HDFS Architecture
Part - 1

HDFS

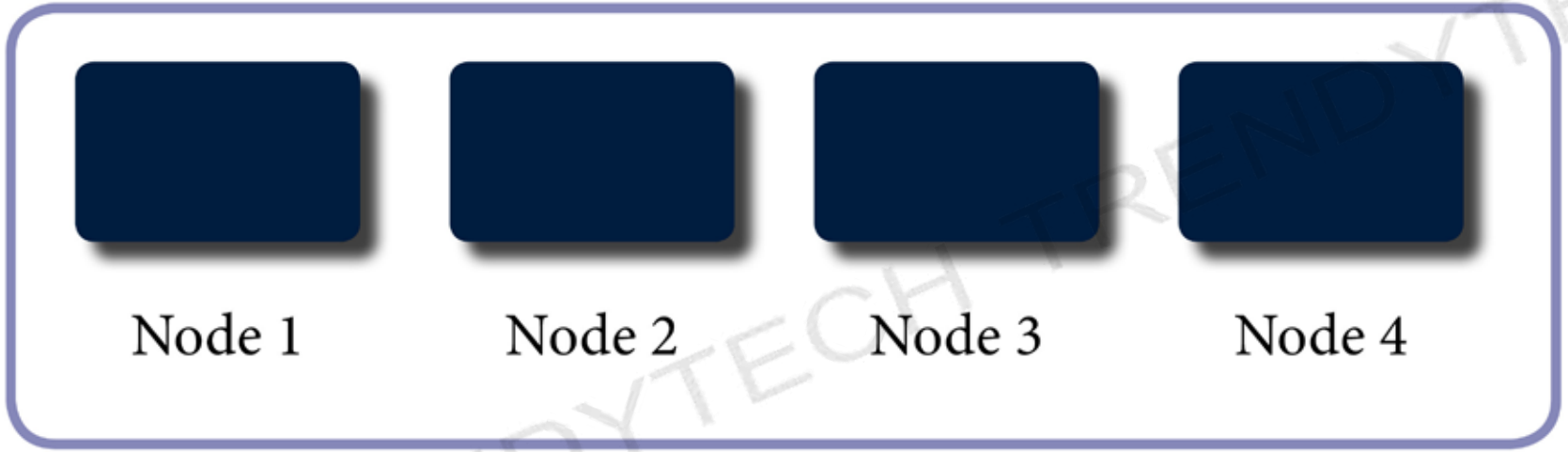
Hadoop Distributed File System

by Sumit Mittal

HDFS ARCHITECTURE

TRENDYTECH TRENDYTECH TRENDYTECH

HDFS ARCHITECTURE

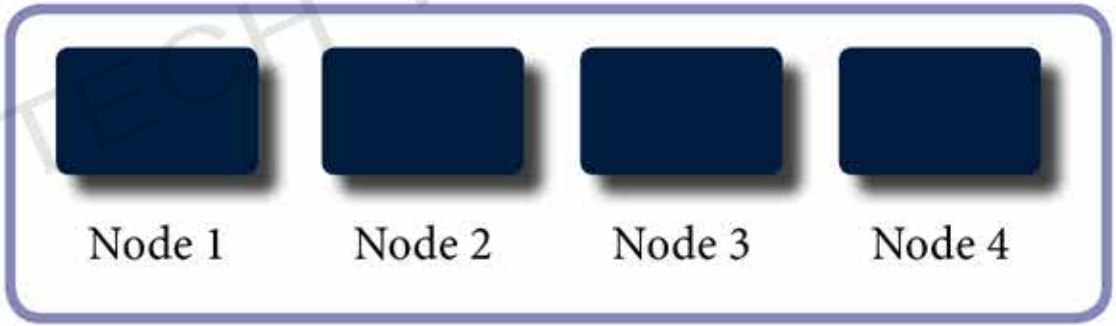


4 Node Cluster

TRENDYTECH TRENDYTECH TRENDYTECH

HDFS ARCHITECTURE

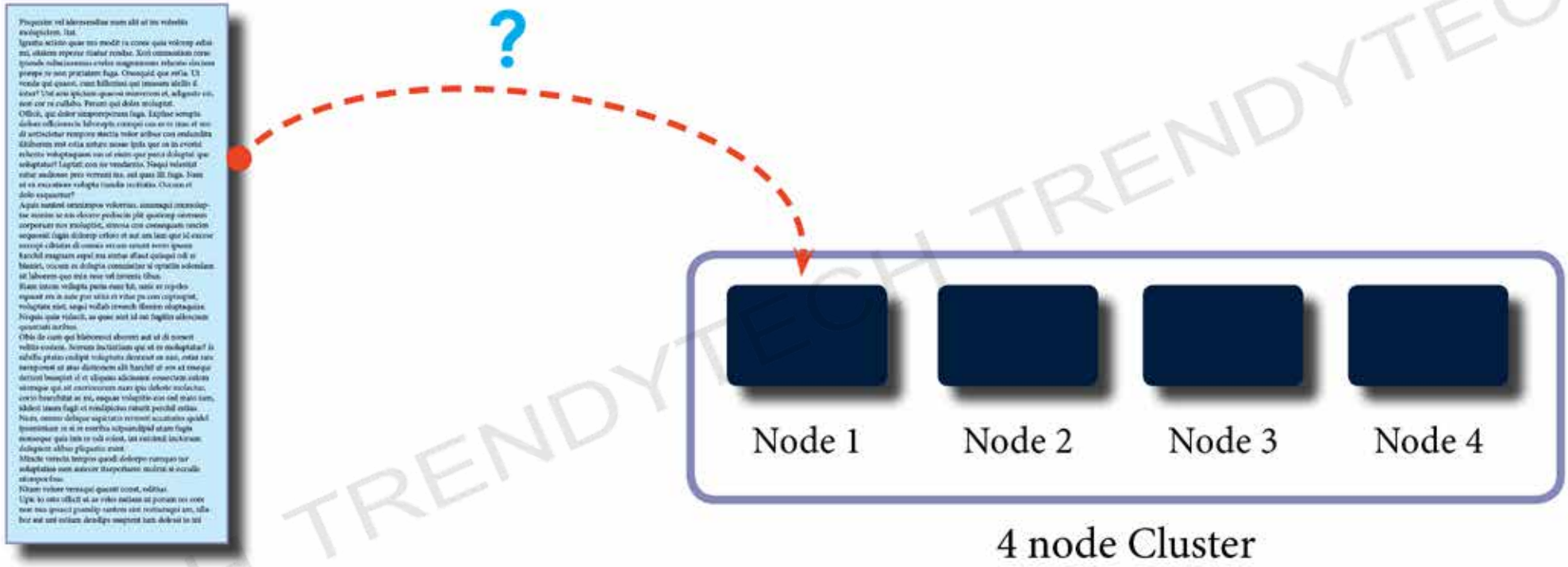
File 1
(500 mb)



