Midterm Questions

CS 4113/5113 FA 19

Table of Contents

HW1 ***1***

cs4113 **1**

cs5113 **7**

HW2 ***15***

cs4113 **15**

cs5113 **20**

HW3 ***28***

cs4113 **28**

cs5113 **32**

HW4 ***39***

cs4113 **39**

cs5113 **44**

HW5 ***50***

cs4113 **50**

cs5113 **54**

HW1

cs4113

According to van Steen and Tanenbaum, a distributed system in which it is easy to add new components or replace existing ones without affecting those components that stay in place is considered:

A. Extensible
B. Interoperable
C. Portable
D. Scalable

The correct answer (answer A) comes from page 13:
"Another important goal for an open distributed system is that it should be easy to configure the system out of different components (possibly from different developers). Also, it should be easy to add new components or replace existing ones without affecting those components that stay in place. In other words, an open distributed system should also be extensible."

Answers B. and C. are other important parts of an open distributed system, but they are defined differently (also on page 13).
"Interoperability characterizes the extent by which two implementations of systems or components from different manufacturers can co-exist and work together by merely relying on each other's services as specified by a common standard."
"Portability characterizes to what extent an application developed for a distributed system A can be executed, without modification, on a different distributed system B that implements the same interfaces as A."

Answer D. is another important part of distributed systems, but it is also defined differently. (page 15)

According to Van Steen, scalability problems could be solved with scaling up. However, when it comes to scaling out, there are three basically three techniques: hiding communication latencies, distribution of work and replication. what is the advantage of Replication?

A. avoid waiting for responses to remote-service requests as much as possible

B. increases availability, balance the load and hide much of the communication latency problems.

C. splitting data into smaller parts, and subsequently spreading those parts across the system.

D. improving network with RAM, CPU,...

The answer is found on pages 22.

A. hiding communication latencies

B. right answer

C. distribution of work

D. improve network

Which of the following does NOT represent a problem with replication in geographically widely dispersed systems?
a.) Replication requires global updating on each modification to keep consistency across files.
b.) Global synchronization makes large-scale solutions improbable due to communication delays and implementation complexity.
c.) Replication across a distributed system can limit communication latency between nodes.
d.) Modifying a copy creates inconsistencies and affects the scalability of the system.
a.) INCORRECT (Van Steen pg. 23)
Modification to files, in the distributed system, requires "immediate propagation" for strong consistency to remain and large updates across the system have a "natural ordering problem" via concurrency.
b.) INCORRECT (Van Steen pg. 23)
Global synchronization suffers from "network latencies" and solutions tend to be naturally "non-scalable."
c.) CORRECT (Van Steen pg. 22)
Copies located close to a user's physical location helps increase "availability and improve performance" through improved communication latency between client and server.
d.) INCORRECT (Van Steen pg. 23)
Tolerating inconsistencies is application dependent. If inconsistencies cannot be tolerated, strong consistency and "other desirable features may be impossible" to implement in a scalable way.

According to Van Steen, which of the following scaling techniques involves taking a component and splitting it up in to smaller parts and subsequently spreading across the system?

A. Partitioning and Distribution

B. Hiding communication latencies

C. Replication

D. Caching

E. Geographical Scalability

A is correct. Chapter 1.2, Page 21. Partitioning and distribution involves the taking of a component, splitting it into smaller parts, and subsequently spreading those parts across the system.

B. is incorrect because Hiding Communication Latencies is done by trying to void waiting for responses to remote-service requests as much as possible.

C. is incorrect because replication replicates the components across the system to reduce performance degradation

D. is incorrect since Caching is a form of replication and therefore makes a copy of the resource to reduce performance degradation

E. is incorrect as it describes a design rather than a technique.

According to Van Steen, an important class of distributed systems is the one used for high-performance computing tasks. What is cluster computing?

A. The underlying hardware consists of a collection of similar workstations or PCs, closely connected by means of a high-speed local-area network. In addition, each node runs the same operating system.

B. This subgroup consists of distributed systems that are often constructed as a federation of computer systems, where each system may fall under a different administrative domain, and may be very different when it comes to hardware, software, and deployed network technology.

C. Providing the facilities to dynamically construct an infrastructure and compose what is needed from available services.

D. A group of computers computing the same problem.

Correct answer is A.

B. Grid computing

C. Cloud computing

D. Chapter 1 didn't cover a topic like this.

I got the questions and answers from the book in chapter 1 page 25.

Newman (1994) says Scalability of a system can be measured along at least three different dimensions. One of them being size scalability. What is size scalability?

A. An increase in users and resources doesn't have a negative impact on performance.

B. users and resources may lie far apart, but the fact that communication delays may be significant is hardly noticed.

C. system can still be easily managed even if it spans many independent administrative organizations.

D. a container of commonly used components and functions that need not be implemented by applications separately

A. is correct; Ch 1.2, Page 15. Size scalability: A system can be scalable with respect to its size, meaning that we can easily add more users and resources to the system without any noticeable loss of performance.

B. is incorrect b/c it describes Geographical scalability

C. is incorrect b/c it describes Administrative scalability

D. is obviously wrong because it describes middleware

Exam Question:

Poslad, roughly describes the core requirements of a Ubiquitous Computing System, and lists them as follows: Distribution, Interaction, Context awareness, Autonomy, and Intelligence.

From a distributed-system perspective, the basic concept of Autonomy is:

a. The system should react automatically to changes, and manual intervention in the system should be kept to a minimum.

b. Systems should employ advanced algorithms and models to handle incomplete input, quickly react to a changing environment, handle unexpected events, etc.

c. A system's ability to characterize the situation of entities that are relevant to the interaction between a user and application.

d. The devices and other computers forming the nodes of a system are simply networked and work together to form a single coherent system.

e. The system's ability to take all implicit and explicit interactions into account and react accordingly.

The answer is found on pages 42-43. Options b-e are taken from the other core requirements found between pages 41-43.

Ad. 4: Autonomy. An important aspect of most ubiquitous computing systems is that explicit systems management has been reduced to a minimum. In a ubiquitous computing environment there is simply no room for a systems administrator to keep everything up and running. As a consequence, the system as a whole should be able to act autonomously, and automatically react to changes. This requires a myriad of

techniques of which several will be discussed throughout this book. To give a few simple examples, think of the following:

Address allocation: In order for networked devices to communicate, they need an IP address. Addresses can be allocated automatically using protocols like the Dynamic Host Configuration Protocol (DHCP) [Droms, 1997] (which requires a server) or Zeroconf [Guttman, 2001].

Adding devices: It should be easy to add devices to an existing system. A step towards automatic configuration is realized by the Universal Plug and Play Protocol (UPnP) [UPnP Forum, 2008]. Using UPnP, devices can discover each other and make sure that they can set up communication channels between them.

Automatic updates: Many devices in a ubiquitous computing system should be able to regularly check through the Internet if their software should be updated. If so, they can download new versions of their components and ideally continue where they left off.

Admittedly, these are very simple examples, but the picture should be clear that manual intervention is to be kept to a minimum. We will be discussing many techniques related to self-management in detail throughout the book.

According to Van Steen, one characteristic of a distributed system is the fact that, to users, all individual computing elements should act as a single coherent system. A way to accomplish this is by distribution transparency, one type of which is relocation transparency. What is an example of relocation transparency?

- A. The idea that the system should not operate differently if workloads were to be moved from one data center to another.
- B. The fact that a user of the system should not be able to determine where an object is physically located based on how they're accessing it logically.
- C. If a system supports users moving from one geographical location to another while their existing processes and communications are not affected by this movement.
- D. The fact that a system might be exist in multiple data centers so that the system can fail over to a different location without the user realizing.

A. is the correct answer. "For example, the entire site may have been moved from one data center to another, yet users should not notice. The latter is an example of relocation transparency...," (Van Steen, pg. 9).

B. describes location transparency.

C. describes migration transparency.

D. describes replication transparency.

According to Van Steen in the Distributed Systems textbook, in the context of the architecture of grid computing, what is the fabric layer?

A) Layer that provides interfaces to local resources at a specific site.

B) Layer that consists of communication protocols for supporting grid transactions that span the usage of multiple resources.

C) Layer responsible for managing a single resource.

D) Layer that consists of the applications that operate within a virtual organization.

A - Correct. "The lowest fabric layer provides interfaces to local resources at a specific site." (Van Steen 29)

B - Refers to the connectivity layer (Van Steen 29)

C - Refers to the resource layer (Van Steen 29)

D - Refers to the application layer (Van Steen 29)

---

cs5113

Which of the following is true for a platform layer, one of the four layers in organization of cloud-computing?

a.) Platform layer provides the means to cloud-computing users to develop and execute user submitted programs. For example, software framework (Java/Python/.Net)

b.) Youtube or google docs is an example of platform layer in cloud computing which allows users to upload and share videos or articles

c.) It provides virtual storage and computing resources to the customers such as virtual machine for computation.

d.) This layer consists of physical components such as processors, memory and hard disk.

Answer:

A is correct: As per textbook, services that an operating system provides to application developers are similar to what platform layer provides to a cloud-computing customer, such as means to develop and deploy applications.

B is incorrect: These are examples of application layer. Applications such as gmail, are executed in this layer.

C is incorrect: Infrastructure layer is responsible for services such as virtual machines.

D is incorrect: Lowest layer, hardware layer is responsible for managing necessary processors, disks or any other required hardware.

Roadside Inc. produces robotic wheels that turn depending on given inputs from a control unit. Mother Brain Industries designs control units to be used with robotic wheels. Both manufacturers agree to a common standard for communication between their products, without publishing the specifics of their codebases implementation of functions. According to Van Steen and Tanenbaum, what is this an example of?

A) Interoperability, because it is an example of two separate systems from different manufacturers operating together by relying on each other's services as defined by a agreed upon standard.
B) Interoperability, because it is an example of an application written for one system being run on a different system with no/minimal changes.

C) Portability, because it is an example of two separate systems from different manufacturers operating together by relying on each other's services as defined by a agreed upon standard.
D) Portability, because it is an example of an application written for one system being run on a different system with no/minimal changes.

Answer:

A is correct. Van Steen and Tanenbaum define interoperability to be "the extent by which two implementations of systems or components from different manufacturers can co-exist and work together by merely relying on each other’s services as specified by a common standard"

B is not correct. This is not an example of an application written for one system being run on another system. The companies do not have access to the source code used by the other.

