Proceedings of the CAS–CERN Accelerator School: Beam Injection, Extraction and Transfer, Erice, Italy, 10-19 March 2017, edited by B. Holzer, CERN Yellow Reports: School Proceedings, Vol. 5/2018, CERN-2018-008-SP (CERN, Geneva, 2018)

Machine Protection and Activation

A. Nordt, R. Andersson, E. Bargalló, S. Kövecses
European Spallation Source ERIC, Lund, Sweden

Abstract

Machine protection can be described as a strategy that modern particle accelerator driven research facilities utilize to achieve high and optimized operational availability. In this context it is important to define what the terms 'operational availability' and 'machine protection strategy' intend to address. Research facilities have become more complex and also more expensive during the past decades, often aiming at unprecedented beam energies or beam power levels. Availability in the context of particle accelerator driven research facilities is often referring to 'availability of beam' that can be used for scientific applications performed by 'users'. This paper provides a general overview of how machine protection can be utilized to increase beam availability of accelerator driven research facilities.

Keywords

CERN report; machine protection; activation; risk management, availability.

1 Machine protection: a term, a system, or a strategy?

In this paper, the authors focus on providing a general overview of what machine protection (MP) is, how to apply machine protection methods and strategies, what criteria to address, and what guidelines potentially to follow when being confronted with the task of implementing or optimizing machine protection at a (complex) accelerator driven research facility. Thus the reader should not expect a detailed description of the mechanisms that can lead to damage, details on the beam physics aspects related to beam-induced damage at high beam energy or high beam power machines, or details on how to mitigate against certain hazards. There are excellent publications on these topics such as from the Joint International Accelerator School on Beam Losses and Accelerator Protection, that took place in 2014, see Ref. [1] and much more, see e.g., Ref. [2].

Instead the reader should be aware that he/she will be guided through the following main topics in a rather strategically inspired way.

- What does high availability of a research facility mean?
- How can we describe availability?
- Why could high availability become more and more important nowadays?
- What could machine protection strategy mean in the context of operational availability?
- How could Machine Protection be implemented?

1.1 What could possibly go wrong?

This section intends to give a brief overview of different types of damage events that have occurred in the past at accelerator driven research facilities. The accidents described are either beam induced or non-beam induced.

https://doi.org/10.23730/CYRSP-2018-005.533

In 2004, during extraction tests with 450 GeV/c protons in the SPS[1] extraction line at CERN, Switzerland, a beam with an energy of 2 MJ was deflected with grazing incidence into a vacuum chamber. This happened after the failure of a septum magnet. The vacuum chamber was cut along a length of 25 cm. A magnet further downstream was damaged due to beam losses arising from this incident. Condensed drops of steel were visible on the opposite side of the vacuum chamber and the vacuum chamber and quadrupole magnet had to be replaced, see Ref. [2].

Another accident happened at Tevatron, Fermilab, USA. A roman pot was moved into the beam and the particle showers generated by the Roman pot's movement quenched superconducting magnets located further downstream. What happened was that the beam moved by 0.005 mm/turn, and eventually touched a collimator jaw surface after about 300 turns. The entire beam was then lost, mostly on the collimator, that was damaged, see Ref. [2].

At SNS[2], Oakridge, USA, superconducting radio-frequency (RF) cavities are used to accelerate the beam. This equipment is operated at very high voltage levels and can be very sensitive to even very small beam losses. Beam losses then can lead to a surface quality degradation of the cavity. This means that operation at the same voltage level is not possible anymore and the probability for arcing is increased. At SNS, errant beam losses led to degradation of a superconducting cavity. Beam current monitors (BCMs) measured beam losses of a few microseconds. After such errant beams, sometimes the cavity gradient needed to be lowered. Conditioning after warm up helped in most cases, but in one case a complete cryo-module had to be exchanged because the degradation was too severe. The energy of beam losses in this case is only about 10−100 J and rather difficult to detect. The damage mechanisms are not yet fully understood; it is assumed that some of the beam hitting the cavity desorbs gas or particulates, creating an environment for arcing, see Ref. [2] and Ref. [3].

One more example from SNS is not related to beam-induced damage. In 2014, the water cooling of the absorber of the medium energy beam transfer (MEBT) chopper failed and water entered the MEBT, causing several weeks of downtime. Luckily, this happened during a maintenance period during which all the vacuum gate valves were closed. The water from the damaged MEBT chopper absorber was 'stopped' in the drift tube linac[3] (DTL) section. The absorber was not interlocked and also not checked periodically until this accident happened.

At LHC[4], CERN, an accident happened during test runs without a beam in 2008. A magnet interconnection was malfunctioning and a resulting electrical arc caused the release of a huge helium pressure wave (superconductivity of the helium was lost). Around 600 m of the ring's equipment were damaged and 2 km of the vacuum chamber had to be cleaned. A total of 53 magnets had to be repaired, see Ref. [2] and references therein.

1.2 What if things have gone wrong?

The initiating events and also the consequences of accidents like the ones mentioned in section 1.1 are manifold.

The obviously not so good consequences of such accidents are, from an organizational viewpoint:

- Unplanned downtime/no time for science;
- Additional cost for repair or replacement of damaged equipment;

[1] Super Proton Synchrotron

[2] Spallation Neutron Source

[3] Linear Accelerator

[4] Large Hadron Collider

- Unnecessary beam losses leading to unnecessary activation of equipment;
- Unnecessary beam losses leading to unnecessary stress of equipment and potential early aging.

The good outcome of such accidents however should not be neglected. Certainly, accidents impact on the awareness of the importance of establishing a robust safety and protection culture within an organization. Such culture, if embedded in the daily life of the designers, scientists, and users of research facilities, forms a very robust medium for safe operation and helps drastically in decreasing the probability of accidents happening. When addressing protection aspects it is important to also think out of the box. One should not neglect the obvious and easy parts of a system's design. Often, engineers and designers tend to focus on the challenging aspects of a new system's design. This means that they prefer focusing on the challenges imposed by using a new technology or design approach, rather than also focussing on the 'normal' equipment that is part of the new system. It is however crucial to keep the global system functionality in mind and to put the system into the context of the whole machine and eventually its role in beam production, acceleration, etc. Many accidents in fact happen because of a non-optimal interplay of several systems and lack of awareness of the global picture.

Machine protection tries to cope with this complexity and uncertainty by providing rather flexible functionality that helps to prevent and mitigate the risks of damage events to an acceptable level.

1.3 What is machine protection?

The latest generation of accelerator driven research facilities are very complex and often a high investment must be made to build these machines. As their purpose is to serve external users, these accelerator driven facilities need to fulfil the demands of the users both in terms of machine performance and availability. Numerous operational risks have to be accounted for when operating such machines. Two of the most imminent risks are damage to equipment and the associated downtime of the facility. The increased beam energies and beam power levels lead to increased damage potential of the beams, both from the particle beam and through the increased stress on the equipment. In addition to the unwanted downtime related to damage, there is also a certain cost to consider.

It is therefore necessary for accelerator facilities to have a strategy and systems that address these risks appropriately. Such systems are generally referred to as machine protection systems (MPS), even though the 'S' has become more of a strategy, or system-of-systems, than a single system that operates independently of other systems.

Based on the complexity, cost, and uniqueness of many research facilities but also more and more the demands for machine and beam availability from users, together with the knowledge that 'things can go wrong', the following two goals can describe what machine protection is there for.

Goal 1: Machine protection shall, in this order, prevent and mitigate damage to the machine, be it beam induced or from other sources, in accordance with beam and facility related availability requirements.

