

# DATABASE

**Data:** Facts, figures, statistics etc.

**Record:** Collection of related data items.

**Table or Relation:** Collection of related records.

**Database:** Collection of related relation/data. In database, data is organized strictly in row and column format. The columns are called **Fields**, **Attributes** or **Domains**. The rows are called **Tuples** or **Records**.

## Features of Data In a Database:

- 1) Security
- 2) Consistency
- 3) Non-Redundancy
- 4) Shared
- 5) Independence
- 6) Persistence

**DBMS(Database Management System)** It is software that allows creation, definition and manipulation of database. It is middle layer between data and program.

## File System:

- Stores permanent records in various files
- Need application program to access and manipulate data.

## Disadvantage of File System:

- Data Redundancy
- Data Inconsistency
- Difficult in accessing data
- Data Integrity
- Low Security

**Data redundancy:** Data redundancy is the repetition or superfluity of data. Data redundancy data is a common issue in computer data storage and database systems. This data repetition may occur either if a field is repeated in two or more tables or if the field is repeated within the table. Data can appear multiple times in a database for a variety of reasons. A positive type of data redundancy works to safeguard data and promote consistency. Many developers consider it acceptable for data to be stored in multiple places. The key is to have a central, master field or space for this data, so that there is a way to update all of the places where data is redundant through one central access point. Otherwise, data redundancy can lead to big problems with data inconsistency, where one update does not automatically update another field. For example, a shop may have the same customer's name appearing several times if that customer has bought several different products at different dates.

## Disadvantages Of Data Redundancy:

- 1) Increases the size of the database unnecessarily.
- 2) Causes data inconsistency.
- 3) Decreases efficiency of database.
- 4) May cause data corruption.

**Data Isolation:** The database must remain in a consistent state after any transaction. No transaction should have any adverse effect on the data residing in the database. If the database was in a consistent state before the execution of a transaction, it must remain consistent after the execution of the transaction.

as well. As an example, if two people are updating the same catalog item, it's not acceptable for one person's changes to be "clobbered" when the second person saves a different set of changes. Both users should be able to work in isolation, working as though he or she is the only user. Each set of changes must be isolated from those of the other users.

**Data Integrity** is the assurance that information is unchanged from its source, and has not been accidentally (e.g. through programming errors), or maliciously (e.g. through breaches or hacks) modified, altered or destroyed. In another words, it concerns with the completeness, soundness, and wholeness of the data that complies with the intention of data creators. *It's a logical property of the DB, independent of the actual data.*

**Data Consistency** refers to the usability of the Data, and is mostly used in single site environment. But still in single site environment, some problems may arise in Data Consistency during recovery activities when original data is replaced by the backup copies of Data. You have to make sure that you data is usable while backing up data.

**Data Abstraction:** To simplify the interaction of users and database, DBMS hides some information which is not user interest is called Data Abstraction. So, developer hides complexity from users and show Abstract view of data.

### **DBMS Architecture/3-Tier Architecture:**

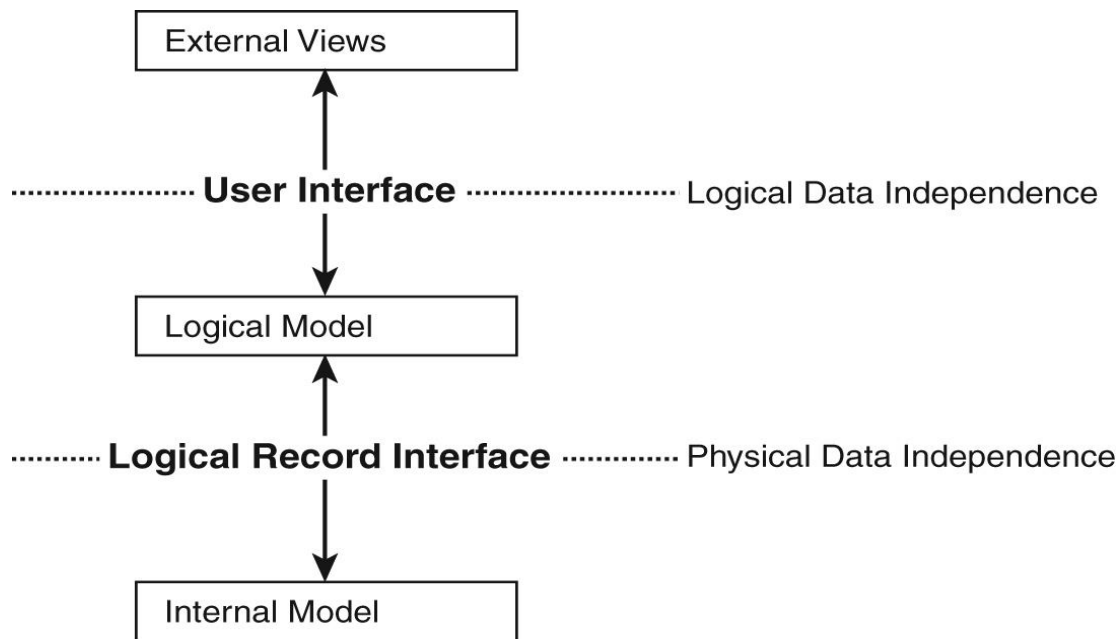
**1)External/View Level:** It is user's view of the database. This level describes the part of the database that is relevant to each user.

