Hinterfrage deine Annahmen: Praxiserfahrungen mit Testautomatisierung und Pentests

Question your assumptions: Practical experience with test automation and pentests

Martin Koop | Leon Hagemann | Alain Kaffo

In der modernen Eisenbahnindustrie spielen Penetrationstests (Pentests) und funktionale Tests eine zentrale Rolle für die Sicherstellung der Sicherheit und Zuverlässigkeit komplexer Betriebs- und Steuerungssysteme (Operational Technology – OT). Diese Tests dienen der Qualitätssicherung und sind entscheidend für den reibungslosen Betrieb sowie den Schutz vor Sicherheitsbedrohungen. Während funktionale Tests die korrekte Funktionsweise der Systeme überprüfen, zielen Pentests darauf ab, potenzielle Sicherheitslücken zu identifizieren. In diesem Beitrag werden die Unterschiede und Bedeutung sowie die Planung, Durchführung und Herausforderungen anhand zweier Projekte rund um das Radio Block Centre (RBC) sowie im bahnbetrieblichen Übertragungssystem des Digitalen Stellwerks (DSTW) detailliert erläutert.

1 Unterschied zwischen Pentest und funktionalen Tests

Funktionale Tests stellen sicher, dass alle OT-Systeme und Software-Komponenten wie vorgesehen funktionieren und sämtliche spezifizierten Anforderungen erfüllen. Sie gewährleisten, dass die komplexen Systeme unter realistischen Betriebsbedingungen korrekt arbeiten und alle sicherheits- und betriebsrelevanten Prozesse zuverlässig ablaufen. Beispielsweise wird überprüft, ob Signale korrekt ausgelöst werden oder die Kommunikation zwischen Systemkomponenten nach Anforderung / Spezifikation funktioniert.
Pentests dagegen zielen darauf ab, potenzielle Sicherheitslücken zu identifizieren, die OT-Systeme gefährden könnten. Diese Tests simulieren die Methoden eines Angreifers, um beispielsweise Schwachstellen in der Netzwerksicherheit, der Software und den Konfigurationen aufzudecken. Dabei werden beispielsweise Funktionen über den „Rand“ ihrer Anforderungsspezifikation gebracht, um die unautorisierte Kontrolle des Systems zu erlangen, was der wesentliche Unterschied zu funktionalen Tests ist. Die Angriffssimulation wird oft auch als Red Teaming bezeichnet und in diesem Beitrag synonym verwendet.

2 Wieso müssen Pentests und funktionale Tests durchgeführt werden?

Angesichts der zunehmenden Digitalisierung und Vernetzung von IT / OT-Systemen wird die Gewährleistung ihrer Sicherheit immer wichtiger. Der Eisenbahnsektor unterliegt hier strengen gesetzlichen Vorschriften und Standards auf europäischer und nationaler Ebene, die spezifische IT / OT-Sicherheitsanforderungen, darunter die Durchführung von Penstests und funktionalen Tests, fordern (siehe Tab. 1).

Penetration tests (pentests) and functional tests play a central role in the modern railway industry when ensuring the safety and reliability of complex operational technology (OT) systems. These tests are used for quality assurance and are crucial for smooth operations and protection against any security threats. While functional tests verify the correct functioning of systems, pentests aim to identify any potential security vulnerabilities. This article explains the differences between these two test types in detail, as well as their importance and implementation, with reference to two projects related to an Radio Block Centre (RBC) and a communication network within a digital interlocking (DSTW).

1 The difference between pentests and functional tests

Functional tests ensure that all the OT systems and software components are functioning as intended and that they meet all the specified requirements. They guarantee that the complex systems are working correctly under realistic operating conditions and that all the safety, security and any operationally relevant processes are running reliably. For example, tests are undertaken to ascertain whether signals are triggered correctly or the communication between system components is functioning according to the requirement specifications.
Pentests, on the other hand, aim to identify any potential security vulnerabilities that could jeopardise the OT systems. These tests simulate the methods used by an attacker to uncover any vulnerabilities in the network security, software and configurations. For example, system functions are taken beyond the “edge” of their requirement specification in order to gain unauthorised control of the system, which is the main difference compared to functional tests. The attack simulation is also often referenced as Red Teaming and used synonymously in this article.

2 Why do pentests and functional tests have to be carried out?

Ensuring the security of IT / OT systems is becoming increasingly important in light of the progressively digitalisation and communication within them. The railway sector is subject to strict legal regulations and standards at the European and national levels that demand specific IT / OT security requirements, including the performance of pentests and functional tests (see tab. 1).

3 The advantages of functional tests in the railway sector

Functional tests are part of the verification process when implementing the defined specification requirements and are therefore crucial for ensuring the safety, reliability and efficiency of

Anforderungen	NIS 2	EU Cyber Resilience Act	EU Cyber-Security Act	TSI ZZS	CENELEC 50128, 50129	BSI IT-SiG / KRITIS	ISO 27001/2:2022	ISO/IEC 62443	TS 50701:2023
Pentest	Nr. 58	Anhang 1	Chapter VI, Nr. 52	Verschiedene (Interoperabilität)	Abschnitt 6 & 7	§ 8a	Nr. 8.29	62443-3-3, 62443-4-1	Abschnitt 9.3.2
Testing	Nr. 15	Nr. 44	Chapter VI, Nr. 51-53	Verschiedene (Interoperabilität)	Abschnitt 6 & 7	§ 8a	Nr. 8.29	62443-3-3, 62443-4-1	Abschnitt 9.3.2

