Understanding the Use, Misuse and Abuse of Safety Integrity Levels[1]

Felix Redmill
Redmill Consultancy
Email: Felix.Redmill@ncl.ac.uk

Abstract

Modern standards on system safety employ the concept of safety integrity levels (SILs). Increasing numbers of system purchasers are expecting their suppliers to demonstrate that they use the concept, so system developers are seeking to apply it. But the standards differ in their derivation of SILs and none explains the concept satisfactorily, with the result that it is often misunderstood and used inconsistently, incorrectly, and inappropriately.

This paper explains the concept and its application, giving examples of how SILs are derived in three current safety standards. It then shows a number of ways in which the SIL concept is misinterpreted and used misleadingly. Further, it considers the relationship between SILs and risk-tolerability decisions.

1 Introduction

The concept of safety integrity levels (SILs) is now prevalent in the field of safety-critical systems, and a number of standards advocate its use in the design and development of such systems. However, not only do the various standards derive SILs differently, but none provides a clear and detailed explanation of how they are derived and applied. The result is that SILs are not well understood. Whereas the concept is intended to facilitate the achievement and demonstration of safety, it is in many cases causing confusion and dismay.

Further, although the derivation and application of SILs is complex and can be confusing, it is also conceptually simple and can be explained simply, with the result that the SIL concept is used inconsistently and often incorrectly and inappropriately.

One purpose of this paper is to explain the SIL concept. To do so, the paper offers not only a general description but also explanations of the ways in which SILs are derived and applied according to three recent standards.

A further purpose of the paper is to draw attention to the ways in which the SIL concept can be misleading and how it is being misunderstood and misused. The concept is a tool and, like any other tool, it can be useful if employed wisely but can cause problems if applied inappropriately.

---

[1] This is an extended version of a paper published in the Proceedings of the Eighth Safety-critical Systems Symposium, Southampton, UK. Springer, 8-10 February 2000.

2 What are Safety Integrity Levels?

The SIL concept has emerged from the considerable effort invested in the safety of systems during the last two decades. Two factors have stood out as principal influences.

The first is a move from the belief that a system can be either safe or unsafe, i.e. that safety is a binary attribute, to the acceptance that there is a continuum between absolute safety and certain catastrophe and that this continuum is a scale of risk. This has led to an emphasis on risk analysis as an essential feature in the development of safety-related systems.

The second influential factor is the huge increase in the use of software (and complex hardware, such as microprocessors) in the field of safety. This has led to a change in the balance between random and systematic faults. Previously, it was normal to assume (often implicitly) that safety could be achieved through reliability, and to deduce a value for the reliability of a system by aggregating, often through a fault tree, the random failure rates of its components. In some cases the failure rates were derived from historic use of the components and in others they were estimated, so the accuracy of the result was never beyond question. In fact, the greatest accuracy that could be achieved was that derivable from considering only random failures, for probabilistic methods are not valid for the analysis of systematic faults — those introduced, for example, through specification and design errors. With software, which does not wear out and in which all faults are systematic, there is no possibility of deducing system reliability by a method that is restricted to the consideration of random failures.

Another feature of software is its inherent complexity. Not only is it impossible to prove the absence of faults, but it would require an impracticably long time to derive high confidence in reliability from testing.

So a number of problems arise for the developer, who needs not only to achieve but also to demonstrate safety. Some of the problems may be summarised by the following questions and brief discussions.

- How do we define a system's safety requirements? These may result from a risk analysis that may be quantitative or qualitative. However, as software failures result from systematic and not random faults, direct measurement of the probability of failure, or the probability of a dangerous failure, is not feasible, so qualitative risk analysis must be employed. While the reduction of a given risk may be defined as the specification of a software safety function, the tolerable failure rate of that function may be defined in terms of a SIL. Depending on the standard in use, the SIL may or may not be equated to numerical ranges of failure rates.
- Given that greater rigour in the development of software is correlated with increased cost, how do we define the level of rigour that is appropriate to any particular case? Once risk analysis has led to a SIL, this is used to define the rigour of the development process. The higher the SIL, the greater the rigour, and tables are used in the standards to identify the methods, techniques, and management processes appropriate to the various SILs.
- If we can measure reliability directly, but not safety, can we define safety targets in terms of reliability measurements? In two of the standards discussed below, SILs are defined as rates of failure, and in one as the rate of dangerous (or unsafe) failures — all of which are reliability-type measurements.
- How do we define criteria against which to make claims of achieved safety? When a SIL has been used to define the level of safety to be achieved, it follows that that SIL should be the criterion against which a claim for the achieved safety would be made

(and judged). But if numerical values for the expected failure rate of software cannot be derived with confidence, it may not be possible to adduce proof of such a claim.

So, the use of SILs is an attempt to address the above questions. The derivation of a SIL may be summarised as the funnelling in of the risk assessment process to a result, the interpretation of that result into the SIL, and then the funnelling out into the development process which is defined by the SIL — as in the 'Bowtie Diagram' of Figure 1.

