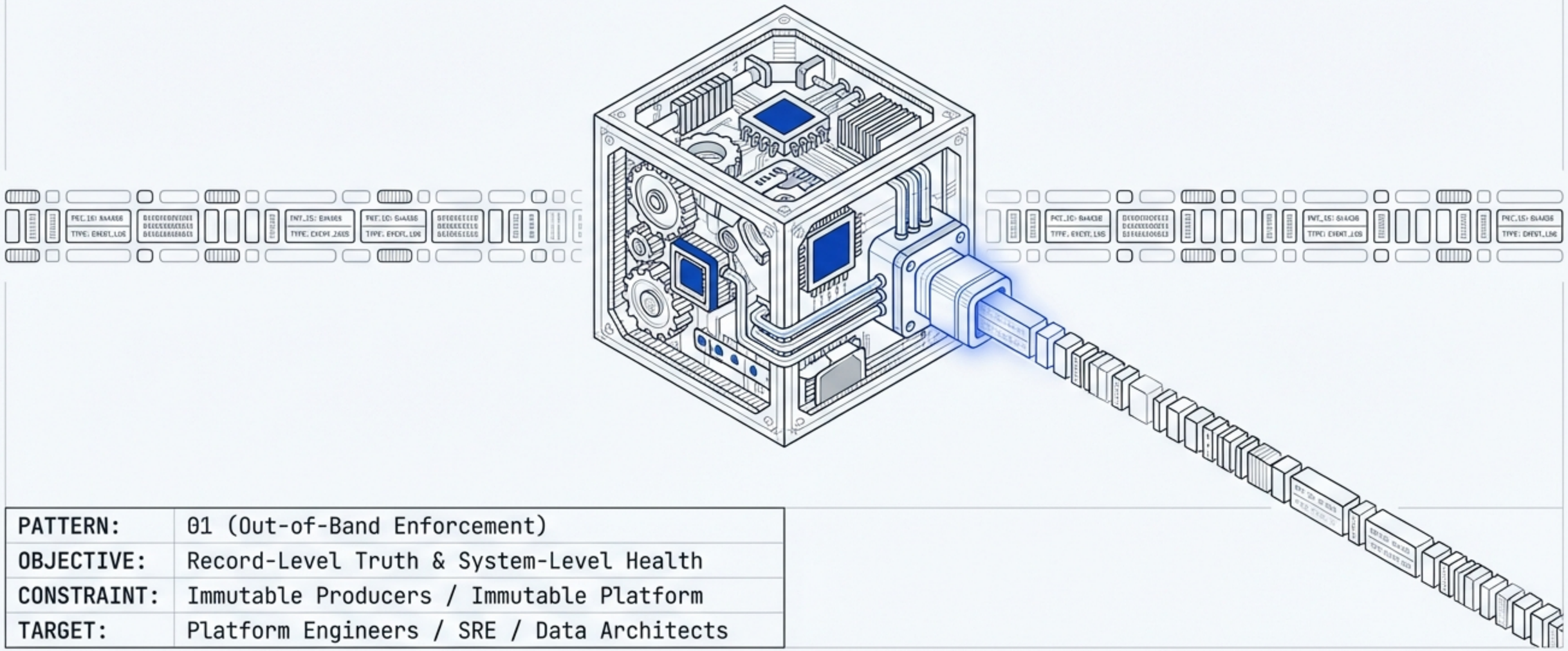
Low-Level Design: Out-of-Band Observability & Policy Enforcement

Establishing Immutable Truth in an Immutable Streaming Environment



PATTERN:	01 (Out-of-Band Enforcement)
OBJECTIVE:	Record-Level Truth & System-Level Health
CONSTRAINT:	Immutable Producers / Immutable Platform
TARGET:	Platform Engineers / SRE / Data Architects

The Constraint: Pattern 1 (Out-of-Band Enforcement)

The Challenge:

The Central Streaming Platform and Producers are immutable. We cannot modify their code, change routing logic, or block writes inline.

