DHANALAKSHMI SRINIVASAN
INSTITUTE OF TECHNOLOGY
(Approved by AICTE, New Delhi & Affiliated to Anna University)
NH - 45, Trichy - Chennai Trunk Road,
SAMAYAPURAM, TRICHY - 621 112.
E.mail:dsit2011@gmail.com Website:www.dsit.ac.in

✆ STD - 0431 - 2670142
FAX - 0431 - 2670143

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
QUESTION WITH ANSWER

UNIT I

1.What is meant distributed system?

    1. We define a distributed system as a collection of autonomous computers linked by a network, with software designed to produce an integrated computing facility.

    2. A system in which hardware or software components located at networked computers communicate and coordinate their actions only by message passing.

    3. A collection of two or more independent computers which coordinate their processing through the exchange of synchronous or asynchronous message passing.

    4. A collection of independent computers that appear to the users of the system as a single computers.

2. What are the significance of distributed system?

    a. Concurrency of computers.

    b. No global clock.

    c. Independent failures.

3. Why we do you need distributed system?

    a. **Functional distribution**: Computers have different functional capabilities (i.e., sharing of resources with specific functionalities).

    b. **Load distribution/balancing**: Assign tasks to processors such that the overall system performance is optimized.

    c. **Replication of processing power**: Independent processors working on the same task.

    d. Distributed system consisting of collections of microcomputers may have processing powers that no supercomputer will ever achieve.

    e. **Physical separation**: Systems that rely on the fact that computers are physically separated (e.g., to satisfy reliability requirements).

    f. **Economics**: Collections of microprocessors offer a better price/performance ratio than large mainframes.mainframes:10 times faster, 1000 times as expensive.

4. Examples of distributed system?

    a. Internet

    b. Intranet

    c. Mobile and ubiquitous computing.

5. What is meant by location aware computing?

    Mobile computing is the performance of computing tasks while the users are on the move and away from their residence intranet but still provided with access to resources via the devices they carry with them. They can continue to access the intranet, they can continue to access resources in their home intranet, and there is increasing provision for users to utilize resources such as printers that are conveniently nearby as they move around. This is known as location aware computing.

6. What are the two type of resource sharing?

      a. Hardware sharing: Printers. plotters and large disks and other peripherals are shared to reduce costs.

      b. Data sharing is important in many applications:

            1. Software developers working in a team need to access each other's code and share the same development tools.

            2. Many commercial applications enable users to access shared data objects in a single active database.

            3. The rapidly growing area of group-ware tools enables users to cooperate with

in a network.

7. List the importance of data sharing?

- Software developers working in a team need to access each other's code and share the same development tools.
- Many commercial applications enable users to access shared data objects in a single active database.
- The rapidly growing area of group- ware tools enables users to cooperate with in a network.

8. Write the technological components of web?

- HTML
- HTTP-request-reply protocol
- URL's

9. List the distributed systems challenges?

      a. Heterogeneity: standards and protocols; middleware; virtual machine;

      b. Openness: publication of services; notification of interfaces;

      c: Security: firewalls; encryption;

      d. Scalability: replication; caching; multiple servers;

      e. Failure Handling. failure tolerance; recover/roll-back; redundancy;

      f. Concurrency. concurrency control to ensure data consistency.

      g. Transparency. Middleware; location transparent naming; anonymity

10. What are the three components of security?

      Security for information resources has three components:

- **Confidentiality:** production against disclosure to unauthorized individuals.
- **Integrity:** production against or corruption.
- **Availability:** production against interference with the means to access the resources.

11. What is the use of firewall?

A firewall can be used to form a barrier around an intranet to protect it from outside users but does not deal with ensuring the appropriate use of resources by users within the intranet.

12. What are the security challenges? List them.

      a. **Denial of service attacks:** Another security problem is that the user may wish to disrupt a service for some reason. This can be achieved by bombarding the service with such a large number of pointless requests that the serious users are unable to use it. This is called a denial of service attack

and there are many on well known web services.

b. **Security of mobile code:** Mobile codes needed to be handled with care. PC users sometimes send executable files as email attachments to be run by the recipient, but a recipient will not be able to run it.