C is not correct. Van Steen and Tanenbaum define portability as the " extent an application developed for a distributed system A can be executed, without

modification, on a different distributed system B that implements the same interfaces as A. "

D is not correct. While that is the correct meaning of portability, this is not an example of that principle in practice.

According to Van Steen and Tanenbaum, which of the following is NOT a property of all transactions?

A) Parallelizable
B) Atomic
C) Durable
D) Isolated

Answer:

A is correct. While it is true that distributed systems often perform transactions as a set of subtransactions that can be run in parallel, this is not required of all transactions.
B is incorrect. All transactions should appear to a user as a single, indivisible operation.
C is incorrect. All transactions should be permanent after they have finished executing.
D is incorrect. All transactions should have no effect on other transactions running concurrently.

Q: In an operating system, which of the following is FALSE about Process Control Block (PCB)?

(A) Process communication

(B) Process priority

(C) Process state

(D) Maintain the data needed for the process

Answer:

(A) - Correct: PCB does not play any role in communication among processes. The OS takes care of inter process communication.

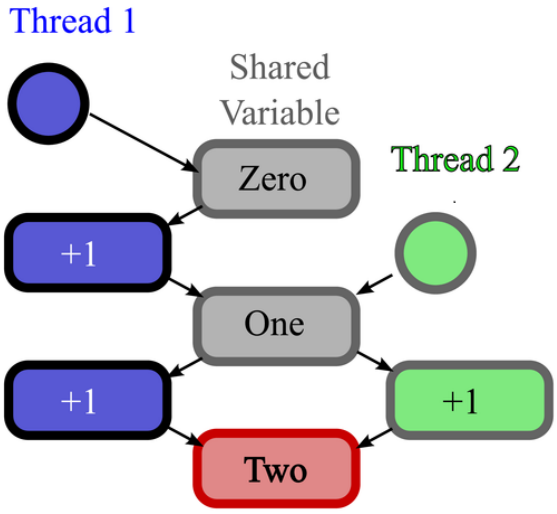
(B) - Incorrect: PCB contains process priority, pointers to scheduling queues for execution.

(C) - Incorrect: PCB contains information about the state (new, ready, running etc)

(D) - Incorrect: PCB stores all the data needed for a process.

Question: In a multi-threading environment, when two or more threads access shared memory and try to update the memory at the same time leading to different values of the shared memory. This leads to which of the following issues?

NOTE: In the figure, the shared variable value must be three(2 times +1 from Thread 1 & one time +1 from Thread 2), instead of two.



a) Race Condition

b) DeadLock

c) LiveLock

d) Starvation

Answer:

Correct :(A) - A race condition occurs when two or more threads try to update the shared memory leading to different results depending on the order of execution.

Incorrect:(B) - When threads cannot proceed because they are waiting on another process acquired resources, a deadlock occurs.

Incorrect:(C) - A livelock occurs when threads are running, but not making any progress.

Incorrect(D) - Due to physical constraints, at least one thread cannot get executed then starvation happens.

Reference: http://www.techenablement.com/intel-xeon-phi-optimization-part-1-of-3-multi-threading-and-parallel-reduction/

Q:

Which of the following is a correct assumption to make when developing a distributed application?

A) The topology of the network does change

B) Latency is 0

C) The network is homogeneous

D) The network is reliable

Answers:

A is correct because as you add or remove different nodes or if nodes fail the topology of the network will change, and if a node is on a sub-network there topology might be different from network that the DS itself uses

B is incorrect because there will always be latency when talking across a network

C is incorrect because different nodes will be talking on different protocols or medians but still can communicate with one another through middle ware or another node.

D is incorrect because the network can fail or drop data.

For the Distributed System to be scalable it is very important to have a decentralized algorithm for collecting and transporting all the input and output information. Why can't it be assumed that there is an implicit global clock for such algorithms?

A) Distributed systems use their own clock, so it is almost Impossible to get all the clocks synchronized.
B) Easy to get all the clocks synchronized.
C) Distributed systems use their own clock.
D) International Laws.

Answer:

A is correct.

Explanation : It is important to get all the clocks synchronized for the decentralized algorithms to work properly. But because of these systems placed in different geographic regions, it becomes difficult to synchronize these systems. The larger the system, larger the uncertainty of time. On a single LAN, with considerable effort it may be possible to
get all clocks synchronized down to a few microseconds, but doing this nationally or internationally is tricky.

Reference: Textbook

According to Van Steen and Tanenbaum, the integration of enterprise applications is an important goal for many middleware products. There are different ways to integrate such applications. Which of the following methods is NOT correct?

A) File transfer

B) Shared database

C) Remote procedure call

D) Bootloader

Answer:

A is correct. An application produces a file containing shared data which is subsequently read by other applications.

B is correct. All applications have access to the same data (for example a SQL database).

C is correct. An application offers a procedure to other applications via RPC.

D is incorrect. A bootloader is used to boot the OS of a system and has nothing to do with the integration of enterprise applications.

According to Van Steen and Tanenbaum, which of the following forms of scalability is the most difficult to solve, and why?

A) Administrative scalability, because it deals with nontechnical issues, such as the politics of an organization and human collaboration.
B) Size scalability, because the machines are incapable of handling the work.
C) Geographical scalability, because the issues arising from network latency.
D) Geographical scalability, because the occurrence of data replication and inconsistency.

Explanation:

A) Correct – Administrative scalability is thus far the only open-ended problem with distributed system. Some steps have been made to resolve administrative scalability using peer-to-peer networks, however not all distributed systems are allowed to be peer-to-peer.
B) Incorrect – Size scalability is considered the easiest to fix, as the solution is to simply throw more money in to upgrade the hardware
C) Incorrect – Geographical scalability was considered the second hardest to fix, however there already exist methods for dealing with network latency through network distribution techniques.
D) Incorrect - Geographical scalability was considered the second hardest to fix,

however “combining distribution, replication, and caching techniques with different forms of consistency generally leads to acceptable solutions.”

According to van Steen, a system's scalability can be measured along three dimensions. Which one of the following is not one of these three dimensions?

A. Scalability across devices

B. Geographical scalability

C. Administrative scalability

D. Scalability of size

Answer:

"A" is correct because although the choice seems reasonable, the term is not one of the three dimensions presented in the book, and is usually presented as a UI design concern (i.e. scaling the size of a UI to fit various display sizes).

B is not a correct answer as geographical scalability is presented in the book and refers to the cloaking of delays in communication between a user and a far-away resource.

C is not a correct answer as administrative scalability is presented in the book and refers to the ability for the system to be managed across multiple independent administrative organizations.

D is not a correct answer because size scalability is presented in the book and refers to the ability for the system to take on additional resources and users without noticeable degradation of performance.

Source: Distributed Systems Third Edition by Maarten van Steen and Andrew S. Tanenbaum. Page 15.

According to van Steen, which of the following is true regarding the degree of distribution transparency:

(A) There exists a trade-off between a high degree of transparency and the performance of a system.

(B) Aiming for distribution transparency should always be the main goal.

(C) When distribution transparency is impossible, we should pretend it is feasible.

(D) If replicas of a system are located very far away (i.e. different continents), for any given system update the user should not be aware of the fact that the system is distributed.

Answer:

A is correct in the example where an Internet application repeatedly tries to contact a server before giving up. Instead of giving up earlier or letting the user cancel the attempts, the system has been slowed down.

B is incorrect because other issues such as performance and comprehensibility should be considered.

C is incorrect because sometimes it is wiser to make distribution explicit for the user and application developer. Therefore, the user will likely better understand the behavior of a distributed system and thus will be better prepared for various possible scenarios.

D is incorrect because an update operation may take several seconds to complete, which will leave the user questioning why the system is taking so long to respond.

According to Van Steen and Tanenbaum, distributed systems should be open. Which of the following is an important step in providing openness?

A) Defining services through interfaces using an Interface Definition Language (IDL)

B) Obscuring where an object is physically located on the system from users

C) Ensuring that applications needing a certain service are all implemented in the same way

D) Never assuming that latency is zero

E) Learning to trust people again even though you've been hurt before

Answer:

(A) is correct. To be open means that components should adhere to standard rules that describe the syntax and semantics of what those components have to offer. An IDL captures the syntax of the services, thus playing an important role in openness.

(B) is incorrect. It is describing location transparency, which is unrelated to openness.

(C) is incorrect. It is the opposite of what we want for openness. Rather, an IDL should not prescribe what an implementations should look like. It should be neutral.

(D) is incorrect. It is true that latency is never zero, but it is unrelated to openness.

(E) is incorrect. This question is about the openness of distributed systems, not emotional openness.

According to Van Steen, which of the following are true about grid-computing systems?

a. For a grid-computing system, there are no assumptions are made about the similarity of hardware, operating systems, networks, administrative domains, or security policies.
b. Grid-computing is used, in virtually all cases, for parallel programming in which a single program is run in parallel on multiple machines.
c. Grid computing is characterized by an easily usable and accessible pool of virtualized resources where resources can be configured dynamically.
d. The layered architecture of a grid-computing system consists of 3 layers, a resource layer, a fabric layer, and a collective layer.

Explanation:
A is correct. Grid computing systems consist of different types of nodes. Different networks are allowed.
B is incorrect. This is a characteristic of cluster computing.
C is incorrect. This is a characteristic of cloud computing.
D is incorrect. Van Steen states that there are 4, not 3, layers in the layered architecture for grid-computing systems: a collective layer, a connectivity layer, a resource layer, and a fabric layer.

Source: Course Textbook

HW2

cs4113

According to the textbook, which of the following levels in an Application Layering architecture is extremely varied between distributed systems?

A. The Processing Level

B. The Application-Interface Level

C. The Data Level

D. The User-Interface Level

A. A is correct. On page 60 of the textbook, it discusses how the processing level "contains the core functionality of the application... there are not many aspects common to the processing level."

B. The application-interface level is a user interface or application interface, which tends to be somewhat standardized for ease of use.

C. The data level tends to be a database or a file system.

D. The user-interface level isn't a typical level of the application layering architecture. User interfaces are already considered a part of the application-interface level.

