

# Modeling Drive Reliability with ReMan

Serkay Ölmez

## Abstract

The goal of this work is to develop a complete model to study the benefits of depopulating heads in the field. These benefits include the extension of the device lifetime and the gain in the data durability for erasure coded systems. We will assume that there will be two failure modes: the first one is head related, causing individual heads to fail, and the other will be at the device level, failing the whole device. The model will be flexible to address various cases. The model is supported by simulations.

## Index Terms

Depop, Weibull Distribution, reliability

## I. MODELING FAILURE DISTRIBUTIONS

Weibull distribution [1] is a versatile function due to its flexibility to fit distributions of various shapes. We first consider a particular device failure mode, and assume it can be accurately fit by a Weibull distribution. A Weibull distribution has two parameters:  $\alpha$  parameter sets the time scale whereas the shape parameter  $\beta$  sets the shape of the distribution. The failure probability density, cumulative failure distributions and the hazard rate (failure rate) are defined as follows:

$$f_{\alpha,\beta}(t) = \frac{\beta}{\alpha} \left( \frac{t}{\alpha} \right)^{\beta-1} e^{-\left(\frac{t}{\alpha}\right)^\beta}, \quad (1)$$

$$F_{\alpha,\beta}(t) = \int_0^t d\tau f_{\alpha,\beta}(\tau) = 1 - e^{-\left(\frac{t}{\alpha}\right)^\beta}, \quad (2)$$

$$h_{\alpha,\beta}(t) = \frac{f_{\alpha,\beta}(t)}{1 - F_{\alpha,\beta}(t)} = \frac{\beta}{\alpha} \left( \frac{t}{\alpha} \right)^{\beta-1}. \quad (3)$$

The hazard rate function,  $h_{\alpha,\beta}(t)$ , is the event rate normalized with respect to the population that survived until  $t$ . The actual values of  $\alpha$  and  $\beta$  vary from product to product:  $\beta$  is expected to be between 1 and 2. When  $\beta \neq 1$ , the failure rates are time dependent. It is certainly possible to model ReMan with  $\beta \neq 1$ , but  $\beta = 1$  is typically a good fit, and that is what we will assume in this work.

e-mail:serkay.olmez@seagate.com

This is a static copy of the material. Find the interactive HTML document here.

Acknowledgements: Iman Anvari, Mike Barrell, John Bent, Drew Cheney, Ian Davies, Cary Johnson, Erin Foley and Matt Shumway

## VISUALIZING FAILURE DISTRIBUTIONS

Figure 1. Animated Weibull functions as  $\beta$  is varied. The animation requires Adobe Acrobat reader (it will not work in a browser viewer.)

The actual values of  $\alpha$  and  $\beta$  vary from product to product:  $\beta$  is expected to be between 1 and 2. When  $\beta \neq 1$ , the failure rates are time dependent. It is certainly possible to model ReMan with  $\beta \neq 1$ , but  $\beta = 1$  is typically a good fit, and that is what we will assume in this work.  $\beta = 1$  gives the exponential distribution,

which has completely random head failure times with a fixed failure(hazard) rate:  $h_{\alpha,\beta}(t) = \frac{1}{\alpha} \equiv \lambda$ . This simplifies Eqs. (1), (2) and (3) to:

$$f_{\lambda}(t) = \lambda e^{-\lambda t}, \quad F_{\lambda}(t) = 1 - e^{-\lambda t}, \quad \text{and} \quad h_{\lambda}(t) = \lambda. \quad (4)$$

It is important to understand how the storage devices fail. Weibull is typically a good fit.

No matter how reliable individual devices are, failures are inevitable.

Therefore, redundancies are needed to preserve data in case of drive failures.

