

CONF42 SRE 2026

AI-Driven

Risk-Aware Decision-Making

for Scalable Reliability Systems

Oreoluwa Omoike · Site Reliability Engineer ·

AGENDA

01

The Problem with Reactive SRE

Why detection alone isn't enough

02

Risk-Aware Decision Framework

Scoring risk in real-time with AI

03

AI Integration Patterns

Anomaly detection, SLO prediction & more

04

Production Case Study

Real results from a live deployment

05

Scaling & Pitfalls

What breaks — and how to avoid it

THE PROBLEM WITH REACTIVE SRE

Detecting failures after they happen is already too late.

14 min

Mean Time to Detect

avg incident detection lag

70%

Alert Fatigue

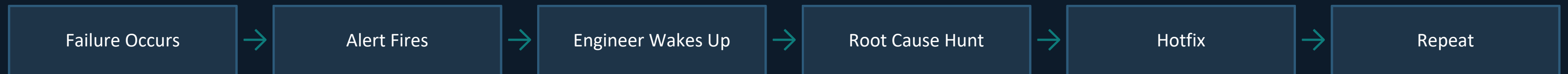
of all alerts are noise

45 min

Manual Runbooks

avg time-to-mitigate

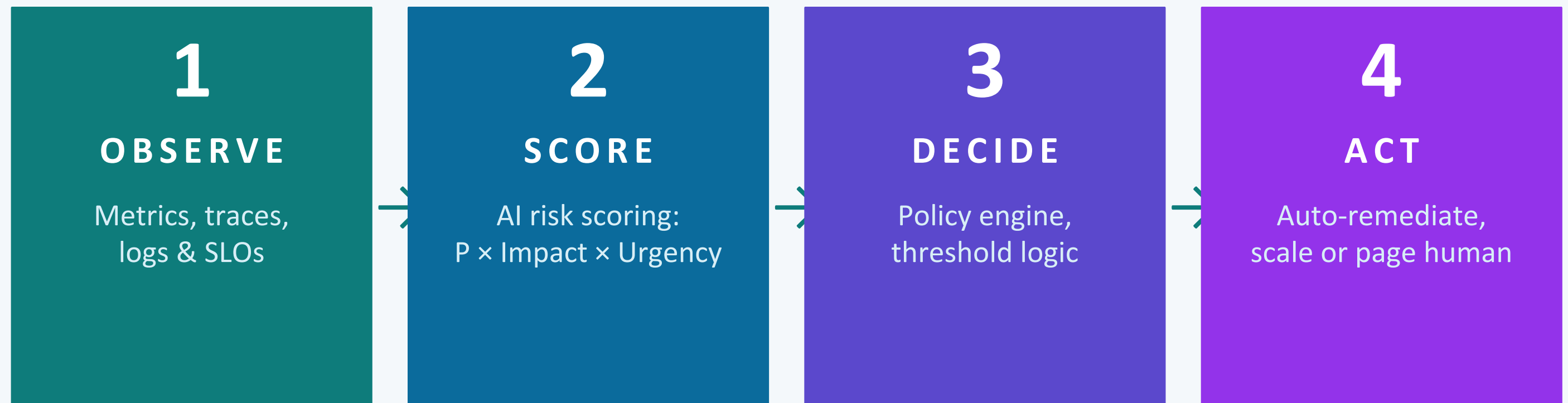
The Reactive Loop



SRE needs to shift from reactive alerting → proactive, risk-aware decision intelligence.

RISK-AWARE DECISION FRAMEWORK

Four pillars: Observe → Score → Decide → Act



Risk Score Formula

$$\text{Risk Score} = \text{P(failure)} \times \text{Business Impact} \times (1 / \text{Time-to-Impact})$$

P(failure) = ML-predicted probability **Business Impact** = revenue / user-hours at risk **TTI** = urgency window

AI INTEGRATION PATTERNS



Anomaly Detection

Isolation Forest / LSTM

Spots metric drift before your SLO actually breaches — trained on baseline traffic, triggers when the score crosses your threshold.



SLO Burn Rate Prediction

Time-series Forecasting

Forecasts error budget exhaustion 30–60 min ahead, giving you time to preemptively scale or shift traffic — not react to a breach.



