

Predictive Workflow Integrity for SRE: Autonomous Triage at Global Scale

ROHIT WADHWA

CONF42 SITE RELIABILITY ENGINEERING (SRE) 2026

WALMART INC



Modern Enterprise Platforms: Scalability and Complexity

EVENT-DRIVEN MICROSERVICES

Today's **distributed microservices architecture** enables organizations to scale rapidly while managing complexities across global deployments effectively.



SRE Challenges at Scale

STATE DIVERGENCE

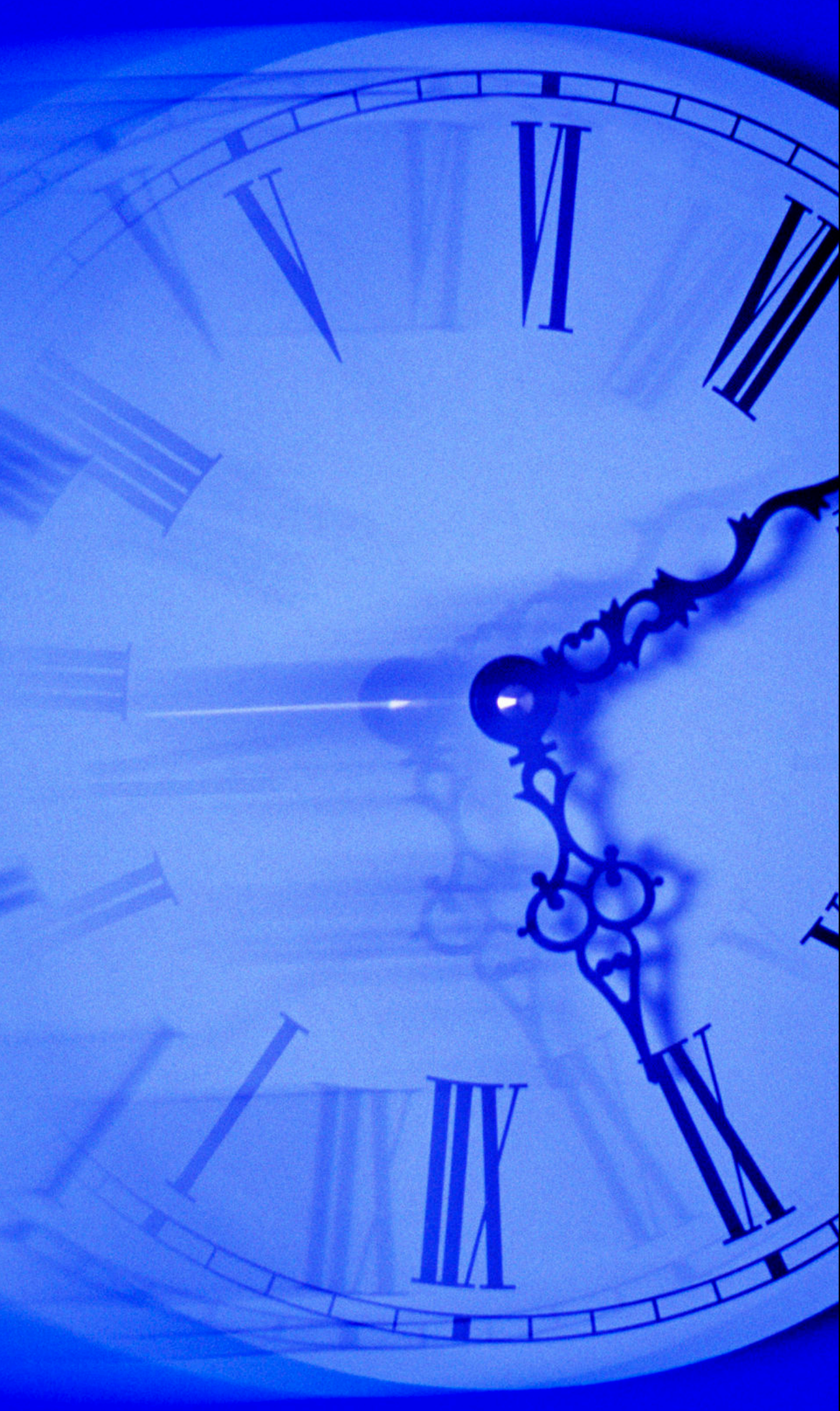
State divergence occurs when distributed services operate inconsistently, leading to unpredictable behavior and complicating troubleshooting across the system.