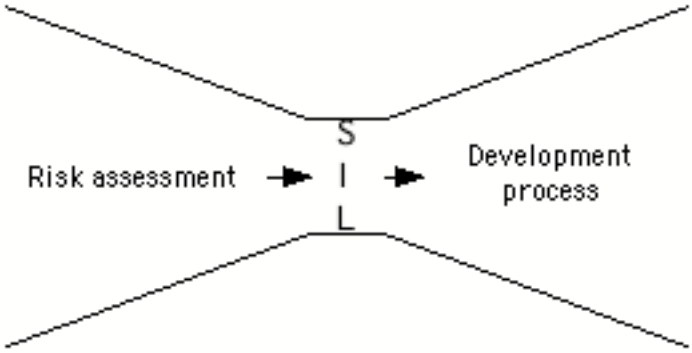

Figure 1: The 'Bowtie Diagram' showing the derivation and application of SILs

In essence, the SIL principle is this. If something is to do an important job, it needs to be reliable, and the more important the job, the more reliable it should be. Thus, there is an inverse relationship between the SIL and the tolerable rate of (dangerous) failures. In the case of a safety-related system, the job is to achieve safety, and the greater the importance to safety of the system whose SIL is under consideration, the lower the rate of unsafe failures should be. Then, the higher the SIL must be so as to indicate this requirement.

For those aspects of systems where random failures dominate, a numeric value of a SIL (as in IEC 61508) may be useful if it is possible to demonstrate quantitatively that a certain architecture or design will satisfy it. With systematic failures, it is currently unlikely to be possible to claim conformity to a numeric value, but the standards' authors believe that numeric values for SILs would be useful in the future if it later became possible to carry out more refined testing and measurement. Then we could have confidence in higher direct measures. Numeric values also set reference points to provide consistency of understanding of SILs across industry sectors.

There are different routes to the derivation of SILs depending on the standard in use, and three of these are examined in the next section.

3 SILs According to the Standards

3.1 SILs According to IEC 61508

The International Electrotechnical Commission's standard IEC 61508 [IEC 2000, Redmill 1998] is a generic international standard intended to provide guidance to all industry sectors. The model on which it is based (see Figure 2) assumes that we are starting out with some equipment, or plant — the 'equipment under control' (EUC) — which is to be used to provide some form of benefit or utility. Complementary to this is a control system, and together the EUC and its control system may pose risks.

The standard recommends that the hazards posed by the EUC and its control system be identified and analysed and that a risk assessment be carried out. Each risk is then tested against tolerability criteria to determine whether it should be reduced. If risks are reduced by redesign of the EUC, we are back to the starting point and hazard identification and analysis and risk assessment should again be carried out.

When it is decided that risk-reduction facilities should be provided in addition to the EUC and its control system, and that these should take the form of one or more electrical, electronic, or programmable electronic systems, then the terms of the standard apply to it or them.

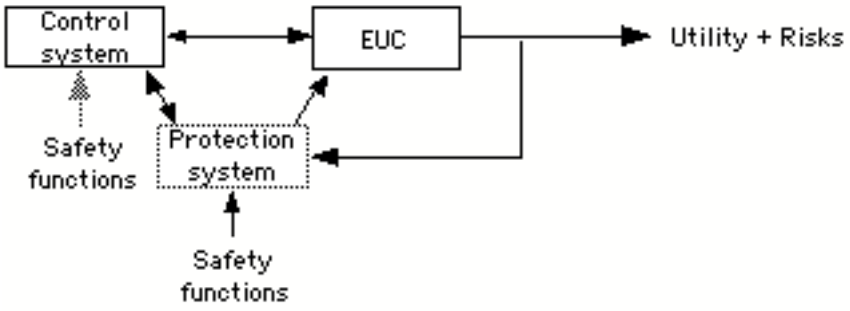

Figure 2: IEC 61508 system model

The risks posed by the EUC and its control system may be contributed to by many hazards, and each must be mitigated until its risk is considered tolerable. The reduction of the risk associated with each hazard is specified as a 'safety requirement' and, according to the standard, each safety requirement must have two components, the functional requirement and the safety integrity requirement. The latter takes the form of a SIL.

In Part 4 of IEC 61508, safety integrity is defined as 'the likelihood of a safety-related system satisfactorily performing the required safety functions under all the stated conditions, within a stated period of time', and a SIL as 'a discrete level (one of 4) for specifying the safety integrity requirements of safety functions'. Thus, a SIL is a target probability of dangerous failure of a defined safety function.

The totality of the safety requirements for all hazards forms the safety requirements specification. Safety requirements are satisfied by the provision of safety functions, and in design these are implemented in 'safety-related systems'. The SILs of the safety requirements become those of the safety functions that will provide them, and then of the safety-related systems on which the safety functions are to be implemented. The separation of safety-related systems from the EUC and its control system (as by the provision of a protection system — see Figure 2) is preferred. However, safety functions may also be incorporated into the control system and, when this is done, certain rules apply to ensure that higher-SIL functions are not affected by the failures of lower-SIL functions. Design is usually an iterative process in which the combination of safety functions in safety-related systems is decided on cost as well as on technical grounds.

The standard equates SILs with probabilities of unsafe failures in two tables, one for on-demand systems whose demand rate is low and one for systems with continuous operation or a high-demand rate (see Tables 1 and 2). Why the difference between them, and how do they relate to each other? The standard defines a low demand mode of

operation as 'no greater than one [demand] per year'. Since in approximate terms a year is taken to consist of 10^4 hours, the tolerable probabilities of failure in the low-demand cases are increased by a factor of 10^4 in order to arrive at the equivalent SIL values for continuous systems. Assuming a failure rate of once per year, the SIL 4 requirement for the low-demand mode of operation is no more than one failure in ten thousand years.

