TPK4120 - Lecture summary

Jørn Vatn
eMail: jorn.vatn@ntnu.no

Updated 2022-10-26

Chapter 13 - Reliability of Safety Systems

The international standard Functional safety of electrical/electronic/programmable electronic safety-related systems (IEC 61508, 2010) is a generic, performance-based standard for safety-related systems. In this chapter we present some fundamental elements addressed in IEC 61508.

Safety Instrumented Systems and Safety Instrumented Functions

A safety instrumented system (SIS) is an independent protection layer that is installed to mitigate the risk associated with the operation of a specified hazardous system, which is referred to as the equipment under control (EUC). An example of an EUC is a process vessel.

A SIS is composed of *sensors* often referred to as *input elements*, *logic solvers* and *actuating items* often referred to as *final elements*.

A safety instrumented function (SIF) is a function that is implemented by a SIS and that is intended to achieve or maintain a safe state for the EUC with respect to a specific process demand such as high pressure in the vessel.

A SIS has two main system functions:

1. When a predefined process demand occurs in the EUC; the deviation shall be detected by the SIS sensors, and the required actuating items shall be activated and fulfil their intended functions.

2. The SIS shall not be activated spuriously, that is, without the presence of a predefined process demand in the EUC.

A *demand* is defined as: An event or a condition that requires a SIF to be activated (i) to prevent an undesired event from occurring or (ii) to mitigate the consequences of an undesired event. In the process industry, a demand is also called a *process upset* or a *process deviation*.

Testing of Safety Instrumented Functions

A SIS is often a passive system that is activated only when a demand occurs. Failures may therefore occur and remain hidden until the system is demanded or tested. We often referred to two main categories of testing.

- *Proof Testing*. To verify that a SIS is able to perform its SIFs, the system is usually proof tested at regular intervals of length τ. The time interval between two consecutive proof tests is often called the proof test interval. Proof testing is also called functional testing. Dangerous failures detected by proof testing are called dangerous undetected (DU) failures.

- *Diagnostic testing*. A diagnostic test is an automatic partial test that uses built-in self-test features to detect failures. Dangerous failures detected by a diagnostic test are called dangerous detected (DD) failures. The identified faults are announced as alarms, locally at the equipment and in the control room.

Safety Integrity Levels (SILs)

IEC 61508 uses safety integrity as a performance measure for a SIF. Safety integrity is the probability of a SIS satisfactorily performing the specified SIFs under all the stated conditions within a stated period of time. IEC 61508 does not specify detailed probability values, but divides the requirements into four safety integrity levels, SIL 1, SIL 2, SIL 3, and SIL 4, with SIL 4 being the most reliable and SIL 1 being the least reliable.

Reliability Metrics

Probability of Failure on Demand

The probability of (dangerous) failure on demand, PFD(t) is the probability that the SIS has a dangerous fault and that it is not able to perform its SIFs at time t. The notion probability of failure on demand may indicate that we are dealing with a conditional probability, given that a demand has occurred. This is not correct and PFD(t) may be expressed as

$$\Pr(\text{The SIS is not able to perform its SIF at time } t)$$

irrespective of whether a demand occurs or not. If a demand should occur at time t, PFD(t) is the probability that the SIS fails to perform its SIF. In many cases, it is not necessary to determine the PFD as a function of time and we can suffice with an average value. If the SIF is proof tested after regular intervals of length τ and the system is considered to be as-good-as-new after each proof test, the long-term average probability of failure on demand can be expressed as

$$\text{PFD} = \frac{1}{\tau}\int_0^\tau \text{PFD}(t)dt$$

Average Frequency of Dangerous Failures per Hour

For SIFs that are operated in high-demand or continuous mode, IEC 61508 requires that the reliability is specified by the average frequency of dangerous failures (PFH) where the frequency is given as number of dangerous failures per hour. The abbreviation PFH is retained from the previous version of IEC 61508 where the metric was called “average probability of (dangerous) failure per hour.” The idea behind using the PFH as a reliability metric is that demands will occur so often that when a dangerous failure of the SIF occurs, it is most likely that a demand will occur and a hazardous event will be manifested before we can bring the EUC to a safe state.