13. List the challenges to be considered for designing scalable distributed system?
- Controlling the cost of physical resources
- Controlling the performance loss

- Preventing software resources running out
- Avoiding performance bottlenecks.

14. What are the types of transparencies?
a. Access transparency
b. Location transparency
c. Concurrency transparency
d. Replication transparency
e. Failure transparency
f. Mobility transparency
g. Performance transparency
h. Scaling transparency

15. What are the failures detected in DS?
Masking failures
Tolerating failures
Recovery from failures

16. List the key design goals of DS?
a. Performance
b. Reliability
c. Scalability
d. Consistency
e. Security

17. List the technical design goals of DS?
a. Naming
b. Communication
c. Software structure
d. Workload allocation
e. Maintenance of consistency.

18 . What is the use of multicast?
a. **Locating an object:** A process multicasts a message containing a name of a resource to a group of server processes. Only the process that holds the resource responds to the message.
b. **Fault tolerance:** A process multicasts its request to a group of identical server processes. The group of servers can continue to provide their service even if one of its members fails.
c. **Multiple update:** Used for example in video conferencing with multiple participants.

19. Write the models used in workload allocation?

The following four models for workload allocation are presented here.
- Processor pool model
- Shared memory multiprocessors
- Parallel virtual machines
- Distributed shared memory.

20. What is meant by PVM?

PVM is an integrated set of software tools and libraries that emulates a general-purpose, flexible heterogeneous concurrent computing framework on interconnected computers of varied architecture. The overall objective of the PVM system is to enable a collection of computers to be used cooperatively for concurrent or parallel computation.

21. List the types of consistencies in DS?
- Update consistency
- Replication consistency
- Cache consistency
- Failure consistency
- Clock consistency

22. List the user requirements used in design of DS?
- **Functionality:** what should the system do for the users.
- **Quality of service:** issue containing performance, reliability and security.
- **Reconfigurability:** the need to accommodate changes without causing disruption of the actual service.

23. List the main types of architectural model?

  a. Software architecture.

  b. System architecture.
- Client server model
- Services provided by multiple servers.
- Proxy servers and cache.
- Peer processes.

24. Enumerate the key features of spontaneous networking?

  a. **Easy connection to a local network:** Wireless links avoid the need for pre-installed cabling and avoid the inconvenience and reliability issues surrounding plugs and sockets.

  b. **Easy integration with local services:** Devices are able to find themselves inserted into existing networks of devices discover automatically what services are provided there, with no special configuration actions by the user.

25. How the fundamental models are categorized?

  a. Interaction

  b. Failure

  c. Security

26. What is asynchronous DS?

1. Many distributed systems, such as the Intranet, qualify as asynchronous system.
2. An asynchronous distributed system is one in which there are no bounds on:

  1. **Process execution speeds-**for example, one process step may take only a picoseconds and another a century; all that can be said is that each step may take an arbitrarily long time.

2. **Message transmission delays-**for example, one message from process A to process B may be delivered in negligible time and another may take several years. In other words, a message may be received after an arbitrarily long time.

3. **Clock drift rates-** again, the drift rate of a clock is arbitrary.

27. What is omission failure?

The faults classified as omission failures refer to cases when a process or communication channel fails to perform actions that it is supposed to do.

UNIT II

1. Draw the Middleware Architecture.

Application
RMI,RPC and events
Request reply protocol
External data representation
Operating system

2. What are the benefits of programming with interface in DS?

I.   As with any form of modular programming, programmers are concerned only with the abstraction offered by the service interface and need not be aware of implementation details.

II. In potentially heterogeneous distributed systems , programmers also do not need to know the programming language or underlying platform used to implementation service.

III. This approach provides natural support for software evolution in this implementation can change as long as the interface remains the same.

