

# CS-425 MP2 Report Group No: 44

Lin Lyu(linlyu2), Ching-Hua Yu(cyu17)

## I. Design

**Master Election:** machines in a ring maintain a master consensus by attaching the master is into the heartbeat. Once a node failed and it is master, machines will elect the node with the smallest id in the ring to be the new master and this message will be attached to heartbeat again to ensure consensus.

**Replication Strategy:** Every file has 4 replicas in the SDFS to service less or equal to 3 failures. Once a node failed, if it is not the master, master will notice its failure according to the change of membership list. Master will lookup all replicas in this node. For each replica, master will select one node(node\_a) in the membership list and one node(node\_b) having this replica, and ask node\_b send node\_a this file. If master is failed, the new master will ask every alive machine to hand-in the sdfs file name list. If there are files which don't have enough replicas(because they have one in previous master), the new master will ask a node with the file to duplicate one on a node who don't have this file.

**PUT:** A node will send put request to master, and master will check whether the file is in the SDFS. If the file does not exist, the master will pick 4 nodes and sends id of those nodes to the node request a put. The node connects those nodes after receiving the acknowledgement from master and send the file. If the file already exists and more than 1 minute past since last put, the master will allow this node to use its replica replace original file. If the update is within 1 minute, the master will ask the machine to confirm its action within 30s. We use Quorum strategy to write and read. A put is regarded as success only if every replica of this file is updated.

**GET:** A node planning to read a file in SDFS will first ask master. The master will return a node id who has replica of this file if the file exists. After that, the node will ask that node to get the file. Since we adopt  $W = N$ , the system can read any replica of a file.

**DELETE:** The node will ask master and master will send the node list back if the file exists. After that the node will ask those nodes to delete the file.

**LS:** The node will ask master and master will send the list back if the file exists.

**STORE:** Print out SDFS replicas one machine has.

## II. How MP1 used in MP3

We logged actions to files(PUT/GET/DELETE) and node fail or join. And we can query logs from any machine to debug easily.

## III. Measurements

- 1) re-replication time and bandwidth (Bps not messages per sec) upon a failure
- 2) times to insert, read, and update, file of size 25 MB, 500 MB