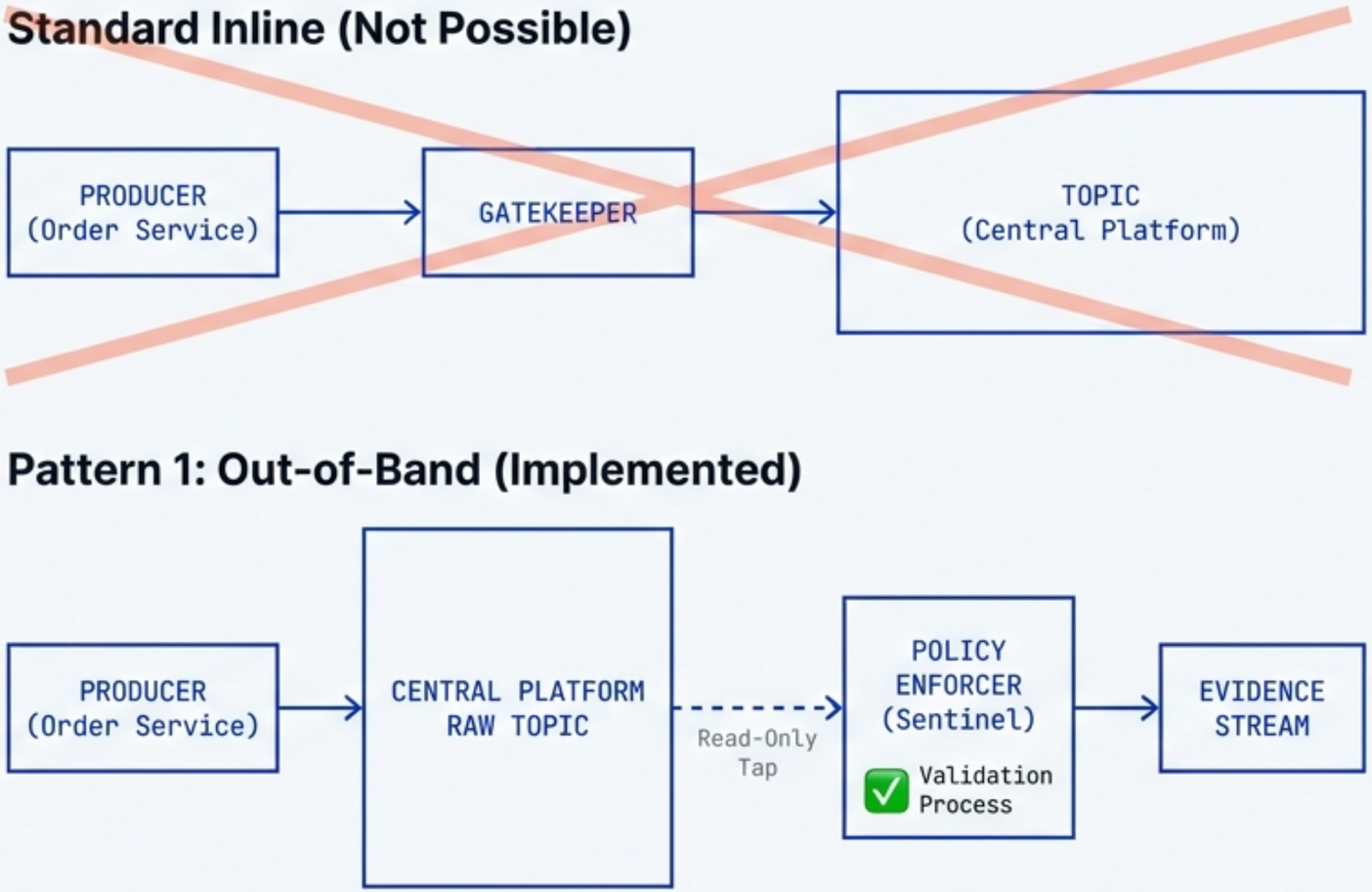
The Architectural Shift:

Move from “Gatekeeper” (blocking the write) to “Sentinel” (validating the read).

