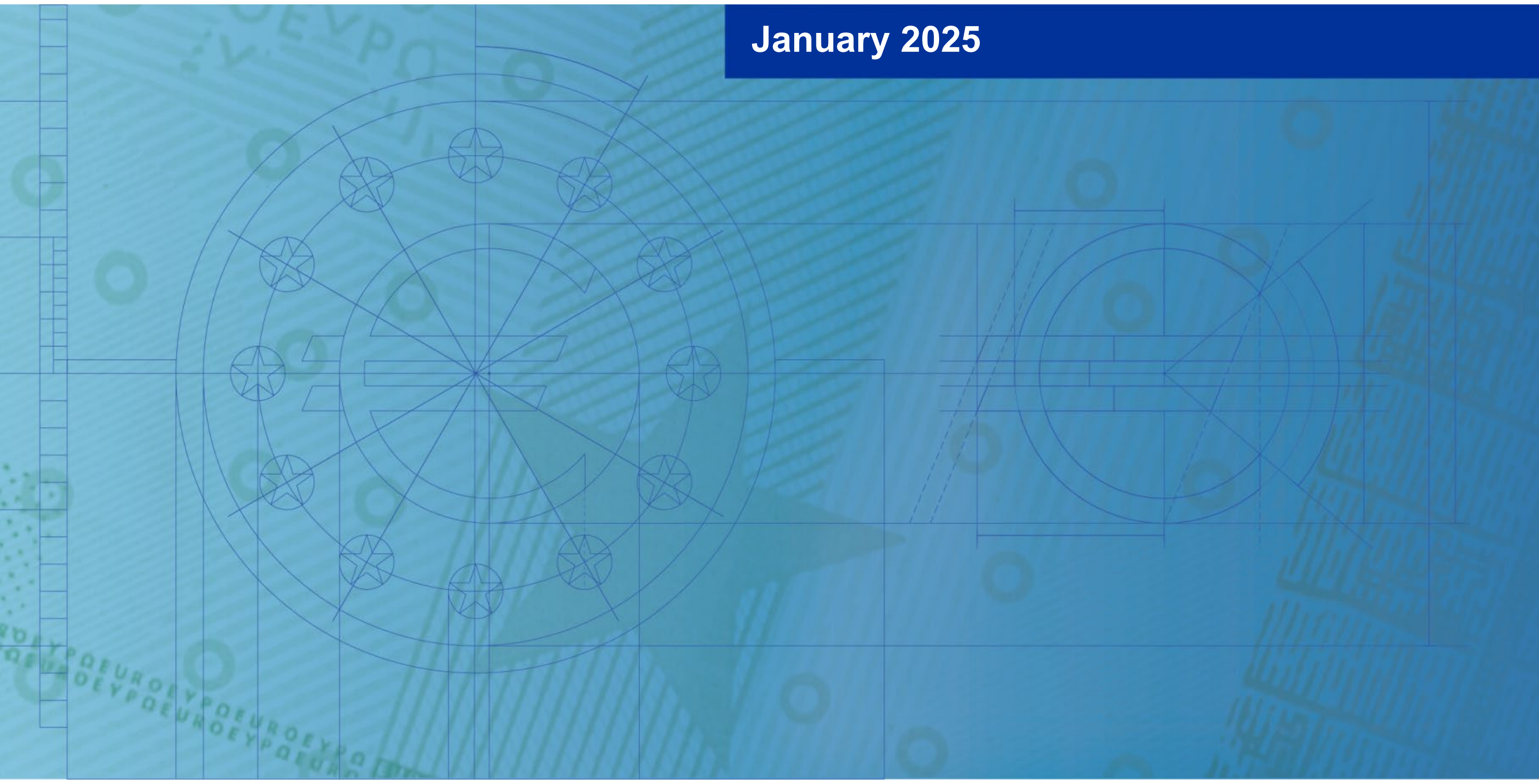
EUROPEAN CENTRAL BANK
EUROSYSTEM

TIBER-EU

Purple Teaming Guidance



January 2025

Contents

1 Introduction

Purple teaming (PT) is a form of a collaborative activity that involves both the red team testers (RTT) and the blue team (BT) in a TIBER-EU test and their corresponding offensive and defensive actions. This can include insights into particular attack phases, detections, defensive actions and test reports. Increased collaboration helps to expand knowledge on the threat actors' tactics, techniques and procedures (TTPs), mitigate certain risks linked to red teaming, and to identify areas and actions that can be improved at people, process and technology level. It also helps to jointly pinpoint weaknesses in protection and detection capabilities so that they can be addressed and incorporated in the Remediation Plan. Such collaborative PT may be undertaken in various ways, ranging from tabletop discussions to specific technical testing activities.