LATENCY VARIABILITY

Latency variability can cause inconsistent response times, impacting user experiences and complicating the performance monitoring of event-driven architectures.

CASCADING FAILURES

Cascading workflow failures result when one service's failure triggers a domino effect, leading to widespread outages across interconnected systems.



Mean time to detection: ~12 minutes

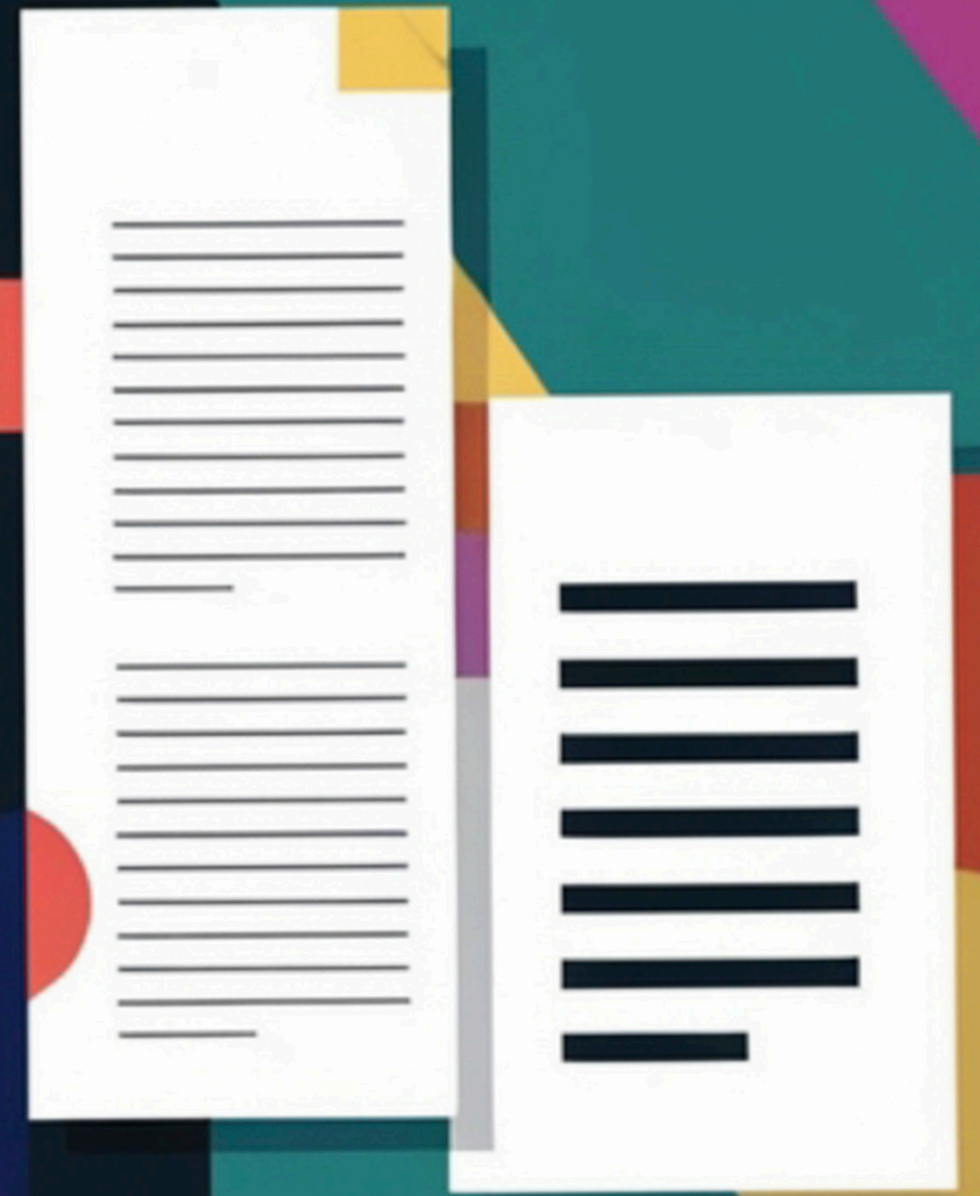
DETECTION TIME

The **mean time to detection** (MTTD) of ~12 minutes highlights the inefficiency of reactive monitoring. This delay can lead to prolonged service disruptions and impacts overall system reliability. By shifting to more proactive monitoring strategies, organizations can significantly improve response times and reduce operational risks.