Tab. 1: Gesetzliche und normative Anforderungen für Pentests und funktionale Tests Quelle / Source: INCYDE

Requirements	NIS 2	EU Cyber Resilience Act	EU Cyber-Security Act	TSI CCS	CENELEC 50128, 50129	BSI IT-SiG / KRITIS	ISO 27001/2:2022	ISO/IEC 62443	TS 50701:2023
Pentest	Nr. 58	Appendix 1	Chapter VI, Nr. 52	Various (Interoperability)	Section 6 & 7	§ 8a	Nr. 8.29	62443-3-3, 62443-4-1	Section 9.3.2
Testing	Nr. 15	Nr. 44	Chapter VI, Nr. 51-53	Various (Interoperability)	Section 6 & 7	§ 8a	Nr. 8.29	62443-3-3, 62443-4-1	Section 9.3.2

Tab. 1: The legal and normative requirements for pentests and functional tests Quelle / Source: INCYDE

3 Vorteile funktionaler Tests im Eisenbahnsektor

Funktionale Tests sind Teil der Nachweisführung zur Umsetzung der definierten Lastenheftanforderungen und damit entscheidend für die Sicherstellung der Sicherheit, Zuverlässigkeit sowie Effizienz der Bahninfrastruktur. Diese Tests ermöglichen die frühzeitige Erkennung von Fehlern, wodurch potenzielle Betriebsstörungen und Unfälle vermieden werden können. Sie minimieren Risiken, indem sie sicherstellen, dass alle sicherheitskritischen Systeme und Komponenten zuverlässig arbeiten.
Durch funktionale Tests wird die Qualität der OT-Systeme kontinuierlich überwacht und verbessert, was zu einer höheren Betriebsqualität führt. Diese Tests tragen auch dazu bei, spätere Kosten zu vermeiden, indem sie Probleme frühzeitig identifizieren und beheben, bevor sie zu teuren Reparaturen und langen Ausfallzeiten führen. Dabei ist es essenziell, dass Testfälle schon früh vor der Integration und auch bei jeder Änderung, vor allem im Betrieb, durchgeführt werden.
Ein weiterer zentraler Aspekt ist die Sicherstellung der Kompatibilität neuer und geänderter Bestandssysteme.

3.1 Warum funktionale Tests automatisiert werden sollten

Die Komplexität moderner OT-Systeme im Eisenbahnsektor, wie im Projektbeispiel des zu testenden DSTW-Übertragungssystems, erfordert eine Automatisierung der Testfälle, um die notwendige Effizienz und Geschwindigkeit der Testzyklen zu gewährleisten. Im Projektbeispiel wäre die enorme Anzahl an mehreren tausend Netzwerkkomponenten im Feld, mit deren resultierender Konfiguration in hunderten von Testfällen, manuell nicht testbar.
Nur durch die Automatisierung ist es möglich, die umfangreichen Testfälle zu bewältigen und schnell auf Änderungen reagieren zu können. Langfristig führt dies zu erheblichen Kosteneinsparungen und einer höheren Qualität der Komponenten. Automatisierte Tests bieten zudem konsistente und wiederholbare Ergebnisse, eliminieren menschliche Fehler und ermöglichen eine umfassendere Testabdeckung. Sie erleichtern Regressionstests und gewährleisten, dass neue Änderungen keine bestehenden Funktionen beeinträchtigen.

3.2 Tools zur Testautomatisierung

Für die Testautomatisierung der Firewall-Komponenten im Übertragungssystem des DSTW wurde ein skalierbares Setup realisiert (Bild 1). Dieses ermöglicht neben der Automatisierung des Testprozesses eine detaillierte Überwachung und Analyse der Testergebnisse in Echtzeit, eine zentrale Steuerung sowie verständliche Visualisierung.
Kibana dient dabei als Benutzeroberfläche für **Elasticsearch**, die zentrale Datenbank zur Speicherung und Abfrage von Logs und Me-

the railway infrastructure. These tests enable faults to be recognised at an early stage, which means that potential disruptions to operations and accidents can be avoided. They minimise risks by ensuring that all the safety-critical systems and components are working reliably.
Functional tests continuously monitor and improve the quality of the OT systems, resulting in a higher operating quality. These tests also help avoid any later costs by identifying and resolving problems at an early stage before they can lead to expensive repairs and long downtimes. It is essential that the test cases are carried out at an early stage before integration and also in relation to every change, especially during operations.
Another key aspect involves ensuring the compatibility of any new and modified existing systems.

3.1 Why functional tests should be automated

The complexity of modern OT systems in the railway sector, as in the project example involving the communication network for the DSTW, requires the automation of the test cases in order to ensure the necessary efficiency and speed of the test cycles. In the project, the enormous number of several thousand network components in the field, with their configuration resulting in hundreds of test cases, could not be tested manually.
Only automation makes it possible to manage the extensive test cases and to react quickly to any changes. This leads to considerable cost savings and higher component quality in the long term. Automated tests also provide consistent and repeatable results, eliminate human error and enable more comprehensive test coverage. They facilitate regression tests and ensure that any new changes do not affect the existing functions.