#### **2)Conceptual/Logical Level:**

- Describes what data is stored in the database and the relationship among the data.
- Represent all entities, their attributes and their relationship
- Constraints on the data
- Security and Integrity information

#### **3)Physical/Internal Level:**

- Describes how the data is stored in the database
- Storage Space allocation for data and indexes
- File System
- Data compression and Data encryption techniques
- Record Placement



### Schemas:

- It is overall description of the database. In three-level architecture, one schema at each level.
- Does not specify relationship among files.

### Instances:

Collection of information stored in the database at a particular moment.

**Sub-schema:** It is a subset of schema and inherits the same property that the schema has. It is an application programmer's or user view of the data items types and record types which he or she uses.

### Data Independence in DBMS:

Upper level are unaffected by changes in lower level.

### Two Types of Data Independence:

#### a)Physical Data Independence:

- Physical storage structure or devices can be changed without affecting conceptual schema.
- Modification done to improve performance.
- It provide independence to conceptual schema and external schema

#### b)Logical Data Independence:

- Conceptual schema can be changed without affecting external schema.
- Structure of database is altered when modification done in conceptual schema.
- It provide independence to external schema.

# ORACLE

An **Oracle database** is a collection of data treated as a unit. The purpose of a **database** is to store and retrieve related information. A **database** server is the key to solving the problems of information management.

- Oracle 9i is an Object/Relational Database Management System specifically designed for **e-commerce**.
- Oracle 9i, a version of Oracle database. The letter “i” refers to the **internet**.
- It can scale ten thousands of concurrent users.
- It includes Oracle 9i Application server and Oracle 9i Database that provide a comprehensive high-performance infrastructure for Internet Applications.
- It supports client-server and web based applications.
- The maximum Database holding capacity of Oracle 9i is upto **512** peta bytes(PB).[1 Peta Byte = 1000 Tera Byte]
- It offers Data warehousing features and also many management features.

We can set primary key on table up to 16 columns of table in oracle 9i as well as in Oracle 10g. The maximum number of data files in Oracle 9i and Oracle 10g Database is 65,536.

**CURSOR:** A cursor is a temporary work area created in the system memory when a SQL statement is executed. A cursor contains information on a select statement and the rows of data accessed by it. This temporary work area is used to store the data retrieved from the database, and manipulate this data. A cursor can hold more than one row, but can process only one row at a time. The set of rows the cursor holds is called the *active* set.

## Two Types of Cursor :

### 1)Implicit Cursor

Implicit cursors are automatically created by Oracle whenever an SQL statement is executed, when there is no explicit cursor for the statement. Programmers cannot control the implicit cursors and the information in it. Whenever a DML statement (INSERT, UPDATE and DELETE) is issued, an implicit cursor is associated with this statement. For INSERT operations, the cursor holds the data that needs to be inserted. For UPDATE and DELETE operations, the cursor identifies the rows that would be affected.

### 2)Explicit Cursor

They must be created when you are executing a SELECT statement that returns more than one row. Even though the cursor stores multiple records, only one record can be processed at a time, which is called as current row. When you fetch a row the current row position moves to next row.

For Example: When you execute INSERT, UPDATE, or DELETE statements the cursor attributes tell us whether any rows are affected and how many have been affected. When a SELECT... INTO statement is executed in a PL/SQL Block, implicit cursor attributes can be used to find out whether any row has been returned by the SELECT statement. PL/SQL returns an error when no data is selected.

In PL/SQL, you can refer to the most recent implicit cursor as the **SQL cursor**, which always has the attributes like %FOUND, %ISOPEN, %NOTFOUND, and %ROWCOUNT. The SQL cursor has additional attributes, %BULK\_ROWCOUNT and %BULK\_EXCEPTIONS, designed for use with the FORALL statement.

