Francesco Stucci – Matricola N° S4234595

Report Backup Simulator

Single Block per server

The scenario consists of a node with 10 servers (*N*) each storing 1 block of data. In the simulation we have the number of needed blocks to recover the entire data (*K*) in the range [7, 8, 9] which corresponds to respectively 1.43, 1.25 and 1.11 of redundancy. The more blocks are needed to recover our data, the less redundancy we have.
We assume from the theory we can recover the whole data even with few more space of the original data size, so we just check if at each block loss we have at least *K* other stored blocks, in local or in remote.

The following are the constants of our environment:

- NODE_LIFETIME = 100 * DAY -> average time before node crashes and loses data
- NODE_UPTIME = 8 * HOUR -> average time spent online by the node
- NODE_DOWNTIME = 16 * HOUR -> average time spent offline
- DATA_SIZE = 100 * GB -> amount of data to backup
- SERVER_LIFETIME = 365 * DAY -> average time before server crashes and loses block
- SERVER_UPTIME = 30 * DAY -> average time spent online by the server
- SERVER_DOWNTIME = 2 * HOUR -> average time spent offline
- MAXT = 100 * YEAR -> time of the simulation

Usually, in a real world scenario, you cannot have control of these variables and for this reason I preferred not to change these values.

The only value I found reasonable to change are the Download and Upload Speed. Since enterprises can deal with different kind of network connection, I find these values reasonable to make them variables.

In the following graphs, we note that for a few values of redundancy, we need much more network speed in order to take safe our data.

