Sample Questions: CS3204

All lectures, from Lecture-1 to Lecture-21, are required for the exam. The following list includes review questions that can help you to prepare for the exam. Apart from the below questions, review all exercises and problems proposed or discussed during the lectures. Exam questions can be phrased differently.

Introduction to Cloud Computing, Computer Systems and Computer Networks

- Provide a definition of cloud computing in your own words.
- Why people prefer to run applications on the cloud instead of on on-premise "bare metal" servers?
- What are the disadvantages of running applications on the cloud?
- Explain the components of a computer system.
- What is the function of an operating system? What is the purpose of a bootstrap program?
- Why is a memory hierarchy necessary in a computer system? Draw a commonly used memory hierarchy diagram.
- Explain capacity and propagation delay.
- What is the difference between circuit switching and packet switching?
- Describe the layered internet protocol stack.
- What causes packet loss in communication networks?

Distributed Computing Systems

- Discuss the merits and demerits of using distributed computing systems.
- Explain the main differences between a cluster and a grid computing.
- Describe the most common computing model for a cluster of computers.
- Explain Amdahl's Law and its equation.
- Solve math exercises associated with Amdahl's Law.
- Provide a graphical explanation of Amdahl's Law with various cases (e.g., $n = 1$, $p = 1$).
- Justify the statement: "Amdahl's Law provides a limit or upper bound on the achievable speedup in a parallel computing system."
- What are the common design goals of a distributed system? Provide a brief description of each.
- What does transparency mean in a distributed system? Explain the various types of transparencies in distributed system (e.g., location, replication).
- Solve math exercises associated with analyzing size scalability and its implications under different utilization values.
- Explain the client-server architecture with a suitable timing diagram and example.
- Illustrate the spectrum of choices for multi-tiered client-server architecture using a diagram.
- What are peer-to-peer architectures in distributed systems? Compare them with client-server architectures.
- Discuss implementation choices in peer-to-peer computing systems.

- What is Chord? List its basic features.
- Explain the insertion and searching steps in Chord.
- Solve example exercises related to Chord.
- What are the basic characteristics of a finger table in Chord?

Virtualization

- What is virtualization? Explain the different types of virtualization.
- What is a hypervisor? Explain the differences between Type 1 and Type 2 hypervisors.
- Define full virtualization, para-virtualization, and emulation.
- What are the benefits of virtualization? How is virtualization used?
- Explain memory and I/O virtualization.
- What is virtual machine (VM) migration? Why is VM migration necessary?
- Explain the different types of VM migration.
- List the steps of pre-copy VM migration.
- List the steps of post-copy VM migration.
- Solve math exercises related to pre-copy and post-copy VM migrations.
- What is container technology in the context of virtualization?
- Explain the motivation behind using containers, highlighting the limitations of virtual environments and virtual machines.
- What are the benefits of container technology?
- Briefly explain namespaces and control groups in container technology.
- What is an image in container technology? What are image registries?
- Show the components of the Docker engine using a diagram.
- Why is a layer-based approach adopted for container images?
- What are Docker volumes and Docker Compose?
- What are the limitations of Docker?
- What is Kubernetes?
- What are the roles of pods and kubectl in Kubernetes (K8s)?
- Explain the Kubernetes architecture with a suitable diagram.
- What is a YAML manifest in Kubernetes? What are some key fields in a YAML manifest?
- What is a ReplicaSet in Kubernetes?

Data Centers & Sustainable Computing

- What is a data center, and what are its purposes?
- State the standard dimensions (width, height, depth) for data centers.
- Explain the spine-leaf architecture in data centers. Why is traditional architecture insufficient?
- Describe top-of-rack (TOR) and end-of-rack (EOR) designs in data centers, listing their advantages and disadvantages.
- What is oversubscription in data center network design? What is switch oversubscription?
- Solve math exercises related to oversubscription.
- What are mice and elephant traffic in data centers?
- What is head-of-line blocking? How can this issue be mitigated?
- What is Power Usage Effectiveness (PUE)?

- What is sustainable computing?
- Explain embodied and operational carbon emissions in the context of cloud computing.
- Define carbon footprint and carbon intensity. What are their units?
- Solve math exercises related to PUE and carbon footprint.

Cloud Computing Services

- Is there a difference between elasticity and scalability in cloud computing? Explain.
- Discuss the service models in cloud computing (IaaS, PaaS, SaaS).
- What is serverless computing?
- Compare cloud deployment models (public, private, hybrid clouds).
- What are SLAs and SLOs?
- What are common performance metrics associated with SLAs?
- Solve math exercises related to SLAs.
- Compare monolithic and microservice software design approaches.
- What are the issues with the monolithic software design approach?
- Define RPC, client stub, and server stub.
- Explain the steps in RPC.
- What challenges may arise with the microservice software design approach?

Fault Tolerance

- What are the requirements for a dependable computing system?
- What are the types of faults that exist in computing?
- Describe failure models in computing.
- Illustrate triple modular redundancy for fault tolerance with a diagram.
- What are Byzantine faults in computing systems?
- Explain Lamport's recursive algorithm.
- What is Lamport's proposal on the minimum number of functional nodes required to achieve agreement in the presence of Byzantine faults?
- Explain the CAP theorem in distributed computing.

Big Data

- What is big data technology?
- Explain the three V's in big data.
- Define batch, stream, and parallel processing in big data, with example frameworks. What are their differences?
- Compare traditional database and data lake approaches for data storage and access.
- What is the difference between block-based and object-based storage?
- List the high-level steps to set up an HDFS cluster.
- Illustrate the overall architecture of HDFS with a diagram.
- What are the main components of the NameNode?
- Illustrate the overall workflow of HDFS with a diagram.
- Explain the process when a client wants to read a file from an HDFS system.

- Explain the process when a client wants to write a file into an HDFS system.
- Is S3 a file system? Justify your answer.

Edge Computing and Computing Applications

- What is the delay issue in cloud computing?
- Explain horizontal and vertical offloading with examples.
- What is Mobile Edge Computing (MEC)?
- What is a cloudlet?
- What are the benefits of web caching?
- Explain push-based and pull-based web caching.
- What are invalidate and update mechanisms in push-based web caching?
- What are the advantages and disadvantages of push-based and pull-based web caching?
- Explain how cloud computing can play a role in sensing, analyzing, and controlling to enhance the services related to following use cases:
    - Building
    - Transportation
    - Agriculture
- What are some cloud computing application areas in smart cities?
- Define crowd sensing. What are participatory sensing and opportunistic sensing?
- What are the requirements and challenges in crowd sensing?