Tagungsband

Scientific Railway Signalling Symposium 2023

Einfach Fahren! – Digitale Transformation im Spannungsfeld Automatisierung, europäische Standardisierung und schneller Rollout

Herausgeber
Prof. Dr.-Ing. Andreas Oetting
Prof. Dr.-Ing. Birgit Milius

14.06.2023

TECHNISCHE UNIVERSITÄT DARMSTADT

TECHNISCHE UNIVERSITÄT BERLIN

Institut für Bahnsysteme und Bahntechnik

bbi Bahnbetrieb & Infrastruktur

Um die hohe wissenschaftliche Qualität und praktische Bedeutung der auf dem SRSS präsentierten Beiträge sicherzustellen, durchliefen alle Papers eine unabhängige Begutachtung durch folgenden Fachexperten:

- Dr.-Ing. Frederik Düpmeier
- Dr. rer. pol. Michael Leining
- Prof. Dr.-Ing. Birgit Milius
- Prof. Dr.-Ing. Andreas Oetting
- Prof. Dr. Björn Scheuermann
- Prof. Dr.-Ing. Jochen Trinckauf

Vorwort zur Begleitpublikation zum SRSS 2023 – Tagungsband SRSS 2023

von

Prof. Dr.-Ing. Andreas Oetting,

Leiter des Instituts für Bahnsysteme und Bahntechnik an der Technischen Universität Darmstadt

Prof. Dr.-Ing. Birgit Milius,

Leiterin des Fachgebiets Bahnbetrieb und Infrastruktur am Institut für Land- und Seeverkehr der Technischen Universität Berlin

Dr. Bernd Elsweiler,

Leiter Entwicklung Digitale Leit- und Sicherungstechnik bei DB Netz AG

Liebe Mitglieder der Fachcommunity und Bahn-Interessierte,

im Rahmen des Programms „Digitale Schiene Deutschland“ wird aktuell an der digitalen Transformation der Infrastruktur gearbeitet. Ziel ist es unter anderem, eine höhere Kapazität sowie optimale Auslastung des Schienennetzes zu erreichen, ohne im größeren Umfang zusätzliche Gleise zu verlegen. Für dieses Ziel bedarf es eines harmonisierten und reibungslosen Zusammenspiels zukunftsweisender Technologien im Bereich der Leit- und Sicherungstechnik, die zunächst umfassend erforscht und anschließend im Feld eingesetzt werden müssen. Mit dieser spannenden Thematik beschäftigte sich das Scientific Railway Signalling Symposium (SRSS) 2023 am 14. Juni, welches vom Institut für Bahnsysteme und Bahntechnik der TU Darmstadt zusammen mit der DB Netz AG sowie erstmals dem Fachgebiet Bahnbetrieb und Infrastruktur am Institut für Land- und Seeverkehr der TU Berlin organisiert wurde. Ca. 100 Fachexpert:innen aus Forschung und Praxis verfolgten die unterschiedlichen Vorträge unter dem Tagungstitel „Einfach Fahren! – Digitale Transformation im Spannungsfeld Automatisierung, europäische Standardisierung und schneller Rollout“ und brachten ihre Expertise in Diskussionsrunden ein.

Zu Beginn der Tagung stellte Volker Hentschel von der DB Netz AG in seiner Keynote das volle Potenzial des digitalen Bahnsystems vor. Dabei gab er einen Überblick über die Innovationen, deren Umsetzung in den kommenden Jahren im Rahmen der Digitalen Schiene zu erwarten sind. Daran anschließend diskutierten die Kleingruppen, welche Innovationen der digitalen LST schnell zu Verbesserungen im Feld führen können. Bevor Markus Montigel, RAILvelation GmbH, zum Abschluss des Tages durch einen interaktiven Vortrag das Motto des Tages „Einfach fahren“ noch einmal aus etwas Distanz reflektierte, boten zahlreiche Beiträge detaillierte Einblicke zu unterschiedlichen LST-bezogenen Themen aus Wissenschaft und Praxis. Aufgrund der gestiegenen Anzahl der Vorträge im Vergleich zu den Vorjahren wurde die Tagung im Nachmittagsblock erstmals in parallele Sessions aufgeteilt.

In dieser Publikation sind diejenigen Beiträge enthalten, die erfolgreich den Review-Prozess durchlaufen haben.

Eine wesentliche Innovation ist das Automatisierte Fahren, dessen Sicherheit vor Einführung geprüft und sichergestellt werden muss. Marco Kinas et al. beschreiben daher in ihrem Beitrag Aspekte des im Rahmen des ATO-RISK-Projekts verfolgten Ansatzes, so z.B. den Einsatz von Regelwerken als Risikoakzeptanzprinzip gemäß der CSM-Verordnung und die Analyse menschlicher Faktoren im Vergleich zum derzeitigen bemannten Betrieb. Neben der funktionalen Sicherheit spielt auch die

Sicherheit vor Angriffen eine wichtige Rolle. Simon Unger et al. erläutern daher in ihrem Beitrag das Zusammenspiel zwischen Safety und Security und arbeiten insbesondere die Relevanz von Security in Safety-kritischen Systemen anhand des RaSTA-Protokolls heraus. Im Gegensatz zu Safety handelt es sich bei Security um einen dynamischen Aspekt, denn die Bedrohungslage durch Angreifer kann sich stetig wandeln. Matthias Drodt von der DB Netz AG und Frank Weber vom Eisenbahn-Bundesamt beschreiben in ihrem Beitrag daher Lösungsansätze zur Zulassung der sehr dynamischen IT-Security in Eisenbahnsystemen.

Für eine schnelle Zulassung und einen schnellen Rollout ist auch eine Beschleunigung des Entwicklungs- und Sicherheitsnachweisprozesses notwendig. Frederic Reiter et al. beschreiben in ihrem Paper eine vollständig virtualisierte Evaluierungsumgebung für digitale Stellwerke, die die Komplexität reduziert und Tests für Ausführung von Funktionen vor der Integration mit physischen Komponenten ermöglicht. Gunnar Smith von Prover Technology aus Stockholm stellt in seinem Beitrag die Methoden vor, die zur Reduzierung des Aufwandes in der Entwicklungsphase der LST-Systeme und dadurch auch zum schnellen Rollout beitragen werden. Julian Lucas von der TU Darmstadt und Markus Rothkehl von der DB Netz beschreiben ein Konzept für ein modulares Nachweisverfahren für verschiedene Komponenten der digitalen Sicherungstechnik.

Auch im Falle eines beschleunigten Rollouts wird jedoch eine längere Übergangsphase zur digitalen LST erforderlich sein, in der es mehrere Migrationsschritte geben wird. In seinem Beitrag erläutert Melvin Zinngrebe von SBB AG zwei Strategien, die von der SBB AG zur Bewältigung der Migrationsherausforderungen angewendet werden. Eine davon ist die Lebensverlängerung der ersten Generation elektronischer Stellwerke, um Investitionen in eine Übergangstechnologie vor der Führerstandssignalisierung (FSS)-Migration zu vermeiden. Die andere Strategie ist eine Nachrüstung mit FSS.

Zuverlässige Infrastrukturdaten sind eine unerlässliche Voraussetzung für die erfolgreiche Durchführung der digitalen Transformation. Benedikt Wenzel et al. haben in ihrem Beitrag die Auswirkungen auf die Qualität von Infrastrukturdaten und geeignete Strategien zur Bewältigung der Herausforderungen, die mit der Erstellung und Pflege zuverlässiger Infrastrukturdaten verbunden sind, herausgearbeitet.

Um Pünktlichkeit und den reibungslosen Zugverkehr sicherstellen zu können, ist die Gewährleistung der Verfügbarkeit von LST-Systemen ein wichtiges Thema. In ihrem Paper beleuchten Simon Hofer und Martin Müller von team Technology Management GmbH verschiedene Strategien, um die Anforderungen an die hohe Verfügbarkeit beim Zugsicherungssystem ETCS gewährleisten zu können. Die Potenziale für Kapazitätssteigerungen und Leistungsverbesserungen, die durch Digitale Stellwerke, ETCS und andere Technologien erzielt werden können, sind in Deutschland nach wie vor enorm. Peter Reinhart von der DB Netz AG fasst in seinem Paper den bisherigen Erkenntnisstand hierzu anhand des Digitalen Knoten Stuttgart zusammen.

Die breite Themenvielfalt der in diesem Tagungsband enthaltenen Artikel zeigt, wie dynamisch das Umfeld in unserer Branche derzeit ist und wie viele spannende Innovationen wir in den nächsten Jahren erwarten dürfen.

Wir wünschen Ihnen daher viel Spaß beim Lesen der genannten Beiträge.

Prof. Dr.-Ing. Andreas Oetting

Prof. Dr.-Ing. Birgit Milius

Dr. Bernd Elsweiler

Risk Analysis for Automatic Train Operation

Jens Braband [1], Marco Kinas [1], Christian Klotz [2], Birgit Milius [3], Hendrik Schäbe [4]

[1] Siemens Mobility

[2] German Centre for Rail Research (DZSF)

[3] TU Berlin

[4] TÜV Rheinland

1 Introduction

The railway is an important backbone for the transportation of passengers as well as freight in Germany and in the EU. As it produces significantly less CO_2 than other means of transport, it also plays a key role in the reduction of emissions in the transportation sector. This means that the amount of goods and passengers transported by railways will increase in the future.

The enhanced modal share requires a competitive, reliable system with increased capacities. This will have to be achieved through the digital transformation of the rail sector. Automated Train Operation (ATO) is one of the key elements in this field and is currently under development in various stages and projects across Europe. It is expected to deliver an increase in efficiency, reliability, and capacity on the existing networks.

Solutions for full automation already exist for several metro lines. For mainline applications, only lower grades of automation (GoA) have been realized so far. In GoA 2 applications, the vehicle is self-driving most of the time while a driver is still present and responsible for safety. Higher grades of automation (GoA 3 and GoA 4) imply that there is no train driver permanently present. Hence, technical systems will take over this part. The perception system is the most discussed functionality in this field. This system will have to detect obstacles and threats in the driveway of the vehicle and assure that the track elements are in proper condition. It is expected that optical sensor technology or other sensors like radar must solve this problem in conjunction with advanced software such as neural networks or, more generally, methods of artificial intelligence. This technology has not yet been applied for this task in the railway sector and various questions are open concerning the system's safety.

In addition to obstacle detection, a variety of other functions must also be evaluated in GoA 3 and GoA 4 operation, including, for example, automatic detection and suppression of fire and smoke in the train. See (DZSF 2023) for a comprehensive list of functions.

In a workshop held by the German Centre for Rail Research (DZSF) in 2019, experts from the rail sector identified two important points which have to be solved in order to allow for highly automated mainline rail operation in the future. Firstly, functional requirements need to be found for the applied technology and secondly, risk acceptance criteria must be derived or formulated to enable the risk assessment for

[1] jens.braband@siemens.com|marco.kinas@siemens.com

[2] KlotzC@dzsf.bund.de

[3] birgit.milius@tu-berlin.de

[4] schaebe@de.tuv.com

the automation systems. Manufacturers as well as railway operators need this information to specify, design, validate, operate, and maintain the future ATO systems. Consequently, two projects (ATO-SENSE and ATO-RISK) have been initiated by the DZSF to address these issues. ATO-RISK addresses the point of risk acceptance criteria. This paper is dedicated exclusively to ATO-RISK.

For lower grades of automation, rules and requirements are already being defined, for instance in the new revision of the TSI CCS. For higher grades of automation, the risk assessment according to CSM Regulation (European Implementing Regulation Nr. 402/2013) gives a basis for the risk evaluation. Since there is no specification or prototype of concrete realizations of the systems yet, the considerations cover a generic level, identifying the system functions and the respective failure scenarios. The results are intended for the use by manufactures, who must design and validate their products, as well as assessment bodies and safety authorities, who need to verify the compliance with the safety requirements. These parties will have to break down the results further to a technology and application specific level.

The present paper has two main objectives. On the one hand, a general overview of the approach that was applied for the ATO-RISK project is provided. On the other hand, this paper explores two aspects of this approach in greater depth: First, the importance of codes of practice as an important risk acceptance principle according to CSM Regulation (European Implementing Regulation Nr. 402/2013) and second, human factor analyses to enable a comparison with the current operation with a driver.

Thereby, in Chapter 2 we discuss the risk acceptance principles according to CSM Regulation (European Implementing Regulation Nr. 402/2013) in general. The general approach of the ATO-RISK project is described in Chapter 3. As illustrated above, Chapters 4 and 5 are dedicated to two different approaches for deducing tolerable risk: code of practice and human factor analyses. The following Chapter 6 is dedicated to validation and in Chapter 7, we discuss the results.

2 Risk acceptance principles according to CSM Regulation

The European Union Implementing Regulation No. 402/2013 (European Implementing Regulation Nr. 402/2013) establishes a common safety method for the evaluation and assessment of risks. These common safety methods are intended to ensure a unified level of safety across all member states. When changes are made to the railway system, the changes must first be evaluated in terms of their significance. Significant changes in the sense of CSM Regulation (European Implementing Regulation Nr. 402/2013) are those changes which:

a) have an impact on safety and
b) are classified as significant based on an expert assessment by the proposer.

This expert assessment shall consider possible consequences of failures ("credible worst-case scenarios"), innovative elements in the introduction of the change, the complexity of the change, the possibility of monitoring the change over the entire life cycle of the system, a reversibility of the change, and the additive effect of several small changes that are individually not classified as significant (European Implementing Regulation Nr. 402/2013, p. 12).

Where significant changes are involved, the risk management process referred to in Article 5 of CSM Regulation (European Implementing Regulation Nr. 402/2013, p. 12) must be applied. Thereby, the CSM Regulation (European Implementing Regulation Nr. 402/2013, p. 18) specifies three different principles of risk acceptance, which are:

1. Application of codes of practice

2. Comparison with similar systems
3. Explicit risk estimation

In the first risk acceptance principle, several or all hazards are adequately covered by the application of relevant codes of practice. In this case, the hazard identification can be limited to the following steps:

1. Verification of the relevance of the code of practice
2. Determination of deviations from the code of practice

To be considered relevant, a code of practice must meet the following requirements:

a) It must be generally recognized in the railway sector. Otherwise, a justification must be provided to explain why the code of practice is applicable, nevertheless. This justification must be accepted by the assessment body.
b) There must be relevance of the code of practice to the control of the hazard in consideration. Successful application of the codes of practice to similar cases and effective control of the hazards is sufficient.
c) The codes of practice must be available to the assessment body for evaluation on request.

In case hazards are covered by relevant regulations, the risks associated with these hazards are considered acceptable. Thus, the relevant risks do not need to be analyzed further. The application of the code of practice must be documented in the hazard log as safety requirements for the corresponding hazards (European Implementing Regulation Nr. 402/2013, p. 19f).

If the risk caused by a particular hazard cannot be limited to an acceptable level by an appropriate code of practice, an alternative risk acceptance principle must be used. In case a similar system is used, hazards are adequately covered by a reference system. Risks associated with the hazards covered by the reference system are then considered acceptable. If the hazards cannot be adequately covered by either code of practice or a reference system, the risk's acceptability must be determined by an explicit risk analysis (European Implementing Regulation Nr. 402/2013, p. 20f).

3 Approach of the ATO-RISK project

The introduction of automated driving on federal railways in Germany is a significant change by the definition of the CSM Regulation (European Implementing Regulation Nr. 402/2013, p. 12), at least when it comes to the driverless variants GoA 3 (driverless operation) or GoA 4 (no operating personnel on board). Therefore, the focus of this study primarily was on the derivation of risk acceptance for train journeys on federal railways with automated systems in the GoA 3 and GoA 4 automation levels. Overall, the process defined in the CSM Regulation (European Implementing Regulation Nr. 402/2013, p. 17ff) had to be followed, where the principle of risk acceptance to be applied may be selected, see Chapter 2.

For hazards that cannot be covered by either a relevant code of practice or a reference system, an explicit risk analysis becomes necessary. The effective and CSM compliant design of the explicit risk analysis is therefore one of the key points of the approach that was conducted in the ATO-RISK project.

Within the framework of the ATO-RISK project, a system for automated railway operation with GoA 3 as well as GoA 4 was first defined. The system definition is based on a complete function list that specifies the overall functionality under the defined boundary conditions. The system definition had to be carried out at a suitable functional level that is compatible with the CSM Regulation (European Implementing Regulation Nr. 402/2013) and (DIN VDE V 0831-103 2021) as the selected method for explicit risk analysis.

For the explicit risk analysis, the German standard (DIN VDE 0831-103 2021) was applied. DIN VDE 0831-103 has been an acknowledged rule of technology in Germany since 2014. It demonstrably meets the requirements of the CSM Regulation. The risk level depicted in it corresponds to the safety level that is accepted in today's long-distance railway operation in Germany for a wide range of control and safety technology applications – from electronic interlockings (ESTW) and level crossings to ETCS. Moreover, this has been constructed and validated against the requirements of (DIN VDE V 0831-101 2021), as the only semi-quantitative method to date. The applicability and appropriateness of the methodology was also confirmed in (Federal Ministry 2016).

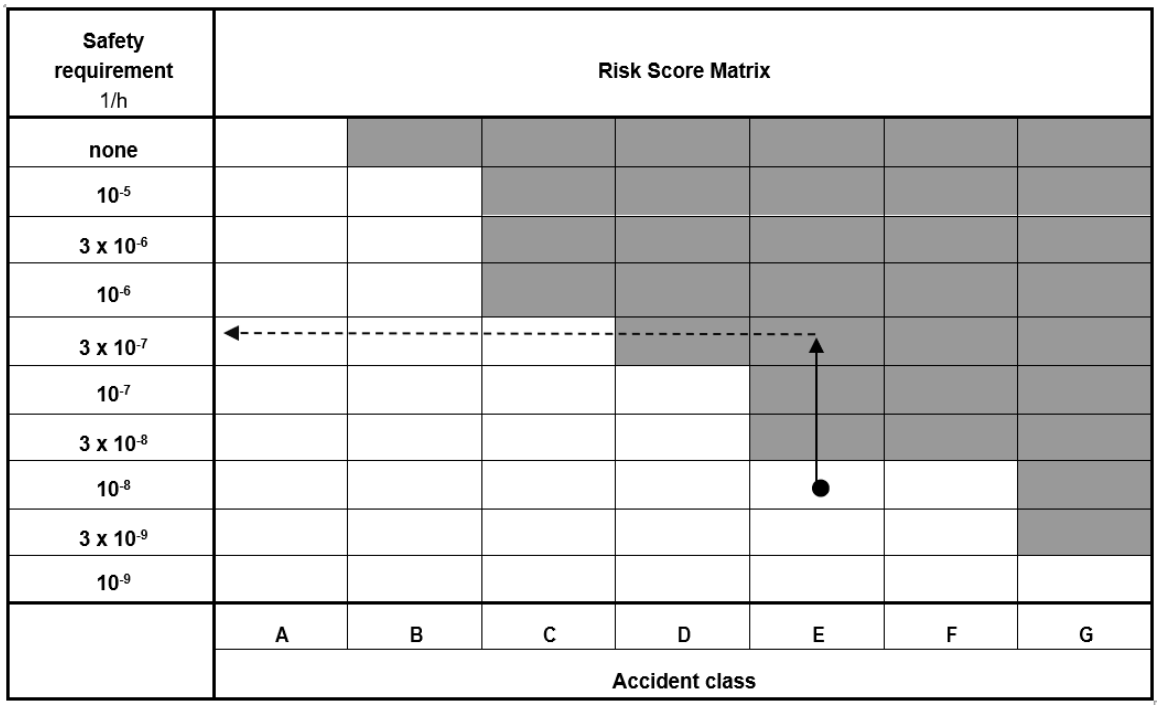


Figure 1: Risk Score Matrix (RSM) in DIN VDE V 0831-103
Source: (DZSF 2023, p. 21)

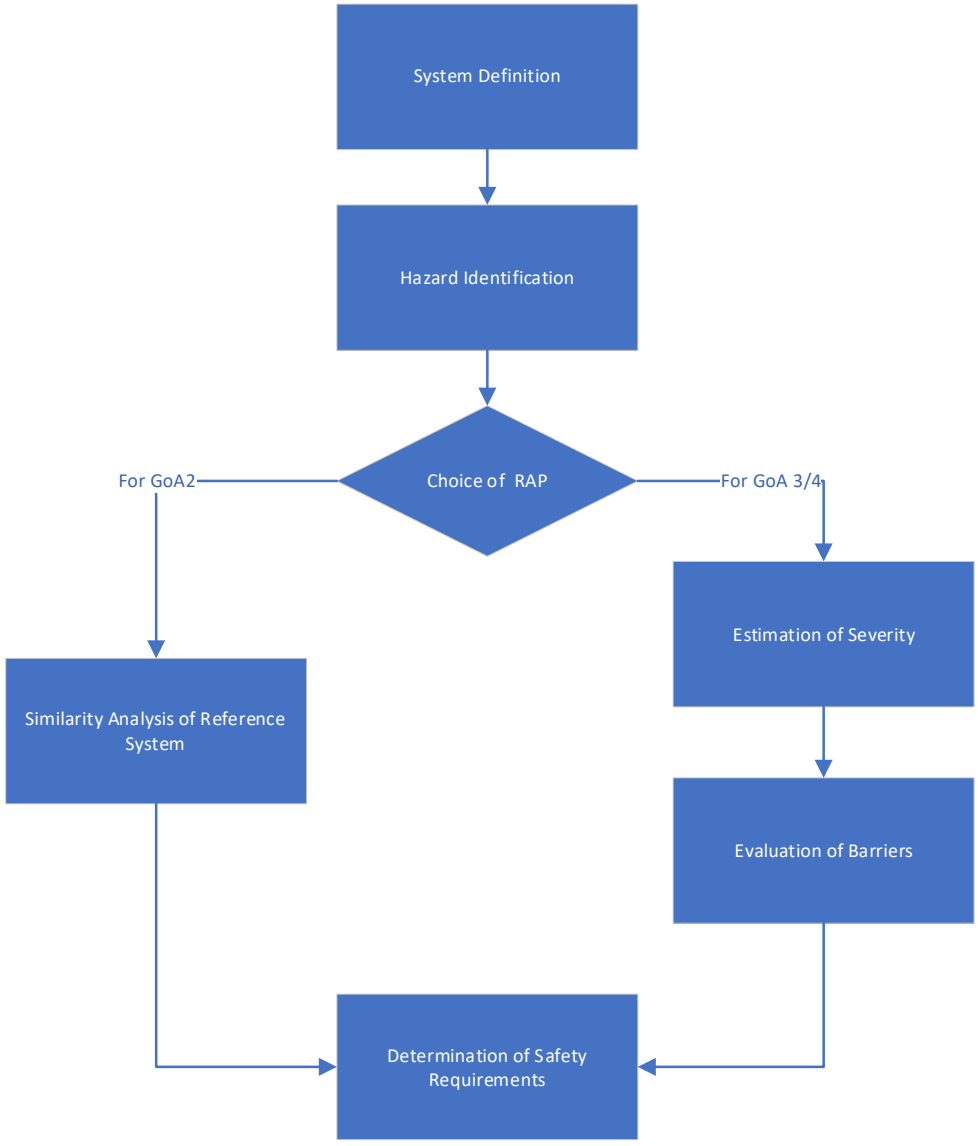
In the ATO-RISK project, the evaluation of different scenarios was carried out according to the semi-quantitative approach presented in DIN VDE 0831-103 with the aid of a so-called Risk Score Matrix (RSM), which qualitatively differentiates the expected severity according to accident classes, see Figure 1. A significant advantage of the RSM is that the estimation of the severity of damage is made qualitatively according to accident classes described by typical event types. Furthermore, these event types were validated based on the DB AG accident database. These classifications have already been estimated towards the safe side so that, i.e., the mean value has not been used but a more severe but still credible severity of damage has been considered. For a more detailed presentation, please refer to (DZSF 2023, p. 13ff).

This semi-quantitative approach was chosen for this purpose, since there is not enough data available for a quantitative risk analysis in Germany to reliably estimate the numerous parameters required with sufficient statistical confidence. This has been shown, among others, by the risk analysis for radio driving operations (FFB) (Braband 2005) and it may be assumed that ATO-RISK has a similar scope or complexity.

To validate the results of the explicit risk analysis, an analysis of human reliability in today's operation with a grade of automation up to GoA 2 was also performed, compare Figure 2. This analysis has been

comprehensively presented in (Adebahr and Schäbe 2023, 21ff). In addition, a crucial point was the consolidation as well as the final validation of the results.

Compared to other approaches, one can see the extended scope and the comprehensive comparison of all aspects for determining the safety integrity requirements from Figure 2. One can also interpret the approach in analogy to the well-known watchdog architecture – a special case of a two-channel system with comparator in the sense of reactive fail-safety according to (DIN EN 50129 2018): The two analyses are performed largely independently based on the common system definition. An independent comparator (validation) makes the results of the risk analysis plausible with the results of the human



reliability analysis, where possible and reasonable.

Figure 2: Procedure ATO-RISK with focus on the comparison of the results of an explicit risk analysis and the determination of human error probabilities
Source: (DZSF 2023, p. 15)

4 Application of code of practice within ATO-RISK

Even when assessing highly novel functions, such as those being investigated in ATO-RISK, code of practice can be of great importance. As a result, the reassessment of functions and systems that are already covered by existing codes of practice can be avoided. Existing specifications can thus be adopted directly.

In the following sections, various examples are given of how the risk acceptance principal “code of practice” is applied in the context of ATO-RISK. While in Section 4.1 and 4.3 functions are directly covered by existing rules and regulations, in Section 4.2 applicable laws are used to limit the dimensioning scope for obstacle detection of persons on the free track.

4.1 Example 1: Detection and fighting of fire/smoke in the train

The European standard EN 45545 "Railway applications - Fire protection in rail vehicles" (DIN EN 45545 2020) applies to both mass transit and mainline applications. The standard defines measures for preventive fire protection in rail vehicles on the one hand and the corresponding verification methods for these measures on the other. There is no coverage of freight wagons by EN 45545 (DIN EN 45545 Part 1 2013, p. 5).

To define measures adapted to the specific operation and type of vehicle, EN 45545 distinguishes between different operation categories and design categories. Design category A refers to vehicles that are part of an automated train and therefore do not have personnel trained for emergency situations on board (DIN EN 45545 Part 1 2013, p. 16). Thus, EN 45545 is directly applicable to GoA 3 and GoA 4 operation, which makes it particularly suitable for ATO-RISK. Requirements relating to systems for fire detection, firefighting, shutdown of technical equipment, and evacuation management are specified in Part 6 of EN 45545 (DIN EN 45545 Part 6 2015). To give an example, EN 45545 prescribes the locations and facilities that must be monitored by fire alarm systems, see Table 1.

Table 1: Prescribed locations of fire detection for design category A

	Operation category	Passenger Areas	Corridors	Toilets	Staff Area	Cooking or catering area	Combustion engines	Technical cabinets containing traction equipment	Other technical cabinets	Luggage compartments
Design Category A	1	nr	nr	nr	nr	nr	X	X	nr	X
	2	nr	nr	X	nr	nr	X	X	nr	X
	3	X[c]	nr	X	nr	nr	X	X	X[b]	X
	4	X[c]	nr	X	nr	nr	X	X	X[b]	X

X: indicates requirement.

nr: indicates no requirement.

b: There are no requirements if there is no electrical traction equipment in the technical cabinet, and if the technical cabinet complies with one of the following conditions: the technical cabinet content is compliant to EN 45545-2, the technical cabinet is contained in a manner compliant to EN 45545-3

c: There are no requirements if the Railway vehicle is not in the field of DIRECTIVE 2008/57/EC

on the interoperability of the rail system within the Community.

Source: (DIN EN 45545 Part 6 2015, p. 8)

Thus, requirements of CSM (European Implementing Regulation Nr. 402/2013) for relevant codes of practice are fulfilled, as (a) the standard is generally recognized in the railway industry throughout Europe; (b) the standard covers all the risks of the hazard under consideration since it is specifically designed for fire protection and covers unattended train operation as well as mainline railways; and (c) the standard is freely accessible.

Fire detection and fighting can thus be covered by code of practice as risk acceptance criterion in GoA 3 and GoA 4 operation. A generic, explicit risk analysis is therefore not required to comply with CSM Regulation.

4.2 Example 2: Dimensioning scope for obstacle detection of persons on the free track

Obstacle detection of persons on the free track is a technically novel field – in particular, for mainline applications. As a result, there are not yet any codes of practice or reference systems that could be used as an applicable risk acceptance principle. Therefore, an explicit risk assessment is required. The procedure for this explicit risk analysis is described in (DZSF 2023, p. 21ff) and in the dedicated paper (Braband et al 2023).

A clear definition and delimitation of the dimensioning scope are crucial. More precisely, one needs to determine, which groups of people the obstacle detection systems are to be dimensioned for. On the free track, the relevant groups of persons are trespassers, construction workers/maintenance personnel, (playing) children, affected/disabled persons, suicides, and injured persons, in particular. See Table 2 and (DZSF 2023, p. 94ff) for more details.

Table 2: Dimensioning scope for obstacle detection of persons on the free track

	Hazard	Risk acceptance principle	Possible protective devices
Person has entered the standard track clearance on the free track from outside the system	Trespasser	Code of practice	Legal regulations, see § 62 EBO and § 19 StVO
	(Playing) children	Explicit risk analysis	On-board obstacle detection, supervisory duty, warning
	Impaired/disabled persons: • Limited perception (language barriers, alcohol, drugs) • Limited mobility (elderly citizens, physically handicapped people) • Mentally handicapped or demented persons • Accompanied by children, animals or personal (bulky) objects. • Hearing and or visually impaired persons	Explicit risk analysis	On-board obstacle detection, warning
	Suicide	Code of practice	Legal regulations, see § 62 EBO and § 19 StVO
	Accidental person (e.g., slipped off hillside)	Explicit risk analysis	On-board obstacle detection
Persons inside the system	Construction workers/maintenance personnel (system internal); e.g., tree pruning or marking work on railway	Explicit risk analysis, (code of practice)	Legal regulations, on-board obstacle detection

Source: (DZSF 2023, p. 95f)

Using the example of trespassers, the question of whether the trespasser should be included in the dimensioning scope of the detection systems is to be clarified. In the German Railway Construction and Operating Regulations (EBO 2019) § 62 "Entering and using railway facilities and vehicles", non-officially authorized persons are prohibited from entering railway facilities unless it is for general traffic use, or a special usage relationship entitles them to do so. In the latter case, however, these individuals would no longer be classified as trespassers. For general traffic use, at level crossings, German Road Traffic Regulations (StVO 2013) § 19 "Level crossings" applies, in which it is expressed that at level crossings with St. Andrew's cross and at level crossings over footpaths, field paths, forest paths or cycle paths (without St. Andrew's cross), rail vehicles have priority. In addition, vehicles or pedestrians must wait at a safe distance in front of the level crossing.

It is cognizable that the requirements for valid code of practice of the CSM Regulation are fulfilled in Germany because (a) EBO and StVO are valid laws in Germany; (b) the laws are directly relevant for the control of the corresponding hazards; and (c) the corresponding laws are publicly accessible.

Thus, as risk acceptance criterion, according to CSM Regulations, codes of practice can be selected that sufficiently mitigate the risk of endangering people on the track, which are trespassers or suicides. As an additional risk reduction, the best possible detection of this group of people should be carried out, for which no explicit risk analysis is required.

In contrast to trespassers, a complete, explicit risk analysis is performed for persons requiring increased levels of protection, like children or affected/disabled persons. Also, it is important to note that a

separate explicit risk analysis must be performed for the obstacle detections of objects (not people). The associated results and preliminary SIL classifications can be found in (DZSF 2023, p. 99).

4.3 Example 3: Person detection in the platform track area

In Germany, fully automated train operation currently only takes place at the Nuremberg metro since June 2008 (apart from people movers). VDV publication 399 (VDV 2000) serves as the basis for this, which could also be applied to German mainline railways as a relevant code of practice, provided the requirements of the CSM Regulation are satisfied.

These are fulfilled because (a) VDV specifications are generally accepted in the railway sector, (b) the successful application of the specifications in similar cases can be shown due to the operational experience with the Nuremberg metro since 2008 and (c) the VDV specifications are distributed openly. The general transferability of the results from mass transit to mainline railways is also confirmed by (Federal Ministry 2016).

For the risk assessment of persons at the platform track area, the speed of the train is particularly decisive. The evaluation of this function involves a distinction between the two operating modes:

1. Train passage
2. Train entry (slow speed is assumed for train entry)

VDV Specification 399 is relevant for the second scenario and the associated hazards, as it represents a relevant code of practice according to CSM Regulation. The train speed applicable at arrival and likewise the vehicles' characteristics, e.g., regarding braking ability, are comparable between mass transit and mainline railways or at least there are no relevant differences with respect to the risk of accidents. Therefore, code of practice can be applied as the risk acceptance principle for train entry and no further explicit risk analysis is required.

For the first scenario, the conditions from mass transit, e.g., speeds during train passage, are not considered comparable to mainline railways. Thus, an explicit risk analysis for the train passage must be performed. The results of this explicit risk analysis can be found in (DZSF 2023, p. 99ff).

All hazards considered in VDV Specification 399 (VDV 2000) lead to requirements that mean SIL 1 according to DIN EN 50129, this corresponds to a TFFR of 10^{-5}/h or smaller. VDV Specification 399 defines a sphere (ball) with a diameter of 30 cm and (if relevant) a weight of 10 kg as the test body. Smaller obstacles do not have to be detected.

5 Human factor analysis

When there was no relevant code of practice or reference system available and, consequently, an explicit risk analysis was performed to determine the safety integrity requirements, the project used a pre-standard aimed at evaluating functions, which are realized by technical systems, as described previously. The comparison, however, needs to be carried out with respect to the current system, in which many of the functions are performed by humans. In order to check to what extent the results fit with today's operation, and especially with human performance in terms of reliability, an assessment of the error probability in the performance of relevant tasks in today's operation up to GoA 2 was carried out.

There are several different methods for estimating the human error probability, which have a different scope. Also, progress regarding human cognition that has emerged in recent decades has changed the approach to estimating human reliability. For ATO-RISK, the dissertation by (Hinzen 1993) and the

RARA manual by (Gibson 2012) were used to derive human error probabilities. The former is widely used in German-speaking countries and the latter is successfully applied in the railway sector in Great Britain.

5.1 Methodology of Hinzen

Hinzen's method can be described as the standard basis for estimating human reliability in the German railway sector within the last decades. However, the data basis is several decades old, and the figures are based on the SHARP method, which emerges from the nuclear industry. Therefore, its applicability today certainly needs to be discussed. Another aspect is that Hinzen mainly evaluates an information processing task, whereas information perception and reaction are mostly not considered.

The method consists of three categories which need to be assessed:

- human behavioral level (skill-based, rule-based, knowledge-based),
- environmental conditions (favorable, unfavorable) and
- stress level (underchallenge, optimal stress level, overchallenge).

Hinzen's classification is applied to a task by assessing its necessary human behavior level, the expected stress level and the expected environmental conditions. Hinzen gives a separate probability of error for each combination. It should be noted that Hinzen's method may lead to misjudgment, if the information processing is not the dominating part for the error of the human action but the information perception or human reaction would dominate. In such cases, the Hinzen method could lead to conservative estimates, e.g., if human perception is the main source of error. Hinzen's method is conservative for typical applications, such as the derivation of barriers, because a comparatively high human error probability is assumed. This was explicitly considered by Hinzen when constructing the method. For the definition of a threshold for a GoA 3 or GoA 4 system as a substitute for a train driver, the approach is not conservative. It would lead to risk acceptance criteria that are less strict compared to an average driver. Therefore, the numerical values for the ATO-RISK project are reduced by a factor of 2, taking into account information used as basis by Hinzen, see Table 3.

Table 3: Error probabilities for ATO-RISK inspired by Hinzen

Human behavioral level	Favorable environmental conditions			Unfavorable environmental conditions		
	Stress due to under-challenge	Optimal stress level	Stress due to overchallenge	Stress due to underchallenge	Optimal stress level	Stress due to overchallenge
Skill-based	1×10^{-3}	5×10^{-4}	1×10^{-3}	5×10^{-3}	$2{,}5 \times 10^{-3}$	5×10^{-3}
Rule-based	1×10^{-2}	5×10^{-3}	1×10^{-2}	5×10^{-2}	$2{,}5 \times 10^{-2}$	5×10^{-2}
Knowledge-based	1×10^{-1}	5×10^{-2}	$2{,}5 \times 10^{-1}$	5×10^{-1}	$2{,}5 \times 10^{-1}$	5×10^{-1}

Source: (DZSF 2023, p. 55)

5.2 Methodology of RARA

The RARA (Railway Action Reliability Assessment) User Manual of the Rail Safety and Standards Board Ltd in London (RSSB) (Gibson 2012) provides an application-oriented guide for the use of the RARA method. The method combines human information perception, information processing, and information implementation into a holistic view of actions. The RARA method is geared towards the railway sector and has been established in England for current investigations of railway accidents. The RARA method is based on the HEART method and adapts it to the railway sector. In recent years, the method has been revised to HEART+ and the resulting changes have been considered in ATO-RISK. An issue of transferability of the RARA method to Germany could be relevant due to the different user behaviors, differences in the training of the operating staff, and differences in the technical implementation of the railway system. However, it was assumed by the project team that the chosen approach is target-oriented and sufficiently accurate for the project.

The error probability for a human action according to RARA is determined in three steps. First, a classification is made according to eight task types (Generic Task Types, GTT for short). Examples are R4 “Skill-based tasks when there are some opportunities for confusion” or R3 “simple response to a dedicated and alarm and execution of actions covered in procedures”. Then an error probability adjustment is carried out according to 27 different "Error Producing Conditions" (EPC). Finally, a weighting of the EPCs is carried out in 10 % steps. The EPCs are divided into the following categories: Task design, Interface design, Persons, Procedures, Competence management and Environment. A resulting error probability is then calculated.

5.3 Application in ATO-RISK

For ATO-RISK, an average train driver is considered. An above-average or below-average train driver is not used as a reference to obtain average values for the overall system and thus to comply with the view on the overall railway system.

The overall process of assessing the individual functions is described in (DZSF 2023, p. 40ff) in detail. In the actual implementation, it became apparent that several assumptions had to be made to achieve comparable results. E.g., it became necessary to decompose some functions that were assessed with (DIN VDE V 0831-103 2021). While the functions follow possible system boundaries of technical systems to be treated with (DIN VDE V 0831-103 2021), individual human actions each comprise several of the mentioned functions. Other assumptions were:

- *Information perception, information processing, and reaction:* Human actions are considered holistically; a division into information intake, information processing, and reaction has not taken place.
- *Simultaneity of human attention:* While different technical systems can simultaneously record and process the same sensory perception redundantly, a human being generally perceives with all senses. A subdivision of functions for the same sense, such as visual obstacle detection at the same time, was not done.
- *Similarity of functions from the point of view of human reliability:* Similar functions are to be presented separately when different results are to be expected. For example, the attention of drivers on the open track and in the station area can be different.
- *Type of information reception:* A piece of information can reach people in several ways, which might influence the error probability. A distinction is made between, for example, the evaluation of faults reported by the vehicle and faults detected by the driver.

For example, the general task "monitoring of track integrity" is divided into the functions "monitoring of infrastructure elements" and "monitoring for environmental influences". The former function is subdivided for the HRA into the two functions "monitor and react to faults of constantly existing infrastructure elements" and "monitor and react to faults of punctually existing infrastructure elements". Again, human action includes not only fault detection but also the subsequent reaction, both of which are mapped and evaluated in one function. It is assumed that the drivers do not constantly monitor existing infrastructure elements, such as the overhead line, whereas a higher level of attention can be assumed for elements that occur at certain points, such as signals. The function "monitoring environmental influences" is rephrased as "monitoring and reacting to environmental influences" for the HRA since human perception is mostly holistic. So, it is not further subdivided. Environmental influences contain, e.g., falling rocks.

The evaluations according to Hinzen and RARA were not used within an overall evaluation or related to each other. This procedure would have generated a combined method according to RARA and Hinzen, which was not the task of this study. For each function, first the method of Hinzen was evaluated and then RARA was used. It should be noted that the separate assessment sometimes yielded very different results between RARA and Hinzen. The assumption suggests that this is caused by possible errors in the perception of the information, which was not considered in Hinzen. Here, RARA delivers more realistic results. This becomes particularly clear when looking at the complementary function to the example above "dealing with obstacles on the free path when it is not immediately clear whether an object is a relevant obstacle" where RARA delivers error probabilities of up to one and Hinzen deviates by up to two orders of magnitude.

All results as well as more detailed explanations to the methods and their applicability are given in (DZSF 2023, p. 40ff).

5.4 Example evaluation

The classification into the GTT for RARA and the classification into skill-based, rule-based, and knowledge-based behavior for Hinzen form the basis of the following assessment. Details can be found in (DZSF 2023, p. 40ff). Here an example function "dealing with obstacles on the open track, provided they are clearly recognizable as obstacles" is presented for better understanding.

For the Hinzen method, normal (favorable) environmental conditions and an increased stress level (excessive demand) of the driver are assumed. An obstacle in front of the train and a potential risk of collision associated with it is likely to trigger stress in the driver. Furthermore, the action is skill-based, since the obstacle is clearly recognizable and does not require the application of rules in the sense of the learned set of rules. The action can be described as sensomotoric. For Hinzen, this results in an error probability of 2×10^{-3} per information processing; for the modified use of Hinzen, this results in 1×10^{-3} per information processing.

For RARA, GTT R3 "One-simple reaction to a warning and the action according to predefined sequences. (Keyword: simplicity)" is chosen, whereby the warning is understood as the obvious visual stimulus of the obstacle. Simplicity implies that rapid braking would be initiated without explicit application of rules. Furthermore, EPC T2 (only very short timeframe for recognizing errors and reacting to it), T6 (very little or no independent supervision of the action by other people) and P2 (shift work or similar organization of work) are applied. P2 is used because the influence of shift work is assumed to be inherent in the system. T6 is used because the collision with smaller obstacles is potentially unnoticed. T2 is used because braking distances in the rail system are long; emergency braking must take place

immediately after detection of the obstacle. With the 25 % and 50 % weighting for all EPCs, the band of error probability per action is derived as 2.94×10^{-3} to 8.64×10^{-3}.

It can be seen that using the conservative limit of the interval for the adjusted RARA method, the results of both methods are still close to each other. However, Hinzen proves to be too conservative in both variants since the information intake is decisive in this example.

6 Validation

For a comprehensive explanation of the methodological approach of ATO-RISK, please refer to (Braband 2021, p. 25ff) or the final report (DZSF 2023).

The validation was carried out in four steps:

1. Review (verification) of the results of the three work packages
2. Validation of the final results by comparing the work results to the status quo GoA 2 as well as risk acceptance GoA 3/4
3. Derivation of significant results using an independent method
4. Discussion of the effects of the statements made on risk acceptance on rail operations after the introduction of GoA 3 or 4

The verification of the results of the work packages was conducted in the form of a review. The review took place in two phases. Firstly, TÜV Rheinland as the validator was present during the presentation and discussion of the interim results. Here, discussions were held to determine whether the work was going into the right direction. This work took place before the relevant reports were prepared. The aspects expressed in discussions were taken into account. The validator expressed his opinion here, but still left enough room for maneuvers so that no decision was imposed on the project team. Such a way of working is important, on the one hand, in order to point out problems as early as possible but also, on the other hand, in order not to endanger the independence of the validator.

No methodological deficiencies were identified during the review. The sources for the function list have been chosen correctly. It can be assumed that the functional list is correct and complete. No deficiencies were found in the further preparations regarding the severity classification of accidents and the efficiency of barriers either.

The determination of human error probabilities was carried out according to common procedures applicable in the railway sector (Adebahr and Schäbe 2023, 21ff). The same applies to the derivation of TFFR values for technical systems that are to replace the driver (DZSF 2023, p. 21ff).

The validation of the final results was performed in two steps. In the first step, a cross-validation of the results was performed. In the second step, a validation of the results against other sources of human error was performed.

As part of the validation of the results, independent calculations and observations were carried out, in particular, using the MEM principle and comparison with accident statistics. In doing so, the results could be confirmed in principle, at least as far as can be expected with four different (semi-) quantitative methods, see Figure 3 for an example.

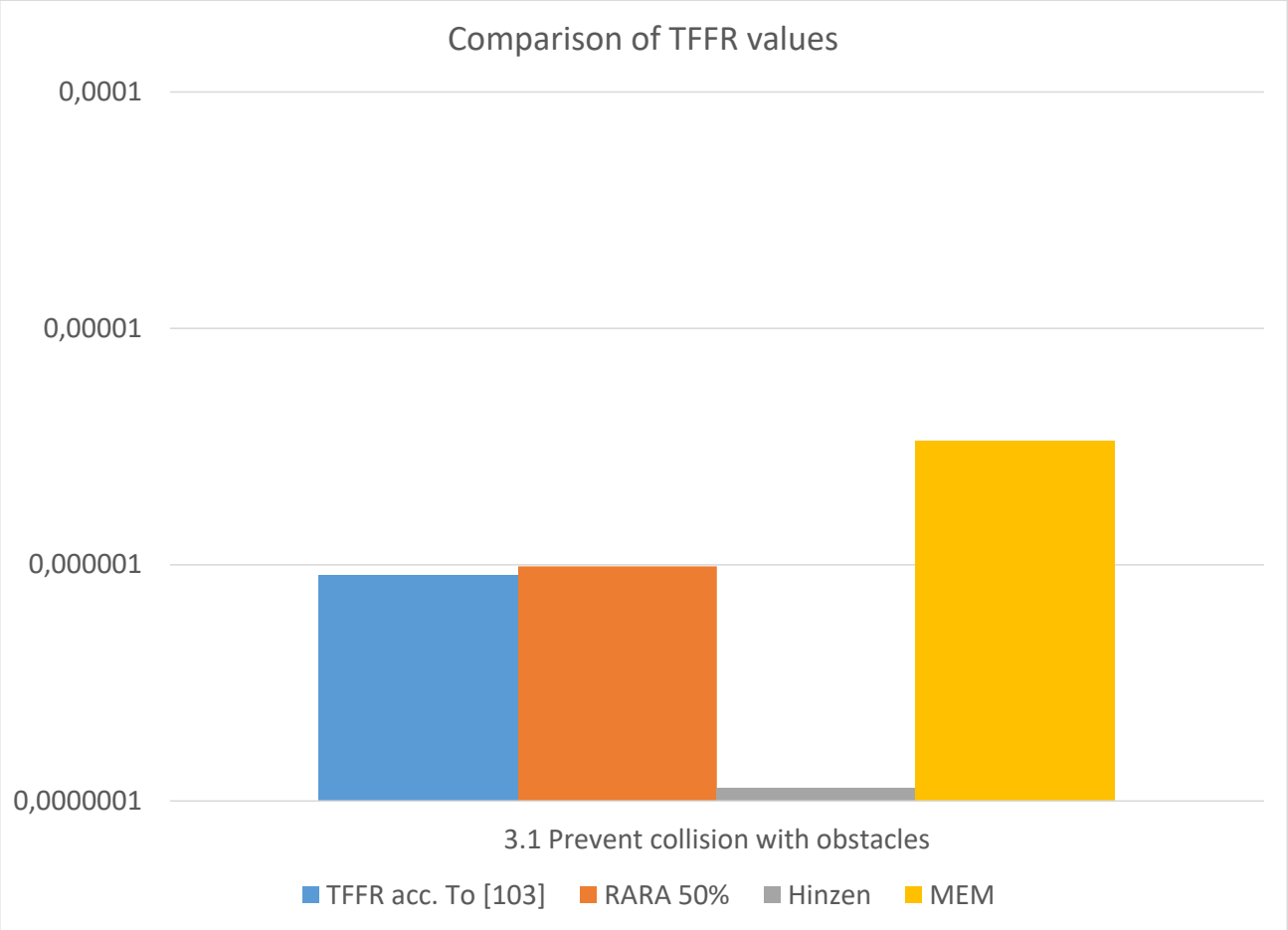


Figure 3: Example of comparison of results

Source: (DZSF 2023, p. 150)

However, especially in comparison with the classical evaluation method according to (Hinzen 1993) there are considerable deviations, which have to be investigated.

The reconciliation of the results with the results for the analysis of the reference system GoA 2 was only possible for functions for which an explicit risk analysis was carried out but this included in particular the important function group of obstacle detection. Here it could at least be shown and confirmed by the validation that from the point of view of the status quo in GoA 2 there are no contradictions with the results of the explicit risk analysis. A more detailed comparison was not meaningful here, since on the one hand the results of the established methods of human reliability analysis already showed partly considerable differences (Adebahr and Schäbe 2023, 21ff), but on the other hand the human reliability analyses also required a different cutting of the function list than in the technical realization (DZSF 2023, p. 60ff). In this context it has to be mentioned that also no better result quality was demanded or expected, but quantitative results should agree in the order of magnitude. This has been achieved in all cases.

The quality of the results can be compared with the results and experiences of (Hinzen 1993), who validated the results of model calculations against accident statistics: e.g. for collisions of trains, the statistical mean value was 3.3×10^{-8} (the units and other parameters are to be neglected here for simplicity), whereas on the basis of the model calculation, the mean value was 1.2×10^{-7} , i.e., the numerical difference was about a factor of 3.6 (where the model was more conservative than the operational statistics, i.e. Hinzen's method gave more pessimistic results). If the numerical difference

had been 10, the statistical confidence would have been lower, but still acceptable for such complex models.

It can thus be stated that the explicit risk analysis according to (DIN VDE V 0831-103 2021) is in principle suitable for deriving risk acceptance criteria for ATO projects. The method according to Hinzen leads in part to stricter requirements, but this is due to the construction or application of the method, which is not suitable for all functions considered (e. g., due to the focus on the information perception for obstacle detection).

Overall, it can be stated that the approach of risk analysis in accordance with the CSM Regulation proved to be very advantageous, as it was possible to adopt existing specifications when using the risk acceptance principles of "code of practice" or "reference system". The codes of practice included building regulations, regulations on fire protection or regulations on the transport of hazardous goods. In addition, the advantage of a semi-quantitative risk analysis based on DIN VDE V 0831-103 was demonstrated for those sub-functions for which neither a set of rules nor a reference system existed.

