HDFS Architecture
Part - 3

HDFS

Hadoop Distributed File System

by Sumit Mittal

IMPORTANT

Copyright Infringement and Illegal Content Sharing Notice

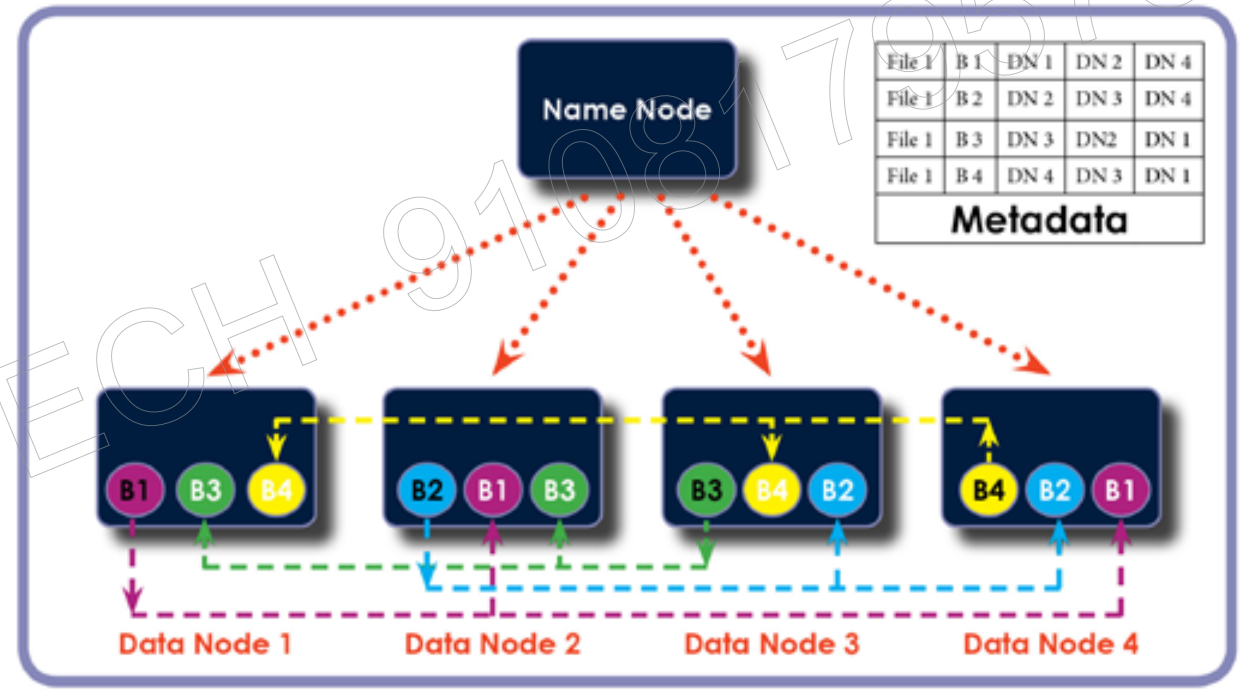
HDFS ARCHITECTURE

Understanding Replication Factor

Each data block in Hadoop has **3** replicas