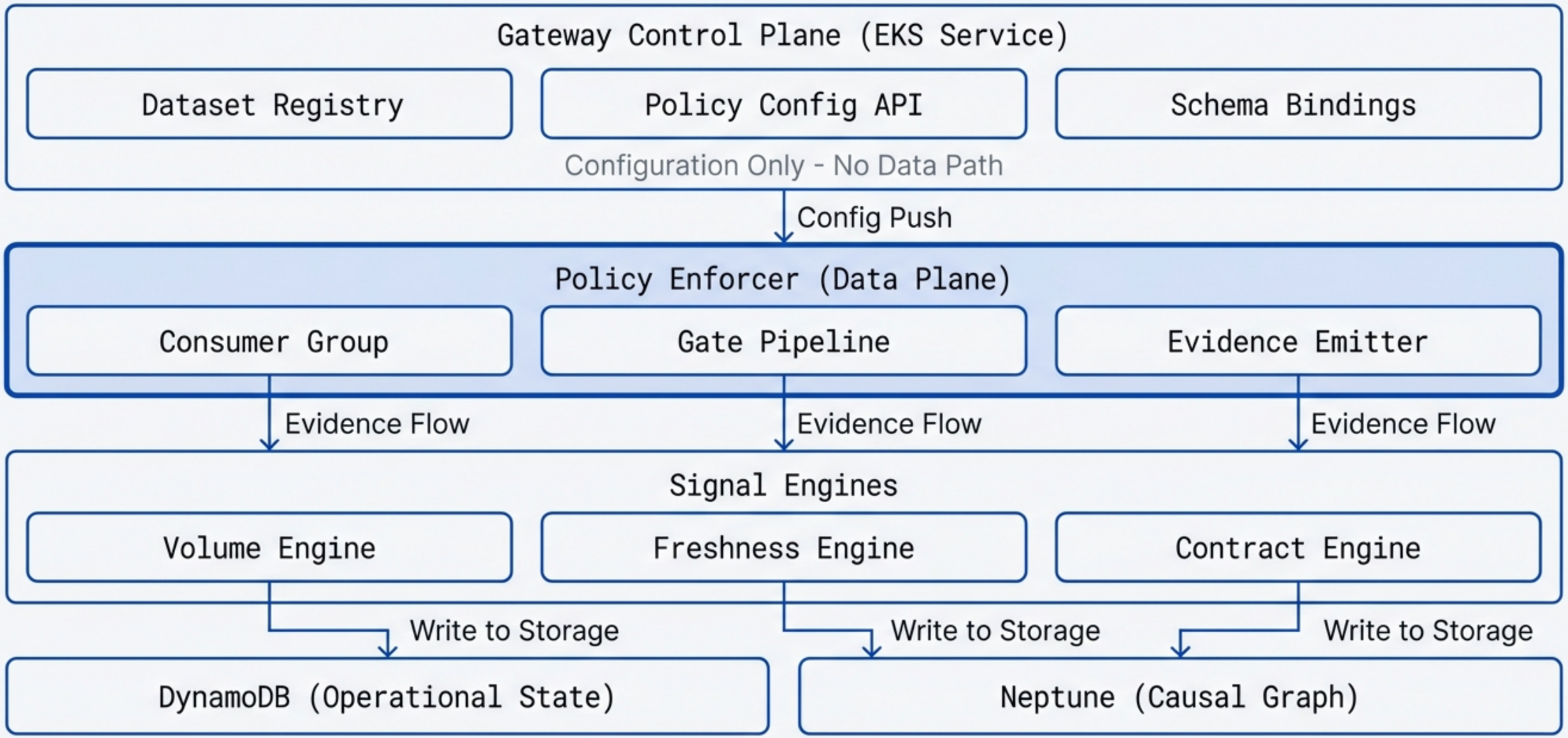
The Mechanism:

Pre-consumption enforcement. We identify and isolate bad data before it is consumed by downstream applications.

