Distributed Systems

30th April, 2024

End Semester Examination

Q1. (6 marks) Distributed deadlock detection algorithms normally have substantial message overhead, even when there is no deadlock. Instead of using a deadlock detection algorithm, we could also handle deadlocks in the distributed systems simply by using "timeouts" i.e. after waiting for a certain time for a lock a process can timeout and release the locks to try again later. Explain the advantages and disadvantages of both methods: using deadlock detection algorithms and using timeouts.

Q2. (7 marks) In the Lamport Shoshtak Pease algorithm, mention what the vector pattern of a non malicious and malicious lieutenant respectively would be like in the case where the general is faulty assuming all lieutenants and general decides on 0's or 1's only. Explain why such a pattern is guaranteed to be formed. What would the vector pattern be like if the general is non faulty and why ?

Q3. GFS algorithm (3+3+5+3)

A. What's a lease in Google File system? Why is this required?
B. Explain the concept of "chunking" in GFS. How does the selection of chunk size impact the performance?
C. Explain the consistency model of GFS
D. Why is atomic record append at-least-once, rather than exactly once?

Q4 (3+5+4+3) Given is a weighted undirected graph $G=(V,E)$ with distinct weights. We need to compute the minimum spanning tree(MST). We would like to distribute the workload over two machines as follows. Split the edge set E into two equal sized halves $E=E1 \cup E2$ where $|E1|,|E2| \leq |E/2|$. On machine 1, we compute the MST $T1$ of $G1=(V,E1)$ and on machine 2, we compute the MST $T2$ of $G2=(V,E2)$. Our task is to compute the minimum spanning tree T for G using $T1$ and $T2$. Note: For a disconnected graph, a minimum spanning forest is composed of a minimum spanning tree for each connected component.

a. Suppose $e \in E1$ but e $\notin$ T1. Can e be part of T? justify your answer

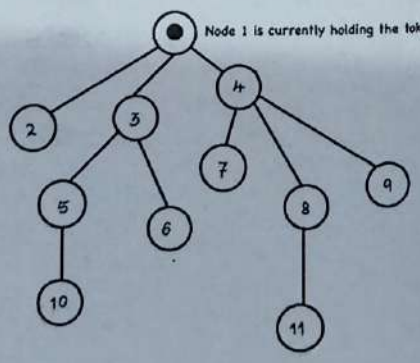
b. Describe an efficient algorithm to compute T using T1 and T2 and justify your correctness. Give its running time (you can ignore the time taken to compute T1 and T2)

c. How can a distributed MST algorithm be used for leader election? Leader election is when the nodes of the distributed system need to come up with one node as their leader.

d. In the GHS algorithm; the lower the level of the final Minimum Spanning Tree output; the more is the distance from the final core edge E_C to the node farthest from E_C. True or False? Explain.

Q5 Mutual Exclusion (5 + 5 + 5)

a. Consider the Raymond Algorithm. In the diagram, assume that node holds the token and is executing its critical section. Node 10, node 8, node 3, node 6 request for critical section at times 0, 2T, 3T, 4T time instant respectively, where T is the message delay from one node to its neighbour node. No further request for critical section is generated. Node 1 releases critical section at time 5T. Show how the token is passed to the requesting nodes with sequence of modified trees and message queue in nodes.



b. Do we require the FIFO property of a channel in both the Lamport's mutual exclusion scheme and in the Ricarta Agarawal's scheme? Explain why it is required or not required in each.

c. Consider a special case where in a system at most one message can get dropped in the running of the algorithm. Compare the Token Based Raymond's Algorithm, with the Message Based Ricart-Agarwala, in a scenario where at most one of the messages can get dropped.

Q6 (5+3+3 marks)

a) What are the benefits and drawbacks to the 3 phase protocol over the 2 phase protocol?

b) In the 2 phase commit protocol, what would happen if the Coordinator fails after sending everyone <prepare T> but before all other sites send <ready T>.

c) In the 3 phase commit protocol, what would happen if the coordinator failed after sending everyone <prepare T> but before all other sites send <ready T>.

Q7 (3+4+6)

a) Will consistent hashing work if my company has more than 360 servers ? Explain.

b) If the consistent hashing algorithm is implemented on three servers, namely A, B, and C, and server C experience a failure, then, it would also impact the keys that are stored on server A. Argue whether true or false.

c) Consider three events a,b,c in a distributed system and no two events belong to the same process. Using Lamport's definition of sequential and concurrent events, argue for each statement below whether it is true or false. Concurrent is shown as '||' and casual dependency by '→'

(i) $(a||b) \wedge (b \rightarrow c) \Rightarrow a \rightarrow c$

(ii) $(a||b) \wedge (b||c) \Rightarrow a||c$