Replicas are stored in different locations

The replica locations are also updated in Metadata stored in the Name Node

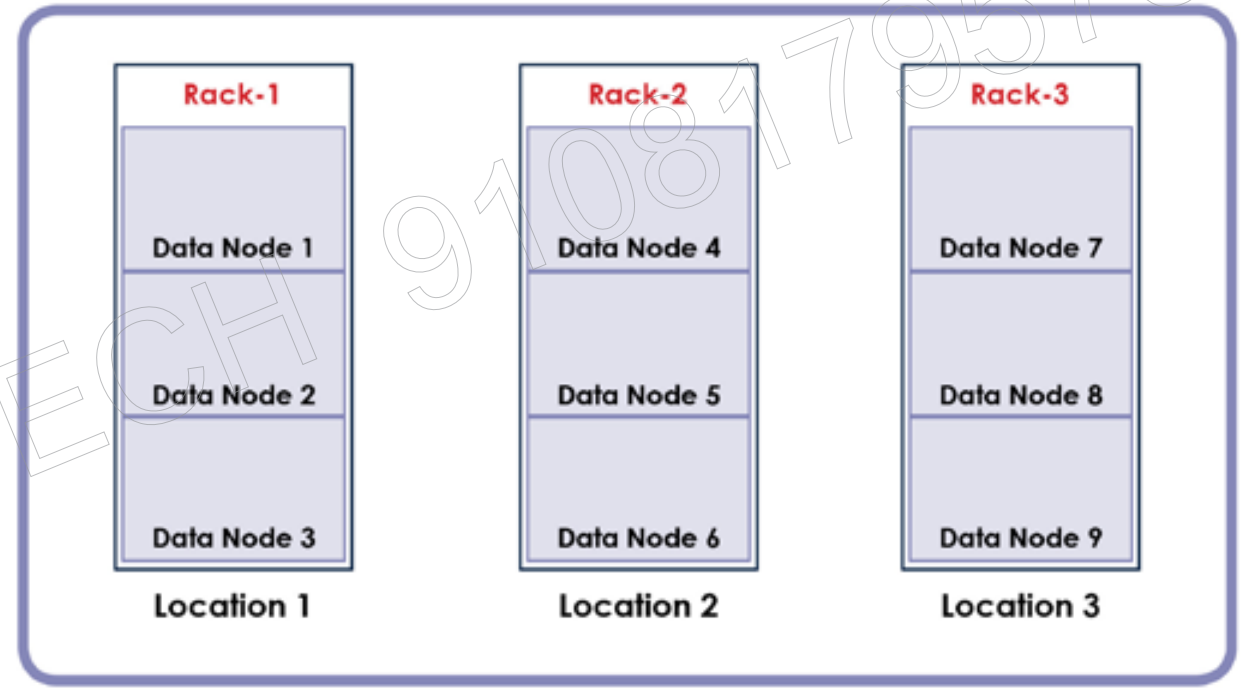


4 node Cluster

HDFS ARCHITECTURE

How Replicas are placed

Rack means group of systems placed in different geographical locations



