



# DIVING THROUGH RCAs: A QUEST FOR RELIABILITY

DIANA TODEA - SRE

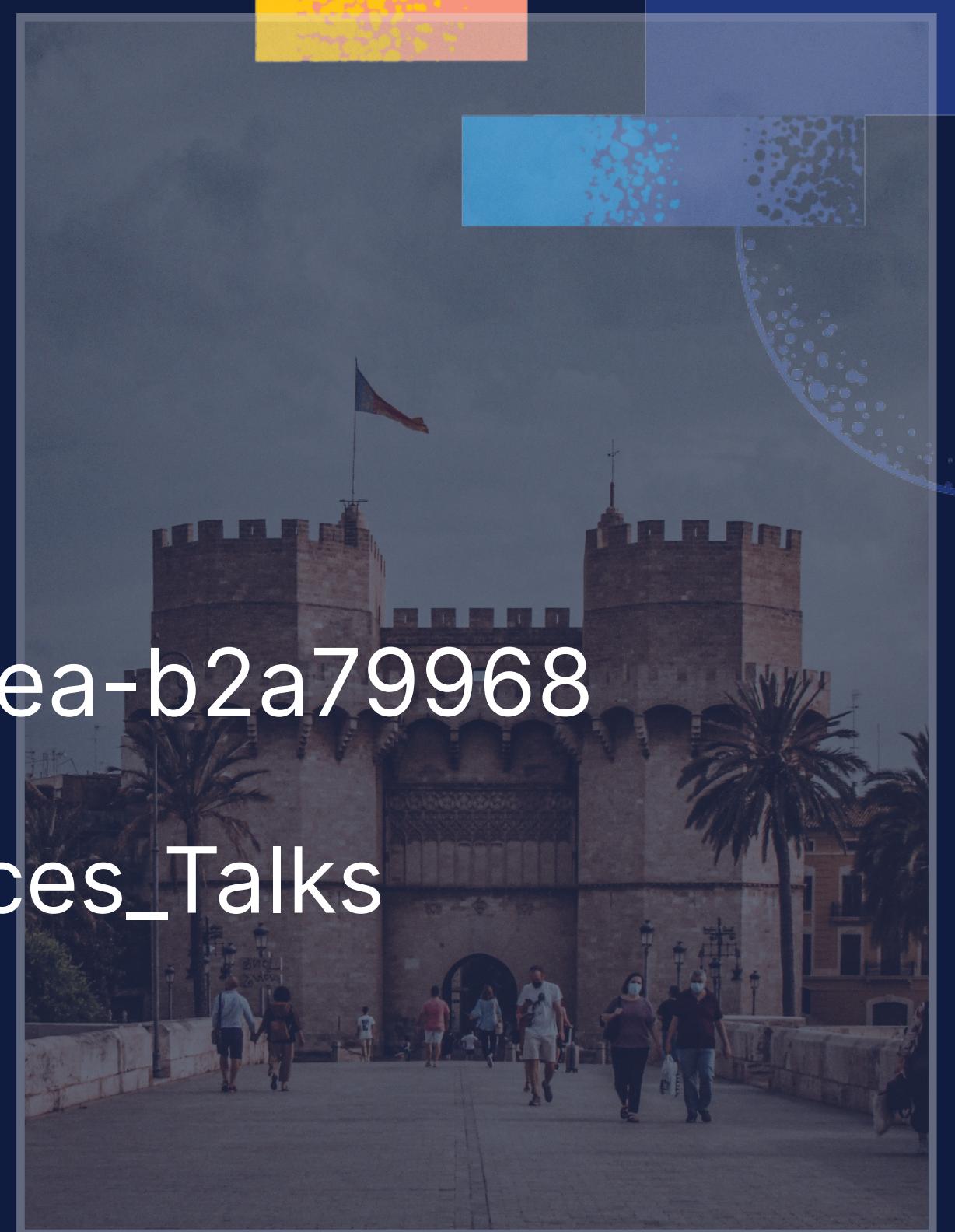
DEVOPSDAYS GENEVA MAY 2024



DIANA TODEA - OBSERVABILITY SRE

<https://www.linkedin.com/in/diana-todea-b2a79968>

[https://github.com/didiViking/Conferences\\_Talks](https://github.com/didiViking/Conferences_Talks)