Goal 2: Machine protection shall protect the machine from unnecessary beam-induced activation having a potential to cause long-term damage to the machine or increase maintenance times. This shall be addressed in accordance with beam and facility related availability requirements.

It should be noted that machine protection cannot be implemented by a single group of people or a single work package/project. A common effort across the whole organization is needed to ensure the right level of protection. Teamwork is vital for implementing machine protection. Awareness, openness, global thinking, as well as understanding of the impact and consequences of certain decisions on a global (machine wide) level, are highly important.

2 When is machine protection needed?

In order to know how and what to protect against what kind of failures it is vital to set up a failure catalogue of your machine. There are a lot of different risk and hazard analysis methods available, however, it is unlikely that a single one of these could cover all aspects needed for machine protection. These methods are in almost all cases targeted towards safety related analysis. In general it should be avoided to perform a 'silo-system' component failure analysis, but rather focus on performing a functional analysis. Functional analyses can become very complex in the case of an accelerator, where for example the function of accelerating the beam is performed by several rather complex systems. The degrees of freedom for such analyses can easily become very large and then the analysis method itself may become unpractical. The interplay between the different systems performing one function is to be analysed with care.

Since particle accelerators used for research are many times very complex and not necessarily built with only standard off-the-shelf equipment, detailed documentation of e.g., failure modes, may not be available, which makes a standard risk analysis difficult to execute. Most of the research facilities operate 'prototype' machines, using a large amount of state-of-the-art equipment, hence one cannot establish the equipment's failure modes with high confidence nor document them, simply because they are not known. By 'prototypes', we mean here that usually there is only one kind of a specific facility/machine; there are e.g., several colliders around the world (Tevatron, LHC, etc.), however they operate at very different beam energy and beam power levels and use different technologies, hence possible equipment failures can lead to different accidents depending on the machine.

In order to assess whether machine protection is necessary, a large set of criteria are considered. These are, among many other aspects, the damage potential of the beam, the expected beam loss levels at different locations along the accelerator, the delicacy of the equipment (susceptibility for damage, location, environmental aspects, availability of spares, etc.), the beam injection and extraction mechanisms, beam stop capabilities, and requirements on beam and system availability. On top of this there is a need to prepare for unexpected events as modern accelerators take unbeaten paths. In this section we will discuss the criteria to be considered when deciding whether machine protection is needed or not.

2.1.1 Damage potential of the beam

It is important to understand how the beam can damage the different materials of the facility's equipment. To understand heating, melting, and damage mechanisms, one has to know about the particle type, its momentum, the stored energy in the beam, beam power, beam size, beam power/energy density, time structure of the beam, cooling conditions, and more. In order to estimate the order of magnitude of possible damage, the following rule of thumb can be used. One MJ can heat and melt ~ 1.5 kg of copper. One MJ corresponds to the energy stored in ~0.25 kg of Trinitrotoluene (TNT). And one MW during one second corresponds to one MJ, see Ref. [2]. Facilities such as the European Spallation Source, Lund, Sweden (currently under construction) intend to operate at 5 MW average beam power and 125 MW peak power per pulse.

2.1.2 Beam losses

Different types of beam losses lead to different aspects of required protection.

Continuous beam losses are inherent during the operation of accelerators. Usually there is a large effort made during the design phase to optimize continuous beam losses and keep these as low as possible.

Accidental beam losses are transient losses with time scales from nanoseconds to many seconds. These can occur due to a multitude of failure mechanisms. Machine protection protects equipment from damage, activation, and downtime due to accidental beam losses. Machine protection includes a large variety of systems performing this function. For example: for 1 MW, a loss of 1% corresponds to 10 kW, not to be lost along the beam line to avoid activation of material, heating, etc. Assuming a length of 200 m, such losses would correspond to 50 W/m. However, the 1 W/m rule is a reasonable limit for hands-on maintenance, see Ref. [2].

2.1.3 Activation – hands-on maintenance

An average beam loss of 1 W/m should be a reasonable limit for hands-on maintenance, where 1 W/m corresponds to 6×10^9 protons per [metres times seconds] of energy 1 GeV/c (uniformly distributed). Simulations show that if the irradiation time of a steel pipe with 1 W/m beam losses is 100 days and cool down time is then 4 hours, the effective dose at 30 cm distance from the beam pipe is about 1 mSv per hour, see Ref. [4].

2.1.4 Environmental conditions for equipment

Another criterion to be considered is the environment in which the equipment and its electronic controls are located. It is advisable to e.g., avoid electronics in radiation areas, or areas where high temperatures and high humidity might occur, exposure to water, fire, etc.

It is very important to supply the equipment with rich diagnostics, such as temperature sensors, flow switches, or other sensors, and it is also very important to archive these data in a good way, such that further analysis can be easily performed. Such offline analysis can help to improve the performance of the machine.

2.1.5 Means to inject, extract, stop beam

It is important to understand how the beam is injected into the machine and how it can be extracted or stopped and on what time scales this can be achieved. In linacs it is common to stop the beam by interrupting the source (stop the extraction of plasma), and to use the choppers to deflect the beam to the absorbers. In ring machines usually kicker magnets are used to kick the beam out of the ring into a dump line. Extraction or interruption of beam operation should preferably be possible before melting, or other beam-induced damage with high severity, can occur. Otherwise passive protection needs to be implemented or a re-design undertaken, see Ref. [2].

2.1.6 Availability requirements

Each facility has different requirements on beam and machine availability. This is often because of different sponsoring schemes and budgets, manifold user communities, and scientific goals set up for the specific machines. User and medical facilities usually have very strict availability requirements due to their nature. Many efforts have been made in defining one common availability measurement and analysis method and metrics. This might be very helpful when comparing several machines with each other. This practice works well when comparing facilities of the same type. For example synchrotron light sources. These are strongly user driven and use similar principles and technologies.

2.1.7 Foresee the unforeseeable: how?

Newer facilities nowadays aim towards unprecedented beam energies and beam power levels, and the upcoming failures or incidents are difficult to foresee. This needs to be considered, and planning for rich diagnostics and/or redundancy at an early stage can be crucial to achieve successful operation.

To illustrate what is meant here, we take the example of unknown falling objects (‘UFOs’) detected in the LHC at CERN during 2010 and existing since. These UFOs are probably dust and dirt particles of a few micrometres size only, creating beam losses when ‘falling’ into or through the beam. This phenomenon was not observed before at any other machine and is hard to predict. However, it was of tremendous advantage to have a very large beam loss monitoring system consisting of 4000 detectors installed around the 27 km long ring. This system allows for a wide coverage and detection of all types of beam losses, from ultra-fast, to fast, to slow. The measurements of this system are taken and archived with high quality and hence it is possible to perform very efficient data analyses. Certain patterns of beam losses caused quenches of magnets and happened very fast (within a few tens of microseconds). The shape of the losses looks similar to the losses created when doing a wire scan, i.e., moving a wire through the beam to measure the beam profile. This is why in the beginning these type of losses were considered to be caused by something falling through the beam. It is very difficult to predict such issues, but reliable beam monitoring systems with a large dynamic measurement range and several integration times are vital to understand them.

