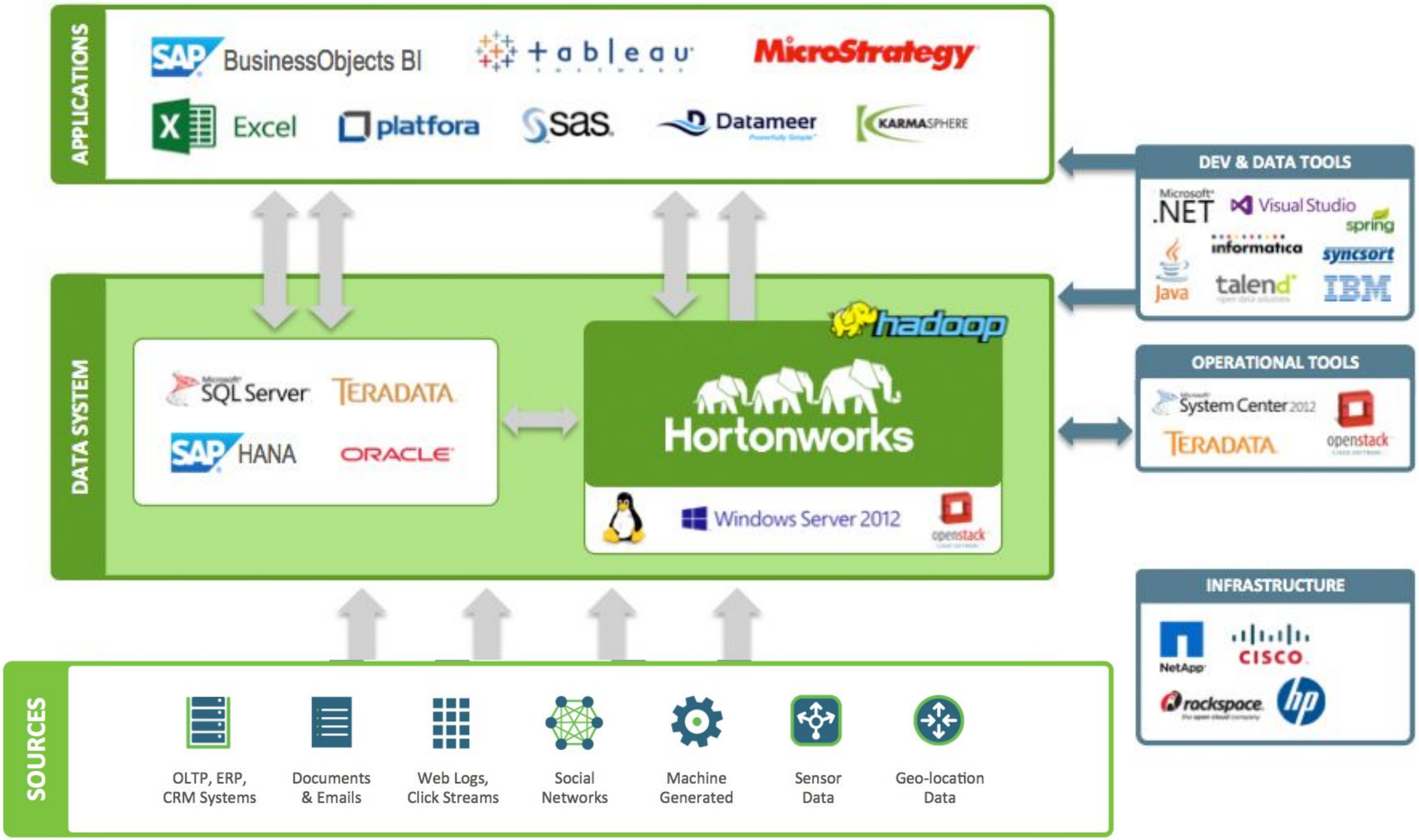
Hadoop as +1 Architecture

- Though it has the potential to replace all others, it can also be used to complement existing systems if they can't be removed due to any constraints.



