HOW MANY SPARE DOES ONE REALLY NEED?

Tony Frisch
Email: tony.frisch@xtera.com

Xtera Communications, Bates House, Church Road, Harold Wood, RM7 0SD, UK

Abstract:
Given current financial conditions the need to have sufficient spares to meet required service levels needs to be balanced against the cost of procuring them. In theory the "correct" quantity of spares can be derived from the number of units, their FIT values, the Mean Time to Repair and Return and the target service level. In practice, the process can be hard to interpret, as FIT values are seldom known precisely, except in the case of units which have been in use for a long time. In the case of recently developed units, the field-proven FIT value may be high simply due to the small number of deployed device-hours, yet with time the unit could prove to be highly reliable – or not! This paper briefly discusses some of the methodology that can be applied to this problem, but aims to focus more on practicality than mathematics.

1. INTRODUCTION

High reliability has always been a feature of submarine systems and today's Service Level Agreement (SLA) environment make reliability/availability a commercial issue which needs serious thought. The huge cost and time impact associated with marine repairs make the reliability of submerged plant the most obvious concern, while failures in terminal equipment can be put right rapidly if a spare unit is available. The consequence of needing a spare unit but not having one available is a long outage that continues until the failed unit is repaired or replaced. Given this, it is clear that the number of spares needs to be considered quite carefully; purchasing a large number can reduce the risk, but adds expense. This paper discusses some of the ways that one might determine the "correct" number of spares for each type of unit.

2. CALCULATION

Nearly all calculations assume a constant failure rate (λ) and a well defined time (T_{rep}) for a failed unit to be repaired. This ignores the effects of infant mortality (which can be largely eliminated by burn-in) and wear-out (which is usually unknown) and is a pragmatic choice which allows relatively simple statistical techniques to be applied. This paper makes these assumptions and uses the following definitions and formulae:

λ is the failure rate of the unit (often expressed as a FIT value or rate, where F fits is an average of F unit failures per 10^9 device hours)

μ is the expected number of failures; for N devices and a time period of T $\mu = \lambda N T$

P(j) is the probability of j failures, in this case given by the Poisson formula $P(j) = e^{-\mu} \mu^j / j!$

There has been a great deal of rapid development in DWDM terminal equipment during the last few years, with several new components being introduced. This means that usually both life test data and field experience may be somewhat limited, so one aim of this paper is to investigate how accurately one can know the failure rate and how best to handle any resultant uncertainty.

3. ACCURACY OF DETERMINING FAILURE RATES

In general failure rates are measured on the basis of laboratory tests or from field experience where one has a number of devices (N) of which there are a known number of failures (F) during a defined period of time (T). Given sufficiently large numbers the failure rate is F/(NT), but a lack of both time and components often means that the numbers of failures is relatively small, thus making the accuracy of determination poor. The following graph shows the relative probability of that the failure rate differs from F/(NT) by a factor X.

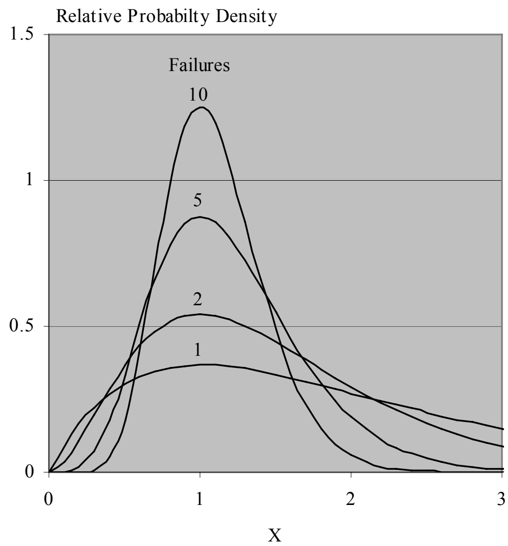

Figure 1: Accuracy of failure rate

From this it is clear that even with 10 failures there is considerable uncertainty in the failure rate; with fewer failures there is a lot of uncertainty. Given this uncertainty it is common to quote/require the 95% Upper Confidence Limit (UCL), which is typically quite a bit larger than the most likely value, as can be seen from the following table.

