Rail Security Expert Group

Security Logging Guideline

23E177
1A
25.04.2024

Modification history

Version	Date	Modification / Description	Editor
1A	25.04.2024	Final version after review	Jorge Gamelas, Oliver Lovric, Richard Poschinger, Nicolas Poyet, Max Schubert

Table of Contents

Table of Figures

1 Introduction

1.1 Scope

1.1.1.1 The purpose of this document is to give guidance on architectural aspects of log data and SIEM infrastructures. Furthermore, corresponding processes are defined and explained.

1.2 References

1.2.1.1 Subsets and EUG publication are referenced directly with their corresponding ID.

1.2.1.2 Other referenced documents:

[1] "RFC 2119," 1997. [Online]. Available: https://datatracker.ietf.org/doc/html/rfc2119.

[2] „Mitre ATT&CK," [Online]. Available: https://attack.mitre.org/.

[3] „NIST SP800-92," [Online]. Available: https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-92.pdf.

1.3 Abbreviations

IM .... *Infrastructure Manager*
OT .... *Operational Technology*
RU .... *Railway Undertaking*
SIEM .... *Security Incident and Event Management*
SOC .... *Security Operations Centre*

ERTMS Abbreviations are listed in SUBSET-023

1.4 Authors

1.4.1.1 The Rail Security Expert Group (RSEG) consists of security experts of the following groups:

- ERTMS Security Expert Group (ESCG) – Part of the EEIG ERTMS Users Group
- EULYNX Security Cluster – Part of the EULYNX Initiative

1.4.1.2 The following members of the Rail Security Expert Group were involved in creating this document:

- ERTMS User Group (EUG) / EULYNX
  - Max Schubert
  - Richard Poschinger
- SBB
  - Oliver Lovric
- SNCF
  - Nicolas Poyet
- Trafikverket
  - Jorge Gamelas

1.5 Applicability and Document Status

1.5.1.1 In order to ensure the usability for tender documents, this document is using classifications and requirement key words. This classification does not result in any binding requirements for members of the EUG or other involved parties. The documents will be updated in the future to be adapted to a changed threat landscape, updated standards, and newly developed security solutions.

1.6 Definition of Requirement Types

1.6.1.1 This document uses key words indicating requirement levels according to RFC 2119 [1]. Each clause in this document is classified as follows:

M	Mandatory	function must be implemented as specified
O	Optional	not mandatory, must be as specified if implemented
I	Informative	included for clarification purposes only
R	Recommendation	included as recommendation

Texts without a tag do not constitute a requirement.

2 Log Management Infrastructure

2.1 Architecture (architecture, not detailed interface description)

2.1.1.1 Before one can start logging information, an overall security logging strategy and architecture shall be developed. **(I)**

2.1.1.2 The following decisions must be made before the start of the implementation: **(I)**

- users or user groups of the SOC/SIEM services
- number of SIEM and SOC services
- log aggregation locally or centrally

2.1.1.3 Typical user groups of SOC and SIEM services in a railway system are: **(I)**

- Business IT
- Operation IT systems for maintenance or monitoring
- OT systems trackside and on-board like interlockings or trains.

2.1.1.4 Per user groups different use-cases are applied depending on their organisational and regulatory conditions and technical needs. **(I)**

2.1.1.5 The use-case definition and tuning should be accomplished by a combined team of Security Analysts and experts of the user group. **(R)**

2.1.1.6 Within one company one centralized view on security incidents should be created. Generally, there are two solutions practically approved and recommendable to achieve this goal: **(R)**

- *One single SIEM*
- *Federated system with multiple SIEM with aggregation to an overarching system, e.g. a CSIRT*

2.1.1.7 The SIEM shall be capable of managing clients to allow different user categories while accessing the same stack of log information. **(M)**

2.2 Evaluation of Variants

2.2.1.1 The following variants of task and infrastructure separation are evaluated in detail. All variants contain a green part consisting of infrastructure and personnel belonging to the operator. In some cases, the SOC personnel (purple) is partly outsourced. Furthermore, the infrastructure (orange) might be provided by an external party as well. Following the Variants are shown and described. **(I)**

