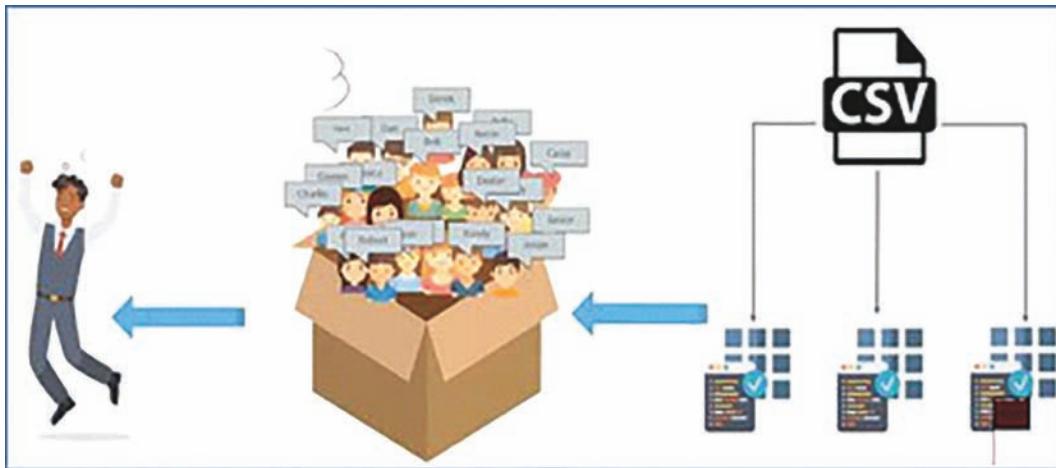


Fig. 4.3(b): CSV as a Solution

CSV files are commonly used because they are easy to read and manage, small in size, and fast to process/transfer. Because of these salient features, they are frequently used in software applications, ranging anywhere from online e-commerce stores to mobile apps to desktop tools. *For example*, Magento, an e-commerce platform, is known for its support of CSV.

Thus, in a nutshell, the several advantages that are offered by CSV files are as follows:

- CSV is faster to handle.
- CSV is smaller in size.
- CSV is easy to generate and import onto a spreadsheet or database.
- CSV is human readable and easy to edit manually.
- CSV is simple to implement and parse.
- CSV is processed by almost all existing applications.

After understanding the concept and importance of using CSV files, we will now discuss the read and write operations on CSV files.

4.15 CSV FILE HANDLING IN PYTHON

For working with CSV files in Python, there is an inbuilt module called **CSV**. It is used to read and write tabular data in CSV format.

Therefore, to perform read and write operation with CSV file, we must import **CSV module**. CSV module can handle CSV files correctly regardless of the operating system on which the files were created.

Along with this module, `open()` function is used to open a CSV file, and return file object. We load the module in the usual way using `import`:

```
>>> import csv
```

Like other files (text and binary) in Python, there are two basic operations that can be carried out on a CSV file.

1. Reading a CSV
2. Writing to a CSV.

Let us discuss these CSV operations.

4.15.1 Reading from CSV File

Reading from a CSV file is done using the reader object. The CSV file is opened as a text file with Python's built-in open() function, which returns a file object. This creates a special type of object to access the CSV file (reader object), using the reader() function.

The reader object is an iterable that gives us access to each line of the CSV file as a list of fields. You can also use next() directly on it to read the next line of the CSV file, or you can treat it like a list in a for loop to read all the lines of the file (as lists of the file's fields).

This is shown in the practical implementation given below.

Before this, enter the student details in spreadsheet and save this file as shown:

Next step is to open the Notepad and enter the data for student.csv, which will be the equivalent for student.xls.

The screenshot shows two windows side-by-side. On the left is Microsoft Excel with a spreadsheet titled 'student.xls'. The spreadsheet has three columns: 'Name', 'Class', and 'Marks'. The data consists of 12 rows of student information. On the right is a Windows Notepad window titled 'student.csv' containing the same data as a CSV file. A red arrow points from the Excel window to the Notepad window, indicating the equivalence between the two formats.

	Name	Class	Marks
1			
2	Name	Class	Marks
3	Rinku	XII	90
4	Veenu	XI	88
5	Ajay	XI	78
6	Sushant	X	70
7	Radhika	IX	80
8	Shaurya	X	90
9	Deepak	XI	60
10	Jatin	XII	70
11	Himesh	XII	78
12	Vinay	XI	99

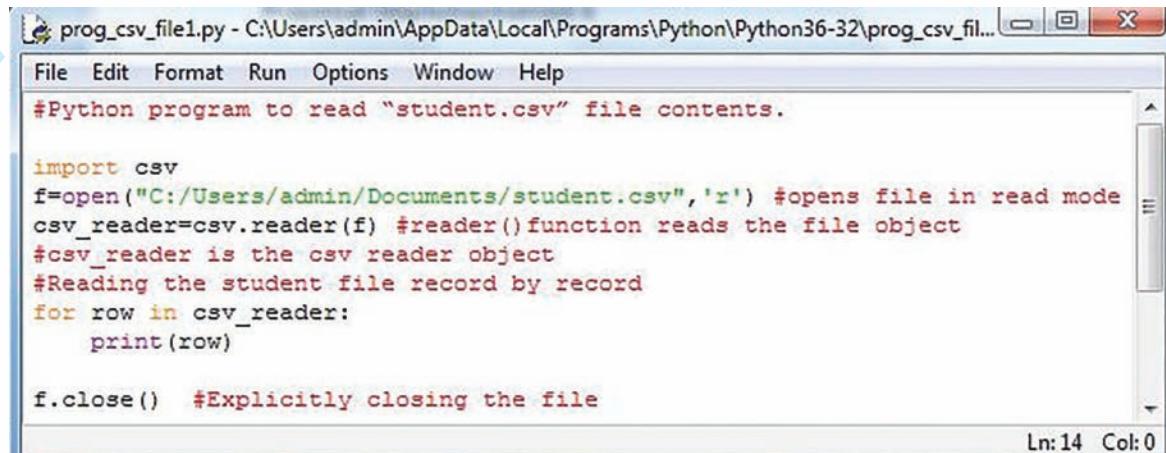
student.xls

student.csv

In student.csv (notepad) file, first line is the header and remaining lines are the data/records. The fields are separated by comma, or we may say the separator character. In general, the separator character is called a delimiter, and the comma is not the only one used. Other popular delimiters include the tab (\t), colon (:) and semi-colon (;) characters.

Practical Implementation-5

Write a program to read the contents of “student.csv” file.



```
# Python program to read "student.csv" file contents.

import csv
f=open("C:/Users/admin/Documents/student.csv",'r') #opens file in read mode
csv_reader=csv.reader(f) #reader() function reads the file object
#csv_reader is the csv reader object
#Reading the student file record by record
for row in csv_reader:
    print(row)

f.close() #Explicitly closing the file
```




```
>>>
RESTART: C:/Users/admin/AppData/Local/Programs/1.py
['Name', 'Class', 'Marks']
['Rinku', 'XII', '90']
['Veenu', 'XI', '88']
['Ajay', 'XI', '78']
['Sushant', 'X', '70']
['Radhika', 'IX', '80']
['Shaurya', 'X', '90']
['Deepak', 'XI', '60']
['Jatin', 'XII', '70']
['Himesh', 'XII', '78']
['Vinay', 'XI', '99']
[]
```

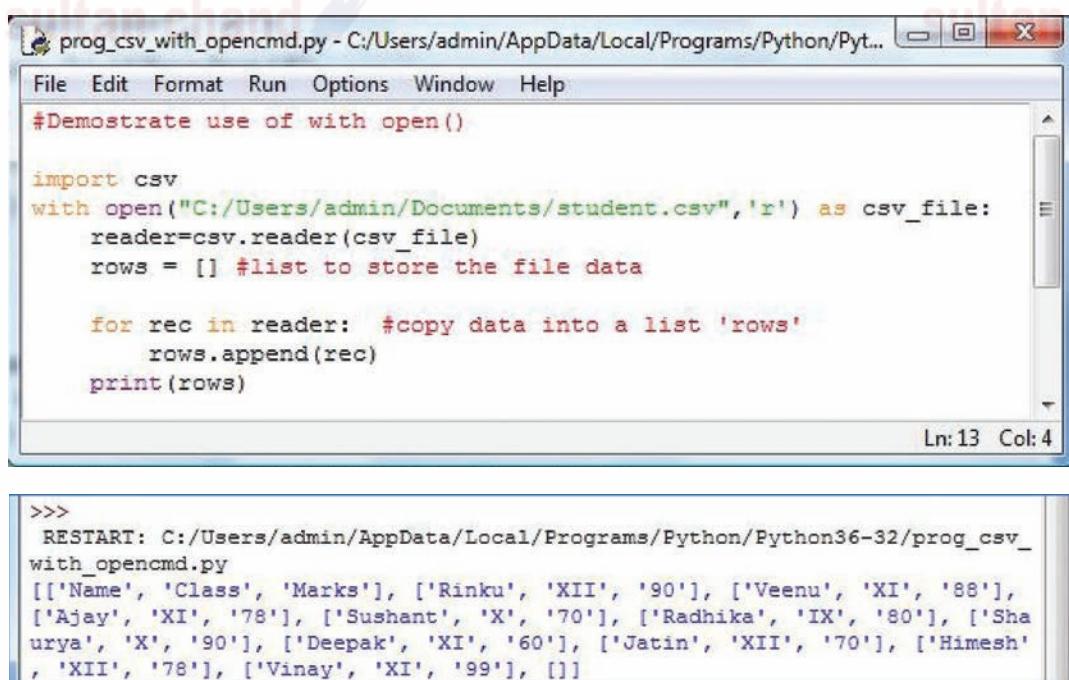
Explanation:

As seen from the above output, every record is stored in reader object in the form of a List. In the above code, we first open the CSV file in READ mode. The file object is named as **f**. The file object is converted to **csv.reader** object. We save the **csv.reader** object as **csv_reader**. The reader object is used to read records as lists from a csv file. Now, we iterate through all the rows using a for loop. When we try to print each row, one can find that row is nothing but a list containing all the field values. Thus, all the records are displayed as lists separated by comma.

In the next implementation, we will count the number of records present inside the student.csv file.

Practical Implementation-6

Write a program to read the contents of “student.csv” file using with open().



```
#Demostrate use of with open()

import csv
with open("C:/Users/admin/Documents/student.csv",'r') as csv_file:
    reader=csv.reader(csv_file)
    rows = [] #list to store the file data

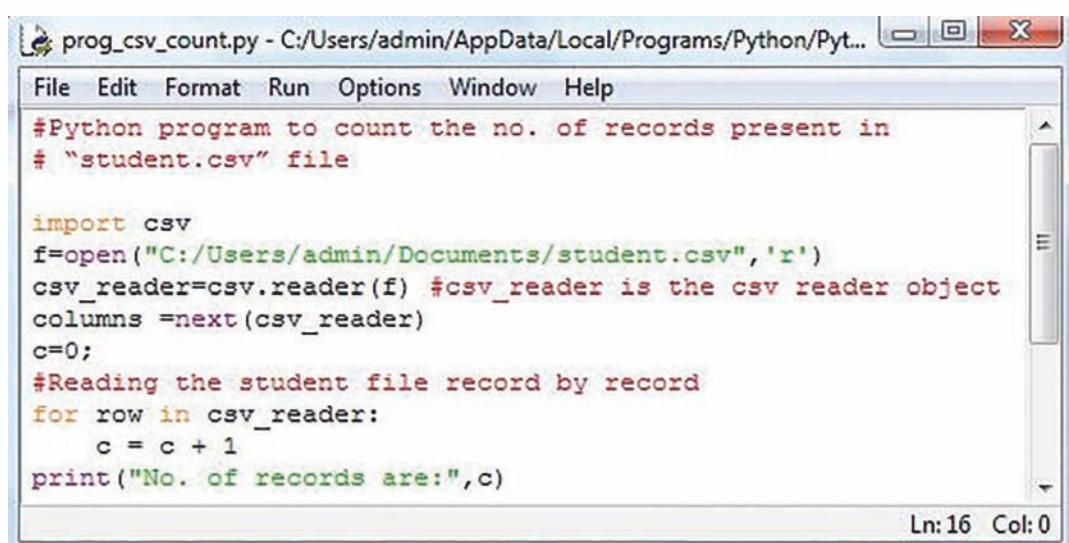
    for rec in reader: #copy data into a list 'rows'
        rows.append(rec)
print(rows)
```

>>>
RESTART: C:/Users/admin/AppData/Local/Programs/Python/Python36-32/prog_csv_with_opencmd.py
[['Name', 'Class', 'Marks'], ['Rinku', 'XII', '90'], ['Veenu', 'XI', '88'], ['Ajay', 'XI', '78'], ['Sushant', 'X', '70'], ['Radhika', 'IX', '80'], ['Shurya', 'X', '90'], ['Deepak', 'XI', '60'], ['Jatin', 'XII', '70'], ['Himesh', 'XII', '78'], ['Vinay', 'XI', '99'], []]

The above modified code uses “with open()” function, the only difference being that the file being opened using with open() gets automatically closed after the program execution gets over, unlike open() where we need to give close() statement explicitly.

Practical Implementation-7

Write a program to count the number of records present in “student.csv” file.



```
#Python program to count the no. of records present in
# "student.csv" file

import csv
f=open("C:/Users/admin/Documents/student.csv",'r')
csv_reader=csv.reader(f) #csv_reader is the csv reader object
columns =next(csv_reader)
c=0;
#Reading the student file record by record
for row in csv_reader:
    c = c + 1
print("No. of records are:",c)
```

>>>
RESTART: C:/Users/admin/AppData/Local/t.py
No. of records are: 11

Explanation:

In the above program, a special type of object is created to access the CSV file (reader object), which is csv_reader using the reader() function. The reader object is an iterable that gives us access to each line of the CSV file as a list of fields. The function next() is used to directly point to this list of fields to read the next line in the CSV file. .next() method returns the current row and advances the iterator to the next row.

The variable 'c' is used as a counter variable to count the number of rows/records present in this file, which is finally printed and thus the output is so obtained.

One of the important observations from the output is the number of records which are being displayed as 11 instead of 10. This is so because the header (first line) in the student csv file is also treated as a record only. This limitation is overcome in the next implementation.

CTM: .next() method returns the current row and advances the iterator to the next row.

Practical Implementation-8

Program to count the exact number of records present in the csv file excluding the header.

```
#Python program to count the no. of records present in
# "student.csv" file
import csv
f=open("C:/Users/admin/Documents/student.csv",'r')
csv_reader=csv.reader(f) #csv_reader is the csv reader object
csvRows = []
value=0
for row in csv_reader:
    if csv_reader.line_num == 1: #skip first row
        continue
    csvRows.append(row)
    value = len(list(csv_reader))
print("No. of records are:",value)
f.close()
```

```
>>>
RESTART: C:/Users/admin/AppData/Local/
t.py
No. of records are: 10
```

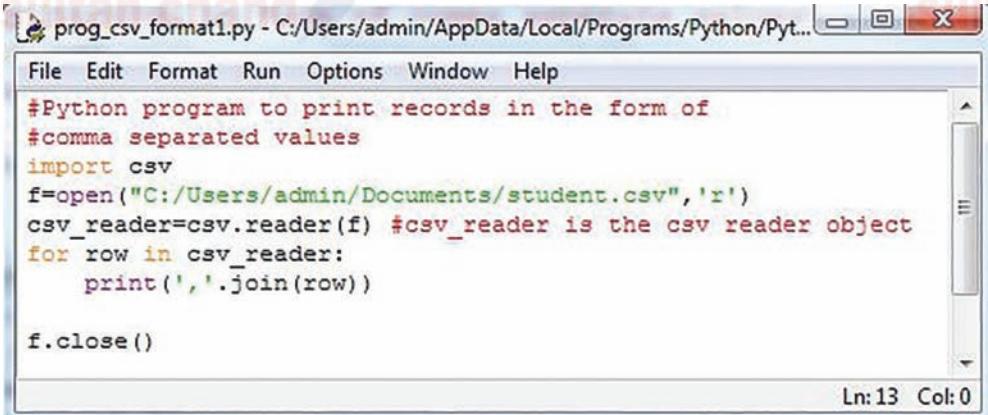
In the above program we have used line_num object of CSV file. Our csv_reader_object has a method called line_num that returns the number of lines in our CSV.

Then, if statement checks if the line is first line or not. If the condition is true, i.e., if it is the header line, then it is ignored using continue statement and the counting of records is resumed from second line onwards. Also, line_num object always stores the current line in consideration and, hence, the correct output for 10 records is so obtained.

CTM: line_num is nothing but a counter which returns the number of rows which have been iterated.

Practical Implementation-9

Program to print the records in the form of comma separated values, instead of lists.



```
File Edit Format Run Options Window Help
#Python program to print records in the form of
#comma separated values
import csv
f=open("C:/Users/admin/Documents/student.csv",'r')
csv_reader=csv.reader(f) #csv_reader is the csv reader object
for row in csv_reader:
    print(','.join(row))

f.close()

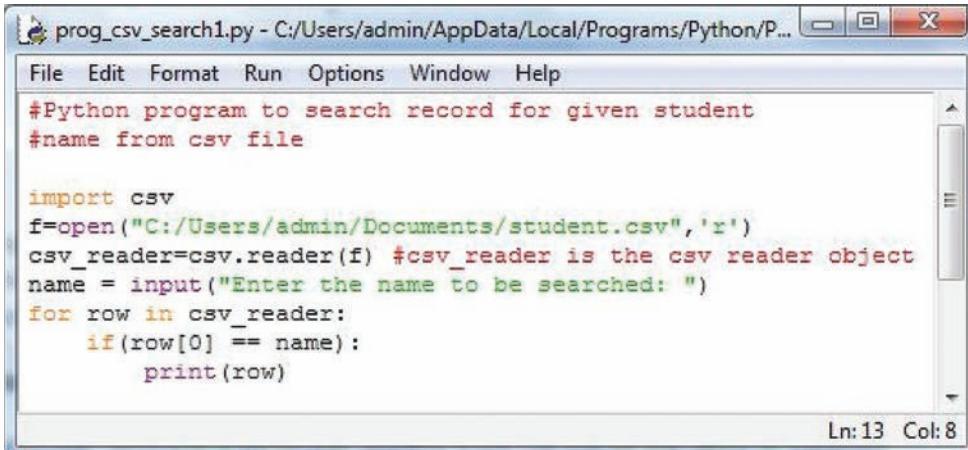
Ln:13 Col:0
```

```
>>>
RESTART: C:/Users/admin/AppData/Local/
ati.py
Name,Class,Marks
Rinku,XII,90
Veenu,XI,88
Ajay,XI,78
Sushant,X,70
Radhika,IX,80
Shaurya,X,90
Deepak,XI,60
Jatin,XII,70
Himesh,XII,78
Vinay,XI,99
```

In the above program, we have used a new function `join()`. `join()` is a string method that joins all values of each row with comma separator. Thus, all the records are displayed as a string separated by a comma separator and not as a list and hence the output is so obtained.

Practical Implementation-10

Program to search the record of a particular student from CSV file on the basis of inputted name.



```
File Edit Format Run Options Window Help
#Python program to search record for given student
#name from csv file

import csv
f=open("C:/Users/admin/Documents/student.csv",'r')
csv_reader=csv.reader(f) #csv_reader is the csv reader object
name = input("Enter the name to be searched: ")
for row in csv_reader:
    if(row[0] == name):
        print(row)

Ln:13 Col:8
```

```
>>>
RESTART: C:/Users/admin/AppData/Local/Programs/
ch1.py
Enter the name to be searched: Jatin
['Jatin', 'XII', '70']
```

Till now we have covered the basics of how to use the CSV module to read the contents of a CSV file. Now, we will discuss about how to write to a CSV file in python.

4.15.2 Writing to CSV File

To write to a CSV file in Python, we can use the **csv.writer()** function. The csv.writer() function returns a writer object that converts the user's data into a delimited string. This string can later be used to write into CSV files using the writerow() function.

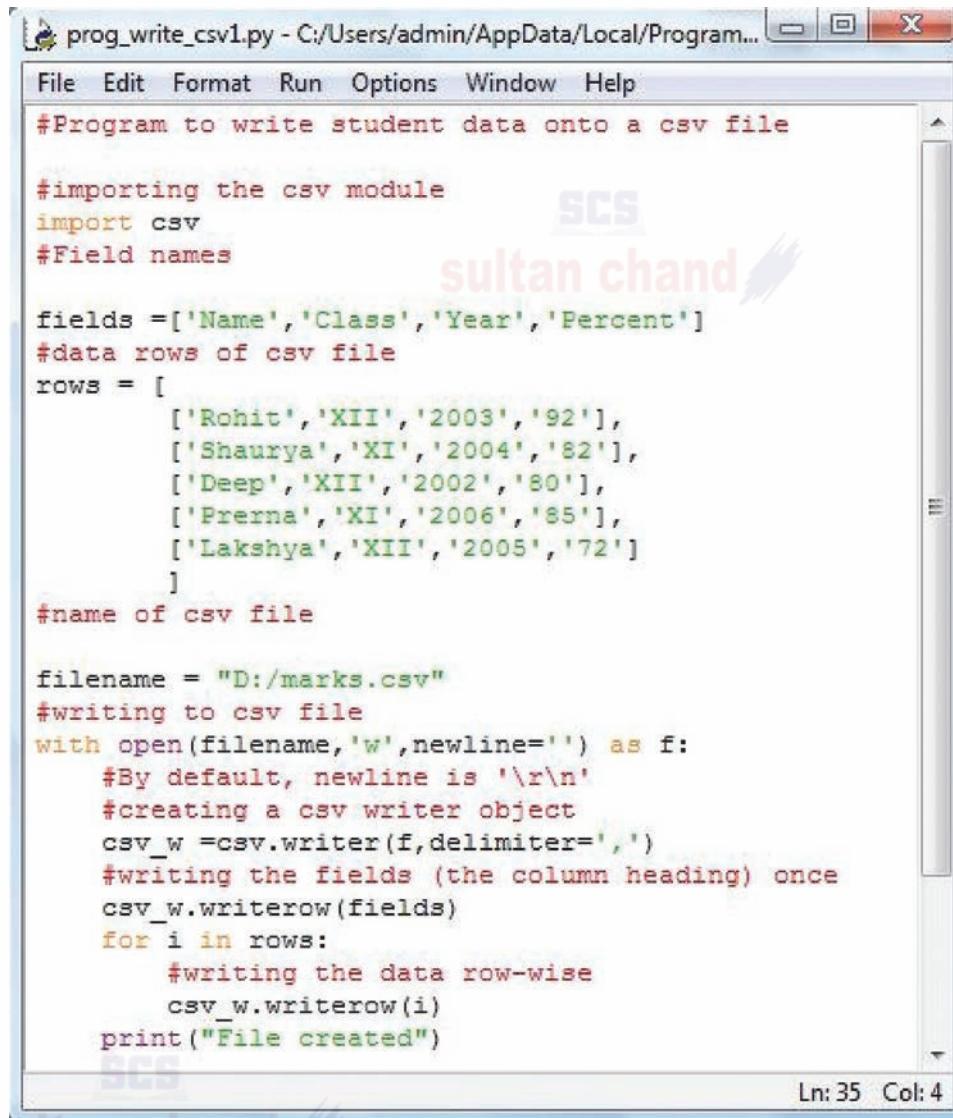
In order to write to a CSV file, we create a special type of object to write to the CSV file "**writer object**", which is defined in the CSV module, and which we create using the writer() function.

The **writerow()** method allows us to write a list of fields to the file. The fields can be strings or numbers or both. Also, while using writerow(), you do not need to add a new line character (or other EOL indicator) to indicate the end of the line, writerow() does it for you as necessary.

Let us write data onto a CSV file using writerow() method.

Practical Implementation-11

Program to write data onto "student" CSV file using writerow() method.

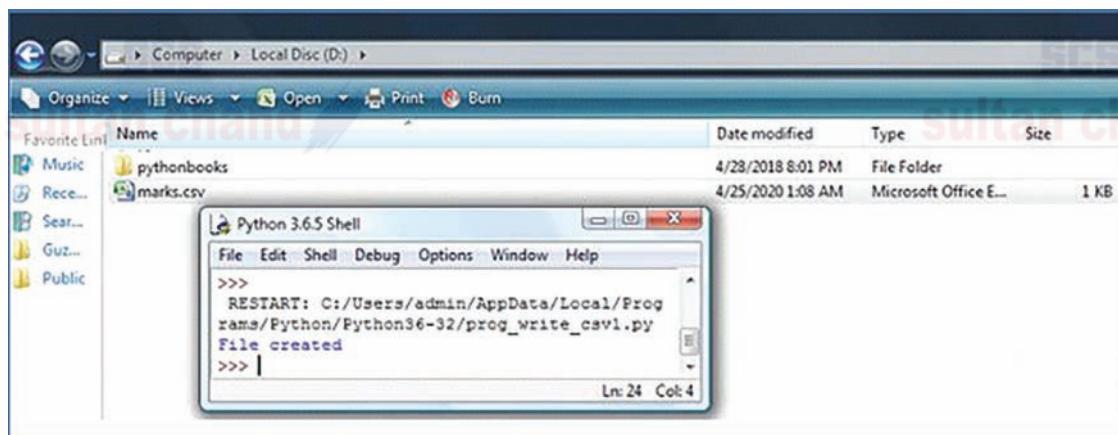


```
prog_write_csv1.py - C:/Users/admin/AppData/Local/Program... X
File Edit Format Run Options Window Help
#Program to write student data onto a csv file

#importing the csv module
import csv
#Field names

fields =['Name','Class','Year','Percent']
#data rows of csv file
rows = [
    ['Rohit','XII','2003','92'],
    ['Shaurya','XI','2004','82'],
    ['Deep','XII','2002','80'],
    ['Prerna','XI','2006','85'],
    ['Lakshya','XII','2005','72']
]
#name of csv file

filename = "D:/marks.csv"
#writing to csv file
with open(filename,'w',newline='') as f:
    #By default, newline is '\r\n'
    #creating a csv writer object
    csv_w =csv.writer(f,delimiter=',')
    #writing the fields (the column heading) once
    csv_w.writerow(fields)
    for i in rows:
        #writing the data row-wise
        csv_w.writerow(i)
print("File created")
Ln: 35 Col: 4
```



Contents of "marks.csv" created:

	A	B	C	D	E	F	G
1	Name	Class	Year	Percent			
2	Rohit	XII	2003	92			
3	Shaurya	XI	2004	82			
4	Deep	XII	2002	80			
5	Prerna	XI	2006	85			
6	Lakshya	XII	2005	72			
7							
8							

Explanation:

In the above program, the very first line is for importing csv file into your program. Next, whatever are the column headings for our data are mentioned as a list in variable called **fields**. All the data stored inside these fields are placed inside variable **rows**.

Now give the name of your file, let us say, **student.csv** and that will be created and stored inside your current working directory or the path that you mentioned (as we have given for D:/) for the attribute "**filename**".

'w' stands for write mode and we are using the file by opening it using "**with open**", since using **with open** does not require file to be closed explicitly. The next statement comprises the most important function used for writing onto csv file, viz., **csv.writer()**, to obtain a writer object and store it in the variable **csv_w** as the name of the variable and this is the CSV object. **writer()** takes the name of file object 'f' as the argument. By default, the delimiter is comma (,).

writerow(fields) is going to write the fields which are the column headings into the file and have to be written only once. Using for loop, rows are traversed from the list of rows from the file. writerow(i) is writing the data row-wise in the for loop and in the last the file is automatically closed.

Also, while giving csv.writer(), the delimiter taken is comma. We can change the delimiter whenever and wherever required by changing the argument passed to delimiter attribute.

For example, delimiter = "|" (pipe symbol). You can put any character as delimiter and if nothing is given, comma is placed by default.

writerow() method is used to write each row.

In this program, we have used for loop for writing data row-wise onto the file using writerow() method. We can avoid using for loop and can write all the rows/records in one go.

This can be done by using writerows() method. writerows() writes all the rows in one go, so you need not use for loop and iterations.

Practical Implementation-12

Program to write data onto "student" csv file using writerows() method (modification of Practical Implementation -11).

```
prog_write_csv1.py - C:/Users/admin/AppData/Local/Programs/Pyth... X
File Edit Format Run Options Window Help
#Program to write student data onto a csv file
# importing the csv module
import csv
#Field names
fields =['Name','Class','Year','Percent']
#data rows of csv file
rows = [
    ['Rohit','XII','2003','92'],
    ['Shaurya','XI','2004','82'],
    ['Deep','XII','2002','80'],
    ['Frerna','XI','2006','85'],
    ['Lakshya','XII','2005','72']
]
#name of csv file
filename = "D:/newmarks.csv"
#writing to csv file
with open(filename,'w',newline='') as f:
    #By default, newline is '\r\n'
    #creating a csv writer object
    csv_w =csv.writer(f,delimiter=',')
    #writing the fields (the column heading) once
    csv_w.writerow(fields)
    #writing the rows all at once
    csv_w.writerows(rows)
    print("All rows written in one go")
Ln: 33 Col: 4

>>>
RESTART: C:/Users/admin/AppData/Local/Programs/Python36-32/prog_write_csv1.py
All rows written in one go
```

OBJECTIVE TYPE QUESTIONS

1. Fill in the blanks.

- (a) format is a text format, accessible to all applications across several platforms.
(b) method is used for random access of data in a CSV file.
(c) method of pickle module is used to write an object into binary file.
(d) method of pickle module is used to read data from a binary file.
(e) statement is given for importing csv module into your program.
(f) is a string method that joins all values of each row with comma separator in CSV.
(g) object contains the number of the current line in a CSV file.
(h) To end all the file contents in form of a list, method is used.
(i) To read all the file contents, method may be used.
(j) To force Python to write the contents of file buffer on to storage file, method may be used.

Answers: (a) CSV (b) seek() (c) dump()
(d) load() (e) import csv (f) join()
(g) line_num (h) readlines() (i) writelines()
(j) flush()

2. State whether the following statements are True or False.

- (a) CSV module can handle CSV files correctly regardless of the operating system on which the files were created.
(b) CSV module gets automatically imported into your program for reading a CSV file.
(c) The type of operation that can be performed on a file depends upon the file mode in which it is opened.
(d) Functions readline() and readlines() are essentially the same.
(e) Every record in a CSV file is stored in reader object in the form of a list.
(f) writerow() method allows us to write a list of fields to the CSV file.
(g) Comma is the default delimiter for a CSV file.
(h) tell() method of Python tells us the current position within the file.
(i) The offset argument to seek() method indicates the number of bytes to be moved.
(j) If the offset value is set to 2, beginning of the file would be taken as seek position.

Answers: (a) True (b) False (c) True (d) False (e) True (f) True
(g) True (h) True (i) True (j) False

3. Multiple Choice Questions (MCQs)

- (a) Which of the following is not a valid mode to open a file?
(i) ab (ii) rw (iii) r+ (iv) w+
- (b) Which statement is used to change the file position to an offset value from the start?
(i) fp.seek(offset, 0) (ii) fp.seek(offset, 1)
(iii) fp.seek(offset, 2) (iv) None of the above
- (c) The difference between r+ and w+ modes is expressed as?
(i) No difference
(ii) In r+ mode, the pointer is initially placed at the beginning of the file and the pointer is at the end for w+
(iii) In w+ mode, the pointer is initially placed at the beginning of the file and the pointer is at the end for r+
(iv) Depends on the operating system
- (d) What does CSV stand for?
(i) Cursor Separated Variables (ii) Comma Separated Values
(iii) Cursor Separated Values (iv) Cursor Separated Version
- (e) Which module is used for working with CSV files in Python?
(i) random (ii) statistics (iii) csv (iv) math



SOLVED QUESTIONS

1. Write a Python code to find the size of the file in bytes, number of lines and number of words.

Ans. # reading data from a file and find size, lines, words

```
f = open('Lines.txt', 'r')
str = f.read()
size = len(str)
print('size of file n bytes ', size)
f.seek(0)
L = f.readlines()
word = L.split()
print('Number of lines ', len(L))
print('Number of words ', len(word))
f.close()
```

2. Consider the following code:

```
f = open("test", "w+")
f.write("0123456789abcdef")
f.seek(-3,2)           //Statement 1
print(f.read(2))       //Statement 2
```

Explain statement 1 and give output of statement 2.

Ans. Statement 1 uses seek() method that can be used to position the file object at a particular place in the file.

It's syntax is:

```
fileobject.seek(offset [, from what])
```

So, `f.seek(-3,2)` positions the fileobject to 3 bytes before end of file.

Output of Statement 2 is:

de

It reads 2 bytes from where the file object is placed.

3. Yogendra intends to position the file pointer to the beginning of a text file. Write Python statement for the same assuming "F" is the Fileobject.

Ans. F.seek(0)

4. Differentiate between file modes r+ and rb+ with respect to Python.

Ans. `r+` opens a file for both reading and writing. The file pointer placed at the beginning of the file.

`rb+` opens a file for both reading and writing in binary format. The file pointer is placed at the beginning of the file.

5. In which of the following file modes the existing data of the file will not be lost?

In which of the following are m , a , b , w , $w+b$, $a+b$, wb , $wb+$, $w+$, $r+$

Ans. In file modes `rb`, `ab`, `a+b` and `r+`, data will not be lost.

In file modes `w`, `w+b`, `wh`, `wbt` and `w+`, data will be truncated i.e. lost.

6. Write a statement in Python to perform the following operations:

- (a) To open a text file "Book.txt" in read mode
- (b) To open a binary file "Book.dat" in write mode

Ans. (a) `f = open("Book.txt", "r")`
(b) `f = open("Book.dat", "wb")`

7. What is a CSV file?

Ans. CSV (Comma Separated Values) is a simple file format used to store tabular data, such as a spreadsheet or database. A CSV file stores tabular data (numbers and text) in plain text.

8. What are the advantages of CSV file formats?

Ans. Advantages:

- (a) A simple, compact and ubiquitous format for data storage.
- (b) A common format for data interchange.
- (c) It can be opened in popular spreadsheet packages like MS-Excel, Cal, etc.
- (d) Nearly all spreadsheets and databases support import/export to csv format.

9. Differentiate between a text file and a binary file.

Ans. A text file stores data as ASCII/UNICODE characters whereas a binary file stores data in binary format (as it is stored in memory). Internal conversion is required in text file and, hence, slower but binary file does not need any translation and is faster.

10. Write a program to add (append) Employee records onto a csv file.

```
Ans. import csv
with open('myfile.csv',mode = 'a') as csvfile:
    mywriter = csv.writer(csvfile, delimiter = ',')
    ans = 'y'
    while ans.lower() == 'y':
        eno = int(input("Enter Employee Number:"))
        name = input("Enter Employee Name:")
        salary = int(input("Enter Employee Salary:"))
        mywriter.writerow([eno, name, salary])
        print("## Data Saved... ##")
        ans = input("Add More?")
```

11. Write user-defined functions to perform read and write operation onto a student.csv file having fields as roll number, name, stream and marks.

```
Ans. import csv
row = ['2', 'Akshat Chauhan', 'Commerce', '98']
def readcsv():
    with open("D:/student.dat", 'r') as f:
        data = csv.reader(f)
        # reader function to generate a reader object
        for row in data:
            print(row)
def writecsv():
    with open("D:/student.dat", 'w', newline='') as fobj:
        # write new record in file
        csv_w = csv.writer(fobj, delimiter=',')
        csv_w.writerow(row)
print("Press-1 to Read Data and Press-2 to Write data: ")
a = int(input())
if a == 1:
    readcsv()
elif a == 2:
    writecsv()
else:
    print("Invalid value")
```

- 12.** Write a Python program to read specific columns from a "department.csv" file and print the content of the columns, department id and department name.

```
Ans. import csv
with open('departments.csv', newline=' ') as csvfile:
    data = csv.reader(csvfile)
    print("ID Department Name")
    print("-----")
    for row in data:
        print(row['department_id'], row['department_name'])
```

- 13.** Explain briefly the CSV format of storing files.

Ans. The acronym CSV is short for Comma-Separated Values, which refers to a tabular data saved as plaintext where data values are separated by commas. In CSV format:

- ☞ Each row of the table is stored in one row, i.e., the number of rows in a CSV file are equal to the number of rows in the table, (or sheet or database table etc.).
- ☞ The field values of a row are stored together with commas after every field value; but after the last field's value, no comma is given, just the end of line.

- 14.** Write a menu-driven program implementing user-defined functions to perform different functions on a csv file "student" such as:

- Write a single record to csv
- Write all the records in one single go onto the csv.
- Display the contents of the csv file.

```
Ans. import csv
# To create a CSV File by writing individual lines
def CreateCSV1():
    # Open CSV File
    Csvfile = open('student.csv', 'w', newline=' ')
    # CSV Object for writing
    Csvobj = csv.writer(Csvfile)
    while True:
        Rno = int(input("Rno:"))
        Name = input("Name:")
        Marks = float(input("Marks:"))
        Line = [Rno, Name, Marks]
        # Writing a line in CSV file
        Csvobj.writerow(Line)
        Ch = input("More(Y/N)?")
        if Ch == 'N':
            break
    Csvfile.close()           # Closing a CSV File

# To create a CSV File by writing all lines in one go
def CreateCSV2():
    # Open CSV File
    Csvfile = open('student.csv', 'w', newline=' ')
    # CSV Object for writing
    Csvobj = csv.writer(Csvfile)
    Lines = []
    while True:
        Rno = int(input("Rno:"))
        Name = input("Name:")
        Marks = float(input("Marks:"))
        Lines.append([Rno, Name, Marks])
        Ch = input("More(Y/N)?")
        if Ch == 'N':
            break
```



```
# Writing all lines in CSV file
Csvobj.writerows(Lines)
Csvfile.close()           # Closing a CSV File

# To read and show the content of a CSV File
def ShowAll():
    # Opening CSV File for reading
    Csvfile = open('student.csv', 'r', newline='')
    # Reading the CSV content in object
    Csvobj = csv.reader(Csvfile)
    for Line in Csvobj:      # Extracting line by line content
        print(Line)
    Csvfile.close()           # Closing a CSV File

print("CSV File Handling")
while True:
    Option = input("1:CreateCSV 2:CreateCSVAll 3>ShowCSV 4:Quit ")
    if Option == "1":
        CreateCSV1()
    elif Option == "2":
        CreateCSV2()
    elif Option == "3":
        ShowAll()
    else:
        break
```

- 15.** (a) Create a binary file “employee” that stores the records of employees and display them one by one.
(b) Display the records of all those employees who are getting salaries between 25000 to 30000.

Ans. (a) import pickle

```
f1 = open('emp.dat', 'rb')
e = pickle.load(f1)
for x in e:
    print(x)
f1.close()

(b) import pickle
f1 = open('emp.dat', 'rb')
e = pickle.load(f1)
for x in e:
    if(e[x]>=25000 and e[x]<=30000):
        print(x)
f1.close()
```

- 16.** What is the output of the following code?

```
fh = open("test.txt", "r")
Size = len(fh.read())
print(fh.read(5))
```

Ans. No output.

Explanation. The fh.read() of line 2 will read the entire file content and place the file pointer at the end of file. For the fh.read(5), it will return nothing as there are no bytes to be read from EOF. Thus print() statement prints nothing.

UNSOLVED QUESTIONS

1. Write appropriate statements to do the following:
 - To open a file named "RESULT.DAT" for output.
 - To go to the end of the file at any time.

2. Write a program to add two more employees details to the file "emp.txt" already stored in disk.

3. How are the following codes different from one another?

(a) `fp = open("file.txt", 'r')`
`fp.read()`

(b) `fp=open("file.txt", 'r')`

4. What is the output of the following code fragment? Explain.

```
fout = file("output.txt", 'w')
fout.write("Hello, world!\n")
fout.write("How are you?")
fout.close()
file("output.txt").read()
```

5. Write the output of the following code with justification if the contents of the file ABC.txt are:

```
Welcome to Python Programming!
f1 = file("ABC.txt", "r")
size = len(f1.read())
print(size)
data = f1.read(5)
print(data)
```

6. Anant has been asked to display all the students who have scored less than 40 for Remedial Classes. Write a user-defined function to display all those students who have scored less than 40 from the binary file "Student.dat".

7. Give the output of the following snippet:

```
import pickle
list1, list2 = [2, 3, 4, 5, 6, 7, 8, 9, 10], []
for i in list1:
    if (i%2==0 and i%4==0):
        list2.append(i)
f = open("bin.dat", "wb")
pickle.dump(list2, f)
f.close()
f = open("bin.dat", "rb")
data = pickle.load(f)
f.close()
for i in data:
    print(i)
```

8. Following is the structure of each record in a data file named "PRODUCT.DAT".

```
{"prod_code": value, "prod_desc": value, "stock": value}
```

The values for prod_code and prod_desc are strings, and the value for stock is an integer.

Write a function in Python to update the file with a new value of stock. The stock and the product_code, whose stock is to be updated, are to be inputted during the execution of the function.

9. Given a binary file "STUDENT.DAT", containing records of the following type:

```
[S_Admno, S_Name, Percentage]
```

Where these three values are:

S_Admno – Admission Number of student (string)

S_Name – Name of student (string)

Percentage – Marks percentage of student (float)

Write a function in Python that would read contents of the file "STUDENT.DAT" and display the details of those students whose percentage is above 75.

- 10.** Write a statement to open a binary file C:\Myfiles\Text1.txt in read and write mode by specifying for file path in two different formats.
- 11.** What are the advantages of saving data in : (i) binary form (ii) text form ?
- 12.** When do you think text files should be preferred over binary files?
- 13.** Write a statement in Python to perform the following operations:

[CBSE D 2016]

- (a) To open a text file "BOOK.TXT" in read mode
- (b) To open a text file "BOOK.TXT" in write mode

- 14.** What is following code doing?

```
File = open ("contacts.csv", "a")
Name = input("Please enter name: ")
Phno = input("Please enter phone number: ")
File.write(name + ", " + phno + "\n")
```

- 15.** Write code to open the file in the previous question and print it in the following form:

Name : <name> Phone: <phone number>

- 16.** Consider the file "contacts.csv" created in above Q and figure out what the following code is trying to do?

```
name = input ("Enter name:")
file = open ("contacts.csv", "r")
for line in file:
    if name in line:
        print (line)
```

Create a CSV file "Groceries" to store information of different items existing in a shop. The information is to be stored w.r.t. each item code, name, price, qty. Write a program to accept the data from user and store it permanently in CSV file.

8

Computer Networks

8.1 INTRODUCTION

The greatest breakthrough in technology and communication over the past 20 years has been the development and advancement of the computer network. From emailing a friend, to online bill payment, to downloading data from the internet, to e-commerce, networking has made our world much smaller and forever changed the way we communicate.

Network provides salient features which have made our life easy and comfortable, be it sending an email, withdrawing money from an ATM machine, online railway or airline reservation, or sharing audio and video files. Apart from these, the most extensively-used feature is the Print command sent from a computer to get a printout from a printer attached to some other computer. All this involves a network.

It is the network that connects various computers to each other and handles a large volume of data.



Fig. 8.1: A Computer Network

8.2 COMPUTER NETWORK—A BRIEF OVERVIEW

Several devices connected to each other for reliable communication/transfer of data constitute a network. A network can consist of a computer, a fax machine, a printer, a camera, a cell phone, etc. A collection of interconnected computers is called a **Computer Network**. Two computers or devices are said to be interconnected if they are capable of sharing and exchanging information with each other by following a protocol (set of rules).

CTM: A computer network is a collection of interconnected computers and other devices to share data and other resources (hardware and software resources).

8.2.1 Advantages of Computer Networks

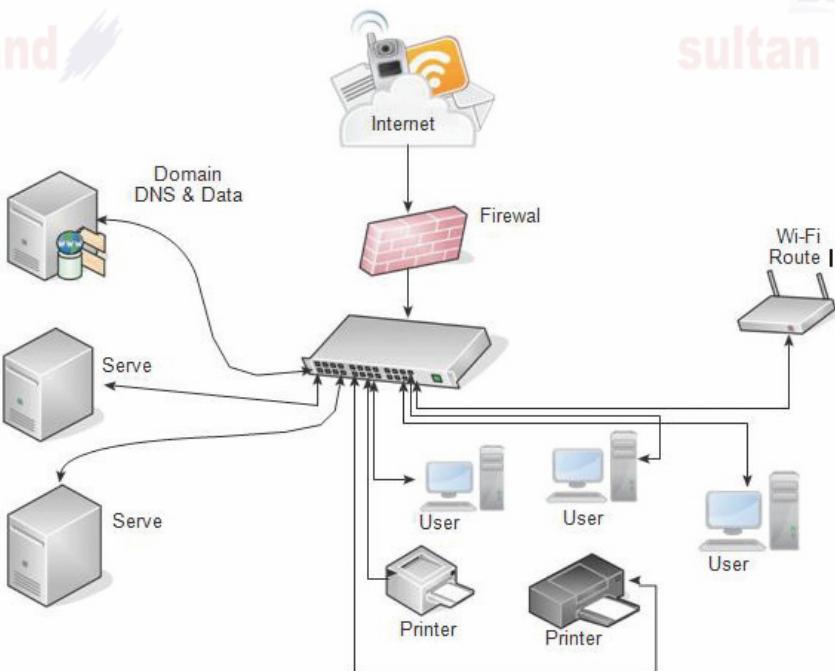


Fig. 8.2: The Network Diagram

Let us now discuss why networks are essential. Are there any advantages of networked computers over stand-alone machines? Yes, networked systems are far better. A network uses a distributed processing system in which a task is divided among several devices which are interconnected with each other. Therefore, instead of a single computer being responsible for completing the entire task, all the interconnected computers are responsible for completing the task assigned to them. This leads to better performance with high processing speed.

Networks have several advantages which are described below:

- Resource Sharing:** The primary use of a network is to share among users programs/applications, data and peripheral devices connected to the network, irrespective of their physical location. You must have noticed in your networked computer labs that when a print command is given on one computer, the document is printed by a printer attached to some other system. This allows printing of documents by several users and, hence, the printer is shared by multiple users on the network. Other resources like hard disk, DVD drive, scanner, etc., can also be shared on a computer network. *For example*, sharing database, audio and video files, antivirus software, application software, printers and scanners, etc.
- Improved Communication:** A computer network enables fast, reliable and secure communication between users. It saves time and offers easy communication methods. *For example*, in an organization, managers work at different locations to make financial reports. While working on a network, any change made by one manager on his/her computer can easily be seen by other managers and employees. Thus, a network allows managers to easily update information. This increases their efficiency and allows them to complete their work quickly.
- Reduced Communication Cost:** Sharing resources also reduces communication cost. Using public networks, we can send a large quantity of data at a low cost. Internet and mobile networks are playing a very important role in sending and receiving text, image, audio and video data at a low cost.
- Reliability of Data:** Reliability means backing up of data, i.e., data can be copied and stored on multiple computers. In a network system, all computers are connected to each



other. Thus, the information or message which is shared by each device is stored on their respective workstations (computers). If, due to some reason (hardware crash, etc.), the data gets corrupted and, thus, becomes unavailable on one computer, a copy of the same data can be accessed from another workstation for future use. This leads to smooth functioning and further processing without disruption.

- (e) **Central Storage of Data:** Files can be stored on a central node (the file server) that can be shared and made available to each and every user in an organization. With centralized processing, data is stored and retrieved from a single central location. Thus, there is no duplication of data and almost no data redundancy.

8.3 EVOLUTION OF NETWORK

The network did not evolve in a single day; rather, it took decades to become more powerful, efficient and reliable. The network has passed through several stages which are described below:

- **ARPANET (Advanced Research Project Agency Network):** ARPANET, which was jointly designed and named by the Advanced Research Projects Agency (ARPA) and US Department of Defence (DoD), was the first network and came into existence in 1969. It was a project that connected a handful of computers at different universities and US DoD for sharing of data and messages and playing long-distance games, and socializing with people to share their views.
- **NSFNET (National Science Federation Network):** In the mid-80's, another federal agency, NSFNET (National Science Federation Network), created a new network which was more capable than ARPANET. Its main aim was to use network only for academic research and not for any private business activity. Later, many private companies combined their own private networks with ARPANET and NSFNET to make a more capable and broad network—the Internet. It is the internet that links two or more networks to make a large network for sharing of information and messages.

