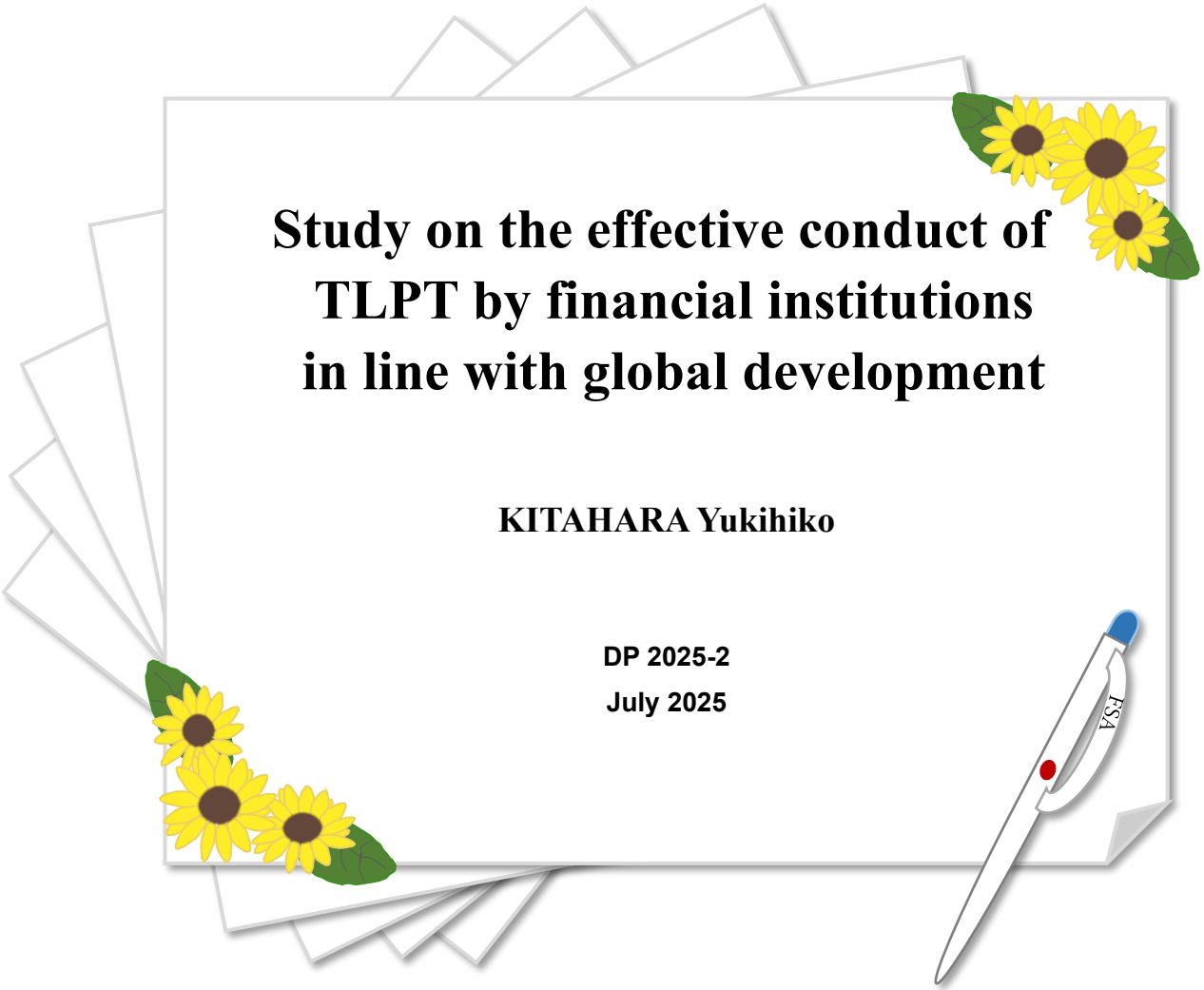
Financial Services Agency

Discussion Paper Series



Study on the effective conduct of TLPT by financial institutions in line with global development

KITAHARA Yukihiko

DP 2025-2

July 2025

Financial Research Center (FSA Institute)
Financial Services Agency
Government of Japan
3-2-1 Kasumigaseki, Chiyoda-ku, Tokyo 100-8967, Japan

Study on the effective conduct of TLPT by financial institutions in line with global development

KITAHARA Yukihiko*

Abstract

The threat of cyber attacks on financial institutions is increasing, and the methods of attack are evolving constantly, becoming more sophisticated and shrewd. In order to respond to such increasingly sophisticated cyber attacks, it is important for financial institutions to identify their cybersecurity risk continuously, and to strengthen their response to cyber attacks by taking measures not merely through system improvement, but from the perspective of strengthening their overall response system, including the corresponding behavior and the development of its response processes against cyber attacks.

In the situation of increasing cybersecurity risk, Threat-led Penetration Testing (TLPT) is considered useful as a tool for assessing how effectively financial institutions' response systems work against potential cyber attacks, for improving their response systems, and for drawing lessons for further improvement. In Japan, the use of TLPT has been increasing, particularly by major financial institutions. In other countries, initiatives related to TLPT are more advanced, especially in Europe and the United States, and relevant legal systems and frameworks for its implementation have also been developing.

This report provides an overview of developments related to initiatives for TLPT taken by financial institutions in other countries, and frameworks for the implementation of TLPT, including rules and regulations and other legal arrangements. Based on a review of developments in other countries, the report discusses how Japanese financial institutions should conduct TLPT and identifies some of the points that need consideration in conducting TLPT.

Keywords: cyber attacks, cybersecurity, TLPT, Threat-led Penetration Testing.

* Research Fellow, Financial Research Center, Financial Services Agency (FSA).
I am very grateful for the assistance from experts within and outside the FSA, including MIURA Shun (Director for IT Risk, Cybersecurity and Economic Security, Risk Analysis Division, Strategy Development and Management Bureau), and their helpful comments in preparing this paper. The views expressed here are solely those of the author, and do not reflect the views of the FSA or the Financial Research Center of the FSA.

1. Introduction

Cyber threats are increasingly growing, with large-scale cyber attacks that states are suspected of being involved occurring constantly by means of sophisticated and persistent tactics. At the same time, as financial information systems in Japan and overseas have become increasingly connected to external networks as a result of globalization, the threat of cyberattacks is likely to increase further.

To address cyber threats, financial institutions need to identify and analyze the threat of cyber attacks, take necessary defensive measures, and verify their effectiveness. However, in reality, many financial institutions have not taken sufficient measures against the ever-increasing cybersecurity risk, and so the potential risks are increasing. Under these circumstances, as an initiative to implement more active cyber defense measures, [the implementation of] Threat-Led Penetration Testing (TLPT) has become an effective means to strengthen cyber resilience of financial institutions. TLPT is a methodology to verify the capabilities of financial institutions to respond to cyber attacks by analyzing real-life cyber threats (threat intelligence) and conducting tests based on the analysis using tactics of real-life cyber threat actors.

Whilst undertaking research as a research fellow at Japan's Financial Services Agency (FSA) since 2022, the author has been engaged in monitoring of financial institutions' cybersecurity as a financial securities inspector. Through monitoring activities, the author had direct dialogue with financial institutions implementing TLPT and identified how they developed their TLPT initiatives and their approaches to TLPT. In addition, the author has taken part in FSA's policy measures to improve the effectiveness of TLPT, under the policy aimed at strengthening cyber resilience of financial institutions. As such measures, the FSA has launched a feedback project in business year 2023 providing information on cases of TLPT conducted at financial institutions, and a pilot project for implementation of TLPT at regional financial institutions in business year 2024. Through these policy measures, the author acknowledged that there is room for improvement in the TLPT performed by Japanese financial institutions.

