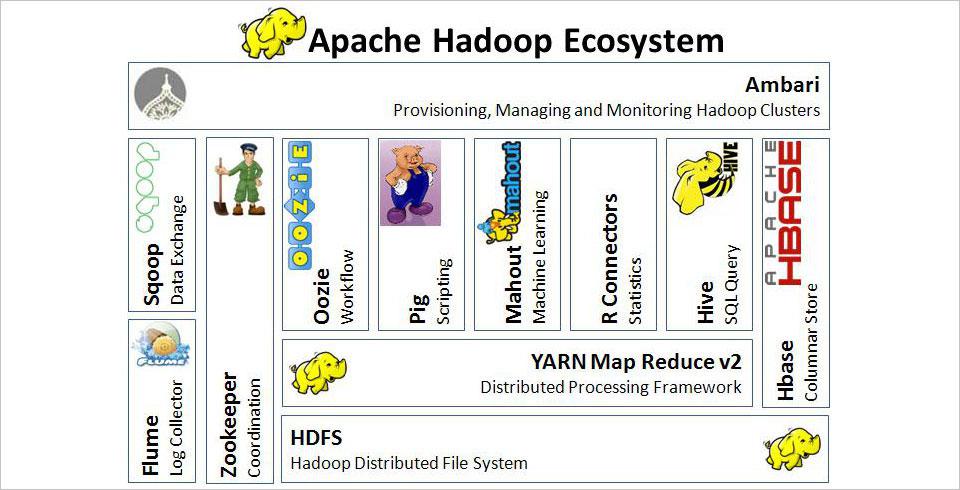
Round2Day13/14 -**Zookeeper**

1. Apache ZooKeeper is a software project of the Apache Software Foundation. It is essentially a service for distributed systems offering a hierarchical key-value store, which is used to provide a distributed configuration service, synchronization service, and naming registry for large distributed systems. ZooKeeper was a subproject of Hadoop but is now a top-level Apache project in its own right
2. In other words, an effort to develop and maintain an open-source server which enables highly reliable distributed coordination
3. Stable release: ZooKeeper Repository
4. Language: Java
5. Operating system: Cross-platform
6. Type: Distributed computing
7. 
8. Originally developed at Yahoo! For streamlining the processes running on big-data clusters by storing the status in local log files on the ZooKeeper servers
9. Prime features of Apache ZooKeeper are:
   1. Reliable system – this system is very reliable as it keeps working even if a node fails
   2. Simple architecture – the architecture of ZooKeeper is quite simple as there is a shared hierarchical namespace which helps coordinating the processes
   3. Fast processing – ZooKeeper is especially fast in ‘read-dominant’ workloads (i.e. workloads in which reads are much more common than writes)
   4. Scalable – the performance of ZooKeeper can be improved by adding nodes
10. Architecture
    1. Node – the systems installed on the cluster
    2. ZNode – the nodes where the status is updated by other nodes in cluster
    3. Client Applications - the tools that interact with the distributed applications
    4. Server Applications – allows the client applications to interact using a common interface
11. Use cases
    1. **Naming service** (directory service) – maps the names of network resources to their respective network addresses
    2. **Configuration management –** Systems engineering process for establishing and maintaining consistency of a product’s performance, functional, and physical attributes with its requirements, design, and operational information throughout its life
    3. **Data Synchronization** - refers to one of two distinct but related concepts: synchronization of processes, and synchronization of data.
       1. Process synchronization refers to the idea that multiple processes are to join up or handshake at a certain point, in order to reach an agreement or commit to a certain sequence of action
       2. Data synchronization refers to the idea of keeping multiple copies of a dataset in coherence with one another, or to maintain data integrity. Process synchronization primitives are commonly used to implement data synchronization
    4. **Leader election**  - the process of designating a single process as the organizer of some task distributed among several computers (nodes). Before the task is begun, all network nodes are either unaware which node will serve as the leader (or coordinator) of the task, or unable to communicate with the current coordinator. After a leader election algorithm has been run, however, each node throughout the network recognizes a particular, unique node as the task leader
    5. **Message queue**  - (Mailboxes) are software-engineering components typically used for inter-process communication (IPC), or for inter-thread communication within the same process. The use a queue for messaging – the passing of control or of content.
    6. **Notification system** – a combination of software and hardware that provides a means of delivering a message to a set of recipients. It commonly shows activity related to an account. Such systems constitute an important aspect of modern web application
12. Referred resources
    1. <https://en.m.wikipedia.org/wiki/Apache_ZooKeeper>
    2. <https://zookeeper.apache.org/>
    3. https://www.tutorialspoint.com/zookeeper/zookeeper\_overview.htm

