31/10

Paxos

-> Paxos, is the algorithm that is used to achieve consensus among a distributed set of computers that communicate via an asynchronous network.

↳ One (or) more clients proposes a value to Paxos and we have consensus when a majority of system running Paxos agrees on one of the proposed value.

↳ Paxos is widely used and is legendary in computer Science since it is the first consensus algorithm that has been rigorously proved to be correct

↳ Paxos simply selects a single value from one (or) more values that are proposed to it and lets everyone know what the value is.

↳ A run of the Paxos protocol results in the selection of single proposed value.

↳ If you need to use paxos to create a replicated log, then you need to run paxos repeatedly. This is called multi-Paxos.

↳ Paxos provide abortable consensus. This means that some proposes abort the consensus if there is contention while other decide on the value

↳ Those processes that decide have to agree on the same value. Aborting allows a process to terminate rather than be blocked indefinitely.

Assumptions for the algorithm

↳ concurrent proposal

↳ validity

↳ Majority rule

↳ Asynchronous n/w

↳ Fail-stop faults

↳ Unicasts

↳ Announcements

RAFT ALGORITHM

↳ This is designed to be easy to understand.

↳ It is equivalent to Paxos in fault-tolerance and performance.

↳ Raft state that each node in a replicate state machine

can stay in any of the 3 states, namely,

↳ leader

↳ candidate

↳ follower

*) Under normal condition, a node can stay in any one of the states

↳ Only a leader can interact with the client, any request to the follower node is redirected to the leader node.

↳ A candidate can ask votes to become a leader.

A follower only responds to candidate(s) or the leader.

↳ To maintain server status, the raft algorithm divides time into small terms of arbitary length. Each term is identified by a monotonically increasing number, called term number.

x——x.