2.2.1.2 Variant A: **(I)**

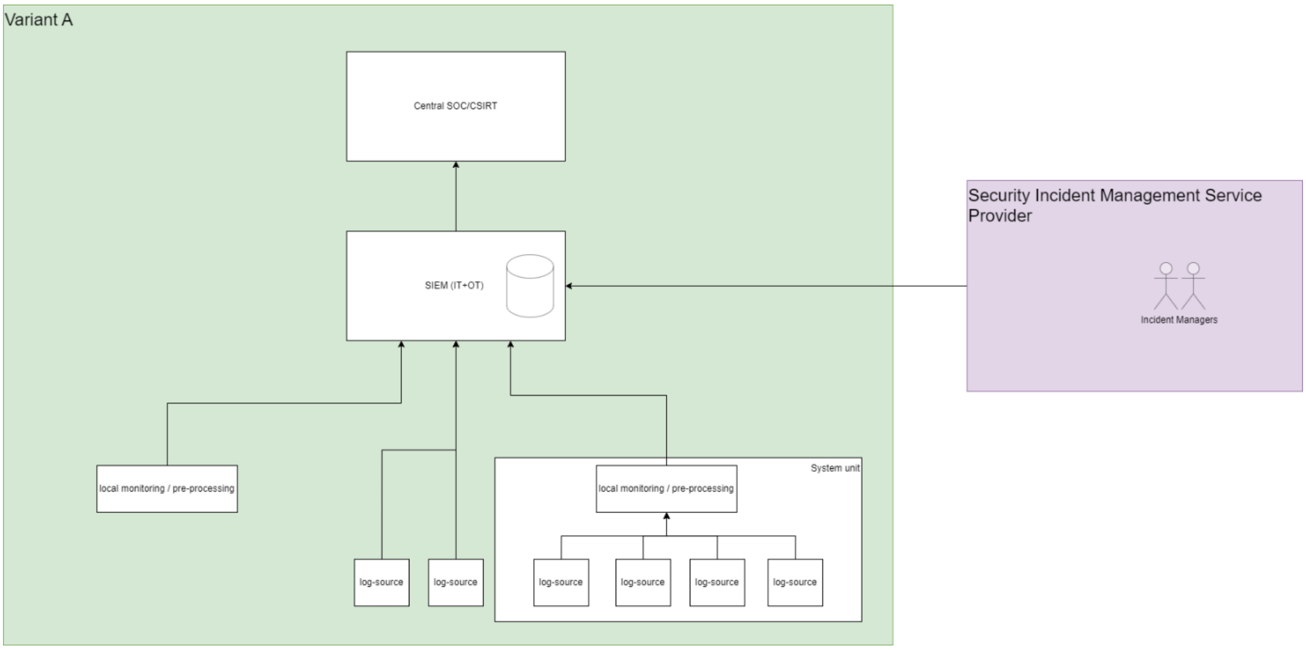

Figure 1: Variant A

Variant A (Figure 1) consists of an internal log infrastructure. The SIEM for IT and OT is provided by the operator. This SIEM is fully operated and maintained by external personnel (Incident Managers). Resulting data is provided to the internal SOC/CSIRT.