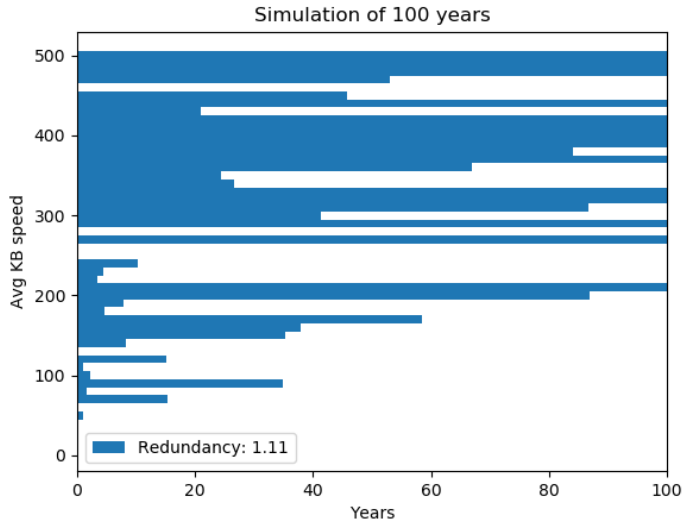

Figure 1: network speed variable and K=9

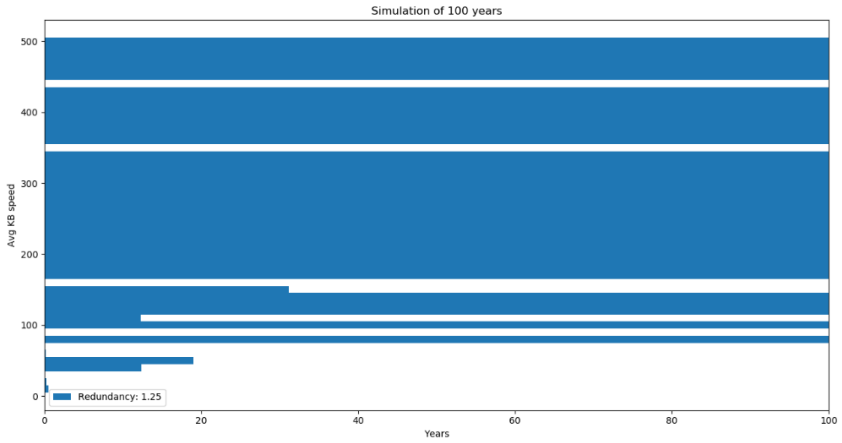

Figure 2: network speed variable and K=8

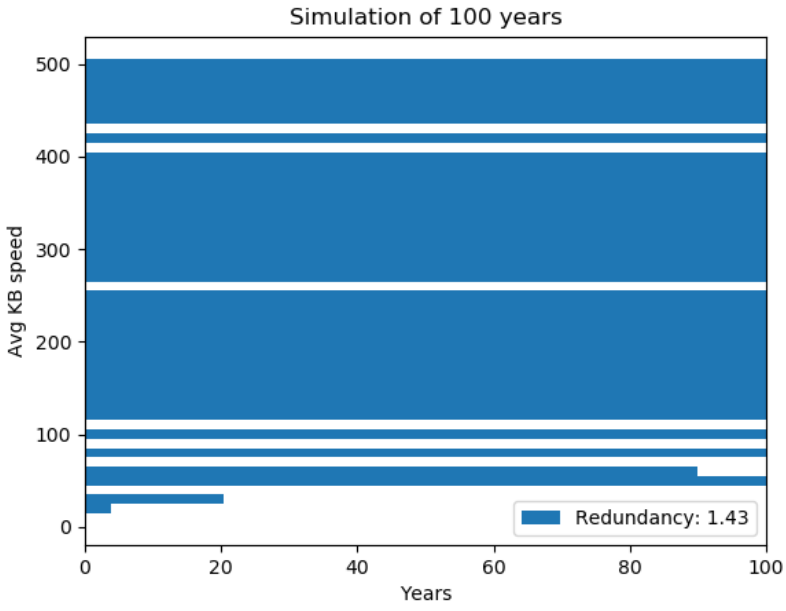

Figure 3: network speed variable and K=7

Multiple Blocks per server

In this case we assume each server can store multiple blocks instead on one as previously.
The tests that has been made on the system are the same: varying the network speed and the data redundancy.

We get the following graphs:

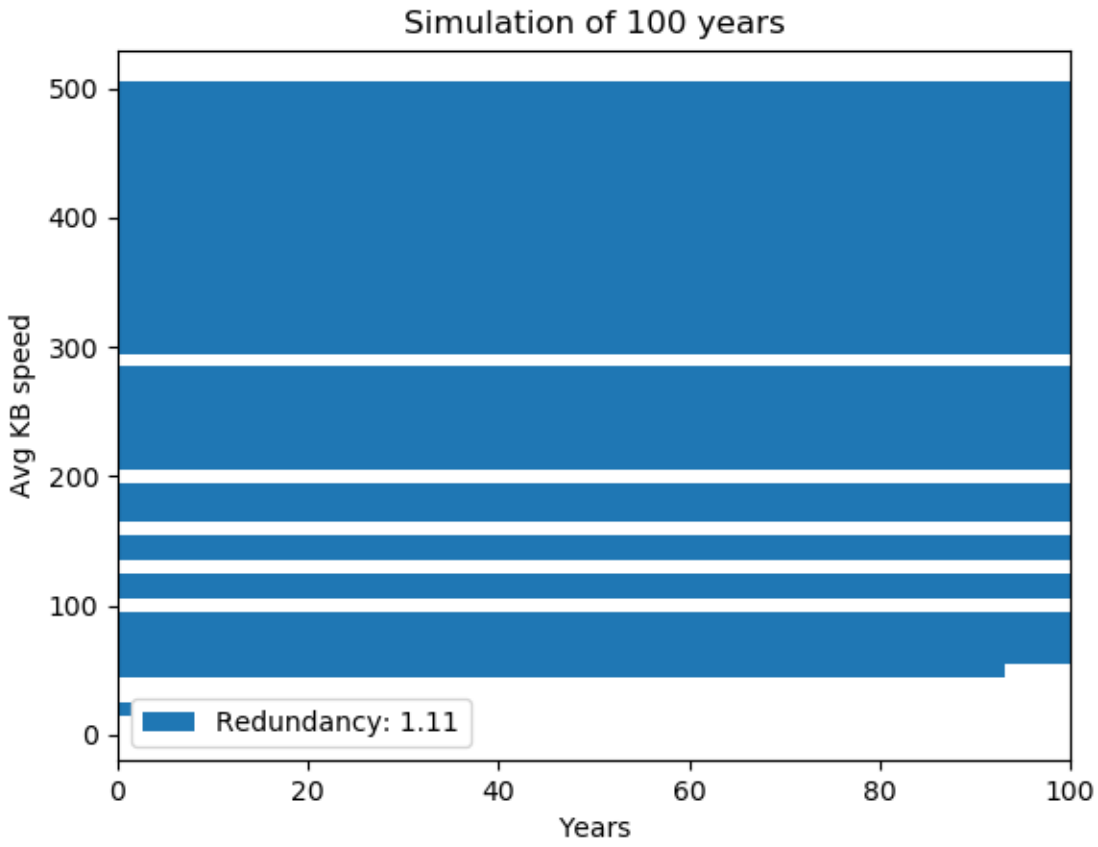

Figure 4: network speed variable and K=7, multiple blocks

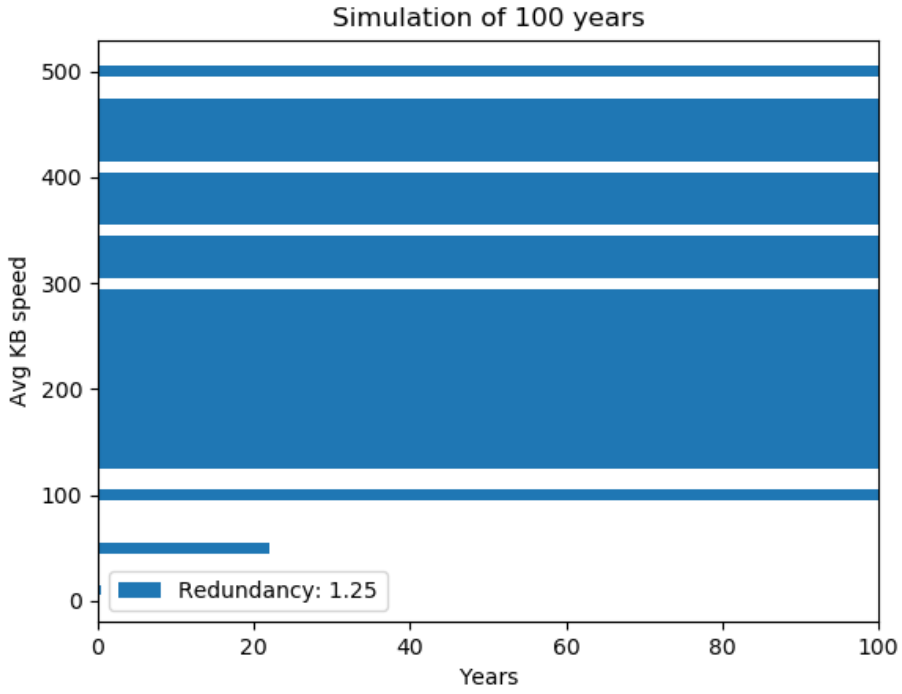

Figure 5: network speed variable and K=8, multiple blocks

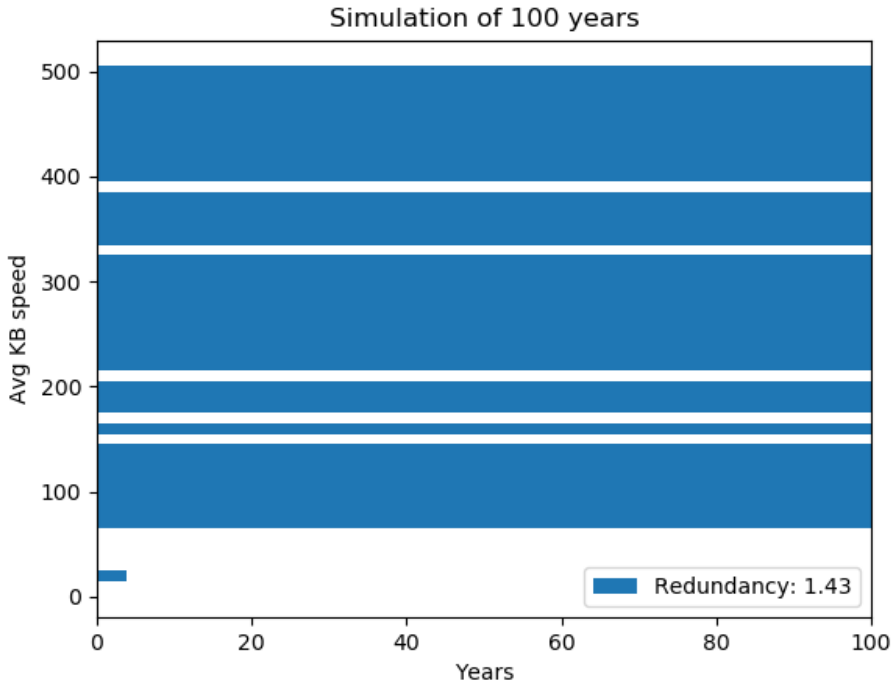

Figure 6: network speed variable and K=9, multiple blocks

With respect to the single block storing case, we can notice that the data are almost always save, even with very slow network connection.
We have some unlucky occasions where node crashed at the beginning hence we lose all data in the first days of the simulation.