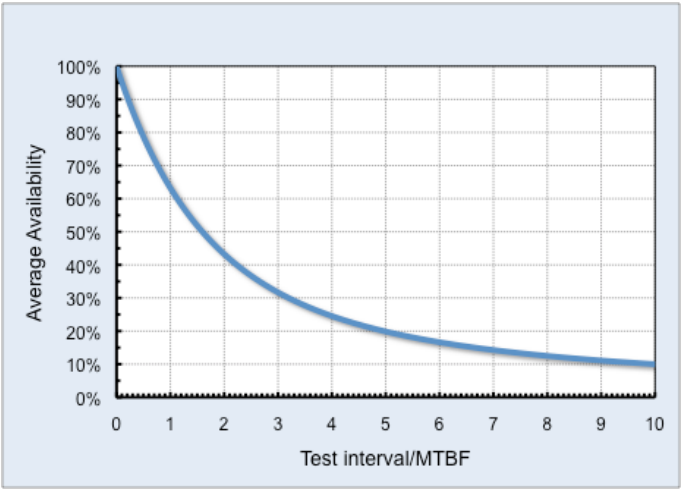
5 The basis of decision-making

5.1 Introduction

In earlier chapters we saw that a protective device's availability is not fixed, but it depends on its reliability and how frequently it is tested. For an idealised single device that is expected to fail at random, its average availability falls continuously from 100% as the testing interval is increased.



This relationship enables us to choose a task interval that delivers the minimum average availability that is needed. This chapter focuses on the question

How do we decide what protective device availability is needed?

The following sections show that there are three different approaches:

- Use an availability figure determined by detailed quantitative modelling
- Specify the minimum allowed mean time between multiple failures
- Choose the availability that delivers the lowest cost to the organisation

5.2 Availability

If it is possible to set a target availability for the protective device, then it can be used directly to calculate the required failure-finding interval. The formula used to work out the device's availability must take into account its configuration and technical characteristics, but in principle it is easy to define the right testing interval.

First we are going to look ahead to a later chapter where we find that the average availability of a simple, single protective device that fails at random is given by the following formula.

$$A = \frac{M_{dev}}{T}\left(1 - \exp - \frac{T}{M_{dev}}\right)$$

The terms in this equation are

A	The device's average availability over time
M	The device's mean time between failures
T	The failure-finding interval

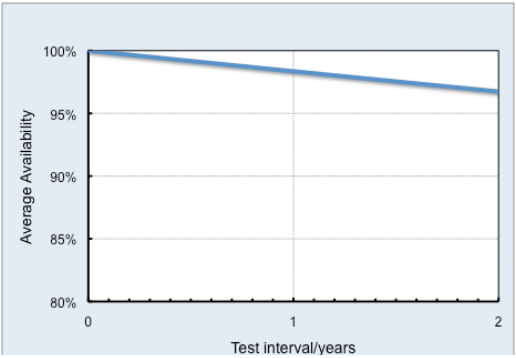
If we know the required average availability and the device's mean time between failures, the required test interval is easily calculated, or it can be found by drawing a graph of availability against test interval and finding the interval that gives the required availability.

Example

A single level switch is used to sound an alarm if the liquid level in a storage tank rises above the permitted high level. The switch is thought to fail randomly and its mean time between failures is at least 30 years. The average required availability is 99%.

Based on these figures, we need to find the failure-finding interval T for which

$0.99 = (30/T)\,(1–\exp(–T/30))$



The required interval is 0.6 years, or about 7 months. In practice the task would probably be carried out every six months to simplify maintenance scheduling.

The calculation is simple and it only requires two items of information: the mean time between failures of the protective system and the required availability.

Although the device's mean time between failures may not be known with absolute certainty, it is usually possible to find a worst case, lower bound by using maintenance records, manufacturers' data, or information that is available from generic industry databases. But where does the required availability come from?

Sometimes the required availability can be found in equipment or system documentation, particularly if the analysis involves an asset that has been subject to a rigorous, quantitative risk analysis using techniques such as fault tree analysis (FTA). Sadly, no easily accessible availability target exists for most industrial equipment, and we have to do a little more work before we can calculate the failure-finding interval.

Availability is the simplest criterion that can be used to derive a failure-finding interval, but it should not be used unless a robust quantitative model is available which justifies the chosen value.

5.3 Tolerable Risk

In the previous section we saw how easy it can be to calculate a failure-finding interval using just two pieces of data:

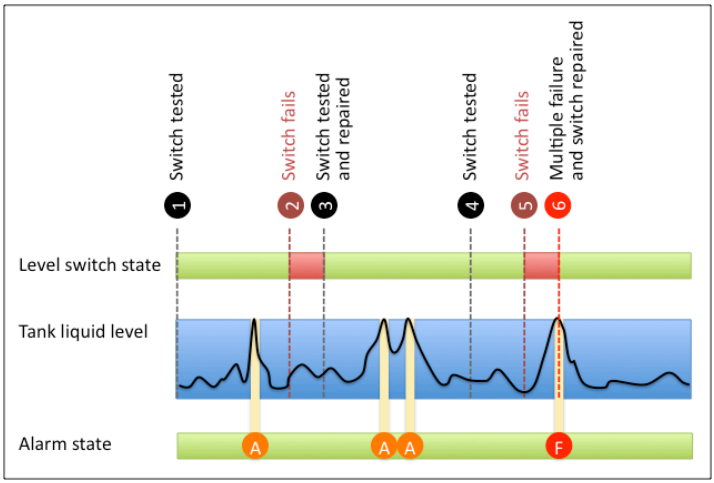
- The device's mean time between failures
- The required average device availability

Unfortunately we also saw that the required availability is not usually known with any certainty. How can we calculate it?

