XP Power

TECHNICAL ARTICLE

RELIABILITY IN ELECTRONICS

CONTENTS

- Introduction

1.1 Failure Rate
- 1.2 Reliability
- 1.3 Mean Time Between Failures (MTBF)
  Mean Time to Failure (MTTF)
- 1.4 Service Life (Mission Life, Life)

- Factors Affecting Reliability

2.1 Design Factors
- 2.2 Complexity
- 2.3 Stress

2.4 Generic (Inherent)

- Estimating The Failure Rate
  3.1 Prediction
  3.1.1 Parts Stress Method
  3.1.2 Parts Count Method
  3.2 Assessment
  3.2.1 Confidence Limits
  3.2.2 PRST
  3.3 Observation
- Prototype Testing
- Manufacturing Methods
- Systems Reliability
  (a) More Reliable Components
  (b) Redundancy
- Comparing Reliabilities

Introduction

Most of us are familiar with the concepts of reliability and MTBF at a superficial level, without considering what lies behind the figures quoted and what significance should be attached to them. The subject deserves a deeper understanding, so, let us start by having a better look at the terminology.

1.1 Failure Rate (λ)

The failure rate is defined as the percentage of units failing per unit time. This varies throughout the life of the equipment and if (lambda) is plotted against time, the characteristic "bathtub" curve is obtained for most electronic equipment (See Figure 1).

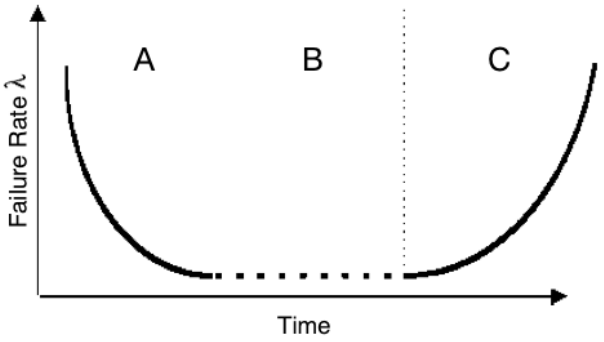

Fig 1. Failure Rate vs. Time

This curve has three regions:

- A Infant mortality.
- B Useful life.
- C Wear out

In region "A", poor workmanship and substandard components cause failures. This period is usually a few hundred hours and a "burn in" is sometimes employed to stop these failures occurring in the field. Note that this does not stop the failures occurring, it just ensures that they
happen in-house and not on the customer's premises.

In region "B", is approximately constant and it is only for this region that the following analysis applies.

In region "C", components begin to fail through having reached their end of life, rather than by random failures.

Examples are electrolytic capacitors drying out, fan bearings seizing up, switch mechanisms wearing out etc. Well implemented preventive maintenance can delay the onset of this region.

1.2 Reliability ($R_{(t)}$)

There are a large number of definitions and one will get different answers from statisticians, engineers, mathematicians and so on, an essentially practical definition is: The probability that a piece of equipment operating under specified conditions shall perform satisfactorily for a given period of time.

Probability is involved since it is impossible to predict the behaviour with absolute certainty. The criterion for "satisfactory performance" must be defined as well as the operating conditions such as input, output, temperature, load etc.

1.3 Mean Time Between Failures (MTBF), Mean Time To Failure (MTTF)

Strictly speaking, MTBF applies to equipment that is going to be repaired and returned to service, MTTF to parts that will be thrown away on failing.

The MTBF is the inverse of the failure rate.

$$MTBF = \frac{}{\lambda} \quad \text{....(1)}$$

Many people, unfortunately, misunderstand MTBF, and tend to assume that the MTBF figure indicates a minimum, guaranteed, time between failures. This assumption is wrong, and for this reason the use of the failure rate rather than the MTBF is highly recommended.

$$R_{(t)} = e^{-\lambda t} = e^{\frac{-t}{m}} \quad \text{....(2)}$$

$$m = \frac{t}{\text{Log}_n\left(\frac{1}{R_{(t)}}\right)} \quad \text{....(3)}$$

Where

$R_{(t)}$ = Reliability
e = Exponential (2.718)
λ = Failure Rate
m = MTBF

Note that for a constant failure rate, plotting reliability against time "t" gives a negative exponential curve (See Figure 2).

