

# The Repair Loop in Operation

*Self-Healing to the Data, Not the Screen: A Field Report from the Reference Implementation*

A COMPANION FIELD REPORT

**Melvin Fahnestock**

Version 1.1 — August 4, 2026

[doi.org/10.5281/zenodo.21798331](https://doi.org/10.5281/zenodo.21798331)

[expectationdrivendevelopment.com](https://expectationdrivendevelopment.com)

[mel@expectationdriven.com](mailto:mel@expectationdriven.com)

Companion to “Expectation-Driven Development: Gold-State Parity™ as a Write-Path Acceptance Gate for Agent-Written Stateful Applications” · Zenodo · Licensed CC BY 4.0

## CONTENTS

What this document is

The central result

Terminology and the method in one page

Record I — the Prove chain, clean

Record II — the Validate session at volume

Record III — the repair loop under seeded faults

Record IV — isolating the parity gate

What the four records show together

Scope and limitations

Closing

Appendix A: the Prove chain, one row per scenario

Appendix B: the Validate session, one row per session entry

Appendix C: the seeded-fault run, one row per execution

Appendix D: the parity run, one row per execution

Appendix E: the redaction policy

## What this document is

Expectation-driven development verifies an application at its persisted state: a trusted, fixed, versioned **gold state** supplies both the inputs each test scenario enters and the exact persisted result the application must produce; the inputs travel only through the application’s real screens and real command paths, never through a back door; and what the application actually persisted is compared against gold’s expected state, value by value, failing closed on any disagreement. The state can be anything structured enough to declare as gold; in the implementation reported here it is relational, and the records below speak of rows.

That standing acceptance test is **the gate**: the complete, versioned procedure just described, yielding one verdict for one identified build, execution configuration, and specification version — pass only when every covered obligation holds, or fail with an artifact naming the unmet obligation or the exact disagreement. It is what makes the self-healing built on it different in kind from the familiar sort. A repair loop can only ever heal to the standard of the verdict that stops it: a loop stopped by interface tests heals until the screens behave, and this one heals until **what you enter, derived from gold, produces a result that matches gold**, verified value by value at the stored state itself.

The main paper specifies the gate — the rules, the two lanes that operate it, and the conditions a pass requires. This report records the gate in operation on the reference implementation through four records. Before those runs, the author completed a dedicated optimization, review, and hardening pass covering the runner, both lane drivers, and the test scenarios.

The interface-evidence lane is named Prove and is invoked with `/prove`. Historical normalized console excerpts retain the lowercase `proof` labels emitted during the recorded runs.

| Record                             | Question                                                                                     | Result                                                                                                                   | Role in the argument                                           |
|------------------------------------|----------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------|
| <b>Record I: Prove</b>             | Can the declared interface obligations run cleanly and mint current evidence?                | 42 active scenarios: 36 first-attempt PASS, 6 pinned capability skips, 0 failures, 0 retries                             | Establishes the clean interface-evidence baseline              |
| <b>Record II: Validate</b>         | Can the full in-scope corpus reproduce gold through the official command path?               | 19 first-attempt replay PASS legs; roughly three-quarters of a million cumulative row comparisons; 0 failures, 0 retries | Establishes clean persisted-state parity at volume             |
| <b>Record III: Repair loop</b>     | Does the chain stop, diagnose, repair, and retest when upper-layer checks can see the fault? | 6 planted faults found and repaired across 5 stops; one sweep-found before its scenario ran; two stacked in one scenario | Demonstrates governed repair behavior and progressive exposure |
| <b>Record IV: Parity isolation</b> | Does parity refuse when every interface-level check passes but stored values are wrong?      | 7 planted create-path mapping faults; 7 interface-level passes; 7 parity convictions                                     | Tests the paper's central write-path claim                     |

Record II was gated by Record I and completed before either seeded demonstration. Records III and IV then ran in a commit-free debug mode that the harness structurally bars from certifying anything.

## The central result

Seven deliberately planted create-path mapping faults produced records that looked normal at the interface. Every interface-level test passed. The submitted commands were valid. The expected number of rows was created. Reading the records back produced plausible values. Yet each record was wrong at the persisted state: a constant replaced an entered value, a boolean was forced true, neighboring columns were transposed, or one field received another field's value.

The full-row comparison against gold supplied all seven convictions. After each repair, the failed scenario was rerun before the chain could advance. When the seventh repair was complete, the working tree was byte-identical to the clean committed baseline except for a disposable run marker. This was an author-run, seeded exercise on one deliberately selected fault class in one proprietary reference implementation. It establishes no general defect-detection rate; it shows what only the comparison caught in these seven executions.

Readers interested primarily in the central result can begin with Record IV. Readers evaluating the evidence chain should read the records in order.

## Terminology and the method in one page

The **gate** returns PASS or refuses with evidence; it does not repair code. The **repair loop** sits above it: diagnose the refusal, change the implementation, rerun the failed scenario, and advance only after the deterministic driver accepts the repair. `/prove` and `/validate` are the operator commands that invoke the Prove and Validate lanes. In this report, a **retry** repeats an execution without an intervening repair; a **retest** is the required execution after a change. The clean records consumed no retries. The seeded records used retests to adjudicate repairs.

The **expected store** holds the materialized expected-state projections — the gold in the form the comparison consumes. The **working store** is the test store the application writes through its declared doors. The expected store is never written by a run; the working store is never seeded with expected data. The comparison closures exercised in the recorded Validate

session ranged from a handful of rows to the low six figures. These closure sizes do not describe or limit the cardinality of the underlying tables. Appendix B records the replay volumes in buckets.

A trusted, fixed, versioned gold state supplies two projections for each scenario:

1. an **input projection** containing only the values allowed to enter through the declared interface or command path; and
2. an **expected-state projection** containing the contract-complete persisted result the application is required to produce.

The working store begins from a declared baseline. It is never seeded with the scenario's expected business data. Inputs travel through the application's supported entry path. The resulting state is then compared with the expected projection, value by value. Every in-scope observable must be compared exactly, handled by a declared structural translation, verified by a named invariant, or explicitly excluded with a mandatory delegated gate. Unknown state fails closed.

The implementation divides that obligation into two mechanically linked lanes:

| Dimension           | Prove lane ( /prove )                                                                                                                                                              | Validate lane ( /validate )                                                                                                                                                         |
|---------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Purpose</b>      | Prove the declared interface obligations and capture the exact command shapes the interface emits                                                                                  | Reproduce the full in-scope corpus through the official command path and compare the resulting state with gold                                                                      |
| <b>Scale</b>        | Coverage-selected interface scenarios; each compares the rows it creates with its scenario-specific expected projection                                                            | Every in-scope expected-state row in the reviewed replay closure                                                                                                                    |
| <b>Mechanism</b>    | Browser automation drives the real screens under deliberately constrained Prove geometry, captures the emitted command shape, and compares the resulting persisted state with gold | Command replay uses only shapes previously captured from Prove, follows the normal application-service and domain-validation path, and compares the full reviewed closure with gold |
| <b>Output</b>       | Current PASS markers, pinned capability skips, captured command-contract evidence, and scenario-level row and field comparisons                                                    | Row- and field-level comparison artifacts, plus one session verdict                                                                                                                 |
| <b>Failure rule</b> | Any unpinned skip, unmet interface obligation, unstable runtime identity, or scenario-level persisted-state disagreement refuses the chain                                         | Any skipped row, rejected command, unmatched command shape, unknown observable, or persisted-state disagreement refuses the session                                                 |

Validate cannot run merely because Prove ran at some earlier time. It requires fresh Prove evidence bound to the current build, scenario manifest, store identity, and runtime profile. Every replayed command must resolve uniquely to a command variant the interface actually demonstrated. A commit change strands prior markers behind the build they attest to. These are executable preconditions, not conventions.

The distinction between gate and loop matters. The loop's quality cannot exceed the quality of the verdict that tells it when to stop.

Names, identifiers, quantities, and infrastructure details are generalized under the redaction policy in Appendix E. The causal sequence, verdicts, attempt counts, and relationships among failures, repairs, and retests are preserved.