According to the textbook, what is the new problem that peer-to-peer networks have to deal with?

A. capacity of the network to manage big system

B. unstable behavior of super peer and its back up issue

C. how to select node to be come super peer

D. flexible for weak peer to connect to different super peer.

The answer is C. we can find an answer in page 88. the peer to peer network help normal node to easily connect to super peer around it. however, when one node becomes bigger and contain more data, it will hardly to decide which node will becomes super peer.

Considering the peer-to-peer file downloading system BitTorrent, choose the answer below that is known for high efficiency when downloading files between users.

A.) Trading Phase

B.) Bootstrap Phase

C.) Last Download Phase

D.) None of the Above

A.) Correct => According to pg. 93, the trading phase is known for "highly efficient" exchanging of blocks between peers.

B.) Incorrect => May need to wait for neighbors to be added to the potential set before blocks can be exchanged.

C.) Incorrect => Dependent on newly arriving peers in-order to get the remaining blocks of a file.

D.) Incorrect

According to the textbook, what is the interface responsible for in an object?

A. Conceals implementation details allowing us to consider an object independent of the environment.

B. Remembers the objects relevant information.

C. Marshals a method and send a stream of bytes to the server to be unmarshaled.

D. Unmarshals a received stream of bytes.

A. is correct because in Section 2.1 of the book it states that "The interface offered by an object conceals implementation details, essentially meaning that we, in principle, can consider an object completely independent of its environment"

B. is incorrect because it describes an objects state.

C. is incorrect because it describes a proxy.

D. is incorrect because it describes a skeleton.

According to the textbook, wrappers are one of the design patterns of middleware to promote openness. What is a wrapper?

A. Is a special component that offers an interface acceptable to a client application, of which the functions are transformed into those available at the component.

B. A software construct that will break the usual flow of control and allow other (application specific) code to be executed.

C. Helps in updating the middleware through dynamically constructing middleware from components.

D. A component that allows applications to invoke remote objects.

A. Is correct according to the textbook on page 72.

B. This is the definition of an interceptor, which is the other design patterns of middleware besides wrappers.

C. Wrappers don't directly get involved in the modification process of a middleware.

D. This is the object adapter that is in the context of object-oriented programming where wrappers come from. This is close but not quite right.

According to the textbook, how do a web server and browser communicate?

A. HTTP request is sent from the browser to the **web server** to fetch a document.

B. HTTP request is sent from the browser to the **server** to fetch a document.

C. URL, specifying the document info, communicates between the Browser and the web server

D. The network file system is responsible for communication between the web server and the browser.

A. A is correct because in section 2.4 of the book, it clearly states, "The communication between a browser and Web server is standardized: they both adhere to the HyperText
Transfer Protocol (HTTP)."

B. B is almost correct. It should say web server, not a server. These are 2 different things.

C. This is only half correct. It does specify where a document is but doesn't help in communication.

D. This is not what a network file system(NFS) does. NFS provides a view of the local file system.

Which of the following statements is **not** a specification of the Transmission Control Protocol (TCP)?

A. The protocol specifies the active nodes that have (chunks of) the requested file.
B. The protocol specifies which messages are to be exchanged for setting up or tearing down a connection.
C. The protocol specifies what needs to be done to preserve the ordering of transferred data.
D. The protocol specifies what both parties need to do to detect and correct data that was lost in transmission.

A. is from the definition of a torrent tracker. pg. 91-92
B. Is part of the TCP. pg. 58-59
C. Is part of the TCP. pg. 58-59
D. Is part of the TCP. pg. 58-59

This kind of service is realized in the Internet by means of the **Transmission Control Protocol (TCP)**. The protocol specifies which messages are to be exchanged for setting up or tearing down a connection, what needs to be done to preserve the ordering of transferred data, and what both parties need to do to detect and correct data that was lost during transmission.

Van Steen describes peer-to-peer systems as "a class of modern system architectures that support horizontal distribution," which is a way to refer to the way that clients and servers are distributed. There are two main types of peer-to-peer systems: structured and unstructured. If you wanted to build a structured peer-to-peer system, how would you implement the process of searching for the data needed?

- A. Implementing a distributed hash table, in which each node has its own unique identifier. A requesting node passes the data item's key through a lookup function, and the node where the data item is stored is returned.
- B. Having a node pass a request for data to all of its neighbors, who repeat the same process until the data is found and is passed up to the original requestor.
- C. Using a broker that keeps data on resource usage and availability for a nodes in a similar geographic region to optimize storage location.
- D. Making a node pass a request for data to a random neighbor, who, if itself does not contain the data, will repeat the process until it is found.

A. is the correct answer. "Following this approach now reduces the essence of structured peer-to-peer systems to being able to look up a data item by means of its key. That is, the system provides an efficient implementation of a function *lookup* that maps a key to an *existing* node: *existing node = lookup(key),*" (Van Steen, pg. 82).

B. describes "Flooding," which is a way of implementing an unstructured P2P system.

C. describes a collaborative Content Delivery Network, which is a way of implementing an unstructured P2P system.

D. describes "Random Walks," which is a way of implementing an unstructured P2P system.

According to the textbook, in a network file system, what is a Remote Procedure Call?

a) A standardized way to let a client machine make an ordinary call to a procedure that is implemented on another machine.

b) Component responsible for handling incoming client requests.

c) Component responsible for handling access to files stored at a remote server.

d) Idea that each file server provides a standardized view of its local file system.

-

a) Correct (Textbook 95-96). The client uses RPCs to tell the file server what operations to perform on the file.

b) Incorrect. This references the NFS Server (Textbook 96). This is what receives the RPCs from the client.

c) Incorrect. This references the NFS Client (Textbook 95). This is what communicates the RPCs to the NFS Server.

d) Incorrect. This is the fundamental idea behind a network file system.

---

cs5113

Question:

Which of the following nodes can be referred as super peers in Hierarchically organized peer-to-peer networks?

a.) Nodes that maintain the data on resource usage and availability for the nodes of the network that are in each other's proximity.

b.) Nodes would not maintain an index of the data items.

c.) Nodes that are not organized in a peer-to-peer network

d.) Nodes referred as super peers are short lived processes with high availability

Answer:

Option a.) is correct, as super peer nodes acts as broker and maintains resource usage or availability of nearby nodes to allow quick selection of the node with sufficient resources. Special nodes provide an alternative many peer-to-peer systems that have scalability problem

Option b.) is incorrect, because super peer maintains an index of the data items, which is an alternative to scalability problem. For unstructured peer-to-peer systems, locating relevant data items becomes a problem with growing networks

Option c.) is incorrect, as per the textbook, super peers are organized leading to a hierarchical organization in a peer-to-peer network

Option d.) is incorrect, super peers are long lived processes with high availability

Reference: Textbook

According to Van Steen and Tanenbaum, which of the following is a potential downside to choosing a hierarchical peer-to-peer model?

A. Choosing which nodes should be super nodes can be a challenging problem.

B. There is no way to handle super peers malfunctioning.

C. Security becomes exponentially harder in a hierarchical architecture.

D. Because managing super peers takes considerable resources, searching in hierarchical architectures actually takes more time than in non-hierarchical ones.

Answers:

A. This is correct. This is similar to a **leader-election problem**, and is discussed later in the book.

B. This is false. Van Steen and Tanenbaum note that backup schemes (such as attaching weak peers to multiple super peers) already exist to mitigate this issue.

C. This is false. Van Steen and Tanenbaum do not say anything regarding to the security of hierarchical architectures.

D. This is false. The whole point of doing a peer-to-peer architecture this way is that it reduces the resources needed to search.

As described by Van Steen and Tanenbaum, which of the following is a characteristic of a RESTful architecture?

A) Messages to and from a service are fully self-described
B) Resources can be named using a variety of different naming schemes
C) The four operations available in a RESTful architecture are PUT, GET, READ, and WRITE
D) Components maintain a history containing information about callers

Answers:

A) It is true that messages to and from a service are self-described
B) Resources should be named using the same, single naming scheme
C) The four operations are PUT, GET, POST, and DELETE
D) Components are stateless, and thus forget everything about a caller after executing an operation

---

According to Van Steen and Tanenbaum, which of the following statements about Unstructured peer-to-peer systems are FALSE?

(A) Flooding is more efficient when compared with the random walk because number of node visits will be less and therefore less time to complete the work.

(B) Flooding can be very expensive because every time we are requesting a data item to all the neighbors and hence a request often has an associated time-to-live (TTL) value.

(C) Random walk imposes much less network traffic, yet it may take much longer before a node is reached that has the requested data.

(D) In policy-based search methods, a node may decide to keep track of peers who responded positively and turn them into preferred neighbors for succeeding queries.

Answer: (A)

(A) False: Flooding actually takes less time to find the data item but we have to visit more number of nodes usually when compared with a random walk. Authors illustrated this in Note 2.6 (Advanced flooding versus random walks).

(B) True: Flooding is expensive because it creates huge network traffic by passing a request to all its neighbors. To terminate at some point, we often have an associated time-to-live (TTL) value.

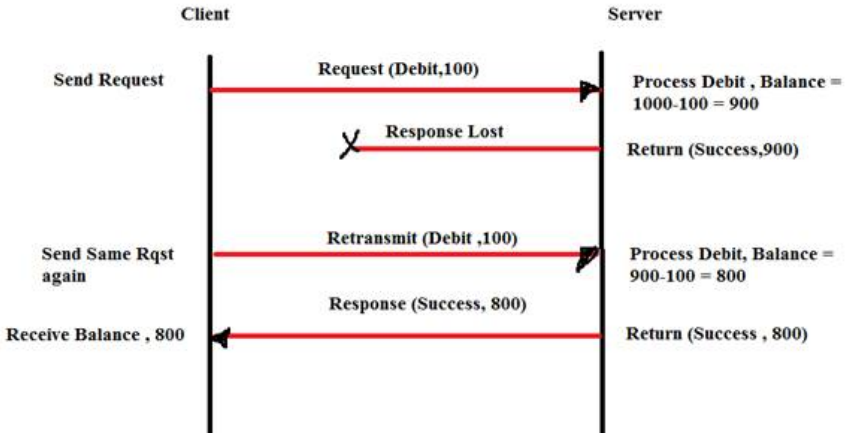
(C) True: Random walk imposes less traffic because it passes a request only to certain number of neighbors usually 16, 64 (according to Lv et al.[2002]) hence it takes more time to find the data item.

