Practical Examples of a New Approach to Creating Clear Operational Safety Cases

Jane Fenn, Richard Hawkins, Mark Nicholson

University of York, York, UK

Abstract *The concept of an 'Operational Claim Point', (OCP), has recently been proposed as a mechanism for improving the structuring and clarity of Operational Safety Cases. OCPs provide a mechanism by which arguments and evidence in the operational domain can be explicitly connected to design-time risk arguments. This gives rise to a number of benefits: ensuring that system operators are able to focus on just the operational aspects of the safety case relevant to them (hiding irrelevant and potentially confusing design details); making sure that, at the same time, the crucial relationship between the operational safety case and the design-time risk argument is explicitly documented and maintained (helping operators to better understand the safety impact of their work); and allowing design-time safety engineers to specify, in the risk argument, safety claims relating to system operation. We provide worked examples of how OCPs can be used in practice. Through these examples we explore some of the challenges in creating operational safety cases, including the link to the operational Safety Management System. We consider the impact of evidence that becomes apparent during operation, indicating unacceptable risk levels, and argument and evidence that may change depending on the specific choices of different operators of the same system.*

1 Introduction

In our paper, (Fenn et al. 2024) we proposed the concept of the Operational Claim Point to prompt debate and discussion around a clearer way to present Operational Safety Cases. The interest generated suggested that the approach was deserving of additional maturation and so this paper presents further examples of usage and begins to answer some of the outstanding questions raised in (Fenn et al. 2024).

1.1 Background of Operational Claim Points

In (Fenn et al. 2024), we explain that the origins of the use of safety cases for engineered systems were intended for discussion of safety of operations. Most safety case guidance, however, is written for those developing systems, assuming the safety case supports the transition from the design phase to an 'operational' or 'in-service' phase. Consequently, it is often unclear in guidance and standards whether the 'operational' safety case is an evolution of the design time safety case or a separate and distinctly differently-focussed case. Our review of practice suggests that it is often the latter.

Current practice in the operational phase is often to revert to a much simpler representation of the risks associated with hazards of a system, typically Bow Ties, (Acfield and Weaver 2012), which bring a less flexible approach to risk mitigation.

We propose that the operational safety case, though distinct, should be much more explicitly linked to the design safety case, in a defined relationship, which:

- Ensures that system operators are able to focus on just the operational aspects of the safety case that are relevant to them (hiding irrelevant and potentially confusing design details).
- Ensures that, at the same time, the crucial relationship between the operational safety case and the design-time risk argument is explicitly documented and maintained (helping operators to better understand the safety impact of their work).
- Allows design-time safety engineers to specify, in the risk argument, safety claims relating to system operation,

1.2 Proposed Approach

The proposed approach is based on earlier concepts around good practice in structuring a safety case, as defined in (ACWG 2021). Cases are structured according to three separate but related arguments, namely:

- A risk argument that records the arguments and evidence used to establish direct claims on the acceptability of safety risk
- A confidence argument that justifies the sufficiency of confidence in the safety risk argument
- A conformance argument that justifies belief in conformance with the requirements of a standard or regulation

The approach also uses the concept of Assurance Claim Points, (ACPs), defined in (Hawkins et al. 2011), where ACPs are used to indicate points in the risk argument where arguments of confidence are required. Comparably, Operational Claim Point, (OCPs), indicate points in the design time risk argument where operational aspects need to be considered. Examples of such points might include: environmental conditions in which a system is expected to operate; assumptions about how the

system will be routinely operated; mitigating procedures that the designer wished to put in place, etc.

OCPs are then addressed by arguments with supporting evidence that exists, or will exist, when the system is in the operational domain. The operational safety case will provide this argument and evidence for each OCP. In this way the operational safety case is separate from, but directly traceable to the risk argument created during system development. For each OCP, the argument and evidence in the operational safety case must relate specifically to the operational aspects associated with the point in the risk argument to which the OCP relates. In this way we can ensure that operational arguments specifically address aspects of risk mitigation in the design safety case.