Turning to developments abroad, frameworks for implementing TLPT have been devised in Europe and Asian countries, such as the UK's Critical National Infrastructure Banking Supervision and Evaluation Testing (CBEST) and the European Red Team Testing Framework (TIBER-EU) in Europe. In January 2025, the Digital Operational Resilience Act (DORA) was enacted in Europe, which has, in effect, made TLPT mandatory for all financial institutions subject to the Act. It is evident from these international developments that financial institutions' initiative to implement effective TLPT are more advanced in other countries compared to Japan, and we believe that developments abroad will provide useful lessons for improving the effectiveness of TLPT at Japanese financial institutions.

In this report, Section 2 provides an overview of TLPT and describes initiatives taken in Japan, and Section 3 presents trends in other countries such as the development of regulations and guidelines related to TLPT. Then in Section 4, we analyze what is required in foreign TLPT regulations and guidelines, and examine points that Japanese financial institutions should consider in the conduct of TLPT based on developments related to TLPT abroad.

2. TLPT Initiatives in Japan's Financial Sector

2.1 TLPT as Cybersecurity Resilience Test

2.1.1 Testing and assessment methodology for cybersecurity

In a situation where an organization is exposed to cyber threats, it is utmost important to test and verify the effectiveness of cybersecurity measures, identify potential vulnerabilities in the organization, and take necessary measures to prevent system compromise by cyber threat actors. It is also extremely important to evaluate and improve cyber resilience, including response processes to incidents.

In general, testing of cybersecurity is divided into tabletop testing[1] and technical testing. Of these, technical testing are further classified into three categories: "vulnerability assessment," "penetration testing," and "TLPT."

[1] Examples include risk assessment using checklists and interviews, document review, and tabletop exercises.

Vulnerability assessment is a method for identifying risks and vulnerabilities in computer systems, networks, application software and other system devices by making simulated attacks via destructive online communication or requests. Penetration testing is a methodology to assess system resilience by simulating cyber threat attempts from the viewpoint of threat actors and verifying whether systems can be compromised. TLPT is a testing method that mimics tactics used by real-life threat actors to verify system resilience and identify what vulnerabilities and challenges exist regarding people, organization and processes, in addition to system vulnerabilities. TLPT is typically the most complex and costly of the three testing approaches.

2.1.2 Definition of TLPT

The G-7 Cyber Expert Group published in 2018 the "G-7 Fundamental Elements for Threat-Led Penetration Testing," which defines TLPT as follows. TLPT is "a controlled attempt to compromise the cyber resilience of an entity by simulating the tactics, techniques and procedures of real-life threat actors. It is based on targeted threat intelligence and focuses on an entity's people, processes and technology, with minimal foreknowledge and impact on operations."

With cyber threats becoming more and more sophisticated, TLPT is an advanced security testing method aimed at strengthening cyber resilience by evaluating an entity's resilience capabilities against real-life cyber threats. Given that TLPT verifies resilience based on scenarios assuming real-life cyber attacks, the test needs to be conducted on the live production systems. TLPT is performed in teams: "Red Team" is the group of testers that execute a simulated attack; "Blue Team" is the group responsible for defending, detecting, and responding to the simulated attack; and "White Team"[2] is the group responsible for managing and coordinating the entire process of the test. Red Team executes a simulated attack on target systems and networks in an attempt to achieve its attack objectives, and Blue Team detects and responds to the attack to verify the effectiveness of its cyber defense capabilities. It would therefore be ideal to conduct the test without the foreknowledge of the Blue Team in general.

While the vulnerability assessment and penetration testing are conducted for the purpose of detecting system-related vulnerabilities, TLPT aims to assess cyber resilience not only in terms of systems, but also of "people," "organization," and "response processes." In light of this, in conducting TLPT, the test is planned and executed focusing on real-life cyber threats, and set the scope and goals of the test based on the expertise of the Red Team, so as to verify resilience against likely threat scenarios.

2.2 Initiatives for Implementation of TLPT in Japan

2.2.1 Authorities' efforts

The FSA has been carrying out a policy to promote the implementation of TLPT to financial institutions, as TLPT is a sophisticated and effective assessment methodology for strengthening their cybersecurity response capabilities.

According to the second edition of the "Policy Approaches to Strengthen Cyber Security in the Financial Sector" the FSA published in 2018, it encouraged financial institutions, particularly large institutions, to implement TLPT. The report states "The FSA will also further boost responsiveness by promoting the utilization of high-level assessment methodologies such as threat-led penetration tests in light of the best practices of G-SIFIs and the developments of international discussions." The third edition of the "Policy Approaches" released in 2022 indicated that "improving the effectiveness of TLPT" and "developing response processes against large-scale cyber incidents" are examples of measures for strengthening cyber resilience. In addition to these, to better understand developments in other countries and regions, the FSA published "report on the conduct of TLPT in major countries and economies"[3] in 2018, which explains how TLPT is conducted and how financial institutions in other countries are utilizing TLPT, thereby encouraging financial institutions in Japan to adopt TLPT.

Regarding policy measures the FSA has taken in recent years, in business year 2023, it undertook an initiative to collect and analyze TLPT cases from banks that have conducted TLPT and

[2] The EU's Digital Operational Resilience Act (DORA) and TIBER-EU, which will be discussed later, define it as a "control team."

[3] https://www.fsa.go.jp/common/about/research/20180516.html (available in Japanese)

give feedback to financial institutions regarding core challenges and good practices related to TLPT.[4] The FSA is promoting further enhancement of cybersecurity management as financial institutions make use of these cases to improve the effectiveness of TLPT. In business year 2024, the FSA launched a similar initiative, collecting TLPT cases for the insurance industry.

In addition, in business year 2024, the FSA guided implementation of TLPT on multiple regional financial institutions in a pilot project aimed for implementation of TLPT to regional financial institutions, demonstrating the usefulness of TLPT to them, and also lowering barriers to TLPT implementation by identifying threat intelligence that are common to regional financial institutions, which is essential in conducting TLPT. The FSA gave feedback to regional financial institutions on common vulnerabilities identified through TLPT, thereby strengthening cybersecurity of regional financial institutions overall.

2.2.2 Japan's regulations and guidelines

As of 2025 when the author is writing this discussion paper, there are no regulations in Japan mandating the implementation of TLPT. As for guidelines and guidance, there is "The Guide for Financial Institutions to Implementing TLPT" (hereafter "FISC Guide") published by the Center for Financial Industry Information Systems (FISC) in 2019. The FISC Guide was prepared so that financial institutions conducting TLPT can refer to, providing process-by-process considerations for its smooth implementation. While it was prepared based on TLPT guidelines abroad, "it was devised tailored to the current situation in Japan, referring to the results of interviews with domestic financial institutions on the implementation of TLPT." The FISC Guide would be useful to financial institutions in Japan in understanding the TLPT implementation process. However, given that it has not been updated since 2019, it may be inadequate compared to the level required by foreign TLPT frameworks such as CBEST and TIBER-EU that have been updated drawing on years of experience. Specifically, in the FISC Guide, test environment and the scope of information provided to related parties prior to TLPT is left to the financial institution to decide, and so the guide is not clear as to the extent it is encouraged to perform TLPT in a live production environment, making it difficult for financial institutions to decide on details of TLPT.

The FSA published the "Guidelines for Cybersecurity in the Financial Sector" in October 2024, which clarifies issues related to governance, [risk] identification, defense, detection, response, recovery, and third-party risk management from a cybersecurity perspective, and clarifies "fundamental measures" and "desirable measures" for financial institutions in each topic. With regard to TLPT, the Guidelines sets out "measures desirable to be taken" as follows:

2.2.4. Vulnerability assessment and penetration testing
[measures desirable to be taken]
b. Conduct TLPT regularly. When conducting TLPT, take note of the following points:

- Select service providers that have necessary experience and skills (this includes checking certifications and experience of the testers);
- Conduct tests using the equivalent level of techniques that real-life cyber threat actors command, based on threat intelligence.
- Consider severe but plausible threat scenarios, which could impact the provision of services to relevant parties, for the test plan.
- Assess incident response capabilities (defense, detection, reporting, containment, etc.) of the defense team (Blue Team) in the test.
- Conduct tests in the live production environment without the foreknowledge of the defense team (Blue Team).
- Report challenges and vulnerabilities identified to the [management/board] and take remediation action.

