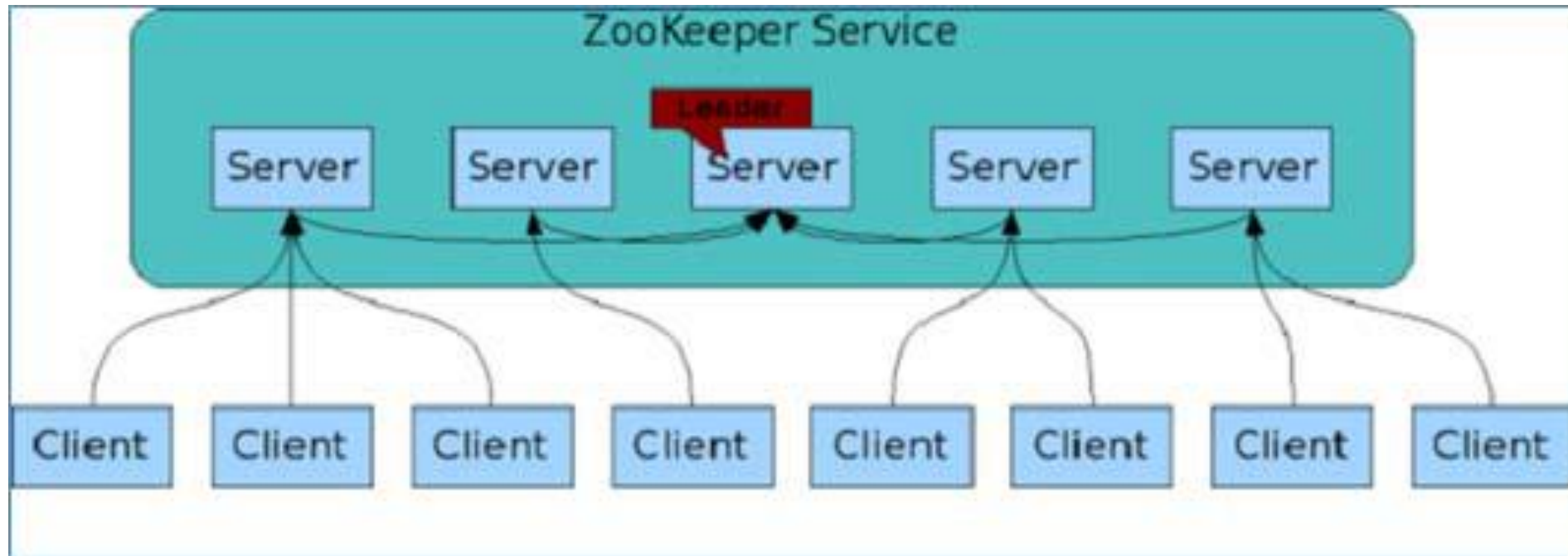


Apache Zookeeper

A top-down view of zookeeper service

- Apache ZooKeeper is a coordination service for distributed applications.
- It aims to solve the tough problems associated with the coordination of components in a distributed application