Failures	1	2	5	10
Multiplier	4.74	3.15	2.10	1.70

Table 1

Using the 95% UCL value provides a safety margin, but generally overestimates the true failure rate by a factor of 2 or more. A further problem is that components/units are usually manufactured in batches and occasionally a batch with lower reliability is produced.

4. TERMINALS AND NETWORKS

Terminal equipment contains a number of traffic-affecting units of widely varying complexity/reliability. A simplified diagram is shown below, with very approximate values for the reliabilities.

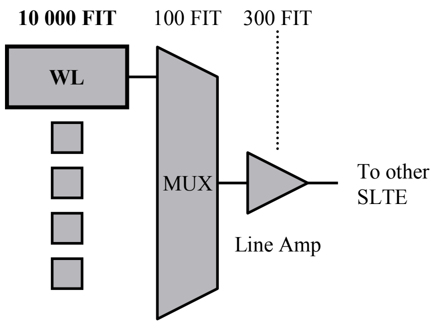

Figure 2: Simplified terminal schematic

The unit labeled WL is the one which contains client interfaces, FEC, laser, modulator, receiver etc. and is thus the most complex and expensive item in the SLTE. It is usually the most significant in terms of sparing calculations.

5. CALCULATION

Most calculations take the approach of setting a target in terms of (1) confidence of having spares available, or (2) meeting an availability target, and then searching for the minimum number of spares needed to achieve the target. There is no simple mathematical formula for this number, although the inverse Poisson function can be applied for simple cases. In general, however, the analysis should be capable of addressing complexities such as transferring spares from one station to another. This requires a slightly more complex model.

For simplicity let us first consider just one type of unit and a site where there are N operating units and K spares. (Handling all the types of unit is then simply a matter of repeating the process with different parameter values.) A useful approach is to consider this as a system which at any point in time is described by the number of units which have failed and are still being repaired. In the following diagram the circles represent these states, with failures shown by solid lines and unit returns by dotted lines.

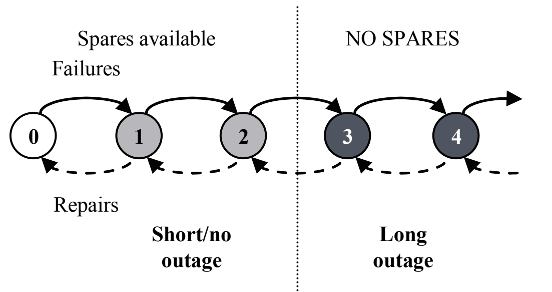

Figure 3: Example state diagram for the case of a location with 2 spares

If there is no protection switching then the first failure results in an outage equal to the average unit replacement time. Once there are no spare units available a failure will produce an outage which lasts until a repaired spare is returned – potentially a long time. Protection switching greatly improves the situation, as there will be no outage until the second failure after all the spares have been sent for repair.

There are a number of ways to analyze this type of system, but most evaluate the probability of being in the different states. One good and flexible approach is to use a computer simulation; this has the benefit that it becomes possible to handle repair times which vary. A potential issue with simulations is that it is hard for anyone other than the person who created the model to check its validity. An alternative, which still offers flexibility with relative ease of validation is a Markov model, using the steady-state equilibrium principle [1]. In this case one equates a repair time to a transition rate; for example if the repair time is 50 days, one would use a repair rate of 1/50 repairs per day. While this might seem reasonable, it is not intuitively obvious that it will yield the correct results. The figure below compares the steady state Markov calculation and a simulation for a location with 16 units of 10,000 FIT reliability and a 60 day repair and return time.