The (EBO 2019) requires the application of recognised rules of technology, this is somewhat stricter than the principle of risk acceptance "application of codes of practice " of the (European Implementing Regulation Nr. 402/2013), since not all rules and regulations are also recognised rules of practice, see (European Implementing Regulation Nr. 402/2013) section 2.3.2 on the criteria as to which document can be regarded as a set of codes of practice. If safety cannot be demonstrated by means of the recognized rules of technology, the (EBO 2019) opens up the possibility of demonstrating the same safety in clause (2), e.g., by means of a reference system.

Proof of the same level of safety can now be provided by setting adequate safety requirements in a first step and proving that these requirements are met in a further step. The first step has now been taken with the ATO-RISK research project.

This ultimately means that ATO-RISK has to derive such safety requirements that in no case a lower safety level is achieved than before the replacement of the driver by technical systems.

It can be assumed that, due to the conservative approach taken in deriving the requirements for technical systems, a higher level of safety can be expected after the replacement of the driver by technical systems, provided that the relevant standards and regulations are observed in the development and implementation of the systems.

7 Summary and discussion

The project ATO-RISK provides an approach for the risk assessment of highly automated train driving functions starting from a generic level. The system definition based on the (DIN VDE V 0831-101 2021) assures a complete overview on the affected system functions.

The CSM Regulation considers three possible risk acceptance principles. The first one – codes of practice – is usually the most effective way and the method of choice if appropriate standards are available. Despite the novelty of ATO mainline application, the examples given in Section 4 illustrate that some aspects are already covered. The availability of standards and regulations for risk acceptance will certainly enhance when first projects will have been realized and experience is growing.

When standards are not available, a reference system can be selected to derive functional and safety requirements. Anyhow, it is not yet determined when such a reference is sufficiently similar and therefore valid. It is not expected that all system functions can be covered in that way.

For the explicit risk analysis, the third risk acceptance principle, the safety requirements were determined in terms of TFFR values, which allow for the derivation of SIL levels. The results of ATO-RISK are the basis for the risk assessment according to the CSM Regulation. It shall be noted that the analysis is based on assumptions that hold for railway operation in Germany and may vary for other countries or special applications. A guideline for the derivation of results for different cases and assumptions is given in the final report (DZSF 2023, p. 34ff). Since the analyses are based on the Risk Score Matrix method as defined in (DIN VDE V 0831-103 2021), the validity of the standard for the application in the ATO context is required.

The explicit risk analysis has been supplemented by human error considerations applying the RARA method and the method of Hinzen, for details see (Adebahr and Schäbe 2023, 21ff). In principle, similar results as those derived from (DIN VDE V 0831-103 2021) have been obtained.

The project can be considered as a starting point for further research on a way to the realization and homologation of highly automated mainline rail traffic. For the use in a specific solution, additional and deeper analysis is necessary in several points, which were outside the predefined project scope of ATO-RISK:

- The functional requirements for the technology need to be further specified. As an example, the resolution and sensitivity of camera systems, the range of Lidar systems and other technical specifications are still an open point of discussion. A contribution in this point will be delivered in the parallel project ATO-SENSE.
- The safety requirements expressed in terms of TFFR values have to be broken down further to a technology and application specific level.
- Especially when AI methods or other data based, probabilistic solutions are applied, i.e., for perception and recognition tasks, measures and threshold values for reliability have to be derived on a per-case basis. This will require a database of test cases, which is assured to be representative and complete, covering all relevant scenarios and conditions that can occur for a specific application as well as a more precise definition of the underlying technical system, see (Braband and Schäbe 2022, p. 14 ff.).
- It might also be helpful to specify the results further for a defined area and or a specific use case in order to quantify the occurrence of certain hazardous scenarios and to consider special risk reducing measures (i.e., reducing the risk of trespassers).

8 References

[1] Adebahr, Frederik; Schäbe, Hendrik: The Probability of Human Error and Failure Rates of Technical ATO Systems, Signal & Datacommunication 03/2023, p. 21-29

[2] DZSF: Risk Acceptance Criteria for Automated Driving on Rail (ATO-RISK), Final Report, 2023

[3] Braband, Jens; Lindner, Luisa; Rexin, Franziska: Risk analysis for obstacle detection with automated driving, Signal & Datacommunication 03/2023, p. 12-20

[4] Braband, Jens: Risk analyses in railroad automation, Eurailpress, 2005

[5] Braband, Jens: Approach for a Risk Analysis for Automated Train Operation, Signal & Datacommunication, 10/2021, p. 25-34

[6] Braband, Jens; Schäbe, Hendrik: The application of artificial intelligence in railway technology for safety-relevant applications – opportunities and problems, Signalling and Datacommunication 05/2022, p. 14-21.

[7] Federal Ministry of Transport and Digital Infrastructure: Autonomous driving - evaluation of potentials, analysis of existing safety requirements and examination of transferability to the German railroad system, final report, research project 97.370/2016, 2018.

[8] Commission Implementing Regulation (EU) No 402/2013 of 30 April 2013 on the common safety method for the evaluation and assessment of risks and repealing Regulation (EU) No 352/2009, Official Journal of the European Union, supplemented by: Regulation (EU) No 2015/1136, Commission Implementing Regulation (EU) No 2015/1136 of 13 July 2015 amending Implementing Regulation (EU) No 402/2013 on the common safety method for the evaluation and assessment of risks; Official Journal of the European Union L 70 of 16 March 2016 L121/10 of 03.05.2013.

[9] DIN EN 45545: Railway applications - Fire protection on railway vehicles, 2020

[10] DIN EN 45545 Part 1: Railway applications - Fire protection on railway vehicles - Part 1: General, 2013

[11] DIN EN 45545 Part 6: Railway applications - Fire protection on railway vehicles – Part 6: Fire control and management systems, 2015

[12] DIN EN 50129: Railway applications - Communication, signalling and processing systems - Safety related electronic systems for signaling, 2018

[13] DIN VDE V 0831-101: Electrical railway signalling systems - Part 101: Semi-quantitative methods for the risk analysis of technical functions in railroad signal technology, 2021

[14] DIN VDE V 0831-103: Electrical railway signalling systems - Part 103: Determination of safety requirements for technical functions in railway signalling, 2020

[15] EBO, Eisenbahn-Bau- und Betriebsordnung of 8 May 1967 (BGBl. II p. 1563), last amended by Article 2 of the Ordinance of 5 April 2019 (BGBl. I p. 479)

[16] Gibson, Huw: Railway Action Reliability Assessment user manual: A technique for the quantification of human error (T270 Manual), London: Rail Safety and Standards Board Ltd (RSSB), 2012

[17] Hinzen, Albrecht.: The Influence of Human Error on Railway Safety, Doctoral Thesis, RWTH Aachen University, 1993

[18] StVO, Straßenverkehrs-Ordnung of 6 March 2013

[19] VDV: Requirements for devices to ensure passenger safety at train stops during driverless operation, VDV-Specification 399, October 2000

Praktisch umsetzbare Angriffe auf das RaSTA Protokoll

Simon Unger[1], Stefan Katzenbeisser[1], Robert Schmid[2], Andreas Polze[2], Eric Ackermann[2], Marion Christl[1], Sven Gebauer[1], Maximilian Seitz[1], Bastian Schindler[3], Johannes Vedder[2]

[1] Universität Passau

[2] Hasso-Plattner-Institut an der Universität Potsdam

[3] Universität Chemnitz

1 Einleitung

Das Rail Safe Transport Application (RaSTA) Protokoll wird für die Kommunikation zwischen Stellwerken und Feldelementen, wie z.B. Lichtsignalen und Weichen, genutzt. RaSTA ist ein Safety-kritisches Netzwerkprotokoll. Wir unterscheiden in Bezug auf die Sicherheit eines Systems zwei verschiedene Begriffe, nämlich Safety und Security. Safety beschreibt den Schutz gegen Betriebsstörungen oder Fehlfunktionen eines Systems, welche Schaden an Personen oder anderen physischen Objekten anrichten können. Security beschreibt den Schutz gegen böswillige Angriffe auf die Vertraulichkeit, die Verfügbarkeit und die Integrität von Systemen und Informationen. In Eisenbahnsystemen sind diese beiden Konzepte heutzutage eng miteinander verknüpft, da diese Teil einer kritischen Infrastruktur (KRITIS) sind, die auch ein beliebtes Angriffsziel darstellen (Giannopoulos 2021).

Wie von der DIN EN 50159 gefordert, stellt das RaSTA-Protokoll sicher, dass Inhalt und Reihenfolge von Nachrichten während der Übertragung zwischen Sender und Empfänger nicht verändert werden können. Allerdings zielt das Protokoll nur auf unbeabsichtigte Fehler in der Kommunikation ab, womit die „Safety“ behandelt wird. Jedoch muss zusätzlich auch die „Security“, also die Sicherheit des Protokolls vor gezielten Angriffen, betrachtet werden.

Wir zeigen in dieser Arbeit erstmals experimentell, dass die im RaSTA Protokoll spezifizierten Schutzmaßnahmen keinen hinreichenden Schutz gegen gezielte Angriffe bieten und schlagen Verbesserungen vor, welche die Security des Protokolls erhöhen. Wir weisen experimentell nach, dass ein aktiver Angreifer beliebige Änderungen an übertragenen Nachrichten vornehmen kann, ohne dass die Kommunikationspartner dies nachweisen können. Dadurch können Safety-Maßnahmen wie Flankenschutz ausgehebelt und unsichere Situationen hervorgerufen werden. Außerdem demonstrieren wir, dass der Angriff mit handelsüblicher PC-Hardware für jedermann durchführbar ist. Dazu beschleunigen wir unter anderem einen Brute-Force-Angriff von Heinrich et al. (Heinrich 2011, S. 199ff) um den Faktor 2.500.

In der vorliegenden Arbeit wird zuerst der Aufbau und die Funktionsweise des RaSTA Protokolls kurz erklärt, sowie auf die im Protokoll vorgegebenen Maßnahmen zur Sicherung der Integrität der Nachrichtenübertragung eingegangen. Anschließend wird der Laboraufbau beschrieben, der zum Testen verschiedener Angriffe auf RaSTA benutzt wurde. Im vierten Kapitel werden die durchgeführten Angriffe gegen RaSTA vorgestellt. Diese basieren auf einer theoretischen Schwachstellenanalyse von

[1] Korrespondierender Autor: Simon.Unger@uni-passau.de

Diese Arbeit wurde im Rahmen des Projekts „FINESSE“ (FKZ16KIS1587) durchgeführt. Wir danken für die finanzielle Unterstützung des Projekts durch das Bundesministerium für Bildung und Forschung (BMBF).

Heinrich et al. (Heinrich 2011, S. 199ff). Sie wurden im Rahmen der Digital Rail Summer School 2022 auf ihre Umsetzbarkeit untersucht und verbessert. Abschließend werden Möglichkeiten vorgestellt, welche die Sicherheit des Protokolls gegen gezielte Angriffe kurz-, mittel- und langfristig erhöhen können.

2 Das Rail Safe Transport Application (RaSTA) Protokoll

Das RaSTA-Protokoll ist ein Transportprotokoll für Safety-kritische Kommunikationsnetze und ist in der DIN VDE V 0831-200 (DKE 2015) standardisiert. Es wird für den Transport von Nachrichten höherer Protokolle zwischen dem Stellwerkskern und den Feldelementen, wie Lichtsignalen (SCI-LS), Achszählern (SCI-TDS) und Weichen (SCI-P) verwendet. RaSTA besteht aus zwei verschiedenen Schichten, einer Sicherheits- und Wiederholungsschicht und einer Redundanzschicht, und setzt seinerseits auf dem verbindungslosen UDP-Protokoll auf. Die Redundanzschicht ist dafür verantwortlich, die richtige Nachrichtenreihenfolge sowie Zeitgenauigkeit und Verfügbarkeit aufrechtzuerhalten. Dies geschieht über eine konfigurierbare Anzahl von redundanten, physisch unabhängigen UDP-Kanälen. Die Sicherheits- und Wiederholungsschicht stellt hingegen sicher, dass verlorene oder beschädigte Pakete erneut angefordert werden. Weiterhin gewährleistet sie die Authentizität und Integrität von Nachrichten (DKE 2015). Authentizität von Nachrichten bedeutet in diesem Kontext, dass eine Nachricht tatsächlich vom angegebenen Sender an den angegebenen Empfänger gesendet wurde. Integrität bedeutet, dass eine Nachricht auf dem Übertragungsweg nicht geändert wurde. Beide Sicherheitseigenschaften werden dabei durch eine 8-Byte-Prüfsumme, welche auch als Sicherheitscode bezeichnet wird, gewährleistet. Die Prüfsumme wird mithilfe eines Message Authentication (MAC)-Verfahrens gebildet, welches einen geheimen und beiden Kommunikationspartnern bekannten Schlüssel nutzt. Durch eine Sequenznummer im geschützten Teil des Pakets wird sichergestellt, dass die Reihenfolge von Paketen nicht verändert werden kann, ohne die Prüfsumme zu verändern. Ebenso kann die Identität von Sender und Empfänger nicht geändert werden und es können keine zusätzlichen Pakete in eine Kommunikation eingebracht werden, ohne die Prüfsumme neu zu berechnen. Dadurch wird die Authentizität der Kommunikationspakete gewährleistet.

Das Protokoll erfüllt dadurch die Systemanforderungen (Authentizität, Integrität, Zeitgenauigkeit, und Reihenfolge), die in EN 50159 für Eisenbahnanwendungen definiert sind, sofern keine gezielten Angriffe durchgeführt werden. Jedoch haben Heinrich et al. (Heinrich 2011, S. 199ff) bereits gezeigt, dass diese Anforderung nicht ausreichen, damit das RaSTA Protokoll als sicher vor aktiven Angriffen eingestuft werden kann. Insbesondere kann ein Angreifer das unterliegende MAC-Verfahren brechen, wodurch die Sicherheitseigenschaften des Protokolls gegenüber aktiven Angriffen nicht mehr gewährleistet sind.

3 Laboraufbau und Voraussetzungen

Um die von Heinrich (Heinrich 2011, S. 199ff) spezifizierten Angriffe auf ihre praktische Umsetzbarkeit zu untersuchen, wurde eine Open-Source Referenzimplementierung des RaSTA Protokolls[2] verwendet. Diese wurde im „EULYNX Live Lab“ installiert, einer Laborumgebung, in der sowohl simulierte als auch reale Weichen und Lichtsignale mit einer Stellwerkssimulation verbunden sind. Das verteilte Labor

[2] https://github.com/Railway-CCS/rasta-protocol

ermöglicht reproduzierbare Tests von EULYNX-kompatiblen digitalen Stellwerkskomponenten, die von verschiedenen Partnern aus Wissenschaft und Industrie bereitgestellt werden.

Für die Security-Analyse des RaSTA-Protokolls stehen eine Stellwerksansicht und eine Zustandsansicht eines simulierten Lichtsignals zur Verfügung. Voraussetzung zum Ausführen der Angriffe ist der Zugriff zum Kommunikationsnetzwerk. In der Realität ist dieser durch Zugriff auf ein Gerät im Netzwerk oder Zugang zu den Kommunikationskabeln relativ leicht zu erlangen. Im Versuchsaufbau hatte der Angreifer Zugriff auf einen Router, der Nachrichten im Netzwerk weiterleitete.

Die RaSTA-Spezifikation ist frei verfügbar und kann von jedermann erworben werden. Außerdem werden standardisierte Bausteine für die Implementierung des Protokolls verwendet. Daher werden diese und ihre Inhalte auch als allgemeines Wissen betrachtet. Auch soll nach dem Kerckhoffs'schem Prinzip (Petitcolas, 2011, S.1) die Security nicht von einer konkreten Implementierung abhängig sein.

4 Praktisch umsetzbare Angriffe

Mithilfe des in Abschnitt 3 beschriebenen Testaufbaus wurde im Rahmen der Digital Rail Summer School 2022 das RaSTA-Protokoll auf mögliche Angriffe getestet. Zunächst spielte es keine Rolle, ob es sich um Implementierungs- oder Einrichtungsfehler handelt, welche schließlich zur Schwachstelle führen. Dadurch konnten wir das Protokoll besser verstehen und mögliche Angriffspunkte für anspruchsvollere Angriffe entdecken, die aufgrund von Schwachstellen im Protokoll möglich sind.

Die folgenden Angriffe wurden je von zwei unabhängigen Teams erfolgreich durchgeführt und sind unabhängig von der spezifischen Implementierung durchführbar. Wir erwähnen jedoch, wo korrekte Implementierung und Einstellungen des Protokolls Angriffe verhindern können.

4.1 Replay Angriff

Bei einem Replay-Angriff zeichnet ein Angreifer die Nachrichten zwischen zwei Kommunikationspartnern auf und sendet sie zu einem späteren Zeitpunkt erneut an eine der beiden Parteien. Dadurch wird beim Empfänger der Nachrichten ein neuer Zustand erreicht, ohne dass dies dem ursprünglichen Sender bekannt ist. Im RaSTA-Kontext kann ein Angreifer etwa die Kommunikation zwischen dem Stellwerk und einem Lichtsignal aufzeichnen, bis eine Nachricht empfangen wird, die das Signal auf „Fahrt" stellt. In einem Angriffsszenario könnte diese zu einem späteren Zeitpunkt erneut gesendet werden, um das Signal fälschlich auf „Fahrt" zu stellen, während das Stellwerk davon ausgeht, dass das Signal auf „Halt" steht. Im schlimmsten Fall kann ein solcher Angriff eine kontrollierte Entgleisung oder sogar eine Zugkollision verursachen.

Sobald ein Angreifer Zugang zu einem Gerät im Netzwerk hat, kann dieser Angriff nicht grundsätzlich vom RaSTA-Protokoll verhindert werden. Nachrichten im RaSTA-Protokoll verwenden zwar Sequenznummern und Zeitstempel, letztere werden jedoch nur für die Überwachung der Kanalgüte verwendet. Sequenznummern können den Angriff erschweren, bieten jedoch allein keinen zuverlässigen Schutz vor Replay-Angriffen. Durch die feste Größe der Sequenznummer in den Paketen kann nur eine endliche Zahl von Paketen übertragen werden, bis sich eine Sequenznummer wiederholt. Wenn sich allerdings eine Sequenznummer wiederholt, kann der Angreifer frühere Kommunikation mit derselben Sequenznummer einspielen, ohne dass der Empfänger dies bemerken kann.

Diese Situation kann auch eintreten, ohne dass alle Sequenznummern einmal gesendet wurden. Eine korrekte Implementierung beginnt bei einer neuen Kommunikation mit einer nicht vorhersagbaren

zufälligen Sequenznummer. In diesem Fall muss ein Angreifer bei üblichen Paketraten etwa 42 Jahre warten, bis sich die erste Sequenznummer wiederholt und dadurch ein Replay-Angriff möglich wird. Falls eine Implementierung allerdings bei einer neuen Kommunikation oder nach einem Geräteneustart mit einer festen Sequenznummer wie 0 beginnt, verhält sie sich so, als wären alle Sequenznummern einmal verwendet worden. Eine solche Implementierung ist also besonders verwundbar gegen den genannten Angriff. Durch dieses Verhalten kann der Angreifer sofort den Replay-Angriff durchführen, ohne abzuwarten, bis sich die Sequenznummern wiederholen.

Wie bereits erwähnt, war im Setup der Zugriff auf einen Router, welcher die Nachrichten im Netzwerk weitergeleitet, bereits gegeben. Das Aufzeichnen der Nachrichten wurde mit Hilfe von Packet-Capture-Programmen, wie Wireshark, realisiert. Mithilfe eines simplen Programmes auf dem Router wurden danach ausgewählte Datenpakete an das Feldelement weitergeleitet. Somit ist es möglich, Datenpakete, welche die passende Sequenznummer beinhalten, am Router zu unterdrücken und stattdessen die zuvor aufgezeichneten Pakete zu senden.

Die einzige den Autoren bekannte Möglichkeit, diesen Angriff zu verhindern, ist durch ein sogenanntes Key-Rotation-Verfahren. Dabei wird, sobald sich die Sequenznummer wiederholen würde, ein neuer geheimer Schlüssel für das MAC-Verfahren ausgehandelt. Wird in diesem Fall alte Kommunikation erneut eingebracht, ist die Prüfsumme unter dem neuen Schlüssel nicht mehr gültig, und der Angriff schlägt fehl. Ein Beispiel für ein solches Verfahren ist der 4-Way Handshake aus dem WPA2-Protokoll.

4.2 Man-in-the-Middle Angriff

Ein Man-in-the-Middle (MitM)-Angriff nutzt nicht nur bereits gesendete Pakete, wie bei einer Replay-Attacke, sondern fügt gänzlich neue oder manipulierte Pakete in die Kommunikation ein. Wenn ein Angreifer in der Lage ist, den Inhalt einer Nachricht zu manipulieren, kann keine der bisher genannten Sicherheitsmaßnahmen die Security der Kommunikation garantieren. Alle Attribute der Pakete, wie etwa Sequenznummern, können so verändert werden, dass der Angriff nicht erkannt werden kann. Für den Angreifer uninteressante Pakete wie Heartbeats werden automatisch weitergeleitet und Steuerbefehle vom Stellwerk oder Antworten vom Feldelement können nach Wunsch des Angreifers manipuliert werden. Letzteres ist besonders interessant, wenn man sicherstellen möchte, dass das Stellwerk den Angriff nicht erkennt.

Die gegenseitige Authentifizierung von Stellwerk und Feldelementen, sowie die Integrität der Kommunikation im RaSTA-Protokoll wird gemäß Spezifikation ausschließlich durch die Kenntnis des geheimen Schlüssels für das MAC-Verfahren garantiert. Der Sicherheitscode in einer RaSTA-Nachricht ist der einzige Wert im Protokollstapel, welcher von einem Angreifer ohne Kenntnis des Schlüssels nicht berechnet werden kann. Daher benötigt ein Angreifer, der Pakete manipulieren möchte, nur Zugriff auf den Netzwerkschlüssel, damit die Prüfsumme neu berechnet werden kann und gültige Pakete gesendet werden können.

An dieser Stelle ist das in RaSTA verwendete MAC-Verfahren die Schwachstelle. Die Autoren der RaSTA-Spezifikation sahen die MD4-Hashfunktion zur Bildung des Sicherheitscodes vor. Dieser entspricht dabei dem MD4-Hash der Nachricht, welcher optional noch auf 8 Byte verkürzt wird. Eine Hashfunktion berechnet eine Prüfsumme fester Länge über eine Nachricht, welche nur mit vernachlässigbar geringer Wahrscheinlichkeit mit der Prüfsumme einer anderen Nachricht kollidiert. Allerdings gilt MD4 bereits seit 1995 als kryptographisch gebrochen, sodass MD4 allein nicht als MAC-Verfahren fungieren kann (Dobbertin, 1998, S. 253ff). Daher haben die Autoren von RaSTA ein MAC-Verfahren auf Basis von MD4 entwickelt, indem sie optional den Initialisierungsvektor der zugrunde liegenden Kompressionsfunktion durch den geheimen Schlüssel ersetzen. Es wird von den Autoren angenommen,

dass ein Angreifer ohne Kenntnis des konkret verwendeten Initialisierungsvektors nicht mehr in der Lage ist, Prüfsummen für veränderte Pakete zu fälschen. Allerdings ist es für einen Angreifer unter Kenntnis des geheimen Schlüssels trivial möglich, Prüfsummen zu fälschen und damit beliebige gefälschte Nachrichten in die Verbindung einzufügen.

Dies kann auf verschiedene Arten erreicht werden: (1) In der Konfiguration der Geräte wird statt eines geheimen Schlüssels der öffentlich bekannte Standardinitialisierungsvektor verwendet. Dieser ist in dem Standard DIN VDE V 0831-200 (DKE, 2015) angegeben. (2) Durch physischen Zugriff auf ein Feldgerät, da alle Geräte den Netzwerkschlüssel kennen müssen. Der Angreifer kann in diesem Fall die legitimen Interfaces zum Speicher benutzen, um den Initialisierungsvektor auszulesen. (3) Durch einen Brute-Force-Angriff, der in Abschnitt 4.4 näher erläutert wird.

Sobald der Netzwerkschlüssel vorliegt, werden die Steuerbefehle des Stellwerks abgefangen und die ursprüngliche Payload gespeichert. Die ursprüngliche Payload wird dann durch den gewünschten Befehl ersetzt und die Prüfsumme wird neu berechnet. Das manipulierte Paket kann dann weitergeleitet werden. Damit das Stellwerk den Angriff nicht bemerkt, wird auch die Antwort des Feldgeräts abgefangen und die Payload wird auf die des ursprünglichen Steuerbefehls geändert. Anschließend wird die Prüfsumme erneut berechnet und die manipulierte Antwort wird an das Stellwerk zurückgesendet. Das Stellwerk glaubt somit, dass der tatsächliche Steuerbefehl implementiert wurde. Im Rahmen des Projektes haben wir diesen Angriff in der Simulationsumgebung implementiert. Dabei wurde zunächst ein bekannter Netzwerkschlüssel angenommen. Im nächsten Abschnitt wird ein Verfahren beschrieben, das auch bei einer unbekannten Passphrase einen Angriff ermöglicht.

4.3 Passwort Angriff

Eine weitere Option besteht darin, die Passphrase mithilfe von handelsüblicher Hardware und frei verfügbarer Software durch einen Brute-Force-Angriff wiederherzustellen. Zunächst wird mithilfe von Netzwerkanalyseprogrammen wie Wireshark der Nachrichtenverkehr aufgezeichnet und nach Paketen mit RaSTA-Heartbeat-PDU (Protocol Data Unit) gefiltert. Es kann jedes beliebige Paket anstelle des Heartbeats verwendet werden, allerdings ist der Angriff mit Heartbeat-Paketen am effizientesten, da diese die kürzesten Pakete im Protokoll sind. Im zweiten Schritt werden die PDU der Sicherheits- und Wiederholungsschicht aus den Paketen extrahiert und in den Sicherheitscode und die verbleibende PDU aufgeteilt. Im letzten Schritt wird ein von uns für hashcat entwickeltes Modul verwendet, um die Passphrase mithilfe eines Brute-Force-Angriffs wiederherzustellen. hashcat ist ein Open-Source-Programm zur Passwortwiederherstellung, das konzeptionell aus zwei Teilen besteht: (1) einem Passwortgenerator, der Passphrasen-Kandidaten aus Wortlisten, kombinierten Wortlisten, vom Benutzer generierten Mustern und anderen Quellen generiert, und (2) OpenCL-Kernen, die die Passphrase-Kandidaten gegen ein gegebenes Ziel validieren.

Durch eine geringfügige Modifikation der OpenCL-Kerne für MD4 können diese verwendet werden, um einen effizienten Brute-Force-Angriff gegen RaSTA-Passphrasen für sowohl 8- als auch 16-Byte-Sicherheitscodes durchzuführen. Dafür muss lediglich der Standardinitialisierungsvektor durch den Passwort-Kandidaten ausgetauscht werden. Wenn derselbe Prüfcode für die aufgezeichneten Nachrichten berechnet wurde, dann entspricht der aktuelle Passwortkandidat dem tatsächlichen geheimen Schlüssel. Die modifizierten Kerne ordnen Passphrase-Kandidaten in ASCII-Codierung den Bytes des MD4-Initialisierungsvektors zu und füllen den Vektor mit den verbleibenden Bits des MD4-Standard-IV auf. Diese Konstruktion des Initialisierungsvektors (IV) basiert auf dem Beispiel aus Anhang A von VDE 0831-200. Andere Konstruktionen des IV ändern den grundlegenden Angriff nicht und

können mit einer geringfügigen Modifikation der Kernel implementiert werden. Durch das Ausführen des Kernels auf GPGPUs kann ein hohes Maß an Parallelisierung des Angriffs erreicht werden. In einem Benchmark auf einer NVIDIA RTX 2080 GPGPU konnten zwischen 3,8 und 4 Milliarden Passwortkandidaten pro Sekunde im langsamsten Angriff validiert werden. Dies ermöglicht es, den gesamten Raum einer alphanumerischen Passphrase mit 8 Byte in etwa 20 Tagen zu generieren und zu validieren. Es sollte beachtet werden, dass Passwörter in der Praxis oft nicht als zufälliger Text gebildet werden, sondern nach Kombinationsregeln. hashcat kann beliebige Regeln effizient abdecken, was auch einen Angriff auf die maximalen 16 Byte der RaSTA-Passphrase realistisch macht. Speziell für Angriffe auf Listen bekannter Passwörter (Wörterbuchangriffe) steht auch ein schnellerer Kernel zur Verfügung, der Schwachstellen in der MD4-Funktion ausnutzt, um Passphrase-Kandidaten mit weniger Anweisungen zu berechnen und somit noch mehr Kandidaten pro Sekunde berechnen kann.

5 Mögliche Security-Verbesserungen

Abgesehen von den tatsächlich durchführbaren Angriffen wurden auch Möglichkeiten untersucht, die Security des Protokolls zu erhöhen. Im Folgenden schlagen wir kurz-, mittel- und langfristige Verbesserungen vor. Die kurzfristige Methode ist am einfachsten zu realisieren, bietet aber langfristig gesehen keinen ausreichenden Schutz. Die kurz- und mittelfristige Methode kann also die Zeit überbrücken, bis die langfristige Verbesserung umgesetzt werden kann, jedoch bietet nur Letztere langfristig einen ausreichenden Schutz.

5.1 Kurzfristige Verbesserung: Erweiterung des Sicherheitscodes

Wie die oben genannten Angriffe zeigen, ist die Verwendung eines alpha-numerischen Passworts mit 8 Byte Länge praktisch in kurzer Zeit angreifbar. Daher empfehlen wir, nur zufällig generierte Zeichenfolgen mit einer Länge von 16 Byte zu verwenden. Da das Passwort für das spezifizierte Verfahren im Klartext im Feldelement verfügbar sein muss und somit daraus extrahiert werden kann, sollte für jedes Feldelement und jede Weiche auch ein individuelles Passwort definiert und regelmäßig in kurzen Abständen geändert werden.

Der in 4.3 durchgeführte Angriff ist damit immer noch umsetzbar, dauert in der Regel jedoch länger, weswegen die Zeitabstände, in denen Passwörter geändert werden, entsprechend gering gewählt werden müssen. Dadurch, dass das Risiko eines erfolgreichen Angriffs nur geringfügig niedriger wird, bietet diese Verbesserung keine zuverlässige dauerhafte Sicherheit.

5.2 Mittelfristige Verbesserung: Tunneling

Eine mittelfristige Mitigierung des Angriffs kann durch das Tunneln von RaSTA durch ein Sicherheitsprotokoll wie (D)TLS oder OpenVPN erreicht werden. Tunneling kann die kryptografische Integrität und Vertraulichkeit der Kommunikation gewährleisten, während Zertifikate die gegenseitige Authentifizierung der Kommunikationspartner ermöglichen. In einem Projekt zur Implementierung und Evaluation von RaSTA-Tunneling durch die TLS1.3- und DTLS1.2-Protokolle des Lehrstuhls für Betriebssysteme und Middleware an der Universität Potsdam wurde nachgewiesen, dass auch durch TLS- und DTLS-Tunnel die Funktionalität des RaSTA-Protokolls garantiert werden kann. Die Verwendung von Tunneling führt jedoch auch zu neuen Problemen. Erstens garantieren weder VPN-Protokolle noch (D)TLS die Einhaltung erforderlicher Timing- oder Zuverlässigkeitsgarantien. Insbesondere bei der Verwendung von TLS kann es durch das Congestion-Protokoll des unterliegenden TCP-Transportprotokolls zu nicht vorhersagbaren Unterbrechungen in der Kommunikation kommen. Es

muss also immer noch nachgewiesen werden, dass Zeit- und Zuverlässigkeitsgarantien eingehalten werden können, wenn alle redundanten Kanäle getunnelt werden.

Insbesondere die Verwendung von TLS macht das RaSTA-Protokoll außerdem besonders verwundbar gegenüber Angriffen auf die Verfügbarkeit. Wenn ein Angreifer ein einzelnes Bit eines TLS-Records ändert, bricht TLS die Verbindung ab. Danach muss der Tunnel neu aufgebaut werden, was einen hohen Kommunikations- und Berechnungsaufwand erfordert (Rescorla 2008).

Besonders für zeitkritische Anwendungen wird die Verwendung von DTLS anstelle von TLS empfohlen, da DTLS auf UDP als Transportprotokoll basiert und somit keine Sendewiederholung implementiert ist. Damit wird im Gegensatz zu TLS die Aufgabe von RaSTA nicht vom Sicherheitsprotokoll dupliziert, was wiederum den Overhead reduziert.

Darüber hinaus werden in TLS und seinen Derivaten die Authentifizierung und Autorisierung nur durch kryptografische Zertifikate sichergestellt. Die Integrität der öffentlichen Schlüsselinfrastruktur (PKI), insbesondere die Geheimhaltung der privaten Schlüssel von Root-Zertifikaten, muss daher sichergestellt werden, um diesen Sicherheitsgarantien zu entsprechen. Eine Möglichkeit der weiteren Absicherung ist das sogenannte „Zertifikats-Pinning“. Dabei wird auf jedem Gerät für alle Kommunikationspartner das korrekte Zertifikat des Partners hinterlegt. Wenn der geheime Schlüssel des Root-Zertifikats kompromittiert wurde, können die Geräte abweichende Zertifikate von den gespeicherten Zertifikaten erkennen und die Verbindung verweigern. Außerdem erfordert TLS eine regelmäßige Aktualisierung der TLS-Implementierung (bei jeder relevanten Revision des TLS-Standards und bei Bekanntwerden von Sicherheitslücken in der Implementierung) und auch regelmäßige Rotation der Root-Zertifikate, um langfristige Sicherheit zu gewährleisten. Darüber hinaus sollten keine Zertifikate ohne Widerrufsinformation ausgestellt werden, um potenziell kompromittierte Zertifikate widerrufen zu können, und es sollte ein Mechanismus für schnelle Updates der TLS-Implementierung eingeführt werden, falls ernste Implementierungsschwachstellen bekannt werden.

5.3 Langfristige Verbesserung: Protokollverbesserungen

Um langfristig Sicherheit zu gewährleisten, sollte das RaSTA-Protokoll selbst verbessert werden. Zum einen empfehlen wir daher, wie bereits von Heinrich et al. vorgeschlagen (Heinrich 2011, S. 199ff), die MD4-Hashfunktion durch ein kryptografisch sicheres MAC-Verfahren zu ersetzen. Man könnte entweder eine MAC-taugliche Hashfunktion wie Blake2B verwenden, deren Sicherheitseigenschaften mit SHA-3 vergleichbar sind und die immer noch eine mit MD4 vergleichbare Leistung bietet, oder ein generisch konstruiertes HMAC auf Basis einer sicheren Hashfunktion. Die Ersetzung der Hashfunktion löst jedoch nicht das grundlegende Sicherheitsproblem, sondern reduziert nur die Wahrscheinlichkeit, noch effizientere Angriffe als die oben beschriebenen zu finden. Da moderne Hashfunktionen, insbesondere Blake, für effiziente Berechnungen modifiziert sind, bleiben reine Brute-Force-Angriffe eine Möglichkeit.

Um das grundlegende Sicherheitsproblem der langfristigen Passwort- und Brute-Force- Verwundbarkeit zu lösen, empfehlen wir, das RaSTA-Protokoll um ein Schlüsselaustauschprotokoll, wie etwa einen Password-Authenticated Key Exchange (PAKE), zu erweitern. Ein PAKE ist ein kryptografisches Primitiv, welches die Erzeugung von geheimen Sitzungsschlüsseln zwischen zwei Parteien und die gleichzeitige gegenseitige Authentifizierung der Parteien mit lediglich einem Passwort ermöglicht. Im Gegensatz zu TLS ist hier kein Zertifikat und keine PKI nötig, um Sicherheit zu gewährleisten. Dies wurde von uns in einer Proof-of-Concept-Implementierung mit dem OPAQUE-PAKE-Protokoll und der vorhandenen Open-Source-RaSTA-Implementierung implementiert und getestet. Hier führen die

RaSTA-Kommunikationspartner nach dem Verbindungsaufbau zuerst das OPAQUE-Protokoll gemäß dem aktuellen IETF-Entwurf aus. Anschließend wird der abgeleitete Sitzungsschlüssel als MD4-Initialisierungsvektor verwendet. Der abgeleitete Sitzungsschlüssel könnte auch als Schlüssel für ein verbessertes MAC-Verfahren verwendet werden.

Neben dem PAKE-Modus kann OPAQUE auch als erweitertes („augmented“) PAKE verwendet werden. In diesem Fall hat der Kommunikationspartner in der Serverrolle keinen Klartextzugriff auf das Passwort, sondern muss nur auf einen „Verifier“ zugreifen können. Aus dem Verifier kann das Passwort nicht effizient abgeleitet werden, selbst wenn es aus einer dem Angreifer bekannten Liste ausgewählt wurde. Durch Ausführen der Feldelemente in der Serverrolle kann in diesem Fall verhindert werden, dass ein Angreifer das Passwort durch Zugriff auf das Feldelement extrahieren kann. In unserer Implementierung leitet ein separates Hilfsprogramm den Verifier aus RaSTA-IDs und einem Passwort ab. Dieser kann dann in das Feldelement importiert werden. Der Sitzungsschlüssel kann auch periodisch automatisch mithilfe desselben PAKE-Protokolls rotiert werden, was ebenfalls in der PoC-Implementierung demonstriert wird. Dadurch kann das Zeitfenster für Brute-Force-Angriffe auf ein unpraktikables Fenster reduziert werden, und wie bereits erwähnt, kann so ein Replay-Angriff ausgeschlossen werden.

6 Schlussfolgerung

Diese Arbeit demonstriert am Beispiel des RaSTA-Protokolls, dass Sicherheit vor gezielten Angriffen („Security“) nicht mit Sicherheit vor unbeabsichtigten Fehlern („Safety“) gleichgesetzt werden kann.

Das RaSTA-Protokoll erfüllt die Safety-Eigenschaften, die von einem Eisenbahn-Transportprotokoll erwartet werden. Wie praktisch demonstriert und theoretisch hergeleitet wurde, kann ein aktiver Angreifer durch gezielte Manipulation der ausgetauschten Nachrichten allerdings die Sicherheitseigenschaften des Protokolls umgehen und gezielt gefährliche Situationen erzeugen. Dies wird durch grundlegende Spezifikationsschwächen des RaSTA-Protokolls ermöglicht und kann durch fehlerhafte Konfiguration noch weiter erleichtert werden. Insbesondere wurde nachgewiesen, dass die Security der RaSTA-Kommunikation nicht ohne Tunneling oder grundlegende Änderungen im Protokoll gewährleistet werden kann. Konkrete Verbesserungsvorschläge am Protokoll, sowie Vor- und Nachteile von Tunneling, wurden dabei diskutiert.

Wir möchten zum Abschluss darauf hinweisen, dass beim Entwurf von Kommunikationssystemen im Eisenbahnwesen immer Safety, Security und Reliability (Zuverlässigkeit) des Gesamtsystems gemeinsam betrachtet werden sollten. Wie mehrfach in dieser Arbeit angesprochen, hängen diese Eigenschaften oft voneinander ab. Eine generische Konstruktion eines Systems, welches alle diese Eigenschaften erreichen soll, aus Schichten, die jeweils nur eine der Eigenschaften betrachten, ist nach Meinung der Autoren im Allgemeinen nicht realisierbar.

7 Literaturverzeichnis

[1] DKE: Electric signalling systems for railways – Part 200: Safe transmission protocol according to DIN EN 50159 (DIN VDE V 0831-159). In: DIN, Standard DIN VDE V 0831-200, Jun. 2015.

[2] Dobbertin, Hans: Cryptanalysis of MD4. In: Journal of Cryptology, vol. 11, pp. 253-271, 1998. doi: 10.1007/s001459900047

[3] European Commission, Joint Research Centre, Giannopoulos, G., Smith, H., Theocharidou, M.: The landscape of hybrid threats : a conceptual model : public version. Giannopoulos, G. (editor), Smith, H. (editor), Theocharidou, M. (editor), Publications Office, 2021,

[4] Heinrich, Markus; Vieten, Jannik; Arul, Tolga; Katzenbeisser, Stefan: Security Analysis of the RaSTA Safety Protocol. In: 2018 IEEE International Conference on Intelligence and Security Informatics (ISI), Miami, FL, USA, 2018, pp. 199-204, doi: 10.1109/ISI.2018.8587371.

[5] Petitcolas, Fabien: Kerckhoffs Principle. In: Encyclopedia of Cryptography and Security, 2nd Edition, 2011, doi: 10.1007/978-1-4419-5906-5.

[6] Rescorla, Erik; Dierks, Tim: The Transport Layer Security (TLS) Protocol Version 1.2. In: RFC 5246, 2008.

Towards an Evaluation Environment for Digital Interlocking Networking

Frederic Reiter[1], Lukas Iffländer[1], Ulrich Maschek[2], Richard Kahl[2]

[1] *Deutsches Zentrum für Schienenverkehrsforschung beim Eisenbahn-Bundesamt*

[2] *Technische Universität Dresden, Professur für Verkehrssicherungstechnik*

1 Introduction

Interlocking faces a major restructuring. Existing heterogenic systems require massive maintenance structures able to support four different generations of interlocking (mechanical, electromechanical, relay and electronic interlocking) but also various types from various manufacturers for each generation. These types rarely share replacement parts and are usually incompatible with each other. Connections between neighboring interlockings require proprietary adapters. Furthermore, once interlocking for a station is commissioned, replacement parts and modifications can only be acquired from the original supplier creating an undesired supply monopoly.

Germany's incumbent infrastructure manager Deutsche Bahn, therefore, decided to switch to a new type of interlocking, termed digital interlocking. While one might call the decision process in electronic interlocking – and generally all states in interlocking systems – digital, the differences to "classical" electronic interlocking is that field elements are connected to the control unit using standardized data busses, protocols and interfaces. Through this standardization, infrastructure managers can exchange field elements (e.g., light signals, switches, axle-counters) from one manufacturer with elements from a competitor, eliminating dependencies. Furthermore, systems can easily be upgraded or exchanged without replacing the remaining systems – today, when replacing an electronic interlocking it is often necessary to also replace all field elements.

Germany intends to digitize its state-owned infrastructure's interlocking systems by 2040. The German railway industry lobbies for an even more ambitious timeline of 2035. Over a hundred thousand field elements require a migration from classical interlocking to digital interlocking. This goal is ambitious not only when it comes to planning processes but also, when it comes to testing and authorization for the used software.

Current approaches to integration testing do not support the planned implementation speed. Deutsche Bahn creates expensive and complicated test environments for digital interlocking projects like the line section between Mertingen and Meitingen. These environments comprise an exact clone of the interlocking except for the physical field elements. Neither is continuing this approach for a large-scale rollout economically feasible, nor does the necessary amount of personnel exist to attain the desired speed. With our contribution in this paper, we want to support the transition from real-world environments to their virtualized equivalents. Additionally, to the requirement for faster testing, the need for faster interlocking arises. Studies found that the interlocking operation speed has a significant impact on the capacity of stations and routes [1]. Therefore, it is necessary to identify potentials to further shorten round setting times.

In this work, we propose and describe a design for a fully virtualized evaluation environment that reduces complexity while keeping close to reality. To our knowledge, similar research with a focus on digital interlocking systems has not been published with the exception of [8], which examines

[1]Korrespondierende Autoren: reiterf@dzsf.bund.de, ifflaenderl@dzsf.bund.de

interlocking interfaces qualitatively but provides no basis for quantitative evaluation of performance. Our environment allows to perform feature and performance testing for all steps before the integration with the physical field elements. Furthermore, the system is designed to incorporate hardware-in-the-loop testing, allowing to add real-life field elements or interlocking into the environment while keeping the remaining environment virtualized.

Our environment uses virtualized machines to create an exact replica of the rail IP System (bbIP) specified by Deutsche Bahn. Field elements and interlocking communicate using the Rail Safe Transport Application (RaSTA) protocol [9]. As application layer protocol, we use the EULYNX specification that builds upon RaSTA [15]. EULYNX is a European initiative by 14 infrastructure managers to standardize interfaces and elements of the signaling systems.

The proposed implementation reaches prototype level and comprises a simplified interlocking, the network between interlocking and field elements (including a redundant connection as specified for RaSTA), sample field elements (switches and light signals) and monitoring functionality. This functionality allows to monitor and benchmark the entire process of the route setting time at every step to identify which parts of the system take up what amount of time.

We evaluate our system regarding the ability to test and verify the functionality of software components and to measure the performance and the theoretically possible route setting times. For the first goal, we found multiple issues in an existing open source RaSTA implementation and validated the correction inside our environment. For the route setting time, we found, that the protocol aspect is nearly negligible and, in theory, significantly shorter route resolution time and route setting times than required by Deutsche Bahn are possible. This result suggests a requirements reevaluation for further projects.

The remainder of this paper is structured as follows: After this introduction, in Section 2, we introduce technical background on bbIP and modular testing. Section 3 shows our approach to system design and implementation. Afterwards, we evaluate our approach in Section 4 and discuss the results in Section 5. Lastly, Section 6 provides a short summary and an outlook on future work.

2 Background

EULYNX is poised to be the rail industries USB standard with respect to field elements and has been adopted by Deutsche Bahn infrastructure company DB Netz AG as successor to NeuPro – Deutsche Bahn's original standardization proposal [10]. Standardization enables a plethora of benefits, the most pronounced of which are opportunities for small and medium sized companies to enter the market, which makes for cheaper acquisition and maintenance of infrastructure, as well as easier replacement of old parts with the components of a different supplier [6]. EULYNX interfaces are specified in semi-publicly available documents, which determine, next to other things, the contents of the exchanged messages and how and when each component should send which kind of message [15]. It also specifies a safe communication network, which for the DBoperated area, will be the bbIP. Quality-of-Service requirements for the network's latency posed by the EULYNX specification are 50 ms for the Standard Wired Profile.

The bbIP network is made up of three layers, each for a different area of operation. They consist of completely redundant blue and grey network planes. The lowermost layer (access layer) connects the site of operation (TSO), which includes interlocking and Maintenance & Data Management (MDM), to

the field element access point, which in turn connects to the field elements' Object Controllers (OC). The OC themselves come with the actual application module and two security components, a crypto-box and a VPN [3].

As mentioned in the opening, the usage of standardized interfaces and commercial-of-the-shelf (COTS) components for command-and-control poses new challenges and opportunities for the process of testing and authorization. Caspar et al. in [5] previously concluded that the interlockings composition of a multitude of modules which originate from different sources would necessarily lead to a change in validation practices, as the conventionally intended test of the system could only be conducted after it had been built in the field. At this time, the validation effort would lead to pronounced losses in time and money, as well as needless delays in project completion. They propose a multi-phased test procedure, called "modular, hierarchical testing", which schedules integration tests of the test unit with an increasing number of interfacing systems and in increasingly realistic environments. An Evaluation environment like ours represents the tool, which is needed to conduct the tests of the intermediate phases in an approach like the one by Caspar et al.

3 Methods

Structurally, a digital interlocking is an arrangement of computers which feature-specific hard- and software, interconnected by the bbIP network which bases on standard IPv4. The behavior of a digital interlocking can be emulated by realizing a representation of such computers in correspondence to the architecture of the target network. By configuring the environment with the respective hardware capabilities and network Quality of Service (QoS) parameters (e. g. maximum throughput, latency, or transmission failure rate), results of experiments in the emulated scenario can approach real world application accuracy. To authentically model the bbIP environment, we requested and received latency measurement data for the newly built TSO Mertingen– Meitingen from DB Netz AG [4]. Even for higher amounts of bytes per packet (1024 bytes) than usually reached by cryptographically secured and RaSTA-headed EULYNX Protocol Data Interface (PDI) messages, the latency didn't surpass 1 ms. We used the respectively measured latency of 0.562 ms for usual packet sizes of 512 bytes in the simulation.

To support a wide range of different use cases, the proposed evaluation environment is based on modern virtualization techniques. Thus, instead of using real hardware for each single machine, we create virtual machines for every component. Virtual machine hypervisors create a hardware-agnostic environment of virtual processors, memory and drives, on which they allow to run entire operating systems without them realizing they do not run on actual hardware. Inside fully virtualized machines, container virtualization allows to run multiple processes on the same machine without them having knowledge of each other or being able to interfere. While the separation is less than for full virtualization the omission of having multiple instances of the operating system saves a significant amount of resources.

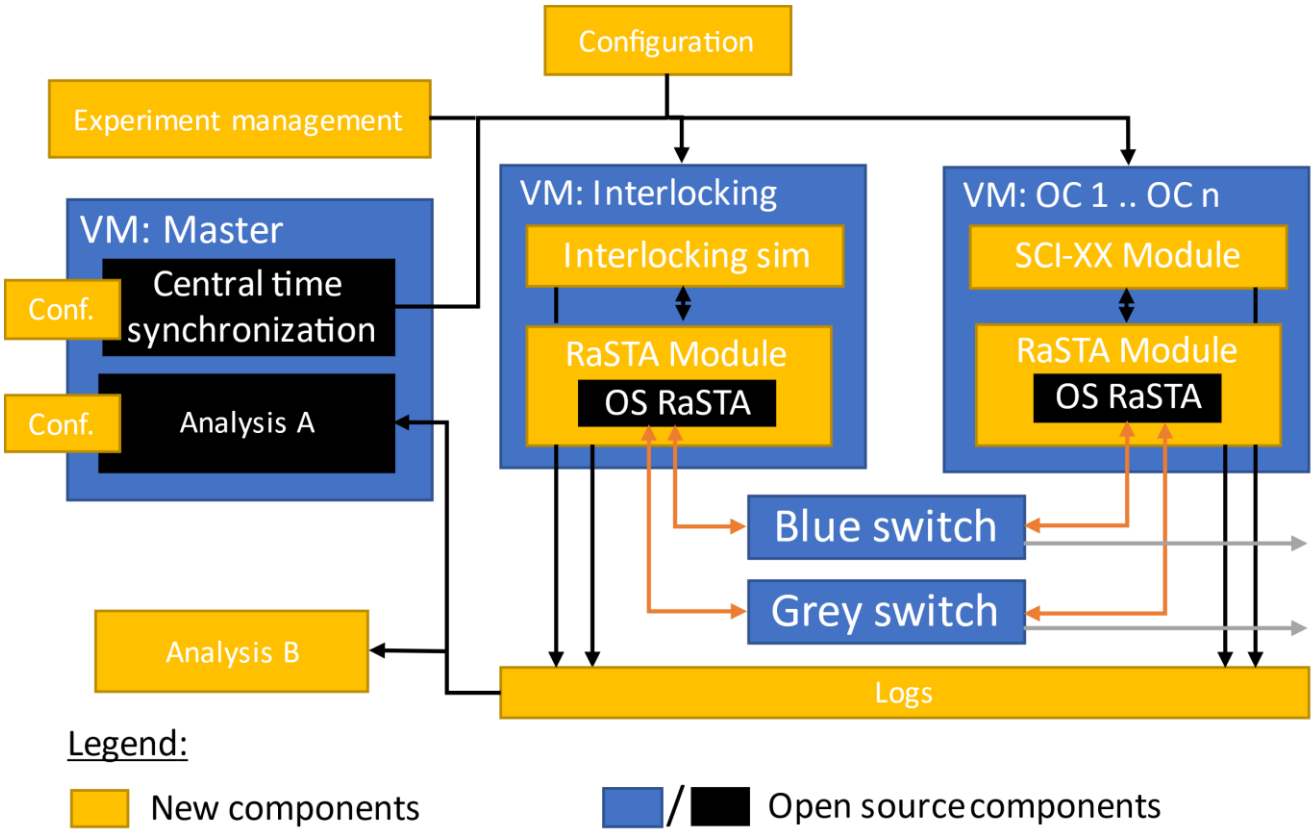


Figure 1: Evaluation environment structure comprising existing (open source) components and specific components for the desired applications. The environment comprises several virtual machines (VMs) to represent the interlocking and the Object Controllers (OC). OC and interlocking each comprise two containers. One container handles the networking using the RaSTA protocol while the other container is responsible for the functional aspects. OC use the varius standard communication interfaces (SCI) specified by Eulynx. Black arrows represent the experiment control communication while orange arrows represent interlocking communication. Grey arrows hint at possible extensions to other components like control stations or radio block centers.

