# A Framework for Threat-Led Penetration Testing in the Financial Services Industry

Version 2
December 2020

gfma afme/ asifma sifma

# 1 Table of Contents

## Disclaimer

These materials are for general informational purposes only, and are not intended to provide, and do not constitute, investment, tax, business or legal advice to any individual or entity. The views and opinions expressed in these materials are solely those of the authors and do not necessarily reflect the official policy or position of GFMA*, SIFMA, AFME, ASIFMA, or their employees, or members. We make no representations warranties or guarantees, expressed or implied, that the information contained herein is up-to-date, accurate, or complete, and we have no obligation to update, correct, or supplement this information, or to otherwise notify you, in the event that any such information is or becomes outdated, inaccurate, or incomplete. To the fullest extent permitted by law, we expressly disclaim all warranties of any kind, whether expressed or implied, including but not limited to implied warranties of merchantability, fitness for a particular purpose, title, and non-infringement. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by GFMA, SIFMA, AFME, ASIFMA, or their employees, or members.

* The Global Financial Markets Association ("**GFMA**") brings together three of the world's leading financial trade associations to address the increasingly important global regulatory agenda and to promote coordinated advocacy efforts. The Association for Financial Markets in Europe (AFME) in London, Brussels and Frankfurt, the Asia Securities Industry & Financial Markets Association (ASIFMA) in Hong Kong and the Securities Industry and Financial Markets Association (SIFMA) in New York and Washington are, respectively, the European, Asian and North American members of GFMA. For more information, please visit http://www.gfma.org.

## Executive Summary

Cybersecurity is a top priority for the financial sector. This has resulted in authorities and the sector developing mechanisms to test the resilience of firms by use of various methodologies like vulnerability assessments, application vulnerability scanning, penetration testing, red-teaming and threat-led penetration testing. Each type of testing has its own unique objective, technique, and scope and this Framework acknowledges that there are many testing types available for firms to assess the effectiveness of their security programs (details and differences are captured in the appendix). As such, this document will focus on threat-led penetration testing due to the high risk of operational disruptions that such testing could cause if undertaken without appropriate planning and involvement from across an organization by a qualified team of practitioners.

Testing allows firms to evaluate their systems and the controls that protect them in order to identify and remediate vulnerabilities, thereby strengthening their infrastructure and organization against cyber threats. Likewise, for regulators, testing can help identify systemic issues and trends of where vulnerabilities might persist. Increased interest by global regulators has led to the proliferation of regulatory-mandated testing regimes. While testing is understandably important to regulatory oversight, it can also introduce risks to firms and consumers—for example, if test results are publicly exposed, inadvertently disclosed, or stolen. This is especially true if testing is not approached in a collaborative and coordinated manner between the institution and the regulator.

To that end, the GFMA and our members jointly developed and published, in July of 2019, a set of principles to guide the development of testing frameworks to harmonize the growing regulatory demand for threat-led penetration testing. In this 2020 version, these principles are updated based on the evolution of industry best practices and guidance from frameworks around the world including Bank of England's CBEST[1], European Union's TIBER[2] , Hong Kong Monetary Authority's - iCAST[3], Saudi Arabian Monetary Authority FEER[4] and the G-7's Fundamental Elements for Threat-Led Penetration Testing[5].

This Framework is 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. The Framework's objectives are to:

- Establish testing best practices that both regulators and individual firms can use to reduce risk and enhance security.
- Engage regulators globally with common principles to facilitate open dialogue about frameworks like CBEST, TIBER, i-CAST, for regulator-led and firm-led testing.
- Ensure both regulatory and financial firm concerns and recommendations are considered.

[1] https://www.bankofengland.co.uk/-/media/boe/files/financial-stability/financial-sector-continuity/cbest-implementation-guide
[2] https://www.ecb.europa.eu/pub/pdf/other/ecb.tiber_eu_framework.en.pdf
[3] https://www.hkma.gov.hk/media/eng/doc/key-information/guidelines-and-circular/2016/20161221e1.pdf
[4] http://www.sama.gov.sa/en-US/Laws/BankingRules/Financial%20Entities%20Ethical%20Red%20Teaming%20Framework.pdf
[5] https://www.boj.or.jp/en/announcements/release_2018/rel181015k.htm/

- Serve as industry-wide guidance whereby emerging technologies, knowledge of threats, industry-leading practices, and regulatory requirements can help drive continued iteration of the Framework.

The target audience for this Framework includes those in the financial services industry who conduct, rely on or call for the execution of threat-led penetration testing, including (but not limited to):

- Financial Industry Regulators / Supervisors
- CISOs, CROs and Senior Management
- Information Security Professionals
- Information Owners / Technology Specialists
- Internal and External Penetration Testers
- Third-party Stakeholders (e.g., in the case of managed systems testing, cloud computing providers, etc.)
- Other Industries like Fintech, Telecommunication, Media, etc.

The Framework documented below outlines a four-phased testing lifecycle (see diagram) for threat-led penetration testing to ensure firms are following industry best practices while simultaneously meeting regulatory demands. The four phases of threat-led penetration testing are:

- **Threat Intelligence Phase** – A firm's intelligence (either internally or externally provided) could be augmented by government agencies and sector level financial industry resources where possible. Final threat intelligence scenarios should be agreed upon between regulatory authorities and financial firms, where applicable.
- **Planning Phase** – Test activities should be prioritized and scheduled according to threat intelligence and regulator input in planning the scope of the exercise. Firms looking at establishing mutual recognition of the test should agree on the scope with the lead regulator and other regulatory regimes for which they are trying to achieve mutual recognition.
- **Testing Phase** – Testing should begin after operational planning and attack methodologies are agreed upon.
- **Analysis and Response Phase** – This phase includes the development of executive / technical reports and associated firm responses. Summary versions of final reports may be distributed internally within the firm and to regulators and would include a sign-off from an organization's, CISO, Chief Risk Officer or Senior Management, where appropriate, on the identified vulnerabilities and associated remediation plan.

## Threat-Led Penetration Testing Lifecycle


While the Framework provides an approach for threat-led penetration testing, it is not intended to serve as a detailed industry playbook for conducting testing, generally. It is primarily focused on the interaction between regulators and firms when conducting threat-led tests and is not intended to provide granular technical details of the testing processes.

This document, while directional in nature, is designed to guide both regulators and the financial services industry to develop a mutually agreeable, safe, secure, and scalable testing program to manage inherent shareholder, investor, market and firm reputation/financial risks arising out of a potential cyber-attack.