2.1.8 Time scales

As mentioned above, failures can be categorized based on severity, but it is also important to categorize them according to other criteria, such as whether the damage is beam induced, non-beam induced, local but impacting on beam, local but not impacting on beam performance. Besides this, it is vital to understand the time scales during which damage can happen. Time evolution scales for damage events can be slow [days, hours, seconds], fast [milliseconds], ultra-fast [nanoseconds, microseconds]. The time leading to a damage event includes several parts: the start time; e.g., a magnet power supply fails and does not provide any voltage to the magnet anymore. Then, there is a certain decay time of the magnetic field during which the effect on the beam will be different. In the beginning, when the magnetic field is only a few per cent lower than anticipated, almost nothing will happen to the beam. At a certain moment in time however, beam losses will occur. Eventually, it can happen that the beam is deflected in a way such that damage to surrounding equipment could happen. Depending on these time scales involved, it is possible to define the timing requirements for the protection functions (PFs). Preferably one can set the goal for machine protection to act in a preventive way first, but always anticipate a layer of protection that can mitigate the event. If damage can happen very quickly, there might be a need to change from active protection to passive protection or to go back and re-design the system, see Ref.[2].

2.1.9 Summary

Machine protection is a facility-wide, availability-driven system of systems, which has equipment protection as its main goal in order to prevent long downtimes of the facility. It should be noted that the primary purpose of machine protection is not to directly handle internal failures of equipment, but to be involved in the prevention of additional damage that can arise, as an extended effect of the equipment failures or system interactions.

3 Means to achieve machine protection

Summarizing the first two chapters, the following ‘means’ help in achieving high operational availability of accelerator driven research facilities.

3.1 General machine protection (MP) requirements

In the following we summarize the most generic requirements for machine protection.

– Designing and operating the equipment under control (EUC) with high inherent reliability and overall low damage potential.

– Minimization of the mean down time (MDT) of EUC by introducing dedicated technical systems preventing and mitigating damage.

– Minimization of the MDT of EUC systems by introducing dedicated operational and preventive maintenance procedures reducing the probability of the need for corrective maintenance.

– Introducing supporting systems dedicated to reducing MDT. These include analysis, management and recovery tools addressing operational activities related to machine protection.

Machine protection functions should be implemented with timing and protection integrity levels (PILs) in accordance with damage risk reduction requirements. This implementation has to be done such that the probability of spurious or false trips is reduced in accordance with the availability and damage risk reduction requirements.

Machine protection shall transmit all necessary information to the responsible staff, allowing them to take adequate action to resume facility operation within a minimum amount of time. All information about detected off-nominal states should be recorded. Performed prevention and mitigation actions should be recorded as well to allow for a posteriori event reconstruction and analysis. Furthermore, machine protection should support operation during all foreseen lifecycle phases of the machine including, but not limited to, assembly and installation, commissioning, tuning, operation, maintenance, and dismantling.

Machine protection should support all foreseen operating modes of the machine, including but not limited to beam up to intermediary targets, beam with reduced beam power, or alternative duty cycles. Machine protection should also support operation in case of a degraded mode of operation of the EUC if required for reaching the availability goals and if compatible with damage risk reduction requirements. Also, it should support operation in case of degraded protection functions, if required for reaching the availability goals, and still be compatible with damage risk reduction requirements.

3.2 The 'P^3' principle – tasks of machine protection

The purpose of machine protection, as mentioned before, is to reduce the scientific and economic losses as much as possible. The task is to optimize damage events such that their occurrence is kept at a level that is as low as reasonably achievable. This can be done by simply interrupting beam operation and adjusting relevant parameters to avoid damage and downtime. On the other hand, simply stopping everything in case of 'any' errant situation in a complex facility would lead to very little operational time. Instead, machine protection needs to be able to handle minor problems in the background while operation continues.

In addition to protecting equipment and avoiding unnecessary downtime, information on the cause of the stop can be collected for further analysis, to allow for continuous operational improvements. The tasks for machine protection for an accelerator facility can then be summarized as the 'P^3' principle, see Ref. [5].

– Protect the equipment.

– Protect the beam.

– Provide the evidence (of what caused the stop).

Protecting the equipment means avoiding damage due to wrong behaviour or a wrong configuration or due to the beam. Protecting the beam means that interruptions of beam operation should be kept at a minimum level. To provide the evidence, a so-called post-mortem system is implemented. This is a system that, in case of a beam stop, collects data on the current machine configuration, time stamps of when the event occurred, and what system of the machine sent the beam stop signal.

3.3 A system of systems approach to managing machine protection

As the systems required for successful research facility operation are typically diverse in both technology and ownership, a traditional approach, where the integrated systems are treated equally, would not handle this diversity satisfactorily. Instead, a system of systems (SoS) approach is more suitable, see Ref. [6], as this appreciates the emergent properties and operational independence of the constituent systems. According to Ref. [7] and Ref. [8], a SoS has the following criteria for the constituent systems, which are found (by the authors) to match well the protection approach of a complex research facility.

- Operational independence.
- Managerial independence.
- Geographical distribution.
- Emergent behaviours.
- Evolutionary behaviours.

The two latter criteria are not apparent if the systems and their interactions are modelled separately. These five criteria are also sometimes referred to as Maier's criteria of a system of systems, see Ref. [8]. Thus, a complex research facility is a typical example of a SoS with many subsystems that jointly perform a task that could not be done by one system alone. It should be noted that SoS is not a tool, method, or practice in itself, but rather a way of thinking. Therefore, it can be well adapted into the functional protection method, see chapter 5. Adapting this way of thinking naturally expands beyond a limiting 'root cause' mentality as the sole solution to protection problems.

4 Machine protection at ESS: a system of systems

This chapter provides an overview on how to apply the SoS approach at a complex research facility such as the European Spallation Source, Lund, Sweden for machine protection purposes.

4.1 Introduction to the European Spallation Source

The European Spallation Source (ESS), currently under construction in Lund, Sweden, will be a user facility for neutron science, aiming to be the brightest neutron source in the world throughout the next decades.

Protons are accelerated to 2 GeV/c along a 600 metres long linac. The linac is operated in pulsed mode and 2.86 millisecond long pulses are injected at a repetition rate of maximum 14 hertz. The average beam power is 5 MW with a peak power of 125 MW per pulse. The 2 GeV/c protons are steered into a rotating, helium cooled, 5 tons tungsten target where the neutron spallation process takes place. Then, the created spallation neutrons are further moderated to appropriate speed and guided towards the neutron science stations where the different experiments can take place. The first proton beam production will take place in 2018 and the first neutron user program will start in 2023.

The ESS proton beam power will be unprecedented and melting of copper or steel can happen within a few (2−10) microseconds, if the beam is deflected perpendicularly into the beam pipe. This is especially an issue in the normal conducting linac (first 50 m), where the energy of the beam is still low (75 keV – 90 MeV).

4.2 ESS availability-driven machine protection concept

Machine protection at ESS has its starting point in the operational goals and user needs for neutron experiments. From Ref. [10], two relevant goals are the following:

– StR-11 – Annual operating periods: “The partnership expects the ESS scientific infrastructure to be available at least 4000 h of each annual operating period when fully operational”.

– StR-12 – Availability: “The partnership expects that the ESS facility will be designed with a goal of 95% availability during its annual operational periods when fully commissioned”.

Based on these, the ESS Neutron Source Reliability and Availability document, see Ref. [10], outlines a strategy to break the overall goals into more concrete and acceptable downtime requirements for the different ESS systems. Machine protection then transfers this into protection needs, plays a key role in its fulfilment, and integrates a facility-wide strategy and analysis effort to reach the ESS goals for neutron availability. This effort will be described in more detail in the next sections and further in chapter 5.