(D) True: Policy-based search methods lie usually between flooding and random walks where we keep track of nodes who responded positively and make them preferred neighbors for upcoming queries

---

Question) According to Van Steen and Tanenbaum, to maintain the state(consistency) of the system, if an operation is requested by a client to the server multiple times, it should not harm the system. These operations are called what?

NOTE: Here the operation(Request $100 withdrawal) by the client affects the system if performed more than once.



A) idempotent operations

B) secured operations

C) reliable operations

D) atomic operations

Answer:

Correct - (A) An idempotent operation does not harm the system after repeating multiple times.

Incorrect - B) A secured operation if performed more than once might/might not affect the system.

C) A reliable operation, if performed more than once might/might not affect the system.

D) An atomic operation is an indivisible request made by a client to server, it might/might not affect the system if performed more than once.

According to Van Steen and Tanenbaum, which of the following design provides the most openness for middleware?

A) Modifiable middleware

B) Wrappers

C) Interceptors

D) Brokers

Answers:

A) Modifiable middleware can change dynamically, based on the environment, during runtime to load and unload different components to be used by the system.

B) Wrappers are used to achieve openness to a middleware but must be updated based on what components are in the system.

C) Interceptors are used to achieve openness to a middleware but must be updated based on what components are in the system.

D) Brokers are a design technique to centralize and process requests for Wrappers and Interceptors.

Which one from the following is the correct steps of establishing a TCP socket communication on the server side?

A) create Socket -> bind the socket to an address -> listen for connections -> accept a connection -> send and receive data -> close the connection

B) create Socket -> send and receive data

C) create Socket -> listen for connections -> accept a connection -> send and receive data -> close the connection

D) create Socket -> bind the socket to an address -> send and receive data

Answer: A

Explanation:

The steps involved in establishing a TCP socket on the server side are as follows:

- Create a socket with the socket() function;
- Bind the socket to an address using the bind() function;
- Listen for connections with the listen() function;
- Accept a connection with the accept() function system call. This call typically blocks until a client connects with the server.
- Send and receive data by means of send() and receive().
- Close the connection by means of the close() function.

Reference : https://www.cs.dartmouth.edu/~campbell/cs60/socketprogramming.html

---

According to Van Steen and Tanenbaum, a Network File System (NFS) is a solution for file server to provide a standardized view of its local system. Which of the following statements about NFS is correct?

A) The client-server communication is done through RPC (Remote Procedure Calls).

B) The upload/download model is used to exchange commands between server and client.

C) Clients can access files on the file server only if the OS is the same.

D) Using the remote access model, the file is transferred from the server to the client and uploaded back again after modification.

Answers:

A) is correct. No explanation needed.

B) is false. Using the upload/download model a client accesses a file locally after downloading it from the server. After the client is finished the file is uploaded back to the server.

C) is false. The basic idea behind NFS is that the access is independent of the underlying server-side local file system.

D) is false. The file stays on the server and is accessed remotely.

According to Van Steen and Tanenbaum, which of the following peer-to-peer architectures relies on a connection between weak peers and super peers?

A) Hierarchical peer-to-peer systems
B) Unstructured peer-to-peer systems
C) Structured peer-to-peer systems
D) Distributed peer-to-peer systems

Explanations:

A) Correct. Select nodes in a hierarchical peer-to-peer networks are designated as super peers. Most other connections, weak peer, will bind themselves to the nearest super peer. The super peers contain information about the week peers connected to them and all other super peers

B) Incorrect. An unstructured network would not have organization such as this.

C) Incorrect. Structured peer-to-peer networks have predefined organizations, where hierarchical are a bit more free formed.

D) Incorrect. This is not actually a term for a peer-to-peer system as defined by the book. Generally all peer-to-peer systems would be distributed.

According to van Steen, which of the following is NOT a type of peer-to-peer system?

A) Multitiered Peer-To-Peer System

B) Hypercube Peer-To-Peer System

C) Structured Peer-To-Peer System

D) Non-Structured Peer-To-Peer System

Answer:

A. is correct as no such term exists in the book. The word 'multitiered' should be recognized as a mechanism supporting vertical distribution, which is in contrast with peer-to-peer systems supporting horizontal distribution, with each process acting as both a server and a client.

B. is incorrect as it is given as an example as a type of structured peer-to-peer system in the text.

C. is incorrect as structured peer to peer systems are presented in the text. It's a P2P system in which each node is assigned a unique identifier, stored in a distributed hash table allowing for efficient lookup.

D. is incorrect as unstructured peer to peer systems are presented in the text. It's a P2P system in which each node contains its own list of neighboring nodes. Data

lookup in unstructured networks is generally less efficient because there is no predetermined route to follow, leading to scalability issues.

Source: Distributed Systems Third Edition by Maarten van Steen and Andrew S. Tanenbaum. Pages 81-86

According to Van Steen, which of the following is true about the RESTful architecture:

A. When distributed transactions are needed, RESTful architectures are not a good idea.

B. RESTful services offer better solutions than traditional services.

C. The simplicity of RESTful architectures provides easy solutions to intricate communication schemes.

D. When using RESTful calls, many syntactical errors can often be caught during compile time.

A is correct because distributed transactions generally require that services keep track of the state execution, scenario where RESTful architectures are not the best option.

B is incorrect because there are scenarios when RESTful services offer a better solution and scenarios when traditional services offer a more elegant solution.

C is incorrect because the simplicity of RESTful architectures can easily prohibit easy solutions to intricate communication schemes.

D is incorrect because usually, due to the structure of a RESTful call, checking is deferred until runtime.

According to Van Steen and Tanenbaum, which of the following is NOT true of a structured peer-to-peer system?

A) Logically different components should be placed on different machines, such as separating user interface, processing components, and data management components.

B) A client or server may be physically split into logically equivalent parts, but each part is operating on its own share of the complete data set.

(C) The system can be represented by a network in which the nodes are formed by the process and the links represent the possible communication channels.

(D) The system usually employs a semantic-free index.

Answers:

(A) is correct. It is describing vertical distribution, which is not a necessary condition for peer-to-peer systems. Rather, we require horizontal distribution.

(B) is incorrect. It is describing horizontal distribution, which a peer-to-peer system should support.

(C) is incorrect. It is describing an overlay network. A structured peer-to-peer system requires an overlay network that adheres to a specific, deterministic topology such as a ring or a grid.

(D) is incorrect. The essence of structured peer-to-peer systems is being able to look up a data item by means of its key. This is usually achieved by uniquely associating each data item with a key by means of a hash function and using that key as an index (i.e., using a semantic-free index).

Van Steen explains 4 coordination models; which of the following is true for the mailbox coordination model?

A. There is no need for communicating processes to be executing at the same time for coordination to occur.
B. The mailbox coordination model is also known as the shared data space model.
C. Communicating processes do not need know each other explicitly.
D. Processes publish notifications describing the occurrence of events.

Explanation:
A is correct. This is the one of the characteristics of the mailbox coordination model. The mailbox coordination model is referentially coupled and temporally decoupled.
B is incorrect. The shared data space model is something different. The shared data space model uses communication that is temporally decoupled and referentially decoupled.
C is incorrect. Processes have to know of each other explicitly in the mailbox model.
D is incorrect. This refers to the Event-based model, which is temporally coupled and referentially decoupled.

Source: Course Textbook

HW3

cs4113

According to Dr. Sudarshan Dhall's lecture on 9/17/20, which of the following types of Deadlock Models requires that a process can have at most one outstanding request for only one unit of a resource?

A) Single Resource Model

B) The AND Model

C) The OR Model

D) The Centralized Model

Answers:

A) is correct. As taken from the lecture "A process can have at most one outstanding request for only one unit of a resource"

B) is another type of Deadlock Model, but in the AND Model "a process can request for more than one resource simultaneously"

C) is also a type of Deadlock Model, but in the OR Model "a process can make requests for multiple resources simultaneously"

D) is not a type of deadlock model. Centralized deadlock detection algorithms are a specific type of Deadlock Detection algorithm in which a control site constructs WFGs and checks for directed cycles.

According to the Lamport's logical clock, if two events A and B happen in two different processes then:

A) can't tell

B) A --> B

C) B --> A

D) A and B happen in the same time

Answer: A. because A and B are in two different processes so they can happen in the same time or A--> B or B --> A. textbook page 311

B: A and B are in 2 different process

C: A and B are in 2 different process

D: A and B are in 2 different process

According to the textbook, which mutual-exclusion algorithm is widely applied due to its simplicity and efficiency?

A.) Centralized

B.) Decentralized

C.) Distributed

D.) GPS

A.) is correct. The author concretely states, on page 328, that the centralized approach "is [the] simplest and most efficient."

B.) is incorrect. This algorithm is slow (in message time) and can lead to resource underutilization(pg. 328).

C.) is incorrect. This algorithm has problems with increased points of failure (N points) and imposes a burden on resource-constrained machines(pg. 325).

D.) is incorrect. Not referring to a mutual-exclusion algorithm.

According to the textbook, which threading model describes a model in which all threads are schedulable entities that are executed in the operating system's kernel?

A. One-to-one threading model

B. One-to-many threading model

C. Many-to-One threading model

D. Many-to-many threading model

A is correct because it is described in Section 3.1 on page 109, "These problems can mostly be circumvented by implementing threads in the operating system's kernel, leading to what is known as the one-to-one threading model in which every thread is a schedulable entity."

B is incorrect because a One-to-many threading model does not exist.

C is incorrect because the threads are at the user-level rather than the kernel level and multiple threads are mapped to a single schedulable entity.

D is incorrect because a Many-to-Many model is a hybrid between the many-to-one threading model and the one-to-one threading model, thus there are user-level threads.

According to the textbook, there is a design issue servers whether a server is stateless or not. Where are stateful, stateless, and soft state servers. What is a stateless server?

A) Meaning does not keep information on the state of its clients, and can change its own state without having to inform any client.

B) The server promises to maintain state on behalf of the client, but only for a limited time.

C) Means maintaining persistent information on its clients.

D) Contains a small piece of data containing client information.