A top-down view of zookeeper service



A ZooKeeper service and how clients connect to the service

A top-down view of zookeeper service

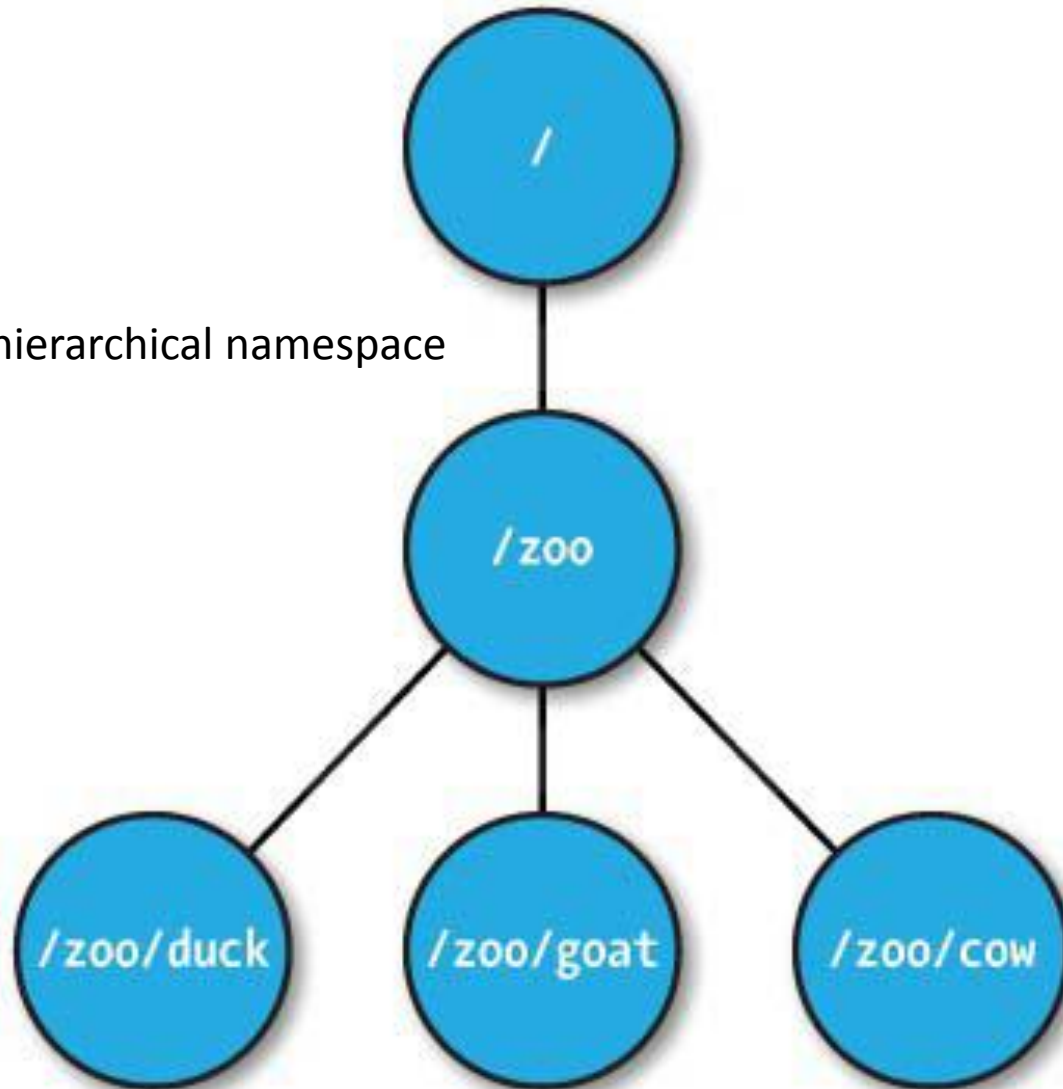
- You can see the replicated set of servers
- Zookeeper service runs on these servers
- This is called an ensemble.
- Clients can connect to a zookeeper service by connecting to any member of the ensemble.

Zookeeper Data Model

- ZooKeeper allows distributed processes to coordinate with each other through a shared hierarchical namespace of data registers
- The namespace looks quite similar to a Unix filesystem.
- The data registers are known as **znodes in the ZooKeeper nomenclature.**

Zookeeper Data Model

A ZooKeeper's hierarchical namespace



Zookeeper Data Model

- znodes are organized hierarchically, much like a tree, as a standard filesystem.
- The root node has one child znode called **/zoo**, **which in turn has three znodes.**
- Every znode in the ZooKeeper tree is identified by a path, and the path elements are separated by **/**.
- The znodes are called data registers because they can store data. Thus, a znode can have children as well as data associated with it. It's analogous to having a filesystem that allows a file to also be a path.

Zookeeper Data Model

- The data in a znode is typically stored in a byte format
- A maximum data size in each znode is not more than 1 MB.

Types of Znodes

- ZooKeeper has two types of znodes:
 - persistent and ephemeral.
- There is a third type that you might have heard of, called a sequential znode, which is a kind of a qualifier for the other two types.
- Both persistent and ephemeral znodes can be sequential znodes as well.
- Note that a znode's type is set at its creation time.

Types of Znodes - Persistent

- Persistent znodes have a lifetime in the ZooKeeper's namespace until they're explicitly deleted
- A znode can be deleted by calling the **delete** API call

Types of Znodes - Ephemeral

- An ephemeral znode is deleted by the ZooKeeper service when the creating client's session ends
- An end to a client's session can happen because of disconnection due to a client crash or explicit termination of the connection.
- Even though ephemeral nodes are tied to a client session, they are visible to all clients, depending on the configured **Access Control List (ACL) policy**.

Types of Znodes - Ephemeral

- An ephemeral znode can also be explicitly deleted by the creator client or any other authorized client by using the **delete** API call.
- An ephemeral znode ceases to exist once its creator client's session with the ZooKeeper service ends.
- Hence, in the current version of ZooKeeper, ephemeral znodes are not allowed to have children.

Types of Znodes - Ephemeral

- The concept of ephemeral znodes can be used to build distributed applications where the components need to know the state of the other constituent components or resources.
- For example, a distributed group membership service can be implemented by using ephemeral znodes.
- The property of ephemeral nodes getting deleted when the creator client's session ends can be used as an analogue of a node that is joining or leaving a distributed cluster.
- Using the membership service, any node is able to discover the members of the group at any particular time.

Types of Znodes – Sequential

- A sequential znode is assigned a sequence number by ZooKeeper as a part of its name during its creation.
- The value of a monotonously increasing counter (maintained by the parent znode) is appended to the name of the znode.

Types of Znodes - Sequential

- Sequential znodes can be used for the implementation of a distributed global queue, as sequence numbers can impose a global ordering.
- They may also be used to design a lock service for a distributed application.

Types of Znodes

Since both persistent and ephemeral znodes can be sequential znodes, we have a total of four modes of znodes:

- persistent
- ephemeral
- persistent_sequential
- ephemeral_sequential