

TEKLRN

HADOOP LEVEL 13

1. Hadoop Node Failures
 - Types of Failures
 - Hadoop Node Failure
 - The DFS Admin Command
 - Cluster setup, Test Files and Block Sizes
 - Fault Tolerance and Elastic MapReduce
2. Killing a DataNode Process
 - NameNode and DataNode Communication
3. Replication Factor
4. Missing Blocks
 - When Data may be lost
 - Block Corruption
5. Killing a TaskTracker Process
 - Comparing the DataNode and TaskTracker failures
 - Permanent Failures
 - Killing the Cluster Masters
6. Killing the JobTracker
 - Starting a replacement JobTracker
7. Killing NameNode Process
 - Starting a Replacement NameNode
 - The role of the NameNode in more details
 - File systems, Files, Blocks and Nodes
 - fsImage
 - DataNode Startup
 - Safe Mode
 - Secondary NameNode
 - Action to take when NameNode process has a critical Failure
 - BackupNode/ CheckPointNode and NameNode HA
 - Hardware Failure
 - Host Failure
 - Host Corruption
 - The Risk of correlated Failures

8. Causing Task Failure

- Hadoops Handling of slow running Tasks
- Speculative Execution
- Hadoop's handling of failing tasks

9. Task Failure due to Data

- Handling Dirty Data through Code
- Using Hadoops Skip Mode

10. Handling Dirty data using Skip Mode