It is important to note that while threat-led penetration testing is a powerful tool, it is not a "silver bullet" used to solely evaluate a firm's cybersecurity risk management program or even the financial/operational health of a firm. It should be used in conjunction with other testing and monitoring methodologies to provide a holistic view of a firm's approach to cybersecurity.

As technologies emerge that provide a safer, more efficient method for evaluating the efficacy of cybersecurity controls, industry should continue to evolve their threat-led testing methodologies into their cybersecurity risk management programs.

## Contributing Organizations

The following financial firms, trade associations and consulting organizations provided substantial support and input towards developing this document and its contents.



















# Introduction

## 1.1 Background

The growing regulatory interest in threat-led penetration testing (TLPT) and the proliferation of various frameworks and approaches drove the development of this Framework to ensure consistent, safe, secure, and scalable testing regimes. Threat-led penetration tests are powerful tools to assess a firm's cyber security program, but due to their invasive nature, testing data and results are particularly sensitive. Testing results with vulnerability data, even if anonymized, can provide a clear roadmap to attack a firm, therefore it is imperative that the distribution of detailed test results be tightly controlled. With more jurisdictions interested in these invasive tests, firms face increasing operational risk without an agreed industry framework for performing tests that can safely fulfill supervisory requirements.

The Global Financial Markets Association (GFMA) together with the Securities Industry and Financial Markets Association (SIFMA), Association for Financial Markets in Europe (AFME), Asia Securities Industry and Financial Markets Association (ASIFMA), in partnership with the financial industry, issued a joint comment letter[6] during July 2016 outlining issues associated with regulatory-driven testing followed by a set of principles[7] issued December 2017 intended to harmonize the growing regulatory demand for TLPT.

A common testing framwork would allow for regulators to guide programs being developed in a particular region and assurance that the tests are conducted by appropriately trained and experienced personnel regardless of whether it is a firm-led or 3rd party tester. The framework would provide for safe, secure and scalable testing operations satisfiying regualtory goals and minimizing risk and burden to the firms.

GFMA then brought together a team of financial sector subject matter experts, to develop a framework for carrying out safe, secure, and scalable testing (see list of Contributing Organizations above). This framework was developed over a period of 6 months by 3 working groups covering the following topics:

- Scoping and scheduling tests
- Executing the tests
- Communicating test results to management and regulators

The working group members have engaged with regulators and carried out a thorough review of existing and emerging regulatory TLPT frameworks and guidelines and have used the findings from this analysis to continuously enhance the industry framework and bring about harmonization of the various regulatory frameworks and guidelines.

[6] https://www.gfma.org/wp-content/uploads/0/83/91/207/183190df-454c-4b99-99ee-3de70ea70d83.pdf

[7] https://www.gfma.org/uploadedfiles/News/GFMA_in_the_News/2017/GFMA-Penetration-Testing-Principles.pdf

## 1.2 Purpose of this Framework

This Framework provides a guide for the development of a safe, secure, and scalable testing program for firms and regulators for conducting effective testing while managing operational risk. This Framework offers guidance and best practices on building threat-led penetration testing aligned with firm needs and regulatory expectations while providing regulators confidence that firms are conducting quality tests that appropriately assess cybersecurity risk management programs. With quality testing programs in place, and jurisdictions mutually recognizing testing performed under a common framework, firms should be able to satisfy different testing requests on processes and technologies used across multiple locations with a single test for a critical process used across multiple locations, thus eliminating wasted resources and, at the same time, keeping sensitive data tightly controlled.

The Framework design was guided by the following core principles which aims to:

- Provide firms and regulators the ability to develop and guide their TLPT programs to meet firm and supervisory objectives through use of scenarios based on current risks.
- Provide regulators with a high degree of confidence that testing is conducted by independent (either third-party or independent internal teams with proper governance in place[8]), trained, experienced personnel with sophisticated tools that can accurately emulate adversaries, as required.
- Provide regulators transparency into the testing process inclusive of firm-led testing and results as well as assurance that firm governance identifies and properly addresses weaknesses.
- Ensure testing activities are conducted in a manner that minimizes operational risks. For instance, testing on live systems greatly increases operational burden and risks on firms. This type of testing should be administered judiciously.
- Certain requirements around testing may be unsafe to firms and the sector, such as: installing untested software or hardware in production networks and providing regulators or third-party testers with unfettered network access to complete tests.
- Regulators must consider results in the appropriate operational and security technology controls context. Improperly interpreting results without proper context may result in incorrect risk assumptions that may lead to impractical expectations around findings management.
- Detailed results should never be shared, however, regulators can play a critical role in collecting high-level testing results from firms and provide the sector with a thematic view of where sector vulnerabilities reside. Results should not be traceable to institutions and the level of detail shared should be considerate of what could happen should these results fall into the hands of malicious actors.

[8] Firms should not be mandated to outsource testing—this should be a choice that each firm can weigh. Firms may have their own testing teams that are familiar with their firm's environment and be able to quickly pivot to more advanced and useful testing. Firms may also rely on external testing and have already budgeted to do so. Mandating one approach over the other will not produce optimal results and may place an unnecessary financial burden on firms.

- Provide firms with a capability to test their cybersecurity risk and controls framework based on global best practices.
- Establish mutual recognition of tests across multiple locations thus reducing the regulatory burden on firms.

Additionally, testing conducted according to this Framework should provide confidence in the testing process regardless of whether the firm or regulator initiated the testing. Firms should be able to test their systems periodically and leverage such test results to satisfy regulatory requirements in multiple jurisdictions around the world.

## 1.3 Testing Options

The Framework acknowledges that there are many testing types available for firms to assess the effectiveness of their security programs ranging from threat-led penetration testing, application vulnerability scanning to red teaming (details and differences are captured in the appendix). This Framework focuses on threat-led penetration testing due to the high risk of operational disruptions that such testing could cause if undertaken without appropriate planning and involvement from across an organization by a qualified team of practitioners.

Threat-led penetration testing options may range from external / internal infrastructure, web / server-client based applications, third-parties, or software / code analysis, whereas adversary emulation testing seeks to replicate the tactics and techniques of sophisticated threat actors. Testing typically begins with agreement between the firm and the testers to use agreed-upon actions and objectives for attacking target systems and processes inside the firm.

Determining which business processes to be tested can be based on a number of factors but should start with those that are relevant to the target organization. For example, a business-based risk assessment can be used to determine the relevance of a business function or process to the particular organization. A business-based risk assessment may include: (1) the importance of the firm and business process to the stability of financial system; (2) potential business impact of a proposed assessment; (3) the effort required for testing; (4) the size and complexity of what is being tested; and (5) the length of time since an assessment was last performed on that particular business process.