If there is to be no more than one demand per year made on a protection system, the EUC and its control system must have a dangerous failure rate of no more than once per year - or 10^{-4}. But arriving at this conclusion can be problematic because doing so is at the very limit of practical testability [Littlewood and Strigine 1993]. It is therefore not obvious how the assumption of low demand will be justified, other than in cases where the EUC and its control system are replicated many times and dangerous failure rates are derived from the combined use in their operational environments.

The failure rates attached to SILs for continuous operation are even more demanding (by a factor of 10^4) and are intended to provide targets for developers. Because a system cannot be shown to have met them - certainly not a software-based system - they are intended to define the rigour to be used in the development processes. SIL 1 demands basic sound engineering practices, such as adherence to a standard quality system, repeatable and systematically documented development processes, thorough verification and validation, documentation of all decisions, activities and results, and independent assessment. Higher SILs, in turn, demand this foundation plus further rigour.

Table 1: Safety integrity levels of low demand operation (from IEC 61508)

Safety Integrity Level	**Low Demand Mode of Operation (Pr. of failure to perform its safety functions on demand)**
4	>= 10^{-5} to 10^{-4}
3	>= 10^{-4} to 10^{-3}
2	>= 10^{-3} to 10^{-2}
1	>= 10^{-2} to 10^{-1}

Table 2: Safety integrity levels for continuous operation (from IEC 61508)

Safety Integrity Level	**Continuous/High-demand Mode of Operation (Pr. of dangerous failure per hour)**
4	>= 10^{-9} to 10^{-8}
3	>= 10^{-8} to 10^{-7}
2	>= 10^{-7} to 10^{-6}
1	>= 10^{-6} to 10^{-5}

Thus, the SIL of a safety-related system reflects the risk reduction that the system must achieve. For example, if the tolerable risk is deemed to be 10^{-9} dangerous failures per hour, and the EUC is calculated to have a probability of 10^{-2} dangerous failures per hour, the difference must be achieved by one or more safety functions. If the risk reduction were provided by a protection system separated from the EUC and its control system (as preferred by the standard - see Figure 2), the protection system would need to have a probability of 10^{-7} dangerous failures per hour (see the simple fault tree of Figure 3). From this, and from Table 1, we can deduce that the safety-related system would need to be of SIL 2. Of course, for this simple subtraction of indices to be a valid means of calculation,

there needs to be a very strong argument that the dangerous failures of the EUC and those of the protection system are independent.

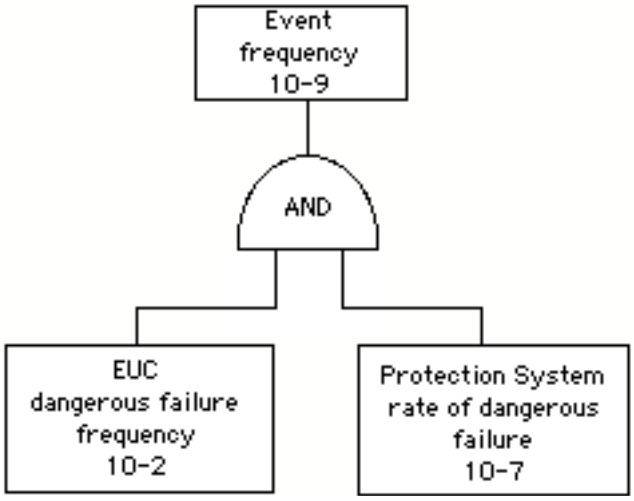

Figure 3: The principle of a protection system

3.2 SILs According to the MISRA Guideline

Whereas IEC 61508 is a generic standard, intended as the basis for preparing more detailed sector-specific standards, the Motor Industry Software Reliability Association's 'Development Guidelines for Vehicle Based Software' [MISRA 1994] is sector-specific. The document was tailored to the use of developers of software-based systems to be employed in motor vehicles.

In this guideline, SILs are based on the consequence of failure of the system in question. For motor vehicles, the ultimate consequence of a system failure (in terms of accidents and their possible outcomes) is speculative, so in the guideline the consequence of failure is defined in terms of something more predictable — the controllability of the vehicle by its occupants. So the guideline advocates that system developers carry out a hazard identification and analysis and determine the worst possible result of the failure of their system — in terms of the controllability of the vehicle. Five levels of uncontrollability are defined (see Table 3) and SIL values are defined to accord with them. (The guidelines provide definitions of the controllability categories and readers are referred to the document itself if further explanation is required.)

Table 3: Consequence-based SILs

Controllability Category	Integrity Level
Uncontrollable	4
Difficult to control	3
Debilitating	2
Distracting	1
Nuisance only	0

The controllability category is directly related not only to an integrity level but also to an 'acceptable failure rate', as in Table 4. But the authors of the guideline, recognising that low failure rates of software cannot be measured with confidence, did not place numeric

values on them. The qualitative terms used can be helpful, for example, by reminding a developer of a system whose failure could render the vehicle uncontrollable to know that failure of the system should be 'extremely improbable'.

