End Semester Examination
Distributed Systems
Monsoon 2024

Instructions

Please read the following instructions carefully.

- The exam is for a duration of 180 minutes and is for a maximum of 90 points.
- The question paper has **FOUR** printed pages. Check your copy to see that you have all the questions printed properly.
- Fully state any assumptions you make while answering the questions.
- Answer all parts of a question contiguously.
- Too verbose answers may attract a penalty.

Section A: Short Answer Questions

Answer the following questions in brief. Usually, one point requries about two to three sentences of written answer. This section is for a total of 28 points.

1. Define the distributed consensus problem. What can be said about the outcome of the agreement problem when the source node is faulty? **(Points: 2+2=4) (COs: 4)**
2. Write the vector time for the events in the process time diagram shown in Figure 1. Use an increment of i for process P_i for $i = 1, 2, 3, 4$. Identify events that are logically concurrent. **(COs: 1)**
3. Describe the MPC model in brief. What aspects of cloud computing does the MPC model address in particular? **(Points: 2+2=4) (COs: 5)**
4. Explain what is meant by causal message delivery using both symbols and words along with a suitable example. **(Points: 2+2=4) (COs: 1)**
5. State and explain the CAP theorem in brief. Consider an LDAP service that runs for an organization across the globe. If it were implemented as a distribtued service, what aspects of the CAP theorem does it adhere to? Justifiy your answer. **(Points: 2+1+1=4)(COs: 6)**

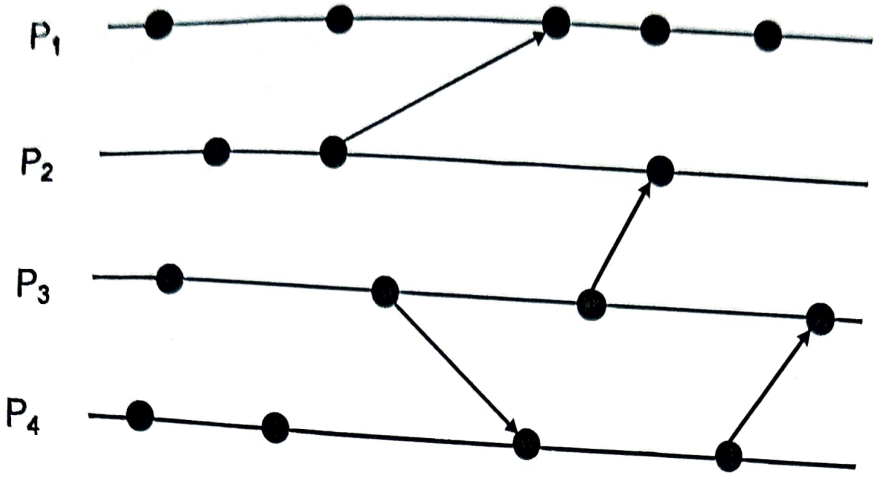

Figure 1: Figure for Question 2, Section A.

6. Explain the semantics that GFS ensures for the WriteAppend operation. Elaborate on aspects of these semantics are markedly different from a typical file system write operation? **(Points: 2+2=4) (COs: 3)**

7. Explain the functions MPI_REDUCE and MPI_ISEND briefly with the help of an example. **(Points: 2+2=4) (COs: 2)**

Section B: Long Answer Questions

1hr

Questions in this section require detailed answers. This section is for a total of 48 points.

1. Consider the problem of obtaining a maximal independent set. Answer the following questions.

   (a) Mention the characterization that the analysis of Luby's algorithm uses for classysyfing nodes as good and bad nodes, and edges as good and bad edges.

   (b) The analysis shows that in every iteration, at least a constant fraction of edges are good. Can such a guarantee be provided for good nodes? In other words, can we show that in every iteration, there are at least a constant fraction of good nodes.

   (c) How is the algorithm of Luby able to work in $O(\log n)$ rounds, even if there do not exist a constant fraction of nodes that are good?

