ARTIFACT SECURITY

Advanced Persistent Testing Methodology

Document identifier	Version	Date of issue	Supersedes
APT2025v1.1	**1.1**	**2025-02-11**	v1, APT2025v1.0

Framework alignment	Attack sequences
MITRE ATT&CK, MITRE D3fend	Intrusion, Infiltration, Propagation

Issuing lab

Artifact Security

Contents

Clause	Title	Test IDs	Page
1	Goals		3
2	Attack Test Corpus		3
3	Environment		3
4	Execution		4
4.1	Red Steps		4
4.2	Blue Responses		6
4.3	Example Test Case Appendix		6
5	Rating		8
5.1	Attack Rating: Detection		8
5.2	Attack Rating: Protection		9
5.3	Attack Rating: Protection capabilities		10
5.4	False Positives		11
5.5	Total Rating		12
6	Configuration disclosure		12
7	Change Log		13

01 Goals

This evaluation aims to test security solutions' response to advanced attack methods. The methodology considers both real attacks and options available to defenders to cope with the attacker. This methodology can be applied against enterprise detection and response solutions.

02 Attack Test Corpus

The attack test corpus is tailored for an overarching goal. This is outlined in the test scheme. It can be a generic goal such as deploying ransomware or replicating the behaviour of a specific real-life APT.

03 Environment

The environment is hosted using Microsoft Azure cloud services. Where necessary, alternative hosting can be provided to ensure compatibility. This will be noted in any accompanying report.

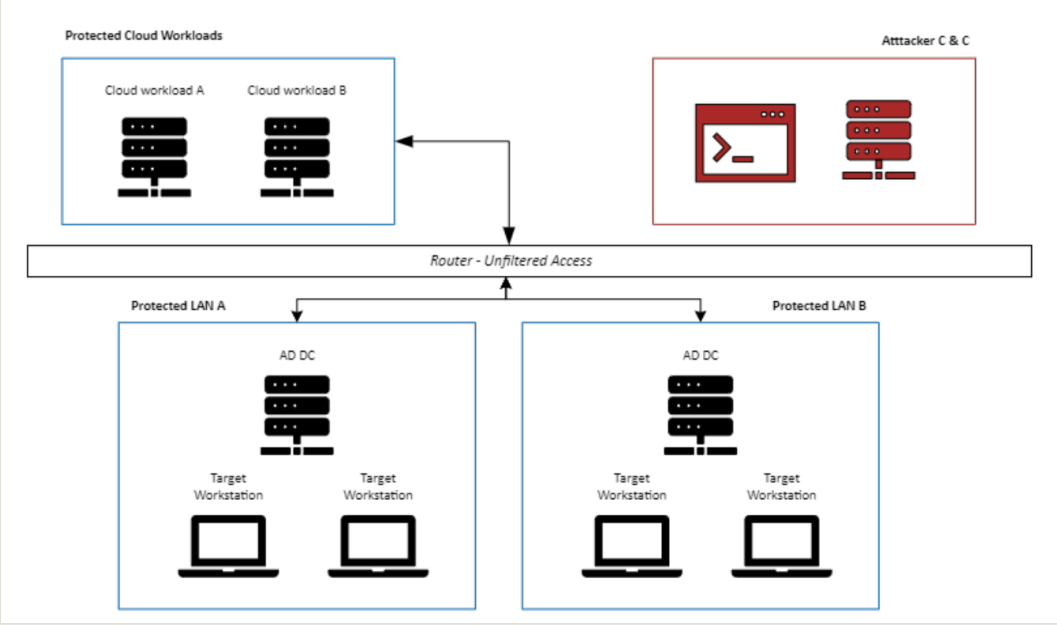

Figure 1. Catch-all environment layout.

The target environment can vary to adapt to the test scheme. The diagram above is a catch-all layout. Exact versions of operating systems are specified in the test schemes and each accompanying report.

