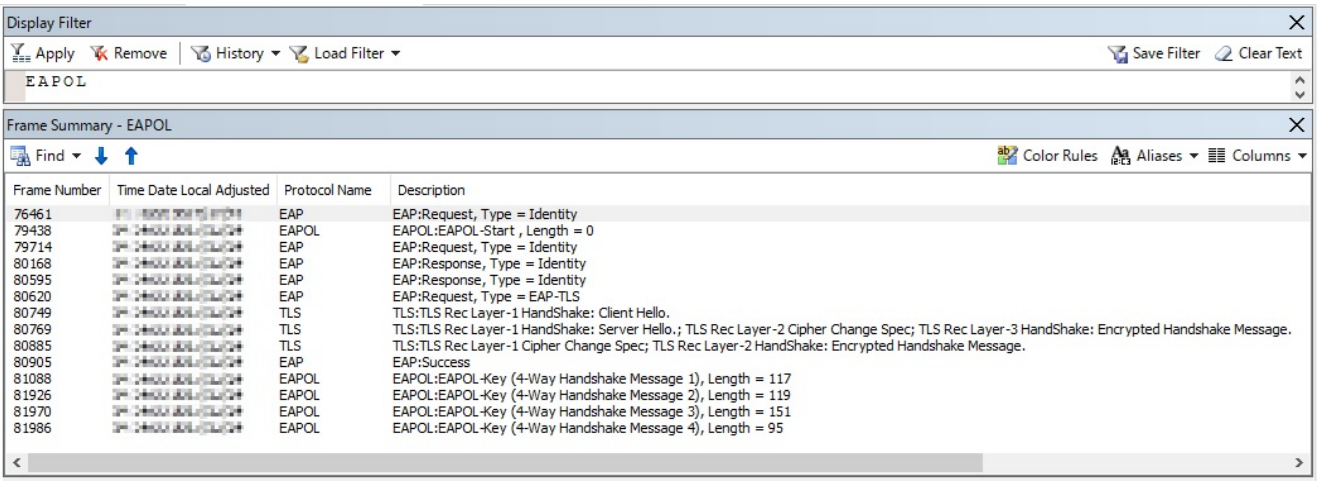
THREATVISOR RESEARCH

What a successful network test actually proves

An evidence model for separating authentication, access and application outcomes, with controls that expose misleading passes and unexplained failures.

TV-R04 · Engineering report · Version 1.0

Last updated · November 7, 2026



Microsoft Network Monitor: a packet trace records the authentication exchange. It does not establish application behavior. Source: Microsoft Learn. Used with permission from Microsoft.

An independent ThreatVisor publication, not affiliated with, authorized, sponsored or approved by Microsoft Corporation.

Microsoft and Windows are trademarks of the Microsoft group of companies. All other trademarks are the property of their respective owners.

Prepared for

Network and security engineers designing checks, reviewing rollout evidence or investigating ambiguous test results.

Sourced analysis and proposed engineering practice, not original experimental results.

Executive summary

- An accepted authentication request, a reachable TCP port and a successful application operation support different claims.
- Negative tests need evidence of the intended cause. An unrelated failure can make a broken security test appear successful.
- Missing, mismatched or stale observations should remain inconclusive. A defensible result connects a bounded claim to a repeatable procedure and reviewable evidence.

Method and scope

This is a proposed evidence and control design informed by Microsoft documentation. The examples are reasoning exercises, not recorded lab measurements. It does not prescribe a finished ThreatVisor implementation or claim support for a particular customer environment.

In this report

1. A result is an observation before it is a conclusion
2. Three boundaries that need separate evidence
3. Design controls that can disprove the interpretation
4. Connect every claim to a receipt
5. Use inconclusive as an engineering result
6. Evaluate the assessment method as well as the change

1 A result is an observation before it is a conclusion

”The test passed” hides the most important question: what did the test actually observe? A process exiting successfully may show only that a command ran. An authentication event may show an accepted request. A completed business operation may show that the application was usable on one path at one time. These observations can all be accurate while supporting very different decisions.

Begin with a bounded claim. ”This approved identity can retrieve a fresh record from this service through this Wi-Fi path after the candidate change” is reviewable. ”The network is safe” is not a useful test specification. The first statement identifies an identity, operation, path and changed state. It also exposes what the test does not cover.