Source: Financial Services Agency, "Guideline for Cybersecurity in the Financial Sector."

[4] For an overview, see the column in "analysis report on financial institutions' IT system failures" (available in Japanese) published by the FSA in June 2024. (https://www.fsa.go.jp/news/r5/sonota/20240626/20240626.html)

This guideline does not require one-size-fits-all approach to financial institutions, and so each institution should take its own "risk-based approach." Given that provisions related to TLPT have many common aspects with foreign guidelines and frameworks explained in a later section, financial institutions are required to take substantial and effective action in light of the objective of the relevant rules, supervisory policy and the guideline.

2.2.3 Challenges for financial institutions in Japan in implementing TLPT

As we have seen, Japanese authorities are promoting TLPT, and guidelines on TLPT that financial institutions can refer to are being devised, but financial institutions performing TLPT are still limited to a small number of institutions, mainly large ones. Further, among financial institutions that have implemented TLPT, the quality of the test is often not adequate. According to cases informed for the feedback initiative of TLPT in business year 2023, there were following cases: (i) threat intelligence is limited to the analysis of general threat information; (ii) the TLPT plan is communicated to the Blue Team in advance, and so the Blue Team knows that a simulated attack will be made; and (iii) an adequate budget that align with the purpose and objectives of TLPT has not be secured, and threat intelligence and Blue Team assessment been omitted, resulting in tests that are inadequate for TLPT.

	Desirable cases	Inadequate cases
Threat intelligence	- Identify threat intelligence specific to the institution and decides on scenarios accordingly.	- Threat intelligence is limited to an analysis of general threat information.
Assessment	- Conduct TLPT without foreknowledge of Blue Team, and assess its capabilities to detect and respond to cyber threats. - Assess Computer Security Incident Response Team's capabilities, such as escalation to security incident from phishing emails, as well as Security Operation Center's detection and response capabilities.	- Since the TLPT plan has been informed to the Blue Team in advance and the Blue Team knows that a simulated attack will occur, there is a risk that its detection and response capabilities is not assessed properly. - Since there are no attempts to enter or bypass attacks from multiple routes that could be considered by an attacker, it is limited to a vulnerability assessment that is not substantially different from the vulnerability assessment that verifies the vulnerability to be verified with the attack method and route assumed in advance.
Reporting to the board Board's response	- The section responsible for cybersecurity reports to the management risks identified from the TLPT results that could have a company-wide impact. - In addition to receiving reports on the results of the TLPT and giving instructions on how to address the issues, the management has also given instructions to enhance the TLPT by expanding the scope of testing and conducting tests without prior notice.	- Similar issues that could have a serious impact on business operations and customers have been repeatedly detected, countermeasures have not been sufficient. It is likely that significant risks remain, but the department in charge of cybersecurity has omitted reporting to the management based on similarities to past findings. - Among the issues detected by the TLPT, those that could have a material impact on the operation of the financial institution were not reported to the management, nor were specific risks reported. Instead, the report simply stated that "the results were favorable."
Utilizing identified points	- The section in charge of cybersecurity instructs the persons in charge of other systems and those in charge of systems of affiliated companies to check and report whether the same vulnerabilities detected	- Checking whether vulnerabilities identified in TLPT are also found in other systems is not made. As a result, similar vulnerabilities are identified in subsequent TLPT performed on other systems.

	in the TLPT are found in systems that were not subject to the TLPT.	
Others	- Conduct TLPT using multiple scenarios to address cyber threats posed by attacks from different actors, including criminal groups, entities suspected of state involvement, and insiders.	- A sufficient budget that align with the objective and purposes of TLPT is not secured, and threat intelligence and blue team assessment are omitted, resulting in inadequate testing for TLPT.

Figure 1: TLPT good practices and challenges
Source: Financial Services Agency, "analysis report on financial institutions' IT system failures" (June 2024).

Even such tests can contribute to enhancing cybersecurity in some way, and therefore the implementation of tests itself should not be denied, but its quality is inadequate in view of the purpose and anticipated effects of TLPT. Not all financial institutions would be in the situation to perform TLPT in a desirable way from the beginning given the difference in their risk profiles and resource constraints. However, they should not be satisfied of merely conducting TLPT for form's sake. Financial institutions need to understand how a desirable TLPT should be conducted and, [where TLPT is not conducted in a desirable way] report to the management appropriately that the TLPT they conducted was not one implemented in a desirable manner, and take action to cover the deficiencies. In cases where resource constraints are the obstacles to the implementation of desirable TLPT, appropriate communication with the management is required, conveying the significance of conducting effective TLPT to the management and securing sufficient resources through budget increases.

The fundamental cause of the problem that the quality of TLPT conducted by financial institutions are in effect inadequate lies in "cost" and "risk."

In terms of cost, compared to a typical penetration test, in TLPT, phases for threat intelligence and blue team assessment are added and this leads to increased cost and time. Particularly for small financial institutions, there is room to consider whether it is truly necessary to rigorously identify threat intelligence from the viewpoint of cost effectiveness. In the FSA's TLPT pilot project in business year 2024, threat information common to financial institutions in the same sector was shared by utilizing [available] threat intelligence. By simplifying the process of TLPT through an arrangement for information sharing, this would lower the barriers to conducting TLPT. These considerations are discussed in section 4.6.

Regarding risk, which is the other cause, financial institutions may be concerned about the risk that live production environment are affected by, for example, a system crash as a simulated attack is made in the production environment without notifying the Blue Team in advance. Such concerns could be stemming from the fact that financial institutions' White Teams have had little experience with TLPT, and that they have uncertainties about managing and conducting TLPT. In such situation, financial institutions can conduct, before testing in the live production environment, tests in an environment where there is little concern about the impact on the production environment, such as a system development environment, with the understanding of the management that such a test is different from a desirable TLPT that the FSA recommends. In this way, financial institutions can enhance the capability to manage and operate TLPT in the planning and testing phases step by step.

The author's thinking on how to approach these issues is explained in Section 4.

3. TLPT in Other Countries

Turning to developments in TLPT in other countries, particularly in Europe, respective authorities are promoting implementation of TLPT by financial institutions and taking policy measures to this end. This section describes TLPT frameworks and implementation procedures developed by financial regulators, focusing on developments observed in 2025 at the time this paper is prepared. We examine the background that led to the development of TLPT frameworks and explore the TLPT frameworks and procedures authorities are developing. We also discuss what kind of terms and standards exist regarding the scope of the test and financial institutions subject to the test. In addition, it addresses how authorities are involved in the tests and how the results of the tests are treated.

3.1 Digital Operational Resiliency Act (DORA)

In the financial sector, the enactment of EU's regulation, the Digital Operational Resilience Act (DORA), is a highlight of recent developments related to TLPT. DORA is an EU regulation promulgated on January 16, 2023, and came into effect on January 17, 2025.

DORA is a comprehensive framework, consisting of five pillars: (i) ICT-risk management and governance; (ii) ICT-related incident reporting; (iii) digital operational resilience testing; (iv) third-party risk management; and (v) information sharing. Of these, "digital operational resilience testing" sets out the basic requirements for resilience testing aimed at assessing preparedness for handling ICT-related incidents, of identifying weaknesses, deficiencies and gaps in digital operational resilience, and of promptly implementing corrective measures. Essentially, this requires financial institutions that are subject to the regulation to conduct TLPT, and EU national authorities are required to oversee and manage the implementation of TLPT by eligible financial institutions in their respective countries. The implementation of TLPT is expected to increase to a wide range of financial institutions with the DORA's enforcement in 2025.