**TRIGGER: Triggers** are stored programs, which are automatically executed or fired when some events occur. Trigger automatically associated with DML statement, when DML statement execute trigger implicitly execute.

# Computer Networks

**Data Communication:** When we communicate, when we share information. It can be local or remote. Local communication occur face to face while remote communication take place over distance. Data communication are the exchange of data between two devices via the some form of transmission medium such as wire cable.

## Characteristics Of Data Communication

The Data communication must have major three fundamental characteristics:

- Delivery
- Accuracy
- Time Line

1) Where Delivery means system must delivered data to correct destination. Data must be received by the intended device.

2) Accuracy mean data delivered in accurately. Means that data should not be altered during transmission.

3) Time line means data should be delivered in time. When data in form of video audio is transfer as they produced at same time to other location is called real time transition.

## Types Of Data Communication

There are two types of data communication

- Serial communication
- Parallel communication

**OSI (Open Systems Interconnection)** is reference model for how applications can communicate over a network. A reference model is a conceptual framework for understanding relationships. The purpose of the OSI reference model is to guide vendors and developers so the digital communication products and software programs they create will interoperate, and to facilitate clear comparisons among communications tools. Most vendors involved in telecommunications make an attempt to describe their products and services in relation to the OSI model. And although useful for guiding discussion and evaluation, OSI is rarely actually implemented, as few network products or standard tools keep all related functions together in well-defined layers as related to the model. The TCP/IP protocols, which define the Internet, do not map cleanly to the OSI model.

## **Frame Relay:**

Frame Relay is a standardized wide area network technology that operates at the physical and logical link layers of OSI model. Frame relay originally designed for transport across Integrated Services Digital Network (ISDN) infrastructure, it may be used today in the context of many other network interfaces. Frame relay is an example of a packet switched technology. Packet switched network enables end stations to dynamically share the network medium and the available bandwidth.

An IP is a 32-bit number comprised of a host number and a network prefix, both of which are used to uniquely identify each node within a network. To make these addresses more readable, they are broken up into 4 bytes, or octets, where any 2 bytes are separated by a period. This is commonly referred to as dotted decimal notation. The first part of an Internet address identifies the network on which the host resides, while the second part identifies the particular host on the given network. This creates the two-level addressing hierarchy. All hosts on a given network share the same network prefix but must have a unique host number. Similarly, any two hosts on different networks must have different network prefixes but may have the same host number

## **Network Masks**

A network mask helps you know which portion of the address identifies the network and which portion of the address identifies the node. Class A, B, and C networks have default masks, also known as natural masks, as shown here:

Class A: 255.0.0.0

Class B: 255.255.0.0

Class C: 255.255.255.0

An IP address on a Class A network that has not been subnetted would have an address/mask pair similar to: 8.20.15.1 255.0.0.0. In order to see how the mask helps you identify the network and node parts of the address, convert the address and mask to binary numbers.

# **NETWORK SECURITY**

**Cryptography** can reformat and transform our data, making it safer on its trip between computers. The technology is based on the essentials of secret codes, augmented by modern mathematics that protects our data in powerful ways.

- Computer Security - generic name for the collection of tools designed to protect data and to thwart hackers
- Network Security - measures to protect data during their transmission
- Internet Security - measures to protect data during their transmission over a collection of interconnected networks.

## **Most commonly, the following tactics are used in to attack these applications:**

- SQL Injection
- XSS (Cross Site Scripting)
- Remote Command Execution

- Path Traversal

**1)SQL Injection:** SQL injection is a type of security exploit in which the attacker adds Structured Query Language (SQL) code to a Web form input box to gain access to resources or make changes to data. An SQL query is a request for some action to be performed on a database.

**IPSec:**IPsec (Internet Protocol Security) is a framework for a set of protocols for security at the network or packet processing layer of network communication. It is an Internet Engineering Task Force (IETF) standard suite of protocols that provides data authentication, integrity, and confidentiality as data is transferred between communication points across IP networks. IPSec provides data security at the IP packet level. A packet is a data bundle that is organized for transmission across a network, and it includes a header and payload (the data in the packet). IPSec emerged as a viable network security standard because enterprises wanted to ensure that data could be securely transmitted over the Internet. IPSec protects against possible security exposures by protecting data while in transit.

**Kerberos** is an authentication protocol and a software suite implementing this protocol. Kerberos uses symmetric cryptography to authenticate clients to services and vice versa. For example, Windows servers use Kerberos as the primary authentication mechanism, working in conjunction with Active Directory to maintain centralized user information. Other possible uses of Kerberos include allowing users to log into other machines in a local-area network, authentication for web services, authenticating email client and servers, and authenticating the use of devices such as printers. Kerberos is a protocol for authenticating service requests between trusted hosts across an untrusted network.

