Autumn 2018

Big Data Platforms
Tutorial 3

In this tutorial we do some calculations of mean time to failure (MTTF) of hypothetical storage configurations. In the following assume a mean time to failure (MTTF) of a hard disk to be 300 000 hours (an estimate based on field data of hard disk replacements). See "Bianca Schroeder, Garth A. Gibson: *Disk Failures in the Real World: What Does an MTTF of 1, 000, 000 Hours Mean to You?* FAST 2007: 1-16" for a more detailed analysis of real life hard disk replacement rates.

For more information on RAID levels, see:
`http://en.wikipedia.org/wiki/Standard_RAID_levels`

In the following, assume hard drive failures form a Poisson process, i.e., they are independent and the time between failures follows an exponential distribution.

1. When using 4 TB hard disks in RAID 0 configuration (striping data over all disks, no parity), compute the mean time to failure in years of a RAID 0 array that consist of the following amounts of storage space:

   a) 16 TB
   b) 64 TB
   c) 256 TB
   d) 1024 TB
   e) 4096 TB

   Hint: The array fails if any one of the hard disk fails in RAID 0 configuration, and the number of hard disk used equals the amount of storage.

2. When using 4 TB hard disks in RAID 5 configuration (n disks for data + 1 disk for parity), compute the mean time to failure in years of a RAID 5 array that consist of the following amounts of storage space. To simplify calculations, we (pessimistically!) assume that failed hard disks are left unrepaired:

   a) 16 TB
   b) 64 TB
   c) 256 TB
   d) 1024 TB

e) 4096 TB

Hint: Use the minimum number of disks, e.g., 16 TB will be obtained using 4 data disks + one parity disk. The array fails if two hard disks fail.

3. Demo exercise, you can also use simulation tools to estimate the results: When using 2 TB hard disks in RAID 10 configuration (also called RAID 1+0, stripe of mirrored disks, see `http://en.wikipedia.org/wiki/Nested_RAID_levels`), compute the mean time to failure in years of a RAID 10 array that consist of the following amounts of storage space. To simplify calculations, we (pessimistically!) assume that failed hard disks are left unrepaired:

   a) 4 TB
   b) 16 TB
   c) 64 TB
   d) 256 TB
   e) 1024 TB

   Hint: Use simple mirrors with two disks per mirror set. The array fails if two hard disks of the same mirror pair fail.

Appendix

Background

Before discussing possible solutions to the three problems, it helps to review some properties of Poisson processes. From Wikipedia we know that *a Poisson process is a stochastic process in which events occur continuously and independently of one another. Examples that are well-modeled as Poisson processes include the radioactive decay of atoms, telephone calls arriving at a switchboard, page view requests to a website, and rainfall. A Poisson process is usually described as a function of time, although it need not be.* In this tutorial we assume that hard disk errors (more precisely, their arrival process) can be modeled as a Poisson process, which is an assumption that helps in finding analytical solutions to many questions in this context. However, it is worth noting that this assumption may not be necessarily true in practice.

Arrival times

Let us define:

$$N(t) := \text{number of arrivals in the time interval } (0, t),$$

where $N(t) \sim \text{Poisson}(\lambda\, t)$, which means that $N(t)$ is a random variable that follows a Poisson distribution with intensity (or arrival/error rate) λ. Later, for our disk error process we will have $\lambda = 1/(3 * 10^5 \text{h})$. Figure 1 visualizes such an arrival process, where arrows indicate discrete events (such as hard disk errors).

Due to the properties of the distribution we know that for any time $t > 0$

$$\text{Prob}\{N(t) = n\} = \frac{(\lambda\, t)^n}{n!} e^{-\lambda\, t}$$

and more generally

$$\text{Prob}\{N(t_0 + \Delta) - N(t_0) = n\} = \frac{(\lambda\, \Delta)^n}{n!} e^{-\lambda\, \Delta}$$

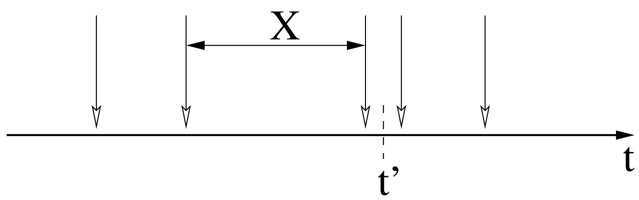

Figure 1: Arrival process of discrete events (e.g., disk failures); the random variable X corresponds to the interarrival time between two consecutive events. For this *realization* of the random process, we have $N(t') = 3$.