3.2 The test automation tools

A scalable setup was created for the test automation of the firewall components used in the DSTW communication network (fig. 1). In addition to automating the test process, this also enables detailed monitoring and analysis of the test results in real time, central control and clear visualisation.
Kibana serves as the user interface for **Elasticsearch**, the central database for storing and querying logs and metrics. **Grafana** provides a dashboard to visualise the data from **Elasticsearch** and **Prometheus**, which especially collects metrics from Jenkins, the automation server for Continuous Integration and Continuous Deployment (CI/CD). **Jenkins** runs automated pipelines that retrieve and execute test scripts from **Gitea**, a self-hosted

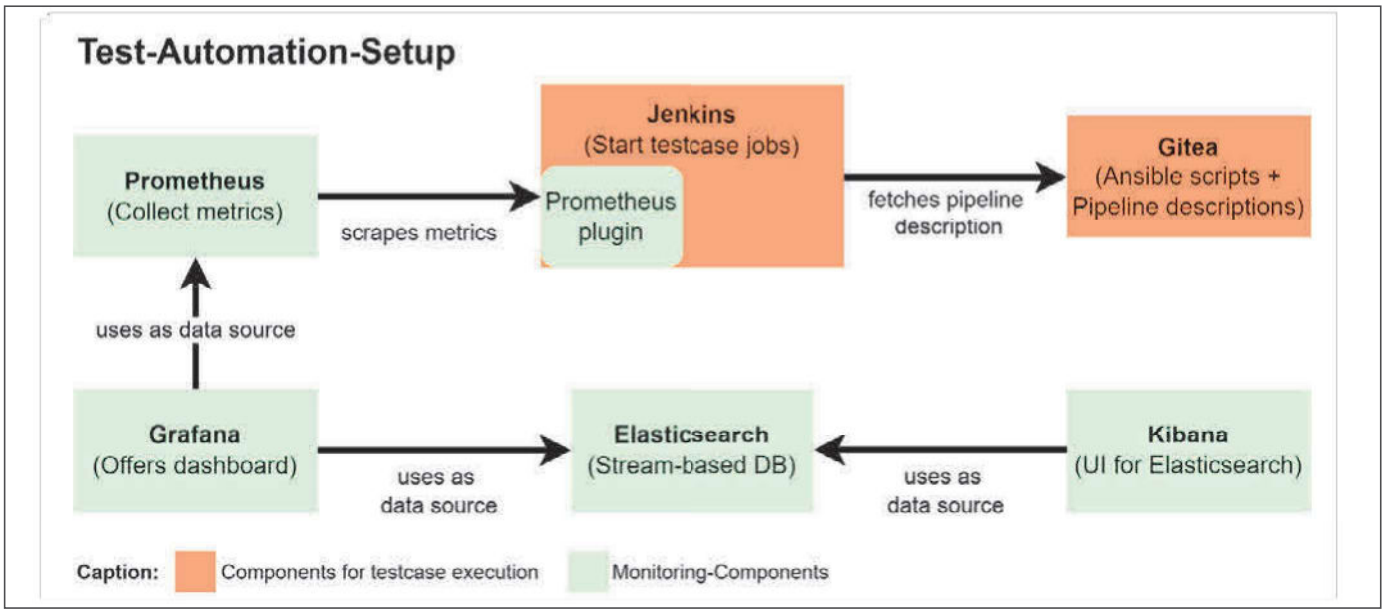

Bild 1: Setup der Testautomatisierungs-Infrastruktur
Fig. 1: The test automation infrastructure setup

Quelle / *Source*: INCYDE

triken. **Grafana** bietet ein Dashboard zur Visualisierung der Daten aus **Elasticsearch** und **Prometheus**, das Metriken sammelt, insbesondere aus Jenkins, dem Automatisierungsserver für Continuous Integration und Continuous Deployment (CI/CD). **Jenkins** führt automatische Pipelines aus, die Testskripte aus **Gitea**, einer selbstgehosteten Git-Diensteplattform, abrufen und ausführen. **Terraform** stellt die Skalierung der Server / Instanzen sowie die Konfiguration der Netzwerkkomponenten der gesamten Infrastruktur bereit.

3.3 Herausforderungen und Lösungen in der Testautomatisierung

Die Einführung und Implementierung funktionaler Tests sowie die Testautomatisierung stehen vor vielfältigen Herausforderungen. Hohe anfängliche Investitionen in Tools und Infrastruktur, Komplexität der zu testenden Systeme und die Erstellung entsprechender Testskripte stellen oft zu Beginn eine finanzielle Hürde dar. Ebenso können fehlende Fachkenntnisse im Team die Effektivität der Tests reduzieren. Maßnahmen zur Bewältigung dieser Herausforderungen umfassen die Auswahl von Open-Source-Lösungen, den Einsatz eines erprobten Automatisierungs-Setups und die Einstellung erfahrener Expertenteams.

Probleme stellen auch andauernde Änderungen von Anforderungen oder Systemupdates dar. Veraltete Testfälle, erhöhter Wartungsaufwand und Synchronisation mit allen Beteiligten sind die Folge. Die Etablierung von Prozessen und Integration einer Testautomatisierungs-Pipeline kompensierten dies im Projekt. Dabei ist es essenziell, Versionierungssysteme einzusetzen, um Änderungen der Lastenhefte und deren Testfälle nachzuverfolgen sowie die Abdeckung nachzuweisen. Zusätzlich zeigt sich die Planung von Code- und Peer-Reviews als sinnvoll, um korrekte und vollständige Tests sicherzustellen. Im optimalen Fall sollten modulare Testskripte parallel zur Komponentenentwicklung erstellt werden, um sicherzustellen, dass Tests auf dem neuesten Stand sind sowie leicht angepasst werden können.