2.2.1.3 Variant B: **(I)**

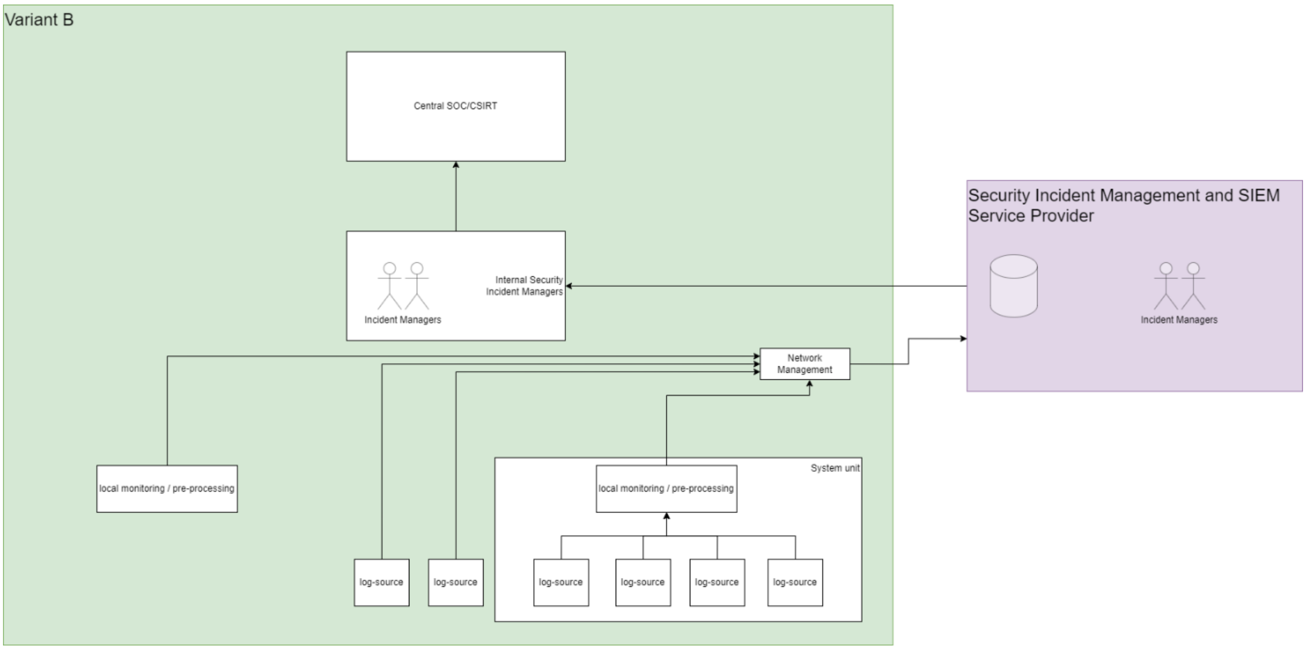

Figure 2: Variant B

In Variant B (Figure 2) the Incident Management is provided by an external provider similar to Variant A. Furthermore, the SIEM infrastructure is outsourced and operated in the infrastructure of the external SIEM provider. The operator provides its own Incident Managers in addition to external personnel.