In (Fenn et al. 2024), we discussed the most likely uses of the OCP would be around asserted context or evidence in the risk argument. Asserted context (which may include assumptions) in the design safety case are declarations within which the system is argued to be safe, and can include, for example, operating limits of the system such as temperature or pressure. An OCP associated with this asserted context would need to demonstrate, as part of the operational safety argument, how adherence to the system limitations would be achieved through-life. An OCP on evidence in the design safety case indicates that the validity of the evidence item may be affected during the operation of the system, (e.g. by a change in personnel during operation affecting evidence of operator competency, or evidence of system servicing history being updated) so the operational safety case must argue the ongoing sufficiency of that evidence as it changes throughout the operational phase. Below we consider potential uses of OCPs and worked examples of the type of argument that might be created to satisfy these OCPs.

1.3 Notation

The concept of OCPs can be used to link operational arguments and evidence to the risk argument however a safety case is presented. In (Fenn et al. 2024), we proposed how OCPs can be represented in safety arguments when using Goal Structuring Notation (GSN). This is shown in Fig 1 and can be seen to mirror the existing notation used for Assurance Claim Points, (ACPs). A key to GSN symbology is provided for those who are unfamiliar, in Fig 2. A key to the GSN symbols used in this paper

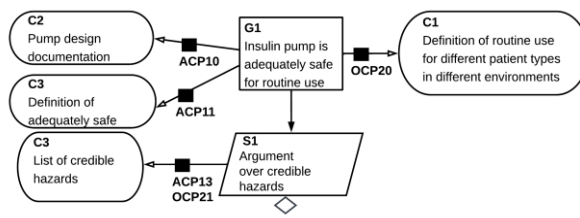

Fig 1. Example of Operational Claim Points and Assurance Claim Points

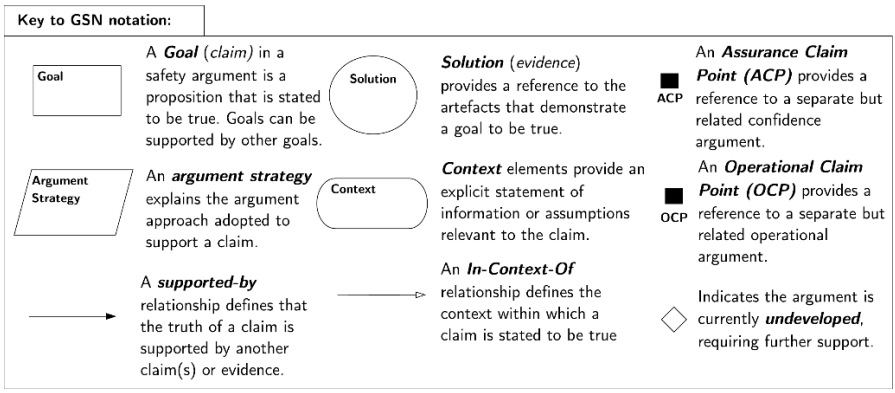

Fig 2. A key to the GSN symbols used in this paper

2 Case Study

The CERN Large Hadron Collider (HLC) is the world's largest and most powerful particle accelerator. It consists of a 27-kilometre ring of superconducting magnets with a number of accelerating structures to boost the energy of the particles along the way. Due to the nature of the facility, many of the materials on the design of the LHC are made available publicly. This has allowed researchers to generate, and publish, a safety case for the operation of the Machine Protection System, (MPS), of the HLC. This has been validated with CERN, and made available for academic research purposes, see (Rees et al. 2023). This has provided the opportunity to test the OCP approach with a larger case study than used previously and one which is authored independently of this paper's authors.

We have focussed on a number of key areas of the LHC MPS[1] where there are clearly discussions about how the system will be used in operation or where there are necessary operational mitigations identified in the argument. The LHC safety case is presented in (Rees et al. 2023) using Eliminative Argumentation, (EA), notation. Where necessary, we have re-factored the argument slightly, using GSN, to illustrate clearly how the argument might be presented when separating out design and operational aspects using OCPs.

[1] From (Rees et al. 2023): The Machine Protection System (MPS) is comprised of inter-dependent components designed to ensure that the LHC does not become damaged during operation. It proactively protects the system by monitoring all conditions that could lead to damage, and issuing a beam dump (i.e., extracting all particles from the LHC rings) before hazardous scenarios occur.