ARPANET + NSFNET + PRIVATE NETWORKS = INTERNET

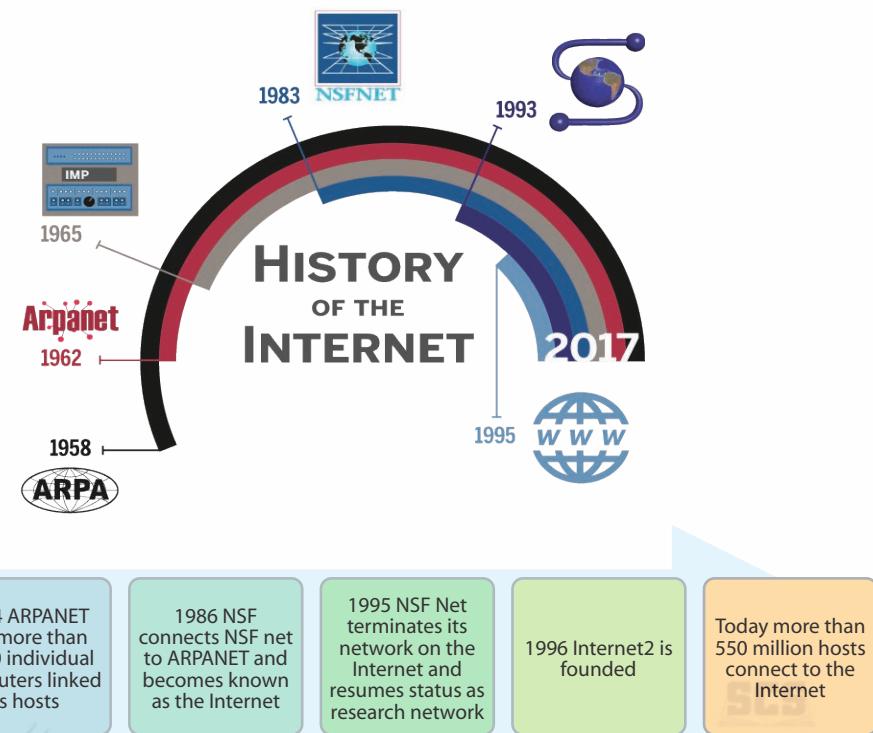


Fig. 8.3: Evolution of Internet

- **Internet:** In the 1990's, internet, which is a network of networks, came into existence. The internet has evolved from ARPANET. The computers are connected through World Wide Web that comprises a large network and shares a common communication protocol (Transmission Control Protocol-Internet Protocol, TCP/IP). It allows computers of different types to exchange information and is known as internet. Millions of domestic, business and government networks are connected with each other for the purpose of sharing files, data, email, etc. Most of the computers are not connected directly to the internet. Instead, they are connected to smaller networks which are further connected to a backbone network through gateways.

CTM: Network of networks makes the internet.

- **Interspace:** Interspace is a software that allows multiple users in a client-server environment to communicate with each other by sending and receiving data of various types such as data files, video, audio and textual data in a 3-D environment. It facilitates online real-time exchange of data. Interspace is the most advanced term of communication available on the internet today.

8.4 HOW DOES INTERNET WORK

One of the greatest things about the internet is that nobody really owns it. It is a global collection of networks, both big and small. These networks connect together in many different ways to form the single entity that we know as internet. In fact, the very name comes from this idea of interconnected networks.

Since its beginning in 1969, the internet has grown from four host computer systems to tens of millions. However, just because nobody owns the internet does not mean that it is not monitored and maintained in different ways. The Internet Society, a non-profit group established in 1992, oversees the formation of the policies and protocols that define how we use and interact with the internet.



Fig. 8.4(a): Working of the Internet

Before we learn about the basic underlying structure of the internet, e.g., domain name servers, network access points and backbones, we first need to understand how our computer connects to others.

Every computer that is connected to the internet is part of a network, even the one in our home. For example, we may use a modem and dial a local number to connect to an **Internet Service Provider** (ISP). At work, a computer may be part of a **Local Area Network** (LAN), but it most likely still connects to the internet using an ISP that the company has contracted with. When it connects to the ISP, it becomes part of their network. The ISP may then connect to a larger network and become part of their network. The internet is simply a network of networks.

Most large communication companies have their own dedicated backbones connecting various regions. In each region, the company has a **Point of Presence** (POP). The POP is a place for local users to access the company's network, often through a local phone number or dedicated line. The amazing thing here is that there is no overall controlling network. Instead, there are several high-level networks connecting to each other through **Network Access Points** or NAPs.

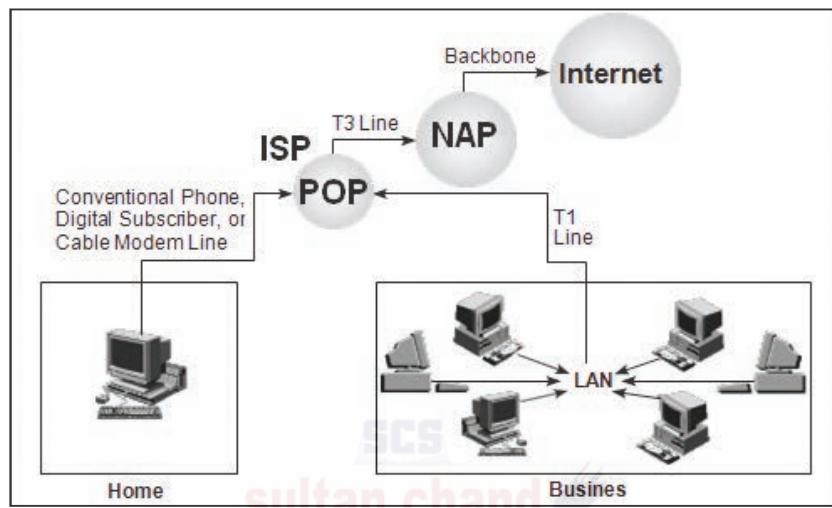


Fig. 8.4(b): POP and NAP

Gateway: Gateway is a device that connects dissimilar networks. A backbone is a central interconnecting structure that connects one or more networks just like the trunk of a tree.

At the source computer the message to be sent is broken down into small parts called packets. Each packet is given a serial number, e.g., 1, 2, 3. All these packets are sent to the destination computer. The destination computer receives the packets in random order (10 may come before 1). The packets are reassembled in the order of their number and message is restored.

How it functions smoothly: Every computer connected to the internet uses the same set of rules for communication. A set of rules is called protocol. Communication protocol used by internet is TCP/IP. The TCP (Transmission Control Protocol) part is responsible for dividing the message into packets on the source computer and reassembling them at the destination computer. The IP (Internet Protocol) is responsible for handling the address of the destination computer so that the packet is sent to its proper destination.

8.4.1 Elementary Terminology of Networks

- Nodes (Workstations):** The term node refers to computers that are attached to a network and are seeking to share resources.
- Server:** A computer that facilitates the sharing of data, software and hardware resources on the network.
- Network Interface Unit (NIU) (MAC Address):** A network interface unit is an interpreter that helps in establishing communication between the server and the client.

4. **IP Address:** Every machine on a TCP bar/IP Network has a unique identifying number called an IP Address.
 5. **Domain Name:** It is a way to identify and locate the computers connected to the internet. It must be unique.

8.5 COMPONENTS OF DATA COMMUNICATION

A network comprises several components along with their functionalities that contribute to its smooth functioning. To form a network, a lot of hardware devices are required which are described as follows:

- **Sender:** A device or a computer that sends the data.
 - **Receiver:** A device or a computer that receives the data.
 - **Message:** Message is the information to be communicated. It may be text, image, audio or video.
 - **Transmission Medium:** A transmission medium is a physical path through which the data flows from sender to receiver. A cable or wire or radio waves can be the medium.
 - **Protocol:** A set of rules that governs data transmission. It represents the communication methods which are to be followed by the sending and receiving devices.

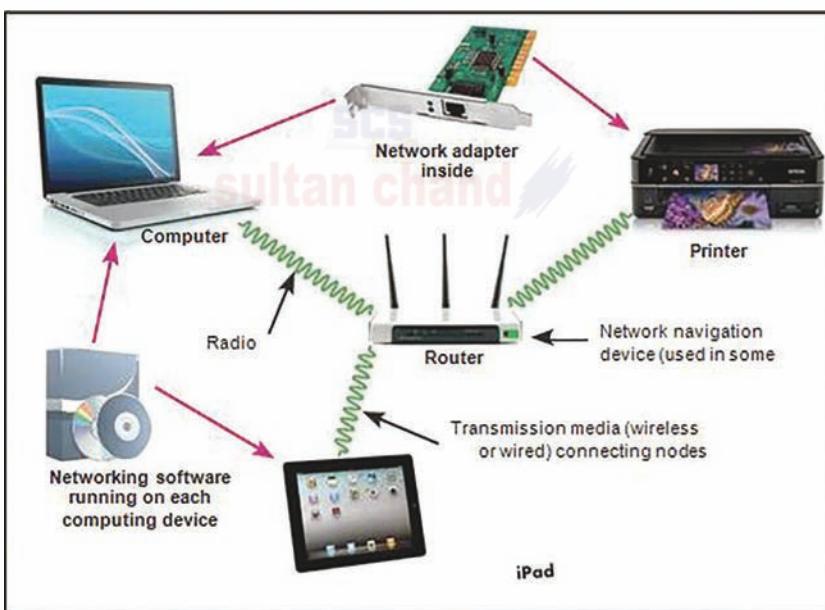


Fig. 8.5: Components of a Network

8.6 MODES OF DATA TRANSMISSION

- (a) Analog or Broadband Transmission

 - The signal is a radio frequency signal or analog, *i.e.*, it can consist of continuous electrical waves that are of varying amplitudes.
 - Telephone networks use this type of transmission.
 - Requires modem for transmitting data over baseband medium.

- (b) Digital or Baseband Transmission
 - No special device for conversion of signal to be transmitted over baseband medium.
 - The signal is a group of discrete electrical units which is transmitted in rapid succession.
- (c) Parallel Communication
 - When data is transmitted through multiple wires with each wire carrying each bit, it is called parallel communication.
- (d) Serial Communication
 - When bits are sent one after another in a series along a wire, it is called serial communication.

10001-----10001

- (e) Synchronous or Asynchronous Transmission
 - When sender and receiver synchronize their checks before transmission, *i.e.*, the sender first sends control characters to the receiver and then sends the actual data, it is called synchronous transmission.
 - Advantage**—Faster than asynchronous mode.
 - Disadvantage**—Costly and complex set-up required.
 - In asynchronous transmission, data is preceded and succeeded by a start bit and stop bit respectively. No synchronization is required.
 - Advantage**—Hardware required is simple and cheap.
 - Disadvantage**—Slower than synchronous mode.

8.7 DIFFERENT WAYS OF SENDING DATA ACROSS NETWORK

There are several ways of sending data from one node to another through network. It can be in the form of calls, messages, *etc.* By using various types of switching techniques, we can establish the connection/communication.

8.7.1 Network Switching

A network is made up of several interconnected nodes. There can be a point-to-point connection or star topology between pairs of devices, but both are not relevant for a large network. Hence, various switching techniques are used to transfer packets of data from one port of a node to another. A switched network is made up of a series of interconnected nodes called switches.

8.7.2 Switching Techniques

The main goal of networking is the reliable exchange of data or information among several interconnected nodes. For the delivery of data with accuracy, various types of switching techniques are used, namely:

1. Circuit Switching
2. Packet Switching
3. Message Switching

CTM: The technique of sending data across the network is known as Switching technique. The three types of techniques used are circuit switching, packet switching and message switching.

Circuit Switching

This provides end-to-end connection between two computers. Circuit switching is established usually in a telephone network where one person is making a call and another is receiving a call. In telephone system, the communication must be established between the two participants, i.e., the sender and the receiver. The circuit is established between these two participants before the transfer of data takes place.

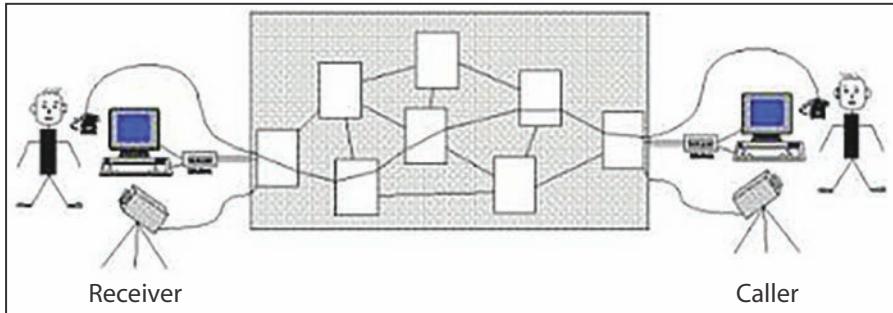


Fig. 8.6: Circuit Switching

In this technique, the entire link remains dedicated and no other user can use it even if the path remains idle. The following actions take place during circuit switching:

1. A request signal is sent by the sender to set up the connection with the receiver. It establishes a physical connection between the two participants.
2. All intermediate nodes are identified. These nodes are also called switching nodes.
3. If the destination node is available, it sends back the acknowledgement of receiving a signal. Hence, data transmission begins.
4. When the data transmission is complete, the call can be terminated.

CTM: Circuit switching is a connection-oriented service. In this technique, there is a dedicated link between the sender and the receiver and no other call can be made during this link, even if the link remains idle.

Packet Switching

In packet switching technique, the entire data is divided into small fragments called packets. Each packet is of a fixed size, usually 128 bytes or 512 bytes. Packet switching is similar to post office operation. Each packet has a source address as well as destination address (IP address) for being transmitted, in the same way as a postman delivers a letter to a specific destination address.

As there is no direct connection established between the sender and the receiver, each packet follows different routes and, therefore, the packets are delivered in a random order at the destination address. It is the TCP protocol which then arranges all received packets in a sequential order. During the transfer of packets, each packet has to pass through several intermediate nodes, so each intermediate node checks for destination IP address. If the packet matches with the node address, it is received; otherwise it is passed on to the next node until it reaches the destination IP address.

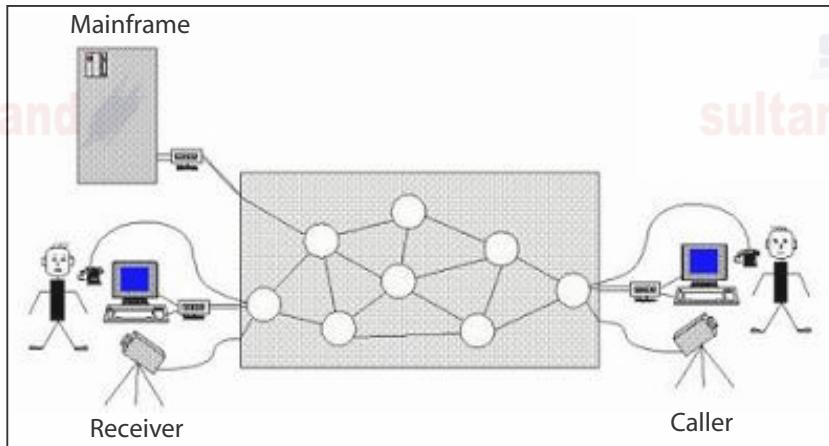


Fig. 8.7: Packet Switching

CTM: Packet switching offers a connectionless service. Data is fragmented into small packets and each packet is of fixed size in packet switching technology.

Message Switching

In message switching, the sender sends the data to a switching office first, which is then stored in its buffer. It then checks the available link and if it is free, the data is relayed to another switching office. This process goes on until the data is sent to the destination (receiver). As the data is first stored in a buffer and then sent to the next switching office, it is also called store and forward switching technique.

CTM: Message switching is a store and forward switching technique where there is no direct connection between the sender and the receiver.

8.7.3 Difference between Circuit Switching and Packet Switching

1. The circuit switching reserves the required bandwidth in advance, whereas packet switching uses bandwidth as and when required by the packets to be transmitted.
2. Circuit switching is a fast technology as compared to packet switching which is a slow mechanism of transferring packets from sender to receiver.
3. Circuit switching requires a dedicated path. Once the connection is established, the communication path is entirely dedicated to it until the data is completely transferred from sender to receiver, whereas in packet switching, packets can use any dynamic path.
4. In circuit switching, if the path is overloaded, the call is blocked and communication is delayed. But in packet switching, packets are allocated to different paths.
5. Circuit-switched networks are used for phone calls and packet-switched networks handle data.
6. Packet switching is more efficient because the cost of the link is shared by many users.
7. In circuit switching, the telephone message is sent unbroken. The message is received in the same order as it is originally sent. In packet switching, the message is broken into small packets which are randomly sent from source and received in random order at destination, which is then sequentially arranged.

8.8 DATA COMMUNICATION TERMINOLOGIES

1. **Channel:** A channel is a communication path through which the data is transmitted from the sender device to the receiver device.
2. **Baud:** The number of changes in a signal per second is known as baud. It is the measuring unit of the data transfer rate. Technically, baud refers to a number of discrete signal elements transmitted per second. 1 baud represents only 1 signal change per second and is equivalent to 1 bit per second.
3. **Bandwidth:** The amount of data that can be passed along a communication channel in a given period of time (1 second) is termed as bandwidth. The measuring unit is hertz (Hz), where $103\text{ Hz} = 1\text{ Kilo Hertz (KHz)}$, $103\text{ KHz} = 1\text{ Mega Hertz (MHz)}$.
4. **Data and Signals:** Information that is stored within computer systems and transferred over a computer network can be divided into two categories—data and signals. Data are entities that are stored in the form of 0's and 1's, which convey some special meaning to the computer system. When this data is transmitted from one place to another, it is converted into signal. Signals are the electric or electromagnetic encoding of data and are used to transmit data.
5. **Communication/Transmission Media:** It is a means of communication or access (lines of communication) set up between two organizations to exchange data/information. Communication media is the way of transmitting the signal from one place to another. Communication media is also known as transmission media. It plays an important role in sending and receiving of data to and from the sender and receiver.
6. **Data Transfer Rate:** It is the amount of data transferred in one direction over a link divided by the time taken to transfer it in bits per second (bps). The various measuring units are Both bits per second and bytes second have been abbreviated as (bps). or baud, kilobits per second (kBps), megabits per second (mBps), gigabits per second (gBps), terabits per second (tBps.)

8.9 NETWORK DEVICES

1. **Modem:** A MODEM (Modulator DEModulator) is an electronic device that enables a computer to transmit data over telephone lines. It is a device used to convert digital signals into analog signals and vice versa. There are two types of modems, namely internal modem and external modem.
2. **RJ-45 Connector:** RJ-45 is a standard type of connector for network cables. The RJ-45 (Registered Jack) connectors are the plug-in devices used in networking and telecommunications applications. They are used primarily for connecting LANs, particularly Ethernet.

CTM: RJ-45 is a short term for Registered Jack-45. It is an eight-wire connector used to connect computers on LANs, especially Ethernets.

3. **Ethernet Card:** It is a hardware device that helps in the connection of nodes within a network. Ethernet card is also known as a network card, network adapter or NIC (network interface card). It is a card that allows computers to communicate over a computer network. On Ethernet card, a physical address of each communicating computer is mentioned. Physical address is known as MAC address.

4. **Hub:** It is multi-port and unintelligent network device which simply transfers data from one port of the network to another. A hub is a hardware device used to connect several computers together with different ports. When the packet reaches one port, it is copied to all other ports of the hub without changing the destination address in the frame. Rather, it simply copies the data to all of the nodes connected to the hub.

Hubs can be either active or passive. Hubs can usually support 8, 12 or 24 RJ-45 ports.



Fig. 8.8: Hub

But the problem with hub is that it is not an intelligent device. It shares bandwidth with all the attached devices and broadcasts the data, *i.e.*, sends the data frames to all the connected nodes, as it does not remember devices/computers connected to it. Also, it cannot filter the data and causes unnecessary traffic jams.

A hub can both send as well as receive information, but only one task at a time. However, a hub is an inexpensive way to connect multiple nodes/devices to network.

CTM: Hub is a device used to connect several computers with each other.

5. **Switch:** A switch (switching hub) is a network device which is used to interconnect computers or devices on a network. It filters and forwards data packets across a network. It is also a multi-port device but with some intelligence and so the data packets received from one port of network are refreshed and delivered to the other port of the network. The main difference between hub and switch is that hub replicates what it receives on one port onto all the other ports, while switch keeps a record of the MAC addresses of the devices attached to it.

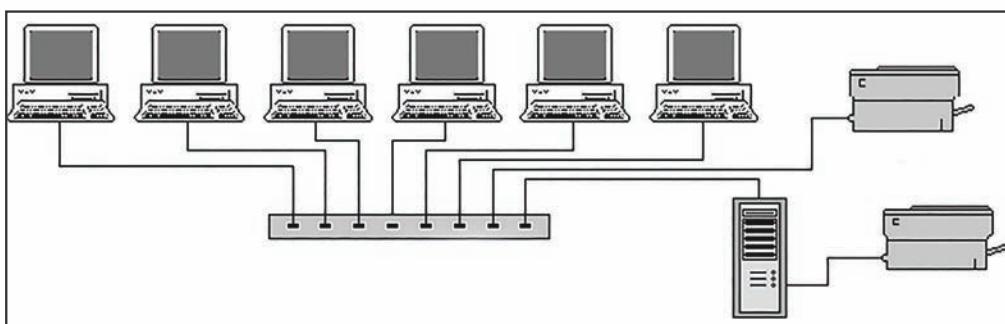


Fig. 8.9: Switch

CTM: A switch is a device that transmits data to the computers in a LAN.

6. **Bridge:** A bridge is a device that works on the physical layer as well as on data link layer. A network bridge connects multiple network segments at the data link layer (layer 2) of the OSI model. Bridges relay frames between two originally separate segments. When a frame enters a bridge, the bridge not only regenerates the signal but also checks the physical address of the destination and forwards the new copy only to that port.

An important advantage of using a bridge is that it is a smarter hub as it can filter network traffic on the basis of the MAC addresses.

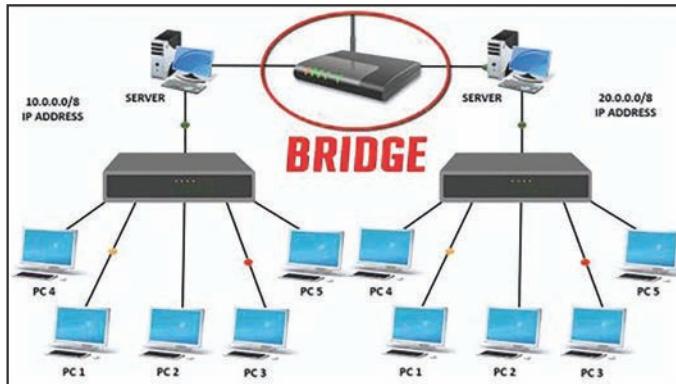


Fig. 8.10: Ethernet Bridge

CTM: A bridge is a device that links two segments together of the original network.

7. **Gateway:** A gateway is a device that connects dissimilar networks. In internet, several networks are communicating with each other and each network has a different configuration. In order to make reliable communication, there must be a device that helps in communicating.

Gateway is a device which establishes an intelligent connection between a local area network and external networks with completely different structures.

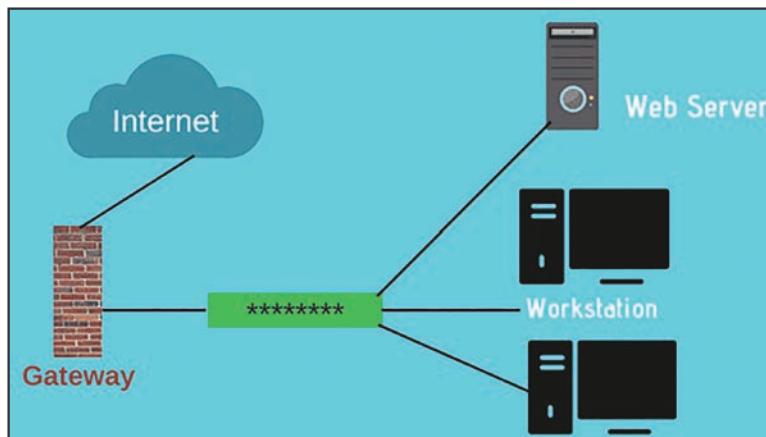


Fig. 8.11: Gateway

CTM: A gateway is a device that connects dissimilar networks.

8. **Repeater:** A repeater is a device that operates only on the physical layer of the OSI model. As a signal travels a fixed distance, before attenuation of the signal, a repeater is used which amplifies and restores signals for long-distance transmission. A repeater is an electronic device that receives a signal before it becomes too weak and regenerates the original signal. Also, it is a two-port network device that strengthens the signal intensity and connects two

identical networks. In most twisted pair Ethernet configurations, repeaters are required for cable runs longer than 100 metres. A repeater does not change the functionality of the network; instead, it makes the signal strong before it degrades.

Repeaters are also extensively used in broadcasting where they are termed as translators or boosters.

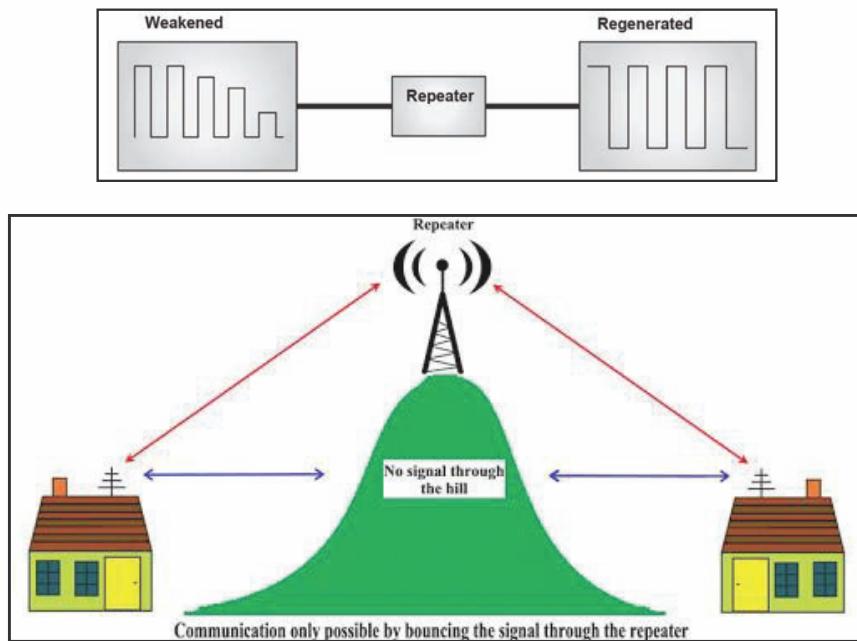


Fig. 8.12: Working of a Repeater

CTM: Repeater is a device that amplifies a signal that is transmitted across the network so that the signal is received in the same way as it is sent.

9. **Router:** A router is a networking device that forwards data packets from the source machine to the destination machine by using the shortest path. Routers are used at the network layer, which is the third layer of the OSI model.

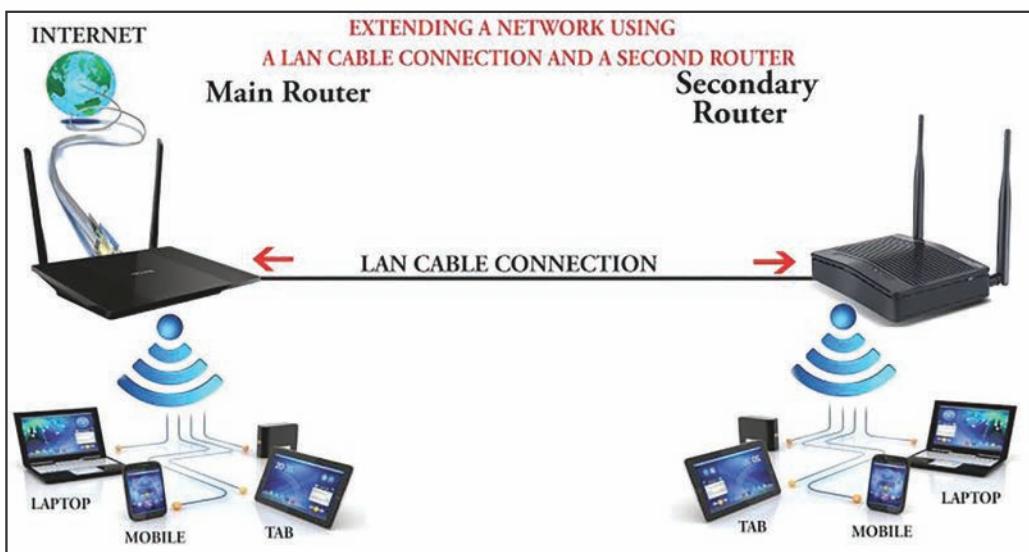


Fig. 8.13: Router

CTM: A router is a networking device that helps in forwarding packets from one machine to another.

10. **Wi-Fi Card:** A Wi-Fi card is either an internal or external Local Area Network adapter with a built-in wireless radio and antenna. A Wi-Fi card is used in a desktop computer that enables a user to establish an internet connection. Wi-Fi cards are known as wireless fidelity cards as they allow the user to set up connection without any wire. Wireless Fidelity (Wi-Fi) cards are widely used in notebook computers due to their highly portable nature. The most common Wi-Fi cards used in desktop computers are PCI-Express Wi-Fi cards made to fit the PCI- Express card slots on the motherboard.

8.10 TYPES OF NETWORKS

A computer network may be small or big depending upon the number of computers and other network devices linked together. Thus, networks vary in size, complexity and geographical spread. A computer network can be on a table, in a room, building, city, country, across continents or around the world.

On the basis of geographical spread, networks may be classified as:

1. PAN
2. LAN
3. MAN
4. WAN

8.10.1 Personal Area Network (PAN)

PANs are small networks used to establish communication between a computer and other handheld devices in the proximity of up to 10 metres using wired USB connectivity or wireless systems like Bluetooth or Infrared. PANs are used to connect computers, laptops, mobiles and other IT-enabled devices to transfer files including emails, digital photos, audio and video, etc. The Bluetooth technology implements PAN. PAN may include wireless computer keyboard and mouse, Bluetooth-enabled headphones, wireless printers and TV remotes.

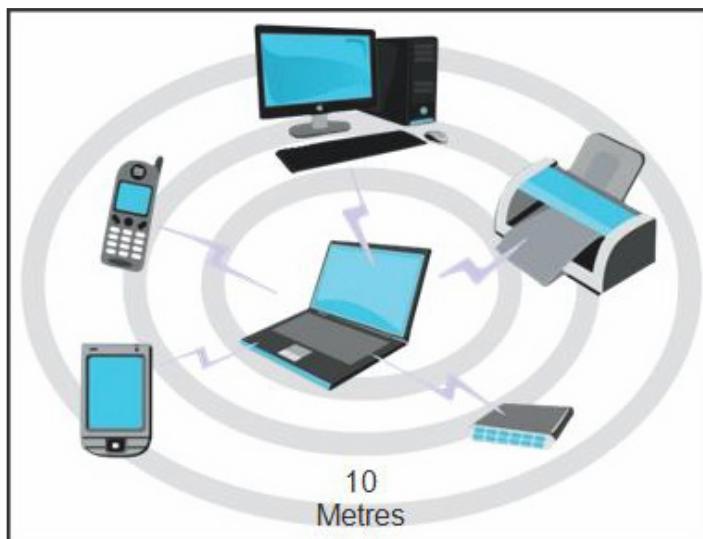


Fig. 8.14: Personal Area Network (PAN)

CTM: The network that belongs to a single person or user is known as PAN.

8.10.2 Local Area Network (LAN)

LAN is a privately owned computer network covering a small geographical area (small physical area), like a home, office or a building such as a school. It can cover an area spread over a few metres to a radius of a few kilometres.

Occasionally, a LAN can span a group of nearby buildings. In addition to operating in a limited space, a LAN is owned, controlled and managed by a single person or organization. A LAN can be set up using wired media (UTP cables, coaxial cables, etc.) or wireless media (Infrared, Radio waves). If a LAN is set up using unguided media, it is known as WLAN (wireless LAN). The key purpose of a LAN is to share resources. LAN users can share data, programs, printer, disk, modem, etc.

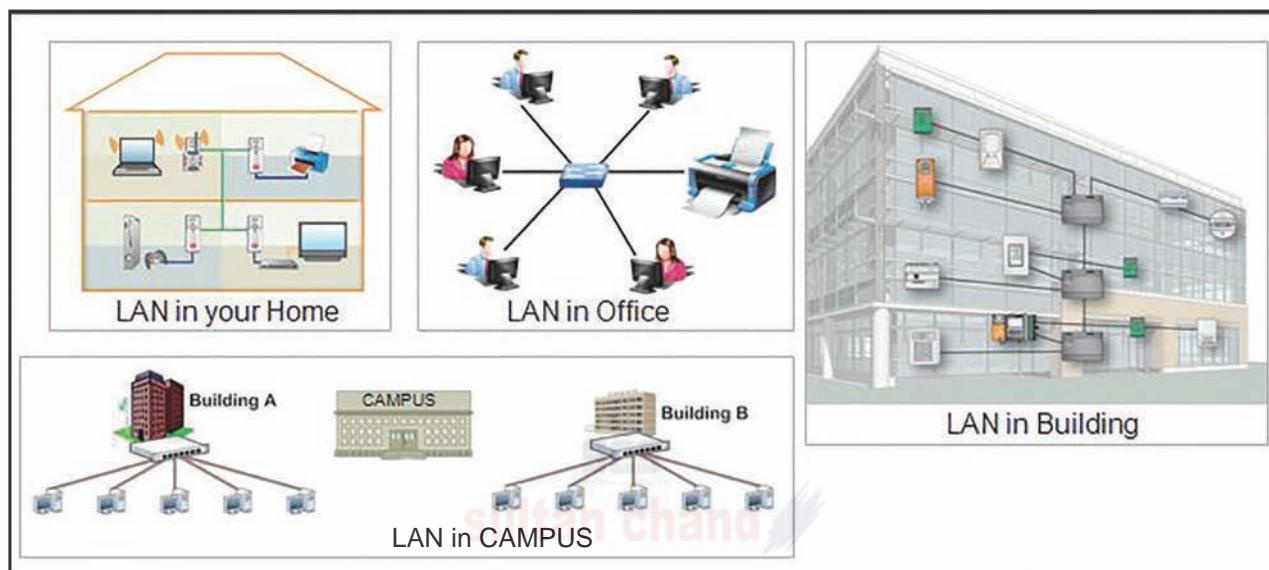


Fig. 8.15: Local Area Network (LAN)

Data transfer rate speed over a Local Area Network can vary from 10 mbps to 1 gbps.

8.10.3 Metropolitan Area Network (MAN)

MAN is larger than a LAN and can cover a city and its surrounding areas. A MAN usually interconnects a number of LANs and individual computers. It also shares the computing resources among users. All types of communication media (guided and unguided) are used to set up a MAN. A MAN is typically owned and operated by a single entity such as a government body or a large corporation. A good example of MAN is the interconnected offices of a Multinational Corporation (MNC) or cable television networks available in the whole city.

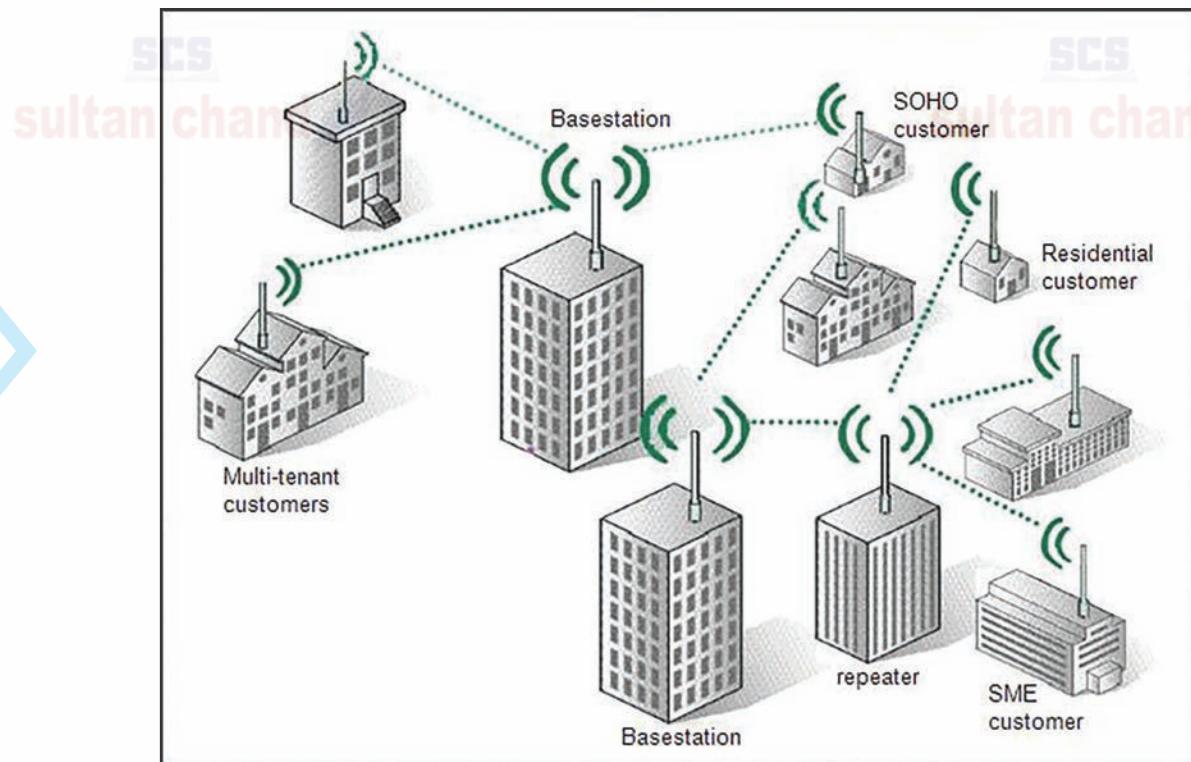


Fig. 8.16: Metropolitan Area Network (MAN)

8.10.4 Wide Area Network (WAN)

WAN is a telecommunication network. This type of network spreads over a large geographical area across countries and continents. WANs are generally used to interconnect several other types of networks such as LANs, MANs, etc. They facilitate fast and efficient exchange of information at a high speed and low cost. A WAN uses common carriers like satellite systems, telephone lines, etc.

WAN can cover an area with a radius spanning hundreds of kilometres. A network of ATMs, banks, government offices, international organizations' offices, etc., spread over a country, continent or covering many continents are examples of WAN.

All types of communication media (guided and unguided) are used to set up a WAN. The best known example of a WAN is the internet. The internet is the largest WAN spanning the entire planet.

CTM: A WAN interconnects all the computers across the world.

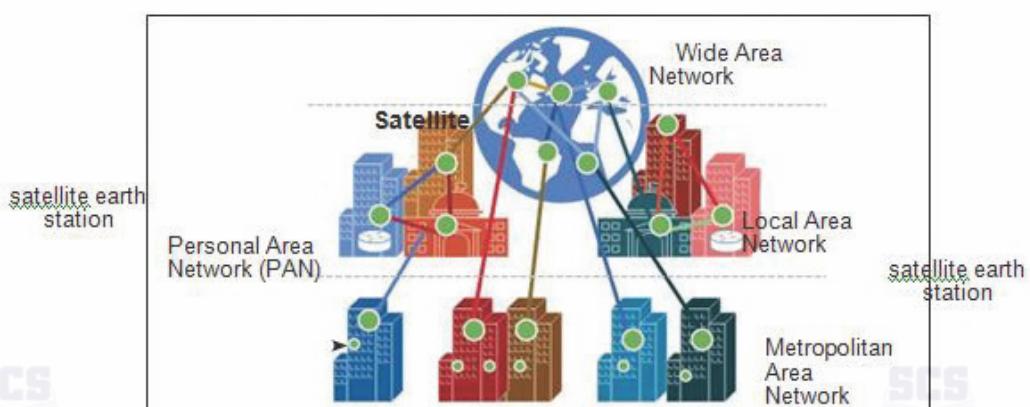


Fig. 8.17: Wide Area Network (WAN)

The following table summarizes the characteristics of PANs, LANs, MANs and WANs.

Parameter	PAN	LAN	MAN	WAN
Area covered	Small area (up to 10 m radius)	A building or campus (up to 10 km)	A city (up to 100 km radius)	Entire country, continent or globe
Networking cost	Negligible	Inexpensive	Expensive	Very expensive
Transmission speed	High speed	High speed	Moderate speed	Low speed
Error rate	Lowest	Lowest	Moderate	Highest
Network devices used	WLAN, USB Dongle	LAN/WLAN, Hub/ Switch, Repeater, Modem	Router, Gateway	Router, Gateway
Technology/media used	Infrared, Bluetooth	Ethernet, Wi-Fi	Optical fibre, Radio-wave, Microwave	Microwave Satellite

CTM: LAN and WAN are the two primary and best-known categories of area networks; the others have emerged with technological advances.

Comparing LAN and WAN

As we have seen, computer networks can be classified into LAN, MAN and WAN categories on the basis of their geographical domains. A WAN extends over a large geographical area, such as states or countries. A LAN is confined to relatively smaller areas, such as an office, a building, etc. A MAN usually covers an entire city. It uses the LAN technology. The most common example of MAN is the cable television network. Thus, the basic points of difference between LAN and WAN are as follows:

1. The distance between the nodes in a LAN connection is limited to a specific range. The upper limit is approximately 10 kilometres and the lower limit is one metre. On the other hand, WANs are spread across thousands of kilometres in different countries or regions.
2. LANs operate between speeds of 1 mega bit per second (mbps) and 10 mbps while WANs operate at speeds of less than 1 mbps. To achieve speeds of several hundred mbps, it is advisable to use the optical fibre technology.
3. The error rate in LANs is lower than that in WANs because of the short distances involved in LANs. The error rate in LANs is approximately one thousand times less than that in WANs.
4. As LANs are limited by distance, an entire LAN is usually under the control of a single entity, such as an organization. On the other hand, WANs are usually operated and controlled by multiple organizations.

Thus, we can say that in comparison to WANs, LANs cover a limited area but they operate with high speed and low error rates.

8.11 NETWORK TOPOLOGIES

Topology is a way of connecting devices with each other either physically or logically. Two or more devices make a link and two or more links form a topology. It is basically a geometrical representation of how a network is laid out.

CTM: Topology is a way of connecting several devices with each other on a network.

Types of Topologies

Basically, there are five types of topologies and each topology has some advantages and disadvantages.

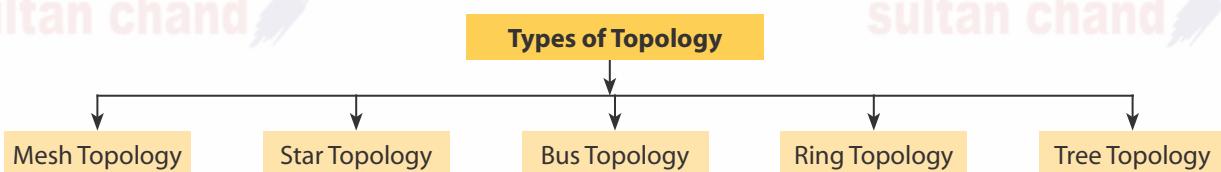


Fig. 8.18: Classification of Network Topologies

8.11.1 Mesh Topology

In mesh topology, each computer is connected with the other computer. There is a point-to-point link between each dedicated node (workstation). In this type of topology, the link carries traffic only between the two connected devices. A fully connected mesh network has $n(n-1)/2$ links, where n is the total number of connecting nodes.

For example, if there are five computers and each is connected with the other one, then there will be $5(5-1)/2=10$ links.

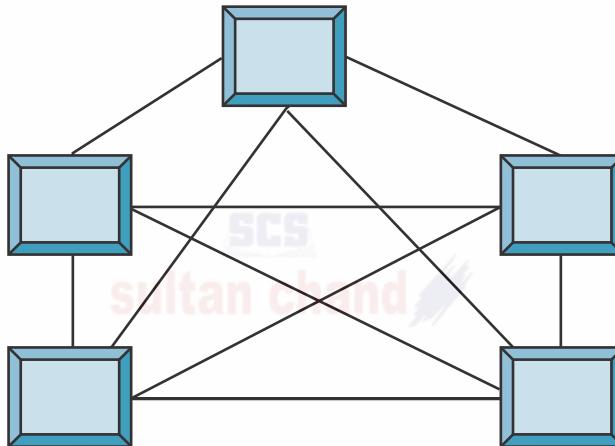


Fig. 8.19: Mesh Topology

Advantages of Mesh Topology

- (a) Each communicating device carries its own data through its own dedicated link, hence eliminating traffic problems.
- (b) A mesh topology is robust. If one link becomes unusable, it does not affect the entire system.
- (c) Expansion and modification in topology can be done without disrupting other nodes.
- (d) There is the advantage of privacy or security of data. When every message travels along a dedicated line, only the intended recipient sees it. Physical boundaries prevent other users from gaining access to messages.
- (e) Point-to-point links make fault identification and fault isolation easy. Traffic can be routed to avoid links with suspected problems. This facility enables the network manager to discover the precise location of the fault and aids in finding its cause and solution.

Disadvantages of Mesh Topology

- (a) Mesh topology is highly expensive to set up and involves high maintenance costs because of the amount of cabling and the number of I/O ports required.

- (b) The hardware required to connect each link (I/O ports and cable) can be prohibitively expensive.
- (c) Set-up and maintenance of this topology is very difficult. Even administration of the network is tough.

CTM: When there is a point-to-point link between each computer in a network, it forms mesh topology.

8.11.2 Star Topology

In star topology, each communicating device is connected to a central controller called **hub**. (*A hub is a hardware device used to connect several computers together.*)

Unlike mesh topology, the devices in star topology send and receive data indirectly; the data passes to and from the hub. If the data is required to be sent from one computer and received by another computer, then this can be accomplished only by the central controller as each data is first sent to the hub, which then relays that data to the destination computer.

Most popular and widely used LAN technology, **Ethernet**, currently operates in Star Topology.

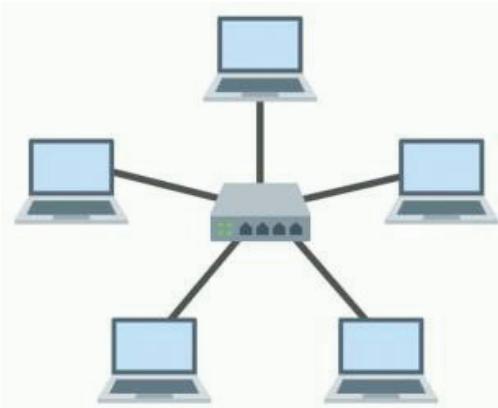


Fig. 8.20: Star Topology

Advantages of Star Topology

- (a) Fewer wires are required as compared to mesh topology which, thus, reduces the number of input/output ports.
- (b) Installation and maintenance of network is easy and takes less time.
- (c) It is easy to detect faults in this network as all computers are dependent on the central hub. This means that any problem which makes the network non-functioning can be traced to the central hub.
- (d) The rate of data transfer is fast as all the data packets or messages are transferred through central hub.
- (e) As the nodes are not connected to each other, any problem in one node does not hamper the performance of other nodes in the network.
- (f) Removal or addition of any node in star topology can take place easily without affecting the entire performance of the network.

Disadvantages of Star Topology

- (a) Extra hardware is required for installation of central controller known as hub.
- (b) All nodes of star topology are dependent on central hub and, therefore, any problem in the hub makes the entire network shut down.
- (c) The performance of the entire network is directly dependent on the performance of the hub. If the server is slow, it will cause the entire network to slow down.
- (d) More cabling is required in star topology as compared to any other topology (ring, bus, tree) as all nodes are directly connected to a central hub.

8.11.3 Bus Topology

Bus topology is a multipoint configuration, i.e., several devices are connected to a main long cable which acts as a backbone. Nodes are connected by drop lines and taps. A drop line is a connection between the long cable and devices and taps are the connectors that are punctured inside the main cable. The data flows from one end of the cable to the other.

However, as the signal travels a long distance, it becomes weaker and weaker.

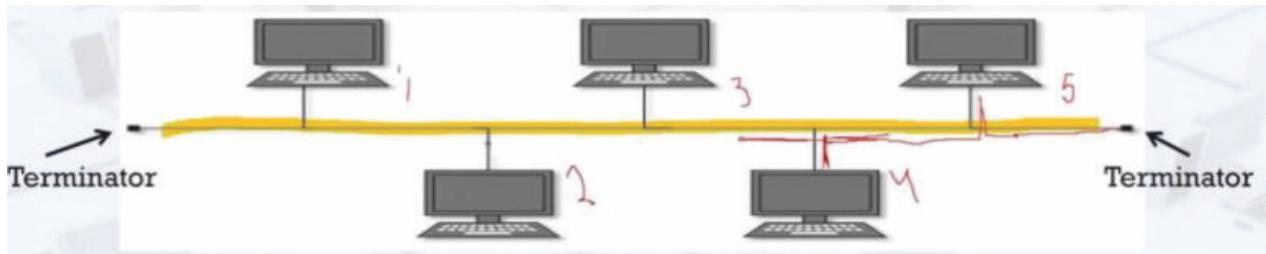


Fig. 8.21: Bus Topology

Therefore, there should be a limited number of nodes connected to a line. Ethernet is a common example of bus topology.

