

Family name: _____ Given name: _____

- (20%) Given N replicas, let's call R the *ReadConcern* parameter of MongoDB and W the *WriteConcern* (which indicate respectively the number of copies it reads and writes, before confirming the operation to the user); give the equation involving those variables that corresponds to the **eventually consistent** configuration.

.....

2. (20%) Given an empty **linear hash** with $f(x) = x$ (i.e., we directly apply the module to the keys), and a capacity of four keys per bucket, draw the result of inserting the following keys in the given order: 12, 4, 10, 49, 42, 60, 63, 53, 47, 27, 26, 28, 13, 52.

3. (20%) Given a file of $3.2GB$ stored in an HDFS cluster of 50 machines, and containing $16 \cdot 10^5$ key-value pairs in a SequenceFile; estimate the execution time of a Spark job containing a single map transformation and an action storing the results in a file. Explicit any assumption you make and consider also the following parameters:

- Chunk size: 128MB (default)
- Replication factor: 3 (default)
- Map function (i.e., the parameter of the transformation) execution time: 10^{-3} sec/call (this is **the only cost** you have to consider)
- Save action execution time: 0sec (do not consider its cost at all)

[illegible]

4. (20%) Analyze (i.e., briefly give **pros and cons**) the following **JSON design** compared to other equivalent JSON designs from the three perspectives:

```

1  "BID-PRODUCT":{
2    "B_ID": int(4), "B_PRICE": int(4), "U_ID": int(4),
3    "PRODUCT": {
4      "P_ID": int(4), "P_INFO": varchar(100)
5    }}
6  "PRODUCT-SELLER-REGION": {
7    "P_ID": int(4), "P_INFO": varchar(100),
8    "USER": {
9      "U_ID": int(4), "U_F_NAME": varchar(20),
10     "REGION": {
11       "R_ID": int(4), "R_NAME": varchar(10)
12     }}}
13 "PRODUCT-COMMENTS": {
14   "P_ID": int(4), "P_INFO": varchar(100),
15   "COMMENTS": [{
16     "C_ID": int(4), "C_TITLE": varchar(20), "U_ID": int(4)
17   ]}]

```

(a) Read (a.k.a. Query)

i. Pros

.....

.....

.....

ii. Cons

.....

.....

.....

(b) Update

i. Pros

.....

.....

.....

ii. Cons

.....

.....

.....

(c) Memory (a.k.a. Space)

i. Pros

.....

.....

.....

ii. Cons

.....

.....

.....

[illegible]