2.2.1.4 Variant C: **(I)**

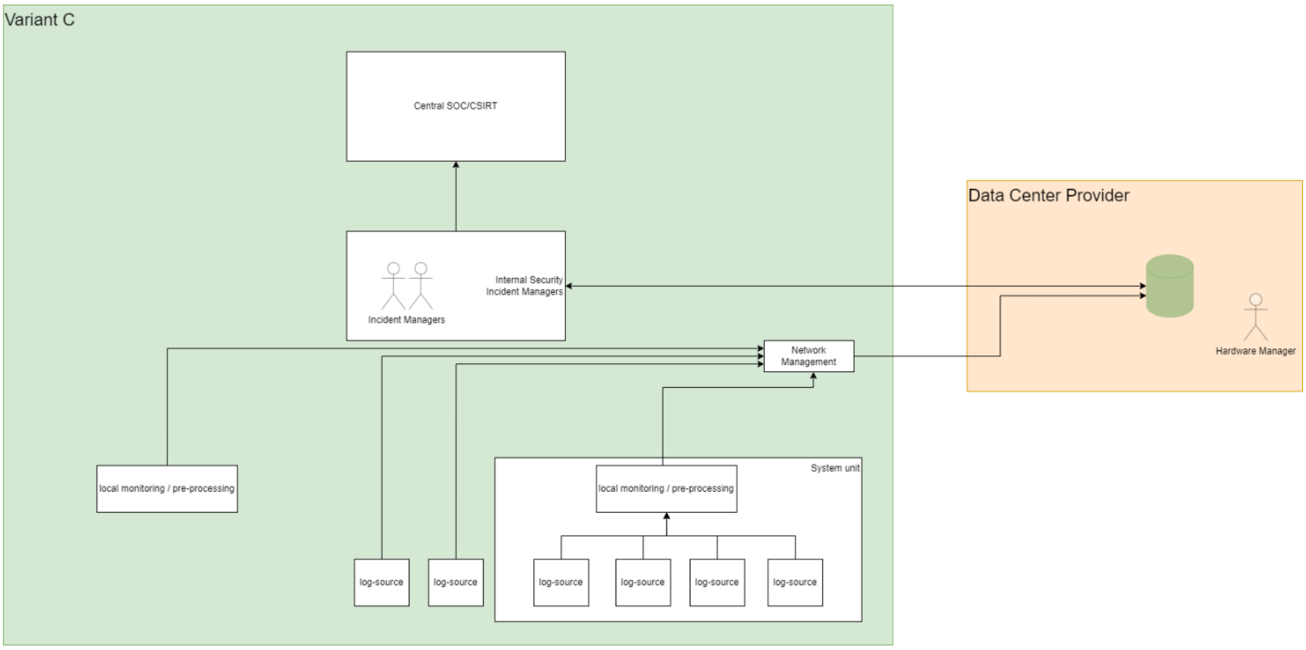

Figure 3: Variant C

The Design and Decision making as well as the Incident Management in Variant C (Figure 3) is performed internally by the operator. Instead of hosting the SIEM solution internally, it is running in an external data centre, so that resources from available, nationally trusted data centre providers can be used. A change of location can be always easily initiated as the full control over the systems stays at the operator site. The data centre provider is only responsible for the management of hardware and has no access to the SIEM data.

2.2.1.5 Variant D: **(I)**

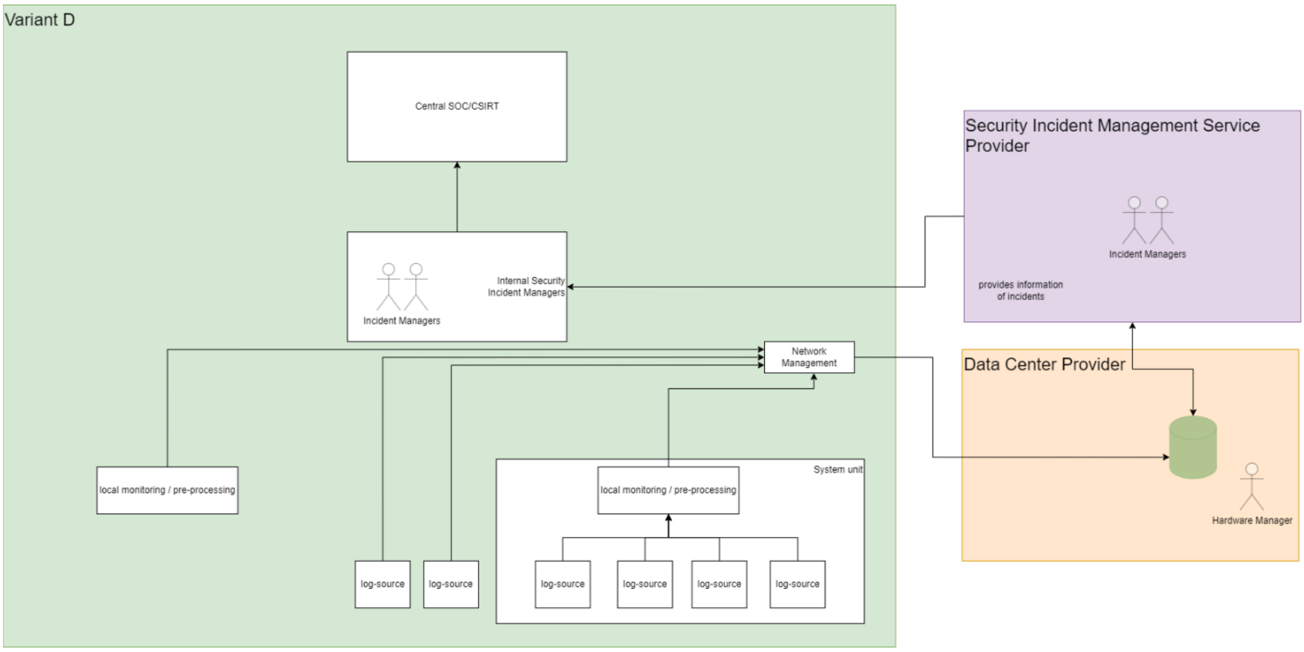

Figure 4: Variant D

Similar to Variant C, the SIEM is hosted in an external data centre (Figure 4) with all its possibilities and the full ownership of the operator. The operator has full control over the data in the SIEM. Internal staff of the operator is designing and leading the SIEM building and development. The internal personnel are supported by external experts. The Incident Management (1st, maybe 2nd level support) is provided by external companies. Playbook design and writing as well as local expertise and reaction after 1st and 2nd level stay with the operator. The external Incident Managers have access to the SIEM. The external Security Incident Management Service provides pre-analysed data to the internal staff of the operator.

