HW 8

1. Two Phase Commit:

Explain the action taken in the two phase commit protocol after restart if the coordinator or subordinate nodes fail.

There are two parts to answering this question. The first section will answer what happens when the one of the subordinate nodes fails. The second part will answer what is done if one of the coordinator node fails.

If a subordinate node fails then there are several steps we may need to take for T depending on what is in the log. If we have an end record for T then we don’t need to do anything. However, if we have a log record for T but we don’t have an end record than we will need to undo or redo T. If we have a prepare log for T but not a commit or abort we can say that this node is a subordinate. We will then contact the coordinator to enquire about the status of T and will then accordingly redo or undo T. Additionally, if we have no prepare log for T then we need to abort T.

If we are the coordinator and we have a log for T and we have an end log for T we don’t need to worry about T because T was completed. However, if we do not have an end log for T then we need to keep sending commit/abort messages to subordinates until the acknowledge the request. Again if we have no prepare log for T we need to abort T along with sending out the abort message to all subordinates for T.

1. Multi-Version concurrency control:

Consider transactions T1, T2 and T3 with timestamps 10, 20 and 30 respectively in a database system that follows multi-version concurrency control protocol: T1 is a reader and T2 and T3 are writer transactions.

* 1. Let T1 read and T2 read and update data item D. if transaction T# updates only D, provide a schedule where the multi-version concurrency control method will not restart T3.
  2. Again, assume that T1 reads and T2 and T3 read and update only data item D provide a schedule where the muli-Version concurrency control method will restart T2.

1. Quorum Reading:

Let the number of copies of a data item in a cluster, N, be 10 and the number of nodes that participate in a successful write, W, be 5.

* 1. What is minimum number of nodes that should participate in a successful read, R?
  2. Now, consider a sloppy quorum with the same values for N and W. What is the minimum number of nodes that should participate in a successful read R?

1. Eventual consistency and vector clocks:

Each row in the following table shows vector clocks of different copies of the same data on a system with three nodes SX, SY and SZ. Explain if the copies in each row have a conflict.

1. Page Rank:

Assume the Following graph depict a part of the Web, where nodes represent pages and edges show hyper-links. Find out the pages whose PageRank values are greater than-zero and their relative PageRank values in the graph. You do not need to preform the fix point computation to determine the PageRank values. Instead, you should guess the PageRank values based on your understanding of the PageRank algorithm and explain why you think they are correct. If it is not possible to make any educated guess for some page(s), you should explain why.

1. MongDB
   1. Write a Query that returns cuisine type of the restaurant name The Dead Rabbit.
   2. Write a query to create an index on the name attribute of the restaurants.
   3. Write a query that uses the index created in part b to return all the restaurants containing the term rabbit in their name.
   4. Write an aggregate query to show total count of restaurants in each borough.