---

## Record I — the Prove chain, clean

### Question

Could the complete active interface chain run at one identified commit, with every declared scenario either passing on its first attempt or producing a capability skip already pinned by the reviewed oracle?

## Result

One operator command launched the full core chain from level 0. Across a fifteen-level dependency ladder, 42 active scenarios ended as 36 first-attempt PASS results and 6 pinned capability skips. There were zero failures, zero retries, and zero fix commits. Scenario time was roughly nineteen minutes. A dedicated gateability check then verified all 42 markers at one commit.

Appendix A records the chain one row per scenario.

## What happened

Nothing was sampled and nothing was retried. Every execution ran with `attempts=1`. The two longest scenarios reflected volume rather than instability: the root reference scenario took about two minutes and the cross-cutting grid sweep about four.

The Prove lane drove every datum through the real interface. “Prove geometry” deliberately reduced page sizes and viewports so that paging, scrolling, and search collisions were forced to occur. The row set was small because the lane selected for behaviors, not because it sampled behavior probabilistically.

The six skips were not silently accepted. Five belonged to one disabled capability module and one to another. Each skip was pinned in the reviewed expected-skips oracle and resolved against the application’s live feature state. Drift fails in both directions: an unexpected skip stops the chain, and a pinned scenario unexpectedly running also stops it.

Five additional scenarios were deferred before the chain began. Two represented append-only event streams with no command surface; three native-transition variants awaited reconstruction on the test-reference pattern. Deferred scenarios never entered the tally and were not counted as passes or skips.

The durable product of Record I was not merely a green console. It was a complete marker set tied to one commit. Because marker identity follows the commit rather than the clock, any later code change invalidates the set for a Validate session until Prove is run again.

## What it establishes

Record I establishes that the active interface-evidence chain completed cleanly, without retries, at one commit, and produced the evidence required to authorize Record II.

## What it does not establish

Record I does not establish full-corpus parity. It exercises the versioned interface coverage model, not every expected-state row. Pinned skips and deferred scenarios remain outside its executed behavior and are reported as such.

---

## Record II — the Validate session at volume

### Question

Given Record I’s current Prove evidence, could the full reviewed replay closure travel through the application’s official command path and reproduce the expected persisted state without disagreement?

## Result

After the preflight steps described below, the run entered a full-ladder Validate session at the same commit. Nineteen replay legs passed on their first attempt. Six capability skips were carried by their Prove markers. The driver reported a cumulative closure sum of roughly three-quarters of a million row comparisons, evaluated row by row and value by value. There were zero failures, zero retries, and zero code changes. The session completed in just over an hour.

Appendix B records the session one row per session entry and reconciles the replayed, skipped, deferred, and classified-out scenarios.

## What happened

Prove and Validate counts cannot be compared without their scope reconciliation:

42 active chain scenarios = 19 replay legs + 6 marker-backed capability skips + 7 classified gaps + 10 scenarios with no distinct replay data. Five additional deferred scenarios sat outside the active chain.

The Validate PASS certifies the 19 replay legs and the 6 marker-backed skips, and nothing else. Every non-replayed scenario belongs to a named category in the reviewed scope ledger:

- a baseline-synchronization entity with no create command;
- workflow-state entities whose terminal gold rows cannot be reproduced through a create path;
- derived event streams without one-to-one commands;
- native-transition variants awaiting rebuild;
- behavior scenarios with no independent comparison closure; or
- interface scenarios whose data already sits inside another replayed closure.

Two obstacles appeared before the replay ladder. The first was a reset refusal enforced by the driver. The second required the run's only operator override.

First, a full Prove chain leaves non-replay rows that restrictively reference session tables. The reset preflight correctly refuses to purge state it does not own. The driver therefore resolved a 26-scenario mechanical closure and re-proved it as a closure-scoped Prove chain before launch. That 26-scenario evidence set authorized the session; it is not the session tally. Appendix B records 25 session entries: 19 replay PASS legs and 6 marker-backed skips.

Second, the closure evidence did not cover the working tree's only newer commit, which contained the Prove run's own documentation record and no source change. The driver's default advice — to sync forward and re-prove — would have split the 42 markers across two commits. The operator instead used the explicit “prove tree as-is” override, recorded that the newer commit was documentation-only, and reran the separate gateability check. All 42 markers remained at one commit. The override bypassed a preflight coverage refusal; it did not bypass marker identity, scenario verdicts, command conformance, or final parity.

The reset then passed. Nineteen replay legs traversed the dependency ladder without failure. The expected store remained read-only. Commands wrote only to the working store, and each command conformed to a Prove-captured shape. The largest closures reached the low six figures.

The reported volume is intentionally qualified. This rounded volume is the driver’s cumulative sum of per-scenario comparison closures, not a distinct-row count. A table included in multiple closures contributes more than once. It describes comparison work performed in this session; it neither measures nor limits the cardinality of any underlying table.

### **What it establishes**

Record II establishes a clean, evidence-tethered full-corpus replay over the reviewed scope: current Prove markers, command-shape conformance, normal application entry paths, and no persisted-state disagreement.

### **What it does not establish**

The PASS says nothing about the classified-out categories, security, permissions, performance, concurrency, external delivery, or business correctness of the gold state. It also does not independently validate the runner that produced the verdict.

---

## **Record III — the repair loop under seeded faults**

### **Question**

When defects are visible to interface assertions, API contracts, or scenario probes, does the governed chain stop, preserve the failure, support diagnosis, retest the repair, and refuse to advance until the failed scenario passes?

### **Result**

The operator planted six single-token mutations in application code. The chain stopped five times across four scenarios. Six faults were found and repaired: five after driver stops and one through a peer sweep before its own scenario ran. Two faults were stacked in one scenario, so the second could not become visible until the first was repaired. Every repair was retested by the driver before the chain advanced. Every attempted scenario finished passing. The exercise took roughly one hour including diagnosis and repair.

The entire exercise ran in a commit-free debug mode. It created no commits, pushed nothing, merged nothing, and minted only scratch markers resting on uncommitted code. Those markers could not authorize a mainline Validate session.

Appendix C carries the execution record, the five stops, and the supporting diagnosis notes.

### **What happened**

Checkpoint resumes limited the cost of each stop: the driver first retested the failed scenario, then continued from the last verified boundary. Each failure created a repair debt associated with the run. The debt cleared only when the corresponding retest passed.

The spine of the run, as the operator’s console relayed it — normalized for names and rounded durations, with status lines faithful to the originals:

```

proof 1/42 · L0 root reference scenario          – PASS (~2 min)
proof 2/42 · L1 reference set C                  – PASS (~30 s)
... fifteen more first-attempt passes, levels 1 through 4 ...
proof 18/42 · L4 structured-document scenario   – FAIL (~55 s)
● CHAIN STOPPED – finalStatus=FAIL (only a clean PASS advances)
  diagnosis from artifacts and live probes; one-token repair applied,
  uncommitted; peer sweep repairs the same fault in a second grid service
proof resumed · retesting the failed scenario first
proof 1/25 · L4 structured-document (retest)    – PASS (~75 s) – fix accepted
... level 5 clean ...
proof 6/25 · L6 child transaction document      – FAIL (~40 s)
● CHAIN STOPPED – approval gate inverted; one-token repair, uncommitted
proof 1/20 · L6 child transaction doc (retest)  – PASS (~70 s) – fix accepted
... level 7: passes and the six pinned dark-module skips ...
proof 10/20 · L7 rate-resolution: applicable rate – FAIL (~10 s)
● CHAIN STOPPED – wrong identifier bound; one-token repair, uncommitted
proof 1/11 · L7 rate-resolution (retest)       – PASS (~50 s) – fix accepted
... siblings on the same endpoint pass; levels 11 and 12 clean ...
proof 9/11 · L13 workspace console scenario     – FAIL (~20 s)
● CHAIN STOPPED – effective-date filter off by one day; repair, uncommitted
proof 1/3 · L13 workspace console (retest 1)   – FAIL (~85 s)
  different test, further along – 17 of 35 passing, up from 10;
  same-failure detector: not triggered; progress: yes
● CHAIN STOPPED – assembler persisted the wrong identifier; repair, uncommitted
proof 1/3 · L13 workspace console (retest 2)   – PASS (~2 min) – all 35 tests
proof 2/3 · L13 cross-cutting grid sweep       – PASS (~4 min)
proof 3/3 · L14 master-entity compliance view  – PASS (~30 s)
CHAIN COMPLETE – every attempted scenario passing; markers are scratch state
resting on uncommitted code; nothing merged; the mainline gate is unaffected