04 Execution

4.1 Red Steps

Attackers' steps taken towards their goal are described under red actions. The techniques used here are linked to the MITRE ATT&CK framework.

Sequence 1: Intrusion

This sequence is defined by the initial delivery mechanisms employed by the attacker against the target organisation.

ATT&CK Tactics applicable: Initial Access

Table 1 Sequence 1: Intrusion

Sequence	Initial Access
Intrusion	T1133: External Remote Services
	T1190: Exploit Public Facing Application
	T1566: Phishing

Sequence 2: Infiltration

This sequence is defined by the attacker executing and taking actions on the initial target.

ATT&CK Tactics applicable: Execution, Persistence, Privilege Escalation, Defense Evasion, Credential Access

Table 2 **Sequence 2: Infiltration**

Sequence	Execution	Persistence	Privilege Escalation	Defense Evasion
Infiltration	T1204: User Execution	T1098.001: Additional Cloud Credentials	T1098.003: Additional Cloud Roles	T1027.002: Software Packing
	T1059.001: Command and Scripting Interpreter: PowerShell	T1098.003: Additional Cloud Roles	T1098.005: Account Manipulation: Device Registration	T1070: Indicator Removal
	T1047: Windows Management Instrumentation	T1133: External Remote Services	T1068: Exploitation for Privilege Escalation	

Sequence 3: Propagation

This sequence is defined by the attacker progressing past the first intrusion.

ATT&CK Tactics applicable: Discovery, Lateral Movement, Collection, Exfiltration, Impact

Table 3 **Sequence 3: Propagation**

Sequence	Discovery	Lateral Movement	Collection	Exfiltration
Propagation	T1087: Account Discovery	T1021.007: Cloud Services	T1560.001: Archive via Utility	T1048.002: Exfiltration Over Asymmetric Encrypted Non-C2 Protocol
	T1482: Domain trust Discovery	T1021.001: Remote Desktop Protocol	T1213.003: Code Repositories	
	T1135: Network Share Discovery	T1021.002: SMB/Windows Admin Shares	T1074.002: Remote Data Staging	

4.2 Blue Responses

Available responses to each phase of the attack are described under blue responses. The options given to defenders are assessed. Desired outcomes are prescribed alongside the red objectives when the test scheme is built.

Where possible, mitigations are tied to the corresponding ID in the ATT&CK framework. A technique can be employed a multitude of ways therefore the mitigation chosen by the evaluated solution must be appropriate to how the technique is employed by the tradecraft chosen by the attacker. Where possible MITRE's D3fend taxonomy is used to describe the defenders' options.

The rating system is applied differently for solutions configured solely for detection compared to those focused on protection testing as described below.

4.3 Example Test Case Appendix

Goal Gain initial foothold, move laterally and exfiltrate sensitive data

Table 4 **Example test case: red actions and blue responses**

Sequence	Detection		Protection	
	Red Actions	Blue Responses	Red Actions	Blue Responses
Sequence 1: *Intrusion*	T1566.001 Phishing: Spearphishing Attachment	DS0029 Network Traffic Content DS0022 File: File Creation DS0015 Application Log: Application Log content	T1566.001 Phishing: Spearphishing Attachment	M1049 Antivirus/Antimalware – automatic quarantine
Sequence 2: *Infiltration*	T1204.002 User Execution: Malicious File	DS0022 File: File Creation DS0009 Process: Process Creation	T1204.002 User Execution: Malicious File	M1040 Behaviour Prevention on Endpoint M1038 Execution Prevention M1017 User Training (Out of Scope) M1049 Antivirus/Antimalware
	T1059.003 Command Scripting Interpreter: Windows Command Shell	DS0017 Command: Command Execution DS0009 Process: Process Creation	T1059.003 Command Scripting Interpreter: Windows Command Shell	
	T1027.002 Obfuscated Files or Information: Software Packing	DS0022 File: File Metadata	T1027.002 Obfuscated Files or Information: Software Packing	