4.3 ESS machine protection as system of systems

At ESS, a single protection function can be performed by different parts of the machine. These parts are managed by different divisions or groups and designed and built by different laboratories around Europe. This requires a way of organizing the responsibilities, which can be achieved by applying the SoS approach. This approach is currently followed at ESS. All systems that belong to the MP-SoS play a role in the protection of the machine.

The ESS project organization is set up to address each of the three sections of the facility through a dedicated division/directorate: the proton accelerator, spallation target, and neutron science systems. Within each division/directorate, there is a set of work packages that deal with specific parts identified in the facility breakdown structure, Fig. 1. The MP-SoS analyses need to merge into the development of the different work packages at the same time as they tie back to the overall goals of ESS. MP-SoS are therefore continuously observing and analysing the progress of the different work packages, in order to cope with the complexity and system interaction that might lead to damage and beam losses that would cause ESS to not fulfil the overall goals connected to operational availability. The selected risk management method has to be flexible enough to handle continuous scope expansion by allowing modular and parallel work for the different work packages, see Ref. [11].

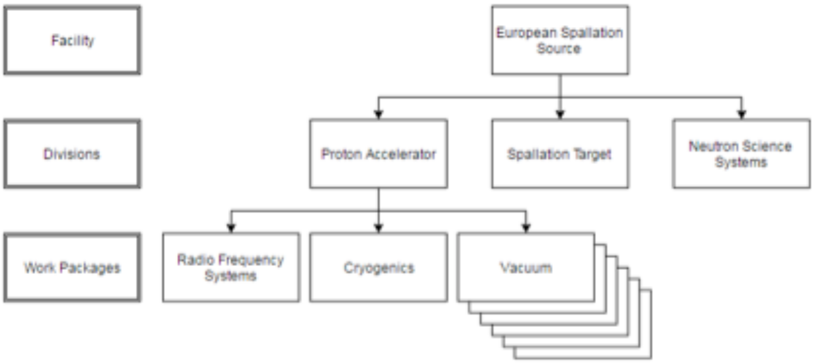


Fig. 1: Part of the facility breakdown structure at ESS

4.4 ESS organizational structure and machine protection responsibilities

The ESS organization has and continues to expand rapidly, hence the culture, communications procedures, and all of the administrative efforts, have to be developed in parallel to the design and construction of the facility. The selected risk management method needs to cope with limited procedures for documentation as well as changing organizational structures internally and externally. In addition this places high demands on flexibility and continuous updates of the information to be analysed. Traceability then becomes an important aspect already in the early stages, in order to verify that the analyses are matching the latest versions of e.g., the hardware.

The rapid expansion can also be viewed from the perspective of scope, where the MP-SoS includes more systems as the project proceeds, and there is a need to inform and set up discussions with new stakeholders. These stakeholders are themselves required to perform concept designs of their systems at the same time as potential interfaces are changing. It is therefore important for the technical risk managers to recognize that the organizational development affects the technical developments, including systems relevant for machine protection, see Ref. [11].

The protection of the machine occurs at multiple levels. At ESS, three main levels are considered.

The first is at component level, where the manufacturer or the group building it takes care of the protection. Some examples are electronic boards, power supplies, or any stand-alone piece of equipment. This level is quite straightforward to distinguish and it is the responsibility of each system owner.

The second level is related to protection in-between components. This protection generally involves direct relationships between components in the form of services and connections. Some examples of direct relationships are water cooling, controls, power supply, etc. In these cases, the components usually belong to the same system and therefore the protection is the responsibility of the system owner. The protection functions involved in this category are called local protection functions. The example in Fig. 2 shows the local protection for a magnet system where the power supply performs the protection of the magnet. The power supply has to stop delivering power if overheating is detected in the magnet coils. Local protection is different for each system, some of them are quite complex and others are relatively simple.

The third level is focusing on protection functions that are located in-between systems. These functions usually cover damage events that require stopping the proton beam operation to bring the machine to a protected state. In Fig. 2, the second protection function (stop beam operation in case the magnet is not powered since the beam could be deflected in an unwanted and potentially critical way) is performed trough the magnet system, the beam interlock systems (BIS) and the different actuator systems. These together form a so-called global protection function. The ESS MP team is in charge of analysing these functions, retrieving the correct information about the systems that participate in performing these functions, and making sure that the requirements are fulfilled.

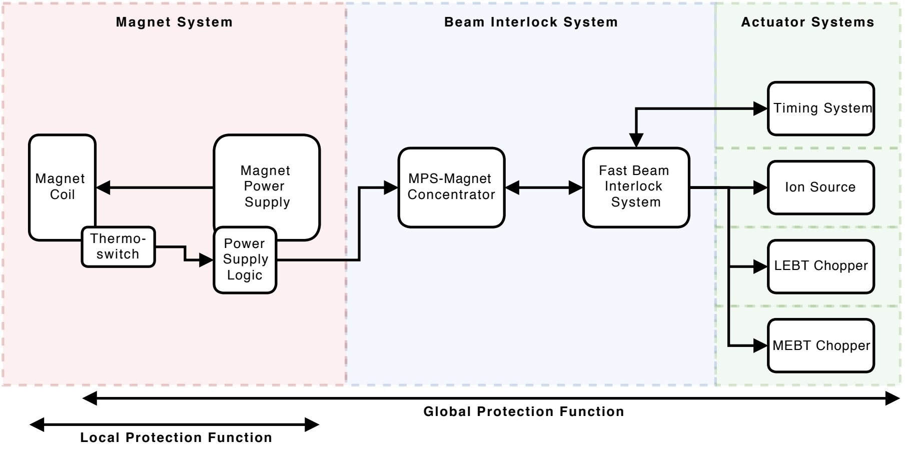


Fig. 2: Example of a global and local protection function as defined at ESS

4.5 Main function of the ESS MP-SoS

The main function of the MP-SoS is the following:

– MP-SoS has to achieve or maintain the protected state of the machine.

If all systems are properly operating and the beam is well controlled, the machine is in a protected state. However, if something that could lead to damage is detected, then the MP-SoS has to act accordingly to reach the protected state again. In order to achieve or maintain this state, the MP-SoS systems have to detect the hazard, interpret if it is acceptable and in case it is not acceptable, stop operation.

4.5.1 Global protection functions and beam related hazards

The global protection functions protect against hazards that are beam related. Some examples of beam related hazards are the following.

– Proton beam is out of the expected position/path and critical beam losses are generated (e.g., beam not well focused or steered or an element is inserted in the beam pipe unexpectedly).

– Beam power exceeds design specifications for a specific beam destination (BD) or insertable device (ID).

– One of the MP-related systems is not operative (e.g., RF transmitter is off and then beam would not be properly accelerated).

For these cases and others, the way to achieve or maintain the protected state is to stop beam operation or to not allow beam operation.

4.5.2 Beam related damage protection

Protection against beam-induced damage can be provided in the following ways.

– Evaluate the status of critical components or systems through the status signals from their local protection systems (LPS) or their dedicated MP systems. Critical components or systems are the ones that might cause beam-induced damage in case of failure (e.g., magnets or cavities).

– Use information from the beam monitoring systems to prevent or mitigate beam-induced damage. For example, in the case that a steerer magnet is not properly operated, the beam can be deflected, but a beam position monitor can detect this and trigger a beam stop before beam losses occur. BCMs or beam loss monitors (BLMs) can be used to detect beam losses. Beam monitoring systems can trigger a beam stop in case critical beam parameters are being detected.

– Ensure that beam parameters are within the limits defined by the configured beam mode (BM) and that the beam reaches the configured BD.