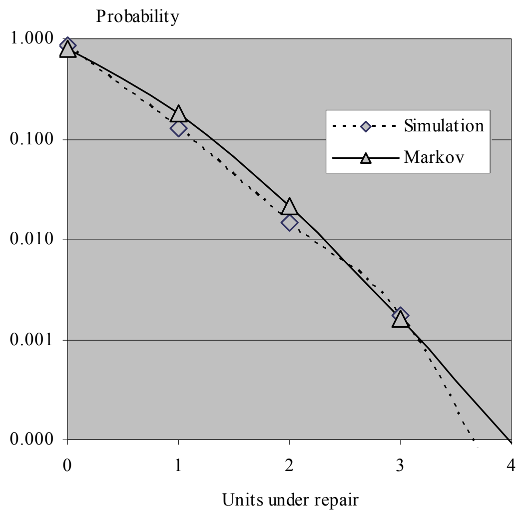

Figure 4: Comparison of Simulation and Calculation

The agreement is good, and proves to be so for other cases, so for the rest of this paper the steady state Markov calculation has been used for convenience. The small deviations are due largely to the restricted sample size used and serve to show that simulations need to include long trial periods in order to be accurate. As an aside, the Poisson approach also gives the same results, but is only applicable to the case of a constant repair and return time.

There are two common criteria used in then determining the number of spares:

1. Choose a number which gives (say) 95% or 99% confidence that no site will run out of spares

2. Choose a number which ensures that the system meets an availability target – usually 99.999%

To determine availability requires the additional step of determining the outage associated with failures in different states. Providing that one has a spare available or can perform a protection switch, the outage is small or zero, as the switch will occur in <10 seconds: once one has run out of spares or the possibility of switching, the outage lasts for the period that the system occupies those states.

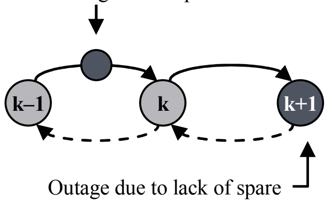

Figure 5: When outages occur

Thus outage is obtained by summing the outages due to replacements, plus the probability of being in states where replacement is not possible. Note that the outage due to replacement may not be fixed. For example one might have a spare unit at each site plus one or two centralized units. In this case the outage due to the first failure might be just a few hours while the outage when units are already on repair would be longer, as units must first be transported to site.

It is important to realize that without some form of protection switching good availability is problematic. Consider a unit with 5,000 FIT and a replacement time of just 2 hours – both rather good values – which gives an outage of 2 x 5,000 x 2E-9 = 2E-5 per unit. Since a point-to-point link will include two such units, the availability would be 1 – 2 x 2E-5 = 0.99994, which is somewhat short of the usual 99.999% objective. Except for sites which are permanently staffed repair times are generally longer.

Also worth consideration is protection and the system topology, which are illustrated in the following two examples. Firstly, an SLTE protected by an N:1 switch.

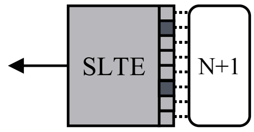

Figure 6: N:1 protection

Even when there are no spares available a single failure doesn't produce an outage, but a second one creates an outage which lasts until a unit is returned. Contrast this with ring protection.

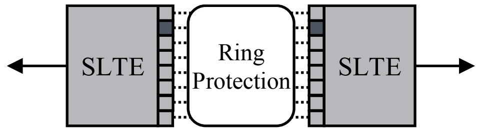

Figure 7: Ring-protected SLTEs

Here the second failure produces an outage only if it affects the same wavelength. In an N wavelength system there are N pairs of failures which create an outage out of N(N+1)/2 possible pairings. In most cases there will be no outage from the second failure and the outage can be of short duration, as the problem can be fixed by swapping wavelength units.

As before, the overall network needs to be considered, with combinations of failures categorized according to whether they create an outage or not. This can become quite a complex combinatorial problem and it may be worth using some form of approximation.

6. EXAMPLES AND STRATEGIES

It is interesting to consider a practical case and contrast different approaches and in this section a simple network of four stations each with two 8 wavelength SLTEs is considered and for simplicity only the wavelength units are considered.
Consider first a 95% UCL FIT value of 10,000 and a 60 day repair and return time.

To achieve 99% confidence of having spares in a given station requires 2 spares per station. To achieve 99.999% availability requires 3 and can only be achieved with protection switching. (Remember that with protection there is no outage until after one runs out of spares.)