```

| Stop         | Fault                                                                                | Convicting evidence                                                                                                 | Outcome                                                      |
|--------------|--------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------|
| 1            | A grid filter required a term to appear in both searchable columns instead of either | Live API truth table: both-column terms matched; single-column terms returned zero                                  | One-token repair; retest PASS                                |
| <b>Sweep</b> | The same conjunction existed in a second grid service                                | Peer review of every related filter predicate                                                                       | Repaired before that scenario ran; first-attempt PASS        |
| 2            | Approval controls were shown only for already-approved records                       | The scenario had already proved the records open; the gate was the inverse of the stated lifecycle contract         | One-token repair; retest PASS                                |
| 3            | The master entity's identifier was bound where the second master entity's belonged   | Symmetric live probes produced "invalid" in one parameter and "missing" in the other                                | One-token repair; retest PASS; sibling scenarios also passed |
| 4            | An effective-date filter admitted ranges starting one day after the query date       | A purpose-built boundary decoy starting exactly one day later was wrongly admitted and outranked the valid baseline | One-token repair; retest progressed further                  |
| 5            |                                                                                      |                                                                                                                     | One-token repair; final retest passed all 35 tests           |

| Stop | Fault                                                                                   | Convicting evidence                                                                                                                                            | Outcome |
|------|-----------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
|      | The assembler persisted the outer selector's identifier into the inner selector's field | The repaired scenario advanced from 10 to 17 passing tests; the trace showed a second submission rejected after rebuilding context from corrupted stored state |         |

## What the stacking establishes

The last two stops form the record's central sequence. Stop 4 prevented the scenario from reaching the code path containing Stop 5. Once the date-boundary defect was repaired, the scenario advanced farther and failed on a different test with a different message. The driver's same-failure detector did not trigger, and the passing count increased from 10 to 17.

The newly exposed defect was a valid identifier stored in the wrong field. It survived every read that merely displayed the stored value. The trace showed two submissions: the first created the corrupted record; the second rebuilt context from that record and was refused because the inner selector now carried the outer selector's identifier. Direct submissions against the same live application ruled out the versioning path. The repair used the correctly resolved local value the assembler already possessed. The final retest passed all 35 tests, and the remaining chain completed cleanly.

## What it establishes

Record III establishes the control behavior of the repair loop on these seeded faults: stop on evidence, diagnose, repair, retest, and advance only on a new deterministic PASS. It also demonstrates progressive exposure. Repairing one blocker did not satisfy the scenario; it allowed the scenario to reach and reveal the fault behind it.

Record III also acts as a positive control for Record IV. Interface assertions, contracts, and scenario probes were capable of convicting faults within their declared jurisdictions. The next record changes the fault class so that those upper layers remain correctly green.

## What it does not establish

All six defects were planted single-token mutations. The record says nothing about general convergence, multi-file faults, distributed logic errors, faults in the oracle, or detection rates. One of the six faults was found by a source sweep rather than convicted by a gate artifact; those are different evidentiary claims. All four failing scenarios inherited already-running application servers, so their diagnoses rely on interface snapshots, traces, live probes, and stored state rather than per-scenario server logs.

## Record IV — isolating the parity gate

### Question

If the submitted command is correct, the expected number of rows is created, the interface tests pass, validation rules accept the write, and the resulting record looks plausible when read back—but one or two persisted fields are wrong—does full-row parity refuse the run?

### Result

Seven create-path mapping faults were planted as working-tree mutations against a clean committed baseline. Their sibling update paths remained correct. Every interface-level test passed in all seven executions. Command conformance, domain validation, and read-path behavior supplied no conviction. The full-row comparison against gold refused all seven executions and identified the differing state. No additional unexplained comparison differences were reported.

The run used nine driver segments and stopped eight times: seven parity convictions and one integrity refusal that exercised no scenario. Every executed attempt count was 1. After the seventh repair, the working tree was byte-identical to the committed baseline except for the disposable run marker.

Appendix D carries the execution record, the diff signature for each conviction, and the integrity refusal that exercised no scenario.

## What happened

The class was chosen deliberately: **a create-path assignment writes the wrong value to one field, or a pair of create-path assignments transposes two fields, while the sibling update path maps those fields correctly.** This isolates the gap between a valid submitted command and the row ultimately stored.

The run as the operator saw it—normalized for names and rounded durations, with verdict lines faithful to the originals:

```
proof 1/42 · L0 root reference scenario      – PASS (~2 min)
    ... three more first-attempt passes ...
proof 5/42 · L1 reference set D              – FAIL: 7 differences, 7 rows vs 7 rows
⊖ CHAIN STOPPED – interface test PASSED; the compare convicted a hardcoded
  display-order constant on every created row; one-token repair
proof retest · reference set D – PASS – fix accepted, chain advances
⊖ reference set E – 1 difference: the single inactive reference row came back
  active – a forced-true flag, invisible on the ten rows that default true
⊖ reference set F – 7 differences: every description exactly its row's name,
  against references that hold none – a fabricated value, not a discard
⊖ master entity – 12 differences in mirrored pairs: two contact-email
  columns transposed, the nulls travelling with the swap
⊖ extension profile – 6 mirrored boolean differences: two payment-control
  flags transposed; the fourth profile invariant because its flags are equal
⊖ second master entity – a dimension column carrying the adjacent catalog
  column's values; the four sibling comparison specs all PASS
⊖ retest refused – checkpoint restore hash mismatch: unprovable state;
  the runner's own remedy taken verbatim – plain relaunch from level 0
proof relaunch · fifteen consecutive first-attempt passes re-witness every
  repair below level 4 at the same baseline
⊖ template entity – 1 difference: the one line carrying a description came
  back empty – a hardcoded empty value on the line create path
proof retest · template entity – PASS – fix accepted
    ... levels 4 through 14 run out clean ...
CHAIN COMPLETE – every attempted scenario passing; markers are scratch state;
nothing merged; the mainline gate is unaffected
```

| Stop | Planted create-path fault               | Comparison signature                                                                  | Repair evidence               |
|------|-----------------------------------------|---------------------------------------------------------------------------------------|-------------------------------|
| 1    | Display order hardcoded to one constant | Seven differences; every created row carried the same value against three gold values | One-token repair; retest PASS |
| 2    | Active flag forced true                 | One difference—the only gold row whose value was false                                | One-token repair; retest PASS |

| Stop | Planted create-path fault                              | Comparison signature                                                                                | Repair evidence                                  |
|------|--------------------------------------------------------|-----------------------------------------------------------------------------------------------------|--------------------------------------------------|
| 3    | Description copied from the neighboring name field     | Seven differences; every description duplicated its row's name                                      | One-token repair; retest PASS                    |
| 4    | Two contact-email columns transposed                   | Twelve mirrored differences across six records, including nulls moving with the swap                | Two-line repair; retest PASS                     |
| 5    | Two payment-control flags transposed                   | Six mirrored boolean differences across three profiles; the equal-valued profile remained invariant | Two-line repair; retest PASS                     |
| 6    | Dimension field filled from the adjacent catalog field | Six differences on one column, each duplicating its neighbor; four sibling specs passed             | One-token repair; verified by a level-0 relaunch |
| 7    | The only populated line description hardcoded empty    | One difference on the single gold line carrying a description                                       | One-token repair; retest PASS                    |

The layers reported exactly what their contracts allowed:

| Layer                                                      | Verdict on the seven planted faults |
|------------------------------------------------------------|-------------------------------------|
| Interface tests: submitted command and resulting row count | Passed all seven                    |
| Command-contract conformance                               | No conviction                       |
| Domain validation                                          | No conviction                       |
| Read-path rendering                                        | No conviction                       |
| <b>Full-row parity against gold</b>                        | <b>Convicted all seven</b>          |

“No conviction” does not mean every upper layer contained an assertion capable of catching this class. The class was designed so that the submitted command remained correct and the defect occurred strictly between acceptance and persistence. The point is not that the upper layers failed at their own jobs. The point is that their jobs ended before the wrong value became observable as wrong.