In the scope of a TIBER test, PT refers to two specific and particular points within the testing process only. During active testing in the testing phase, limited PT (LPT) might be used as a way to continue a test where it could otherwise not be continued in a safe or meaningful way. In the closure phase, the PT exercise is a fixed element to theoretically or practically enquire into selected aspects not having been investigated elsewhere within the test process. As such, PT is crucial within any TIBER test to stimulate learning in a most efficient way, but it is of importance to always be clear about which of the two specific variants of PT one refers to.

PT is not intended to replace the confidential active red teaming phase of a TIBER test. Rather, it is intended as a collaborative activity in particular circumstances, to increase the learning experience of the test. During PT in the closure phase, the BT of the entity undergoing the test is fully aware of the ongoing activities and possibly even cooperates with the RTT during their execution.

1.1 Purpose of this document

The purpose of this document is to provide the relevant stakeholders with information on the requirements[1] for conducting LPT during the testing phase, as well as the PT exercise during in the closure phase. It also provides guidance on important aspects to be considered during these efforts.

[1] In addition to the minimum requirements for complying with the TLPT obligations under DORA, this document also includes operational TIBER-EU guidance based on best practices, knowledge and experience from numerous previous tests.

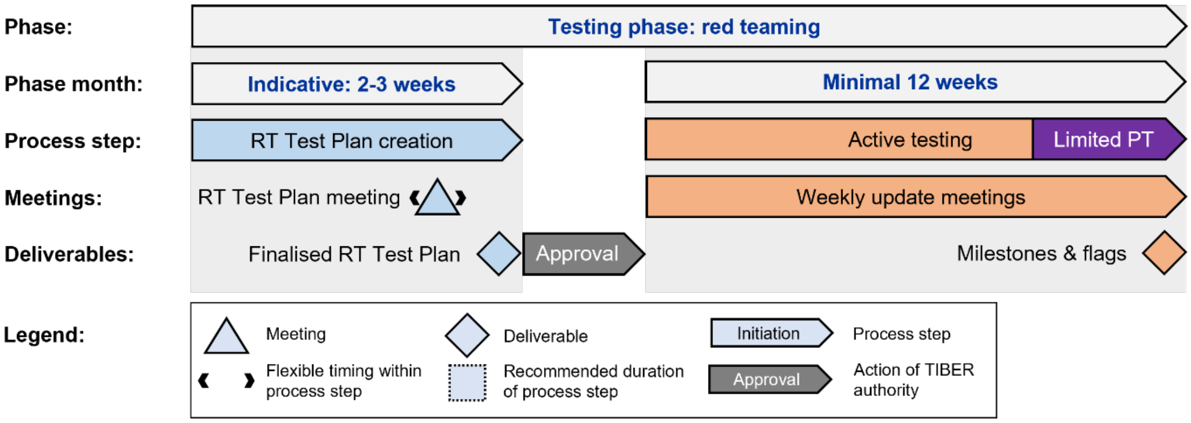
1.2 Target audience

This TIBER-EU Purple Teaming Guidance is aimed at all stakeholders of a TIBER test involved either in LPT during the testing phase or in the PT exercise during the closure phase.

1.3 Location within testing process

LPT in the active testing process step of the testing phase is used as a last resort, once all other options have been exhausted. The use of LPT is subject to a proposal from the control team (CT) and to the agreement of the test manager (TM). It can serve as a response to continue or unblock a TIBER testing phase in a situation where the test would otherwise end prematurely. In such cases, the scope of the PT is limited and very specific. It is conducted to supplement particular parts of the attack scenarios, with the sole aim of ensuring the value of the test and the return on investment in terms of learning opportunities. The reasoning and rationale put forward by the CT to continue the TIBER test using LPT requires careful assessment by the TM to ensure alignment with the TIBER-EU requirements and the spirit of the framework.

