# **Distributed and Cloud Systems Programming**

## MPI

A distributed system is defined as a collection of autonomous computers linked by a network with software designed to produce an integrated computing facility. True or False?

1. False
2. True

If one site fails in a distributed system …………..

1. The remaining sites can continue operating
2. All the sites will stop working
3. Directly connected sites will stop working
4. None of the mentioned

MPI stands for

1. Method Passing Interface
2. Message Passing Interface
3. Message Passing Information
4. Method Passing Information

In MPI global communication functions that can involve more than two processes. These functions are called…

1. Synchronous Communication
2. Asynchronous Communication
3. Collective Communication
4. None of the above

MPI is specifically used for parallel programming.

1. TRUE
2. FALSE

In MPI, which function is commonly used to send a message from one process to another?

1. MPI\_Receive
2. MPI\_Send
3. MPI\_Comm\_rank
4. MPI\_Bcast

Which MPI function is used to initialize the MPI environment?

1. MPI\_Initialize
2. MPI\_Init
3. MPI\_Comm\_init
4. MPI\_Start

What is the purpose of the MPI\_Comm\_size function in MPI?

1. To initialize MPI communication
2. To get the rank of the process
3. To determine the size of the communicator
4. To finalize MPI communication

In MPI, what is the purpose of the MPI\_Comm\_rank function?

1. To get the rank of the process
2. To determine the size of the communicator
3. To initialize MPI communication
4. To finalize MPI communication

In MPI, what does the term “rank” refer to?

1. The size of the communicator
2. The process identifier
3. The message size
4. The type of data being sent

What is the purpose of the MPI\_Finalize function?

1. To close MPI communication channels
2. To signal the end of the MPI program
3. To synchronize MPI processes
4. To free up memory allocated for MPI communication

Which of the following command are used to compile MPI program

1. mpiexec
2. mpicc
3. mpcc
4. mpexec

Which of the following command are used to run MPI program

1. mpiexec
2. mpicc
3. mpexec
4. mpcc

Which of the following is/are predefined data type?

1. MPI\_CHAR
2. MPI\_SHORT
3. MPI\_LONG
4. All of the above

What is the format for executing MPI programs?

1. mpiexec <number of processes>-n ./mpi\_hello
2. mpiexec ./mpi\_hello -n <number of processes>
3. mpiexec <number of processes> ./mpi\_hello -n
4. mpiexec -n <number of processes> ./mpi\_hello

What is the command used to execute the mpi\_hello program with 2 processes?

1. mpiexec 2 -n ./mpi\_hello
2. mpiexec -n 2 ./mpi\_hello
3. mpiexec ./mpi\_hello -n 2
4. mpiexec ./mpi\_hello -n 2

Which of the following commands determines the size of the group associated with a communicator?

1. MPI\_COMM\_WORLD
2. MPI\_COMM\_SIZE
3. MPI\_COMM\_RANK
4. MPI\_COMM\_GROUP

The format of MPI\_COMM\_SIZE is \_\_\_\_\_\_\_\_\_

1. int MPI\_Comm\_size(MPI\_Comm comm, int \*size)
2. int MPI\_Comm\_size(int \*size, MPI\_Comm comm)
3. int MPI\_Comm\_size(MPI\_Comm comm)
4. int MPI\_Comm\_size(int \*size)

Which of the following commands determines the rank of the calling process in the communicator.

1. MPI\_COMM\_WORLD
2. MPI\_COMM\_SIZE
3. MPI\_COMM\_RANK
4. MPI\_COMM\_GROUP

The format of MPI\_COMM\_RANK is \_\_\_\_\_\_\_\_\_

1. int MPI\_Comm\_ rank (MPI\_Comm comm)
2. int MPI\_Comm\_ rank (int \*rank, MPI\_Comm comm)
3. int MPI\_Comm\_rank (MPI\_Comm comm, int \*rank)
4. int MPI\_Comm\_ rank (int \*rank)

Which of the following statements is/are true about MPI\_INIT and MPI\_FINALIZE?