Table 4: SIL relationships

Controllability Category	Acceptable Failure Rate	Integrity Level
Uncontrollable	Extremely improbable	4
Difficult to control	Very remote	3
Debilitating	Remote	2
Distracting	Unlikely	1
Nuisance only	Reasonably possible	0

Once a SIL has been derived, it is applied as in IEC 61508 and according to the Bowtie Diagram of Figure 1: to determine the rigour of the system-development processes.

In this guideline, SILs are referred to merely as 'integrity levels'. They are related to the reliability of the system to which they are applied, according to the danger attached to the failure of the system. In the context in which the guideline applies, no attempt is made to determine the relative probabilities of benign and dangerous failures of a given system.

3.3 SILs According to Defence Standard 00-56

'Def Stan 00-56' [MoD 1996] is a UK defence standard for system safety management. It states that 'it is widely accepted that the estimation of the probability of random events can be predicted to a reasonable degree of accuracy'. However, recognising the difficulty of estimating systematic failure integrity, it defines the SIL concept as 'an indicator of the required level of protection against systematic failure'. It allocates a SIL to 'each abstract function' at the 'early design phases' and calls for this to be inherited by the components that implement the function.

The standard recommends a risk analysis process which, for a given risk, places the consequence in one of four categories and the probability of occurrence in one of six. It combines these two sets of criteria in a matrix which it refers to as a 'risk classification scheme' and populates this with four tolerability classes (see Table 5). Then the standard uses both consequence and probability of failure in determining SILs and defining what they should achieve. The SIL is determined according to the consequence to which the hazard could give rise, and the requirement of the SIL is defined in terms of the worst probability of failure of the function involved.

Table 5: An example risk classification scheme

	Catastrophic	Critical	Marginal	Negligible
Frequent	A	A	A	B
Probable	A	A	B	C
Occasional	A	B	C	C
Remote	B	C	C	D
Improbable	C	C	D	D
Incredible	C	D	D	D

But the standard adds a complication, that of distinguishing between the first or only, and

subsequent functions on which safety depends (in the context of the risk in question). The SIL of the first or only function is based on the estimated accident severity, as defined in Table 6; that of the second function and any subsequent functions is based on the accident severity plus the failure probability of the first function (see Table 7).

Table 6: SIL for the only or first function

Accident severity			
Catastrophic	**Critical**	**Marginal**	**Negligible**
SIL 4		SIL 3	SIL 2

Table 7: SIL for the second and subsequent functions

Failure probability of first function	**Accident severity**			
	Catastrophic	**Critical**	**Marginal**	**Negligible**
Frequent	SIL 4			
Probable		SIL 3		
Occasional				
Remote		SIL 2		
Improbable			SIL 1	

The decision of whether to have a second function is not merely a design preference, for the standard adds a constraint that would in many cases prescribe the need for one. It defines 'claim limits' (see Table 8) which limit the claim that can be made for the probability of failure of a function (however reliable it may be thought to be), depending on the SIL — by implication, on the possible consequence of failure. Thus, if it were considered necessary to reduce a 'catastrophic' risk (see Table 5) from 'probable' to 'improbable', a second function would be essential as a first function of SIL 4 could not be claimed to have reduced the probability any lower than 'remote'.

Table 8: Claim limits

Safety Integrity Level	**Minimum failure rate that can be claimed**
SIL 4	Remote
SIL 3	Occasional
SIL 2	Probable
SIL 1	Frequent

The standard also allows for a higher SIL to be achieved by the combination of components of lower SILs. For example, a SIL 4 function may be provided by two independent SIL 3 components with a SIL 4 'combinator'.

Whereas in IEC 61508 SILs are based on risk reduction, and in the MISRA Guideline on consequence severity, in Def Stan 00-56 they are based on both. As in the other standards, SILs are used to define development processes. However, it is noticeable that here only 'design rules and techniques' are mentioned as being subject to SIL control rather than all safety management processes.

4 Some SIL Problems

4.1 Confusion between Standards

As seen in Section 3, standards that use the SIL concept apply different interpretations to it and derive SILs in different ways. Unless one states the standard that forms the context of a reference to SILs, misunderstanding can arise. Further, if the recipient of the information is unfamiliar with the standard in question, confusion is almost guaranteed.

4.2 Claim of Achievement against a SIL

In the first instance, SILs define what we expect of our safety functions and safety-related systems and are therefore targets. What confidence can we have that the systems that perform the defined safety functions really do satisfy the SIL requirements?

For simple systems with known fault histories in the application under consideration, a claim to have met a SIL may be deemed justifiable. Similarly, for systems composed of simple hardware components with known fault histories, in simple architectures, it may be credible to deduce worst-case failure rates by probabilistic means. But when a system is based on software or more complex hardware (e.g. microchips), so that systematic rather than random faults predominate, and testing cannot in practical time offer reliable predictions of the rates of dangerous failure, a claim to have met a SIL cannot, in the present state of the art, be supported by measurement.

The value of the SIL is in providing a target failure rate for the safety function or safety-related system. It places constraints on the processes used in system development, such that the higher the SIL, the greater the rigour which must be applied. The processes defined as being appropriate to the various SILs are the result of value judgements regarding what needs to be done in support of a reasonable claim to have met a particular SIL.

However, the development processes used, however good, appropriate, and carefully adhered to, do not necessarily lead to the achievement of the defined SIL. And, even if in a particular case they did, the achievement could not be proved. So a SIL could not normally be said to define the actual rate of dangerous failures of the product.