## II. MODELING ONE-HEAD REMAN

The probability of having no failures in an  $N$ -header drive is:

$$p_{head}^{all\ pass} = (p_{head}^{pass})^N = e^{-N \left(\frac{t}{\alpha}\right)^{\beta}}, \quad (5)$$

where  $\alpha$  and  $\beta$  are head-level Weibull parameters. The probability of having **exactly** one head failure is

$$p_{drive}^{exactly\ one\ head\ fail} = N \left(p_{head}^{fail}\right) (p_{head}^{pass})^{N-1} = N \left(1 - e^{-\left(\frac{t}{\alpha}\right)^{\beta}}\right) e^{-(N-1) \left(\frac{t}{\alpha}\right)^{\beta}}. \quad (6)$$

Then probability of having two or more failures in an  $N$ -header drive ( $CFR_R$ ) reads:

$$CFR_R = N \left(1 - e^{-(N-1) \left(\frac{t}{\alpha}\right)^{\beta}}\right) - (N-1) \left(1 - e^{-N \left(\frac{t}{\alpha}\right)^{\beta}}\right). \quad (7)$$

Eq. (7) is a mixture of two Weibull modes. The first term describes  $N-1$  header drives, and the second describes  $N$  header drives. This equation still uses the time per head, and we need to convert that to drive level time. After the first head failure, the head time will increase as  $t_d/(N-1)$ . Scaling the time in Eq. (7), we get

$$CFR_R^{drive}(t_d) = N \left(1 - e^{-(N-1) \left(\frac{t_d}{\alpha(N-1)}\right)^{\beta}}\right) - (N-1) \left(1 - e^{-N \left(\frac{t_d}{\alpha(N-1)}\right)^{\beta}}\right). \quad (8)$$

## III. MODELING K-HEAD REMAN

Consider the case where at most  $k$ -head failure is acceptable. The probability of having up to  $k$  (including  $k$ ) failures is

$$p_{drive}^{up\ to\ k\ head\ failures} = \sum_{j=0}^k \binom{N}{j} \left(p_{head}^{fail}\right)^j (p_{head}^{pass})^{N-j} = \sum_{j=0}^k \binom{N}{j} \left(1 - e^{-\left(\frac{t}{\alpha}\right)^{\beta}}\right)^j e^{-(N-j) \left(\frac{t}{\alpha}\right)^{\beta}}. \quad (9)$$

Then the CFR for having more than  $k$  head failures becomes

$$\begin{aligned} CFR(k - \text{ReMAN}) &= p_{drive}^{more\ than\ k\ head\ failures} = 1 - p_{drive}^{up\ to\ k\ head\ failures} \\ &= 1 - \sum_{j=0}^k \binom{N}{j} \left(1 - e^{-\left(\frac{t}{\alpha}\right)^{\beta}}\right)^j e^{-(N-j) \left(\frac{t}{\alpha}\right)^{\beta}}. \end{aligned} \quad (10)$$

## IV. REMAN ELIMINATES HEAD RELATED FAILURES

There are two types of failures:

Head specific failures, which can be handled by ReMan: Just depop the failing head, and move on. We will show that this practically eliminates head related failures.

Entire device failures: motor/PCBA/HSA failures, etc. These cannot be addressed by ReMan.

Data suggests that a significant portion of the failures fall into the ReMan'able bucket. The overall reliability of a device can be computed by the multiplying the individual reliability functions,  $R_i$ :

$$R(t) = \prod_i R_i = R_{\text{ReMan'able}} \times R_{\text{non-ReMan'able}}. \quad (11)$$

$R_{\text{non-ReMan'able}}$  is the reliability associated with the entire device failures, and it can simply be defined by a failure rate  $\lambda_{\text{NR}}$ .  $R_{\text{ReMan'able}}$ , on the other hand, is the associated with the head-specific failures. They are fixable by depoping heads.

See Fig. 2 for an example simulation. See the caption for the parameters.