3. Define IDL.

Interface Definition Languages (IDLs) are designed to allow procedures implemented in different languages to invoke one another. An IDL provides a notation for defining interfaces in which each of the parameters of an operation may be described as for input or output in addition to having its type specified.

4. List the used of IDL in web services.

The concept of an IDL was initially developed for RPC systems but applies equally to RMI and also web service. Some of them are:

I.   Sun XDR as an example of an IDL for RPC

II.   CORBA IDL as an example of an IDL for RMI

III.   The web service Description Language (WSDL),which is designed for an Internet wide RPC supporting web service.

5. What are the main choices to be considered in design of RMI?

RMI invocation semantics

a. Retry-reply protocols, where we showed that doOperation can be implemented in different ways to provide different guarantees.

b. The main choices are:
- Retry request message:

- Duplicate filtering

- Retrasmission of results

6. Sketch the RMI reply-request message structure.

Message Type
Request Id
Object Reference
Method Id
arguments

7. What is persistent object store?

An object that is guaranteed to live between activations of processes is called a persistet object. Persistent objects are generally managed by persistent object stores, which store their state in a marshaled from on disk. In generally a persistent object store will manage very large numbers of persistent object which are stored in a disk or in a database until they are needed.

8. What is event notification?

The distributed event based system extend the local event model by allowing multiple object at different location to be notified of events takes place at an object. They use the publish subscribe paradigm. A publish subscribe system is a system where publishers publish structured events to an event service and subscriber express interest in particular events through subscriptions which can be arbitrary patterns over the structure events.

9. **List the example of publish subscribe system.**

> Publish-subscribe system is used in a wide variety of application domains particularly those related to a large scale dissemination of events.

- Financial information systems.
- Other area with live feeds of real time data(including RSS feeds)
- Supports for cooperative working where a number of participants need to be informed of events of shared interest
- Support for computing including the management of events from the infrastructure

10. **List the characteristics of publish-subscribe system.**

> **Heterogeneity:** When event notifications are used as a means of communication components in a distributed system that were not designed to interoperate can be made to work together. All that is required is that event generating objects publish the types of events they offer and that other objects subscribe to patterns of events and provide an interface for receiving and dealing with the resultant notification.

11. **Define callbacks.**

> The general idea behind callbacks is that instead of clients polling the server to find out whether some event has occurred, the server should inform its clients whenever that event occurs. The term callback is used to refer to a server's action of notifying clients about an event.

12. **How callback is implemented in RMI.**

> Callback can be implemented in RMI as follows;

i. The client creates a remote object that implements an interface that contains method for the server to call. We refer to that as a callback object.

ii. The server provides an operation allowing interest clients to inform it of the remote object references of their callback objects. It records these in a list.

iii. Whenever an event of interest occurs, the server calls the interested clients, For example, the whiteboard server would call its clients whenever a graphical object is added.

13. Define process.

A process consists of an execution environment together with one or more threads.

14. Define thread.

A thread is the operating system abstraction of an activity (the term derives from the phrase 'thread of execution'). An execution environment is the unit of resource management: a collection of local kernel managed resources to which its threads have access.

15. Define Unix address space.

This representation of an address space as a sparse set of disjoint regions is a generalization of the UNIX address space, which has three regions: a fixed, unmodifiable text region containing program code; a heap, part of which is initialized by values stored in the program's binary file , and which is extensible towards higher virtual addresses ; and a stack, which is extensible towards lower virtual addresses.

16. List the architecture of multi threaded server.

- Working pool Architecture
- Thread-per-request Architecture;
- Thread-per-connection Architecture
- Thread-per-object Architecturre:

17. Compare process and threads.

a. Creation a new thread within an existing process is cheaper than creating a process.

b. More importantly switching to a different thread within the same process is cheaper than switching between threads belonging to different processes.

c. Threads within a process may share data and other resources conveniently and efficiently compared with separate processes.

d. But by the same token threads within processes are not protected from one another.

