Hadoop Documentation

Level-3

1. NoSQL – Hbase – Architecture
   1. Data storage
   2. Write –Ahead Log
   3. Read path
   4. Region Lookups
   5. The Region Life cycle
   6. Zookeeper
   7. Replication
2. Hbase – Client APIs

2.1 CRUD operations –PUT,GET and DELETE

2.1.1 Batch operations

2.1.2 Row Locks

2.2 Scan operation

2.3 Miscellaneous Features (HBASE Java API’s)

1. HBase Clients

3.1 Interactive Clients – REST, Thrift, Avro and others

3.2 Batch Clients

3.3 HBase Shell

3.4 Web – based UI

1. Hbase – MapReduce Integration

4.1 MapReduce Overview

4.2 MapReduce Over HBase

YARN

1. Yarn – Core Concepts
   1. Beyond MapReduce
   2. Apache Hadoop Mapreduce
   3. Apache Hadoop YARN Components
2. Functional Overview of Yarn Components
   1. Architecture Overview
   2. YARN Schedulling Components
   3. Containers
   4. NodeManager
   5. Application Master
   6. YARN Resource Model
   7. Managing Application Dependencies
3. YARN Architecture
   1. Resource Manager
   2. Node Manager
   3. Application Master
   4. YARN Containers
4. Capacity Scheduler In YARN
   1. Introduction
   2. Capacity Scheduler Configuration
   3. Queue & Hierachical Queues
   4. Queue Access control
   5. Capacity Management with Queues
   6. User Limits & Reservations
   7. State of the Queues
   8. Limits on Applications
   9. User Interface
5. MR With Yarn
   1. Running Hadoop YARN MapReduce Examples
   2. MapReduce Compatibility
   3. MapReduce Application Master
   4. Calulating The Capacity of a node
   5. Changes to the Shuffle Service
   6. Running Existing Hadoop Version 1 Applications
   7. Running MapReduce version 1 Existing Code
   8. Advanced Features
6. Oozie – Automating data processing
   1. Oozie workflow
   2. Oozie Coordinator
   3. Oozie bundle
   4. Oozi parameterization with Expression Language
   5. Job Execution Model
   6. Different ways for accessing the Oozie server
7. Using Oozie
   1. Designing an Oozie application
   2. Implementing Oozie Workflows
   3. Implementing Oozie Coordinator applications
   4. Implementing an Oozie Bundle
   5. Understanding how to deploy, test, and execute Oozie applications
8. Advanced Oozie features
   1. Building custom Oozie actions
   2. Building a Dynamic workflow
   3. Using the Oozie Java ApI
   4. Using uber jars with Oozie
   5. Building Dynamic data pipelines