The class also conceals itself during ordinary investigation. Reading a record back shows a valid value. Because the sibling update path was correct by construction, editing and saving the record would have overwritten the wrong value with the right one. That consequence follows from the planted fault design; it was not exercised in this run. A user investigating through the screen could therefore erase the evidence without learning why the original row was wrong. The comparison between the persisted row and gold held outside the application under test remained the only witness left standing.

Two faults make the operational stakes concrete even under the report's generalized vocabulary. One transposed two contact-email columns. Another transposed two payment-control flags that determine whether a downstream payment action proceeds or is held. Neither defect made a screen fail. Both could survive until a downstream process consumed the field for a purpose different from display.

## The detector that had to learn its own lesson

After the first conviction, the operator began a blast-radius sweep. Its first two formulations were wrong. A search for literal right-hand sides missed the forced boolean. A wider search for assignments referencing no input missed the fabricated description because it referenced an input — the wrong one. After the third conviction, the successful detector compared create-path assignments with their corresponding update-path mappings. It returned the four faults then remaining and six reviewed non-faults.

Those four predicted faults were deliberately left in place until the chain reached them. Removing them early would have made the detector, rather than parity, the source of the result. Leaving them cost four additional stop-repair-retest cycles and preserved the question the exercise was designed to answer. Each was later convicted by the comparison on its own execution evidence.

## **The refusal in the middle and the final baseline check**

One stop in the middle was not a defect conviction. A checkpoint restore produced state whose full-scope hash did not match the checkpoint's recorded hash. The driver refused before exercising the scenario and instructed a plain relaunch from level 0. The operator followed that remedy. Fifteen consecutive first-attempt passes then re-witnessed every earlier repair at the same baseline before the chain continued.

The run began with seven mutations recorded in the operator's manifest against a clean commit. Each repair restored the baseline implementation. At completion, a byte-for-byte comparison found only the disposable run marker. The manifest establishes the count of seven; the final comparison confirms that no planted mutation remained.

## **What it establishes**

Within the tested class — a wrong single-field assignment or a two-field transposition on the create path, with the sibling update path correct — the persisted-state comparison convicted all seven while every interface-level test passed. The run demonstrates a concrete difference between “the application accepted and displayed the write” and “the application persisted the expected state.”

It also demonstrates the stopping standard inherited by the repair loop. An interface-driven loop would never have opened on these executions because its verdict was already green. The parity-driven loop opened on the stored disagreement and did not stop until the failed scenario reproduced gold.

## **What it does not establish**

The seven faults are a designed isolation exercise, not a sample from production. They do not establish a general detection rate or coverage for distributed faults, plausible but incorrect calculations, concurrency defects, security failures, external side effects, or oracle errors. The exercise was run by the author, against an author-built instrument, with unpublished proprietary artifacts.

## **What would refute it**

A direct refutation remains possible. A conforming reproduction in which the upper layers pass, the persisted state differs from gold inside the declared comparison closure, and parity nevertheless returns green would refute this result. A conviction on an unmodified baseline would challenge the comparison's validity. A stronger replication would use mutations sealed by someone other than the diagnosing operator, an independently reviewed runner, and retained artifacts for adjudicating both false-green and false-red outcomes.

---

## **What the four records show together**

Records I and II establish the clean state: the interface evidence was current, the full replay closure ran through the official command path, and the reviewed comparison scope reproduced gold without disagreement.

Record III establishes the repair discipline: when a declared upper-layer obligation failed, the chain stopped, the evidence supported diagnosis, and the driver required a retest before advancing. The stacked pair shows why a repair loop must continue until the scenario is actually clean rather than until one plausible defect has been fixed.

Record IV isolates the claim unique to this report. The screens were not broken. The commands were not malformed. The writes were accepted. The records could be read back. Yet the stored values were wrong, and only the comparison with the separately maintained expected state detected the error.

The distinction can be stated without metaphor:

- An interface test can establish that a user action was accepted and produced the behavior it asserts.
- A read-back test can establish that the system returns what it stored.
- Gold-State Parity™ can establish, over its declared comparison closure, that what the system stored matches the approved expected state maintained outside the application under test.

None of those claims subsumes the others. The expected state was maintained separately from the application under test. That separation is architectural; it does not imply independent authorship, governance, or review. What the report puts on the record is the third claim operating as the stop condition for a governed repair loop.

---

## Scope and limitations

This is a practitioner field report from one proprietary reference implementation. The four records come from one hardened build line and were reported by the implementation's author. The underlying run artifacts are proprietary and unpublished. The report therefore carries the evidentiary weight of a disclosed practitioner account, not an independently auditable dataset.

The clean records contain no defect. Records I and II show the machinery running green; they do not show what it detects. The seeded records contain no organically discovered defect. Records III and IV show the machinery responding to thirteen operator-planted mutations across two bounded fault families; they do not establish field prevalence or a general detection rate.

The runner is the measurement instrument, and its review was not independent. The author performed the optimization and hardening pass that preceded these records. No independent party has reviewed the comparison and marker paths against the written specification or run them against an independently authored, sealed known-answer corpus. Record III's failures were corroborated through live probes, traces, and stored state outside the runner. Records I and II necessarily rely more heavily on the instrument's own PASS verdicts; Record IV's seven convictions likewise rely on the comparison engine, its diff artifacts, and the reported clean-baseline identity rather than an independent reproduction.

The Validate PASS applies only to its reviewed scope. Classified-out categories are named in Record II and Appendix B. The reported volume is a cumulative closure sum, not a distinct-row count. Timing and orchestration figures describe this implementation on this hardware; they are not intrinsic performance claims about the method.

The gate does not determine whether gold reflects correct business intent. It verifies reproduction of approved gold over the declared closure. Governance of the reference, selection of scenarios, sufficiency of coverage, architecture, security, privacy, permissions, performance, resilience, concurrency, and external outcomes remain separate responsibilities and gates.

The next confidence-building step is not a larger author-run tally. It is independent replication: a reviewed instrument, sealed mutations or known-answer data, an operator who did not plant the faults, and retained artifacts that permit adjudication of both passes and refusals.

---

## Closing

The same instrument produced four complementary records. It proved the declared interface behaviors, validated the reviewed corpus at volume, stopped and recovered under seeded upper-layer faults, and refused seven runs after the application persisted writes that had passed every implemented interface-level check.

Together, the four records provide both halves of the argument. Records I and II show deliberately uneventful green runs supported by current evidence. Records III and IV show what the same gates do when the code is wrong: first against faults visible to upper-layer checks, then against seven designed to escape the implemented interface-level checks while remaining visible in persisted state.

The final record is the reason for the report. Each bad write looked ordinary from the screen. Each contained a valid or plausible value. By construction, several would have been overwritten with the correct value if a user opened the record and saved it; that behavior was not exercised in the run. The wrongness existed not in whether the application could display the row, but in whether the row matched the state it was supposed to persist.

The repair loop did not make that judgment. The comparison did. The loop inherited the comparison's stopping standard and kept working until the persisted result matched the expected state.