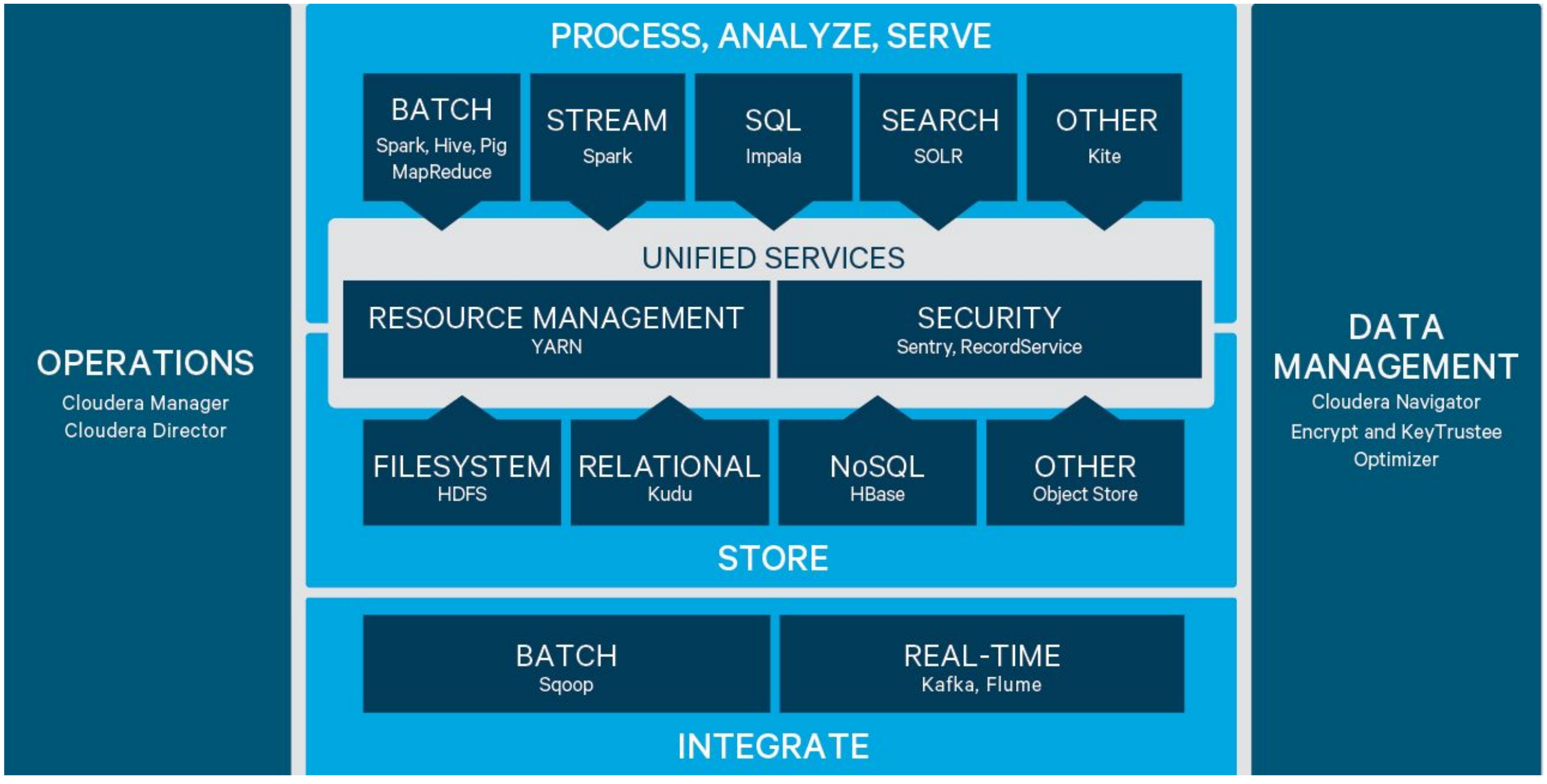
Hadoop Distributions

cloudera

MAPR®

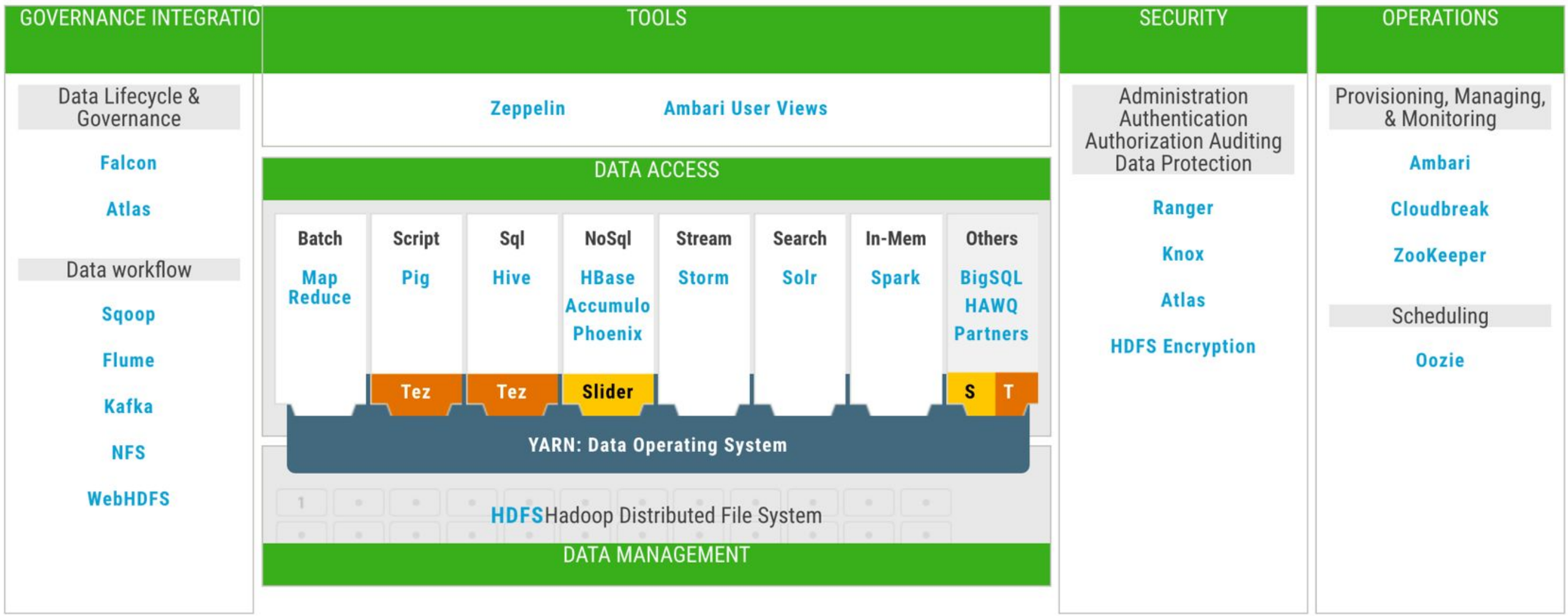
DICE ANALYTICS

Cloudera



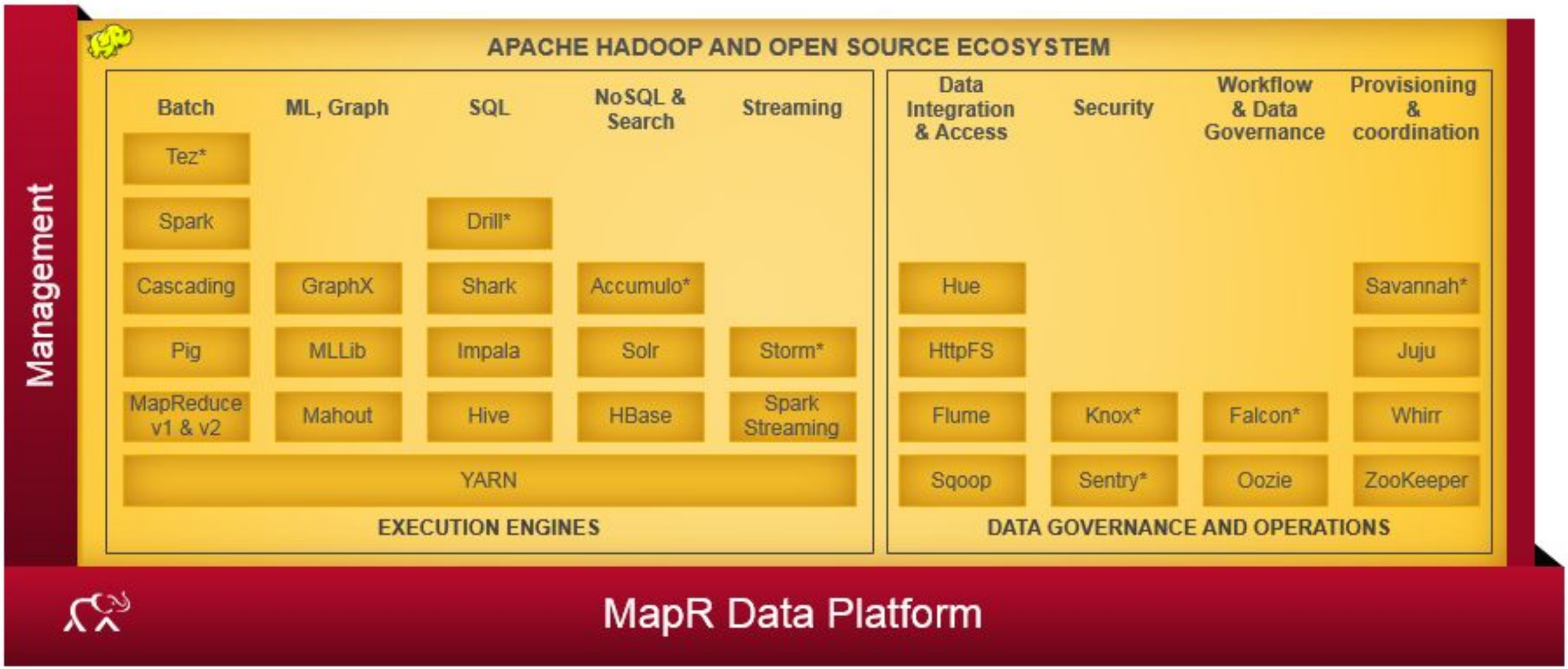
DICE ANALYTICS

Hortonworks Data Platform (HDP)



DICE ANALYTICS

MapR



Cloudera vs Hortonworks vs MapR (1)

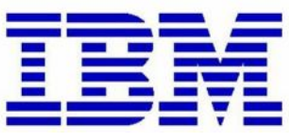
Hadoop Distribution	Advantages	Disadvantages
Cloudera Distribution for Hadoop (CDH)	CDH has a user friendly interface with many features and useful tools like Cloudera Impala	CDH is comparatively slower than MapR Hadoop Distribution
MapR Hadoop Distribution	It is one of the fastest hadoop distribution with multi node direct access.	MapR does not have a good interface console as Cloudera
Hortonworks Data Platform (HDP)	It is the only Hadoop Distribution that supports Windows platform.	The Ambari Management interface on HDP is just a basic one and does not have many rich features.

https://www.dezyre.com/article/cloudera-vs-hortonworks-vs-mapr-hadoop-distribution-comparison-/190

Cloudera vs Hortonworks vs MapR (2)

	Hortonworks	Cloudera	MapR
Dependability			
High Availability	Single failure recovery	Single failure recovery	Self healing across multiple failures
MapReduce HA	Restart jobs	Restart jobs	Continuous without restart
Upgrading	Planned downtime	Rolling upgrades	Rolling upgrades
Replication	Data	Data	Data + metadata
Snapshots	Consistent only for closed files	Consistent only for closed files	Point-in-time consistency for all files and tables
Disaster Recovery	No	File copy scheduling (BDR)	Mirroring

Companies using Hadoop



HDP Sandbox Tour

https://www.youtube.com/watch?v=7sxqHgBdxB8

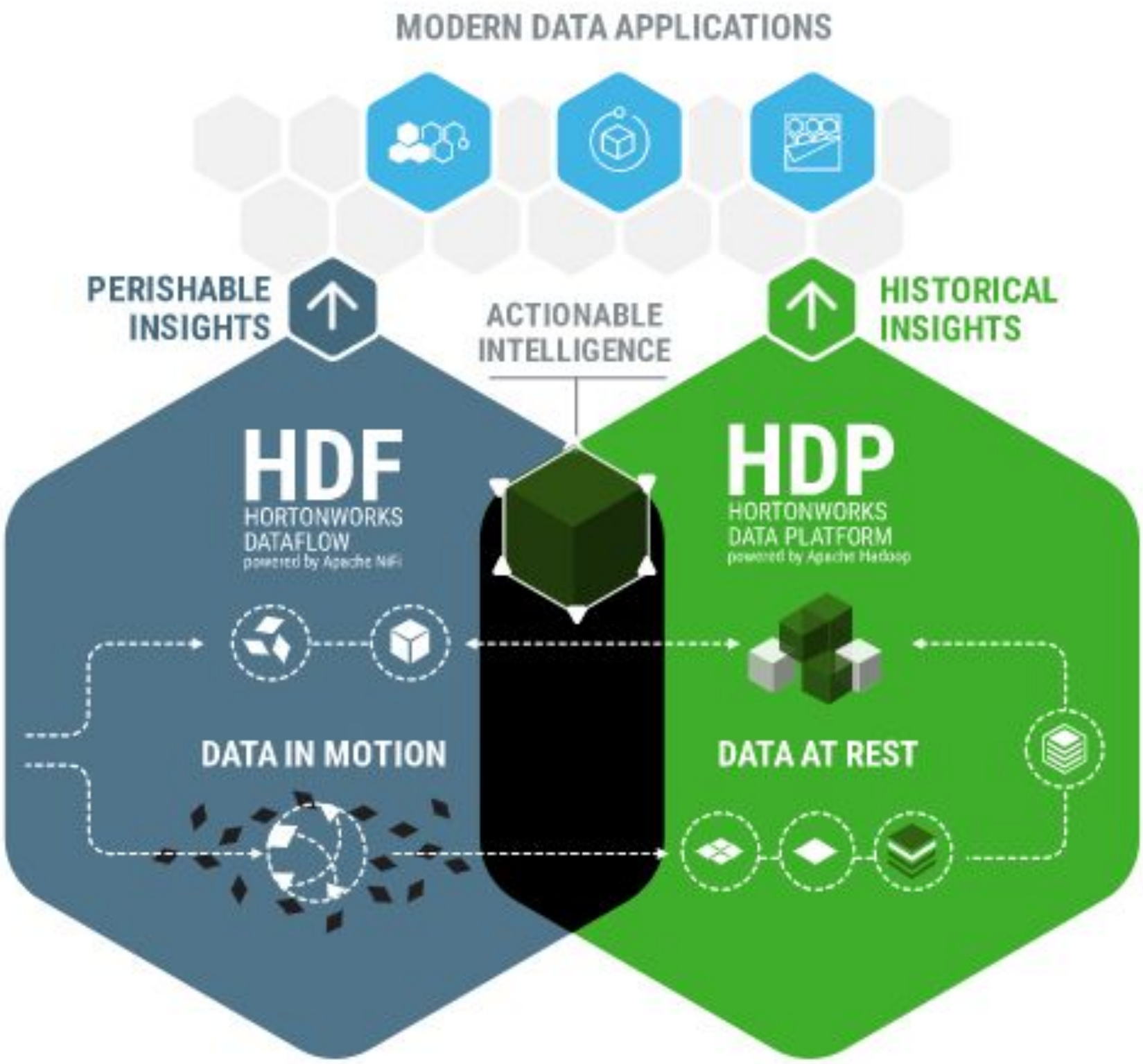
DICE ANALYTICS

HDP Sandbox Tour

https://www.youtube.com/watch?v=7sxqHgBdxB8

https://hortonworks.com/tutorial/learning-the-ropes-of-the-hortonworks-sandbox/

DICE ANALYTICS



Hortonworks Two Major Products

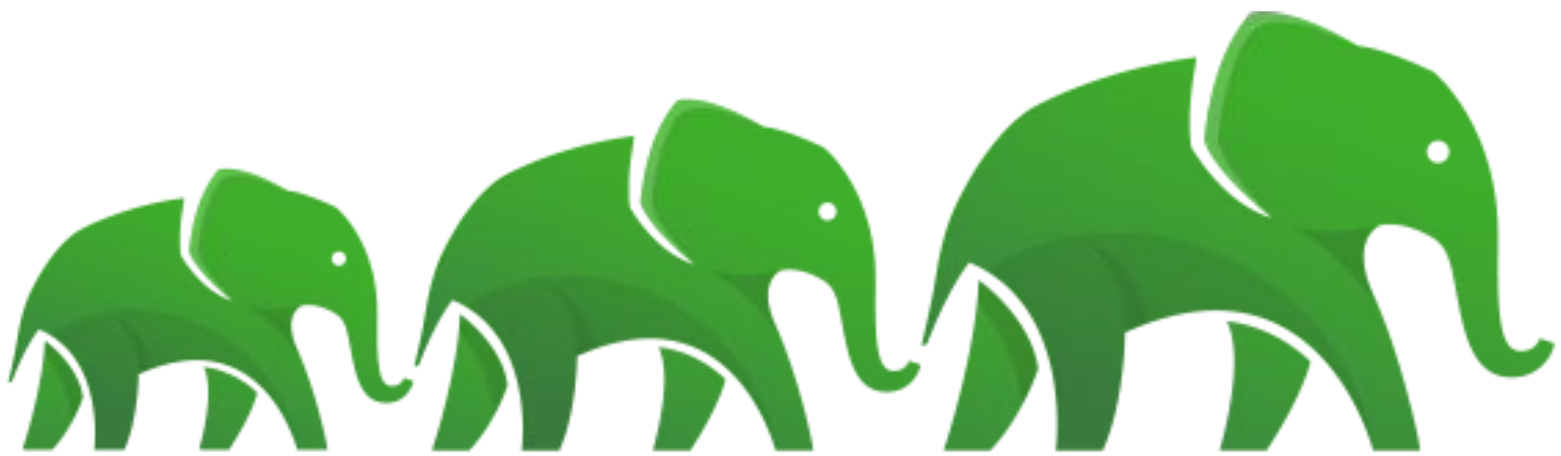
DICE ANALYTICS

Supporting Videos

https://www.youtube.com/watch?v=l6oKriR-RjM

https://www.youtube.com/watch?v=bIY3LUZ7i8Y

DICE ANALYTICS



HORTONWORKS®

https://docs.hortonworks.com/HDPDocuments/Ambari-2.6.0.0/bk_ambari-administration/bk_ambari-administration.pdf

