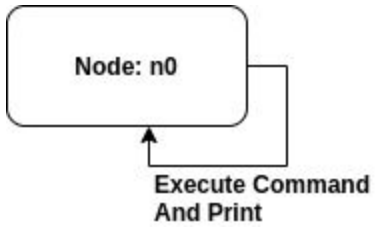
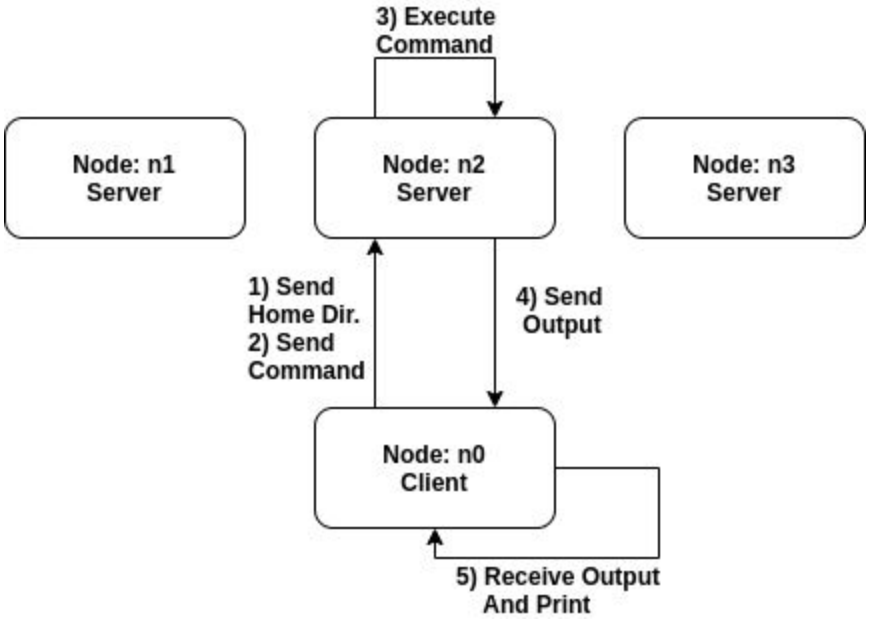
P2. Cluster Shell

Local Commands: These commands are simply executed on the local machines and their output is printed on the screen. Implemented completely as a part of client code as no interaction with the server is needed.



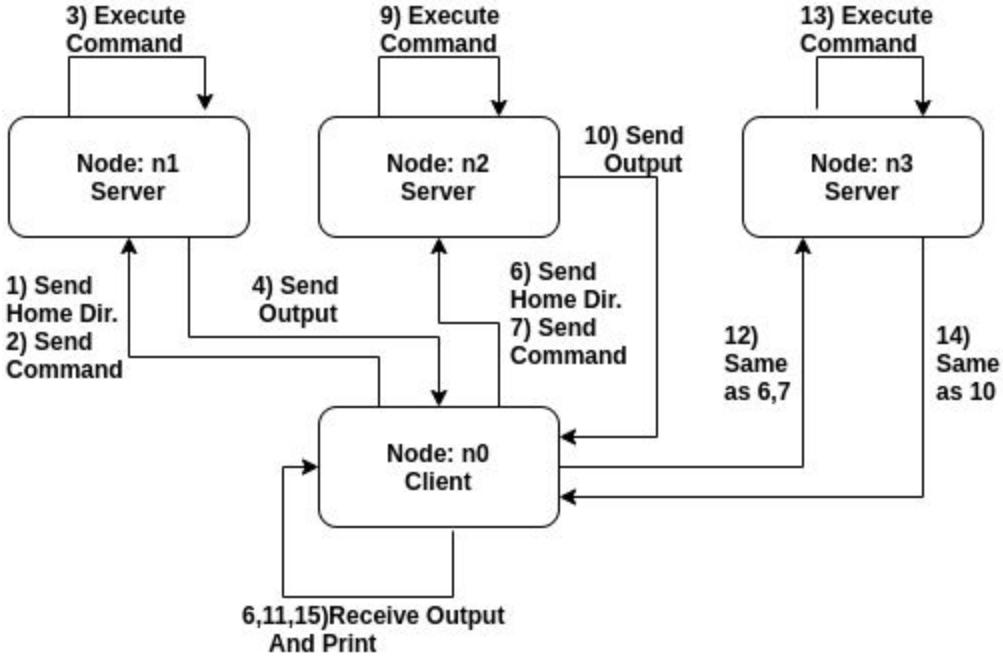
Remote Commands:

Type 1 - Execution on Single machine - E.g. n2.ls



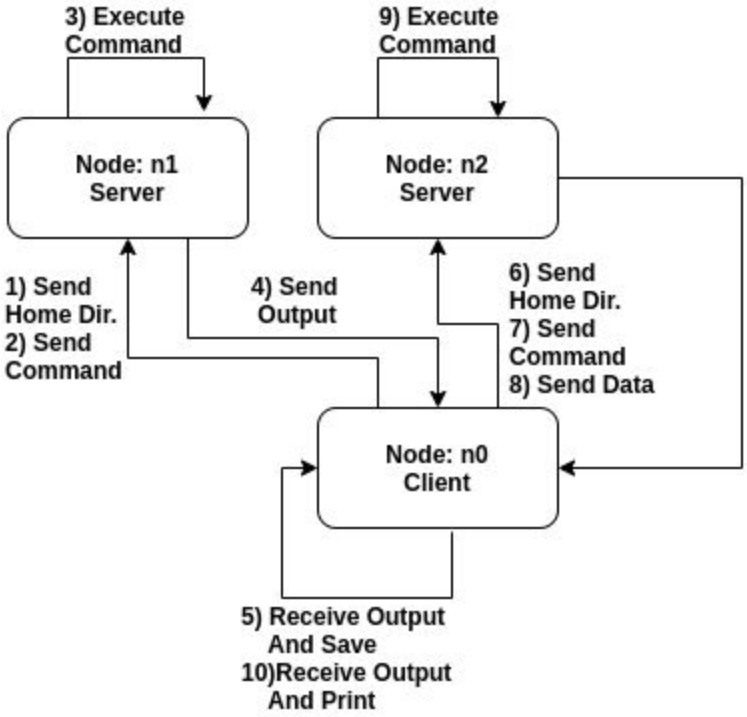
The IP Address of the remote node is queried from the configuration file. Once the TCP connection is established, the client node sends the home directory and the command to be executed to the remote node. Remote node before executing dup the socket connection fd with the stdout, stdin and stderr before execution. By performing the above step the executing command directly writes the result in the socket which is read at the client node and printed on the screen.
Here the client doesn't maintain the state of remote nodes to which it has sent its home directory, it is the job of the server to take action of changing the directory as per it is in default state or has already changed the directory to client's home directory.

Type 2 - Execution on All Machines



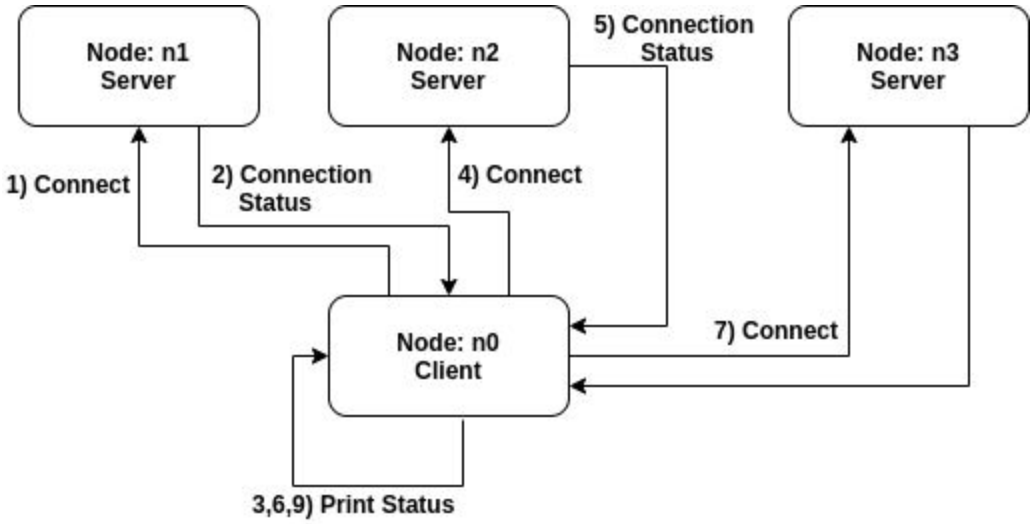
For executing a command on all the machines we follow an iterative approach which can later be extended with the pipeline of commands. The client node sends the home directory and the command to each of the nodes, receive the output from the node and prints it on the screen corresponding to each node.

A pipeline of Commands:



For pipeline of commands, a similar iterative approach is followed but along with sending the commands, we also send the data on the stdout of the previous command if the current remote node needs to read the data it can read it from the socket which is duped on its stdin. The output of the command on stdout is saved for sending to the next node.
For the last command in the pipeline, the final output is displayed on the screen.
We followed an iterative approach so that even if some node is down in between we can still have the intermediate data stored somewhere.

Nodes:



To implement nodes we try to establish a TCP connection with each node in the configuration file if the connection is successfully established the “Node Online” status is printed or if the connection fails “Node Offline” status is printed. After getting the status the connection is closed.