**That is the difference between healing until the screens behave and healing until what you enter, derived from gold, produces a result that matches gold.**

---

## Appendix A: the Prove chain, one row per scenario

One verdict line per scenario, as the driver stamped them; roles per the redaction policy, durations rounded. Every row ran at `attempts=1`. Levels holding no scenarios in this chain do not appear.

| Level | Scenario (role)                                         | Verdict        |
|-------|---------------------------------------------------------|----------------|
| L0    | root reference scenario                                 | PASS ( ~2 min) |
| L1    | reference set C                                         | PASS ( ~30 s)  |
| L1    | reference set B                                         | PASS ( ~15 s)  |
| L1    | reference set A                                         | PASS ( ~15 s)  |
| L1    | reference set D                                         | PASS ( ~20 s)  |
| L1    | reference set E                                         | PASS ( ~20 s)  |
| L1    | reference set F                                         | PASS ( ~20 s)  |
| L2    | reference set G                                         | PASS ( ~20 s)  |
| L2    | master entity                                           | PASS ( ~35 s)  |
| L2    | reference set H                                         | PASS ( ~20 s)  |
| L3    | dark-capability scenario (the always-on OFF-side check) | PASS ( ~10 s)  |

| Level | Scenario (role)                             | Verdict                                 |
|-------|---------------------------------------------|-----------------------------------------|
| L3    | extension profile                           | PASS (~35 s)                            |
| L3    | second master entity                        | PASS (~25 s)                            |
| L4    | duplicate-rejection smoke (reference set G) | PASS (~10 s)                            |
| L4    | lookup-cache invalidation scenario          | PASS (~15 s)                            |
| L4    | template-entity scenario                    | PASS (~30 s)                            |
| L4    | entity-group scenario                       | PASS (~30 s)                            |
| L4    | structured-document scenario                | PASS (~25 s)                            |
| L5    | entity-group clone smoke                    | PASS (~10 s)                            |
| L5    | entity-group edit-save smoke                | PASS (~10 s)                            |
| L5    | entity-group readiness-block smoke          | PASS (~10 s)                            |
| L5    | transaction document                        | PASS (~40 s)                            |
| L6    | child transaction document                  | PASS (~30 s)                            |
| L7    | derived-valuation scenario                  | PASS (~35 s)                            |
| L7    | bulk-input resolver scenario                | <b>SKIPPED</b> — dark module M1, pinned |
| L7    | hold-placement UI smoke                     | <b>SKIPPED</b> — dark module M2, pinned |
| L7    | transaction-document detail-tabs smoke      | PASS (~15 s)                            |
| L7    | negotiation-flow scenario (base)            | <b>SKIPPED</b> — dark module M1, pinned |
| L7    | negotiation-flow scenario (variant)         | <b>SKIPPED</b> — dark module M1, pinned |
| L7    | seed-data scenario                          | <b>SKIPPED</b> — dark module M1, pinned |
| L7    | hub scenario                                | <b>SKIPPED</b> — dark module M1, pinned |
| L7    | rate-resolution: applicable rate            | PASS (~10 s)                            |
| L7    | rate-resolution: bridge                     | PASS (~15 s)                            |
| L7    | rate-resolution: legacy/future              | PASS (~10 s)                            |
| L7    | child-transaction approval smoke            | PASS (~15 s)                            |
| L11   | configuration-entity scenario               | PASS (~30 s)                            |
| L12   | defaults-profile scenario                   | PASS (~30 s)                            |
| L13   | aggregate-worksheet scenario                | PASS (~20 s)                            |
| L13   | derived transaction documents               | PASS (~20 s)                            |
| L13   | workspace console scenario                  | PASS (~70 s)                            |
| L13   | cross-cutting grid sweep                    | PASS (~4 min)                           |
| L14   | master-entity compliance view               | PASS (~30 s)                            |

#### CHAIN COMPLETE — PARTIAL

42 active scenarios: 36 PASS + 6 pinned capability-skips (5 on dark module M1, 1 on M2)

5 statically deferred, excluded at discovery · 0 FAIL · 0 retries consumed

GATEABLE: all 42 active markers PASS/SKIPPED at one commit

Statically deferred, visible in the discovery record rather than the tally: an account-provisioning scenario and a message-thread scenario at L3, both append-only event streams with no command surface; and native-transition variants of the child transaction document (L6), the transaction document (L7), and the compliance view (L14), awaiting rebuild on the test-reference pattern, their expected-state seeding having been removed as an invalid concept.

## Appendix B: the Validate session, one row per session entry

The session's closing evidence box, normalized: roles per the redaction policy, row counts bucketed, durations rounded. Every replay leg passed on its first attempt.

| Level | Scenario (role)                     | Result                          | Rows compared     | Duration |
|-------|-------------------------------------|---------------------------------|-------------------|----------|
| L1    | reference set C                     | PASS                            | a handful         | seconds  |
| L1    | reference set B                     | PASS                            | several hundred   | seconds  |
| L1    | reference set A                     | PASS                            | ~100              | seconds  |
| L1    | reference set D                     | PASS                            | ~150              | seconds  |
| L1    | reference set E                     | PASS                            | a handful         | seconds  |
| L1    | reference set F                     | PASS                            | dozens            | seconds  |
| L2    | reference set G                     | PASS                            | ~300              | seconds  |
| L2    | master entity (with child records)  | PASS                            | over ten thousand | ~1 min   |
| L2    | reference set H                     | PASS                            | dozens            | seconds  |
| L3    | extension profile                   | PASS                            | several thousand  | ~30 s    |
| L3    | second master entity                | PASS                            | tens of thousands | ~2 min   |
| L4    | template-entity scenario            | PASS                            | low six figures   | ~26 min  |
| L4    | entity-group scenario               | PASS                            | tens of thousands | ~10 s    |
| L4    | structured-document scenario        | PASS                            | low six figures   | ~4 min   |
| L7    | derived-valuation scenario          | PASS                            | low six figures   | ~17 min  |
| L7    | bulk-input resolver scenario        | SKIPPED — M1 off, marker-backed | —                 | seconds  |
| L7    | hold-placement UI smoke             | SKIPPED — M2 off, marker-backed | —                 | seconds  |
| L7    | negotiation-flow scenario (base)    | SKIPPED — M1 off, marker-backed | —                 | seconds  |
| L7    | negotiation-flow scenario (variant) | SKIPPED — M1 off, marker-backed | —                 | seconds  |
| L7    | seed-data scenario                  | SKIPPED — M1 off, marker-backed | —                 | seconds  |
| L7    | hub scenario                        | SKIPPED — M1 off, marker-backed | —                 | seconds  |
| L11   | configuration-entity scenario       | PASS                            | low six figures   | ~9 min   |
| L12   | defaults-profile scenario           | PASS                            | thousands         | ~30 s    |
| L13   | aggregate-worksheet scenario        | PASS                            | tens of thousands | ~2 min   |
| L14   | master-entity compliance view       | PASS                            | tens of thousands | ~2 min   |

```

VALIDATE SESSION COMPLETE (REVIEWED SCOPE) – 19/25 PASS, 6 SKIPPED
scope: 42 chain = 19 replay + 6 dark-pinned + 7 reported gap + 10 no-distinct-data (silent)
plus 5 deferred classified out but not in the chain
roughly three-quarters of a million cumulative row comparisons against gold through the
application command path, row by row, value by value
0 FAIL · 0 retries · 0 code changes · session reset PASS

```

Classified out of replay scope, by category: the root reference scenario (baseline-sync entity); the transaction document, child transaction document, and derived transaction documents (workflow-state machine); the account-provisioning and message-thread scenarios (derived event streams — deferred); the three rate-resolution scenarios (no compare closure); the three native-transition variants (awaiting rebuild — deferred). Categories and exit criteria are held in the reviewed scope ledger described in Record II.