Hadoop Distributed File System

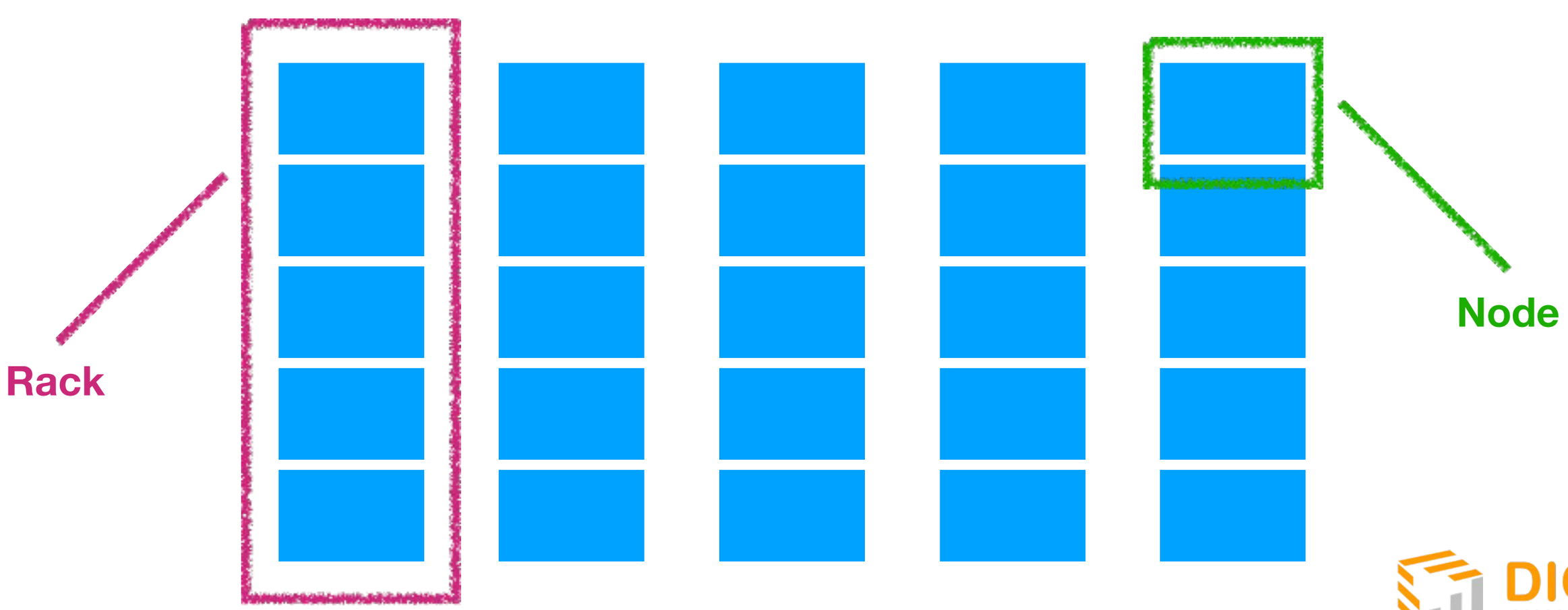
http://www.informit.com/articles/article.aspx?p=2460260

https://www.youtube.com/watch?v=4Gfl0WuONMY

DICE ANALYTICS

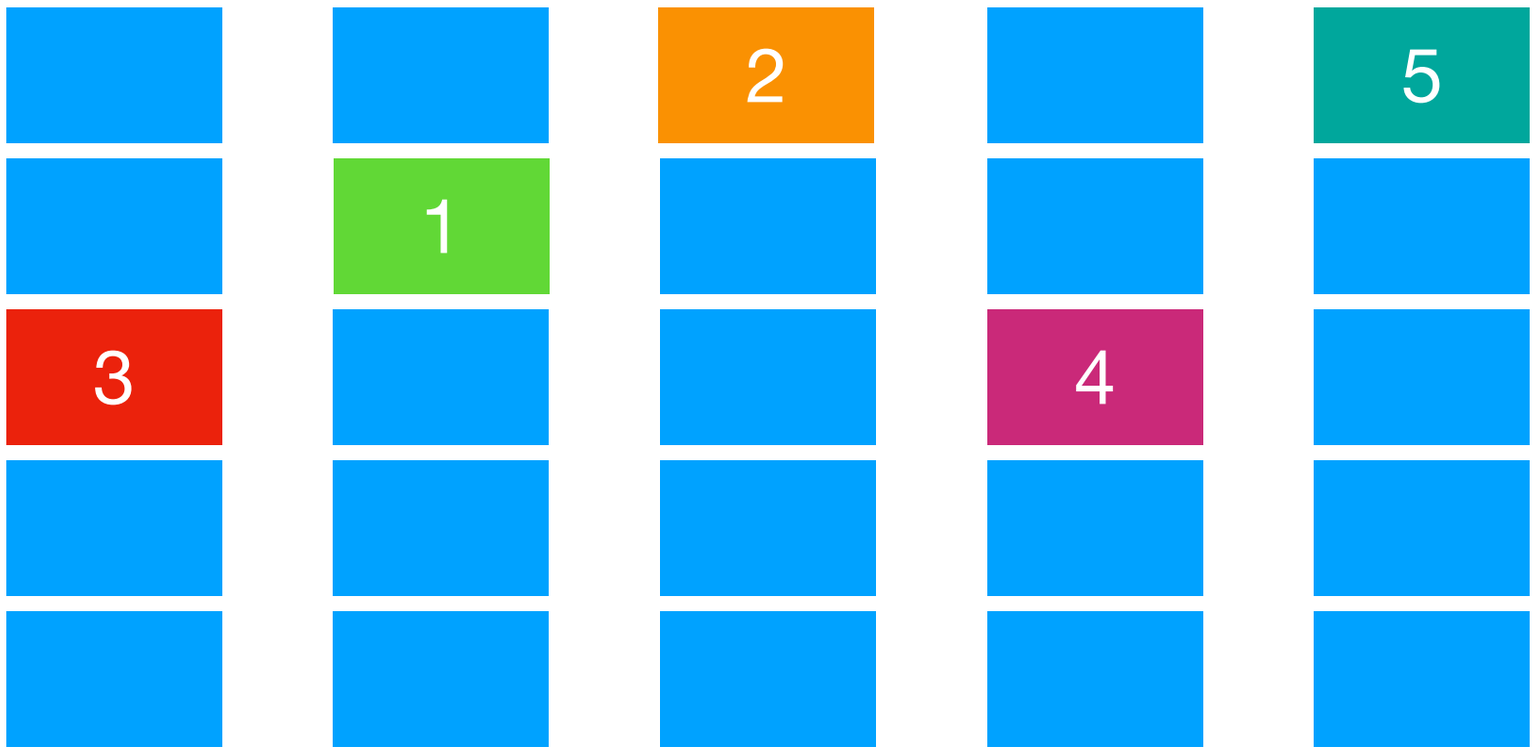
Distributed File System

- We said Big Data has a lot of Volume. Is it then possible to store Big Data in a single system?
- It’s not. We need to distribute Big Data into multiple systems, for which we need a **Distributed File System (DFS)**.



Distributed File System

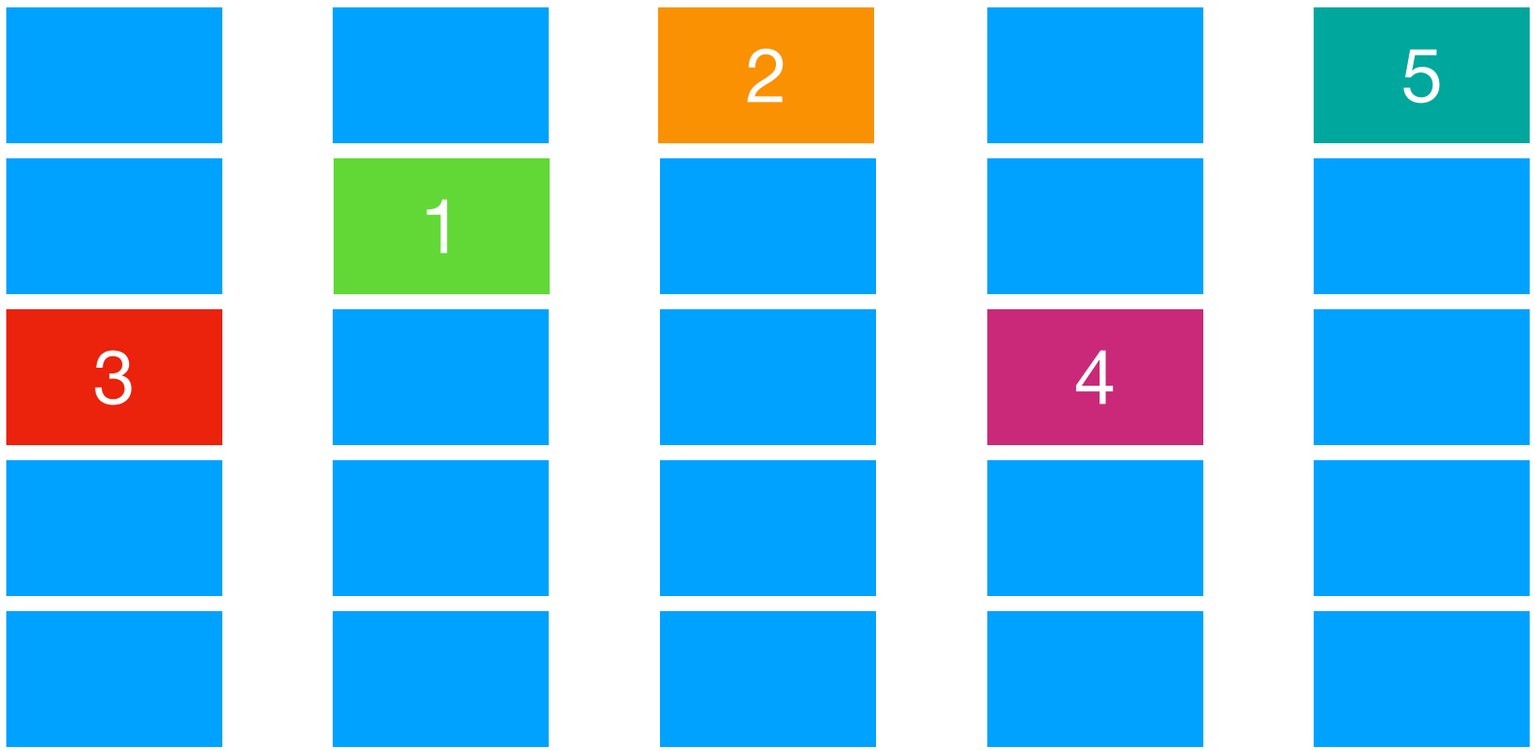
- To achieve parallelization, we distribute data across nodes and also move computation to each node.



DICE ANALYTICS

Distributed File System

- To achieve parallelization, we distribute data across nodes and also move computation to each node.