**Virus:** A computer **virus** is a program, script, or macro designed to cause damage, steal personal information, modify data, send e-mail, display messages, or some combination of these actions.

**SAML (Security Assertion Markup Language)** is an open standard for exchanging authentication information between a service provider and an identity provider (IdP). A third-party IdP is used to authenticate users and to pass identity information to the service provider in the form of a digitally signed XML(Extensible Mark-up language) document. Tableau Server is a service provider. Examples of IdPs include PingOne and OneLogin. SAML is designed for business-to-business (B2B) and business-to-consumer (B2C) transactions.

A firewall is a network security device that monitors incoming and outgoing network traffic and decides whether to allow or block specific traffic based on a defined set of security rules.

## DATA MINING AND DATA WAREHOUSING

**1)Data Mining:** Data mining, *the extraction of hidden predictive information from large databases*, is a powerful new technology with great potential to help companies focus on the most important information in their data warehouses. Data mining tools predict future trends and behaviors, allowing businesses to make proactive, knowledge-driven decisions. The automated, prospective analyses offered by data mining move beyond the analyses of past events provided by retrospective tools typical of decision support systems.

### The most commonly used techniques in data mining are:

- **Artificial neural networks:** Non-linear predictive models that learn through training and resemble biological neural networks in structure.
- **Decision trees:** Tree-shaped structures that represent sets of decisions. These decisions generate rules for the classification of a dataset. Specific decision tree methods include Classification and Regression Trees (CART) and Chi Square Automatic Interaction Detection (CHAID) .
- **Genetic algorithms:** Optimization techniques that use processes such as genetic combination, mutation, and natural selection in a design based on the concepts of evolution.
- **Nearest neighbor method:** A technique that classifies each record in a dataset based on a combination of the classes of the k record(s) most similar to it in a historical dataset (where  $k \geq 1$ ). Sometimes called the k-nearest neighbor technique.
- **Rule induction:** The extraction of useful if-then rules from data based on statistical significance.

### Types of OLAP:

- **MOLAP – Multidimensional OnLine Analytical Processes**  
MOLAP is the more traditional OLAP type. In MOLAP, both the source data and the aggregation calculations are stored in a multidimensional format. This type is the fastest option for data retrieval, but it also requires the most storage space. MOLAP systems are more optimized for fast query performance and retrieval of summarized data. The limitations in MOLAP are that it is not very scalable and can only handle limited amounts of data since calculations are predefined in the cube.
- **ROLAP – Relational OnLine Analytical Processes**  
ROLAP stores all data, including aggregations, in the source relational database. This type of storage is good for enterprises that need larger data warehousing. ROLAP uses an SQL reporting tool to query data directly from the data warehouse. ROLAP's advantages include better scalability, enabling it to handle huge amounts of data, and the ability to efficiently manage both numeric and textual data.

**Data Lake:** Data warehousing applies the structure on the way in, organizing it to fit the context of the database schema. Data lakes facilitate a much more fluid approach; they only add structures to data as it dispenses to the application layer. In storage, data lakes preserve the original structures or unstructured forms to remain; it is a Big Data storage and retrieval system that could conceivably scale upward indefinitely.

## SOFTWARE ENGINEERING

**Software Engineering** is an engineering approach for software development. The basic principle of software engineering is to use structured, formal and disciplined methods for building and using systems. The outcome of software engineering is an efficient and reliable software product.



Without using software engineering principles it would be difficult to develop large programs. In industry it is usually needed to develop large programs to accommodate multiple functions. A problem with developing such large commercial programs is that the complexity and difficulty levels of the programs increase exponentially with their sizes. Software engineering helps to reduce this programming complexity. Software engineering principles use two important techniques to reduce problem complexity: abstraction and decomposition. The principle of abstraction implies that a problem can be simplified by omitting irrelevant details. In other words, the main purpose of abstraction is to consider only those aspects of the problem that are relevant for certain purpose and suppress other aspects that are not relevant for the given purpose. Once the simpler problem is solved, then the omitted details can be taken into consideration to solve the next lower level abstraction, and so on. Abstraction is a powerful way of reducing the complexity of the problem.

The other approach to tackle problem complexity is decomposition. In this technique, a complex problem is divided into several smaller problems and then the smaller problems are solved one by one. However, in this technique any random decomposition of a problem into smaller parts will not help. The problem has to be decomposed such that each component of the decomposed problem can be solved independently and then the solution of the different components can be combined to get the full solution. A good decomposition of a problem should minimize interactions among various components.