4 node Cluster

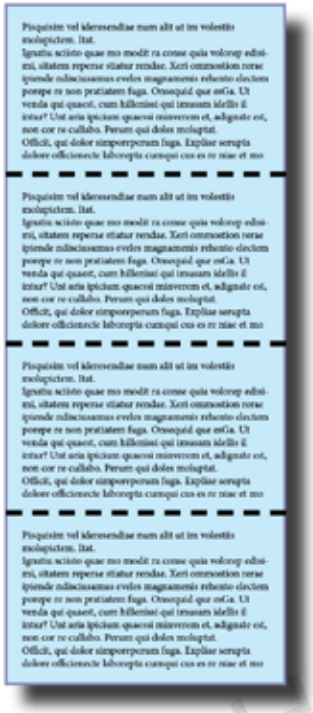
TRENDYTECH TRENDYTECH TRENDYTECH

HDFS ARCHITECTURE



TRENDYTECH TRENDYTECH TRENDYTECH

HDFS ARCHITECTURE



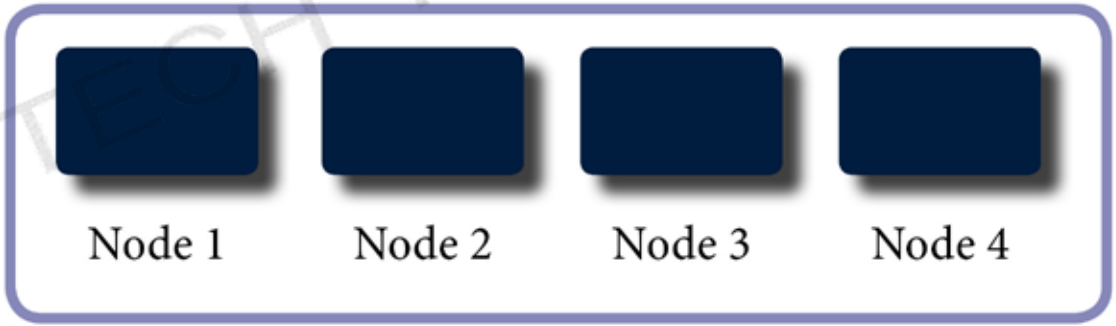
block 1

block 2

block 3

block 4

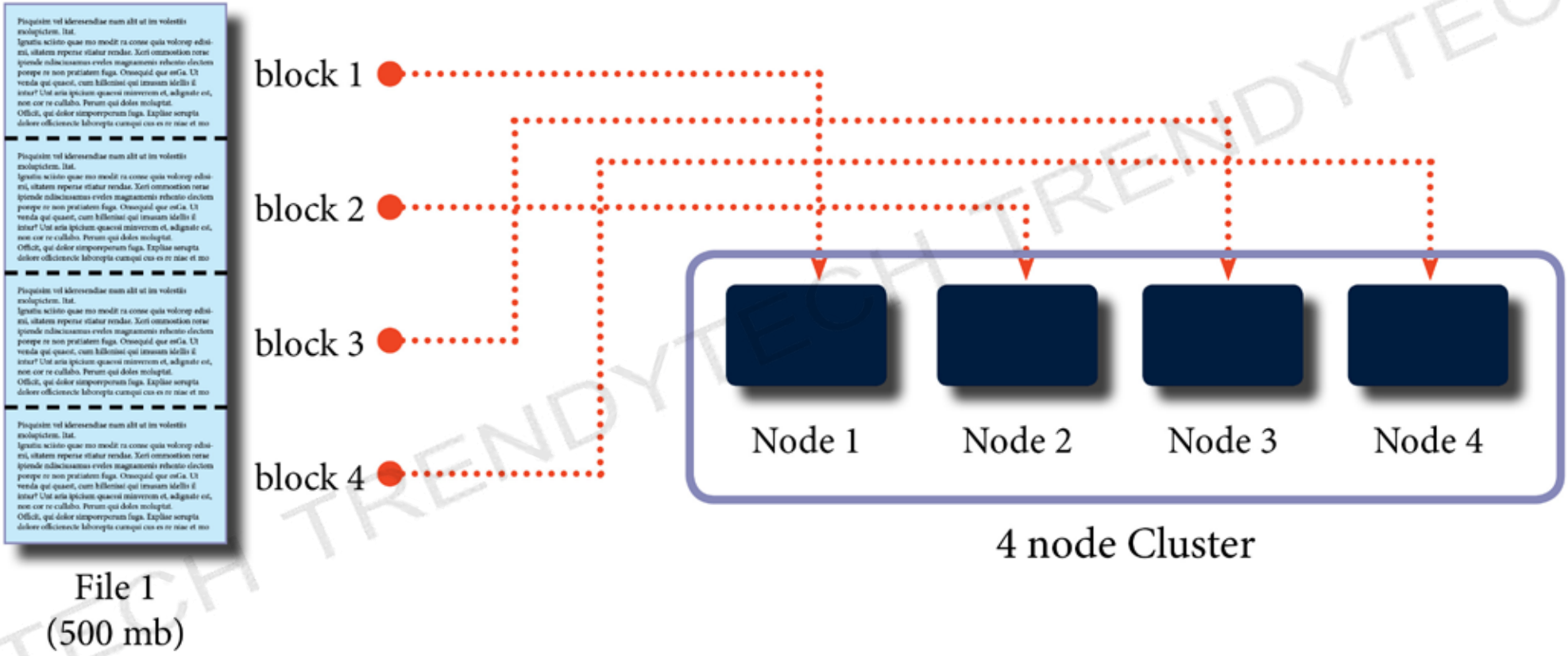
File 1
(500 mb)