Die Abhängigkeit von anderen Teilsystemen, eine gemeinsam genutzte Testumgebung und die resultierende Koordination mit unterschiedlichen Teams stellen eine zusätzliche Herausforderung dar. Diese spiegelt sich in Ressourcenkonflikten und instabiler Testum-

Git service platform. **Terraform** provides the scaling of the servers / instances and the configuration of the network components from the entire infrastructure.

3.3 The challenges and solutions in test automation

The introduction and implementation of functional tests and test automation face a variety of challenges. High initial investments in tools and infrastructure, the complexity of the systems to be tested and the creation of corresponding test scripts often represent a financial hurdle at the beginning. A lack of expertise in the team can also reduce the effectiveness of the tests. Measures to overcome these challenges include the selection of open source tools and the use of a proven and tested automation setup, as well as an experienced team of experts.

Ongoing changes to requirements or system updates also pose problems. This results in outdated test cases, increased maintenance costs and synchronisation with all the parties involved. The establishment of processes and the integration of a test automation pipeline has compensated for this. It is essential to use versioning systems in order to track the changes to the specifications and their test cases, as well as to prove coverage. In addition, the planning of code and peer reviews is also useful in order to ensure correct and complete tests. Ideally, modular test scripts should be created in parallel with the component development in order to guarantee that the tests are up to date and can be easily adapted.

The dependency on other subsystems, a shared test environment and the resulting coordination with different teams poses an additional challenge. This is reflected in resource conflicts and an unstable test environment.

In addition to careful planning and test strategies, the solution requires the implementation of a comprehensive monitoring and alerting system as well as the creation of an isolated test infrastructure on virtual container systems.

Last but not least, limited resources such as time, budget and a lack of support from management can also affect the implementation and acceptance of test automation. These challenges require a holistic approach, clear objectives and the development

gebung wider. Die Lösung erfordert eine sorgfältige Planung und Teststrategien, die Implementierung eines umfassenden Monitoring- und Alarmierungs-Systems sowie den Aufbau einer isolierten Testinfrastruktur auf virtuellen Container-Systemen.
Nicht zuletzt können beschränkte Ressourcen wie Zeit, Budget und die mangelnde Unterstützung seitens des Managements die Umsetzung und Akzeptanz der Testautomatisierung beeinträchtigen. Diese Herausforderungen erfordern eine ganzheitliche Herangehensweise, klare Zielsetzungen und die Entwicklung von Business Cases, welche die Qualitätsverbesserung und Kostenersparnis darstellen. Darüber hinaus wird die Akzeptanz verbessert durch die Einführung von kleinen Pilotprojekten, welche den Nutzen und die Effektivität der Testautomatisierung demonstrieren. Regelmäßige Reports, welche die kontinuierliche Verbesserung der Testprozesse darstellen, zeigen zusätzlich die Qualität und Effizienz der Testautomatisierung.

4 IT- vs. OT-Pentesting

IT-Pentesting konzentriert sich auf die Sicherheit von Informationssystemen, wie Servern, Netzwerken und Anwendungen, und kann häufig remote durchgeführt werden. OT-Pentesting hingegen befasst sich mit industriellen Steuerungssystemen wie SCADA (Supervisory Control and Data Acquisition)-Systemen, die spezifische Kenntnisse und Fähigkeiten erfordern. OT-Pentests müssen oft vor Ort durchgeführt werden, um die Auswirkungen auf die physischen Prozesse zu berücksichtigen.
Insgesamt erfordert OT-Pentesting eine spezialisiertere Herangehensweise als IT-Pentesting, da selten automatisierte Tools auf spezifischen Hardwarekomponenten und Protokollen eingesetzt werden können.

4.1 Warum Pentesting bei der Bahn?

Mit der fortschreitenden Digitalisierung und Einführung neuer Technologien wie dem DSTW und dem bahnbetrieblichen IP-Netz (bbIP) steigt die Komplexität der Systeme und damit auch das Risiko von Sicherheitslücken.
Diese neuen Technologien bilden die Grundlage für weitere Folgeprojekte, wie beispielsweise Automatic Train Operation (ATO), die eine noch engere Verzahnung von IT- und OT-Systemen erfordern. Pentesting trägt durch intiale und regelmäßige Überprüfung der Systeme und Netzwerke dazu bei, die Sicherheit und Zuverlässigkeit des Bahnverkehrs zu gewährleisten, potenzielle Schwachstellen zu identifizieren und das Vertrauen der Fahrgäste und anderer Stakeholder zu stärken. So können Auswirkungen von potenziellen Angriffen simuliert und geeignete Gegenmaßnahmen rechtzeitig ergriffen werden. Für eine realistische Bewertung ist es von entscheidender Bedeutung, die Systeme mit Angriffsmethoden zu penetrieren, anstatt nur bekannte Listen von Schwachstellen zu prüfen. Es ist wie Schattenboxen und Sparring. Wer nur Schwachstellen-Scans (Schattenboxen) durchführt, findet erst im echten Kampf (Angriff) heraus, wie viel man einstecken kann. Dann ist es zu spät, um zu reagieren. Die Durchführung von Angriffen (Sparring) ist daher die beste Möglichkeit, den Ernstfall zu simulieren, um vorbereitet zu sein.