DICE ANALYTICS

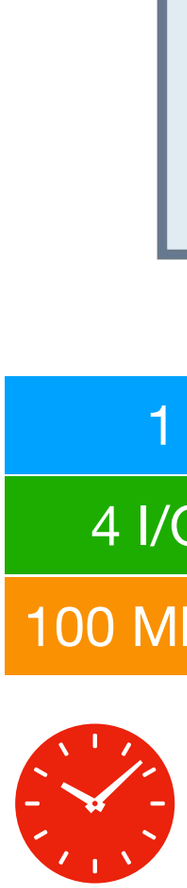
Distributed File System

Reading 1 TB Data

1 Machine

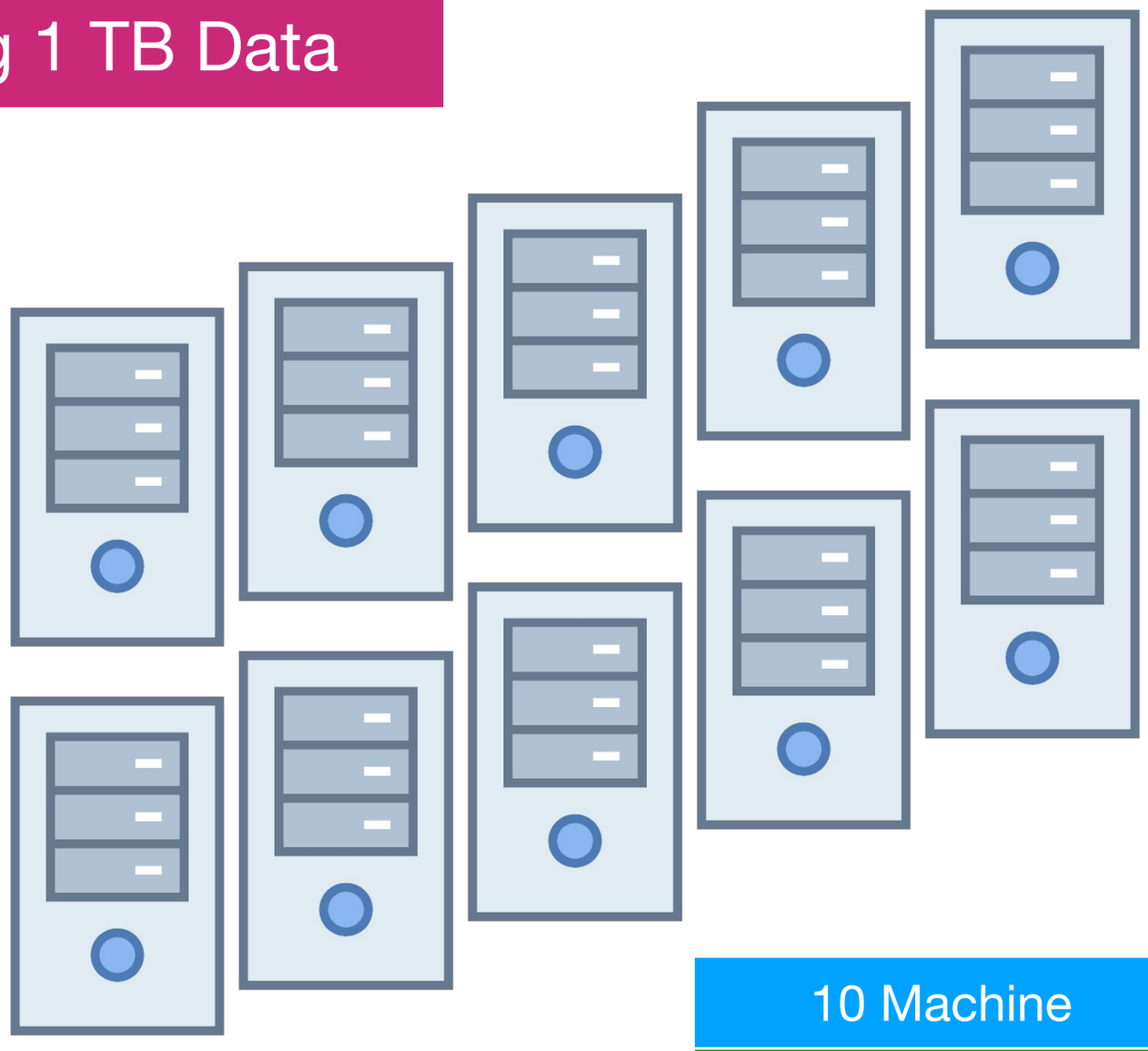
4 I/O Channels

100 Mbps / Channel



43 Minutes

10 Times Faster!



10 Machine

4 I/O Channels

100 Mbps / Channel

4.3 Minutes

Distributed Computing

- It can be defined as the use of a distributed system to solve a **single large problem** by breaking it down into **several tasks** where each task is computed in the individual computers of the distributed system.
- All the computers connected in a network communicate with each other to attain a common goal by making use of their own local memory.
- **Hadoop makes use of Distributed Computing.**

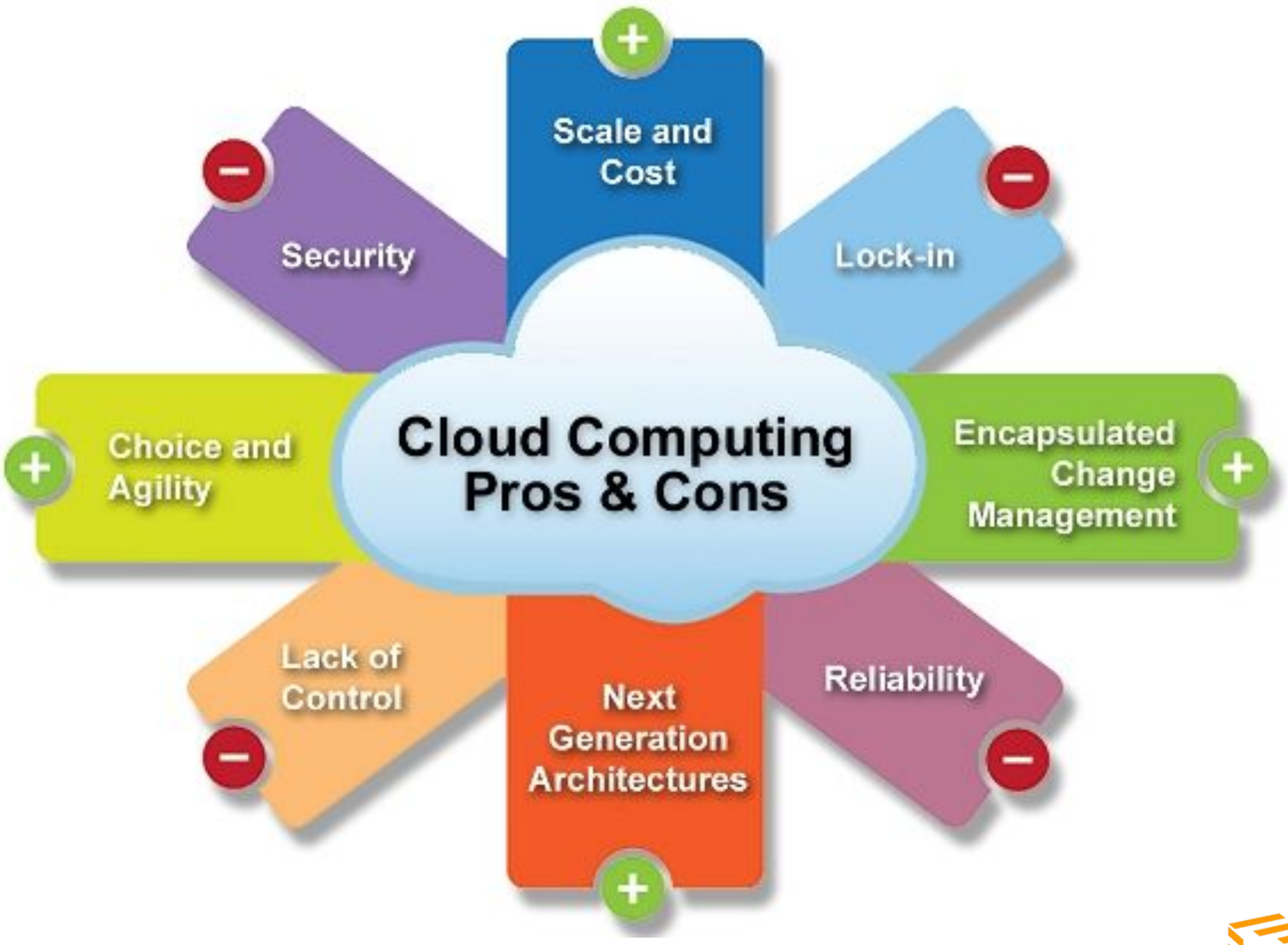
Cloud Computing

- **Cloud Computing** is an information technology paradigm, a model for enabling ubiquitous access to shared pools of configurable resources (such as **computer** networks, servers, storage, applications and services).

- These can be rapidly provisioned with minimal management effort, often over the Internet.