   (d) Does the algorithm of Luby always identify an MIS of the largest size? Explain your answer.

   (Points: 2+2+2+2=8) (COs: 5)

2. Consider a system with seven nodes running the algorithm of Maekawa for mutual exclusion. Each node, P_i, for $i = 1$ to 7 chooses the set S_i as follows.

   $S_1 = \{1, 5, 7\}$ $S_2 = \{2, 1, 5\}$ $S_3 = \{3, 4, 1\}$
   $S_4 = \{4, 7, 5\}$ $S_5 = \{5, 3, 7\}$ $S_6 = \{6, 2, 4, 7\}$
   $S_7 = \{7, 4, 2\}$

   Answer the following questions.

   (a) What are the conditions needed on the sets S_i to ensure that sites gain access to the criticial section in a mutually exclusive manner.

(b) In the given example, if node P_i seeks a permission from the nodes in its set S_i to gain entry to the critical section, is there a possibility of a deadlock.

(c) How does the algorithm of Maekawa prevent such a possibility? Explain in the context of the given example.

(Points: 2+3+3=8) (COs: 4)

3. **Define the problem of distributed agreement. Show that agreement cannot always be reached among four processors if two processors are Byzantine faulty. Does the answer change if the initiator is not faulty? (Points: 2+5+1=8) (COs: 4)**

4. **What are some fundamental design consdierations for distributed file systems? Map NFS, GFS, and Haystack with respect to the above considerations. How does the support or the lack of support for shared write of files influence distributed file system design? (Points: 3+2+3=8) (COs: 3)**

5. Recall the 2-phase commit protocol in the context of distributed transaction processing. What kind of faults does the protocol tolerate? What are the assumptions made by the protocol? What would happen in the situation when the coordinator receives an "OK to Commit" from all participants, sends the "doCommit" message to all participants, but one of the participants fail to receive this message. What would happen in the situation when the coordinator fails to receive a reply to its "canCommit" message from one of the participants? Justify your answers. **(Points: 2+2+2+2=8) (COs: 4)**

6. State the properties of consistent hashing. How does the Chord overlay network make use of consistent hashing. Assume that in a Chord network with eight nodes, the hash values of nodes are 0.45, 0.22, 0.11, 0.58, 0.36, 0.88, 0.92, and 0.65. Find the fingers that the node with hash value 0.36 in the network. If node with hash value 0.88 leaves the network, mention the hash values of objects that have to be moved to a different bucket. Similarly, if a new node with a hash value of 0.75 joins the network, mention the hash values of objects that have to be moved to a different bucket.

(Points: 2+2+2+1+1=8) (COs: 3)

Section C: Design Question 1-5hrs.

In this section, there is one questions that spans multiple aspects of distributed system design. The system requirements are mentioned in brief and you are asked to suggest appropriate design chocies for the questions asked. This section is for a total of 14 points.

Case Study: Consider a system that a television news channel wants to develop for use in its coverage of Indian elections. The channel expects their reports, numbering in thousands and spread across the country, to post information on events such as election rallies, interviews with candidates, counting information as counting is in progress, statistics and surveys, and the like. The infornation can be as plain text, audio files, video files, and images. The anchors in the news studio consume this information and present it to viewers. We assume that the amount of information will be very large and the channel wants to also use the system across years with the data from past years available for use.

From a distributed sysems perspective, answer the following questions that this system design has to address.

1. What database model and semantics should the system designers choose?
2. Does the system need support for transactions? Does this simplify your database model? (4 points)
3. What guarantees do you want the system to provide with respect to the CAP theorem? Justify your answer.
4. How does the system best deal with data about the same phenomena but from across different times? For instance, the tally of votes that each candidate gets at a given place varies with time with all other parameters fixed.
5. Do you advocate using physical time or logical time for the various events in the system? Discuss your choice with respect to the above questions as well.
6. Briefly outline the overall system architecture and mention some advantages and disadvantages of your design.

(Points: 3+3+3+2+3=14) (COs: 1,3,4,6) 0.5 hrs