Figure 1 shows the architecture for our environment called DSTW-Sim. Besides fixed virtual machines for the interlocking and an experiment controller, virtualization enables the deployment of an arbitrary number of field elements. Furthermore, the incorporation of physical elements is possible, but not mandatory, allowing for so called hardware-in-the-loop testing where physical objects are validated and evaluated using simulated or virtualized environments.

The architecture of the bbIP infrastructure is reproduced accurately in DSTW-Sim, meaning every object controller, the interlocking, as well as all switches of the redundant network planes are represented by their own virtual machine (VM). To keep the behavior of the test environment as close as possible to the real system, we use hypervisor-based virtualization provided by **VirtualBox** v6.1.38 [16] for each individual VM. Where applicable, container-based virtualization provided by **Docker** v20.10.12 [17] is used inside the VM to attain a modular design and make for an easy replacement of subcomponents [2].

For provisioning, we use the open source tool **Vagrant** v2.3.4[14] which is configured automatically according to the user-defined setup of the experiment. Vagrant is a tool for building complete environments using virtualized machines. Once configured, the entire evaluation environment can be easily provisioned or destroyed without manual intervention.

A newly designed configuration language based on Yet Another Markup Language (YAML) [13] provides a way for the user to specify the setup of any combination of field elements. The elements are given by category and name. We use a script written in **Python** [18] to translate the experiment configuration to a Vagrant configuration, create all experiment specific files, control the experiment and analyze the results.

The OC in DSTW-Sim consists of two modules: The application module is implemented in the memory-safe language **Rust** [19] and processes PDI messages according to the EULYNX interface specification, while the RaSTA module handles the network communication as intended by EULYNX. The two modules communicate locally via User Datagram Protocol (UDP).

The time and content of all messages to be sent and received by the application layer in the process of route setting is implemented to the exact EULYNX requirements specification to set the conditions for realistic measurement results. Specifically, the message processing features of the Standard Communication Interface for Point (SCI-P), Light Signal (SCI-LS) and electronic interlocking were realized. Simplifications in this initial implementation include omitting the cryptographic components of the OC otherwise used in the bbIP architecture and assuming the PDI connection between interlocking and OC to be already established. Additionally, the simulation of the interlocking supports only the setting of one route at a time.

Crucially, the evaluation environment is executable on every computer that runs a standard operating system (Windows, Linux, MacOS) after installing the open source utilities Vagrant, VirtualBox and Docker. This property yields excellent reproducibility of experiments. We realized all implementations solely by using open source components.

Our environment allows us to instrument various experiments, collect logging information and experiment results and analyze these results. As a method of validating our system, we designed test cases for the network protocol RaSTA or, more specifically, it's open source implementation by Railway-CCS [7]. RaSTA features mechanics to detect the untimely delivery of a message, as well as detection and correction of packet loss. To test these features, we use failure injection by the network emulation software **netem** [11] to simulate degraded QoS connectivity profiles.

While not required for the functionality of the system, we implement internal clock synchronization for all participating computers using the Linux kernel-integrated tool **linuxPTP** [12] for the purpose of comparing logged timestamps across the systems in the environment.

4 Results

We evaluated our environment's functionality by setting up a sample railway station with two light signals and two switches (see Figure 2). We measured the performance of the modelled systems and the RaSTA implementation by repeatedly setting the specified routes in the interlocking simulation and logging the exchange of all relevant messages. The necessary steps for route setting according to the EULYNX specification are pictured in Figure 3. These include all steps of the interlocking domain before a movement authority would be passed to the radio block center (RBC). The steps represented by white boxes are not included in the calculation for the route setting time.

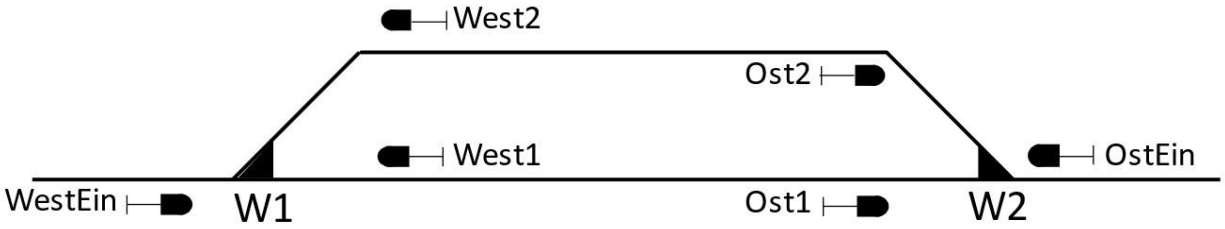


Figure 2: Sample railway station used for this paper comprising six light signals and two points.

EULYNX interaction between all components performed flawlessly. However, we found several issues within the used open source RaSTA implementation. For example, our environment showed implementation errors in RaSTA's check for untimely delivery with respect to the clock independency mechanism that led to unintentionally terminated connections (see Figure 4). Specifically, the system clocks don't have to be synchronized between RaSTA communication partners, because they are never compared with each other. At the time of initialization, the local clocks' time is used as a placeholder value. Erroneously, this initialization happened for every received heartbeat message (HB), making the system race its own clock. The result of this were messages, that appeared to have been sent even before the previous message which had already confirmed the current timestamp (CTS). Please refer to [9] for details on the untimely delivery of messages in RaSTA.

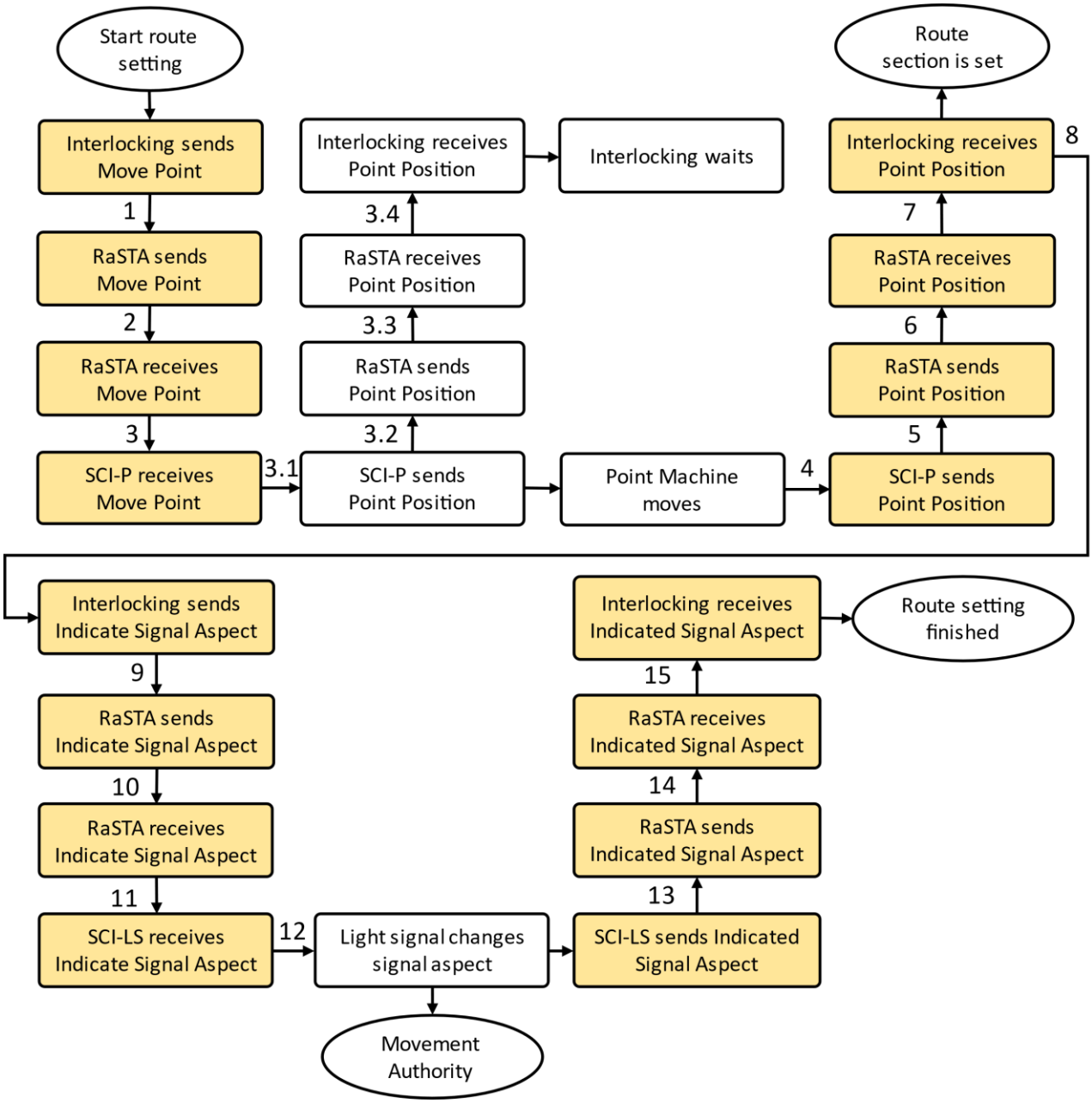


Figure 3: Phases of the route setting process with respect to exchanged EULYNX messages.

Furthermore, the interlocking's RaSTA module was initially unable to connect to more than one field element at the time, because the implementation relied on the field elements establishing connections, while EULYNX specification assigns the connection startup to the interlocking. We fixed both issued and submitted our patch to the original authors. The corresponding pull request [20] was accepted and is now part of the public repository and available to further researchers. These results show that our environment allows us to validate the functionality of protocol implementations in realistic scenarios.

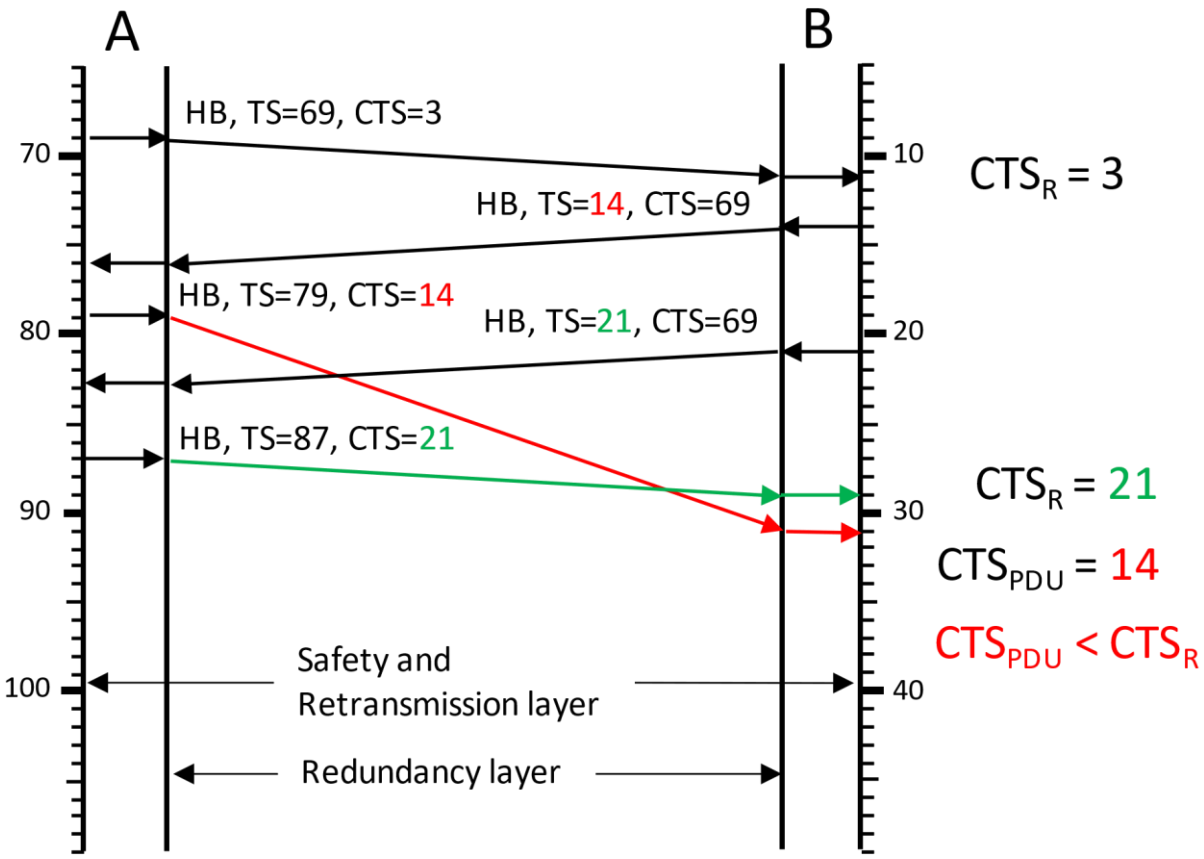


Figure 4: Observed behavior during debugging and the associated messages' relevant data fields. Heartbeat messages (HB) are exchanged, which include a timestamp (TS) and a confirmed timestamp (CTS). The last relevant CTS is stored in CTS_R, to be compared with newly arrived CTS in CTS_{PDU}.

Additionally, we instrumented our environment to measure the time spent in different communication steps, e.g., when moving a point machine to a new end position under different network conditions (see Figure 6). We confirm that the impact of increasing network latency is linear as expected, validating our environment's ability to perform quantitative assessments. Due to more apparent limitations with the used RaSTA implementation, we couldn't test the RaSTA protocols' behavior in networks under influence of packet loss since the loss of packets and subsequent arrival of unexpected packet sequence numbers lead to crashes of the RaSTA module. Addressing these was out of the scope of this paper.

The route setting time was averaged over 100 repetitions with a network parametrized to the measurements of the DB Netz AG and resulted in 26 ms. In our scenario, the point machine move duration is disregarded, meaning the point machine is assumed to instantly arrive at the new end position. Table 1 displays the mean and the standard error of the mean of each steps' duration. The distribution of durations of specifically networkdependent steps is shown in Figure 5.

Tab. 1 Mean and standard error of the mean for individual durations of route setting steps

Step	Mean [ms]	SEM [ms]	Actor	Process	Message
0	-	-	Interlocking	System	Move Point
1	0.637	0.021	RaSTA sends	System	Move Point
2	4.403	0.358	RaSTA receives	Network	Move Point
3	0.040	0.001	SCI-P receives	System	Move Point
3.1	0.555	0.011	SCI-P sends	System	Point Position
3.2	0.948	0.020	RaSTA sends	System	Point Position
3.3	5.445	0.288	RaSTA receives	Network	Point Position
3.4	0.101	0.004	Interlocking receives	System	Point Position
4	-	-	Point Machine moves/ SCI-P sends	System	Point Position
5	1.216	0.286	RaSTA sends	System	Point Position
6	5.669	0.325	RaSTA receives	Network	Point Position
7	0.117	0.002	Interlocking receives	System	Point Position
8	1.108	0.015	Interlocking sends	System	Indicate Signal Aspect
9	0.629	0.008	RaSTA sends	System	Indicate Signal Aspect
10	4.284	0.299	RaSTA receives	Network	Indicate Signal Aspect
11	0.048	0.001	SCI-LS receives	System	Indicate Signal Aspect
12	0.662	0.014	SCI-LS sends	System	Indicated Signal Aspect
13	0.931	0.012	RaSTA sends	System	Indicated Signal Aspect
14	6.155	0.298	RaSTA receives	Network	Indicated Signal Aspect
15	0.101	0.009	Interlocking receives	System	Indicated Signal Aspect

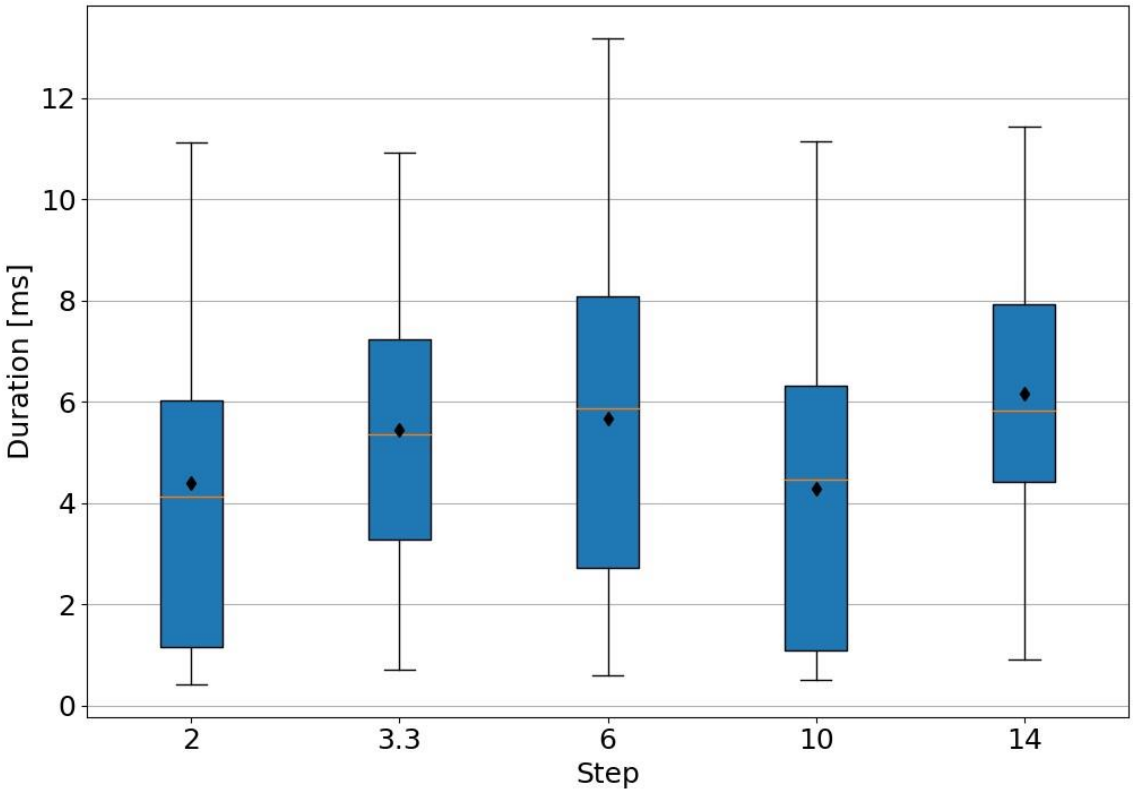


Figure 5: Distribution of durations of network-dependent steps during route setting

We also measured the bandwidth load on our environment's internal network using a tool provided by VirtualBox. The experiment entails the continuous exchange of messages for the sequential setting of two alternating routes and thus represents a worst-case scenario load-wise. During the procedure, we measured the bit rate going through both the blue and the grey switch of the interlocking to be 9.6 kbit/s for the SCI-P connection and the 7.47 kbit/s SCI-LS connection respectively. The difference is due to one additional message during route setting for SCI-P.

Although hardware-in-the-loop functionality is missing from the prototype, the resource monitor of VirtualBox allows for an evaluation of the required hardware capabilities to run the VM with respect to the host system. By extrapolation based on CPU benchmarks we estimate the OC module to comfortably run on a low powered single-board computer like the Raspberry Pi Zero 2 W.

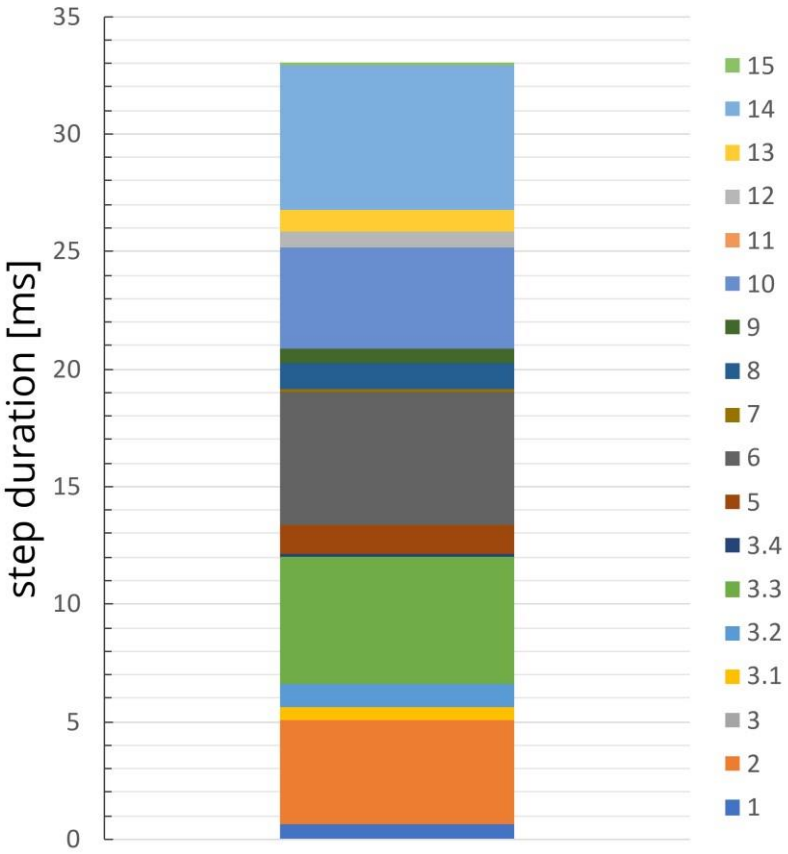


Figure 6: Aggregated route setting duration for all steps

5 Discussion

The test subject for the validation of our evaluation environment, the open source RaSTA implementation [7] in the version published 18th of July 2022 included some breaking bugs. After our changes described above, it managed to perform according to expectations in the limits set for the network latency by EULYNX of 50ms. A possible explanation for the discovered errors in the used RaSTA implementation is that the authors only used environments based on local, container-based virtualization for their integration tests. Environments like the one proposed by us achieve a higher degree of realism and can help to reveal further flaws in IP based command and control systems.

The measured route setting time of 26ms is multiple magnitudes lower than other figures found in literature for DSTW route setting times. Within the scope of the research conducted for implementation of ETCS for the new suburban railway network in and around Stuttgart, suppliers estimated route setting times of 7 to 9 s, disregarding point machine movement durations. This striking difference can only partially be explained by the simplifications made for our prototype. Properties of safe systems, whose omission might lead to a lower measurement of latency, include slower clocks in processors for safety-critical applications and the overhead required for M out of N voting algorithms. Still, the latencies provided by the suppliers at the time most likely relate to guarantees which they are ready to make and do not push the technologies limits. It is up to the infrastructure managers to request and reward the exploitation of the remaining optimization potential.

The experiment setup also allows for an estimation of the maximum amount of field elements connectable to a TSO at a time via one access network in terms of bandwidth. According to the

information provided by DB Netz, the bbIP network will be capable of full-duplex, gigabit speed [4]. In the worst-case scenario of continuously setting new routes, one TSO could serve upwards of 104,000 field elements, excluding cryptographic overhead. Considering the planned outfit of OC with both crypto-box and VPN components, we estimate the traffic to increase by 596% in the worst-case scenario, resulting in 17,477 served field elements. However, we regard the usage of both crypt-box and VPN to be a questionable decision, since RaSTA based on TLS over TCP is already seen as sufficiently secure by the EULYNX specification. In terms of the number of required TSO, we estimate that the approximately 250,000 installed field devices in Germany could be served from between 3 and 15 TSO locations, depending on the scenario. However, additional redundancy might be desirable.

6 Conclusion and Future Work

This work presents a first step towards a fully virtualized evaluation of components and protocols for digital interlocking. Instead of creating a physical replica of the entire network, we create a virtualized replica that allows for simple reconfiguration. We showed that our environment could validate basic OC functionality and perform quantitative assessments.

However, further steps are necessary to perform unit and integration testing inside such an environment. For example, currently, we can only evaluate virtualizable components. Thus, the evaluation of, e.g., object controllers in combination with the controlled field elements is necessary for integration testing. Therefore, we plan to implement the hardware-in-a-loop capability that was accounted for in the infrastructure design to integrate real hardware inside our virtualized system. Furthermore, for now, the interlocking configuration is done manually. We intend to automatically generate the entire digital interlocking configuration files based on EULYNX or PlanPro configuration files.

For now, we use Vagrant on a single machine to provision our environment. Thus, this single machine limits the scale of our environment (for now, the environment is mainly limited by memory capacity). We plan to migrate our environment to a private cloud environment and add functionality to run in public clouds provided by, e.g. Amazon, Google or Microsoft to further simplify development.

The application and impact of the route setting times measured within the scope of our prototypal implementation is limited due to its simplifications, most importantly the missing safety features. An evaluation environment with fewer or no simplifications with respect to the actual interlocking, running on SIL4-certifiable hardware, can demonstrate the practical limits of processing speed for safe rail operation. Which we estimate will not deviate meaningfully from the results achieved in this paper.

7 Bibliography

[1] Marc Behrens, Mirko Caspar, Andreas Distler, Nikolaus Fries, Sascha Hardel, Jan Kressner, Ka-Yan Lau, and R Pensold. 2021. Schnelle Leit- und Sicherungstechnik für mehr Fahrwegkapazität. *EI-Der Eisenbahningenieur* (June 2021), 50–55.

[2] Jossekin Beilharz, Philipp Wiesner, Arne Boockmeyer, Lukas Pirl, Dirk Friedenberger, Florian Brokhausen, Ilja Behnke, Andreas Polze, and Lauritz Thamsen. 2021. Continuously Testing Distributed IoT Systems: An Overview of the State of the Art. Retrieved from http://arxiv.org/pdf/2112.09580v1

[3] Dr. Bernd Elsweiler, Nikolaus Fries, Christian Summen, and EBA Herausgeber. 2021. EBA-Infotage: Einordnung und Übersicht über das Programm Digitale Schiene Deutschland (DSD).

[4] Petros Matios. 2022. Persönliche Kommunikation.

[5] Mirko Caspar, Hardi Hungar, and Daniel Schwencke. Effizientes Testen modularisierter und standardisierter Stellwerkskomponenten. *SIGNAL+DRAHT* 110, 9/2018.

[6] Silvia Pascual and Frits Makkinga. Reduzierung der Betriebskosten für ERTMS Level 2 durch Implementierung einer standardisierten Stellwerksschnittstelle. *SIGNAL+DRAHT (109) 9/2017*, 8.

[7] Railway-Ccs. 2023. rasta-protocol. *GitHub*. Retrieved from https://github.com/Railway-CCS/rasta-protocol

[8] Robert Schmid, Arne Boockmeyer, Lukas Pirl, Randolf Berglehner, Ibtihel Cherif, Andreas Korff, Bernd Elsweiler, and Andreas Polze. 2021. EULYNX-Live: Eine Methodik zum Validieren von

[9] Systemspezifikationen in hybriden Feldtests. *Signal+Draht* 6, 113 (2021), 24–31.

[10] 2015. *DIN VDE V 0831-200: Elektrische Bahn-Signalanlagen – Teil 200: Sicheres Übertragungsprotokoll RaSTA nach DIN EN 50159 (VDE 0831-159)*. Deutsche Kommission Elektrotechnik Elektronik Informationstechnik in DIN und VDE.

[11] 2017. Report on Industry Workshop held 25th January 2017. Retrieved from

https://eulynx.eu/index.php/documents/presentations-given/39-20170125-workshop-industryreport/file

[12] 2017. NetEm - Network Emulator at Linux.org. Retrieved from https://www.linux.org/docs/man8/tcnetem.html

[13] 2019. The Linux PTP Project. Retrieved from https://linuxptp.sourceforge.net

[14] 2021. The Official YAML Web Site. Retrieved from https://yaml.org

[15] 2023. Vagrant by HashiCorp. *Vagrant by HashiCorp*. Retrieved from https://www.vagrantup.com

[16] 2023. EULYNX documents. Retrieved from https://eulynx.eu/index.php/documents/publisheddocuments

[17] 2023. Oracle VM VirtualBox. Retrieved from https://www.virtualbox.org

[18] 2023. Docker: Accelerated, Containerized Application Development. *Docker*. Retrieved from https://www.docker.com

[19] 2023. The official home of the Python Programming Language. *Python*. Retrieved from https://www.python.org

[20] 2023. Rust Programming Language. Retrieved from https://www.rust-lang.org

[21] FixCTS_R bug SgtChrome/rasta-protocol@24ce71d. Retrieved November 24, 2022 from https://github.com/SgtChrome/rastaprotocol/commit/24ce71d5a24a32bb906a04b7a8101b04b5fa236c

Innovative Qualifizierungsstrategien für modulare Technologien: Sicherheit und Effizienz in der Bahninfrastruktur

Julian Lucas[1] und Markus Rothkehl[2]

[1] Lehrstuhl für Bahnsysteme und Bahntechnik, TU Darmstadt

[2] DB Netz

Abstract

Der vorliegende Beitrag stellt ein Konzept für die Sicherheitsnachweisführung und Genehmigung in modularen Systemarchitekturen im Bereich der Digitalen Schiene Deutschland vor. Dabei werden ein Sicherheitsnachweisverfahren, ein Genehmigungsverfahren und ein Systemqualifizierungsplan (SQP) sowie die Logik verschiedener Teilqualifizierungspläne (TQP) entwickelt, die den Anforderungen an die neu entstandene Rolle des Integrators gerecht werden. Dies ermöglicht die sichere Integration von Komponenten mehrerer Hersteller über standardisierte Schnittstellen und die Anpassung der Alttechnik an die DSTW/Neupro-Architektur. Die modulare und standardisierte Systemarchitektur, die von der DB Netz vorgegeben wird, spielt eine entscheidende Rolle für die Sicherheitsnachweispflicht nach den Vorgaben der EN 50126 ff.

Insgesamt trägt diese Arbeit durch die Konzepte zur Nachweisführung und Genehmigung dazu bei, die Sicherheit, Zuverlässigkeit und Wirtschaftlichkeit des Eisenbahnverkehrs in Deutschland zu erhöhen.

1 Einleitung

Eine zuverlässige Bahninfrastruktur ist von zentraler Bedeutung für einen planmäßigen Eisenbahnverkehr. Infolgedessen ist eine Modernisierung der Infrastruktur notwendig, da viele Strecken noch mit veralteten Leit- und Sicherungstechnologien ausgestattet sind. Insbesondere der von der EU angestrebte grenzübergreifende Schienenverkehr mittels ETCS kann nur durch den Einsatz elektronischer und digitaler Stellwerkstechnik ermöglicht werden. Die Herausforderung besteht darin, die neue Systemarchitektur der digitalen Leit- und Sicherungstechnik (DLST) so zu gestalten, dass sie einerseits technische Standards vereinheitlicht und andererseits den Ausbau beschleunigt. Dabei stellt sich die Frage, wie Sicherheitsnachweise in modularen Systemarchitekturen effizient und effektiv geführt werden können, um die Umsetzung sicherheitskritischer Bahninfrastrukturprojekte zu beschleunigen. Der vorliegende Beitrag erörtert daher verschiedene Ansätze und Verfahren, die zur Beschleunigung der Sicherheitsnachweisführung beitragen können. Die gesamten Digitalisierungsmaßnahmen der DB werden unter der Konzernstrategie "Digitale Schiene Deutschland" gebündelt, in deren Rahmen auch die DLST gefördert wird [1].

[1] Julian Lucas: lucas@verkehr.tu-darmstadt.de

[2] Markus Rothkehl: Markus.Rothkehl@deutschebahn.com

2 Herausforderungen bei der Sicherheitsnachweisführung und Genehmigung

2.1 Problemstellung

ESTW und die DSTW Vorläuferprojekte sind bisher als monolithische Systeme von einem Hersteller geliefert worden. Dies bedeutet, dass das gesamte System inklusive Teilsysteme von nur einem Hersteller geplant und geliefert wird. Bisherige Zulassungsverfahren sind dementsprechend auf diese ein Hersteller Systeme ausgelegt, der die Verantwortung für die Sicherheitsnachweisführung trägt. Das monolithische Verfahren ist aufgrund der Modularisierung der DLST jedoch nicht mehr anwendbar, da es nicht mehr nur noch einen Hersteller gibt, der das Gesamtsystem liefert. Die jeweiligen Hersteller erbringen nur noch Nachweise für das von Ihnen gelieferte Teilsystem und Komponenten. Hieraus entsteht eine Lücke in der Integration der Nachweisführung zum Gesamtsystem und dessen Verantwortlichkeit. Die Modularisierung und generische Betrachtung ermöglicht es jedoch, die Effizienz der Zulassungsverfahren zu erhöhen und somit dessen Zeitaufwand über lange Sicht zu verringern, da Teile der Nachweise wiederverwendet werden können.

2.2 Ziel

Ziel der neuen modularen Systemarchitektur ist es, die Digitalisierung der Leit- und Sicherungstechnik (DLST) des deutschen Schienennetzes zu beschleunigen. Damit sollen die Kapazität und die europäische Interoperabilität des Schienennetzes erhöht werden. Des Weiteren soll durch den Einsatz der DLST die Herstellervielfalt gefördert werden, um den Wettbewerb zu erhöhen und damit die Wirtschaftlichkeit der infrastrukturseitigen LST in Zukunft zu verbessern.

Für eine erfolgreiche Zulassung der modularen Systemarchitektur ist eine Systematisierung der Erstellung des Sicherheitsnachweises nach EN 50129 erforderlich [5]. Des Weiteren müssen die Legitimations- und Genehmigungsprozesse auf die modulare Systemarchitektur angepasst werden. Daraus ergibt sich die Notwendigkeit der Systemqualifizierung, welche sowohl eine effizientere Sicherheitsnachweisführung ermöglicht als auch die beschleunigte Umsetzung von Projekten.

Ziel der Systemqualifizierung ist die Beschreibung des Verfahrens zur Erstellung des Sicherheitsnachweises nach EN 50129 [5]. Erforderliche Legitimations- und Genehmigungsprozesse des Gesamtsystems von Hersteller und Betreiber werden damit systematisiert sowie klare Verantwortlichkeiten zugeteilt.

2.3 Anforderungen an das Nachweis- und Genehmigungskonzept

Die Anforderungen an die Systemarchitektur und den Nachweisprozess ergeben sich aus den CENELEC Normen DIN EN 50126, DIN EN 50128 und DIN EN 50129 [3-5]. Die Normen beschreiben die Sicherheitsanforderungen im Entwicklungsprozess und geben die erforderlichen Nachweise an. Für eine Wiederverwendung der Nachweise und somit eine effiziente Nachweisführung, muss das Verfahren vom EBA anerkannt werden. Die Umsetzbarkeit der Verfahren in der Praxis ist eine notwendige Voraussetzung für die Realisierung von Kosteneinsparungen und einer Beschleunigung des Roll-Outs.

Eine allgemeine Anforderung der DB an den Nachweisprozess ist, dass Aktivitäten und Prozesse für alle Beteiligten der Branche möglichst transparent dargestellt werden und die zu definierenden Verfahren von der Bauartbetreuung des Betreibers durch technische und formale Beherrschung des Systems unterstützt werden. Außerdem sollen bestehende Umsysteme (z.B. Bahnübergangsicherungsanlagen)

nicht komplett generalsaniert werden, sondern in die DLST integriert werden. Die Anforderungen sind in Tabelle 1 zusammengefasst.

Tabelle 1: Anforderung an die Systemarchitektur und den Nachweisprozess der DLST

	ANFORDERUNG:
1.	**Einhaltung der CENELC Normen: DIN EN 50126, DIN EN 50128 und DIN EN 50129**
2.	**Anerkennung des Prozesses durch das EBA**
3.	**Wiederverwendung von Nachweisen (Anerkennung der System-GIuV)**
4.	**Umsetzbarkeit in der Praxis**
5.	**Transparente Prozesse und Aktivitäten für Herstellerfirmen**
6.	**Migration von Umsystemen nach MÜ8004 (Anbindung bereits existierender Umsysteme möglich)**

2.4 Methode und Vorgehensweise

Die Entwicklung des gesamten Nachweisführungsverfahrens basiert auf einer Inhaltsanalyse von Dokumenten und Quellen, wie Normen, Verordnungen, Gesetzen und Experteninterviews. Diskussionen mit Fachexperten ermöglichen eine kontinuierliche Überprüfung und Anpassung der Vorgehensweise. Durch agiles Projektmanagement können auf sich ändernde oder zusätzliche Anforderungen flexibel reagiert werden.

Die Vorgehensweise ist durch die erste Anforderung (Einhaltung der CENELEC Normen) vorgegeben. Nach EN 50129 besteht die Sicherheitsnachweisführung für ein System aus dem generischen Anwendungssicherheitsnachweis (GASC) und dem anwendungsspezifischen Sicherheitsnachweis (SASC) [5]. Da die Anwendung als Stellwerk gleichbleibt, wird sich der SASC bei DSTW-Systemen strukturell nicht von älteren Bauformen unterscheiden. Für die Anwendung des SASC kann daher auf die bestehenden Verfahren zurückgegriffen werden. Im Fall des GASC werden Änderungen notwendig. Da die Systemarchitektur vom Betreiber vorgegeben wird und mehrere Hersteller am Projekt mitwirken können, hat die DB als Betreiber eine Integrationsverantwortung der nach Ihrer Definition gelieferten technischen Systeme. Die Systemqualifizierung beschreibt demnach den veränderten Prozess zum Erlangen des GASC und derer dazugehörigen Genehmigung des generischen DSTW-Systems.

3 Die Systemarchitektur der digitalen Leit- und Sicherungstechnik

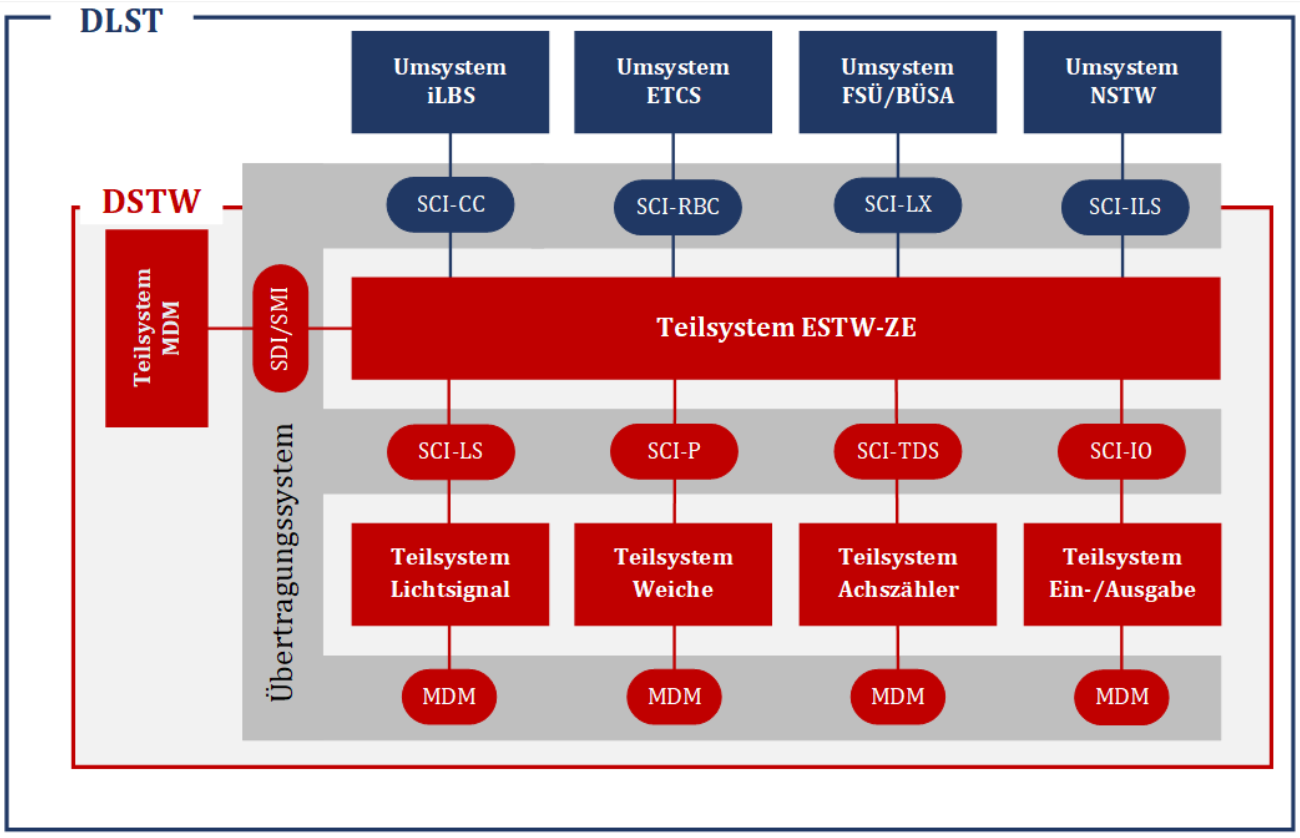
3.1 Einführung in die DLST-Architektur

Im Gegensatz zu früheren Leit- und Sicherungstechnik (LST) Projekten, bei denen hauptsächlich betriebliche Anforderungen an die Hersteller gestellt wurden, stellt die digitale Leit- und Sicherungstechnik (DLST) eine von der DB Netz AG vorgegebene technische Systemarchitektur dar. Diese Architektur soll von einem oder mehreren Herstellern implementiert werden. Die zentrale Komponente der DLST ist das Digitale Stellwerk (DSTW).

3.2 Komponenten und Struktur der DLST-Architektur

Abbildung 1 veranschaulicht die spezifizierte DLST-Systemarchitektur, wobei der Fokus auf die wesentlichen Komponenten und deren Zusammenspiel aus der Perspektive des DSTW gelegt wird. Die Architektur gliedert sich in Teilsysteme (rot) und Umsysteme (blau), die über standardisierte

Schnittstellen (SCI) miteinander verbunden sind. Die Kommunikation zwischen den Systemen erfolgt über das in grau dargestellte Übertragungssystem (ÜS). Die standardisierten Schnittstellen (SCI) spielen eine entscheidende Rolle, da sie eine nahtlose Integration und Interoperabilität zwischen den Systemen unterschiedlicher Hersteller ermöglichen. Die Spezifikationen der zu übertragenden Daten werden genau festgelegt, um eine reibungslose Kommunikation zwischen den Systemen sicherzustellen.



Legende

NSTW, Nachbarstellwerk
FSÜ, Fahrstraßenüberwachung
BÜSA, Bahnübergangssicherungsanlage
ETCS, European Train Control System
MDM, Maintenance Data Management
iLBS, integriertes Leit- und Bediensystem
RBC, Radio Block Center
ZE, Zentraleinheit
ILS, Interlocking System
IO, Input/Output

ESTW, Elektronisches Stellwerk
SCI, Standard Communication Interface
SDI, Standard Diagnostic Interface
SMI, Standard Maintenance Interface
CC, Command Control
LX, Level Crossing
LS, Lichtsignal
P, Weiche
TDS, Train Detection System

Abbildung 1: Schematische Systemarchitektur DSTW. Aus [2]

3.3 Potenziale und Herausforderungen der DLST-Architektur

Die vorgegebene technische Systemarchitektur der DLST bietet verschiedene Potenziale. Sie ermöglicht eine größere Flexibilität und Modularität, da verschiedene Teilsysteme und Umsysteme von unterschiedlichen Herstellern entwickelt und implementiert werden können. Dies führt zu einer erhöhten Wettbewerbsfähigkeit und verbesserten Wirtschaftlichkeit der Infrastrukturseitigen LST. Die modulare Architektur stellt auch eine Chance dar, den Digitalisierungsprozess der Leit- und Sicherungstechnik des deutschen Schienennetzes zu beschleunigen, die Kapazität und die europäische Interoperabilität des Schienennetzes zu erhöhen, sowie die Herstellervielfalt zu fördern. Jedoch birgt die Einführung einer solchen Architektur auch Herausforderungen, die über die in Kapitel 2 bereits diskutierten Nachweis- und Genehmigungsaspekte hinausgehen. Eine der zentralen Herausforderungen besteht darin, eine effektive Kommunikation und Zusammenarbeit zwischen den beteiligten Herstellern

und dem Betreiber zu gewährleisten. Es ist wichtig, dass alle Beteiligten über klar definierte Rollen und Verantwortlichkeiten verfügen und transparente Prozesse etabliert werden.

Zusammenfassend betrachtet, stellt die DLST-Architektur eine bedeutende Veränderung im Bereich der Leit- und Sicherungstechnik dar, die neue Potenziale und Herausforderungen mit sich bringt. Eine sorgfältige und systematische Analyse der Architektur und ihrer Komponenten ist notwendig, um die Vorteile der DLST zu nutzen und die damit verbundenen Herausforderungen erfolgreich zu bewältigen.

4 Systemqualifizierung des DSTW

4.1 Struktur des Systemqualifizierungsplans

Das DSTW zeichnet sich dadurch aus, dass es auf Systemebene keine eigene Hard- oder Software besitzt. Stattdessen ergeben sich die Funktionen aus der Kombination der einzelnen Teilsysteme [6]. Diese Teilsysteme verfügen bereits über Sicherheitsnachweise und Genehmigungen zum Inverkehrbringen und Verwenden (GIuV), wovon das DSTW profitieren kann. Um die Aktivitäten von Herstellern, Integratoren und Systembetreibern voneinander abzugrenzen, wird der Sicherheitsnachweis des DSTW als Systemqualifizierungsnachweis (SQN) bezeichnet.

Der SQN wird innerhalb eines Systemqualifizierungsplans (SQP) festgelegt, der die Verantwortlichkeiten aller beteiligten Akteure klar verteilt. Sowohl der SQP als auch der SQN basieren auf der Struktur der Norm EN 50129. Im SQP sind die verschiedenen Prüfaspekte aufgeführt, wobei für jeden Aspekt die Verantwortlichkeiten hinsichtlich Umsetzung und Verifizierung tabellarisch festgelegt sind. Eine Teilansicht von Kapitel 4.2.5 im SQP ist in Abbildung 2 dargestellt.

4.2.5 Nachweis der korrekten Funktionalität		**Erstellung**	**Prüfung**
Inhalt [EN 50129, Kapitel 7.2, Abschnitt 2.5 und 2.6]	Der Nachweis der korrekten Hard- und Softwarefunktionalität ist überwiegend nur relevant, wenn für die Systemebene eine eigene Hard- und Software erzeugt wird. Ansonsten erfolgt diese Prüfung beim Hersteller des Teil- oder Umsystems. Da das DSTW-System keine Hard- oder Software besitzt, wird der Fokus auf den Nachweis der korrekten Funktionalität des Gesamtsystems auf Grundlage der Teilsystem-funktionalitäten gelegt. Dies beinhaltet auch Betrachtungen auf Systemebene von abgestuften Funktions-einschränkungen der Teilsysteme.		
Anmerkung	Nachweis durch Teilsystemtests, Integrationstests, Systemtests und Sicherheitserprobung		
Anforderungen	▪ Lastenhefte DSTW ▪ Systemintegrationsplan ▪ Systemvalidierungsplan	Betreiber	FGV PSV Betreiber
Referenzen	▪ Gutachten / Prüfbescheinigungen der Teilsysteme ▪ Bestätigungen der technischen Integrierfähigkeit (Teilsysteme)	Hersteller	PSV Hersteller
	▪ Systemintegrationsbericht ▪ Sicherheitserprobungsbericht ▪ Konformitätsbewertung der Teilsysteme	Betreiber	PSV Betreiber FGV
Nachweis	▪ Systemvalidierungsbericht	Betreiber	PSV Betreiber

Abbildung 2: Auszug aus SQP: Kapitel 4.2.5. – Nachweis der korrekten Funktionalität. Aus [2]

Die Validierung des DSTW kann durch verschiedene Methoden erfolgen, wie beispielsweise Funktionsnachweise im Labor, im Feld oder durch gezielte Erprobungen. Die Validierung der

Teilsysteme orientiert sich an den Anforderungen der Norm EN 50126 und der Eisenbahn-Inbetriebnahmegenehmigungsverordnung (EIGV). Da das DSTW und seine Teilsysteme im Sicherheitsintegritätslevel (SIL) 4 eingestuft sind, müssen die Hersteller bei ihrer Nachweisführung die entsprechenden Anforderungen von SIL 4-Systemen einhalten [3].

Der SQN leitet sich hauptsächlich aus den Qualifizierungsergebnissen der Teil- und Umsysteme ab und wird durch integrative Betrachtungen auf System-Ebene ergänzt. Voraussetzung hierfür ist die Umsetzung des Lebenszyklusmodells gemäß EN 50126. Die Vorlegitimierung der Teilsysteme trägt zu einer effektiven Sicherheitsnachweisführung für das DSTW bei. Durch die klare Definition von SQP und SQN werden die Verantwortlichkeiten der verschiedenen Akteure abgegrenzt und die Einhaltung gemeinsamer Standards gewährleistet. Die Verwendung von Funktionsnachweisen und die Beachtung der SIL 4-Anforderungen stellen sicher, dass das DSTW höchsten Sicherheitsanforderungen genügt.