2.2.1.6 Variant E: **(I)**

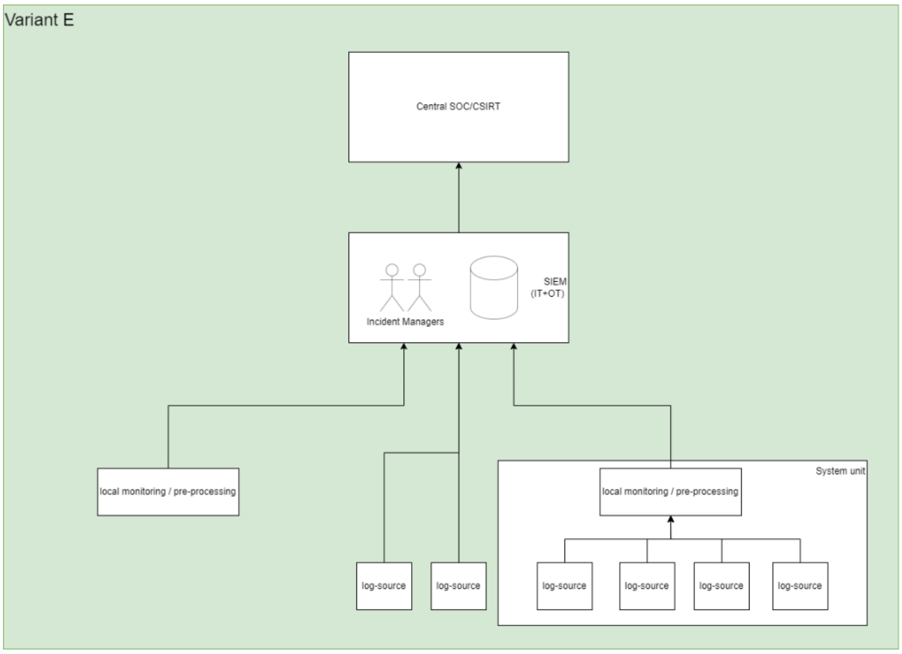

Figure 5: Variant E

Variant E (Figure 5) consists of a logging and SIEM infrastructure as well as personnel provided fully internally. No external organisation or data centre is involved.

2.2.1.7 Rationale on SIEM ownership and operation: **(I)**

Criteria	Weight	Variant A	Variant B	Variant C	Variant D	Variant E
Requirement of ownership	2	3	1	3	3	3
Time to operation	1	2	3	2	2	1
Security constraints (trust/confidentiality)	3	2 (national certificatio n required)	1 (national certificatio n required)	2	2 (national certificatio n required)	3
Internal Personnel required	1	3	2	1	2	1
Maintain knowledge	1	1	2	3	2	3
Maintain operational readiness	2	3	2	2	2	2
Expandability (scalability)	2	2	3	2	3	1
CAPEX	1	2	1	2	3	1
OPEX	2	2	3	2	3	2
Weighted Sum		**34**	**29**	**32**	**37**	**31**

2.2.1.8 Criteria definition: **(I)**

- Requirement of ownership:
  Measures the variant's ability to comply with regulations requiring the ownership of the solution by the infrastructure manager/railway undertaking. This is a legal requirement in some countries. Furthermore, the ownership of the solution is improving the security regarding the control over data and functionalities and long-term planning.
  1 = not completely owned by IM/RU
  2 = not defined
  3 = completely owned by IM/RU
- Time to operation:
  Measures the variant's time to be put into operation.
  1 = Long
  2 = Middle
  3 = Short