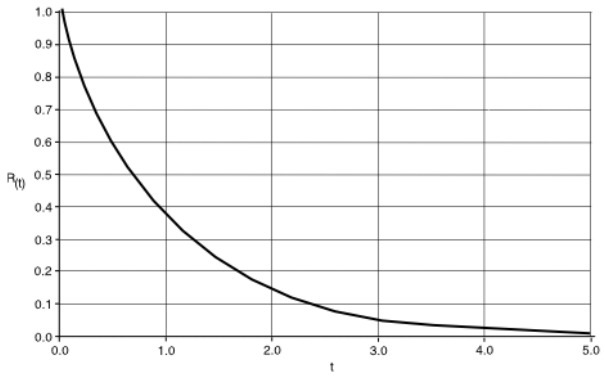

Fig 2. Reliability R(t) Plotted Against (λ t) for a unit with a constant failure rate λ

When t/m = 1, i.e., after a time “t”, numerically equal to the MTBF figure “m”:

$$R(t) = e^{-1} = 0.37 \quad \text{....(4)}$$

Equation (4) can be interpreted in a number of different ways:

(a) If a large number of units are considered, only 37% of them will survive for as long as the MTBF figure.
(b) For a single unit, the probability that it will work for as long as its MTBF figure, is only 37%.
(c) We can say that the unit will work for as long as its MTBF figure with a 37% Confidence Level.

In order to put these numbers into context, let us consider a power supply with a MTBF of 500,000 hours, (a failure rate of 0.2%/1000 hours), or as the advertising would put it "an MTBF of 57 years!"

From eq.(2), R(t) for 26,280 hours (3 years) is approximately 0.95, i.e., if such a unit is used 24 hours a day for 3 years, the probability of it surviving that time is 95%. The same calculation for a ten year period will give a R(t) of 84%.

Now let us consider a customer who has 700 such units. Since we can expect, on average, 0.2% of units to fail per 1000 hours, approximately one unit per month will fail on average, since the number of failures per year is:

$$\frac{0.2}{100} * \frac{1}{1000} * 700 * 24 * 365 = 12.26$$

1.4 Service Life (Mission Life, Life)

Note that there is no direct connection or correlation between service life and failure rate. It is perfectly possible to design a very reliable product with a short life. A typical example is a missile for example: it has to be very, very reliable (MTBF of several million hours), but its service life is only 0.06 hours (4 minutes). 25 year old humans have a MTBF of about 800 years, (FR about 0.1%/year), but not many have a comparable "service life". Just because something has a good MTBF, it does not necessarily have a long service life as well (See Figure 3).

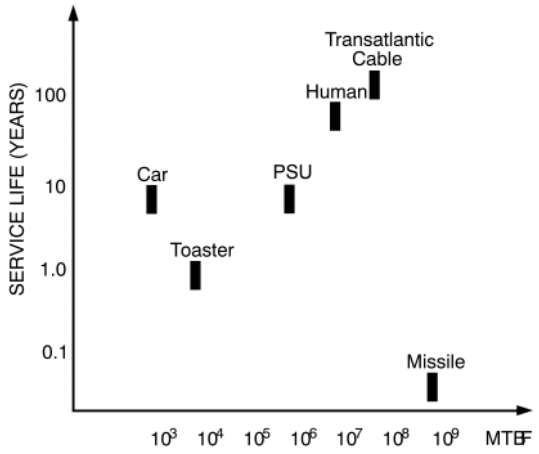

Fig 3. Examples of Service Life vs. MTBF

Factors Affecting Reliability

2.1 Design Factors

The most important factor is good, careful design based on sound experience, resulting in known safety margins. Unfortunately, this does not show up in any predictions, since they assume a perfect design!

It has to be said that a lot of field failures are not due to the classical random failure pattern discussed here, but to shortcomings in the design and in the application of the components, as well as external factors such as occasional voltage surges, etc. These may well be 'outside
specification' but no one will ever know, all that will be seen is a failed unit. Making the units rugged through careful design and controlled overstress testing is a very important part of making the product reliable.

The failure rate of the equipment depends on three other factors:

- Complexity
- Stress
- Inherent (generic) reliability of the components used

2.2 Complexity

Keep things simple - what isn't there, can't fail, but be careful: what isn't there can cause a failure! A complicated or difficult specification will, invariably result in reduced reliability. This is not due to the shortcomings of the design staff, but to the resultant component count. Every
component used will contribute to the equipment's unreliability.

2.3 Stress

In electronic equipment, the most prominent stresses are temperature, voltage, vibration, and temperature rise due to current. The effect of each of these stresses on each of the components must be considered. In order to achieve good reliability, various derating factors have to be applied to these stress levels. The derating has to be traded off against cost and size implications.

Great care and attention to detail is necessary to reduce thermal stresses as far as possible. The layout has to be such that heat-generating components are kept away from other components and are adequately cooled.

