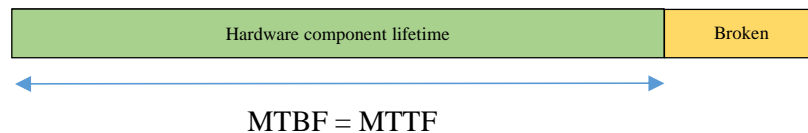


Network services availability and failure metrics within a time period

Mean Time Between Failures (MTBF)

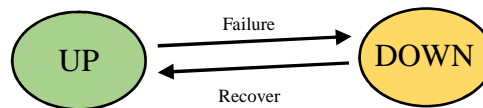
One of the most widely used metric to express something's reliability is MTBF, though different interpretations exist on MTBF.

Traditionally MTBF is used to express the expected lifetime of a hardware component. This assumes it is non-repairable, for instance, a lamp. In such a scenario, MTBF represents the same as the **Mean Time to Failure** (MTTF).

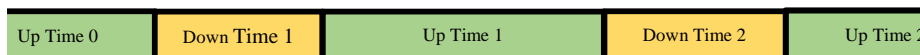


Typical behaviour for network services, on the other hand, is rather different. They are most often available but may become unavailable for short periods of time. In other words, they are repairable.

Under availability point of view, a service or host may be in one of two states: available or unavailable. For the sake of simplicity will call these two states, respectively, **UP** and **DOWN**.



Each time the service goes DOWN, it will be DOWN for a period of time and then will recover back to the UP state. The average duration of DOWN periods is called **Mean Down Time** (MDT), usually regarded as being the same as **Mean Time to Repair** (MTTR). One possible distinction between these two metrics is MTTR start counting only when the actual repair job starts. Otherwise, they are pretty the same.



MTTR (or MDT) represents the average down time periods duration:

$$MTTR = \frac{\sum_{i=1}^d DownTime_i}{d}, \quad d = \text{number of DOWN periods}$$

On the other hand, MTTF represents the average up time periods duration:

$$MTTF = \frac{\sum_{i=1}^u UpTime_i}{u}, \quad u = \text{number of UP periods}$$

We can express the service availability as:

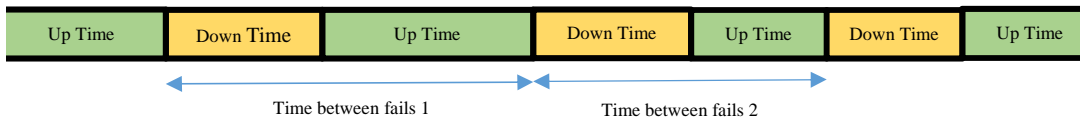
$$Availability = \frac{\sum_{i=1}^u UpTime_i}{\sum_{i=1}^u UpTime_i + \sum_{i=1}^d DownTime_i}$$

If, and only if, the number of DOWN periods is the same as the number of UP periods ($u = d$), then we can also express the service availability as:

$$Availability = \frac{MTTF}{MTTF + MTTR} ; \quad u = d$$

It's simply a matter of dividing both fraction members by the number of periods (u).

Analysing the service state along the time results now in a different interpretation for MTBF. A failure is a state transition from UP to DOWN, therefore MTBF is the average time between those transitions. The time between fails comprehends a DOWN period and the following UP period (DOWN/UP cycle).



Given a sequence of DOWN/UP cycles, we can express MTBF within that period as:

$$MTBF = \frac{\sum_{i=1}^f TimeBetweenFails_i}{f} , \quad f = \text{number complete DOWN /UP cycles}$$

If, and only if, we focus the analysis in a time period going from one fail start until another fail start, in other words, a sequence of complete DOWN/UP cycles, then:

$$MTBF = MTTR + MTTF ; \quad Availability = \frac{MTTF}{MTBF}$$

This is otherwise imprecise.

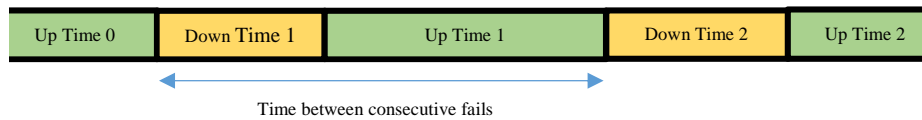
One other approach to calculate a value for MTBF in a given period is simply by dividing the total period time by the number of UP/DOWN transitions (failures).

$$MTBF = \frac{TotalTime}{f} , \quad f = \text{number of UP / DOWN transitions}$$

Availability analysis within a time period

The availability analysis within a time period presents some challenges. We must try to take the most of the provided data within the time window. Nevertheless, we must face the fact our data is incomplete, we will know the initial service state, but not for how long it has been in that state. Likewise, we will know the final service state, but not for how long it will remain in that state afterwards.

The time period start and time period end cannot be regarded as state transition points, this impacts on availability metrics calculation and inevitably introduces some inaccuracy. Of course, this inaccuracy can be reduced by enlarging as far as possible the analysis time period.



On the image above respecting a sample time period we can point three up time periods (in green), two down time periods (in orange) and four state transitions only. Notice that from the point of view of MTBF calculation, within this time interval, we have one single useable value.

Within this time period analysis, metrics should be calculated as follows:

Mean Time to Failure (MTTF) is the average up time:

$$MTTF = \frac{UpTime0 + UpTime1 + UpTime2}{3}$$

Mean Time to Repair (MTTR) or Mean Down Time (MDT) is average down time:

$$MTTR = \frac{DownTime1 + DownTime2}{2}$$

Availability is the relative total UP time, the exact value can be achieved by:

$$Availability = \frac{UpTime0 + UpTime1 + UpTime2}{UpTime0 + UpTime1 + UpTime2 + DownTime1 + DownTime2}$$

In this case, an exact value for availability cannot be calculated from MTTF and MTTR because the number of DOWN periods is not the same as the number of UP periods.

Mean Time Between failures (MTBF) – average time between consecutive failures (UP to DOWN transitions):

$$MTBF = DownTime1 + UpTime1$$

With the provided data, this is the best MTBF value we can calculate because we only have one complete DOWN/UP cycle.

We might be tempted to calculate availability by using this MTBF value, that would be:

$$Availability = \frac{UpTime1}{MTBF}$$

Yet, **this is a bad idea**, it will result in a rather different value because it is using only part of the provided data (Down Time 1 and Up Time 1), thus this value is less accurate.

A more representative value for MTBF can be considered by using all provided data:

$$MTBF = \frac{TotalTime}{fails} = \frac{UpTime0 + UpTime1 + UpTime2 + DownTime1 + DownTime2}{2}$$

Now availability calculated from MTBF comes as:

$$Availability = \frac{MTTF}{MTBF} = \frac{\frac{UpTime0 + UpTime1 + UpTime2}{3}}{\frac{UpTime0 + UpTime1 + UpTime2 + DownTime1 + DownTime2}{2}}$$

Nevertheless, we can see this value is two-thirds of the exact value previously calculated for availability. This error is due to the fact that in the time period there are three up time periods and only two failures. Of course, the longer is the analysis period the smaller will this error tend to be.