Governance and guidance for firm-led testing:

a. In a firm-led test, the project management and engagement with various other regulators including the agreement of the scope of testing would need to be carried out by the firm with regular and planned inputs / checkpoints from the regulator / supervisors.
b. Firms to utilize the 2nd line of defense to review the scope and attestation before it is signed off on by Senior Management.
c. Firms to have a documented governance model which captures the following points:
    - What techniques are being used and if firms follow any framework for attack?
    - What channels are being tested and processes to prioritize them?
    - What are a firm's technical capabilities and the process they use to analyze and track findings?
    - What do firms do to fix the findings and how are they are tracked?
    - How do firms decide on the timeframes to mitigate these findings?
    - Is the testing team sufficiently independent—in the case of an internal testing team?

d. Senior Management sign off on scope and an attestation that the test was conducted in a transparent and fair method.

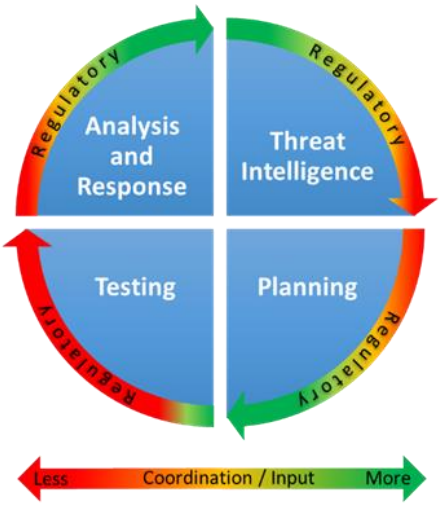
Due to the complicated and potentially disruptive nature of testing, this framework focuses on the necessary steps to successfully leverage this capability to assess a financial firms' cybersecurity defenses.

## 1.4 Testing Lifecycle

Regardless of the type of testing firms choose to perform, they are encouraged to follow a four-phased Testing Lifecycle (see diagram) to ensure they are adhering to industry best practices while simultaneously meeting regulatory demands.

It is important to note that different types of testing may emphasize different aspects of the lifecycle, e.g., TLPT will require a more rigorous threat intelligence phase than application penetration testing. Firms should select their type of testing and include an evaluation of how much time and effort they will spend on each phase to ensure their testing results are the product of a thorough testing lifecycle process.

**Testing Lifecycle**




### 1.4.1 Threat Intelligence Phase

The threat intelligence phase identifies key threats facing the financial sector. Firms, regulators and/or government agencies responsible for cybersecurity should identify and validate key threats which will act as the basis for threat scenarios. These key threat scenarios should be used by individual firms when scoping and executing TLPT.

### 1.4.2 Planning Phase

The planning phase covers test preparation. For regulator-led testing, firms should collaborate with regulators to determine testing expectations, and review business and technical processes to ensure assessment of appropriate systems. Firms will articulate the scope of the test based on their own business processes and the business functions associated with the testing expectations. A business level description of scope will enable the firm to test systems based upon on their own evaluation of that business function. Regulators planning to utilize mutual recognition of tests across multiple locations would need to find the common systems/line of businesses and technologies which they want to test across multiple geographies governed by various regulators. Firms would need to agree with the lead supervisor and supervisors of other locations on the scope of the test.

### 1.4.3 Testing Phase

The testing phase covers test execution within each firm. Firms are responsible for executing tests based upon the scope and the procedures identified in the Planning Phase. Tests performed in accordance with the Framework should give regulators confidence that the executed tests provide quality results and guide firms on how to improve their cybersecurity programs.

### 1.4.4 Analysis and Response Phase

The Analysis and Response Phase covers evaluation of the test results, confirms identified risks, and prepares and tracks firm responses to said risks. As part of this phase, regulators will be provided the ability to review the sensitive test results and responses prepared by the firm in a controlled environment at the firm's premises.

## 1.5 Regulatory Role

Regulatory coordination and input play a key role in regulator-led testing. Firms and regulators should engage in collaborative discussions on testing requirements including scheduling, scope, and goals. When regulators request that financial institutions complete testing activities, both they and the industry should be clear on the scoping of the testing activity, the timing of the testing, the testing requirements and expected test outcomes. Firms should collaborate with regulators to determine the appropriate systems within scope for testing.

Due to the sensitivity of these tests, mutual recognition and acceptance by different jurisdictions of TLPT performed under a common framework would greatly reduce risk and burden to firms, their clients and the sector as a whole. Similarly, mutual recognition would enable authorities to recognize and rely on each other's assessment, producing a common format to review test results promoting regulatory efficiency. It will be the responsibility of the firm to identify the locations in which they want to carry out a mutually recognized test; and then work with their primary authority/supervisor and other relevant authorities to agree on a common scope, geographies and timing. Firms that follow the Framework are likely to produce high quality testing and have appropriate procedures in place to improve their security programs based upon previous test results.

## The Testing Lifecycle

The following sections provide detailed information on each phase of the Testing Lifecycle.

# 2 Threat Intelligence

Threat intelligence-based scenarios mimicking real-life cyber adversaries are essential to the success of testing activities. Threat intelligence scenarios should reflect the most significant risks faced by the financial sector. It is recommended that industry and regulators jointly identify and prioritize the threats in order to develop test scenarios at the financial sector level. This will ensure firms are consistently addressing current and relevant threats. Developing sector-level scenarios benefits the industry by bringing together key stakeholders and subject matter experts to provide a high quality, consistent view of the inherent threats industrywide.


Threat intelligence-based scenarios should consider key financial market participants, including banks, broker-dealers, investment banks, asset managers, financial market infrastructures, financial market utilities, critical third parties and the geographies in which they operate.

## 2.1 Scenario Development

To provide a broader view to the sector, threat intelligence scenarios should be prepared by a financial sector representative group ("Sector Representatives" or "Sector Group") with appropriate financial sector threat intelligence expertise. Appropriate threat intelligence expertise is sourced from financial firms, regulators, Subject Matter Experts (SMEs), Information Sharing and Analysis Centers (ISACs) and government agencies. Output from the Sector Group will include threat intelligence scenarios prioritized by risk to the financial sector that reflect the current threat landscape. Cyber threat scenarios should be detailed enough to provide penetration testing providers with all relevant approaches and information required to formulate effective test plans. For example, if the current threat trends involve payment systems, the financial sector group could create a testing scenario based around payment processes. Government agencies with a responsibility for cybersecurity will have an opportunity to validate the threat scenarios (e.g., National Cyber Security Center in the UK, Netherlands or the Department of the Treasury Office of Cybersecurity and Critical Infrastructure Protection in the U.S.).