Correct Answers:

A) Is correct according to the textbook on page 131 section Stateless vs Stateful Servers. The definition of stateless came from Birman 2012.

B) Is not correct because this is soft state which is a type of stateless server but it stores information only for a limited time.

C) Is not correct because this is the definition of stateful servers.

D) Is not correct because this describes a cookie.

What is a solution for the problem of bottleneck in the context of stateful servers?

A. A TCP handoff

B. Using threads instead of processes

C. mix user-level and kernel-level threads

D. Use a separate port for urgent data

A. Correct. According to slide 47 of the Ch. 3 slides, a TCP handoff is one solution resolving a bottleneck

B. using threads is a solution for the downsides of processes since it is easier to delete, create threads than processes

C. Used to improve performance; not related w/ the bottleneck issue

D. This is a solution for interrupting a server after it has accepted a request

---

According to the text book, the mutual-exclusion algorithm described in the table is?

Messages per entry/exit	Delay before entry (in message times)
2•m•k+m, k=1,2,…	2•m•k

A. Decentralized
B. Token ring
C. Distributed
D. Centralized

The answer is easily found on page 328 in Figure 6.19.

Algorithm	Messages per entry/exit	Delay before entry (in message times)
Centralized	3	2
Distributed	$2 \cdot (N-1)$	$2 \cdot (N-1)$
Token ring	$1, \ldots, \infty$	$0, \ldots, N-1$
Decentralized	$2 \cdot m \cdot k + m, k = 1, 2, \ldots$	$2 \cdot m \cdot k$

Figure 6.19: A comparison of four mutual exclusion algorithms.

Choices A-D are names of the mutual-exclusion algorithms in the table.

According to the textbook, all hardware clocks (and consequently software clocks) are subject to "clock drift." What is the typical clock drift rate?

- A. 10^{-6} seconds per second
- B. 10^{-9} seconds per second
- C. 10 milliseconds per second
- D. 50 nanoseconds per second

A is the correct answer. "A typical quartz-based hardware clock has a clock drift rate of some 10^{-6} seconds per second," (Van Steen, pg. 303).

B is a distractor answer.

C is the practical accuracy of the UTC service.

D is the accuracy of ground time servers that connect in with the satellite UTC service.

According to the textbook, a server cluster is typically organized into 3 tiers. What does the 3rd tier consist of?

A) Distributed file/database system

B) Application/compute servers

C) Logical switch

D) Front end

A) Correct. According to the image and text on page 142 of the textbook. "... the third tier, which consists of data-processing servers, notably file and database servers."

B) Incorrect. The application/compute servers are the second tier. These servers are "typically running on high-performance hardware dedicated to delivering compute power." (Textbook 141)

C) Incorrect. "The first tier consists of a (logical) switch through which client requests are routed." (Textbook 141)

D) Incorrect. Front end is just another name for the logical switch, the first tier. (Textbook 142)

---

cs5113

Question:

As per van Steen and Tanenbaum, the most important application of virtualization is in cloud computing. Which of the following is true for Infrastructure-as-a-Service (IaaS), one of the three different services offered by cloud providers?

A. In case of Infrastructure-as-a-Service (IaaS), virtualization allows almost complete isolation between customers of cloud providers and gives an illusion of renting a dedicated physical machine to each customer.

B. The isolation among customers leads to high performance

C. Virtualization is key factor in IaaS, which will allow a customer to create no more than single virtual machine configured with resources

D. IaaS providers cannot provide a service to configure a system with networked virtual servers who can communicate using IP networks.

Answer A. is correct, virtualization plays a key role in distributed systems. A programming interface can be extended using virtualization to imitate behavior of another system. For cloud providers, even with shared physical machine, virtualization allows renting virtual machines to customers with almost complete isolation.

Answer B. is incorrect, with the fact that actual physical resources are shared, isolation is never complete, and it leads to lower performance.

Answer C. is incorrect because IaaS allows a customer to create a large number of virtual machines, that are networked together

Answer D. is incorrect because many IaaS providers with the use of virtualization provides service to the customers to create a complete distributed system

---

According to van Steen and Tanenbaum, what is the key difference between the concepts of **precision** and **accuracy**?

A. Precision measures the deviation between the clocks of two machines inside of a distributed system, while accuracy refers to the measure of deviation from an outside reference point, like UTC.

B. Accuracy measures the deviation between the clocks of two machines inside of a distributed system, while precision refers to the measure of deviation from an outside reference point, like UTC.

C. Precision and accuracy both measure the derivation between the clocks of two machines inside a distributed system, but accuracy is more strict.

D. Precision and accuracy both measure the derivation between the clocks of two machines inside a distributed system, but precision is more strict.

E. It's all meaningless because time is a tool invented by the state to suppress us.

Answers:

A. is correct, as this is how it is presented in the book.

B. is incorrect, as the definitions are reversed.

C. is incorrect as accuracy does not measure the difference between two machines in the same DS.

D. is incorrect for the same reason as C was.

E. is incorrect. Probably. Or maybe not, but van Steen and Tanenbaum never claimed this so it doesn't answer this question.

According to van Steen and Tanenbaum, a clock that captures not only the ordering of events but also the causality of events is a:

A) Vector clock

B) Logical clock

C) Global clock

D) Reference clock

Answers:

A: correct

B: A logical clock captures only the order of events

C: A global clock is one that all processes are able to access

D: A reference clock is one that synchronizes to another clock, such as an atomic clock

---

According to Van Steen and Tanenbaum, which of the following statements about threads are not correct?

(A) Multithreading always leads to a performance gain.

(B) In the case of threads, no attempt has been made to achieve a high degree of concurrency transparency.

(C) Multithreading helps in exploiting parallelism when executing the program on a multiprocessor or multicore system.

(D) Large applications can also be constructed such that different parts are executed by separate threads.

Answers:

A) is false. Usually multithreading leads to a performance gain.

B) is true. A high degree of concurrency transparency is only seen in processes.

C) is true. The main advantage of multithreading is to exploit parallelism.

D) is true. If we are constructing large applications as a collection of cooperating programs, each to be executed by a separate process. We have a drawback of context switching.

Q) According to van Steen and Tanenbaum, which of the following is FALSE about a stateful server.

A) A stateful server can resume effectively to its previous state with no loss of information after any sudden crash.

B) A stateful server maintains persistent information about its clients. This information about the clients is explicitly deleted by the stateful server

C) A stateful server keeps track of the permissions it has issued to its set of clients which improves the performance of read & write operations.

D) A file server is a stateful server and it maintains a table of file,client entries.

Answer:

A) Correct - A stateful server cannot guarantee that it processed all the requests it has received from its clients after a crash. It simply starts running again and waits for requests from clients.

B) Incorrect - A stateful server has to explicitly add or delete the information about its clients

C) Incorrect - The performance improvement is achieved by a stateful server by keep tracking the permissions.

D) Incorrect - A table is maintained by stateful servers to handle all the client info.

Q: Which of the following is not a Clock synchronization algorithm mentioned by Van Steen and Tanenbaum?

A) Precision Time Protocol (PTP)

B) Network Time Protocol (NTP)

C) The Berkeley algorithm

D) Reference broadcast synchronization(RBS)

Answers:

A is correct, even though Precision Time Protocol (PTP) is used to sync clocks on a network by having a master clocks who will get there time from a grand master clock and slaves will sync to a master clock, but it is not mentioned by Van Steen and Tanenbaum.

B is incorrect, Network Time Protocol (NTP) is used to sync clocks on a network by having clients contact a time server and to get the updated time but must also correct for delay in transmission, and is mentioned by Van Steen and Tanenbaum.

C is incorrect, The Berkeley algorithm the time server polls all the devices and will the tell the device to speed up or slow down its clock to achieve clock sync, and is mentioned by Van Steen and Tanenbaum.

D is incorrect, Reference broadcast synchronization(RBS) will broadcast the time out allowing receivers to adjust their clocks as needed, and is mentioned by Van Steen and Tanenbaum.

According to van Steen and Tanenbaum which one of the following is right about a Layered Architecture?

A) Components are organized in a layered fashion where a component at layer Lj can make a downcall to a lower layer Li (i < j) and expects a response. In exceptional cases upcall is made.

B) Lower layer Li components always make a upcall to a higher layer Lj where (i < j).

C) In a downcall, a higher layer always makes a request to a layer just below it.

D) In a Layered communication-protocol stack, the layer interface should not hide the implementation of the service.

Answer: A

B is wrong because there is an option for doing an upcall as well.

C is wrong because in a Mixed Layered architecture a higher layer can make a request to any layer just below it.

D is wrong because the layer interface should always hide the details.

Source : Textbook. Page: 57- 58

According to van Steen and Tanenbaum a process is a instance of a computer program which is being executed. Which of the following statements about processes is correct?

A) In order to execute a program, an OS creates virtual processors, each one for running a different program.

B) The program counter is stored in a thread.

C) One thread can consist of multiple processes.

D) After a hardware interrupt, the process is shut down.

Answers:

A) is correct.

B) is false. The process context always contains the program counter.

C) is false. A process cant consist of multiple threads but not otherwise.

D) is false. The information about the interrupt is stored so the program can continue to run afterwards.

According to Van Steen and Tanenbaum, which of the following values of d C p (t) d tcorresponds to a fast clock?

A) 2

B) 1

C) 0.5

D) 0

Explanation:

According to the book, a perfect clock has $\frac{dC_p(t)}{dt}$ of 1, fast being > 1, and slow being < 1.

According to van Steen, which of the following is most true concerning the election of a super peer in a peer to peer network?

A. Super peers should be evenly distributed across the network

B. All normal peers should have access to a super peer, regardless of latency

C. Super peers should be able to serve any number of nodes

D. A super peer should be created for, at most, every five nodes

Answer:

A. is correct, super peers should be evenly distributed for the purposes of efficiency.

B. is incorrect, low-latency between a node and its assigned super peer is required.

C. is incorrect, super peers should not need to serve more than some fixed maximum of nodes.

D. is incorrect, although the amount of super peers should be proportional to the total number of nodes, it does not necessarily need to be a 1:5 ratio or under.

Source: Page 335

According to Van Steen, which of the following is true about the Berkeley algorithm?

A) The time server polls every machine from time to time to ask what time it is there.

B) It is suitable for a system in which machines have UTC receivers.