Advantages of Bus Topology

- (a) Nodes can be connected or removed easily from bus network.
- (b) It requires less cable length than a star topology.
- (c) Bus network is easy to implement and can be extended up to a certain limit.
- (d) It works well for small networks.

Disadvantages of Bus Topology

- (a) If there is a fault or break in the main cable, the entire network shuts down.
- (b) Terminators are required at both ends of the backbone cable.
- (c) Fault isolation is difficult to detect if the entire network shuts down.
- (d) When the network is required in more than one building, bus network cannot be used.
- (e) The signal becomes weaker if number of nodes becomes large.
- (f) Performance degradation occurs with the increased number of nodes.
- (g) Collision of data can take place because several nodes can transmit data to each other at one time.

CTM: There is a main cable which is connected to several workstations through taps. Collision of data can take place in bus topology.

8.11.4. Ring Topology

In ring topology, each node is connected to two other nodes on either side of it, forming a ring network. It shows the line configuration in which each node is connected to one predecessor node and one successor node. Signal is transmitted only in one direction along the entire ring in a circular fashion. In ring topology, each device is incorporated with a repeater to strengthen the signal as a signal passes through all nodes in the entire network. When the data is transmitted from one node

to its recipient node, then the intermediate node regenerates the signal and passes the signal to the destined node.

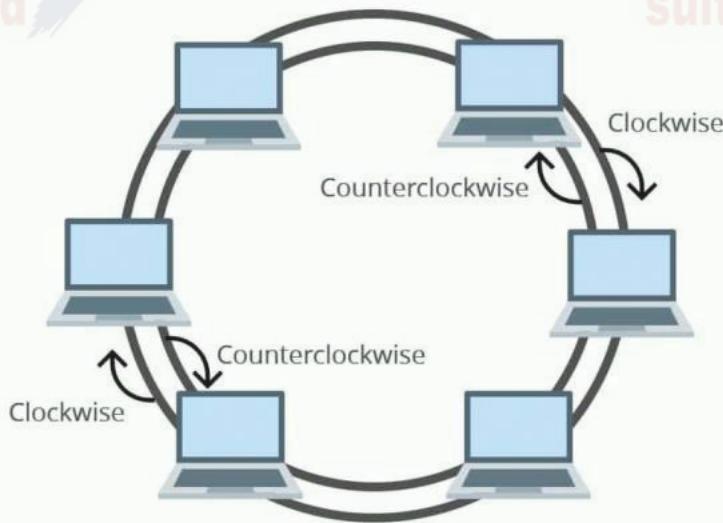


Fig. 8.22: Ring Topology

Token Ring is an example of ring topology.

Advantages of Ring Topology

- (a) A central server is not required in ring topology as the data is passed between two nodes which then pass through the entire network.
- (b) The data is transmitted in one direction only and, hence, the transmission rate increases.
- (c) The adding or removing of network nodes is easy as the process requires changing only two connections.
- (d) The configuration makes it easy to identify faults in network nodes.
- (e) In this topology, each node transmits the data to its next node in a ring.
- (f) It is relatively cheaper as compared to star topology.

Disadvantages of Ring Topology

- (a) If there is a fault in a single node, it can cause the entire network to fail.
- (b) The movement or changes made to network nodes affect the entire network's performance.
- (c) Transmission speed becomes slower with an increase in the number of nodes.
- (d) If there is a fault or break in a cable to which all other nodes are connected, the entire network shuts down.
- (e) For proper communication between each node, it is required that each computer must be turned on.

CTM: In ring topology, each workstation is connected with the predecessor node as well as with the successor node and, thus, forms a ring. Data is transmitted only in one direction.

8.11.5 Tree Topology

In tree topology, all or some of the devices are connected to the central hub, called an active hub, and some of the devices are connected to the secondary hub, which may be an active hub or passive hub. An active hub contains the repeater that regenerates the signal when it becomes weaker with longer distances. A passive hub simply provides a connection between all other connecting nodes.

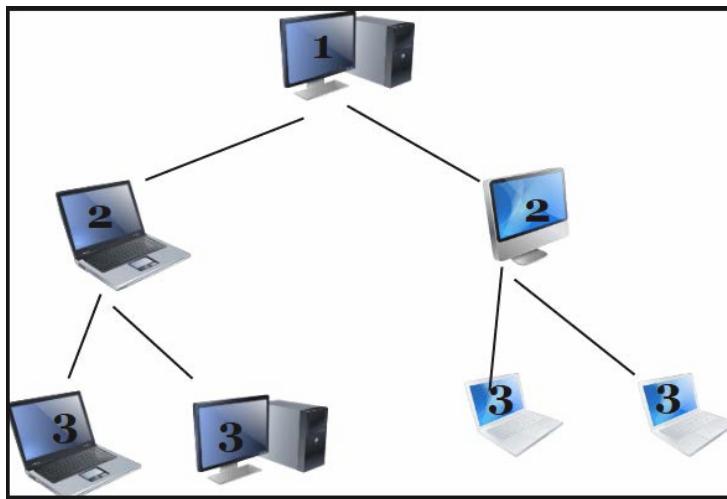


Fig. 8.23: Tree Topology

Advantages of Tree Topology

- (a) The tree topology is useful in cases where a star or bus cannot be implemented individually.
- (b) It is most suited in networking multiple departments of a university or corporation, where each unit (star segment) functions separately, and is also connected with the main node (root node).
- (c) The advantages of centralization that are achieved in a star topology are inherited by the individual star segments in a tree network.
- (d) Each star segment gets a dedicated link from the central bus. Thus, failing of one segment does not affect the rest of the network.
- (e) Fault identification is easy.
- (f) The network can be expanded by the addition of secondary nodes. Thus, scalability is achieved.

Disadvantages of Tree Topology

- (a) As multiple segments are connected to a central bus, the network depends heavily on the bus. Its failure affects the entire network.
- (b) Owing to its size and complexity, maintenance is not easy and costs are high. Also, configuration is difficult in comparison to other topologies.
- (c) Though it is scalable, the number of nodes that can be added depends on the capacity of the central bus and on the cable type.

CTM: In tree topology, the main central hub is connected to the secondary hub through which all devices are connected with each other. Tree topology is just like a parent-child relationship.

The decision to select a particular topology for a network does not take place in isolation as the topology determines the type of media and access methods that would be used on the network. Therefore, several factors are taken into consideration before selecting a topology which are as follows:

- **Cost:** Cost signifies that the network should be cost-effective. For this, it is required to minimize the installation cost. This can be done by selecting an economical transmission medium (that is, wires) for data exchange between the network nodes. The network cost can also be minimized by reducing the distances between the nodes.
- **Flexibility:** Flexibility signifies that the network must be flexible enough, *i.e.*, it should be easy to reconfigure. Reconfiguring a network means to rearrange the existing nodes or add new ones on the network.
- **Reliability:** It refers to the degree of trust that can be placed on a network. Like any other system, a network can also encounter failure. A network failure can occur due to the following causes:
 1. When an individual node stops working.
 2. When the entire network fails. This is caused by a more serious fault that stops the working of an individual node.

What is Point-to-Point (P-P) Link

A P-P link, also known as a dedicated link, is used to connect two nodes in a network. The function of P-P link is to send and receive data over a network. In a P-P network, each workstation receives data from only one transmitter and each transmitter transmits data only to one receiver. Fig. 8.23 shows a P-P link in a network.

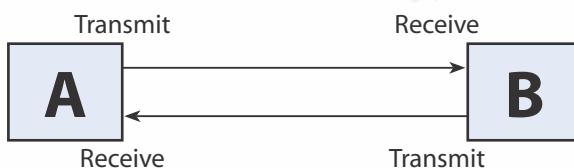


Fig. 8.24: A Point-to-Point Link

For better performance, the transmit and receive operations can occur over separate cables or wires, or they can occur in turns over the same wire by using different transmission techniques. A P-P link can be established in several ways. The simplest way is to install a P-P link between each pair of computers over a network.

8.12 COMMUNICATION MEDIA

Communication media is also known as transmission media through which data or signal is transferred between two communicating devices, *i.e.*, from one system to another system, through wires or without wires. If the data is sent across network through wires, it is called guided media and if the data is sent without wires, it is called unguided media.

CTM: Communication media is a transmission media used for sending data or signal across the network.

Types of Communication/Transmission Media

All communication/transmission media can be divided into two categories:

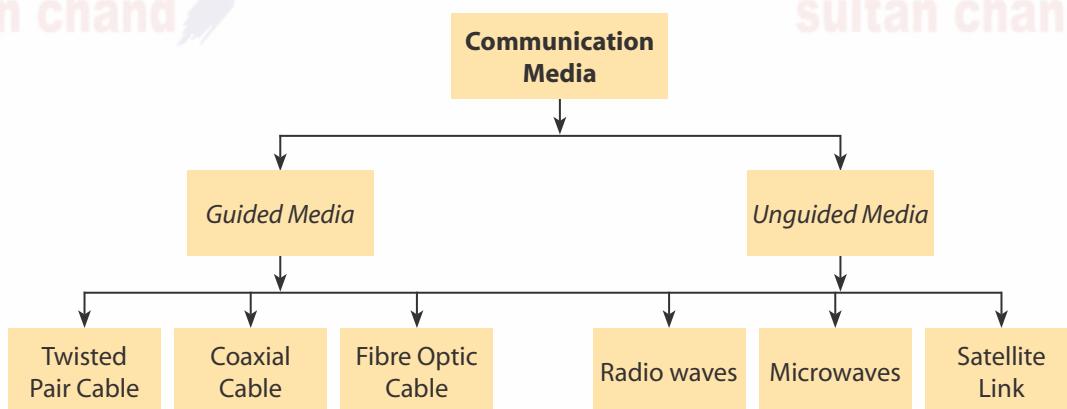


Fig. 8.25: Types of Communication Media

- A. **Guided Media (Wired Media):** Guided media is also known as physical or conducted media. These media use wires for transmitting data. Various wired connections are twisted pair wire, coaxial cable and fibre optic cable.
- B. **Unguided Media (Wireless Media):** A transmission media that does not require the use of cables for transmission of data is known as unguided media. Wireless transmission is also known as unguided media or non-physical media as the transmission takes place through various types of electromagnetic waves, such as radio waves, terrestrial microwave transmissions, satellite transmissions, cellular radio systems, infrared transmissions, etc.

8.12.1 Guided Media or Wired Networking Technologies

1. Twisted Pair Cable

A twisted pair cable is the oldest, simplest and the most common type of conducted media. It is made of two plastic insulated copper wires which are twisted together to form a single wire. Each wire is 1 mm thick. Out of these two wires, only one carries the actual signal while the other is used for ground reference. The wires so twisted are helpful in avoiding interference from the nearby similar pairs, which is known as crosstalk. Twisted pair can be specified as Category 1–5 and is abbreviated as Cat 1–5. Category 6 twisted pair can support data transmission as high as 200 mbps for 100 metres while Category 7 twisted pair can support still higher data rates.

Twisted pair comes in two varieties:

- (a) **Shielded twisted pair (STP):** STP cables are covered in metal foil. This makes them indifferent to noise and crosstalk.
- (b) **Unshielded twisted pair (UTP):** UTP has seven categories, each suitable for a specific use. In computer networks, mostly Cat-5, Cat-5E and Cat-6 cables are used. UTP cables are connected by RJ-45 connectors.

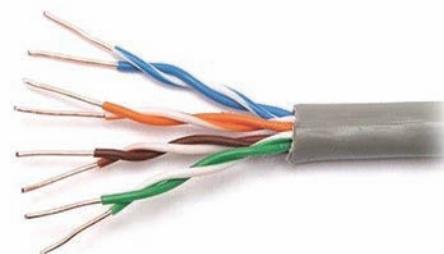


Fig. 8.26: Twisted Pair Cable

Advantages of Twisted Pair Cable

- (a) It is simple to use.
- (b) It is inexpensive and does not require skilled personnel.
- (c) It is less susceptible to electrical interference caused by nearby equipment or wires in a telephone system. Signals can travel several kilometres without amplification when twisted pair wires are used.
- (d) These media can be used for both analog and digital data transmission. The bandwidth depends on the thickness of the wire and the distance travelled, but several megabits per second can be achieved for a few kilometres in many cases.
- (e) If a portion of a twisted pair cable is damaged, the entire network is not shut.

Disadvantages of Twisted Pair Cable

- (a) STP wire is physically larger and more expensive than twisted pair wire.
- (b) STP is more difficult to connect to a terminating block.
- (c) It easily picks up noise signals which results in higher error rates when the line length exceeds 100 metres.
- (d) Being thin in size, it is likely to break easily.
- (e) It can support 19,200 bps up to 50 feet on RS-232 port.

2. Coaxial Cable

A coaxial cable is generally called a coax wire. It consists of insulated copper wires surrounded by a braided metal shield and covered in a plastic jacket. Coax cables are capable of carrying higher frequency signals than that of twisted pair cables. The wrapped structure provides it a good shield against noise and crosstalk. Coaxial cables provide high bandwidth rates of up to 450 mbps. Coaxial cable also comes in two primary physical types: **thin coaxial cable** and **thick coaxial cable**. There are three categories of Coax cables, namely RG-59 (Cable TV), RG-58 (Thin Ethernet) and RG-11 (Thick Ethernet). RG stands for Radio Guide. Coax cables are widely used for internet connections and cable televisions.

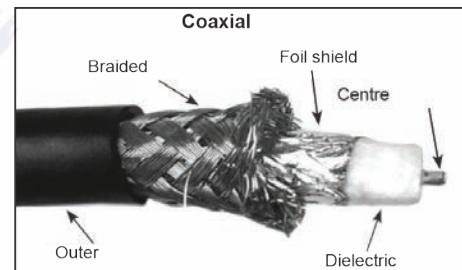


Fig. 8.27: Twisted Pair Cable

Advantages of Coaxial Cable

- (a) Coaxial cable can support greater cable lengths between network devices than twisted pair cable.
- (b) It is useful for transmitting analog as well as digital data across the network. For analog data transmission, 75 ohm broadband coaxial is used and for digital transmission, 50 ohm baseband cable is used.
- (c) It is widely used for cable television and internet connections.
- (d) Coax are used for transmitting several channels simultaneously, i.e., they are helpful in broadband transmission.
- (e) Coaxial cable has excellent noise immunity because of thick covering outside the insulated wires.

- (f) Thick coaxial cable has an extra protective plastic cover that helps keep moisture away.
- (g) It is relatively inexpensive as compared to fibre optic cable.

Disadvantages of Coaxial Cable

- (a) Thick coaxial cables do not bend easily and thus are difficult to install.
- (b) It is expensive as compared to twisted pair cable.

3. Fibre Optic Cable

Fibre optic transmits light signals rather than electrical signals. It is the newest form of guided media. Several glass fibres are bundled together and are encased in an insulated covering. Light signals travel into the fibre optic cable at one end and are received at the other end. When the light enters the fibre optic, the light pulse inside the cable hits the outer walls of the wire at a similar angle, which helps in moving the light wave forward. The outer surface of the glass wire provides just the right angle reflection to keep the light bouncing back and forth along the length of cable. The light source used in this process is light emitting diode (LED).

Optical fibres may be single mode or multimode. Single mode optic fibre transmits only single light width but multimode fibre uses multiple light paths.

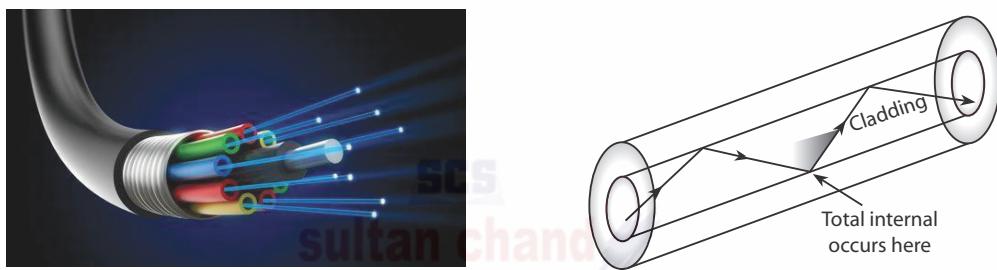


Fig. 8.28: Fibre Optic Cable

Advantages of Fibre Optic

- (a) Fibre optic typically offers better bandwidth and can carry more information at once.
- (b) As the signal travels in the form of light, there is less attenuation and signal degradation.
- (c) Optical fibre wires are made of glass, so there is little risk of fire because of absence of spark hazards.
- (d) Fibre optic cables are much thinner and lighter than metal wires.
- (e) Lighter weight makes fibre easier to install.
- (f) It does not leak signals, so it is immune to eavesdropping.
- (g) A signal can run for 50 km without requiring regeneration.
- (h) Fibre optic cables are also used in research and development.

Disadvantages of Fibre Optic

- (a) A highly skilled labour is required for its installation and maintenance.
- (b) It is relatively expensive as compared to other guided media.
- (c) As fibre optic is made of glass, they can be easily broken.
- (d) As light travels in a straight line, two fibres are needed if we need bidirectional communication.

8.12.2 Unguided Media or Wireless Networking Technologies

Features of Unguided Media:

1. Transmission and reception are achieved by means of an antenna.
2. Transmission can be either directional or omnidirectional.
 - (a) Directional
 - transmitting antenna puts out focused beam
 - transmitter and receiver must be aligned
 - (b) Omnidirectional
 - signal spreads out in all directions
 - can be received by many antennas

1. Microwave

Microwave signals are used to transmit data without the use of cable. It is a line-of-sight transmission as signal travels in a straight line. In microwave communication, two directional parabolic antennas are mounted on towers, buildings or hills to send and receive signals through air. However, they must be properly aligned with each other, otherwise the signal will not be focused well at the receiving antenna.

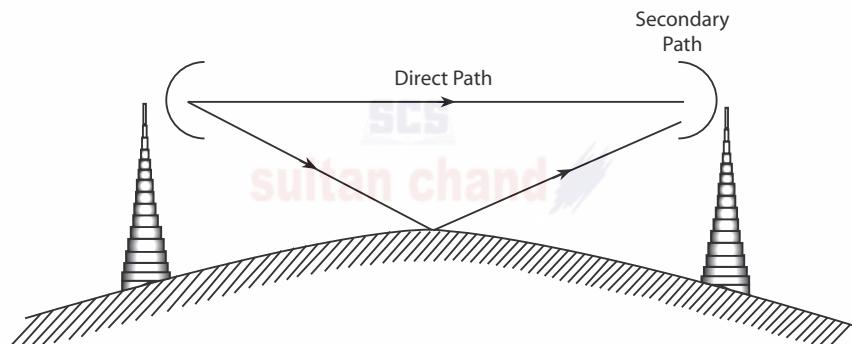


Fig. 8.29: Microwave

Advantages of Microwave Transmission

- (a) It is a cheaper source of communication as it avoids using cables and maintaining repeaters.
- (b) Communication through microwave is much easier over difficult terrain.
- (c) Microwave system permits data transmission rate of about 16 gigabits per second.

Disadvantages of Microwave Transmission

- (a) It is an insecure mode of communication.
- (b) Signals can be split and propagated in different directions in air and received by the receiver antenna with a reduced strength.
- (c) Microwave propagation is affected by weather conditions such as rain, thunderstorm, etc.
- (d) The cost of implementing towers, antennas is relatively high.

2. Radio Waves

Radio waves use radio frequencies which are allocated to private businesses for direct voice communication. A radio set-up uses transmitter and receiver. A transmitter sends radio waves and encodes them into sine waves which, when received by a receiver, are decoded and the message is received. Both the transmitter and receiver use antennas to radiate and fetch radio signals. They are not line-of-sight transmission and, hence, can penetrate buildings easily.

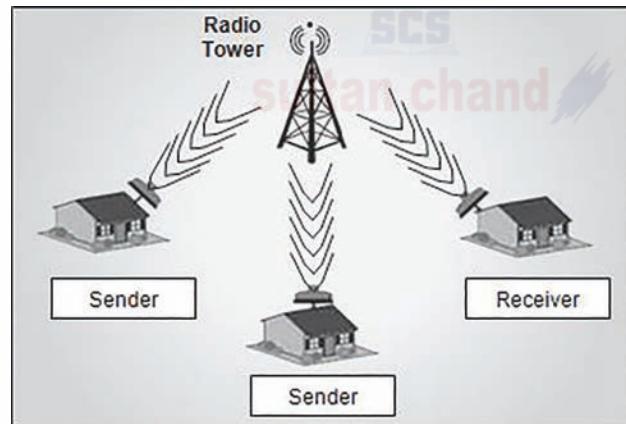


Fig. 8.30: Radio Waves

Advantages of Radio Waves

- (a) They can be used indoors or outdoors.
- (b) They are omnidirectional and can travel in any direction.
- (c) Transmitter and receiver antenna do not need to be physically aligned.
- (d) Radio wave transmission offers mobility.
- (e) It is cheaper than laying cables and fibres.
- (f) It offers ease of communication over difficult terrain.

Disadvantages of Radio Waves

- (a) Radio wave communication is an insecure mode of communication.
- (b) Radio wave propagation is susceptible to weather effects like rain, thunderstorm, etc.

3. Satellite Link

Satellite transmission is also a kind of line-of-sight transmission that is used to transmit signals throughout the world. When the frequency is greater than 3 GHz, the transmission is known as microwave. Satellite is a special type of microwave transmission medium.

It provides various types of services such as transmitting fax, voice data, video, email and other internet applications. In satellite communication, an earth station has a satellite dish, which functions as an antenna to transmit and receive data from satellites.

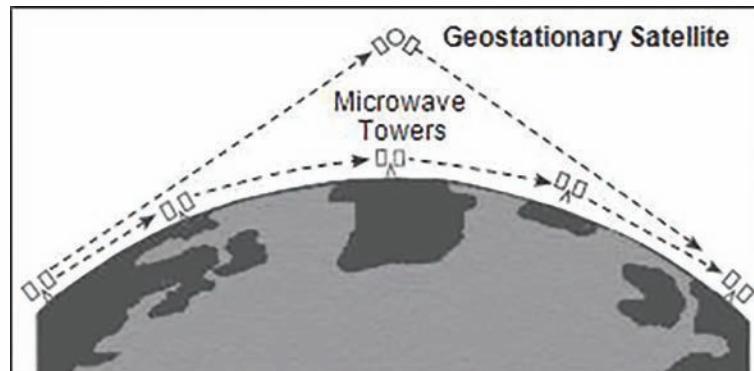


Fig. 8.31: Satellite Link

When data is transmitted from an earth station to a satellite, it is known as uplink and when transmission takes place from a satellite to an earth station, it is known as downlink. In satellite, there are transponders that send and receive signals from/to the earth station.

Advantages of Satellite Link

- (a) The area covered is quite large.
- (b) No line-of-sight restrictions such as natural mountains, tall buildings, towers, etc.
- (c) Earth station which receives the signals can be at a fixed position or can be relatively mobile.

Disadvantages of Satellite Link

- (a) It is very expensive as compared to other transmission mediums.
- (b) Installation is extremely complex.
- (c) Signals sent to the stations can be interrupted by external interference.
- (d) Low antenna gains result in overcrowding of available bandwidth.

4. Infrared

The type of transmission that uses infrared light to send data is known as infrared transmission. The data is transmitted through air and can propagate in the open space; however, it cannot penetrate the walls of the room. It is an example of short range wireless network. Infrared speed varies from 2.4 kbps to 16 mbps. A very good example of infrared transmission is a handheld remote control such as remote control of a TV or AC, etc.

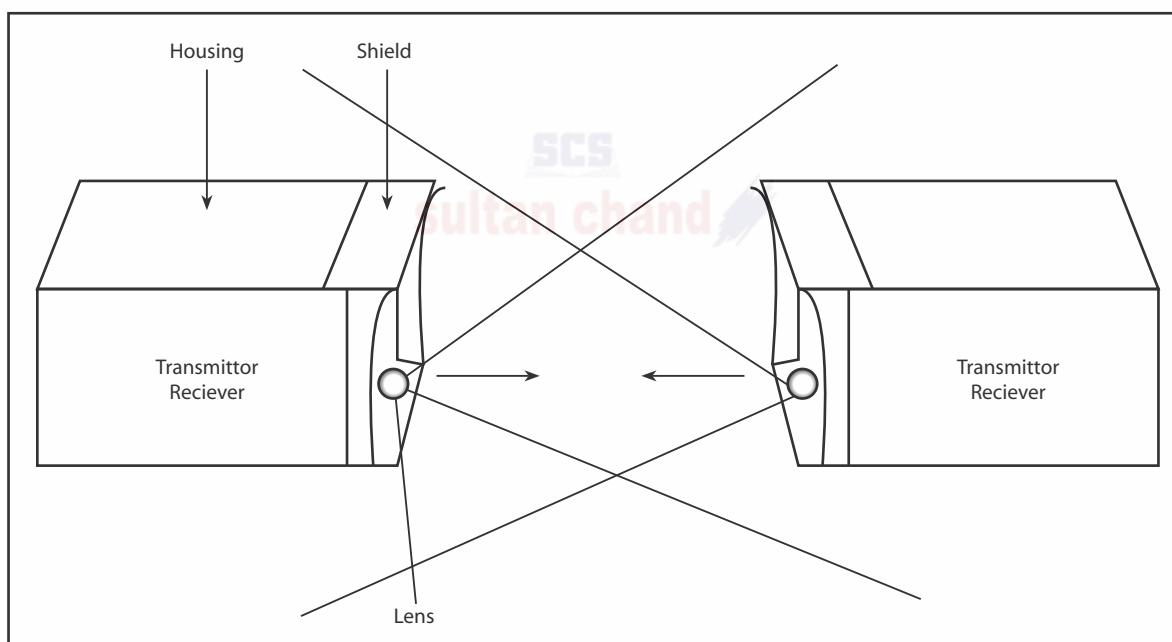


Fig. 8.32: Infrared Transmission System

Advantages of Infrared Transmission

- (a) It is a secure medium of transmitting data.
- (b) It is a cheap mode of transmission.

Disadvantages of Infrared Transmission

- (a) It can work only for short distances.
- (b) It cannot penetrate walls and is affected by distance, noise and heat.



5. Wireless Technology

Wireless technology is the process of sending information through invisible waves in the air. Information such as data, voice and video are carried through the radio frequency of the electromagnetic spectrum. Thus, wireless communication is the transfer of information between two or more points that are not connected by an electrical conductor. The various wireless technologies available are described as follows:

- (a) **Wi-Fi (Wireless Fidelity):** Wi-Fi is wireless fidelity where communication takes place without wires. The users share the data with each other by connecting to the broadband internet service without using cables. As it is not a secured media of transmitting data, the user must use privacy methods such as using passwords and make the connection security enabled so that it does not become susceptible to hackers. For transmission of data through Wi-Fi, a user must have a broadband connection, a wireless router and a laptop or desktop.
- (b) **Wi-Max (Worldwide Interoperability for Microwave Access):** Wi-Max is a wireless communication system that provides broadband internet accessibility up to 30 miles. The data transmission takes place without wires. It provides data rates up to 30 to 40 megabit-per-second and up to 1 Gbit/s for fixed stations. Wi-Max is based on standard IEEE 802.16 technology that provides users with access to high-speed voice, data and internet connectivity. Wi-Max uses broadband internet connection and requires a tower known as base station to communicate with the user. Instead of wires, it uses a microwave link to establish a connection. Internet connectivity to end-user is provided by a microwave link between the tower and the end-user.

6. Other Unguided/Wireless Media

In this section, we will discuss two other important and extensively used wireless media— laser waves and Bluetooth.

- (a) **Laser Technology:** Laser refers to a transmission medium that requires direct line of sight. It provides point-to-point transmission between points located at short distances, such as buildings. Similar to microwave, a laser is unidirectional but much faster. Lasers require transmitters and photosensitive receivers for transmission. A disadvantage associated with laser transmission is that it is badly affected by weather conditions.
- (b) **Bluetooth Technology:** Bluetooth refers to a telecommunication industry specification that defines how different devices can be connected virtually and transfer information among each other. Bluetooth technology is commonly used in various portable devices such as laptops and PDAs to establish a wireless connection in the form of Wireless LAN (WLAN).



Fig. 8.33: Bluetooth Technology

Transmission of information using Blue-tooth requires a low-cost transceiver chip in each of the devices that need to be connected. Data is transferred at the frequency of 2.45 GHz, which is available globally. The rate of data transfer in this type of transmission is 1 mbps. Each device has a 48-bit address as per the IEEE 802 standard. The connection between the devices can be either point-to-point or multipoint. Bluetooth is equipped with various features such as encryption and verification to provide security.

8.13 NETWORK TERMINOLOGIES

1. **Data Channel:** A channel is a communication medium through which data or message is transferred from the sender to the receiver.
2. **Baud:** The number of changes in a signal per second is called a baud. It is the unit of measurement for the information-carrying capacity of a communication channel.
3. **Bits per second:** It is the measuring unit of speed at which data transfer takes place.
4. **Bandwidth:** The bandwidth measures the information-carrying capacity of a line or a network. It is the difference between the highest and lowest frequencies allowed on a transmission media.

$$B = f_h - f_l$$

where f_h and f_l are the highest and lowest frequencies.

For example, If the highest frequency is 80 Hz and lowest frequency is 50 Hz, what is the bandwidth of a signal?

Sol. $B = f_h - f_l$

$$B = 80 - 50$$

$$B = 30 \text{ Hz.}$$

5. **Protocol:** A protocol is an agreement between the communicating parties on how communication is to proceed. Protocol means a set of rules that governs a network. A protocol is a formal description of message formats and the rules that two or more machines must follow to exchange those messages.

For example, using library books.

CTM: Protocol is a set of rules that governs the network.

Types of Protocols

Protocol specifies what is communicated and how. Let us take an example to explain this concept. In India, different people speak different languages. Now, a person from Tamil Nadu and a person from Punjab may not be able to communicate with each other because of the language difference. However, they can exchange their ideas and communicate with each other using English as their common language. Similarly, in case of computers, the hardware, software or even a combination of the two might be required to implement and carry out the protocol. Thus, the protocol will help in setting up a channel of communication or a connection between two computers; in other words, a hardware connection between two computers.

There are multiple protocols defined for computer networks, which are as follows:

- (a) TCP (Transmission Control Protocol)
- (b) IP (Internet Protocol)
- (c) FTP (File Transfer Protocol)
- (d) PPP (Point-to-Point Protocol)
- (e) SMTP (Simple Mail Transfer Protocol)
- (f) POP3 (Post Office Protocol)
- (g) TELNET (Remote Login)



(a) TCP/IP (Transmission Control Protocol/Internet Protocol)

TCP is one of the main protocols in TCP/IP networks. The IP protocol deals only with packets but TCP enables two hosts to establish a connection and exchange streams of data. TCP works with the Internet Protocol (IP), which defines how computers send packets of data to each other. The packets are delivered not in a sequential order; instead, they are delivered randomly. Now, TCP at the receiver side collects all packets and arranges them in a sequential order. TCP is a reliable stream delivery service that guarantees that all bytes received will be identical with bytes sent and in the correct order. Since packets are transferred over many networks and, thus, not reliable, a technique known as positive acknowledgement with retransmission is used to guarantee reliability of packet transfers.

(b) IP (Internet Protocol)

IP is the primary protocol present in the internet layer of the internet protocol suite. It is responsible for delivering packets from the source host to the destination host based on the IP addresses mentioned in the packet headers. IP forwards each packet based on a four byte destination address (the IP number). The packets are moved randomly from source to destination, which are then assembled in a sequential order at the destination computer. IP stores destination addresses in the form of IP addresses so that the packets will move to the destined address only by following the shortest route.

(c) FTP (File Transfer Protocol)

FTP is the simplest and most secure way to exchange files over the internet. The main objectives of FTP are:

- Transmitting and sharing of files (computer programs and/or data).
- Indirect or implicit use of remote computers.
- To shield a user from variations in file storage systems among different hosts.
- To transfer data reliably and efficiently.
- FTP uses the internet's TCP/IP protocols to enable data transfer.

FTP is most commonly used to download a file from a server using the internet or to upload a file to a server (e.g., uploading a web page file to a server).

While sharing files from one system to another, we may encounter several problems—two systems may have different directory structures, two systems may have different file-naming conventions, or two systems may have different ways to represent text and data. All these problems are solved by FTP.

CTM: File transfer protocol is used to transfer files from server system to requesting node, primarily for information sharing.

(d) PPP (Point-to-Point Protocol)

PPP is the most commonly used data link protocol. It is a protocol used to establish a direct connection between two nodes. It can provide connection authentication, transmission encryption and compression. It is used to connect the Home PC to the server of ISP. The communication takes place through high speed modem. PPP is used to connect



telephone dial-up lines to the internet. Internet service provider may provide you with a PPP connection so that when you send any request, the provider's server can respond to your requests, forward it to the internet server, and then forward responses back to you. For communicating through PPP, the user needs to install PPP drivers assigned by the internet to the computer.

The following steps are required to connect to the internet:

- Double click on the dial-up icon.
- Dial the number provided to you by the ISP.
- Enter the login and password.
- Open a browser like Internet Explorer or Netscape Navigator.

The various features of PPP are:

- Establishing the link between two devices.
- Maintaining this established link.
- Configuring this link.
- Terminating this link after the transfer.
- PPP provides error detection.
- PPP supports multiple protocols.
- It also defines how two devices can authenticate each other.

(e) **SMTP (Simple Mail Transfer Protocol)**

As the name suggests, SMTP is used for sending email messages to other networks or computers. It handles only outgoing messages and not incoming messages. SMTP does not create messages: rather, it helps in forwarding messages between client servers. It uses TCP to send messages to another host. For receiving messages, another protocol POP (post office protocol) is used. Thus, email clients require addresses of both SMTP server and another server that processes incoming messages (usually POP or IMAP). SMTP uses TCP port number 25 for standard communication.

(f) **POP3 (Post Office Protocol 3)**

The POP (Post Office Protocol 3) is a simple and standard method to access mailbox and download messages to the local computers. The user can receive messages with the help of POP protocol. The advantage is that once the messages are downloaded, an internet connection is no longer needed to read the mail. A user can read all emails offline as these are saved on the computer's hard disk.

Just like with the SMTP protocol, the POP protocol (POP2 and POP3) sends text commands to the POP server. There are two main versions of this protocol—POP2 and POP3—to which ports 109 and 110 respectively are allocated and which operate using radically different text commands. To get a mail from POP server, a user must enter a valid username and password for their email account. The POP3 protocol thus manages authentication using the user name and password; however, it is not secure because the passwords, like the email, circulate in plain text over the network. POP3 protocol blocks inbox during its access which means that simultaneous access to the same inbox by two users is impossible



(g) TELNET (Remote Login)

Telnet is a remote login that helps a user to log on to another user's terminal without being its original user. A user who is logging in to their own system can also get access to log on to another user system and perform various functions such as accessing files or sharing files to/from the remote system. With TELNET, a user logs in as a regular user with whatever privileges they may have been granted to the specific application and data on that computer.

Working of Telnet

- (i) A user is logged in to the local system and invokes a TELNET program (the TELNET client) by typing **telnet<host address> or telnet <IP address>**
- (ii) The TELNET client is started on the local machine (if it isn't already running). The client then establishes a TCP connection with the TELNET server on the destination system.
- (iii) Once the connection has been established, the client program accepts characters from the keyboard feed by the user and passes one character at a time, to the TELNET server.
- (iv) The server on the destination machine accepts the characters sent to it by the client and passes them to a terminal server.
- (v) The terminal server gives outputs back to the TELNET server and displays them on the user's screen.

The user can terminate the telnet session by typing LOGOFF or LOGOUT on the system prompt.

8.14 WIRELESS/MOBILE COMMUNICATION

8.14.1 GSM (Global System for Mobile Communication)

GSM stands for Global System for Mobile communication. It provides its subscribers with roaming facility so that they can use their mobile phones all over the world to communicate with each other. GSM provides digital signalling as well as digital call facility and is so considered as a second generation (2G) mobile phone system. It provides consumers with better voice quality and low-cost alternatives to making calls such as short message service (SMS). It has an ability to carry 64 kbps to 120 mbps of data rates. The key feature of GSM is the Subscriber Identity Module (SIM), generally known as SIM card. It is a detachable smart card that contains the subscriber's information along the phone book. This allows the user to use the phone book and other information even after changing the handset.

CTM: GSM is a wireless communication medium that provides the user with roaming facility, good voice quality, SMS, etc., through digital signals.

8.14.2 CDMA (Code Division Multiple Access)

CDMA stands for Code Division Multiple Access. CDMA is an example of multiple access, where several transmitters can send information simultaneously over a single communication channel. It is a channel access method used by various radio communication technologies. It allows the

division of transmission medium into different channels so that transmission from different stations is separated from each other. It employs spread spectrum digital technology where the data is fragmented into small chunks over a number of different frequencies available for use. Analog to digital conversion takes place where audio input is first digitized into binary elements. The CDMA system works directly on 64 kbit/sec digital signals.

CTM: CDMA is a digital cellular technology that uses spread spectrum technique where the entire bandwidth is divided among several users for transmission of data.

8.14.3 GPRS (General Packet Radio Services)

GPRS or General Packet Radio Services provides various features over 2G phones with respect to high speed data transfer. A user can send and receive data at the same time and thus uses the same bandwidth for both purposes. Using GPRS technology, a user can make a call and at the same time receive a message without disconnecting the call. However, GPRS usage is charged for the amount of data sent or received. GPRS can provide data rates up to 32kbps to 48kbps. With this data rate, email messages, video streaming, audio files, etc., can be downloaded and, therefore, can be called 2.5G technology as it lies between the second (2G) and third (3G) generations of mobile telephony.

Services Provided By GPRS

- (a) Sending and receiving text messages.
- (b) Internet access.
- (c) Multimedia Messaging Service (MMS).
- (d) Internet applications for smart devices through Wireless Application Protocol (WAP).
- (e) Networking facility with one person or with several persons in a group, i.e., videoconferencing.

CTM: GPRS provides high speed data transfer. A user is allowed to download video streaming, audio files, email messages, etc.

8.14.4 Wireless in Local Loop (WLL)

WLL provides the subscribers with wireless phone facility to communicate with each other in order to get better voice quality. It employs the use of electromagnetic radiation to connect subscribers to the local exchange without the use of wires. The user can use wireless phone, speaker phone and parallel phones for communication with each other. In traditional telephone networks, phone would be connected to the nearest exchange through a pair of copper wires. But in Wireless Local Loop (wireless in local loop) technology, the subscriber is connected to the nearest exchange through a radio link instead of copper wires. Wireless in local loop is cheaper and quicker than copper wire connectivity. As the cost of copper along with the cost of digging increases over time, this method proves cheaper than using copper wires. It is used in remote areas where digging for copper wires is not possible.

There are various technologies like Frequency Division Multiple Access (FDMA), Time Division Multiple Access (TDMA) and Code Division Multiple Access (CDMA) used for wireless in local loop. In crowded urban localities where permission to dig may be almost impossible to get, wireless technology is advised. However, there are also some disadvantages of using wireless in local loop. First, as the distance between a receiver and transmitter increases in a wireless system, the strength

of the signal at the receiving end decreases even if there are no obstacles in the way. Second, as the electromagnetic waves are propagated through air, a signal travelling from a transmitter may take alternative paths on its way to the receiver. These paths may cause delays in the received signal due to the extra distance travelled.

CTM: Wireless in local loop is a system that is similar to telephone system which provides wireless telecommunication by deploying a multiplicity of multichannel transceivers.

8.15 MOBILE TELECOMMUNICATION TECHNOLOGIES—1G, 2G, 3G & 4G

➤ 1G Technology

1G technology was used in the first mobile phones. 1G was introduced in the 1980's. Transmission of voice data took place through analog radio signals. 1G network was used for voice calls and not for transmitting text messages. *For example, NMT, C-Nets, AMPS, TACS.*

Salient features of 1G technology:

- It provides data rates up to 2.4 kbps.
- It uses analog signals.
- Voice quality is not very good.
- It does not support transmission of text messages.
- It does not provide security.

➤ 2G Technology

2G technology is the first digital cellular system that was launched in the early 1990's that provides high data transmission rate in digital format. 2G also introduced data services for mobiles, starting with SMS. *For example, D-AMPS, GSM/GPRS, CDMAOne.*

Salient features of 2G technology:

- Good quality of sound.
- Higher data rates up to 64 kbps.
- Improved security mechanism.
- Transmission of data such as text messages in digital format.
- Support transfer of picture messages and MMS.
- It does not support transfer of complex data such as videos.
- It allows multiple users per radio channel with each user talking one at a time.
- Digital transmissions enable compressed voice and multiplexing multiple users per channel.

➤ 3G Technology

3G technology was introduced in the year 2000 and provides much higher data rates with speed from 144kbps to 2mbps.

3G technology has introduced more efficient ways of carrying data, making it possible to have faster web services, live chat, fast downloading, videoconferencing, etc., over mobile phones. *For example, CDMA2000/EV-DO,WCDMA/HSPA+,TD-SCDMA.*

Salient features of 3G technology:

- It has introduced more efficient ways of carrying data with faster web services.
- Live chat, fast downloading, videoconferencing, etc., are also possible over mobile phones.
- It allows the user to play 3D games.
- A user can see live streaming on smartphones.
- It provides broadband internet services.
- It allows the user to send and receive large email messages.
- High bandwidth is required.
- Splits channel into time intervals enabling a single user to get all the resources at once.

➤ 4G Technology

Unlike previous generations of mobile technology, 4G mobile technology uses ultra-high broadband internet services with much faster data speed, typically between 100 mbps-1gbps. Now, 4G rules the mobile market. Unlike previous generations of mobile technology, 4G mobile technology is used for internet access on computers also, and it is totally wireless. 4G provides internet access, high quality streaming video and “anytime, anywhere” voice and data transmission at a much faster speed than 3G. The “anytime, anywhere” feature of 4G is also referred to as “MAGIC” (Mobile multimedia; Anytime/anywhere; Global mobility support; Integrated wireless solution; Customized personal services).

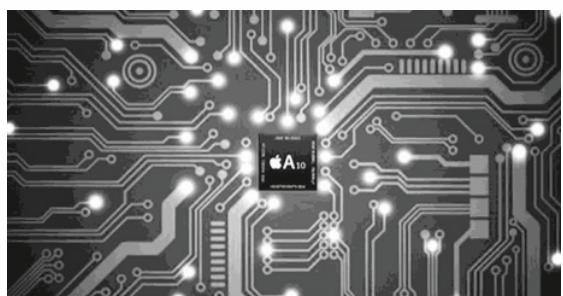
Salient features of 4G technology:

- It is used for internet access on computers also and is totally wireless.
- 4G provides internet access, high quality streaming video and “anytime, anywhere” voice and data transmission at a much faster speed than 3G.
- It delivers faster and better mobile broadband experiences.
- It provides more data capacity for richer content and more connections.

The “anytime, anywhere” feature of 4G is also referred to as “MAGIC” (Mobile multimedia) anytime/anywhere.

8.16 MOBILE PROCESSORS

Processors are required to run an operating system, be it a desktop, laptop or a mobile. Processors provide the necessary resources to start an operating system, run applications and do certain tasks. Today's smartphones and mobile processors are very powerful, so much so that they can compete with desktop computers.



Processors are now available in many cores. First it was single core, then came the dual core, and we now have quad core, hexa core and even octa core processors. Most processors available today are 64-bit as against 32-bit earlier. The processing speed has also touched 3.0–3.5 GHz now. The ability to equip mobile processors with GPU (Graphics Processing Unit) has enabled the devices to churn out best graphic pictures, have 3-D capability, Virtual Reality and 4K recording capability. The improved processor technology has also made modern mobile devices more power-efficient.

Today, there are many processors available in the market. QUALCOMM, Apple mobile processors, Intel mobile processors and some other giants are ruling the market. Let us discuss these mobile processors.

8.16.1 Qualcomm Snapdragon

Qualcomm Technologies is a US-based company. Qualcomm first became a known brand when they introduced the CDMA technology. Qualcomm is actively involved in technology related to semiconductor designing for mobile devices, tracking devices, satellite phones, virtual reality, wireless charging, communications, etc. Qualcomm is now known for its Snapdragon brand which is responsible for marketing mobile processors and LTE modems (4G).



Snapdragon became a big name in the processors market after it introduced the first 1 GHz processor when the average speed of most smartphones was only 512 MHz.

Since 2005, Snapdragon has come out with a number of series—S1, S2, S3, S4, S200, S400, S600, and S800. S800 series has already released S800, S801, S805, S808, S810, S820 and S821. The most common Qualcomm processor in the medium-to high-end phone market is either S820 or S821, which is the latest version.

8.16.2 Apple Mobile Processors

Apple does not manufacture microprocessors. Instead, it enters into contracts with processor-manufacturing companies, mainly Samsung and TSMC, for making custom-built processors that suit its design and performance expectations. For instance, A9 14nm processor was built by Samsung, while the A9 16nm version was built by TSMC.

Apple A series is designed for processors to be used in iPhone, iPad, iPad Touch and Apple TV.



Some of the processors in the series are A4, A5, A5X, A6, A6X, A7, A8, A8X, A9, A9X and A10.

- **Apple A10 Fusion** is the latest processor which is used in iPhone 7 and iPhone 7 Plus. A10 is a quad core built on 16 nm FinFET processor capable of running at 2.4GHz speed and a hexa core PowerVR GPU. A10 is twice as fast as its predecessor A9 and improves graphic processing by 50%. This processor is manufactured by TSMC.
- **Apple S series** is designed for processors to be used in Apple Watch. Some of the processors in the series are Apple S1, Apple S1P and Apple S2.

The current version, Apple S2, is a dual core processor with built-in GPS used in Apple Watch Series 2. The processor is manufactured by Samsung under a contract with Apple.

- **Apple W series** is used in headphones for wireless audio connectivity. The current series, Apple W1, is used in wireless headphones and AirPods.
- **Apple T series** is designed to be used in TouchID sensors in MacBook Pro. The only version released till date is Apple T1.

8.16.3 Intel Atom and Core M Processors

Intel is an American multinational company synonymous with PC and microprocessors. Atom is the brand name given for the low power-consuming and low-cost 32-bit and 64-bit chips manufactured for smartphones and tablets.

Intel processors are based on X86 architecture which is more powerful than ARM but consumes more power compared to ARM architecture. The latest versions of Intel processors have reduced the power consumption, bringing it down to less than 5 watts, which is ideal for all mobile devices. Though Atom processors in the beginning supported only Windows, they now support all major mobile operating systems.



Intel Atom processors are currently used in Atom X5 and X7 series. These chips are 64-bit quad core processors in 14 nm size with speeds of up to 1.6 GHz that can be scaled up to 2.4 Ghz. Intel also released **Intel Core M** ultra low-voltage microprocessors designed for ultra-thin notebooks, mobile devices and 2-in-1 convertibles. The processor consumes 4.5 watts or less power, making it ideal for long battery life. These are dual core processors with a speed of about 1.5 GHz which can be scaled up to 3.2 GHz. Intel Core M processors offer 40% boost in CPU and graphics performance as compared to the earlier versions.

8.16.4 Nvidia Tegra

Nvidia Corporation is a US-based technology company which specializes in making processing units for graphics, gaming units and mobile devices. Nvidia develops chips for smartphones, such as transmitting fax, voice data, tablets and mobile devices under the brand Tegra.



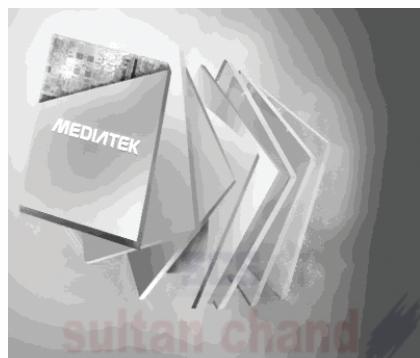
Tegra processors are built on 64-bit ARM architecture. Tegra has already marketed Tegra 1, Tegra 3, Tegra 4, Tegra 4i, Tegra K1, Tegra X1.