4.2 Sicherheitsnachweisführung

Bei der Erstellung einer Sicherheitsnachweisstruktur gemäß EN 50126 ff. ist es wichtig, normative Prozesse und Nachweisverfahren zu berücksichtigen, die bei der Entwicklung der Teilsysteme angewendet werden. Dies ermöglicht eine geeignete Auswahl, abhängig von den Anforderungen des Herstellers und den Vorgaben im Entwicklungsprozess. Es wird angenommen, dass die sicherheitsrelevanten und nicht sicherheitsrelevanten Teilsysteme gemäß EN 50126 ff. und EIGV entwickelt und genehmigt werden. Da DSTW-Teilsysteme neu entwickelt oder erweitert werden, wird die Mü 8004 bei der Sicherheitsnachweisführung nicht mehr berücksichtigt. Die Systemintegration basiert auf nach EN 50126 ff. bewerteten Teilsystemen, wobei die Zentraleinheit die meisten Nachweisaspekte für die Systemintegration liefert.

Da auf DSTW-Systemebene keine separate Hard- und Software vorhanden ist, leiten sich die Nachweise des Systemqualifizierungsnachweises (SQN) aus den Qualifizierungsergebnissen der Teil- und Umsysteme ab. Voraussetzung dafür ist die Umsetzung des Lebenszyklusmodells gemäß EN 50126. Da das DSTW und seine Teilsysteme im Sicherheitsintegritätslevel (SIL) 4 eingestuft sind, müssen die Hersteller die entsprechenden Anforderungen von SIL 4-Systemen einhalten. Im Rahmen der Teilsystemqualifizierung werden die Anforderungen an das Teilsystem verifiziert und bewertet. Bei Nachweisdefiziten müssen diese auf Systemebene weiterbearbeitet werden. Sicherheitsbezogene Anwendungsvorschriften (SAV) werden vom Hersteller an das übergeordnete DSTW-System weitergegeben. Änderungen an Hard- und Software müssen dokumentiert werden, um Regressionen zu vermeiden. Diese Änderungen werden in die Auswirkungsanalyse aufgenommen, die in den Release Notes dokumentiert ist.

Die Verifizierung und Validierung der Nachweise ist ein wichtiger Bestandteil der Sicherheitsnachweisführung. Dafür müssen geeignete Test- und Bewertungsmethoden angewendet werden. Im Falle von DSTW erfolgt dies auf Teilsystemebene und auf Systemebene. Die Nachweisführung ist nicht nur auf die Entwicklungsphase beschränkt, sondern muss während des gesamten Lebenszyklus des Systems aufrechterhalten werden. Dies erfordert geeignete Maßnahmen zur Instandhaltung und Überwachung des Systems.

Zusammenfassend ist die Sicherheitsnachweisführung ein kontinuierlicher Prozess, der während des gesamten Lebenszyklus des Systems aufrechterhalten werden muss und eine sorgfältige Planung, Umsetzung und Überwachung erfordert, um einen sicheren und zuverlässigen Betrieb zu gewährleisten.

5 Teil-Qualifizierung von Teil- und Umsystemen

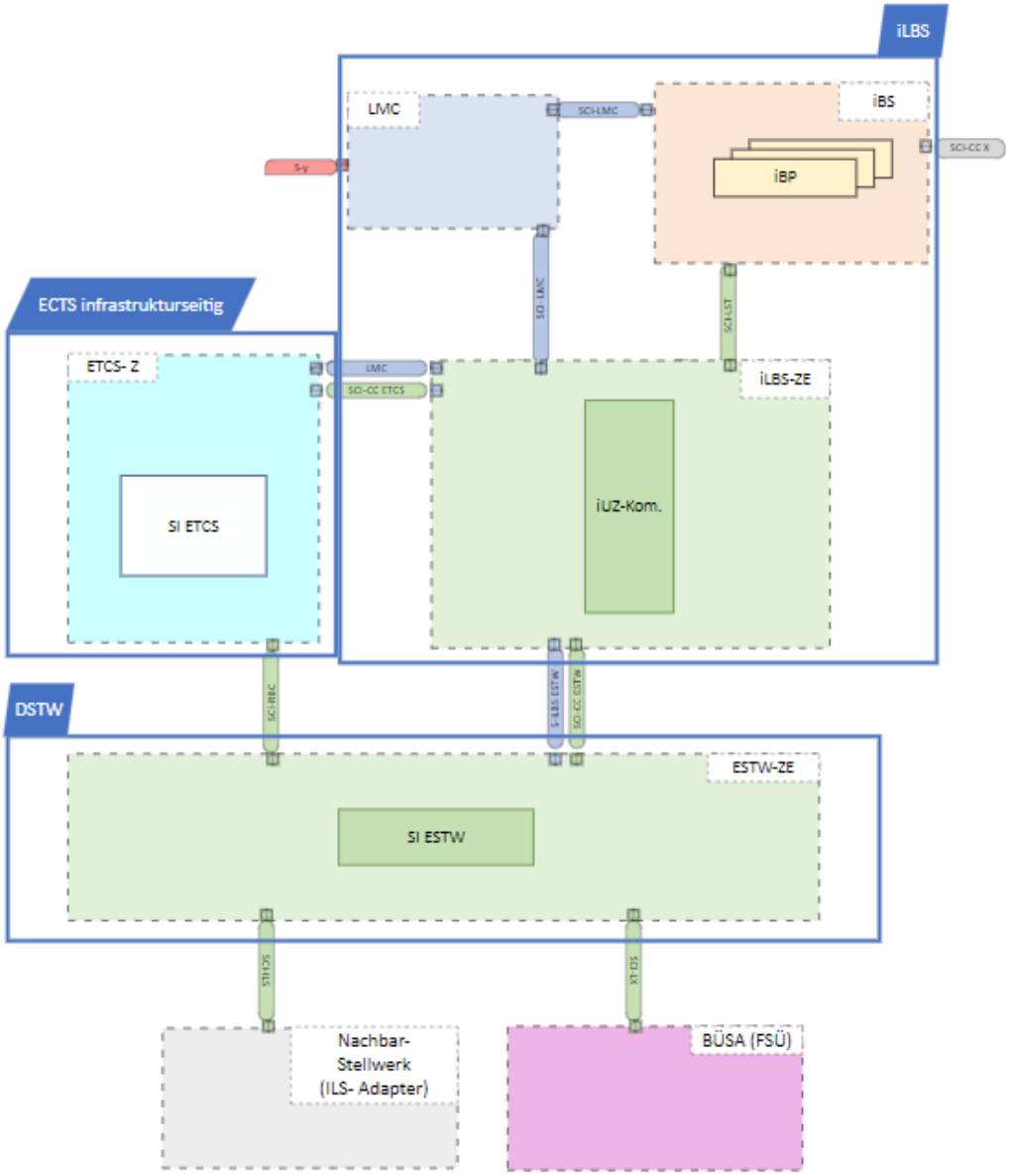
Das vorliegende Kapitel erläutert die Prozesse zur Qualifizierung von eigenständigen Systemen der DLST sowie die Integration von bereits qualifizierten Systemen und Komponenten mit dem DSTW. Es werden alle Systeme der DLST berücksichtigt, die an einer Zugfahrt beteiligt sein müssen, um diese sicher durchführen zu können. Zusätzlich des ÜS als Teilsystem werden alle nicht dem Stellwerk zugeordneten Systeme der DLST (Umsysteme) im sogenannten Teil-Qualifizierungsplan bzw. den daraus entstehenden Teil-Qualifizierungsnachweisen qualifiziert. Tabelle 2 gibt eine Auflistung der zu qualifizierenden Teil- und Umsysteme.

Tabelle 2: Aufteilung der in TQP zu qualifizierenden Systemen nach Teil- und Umsystem

Zu Qualifizierende Teil – und Umsysteme:	**Art des Systems:**
Übertragungssystem (ÜS)	Teilsystem
European Train Control System (ETCS)	Umsystem
Fahrstraßenüberwachung/ Bahnübergangsicherungsanlagen (FSÜ/ BÜSA)	Umsystem
Nachbarstellwerk (NSTW)	Umsystem
Integriertes Leit- und Bediensystem (iLBS)	Umsystem

5.1 Zu qualifizierende Umsysteme des DSTW

Die Verschaltung und Schnittstellen der Umsysteme und des DSTW sind in Abbildung 3 zusammengefasst. Standardisierte Schnittstellen sind mit der Vorsilbe SCI gekennzeichnet. Unterschiedliche Farbgestaltung klassifiziert die Unabhängigkeit der Systeme, sodass für jede Farbmarkierung eine andere Herstellerfirma beteiligt sein könnte. Es ergibt sich, dass derzeit nur die ESTW-ZE und die iLBS-ZE vom selben Hersteller kommen muss, da es zusätzlich zur standardisierten SCI-CC eine herstellerabhängige proprietäre Schnittstelle gibt (S-iLBS ESTW). Dies ist von den derzeitigen Herstellern des DSTW gewünscht worden, um die Lastenheftanforderungen der Schnittstellenfunktionen zu erfüllen. Das Übertragungssystem ist in Abbildung 3 nicht zu sehen, da es Teilsystem des DSTW ist und das DSTW mit den Umsystemen technisch miteinander verbindet.



Abkürzungen:

BÜSA - Bahnübergangsicherungsanlage
FSÜ - Fahrstraßenüberwachung
iBP – integrierter Bedienplatz
iBS – integriertes Bediensystem
ILS – Interlocking System
iUZ – integrierte Unterzentrale
LMC – Leit- und Sicherungstechnik Management Center
LS – Level Crossing
RBC – Radio Block Center

S – Schnittstelle (nicht näher spezifiziert)
SCI – Standard Communication Interface
SI – Service Interface
Z - Zentrale
ZE – Zentraleinheit

Abbildung 3: Grobarchitektur der DLST: Schnittstellen des DSTW sowie den betrachteten Umsystemen. Aus [7]

5.2 Struktur eines Teil-Qualifizierungsplans am Beispiel des iLBS

Bisher existieren Strukturen sowie Prüfschritte für den TQP ÜS sowie den TQP iLBS. Die Qualifizierung erfolgt in beiden Fällen über einen mehrstufigen Prozess. Da der TQP ÜS als Teilsystem des DSTW eng an den SQP gegliedert ist, wird die Vorgehensweise für die Teil-Qualifizierung der Umsysteme vom TQP iLBS aus [7] abgeleitet. Die verallgemeinerten Kurzinhalte und fünf Qualifizierungsschritte sind in Abbildung 4 ersichtlich. In jedem der fünf Schritte entstehen auf das Umsystem angepasste Nachweise. Die Inhalte der Nachweise sind darin in Abbildung 4 stichpunktartig aufgezählt. Im eigentlichen TQN sind bzw. werden die geforderten Nachweise tabellarisch detaillierter definiert. Die Teilschritte sind so definiert, dass die Integration im Verlauf zunimmt. So können Probleme auf Einzelsystemebene auch erstmal dort gelöst werden, bevor die Komplexität der Integration die Lösungsfindung erschwert. Je dunkler die Hintergrundfarbe in Abbildung 4, umso mehr Systeme sind am Nachweisprozess beteiligt. Schritt eins ist alleinige Aufgabe des Lieferanten. Dieser muss nachweisen, dass das von ihm gelieferte Umsystem den Anforderungen aus den CENELC Normen sowie Pflichtenheften genügt. Bei Bestandsystemen muss in diesem Schritt geprüft werden, ob die Gefährdungsraten des Bestandsystems die Anforderungen der aktuellen Normen einhalten. Ebenso muss die Integrationsfähigkeit in die DLST bestätigt werden. Die Schritte zwei und drei beschäftigen sich mit der spezifischen Interaktion des nachzuweisenden Umsystems mit dem DSTW. Es wird zwischen Funktionen mit Sicherheitsbezug und Funktionen ohne Sicherheitsbezug unterschieden, um die Anzahl der Nachweise so gering als möglich zu halten. Gibt es Abweichungen bei der Umsetzung des Pflichtenhefts zum Lastenheft bei der Umsetzung von Funktionen, kann es sein, dass durch den Hersteller oder im Laufe der ersten drei Nachweisschritte sicherheitsbezogene Anwendungsvorschriften eingebracht werden. Diese werden im Schritt vier behandelt und technisch (durch den Integrator DB) oder betrieblich (durch den Betreiber) aufgelöst.

In Schritt fünf erfolgt die Validierung über alle Umsysteme hinweg inklusive der Integration in das DSTW. Das heißt, dass das komplette System gemeinsam getestet wird. Es müssen demnach auch schon das DSTW inklusive aller angeschlossenen Umsysteme (iLBS, ETCS, NSTW, BÜSA) betriebsbereit sein.

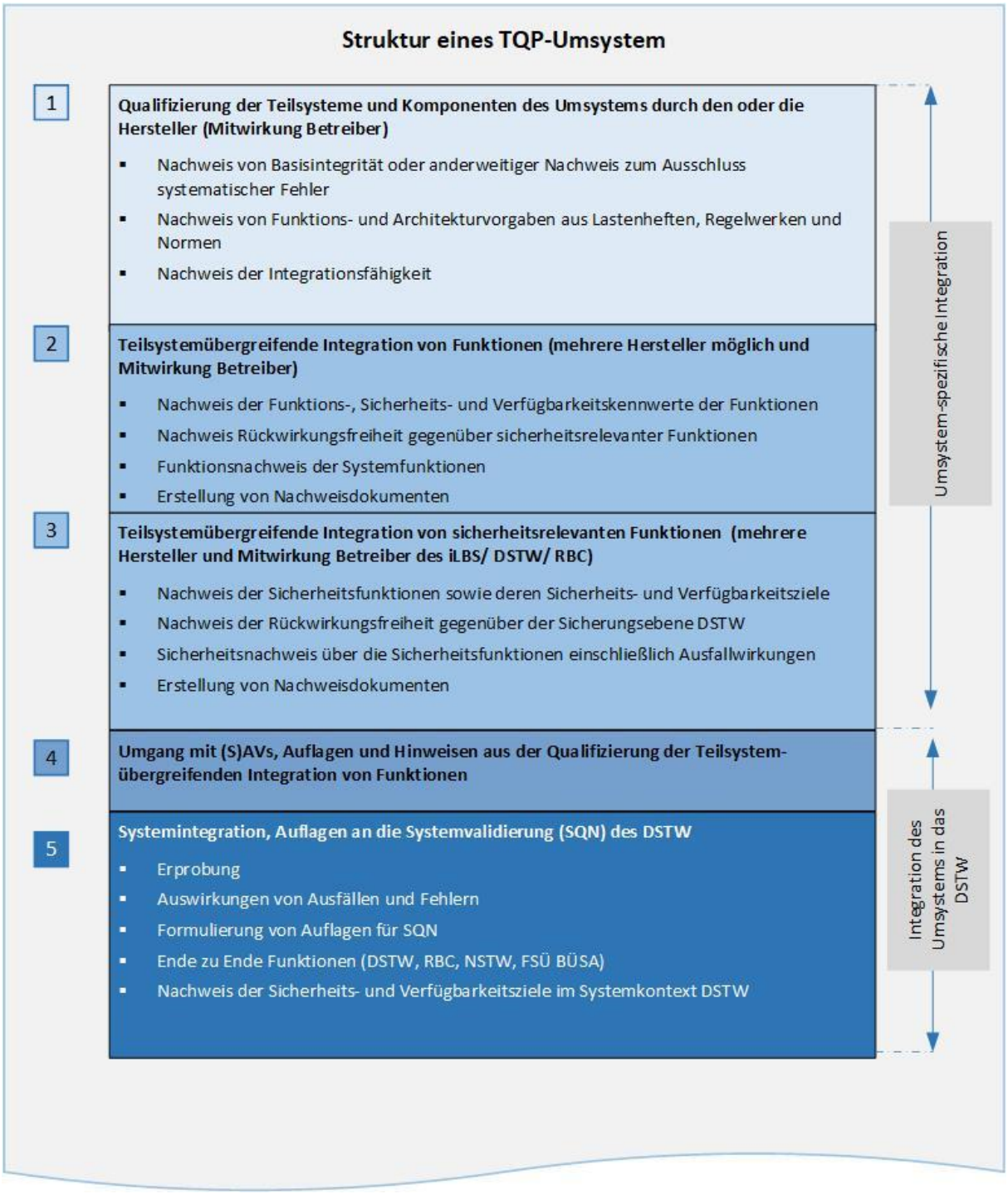


Abbildung 4: Struktur eines TQP für Umsysteme

6 System-Genehmigung

Die erreichte Tiefe der Nachweisführung und Integration sollte für künftige Projekte mit ähnlichen Konstellationen weitergeführt werden können. Hierfür ist ein Verfahren notwendig, das unveränderte, gemeinsame Aspekte überträgt. Gemäß den Bestimmungen der EIGV und den Vorgaben aus der Sektorleitlinie erlaubt § 27 EIGV die Anwendung einer Genehmigung zum Inverkehrbringen und Verwenden (GIuV), die im weitesten Sinne einer generischen Zulassung gleichkommt.

Mit der Integration der Teilsysteme entsteht durch das inhärente Übertragungssystem ein neues System, das den Zulassungsanforderungen gerecht wird. Daher haben DSTW-Integratoren die Möglichkeit, beim Eisenbahn-Bundesamt (EBA) eine GIuV für das gewünschte DSTW-System zu erlangen. Diese GIuV ist optional und geht der obligatorischen (spezifischen) Inbetriebnahmegenehmigung (IBG) voraus.

6.1 Regulatorische Vorgaben

6.1.1 Eisenbahn-Inbetriebnahmegenehmigungsverordnung

Die Eisenbahn-Inbetriebnahmegenehmigungsverordnung (EIGV) legt die Rahmenbedingungen für Genehmigungen im Eisenbahnsektor fest. Ein zentrales Element für DSTW ist die Inbetriebnahmegenehmigung (IBG), die gemäß § 14 EIGV für jedes spezifische Bauprojekt im Zusammenhang mit Modernisierung und Erneuerung obligatorisch ist (siehe S. 1277, [8]). Bei gegenwärtigen Infrastrukturprojekten handelt es sich in der Regel um Modernisierungen und Erneuerungen bestehender Strecken, während die Inbetriebnahme neuer Verbindungen nach § 8 EIGV eher selten ist [10, S.26ff.].

Um die Wirtschaftlichkeit zu steigern, sollten nachgewiesene Komponenten auf zukünftige Systeme übertragbar sein. Die IBG ist jedoch spezifisch und erfordert bei jeder Inbetriebnahme eine vollständige Nachweisführung, was die Wirtschaftlichkeit nicht allein steigern kann. Abhilfe schafft hier § 27 EIGV, der eine Genehmigung zum Inverkehrbringen und Verwenden (GIuV) beschreibt. Im Gegensatz zur verpflichtenden IBG ist die GIuV optional und kann bei mehrfacher Anwendung eines Eisenbahnsystems beantragt werden. Die GIuV vermeidet die Wiederholung von Nachweisverfahren in Folgeprojekten und kann so den langfristigen Aufwand reduzieren, insbesondere wenn der generische Anteil groß ist. Bei signifikanten Änderungen des Eisenbahnsystems ist laut Verordnung 402/2013 ein Sicherheitsbewertungsbericht von einer unabhängigen Bewertungsstelle zu erstellen. Das Eisenbahn-Bundesamt (EBA) erteilt die GIuV gemäß § 27 Abs. 1 EIGV für jedes System, wobei das EBA nur die Vollständigkeit und Eindeutigkeit der eingereichten Unterlagen überprüft, sofern keine begründeten Zweifel bestehen. Die Genehmigung wird gemäß § 27 Abs. 3 erteilt, wenn die Anforderungen des § 9 Abs. 1 in Verbindung mit einer Prüfbescheinigung eines Prüfsachverständigen erfüllt sind (siehe S. 1282 [8]), [10, S.26ff.].

Eine vorgeschaltete GIuV erfordert eine frühzeitige Dokumentation und Nachweisführung, sodass Komponenten deutlich vor der IBN vorhanden sein müssen. Dies kann den Zeitdruck für das erste Projekt erhöhen, ist aber akzeptabel, da die Wiederholung der Nachweisarbeit vermieden werden soll und somit Folgeprojekte davon profitieren. Wenn die Deutsche Bahn auf die GIuV verzichtet, müssen dieselben Nachweise in die IBG eingefügt werden. Ausschließlich projektspezifische Nachweisführungen lassen sich nicht ohne weiteres auf veränderte Situationen übertragen [10, S.26ff.].

6.1.2 Sektorleitlinie

Die Sektorleitlinie (SLL) legt die Verfahren der Zulassungsbewertung in den Gebieten der Signaltechnik, Telekommunikationstechnik und Elektrotechnik auf Grundlage der EIGV neu fest und ersetzt die

Verwaltungsvorschrift Neue Typzulassung (VV NTZ) [9]. Im Unterschied zu den bisherigen Verfahrensvorschriften vereint die SLL Aspekte der Signaltechnik, Telekommunikationstechnik und Elektrotechnik gleichermaßen. Sie wurde in Kooperation von Industrie, Betreibern und dem Eisenbahn-Bundesamt entwickelt. Anhang 1 der SLL visualisiert den Prozessablauf in den Phasen Lastenheft, Pflichtenheft und Produkt und weist die Verantwortlichkeiten grob zu [10, S.26ff.].

Zum Zeitpunkt der Erarbeitung der SLL war das Verfahren zur Nachweisführung von modularem DSTW noch nicht vollständig ausgearbeitet, daher konnte es in der SLL nicht berücksichtigt werden. Der modulare Aufbau des DSTW lässt Systemarchitekturen zu, die noch nicht durch Anlage 17 "Herstellerübergreifende Integration" abgedeckt sind. Oft geht man davon aus, dass ein System oder Produkt von höchstens zwei Herstellern stammen kann und integriert werden muss. Die Architektur lässt aber zu, dass mehr als zwei Hersteller beteiligt sein könnten. Dies setzt eine klare Trennung der Rollen von Betreiber und Hersteller voraus [10, S.26ff.].

Unter der Annahme, dass das systemgemäß als Teilsystem definierte Übertragungssystem von der DB bereitgestellt wird, gibt es in einem DSTW-System mindestens drei Parteien. Dabei spielt es keine Rolle, ob die weiteren Teilsysteme (z.B. Weiche und Zentraleinheit) von einem oder mehreren Herstellern stammen, da jedes Teilsystem eigenständig ist. Das in der SLL beschriebene Verfahren ist grundsätzlich auf DSTW anwendbar, jedoch fehlt die Methode der modularen Stellwerkstechnik, sodass dem Leitfaden an einigen Stellen die notwendigen Details fehlen. Dazu gehören zum Beispiel spezifische Informationen über nachgewiesene oder nicht nachgewiesene Offenbarungen von Ausfällen im Fehlerfall und das konkrete Ausfallverhalten, um beurteilen zu können, welche zusätzlichen Betrachtungen auf Systemebene ergänzt werden müssen [10, S.26ff.].

6.2 System-GIuV

6.2.1 Umfang und Voraussetzungen der Betrachtung

Die bisherigen Untersuchungen zu einer GIuV wurden auf der Teilsystemebene durchgeführt, daher wird die notwendige Genehmigung für das DSTW-System als System-GIuV bezeichnet, um eine klare Unterscheidung zu gewährleisten. Die Vorteile einer System-GIuV sind stark von der Wiederverwendbarkeit der Nachweisergebnisse in verschiedenen Systemkombinationen abhängig [10, S.26ff.].

Der Untersuchungsumfang der System-GIuV beinhaltet Teil- und Umsysteme des DSTW. Die Einbindung von Umsystemen in die System-GIuV erfolgt stufenweise, je nach den Anforderungen des Projekts. Bei den Nachweisaktivitäten wird zwischen der Erstgenehmigung einer Systemfreigabe für einen Neubau und der Aktualisierung einer zugelassenen Systemkonfiguration unterschieden. Auf dieser Basis wird in einem weiteren Schritt analysiert, welche Systemänderungen Einfluss auf eine System-GIuV haben können und durch welche Nachweisschritte diese ausgeschlossen oder minimiert werden können [10, S.26ff.].

6.2.2 Verwaltung der Kompatibilität

Das Kompatibilitätsmanagement basiert auf der Annahme, dass ein DSTW aus bereits qualifizierten Teil- und Umsystemen mit eigener GIuV besteht und dass die Hersteller ein starkes Interesse am Fortbestand dieser Genehmigungen haben. Genehmigungen werden in der Regel sehr präzise formuliert. Änderungen an den Teil- und Umsystemen wirken sich stets direkt auf die System-GIuV aus. Wenn solche Änderungen häufig auftreten, können die Vorteile der System-GIuV verloren gehen und die

notwendigen Systemüberlegungen müssen in das Inbetriebnahmegenehmigungsverfahren für ein bestimmtes DSTW-System einbezogen werden. Dies ist nach der EIGV erlaubt, da eine Genehmigung für das Inverkehrbringen und Verwenden von sicherungstechnischen oder elektrotechnischen Systemen gemäß § 27 der EIGV nur optional ist [10, S.26ff.].

Daher sollte berücksichtigt werden, dass ein erhöhter Nachweisaufwand die Vorteile der System-GIuV ausgleicht und somit verringert. Die System-GIuV sollte daher so wenig wie möglich geändert werden. In der System-GIuV sind deshalb unter anderem anstelle fester Hard- und Softwareversionen von Teil- und Umsystemen Verweise auf eine qualifizierte Kompatibilitätsdatenbank enthalten. Der Nutzen einer Datenbank liegt in ihrer Skalierbarkeit. Die Datenbank wird stets alle vorhandenen Kompatibilitäten einer Zentraleinheit mit den integrierten Feldelementen transparent darstellen [10, S.26ff.].

6.2.3 Systematisierung von Änderungsbetrachtungen

Um zu ermitteln, ob eine Änderung in einem Teilsystem dazu führt, dass eine bestehende Prüfbescheinigung des DSTW-Systems und damit auch die System-GIuV ungültig wird, stellt Tabelle 3 Änderungskategorien für Teilsysteme dar [10, S.26ff.].

Tabelle 3: Änderungskategorien für Teilsysteme des DSTW. Aus [10].

Kategorie A	Fehlerkorrektur an einem Teilsystem (Feldelement) ohne Änderung der Anforderungen
Kategorie B	Fehlerkorrektur an einem Teilsystem (Feldelement) mit Änderung der Anforderungen – ohne Funktionshub
Kategorie C	Fortschreibungen an einem Teilsystem (Feldelement) mit Änderung der Anforderungen – mit Funktionshub
Kategorie D	Fehlerkorrektur / Fortschreibungen am Teilsystem Zentraleinheit mit Änderung der Anforderungen
Kategorie E	Fehlerkorrektur / Fortschreibungen an einem Teilsystem (Feldelement) mit Änderung nicht funktionaler Anforderungen
Kategorie F	Anpassungen/Veränderungen am Sicherheits- und Qualitätsmanagement

Hinweise:

- *Bei einem Funktionshub nach Kategorie C wird eine neue Systemfunktion (Lastenheft-Funktionalität) oder ein neues System-Release eingebracht; bei Änderungen an einer bestehenden Systemfunktion gilt Kategorie B*
- *Die Änderung von nicht funktionalen Anforderungen nach Kategorie E betrifft z.B. die Sicherheits- und Verfügbarkeitskennwerte (RAMS) eines Teilsystems*

Aus diesen Kategorien ergeben sich Randbedingungen und Prüfkriterien, die letztlich zu Nachweisaktionen und deren Einhaltung führen, um bei Veränderungen die anhaltende Gültigkeit der System-GIuV bestätigen zu können. Eigene Änderungskategorien werden auch für das Übertragungssystem und die Umsysteme abgeleitet, die jedoch im Rahmen der jeweiligen Teil-Qualifizierung berücksichtigt werden. Fallspezifische Ergänzungen werden in den Prüf- und Validierungsschritten der jeweiligen Änderungskategorie angegeben [10, S.26ff.].

Als Beispiel wird in der Kategorie B (aus Tabelle 3) die ergänzenden Prüf- und Validierungsschritte bei einer Fehlerkorrektur an einem Teilsystem (hier bei einem Feldelement) mit Änderung der Anforderungen in Tabelle 4 dargestellt. Im Zuge dieser Änderung am Teilsystem werden Anforderungen hinzugefügt oder entfernt. Sicherheitsrelevante Anwendungsvorschriften (SAV) können erforderlich

werden. Abweichungen von den genannten Prüf- und Validierungsschritten erfordern die Zustimmung des Betreibers. Generell werden folgende Randbedingungen definiert [10 S.26ff.]:

- *Das betroffene Teilsystem war vor der Änderung legitimiert und die Systemkompatibilität bestätigt.*
- *Ggf. erforderliche Änderungen an Lastenheften oder Pflichtenheften betreffen ausschließlich Fehlerkorrekturen und keine neuen Systemfunktionen*
- *Der Hersteller erstellt eine Auswirkungsanalyse zur Änderung und benennt ggf. erforderliche Anpassungen im SQN.*

Im Folgenden sind die allgemeinen Prüf- und Validierungsschritte (grau hinterlegt) und zusätzlich die spezifischen Prüf- und Validierungsschritte für Kategorie B aufgeführt [10, S.26ff.].

Tabelle 4: Prüf- und Validierungsschritte der Kategorie B. Aus [10].

<table>
<tr><td colspan="2">1. Erstellung einer zur Änderungsankündigung gehörenden Auswirkungsanalyse für das betreffende Teilsystem
a) Gegenstand der Auswirkungsanalyse
▪ Angabe des Änderungsgegenstands
▪ Beschreibung des Änderungsgrunds
▪ Abgrenzung des Betrachtungsumfangs
b) Erläuterung und Bewertung der Auswirkungen der Änderungen (die nachstehenden Angaben stellen den Mindestumfang dar, Angaben hierzu sind immer erforderlich)
▪ Schnittstellen
▪ Laufzeiten, Lastverhalten
▪ Sicherheits- und Zuverlässigkeitskennwerte
▪ Regelwerke (Planung, Projektierung, Abnahme, Instandhaltung, Betrieb…)</td></tr>
<tr><td colspan="2">2. Identifikation des betreffenden Kapitels im SQN und Prüfung der Auswirkung auf Systemebene</td></tr>
<tr><td colspan="2">3. Prüfung eventueller Auswirkungen auf das Gefährdungslogbuch</td></tr>
<tr><td>B1.</td><td>Konformitätsnachweis/Bestätigung zur Anforderungsabdeckung der Lastenhefte</td></tr>
<tr><td>B2.</td><td>Nachweis/Bestätigung der technischen Integrierfähigkeit</td></tr>
<tr><td>B3.</td><td>Nachweis/Bestätigung, dass keine Auswirkungen auf die betriebliche Systemintegration des Teilsystems bestehen</td></tr>
<tr><td>B4.</td><td>Bestätigung der inhaltlichen Gültigkeit der bestehenden Kompatibilitätsbewertung (trotz abweichendem Ausgabestand) aus Herstellersicht</td></tr>
<tr><td>B5.</td><td>Nachweis der herstellerseitigen Systemintegration (mit Zentraleinheit des Herstellers, ohne Übertragungssystem und ohne Umsystem) und Bestätigung der Ergebnisse der bisherigen Systemintegration
Alternativ:
▪ Nachweis, dass keine Auswirkungen auf die herstellerseitige Systemintegration bestehen oder
▪ Abstimmung mit dem Betreiber über vergleichbare Maßnahmen zur Systemintegration (z.B. bei Hersteller ohne Zentraleinheit durch Anschaltung an Fremdlabore, Feldversuche (ohne Sicherheitsverantwortung), Anschaltung an Testeinrichtungen des Betreibers)</td></tr>
<tr><td>B6.</td><td>Ggf. Erstellung sicherheitsbezogener Anwendungsvorschriften (SAV)</td></tr>
</table>

Eine zuvor bestätigte Systemkonformität kann, unter Beachtung der genannten Randbedingungen, bestehen bleiben, wenn die Ergebnisse aus den genannten Prüf- und Validierungsschritten bestätigt werden [10, S.26ff.].

7 Evaluation - Erstanwendung des modularen DSTW Thales – Release 1.2 in Mertingen-Meitingen

Das vorliegende Kapitel widmet sich der Bewertung der erstmaligen Implementierung des modularen DSTW Thales – Release 1.2 in Mertingen-Meitingen. Hierbei handelt es sich um das erste modulare DSTW, das unter der Ägide der DB systemintegriert und validiert wurde. Dies stellt eine bemerkenswerte Errungenschaft dar, da die technische Systemverantwortung nach mehreren Jahrzehnten erstmals wieder bei der DB lag.

Die Realisierung des Projekts brachte diverse Herausforderungen mit sich, insbesondere die komplexe Systemintegration und -validierung, die erstmals unter der Verantwortung der DB durchgeführt wurden. Hierbei waren verschiedene Fachbereiche innerhalb der DB sowie der Hersteller der LST beteiligt. Darüber hinaus wurden das Verfahren zur Erstellung des SQN und für das TQN-Übertragungssystem erstmalig eingesetzt und umgesetzt. Die Implementierung dieser neuen Verfahren bedurfte mehrerer Iterationen seitens des Herstellers, der DB und des PSV, um die Nachweisführung zu finalisieren. Während dieses Prozesses wurden unpräzise Formulierungen korrigiert und notwendige Details ergänzt.

Die Begutachtung durch den PSV stellte einen wichtigen Meilenstein für das Projekt dar und wurde positiv durchgeführt. In diesem Zusammenhang wurden die Funktionen und Eigenschaften des DSTW auf Gesamtsystemebene geprüft. Das DSTW musste dabei eine Reihe von Tests bestehen, die sowohl die normativen als auch die betrieblichen Anforderungen an ein Stellwerk erfüllen. Insbesondere wurden die sicherheits- und verfügbarkeitsrelevanten Aspekte des DSTW sorgfältig überprüft. Die Inbetriebnahme des DSTW selbst steht zum gegenwärtigen Zeitpunkt noch aus.

Eine Vereinbarung mit dem EBA besagt, dass das Verfahren zur System-GIuV nach Abschluss des Projekts, also nach Inbetriebnahme des DSTW, erprobt werden soll.

8 Zusammenfassung

Die praktische Anwendung der vorgestellten Methoden wurde bei der Erstimplementierung des modularen DSTW Thales – Release 1.2 in Mertingen-Meitingen erfolgreich demonstriert, auch wenn die Inbetriebnahme des DSTW zum gegenwärtigen Zeitpunkt noch aussteht. Dieses DSTW markiert einen entscheidenden Fortschritt bei der Modernisierung der Bahninfrastruktur in Deutschland und trägt wesentlich zur Steigerung von Sicherheit und Effizienz im Schienenverkehr bei.

Mit der erfolgreichen Erstanwendung dieser Methoden ist ein wichtiger Grundstein für weitere Projekte dieser Art gelegt worden. Die eingeführten Verfahren und Strukturen stellen sich als flexible und robuste Werkzeuge für die Integration von Komponenten verschiedener Hersteller in einem komplexen System dar. Sie ermöglichen eine effiziente und transparente Nachweisführung im Sinne der Sicherheitsanforderungen.

Die Erkenntnisse aus der praktischen Umsetzung liefern wertvolle Ansätze zur Optimierung des Verfahrens und zur weiteren Verbesserung der Systemintegration. Ein kontinuierlicher Prozess des Lernens und der Weiterentwicklung wird dazu beitragen, das hohe Sicherheitsniveau des Schienenverkehrs in Deutschland auch in Zukunft zu gewährleisten und weiter zu verbessern.

Insgesamt zeigt dieser Artikel, dass eine systematische und strukturierte Herangehensweise an die Sicherheitsqualifizierung und Genehmigung von komplexen Bahninfrastruktursystemen, wie dem DSTW, eine wesentliche Grundlage für einen sicheren und effizienten Schienenverkehr bildet. Dabei unterstreicht er die entscheidende Rolle der DB Netz als Integrator und Betreiber in diesem Prozess.

9 Ausblick

Um die Effizienz der generischen Nachweisführung zu steigern und eine rasche Integration neuer, modifizierter oder korrigierter Teil-/Umsysteme in den Geltungsbereich der System-GIuV zu gewährleisten, wird empfohlen, eine klare Abgrenzung zwischen der generischen Systemintegration und spezifischen Bauprojekten vorzunehmen. In diesem Zusammenhang plant die DB die Errichtung eines eigenen Referenzlabors inklusive Referenzstrecke. Durch diese Separation von generischen und spezifischen Aspekten entstehen zeitliche Flexibilitäten (z.B. hinsichtlich Inbetriebnahme-Terminen) und es eröffnen sich Möglichkeiten für alle Beteiligten, neue Funktionen gemeinschaftlich zu erproben [10, S.29].

Das übergeordnete Ziel ist es, in den kommenden Jahren die Anzahl der Testfälle erheblich zu steigern und die Projektszenarien des Referenzlabors progressiv zu erweitern, um das gesamte Funktionsspektrum der DLST repräsentieren zu können. Analog zum Baukastensystem können vorab geprüfte und damit validierte Kombinationen zügiger in Praxisprojekten implementiert werden, was wiederum die Inbetriebnahme beschleunigt [10, S.29].

10 Literaturverzeichnis

[1] https://fahrweg.dbnetze.com/fahrweg-de/kunden/nutzungsbedingungen/digitale_lst,

12.04.2023 um 08:41 Uhr, 15.06.2021 um 08:41 Uhr

[2] Rothkehl, Markus; Oetting, Andreas (11|2022): Sicherheitsnachweisführung. Technische Systemintegration in Verantwortung des Betreibers. In: *Eisenbahningenieur*, S. 29 - 33.

[3] EN 50126-1:2017, 2017: Bahnanwendungen – Spezifikation und Nachweis von Zuverlässigkeit, Verfügbarkeit, Instandhaltbarkeit und Sicherheit (RAMS)

[4] EN 50128:2011, 2011: Bahnanwendungen – Telekommunikationstechnik, Signaltechnik und Datenverarbeitungssysteme – Software für Eisenbahnsteuerungs- und Überwachungssysteme; Deutsche Fassung EN 50128:2011

[5] EN 50129:2018, 2018: Bahnanwendungen – Telekommunikationstechnik, Signaltechnik und Datenverarbeitungssysteme – Sicherheitsbezogene elektronische Systeme für Signaltechnik;

[6] Wallasch, Stephan (2017). Systemdefinition ESTW-NeuPro, Frankfurt a. Main. NeuPro.22, Baseline 4.0.

[7] Lucas, Julian; Rothkehl, Markus; Oetting, Andreas (01|2023): Teil-Qualifizierung in der digitalen Leit- und Sicherungstechnik. Integration von Neu- und Bestandsystemen mit dem modularen DSTW. In: *Eisenbahningenieur*, S. 14–17.

[8] EIGV. Bundesgesetzesblatt, 1270–1310.

[9] Eisenbahn-Bundesamt (2021). Sektorleitlinie für die Zulassungsbewertung von Signal-, Telekommunikations- und Elektrotechnischen Anlagen (Technische Vorschrift). München.

[10] Rothkehl, M.; Oetting, A. (2023): System-GIuV. Genehmigungen in Verantwortung des Betreibers. In: Eisenbahningenieur, 1/2023 S. 26- 30

Strategies for Reliable Infrastructure Data

Wenzel, Benedikt[1]; Wenz, Andreas[2]; Nitzschke, Henning[3]

[1] NEXTRAIL GmbH; [2] SBB AG; [3] Digitale Schiene Deutschland, DB Netz AG

Summary

Planning, building and operating railway systems depend on reliable infrastructure data – with a growing number of consumers and increasing quality requirements. These requirements are contrasted by today's data preparation or management processes, which are too often based on insufficient inventory data, not machine-readable formats like paper, interrupted toolchains, low automation, and poor standardisation. The impacts on the quality of infrastructure data and appropriate strategies to meet the challenges associated with generating and maintaining reliable infrastructure data are described.

Keywords: Infrastructure Data, Data Management, Standardisation, Data Acquisition, Data Quality

1 Introduction

The sufficient quality of infrastructure data as an essential input of engineering data is crucial to achieving highly reliable trackside or onboard system functions. While this is already true for today's systems such as ETCS (European Train Control System), it even becomes more critical in the context of new architectures, as specified by RCA (Reference CCS architecture), Shift2Rail or the new ERJU (Europe's Rail Joint Undertaking) program, since the number of subsystems and functions that directly rely on reliable infrastructure data increases, e.g. map-supported localisation 0000, Interlocking or Advanced Protection Systems, ATO (Automatic Train Operation) with precision stop balises and segment profiles (Subset 126), or ATO GoA4 with perception and incident management systems.

This paper addresses the following aspects regarding reliable infrastructure/map data:

- What are the aspects of infrastructure data quality? What does reliable infrastructure data mean?
- What are the impacts on data quality over the life cycle of infrastructure data?
- Which strategies can be applied to achieve and maintain reliable infrastructure data during engineering and operation?
- What are the possibilities, challenges, but also limitations of infrastructure data acquisitions during engineering or operation to achieve reliable infrastructure data?
- How can standardisation support the strategies for reliable infrastructure data?

Parts of this paper are based on the RCA.doc.77 Digital Map Quality Framework 0, published as part of the RCA baseline 1 release. Other influences are discussions and specification work in the ERJU context (e.g. System Pillar Domain Transversal CCS) or national programs (e.g. Digitale Schiene Deutschland, Digitalisierung Bahnsystem) in combination with many years of practical experience in the field of infrastructure data acquisition and engineering processes.

2 Definition of "Reliable Infrastructure Data"

In published documents of the RCA Digital Map cluster, e.g. RCA Digital Map Quality Framework 0, the term *reliable infrastructure data* (synonym to *reliable map data*) has been introduced, which implies a certain quality level of infrastructure data.

Fundamentally, *reliable* infrastructure data is <c since a sufficient level of quality is also required to guarantee a good performance, high availability, and reliability of non-safe functions such as realised by ATO or TMS (Traffic Management System). However, in terms of safety-related applications, the provided data must be trustworthy in the sense of a sufficiently low probability of incorrect information and compliant, e.g. with EN 50129 0.

The attribute "reliable" refers to a typical characteristic of data that satisfies the qualities of being consistent, current, complete, and correct (incl. accurate), as it is similarly defined in the smartrail 4.0 project Topo4 0 or the SafeRailMap project of Digitale Schiene Deutschland (0, annexe C):

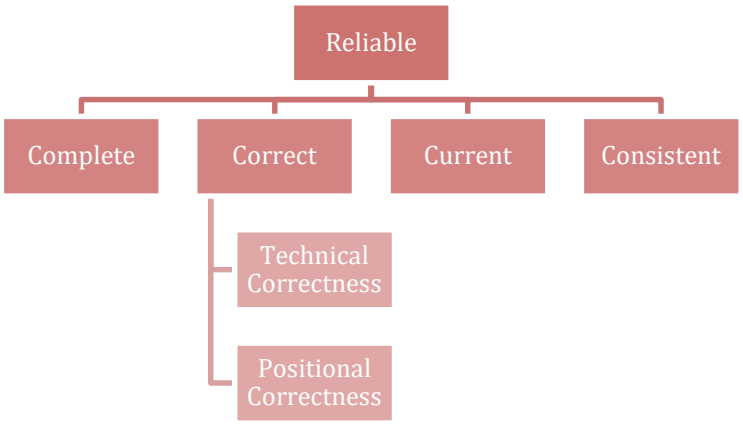


Figure 1 Aspects of Reliable Map Data (source: 0)

- *Complete*: All data (e.g., tracks and elements/objects) within the scope of the application are present, and all attributes required for the intended use case of data are populated
- *Correct*[3]*:* All data is correct regarding the following sub-attributes:
  - *Technical correctness:* the data is valid in the sense that they conform to the rules (e.g., right element type, the minimum distance between balise groups is respected)
  - *Positional correctness*: the data is correct in the sense that they conform to the real existing infrastructure (positional error ≤ accepted positional error)
- *Current*: The data represent valid information for the right time when it is published to the systems
- *Consistent*[4]: No ambiguities or contradictions between different topical aspects, versions, or adjacent regions of data.

While the following paper will emphasise the infrastructure-related data with its positional correctness, completeness and currentness, some of the presented measures can also improve the general quality (including consistency, technical correctness) of data.

[3] The term "correct" is used instead of the formerly used term "accuracy" to avoid collisions with other meanings within this context

[4] "Consistency" is mentioned explicitly but could also be part of e.g. "technical correctness"

3 Impacts on Quality of Infrastructure Data

At the point in time when a consuming function uses infrastructure data during railway operation, the infrastructure data's quality attributes are determined by several impacts and influences that emerge during the life cycle of the data before its consumption and which might enhance or degrade the data quality regarding the mentioned aspects of reliable infrastructure data. These effects are called "contributions". The main effects can be classified into the phases of

- *engineering (*including construction, mounting) within a given engineering project and
- *maintenance* during railway *operation*.

Figure 2 presents the relevant contributors to the infrastructure data quality attributes from these two life cycle phases. The identified contributors are picked up in the subsequent chapters to estimate their impacts on the overall quality of the infrastructure data.

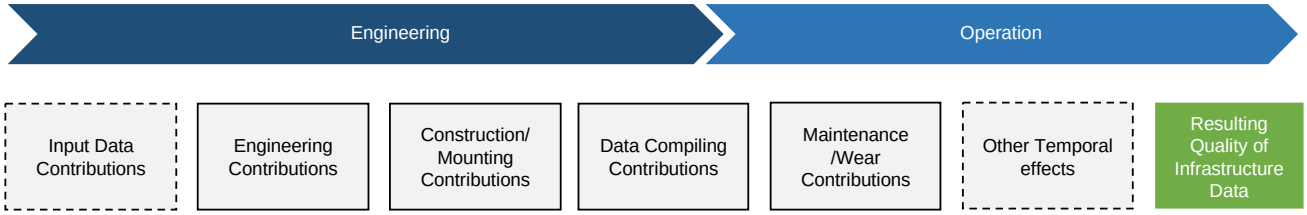


Figure 2 Infrastructure Data Life Cycle from a Quality Point of View (modified based on 0)

3.1 Contributions from Engineering Phase

3.1.1 Input Data

Regarding Planning and Engineering, it is to be divided between *brownfield* projects to upgrade existing tracks, e.g. with CCS (Control, Command, Signalling) systems (like ETCS), and new-build lines in so-called *greenfield* projects. The reason is that, especially in brownfield projects, the quality of inventory data as an impact on the engineering process plays a dominant role. The potential error contribution largely depends on the specific situation of the infrastructure manager regarding the proper management of inventory data, including the management of tracks and assets during operation and the provided data quality from former engineering projects.

Today, in many situations, the formal binding version of inventory data is available on paper or non-machine-readable formats only, e.g. regarding the element positions and functional information based on CCS planning. Usually, the information from plans is digitised to deliver the required input data for the planning. The actual error contribution of digitising paper plans depends on the drawings' quality and the proper use of tooling, so the estimated contributions can be even worse. Also, the aggregated information from several (paper) plans is often inconsistent, incomplete, or outdated.

3.1.2 Engineering

Ideally, there should be no additional error contribution from the engineering itself, e.g. if the tolerated accuracy level is in the sub-meter or decimetre range. However, the engineering processes and tools of today still provide error sources like:

- Legacy-driven tool chains and processes do not fit the system functionality, e.g. the typical reference is still the line kilometre, even for the engineering of track-sensitive ETCS supervision or ATO segment profiles.
- The lack of continuous toolchains with digitised interfaces over the whole process requires manual data transfers, which cause efforts and are prone to errors.
- The engineering processes and tools, especially for CCS, are usually based on drawings and plans, not data-centred models.
- The engineering data is not geo-referenced, so the validation against measurement data causes high efforts, and the connection to BIM methods is challenging.
- The planning decisions are not completely documented or rather capsulated in additional documents, so knowledge transfer is difficult and additional efforts are caused in follow-up projects.
- Due to the different requirements of each infrastructure manager and their planning guidelines, the potential of generic and automated solutions with a valid business case is very limited.
- The tools and, e.g. automation tasks, are usually not developed as reliable tools according to EN50128 0 and EN50129 0, so the safety responsibility and main effort are still transferred to the planner.
- The transformation of coordinate reference systems can lead to additional inaccuracies

3.1.3 Construction and Mounting

Additional discrepancies between the engineered track axis and the constructed track arise from the positional errors introduced during construction: Even though the commissioning process requires the constructed track axis to match the engineered track axis closely, specific tolerances are accepted, which contribute to the positional error of the engineering data.

The same applies to mounting trackside CCS elements like axle counters or balises, usually with higher tolerances than track construction.

Due to insufficient feedback loops, the engineering process might be unaware of potential ad-hoc changes in the field, leading to outdated/incomplete data (even if the safety is ensured by alternative solutions, e.g. additional paper documentation.

3.1.4 Data Compiling

The compilation step transforms the engineering data, including infrastructure information, into the data structure as required by consumers (e.g. ETCS, localisation,…). When the data is compiled, additional contributions can be caused by sampling the original data with simplified structures that the systems can process during runtime, e.g. sampling the alignment elements (curves, straight, transition curve,..) of the track axis by track points with variable distances (e.g. higher point density for track sections with small radius). Another example is the transformation of gradient data to a simplified gradient profile according to the restrictions of the ETCS airgap between the trackside and onboard subsystem and as defined by Subset 026, Packet 21.

Other effects, like the transformation of coordinates to another reference system, can be neglected if proper data and reference systems are used. Besides the mentioned aspects, the transformations during compile process should be (nearly) lossless.

3.2 Contributions from Operational Phase

3.2.1 Track Maintenance/Wear

During operation, a track axis can change over time due to wear and maintenance activities like tamping. The change leads to a discrepancy between the infrastructure data and the actual track axis. During the regularly applied tamping, the track axis is corrected to match the nominal axis from engineering, but a residual error remains due to tolerances.

The possible contribution depends on impacts like track tamping interval, season, and track usage/axle load. However, the contributor is limited by the track tolerances of the infrastructure managers or national regulations.

3.2.2 Element Maintenance/Wear

During operation, the position of trackside elements can change over time due to track or element maintenance activities, which can include remounting of elements. The change leads to a discrepancy between the infrastructure data and the actual position of the elements. The error contribution usually depends on the national guidelines and processes of element maintenance. The actual tolerance today is assumed to be element-type specific.