In some instances, firms may have their own threat intelligence team that tracks threats that are more tailored to their risk profile and identifies the most significant risks to the firm. Under those circumstances, the findings from a firm's threat intelligence team would be better suited to drive the scenarios used by those firms for testing.

## 2.2 Select and Prioritize Testing Scenarios

A sector group should select and prioritize scenarios in accordance with their evaluation of the respective risk to the financial sector. Example sector groups are the Financial Services Information Sharing and Analysis Center (FS-ISAC), the Critical Infrastructure Partnership Advisory Council (CIPAC - US), the Association of Banks in Singapore (ABS-SG) and the Cyber Cross Market Operational Resilience Group (CCG - UK).

The sector group would identify a set of relevant threat scenarios based on the information sources. The sector group could also identify and define sub-scenarios, extending one or more of these variables, to provide more detailed threat scenarios. The FS-ISAC is producing a bi-annual financial sector specific threat intelligence scenario which meets the suggested requirements of this Framework. Firms and regulators may be able to use the FS-ISAC sector-wide intelligence to inform the scenarios used for their testing purposes and develop firm or country-specific sector threat profiles.

Firms are encouraged to develop multiple scenarios for testing purposes. Scenarios should incorporate various Tactics, Techniques and Procedures (TTPs) reported to be in use by real world adversaries. Testers should be given flexibility, within the bounds of the planned scope of the test, to pursue different attack methods to ensure a realistic testing environment while minimizing operational risk.

## 2.3 Validation

A financial sector group with expertise in threat intelligence, led by the FS-ISAC, should validate any identified threat intelligence scenarios. Once the sector group completes their review, they should create and distribute a sector-wide assessment to financial institutions, government agencies and regulators. We encourage regulators and government agencies to engage in a broader discussion with the sector after they complete their review of the assessment and propose modifications to the sector testing scenarios.

The Framework highlights the need for industry and regulators to have a formal approach for the review and validation process. This approach would help build a consensus around the validated testing scenarios that can then be used at a national level. A successful validation process should define industry participation and engagement.

## 2.4 Maintenance

The threat landscapes change rapidly. For the purposes of a regulatory sponsored security assessment, the financial sector group should re-assess and re-prioritize threat intelligence scenarios at least bi-annually.

# 3 Planning Phase

The Planning Phase covers test preparation activities. Rigorous test preparation is required to ensure operational risks are effectively managed and test objectives are achieved.

The Planning Phase includes project management of testing activities, risk management considerations, and the scope and timing of tests (e.g., threat scenarios, testing options, rules of engagement), and necessary resourcing of testers. This phase is designed to establish the governance structure, prioritization of tests, the scope and related entity/assessment types, and testing scenarios. Targets for testing are identified based on input from threat intelligence, business owners, as well as information and data security criteria. Selection may be based on risk type, prior test results, or other criteria as appropriate.


Organizations should define the purpose and associated limits of the tests and document them in a detailed plan. Technology owners should be involved in selecting the test purpose and limitations and may provide technical information pertaining to the target system such as architecture diagrams, process documentation and technological configurations.

Organizations should establish agreements regarding scope, and assignment of teams and resources during the early stages of the overall process. This includes designating appropriate control and working groups. The working group should provision the scoping documents and background information relating to the functions and systems in-scope.

It is recommended that the control group numbers be kept to a minimum. Members of either the control or working group are reminded to re-affirm their contractual confidentiality obligations. This is to ensure that testing details remain confidential. Minimizing the numbers in the control group and adhering to confidentiality maintains the integrity of the test and provides an excellent opportunity for the organization to test the preparedness of their controls and response teams. Firms may consider having the working group sign a non-disclosure agreement to protect the integrity of the test, if external third parties are engaged.

## 3.1 Project Management

Organizations should use project management best practices during a firm-led TLPT. A project plan should outline the schedule for different testing phases. The plan should be communicated to stakeholder groups, be closely adhered to and should have protocols in place if deviation from the plan

becomes necessary. Since testing may occur in a production or non-production environment, project management planning is essential to avoid and mitigate operational risks.

## 3.2 Risk Management

Risks are inherent during any type of test but are particularly evident in an adversary emulation test. The possibility of causing a denial-of-service incident, unexpected system crash, damage to live critical systems, or the loss, modification, or disclosure of production data highlight the need for active and robust risk management.

To reduce the risks associated with testing, sufficient planning and coordination must take place beforehand. This planning should include agreements between the test subjects and testers on the Rules of Engagement. This would also include scope and, where required, a contractual arrangement covering the testing – including indemnification and liability provisions. Firms should develop a control group during the planning stages to manage the overall risk of the firm. Firms should also conduct thorough due-diligence of in-scope systems prior to any testing to ensure that backups systems are in place, and recovery procedures are up to date and have been recently tested. A rules of engagement (ROE) document should be created that clearly defines the testing parameters, approvals, and escalations (see Section 3.6 for more detail relating to ROE).

Protecting the confidentiality of the test is crucial to its effectiveness. To that end, the firm should limit awareness of the test to a small trusted group whose members have the appropriate levels of seniority to make risk-based decisions regarding the test.

The firm should clearly define which measures are to be taken to ensure that only selected members are informed about the test (e.g., members may sign a non-disclosure agreement (NDA) to ensure their confidentiality throughout the test). The firm should also agree on escalation procedures to avoid the triggering of actions that would be mandatory in the case of a real event. Such actions include communicating with an external party (e.g., declaring an incident to a computer security incident response team, sharing information on a platform, etc.) or calling law enforcement.

Risks should be carefully managed throughout the Planning Phase. Higher risk activities can be managed by conducting tests off-hours, or against non-production systems (however, testing parties should ensure these systems are operating with the same parameters of the production systems to ensure a valid test).

The standard language and structure requirements for contracts with TLPT vendors should include a requirement for third-party testers to meet security and confidentiality requirements at least as stringent as those followed by the underlying institution regarding confidential information handling, including PII (i.e., personally identifiable information). The contracts should also include a clause related to data destruction requirements and breach notification provisions.

## 3.3 Scoping

The scope of a test should focus on a firm's business functions, prioritizing testing on those functions based upon their criticality to the firm and to the larger sector and designed in a way that ensures the testing activity is completed within an agreed-upon timeframe. Communication between a firm and its regulator should be clear as to the scope of the test for the chosen function focused on the business

processes of the firm and the criticality to the sector. Firms should determine the test scope and share the test parameters with regulators.

During the Scoping Phase, stakeholders, including the threat-led penetration testing providers, should reach agreement on the classification schema for vulnerabilities discovered during testing, as well as on objectives that demonstrate successful compromise of the firm. The classification schema aims to show criticality and priority of discovered vulnerabilities in line with the firm's risk management framework.