Tegra X1 is currently the most advanced Tegra chip in the market. The processor is Quad Core with 256 GPU cores and 4K video capabilities. The chips are built on 20 nm technology. The processor is currently used in Nvidia SHIELD Android TV.

The Tegra processors mainly used in smartphones and tablets are Tegra 4, Tegra 4i and Tegra K1.

8.16.5 MediaTek

MediaTek is a Taiwanese semiconductor company providing chips for mobile devices, HDTVs and other electronic devices.



MediaTek processors are built on 64-bit ARM architecture. The latest MediaTek processor supports up to 3 GHz speed. They come in a variety of cores such as dual core (2 core), quad core (4 core), hexa core (6 core) and deca core (10 core).

The latest processors from MediaTek, **Helio X20** and **Helio X25**, are used in smartphones and tablets. MediaTek processors are mostly popular with Chinese manufacturers. Xiaomi, Meizu,

LeEco Le, Yu, etc., use them in smartphones. Acer, Asus, Lenovo, Amazon Fire HD, QMobile are some of the other manufacturers that use MediaTek processors in their tablets.

Helio X30 and Helio X27, the latest from the company's stable, use 10 nm and 20 nm processors respectively. Both are deca core with 2 dual core and a single dual core built inside the processor.

8.16.6 HiSilicon

HiSilicon is a Chinese company specializing in semiconductor technology. The company, owned by Huawei, creates chips based on ARM architecture. It is the largest domestic integrated circuit designer in China.

Some of the processors released by HiSilicon are K3V1, K3V2, Kirin 620, Kirin 650, Kirin 910, Kirin 920, Kirin 930, Kirin 950 and Kirin 960. Some of the devices with Kirin 950 are Honor 8, Huawei mate 8 and Huawei MediaPad M3.



HISILICON

Kirin 960 is the latest model to be released in the series. It is built on 64-bit ARM architecture on 16 nm FinFET technology. The processor is quad core capable of attaining a speed of 2.4GHz.

8.16.7 Samsung Exynos

Exynos is a brand of Samsung Electronics which makes processors based on ARM architecture. Some of the processors in the series are Exynos 7 Dual, Exynos 7420, Exynos 7 Octa 7580, Exynos 7 Octa 7870.



Exynos 8 Octa 8890 is the latest processor from Exynos. The processor is equipped with Octa Core on 64-bit ARM architecture with Mali GPU. The processor is capable of running at a speed of 2.3 GHz with support for 3D gaming and 4K UHD resolution. The chips are built on 14 nm technology. **Exynos 8 Octa 8890** is used in Samsung Galaxy S6 and S6 edge.

8.17 ELECTRONIC MAIL (EMAIL)

Email is a method of exchanging digital messages from a sender to one or more recipients. Some common email protocols are:

- **Internet Message Access Protocol (IMAP):** A protocol for receiving email messages on the internet.
- **Post Office Protocol Version 3 (POP3):** A protocol used by email clients to retrieve messages from remote servers.
- **Simple Mail Transfer Protocol (SMTP):** A protocol used for sending email messages on the internet.
- **GOPHER:** Another tool of the internet is Gopher, a menu-based program that enables us to browse for information without knowing where the material is located. It lets us search a list of resources and then sends the material to us.

Some other networking-related terms are:

1. **Domain Name System (DNS):** It translates network address (such as IP addresses) into terms understood by humans (such as domain names) and vice versa.
2. **Domain Host Configuration Protocol (DHCP):** It automatically assigns internet addresses to computers and users.

8.18 VIDEOCONFERENCING

Videoconferencing is a communications technology that integrates video and audio to connect users anywhere in the world as if they were in the same room. This term usually refers to communication between three or more users who are in at least two locations, rather than one-on-one communication and it often includes multiple people at each location. Each user or group of users who is participating in a videoconference typically must have a computer, a camera, a microphone, a video screen and a sound system.

Basically, this is a system that allows us to conduct meetings or trainings in different places simultaneously. So, this technology is especially popular in the field of business because it allows meetings or conferences to be held without the need for all the participants to travel to a single location, so it saves time and money.

The most popular software used for videoconferencing is:

- (a) TCP Cam
- (b) Ekiga
- (c) Skype

8.19 PROTOCOLS FOR CHAT AND VIDEOCONFERENCING

With the arrival of internet, communication formats such as chat and videoconferencing, etc., have gained popularity. In this section, we shall talk about some common chat and videoconferencing protocols. The most common chat protocol is IRC (Internet Relay Chat) while the most common videoconferencing protocols are H.323 and SIP (Session Initiation Protocol).

8.19.1 IRC (Internet Relay Chat)

IRC is an application layer protocol that allows users to share text messages. It uses client server model where the clients install IRC client program on their system so that they can communicate with IRC chat server to transfer messages to other clients. In fact, the IRC client sends request to IRC client server and the server forwards this request to another client to enable them to communicate with each other. It provides one-to-one communication as well as group communication for chatting and file sharing, such as Talk City. IRC networks such as the Undernet provide servers and help us download an IRC client on our PC.

The IRC protocol was developed over four years since it was first implemented as a means for users on a BBS to chat amongst themselves.

CTM: IRC is used for chatting by sending and receiving text messages. The sender sends request to IRC server, which then forwards this request to another client to communicate with each other.

8.20 PROTOCOL FOR VoIP

VoIP stands for Voice over Internet Protocol. For transferring of voice, voice over internet protocol is used. The voice calls are first digitized, compressed and then fragmented into small packets, which are then relayed by Internet Protocol (IP) cross network. Voice-over-IP (VoIP) implementation enables users to carry voice traffic (*For example, telephone calls and faxes*) over an IP network. So, VoIP can be achieved on any data network that uses IP, like the internet, intranets and Local Area Networks (LAN). As the data is transmitted in the form of packets, VoIP uses packet switching technology where each packet follows best route to reach its destination. VoIP allows both voice and data communications to be run over a single network, which can significantly reduce infrastructure costs.

There are 3 main causes for the evolution of the voice-over IP market:

1. Low-cost phone calls
2. Add-on services and unified messaging
3. Merging of data/voice infrastructures

Services provided by VoIP are:

Phone to phone, PC to phone, phone to PC, fax to email, email to fax, fax to fax, voice to email, IP Phone, transparent CCS (TCCS), toll free number (1-800), class services, call centre applications, VPN, Unified Messaging, Wireless Connectivity, IN Applications using SS7, IP PABX and soft switch implementation.

The various protocols used for VoIP are:

1. H.323
2. Session Initiation Protocol (SIP)

CTM: VoIP is a protocol that is used for transmitting voice data and multimedia data over internet protocol. It uses high speed broadband internet connection.

1. H.323 Protocol for VoIP

H.323 is a protocol that provides communication for multimedia services such as audio, video and data communication over packet-based network. It specifies the standards and protocols for all these services. As H.323 provides various communication facilities, it can be applied in a wide variety of areas—consumer, business and entertainment applications. H.323 supports call set-up, teardown and forwarding/transfer. A key feature of H.323 is Quality of Service (QoS). QoS technology allows real-time prioritization and traffic management constraints to be placed on “best-effort” packet delivery systems like TCP/IP over Ethernet. It can be applied in a variety of mechanisms:

- Audio only
- Audio and data
- Audio and video
- Audio, video and data

2. SIP Protocol for VoIP

The expanded form of SIP is Session Initiation Protocol. SIP is a communication protocol (more specifically, a signalling protocol) originally developed in 1996. It uses IP protocol that establishes, modifies and terminates VoIP telephone calls. It provides videoconferencing service to the users, so that a user can communicate with more than one person at a time. Other SIP applications include streaming multimedia distribution, instant messaging and information, file transfer, fax over IP and online games. SIP transparently supports name mapping and redirection services which support personal mobility.

8.21 NETWORK SECURITY CONCEPTS

Network is used for sharing, messaging and collaboration of data. However, more and more network is used for this purpose, the lesser is the security of data, either in terms of viruses or hacking and other cyberattacks. To prevent the network from these malicious and/or unethical practices, various strategies and choices are available that work as building blocks of network security. These include password authentication, digital signature, challenge handshake authentication protocol, etc.

8.22 Types of Threats and Prevention

1. Viruses

Viruses are small programs that are written intentionally to damage the data and files on a system. These programs are spread from one computer system to another which interrupts the normal functioning of a computer. Viruses can attack any part of a computer's software such as boot block, operating system, system areas, files and application program macros. Viruses are most easily spread with email attachments.

Viruses are broadly classified into three types:

- (a) **File Infector Viruses:** Viruses that attach themselves to a program file.
- (b) **Boot Sector Viruses:** Viruses that install themselves in boot sectors of hard drive.
- (c) **Macro Viruses:** They infect data files and corrupt the data.

Characteristics of Viruses:

- (a) Speed of a computer system becomes slower than normal.
- (b) Computer system frequently hangs up.
- (c) Computer restarts automatically after every few minutes.
- (d) Various applications of computer do not function properly.
- (e) Dialog boxes, menus and other error message windows are distorted.

Damage caused by Viruses:

- (a) They corrupt file allocation table which results in corrupting the entire file system.
- (b) Replication of files occurs that decreases space in hard disk.
- (c) They can destroy system programs and files.
- (d) They can cause destruction of specific executable files and alteration in data files resulting in reinstallation of a system.
- (e) They can format specific tracks on the disks or format the entire disks.

Prevention from Viruses:

- (a) Use antivirus software to permanently remove viruses.
- (b) Always scan pen drives and other secondary storage media in order to detect viruses and safeguard your system.
- (c) Frequent update your computer system.

CTM: A virus is a malicious program that damages data and files of a system and can also corrupt the file allocation table.

2. Worms

A worm is a self-replicating program that runs independently and travels across network connections. The characteristics of viruses and worms are more or less same, but a worm causes more damage.

CTM: A worm is a computer program which can self-replicate and propagate over the network, with or without human intervention, and has malicious content.

Characteristics of Worm:

- (a) It replicates itself.
- (b) Unlike virus, worm does not require host and it is self-contained.
- (c) It spreads across networks through email, instant messaging or junk mails.
- (d) Worms run independently.

The various types of Worms are:

- (a) **Email Worms:** Worms spread through any email which contains any file attachment or link to any infected website.
- (b) **Instant Messaging Worms:** Worms spread through instant messaging across the network by sending mails to infected website.
- (c) **Internet Worms:** The infected worm will use the contact list of the user's chat-room profile or instant-message program to send links to infected websites. These are not as effective as email worms as the recipient needs to accept the message and click the link. The users of the particular program are affected by it.
- (d) **File Sharing Network Worms:** These types of worms are downloaded along with the required files downloaded by the user. A user is not aware about this worm and, therefore, when the user downloads any file, the worm will copy itself into a shared folder with an unassuming name. When another user on the network downloads files from the shared folder, the worm gets downloaded on their system also. In this way, the worm copies itself and repeats the process. In 2004, a worm called "Phatbot" infected millions of computers in this way and had the ability to steal personal information, including credit card details, and send spam on a large scale.

Damage caused by Worms:

- (a) A worm may corrupt the files on the host computer.
- (b) It may affect communication between the host and other systems.

- (c) It may disable the antivirus software on the host, which will enable it to cause more damage.
- (d) Bulk email chaining can be created with an intention to guess email passwords.
- (e) A worm consumes too much system memory (or network bandwidth), causing web servers, network servers and individual computers to stop responding.

3. Trojan Horse

A Trojan horse is a kind of virus that looks safe but has hidden effects. It is a hidden code in a program such as a game or a spreadsheet that can damage the system when running these applications. It can destroy or alter information on a computer system in the background. Unlike viruses, Trojans do not replicate themselves but they are destructive. Trojans are executable programs, which means that when a user runs any application or plays games, they work behind that application and can damage the system completely.

There are several types of Trojan Horse:

- (a) Remote access Trojan horse
- (b) Data sending Trojans
- (c) Destructive Trojans
- (d) Proxy Trojans
- (e) FTP Trojans
- (f) Denial-of-service attack Trojans

Trojans are generally spread through email attachments and exchange of disks and information between computers. Worms can also spread Trojans. The damage caused by Trojans is similar to that caused by viruses. Sometimes the user is unaware about a Trojan because of its masking effect as it runs as a hidden code.

CTM: Trojan horse is a hidden code that looks safe but it has some hidden effects while running applications.

4. Spams

Spam is an unwanted bulk mail which is sent by an unauthorized or unidentified person in order to eat the entire disk space. In non-malicious form, it floods the internet with many copies of the same message to be sent to a user which he may not otherwise receive. Generally, it is in the form of pampering the user with various discount schemes, or commercial advertising, often for dubious products, get rich quick schemes, etc. Mobile phone spam is a form of spamming that uses text message service of mobile phone.

Spams can be avoided by using email filtering, spam traps, etc.

CTM: Spam refers to electronic junk mail that eats up the entire computer's space.

8.23 COOKIES

A cookie is also known as an HTTP cookie, web cookie, internet cookie or browser cookie. When the user browses any website, a cookie identifies users and prepares web pages for them which are then sent to the web server for later use. Cookie is basically a piece of data that is stored by the

website on the user's hard disk. The information is stored in the form of name value pair. Generally, the cookies folder is stored in c:\windows\cookies.

Working of Cookies

1. The user enters the name of the website on the browser.
2. The browser contacts the DNS server to convert the domain name into IP address. The browser looks for the cookies on the hard disk.
3. The IP address is used to contact the corresponding server along with cookies data.
4. If no cookies data is supplied, then the website comes to know that the user is visiting the website for the first time.
5. The server creates an ID of a person which is then stored by cookies on the hard disk.

The information which is stored by a website is known as state information. The information can be of the following types:

1. How many visitors have visited the site?
2. How many are new visitors?
3. How many are repeat visitors?
4. What is the frequency of a particular visitor?

CTM: Cookies are the messages which are stored by the website on a user's hard disk whenever they visit any website.

8.24 FIREWALL

A firewall is a software that protects the private network from unauthorized user access. The firewall filters the information coming from the internet to the network or a computer to protect the system. Firewall exists both as a software solution and as a hardware application. In the form of hardware firewall such as router, it protects the network, and in terms of software firewall, it helps in preventing the spread of virus from one computer to another. A firewall is a network security system, either hardware- or software-based, that controls incoming and outgoing network traffic based on a set of rules. Various examples of firewalls are CISCO firewall, NetGear firewall, Netscreen 25, etc.

A firewall can use various methods for filtering the information such as:

1. **Packet filtering:** In packet filtering, the data, which is outgoing or incoming in the form of packets, is filtered. Packet filter looks at each packet entering or leaving the network and accepts or rejects it based on user-defined rules. Packet filtering is fairly effective and transparent to users, but it is difficult to configure. In addition, it is susceptible to IP spoofing.
2. **Proxy service:** The information which is requested is not directly sent to the person who makes the request; rather, the information is first received in firewall and then sent to the proxy server. The proxy server intercepts all messages entering and leaving the network and effectively hides the true network addresses.
3. **IP address blocking:** If the data is coming from a network or website that contains some unwanted data, then the data from that particular IP address or domain name is blocked by firewall.

4. **Protocol blocking:** The firewall can be set to disallow a particular protocol service to a particular user or group of users.
5. **Application Gateway:** It applies security mechanisms to specific applications, such as FTP and Telnet servers. This is very effective but can impose performance degradation.
6. **Circuit-level Gateway:** It applies security mechanisms when a TCP or UDP connection is established. Once the connection has been made, packets can flow between the hosts without further checking.
7. **Port blocking:** The firewall can be used to block a particular port. Generally, HTTP and FTP services use port 80 and port 21.

Firewall protects the user from the following:

1. Remote login by others who are not authorized to gain access to the system
2. Application backdoors
3. SMTP session hijacking
4. Email bombs
5. Viruses and macros
6. Spam or junk mails

CTM: Firewall is a system that is designed to protect the network from illegal use by an unauthorized person.

8.25 INDIA'S IT ACT

The **Information Technology Act 2000** (also known as **ITA-2000** or the **IT Act**) is an Act of the Indian Parliament (No. 21 of 2000) notified on October 17, 2000.

IT Act 2000 has been defined as:

"An Act to provide legal recognition for transactions carried out by means of electronic data interchange and other means of electronic communication, commonly referred to as 'electronic commerce', which involve the use of alternatives to paper-based methods of communication and storage of information, to facilitate electronic filing of documents with the Government agencies and further to amend the Indian Penal Code, the Indian Evidence Act, 1872, the Bankers' Books Evidence Act, 1891 and the Reserve Bank of India Act, 1934 and for matters connected therewith or incidental thereto."

8.26 CYBER LAW

The law that governs the cyber space is known as cyber law. It is the law which deals with various computer-related activities. In essence, cyber law is an attempt to apply laws designed for the physical world to human activity on the internet. In India, the IT Act, 2000, as amended by the IT (Amendment) Act, 2008, is known as the cyber law. It has a separate Chapter XI entitled "Offences" in which various cyber crimes have been declared as penal offences punishable with imprisonment and fine. A large number of cyber-criminal activities such as hacking passwords or accessing files by an unauthorized person or getting private information, etc., have been prevented by implementing cyber laws.

As the usage of internet has been growing at an exponential rate, it has become very important to protect individual users and organizations from unauthorized intruders into the system.

8.27 CYBER CRIMES

When any crime is committed over the internet, it is referred to as cyber crime. There are many types of cyber crimes and the most common ones are explained below:

1. **Hacking:** Gaining knowledge about someone's private and sensitive information by getting access to their computer system illegally is known as hacking. This is different from ethical hacking, which many organizations use to check their internet security protection. In hacking, a criminal uses a variety of software so as to enter a person's computer and that person may not be aware of his computer being accessed from a remote location.
2. **Theft:** Theft occurs when a person downloads music, movies, games and software by violating copyright. There are even peer-sharing websites which encourage software piracy and many of these websites are now being targeted by the Federal Bureau of Investigation (FBI).
3. **Cyber Stalking:** Cyber stalking is a kind of online harassment where the victim gets unwanted abusive online messages and emails. Typically, these stalkers know their victims and instead of resorting to offline stalking, they use the internet to stalk. If they notice that cyber stalking is not having the desired effect, they begin offline stalking along with cyber stalking to make their victim's life miserable.
4. **Identity Theft:** This has become a major problem with people using the internet for cash transactions and banking services. In this cyber crime, a criminal accesses data about a person's bank account, credit card, social security card, debit card and other sensitive information to gain money or to buy things online in the victim's name that can result in major financial loss for the victim and even spoil the victim's credit history.
5. **Malicious Software:** These are internet-based software or programs known as pirated software that are used to disrupt proper functioning of the network. The software is used to steal sensitive information or data that can cause damage to existing software in a computer system.
6. **Child Pornography:** In this cyber crime, defaulters create, distribute or access materials that sexually exploit underage children. The criminals solicit minors via chat rooms for the purpose of child pornography. The FBI has been spending a lot of time monitoring chat rooms frequented by children with the hope of reducing and preventing child abuse and soliciting.
7. **Sales and Investment Fraud:** With the increase in e-commerce, the application of digital technology to fraudulent endeavours has become that much greater. The use of telephone for fraudulent sales pitches, deceptive charitable solicitations or bogus investment overtures is becoming increasingly common. There are some fraudulent sites that sell poor quality products at cheaper rates. They also promise the customers heavy discounts and freebies.
8. **Electronic Funds Transfer Fraud:** A cyber crime occurs when there is a transfer of funds which may be intercepted and diverted. Valid credit card numbers can be hacked electronically and then misused by a fraudulent person or organization.

9. **Defamation:** It involves a cyber crime with the intent of lowering the dignity of someone by hacking into their email account and sending mails using vulgar language to an unknown person's account.

CTM: Cyber law defines all the legal and regulatory aspects of internet and the World Wide Web.

10. **Assault by Threat:** It refers to threatening a person or his family members with dire consequences through the use of a computer network, *i.e.*, email, videos or phones.
11. **Denial of Service (DoS) Attacks:** A DoS attack is an attack by which legitimate users of a computer are denied access or use of the resources of that computer. Generally, DoS attacks do not allow the attacker to modify or access information on the computer. A DoS attack can be of the following types:
- **Denial of Access to Information:** Refers to an attack in which information sought by a legitimate user is either destroyed or changed to some unsubtle form.
 - **Denial of Access to Applications:** Refers to an attack that denies a user from accessing an application by making it either unusable or unavailable. This is usually done to prevent the user (or organization) from using applications to perform any task.
 - **Denial of Access to Systems:** Refers to the unavailability of the system, including all the applications installed on the system or any data stored on the system.
 - **Denial of Access to Communications:** Refers to a common type of attack that can range from cutting wires and jamming radio communications to flooding networks with excessive traffic. An example of this type of attack is flooding a computer with junk mail.

8.28 IPR ISSUES

IPR stands for intellectual property rights which is the right to intangible property such as music, literature and artistic work created by a person. **Intellectual property (IP)** is a legal term that refers to creations of the mind. Intellectual property rights may be protected by patents, copyrights, industrial design rights, trademarks, trade dress and, in some jurisdictions, trade secrets. The owner of intellectual property is the person who has developed the product or the organization which has funded it. Safeguarding intellectual property from illegal use can be done by giving some exclusive rights to the owner of that property. These rights also promote creativity and dissemination and application of its result and encourage fair trading which helps in developing social and economic areas of a country.

IPR-related issues in India like patents, trademarks, copyrights, designs and geographical indications are governed by the Patents Act, 1970 and Patent Rules, 2003, Trademarks Act, 1999 and the Trademarks Rules, 2002, Indian Copyrights Act, 1957, Design Act, 2000 and Design Rules, 2001, and the Geographical Indications of Goods (Registration & Protection) Act, 1999 and the Geographical Indications of Goods (Registration & Protection) Rules, 2002, respectively.

Prevention from Security Threats

There are various methods to protect network threats. These protection methods are as follows:

1. **Authorization:** Authorization means Intrusion Detection. Authorization means to grant a person access to a network for legal use. It is an act of giving authority or legal identity to a user to become an authorized user of the system. Authorization checks can also be

implemented to a program or process to make data free of risk, such as entering of viruses, worms or Trojan horse.

2. **Authentication:** To determine the identity of a person before granting access to private or sensitive data or information is known as authentication. Verifying the identity of an intruder or a person is difficult and needs complex protocols based on cryptography.
3. **Privacy:** The data which is accessible only to an authorized person is known as private data.
4. **Secrecy:** Hiding some relevant information from an unauthorized person is called secrecy.
5. **Biometric System:** Biometric system forms the most secure level of authorization. It involves digital signature, finger prints, retinal patterns, etc., to establish identity.
6. **Password Protection:** To protect the system or network from an unauthorized person, a system must be password protected. A password protected system allows access to resources based upon a secret word entered by the user.
7. **File Permission:** A user can give access to a person to read a file, write to a file, open a file, modify a file, etc. Different types of permissions can be given to a specific person according to their authorization. Each file has an access control list attribute that describes which user or group accounts have what type of access to the file.

Three types of file access permissions that are granted to a user:

- (a) *Read:* Allows a user to view and read a file.
- (b) *Write:* Allows a user to edit and write on a file.
- (c) *Execute:* Allows a user to execute a file.

File access permission is granted to three types of users:

- (a) *Owner:* Refers to the user who has created the file.
- (b) *Group:* Refers to the group of users which is working with the file owner.
- (c) *Others:* Refers to all other users.

8. **Firewall:** A firewall is a part of a computer system or network that is designed to block unauthorized access while permitting authorized communication. It is a programmer software or device or set of devices configured to permit, deny, encrypt, decrypt, or proxy wall (in and out) computer traffic between different security domains based upon a set of rules and other criteria.
9. **Proper Security Policy:** An organization's security policy is a formal statement consisting of the rules that its employees need to follow to access information about the organization. The policy should clearly communicate the security goals to all the users, administrators and managers of the organization. A good security policy must be:
 - enforced with adequate security tools.
 - able to define the areas of responsibility for a user, an administrator or a manager.
 - able to adjust itself according to the changing configurations of computer networks.

8.29 HACKING

Hacking is the practice of modifying the features of a system in order to accomplish a goal outside the creator's original purpose. A person who consistently engages in hacking activities and has accepted hacking as a lifestyle and philosophy of their choice is called a hacker.

Computer hacking is the most popular form of hacking nowadays, especially in the field of computer security. However hacking exists in many other forms also, such as phone hacking, brain hacking, etc., and it is not limited to either of them.

Due to the mass attention given to black hat hackers by the media, the whole hacking term is often mistaken for any security-related cyber crime.

8.30 INTRODUCTION TO WEB SERVICES

Internet offers several important and extensively used features or services which are described as follows:

8.30.1 WWW (World Wide Web)

WWW is an information service that can be used for sending and receiving information over the internet through interlinked hypertext documents. Web pages may contain text, images, videos and other multimedia components as well as web navigation features consisting of hyperlinks. The documents are formatted in a markup language called HTML (Hyper Text Markup Language) that provides links to other documents as well as graphics, audio and video files. The World Wide Web is based upon client-server architecture where a client sends a request and the server processes that request and sends responses. A WWW client is called a web browser and a WWW server is called a web server.

CTM: WWW is a set of programs and protocols that allows the user to create and display multimedia web pages and is linked to the internet.

8.30.2 Hyper Text Markup Language (HTML)

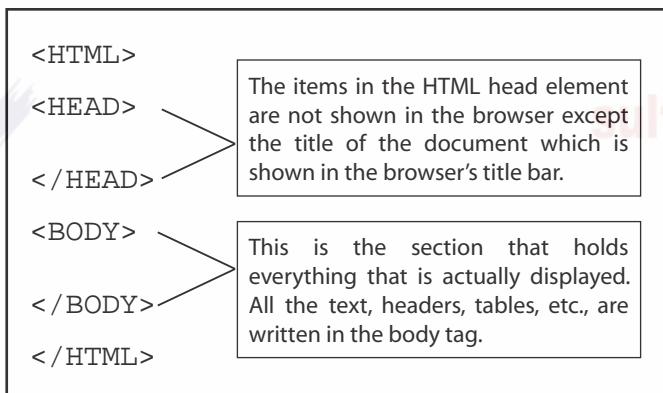
HTML is a language that is used to create web pages which are then displayed by web browsers. This language tells the browser how to display text, pictures and links on the screen. HTML is a language that supports multimedia documents and consists of audio, video, graphics, pictures, etc. It also helps in creating hyperlinks so that various documents can be linked together. When a person clicks on a specified link, the document related to that link displays. HTML is a document layout and hyperlink specification language, *i.e.*, a language that uses various coded elements known as tags for formatting the document and to specify the hyperlinks.

HTML document can be written using any text editor such as NOTEPAD or NOTEPAD2 and save a file with extension either .HTM or .HTML

Any HTML document, in general, contains at least three elements—HTML, HEAD, and BODY.

These elements are specified by the following respective tags:

1. <HTML> . . . </HTML>
2. <HEAD> . . . </HEAD>
3. <BODY> . . . </BODY>



The structure of HTML document is as follows:

```
<HTML>
  <HEAD>
    <TITLE> This is my first page</TITLE>
  </HEAD>
  <BODY> I am writing my first page using html.
  </BODY>
</HTML>
```

Now, execute this file using any web browser and it will display a page as below:



CTM: HTML is a Markup language that enables users to create web pages and format them using predefined tags. Tags are called coded elements.

8.30.3 Extensible Markup Language (XML)

Extensible Markup Language is a text-based Markup Language that allows the user to create their own tags to store data in a structured format. However, these structured formats can be represented in different ways. In XML, the tags are not predefined; rather, they are created by the user for their own purpose. Unlike HTML, in XML, tags are case-sensitive and each tag must have a corresponding closing tag. It is a general-purpose specification that allows users to create custom Markup language. XML was designed to carry data and not to display data. For formatting data, a separate style sheet known as cascading style sheet is used.

XML is recommended by the World Wide Web Consortium (W3C). It is a free open standard. The W3C recommendation specifies both the lexical grammar and the requirements for parsing.

Structure of XML document:

```
<?xml version="1.0">
<Client>
  <Clientid>C100</Clientid>
  <Clientname>Johnson</Clientname>
  <Company>APPLE</Company>
</Client>
<Client>
  <Clientid>C101</Clientid>
  <Clientname>McGraw</Clientname>
  <Company>HCL</Company>
</Client>
</xml>
```

CTM: XML is a Markup Language for creating documents in a structured format. Users can create their own tags along with predefined tags already defined by HTML.

8.30.4 Hyper Text Transfer Protocol (HTTP)

HTTP is used to transfer all files and other data (collectively called resources) from one computer to another on the World Wide Web. This protocol is used to transfer hypertext documents over the internet. HTTP defines how the data is formatted and transmitted over the network. When an HTTP client (a browser) sends a request to an HTTP server (web server), the server sends responses back to the client. This transfer of requests and responses is done following HTTP protocol.

The main features of an HTTP document are:

1. It is a stateless protocol; this means that several commands are executed simultaneously without knowing the command which is already executing before another command.
2. It is an object-oriented protocol that uses client server model.
3. The browser (client) sends request to the server, the server processes it and sends responses to the client.
4. It is used for displaying web pages on the screen.

8.30.5 Domain Names

To communicate over the internet, we can use IP addresses. But it is not possible to remember the IP address of a particular website or computer every time. Domain names make it easier to resolve IP addresses into names, *for example*, cbse.nic.in, google.com, meritnation.com, etc. It is the system which assigns names to some computers (web servers) and maintains a database of these names and corresponding IP addresses. Domain names are used in URLs to identify particular web servers, *for example*, in the URL <https://www.cbse.nic.in/welcome.htm>, the domain name is cbse.nic.in.

A domain name consists of the following parts.

1. Top-level domain name or primary domain name, and
2. Sub-domain name(s).

For example,

In the domain name **cbse.nic.in**:

in is the primary domain name

nic is the sub-domain of in

cbse is the sub-domain of nic.

The top-level domains are categorized into following domain names:

Generic Domain Names

- com - commercial business
- edu - Educational institutions
- gov - Government agencies
- mil - Military
- net - Network organizations
- org - Organizations (non-profit)

Country Specific Domain Names

- in - India
- au - Australia
- ca - Canada
- ch - China
- nz - New Zealand
- pk - Pakistan
- jp - Japan
- us - United States of America

8.30.6 URL

URL stands for uniform resource locator that helps in locating a particular website or a web page, *for example*, <http://www.cbse.nic.in/academics.html> is a URL for a specific website. In this URL, 'http' stands for hypertext transfer protocol, and 'www.cbse.nic.in' indicates the IP address or the domain name where the source is located. 'academics.html' specifies the name of the specified html document on the website of CBSE.

CTM: URL stands for uniform resource locator that stores the address of a web page.

8.30.7 IP Address

The computers connected to a network also need to follow some rules to communicate with each other. These sets of rules are known as protocols. Several types of protocols are used for communication over networks. However, the most common one is the Transmission Control Protocol/Internet Protocol or TCP/IP. A network using TCP/IP is known as a TCP/IP network. The internet is an example of the TCP/IP network. Therefore, it becomes important that each device should have a unique address to identify it on a TCP/IP network. This unique address is known as IP address. IP address is short for Internet Protocol (IP) address. An IP address is an identifier for

a computer or device on a TCP/IP network. Networks using the TCP/IP protocol route messages based on the IP address of the destination. The format of IP address is a 32-bit numeric address written as four numbers separated by periods. Each number can be 0 to 255. Some examples of IP addresses are: 192.168.1.2, 10.324.1.3 and 109.134.2.2.

As seen in the above examples, the numbers in an IP address are in the decimal form. When an IP address is processed, the computer converts these numbers into binary form internally. The following is an example of IP address in the decimal form: 192.168.1.10

The preceding IP address in its binary form is: 11000000.10101000.00000001.00001010

The IP address of a computer is assigned by the Internet Service Provider (ISP) whose internet connection is used on that computer.

You can find out the IP address of your computer by performing the following steps:

1. Click the **Start** button. A pop-up menu containing various options appears.
2. Open the **Run** dialog box.

Type cmd in the **Run** dialog box and press Enter. The command prompt (cmd) window opens.

3. Type **ipconfig** and press Enter. The output appears, displaying the IP address of your computer (as shown in the output window given below).

```
C:\Windows\system32\cmd.exe
Microsoft Windows XP [Version 5.1.2600]
(C) Copyright 1985-2001 Microsoft Corp.

C:\Documents and Settings\Owner>ipconfig
Windows IP Configuration

Ethernet adapter Local Area Connection:

      Connection-specific DNS Suffix  . : 
      IP Address. . . . . : 192.168.1.2
      Subnet Mask . . . . . : 255.255.255.0
      Default Gateway . . . . . : 192.168.1.1

C:\Documents and Settings\Owner>
```

8.30.8 Website

A website is a collection of various web pages, images, videos, audios or other kinds of digital assets that are hosted on one or several web servers. The first page of a website is known as home page where all the links related to other documents are displayed. The web pages of a website are written using HTML and the information is transferred over the internet through HTTP protocol. The HTML documents consist of several hyperlinks that are accessed through HTTP protocol. Examples of various websites are: cbse.nic.in, google.com, amazon.in, etc.

CTM: A website is a collection of several web pages which are related to each other through hyperlinks.

8.30.9 Web Page

A web page is an electronic document/page designed using HTML. It displays information in textual or graphical form. Traversal from one web page to another web page is possible through hyperlinks.

A web page can be classified into two types:

- **Static web page:** A web page which displays same kind of information whenever a user visits it is known as a static web page. A static web page generally has .htm or .html as extension.
- **Dynamic web page:** An interactive web page is a dynamic web page. A dynamic web page uses scripting languages to display changing content on the web page. Such a page generally has .php, .asp, or .jsp as extension.

8.30.10 Web Browser

It is a software that helps in accessing web pages and, thus, is also called web client. It helps the user to navigate through the World Wide Web and display web pages. Some popular web browsers are: Mozilla Firefox, Opera, AOL, Webkit, Iceweasel, etc.

CTM: A web browser is a WWW client that navigates through the WWW and displays web pages.

8.30.11 Web Server

A web server is a server that stores web pages and when a web client is sending any request to a server, a server responds to the request and displays the requested web pages. A web server is a program that runs on a computer connected to the internet. Web server waits for a request, finds the documents and generates information, if required, and sends back to the browser that requested for it. A single web server may support multiple websites, or a single website may be hosted on several linked or mirrored web servers.



Fig. 8.34: Sending and Receiving Request

Some popular web servers are: Apache web server, Netscape enterprise web server, Microsoft internet information server, etc.

CTM: A web server stores web documents and responds to the requests made by web browsers.

8.30.12 Web Hosting

Web hosting is a service which is provided by companies to its clients to allow them to construct their own websites which are accessible to the internet users via World Wide Web. Such companies are known as web host. These companies provide space on a web server they own for use by their clients as well as provide internet connectivity. The websites which are constructed display information for their organization in the form of web pages. The host may also provide an interface or control panel for managing the web server so as to add news and events related to their organization or for uploading some information which may be valuable for the internet users. A client can also use control panel for installing scripts as well as other modules and service applications like email. webhostingsitesindia.co.in is one of the top domain name registration and web hosting companies in India. It is the only hosting company which provides support in regional languages.

CTM: Web hosting is a service that is provided by the company to users to create web-based applications.

8.30.13 Domain Name Resolution (DNR)

You already know that communication between computers on a network takes place with the help of IP addresses. However, to access a particular website, you use its URL because it is much easier to remember than the IP address. When you use the URL to access a website, your computer needs to find its matching IP address. How does it find this address? The answer is: by using the Domain Name Resolution (DNR). DNR is the name given to the process by which your computer finds the IP addresses of domain names.

The DNR process takes place in the background, and can be explained in the following steps:

1. Enter the domain name of the website that you want to access in the Web browser.
2. The Web browser issues a command to the operating system of your computer to generate the IP address of the given domain name.

The domain name is resolved according to the configuration of the operating system you are using. Different operating systems, such as Windows, XP, Windows 7, Linux, and Unix have different configurations.

This is done in the following manner:

- Generally, the operating system maintains a HOSTS file, which contains a list of the IP addresses of some domain names. Therefore, the operating system first checks this file to find the IP address of the given domain name.
 - If the IP address is not found in the HOSTS file, the operating system connects to the DNS server on a network. The DNS server maintains a directory containing a list of all the domain names and IP addresses that are registered on the internet. The DNS server finds the IP address of the given domain name and returns it to the operating system.
3. After obtaining the IP address, the operating system sends it to the Web browser that has requested it.

CTM: A web browser is a WWW client that navigates through the WWW and displays web pages.

8.30.14 Web Scripting

Website Scripting is used for creating web pages to publish them on the web interactively. The communication between web browser and web server happens through small programs called web scripts. Script is a programming language which, when executed, displays the web page. The written codes for a script may be used by the server side or may be used by the client side as per the requirement. The tasks which are executed by a web server are interpreted and automated through the web scripts written in Web Scripting Language.

CTM: A script is a list of commands embedded in a web page which are executed by a certain program or scripting engine.

(a) Client-Side Web Scripting Languages

Client-side scripting enables the user to interact with web pages. The client-side scripts are downloaded at the client end and then interpreted and executed by the web browser. The client-side scripting is browser dependent and, therefore, the browser must be scripting-enabled that can interpret the script code. Examples of client-side scripting where it is used are online games, downloading data from the server, etc.

Some Popular Client-Side Scripting Languages are:

1. JavaScript

2. VBScript

3. PHP

1. JavaScript

JavaScript was the first web scripting language to be introduced and it is still by far the most popular. The original name of JavaScript was “LiveScript” and it was first introduced in Netscape Navigator 2.0 in 1995 and was renamed JavaScript so as to correlate with Java Language. JavaScript is primarily used for client-side processing. It is a scripting language and not a programming language. It can easily be embedded in HTML tag and executes immediately as the page is displayed. JavaScript, along with languages like XML, can call in only the required data from the server without receiving a full web page.

Example of Java Script is as follows:

```
<HTML>
<HEAD>
<TITLE>My first java script</TITLE>
</HEAD>
<BODY>
<SCRIPT LANGUAGE= "JAVASCRIPT" >
document.write("Welcome")
</SCRIPT>
</BODY>
</HTML>
```

JavaScript allows the user to perform several functions such as:

- (i) Add scrolling or changing messages to the browser's status line.
- (ii) Update the contents of a form by using validation checks and make calculations.
(For example, on entering the marks of 5 subjects of a student, it will calculate total marks and percentage.)
- (iii) Display messages to the user, either in a new web page, which is a sub-part of the main web page, or in alert boxes.
- (iv) Create and animate images that change when the user hovers the mouse over them.
- (v) Detect the browser in use and display different contents for different browsers.
- (vi) Detect installed plug-ins and notify the user if a plug-in is required.

2. VBScript (Visual Basic Script)

VBScript is a light weight programming language developed by Microsoft Corporation. VBScript is the default scripting language for ASP (Active Server Pages). VBScript is a server-side scripting language. However, it may also be used for client-side scripting (although it is currently only supported by Internet Explorer).



Some of the useful features of visual basic are not supported by VBScript such as strong typing, extended error trapping and the ability to pass a varied number of parameters to a sub-routine. However, its use is relatively widespread because it is easy to learn. VBScript can be effectively used for automating day-to-day office tasks as well as monitoring in Windows-based environment. When used for client-side web development in Microsoft Internet Explorer, VBScript is similar in function to JavaScript. It is used to write executable functions that are embedded in or included from HTML pages and interact with the Document Object Model (DOM) of the page, to perform tasks not possible in HTML alone. VBScript is simple to create and can be written using text editor like Notepad. A simple VBScript document is saved with ".vbs" extension.

Example of VBScript is as follows:

```
<HTML>  
<HEAD>  
<TITLE>My first VBScript</TITLE>  
</HEAD>  
<BODY>  
<SCRIPT TYPE="text/vbscript">  
document.write("Welcome")  
</SCRIPT>  
</BODY>  
</HTML>
```

3. PHP (Hypertext Pre-Processor)

PHP stands for Hypertext Pre-processor. It is a server-side scripting language that is used to enhance web pages. With PHP, a user can do things like create username and password login pages, check details from a form, create forums, picture galleries, etc. It was created in 1994 by Rasmus Lerdorf to add dynamic content to an HTML page. PHP initially stood for 'Personal Home Page', but now it is translated as 'PHP Hypertext Pre-processor'. The PHP code is embedded within the HTML code between special tags. When the page is accessed, the server processes the PHP code and then sends the output from the script as HTML code to the client.

The salient features of PHP are as follows:

- (i) A user can create dynamic web pages with the PHP scripting language.
- (ii) It is a server-side scripting language and, therefore, the PHP scripts are executed on the server.
- (iii) PHP is free and an open source software product.
- (iv) PHP provides connectivity with many databases (MySQL, Sybase, Oracle and many others).
- (v) PHP runs on different platforms (UNIX, Linux, Windows).
- (vi) PHP is compatible with almost all web-servers used today (Apache, IIS, etc.).
- (vii) PHP commands are embedded within a standard HTML page.

- (viii) Its syntax is similar to that of C and Perl languages which makes it easy to use.
(ix) PHP files can have one of the following extensions: PHP, PHP3 or PHTML.

The structure of PHP Script is as follows:

```
<HTML>  
<HEAD>  
<TITLE>PHP Test</TITLE>  
</HEAD>  
<BODY>  
<?php echo '<p>Hello World</p>' ?>  
</BODY>  
</HTML>
```

(b) Server-Side Script

Server-side scripting gets executed on the server before displaying the information requested.

1. ASP (Active Server Pages)

It is a technology that is used to create dynamic web pages so that the user can see these pages without any specific web browser. The user does not require any specific web browser to view these pages. To create active server pages, a default scripting language is used, viz. VBScript and sometimes JSCRIPT. ASP pages are saved with the extension .ASP and not by .html. Any web page that contains ASP code cannot be run simply by executing it through web browser; instead, the page must be requested through a web server that supports ASP. When a browser requests an ASP, the web server generates a page with HTML code and sends it back to the browser. ASP is also the short term used for Application Service Provider.

Various services offered by ASP are:

- (i) It helps in creating interactive and dynamic web pages.
- (ii) It allows access to any data or databases and returns the result to web browser.
- (iii) It allows the user to dynamically edit, change or add any content of a web page.
- (iv) It provides security to a page as it is not executed by a web browser.

2. JSP (Java Server Pages)

Java Server Pages (JSP) is a technology that helps software developers to create dynamic web pages based on HTML, XML or other document types. It was released in 1999 by Sun Microsystems. JSP is similar to PHP, but it uses the Java programming language. JSPs have dynamic scripting capability that is embedded in an HTML code. Java server pages can be run by a compatible web server with a servlet container, such as Apache Tomcat or Jetty. To process a JSP file, developers need a JSP engine, which is connected to a Web server. The JSP page is then compiled into a servlet, which is handled by the servlet engine. This phase is known as translation. The servlet engine then loads the servlet class and executes it to create dynamic HTML, which is then sent to the browser. A Java Server Pages compiler is a program that parses JSPs and converts them into executable Java Servlets.

The various features of JSP are as follows:

- (i) **Platform Independent:** JSP can be deployed across many platforms. All these components can be run across web servers.
- (ii) **Configured for reusability:** JSP components can be reused across servlets, JavaBeans and Enterprise JavaBeans (EJB).
- (iii) **Simple and easy to use:** JSP is simple in the process of development and maintenance.

CTM: Scripting languages are also often used by applications as control or configure languages. An example: Firefox is written in C/C++ and can be controlled with JavaScript.

Table: 14.1: Difference between Client-Side Scripting and Server-Side Scripting

S.No.	Client-Side Scripting	Server-Side Scripting
1.	Scripting runs through web browser.	Scripting runs through web server.
2.	The processing takes place on the end-user's computer.	The processing takes place on the server side.
3.	The browser receives the page sent by the server and executes the client-side scripts.	Server executes server-side scripts to send out a page but it does not execute client-side scripts.
4.	Client-side scripting cannot be used to connect to the databases on the web server.	Server-side scripting is used to connect to the databases that reside on the web server.
5.	Client-side scripting can be blocked by the user.	Server-side scripting cannot be blocked by the user.
6.	Response from a client-side script is faster as compared to a server-side script because the scripts are processed on the local computer.	Response from a server-side script is slower as compared to a client-side script because the scripts are processed on the remote computer.
7.	Examples of client-side scripting languages are Javascript, VBScript, etc.	Examples of server-side scripting languages are PHP, JSP, ASP, ASP.Net, Ruby, Perl, etc.

8.31 WEB 2.0

Web 2.0 is often known as World Wide Web for second generation where people collaborate with each other through social networking sites, blogs, wikis, folksonomies, video-sharing sites, hosted services, web applications and mashups. Web 2.0 offers more dynamic pages instead of static pages that allows users to create online applications. Web 2.0 also allows groups of people to work on multiple applications. *For example*, a user can work on a document or spreadsheet simultaneously, while in the background a computer keeps track of who made what changes, where and when.

Web 2.0 tools are also less expensive than traditional software—many are even free. Because they are web-based, the need is only to update the browser.

The various features of Web 2.0 are:

- Web-based applications can be accessed from anywhere.
- It provides multitasking.
- It is less expensive and provides dynamic pages.
- Specific problems can be solved by using simple applications.
- Value lies in content, not in the software used to display content.
- Transfer of data is readily shared.
- Employees and customers can create their own newsgroup.
- Social tools encourage people to create, collaborate, edit, categorize, exchange and promote information.

Web 2.0 website uses a new programming language called AJAX (Asynchronous JavaScript and XML).

Applications supported by Web 2.0 are as follows:

- Blogging
- Social bookmarking
- RSS
- Wikis and other collaborative applications
- Interactive encyclopaedias and dictionaries
- Advanced Gaming

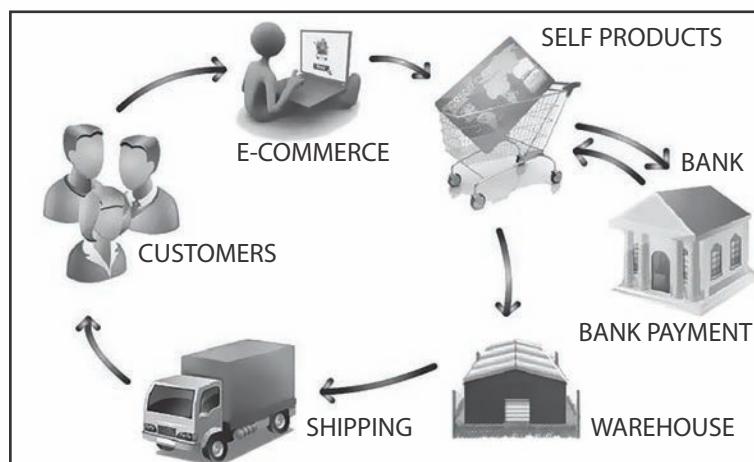
CTM: Web 2.0 refers to added features and applications that make the web more interactive and easy to provide information using newsgroups, social networking sites, RSS, etc.

8.32 E-COMMERCE

E-commerce (electronic commerce or EC) is the buying and selling of goods and services, or the transmitting of funds or data over an electronic network, primarily the internet. These business transactions occur either as business-to-business, business-to-consumer, consumer-to-consumer or consumer-to-business. The terms e-commerce and e-business are often used interchangeably.



E-commerce is conducted using a variety of applications, such as email, online catalogues and shopping carts, EDI, File Transfer Protocol and web services. This includes business-to-business activities and outreach such as using email for unsolicited ads (usually viewed as spam) to consumers and other business prospects, as well as to send out e-newsletters to subscribers. More companies now try to entice consumers directly online, using tools such as digital coupons, social media marketing and targeted advertisements.



The advantages of e-commerce include its round-the-clock availability, speed of access, wide availability of goods and services to the consumer, easy accessibility and international reach.

8.32.1 Payment Transactions through E-Commerce

1. Online Banking

Online banking, also known as internet banking, e-banking or virtual banking, is an electronic payment system that enables customers of a bank or a financial institution to conduct a range of financial transactions through the financial institution's website.



2. Mobile Banking

Mobile banking is a service provided by a bank or a financial institution that allows its customers to conduct a range of financial transactions remotely using a mobile device such as a **mobile phone** or **tablet**, and using software, usually called an app, provided by the financial institution for the purpose.



8.32.2 Payment Apps and Services

Payment apps are also referred to as Mobile Wallets. There are many payment apps available in the market which are very efficient to use and convenient to handle.

The following are some of the wallets which are frequently used in the market.