C) The time daemon's time is set automatically.

D) It is not sufficient that all machines agree on the same time.

A is correct.

B is incorrect because the Berkeley algorithm is suitable for systems in which no machine has UTC receivers.

C is incorrect because the time daemon's time must be set manually by the operator periodically.

D is incorrect because for many purposes it is sufficient that all machines agree on the same time.

Suppose there is a distributed system that uses a distributed algorithm mutual exclusion technique, so that a process needs permission from every other process in order to access a shared resource, and suppose that process A sends out a message requesting the use of resource x. Suppose process B also wants to access resource x but has not done so yet. According to Van Steen and Tanenbaum, how should B respond to A's message?

A) Compare the timestamp of the incoming message with the timestamp of the message it sent in order to determine the appropriate action

B) Queue the request and do not reply

C) Send back an OK message

D) Send back a Permission Denied message

Answer:

A) is correct. B should compare timestamps of A's message and the message that B sent to all the processes. If A's message has a lower timestamp, send back an ok message. If B's message has a lower time stamp, queue the incoming request and don't reply.

B) is incorrect. This is the appropriate behavior for a process that currently has access to the resource.

C) is incorrect. This is the correct behavior for a process that isn't accessing the resource and doesn't want to access the resource.

D) is incorrect. The algorithm described only involved sending out affirmative replies, not negative ones, and either way B would need to check whether its message has a lower timestamp before knowing which process would have next access to the resource.

Which of the following is true about token-based algorithms to mutual exclusion?

A. It is one of the two classifications of distributed mutual exclusion algorithms, the other being permission based.
B. For token based algorithms, a process wanting to access a shared resource first requires the permission from other processes.
C. Token-based algorithms are commonly cause starvation.
D. Token-based algorithms often have multiple tokens to access a shared resource.

Explanation:
A is correct. Van Steen classifies distributed mutual exclusion algorithms in two categories: token-based and permission based.
B is incorrect. This is false. This is for a permission-based algorithm.
C is incorrect. This is incorrect. Token-based algorithms avoid starvation.
D is incorrect. This is false. There is only one token in a token-based solution.

Source: Course Textbook

HW4

cs4113

According to the textbook, which general type of mutual exclusion algorithm functions by passing a special message between the processes?

A) Token-Based Algorithms

B) Permission-Based Algorithms

C) Centralized Algorithm

D) Distributed Algorithm

A) is correct. On page 322, it says "In token-based solutions mutual exclusion is achieved by passing a special message between the processes, known as a token."

B) Permission-Based algorithms work by requiring permission from other processes first (page 322).

C) A centralized algorithm designates a single process as the coordinator, with all processes that request resources having to go through the coordinator for permission (page 322-323).

D) A distributed algorithm is a specific type of algorithm that involves sending messages between processes containing the name of the resource, the process number, and the current logical time (pages 323-325).

---

there are two important types of design patterns that often applied to the organization of middleware: wrappers and interceptor. what does wrapper use for?

A. offer an interface acceptable to a client application, of which the functions are transformed into those available at the component.

B. handles all the accesses between different applications

C. break the usual flow of control and allow other code to be executed

D. assist in transferring the invocation to the target object

A is correct. as stated on page 72 about the wrapper that works as an adapter for the client application

B is incorrect because it is used of broker

C is incorrect because it is used of interceptor

D is incorrect because it is used of message-level-interceptor

---

Enterprise Java Beans(EJB), described under Note 3.7 in the textbook, provides an interface for underlying services implemented by the application server. Which definition below accurately defines a **stateless session bean** found from the text?

A.) "is an object that is invoked once, does its work, after which it discards any information it needed to perform the service is offered to a client."

B.) "maintains client-related state...when the client is finished (possibly having invoked the object several times), the bean will automatically be destroyed."

C.) "considered to be a long-living persistent object...stores information that may be needed a next time a specific client accesses the server."

D.) "is used to program objects that react to sending and receiving messages...contains application code for handling the message, after which the server simply
discards it."

A.) is correct. The definition and more information can be found on page 138.

B.) is incorrect. This definition defines **stateful session beans** and further information can be found on page 138.

C.) is incorrect. This definition defines **entity beans** and further information can be found on page 138-139.

D.) is incorrect. This definition defines **message-driven beans** and further information can be found on page 139.

According to the book, in most cases a server cluster is logically organized into how many tiers?

A. 3

B. 2

C. 4

D. 5

A is the correct answer because according to Section 3.4 the server cluster consists of the first tier where there is a" (logical) switch where client requests are routed", a tier that is dedicated for application processing, and the third tier which consists of data processing servers.

B is incorrect because it is not 3

C is incorrect because it is not 3

D is incorrect because it is not 3

According to the textbook, where are several ways to organize a server. How is the iterative server organized?

A) The server itself handles the request and, if necessary, returns a response to the requesting client.

B) Does not handle the request itself, but passes it to a separate thread or another process, after which it immediately waits for the next incoming request.

C) Doesn't keep information about the client and can change its own state without letting the client know.

D) By storing a piece of information in the client's browser.

Explanations:

A) Is correct as explained in section 3.4 page 129.

B) Is not correct as this is the concurrent server. It is explained in section 3.4 page 129.

C) Is not correct as this talks about a stateless server.

D) Is not correct as this explains a cookie.

According to section 3.2 in the textbook, why is virtualization important?

A) Because hardware changes faster than software

B) modify guest OS, either by preventing nonprivileged sensitive instructions or making them nonsensitive

C) does not isolate failed components

D) hide network latency by reacting to the next request while the previous one is being replied

A) A is correct. As stated on page 116, "One of the most important reasons for introducing virtualization back in
the 1970s was to allow legacy software to run on expensive mainframe hardware."

B) This is actually about paravirtualization

C) Virtualization isolates failed components

D) This is not about virtualization but a benefit of using threads

---

Which of the following is ***not*** one of the three models discussed in section 3.1 of the book for constructing a server?

A) Variable-state machine

B) Multithreading

C) Single-threaded process

D) Finite-state machine

A) Is correct since it is not discussed in the book (and is not a real server model).

B), C), and D) are found in figure 3.5 and discussed in the surrounding text in section 3.1.

Model	Characteristics
Multithreading	Parallelism, blocking system calls
Single-threaded process	No parallelism, blocking system calls
Finite-state machine	Parallelism, nonblocking system calls

Figure 3.5: Three ways to construct a server.

According to Dr. Pierre St. Juste's lecture, which of the requirements for distributed design (Network, Functions, Data) are protocol buffers directly tie to, as opposed to gRPC?

A) Data

B) Network & Functions

C) Data & Functions

D) Network

-

A) Correct. Protobufs are directly tied with only the Data requirement for distributed design. There was a slide in his presentation that said this, but those slides haven't been posted yet.

B) Incorrect. gRPC handles the Network and Function requirements. This was also covered in his presentation.

C) Incorrect.

D) Incorrect.

cs5113

Question: As per the textbook, basic idea behind Remote Procedure Call is to make a remote procedure call look like a local one. Which of the following is true about Remote Procedure Calls (RPC)?

A. For RPC to work properly, the caller and the callee should follow same protocol. Both the sides should agree on steps to follow and format of exchanged messages.

B. When a client calls a remote procedure, the client is not blocked until a reply is returned from server.

C. The client and the server don't have to agree on representation of simple data structures, such as integers, characters, Booleans, etc.

D. An RPC is not ideal for client-server applications

Answer A. is correct. As per the book, An RPC aims at hiding the complexities of message passing and is ideal for client-server applications. Both sides (Client and server) in an RPC should follow the same protocol

Answer B is incorrect: As quoted in textbook – "In conventional procedure calls, when a client calls a remote procedure, the client will block until a reply is returned. This request-reply behavior is unnecessary when there is no result to return or may hinder efficiency when multiple RPCs need to be performed."

Answer C is incorrect: both sides in an RPC should follow the same protocol including representation of simple data

Answer D is incorrect: As RPC is used in hiding most of complex steps of message passing and is ideal for client-server applications.

According to the information presented in the book, what is a downside to using the Berkeley algorithm for time synchronization?

A. An operator must manually enter the correct time periodically for accurate time keeping.

B. Every node in the network must have a UTC receiver.

C. Every node must periodically ask the time server for the correct time.

D. The algorithm assumes a strict hierarchy of nodes.

Answers:

A. Is correct; although the oversight of the other nodes is handled by the time server and **internally** the nodes will agree, the operator must provide accurate time if it is necessary.

B. This is not correct. In fact, the algorithm works if no nodes have UTC receivers.

C. This is not correct. The server is the active participant in this algorithm, and it periodically checks the nodes.

D. This is not correct. The books states no such requirement. Besides, all that is necessary is for the time server to be able to contact every node in the network in some way.

According to Van Steen and Tanenbaum, a server that accepts requests, passes them off to another thread or process, and then continues waiting for incoming requests is known as a:

A) concurrent server

B) iterative server

C) stateful server

D) stateless server

Answers:

A: Correct.

B: An iterative server handles requests itself and then continues to listen for incoming requests

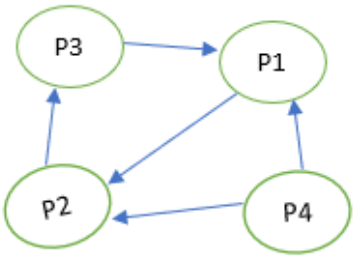
C: A stateful server could be set up concurrently, but the term refers to a server that maintains information about its clients

D: A stateless server could also be set up concurrently, but the term refers to a server that does not maintain information about its clients

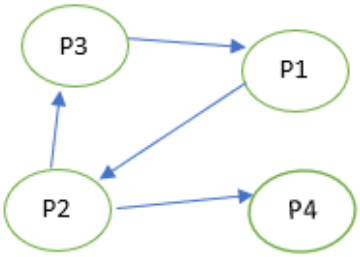
---

According to Dr. Dhall, identify the OR Model with deadlock.