3.2.3 Further temporal effects

Over time, continental drift can induce positional discrepancies between infrastructure data and reality, which can be neglected with a suitable geo-reference system such as ETRS89 for Europe.

Another contribution can result from weather or temperature-related influences, e.g. on buildings such as bridges and the track axis. This effect is already covered by the tolerances of track maintenance (section 3.2.1).

3.3 Summary of contributions

All the contributions mentioned can have an effect on the reliability of data, e.g. regarding

- (Positional) Correctness: Accuracy of track or element 2D position (longitudinal, transversal) and height. Regarding the track axis, additional attributes like gradient and CANT are relevant.
- Completeness
- Currentness
- Consistency: is implicitly included in the other aspects or less relevant if we focus on the correct representation of infrastructure

Cost efficiency should be considered as an additional parameter since all lack of data quality also decreases process efficiency. On the other side, potential improvements to mitigate contributions induce additional costs, which must also be part of the equation.

Based on RCA Digital Map Quality Framework 0, Table 1 summarises the estimated contributions of the mentioned effects, which are considered to be relevant[5]. The columns of the table represent the following estimations:

- Track/Element: Where applicable, it is distinguished between contributors to the quality of track geometry (track axis) and element position.
- Longitudinal N-E [m], Transversal N-E [m], Height [m]: estimation of positional correctness (quantitatively: positional error)
- Outdated or Incomplete: estimation of contribution to
  - *Currentness*: added possibility for outdated data, e.g. 50% means every second project faces the problem of outdated infrastructure data and
  - *Completeness*: added probability of incompleteness, e.g. 50% means every second project faces the problem of incomplete infrastructure data
- Cost Increase/invest: to support a trade-off analysis between achievable data quality and economic aspects, the analysis aspects are supplemented by one-time investment costs (estimated cost increase, e.g. by process change, calculated in the percentage of overall project costs) and ongoing costs (qualitatively: no, low, medium, high, very high, e.g. caused by additional project or maintenance efforts)
- Criticality of contribution: this is a qualitative categorisation of contributors regarding their potential negative effect on data quality (e.g. “high” means big a big quality decrease)

The values of Table 2 are derived from guidelines, technical analysis (e.g. data compiling) and estimations based on experiences of the involved experts in the cluster (detailed rationales for values to be found in 0).

[5] additional track attributes gradient and CANT have been neglected in the analysis of the current release but can be added following the introduced methodology of RCA Digital Map Quality Framework 0

Table 1: Estimated, relevant contributions based on RCA Digital Map Quality Framework 0

Contributor	Track / Element	Longitudinal N-E [m]	Transversal N-E [m]	Height [m]	Outdated or Incomplete	Cost Increase/ invest	Criticality of contribution
Engineering Input Data: Import existing IM Data	both	2	2	1	50%	50% /no	high
Engineering Input Data - Alternative: Digitalization of Inventory Data	both	5	5	5	50%	50% /no	high
Engineering process of today (not estimated in 0)	both	1	1	1	50%	50% / no	High
Construction and Mounting: Track construction error (tolerance)	track	0.015	0.015	0.025	10%	0% /no	Medium
Construction and Mounting: Element mounting error	element	0.6	0.1	0.1	10%	0% /no	High
Data Compiling: e.g. Point based representation with sufficiently low point distance	track	0.01	0.01	0.01	0 %	0% / low	Low
Data Compiling - Alternative: e.g. Vector based representation of track axes	track	0	0	0	0 %	0% / low	-
Track Maintenance/Wear: Standard Track Maintenance	track	0.025	0.025	0.030	10%	0% /no	Medium
Element Maintenance/Wear: Standard, e.g., remount balise at different sleeper	element	0.60	0.60	0.03	10%	0% /no	High

In the following, some of the critical contributors are analysed regarding effective mitigation measures, such as processual or methodical improvement (section 0). In addition, the potential effect of newly acquisitioned data to improve the data quality is analysed in section 3.5. The final effects of potential combinations of contributions – with or without improvements – in possible process scenarios are summed up by the quality evaluation in Chapter 4.

3.4 Improvements

3.4.1 Engineering: Improved tool-supported process

To overcome the negative influences of today's engineering processes, as pointed out in 3.1.2, the following topics should be addressed:

- Build a digital toolchain that connects across the trades and synchronises the planning results efficiently
- Avoid manual data transfers and replace them with standardised data flows
- Introduce digital workflows also in communication and verification/validation activities (e.g. digital signature, digital plan verification)
- Automate repetitive tasks as much as economically and technically possible to minimise the potential for human error

- Ensure acceptance/approval/trust of tools incl. automation so no project engineer needs to check the results of automated functions
- Implement data-centred instead of plan-centred workflows so the digital database is the master instead of the drawings. The “plan” becomes just a view of the data.
- Drawing activities, in general, should be avoided within the context of CCS/TMS engineering and replaced by data modelling.
- Support consequent georeferenced planning/engineering to simplify the processes of exact mounting and synchronisation with acquisition data.
- A sufficient CRS (Coordinate Reference System) must be used to mitigate the continental drift effect (e.g. ETRS89 for European Countries) and avoid positional incorrectness due to transformation or temporal effects.
- Connect the engineering process with a centralised master database that ensures the achieved data quality during and beyond the project
- Introduce appropriate engineering tools that follow the mentioned principles and support the engineer in the best way with a high grade of usability
- Ensure sufficient engineering input data (IM Data) quality to avoid project risks, additional costs, and efforts spent for correction (see section 3.5 for the potential benefit of data acquisition).

This list of optimisations is, of course, not exhaustive (more details and structure in 0 and its references). Still, it covers important aspects to bring the engineering processes to a “state of the art” level and speed up the implementation projects, especially in the context of CCS or TMS. The measures require a high investment. How standardisation can support this transformation and reduce the efforts is explained in Chapter 5.

3.4.2 Construction and Mounting: Improved element mounting

With this improvement, the mounting tolerance for elements is reduced and controlled by sufficient measures. First, an improved, geo-referenced engineering process, as described in 3.4.1, can already consider actual mounting possibilities for elements like balises and deliver exact mounting positions as planned coordinates.

Vice versa, the engineering process is aware of all remaining ad-hoc changes during mounting by the fully digital process, which integrates the actual mounting position into the engineering dataset. The position is captured by proper measurement tooling. For validation purposes, several measurements can be taken into account and compared (e.g. distance to nearest object plus geolocation as coordinate).

The extra costs for such optimisations are fully compensated by reduced project risks and faster reaction to spontaneous changes during mounting.

3.4.3 Track Maintenance/Wear: Improved track maintenance

The track wear contribution accumulating over time can be reduced if the preventive maintenance of tracks (tamping) is done with higher frequency (e.g. doubled), which comes with substantial additional costs (incl. impact on regulations, ...). Also, the initial investment is considered high due to additional tamping/maintenance machines.

3.4.4 Element Maintenance/Wear: Improved element maintenance

Like the measures of improved element mounting (3.4.2), the error contribution of element maintenance can be limited by more restrictive element tolerances for maintenance works, which are confirmed by appropriate methods (4 eyes principle, measurement with proper tools...). If the deviation exceeds this lower limit, a correction of the element position or an update of the infrastructure data with the newly determined element position must be performed. In addition, the maintenance work should be digitally connected to a single source of truth for infrastructure data so that the engineering processes and other data users are aware of relevant events during maintenance (investment).

3.4.5 Summary of improvements

The following table shows the contributions of the improved process steps. If the improvement is implemented, the new contributions replace the old contributions of the assigned process step from Table 1. The last column sums up a qualitative cost-benefit estimation. A good cost-benefit ratio indicates that the measure for improvement is recommended. The final effect will also be demonstrated by the evaluation in Chapter 4.

Table 2: Estimated, contributions of improvements based on RCA Digital Map Quality Framework 0

Contribution of Improvements	Track / Element	Longitudinal N-E [m]	Transversal N-E [m]	Height [m]	Outdated or Incomplete	Cost Increase/ invest	Cost-Benefit Ratio
Engineering: improved tool-supported process	both	0	0	0	0%	0% / high	Good
Construction and Mounting*:* Improved element mounting	element	0,025	0,025	0,030	0%	10% /med.	Good
Track Maintenance/Wear: Improved, e.g. Track Tamping interval reduced by 50%	track	0.013	0.013	0.015	-	100% /high	Bad
Element Maintenance/Wear: Improved, e.g. limitation of element maintenance tolerance	element	0.013	0.013	0.015	0%	20% /med.	Good

3.5 Further Improvement by Data Acquisition

The previous sections listed some possible negative contributions to infrastructure data quality along the processes of engineering and operation. In both phases, data acquisitions in the field could be performed to approve or enhance the data quality. To create and maintain the value of acquisition data over the life cycle, the following is required:

- a proper toolchain and databases with lossless data flows along the whole life cycle of data, e.g. see the improved engineering process described in 3.4.1 and
- a seamless process for detecting all relevant changes between two surveys, see 3.4.2 improved element mounting or 3.4.4 improved element maintenance.

The challenges and potential effects of data acquisition are evaluated in more detail for each phase.

3.5.1 Data Acquisition during Engineering

During the engineering phase, the acquisition of the existing infrastructure can be performed to confirm or improve the quality of infrastructure-related Engineering Data. Hence, the acquisition can sufficiently address the quality aspects of positional correctness, completeness, and currentness.

While the infrastructure acquisition as a control measure can mitigate the contributions of the previous engineering activities, it also comes with its own error contributions. For an estimation of the introduced positional errors, see estimations in 0.

Moreover, the integration of acquisition into the project also comes with additional efforts and costs, e.g. for the measurement itself, the post-processing of measurement data, the feature extraction and comparison against existing data.

The acquisition should be integrated into the data preparation phase to provide good engineering input data or validate engineering/infrastructure data against the existing infrastructure. A vital pre-requisite and also one of the biggest challenges for the successful integration of infrastructure acquisition (also referred to as “digitisation campaigns”) is the rapid and reliable analysis of measurement data (point clouds, videos, trajectories, ...) to the structured input data, that is required for data preparation or maintaining data quality (tracks, nodes/edges, elements like signals, balises, …). The key challenges to be solved are:

- Apply measurement methods that fulfil the required accuracy level:
  - The required accuracy must be within a technical, economically feasible range (e.g. consider acquisition efforts, avoid too restrictive mm/cm accuracy)
  - The provided accuracy of the applied measurement tool must be within the tolerated range
  - The suitability of the measurement method must be confirmed within the framework of appropriate tests/certifications/approvals.
  - Control points should be processable if necessary and reasonable from a technical/economic point of view to confirm the measurement during application. Note: additional maintenance efforts due to detectable control points should be avoided
  - Regular maintenance and calibration of the measuring system shall continuously ensure its suitability.
  - The recorded data (e.g. point cloud) must provide sufficient density to minimise the error of feature extraction (analysis)
- Apply acceptable analysis methods to extract the tracks and elements from the measurement data:
  - The toolchain implementing the analysis functions (incl. extraction of tracks and elements) shall be implemented safely and robustly (incl. error handling…).
  - a high degree of automation regarding post-processing and analysis is required to avoid process barriers
  - the exact reference position must be defined for each measured element type
  - The extraction shall allow precise positioning of the tracks and elements within the permissible tolerances

- Efficient validation and merging of engineering data with acquisitioned information
- All tool support (e.g. automatisation of feature extraction, automated comparison or data transformation) must be done in a reliable way to make the resulting data acceptable for SIL4 functions. (e.g. tool certification according to EN 50128/9)

3.5.2 Data Acquisition during Operation

Similar to the engineering phase, the error contribution might be limited by the acquisition of infrastructure, which is applied continuously during operation. In this approach, e.g. operational trains are equipped with extended recording systems that (permanently) collect data during operation.

However, continuous data collection is only suitable to a very limited extent for the reliable and timely detection of potential maintenance faults. Hence, the acquisition should also be locally triggered in case of defined events (e.g. finished maintenance works, environmental impacts, …).

Concerning the track axes position (geometry), the dynamic method would have to have a very high accuracy (<= cm range) to provide more up-to-date data than the original alignment (which comes with high costs and time expenditure!). The required accuracy is very challenging for dynamical measurement methods, as they are recommended here to avoid impacts on the line capacity.

On the other hand, this method of continuous acquisition is suitable for the gradual adjustment of element information in the infrastructure data, e.g. to compensate for the seasonal changes if relevant for landmarks of perception systems. Moreover, due to the high amount of remeasurement repetitions, a high level of trust can be built up in the infrastructure data (a counterargument, besides the high efforts, is that the data already must comply with the required properties when the infrastructure is put into operation).

4 Achievable data quality

The RCA Digital Map Quality Framework 0 combines all contributions with potential improvements, including the acquisition at different project stages (e.g. beginning of the project, after mounting elements, during operation), to relevant scenarios for further investigation. For each scenario, the relevant contributions have been selected and summed up to express the achievable data quality and evaluate the resulting efforts/costs regarding efficiency[6].

For example, for the engineering phase, the following scenarios have been analysed (applied improvements incl. acquisition are *highlighted*):

- ES0: Import Input Data + *(Improved) Engineering* + Mounting + Compile/Publish
- ES1: Import Input Data + *(Improved) Engineering* + Mounting + *Acquisition* + Compile/Publish
- ES2: *Acquisition* + *(Improved) Engineering* + *Improved Mounting* + Compile/Publish
- ES3: Engineering with reliable input data (e.g. new build or reliable inventory data) + Construction + *Improved Mounting* + Compile/Publish

[6] The contributions are clustered in a way that additional interrelations between contributions can be neglected (e.g. if “Acquisition” replaces “Import Input Data” at the beginning of the scenario, then the negative effects of insufficient data for engineering are also considered)

Based on these scenarios of the engineering phase, the following contributions are summed up for tracks and elements (Figure 3):

Implementation (Mitigation)	2D Pos Track Imprecision [m]	2D Pos Elem Imprec. [m]	Height Track Imprec. [m]	Height Elem. Imprec. [m]	Probability Incomplete	Probability Outdated	Cost Increase Project	Invest
ES0: Import+Engineer.+Mount.+Comp/Pub	2,83	3,44	1,00	60%	60%	50%	no	
ES1: Import+Engineer.+Mount.+Acquisition+Comp	0,07	0,08	0,07	0%	0%	55%	medium	
ES2: Acquisition+Engineer.+Impr.Mount+Comp/Pu	0,07	0,10	0,07	0%	0%	40%	high	
ES3: Engineer.+Constr+Impr.Mount+Comp/Pub	0,02	0,02	0,03	0%	0%	10%	low	

Figure 3: Summarized contributions for Engineering Scenarios

After that, the resulting error contributions for tracks and elements are compared against *predefined target values* defined as follows:

- For the position, the total 2D positional error is calculated by applying the vector norm (the root of summed squares) based on the following components:
  - position error longitudinal N-E $\Delta p^{\parallel N-E} = 0{,}1\text{m}$
  - position error transversal N-E $\Delta p^{\perp N-E} = 0{,}1\text{m}$
- The tolerated height error is as high as the positional error
- There is no accepted probability of incomplete or outdated data.
- The maximum cost increase for both phases is 40% (20% per engineering project, 20% per maintenance interval)

The calibration follows typical requirements from ETCS or ATO implementation projects but can be changed in the dynamically calculated quality model that is provided as an attachment to 0.

To allow a separate analysis, the defined target values are split into budgets for the phases of engineering and operation, e.g. like shown in Figure 4:

Accepted Quality (Positional Correctness: 1 sigma)	2D Track position [m]	2D Element position[m]	Height Track Imprec. [m]	Height Elem Imprec. [m]	Probability Incomplete	Probability Outdated	Cost Increase Project/ Interval	Invest
Overall	0,14	0,14	0,1	0,1	0%	0%	40%	-
Budget "Engineering"	0,10	0,09	0,07	0,07	0%	0%	20%	-
Budget "Operation"	0,04	0,05	0,03	0,03	0%	0%	20%	-

Figure 4: Quality Budgets for Engineering and Operation Phase

4.1 Results and Recommendations for Engineering Phase

The evaluation of the engineering scenarios leads to the following grade of fulfilment for the required acceptance levels (around 100% is required for fulfilment, see Figure 5):

Evaluation Engineering Scenarios	Correct 2D Track Pos	Correct 2D Elem Pos	Correct Track Height	Complete	Current	(Cost efficient)	Invest	Acceptable
Reference	*100%*	*100%*	*100%*	*100%*	*100%*	*100%*	-	-
ES0: Import+Engineer.+Mount.+Comp/Pub	4%	3%	7%	40%	40%	70%	no	
ES1: Import+Engineer.+Mount.+Acquisition+Comp	130%	122%	108%	100%	100%	65%	medium	
ES2: Acquisition+Engineer.+Impr.Mount+Comp/Pu	130%	96%	108%	100%	100%	80%	high	
ES3: Engineer.+Constr+Impr.Mount+Comp/Pub	130%	130%	130%	100%	100%	110%	low	

Figure 5: Evaluation Engineering Scenarios 0

Based on Figure 5 and the chart in Figure 6, the following scenarios can fulfil the criteria:

- ES0: quality criteria not achieved
- ES1: relevant quality criteria fulfilled, but the most expensive solution
- ES2: relevant quality criteria narrowly fulfilled with improved cost efficiency due to avoided costs from insufficient input data
- ES3: all quality criteria are easily met with the lowest project costs due to available high-quality engineering input data (new build line, optimised inventory data, …)

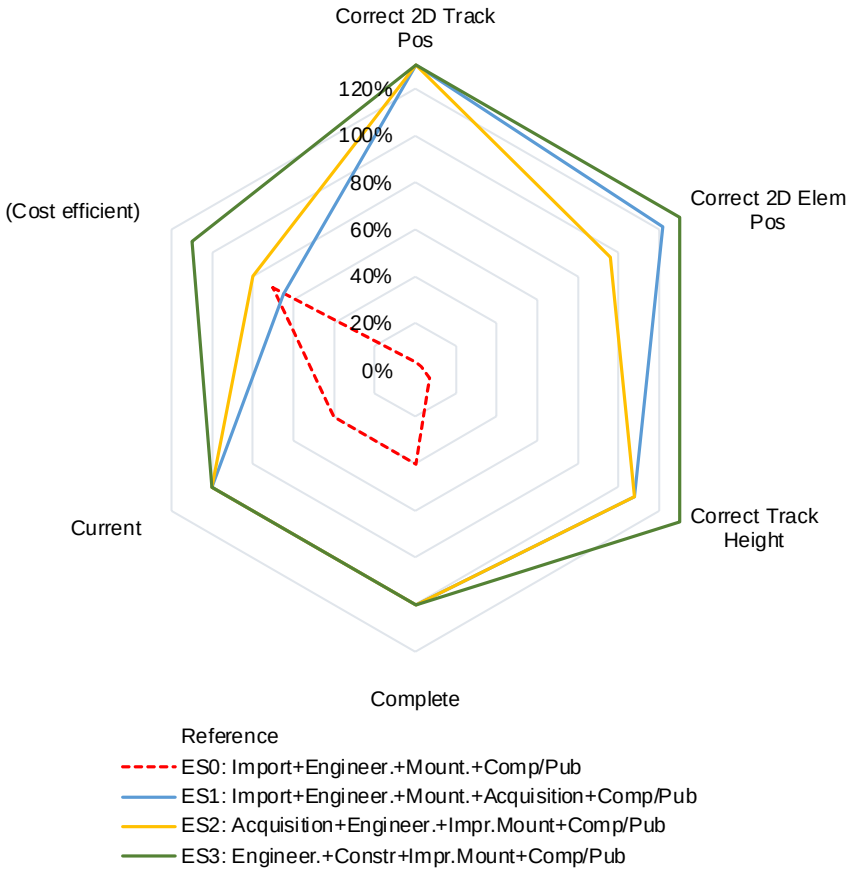


Figure 6: Chart Evaluation Engineering Scenarios 0

4.2 Results and Recommendations for Operation Phase

The same methodology has been applied to the operational phase. Figure 7 summarises the analysed scenarios and their criteria fulfilment. The best criteria fulfilment and suitable cost-benefit-ratio is provided by scenario "OS1: Standard Track Maintenance + *Improved Element Maintenance*". Due to the high efforts and limitations regarding accuracy and positive effects, continuous acquisition is not recommended, as already indicated in section 3.5.2.

Evaluation Operation Scenarios	Correct 2D Track Pos	Correct 2D Elem Pos	Correct Track Height	Complete	Current	(Cost efficient)	Invest	Acceptable
Reference	*100%*	*100%*	*100%*	*100%*	*100%*	*100%*	-	-
OS0: Standard Maintenance	120%	5%	100%	90%	90%	120%	no	
OS1: Standard Track + Impr. Elem. Maint.	120%	130%	100%	100%	100%	100%	medium	
OS2: Improved Track + Impr. Elem Maint.	130%	130%	130%	100%	100%	40%	very high	
OS3: Continuous Acquisition	71%	71%	55%	90%	90%	40%	high	

Figure 7: Evaluation Operation Scenarios 0

It is referred to [1] and its attached quality model for detailed results.

5 Chances of Standardisation

A new standardised architecture, as intended by the former RCA or the new ERJU System Pillar, offers new possibilities to improve the situation in a more efficient and economical way, such as:

- A standardised engineering input data format could be used to import inventory data from existing IM databases or new acquisitions, e.g. based in IFC rail, RSM or other standardised structures.
- A standardised data model format can be defined to harmonise the information that enables the functionality of all covered systems and their use cases.
- Instead of digitising the actual processes, the data needs and target processes can be derived functionally to focus on the generic core and avoid expensive variations.
- Standardised engineering rules should be part of the overall system design, including minimising complexity and maximising the generic rules.
- The design of new architectures should also aim for reduced engineering effort by replacing fixed configuration with dynamic system function (e.g. avoiding pre-configured routes)
- Standardised infrastructure/engineering data provisioning procedure via functional interfaces to the systems, incl. loading and synchronised activation of incremental data version updates.
- The standardised data provisioning also requires harmonised version- and id-management (e.g. for assets), which simplifies many processes, especially during operation (e.g. stable connection to asset management with clearly defined id life cycle)

Each mentioned standardisation aspect increases the room for generic and cost-efficient solutions to optimise and automate data engineering, validation, and transformation/compiling. This is crucial to achieving economical solutions, especially in the safety-related system context. Consequently,

subdomain 1 of the Transversal CCS domain within the System Pillar aims to address these issues for standardisation – in coordination with the other domains and the Flagship Areas of the Innovation Pillar.

6 Conclusion

The paper shows that reliable infrastructure data is achievable in different scenarios. The acquisition of real infrastructure is especially recommended in “brownfield” projects. It has been shown how other processual and methodical improvements can substantially increase the quality of infrastructure data. Other measures, such as continuous data acquisition during operation, did not show sufficient positive effects to justify the induced efforts by this analysis.

Proper data management is generally required to create and maintain high data quality over the life cycle. The basic attributes of “state-of-the-art” tool-based processes have been summarised. Finally, the standardisation as the ERJU program intends, offers the possibility for generic, highly efficient solutions to improve the situation and accelerate the realisation projects.

7 Literature

[1] Benedikt Wenzel, Andreas Wenz, Henning Nitzschke: Digital Map - Quality Framework. RCA.doc.77 v0.2 as part of RCA Release Baseline 1, 2022

[2] EN 50128: Railway applications - Communication, signalling and processing systems - Software for railway control and protection systems, CENELEC, Apr 2011

[3] EN 50129: Railway applications - Communication, signalling and processing systems - Safety related electronic systems for signalling, CENELEC, Nov 2019

[4] SBB smartrail 4.0 Topo4 Functional Concept (published filename: ES TOPO4 45-Published Functional Concept.pdf), https://www.voev.ch/de/Service/content_?download=18086, 29.06.2022

[5] Wenz, A., Ehrler, R., & Ohrendorf-Weiss, S. (2022). Map Supported Train Localization. Proceedings of the 35th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2022), 1978–1988. https://doi.org/10.33012/2022.18547

[6] Roth, M., Baasch, B., Havrila, P., & Groos, J. (2018). Map-Supported Positioning Enables In-Service Condition Monitoring of Railway Tracks. 2018 21st International Conference on Information Fusion, FUSION 2018, 2346–2353. https://doi.org/10.23919/ICIF.2018.8455377

[7] Böhringer, F., & Geistler, A. (2006). Location in railway traffic: Generation of a digital map for secure applications. WIT Transactions on the Built Environment, 88, 459–468. https://doi.org/10.2495/CR060461

[8] Henning Nitzschke: CLUG publication - Definition of the required Maps for Localisation. Version 1.3, 23.06.2022

IT-Security im System Bahn: Zulassungsbewertung & Nachweisführung im Rahmen der Digitalen Transformation (Erweiterung der Sektorleitlinie)

Dr. Matthias Drodt[1] und Dr. Frank Weber[2]

[1] DB Netz AG

[2] Eisenbahn-Bundesamt (EBA)

Disclaimer:

In diesem Paper werden Lösungsansätze aufgezeigt für die Zulassung von IT-Security in Eisenbahnsystemen unter Einhaltung der gesetzlichen und regulatorischen Rahmenbedingungen (Erweiterung der entsprechenden Sektorleitlinie 0). Diese Ansätze wurden in einer sektorübergreifenden Arbeitsgruppe erarbeitet[3]. Ein sich anschließendes Review steht aktuell noch aus. In der später in Kraft gesetzten Version der Sektorleitlinie sind diesbezüglich abweichende Regelungen möglich.

Im Rahmen dieses Papers sowie auch in der Sektorleitlinie ist mit dem Begriff „Sicherheit" grundsätzlich die funktionale Sicherheit (Safety) gemeint. IT-Security bezeichnet im Gegensatz hierzu die Cybersecurity bzw. IT-Sicherheit.

1 Einleitung

Bei Systemen im Eisenbahnumfeld werden regelmäßig Aktivitäten im Kontext RAMS beschrieben. RAMS beschreibt alle Prozesse im Lebenszyklus eines Produktes oder Systems, die sich mit der Zuverlässigkeit (**R**eliability), Verfügbarkeit (**A**vailability), Wartbarkeit (**M**aintainability) und funktionalen Sicherheit (**S**afety, nicht Security!) beschäftigen. Bei allen diesen Prozessen wird i.d.R. davon ausgegangen, dass das System Bahn regelkonform betrieben wird und keine mutwilligen und bewussten Veränderungen am System oder an Teilsystemen vorgenommen werden, d.h. es werden ausschließlich systematische und zufällige Fehler oder Fehlhandlungen betrachtet. In § 4 AEG[4] [2] sind hierzu Anforderungen an die Sicherheit (im Sinne von Safety) von Eisenbahninfrastrukturen und Fahrzeugen und in § 2 EBO[5] [3] Anforderungen an die Betriebssicherheit von Bahnanlagen und Fahrzeugen definiert. Bewusste Veränderungen (Manipulationen), unberechtigte Zugriffe oder gar Cyberangriffe wurden bisher nicht vollumfänglich berücksichtigt.

Obwohl sowohl Safety- als auch Security-bezogene Risiken zu analysieren sind und sich die Herangehensweisen nicht grundsätzlich unterscheiden, werden die Ursprünge der Risiken dennoch unterschiedlich betrachtet. Quellen für Safety-Risiken sind statistische Fehlfunktionen der eingesetzten Technik (z.B. durch Bauteilfehler), systematische Fehler im Entstehungsprozess oder Fehlhandlungen von Personen. Safety charakterisiert also eine bestimmte Robustheit und Zuverlässigkeit im Gebrauch

[3] Beide Autoren sind Mitglieder dieser Arbeitsgruppe

[4] Allgemeines Eisenbahngesetz

[5] Eisenbahn-Bau- und Betriebsordnung

von Technik und dies auch im Rahmen zufälliger Fehlerereignisse (Fehlbedienung bzw. Technikversagen). Ziel ist hierbei der Schutz von „Mensch und Umwelt“ und dies gelingt unter anderem durch das Einnehmen von Ausfallzuständen im Fehlerfall (auf Kosten der Verfügbarkeit). Ein Beispiel im Eisenbahnbetrieb ist hier das „Fahren auf Befehl“, bei dem mit reduzierter Verfügbarkeit (eingeschränkte Leistungsfähigkeit) ein sicherer Bahnbetrieb möglich ist.

Im Gegensatz dazu geht man bei IT-Security in Anlehnung an die EN ISO/IEC 27000 [4] in der Regel von absichtlich erschaffenen Bedrohungen aus, die zielgerichtet auf Schwachstellen sogenannter Assets (zu schützende informationstechnische Werte und/oder Güter) einwirken und somit deren Schutzwerte (in erster Linie Vertraulichkeit, Verfügbarkeit oder Integrität) gefährden, wobei es Abstufungen Ressourceneinsatz und Motivation der angreifenden Instanzen gibt. In der Praxis kann dies unterschiedlichste Auswirkungen haben, z.B. komplette(r) oder teilweise(r) Ausfall bzw. Fehlfunktion des betroffenen Systems, ggf. mit Folgewirkungen auf weitere angeschlossene, nicht unmittelbar manipulierte Systeme. Bei der Betrachtung von IT-Security geht es also um den Schutz IT-basierter technischer Systeme vor zielgerichteten Einwirkungen von außen. Hierbei findet z.B. das Verbreiten von Schadcode durchaus zielgerichtet statt (z.B. abzielend auf einen bestimmten IT-System-Typ bzw. eine bestimmte Software), allerdings kann das Einbringen des Schadcodes in eine konkrete Anlage durch Personen auch unabsichtlich (d.h. ohne aktives Zutun der eigentlichen angreifenden Instanz) erfolgen (z.B. durch einen legitimierten Anlagenbediener mittels eines mit Schadcode infizierten USB-Sticks).

Aufgrund der fortschreitenden Digitalisierung und ergo Verwendung vernetzter IT-basierter operativer Betriebstechnik im Eisenbahnwesen sowie Kommunikation über offene Kommunikationsnetze müssen nun auch Aspekte der IT-Security berücksichtigt werden. Einher geht dies mit einer stetig zunehmenden Bedrohung durch Cyberangriffe sowie mit gesetzlichen Anforderungen aus dem IT-Sicherheitsgesetz [5]. D.h. IT-Security trifft auf Safety und bedingt enorme Herausforderungen für den Eisenbahnsektor.

Während „klassische“ Bahnsysteme mittels aufwendiger Verfahren für mehrjährige Lebenszyklen zugelassen (und somit quasi fixiert) werden, müssen IT-Security-Systeme ständig aktuell gehalten werden, um IT-basierte Gefährdungen zuverlässig beherrschen zu können.

Für die Inbetriebnahme von Eisenbahninfrastrukturen sind bereits heute stringente Prozesse der Zulassungsbewertung umzusetzen und die Ergebnisse in einem Sicherheitsnachweis (Safety Case) zu dokumentieren. Dieser berücksichtigt Safety-Anforderungen und wird während der Lebenszeit des betreffenden Systems üblicherweise nicht angepasst. IT-Security-Anforderungen werden bisher nur im geringen Maße berücksichtigt, sind aber normativ gefordert. In der DIN EN 50129 [6] ist in Kapitel 6.4 beschrieben, dass Bedrohungen der IT-Security (wörtlich „IT-Sicherheit“) in den Prozessschritten Risikobewertung und Gefährdungsbeherrschung behandelt werden müssen, sobald eine Auswirkung der IT-Security auf die Safety nicht durch einfache Argumentation ausgeschlossen werden kann. Maßnahmen zur Behandlung der IT-Security müssen im Sicherheitsnachweis oder durch Verweis aufgeführt werden.

In diesem Paper werden Lösungsansätze aufgezeigt für die bisher nicht geregelte Zulassung von IT-Security in Eisenbahnsystemen unter Einhaltung der gesetzlichen und regulatorischen Rahmenbedingungen. Betrachtet werden hierbei sowohl dedizierte IT-Security-Systeme als auch IT-Security-Funktionen in einem Gesamtsystem. Eine wesentliche Herausforderung ist hierbei eine – ggf. auch kurzfristig – erforderliche Anpassung des betreffenden Systems (z.B. Einspielen eines IT-Security-Patches) vorzunehmen, ohne dass diese einer Änderung des Sicherheitsnachweises bedarf.

Auf prozessuale bzw. technische Aspekte der IT-Security kann an dieser Stelle nicht eingegangen werden. Hierzu wird auf die umfassend verfügbare Fachliteratur verwiesen (z.B. [7] und [8]).

2 Sektorleitlinie

Die Prozesse der Zulassungsbewertung von Signal-, Telekommunikations- und Elektrotechnischen Anlagen (STE) gemäß EIGV[6] [9] sind aktuell bereits in der Sektorleitlinie [1] beschrieben. Eine solche Zulassung ist erforderlich für die Inbetriebnahme von STE-Anlagen bzw. zur Beantragung einer ‚*Genehmigung zum Inverkehrbringen und Verwenden*' (GIuV). Die Sektorleitlinie stellt hierbei eine nationale technische Vorschrift gemäß (AEG) [2] sowie der EIGV [9] dar.

Die Sektorleitlinie ist seit 01.09.2021 in Kraft und löst die Prozesse der NTZ[7] ab. Neben übergreifenden Festlegungen in Abschnitt 1 sind in Abschnitt 2 die Prozesse für signaltechnische Anlagen und in Abschnitt 4 für Elektrotechnische Anlagen beschrieben. Abschnitt 3 ist reserviert für Telekommunikationsanlagen (aktuell in Erstellung).

Übergreifend sind u.a. in Abschnitt 1 der Sektorleitlinie die wesentlichen Rollen im Rahmen der Zulassungsbewertung beschrieben (Freigabeverantwortliche (FGV) sowie Prüfsachverständige (PSV)), Regelungen bzgl. CSM-RA[8] sowie Schnittstellen zwischen Herstellern, Betreibern und Prüfinstanzen auf der einen und der Aufsichtsbehörde (EBA) auf der anderen Seite.

Zusätzlich geregelt werden in der Sektorleitlinie:

- Es wird festgelegt, dass für nicht signifikante Änderungen am Bahnsystem eine Prüfung durch einen anerkannten Prüfsachverständigen auf Einhaltung der Technischen Vorschriften ausreichend ist (Konkretisierung der Signifikanzkriterien für den Zulassungsbewertungsprozess).
- Sowohl eine Prüferklärung eines Freigabeverantwortlichen bzw. eine Prüfbescheinigung eines Prüfsachverständigen kann als Grundlage für eine Plan- und Abnahmeprüfung zu Grunde gelegt werden (Gleichstellung gemäß § 27 Abs. 3 EIGV [9]). Ein Antrag auf eine GIuV ist hierbei optional.
- Zulassungsbewertungsprozesse, bei denen keine inhaltliche Beteiligung des Betreibers notwendig ist, können ohne Beteiligung des Betreibers durchgeführt werden. Hierzu legt die Sektorleitlinie eindeutige Kriterien und Verfahren fest.
- Die Einbindung des Eisenbahn-Bundesamtes über Anzeigen ist festgelegt. Damit wird eine Information des EBA in der Funktion als Aufsichtsbehörde sichergestellt, insbesondere wenn kein Antrag auf eine GIuV gestellt wird.

Die Sektorleitlinie beschreibt die generischen Prozesse der Zulassungsbewertung von STE-Systemen und beschränkt sich ausschließlich auf die Abläufe in den Phasen Lastenheft, Pflichtenheft und Produkt. Diesen Phasen vor- oder nachgelagerte Aspekte (z.B. Systemdefinition, Betriebsprozesse) werden in der Sektorleitlinie nicht betrachtet. Zu berücksichtigen ist, dass Produktfreigaben o.ä. sowie Genehmigungen zur Inbetriebnahme nachgelagert erfolgen und ebenfalls nicht in den generischen Prozessen der Zulassungsbewertung der Sektorleitlinie betrachtet werden.

[6] Eisenbahn-Inbetriebnahmegenehmigungsverordnung

[7] NTZ: Neue Typzulassung von Signal-, Telekommunikations- und Elektrotechnischen Anlagen

[8] CSM-RA: Common Safety Method – Risk Assessment (gemeinsame Sicherheitsmethode für die Risikobewertung) 0

3 Zulassungsbewertung von IT-Security-Systemen

Für signaltechnische Systeme (z.B. Stellwerke), bei denen die Kommunikation ausschließlich in geschlossenen Kommunikationsnetzen gemäß DIN EN 50159 [10] stattfindet, kann gemäß den aktuellen Regelungen der Sektorleitlinie auf eine dedizierte Betrachtung hinsichtlich IT-Security verzichtet werden (siehe Kapitel 2.1.15 der Sektorleitlinie). Dies betrifft im Wesentlichen die sicherheitsrelevanten Einrichtungen der Bestandstechnik. Zusätzlich ist für nicht sicherheitstechnische Komponenten oder Systeme, die eine Verbindung zu sicherheitstechnischen Komponenten haben, ein Nachweis der Rückwirkungsfreiheit ausreichend. Eine dezidierte Betrachtung zur IT-Security ist im Rahmen der Zulassungsbewertung nicht erforderlich.

Für STE-Systeme und -Komponenten, die über offene Kommunikationsnetze gemäß DIN EN 50159 [10] verbunden sind, ist hingegen eine dezidierte Nachweisführung hinsichtlich IT-Security erforderlich. Entsprechende Regelungen sind bisher nicht vorhanden. Lösungsansätze bzgl. IT-Security zur Erweiterung der Sektorleitlinie – für offene Kommunikationsnetze – wurden im Rahmen einer Arbeitsgruppe zwischen Eisenbahn-Bundesamt, VDB (Verband der Bahnindustrie in Deutschland e.V.) und DB Netz AG erarbeitet. Die beiden Autoren sind Mitglieder dieser Arbeitsgruppe. Geplant ist die Beschreibung dieser Regelungen im neuen Abschnitt 5 der Sektorleitlinie.

Wesentlich für die Zulassungsbewertung sind die beiden nachfolgenden Definitionen gemäß DIN EN 50126 [12]:

- Sicherheitsfunktion ist eine Funktion, deren alleiniger Zweck die Sicherstellung der Sicherheit ist (auch als funktionale Sicherheit oder Safety bezeichnet)
- Systeme weisen sicherheitsbezogene Funktion auf, wenn durch deren Ausfall oder Störungen (z.B. mittels Manipulationen oder Cyberangriffe) der sichere Bahnbetrieb gefährdet sein kann.

Alle Sicherheitsfunktionen sind sicherheitsbezogene Funktionen, aber nicht alle sicherheitsbezogene Funktionen sind Sicherheitsfunktionen.

Grundlegend ist hierbei, dass Systeme der IT-Security selbst keine Sicherheitsfunktionen haben, aber sicherheitsbezogene Funktionen aufweisen können, indem z.B. durch vorhandene Schwachstellen mittels konkreter Bedrohungen auch Auswirkungen auf die Sicherheit des Bahnbetriebs möglich sind. Demzufolge werden Systeme der IT-Security selbst nicht mit einem Safety-Integritätslevel (SIL) gemäß DIN EN 50129 [9] beaufschlagt, sondern es müssen lediglich Maßnahmen zur Basisintegrität umgesetzt werden. Dies bedeutet, dass für solche Systeme kein Sicherheitsnachweis (Safety Case) gemäß DIN EN 50126 [12] notwendig ist, aber ein Nachweis bzgl. IT-Security, z.B. ein Cyber Security Case (CSC) gemäß DIN CLC/TS 50701 [13].

Für elektrotechnische Komponenten, Anlagen und Systeme (Bahnenergieversorgung mit den Frequenzen 50 Hz und 16,7 Hz sowie Gleichstrom), die über Kommunikationsprotokolle kommunizieren, sind die Aspekte der IT-Security im Rahmen der Zulassungsbewertung gesamthaft zu betrachten (keine getrennte Nachweisführung), da hier meist hochintegrierte Systeme vorhanden sind. Für die Bewertung und Beherrschung der Risiken in Bezug auf IT-Security findet die DIN CLC/TS 50701 [13] durch den Cyber Security Case in den Phasen Lastenheft, Pflichtenheft und Produkt Anwendung. Details sind bereits heute in Abschnitt 4 der Sektorleitlinie geregelt. Der geplante Abschnitt 5 findet hierauf keine Anwendung.

3.1 Wesentliche Rahmenbedingungen

Die nachfolgenden Rahmenbedingungen sind bei der Ausarbeitung der Lösungsansätze für IT-Security-Systeme oder -Funktionen wesentlich:

- Prozessabläufe im bewährten Drei-Phasen-Modell (Lastenheft, Pflichtenheft und Produkt)
- Einsatz von sogenannten COTS[9]-Komponenten (ggf. auch unterschiedlicher Hersteller), die nicht speziell für den Bahnsektor entwickelt werden
- Prozessuale und technische (logische!) Trennung von IT-Security und Safety inkl. Trennung der beiden Zulassungsbewertungsprozesse, um erteilte Safety-Zulassungsbewertungen nicht zu gefährden.
    - Dies beinhaltet auch getrennte Nachweisführungen und -dokumente (Safety Case und Cyber Security Case).
    - Damit ist eine getrennte Weiterentwicklung möglich.
    - Die Notwendigkeit zur Trennung ist in der sehr unterschiedlichen Veränderungsgeschwindigkeit der Systeme begründet (die Systeme der funktionalen Sicherheit entwickeln sich relativ langsam weiter, die Systeme der IT-Security auf Grund der sich laufend verändernden Bedrohungen wesentlich schneller). Somit ist in einem Gesamtsystem (z.B. DSTW[10]) eine schnelle Reaktion auf IT-Security-Änderungen möglich (z.B. Notwendigkeit zur zeitnahen Einspielung eines IT-Security-Patches als Maßnahme gegen eine kritische Schwachstelle), ohne dass erstellte Sicherheitsnachweise ihre Gültigkeit verlieren.
- Beschreibung der für IT-Security relevanten Rollen
- Berücksichtigung des „Stands der Technik“[11] für IT-Security (im Gegensatz zu den „anerkannten Regeln der Technik“[11] für safety-bezogene Systeme) als Prüfgrundlage für Prüfsachverständige IT-Security
    - Die große Dynamik im Bereich der IT-Security ebenso wie das IT-Sicherheitsgesetz erfordern eine Nachweisführung auf Basis des Stands der Technik.
- Beschreibung von Struktur und Inhalten des IT-Security-Nachweises gemäß DIN CLC/TS 50701 [13].
    - Nachweisführung, dass alle Anforderungen bzgl. IT-Security umgesetzt sind und Gefährdungen aus der IT-Security auf die Safety beherrscht werden
    - Dokumente im Cyber Security Case
- Anforderungen an einen Patch-Prozess (siehe Kapitel 3.7), sodass bei einer rein auf IT-Security bezogenen Software-Aktualisierung (z.B. Einspielen von IT-Security-Patches) in der Regel keine neuen Nachweise, weder für IT-Security und schon gar nicht für Safety, erforderlich sind
- Regelungen zur Kompatibilität bei Austausch von Komponenten oder Produkten
- Ggf. Berücksichtigung von Produktzertifikaten

[9] COTS: commercial-off-the-shelf: seriengefertigte, i.d.R. hochstandardisierte Produkte

[10] DSTW: Digitales Stellwerk

[11] Begriffe gemäß „Kalkar-Entscheidung“ des Bundesverfassungsgerichtes aus dem Jahr 1978

3.2 Trennung der Nachweisführung

Die logische Trennung der Zulassungsbewertung zwischen den Systemen der funktionalen Sicherheit (z.B. der Stellwerksfunktionalität) und denen der IT-Security (z.B. Security Gateways) wird im neu hinzukommenden Abschnitt 5 der Sektorleitlinie beschrieben.

Abb. 1 zeigt eine vereinfachte grafische Aufbereitung dieser Trennung bzw. den Zusammenhang der zugehörigen Prozesse. Der für die separate Nachweisführung erforderliche Anforderungsabgleich zwischen Safety- und IT-Security-Funktionen bzw. -Systemen erfolgt bereits in der Lastenheftphase. Hierbei werden Security-bezogene sowie ggf. funktionale Lasten (wie z.B. maximal zulässige Laufzeiten im Datentransport), die an einzubeziehende IT-Security-Systeme bestehen, anhand der allgemeinen Systemanforderungen identifiziert und in das Lastenheft des Security-Systems übernommen. So werden gegenseitige Abhängigkeiten adressiert. Die Funktionen werden aber im weiteren Verlauf in getrennten Prozessen betrachtet, wie in Abb. 1 in den beiden parallel verlaufenden generischen Prozessabläufen dargestellt (rechte Seite: Safety, linke Seite: IT-Security). Die nachfolgenden Zulassungsbewertungen der Safety- und Security-Funktionen bzw. -Systeme erfolgen separat voneinander auf Basis der jeweils relevanten Abschnitte der Sektorleitlinie.

Dies gilt auch für Systeme, die sowohl Safety- als auch IT-Security-Funktionalitäten kombiniert aufweisen; für derartige Systeme werden also ebenfalls separate Safety- und IT-Security-Zulassungsbewertungen erforderlich. Die für die Zulassungsbewertung sinnvolle Trennung zwischen Safety und IT-Security, muss in einem solchen Fall auf funktionaler bzw. logischer Ebene erfolgen.

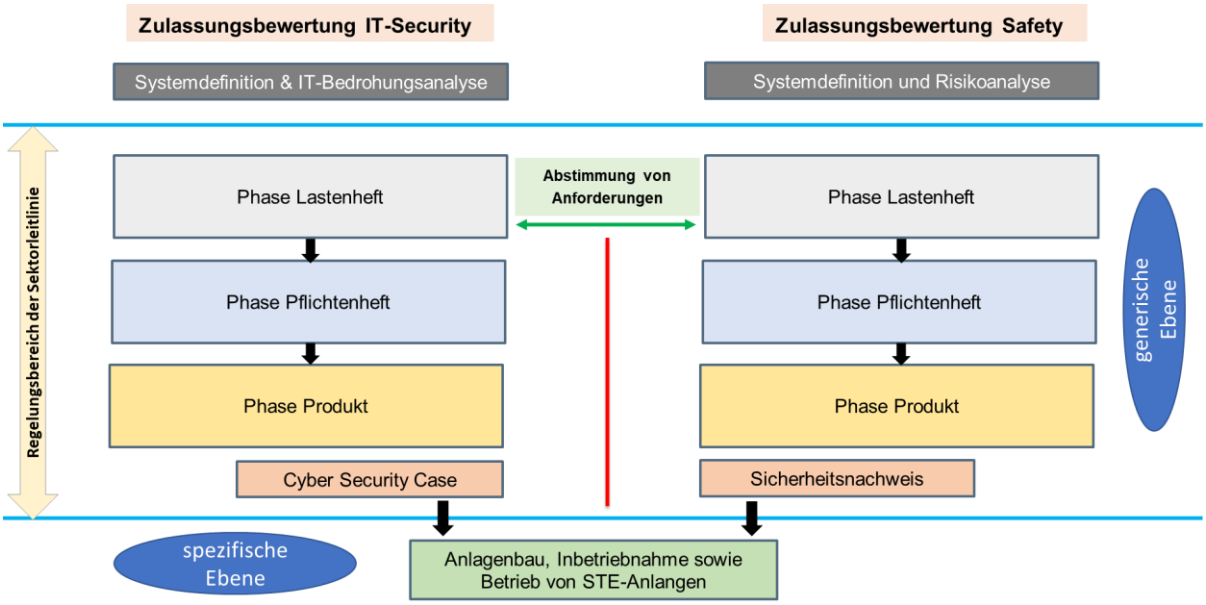


Abb. 1 Zusammenhang der Zulassungsprozesse Safety und IT-Security

Im nachfolgenden Abschnitt dieses Papers sind die wesentlichen Eckpunkte des derzeit noch im Review-Prozess befindlichen Abschnitts 5 der Sektorleitlinie für die Zulassungsbewertung von IT-Security-Systemen dargestellt.

3.3 Stand der Technik

Ein wesentlicher, technologiebedingter Unterschied in der Bewertung von IT-Security-Systemen gegenüber Signal- (S), Telekommunikations- (T) und Elektrotechnischen Anlagen (E) ist der zugrunde gelegte Technologiestand. Während in den bewährten Gewerken S, T und E im Eisenbahnsektor

üblicherweise „anerkannte Regeln der Technik“ (a.R.d.T.) als Beurteilungsgrundlage zum Einsatz kommen, wird im deutlich schnelllebigeren Bereich der IT-Security typischerweise der „Stand der Technik“ (S.d.T.) referenziert. Dieser berücksichtigt im Vergleich zu a.R.d.T. potenziell aktuellere technische Erkenntnisse, verfügt typischerweise jedoch über einen weniger gefestigten Erfahrungsschatz. Im Kontext IT-Security werden die resultierenden potenziellen Risiken jedoch durch deutlich flexiblere Handlungsmöglichkeiten ausgeglichen, insbesondere was Detektion und Abwehr von Cybergefahren angeht.

Als Prüfgrundlage für die Prüfsachverständigen IT-Security wurde eine Liste mit Referenzdokumenten erstellt, die als Technische Vorschriften zum Stand der Technik für die Zulassungsbewertung der IT-Security anzuwenden sind. Dies erleichtert auf der einen Seite die Arbeit der Prüfsachverständigen und erhöht auf der anderen Seite die Standardisierung.

3.4 Rollen und Verantwortlichkeiten

Neben den Zulassungsbewertungsprozessen werden auch die Schnittstellen zwischen den hauptsächlich beteiligten Organisationen und Rollen festgelegt. Die hauptsächlich Beteiligten sind:

- **Betreiber** und damit Assetverantwortlicher
  - primär bei Eisenbahninfrastrukturunternehmen (Bedrohungsermittlung, Anforderungsdefinition und Betrieb)
  - verantwortlich für den CSC
- **Komponentenhersteller**
  - Unternehmen, die einzelne, z.B. nicht bahnspezifische COTS-Komponenten zum Schutz gegen Cyberangriffe herstellen; z.B. Hersteller von Verschlüsselungskomponenten
- **IT-Security-Integrator**
  - Betreiber oder Hersteller
  - Unternehmen, das die Zusammenschaltung aller Komponenten der IT-Security, die Erprobung und den Nachweis für den ausreichenden Schutz also der Erfüllung der Anforderungen erbringt (nicht zu verwechseln mit dem Gesamtsystemintegrator!)