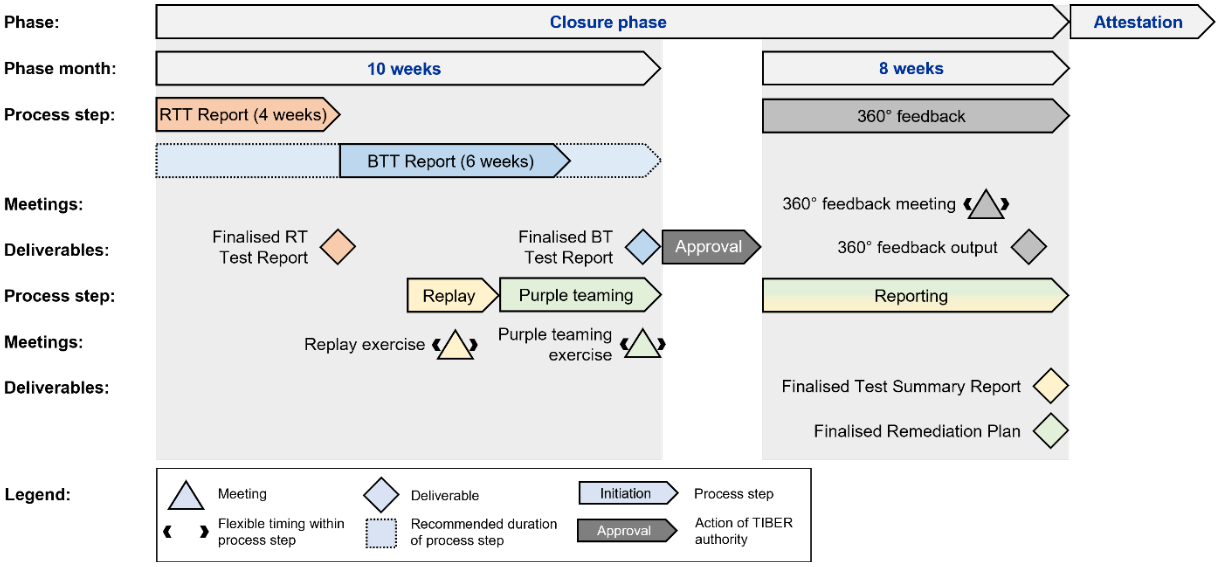
Figure 1[2]
Limited purple teaming in the testing phase



In the PT exercise process step of the closure phase, PT is used to enhance the learning from the testing activities and its results. It shall focus on selected topics jointly identified by the BT and RTT. Such topics might include particular vulnerabilities identified during the test, issues that could not be tested during the active testing or other topics of relevance to the BT.

[2] Note that only the actions of the TIBER authority are included in the figure that have an impact on the timelines of the test. The figure is not an exhaustive overview of all actions to be undertaken by the involved stakeholders.

Figure 2[3]
Purple teaming in the closure phase



[3] Note that only the actions of the TIBER authority are included in the figure that have an impact on the timelines of the test. The figure is not an exhaustive overview of all actions to be undertaken by the involved stakeholders.

2 Requirements for purple teaming

There are different requirements for the two variants of PT. While LPT in the testing phase is to be conducted only under particular conditions (and in agreement with the TM), PT in the closure phase is a standard element of the testing process:

Limited purple teaming in the testing phase might be conducted:

- under exceptional circumstances, to avoid:
  - risks of impact on data;
  - damage to assets;
  - disruption to critical or important functions (CIFs), services or operations of the financial entity itself, its ICT third-party service providers or ICT intragroup services providers;
  - disruptions to its counterparts or to the financial sector;
- after the control team lead (CTL) has suspended the test;
- if the continuation of the test is not otherwise possible;
- after prior approval by the TM;
- to fulfil the twelve-week minimum duration of the active red team testing phase.

Purple teaming exercise in the closure phase must:

- be conducted no later than ten weeks after the end of the active red team testing phase;
- include topics jointly identified by the key stakeholders, which may include:
  - vulnerabilities identified during the test;
  - issues that could not be tested during the active red team testing phase;
  - other steps which could have been taken by the RTT and potential BT responses;
  - alternative scenarios and their potential consequences;
  - proof of concepts;
  - discussion of anticipated remediation measures with the RTT;
  - business continuity exercises.

3 Considerations when conducting purple teaming

3.1 General considerations

3.1.1 Scope and objectives

For PT to be successful, the stakeholders involved must clearly define its scope, goals, objectives, timing and rules of engagement for the actual PT activities. The risk management controls should be reviewed and where necessary adapted, as it is possible that different and more elaborated attack scenarios can be tested under PT.