1. Google Wallet
2. Apple Passbook
3. Paytm Wallet
4. Freecharge Wallet
5. MobiKwik Wallet
6. PayU
7. Square Wallet
8. PayPal
9. Dwolla
10. Venmo
11. Bharat Interface for Money (BHIM) App



The online mode of payment has brought about a drastic change to marketing. There are many start-ups which are doing very well thanks to the online mode of payment. This revolution has taken the world by storm.

8.33 CLOUD COMPUTING

Cloud computing provides computing and storage capacity services to the heterogeneous community of end-recipients. The name comes from the use of clouds as an abstraction for the complex infrastructure.

Cloud Computing = Software as a Service + Platform as a Service + Infrastructure as a Service
+ Data as a Service

Therefore, cloud computing:

- Provides a shared pool of configurable computing resources
- Provides on-demand network access
- Is provisioned by the Service Provider

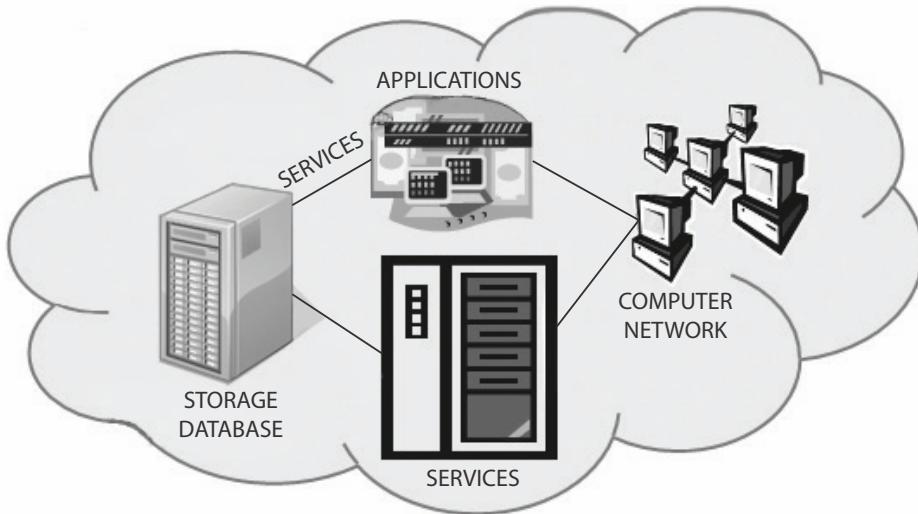


Fig. 8.35: Components of Cloud Computing

It entrusts services with a user's data, software and computation over a network. It has considerable overlap with Software as a Service (SaaS).

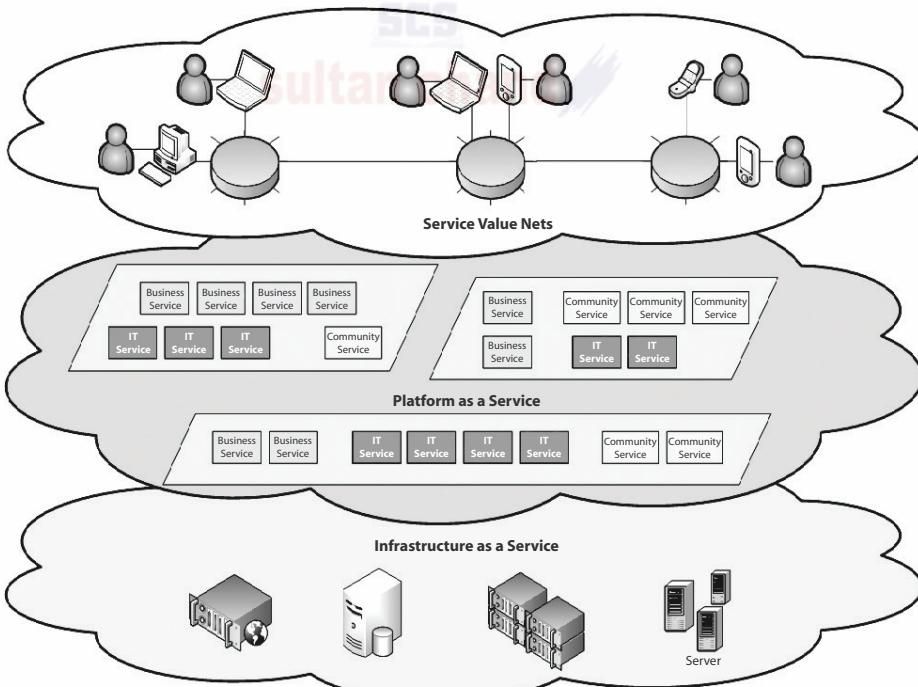


Fig. 8.36: Cloud Architecture

➤ **Software as a Service (SaaS)**

- From the end-user's point of view
- Apps are located in the cloud
- Software experiences are delivered through the internet

- Platform as a Service (PaaS)
 - From the developer's point of view (*i.e.*, cloud users)
 - Cloud providers offer an internet-based platform to developers who want to create services but do not want to build their own cloud.

➤ Infrastructure as a Service (IaaS)

- Cloud providers build datacentres
- Power, scale, hardware, networking, storage, distributed systems, etc.
 - Datacentre as a service
- Cloud users rent storage, computation and maintenance from cloud providers (pay-as-you-go-like utility)

➤ Data as a Service (DaaS)

Data → Information → Knowledge → Intelligence

- Infrastructure for web-scale data mining and knowledge discovery
- Empower people with knowledge
- Empower applications and services with intelligence

Benefits of cloud computing: There are several benefits of cloud computing because of which it has become essential and popular these days.

➤ Reduces capital and operational costs

- No longer required to make large upfront capital investment on datacentres
- Eliminates the need to plan ahead for provisioning
- Allows companies to start small and increase their resource investment as needed (pay-as-you-go)

➤ Simplifies app deployment & management

- Common programming model across mobile, browser, client, server, cloud
- Access to strong ecosystem of widely deployed applications
- Integration with existing IT assets (Software + Services)

8.34 SETTING UP A COMPUTER NETWORK—AN EXAMPLE

The network functioning is based on Client-Server architecture which requires effective and efficient network design. It defines how clients are connected to server machine(s) on a network. The most important rule or methodology for network (LAN) design is the 80:20 rule.

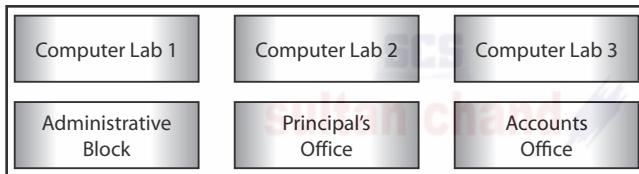
The 80:20 Thumb Rule

This thumb rule states that in a well-organized and designed network, 80 percent of the traffic on a given network segment is local (*i.e.*, destined for a destination system in the same workgroup), and not more than 20 per cent of the network traffic should move across a backbone. The backbone in a network which violates this 80:20 rule leads to network congestion and traffic jams.

Keeping the above significant rule in mind, let us take an example to understand how to go about LAN design.

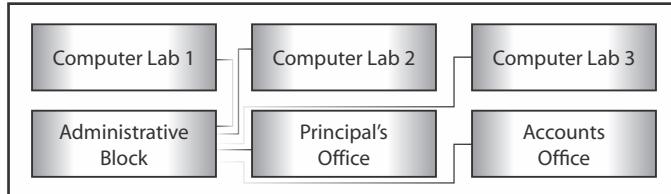
An educational society (say XYZ Educational Society), with its head office in Chennai (Tamil Nadu) and schools in various parts of the globe, is setting up a new senior secondary school, 'SF School', in Bahadurgarh (Haryana).

The 'SF School' will have 3 computer labs with 30 computers in each lab, one Accounts office with three computers, one Administrative block with five computers, and a Principal's office with one computer.



Let us see how a computer network can be set up in the school. First of all, we can draw a rough sketch of the school with computers at various locations as follows:

1. Independent LANs can be set up in each of the following buildings: Computer Lab1, Computer Lab2, Computer Lab3, Administrative Block, and Accounts Office.
2. These LANs can be set up in STAR topology using UTP cable (economical, reliable and easily available).
3. For this, one switch (with suitable number of ports) will be required in each of these buildings. More than one switch can be used in computer labs if a switch with more than 30 ports is not available.
4. Two internet connections (broadband for high speed) can be procured in Administrative Office. Two connections should be procured from two different ISPs so that:
 - (a) Internet connection in Administrative office can be kept separate from the other Computer labs where students may do a lot of experimentation.
 - (b) If one internet connection is not working, the other can be used in case of urgent requirements.
5. These buildings can then be connected as follows:



6. This interconnection will ensure that each building is directly connected to Admin block.
7. This way, internet connection will be available in each building irrespective of the status of the other building.
8. Server (if any) of the school may be placed in the Admin block so that it remains safe (physically) and a firewall can be set up so that the whole network remains safe from any kind of virus or intrusion attacks.

There is no need to put in any extra efforts or expenses to link the school to its head office. This can be taken care of using the internet connections

Tips to solve technical questions based on Networking

Where Server should be placed: Server should be placed in the building where the number of computers is maximum.

1. **Suggest a suitable cable layout of connection:** A suitable cable layout can be suggested in the following two ways:



- (a) **On the basis of Server:** First, the location of the Server is found out. Server should be placed in that building where the number of computers is maximum (according to the 80:20 rule). After finding the server position, each building distance is compared with the Server building directly or indirectly (taking other building(s) in between). The shortest distance is counted, whether it is directly or indirectly calculated.
- (b) **On the basis of distance from each building:** The distance between each building is compared to all other buildings, either directly or indirectly. The shortest distance is calculated, whether it is direct or through some other building.

2. Where the following devices should be placed:

<i>Server</i>	: Large number of computers in the building
<i>HUB/Switch</i>	: Each building
<i>Modem</i>	: In the server room
<i>Repeater</i>	: It is used if the distances are higher than 70m. It regenerates data and voice signals.
<i>Router</i>	: When one LAN is required to be connected to the other LAN
<i>Best Layout</i>	: Star (from Server), BUS topology
<i>Best Cable</i>	: Twisted Pair, Ethernet Cable, coaxial cable (when distance is in metres); For large distances—Fibre optics cable.

Best connecting technique : In hilly regions, radio waves should be used and city-to-city, state-to-state satellite should be used.



MEMORY BYTES

- Internet is a network of networks that spreads all over the world.
- ARPANET was the first internet followed by NSFNET and other small networks.
- A gateway is a device that connects dissimilar networks.
- A backbone is a central interconnecting device that connects two or more computers.
- Topology is a way of connecting computers physically or logically.
- Star topology uses a central hub where each computer indirectly connects with the other computer on the network.
- A set of rules that governs internet is called protocol.
- TCP protocol is responsible for sequential arrangement of packets.
- IP protocol is responsible for fragmentation of data into packets and sends those packets in random order.
- FTP protocol is used to share files across networks.
- TELNET is a remote login where a user can login on another user's system.
- HTTP is used for displaying web pages.
- Web browser is an application program that helps in opening web pages.
- The first page of any website is known as home page.
- Communication media is a transmission media for transmitting data across the network.
- Guided media is also known as wired media while unguided media is also known as wireless media.
- LAN, MAN, WAN and PAN are the four types of networks.
- Viruses are malicious programs that can damage files, disks, file allocation table, etc.
- Spams are unsolicited mails that eat up the disk space.

- Hub refers to a networking component which acts as a convergence point of a network allowing the transfer of data packets.
- Switch refers to a device which filters and forwards data packets across the network.
- Web hosting service is a type of internet hosting service that allows individuals and organizations to host their own website and users with online systems to store information such as images, videos, etc.
- A data channel is the medium used to carry information or data from one point to another.

OBJECTIVE TYPE QUESTIONS

1. Fill in the blanks.

- (a) Through you can establish contact with anyone in the world.
- (b) The main function of is to divide the message or data into packets of a definite size on the source computer.
- (c) refers to wireless fidelity which enables us to connect to the ISP without any cable.
- (d) is a software that enables us to access the internet and explore websites.
- (e) Web page constitutes the
- (f) A is someone with a strong interest in how things work, who likes to create and modify things for their own enjoyment.
- (g) A computer is a small software program that spreads from one computer to another and interferes with the normal functioning of computer.
- (h) Electronic junk mail or junk newsgroup postings are known as
- (i) Digital signature meets the need for and integrity.
- (j) The first network that planted the seed of internet was
- (k) The protocol used for internet is
- (l) A device used to connect dissimilar networks is called
- (m) is responsible for handling the address of the destination computer so that each packet is delivered to its proper destination.
- (n) Tricking people through authentic-looking emails or websites is called
- (o) A program designed to replicate and eat up a computer's storage is called
- (p) A digital document issued to a site by a certification authority of the internet is called a
- (q) To connect computers located in a specific building or campus is known as
- (r) Wi-Fi, infrared and Bluetooth are examples of
- (s) Interspace is a
- (t) A server that provides its services to other workstations on a network is a
- (u) The techniques of switching in which data is fragmented into smaller techniques is called
- (v) is a dedicated line between the caller and the sender.
- (w) is the measuring unit of speed at which the data transfer takes place.
- (x) All the computers are connected with each other in an unorganized manner in topology
- (y) In, all computers share equivalent responsibility for processing data..

- Answers:**
- | | | |
|---------------------------|--------------------------|----------------------|
| (a) Internet | (b) TCP | (c) Wi-Fi |
| (d) Browser | (e) World Wide Web | (f) Hacker |
| (g) Virus | (h) SPAM | (i) Authentication |
| (j) ARPANET | (k) TCP/IP | (l) Gateway |
| (m) IP | (n) Hacking | (o) WORM |
| (p) Digital certificate | (q) LAN | |
| (r) Communication Mediums | (s) Network | (t) Dedicated server |
| (u) Packet switching | (v) Circuit switching | (w) Bits/Second |
| (x) Mesh | (y) Peer-to-peer network | |



2. State whether the following statements are True or False.

- (a) A set of rules that governs internet is called protocol.
 - (b) A repeater handles different protocols.
 - (c) A hub is known as an intelligent device on the network.
 - (d) A location on a net server is called a website.
 - (e) A document that uses HTTP is called a web page.
 - (f) A switch is a device used to segment networks into sub-networks or subnets.
 - (g) Email is sending and receiving messages through videoconferencing.
 - (h) The degeneration of a signal over a distance on a network is called attenuation.
 - (i) Coaxial cable possesses higher tensile strength than optical fibre.
 - (j) When two entities are communicating and do not want a third party to listen, this situation is defined as secure communication.

Answers: (a) True (b) False (c) False (d) True (e) True (f) True
(g) False (h) True (i) False (j) True

3. Multiple Choice Questions (MCQs)

- (a) A computer network:

 - (i) Is a collection of hardware components and computers
 - (ii) Is interconnected by communication channels
 - (iii) Allows sharing of resources and information
 - (iv) All of the above

(b) What is a firewall in computer network?

 - (i) The physical boundary of network
 - (ii) An operating system of computer network
 - (iii) A system designed to prevent unauthorized access
 - (iv) A web browsing software

(c) What is the use of Bridge in the network?

 - (i) To connect LANs
 - (ii) To separate LANs
 - (iii) To control network speed
 - (iv) All of the above

(d) Each IP packet must contain:

 - (i) Only Source address
 - (ii) Only Destination address
 - (iii) Source and Destination address
 - (iv) Source or Destination address

(e) Which of these is not a communication channel?

 - (i) Satellite
 - (ii) Microwave
 - (iii) Radio wave
 - (iv) Wi-Fi

(f) MAN Stands for

 - (i) Metropolitan Area Network
 - (ii) Main Area Network
 - (iii) Metropolitan Access Network
 - (iv) Metro Access Network

(g) Which of these is not an example of unguided media?

 - (i) Optical Fibre Cable
 - (ii) Radio wave
 - (iii) Bluetooth
 - (iv) Satellite

(h) In which topology are all the nodes connected through a single Coaxial cable?

 - (i) Star
 - (ii) Tree
 - (iii) Bus
 - (iv) Ring

(i) Which of the following is the smallest network?

 - (i) WAN
 - (ii) MAN
 - (iii) LAN
 - (iv) Wi-Fi

(j) Which protocol is used for the transfer of hypertext content over the web?

 - (i) HTML
 - (ii) HTTP
 - (iii) TCP/IP
 - (iv) FTP

Answers: (a) (iv) (b) (iii) (c) (i) (d) (iii) (e) (iv)
(g) (i) (h) (iii) (i) (iii) (j) (ii)

Answers: (a) (iv) (b) (iii) (c) (i) (d) (iii) (e) (iv) (f) (i)
(g) (i) (h) (iii) (i) (iii) (j) (ii)

SOLVED QUESTIONS

1. Define a network. What is its need?

Ans. A network is an interconnected collection of autonomous computers that can share and exchange information.

Need for networking:

- (a) *Resource sharing:* Resources are shared by all computers over the network for effective utilization.
- (b) *File sharing:* A file in a network can be accessed from anywhere.

2. Write **two** advantages and **two** disadvantages of network.

Ans. *Advantages of network:*

- (a) We can share resources such as printers and scanners.
- (b) We can share data and access files from any computer.

Disadvantages of network:

- (a) If there is any problem in the server, then no communication can take place.
- (b) Network faults can cause loss of data.
- (c) If there is no privacy mechanism used then the entire network can be accessed by an unauthorized person.

3. What is ARPANET? What is NSFNET?

Ans. ARPANET (Advanced Research Project Agency Network) is a project sponsored by US Department of Defence.

NSFNET was developed by the National Science Foundation which was a high capacity network and strictly used for academic and engineering research.

4. What are the various types of networks?

Ans. A network is an interconnection of several nodes through some communication media with the goal of sharing data, files and resources. There are three types of networks:

- (a) Local Area Network (LAN)
- (b) Metropolitan Area Network (MAN)
- (c) Wide Area Networks (WAN)

5. Name the various layers of coaxial cable.

Ans. Coaxial cable consists of the following layers:

- (a) A metallic rod-shaped inner conductor.
- (b) An insulator covering the rod.
- (c) A metallic outer conductor called shield.
- (d) An insulator covering the shield.
- (e) A plastic cover.

6. What is a spam mail?

[CBSE D 2015]

Ans. Spam is the abuse of electronic messaging systems (including most broadcast media, digital delivery systems) to send unsolicited bulk messages indiscriminately.

7. Differentiate between FTP and HTTP.

[CBSE D 2015]

Ans. FTP is a protocol to transfer files over the internet. HTTP is a protocol which allows the use of HTML to browse web pages in the World Wide Web.

8. Out of the following, which is the fastest (a) Wired, and (b) Wireless medium of communication? Infrared, Coaxial Cable, Ethernet Cable, Microwave, Optical Fibre.

[CBSE D 2015]

Ans. (a) *Wired:* Optical Fibre

(b) *Wireless:* Infrared or Microwave

9. What is a worm? How is it removed?

[CBSE D 2015]

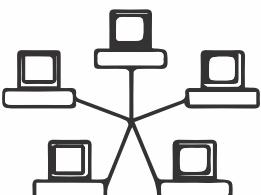
Ans. A worm is a self-replicating computer program. It uses a network to send copies of itself to other computers on the network and it may do so without any user intervention. Most of the common antivirus (anti-worm) software remove worm.

10. Illustrate the layout for connecting five computers in a Bus and a Star topology of Networks.

Ans. Bus topology



Star topology



11. In networking, what is WAN? How is it different from LAN?

[CBSE D 2011]

Ans. Internet is an example of WAN (Wide Area Network). Most WANs exist to connect LANs that are not in the same geographical area. WAN is different from LAN due to its network range. WAN is for connecting computers anywhere in the world without any geographical limitation whereas LAN is confined within a range of 100m to 500m.

12. What is meant by topology? Name some popular topologies.

Ans. Topology is the arrangement by which computers are connected with each other, either physically or logically.

The popular topologies are:

- (a) Bus or Linear Topology
- (b) Ring Topology
- (c) Star Topology
- (d) Tree Topology

13. Why are fibre optic cables becoming popular?

Ans. Fibre optic transmission is becoming increasingly popular due to its noise resistance, low attenuation and high bandwidth capabilities.

14. What factors should be taken into consideration while opting for a particular topology?

Ans. There are a number of factors which are to be considered:

- (a) Cost
- (b) Flexibility
- (c) Reliability

15. What is a modem? What are the two types of modems?

Ans. Modem stands for modulator demodulator that converts analog signals to digital signals at the sender's end. It converts digital signals back to analog signals at the receiver's end.

The two types of modems are: internal modem and external modem.

16. What is remote login? What is Telnet?

Ans. Remote login is the process of accessing a network from a remote place without actually being present at the place of working. Telnet is an internet utility that lets us log on to a remote computer system. A user is able to log in the system for sharing of files without being the actual user of that system.

17. Briefly explain FTP.

Ans. FTP stands for File Transfer Protocol. It is the standard mechanism provided by TCP/IP for copying a file from one host to another. While sharing files from one system to another, we may encounter several problems—two systems may have different directory structures, two systems may have different file-naming conventions, or two systems may have different ways to represent text and data. All these problems are solved by FTP.



18. What is protocol? Name some commonly used protocols.

Ans. A protocol means the rules that are applicable for a network, or we can say the common set of rules used for communication in network.

Different types of protocols are:

- (a) **HTTP:** Hyper Text Transfer Protocol
- (b) **FTP:** File Transfer Protocol
- (c) **SLIP:** Serial Line Internet Protocol
- (d) **PPP:** Point-to-Point Protocol
- (e) **TCP/IP:** Transmission Control Protocol/ Internet Protocol
- (f) **SMTP:** Simple Mail Transfer Protocol
- (g) **POP:** Post Office Protocol
- (h) **IMAP:** Internet Mail Access Protocol

19. What is TCP/IP?

Ans. TCP/IP (Transmission Control Protocol/Internet Protocol) is a protocol for communication between computers used as a standard for transmitting data over networks and is the basis for standard Internet protocols. It is also responsible for assembling packets at the receiver's side.

20. How is FTP different from HTTP?

Ans. FTP is a protocol used to upload files from a workstation to an FTP server or download files from an FTP server to a workstation, whereas HTTP is a protocol used to transfer files from a Web server on to a browser in order to view a web page that is on the internet.

21. Define Mobile Communication and Wireless Communication.

Ans. Mobile Communication essentially refers to a computing device that is not continuously connected to the base or central network. This may include laptops, newly-created smartphones and also PDAs. Wireless Communication is simply data communication without the use of a landline. This may involve a cellular telephone, a two-way radio, a fixed wireless connection, a laser or satellite communication.

22. What is Mobile Processor?

Ans. Mobile Processors are required to run an operating system, be it desktop, laptop or mobile. They provide necessary resources to start an operating system, run applications and do tasks.

23. Name any four popular mobile processors in the market.

Ans. Four popular mobile processors are Qualcomm Snapdragon, Apple's Mobile Processors, HiSilicon and Samsung Exynos.

24. What are the advantages of e-commerce applications?

Ans. The advantages of e-commerce applications include their round-the-clock availability, the speed of access, the wide availability of goods and services for the consumer, easy accessibility, and international reach.

25. Define web browser and web server.

Ans. **Web Browser:** A web browser is a software which is used for displaying the content on web page(s). It is used by the client to view websites. Examples of web browser—Google Chrome, Firefox, Internet Explorer, Safari, Opera, etc.

Web Server: A web server is a software which entertains the request(s) made by a web browser. A web server has different ports to handle different requests from web browser, like generally FTP request is handled at Port 21 and HTTP request is handled at Port 80. Example of web server is Apache.

26. Differentiate between XML and HTML.

Ans. In HTML (Hyper Text Markup Language), both tag semantics and the tag set are fixed, whereas XML (Extensible Markup Language) is a meta-language for describing markup languages. XML provides the facility to define tags and the structural relationships between them. All the semantics of an XML document will either be defined by the applications that process them or by style sheets.

27. What is web hosting?

Ans. Web hosting is a means of hosting web server applications on a computer system through which electronic content on the internet is readily available to any web-browser client.

28. What is hacking?

Ans. Hacking is a process of accessing a computer system or network without knowing the access authorization credential of that system. Hacking can be illegal or ethical depending on the intention of the hacker.

29. What are cookies?

Ans. Cookies are messages that a web server transmits to a web browser so that the web server can keep track of the user's activity on a specific website. Cookies are saved in the form of text files in the client computer.

30. Differentiate between cracking and hacking.

Ans. Cracking is defined as an attempt to remove the copy protections inserted into software programs. A program successfully stripped of protections is then known as having been "Cracked". Hacking can be ethical/legal but cracking is a totally illegal method and is also called piracy.

31. What is web scripting?

Ans. A script is a small bit of code that enables web browsers to do something rather than just displaying static results. Scripts are used in web design to create dynamic pages. There are two categories of web scripts: Client-Side Script which can be written by using JavaScript, VBScript, and Server-Side Script, which can be written in PHP (used for client-side scripting also) and JSP.

32. Name some web scripting languages.

Ans. There are many scripting languages available today. Most common are VBScript, JavaScript, ASP, PHP, PERL and JSP.

33. What is Cyber Crime?

Ans. When any crime is committed over the internet, it is referred to as Cyber Crime.

34. What is Web 2.0?

Ans. Web 2.0 is a concept that takes the network as a platform for information sharing, interoperability, user-centred design, and collaboration on the internet or World Wide Web. A Web 2.0 site allows users to interact and collaborate with each other. Examples of Web 2.0 include social networking sites, facebook, google+, twitter, etc.

35. Give one advantage of bus topology of network. Also state how four computers can be connected with each other using star topology of network.

Ans. In bus topology, the workstations can easily be extended or removed. In star topology, four computers can be connected with each other through a server.

36. Write two advantages of using an optical fibre cable over an Ethernet cable to connect two service stations which are 200m away from each other. [CBSE D 2013]

Ans. Optical fibre cable guarantees secure transmission and a very high transmission capacity. Optical fibre cable is immune to electrical and magnetic interference.

37. Write two characteristics of Wi-Fi.

[CBSE D 2014]

Ans. (a) It allows an electronic device to exchange data or connect to the internet wirelessly using microwaves.
(b) Network range of Wi-Fi is much less than other network technologies like wired LAN.

38. What is the difference between Email and Chat?

[CBSE D 2014]

Ans. (a) Chat is a type of software while Email is a protocol.
(b) Chat requires the permission of both parties while Email does not.
(c) Chat is typically software dependent while Email is not.
(d) Chat needs accounts on the same provider while Email does not.

39. What are VoIP?

Ans. VoIP are communication protocols and transmission technologies for delivery of voice communication and multimedia sessions over Internet Protocol (IP) networks, such as the internet. Also, we can say that VoIP are IP technology, internet telephony and broadband telephony.



40. Expand the following terms:

- | | |
|---------|---------|
| (a) XML | (b) GSM |
| (c) SMS | (d) MAN |

Ans. (a) XML– Extensible Markup Language
 (b) GSM– Global System for Mobile communication
 (c) SMS– Short Messaging Service
 (d) MAN– Metropolitan Area Network

41. How many switching techniques are there? Explain any one.

Ans. There are three switching techniques:

- (a) Circuit Switching
- (b) Packet Switching
- (c) Message Switching

Circuit Switching: In this technique, first the complete physical connection between two computers is established and then data is transmitted from the source computer to the destination computer. The entire dedicated line is used by the caller and the receiver and no other user can use it even if the line becomes idle. When the data transmission is over, the line is disconnected and is available for the next communication.

42. How are Trojan horses different from Worms? Mention any one difference. [Sample Paper]

Ans. A Trojan horse is a term used to describe malware that appears to the user to perform a desirable function but which, in fact, facilitates unauthorized access to the user's computer system.

A computer Worm is a self-replicating program. It uses a network to send copies of itself to other nodes and that too without human intervention.

43. What is a communication channel? Name the basic types of communication channels available.

Ans. A communication channel is also known as communication media or transmission media. Communication media can be wireless or wired. Wireless media is also known as unguided media while wired media is also known as guided media.

Following are three basic types of communication channels available:

- (a) Twisted Pair Cables
- (b) Coaxial Cables
- (c) Fibre-optic Cables

44. Define baud, bps and Bps. How are these interlinked?

Ans. **Baud** is a unit of measurement for the information-carrying capacity of a communication channel.

bps (bits per second) refers to a thousand bits transmitted per second.

Bps (Bytes per second) refers to a thousand bytes transmitted per second.

All these terms are measurement units used to refer to the amount of information travelling through a single channel at any one point of time.

45. Differentiate between star topology and bus topology.

Ans. The main points of difference between star and bus topology are:

Star topology	Bus topology
A central hub is required to connect all computers with each other.	A long cable known as backbone is used to connect all computers with each other.
The data is transmitted from the sender to the receiver by passing through the hub.	The data is transmitted through a long cable from the sender to the receiver.
No collision takes place through transmission of data.	Collision can take place as the data can be transmitted from both ends at the same time.
If the central hub fails, the entire network shuts down.	If there is a break in a cable, no transmission takes place.



46. Define the following terms:

- | | |
|-------------------|--------------|
| (a) RJ-45 | (b) Ethernet |
| (c) Ethernet card | (d) Hub |
| (e) Switch | |

Ans. (a) **RJ-45:** RJ-45 is a standard type of connector for network cables and networks. It is an 8-pin connector usually used with Ethernet cables.

- (b) **Ethernet:** Ethernet is a LAN architecture developed by Xerox Corp. along with DEC and Intel. It uses a bus or star topology and supports data transfer rates of up to 10 mbps.
- (c) **Ethernet card:** The computer parts of Ethernet are connected through a special card called Ethernet card. It contains connections for either coaxial or twisted pair cables
- (d) **Hub:** In computer networking, a hub is a small, simple, low-cost device that joins multiple computers together.
- (e) **Switch:** A switch is a small hardware device that joins multiple computers together within one local area network (LAN).

47. Define the following data communicating devices:

- | | |
|--------------|-------------|
| (a) Repeater | (b) Bridge |
| (c) Router | (d) Gateway |

Ans. (a) **Repeater:** It is a device that amplifies and restores the signal before it gets degraded and transmits the original signal back to the destination. A repeater is a regenerator and not an amplifier.

(b) **Bridge:** A bridge is a device designed to connect two LAN segments. The purpose of a bridge is to filter traffic on a LAN. Bridge relays frames between two originally separate segments. When a frame enters a bridge, the bridge not only regenerates the signal but also checks the physical address of the destination and forwards the new copy only to that port.

(c) **Router:** Routers operate in the physical, data link and network layers of the OSI model. They decide the path a packet should take. A router is a networking device whose software and hardware are usually tailored to the tasks of routing and forwarding data packets across network.

(d) **Gateway:** A gateway operates on all the seven layers of OSI model. A network gateway is a computer which has internet-working capability of joining together two networks that use different base protocols. Gateway converts one protocol to another and can, therefore, connect two dissimilar networks.

48. What is HTML? Where is it used?

Ans. HTML (Hyper Text Markup Language) is used to create Hypertext documents (web pages) for websites.

HTML is the static mark-up language which is used for the following purposes:

- It is used to create web pages.
- It tells the browser how to display text, pictures and other support media.
- It supports multimedia and new page layout features.
- It provides many tags for controlling the presentation of information on the web pages, such as: <BODY>, , <HR>, etc.

49. Define GSM, CDMA and WLL.

Ans. **GSM:** Global system for mobile communication (GSM) is a wide area wireless communications system that uses digital radio transmission to provide voice data and multimedia communication services. A GSM system coordinates the communication between mobile telephones, base stations and switching systems.

CDMA: Code Division Multiple Access (CDMA) is a digital wireless telephony transmission technique which allows multiple frequencies to be used simultaneously—Spread Spectrum.

WLL: Wireless in Local Loop (WLL) is a system that connects the subscriber to the public switched telephone network (PSTN) using radio signals as alternative to other connecting media.

UNSOLVED QUESTIONS

1. What is internet?
2. What is network?
3. What are the various types of topologies?
4. Describe bus topology and star topology.
5. Define the following terms:

(a) Baud	(b) Communication channel
(c) Hubs	(d) Repeaters
6. Define GSM and GPRS wireless communication system.
7. What is modem? Define the functioning of internal modem and external modem.
8. Expand and explain the following terms:

(a) PPP	(b) POP3
(c) VoIP	(d) IRC
9. What is the significance of cyber law?
10. Describe the following networking devices:

(a) Hubs	(b) Repeaters	(c) Routers
(d) Bridges	(e) Gateways	
11. What are Wi-Fi cards? Explain.
12. What is the significance of using firewalls and authentication?
13. What is a communication protocol? What is its role in networking?
14. What is https? How does it work?
15. What is Ethernet? What is Ethernet Card?
16. What are hubs? How are active hubs different from passive hubs?
17. What are the facilities provided by the Server in a network environment?
18. Which communication medium is to be suggested for very effective and fast communication in guided medium?
19. In a harsh industrial environment, which cable would you like to use?
20. Which media have the ability to communicate over oceans?
21. What is the difference between microwave and radio wave transmission?
22. Which transmission medium is useful for sparsely populated areas?
23. Which network is easy to expand?
24. Which device filters the data and which device can handle different protocol?
25. What is a network? What are its goals and applications?
26. Write some advantages and disadvantages of the following:

(a) Optical fibres	(b) Coaxial cables	(c) Twisted pair cables
(d) Radio waves	(e) Microwaves	(f) Satellites
27. Explain the role of HTTP protocol.
28. Define email.
29. What do you understand by data transfer rates?
30. What are hubs? What are its types?
31. What is the role of a switch in a network?
32. Briefly discuss the role of the following devices in the context of networking:

(a) Repeater	(b) Router
(c) Bridge	(d) Gateway
33. Briefly

(a) HTTP	(b) TCP/IP	(c) FTP
----------	------------	---------
34. What is computing? How is it different from mobile computing?
35. When would you prefer (i) hubs over repeaters, (ii) bridges over hubs, and (iii) switches over other networking devices?

12

Relational Database and SQL

12.1 INTRODUCTION

When we speak about an organization, large amount of data is required to be processed and handled. This data handling is performed by arranging data in the form of tables and databases.

A **database** is defined as an organized collection of data (information) about an entity (something that exists) or things. It is a shared collection of related data/information used to support the activities and decision-making of a particular organization. It also allows the users to enter, access and analyze their data quickly and easily. It serves as a container which may contain various database objects. Database is **integrated** as well as **shared**. *For example*, all files belonging to one organization will be treated as the database of that organization. A database, therefore, is considered as a repository of stored data. We will now discuss some components like files, tables, records, fields, etc., that are an important part of a database.

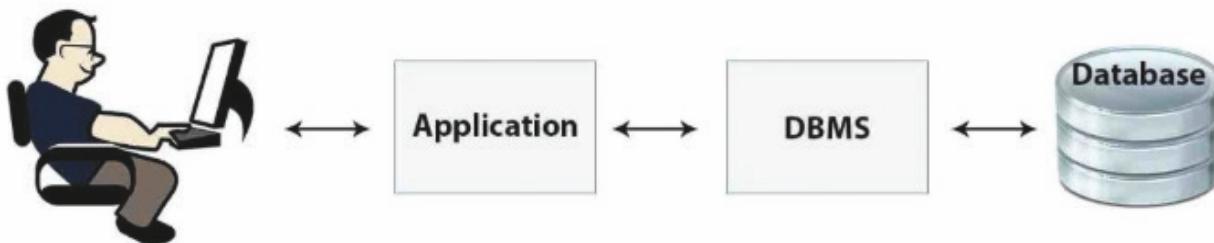
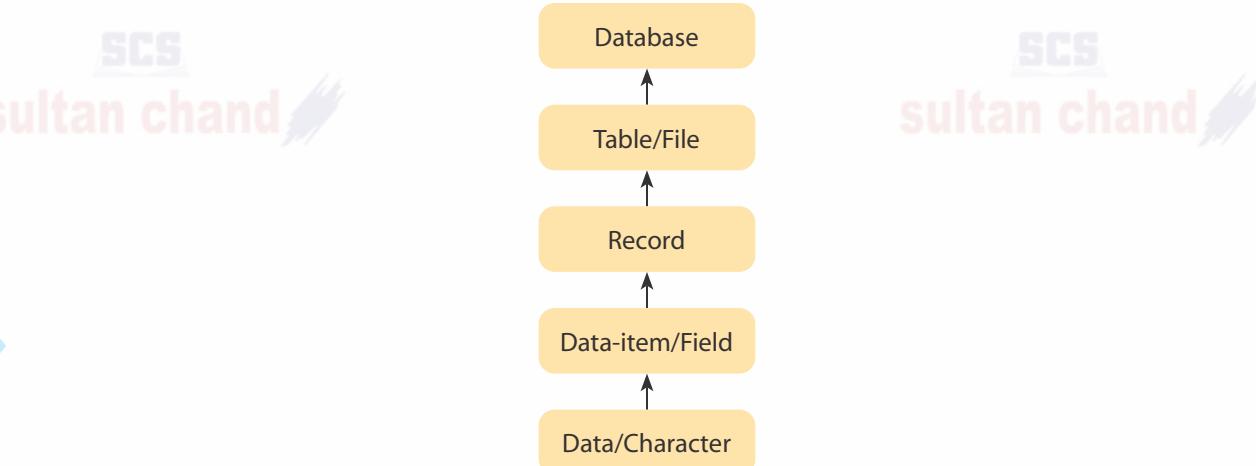


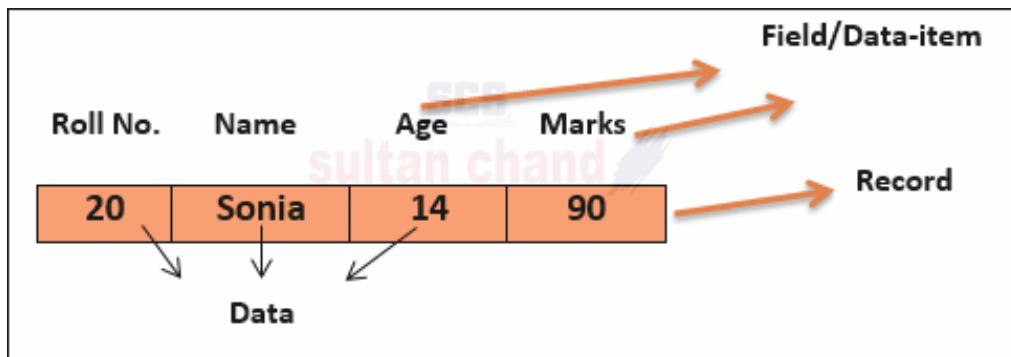
Fig. 12.1: Database and Database Management System

CTM: Database is an organized collection of interrelated data that serves many applications.

Consider the example of a “School” database. This database shall constitute tables related to student, teacher, result, etc. The data is arranged inside a database as per the file organization hierarchy as shown in Fig. 12.2.

**Fig. 12.2:** File Organization

- A **field** is a set of characters which are used together to represent specific data elements. It is also termed as a data item. A specific or an individual data item within a record is known as a field.
For example, roll number, name, age and marks are the fields in a student's record.
- A collection of fields is termed as a **Record**. *For example*, a student record consists of the fields Roll No, Name, Age and Marks as shown in Fig. 12.3.

**Fig. 12.3:** Student Table

- A collection of logically related records is called a file. A **file is also termed as a table or a relation**. A table has rows and columns, where rows represent records or tuples and columns represent the attributes or fields. *For example*, the entire information about all the students (in the form of records) in a class is kept in a file or table named "student" (Fig. 12.3).
- Database is, therefore, a place where related information is stored and various operations can be performed on it. It is the highest unit of file organization.

These databases are generally managed by special software known as **Database Management System (DBMS)**.

12.2 DATABASE MANAGEMENT SYSTEM (DBMS)

Database Management Systems are specially-designed applications that connect the user and program, and store data in an organized manner. The purpose of DBMS software is to allow the user to create, modify and control a database.

A DBMS stores data in such a manner that it becomes easier and highly efficient to retrieve, manipulate and produce information. Thus, a DBMS is an electronic or computerized record-keeping system. It maintains the various pieces of information in an integrated and summarized form instead of keeping them in separate independent files.

Examples of Database Management Systems are MS Access, MySQL, PostgreSQL, SQLite, Microsoft SQL Server, Oracle, SAP, dBase, FoxPro, etc.

Few customized DBMSs are computerized library systems, automated teller machines, flight reservation systems, computerized parts, inventory systems, etc.

A DBMS gives us tools to:

- Store data in a structured way.
- Query the database (*i.e.*, ask questions about the data).
- Sort and manipulate the data in the database.
- Validate the data entered and check for inconsistencies.
- Produce flexible reports, both on screen and on paper, that make it easy to comprehend the information stored in the database.

Also, it maintains data consistency in the case of multiple users.

CTM: A DBMS is a general purpose software system that facilitates the process of defining, constructing and manipulating databases for various applications.

12.2.1 Need for DBMS

The database system is used to eliminate the problems of data redundancy and data inconsistency. It does not maintain separate files for different applications. Rather, it works on the centrally maintained database, which means that the data is kept at one place and all the applications that require the data may refer to this database. Whenever any file gets updated, the updated version of the file is available to all applications using the database system, as shown in Fig. 12.5. So, data redundancy and data inconsistency are controlled to a large extent.

However, at times, there might be data redundancy due to some technical requirements in business applications. In such cases, we are required to maintain same data for different files but this is not recommended.

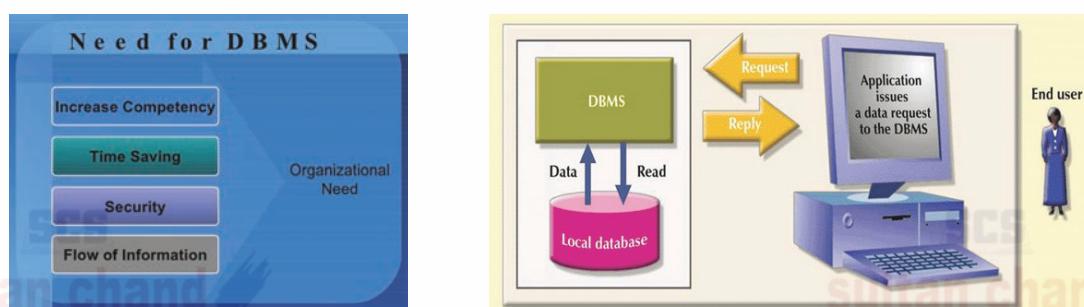


Fig. 12.5: Purpose of a Centralized Database System

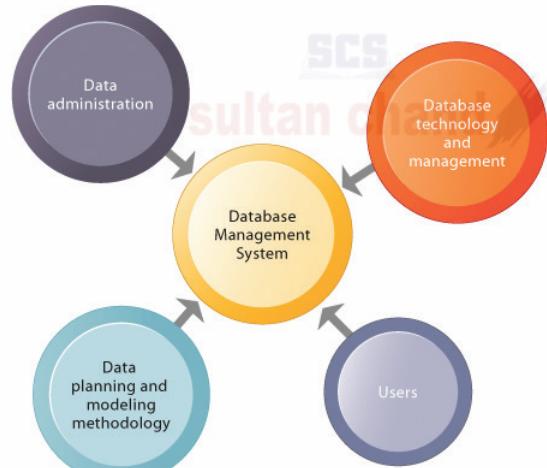


Fig. 12.4: Role of Database Management System

12.2.2 Components of a Database System

The various components of a database system are described in Fig. 12.6 below:

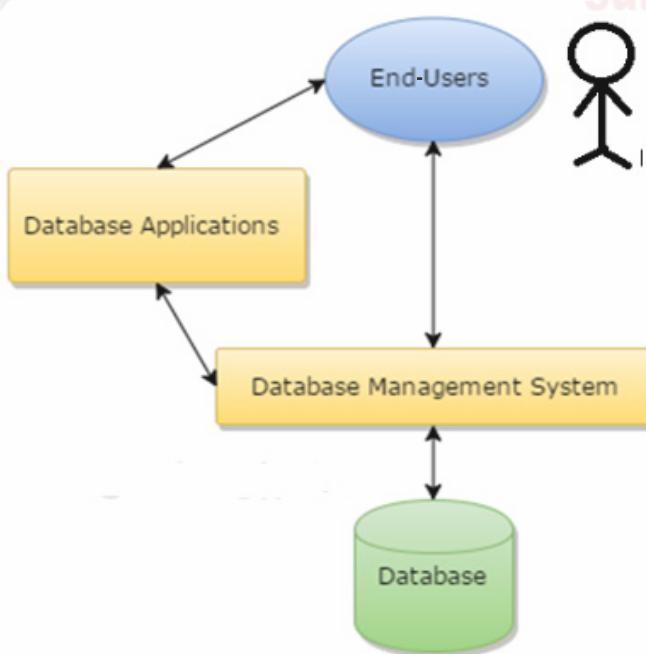


Fig. 12.6: Components of a Database System

Let us discuss these various components.

1. **Users:** Users can be of varied types, usually a DB administrator, System or Application developers and End-users. DBMS provides the following critical services to the user:
 - (a) **Database Creation:** A DBMS helps the user in creating and defining the required data or, in turn, a database. It manages and organizes the required data and databases.
 - (b) **Database Maintenance:** It helps in maintenance of data and database by addition, deletion, modification and regular updation of the tables and its records.
 - (c) **Database Processing:** A DBMS performs one of the major tasks of query processing—it processes the queries or the information requirement of users and retrieves necessary information from the database.
2. **Database Application:** Database application may be Personal, Departmental, Enterprise and Internal. It may be general-purpose or customized as per the needs of a user.
3. **DBMS:** Software that allows users to define, create, access and manage database(s) is termed as a DBMS. *For example, MySQL, Oracle, etc.*
4. **Database:** It is a collection of logical data.

12.2.3 Advantages of a DBMS

Apart from providing various salient features described above, a DBMS has several advantages over traditional data processing techniques.

1. **Elimination of Data Redundancy:** Duplication of data leads to wastage in storage space. A DBMS eliminates data redundancy (duplication of data) by integrating the files so that multiple copies of the same data are not stored.

2. **Data Consistency:** A DBMS provides data consistency to a larger extent as the changes made at one place are reflected at all other places or to all the users.
3. **Sharing of Data:** By using a DBMS, not only can existing applications share data in the database, but new applications can also be developed to operate against the same stored data.
4. **Reduced Programming Effort:** A DBMS saves a lot of programming effort since a user need not write programs for query processing involving several tables or files, report generation, addition, modification and deletion of data, etc. Thus, it provides easy retrieval of data.
5. **Database Enforces Standards:** With centralized control of the database, the DBA (Database Administrator) can ensure that all applicable standards are followed in the representation of data, i.e., format, documentation standards and conventions, etc.
6. **Improved Data Integrity:** Data integrity refers to the validity and consistency of stored data. *For example*, the system itself checks for the correct information to be entered by the user in the correct format. It consists of various constraints.
7. **Privacy and Security:** Data security refers to protection of data against accidental or intentional disclosure to unauthorized persons. Since there is centralized control, the data is protected.
8. **Economical:** Combining all the organization's operational data into one database and creating a set of applications that work on this single source of data can result in cost savings. The overall maintenance cost of data is reduced.
9. **Improved Backup and Recovery System:** A database system provides facilities for recovery from hardware or software failures.
10. **Meeting Enterprise Requirements than Individual Requirements:** Since many types of users with varying levels of technical knowledge use a database, a DBMS should provide a variety of user interfaces.

CTM: The repetition (duplication) of same data at multiple places in a database is known as **data redundancy**.

12.3 DBMS MODELS

Data models define how the logical structure of a database is modelled. A data model is an integrated collection of conceptual tools that can be used to describe the structure of the database along with the appropriate data types, relationships and constraints required to be applied on the data.

Data Models are used to implement abstraction in a DBMS. They are a communication tool. Data models define how data is connected and how it is processed and stored inside the system. They organize data for various users. A data model should be able to give best data representation and should possess the following desirable characteristics:

- (a) Data models should be presented graphically using diagrams and symbols.
- (b) Data representation in a data model should have no data redundancy.
- (c) A data model should be made available and shared by various applications.
- (d) Data represented should be consistent, stable and valid in all aspects.