3.5 Ausprägung der Phasen Lastenheft, Pflichtenheft und Produkt

Die Zulassungsbewertung der IT-Security orientiert sich gemäß aktuellem Entwurf an den drei Phasen Lastenheft, Pflichtenheft und Produkt (wie bereits auch für die sonstigen in der Sektorleitlinie betrachteten Gewerke festgelegt). Hinsichtlich der Inhalte und Abläufe der einzelnen Phasen bestehen Parallelen insbesondere zur signaltechnischen Zulassungsbewertung, z.B. im Hinblick auf die Beauftragung (und deren Anzeige beim EBA) von Freigabeverantwortlichen und Prüfsachverständigen, ggf. nötige Signifikanzentscheidungen und zugehörige Prozesse im Falle neuer Anforderungen oder Regelwerksabweichungen sowie die jeweils abschließende Erstellung von Teil-Prüferklärungen, Inspektionsberichten, Gutachten etc.

Ein wesentlicher Unterschied zu anderen Gewerken besteht pro Phase in der Notwendigkeit der Auskopplung von Dokumenten bzw. Nachweisen für den Cyber Security Case (siehe Kapitel 3.9). Des Weiteren erfolgt im Rahmen der Pflichtenheftphase die Überprüfung, ob das zuzulassende IT-Security-System bzw. -Produkt voraussichtlich während des laufenden Betriebs mit IT-Security-Patches versorgt

werden muss. Ist dies der Fall, wird ein generischer „Teilprozess Patch“ angestoßen, der zur Erstellung eines individuellen Patch-Prozesses führt (siehe Kapitel 3.7).

Für kleinere Projekte (z.B. Austausch einzelner IT-Security-Produkte aus Obsoleszenzgründen oder Änderungen lediglich einer Funktion in einem bereits genehmigten IT-Security-System) besteht nach Abstimmung mit dem Prüfsachverständigen die Möglichkeit der Zusammenfassung von Pflichtenheft- und Produktphase.

3.6 Plan- und Abnahmeprüfungen

Die nachfolgenden Bauprozesse (Prozess gemäß VV BAU-STE [14]) können unverändert beibehalten werden. Die betreffenden Unterlagen für die IT-Security (Planungs-, Projektierungs-, Prüfunterlagen, etc.) werden dahingehend geprüft, dass diese ausreichen, um bisherige Plan- und Abnahmeprüfer für eine Anlage einsetzen zu können. Die entsprechende Qualität der Unterlagen ist neben dem IT-Security-Integrator auch durch den Prüfsachverständigen IT-Security, der die Zulassungsbewertung prüft, sicherzustellen. Damit ist der Einsatz der aktuell anerkannten Plan- und Abnahmeprüfer weiterhin möglich.

3.7 Patch-Prozess

Im Rahmen der Sektorleitlinie wurden Anforderungen an einen generischen Patch-Prozess definiert, sodass bei einer rein auf IT-Security-Funktionen bezogenen Software-Aktualisierung (z.B. Einspielen von IT-Security-Patches) in der Regel keine neuen Nachweise, weder für IT-Security und schon gar nicht für Safety, erforderlich sind. Der resultierende systemspezifische Patch-Prozess ist Bestandteil der Zulassungsbewertung. Er wird parallel zur Pflichtenheft- und Produktphase durch den IT-Security-Integrator, vorzugsweise unter Beteiligung des Betreibers, erstellt, durch den Prüfsachverständigen geprüft und im Cyber Security Case dokumentiert. Ziel des Patch-Prozesses ist es, unabhängig von der Zulassungsbewertung unter den im Prozess zu definierenden Rahmenbedingungen, Möglichkeiten für das Patching (z.B. für die Beseitigung von Schwachstellen in IT-Security-Systemen und -Produkten) vorzusehen; das Patching selbst benötigt in diesem Fall keine Zulassungsbewertung. Die Freigabe für einen Patch ist damit wesentlich kurzfristiger möglich als die Erstellung einer Prüfbescheinigung oder Prüferklärung für eine Produkt- oder Systemänderung. Wesentliche Voraussetzung hierfür ist, dass der Patch keinerlei Auswirkung auf Safety-Funktionen hat, was im Rahmen des Patch-Prozesses definiert wird und nachzuweisen ist. Mit der geplanten Vorgehensweise ist nicht nur ein Patch relativ schnell freigebbar, der Patch ist auch schnell auf den entsprechenden Anlagen implementierbar (z.B. kein Bauprozess, kein Abnahmeprüfer erforderlich).

3.8 Austausch von Produkten/Komponenten

Bei Notwendigkeit des Austauschs einzelner Produkte/Komponenten der IT-Security ist durch einen Prüfsachverständigen ein Nachweis der Erfüllung aller Anforderungen im Rahmen der Produktrealisierung zu bewerten. Dies beinhaltet u.a. Prüfung auf Einhaltung der relevanten Regelwerke und technischen Vorschriften. Der Prüfsachverständige bestätigt hierbei, dass das neue Produkt bzw. die neue Komponente kompatibel zum Gesamtsystem ist, ggf. durch Überprüfung der Abwärtskompatibilität zu einem vorhandenen Vorgängerprodukt/Vorgängerkomponente. Somit darf das neue Produkt bzw. die neue Komponente ohne weitere Prüfung auf Basis des jeweils aktuellen IT-Security-Patch-Standes verwendet werden.

Sofern entsprechende Nachweise vorliegen, kann durch die Sicherstellung der Abwärtskompatibilität gewährleistet werden, dass Komponenten/Produkte, die in vorhandenen Kennblättern genannt sind, ausgetauscht und/oder verwendet werden können. Um dieses formal zu ermöglichen, ist in einem Kompatibilitätsblatt die Abwärtskompatibilität zu der zu ersetzenden Komponente/des Produktes ggf. darzustellen. Das ursprüngliche Kennblatt im Zusammenhang mit dem Kompatibilitätsblatt behält hierbei seine Gültigkeit.

3.9 Cyber Security Case (CSC)

Für den gemäß DIN EN 50129 [6] geforderten Nachweis der IT-Security wird gemäß DIN CLC/TS 50701 [13] ein Cyber Security Case erstellt. Es erfolgt hierbei eine Orientierung am Stand der Technik (siehe Kapitel 3.3).

Zusätzlich ist auf Grund der sicherheitsbezogenen Funktion des Teilsystems IT-Security die Basisintegrität gemäß DIN EN 50126 [12]. zu Grunde zu legen. Im Gegensatz zum statischen Sicherheitsnachweis (Safety Case) erstreckt sich die Erstellung des Cyber Security Cases über den gesamten Lebenszyklus des Betrachtungsgegenstandes. Damit kann der Cyber Security Case nicht vollständig im Rahmen der Zulassungsbewertung dieser Sektorleitlinie erstellt werden. Vielmehr beginnt die Erstellung zeitlich bereits vor dem Start der Prozesse dieser Sektorleitlinie und endet nicht mit dem Abschluss des Zulassungsbewertungsprozesses, sondern zieht sich bis zum Ende der Betriebsphase. Dieses bedeutet konkret, dass die Dokumentation zum Cyber Security Case bereits im Rahmen der Systemdefinition und der IT-Bedrohungsanalyse (z.B. im Rahmen des ISMS[12] beim Betreiber) begonnen (also bereits vor den in der Sektorleitlinie beschriebenen Phasen), die Dokumentation über die drei Phasen der Zulassungsbewertung vervollständigt und über den Betrieb des Produktes gepflegt werden muss. Über die IT-Bedrohungsanalyse werden u.a. die Lastenheftanforderungen ermittelt. Die Gesamtverantwortung für den CSC trägt der Betreiber (Assetverantwortlicher).

Im Cyber Security Cases wird im Wesentlichen die Nachweisführung bzgl. der nachfolgenden Themen dokumentiert:

1. Gefährdungsbeherrschung inkl. Rückwirkungsfreiheit der IT-Security auf die betriebliche Sicherheit (keine Auswirkungen Security auf Safety) sowie
2. Erfüllung der allgemeinen Ziele und Anforderungen der IT-Security nach dem Stand der Technik (z.B. notwendige Schutzfunktionen, Berücksichtigung von aktuellen Cyberbedrohungen, Erkennung von Unregelmäßigkeiten).

Der erste Aspekt ist bereits beim Einbau/im Rahmen der Grundkonfiguration erforderlich, der zweite Aspekt ist im laufenden Lebenszyklus sicherzustellen und ggf. dynamisch anzupassen (dies bedingt dann die Fortschreibung des Cyber Security Cases).

Im Gegensatz zum Sicherheitsnachweis ist für den Cyber Security Case in seiner Gesamtheit selbst keine Begutachtung durch Prüfsachverständige oder eine Unabhängige Bewertungsstelle erforderlich. Die Bestandteile und einzelne Dokumente des Cyber Security Cases werden im Rahmen der Zulassungsbewertung ohnehin durch einen Prüfsachverständigen begutachtet.

Im Cyber Security Case sind beispielsweise die nachfolgenden Dokumente, Berichte oder Verweise enthalten (keine abschließende Aufzählung):

- Lastenheft sowie lastenheftbegleitende Dokumente,

[12] ISMS: Information Security Management System

- Nachweis- und Bewertungsdokumente,
- Pflichtenheft sowie pflichtenheftbegleitende Dokumente,
- produktbegleitende Dokumente,
- Inspektionsberichte/Prüfbescheinigungen zum Pflichtenheft bzw. Produkt sowie
- Erklärung zur Beherrschung der IT-Security-Risiken

4 Zusammenfassung

Dieses Paper beschreibt eine Prozessmethodik und zeigt Lösungsansätze auf, wie Integration und Aktualisierung von IT-Security-Funktionen in zuzulassenden Bahnsystemen ohne Anpassung der Safety-bezogenen Sicherheitsnachweise ermöglicht werden können.

Aktuell befinden sich die ergänzenden Ausführungsbestimmungen im Review-Prozess zwischen EBA, VDB und DB. Abschließend werden die aufgezeigten Ansätze finalisiert und sollten anschließend in Anwendungsprojekten verprobt werden, um ihre Praxistauglichkeit nachzuweisen. Geplant ist die Finalisierung und Inkraftsetzung der fortgeschriebenen Sektorleitlinie bis Ende 2023; sie steht dann allen Beteiligten im Sektor Eisenbahninfrastruktur zur Verfügung. Für zukünftige Inbetriebnahmen ist damit die erforderliche Rechtssicherheit bzgl. IT-Security gewährleistet.

5 Literaturverzeichnis

[1] Sektorleitlinie für die Zulassungsbewertung von Signal-, Telekommunikations- und Elektrotechnischen Anlagen (Technische Vorschrift), Ausgabe 1.0 vom 07.07.2021. Available at: https://www.eba.bund.de/SharedDocs/Downloads/DE/Infrastruktur/AllgemeineVorschriften/VV_GIuV/Sektorleitlinie/22_Sektorleitlinie.html (Accessed: 03.03.2023)

[2] Allgemeines Eisenbahngesetz (AEG), 27.12.1993 und zuletzt geändert am 10.09.2021. Available at: https://www.gesetze-im-internet.de/aeg_1994/ (Accessed: 03.03.2023)

[3] Eisenbahn-Bau- und Betriebsordnung (EBO); zuletzt geändert im April 2019. Available at: https://www.gesetze-im-internet.de/ebo/ (Accessed: 20.03.2023)

[4] DIN EN ISO/IEC 27000:2020 Informationstechnik - Sicherheitsverfahren - Informationssicherheits¬managementsysteme - Überblick und Terminologie (ISO/IEC 27000:2018)

[5] Gesetze zur Erhöhung der Sicherheit informationstechnischer Systeme des BSI: IT-SiG 1.0 von 2015 sowie das IT-SiG 2 von 2021. Available at: https://www.bsi.bund.de/DE/Das-BSI/Auftrag/Gesetze-und-Verordnungen/IT-SiG/it_sig_node.html (Accessed: 03.03.2023)

[6] DIN EN 50129: 2019: CENELEC: Bahnanwendungen - Telekommunikationstechnik, Signaltechnik und Datenverarbeitungssysteme, Sicherheitsbezogene elektronische Systeme für Signaltechnik

[7] Nobert Pohlmann: Cyber-Sicherheit: Das Lehrbuch für Konzepte, Prinzipien, Mechanismen, Architekturen und Eigenschaften von Cyber-Sicherheitssystemen in der Digitalisierung, Springer Vieweg, 2022

[8] Claudia Eckert: IT-Sicherheit: Konzepte – Verfahren – Protokolle, De Gruyter Studium, 2023

[9] EIGV: Verordnung über die Erteilung von Inbetriebnahmegenehmigungen für das Eisenbahnsystem (Eisenbahn-Inbetriebnahmegenehmigungsverordnung), Stand 26.07.2018. Available at: https://www.gesetze-im-internet.de/eigv/ (Accessed: 03.03.2023)

[10] DIN EN 50159: 2010: CENELEC: Bahnanwendungen - Telekommunikationstechnik, Signaltechnik und Datenverarbeitungssysteme, Sicherheitsrelevante Kommunikation in Übertragungssystemen

[11] VO CSM-RA: Durchführungsverordnung (EU) Nr. 402/2013 zuletzt geändert durch Nr. 2015/1136

[12] DIN EN 50126: 2017: CENELEC: Bahnanwendungen - Spezifikation und Nachweis von Zuverlässigkeit, Verfügbarkeit, Instandhaltbarkeit und Sicherheit (RAMS) – Teil 1: Generischer RAMS-Prozess

[13] DIN CLC/TS 50701 (VDE V 0115-701), Bahnanwendungen – IT-Sicherheit, Version November 2022

[14] VV BAU-STE: Verwaltungsvorschrift für die Überwachung der Erstellung von Signal-, Telekommunikations- und Elektrotechnischen Anlagen, Ausgabe 5.1, 05.07.2020

A formal approach to developing rail control software – Using automation, formal specifications, and digital twins

Gunnar Smith[1]

[1] Prover Technology

1 Introduction

Projects delivering rail control systems, including software, often suffer from budget overruns and missed schedules. This is especially true for brownfield projects. Common root causes for this include a lack of understanding and documentation of existing systems and interfaces, vague and ambiguous requirement specifications, together with outdated tools, processes, and methods for development and verification of the systems.

In this paper we will present a solution for how this situation can be overcome, by using a formal approach that focuses on getting the requirements right from the start, and then applying automation tools to develop, test, and verify the software based on the requirements. We will illustrate how this can be realized, and what the benefits are, by looking at a case study from Stockholm Metro.

2 Background: Formal Methods

Formal methods are often defined as *using rigorous mathematical languages, techniques, and tools for specification, design, and verification of software and hardware systems.* There are many success stories of applying formal methods in various industries, but it is often perceived as a complex and costly approach that requires specific skills from the engineers applying the tools and methods. For this reason, such methods are mostly used in the development of safety critical applications, such as control systems within rail or avionics, or in applications where errors that are not found before production may become very costly to resolve, as is often the case for e.g., integrated circuits.

The most common applications of formal methods, at least within the rail control domain, are within the specification and verification phases, with the goal of ensuring the correctness (and safety) of the developed system. In this paper we will also present how formal methods can be used to bring more automation to the implementation phase, thus also improving delivery capacity and decreasing costs.

2.1 Formal Specification

Formal Specification aims at improving the quality of specifications and systems by providing more insight and understanding of the requirements before the systems are built. Typically, the formal specification process starts with a set of properties expressed in an informal language, these properties are then formalized in a formal language. A key property of a formal language is that it has a precise definition of its syntax and semantics, so that specifications in this language can be processed automatically by software tools, e.g., to model the requirements for debugging and validation. Having a formal definition of the requirements is also a prerequisite for performing *Formal Verification*.

[1] gunnar.smith@prover.com

2.2 Formal Verification

Formal Verification is a method to prove with 100% certainty that some explicitly defined requirements are fulfilled by a system. You build a mathematical model of the system and formalize your requirements in a logic language, then you let a computer use mathematical algorithms to prove that your system model satisfies your requirements. You reach 100% certainty because you do not check your requirements scenario by scenario, you check them with mathematical proof. You may compare this to the certainty you have about the sum of the angles in a triangle always being 180 degrees – you do not need to verify this by measuring all possible triangles, because there is a mathematical proof that guarantees that this will always be the case, also for triangles we never encountered before. This, the ability to *prove* the absence of errors, is the main benefit of Formal Verification over traditional verification methods such as testing or manual reviews.

2.3 Formal Design and Development

Formal methods can also be applied in the system implementation phase, using software tools that automate various design and development tasks, based on the formal specifications. One example of this is *code generation* where detailed implementations can be generated from high-level formal specifications and configuration data.

The implementation process can also play an important role for the success of formal verification in a project. Regardless of the exact process for the implementation, it is important that it is developed with the formal specification in consideration, to ensure that the formal verification will succeed. It is also often a good idea to use formal verification during the implementation, both to guide and assist the implementation and to simplify the verification phase.

3 Challenges in delivery of rail control software

Unfortunately, many rail control projects, especially brownfield projects, suffer from delays and budget overruns before commissioning. There are even several examples where major rail control projects have been cancelled by the buyer organization due to failure to complete years after project start. It is also not uncommon that there is an extended period of trouble shooting and debugging of system functionality even after commissioning.

Figure 8. Issues related to the specification and development of rail control systems are often manifested in delayed installations and an extended period before all desired functionality is in place.

A summary of these challenges and issues from the buyer perspective are:

1. **Unpredictable schedules and costs.** Over the course of the development process, systems can turn out to be more costly to procure, develop and maintain than expected. System errors and omissions are discovered late in the process. Or perhaps the lifespan of a system isn't as long as hoped for compared to more traditional systems, resulting in increased life cycle costs.

2. **Lack of control over your systems.** Many infrastructure managers experience a general lack of control over their systems. This is often due to being locked to a chosen vendor over a long period, with the vendor sitting on all the knowhow of the system. This may make it difficult and overly complicated to replace or modernize parts of the system down the line.
3. **Dominant issues in the industry.** The industry at large hasn't fully evolved to adopt a more modern and future-proof production process that is efficient and takes the specification, development, verification, validation, and the commissioning of systems into consideration. There is also the prevalence of an oligopoly of system suppliers who deliver complete systems without consideration for open standards and interfaces. While there is a number of ongoing activities to address this, these have yet to make significant impact in the industry, especially on the supplier side. Examples of such activities include the EULYNX standardization initiative and the work within Shift2Rail and its continuation Europe's rail (these also consider the use of formal methods). There is also a growing interest in alternative signalling solutions built on commercial-off-the-shelf (COTS) components, such as industry standard safety PLCs, supported by a group of smaller suppliers.

This situation was discussed at the *Shift2Rail Innovation Days (Oct 22, 2020), TD2.7 Formal Methods and Standardisation for Smart Signalling Systems,* and it was concluded that its main root causes are:

Root cause #1: Tender requirements tend to be vague.

- Leads to design choices whose drawbacks are understood late (expensive to change).
- Design choices lead to systems based on proprietary solutions ("vendor lock-in")

Root cause #2: Verification based on traditional methods.

- Time-consuming and error-prone (review, test).
- Verification coverage poor; can only detect issues, not prove absence of issues.
- Vague and imprecise requirements complicate verification.

Root cause #3: Lack of standard architecture and interfaces.

- Duplicates efforts (multiple suppliers and systems).
- Prevents or limits reuse (in V&V), "plug-and-play" and use of commercial off-the-shelf (COTS) components.

3.1 Requirement Specifications

Of all the steps involved in the procurement and development of a rail control system, writing the specifications that you will later build your system to is one of the most consequential when it comes to determining how smooth and successful the rest of your project will be. Any errors or omissions in your specifications will have an impact on the later steps in the development process. And since they were introduced early and discovered late, such errors will be costly to correct. Clearly specifying your needs, requirements, and expectations is a critical first step in the process.

Common specification issues include:

- **Incomplete:** Specified requirements fail to mention necessary parts of the system or lacks sufficient details.
- **Inconsistent:** Specified requirements are contradictory, making realization impossible.
- **Incorrect:** Specified requirements are simply wrong.
- **Ambiguous:** Specified requirements are open to multiple interpretations.

When you create specifications for a tender of a rail control system it is important that you focus on the right things; some aspects of the system may be important that you specify (to get a system that meets your expectations) whereas others may be better left to the suppliers to detail, so that they can provide the optimal solution at the best price, based on their generic products. So, while some parts of the specifications may need to be very detailed, it is also important not to over-specify the system.

High-quality specifications make it possible to:

- Procure the best solution and service for the best price.
- Efficiently handle change requests.
- Minimize risk for project delays.
- Validate that the delivered system meets requirements and expectations.
- Automatically prove safety and minimize on-site testing.
- Simplify and automate the development process.
- Simplify maintenance and upgrades after commissioning (addressing the life cycle of the system).

Developing such high-quality specifications can be a difficult task, especially for complex systems that need to be integrated in an existing environment that may not be fully understood or specified. One way to deal with this is to build a model of the systems and their environment and use this model to validate the specifications. This model can include abstractions as well as very detailed models, but it can also interface directly to existing implementations and sub-systems (hardware and software). At Prover, we refer to such a model of a rail control system as a *Digital Twin*.

4 Digital Twins for Rail Control Systems

A digital twin in general can be described as a virtual, interactive model of a real physical system, asset, or process, including its real-time characteristics and behaviors. Applied to the railway sector, a digital twin may encompass the entire infrastructure – from stations, rolling stock, and signals, to the coordinating IT systems. At Prover we focus on modelling various parts of rail control systems as digital twins, supporting the complete process, covering development of concepts and prototypes, requirement definition, implementation, verification, and maintenance. The biggest impact is perhaps on the earlier

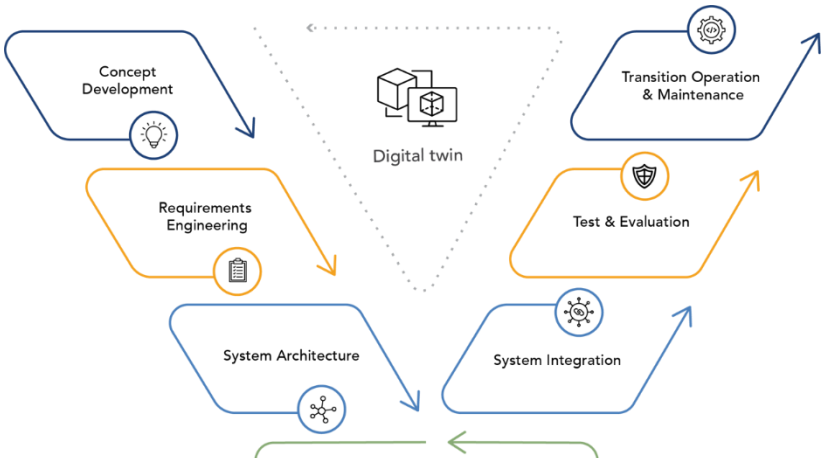


Figure 9. The Digital Twin supports the entire V-model, from concept to operation and maintenance.

stages, where the Digital Twin gives the opportunity to validate the requirement specifications already before tendering the system.

The digital twin is best developed with an iterative, test-driven, and agile process based on formal methods and automation. Formal methods enable the development of specifications, digital twins, and actual systems, using a set of automation tools and processes that we call *Signaling Design Automation* (SDA). Being formal, such methods offer enough detail for the digital twin to be analyzed for completeness, consistency, and correctness. This process can be summarized in the following three steps:

1. **Gather and analyze input.** The process starts with an analysis of the needs and information available, which helps define the test and safety requirements on a high level. Tender requirements, use cases, legacy systems, applicable standards, interfaces, rules, regulations, and project scope all provide input to this task.
2. **Formulate requirements and define the Object Model.** Formulate test cases and safety requirements in natural language using an Object Model that defines the objects in the system and how they interact. The test and safety specifications are then used to develop the design specification, for the implementation of the system. The specifications are formalized so that they can be interpreted by automation tools for configuration, design, implementation, testing, and verification.
3. **Configure your applications and validate the requirements.** Specify the configuration data that will be used to create instances, or Specific Applications, of the Object Model. The formalized requirement specifications are then validated with formal verification and simulation using the Specific Applications. Specifications are then updated based on these results, in an iterative and agile process.

5 Signaling Design Automation

As discussed, formal methods typically refer to the use of mathematical methods to describe, analyze and develop software systems. They have been in use for the verification of safety aspects of rail control systems for 20 or more years. However, there is room for improvement – formal methods can do more! There are many tools based on formal methods that you can use for various specific tasks. But the complete processes often sit with the suppliers and are bespoke to each respective one. This doesn't maximize the value for the infrastructure manager or the buyer side of these projects or systems. What could make the process complete and fill in the value gaps is automation. This is where Signaling Design Automation, or SDA, comes in. This process does not replace formal methods, but rather provides a way to integrate them into an automated development process.

Signaling Design Automation combines automation with the tools that formal methods offer to ensure that you can efficiently develop rail control systems and verify that they meet requirements during Verification and Validation (V&V). Automation tools can be used to your advantage at every phase of the software development process – from establishing requirements to system development and maintenance.

Capturing principles and formalizing requirements for a set of systems that are to be developed for multiple configurations are key concepts in the Signaling Design Automation process. This includes a clear separation of different types of requirements (configuration, interfaces, safety, and test). SDA also promotes the use of open and standardized formats and interfaces. As a result, infrastructure managers

will gain more control over their systems and acquire the ability to quickly identify and debug requirement issues before they become a major problem.

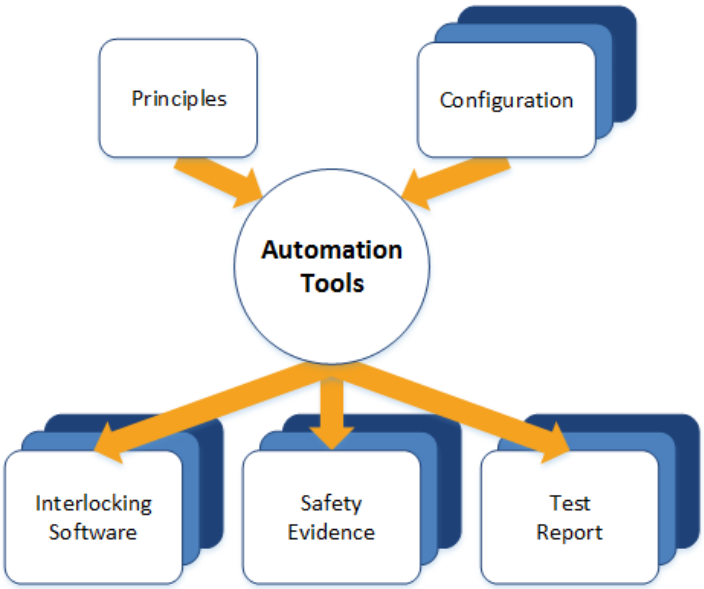


Figure 10. The SDA process uses software tools based on formal methods to automate the development, test, and verification processes for rail control software.

The Signaling Design Automation process can be summarized as consisting of the following four parts:

1. **Generic requirement specifications.** Generic requirements for a given type of system are specified and formalized, with a clear separation of object model (interfaces), design, safety, and test requirements. Design, test cases, and safety requirements are generated for individual, specific, applications.
2. **Configuration.** The data is split into Generic and Specific configuration data. From there, you can configure the individual signaling system that you are developing. Automation tools will take care of creating the specific requirements from your specifications for a particular configuration.
3. **Automated verification and validation.** The safety requirements are formally verified, and test cases simulated. Detailed verification and test reports are generated.
4. **Code generation.** Code generation tools convert your design into software code to be compiled, configured, and installed on the computing platform of the rail control system. This is the software that will be executed on the actual system in the field.

The key factor to the success of this process is that it provides a pipeline of automated steps which significantly reduces complexity and required effort. It capitalizes on knowledge by reusing structured requirements, then follows an iterative process (Specify, Configure, Test, and Verify) to develop and verify software applications with minimal effort. The high degree of automation, and the fact that it is very easy to evaluate the impact of changed requirements, or new functionality, also makes it the ideal process for developing prototypes, or digital twins.

6 Case study: Stockholm Metro

Stockholm Metro's system consists of a central traffic management system (TMS) and a relay-based vital and non-vital signaling system distributed over several relay rooms. The plan is to replace the mechanical control panels in the TMS used by the operators today with a modern computerized traffic management system, as well as to replace the non-vital relays with a computerized (PLC) solution, and, at least for the time being, leave the vital relays unchanged.

In order to analyze the existing system and validate this approach, a digital twin of the system was developed using formal methods. These formal methods included formal specifications with an emphasis on interfaces, separation of configuration data and generic requirements, automated simulation-based testing, and formal safety verifications.

In this case, the Stockholm metro had three primary goals. The first was to discover any dependencies that would make it difficult to leave the vital relays untouched when introducing new functionality in the overall system. The second was to identify any safety critical functions that are dependent on the existing non-vital relay or control panel design, to include such requirements in the specifications for the new systems. For example, physical wiring may prevent certain commands from being received simultaneously, or safety standards and best practices may have changed since the original system was commissioned. And finally, they wished to try out the concept to avoid any surprises that could delay the project and make it more expensive.

At the start of this project, the Stockholm Metro's architecture consists of the following components:

- Physical maneuver panel – Push/pull buttons and switches for controls, lamps for indications.
- Non-vital relays – Interface between panel and interlocking, with additional (non-safety) logic.
- Vital relays – Safety related signaling logic, locking of routes and points, signal aspects.

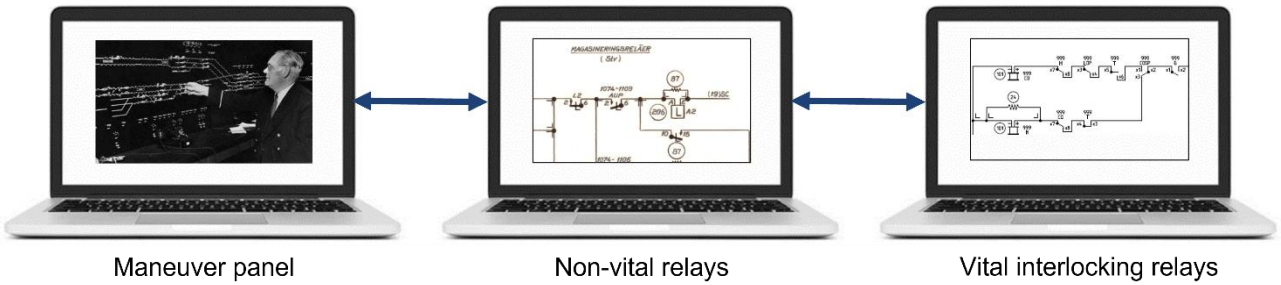


Figure 11. Digital Twin of the existing system.

The end goal of the project is a future architecture consisting of:

- A computerized traffic management system with an operator interface.
- A set of PLCs with (more or less) the same functionality as the current non-vital relays.
- The vital relays from the original system left unchanged, interfacing with the PLCs.

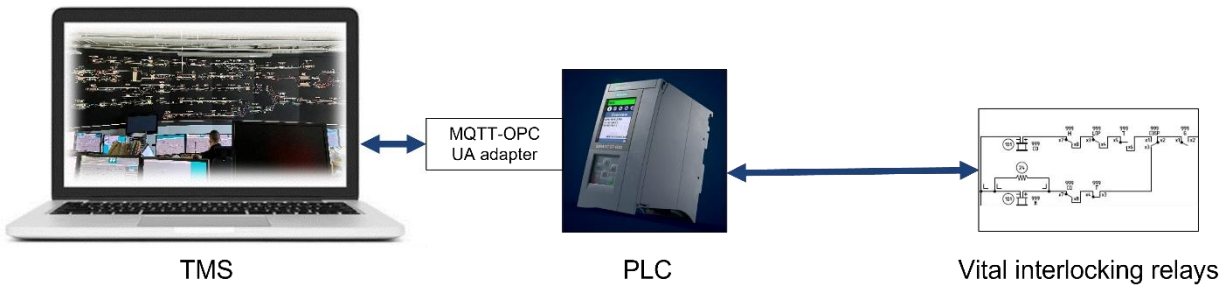


Figure 12. In the last phase of the project the Digital Twin of the future system was connected to the vital interlocking relays in the field.

Four different levels of the system were modeled with digital twins, and in each level the functionality, requirements and interfaces were validated with simulation and formal verification:

1. **A digital twin of existing system.** The SDA process was used to build the models for the maneuver panel and the non-vital relay system, and the model of the vital interlocking relays was generated directly from the (digitized) relay schematic drawings.
2. **A digital twin of the future system.** The model for the maneuver panel was extended with functionality of the future system to model the TMS. The non-vital relay model was replaced with a model of a PLC, with the additional functionality required by the new system.
3. **PLC connected to the model.** Replaced the model of the PLC with an actual PLC, connected to the models of the TMS and the vital relays.
4. **Vital relays connected to the model.** Replaced the model of the interlocking with the actual relay system. In this phase the model was connected to the relays in the field and tests were performed overnight.

Following this process, the Stockholm Metro was able to successfully validate the modularity of the proposed system and establish that it is feasible to replace the non-vital relays with a PLC system. They were also able to identify a few unexpected details in the existing system. Furthermore, the digital twin validation approach simplified the transition to the physical environment relay system; testing with the PLC installed in the field was conducted over the course of a single night.

In the future, the Stockholm Metro can use the digital twin for preparing tenders, validation and verification, reproducing issues from field error reports/logs, replacing testing in the field, training, and documentation.

7 Conclusions and Discussion

Our experience from working with Stockholm Metro and other clients shows that the proposed process indeed delivers the expected benefits of reduced effort and time for development and verification as well as increased quality and predictability. Formal verification can also raise the level of trust in the system safety, and automation can reduce the number of software issues found in commissioned systems.

With the right tools and processes, applying formal methods need not be as daunting as it is often perceived, it is simply a good engineering practice among others. Engineers working with the tools do not need to have a deep knowledge of the theories they are based on, basic training on the tools and some degree of knowledge of logic and software engineering is all it takes.

While formal methods have been used in the rail control domain for more than 20 years and is being promoted by some of the world's leading infrastructure managers and signalling suppliers, it still has a long way to go before it can be considered mainstream. This has maybe less to do with the technology itself than with the, often conservative, mindset of the community – more education and dissemination of success stories and best practices are needed to address this. There is also more work to be done on providing, and promoting, commercially available and standardized languages and tools, to make formal methods easily accessible for the rail control community.

The use of digital twins to model systems is particularly valuable when new systems need to be integrated in an existing environment, that is not always well documented or understood. It can be a great help to increase the knowledge of the systems that you, as an infrastructure manager, have installed, thus reducing dependency on single vendors, and reducing life cycle costs.

How can we further expand the digital twin solution in the context of rail control systems, to connect it to other areas such as Building Information Modelling and predictive maintenance? What would be the benefits? This is something that will have to be explored in the communities of rail control and digital twins. There are a number of digital twin applications and projects emerging within the rail domain, that may help answer these questions. What is certain is that there is a need to use more modelling to increase the value of the huge investments that are made in railway infrastructure, including control systems and signaling.

In this paper we have argued for the need for infrastructure managers to take more control of their systems and increase their knowledge of them. The purposes of doing this are to make them better buyers that can clearly communicate their needs to the suppliers and to increase standardization and give more modularized systems with reduced life cycle costs. However, it also comes with a risk of overspecification, resulting in more costly customer adaptations of generic supplier solutions. More standardization and collaboration between suppliers and infrastructure managers seems to be the way forward.

Lebensverlängernde Massnahmen und Upgradability als Werkzeuge zur netzweiten Migration auf Führerstandsignalisierung in der Schweiz

Melvin Zinngrebe[1]

[1] SBB AG

1 Einleitung

Wie in anderen europäischen Ländern ist die Landschaft der Sicherungsanlagen und Zugbeeinflussungssysteme in der Schweiz historisch gewachsen. Neben modernen elektronischen Stellwerken, die bspw. auf dem Transeuropäischen Korridor (TEN) Rotterdam-Genua mit dem European Train Control System Level 2 (ETCS L2) zum Einsatz kommen, sichern über die Hälfte der Stellwerke in der Schweiz mittels Relaistechnik.

Mit der Strategie „European Rail Traffic Management System“ (ERTMS) hat das Bundesamt für Verkehr im Jahr 2021 die Weichen in Richtung des netzweiten Rollouts der Führerstandsignalisierung (FSS) gestellt.

In diesem Paper werden zwei Ansätze vorgestellt, mit denen die SBB AG den Herausforderungen dieser Migration begegnet: Mit der Lebensverlängerung von elektronischen Stellwerken der ersten Generation werden Investitionen in eine Zwischengeneration vor der FSS-Migration vermieden. Gleichzeitig werden die aktuell nicht FSS-fähigen elektronischen Stellwerke der ersten Generation für den Einsatz mit FSS ertüchtigt.

Neben der Lebensverlängerung verfolgt die SBB mit dem als «Upgradability» bezeichneten Ansatz der Homogenisierung der Projektierungen von optischer Signalisierung und FSS. Am Beispiel der Strecke Dagmersellen – Emmenbrücke konnten Kapazitätsentwicklungen simuliert werden, die die Migration weiter vereinfachen könnten.

2 Herausforderungen bei der Migration

Um den flächendeckenden Rollout von FSS zu ermöglichen, sind eine Vielzahl von ineinandergreifenden Massnahmen notwendig. Eine besondere Rolle spielt dabei die Integration der europäischen Standardisierung auf technologischer und betrieblicher Ebene, die im Rahmen des Europe’s Rail Joint Undertaking (ERJU) zwischen den europäischen Bahnen und der Industrie vorangetrieben wird. Hinzu kommen die technischen Spezifikationen für Interoperabilität, die das Zusammenwirken der verschiedenen ERTMS-Bausteine grenzüberschreitend sicherstellen sollen. (Verordnung (EU) 2021/2085 des Rates, Ziffer 7)

Das Zusammenspiel von ETCS, Fahrzeugausstattung und neuem Funkstandard „Future Railway Mobile Communication System“ stellt mit seinen gegenseitigen Abhängigkeiten im Rahmen des System Versionings eine besondere Herausforderung dar. Ein abgestimmter Rollout einer auch auf zukünftige Anforderungen anpassbaren und wirtschaftlich sinnvollen ERTMS-Baseline, ist vor Mitte der 2030er Jahre aktuell nicht vorstellbar (Verband des öffentlichen Verkehrs (Hrsg.), 2022a). Bereits heute stellt die Finanzierung der Fahrzeugausstattung die Migrationsplanung vor eine grosse Herausforderung. Ein Upgrade der Fahrzeuge auf die aktuell verfügbare Baseline 3 würde spätere Aufwände für die für den Einsatz mit FRMCS benötigte Baseline 4 nach sich ziehen. Die bisher mangelhaft spezifizierte Auf- und Abwärtskompatibiltät zwischen den Baselines führt auf Grund der langen Umrüstzeiträume (in der

Schweiz wird für den Rollout der gesamten Flotte auf Grund knapper Werkstattkapazitäten und wenigen Fahrzeugreserven von einem Zeitraum von sieben Jahren ausgegangen) zu einer vollständigen Migration der Strecke erst nach Upgrade des letzten verkehrenden Fahrzeugs. Aus der für die Baseline 4 benötigten TSI 2026 ergibt sich durch die langen Umrüstzeiten der Fahrzeuge damit ein Migrationszeitraum ab 2034. Vorherige Installationen von FRMCS oder ETCS Level R können in dieser Zeit nicht oder nur eingeschränkt eingesetzt werden – bei vollen Kosten.

Aus dieser Abhängigkeit heraus entsteht die Herausforderung, den Zeitraum bis zur Verfügbarkeit von technischen Lösungen zur Realisierung des Zielbilds zu überbrücken und für die bis dahin zu ersetzenden Anlagen einen zukunftsfähigen Umgang zu definieren.

2.1 Obsoleszenz und fehlende Zielbild-Konformität

Bis zum Zeitpunkt, an dem die für einen nachhaltigen Umstieg auf ERTMS notwendigen Produkte zur Verfügung stehen, erreichen grosse Teile der bestehenden Sicherungsanlagen ihr Lebenszyklusende. Dies betrifft zum einen die Relaisstellwerke (rStw), die gemessen am Wiederbeschaffungswert heute über 50% der Stellwerkslandschaft der SBB ausmachen. Hinzu kommen weitere rund 20% der Stellwerke, die zu den elektronischen Stellwerken der ersten Generation (eStw 1. Gen) gezählt werden. (SBB, 2023)

Bei den rStw besteht aktuell weder ein Know-How- noch ein Obsoleszenz-Problem. Allerdings sind diese Anlagen auf Grund ihrer Technologie nicht FSS-fähig und müssen zur Erreichung des FSS-Zielbilds ersetzt werden.

rStw werden in der Schweiz mit einem Lebenszyklus von 60 Jahren kalkuliert (Verband des öffentlichen Verkehrs (Hrsg.), 2018, S. 27). Der gemessen an ihrer Anzahl grösste Relaisstellwerkstyp – das Domino 67 – befindet sich Stand 2023 mit durchschnittlich 37 Jahren knapp über der Hälfte dieses Lebenszyklus. (SBB, 2023)

Bis zu einem denkbaren ERTMS-Rollout Mitte der 2030er Jahre werden diese Stellwerke ihr technisches Lebenszyklusende erreicht haben. Dies bietet zum einen die Chance einer direkten Ersatzinvestition auf das Zielbild ERTMS. Sollte dieses jedoch bis zum Lebenszyklusende der rStw nicht zur Verfügung stehen, drohen massive Investitionen in konventionelle Technologie, die für eine spätere Migration auf ERTMS im Sinne des ERJU abgeschrieben werden müssten.

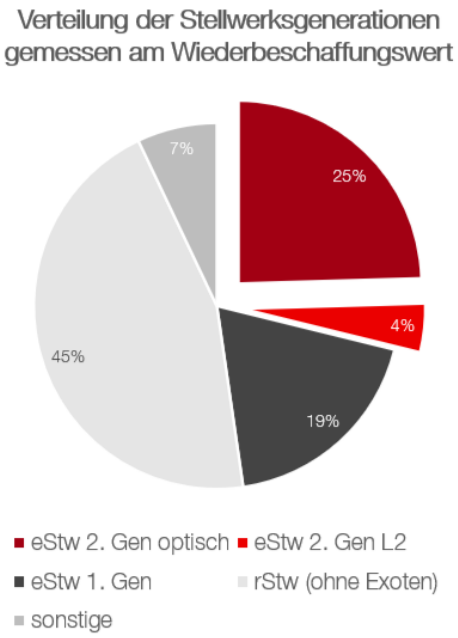


Abb. 1: Verteilung der Stellwerkstypen nach Wiederbeschaffungswert
Quelle: SBB 2023

Bei der ersten Generation der elektronischen Stellwerke, die zwischen 1990 und 2010 gebaut wurden (SBB 2023), bestehen bereits heute akute Obsoleszenz-Probleme in den Bereichen Ersatzteilversorgung und Know-How. Getrieben durch diese Obsoleszenz werden Lösungen für den Stellwerksersatz vor 2032 und damit vor der Verfügbarkeit einer stabilen ERTMS-Baseline benötigt (Verband des öffentlichen Verkehrs (Hrsg.), 2022b). Damit droht das Risiko einer nicht-zielbild-konformen Ersatzinvestition in Höhe von rund 20% des Wiederbeschaffungswerts aller Sicherungsanlagen, die später auf den ERJU-Standard migriert werden müssen.

2.2 Hohe Migrationskosten bestehender Anlagen auf FSS

Der Ersatz von Sicherungsanlagen mit nicht-zielbild-konformen Anlagen verschärft ein anderes, bereits existierendes Problem: Als einziges Land neben Luxemburg hat die Schweiz ETCS flächendeckend im Hauptnetz ausgerollt (Jung, S. 7). Dabei kommt überwiegend der Betriebsmodus limited Supervision zum Einsatz, bei dem der Zug nur teilweise von ETCS überwacht wird und der Triebfahrzeugführer weiterhin verpflichtet ist, auch die streckenseitige Signalisierung zu beachten. (ETCS Spezifikation der ERA, Subset 023, Version 3.3.0). Bis auf wenige Ausnahmen wie dem Gotthard-Basistunnel, in dem ETCS Level 2 eingesetzt wird, erfolgt die Signalisierung in der Schweiz optisch.

Sowohl für den Einsatz mit FSS als auch für den Einsatz mit optischer Signalisierung werden elektronische Stellwerke der 2. Generation (eStw 2. Gen) eingesetzt. Dabei machen die optischen Stellwerke mit 25% der Gesamtanlagen den deutlich grösseren Anteil gegenüber den bereits mit FSS eingesetzten Anlagen (4%) aus (SBB 2023).

Obwohl es sich um identische Stellwerksprodukte handelt, haben Erfahrungen in der Vergangenheit gezeigt, dass bei der Migration bestehender elektronischer Stellwerke von optischer Signalisierung auf Führerstandsignalisierung etwa Kosten in Höhe eines kompletten Stellwerksersatzes anfallen. Den getätigten Investitionen in die Anlagen mit eStw 2. Gen und optischer Signalisierung droht damit eine vorzeitige Abschreibung. Gleiches gilt für alle Anlagen, die vor Verfügbarkeit von Zielbild-konformen

Produkten aus Obsoleszenz-Gründen oder auf Grund von Leistungssteigerungen durch eStw ersetzt werden müssen.

Die Gründe für diesen Kosteneffekt liegen in den unterschiedlichen Philosophien von optischer Signalisierung und FSS. Während die optische Signalisierung eine Geschwindigkeitssignalisierung ist, handelt es sich bei FSS um eine Abstandssignalisierung. Konkret bedeutet das, dass bei optischer Signalisierung die Fahrtgeschwindigkeit für Streckenabschnitte unabhängig von der Verkehrsdichte fest vorgegeben ist. Bei FSS ist die Fahrtgeschwindigkeit abhängiger von der eigenen Zugkomposition und von vorausfahrenden Zügen. Spätestens mit dem angestrebten „Moving Block“, der eine Zugfolge im Bremsabstand möglich machen soll, wird das Paradigma einer festen Geschwindigkeit-Strecke-Zuordnung aufgegeben und bisher ungenutzte Kapazitäten realisiert (Zilch (Hrsg.), 2002, S. 63).

Der Einsatz von Moving Block scheitert aktuell vor allem an zwei Aspekten: Insbesondere bei Güterzügen, die im Schweizer Mischverkehr auf den gleichen Strecken verkehren wie Personenzüge, ist die Feststellung der Zugintegrität nicht flächendeckend möglich. Um diesem Problem Herr zu werden, kommen Achszähler zum Einsatz. Diese zählen sowohl bei Einfahrt als auch bei Ausfahrt in einen Block die ein- und ausgehenden Zugachsen und können so die Zugintegrität sicherstellen. Durch die zudem fehlende sichere Ortung der Fahrzeuge müssen zudem Gleisfreimeldeeinrichtungen eingesetzt werden, die feststellen, in welchem Blockabschnitt sich ein Zug befindet (Pachl, 2011, S. 82).

Bei der Migration von optischer Signalisierung auf FSS kann im Elementmix der Schweiz auf ca. 20% der Elemente verzichtet werden (Haupt- und Vorsignale). Die im Zuge der FSS-Migration vorgenommene Blockverdichtung erfordert jedoch zusätzliche Achszähler und Gleisfreimeldeeinrichtungen, die den Wegfall der Aussensignale nahezu kompensieren. Bedingt durch unterschiedlich normierte Bremsdistanzen müssen die übrigen Aussenelemente teilweise neu platziert werden. Um diesen Umbau ohne Betriebsunterbrüche zu ermöglichen, werden dazu neue Aussenelemente am Zielstandort platziert, bevor auf diese umgeschaltet wird und die bestehenden Elemente zurückgebaut werden. Auf eine Weiterverwendung bestehender Aussenelemente und die damit verbundenen Kosteneinsparungen wird somit zugunsten der Fahrbarkeit verzichtet. Der Begriff «Fahrbarkeit» meint in diesem Zusammenhang die Beförderung von Fahrgästen entsprechend dem Regelfahrplan und ohne baustellenbedingter Betriebsunterbrechungen. Noch nicht berücksichtigt in dieser Aufstellung sind Mehrkosten, die durch das Radio Block Center als zusätzliche Komponente entstehen.

3 Lösungsansätze

Um den in Kapitel 2 ausgeführten Herausforderungen zu begegnen und gleichzeitig die ERTMS-Strategie des Bundesamts für Verkehr zu berücksichtigen, wurde 2021 das Portfolio „Evolution ERTMS Sicherungsanlagen“ ins Leben gerufen. Dieses Portfolio bündelt alle Projekte zur Weiterentwicklung der Stellwerkslandschaft der SBB. Nachfolgend werden Teile der Ergebnisse der Projekte „Lebensverlängernde Massnahmen Simis-C“ sowie „Industrialisierter Rollout“ vorgestellt.

3.1 Lebensverlängernde Massnahmen bei elektronischen Stellwerken der 1. Generation

Die elektronischen Stellwerke der ersten Generation stellen rund 20% des Anlagenwerts der SBB dar. Den grössten Anteil an dieser Anlagengeneration macht das Produkt «Simis-C» von Siemens aus. Aktuell betreibt die SBB noch 44 zentrale und 21 abgesetzte Stellwerke dieser Generation in nahezu allen grösseren Betriebspunkten wie Basel, Zürich, Bellinzona oder Biel. (SBB 2023)

Industrieseitig ist die Simis-C Stellwerkstechnologie getrieben durch die Pensionierung von Know-How-Trägern auf 2032 abgekündigt. Da die meisten der Anlagen ihren mit 30 Jahren kalkulierten Lebenszyklus zu diesem Zeitpunkt bereits überschritten haben werden, ist der Know-How-Erhalt über diesen Zeitpunkt hinaus nicht wirtschaftlich abbildbar. (Zinngrebe, 2022, S. 22 f.)

Um die Investitionen in eine nicht-zielbildkonforme Zwischengeneration zu vermeiden, arbeitet die SBB seit 2021 an einer Möglichkeit zur investitionsschonenden Lebensverlängerung der Simis-C Stellwerke. Gemeinsam mit Siemens wurde dazu ein Ansatz entwickelt, der es ermöglicht, die obsolete Sicherungslogik der Simis-C Stellwerke durch die moderne Logik des Nachfolgeprodukts «Simis W» zu ersetzen. Der so entstehende Hybride wird als «Simis-C Plus» bezeichnet. Der Vorteil dieser Lösung liegt darin, dass die bestehenden Aussenelemente unverändert zu übernehmen.