The CT should also approach PT with an exploratory mindset to delve deeper into the attack scenarios and examine additional techniques and possibly additional attack scenarios. These scenarios may have a forward-looking, outside-the-box perspective that is extreme but plausible, and should resemble attacks which could occur in the (near) future.

3.1.2 Cooperation *among* stakeholders

During PT, the TM should continue to provide advice to the CT to support its management of the test. The objective should be to continue deriving the maximum possible learning value.

For PT to be successful, the BT and RTT are expected to forge a working relationship and maximise their collaboration, in order to create a unique learning experience and enhance each other’s understanding of the test events. It should also take into account barriers between the various stakeholders that may hinder understanding due to different types of knowledge and expertise. Since information must flow between the different teams, language should be adapted so that all stakeholders have a common understanding. To facilitate this cooperation, the RTT should lead by example, clearly and thoroughly explaining their tactics and objectives to the BT, acknowledging the areas of strength, and opening up the conversation to the areas that need improving. This can be done for example by conducting remediation to refine existing controls or implementing new ones.

Close cooperation between the CT, RTT and BT is crucial for the success of any PT. There should be regular checkpoints to confirm the RTT and BT understand each other’s actions. The CT is instrumental in establishing a good basis for cooperation and needs to make the roles and responsibilities within the PT setting explicit.

For communication channels to be efficient and effective and to avoid misunderstandings, the CT should clearly define communication frequency and secure channels in advance, for example by establishing formal (real-time) communication between the RTT and BT via secured channels (e.g. involving end-to-end encrypted email or chat). Effective, efficient and transparent communication among stakeholders is a critical success factor for any TIBER test, and all the more so for PT.

3.1.3 Roles and responsibilities

The stakeholders involved in PT remain the same for both LPT in the testing phase as well as the PT exercise in the closure phase.

3.1.3.1 The test manager

The TM serves as an adviser for all parties during PT. In particular, the TM should ensure that the spirit, principles and processes envisaged in the TIBER-EU framework are maintained and observed. Moreover, the TM has the power to invalidate a test if it has not been conducted in line with both the requirements set out in the TIBER-EU framework and the spirit of the framework, or any applicable regulation(s).

3.1.3.2 The control team

The CT is responsible for making all the necessary decisions as circumstances arise and for ensuring that proper risk management controls are in place for the test to be conducted in an appropriate manner. In addition to making sure that risk management controls remain effective in PT, the CT must also ensure that:

- stakeholders fully comprehend the agreed scope, goals and objectives, especially when switching to (L)PT during the testing or closure phases;
- stakeholders are aware of, and agree on the communication channels to be used, including between the RTT and BT;
- appropriate arrangements are in place to facilitate the shift to PT and provide the clarity required by the RTT and BT to be able to adapt to this new collaborative way of working;
- the RTT and BT adapt their behaviour when engaging in different types of PT, and cultivate cooperation and mutual support.

3.1.3.3 The threat intelligence provider

The threat intelligence provider (TIP) provides expert judgement on the (alternative) scenarios and TTPs to be used in PT. The involvement of the TIP is crucial in both testing and closing phases, as scenarios may need to be adapted. Additional and more advanced scenarios or TTPs may be added, depending on test specificities and planning, resourcing and timing.

3.1.3.4 The red team testers

The RTT expert judgement should also be sought when considering and planning PT. The RTT should work with the TIP to validate the plan and provide a list of new and alternative TTPs to be used during PT.

3.1.3.5 The blue team

The BT is in charge of all the defensive aspects of the scenarios being executed. The BT may also contribute to additional scenarios and/or variations by providing interesting leads and feeding information back to the RTT in the course of PT. Note that right after the active red teaming phase the BT might have difficulty shifting from a defensive attitude to a more cooperative one, particularly if the actions it has to engage in may not be clear. The CT should communicate with the BT regularly to ensure the constructive nature of PT is maintained.

3.2 Considerations for limited purple teaming in the testing phase

3.2.1 Rationale

While the TIBER-EU methodology ensures thorough planning, circumstances may arise during a live test that force the stakeholders to act pragmatically to balance the objective of maximising the learning outcome against maintaining a strict interpretation of the framework.

