1. **Define what is meant by the term distributed system.**

A distributed system is a collection of autonomous computing elements that appears to its users as a single coherent system. These computing elements need communicate through the network and coordinate work in order to complete common tasks. The purpose of distributed system is to use more machines to process more data.

Describe the relationship between computer networks and distributed systems.

Computer networks define connectivity between computers, and distributed systems focus on the services base on top of these networks. Therefore, the physical structure of the distributed system and the computer network are basically the same. Computer networks provide a technical foundation for distributed system , and distributed systems are an advanced stage of computer network technology development.

1. **A distributed system typically satisfies certain criteria. What are these criteria?**

A distributed system typically satisfies four criteria:

• Multiple processes – more than one sequential process. Each process should have an independent thread of control.

• Interprocess communication – processes communication using messages. Message links are known as channels.

• Disjoint address spaces – Process have disjoint address spaces. Shared memory multiprocessor is not considered a true representative of a distributed system.

• Collective goal – processes must interact to meet a common goal.

1. **Give three reasons for the growing importance of distributed systems.**

* Resource sharing - this includes hardware (e.g. printer, computation power), software (e.g. Google Docs) and information.

• Geographical distributed environments - in many situations, the computing environment itself is geographically distributed (e.g. banking, the WWW).

• Speed up - the speed of computation in traditional uniprocessors is fast approaching the physical limit. More computational power can be achieved using multiple processors.

• Fault tolerance - when a fraction of the processors fail, the remaining processors keep the application running. Allows for graceful degradation.

1. **Replica management is a challenge when designing a distributed system. Describe this challenge.**

The challenge is that replication requires that the replicas be appropriately updated.

Update cannot be done instantaneously; therefore, inconsistent replicas may occur.

1. **What is a failure?**

A failure is that the system is not adequately providing the services it was designed for.

1. **Name and describe three types of failures which can occur in a distributed system.**

Crash failure. It can occur halts, but is working correctly until it halts

Omission failure, it can occur fail to respond to incoming requests, which includes 2 parts: the receive omission which fails to receive incoming messages and the send omission which fails to send messages.

Timing failure, it can occur that the response lies outside a specified time interval.

Response failure, it can occur that response is incorrect, which includes 2 parts: the value failure which made the value of the response is wrong and the state-transition failure which deviates from the correct flow of control.

1. **Synchronisation of processes is a challenge when designing a distributed system. Describe this challenge. Illustrate this challenge with an example.**

Each process in a distributed system has a local clock. If these clocks are all correct, the system can exhibit synchronous behavior. Unfortunately, this is difficult to achieve because drift in local clocks is inevitable. One solution is to use a time server that keeps all local clocks synchronized (resynchronization).

Example: In Unix systems the make command is used to compile new or modified code without the need to recompile unchanged code. The make command uses the clock of the machine it runs on to determine which source files need to be recompiled. If the sources reside on a separate file server and the two machines have unsynchronized clocks, the make program may not produce the correct results.

1. **What does it mean to say that a distributed system is size scalable?**

System is size scalable if can add more users/processes without impairing performance. Additional resources required to cope with the increased size should be manageable.

1. **What does it mean to say that a distributed system is geographical scalable?**

System is geographical scalable if users/resources lie far apart but communication delays hardly noticed.

1. **What is middleware and what is the purpose of middleware?**

Middleware is a type of computer software that connects software components and applications. It includes a set of services.

The purpose is to facilitate multiple software running on one or more machines to interact through the network. The interoperability provided by this technology promotes the evolution of a consistent distributed architecture, which is commonly used to support and simplify those complex distributed applications

1. **What is the aim distribution transparency?**

Aim of distribution transparency is to make the distribution of processes and resources transparent, that is invisible, to users and applications.

1. **Name and describe three types of distribution transparency.**

Access transparency: Hide differences in data representation and how an

object is accessed.

Location transparency: Hide where an object is located.

Migration transparency: Hide that an object may be moved to another

location.

Replication transparency: Hide that an object is replicated.

1. **Describe the relationship between distributed and parallel systems?**

When a task needs to be executed, the distributed system and the parallel system will split the task into several sub-tasks, and then allocate them to all machines / threads for execution according to the appropriate scheduling algorithm, and finally integrate the results together and finished this general task. The different is just that the distributed system assigns tasks to each computer, while the parallel system assigns tasks to different threads.

1. **Define what is meant by the term cluster computing.**

Cluster computing is collection of compute nodes (computers) controlled and accessed by means of a single master node which have similar hardware and OS. And it connected by high-speed local-area network and each node performs the same task.

1. **What is a sensor network?**

Sensor network consists of large number of relatively small nodes, each equipped with sensors and often actuators. Many sensor networks use wireless communication, and the nodes are often battery powered.

1. **Define what is meant by the term cloud computing.**

Cloud computing is a type of distributed computing, which refers to the decomposition of huge data computing processing programs into countless small programs through the network "cloud", and then through a system composed of multiple servers for processing and analysis The applet gets the result and returns it to the user.

1. **Cloud computing employs a measured service model. Define this model.**

In cloud computing clients outsource their software usage, data storage and computing

infrastructure to third-party companies (e.g. Amazon). The main advantage of this to the

client is that they need not acquire or maintain expensive hardware, software or technical

staff. The client pays the cloud service provider for this service on a pay-per-use basis; this

is known as a measured service model.

1. **There are three main types of cloud computing. What are these?**

There are three main types of cloud computing services:

Software-as-a-service (SaaS) – allows the usage of cloud apps e.g. Google Docs.

Platform-as-a-service (PaaS) – virtualized environment with OS installed is rented.

Infrastructure-as-a-service (IaaS) – includes hardware, storage, IP addresses and firewalls.

1. **Mutual exclusion is a challenge when designing a distributed system. Describe this challenge.**

In many cases different processes will want to simultaneously access the same resource.

Concurrent accesses may corrupt the resource or make it inconsistent. Solutions are needed to grant mutual exclusive access by processes. Algorithms for achieving mutual exclusion include centralized coordinator and token-ring algorithms.