Chapter 4 introduced the concept of *multiple failures*. A multiple failure occurs if a demand is made on the protective system while it is in a failed state; in other words, while it is unavailable.

In the example above, a demand on the protective device occurs if the liquid level in the tank rises above the alarm level. An alarm sounds if the level switch is working; if it has failed, then we have a multiple failure that might lead to a process trip or liquid escaping from the tank.

The diagram below shows a possible history of the tank alarm system and tank liquid level.



The switch is tested regularly (at points 1, 3 and 4 in the diagram). It is working at the start of the time line, but at point 2 it fails. When it is tested at point 3 it is repaired; during the interval between points 2 and 3 it is in a failed state and could not sound an alarm if the liquid level were to rise, but there is no abnormally high level and so there is no multiple failure.

The tank level rises above the alarm level at four different times. The first three times, the switch is working (available) and an alarm is sounded each time.

The switch is tested at point 4, but it fails at point 5. From this time onward it is unavailable, and before it is tested again, the liquid level rises but no alarm is sounded. The multiple failure occurs at point 6.

We know that the availability of the level switch can be increased by testing it more often, but it is impossible to make it function continuously unless we also check it continuously. One way to decide what level of availability is needed is to set a maximum tolerable rate of multiple failures.

Example

Suppose that the tank is overfilled on average three times a year. These occasions occur at random, so it is not possible to know when the alarm system might be needed. If the availability of the level switch is 90%, then the probability that the alarm sounds each time is 90%, and the average number of alarms per year is

90% x 3 = 2.7 per year

Conversely, the average number of multiple failures—tank overfills that do not result in an alarm—is

10% x 3 = 0.3 per year

Of course this is just the average number of multiple failures per year. In reality there may be zero, one, two, three or even more per year, but the average over a long period of time should be 0.3 per year.

We can use this multiple failure rate (0.3 per year in the example above) to set the required availability. If we increase the availability of the protective device, then the number of multiple failures per year is decreased. So instead of choosing the availability of the protective device directly, we ask

How often, on average, are we willing to tolerate a multiple failure?

If the mean time between demands (the average time between tank high levels in this example) is M_{dem}, and the average protective device availability is A, then the average number of multiple failures (undetected tank high levels) per unit time is

$$\frac{1}{M_{dem}}(1-A)$$

If we decide that the minimum mean time between multiple failures that we are willing to tolerate is M_{mf}, then by rearranging the equation above, the required availability is

$$A = 1 - \frac{M_{dem}}{M_{mf}}$$

Now it is easy to see why availability is a poor criterion for setting failure-finding intervals unless it is based on a robust quantitative model. If we pluck a required availability level out of the air, then we make the assumption that the resulting multiple failure rate is tolerable. But the mean time between multiple failures achieved depends on both the protective device availability and on the demand rate; because the multiple failure rate is what is ultimately important to us, we always need to take into account the demand rate.

For multiple failures that have safety or environmental consequences, use the required mean time between multiple failures to determine the device availability that is necessary

Example

A crane's overhoist protection switch is designed to stop upward movement of the load if it goes beyond a set position. If this switch were to fail when required, the crane would be damaged and its load could drop 20m to the floor below, possibly injuring or killing several workers.

The manufacturer's data suggest that the minimum mean time between failures of the switch (failure to operate when required) is 150 years. An overhoist condition that activates the switch occurs about once every five years.

After discussion, the analysis group agrees that the multiple failure should occur no more often than once every million years.

The required average availability is therefore

$$A = 1 - \frac{M_{dem}}{M_{mf}} = 1 - \frac{5}{1000000} = 99.9995\%$$

Using the formula discussed later in this book, the group concludes that the switch should be tested twice per day.

…and the data?

Now we have a way to calculate the required device availability, but at the cost of needing two numbers rather than one:

- The rate of demands on the protective device (or the mean time between demands)
- The minimum tolerated mean time between multiple failures

Demand rates can vary over a huge range. Some protective systems are activated several times per day, while others may never be used over a period of decades. It is obviously relatively easy to find the demand rate for systems that are activated frequently, but data for rare demands may need research or may have to be estimated.

Specifying the shortest tolerated mean time between multiple failures can be far more challenging. How do we decide whether we should tolerate one failure per year, per century, per millennium, or per million years?