2.1 Magnets

An important part of the safety case for the LHC revolves around the magnets which control the beam. A claim is made: "C0079: When active, the MSD magnets[2] correctly divert the beam vertically towards the extraction lines and ultimately the TDE dumping block[3]". The safety case makes arguments about the design and maintenance of the magnets, including their power supply, which is monitored via an Energy Tracking System, (ETS). In Fig 3 we refactor the argument, using GSN, to separate out the design and maintenance activities:

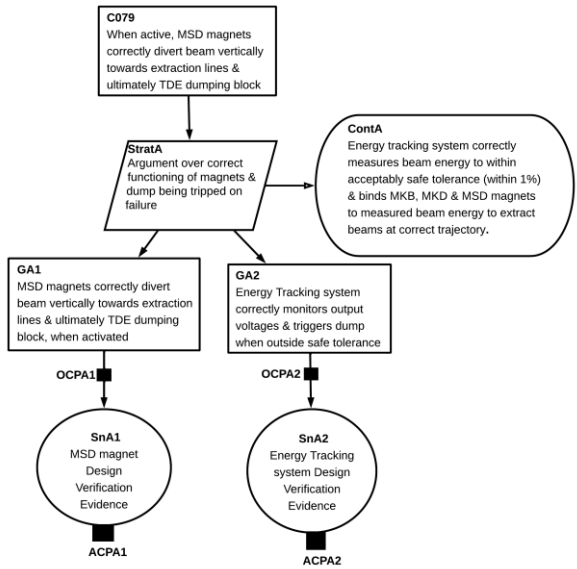

Fig 3. Fragment of argument related to magnets

An ACP is shown related to each of the design mitigation evidence items (SnA1 and SnA2). These ACPs provide a confidence argument regarding the rigour of the evidence, which is crucial to the validity of this argument. This evidence is created on the basis of assumptions in the design standard regarding the maintenance of the magnets and ETS. If the maintenance that occurs during operation does not satisfy this assumption (the maintenance interval is too long or maintenance activities are missed) then this could also challenge the validity of the argument. The requirement to provide an argument regarding the sufficiency of this maintenance for the ETS is captured in Fig 3 as an OCP (OCPA2). Fig 4. - OCP satisfaction for magnet and ETS maintenance shows the operational argument that relates to OCPA2.

[2] MSD magnets are a set of magnets intended to vertically divert an extracted LHC beam towards the target dump block.

[3] "The Target Dump External (TDE) block absorbs beams that pass through the MKB kicker magnets"