Maintain a distinction between a measurement, its interpretation and the release decision. A collector gathers observations. An evaluator checks whether the evidence satisfies a criterion. An accountable owner decides whether the resulting scope and remaining gaps are acceptable. Combining these responsibilities into one unexplained success label makes disagreement difficult to resolve.

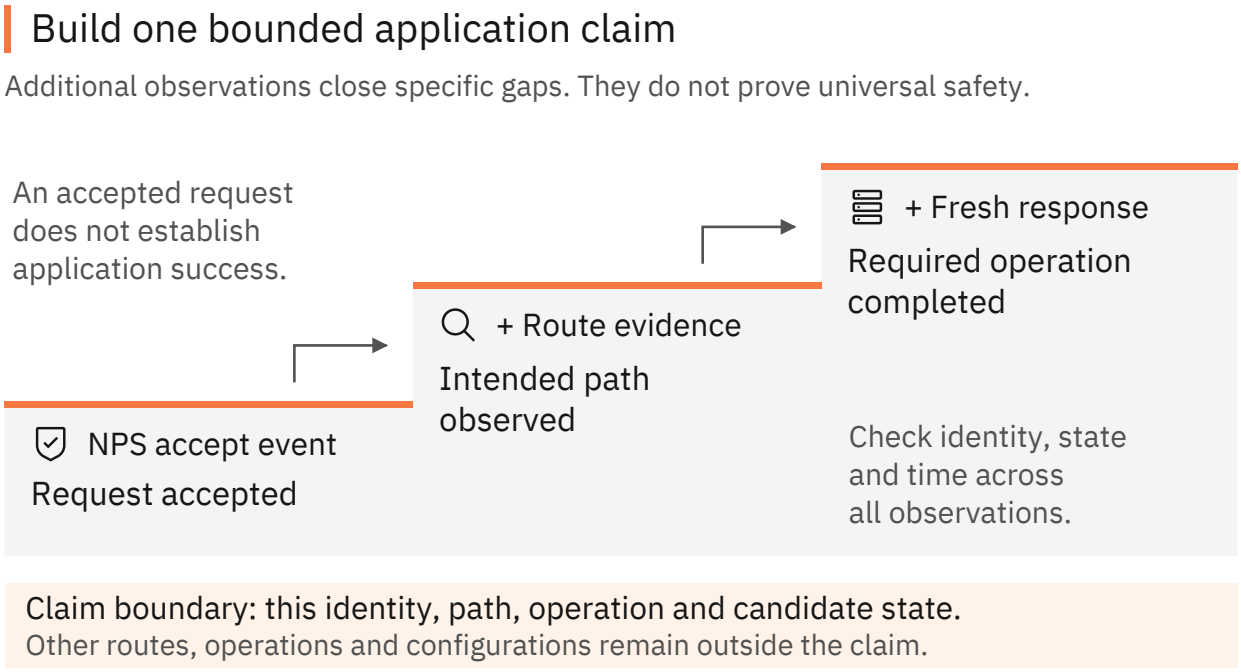
2 Three boundaries that need separate evidence

NPS authentication and authorization are one boundary. Application use is another. Microsoft’s NPS architecture places access points in the RADIUS-client role, while NPS makes network-access decisions. An accepted request therefore does not establish that address assignment, name resolution or a required application operation succeeded afterward. [2]

NPS event 6272 is evidence of acceptance and 6273 of rejection. The associated identity, policy and reason matter when interpreting either. A nearby event for another identity cannot stand in for the request under review. Client wireless logs can add context, but proximity in time is only one correlation aid. [3]

Test-NetConnection can supply useful connection and route diagnostics. A successful TCP connection establishes less than an authenticated application transaction with the required output. Conversely, an application result without route evidence can come from an alternate interface. Neither observation should silently fill the other’s gap. [4]

A practical evidence record therefore links three boundaries: the intended authentication attempt, the resulting network path and the fresh application operation. A complete record does not require pretending these components emit a common run identifier. It requires stating how the observations were associated and where that association remains uncertain.