Spurious Trip Rate and related concepts

The spurious trip rate, (STR) is the rate of spurious trips of a specified SIF per hour.

There are three main types of spurious activation: (i) spurious activation of individual components, (ii) spurious activation of a SIF, and (iii) spurious shutdown of the process. To use the same concept to describe all the three types may lead to misunderstanding and confusion. To distinguish the different types of spurious activation, the following terms and definitions are often used (deviates from the presentation in the textbook):

- *Spurious operation.* A spurious operation (SO) is an activation of the safety function of a channel without the presence of a specified process demand. A spurious operation of a channel is said to be an SO-failure and the SO-failure rate is denoted λ_{SO}.
- *Spurious trip.* A spurious trip (ST) is an activation of a SIF without the presence of a specified process demand.
- *Spurious shutdown.* A spurious shutdown is a partial or full process shutdown without the presence of a specified process demand.

Reliability Metrics and SIL

To fulfil the requirements of a safety integrity level, a SIF in low-demand mode must have a PFD in the corresponding interval specified in Table 1. Similarly, a SIF in high-demand or continuous mode must have a PFH in the corresponding interval specified in Table 1.

Table 1: SIL requirements vs PFD/PFH

SIL	**Low demand mode of operation** (Average probability of failure to perform its design function on demand)	**High demand mode of operation** (Average probability of failure per hour to perform its design function)
4	$10^{-5} \leqslant$ PFD $<10^{-4}$	$10^{-9} \leqslant$ PFH $<10^{-8}$
3	$10^{-4} \leqslant$ PFD $<10^{-3}$	$10^{-8} \leqslant$ PFH $<10^{-7}$
2	$10^{-3} \leqslant$ PFD $<10^{-2}$	$10^{-7} \leqslant$ PFH $<10^{-6}$
1	$10^{-2} \leqslant$ PFD $<10^{-1}$	$10^{-6} \leqslant$ PFH $<10^{-5}$

Classification of Failures Based on Consequence and Detectability

Hardware failures can be classified as:

- *Dangerous* (D) failure. A dangerous failure is a failure that brings the item into a state where it is not able to perform its safety function(s). When the item is in such a state, it is said to have a dangerous (D) fault.
- *Safe* (S) failure. A safe failure is a failure that does not leave the item in a state where it is not able to perform its safety function(s). When the item is in such a state, it is said to have a safe (S) fault.

Dangerous and safe hardware failures/faults may also be categorized as detected or undetected.

- *Detected*. A fault that is detected by automatic diagnostic testing, internal in the item or connected to a logic solver.
- *Undetected*. A fault that is not detected (not diagnosed) by automatic diagnostic testing, internal in the item or connected to a logic solver.

Combining the two principles of categorization yields:

- *Dangerous undetected* (DU) faults. DU-faults are preventing activation on demand and are revealed only by proof testing or when a demand occurs. DU-faults are sometimes called dormant or hidden faults. The DU-faults are of vital importance when calculating the SIF reliability as they are a main contributor to SIF unavailability.
- *Dangerous detected* (DD) faults. DD-faults are detected short time after they occur, by automatic diagnostic testing. The average period of unavailability due to a DD-failure is called the mean time to restoration (MTTR), the mean time elapsing from the DD-failure occurs until the function is restored.

- *Safe undetected* (SU) failures. Non-dangerous failures that are not detected by automatic self-testing.

- *Safe detected* (SD) failures. Non-dangerous failures that are detected by automatic self-testing. In some configurations, early detection of failures may prevent an actual spurious trip of the system.

PFD calculations for systems

To obtain PFD for a system we may follow the following procedure:

1. Find PFD for the system as a function of t in an interval, i.e., $0 \leqslant t \leqslant \tau$, and denote the result PFD(t)

2. To obtain PFD(t) we often utilize the system survivor function, say $R(t)$

3. Find the average PFD(t) by integration: $\text{PFD} = \frac{1}{\tau}\int_0^\tau \text{PFD}(t)dt = 1 - \frac{1}{\tau}\int_0^\tau R(t)dt$

The classical example is one component proof tested at point of times τ,2τ,3τ,..., and time to failure is exponentially distributed, i.e., $R(t) = e^{-\lambda t}$.

