SAMPLE CLOUD COMPUTING QUESTIONS

1. Is it possible to have an application-level virtualization? Support your answer by explaining the reasons and examples.

2. Design a WordCount system using the MapReduce techniques. Draw a figure and explain how your design works based on the given two sentences.

   <u>Input:</u>

   ```
   This is an apple.
   Apple is red in color.
   ```

3. What is serverless computing? Is it the same as function as a service? Why is it called serverless?

4. Considering using a control theory to optimize cloud resource management, what are the sources of instability when applying the control theory?

Sample Multiple Choice Question

1. Which component in GFS is primarily responsible for managing metadata, such as namespace and file-to-chunk mapping?
   - A. The chunk server
   - B. The client interface
   - C. The replication manager
   - D. The single master node
   - E. The distributed lock service
   - F. The metadata proxy

2. Which statement best describes the role of chunk servers in GFS?
   - A. They maintain the file namespace and metadata
   - B. They manage client authentication and access control
   - C. They store the actual data chunks and handle read/write requests
   - D. They provide encryption services for data in transit
   - E. They balance the overall network load
   - F. They directly communicate with the master for all operations

3. Why does GFS use a single master node?
   - A. To eliminate the need for distributed consensus
   - B. To centralize metadata management and simplify coordination
   - C. To increase data replication performance
   - D. To allow clients to access data directly without authentication
   - E. To balance network traffic between different data centers
   - F. To store both metadata and actual file contents

4. What is the primary goal of the Ray framework?
   - A. To provide a specialized deep learning framework for neural networks
   - B. To optimize traditional batch processing for big data analytics
   - C. To support distributed execution of AI and machine learning applications
   - D. To replace Kubernetes for container orchestration
   - E. To provide a real-time streaming system for IoT data processing
   - F. To offer a cloud-based service for GPU-based computations

5. What component in Ray is responsible for managing distributed task execution?
   - A. The Ray Client
   - B. The Ray Master Node
   - C. The Global Control Store (GCS)
   - D. The Ray Trainer Module
   - E. The Cluster State Manager
   - F. The AI Execution Engine

6. Which of the following is NOT a key feature of Ray?

    A. Support for heterogeneous computing environments

    B. Built-in distributed reinforcement learning support

    C. Automatic scheduling of machine learning training jobs

    D. Direct execution of SQL queries on large datasets

    E. Fault-tolerant task execution

    F. Actor-based execution for dynamic workloads

---

7. What is a Borgcell?

    A. A collection of machines managed by a single Borg master

    B. A lightweight virtual machine used in Borg

    C. A distributed file system for storing application data

    D. A backup cluster used for fault tolerance

    E. A specialized scheduling algorithm within Borg

    F. A network partitioning mechanism

---

8. What component of Borg is responsible for tracking job and task states?

    A. The Borgmaster

    B. The Node Agent

    C. The Borg UI

    D. The Borg Database

    E. The Priority Queue Manager

    F. The Job Monitor