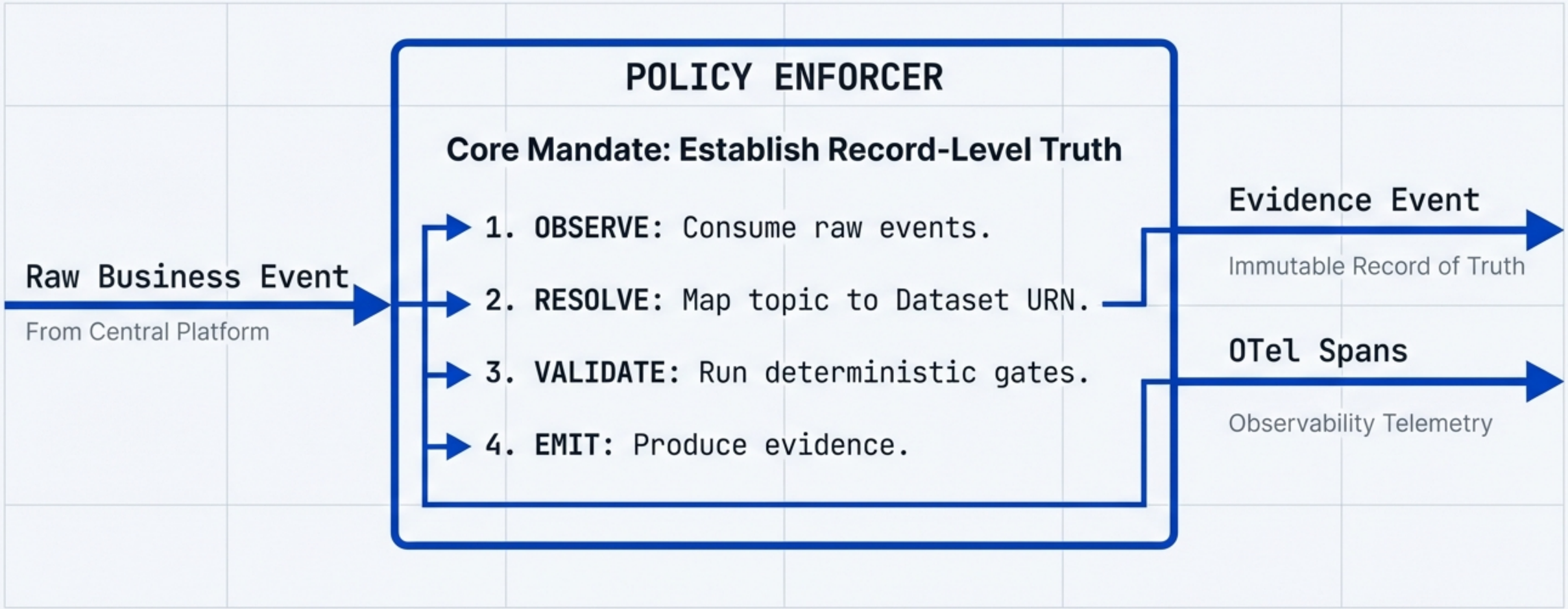
SCHEMATIC DIAGRAM: INLINE VS. OUT-OF-BAND



High-Level Component Architecture



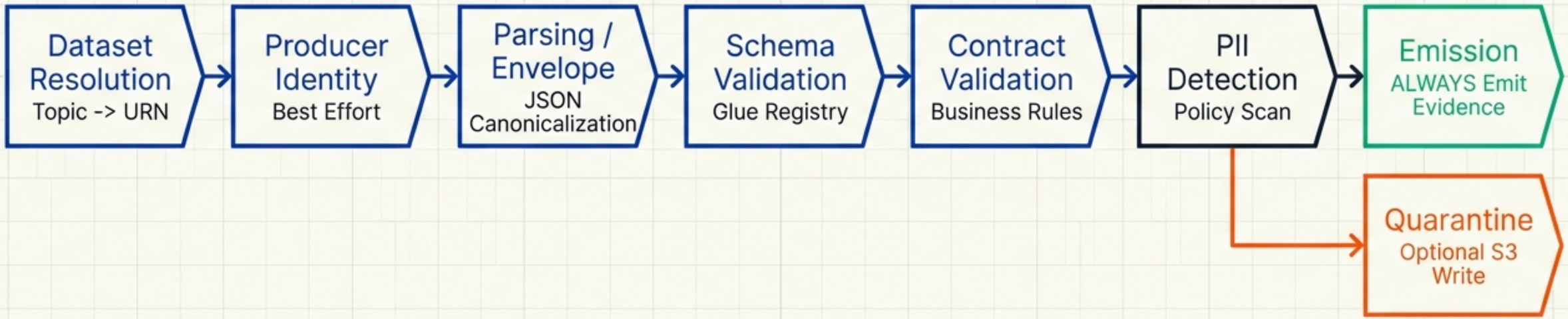
Deep Dive: The Policy Enforcer Responsibilities



HARD BOUNDARIES (What the Enforcer does NOT do):

✕ - No Aggregation (Signal Engine job) ✕ - No Alerting (Incident Mgmt job) ✕ - No Subjective Judgment (Only Facts)

The Gate Pipeline: Deterministic Evaluation



Sequential, deterministic processing for every single record.

The API Contract: Evidence Event Schema

The only stable API between the Enforcer and the rest of the world.

EVIDENCE EVENT SCHEMA

```
{
  "evidence_id": "evd-01JH...",
  "timestamp": "2026-01-14T10:01:02Z",
  "dataset_urn": "urn:dp:orders:created",
  "source": {
    "topic": "raw.orders.events",
    "partition": 7,
    "offset": 1882341
  },
  "validation": {
    "result": "FAIL",
    "failed_gates": ["CONTRACT"],
    "reason_codes": ["MISSING_FIELD:customer_id"]
  },
  "otel": {
    "trace_id": "ab91f...",
    "span_id": "c301..."
  }
}
```