### **Waterfall Model:**

The Waterfall Model was first Process Model to be introduced.

The Waterfall Model is a linear sequential flow. In which progress is seen as flowing steadily downwards (like a waterfall) through the phases of software implementation. This means that any phase in the development process begins only if the previous phase is complete. The waterfall approach does not define the process to go back to the previous phase to handle changes in requirement. The waterfall approach is the earliest approach that was used for software development.

**Black Box Testing:** Black Box Testing, also known as Behavioral Testing, is a software testing method in which the internal structure/ design/ implementation of the item being tested is not known to the tester. These tests can be functional or non-functional, though usually functional.

**Data Flow Diagram:** Data Flow Diagram (DFD) is a graphical representation of flow of data in an information system. It is capable of depicting incoming data flow, outgoing data flow, and stored data. The DFD does not mention anything about how data flows through the system. There is a prominent difference between DFD and Flowchart. The flowchart depicts flow of control in program modules. DFDs depict flow of data in the system at various levels. It does not contain any control or branch elements.

### **Types of DFD Data Flow Diagrams are either Logical or Physical.**

- Logical DFD - This type of DFD concentrates on the system process, and flow of data in the system. For example in a banking software system, how data is moved between different entities.
- Physical DFD - This type of DFD shows how the data flow is actually implemented in the system. It is more specific and close to the implementation.

**Software Quality Assurance (SQA)** is a set of activities for ensuring quality in software engineering processes (that ultimately result in quality in software products).

It includes the following activities:

- Process definition and implementation
- Auditing
- Training
- **Test Case:**

A **test case** is a document, which has a set of test data, preconditions, expected results and post conditions, developed for a particular test scenario in order to verify compliance against a specific requirement. Test Case acts as the starting point for the test execution, and after applying a set of input values, the application has a definitive outcome and leaves the system at some end point or also known as execution post condition.

Data Structure is a way of collecting and organising data in such a way that we can perform operations on these data in an effective way. Data Structures is about rendering data elements in terms of some relationship, for better organization and storage. In simple language, Data Structures are structures programmed to store ordered data, so that various operations can be performed on it easily.

**Data structure can be subdivided into two types:**

**1)Linear Data Structure**

**2)Non Linear Data Structure**

**Linked List:** A linked list is a linear collection of data elements, called nodes, where the linear order is given by means of pointers.

**Queue:**

- Abstract data type
- First element is inserted from one end called **REAR**(also called tail), and the deletion of existing element takes place from the other end called as **FRONT**(also called head). This makes queue as FIFO data structure, which means that element inserted first will also be removed first.

**There are many types of Sorting techniques, differentiated by their efficiency and space requirements. Following are some sorting techniques which we will be covering in next sections.**

1. Bubble Sort
2. Insertion Sort
3. Selection Sort
4. Quick Sort
5. Merge Sort
6. Heap Sort

Algorithm	Best Time Complexity	Average Time Complexity	Worst Time Complexity	Worst Space Complexity
Linear Search	$O(1)$	$O(n)$	$O(n)$	$O(1)$
Binary Search	$O(1)$	$O(\log n)$	$O(\log n)$	$O(1)$
Bubble Sort	$O(n)$	$O(n^2)$	$O(n^2)$	$O(1)$
Selection Sort	$O(n^2)$	$O(n^2)$	$O(n^2)$	$O(1)$
Insertion Sort	$O(n)$	$O(n^2)$	$O(n^2)$	$O(1)$
Merge Sort	$O(n \log n)$	$O(n \log n)$	$O(n \log n)$	$O(n)$
Quick Sort	$O(n \log n)$	$O(n \log n)$	$O(n^2)$	$O(\log n)$
Heap Sort	$O(n \log n)$	$O(n \log n)$	$O(n \log n)$	$O(n)$
Bucket Sort	$O(n+k)$	$O(n+k)$	$O(n^2)$	$O(n)$
Radix Sort	$O(nk)$	$O(nk)$	$O(nk)$	$O(n+k)$
Tim Sort	$O(n)$	$O(n \log n)$	$O(n \log n)$	$O(n)$
Shell Sort	$O(n)$	$O((\log(n))^2)$	$O((\log(n))^2)$	$O(1)$

**The following computer algorithms are based on divide-and-conquer programming approach**

- Merge Sort
- Quick Sort
- Binary Search
- Strassen's Matrix Multiplication
- Closest pair (points)

**AVL Tree:** One of the more popular balanced trees, known as an AVL tree in Data Structures, was introduced in 1962 by **Adelson-Velski and Landis**. An Avl trees is a binary search tree in which for every node in the tree, The height of the left and right Sub trees differ by at most 1.