Thermal barriers are used where necessary and adequate ventilation needs to be provided. The importance of these provisions cannot be overstressed since the failure rate of some components will double for a 10 °C increase in temperature. Note that decreasing the size of a unit without increasing its efficiency will make it hotter, and therefore less reliable!

2.4 Generic (Inherent) Reliability

Inherent reliability refers to the fact that film capacitors are more reliable than electrolytic capacitors, wirewrap connections more reliable than soldered ones, fixed resistors more reliable than pots, and so on. Components have to be carefully selected to avoid the types with high generic failure rates. Quite often, there is a cost trade off - more reliable components are usually more expensive.

3. Estimating the Failure Rate

The Failure Rate should be estimated and measured throughout the life of the equipment:

- During design, it is predicted
- During manufacture, it is assessed
- During the service life, it is observed

3.1 Prediction

Predicting the failure rate is done by evaluating each of the factors effecting reliability for each component and then summing these to get the failure rate of the whole equipment. It is essential that the database used is defined and used consistently. There are three databases in common use: MIL-HDBK-217, HRD5 and Bellcore. These reflect the experiences of the US Navy, British Telecom and Bell Telephone, respectively.

Other sources of data are component manufacturers and some large companies like Siemens, Philips, France Telecom or Italtel. Data from these should not be used unless specifically requested by the customer.

In general, predictions assume that:

- The design is perfect, the stresses known, everything is within ratings at all times, so that only random failures occur
- Every failure of every part will cause the equipment to fail.
- The database is valid

These assumptions are wrong. The design is less than perfect, not every failure of every part will cause the

equipment to fail, and the database is likely to be at least 15 years out-of-date. However, none of this matters much, if the predictions are used to compare different topologies or approaches rather than to establish an absolute figure for reliability. This is what predictions should be used for.

3.1.1 Parts Stress Method

In this method, each factor affecting reliability for each component is evaluated. Since the average power supply has over 100 components and each component about 7 factors (Typically: stress ratio, generic, temperature, quality, environment, construction, and complexity) this method requires a considerable effort and time. Predictions are usually done in order to compare different approaches or topologies, i.e. when detailed design information is not available and the design itself is still in a fluid state. Under such circumstances, it is hardly worthwhile to spend this effort, and the much simpler and quicker Parts Count Method is used.

3.1.2 Parts Count Method

In this method, all like components are grouped together, and average factors allocated for the group. So, for example instead of working out all the factors for each of the 15 electrolytic capacitors used, there is only one entry of cap. electr.' and a quantity of 15. Usually only two factors are allocated: generic and quality. The other factors, including stress levels, are assumed to be at some realistic level and allowed for in the calculation. For this reason, the factors are not interchangeable between the two methods. In general, for power supplies, HRD5 gives the most favourable result, closely followed by Bellcore, with MIL-HDBK-217F the least favourable. This depends on

Great care and attention to detail is necessary to reduce thermal stresses as far as possible. The layout has to be such that heat-generating components are kept away from other components and are adequately cooled.

Thermal barriers are used where necessary and adequate ventilation needs to be provided. The importance of these provisions cannot be overstressed since the failure rate of some components will double for a 10 °C increase in temperature. Note that decreasing the size of a unit without increasing its efficiency will make it hotter, and therefore less reliable!

2.4 Generic (Inherent) Reliability

Inherent reliability refers to the fact that film capacitors are more reliable than electrolytic capacitors, wirewrap connections more reliable than soldered ones, fixed resistors more reliable than pots, and so on. Components have to be carefully selected to avoid the types with high generic failure rates. Quite often, there is a cost trade off - more reliable components are usually more expensive.

3. Estimating the Failure Rate

The Failure Rate should be estimated and measured throughout the life of the equipment:

- During design, it is predicted
- During manufacture, it is assessed
- During the service life, it is observed

3.1 Prediction

Predicting the failure rate is done by evaluating each of the factors effecting reliability for each component and then summing these to get the failure rate of the whole equipment. It is essential that the database used is defined and used consistently. There are three databases in common use: MIL-HDBK-217, HRD5 and Bellcore. These reflect the experiences of the US Navy, British Telecom and Bell Telephone, respectively.

Other sources of data are component manufacturers and some large companies like Siemens, Philips, France Telecom or Italtel. Data from these should not be used unless specifically requested by the customer.

In general, predictions assume that:

- The design is perfect, the stresses known, everything is within ratings at all times, so that only random failures occur
- Every failure of every part will cause the equipment to fail.
- The database is valid

These assumptions are wrong. The design is less than perfect, not every failure of every part will cause the

equipment to fail, and the database is likely to be at least 15 years out-of-date. However, none of this matters much, if the predictions are used to compare different topologies or approaches rather than to establish an absolute figure for reliability. This is what predictions should be used for.