4 node Cluster

TRENDYTECH TRENDYTECH TRENDYTECH

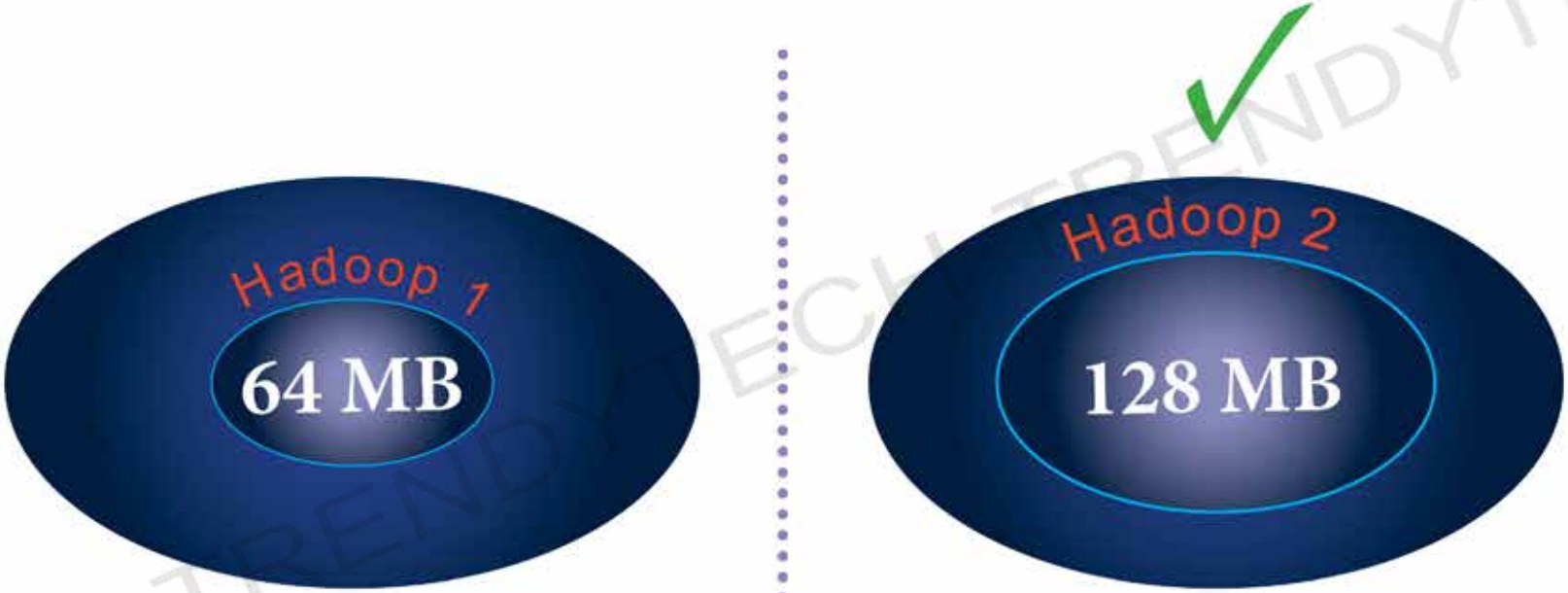
HDFS ARCHITECTURE



TRENDYTECH

HDFS ARCHITECTURE

HDFS Block Size

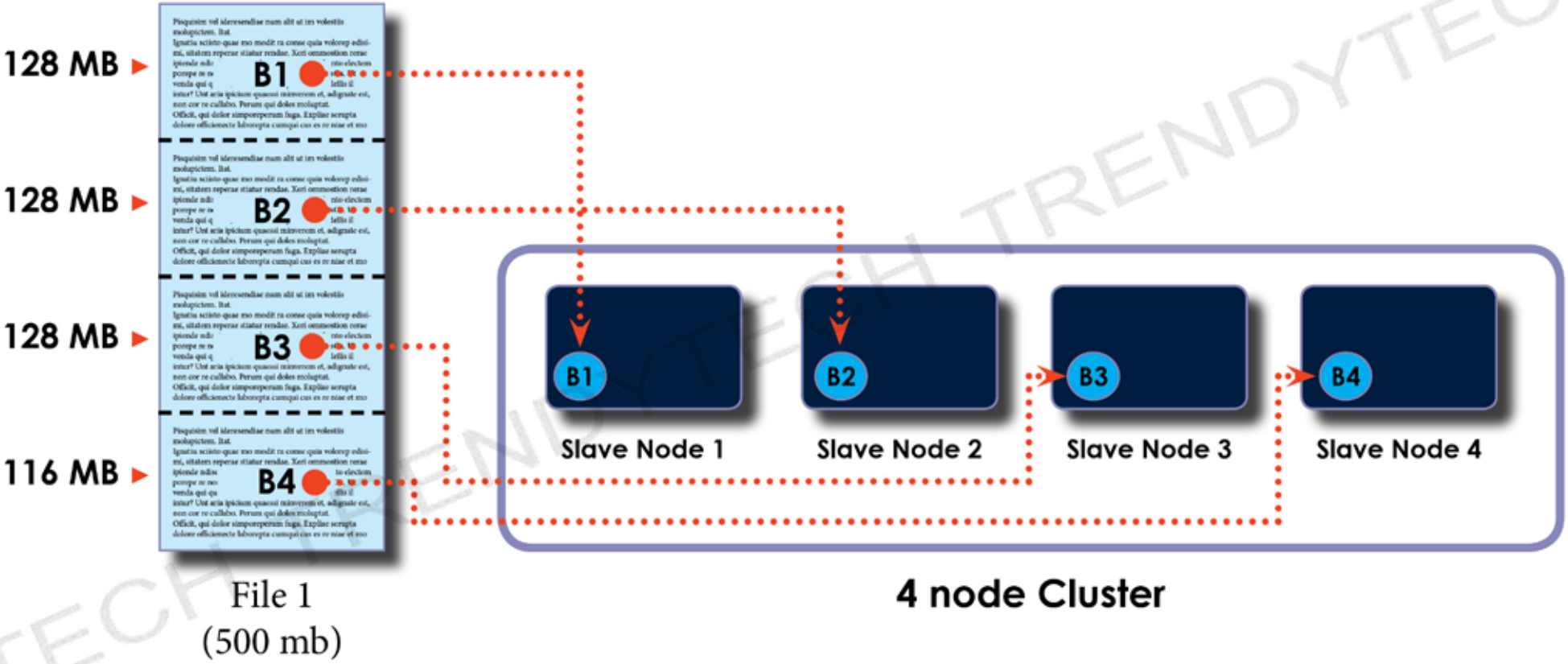


older version

current version

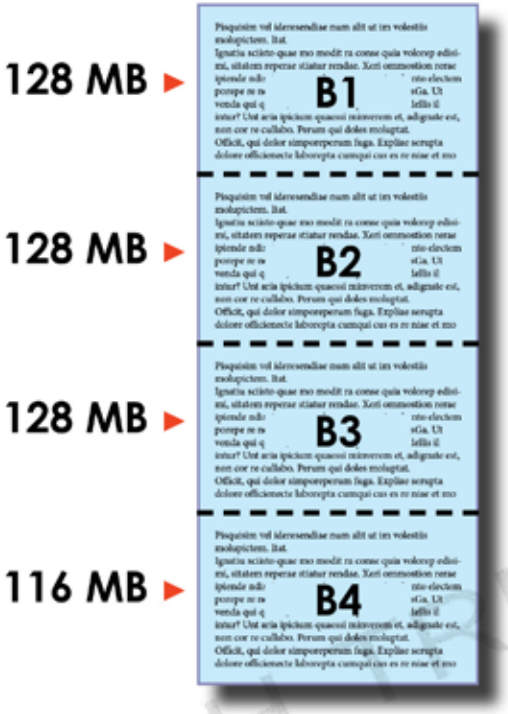
TRENDYTECH

HDFS ARCHITECTURE

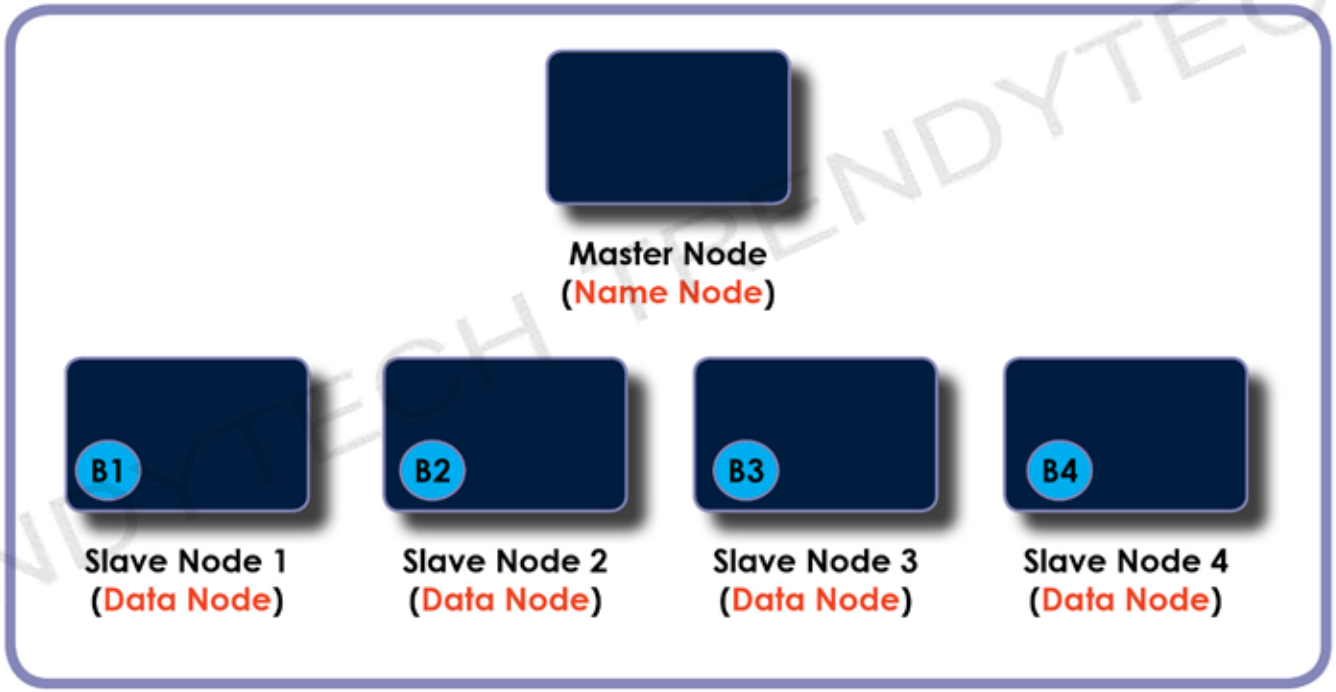


TRENDYTECH

HDFS ARCHITECTURE

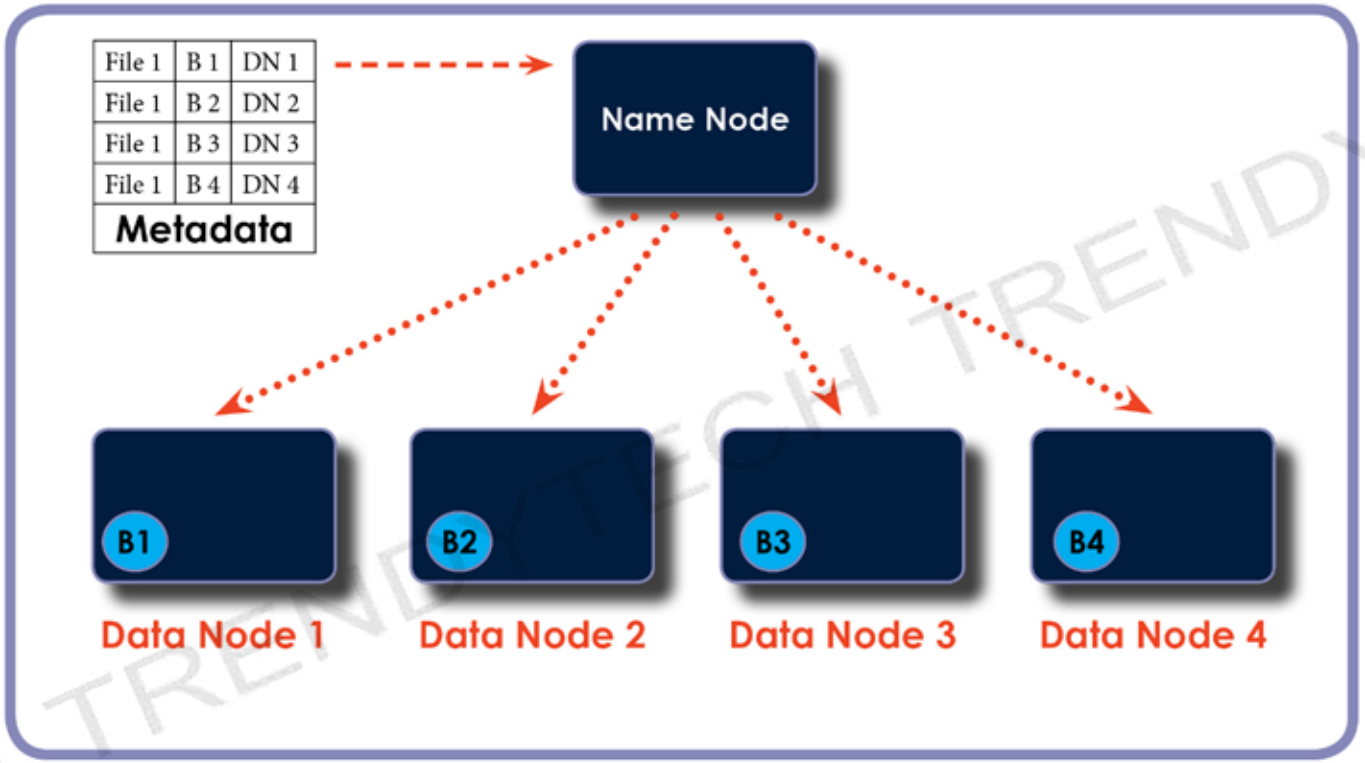


File 1
(500 mb)