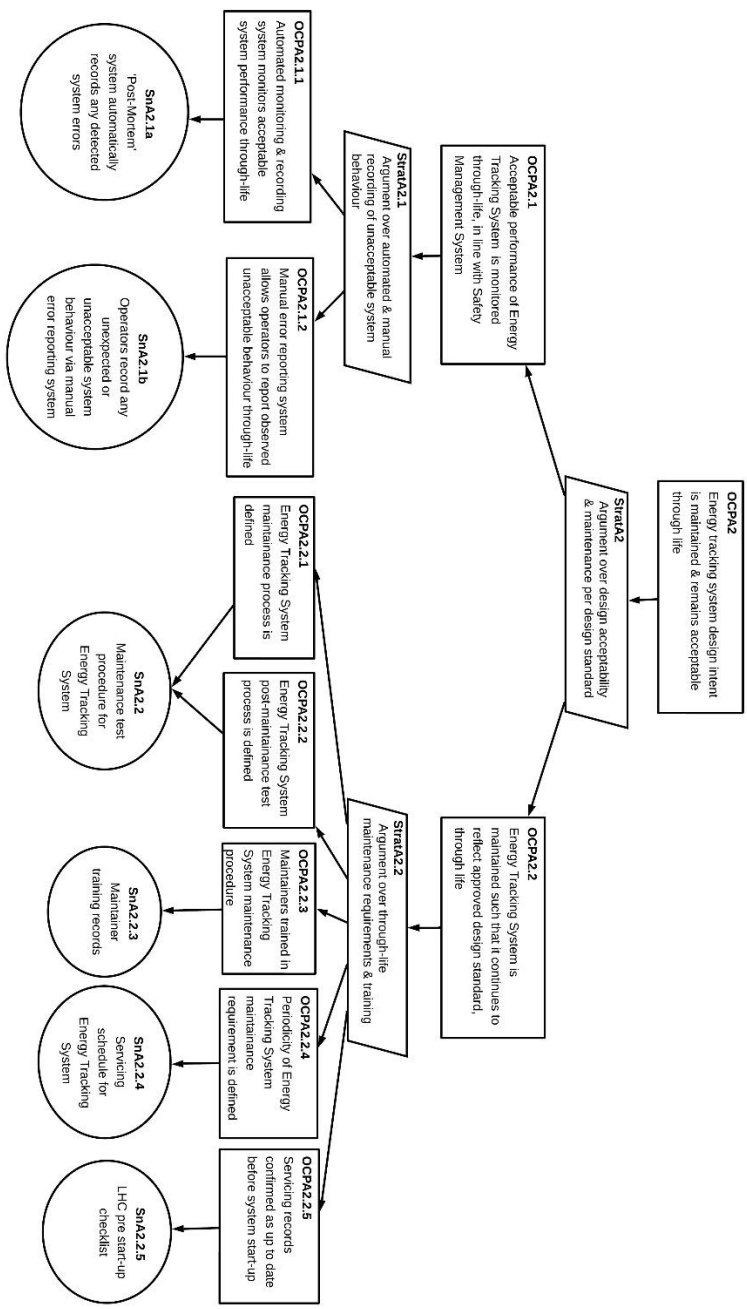

Fig 4. - OCP satisfaction for magnet and ETS maintenance

It was not possible to view the actual procedures used at CERN, however, per our previous paper, (Fenn et al. 2024), we speculate that maintenance procedures would typically exist, including post-maintenance test procedures which would be scheduled to occur at a set period. It is expected that maintainers would be appropriately trained and perhaps servicing records are confirmed before each experiment is undertaken using the LHC. The left-hand side of the argument in Fig 4, under Goal OCPA2.1 considers the role of the Operational Safety Management System in the operational argument. We take account of known automated error monitoring systems (SnA2.1a) and also the manual human error reporting system (SnA2.1b) that we assume would be used for most safety critical systems.

2.2 Target Dump External Block

Another interesting argument in the LHC safety case is around the 'Target Dump External Block'. This is a block that absorbs energy from the beam. The LHC must be operated such that the temperature of the block stays within safe limits. This is achieved by limiting how many and how often individual experiments are undertaken. There is evidence in the published safety case that this control mechanism may have changed over time from a purely manual control measure (where the timing between experiments is controlled manually by the operators using a minimum timing interval) to an automated control system function which will not allow a new experiment to commence until the block has cooled to a sufficient measured temperature.

To demonstrate the OCP concept, in this paper we have therefore considered the impact on operational safety, the Safety Management System and the Operational Safety Case of three change scenarios. We start with the safety control being a manual operator procedure. We then speculate about what may have happened in service to indicate that the operator control alone was insufficient. We discuss the addition of an automated system that might be developed in lieu of the operator control procedure. The third scenario speculates about the utility of using both automated control procedure and manual control procedures. The design safety argument fragment related to all three scenarios is represented in Fig 5. Options for TDE temperature management In doing so we do not imply that any of these situations arose at CERN; we use them simply to explore the utility of the OCP approach.