A considerable amount of time, cost, and effort goes into planning and executing a TIBER test. So, the invalidation of a TIBER test is not desirable, unless the test fails to meet the requirements and spirit of the framework. Hence, it is reasonable that in certain circumstances, it is possible to carry out limited PT during the testing phase to continue the TIBER test and to maximise the return on investment, i.e. the learning experience.

When planning for, or transitioning to, limited PT during the testing phase, it is advisable to re-evaluate the attack scenarios, ensure they fit the PT setting – but remain close to the original. The PT activity may cover one or more of the attack scenarios, depending on the situation. It is however conceivable that PT will be applied to a specific attack scenario, whereas other attack scenarios will continue normally under red teaming.

3.2.2 Circumstances leading to limited purple teaming

Alternative ways of progressing have to be thoroughly examined by the CT before proposing to move into PT. For example, pausing the test should be considered to see if this would be an equally suitable measure to maximise the lessons learnt. Together with the TMs, the stakeholders evaluate each case individually and assess whether limited PT is an option.

Examples of potential circumstances that may lead to PT during the testing phase are described below.

- When the BT has detected the RTT in such a way that the secrecy of the test is irreparably compromised. Note that it is possible that during a test, the BT may detect some RTT actions; however, this alone does not necessarily mean that PT is the right way forward, and it is possible to still continue the test in its original manner using a cover story (e.g. a local penetration test) to explain certain detections to the BT or to only introduce PT for the detected attack scenarios. In addition, it may be possible that a test is partially detected by the BT and the CT can then instruct a freeze on RTT activities to allow the elevated threat level to subside. In such cases, it is crucial to have alternative approaches and techniques at hand, as it is a common mistake to pause only to reuse the same attack vectors that have already been detected.

- PT can also be triggered by difficult to foresee situations, where there is a high degree of risk that the emulated attack could lead to business impact or disruption to CIFs. In these cases, it is advisable to discontinue testing and to introduce PT for these systems instead. This would enable the BT, once informed, to take timely action to prevent and minimise any impact, or advise on a safe alternative TTPs.

- In the case of a real cyberattack during the same period of the TIBER test, where the BT has to fully shift its focus to disruption prevention and containment. This may result in the TIBER test being revealed to ensure the BT can differentiate TIBER activities from the malicious attack. Among other possibilities, the test may be postponed to a later date, possibly utilising PT.

- When there is a high probability or clear signs that the response of the uninformed BT to contain the detected emulated attack will have a disruptive impact on CIFs. This potential overreaction might be appropriate in the event of a real attack, but not in the context of a TIBER test, given that the RTT will never deliberately cause real harm. If the BT is not aware of the TIBER test, they have no way of knowing if their response is adequate.

- To prevent situations that can lead to the BT straying from normal response procedures when suspecting the detected attack in a test. This would both reduce the realism of the test and hamper its learning outcome.

- When the CT is unable to stop escalation by the BT, and the BT has involved external parties such as the police, intelligence services, government authorities, industry bodies or financial institutions, for example, due to the perceived severity of the incident. Involving these parties will put an unnecessary strain on those authorities and could have a severe impact on current and future testing activities. The test should be halted immediately, (partially) revealed, and PT may be considered.

3.2.3 Considerations when moving into limited purple teaming

It is not possible to provide an exhaustive list of circumstances that could result in shifting to PT during the testing phase. However, one of the main criteria should be that the testing phase cannot continue in a secret and/or secure manner due to an event outside of the control of the CT, RTT or TM.

Entering into LPT during the testing phase should involve the following steps:

- the CT formally proposing PT, detailing specific scope and objectives;
- the TM agreeing to move to LPT;
- the test otherwise still being conducted in accordance with the spirit of the framework (i.e. PT should be considered an option of last resort rather than a relaxation of the TIBER-EU requirements), focusing on maximising the learning experience and outcome;
- the CT liaising with the TIP and RTT as necessary to adapt existing scenarios or implement alternative scenarios so as to maximise the value of the test for the tested entity;
- agreeing in advance on expectations regarding the outcome, communication channels, response and recovery activities, confidentiality boundaries, start and end, escalation paths, allocated resources (including budget) and reporting formats;
- agreeing that the outcomes of PT be clearly documented and form an integral part of the Remediation Plan.

3.3 Considerations for purple teaming in the closure phase

3.3.1 Rationale

The PT element in the closure phase helps to optimise RTT and BT collaboration and maximises learning opportunities, defence capabilities, situational awareness and ultimately the return on investment of the whole test.