DORA sets out regulatory technical standards (RTS) that define specific requirements and implementation methods for TLPT. The "RTS on threat-led penetration testing" was devised as the RTS for digital operational resilience testing, which provides detailed requirements and procedures for performing TLPT so that financial institutions can strengthen their cyber resilience. Specific and rigorous requirements are set out for financial institutions that meet certain criteria, such as global systemically important insurers (G-SIIs) and other systemically important institutions (O-SIIs), that are subject to the RTS.[5] For example, those subject to RTS are expected to perform TLPT once in three years at a minimum, and the duration of the active red team testing phase should last for at least 12 weeks. These requirements need to be recognized not only by EU financial institutions but also foreign financial institutions operating in the EU. Therefore, such requirements would become important standards for how TLPT should be implemented by financial institutions in the future.

3.2 TLPT framework led by authorities

TLPT frameworks applicable to major financial institutions have been developed in the UK and other European countries, with the involvement of authorities. The trend is spreading to Asia and countries outside Europe. Their frameworks have been devised based on UK's CBEST and EU's TIBER-EU, with consideration of each country's situation. Japanese financial institutions can also make use of frameworks developed overseas like CBEST and TIBER-EU, learning implications for how TLPT should be conducted. For the details of TLPT frameworks led by authorities such as CBEST developed before May 2018, please see "report on the conduct of TLPT in major countries and economies" referred to in section 2.2.1.

3.2.1 CBEST by the Bank of England

Amongst authorities, the Bank of England was the first to develop a TLPT framework for financial institutions, the CBEST, in 2014 jointly with the Council for Registered Ethical Security Testers (CREST), a non-profit organization that assures quality of services in the information security sector.

The primary objective of CBEST is to simulate realistic cyber threats that financial institutions are exposed to and to assess and strengthen their capabilities to defend their systems and business operations. This is expected to help financial institutions improve their resilience against cyber attacks.

The feature of CBEST is that the quality and reliability of the test are assured because the testers are selected from professional organizations accredited by CREST. CBEST is the first framework for TLPT presented by regulatory authorities, and other frameworks developed and/or revised after it have referenced CBEST as a best practice. CBEST is considered as a de facto standard

[5] "RTS on threat-led penetration testing" sets out specific criteria regarding financial institutions subject to TLPT in Article 2 "Identification of financial entities required to perform TLPT."
https://www.dora-info.eu/rts-tlpt/article-2/

for guidelines and frameworks concerning TLPT for financial institutions.

CBEST published an "Implementation Guide for CBEST participants" (hereafter "CBEST Implementation Guide") that explains the key phases, activities, deliverables and interactions involved in a CBEST assessment. This guide was revised in 2024, with three main changes. First is the "change in timeline," extending the typical project duration to "around 9 to 12 months" from the previous "9 months." Second is "strengthening of CG responsibilities," requiring control group (CG), the role assumed by White Team, to report to the regulator immediately any significant concerns in relation to the project plan and the technical execution of TLPT. And third is "strengthening of risk management," stipulating completion of risk assessment prior to testing that CBEST risk assessment process should ensure that the CG remains in technical and operational control of the CBEST during all phases. All three changes are intended to make the conventional CBEST framework further advanced.

3.2.2 "TIBER-EU" by the European Central Bank

TIBER-EU was issued in May 2018 as a TLPT framework that can be used by the financial sector in the 27 countries participating in the European System of Central Banks. This framework aims to assess and improve institutions' defense capabilities through cyber attacks that mimic tactics, techniques and procedures real-life threat actors.

TIBER-EU was developed based on CBEST and other frameworks prepared by other authorities, and has been improved drawing lessons from the implementation of CBEST.

As DORA came into effect in January 2025 in EU, TIBER-EU was updated in February 2025 to ensure alignment with TLPT defined in "RTS on threat-led penetration testing." The purpose of this revision is to ensure consistency with DORA, integrating the terminology and adding provisions on making purple teaming mandatory, which is required by DORA's RTS.

TIBER-EU provides guidance that complement the framework. For example, in the "TIBER-EU for the Blue Team Test Report," key considerations for preparing an assessment report in conducting blue team assessment are specified. Guidance listed below is useful in performing TLPT, so financial institutions are advised to refer to it when necessary.

Table 1: Guidance provided by TIBER-EU

Title	Content
TIBER-EU Framework	framework
TIBER-EU Control Team Guidance	guidance on control team
TIBER-EU Initiation Documents Guidance	getting started guidance
TIBER-EU Guidance for Service Provider Procurement	procurement guidance on service providers
TIBER-EU Scope Specification Document Guidance	scope specification
Targeted Threat Intelligence Report Guidance	reporting guidance on threat intelligence
TIBER-EU Red Team Test Plan Guidance	guidance on test plans of Red Team
TIBER-EU Red Team Test Report Guidance	guidance on test report of Red Team
TIBER-EU for the Blue Team Test Report	guidance on test plans of Blue Team
TIBER-EU Purple-Teaming Guidance	guidance on Purple Teaming
TIBER-EU Test Summary Report Guidance	guidance on preparation of Summary Report
TIBER-EU Remediation Plan Guidance	guidance on remediation plans
TIBER-EU Attestation Guidance	guidance on attestation

Source: European Central Bank (https://www.ecb.europa.eu/paym/cyber-resilience/tiber-eu/html/index.en.html)

3.2.3 TLPT framework led by other authorities

Since the development of CBEST in 2014 and TIBER-EU in 2017, authorities in other countries have also been leading development of TLPT programs.

The Hong Kong Monetary Authority (HKMA) devised "Intelligence-led Cyber Attack Simulation Testing" (iCAST) in December 2016. The feature of iCAST is that, based on risk assessment by HKMA, banks with "high" and "medium" level inherent risk profiles are required to carry out iCAST exercises.

In Singapore, "Red Team: Adversarial Attack Simulation Exercises−Guidelines for the Financial Industry in Singapore" (AASE, Version 1.0) was published in 2018. This document provides guidelines for the financial industry in Singapore for executing Red Team (RT) exercises and contains the guidelines and best practices to help organizations plan, execute and report such exercises. The "AASE" was updated in September 2024, as version 2.0, adding new contents related to Purple Teaming (Purple-Teaming is explained in Section 4.6).

In Saudi Arabia, the Saudi Arabian Monetary Authority (SAMA) published "Financial Entities Ethical Red-Teaming Framework (FEER)" in 2019. In addition to Red Team, Blue Team, and White Team, SAMA defines "Green Team" that performs the supervisory role in the conduct of TLPT, whose function is provided by the IT Risk of Financial Sector Supervision Department.

In Australia, the Council of Financial Regulators published "Cyber Operational Resilience Intelligence-led Exercises" (CORIE) in 2020. This is a program guide for TLPT for financial institutions (including Financial Market Infrastructure) in Australia. It was updated to version 2.0 in July 2022. The unique feature of CORIE is that senior executives of financial institutions participate in the "Gold Team" and take part in "crisis simulation table top" exercises to assess and improve crisis management procedures and senior management decision-making ability in preparation for a real cyber incident.

In EU member countries, each country has developed TIBER that corresponds to the situation of the country based on TIBER-EU. For example, Sweden is formulating and Iceland TIBER-IS. With the update of TIBER-EU in 2025, each country's TIBER program is expected to be updated regarding its implementation procedures to align with requirements for TLPT under DORA.

In 2024, after a decade since CBEST was developed, two frameworks by the authorities were published as TLPT frameworks. One is STAR-FS by CREST and the other is ART formulated by the De Nederlandsche Bank (DNB, the central bank of the Netherlands).

CREST published "Simulated Targeted Attack & Response for the Finance Sector" (STAR-FS) in March 2024. This is a framework that requires regulators to take the same level of rigorous approach as CBEST whilst reducing impact on regulatory resources. Hence, a wider range of financial institutions are accessible relative to CBEST, bringing the benefits to a wider set of firms.