Hamilton and Rees [Hamilton 1999] warn that relating SILs to process requirements can lure the practitioner into the following false safety argument: 'The requirement was for a SIL X system and I have adhered to the standard's process for a SIL X system, therefore I have developed a SIL X system.' They point out that the confusion stems from failing to recognise the twin goals of a safety engineering activity: to engineer a safe system and, while doing so, to build up evidence that the system is as safe as it is required to be. So, to be valid, the argument needs to be: 'The requirement was for a SIL X system, and good practice decreed that I adhered to the standard's processes for a SIL X system. In doing so, I have generated the evidence appropriate to a SIL X system, and assessment of the evidence has found that I have adhered to the defined processes.' Unfortunately, in practice the evidence is usually insufficient to show that the SIL requirement has been met, but it does increase confidence in the system and its software.

4.3 S is for Safety

The 'S' in 'SIL' refers to 'safety', so it is misguided to use the acronym 'SIL' outside the

context of safety.

Yet, there is a move, in some industry sectors at least, to use it in all contexts (e.g., 'This is a SIL 3 pump'). This is misleading, and there is already confusion in the application of the term. An added problem, or at least a factor which compounds the problem, is that those who are misled are, in the main, not aware of the error.

Of course, the standards vary in their application of the SIL concept. In the MISRA Guideline, SIL does in fact refer to the overall reliability of a system — but the system has previously been identified as having an impact on safety, and the guideline is industry-specific and defines its use of the SIL concept clearly for its users.

In usage according to IEC 61508, it is not sufficient to relate the SIL to the failure rate of a safety-related system; it must be related to the dangerous failure rate. For this, we must distinguish between dangerous and non-dangerous failures. If SILs are used as indicators of reliability (the probability of a failure) rather than of the probability of a dangerous failure, systems will cost more to develop than they need to.

But how can we distinguish between dangerous and non-dangerous failures? Only by carrying out a safety analysis, identifying all the system's failure modes, and determining which are dangerous and which are not. Particularly for a control system, this is crucial. But too often the distinction is not observed.

The SIL concept appears to offer simple rules for the reliability requirements of safety-related systems. But the rules are not simple, and their apparent simplicity, combined with the perceived importance of SILs, seems to be encouraging some practitioners to neglect thorough safety analysis in favour of deducing SILs. But the proper deduction of SILs can only be based on thorough safety analysis.

4.4 Claims Based on Reliability Estimates rather than Process

SIL values are used to define development processes, and it is almost always impossible to prove that a SIL has been met. Thus, one would not expect an unqualified statement such as 'The system is SIL 2' to be made. But if it were made, one would expect it to refer to the fact that the processes appropriate to SIL 2 were rigorously employed during the system's development - and that independent safety assessment has confirmed that they were. Yet, such statements are being used in cases where the system in question has merely been calculated or estimated to have a particular rate of dangerous failures.

If SIL claims based on optimistic reliability estimates are taken at face value, they can be misleading and could lead to inappropriate equipment being used in safety-related applications. We need to be questionning when SIL claims are made.

4.5 Result of Loose Use of the Term 'SIL'

That the term 'SIL' is often used loosely has already been pointed out. So, can we be sure that we know what is meant when the term is used? Does it offer a guarantee that the system's probability of dangerous failure is appropriate to the SIL? Or does it say that the system's probability of dangerous failure is thought or assumed to be so? Or does it state that processes appropriate to the SIL have been used in the development of the system?

We need to be careful to enquire exactly what is meant when we are given SIL information about a system.

4.6 Hazards Introduced by a Safety-related System

A casual use of SILs often neglects the hazards posed by the safety-related system itself. In fact, most advice on SILs makes no mention of the possibility of such systems introducing new hazards and appears to carry the assumption that they do not do so. But let us consider the example of a fire control system intended to protect against fire by detecting heat or smoke and, perhaps, to act to control the fire by dumping a dousing or smothering agent onto it. Not only would there be a hazardous situation if the system did not detect the fire, but there would also be one if it incorrectly emitted an alarm and caused frenzied evacuation from premises such as a night club, or if it deposited its dousing or smothering agent on a room-full of people when there was no fire.

Analysts may fail to recognise the hazards posed by safety-related systems, and the standards are not helpful in this respect. For example, IEC 61508 does not offer explicit advice on how to address safety functions that are incorporated into control systems.

4.7 Reference to a Component

Safety is application-dependent, and to attach SIL values to products outside the context of the systems in which they will function can be misleading - and dangerous. For instance, it would be incorrect to speak of 'a SIL X component'.

Yet, now that the term 'SIL' is in currency, and given that its implication is for certain processes to be used in a system's development, it is not unnatural for suppliers to want to apply the term to their products regardless of how or where they will be used. It would be possible to make clear and specific statements about an item's development processes, or about its estimated or historic failure rate, or about the intention or purpose behind its design, without reference to SILs. But it is becoming more common for the SIL concept to be used (often unspecifically) in support of products.

To avoid misunderstandings, misrepresentations, and resulting unsafe systems, a convention is needed for the derivation and communication of confidence in safety-related products, both hardware and software. Issues to be covered should include:

- Indications of the rigour that has been applied throughout development;
- The nature and the independence of the assessment to which the development processes were subjected;
- The testing and test results which provided confidence in the developed product;
- The history of use, if any, which confirmed that confidence.