A well-executed PT exercise can provide the entity with a comprehensive review of the effectiveness at each layer of its infrastructure in scope. Moreover, it aims to improve the detection controls that are crucial to shed light on suspicious activity.

3.3.2 Planning

PT in the closure phase should be scheduled to take place close to the delivery of the final RTT and BT Test Reports, and after the replay exercise. This timeline ensures that PT is carried out while the details and observations noted during the testing phase are still fresh in the minds of the BT and RTT.

3.3.3 Results

The PT exercise in the closure phase allows for more detailed examination and evaluation of particular aspects of a TIBER-EU test, without the constraints present in the testing phase, such as BT detection or limited amount of RTT and BT collaboration. It makes it possible to directly leverage the expert knowledge of the RTT and BT to revisit and address specific areas deemed important by the tested entity.

PT can therefore result in a deeper understanding of the interconnections and implications of the most relevant offensive and/or defensive measures for the tested entity. It might help to demonstrate and highlight the potential consequences from both a technical and a business perspective (e.g. remediation, recovery time/point-, business continuity, etc.) and hence inform considerations beyond the technical realm. As a result, a PT exercise might facilitate a better understanding of the consequences of an attack, further proliferation of an attack and alternative ways to enhance protection and detection.

The results of PT will greatly benefit the further refinement of recommendations and remediation planning, which will in turn enhance the cyber resilience of the tested entity. In addition, they might feed into other operational resilience exercises and improve the entity’s operational risk and information security/cyber resilience programme or framework. One such example is to utilise the scenarios in crisis simulation and coordination exercises.

3.4 Types of purple teaming

PT can vary in its purpose, learning experience, form, level of BT involvement and specificities. This section describes some possible types of PT that can be used alone or in combination. However, since each TIBER test is different, they might be expanded or adapted according to the specific circumstances of a given test. The best approach when deciding which type of PT to select is for the relevant stakeholders to openly discuss the different alternatives.

3.4.1 Types of limited purple teaming in the testing phase

3.4.1.1 Catch-and-release

"Catch-and-release" is a useful way of testing an entity's defensive capabilities when there have been repeated detections by the BT during the final stage of a test. Because of such detection and consequent blocking of the accounts and tools involved, further progress in testing activities might not be possible without a significant change in TTPs. While different TTPs, such as choosing alternative routes into or through a network or different attack techniques, could be employed in the early stages of a test, this might not be feasible during the final stage. For example, if production systems have been successfully breached but exfiltration or modification of data is still outstanding, or when a route to the critical functions cannot be identified by the RTT, a sudden change in TTPs might be counterproductive. In such situations, a catch-and-release approach might be carried out, enabling lessons to be learnt about very specific aspects of an attack that could not otherwise be achieved.

Catch-and-release is initiated by revealing to or detection by the BT that a test is being performed on the systems and installing a dedicated communication channel between the RTT, BT and CT. In the event of any further detection within well-defined boundaries (e.g. certain machines or subnets), the BT uses this channel to report the detected Indicators of Compromise (IoC) to the RTT, which confirms or refutes those as being part of their test. Note that special care should be taken by the CT to ensure that the BT blocks out confidential information that might not be related to the TIBER test. Should the identified IoCs indeed be part of the TIBER test, the BT will then perform the agreed measures to allow the test to continue (e.g. releasing an isolated machine or account). Alternatively, other previously agreed actions might also be taken, such as shutting down a machine, escalating the incident or starting a forensic investigation with the aim of evaluating these processes as part of the test.

Importantly, specific guardrails should be defined in advance, specifying which machines (or subnets) are in scope of such types of PT as well as which responses actions cannot or should not be skipped.

Although the BT is aware of a test being performed on its production systems, this does not necessarily mean that the BT is aware that it is a TIBER test or what scenarios are covered by the test. During the complete PT phase, regular updates should take place to ensure adequate risk control and facilitate mitigation measures.

3.4.1.2 Collaborative proof-of-concept

A collaborative "proof-of-concept" can be a very helpful activity to provide evidence of a weakness discovered during TIBER tests in situations where practical testing on the production systems by the RTT alone is not feasible (e.g. because of being out of scope, unjustified high risk of impacting critical systems, etc.). In some settings, a proof-of-concept might require the explicit involvement of the BT to provide a particular part of infrastructure expertise or risk control, for example. A collaborative proof-of-concept offers a very detailed learning experience related to a particular attack step or vulnerability.