UNSPASH [Juan Puyo](#)

**ALERTS**



**INCIDENT**

**INCIDENT CHANNEL**

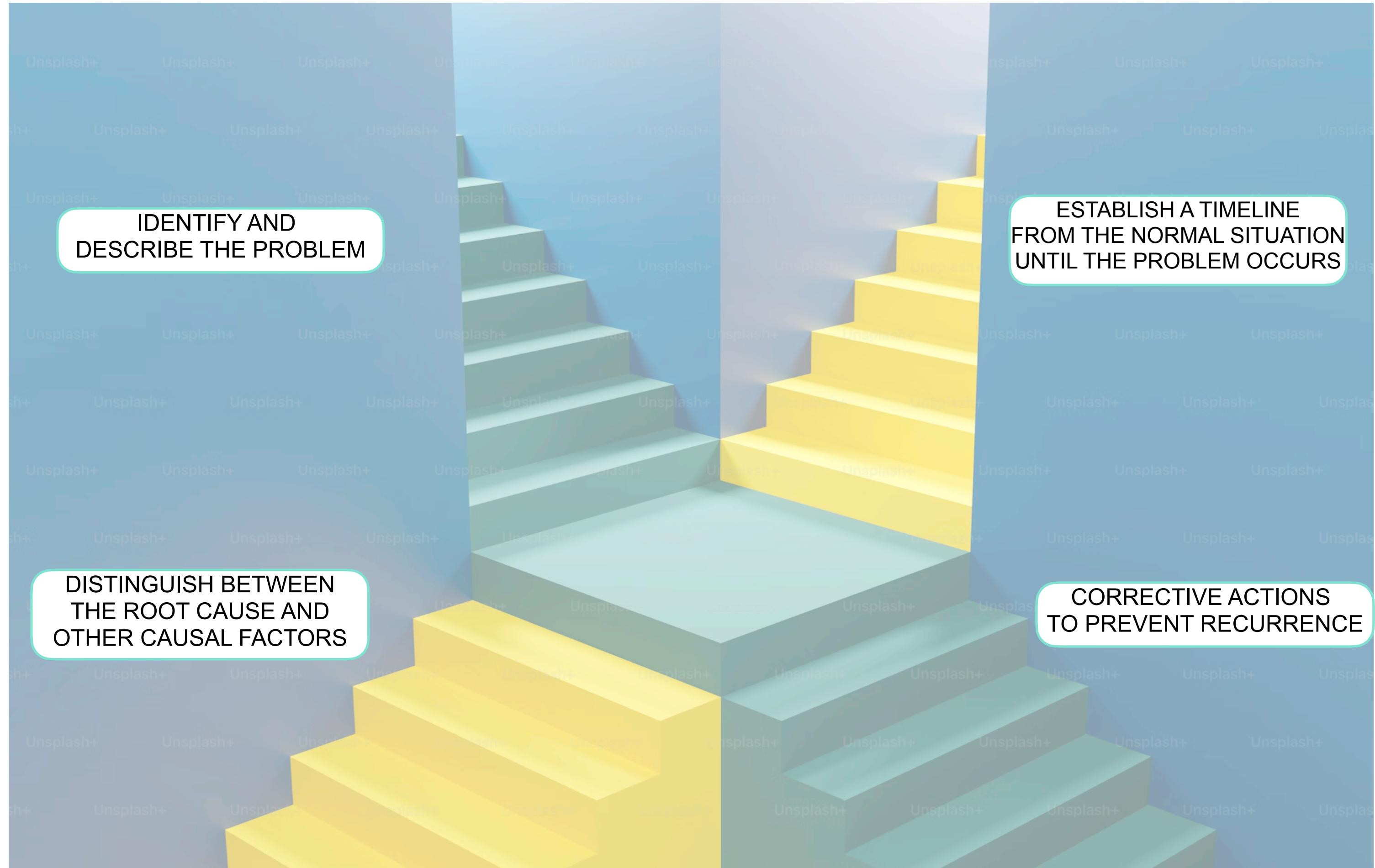
**CUSTOMERS**

**MANAGERS**

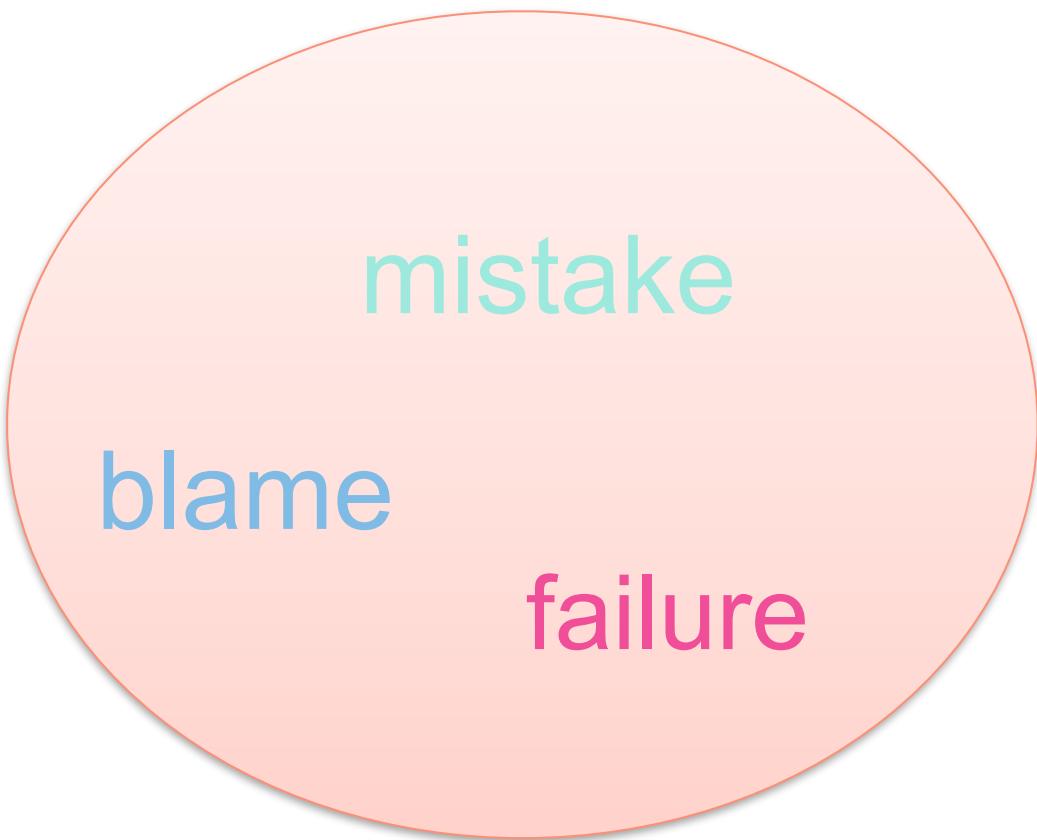
# WHAT IS ROOT CAUSE ANALYSIS?

Root cause analysis (RCA) is the process of discovering the root causes of problems in order to identify appropriate solutions.