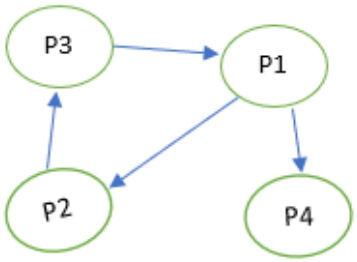
(A)



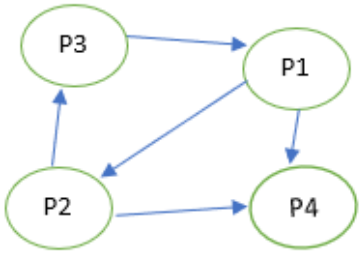
(B)



(C)



(D)



(A) is correct. Consider subgraph with nodes/Processes: {P1, P2, P3}. Reachability from P1 = {P2, P3, P1}, Reachability from P2 = {P3, P1, P2}, Reachability from P3 = {P1, P2, P3}. A knot is identified implies deadlock is present.

(B) is incorrect. We cannot identify the subgraph because reachability from P1, P2, P3 contains all nodes.

(C) is incorrect. No subgraph can be identified because reachability from P1, P2, P3 contains all nodes.

(D) is incorrect. Pick any subgraph, you cannot identify knot.

Source: Dr. Dhall's lecture material.

Q) According to Lecture notes, which of the following statements is not TRUE.

(A) In the AND model, cycle gives all the deadlocked processes.
(B) In the OR model, the presence of knot indicated a deadlock.
(C) Centralized deadlock detection algorithm detects phantom deadlocks.
(D) Ho-Ramamoothy 1-phase algorithm does not detect phantom deadlock.

(A) is false. A cycle does not give all the deadlocked processes in AND model refer slides 23, 24.
(B) is true, If we identify a knot implies deadlock in OR model.
(C) is true, centralized deadlock detection algorithm detect phantom deadlocks because of large communication delays.

(D) is true, because it eliminates inconsistency in state information by using only information common to both the tables. (Process and Resource Tables)

According to Van Steen and Tanenbaum, what is the main reason for performing a process migration?

A) Performance

B) Flexibility

C) Security

D) Mobility

Answers:

A) Correct

B) Incorrect, although this is a reason you might want to perform process migration this is not the main reason

C) Incorrect, this is an issue that is caused by dynamically downloading client-side software

D) Incorrect, this offers an advantage, but is not main the reason for process migration.

According to Dr. Dhall's lecture, which of the following statement is FALSE.

(A) Obermarck's Path Pushing Algorithm doesn't detect phantom deadlock.

(B) In Ho-Ramamoorthy 2-phase algorithm, a cycle in WFG doesn't necessarily imply a deadlock exists.

(C) Efficient resolution of deadlock requires knowledge of all processes and resources involved.

(D) In centralized deadlock detection algorithm, we use central coordinator to maintain graph for the entire state.

(A) is false. Because each site takes an asynchronous snapshot of the global view.
(B) is true, we may get consistent view of the state because of more communication but later proved that phantom deadlocks are not completely eliminated.
(C) is true, both are needed to decide the presence of deadlock.
(D) is true, central coordinator maintain the graph for the entire site.

According to Dr. Dhall's lectures, which class of deadlock detection suffers from large communication delays?

A) Centralized Deadlock Detection

B) Distributed Deadlock Detection

C) Hierarchical Deadlock Detection

D) Global State Deadlock Detection

Explanations:

A) Correct. Centralized deadlock detection, while easy to implement, incurs a great communication cost. This is because the central coordinator must retrieve information from **all** of the nodes to construct the corresponding wait-for-graph.

B) Incorrect. Distributed deadlock detection is controlled by each individual node, where each node tracks considerably more information locally. This information storage in turn reduces the amount of communication that has to be completed.

C) Incorrect. Hierarchical mitigates some of the communication errors present in centralized systems by limiting the search space for each node to just its descendants.

D) Incorrect. This is not a class, rather a specific method.

According to van Steen, which of the following is NOT one of the three segments of a process?

A. Stack segment

B. Code segment

C. Resource segment

D. Execution segment

Answer:

A. is correct as although a stack is stored in a process, it's handled by the execution segment

B. is incorrect. The code segment contains the instructions making up the program

C. is incorrect. The resource segment contains references to any external resources needed by the process.

D. is incorrect. The execution segment stores the execution state of the process.

Source: Page 155 Note 3.11

According to Van Steen, which of the following is true about Web browsers in the context of multithreaded clients?

A) Developing a browser as a multithreaded client is a good idea.

B) If a browser is developed as a multithreaded client, separate threads can be activated before the main HTML file is fetched.

C) When using a multithreaded client, browser connections cannot be set up to different replicas , therefore the data cannot be transferred in parallel.

D) When using a multithreaded client and a replicated server, the entire Web document takes more time to be displayed than in the case with a nonreplicated server.

A is correct since the browser is doing several tasks simultaneously, and it turns out that developing it as a multithreaded client simplifies matters considerably.

B is incorrect because separate threads can be activated only after the HTML file is fetched.

C is incorrect because browser connections can be set up to different replicas which allows the data to be transferred in parallel.

D is incorrect because when using a multithreaded client and a replicated server, the entire Web document is fully displayed in a much shorter time by taking advantage of the replicated servers (parallel data transfer).

According to Van Steen and Tanenbaum, which of the following pieces on information would NOT be part of a thread context?

A) Memory maps

B) The program counter

C) Whether the thread is currently blocked by a mutex

D) The stack pointer

Answer:

A) is correct. The current values held in memory is part of a process context but not a thread context.

B) is incorrect. The program counter is part of a processor context, and is therefore part of the thread context.

C) is incorrect. Information for thread management is part of a thread context.

D) is incorrect. The stack pointer may be part of a processor context and therefore may be part of the thread context.

Which of the following is true about user-level threads?
A. Using user-level threads makes it cheap to create and destroy threads.
B. Using user-level threads makes switching thread contexts take several instructions.
C. In user-level threads, multiple threads are mapped to multiple schedulable entities.
D. In user-level threads, the kernel is aware of threads and schedules them.

Explanation:
A is correct. This is the main advantage of using user-level threads.
B is incorrect. User-level threads make switching thread contexts takes a few instructions.
C is incorrect. In user-level threads, multiple threads are mapped to a single schedulable entity.
D is incorrect. This is true for kernel level threads, not user-level threads.

Source: Course Textbook

HW5

cs4113

According to the guest lecture by Dr. Pierre St. Juste, what is the main difference between JSON and Protobufs?

A) Protobufs are static, which means they can be optimized at compile time. JSON is dynamic.

B) Protobufs are abstractions from objects implemented in other languages.

C) JSON is more computationally efficient than Protobufs

D) Protobufs are language agnostic, while JSON is language dependent

A) is correct because Protobufs can be precompiled, whereas JSON is dynamic.

B) is incorrect because both Protobufs and JSON can be considered abstractions of objects

C) is incorrect in most cases because protobufs are compiled into objects of various languages, while JSON has to be parsed into objects.

D) is incorrect since JSON is also a language independent implementation of objects (despite JSON standing for Javascript Object Notation). There are implementations of JSON in most modern languages.

---

what will happen with the message after submitted by persistent communication combine with synchronous communication?

A. stored by the communication middleware as long as it takes to deliver it to the receiver and it will block the sender from submitting a new message

B. stored by the communication system only as long as the sending and receiving application are executing and it will block the sender from submitting a new message

C. stored by the communication middleware as long as it takes to deliver it to the receiver and sender can keep submitting a new message

D. stored by the communication system only as long as the sending and receiving application are executing and sender can keep submitting a new message

A is the right answer. we can find the answer on page 172, 173 in textbook. this combination is one of the most popular in the distributed system that helps to make sure the message will be sent and the sender could not send multiple messages at one time

B. combine of transient communication and synchronous communication

C. combine of persistent communication combine with asynchronous communication

D. combine of transient communication combine with asynchronous communication

---

According to the OSI model, beginning on page 164 in the textbook, communication protocols are divided into seven layers. Which definition below accurately defines the **Physical Protocol?**

A.) "Deals with standardizing how two computers are connected and how 0s and 1s are represented."

B.) "Provides support for sessions between applications."

C.) "Prescribes how data is represented in a way that is independent of the hosts on which communicating applications are running."

D.) "Provides the means to detect and possibly correct transmission errors, as well as protocols to keep a sender and receiver at the same pace"

A.) is correct

B.) Defines **Session Layer** and description can be found on pages 165-166 in the text.

C.) Defines **Presentation Layer** and description can be found on pages 165-166 in the text.

D.) Defines **Data Link Layer** and description can be found on pages 165-166 in the text.

According to the book in page 226, what is probabilistic flooding?

A) When a node is flooding a message m and needs to forward m to a specific neighbor, it will do so with a probability pflood.

B) Nodes form a bridge to other remote parts of the network; therefore, they should be contacted as soon as possible.

C) If node P has just been updated for data item x, it contacts an arbitrary other node Q and tries to push the update to Q. However, it is possible that Q was already updated by another node.

D) A way to rapidly propagate information among a large collection of nodes using only local information.

Correct Answers and Explaination:

A) Is correct according to the book's definition in page 226.

B) Not correct because it is directional gossiping in page 233.

C) Not correct because that is rumor spreading in page 231.

D) Not correct because that is epidemic protocols in page 229.

According to section 4.1 in the textbook, which of the following is the best description of persistent communication?

A) neither the sender nor the receiver needs to be up and running for message transmission to take place

B) no storage facilities are offered, so the receiver must be prepared to accept the message when it is sent.

C) the sender is allowed to continue immediately after the message has been submitted for transmission

D) the sender is blocked at least until a message has been received

A) is correct. "With persistent communication, a message that has been submitted for transmission is stored by the communication middleware as long as it takes to deliver it to the receiver. As a consequence, it is not necessary for the sending application to continue execution after submitting the message. Likewise, the receiving application need not be executing when the message is submitted."

B) is wrong since this is a description of transient communication

C) is wrong since this is a description of Asynchronous communication

D) is wrong since this is a description of Synchronous communication

---

```
According to the text book, which of the following is not one of the three
most important communication patterns supported by ZeroMQ?

  A. message-passing
  B. request-reply
  C. publish-subscribe
  D. pipeline


Answer found on page 200.

The three most important communication patterns supported by ZeroMQ are
request-reply, publish-subscribe, and pipeline.

  A. message-passing was taken from Message Passing Interface(MPI), and is
     not listed as one of the three important communication patterns
     supported by ZeroMQ.

  A. The request-reply pattern is used in traditional client-server
     communication,
     like the ones normally used for remote procedure calls.

  A. In the case of a publish-subscribe pattern, clients subscribe to
     specific
     messages that are published by servers.

  A. The pipeline pattern is characterized by the fact that a process
     wants to push out its results, assuming that there are other processes
     that
     want to pull in those results.
```

According to Van Steen, what is something that Interface Description Languages (IDLs) help with?