Sequence	Detection		Protection	
	Red Actions	Blue Responses	Red Actions	Blue Responses
	T1057 Process Discovery	DS0017 Command: Command Execution DS0009 Process: Process Creation	T1057 Process Discovery	
	T1082 System Information Discovery	DS0017 Command: Command Execution DS0009 Process: Process Creation OS API Execution	T1082 System Information Discovery	
Sequence 3: *Propagation*	T1135 Network Share Discovery	DS0017 Command: Command Execution DS0009 Process: Process Creation OS API Execution	T1135 Network Share Discovery	M1208 Operating System Configuration M1037 Filter Network Traffic M1035 Limit Access to Resource Over Network M1027 Password Policies M1026 Privileged Account Management M1057 Data Loss Prevention
	T1021.002 Remote Services: SMB/Windows Admin Shares	DS0033 Network Share: Network Share Access DS0029 Network Traffic: Network Connection Creation Network Traffic Flow DS0009 Process: Process Creation	T1021.002 Remote Services: SMB/Windows Admin Shares	
	T1083 File and Directory Discovery	DS0017 Command: Command Execution DS0009 Process: Process Creation OS API Execution	T1083 File and Directory Discovery	
	T1039 Data from Network Shared Drive	DS0033 Network Share: Network Share Access DS0022 File: File Access DS0029 Network Traffic: Network Connection Creation	T1039 Data from Network Shared Drive	
	T1005 Data from Local System	DS0022 File: File Access DS0012 Script: Script Execution DS0017 Command: Command Execution	T1005 Data from Local System	
	T1560.002 Archive Collected Data: Archive via Library	DS0012 Script: Script Execution DS0022 File: File Creation	T1560.002 Archive Collected Data: Archive via Library	

05 Rating

5.1 Attack Rating: Detection

Ratings are given per attack sequence. While direct references to the corresponding ATT&CK technique are useful they are not required to earn detection points. Data sources are marked and mapped to the corresponding source but do not influence the final accuracy rating. Within each defined tactic scope of a test case the solution must detect 50% of the techniques used to earn the maximum grade. For example:

Table 5 Detection rating example

Sequence	Technique	Tactic	Min # Ts	Rating
Intrusion	T1566.001	Initial Access (1)	1	+10
Infiltration	T1204 T1059.001 T1047	Execution (3)	1	+10
	T1053.005 T1078.002 T1078.003 T1078.004 T1547.001	Persistence (5)	2	
	T1053.005 T1078.002 T1078.003 T1078.004	Privilege Escalation (4)	2	
	T1070.004 T1070.006 T1036.005 T1140 T1484.002 T1562.004 T1562.001 T1562.002 T1070.008	Defense Evasion (9)	4	
	T1555.003 T1606.001 T1003.006 T1558.003 T1539	Credential Access (5)	2	

Sequence	Technique	Tactic	Min # Ts	Rating
Propagation	T1087.002 T1482 T1018	Discovery (3)	1	+10
	T1021.006	Lateral Movement (1)	1	
	T1005 T1114.002 T1560.001 T1213.003 T1074.002	Collection (5)	2	
	T1048.002	Exfiltration (1)	1	

5.2 Attack Rating: Protection

Solutions configured for protection are expected to stop the attacker as early in the attack chain as possible. The responses given by the solution are matched to the mitigations available for the techniques executed. Custom and bespoke responses available to a responder will be awarded with wildcards and where possible will be communicated in the final report.

A full attack chain is executed to measure the protection capabilities of the solution. Where necessary, the attack will revert and resume from the last possible sequence to showcase the capabilities of the solution when one sequence successfully bypasses the solution.

