zookeeper service

Made up of individual servers

Represents the ensemble

Coordination services

Mechanisms enabling distribution of processing

Such mechanisms include synchronization of individual node processing such that nodes operate according to a pre-defined coordination specification.

Nodes coordinate via event notification

znode

the primary data structure in zookeeper. May contain both data and child nodes

nodes are always referred to absolutely, never by relative reference.

Created by application clients

Like a file that persists in memory on zookeeper servers

Can be updated by any node in the cluster

Any node in the cluster can register to be notified of changes in a znode

Used by applications to store the state of their processing

Any change in one znode is notified to other servers in the cluster who are watching each other’s copies

A data node

The main entity with which a programmer interacts.

Parent znode

The parent of a particular znode

Each znode may only have a single parent

Zookeeper quorum

All the zookeeper servers which make up a zookeeper service

Same as a zookeeper ensemble

Quorum peer

A zookeeper server

leader election

Zookeeper server

Machine which keeps a copy of the entire state of the system and stores on log files

A member of the zookeeper quorum

zookeeper tree

a structure of znodes

represents the data managed by zookeeper

znode data

TODO

zookeeper cache

TODO

stat structure

Properties of a znode

Each znode contains a stat structure containing properties that support various zookeeper functionality, like

Properties include

Ephemeral node

A znode which may be created as part of a client session.

Ephemeral nodes may not have any children

Ephemeral nodes are deleted when the session is destroyed.

Ephemeral nodes are visible to other severs / clients

Sequential node

Node which is named using a sequence unique to the znode’s parent

Node has %010d appended to it’s name

Ensemble - a cluster of zookeeper nodes (znodes)

Represents the high availibility / fault-tolerant aspects of the system

Each node in an ensemble knows about every other node

Strict ordering

Refers to updates in the system being ordered.

Watch -

A listener that can be registered by a client interested in being notified of changes to a particular znode.

Only last for a single detected change

Replica -

data snapshot -

myid file -

distributed filesystem -

hierarchical namespace -