In other words, we need a mini-safety case for components for which a SIL is claimed.

It would be stretching the use of the SIL concept to suggest that a product would be suitable for all SIL X applications because one, or even all, of these criteria met SIL X requirements. Further assessment of any new system of which the component became a part would always be necessary, but an accepted convention would make this more feasible than at present.

4.8 Beware of reuse

Beware of thinking that if you have achieved a system of a given SIL for one application it will be effective in another application that calls for a system of the same SIL. The safety of

an item is application-specific and a single, seemingly trivial or even unrecognised, variation in the design or use between the new and old applications can have considerable safety implications. Reuse is dangerous. Yet, there is a move, even in the context of safety, to more extensive use of commercial off-the-shelf (COTS) systems and components, so there is an increasing need for care (and guidance) in this matter.

4.9 SILs Say Nothing about System Attributes

Reuse (of software or any component) in a new system of the same SIL as the one in which it was previously used is equivalent to applying SILs to components without reference to their safety application. There is a strong case against both. In reuse the critical characteristics required of a system are different between applications, and a SIL says nothing about which attributes of a system are of concern. It may not be obvious that those that were appropriate to its former application are not significant in its later context and those necessary in the latter were not emphasised in its development for the former.

However appropriate the development processes are to a particular SIL, they do not guarantee that the product is bestowed with the attributes necessary for its application. Only design and development in accordance with a good specification, with attention to the objectives of use, can approach this.

For example, in the case of an emergency communications link it may be availability rather than reliability that is the critical characteristic because satisfactory communication may be possible in spite of intermittent failures, as long as those failures are short-lived. Yet, the SIL concept is reliability-based and is intended to provide confidence in reliability rather than availability. In another case it may be crucial not only to have a highly reliable system but also one that can quickly and easily be reconfigured if it did fail. Consider, for example, a recently reported air traffic control (ATC) event. During routine testing, an ATC system crashed because of a momentary loss of power. But the system was incapable of rapid recovery, and return to service required rebooting and re-calibration that took about an hour, during which time all computer-based systems, including radar and back-up systems, were inoperative. Only an antiquated radio system allowed communication between AT controllers and pilots.

SILs do not provide clear indicators of the quality of any given attributes of the system, so it is not enough to develop a system to a given SIL, using appropriate processes. It is essential to identify the critical attributes and ensure that they are of the required integrity.

4.10 Combining SILs

Recently I was told that 'two SIL 2s make a SIL 3'. Combining SILs is catered for in Defence Standard 00-56, but the fact that the statement was made in the context of IEC 61508 demonstrates the confusion that is being caused by different standards using the same term or concept in different ways. In standards other than Defence Standard 00-56 there is no acceptance of creating a system of one SIL by means of a number of components of lower SILs. To make such a statement in a general sense is dangerous.

Another point worth noting is that Def Stan 00-56 does not allow indiscriminate combinations but lays down strict rules, for example concerning independence, for the conditions under which combination is acceptable. To ignore the rules is dangerous, as is the generalisation of combination formulae, even within the context of Def Stan 00-56.

4.11 The Need for Assessment Criteria

It may not be possible to measure the rate of dangerous failures of a system, particularly one in which systematic faults predominate, but it is still worth defining how we should derive confidence that the system's performance approaches the demands placed on it by the SIL.

Lindsay and McDermid point out that a shortcoming of current approaches to SILs is that they offer little guidance on how to assess whether desired levels of safety integrity have actually been achieved [Lindsay 1997]. The emphasis is almost entirely on the process of development. So even if the belief of those claiming the SIL is that it has been achieved, the claim should be considered doubtful until their evidence has been examined.

Lindsay and McDermid point out that 'knowing what process was followed in developing a system is not assurance enough on its own: evaluation criteria should be defined, and related back to integrity requirements allocation, to assess how thoroughly the process was followed and how thoroughly the product was checked.'

4.12 Failure Modes of Software

Conducting a development process in accordance with a SIL provides some confidence in the product. However, it has already been pointed out that it is not normally possible to conclude that the rate of dangerous failures designated by the SIL has been achieved. Moreover, an overall level of confidence does not offer any guidance on how the software is most likely to fail or which types of failure would be most dangerous. It is therefore important to carry out studies, perhaps using techniques such as hazard and operability studies (HAZOP) and failure mode and effect analysis (FMEA) on the software to increase the knowledge of its likely failure modes and their effects. This could in some cases lead to additional safeguards to increase confidence that the worst types of failures were protected against.

5 SILs and Risk-tolerability Decisions

In the standards, the SIL of a safety function or system is defined as part of its requirements. At the time of stating requirements, what is identified is the specifiers' opinion of a tolerable level of risk, and the SIL is defined with respect to it. For example, see Figure 4, which shows the IEC 61508 means of deriving SILs. Here, Ra represents the assessed value (quantitative or qualitative) of a risk and Rt represents the 'tolerable' level to which it should be reduced; the difference between them (assuming Ra to be greater than Rt) gives rise to the SIL.

However, what is thought to be a tolerable level of risk at the specification stage of a project is not necessarily what is (or would be) deemed to be tolerable later. Risk-tolerability is a touchy matter to decide on. As an example, let us consider the ALARP (as low as reasonably practicable) principle — a legal requirement in the UK [HSE 1992]. Under this, the risk must be reduced not merely to a tolerable level but to a level which is as low as reasonably practicable (represented by Ralarp in Figure 4).