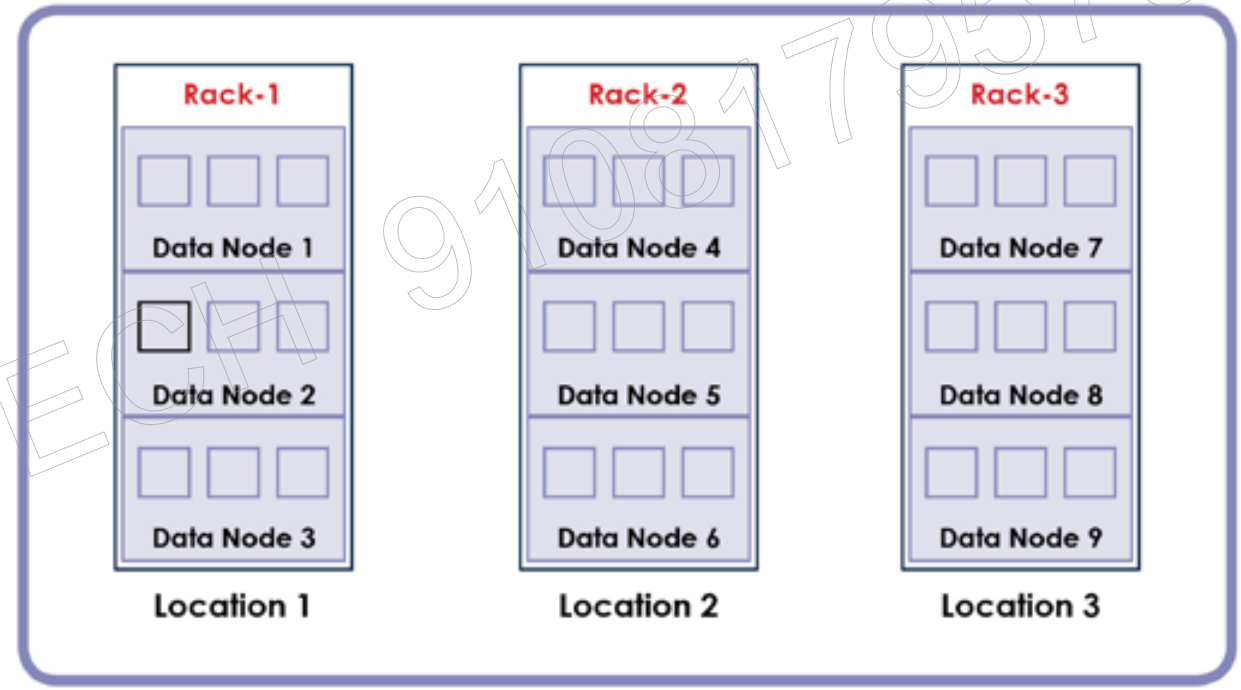
HDFS ARCHITECTURE

Replica Pipeline

Replicas of data block are forwarded to new locations

Forwarding data blocks requires large amount of network bandwidth

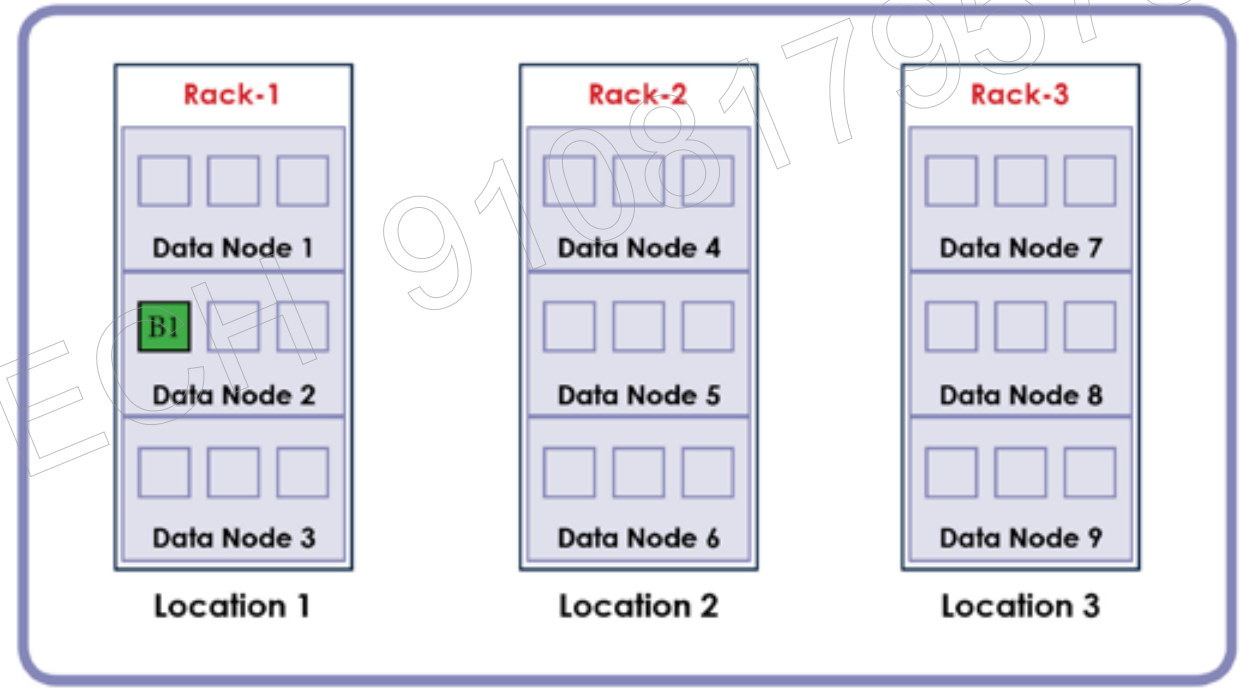
Involves more input-output operations



HDFS ARCHITECTURE