Authentication acceptance, intended-path evidence and a fresh application response support different parts of one bounded claim. Correlate identity, candidate state and time across them. Additional observations do not establish untested operations or configurations.

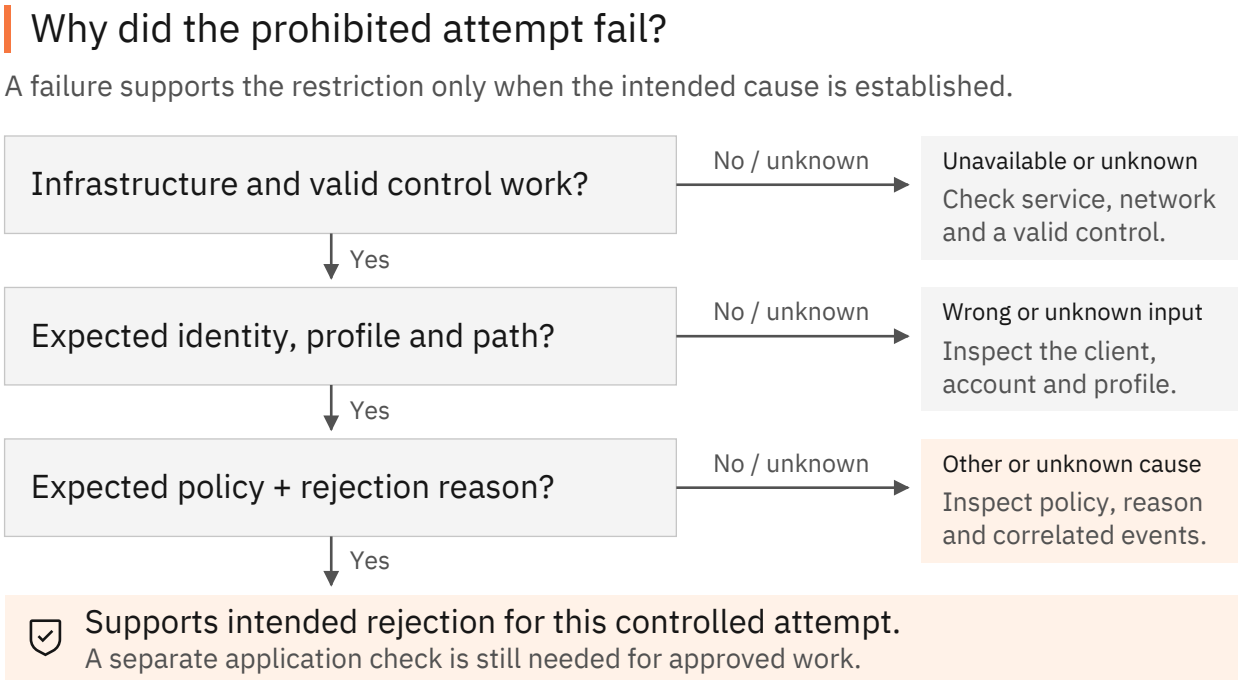
3 Design controls that can disprove the interpretation

A useful control tests an alternative explanation. Suppose an intended security check requires an old authentication method to be rejected. A failed connection is compatible with that claim, but it is also compatible with an unreachable server, a wrong identity or an unrelated certificate error. Unless the procedure distinguishes these explanations, it can pass when the intended restriction is absent.

Pair the prohibited attempt with an approved attempt where the scenario permits. Observe the expected rejection reason and the relevant client and server context. If both attempts fail because infrastructure is unavailable, the security outcome is unresolved. If the approved attempt succeeds through an alternate route, the workflow outcome is unresolved for the target path.

Controls must remain authorized and isolated. Use a permitted test identity, profile and network segment. Agree recovery with the infrastructure owner. A control designed to expose a weak interpretation is not permission to weaken a shared production policy or interrupt other users.

Also test the evidence pipeline itself. Deliberately omit a required observation in a controlled run and verify that the evaluator returns inconclusive. Supply a receipt from the wrong configuration and verify that it is not accepted. These checks evaluate the assessment method rather than the network's security.



A failed prohibited attempt supports the intended restriction only after service and positive controls, identity/profile/path, and the expected policy and rejection reason are checked. No or unknown evidence leads to investigation. Approved application work needs a separate check.