4 node Cluster

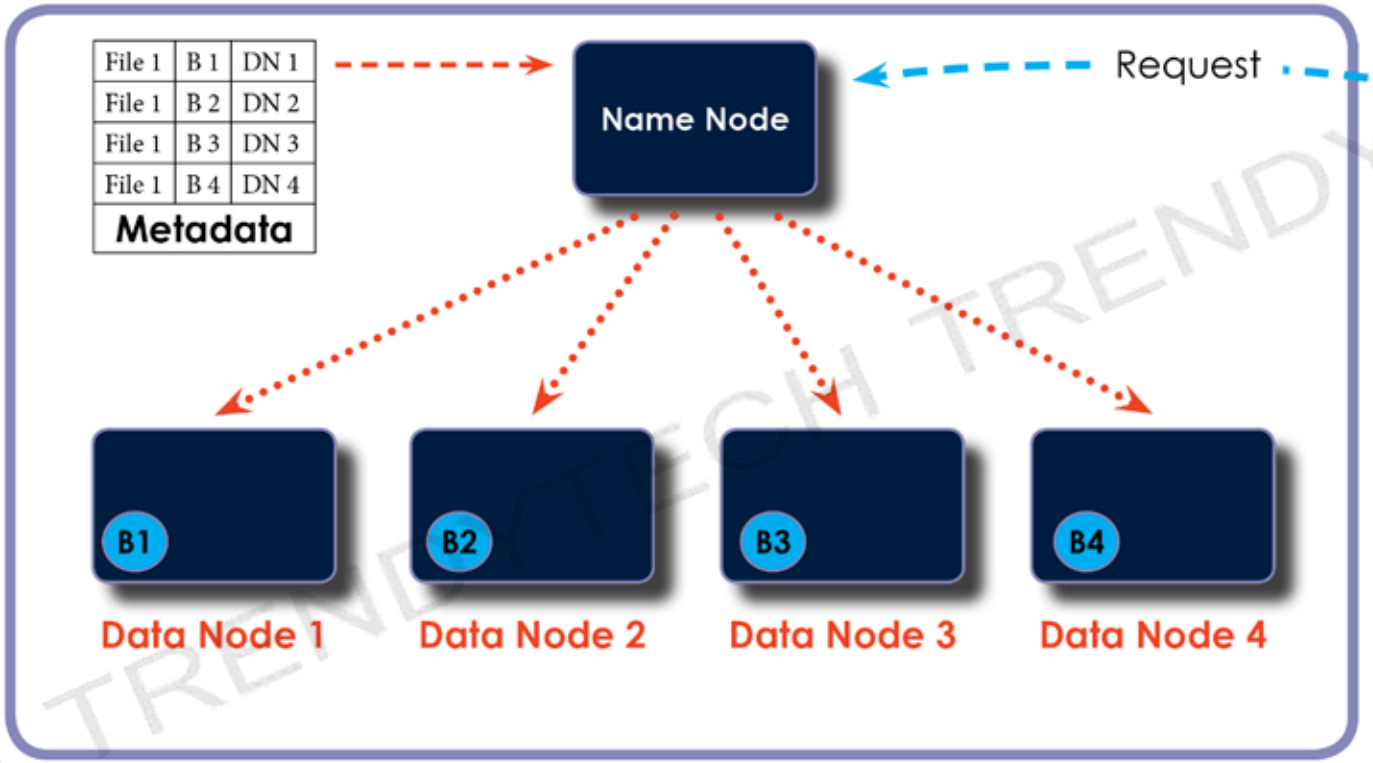
HDFS ARCHITECTURE



4 node Cluster

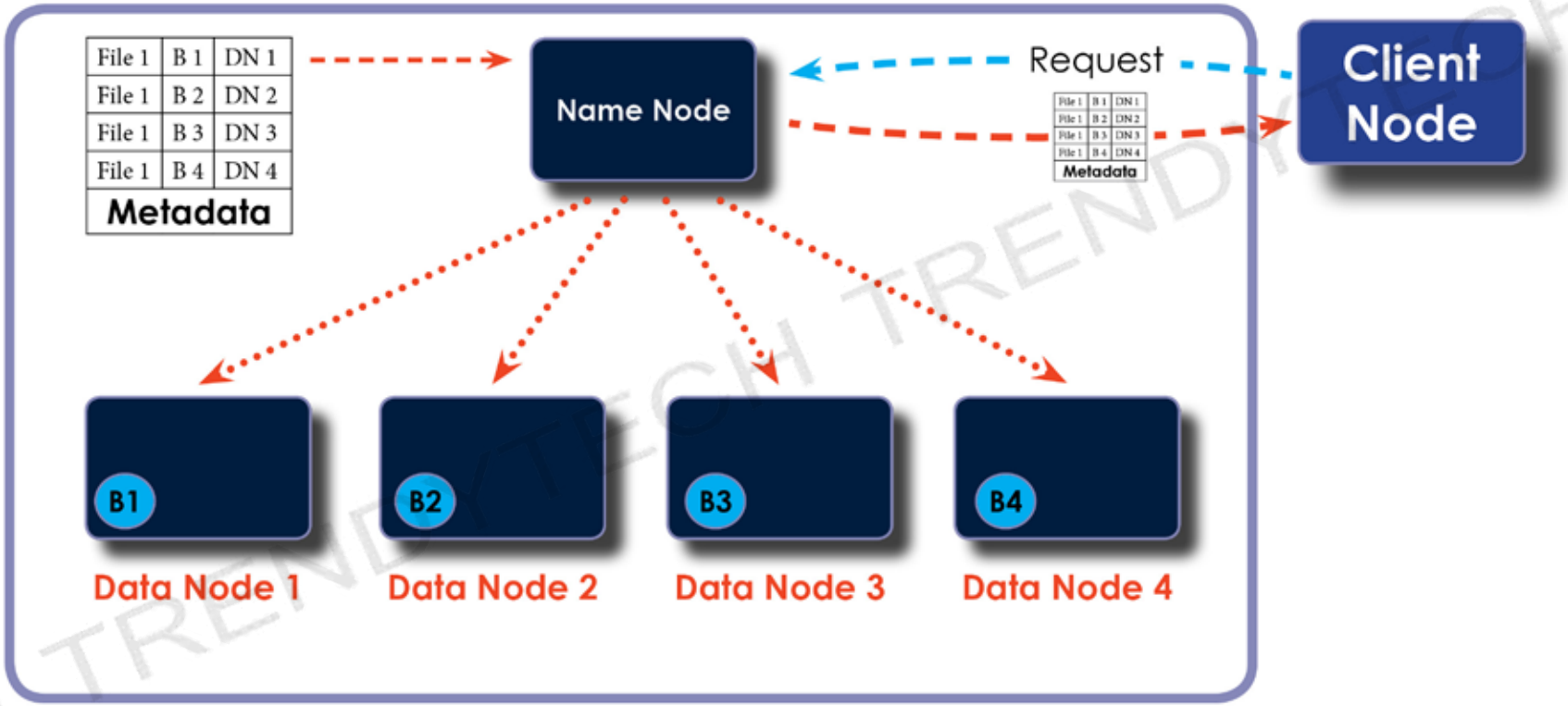
TRENDYTECH

HDFS ARCHITECTURE



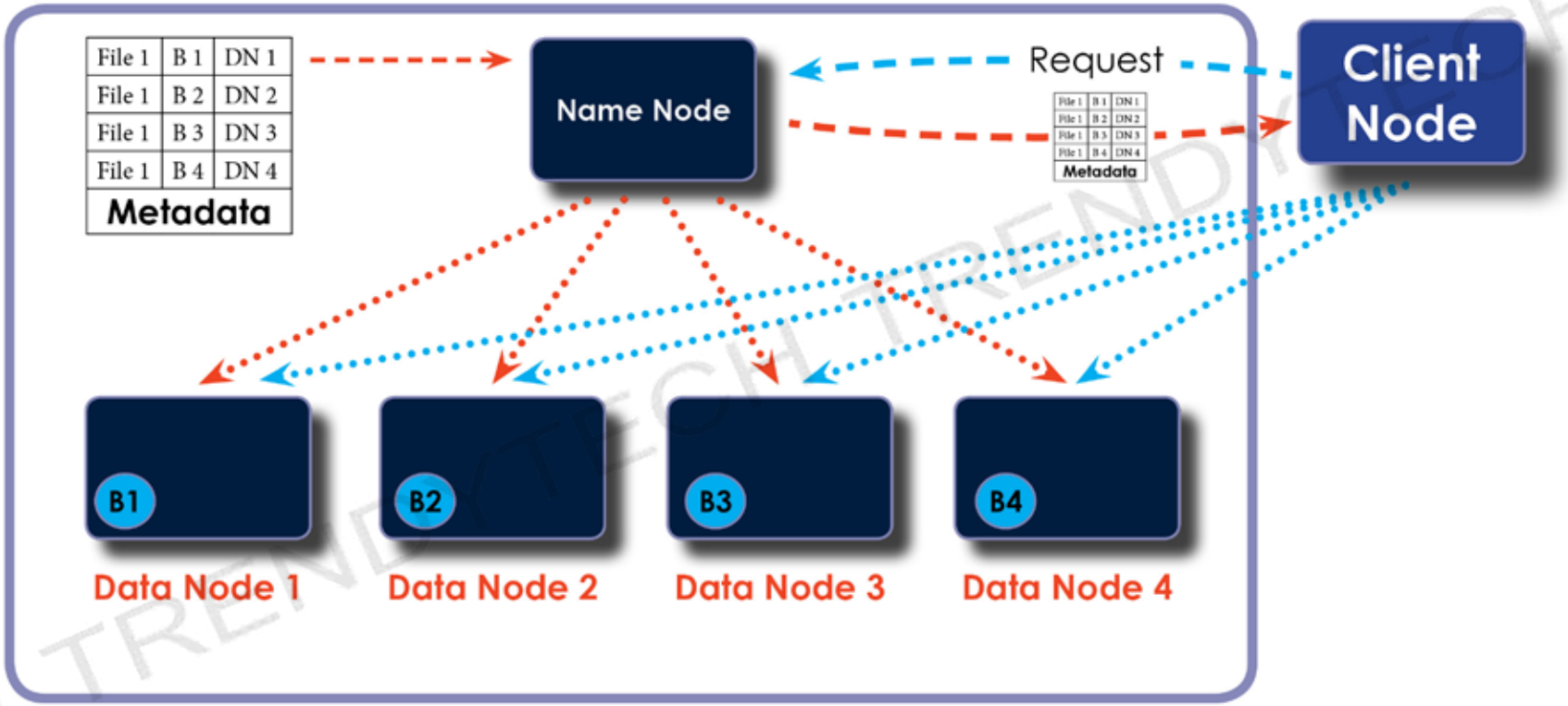
4 node Cluster

HDFS ARCHITECTURE



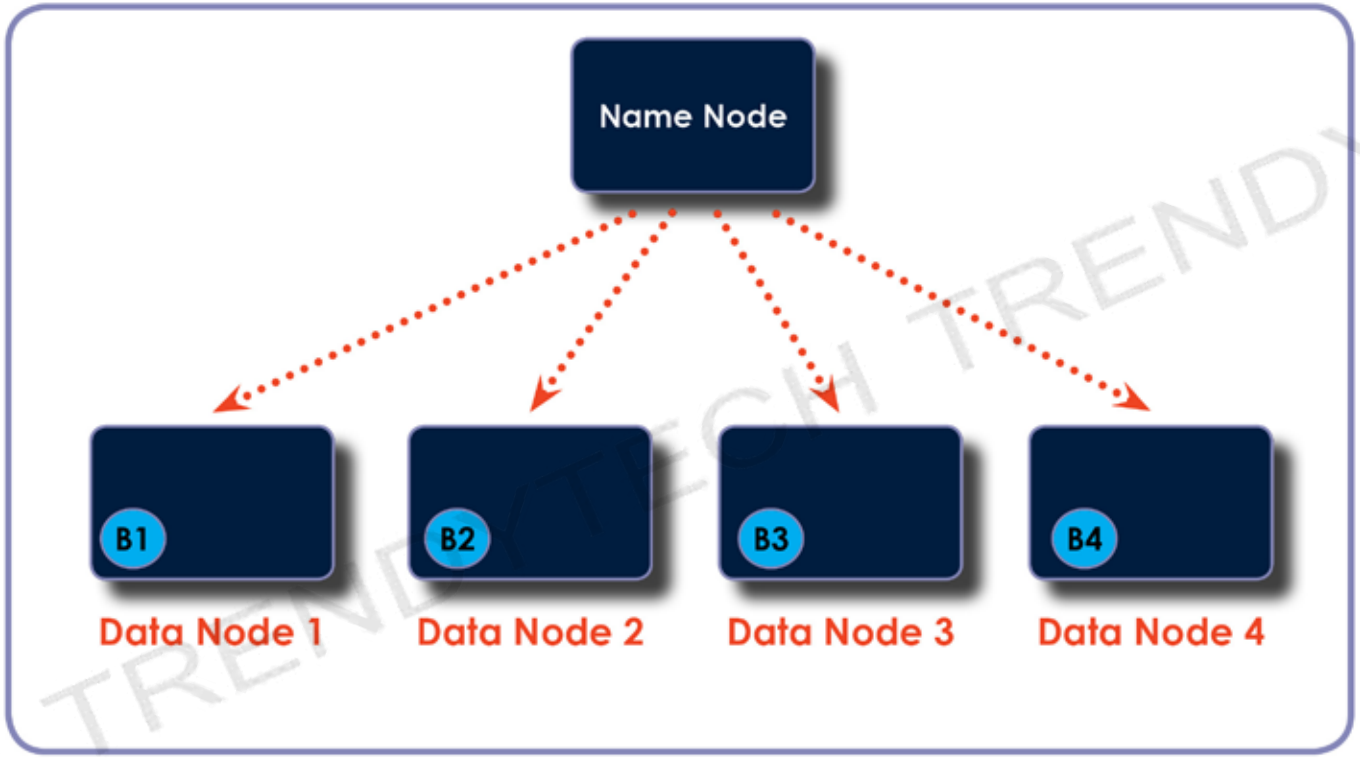