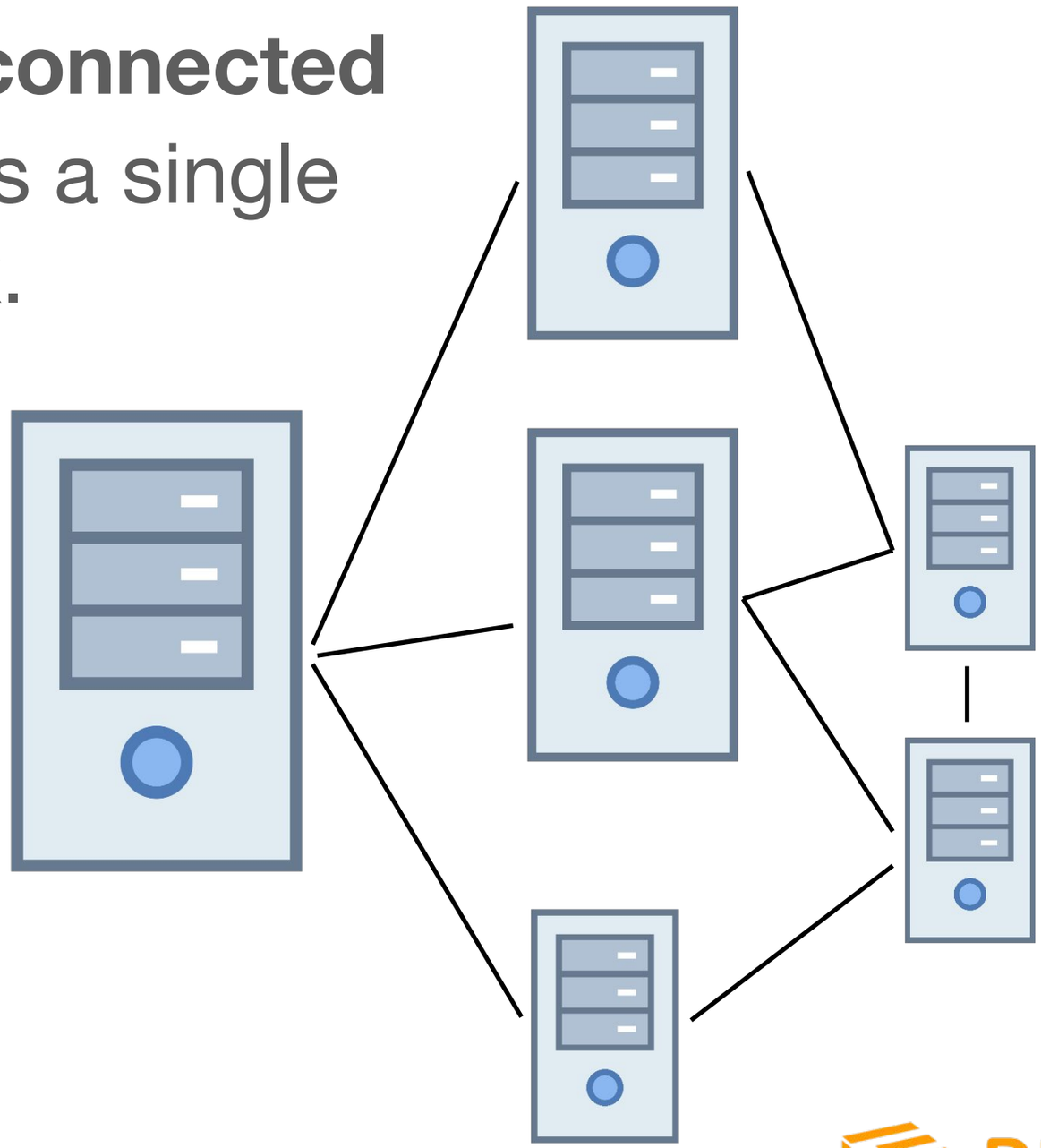
DICE ANALYTICS

Cloud Computing



Computer Cluster

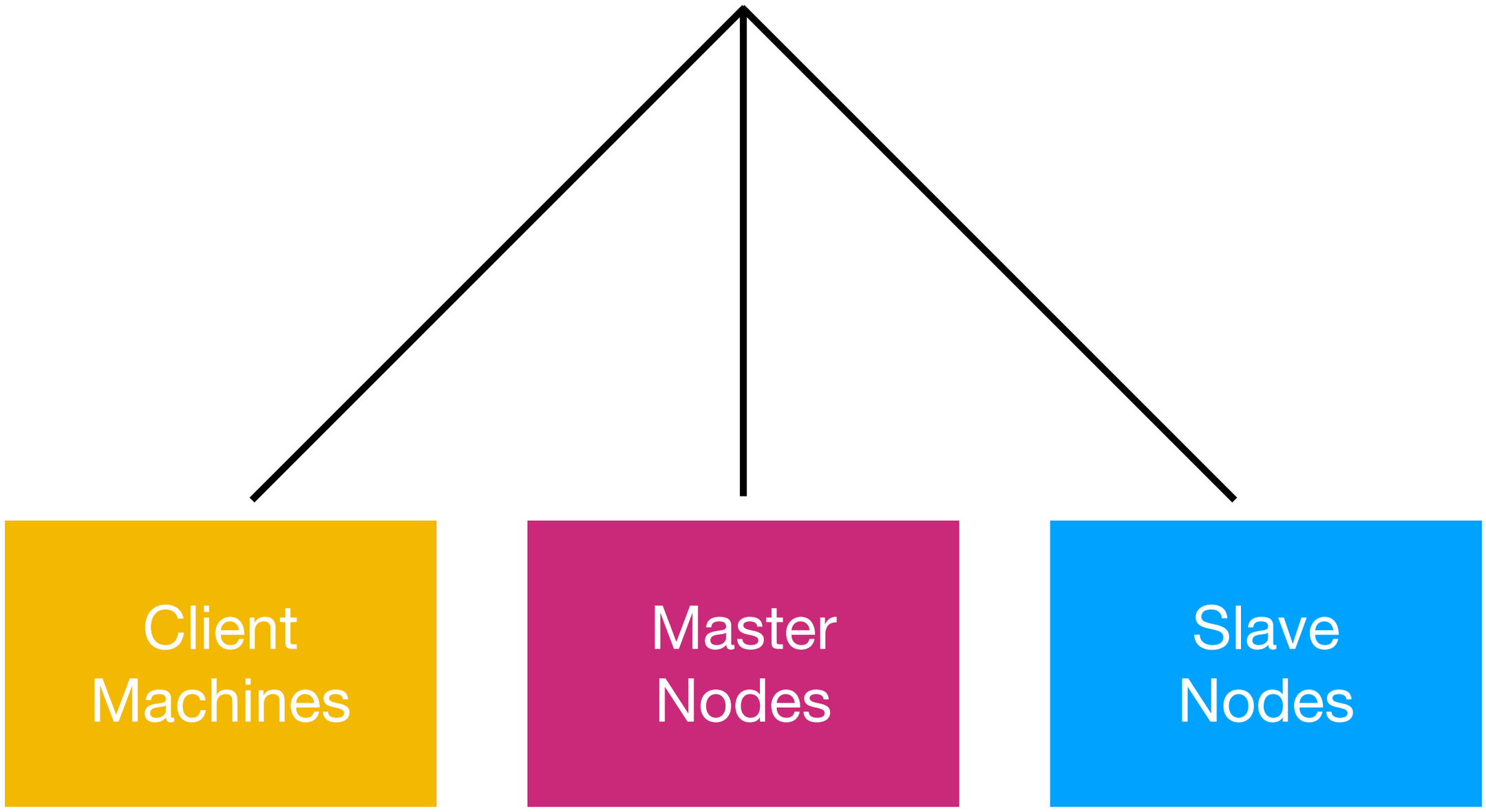
- A group of **network connected** computers working as a single unit to perform a task.



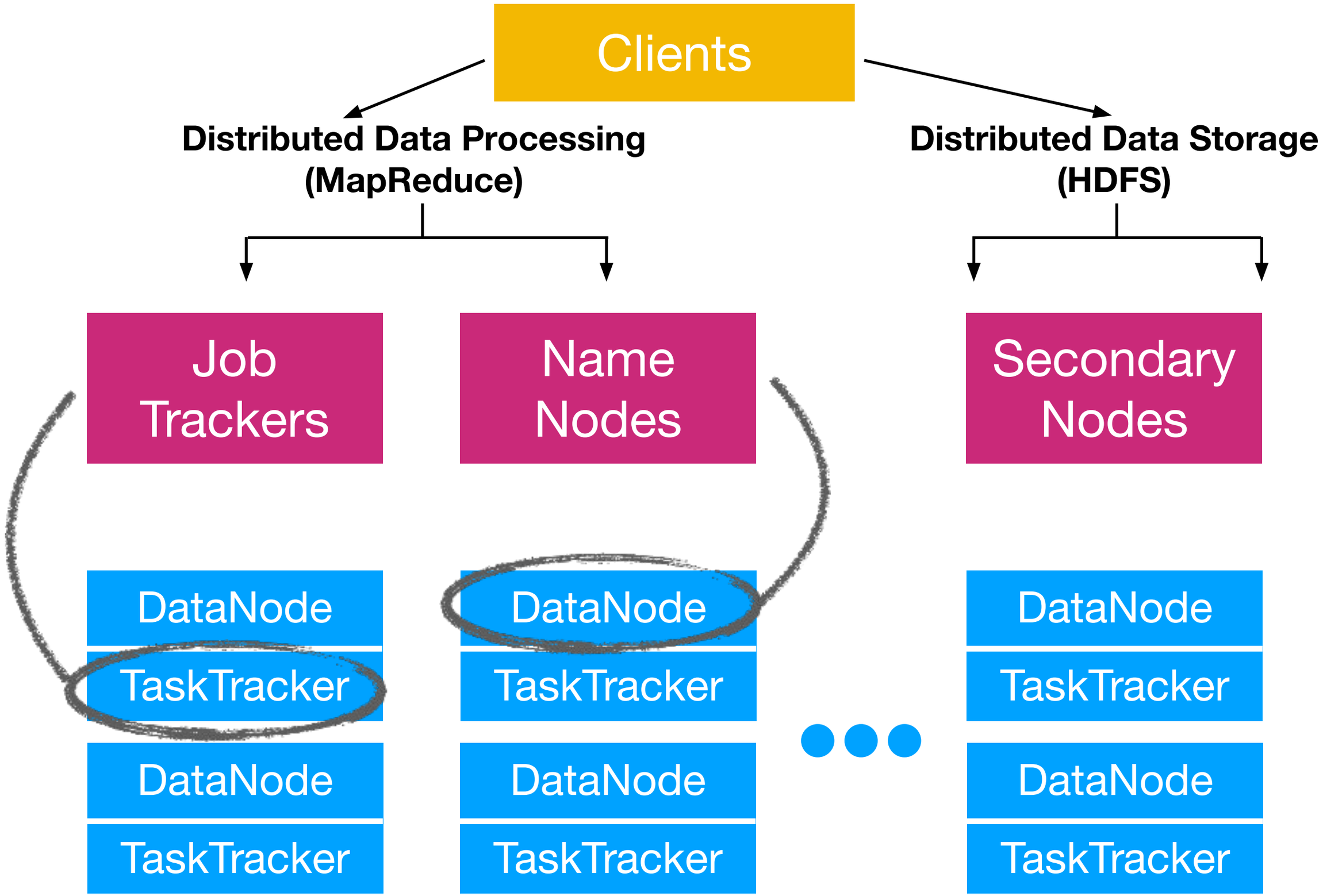
Hadoop Cluster

DICE ANALYTICS

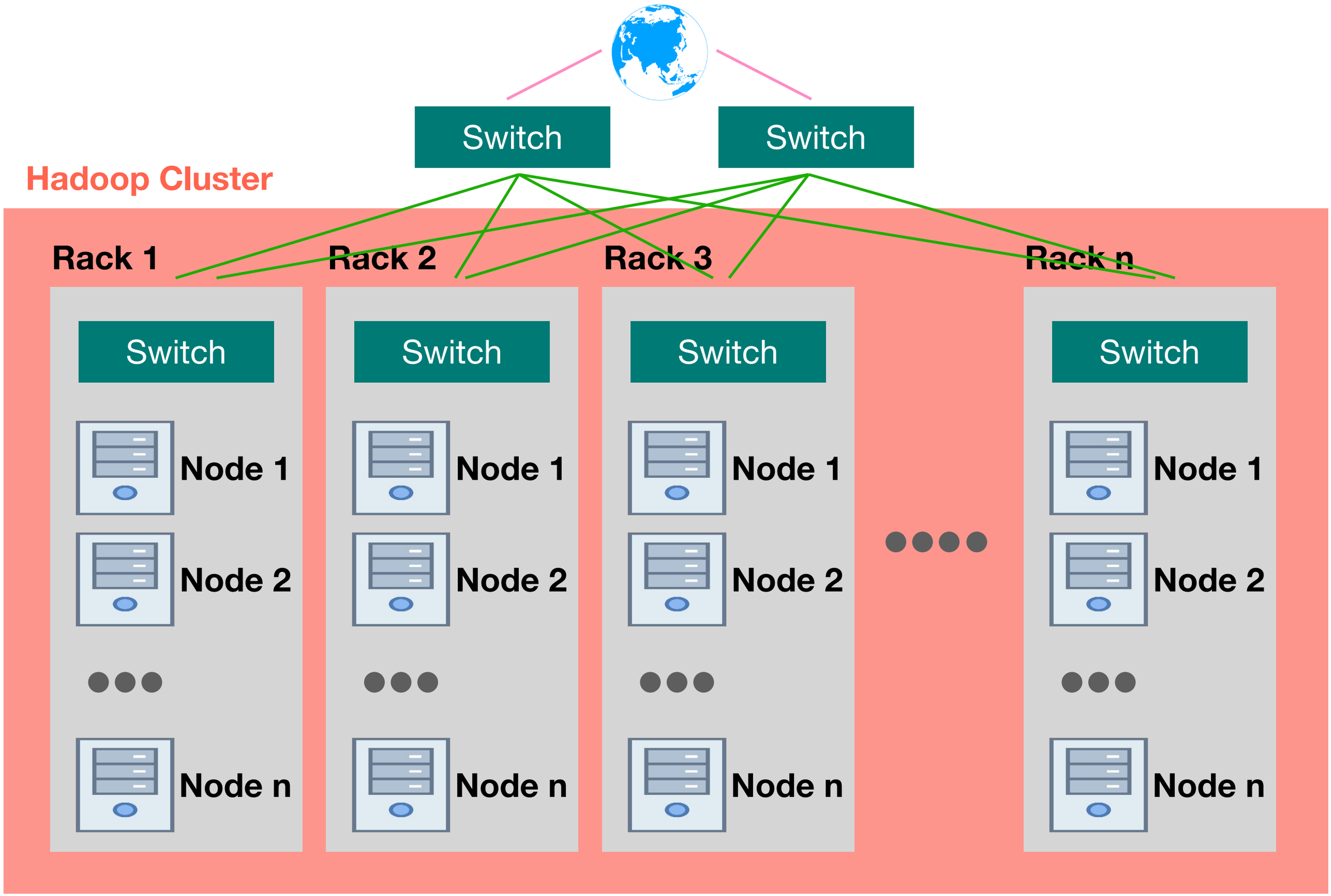
Hadoop Server Roles



Hadoop Server Roles



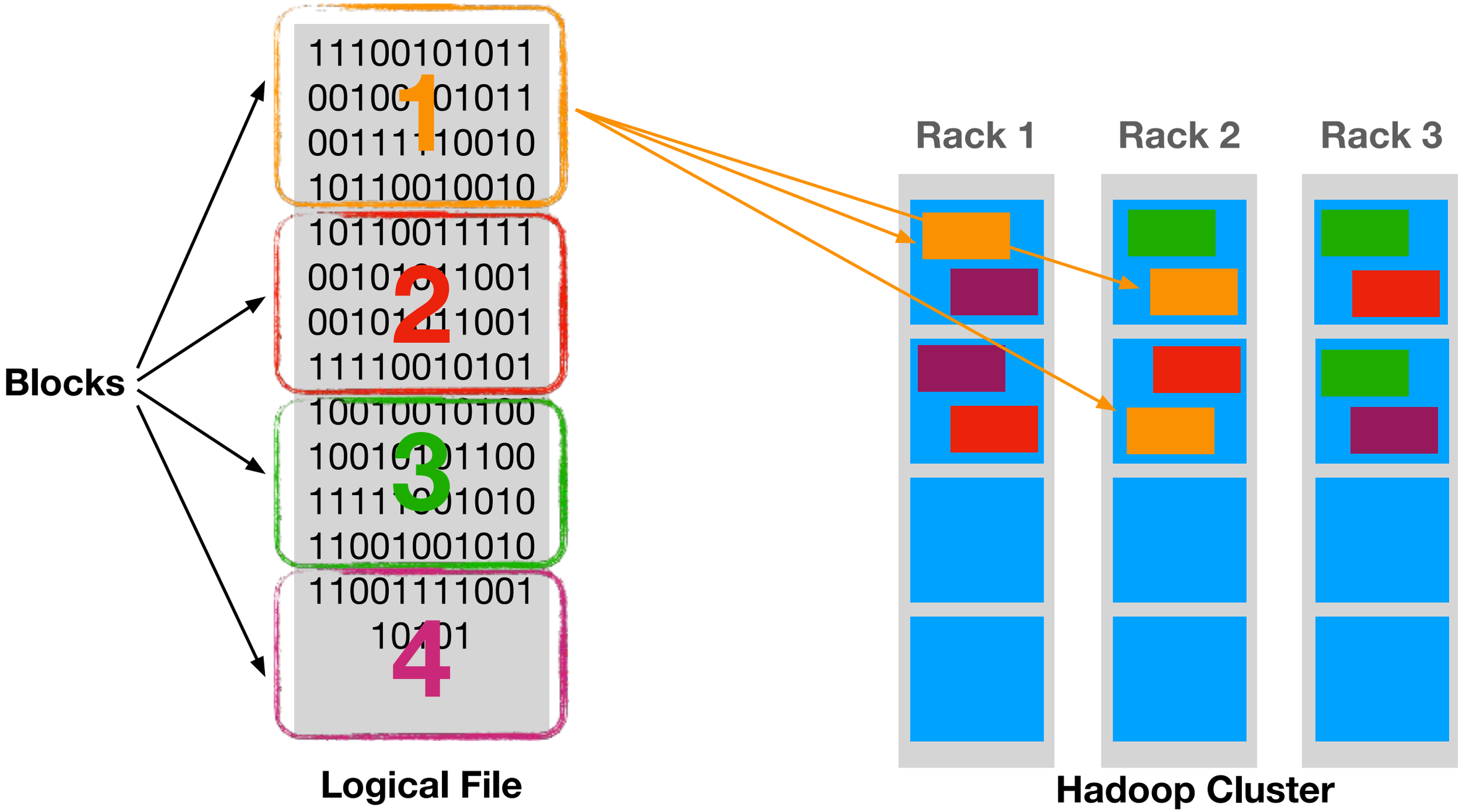
Typical Hadoop Cluster



HDFS Overview

- A fault-tolerant distributed file system for big large files.
- Write Once, Read Many Times (WORM)
- Divide files into big blocks and distribute across the cluster.
- Store multiple replicas of each block for reliability.
- Programs can ask “Where do the pieces of my file live.