## Appendix C: the seeded-fault run, one row per execution

Commit-free mode; all repairs uncommitted by contract; roles per the redaction policy, durations rounded. One row per execution, in execution order across the run's six driver segments; a scenario that failed and was repaired appears again as a retest, so the four failing scenarios occupy more than one row each. The **Segment** column names which driver segment carries the row; each of the five resumes restored the last verified checkpoint rather than re-proving from level 0, so no stop cost a full relaunch. Five scenarios are statically deferred and never entered the chain; they are the same five listed under Appendix A.

| Level | Scenario (role)                                | Verdict       | Segment | Fault / fix note |
|-------|------------------------------------------------|---------------|---------|------------------|
| L0    | root reference scenario                        | PASS (~2 min) | 1       |                  |
| L1    | reference set C                                | PASS (~30 s)  | 1       |                  |
| L1    | reference set B                                | PASS (~15 s)  | 1       |                  |
| L1    | reference set A                                | PASS (~15 s)  | 1       |                  |
| L1    | reference set D                                | PASS (~20 s)  | 1       |                  |
| L1    | reference set E                                | PASS (~20 s)  | 1       |                  |
| L1    | reference set F                                | PASS (~20 s)  | 1       |                  |
| L2    | reference set G                                | PASS (~20 s)  | 1       |                  |
| L2    | master entity                                  | PASS (~35 s)  | 1       |                  |
| L2    | reference set H                                | PASS (~20 s)  | 1       |                  |
| L3    | dark-capability scenario                       | PASS (~10 s)  | 1       |                  |
| L3    | extension profile                              | PASS (~35 s)  | 1       |                  |
| L3    | second master entity                           | PASS (~25 s)  | 1       |                  |
| L4    | duplicate-rejection<br>smoke (reference set G) | PASS (~10 s)  | 1       |                  |
| L4    | lookup-cache<br>invalidation scenario          | PASS (~15 s)  | 1       |                  |
| L4    | template-entity scenario                       | PASS (~30 s)  | 1       |                  |

| Level | Scenario (role)                           | Verdict                                           | Segment | Fault / fix note                                                                                                                                                                                                      |
|-------|-------------------------------------------|---------------------------------------------------|---------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| L4    | entity-group scenario                     | PASS (~25 s)                                      | 1       |                                                                                                                                                                                                                       |
| L4    | structured-document scenario              | <b>FAIL</b> (~55 s)                               | 1       | <b>Stop 1</b> — grid free-text filter conjoined its two searchable columns. One-token repair, uncommitted; the peer sweep repaired the same fault in the defaults-profile scenario's service at L12 in the same round |
| L4    | structured-document scenario (retest)     | PASS (~75 s)                                      | 2       | fix accepted                                                                                                                                                                                                          |
| L5    | entity-group clone smoke                  | PASS (~15 s)                                      | 2       |                                                                                                                                                                                                                       |
| L5    | entity-group edit-save smoke              | PASS (~10 s)                                      | 2       |                                                                                                                                                                                                                       |
| L5    | entity-group readiness-block smoke        | PASS (~10 s)                                      | 2       |                                                                                                                                                                                                                       |
| L5    | transaction document                      | PASS (~40 s)                                      | 2       |                                                                                                                                                                                                                       |
| L6    | child transaction document                | <b>FAIL</b> (~40 s)                               | 2       | <b>Stop 2</b> — approval controls gated on the approved status where open belonged. One-token repair, uncommitted                                                                                                     |
| L6    | child transaction document (retest)       | PASS (~70 s)                                      | 3       | fix accepted                                                                                                                                                                                                          |
| L7    | derived-valuation scenario                | PASS (~35 s)                                      | 3       |                                                                                                                                                                                                                       |
| L7    | bulk-input resolver scenario              | <b>SKIPPED</b> (seconds) — dark module M1, pinned | 3       |                                                                                                                                                                                                                       |
| L7    | hold-placement UI smoke                   | <b>SKIPPED</b> (seconds) — dark module M2, pinned | 3       |                                                                                                                                                                                                                       |
| L7    | transaction-document detail-tabs smoke    | PASS (~15 s)                                      | 3       |                                                                                                                                                                                                                       |
| L7    | negotiation-flow scenario (base)          | <b>SKIPPED</b> (seconds) — dark module M1, pinned | 3       |                                                                                                                                                                                                                       |
| L7    | negotiation-flow scenario (variant)       | <b>SKIPPED</b> (seconds) — dark module M1, pinned | 3       |                                                                                                                                                                                                                       |
| L7    | seed-data scenario                        | <b>SKIPPED</b> (seconds) — dark module M1, pinned | 3       |                                                                                                                                                                                                                       |
| L7    | hub scenario                              | <b>SKIPPED</b> (seconds) — dark module M1, pinned | 3       |                                                                                                                                                                                                                       |
| L7    | rate-resolution: applicable rate          | <b>FAIL</b> (~10 s)                               | 3       | <b>Stop 3</b> — the master entity's identifier bound into the field reserved for the second master entity's. One-token repair, uncommitted                                                                            |
| L7    | rate-resolution: applicable rate (retest) | PASS (~50 s)                                      | 4       | fix accepted                                                                                                                                                                                                          |
| L7    | rate-resolution: bridge                   | PASS (~15 s)                                      | 4       | sibling on the same endpoint; covers the stop 3 fix                                                                                                                                                                   |
| L7    | rate-resolution: legacy/future            | PASS (~10 s)                                      | 4       | sibling on the same endpoint; covers the stop 3 fix                                                                                                                                                                   |

| Level | Scenario (role)                       | Verdict             | Segment | Fault / fix note                                                                                                                                                                                                          |
|-------|---------------------------------------|---------------------|---------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| L7    | child-transaction approval smoke      | PASS (~15 s)        | 4       |                                                                                                                                                                                                                           |
| L11   | configuration-entity scenario         | PASS (~35 s)        | 4       |                                                                                                                                                                                                                           |
| L12   | defaults-profile scenario             | PASS (~30 s)        | 4       | first-attempt PASS carrying the sweep-found repair; applied before the chain reached it                                                                                                                                   |
| L13   | aggregate-worksheet scenario          | PASS (~20 s)        | 4       |                                                                                                                                                                                                                           |
| L13   | derived transaction documents         | PASS (~20 s)        | 4       |                                                                                                                                                                                                                           |
| L13   | workspace console scenario            | <b>FAIL</b> (~20 s) | 4       | <b>Stop 4</b> — shared effective-date range filter admitted ranges starting the day after the query date. Repair; uncommitted                                                                                             |
| L13   | workspace console scenario (retest 1) | <b>FAIL</b> (~85 s) | 5       | <b>Stop 5</b> , the stacked fault — the assembler persisted the outer selector's identifier into the inner selector's field. 17 of 35 tests passing, up from 10; same-failure detector not triggered. Repair; uncommitted |
| L13   | workspace console scenario (retest 2) | PASS (~2 min)       | 6       | fix accepted; all 35 tests passing                                                                                                                                                                                        |
| L13   | cross-cutting grid sweep              | PASS (~4 min)       | 6       |                                                                                                                                                                                                                           |
| L14   | master-entity compliance view         | PASS (~30 s)        | 6       |                                                                                                                                                                                                                           |

**The stops in detail.** The same five stops with the evidence that convicted each fault and the repair that cleared it.

| Stop | Level / scenario                     | Planted fault                                                                       | Convicted by                                                                                                                                                   | Repair, retested                                                           |
|------|--------------------------------------|-------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------|
| 1    | L4 structured-document               | Grid free-text filter conjoined its two searchable columns (AND for OR)             | Live-API truth table: both-columns terms matched, single-column terms returned zero rows                                                                       | One token; retest PASS                                                     |
| —    | L12 defaults-profile                 | Same conjunction, second grid service                                               | A peer sweep across every grid service, before the scenario ran                                                                                                | Same round; first-attempt PASS at L12                                      |
| 2    | L6 child transaction document        | Approval controls gated on the approved status where open belonged                  | The scenario's own probes: all documents proven open; its negative assertion requires the control gone once approved — the gate was the exact inverse          | One token; retest PASS                                                     |
| 3    | L7 rate-resolution (applicable rate) | Master entity's identifier bound where the second master entity's belonged          | Live-probe symmetry: the bogus value produced the invalid-entity refusal in one parameter and "missing" in the other                                           | One token; retest PASS; sibling scenarios covered the fix                  |
| 4    | L13 workspace console (iteration 1)  | Effective-date range filter admitted ranges starting the day after the query date   | A purpose-built boundary decoy starting exactly one day later was wrongly admitted and outranked the baseline on the tie-break                                 | One token; retest advanced 10 → 17 tests, then failed on the stacked fault |
| 5    | L13 workspace console (iteration 2)  | Assembler persisted the outer selector's identifier into the inner selector's field | The console resuming from its own persisted state: the re-posted assemble refused for matching no baseline — the corrupted write re-consumed by the write path | One token; final retest all 35 tests PASS; chain ran out clean             |

**Diagnosis notes.** The evidence chain behind each conviction is held here so Record III can carry the run and this appendix can carry the forensics. Stop 5's diagnosis remains in Record III because the stacked sequence is the point.