– Define procedures (e.g., for ramp up in beam power), restrictions in the control system (e.g., steerers maximum configurable current), or configuration control (e.g., firmware versions for critical electronic devices).

4.5.3 Interplay of MP systems

Figures 3 and 4 show the main types of systems that can generate the so-called beam interlocks and how they are connected to the BIS. The BIS consists of the MPSs and the fast beam interlock system (FBIS). The BIS is responsible for interpreting inputs and deciding if, based on the BM and BD, the beam has to be stopped or not. Finally, the FBIS requests the different actuators to trigger a beam stop. Since there are different ways to trigger the actuators, the status of the actuators is used to escalate in case of a failed trigger.

Systems that connect to the FBIS might generate ready signals (where the FBIS does not require an external reset) or beam permits (where the FBIS requires external reset). The decision of generating either a beam permit or a ready signal for each system might change through different phases of the project. These signals are grouped here under the umbrella name interlock.

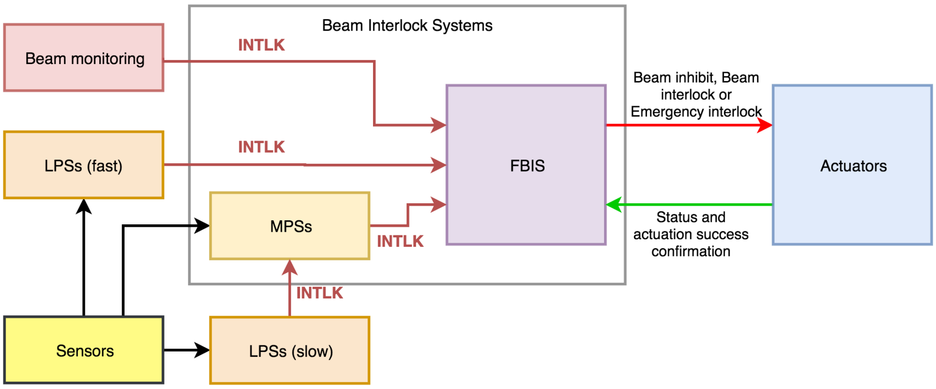


Fig. 3: Interaction of systems involved in achieving or maintaining the protected state of the machine

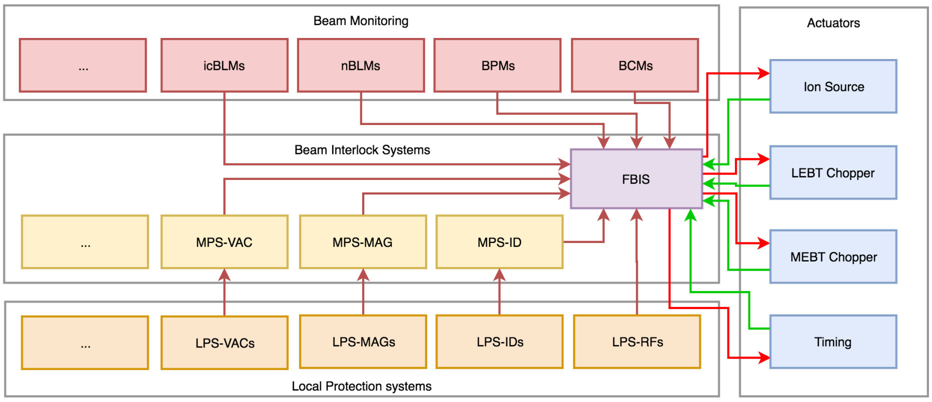


Fig. 4: Examples of systems involved in achieving or maintaining the protected state of the machine and their interaction.

5 A functional protection method and its lifecycle

5.1 The need for a functional protection method

In the field of safety engineering, formal processes for ensuring the correct level of safety are required and there are several existing standards tailored for the operation of large facilities, see Refs.[12–15]. Relevant safety systems and their product lifecycle claim compliance with a safety standard and culminate in a certification process, which is necessary to receive the green light for commencing operation. When protecting hardware systems and equipment from damage, there is no legal obligation to follow any international or national standard. Still, some facilities apply a formalized process for ensuring a protected system has been implemented, see Ref.[16] and Ref.[17]. In parallel with pushing the technology of modern research facilities forward, the demands on reliability and availability of the equipment increase as well.

This also goes for ESS. Machine protection at ESS has been identified as an important driver for successfully reaching the availability goals of the facility. The MP strategy is to identify and analyse systems and devices that play a role in this goal and adapt their functional behaviour accordingly. Hence it was required to develop a robust method that allows management and analysis of identified risks in the context of availability driven machine protection.

Due to these demanding requirements, it is suitable to put in place a method that efficiently deals with a vast number of interacting components and systems, but at the same time incorporates flexibility and resilience in order to meet the challenges in changing project scopes, which are typical for the design and construction of complex research facilities. Traditional safety methods follow a classic waterfall progression where each specific item is dealt with before moving to the next, see Ref. [7]. While this makes sure that the method is applied in such a way that the risks are accurately handled, it would be difficult to enforce alongside the actual development of the systems and components themselves. Machine protection, as opposed to safety, needs to be intertwined with the systems themselves and often occurs on the system level rather than as a standalone.

The functional protection method, described here, has been developed at ESS and is an abductive, holistic approach that deals with the various hazards and protection-relevant subsystems, and derives a set of protection functions that treat risks in an appropriate and reliable way. The method of defining protection functions can be applied at any stage during the facility lifecycle, being it the design, commissioning, or operational phases.

5.2 What is functional protection?

Functional protection is a technical risk management method suitable for application to a SoS or other complex system. The method was developed at ESS together with the team of functional safety experts from Zurich University of Applied Sciences (ZHAW), Winterthur, Switzerland and can be integrated into the design, early commissioning, and operations phases of accelerator driven facilities to enhance their reliability and availability, see Ref.[18]. The risk management method for MP-SoS follows the applicable ISO standards 31000 and 16085. These standards outline the process of establishing a risk management context followed by identifying, analysing, evaluating, and treating the risks in a traceable and transparent way. Additionally, the IEC 61508 [12] and 61511 [13] standards for functional safety of electric, electronic, and programmable electronic (E/E/PE) systems are guiding the lifecycle approach as well as the classification of protective functions. This places ESS risk management within a robust framework that allows for efficient and purposeful analyses and accurate treatment of the technical risks associated with running a complex high power accelerator facility. The method is currently applied at ESS and builds on close collaboration between system owners and risk analysts, see Ref. [11].

5.3 The functional protection lifecycle

The lifecycle is shown in Fig. 7. It covers the concept and scope definition, analysis, as well as the integration, specification, design, implementation, installation, testing, operating, and adjusting phases.

Three interconnected teams are identified as vital for the functional protection lifecycle progress and execution of MP-SoS related work, see Fig. 8.

- Protection analysis team (PAT).
- Integrated protection team (IPT).
- Implementation and design team (IDT).

The three groups have continuous interaction and follow ups during the lifecycle of the system. At ESS the PAT and IPT are fixed but the IDT varies depending on the analysed system (work package) [19]. The teams are involved in all lifecycle steps but are the mainly responsible for certain parts in the lifecycle, which is described in more detail in the following.