- Security constraints (trust/confidentiality):
  Measures the variant's ability to comply with security constraints related to e.g., trust and confidentiality.
  1 = Hard to comply
  2 = Middle to comply
  3 = Easy to comply
- Internal Personnel required:
  Measures the number of Incident Managers for SIEM needed to be hired by the operator to build and operate the variant under evaluation.
  1 = Many
  2 = Some
  3 = None
- Maintain knowledge:
  Possession and ability to apply needed knowledge to operate, maintain and improve the system now and in the future, including transfer of knowledge under circumstances of changing responsibilities/personnel manager/railway.
  1 = Low level of knowledge possessed
  2 = Middle level of knowledge possessed
  3 = High level of knowledge possessed
- Maintain operational readiness of SOC staffing:
  Measures the level of availability and flexibility of SOC staffing.
  1 = Low availability and flexibility
  2 = Average availability and flexibility
  3 = High availability and flexibility
- Expandability/scalability:
  Measures the variant's ability to expand/scale for future needs.
  1 = Not expandable / scalable
  2 = Difficult to expand / scale
  3 = Easy to expand / scale
- CAPEX:
  Measures the cost of building and commissioning the variant.
  1 = High
  2 = Medium
  3 = Low
- OPEX:
  Measures the cost of operating the variant.
  1 = High
  2 = Medium
  3 = Low

2.2.1.9 Conclusion: **(R)**

Based on the evaluation, variant D is the preferred one. It gives the best combination of ownership, control over the system, flexibility of service providers and time to operation. The main factors in favour of variant D are:

- The design and main knowledge of the system is in house of the operator.
- The hardware is owned by the operator, so no law constraints should occur and moving is always possible.
- The time to operation is reduced due to the use of existing, professional data centre services and security incident monitoring teams.
- The operator has full control over knowledge, ownership and contract partners, vendor lock-in risk is reduced to a minimum.

2.2.1.10 The operator shall define requirements for the security clearance of the people that are hired for the security services. **(R)**

3 Log Management Operational Processes and Technical Solutions

3.1 Security Use-Case Definition

3.1.1.1 A use case is a mix of technical rules and/or actions within a SIEM tool. It converts threats to business processes and activities into SIEM technical rules and actions which can detect events of interest such as possible compromise of user credentials, escalation of privileges, non-compliances, etc. **(I)**

3.1.1.2 The definition of use-cases should be based on identified threats and according attack strategies. **(R)**

3.1.1.3 The definition of the use-case should be linked to best-practice frameworks like Mitre ATT&CK [2]. **(R)**

3.1.1.4 There are mainly two different types of use-cases. First there are generic use-cases that are widely used, independent from the system environment. Second there are system or component specific use-cases. **(I)**

3.1.1.5 It is recommended to implement and use generic use-cases first, to quickly gain a view on the system. These use-case are usually easy to implement and configure as they are prepared by the SIEM software used already. **(R)**

3.1.1.6 In a second stage, the more complicated, system or component specific use-cases can be defined and applied. **(R)**

3.1.1.7 For legacy systems in many cases the standard use-cases can’t be applied as the legacy system does not provide the needed log information. **(I)**

3.1.1.8 The use-case definition should always contain: **(R)**

- Use-Case Name
- System/Component relation
- Goal of the Use-Case (the threat/attack that can be covered)
- Actors, stakeholders, responsibilities
- Pre-conditions
- Required log information and frequency
- Threshold/Value/Anomaly value that triggers the alarm
- Planned reaction / Play book

3.2 Create security logs for legacy systems

3.2.1.1 Existing system differ in their capabilities to generate and send log information. In general, one can distinguish between the following capability classes: **(I)**

- Variant 1. The component can generate logs and provides them at an interface in directly usable format.
- Variant 2. The component can generate logs but does not provide them at an interface or not in directly usable format.
- Variant 3. The component can't generate logs.

3.2.1.2 Variant 1 are usually routers, switches, firewalls or other standard IT-systems in the network. **(I)**

3.2.1.3 For Variant 1 the following steps to integrate the logs in the SIEM need to be performed: **(R)**

- Configure the component.
- Configure the network to allow the connectivity to the SIEM.

3.2.1.4 For Variant 1 homologation constraints may apply, if the components are part of safety homologated system. Constraints could be additional ports or more traffic that might generate delay on the network. This needs to be analysed and managed before applying the measure. **(R)**

3.2.1.5 Variant 2 are usually maintenance or monitoring system for existing systems like interlockings or level crossings. **(I)**

3.2.1.6 For Variant 2 the following steps to integrate the logs in the SIEM need to be performed: **(R)**

- If the component is capable of sending information, it is recommended to send these in the format they are available and parse (normalise) them outside of the component. The parsing should be managed at decentralised pre-processing system or at a central location (SIEM).
- If the component is not capable of sending information, a modification of the capability might be possible. In this case the need of a re-homologation is very likely. This is not recommendable. In this case the application of variant 3 measure is proposed.