Deterministic Join Key

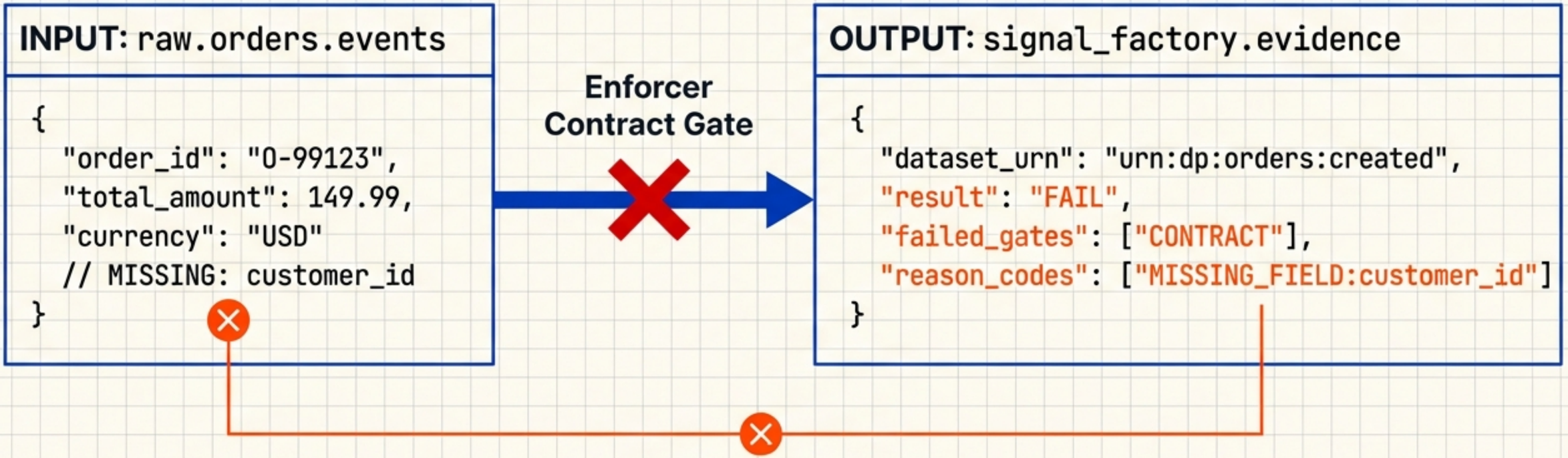
Immutable Fact of Quality

RCA Anchor (Trace ID)

Key Takeaway: One Evidence event is emitted for every raw record, regardless of validity. This creates a complete audit trail.

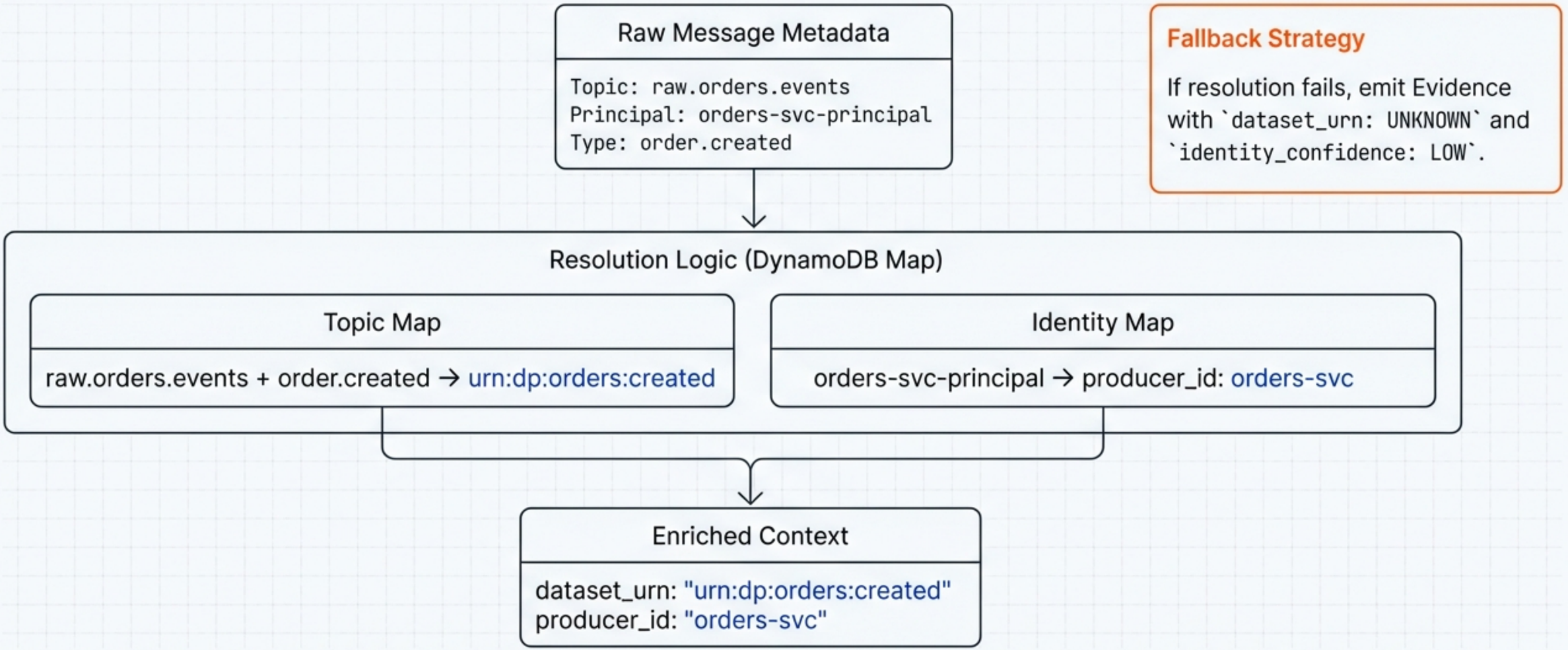
Steel Thread: The 'Order Created' Failure

Scenario: Order Service deploys bad code at 09:58 AM.
The 'customer_id' field is missing.