4 Connect every claim to a receipt

Use one row per criterion rather than one row per tool. A tool inventory can show substantial activity without explaining what has been established. A criterion-based record makes it possible to ask which evidence supports the intended security effect and which supports continued work.

Preserve the procedure version, configuration reference, observation time, raw receipt location and reviewer interpretation. Keep access-controlled raw logs separate from a shareable summary where necessary. Event exports can contain identities and infrastructure details, so the report should reference evidence without indiscriminately copying it.

Define the required freshness of the application operation. One approach is to create a harmless new record for the run and retrieve it through the target service, preserving a response identifier. The exact method depends on the application and permission available. If a fresh transaction cannot be observed, report the narrower connectivity result rather than inventing application coverage.

A claim-to-evidence record

Claim	Evidence to retain	Unresolved when
Correct attempt was assessed	Identity, profile, interface and bounded correlation window	A different identity or ambiguous request could explain the event.
Intended restriction took effect	Relevant attempt and expected rejection evidence	Failure reason is absent or infrastructure is unreachable.
Approved work completed	Fresh application output plus intended-path evidence	Output may be cached or returned through another route.
Candidate was evaluated	Build/configuration references and procedure version	Evidence describes a different or unknown state.

5 Use inconclusive as an engineering result

A three-state assessment is often more honest than a binary pass or fail. Satisfied means the required observations support the bounded criterion. Not satisfied means adequate observations show that the criterion was not met. Inconclusive means the available evidence cannot resolve it. The distinction prevents collection failure from masquerading as system failure or success.

Consider a proposed certificate-authentication change. The approved request is accepted by NPS, but the application transaction fails. The authentication criterion may be satisfied while the workflow criterion is not. If the application collector never produced a receipt, the workflow result instead remains inconclusive. Those cases require different investigations.

Now consider a negative control that is rejected because the client certificate has the wrong purpose. Microsoft documents separate client and server certificate requirements. That rejection may establish a certificate problem, but it does not prove that an unrelated prohibited authentication method was correctly blocked. The conclusion must follow the observed mechanism. [1]

When a reviewer overrides a proposed verdict, preserve the original observation, the revised interpretation and the reason. Do not rewrite the evidence to make the decision look inevitable. This history is useful when a later rollout exposes a condition the initial assessment did not cover.

6 Evaluate the assessment method as well as the change

The first question for an automated method is whether it reaches the right conclusion on known cases. Establish an approved baseline, a deliberately broken candidate and a corrected candidate in an authorized environment. Include controls that challenge identity, path and evidence completeness. A method that cannot distinguish these cases is not ready to support a rollout decision.

Compare it with a competent manual or scripted workflow using the same criteria. Count preparation, evidence collection, diagnosis, review and maintenance. Faster command execution is not the same as reduced assessment effort. A method that needs extensive correction after each configuration change may move work rather than remove it.

The companion evidence worksheet provides a place to record claims, observations, alternative explanations and bounded conclusions. The intended output is a reviewable argument about one change. It is not a certificate that every dependent system is safe, nor evidence that a proposed automation has already been validated.

References

Sources reviewed November 7, 2026.

[1] Microsoft: Certificate requirements for PEAP and EAP

Describes distinct server and client certificate requirements, including purpose and identity checks.

https://learn.microsoft.com/en-us/windows-server/networking/technologies/nps/nps-manage-cert-requirements

[2] Microsoft: Network Policy Server overview

Establishes the role of NPS and RADIUS clients in network-access authentication and authorization.

https://learn.microsoft.com/en-us/windows-server/networking/technologies/nps/nps-top

[3] Microsoft: Troubleshoot 802.1X authentication

Documents NPS acceptance/rejection events and client-side wireless, wired and certificate diagnostics.

https://learn.microsoft.com/en-us/troubleshoot/windows-client/networking/802-1x-authentication-issues-troubleshooting

[4] Microsoft: Test-NetConnection

Documents connectivity and route diagnostics. A successful TCP check is not an application transaction.

https://learn.microsoft.com/en-us/powershell/module/nettcpip/test-netconnection?view=windowsserver2025-ps