12.3.1 Types of Data Models

Data models are categorized into three different categories:

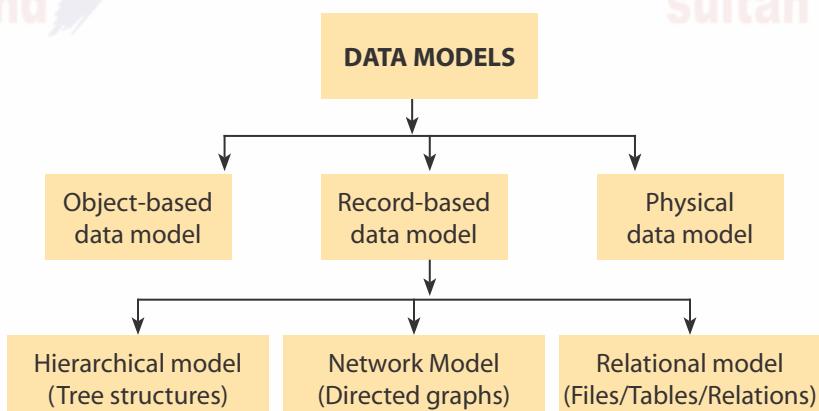


Fig. 12.7: Classification of Data Models

12.3.1.1 Hierarchical Data Model

In a hierarchical model, records are organized as trees rather than graphs, *i.e.*, hierarchical model represents a hierarchy of parent and child data segments. It is represented by an upside-down “tree”. Hierarchical model depicts one-to-many (1:M) relationships between a parent and child segment. Each parent can have many children, but vice versa does not hold true, *i.e.*, each child has only one parent.

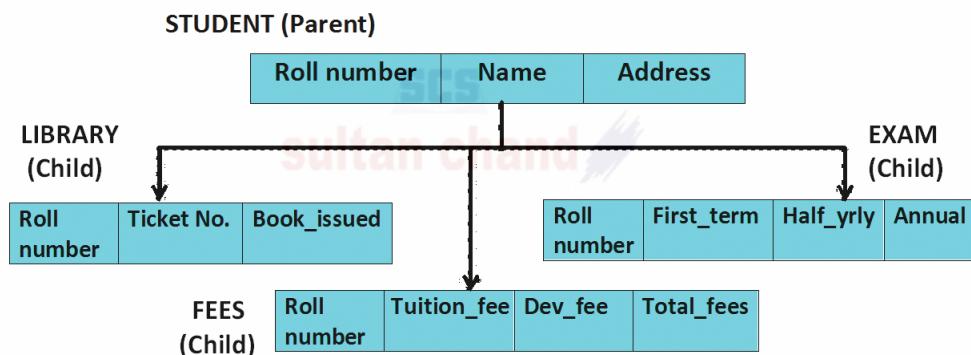


Fig. 12.8: Hierarchical Model

From Fig. 12.8, we can see that the student database contains various files related to a student. In this kind of arrangement, a file higher in the hierarchy is known as parent of the files contained inside it. Here, Student is a parent of the Library, Fees and Exam files respectively. Thus, this model represents a one-to-many relationship between a parent and its children in the form of a hierarchical (upside-down) tree. The **advantages** of this model are:

1. It is simple in operation and in concept.
2. It promotes data sharing.
3. Parent-child relationship promotes data integrity.
4. It is efficient and can be naturally mapped for 1: M (one-to-many) relationships.

The **limitations** of this model are as follows:

1. It is suitable for hierarchical relationships only; otherwise, it is inflexible for non-hierarchical relationship mapping.
2. Implementation is complex and difficult since it is done using pointers from a parent to its children which require extra memory space.

3. Dependency on parent node, which is the root node, and its deletion can cause deletion of all child nodes and, in turn, the entire model.
4. Changes in the structure require changes in all the applications.

IBM's IMS (Information Management System) and system2000 are examples of hierarchical database management system.

12.3.1.2 Network Data Model

The network model is a database model conceived as a flexible way of representing objects and their relationships. It consists of a collection of records connected to one another through **links**. The network model replaces the hierarchical model with a graph, thus allowing multiple connections among the nodes. The major point of difference between the two models is that unlike the hierarchical model, the network model provides the capability to handle many-to-many (M:M) relationships. Thus, a child can have more than one parent.

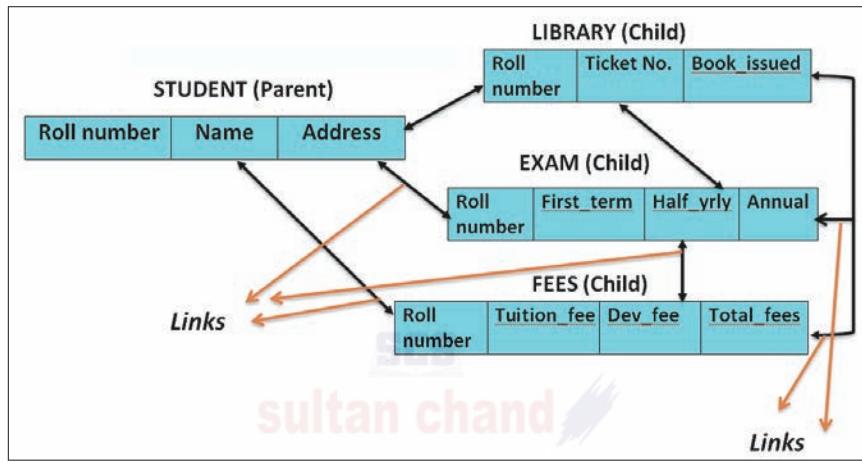


Fig. 12.9: Network Model

From Fig.12.9, it is clear that this model represents many-to-many relationships. It can be observed that the file EXAM is associated with three tables, viz. LIBRARY, FEES and STUDENT. Thus, the figure shows that a child can have more than one parent. This model is an improved version of hierarchical model having records along with pointers. These pointers establish many-to-many relationships and are termed as **Links**. Each record has its respective pointer or link with which it is associated.

The **advantages** of network model are:

1. It can handle more relationship types, *i.e.*, M:M (many-to-many) multi-parent relationships.
2. Data access is easy and more flexible.
3. It includes DDL (Data Definition Language) as well as DML (Data Manipulation Language).
4. It promotes better data integrity.

The **limitations** of this model are:

1. The system is quite complicated in structure since there are several number of links which also limits efficiency.
2. Structural changes require changes to be made in all application programs.
3. A high-level language interface is required to interact with the database.
4. Extra memory is consumed for holding links (pointers).

UNIVAC's DMS 1100 and DBMS 10-20 from DEC are examples of network database management system.

12.3.1.3 Relational Data Model

The most popular data model in DBMS is the relational model. It is a more scientific and stable model than hierarchical or network model since there are no pointers involved with the records, and in case of any fault or error these pointers may result in inconsistency among the records and can lead to reduced data integrity.

In this model, data is organized in two-dimensional tables called **relations**. The tables or relations are related to each other. A relational database is based on this relational model developed by E.F. Codd. In this type of database, the data and relations between them are organized into tables, having logically related records containing the same fields. The contents of a table or relation can be permanently saved for future use. Thus, this model is simple and has all the properties and capabilities required to process data with storage efficiency.

A relational model consists of a collection of tables, each of which is assigned a unique name, *i.e.*, it represents the database as a collection of relations or tables. **Characteristics** of relational database are:-

1. Data is conceptually represented as an orderly arrangement of data into rows and columns, termed as a relation.
2. Each relation is represented as a table.
3. Columns described in a table are the attributes that belong to an entity which is modelled as a table.
4. Every row in a table represents a single entity.
5. All the values in a relation are scalar, *i.e.*, at any given row or column position, there is one and only one value.
6. All operations are performed on the entire relation and the result is an entire relation.

Advantages of a relational model are as follows:

1. A relational model provides structural independence by using independent tables.
2. Changes made in the table structure do not affect the data access or other application programs.
3. It is represented in the form of tables; so, it is simple and easier to understand.
4. Tabular view also provides easier database design, use, implementation and management.
5. Built-in query support based on SQL is provided by RDBMS (Relational Database Management System).
6. Data organization and manipulation is easy, flexible and faster.
7. Powerful structure designing and processing capabilities of RDBMS isolate the end-user from physical-level details and, thus, improve implementation and management simplicity.
8. Mathematical operations can be successfully carried out using RDBMS.

The **limitations** of relational model are:

1. RDBMS incurs hardware and system software overheads.
2. The size of database becomes very large.

12.4 RELATIONAL DATABASE

A **relational database** is a type of database that stores and provides access to data points that are related to one another. Relational databases are based on the relational model, an intuitive, straightforward way of representing data in tables. In a relational database, each row in the table is a record with a unique ID called the key. The columns of the table hold attributes of the data and each record usually has a value for each attribute, making it easy to establish the relationships among data points.

Often, data in a relational database is organized into tables.

Basic Terminologies related to a Relational Database

- Entity:** An entity is something that exists and about which we can store some information. It is an object which can be distinctly identified. *For example*, student entity, employee entity, item entity, etc. Entity becomes the name of the table.
- Attribute:** In a relational table, an attribute is a set of values of a particular type. The term attribute is also used to represent a column. A table consists of several records (row); each record can be broken into several smaller entities known as fields or attributes or columns. A set of attributes defines the characteristics or properties of an entity. In the given table, Student relation (Fig. 12.11) consists of four fields or attributes—Roll number, Name, Address and Gender.
- Tuple:** Each row in a table is known as tuple. It is also called a row/record. A single entry in a table is called a record or row. A record in a table represents a set of related data. *For example*, the Student table given below has 10 records.
- Cardinality of Relation:** It is the number of records or tuples in the relation. Thus, the cardinality of Student relation is 10.
- Degree of Relation:** Number of columns or attributes is known as degree of a relation. Thus, the degree of Student relation is 4.
- Domain of Relation:** It defines the kind of data represented by the attribute. It is the set of all possible values that an attribute may contain. *For example*, in the given table Student, domain for the field Gender is two since it can have either 'M' or 'F' as the possible and available values that it may contain.
- Body of the Relation:** It consists of an unordered set of 0 or more tuples.

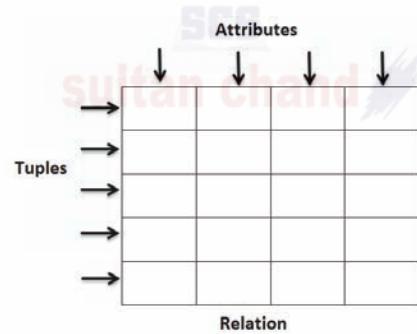


Fig. 12.10: Tuples and Attributes

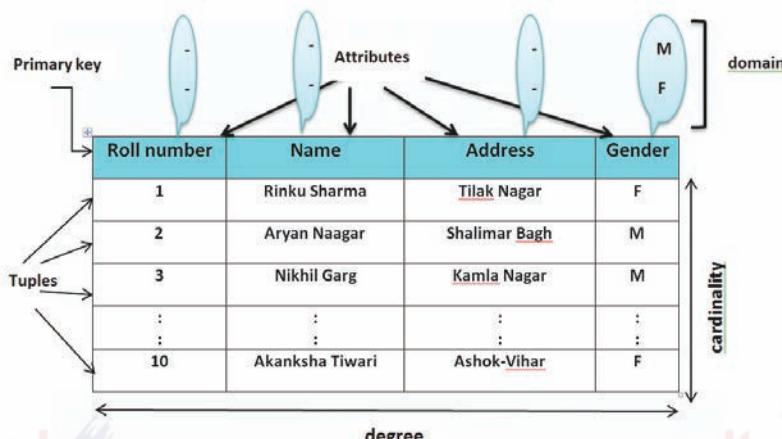


Fig. 12.11: Relational Model (Student Relation)

Therefore, in reference to Fig. 12.11, in the above Student relation:

- There are 10 tuples (*i.e.*, cardinality=10) and 4 attributes (*i.e.*, degree=4).
- Roll number, Name, Address and Gender are the attribute names.
- The first tuple contains the values (1, "Rinku Sharma", "Tilak Nagar", "F").
- The domain of the attribute Gender is (M,F).

12.5 DATABASE KEYS

Keys are a very important part of relational database. They allow us to identify an attribute or a set of attributes on the basis of which a table is identified. They are used to establish and identify relation among two or more tables. They also ensure that each record within a table can be uniquely identified by a combination of one or more fields within a table.

The different types of keys in a RDBMS are as follows:

KEY	DESCRIPTION
Primary Key	A primary key is an attribute or a group of attributes that can uniquely identify tuples within the relation.
Candidate Key	A candidate key is one that is capable of becoming the primary key. (<i>i.e.</i> , candidate for primary key position).
Alternate Key	A candidate key that is not the primary key is called an alternate key.
Foreign Key	A non-key attribute, whose value(s) are derived from the primary key of some other table, is known as foreign key in its current table.

Let us discuss these keys in detail.

1. **Primary Key:** A primary key is a set of one or more attributes/fields which uniquely identifies a tuple/row in a table. The salient features of a primary key are as follows:

- (a) It is always unique in nature, *i.e.*, non-redundant. It does not have duplicate values in a relation.
- (b) It arranges the table in its own order.
- (c) It cannot be re-declared or left null.
- (d) One table can have only one primary key; however, primary key can be a combination of more than one field. *For example*, roll number along with admission_no can be combined together and can be declared as a primary key in the relation Student. In the table Item given below, Item_id is the primary key while Supp_id (supplier id) is the primary key in the table Supplier.

Table: Item

Item_id	Item_name	Qty
I101	Printer	400
I102	CD	200
I104	DVD	150
I105	Mouse	300
I103	Keyboard	180
I109	Cable	500

Fig. 12.12: Keys in a Database (Tables)

Table: Supplier

Supp_id	Area
S01	Ashok-Vihar
S02	Noida
S04	CP
S05	Punjabi Bagh



2. **Candidate Key:** A candidate key refers to all the attributes in a relation that are candidates or are capable of becoming a primary key. We mark it virtually.

In the above Item table, Item_id and Item_name are the candidate keys. Out of these keys, Item_id is the primary key and Item_name becomes the alternate key. Similarly, in the case of Supplier relation, Supp_id and Area are the candidate keys, Supp_id is the primary key and Area becomes the alternate key. Thus, the equation becomes:

Candidate Keys – Primary Key = Alternate Key

3. **Alternate Key:** A candidate key that is not the primary key is called an alternate key. In other words, any attribute that is a candidate for the primary key, i.e., which is capable of becoming a primary key but is not a primary key, is an alternate key. *For example*, in a Customer table, cust_name is the alternate key. Similarly, in the above table Item, Item_name becomes the alternate key.
4. **Foreign Key:** A foreign key is a non-key attribute whose value is derived from the primary key of another table; in other words, a primary key in some other table having relationship with the current or original table.

Foreign Key



Table: Employee

Item_id	Emp_name	Desig_code
E01	Ankur Mehta	Mgr
E02	Deepika Gupta	Dir
E04	Arnav Bansal	Asst_mgr
E03	Harshit Singh	Acc
E05	Kirti Dubey	Mgr

Table: Department

Desig_code	Designation
Mgr	Manager
Dir	Director
Acc	Accountant
Asst_mgr	Assistant Manager

In the above table, Desig_code is the primary key in the table Department which, when related with the table Employee, becomes a foreign key to it.

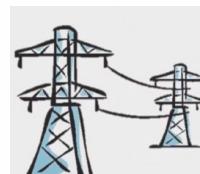
This was all about the database concepts. Now, we will be moving on to their implementation using SQL.

12.6 INTRODUCTION TO SQL

Database Management Systems (DBMS) are not a new concept and had been first implemented in the 1960's. With time, database technologies evolved a lot while usage and expected functionalities of databases increased immensely. These days, there is no organization which does not manage its data and records using some type of DBMS. An online telephone directory uses DBMS to store data pertaining to people, phone numbers and other contact details. Apart from this, your electricity service provider uses a DBMS to manage billing, client-related issues, to handle fault data, etc. Not to forget, Facebook—it needs to store, manipulate and present data related to its members, their friends, member activities, messages, advertisements and a lot more.



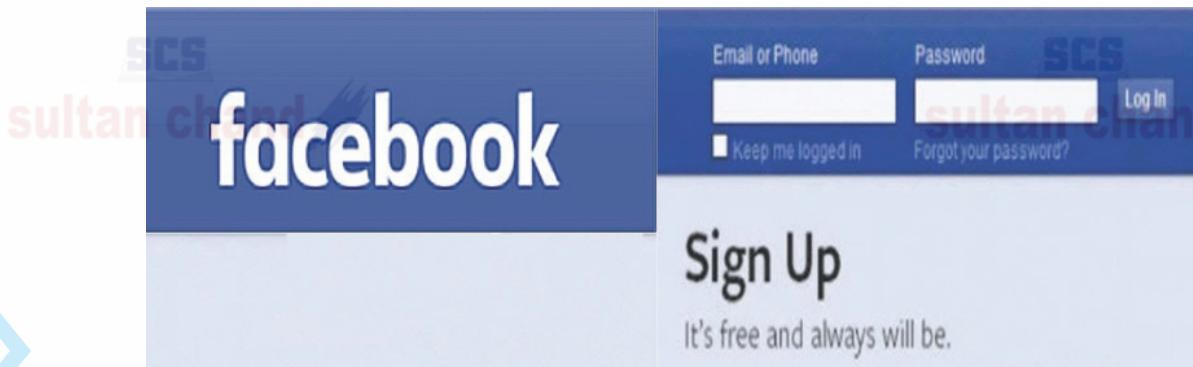
Online
Telephone Directory



Electricity
Billing System

12.11

Relational Database and SQL



All these real-life applications require a DBMS to manipulate and handle this enormous data. A DBMS requires some language to handle and manipulate its data, which is known as Structured Query Language (SQL). The following topic deals exclusively with relational databases, their tables and retrieving data using Structured Query Language (SQL).

12.7 OVERVIEW OF SQL AND MySQL

SQL (Structured Query Language) is a standard language for accessing and manipulating databases. SQL commands are used to create, transform and retrieve information from Relational Database Management Systems and are also used to create interface between a user and database. By using SQL commands, one can search for any data in the database and perform other functions like creating tables, adding records, modifying data, removing rows, dropping tables, etc.

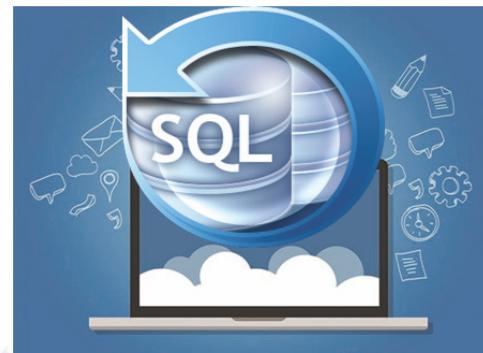


Fig. 12.13: Structured Query Language



MySQL is an open-source and freely-available Relational Database Management System (RDBMS) that uses Structured Query Language (SQL). It provides excellent features for creating, storing, maintaining and accessing data, stored in the form of databases and their respective tables. A single MySQL database can store several tables at a time and can store thousands of records in it.

Being an open-source software, it can be freely and easily downloaded from the site www.mysql.org. MySQL is developed and supported by Sweden-based company, MySQL AB. It is fully secured, reliable, and fast, and possesses far better functionalities than many other commercial RDBMs available in the market.

12.8 FEATURES OF SQL

SQL is the most common language used to create, operate, update, manipulate and communicate with a database.

In 1970, SQL was developed by Donald D. Chamberlin and Raymond F. Boyce at IBM. Thus, SQL is a fourth generation non-procedural language that is used to create, manipulate and process the databases (relations). In other words, these languages describe what data is to be retrieved, inserted and updated or deleted from a database.

It has the following salient features and strong processing capabilities:

- It can retrieve data from a database through Query processing.
- It can insert records in a database.
- It can update records in a database.
- It can create new databases and modify the existing ones.
- It can create new tables in a database.
- It can create views in a database.
- It allows modifying the security settings of the system.

CTM: SQL (Structured Query Language) is a unified, non-procedural language used for creating, accessing, handling and managing data in relational databases.

12.9 ADVANTAGES OF SQL

SQL has the following advantages:

1. **Ease of use:** It is very easy to learn and use and does not require high-end professional training to work upon it.
2. Large volume of databases can be handled quite easily.
3. **No coding required:** It is non-procedural and a unified language, *i.e.*, we need not specify the procedures to accomplish a task but only need to give a command to perform the activity.
4. SQL can be linked to most of the other high-level languages which makes it the first choice for database programmers.
5. **Portable:** It is compatible with other database programs like Dbase IV, FoxPro, MS Access, DB2, MS SQL Server, Oracle, Sybase, etc.
6. SQL is not a case-sensitive language, *i.e.*, both capital and small letters are recognized.

12.10 CLASSIFICATION OF SQL STATEMENTS

SQL is the language used to interact with the database. The SQL statements or commands that we type are the statements that are regarded as the instructions to the database.

SQL provides different types of statements or commands for different purposes. These statements are classified into the following categories:

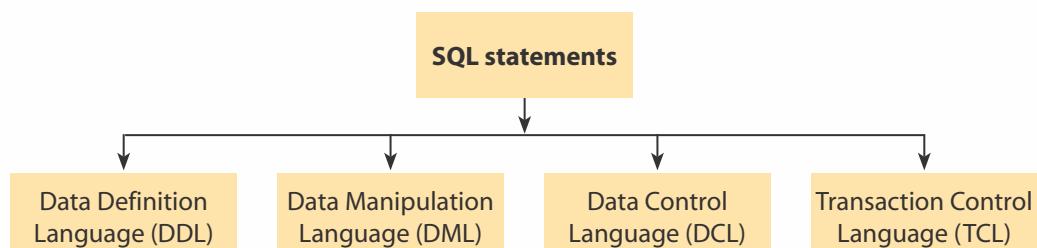


Fig. 12.14: Classification of SQL statements

DCL and TCL are beyond the scope of this book. So, we shall be discussing only DDL and DML commands in detail.

12.10.1 Data Definition Language (DDL) Commands

The DDL part of SQL permits database tables to be created or deleted. It also defines indices (keys), specifies links between tables and imposes constraints on tables. It contains the necessary statements for creating, manipulating, altering and deleting the table.

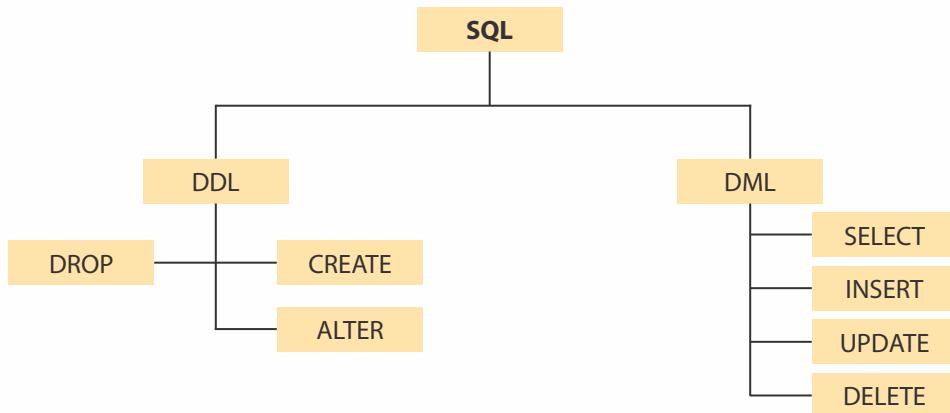


Fig. 12.15: Most Commonly Used SQL Commands

Examples of DDL commands in SQL are:

- **CREATE DATABASE:** creates a new database.
- **USE command:** to select and open an already existing database.
- **CREATE TABLE:** creates a new table.
- **ALTER TABLE:** modifies a table.
- **DROP TABLE:** deletes a table.

CTM: The DDL command lets us define the database structure and its related operations.

The DDL provides a set of definitions to specify the storage structure and access methods used by the database system and also defines proper and relevant data types.

12.10.2 Data Manipulation Language (DML) Commands

A Data Manipulation Language (DML) is a part of SQL that helps a user access or manipulate data. The DML statements are executed in the form of queries which are handled by the DML compiler. It contains the set of statements to:

1. Retrieve data from the tables of the database.
2. Insert data into the tables of the database.
3. Delete data from the tables of the database.
4. Update data among the rows/records in the tables of the database.

DML commands carry out query-processing operations and manipulate data in the database objects. Several DML commands available are:

1. **SELECT statement:** To extract information from the table; may or may not be on the basis of certain conditions/criteria.
2. **INSERT INTO statement:** To insert new data (record) into a table.

Learning Tip: The query and update commands form the DML part of SQL. They enable the user to access or manipulate data stored in a database.

3. **UPDATE statement:** To modify or change the data (tuple) in a table (not modifying the data type of column).
4. **DELETE:** To delete data (tuple) from a table (not deleting a column).

Table 12.1 Difference between DDL and DML commands

DDL Commands	DML Commands
1. DDL stands for Data Definition Language.	1. DML stands for Data Manipulation Language.
2. These commands allow us to perform tasks related to data definition, i.e., related to the structure of the database objects (relations/databases).	2. These commands are used to manipulate data, i.e., records or rows in a table or relation.
3. The examples of DDL commands are Create, Alter, Drop, Grant, Revoke, etc.	3. The examples of DML commands are Insert into, Update, Delete, Select, etc.
4. DDL is not further classified.	4. DML commands are further classified into two types: (a) Procedural DMLs (b) Non-Procedural DMLs

Let us start implementing SQL using MySQL as the platform.

12.11 MYSQL

MySQL is an open-source and freely available Relational Database Management System (RDBMS) that uses Structured Query Language (SQL). It provides excellent features for creating, storing, maintaining and accessing data, stored in the form of databases and their respective tables. A single MySQL database can store several tables at a time and can store thousands of records in it.

Being an open-source software, it can be freely and easily downloaded from the site www.mysql.org. MySQL is developed and supported by a Sweden-based company, MySQL AB. It is fully secured, reliable, and fast, and possesses far better functionalities than many other commercial RDBMs available in the market.

MySQL database system works upon Client/Server architecture. It constitutes a MySQL server which runs on a machine containing the databases and MySQL databases (clients), which are connected to these server machines over a network.

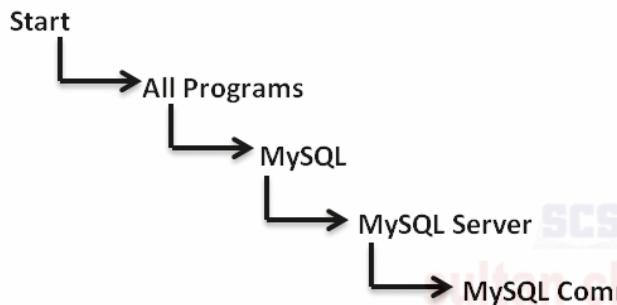
☞ **Advantages of MySQL:** MySQL provides the following salient features and advantages:

1. **Reliability and Performance:** MySQL is a very reliable and high performance Relational Database Management System.
2. **Modifiable:** Being an open-source software, MySQL comes with its source code; so, it is easily modifiable and we can recompile its associated source code.
3. **Multi-Platform Support:** MySQL supports several different platforms like UNIX, Linux, Mac OS X and Microsoft Windows.
4. **Powerful Processing Capabilities:** MySQL is a powerful, easy, compatible and fast Relational Database Management System. It can handle complicated corporate applications and processing requirements.
5. **Integrity (Checks):** MySQL provides various integrity checks in order to restrict the user input and processing.
6. **Authorization:** MySQL provides DDL commands to check for user authentication and authorization by restricting access to relations and views.

7. **Powerful Language:** All SQL operations are performed at a prescribed and fixed level, i.e., one SELECT command can retrieve data from multiple rows and one MODIFY command can edit multiple rows at a time. These features make SQL a very powerful language as compared to other languages where one command can process only a single record at a time.
8. **Reliable:** SQL provides a high level of well-defined set of commands that provides the desirable results without any ambiguity.
9. **Freedom of Data Abstraction:** SQL provides a greater degree of abstraction freedom compared to any other procedural language.
10. **Complete Language for a Database:** Apart from being a strong query processing language, it can also be used to create, insert, delete and control access to data in databases.

12.11.1 Starting the MySQL Database

MySQL is an open-source database system. You can download and install it directly from the internet. After installing, you need to start working with MySQL by following the given steps:



Alternatively, (for Windows 7 and above)

Start → Apps by name

→ MySQL Command Line Client

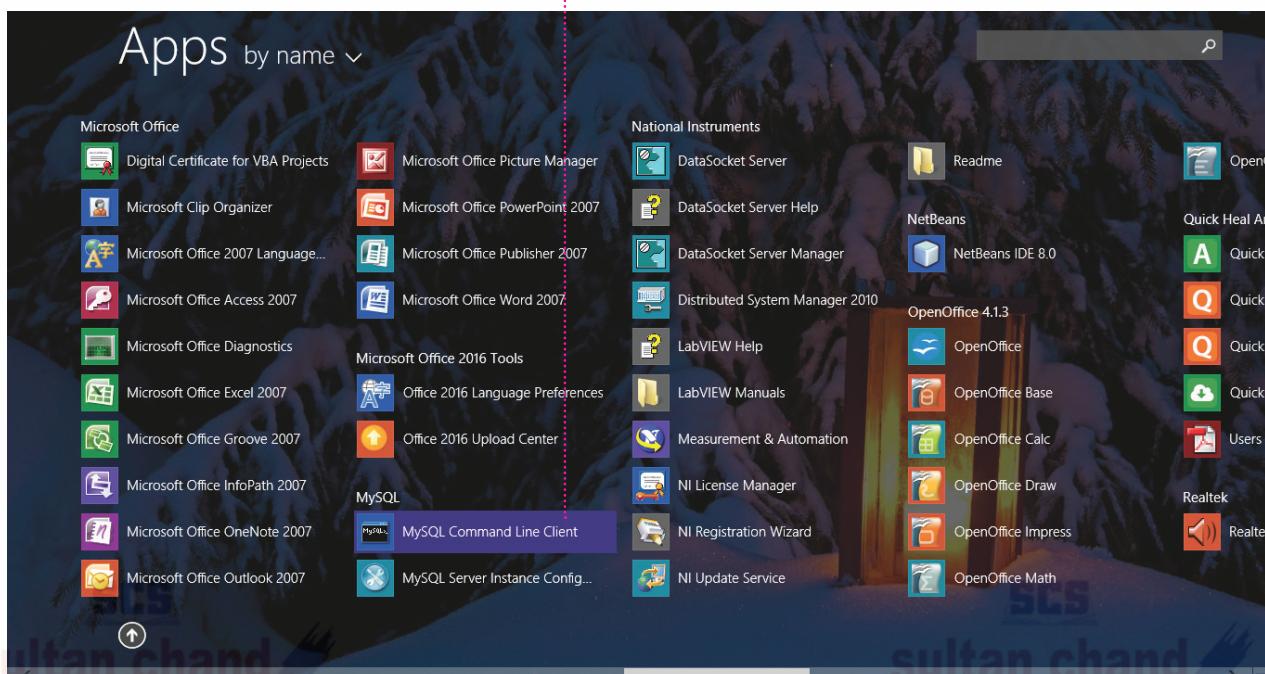
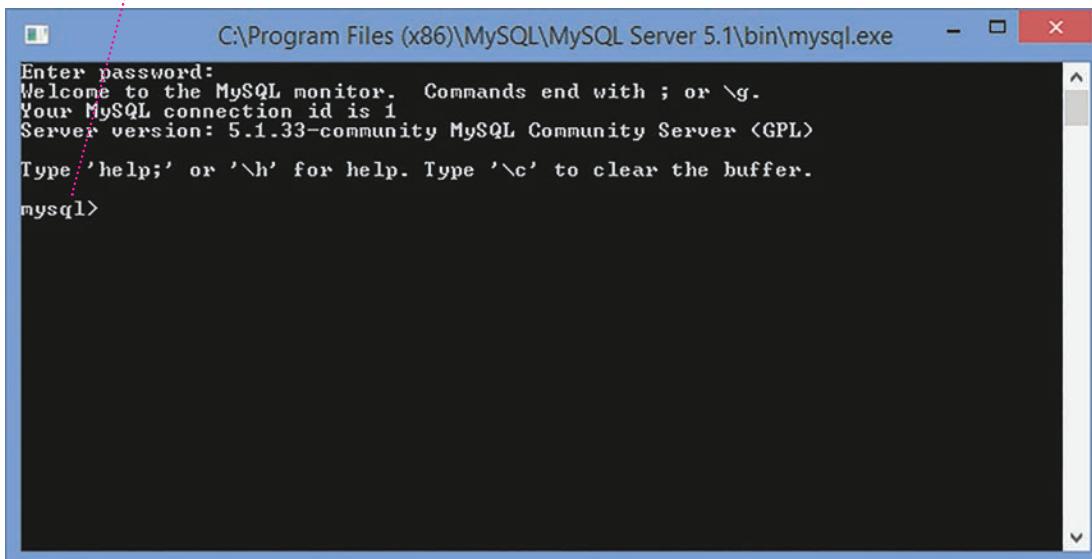


Fig. 12.16: Steps to Start MySQL

After opening MySQL, the screen of the MySQL command prompt appears where you need to specify a password to work with it.

After entering the password, the MySQL prompt appears, where you start typing the SQL commands, as shown in Fig. 12.17. In order to come out of MySQL application, you can type quit in front of the mysql> command prompt, as shown in Fig. 12.18.

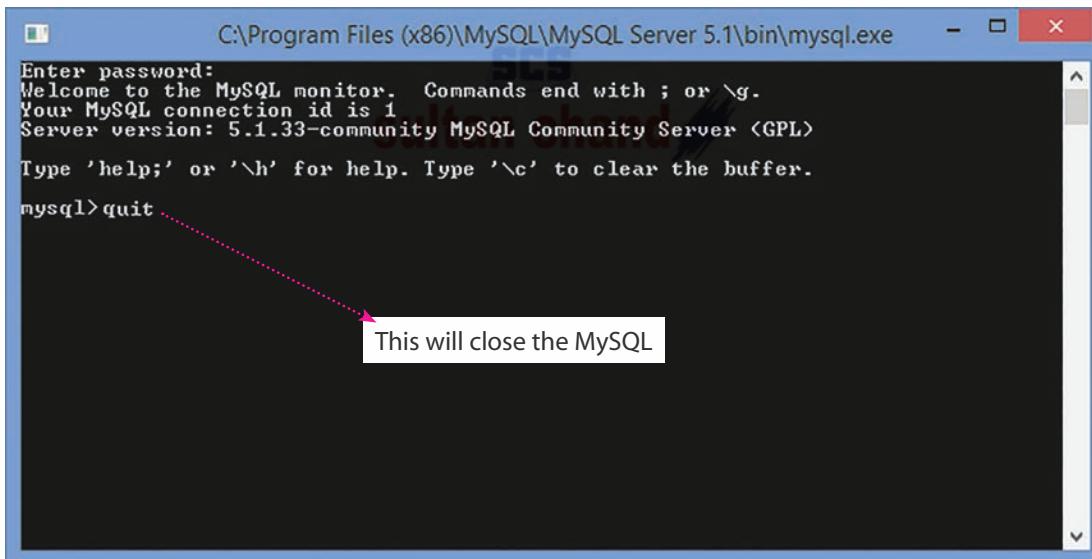


```
C:\Program Files (x86)\MySQL\MySQL Server 5.1\bin\mysql.exe
Enter password:
Welcome to the MySQL monitor. Commands end with ; or \g.
Your MySQL connection id is 1
Server version: 5.1.33-community MySQL Community Server (GPL)

Type 'help;' or '\h' for help. Type '\c' to clear the buffer.

mysql>
```

Fig. 12.17: MySQL Prompt



```
C:\Program Files (x86)\MySQL\MySQL Server 5.1\bin\mysql.exe
Enter password:
Welcome to the MySQL monitor. Commands end with ; or \g.
Your MySQL connection id is 1
Server version: 5.1.33-community MySQL Community Server (GPL)

Type 'help;' or '\h' for help. Type '\c' to clear the buffer.

mysql> quit
This will close the MySQL session.
```

Fig. 12.18: Closing MySQL

Learning Tips:

1. Some database systems require a semicolon (;) at the end of each SQL statement.
2. Semicolon is the standard way to separate each SQL statement in database systems that allows more than one SQL statement to be executed in the same call to the server.
3. SQL is NOT case-sensitive; select is the same as SELECT.

12.12 SQL DATATYPES

Just like any other programming language, the facility of defining data of various types is available in SQL also. SQL supports the following datatypes for the specification of various data-items or fields of a relation/table. In SQL, each column of the table is assigned a datatype which conveys the kind of value that will be stored in the column.

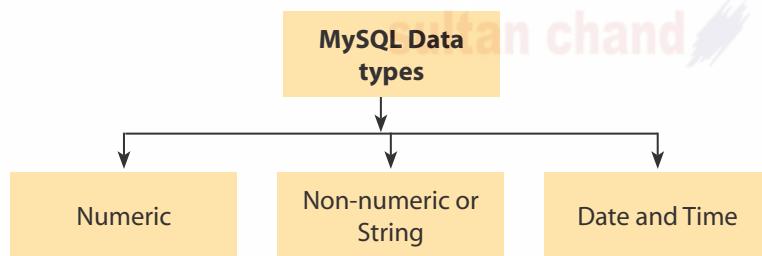


Fig. 12.19: MySQL Datatypes

Datatype	Syntax	Description
INTEGER (Numeric)	INTEGER or integer	It stores/represents positive whole numbers up to 11 digits and negative whole numbers up to 10 digits. The range of integer is from -2,147,483,648 to 2,147,483,647.
SMALLINT	SMALLINT	It is a 16-bit signed integer value that stores whole numbers in the range -32768 to 32767. Its width is up to 5 digits.
NUMERIC	NUMERIC(x,y)	<p>Numbers are stored in the given format, where x is the total number of digits and y is the number of places to the right of the decimal point. x must include an extra place for the decimal point.</p> <p><i>For example, Numeric(8,2)</i></p> <p>In the given example, numeric datatype stores a number that has 5 places before the decimal and 2 digits after the decimal and 1 digit place for the decimal point. Numeric holds up to 20 significant digits. A negative number holds one place for the sign, i.e., (-).</p>
DECIMAL	DECIMAL(x,y) Or DECIMAL(size, precision) decimal point	<p>Numbers stored in the DECIMAL format, where x is the size, i.e., total number of digits, and y is precision, i.e., it is the number of places to the right of the decimal point.</p> <p><i>For example, Decimal(8,2)</i></p> <p>In the above example, decimal datatype stores a number that has 6 digits before the decimal and 2 digits after the decimal.</p> <p>Decimal holds up to 19 significant digits. A negative number uses one place for its sign (-).</p>
CHARACTER (fixed length)	CHAR(x) or CHAR(size)	This datatype stores 'x' number of characters in the string which has a fixed length. A maximum of 254 characters can be stored in a string. If you store strings that are not as long as the 'size' or 'x' parameter value, the remaining spaces are left unused. <i>For example, if you specify CHAR(10), strings such as "ram" and "technology" are each stored as 10 characters. However, a student admission_no is 6 digits long in a school, so CHAR(6) would be appropriate to store the admission_no of all the students.</i> This datatype is suitable where the number of characters to store is fixed. The value for CHAR data type has to be enclosed in single or double quotation marks.



CHARACTER (variable length)	VARCHAR(x) or VARCHAR2(x)	This datatype is used to store variable length alphanumeric data. <i>For example</i> , address of a student can be declared as VARCHAR(25) to store the address up to 25 characters long. The advantage of using this datatype is that VARCHAR will not leave unused spaces. It releases the unused memory spaces.
DATE	DATE	This datatype is used to store a date in 'yyyy/mm/dd' format. It stores year, month and date values. DATE values can be compared with each other only. The date values to be entered are to be enclosed in {} or with single quotation marks.
TIME	TIME	This datatype is used to store time in hh:mm:ss format. It stores hour, minute and second values. <i>For example</i> , a time of day can be taken as 12:30:45p.m. where 12 means hours, 30 means minutes and 45 refers to seconds.
BOOLEAN (logical)	BOOLEAN	This datatype is used for storing logical values, either true or false. In both upper and lower case, T or Y stands for logical true and F or N stands for logical false. The fields with Boolean (logical) datatype can be compared only with other logical columns or constants.
BLOB/RAW/ LONG RAW	BLOB or RAW or LONG RAW	This datatype can store data up to a maximum length of 65535 characters. BLOBs are "Binary Large Objects" and used to store a large amount of data such as images, animations, clips or other types of files. <i>For example</i> , image raw(2000);
MEMO/LONG	MEMO or LONG	This datatype allows storing characters or remarks up to 2 GB per record.

Table 12.2 Difference between CHAR and VARCHAR datatypes

CHAR	VARCHAR
1. CHAR datatype provides fixed length memory storage. It specifies a fixed length character string.	1. VARCHAR datatype provides variable length memory storage. It specifies a variable length string (changeable).
2. The CHAR datatype can store a maximum of 0 to 255 characters.	2. The VARCHAR datatype can store a maximum number up to 65,535.
3. CHAR datatype is used when the data entries in a column are expected to be of the same size.	3. VARCHAR datatype is used when the data entries in a column are expected to vary considerably in size.
4. CHAR(x) will take x characters of storage even if you enter less than x characters to that column.	4. VARCHAR(x) will take only the required storage for the actual number of characters entered to that column.
5. If a value entered is shorter than its length x, then blanks are added.	5. No blanks are added if the length is shorter than the maximum length, x.
6. CHAR datatype takes memory space of 1 byte per character.	6. VARCHAR takes up memory space of 1 byte per character, +2 bytes to hold variable length information.
7. Search operation is faster with CHAR datatype column.	7. Search operation works slower in VARCHAR datatype column as compared to CHAR type.
8. <i>For example</i> , name char(10); name="anu"; name field occupies 0 bytes, with the first three bytes with values and the rest with blank data.	8. <i>For example</i> , name varchar(10); name="anu"; then name occupies only 3+2=5 bytes, the first three bytes for value and the other two bytes for variable length information.

CTM: While defining datatype for columns or attributes in a relation, two points should be kept in mind:

- When using fixed length data in columns like phone number, area code, use character datatype.
- When using variable length data in columns like name, address, designation, use varchar datatype.

12.13 SQL COMMANDS

SQL provides a predefined set of commands that help us to work with relational databases. Before discussing these commands, we must be familiar with the conventions and basic terminologies related to SQL. Throughout this chapter, the words keyword, clause and statement have been used.

A keyword refers to an individual SQL element that has a special meaning in SQL. *For example*, SELECT and FROM are keywords. A clause is a distinct logical part of an SQL statement. Clauses begin with a keyword for which they are named and consist of arguments as well. *For example*, SELECT empno, ename, FROM employee WHERE salary>45000, are clauses in SQL. Here, arguments complete or modify the meaning of a clause, which is salary in the given example.

A statement or command is a combination of two or more clauses. Statements are basically the instructions given to SQL database for executing any task. *For example*, SELECT * FROM employee is an SQL statement. An important point to remember here is that all the statements in SQL terminate with a semi-colon (;). Also, SQL is not case sensitive; therefore, we can type commands in either upper case or lower case.

Let us now learn how a database and tables in a database are created in SQL. A database is used to house data in the form of tables. Therefore, before creating a table, it is mandatory to create a database first.

We shall create a sample database School and then create a table Student in it.

Database: School

Tables_in_School
Student
Fees

Table: Student

Rollno	Name	Gender	Marks	DOB
1	Raj Kumar	M	93	17-Nov-2000
2	Deep Singh	M	98	22-Aug-1996
3	Ankit Sharma	M	76	02-Feb-2000
4	Radhika Gupta	F	78	03-Dec-1999
5	Payal Goel	F	82	21-April-1998
6	Diksha Sharma	F	80	17-Dec-1999
7	Gurpreet Kaur	F	65	04-Jan-2000
8	Akshay Dureja	M	90	05-May-1997
9	Shreya Anand	F	70	08-Oct-1999
10	Prateek Mittal	M	75	25-Dec-2000

Table: Fees

Rollno	Name	Fees
1.	Raj Kumar	8000
2.	Deep Singh	9000
5.	Payal Goel	7500
6.	Diksha Sharma	8000
9.	Shreya Anand	8020
10.	Prateek Mittal	9200

To get started on our own database, we can first check which databases currently exist in MySQL server. Use the SHOW DATABASES statement to find out which databases currently exist on the server:

mysql> show databases;

```
+-----+
| Database   |
+-----+
| mysql      |
| test       |
+-----+
2 rows in set (0.01 sec)
```

1. Creating Databases

The **CREATE DATABASE** command is used to create a database in RDBMS.

Syntax for creating a database:

```
CREATE DATABASE <database_name>;
```

For example,

CREATE DATABASE school; ←— Creates database with the name school.

When the above-mentioned command gets executed, a database with the name school will be created on the system.

2. Opening Databases

Once a database has been created, we need to open it to work on it. For this, **USE** command is required.

Syntax for opening a database:

```
USE <database_name>;
```

For example,

```
mysql>USE school;
```

Database changed

3. Removing Databases

To physically remove/delete a database along with all its tables, **DROP** command is used.

Syntax for removing a database:

```
DROP DATABASE <database_name>;
```

For example,

```
mysql>DROP DATABASE school;
```

Database deleted

4. Creating a Table

The **CREATE TABLE** statement is used to create a table in a database. Tables are organized into rows and columns, and each table must have a name. It is the most extensively used DDL command. A table must have at least one column.

Syntax for creating a table:

```
CREATE TABLE <table_name>
(
<column_name1><data_type> [(size)],
<column_name2><data_type> [(size)],
<column_name3><data_type> [(size)],
.....
);
```

For example,

mysql>CREATE TABLE student

```
(1) Rollno    integer      NOT NULL PRIMARY KEY,
     Name      varchar(20)   NOT NULL,
     Gender    char(1),
     Marks     integer(11)
     DOB       date );
```

Query OK, 0 rows affected (0.04 sec)

For each column, a name and a datatype must be specified and the column name must be unique within the table definition. Column definitions are separated by comma. Upper case and lower case letters make no difference in column names; the only place where upper and lower case letters matter are string comparisons.

5. Viewing a Table

To verify that the table has been created, **SHOW TABLES** command is used.

mysql> show tables;

```
+-----+
| Tables_in_school |
+-----+
| student          |
| fees             |
+-----+
```

2 rows in set (0.01 sec)

6. Viewing a Table Structure

To view a table structure, **DESCRIBE** or **DESC** command is used.

Syntax: DESCRIBE <tablename>; or DESC <tablename>;

For example,

mysql> describe student;

```
+-----+-----+-----+-----+-----+-----+
| Field    | Type      | Null   | Key    | Default | Extra   |
+-----+-----+-----+-----+-----+-----+
| Rollno   | integer(11) | YES    |        | NULL    |          |
| Name     | varchar(20)  | YES    |        | NULL    |          |
| Gender   | char(1)     | YES    |        | NULL    |          |
| Marks    | number(11)   | YES    |        | NULL    |          |
| DOB      | date       | YES    |        | NULL    |          |
+-----+-----+-----+-----+-----+-----+
```

5 rows in set (0.02 sec)

7. Inserting Data into a Table

The **INSERT INTO** command is used to insert a new record/row/tuple in a table.

It is possible to write the INSERT INTO statement in the following different forms:

- (a) **Inserting the data (for all the columns) into a table:** In the first method, it does not specify the column names where the data will be inserted, only their values.

Syntax for SQL INSERT is:

Syntax:

INSERT INTO *table_name*

VALUES (*value1, value2, value3...*);

For example, **INSERT INTO** student

VALUES (1,'Raj Kumar', 'M', 93, '2000-11-17');

While inserting a row, if we are adding value for all the columns of the table, we need not specify the column(s) name in the SQL query. But we need to make sure that the order of the values is in the same order as the columns represented in the structure of the table. The following points should be kept in mind while inserting data in a relation:

- When values are inputted using **INSERT INTO** command, it is termed as single row insert since it adds one tuple at a time into the table.
- The **INTO** clause specifies the target table and the **VALUES** clause specifies the data to be added to the new record of the table.
- The argument/values of character datatype are always enclosed in double or single quotation marks.
- Column values for the datatype of a column are provided within curly braces { } or single quotes.
- **NULL** values are stored and displayed as **NULL** only without any quotes.
- If the data is not available for all the columns, then the column-list must be included following the table name.

CTM: In SQL, we can repeat or re-execute the last command typed at SQL prompt by typing "/" key and pressing enter.

- (b) **Inserting the data directly into a table:** The second form specifies both the column names and the values to be inserted.

Syntax: **INSERT INTO** *table_name* (*column1, column2, columnN, ...*)

VALUES (*value1, value2, valueN, ...*);

Here, *column1, column2, ...columnN*—the names of the columns in the table for which you want to insert data.

For example, **INSERT INTO** student(**RollNo, Name, Gender, Marks, DOB**)

VALUES(2,'Deep Singh', 'M', 98, '1996-08-22');

CTM: When adding a row, only the characters or date values should be enclosed within single quotes.