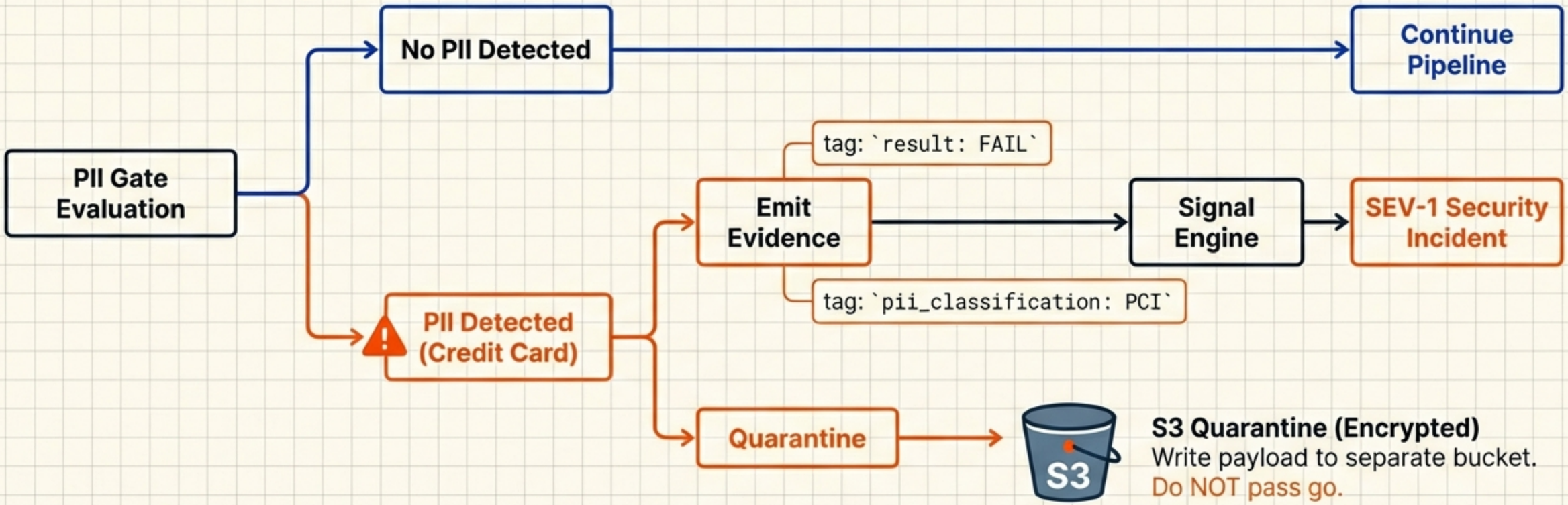
Dataset Resolution & Identity Inference

The Challenge: Out-of-band messages don't always carry headers declaring "I belong to Dataset X".



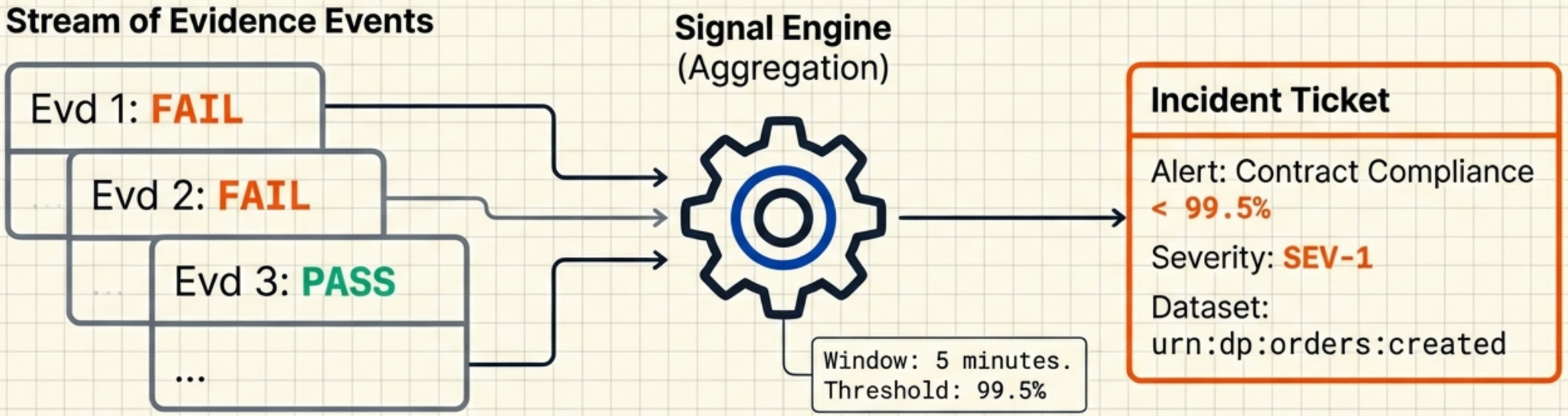
Policy Actions: PII Detection & Quarantine

Scenario: Developer logs a credit card number.
Policy is strict PCI blocking.



