The Business of Stochastic Multi Level Disk Risk Modeling for Storage Arrays

Rachel Traylor (and Craig Struble)
www.themathcitadel.com
@Mathpocalypse

August 15, 2017
Introduction and Background

- Mathematician, with research interests in
  - probability theory, especially sequences of dependent random variables
  - stochastic analysis
  - system reliability and queuing theory
  - statistics, particularly time series
- Hobbies: marathon swimming, surfing, and eating bacon for as many meals as possible
Actual Introduction: The Problem

Disks fail, but we don’t really care that much about it.

- Modern storage systems have redundancy (like RAID) to handle a single disk failure
- What we’re really interested in is the probability of data loss
How Project Risky Business works

1. Calculate the cumulative probability of each disk failing in the next X days
2. Calculate the probability of data loss in each individual RAID group
3. Calculate the probability of data loss in a customer system
4. Calculate a customer’s probability of data loss overall
Step 1: Calculate the cumulative probability of disk failure in the next $X$ days

For the number of transient states $K$, 

$$ P = \begin{bmatrix} p_{00} & p_{01} & p_{02} & \cdots & \cdots & p_{0F} \\ 0 & p_{11} & p_{12} & \cdots & \cdots & p_{1F} \\ 0 & 0 & p_{22} & \cdots & \cdots & p_{2F} \\ 0 & \cdots & 0 & \cdot & \cdots & \cdot \\ \vdots & \cdots & \cdots & \cdots & \cdots & \cdot \\ 0 & \cdots & \cdots & \cdots & p_{KK} & p_{KF} \\ 0 & \cdots & \cdots & \cdots & 0 & 1 \end{bmatrix} $$
Step 1: Calculate the cumulative probability of disk failure in the next $X$ days

The probability that a disk with $i$ medium errors today will fail in 30 days is given by

$$P_{iF}^{30}$$

In other words, the last column of the $X$th power of the transition matrix gives the information we need.

That’s it!
Step 2: Calculate the probability of data loss in each individual RAID group

The probability of data loss in $X$ days (RAID 6) is $P(F \geq 2)$, or,

$$1 - P(F = 0) - P(F = 1)$$
Step 2: Calculate the probability of data loss in each individual RAID group

Suppose there are 10 disks in a RAID group. What is $P(F_5 = 0)$?

Individual failure probabilities: $p_1, \ldots, p_{10}$

Then the probability they will not fail in the next 5 days is

$$1 - p_i, \ i = 1, \ldots, 10$$

$$P(F = 0) = (1 - p_1)(1 - p_2)(1 - p_3)\ldots(1 - p_{10})$$
Step 2: Calculate the probability of data loss in each individual RAID group

What is $P(F = 1)$? We need to "loop" through selecting one disk to fail, and the other 9 don’t.

If the failure probabilities were all the same, the distribution would be binomial. (But they generally are not.)
Step 3: Calculate the probability of data loss in a customer system

A customer system loses data if any RAID group loses data. So

\[ P(S) = P(E \geq 1) = 1 - P(E = 0) \]

Each enclosure has data loss probability \( \pi_i, i = 1, ..., N \). Then

\[ P(S) = 1 - ((1 - \pi_1)(1 - \pi_2) \cdots (1 - \pi_N)) \]
Step 4: Calculate a customer’s probability of data loss overall

We use the exact same concept as Step 3. Let $S$ be the number of systems in a customer’s profile that lose data.

$$P(C) = 1 - P(S = 0)$$

Each system has data loss probability $s_i, i = 1, \ldots, M$. Then

$$P(C) = 1 - ((1 - s_1)(1 - s_2) \cdots (1 - s_M))$$
In practice
In practice

- This whole model runs on the entire install base every night and updates
- Validation analysis shows that this method is quite successful
Lessons

- Sometimes an elegant model works best
- Simplicity allowed for interpretability and granularity
- This model can be used on any error type you like, or any RAID algorithm.
Questions?