3.1.1 Parts Stress Method

In this method, each factor affecting reliability for each component is evaluated. Since the average power supply has over 100 components and each component about 7 factors (Typically: stress ratio, generic, temperature, quality, environment, construction, and complexity) this method requires a considerable effort and time. Predictions are usually done in order to compare different approaches or topologies, i.e. when detailed design information is not available and the design itself is still in a fluid state. Under such circumstances, it is hardly worthwhile to spend this effort, and the much simpler and quicker Parts Count Method is used.

3.1.2 Parts Count Method

In this method, all like components are grouped together, and average factors allocated for the group. So, for example instead of working out all the factors for each of the 15 electrolytic capacitors used, there is only one entry of cap. electr.' and a quantity of 15. Usually only two factors are allocated: generic and quality. The other factors, including stress levels, are assumed to be at some realistic level and allowed for in the calculation. For this reason, the factors are not interchangeable between the two methods. In general, for power supplies, HRD5 gives the most favourable result, closely followed by Bellcore, with MIL-HDBK-217F the least favourable. This depends on

the mix of components in the particular equipment, since one database maybe "unfair" on ICs, and an other on FETs. Hence, the importance of comparing results from like databases only.

3.2 Assessment

This is the most useful and accurate way of predicting the Failure Rate. A number of units are put on "life test" (more correctly described as a Reliability Demonstration Test), usually at an elevated temperature, and so the stresses and the environment is controlled. Note, however, that it is not always possible to model the real environment accurately in the laboratory.

During life-tests and reliability demonstration tests, it is usual to apply greater stresses than normal, so that we get to the desired result quicker. Great care has to be applied to ensure that the effects of the extra stress is known and proven to be calculable, and that no hidden, additional failure mechanisms are activated by the extra stress. The usual "extra stress" is an increase of temperature, and its effect can be calculated from the Arrhenius equation, as long as the maximum ratings of the device are not exceeded.

Note that the accelerating effect depends on the activation energy that applies for the chemistry of the particular component. This would indicate that the Acceleration Factor from 25 °C to 50 °C is approx. 5.25, so be suspicious of results based on 0.7 eV, or even 1 eVn.

At the beginning of such a test it is sometimes difficult to distinguish between early failures ("infant mortality", region A) and the first failures belonging to the "constant failure rate" region (region B). In such cases, the Cumulative Distribution Function is plotted on Weibull paper. This paper has double logarithmic scaling such that a constant failure rate will result in a straight line at an indicated gradient of 1. Decreasing FR (region A) will give a smaller gradient, increasing FR (wearout, region C) a higher gradient. Both the available time and the number of units on test are limited, and so it is of the utmost importance that the maximum amount of useful information is extracted from a limited amount of data. Statistical methods are used to achieve this.

3.2.1 Confidence limits

What we are attempting to do is to predict the behaviour of the large number of units in the field (called the population) from the behaviour of a small number of randomly selected units (called the sample). This process is called Statistical Inference. The results obtained by such means cannot, of course, be completely accurate, and it is therefore essential to establish the degree of accuracy that applies. This is done by estimating the mean value and defining a band or an interval around this estimated mean that will include the actual, true mean value of the complete population. Such an interval is defined by a Confidence Limit, i.e., if we establish that the failure rate is between 1%/1000 hours and 2%/1000 hours with a Confidence Limit of 90%, this means that we expect 90% of the units in the field to exhibit failure rates between these limits, and the other 10% of units to have a lower or higher failure rate. For a population exhibiting a constant failure rate,

$$\lambda = \frac{X^2 (2r+2), (1-\phi)}{2tN} \quad \ldots.(5)$$

where λ = demonstrated failure rate with a one-sided higher confidence limit of Ø (phi)

t = test time
N = number of units on test
r = number of failures
$X^2 (2r+2), (1-Ø)$ = value of the X^2 distribution

r	Ø = 0.6	Ø = 0.9
0	93×10^3	230×10^3
1	200×10^3	390×10^3
2	310×10^3	530×10^3
3	420×10^3	670×10^3
4	530×10^3	790×10^3
5	630×10^3	910×10^3
6	730×10^3	1040×10^3
7	830×10^3	1160×10^3
8	930×10^3	1300×10^3
9	1040×10^3	1410×10^3
10	1140×10^3	1530×10^3

with probability (1 - Ø) of not being exceeded in random sampling where (2r + 2) is the number of degrees of freedom.

