**Course: Advanced Bioinformatics**

**Module title: DB Architecture**

**Module no. :11**

Database System Architectures

There are four major types of architectures for database systems. Details of these are provided in this module.

* Centralized
* Client-Server Systems
* Parallel Systems
* Distributed Systems

**Centralized System:** Run on a single computer system and do not interact with other computer systems. General-purpose computer system: one to a few CPUs and a numberν of device controllers that are connected through a common bus that provides access to shared memory.

Single-user system (e.g., personal computer or workstation): desk-topν unit, single user, usually has only one CPU and one or two hard disks; the OS may support only one user.

Multi-user system: more disks, more memory, multiple CPUs, and aν multi-user OS. Serve a large number of users who are connected to the system vie terminals. Often called server systems.

**Client-Server Systems:** Server systems satisfy requests generated at m client systems.

Database functionality can be divided into:

 Back-end: manages access structures, query evaluation andoptimization, concurrency control and recovery.

 Front-end: consists of tools such as forms, report-writers, andgraphical user interface facilities.

The interface between the front-end and the back-end is through SQL orthrough an application program interface.

Advantages of replacing mainframes with networks of workstations orpersonal computers connected to back-end server machines:

 Better functionality for the cost

 Flexibility in locating resources and expanding facilities

 Better user interfaces

 Easier maintenance

**Parallel database systems** consist of multiple processors and multipledisks connected by a fast interconnection network.

A coarse-grain parallel machine consists of a small number ofpowerful processors

A massively parallel or fine grain parallel machine utilizesthousands of smaller processors.

Two main performance measures:

 **Throughput** --- the number of tasks that can be completed in agiven time interval

 **Response time** --- the amount of time it takes to complete a singletask from the time it is submitted

**Distributed databases**

 Same software/schema on all sites, data may be partitionedamong sites

 Goal: provide a view of a single database, hiding details of distribution

 Heterogeneous distributed databases

 Different software/schema on different sites

 Goal: integrate existing databases to provide useful functionality

 Differentiate between local and global transactions

 A local transaction accesses data in the single site at which thetransaction was initiated.

 A global transaction either accesses data in a site different fromthe one at which the transaction was initiated or accesses data inseveral different sites.