4 node Cluster

HDFS ARCHITECTURE



4 node Cluster

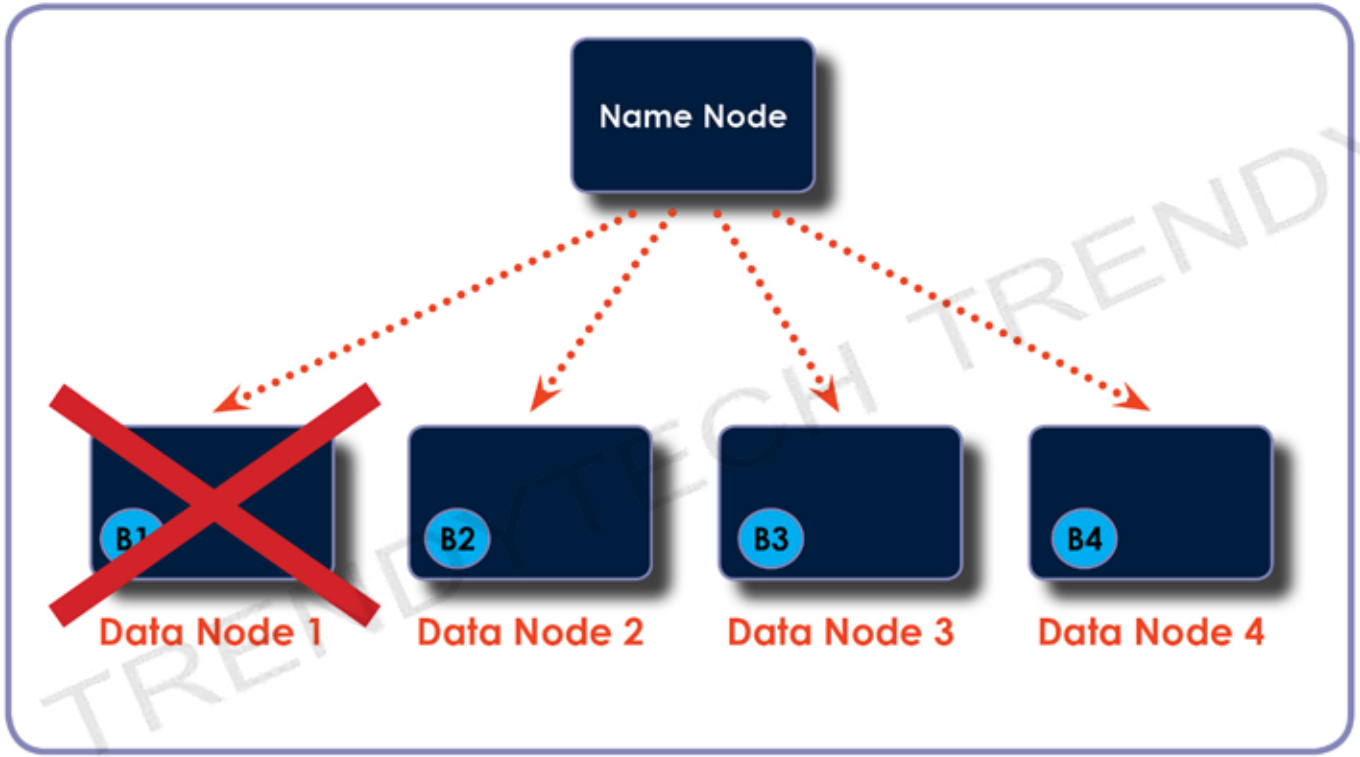
HDFS ARCHITECTURE



4 node Cluster

TRENDYTECH TRENDYTECH TRENDYTECH

HDFS ARCHITECTURE



?

4 node Cluster

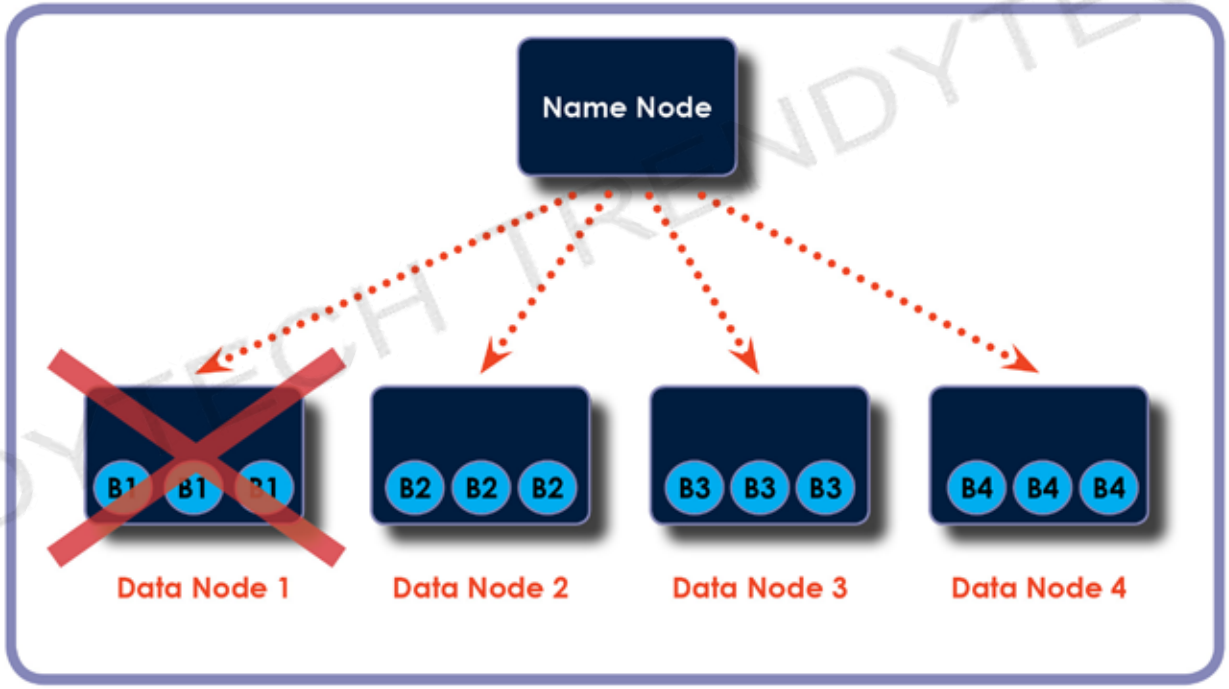