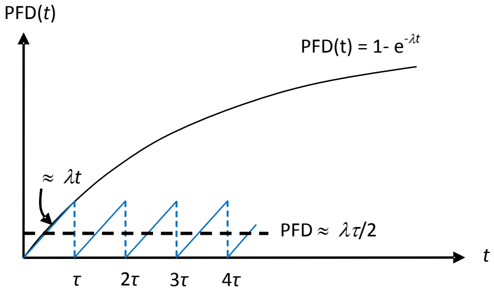
$$\text{PFD} = 1 - \frac{1}{\tau}\int_0^\tau R(t)dt = 1 - \frac{1}{\tau}\int_0^\tau e^{-\lambda t}dt = 1 + \frac{1}{\lambda\tau}\int_0^\tau -\lambda e^{-\lambda t}dt = 1 + \frac{1}{\lambda\tau}e^{-\lambda t}\Big|_0^\tau = 1 + \frac{1}{\lambda\tau}\left(e^{-\lambda\tau} - 1\right)$$

If $\lambda\tau$ is small, i.e., (<0.01) we utilize that $e^{-x} \approx 1 - x + x^2/2$, and inserting in the expression for PFD(t) yields:

$$\text{PFD} = 1 + \frac{1}{\lambda\tau}\left(e^{-\lambda\tau} - 1\right) \approx 1 + \frac{1}{\lambda\tau}\left(1 - \lambda\tau + (\lambda\tau)^2/2 - 1\right) = \lambda\tau/2$$

Note that λ is the rate of DU-failures. In some presentations the notation λ_{DU} is used, but for simplicity we only use λ. In general we allso need to add contribution of DD failures, but this is not further discussed in this presentation.

Another way to obtain the same result is to use that $e^{-x} \approx 1 - x$ for small x-values directly, hence we have that $\text{PFD}(t) = 1 - e^{-\lambda t} \approx \lambda t$ in an each proof test interval:



which yields PFD $\approx \lambda\tau/2$.

Such an argument we may also use for two identical components in parallel that are proof tested at the same time. The time dependent PFD of the two components is found by

$$\mathrm{PFD}(t) = \mathrm{PFD}_1(t) \cdot \mathrm{PFD}_2(t) \approx (\lambda t)^2$$

yielding:

$$\mathrm{PFD} = 1/\tau \int_0^\tau \mathrm{PFD}_1(t) \cdot \mathrm{PFD}_2(t) dt \approx 1/\tau \int_0^\tau (\lambda t)^2 dt = \frac{(\lambda\tau)^2}{3}$$

The PFDs of some $k$oo$n$ systems of identical and independent components with constant failure rate λ and test interval τ are found to be:

$k\backslash n$	1	2	3	4
1	$\frac{\lambda\tau}{2}$	$\frac{(\lambda\tau)^2}{3}$	$\frac{(\lambda\tau)^3}{4}$	$\frac{(\lambda\tau)^4}{5}$
2	–	$\lambda\tau$	$(\lambda\tau)^2$	$(\lambda\tau)^3$
3	–	–	$\frac{3\lambda\tau}{2}$	$2(\lambda\tau)^2$
4	–	–	–	$2\lambda\tau$

The general formula for PFD is

$$\mathrm{PFD} = \binom{n}{n-k+1} \frac{(\lambda\tau)^{n-k+1}}{n-k+2}$$

The argument for this formula is as follows. We have a $koon : G$ system corresponding to an $n - k + 1oon : F$ system. This means that if $n - k + 1$ components are in a fault state, the system will be in a fault state. The minimal cut sets will all contain $n - k + 1$ components, and there are $\binom{n}{n-k+1}$ such minimal cut sets. At time $t, 0 \leq t < \tau$ the probability that one such cut set is in a fault state is $\mathrm{PFD}_j(t) \approx (\lambda t)^{n-k+1}$ with an average $\mathrm{PFD}_j = \frac{(\lambda\tau)^{n-k+1}}{n-k+2}$. Multiplying with the number of minimal cut sets gives the above formula.