Introducing Predictive Workflow Integrity for SRE

SHIFT TO WORKFLOW-LEVEL OBSERVABILITY

Predictive Workflow Integrity (PWI) moves from traditional service-level monitoring to **workflow-level observability**, enabling better detection and management of complex interactions.



PWI Anomaly Detection

DELAYED CONVERGENCE

Delayed convergence occurs when workflows do not reach their expected states within the designated time, leading to potential operational disruptions.

CONFLICTING OWNERSHIP

Conflicting ownership arises when multiple teams manage the same workflows, causing confusion and inefficiencies in resolving issues that may occur.

ABNORMAL EXECUTION PATHS

Abnormal execution paths refer to unexpected workflow routes that can create instability, resulting in failures or degraded performance of services.



MTTD reduced to 58 seconds

MTTD

The implementation of Predictive Workflow Integrity (PWI) has significantly improved detection times. By reducing the **mean time to detection (MTTD)** from around 12 minutes to just **58 seconds**, teams can respond to incidents much more rapidly, enhancing overall operational efficiency and reliability in service delivery.

Understanding Autonomous Incident Triage Mechanisms

REAL-TIME EVENT CORRELATION

Autonomous incident triage correlates logs, metrics, and traces, building a unified event-driven causal graph for enhanced real-time impact assessment. By continuously analyzing these diverse data sources, the system can identify patterns and anomalies that indicate potential issues or threats. This proactive approach helps in prioritizing incidents based on their severity and potential impact, allowing for faster and more efficient resolution. The integration of machine learning algorithms further enhances the system's capability to learn from past incidents, improving its accuracy and adaptability over time. As a result, organizations can achieve a more resilient and responsive IT infrastructure, minimizing downtime and optimizing resource allocation.

Autonomous Incident Triage

UNIFIED GRAPH

The **unified causal graph** integrates logs, metrics, and traces, providing a comprehensive view of system interactions for improved diagnostics.

REAL-TIME PRIORITIZATION

Real-time impact prioritization allows **SRE teams** to address critical incidents promptly, ensuring operational efficiency in high-demand environments.

AUTOMATED REMEDIATION

Supports automated remediation aligned with **SRE goals**, reducing manual intervention and enabling faster recovery from incidents to maintain service continuity.

Ring-Based Geolocation Routing Strategy

ENHANCING EVENT PROCESSING

This approach **localizes event processing** to improve determinism during regional degradation, ensuring swift response times and enhanced service reliability.





80%

LATENCY IMPROVEMENT

The implementation of geolocation routing has achieved a **significant reduction** in edge-to-cloud latency, decreasing it from approximately 5 seconds to under 1 second. This improvement enhances overall system responsiveness and user experience, ensuring that services are delivered **efficiently** and reliably across global platforms.

Highlights of Large-Scale Enterprise Deployment

COMPLEXITY AT SCALE

The deployment features **thousands of nodes** and manages high-volume event streams, showcasing the intricate architecture of modern enterprise systems.



Practical SRE Takeaways

WORKFLOW OBSERVABILITY

Focus on **end-to-end workflow observability** to gain insights into system performance and dependencies, leading to improved operational efficiency.

PREDICTIVE DETECTION

Integrate **predictive anomaly detection** into your monitoring systems for faster response times and to minimize downtime during critical events.

AUTONOMOUS TRIAGE

Leverage **autonomous triage mechanisms** to effectively reduce alert noise, streamline incident management, and enhance operator trust and efficiency.

Summary

EMPHASIZING PREDICTIVE RESILIENCE

Predictive Workflow Integrity transforms SRE practices, enabling **self-regulating systems** that enhance operational efficiency and improve service reliability across global platforms.



Thank You for Your
Attention!