1. MPI\_INIT must be called before any other MPI function and called exactly once per process
2. No further MPI function can be called after MPI\_FINALIZE
3. Both a. and b.
4. None of the above

The format of MPI\_SEND is \_\_\_\_\_\_\_\_\_\_\_\_

1. int MPI\_Send(int count, int dest, int tag, const void \*buf, MPI\_Datatype datatype, MPI\_Comm comm)
2. int MPI\_Send(int count, MPI\_Datatype datatype, const void \*buf, int dest, int tag, MPI\_Comm comm)
3. int MPI\_Send(const void \*buf, int count, MPI\_Datatype datatype, int dest, int tag, MPI\_Comm comm)
4. int MPI\_Send(const void \*buf, int tag, MPI\_Datatype datatype, int dest, int count, MPI\_Comm comm)

## 

## Advanced MPI

In MPI, what does the term “blocking communication” mean?

1. Processes are not allowed to communicate
2. Processes can only communicate within a communicator
3. Processes are synchronized until the communication completes
4. Processes can communicate concurrently without synchronization

In MPI, what does the term “non-blocking communication” mean?

1. Processes are not allowed to communicate
2. Processes can only communicate within a communicator
3. Processes can communicate concurrently without synchronization
4. Processes are synchronized until the communication completes

In MPI, which function is used for collective communication to broadcast data from one process to all others?

1. MPI\_Scatter
2. MPI\_Gather
3. MPI\_Bcast
4. MPI\_Reduce

Which MPI function is used for blocking point-to-point communication to receive a message?

1. MPI\_Recv
2. MPI\_Send
3. MPI\_Irecv
4. MPI\_Isend

Which MPI function is used for collective communication to scatter data from one process to all others?

1. MPI\_Scatter
2. MPI\_Gather
3. MPI\_Bcast
4. MPI\_Reduce

In MPI, what is the purpose of MPI\_Reduce function?

1. To initialize MPI communication
2. To synchronize processes within a communicator
3. To reduce data across all processes
4. To broadcast data from one process to all others

In MPI, what is the purpose of MPI\_Barrier function?

1. To initialize MPI communication
2. To synchronize processes within a communicator
3. To broadcast data from one process to all others
4. To reduce data across all processes

In MPI, what does the term “communicator” refer to?

1. The message size
2. The process identifier
3. A group of processes that can communicate with each other
4. The size of the communicator

What is the purpose of MPI\_Gather function in MPI?

1. To scatter data from one process to all others
2. To gather data from all processes within a communicator
3. To broadcast data from one process to all others
4. To reduce data across all processes

## AKKA Framework (Actor Model)

What is Akka in the context of Java distributed systems?

1. A distributed database management system
2. An actor-based concurrency toolkit
3. A message queuing system
4. A distributed file system

Which of the following best describes an actor in Akka?

1. An object that represents a user in a distributed system
2. A lightweight process that communicates asynchronously via messages
3. A data structure used for caching in distributed systems
4. A network protocol for inter-process communication

In Akka, how are messages typically sent between actors?

1. Through shared memory
2. Via RESTful APIs
3. Using synchronous function calls
4. Asynchronously through mailboxes

What is the purpose of the actor hierarchy in Akka?

1. To define the network topology of the distributed system
2. To manage the lifecycle of actors and their supervision
3. To enforce security policies within the system
4. To optimize message routing between actors

How does Akka ensure fault tolerance in distributed systems?

1. By using distributed consensus algorithms
2. By replicating data across multiple nodes
3. By isolating failures and supervising actor lifecycles
4. By using strong encryption for data transmission

Which Akka component is responsible for managing actor lifecycles and supervising child actors?

1. Mailbox
2. Router
3. Supervisor
4. Dispatcher

How does Akka handle message delivery guarantees between actors?

1. It guarantees exactly-once message delivery
2. It guarantees at-least-once message delivery
3. It guarantees at-most-once message delivery
4. It does not provide any delivery guarantees

Which Java library provides support for building distributed systems using the Akka actor model?

1. Apache Kafka
2. Netty
3. Spring Framework
4. Akka Framework

Can actors modify their internal state?