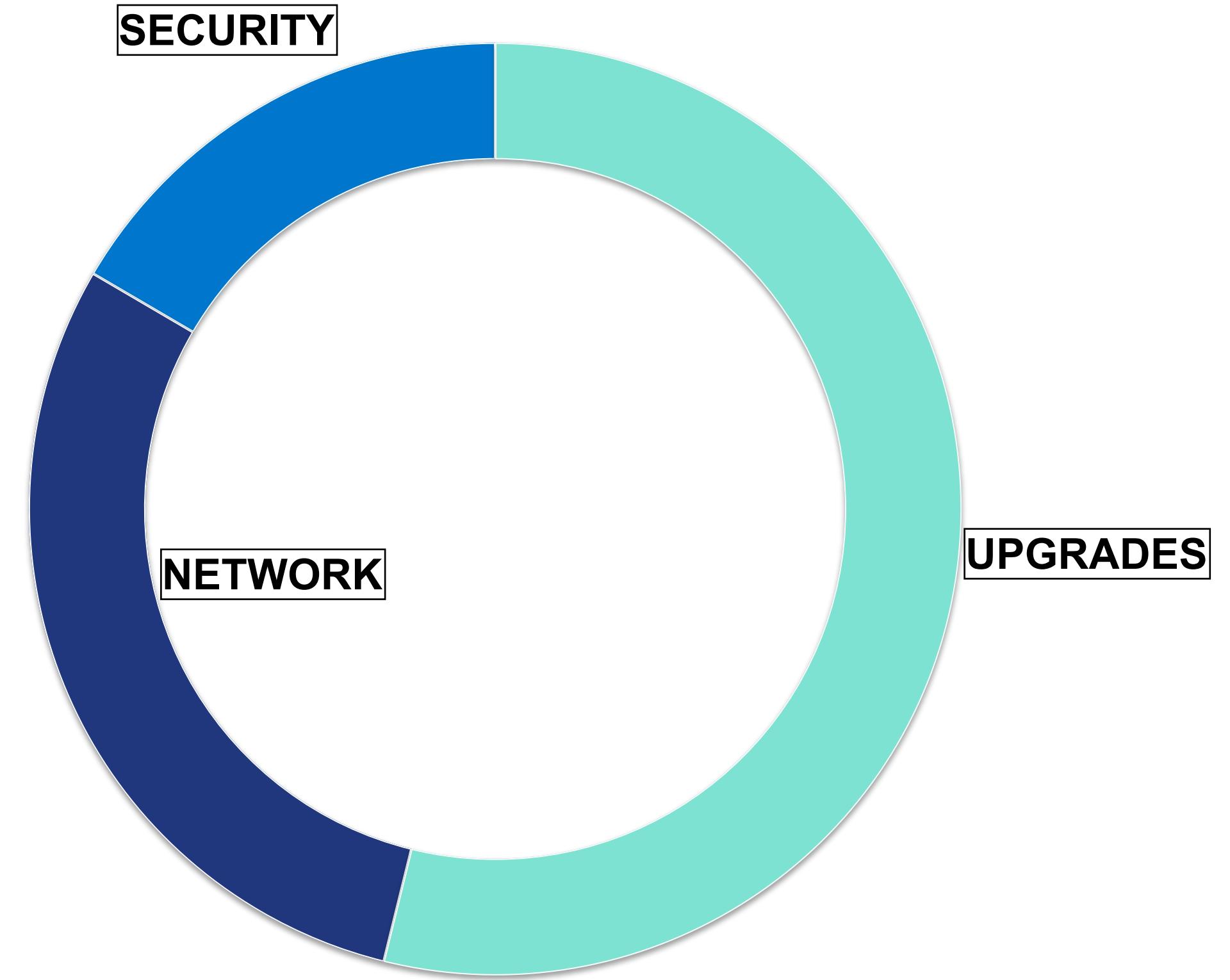
SOURCE: TABLEAU



# POSTMORTEM CULTURE



# TOP SEVERITY 1 AND 2 INCIDENTS' ROOT CAUSES



## 5 WHYS METHODOLOGY

1K nodes affected in a region

Incident duration: 5h

Impact duration: 7h

Impacted allocators: 96

Affected customer  
deployments: 1042

Affected organizations: 750

**Live Linux kernel patch**

**Affected network functionality on allocators**

**Impact on inter-node communication for deployments**



CUS  
TOM  
ERS

OBS  
ERV  
ABIL  
ITY

SEC  
URI  
TY

# FOCUS ON OBSERVABILITY

MOVE AWAY FROM  
3RD PARTY TOOLS

CREATE UPTIME  
CHECKS  
(SYNTHETIC  
MONITORING)