4.2 Herausforderungen und Besonderheiten – Pentesting im Bahnwesen

Pentesting im Bahnwesen stellt aufgrund der teilweise unvollständigen Dokumentenlage, der heterogenen System- und Komponentenlandschaft sowie dem häufigen Auftreten von proprietären Protokollen eine anspruchsvolle Aufgabe dar, da häufig keine Standardtools eingesetzt werden können. So ist eine individuelle und aufwendige Entwicklung von Testmethoden und -werkzeugen erof business cases that represent quality improvement and cost savings. Moreover, acceptance is improved by the introduction of small pilot projects that demonstrate both the benefits and effectiveness of the test automation. Regular reports, which show the continuous improvement of the test processes, also demonstrate the quality and efficiency of the test automation.

4 IT vs OT pentesting

IT pentesting focuses on the security of information systems, such as servers, networks and applications, and can often be carried out remotely. OT pentesting, on the other hand, deals with industrial control systems such as SCADA systems that require specific knowledge and skills. OT pentesting often needs to be carried out on-site in order to consider the impact on the physical processes.
Overall, OT pentesting requires a more specialised approach than IT pentesting, as automated tools can rarely be used on specific hardware components and protocols.

4.1 Why pentesting on the railway?

As digitalisation progresses and new technologies such as the DSTW and the railway IP network (bbIP) are introduced, the complexity of the systems increases and with it the risk of security gaps.
These new technologies form the basis for further follow-up projects, such as Automatic Train Operation (ATO), which require an even closer integration of the IT and OT systems. Pentesting helps ensure the security and reliability of rail transport, identify any potential vulnerabilities and strengthen the trust of passengers as well as other stakeholders by initial and regularly checking the systems and networks. This allows the effects of any potential attacks to be simulated and suitable countermeasures to be implemented in good time. Attacking the systems by adversarial means instead of just checking known lists of vulnerabilities is crucial for realistic evaluation. It is like shadow boxing and sparring. If you perform vulnerability scans (shadow boxing) only, you will only find out in the real fight (attack) how much you can take. This will be too late to react. So, performing attacks (sparring) is the best way to simulate the emergency case to be prepared.

4.2 The challenges and special features of pentesting in the railway sector

Pentesting in the railway sector is a challenging task due to the often incomplete documentation, the heterogeneous system and the component landscape, as well as the frequent occurrence of proprietary protocols, as it is often not possible to use standard tools. This means that the customised and complex development of test methods or tools is required.
The dependence on manufacturers to obtain detailed information, as well as a lack of or shared test environments, also make implementation more difficult. These challenges require thorough preparation and careful coordination between all the users at a test facility. Furthermore, legacy and proprietary systems as well as protocols are often meant to be secure as the source code is not known to public or widely spread. Practical experience proofs the opposite. That's why, the main attitude in pentesting should be: "Question your assumptions."
In addition to IT driven pentests, it can also be useful to pursue red teaming activities on the organisation. This means, for example, using social engineering as an additional attack vector in

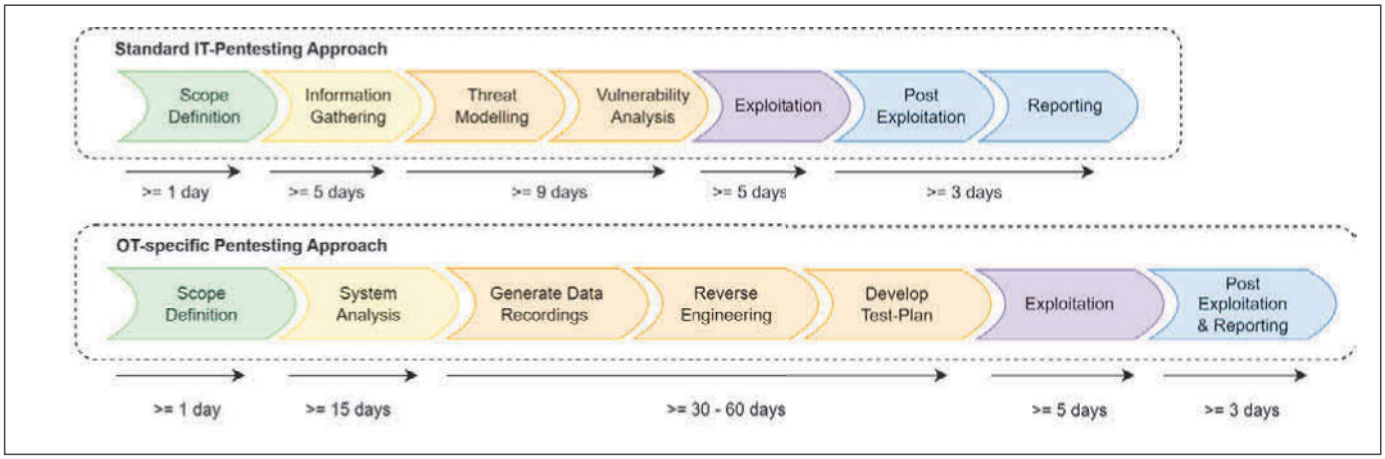

Bild 2: Beispielablauf OT-Pentest
Fig. 2: An example of an OT pentest procedure

Quelle / *Source*: INCYDE

forderlich. Die Abhängigkeit von Herstellern, um Detailinformationen zu erhalten, genauso wie mangelnde oder gemeinsam genutzte Testumgebungen erschweren zusätzlich die Durchführung. Diese Herausforderungen erfordern eine gründliche Vorbereitung und sorgfältige Koordinierung zwischen allen Nutzern einer Testanlage. Außerdem gelten alte und proprietäre Systeme sowie Protokolle oft als sicher, da der Quellcode nicht öffentlich bekannt oder weit verbreitet ist. Die praktische Erfahrung beweist das Gegenteil. Deshalb sollte die wichtigste Einstellung beim Pentesting sein: „Hinterfrage deine Annahmen".