- A. Generating client and server stubs.

- B. Defining the message format for an RPC.
- C. Determining encoding rules for the RPC message.
- D. Implementing a communication service.

A. is the correct answer. "An interface specified in such an IDL is then subsequently compiled into a client stub and a server stub, along with the appropriate compile-time or run-time interfaces," (Van Steen, pg. 184).

B., C., and D. are all decisions made by the application developer.

According to the textbook, which of the following is not an operation of the Message Queue Interface (MQI)?

A) MQOPEN

B) MQPUT

C) MQGET

D) MQSEND

A, B, and C are incorrect as they are a part of the MQI

D) is incorrect as it is not an operation within the MQI

---

cs5113

Question: As per text book, Remote procedure calls have been widely adopted as the basis of middleware and distributed systems in general. Which of the following is true about the DCE RPC (Distributed Computing Environment RPC)

A. DCE is a true middleware system. It is designed to be executed as a layer of abstraction between distributed applications and existing network.

B. The existing applications would be affected if customer add the DCE to a collection of machines and try to execute it as a distributed applications.

C. The programming model underlying DCE is not a client-server model.

D. As RPC can hide details, clients and servers are highly dependent of one another and can't be written in different programming languages

Solution:

A. is correct answer: As per textbook, DCE RPC closely represents the idea behind traditional RPC systems that uses a combination of interface specifications and explicit bindings to various programming languages. The RPC system would allow a client to call a local procedure for accessing a remote service.

B. is incorrect: DCE software would allow the customer to use a collection of machines as a distributed applications without affecting individual functionality

C is incorrect: DCE is based on a client-server model.

D. is incorrect. As per text book, A client and server can run on different hardware platforms and use different operating systems using the ability of RPC to hide details effectively

---

According to van Steen and Tanenbaum, which of the following statements is true about RPCs?

A. The client should not be able to tell that a procedure is not implemented locally by its usage alone.

B. It's important that the client and server have the same memory so that pointers work on both machines.

C. RPCs must be asynchronous if they are to be transparent to the user.

D. RPCs must only send one parameter at a time.

Answers:

A. This is correct. Th RPC should appear to be a locally defined function to the programmer, with its *send* and *receive* abstracted away from the user.

B. This is false. Pointers are only locally meaningful. RPCs must implement ways around pointer parameters.

C. This is false. RPCs can be synchronous or asynchronous.

D. This is false. There is no inherent limitation to the number of parameters an RPC may operate on.

According to van Steen and Tanenbaum, an RPC that can be be repeated multiple times without harm (because it does not modify the state of the system) is referred to as:

A) idempotent

B) at-most-once

C) a multicast RPC

D) an asynchronous RPC

Answers:

A: correct

B: An at-most-once operation is one that can be performed only once, and if a crash occurs, a client cannot repeat the operation in case it had already been carried out before the crash

C: A multicast RPC is one that is sent to a group of servers in parallel, which process the call in parallel.

D: An asynchronous RPC is one where the server responds immediately to indicate that it has received the request, and the client is able to continue while the server processes the request. This is often combined with a callback that is called when the request is finished.

According to van Steen and Tanenbaum, which of the following statement about OSI layers is incorrect.

(A) Mobility layer: Mainly contains the protocols for directly supporting applications.

(B) Network layer: Contains the protocols for routing a message through a computer network.

(C) Presentation layer: Prescribes how data is represented in a way that is independent of the hosts on which communicating applications are running.

(D) Data link layer: Provides the means to detect and possibly correct transmission errors.

Answer:

(A) Incorrect: Transportation layer contains the protocols for directly supporting applications.

(B) Correct: Protocols for routing a message through a computer network is the duty of network layer.

(C) Correct: Data representation is taken care by presentation layer.

(D) Correct: Data link layer contains protocols to keep sender and receiver in the same pace, in addition to detecting and correcting transmission error.

Source: Course Textbook, Chapter 4, P166.

Q) According to van Steen and Tanenbaum, _____ is the service in the distributed environment to find a network address with regards to the name of the network.

A) Domain Name System (DNS)

B) Internet Protocol System (IPS)

C) Transport Protocol System (TPS)

D) Transmission Control Protocol (TCP)

Answer:

A) Correct - DNS a distributed service that is used to look up a network address associated with the name of the network.

B) Incorrect - IP is used to transmit the data packets across the network without any setup

C) Incorrect - TCP implements services that are not provided at the interface of the network layer, but which are reasonably needed to build network applications

D) Incorrect - TCP is used as a standard for network communication

---

According to van Steen and Tanenbaum which of the following is the Internet Transport Protocol?

A) Transmission Control Protocol (TCP)

B) Universal Datagram Protocol (UDP)

C) File Transfer Protocol (FTP)

D) Trivial File Transfer Protocol (TFTP)

Answers:

A) Correct

B) Incorrect, UDP is part of the Internet sweet but is connectionless communication.

C) Incorrect, FTP is not a part of the transport layer and is part of the application layer, but does use TCP as its transport layer.

D) Incorrect, TFTP is not a part of the transport layer and is part of the application layer, and uses UDP as its transport layer.

According to van Steen and Tanenbaum, which of the following is correct about connection-oriented service :

A) Before sending data the sender and receiver first explicitly establish a connection, and possibly negotiate specific parameters of the protocol they will use.

B) The connection stays even after the exchange of data is done.

C) The sender sends message whenever it is ready.

D) There is no need to setup before starting the communication process.

Answers:

A: correct

B: When the sender and receiver are done exchanging messages, they release (terminate) the connection.

C: The sender first has to setup a connection with the receiver to send a message.

D: The sender and receiver explicitly establish a connection before sending a message.

According to Van Steen and Tanenbaum, a multicast/broadcast communication is used to send data to multiple receivers in distributed systems. Which of the following statements about multicasts/broadcasts is correct?

A) A naive approach to send data to every node in the network is to apply flooding.

B) Multicast means that the message is sent to every node in the network.

C) Epidemic protocols are the same as a simple Broadcast.

D) Deleting a data item in an epidemic algorithm is simple.

Answers:

A) is correct. Each node forwards the message to all neighbors except to the one which it received the message.

B) is false. That is broadcasting. Multicast is sending a message to a subset of all nodes in the network.

C) is false. Epidemic protocols rely on the spread of the message by sending the message to some specific neighbors.

D) is false. The deletion of the data item destroys all information on that item. Therefore, a node may interpret old copies as a update.

According to van Steen and Tanenbaum, in an epidemic model for information dissemination, which of the following nodes has yet to see the data?

A) Susceptible nodes

B) Infected nodes

C) Removed nodes

D) Quarantined nodes

Explanation:

A) Correct. A susceptible node has not yet received the information.

B) Incorrect. An infected node has received the data and is willing to spread it to other nodes

C). Incorrect. A removed node is a node that is not willing or is not able to spread the data

D) Incorrect. A Quarantined node is not really an node type, but it fits the epidemic theme.

According to van Steen and Tanenbaum, which of the following is not listed as one of the three most important communication patterns in ZeroMQ?

A. One-Way

B. Pipeline

C. Request-Reply

D. Publish-Subscribe

Answer:
A. is correct. One-Way messaging is not a communication pattern present in ZeroMQ.

B. is incorrect, The Pipeline pattern is used in ZeroMQ to push results from processes out to the first available process requesting to pull those results, with the goal of having as many processes working as possible.

C. is incorrect. The Request-Reply pattern is used in ZeroMQ to facilitate two way communication between a server and client without the need for a client to constantly listen. The client sends a request message to the server, which sends back a reply and a call to send.

D. is incorrect, The Publish-Subscribe pattern is used in ZeroMQ to let clients listen to certain servers they subscribe to, with messages from servers only being pushed out to clients that are subscribed to it.

Source: Pages 200-202

According to van Steen and Tanenbaum, which of the following is not listed as one of the three most important communication patterns in ZeroMQ?

A. One-Way

B. Pipeline

C. Request-Reply

D. Publish-Subscribe

Answer:
A. is correct. One-Way messaging is not a communication pattern present in ZeroMQ.

B. is incorrect, The Pipeline pattern is used in ZeroMQ to push results from processes out to the first available process requesting to pull those results, with the goal of having as many processes working as possible.

C. is incorrect. The Request-Reply pattern is used in ZeroMQ to facilitate two way communication between a server and client without the need for a client to constantly listen. The client sends a request message to the server, which sends back a reply and a call to send.

D. is incorrect, The Publish-Subscribe pattern is used in ZeroMQ to let clients listen to certain servers they subscribe to, with messages from servers only being pushed out to clients that are subscribed to it.

Source: Pages 200-202

According to Van Steen and Tanenbaum, which of the following is an example of a connection-oriented service?

A) TCP

B) IP

C) Sending a letter

D) UDP

Answer:

A) is correct. TCP establishes a connection for the whole message, then the connection is terminated at the end.

B) is incorrect. IP is only concerned with routing each packet to its destination independent of all others. No internal path is selected and remembered.

C) is incorrect. The sender transmits the message when it is ready and there is no connection or setup needed in advance.

D) is incorrect. UDP only has minor additions to IP, and is a connectionless transport protocol.

Which of the following is true about persistent communication?
A. In persistent communication a message that has been submitted for transmission is stored by the communication middleware as long as it takes to deliver it to the receiver.

B. In persistent communication, it is necessary for the sending application to continue execution after submitting the message.
C. Persistent communication is the same as synchronous communication.
D. In persistent communication, a receiving application must be executing when the message is submitted.

Explanation:
A is correct. This is the definition of persistent communication.
B is incorrect. There is no requirement for the sender to continue in execution after sending the message.
C is incorrect. Synchronous communication is different. In synchronous communication, the sender is blocked until its request is known to be accepted.
D is incorrect. There is no requirement for this in persistent communication.

Source: Course Textbook, Chapter 4 Section 1, p172-173