CS 425 Distributed Systems, Fall 2010, University of Illinois at Urbana-Champaign

Homework 5: Solution

1. Problem 15.13.

   Solution:

   The only replica manager that can satisfy a query from this front end is the third, with (value) timestamp $(4, 5, 8)$. The others have not yet processed at least one update seen by the front end. The resultant time stamp of the front end will be $(4, 5, 8)$.

   Similarly, only the third replica manager could incorporate an update from the front-end immediately.

2. In Dijkstra's self-stabilizing mutual exclusion protocol for a unidirectional ring, show that if $k = 3, n = 4$, then the system does not satisfy convergence. Give a concrete counter-example.

   Solution:

   A counter-example is shown in Figure 1. Note that in an illegal configuration multiple nodes can hold the lock. So, they can change their states simultaneously.

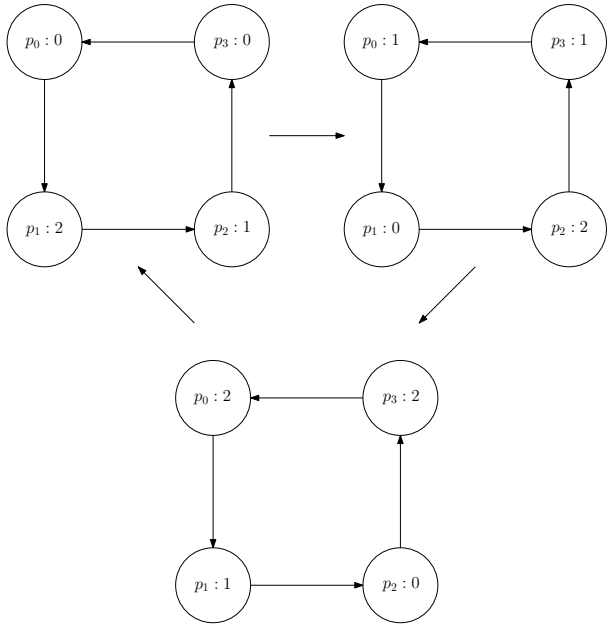

Figure 1: A run of Dijkstra's self-stabilizing mutual exclusion with $k = 3$, $n = 4$

3. A sensor network needs to periodically calculate the average of all temperature readings across the sensor node population. There are two options to do this: (1) all sensors send their value to the far-away base station, which then calculates the average; (2) the sensors run an distributed protocol that aggregates data amongst the sensors themselves and then reports the final average to the base station. Which of these two approaches would you prefer, and what is the most important reason for your choice?

   Solution:

   If the far-away base station is within the radio range of all the sensors, then the sensors can send their values directly to the base station in approach (1). In approach (2) sensors

can talk to their neighboring nodes to aggregate information. Therefore, the communication distance is significantly reduced in (2). This increases the battery-life of the sensors. So, approach (2) is preferable in this scenario.

In case the far-away base station is beyond the radio range of most of the sensors, the sensors will have to use multiple hops to reach the destination in approach (1). This suggests that the sensors closer to the base station will have to forward a large number of other sensors' data. On the other hand, in approach (2), the number of messages is much smaller since each sensor aggregates the data before reporting it to the neighboring nodes. So, approach (2) is preferable in this scenario as well.

Note that averaging is a low-overhead operation and carrying out this operation within the sensor nodes should not consume much power.

4. Web site popularity is well-known to follow a Zipf distribution. Someone measuring top-10 traffic to Websites gives you the following number of monthly visits (in millions) for their top 10 list:
1. 375 M; 2. 200 M; 3. 140 M; 4. 110 M; 5. 92 M; 6. 80 M; 7. 75 M; 8. 70.5 M; 9. 68 M; 10. 63 M; Would you say this data is valid, i.e., does it (approximately) satisfy the Zipf distribution?

   Solution:

   Figure 2 shows that the data approximately satisfy the Zipf distribution.

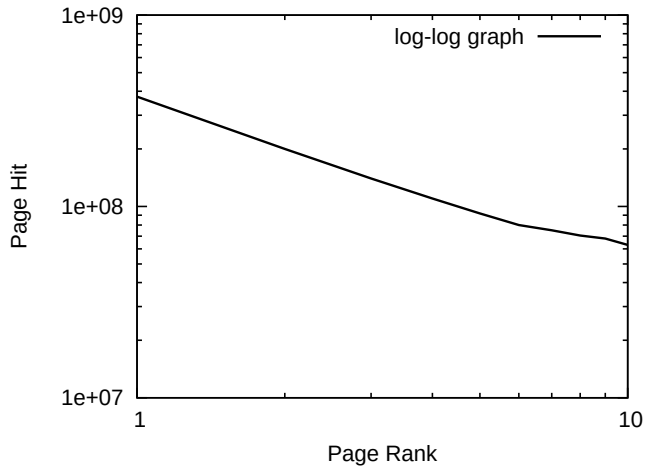

Figure 2: Plot

5. You are working in a company which stores credit card records for many individuals. You are asked to write down the basic confidentiality, integrity, and availability properties for this service. Write down one sentence for each. You don't need to be comprehensive - just cover the most important concern in each.

   Solution:

   Confidentiality: credit card records should not be disclosed to unauthorized parties.
   Integrity: credit card records should not be modified inappropriately by unauthorized parties.
   Availability: credit card records should be available for access at all times by authorized parties.

6. A file system you are implementing will have files whose access permissions change very rapidly and frequently. Would you prefer ACLs or capabilities for this file system (Access Control Matrix not allowed)?

   Solution:

   ACLs are preferred since a capability-based system would require iteration over all the users' capabilities for the access permission change of a single file. In contrast, in an ACL-based system, for each access permission change, only the ACL of the associated file needs to be modified. Also, access permission revocation is hard with capabilities and rapidly changing permissions make it even harder.

7. (a) Consider a PBFT system with 2 faulty nodes. What is the minimum number of non-faulty nodes required in the system to support Byzantine fault tolerance? (b) In this system, a client issues a request to the primary and the primary starts the three phase protocol (you can assume that the primary is non-faulty). Calculate the total number of messages exchanged by the non-faulty replicas until all of them reply to the client. It is given that the system only supports unicast messages.

   Solution:

   (a) Minimum number of non-faulty nodes = $2f + 1 = 5$

   (b) Pre-prepare phase messages (primary to all the replicas) = 6
   Prepare phase messages (Non-faulty replicas, except the primary, to all other replicas) = $4 \times 6 = 24$
   Commit phase messages (All non-faulty replicas to all other replicas) = $5 \times 6 = 30$
   Total messages = $6 + 24 + 30 = 60$