Mean Time Between Failure: Explanation and Standards

By Wendy Torell

Victor Avelar

White Paper #78

APC
Legendary Reliability®

Executive Summary

Mean Time Between Failure is a reliability term used loosely throughout many industries and has become widely abused in some. Over the years the original meaning of this term has been altered which has led to confusion and cynicism. MTBF is largely based on assumptions and definition of failure and attention to these details are paramount to proper interpretation. This paper explains the underlying complexities and misconceptions of MTBF and the methods available for estimating it.

Introduction

Mean Time Between Failure (MTBF) has been used for over 60 years as a basis for various decisions. Over the years more than 20 methods and procedures for lifecycle predictions have been developed. Therefore, it is no wonder that MTBF has been the daunting subject of endless debate. One area in particular where this is evident is in the design of mission critical facilities that house IT and telecommunications equipment. When minutes of downtime can negatively impact the market value of a business, it is crucial that the physical infrastructure supporting this networking environment be reliable. The business reliability target may not be achieved without a solid understanding of MTBF. This paper explains every aspect of MTBF using examples throughout in an effort to simplify complexity and clarify misconception.

What is a Failure? What are the Assumptions?

These questions should be asked immediately upon reviewing any MTBF value. Without the answers to these questions, the discussion holds little value. MTBF is often quoted without providing a definition of failure. This practice is not only misleading but completely useless. A similar practice would be to advertise the fuel efficiency of an automobile as "miles per tank" without defining the capacity of the tank in liters or gallons. To address this ambiguity, one could argue there are two basic definitions of a failure:

1) The termination of the ability of the product as a whole to perform its required function.[1]

2) The termination of the ability of any individual component to perform its required function but not the termination of the ability of the product as a whole to perform.[2]

The following two examples illustrate how a particular failure mode in a product may or may not be classified as a failure, depending on the definition chosen.

Example 1:
If a redundant disk in a RAID array fails, the failure does not prevent the RAID array from performing its required function of supplying critical data at any time. However, the disk failure does prevent a component of the disk array from performing its required function of supplying storage capacity. Therefore, according to definition 1, this is not a failure, but according to definition 2, it is a failure.

Example 2:
If the inverter of a UPS fails and the UPS switches to static bypass, the failure does not prevent the UPS from performing its required function which is supplying power to the critical load. However, the inverter failure does prevent a component of the UPS from performing its required function of supplying conditioned power. Similar to the previous example, this is only a failure by the second definition.

---

[1] IEC-50
[2] IEC-50

If there existed only two definitions, then defining a failure would seem rather simple. Unfortunately, when product reputation is on the line, the matter becomes almost as complicated as MTBF itself. In reality there are more then two definitions of failure, in fact they are infinite. Depending on the type of product, manufacturers may have numerous definitions of failure. Manufacturers that are quality driven track all modes of failure for the purpose of process control which, among other benefits, drives out product defects. Therefore, additional questions are needed to accurately define a failure.

Is customer misapplication considered a failure? There may have been human factors that designers overlooked, leading to the propensity for users to misapply the product. Are load drops caused by a vendor's service technician counted as a failure? Is it possible that the product design itself increases the failure probability of an already risky procedure? If an LED (Light Emitting Diode) on a computer were to fail is it considered a failure even though it hasn't impacted the operation of the computer? Is the expected wear out of a consumable item such as a battery considered a failure if it failed prematurely? Are shipping damages considered failures? This could indicate a poor packaging design. Clearly, the importance of defining a failure should be evident and must be understood before attempting to interpret any MTBF value. Questions like those above provide the bedrock upon which reliability decisions can be made.

It is said that engineers are never wrong; they just make bad assumptions. The same can be said for those who estimate MTBF values. Assumptions are required to simplify the process of estimating MTBF. It would be nearly impossible to collect the data required to calculate an exact number. However, all assumptions must be realistic. Throughout the paper common assumptions used in estimating MTBF are described.

Reliability, Availability, MTBF, and MTTR Defined

MTBF impacts both reliability and availability. Before MTBF methods can be explained, it is important to have a solid foundation of these concepts. The difference between reliability and availability is often unknown or misunderstood. High availability and high reliability often go hand in hand, but they are not interchangeable terms.