From Record Truth to System Health

Separation of Concerns: The Enforcer creates facts; the Signal Engine creates judgments.



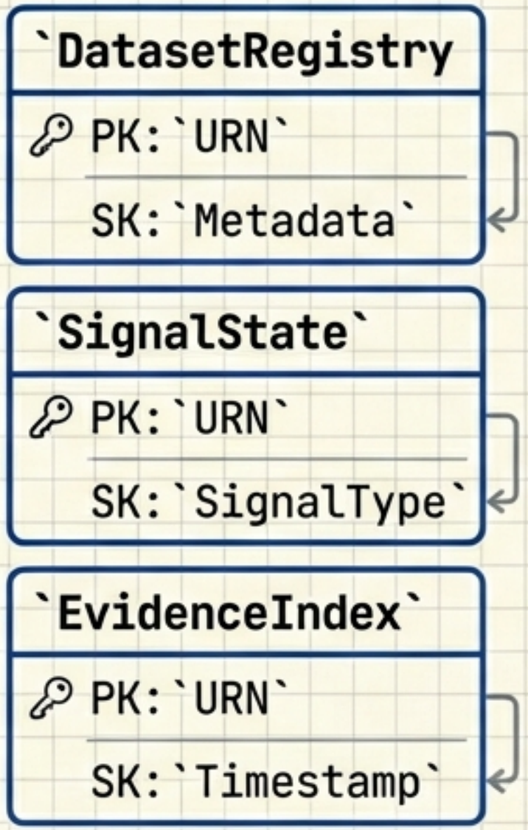
Signal Types

1. **Volume** (Rate drops/spikes)
2. **Freshness** (Time since last valid event)
3. **Contract** (% Compliance)

Data Modeling: State vs. Causality

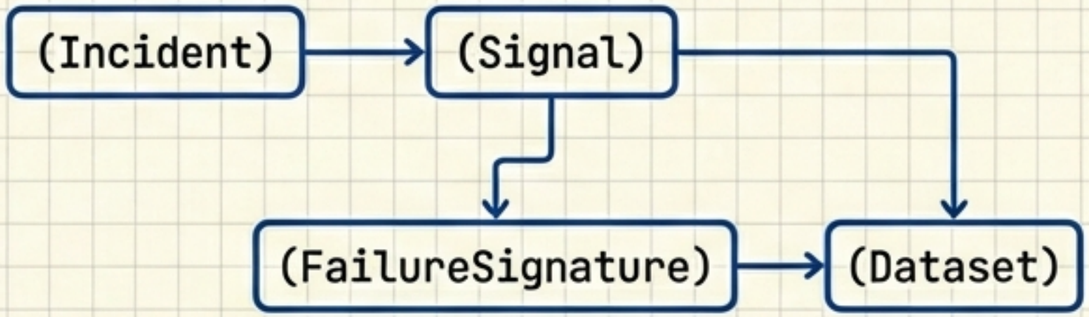
DynamoDB (Operational State)

What is happening right now?



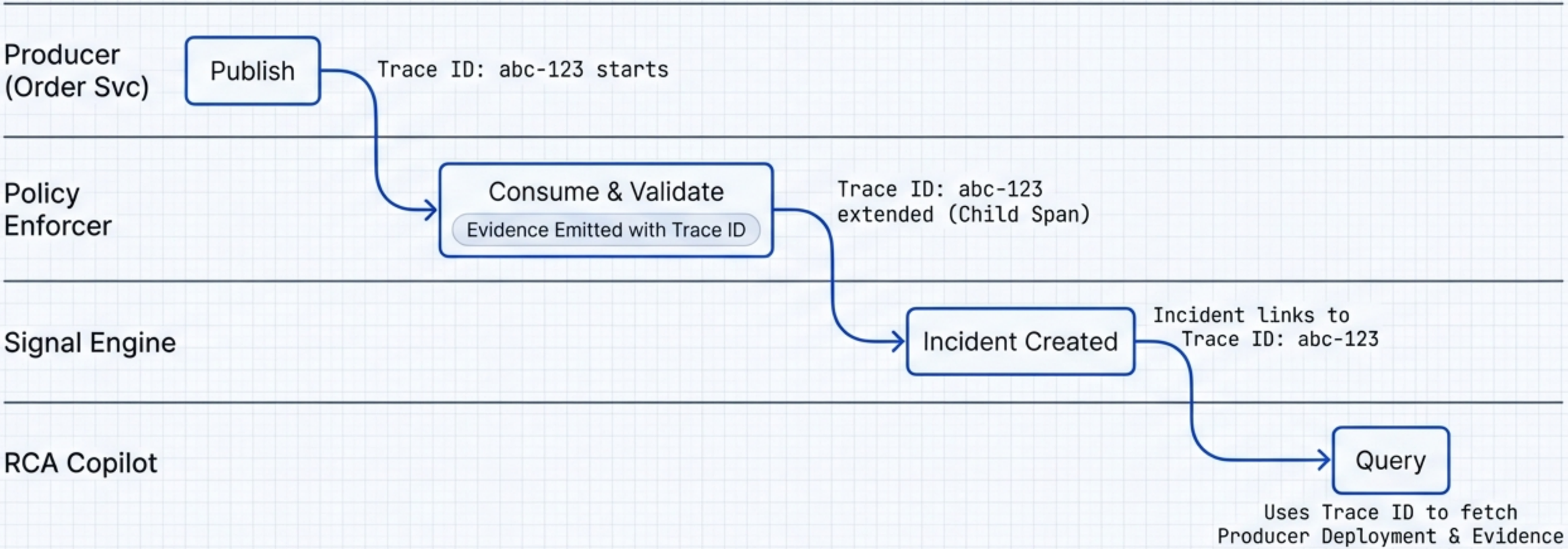
Neptune (Causal Graph)