**Stop 1 — the structured-document scenario (L4): a filter that demanded both.** The grid rendered “no records found” for a sentinel name that verifiably existed. Live API probes built the truth table: the shared sentinel prefix, in both name and

description, returned all six rows; the full name, in the name only, returned zero; a description-only term returned zero, ruling out “searches the wrong column”; every both-columns substring passed and every single-column substring failed, down to a two-character term matching exactly the one row carrying it in both. The filter conjoined its two searchable columns where either should suffice. One token. A peer sweep then examined every grid service’s filter predicate and every conjunction within them: ten services carried one in their text, eight were null-guards inside a disjunct and correct, and **exactly two were genuine — this one and the defaults-profile scenario’s service at L12** — fixed in the same round; that scenario later passed on its first attempt, its planted fault repaired before the chain ever reached it. Rejected alternative, on the record: rewriting the test’s row locator. It is the suite’s canonical row selector, and the same test had just driven pagination and multi-sort through it successfully — innocent; only the narrowed search made it resolve to nothing.

**Stop 2 — the child transaction document (L6): an approval gate inverted.** The detail page’s approval controls never rendered. The test’s own preceding probe had just proven every created document open, and its later assertion requires the approval control to vanish once a document is approved; the live gate offered the controls *only* on already-approved documents — the exact inverse of the contract the scenario states. The same block corroborates: it also renders the submit-for-review control, coherent on an open record and incoherent on an approved one, and the page’s own attachment guard already treats approved and denied as terminal. The fix moves the gate from the approved status to the open status, satisfying both directions of the stated contract. Rejected alternative: pointing the test at the review lane’s approval control — a different lifecycle transition on a different endpoint, whose substitution would have deleted the open-to-approved coverage rather than restored it.

**Stop 3 — the rate-resolution scenario (L7): the wrong identifier bound.** A request carrying a deliberately invalid identifier for the second master entity came back with the wrong refusal code — “missing” where “invalid” belonged. Live probes convicted the endpoint’s controller by symmetry: the bogus value in one parameter produced the invalid-entity refusal quoting it, in the other it produced “missing,” and a third probe put a bogus value in an unrelated scope parameter and got the correct scope refusal back — not a blanket binding failure but one line, with the resolver’s own guards and queries reading innocent on inspection. The controller bound the master entity’s identifier into the field reserved for the second master entity’s and never read the real parameter. One token. Rejected alternative: accepting the observed status as the contract. The two refusal codes are distinct diagnostics the endpoint exists to separate; accepting either would have made the scenario’s whole status contract unassertable. The sibling scenarios on the same endpoint passed on the next launch, covering the fix.

**Stop 4 — the workspace console (L13): the boundary-decoy case.** The console resolved the wrong template. The shared effective-date range filter admitted ranges beginning the day *after* the query date — off by one day. The scenario’s fixture contained a decoy built for that exact boundary; it was wrongly admitted and outranked the baseline on the declared tie-break. A second decoy, whose range ended on the effective date, contained neither the true date nor the shifted date and was correctly excluded either way. That result isolated a one-day-forward shift. Every caller of the filter was checked for a compensating shift before the fix; none compensated. Moving the decoy’s window would have removed the boundary assertion while leaving the faulty filter in place. Because the planted mutation matched a boundary the fixture explicitly covered, this stop demonstrates the loop’s control behavior more strongly than the reach of the coverage model.

Six faults, five stops, one sweep-found repair, two faults stacked in one scenario and peeled in order. The chain advanced only on driver-witnessed retests; the run minted nothing that could reach the mainline or gate a session.

## Appendix D: the parity run, one row per execution

Commit-free mode; all repairs uncommitted by contract; roles per the redaction policy, durations rounded. Every conviction came from the final comparison; the interface test passed in all seven cases. Diff counts describe the small comparison slice exercised by each Prove scenario, not the larger Validate replay closure in Appendix B. One row appears for each execution across the run's nine driver segments. Every executed row ran at `attempts=1`; the integrity refusal in segment 7 records `attempts=0` because no scenario was exercised. Segment 8 is a plain relaunch from level 0 rather than a resume, which is why levels 0 through 3 appear twice. Its fifteen consecutive first-attempt passes re-witnessed every repair below level 4 at the same baseline. Five scenarios are statically deferred and never entered the chain; they are the same five listed under Appendix A.

| Level | Scenario (role)            | Verdict        | Segment | Fault / fix note                                                                                                                                                                                                         |
|-------|----------------------------|----------------|---------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| L0    | root reference scenario    | PASS (~2 min)  | 1       |                                                                                                                                                                                                                          |
| L1    | reference set C            | PASS (~30 s)   | 1       |                                                                                                                                                                                                                          |
| L1    | reference set B            | PASS (~15 s)   | 1       |                                                                                                                                                                                                                          |
| L1    | reference set A            | PASS (~15 s)   | 1       |                                                                                                                                                                                                                          |
| L1    | reference set D            | FAIL (~20 s)   | 1       | <b>Stop 1</b> — 7 differences, 7 rows vs 7 rows: a display-order constant on every created row. Interface test PASSED. One token                                                                                         |
| L1    | reference set D (retest)   | PASS (~60 s)   | 2       | fix accepted                                                                                                                                                                                                             |
| L1    | reference set E            | FAIL (~20 s)   | 2       | <b>Stop 2</b> — 1 difference: the single inactive reference row came back active, a forced-true flag invisible on the rows that default true. Interface test PASSED. One token                                           |
| L1    | reference set E (retest)   | PASS (~60 s)   | 3       | fix accepted                                                                                                                                                                                                             |
| L1    | reference set F            | FAIL (~20 s)   | 3       | <b>Stop 3</b> — 7 differences: every description exactly its own row's name, against references that hold none. Interface test PASSED. One token                                                                         |
| L1    | reference set F (retest)   | PASS (~55 s)   | 4       | fix accepted                                                                                                                                                                                                             |
| L2    | reference set G            | PASS (~20 s)   | 4       |                                                                                                                                                                                                                          |
| L2    | master entity              | FAIL (~35 s)   | 4       | <b>Stop 4</b> — 12 differences in mirrored pairs across six records: two contact-email columns transposed, the nulls travelling with the swap. The two sibling specs PASS. Interface test PASSED. Two lines              |
| L2    | master entity (retest)     | PASS (~70 s)   | 5       | fix accepted                                                                                                                                                                                                             |
| L2    | reference set H            | PASS (~20 s)   | 5       |                                                                                                                                                                                                                          |
| L3    | dark-capability scenario   | PASS (~10 s)   | 5       |                                                                                                                                                                                                                          |
| L3    | extension profile          | FAIL (~35 s)   | 5       | <b>Stop 5</b> — 6 mirrored boolean differences across three of the four profiles: two payment-control flags transposed; the profile whose flags are equal is invariant. Interface test PASSED. Two lines                 |
| L3    | extension profile (retest) | PASS (~80 s)   | 6       | fix accepted                                                                                                                                                                                                             |
| L3    | second master entity       | FAIL (~25 s)   | 6       | <b>Stop 6</b> — 6 differences on one column: a dimension column carrying the adjacent catalog column's values. The four sibling comparison specs all PASS. Interface test PASSED. One token                              |
| L3    | second master entity       | REFUSED (~8 s) | 7       | <i>Not a fault</i> — the integrity refusal: restored state did not match the checkpoint's recorded full-scope hash. attempts=0, nothing exercised; the runner's own remedy taken verbatim, a plain relaunch from level 0 |
| L0    | root reference scenario    | PASS (~2 min)  | 8       |                                                                                                                                                                                                                          |