> *Reliability is the ability of a system or component to perform its required functions under stated conditions for a specified period of time [IEEE 90].*

In other words, it is the likelihood that the system or component will succeed within its identified mission time, with no failures. An aircraft mission is the perfect example to illustrate this concept. When an aircraft takes off for its mission, there is one goal in mind: complete the flight, as intended, safely (with no catastrophic failures).

> *Availability, on the other hand, is the degree to which a system or component is operational and accessible when required for use [IEEE 90].*

It can be viewed as the likelihood that the system or component is in a state to perform its required function under given conditions at a given instant in time. Availability is determined by a system's reliability, as well as its recovery time when a failure does occur. When systems have long continuous operating times (for example, a 10-year data center), failures are inevitable. Availability is often looked at because, when a failure does occur, the critical variable now becomes how quickly the system can be recovered. In the data center example, having a reliable system design is the most critical variable, but when a failure occurs, the most important consideration must be getting the IT equipment and business processes up and running as fast as possible to keep downtime to a minimum.

MTBF, or Mean Time Between Failure, is a basic measure of a system's reliability. It is typically represented in units of hours. The higher the MTBF number is, the higher the reliability of the product. Equation 1 illustrates this relationship.

$$\text{Reliability} = e^{-\left(\frac{Time}{MTBF}\right)}$$

Equation 1

A common misconception about MTBF is that it is equivalent to the expected number of operating hours before a system fails, or the "service life". It is not uncommon, however, to see an MTBF number on the order of 1 million hours, and it would be unrealistic to think the system could actually operate continuously for over 100 years without a failure. The reason these numbers are often so high is because they are based on the rate of failure of the product while still in their "useful life" or "normal life", and it is assumed that they will continue to fail at this rate indefinitely. While in this phase of the products life, the product is experiencing its lowest (and constant) rate of failure. In reality, wear-out modes of the product would limit its life much earlier than its MTBF figure. Therefore, there should be no direct correlation made between the service life of a product and its failure rate or MTBF. It is quite feasible to have a product with extremely high reliability (MTBF) but a low expected service life. Take for example, a human being:

> There are 500,000 25-year-old humans in the sample population.
> Over the course of a year, data is collected on failures (deaths) for this population.
> The operational life of the population is 500,000 x 1 year = 500,000 people years.
> Throughout the year, 625 people failed (died).
> The failure rate is 625 failures / 500,000 people years = 0.125% / year.
> The MTBF is the inverse of failure rate or 1 / 0.00125 = 800 years.
> *So, even though 25-year-old humans have high MTBF values, their life expectancy*
> *(service life) is much shorter and does not correlate.*

The reality is that human beings do not exhibit constant failure rates. As people get older, more failures occur (they wear-out). Therefore, the only true way to compute an MTBF that would equate to service life would be to wait for the entire sample population of 25-year-old humans to reach their end-of-life. Then, the

average of these life spans could be computed. Most would agree that this number would be on the order of 75-80 years.

So, what is the MTBF of 25-year-old humans, 80 or 800? It's both! But, how can the same population end up with two such drastically different MTBF values? It's all about assumptions!

If the MTBF of 80 years more accurately reflects the life of the product (humans in this case), is this the better method? Clearly, it's more intuitive. However, there are many variables that limit the practicality of using this method with commercial products such as UPS systems. The biggest limitation is time. In order to do this, the entire sample population would have to fail, and for many products this is on the order of 10-15 years. In addition, even if it were sensible to wait this duration before calculating the MTBF, problems would be encountered in tracking products. For example, how would a manufacturer know if the products were still in service if they were taken out of service and never reported?

Lastly, even if all of the above were possible, technology is changing so fast, that by time the number was available, it would be useless. Who would want the MTBF value of a product that has been superceded by several generations of technology updates?

MTTR, or Mean Time to Repair (or Recover), is the expected time to recover a system from a failure. This may include the time it takes to diagnose the problem, the time it takes to get a repair technician onsite, and the time it takes to physically repair the system. Similar to MTBF, MTTR is represented in units of hours. As Equation 2 shows, MTTR impacts availability and not reliability. The longer the MTTR, the worse off a system is. Simply put, if it takes longer to recover a system from a failure, the system is going to have a lower availability. The formula below illustrates how both MTBF and MTTR impact the overall availability of a system. As the MTBF goes up, availability goes up. As the MTTR goes up, availability goes down.