The "Advanced Red Teaming (ART) Framework" published in April 2024 by DNB is a comprehensive framework that empowers a wide range of financial institutions to conduct advanced ethical red teaming tests driven by high-level threat intelligence. The ART framework is developed based on TIBER, but it also targets financial institutions that are not yet ready for a full TIBER test or are not subject to DORA/TLPT. It is a framework that financial institutions that are not large enough to be subject to DORA or TIBER can refer to when conducting Red Teaming.

In recent years, rules and standards related to the implementation of TLPT have been strengthened for large financial institutions that play an important role in providing financial services, as seen in the enactment of DORA. At the same time, frameworks that can be applied to financial institutions that are not yet ready for advanced TLPT are being developed, such as CREST's STAR-FS and the Netherlands's ART.

3.3 Other Guidelines and Frameworks

The previous section highlighted some of the authority-led TLPT frameworks. This section focuses on other guidelines, namely "G-7 Fundamental Elements for TLPT" and the framework by GFMA.

Published in 2018 by the G7 Cyber Expert Group is the "G-7 Fundamental Elements for

Threat-Led penetration test" (hereafter "G7FE-TLPT"),[6] discussed in section 2.1.2. The "G7FE-TLPT" provides "entities with a guide for the assessment of their resilience against malicious cyber incidents through simulation and a guide for authorities considering the use of Threat-Led Penetration Testing (TLPT) within their jurisdictions." The provisional Japanese translation of the "G7FE-TLPT" is available on the website of FSA. Along with the FISC Guide introduced in section 2.2.2, it is one of very few guidelines that are available in Japanese, so Japanese financial institutions considering implementing TLPT are encouraged to refer to it. The "G7FE-TLPT" consists of six elements. In the sixth element, "Thematic data," it is noted that the production of thematic data relating to TLPT engagements is the responsibility of the relevant authorities, and that thematic data should identify common sector findings and vulnerabilities, and by sharing of thematic data amongst authorities and entities, this would contribute to the improvement of the cyber resilience of entities and the financial sector more generally. In light of this statement, the FSA undertook a project collecting data from cases of TLPT from financial institutions in 2023 and provided feedback as thematic data on good practices and challenges of TLPT.

In 2019, the Global Financial Markets Association (GFMA) developed a commonly accepted framework for cybersecurity penetration testing, and published it as the Version 2 in 2020.[7] This Framework was designed to create an agreed upon approach for regulators and financial services firms to conduct effective testing to satisfy both supervisory and firm originated requirements. Further, the framework is primarily focused on the interaction between regulators and firms when conducting tests and is not intended to provide granular technical details of the testing process. It is interesting to note that many of the authorities-led TLPT frameworks we looked at in section 3.2 were developed from the perspectives of authorities, but the GFMA framework was created by the financial industry and so it reflects the viewpoint of the financial industry. Specifically, many of the authorities-led frameworks like DORA and TIBER-EU require TLPT to be conducted in the live production environment, but the GFMA framework anticipates cases where higher risk activities are managed by conducting tests off-hours, or against non-production systems.

4. Effective TLPT for Financial Institutions in Japan

So far, we have examined the initiatives related to TLPT by the Japanese authorities in Section 2, and overseas regulation, frameworks, and guidelines related to TLPT in Section 3. For those financial institutions that operate globally, addressing and complying with overseas regulation and guidelines is essential. Even for small and medium-sized financial institutions, the guidelines used by foreign financial institutions would greatly benefit them in addressing cybersecurity risk since there are no borders to cyber attacks. In this section, we explore issues related to how financial institutions in Japan should conduct TLPT based on current initiatives in Japan and on frameworks and guidelines in other countries.

4.1 Test Providers

When implementing TLPT, a key point for financial institutions is the selection of entities who conduct the test. Foreign financial institutions that operate globally have formed their own in-house red teams to conduct tests, but most of the Japanese financial institutions outsource testing to service providers such as external security vendors. Since TLPT is supposed to mimic real-life threat actors, including state-sponsored criminal groups, that use sophisticated techniques, it is utmost important for financial institutions to carefully assess service providers' skills when selecting test providers.

In TLPT frameworks abroad, CBEST requires service providers implementing TLPT to meet certain requirements regarding experience, and they need to be accredited by CREST. To go through an accreditation process of CREST, an operator must have at least a defined period of

[6] https://www.fsa.go.jp/inter/etc/20181015/20181015.html

[7] "A Framework for Threat-Led Penetration Testing in the Financial Services Industry" (https://www.gfma.org/correspondence/updated-gfma-framework-for-the-regulatory-use-of-penetration-testing-in-the-financial-services-industry/)

experience. DORA stipulates the requirements for entities that conduct TLPT in Article 27, and TIBER-EU has also formulated the "TIBER-EU: Guidance for Service Provider Procurement," which was revised in January 2025 to be consistent with the DORA.

Regarding testing operators, the credibility of the institution or of individuals is an important factor. In DORA, "requirements for testers for the carrying out of TLPT" are set out in Article 27. The article requires that testers are of highest suitability and reputability, and are certified by an accreditation body in a Member State or adhere to formal codes of conduct or ethical frameworks. In addition, DORA's RTS requires that the staff of the red team assigned to the TLPT is composed of at least a manager with at least 5 years' experience in penetration testing and red team testing and at least two additional testers, each with penetration testing and red team testing of at least two years. In order to conduct effective TLPT, it is important to clarify the requirements of the organizations that carry out the tests, and also requirements regarding individuals that compose the team, particularly a project manager of the team, and select test providers that meet the requirements.

Selection of inappropriate test providers can result in underestimation of risk, poor test quality, and cases of non-compliance. Financial institutions should avoid making an easy choice by simply choosing the test provider from the most reasonable cost estimates provided by multiple vendors. They should determine the test provider by carefully examining vendors whether they have credibility appropriate for operating tests.

As for Japanese documents laying down requirements regarding test providers, there is FISC Guide, and Section 3 titled "Considerations to be made in selecting service providers" presents viewpoints for evaluating test providers. The guide would be useful for financial institutions in Japan to refer to when selecting service providers.

Given that no arrangement is in place in Japan where a body like CREST certifies service providers, it is difficult to assess the technical capabilities of test providers in an objective manner. However, since there are many guidelines at home and abroad regarding the selection of service providers, financial institutions should refer to these guidelines and select highly-skilled service providers based on past performance (including the years of experience and details of past assignments) and the references proving capabilities to provide effective and high-quality services (including the certifications held).

Regarding in-house test teams, conventional frameworks such as CBEST and TIBER-EU (2018 version) do not assume in-house testing. However, more recent frameworks such as DORA, TIBER-EU (revised in 2025), and the Netherlands's ART, were developed assuming that some tests are operated by in-house teams. DORA stipulates requirements for using internal testers that such use should be approved by the relevant competent authority and that conflicts of interest should be avoided. The TIBER-EU framework also writes "In TIBER-EU, it is mandatory to use an external TIP [Threat Intelligence Providers] and strongly encouraged to use external RTT [Red Team Testers]," and attach conditions for cases where an internal team is used for a TIBER-EU test. For tests like vulnerability assessment and penetration testing that aim at identifying systems' vulnerabilities and are relatively simple to conduct, there are benefits to conducting such tests internally in terms of efficiency and cost effectiveness. However, for the implementation of TLPT, which requires advanced technical capabilities and independent perspective, internal testers need to adhere to certain conditions. It is important to distinguish tests by external service providers and those by in-house teams appropriately.

4.2 Target Environment

When performing TLPT, to test in a live production environment or in non-production environment (such as against systems in development or staging environment) is a key point to consider. The live production environment and non-production environment do not have the same security level. In addition, the same level of assessment as in the production environment may not be obtained for Blue Team assessment if tested in non-production environment. On the other hand, if testing is conducted in a live production environment, certain attack simulation may include high-risk activities, and if defense capabilities are inadequate and a system disruption occurs at a financial institution, this could cause significant impact, possibly inducing large damages that exceed the benefits obtained from TLPT. Therefore, financial institutions should thoroughly evaluate whether it is appropriate to conduct tests in the live production environment by conducting risk analysis at the scope specification phase.