The constants given by this equation are tabulated below for values of r between 0 and 10, and for of 0.6 and 0.9. (These are the usual values for Confidence Limits used in industry).

To use this table divide the factor given by the total number of unit-hours to get the failure rate in %/1000 hours. Let us consider the case when we have 50 units on test and one fails after 4 months (2920 hours):

t = 2920, N = 50, r = 1

From the table, we can say with 60% confidence that the failure rate will be less than:

$$\frac{200{,}000}{50 \times 2920} = 1.37\%/1000 \text{ hrs}$$

Alternatively, we can say with 90% confidence that the failure rate will be less than:

$$\frac{390{,}000}{50 \times 2920} = 2.67\%/1000 \text{ hrs}$$

In the parent population, therefore, we expect 60% of the units to exhibit a failure rate better than 1.37%/1000 hrs (an MTBF of 73,000 hrs.), and therefore 40% of units to have a FR worse than that; or 90% of the units to be better than 2.67%/1000hrs (an MTBF of 37,400 hrs.), and
therefore 10% of units to have a FR worse than that (See Figure 4).

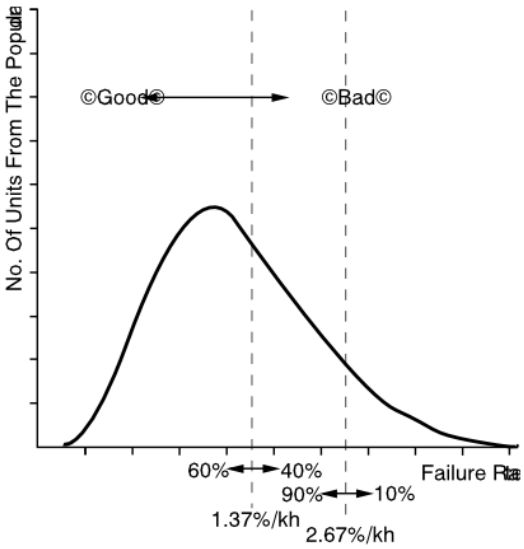

Fig 4. Confidence Limit

However, there is a practical problem with this method: although we get valid answers, the length of time for that answer is a function of

values for Confidence Limits used in industry).

To use this table divide the factor given by the total number of unit-hours to get the failure rate in %/1000 hours. Let us consider the case when we have 50 units on test and one fails after 4 months (2920 hours):

t = 2920, N = 50, r = 1

From the table, we can say with 60% confidence that the failure rate will be less than:

$$\frac{200{,}000}{50 \times 2920} = 1.37\%/1000 \text{ hrs}$$

Alternatively, we can say with 90% confidence that the failure rate will be less than:

$$\frac{390{,}000}{50 \times 2920} = 2.67\%/1000 \text{ hrs}$$

In the parent population, therefore, we expect 60% of the units to exhibit a failure rate better than 1.37%/1000 hrs (an MTBF of 73,000 hrs.), and therefore 40% of units to have a FR worse than that; or 90% of the units to be better than 2.67%/1000hrs (an MTBF of 37,400 hrs.), and
therefore 10% of units to have a FR worse than that (See Figure 4).

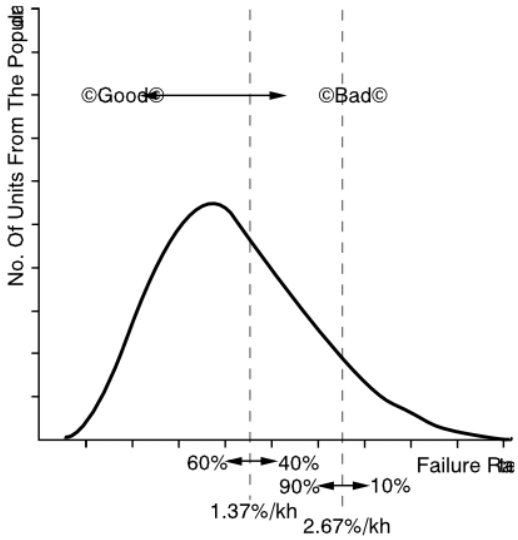

Fig 4. Confidence Limit

However, there is a practical problem with this method: although we get valid answers, the length of time for that answer is a function of

1. The sheer mechanics of actually collecting and collating the data.
2. The uncertainty of the duty, conditions of use and stresses, or abuse, that the units were subjected to.

Great care has to be exercised in drawing conclusions due to the difficulty of distinguishing between true random failures and misuse in the field (accidental or otherwise).

