Homework 2

1. **Why is it not possible to ensure 100% accuracy of the current real time using Network Time Protocol (NTP)? [0.25 marks]**
    - **a.** NTP cannot ensure 100% accuracy as it relies on the assumption that the network delay from the client to server is the same as the network delay from server to client. This assumption may not hold in practice. NTP is known to have errors in the range of 1-50 msec.

2. **Mention the two situations in which the "happens-before" relation holds true for two events *a* and *b* in a distributed system. [0.25 marks]**
    a. If a and b are events in the same process, and a occurs before b, then a can be said to **happen-before** b.
    b. If a is the event of a message being sent by one process, and b is the event of the message being received by another process, then a can be said to **happen-before** b. A message cannot be received before it is sent, or even at the same time it is sent, since it takes a finite, nonzero amount of time to arrive.
3. **Describe Lamport's algorithm for logical clocks in 3-4 steps. [0.25 marks]**
    a. To implement Lamport's logical clocks, each process Pi maintains a local counter C_i. These counters are updated according to the following steps:
    b. Before executing an event (i.e., sending a message over the network, delivering a message to an application, or some other internal event), P_i increments as $C_i = C_i + 1$.
    c. When process P_i sends a message m to process P_j, it sets m's timestamp ts(m) equal to C_i after having executed the previous step.
    d. Upon the receipt of a message m, process P_j adjusts its own local counter as $C_j = \max\{C_j, ts(m)\}$ after which it then executes the first step and delivers the message to the application.
4. **What is the limitation of logical clocks (Lamport's) which is overcome by vector clocks? [0.25 marks]**
    a. Lamport clocks do not capture causality. That is, just by looking at the Lamport timestamps of two events A and B, we cannot conclude anything about their causal relationship. But just by looking at the vector clocks between A and B, we can comment on the possibility of causal relationship between A and B.

5. **Describe how each process maintains/updates its vector clock in 3-4 steps. [0.25 marks]**

1. Before executing an event (i.e., sending a message over the network, delivering a message to an application, or some other internal event), P_i executes $VC_i[i] \leftarrow VC_i[i] + 1$. This is equivalent to recording a new event that happened at P_i.
2. When process P_i sends a message m to P_j, it sets m's (vector) timestamp $ts(m)$ equal to VC_i after having executed the previous step (i.e., it also records the sending of the message as an event that takes place at P_i).
3. Upon the receipt of a message m, process P_j adjusts its own vector by setting $VC_j[k] \leftarrow \max\{VC_j[k], ts(m)[k]\}$ for each k (which is equivalent to merging causal histories), after which it executes the first step (recording the receipt of the message) and then delivers the message to the application.

6. **Mention one advantage and disadvantage of the centralized algorithm for ensuring mutual exclusion. [0.25 marks]**
    a. **Advantage:** It is easy to implement, too, and requires only three messages per use of resource (request, grant, release).
    b. **Disadvantage:** The coordinator is a single point of failure, so if it crashes, the entire system may go down.
7. **Mention two disadvantages of the distributed algorithm for ensuring mutual exclusion. [0.25 marks]**
    a. This algorithm has N points of failure. If any process crashes, it will fail to respond to requests. This silence will be interpreted (incorrectly) as denial of permission, thus blocking all subsequent attempts by all processes to enter any of their respective critical regions.
    b. If the total number of processes is N, then the number of messages that a process needs to send and receive before it can enter its critical region is 2 * (N −1). This is much more than the number of messages required by the centralized algorithm.
    c. Another problem with this algorithm is that either a multicast communication primitive must be used, or each process must maintain the group membership list itself, including processes entering the group, leaving the group, and crashing.
8. **What kind of data structure is used by Apache Zookeeper to facilitate different coordination services? [0.25 marks]**
    a. In-memory data structure organized as a tree.