We examine how various guidelines lay down the target environment for testing. First, in DORA, paragraph 2 of Article 26 reads "Each threat-led penetration test shall cover several or all critical or important functions of a financial entity, and shall be performed on live production systems supporting such functions." The clause requires TLPT to be conducted in a live production environment. The conventional frameworks used in Europe such as CBEST and TIBER-EU also require tests to be performed in a production environment.

Under DORA and CBEST, which are frameworks led by authorities that require implementation of advanced TLPT, tests are to be conducted in the live production environment. However, as described in section 2.2.3, considering the risks in reality, there may be a number of financial institutions that would hesitate to conduct tests in the production environment, even if they have concluded contracts with vendors that have experienced Red Teams. This may be due to a lack of experience of the White Team or its uncertainties about its management skills stemming from lack of trust with Red Team. White Team needs to have a thorough planning[8] at the preparation phase to prevent risks from materializing and to assess the skills of Red Team. In this regard, it may be one approach for financial institutions to take a step to be matured for conducting TLPT by verifying capabilities of White Team and Red Team through TLPT in a non-live environment (systems in development) or in a staging environment to gain experience of testing, although such testing is different from what the FSA recommends. It must be noted that testing in a non-production environment and judging that there was no problem from the outcome of such testing is not adequate to assess "people," "organization," and "processes" of the financial institutions. In light of the objectives of TLPT, to make the test results meaningful and effective, the final goal should be to perform testing in the live production environment. In the implementation of TLPT in production environment, it should be noted that there can be certain operational risks, and so there must be necessary preventive measures and measures to address operational incidents in place.

4.3 Test Approach

There are generally three approaches to TLPT: black box test; grey box test; and white box test.

A black box test is a method of test in which the tester is not provided with any information in advance about the target systems. Since the objective of TLPT is to conduct tests by simulating real-life cyber threats, this approach is very much in line with the objective but there are drawbacks to it because, given that little information is provided beforehand, it may take time to identify risks, and in some cases significant impacts may be exerted on live production environment.

A grey box test is a method of tests in which the minimum necessary information is provided to the test provider, and a white box test is a test method in which all the necessary information for the test is provided to the test provider. These methods have the benefits in terms of efficiency because, the wider the scope of information disclosed, the more efficient it is with respect to time and cost, but there are disadvantages as it becomes difficult to assess the real-life threats. Grey box testing is effective in shortening the reconnaissance phase that takes time and the phase of collecting information through attacks, but it should be noted that there is a concern that an attack made upon giving necessary information to the Red Team may be different from the situation expected to arise in real-life cyber attacks. When a grey box test is chosen, it is important to ensure that only the necessary information is provided to the tester.

Regarding test method, CBEST and G7FE-TLPT do not assume a full black box test for TLPT, as a target financial institution is supposed to provide certain information to the test provider prior to the testing in threat intelligence. In addition, TIBER-EU (2025 updated edition) advocates selecting grey box approach, that the financial institution may deliver additional information to the Red Team (test provider) to facilitate an effective and efficient test. As evident from these guidelines, TLPT frameworks of other countries do not necessarily require a "black box" testing approach, despite the thinking that tests should simulate real-life cyber threats.

It would be possible to take an approach of starting the test first as a black box test, and depending on the outcome of a simulated attack and the progress in the overall test schedule,

[8] It is considered that in the practice of TLPT, a breach of system administrators' authority is deemed a successful compromise of systems, and an act of breaking further into systems using that authority is not made.

information may be provided partially, moving to a "grey box" approach flexibly according to the situation. However, such a flexible approach would require a highly-skilled Red Team and also close cooperation between White Team and Red Team. In addition, it would be one option to choose a grey box test so as to conduct the test effectively and efficiently from the cost perspective as well as to control the impact on the live production systems. However, it should be noted that, when Red Team has been successful in making a simulated attack, if the financial institution underestimates the vulnerabilities that were identified by the test based on the fact that information that was supposed to be undisclosed had been provided to the Red Team, it would be a waste of money to conduct TLPT that costs so much.

4.4 White Team (Control Team)

White Team plays key roles in the overall management and coordination of the test in TLPT. In the past, the team was called "White Team" to distinguish it from "Red Team" and "Blue Team." CBEST calls it "Control Group," and TIBER-EU and DORA call it "Control Team." In Europe, the White Team is called according to the roles it performs.

TIBER-EU has published "TIBER-EU Control Team Guidance" that defines requirements for setting up a control team. The guidance defines that the control team plays a number of important roles necessary for implementing effective TLPT. Control team performs the following roles: planning the test; selecting and managing external test operators; risk assessment; ensuring that the test is conducted in a safe and controlled manner; communicating to and between the involved stakeholders; overseeing the implementation of test; responding to problems; and reporting the test assessment after the test. White Team plays central roles in TLPT and also plays important roles to ensure that the test is conducted in a safe and efficient manner and maximize learning experience from the test by controlling the risks. Therefore, the role of White Team is pursued by a control team lead that has a skillset that includes risk management skills, project management skills, ability to communicate with the different levels of staff, and also by external security experts where necessary.

If the skills of White Team such as knowledge of cyber attacks, risk management skills, and ability to communicate are inadequate, this would result in poor quality of testing due to inappropriate decision making and delay in reporting to stakeholders. There may be cases where considerable damage is caused to the live production environment due to poor planning and inappropriate decision making, as described in section 4.2. White Team should therefore be set up carefully so that its responsibilities are fulfilled by members with sufficient skillset.

4.5 Threat Intelligence

TLPT is generally classified into the phase of threat intelligence and the phase of penetration test (Red Team Testing). One of the differences between TLPT and a general penetration test is that TLPT includes a process for analyzing potential real-life cyber threats and conducting the test based on the analysis, by having a threat intelligence phase before carrying out the test. Thus, in moving from penetration testing to TLPT, this threat intelligence phase becomes extremely important.

The threat intelligence phase can be, as noted in section 2.2.3, a burden for financial institutions. For small and medium-sized financial institutions, in particular, it is not clear whether there would be specific threats that target them and therefore it would be difficult for them to find benefits in bearing costs for such analysis. It may also be difficult for these institutions to imagine how the analysis of threat intelligence is reflected in cyber threat scenarios. In conducting TLPT, where there are budget constraints, threat intelligence phase could be the very reason financial institutions consider giving up the implementation of TLPT.

Regarding this point, it is interesting that while many of the TLPT frameworks led by authorities require service providers to conduct threat intelligence, the TLPT Framework developed by GFMA, the industry association, reads "It is recommended that industry and regulators jointly identify and prioritize the threats in order to develop test scenarios at the financial sector level." By sharing with the financial sector threat information that industry associations and regulators identified, individual institutions can obtain necessary information without incurring too much cost, so this would be a reasonable way to address threat intelligence.

CBEST, STAR-FS, and "G7FE-TLPT," consider the "intelligence phase" by classifying it

into “threat intelligence” and “targeting.” “Threat intelligence” aims at creating realistic attack scenarios by analyzing cyber threat actors’ attack techniques, while “targeting” aims at scope specification, selecting target attack surfaces, and identifying attack scenarios to be used in TLPT. The threat intelligence mentioned here may vary depending on the industry or the type of business, but it would not differ so much on the firm level. The threat intelligence could be conducted with the initiative of industry associations and regulators, and by doing so the overall cost for threat intelligence can be contained. In the pilot project the FSA conducted in business year 2024, one of the objectives was to share threat information that is common to financial institutions across the sector, based on threat intelligence by regional financial institutions. Through these initiatives, it is hoped that the scope of financial institutions that can conduct TLPT will broaden.