Neben klassischen IT-Pentests kann es auch sinnvoll sein, Red-Teaming-Ansätze innerhalb der Organisation durchzuführen. Dies bedeutet z. B. Social Engineering als zusätzlichen Angriffsvektor in Form von Phishing- oder Tailgating-Attacken einzusetzen. Red Teaming hat dabei das Ziel, die gesamten Sicherheitsmechanismen der Organisation zu adressieren, einschließlich der Fähigkeit, Angriffe zu erkennen und zu verhindern. Ein Beispiel wäre die Prüfung von Zugangs- und Zugriffskontrollen an physischen Standorten. Das nachfolgende Szenario beschreibt einen Pentest, welcher in einer Laborumgebung durchgeführt wurde.

4.3 Ablauf eines Penetrationstests

Der Ablauf eines OT-Pentests wird am Beispiel eines tatsächlich durchgeführten Pentests innerhalb von Bild 2 skizziert. Der Pentest wurde in einer ETCS-Testumgebung (ETCS, European Train Control System) bestehend aus RBC inklusive Anbindung an ein LBS (Leit- und Bediensystem) sowie Elektronisches Stellwerk (ESTW) durchgeführt. Der Beispielablauf erfüllt hierbei nicht den Anspruch, als Blaupause für OT-Pentests zu dienen. Hierzu sind OT-Pentests im Allgemeinen zu abhängig vom vorliegenden Kontext. Nichtsdestotrotz kann der dargestellte Beispielablauf herangezogen werden, um Unterschiede zwischen IT- und OT-Pentests herauszuarbeiten. Bild 2 stellt den Zusammenhang der durchgeführten Schritte zu den allgemein anerkannten IT-Pentesting-Phasen heraus. Hierbei wird vor allem der Unterschied mit Blick auf den zeitlichen Horizont zwischen IT -und OT-Pentesting deutlich.

1. Scope Definition: Einschränkungen und Abgrenzungen müssen klar definiert werden, um den Pentest effektiv und reibungslos durchzuführen. Dies umfasst die Festlegung von Angriffsarten, die einbezogen oder ausgeschlossen werden, wie z. B. den Ausschluss von Denial-of-Service (DoS)-Angriffen, die aufwendige Neustarts erfordern könnten. Besonders bei komplexen OT-Anlagen wie beim ETCS oder DSTW ist es nicht immer möglich, diese schnell wieder in den Ausgangszustand zu versetzen. Zusätzlich müssen Annahmen getroffen werden, um Detailinformationen zu erhalten, genauso wie mangelnde oder gemeinsam genutzte Testumgebungen erschweren zusätzlich die Durchführung. Diese Herausforderungen erfordern eine gründliche Vorbereitung und sorgfältige Koordinierung zwischen allen Nutzern einer Testanlage. Außerdem gelten alte und proprietäre Systeme sowie Protokolle oft als sicher, da der Quellcode nicht öffentlich bekannt oder weit verbreitet ist. Die praktische Erfahrung beweist das Gegenteil. Deshalb sollte die wichtigste Einstellung beim Pentesting sein: „Hinterfrage deine Annahmen".

Neben klassischen IT-Pentests kann es auch sinnvoll sein, Red-Teaming-Ansätze innerhalb der Organisation durchzuführen. Dies bedeutet z. B. Social Engineering als zusätzlichen Angriffsvektor in Form von Phishing- oder Tailgating-Attacken einzusetzen. Red Teaming hat dabei das Ziel, die gesamten Sicherheitsmechanismen der Organisation zu adressieren, einschließlich der Fähigkeit, Angriffe zu erkennen und zu verhindern. Ein Beispiel wäre die Prüfung von Zugangs- und Zugriffskontrollen an physischen Standorten. Das nachfolgende Szenario beschreibt einen Pentest, welcher in einer Laborumgebung durchgeführt wurde.

4.3 The penetration test procedure

The OT pentest process is illustrated in fig. 2 using the example of an actual performed pentest. The pentest was carried out in an European Train Control System (ETCS) test environment consisting of an RBC, including a connection to an LBS (control and operating system) and electronic interlocking (ESTW). The example procedure is unable to serve as a blueprint for OT pentests, as OT pentests are generally too dependent on the context at hand. Nevertheless, the shown example procedure can be used to work out the differences between IT and OT pentests. Fig. 2 shows the relationship between the performed steps and the generally recognised IT pentesting phases. The difference in the time horizon between IT and OT pentesting also becomes particularly clear here.

1. Scope definition: Any restrictions and boundaries must be clearly defined in order for the pentest to be carried out effectively and smoothly. This includes the definition of the included or excluded attack types, such as the exclusion of Denial-of-Service (DoS) attacks that could require costly restarts. It is not always possible to quickly restore systems to their original state, especially in the case of complex OT systems such as ETCS or DSTW. Furthermore, certain assumptions must also be made, such as the availability of given functions, e.g. user accounts or accessibility via networks and systems. Finally, the target system and its relevant components and interfaces must be clearly defined in order to ensure that the pentest is focussed on the essential areas and has no unintended effects on the other systems.
2. The system analysis: Reviewing system documentation, architecture plans and configuration files is an important first step in understanding the target system and its components and identifying any vulnerabilities at an early stage. The existing security measures such as firewalls, Intrusion Detection System (IDS), Intrusion Prevention System (IPS) and encryption technologies are reviewed on this basis in order to assess their effectiveness and identify any vulnerabilities. This forms the

fen werden über die Verfügbarkeit bestimmter Funktionen, wie z. B. Benutzerkonten oder Zugänglichkeit über Netzwerke und Systeme. Schließlich müssen das Zielsystem und seine relevanten Komponenten und Schnittstellen klar definiert werden, um sicherzustellen, dass der Pentest auf die wesentlichen Bereiche fokussiert ist und keine unbeabsichtigten Auswirkungen auf andere Systeme hat.