The tolerated multiple failure rate depends on a number of factors, including:

- The effects of the multiple failure
- The number of possible serious failures for which the organisation is responsible
- Who would be exposed to the multiple failure
- Constraints imposed by law and statutory bodies

The issue of tolerable risk is the subject of the next chapter.

5.4 Economic Basis

The previous section considered how the required protective device availability can be calculated if we know the demand rate on the system and if we specify a minimum tolerated mean time between multiple failures. The issue of how to determine the level of risk that can be tolerated was left for a later chapter.

The concept of "tolerated risk" can be applied to a range of failures that have safety effects.

Boiler and relief valves

Boiler pressure is limited by two relief valves. The required relief valve availability is determined by how often the boiler pressure exceeds a safe level and the tolerated mean time between unrelieved pressure excursions that could result in a boiler explosion.

Turbine overspeed system

A turbine overspeed system should shut down the turbine if its speed exceeds a safe level. The overspeed system availability is determined by how often overspeed events occur and the tolerated mean time between undetected overspeed events which may lead to serious damage and possible injury.

It can also be applied to failures that have environmental consequences.

Tank ultimate level switch

A tank ultimate level switch should shut down the supply pump and the upstream process if the tank level exceeds 30cm below the overflow. Overflowing effluent from the tank could lead to a reportable environmental incident. The level switch availability is determined by the rate of demands on the switch and the minimum tolerated mean time between environmental incidents.

Now consider applying the same technique to this example.

Pump low supply pressure switch

A low pressure switch is intended to shut down a pump if the suction line pressure drops below a set level. If it fails to trip when it is required, the pump could be damaged with a potential cost of about $1500, and about two hours' production would be lost, with a value of about $3500.

> The low pressure switch availability is determined by the rate of demands (how often the suction line pressure is low) and the minimum tolerated mean time between undetected low pressure events.

Although it is difficult to answer the question,

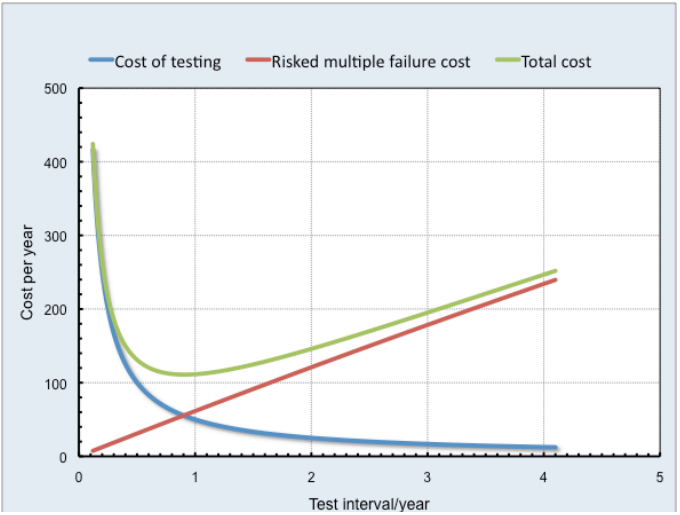
"How often are we willing to allow boiler explosions to occur?"

it is at least possible, perhaps after some discussion, to define an intolerable range of risk. However, if we now ask the question,

"How often are we willing to experience undetected pump low pressure events?"

it is not at all clear where the range of "tolerable risk" lies. On one hand, it is obvious that we should ensure that there is some level of protection against these events, because the potential economic costs are not insignificant. On the other hand, if the risk of a multiple failure were reduced to a very low level, the organisation would spend far too much on frequent testing of the pressure switch. Although we know that the extreme limits (high availability with too much testing or low availability with too little testing) are both undesirable, it is not possible to be sure where the right availability level is to be found.

This example suggests a different way to deal with multiple failures that have only economic (monetary) consequences. If testing infrequently results in unacceptable damage and downtime penalties, but the cost of very frequent tests is too high, then presumably there is a testing interval that results in the a lower expenditure than the two extremes. This is a balance between testing costs (which increase directly with testing frequency) and the risked costs of the multiple failure, which increase with lower protective device availability.



The graph above shows the relationship between testing costs and risked downtime costs as the failure-finding interval is increased. The total cost—the cost of testing plus the risked cost of downtime—has a broad minimum, in this case at a test interval of just under one year. This represents the lowest cost to the organisation, and is therefore the optimum testing interval.

To calculate the optimum failure-finding interval of a simple protective system we need four pieces of information.

- The mean time between failures of the protective device
- The mean time between demands on the protective device
- The cost of a single failure-finding task
- The cost of a single multiple failure event

Details of the calculation are given in a later chapter.

5.5 Key Points and Review

Three criteria can be used to set failure-finding intervals for protective devices.

Availability	The average availability required from the protective device
Mean time between multiple failures	The minimum tolerable mean time between multiple failures. This is usually applied to failures that have safety or environmental consequences
Lowest overall cost	The interval selected minimises the overall cost by balancing the cost of testing the protective device against the risked cost of damage and downtime if the multiple failure occurs

Availability is generally a poor criterion for setting failure-finding intervals unless there is a pre-existing detailed design or some other robust justification for selecting a specific minimum level of availability.