$$\text{Availability} = \frac{MTBF}{(MTBF + MTTR)}$$
Equation 2

For Equation 1 and Equation 2 above to be valid, a basic assumption must be made when analyzing the MTBF of a system. Unlike mechanical systems, most electronic systems don't have moving parts. As a result, it is generally accepted that electronic systems or components exhibit constant failure rates during the useful operating life. Figure 1, referred to as the failure rate "bathtub curve", illustrates the origin of this constant failure rate assumption mentioned previously. The "normal operating period" or "useful life period" of this curve is the stage at which a product is in use in the field. This is where product quality has leveled off to a constant failure rate with respect to time. The sources of failures at this stage could include undetectable defects, low design safety factors, higher random stress than expected, human factors, and natural failures. Ample burn-in periods for components by the manufacturers, proper maintenance, and proactive replacement of worn parts should prevent the type of rapid decay curve shown in the "wear out period". The discussion above provides some background on the concepts and differences of reliability and

availability, allowing for the proper interpretation of MTBF. The next section discusses the various MTBF prediction methods.

Figure 1 – Bathtub curve to illustrate constant rate of failures
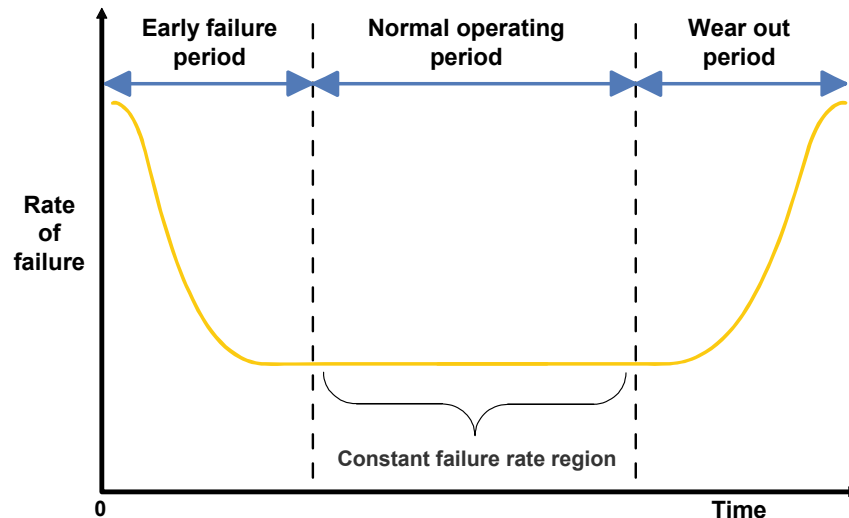
Methods of Predicting and Estimating MTBF

Oftentimes the terms "prediction" and "estimation" are used interchangeably, however this is not correct. Methods that *predict* MTBF, calculate a value based only on a system design, usually performed early in the product lifecycle. Prediction methods are useful when field data is scarce or non-existent as is the case of the Space Shuttle or new product designs. When sufficient field data exists, prediction methods should not be used. Rather, methods that *estimate* MTBF should be used because they represent actual measurements of failures. Methods that *estimate* MTBF, calculate a value based on an observed sample of similar systems, usually performed after a large population has been deployed in the field. MTBF estimation is by far the most widely used method of calculating MTBF, mainly because it is based on real products that are experiencing actual usage in the field.

All of these methods are statistical in nature, which means they provide only an approximation of the actual MTBF. No one method is standardized across an industry. It is, therefore, critical that the manufacturer understands and chooses the best method for the given application. The methods presented below, although not a complete list, illustrate the breadth of ways MTBF can be derived.

Reliability Prediction Methods

The earliest methods of reliability prediction came about in the 1940's with a German scientist named Von Braun and a German mathematician named Eric Pieruschka. While trying to improve numerous reliability problems with the V-1 rocket, Pieruschka assisted Von Braun in modeling the reliability of his rocket thereby creating the first documented modern predictive reliability model. Subsequently, NASA along with the

growth of the nuclear industry prompted additional maturation in the field of reliability analysis. Today, there are numerous methods for predicting MTBF.

MIL-HDBK 217

Published by the U.S. military in 1965, the Military Handbook 217 was created to provide a standard for estimating the reliability of electronic military equipment and systems so as to increase the reliability of the equipment being designed. It sets common ground for comparing the reliability of two or more similar designs. The Military Handbook 217 is also referred to as Mil Standard 217, or simply 217. There are two ways that reliability is predicted under 217: Parts Count Prediction and Parts Stress Analysis Prediction.