4. Prototype Testing

With all the sophisticated computer analysis, simulation and tolerancing methods available, there is still no substitute for thoroughly testing the maximum number of prototypes. An effort should be made to locate and use components from different batches, especially for critical components. These units must be tested under dynamic conditions to ensure reliability. An effective test is to cycle the temperature, the input, and the load independently. The units should be tested at both maximum and minimum temperature cycling according to this plan. Cpk analysis of the results is used to ensure that the specification parameter margins are adequate. After testing, these units are normally used as the first batch on the reliability demonstration tests.

At least one unit should be subjected to HALT testing, and several to destructive overstress tests to establish the safety margins.

The timing of these tests is critical - it must not be so early in the development phase that the final circuit is radically different, and it must not be so late that production starts before the results are evaluated. A pitfall to watch out for, if changes are proposed as a result of these tests is that the up-dated units must be subject to long term testing themselves.

5. Manufacturing Methods

This is a separate subject in itself, but there are three main factors contributing to unreliability in manufacture:

- Suppliers
- Manual assembly methods
- Tweaking of settings and parameters

Suppliers must be strictly controlled to deliver consistently good devices, with prior warning of any process changes and any other changes. These days, with modern QA practices and JIT manufacturing methods, this level of reliability is achieved by dealing with a small number of trusted suppliers. Manual assembly is prone to errors and to some random, unintentional abuse of the components by operators. This creates latent defects, which show up later.

Tweaking produces inconsistency and side effects. A good motto is: if it works, leave it alone; if it does not, find the root cause and do not tweak. There must be a root cause for the deviation, and this must be found and eliminated, rather than masked by the tweak. There are well-established TQM and SPC methods to achieve this. Testing and Quality Assurance has a major part to play. Testing must be appropriate to ensure that the units perform well in the application. Cpk analysis ensures that the specification parameter margins are adequate and
controlled.

6. System Reliability

There are two further methods of increasing system reliability. Firstly, more reliable components. MIL standard or other components of assessed quality could be used, but in industrial and commercial equipment, the expense is not normally justified.

Secondly, redundancy. In a system where one unit can support the load, and two units are used in parallel, the system is much more reliable since the system will still work even with one unit failed. Clearly, the probability of two units failing simultaneously is much less than that of one unit failing. This system would have a big size and cost penalty, (twice as big and twice as much) so normally an N+1 system is used, where N units can support the load, but N+1 units are used in parallel, "2+1" or "3+1" being the usual combinations. Supposing the reliability of each unit under the particular conditions is 0.9826, (m=500,000h, t=1year) the system reliability for an "N+1" system where N = 2 would be 0.9991, an improvement of 20 times. (Nearly 60 times in a 1+1 system).

However, there are many pitfalls in the system design, such as:

1. N units must be rated to support full load.
2. Any part failing must not make the system fail.
3. If any part fails this must be brought to the operator's notice so that it can be replaced.
4. Changing units must not make the system fail (hot plugging).

It is very difficult and tricky to design the system to satisfy items 2 & 3. For example the failure of components that do not effect system

1. The sheer mechanics of actually collecting and collating the data.
2. The uncertainty of the duty, conditions of use and stresses, or abuse, that the units were subjected to.

Great care has to be exercised in drawing conclusions due to the difficulty of distinguishing between true random failures and misuse in the field (accidental or otherwise).

4. Prototype Testing

With all the sophisticated computer analysis, simulation and tolerancing methods available, there is still no substitute for thoroughly testing the maximum number of prototypes. An effort should be made to locate and use components from different batches, especially for critical components. These units must be tested under dynamic conditions to ensure reliability. An effective test is to cycle the temperature, the input, and the load independently. The units should be tested at both maximum and minimum temperature cycling according to this plan. Cpk analysis of the results is used to ensure that the specification parameter margins are adequate. After testing, these units are normally used as the first batch on the reliability demonstration tests.

At least one unit should be subjected to HALT testing, and several to destructive overstress tests to establish the safety margins.

The timing of these tests is critical - it must not be so early in the development phase that the final circuit is radically different, and it must not be so late that production starts before the results are evaluated. A pitfall to watch out for, if changes are proposed as a result of these tests is that the up-dated units must be subject to long term testing themselves.

5. Manufacturing Methods

This is a separate subject in itself, but there are three main factors contributing to unreliability in manufacture:

- Suppliers
- Manual assembly methods
- Tweaking of settings and parameters

Suppliers must be strictly controlled to deliver consistently good devices, with prior warning of any process changes and any other changes. These days, with modern QA practices and JIT manufacturing methods, this level of reliability is achieved by dealing with a small number of trusted suppliers. Manual assembly is prone to errors and to some random, unintentional abuse of the components by operators. This creates latent defects, which show up later.