# OPERATING SYSTEM

**An operating system** is a program that manages the computer hardware. It also provides a basis for application programs and acts as an intermediary between the computer user and the computer hardware. An amazing aspect of operating systems is how varied they are in accomplishing these tasks. Mainframe operating systems are designed primarily to optimize utilization of hardware. Personal computer (PC) operating systems support complex games, business applications and everything in between. Operating systems for handheld computers are designed to provide an environment in which a user can easily interface with the computer to execute programs. Thus some operating system are designed to be convenient others to be efficient and others some combination of the two. The operating system acts as a manager of the above resources and allocates them to specific programs and users as necessary for their task. Therefore operating system is the resource manager i.e. it can manage the resource of a computer system internally. The resources are processor, memory, files, and I/O devices.

## Context Switch

- When CPU switches to another process, the system must save the state of the old process and load the saved state for the new process via a context switch.
- Context of a process represented in the PCB
- Context-switch time is overhead; the system does no useful work while switching. The more complex the OS and the PCB -> longer the context switch
- Time dependent on hardware support. Some hardware provides multiple sets of registers per CPU -> multiple contexts loaded at once

## **Process vs Thread?**

The typical difference is that threads within the same process run in a shared memory space, while processes run in separate memory spaces.

Threads are not independent of one other like processes as a result threads shares with other threads their code section, data section and OS resources like open files and signals. But, like process, a thread has its own program counter (PC), a register set, and a stack space.

## ***Advantages of Thread over Process***

1. *Responsiveness*: If the process is divided into multiple threads, if one thread completed its execution, then its output can be immediately responded.
2. *Faster context switch*: Context switch time between threads is less compared to process context switch. Process context switch is more overhead for CPU.
3. *Effective Utilization of Multiprocessor system*: If we have multiple threads in a single process, then we can schedule multiple threads on multiple processor. This will make process execution faster.
4. *Resource sharing*: Resources like code, data and file can be shared among all threads within a process.

Note : stack and registers can't be shared among the threads. Each thread have its own stack and registers.

**Semaphore:**

A semaphore is hardware or a software tag variable whose value indicates the status of a common resource. Its purpose is to lock the resource being used. A process which needs the resource will check the semaphore for determining the status of the resource followed by the decision for proceeding. In multitasking operating systems, the activities are synchronized by using the semaphore techniques.

**Deadlock:** Is it a state where two or more operations are waiting for each other, say a computing action 'A' is waiting for action 'B' to complete, while action 'B' can only execute when 'A' is completed. Such a situation would be called a deadlock. In operating systems, a deadlock situation is arrived when computer resources required for complete of a computing task are held by another task that is waiting to execute. The system thus goes into an indefinite loop resulting into a deadlock. The deadlock in operating system seems to be a common issue in multiprocessor systems, parallel and distributed computing setups.

**Dijkstra Banking Algorithm :**

The Banker's Algorithm is a strategy for **deadlock** prevention. In an operating system, deadlock is a state in which two or more processes are "stuck" in a **circular wait** state. All deadlocked processes are waiting for resources held by other processes. Because most systems are **non-preemptive** (that is, will not take resources held by a process away from it), and employ a **hold and wait** method for dealing with system resources (that is, once a process gets a certain resource it will not give it up voluntarily), deadlock is a dangerous state that can cause poor system performance.

# UNIX/Linux

**UNIX** is an operating system which was first developed in the 1960s, and has been under constant development ever since. By operating system, we mean the suite of programs which make the computer work. It is a stable, multi-user, multi-tasking system for servers, desktops and laptops.

UNIX systems also have a graphical user interface (GUI) similar to Microsoft Windows which provides an easy to use environment. However, knowledge of UNIX is required for operations which aren't covered by a graphical program, or for when there is no windows interface available, for example, in a telnet session.

## **cd**

As you might already have guessed, the **cd** command changes directories. It's a very common navigation command that you'll end up using, just like you might have done in MS-DOS.

You must put a space between **cd** and the `".."` or else it won't work; Linux doesn't see the two dots as an extension to the **cd** command, but rather a different command altogether. It'll come to make sense if it doesn't already.

## **ls**

The **ls** letters stand for **list**. It basically works the same way as the **dir** command in DOS. Only being a Unix command, you can do more with it. :-)

Typing **ls** will give you a listing of all the files in the current directory. If you're new to Linux, chances are that the directories you are commonly in will be empty, and after the **ls** command is run, you aren't given any information and will just be returned to the command prompt (the shell).