(c) **Inserting data into specific columns of a table:**

Syntax for SQL INSERT is:

INSERT INTO <table_name>[(column1, column2, ... columnN)]

values [(value1, value2,... valueN)];

For example, to insert a record into the student table for the columns Rollno, Name and Marks only, the SQL insert query is:

INSERT INTO student (Rollno, Name, Marks) values (4,"Radhika Gupta",78);

The above statement shall insert the values for specific columns—Rollno , Name and Marks respectively.

- (d) **Inserting NULL values into a table:** If a column in a row has no value or missing value, then the column is said to be null or holding NULL value. Null value can be given to any column other than being assigned as primary key or Not Null constraint. It is advisable to use Null when the actual value is not defined or unavailable. NULL values are treated differently from other values as they represent missing unknown data. By default, a column in a table can hold NULL values.

If a column in a table is optional, we can insert a new record or can modify an existing tuple without adding values to this column. In other words, the values in every record for this column/field shall be stored as NULL. We can insert NULL value into any column in a table. It can be done by typing NULL without quotes.

Null is not equivalent to 0, i.e., $\text{NULL} \neq 0$. It acts as a placeholder for unknown or inapplicable values.

*For example, **INSERT INTO student(Rollno, Name, Gender, Marks, DOB)***

VALUES(12,'Swati Mehra', 'F', NULL, NULL);

After the execution of the above command, NULL value shall be inserted for the fields Marks and DOB respectively.

CTM: Null means unavailable or undefined value. Any arithmetic expression containing a NULL always evaluates to null.

8. Modifying Data in a Table

To modify data in a table or to make changes for some or all of the values in the existing records in a table, we use the UPDATE statement. The UPDATE command specifies the rows to be modified using the WHERE clause and the new data is written into the respective record using the SET keyword.

Syntax for UPDATE:

UPDATE <table_name>

SET <column1> = <value1>, <column2> = <value2>,....

WHERE <condition>;

For example, UPDATE student
 SET Marks = 90
 WHERE Rollno = 8;

The above statement shall change the value of Marks field to 90 for the student whose roll number is 8.

(a) Updating multiple columns

Modifying the values in more than one column can be done by separating the columns along with the new values using SET clause, separated by commas.

For example, UPDATE student
 SET Marks = 70, DOB='1998-08-11'
 WHERE Name="Payal";

The above statement shall change the values for both the fields Marks and DOB to 70 and '1998-08-11' for the student whose name is Payal.

(b) Updating to NULL values

The values for the attributes in a relation can also be entered as NULL using UPDATE command.

For example, UPDATE student
 SET Marks = NULL
 WHERE Rollno = 9;

The above statement shall change the value of the field Marks to 0 for the student whose roll number is 9.

(c) Updating using an expression or formula

For example, UPDATE student
 SET Marks = Marks + 10
 WHERE (Rollno = 5 or Rollno =10);

The above statement shall increment the value of Marks by 10 for all the records with Roll number 5 or 10.

9. Removing Data from a Table

The **DELETE** statement is used to delete rows from a table.

Syntax for DELETE Statement:

DELETE FROM <table_name> WHERE <condition>;

here <table_name> is the table whose records are to be deleted.

POINT TO REMEMBER

The WHERE clause in the SQL delete command is optional and it identifies the rows in the column that get deleted. If you do not include the WHERE clause, all the rows in the table are deleted.

For example, DELETE FROM student WHERE Rollno = 10;

The above statement shall delete the record only for roll number 10.

- To delete all the rows from the student table, the DELETE statement will be: DELETE FROM student;

➤ SQL TRUNCATE Statement

The **SQL TRUNCATE** command is used to delete all the rows from the table and free the space containing the table.

Syntax to TRUNCATE a table:

```
TRUNCATE TABLE<table_name>;
```

For example,

To delete all the rows from student table, the statement will be:

```
TRUNCATE TABLE student;
```

Difference between DELETE and TRUNCATE Statements

DELETE Statement: This command deletes only the rows from the table based on the condition given in the where clause or deletes all the rows from the table if no condition is specified. But it does not free the space containing the table.

TRUNCATE Statement: This command is used to delete all the rows from the table and free the space containing the table.

10. ALTER TABLE Command

The **ALTER TABLE** command is used to modify the definition (structure) of a table by modifying the definition of its columns. The ALTER TABLE command is used to perform the following operations:

- To add a column to an existing table.
- To rename any existing column.
- To change the datatype of any column or to modify its size.
- To remove or physically delete a column.

A. Adding a column to an existing table:

Once a table has been created, new columns can be added later on, if required. The new column is added with NULL values for all the records/rows in the table. It is possible to add, delete and modify columns with ALTER TABLE statement.

Syntax for adding a new column:

```
ALTER TABLE <table_name> ADD(<column_name><datatype> [size]);
```

For example, to add a new column Mobile_no of type integer in the table student:

```
ALTER TABLE student ADD (Mobile_no integer);
```

Thus, the above statement shall add a new column Mobile_no into the table student with NULL value in it.

POINT TO REMEMBER

We have just added a column and there will be no data (NULL) under this attribute. UPDATE command can be used to supply values/data to this column.

B. Adding a column with default value:

ALTER TABLE command can be used to add a new column to an existing table with default values.

Syntax for adding a column with a default value:

```
ALTER TABLE <table_name>
ADD ([column_name1]<datatype1>default data);
```

For example, ALTER TABLE student ADD(City char(6) DEFAULT "DELHI");

The above command will add a new column City with default value as "DELHI" to the student table.

Resultant table: student

Rollno	Name	Gender	Marks	DOB	Mobile_no	City
1.	Raj Kumar	M	93	17-Nov-2000	NULL	DELHI
2.	Deep Singh	M	98	22-Aug-1996	NULL	DELHI
3.	Ankit Sharma	M	76	02-Feb-2000	NULL	DELHI
4.	Radhika Gupta	F	78	03-Dec-1999	NULL	DELHI
5.	Payal Goel	F	82	21-April-1998	NULL	DELHI
6.	Diksha Sharma	F	80	17-Dec-1999	NULL	DELHI
7.	Gurpreet Kaur	F	65	04-Jan-2000	NULL	DELHI
8.	Akshay Dureja	M	90	05-May-1997	NULL	DELHI
9.	Shreya Anand	F	70	08-Oct-1999	NULL	DELHI
10.	Prateek Mittal	M	75	25-Dec-2000	NULL	DELHI

C. Modifying an existing column definition:

The MODIFY clause can be used with ALTER TABLE command to change the datatype, size, constraint related to any column of the table.

Syntax for modifying existing column datatype:

ALTER TABLE <table_name>

MODIFY([column_name1] <datatype1>);

For example,

ALTER TABLE student MODIFY (Name varchar(25));

The above command will modify the datatype size for the Name field from 20 to 25 characters.

D. Renaming a column:

The existing column in a relation can be renamed using ALTER TABLE command.

Syntax for renaming an existing column:

ALTER TABLE <table_name>

CHANGE [COLUMN] <old-column-name> <new-column-name> column_definition;

For example,

ALTER TABLE student

CHANGE City State varchar(10);

The above command shall rename the City column to State.

Rollno	Name	Gender	Marks	DOB	Mobile_no	City
1.	Raj Kumar	M	93	17-Nov-2000	NULL	DELHI
2.	Deep Singh	M	98	22-Aug-1996	NULL	DELHI
3.	Ankit Sharma	M	76	02-Feb-2000	NULL	DELHI
4.	Radhika Gupta	F	78	03-Dec-1999	NULL	DELHI
5.	Payal Goel	F	82	21-April-1998	NULL	DELHI
6.	Diksha Sharma	F	80	17-Dec-1999	NULL	DELHI
7.	Gurpreet Kaur	F	65	04-Jan-2000	NULL	DELHI
8.	Akshay Dureja	M	90	05-May-1997	NULL	DELHI
9.	Shreya Anand	F	70	08-Oct-1999	NULL	DELHI
10.	Prateek Mittal	M	75	25-Dec-2000	NULL	DELHI

E. Removing a column:

To remove or drop a column in a table, ALTER TABLE command is used.

Syntax for removing a column:

ALTER TABLE <table-name> DROP

<column-name>;

For example,

ALTER TABLE student DROP (State);

The above command will drop State column from the student table.

Resultant table: student

Rollno	Name	Gender	Marks	DOB	Mobile_no
1.	Raj Kumar	M	93	17-Nov-2000	NULL
2.	Deep Singh	M	98	22-Aug-1996	NULL
3.	Ankit Sharma	M	76	02-Feb-2000	NULL
4.	Radhika Gupta	F	78	03-Dec-1999	NULL
5.	Payal Goel	F	82	21-April-1998	NULL
6.	Diksha Sharma	F	80	17-Dec-1999	NULL
7.	Gurpreet Kaur	F	65	04-Jan-2000	NULL
8.	Akshay Dureja	M	90	05-May-1997	NULL
9.	Shreya Anand	F	70	08-Oct-1999	NULL
10.	Prateek Mittal	M	75	25-Dec-2000	NULL

11. DROP TABLE Command

Sometimes, we may need to physically remove a table which is not in use. **DROP TABLE** command is used to remove/delete a table permanently. It should be kept in mind that we cannot drop a table if it contains records. That is why all the rows of the table have to be deleted first and only then can the table be dropped. This command will completely destroy the table structure.

Syntax for removing a table:

DROP TABLE <table-name>;

For example,

mysql>DROP TABLE student;

This command will permanently remove the table student from the database school.

12.14 SQL QUERY PROCESSING

After creating the database, the table is created and the data is stored in it. Now, it is time to perform query processing on the already-created tables to retrieve and view the data on the screen. Retrieving information from the tables is done mainly using the SELECT command. The SQL SELECT statement is used to fetch data from one or more database tables. It is used to select rows and columns from a database/relation.

12.14.1 SQL SELECT Statement

This command can perform selection as well as projection. It is the most extensively used SQL command. The SELECT statement can be used to retrieve a subset of rows or columns from one or more tables present in a database.

1. Selection

This capability of SQL returns the tuples from a relation with all the attributes.

Syntax:

SELECT <column-name1> [, <column-name2>...]

FROM <table-name>;

**OR SELECT <what_to_select> FROM
<which_table>**

WHERE <conditions_to_satisfy>;

For example,

SELECT Name, Gender FROM student;

The above command displays only name and gender attributes from the student table.

2. Projection: Selecting Specific Rows—WHERE Clause

This is the capability of SQL to return only specific attributes from the relation. Use of WHERE clause is required when specific tuples are to be fetched or manipulated. To select all the columns from a table, the asterisk (*) can be used.

For example,

SELECT * FROM student;

The above command will display all the tuples (rows) from the relation student.

CTM: The asterisk (*) means “All”. SELECT * means displaying all the columns from a relation.

Resultant table: student

Rollno	Name	Gender	Marks	DOB	Mobile_no
1.	Raj Kumar	M	93	17-Nov-2000	NULL
2.	Deep Singh	M	98	22-Aug-1996	NULL
3.	Ankit Sharma	M	76	02-Feb-2000	NULL
4.	Radhika Gupta	F	78	03-Dec-1999	NULL
5.	Payal Goel	F	82	21-April-1998	NULL
6.	Diksha Sharma	F	80	17-Dec-1999	NULL
7.	Gurpreet Kaur	F	65	04-Jan-2000	NULL
8.	Akshay Dureja	M	90	05-May-1997	NULL
9.	Shreya Anand	F	70	08-Oct-1999	NULL
10.	Prateek Mittal	M	75	25-Dec-2000	NULL

10 rows in a set (0.02 sec)

The above command displays all the rows of all the columns according to the column-list defined in the table structure. The salient features of SQL SELECT statement are as follows:

- SELECT command displays the columns of the table in the same order in which they are selected from the table.
- In order to retrieve all the columns in the column-list from a table using SELECT command, asterisk (*) is used and the columns are displayed in the same order in which they are stored in the table.
- All the statements (inclusive of SELECT statement) in SQL are terminated with a semicolon (;). Use of semicolon is dependent on the version in use.

Using WHERE clause

```
SELECT * FROM student WHERE Rollno<=8;
```

The above command shall display only those records whose Rollno is less than or equal to 8.

Resultant table: student

Rollno	Name	Gender	Marks	DOB	Mobile_no
1.	Raj Kumar	M	93	17-Nov-2000	NULL
2.	Deep Singh	M	98	22-Aug-1996	NULL
3.	Ankit Sharma	M	76	02-Feb-2000	NULL
4.	Radhika Gupta	F	78	03-Dec-1999	NULL
5.	Payal Goel	F	82	21-April-1998	NULL
6.	Diksha Sharma	F	80	17-Dec-1999	NULL
7.	Gurpreet Kaur	F	65	04-Jan-2000	NULL
8.	Akshay Dureja	M	90	05-May-1997	NULL

8 rows in a set (0.02 sec)

When a WHERE clause is used with a SELECT statement, the SQL query processor goes through the entire table one row/record at a time and checks each row to determine whether the condition specified is true with respect to that row or not. If it evaluates to True, the corresponding row is selected, retrieved and displayed, else it returns an empty set (*i.e.*, no data found).

CTM: SQL is case-insensitive, which means keywords like SELECT and select have same meaning in SQL statements. On the contrary, MySQL makes difference in table names. So, if you are working with MySQL, then you need to give table names as they exist in the database.

3. Recording Columns while Displaying Query Results

While displaying the result for a query, the order of the columns to be displayed can be changed according to the user's requirement. But this is done only for the display purpose and no actual (physical) rearrangement of the columns is done.

For example,

```
SELECT Name, Rollno, DOB, Marks from student;
```

After executing the above statement, the column shall be displayed in the changed order as Name shall be displayed as the first column, Rollno as the second column, DOB as the third column and Marks as the fourth column respectively.

Resultant table: student

Name	Rollno	DOB	Marks
Raj Kumar	1.	17-Nov-2000	93
Deep Singh	2.	22-Aug-1996	98
Ankit Sharma	3.	02-Feb-2000	76
Radhika Gupta	4.	03-Dec-1999	78
Payal Goel	5.	21-April-1998	82
Diksha Sharma	6.	17-Dec-1999	80
Gurpreet Kaur	7.	04-Jan-2000	65
Akshay Dureja	8.	05-May-1997	90
Shreya Anand	9.	08-Oct-1999	70
Prateek Mittal	10.	25-Dec-2000	75

10 rows in a set (0.02 sec)

CTM: The order in which the columns are displayed using the SELECT command is in accordance with the order in which they are actually stored in the table.

2. Eliminating Duplicate/Redundant Data—DISTINCT clause

DISTINCT clause is used to remove duplicate rows from the results of a SELECT statement. It is used to retrieve only unique values for a column in the table. The DISTINCT keyword can be used only once with a given SELECT statement.

Syntax: `SELECT DISTINCT <column-name> from <table-name>;`

For example,

Suppose we have added a new column Stream to the table student:

Resultant table: student

Rollno	Name	Gender	Marks	DOB	Mobile_no	Stream
1.	Raj Kumar	M	93	17-Nov-2000	9586774748	Science
2.	Deep Singh	M	98	22-Aug-1996	8988886577	Commerce
3.	Ankit Sharma	M	76	02-Feb-2000	NULL	Science
4.	Radhika Gupta	F	78	03-Dec-1999	9818675444	Humanities
5.	Payal Goel	F	82	21-April-1998	9845639990	Vocational
6.	Diksha Sharma	F	80	17-Dec-1999	9897666650	Humanities
7.	Gurpreet Kaur	F	NULL	04-Jan-2000	7560567890	Science
8.	Akshay Dureja	M	90	05-May-1997	9560567890	Commerce
9.	Shreya Anand	F	70	08-Oct-1999	NULL	Vocational
10.	Prateek Mittal	M	75	25-Dec-2000	9999967543	Science

10 rows in a set (0.02 sec)

With reference to the above table, if we write the SELECT statement as:

`SELECT Stream from student;`

this statement shall return all the tuples for field Stream from table student. It will return duplicate values also. Thus, in order to remove these duplicate values, DISTINCT clause is used.

Now, we write a query for displaying the distinct contents on the basis of the field Stream from student table:

Resultant table: student

Stream
Science
Commerce
Science
Humanities
Vocational
Humanities
Science
Commerce
Vocational
Science

Science displayed 4 times

10 rows in a set (0.02 sec)

For example,

SELECT DISTINCT Stream from student;

Resultant table: student

Stream
Science
Commerce
Humanities
Vocational

Science displayed once only

4 rows in a set (0.02 sec)

12.14.2 SQL Operators

While working with SELECT statement using WHERE clause, condition-based query can be carried out using four types of SQL operators:

- (a) Arithmetic Operators
- (b) Relational Operators
- (c) Logical Operators
- (d) Special Operators

Table 12.3 SQL Operators and their Functions

OPERATOR/FUNCTION	DESCRIPTION
COMPARISON OPERATORS	
=, >, <, >=, <=, <>	Used in Conditional expressions.
LOGICAL OPERATORS	
AND/OR/NOT	Used in Conditional expressions.
SPECIAL OPERATORS	
BETWEEN	Checks whether an attribute value is within a range.
IS NULL	Checks whether an attribute value is null.
LIKE	Checks whether an attribute matches a given string pattern.
IN	Checks whether an attribute value matches any value with a given list.
DISTINCT	Permits only unique values. Eliminates duplicate ones.
AGGREGATE FUNCTIONS	Used with SELECT to return mathematical results/values on the basis of the operation performed on the columns.
COUNT	Returns the total number of records with non-null values for a given column.
MIN	Returns the minimum/lowest attribute value found in a given column.
MAX	Returns the maximum/highest attribute value found in a given column.
SUM	Returns the sum of all the values for a given column.
AVG	Returns the average of all the values for a given column.

(a) Arithmetic Operators

Arithmetic operators are used to perform simple arithmetic operations like addition (+), subtraction (-), multiplication (*), division (/) and modulus (%). These operators are used with conditional expressions and for performing simple mathematical calculations. The arithmetic operators with SELECT command are used to retrieve rows computed with or without reference to any table.

mysql> SELECT 5 + 10 FROM DUAL;

The above statement shall evaluate the expression 5 + 10 and returns the value 15 as the result.

```

+-----+
| 5 + 10   |
+-----+
| 15       |
+-----+
+-----+-----+
| SIN(PI()/4) | (4+1)*5   |      mysql> SELECT SIN(PI()/4), (4+1)*5;
+-----+-----+
| 0.707107  |    25    |
+-----+-----+
mysql> SELECT 5 * 4 FROM DUAL;
+-----+
| 5 * 4   |
+-----+
| 20      |
+-----+
mysql> SELECT 47 % 5 FROM DUAL;
+-----+
| 47 % 5   |
+-----+
| 2        |
+-----+

```

The modulus (%) operator returns the remainder as the answer after performing the division operation. Hence, the above statement, 47 % 5 shall return the value 2 as the output.

POINT TO REMEMBER

In the above statement, DUAL is the default table in MySQL. It is a one-row, one-column dummy table.

Evaluating Scalar expression with SELECT statement

MySQL permits calculations on the contents of the columns and then displays the calculated result using SELECT statement. We can write scalar expression and constant values for the selected columns. If we are taking NULL value in the expression, it shall result in NULL only. Along with NULL, arithmetic operators can be used while evaluating scalar expressions.

For example, mysql> SELECT Rollno, Name, Marks + 10 FROM student;



The above command, on execution, shall increment the value for all the rows of the field Marks by 10 and shall display the Rollno, Name and Marks for all the students, increased by 10.

Resultant table: student

Rollno	Name	Marks + 10
1.	Raj Kumar	103
2.	Deep Singh	108
3.	Ankit Sharma	86
4.	Radhika Gupta	88
5.	Payal Goel	92
6.	Diksha Sharma	90
7.	Gurpreet Kaur	75
8.	Akshay Dureja	100
9.	Shreya Anand	80
10.	Prateek Mittal	85

10 rows in a set (0.02 sec)

(b) Relational Operators

A relational (comparison) operator is a mathematical symbol which is used to compare two values. It is used to compare two values of the same or compatible data types. Comparison operators are used for conditions where two expressions are required to be compared with each other, which results in either true or false. They are used with WHERE clause.

The following table describes different types of comparison operators in SQL:

OPERATOR	DESCRIPTION
=	Equal to
>	Greater than
<	Less than
>=	Greater than or equal to
<=	Less than or equal to
<>, !=	Not equal to

For comparing character data type values, < means earlier in the alphabetical sequence and > means later in the alphabetical sequence.

For example, **mysql> SELECT Rollno, Name, Marks FROM student**

WHERE Marks>=90;

The above command shall display the Rollno, Name and Marks of all the students with marks either equal to or greater than 90.

Resultant table: student

Rollno	Name	Marks
1.	Raj Kumar	93
2.	Deep Singh	98
3.	Akshay Dureja	90

3 rows in a set (0.02 sec)

For example, `mysql> SELECT * FROM student
WHERE Stream <> 'Commerce';`

The above command shall display the records of all the students who are not from Commerce stream.

Resultant table: student

Rollno	Name	Gender	Marks	DOB	Mobile_no	Stream
1.	Raj Kumar	M	93	17-Nov-2000	9586774748	Science
3.	Ankit Sharma	M	76	02-Feb-2000	8567490078	Science
4.	Radhika Gupta	F	78	03-Dec-1999	9818675444	Humanities
5.	Payal Goel	F	82	21-April-1998	9845639990	Vocational
6.	Diksha Sharma	F	80	17-Dec-1999	9897666650	Humanities
7.	Gurpreet Kaur	F	65	04-Jan-2000	7560567890	Science
9.	Shreya Anand	F	70	08-Oct-1999	8876543988	Vocational
10.	Prateek Mittal	M	75	25-Dec-2000	9999967543	Science

8 rows in a set (0.02 sec)

Thus, while using relational operators in a WHERE clause with a SELECT statement, the database program goes through the entire table checking each record one by one and compares with the condition specified. If it is true, the corresponding row is selected for display, otherwise ignored.

CTM: While comparing character, date and time data using relational operators, it should be enclosed in single quotation marks.

(c) Logical Operators

The SQL logical operators are the operators used to combine multiple conditions to narrow the data selected and displayed on the basis of the condition specified in an SQL statement. Logical operators are also known as Boolean operators. The three logical operators in SQL are—AND, OR and NOT operator. Out of these, AND and OR operators are termed as Conjunctive operators since these two operators combine two or more conditions. The AND and OR operators are used to filter records based on more than one condition.

These operators provide a means to make multiple comparisons with different operators in the same SQL statement.

CTM: The order of precedence for logical operators (AND, OR, NOT operator) is NOT(!), AND(&&) and OR(||).

1. AND operator

The AND operator displays a record and returns a true value if all the conditions (usually two conditions) specified in the WHERE clause are true.

Condition 1	Condition 2	Result (AND operation)
True	True	True
True	False	False
False	True	False
False	False	False

As shown in the table, when both condition 1 and condition 2 are true, then only is the result true. If either of them is false, the result becomes false.

For example, to list the details of all the students who have secured more than 80 marks and are male.

```
mysql> SELECT * FROM student
      WHERE Marks > 80 and Gender= 'M';
```

Resultant table: student

Rollno	Name	Gender	Marks	DOB	Mobile_no	Stream
1.	Raj Kumar	M	93	17-Nov-2000	9586774748	Science
2.	Deep Singh	M	98	22-Aug-2000	8988886577	Commerce
8.	Akshay Dureja	M	90	05-May-1997	9560567890	Commerce

3 rows in a set (0.02 sec)

2. OR operator

The OR operator displays a record and returns a true value if either of the conditions (usually two conditions) specified in the WHERE clause is true.

Condition 1	Condition 2	Result (OR operation)
True	True	True
True	False	True
False	True	True
False	False	False

As shown in the table, when either condition 1 or condition 2 is true, the result is true. If both of them are false, then only the result becomes false.

For example, to display the roll number, name and stream of all the students who are in either Science or Commerce stream.

```
mysql> SELECT Rollno, Name, Stream FROM student
      WHERE Stream= 'Science' or Stream= 'Commerce';
```

Resultant table: student

Rollno	Name	Stream
1.	Raj Kumar	Science
2.	Deep Singh	Commerce
3.	Ankit Sharma	Science
7.	Gurpreet Kaur	Science
8.	Akshay Dureja	Commerce
10.	Prateek Mittal	Science

6 rows in a set (0.02 sec)

3. NOT operator

NOT operator is also termed as a negation operator. Unlike the other two operators, this operator takes only one condition and gives the reverse of it as the result. It returns a false value if the condition holds true and vice versa.

The NOT operator displays a record and returns a true value if either of the conditions (usually two conditions) specified in the WHERE clause is true.

Condition 1	Result (NOT operation)
True	False
False	True

As shown in the table, when the condition is true, the result is false. If the condition is false, then the result becomes true.

For example, to display the name and marks of all the students who are not in the vocational stream.

```
mysql> SELECT Name, Marks FROM student  
      WHERE NOT (Stream = 'Vocational');
```

Resultant table: student

Name	Marks
Raj Kumar	93
Deep Singh	98
Ankit Sharma	76
Radhika Gupta	78
Diksha Sharma	80
Gurpreet Kaur	NULL
Akshay Dureja	90
Prateek Mittal	75

8 rows in a set (0.02 sec)

12.14.3 Comments in SQL

A comment is a text which is ignored by the SQL compiler and is not executed at all. It is given for documentation purpose only. A comment usually describes the purpose of the statement given within an application.

SQL Comments are used to understand the functionality of the program without looking into it. The comments give us an idea about what is written in the given SQL statement and how it works. The comments can make an application or code easier to read as well as to maintain. A comment can be placed between any keywords, parameters or punctuation marks in a statement. Comments can be either single-line comments or multiple-line comments.

SQL or MySQL supports three comment styles:

- ☞ **Comments beginning with - (followed by a space):** The two dash lines indicate a single-line comment in SQL statements. These single-line comments are basically used to show the comments at the start and end of a program. A user can easily use this comment type to explain the flow of program. This text cannot extend to a new line and ends with a line break.
- ☞ **Comments beginning with #:** The comments begin with '#' symbol followed by the text to be displayed for the user's information.. This text cannot extend to a new line and ends with a line break.
- ☞ **Comments beginning with /*:** Multi-line comments begin with a slash and an asterisk /*) followed by the text of the comment. This text can span multiple lines. The comment ends with an asterisk and a slash (*). The opening and terminating characters need not be separated from the text by a space or a line break.

For example,

```
SELECT Rollno, Name, Stream
```

```
/* This statement shall display the records of all those students who are in Science stream and have  
secured marks more than 75. */
```

```
FROM student # student table in use
```

```
WHERE Stream= 'Science' and Marks > 75; --condition for projection
```

12.15 SQL ALIASES

SQL aliases are used to give an alternate name, *i.e.*, a temporary name, to a database table or a column in a table. We can rename a table or a column temporarily by giving another name called alias name which leads to a temporary change (renaming) and does not change the actual name in the database.

Aliases can be used when

- more than one table is involved in a query.
- functions are used in the query.
- column names are big or not very readable.
- two or more columns are combined together.

Using column alias name, we can give different name(s) to column(s) for display (output) purpose only. They are created to make column names more readable. SQL aliases can be used both for tables as well as columns.

- **COLUMN ALIASES** are used to make column headings in the query result set easier to read.
- **TABLE ALIASES** are used to shorten a table name by giving an easy alternate name, making it easier to read or when performing a self-join (*i.e.*, listing the same table more than once in the FROM clause).

Syntax for table alias:

```
SELECT <columnname1>, <columnname2>....  
      FROM <table_name> AS <alias_name>;  
      WHERE [<condition>];
```

Syntax for column alias:

```
SELECT <column-name> AS <"alias_name">  
      FROM <table_name>  
      WHERE [<condition>];
```

For example,

Table: Student

Student_name	Date-of-birth
Raj Kumar	17-Nov-2000
Deep Singh	22-Aug-1996
Ankit Sharma	02-Feb-2000
Radhika Gupta	03-Dec-1999
Payal Goel	21-April-1998
Diksha Sharma	17-Dec-1999
Gurpreet Kaur	04-Jan-2000
Akshay Dureja	05-May-1997
Shreya Anand	08-Oct-1999
Prateek Mittal	25-Dec-2000

```
SELECT Name AS "Student_name", DOB AS "Date_of_birth"  
      FROM student;
```

→ Alias name for fields, Name and DOB

10 rows in a set (0.02 sec)

POINTS TO REMEMBER

- If the **alias_name** contains spaces, you must enclose it in quotes.
 - It is acceptable to use spaces when you are aliasing a column name. However, it is not generally a good practice to use spaces when you are aliasing a table name.
 - The **alias_name** is only valid within the scope of the SQL statement.

- Alias name can be given to a mathematical expression also:

For example,

SELECT 22/7 as PI;

Output: +-----+
| PI |
+-----+
| 3.1429 |
+-----+



12.16 PUTTING TEXT IN THE QUERY OUTPUT

In order to get an organized output from a SELECT query, we can include some user-defined columns at runtime. These columns are displayed with a valid text, symbols and comments in the output only. These included columns will appear as column heads along with the contents for that column.

This makes the query output more presentable by giving a formatted output.

For example, SELECT Rollno, Name, ‘was born on’,DOB
 FROM student;

The above command, on execution, shall display the text “was born on” with every record (tuple) of the table.

Resultant table: student

Rollno	Name	was born on	DOB
1.	Raj Kumar	was born on	17-Nov-2000
2.	Deep Singh	was born on	22-Aug-1996
3.	Ankit Sharma	was born on	02-Feb-2000
4.	Radhika Gupta	was born on	03-Dec-1999
5.	Payal Goel	was born on	21-April-1998
6.	Diksha Sharma	was born on	17-Dec-1999
7.	Gurpreet Kaur	was born on	04-Jan-2000
8.	Akshay Dureja	was born on	05-May-1997
9.	Shreya Anand	was born on	08-Oct-1999
10.	Prateek Mittal	was born on	25-Dec-2000

10 rows in a set (0.02 sec)

12.17 SQL SPECIAL OPERATORS

Apart from several standard library functions discussed earlier, there are some special operators in SQL that perform some specific functions.

12.17.1 Conditions Based on a Range—BETWEEN...AND

SQL provides a BETWEEN operator that defines a range of values that the column value must fall within for the condition to become true. The range includes both the lower and upper value. The values can be numbers, text or dates.

Syntax for BETWEEN:

```
mysql>SELECT <column_name(s)>
   FROM <table_name>
 WHERE <column_name> BETWEEN <value1> AND <value2>;
```

For example,

```
mysql> SELECT Rollno, Name, Marks FROM student WHERE Marks BETWEEN 80 AND 100;
```

The above command displays Rollno, Name along with Marks of those students whose Marks lie in the range of 80 to 100 (both 80 and 100 are included in the range).

Resultant table: student

Rollno	Name	Marks
1.	Raj Kumar	93
2.	Deep Singh	98
5.	Payal Goel	82
6.	Diksha Sharma	80
8.	Akshay Dureja	90

5 rows in a set (0.02 sec) NOT BETWEEN

The NOT BETWEEN operator works opposite to the BETWEEN operator. It retrieves the rows which do not satisfy the BETWEEN condition.

For example,

```
mysql> SELECT Rollno, Name, Marks FROM student WHERE Marks NOT BETWEEN 80 AND 100;
```

Resultant table: student

Rollno	Name	Marks
3.	Ankit Sharma	76
4.	Radhika Gupta	78
9.	Shreya Anand	70
10.	Prateek Mittal	75

4 rows in a set (0.02 sec)

12.17.2 Conditions Based on a List—IN

To specify a list of values, IN operator is used. This operator selects values that match any value in the given list. The SQL IN condition is used to help reduce the need for multiple OR conditions in a SELECT statement.

Syntax for IN:

```
SELECT <column_name(s)>
   FROM <table_name>
 WHERE <column_name> IN (<value1,value2,...>);
```

For example,

mysql> SELECT * FROM student WHERE Stream IN ('Science', 'Commerce', 'Humanities');

Resultant table: student

Rollno	Name	Gender	Marks	DOB	Mobile_no	Stream
1.	Raj Kumar	M	93	17-Nov-2000	9586774748	Science
2.	Deep Singh	M	98	22-Aug-1996	8988886577	Commerce
3.	Ankit Sharma	M	76	02-Feb-2000	NULL	Science
4.	Radhika Gupta	F	78	03-Dec-1999	9818675444	Humanities
6.	Diksha Sharma	F	80	17-Dec-1999	9897666650	Humanities
7.	Gurpreet Kaur	F	NULL	04-Jan-2000	7560875609	Science
8.	Akshay Dureja	M	90	05-May-1997	9560567890	Commerce
10.	Prateek Mittal	M	75	25-Dec-2000	9999967543	Science

8 rows in a set (0.02 sec)

The above command displays all those records whose Stream is either Science or Commerce or Humanities.

NOT IN

The NOT IN operator works opposite to IN operator. It matches, finds and returns the rows that do not match the list.

For example,

mysql> SELECT * FROM student WHERE Stream NOT IN ('Science', 'Commerce', 'Humanities');

Resultant table: student

Rollno	Name	Gender	Marks	DOB	Mobile_no	Stream
5.	Payal Goel	F	82	21-April-1998	9845639990	Vocational
9.	Shreya Anand	F	70	08-Oct-1999	NULL	Vocational

2 rows in a set (0.02 sec)

12.17.3 Conditions Based on Pattern—LIKE

The LIKE operator is used to search for a specified pattern in a column. This operator is used with the columns of type CHAR. The LIKE operator searches the column to find if a part of this column matches the string specified in the parentheses after the LIKE operator in the command.

Conditions Based on Pattern—WILD CARD CHARACTERS

The SQL LIKE condition allows you to use wild cards to perform pattern matching. SQL provides two wild card characters that are used while comparing the strings with LIKE operator:

- a. **Percent(%)** Matches any string
- b. **Underscore(_)** Matches any one character

Syntax for LIKE:

SELECT <column_name(s)>

FROM <table_name>

WHERE <column_name> LIKE <pattern>;

For example,

- `mysql> SELECT * FROM student WHERE Name LIKE "D%";`

Resultant table: student

Rollno	Name	Gender	Marks	DOB	Mobile_no	Stream
2.	Deep Singh	M	98	22-Aug-1996	8988886577	Commerce
6.	Diksha Sharma	F	80	17-Dec-1999	9897666650	Humanities

2 rows in a set (0.02 sec)

The above command shall display those records where the name begins with character 'D'.

- `mysql> SELECT * FROM student WHERE Name LIKE "%a";`

Resultant table: student

Rollno	Name	Gender	Marks	DOB	Mobile_no	Stream
3.	Ankit Sharma	M	76	02-Feb-2000	NULL	Science
4.	Radhika Gupta	F	78	03-Dec-1999	9818675444	Humanities
6.	Diksha Sharma	F	80	17-Dec-1999	9897666650	Humanities
8.	Akshay Dureja	M	90	05-May-1997	9560567890	Commerce

4 rows in a set (0.02 sec)

The above command shall display the records for those students whose name ends with the letter 'a'.

- `mysql> SELECT * FROM student WHERE Name LIKE "%e%";`

Resultant table: student

Rollno	Name	Gender	Marks	DOB	Mobile_no	Stream
2.	Deep Singh	M	98	22-Aug-1996	8988886577	Commerce
5.	Payal Goel	F	82	21-April-1998	9845639990	Vocational
7.	Gurpreet Kaur	F	NULL	04-Jan-2000	7560875609	Science
8.	Akshay Dureja	M	90	05-May-1997	9560567890	Commerce
9.	Shreya Anand	F	70	08-Oct-1999	NULL	Vocational
10.	Prateek Mittal	M	75	25-Dec-2000	9999967543	Science

6 rows in a set (0.02 sec)

As the resultant table shows, the above command displays the records of all the students whose Name contains the character 'e' in it.

- `mysql> SELECT * FROM student WHERE Name LIKE "_e%";`

Resultant table: student

Rollno	Name	DOB
2.	Deep Singh	22-Aug-1996

1 row in a set (0.02 sec)

This command shall display the Rollno, Name and DOB of all the students whose Name contains the letter 'e' at the second place.

- `mysql> SELECT Rollno, Name, Marks, DOB FROM student WHERE Name LIKE "%r_";`

Resultant table: student

Rollno	Name	Marks	DOB
3.	Ankit Sharma	76	02-Feb-2000
6.	Diksha Sharma	80	17-Dec-1999

2 rows in a set (0.02 sec)

This command shall display the Rollno, Name, Marks and DOB of all the students whose Name contains the letter 'r' from the third last position.

12.18 SORTING IN SQL—ORDER BY

The SQL **ORDER BY** clause is used to sort the data in ascending or descending order based on one or more columns. The ORDER BY keyword is used to sort the result-set by one or more fields in a table. This clause sorts the records in the ascending order (ASC) by default. Therefore, in order to sort the records in descending order, DESC keyword is to be used. Sorting using ORDER BY clause can be done on multiple columns, separated by comma.

Syntax for ORDER BY clause:

```
SELECT <column-list> FROM <table_name> [WHERE  
<condition>] ORDER BY <column_name> [ASC|DESC];
```

Here, WHERE clause is optional.

For example,

- To display the roll number, name and marks of students on the basis of their marks in the ascending order.
- **mysql>** SELECT Rollno, Name, Marks FROM student ORDER BY Name;
- To display the roll number, name and marks of all the students in the descending order of their marks and ascending order of their names.
- **mysql>** SELECT Rollno, Name, Marks FROM student ORDER BY Marks DESC, Name;

Rollno	Name	Marks
8.	Akshay Dureja	90
3.	Ankit Sharma	76
2.	Deep Singh	98
6.	Diksha Sharma	80
7.	Gurpreet Kaur	NULL
5.	Payal Goel	82
10.	Prateek Mittal	75
4.	Radhika Gupta	78
1.	Raj Kumar	93
9.	Shreya Anand	70

Sorting on Column Alias

If a column alias is defined for a column, it can be used for displaying rows in ascending or descending order using ORDER BY clause.

For example, SELECT Rollno, Name, Marks AS Marks_obtained

FROM student

ORDER BY Marks_obtained;

Alias name

AGGREGATE FUNCTIONS

Till now, we have studied about single-row functions which work on a single value. SQL also provides multiple-row functions which work on multiple values. So, we can apply SELECT query on a group of records rather than the entire table. Therefore, these functions are called Aggregate functions or Group functions.

Generally, the following aggregate functions are applied on groups as described below:

Table 12.4 Aggregate Functions in SQL

S.No.	Function	Description/Purpose
1.	MAX()	Returns the maximum/highest value among the values in the given column/expression.
2.	MIN()	Returns the minimum/lowest value among the values in the given column/expression.
3.	SUM()	Returns the sum of the values under the specified column/expression.
4.	AVG()	Returns the average of the values under the specified column/expression.
5.	COUNT()	Returns the total number of values/records under the specified column/expression.

Consider a table Employee (employee code, employee name, salary, job and city) with the following structure:

Ecode	Name	Salary	Job	City
E1	Ritu Jain	5000	Manager	Delhi
E2	Vikas Verma	4500	Executive	Jaipur
E3	Rajat Chaudhary	6000	Clerk	Kanpur
E4	Leena Arora	7200	Manager	Bangalore
E5	Shikha Sharma	8000	Accountant	Kanpur

- **MAX()**

MAX() function is used to find the highest value among the given set of values of any column or expression based on the column. MAX() takes one argument which can be either a column name or any valid expression involving a particular column from the table.

For example,

```
mysql> Select MAX(Salary) from EMPLOYEE;
```

Output:

```
+-----+
| MAX(Salary) |
+-----+
| 8000      |
+-----+
```

This command, on execution, shall return the maximum value from the specified column (Salary) of the table Employee, which is 8000.

- **MIN()**

MIN() function is used to find the lowest value among the given set of values of any column or expression based on the column. MIN() takes one argument which can be either a column name or any valid expression involving a particular column from the table.

For example,

```
mysql> Select MIN(Salary) from EMPLOYEE;
```

Output:

```
+-----+
| MIN(Salary) |
+-----+
| 4500      |
+-----+
```

This command, on execution, shall return the minimum value from the specified column (Salary) of the table Employee, which is 4500

• **SUM()**

SUM() function is used to find the total value of any column or expression based on a column. It accepts the entire range of values as an argument, which is to be summed up on the basis of a particular column, or an expression containing that column name. The SUM() function always takes argument of integer type only. Sums of String and Date type data are not defined.

For example,

```
mysql> Select SUM(Salary) from EMPLOYEE;
```

Output:

```
+-----+
| SUM(Salary) |
+-----+
| 30700      |
+-----+
```

This command, on execution, shall return the total of the salaries of all the employees from the specified column (Salary) of the table Employee, which is 30700.

• **AVG()**

AVG() function is used to find the average value of any column or expression based on a column. Like sum(), it also accepts the entire range of values of a particular column to be taken average of, or even a valid expression based on this column name. Like SUM() function, the AVG() function always takes argument of integer type only. Average of String and Date type data is not defined.

For example,

```
mysql> Select AVG(Salary) from EMPLOYEE;
```

Output:

```
+-----+
| AVG(Salary) |
+-----+
| 6140       |
+-----+
```

Ecode	Name	Salary	Job	City
E1	Ritu Jain	NULL	Manager	Delhi
E2	Vikas Verma	4500	Executive	Jaipur
E3	Rajat Chaudhary	6000	Clerk	Kanpur
E4	Leena Arora	NULL	Manager	Bangalore
E5	Shikha Sharma	8000	Accountant	Kanpur

This command, on execution, shall return the average of the salaries of all the employees from the specified column (Salary) of the table Employee, which is 6140.

• **COUNT()**

COUNT() function is used to count the number of values in a column. COUNT() takes one argument, which can be any column name, or an expression based on a column, or an asterisk (*). When the argument is a column name or an expression based on the column, COUNT() returns the number of non-NULL values in that column. If the argument is asterisk (*), then COUNT() counts the total number of records/rows satisfying the condition, if any, in the table.

For example,

- mysql> Select COUNT(*) from EMPLOYEE;

Output:

```
+-----+
| COUNT(*) |
+-----+
| 5         |
+-----+
```

This command, on execution, shall return the total number of records in the table Employee, which is 5.

- mysql> Select COUNT(DISTINCT City) from EMPLOYEE;

Output:

```
+-----+
| COUNT(DISTINCT City) |
+-----+
| 4                   |
+-----+
```

This command, on execution, shall return the total number of records on the basis of city with no duplicate values, *i.e.*, Kanpur is counted only once; the second occurrence is ignored by MySQL because of the DISTINCT clause. Thus, the output will be 4 instead of 5.

- **Aggregate Functions & NULL Values**

Consider the following table Employee with NULL values against the Salary field for some employees:

None of the aggregate functions takes NULL into consideration. NULL values are simply ignored by all the aggregate functions as clearly shown in the examples given below:

mysql> Select Sum(Salary) from Employee;

Output: 18500

mysql> Select Min(Salary) from Employee;

Output: 4500 (NULL values are not considered.)

mysql> Select Max(Salary) from Employee;

Output: 8000 (NULL values are not ignored.)

mysql> Select Count(Salary) from Employee;

Output: 3 (NULL values are not ignored.)

mysql> Select Avg(Salary) from Employee;

Output: 6166.66 (It will be calculated as 18500/3, *i.e.*, sum/total no. of records, which are 3 after ignoring NULL values.)

mysql> Select Count(*) from Employee;

Output: 5

mysql> Select Count(Ecode) from Employee;

Output: 5 (No NULL value exists in the column Ecode.)

mysql> Select Count(Salary) from Employee;

Output: 3 (NULL values are ignored while counting the total number of records on the basis of Salary.)

a. SQL JOINS

An SQL JOIN clause is used to combine rows from two or more tables, based on a common field between them. While querying for a join, more than one table is considered in FROM clause. The process/function of combining data from multiple tables is called a JOIN. SQL can extract data from two or even more than two related tables by performing either a physical or virtual join on the tables using WHERE clause. The types of SQL joins are as follows:

1. Cartesian Product (Cross Product)
2. Equi Join
3. Self Join
4. Non-Equi Join
5. Natural Join

12.18.1 Cartesian Product (Cross Product)

The Cartesian product is also termed as cross product or cross-join. The Cartesian product is a binary operation and is denoted by (X) . The degree of the new relation formed is the sum of the degrees of two relations on which Cartesian product is performed. The number of tuples in the new relation is equal to the product of the number of tuples of the two tables on which Cartesian product is performed.

For example,

If $A = \{1, 2, 3\}$ and $B = \{a, b, c\}$, find $A X B$.

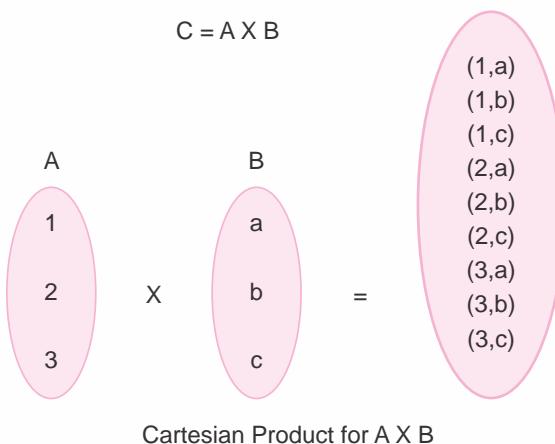


Table: student

Rollno	Name
1	Rohan
2	Jaya
3	Teena

Table: games

gameno	gname
10	Football
11	Lawn tennis

Cartesian product for student X games:

```
mysql> SELECT Name, gname FROM student, games;
```

Therefore, a Cartesian product is formed when no join conditions exist or are invalid. When we perform Cartesian product between two tables, all the rows in the first table are joined to all the rows in the second table. Using Cartesian product operation results in a large number of rows as the output, so it is seldom used.

student X games

Rollno	Name
Rohan	Football
Jaya	Football
Teena	Football
Rohan	Lawn Tennis
Jaya	Lawn Tennis
Teena	Lawn Tennis

Equi Join

An Equi join is a simple SQL join condition that uses the equal to sign (=) as a comparison operator for defining a relationship between two tables on the basis of a common field.

Syntax for Equi Join:

```
SELECT <column1>, <column2>,....  
FROM <table1>, <table2>  
WHERE <table1.column1> =  
<table2.column2>;
```

For example,

Table: student

Rollno	Name
1	Rohan
2	Jaya
3	Teena
4	Diksha

Table: fees

Rollno	Fee
4	4500
2	5500
3	5000

Resultant Table

Rollno	Name	Fee
4	Diksha	4500
2	Jaya	5500
3	Teena	5000

In the given SELECT statement, A and B are the alias names.

1. **Inner Join:** The Inner Join is a classification of equi join where either of the equivalent queries gives the intersection of two tables, i.e., it returns the rows which are common in both the tables.

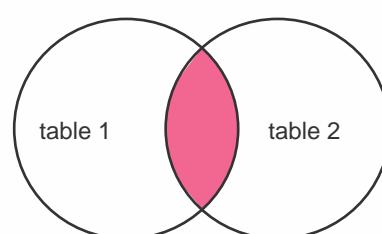
```
mysql> SELECT student.Rollno, Name, fee FROM student, fees
```

WHERE student.Rollno = fees.Rollno ORDER BY student.Rollno;

OR

```
mysql> SELECT A.Rollno, Name, fee  
FROM student A, fees B WHERE A.Rollno =  
B.Rollno ORDER BY A.Rollno;
```

INNER JOIN



Thus, the output for the above command will be:

Resultant Table

Rollno	Name	Fee
2	Jaya	5500
3	Teena	5000
4	Diksha	4500

2. **Outer Join:** The Outer Join keyword returns all rows from the right table (table 2), with the matching rows in the left table (table 1). The result is NULL in the left side when there is no match.

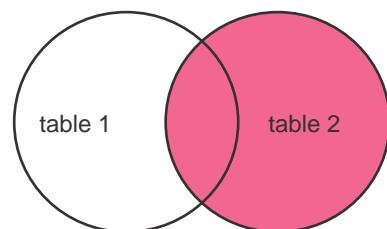
```
mysql> SELECT A.Rollno, Name, fee
   FROM student A, fees B WHERE
A.Rollno = B.Rollno ORDER BY B.fee desc;
```

Resultant Table

Rollno	Name	Fee
2	Jaya	5500
3	Teena	5000
4	Diksha	4500



OUTER JOIN



Note: Self, Non-equi and Natural join are beyond the scope of this book.