18. List out the design issues for RPC?

- The style of programming promoted by RPC programming with interfaces;
- The call semantics associated with RPC;
- The key issue of transparency and how it relates to remote procedure calls.

19. What is indirect communication?

Indirect communication is defined as communication between entities in adistributed system through an intermediary with no direct coupling between the sender and the receiver(s).

20. List the different switching scheme?

- Broadcast
- Circuit switching
- Packet switching
- Frame relay

Unit III

1. What is peer to peer system?
Peer-to-peer systems aim to support useful distributed services and applications using data and computing resources available in the personal computers and workstations that are present in the
Internet and other networks in ever-increasing numbers.

2. What is goal of peer to peer system?
The goal of peer-to-peer systems is to enable the sharing of data and resources on a very

large scale by eliminating any requirement for separately managed servers and their

associated infrastructure

3. What are the characteristics of peer to peer system? MAY/JUNE 2016
Their design ensures that each user contributes
resources to the system.
•   Although they may differ in the resources that they contribute, all the nodes in a peer-to-peer system have the same functional capabilities and responsibilities.
• Their correct operation does not depend on the existence of any centrally administered systems.
• They can be designed to offer a limited degree of anonymity to the providers and users of resources.

4. **What is napster file system? NOV/DEC 2018**

The need for and the feasibility of a peer-to-peer solution were first demonstrated by the

Napster file sharing system which provided a means for users to share files. Napster became

very popular for music exchange. At its peak, several million users were registered and

thousands were swapping music files simultaneously.

5. What is the role of routing overlays in peer to peer system? APR/MAY 2017
Peer-to-peer systems usually store multiple
replicas of objects to ensure availability. In that case, the routing overlay maintains

Knowledge of the location of all the available replicas and delivers requests to the nearest

'live' node (i.e. one that has not failed) that has a copy of the relevant object.

6. What are the tasks performed by routing overlay?
- o Insertion of objects
- o Deletion of objects
- o Node addition and removal

7. What are the case studies used in overlay?NOV/DEC 2017
- o **Pastry** is the message routing infrastructure deployed in several applications including PAST.
- o **Tapestry** is the basis for the Ocean Store storage system

8. What is meant by distributed file system?

A distributed file system enables programs to store and access remote files exactly as they do local ones, allowing users to access files from any computer on a network.

9. **What are the two basic file system used in distributed system?**
- • The Sun Network File System, NFS.
- • The Andrew File System, AFS.

10.Write the Characteristics of file systems?

File systems are responsible for the organization, storage, retrieval, naming, sharing and protection of files. They provide a programming interface that characterizes the file abstraction, freeing programmers from concern with the details of storage allocation and layout.

11.**Define metadata.**

The term metadata is often used to refer to all of the extra information stored by a file system that is needed for the management of files. It includes file attributes, directories and all the other persistent information used by the file system.

12.What are the different forms of transparency are partially or wholly addressed by current file services?
- o Access transparency
- o Location transparency
- o Mobility transparency
- o Performance transparency
- o Scaling transparency

13.Define File group identifier.

- File group identifiers must be unique throughout a distributed system.a unique identifier can be generated by concatenating the 32-bit IP address of the host creating the new group with a 16-bit integer derived from the date, producing a unique 48-bit integer: *32 bits   16 bits*
- *File group identifier*: IP address date

14.What is the need of Name service?

In a distributed system, names are used to refer to a wide variety of resources such as computers, services, remote objects and files, as well as to users. A name is needed to request a computer system to act upon a specific resource chosen out

of many.

15. **Define object address.**

object's address: a value that identifies the

location of the object rather than the object itself.

16. **What is Uniform Resource Locator?**

Uniform Resource Locator (URL) is often used

for URIs that provide location information and specify the method for accessing the resource

17.What is Name resolution?

Name resolution is an iterative or recursive process whereby a name is repeatedly

presented

to naming contexts in order to look up the attributes to which it refers.