Replica Pipeline

Name Node stores data block in a data node of a rack



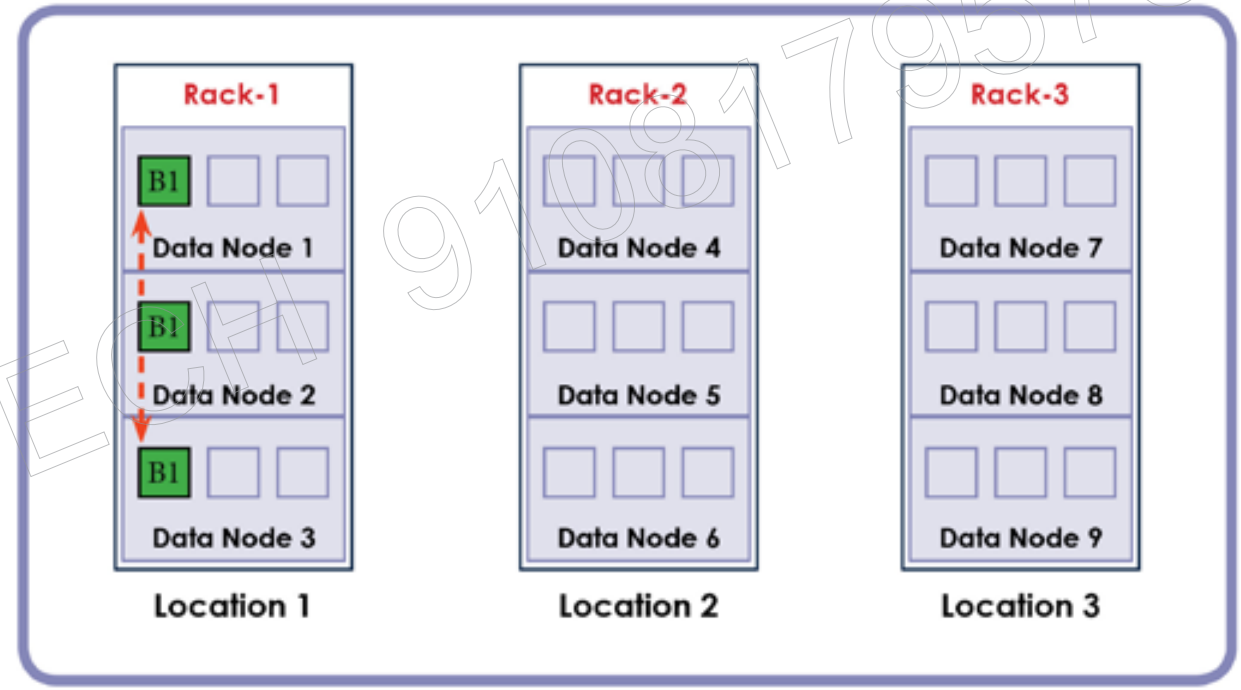
HDFS ARCHITECTURE

Replica Pipeline

Replicas of data block are forwarded to new locations

Forwarding data blocks within the same rack requires small amount of network bandwidth