What is reasonably practicable, Ralarp, cannot be discovered until the design stage when it is decided how to implement the necessary risk reduction. At that time, trades-off involving the risks involved, the technologies available, the possible design options, and

the various costs, are proposed and considered, and often early decisions are overturned by later ones as an iterative process proceeds. Further, the process needs to be carried out for each risk.

So, the risk level that is finally deemed tolerable under the ALARP principle may be different from that defined by the SIL — and it cannot be arrived at until after the SIL has been determined. Yet, system developers define essential safety-related system parameter targets (rate of dangerous failures in the case of IEC 61508, reliability in the case of the MISRA guideline, and so on) according to the SIL. In effect, the SIL currently defines the risk-tolerability decision - while not necessarily complying with the legal requirement for doing so.

There is a further point. ALARP requires that the level of accepted risk be constantly reviewed because if (for example because of changed or cheaper technology) it later becomes (or turns out to be) practicable to reduce the level of risk further, then that further reduction should be made. Should the SIL change to reflect this?

It seems that the use of SILs virtually demands that the choice of a SIL at the requirements specification stage defines the risk-tolerability decision which should, according to the ALARP principle, be taken much later. How can this anomaly be removed? Having a SIL can be useful, for it imposes constraints on the designers of safety functions. Is there a way in which the derivation of a SIL can be supported by risk-tolerability decision-making principles (for example, the ALARP principle)? If not, the SIL will be misleading to developers and could be costly to system owners as later changes would have to be made to comply with legal requirements. Should the SIL be defined as the difference between Ra and Ralarp (see Figure 4) rather than that between Ra and Rt? If so, how can this be facilitated? On the face of it, for this to happen, the SIL could not be defined until the design stage and it would not be available to place initial constraints on the design, as at present.

There needs to be research into the relationship between SILs and risk-tolerability decision making. In the UK, this needs to be based on the ALARP principle.

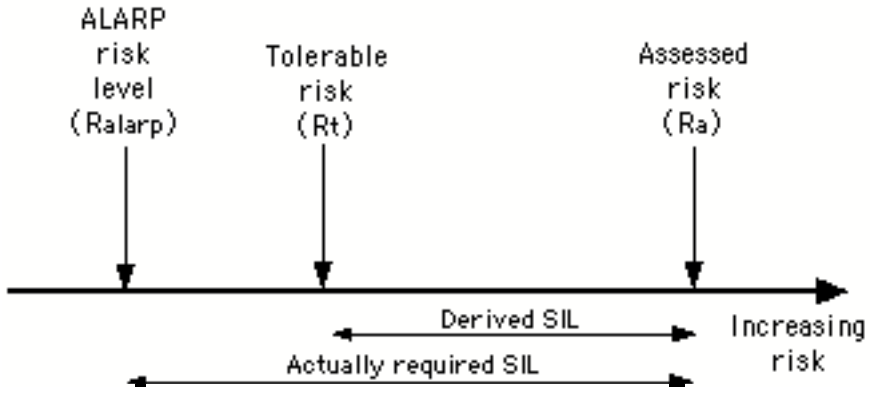

Figure 4: Relation between SIL and risk reduction as in IEC 61508

6 The SIL Concept is a Tool

The term 'SIL' can be misleading as well as helpful. The fact is that the SIL concept is just a tool. Every tool is designed and built for a certain purpose and to be used within certain

constraints, and the SIL concept is no exception. Using a tool outside its design constraints or out of context can lead to results that are incorrect or misleading. In the case of SILs, where safety is the central issue, the results could be dangerous.

No tool is indispensable, and here too the SIL concept is not an exception. Let us examine this. SIL has three main purposes: to define the safety integrity requirements for a safety-related system or function, to provide a guide to appropriate development processes, and to provide a basis for claiming achievement of safety requirements. In these respects it can be useful and convenient. But let us consider the necessity of the SIL concept to each of these.

With regard to defining safety integrity requirements, it may be simpler to say 'SIL 2' than 'A reliability of 10^{-6} dangerous failures per hour.' But it is quite possible to say, 'A reliability of 10^{-6} dangerous failures per hour'. Indeed, it is the required risk reduction (as in IEC 61508), the required reliability (as in the MISRA Guideline), or the accident consequence (as in Def Stan 00-56) which is first deduced, and from it the SIL is determined. So, representing the required probability of dangerous failures by a SIL is convenient short-hand but not essential.

Regarding the provision of a guide to appropriate development processes, it is convenient to use a SIL as an intermediate point, as in the Bowtie Diagram (Figure 1). But, again, it is quite possible to use the reliability figures directly as a guide to the required process.

Similarly, with regard to the third purpose of SILs, it is just as easy to claim that the development processes 'appropriate to a reliability of 10^{-6} dangerous failures per hour' have been carried out as it is to claim that those 'appropriate to SIL 2' have been carried out.

We may summarise these conclusions by reference to Figure 5, which is an extension of the Bowtie Diagram. Depending on which standard we use, our risk analysis process results in a value R from which a SIL is derived. The SIL is then used to inform the development process. But if no SIL were derived, the development process would have to be informed directly, rather than indirectly, by R. Thus, the SIL concept is useful but not indispensable.