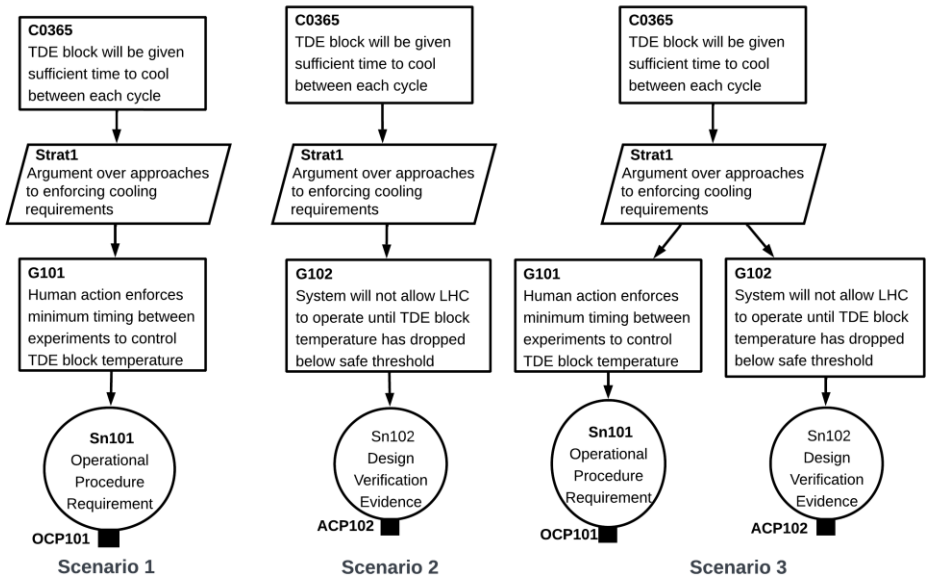

Fig 5. Options for TDE temperature management

2.2.1 Scenario 1 – Initial Design, Test and Commissioning

We imagine that, during development and testing / commissioning of the system, it was envisaged that system operators would have been heavily involved in developing the operating concepts for the system and understand the implication of the TDE block overheating. Consequently, they would be very cognisant and confident in enforcing the time delay between experiments as a mechanism of preventing overheating. In commissioning, perhaps there was not much pressure to undertake multiple experiments so the operator-only control mechanism was successful. We offer a potential confidence argument for OCP101, about the manual control in Fig 6. OCP101 - Manual Control

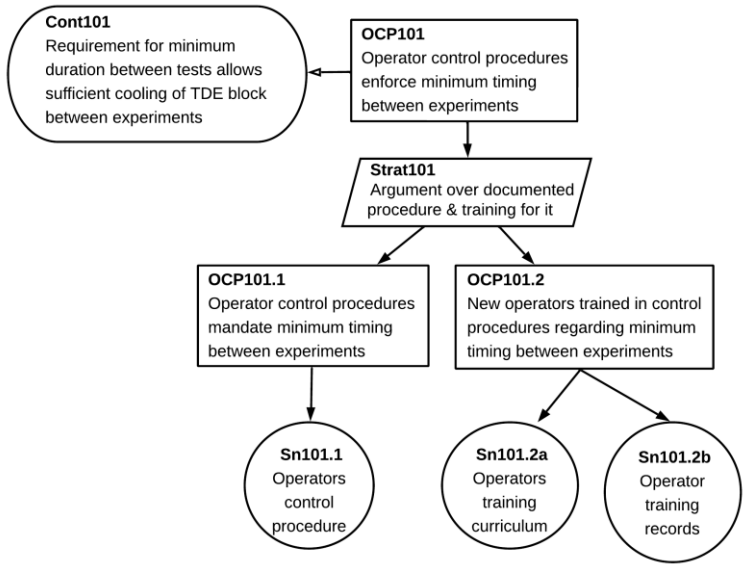

Fig 6. OCP101 - Manual Control

2.2.2 Scenario 2 – Moving into Operational Service

Based on feedback from engineers, we have noted that, in some industries, it is not possible to determine a single point in time where a system moves from a system in design phase to a system in operational phase. We acknowledge that situation but assume that, whenever a safety-critical system is operating, it should be monitored for acceptable safety performance. This is expected to be a core principle of the Safety Management System, (SMS), whether that SMS is still under the control of a developer, commissioning authority or under the authority of the operator's SMS. In that sense, detecting unacceptable safety performance should take place whichever SMS is enacted, and unacceptable safety performance should be rectified.