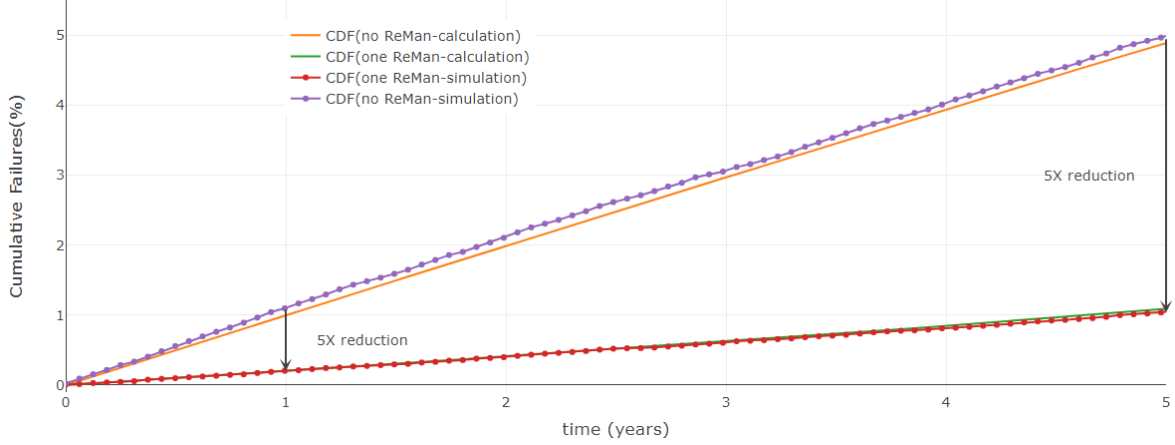


Figure 2. Comparison of cumulative failures with and without ReMan. Total failure rate is taken as 1% split into 0.8% ReManable and 0.2% Non-ReMan'able. We also take  $\beta_R = 1$ , and  $\beta_{NR} = 1$ .

Even one-head ReMan absorbs practically all head specific failures.

In this example, the Annualized Failure Rate (AFR) is effectively reduced down to 0.2% from the original value of 1%.

## V. REMAN PRESERVES CAPACITY AT THE EDGE

Consider an edge application in which failed drives are not replaced immediately. ReMan will be an indispensable tool to preserve the capacity of the storage system:

Entire device failures will take down the full drive causing the loss of the full capacity.

Head specific failures will be ReMan'ed, losing  $1/N$  of the full cap (5% for a 20-header drive).

ReMan helps preserve capacity much longer.

When no ReMan is allowed the average capacity will drop with drive failures, which can be written as follows:

$$\mathcal{CAP}_{0R} = \mathcal{CAP}_0(1 - CFR_{0R}) \quad (12)$$

When at most one ReMan is allowed, the average capacity will be

$$\mathcal{CAP}_{1R} = \mathcal{CAP}_0(1 - CFR_{1R}) - \mathcal{CAP}_0(CFR_{0R} - CFR_{1R})/N, \quad (13)$$

where the second term is the ratio of the drives with exactly one head failure [they lost a capacity of  $\mathcal{CAP}_0/N$ ].

Figure 3 shows the CFR and capacity.