| Level | Scenario (role)                             | Verdict                                           | Segment | Fault / fix note                                                                                                                                           |
|-------|---------------------------------------------|---------------------------------------------------|---------|------------------------------------------------------------------------------------------------------------------------------------------------------------|
| L1    | reference set C                             | PASS (~20 s)                                      | 8       |                                                                                                                                                            |
| L1    | reference set B                             | PASS (~15 s)                                      | 8       |                                                                                                                                                            |
| L1    | reference set A                             | PASS (~15 s)                                      | 8       |                                                                                                                                                            |
| L1    | reference set D                             | PASS (~20 s)                                      | 8       | re-witnesses the stop 1 repair at the same baseline                                                                                                        |
| L1    | reference set E                             | PASS (~20 s)                                      | 8       | re-witnesses the stop 2 repair                                                                                                                             |
| L1    | reference set F                             | PASS (~20 s)                                      | 8       | re-witnesses the stop 3 repair                                                                                                                             |
| L2    | reference set G                             | PASS (~20 s)                                      | 8       |                                                                                                                                                            |
| L2    | master entity                               | PASS (~35 s)                                      | 8       | re-witnesses the stop 4 repair                                                                                                                             |
| L2    | reference set H                             | PASS (~20 s)                                      | 8       |                                                                                                                                                            |
| L3    | dark-capability scenario                    | PASS (~10 s)                                      | 8       |                                                                                                                                                            |
| L3    | extension profile                           | PASS (~35 s)                                      | 8       | re-witnesses the stop 5 repair                                                                                                                             |
| L3    | second master entity                        | PASS (~25 s)                                      | 8       | the stop 6 fix accepted here, by the relaunch rather than by a resume                                                                                      |
| L4    | duplicate-rejection smoke (reference set G) | PASS (~10 s)                                      | 8       |                                                                                                                                                            |
| L4    | lookup-cache invalidation scenario          | PASS (~15 s)                                      | 8       |                                                                                                                                                            |
| L4    | template-entity scenario                    | <b>FAIL</b> (~30 s)                               | 8       | <b>Stop 7</b> — 1 difference on the line comparison spec: the only reference line carrying a description came back empty. Interface test PASSED. One token |
| L4    | template-entity scenario (retest)           | PASS (~65 s)                                      | 9       | fix accepted                                                                                                                                               |
| L4    | entity-group scenario                       | PASS (~30 s)                                      | 9       |                                                                                                                                                            |
| L4    | structured-document scenario                | PASS (~30 s)                                      | 9       |                                                                                                                                                            |
| L5    | entity-group clone smoke                    | PASS (~10 s)                                      | 9       |                                                                                                                                                            |
| L5    | entity-group edit-save smoke                | PASS (~10 s)                                      | 9       |                                                                                                                                                            |
| L5    | entity-group readiness-block smoke          | PASS (~10 s)                                      | 9       |                                                                                                                                                            |
| L5    | transaction document                        | PASS (~40 s)                                      | 9       |                                                                                                                                                            |
| L6    | child transaction document                  | PASS (~35 s)                                      | 9       |                                                                                                                                                            |
| L7    | derived-valuation scenario                  | PASS (~40 s)                                      | 9       |                                                                                                                                                            |
| L7    | bulk-input resolver scenario                | <b>SKIPPED</b> (seconds) — dark module M1, pinned | 9       |                                                                                                                                                            |
| L7    | hold-placement UI smoke                     | <b>SKIPPED</b> (seconds) — dark module M2, pinned | 9       |                                                                                                                                                            |
| L7    | transaction-document detail-tabs smoke      | PASS (~15 s)                                      | 9       |                                                                                                                                                            |

| Level | Scenario (role)                     | Verdict                                           | Segment | Fault / fix note |
|-------|-------------------------------------|---------------------------------------------------|---------|------------------|
| L7    | negotiation-flow scenario (base)    | <b>SKIPPED</b> (seconds) — dark module M1, pinned | 9       |                  |
| L7    | negotiation-flow scenario (variant) | <b>SKIPPED</b> (seconds) — dark module M1, pinned | 9       |                  |
| L7    | seed-data scenario                  | <b>SKIPPED</b> (seconds) — dark module M1, pinned | 9       |                  |
| L7    | hub scenario                        | <b>SKIPPED</b> (seconds) — dark module M1, pinned | 9       |                  |
| L7    | rate-resolution: applicable rate    | PASS (~10 s)                                      | 9       |                  |
| L7    | rate-resolution: bridge             | PASS (~15 s)                                      | 9       |                  |
| L7    | rate-resolution: legacy/future      | PASS (~10 s)                                      | 9       |                  |
| L7    | child-transaction approval smoke    | PASS (~10 s)                                      | 9       |                  |
| L11   | configuration-entity scenario       | PASS (~30 s)                                      | 9       |                  |
| L12   | defaults-profile scenario           | PASS (~30 s)                                      | 9       |                  |
| L13   | aggregate-worksheet scenario        | PASS (~20 s)                                      | 9       |                  |
| L13   | derived transaction documents       | PASS (~20 s)                                      | 9       |                  |
| L13   | workspace console scenario          | PASS (~70 s)                                      | 9       |                  |
| L13   | cross-cutting grid sweep            | PASS (~4 min)                                     | 9       |                  |
| L14   | master-entity compliance view       | PASS (~25 s)                                      | 9       |                  |

**The stops in detail.** The same seven fault stops and the integrity refusal, with the diff signature that convicted each fault and the repair that cleared it.

| Stop | Level / scenario   | Planted fault (create path only; update path correct) | Diff signature that convicted it                                                      | Repair                 |
|------|--------------------|-------------------------------------------------------|---------------------------------------------------------------------------------------|------------------------|
| 1    | L1 reference set D | Display-order field hardcoded to a constant           | The same value on all seven rows against three distinct reference values              | One token; retest PASS |
| 2    | L1 reference set E | Active flag forced true, discarding the input         | Exactly one diff — the single inactive reference row — always in the forced direction | One token; retest PASS |
| 3    | L1 reference set F | Description fabricated from the name field            | Every description exactly its own row's name, against references holding none         | One token; retest PASS |
| 4    | L2 master entity   | Two contact-email columns transposed                  | Twelve diffs in mirrored pairs across six records, nulls travelling with the swap     | Two lines; retest PASS |
| 5    |                    | Two payment-control flags transposed                  |                                                                                       | Two lines; retest PASS |

| Stop | Level / scenario        | Planted fault (create path only; update path correct)             | Diff signature that convicted it                                                          | Repair                                                 |
|------|-------------------------|-------------------------------------------------------------------|-------------------------------------------------------------------------------------------|--------------------------------------------------------|
|      | L3 extension profile    |                                                                   | Six mirrored boolean diffs across three profiles; the equal-flags profile invariant       |                                                        |
| 6    | L3 second master entity | Dimension column filled with the adjacent catalog column's values | All six diffs on one column, each value duplicating its neighbor; four sibling specs PASS | One token; verified by the level-0 relaunch            |
| —    | L3 (retest)             | <i>Not a fault</i> : checkpoint-restore hash mismatch             | The runner refused unprovable state, eight seconds, nothing exercised                     | Plain relaunch from level 0, as the refusal instructed |
| 7    | L4 template entity      | Line description hardcoded to empty                               | Exactly one diff — the only reference line carrying a description                         | One token; retest PASS; chain ran out clean            |

**Diagnosis note — the blast-radius sweep.** The sweep began after the first conviction and was wrong twice before it worked. A scan for literal right-hand sides missed the forced boolean. Widened to assignments referencing no input at all, it still missed the fabricated description, whose right-hand side references an input — the wrong one. After the third conviction, a detector written against the fault *class* returned the four faults then remaining, as recorded in Record IV. The per-stop diff signatures that convicted each fault are in the table above; they were sufficient on their own, and no probe chain was needed for any of the seven.

Seven faults, seven convictions by the comparison alone, one integrity refusal honored verbatim. After the repairs the working tree was byte-identical to the committed baseline: the tree is the injection manifest.

## Appendix E: the redaction policy

The author wrote the proprietary reference implementation and operated the reported runs. The organization and product are not identified. Business names, scenario names, schema identifiers, tool names, product names, commit identifiers, repository references, file paths, and infrastructure details have been removed or generalized. Database-scale row counts and durations have been rounded or bucketed. Small structural and evidentiary counts that carry method content — including scenario, level, stop, fault, attempt, and per-stop comparison counts — are retained deliberately. The causal structure is preserved exactly: the sequence of refusals, diagnoses, repairs, stops, and passes, and the verdict of every event. Runner output appears as **normalized excerpts**: structure and verdict lines are faithful to the originals, while identifiers and scale-revealing quantities are neutralized as described here.