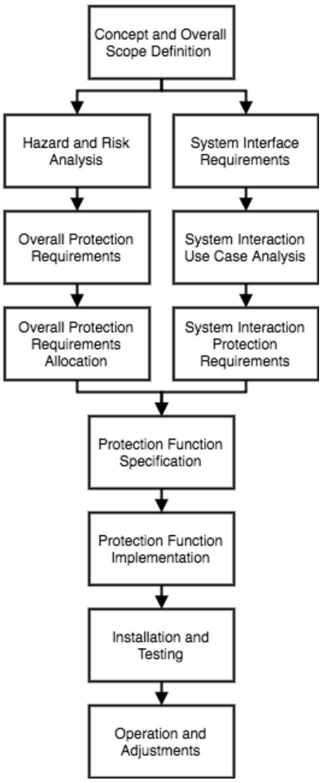

Concept and Overall Scope Definition
Hazard and Risk Analysis
System Interface Requirements
Overall Protection Requirements
System Interaction Use Case Analysis
Overall Protection Requirements Allocation
System Interaction Protection Requirements
Protection Function Specification
Protection Function Implementation
Installation and Testing
Operation and Adjustments

Fig. 5: The functional protection lifecycle. All steps of the process are iterative and might need to be revisited multiple times.

5.3.1 Concept and overall scope definition

The functional protection lifecycle starts by defining a scope and concept for a specific system that plays a role for machine protection.

This step defines the equipment to be protected (EUC), its environment, and what goals are to be met for its protection and availability. It is also useful to specify the types of damage events and hazards that are to be studied in the next step of the lifecycle, in order to create a clear boundary for the analysis. Like this one defines the systems and devices to be included in the analysis, their environment (where appropriate), the damage sources to be considered (e.g., electrical, mechanical, thermal, radiation), and the operational modes where the damage can occur.

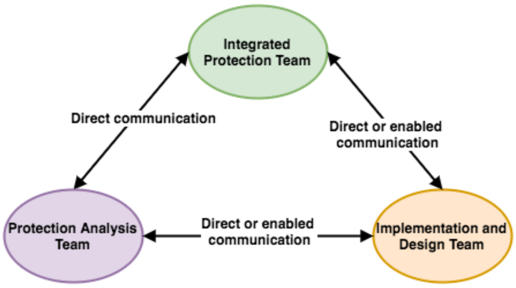

Integrated Protection Team
Direct communication
Direct or enabled communication
Protection Analysis Team
Direct or enabled communication
Implementation and Design Team

Fig. 6: The three groups and their interactions as identified by the functional protection method

5.3.2 PAT and its role

The PAT is in charge of translating global ESS requirements into manageable requirements and functions to be implemented into the hardware and software of the facility. The PAT is responsible for the hazard and risk analysis, the overall protection requirements, and the overall protection requirements allocation.

As the ‘start’, or root cause, of a damage event is purely subjective [20], attempting to define this is completely left out of the analysis. Instead, the functional protection analysis technique has the goal of identifying how each device can be damaged, as opposed to how the device can damage something else.

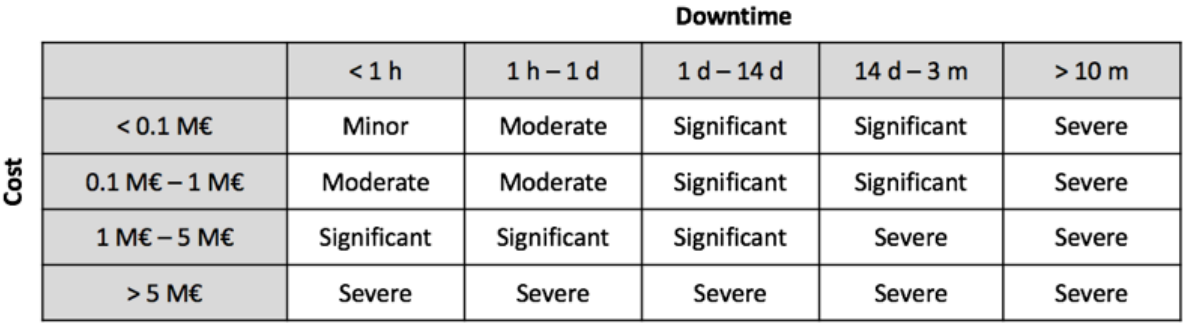
It is important that the teams involved in functional protection are aware of and consistent in the concept and scope of the analysis and risk management. This is primarily done in the risk identification, where the systems and devices that are agreed to be involved are documented.

5.3.2.1 Risk management process

The risk management process for MP at ESS identifies and analyses the damage events that are to be prevented or mitigated by machine protection. However, it is still the responsibility of the system owners to design robust and reliable systems in line with the ESS requirements. Due to this some of the systems and equipment contain certain protection barriers of their own, categorized as other risk reduction measures in the MP risk management framework, in accordance with the IEC 61508 standard. The remainder of the protection functionality is what is required to be carried out by MP-specific PFs, see Ref. [21]. The risk management process applies the same approach, coordinated centrally by the MP personnel, to be taken towards all of the systems and equipment present at ESS. This unifies the analysis efforts and allows for a natural overview of the MP-relevant systems and their role in fulfilling the overall ESS machine protection goals.

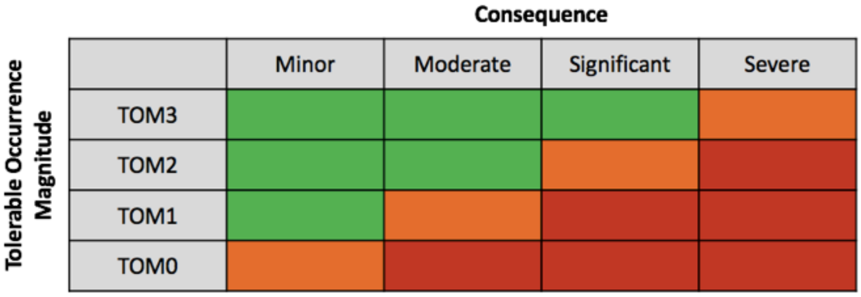
5.3.2.2 Damage event and risk ranking

Damage events are ranked according to their consequence for ESS operation. The consequence is defined through two risk matrices, as seen in Figs. 9 and 10. Risk matrix 1 defines a consequence category based on the cost and facility downtime associated with the damage event. Risk matrix 2 then takes this consequence category and associates a tolerable occurrence magnitude (TOM) with this event, where TOM0 is most relaxed and TOM3 has the highest demands [22].



Cost		Downtime				
		< 1 h	1 h – 1 d	1 d – 14 d	14 d – 3 m	> 10 m
	< 0.1 M€	Minor	Moderate	Significant	Significant	Severe
	0.1 M€ – 1 M€	Moderate	Moderate	Significant	Significant	Severe
	1 M€ – 5 M€	Significant	Significant	Significant	Severe	Severe
	> 5 M€	Severe	Severe	Severe	Severe	Severe

Fig. 7: Risk matrix 1, defining the risk category based on downtime and cost of a damage event



Consequence

Tolerable Occurrence Magnitude	Minor	Moderate	Significant	Severe
TOM3	🟩	🟩	🟩	🟧
TOM2	🟩	🟩	🟧	🟥
TOM1	🟩	🟧	🟥	🟥
TOM0	🟧	🟥	🟥	🟥

Fig. 8: Risk matrix 2, defining the tolerable occurrence magnitude for a damage event of a specified risk category.

Each level corresponds to an order of magnitude for tolerable mean time between occurrences of the damage event. In the case of ESS, TOM0 is once every 5 years, TOM1 every 50 years, TOM2 every 500 years, and TOM3 every 5000 years. The executive aim of the following risk management process can then be viewed as 'moving' damage events that are outside of the TOM into the tolerable region by defining PFs that mitigate the risks.