CREATE SLOS AND  
BURN RATE ALERTS

INTEGRATE WITH  
ELASTIC  
OBSERVABILITY AI  
ASSISTANT



# Observability

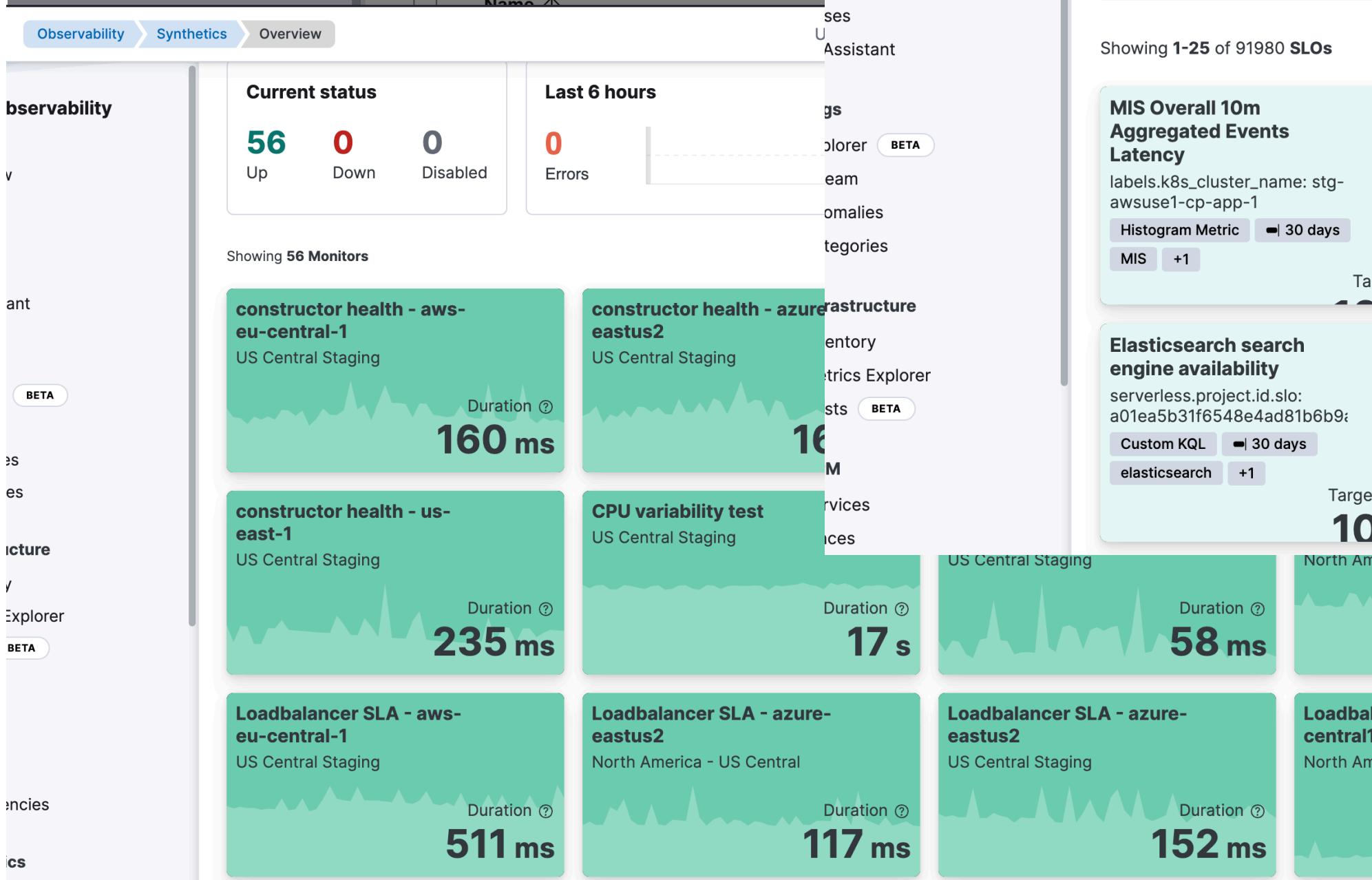
## Rules

Rules Logs

Error found in 13 rules. Show rules with errors

Succeeded: 285 • Failed: 13 • Warning: 0

298 rules



## Create rule

### Custom threshold

Alert when any Observability data type reaches or exceeds a given value. [Learn more](#)