As we have examined above, the desirable way to conduct TLPT would be for financial institutions to analyze their own threat information using the knowledge of external threat intelligence providers, and conduct TLPT based on that threat information. However, for small and medium-sized financial institutions where threats specific to them are difficult to identify, if the process of threat intelligence identification and scenario preparation can be simplified by creating a common scheme, this could lower the barriers to TLPT. On the other hand, if threat intelligence specification is omitted from the viewpoint of reducing costs and time, and the scope of test target is limited to specific systems or cyber threat tactics used for attacks are decided in advance based on general threat intelligence information,[9] this would undermine the effectiveness of TLPT because potential cyber threats that the institution is exposed to are not sufficiently considered in conducting the test.

4.6 Red Team Test

Red team testing is conducted based on the scenarios developed through the threat intelligence phase. A typical attack scenario by an organized threat actor would compromise an institution’s internal network via the Internet and attempts to breach it by social engineering, such as phishing, or by an ingenious method to exploit system vulnerabilities. However, more advanced methods could be attack scenarios starting from internal fraud or physical intrusions into buildings.

In this penetration testing phase, we would like to focus on the timeline prescribed in various frameworks developed by foreign countries. In DORA’s RTS and TIBER-EU that is based on DORA, time allocated to the active red team testing phase should be at least 12 weeks. CBEST provides an indicative timeline of 14 weeks, STAR-FS expects a timeline of 4 to 6 weeks, and the GFMA framework developed by the industry association requires 10 to 12 weeks from the start to the end of TLPT.

Table 2: Timelines penetration testing in selected frameworks on TLPT

	TIBER-EU	CBEST
Total duration	9-12 months	9-12 months
Preparation phase	up to 6 months	up to 6 weeks
Threat intelligence phase	4-6 weeks	up to 10 weeks
Red team testing	at least 12 weeks	up to 14 weeks
Closure phase	up to 18 weeks	up to 4 weeks

Source: TIBER-EU FRAMEWORK (https://www.ecb.europa.eu/paym/cyber-resilience/tiber-eu/html/index.en.html), Implementation Guide for CBEST participants (https://www.bankofengland.co.uk/-/media/boe/files/financial-stability/financial-sector-continuity/cbest-implementation-guide.pdf)

[9] For example, when a financial institution considers that the threat of fraudulent emailing is increasing at the institution and conducts a cyber test by sending emails with strange Japanese expressions or emails with attached files that can easily be detected with an anti-virus software and pattern matching, such a test could be considered as something that does not match the actual level tactics in real-life cyber threats, merely conducted for the purpose of doing a test.

If compliance with DORA is not necessary, the testing phase does not necessarily adhere to "at least 12 weeks." However, to conduct TLPT that assumes cyber attacks that could occur in reality, such as a threat actor hiding in after compromising the system and taking time before moving to make attacks or carrying out persistent attacks, it needs to be understood that this testing phase requires considerable time. When financial institutions conduct TLPT, if it is completed within several days or in a week, such a test would in fact be similar to vulnerabilities assessment or a penetration testing where the scope of testing is limited to specific systems, or the testing method is often conducted for the purpose of detecting specific vulnerabilities. The quality of such testing is considered inadequate in terms of TLPT. It should be noted that we should differentiate TLPT from vulnerabilities assessment and penetration testing.

Assessing vulnerabilities or verifying the possibility of compromise to specific systems based on pre-defined assessment methods is useful as vulnerabilities assessment and general penetration testing, but they do not simulate cyber threats that could happen in reality such as attacks that change flexibly in line with the situation, and are therefore not adequate as an effective TLPT. To regard the implementation of simplified tests similar to vulnerabilities assessment or general penetration testing as TLPT, and to conclude that no significant problem was found by such testing would mislead the management's decision making. In conducting TLPT, the quality of red team testing is important for the selection of test providers.

4.7 Blue Team Assessment

One of the important points when conducting TLPT is the blue team assessment. Regarding this blue team assessment, we need to consider whether to conduct the test in the live production environment or not, as discussed in section 4.2 "Target Environment." Given that the objective of TLPT is to evaluate how developed the institution's people and response processes are, it would be desirable to test in the live production environment without prior notice to Blue Team to assess its detection and response capabilities.

DORA and TIBER-EU clearly stipulate that TLPT is to be conducted without prior notice to the Blue Team. The G7EF-TLPT also specifies that the test would be conducted without the foreknowledge of the Blue Team.

As discussed above, in light of the objective and purposes of TLPT, it is important to conduct it without the Blue Team knowing about it in advance. In the meantime, until the Blue Team has matured adequately to conduct TLPT, it would be effective to assess the current level of the institution's response capabilities against cyber threats by taking steps gradually, such as by informing the Blue Team in advance or informing partially (limited to the timing of the test or address information of the attacker). However, ultimately, more advanced and effective assessment of Blue Team can be made by conducting TLPT without the foreknowledge of the Blue Team.

In addition, DORA and TIBER-EU clearly state that the institution should carry out a purple teaming exercise, which is a collaborative testing activity that involves Red Team and Blue Team aimed at improving the institution's response management against cyber threats. A purple teaming exercise is carried out in closure phase after Red Team testing is completed, and Red Team and Blue Team collaborate to replay the test and exchange feedback on details of the attack, detection status, and gaps between threats and response capabilities, etc. Through this exercise, both teams' knowledge and experience are shared for the assessment of Blue Team's defense capabilities, aiming for their improvement and devising effective security measures. Purple teaming exercises are useful for improving the effectiveness of TLPT and financial institutions' response system against cyber attacks.[10]

4.8 Authorities' Involvement

[10] DORA and TIBER-EU also stipulate limited purple teaming employed under exceptional circumstances in the test phase where triggering of significant risks, such as disruption to critical services, impact on data integrity, impact on the financial sector, need to be avoided.

One of the major differences between Japanese and European TLPT initiatives is the involvement of authorities. In Japan, as described in section 2.2.1, in business year 2023 FSA took initiatives in creating a scheme to collect cases of TLPT by financial institutions and to feedback data on good practices and challenges, and in business year 2024, FSA implemented pilot projects to test the effectiveness of TLPT on regional financial institutions. Furthermore, as discussed in section 2.2.2, FSA published in 2024 the "Guidelines for Cybersecurity in the Financial Sector," which identified viewpoints desirable for implementing TLPT.

Meanwhile, the FSA has not adopted an approach where, as major regulators in Europe do, develop TLPT frameworks and engage in TLPT conducted by financial institutions as an overseer. Similarly, this European approach is taken in the United States, which is an advanced cybersecurity country. Authorities encourage financial institutions to take their own initiatives for TLPT, and support them by providing information where necessary.

For financial institutions in Japan, especially large financial institutions and regional financial institutions that represent the regions, they should understand the objective of the "Guidelines for Cybersecurity in the Financial Sector" and take necessary action to carry out TLPT in the way it should be conducted. As for the authorities, continuous efforts of the FSA in providing useful information for TLPT to financial institutions would lead to effective implementation of TLPT by financial institutions in Japan.

5. Conclusion

In concluding this report, I would like to discuss two points that I consider are important in implementing TLPT, upon examining frameworks and guidelines in other countries mentioned in section 4 and through interviews I had with a number of financial institutions that have conducted TLPT.

The first point is "involvement of the management." The management needs to consider cybersecurity as investment and not as cost, and it is extremely important that they allocate necessary and adequate resources to mitigate the risks. This point applies to cyber security issues overall. To this end, person in charge of system security on the front line should discuss with the management about the importance and the necessity of TLPT, and promote TLPT with the involvement of the management. The management should have an active approach to the implementation and the outcome of TLPT. Furthermore, it is extremely important that management ensures that there is an organizational culture in which any vulnerabilities or challenges discovered through TLPT are reported from staff on the front line to the management without hesitation, and the front line staff and the management communicate thoroughly to understand the risks, and advance works to improve cybersecurity measures for better risk management system with allocation of necessary costs and resources under the leadership of the management.