18. **What is Global Name Service?**

A Global Name Service (GNS) was designed and implemented by Lampson and colleagues at the DEC Systems Research Center [Lampson 1986] to provide facilities for resource location, mail addressing and authentication.

19. **Label the different forms of transparency in file services APRIL/MAY 2018**

Access Transparency Location Transparency Concurrency

Transparency Replication Transparency Fault Transparency

Migration Transparency

Performance Transparency Scaling Transparency

20. **Discuss on LDAP. MAY /JUNE 2016**

LDAP (Lightweight Directory Access Protocol) is a software protocol for enabling anyone to locate organizations, individuals, and other resources such as files and devices in a network, whether on the public Internet or on a corporate intranet

UNIT IV

1. **What is clocks drift rate?**

A clock's *driftrate* is the change in the offset (difference in reading) between the clock and a nominal perfect reference clock per unit of time measured by the reference clock.

2. **What is Coordinated Universal Time?**

Coordinated Universal Time – abbreviated as UTC (from the French equivalent) – is an international standard for timeKeeping. It is based on atomic time, but a so-called 'leap second' is inserted – or, more rarely, deleted – occasionally to Keep it in step with astronomical time. UTC signals are synchronized and broadcast regularly from landbased radio stations and satellites covering many parts of the world.

3. **How the clock synchronization done in Cristian's method?**

A single time server might fail, so they suggest the use of a group of synchronized servers .

It does not deal with faulty servers

4. **What are the issues resolved by BerKley's algorithm?**

The collection of computers whose clocks are to be synchronized are categorized as masters and slaves. The averaging of the clock values cancels out the individuals clocks tendencies to run fast   or slow. This overcomes the uncertainty due to message transmission time introduced in the Synchronized clock values returned by the master.

5. **Write the features of Network Time Protocol**

To provide a service enabling clients across the Internet to be synchronizedaccurately to UTC

To provide a reliable service that can survive lengthy losses of connectivity:

To enable clients to resynchronize sufficiently frequently to offset the rates of driftfound in most computers:To provide protection against interference with the time service, whether maliciousor accidental

6. Difference between reliable and unreliable failure detector?

Reliable failure detector is one that is always accurate n detecting a process failure. It answers processes queries with either a response of unsuspected-which, as before can only be a hint- or failed.

U Reliable failure detector may produce one of two values when given the identity of a process: Unsuspected or suspected. Both of these resultsare hints, which may or may not accurately reflect whether the process has actually failed.

7. Explain global states and consistent cuts with example.

Global state of a distributed system consists of

–Local stateof each process: messages sent and messages received

–State of each channel:messages sent but not received

What is a failure detector? What are its types? Explain.

8. A *failure detector* is a service that processes queries about whether a particular process has failed. It is often implemented by an object local to each process (on thesame computer) that runs a failure-detection algorithm in conjunction with its counterparts at other processes. The object local to each process is called a *local failure detector*.

9. Explain the central server algorithm for mutual exclusion.

   The central server algorithm •

The simplest way to achieve mutual exclusion is to employ a server that grants permission to enter the critical section.

10. Explain the ring based algorithm for mutual exclusion

 A ring-based algorithm

One of the simplest ways to arrange mutual exclusion between the *N* processes without requiring an additional process is to arrange them in a logical ring.

11. What is CUT?

Coordinated Universal Time-abbreviated as UTC (From the French equivalent)-is as international standard for timekeeping. It is based on atomic time, but a so-called 'leap second' is inserted-or, more rarely, deleted-occasionally to keep it in step with astronomical time.

12. What is filter dispersion?

NTP servers apply a data filtering algorithm to successive pairs which estimates the offset o and calculates the quality of this estimates as a statistical quantity called the filter dispersion

13. What is meant by HB relation?

Lamport called the partial ordering obtained by generalizing these two relationships the ♣ happened-before relation. It is also sometimes known as the relation of causal ordering or potential causal ordering. We can define the happened-before relation, denoted -♣> by as follow: HB1: If process pi:e->ie' , then e->e' HB2: For any message m, send (m) receive (m) HB3: IF e, e' and e" are events such that e->e' then e->e"

14. What do you meant by distributed garbage

An object is considered to be garbage if there are no longer any reference to it anywhere in the distributed system. The memory taken up by that object can be reclaimed once it is known as to be garbage.