For example:

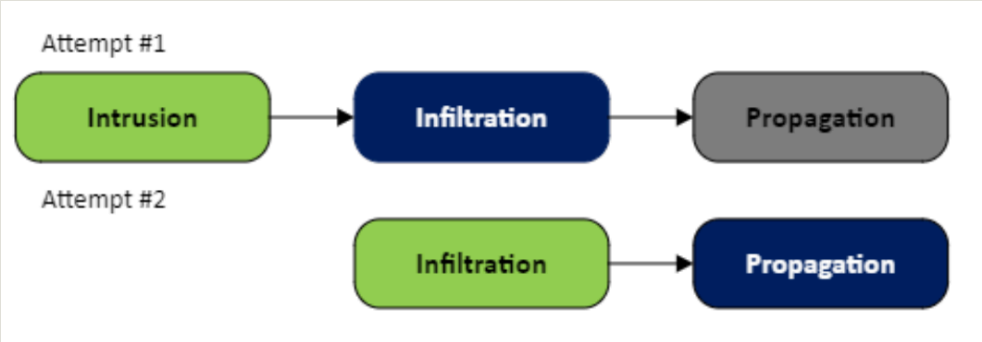

Figure 2. Attack attempts resuming from the last possible sequence.

In Attempt #1 the solution can prevent during the Infiltration sequence not allowing Propagation to be tested. To further showcase the capabilities of the tested solution later in the attack chain the attack will re-start from the Infiltration stage as in Attempt #2.

Efficacy score – the solution's ability to prevent all sequences of the attack.

Efficacy score is calculated as follows. Each successful sequence execution leads to a potential loss of points:

Table 6 **Efficacy score**

Sequence	Intrusion	Infiltration	Propagation
Score	+10/-10	+10/-10	+10/-10
Note	-	Lack of remediation of significant traces of the attack earn +5	

5.3 Attack Rating: Protection capabilities

The protection capabilities will be assessed under the following categories.

Breadth of options – measure the variety of responses during the attack sequence.

> *Example: File Quarantine, Network Isolation, Block Domain or IP*

Depth of options – how specific the responses during the attack sequence are.

> *Example: Assuming a Detection of TrojanXYZ on file ABC.exe at <path>. Action to be taken based on file hash/name or path?*

Detection Engineering/Customisability – how well the response options can empower the responders to improve their future detection.

Proactive vs Reactive – percentage split of proactive and reactive mitigations. Across the whole test what percentage of Proactive and Reactive responses were employed.

5.4 False Positives

Legitimate scenarios are crafted to reflect common behaviour in an organisation. These are provided during the deployment phase of the product. The legitimate rating considers the need of baseline behaviour when necessary. Tested solutions have the right to request a learning period about the environment of up to two weeks. Modern solutions may consider each identity behaviour in the context of the organisation and build the base configuration before the test execution starts. Any learned behaviour that affects the standard practices of the solution is reflected in the final report.

Table 7 Rating

Severity Level	Unknown behaviour	Learned/Established Behaviour
Informational/No/Low Priority	+10	+10
Amber/Medium	+5	-5
Red/High	0	-10
Configuration change required (modifier)		+5

The configuration change modifier can be applied to any sub-optimal rating achieved.

5.5 Total Rating

The total rating for each category is reflected in the accompanying report based on the following thresholds:

Table 8 **Grade thresholds**

Grade	Threshold
S	91% - 100%
A	81% - 90%
B	71% - 80%
C	61% - 70%
D	51% - 60%

The pre-defined categories that are combined for a total accuracy rating:

- Detection
- Protection Efficacy
- False Positive

The protection capabilities rating is given as a separate grade alongside Total Rating.

06 Configuration disclosure

A full configuration and licencing disclosure is taken as part of any public report. If possible, this will be hosted under Artifact Security website. Linked references to the tested vendor resource are also acceptable.

07 Change Log

Table 9 Revision history

Date	Version · identifier	Change
2026-11	v1.1 · APT2025v1.1	Reissued in the new methodology layout. Date of the v1.1 entry corrected from 11/02/2024 to 11/02/2025. Typographical corrections. No change to tests or scoring.
11/02/2025	v1.1 · APT2025v1.1	Added reference to MITRE's D3fend in the taxonomy used
24/10/2024	v1 · APT2025v1.0	Document created