Distributed System
Project Proposal

By: Sravani Namburi, Shriya Jain

Architecture:

Our architecture will follow the techniques of the Dynamo System.

- Replication of data will be done to achieve high availability. Each write will go to a random node. This node will be responsible for replication data to the rest of the nodes asynchronously. This means the write request will be returned to the caller before it is replicated to all the nodes. The write conflicts will be resolved using Data versioning and Vector clocks.

- Most of the time for any data, the new version will subsume the old version but in case of concurrent requests, all available versions of the data will be returned. The resolution takes place in the application layer in such cases to collapse multiple versions into a single version.

- The read request can go to any node. As mentioned above if the data is unresolved due to concurrent writes or failures, all versions available of the data will be returned.

- A quorum configuration of R + W < N will be followed, where R is the minimum number of nodes that must participate in a successful read operation. W is the minimum number of nodes that must participate in a successful write operation and N is the total number of Nodes. Initial configuration is R=1,W=1,N=3.

- Anti Entropy will be implemented using Merkle Trees to detect inconsistencies between replicas.

- The membership information among the nodes will be propagated using the gossip-based protocol.

Metrics:

Measuring the Consistency using the probability of reading a write Δ seconds after it returns in these three cases:

1) Issuing concurrent writes and measuring the consistency and read behaviour.
2) Introducing delays between replication of data among Nodes and measuring consistency.
3) Increasing the number of nodes while maintaining constant quorum configuration.

Testing Scenarios:

1) Multiple clients will be used to invoke read and write requests.
2) Multiple read and write requests (in ratio 10:1, to model a typical read-heavy application) to analyse the behaviour (in terms of consistency of data returned).
3) Introduce Read and Write (among replica nodes) delays and analyse consistency.
4) Changing the total number of nodes and measuring consistency.
5) Change values of R and W and measure impact on consistency.