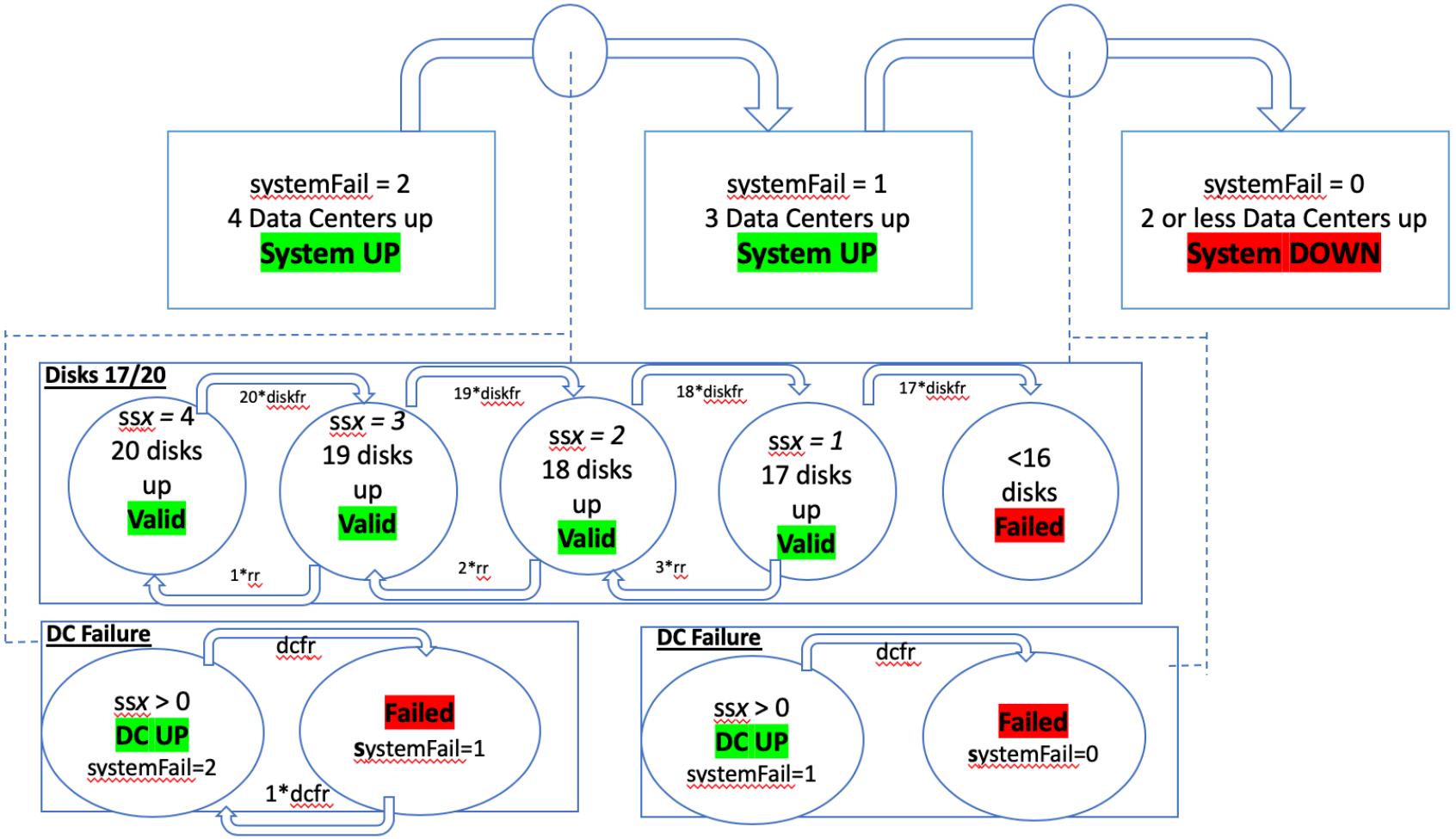
systemFail = 2
4 Data Centers up
System UP
systemFail = 1
3 Data Centers up
System UP
systemFail = 0
2 or less Data Centers up
System DOWN
Disks 17/20
ssx = 4
20 disks
up
Valid
ssx = 3
19 disks
up
Valid
ssx = 2
18 disks
up
Valid
ssx = 1
17 disks
up
Valid
<16
disks
Failed
20*diskfr
19*diskfr
18*diskfr
17*diskfr
1*rr
2*rr
3*rr
DC Failure
ssx > 0
DC UP
systemFail=2
dcfr
Failed
systemFail=1
1*dcfr
DC Failure
ssx > 0
DC UP
systemFail=1
dcfr
Failed
systemFail=0

Description of the model

The model was composed through 4 modules, where each one represented a Data Center. A global variable (systemFail) with values of [0..2] which represented the system state:

- systemFail = 2: represents all data center working correctly, a valid state.
- systemFail = 1: represents 1 data center failing but a valid state due to ¾ coding.
- systemFail = 0: represents an invalid state where 2 or more data centers failed.

Each module could go change the system state within the following scenarios:

- If the disks failed to fulfill the 17/20 coding property and the system was in a “2” state, it would transition to a 1 state due to the ¾ coding property.
- If the module was in a state where 17/20 disks where working and the system was in a “1” state, it would transition to a “0” state by the rate given by 17*diskFailureRate.
- If the disks were in a valid state, and the system was in a “2” state, it would transition to a 1 state with the rate given by the DC failure rate.
- If the disks were in a valid state, and the system was in a “1” state, it would transition to a 1 state with the rate given by the DC failure rate.

The repairing stages where given as followed:

- If “x” number of disks were failing, they would repair with the rate based on the product of the number of failed disks by the disk repair rate. This stage was only available when 3 out of 20 disks were failing.
- If the system was in a “1” state, where only 1 data center was failing, the system would change to “2” with the product of the DC repair rate and the amount of permitted DC failures.

Interpretation of results

The following chart exposes the results of the tests depending on the input repair rate:

Repair Rate (Disk, DC)	Result
1 Week, 1 Week	99.9999991737272
1 Week, 1 Day	99.9999998190262
1 Day, 1 Week	99.9999999230081
1 Day, 1 Day	99.9999999887194

The chart is sorted from the least durable to the most durable. As it was expected, the file will last the least on the repair rate 1 week / 1 week. Since the failure state of the system is reached when more than 1 datacenter fails, and this can be given either through the 17/20 disk coding failure inside either of the 4 datacenters or directly through a complete datacenter failure.

Since repairs can be made, the highest the repair rate the least the probability that the system will fall into a failure state. This is proved in the disk and data center recovery stage, where the higher the product of the failed disk/data center with the repair rate, the highest the probability that the system will stay in an up state (given with ss1>0 and systemFail > 0):

```
//disk repairs
[] (ss1=3) -> 1*rr: (ss1'=4);
[] (ss1=2) -> 2*rr: (ss1'=3);
[] (ss1=1) -> 3*rr: (ss1'=2);
```

```
//datacenter repair
[]systemFail=1 -> 1*dr: (systemFail'=2);
```

Both of these products also show why the highest availability and durability is given with the rate 1 day / 1 day. Since the product will be the highest of all the inputs given, the probability that the systems stay in an up state is higher.

In conclusion, this modeled proved that for cold storage systems, the factors which affects most the durability of a file is how fast either a storage device or the infrastructure that holds them can change its properties from being down (failing) to being up (working) and how probable are that they will failed within a year (AFR).