

Recap

Saturday, July 30, 2022 7:51 PM

1. Structure of data today
2. Big Data
3. Transportation Services Use Case
4. Big Data Solution Architecture
5. Hadoop
6. Distributed File System
7. Hadoop Cluster Architecture
8. Block Replication
9. Rack Awareness

Topics for today

Sunday, April 24, 2022 6:29 PM

1. Hadoop Ecosystem
2. Yarn Components
3. Hadoop Configuration File
4. Hadoop HDFS Commands
5. Big Data Analytics
6. Introduction to Scala
7. Scala Datatypes
8. Scala Variables
9. Control Structures
10. Scala Functions

Hadoop Ecosystem

23 June 2020 00:16

Slide 49/50

Hadoop YARN

23 June 2020 23:45

YARN - Yet Another Resource Negotiator

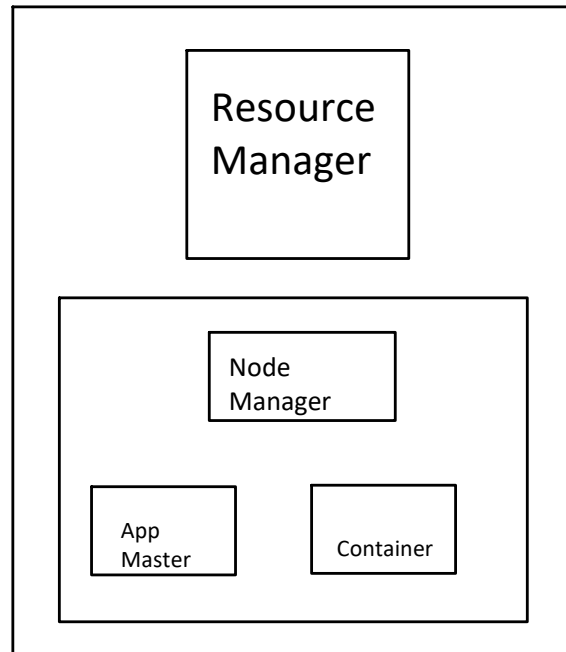
YARN Components

23 June 2020 23:50

Master Daemon , It manages
all other daemons
Accepts Jobs Submission

Responsible for Containers,
monitoring their resources
usage

App Master
One per application
Negotiates Resources



Container

CPU, RAM
Actuals task run

Slide 56

Files location

`/opt/cloudera/parcels/CDH/lib/hadoop/etc/hadoop/`

Core - Core-site.xml - Hdfs and MapReduce

HDFS - hdfs-site.xml - HDFS specific

YARN - YARN-site.xml - YARN Specific

Mapreduce - mapred-site.xml - MapReduce Specific

Hadoop /Hdfs Commands

Sunday, August 30, 2020 6:12 PM

hadoop version

Check HDFS health

```
hdfs fsck /user/edureka_967855
```

List the directories in HDFS

```
hdfs dfs -ls /user/edureka_967855
```

TO create a Directory in hdfs

```
hdfs dfs -mkdir new_edureka
```

To Delete the directory

```
hdfs dfs -rmdir new_edureka
```

```
hdfs dfs -rm -r -f -skipTrash new_edureka
```

to Copy data from Local Machine to HDFS

```
hdfs dfs -copyFromLocal /mnt/home/edureka_  
967855/csvsample/csvfile.txt /user/edureka_  
967855/textfile
```


1) Batch Processing

Interest amount credited into our account after every quarter

2) Real Time Processing

Introduction to Scala

27 June 2020 18:53

Sca - Scalable

La - Language

Modern Multi Paradigm Language

Cobol - Procedural Language

Java - OOps

Haskel - Functional

Scala Supports both OOPs and Functional Paradigm

Scala For Spark

Spark is Written in Scala Language

Matei Zaharia

Play - Web Application

Akka - ConcurrentApplications

Spark - Big data Processing

Graph QL - No SQL Graphical Database Build over
scala Language

Scala REPL

Read Evaluate Print Loop

Scala Datatypes

27 June 2020 19:04