Parts Count Prediction is generally used to predict the reliability of a product early in the product development cycle to obtain a rough reliability estimate relative to the reliability goal or specification. A failure rate is calculated by literally counting similar components of a product (i.e. capacitors) and grouping them into the various component types (i.e. film capacitor). The number of components in each group is then multiplied by a generic failure rate and quality factor found in 217. Lastly, the failure rates of all the different part groups are added together for the final failure rate. By definition, Parts Count assumes all components are in series and requires that failure rates for non-series components be calculated separately.

Parts Stress Analysis Prediction is usually used much later in the product development cycle, when the design of the actual circuits and hardware are nearing production. It is similar to Parts Count in the way the failure rates are summed together. However, with Parts Stress, the failure rate for each and every component is individually calculated based on the specific stress levels the component is subjected to (i.e. humidity, temperature, vibration, voltage,). In order to assign the proper stress levels to each component, a product design and its expected environment must be well documented and understood. The Parts Stress Method usually yields a lower failure rate then the Parts Count Method. Due to the level of analysis required, this method is extremely time consuming compared to other methods.

Today, 217 is rarely used. In 1996 the U.S. Army announced that the use of MIL-HDBK-217 should be discontinued because it "has been shown to be unreliable, and its use can lead to erroneous and misleading reliability predictions"[3]. 217 has been cast off for many reasons, most of which have to do with the fact that component reliability has improved greatly over the years to the point where it is no longer the main driver in product failures. The failure rates given in 217 are more conservative (higher) then the electronic components available today. A thorough investigation of the failures in today's electronic products would reveal that failures were most likely caused by misapplication (human error), process control or product design.

Telcordia

The Telcordia reliability prediction model evolved from the telecom industry and has made its way through a

---

[3] Cushing, M., Krolewski, J., Stadterman, T., and Hum, B., 1996, "*U.S. Army Reliability Standardization Improvement Policy and Its Impact*", IEEE Transactions on Components, Packaging, and Manufacturing Technology, Part A, Vol. 19, No. 2, pp. 277-278.

series of changes over the years. It was first developed by Bellcore Communications Research under the name Bellcore as a means to estimate telecom equipment reliability. Although Bellcore was based on 217, its reliability models (equations) were changed in 1985 to reflect field experiences of their telecom equipment. The latest revision of Bellcore was TR-332 Issue 6, dated December 1997. SAIC subsequently bought Bellcore in 1997 and renamed it Telcordia. The latest revision of the Telcordia Prediction Model is SR-332 Issue 1, released in May 2001 and offers various calculation methods in addition to those of 217. Today, Telcordia continues to be applied as a product design tool within this industry.

HRD5

HRD5 is the Handbook for Reliability Data for Electronic Components used in telecommunication systems. HRD5 was developed by British Telecom and is used mainly in the United Kingdom. It is similar to 217 but doesn't cover as many environmental variables and provides a reliability prediction model that covers a wider array of electronic components including telecom.

RBD (Reliability Block Diagram)

The Reliability Block Diagram or RBD is a representative drawing and a calculation tool that is used to model system availability and reliability. The structure of a reliability block diagram defines the logical interaction of failures within a system and not necessary their logical or physical connection together. Each block can represent an individual component, sub-system or other representative failure. The diagram can represent an entire system or any subset or combination of that system which requires failure, reliability or availability analysis. It also serves as an analysis tool to show how each element of a system functions, and how each element can affect the system operation as a whole.

Markov Model

Markov modeling provides the ability to analyze complex systems such as electrical architectures. Markov models are also known as state space diagrams or state graphs. State space is defined as a collection of all of the states a system can be in. Unlike block diagrams, state graphs provide a more accurate representation of a system. The use of state graphs accounts for component failure dependencies as well as various states that block diagrams cannot represent, such as the state of a UPS being on battery. In addition to MTBF, Markov models provide various other measures of a system, including availability, MTTR, the probability of being in a given state at a given time and many others.

FMEA / FMECA

FMEA (Failure Mode and Effects Analysis) is a process used for analyzing the failure modes of a product. This information is then used to determine the impact each failure would have on the product, thereby leading to an improved product design. The analysis can go a step further by assigning a severity level to each of the failure modes in which case it would be called a FMECA (Failure Mode, Effects and Criticality Analysis). FMEA uses a bottom to top approach. For instance, in the case of a UPS, the analysis starts with the circuit board level component and works its way up to the entire system. Apart from being used as a product design tool, it can be used to calculate the reliability of the overall system. Probability data needed

for the calculations can be difficult to obtain for various pieces of equipment, especially if they have multiple states or modes of operation.