Executing a collaborative proof-of-concept includes collecting all the evidence required to illustrate the feasibility of a certain attack. This might include a theoretical discussion of the expected outcome of executing a given step as well as the protective measures in place. A practical test of partial aspects of the attack is usually also carried out (e.g. sub-steps, using dummy data, execution on testing systems, etc.).

Proofs-of-concepts might equally be conducted during the closure phase to avoid undesired effects during the test itself, such as putting the BT on high alert. However, at a very late stage of testing, the remaining activity might depend on a particular proof-of-concept. BT staff can thus be asked to contribute when drawing up the collaborative proof-of-concept.

Other stakeholders, such as the CT and the TMs, should be closely involved to ensure the proof does indeed provide evidence of a weakness.

3.4.1.3 War game

A war game is an activity in which the RTT and BT are fully aware of each other's respective goal to capture the respective flags or to protect the entity's critical assets and terminate the attackers' access to the network. As such, war games differ from the spirit of a normal TIBER test in the sense that the BT knows what the RTT flags are. This is the main difference between catch and release and war gaming. If a war game is performed, flags are usually placed in systems underpinning CIFs included in the test scope.

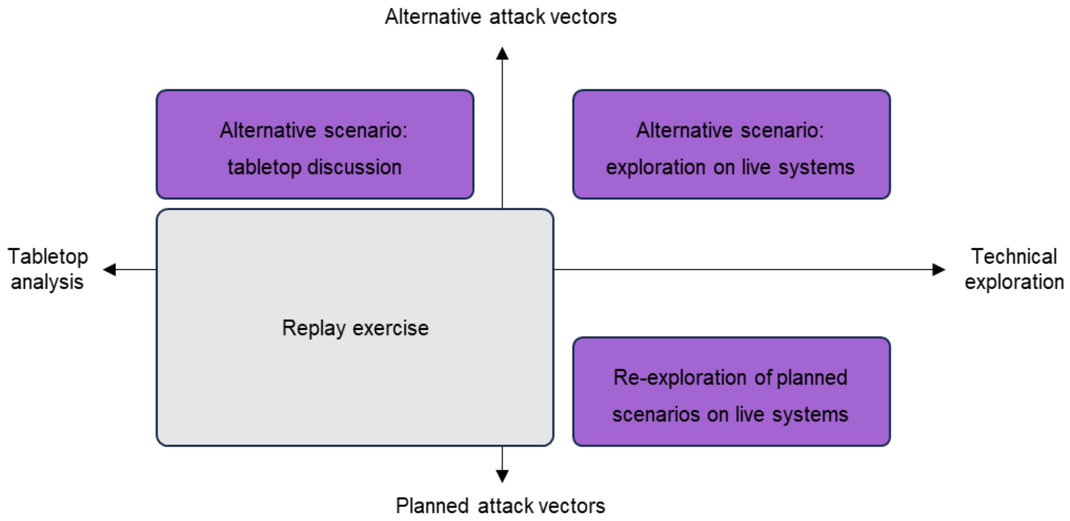
In very particular situations, such as when a TIBER test is known to the BT at a very early stage of the RTT, a war game might be a suitable option to continue the test and still enable a learning experience.

3.4.2 Types of purple teaming exercises in the closure phase

The PT exercise in the closure phase serves the purpose of investigating additional aspects in regard to a TIBER test, which have not been studied during the active testing or the replay exercise. As such, they can greatly enhance the learning of a test.

The PT exercise in the closure phase can take different forms and might focus on attack scenarios planned and executed during the testing phase as well as on alternative and more explorative approaches. Furthermore, it can take the form of tabletop discussions as well as activities on the technical systems (see figure 3).

Figure 1
Types of PT exercise modi in the closure phase


Alternative attack vectors
Alternative scenario: tabletop discussion
Alternative scenario: exploration on live systems
Tabletop analysis
Replay exercise
Technical exploration
Re-exploration of planned scenarios on live systems
Planned attack vectors

Deciding on which types of PT are best suited for a particular test depends on many factors, such as the results of the test, risk considerations, the intended type of learning experience and available resources. The types of PT exercises described below can serve as a reference to ensure the consistent use of terminology by all parties involved. These are likely to be combined and blended in a way that best fits the situation at hand and the specificities of the entity being tested. Additional stakeholders, such as the internal RTT of the entity tested, might also be included in a passive or active role. This can lead to additional learnings for all parties.