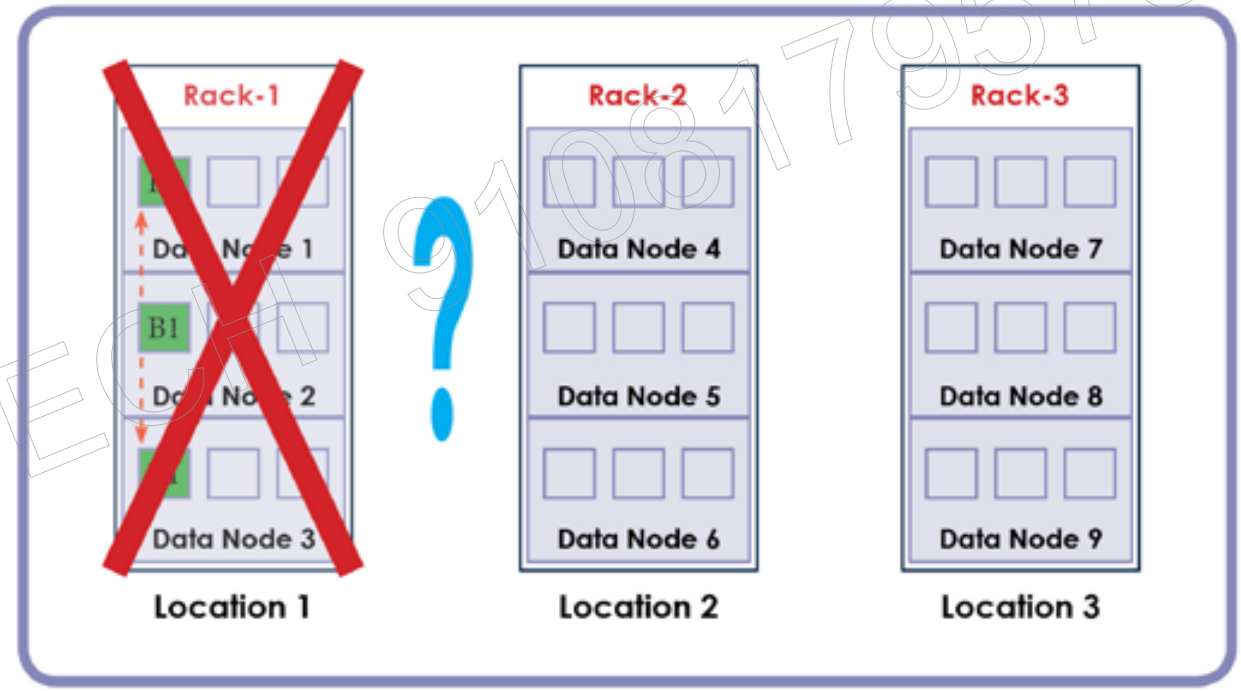
Involves less input-output operations



HDFS ARCHITECTURE

Rack Failure

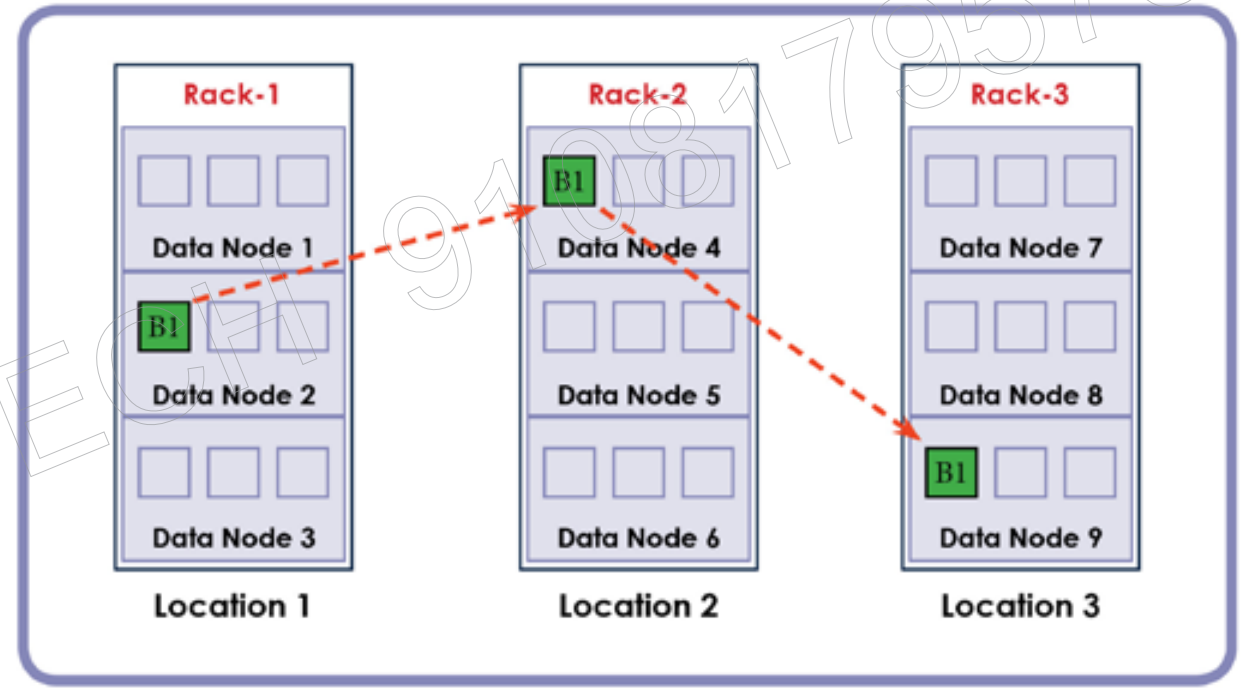
If we place all replicas in a single rack, there are high chances of **data loss** if the rack goes down