Particular attention should be paid to the risk that, when reporting the results of TLPT to the management, if the quality of TLPT implemented was inadequate but the test result was reported as if TLPT was conducted in a manner that was required, this could mislead the management's decision making. For example, even though TLPT was carried out with advance notice to Blue Team, this was not informed to the management, and merely reporting an assessment that Blue Team's detection and response capabilities was satisfactory could compel the management to make incorrect decision for the improvement of risk management system against cyber threats. If any concessions were made to the implementation of TLPT, these should also be reported to management.

Major regulations and guidelines such as DORA and TIBER-EU do not specify the involvement of the management clearly, but the Netherlands's ART requires one of the board members to be part of the control team (White Team). In addition, ART specifies that the board is informed of the threats, test results and the remediation plan (risk mitigation measures), and expresses the importance of board involvement, support and accountability in executing the remediation plan.

The second point is "sharing of findings from TLPT across the institution." There are cases where, even if TLPT is conducted every year, the same vulnerabilities or findings become apparent each year. This could be due to remediation being made to the specific systems for which TLTP was

targeted, and the same problem becomes evident the following year when a different system is tested. To avoid such a case, when TLPT findings that are particularly important in terms of the level of risks, or vulnerabilities that could exist in other systems (e.g., unnecessary password authority or vulnerable password setting, inappropriate authorization process) become evident, in addition to the remediation of the relevant system, it is desirable to implement initiatives to reduce risks for the entire institution by requiring checks and reports on if there are similar vulnerabilities in other systems. This should not be done by just using a checklist and requiring reporting, but by using methodologies that enable system verification across all the systems of the institution to make the remediation more effective. Considering that TLPT is generally costly and so the frequency of its implementation can be limited, we would like to use the lessons learnt from TLPT effectively.

As we have observed, perspectives on how TLPT is conducted vary from country to country. Large financial institutions that are considered more mature in their cybersecurity management systems, particularly those that are subject to DORA and CBEST, need to understand and comply with the regulation. As for small and medium-sized financial institutions, their cybersecurity management systems are not necessarily mature enough, and when a number of high-risk factors s are detected even in the vulnerability assessment or penetration testing, it seems there is little significance in conducting TLPT straightaway. They should first address vulnerabilities and strengthen their system security, and also launch appropriate incident response management systems. After taking these measures, it would be ideal to enhance defense systems against cyber threats by conducting effective TLPT as a final step.

This study examined issues and points that financial institutions in Japan should consider in implementing TLPT, by referring to TLTP frameworks of selected countries. We hope it will be useful for financial institutions in Japan in conducting effective TLPT.

Schedule

Table 3: Regulation and frameworks on TLPT introduced in this paper (in the order of publication or revision)

Title	Country/ Region	Issuance/ Last revision
Intelligence-led Cyber Attack Simulation Testing (iCAST)	Hong Kong	2016
G7 Fundamental Elements for Threat-Led Penetration Testing	G7	2018
Financial Entities Ethical Red-Teaming Framewosrk (FEER)	Saudi Arabia	2019
Guideline on the implementation of TLPT by financial institutions	Japan	2019
A Framework for Threat-Led Penetration Testing in the Financial Services Industry	GFMA (industry association)	2020
Cyber Operational Resilience Intelligence-led Exercises (CORIE)	Australia	2022
Simulated Targeted Attack & Response for the Finance Sector (STAR-FS)	UK	2024
Advanced Red Teaming Framework (ART)	The Netherlands	2024
Adversarial Attack Simulation Exercise Guidelines for the Financial Industry in Singapore (AASE)	Singapore	2024
Cybersecurity Guidelines in the Financial Services Sector	Japan	2024
Critical National Infrastructure Banking Supervision and Evaluation Testing (CBEST)	UK	2024
Digital Operational Resilience Act (DORA)	EU	2025
European Red Team Testing Framework (TIBER-EU)	EU	2025

References

The Association Banks in Singapore (2024), "Adversarial Attack Simulation Exercises 2.0" (https://www.abs.org.sg/docs/library/abs-red-team-adversarial-attack-simulation-exercises-guidelines----september-2024.pdf)

Bank of England, "CBEST Threat Intelligence-Led Assessments" (https://www.bankofengland.co.uk/financial-stability/operational-resilience-of-the-financial-sector/cbest-threat-intelligence-led-assessments-implementation-guide)

Bank of England, "Implementation Guide for CBEST participants" (https://www.bankofengland.co.uk/-/media/boe/files/financial-stability/financial-sector-continuity/cbest-implementation-guide.pdf)

Bank of England (2024), "STAR-FS UK Implementation Guide" (https://www.bankofengland.co.uk/-/media/boe/files/financial-stability/star-fs-implementation-guide-march-2024.pdf)

Council of Financial Regulators, Australia (2022), "Cyber Operational Resilience Intelligence-led Exercises (CORIE) Framework Ver.2.0" (https://www.cfr.gov.au/publications/policy-statements-and-other-reports/2022/revised-corie-framework-rollout/pdf/corie-framework.pdf)

European Central Bank, "TIBER-EU FRAMEWORK" (https://www.ecb.europa.eu/paym/cyber-resilience/tiber-eu/html/index.en.html)

European Insurance and Occupational Pensions Authority, "Digital Operational Resilience Act (DORA)" (https://www.eiopa.europa.eu/digital-operational-resilience-act-dora_en)

Financial Information Systems Center (2019), "The Guide for Financial Institutions to Implementing TLPT" (in Japanese) (https://www.fisc.or.jp/publication/book/004197.php)

Financial Servies Agency (2018), Japanese translation of "G-7 Fundamental Elements for TLPT" (https://www.fsa.go.jp/inter/etc/20181015/02.pdf)

Financial Services Agency (2024), "analysis report on financial institutions' IT system failures" (in Japanese) (https://www.fsa.go.jp/news/r5/sonota/20240626/20240626.html)

Financial Services Agency (2018), the second edition of "Policy Approaches to Strengthen Cyber Security in the Financial Sector" (in Japanese) (https://www.fsa.go.jp/news/30/20181019-cyber.html)

Financial Services Agency (2022), the third edition of "Policy Approaches to Strengthen Cyber Security in the Financial Sector" (in Japanese) (https://www.fsa.go.jp/news/r3/cyber/torikumi2022.html)

Financial Services Agency (2024), "Guideline for Cybersecurity in the Financial Sector" (https://www.fsa.go.jp/news/r6/sonota/20241004/18.pdf)

Financial Services Agency (2018), "report on the conduct of TLPT in major countries and economies" (in Japanese) (https://www.fsa.go.jp/common/about/research/20180516/TLPT.pdf)

Global Financial Markets Association (2020), "A Framework for Threat-Led Penetration Testing in

the Financial Services Industry"
(https://www.gfma.org/wp-content/uploads/2020/12/gfma-penetration-testing-guidance-for-regulators-and-financial-firms-version-2-december-2020.pdf)

Hong Kong Monetary Authority, "Intelligence-led Cyber Attack Simulation Testing (iCAST)"
(https://www.hkma.gov.hk/eng/data-publications-and-research/guide-to-monetary-banking-and-financial-terms/iCAST/)

De Nederland Bank (2024), "Advanced Red Teaming (ART) Framework"
(https://www.dnb.nl/media/cxjjcc4b/art-framework-april-2024.pdf)

Saudi Arabian Monetary Authority (2019), "Financial Entities Ethical Red-Teaming"
(http://sama.gov.sa/en-US/Laws/BankingRules/Financial%20Entities%20Ethical%20Red%20Teaming%20Framework.pdf)

Financial Research Center（FSA Institute）
Financial Services Agency
Government of Japan

3-2-1 Kasumigaseki, Chiyoda-ku, Tokyo 100-8967, Japan

TEL：03-3506-6000
URL：https://www.fsa.go.jp/frtc/english/index.html