

Proposal Project Distributed Systems:

Replication System

Student: Samuel Esposito - s1597183

For the Distributed Systems project I would like to build a replication system that synchronizes data between different nodes of a key-value based web oriented data store. Data nodes can be added and removed dynamically to the system. A minimal spanning tree of the available nodes is calculated at regular intervals in order to optimize the synchronization communication between the set of nodes at a specific point in time. Leader election will be used to dynamically define one or multiple master nodes, depending on the consistency constraints at a specific point in time. If a write requires storing the data on at least two nodes, two master nodes will be elected to which the data is routed for storage. If strong consistency is required, the data will always be written to a set of synchronous master nodes. If only eventual consistency is required, the data can be replicated asynchronously. In the occasion of conflicts during replication the system will use a voting system to decide which of the conflicting pieces of data should remain.