Tweaking produces inconsistency and side effects. A good motto is: if it works, leave it alone; if it does not, find the root cause and do not tweak. There must be a root cause for the deviation, and this must be found and eliminated, rather than masked by the tweak. There are well-established TQM and SPC methods to achieve this. Testing and Quality Assurance has a major part to play. Testing must be appropriate to ensure that the units perform well in the application. Cpk analysis ensures that the specification parameter margins are adequate and
controlled.

6. System Reliability

There are two further methods of increasing system reliability. Firstly, more reliable components. MIL standard or other components of assessed quality could be used, but in industrial and commercial equipment, the expense is not normally justified.

Secondly, redundancy. In a system where one unit can support the load, and two units are used in parallel, the system is much more reliable since the system will still work even with one unit failed. Clearly, the probability of two units failing simultaneously is much less than that of one unit failing. This system would have a big size and cost penalty, (twice as big and twice as much) so normally an N+1 system is used, where N units can support the load, but N+1 units are used in parallel, "2+1" or "3+1" being the usual combinations. Supposing the reliability of each unit under the particular conditions is 0.9826, (m=500,000h, t=1year) the system reliability for an "N+1" system where N = 2 would be 0.9991, an improvement of 20 times. (Nearly 60 times in a 1+1 system).

However, there are many pitfalls in the system design, such as:

1. N units must be rated to support full load.
2. Any part failing must not make the system fail.
3. If any part fails this must be brought to the operator's notice so that it can be replaced.
4. Changing units must not make the system fail (hot plugging).

It is very difficult and tricky to design the system to satisfy items 2 & 3. For example the failure of components that do not effect system

operation when all units are OK, but would effect operation if there was a fault (such as an isolating diode going short circuit, or a parallelling wire or connector going open circuit), must be signalled as a problem, and must be repaired. The circuitry necessary to arrange for all this, (isolating diodes, signalling logic, hot plugging components, current sharing, etc.) has its own failure rate, and so degrades the overall system failure rate. In the following illustrations, this is ignored for simplicity, but in a real calculation, it must be taken into account. In many applications, the only way to detect such latent faults is to simulate a part failing by shutting it down remotely for a very short time. This circuitry will, of course, increase complexity and decrease reliability further still, as well as being dangerous: a system failure could be caused by the test circuit shutting the system down.

Calculating system reliability involves the use of the binomial expansion, as follows:

$$(R + Q)^T = [(R^T + TR^{(T-1)} Q + (T(T - 1)/2!) R^{(T-2)} Q^2 + (T (T - 1) (T - 2)/3!) R^{(T-3)} Q^3 + ... + Q^T] \quad ....(6)$$

where:

T = Total No. of Units

R = Probability of Success

Q = Probability of Failure = (1-R)

The 1st term is the probability that 0 units will fail,

The 2nd term is the probability that 1 unit will fail,

The 3rd term is the probability that 2 units will fail,

The 4th term is the probability that 3 units will fail,

The 5th term is the probability that 4 units will fail, ... and so on.

These terms must be summed as appropriate, based on what combination of part failures gives a system failure.

For example, with 4 units of R = 0.8, the probability of failures is:

0 failures	:	0.8^4	= 0.4096
1 failure	:	$4 \times 0.8^3 \times 0.2$	= 0.4096
2 failures	:	$(4 \times 3/2) \times 0.8^2 \times 0.2^2$	= 0.1536
3 failures	:	$(4 \times 3 \times 2/3 \times 2) \times 0.8^1 \times 0.2^3$	= 0.0256
4 failures	:	0.2^4	= 0.0016

So if 1 unit is enough to supply the load, then if there are 0 Failures, or 1F, or 2F, or 3F, the system is still working, hence the system reliability is:

0.4096 + 0.4096 + 0.1536 + 0.0256 = 0.9984

This particular result could have been obtained from special case 2 (any one is OK, this would be a "n+3" system): $1 - 0.2^4 = 0.9984$

If two units are needed to maintain the system, then only 0, 1 and 2 failures are OK (this would be a "n+2" system):

The system reliability is: 0.4096 + 0.4096 + 0.1536 = 0.9728

If three units are needed to maintain the system, then only 0 and 1 failures are OK:

The system reliability is: 0.4096 + 0.4096 = 0.8192

This particular result could have been obtained from special case 1 ("n+1"): $0.8^4 + 0.2 \times 4 \times 0.8^3 = 0.8192$

Note that the improvement over one unit is only marginal for such a low reliability (0.8), however this is an effective solution in cases where R > 0.9. If there is no redundancy, the only acceptable case is that of 0 failures: $0.8^4 = 0.4096$. This particular case is the same as the series situation (any part failure causes a system failure).