Firms may apply risk and firm-based criticality ratings to in-scope systems, infrastructure, applications, functions, and data. Firms should examine these assets to effectively select the final target set. Financial institutions should create a scoping document that allows for a consensus between firms and regulators.

The scope of the testing should be based on the threat intelligence scenarios and key business processes and could range beyond traditional endpoint and network level testing to include physical threats or threats from third parties / suppliers. Threat intelligence is essential to establish and maintain a library of testing scenarios. The agreed-upon key business processes tested should be matched to one or more of the sector-level identified threat scenarios.

Testing Phase (Section 4) objectives should be defined and linked to the threat scenarios previously identified. Mapping objectives to threat intelligence scenarios will allow firms to identify successful outcomes and calculate actual impact to firms. This in turn will help drive prioritization of remediation activities.

## 3.4 Testing Options

As mentioned previously, this Framework acknowledges that there are many testing types ranging from application and/or network penetration testing to Threat-led penetration testing. Different types of testing present different risks to different systems, such as in production systems and those emulated in testing / development environments. Firms should understand the benefits of and risks associated with the different types of testing and choose the method that best suits their needs in evaluating their cybersecurity programs.

## 3.5 Timing of Tests

Firms and regulators should determine the appropriate timing of test execution based upon the type of test, the criticality of the system or process to be tested, the potential for operational disruption caused by the testing and how recently the systems or processes have been tested. Frequency of testing for operationally sensitive systems should be considered carefully to weigh the value of testing against the risk of any potential disruption. It is also important to note that tests like TLPT are far more resource intensive from a time, personnel, and expense aspect. Regulators should try to be mindful of other TLPT requirements imposed on a single firm. Mutual recognition of testing frameworks seeks to reduce the demand on firm and regulatory resources.

To maximize the utility and scalability of the regulatory use of threat-led penetration testing, a standard schedule is recommended that allows both regulators and firms to adequately plan and execute the necessary type and scope of testing. Firms are often planning their testing activities late in the year for execution the following year. One proposed schedule could be as follows:

**Month 1**: Regulators and Firms, through bi-annual reports provided by the FS-ISAC, engage and agree on the most critical identified threats, vulnerabilities and adversary tactics.

**Month 2-Month 3**: Regulators and Firms meet to discuss Firm-planned testing for the following year in light of institution needs and current threat intelligence.

**Month 4**: Regulators, based upon threat intelligence received and input from Firms, provide guidelines for systems to be tested and scope of testing.

**Month 5**: Regulators provide a schedule for testing to be performed in order for results of previous years testing and current threat intelligence to guide the following testing cycles.

**Month 5- Month 11**: Firms perform testing during this time period.

## 3.6 Rules of Engagement (ROE)

The ROE are an agreement between the firm and testers to provide assurance that tests and associated risks are being managed in a controlled manner. This planning document is essential in testing and should be agreed upon between all parties involved. An effective ROE outlines at a minimum the following:

- Technical Scope and Span to be Tested:
  - Functions
  - People
  - Firm assets
  - Specific out-of-scope areas
  - Specific out-of-scope periods of time
  - Activities to be conducted during testing
  - Locations
  - Specific Technologies
  - Third parties / suppliers
- Administrative Details:
  - Points of contact – trusted agent(s), testers, and leadership
  - Testing guidelines
  - Projected timelines
  - Communication Plan – guidelines between trusted agents and testers
  - Reporting mechanism / frequency
  - Acceptance of liabilities, responsibilities, and risks
  - Test termination criteria

The ROE are typically captured in a project document that may be altered during the course of the test based on approval from control group(s) and the tester(s). The ROE's fluidity is essential to maintain testing flexibility and allows the tester to remain within the scope and boundaries of the ROE, and at the same time practice best judgment when examining the environment. The scope and boundaries of the ROE should only be altered with the permission of the firm.

## 3.7 Resourcing / Qualifications

Testers, both individuals and teams, must: (1) possess a minimum level of expertise, measured level of experience or numerous certification criteria established by international standard-setting bodies, such as CREST, Offensive Security, GIAC et al; (2) gain qualification through a vetted industry third-party due

diligence processes; and (3) be guided by a strict code of conduct. The aforementioned accreditation standards are globally recognized and should be acceptable in determining the competency of testing personnel. However, other recognized accreditation options with similarly rigorous standards may be considered as well. Firms should demonstrate and verify to regulators that their testers and teams meet these baseline standards.

The Framework suggests the use of internal resources for security assessments, and particularly for adversary emulation testing if the accreditation requirements are met. Furthermore, organizations should grant internal resources the appropriate levels of independence and oversight.

The Framework suggests the creation of a commonly accepted accreditation standard. Adopting an industry-wide accreditation standard, enables firms and regulators to:

- Establish a set of practices that defines and prescribes strict adherence to levels of competence, skill, experience, and knowledge requirements for practitioners at different levels (e.g., foundational, moderate, expert)
- Accredit a testing organization for their ethics, reputation, credibility and delivery oversight, covering the methods for assessment, risk management, quality assurance, provision of ongoing training and development, innovation, and research
- Ensure teams can provide adequate information protection, operate within appropriate testing facilities, use an approved suite of tools, and have adequate de-confliction procedures
- Establish clear criteria for trustworthiness of results through meeting professional standards, obtaining relevant qualifications, and adopting a Code of Conduct committing to professional and ethical responsibilities
- Establish certification criteria specifically as it relates to third-party vendors
- Ensure that legal and regulatory requirements of the firm are met when testers receive access to client data during testing (e.g., cross-border requirements)

The application of an internationally agreed upon accreditation standard to both internal and third-party teams would allow for multiple regulators to benchmark internal teams against peers with consideration for firm size, scale, business, and risk profile, and against third-party security assessment organizations. A key goal of the Framework is to establish confidence in a testing team's (internal or third party) ability to deliver the levels of assurance in the execution of regulator-driven security assessments and particularly adversary emulation testing activities.

# 4 Testing Phase

The Testing Phase brings together Industry Threat Intelligence, testing scenarios and financial institutions' scope to deliver a practical assessment of defensive security controls and detection and response capabilities. During testing, the Tactics, Techniques and Procedures (TTPs) are representative of the threat actors within a threat landscape and will be used to deliver a realistic simulation.

The activities within the Testing Phase are:

- **Operational Planning**
- **Execution**
- **Review**


The time allocated for testing is determined by the scope and resources of the financial institution, any external requirements for a given engagement, and availability of supporting information supplied by the financial institution (e.g., regular vulnerability reports or previous assessment data). A test could have limitations that would not apply to genuine threat actors and therefore limitations influencing adversary emulation testing should be considered.