To illustrate, we speculate here about detected unacceptable safety behaviour with respect to the TDE temperature control and potential root causes. Increased occurrences of an experiment being started before the safe TDE temperature was reached, breaching the safety claim, might have been identified. In line with monitoring requirements in the SMS, this would have driven review activity which may have identified concerns about the adequacy of the manual approach to temperature control of the TDE block. Potential root causes of this non-compliance to safety procedures could be:

- The operators moved away from being those involved with system design and became people who were less knowledgeable about the consequences of breaching the minimal time period.

- Experienced operators became complacent because there hadn't been any problems to date and were less rigorous in enforcing the minimum time interval.

These root causes would challenge the argument and evidence in the OCP as to whether the training in the procedure was adequate and whether it needed to be 'refreshed' for each operator, or whether, even with increased training, a manual-only approach was sufficient. But there are also other potential root causes:

- It may have become apparent that the calculations around the cooling rate of the block were inaccurate and actually more time was needed to allow for cooling.

Changes in operating context might also drive concerns in this area too. For example:

- Commercial pressure to run more frequent experiments might result in the operator being challenged whether the cooling time between experiments could be shortened.

Any of these challenges appear credible and not un-typical of operational usage. We speculate that it was decided to use the measured temperature of the block in determining if it was safe to commence the next experiment, rather than infer how much it would have cooled using the proxy of time elapsed since the previous experiment. It may have been decided to automate that control, using measured temperature, either for safety reasons, or a safety change driven by commercial decisions, in the latter case.

In our scenario, this would have driven the need for a system change to measure the block temperature directly and limit operations only on when the block temperature was below safe limits. The system design would be subject to an Assurance Claim Point, though we choose not to expand on the ACP in this paper. Note that we have removed the human operator control action in this option, and hence the OCP would no longer be required.

2.2.3 Scenario 3 – Hybrid Approach

This option is introduced to illustrate how different operators may wish to operate the same system, (although it is accepted that there is only one LHC, at this point in time).

One operator may choose the first option, another the second, and another still may choose a hybrid combination of both manual and automatic control measures. The emergent behaviour in each operational situation would need to be considered thoroughly, including which option took priority when the system is re-started, i.e. if the automatic system says the TDE block is sufficiently cooled, but the minimum time period has not elapsed, is it acceptable to go ahead, and vice versa. It may be considered a safer option if it is necessary to wait until both options consider it safe to recommence experimenting, but this would need to be an additional training point for operators.

3 Conclusions and Further Work

We have presented an exploratory expansion of the concept of the OCP, introduced in our paper (Fenn et al. 2024) as a way of unambiguously linking operational safety arguments to the risk argument in the design safety case. In particular, in this paper, we have expanded on the link between the OCP and the Operational Safety Management System and considered the potential for options in terms of supporting an argument operationally. The nature of this support could depend on operational decisions as well as the choice of different operators. We show an example where operational monitoring, in service, could necessitate a design change to control unacceptable risk that was detected during operation. We anticipate further work around the role of the OCP with respect to the operational Safety Management System and more robust recording of the OCP interface.

References

Acfield A, Weaver R (2012). Integrating safety management through the bowtie concept a move away from the safety case focus. In: Proceedings of the Australian System Safety Conference-Volume 145. pp. 3–12

ACWG: Assurance Case Guidance. Tech. Rep. SCSC-159 v1, Safety Critical Systems Club (2021), https://scsc.uk/r159:1

ACWG: Goal Structuring Notation Community Standard. Tech. Rep. SCSC-141C v3.0, Safety Critical Systems Club (2021), https://scsc.uk/scsc-141C

Fenn J, Hawkins R, Nicholson M (2024). A New Approach to Creating Clear Operational Safety Arguments. In International Conference on Computer Safety, Reliability, and Security (pp. 227-238). Cham: Springer Nature Switzerland.

Hawkins, R, Kelly, Knight J, Graydon P (2011) A new approach to creating clear safety arguments. In: Advances in Systems Safety: Proceedings of the Nineteenth Safety-Critical Systems Symposium, Southampton, UK, 8-10th February 2011. pp. 3–23. Springer

Rees C, Viger T, Delgado M, Lippelt R et al. (2023). Assessing the Usefulness of Assurance Cases: an Experience with the CERN Large Hadron Collider (No. CERN-ACC-2023-0002).