15. Define Global History

Let us return to our general system p of N processes pi(i=1,2,3,.....N) Here a series of events occurs at each process, and that we may characterize the execution of each process by its history

16. Define Failure detector.

A failure detector is a service that processes queries about whether a particular process has failed .It is often implemented by an object local to each process that runs failure detection algorithms in conjunction with its counterparts at the other processes.

17. List the properties of failure detector

A failure detector is not necessarily accurate. Most fails into the category of unreliable failure detectors.

- A result of unsuspected
- A result of Suspected

18. Define critical section problem

The application – level protocol for executing a critical section is as follows

- enter() - enter critical section – block if necessary
- resurceAccesses() - access shared resources in critical section
- exit() - leave critical section other processes may now enter

19. What is meant by election

Election: choosing a unique process for a particular role is called an election – All the processes

agree on the unique choice – For example, server in dist. mutex

20. List the famous mutual exclusion algorithms

- Center server algorithm
- Ring- Based algorithms
- Mutual Exclusion using multicast and Logical Clocks

Unit V

1.What is process?

Process means a program in execution. Process execution must progress sequential order.

2.What is process migration?

The phenomenon of shifting a process from one machine to another one which is called process migration.

3.What is Load?

Load may be define as number of tasks are running in queue, CPU utilization, load average, I/O utilization, amount of free CPU time/memory, etc.

4. List desirable features of good process migration mechanism.

- Transparency
- Efficiency
- Minimal interference
- Minimize freezing time
- Minimal residual dependencies

5. What are strategies for the migration of files?

If the file is locked by the migrating process and resides on the same system, then transfer file with the process If the process is moved temporarily, transfer the file only after an access request was made by the migrated process.

6. Define thread.

A minimal software processor in which context a series of instructions can be executed. Saving a thread context implies stopping the current execution and saving all the data needed to continue the execution at a later stage

7. Explain the benefit of process migration

Better response time and execution speed – up

Reducing network traffic

Improving system reliability

Higher throughput and effective resource utilization

8. **List the types of process scheduling techniques.**
   - Task management approaching
   - Load balancing approaching
   - Load – Sharing approaching

9. **What is kernel level thread?**

   In kernel level thread, thread management is done by kernel. OS support the kernel level thread. Since kernel managing threads, kernel can schedule another thread if a given thread blocks rather than blocking the entire processes.

10. **What is user level thread?**

    User level thread uses user space for thread scheduling. These threads are transparent to the operating system. User level threads are created by runtime libraries that cannot execute privileged instructions.

11. **What is preemptive process migration?**

    Preemptive process transfer involve the transfer of a process that is partially executed. This transfer is an expensive operation as the collection of a process's state can be difficult.

12. **What is non preemptive process migration?**

    Non –preemptive process transfers involve the transfer of process that have not begun execution an hence do not require the transfer of the proces's state. In both types of transfers, information about the environment in which the process will execute must be transferred to the receiving node.

13. **What is strata?**

    The NTP service is provided by a network of servers located across the Internet. Primary servers are connected directly to a time source such as a radio clock receiving UTC; secondary servers are synchronized, ultimately, with primary servers. The servers are connected in a logical hierarchy called a synchronization subnet whose levels are called strata.

14. **What is filter dispersion?**

    NTP servers apply a data filtering algorithm to successive pairs which estimates the offset o and calculates the quality of this estimates as a statistical quantity called the filter dispersion

15. **Define Global History**

    Let us return to our general system p of N processes $p_i(i=1,2,3,.....N)$ Here a series of events occurs at each process, and that we may characterize the execution of each process by its history