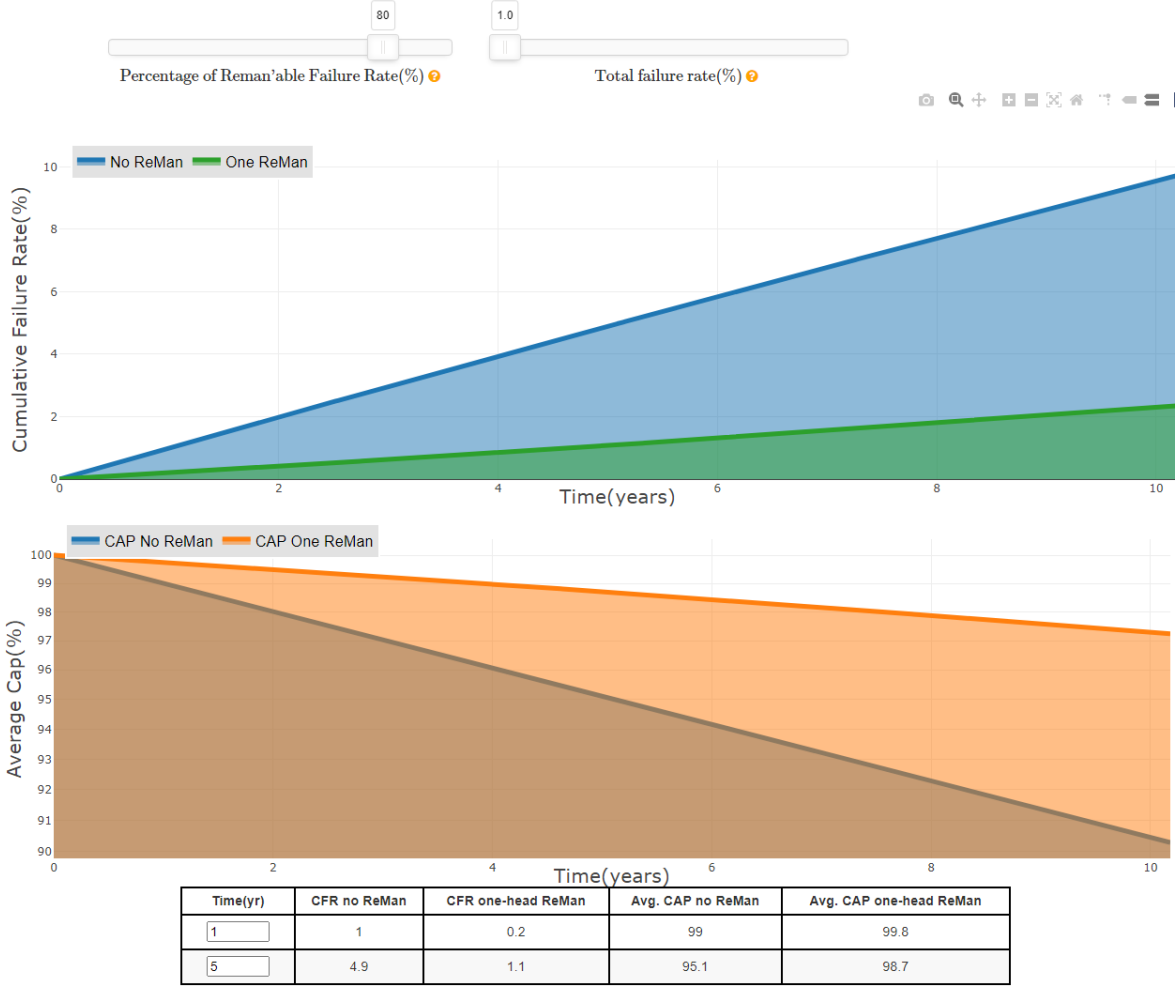


Figure 3. Cumulative Failures/Average Capacity vs time with number of heads  $N = 20$ . We also take  $\beta_R = 1$ , and  $\beta_{NR} = 1$

## VI. REMAN REDUCES SWAP RATE AND TCO IN DATA CENTERS

As shown earlier, ReMan substantially reduces drive failures and, in turn, the cost of replacing them. The only drawback of ReMan is that ReMan'ed drives will have one-less heads, hence slightly reduced capacity. In the case of maximum one-head ReMan, the overall capacity of the system will reduce only slightly. Qualitatively speaking, the ratio of ReMan'ed drives will increase slowly, because the increase is suppressed by two factors:

As drives fail for the NonReMan'able mode, they are replaced. This removes already ReMan'ed drives out of the population.

If a second head in an already-ReMan'ed drive fails, it is replaced. This also removes ReMan'ed drives out of the population.

The precise mathematical model is not trivial[2] due to the factors above, see the details below. The final expression for the ReMan'ed drive density is given in Eq. (16) and it is plotted and compared with simulation results in Figure 4 for illustration purposes[see the caption for details]. Once the ReMan'ed drive density is calculated, the degradation in the overall capacity can be computed as in Eq. (18). How average capacity will change with time is shown in Fig. 5.

When a head fails, it is "turned off," and the drive continues to operate with reduced capacity. We will denote the ratio of devices that lost  $i$  head as  $r_i$ , and compute how these ratios evolve with time. We first consider a replacement policy which allows for maximum one-head ReMan, which implies that

$r_0(t) + r_1(t) = 1$ . Furthermore, since the system starts with new devices, the initial values of the ratios are  $r_0(0) = 1$  and  $r_1(0) = 0$ . We need to calculate how  $r_0$  evolves with time given that devices have two distinct failure modes: **i)** a failure mode that affects individual heads, **ii)** another failure mode that affects the whole device and results in device replacement. We define the failure rate that affects heads as  $\lambda_R$ , where R stands for *Re-manufacturable*, and the other mode as  $\lambda_{NR}$ , where NR stands for *NonRe-manufacturable*. As heads fail, the ratio of the devices with no head failures ( $r_0$ ) will decrease:  $\dot{r}_0 \propto -\lambda_R r_0$ . On the other hand, devices that already have one head failure ( $r_1$ ) will be replaced with new devices once another head fails. The replacement will move these devices to the  $r_0$  bucket. The rate of this event is proportional to the total number of heads in  $r_1$  bucket. Adding this term yields  $\dot{r}_0 \propto -\lambda_R r_0 + \lambda_R \frac{N-1}{N} r_1$ , where the factor  $\frac{N-1}{N}$  normalizes the failure rate since there are  $N-1$  heads per device in the  $r_1$  bucket. There is one more term to add: the device level failures with rate  $\lambda_{NR}$ . Such failures require device replacements, increasing  $r_0$  only if the failure happens to be in the  $r_1$  bucket. Therefore the rate of such events will be  $r_1 \lambda_{NR}$ . Adding all the terms, and using  $r_1 = 1 - r_0$  result in the following differential equation:

$$\dot{r}_0 = -\lambda_R r_0 + \lambda_R \frac{N-1}{N} (1 - r_0) + \lambda_{NR} (1 - r_0). \quad (14)$$

Eq. (14) has the following solution:

$$r_0(t) = \frac{N-1 + N\kappa + Ne^{-(2+\kappa-1/N)\lambda_R t}}{2N-1+N\kappa} = 1 - \frac{1}{2+\kappa-1/N} \left(1 - e^{-(2+\kappa-1/N)\lambda_R t}\right), \quad (15)$$

where we defined  $\kappa = \frac{\lambda_{NR}}{\lambda_R}$ . Since  $r_0 + r_1 = 1$ , we get

$$r_1(t) = 1 - r_0(t) = \frac{1}{2+\kappa-1/N} \left(1 - e^{-(2+\kappa-1/N)\lambda_R t}\right), \quad (16)$$

which is the relative ratio of the ReMan'ed drives. We can now write the average capacity of the system:

$$\overline{CAP} = C(r_0 + r_1 \frac{N-1}{N}) = C(r_0 + r_1 - \frac{r_1}{N}) = C(1 - \frac{r_1}{N}) \quad (17)$$

Equivalently, we can define the relative the loss in the capacity:

$$\frac{\Delta \overline{CAP}}{C} = \frac{C - \overline{CAP}}{C} = \frac{r_1}{N}, \quad (18)$$

which basically is the density of the ReMan'ed drives divided by  $N$ , since they lose  $1/N$  of their capacity.

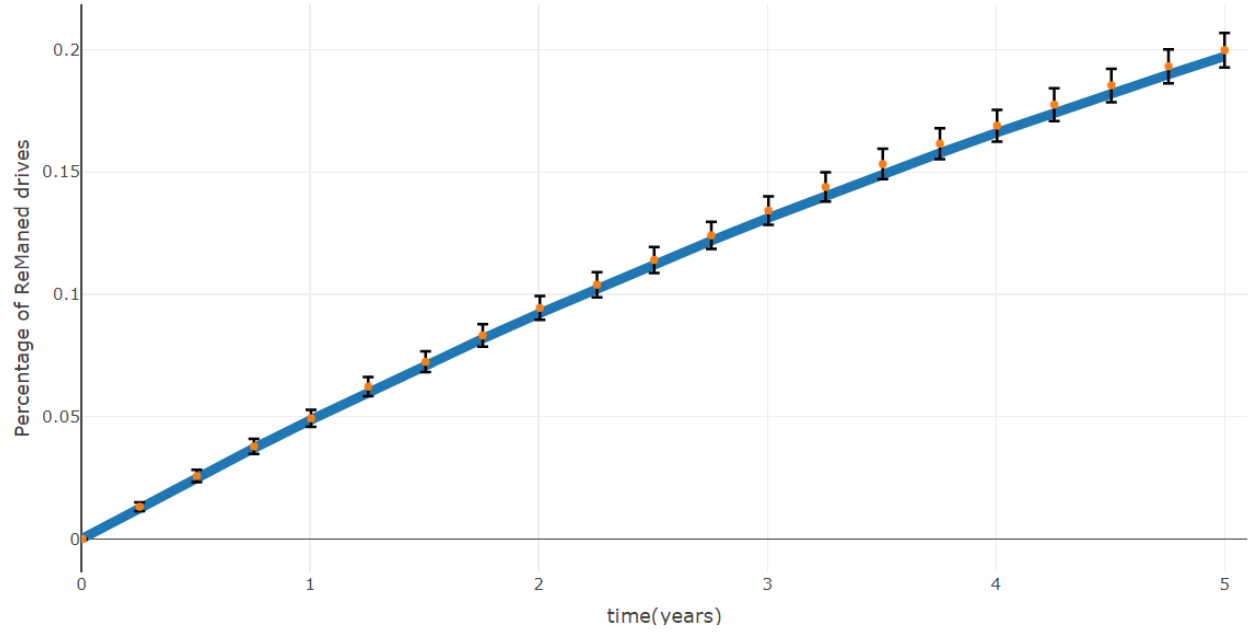


Figure 4. Cumulative Failures/Average Capacity vs time with number of heads  $N = 20$ . We also take  $\beta_R = 1$ , and  $\beta_{NR} = 1$