2. Systemanalyse: Die Überprüfung von Systemdokumentationen, Architekturplänen und Konfigurationsdateien ist ein erster wichtiger Schritt, um das Zielsystem und seine Komponenten zu verstehen und Schwachstellen frühzeitig zu identifizieren. Auf dieser Basis werden die vorhandenen Sicherheitsmaßnahmen wie Firewalls, Intrusion Detection System (IDS), Intrusion Prevention System (IPS) und Verschlüsselungstechnologien überprüft, um ihre Wirksamkeit zu bewerten und Schwachstellen zu identifizieren. Dies bildet das übergeordnete Ziel, mögliche Angriffsvektoren zu bewerten, um den Pentest auf die kritischsten Bereiche zu konzentrieren. Wenn eine Risikoanalyse vorliegt, kann diese dabei helfen; die wahrscheinlichsten Angriffe zu identifizieren.
3. Anfertigung von Datenmitschnitten: Diese sind essenziell, um proprietäre oder unbekannte Protokolle zu analysieren. Durch das Aufzeichnen des Datenverkehrs kann die Funktionsweise des Systems besser verstanden und können Schwachstellen identifiziert werden. Datenmitschnitte helfen auch bei der Erstellung von Test-Skripten für Angriffe wie Man-in-the-Middle, bei denen der Datenverkehr manipuliert wird. Diese Umsetzung konnte im Projektbeispiel am RBC gezeigt werden.
4. Reverse Engineering: Datenpakete werden identifiziert und durch zeitliche Korrelation durchgeführter Aktionen am Zielsystem den Schnittstellen zugeordnet. Hierdurch können Datenpakete gezielt abgefangen und bei Bedarf manipuliert werden. Dabei werden Checksummen und Paritäten untersucht, um valide Pakete zu erzeugen. Diese Methoden dienen primär nicht als IT-Sicherheitsmechanismen, sondern helfen lediglich dabei, die Integrität und Korrektheit der übertragenen Daten aus funktionaler Sicht sicherzustellen und Übertragungsfehler zu verhindern. Im Gegensatz dazu dienen die implementierten IT-Sicherheitsmechanismen wie Verschlüsselung, Authentifizierung und Zugriffskontrolle dazu, das System vor Angriffen zu schützen. Durch die Analyse

overarching goal of evaluating possible attack vectors in order to focus the pentest on the most critical areas. If a risk analysis is available, this can help identify the most likely attacks.

3. Creating data recordings: These are essential for analysing proprietary or unknown protocols. The way the system functions can be better understood and any vulnerabilities can be identified by recording the data traffic. Data recordings also help create test scripts for attacks such as man-in-the-middle, in which data traffic is manipulated. This implementation was demonstrated in the RBC project example.
4. Reverse engineering: Data packets are identified and assigned to the interfaces through the temporal correlation of the actions performed on the target system. This allows data packets to be specifically intercepted and manipulated if necessary. Checksums and parities are analysed in order to generate valid packets. These methods do not primarily serve as IT security mechanisms, but merely help to ensure the integrity and correctness of the transmitted data from a functional perspective and to prevent any transmission errors. In contrast, the implemented IT security mechanisms, such as encryption, authentication and access control, serve to protect the system from attack. Analysing these mechanisms enables the vulnerabilities in the system to be identified and exploited in order to highlight any risks and recommend suitable countermeasures.
5. Developing a test plan: A detailed test plan that includes specific attack scenarios and standard scenarios is created based on the findings. Standard tools are used and specific test scripts that are tailored to the context of the system are developed. Specific test cases, such as in the RBC pentest project, are often written using scripting programming languages such as Python. Standard tools, on the other hand, can also be used to check user authentication, DoS protection or the extension of authorisations on systems that have already been compromised, for example. A heterogeneous (IT and OT-related) test plan is created, which should be agreed with the relevant stakeholders before testing begins in order to ensure a targeted test execution.

Gemeinsamkeiten	Unterschiede
• Den allgemein anerkannten Workflow-Phasen folgen • Standard-Tools kommen zum Einsatz • Übergeordnetes Ziel ist das Identifizieren von Schwachstellen • Ähnliche Fähigkeiten und Denkweisen, die von den Testern verlangt werden	• Workflow-Phasen sind bei OT-Pentests über einen größeren Zeitraum erstreckt • OT-Pentesting weist einen erhöhten Reverse-Engineering-Anteil auf und basiert verstärkt auf der Eigenentwicklung von Scripten / Tools zur Testfalldurchführung • OT-Pentesting weist einen höheren Bezug zu Risikoszenarien auf • OT → Linux-zentriert, IT → Windows-zentriert • IT-Pentesting in der Literatur und Praxis klar definiert, OT-Pentesting befindet sich in der Entwicklung

Tab. 2: IT- vs. OT-Pentesting

Similarities	Differences
• Follow the generally recognised workflow phases • Standard tools are used • The overriding objective is to identify any weak points • Similar skills and mind sets are required of the testers	• The workflow phases are extended over a longer period of time in OT pentests • OT pentesting has an increased proportion of reverse engineering and is increasingly based on the in-house development of scripts / tools for test case execution • OT pentesting is more closely related to risk scenarios • OT→ Linux centred, IT→ Windows centred • IT pentesting is clearly defined in the literature and practice, while OT pentesting is still under development

Tab. 2: IT vs OT pentesting

Quelle / *Source*: INCYDE

dieser Mechanismen können Schwachstellen im System identifiziert und ausgenutzt werden, um Risiken aufzuzeigen und geeignete Gegenmaßnahmen zu empfehlen.