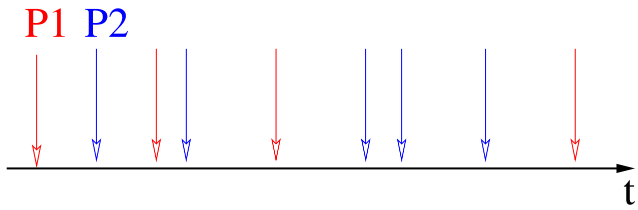

Figure 2: Plot showing the superposition of two (independent) Poisson processes P_1 and P_2.

for any time $t_0, \Delta > 0$. Moreover, considering the *expected value* (or *mean*) $E[N(t)]$ of $N(t)$, we know that

$$E[N(t)] = \lambda\, t.$$

An important property of the Poisson process is that if one adds two (or any number) of (independent) Poisson processes to form a *superposition*, one again ends up with a Poisson process, which has as intensity the sum of intensities of the two original processes. This is depicted in Figure 2. More precisely, if we have a family of K Poisson processes with intensities λ_i, $1 \leq i \leq K$, and define

$$Z(t) := \sum_{i=1}^{K} N_i(t),$$

then $Z(t) \sim \text{Poisson}(t\ \sum_i \lambda_i)$. This fact will be useful in Problem 1, when we combine several hard disks into a single one via a RAID 0 setup.

Interarrival times

An important fact is that if the arrival process is a Poisson process, then the interarrival times (the time between two consecutive arrival events) are independent and follow an *exponential distribution*. Let X be the random variable that corresponds to the interarrival time (see Figure 1), then we have $X \sim \text{Exp}(\lambda)$, where the λ is the intensity of the corresponding arrival process. Note that in our problems the variable X corresponds to the time between two consecutive hard disk failures (which means that it models the lifetime of a disk). Also note that while the arrival process is discrete, the interarrival times are continuous. Due to the properties of the exponential distribution we have

$$\text{Prob}\{X \leq t\} = 1 - e^{-\lambda\, t},$$

which is equivalent to

$$\text{Prob}\{X > t\} = e^{-\lambda\, t}.$$

We further can express the expected value (or mean) of X by

$$E[X] = \frac{1}{\lambda}.$$

An important property of the interarrival-time process here is that it is *memoryless*. Applied to our setting of hard-disk errors, this means that a hard disk that was taken into use 2 months ago has the same probability to fail within the next month from now as it had for the very first month. More precisely,

$$\text{Prob}\{X > t_o + \Delta | X > t_0\} = \text{Prob}\{X > \Delta\},$$

for any time t_0 and time difference Δ. This property is visualized by the plot of $\text{Prob}\{X > t\}$ shown in Figure 3, which also shows the plot for the remaining interarrival time after time t_0 has passed.

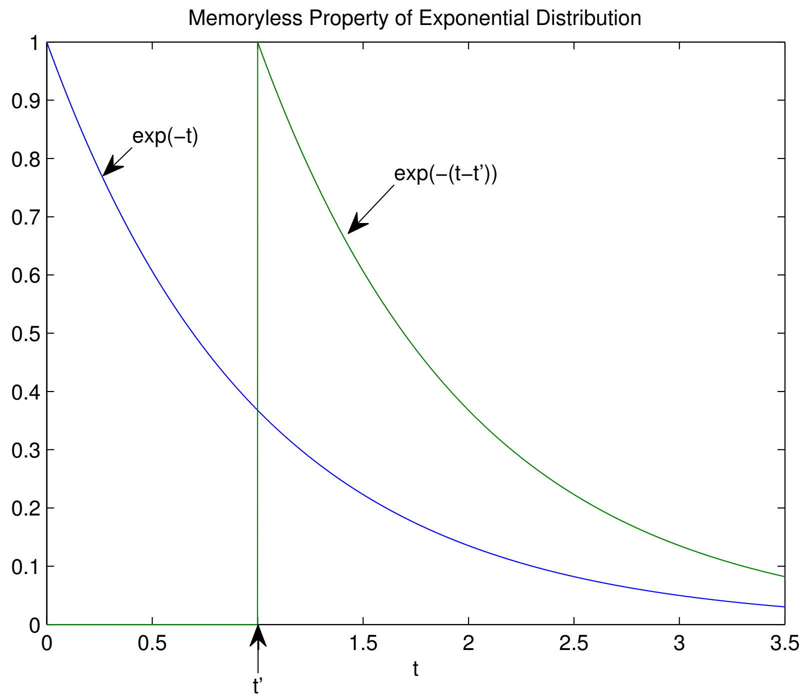

Figure 3: Plot showing the *right tail distribution function* $\text{Prob}\{X > t\}$ for the initial interarrival times and at time t'.