HDFS Overview



DICE ANALYTICS

HDFS Overview

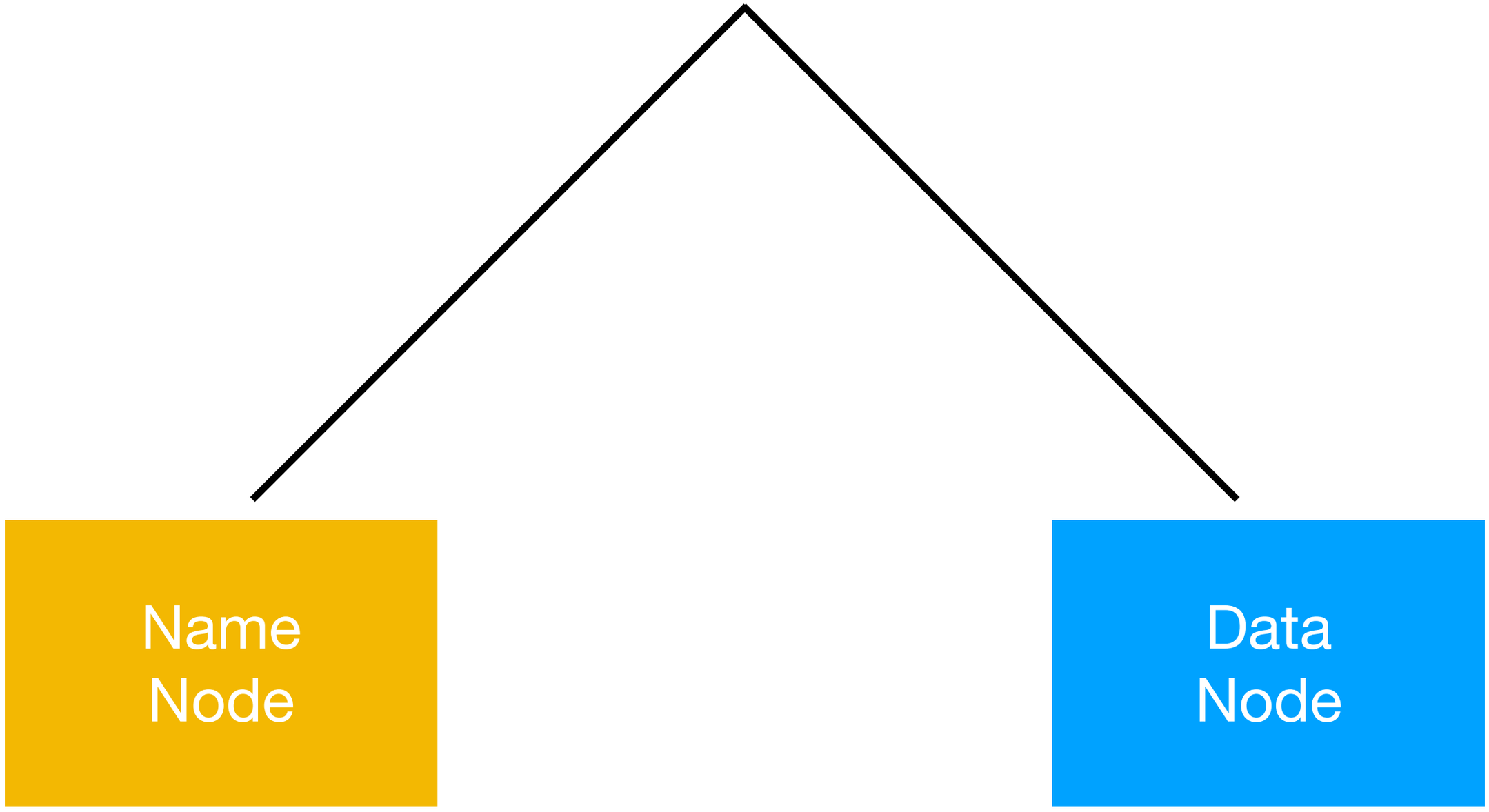
- It looks & acts just like a file system.

```
hdfs dfs -command[args]
```

- A few of almost 30 HDFS Commands:
  - -cat :display file content (uncompressed).
  - -text :just like cat but works on uncompressed files.
  - -ls, -lsr :
  - -chgrp, -chmod, -chown :changes file permissions.
  - -put, -get, -copyFromLocal, -copyToLocal :copies files from the local file system to HDFS and vice versa.
  - -rm, -rmr :
  - -ls -R :list file/directories.
  - -mv, -moveFromLocal, -moveToLocal :moves files.
  - -stat :statistical info for any given file.
  - -tail :

DICE ANALYTICS

HDFS Components



NameNode

- It acts as HDFS Master Component.
- It determines and maintains how chunks of data are distributed and replicated across the DataNodes.
- It maintains critical HDFS information/system state information.
- To enhance HDFS performance, it maintains and serves this information from memory.
- Therefore it is critical to ensure NameNode always has sufficient memory.
- If NameNode fails, HDFS fails.

NameNode

- Overview of information stored by NameNode.

Namespace

- Hierarchy
- Directory names
- Files names

Metadata

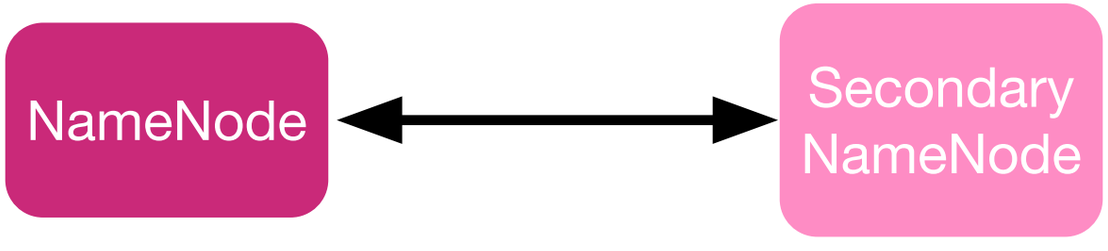
- Permissions & Ownership
- ACLs
- Block Size & Replication levels
- Access & Modification times
- User quotas

Block Map

- Files names > Block IDs

DICE ANALYTICS

Secondary NameNode

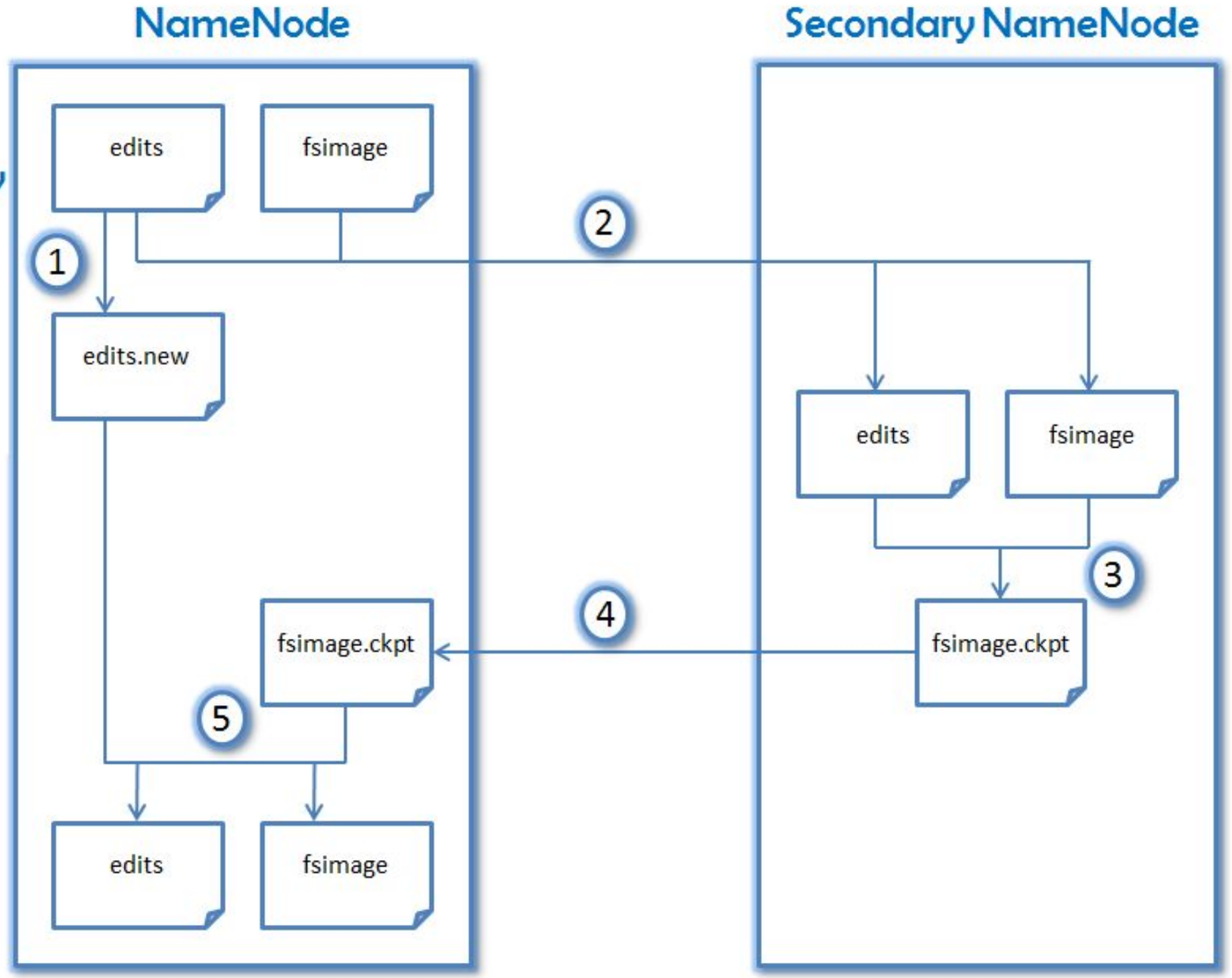


- It does housekeeping and backup of NameNode namespace and metadata.
- It is not a hot-standby for NameNode.
- It connects to NameNode every hour.
- Saved metadata can rebuild a failed NameNode.

https://blog.cloudera.com/wp-content/uploads/2010/03/HDFS_Reliability.pdf

Secondary NameNode

1. Roll edits file
2. Retrieve fsimage & edits from primary
3. Merge
4. Transfer checkpoint to primary
5. Roll fsimage.ckpt and edits.new



http://www.facebook.com/hadoopers

http://myhadoopadmin.blogspot.com/2014/11/hadoop-administration-check-point-node.html

http://pramodgampa.blogspot.com/2013/06/the-building-blocks-of-hadoop.html

Why 3 Name Nodes in Hadoop ?