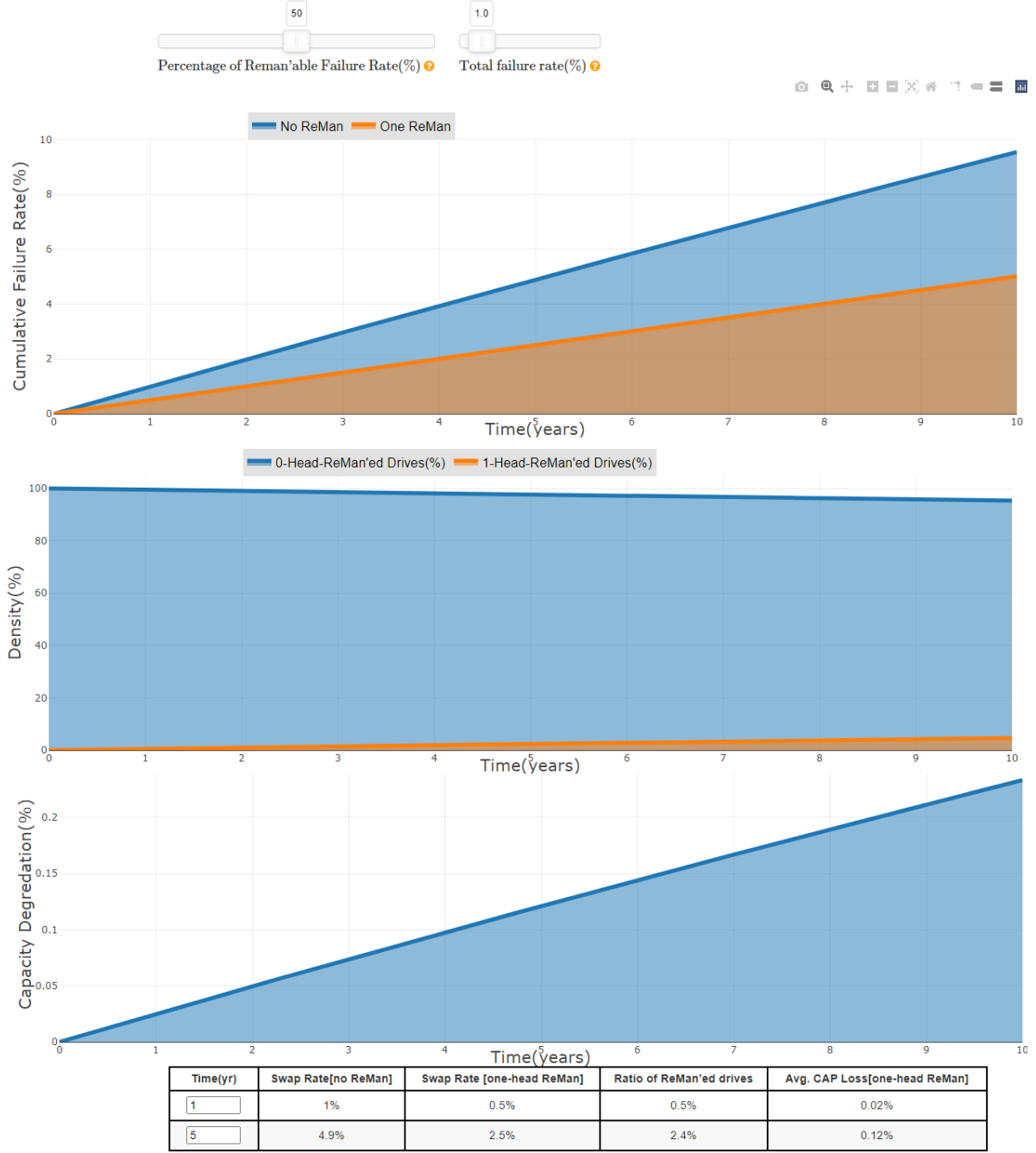


Figure 5. CFR/ReMan Density/Capacity degradation vs time,  $\beta_R = 1$ , and  $\beta_{NR} = 1$ .