There are "hidden" files in Linux, too. Their file names start with a dot, and doing a normal **ls** won't show them in a directory. Many configuration files start with a dot on their file names because they would only get in the way of users who would like to see more commonly used items. To view hidden files, use the **-a** flag with the **ls** command, i.e. **ls -a**.

To view more information about the files in a directory, use the **-l** flag with **ls**. It will show the file permissions as well as the file size, which are probably what are the most useful things to know about files.

You might occasionally want to have a listing of all the subdirectories, also. A simple **-R** flag will do, so you could look upon **ls -R** as a *rough* equivalent of the **dir /s** command in MS-DOS.

You can put flags together, so to view all the files in a directory, show their permissions/size, and view all the files that way through the subdirectories, you could type **ls -laR**.

**pwd**

This command simply shows what directory you're in at the moment. It stands for "Print Working Directory". It's useful for scripting in case you might ever want to refer to your current directory.

**Read this Completely before  
an exam and**

**Download this ebook from:  
[studyregular.in](http://studyregular.in)**

## **C/C++**

The **C language** was developed in 1972 by Dennis Ritchie at Bell Telephone laboratories, primarily as a systems programming language. That is, a language to write operating systems with. Ritchie's primary goals were to produce a minimalistic language that was easy to compile, allowed efficient access to memory, produced efficient code, and did not need extensive run-time support. Thus, for a high-level language, it was designed to be fairly low-level, while still encouraging platform-independent programming.

C++ (pronounced see plus plus) was developed by Bjarne Stroustrup at Bell Labs as an extension to C, starting in 1979. C++ adds many new features to the C language, and is perhaps best thought of as a superset of C, though this is not strictly true as C99 introduced a few features that do not exist in C++. C++'s claim to fame results primarily from the fact that it is an object-oriented language. As for what an object is and how it differs from traditional programming methods, well, we'll cover that in chapter 8 (Basic object-oriented programming).

C++ is an **Object Oriented Programming language** but is not purely Object Oriented. Its features like Friend and Virtual, violate some of the very important OOPS features, rendering this language unworthy of being called completely Object Oriented. Its a middle level language.

**Storage class** of a variable defines the lifetime and visibility of a variable. Lifetime means the duration till which the variable remains active and visibility defines in which module of the program the variable is accessible. There are five types of storage classes in C++. They are:

1. Automatic
2. External
3. Static
4. Register
5. Mutable

### **Function:**

A function is a block of code that performs some operation. A function can optionally define input parameters that enable callers to pass arguments into the function. A function can optionally return a value as output. Functions are useful for encapsulating common operations in a single reusable block, ideally with a name that clearly describes what the function does. Every C++ program has at least one function, which is **main()**, and all the most trivial programs can define additional functions.

### **OOPs (Object Oriented Programming System)**

**Object** means a real word entity such as pen, chair, table etc. **Object-Oriented Programming** is a methodology or paradigm to design a program using classes and objects. It simplifies the software development and maintenance by providing some concepts:

- Object
- Class
- Inheritance
- Polymorphism
- Abstraction
- Encapsulation

## **JAVA**

Java technology is both a programming language and a platform. The Java programming language is a high-level language. In the Java programming language, all source code is first written in plain text files ending with the .java extension. Those source files are then compiled into .class files by the javac compiler. A .class file does not contain code that is native to your processor; it instead contains



bytecodes the machine language of the Java Virtual Machine (Java VM). The java launcher tool then runs your application with an instance of the Java Virtual Machine.

James Gosling - founder of java.

Java is a high-level programming language originally developed by Sun Microsystems and released in 1995.

Java team members (also known as Green Team), initiated a revolutionary task to develop a language for digital devices such as set-top boxes, televisions etc.

For the green team members, it was an advance concept at that time. But, it was suited for internet programming. Later, Java technology was incorporated by Netscape.

Currently, Java is used in internet programming, mobile devices, games, e-business solutions etc. There are given the major points that describes the history of java.

### **JVM:**

Java virtual Machine(JVM) is a virtual Machine that provides runtime environment to execute java byte code. The JVM doesn't understand Java type, that's why you compile your \*.java files to obtain \*.class files that contain the bytecodes understandable by the JVM. JVM controls execution of every Java program. It enables features such as automated exception handling, Garbage-collected heap.

**JRE** : The Java Runtime Environment (JRE) provides the libraries, the Java Virtual Machine, and other components to run applets and applications written in the Java programming language. JRE does not contain tools and utilities such as compilers or debuggers for developing applets and applications.