3.4.2.1 Alternative scenario: tabletop discussion

A tabletop discussion can provide valuable learning experiences when technical systems are not required or available. This type of activity allows a less technical audience, including management, to be included more widely in the discussion. While the planned attack vectors have already been analysed in the replay workshop, a tabletop discussion is a great way to investigate alternative attack vectors and discuss or simulate the "what ifs" without a strict focus on technical systems. For example, a simulation might be used to discuss the entity's response in case the method or pathway used by the attacker to infiltrate the system was successful. Such tabletop discussions offer high flexibility with regard to the tools used.

A tabletop discussion can be carried out in a variety of ways. These might include:

- a role-play to discuss and simulate alternative offensive and defensive measures and their consequences;
- the theoretical evaluation of scenarios that are closely related but out of scope of a TIBER test;

- the simulation of potential consequences reaching far beyond the test (e.g. restoring business continuity after a successful ransomware attack);
- the inclusion of senior management.

Tabletop discussions might include many different stakeholders (such as business process owners) and therefore require thorough planning and moderation. However, they facilitate an all-round view of the wider aspects of an entity's security and can even go far beyond the initial scope of the test.

3.4.2.2 Re-exploration of planned scenarios on live systems

A technical re-exploration of the attack vectors planned and executed during a TIBER test is an effective way to combine the expertise of the RTT and the BT to practically show the offensive and defensive potential of an attack step or attack chain. Although this type of activity is quite resource-intensive in terms of preparation, it can deliver a highly comprehensive and detailed hands-on learning experience for the BT. For example, it might be very helpful in cases where the BT struggled to detect RTT activities during a test. In this case, the RTT and BT might engage in a collaborative attack recollection. Running through a chosen attack sequence step by step, the RTT execute the respective TTPs while the BT can simultaneously provide corresponding defensive information (e.g. event notifications received, alerts, blocked executions, etc.). The two teams can together discuss (and document) the consequences of such an attack sequence and draw up preventive and/or reactive measures from their respective points of view.

This type of activity might include:

- walking through an RTT activity that was not visible in the BT logs during the testing phase;
- walking through an RTT activity that was visible in log entries, but the malicious activity was not detected by the BT during the testing phase;
- walking through an RTT activity that triggered an alert during testing but was not triaged properly;
- walking through activities to which defensive responses were ineffective during testing;
- walking through activities that triggered a defensive response that effectively closed the attack vector but was unable to prevent the attackers from meeting their objectives.

Alternative defensive measures and potential offensive countermeasures might be discussed and practically evaluated to achieve a good understanding of the different possibilities of attack and defence.

3.4.2.3 Alternative scenario: exploration on live systems

Technical explorations of relevant attack scenarios during the closure phase might not be limited to merely evaluating scenarios conducted during the test phase. Variations of tested attack scenarios as well as novel or more elaborate scenarios can be thought of during the active testing phase. These can often not be comprehensively evaluated or executed during the testing phase due to time limit and other constraints. For example, due to late discovery or a high risk of impact or BT detection.

Although it could be done in a theoretical manner, more detailed insights are often gained when testing is carried out on the actual systems. Since this is done in a collaborative manner, even aspects entailing a high degree of risk can be investigated to obtain very realistic results with a high level of technical detail. Alternative scenario explorations on live systems might contain:

- a technical exploration of attack scenarios deviating from those conducted during the testing phase;
- a technical exploration of tested attack scenarios applied to alternative target environments (e.g. execution in a Citrix environment instead of on a company laptop);
- a technical exploration of alternative attack scenarios making it possible to take an in-depth look at the consequences, potential detection and response measures for novel kinds of attacks.

Postal address 60640 Frankfurt am Main, Germany
Telephone +49 69 1344 0
Website www.ecb.europa.eu

For specific terminology please refer to the ECB glossary (available in English only).

PDF ISBN 987-xx-xxx-xxxx-x, ISSN xxxx-xxxx, doi:xx.xxxx/xxxxxx, QB-xx-xx-xxx-EN-N
HTML ISBN 987-xx-xxx-xxxx-x, ISSN xxxx-xxxx, doi:xx.xxxx/xxxxxx, QB-xx-xx-xxx-EN-Q