Common cause failures

The equation above assumes that components in a SIS fail independent of each other. In practice components may fail due to common causes. Common cause failure may be due to maintenance introduced failures, design failures, excessive stress etc. To model common cause failures the total failure rate of one component (i.e., rate of DU failures) is split into an independent part and a dependent part:

$$\lambda = \lambda_{(i)} + \lambda_{(c)} = (1-\beta)\lambda + \beta\lambda$$

where $\beta = \lambda_{(c)}/\lambda$ is the common cause factor. This (beta factor) model now yields:

- For the dependent part, use PFD = $\frac{\beta\lambda\tau}{2}$
- For the independent part, use the independent failure rate $(1-\beta)\lambda$ in the PFD formulas of the $k$oo$n$ system of identical and independent components
- Add the contributions:

$$\text{PFD} = \frac{\beta\lambda\tau}{2} + \binom{n}{n-k+1}\frac{\left[(1-\beta)\lambda\tau\right]^{n-k+1}}{n-k+2}$$

Staggered Testing

Now, consider the case with the two components in a 1oo2 voting having the same λ and τ, but where the testing is not carried out simultaneously. The situation is illustrated in Figure 1.

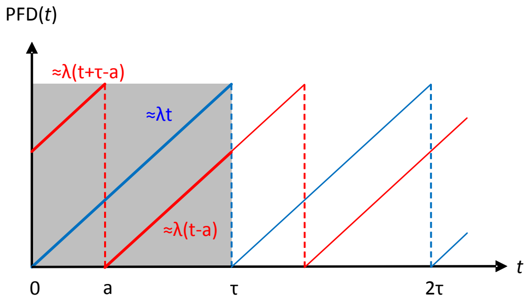


Figure 1: Staggered testing

Assuming that component 2 is tested at time (a) inside the test interval of component 1, it can be shown that:

$$\text{PFD}(a) \approx \frac{(\lambda\tau)^2}{3}\left(1 - \frac{3a}{2\tau} + \frac{3a^2}{2\tau^2}\right)$$

PFD(a) attains its maximum value

$$\text{PFD}_{max} \approx \frac{(\lambda\tau)^2}{3} = \frac{4}{3}\cdot\frac{\lambda\tau}{2}\cdot\frac{\lambda\tau}{2}$$

when $a = 0$ or $a = \tau$, i.e., when the components are tested simultaneously.

PFD(a) attains its minimum value when $a = \tau/2$, i.e., when component 2 is tested in the middle of the test interval of component 1:

$$\text{PFD}_{min} \approx \frac{5}{8}\frac{(\lambda\tau)^2}{3} = \frac{5}{6} \cdot \frac{\lambda\tau}{2} \cdot \frac{\lambda\tau}{2}$$

Note that this minimum PFD is actually smaller than the PFD obtained when simply multiplying the average PFD values of the individual components. Compared to the case of simultaneous testing, we obtain a PFD reduction of 38% in the case of "optimal" testing. Hence, there is a great potential for improvement in the total PFD if components are tested at different times. This is exploited in staggered testing. Also note that the minimum value is obtained when $a = \tau/2$ for a 1oo2 system, for general configuration it is more complicated to set up the optimal staggered testing regime.

More about the test regime

There are three different test regimes that are considered

- Simultaneous testing, i.e., $a = 0$ in Figure 1
- Optimal staggered testing, i.e., $a = \tau/2$ in Figure 1
- Independent testing

If the components are tested independently we can calculate PFD for each component by the formula $\text{PFD}_i = \lambda_i\tau_i/2$ and proceed with the structure function. Due to the independent test regime, it is reasonable to argue that the components are independent, and we proceed with the standard approach which here means to replace x_i in the structure function with $p_i = 1 - \text{PFD}_i$. If common cause failures are relevant, we may add an artificial block to represent the common cause "part" of the components.