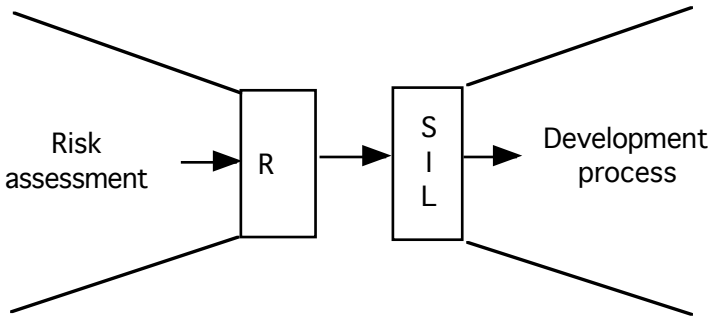

Figure 5: The SIL concept is not indispensable

What is said here is not a proposal to dispense with the SIL concept, but a reminder and a warning to employ it within its intended scope and with professional judgement. The term 'SIL' is now common currency and is unlikely to go away, so we need to understand it. To use it without understanding it is to use it dangerously. Likewise, to use it blindly,

simply as a means of determining suitable development processes, is to use it inadequately and riskily. We need to be extremely professional in our use of this tool.

6 Conclusions

This paper has offered an introduction to safety integrity levels, explaining them and showing how three existing standards derive and use them.

A number of examples were given of ways in which the SIL concept is confusing and misleading and the term 'SIL' is misused and misinterpreted. It was proposed that we need a convention for what information is made available when a SIL claim is made. Some inadequacies of the SIL concept were also discussed. It was then emphasised that the SIL concept is a tool and that, like all tools, it can be useful when used within its valid scope but problematic and dangerous when used outside it.

The SIL concept appears to offer simple rules for the development of safety-related systems. But to derive SILs correctly, we need to start from first principles and carry out thorough safety analyses. However, having done this, we find that we are also in possession of sufficient information to be able to carry out the development without SILs. So, although convenient, the SIL concept is not essential, and it can be replaced with the parameters that it represents. That is not to say that we should abandon its use. It is promoted by the standards, it is already employed, and it is now too late to ignore it. But we need to understand it and use it properly. Because it is not well understood, and there is already a tendency to use it incorrectly and inappropriately, it is all the more important that we attempt to avoid its misuse and recognise its misleading use by others. Indeed, it is as important to be aware of the dangers of misuse of the SIL concept as it is to understand its appropriate use, and this paper has attempted to highlight the dangers.

So far, the recorded use of the SIL concept is small, and it appears that a lack of training has contributed to the fact that many of the problems (for example, of misunderstanding) have not been recognised by those employing it. There is an urgent need for documentation and open reporting of the use of the SIL concept, the difficulties experienced, and the benefits gained.

But the problems in the use of SILs are not all of the users' making. There is a need for harmonisation of the SIL concept across standards, and for improved guidance in the standards themselves. There is also an urgent requirement for documented guidance for managers who need to understand the SIL concept so that they can effectively manage and make judgements about its use. This guidance should be supported by high-quality training, and the managers themselves need to accept their responsibilities in defining and controlling the introduction and use of a new, difficult and, as yet, unproven tool.

If managers, through failure to understand or reluctance to take time to learn, abdicate their responsibilities and leave decisions on the derivation and use of SILs to their subordinates, the current misunderstanding and misuse will continue - to the detriment of suppliers and customers alike. Improvement needs to be led, not only by the standards bodies but also by the management of organisations which seek to apply the SIL concept.

It is also proposed in the paper that there is a need for research into how the derivation and use of SILs interacts with other concepts of the tolerability and reduction of risk, such as the ALARP principle.

References

[Hamilton 1999] Hamilton V and Rees C: *Safety Integrity Levels: An Industrial Viewpoint*. In Redmill F and Anderson T (eds), Towards System Safety — Proceedings of the Seventh Safety-critical Systems Symposium 1999. Springer Verlag, London, 1999

[HSE 1992] Health and Safety Executive: *The Tolerability of Risk from Nuclear Power Stations*. Revised edition, HMSO

[IEC 2000] IEC 61508 — *Functional Safety of Electrical/Electronic/Programmable Electronic Safety Related Systems*. International Electrotechnical Commission, Geneva, 1999

[Lindsay 1997] Lindsay P and McDermid J: *A Systematic Approach to Software Safety Integrity Levels*. In Daniel P (ed), Proceedings of the 16th International Conference on Computer Safety, Reliability and Security, York, 7-10 September 1997. Springer Verlag, London, 1997

[Littlewood B and Strigine L 1993] Validation of Ultrahigh Dependability for Software-based Systems. *CACM* 36 (11) 69-80

[MISRA 1994] The Motor Industry Software Reliability Association: *Development Guidelines for Vehicle Based Software*. The Motor Industry Research Association, UK, 1994

[MoD 1996] Ministry of Defence: *Defence Standard 00-56, Safety Management Requirements for Defence Systems*. Ministry of Defence, UK, Issue 2, December 1996

[Redmill 1998] Redmill F: *IEC 61508: Principles and Use in the Management of Safety*. Computing & Control Engineering Journal, Vol 9, No. 5, IEE, London, October 1998