https://data-flair.training/blogs/hadoop-hdfs-namenode-high-availability/

NameNode High Availability

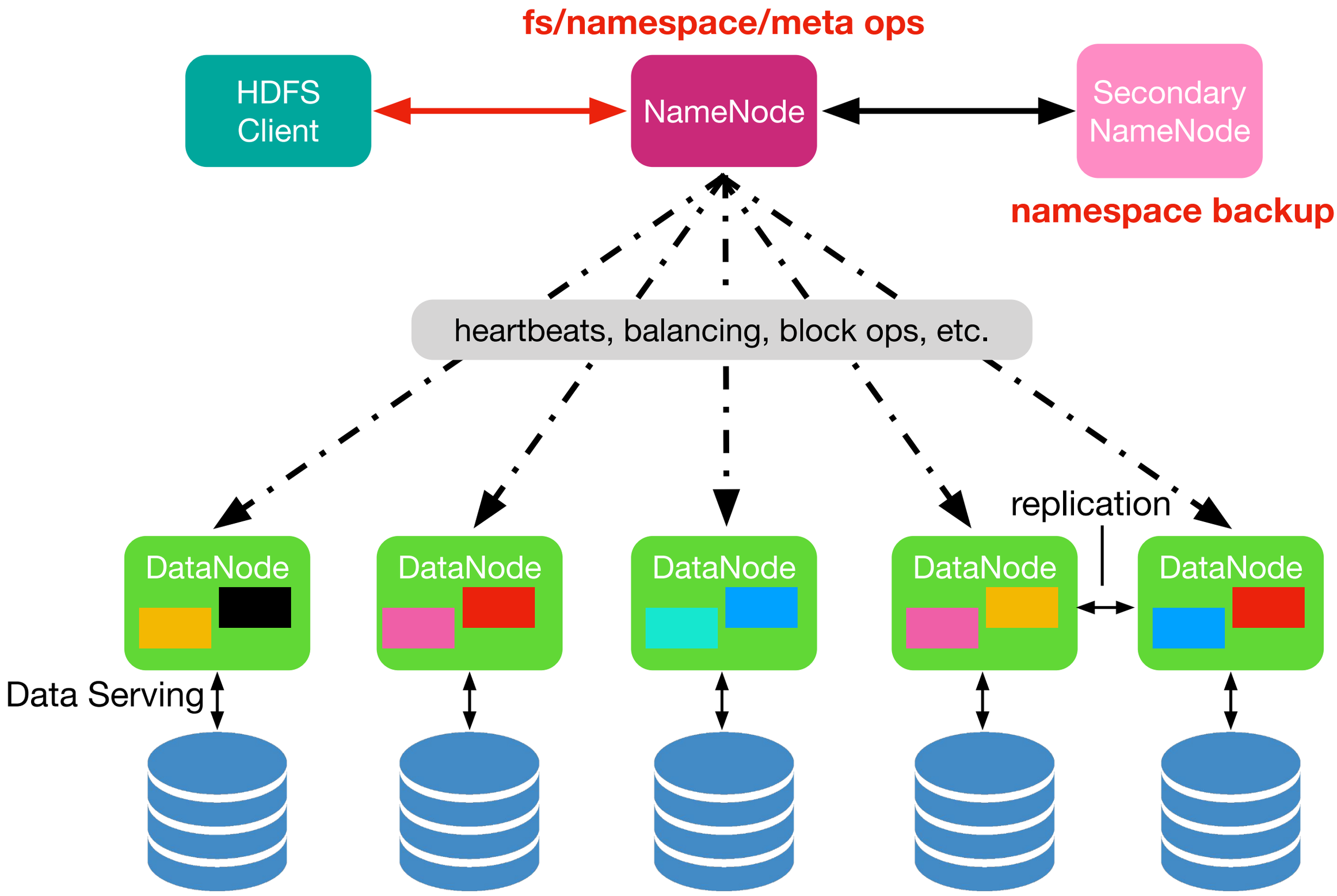
- HDFS NameNode is Single Point of Failure.
- NameNode HA:
  - Uses a redundant NameNode.
  - Enables fast fail-over in response to NameNode failure.
  - Is configured by Ambari.
  - Is configured in an active/standby configuration.
  - Permits administrator-initiated failover for maintenance.

DICE ANALYTICS

DataNode

- It acts as HDFS Slave Component.
- It is the only place where chunks of data are actually stored.
- Other than storing data, it is also responsible for replicating data.
- It keeps on sending its heartbeat to NameNode to tell about it’s availability.
- Every 10th heartbeat is a Block Report.
- DataNodes are heterogeneous: supports different types of storages: Disks, SSDs, Memory.

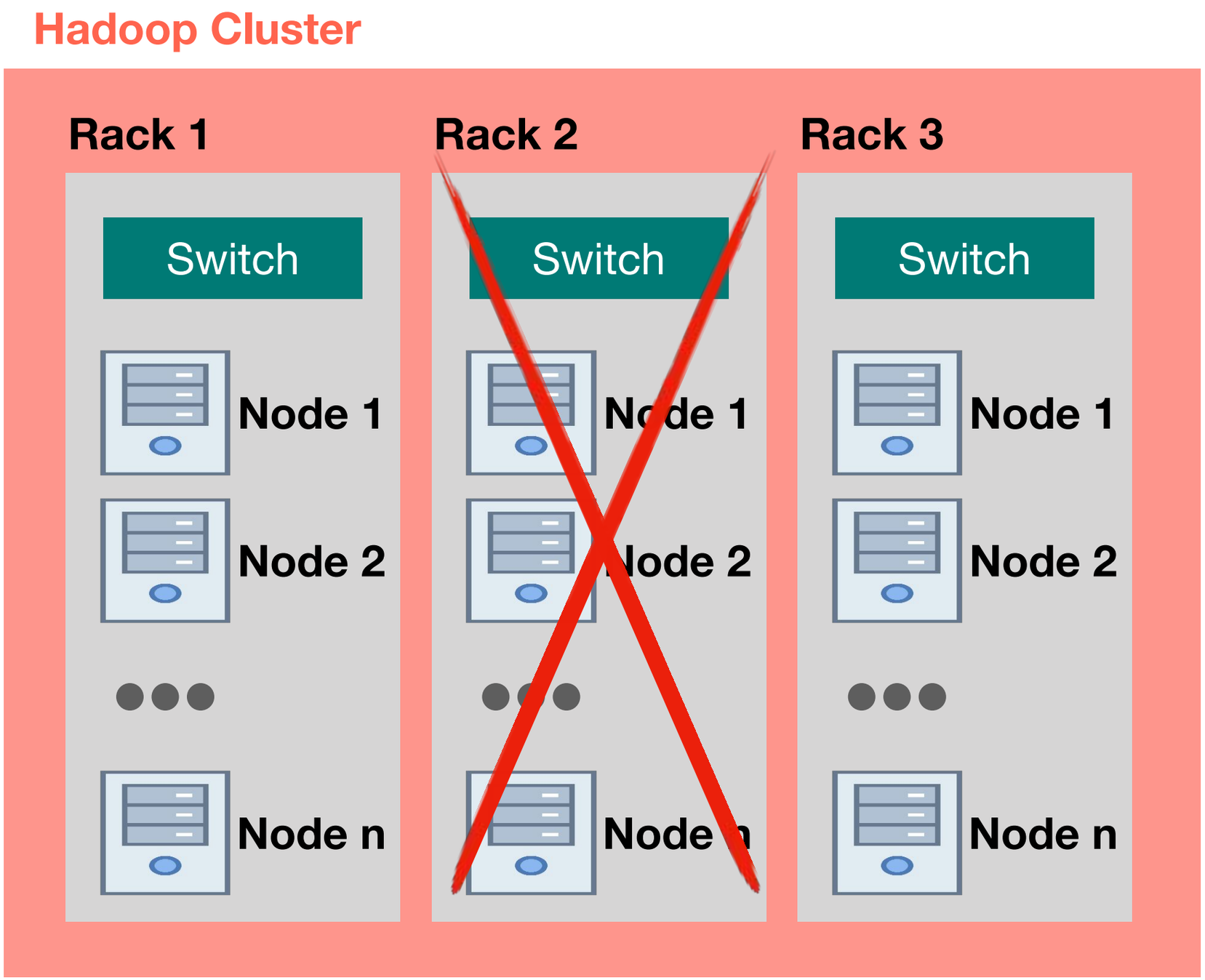
HDFS Architecture



HDFS Rack Awareness

- Never loose all data if entire rack fails. How?

- Store replicas on multiple racks.
- Keep bulky flows in-rack.
- There is higher bandwidth and lower latency in-rack.



In Rack Awareness

First Replica will be written in First Rack

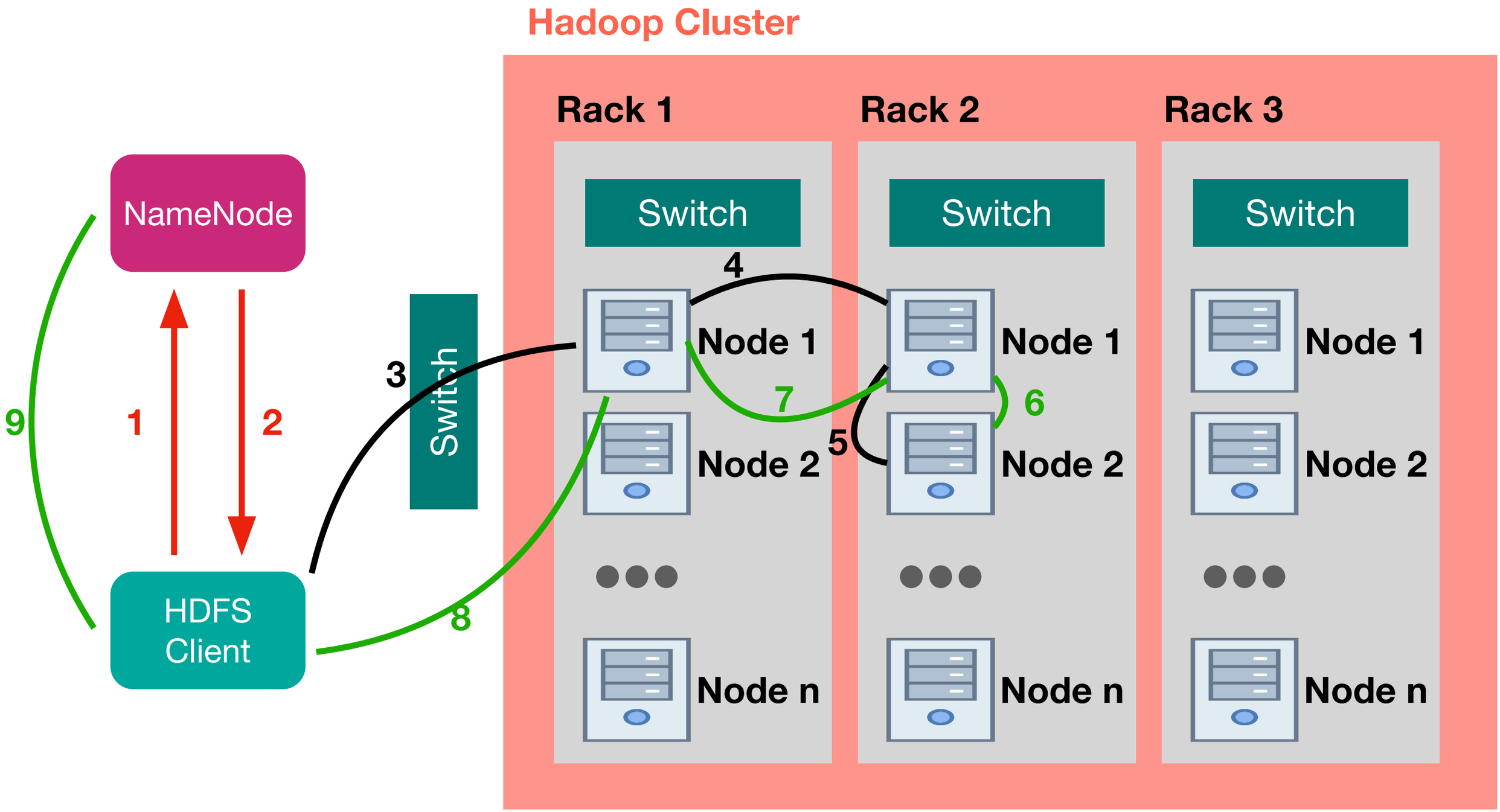
Second & Third Replica will be written in Second Rack

To Save Network Bandwidth & Reduce the Write Time

DICE ANALYTICS

HDFS Write Pipeline

Note: All DataNodes are in constant communication with NameNode so no arrows are drawn for that.