3.2.1.7 Variant 3 are usually safety systems, like the interlocking or level crossing. **(I)**

3.2.1.8 For Variant 3 the following steps to integrate the logs in the SIEM need to be performed: **(R)**

- Generate the logs on network basis as the estimation is that there is no encryption applied and thus information are available when “listening” to the network.
- For the application of a network tap the homologation needs to be checked, if it is required by the national authority. Technically a network tap must fulfil the requirement of “no change to data stream” which fulfils the safety requirement of “feedback free”.
- To allow feedback free extraction of information a data diode might be required.

3.2.1.9 For variant 2 and 3, additional software and/or hardware is needed to generate logs. For this purpose, the following requirements apply: **(R)**

- The software/hardware should be configurable and updateable over the time to support new upcoming threats or vulnerabilities and thus support new use-cases.
- The configuration and update process should be under the responsibility and access of the railway undertaker or operator.
- The information of the log generation and analysis should be fully transparent to the railway undertaker or operator.

3.2.1.10 For all variants the following additional log sources support the overall view on the system for security purposes. **(I)**

- Integration of physical system log information, like access to the data centre or interlocking room.
- Extraction of network data in general (for all variants) may support the overall view on the system. The network data might be extracted in different network zones of the rail system to get a “complete” overview of the system.

3.3 Log Source Configuration

3.3.1.1 Based on the defined use-cases and the analysis of the log sources, the configuration can be defined. **(I)**

3.3.1.2 If the required log information to perform the use-case can be made available by the log sources, the log source should be configured accordingly. **(R)**

3.3.1.3 If the required log information to perform the use-case cannot be made available by the log sources, the use-case should be adapted to the available information. **(R)**

3.3.1.4 Providing the log information is not a continuous data stream but needs to be configured according to the use-case. The following criteria should be taken into consideration and documented before the configuration of the system. **(R)**

- Frequency of sending the log information. This might be time-based or event-based
- Criticality level definition of the log message

3.4 Local collection and pre-processing

3.4.1.1 There are multiple options for extracting and pre-processing security related and relevant information from components. Following the different possibilities are described. **(I)**

- **Pre-Collector** can collect all log information made available from a component and send it to the SIEM. It has no own use-case capabilities.
- **Pre-Processing Unit** can collect all data that is made available from a component and send it to the SIEM. An integration of filtering, use-cases, intrusion detection (IDS) -> only sending relevant and already analysed data to the SIEM is possible. Flexible adjustment of use-cases/Filtering possible as the system is usually not vendor bound.
- **Host based IDS** agent on the (Safety-) Component can do intrusion detection (mostly based on netflow) on the component directly. If accessible, the IDS agent can also access further logging and monitoring data of the component (might fail, if the application sends encrypted data, ...)
- The component could also directly send security and log information to the **SIEM**. The SIEM is the central security information and event management which can process huge amount of data and logs. It integrates IDS, use-cases and further capabilities. Further it does alarming, provides output to the user and integration of playbooks for reaction.
- **Switch IDS** is an agent on the Switch (host based) that can do intrusion detection (mostly based on netflow) on the network. If accessible, the IDS agent can also access further logging and monitoring data of the Switch (might fail, if the application sends encrypted data, ...)
- An **IDS** is a dedicated hardware, connected to the network. It needs a network TAP capability, a connection to a Network TAP or a mirror port of a switch to get access to the relevant network data. It can do intrusion detection (mostly based on netflow) on the network. If accessible (for legacy it's usually not), the IDS can also access further logging and monitoring data of the relevant components. (might fail, if the application sends encrypted data, ...).
- A **Network TAP** is a dedicated hardware that is connected to the network for feedback free reading and providing all network data to another analysis system.

