cluster-forget

- Details on why the ban-list is needed
- Special conditions not allowing the command execution

The command is used in order to remove a node, specified via its node ID, from the set of *known nodes* of the Valkey Cluster node receiving the command. In other words the specified node is removed from the *nodes table* of the node receiving the command.

Because when a given node is part of the cluster, all the other nodes participating in the cluster knows about it, in order for a node to be completely removed from a cluster, the `CLUSTER FORGET` command must be sent to all the remaining nodes, regardless of the fact they are primaries or replicas.

However the command cannot simply drop the node from the internal node table of the node receiving the command, it also implements a ban-list, not allowing the same node to be added again as a side effect of processing the *gossip section* of the heartbeat packets received from other nodes.

Details on why the ban-list is needed

In the following example we'll show why the command must not just remove a given node from the nodes table, but also prevent it for being re-inserted again for some time.

Let's assume we have four nodes, A, B, C and D. In order to end with just a three nodes cluster A, B, C we may follow these steps:

1. Reshard all the hash slots from D to nodes A, B, C.
2. D is now empty, but still listed in the nodes table of A, B and C.
3. We contact A, and send `CLUSTER FORGET D`.
4. B sends node A a heartbeat packet, where node D is listed.
5. A does no longer known node D (see step 3), so it starts a handshake with D.
6. D ends re-added in the nodes table of A.

As you can see in this way removing a node is fragile, we need to send `CLUSTER FORGET` commands to all the nodes ASAP hoping there are no gossip sections processing in the meantime. Because of this problem the command implements a ban-list with an expire time for each entry.

So what the command really does is:

1. The specified node gets removed from the nodes table.
2. The node ID of the removed node gets added to the ban-list, for 1 minute.
3. The node will skip all the node IDs listed in the ban-list when processing gossip sections received in heartbeat packets from other nodes.

This way we have a 60 second window to inform all the nodes in the cluster that we want to remove a node.

Special conditions not allowing the command execution

The command does not succeed and returns an error in the following cases:

1. The specified node ID is not found in the nodes table.
2. The node receiving the command is a replica, and the specified node ID identifies its current primary.
3. The node ID identifies the same node we are sending the command to.

□□□ #1
□ □□□ □□□ 2026-10-31 15:11:49 UTC
□ □□□ □□□ 2026-10-31 15:11:49 UTC