INTRODUCTION TO ZOOKEEPER

Apache Zookeeper ™

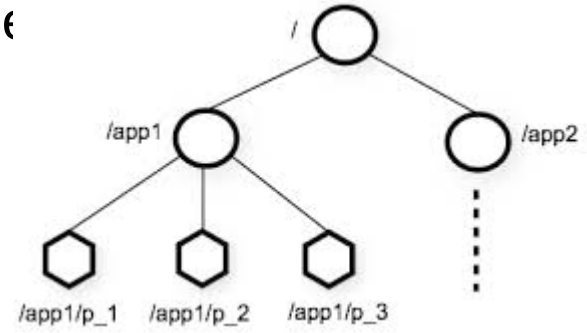
What is a Distributed System

"A **Distributed system** consists of multiple computers that communicate and coordinate their actions by passing messages. The components interact with each other in order to achieve a common goal."

- **Wikipedia**

What is Zookeeper

- An Open source, High Performance coordination service for distributed applications
- Centralized service for
  - Configuration Management
  - Locks and Synchronization for providing coordination between distributed syste
  - Naming service (Registry)
  - Group Membership
- Features
  - hierarchical namespace
  - provides watcher on a znode
  - allows to form a cluster of nodes
- Supports a large volume of request for data retrieval and update
- http://zookeeper.apache.org/



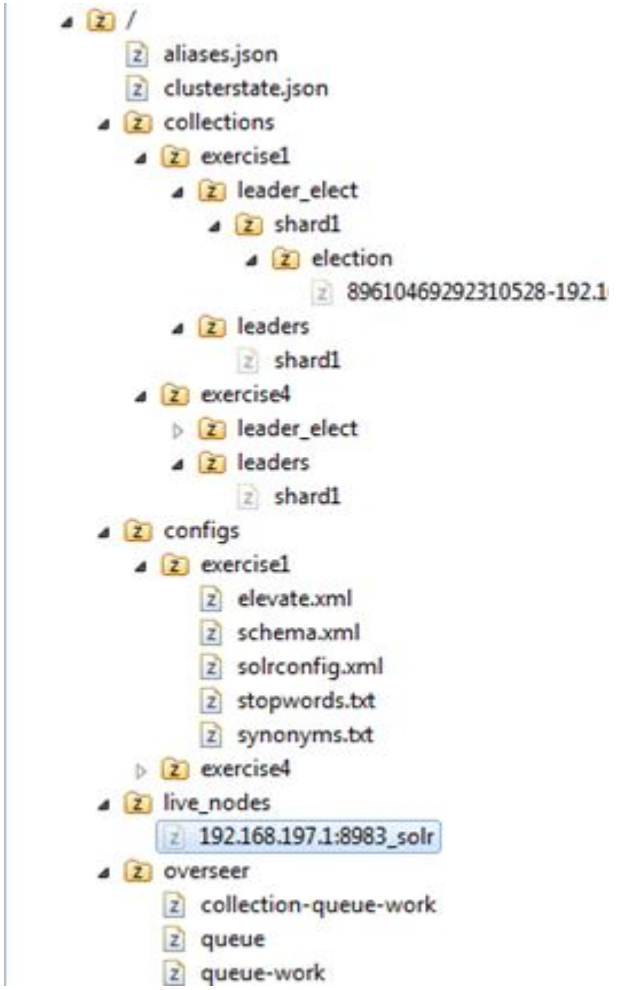
Source : http://zookeeper.apache.org

Zookeeper Use cases

- Configuration Management
  - Cluster member nodes Bootstrapping configuration from a central source
- Distributed Cluster Management
  - Node Join/Leave
  - Node Status in real time
- Naming Service – e.g. DNS
- Distributed Synchronization – locks, barriers
- Leader election
- Centralized and Highly reliable Registry

Zookeeper Data Model

- Hierarchical Namespace
- Each node is called “znode”
- Each znode has data(stores data in byte[] array) and can have children
- znode
  - Maintains “Stat” structure with version of data changes , ACL changes and timestamp
  - Version number increases with each changes



Thank You