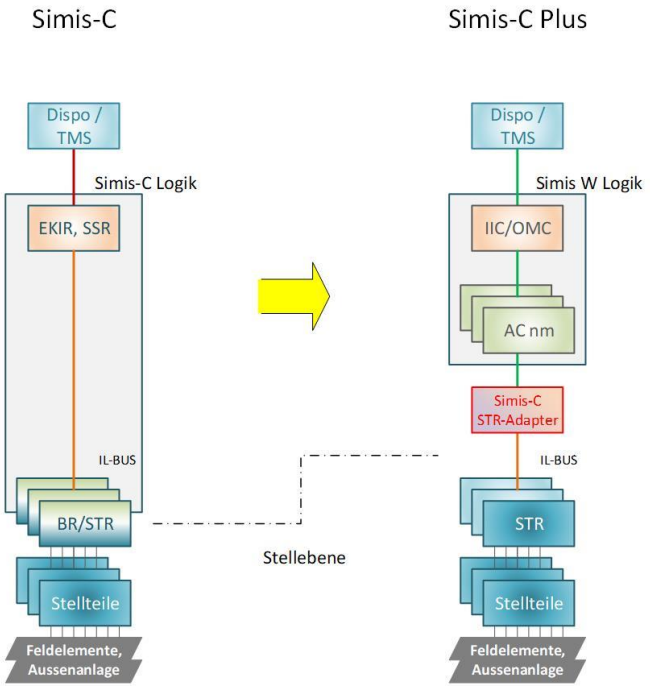


Abb. 2: Architektur für die Migration von Simis-C zu Simis-C Plus
Quelle: Zinngrebe, 2022 S. 24

Da die Aussenelemente gemeinsam mit ihrer Verkabelung für rund 2/3 der Gesamtkosten des Stellwerksersatz verantwortlich sind, entsteht bei der angestrebten Verlängerung bis 2045 ein attraktiver Business Case. Die Verlängerung bis 2045 bedeutet für die meisten Simis-C eine Verlängerung des Lebenszyklus um ca. 15 Jahre, was der Hälfte des ursprünglichen Lebenszyklus' entspricht. Durch die Übernahme der Aussenelemente sinken die Kosten verglichen mit einem Stellwerkersatz auf durchschnittlich etwa 30% des Wiederbeschaffungswerts. (Zinngrebe, 2022, S. 27)

Während für die technische Lösung des Obsoleszenz-Problems eine Adapterlösung zwischen der Sicherungslogik des Simis W und den Stellelementen des Simis-C entwickelt wird, stellt der Umgang mit den funktionalen Unterschieden zwischen Simis-C und Simis W betreiberseitig die grössere Herausforderung dar.

Eine Analyse des Funktionsumfangs der Simis-C identifizierte über 70 Sonderfunktionalitäten, die im Simis W nicht generisch zur Verfügung stehen. Um eine Entwicklung von nicht mehr benötigten Funktionalitäten zu verhindern, wurde in einem cross-funktionalen Team über die gesamte Infrastruktur eine Bewertung vorgenommen. Im Ergebnis konnte auf die Entwicklung von generischen Funktionen weitestgehend verzichtet und viele Funktionen über die Projektierung abgebildet werden.

Bis 2028 plant die SBB mit der Lebensverlängerung einer Pilotanlage und 13 weiterer Anlagen, bei denen spätere Anpassungen an den Anlagen bereits geplant sind und die alternativ hätten ersetzt werden.

3.2 Angleichung der Projektierungsregeln für ETCS L1LS und ETCS L2 zur Ermöglichung von Upgradability

3.2.1 Studiendesign

Durch die unterschiedlichen Betriebsphilosophien von ETCS L1LS und L2 und der Notwendigkeit zum Umbau während des laufenden Betriebs, stellt die Migration von Stellwerksanlagen von L1LS auf L2 eine besondere Herausforderung dar. Um die drohenden Kapazitätsverluste allfälligen Minderkosten durch die Angleichung der Projektierungsregeln zu beurteilen, wurde am Beispiel der Strecke Olten-Luzern die Kapazitätsentwicklung genauer untersucht. Ziel dieser als «Upgradability» bezeichneten Angleichung ist die Reduktion von Umbauaufwänden durch die Nutzung gleicher Signalstandorte sowohl beim Betrieb mit L1LS als auch mit L2.

Tab. 2 Übersicht der betrachteten Szenarien für Upgradability

	L1LS	L2
Kompromisssignalisierung auf Basis L2	Eingeschränkt	Vollwertig
Kompromisssignalisierung auf Basis L1LS	Vollwertig	Eingeschränkt

Da der Streckenabschnitt Dagmersellen – Emmenbrücke ab dem Jahr 2029 mit ETCS L2 betrieben werden soll, lag zu Studienbeginn neben der bestehenden L1LS-Projektierung bereits die optimierte L2-Projektierung vor. Anhand dieser Projektierungen wurden zwei «Kompromisssignalisierungen» genauer untersucht:

- Die «Kompromisssignalisierung auf Basis L2» ging davon aus, dass es möglichst wenig Anpassungen an der späteren L2-Ausstattung geben sollte, jedoch Abstriche und Regelwerksverletzungen in der L1LS-Projektierung zulässig seien. Annahme hinter diesem Szenario war, dass die Nutzung mit L2 die Zielsignalisierung darstelle und es in dieser keine kapazitären Einschränkungen geben solle. Dafür wird die L2-Projektierung umgesetzt und um die für optische Signalisierung erforderlichen Signale erweitert. Die Gleisfreimeldeabschnitte, Achszählerstandorte etc. befinden sich bereits an ihrem Zielstandort im L2-Signalisierungskonzept.
- Die «Kompromisssignalisierung auf Basis L1LS» ging von dem Szenario aus, dass eine Migration auf L2 erst zu einem späteren Zeitpunkt erfolgen könne, und deshalb die kapazitären Einschränkungen im L1LS nicht zu tragen seien. Deshalb wurden die Abstriche in der L2-Projektierung vorgenommen. In diesem Fall wird die L1LS-Projektierung als Grundlage verwendet und auf Optimierungen der Blockabschnitte durch zusätzliche bzw. anders platzierte Gleisfreimelder und Achszähler verzichtet.

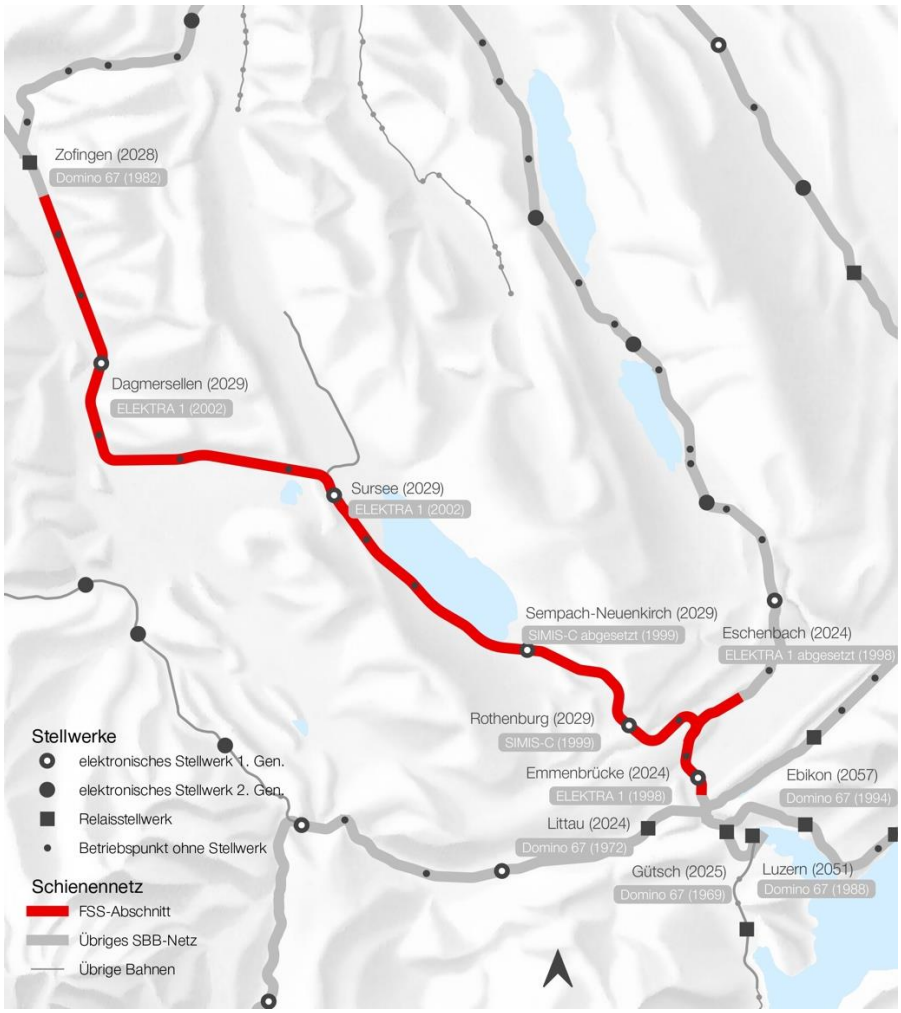


Abb. 3: Übersichtskarte Streckenabschnitt Dagmersellen – Emmenbrücke
Quelle: SBB 2022, S. 7

Ausgewertet wurden die verschiedenen Szenarien hinsichtlich zwei Kapazitätskennzahlen: der verkette Belegungsgrad nach UIC-406 (Mischverkehr, Peak Hour) wurde genutzt, um in einer ersten Indikation eine Aussage bezüglich der Auslastung ohne Berücksichtigung der Fahrbarkeit zu erlangen. Da die Studienautoren davon ausgingen, dass die Kapazitätsdefizite derart gross seien, dass eine vertiefte Analyse nicht notwendig werden würde, diente diese als Früherkennungsmerkmal.

Der verkettete Belegungsgrad nutzt fahrplanabhängig Sperrzeitentreppen und ermittelt unter Berücksichtigung von Geschwindigkeiten, Zugmix und -zahlen sowie Fahrplanrobustheit eine theoretische Kapazität. Dazu wurde pro Fahrtrichtung und Signalisierungskonzept über einen Verkettungszeitraum von ca. 10h der Belegungsgrad komprimiert.

Für die Ermittlung der Fahrbarkeit wurde zudem eine Betriebssimulation durchgeführt. In dieser wurde der Zielfahrplan über die jeweilige Kompromisssignalisierung simuliert, wobei der Simulationszeitraum 8 Stunden betrug.

3.2.2 Studienergebnisse – Verketteter Belegungsgrad

Die Analyse des Belegungsgrads kam für die verschiedenen Signalisierungskonzepte für die Fahrtrichtung Luzern zu folgendem Ergebnis:

Tab. 3 Verkettete Belegungsgrade Fahrtrichtung Luzern

L1LS	Kompromisssignalisierung auf Basis L1LS	Kompromisssignalisierung auf Basis L2	L2
0.685	0.70	0.67	0.645

Im direkten Vergleich der Signalisierungskonzepte L1LS und L2 weist L2 einen um 0.04 Punkte geringeren Belegungsgrad aus. Diese wenn auch geringfügige Unterschreitung des Belegungsgrads von L1LS kann als Indikator für höhere Kapazität herangezogen werden.

Erstaunlich an den Ergebnissen ist jedoch, dass die Kompromisssignalisierung auf Basis L1LS um 0.015 Punkte besser abschneidet, als das auf L1LS optimierte Signalisierungskonzept (stark zu Lasten der Fahrbarkeit, siehe unten). Mit einem verketteten Belegungsgrad von 0.7 schneidet die Kompromiss-signalisierung auf Basis L2 am schlechtesten ab.

Für die möglichen Migrationsszenarien der Upgradeability konnten die befürchteten Kapazitätsverluste durch eine Angleichung der Projektierungsregeln in Fahrtrichtung Luzern nicht bestätigen lassen. Mit der Kompromisssignalisierung L1LS würde die Belegung vor der L2-Migration sogar etwas besser ausfallen als beim unveränderten L1LS, bevor die Zielbelegung mit L2 am niedrigsten wäre.

Eine Ausstattung mit der Kompromisssignalisierung L2 wäre zwar schlechter als L1LS, doch auch hier wären die Kapazitätsverluste vor der Migration auf das Zielkonzept marginal.

Auch die Studienergebnisse in Fahrtrichtung Olten zeichneten bestätigte die vorher beschriebenen Ergebnisse:

Tab. 4 Verkettete Belegungsgrade Fahrtrichtung Olten

L1LS	Kompromisssignalisierung auf Basis L1LS	Kompromisssignalisierung auf Basis L2	L2
0.71	0.73	0.72	0.72

In dieser Fahrtrichtung schnitt L2 gemessen am verketteten Belegungsgrad sogar leicht schlechter ab als L1LS. Insgesamt bewegen sich jedoch auch die Belegungsgrade in einem marginalen Bereich.

3.2.3 Studienergebnisse - Betriebssimulation

Um zu validieren, ob die erstellten Signalisierungskonzepte auch den Anforderungen des Fahrplans genügen, wurde eine Betriebssimulation durchgeführt. In dieser wurden Verspätungen beim Einbruch in den Untersuchungszeitraum und zur Abbildung des Fahrgastwechsels eingestreut. Zudem wurden resultierende Folgeverspätungen im Halt und auf der Strecke durch die Interaktion mit anderen Zügen simuliert. Dem entgegen wirkten Verspätungsabbau im Halt und auf der Strecke. Basierend auf diesen Effekten liess sich eine resultierende bzw. nicht kompensierte Verspätung ermitteln, die im jeweiligen Signalisierungskonzept verglichen wurde.

Die Ergebnisse dieser relativierten die zuvor dargestellten Berechnungen der verketteten Belegungs-grade. Die aus der Einbruchsverspätung und der Primärverspätung Halt resultierenden Sekundärverspä-tungen sowie deren möglicher Abbau fiel im Szenario der Kompromisssignalisierung auf Basis L2 signifikant schlechter aus, als in allen anderen Signalisierungskonzepten. Bemerkenswert bei diesen

Ergebnissen ist, dass sowohl beim Aufbau der Verspätung als auch beim Abbau signifikant höhere Effekte erzielt werden.

Ein L1LS-Signalisierungskonzept mit möglichst vielen Gleisfreimelde-Standorten des L2-Zielkonzepts erweist sich auf Basis dieser Ergebnisse als das schlechteste aller Szenarien. Gleichzeitig scheint sich mit der Weiterverwendung der Gleisfreimelde-Standorte aus dem L1LS-Konzept auch unter L2 eine zu vertiefende Alternative gegenüber dem Komplettersatz aufzutun, auch wenn diese leicht schlechter zu sein scheint, als ein ideales L2.

Tab. 5 Ergebnisse der Betriebssimulation in Minuten

	L1LS	**Kompromisssignalisierung auf Basis L1LS**	**Kompromisssignalisierung auf Basis L2**	**L2**
Einbruchsverspätung	2.3	2.3	2.3	2.3
Primärverspätung Halt	6.0	6.0	6.0	6.0
Sekundärverspätung Halt	14.3	14.3	22.4	13.8
Sekundärverspätung Strecke	8.6	9.2	20.4	9.2
Verspätungsabbau Halt	-4.8	-4.8	-10.3	-4.6
Verspätungsabbau Strecke	-16.0	-16.4	-24.1	-16.3
Summe	**10.4**	**10.7**	**16.7**	**10.4**

4 Zusammenfassung, Fazit, nächste Schritte

Das Fehlen von europäisch standardisierten Produkten und eine Lösung für den Umgang mit den bisher inkompatiblen Systemversionen der verschiedenen ERTMS-Komponenten, stellt für die SBB ein immenses Investitionsrisiko dar. Auf Grund des anstehenden Lebenszyklusendes von grossen Teilen der Stellwerkslandschaft sind akut Massnahmen erforderlich, um die bestehende Bahnproduktion aufrecht zu erhalten.

Die auf ETCS L2 aufwärtskompatible Lebensverlängerung der Simis-C Stellwerke ist ein wesentlicher Baustein dafür. Der Ersatz der obsoleten Sicherungslogik ermöglicht den späteren Anschluss an ein Radio Block Center und damit die Einbindung in L2-Perimeter. Mit der Bereinigung von Sonderfunktionen werden die Anlagen zudem vereinfacht, was eine spätere Ablösung durch standardisierte Produkte beschleunigen soll. Die aktuell grösste Herausforderung in diesem Thema besteht in der Beschleunigung der Rollout-Prozesse, sodass neben den parallel stattfindenden Ausbau- und Erneuerungsprojekten ein Grossteil der 43 zentralen und 21 abgesetzten Simis-C bis zum Lebenszyklusende 2032 migriert werden können. Aktuell ist die Inbetriebnahme der Pilotanlage für Mitte 2026

geplant, der Rollout soll ab 2027 starten. Bis 2032 müssten bei Migration aller Anlagen somit über 12 Anlagen pro Jahr migriert werden.[1]

Die Erfahrungen aus der Analyse der Sonderfunktionalitäten der Simis-C erhöhte die Sensibilität für den Aufwand, den die Analyse und die Erarbeitung von Handlungsoptionen für den Umgang mit ihnen erfordert. Aus diesem Grund analysiert die SBB aktuell die Sonderfunktionalitäten ihrer grössten Relaisstellwerkgenerationen. Damit soll ein Überblick über notwendige Migrationsmassnahmen und Entwicklungsbedarfe erstellt werden, die vorgängig zum angestrebten industrialisierten Rollout einer neuen Generation von Sicherungsanlagen umgesetzt werden müssen.

Durch die Auseinandersetzung mit dem Themengebiet Upgradability konnte ein Weg zum investitionsschonenden Umgang mit Anlagen aufgezeigt werden, die auf Grund von zwingendem Substanzerhalt oder Ausbauvorhaben bereits vor der Verfügbarkeit neuartiger Technologien abgelöst werden müssen. Das erhoffte Szenario, bei dem ETCS L1LS um zusätzliche Gleisfreimelde-Standorte des L2-Zielsystems erweitert wird, erwies sich dabei jedoch als betrieblich deutlich instabiler als mögliche Alternativen. Die Übernahme von Gleisfreimelde-Standorten von L1LS beim Einsatz von L2 hingegen könnte ein Ansatz sein, der einen kostengünstigen Umstieg auf Führerstandsignalisierung ermöglicht und gleichzeitig betrieblich wie kapazitär stabil zu sein scheint.

In einem nächsten Schritt sollen die Ergebnisse der Studie auf die Strecke Lausanne – Genf übertragen werden. Diese Strecke ist betrieblich anspruchsvoller – und zudem ein Paradebeispiel für die Grundherausforderung des System Versionings bei der Migration auf FSS: Auf Grund der Grenznähe zu Frankreich stehen beim Betrieb mit dem bestehenden GSM-R keine ausreichenden Funkkapazitäten für eine Migration auf L2 mit bestehender Technologie zur Verfügung. Eine mögliche Lösung wäre FRMCS – was auf Grund der noch ausstehenden Technischen Spezifikation für Interoperabilität jedoch nicht zur Verfügung steht und eine Fahrzeugausstattung mit der ebenfalls noch nicht verfügbaren Baseline 4 erfordern würde. Die Relaisstellwerke entlang der Strecke nähern sich jedoch ihrem Lebenszyklusende, wodurch ein konkreter Anwendungsfall für Upgradability entsteht.

[1] Zum Vergleich: in den letzten Jahren nahm die SBB pro Jahr durchschnittlich sechs neue Stellwerke in Betrieb

5 Literaturverzeichnis

[1] Jung, Bernd u.a.: Back on track – Solving the digitalization challenge for Europe´s rail sector [online]. Available at: https://www.strategyand.pwc.com/de/en/industries/transport/railway-digitization.html (Accessed 10th April 2023)

[2] Pachl, Jörn (2011): Systemtechnik des Schienenverkehrs. Bahnbetrieb planen, steuern und sichern. 6., überarbeitete Auflage. Wiesbaden: Vieweg + Teubner Verlag

[3] SBB AG (2022, 11. August), Bedarfsorientierte Umsetzung ETCS L2: Variantenentscheid LSS/FSS Dagmersellen - Emmenbrücke. (unternehmensinterne Quelle)

[4] SBB AG (2023, 05. April), „Stellwerkliste aktuell“ [Power-BI Datenbank]. Available at: https://app.powerbi.com/groups/4794f94c-7be1-45b7-b3bf-46cff869142c/reports/4ec59635-a3ee-46b2-a5c2-ec95e2a4b7b1/ReportSectionb05676c02c9ec205e204?clientSideAuth=0 (Accessed 10th April 2023) (unternehmensinterne Quelle der SBB AG)

[5] Verband des öffentlichen Verkehrs (Hrsg.) (2018): RTE 29900: Netzzustandsbericht – Minimalanforderungen. Bern: Verband des öffentlichen Verkehrs

[6] Verband des öffentlichen Verkehrs (Hrsg.) 2022a, „9. Sitzung VöV-Forum Umsetzung ERTMS-Strategie“, 12 [online]. Available at: https://www.voev.ch/de/Service/content_?download=18696 (Accessed: 9th April 2023)

[7] Verband des öffentlichen Verkehrs (Hrsg.) 2022b, „5. Sitzung VöV-Forum Umsetzung ERTMS-Strategie“, 19 [online]. Available at: https://www.voev.ch/de/Service/content_?download=18330 (Accessed: 9th April 2023)

[8] Zilch, Konrad u.a. (Hrsg.) (2002): Handbuch für Bauingenieure. Technik, Organisation und Wirtschaftlichkeit – Fachwissen in einer Hand. Heidelberg: Springer-Verlag

[9] Zinngrebe, Melvin u.a.: Lebensverlängernde Maßnahmen für die erste Generation elektronischer Stellwerke. In: Signal + Draht 114 (2022), S. 22 – 27

Seamless ETCS operation under degraded conditions

Simon Hofer[1] und Dr. Martin Müller[1]

[1] team Technology Management GmbH

1 Einleitung

Die Verfügbarkeit im Bereich der Eisenbahn Sicherungssysteme kommt eine wesentliche Bedeutung für den zuverlässigen Schienenverkehr zu. Die bisher eingesetzten – konventionellen -- Stellwerke und Zugsicherungssysteme sind in der Regel – bezogen auf das Streckennetz – für einen kleinräumigen geographischen Bereich zuständig und dezentral neben der Strecke aufgebaut. Bei Ausfall des Installationsstandortes der Systemtechnik z.B. durch Stromausfall, Brand etc. sind die Auswirkungen somit auf einen lokalen Bereich beschränkt. Ersatzmaßnahmen wie z.B. Weichensperren, Schienenersatzverkehr etc. können einen solchen Ausfall zwar nicht komplett verhindern, die Auswirkungen aber erheblich reduzieren.

Bei der Streckenausrüstung mit ETCS L2 werden die Systemkomponenten (RBC) für einen bestimmten geographischen Bereich, oder das gesamte Gebiet eines EIU – im Gegensatz zu konventionellen Stellwerken – zentral an einem Standort errichtet. Sofern bei der Ausrüstung der Strecke zusätzlich zu ETCS L2 eine Außenlichtsignalisierung besteht, kann diese – analog zu den oben genannten Ausführungen – als eine mögliche Rückfallebene im nationalen Zugsicherungssystem genutzt werden.

Auf Strecken ohne Außenlichtsignalisierung existiert diese Rückfallebene nicht mehr. Somit kann bei Ausfall des zentralen RBC-Standorts nur ein eingeschränkter Zugbetrieb in SR oder Level 0 gemäß den nationalen und betrieblichen Vorschriften und daher meist mit erheblich geringeren Geschwindigkeiten und reduzierter Überwachung durch das Zugsicherungssystem stattfinden. Die Auswirkungen solcher Ausfälle sind besonders aufgrund der geographischen Ausdehnung erheblich für den laufenden Bahnbetrieb insbesondere aus dem Blickwinkel kritischer Infrastruktur.

In der Situation ohne Außenlichtsignalisierung und somit ohne Rückfallebene wird es somit bei einem Ausfall des RBCs (Disaster Fall) zu erheblichen Betriebseinschränkungen kommen, die zum Teil sogar zu Betriebseinstellungen führen werden. Aufgrund der großen geographischen Ausdehnung wird für den Personenverkehr ein Schienenersatzverkehr nur begrenzt möglich sein. Der Verfügbarkeit des RBCs kommt daher eine wesentliche Bedeutung für die Betriebsqualität zu. Der Ausfall einzelnen Komponenten ist durch MTBF (mean time bevor failure) Werte etc. sehr gut beschreibbar und durch Redundanz sowie prädiktive Instandhaltung auch ohne Ausfall des Gesamtsystems behebbar. Für den – sehr selten auftretenden - - Ausfall eines gesamten Standortes (z.B. durch Beschädigung der Infrastruktur, Stromausfälle, Brände, etc.) bedarf es jedoch zusätzlicher Maßnahmen in Form einer georedundanten Ausführung des RBCs.

2 Realisierungsvarianten Georedundanz

2.1 Kategorien Georedundanzstandorte

Gemäß Bauer, Adams, Eustace; 2012; S32 [1] lassen sich die Georedundanzstandorte, die zur Wiederherstellung der Systemverfügbarkeit geplant sind, grob in fünf Realisierungs-Kategorien einteilen:

2.1.1 Ad-hoc-Standort

Um den Ausfall eines Standortes zu kompensieren, kann nach dem Eintreten des Ausfalls kurzfristig ein neuer Standort gesucht und die Ersatzkomponenten dort installiert werden. Denkbar ist hier z.B. der Einsatz von Containern bis zur Wiederherstellung des ursprünglichen Standortes.

2.1.2 Cold Recovery Site

ISO/IEC 24762:2008 [2] definiert eine Cold Recovery Site als einen Standort, der mit angemessenem Platz und zugehöriger Infrastruktur – Stromversorgung, Telekommunikationsverbindungen, Zugangskontrollen usw. – eingerichtet ist. Diese Grundinfrastruktur dient zur Unterstützung von IKT Systemen der jeweiligen Organisation. Diese werden nur im Schadensfall an dem Ausfallstandort im Zuge des Disaster Recovery installiert.

2.1.3 Warm Recovery Site

ISO/IEC 24762:2008 [2] definiert eine Warm Recovery Site als einen Standort, der zumindest teilweise mit der notwendigen Hard- und Software und dem notwendigen Personal ausgestattet ist. Bei Aktivierung des Disaster Recovery wird zusätzliche Hard- und Software am Recovery Standort installiert und gegebenenfalls weiters Personal am Standort tätig.

2.1.4 Vereinbarung über gegenseitige Unterstützung

Unternehmen können gegenseitige Vereinbarungen zur Unterstützung im Schadensfall schließen. So würde z.B. die Grundinfrastruktur wie Rechnerräume etc. für das Disaster Recovery für das jeweilige andere Unternehmen bereitgestellt. Die Systemkomponenten wie Hard- und Software müssten dann – analog der Warm Revovery Site – bei Ausfall des ursprünglichen Standortes erst installiert werden. Möglich ist auch, dass Behörden einen solchen Standort für KRITIS relevante Unternehmen bereithalten und im Schadensfall zur Verfügung stellen.

2.1.5 Hot Site

ISO/IEC 24762:2008 [2] definiert Hot Recovery Site als einen Standort, der vollständig mit der erforderlichen Ausrüstung, der Hard- und Software sowie unterstützendem Personal ausgestattet und rund um die Uhr voll funktionsfähig ist, um jederzeit im Falle eines Disaster Recovery den Betrieb zu übernehmen.

2.1.6 Bewertung der Kategorien in Bezug auf die Wiederherstellungszeit

Die bereits in der Einleitung gezeigt sind bei einer RBC Installation ohne Außenlichtsignalisierung keine längeren Betriebsunterbrechungen zulässig. Daher kommt der Betrachtung des Kriteriums Umschalt- bzw. Wiederherstellungszeit der Systeme eine besondere Bedeutung zu. Die Variante 2.1.1 Ad-hoc-Standort erscheint als mögliche Lösung für ein Desaster Recovery ungeeignet, da für die Standortsuche und die Herstellung der notwendigen Infrastruktur (Stromversorgung, Telekommunikation, Netzwerk etc.) im optimalen Fall mit mehreren Tagen, in der Regel aber mehreren Wochen, Betriebsunterbrechung zu rechnen ist. Hinzu kommt außerdem die Installation der Hard- und Softwarekomponenten sowie die dazugehörige Prüfung der Funktionsfähigkeit. Voraussetzung dafür ist die vollständige und unmittelbare Verfügbarkeit aller Komponenten an einem anderen Standort als dem Betriebsstandort. Damit die Komponenten möglichst schnell in Betrieb genommen werden können ist ein Grundaufbau der Systemschränke (interne Verkabelung etc.) Voraussetzung. Denkbar wäre hier, dass z.B. die Labor-

anlage des Herstellers im Recovery Fall als Ersatzanlage dient. Da in der Regel an einem Betriebsstandort mehrere RBC-Systeme installiert sind, ist dann für jedes dieser RBCs eine Laboranlage notwendig bzw. müssten die fertig verkabelten Systemschränke und alle Komponenten bereitgehalten werden. Für den Transport, die Installation und Prüfung der Komponenten am Recovery Standort ist auch hier von mehreren Tagen Wiederherstellungszeit auszugehen. Der Zeitaufwand für die Installation trifft auch bei den Punkten 2.1.2 Cold Recovery Site und 2.1.4 Vereinbarung über gegenseitige Unterstützung zu. Bei der in Punkt 2.1.3 Warm Recovery Site dargestellten Variante sind die Komponenten zum Teil bereits installiert. Für die zusätzliche Installation weiterer notwendiger Hard- und Softwarekomponenten, deren Inbetriebnahme und Tests wird jedoch auch erhebliche Zeit im Bereich von zumindest einem Tag oder mehr benötigt. Diese Zeiten treten beim Punkt 2.1.5 Hot Site nicht auf, da bei dieser Variante bereits alle notwendigen Hard- und Softwarekomponenten installiert, in Betrieb genommen und getestet sind.

Allein aus dem Kriterium Umschalt- bzw. Wiederherstellungszeit ist daher als Disaster Recovery Strategie für die Realisierung eines georedundanten Standortes einzig die Variante „Hot Site“ sinnvoll. Zur Reduzierung von (Betriebs-)Kosten ist es denkbar die Systeme am Standort – analog Punkt 2.1.5 Hot Site - vollständig betriebsbereit zu installieren und zu testen und die Systeme dann herunterzufahren und nur im Recovery Fall zu starten. Ebenso ist es möglich, dass das Betriebspersonal nur eingeschränkt (z.B. Werktags) am Standort ist und die 24/7 Betreuung ausschließlich im Recovery Fall aktiviert wird. In der weiteren Betrachtung wird daher immer von der Variante „Hot Site“ bzw. einer analogen Abwandlung – gemäß den oben dargestellten Ausführungen – ausgegangen.

2.2 Service-Levels bei „Hot Site“ Standorten

Die installierten RBC und Umsystem-Komponenten können an einem georedundanten Standort in verschiedenen betrieben werden. Bauer, Adams, Eustace; 2012; S33 [1] bezeichnen diese als Servicelevels. In Anlehnung an diese Ausführungen ergeben sich für die mögliche Realisierung einer RBC-Georedundanz fünf verschiedene Kategorien von Service-Levels:

2.2.1 Active

Die RBC- bzw. Umsystem-Komponenten sind permanent – also 24/7 – in Betrieb und stehen für die Nutzung zur Verfügung. Die Schnittstellen zu den anderen Systemen sind dauerhaft in Betrieb und können somit jederzeit von den Fahrzeugen, Bedienern etc. genutzt werden. Die Informationen vom KMC, Stellwerk etc. werden laufend an beiden Standorten abgefragt und verarbeitet. Das primäre sowie das georedundante RBC sind in Hard- und Software nahezu identisch aufgebaut und daher „doppelt“ vorhanden. Bestimmte Daten wie Session-Informationen, Einschränkungen, etc. sind zwischen den RBCs der jeweiligen Standorte zu synchronisieren. Eine mögliche Realisierungsvariante für das RBC ist hier eine Loadsharing-Lösung. Dabei werden die Fahrzeuge zwischen den RBC-Standorten aufgeteilt. Die Informationen stehen jederzeit an beiden RBC-Standorten vollständig zur Verfügung. Bei Ausfall eines Standortes kann der zweite Standort sämtliche Fahrzeuge uneingeschränkt („seamless“) übernehmen. Denkbar ist aber auch sämtliche Komponenten des RBC vollständig redundant aufzubauen und sämtliche Handlungen, Berechnungen etc. an beiden Standorten zeitgleich durchzuführen. Zudem müssen sämtliche relevante Daten an beiden RBCs vollständig zur Verfügung stehen. Die Kommunikation mit sämtlichen Fahrzeugen übernimmt jedoch ausschließlich ein Standort. Im Fall des Disaster Recovery verbinden sich – nach der Umsystem Umschaltung – dann sämtliche Fahrzeuge am

anderen Standort und kann dort die Session übernehmen. Der Aufwand zur Realisierung ist bei dieser Variante am höchsten.

2.2.2 Active mit eingeschränkter Kapazität

Wird das georedundante RBC wie unter Punkt 2.2.1 „Active“ beschrieben als Loadsharing-System realisiert, ist es möglich dieses nur mit eingeschränkter Kapazität zu realisieren. Das bedeutet die volle Kapazität der maximal anmeldbaren Fahrzeuge etc. steht nur im Normalbetrieb mit zwei Standorten zur Verfügung. Bei Ausfall eines Standortes reduziert sich die Kapazität z.B. auf die Hälfte. Zugunsten geringerer Realisierungskosten kann also im Störungsfall mit nur einem Standort lediglich ein eingeschränkter Betrieb mit weniger angemeldeten Tfz. durchgeführt werden.

2.2.3 Hot Standby

Bei dieser Variante ist sowohl der primäre Standort als auch der georedundante Standort in Betrieb. Die Fahrzeuge verbinden sich im Normalbetrieb ausschließlich mit dem primären RBC. Dieses führt auch die entsprechenden Berechnungen, MA-Vergaben etc. durch. Eine parallele Berechnung am georedundanten Standort erfolgt nicht. Die Umsystem-Schnittstellen – wie Verbindung zum KMC, zum Stellwerk etc. – sind permanent aktiv, um sämtliche relevanten Daten bereits verfügbar und aktuell zu haben. Somit wird die Dauer des Umschaltvorgangs sehr kurzgehalten, da einen zusätzlicher – zeitaufwändiger – Abgleich nicht notwendig ist. Einschränkungen, ohne ähnliche Daten, sind zwischen den RBCs der jeweiligen Standorte zu synchronisieren. Die Session des Fahrzeugs wird je nach Realisierungsvariante ggf. nicht synchronisiert. Fehlt die Synchronisierung muss sich das Fahrzeug nach einer Umschaltung wieder neu am RBC anmelden. Ebenso wird die Verbindung zum Nachbar RBC (NRBC) erst mit der Umschaltung auf den georedundanten Standort aktiviert. Der Aufwand für diese Lösung ist relativ hoch, da sämtliche Systeme dauerhaft in Betrieb sind und daher auch entsprechend gewartet werden müssen. Auch bei dieser Variante ist sowohl das primäre als auch das georedundante RBC in Hard- und Software nahezu identisch aufgebaut und daher „doppelt“ vorhanden. Gegenüber den vorangegangenen „Active“ Lösungen – wie z.B. Loadbalancer – ist demgegenüber die Anforderung an die Hard- und Software geringer und daher auch etwas kostengünstiger.

2.2.4 Warm Standby

Bei der „Warm Standby“ Variante ist das RBC inklusive Software betriebsbereit installiert. Im Normalbetrieb ist ausschließlich der primäre Standort für die Verbindungen zu Fahrzeugen, KMC und Stellwerk sowie die Vergabe der MAs verantwortlich. Laufende Daten werden – gengenüber den vorigen Varianten – am georedundanten RBC nicht aktuell gehalten. Daher besteht auch keine Notwendigkeit die Umsysteme – wie z.B. das Stellwerk – mit dem georedundanten RBC zu verbinden. Ebenso entfällt bei dieser Variante die Synchronisierung zwischen den Standorten. Dadurch entfällt ein wesentlicher Implementierungs- und Integrationsaufwand und die Anzahl der notwendigen Umsystem-Schnittstellen die zeitgleich in Betrieb sein müssen verringert sich erheblich. Die Hard- und Software ist jedoch auch hier – ähnlich wie bei den Varianten zuvor -- sowohl für das primäre als auch das georedundante RBC nahezu identisch aufgebaut und daher „doppelt“ vorhanden. Bei der Inbetriebnahme des georedundanten Standorts im Zuge des Disaster Recovery müssen zuerst alle Daten entsprechend abgefragt werden. Eingabedaten wie z.B. Einschränkungen müssen erst manuell eingespielt oder ggf. sogar händisch eingegeben werden. Dadurch steigt die Umschalt- bzw. Wiederherstellungszeit an. Da das RBC keine Daten aktuell halten muss besteht bei dieser Variante die Möglichkeit das RBC am georedundanten Standort auszuschalten. Den Einsparungen bei den Betriebskosten steht der Nachteil

gegenüber, dass die Komponenten in dieser Situation nicht dauerhaft überwacht werden und ein Komponentenausfall somit erst bei Systemstart offenbart wird.

2.2.5 Cold Standby

Ebenso wie bei der Variante 2.2.4 Warm Standby werden auch bei der Cold Standby Variante die Daten am georedundanten Standort nicht aktuell gehalten. Im Unterschied zur Warm Standby Variante ist zudem auch die aktuelle Software am RBC nicht aufgespielt. Vor der Inbetriebnahme des georedundanten RBCs im Zuge des Disaster Recovery ist daher zuerst die Installation der aktuellen Software und die Integration der Systeme notwendig. Die Umschalt- und Wiederherstellungszeiten erhöhen sich dementsprechend. Voraussetzung ist daher ein vorab definierter und standardisierter Prozess, der eine rasche für die rasche Integration ermöglicht. Da in dieser Variante keine aktuelle Software am RBC installiert ist kann das RBC ausgeschaltet werden. Auch hier fehlt dann – wie schon bei der Variante zuvor – die permanente Überwachung der Komponenten. Ein möglicher Einsatzzweck ist die N+1 Georedundanz, sofern die RBCs eines EIU bereits an mehreren Standorten installiert sind und ein zusätzlicher Georedundanz-Standort realisiert wird. In diesem Fall wird an dem Georedundanz-Standort nur einmal die größte Anzahl aller RBCs pro Standort benötigt. Im Gegensatz zu allen vorhergehenden Varianten können in so einer Fallkonstellation also erheblich Hardware- und Umsystem-Kosten eingespart werden. Diesem Kostenvorteil stehen aber die erheblich längeren Umschalt- bzw. Wiederherstellung-Zeiten gegenüber.

3 Umschaltprozess

Ein wesentlicher Faktor beim Disaster Recovery ist die Umschaltung zwischen den verschiedenen Standorten. Abhängig vom eigesetzten System und dem geplanten Service Level kann dieser unterschiedlich gestaltet sein. Wichtige Voraussetzung ist ein standardisierter Ablauf der Umschaltung. Gerade in Stresssituationen wie dem Ausfall eines Standortes – der in der Regel auch andere Systeme betrifft – bedingen Tätigkeiten, die außerhalb der Regelabläufe notwendig sind, ein hohes Potential für Fehler. Die Behebung dieser Fehler führt dann in weiterer Folge zu einer Vergrößerung der Umschalt- bzw. Wiederherstellungszeiten. Idealerweise passiert die Umschaltung daher – sofern dies nicht vollautomatisch abläuft – anhand von vorab erstellten und mehrfach geübten Checklisten. Unabhängig von der Art der Umschaltung wird es notwendig sein den Umschaltprozess in regelmäßigen Abständen zu testen, um notwendige Änderungen im Prozess, den Konfigurationen (z.B. geänderte IP-Adressen), etc. festzustellen und somit im Fehlerfall ein lauffähiges System am georedundanten Standort zu haben.

3.1 Manuelle Umschaltung der einzelnen Systeme

Bei Ausfall des primären Standortes werden bei dieser Realisierungsvariante die Systeme manuell durch das Betriebspersonal umgeschaltet. Die Umschaltung kann dabei manuell per Fernwartung erfolgen. Alternativ besteht auch die Möglichkeit die Verbindungen hardwaremäßig umzustecken oder durch Umschaltung der Konfiguration am jeweiligen System (z.B. mittels Konfigurationsstecker) umzustellen. Der initiale Implementierungsaufwand für diese Lösung ist sehr gering. Demgegenüber sind die Umschaltzeiten sehr hoch, da zuerst das Betriebspersonal tätig werden muss. Die Prozesse und Anleitungen für das Bedienpersonal müssen sehr präzise beschrieben werden, da die Hinzuziehung eines Systemspezialisten in der Regel außerhalb der Bürozeiten einen zusätzlichen Zeitaufwand darstellt und daher die Umschaltung allein durch das 24/7 Betriebspersonal gewährleistet sein muss. Bei der

manuellen Umschaltung besteht die Gefahr, dass Konfigurationsänderungen (z.B. IP-Adressen, Firewall Regeln etc.) auf der georedundanten Schnittstelle nicht umgesetzt wurde. Durch regelmäßige Prüfung der Umschaltung kann sichergestellt werden, dass auch im Störungsfall die Umschaltung reibungslos abläuft.

3.2 Manuelle Umschaltung über ein zentrales System

Analog der Variante 3.1 (Manuelle Umschaltung der einzelnen Systeme) werden auch hier die Schnittstellen manuell umgeschaltet. Anstelle der Umschaltung an den einzelnen Komponenten wird hier die Umschaltung über ein zentrales System vorgenommen. Die zusätzliche Realisierung bedeutet einen höheren Initialaufwand für der Realisierung des zentralen Systems. Demgegenüber verringert sich die Umschaltzeit sowie der Aufwand für die Erstellung der Umschaltprozesse erheblich. Durch die zentrale Umschaltung wird das Risiko von Fehlbedienungen durch das Betriebspersonal minimiert, da die Umschaltungen bei den einzelnen Systemen automatisch vom Zentralsystem ausgelöst werden. Auch bei dieser Variante ist durch regelmäßige Prüfung der Umschaltung sicherzustellen, dass sämtliche Konfigurationen aktuell sind und somit im Störungsfall ein reibungsloser Ablauf gewährleistet ist. Auch das zentrale Umschaltsystem muss entsprechend georedundant ausgeführt sein, um SPOF zu vermeiden und daher im Disaster Recovery Fall die Umschaltung sicherzustellen.

3.3 Automatische Umschaltung

Die Umschaltung der Schnittstellen zum georedundanten Standort erfolgt bei dieser Variante automatisch, also ohne zusätzliche Handlung des Betriebspersonals. Diese Umschaltung kann z.B. durch die Aktivierung des georedundanten RBC-Systems erfolgen. Dabei löst das RBC ähnlich der Implementierung in Variante 3.2 (Manuelle Umschaltung über ein zentrales System) die Umschaltung bei den einzelnen Komponenten aus. Daher sind vom RBC-System entsprechende Schnittstellen zur Umschaltung der Umsystem-Komponenten zu implementieren. Wesentlicher Vorteil ist dabei, dass keine weitere manuelle Handlung des Betriebspersonals notwendig ist und daher die Umschaltzeiten im Störungsfall wesentlich verkürzt werden. Auch die Anforderung an die Umschaltprozesse wird minimiert. Ebenso wie bei den vorhergehenden Varianten kann durch eine regelmäßige Test-Umschaltung sichergestellt werden, dass die Umsystem-Schnittstellen am georedundanten System im Störungsfall fehlerfrei umgeschaltet werden können.

3.4 Loadbalancing

Bei dieser Variante erfolgt keine explizite Umschaltung im Falle des Disaster Recovery. Stattdessen werden die ankommenden Daten automatisch per Loadbalancer oder funktionsähnlicher Systeme (z.B. im GSM-R Core) per Session, Datenpakt usw. automatisch zum jeweiligen aktiven System transferiert. Kann der Sessionaufbau z.B. aufgrund eines Standortausfalls nicht zum geplanten System erfolgen, so wird der Loadbalancer automatisch die Session zum anderen Standort transferieren. Im Falle eines als Loadsharing realisierten RBCs kann der Loadbalancer zudem die Sessions z.B. mittels Round Robin Verfahrens auf die RBCs verteilen. Dadurch sind im Störungsfall nicht alle Sessions oder Teilnehmer zeitgleich betroffen. Als zentrale Komponenten und daher auch Single Point of Failure (SPoF) ist auch das gesamte Loadbalancing System georedundant auszuführen, um auch bei diesen Komponenten die notwendige Verfügbarkeit zu gewährleisten. Der Aufwand in der Umsetzung und im laufenden Betrieb ist hoch. So sind z.B. Prozesse für Upgrades und Konfigurationsänderungen notwendig, damit es zu keiner Betriebsunterbrechung kommt.

4 Georedundanz Varianten in Kombination mit Umsystem-Umschaltungen

Das RBC-System ist mit verschiedenen Umsystemen z.B. Stellwerk, KMC, GSM-R, Bedienoberfläche, etc. verbunden. Eine Möglichkeit der Realisierung für die Umsysteme ist, die Schnittstellen im jeweiligen Umsystem für die Georedundanz gesondert auszuführen. Dies ist jedoch nicht bei jedem Umsystem implementierbar bzw. dürfen in bestimmten Varianten erst mit der Aktivierung des georedundanten RBCs auch die Schnittstellen aktiviert werden. Daher sind Umschalt-System notwendig, um die technischen Anforderungen zu erfüllen. Aufgrund der Eigenschaften und dem Verhalten der verschiedenen Umsetzungs-Varianten der Georedundanz sind nicht alle Umsystem-Umschaltungsmöglichkeiten für die jeweilige Umsetzungs-Varianten geeignet. In der Tabelle Tab. 6 (Mögliche Kombination Umsystem-Umschaltungen) ist für die jeweilige Einsetzbarkeit der verschiedenen Varianten dargestellt. Für die Analyse wurden typische Realisierungen angenommen, mögliche Sonderbauformen oder eine Kombination von verschiedenen Umschaltungen bleiben zugunsten der Klarheit der Darstellung außer Betracht. Die Bewertung erfolgt dabei nach den Kriterien uneingeschränkt Einsetzbar (einsetzbar), bedingt Einsetzbar (bedingt) sowie nicht geeignet (ungeeignet).

4.1 Auflistung der möglichen Umsystem-Umschaltungen

Tab. 6 Mögliche Kombination Umsystem-Umschaltungen und Georedundanz-Varianten

	Active	Active mit eingeschränkter Kapazität	Hot Standby	Warm Standby	Cold Standby
Manuelle Umschaltung der einzelnen Systeme	ungeeignet (4.2.1)	ungeeignet (4.2.1)	bedingt (4.2.2)	einsetzbar (4.2.3)	einsetzbar (4.2.3)
Manuelle Umschaltung über ein zentrales System	ungeeignet (4.2.1)	ungeeignet (4.2.1)	bedingt (4.2.2)	einsetzbar (4.2.3)	einsetzbar (4.2.3)
Automatische Umschaltung	ungeeignet (4.2.1)	ungeeignet (4.2.1)	bedingt (4.2.4)	einsetzbar (4.2.5)	einsetzbar (4.2.5)
Loadbalancing	einsetzbar (4.2.6)	einsetzbar (4.2.6)	einsetzbar (4.2.6)	einsetzbar (4.2.6)	einsetzbar (4.2.6)

4.2 Bewertung der möglichen Umsystem-Umschaltungen

4.2.1 Manuelle sowie automatische Umschaltung mit Varianten Active

Wie unter 2.2.1 (Active) beschrieben seht das georedundante RBC permanent – also 24/7 – zur Verfügung. Die Informationen von Umsystemen wie z.B. des Stellwerks werden permanent verarbeitet. Daher ist es notwendig, dass sämtliche Schnittstellen permanent in Betrieb sind und die entsprechenden Daten an das RBC übermittelt werden. Nachdem bei einer manuellen bzw. automatischen Umschaltung die Schnittstellen erst mit der Umschaltung eine Verbindung zum georedundanten RBC hergestellt wird, kann diese Anforderung an eine permanente Übermittlung an das RBC in diesen Umschaltungsvarianten nicht erfüllt werden. Dabei ist es unerheblich, ob die Umschaltung über ein zentrales System oder an

den einzelnen Komponenten erfolgt. Der Fall trifft ebenso für die Realisierungsvariante „Active“ als auch für die Variante „Active mit eingeschränkter Kapazität“ zu.

4.2.2 Manuelle Umschaltung mit Variante Hot Standby

Die manuelle Umschaltung (unabhängig ob pro Komponente oder über ein zentrales System) ist grundsätzlich für die georedundante Variante Hot Standby nicht geeignet, da wie im Kapitel 2.2.3 beschrieben auch bei dieser Variante die Schnittstellen permanent abgefragt werden und daher die Funktionalität durch eine erst manuelle Umschaltung im Störungsfall die funktionalen Anforderungen nicht erfüllt. Nachdem sich das Fahrzeug nach Umschaltung auf den georedundanten Standort die Session mit dem RBC erneut aufbauen muss, könnten die Schnittstellen RBC-Fahrzeug manuell umgestellt werden. Ebenso triff dies auf die NRBC-Schnittstellen zu. Auch diese können manuelle umgeschaltet werden. Das Einsparungspotential ist jedoch begrenzt, da alle anderen Schnittstellen nicht mit einer manuellen Umschaltung ausgeführt werden können.

4.2.3 Manuelle Umschaltung mit Varianten Warm oder Cold Standby

Sowohl in der Variante Warm Standby (2.2.4) als auch in der Variante Cold Standby (2.2.5) hat das RBC am georedundanten Standort im Standard-Betriebsfall keine Verbindung zu den Umsystemen. Daher ist die Umsetzung mittels manueller Umschaltung möglich.

4.2.4 Automatische Umschaltung mit Varianten Hot Standby

Analog der Beschreibung der manuellen Umschaltung mit der Variante Hot Standby (4.2.2) ist auch die automatische Umschaltung nur bedingt geeignet, da der überwiegende Teil der Schnittstellen laufend aktiv sein muss. Somit bleibt auch hier nur die automatische Umschaltung der Fahrzeug- und der NRBC-Schnittstelle. Die automatische Initiierung der Umschaltung durch das RBC bietet jedoch gegenüber der manuellen Umschaltung erhebliche Vorteile.