**JDK** : The JDK also called Java Development Kit is a superset of the JRE, and contains everything that is in the JRE, plus tools such as the compilers and debuggers necessary for developing applets and applications.

**JIT**: It is the set of programs developed by SUN Micro System and added as a part of JVM, to speed up the interpretation phase.

In the older version of java compilation phase is so faster than interpretation phase. Industry has complained to the SUN Micro System saying that compilation phase is very faster and interpretation phase is very slow.

To solve this issue, SUN Micro System has developed a program called JIT (just in time compiler) and added as a part of JVM to speed up the interpretation phase. In the current version of java interpretation phase is so faster than compilation phase. Hence java is one of the highly interpreted programming languages.

## Object Cloning:

The **object cloning** is a way to create exact copy of an object. For this purpose, clone() method of Object class is used to clone an object. The **java.lang.Cloneable interface** must be implemented by the class whose object clone we want to create. If we don't implement Cloneable interface, clone() method generates **CloneNotSupportedException**.

The **clone() method** is defined in the Object class.

*Java StringBuffer class is used to create mutable (modifiable) string. The StringBuffer class in java is same as String class except it is mutable i.e. it can be changed. Java StringBuffer class is thread-safe i.e. multiple threads cannot access it simultaneously. So it is safe and will result in an order.*

**Java Networking:** Java is a premier language for network programming. **java.net** package encapsulate large number of classes and interface that provides an easy-to use means to access network resources.

Socket is foundation of modern networking, a socket allows single computer to serve many different clients at once. Socket establishes connection through the use of port, which is a numbered socket on a particular machine. Socket communication takes place via a protocol. Socket provides communication mechanism between two computers using TCP. There are two kind of TCP sockets in Java. One is for server and other is for client.

- **ServerSocket** is for servers.
- **Socket** class is for client.

Java programs can be divided into two categories :-

- i) **Applications and**
- ii) **Applets.**

Applications are the programs that contain main( ) method and applets are the programs that do not contain main( ) method. Applications can be executed with a Java interpreter from the command line( with java command). Applets need a browser to execute.

**Applet:** Applet is a Java program that runs on a browser.

**Cloud Computing:** Cloud computing is a type of computing that relies on *sharing computing resources* rather than having local servers or personal devices to handle applications.

In cloud computing, the word cloud (also phrased as "the cloud") is used as a metaphor for "*the Internet*," so the phrase *cloud computing* means "a type of Internet-based computing," where different services — such as servers, storage and applications —are delivered to an organization's computers and devices through the Internet.

### **Types of cloud services: IaaS, PaaS, SaaS**

Most cloud computing services fall into three broad categories: infrastructure as a service (IaaS), platform as a service (PaaS) and software as a service (SaaS). These are sometimes called the cloud computing stack, because they build on top of one another. Knowing what they are and how they are different makes it easier to accomplish your business goals.

#### **Infrastructure-as-a-service (IaaS)**

The most basic category of cloud computing services. With IaaS, you rent IT infrastructure—servers and virtual machines (VMs), storage, networks, operating systems—from a cloud provider on a pay-as-you-go basis.

#### **Platform as a service (PaaS)**

Platform-as-a-service (PaaS) refers to cloud computing services that supply an on-demand environment for developing, testing, delivering and managing software applications. PaaS is designed to make it easier for developers to quickly create web or mobile apps, without worrying about setting up or managing the underlying infrastructure of servers, storage, network and databases needed for development.

#### **Software as a service (SaaS)**

Software-as-a-service (SaaS) is a method for delivering software applications over the Internet, on demand and typically on a subscription basis. With SaaS, cloud providers host and manage the software application and underlying infrastructure and handle any maintenance, like software upgrades and security patching.

### **Management Information System:**

MIS is the use of information technology, people, and business processes to record, store and process data to produce information that decision makers can use to make day to day decisions.

MIS is the acronym for **Management Information Systems**. In a nutshell, MIS is a collection of systems, hardware, procedures and people that all work together to process, store, and produce information that is useful to the organization.

### **The need for MIS**

The following are some of the justifications for having an MIS system

- **Decision makers need information to make effective decisions.** Management Information Systems (MIS) make this possible.
- **MIS systems facilitate communication within and outside the organization** – employees within the organization are able to easily access the required information for the day to day operations. Facilitates such as Short Message Service (SMS) & Email make it possible to communicate with customers and suppliers from within the MIS system that an organization is using.
- **Record keeping** – management information systems record all business transactions of an organization and provide a reference point for the transactions.

**Subscribe Study Regular YouTube Channel and  
Join Our Facebook Group  
For MCQ and Understand these Topic Concepts**