Why did it happen?



Do NOT write every record to Neptune.
Use **FailureSignature** nodes to compress millions of failures.

Trace Propagation: The Causal Spine



Pattern 1 Implementation: If Trace exists in Raw, extend it. If not, start it. Always anchor Evidence to Trace.

Failure Modes & Reliability

Scenario	Impact	Mitigation
Enforcer Down	Evidence Gap (Blind spot)	Signal Engine detects 'No Evidence' -> Triggers `OBSERVABILITY_PIPELINE_DOWN` incident.
Schema Registry Down	Cannot validate schemas	Emit Evidence with `reason: SCHEMA_REGISTRY_UNAVAILABLE`. Fail open or closed based on policy.
Consumer Lag	Stale Signals	HPA on CPU + Lag metrics. Signal Engines must be lag-aware.

Implementation Blueprint (SRE View)

1 Deployment Stack

- ✓ EKS Managed
- ✓ HPA (Scale on Lag + CPU)
- ✓ Multi-AZ Distribution

2 Service Level Objectives (SLOs)

Completeness:	**99.99%** raw records generate evidence.
Latency:	Evidence emitted **< 2s** from ingest.

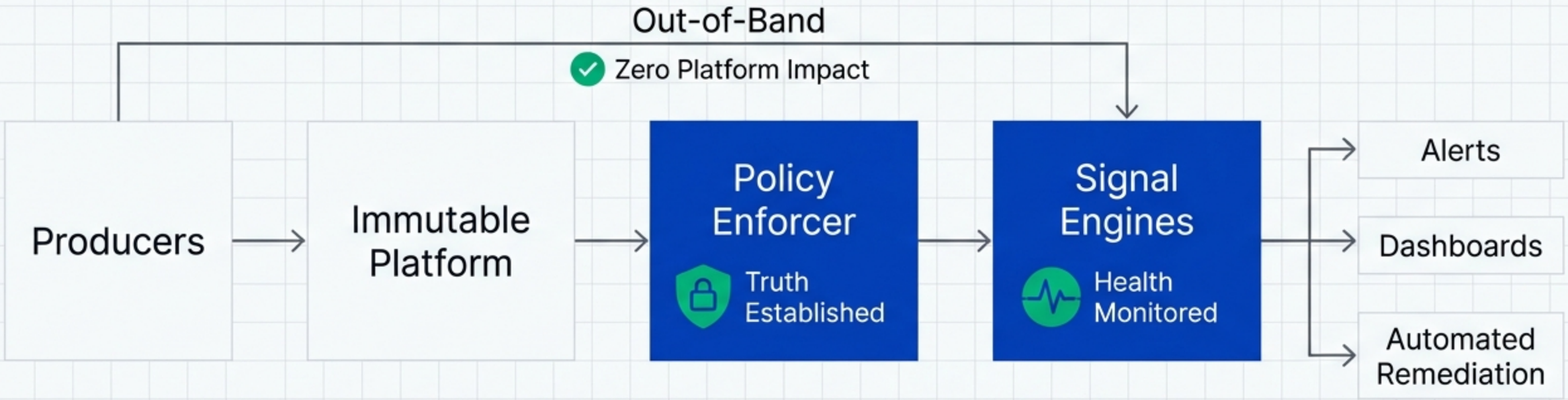
3 Cost & Storage

DynamoDB:	TTLs enabled (7 days).
Neptune:	Async writes. Sampled/Failures only.

4 Definition of Done

- ✓ Evidence schema versioned.
- ✓ Trace keys present.
- ✓ Signal Engines trusting evidence.

Summary: Immutable Truth, Decoupled Action



We achieved rigorous observability and governance without modifying the immutable platform or producers. The Policy Enforcer provides the Immutable Truth that powers the entire pipeline.