## VII. REMAN IMPROVES DATA DURABILITY

Consider the simplest case[2]: one redundancy. Given a head/drive failure, there are two paths to losing data:

A second failure while recovering from a head failure: less likely due to faster recovery.

A second failure while recovering from a drive failure: more likely due to longer recovery.



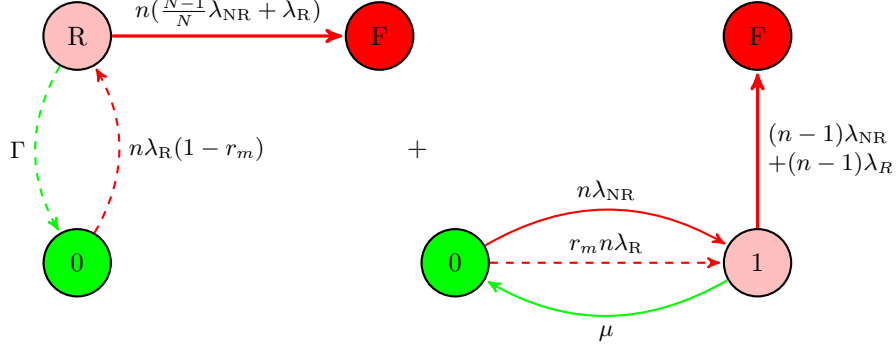


Figure 6. Markov chain split into two parallel paths: recovery from subcomponent failures and recovery from device failures.

The plot on the left shows two parallel Markov chains:

Left-most shows data recovery for a ReMan'able failure.

Right-most shows the standard recovery when a drive fails.

$r_m$  term represents the maximally ReMan'ed drive population.

ReMan'able failures on maximal drives trigger replacement.

This is represented by the dashed red curves.

The transition rates are time dependent, and involve  $r_m$  functions.

$$\mathcal{R}_1^{(1)}(t) = \exp\left(-\frac{\mathcal{C}}{2\Gamma} \left[t + \frac{1-e^{-2\lambda_R t}}{2\lambda_R}\right] - \frac{\mathcal{C}}{2\mu} \left[(1+2\kappa)t - \frac{1-e^{-2\lambda_R t}}{2\lambda_R}\right]\right) \text{ where } \mathcal{C} \equiv n^2 \lambda_R^2 (1+\kappa), \text{ and } \kappa \equiv \frac{\lambda_{NR}}{\lambda_R}$$

$n$  : number of drives in the EC scheme,  $N$  : number of heads per drive,

$\Gamma$  : ReMan Repair rate,  $\mu$  : drive repair rate,  $\lambda_R$ : ReMan'able failure rate,  $\lambda_{NR}$ : non-ReMan'able failure rate.

This is the expression when we allow for 1 ReMan/drive. Similar formulas are calculated for  $R_{max} = 2, 3$ .

We have an analytical model of Erasure Coded systems that support ReMan.

The closed mathematical form can be computed instantly enabling a real-time web application.

## VIII. SIMULATION WITH ONLINE REMAN

- Below is a comparison of Monte Carlo Simulation and theoretical results[Assuming 1-ReMan/drive is allowed.]:
- ReMan'able AFR= 8% and non ReMan'able AFR=2%.[Taken unrealistically high to show the functional behavior]
- The plot shows the theoretical prediction is remarkably accurate: The gain in durability is about 4x.