Special case 1: "n+1" redundancy, identical units.

In this case, 0F and 1F will not cause a system failure, and the reliability is given by the sum of the first two terms of the expansion:

$$R_T = R^T + QT\{R^{(T-1)}\} \quad ....(7)$$

Special case 2: Redundancy where any one unit is capable of supplying the load:

$$R_T = 1-[(1-R_A)(1-R_B)\overline{(1-R_C)(1-R_D)}] \quad ....(8)$$

Parts in series:

(any part failure will cause a system failure)

$$R_T = [(R_A)(R_B)(R_C)(R_D) \ldots] \quad \ldots.(9)$$

Availability

Availability is sometimes mentioned in this context, this is defined as:

$$\text{Availability} = \frac{\text{MTBF}}{\text{MTBF} + \text{MTTR}}$$

Where MTTR is the mean time to repair.---

For good, reliable systems, Availability tends to be 0.99999....., where the mathematics gets tedious and the number difficult to interpret.

In such cases Unavailability is more meaningful, this being (1- Availability) and usually expressed in minutes/year.
Consider the previous example (m=500,000h, t=1year), and assume that MTTR is 3 hours.
Availability is 0.999 994, and Unavailability is 0.000 006 or 3.15 minutes/year.
Now consider the "N+1" system shown (N=2).
Availability will be 0.999 999 694, and Unavailability 0.000 000 306 or 10 seconds/year.

Note however, that we now have 3 units in the system, so service calls will be 3 times as frequent, or in other words the MTBF for service calls = 500,000/3 = 166,700hours.
It is an interesting fact that when using redundancy to improve availability, the service calls to repair system failures gets much less frequent, but the service calls to repair part failures gets more frequent. Since the object of the exercise is to maintain system availability, this is a small price to pay, but the costs of system failure should be weighed against the costs of service maintenance.
In some cases it is possible to either reduce costs or improve system availability further by partitioning, i.e. have different load-groups fed by different power-supply-groups. This is a subject in itself, but as an illustration the level of redundancy in a typical telephone exchange is as follows:

- Each switching card is powered by 1+1 redundant dc/dc inverters.
- Each card is duplicated in 1+1 redundancy
- Each bay and its supplies are partitioned.
- The AC/DC supplies feeding a bay are 1+1 redundant.
- The power cables and connections are 1+1 redundant.
- There is a battery backup system at the output of the ac/dcs, feeding independent busbars.
- There is a diesel generator system to back up the mains supply.

The usual design criteria is that since batteries are large, expensive, dangerous and require maintenance, only about 20 minutes of battery backup is provided, which gives enough time for several attempts to start up the diesel generator. (An automatic sequence of 10 attempts). Since there is, on average, a short failure of the mains every week (MTBF of 170 hours (!)), this is a very necessary precaution.

7. Comparing Reliabilities

The real use of reliability predictions is not for establishing an accurate level of reliability, but for comparing different technical approaches, possibly from different manufacturers, on a relative (comparative) bases. Hence the importance of using the same database, environment etc.

When such comparisons are made, always check that all of the following are satisfied, otherwise the comparison is completely meaningless:

- The database must be stated, and must be identical. Comparing a MIL-HDBK-217F prediction with a MIL-HDBK-217E prediction or an HRD5 prediction is meaningless – there is no correlation.
- The database must be used consistently and exclusively. The result is meaningless if a different database is used for some component. The justification may be reasonable, but the result is meaningless.
- The external stresses and environment must be stated and must be identical. (Input, load, temperature, etc.) The result is meaningless if all the environmental details are not stated, or are different.
- The units must be FFF interchangeable in the application. If one is rated at 10A and the other at 5A, the comparison is fair, as long as the load is less than 5A. If the ratings are identical, but one needs an external filter and the other does not, then there is no comparison. (Although, it is possible, sometimes, to work out the failure rate of the external filter and add it to the FR of the unit, using the same database, environment and stress.)
- Comparing a predicted reliability figure with the results of a reliability demonstration test (lifetest) is also meaningless. One could argue that the results of the reliability demonstration are more meaningful, but that depends on the details of the test, the environment and the acceleration factors used. All these factors must be identical when comparing two test results, but in any case comparing test results with predictions is a meaningless comparison.

There are no miracles: if we predict 200,000 hours and an other manufacturer states 3,000,000 hours for a comparable product, then they must have used either a different database, or a different stress level, or a different environment, etc.