Identifying and analysing each damage event through the two risk matrices allows for the addressing of each hazard that could lead to this damage event. The hazards are defined top-down, meaning that the hazards that directly lead to a damage event, need to be defined first. Based on the expected occurrence (EO) of a hazard, each hazard frequency is estimated qualitatively by the PAT and IDT. These estimations are named EO0 (normal operations), EO1 (lifetime of facility), and EO2 (unexpected), and are specifically used to avoid unverified expert opinion on the probabilities of hazards, but at the same applying useful pragmatism. Normal operations are events within the specified and expected conditions and restrictions of the system, including all operating conditions; lifetime of facility refers to events that are outside the specified conditions and restrictions, but can be expected during the facility's lifetime, also including frequent disruptions that are overcome by the control systems; and unanticipated implies events outside the specified conditions and restrictions which are not expected to occur during the facility's lifetime.

5.3.2.3 Analysis process

Each hazard then receives one corresponding overall protection function (OPF), which is a generic and technology-free set of objectives compiled and associated with each system, 'negating' the associated hazard. The OPF is in itself only an intermediate step in defining the PF, but is useful in framing the protection efforts and should be checked for its objective in the concept and overall scope definition. The OPF also receives a functional integrity magnitude (FIM), which is the difference of the damage event's TOM and the EO level. That is, FIM=TOM−EO. As an example, the OPF for a TOM3 damage event and an EO1 hazard receives a FIM=3−1=2. The EO level thus corresponds to a reduction in the (occurrence) order of magnitude that needs to be treated, where 0 is no reduction, 1 is a reduction of one order of magnitude (10), and 2 is a reduction of two orders of magnitude (100). This way, the OPFs can be traced back to the corresponding damage event to verify that the TOM is fulfilled by the OPFs, before proceeding to the final step below.

The OPFs are then subjected to audit by the PAT together with the system experts, in order to derive technology-specific PFs, containing the sensors that monitor the hazard, the logic elements that take the decision on what action is required, and the actuators that carry out this action. The hierarchical analysis flow is seen in Fig. 11, starting at the system level. In addition, the PFs include a timing requirement for how quickly the PF needs to be carried out, and a PIL that gives requirements on the quality of the PF [22], [12].

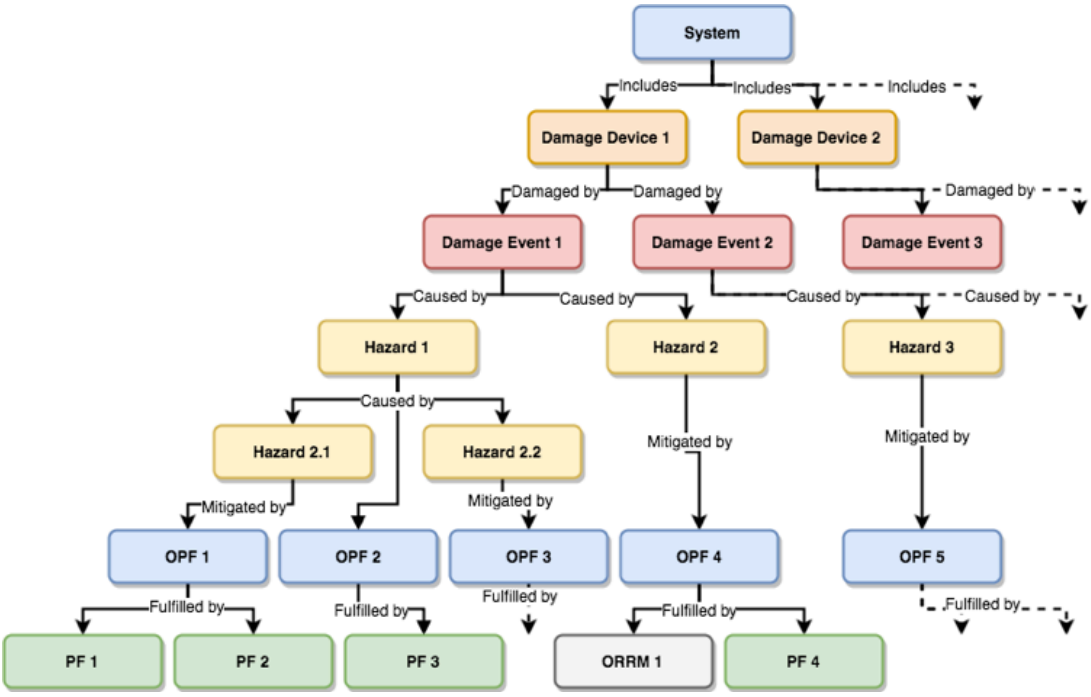


Fig. 9: The Functional protection abductive risk management process, starting at the system level and ending at the level of PFs.

For all of the identified PFs, the MP-related sensor signals are sent to a system specific logic element. This logic element performs the sensor data analysis and further distributes the signal to the ESS fast beam interlock system (FBIS) when a beam stop is required to prevent or mitigate a damage event. The FBIS receives inputs from the entire ESS facility and makes the final decision on when to stop the proton beam. If a beam stop is required, the signal is sent to the actuator systems. There are four systems at ESS that are used to stop beam operation. They are the ion source, the low energy beam transfer (LEBT) chopper, the MEBT chopper and the timing system.

All of the information and risk management process steps are required to be traceable and readily available for all relevant parties. This allows for a continuous workflow where needed, and all associated parts can follow and, where necessary, contribute to the analysis process. Once a set of PFs has finished its internal iterations and suitability checks, it is documented and uploaded to the official ESS document management system for approval by the ESS machine protection committee (MPC).

5.3.3 IPT and its role

The IPT analyses interfaces between protection-related systems and the rest of the facility. The analysis and integration work is mainly done in parallel, but it is beneficial to have access to a draft of the overall protection requirements and an idea of the architecture of the involved system (functional architecture) before starting the integration work. The IPT defines system interface requirements, coordinates the system interaction use case analysis and defines system interaction protection and behavioural requirements.

To obtain the system interaction model, behavioural requirements and to allocate PFs, system interaction use case workshops are held. The feasibility of different system architectures and PF implementations are discussed and simulated by going through foreseen operational sequences, use cases. The different architectures and use cases are documented using enterprise architect. Before the work of the IPT starts, a representative for each protection relevant system has to be identified. The representative could be the designer, owner, or contact person for that system.

The representatives are then gathered together with the IPT during a series of workshops. During the workshops, the architecture, functionality, and possible interfaces of each system are discussed. Based on the concept and overall scope definition, certain PFs have to be performed. Depending on the level of progress of the development, the design and interfaces are more or less constrained. Typically, a more developed system is less flexible. This could cause difficulties if certain functions or hardware have not already been foreseen. When everybody has a clear view of the involved systems architecture, the allocation of PFs is discussed. Questions such as, 'Which sensors are available?', 'Where should the logic be implemented?', and 'How is the information transmitted?' are asked and answered.

A certain architecture and interface set is suggested and fixed for the coming system interaction use cases. The purpose of the use cases is to examine and document the intended interaction between the involved systems from a machine protection perspective. The goal is to derive appropriate and robust solutions for the signal types and interfaces. The use cases are one way to specify interfaces and requirements (especially with respect to PFs), however they are not the unique source. The parallel work of PAT also results in a set of PFs and requirements.