The Testing Phase will involve specialist security testers experienced in executing adversary emulation exercises, which may exist either internally within the financial institution or as part of a third-party engagement with a specialist security consultancy firm. Specialist accreditations are available for both internal and external professionals as discussed in Section 3.7 of this document.

## 4.1 Operational Planning

Operational Planning includes the process of ensuring the outputs from the threat intelligence and wider Planning Phase drive test execution. Furthermore, the Framework suggests that the ownership of Operational Planning should reside within the firm. Appropriate oversight should be provided by the firms' governance, control and working groups.

Inputs, activities, and outputs of the Operational Planning phase for adversary emulation testing are detailed below:

- **Inputs**: Industry Threat Intelligence sources, vetted scope, and developed scenarios
- **Activities**: Test preparation and firm specific intelligence gathering

- **Outputs**: Targeting Report[9], Threat-led penetration Testing Plan, and possible modifications to Rules of Engagement

The following testing elements are an example of what may be examined during the Operational Planning phase:

**Information Gathering:** Reconnaissance against a firm to discover publicly available information such as company email addresses, document metadata, IP addresses, and domain name servers.

**Enumeration:** Collection of data related to hosts/servers on the network. Some of the data that may be discovered includes users, ports, running services, and operating systems.

**Vulnerability Identification:** Enumeration findings are examined to discover vulnerabilities that may be exploited by the tester. Manual analysis as well as automated scanning tools may be used to identify network vulnerabilities.

**Exploit:** Defined way to breach the security of information systems through vulnerability.

**Compromise:** Violation of the security of an information system.

**Post-Exploitation:** After gaining unauthorized access to a host/server, the tester aims to escalate privileges (if the tester is using a low privileged account) and pivot to other hosts/servers to steal more information.

**De-chain:** A point in the test execution where a scenario is artificially progressed to compensate for time limitations or replicate control failures (e.g., if a phishing exercise has not resulted in compromise of a system within a given time frame, the testers may be given access to allow the testing to progress).

As a verification of the testing planning phase it should be possible to map all of the testing activities to the threat scenarios developed during the Threat Intelligence Phase (Scenario Development Section 2.1).

[9] Targeting consists of industry level threat intelligence developed scenarios, scope, and reconnaissance [e.g., open-source intelligence (OSINT), social-media intelligence (SOCINT), passive and/or active information gathering or enumeration]. The goal of targeting is to utilize this information to formulate a targeting report that adheres to the validated scenarios and focuses on systems that may provide a cyber advantage (e.g., a list of potential targets for phishing attacks, or previously publicized vulnerabilities of in-scope assets). Targeting will be undertaken by the testing team prior to executing the penetration test.

## 4.2 Execution

Threat-led penetration testing and, in particular, red-team testing utilizes numerous techniques and approaches. This document will not mandate or describe a specific methodology. However, experienced testers are expected to replicate the modus-operandi of threat actors and their TTP's according to the scenarios selected for testing.

The key inputs, activities and outputs of the Execution Phase are outlined below:

- **Inputs**: Targeting Report, Test Plan, Rules of Engagement
- **Activities**: Scenario driven testing as detailed in the test plan
- **Outputs**: Initial testing results

In this phase, testers are positioned to test against a targeted business process within the agreed-upon timeframe. The Targeting Report and results of the testing will drive the next test activity.

Regular updates on progress will be provided to trusted agents as detailed in the Communications Plan, and in accordance with the ROE.

Test execution can be enhanced using results from similar activities, which may be undertaken as part of a business-as-usual process within the financial institution; for example, if monthly vulnerability assessments are undertaken, this data can feed into the test execution process without needing to repeat previous test elements.

In testing the agreed scenarios, there may be common stages throughout that do not need to be replicated each time. For example, if more than one scenario suggests phishing as an attack vector, this only needs to be undertaken once. There may be common themes (see table below) across the scenarios, and utilizing previous results ensures the most effective use of available time and resources within a testing cycle.

**Infiltration:** Gaining access to and persistence on systems within the scope of the test; this could occur via social engineering, exploiting vulnerabilities, or by taking advantage of misconfigurations.

**Lateral Movement:** Moving around the network to identify key assets within the scope of the test, building up information of the financial institution's systems and processes.

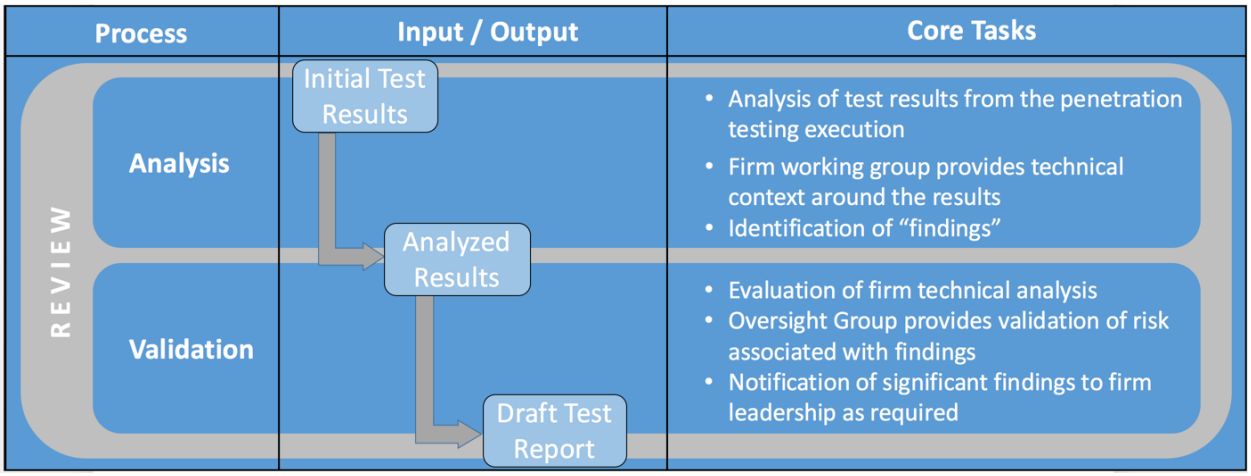
**Action on Objective:** The successful execution of an attack which could result in user / privileged-level access, sensitive information leakage, or other follow-on malicious activity.

Upon completion of execution, initial details such as host information, vulnerability, exploitability, impact, recommended actions for mitigation / prevention, and follow-up actions should be recorded as inputs into the reporting cycle.