3.4.1.2 Criteria definition: **(I)**

- CAPEX:
  Measures the cost of building and commissioning the solution.
  1 = High
  2 = Medium
  3 = Low
- OPEX:
  Measures the cost of operating the solution.
  1 = High
  2 = Medium
  3 = Low

- Ownership:
  Measures the solution's ability to comply with regulations requiring the ownership of the solution by the infrastructure manager/railway undertaking. This is a legal requirement in some countries. Furthermore, the ownership of the solution is improving the security regarding the control over data and functionalities and long-term planning.
  1 = not completely owned by IM/RU
  2 = not defined
  3 = completely owned by IM/RU
- Fit for use:
  Measures the level of adaptation/changes to the existing system required by the solution before it can be operational.
  1 = High (More Adaptations)
  2 = Moderate
  3 = Low (Less Adaptations)
- Legacy:
  Measures the solution's ability to be implemented in legacy systems (existing systems). Is the solution capable of retrieving valuable information from legacy systems that support security monitoring (high) or is it only designed to work with state-of-the-art systems and gives no compatibility with legacy systems (low).
  1 = Low
  2 = Moderate
  3 = High
- Future:
  Measures the solution's ability to be implemented in future environments already considering logging and security standards.
  1 = Hard
  2 = Moderate
  3 = Easy

3.4.1.3 Then these solutions are evaluated against pre-defined criteria to provide a recommendation: **(I)**

First, they are analysed for legacy set-up under the pre-condition that the components are not able of sending security relevant log information (variant 3):

Solution	**Weight**	**Pre-Collector**	**Pre-Processing + Network TAP**	**IDS host based**	**SIEM direct**	**Switch IDS**	**IDS + Network TAP**
CAPEX	1	2	2	1	1	2	1
OPEX	2	3	3	2	2	2	2
Ownership	2	3	3	1	3	1	1
Fit for use	1	2	2	1	1	2	2
Legacy	3	1	2	1	1	1	2
Future	2	-	-	-	-	-	-
Weighted Sum		**19**	**22**	**11**	**15**	**13**	**15**

Secondly, they are analysed for future set-up under the pre-condition that the components in the future are able to send security relevant log information (variant 1):

Solution	Weight	Pre-Collector	Pre-Processing	IDS host based	SIEM direct	Switch IDS	IDS
CAPEX	1	2	2	2	3	2	2
OPEX	2	3	3	2	3	2	2
Ownership	2	3	3	1	3	1	1
Fit for use	1	2	2	2	3	2	2
Legacy	3	-	-	-	-	-	-
Future	2	2	3	2	3	2	2
Weighted Sum		**20**	**22**	**14**	**24**	**14**	**14**

3.4.1.4 The integration of the use-cases is crucial for the security capability of detection and further security related use-cases. This changes constantly over the life-cycle. Thus, it is recommended to use solutions that can be easily adopted over the life cycle. That is why, the following solution is recommended for legacy systems: **(R)**

- Network TAP with Pre-Processing Unit

3.4.1.5 For future systems, the following set-up is recommended: **(R)**

- SIEM directly
- Pre-Processing Unit

3.5 Secure information locally and on transit

3.5.1.1 The component shall retain log data locally until it gets acknowledgement that the central system has received it. **(M)**

3.5.1.2 If a decentralised collector and pre-processing system is used, the system should retain log data locally. **(R)**

3.5.1.3 The time for local retention for decentralised collectors and pre-processing systems should be based on the criticality of the device. A guideline is given by NIST SP800-92 [3]: **(R)**

- For low impact systems the collector shall retain information for 1 to 2 weeks.
- For medium impact systems the collector shall retain information for 1 to 3 months.
- For high impact system the collector shall retain information for 3 to 12 months.

3.5.1.4 The local collector and pre-processing unit shall integrity protect the data at rest. **(M)**

3.5.1.5 The local collector and pre-processing unit shall integrity protect the data in transit. **(M)**

3.5.1.6 For variant 1 the component shall protect the integrity of the data at rest. **(M)**

3.5.1.7 For variant 1 the component shall protect the integrity of the data in transit. **(M)**

3.5.1.8 For variant 2 the railway operator or undertaker shall apply measures or procedures to protect the security logs against manipulation. **(M)**

3.5.1.9 For variant 3 the requirements for local collector and pre-processing unit apply. **(M)**