Let us take an example of two tables to explain the concept of MySQL Joins: (Question taken from CBSE 2015-16 Board Paper)

Table: Vehicle

CODE	VTYPE	PERKM
101	VOLVO BUS	160
102	AC DELUXE BUS	150
103	ORDINARY BUS	90
105	SUV	40
104	CAR	20

Note:

- PERKM is Freight Charges per kilometre
- VTYPE is Vehicle Type

Table: Travel

NO	NAME	TDATE	KM	CODE	NOP
101	Janish Kin	2015-11-13	200	101	32
103	Vedika Sahai	2016-04-21	100	103	45
105	Tarun Ram	2016-03-23	350	102	42
102	John Fen	2016-02-13	90	102	40
107	Ahmed Khan	2015-01-10	75	104	2
104	Raveena	2016-05-28	80	105	4

Note:

- NO is Traveller Number
 - KM is Kilometres travelled
 - NOP is number of travellers in vehicle
 - TDATE is Travel Date
- (a) To display NO, NAME, TDATE from the table TRAVEL in descending order of NO.

Ans. SELECT NO, NAME, TDATE FROM TRAVEL ORDER BY NO DESC;

- (b) To display the NAME of all the travellers from the table TRAVEL who are travelling by vehicle with code 101 or 102.

Ans. SELECT NAME FROM TRAVEL
WHERE CODE='101' OR
CODE='102'; OR
SELECT NAME FROM TRAVEL
WHERE CODE=101 OR
CODE=102;

- (c) To display the NO and NAME of those travellers from the table TRAVEL who travelled between '2015-12-31' and '2016-04-01'.

Ans. SELECT NO, NAME from TRAVEL
WHERE TDATE >= '20160401' AND TDATE <= '20151231';
OR
SELECT NO, NAME from TRAVEL
WHERE TDATE BETWEEN '20160401' AND '20151231';

- (d) To display the CODE, NAME, VTYPE from both the tables with distance travelled (Km) less than 90Km.

Ans. SELECT A.CODE, NAME,
VTYPE FROM TRAVEL A,
VEHICLE B
WHERE A.CODE=B.CODE AND KM<90;

JOIN Operation using
Table alias names

- (e) To display the NAME and amount to be paid for vehicle code as 105. Amount to be paid is calculated as the product of KM and PERKM.

Ans. SELECT NAME,
KM*PERKM FROM TRAVEL A,
VEHICLE B
WHERE A.CODE=B.CODE AND A.CODE='105';

12.19 UNION

The UNION operator is used to combine the result-set of two or more SELECT statements. The UNION operation is used to return all the distinct rows selected by either query. For executing Union between two tables, the number of columns selected from each table should be the same. Also, the datatypes of the corresponding columns selected from each table should be the same.

Table: boys

Rollno	Name
1	Rohan
12	Jayant
3	Tinku

Table: girls

Rollno	Name
6	Reema
10	Jaya

```
mysql> SELECT Name FROM boys WHERE
Rollno < 12 UNION
SELECT Name FROM girls WHERE Rollno > 6;
```

The resultant table is:

Name
Rohan
Tinku
Jaya

```
mysql> SELECT Rollno,
Name FROM boys
UNION
SELECT Rollno, Name FROM girls ORDER BY girls.name;
```

The resultant table is:

Rollno	Name
10	Jaya
12	Jayant
6	Reema
1	Rohan
3	Tinku



MEMORY BYTES

- SQL is a language that is used to create, modify and access a database.
- The various processing capabilities of SQL are Data Definition Language (DDL), Data Manipulation Language (DML) and Data Control Language (DCL).
- DDL is used to create and delete tables, views or indexes.
- DML is used to modify and update the database.
- DESCRIBE or DESC is used to show the structure of a table.
- The SELECT statement is used to fetch data from one or more database tables.
- SELECT * means display all columns.
- The WHERE clause is used to select specific rows.
- The DESCRIBE statement is used to see the structure of a table.
- We can change the structure of a table, i.e., add, remove or change its column(s) using the ALTER TABLE statement.
- The keyword DISTINCT is used to eliminate redundant data from display.
- Logical operators OR and AND are used to connect relational expressions in WHERE clause.
- Logical operator NOT is used to negate a condition.

- The BETWEEN operator defines the range of values that the column values must fall within to make the condition true.
- The IN operator selects values that match any value in the given list of values.
- % and _ are two wild card characters. The percent (%) symbol is used to represent any sequence of zero or more characters. The underscore (_) symbol is used to represent a single character.
- NULL represents a value that is unavailable, unassigned, unknown or inapplicable.
- The results of the SELECT statement can be displayed in the ascending or descending order of a single column or columns using ORDER BY clause.
- DROP DATABASE drops all tables in the database and deletes the database. Once the DROP command is used, then we cannot use that database. So, we should be careful with this command.
- The CREATE statement is used to create a table in MySQL with constraint. A constraint is a restriction on the behaviour of a variable.
- INSERT query is used for inserting new rows or data into an existing table.
- The ORDER BY keyword in MySQL is used to sort the result-set in ascending or descending order.
- The ORDER BY keyword sorts the records in ascending order by default. To sort the records in descending order, use the DESC keyword.
- An aggregate function is a function where the values of multiple rows are grouped together as input based on certain criteria to form a single value of more significant meaning.
- The COUNT() function returns the number of rows that matches a specified criteria.
- The AVG() function returns the average value of a numeric column.
- The SUM() function returns the total sum of a numeric column.

OBJECTIVE TYPE QUESTIONS

1. Fill in the blanks.

- is a freely-available open-source RDBMS that implements SQL.
- MySQL provides a dummy table named
- The keyword eliminates duplicate records from the results of a select statement.
- Patterns in MySQL are described using two special wild card characters such as and
- The keyword is used to select rows that do not match the specified pattern of characters.
- The default order of ORDER BY clause is
- The function is used to count the number of records in a column.
- The rows of the table (relation) are referred to as
- A virtual table is called a
- The non-key attribute which helps to make relationship between two tables is known as

Answers: (a) MySQL (b) dual (c) distinct
 (d) % (percent), _ (Underscore) (e) NOT LIKE (f) ascending
 (g) count() (h) tuples (i) view
 (j) Foreign key

2. State whether the following statements are True or False.

- (Duplication of data is known as Data Redundancy.)
- An Attribute is a set of values of a dissimilar type of data.
- MySQL supports different platforms like UNIX and Windows.
- UPDATE TABLE command is used to create table in a database.

- (e) Null (unavailable and unknown) values are entered by the following command:
 `INSERT INTO TABLE_NAME VALUES ("NULL");`

(f) ALTER TABLE command is used to modify the structure of the table.

(g) Each SQL table must have at least one column and one row.

(h) Foreign key column derives its value from the primary key of the parent table.

(i) DISTINCT clause is used to remove redundant rows from the result of the SELECT statement.

(j) SELECT MIN (salary) FROM Employee will return the highest salary from the table.

Answers: (a) True (b) False (c) True (d) False (e) False (f) True
(g) False (h) True (i) True (j) False

3. Multiple Choice Questions (MCQs)

SOLVED QUESTIONS

1. What is SQL?

Ans. Structured Query Language (SQL) is a language used for accessing databases.

2. Define the following terms: Field, Record, Table.

Ans. Field: A field is the smallest unit of a table which is also known as a column. Columns are called attributes in a table. *For example*, Employee Name, Employee ID, etc.

Record: A record is the collection of values/fields of a specific entity. *For example, record of an employee that consists of Employee name, Salary, etc.*

Table: A table is called a relation. It is a collection of records of a specific type of data. For example, employee table, salary table, etc., that can consist of the records of employees and records of salary respectively.

- 3.** What are DDL and DML statements?
- Ans.** DDL statements are used for creating or deleting tables, views, etc. DML statements are used for manipulating values for records in a table.
- 4.** What is a base table?
- Ans.** A base table is a table from which the values can be derived for another table. *For example*, in case of views, the values are extracted from the base table on which the view depends.
- 5.** What is NULL value?
- Ans.** A NULL value in a table is a value in a field which is blank. This means that a field with a NULL value is a field with no value; not even zero is entered.
- 6.** What is a NOT NULL constraint?
- Ans.** If you do not want a column to have a NULL value, then you need to define such a constraint on this column specifying that NULL is now not allowed for that column.
- 7.** (a) What do you mean by Primary Key? Give a suitable example of a Primary Key from a table containing some meaningful data.
- Ans.** An attribute or a set of attributes which is used to identify a tuple (row) uniquely is known as Primary Key.

Table: Students

Admission_No	First Name	Last Name	DOB
27354	Jatin	Kumar	05-02-1998
25350	Mona	Sinha	24-09-2004
26385	George	Moun	19-05-1997
16238	Mukesh	Kumar	24-09-2004

Admission_No. is the Primary Key because this column will contain unique data for each record.

(b) Write a query that displays city, salesman name, code and commission from salesman table.

Ans. SELECT city, salesman_name, code, commission from salesman;

(c) Write a query that selects all orders except those with zero or NULL in the amount field from table orders.

Ans. SELECT * from orders where amount IS NOT NULL;

(d) Write a command that deletes all orders for the customer SOHAN from table customer.

Ans. DELETE from customer where customer_name is LIKE 'SOHAN';

(e) Differentiate between DROP and DELETE command.

Ans. DROP command is used to drop a table along with all the records stored in it whereas DELETE command is used to delete all the records or some of the records from a table without deleting the table.

(f) How can you add a new column or a constraint in a table?

Ans. If you want to add a new column in an existing table, ALTER command is used. *For example*, to add a column bonus in a table emp, the statement will be given as:

```
ALTER table emp ADD  
(bonus Integer);
```

(g) Can you add more than one column in a table by using the ALTER TABLE command?

Ans. Yes, we can add more than one column by using the ALTER TABLE command. Multiple column names are given, which are separated by commas, while giving with ADD. *For example*, to add city and pin code in a table employee, the command can be given as:

```
ALTER table employee  
ADD (city CHAR(30), PINCODE INTEGER);
```

(h) What are JOINS?

Ans. A JOIN is a query that combines tuples from more than one table. In a join query, the table names are given with FROM clause, separated by a comma. *For example,*

```
SELECT name, salary  
      FROM emp1, emp2;
```

In this statement, the two tables are emp1 and emp2 from which the column name and salary are extracted.

8. How can you eliminate duplicate records in a table with select query?

Ans. The DISTINCT clause is used with SELECT statement to hide duplicate records in a table. *For example,* to display cities from table suppliers.

```
SELECT DISTINCT city FROM suppliers;
```

9. How can you DROP a UNIQUE Constraint?

Ans. To drop a UNIQUE constraint, use the following SQL statement:

```
ALTER TABLE SUPPLIER  
DROP CONSTRAINT chkcommission;
```

Here, the constraint chkcommission is deleted from the table supplier.

10. Define a Foreign Key.

Ans. A foreign key is a key which is used to link two tables together. It is also called a referencing key. Foreign key is a column or a combination of columns whose values match a primary key in a different table. The relationship between two tables matches the primary key in one of the tables with a foreign key in the second table. If a table has a primary key defined on any field(s), then you cannot have two records having the same value of that field(s).

11. Define the various SQL Constraints.

Ans. Constraints are the rules enforced on data or columns on a table. These are used to restrict the values that can be inserted in a table. This ensures data accuracy and reliability in the database.

Following are the most commonly used constraints available in SQL:

- (a) NOT NULL Constraint: Ensures that a column cannot have NULL value.
- (b) DEFAULT Constraint: Provides a default value for a column when no value is specified.
- (c) UNIQUE Constraint: Ensures that all values in a column are unique. There should not be any redundant value in a column which is being restricted.
- (d) PRIMARY Key: Uniquely identifies each row/record in a database table.
- (e) FOREIGN Key: Uniquely identifies a row/record in any other database table.
- (f) CHECK Constraint: The CHECK constraint ensures that all values in a column satisfy certain conditions.

For example, to restrict the salary column that it should contain salary more than ₹ 10,000.

12. Consider the following tables: COMPANY and MODEL.

Table: Company

Comp_ID	CompName	CompHO	ContactPerson
1	Titan	Okhla	C.B. Ajit
2	Ajanta	Najafgarh	R. Mehta
3	Maxima	Shahdara	B. Kohli
4	Seiko	Okhla	R. Chadha
5	Ricoh	Shahdara	J. Kishore

Note:

- (a) Comp_ID is the Primary Key.

Table: Model

Model_ID	Comp_ID	Cost	DateOfManufacture
T020	1	2000	2010-05-12
M032	4	7000	2009-04-15
M059	2	800	2009-09-23
A167	3	1200	2011-01-12
T024	1	1300	2009-10-14

Note:

- (b) Model_ID is the Primary Key.
- (c) Comp_ID is the Foreign Key referencing Comp_ID of Company table. Write SQL commands for queries (i) to (iv) and output for (v) and (vi).
 - (i) To display the details of all the models in the Model table in ascending order of DateOfManufacture.
 - (ii) To display the details of those models manufactured in 2011 and whose Cost is below 2000.
 - (iii) To display the Model_ID, Comp_ID, Cost from the table Model, CompName and ContactPerson from Company table, with their corresponding Comp_ID.
 - (iv) To decrease the cost of all the models in Model table by 15%.
 - (v) Select COUNT(DISTINCT CompHO) from Company;
 - (vi) Select CompName, contact('Mr.',ContactPerson) from Company where CompName ends with 'a';

- Ans.**
- (i) select * from model
order by DateOfManufacture;
 - (ii) select * from model
where year(DateOfManufacture) = 2011 and cost < 2000;
 - (iii) select Model_ID, Comp_ID, cost, CompName, ContactPerson from Model, Company
where Model.Comp_ID = Company.Comp_ID;
 - (iv) Update Model
set Cost = Cost - 0.15*Cost;
 - (v) 3
 - (vi) Ajanta Mr. R. Mehta
Maxima Mr. B. Kohli

13. Consider the following two tables: PRODUCT and CLIENT.

Table: Product

P_ID	ProductName	Manufacturer	Price	ExpiryDate
TP01	Talcum Powder	LAK	40	2011-06-26
FW05	Face Wash	ABC	45	2010-12-01
BS01	Bath Soap	ABC	55	2010-09-10
SH06	Shampoo	XYZ	120	2012-04-09
FW12	Face Wash	XYZ	95	2010-08-15

Note:

- P_ID is the Primary Key.

Table: Client

C_ID	ClientName	City	P_ID
1	Cosmetic Shop	Delhi	FW05
6	Total Health	Mumbai	BS01
12	Live Life	Delhi	SH06
15	Pretty One	Delhi	FW05
16	Dreams	Bengaluru	TP01
14	Expressions	Delhi	NULL

**Note:**

C_ID is the Primary Key.

P_ID is the Foreign Key referencing P_ID of Client table.

Write SQL statements for the queries (i) to (iv) and output for (v) and (vi):

- (i) To display the ClientName and City of all Mumbai and Delhi-based clients in Client table.
- (ii) Increase the price of all the products in Product table by 10%.
- (iii) To display the ProductName, Manufacturer, ExpiryDate of all the products that expired on or before '2010-12-31'.
- (iv) To display C_ID, ClientName, City of all the clients (including the ones that have not purchased a product) and their corresponding ProductName sold.
- (v) Select COUNT(DISTINCT Manufacturer) from Product;
- (vi) Select C_ID, Client_Name, City from Client where City Like 'M%';

Ans. (i) Select ClientName, City from Client

where City = 'Mumbai' or City = 'Delhi';

(ii) Update Product

set Price = Price + 0.10 * Price;

(iii) Select ProductName, Manufacturer, ExpiryDate from Product

where ExpiryDate <= '2010-12-31';

(iv) Select C_ID, ClientName, City, ProductName from Client Left Join Product
on Client. P_ID = Product.P_ID;

(v) 3

(vi) 6 Total Health Mumbai

14. Consider the following two tables: STATIONERY and CONSUMER.

Table: Stationery

S_ID	StationeryName	Company	Price	StockDate
DP01	Dot Pen	ABC	10	2011-03-31
PL02	Pencil	XYZ	6	2010-01-01
ER05	Eraser	XYZ	7	2010-02-14
PL01	Pencil	CAM	5	2009-01-09
GP02	Gel Pen	ABC	15	2009-03-19

Note:

(a) S_ID is the Primary Key.

Table: Consumer

C_ID	ConsumerName	Address	P_ID
01	Good Learner	Delhi	PL01
06	Write Well	Mumbai	GP02
12	Topper	Delhi	DP01
15	Write & Draw	Delhi	PL02
16	Motivation	Bengaluru	PL01

Note:

(b) C_ID is the Primary Key.

(c) P_ID is the Foreign Key referencing S_ID of Stationery table.

Write SQL statements for the queries (i) to (iv) and output for (v) and (vi):

- (i) To display details of all the Stationery Items in the Stationery table in descending order of StockDate.
- (ii) To display details of that Stationery item whose Company is XYZ and price is below 10.
- (iii) To display ConsumerName, Address from the table Consumer and Company and Price from Stationery table, with their corresponding S_ID.

- (iv) To increase the price of all the stationery items in Stationery table by ₹ 2.
 (v) Select COUNT(DISTINCT Address) from Consumer;
 (vi) Select StationeryName, price * 3 from Stationery
 where Company = 'CAM';

- Ans.**
- (i) Select * from Stationery
 order by StockDate desc;
 - (ii) Select * from Stationery
 where Company = 'XYZ' and Price < 10;
 - (iii) Select ConsumerName, Address, Company, Price from Stationery, Consumer
 where Stationery.S_ID = Consumer.P_ID;
 - (iv) Update Stationery
 Set Price = Price + 2;
 - (v) 3
 - (vi) Pencil 15

15. Consider the following tables: STOCK and DEALER.

Table: Stock

ItemNo	Item	Dcode	Qty	UnitPrice	StockDate
5005	Ball Pen 0.5	102	100	16	2011-03-31
5003	Ball Pen 0.25	102	150	20	2010-01-01
5002	Gel Pen Premium	101	125	14	2010-02-14
5006	Gel Pen Classic	101	200	22	2009-01-09
5001	Eraser Small	102	210	5	2009-03-19
5004	Eraser Big	102	60	10	2010-12-12
5009	Sharpener Classic	103	160	8	2010-01-23

Note:

- (a) ItemNo is the Primary Key.
- (b) Dcode is the Foreign Key referencing Dcode of Dealer table.

Table: Dealer

Dcode	DName
101	Reliable Stationers
103	Class Plastics
104	Fair Deals
102	Clear Deals

Note:

- (c) Dcode is the Primary Key.

Write SQL statements for the queries (i) to (iv) and output for (v) and (vi):

- (i) To display details of all the Items in the Stock table in ascending order of StockDate.
- (ii) To display details of those Items in Stock table whose Dealer Code(Dcode) is 102 or quantity in Stock(Qty) is more than 100.
- (iii) To insert a record in the Stock table with the values:
 (5010, 'Pencil HB', 102, 500, 10, '2010-01-26')
- (iv) To display Dcode, Dname from Dealer table and Item, UnitPrice from Stock table of all the Dealers
 (including the dealer details that have not sold any item)
- (v) Select COUNT(DISTINCT Dcode) from Stock;
- (vi) Select Qty * UnitPrice from Stock where ItemNo=5006;

- Ans.**
- (i) Select * from Stock
order by StockDate;
 - (ii) Select * from Stock
where Dcode = 102 or Qty > 100;
 - (iii) Insert into Stock
values (5010, 'Pencil HB', 102, 500, 10, '2010-01-26');
 - (iv) Select Dealer.Dcode, Dname, Item, UnitPrice from Dealer left join Stock on Dealer.Dcode = Stock.Dcode;
 - (v) 3
 - (vi) 4400

16. (a) Explain the concept of Cartesian product between two tables with the help of an example.

Note: Answer the questions (b) and (c) on the basis of the following tables SHOP and ACCESSORIES.

Table: SHOP

Id	SName	Area
S01	ABC Computronics	CP
S02	All Infotech Media	GK II
S03	Tech Shoppe	CP
S04	Geek Tenco Soft	Nehru Place
S05	Hitech Tech Store	Nehru Place

Table: ACCESSORIES

No	Name	Price	Id
A01	Motherboard	12000	S01
A02	Hard Disk	5000	S01
A03	Keyboard	500	S02
A04	Mouse	300	S01
A05	Motherboard	13000	S02
A06	Keyboard	400	S03
A07	LCD	6000	S04
T08	LCD	5500	S05
T09	Mouse	350	S05
T010	Hard Disk	450	S03

Ans. When you join two or more tables without any condition, it is called Cartesian product or Cross Join.

Example: SELECT * FROM SHOP, ACCESSORIES;

(b) Write the SQL queries:

- (i) To display Name and Price of all the Accessories in ascending order of their Price.
- (ii) To display Id and SName of all Shops located in Nehru Place.
- (iii) To display Minimum and Maximum Price of all the accessories.

Ans. (i) SELECT Name, Price FROM ACCESSORIES ORDER BY Price;

(ii) SELECT Id, SName FROM SHOP WHERE Area='Nehru Place';

(iii) SELECT Name, MAX(Price), MIN(Price) FROM ACCESSORIES;

(c) Write the output of the following SQL commands:

- (i) SELECT DISTINCT NAME FROM ACCESSORIES WHERE PRICE>=5000;
- (ii) SELECT AREA, COUNT(*) FROM SHOP GROUP BY AREA;
- (iii) SELECT COUNT(DISTINCT AREA) FROM SHOP;

- Ans.** (i) Name
 Motherboard
 Hard Disk
 LCD
 (ii) AREA COUNT
 CP 2
 GK II 1
 Nehru Place 2
 (iii) COUNT 3

- 17.** (a) Define a candidate key with example and write SQL queries for (b) to (f) and write the output for the SQL queries mentioned in parts (g1) to (g3) on the basis of tables PRODUCTS and SUPPLIERS.

Table: PRODUCTS

PID	PName	QTY	Price	COMPANY	SUPCODE
101	DIGITAL CAMERA 14X	120	12000	RENIX	S01
102	DIGITAL PAD 11i	100	22000	DIGI POP	S02
104	PEN DRIVE 16 GB	500	1100	STOREKING	S01
106	LED SCREEN 32	70	28000	DISPEXPERTS	S02
105	CAR GPS SYSTEM	60	12000	MOVEON	S03

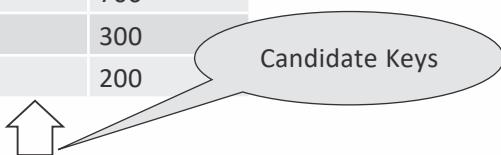
Table: SUPPLIERS

SUPCODE	SNAME	CITY
S01	GET ALL INC	KOLKATA
S03	EASY MARKET CORP	DELHI
S02	DIGI BUSY GROUP	CHENNAI

- Ans.** (a) A table may have more than one such attribute/group of attributes that identifies a tuple uniquely; all such attribute(s) are known as Candidate Keys.

Table: Item

Ino	Item	QTY
101	Pen	560
102	Pencil	780
104	CD	450
109	Floppy	700
105	Eraser	300
103	Duster	200



Candidate Keys

- (b) To arrange and display all the records of table Products on the basis of product name in the ascending order.

Ans. SELECT * FROM PRODUCTS ORDER BY PNAME;

- (c) To display product name and price of all those products whose price is in the range of 10000 and 15000 (both values inclusive).

Ans. Select PNAME, PRICE FROM PRODUCTS WHERE PRICE>=10000 and PRICE<=15000;

- (d) To display the price, product name and quantity (*i.e.*, qty) of those products which have quantity more than 100.

Ans. SELECT PRICE, PNAME, QTY FROM PRODUCTS WHERE QTY>100;

- (e) To display the names of those suppliers who are either from DELHI or from CHENNAI.

Ans. SELECT SNAME FROM SUPPLIERS WHERE CITY="DELHI" or CITY="CHENNAI";

- (f) To display the names of the companies and the names of the products in descending order of company names.

Ans. SELECT COMPANY, PNAME FROM PRODUCTS ORDER BY COMPANY DESC;

- (g) Obtain the outputs of the following SQL queries based on the data given in tables PRODUCTS and SUPPLIERS above.

(g1) SELECT DISTINCT SUPCODE FROM PRODUCTS;

(g2) SELECT MAX(PRICE), MIN(PRICE) FROM PRODUCTS;

(g3) SELECT PRICE*QTY AMOUNT FROM PRODUCTS WHERE PID=104;

Ans. (g1)

DISTINCT SUPCODE
S01
S02
S03

(g2)

MAX (PRICE)	MIN (PRICE)
28000	1100

(g3)

PRICE*QTY
550000

18. (a) Give a suitable example of a table with sample data and illustrate Primary and Alternate Keys in it.

Ans. Primary Key: Primary key is a set of one or more fields/columns of a table that uniquely identifies a record in database table. It cannot accept null, duplicate values. Only one Candidate Key can be Primary Key.

Alternate key: Alternate key is a key that can work as a primary key. Basically, it is a candidate key that currently is not a primary key.

Example: In the table given below, AdmissionNo becomes the Alternate Key when we define Registration No as the Primary Key.

Student Registration Table:

RegistrationNo	AdmissionNo	Name	Phone	Gender	DOB
CBSE4554	215647	Mihir Ranjan	9568452325	Male	1992-04-15
CBSE6985	265894	Amita Guha	8456985445	Female	1993-03-24
CBSE5668	458961	Rajesh Singh	9654212440	Male	1992-12-04
CBSE3654	469799	Mohit Patel	7421589652	Male	1992-05-16

Primary Key – RegistrationNo

Alternate Key – AdmissionNo

(b) Consider the following table CARDEN given below:

Table: CARDEN

Ccode	CarName	Make	Color	Capacity	Charges
501	A-Star	Suzuki	RED	3	14
503	Indigo	Tata	SILVER	3	12
502	Innova	Toyota	WHITE	7	15
509	SX4	Suzuki	SILVER	4	14
510	C Class	Mercedes	RED	4	35

Write SQL commands for the following statements:

- (i) To display the names of all the silver-coloured cars.
- (ii) To display name, make and capacity of cars in descending order of their sitting capacity.
- (iii) To display the highest charges at which a vehicle can be hired from CARDEN.

- Ans.**
- (i) `SELECT CarName FROM carden WHERE Color LIKE 'Silver';`
 - (ii) `SELECT CarName, Make, Capacity FROM carden ORDER BY Capacity DESC;`
 - (iii) `SELECT MAX(Charges) FROM carden;`
 - (c) Give the output of the following SQL queries:
 - (i) `SELECT COUNT(DISTINCT Make) FROM CARDEN;`
 - (ii) `SELECT MAX(Charges), MIN(Charges) FROM CARDEN;`
 - (iii) `SELECT COUNT(*), Make FROM CARDEN;`
 - (iv) `SELECT CarName FROM CARDEN WHERE Capacity=4;`

- Ans.**
- (i) `COUNT(DISTINCT Make)`

4

- (ii) `MAX(Charges) MIN(Charges)`
- 35 12
- (iii) `COUNT(*) Make`
- 5 Suzuki
- (iv) `CarName`
- SX4
- C Class

19. Consider the tables EMPLOYEE and SALGRADE given below and answer (a) and (b) parts of this question.

Table: EMPLOYEE

Ecode	Name	Desig	SGrade	Doj	DOB
101	Abdul Ahmad	EXECUTIVE	S03	23-Mar-2003	13-Jan-1980
102	Ravi Chander	HEAD-IT	S02	12-Feb-2010	22-Jul-1987
103	John Ken	RECEPTIONIST	S03	24-Jun-2009	24-Feb-1983
105	NazarAmeen	GM	S02	11-Aug-2006	03-Mar-1984
108	PriyamSen	CEO	S01	29-Dec-2004	19-Jan-1982

Table: SALGRADE

SGrade	Salary	HRA
S01	56000	18000
S02	32000	12000
S03	24000	8000

- (a) Write SQL commands for the following statements:

- (i) To display the details of all EMPLOYEES in descending order of DOJ.
- (ii) To display NAME and DESIG of those EMPLOYEES whose SALGRADE is either S02 or S03.

(iii) To display the content of the entire EMPLOYEES table, whose DOJ is in between '09-Feb-2006' and '08-Aug-2009'.

(iv) To add a new row with the following content:

109,'Harish Roy','HEAD-IT','S02','9-Sep-2007','21-Apr-1983'

Ans. (i) SELECT * FROM employee ORDER BY DOJ DESC;

(ii) SELECT name, design FROM employee WHERE sgrade=S02 OR sgrade=S03;

(iii) SELECT * FROM employee WHERE DOJ BETWEEN '09-Feb-2006' AND '08-Aug-2009';

(iv) Insert into employee values(109, 'Harish Roy', 'HEAD-IT', 'S02', '9-Sep-2007', '21-Apr-1983');

(b) Give the output of the following SQL queries:

(i) SELECT COUNT(SGRADE),SGRADE FROM EMPLOYEE GROUP BY SGRADE;

(ii) SELECT MIN(DOB),MAX(DOJ) FROM EMPLOYEE;

(iii) SELECT SGRADE, SALARY+HRA FROM SALGRADE WHERE SGRADE ='S02';

Ans. (i) COUNT (SGRADE) SGRADE

2 S03

2 S02

1 S01

(ii) MAX (DOB) MIN (DOJ)

22-Jul-1987 23-Mar-2003

(iii) SGRADE SALARY+HRA

P003 440000

20. Consider the following tables STOCK and DEALERS and answer (B1) and (B2) parts of this question:

Table: STOCK

ItemNo	Item	Dcode	Qty	UnitPrice	StockDate
5005	Ball Pen 0.5	102	100	16	31-Mar-10
5003	Ball Pen 0.25	102	150	20	01-Jan-10
5002	Gel Pen Premium	101	125	14	14-Feb-10
5006	Gel Pen Classic	101	200	22	01-Jan-09
5001	Eraser Small	102	210	5	19-Mar-09
5004	Eraser Big	102	60	10	12-Dec-09
5009	Sharpener Classic	103	160	8	23-Jan-09

Table: DEALERS

Dcode	Dname
101	Reliable Stationers
103	Classic Plastics
102	Clear Deals

B1. Write SQL commands for the following statements:

(i) To display details of all Items in the Stock table in ascending order of StockDate.

Ans. SELECT * FROM STOCK ORDER BY StockDate;

(ii) To display ItemNo and Item name of those items from Stock table whose UnitPrice is more than ₹ 10.

Ans. SELECT ItemNo, Item FROM STOCK WHERE UnitPrice>10;

(iii) To display the details of those items whose dealer code (Dcode) is 102 or Quantity in Stock (Qty) is more than 100 from the table Stock.

Ans. SELECT * FROM STOCK WHERE Dcode=102 OR Qty>100;

(iv) To display Maximum UnitPrice of items for each dealer individually as per Dcode from the table Stock.

Ans. SELECT Dcode, MAX (UnitPrice) FROM STOCK GROUP BY Dcode;

B2. Give the output of the following SQL queries:

(i) SELECT COUNT(DISTINCT Dcode) FROM Stock;

Ans.

Count(DISTINCT Dcode)
3

Ans.

(ii) SELECT Qty*UnitPrice FROM Stock WHERE ItemNo=5006;

Qty*UnitPrice

4400

Ans.

(iii) SELECT MIN(StockDate) FROM Stock;

MIN (StockDate)

01-Jan-09

21. Write the SQL commands for (a) to (d) and output for (e):

Table: GRADUATE

S.NO	NAME	STIPEND	SUBJECT	AVERAGE	DIV
1	KARAN	400	PHYSICS	68	I
2	DIWAKAR	450	COMP. Sc.	68	I
3	DIVYA	300	CHEMISTRY	62	I
4	REKHA	350	PHYSICS	63	I
5	ARJUN	500	MATHS	70	I
6	SABINA	400	CHEMISTRY	55	II
7	JOHN	250	PHYSICS	64	I
8	ROBERT	450	MATHS	68	I
9	RUBINA	500	COMP.Sc.	62	I
10	VIKAS	400	MATHS	57	II

- (a) List the names of those students who obtained DIV I sorted by NAME.
- (b) Display a report, listing NAME, STIPEND, SUBJECT and amount of stipend received in a year assuming that the STIPEND is paid every month.
- (c) To count the number of students who are either PHYSICS or COMPUTER SC graduates.
- (d) To insert a new row in the GRADUATE table: 11, "KAJOL", 300, "computer sc", 75, 1
- (e) Give the output of the following sql statement based on table GRADUATE:
 - (i) Select MIN(AVERAGE) from GRADUATE where SUBJECT="PHYSICS";
 - (ii) Select SUM(STIPEND) from GRADUATE WHERE div=2;
 - (iii) Select AVG(STIPEND) from GRADUATE where AVERAGE>=65;
 - (iv) Select COUNT(distinct SUBJECT) from GRADUATE;

Ans.

- (a) SELECT NAME from GRADUATE where DIV = 'I' order by NAME;
- (b) SELECT NAME,STIPEND,SUBJECT,STIPEND*12 from GRADUATE;
- (c) SELECT SUBJECT,COUNT(*) from GRADUATE group by SUBJECT having SUBJECT='PHYSICS' or SUBJECT='COMPUTER SC';
- (d) INSERT INTO GRADUATE values(11,'KAJOL',300,'COMPUTER SC',75,1);
- (e) (i) 63
 (ii) 800
 (iii) 475
 (iv) 4

22. Write SQL commands for (i) to (vii) on the basis of table SPORTS:

Table: SPORTS

StudentNO	Class	Name	Game1	Grade1	Game2	Grade2
10	7	Sammer	Cricket	B	Swimming	A
11	8	Sujit	Tennis	A	Skating	C
12	7	Kamal	Swimming	B	Football	B
13	7	Venna	Tennis	C	Tennis	A
14	9	Archana	Basketball	A	Cricket	A
15	10	Arpit	Cricket	A	Athletics	C

- (a) Display the names of the students who have grade 'C' in either Game1 or Game2 or both.
- (b) Display the number of students getting grade 'A' in Cricket.
- (c) Display the names of the students who have the same game for both Game1 and Game2.
- (d) Display the game taken up by the students, whose name starts with 'A'.
- (e) Add a new column named 'Marks'.
- (f) Assign a value 200 for marks for all those who are getting grade 'B' or grade 'A' in both Game1 and Game2.

Ans. (a) SELECT Name for SPORTS where grade='C' or Grade2='C';

(b) SELECT Count(*) from SPORTS where grade='A';

(c) SELECT name from SPORTS where game1 = game2;

(d) SELECT game,game2 from SPORTS where name like 'A%';

(e) ALTER TABLE SPORTS add (marks int(4));

(f) UPDATE SPORTS set marks=200 where grade='A';

23. (a) Study the following table and write SQL queries for questions (i) to (iv) and output for (v) and (vi).

Table: Orders

Orderid	Pname	Quantity	Rate	Sale_date	Discount
1001	Pen	10	20	2019-10-05	
1002	Pencil	20	10	2019-10-21	
1003	Book	10	100	2019-11-02	50
1004	Eraser	100	5	2019-12-05	25
1005	Copy	50	20	2019-12-10	

(i) Write SQL query to display Pname, Quantity and Rate for all the orders that are either Pencil or Pen.

(ii) Write SQL query to display the orders which are not getting any Discount.

(iii) Write SQL query to display the Pname, Quantity and Sale_date for all the orders whose total cost (Quantity * Rate) is greater than 500.

(iv) Write SQL query to display the orders whose Rate is in the range 20 to 100.

(v) SELECT Pname, Quantity from Orders WHERE Pname LIKE('_e%');

(vi) SELECT Pname, Quantity, Rate FROM Orders ORDER BY Quantity DESC;

Ans. (i) SELECT Pname, Quantity, Rate FROM Orders WHERE Pname IN('Pencil','Pen');

(ii) SELECT * FROM Orders WHERE Discount is NULL;

(iii) SELECT Pname, Quantity, Sale_date FROM Orders WHERE Quantity * Cost > 500;

(iv) SELECT * FROM Orders WHERE Rate BETWEEN 20 AND 100;

(v)

Pname	Quantity
Pen	10
Pencil	20

(vi)		
Pname	Quantity	Rate
Eraser	100	5
Copy	50	20
Pencil	20	20
Book	10	100
Pen	10	20

- (b) Based on the Orders Table given above, write SQL query to display the minimum quantity, maximum quantity and total quantity of Orders.

Ans. SELECT MIN(Quantity), MAX(Quantity), SUM(Quantity) FROM Orders;

- (c) What is the difference between COUNT(*) and COUNT(Column Name)?

Ans. COUNT(*) includes the cells in the column containing NULL values.

COUNT(Column name) ignores the cells in the column containing NULL values.

UNSOLVED QUESTIONS

1. What is an Alternate Key?
2. What are views? How are they useful?
3. Define the following terms:
 - (a) Relation
 - (b) Tuple
 - (c) Attribute
 - (d) Domain
4. What do you understand by the terms candidate key and cardinality of a relation in a relational database?
5. What is SQL? What are different categories of commands available in SQL?
6. What is a database system? What is its need?
7. Differentiate between DDL and DML commands.
8. What is a datatype? Name some datatypes available in MySQL.
9. Differentiate between char and varchar datatypes.
10. Which operator concatenates two strings in a query result?
11. How would you calculate 13×15 in SQL?
12. Which keywords eliminate redundant data from a query?
13. Write SQL queries to perform the following based on the table PRODUCT having fields as (prod_id, prod_name, quantity, unit_rate, price, city)
 - (i) Display those records from table PRODUCT where prod_id is more than 100.
 - (ii) List records from table PRODUCT where prod_name is 'Almirah'.
 - (iii) List all those records whose price is between 200 and 500.
 - (iv) Display the product names whose price is less than the average of price.
 - (v) Show the total number of records in the table PRODUCT.
14. Define the following terms:
 - (i) Database
 - (ii) Data Inconsistency
 - (iii) Primary Key
 - (iv) Candidate Key
 - (v) Indexes

15. Consider the following tables STORE and SUPPLIERS. Write SQL commands for the statements (i) to (iv) and give outputs for SQL queries (v) to (viii).

Table: Store

ItemNo	Item	Scode	Qty	Rate	LastBuy
2005	Sharpener Classic	23	60	8	31-Jun-09
2003	Ball Pen 0.25	22	50	25	01-Feb-10
2002	Gel Pen Premium	21	150	12	24-Feb-10
2006	Gel Pen Classic	21	250	20	11-Mar-09
2001	Eraser Small	22	220	6	19-Jan-09
2004	Eraser Big	22	110	8	02-Dec-09
2009	Ball Pen 0.5	21	180	18	03-Nov-09

Table: SUPPLIERS

Scode	Sname
21	Premium Stationery
23	Soft Plastics
22	Tetra Supply

- (i) To display details of all the items in the Store table in ascending order of LastBuy.
- (ii) To display Itemno and item name of those items from Store table whose rate is more than 15 rupees.
- (iii) To display the details of those items whose supplier code is 22 or quantity in store is more than 110 from the table Store.
- (iv) To display minimum rate of items for each Supplier individually as per Scode from the table Store.
- (v) SELECT COUNT(DISTINCT Scode) FROM STORE;
- (vi) SELECT Rate*Qty FROM STORE WHERE Itemno=2004;
- (vii) SELECT Item, Sname FROM STORE S, SUPPLIER P WHERE S.Scode=P.Scode AND ItemNo=2006;
- (viii) SELECT MAX(LastBuy)FROM STORE;

16. Write SQL commands for (i) to (vi) on the basis of relations given below:

BOOKS

book_id	Book_name	author_name	Publishers	Price	Type	qty
k0001	Let us C	Sanjay Mukharjee	EPB	450	Comp	15
p0001	Genuine	J. Mukhi	FIRST PUBL.	755	Fiction	24
m0001	Mastering C++	Kantkar	EPB	165	Comp	60
n0002	VC++ advance	P. Purohit	TDH	250	Comp	45
k0002	Programming with Python	Sanjeev	FIRST PUBL.	350	Fiction	30

ISSUED

Book_ID	Qty_Issued
L02	13
L04	5
L05	21

- (i) To show the books of FIRST PUBL. Publishers written by P. Purohit.
- (ii) To display cost of all the books published for FIRST PUBL.
- (iii) Depreciate the price of all books of EPB publishers by 5%.
- (iv) To display the BOOK_NAME and price of the books, more than 3 copies of which have been issued.
- (v) To show total cost of books of each type.
- (vi) To show the details of the costliest book.

17. Write SQL commands for (i) to (vi) and write output for (vii) on the basis of PRODUCTS relation given below:

PRODUCTS TABLE

PCODE	PNAME	COMPANY	PRICE	STOCK	MANUFACTURE	WARRANTY
P001	TV	BPL	10000	200	12-JAN-2018	3
P002	TV	SONY	12000	150	23-MAR-2017	4
P003	PC	LENOVO	39000	100	09-APR-2018	2
P004	PC	COMPAQ	38000	120	20-JUN-2019	2
P005	HANDYCAMS	SONY	18000	250	23-MAR-2017	3

- (i) To show details of all PCs with stock more than 110.
- (ii) To list the company which gives warranty of more than 2 years.
- (iii) To find stock value of the BPL company where stock value is the sum of the products of price and stock.
- (iv) To show number of products from each company.
- (v) To count the number of PRODUCTS which shall be out of warranty on 20-NOV-2020.
- (vi) To show the PRODUCT name of the products which are within warranty as on date.
- (vii) Give the output of the following statements:
 - (a) Select COUNT(distinct company) from PRODUCT;
 - (b) Select MAX(price) from PRODUCT where WARRANTY <= 3;

18. What are DDL and DML?

19. Differentiate between primary key and candidate key in a relation.

20. What do you understand by the terms Cardinality and Degree of a relation in relational database?

21. Differentiate between DDL and DML. Mention the two commands for each category.

22. Consider the given table and answer the questions.

Table: SchoolBus

Rtno	Area_Covered	Capacity	Noofstudents	Distance	Transporter	Charges
1	Vasant Kunj	100	120	10	Shivam travels	100000
2	Hauz Khas	80	80	10	Anand travels	85000
3	Pitampura	60	55	30	Anand travels	60000
4	Rohini	100	90	35	Anand travels	100000
5	Yamuna Vihar	50	60	20	Bhalla travels	55000
6	Krishna Nagar	70	80	30	Yadav travels	80000
7	Vasundhara	100	110	20	Yadav travels	100000
8	Paschim Vihar	40	40	20	Speed travels	55000
9	Saket	120	120	10	Speed travels	100000
10	Janakpuri	100	100	20	Kisan Tours	95000

- (i) To show all information of students where capacity is more than the no. of students in order of rtno.
- (ii) To show area_covered for buses covering more than 20 km., but charges less than 80000.
- (iii) To show transporter-wise total no. of students travelling.
- (iv) To show rtno, area_covered and average cost per student for all routes where average cost per student is—charges/noofstudents.
- (v) Add a new record with the following data:
(11, "Motibagh", 35, 32, 10, "kisan tours", 35000)
- (vi) Give the output considering the original relation as given:
 - (a) select sum(distance) from schoolbus where transporter= "Yadav travels";
 - (b) select min(noofstudents) from schoolbus;
 - (c) select avg(charges) from schoolbus where transporter= "Anand travels";
 - (d) select distinct transporter from schoolbus;

23. Write SQL Commands for (i) to (v) and write the outputs for (vi) to (viii) on the basis of the following table:

Table: FURNITURE

NO	ITEM	TYPE	DATEOFSOCK	PRICE	DICOUNT
1	WhiteLotus	DoubleBed	2002-02-23	3000	25
2	Pinkfeathers	BabyCot	2002-01-29	7000	20
3	Dolphin	BabyCot	2002-02-19	9500	20
4	Decent	OfficeTable	2002-02-01	25000	30
5	Comfortzone	DoubleBed	2002-02-12	25000	30
6	Donald	BabyCot	2002-02-24	6500	15

- (i) To list the details of furniture whose price is more than 10000.
- (ii) To list the Item name and Price of furniture whose discount is between 10 and 20.
- (iii) To delete the record of all items where discount is 30.
- (iv) To display the price of 'Babycot'.
- (v) To list item name, type and price of all items whose names start with 'D'.
- (vi) Select Distinct Type from Furniture;
- (vii) Select Max(Price) from Furniture where DateofStock>'2002-02-15';
- (viii) Select Count(*) from Furniture where Discount<25;

24. Write the outputs of the following:

- (i) Select POWER(2,5);
- (ii) Select ROUND(10.195,2);
- (iii) Select TRUNCATE(10.195,2);
- (iv) Select SUBSTR('COMPUTER SCIENCE',3,4);
- (v) Select INSTR('Technical Teaching','ch');

25. Write SQL Commands for the following on the basis of the given table GRADUATE:

Table: GRADUATE

SL.No.	NAME	STIPEND	SUBJECT	AVERAGE	RANK
1	KARAN	400	PHYSICS	68	1
2	RAJ	450	CHEMISTRY	68	1
3	DEEP	300	MATHS	62	2
4	DIVYA	350	CHEMISTRY	63	1
5	GAURAV	500	PHYSICS	70	1
6	MANAV	400	CHEMISTRY	55	2
7	VARUN	250	MATHS	64	1
8	LIZA	450	COMPUTER	68	1
9	PUJA	500	PHYSICS	62	1
10	NISHA	300	COMPUTER	57	2

- (i) List the names of those students who have obtained rank 1 sorted by NAME.
- (ii) Display a list of all those names whose AVERAGE is greater than 65.
- (iii) Display the names of those students who have opted COMPUTER as a SUBJECT with an AVERAGE of more than 60.
- (iv) List the names of all the students in alphabetical order.
- (v) SELECT * FROM GRADUATE WHERE NAME LIKE "% I %";
- (vi) SELECT DISTINCT RANK FROM GRADUATE;

26. (a) What is the difference between Candidate key and Alternate key?

(b) What is the degree and cardinality of a table having 10 rows and 5 columns?

(c) For the given table, do as directed:

Table: STUDENT

Column name	Data type	size	Constraint
ROLLNO	Integer	4	Primary Key
SNAME	Varchar	25	Not Null
GENDER	Char	1	Not Null
DOB	Date		Not Null
FEES	Integer	4	Not Null
HOBBY	Varchar	15	Null

- (i) Write SQL query to create the table.
- (ii) Write SQL query to increase the size of SNAME to hold 30 characters.
- (iii) Write SQL query to remove the column HOBBY.
- (iv) Write SQL query to insert a row in the table with any values of your choice that can be accommodated there.

27. Write SQL queries based on the following tables:

PRODUCT:

P_ID	ProductName	Manufacturer	Price	Discount
TP01	Talcum Powder	LAK	40	
FW05	Face Wash	ABC	45	5
BS01	Bath Soap	ABC	55	
SH06	Shampoo	XYZ	120	10
FW12	Face Wash	XYZ	95	

CLIENT:

C_ID	ClientName	City	P_ID
01	Cosmetic Shop	Delhi	TP01
02	Total Health	Mumbai	FW05
03	Live Life	Delhi	BS01
04	Pretty Woman	Delhi	SH06
05	Dreams	Delhi	FW12

- (i) Write SQL Query to display ProductName and Price for all products whose Price is in the range 50 to 150.
- (ii) Write SQL Query to display details of products whose manufacturer is either XYZ or ABC.
- (iii) Write SQL query to display ProductName, Manufacturer and Price for all products that are not giving any discount.
- (iv) Write SQL query to display ProductName and price for all products whose ProductName ends with 'h'.
- (v) Write SQL query to display ClientName, City, P_ID and ProductName for all clients whose city is Delhi.
- (vi) Which column is used as Foreign Key and name the table where it has been used as Foreign key.

28. Answer the question based on the table given below:

Table: HOSPITAL

SL.No.	Name	Age	Department	Dateofadm	Charges	Sex
1	Arpit	62	Surgery	21-01-98	300	M
2	Zareena	22	ENT	12-12-97	250	F
3	Kareem	32	Orthopedic	19-02-98	200	M
4	Arun	12	Surgery	11-01-98	300	M
5	Zubin	30	ENT	12-01-98	250	M

6	Ketaki	16	ENT	24-02-98	250	F
7	Ankit	29	Cardiology	20-02-98	800	F
8	Zareen	45	Gynecology	22-02-98	300	F
9	Kush	19	Cardiology	13-01-98	800	M
10	Shilpa	23	Nuclear Medicine	21-02-98	400	F

- (a) To list the names all the patients admitted after 15/01/98.
- (b) To list the names of female patients who are in ENT department.
- (c) To list the names of all patients with their date of admission in ascending order.
- (d) To display Patient's Name, Charges, Age for only female patients.
- (e) Find out the output of the following SQL commands:-
 - (i) Select COUNT(DISTINCT charges) from HOSPITAL;
 - (ii) Select MIN(Age) from HOSPITAL where Sex="F";