TRENDYTECH TRENDYTECH TRENDYTECH TRENDYTECH

HDFS ARCHITECTURE

Replication Factor

Default Hadoop Replication Factor:

3



4 node Cluster

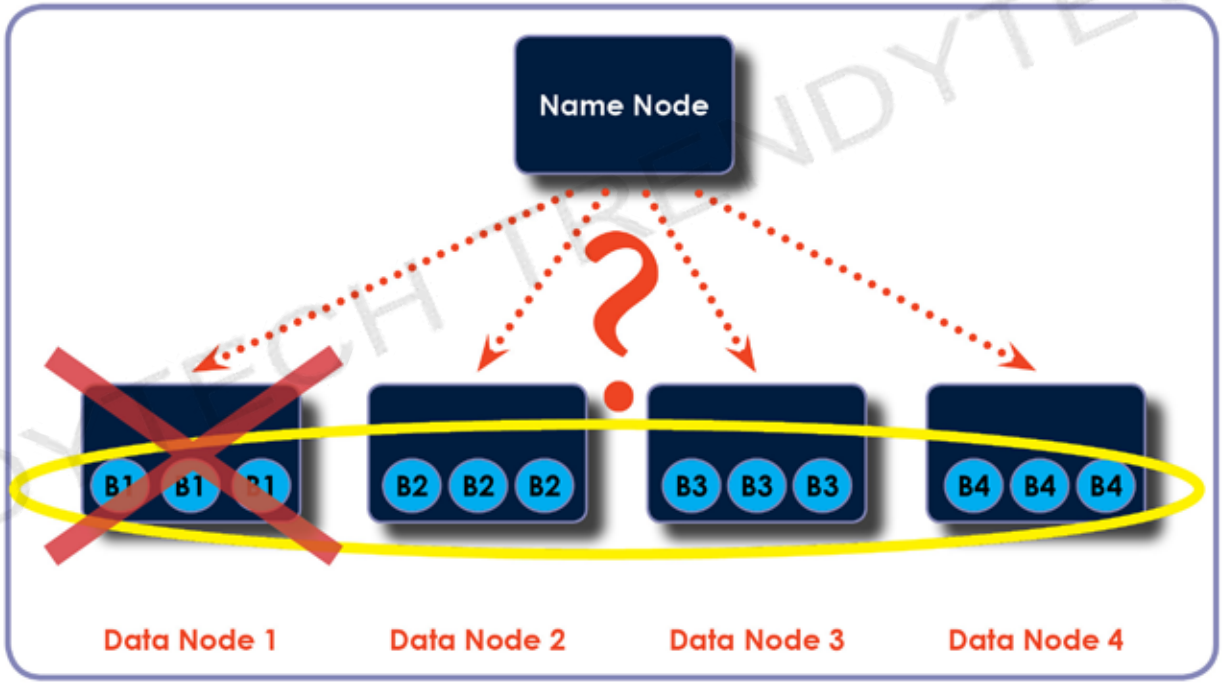
TRENDYTECH TRENDYTECH TRENDYTECH

HDFS ARCHITECTURE

Replication Factor

Default Hadoop Replication Factor:

3



4 node Cluster

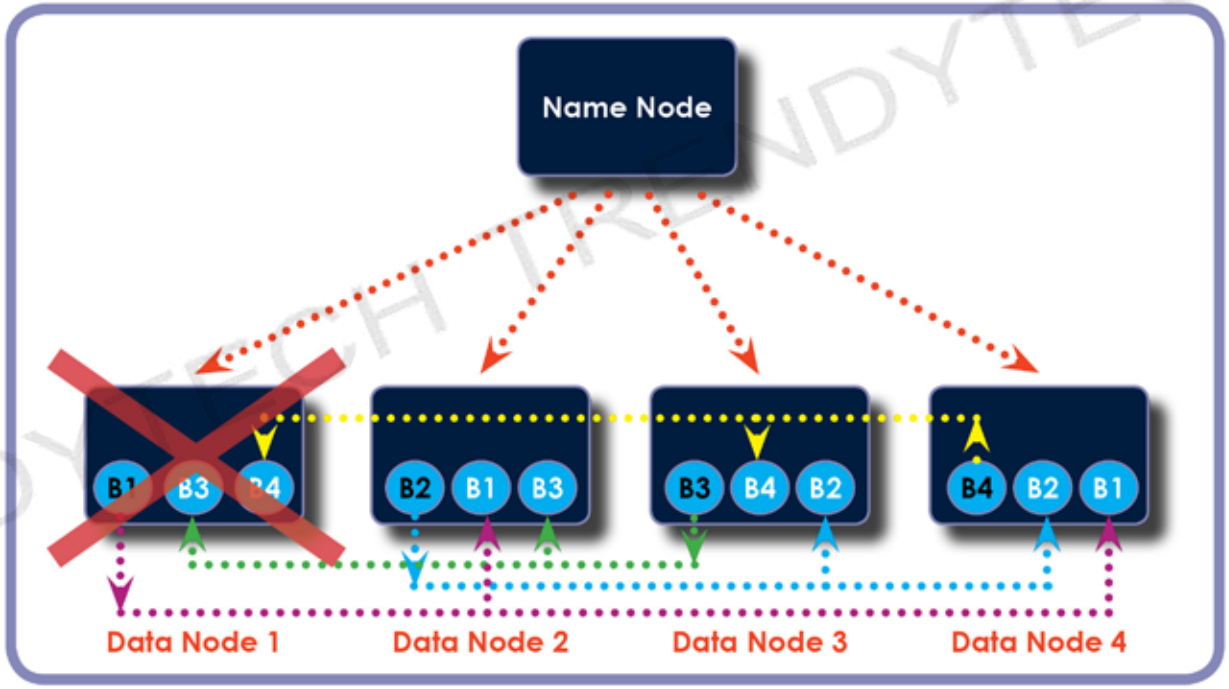
TRENDYTECH TRENDYTECH TRENDYTECH

HDFS ARCHITECTURE

Replication Factor

Default Hadoop Replication Factor:

3



4 node Cluster

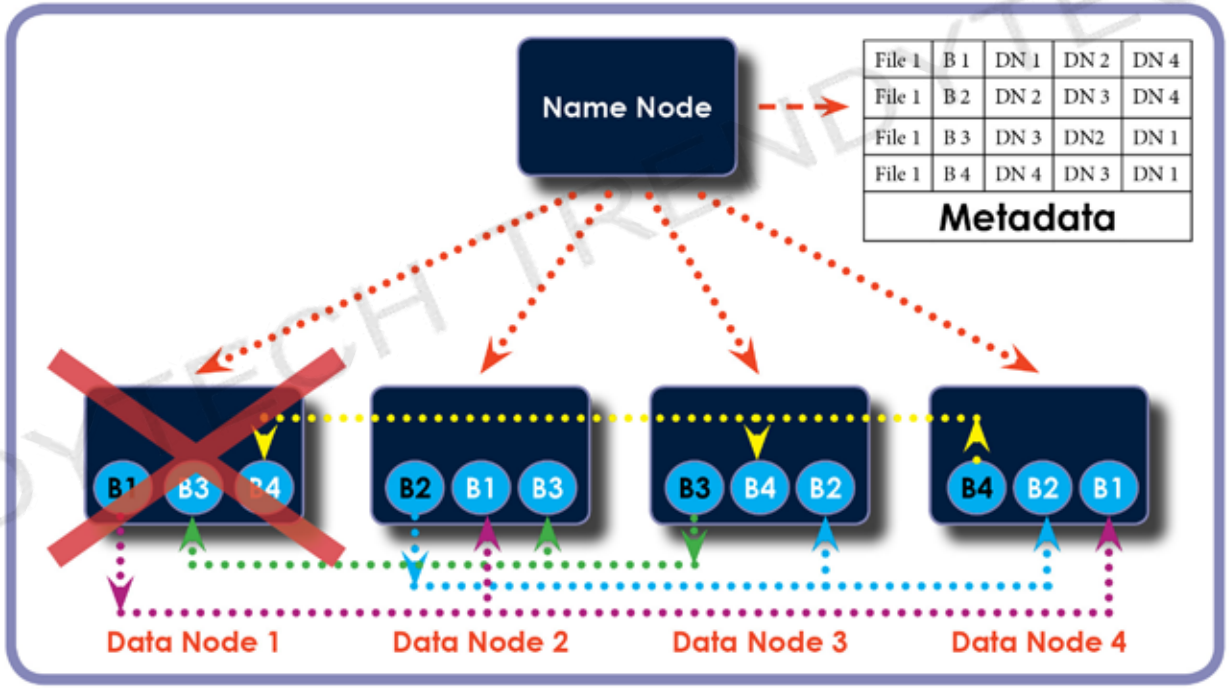
TRENDYTECH TRENDYTECH TRENDYTECH

HDFS ARCHITECTURE

Replication Factor

Default Hadoop Replication Factor:

3



4 node Cluster

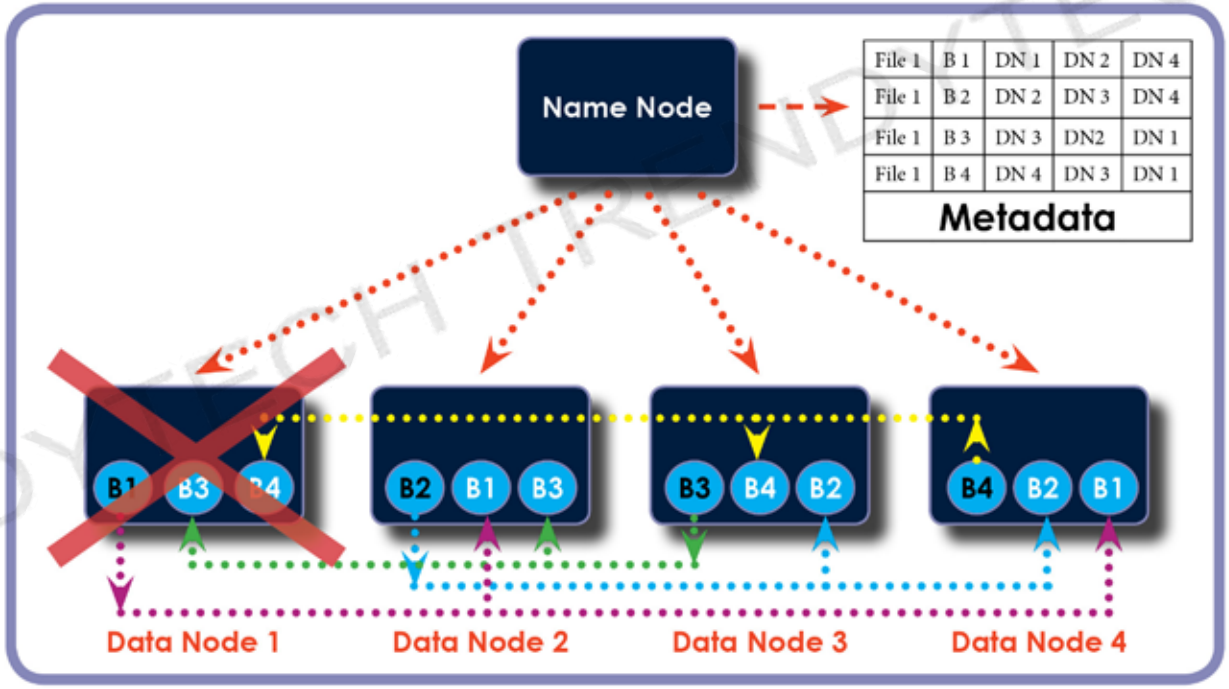
TRENDY TECH

HDFS ARCHITECTURE

Replication Factor

Default Hadoop Replication Factor: **3**

We can change Replication Factor as per our requirements



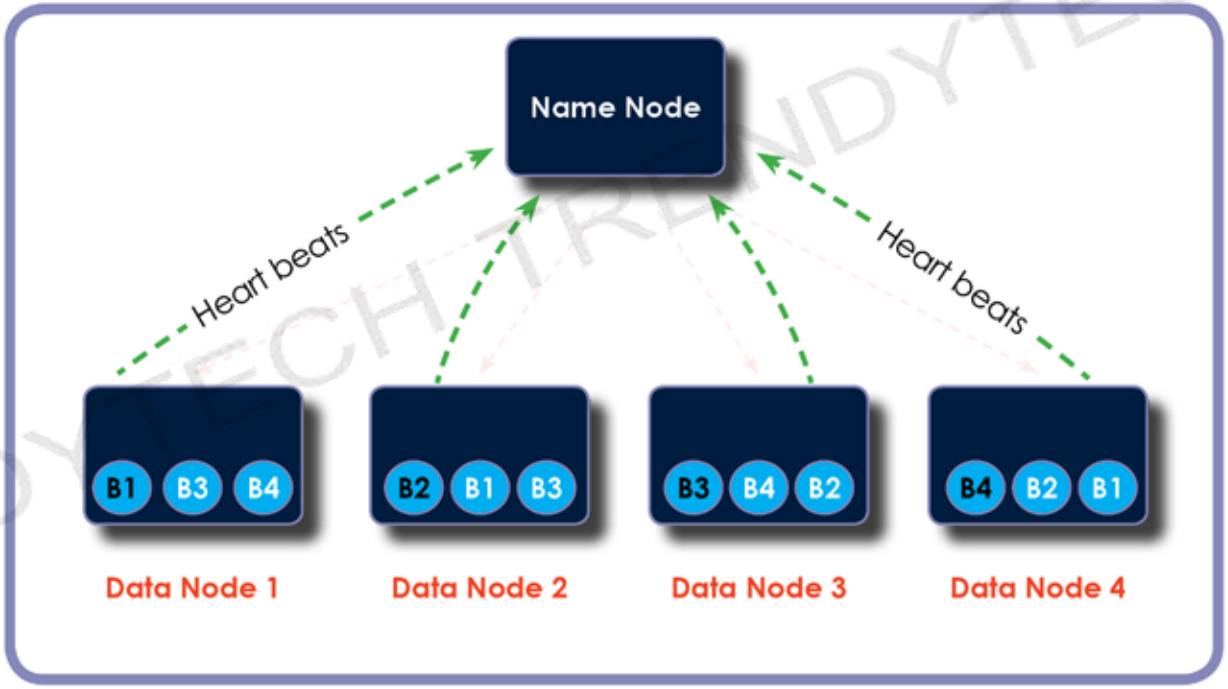
4 node Cluster

HDFS ARCHITECTURE

Heart Beat

Each Data Node sends heart beats to Name Node in every **3** seconds

If a Name Node doesn't receive **10** consecutive heart beats, it assumes that the Data Node is dead or running very slow



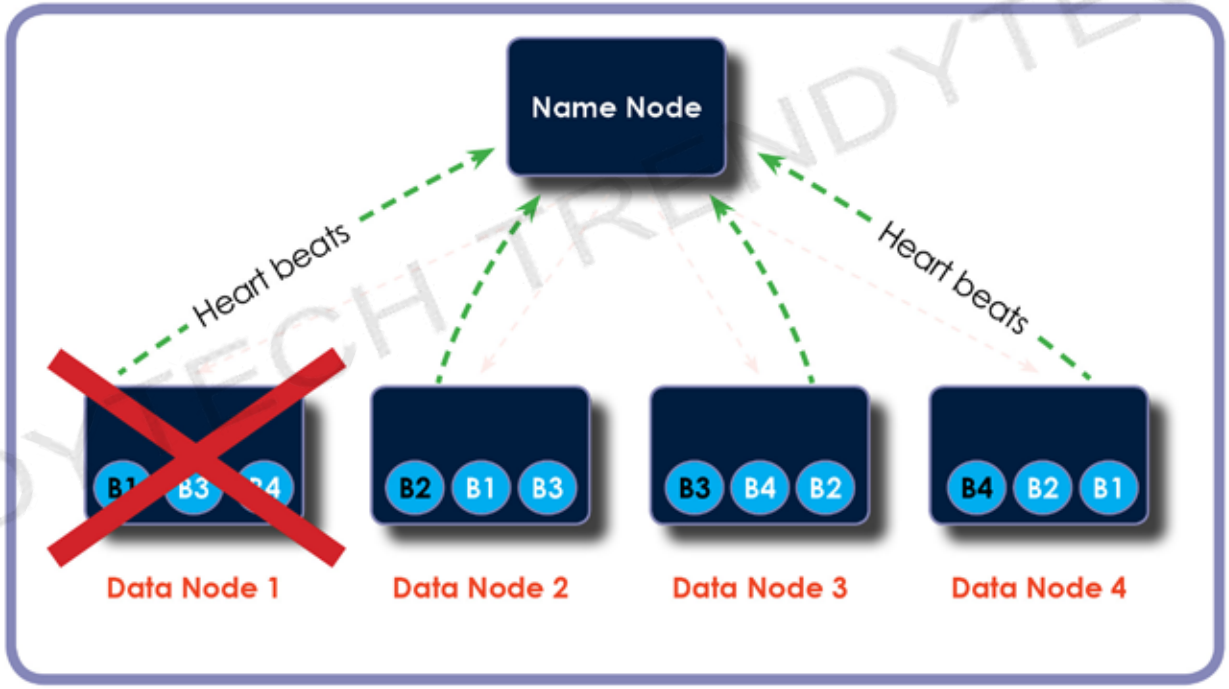
4 node Cluster

TRENDYTECH TRENDYTECH TRENDYTECH

HDFS ARCHITECTURE

Fault Tolerance

If a Data Node goes down the replication factor came down to **< 3**



4 node Cluster

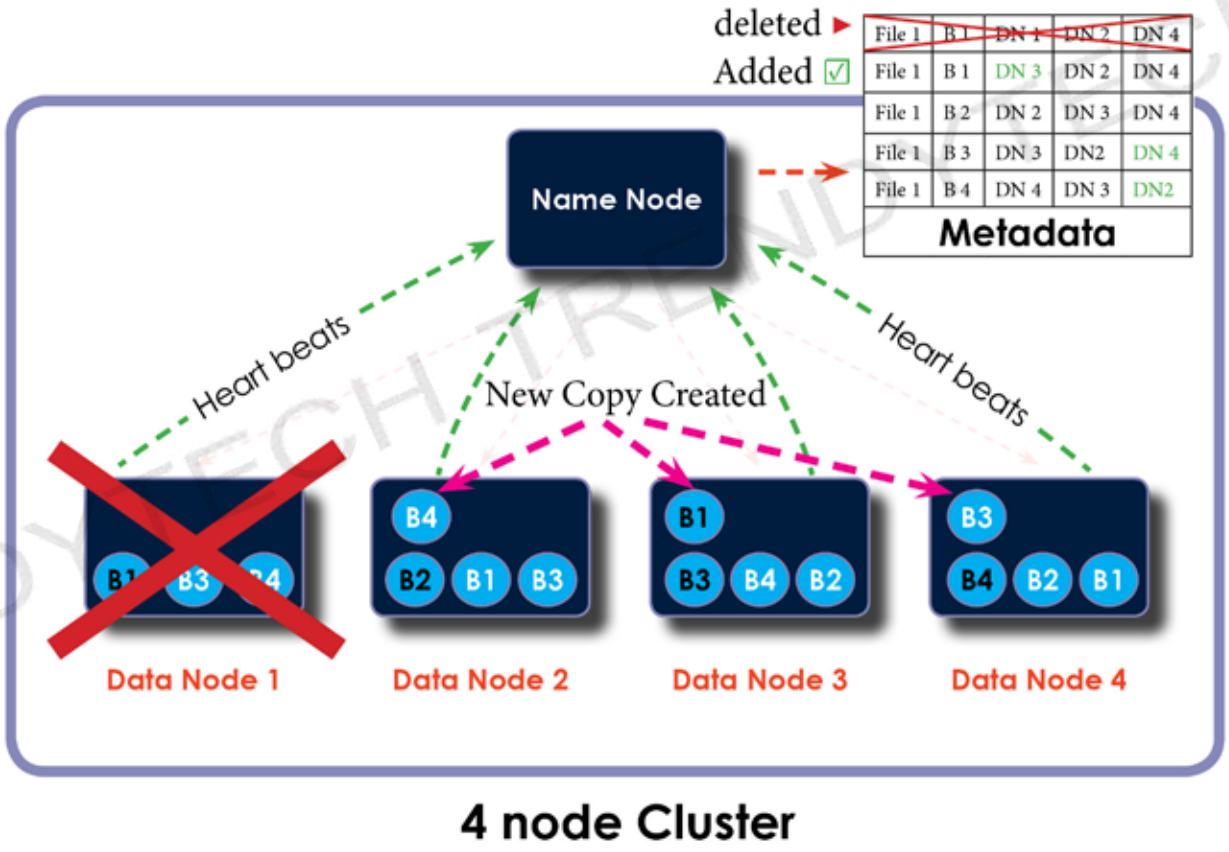
TRENDYTECH TRENDYTECH TRENDYTECH

HDFS ARCHITECTURE

Fault Tolerance

If a Data Node goes down the replication factor came down to **< 3**

Name Node will create one more copy to maintain the replication factor



4 node Cluster

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