These calculations, however, were performed with 95% UCL figures which probably overestimate the true failure rate significantly and it is instructive to repeat them using 5,000 FIT. This suggests 1 or 2 spares depending on the objective chosen.

An even better solution might be to share the spares using a central pool. This approach suggests that only 3 or 4 spares are needed. In this case the calculations need to include the transport times, so the table of cumulative probabilities provided below is not strictly applicable. However, it shows the order of the numbers and for transport periods of a few days the number of required spares is the same.

Failures	Cumulative Probability of failures		
	16 units	16 units	64 units
	10,000 FIT	5,000 FIT	5,000 FIT
0	79.422%	89.119%	63.078%
1	97.720%	**99.385%**	92.144%
2	**99.828%**	**99.977%**	98.841%
3	**99.990%**	99.999%	**99.870%**
4	100.000%	100.000%	**99.988%**
5	100.000%	100.000%	99.999%

Table 2

Sharing spares between sites certainly appears attractive, but it is important to recognize some practical issues. The spares may be owned by the individual landing parties, making sharing problematic, although one can envisage commercial solutions in this case. More difficult is the case where customs clearance takes a a long time; in this case certain stations may not be able to benefit from shared spares.

An alternative way to reduce the spares requirement is to demand a faster return time. Achieving really short times normally means replacing, rather than returning units and the effect is simply to shift the spares holding to the supplier, who will probably suggest that this will add to the cost. Apart from the cost issue system owners will inevitably question whether they can afford the risk that the spares will not be available when they are needed; mathematics does not provide an answer to this question, although commercial experience may. This solution also does not solve the issue of slow customs clearance unless the supplier holds the replacement units in the country concerned.

The examples show some possible ways to reduce spares requirements, but a pessimist would surely suggest that these are all achieved on the basis of ignoring the potential for things to be worse than expected. What would be the effect, for example, of discovering that due to a bad batch of components the failure rate is significantly higher than the expected value? At first sight this might seem to be an argument for purchasing more spares but doing so would mean buying more spares from a batch which is of poor reliability. It also involves a significant additional expense to cover what should be a relatively rare occurrence – and one where the probability is difficult to assess.

Using availability or confidence of spares being available as objectives, as described above, is clearly pragmatic, but maybe an economic view should be considered? Is there a way to balance the risk of outages against the cost of additional spares? As a first approach one might evaluate the cost/benefit balance by calculating the expected cost of payments anticipated under service level agreements (SLAs). Typically such SLAs pay nothing until a specific availability target is not met, but then pay a penalty which may depend on the degree of outage and the number of channels affected. Calculating the expected payment can be done with minor modification of the mathematics described previously, with the details being adjusted to fit the way that the SLA penalties are calculated. For obvious reasons it would not be appropriate to discuss the details of specific SLA agreements here, so no examples are given. This simple cost/benefit analysis is somewhat simplistic; it is clearly important to consider whether the expected outage levels would be acceptable to end-users and to consider whether such a policy sits well with the company implementing it.

7. SUMMARY / CONCLUSIONS

This paper has discussed some of the simplified mathematics used in reliability analysis. Given the difficulty of assessing failure rates with precision, the use of some simple approximations seems to be justified and simple calculations are shown to be in reasonable agreement with simulations. Considering the overall network and any protection scheme is very important, as protection generally reduces the need for spares.

Given the cost issue of purchasing a large number of spares, the paper suggests:

1. that calculations should be done with calculated failure rates rather than with 95% UCL values
2. that (where possible) spares should be shared between sites
3. service agreements which offer faster return times should be considered
4. the objectives used in the calculations should be re-considered and based more on economics

8. REFERENCES

[1] "Parameters and calculation methodologies for reliability and availability of fibre optic systems" ITU-T Recommendation G.911, Para 6.2.2

9. ACKNOWLEDGEMENTS

The author would like to thank various colleagues, in particular Igor Czajkowski, Andrew Robinson and Herve Fevrier for help at various times and would also like to acknowledge the value of discussions with Jacques Gros of AT&T.