HDFS ARCHITECTURE

Choosing Multiple Racks

Each Replica placed in different rack is also not a ideal solution

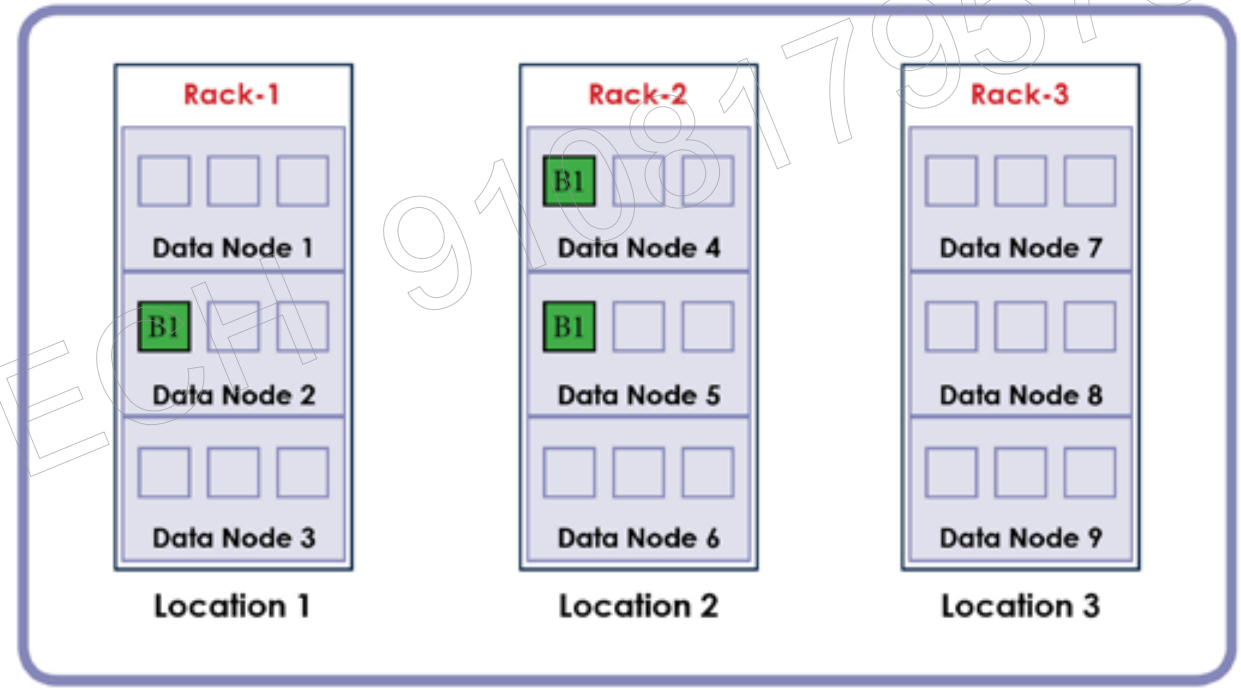


HDFS ARCHITECTURE

Rack Awareness Mechanism

The balanced approch is to place replicas in two different racks

One replica in one rack and other two in a different rack or vice versa



SUMMARY

- Data Nodes
- Name Node
- Fault Tolerance
- Replication Factor
- Checkpointing
- Rack Awareness Mechanism

Thank You !!

9108179578

Visit : *https://trendytech.in/courses/big-data-online-training/*

Google REVIEWS

5 Star Google Rated
Big Data Course

LEARN FROM THE EXPERT

9108179578

Call for more details

Follow US

Trainer	Mr. Sumit Mittal
Phone	9108179578
Email	trendytech.sumit@gmail.com
Website	https://trendytech.in/courses/big-data-online-training/
LinkedIn	https://www.linkedin.com/in/bigdatabysumit/
Twitter	@BigdataBySumit
Instagram	bigdatabysumit
Facebook	https://www.facebook.com/trendytech.in/
Youtube	TrendyTech