## 4.3 Review

During the Review portion of the Testing Phase, the initial results from the testing execution should be reviewed jointly by the testing team and the financial institution control groups articulated in the Planning Phase (Section 3).

The inputs, outputs and core tasks are summarized below:




Firms should provide additional context to properly apply a risk rating for each finding (technically verified vulnerability associated with an in-scope target). A "significant finding" is a finding that potentially possesses a substantial damaging effect of the firm's security, financial, reputational or legal posture. A draft test report of initial findings should be provided to each appropriate audience (e.g., technical findings are provided to the team that will take appropriate action to resolve the findings).

Distribution and discussions of initial findings will be detailed as part of the engagement. If a third party conducts the test, it may provide out-brief workshops or walk-throughs of the test findings / results for validation. The oversight (control) group will validate and accept these findings. Many organizations have found value in not only reviewing the initial findings via a replay of the test engaging both the red and blue teams (referred to as purple teaming) to walk through the testing procedures and articulate how the testing progressed to identify the strengths and weaknesses of the firm's defense and enhance the quality of future tests.

Any significant finding must be communicated as described within the ROE to ensure risk exposure can be managed at or close to the time of the finding being discovered. Firms have the responsibility to share any significant findings with other firms which might be impacted with the same vulnerability that may have been discovered during testing. At the time of review, these findings may have already been remediated. All findings revealed at the time of the test should be captured in the draft Test Report, including any remediation efforts or actions taken by the firm related to the finding.

Tests may identify areas outside of scope that cannot be explored within the agreed-upon timeframe. These potential targets of opportunity should be recommended for review as part of an additional work-stream or follow-on testing.

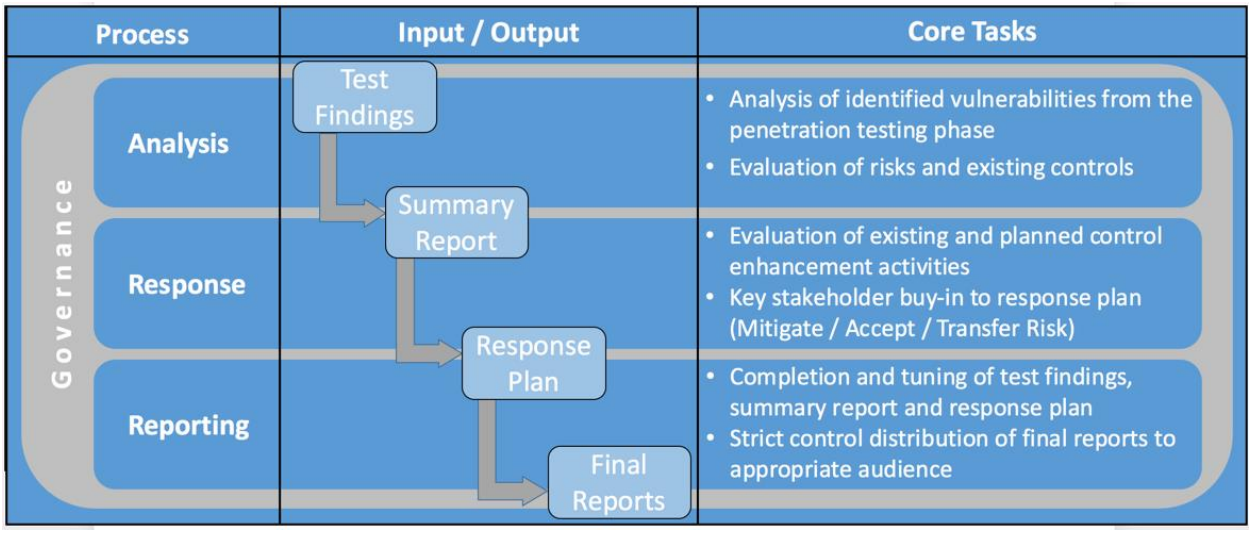
The draft results from the review phase will form the core elements of the Analysis and Response Phase (Section 5).

# 5 Analysis and Response Phase

The Analysis and Response Phase covers analysis of current test results, the appropriate firm response actions to the test, and protocols for reporting testing results. IT firm management and senior manager involvement is crucial, where appropriate, at different phases, particularly throughout the Analysis and Response portions. Relevant stakeholders should be involved during the acceptance, mitigation or transfer of the risks uncovered during the Testing Phase.

IT and risk management personnel should be involved in developing the appropriate responses, which may involve allocating firm resources and prioritization of remediation efforts.


The graphic below represents a typical process flow of the Analysis and Response Phase. Core tasks, as well as common input and output documents are defined for each process. "Governance" represents oversight provided by firm leadership up to and including the Board when appropriate throughout the entire phase.




## 5.1 Analysis

The test findings represent potential weaknesses or vulnerabilities which will be evaluated by firms to establish the actual level of risk they pose. In the Analysis section, firms should utilize existing controls to determine the actual level of risk to the firm. This risk should be described using firms' internal risk

definitions and framework. Based on the level of actual risk, firms may determine that additional controls are necessary.

After analysis is complete, a Test Summary should be prepared, which includes a summary view of the test activity, categorized test findings and risks aligned to existing mitigating controls. The Framework suggests the usage of the Financial Services Sector Profile[10] of the NIST Cyber Security Framework for synchronizing the language used by firms when categorizing test findings.

Often an evaluation of the test by all stakeholders to capture feedback: testers, oversight team and defenders, will provide valuable insight on the efficacy of the testing and how to improve testing procedures in the future. This evaluation should be documented by the firm and used to guide any future testing.

## 5.2 Response

During the Response process, firms should utilize the Summary Report from the Analysis process to consider actionable responses to the identified risks. These responses may include the enhancement of existing controls, the development of new controls, or the update of application code to fix the identified issue. The control enhancements for the identified risks should be reviewed and approved by appropriate senior stakeholders in the firm to ensure organizational accountability.

Risks identified by tests and the firm's response should be coordinated with senior technology leaders, operational leaders and risk committees. They should provide oversight for the identified control enhancements and ensure that they are fulfilled. Firm's operational risk oversight and tracking mechanisms should be used.

Firms should follow accepted protection guidelines when managing data that contains vulnerability information. Priorities may be formulated during this phase based on the risk level of the findings. For instance, priorities may be defined based on the risk levels (e.g., critical, high, medium, and low).

A Response Plan should be prepared and encompass the identified risks from the Summary Report and actions to be taken by the firm, including the evaluation of existing controls and planned control enhancements with target dates.

## 5.3 Notification

Determinations regarding what findings are presented to executive management should be made consistent with documented policies within the firm for alerting management to risks. It is important that findings be brought to the attention of firm management consistent with the firm's overall risk reporting procedures.