1. Yes, actors can modify their state freely
2. No, actors have immutable state
3. Only if they are explicitly granted permission by the supervisor
4. Only during actor initialization

How does Akka ensure that messages are processed in the order they are received by an actor?

1. By enforcing strict message ordering at the network level
2. By using a single-threaded execution model for each actor
3. By implementing a FIFO queue for each actor's mailbox
4. By periodically reordering messages based on priority levels

Which of the following is NOT a characteristic of the Akka actor model?

1. Location transparency
2. Concurrency control
3. Actor isolation
4. Shared memory access

What happens when an actor receives a message?

1. It immediately processes the message synchronously
2. It adds the message to its mailbox for later processing
3. It forwards the message to the supervisor actor

How are actors created in the Akka actor model?

1. By instantiating new objects directly
2. By sending a "create actor" message to a predefined system actor

How do actors communicate in the Akka actor model?

1. Through synchronous method calls
2. Via asynchronous message passing
3. By using a centralized message bus

In the Akka actor model, how does an actor handle failure?

1. By crashing the entire system
2. By ignoring the failure and continuing execution
3. By restarting itself or its children according to a supervision strategy

## 

## 

## Apache Spark

What is Apache Spark?

1. A big data processing engine
2. A machine learning algorithm
3. A data visualization tool
4. A database management system

Which programming languages are supported by Apache Spark?

1. Java, Python, and Scala
2. Java, Python, and C#
3. Java, PHP, and Scala
4. Java, Python, and Ruby

What is an RDD?

1. A Resilient Distributed Dataset
2. A Remote Data Depot
3. A Replicated Database Descriptor
4. A Recursive Data Definition

What is a DAG in Spark?

1. A Directed Acyclic Graph
2. A Distributed Analysis Graph
3. A Data Access Graph
4. A Dynamic Aggregation Graph

What is a Spark SQL?

1. A Spark library for machine learning
2. A Spark library for graph processing
3. A Spark library for stream processing
4. A Spark module for working with structured and semi-structured data

Which of the following is a benefit of using Spark SQL?

1. Faster data processing
2. Better support for unstructured data
3. Easier integration with NoSQL databases
4. More flexibility in data processing operations

What is Spark Streaming?

1. A Spark module for batch processing
2. A Spark module for real-time stream processing
3. A Spark module for graph processing
4. A Spark module for machine learning

What is a Spark ML?

1. A library for machine learning in Spark
2. A module for stream processing in Spark
3. A module for working with structured and semi-structured data in Spark
4. A module for batch processing in Spark

Which of the following is a feature of Spark SQL?

1. Ability to process data in real-time
2. Support for machine learning algorithms
3. Integration with Hadoop Distributed File System (HDFS)
4. Support for SQL-like queries

Which of the following is NOT a supported storage level for RDDs in Apache Spark?

1. MEMORY\_ONLY
2. DISK\_ONLY
3. MEMORY\_AND\_DISK\_SER
4. HDFS

What is the main advantage of using Apache Spark over traditional MapReduce for data processing?

1. Spark provides a simpler programming model
2. Spark is optimized for batch processing only
3. Spark has better fault tolerance mechanisms
4. Spark runs on a single node, making it easier to manage

What is Apache Spark Core?

1. A distributed data processing engine
2. A relational database management system
3. A stream processing framework
4. A machine learning library

## 

## 

## Virtualization and Cloud Computing

“Cloud” in cloud computing represents what?

1. Wireless
2. Hard drives
3. People
4. Internet

Which of the following is the deployment model?

1. Public
2. Private
3. Hybrid
4. All of the mentioned

Which of the following is the best known service model?

1. SaaS
2. IaaS
3. PaaS
4. All of the mentioned

What is the full form of AWS?

1. Amazon web-based service
2. Amazon web-store service
3. Amazon web service
4. Amazon web-data service

Which of these is not a service model in the cloud?

1. Hardware as a service
2. Infrastructure as a service
3. Software as a service
4. Platform as a service

Amazon AWS consists of the following services

1. IaaS
2. PaaS
3. SaaS
4. All of the above

Virtual Machines in clouds are offered by?