It is important to understand the difference between independent *testing* and independent *components*. Independent *testing* means that if we know that one of the component is in a fault state this will have no information regarding if the other component is in a fault state. This is not true for e.g., simultaneous testing. For simultaneous testing we have: If one component is known to be in a fault state it is more likely that we are at the end of the test interval compared to if it was functioning. Hence, it is also more likely that the other component is in a fault state because the likelihood of being at the end of the test interval is higher. The components performance are dependent not because of any physical reasons, only due to the testing regime.

The PDS method

The PDS method is developed by SINTEF Safety. The method has two main features:

1. It proposes a more realistic way to model common cause failures (CCF). In the β-factor model a CFF will always cause all components to fail. This is often not realistic. The PDS method therefore proposes a correction factor to adjust the β-factor to account for the situation where not all components fail due to the CCF situation.

2. In IEC 61508 only random hardware failures are quantified. The PDS method also quantifies systematic failures by the so-called test independent failure term, p_{TIF}. Systematic failures are not treated in this presentation.

The idea behind adjusting the β-factor is that if we have a CFF causing two components to fail, it is not certain that the remaining components will fail. Some assumptions are then made regarding the probability that a third component fails given that two components have failed due to a CFF and so on. The correction factor is dependent on k and n, and is generally denoted C_{koon}, and is presented in Table 2 for some combinations:

Table 2: C_{koon} correction factors

$k\backslash n$	$n=2$	$n=3$	$n=4$	$n=5$	$n=6$
$k=1$	$C_{1oo2}=1.0$	$C_{1oo3}=0.5$	$C_{1oo4}=0.3$	$C_{1oo5}=0.20$	$C_{1oo6}=0.15$
$k=2$	-	$C_{2oo3}=2.0$	$C_{2oo4}=1.1$	$C_{2oo5}=0.8$	$C_{2oo6}=0.6$
$k=3$	-	-	$C_{3oo4}=2.8$	$C_{3oo5}=1.6$	$C_{3oo6}=1.2$
$k=4$	-	-	-	$C_{4oo5}=3.6$	$C_{4oo6}=1.9$
$k=5$	-	-	-	-	$C_{5oo6}=4.5$

A configuration specific β-factor is now calculated by multiplying the original β-factor with the correction factor C_{koon} found in Table 2. Note that the baseline β-factor is assumed to be specified for a 1oo2 system.

PFH calculations

In the following we present the simple approximation formula proposed in the PDS method for the probability of failure per hour, PFH:

$$\mathrm{PFH} = C_{\mathrm{koon}}\beta\lambda + \frac{n![\lambda(1-\beta)\tau]^{n-k+1}}{(n-k+1)!(k-1)!\tau} \tag{1}$$

Note that the correction factor C_{koon} only applies for the PDS method. To obtain eq. (1) we treat CCF failures and independent failures individually. For the CCF failures the results is rather obvious. For independent failures, let $p(t,k,n)$ be the probability that the first $n-k$ components are in a fault state assuming they are numbered $1,2,\ldots,n$. If the the first $n-k$ components are in a fault state and the remaining k components are functioning, then they are *critical*. A system failure will occur if one of remaining k components fails. We have that $p(t,k,n) \approx \left[\lambda(1-\beta)t\right]^{n-k}$. To find the contribution to the PFH for this situation we calculate the average of $p(t,k,n)$ in a proof test period, and multiply with $f_k = k\lambda(1-\beta)$. f_k is the total frequency of the event that one of the remaining k components fails. The above argument is valid when we consider the numbered $n-k$ components. There are $\binom{n}{n-k} = \frac{n!}{(n-k)!k!}$ ways to chose $n-k$ components, and adding up we obtain eq. (1). Note that the probability that the k remaining components being in a *functioning* state is considered to be close to one, so we do not take this into account.

STR calculations

In the following we present the simple approximation formula proposed in the PDS method for the spurious trip rate, STR:

$$\text{STR} = C_{\text{(n-k+1)oon}}\beta\lambda_{\text{SU}}$$

Note that the correction factor $C_{\text{(n-k+1)oon}}$ only applies for the PDS method. We have here explicitly indicated that the failure rate to go into the formula is the rate of SU failures. In some presentations λ_{SO} is used to reflect the rate of spurious operations on component level.