The use cases are set up to follow scenarios which will appear during regular machine operation such as a change of BD or a rearm after an interlock. Besides the 'normal flow', which documents the operational scenarios as they are expected to happen, so-called 'alternative flows' are examined. Alternative flows document how the systems react if the scenario deviates from the normal flow. Typically, this is due to a fault or failure of a system or component.

Each action or information exchange failure would cause a different sequence of events. It is not feasible to go through every single alternative flow due to time constraints. Instead, a set of alternative flows is selected based on the events that are most likely to fail or events that seem to have the largest damage potential. The selection of alternative flows is done by estimations of the involved experts. If unacceptable behaviour of the systems is observed due to a failure in the normal flow, it has to be adjusted. The normal flow should be adjusted to increase the robustness and decrease the vulnerability by changing the normal operational sequence.

The system interaction protection requirements will, similar to the functional protection analysis technique, propose a set of PFs. The requirements derived by the IPT are behavioural requirements. They are defined on the system level and based on the interfaces and interactions between protection-related and other systems. This analysis path completes the picture of the protection requirements in a way that is not possible through the damage-based analysis alone.

One or more use cases are selected and played through with the selected architecture. During the use case studies, it may be identified that systems perform well as standalone systems but that the signal exchange with other systems is flawed or causes a different action than expected. This analysis may also derive additional interfaces that are then added to the system interface requirements.

During each of the lifecycle phases previously defined properties might need to be adjusted. This involves the architecture, the allocation of PFs, system behaviour and interaction.

5.3.4 PF specification

The two tracks that merge into the PF specification in Fig. 7 apply two different approaches to protection. While the leftmost track uses a top-down, abductive analysis that yields requirements for the quality of the PFs, the rightmost track applies inductive reasoning to verify that possible failure modes, including undesired control actions, are handled correctly. This reasoning viewpoint is seen graphically in Fig. A, where the abductive analysis is ‘top-down’ oriented and the inductive analysis is ‘bottom-up’.

5.3.5 IDT and its role

When the work of the two teams has reached a satisfying maturity, the IDT starts the implementation work. Finally, the IDT undertake the implementation of the requirements and serve as system experts in the discussions on possible and relevant designs and functionalities. The IDT plays an important role in the design and analysis phase by contributing with experience and judging the feasibility of different designs.

The installation of the protection-related systems is detailed in the system requirement specification for each system. Before installation of protection-relevant equipment, separate tests (commonly referred to as factory acceptance tests) are performed in the laboratories or factories where the equipment for the functions is assembled. These tests need to verify the functionality and the timing requirements of the systems, as well as their correct interface to the input and output systems. Once the complete functions are installed, a complete test of the functionality (site acceptance test) is carried out to verify that the requirements are met.

It is not possible to practically verify every aspect of a control or protection system before the facility is taken into operation, and therefore the first years of the facility operation (the initial operation) need to be dedicated to continuous follow ups and adjustments of the equipment and its behaviour. The theoretical numbers based on analysis assumptions do not always agree with what is observed in the first stages of the facility’s operational lifetime, but become more relevant as the facility and its organization and procedures mature into a steady state. This also displays the importance of an up-to-date risk record to be traceable and easily available.

6 Summary and outlook

As discussed in this paper, modern particle accelerators are increasingly complex and require robust strategies for them to be operated without damage and with the desired availability. Due to their complexity and vast number of systems and sub-systems, it is argued that these facilities cannot be analysed and treated accurately by applying traditional reliability approaches alone. Instead they need a system of systems mentality as well as recognition of the emergent and often incomplete descriptions of its complex properties.

The functional protection method, presented in chapter 5 aims at being the link between complex systems and acceptable damage risk. Following functional safety standards for protection purposes and connecting the lifecycle steps to proven-in-use risk management and analysis methods would give partial confidence in this link.

However, it is important to keep in mind that the machine protection objectives can only be achieved if a large number of systems, developed by different groups and divisions, interact in a well-orchestrated way.

Acknowledgement

We wish to thank our colleagues from CERN, SNS, ZHAW, and ESS for enlightening comments on the present topic.

References

[1] R. Schmidt, Proceedings of the 2014 Joint International Accelerator School: Beam loss and Accelerator Protection, Newport Beach, United States, 5–14 November 2014, edited by R. Schmidt, CERN 2016-002, https://doi.org/10.5170/CERN-2016-002.

[2] R. Schmidt, in Proceedings of the CAS-CERN Accelerator School: Advanced Accelerator Physics, Trondheim, Norway, 18-29 August 2013, edited by W. Herr, CERN-2014-009 (CERN, Geneva, 2014), pp. 221-244, https://doi.org/10.5170/CERN-2014-009.221.

[3] W. Blokland and C. Peters, A new differential and errant beam current monitor for the SNS accelerator, 2nd International Bean Instrumentation Conf., Oxford, UK, 16−19 September 2013.

[4] N.V. Mokhov and W. Chou, 7th ICFA workshop on High intensity high brightness hadron beams, USA, 1999.

[5] R. Schmidt *et al.*, *New J. Phys.* **8** (2006) 290, https://doi.org/10.1088/1367-2630/8/11/290.

[6] T.Friedrich, C. Hilbes and A. Nordt, Systems of systems engineering for particle accelerator based research facilities: A case study on engineering machine protection, 2017, https://doi.org/10.1109/SYSCON.2017.7934806.

[7] A.P. Sage and C.D. Cuppan, *Inf. Knowl. Syst. Manag.*, **2** (2001) 325.

[8] D. DeLaurentis and R.K. Callaway, *Rev. Policy Res.* **21** (2004) 829, https://doi.org/10.1111/j.1541-1338.2004.00111.x.

[9] M.W. Maier, *Syst. Eng.*, **1** (1998) 267, https://doi.org/10.1002/(SICI)1520-6858(1998)1:4<267::AID-SYS3>3.0.CO;2-D.

[10] E. Bargalló, "ESS reliability and availability requirements," ESS Document (ESS-0008886), 2015.

[11] R. Andersson, S. Kövecses, E. Bargalló and A. Nordt, Challenges in technical risk management for high-power accelerators, ICANSXXII, 2017.

[12] IEC 61508:2010: Functional safety of electrical/electronic/programmable electronic safety-related systems. 2010.

[13] IEC 61511:2003: Functional safety - Safety instrumented systems for the process industry sector. 2004.

[14] IEC 61513:2011: Nuclear power plants - Instrumentation and control important to safety. 2011.

[15] IEC 62061:2015: Safety of machinery - Functional safety of electrical, electronic and programmable electronic control systems, 2015.

[16] M. Kwiatkowski, Methods for the application of programmable logic devices in electronic protection systems for high energy particle accelerators, 2013.

[17] C. Sibley, Machine protection strategies for high power accelerators, Proc. 2003 Bipolar/BiCMOS Circuits Technol. Meet. (IEEE Cat No03CH37440) 2003, https://doi.org/10.1109/PAC.2003.1288989..

[18] R. Andersson, University of Oslo, Oslo, 2017, https://www.duo.uio.no/handle/10852/590699.

[19] S. Kövecses, R. Andersson, A. Nordt, E. Bargalló and M. Rejzek, Applying the functional system interaction process at ESS, ICALEPS, 2017.

[20] N. Leveson, An STPA primer, Version 1, 2013.

[21] C. Hilbes and A. Nordt, Machine protection - systems requirements and architectural framework, ESS Internal Document (ESS-0057251), 2015.

[22] R. Andersson, E. Bargalló, C. Hilbes and A. Nordt, Machine protection risk management process, ESS Internal Document (ESS-0095000), 2017.