Change Risk Scoring

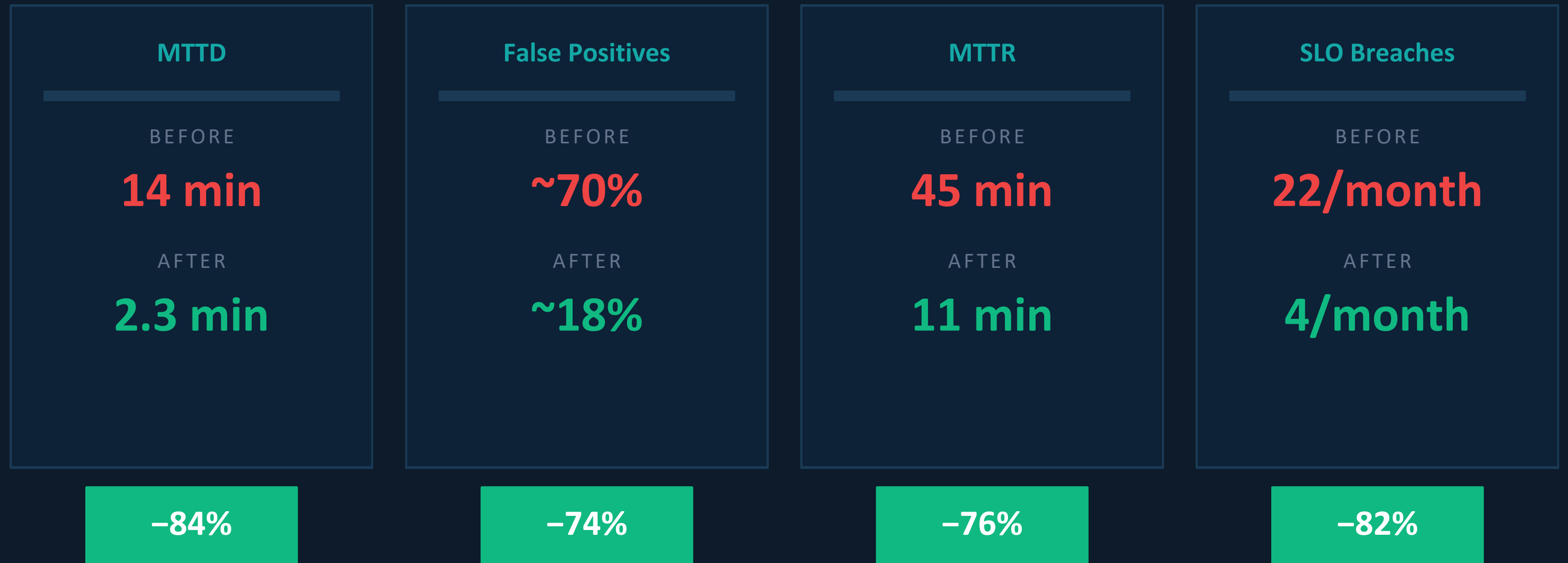
Gradient Boosted Trees

Every deploy, config change, or feature flag gets a blast-radius score before it ever touches production.

All models run as sidecars alongside your existing observability stack — Prometheus, OpenTelemetry, Grafana.

PRODUCTION CASE STUDY

E-commerce platform · 50M events/day · 6-month deployment



“*The AI risk scorer became our most trusted on-call team member. It pages us before our customers notice anything.*

— SRE Lead, 2025

SCALING & PITFALLS



Model Drift

Retrain weekly on a rolling window. Track drift with PSI (Population Stability Index) before degraded scoring sneaks in.



Alert Storm Risk

Start in log-only mode. Graduate to auto-scale before page-oncall. Always keep a human-in-the-loop override.



Latency Overhead

Batch score asynchronously. Target under 5ms per event. Use feature caching for high-cardinality label lookups.



Data Quality

Garbage in, garbage out. Validate metric completeness before inference — 30% missing means don't score.

KEY TAKEAWAYS

1

Reactive SRE is technical debt. AI-driven risk scoring shifts you from detection to prevention.

2

$P(\text{failure}) \times \text{Impact} / \text{TTI}$ gives you one actionable signal — not a dashboard full of noise.

3

Start simple. One Isolation Forest model + a Prometheus metric is enough to begin.

4

Automate low-risk actions first. Build trust before granting full autonomy.

5

Measure everything: MTTD, MTTR, false-positive rate, and error budget burn.