Fault Tree

Fault tree analysis is a technique that was developed by Bell Telephone Laboratories to perform safety assessments of the Minuteman Launch Control System. It was later applied to reliability analysis. Fault trees can help detail the path of events, both normal and fault related, that lead down to the component-level fault or undesired event that is being investigated (top to bottom approach). Reliability is calculated by converting a completed fault tree into an equivalent set of equations. This is done using the algebra of events, also referred to as Boolean algebra. Like FMEA, the probability data needed for the calculations can be difficult to obtain.

HALT

Highly Accelerated Life Testing (HALT) is a method used to increase the overall reliability of a product design. HALT is used to establish how long it takes to reach the literal breaking point of a product by subjecting it to carefully measured and controlled stresses such as temperature and vibration. A mathematical model is used to estimate the actual amount of time it would have taken the product to fail in the field. Although HALT can estimate MTBF, its main function is to improve product design reliability.

Reliability Estimation Methods

Similar Item Prediction Method

This method provides a quick means of estimating reliability based on historical reliability data of a similar item. The effectiveness of this method is mostly dependent on how similar the new equipment is to the existing equipment for which field data is available. Similarity should exist between manufacturing processes, operating environments, product functions and designs. For products that follow an evolutionary path, this prediction method is especially useful since it takes advantage of the past field experience. However, differences in new designs should be carefully investigated and accounted for in the final prediction.

Field Data Measurement Method

The field data measurement method is based on the actual field experience of products. This method is perhaps the most used method by manufacturers since it is an integral part of their quality control program. These programs are often referred to as Reliability Growth Management. By tracking the failure rate of products in the field, a manufacturer can quickly identify and address problems thereby driving out product defects. Because it is based on real field failures, this method accounts for failure modes that prediction methods sometimes miss. The method consists of tracking a sample population of new products and gathering the failure data. Once the data is gathered, the failure rate and MTBF are calculated. The failure rate is the percentage of a population of units that are expected to "fail" in a calendar year. In addition to using this data for quality control, it also is used to provide customers and partners with information about their product reliability and quality processes. Given that this method is so widely used by manufacturers, it provides a common ground for comparing MTBF values. These comparisons allow users to evaluate

relative reliability differences between products, which provide a tool in making specification or purchasing decisions. As in any comparison, it is imperative that critical variables be the same for all systems being compared. When this is not the case, wrong decisions are likely to be made which can result in a negative financial impact. For more information on comparing relative MTBF values see APC White Paper #112, "Performing Effective MTBF Comparisons for Data Center Infrastructure".

Conclusions

MTBF is a "buzz word" commonly used in the IT industry. Numbers are thrown around without an understanding of what they truly represent. While MTBF is an indication of reliability, it does not represent the expected service life of the product. Ultimately an MTBF value is meaningless if failure is undefined and assumptions are unrealistic or altogether missing.

References

1. Pecht, M.G., Nash, F.R., "*Predicting the Reliability of Electronic Equipment*", Proceedings of the IEEE, Vol. 82, No. 7, July 1994
2. Leonard, C., "*MIL-HDBK-217: It's Time To Rethink It*", Electronic Design, October 24, 1991
3. http://www.markov-model.com
4. MIL-HDBK-338B, Electronic Reliability Design Handbook, October 1, 1998
5. IEEE 90 – Institute of Electrical and Electronics Engineers, IEEE Standard Computer Dictionary: A Compilation of IEEE Standard Computer Glossaries. New York, NY: 1990

About the Authors:

Wendy Torell is an Availability Engineer with APC in W. Kingston, RI. She consults with clients on availability science approaches and design practices to optimize the availability of their data center environments. She received her Bachelors degree in Mechanical Engineering from Union College in Schenectady, NY. Wendy is an ASQ Certified Reliability Engineer.

Victor Avelar is an Availability Engineer for APC. He is responsible for providing availability consulting and analysis for clients' electrical architectures and data center design. Victor received a Bachelor's degree in Mechanical Engineering from Rensselaer Polytechnic Institute in 1995 and is a member of ASHRAE and the American Society for Quality.