10. Gossip Protocol

Let's learn about gossip protocol and its usage.

We'll cover the following

- Background
- Definition
- Solution
- Examples

Background#

In a large distributed environment where we do not have any central node that keeps track of all nodes to know if a node is down or not, how does a node know every other node's current state? The simplest way to do this is to have every node maintain a heartbeat with every other node. Then, when a node goes down, it will stop sending out heartbeats, and everyone else will find out immediately. But, this means $O(N^2)$ messages get sent every tick (N being the total number of nodes), which is a ridiculously high amount and will consume a lot of network bandwidth, and thus, not feasible in any sizable cluster. So, is there any other option for monitoring the state of the cluster?

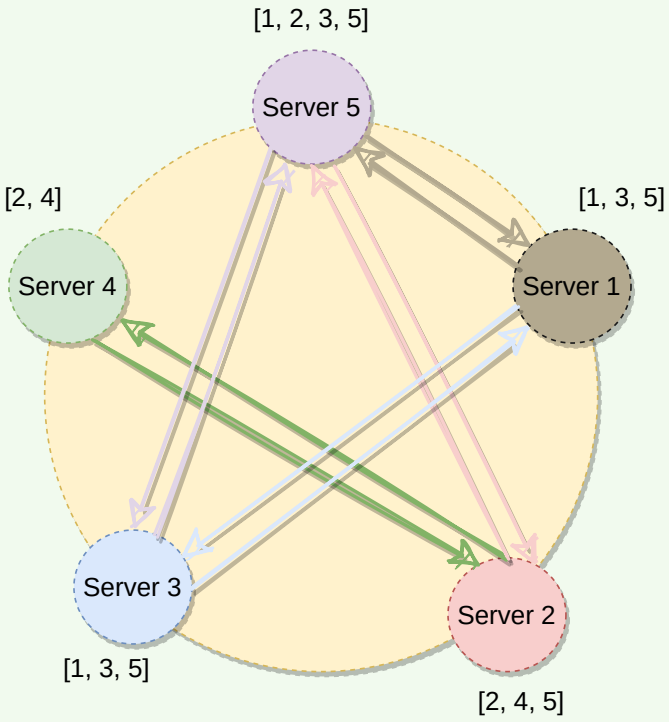
Definition#

Each node keeps track of state information about other nodes in the cluster and gossip (i.e., share) this information to one other random node every second. This way, eventually, each node gets to know about the state of every other node in the cluster.

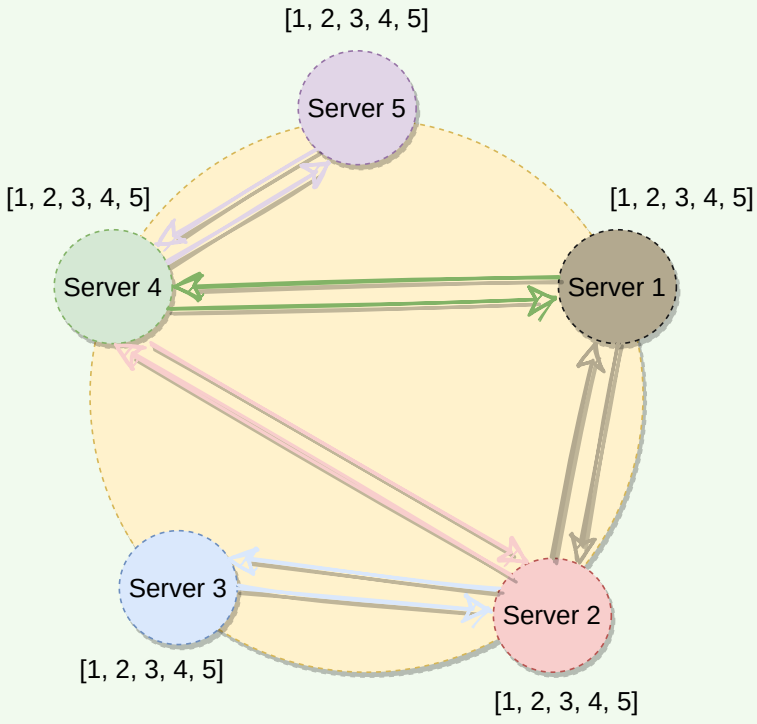
Solution#

Gossip protocol is a peer-to-peer communication mechanism in which nodes periodically exchange state information about themselves and about other nodes they know about. Each node initiates a gossip round every second to exchange state information about themselves and other nodes with one other random node. This means that any state change will eventually propagate through the system, and all nodes quickly learn about all other nodes in a cluster.

Every second each server exchanges information with one randomly selected server



Every second each server exchanges information about all the servers it knows about



Examples#

Dynamo & Cassandra use gossip protocol which allows each node to keep track of state information about the other nodes in the cluster, like which nodes are reachable, what key ranges they are responsible for, etc.

← Back

Next →

9. Heartbeat

11. Phi Accrual Failure Detection

Mark as Completed

Report an Issue