4.2.5 Automatische Umschaltung mit Varianten Warm oder Cold Standby

Ebenso wie unter 4.2.3 beschrieben ist auch die automatische Umschaltung in den Varianten Warm- oder Cold Standby möglich. Der Vorteil liegt gemäß Kapitel 3.3 in den wesentlich vereinfachten Prozessen bei der Umschaltung im Disaster Recovery Fall. Demgegenüber ist aufgrund der langen RBC Inbetriebnahme Zeiten kein wesentlicher Vorteil in der Umschaltzeit zu erwarten.

4.2.6 Loadbalancing mit allen Varianten

Die Umschalt-Variante Loadbalancing ist mit jeder der Georedundanz-Varianten einsetzbar, da der Loadbalancer vollautomatisch (seamless) die Daten zwischen den Standorten umschaltet. Somit ist es unerheblich, ob das RBC-System sich vor dem Fehlerfall bereits in Betrieb befindet, oder erst im Disaster Recovery Fall in Betrieb genommen wird. Zu beachten ist dabei, dass die geringen Umschaltzeiten bei den Varianten Warm oder Cold Standby kaum Auswirkungen haben, die RBC Inbetriebnahme Zeiten wesentlich höher sind.

5 Vergleich

Für den Vergleich werden alle Kombinationen gemäß Tabelle Tab. 6 verglichen, die den Status „bedingt einsetzbar“ oder „einsetzbar“ haben. Als Bewertungskriterien werden quantifizierbare Eigenschaften herangezogen und diese anhand der Skala „gering“ „moderat“ „erheblich“ „hoch“ bewertet. Um einen

standardisierten Vergleich zu ermöglichen, wird von einer Neuimplementierung der RBC-Georedundanz inkl. Umschaltung ausgegangen. EIU können demgegenüber bestimmte Umschaltmechanismen wie z.B. Loadbalancer bereits für andere System in Einsatz haben und somit den Aufwand der Umsetzung minimieren. Hierfür ist eine Betrachtung im Einzelfall notwendig und wird daher im untenstehenden Verglich nicht weiter betrachtet.

5.1 Bewertungskriterien

- Umschalt- bzw. Wiederherstellungszeit
- Systemkomplexität
- Prozesskomplexität
- Personalaufwand Betriebspersonal
- Kosten (Errichtungskosten)
- Fehleranfälligkeit im Umschaltfall
- Automatisierbarkeit
- Synergie – Testsystem
- Übungsaufwand
- Notwendigkeit der Neuanmeldung am RBC (Seamless)

5.2 Vergleichsübersicht

Tab. 7 Bewertung der Realisierungsvarianten

Variante	Umschaltung	Wieder-herstellungs-zeit	System-komplexität	Prozess-komplexität	Personal-aufwand	Kosten	Fehler-anfälligkeit	Automatisier-barkeit	Synergie	Übungs-aufwand	Seamless
Cold Standby	**Manuelle Umschaltung der einzelnen Systeme**	hoch	gering	hoch	hoch	gering	hoch	gering	hoch	hoch	gering
Cold Standby	**Manuelle Umschaltung über ein zentrales System**	hoch	moderat	hoch	hoch	moderat	erheblich	gering	hoch	erheblich	gering
Cold Standby	**Automatische Umschaltung**	erheblich	moderat	erheblich	erheblich	moderat	moderat	moderat	erheblich	moderat	gering
Cold Standby	**Loadbalancing**	erheblich	erheblich	erheblich	erheblich	erheblich	moderat	moderat	erheblich	moderat	gering
Warm Standby	**Manuelle Umschaltung der einzelnen Systeme**	hoch	gering	hoch	hoch	moderat	hoch	gering	hoch	hoch	gering
Warm Standby	**Manuelle Umschaltung über ein zentrales System**	erheblich	moderat	erheblich	erheblich	moderat	erheblich	gering	hoch	erheblich	gering
Warm Standby	**Automatische Umschaltung**	erheblich	moderat	moderat	erheblich	moderat	moderat	moderat	erheblich	moderat	gering
Warm Standby	**Loadbalancing**	erheblich	erheblich	moderat	erheblich	erheblich	moderat	moderat	erheblich	moderat	gering
Hot Standby	**Manuelle Umschaltung der einzelnen Systeme**	erheblich	moderat	hoch	erheblich	erheblich	hoch	moderat	moderat	hoch	gering
Hot Standby	**Manuelle Umschaltung über ein zentrales System**	erheblich	moderat	erheblich	moderat	erheblich	erheblich	moderat	moderat	erheblich	gering
Hot Standby	**Automatische Umschaltung**	moderat	erheblich	moderat	gering	erheblich	gering	hoch	moderat	moderat	gering
Hot Standby	**Loadbalancing**	moderat	erheblich	moderat	gering	erheblich	gering	hoch	moderat	moderat	gering
Active mit eingeschränkter Kapazität	**Loadbalancing**	gering	hoch	gering	gering	erheblich	gering	hoch	erheblich	gering	hoch
Active	**Loadbalancing**	gering	hoch	gering	gering	hoch	gering	hoch	gering	gering	hoch

5.3 Beschreibung ausgewählter Bewertungen

5.3.1 Active mit Loadbalancing

Diese Kombination hat die geringste Wiederherstellungszeit mit zugleich der geringsten Anforderung an die Umschaltprozesse. Dadurch wird auch die Fehleranfälligkeit des Umschaltprozesses im Störungsfall minimiert. Demgegenüber ist das System sehr komplex aufgebaut und benötigt daher bei (Konfigurations-)Änderungen oder Softwareupdates eine entspreche sorgfältige Vorgehensweise. Ebenso sind die (initialen) Kosten für die Umsetzung sehr hoch.

5.3.2 Warm Standby mit manueller Umschaltung über ein zentrales System

Wird eine kostengünstige Lösung mit geringer Komplexität gewünscht, so bietet sich die Warm Standby Variante mit manueller Umschaltung über ein zentrales System an. Diese Variante hat gegenüber der Cold Standby Variante den wesentlichen Vorteil, dass die Software bereits installiert ist und daher die Wiederherstellungszeit minimiert wird. Die Anforderung an die Prozesse und damit auch die Fehleranfälligkeit ist durch den Wegfall der Software-Installation sowie durch eine zentrale Umschaltung minimiert.

5.3.3 Hot Standby mit Loadbalancing

Eine ausgewogene Kombination bietet die Variante Hot Standby mit Loadbalancing. Durch die vollautomatische Schnittstellenumschaltung durch den Loadbalancer und die Verfügbarkeit aller betriebsrelevanten Informationen kann die Umschaltung innerhalb von sehr kurzer Zeit erfolgen. Zugleich kann aber im Wesentlichen auf bestehende Standard-RBC Komponenten zurückgegriffen werden. Lediglich die Synchronisierung ist zusätzlich zu implementieren. Dies bedeutet einen wesentlich geringen Implementierungs- und Komplexitätsaufwand gegenüber der Active (Loadsharing) Variante.

6 Zusammenfassung

In diesem Dokument wurden die grundsätzliche Notwendigkeit eines georedundanten RBC bei ETCS L2 Strecken ohne Außenlichtsignalisierung gezeigt. Bezüglich der Realisierung der Georedundanz zur Erhöhung der Ausfallssicherheit und Betriebsgüte auf ETCS L2 Strecken ohne Außenlichtsignalisierung wurden verschiedene Aspekte bei der möglichen Umsetzungsvarianten betrachtet. Bei der Realisierung der Standortvariante ist aufgrund der Anforderung an die Umschalt- bzw. Wiederherstellungszeit nur die Variante „Hot Site“ sinnvoll. Beim Service-Level und der Umsystem-Umschaltung gibt es keine „Ideallösung“. Vielmehr hängt die optimale Lösung von den Voraussetzungen und Umgebungsbedingungen des jeweiligen EIU ab. Generell ist dabei zu sagen, dass eine Lösung mit möglichst geringer Umschalt- und Wiederherstellungszeit auch mit wesentlich höheren Kosten verbunden ist und eine grundsätzliche Automatisierung der wesentlichen Abläufe erfordert. Grundsätzlich besteht für ein EIU auch die Möglichkeit mit einem günstigeren System mit höherer Umschalt- bzw. Wiederherstellungszeit die Einführung der Georedundanz am RBC zu beginnen und mit dem (Flächen-)Ausbau und somit steigenden Anforderungen sukzessive das Service Level zu steigen.

7 Literaturverzeichnis

[1] Bauer, Adams, Eustace (2012) Beyond redundancy : how geographic redundancy can improve service availability and reliability of computer-based systems. John Wiley & Sons, Ltd

[2] Standard ISO/IEC 24762:2008 Information technology — Security techniques — Guidelines for information and communications technology disaster recovery services

Viel mehr Kapazität mit ETCS (& Co.) – aber wie? Aktuelle Erkenntnisse aus dem Digitalen Knoten Stuttgart

Peter Reinhart[1]

[1] DB Netz AG

1 Motivation

Die Erwartungen an die mit Digitalen Stellwerken, ETCS und weiteren Techniken erzielbaren Kapazitäts- und Leistungsfähigkeitseffekte bleiben auch in Deutschland gewaltig. Werte von 20 oder 35 Prozent mehr Kapazität oder Zügen, ohne neue Gleise, werden häufig genannt.

Das landläufige Verständnis der mit „digitaler“ Technik möglichen Leistungssteigerungen beschränkt sich dabei immer noch weitgehend auf bloße Blockverdichtungen bzw. „Moving Block“. Während deren Spielräume vielfach an Grenzen stoßen, werden unerwünschte Nebenwirkungen neuer Leit- und Sicherungstechnik in Deutschland vielfach noch hingenommen: So wird ETCS in der Regel Fahrzeugen und Infrastruktur ohne Optimierungen übergestülpt, werden beispielsweise Möglichkeiten für optimierte Blockteilungen noch nicht konsequent genutzt. Im Ergebnis führen beispielsweise bereits verlängerte Systemlaufzeiten, weniger flexible Stellwerke und vergleichsweise flache Bremskurven vielfach gerade nicht zu Leistungssteigerungen. (Vogel 2023)

Im Digitalen Knoten Stuttgart (DKS) treten die Deutsche Bahn und ihre Partner an, die Möglichkeiten und Grenzen „digitaler“ Optimierungen für die Fahrwegkapazität auszuloten, unter anderem im Hinblick auf eine räumlich und zeitlich koordinierte Einführung von ETCS (BMVI 2021). Mit der fortschreitenden Realisierung des Projekts mehren sich die Erkenntnisse, rückt das Gesamtsystem Bahn immer mehr in den Fokus. Der vorliegende Beitrag fasst den bisherigen Erkenntnisstand in geraffter Form zusammen.

[1] peter.reinhart@deutschebahn.com

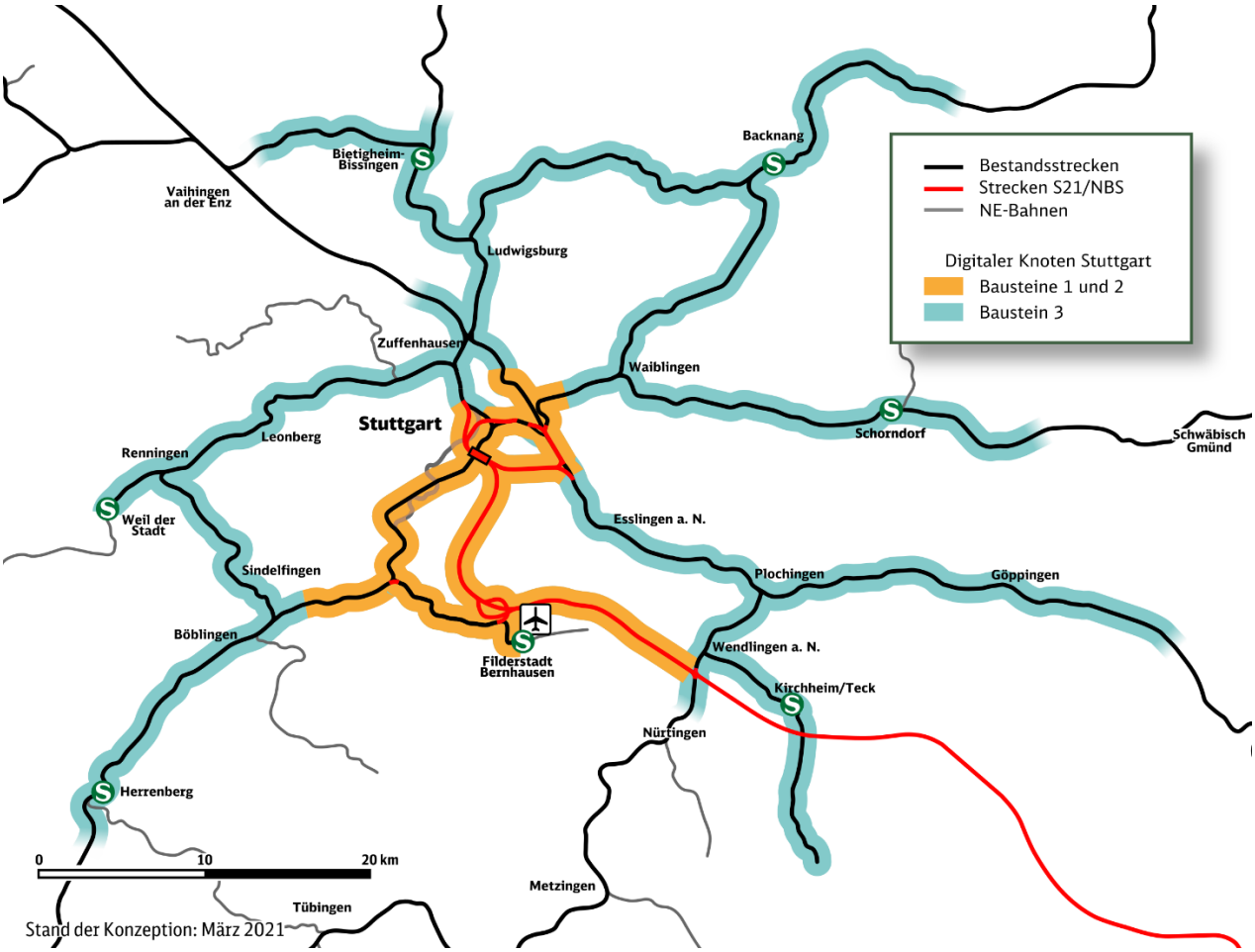


Abb.1 Bis 2025 werden die Bausteine 1 und 2 realisiert, bis in den Horizont 2030 folgt der Baustein 3
Quelle aller Abbildungen: Deutsche Bahn

Zwei große Meilensteine des Projektes liegen im Jahr 2025, wenn zunächst (im September) der S-Bahn-Kernbereich mit ETCS L2oS in Betrieb genommen werden soll, bevor darauf (im Dezember) der neue Hauptbahnhof mit einem Großteil der Infrastruktur von Stuttgart 21 folgt.[2] Eng auf diese Ziele abgestimmt werden rund 333 Triebzüge sowie Nebenfahrzeuge mit einer DSD-/ETCS-Fahrzeugausrüstung nachgerüstet sowie 130 neue Triebzüge „ab Werk“ damit beschafft. (Dietrich 2021, Druckenbrod 2023, Cyril 2023) Im Rahmen eines Modellprojekts fördert der Bund die DSD-Fahrzeugausrüstung, unter anderem geknüpft an 24 technische Bedingungen. (Dietrich 2023)

In den Folgejahren steht die Ausrüstung der weiteren Region an. Im Endausbau wird der DKS dann die gesamte DB-Infrastruktur in den Netzbezirken Stuttgart und Plochingen umfassen (Abb. 1) – insgesamt rund 500 Netzkilometer und somit rund anderthalb Prozent des DB-Netzes. Damit einher werden weitere Techniken in Betrieb genommen und optimiert, beispielsweise das Verkehrsmanagementsystem CTMS, das hochautomatisierte Fahren mit Triebfahrzeugführer (ATO GoA 2), das GSM-R-Nachfolgesystem FRMCS und auch ETCS Level 3. (Beyer 2023)

[2] Tunnel Cannstatt zunächst eingleisig (wegen laufender Arbeiten für die P-Option), östlicher Flughafentunnel folgt nachgelagert (wegen Anbindung des Pfaffensteigtunnels), ohne Gäubahnanbindung

2 Erkenntnisstand der Leistungsfähigkeitsoptimierungen

Die Ausrüstung von Fahrzeugen und Infrastruktur mit „digitaler“ Technik kann und darf freilich kein Selbstzweck sein. Mehr Züge auf dem bestehenden Netz zu fahren ist ein wesentliches Ziel und auch eine wesentliche Triebkraft für den DKS: Bereits im Rahmen der S-Bahn-ETCS-Untersuchung von 2017/2018 (InGe 2019) sowie begleitenden Diskussionspapieren wie (Goers 2019) wurden zahlreiche Möglichkeiten und Potenziale dafür dargelegt. In der weiter Fahrt aufnehmenden Planung und Umsetzung werden immer wieder weitere Potenziale erkannt, beispielsweise die kürzlich in (Chavalier 2023) beschriebenen schneller ladenden Fahrpläne im Zusammenhang mit FRMCS.

2.1 Zugfolge (Stammstrecke)

In der S-Bahn-Stammstrecke liegt nicht nur der Zündfunke für den DKS (Beyer 2019), sondern auch ist auch der Bedarf für Kapazitätssteigerungen am größten. An einem Werktag (Mo-Fr) fahren planmäßig rund 940 Züge pro Tag über die Strecke, wird von 6 bis 20 Uhr ein 2,5-Minuten-Rhythmus angeboten, mit einem Viertelstundentakt auf den sechs in das Umland ausstrahlenden Linien. Sie ist durch einen weitgehend homogenen Betrieb (in Bezug auf Fahrdynamik, Halte, Zuglängen) gekennzeichnet.

In den Spitzenstunden verkehren zumeist Langzüge mit drei Einheiten (rund 205 m) über die Strecke. Die planmäßigen Haltezeiten von 30 Sekunden werden vielfach überschritten, womit sich Verspätungen insbesondere in Spitzenstunden fast unweigerlich auf nachfolgende Züge übertragen. Immer mehr und längere Züge und mehr Fahrgäste führen zu einem Abwärtstrend in der Pünktlichkeit, die nur kurzzeitig durch die Covid-19-Pandemie unterbrochen wurde.

Auf der Stammstrecke hat jede Sekunde Zugfolgezeit, die mit technischen Mitteln herausgearbeitet werden kann, einen immensen Wert. Entsprechend wurden und werden hier alle irgendwie greifbaren Register gezogen, beispielsweise im Hinblick auf die Blockteilung, aber auch auf kleine Potenziale: so wird beispielsweise noch darum gerungen, bereits im heutigen GSM-R die Übertragungsrate von 4,8 auf 9,6 kbit/s (!) zu erhöhen, um ETCS-Fahrterlaubnisse wenige Zehntelsekunden schneller zu übertragen und somit die Mindestzugfolgezeit zu verkürzen.

Im Vergleich von ETCS Level 2 (mit Hochleistungsblock) und ATO GoA 2 gegenüber konventioneller LST (Ks-Signale mit PZB) wies die S-Bahn-ETCS-Untersuchung eine Verkürzung der mittleren Mindestzugfolgezeiten je Richtung von 17 bzw. 23 Prozent aus, auf 120 bzw. 107 Sekunden. Bei unveränderter Leistungsanforderung von 24 Zügen pro Stunde und Richtung wurde ging die Verspätungsänderung über alle Zugfahrten im Auswerteraum um durchschnittlich 26 Sekunden zurück. (InGe 2019)

Der Untersuchung lagen dabei konservative Prämissen zu Grunde, beispielsweise eine Ende-zu-Ende-Systemlaufzeit (ohne umlaufende Weichen) von wenigstens 14 Sekunden, eine nicht gesondert optimierte Blockteilung mit rund 50 m langen Zugfolgeabschnitten entlang des Bahnsteigs sowie nicht optimierte ETCS-Bremskurven (Lambda-Modell).

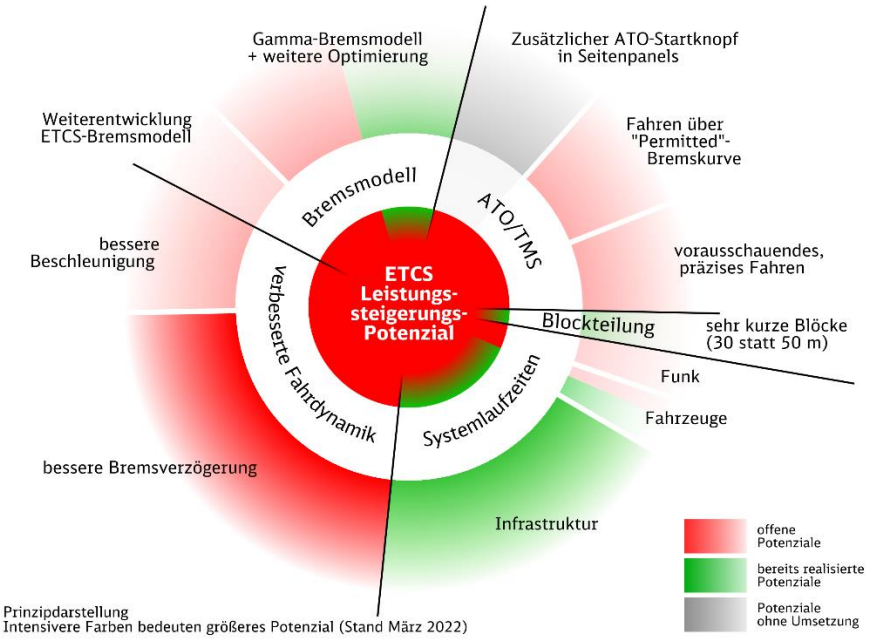


Abb. 2 Kapazitätspotenziale aus der S-Bahn-Untersuchung von 2017/2018 und deren Status

Seither wurden weitere Potenziale gehoben (vgl. Abb. 2):

- Die erwarteten Systemlaufzeiten von Stellwerk und ETCS konnten nach einem Vergabeverfahren, in dem derartige Optimierungen in den Fokus gerückt wurden, erheblich verkürzt werden: In den für die Mindestzugfolgezeiten der S-Bahn-Stammstrecke maßgebenden Abschnitten (Bahnsteig, ohne Weichen) um neun Sekunden. (Behrens 2021) Auch die Verarbeitungszeit auf dem Fahrzeug wurde, gegenüber der Mindestanforderung der ETCS-Spezifikation, um eine Sekunde verkürzt. (Dietrich 2022)
- Durch die mikroskopisch, über alle Zugfolgefälle hinweg optimierte Blockteilung wurde im Modell eine weitere Verkürzung von zwei Prozent erreicht, bei gleichzeitig zehn Prozent weniger Signalen bzw. Blockkennzeichen. (Denißen 2021)
- Durch optimierte ETCS-Bremskurven (insbesondere Gamma- statt Lambda-Modell) werden mit Langzügen wenige Sekunden Verkürzung erwartet. (Förster 2023)
- Zu wenige Sekunden späteren Bremseinsatzpunkten führt auch die erwartete, erheblich genauere Lokalisierung: Die (teils auch für ATO-Präzisionshalte) genau eingemessenen Balisen führen in Verbindung mit einer in der Regel auf wenige Meter genau arbeitenden Odometrie (Dietrich 2022) zu einem erwarteten Ortungsfehler von etwa 5 m, während in der Studie von 2017/2018 noch ein pauschaler Ortungsfehler von 55 m unterstellt wurde.

Mit diesen als gesichert umsetzbar geltenden Optimierungen können die mittleren Mindestzugfolgezeiten um rund 35 Prozent (gegenüber Ks/PZB) verkürzt werden. Rein auf die Betriebsqualität wirkt die in der Studie noch als Potential ausgewiesene und nun in Umsetzung befindliche Erhöhung der zulässigen Geschwindigkeit von 60 auf 80 km/h auf der der heutigen Stammstrecke sowie von 80 auf 100 km/h im südlichen Anschluss daran.[3] Mit mittleren Mindestzugfolgezeiten von rund 100 s wäre

[3] von der Station Universität zur Schwabstraße

beispielsweise ein Betriebsprogramm mit 30 Züge pro Stunde und Richtung (120 s planmäßige Zugfolge) stabil fahrbar.

In der Region Stuttgart wird inzwischen diskutiert, langfristig einen Zehn-Minuten-Takt auf allen sechs Linien einzuführen. Um mittlere planmäßige Zugfolgen von 100 s über die Stammstrecke stabil zu ermöglichen, wären die Mindestzugfolgezeiten gegenüber dem als erreicht geltenden Niveau nochmals um eine Viertelminute abzusenken. Dazu sind drei Stoßrichtungen erkennbar:

- In den verbleibenden „digitalen“ Potenzialen (Abb. 2 in rot, Sektoren ATO/TMS/Systemlaufzeiten) liegt eine Viertelminute Spielraum für die Mindestzugfolgezeit: dazu zählen insbesondere weitere Optimierungen an Bremskurven (innerhalb und unter Weiterentwicklung der ETCS-Spezifikation) (Förster 2023) sowie vorausschauendes Fahren (mit ATO GoA 2 und CTMS) nahe der Schnellbremseinsatzkurve (EBI) (Kümmling 2021). Durch die ohnehin längerfristig geplante Einführung von FRMCS werden im Modell weitere 1,7 s Verkürzung erwartet. (Chavalier 2023)
- Mit Voll- statt Langzügen (zwei statt drei Einheiten) könnten Mindestzugfolgezeiten um rund zehn Sekunden verkürzt werden, da diese früher am Bahnsteig zum Halt kommen und diesen nach Abfahrt schneller räumen.
- Als größtes und langfristigstes Potenzial verbleiben schlicht neue, fahrdynamisch optimierte Fahrzeuge (Abb. 2 links). Mehr als 15 Sekunden Spielraum für die Mindestzugfolgezeit könnte gehoben werden, wenn diese ähnlich beschleunigen und bremsen würden wie Stadtbahnen der Stuttgarter Straßenbahnen (SSB) und zusätzlich (mit ATO und CTMS) vorausschauend kurzzeitig nahe der EBI führen.[4] (Neufahrzeuge sind langfristig ohnehin erforderlich, um beispielsweise die Baureihe 423 abzulösen oder schlicht um noch mehr und längere Züge sowie Linien zu verlängern. Auch ein Betrieb mit 36 Langzügen pro Stunde würde einen massiven Aufwuchs erfordern.)

Zu begleitenden, rein auf die Betriebsqualität wirkenden Potenzialen zählen eine bessere Disposition (mittels CTMS) und die Hebung von Trassierungspotentialen in der Zuführung. Ferner wird bereits heute an vielen Stellen an einer resilienten LST gearbeitet (Behrens 2022). Die Frage scheint insofern nicht mehr, *ob* mit Mitteln der LST langfristig die Grundlage für einen Zehn-Minuten-Takt im Herzen des S-Bahn-Systems ermöglicht werden kann, sondern nur noch *wie*. Eine Kombination aus diesen drei Stoßrichtungen scheint wahrscheinlich.

Im Übrigen ist das viel diskutierte Level 3 mit Moving Block ohne erkennbares Potenzial: Zwar könnte die Freimeldung unter Berücksichtigung des Vertrauensintervalls der Odometrie beispielsweise in 30-m-Abschnitten bis zu etwa 25 m früher erfolgen – dem wirkt jedoch die Offenbarungszeit der Zugtrennung entgegen. Sie liegt zumindest bei der Nachrüstung der Bestandsflotten bei 3,5 Sekunden (Flöter 2022), zuzüglich der weiteren Verarbeitungs- und Funklaufzeit. Bei einer erwarteten Stellwerkslaufzeit (vom physischen Freifahren bis zur virtuellen Fahrtstellung) von einer einzigen Sekunde ist die konventionelle Gleisfreimeldung erheblich schneller und erfolgt früher.

[4] 1,2 statt 0,7 m/s² Betriebsbremsverzögerung, 1,2 statt 0,9 m/s² Anfahrbeschleunigung, Schnellbremsverzögerung von 2,5 m/s²

2.2 Hauptbahnhof

Im Fokus des öffentlichen Interesses liegt auch die Leistungsfähigkeit des im Rahmen von Stuttgart 21 entstehenden neuen Hauptbahnhofs und seiner Zulaufstrecken. Hierfür hat die DB im Rahmen von Diskussionen um die Umsetzbarkeit des Deutschlandtakts 2019 das klare Ziel ausgerufen, bei Bedarf im Hauptbahnhof – unter Praxisbedingungen – auf jedem der acht Bahnsteiggleise alle fünf Minuten einen Zug fahren zu können sowie auf den anschließenden Streckengleisen mittlere Zugfolgezeiten (im Fahrplan) von zwei Minuten zu ermöglichen. (Deutsche Bahn 2019)

Im Gegensatz zur Stammstrecke ist der Betrieb im Hauptbahnhof weniger homogen: Zwar sollen alle Züge im Hauptbahnhof halten, weisen jedoch u. a. unterschiedliche Längen, Fahrdynamiken, Bremskurven, Haltezeiten und Abfertigungsregime auf. Für praktisch jede Zugfahrt laufen einige Weichen um, ferner sollen Bahnsteiggleise teils doppelt belegt werden. Über die homogenen Betrieb der S-Bahn-Stammstrecke dargestellten Ansätze hinaus zeigt sich daher eine Reihe weiterer Optimierungspotenziale.

So liegt ein weitreichendes Hemmnis für die Leistungsfähigkeit in den bisherigen Restriktionen der Blockteilung, insbesondere an elektrischen Schaltabschnittsgrenzen in den Bahnhofsköpfen und entlang des Bahnsteigs. (DB Netz 2023, Hernández 2023) Um diese Restriktionen zu überwinden, bedarf es insbesondere einem Leitsystem wie CTMS, das anhand von Kriterien wie Zuglängen, Stromabnehmerpositionen oder der Bewegung des vorausfahrenden Zuges unerwünschte Teilzugstraßenziele verhindert. (Denißen 2021)

Von wesentlicher Bedeutung ist auch das Abfertigungsverfahren: Anstatt die Abfertigung erst mit der Fahrtstellung des (virtuellen) Ausfahrsignals zu beginnen, könnte diese mithilfe der ohnehin im Betrieb auflaufenden Daten (aus DSTW, ETCS, ATO und Fahrzeugzustandsdaten/TCR (Flöter 2022)) in Verbindung mit einem Leitsystem wie CTMS zukünftig vorausschauend so erfolgen, dass der Zug im Moment der Fahrtstellung bereits zu rollen beginnt. (Ohmeyer 2022)

Bei manchen Zügen führen flache, vom Fahrzeugbetreiber gesetzte Sollbremskurven (Guidance Curve) zu wesentlich längeren Belegungszeiten als mit konventioneller Technik zu erwarten wäre. (Vogel 2022) Der erwartete Effekt des vorausschauenden Fahrens ist bei diesen Zügen besonders groß. Züge mit besonders optimierter Fahrdynamik und optimierten Bremskurven lassen hingegen besonders kurze Zugfolgezeiten erwarten. (Förster 2023)

Zu den weiteren speziell im Hauptbahnhof wirksamen Optimierungen zählt beispielsweise die Einfahrt in teilbesetzte Gleise mit zukünftig voraussichtlich 40 statt bislang 20 km/h. (Beyer 2023)

2.3 Mischverkehrsabschnitte

In dem vielfach zweigleisigen Mischverkehrsstrecken geprägten Umfeld stoßen die auf der Stammstrecke und im Hauptbahnhof angewendeten Optimierungen an enge Grenzen: Zwar kann mit diesen Mitteln auch dort der Abstand zweier Züge verkürzt werden. Da jedoch insbesondere die S-Bahn an vielen Zwischenstationen hält, der übrige Verkehr dort jedoch in der Regel durchfährt, können zwischen zwei S-Bahnen kann nach bisheriger Lesart oft nur ein weiterer (an den S-Bahn-Halten durchfahrender) Fern- oder Regionalverkehrszug geplant werden. Mit zwei Zügen je 15 Minuten und Richtung gelten derartige Streckenabschnitte als ausgelastet, obwohl insbesondere mit Triebzügen des Personenverkehrs durchaus Mindestzugfolgezeiten von einer Minute (Förster 2023) realisiert werden könnten. Trotz Digitalisierung müssten nach bisheriger Lesart derartige Streckenabschnitte aufwendig drei- und viergleisig ausgebaut werden, um neben dem 10-Minuten-Takt der S-Bahn auch den weiteren, massiven Angebotsausbau (insbesondere im Regionalverkehr) zu bewältigen. (VMBaWü 2023)

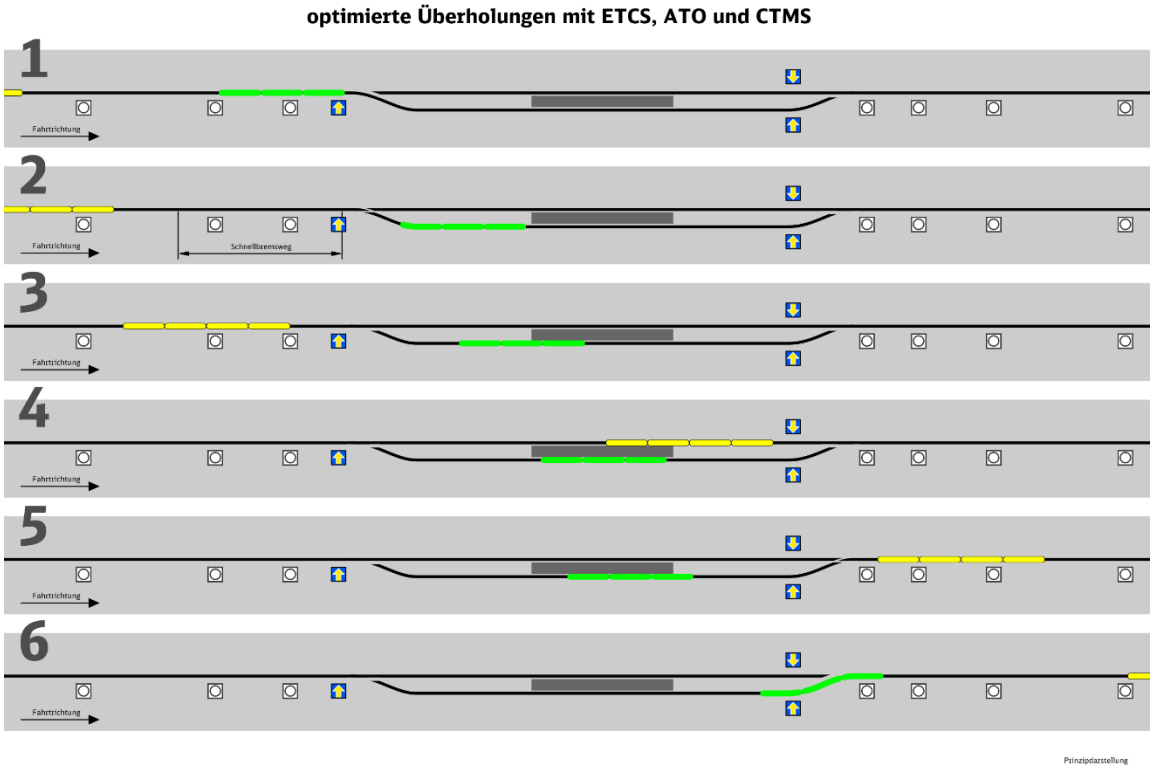


Abb. 3 Schlanke Überholungen am Beispiel einer S-Bahn (grün), die während ihrer 30-sekündigen Verkehrshaltezeit von einem 130 km/h schnellen Regionaltriebzug (gelb) überholt wird

Um diesen Knoten zu durchschlagen, wurde eine Kombination zahlreicher „digitaler“ und überschaubarer konventioneller Optimierungen konzipiert: Im Modell konnte inzwischen nachgewiesen werden, wie beispielsweise eine S-Bahn während einer rund 30-sekündigen Verkehrshaltezeit von einem Regionaltriebzug schlank überholt werden kann (Abb. 3). Dem lagen neben „digitalen“ Optimierungen an Fahrzeugen und Infrastruktur insbesondere auch rund 700 m lange und mit 100 km/h befahrbare Überholgleise zu Grunde. (Ohmeyer 2022)

Bei unveränderten kommerziellen Fahrzeiten werden durch begleitende Geschwindigkeitserhöhungen (insbesondere durch Trassierungsoptimierungen) zusätzliche mehrminütige Reserven geschaffen, die als Spielmassen für CTMS dienen und einen stabilen Betrieb mit schlanken Überholungen erwarten lassen. Dazu fließen die in der „Task Force Trassierung“ des begleitenden Projekts Stuttgart—Ulm gesammelten Erfahrungen (Enzmann 2021, Reinhart 2021, Castano 2022) in die Planung des DKS mit ein. Dazu zählen „kleine“ Trassierungsoptimierungen wie Weichengeschwindigkeiten oder angepasste Überhöhungen ebenso wie schlicht die meter- und 5-km/h-genaue Ausnutzung der Bestandstrassierung mit den Möglichkeiten von ETCS. Diese „kleinen“ Maßnahmen lassen in den vier von der S-Bahn im Mischverkehr befahrenen Zulaufstrecken im Knoten Stuttgart im Mittel etwa 20 km/h größere zulässige Geschwindigkeiten erwarten. Bei gegenüber heute unveränderten kommerziellen Fahrzeiten entstehen so mehrere Minuten Fahrzeitüberschuss, die dem Fern- und Regionalverkehrszügen als zusätzliche Reserve sowie CTMS als Spielmassen dienen können.

Ob dieses Konzept letztlich aufgehen kann, wird in weiteren Untersuchungen (u. a. mit einer optimierten Trassierung und einem TMS) zu klären sein. Falls es gelingt, könnte sich eine Reihe

vergleichsweiser „kleiner“ konventioneller Maßnahmen in Summe durchaus als genauso kapazitativ wirksam erweisen wie beispielsweise Überwerfungsbauwerke, zusätzliche Streckengleise oder gar ganze Neubaustrecken.

3 Resümee und Ausblick

Acht Jahre, nachdem in Stuttgart die ersten Diskussionen um eine leistungsoptimierte ETCS-Ausrüstung der S-Bahn-Stammstrecke aufgekommen sind, zeigen die im und um den Digitalen Knoten Stuttgart gesammelten Erkenntnisse, wie mit der Digitalisierung der Leit- und Sicherungstechnik tatsächlich substanziell mehr Züge auf dem Bestandsnetz gefahren werden können. Dazu reicht es jedoch nicht aus, einfach ETCS Fahrzeugen und Infrastruktur überzustülpen. Vielmehr sind Fachleute mehr denn je gefordert, das System Bahn in seiner ganzen Breite und in all seinen Wechselwirkungen so gut es geht zu verstehen und gesamthaft zu optimieren – nicht nur Fahrzeuge, sondern auch Infrastruktur und Betrieb; nicht nur „digital“, sondern auch mit den übrigen Gewerken; nicht nur mit ETCS, sondern auch mit den weiteren Techniken der Digitalen Schiene Deutschland.

„Digitale“ Technik sollte dabei nicht irgendwie und um der Digitalisierung Willen eingeführt, sondern vielmehr umfassend optimiert werden – auch und gerade in wenig beachteten vermeintlichen Details wie Bremskurvenparametern, Systemlaufzeiten und Geschwindigkeitsprofilen. Neben großen Hebeln wie einem Hochleistungsblock summieren sich dabei auch viele kleine, für sich kaum wahrnehmbare Einzeleffekte auf. In Verbindung mit vergleichsweise einfachen „konventionellen“ Optimierungen werden inzwischen sogar Kapazitätspotenziale offenbar, mit denen selbst auf hochbelasteten Mischverkehrsstrecken Kapazitätssprünge mit vergleichsweise überschaubaren Mitteln und in vergleichsweise kurzen Zeiträumen möglich sein könnten.

Der DKS konnte dabei so nur entstehen, weil bis heute über alle Organisationsgrenzen eher um vernünftige Lösungen gerungen als Probleme gesucht werden; weil – alles in allem – offen und miteinander statt übereinander gesprochen wird; weil nach Kräften informiert, kommuniziert und transparent gemacht wird und keiner Sachdiskussion aus dem Weg gegangen wird.

Somit werden nun Schritt für Schritt Lösungen Wirklichkeit, die viele noch vor wenigen Jahren oder sogar bis heute für kaum möglich hielten: darunter ETCS (& Co.) auf einer S-Bahn-Stammstrecke und mit massiven Leistungssteigerungen, 30-m-Block, ETCS Level 3 mit Zugintegritätsüberwachung, 2 Sekunden Ende-zu-Ende-Systemlaufzeit oder auch ETCS-Bremskurven die sein werden als jene der PZB.

Ein wesentliches Grundprinzip sind dabei Lösungen aus möglichst einem Guss: Anstatt Fahrzeuge und Infrastruktur häppchenweise auszurüsten, wird so weit wie möglich ein über Jahre gereiftes Gesamtkonzept umgesetzt. Anstatt beispielsweise Fahrzeuge zunächst mit ETCS, später mit ATO GoA 2 und anschließend mit FRMCS auszurüsten, wird der tiefgreifende Eingriff für ETCS genutzt, auch die weiteren Techniken weitgehend mit einzubauen, teils jedoch erst später in Betrieb zu setzen. (Dietrich 2023) Während dabei fahrzeug- wie infrastrukturseitig etwa 90 Prozent der Kosten für den Kern des DKS auf eine ohnehin notwendige Minimallösung (Stellwerk und ETCS) entfallen, machen sämtliche Optimierungen (wie optimierte Blockteilung und Bremskurven oder ATO GoA 2) in Summe nur etwa 10 Prozent aus. (Bitzer 2021)

Produktive Sachdiskussionen um eine gleichsam effiziente wie effektive Gestaltung der Digitalisierung der Eisenbahn in Deutschland werden dabei vielfach noch erschwert, weil wesentliche Diskussionsgrundlagen (wie Infrastrukturdaten, Systemlaufzeiten oder Bremskurven) vielfach – obwohl oft digital vorhanden und ohne erkennbare objektive Schutzbedürfnisse – noch nicht zur Verfügung

stehen. Dabei ist es unabdingbar, weitmöglichste Transparenz herzustellen, zu lernen, notwendige offene Diskussionen zu führen und auch Konsequenzen zu ziehen. Dabei könnte es durchaus zu Debatten kommen, welche großen Infrastrukturmaßnahmen durch vielschichtige und vergleichsweise kleine Optimierungen vielleicht nicht mehr, nicht mehr im geplanten Umfang oder Zeithorizont erforderlich sein könnten. Derartige Diskussionen laufen nun im Rahmen des Konzepts „Eisenbahnknoten Stuttgart 2040“ an (VMBaWü 2023), von dessen über 100 geplanten Einzelmaßnahmen (mit einem Gesamtumfang von etwa vier Milliarden Euro) möglichst viele durch schlankere und auf den DKS aufsetzende Alternativen ersetzt werden sollen.

Eine von vielen im Gesamtsystem Bahn klug gestaltete Digitalisierung der Leit- und Sicherungstechnik ist dabei alternativlos. Angesichts eines inzwischen für den rein konventionellen Ausbau aufgetürmten Mittelbedarfs von mehreren hundert Milliarden Euro wäre es – selbst bei maximaler Mittelbereitstellung – schlicht aussichtslos, die hochgesteckten Erwartungen an die Eisenbahn in Deutschland für Verkehrswende und Klimaschutz in den noch verbleibenden Zeiträumen auch zu erfüllen.

Literaturverzeichnis

[1] Behrens, Marc u. a.: Schnelle Leit- und Sicherungstechnik für mehr Fahrwegkapazität. Der Eisenbahningenieur (2021), Nr. 6 (https://bit.ly/2SlQvjY).

[2] Behrens, Marc u. a.: Robuste Leit- und Sicherungstechnik im Digitalen Knoten Stuttgart. Der Eisenbahningenieur 11/2022 (https://bit.ly/3hiu0ZL).

[3] Beyer, Martin u. a.: ETCS als Trägersystem zur Leistungssteigerung bei der S-Bahn Stuttgart. Signal+Draht (2019), Nr. 6 (https://bit.ly/2MJ4zAY).

[4] Beyer, Martin u. a.: Der Digitale Knoten Stuttgart wird Realität. Der Eisenbahningenieur (2023), Nr. 1 (https://bit.ly/3RCeqFR).

[5] Bitzer, Florian u. a.: Quo vadis Digitale Leit- und Sicherungstechnik? Der Eisenbahningenieur 11/2021 (https://bit.ly/3Hv72X6).

[6] Bundesministerium für Verkehr und digitale Infrastruktur: Bekanntmachung der Richtlinie zur Förderung der Ausrüstung von Schienenfahrzeugen mit Komponenten des Europäischen Zugsicherungssystems ERTMS (European Rail Traffic Management System) und des automatisierten Bahnbetriebs (ATO) im Rahmen der infrastrukturseitigen Einführung von ERTMS im „Digitalen Knoten Stuttgart“. Bundesanzeiger, BAnz AT 05.02.2021 B2 (https://bit.ly/3hX5CJx), 1. (1), 2. Anstrich.

[7] Cyril, Gabriel u. a.: Nachrüstung von Nebenfahrzeugen für den Digitalen Knoten Stuttgart. Der Eisenbahningenieur (2023), Nr. 6.

[8] Castano, Alfred u. a.: Umtrassierung des Nordkopfs Ulm während der Bauausführung. Der Eisenbahningenieur 12/2022 (https://bit.ly/40r0x0P).

[9] Chavalier, Didier u. a.: FRMCS-Ausrüstung von 463 Triebzügen für den Digitalen Knoten Stuttgart. Signal+Draht (2023), Nr. 5 (https://bit.ly/3C5ZetG).

[10] DB Netz 2023: DB Netz: ETCS & Co. für maximale Leistungsfähigkeit. Eine Einführung in den Digitalen Knoten Stuttgart. Vortrag auf der FBS-Anwendertagung, 20. April 2023 (https://bit.ly/43yDNxs), S. 28, 42 f.

[11] Denißen, Jonas u. a.: Optimierung der Blockteilung mit ETCS Level 2 im Digitalen Knoten Stuttgart. Signal+Draht (2021), Nr. 7+8 (https://bit.ly/3Ai0gQR).

[12] Deutsche Bahn: Stuttgart 21 ist wesentliche Voraussetzung für den geplanten Deutschland-Takt. Statement des Konzernbevollmächtigten Thorsten Krenz vor dem S21-Ausschuss des Stuttgarter Gemeinderats, 16. Juli 2019 (bit.ly/43NSwVL).

[13] Dietrich, Frank u. a.: Fahrzeugnachrüstung für den Digitalen Knoten Stuttgart. In: Der Eisenbahningenieur (2021), Nr. 9 (https://bit.ly/3tFQWUB).

[14] Dietrich, Frank u. a.: Nachrüstung von 333 Triebzügen für den Digitalen Knoten Stuttgart. ZEVrail (2022), Nr. 5 (https://bit.ly/3DHZl0S).

[15] Dietrich, Frank u. a.: Förderung der DSD-Fahrzeugausrüstung im Digitalen Knoten Stuttgart. Der Eisenbahningenieur (2023), Nr. 4 (https://bit.ly/3N24h5o).

[16] Druckenbrod, Constantin u. a.: Neue Doppelstocktriebzüge für den Digitalen Knoten Stuttgart. Der Eisenbahningenieur (2023), Nr. 2.

[17] Enzmann, Andre: Trassierungsfeinschliff: Millimeterarbeit mit großem Nutzen. Der Eisenbahningenieur (2021), Nr. 4 (http://www.bsu.link/task-force-trassierung).

[18] Flöter, Christian: Innovationskooperation Fahrzeugausrüstung im Digitalen Knoten Stuttgart. Signal+Draht (2022), Nr. 9 (https://bit.ly/3dxD0Z6).

[19] Förster, Jonas u. a.: ETCS-Bremskurven im Spiegel der Praxis. Der Eisenbahningenieur (2023), Nr. 6.

[20] Goers, Hannes u. a.: ETCS als Träger für Leistungs- und Qualitätssteigerungen. 9. Januar 2019 (https://bit.ly/3nHFB7h).

[21] Hernández, Lylly u. a.: Schaltabschnittsgrenzen und Bahnübergänge schränken Kapazitätseffekt von ETCS Level 2 ein. Signal+Draht 1+2/2023 (https://bit.ly/40AY6br).

[22] Ingenieurgemeinschaft Machbarkeitsstudie ETCS S-Bahn Stuttgart: Untersuchung zur Einführung von ETCS im Kernnetz der S-Bahn Stuttgart (https://bit.ly/2Yyaw6h), insbes. S. 279, 299.

[23] Kümmling 2021: Kümmling, Michael u. a.: Maximierung der Fahrwegkapazität mit Digitaler Leit- und Sicherungstechnik, Eisenbahntechnische Rundschau (2021), Nr. 7+8 (https://bit.ly/3eYOapT).

[24] Ohmayer, Roman: Optimierung von Überholvorgängen mit digitaler Leit- und Sicherungstechnik. Bachelorarbeit Mai 2022 (https://bit.ly/3BbuPJR).

[25] Reinhart, Peter u. a.: Kleiner Aufwand – großer Nutzen: Trassierungsfeinschliff am Beispiel des Projekts Stuttgart–Ulm. Foliensatz zum Gleisbauseminar 2021, 16. April 2021 (http://www.bsu.link/task-force-trassierung).

[26] Ministerium für Verkehr Baden-Württemberg: Vorschlag eines Ausbaukonzepts für den Eisenbahnknoten Stuttgart 2040. https://vm.baden-wuerttemberg.de/fileadmin/redaktion/m-mvi/intern/Dateien/PDF/230313_Zukunftspapier_Eisenbahnknoten_Stuttgart.pdf, abgerufen am 20. Juni 2023.

[27] Vogel, Thomas u. a.: Kleiner Aufwand: große Wirkung: Fahrzeugausrüstung im Digitalen Knoten Stuttgart. Vortrag auf dem 22. SIGNAL+DRAHT-Kongress (https://bit.ly/3F9Smht), PDF-Seite 16.

[28] Vogel, Thomas: An einer klugen Digitalisierung führt kein Weg vorbei. In: Der Eisenbahningenieur (2023), Nr. 2.

Dieses Projekt wird kofinanziert durch die Europäische Union.