[10] The Financial Services Sector Profile of the NIST Cybersecurity Framework was developed by the U.S. Financial Services Sector Coordinating Council to act as a harmonizing element for the financial services sector members as well as financial services regulators. The Profile provides a single vocabulary, taxonomy and approach for managing cyber risk within the financial services sector. It can act as a common link for organizations to articulate their cybersecurity risk management program with other firms and their regulators and if used as a basis for financial supervision can coordinate regulatory activity across jurisdictions leading to sector wide improvement in risk management as well as broad efficiencies in supervision.

## 5.4 Reporting

Final reports will be produced from a combination of test findings, summary reports, and response plans. The key final reports should be at the following levels:

- Executive Level - Senior stakeholders in firms should be provided with Test Summaries and Response Plans. Executive management should receive a high-level report that outlines the threatened assets, impacts to the organization, and recommended remediation actions.
- Regulatory Level - Regulators will be provided access to a summarized version of the Test Response Plan and a summary of test results, which will include key risks identified and firm responses.
- Technical Level - The most detailed / technical versions of the report should be reviewed with the firm's IT security team on firm premises and not leave control of the firm.

## 5.5 Data Protection

Testing information and reports must be protected with appropriate technical, physical and administrative safeguards during transit and at rest in line with a firm's own processes and procedures for handling confidential data. Test reports should be sent and stored in a manner that protects their confidentiality and integrity (e.g., using encrypted email, secure files transfer mechanisms, use of multi-factor authentication).

## 5.6 Distribution

Due to the sensitive nature of the data, test reports should be distributed on a need-to-know basis, strictly controlled and appropriately protected. Distribution of reports outside of the firm should happen infrequently and highly sensitive data should not leave the firm at all. Regulators should be provided reasonable access to review such sensitive data within the firms' control.

Specific details about vulnerabilities and systems affected in the test report could cause significant damage and follow-on information security concerns and/or reputational harm. Detailed results are particularly sensitive and should be very tightly controlled to internal review.

General findings and especially best practices the testing have identified may be shared as appropriate for the benefit of the sector.

# 6 Conclusion

This Framework is intended to serve as a high-level directional guide for both firms and regulators during firm-led or regulator-driven Threat-led penetration testing. The four-phased Testing Lifecycle provides assurance of quality by firm designated testers and encourages clear communication amongst the regulatory community and financial firms.

While the goal of this Framework is to promote a better understanding of the requirements and expectations of both firms and regulators, it also assists the industry in reducing risk by defining foundational processes globally. The reduction in risk extends beyond the firms; as clients, shareholders and investors will have an understanding that the industry and regulators work together via a common framework.

From a regulatory viewpoint, it suggests timely collaboration from regulators during each phase of the threat-led penetration testing process. This Framework encourages regulatory involvement during the threat scenarios development phase, as well as Test Plan and Response Plan reviews. Involvement during key milestones in the process ensures that firm-led Threat-led penetration tests have met industry and global standards.

As cyber testing continues to evolve, there may be a need for a more in-depth playbook that examines the details surrounding the testing process. Such details could provide more uniformity among the firms and standardize how the testing process is conducted throughout the financial industry.

There is additional work in ensuring these evaluation tools are used as effectively and safely as possible. As powerful as these tests are, they provide a view of vulnerabilities for a given moment of time, technology and enhanced processes will likely develop methods to continuously monitor the effectiveness of controls. As these methods improve and mature the financial services sector must be ready to develop industry best practices to ensure the continued safety and security of our critical infrastructure by utilizing such continuous control monitoring.

# 7 Glossary: key cyber terms aligned with the Financial Stability Board Lexicon[11].

| Term | Abbreviation | Description |
|---|---|---|
| **Threat-Led Penetration Testing (TLPT) also known as** Adversary Emulation or Red Teaming | N/A | 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. |
| Threat Intelligence | TI | Threat information that has been aggregated, transformed, analyzed, interpreted or enriched to provide the necessary context for decision-making processes. |
| National Institute of Standards and Technology | NIST | An organization of the United States Commerce Department that develops and facilitates the development of standards for the industry. |
| Trusted Agent | TA | A person in the firm responsible for communicating with the tester while maintaining confidentiality of the ongoing test. Topics for discussion between the trusted agent and tester include changing the attack methodology, evidence of past or ongoing breach, and status updates. |
| Open-Source Intelligence | OSINT | A type of intelligence wherein information is gathered from publicly available sources. |
| Social-Media Intelligence | SOCINT | A type of intelligence wherein information is gathered from popular social media sites. |
| Test Findings | N/A | Confirmed vulnerabilities to be presented in a report that outlines their severity based on risk and impact to the firm. |
| Vulnerability | N/A | A weakness, susceptibility or flaw of an *asset* or control that can be exploited by one or more threats. |

[11] http://www.fsb.org/wp-content/uploads/P121118-1.pdf

## Appendix: Difference between Vulnerability Assessment, Pen Testing, Red Teaming and Threat-led Penetration Testing

| | Vulnerability Assessment | Pen Testing | Red Teaming (RT) | Threat Lead Penetration Testing |
|---|---|---|---|---|
| **Objective** | Broad scanning for vulnerabilities or information gathering | Compliance finding as many vulnerabilities as possible | Improving the cyber resilience of the entity by testing not only preventative measures but also detective and reactive controls. Helps to train the blue team and is led by TLPT | Objective is similar to red teaming |
| **Techniques** | Automated software identifying publicly known vulnerabilities | Specialized team of pen testers often without scenarios testing specific systems | Specialized teams with skills emulating (relevant) threat actors with scenarios | Techniques are simulated that are likely by threat actors and intelligence |
| **Scope** | Single system | Specific application system or business practice | Whole entity including people, processes, and technology | Like Red Teaming, everything is in scope |
| **Costs** | Low | Medium | High to very high | Like RT high to very high |
| **Given Knowledge of Entity** | High | Medium to high | None to medium | Like RT none to medium |
| **Vulnerabilities Action** | List them for client | Exploit them (and chain them) | Only exploit what is necessary to reach goal as discretely as possible | Like RT only what is necessary as discreetly as possible |
| **Length** | Day | 1-2 weeks | 6-12 weeks | 10-12 weeks |
| **Threat Actor Simulation** | None | Possible, but partial at best | Specific threat actor, but mostly APT simulation | High end threat actor/APT simulation |
| **Reporting** | Mostly automatic technical generated report | Technical report focuses on software vulnerabilities | Tactical and strategic management report against business processes and additional technical reporting about vulnerabilities identified and how to mitigate | Same as RT |