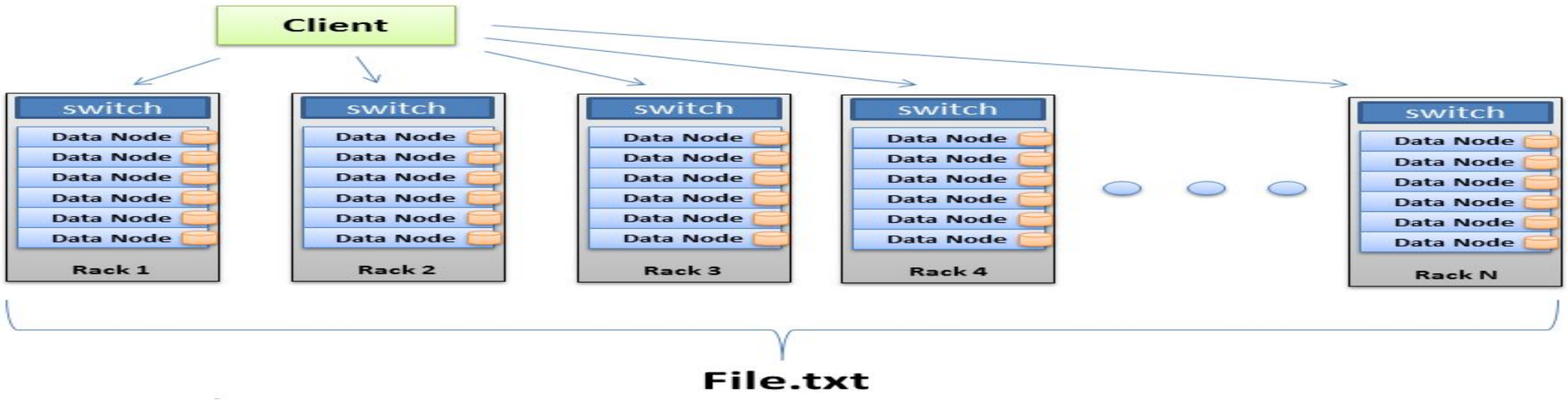


HDFS Write Pipeline

- HDFS manages writing of file block by block.
- Many files are being written in parallel to save time.
- All communication happens through TCP connections so high bandwidth is required.
- And HDFS Write Operation has to major cycles: Acknowledgements and Data Transfer.
- Starting node for each block isn’t necessarily same.
- NameNode updates metadata with the help of Block Reports sent by DataNodes.

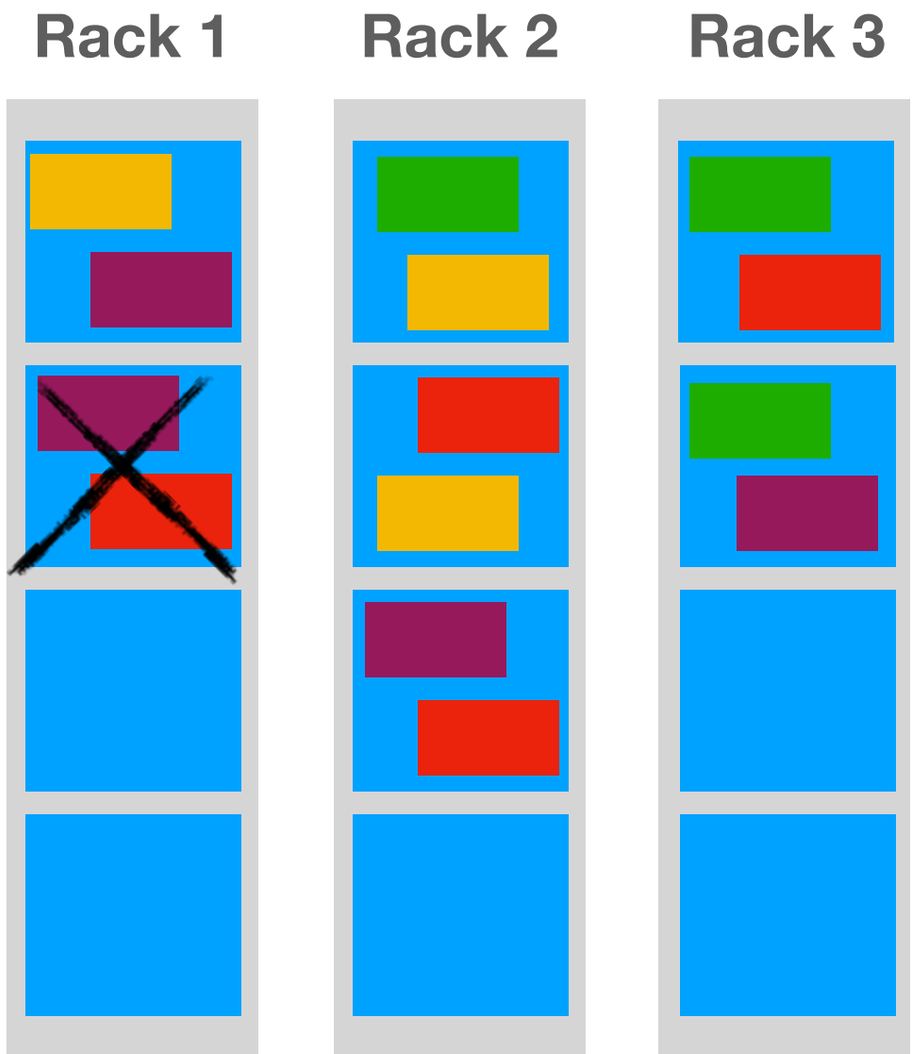
Spanning HDFS Cluster

- Keep Block Size small.
- Small Block Size means more Blocks for a file.
- More Blocks means file is spread on more machines.
- More CPU cores and disk drives that have a block of file mean more parallel processing power and faster results.
- This is why we build large wide clusters.



Re-replicating in HDFS Cluster

- NameNode automatically takes care of recovering missing and corrupted blocks.
- Missing heartbeats signify lost Nodes.
- NameNode consults metadata, finds affected data.
- NameNode consults Rack Awareness script.
- NameNode tells a DataNode to re-replicate itself to a specified DataNode that is available.

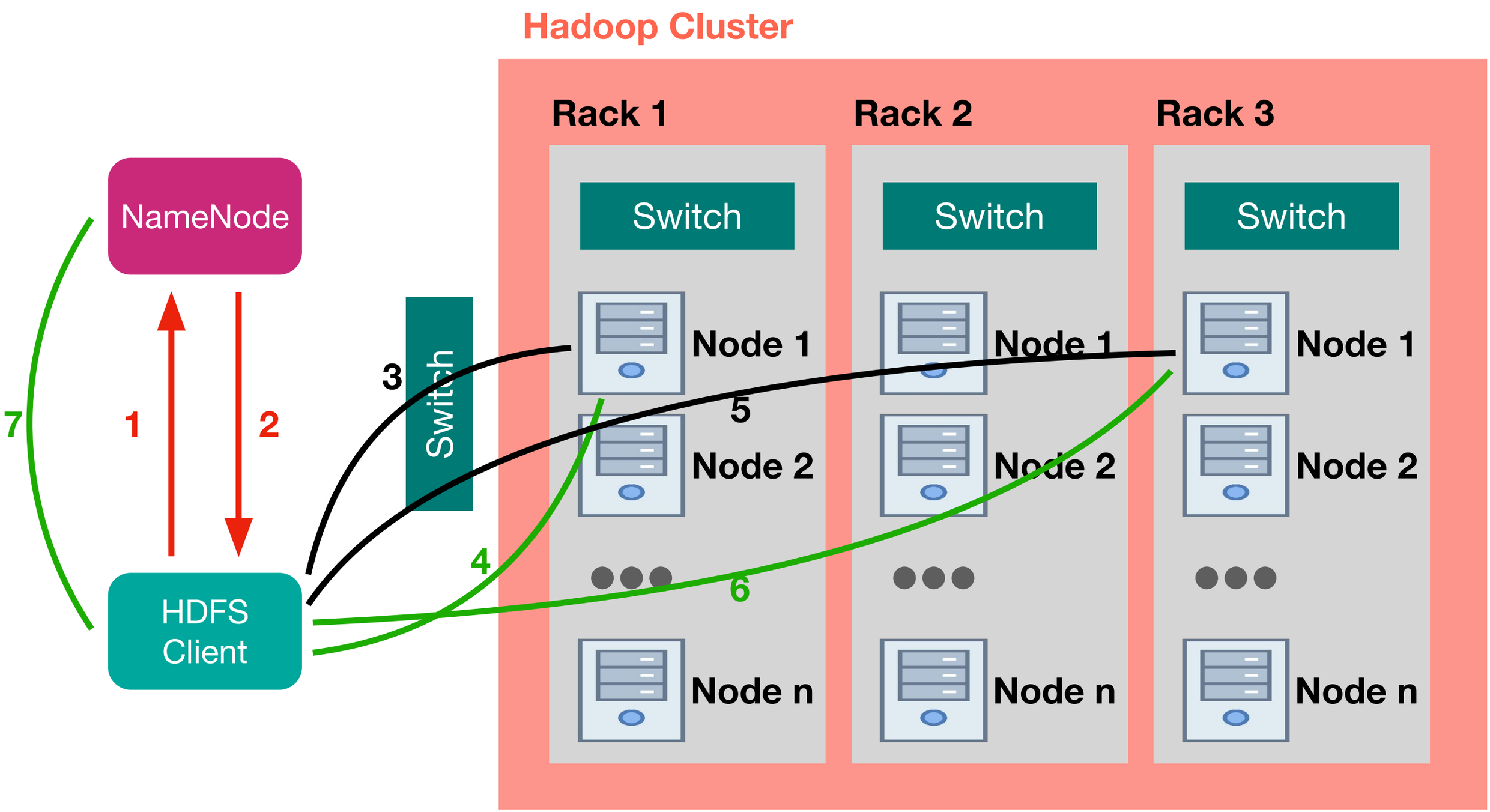


Hadoop Cluster

DICE ANALYTICS

HDFS Read

Note: All DataNodes are in constant communication with NameNode so no arrows are drawn for that.



HDFS Read

- HDFS manages reading of file block by block.
- Many files are being read in parallel to save time.
- All communication happens through TCP connections so high bandwidth is required.
- NameNode updates metadata with the help of Block Reports sent by DataNodes.
- In case a DataNode needs a block that it does not have, the NameNode provides rack local DataNodes first to leverage in-rack bandwidth.

Replication Factor

Block Size

Two Major Configurations of HDFS

DICE ANALYTICS

Block Size Should Be Configures With a Multiple of 512 KB

HDFS

is made to

Write Once

&

Read Many Times

(WORM)

DICE ANALYTICS

Configure Block Size For Particular File Movement

HDFS Tutorials

https://hortonworks.com/hadoop-tutorial/using-commandline-manage-files-hdfs/
https://hadoop.apache.org/docs/r2.4.1/hadoop-project-dist/hadoop-common/FileSystemShell.htm
l
http://hadooptutorial.info/hdfs-file-system-commands/
https://intellipaat.com/tutorial/hadoop-tutorial/hdfs-overview/

DICE ANALYTICS

Remove a File From HDFS

Deleted Files Moves to Trash Directory in HDFS

DICE ANALYTICS