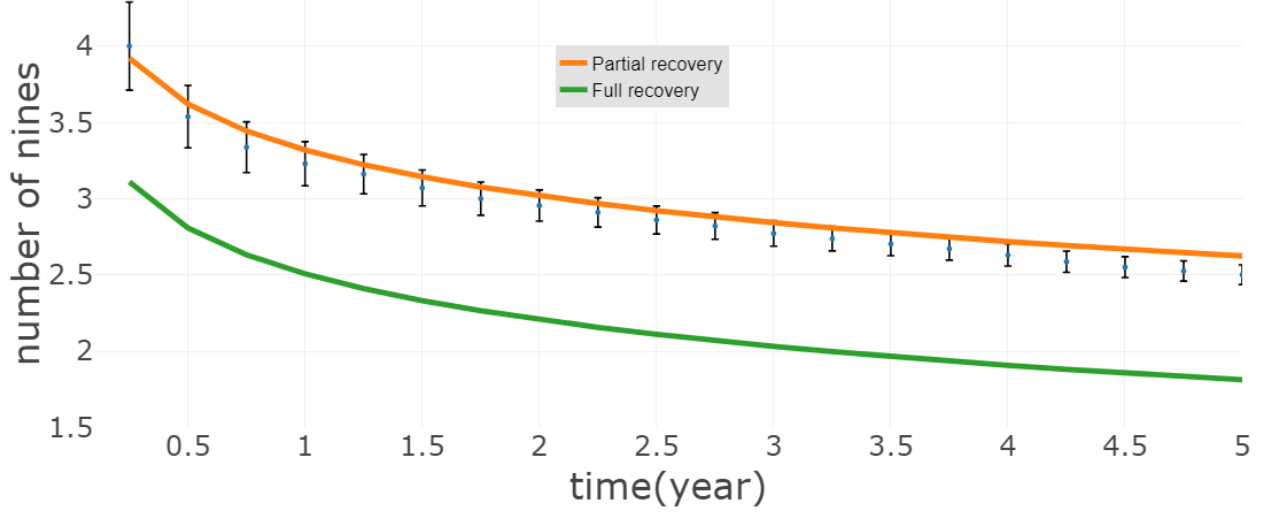


Figure 7. The calculated values of number of nines without ReMan(green curve), with ReMan(orange curve) and the simulation results for number of nines with ReMan (blue dots). Input parameters are taken as  $\lambda_R = 1/12$  and  $\lambda_{NR} = 1/50$  (1/year), which correspond to 8% and 2% AFR respectively.

## IX. REFERENCES

- [1] W. Padgett, “Weibull distribution,” 2011, pp. 1651–1653.
- [2] S. Olmez, “Reliability analysis of storage systems with partially repairable devices,” 2021 [Online]. Available: Accepted to be published in Transactions on Device and Materials Reliability, IEEE
- [3] D. A. Patterson, G. Gibson, and R. H. Katz, “A case for redundant arrays of inexpensive disks (RAID),” *SIGMOD Rec.*, vol. 17, no. 3, pp. 109–116, Jun. 1988, doi: 10.1145/971701.50214. [Online]. Available: <https://doi.org/10.1145/971701.50214>
- [4] L. Hellerstein, G. A. Gibson, R. M. Karp, R. H. Katz, and D. A. Patterson, “Coding techniques for handling failures in large disk arrays,” *Algorithmica*, vol. 12, no. 2, pp. 182–208, Sep. 1994, doi: 10.1007/BF01185210. [Online]. Available: <https://doi.org/10.1007/BF01185210>
- [5] A. G. Dimakis, P. B. Godfrey, Y. Wu, M. J. Wainwright, and K. Ramchandran, “Network coding for distributed storage systems,” *IEEE Trans. Inf. Theor.*, vol. 56, no. 9, pp. 4539–4551, Sep. 2010, doi: 10.1109/TIT.2010.2054295. [Online]. Available: <https://doi.org/10.1109/TIT.2010.2054295>
- [6] G. Wang, L. Xiao-Guang, and L. Jing, “Parity declustering data layout for tolerating dependent disk failures in network RAID systems,” 2002, pp. 22–25, doi: 10.1109/ICAPP.2002.1173547.
- [7] V. Venkatesan and I. Iliadis, “Effect of codeword placement on the reliability of erasure coded data storage systems,” in *Quantitative evaluation of systems*, 2013, pp. 241–257.
- [8] A. Thomasian and M. Blaum, “Higher reliability redundant disk arrays: Organization, operation, and coding,” *ACM Trans. Storage*, vol. 5, no. 3, Nov. 2009, doi: 10.1145/1629075.1629076. [Online]. Available: <https://doi.org/10.1145/1629075.1629076>
- [9] T. Kawaguchi, “Reliability analysis of distributed RAID with priority rebuilding,” 2013.