1. SaaS
2. PaaS
3. IaaS
4. None of these

Google Apps is a suite of cloud computing ..... applications that include e-mail (Gmail), Organizer (Google Calendar), Word Processing documents ( Google Docs) and others.

1. IaaS
2. PaaS
3. SaaS
4. None of the above

What is a private cloud?

1. A standard cloud service offered via the Internet
2. A cloud architecture maintained within an enterprise data center.
3. A cloud service inaccessible to anyone but the cultural elite
4. None of these

Which of the following cloud concepts is related to pooling and sharing of resources?

1. Polymorphism
2. Abstraction
3. Virtualization
4. None of the mentioned

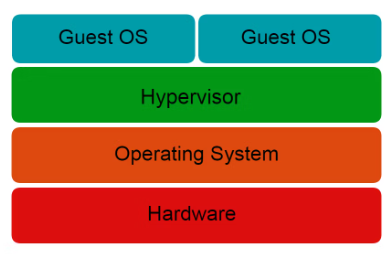
Bare metal hypervisor sits on top of

1. Physical hardware
2. Virtualized hardware
3. Operating system
4. None of the above

Which one of the following cloud concepts is related to sharing and pooling the resources?

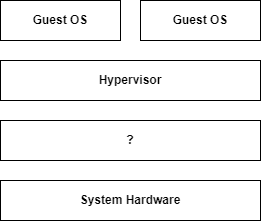
1. Polymorphism
2. Virtualization
3. Abstraction
4. None of the mentioned

Which type of Hypervisor is shown in the following figure?



1. Type 1
2. Type 2
3. Type 3
4. All of the mentioned

Which of the following should be placed in the second lowermost layer of the following figure?



1. Host Operating System
2. Software
3. VM
4. None of the mentioned

## 

## Virtual Machines (EC2/ Azure Virtual Machines)

Amazon Elastic Compute Cloud (Amazon EC2) is a web service based on

1. IaaS
2. PaaS
3. SaaS
4. All of the mentioned

Amazon EC2 provides virtual computing environments, known as:

1. chunks
2. instances
3. messages
4. None of the mentioned

Which of the following operating systems is NOT supported by Azure Virtual Machines?

1. Windows Server
2. Ubuntu
3. CentOS
4. MacOS

What is the Azure service that allows you to automatically scale the number of Virtual Machine instances based on demand?

1. Azure Autoscale
2. Azure Virtual Machine Scale Sets
3. Azure Load Balancer
4. Azure Traffic Manager

Which of the following is NOT a valid EC2 instance purchasing option?

1. On-Demand Instances
2. Spot Instances
3. Reserved Instances
4. Dynamic Instances

## AWS Lambda (Serverless Functions)

What programming languages are supported by AWS Lambda?

1. Java and Python
2. Python, Node.js, Java, C#
3. Python, Node.js, Ruby
4. Python, Node.js, C++

What is the maximum execution time for a single invocation of an AWS Lambda function?

1. 1 minute
2. 5 minutes
3. 15 minutes
4. 30 minutes

Which AWS service can directly trigger an AWS Lambda function?

1. Amazon EC2
2. Amazon S3
3. Amazon DynamoDB
4. Amazon RDS

How is AWS Lambda priced?

1. Based on the number of function invocations
2. Based on the duration of function execution
3. Based on the amount of memory allocated
4. All of the above

What is Amazon CloudWatch primarily used for?

1. Managing IAM users and roles
2. Monitoring and logging AWS resources
3. Storing relational data in the cloud
4. Deploying serverless applications

What type of data can CloudWatch collect and monitor?

1. Metrics
2. Logs
3. Events
4. All of the above

## 

## Content Delivery Network (CDN)

Which of the following is an edge-storage or content-delivery system that caches data in different physical locations?

1. Amazon Relational Database Service
2. Amazon SimpleDB
3. Amazon Cloudfront
4. Amazon Associates Web Services

Which of the following is also referred to as edge computing?

1. CloudWave
2. CloudFront
3. CloudSpot
4. All of the mentioned

Which of the following can be considered as a distributed caching system?