5. Entwicklung eines Testplans: Auf Basis der Erkenntnisse wird ein detaillierter Testplan erstellt, der spezifische Angriffsszenarien und Standardszenarien umfasst. Dabei werden Standard-Tools verwendet sowie spezifische Test-Scripte entwickelt, die auf den Kontext der Anlage zugeschnitten sind. Spezifische Testfälle, wie im Pentest-Projekt zum RBC, werden oft mit Scripting-Programmiersprachen wie z. B. Python geschrieben. Standard-Tools hingegen können verwendet werden, um beispielsweise Nutzer-Authentifizierung, DoS-Schutz oder die Ausweitung von Berechtigungen auf bereits kompromittierten Systemen zu überprüfen. Es entsteht ein heterogener (IT und OT-Bezug) Testplan, der vor dem Testbeginn mit den relevanten Stakeholdern abgestimmt werden sollte, um eine zielgerichtete Testdurchführung sicherzustellen.
6. Testdurchführung: Die Testanlage muss der Produktivanlage physisch und logisch entsprechen, um valide Ergebnisse zu erzielen. Die Projektierung sollte ebenfalls übereinstimmen.
7. Aufarbeitung und Bewertung: Nach Abschluss der Tests werden die Ergebnisse ausgewertet und Schwachstellen dokumentiert. Die Bewertung erfolgt nach definierten Kriterien wie Common Vulnerability Scoring System (CVSS) und der Risikoanalyse. Der Bericht fasst die Tests, identifizierten Schwachstellen und die empfohlenen Maßnahmen zusammen. Er dient als Grundlage für die weitere Planung und Umsetzung von Sicherheitsmaßnahmen und enthält klare Priorisierungen zur Risikominimierung.

Tab. 2 zeigt abschließend etwaige Gemeinsamkeiten und Unterschiede von IT und OT-Pentests auf.

5 Fazit

Pentests und funktionale Tests sind unerlässlich für die Sicherheit und Zuverlässigkeit von OT-Systemen in der Eisenbahnindustrie. Funktionale Tests gewährleisten die korrekte Funktionsweise und Kompatibilität der Systeme, während Red Teaming Pentests potenzielle Sicherheitslücken aufdecken und die Resilienz gegenüber Cyberangriffen durch reale Angriffssimulationen erhöhen. Proprietäre Lösungen sind keine Sicherheitsmaßnahme, wie sich oft gezeigt hat. Das übergreifende Mantra, das wir anwenden sollten lautet daher: „Hinterfrage deine Annahmen". Die Automatisierung von Tests ist dabei entscheidend, um Effizienz und Genauigkeit zu gewährleisten. Expertenteams lösen aufkommende Herausforderungen bei der Implementierung sowie Durchführung dieser Tests durch erprobte Maßnahmen und bieten einen erheblichen Mehrwert für die Sicherheit und Qualität der Bahninfrastruktur. ■

6. Test execution: The test system must correspond physically and logically to the production system in order to achieve valid results. The project planning should also match.
7. Processing and evaluation: Once the tests have been completed, the results are analysed and the weak points are documented. The evaluation is carried out according to defined criteria such as Common Vulnerability Scoring System (CVSS) and the risk analysis. The report summarises the tests, identified vulnerabilities and recommended measures. It serves as the basis for the further planning and implementation of security measures and contains clear prioritisations for risk minimisation.

Finally, tab. 2 shows any similarities and differences between IT and OT pentests.

5 Conclusion

Pentests and functional tests are essential for the safety and reliability of OT systems in the railway industry. Functional tests ensure the correct functioning and compatibility of the systems, while red teaming pentests uncover any potential security vulnerabilities and increase the resilience to cyber-attacks by real attack simulation. Proprietary solutions are no security measure, as often proven. So, "Question your assumption" is the overarching mantra we should apply. The automation of tests is crucial to ensure efficiency and accuracy. Expert teams solve emerging challenges in the implementation and execution of these tests using proven measures and offer significant added value to the safety and quality of the railway infrastructure. ■

AUTOREN | AUTHORS

Dr.-Ing. Martin Koop
Principal IT / OT Security Expert
INCYDE GmbH
Anschrift / *Address*: Rheinstraße 16a, D-64283 Darmstadt
E-Mail: martin.koop@incyde.com

Leon Hagemann
IT / OT Security Expert
INCYDE GmbH
Anschrift / *Address*: Rheinstraße 16a, D-64283 Darmstadt
E-Mail: leon.hagemann@incyde.com

Alain Kaffo
Testmanager Übertragungssystem DKS / *Test Manager Transmission System DKS*
DB InfraGO AG
Anschrift / *Address*: Adam-Riese-Straße 11-13, D-60327 Frankfurt am Main
E-Mail: alain.kaffo@deutschenbahn.com