Markov approach

This presentations deviates from the presentation in the textbook. We have seen that the Markov equations may be written on matrix form:

$$\mathbf{P}(t) \cdot \mathbf{A} = \dot{\mathbf{P}}(t)$$

which may be approximated by:

$$\dot{\mathbf{P}}(t) = \frac{\mathbf{P}(t+\Delta t) - \mathbf{P}(t)}{\Delta t} = \mathbf{P}(t) \cdot \mathbf{A}$$

yielding

$$\mathbf{P}(t+\Delta t) = \mathbf{P}(t)[\mathbf{A}\Delta t + \mathbf{I}]$$

where $\mathbf{I}$ is the identity matrix. This equation may now be used iteratively with a sufficient small time interval Δt and starting point $\mathbf{P}(0)$ to find the time dependent solution. Only simple matrix multiplication is required for this approach.

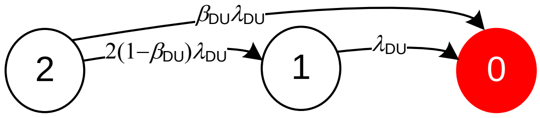
PFD

Assume that we know the state vector $\mathbf{P}(0)$ just after a proof test, and that we have established a Markov transition model for the SIS with respect to a given SIF. Then it is straight forward to find $\mathbf{P}(t)$ within a proof test interval by the approach presented above. Typically the probability of being in a state where all components are functioning (state r) is assumed to be one, and probabilities for the other states are equal to zero. Let F be the set of failed states with respect to the actual safety function of the SIS. We then have:

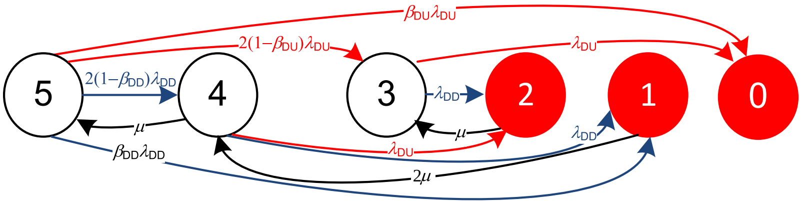
$$\text{PFD} = \frac{1}{\tau}\int_0^\tau \text{PFD}(t)dt = \frac{1}{\tau}\int_0^\tau \sum_{i\in F} P_i(t)dt$$

The integral is replaced by a sum in the numerical calculations since we are already solving the time dependent solution iteratively by time steps Δt.

The following figure shows the Markov diagram for a 1oo2 system considering DU-failures only:



Note that whereas the closed form formulas for PFD presented earlier only takes DU failures into account. With the Markov approach, DD failures may also be included. The Markov diagram for the 1oo2 systems now reads:



where the following system states are defined:

5: Both components OK

4: One OK, one DD-failure

3: One OK, one DU-failure

2: One DU-failure and one DD-failure

1: Two DD-failures

0: Two DU-failures

Red arrows are DU failures, blue arrows are DD failures, and black arrows are repairs.

PFH

The procedure is now similar to the approach for PFD, but we are seeking a rate, i.e., the rate of transition from a functioning state to a fault state for the SIS safety function is found by averaging:

$$\text{PFH} = \frac{1}{\tau}\int_0^{\tau} \text{PFH}(t)dt = \frac{1}{\tau}\int_0^{\tau} \sum_{i\notin F}\sum_{j\in F} a_{ij}P_i(t)dt$$

where a_{ij} is the transition rate from state i to state j measured in expected number of transitions *per hour*. Note that we can interchange the integration and summation operators, i.e., we may first calculate the average state probabilities, then calculate the appropriate transition rates.

STR

The procedure for the spurious trip rate is now similar to the approach for PFH, but we need to consider the spurious trip system failure mode. Therefore, we typically need to draw a new Markov diagram. Let F be the set of system failure states representing a spurious trip state. STR is found by averaging:

$$\text{STR} = \frac{1}{\tau}\int_0^{\tau} \text{STR}(t)dt = \frac{1}{\tau}\int_0^{\tau} \sum_{i\notin F}\sum_{j\in F} a_{ij}P_i(t)dt$$