System Y: portfolio

Jannes Van Goeye

Robbe Pauwels

Imre Liessens

# Introduction

System Y is a distributed filesystem: it’s task is to store files in a network. It’s important that the system is robust and fail-safe, otherwise important files could be lost.

# Assignment 1: Naming Server

The naming server’s task is to keep an overview of all the nodes in the network. It has a supporting role to the nodes and doesn’t store any files. In our design of the system we assume the nameserver is always on and reachable.

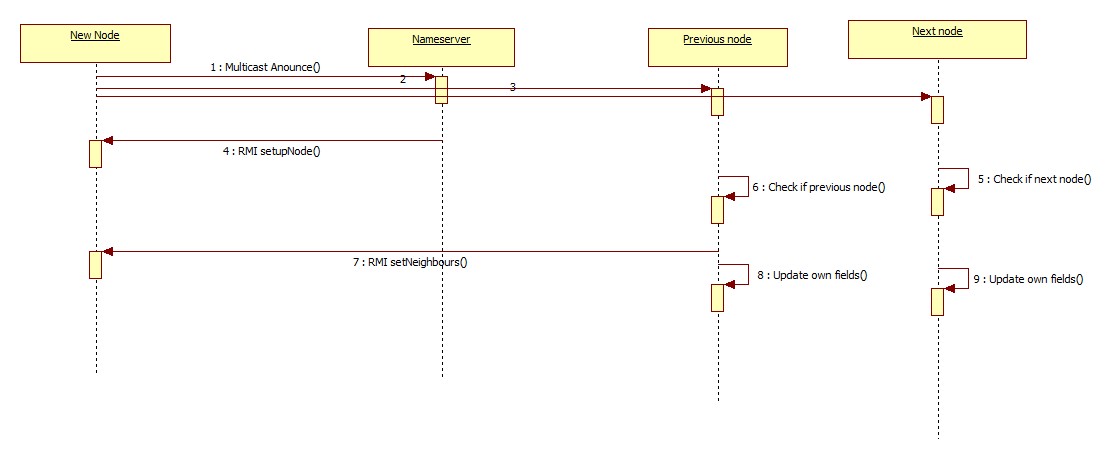
When a change in the node-map occurs, it is also immediately saved to disk. This happens in a comma separated list (CSV). In this manner, the list is maintained when the namingserver shuts down, or experiences a failure. When we restart the program the CSV-file is read and the data map pre-populated.

When in memory, the list of nodes is saved in a TreeMap from the Java Collections framework. This collection has some useful properties. It can’t contain duplicate keys (hashes from the node name), which is exactly what we want. The keys are sorted in ascending order, which makes searching for the next and previous node when given a hash a breeze.



# Assignment 2: Discovery & Bootstrap

The discovery and bootstrap process is executed when a new node is started. Its tasks are manifold, lots of things need to be set up. To accomplish this many classes need to work together, even between different hosts. The function calls are send over the network trough Remote Method Invocation (RMI).



The first announcement is sent via multicast. As internet routers do not forward this type of packet, the discovery process will only work in a LAN. Once the first announcement is sent, the rest of the communication happens trough RMI or TCP (for the file transfers). This has the advantage of a more robust and reliable communication.

# Assignment 3: Replication

The purpose of System Y is to store files in a distributed way, so the replication is a core part of the functionality. When a node joins, where should the files be stored? How should they be transferred?

Each file should also have some metadata associated with it. This data is always stored on the “owner” node and should be transferred if ownership changes.

We should also take care to transfer the file name and extension, so that the replicated files can have the correct name. The solution for this is to have some bytes at the beginning of the TCP stream allocated for this purpose. As the maximum file name length in most platforms (Windows with NTFS, MacOS wit HFS+ and Linux with ext4) is 255 characters, this seems like a good choice.

Most of the bytes will remain unused, but this is a small overhead to pay for the added simplicity. Especially because this happens only once per file transfer.