1. CND
2. CDN
3. CWD
4. All of the mentioned

Which HTTP status code indicates that a resource has been successfully retrieved from a CDN cache?

1. 200 OK
2. 404 Not Found
3. 503 Service Unavailable
4. 304 Not Modified

Which of the following is a feature of CDNs that helps protect against Distributed Denial of Service (DDoS) attacks?

1. Web Application Firewall (WAF)
2. Domain Name System (DNS)
3. Content Distribution Network (CDN)
4. Virtual Private Network (VPN)

What is the primary purpose of a Content Delivery Network (CDN)?

1. Storing relational data
2. Accelerating content delivery
3. Managing virtual machines
4. Processing real-time data

What is the name of the Amazon Web Services (AWS) service that provides a CDN solution?

1. Amazon S3
2. Amazon CloudFront
3. Amazon Route 53
4. Amazon Elastic Load Balancing

Which protocol is commonly used by CDNs to deliver content?

1. FTP
2. SMTP
3. HTTP(S)
4. SSH

## 

## Storage (S3)

Amazon S3 is also known as

1. Secured Storage Service
2. Simple Storage Service
3. Self Storage Service
4. Safe Storage Service

By default, S3 buckets and the objects in them are \_\_\_\_\_\_\_.

1. Public
2. Private
3. Both Public and Private
4. None

After you create a bucket, you can change the name of the bucket or its Region later on.

1. True
2. False

You can store \_\_\_\_\_\_\_\_ objects in a bucket and can have up to \_\_\_\_\_\_\_ buckets in your account.

1. 1 Million, 1000
2. 1 Billion, 10000
3. Any number of, 100
4. I Billion, Infinite

What is Amazon S3 primarily used for?

1. Virtual machine management
2. File storage and object storage
3. Relational database management
4. Container orchestration

What is the maximum size for an individual object stored in Amazon S3?

1. 1 GB
2. 5 GB
3. 10 GB
4. 100 GB

Which of the following is NOT a valid storage class in Amazon S3?

1. S3 Standard
2. S3 Amazon Glacier
3. Cold Storage
4. S3 standard insufficient access

How can you change the access permissions of an Amazon S3 bucket from private to public?

1. Using Bucket Policies
2. Modifying Access Control Lists (ACLs)
3. Setting up Signed URLs
4. All of the above

## 

## AWS Elastic Beanstalk

## Amazon Web Services (AWS) Elastic Beanstalk, Google App Engine and Heroku, are some of the example of providing services for:

1. IaaS
2. SaaS
3. PaaS
4. None

What is AWS Elastic Beanstalk primarily used for?

1. Database management
2. Code deployment and scaling web applications
3. Virtual machine management
4. File storage

Which of the following programming languages is NOT supported by AWS Elastic Beanstalk?

1. Java
2. Python
3. Ruby
4. C++

Which of the following AWS services can trigger an Elastic Beanstalk application deployment?

1. Amazon S3 events
2. Amazon SQS messages
3. Amazon SNS notifications
4. All of the above

What is the name of the configuration file used in Elastic Beanstalk to customize the environment settings and application deployment?

1. eb.ini
2. app.config
3. .ebextensions
4. config.yaml

Which protocol is commonly used by Elastic Beanstalk as the default for communication between the client and the application servers?

1. FTP
2. SMTP
3. HTTP
4. SSH

What options can be used to host an application that uses Apache / NGINX and is scalable at any point in time? Choose 2 correct answers.

1. AWS EC2
2. AWS Elastic beanstalk
3. AWS SQS
4. AWS ELB

## 

## AWS DynamoDB

DynamoDB is \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ type of storage

1. NonSQL
2. SQL
3. Simple Storage
4. All of the above

What type of NoSQL database is Amazon DynamoDB?

1. Document-based
2. Key-value
3. Columnar
4. Graph-based

What is the primary key structure in DynamoDB?

1. Partition key
2. Sort key
3. Composite key
4. All of the above

Which data types are supported by DynamoDB?

1. String, Number, Boolean
2. String, Number, Binary
3. String, Number, List, Map
4. String, Number, Set