ON COVERAGE, MEASUREMENT AND THE GAPS WORTH PUBLISHING

# A vendor who publishes no gaps has not found none

*Coverage claims are the least reliable numbers in security, and they are reliably the ones on which platforms are chosen. This paper is about how to read them, and what an honest one would look like.*

IN BRIEF

"We cover 94% of the framework" is a sentence with no falsifiable content: it does not say what covering means, on which telemetry, in whose estate, measured when. This closing paper in the series proposes four properties a coverage claim needs to be worth anything, argues that a platform's gaps should be published alongside its coverage — including the ones that exist by construction and will not be closed — and ends with the questions this whole series was written to make askable.

## 1 · What a coverage percentage is actually counting

A framework enumerates techniques. A vendor maps rules to techniques. A technique with at least one mapped rule is marked covered, the marks are counted, and the count becomes a percentage on a slide. Every step of that is defensible and the result is close to meaningless, for four reasons that compound.

First, **a technique is not a unit**. Some describe a single system call; others describe a family of behaviours with dozens of practical variants. One rule against the first is comprehensive; one rule against the second is a token. Counting them equally is the arithmetic error at the centre of the whole practice.

Second, **coverage is conditional on telemetry the estate may not be sending**. A rule requiring command-line arguments, or DNS logging, or an audit subsystem that is off by default, is covered in the lab and absent in the customer. The percentage describes the product; the customer's exposure depends on their own collection.

Third, **detecting is not the same as detecting usefully**. A rule that fires on the technique and also on eleven thousand benign events per day will be disabled within a fortnight. It remains mapped. Coverage counts the rule's existence, not its survival in production.

Fourth, **the number has no date**. Attacker tradecraft moves, sources change, rules are disabled during a noisy quarter and never re-enabled. A coverage figure without the day it was measured describes a moment that may be two years gone.

FIGURE 1 · THE SAME ESTATE, THREE HONEST NUMBERS

| 94% | 61% | 43% |
| --- | --- | --- |
| techniques with at least one rule | …whose required telemetry this estate actually sends | …still enabled and firing at usable precision |

Illustrative, not measured — the shape is the point. Only the third number describes what would happen tonight, and only the vendor who computes it from your estate's runtime can produce it. The first is the one on the slide.

## 2 · Four properties of a claim worth reading

**Generated, not written.** The figure is computed from the running system at a stated moment, not retyped into a document each quarter. A published number that cannot be regenerated on demand is a historical artefact within weeks, and everyone involved knows it.

**Conditional on named telemetry.** Each claim states what it requires. “Covered where process command lines are collected” is a useful sentence; “covered” is not, because it silently transfers a prerequisite onto the reader.

**Qualified by observed precision.** A detection's measured firing rate and adjudication outcomes are part of its coverage claim. A rule nobody can afford to leave enabled is not coverage; it is an entry in a table.

**Dated, and stale by default.** Every figure carries the moment it was computed, and a figure past its refresh window is withdrawn rather than left standing. The absence of a number is more honest than an undated one, and considerably easier to explain later.

These are the same properties Paper 06 asked of any figure that leaves the platform, and coverage is where they are hardest to hold — because the marketing value of an unqualified percentage is real, and the qualified one is always smaller.

**The regeneration test.** Ask when the published figure was last computed, and by what. If the answer is a person and a spreadsheet, the number is an opinion with a decimal point. If it is a job with a runtime source and a build that fails when the document drifts from it, the number is a measurement — and the difference is visible in how quickly they can answer.

## 3 · The gaps worth publishing

Some gaps are work not yet done, and a vendor should say so with a date. Others exist *by construction* — direct consequences of a design decision that would be made again, and that will therefore never be closed. Publishing the second kind is unusual, and it is the more useful disclosure, because it tells a buyer what compensating control they need rather than what to wait for.

| | |
|---|---|
| `below the agent` | Firmware and hardware implants beneath the operating system are outside what a host agent observes. This is not a roadmap item; it is where the sensor sits. The compensating controls are attestation, supply-chain assurance and firmware integrity measurement, and a buyer should know to require them. |
| `inside opaque encryption` | Content within properly terminated end-to-end encryption is not inspectable, by design and by preference. Detection there works from endpoint behaviour and metadata, and any vendor claiming payload visibility should be asked precisely how they obtained it. |
| `unmanaged estate` | A device with no agent and no network path through an enforcement point is invisible to a platform built on both. Discovery finds its traces; it does not cover it. Every platform has this gap and few state it. |
| `authorised misuse` | A person doing exactly what their role permits, at ordinary volume, for the wrong reason, produces no anomaly and matches an approved pattern. Behavioural baselining narrows it; legitimacy adjudication cannot resolve it, because the authorisation is genuine. This is a governance problem wearing a detection costume. |
| `the first instance` | A technique nobody has seen has no rule and no baseline deviation on its first use. Provenance shortens the time to understand it afterwards; nothing shortens it to zero, and a vendor implying otherwise is describing a product that does not exist. |

Published this way, a gap table reads as confidence rather than apology — but only if it is specific. “We may not detect novel attacks” is a disclaimer. “Firmware below the agent, and here is the attestation control we expect you to run alongside us” is a position.

## 4 · How to test a claim in an hour

01 When was this figure computed, by what, and can you regenerate it while I watch?

02 Which of these covered techniques require telemetry my estate is not currently collecting?

| | |
|---|---|
| 03 | For your three noisiest rules, what proportion of customers still have them enabled? |
| 04 | Show me a technique you cover with one rule that has a dozen practical variants. How is that a 1? |
| 05 | What do you not detect by construction, and what compensating control do you expect me to run? |
| 06 | What performance figure can you cite, and on which machine was it measured? |
| 07 | When a published number goes stale, what happens to it — is that automatic or a decision? |
| 08 | Show me your last three coverage regressions and what caused them. |

## 5 · What the series argued

Eight papers, one thread. The gap between knowing and acting closes only when the sensor and the enforcement point are one process (01). Acting unattended is defensible only when the alert carries the specific observations that justified it (02) and when something has judged that nobody authorised the behaviour (03). What the platform learns has to become what it enforces, through validation rather than transcription (04), within bounds the operator can state before arming it (05). Everything it observed, judged and did has to remain retrievable, and erasable, years later (06). And the identity plane has to be part of the same system, because the signal that revokes a session belongs to the endpoint (07).

None of these is a novel capability. Each exists, well built, in some product. The argument of the series is about the seams: every one of these decisions is cheap to make once, at design time, and expensive to retrofit — and the cost of not making them is paid in the minutes between a signal and a response, over and over, by people who did nothing wrong.

Which is why this last paper is about measurement. A platform architected this way should be willing to be measured on it, in your estate, with dated figures and published gaps. If it is not willing, the architecture was a story.

> *Ask for the date, the telemetry it assumed, and the gaps. A vendor who can answer all three has been measuring; the rest have been counting.*

EventLake

## ABOUT THIS PAPER

Paper 08, the last in a series on the architecture of active defense, published by **EventLake Research**. Papers 01–07 cover enforcement at the edge, explainable alerts, legitimacy, compiling findings into rules, bounded autonomy, evidence that survives, and continuous trust. The series is deliberately vendor-neutral, and the questions in §4 are meant to be asked of us as readily as of anyone else.

EventLake publishes coverage as a generated artefact with its measurement date and required telemetry stated, alongside the gaps that exist by construction. **eventlake.com**