# Observability

## SLOs

Showing 1-25 of 91980 SLOs

Sort by SLI value ▾ Group by None ▾

MIS Overall 10m Aggregated Events Latency labels.k8s_cluster_name: stg-awsuse1-cp-app-1 Histogram Metric   30 days MIS +1 Target 99% 100%	MIS Overall 60m Aggregated Events Latency labels.k8s_cluster_name: stg-awsuse1-cp-app-1 Histogram Metric   30 days MIS +1 Target 99% 100%	MIS Overall 1m Aggregated Events Latency labels.k8s_cluster_name: stg-awsuse1-cp-app-1 Histogram Metric   30 days MIS +1 Target 99% 100%	MIS Index Service Latency labels.k8s_cluster_name: stg-awsuse1-cp-app-1 Histogram Metric   30 days MIS +1 Target 99.6% 100%
Elasticsearch search engine availability serverless.project.id.slo: a01ea5b31f6548e4ad81b6b9a Custom KQL   30 days elasticsearch +1 Target 99.95% 100%	Elasticsearch search engine availability serverless.project.id.slo: a056f5032e5040c5a41824ac Custom KQL   30 days elasticsearch +1 Target 99.95% 100%	Elasticsearch search engine availability serverless.project.id.slo: a15ce1692d27460195faedade Custom KQL   30 days elasticsearch +1 Target 99.95% 100%	Elasticsearch search engine availability serverless.project.id.slo: a19ec3d18dbb4fdbba3c085e2 Custom KQL   30 days elasticsearch +1 Target 99.95% 100%
US Central Staging Duration 58 ms	North America - US Central Duration 503 ms	US Central Staging Duration 17 s	US Central Staging Duration 152 ms

elastic

elastic

Find apps, content, and more.

Add integrations AI Assistant

Observability SLOs

SLOs

Tell us what you think! Stop refreshing Create new SLO

Search your SLOs... Sort by slo status

Compact view

cartservice-otel APM availability cartservice-... 7 days Target 99.9% 99.519%

Checkout SLO APM availability checkoutSe... 7 days Target 95% 0%

cartservice-otel-latency APM latency cartservice-... 7 days Target 99% 99.498%

Rows per page: 25 1

Logs

- Explorer BETA
- Stream
- Anomalies
- Categories

Infrastructure

- Inventory
- Metrics Explorer
- Hosts BETA

APM

- Services
- Traces
- Dependencies

Synthetics

- Monitors
- TLS Certificates

Uptime

- Uptime Monitors
- TLS Certificates

User Experience

- Dashboard

Universal Profiling

The screenshot shows the Elastic Observability SLOs dashboard. On the left, there's a sidebar with various monitoring categories like Logs, Infrastructure, APM, Synthetics, Uptime, User Experience, and Universal Profiling. The main area displays three SLO cards. The first card, 'cartservice-otel', has a red background and shows a performance metric of 99.519% against a target of 99.9%. It includes tags for APM availability and cartservice-. The second card, 'Checkout SLO', has a pink background and shows 0% against a target of 95%, with tags for APM availability and checkoutSe... The third card, 'cartservice-otel-latency', has a green background and shows 99.498% against a target of 99%, with tags for APM latency and cartservice-. There are also buttons for creating new SLOs and stopping refreshes.



# FOCUS ON INFRASTRUCTURE

REMOVE UNUSED  
ALLOCATORS

UPGRADE TO  
HIGHER INSTANCE  
FAMILIES

CHECK THE BILL AND  
CLEAN UP



# FOCUS ON DOCUMENTATION

UPDATE  
DOCUMENTATION

UPDATE  
RUNBOOKS

WRITE NEW  
PROCESSES

GET INVOLVED  
WITH THE  
COMMUNITY



# FOCUS ON INCIDENT MANAGEMENT

SPEND **LESS TIME**  
IN IDENTIFYING  
INCIDENT  
SEVERITY

INTEGRATE  
RUNBOOKS AND  
DOCUMENTATION

SPEND **MORE TIME**  
IN FIXING THE  
ISSUE

AUTOMATE YOUR  
INCIDENT  
MANAGEMENT  
PLATFORM

# FOCUS ON SYSTEM DESIGN

SERVERLESS  
ES|QL  
eBPF-BASED PROFILING AGENT DONATION TO OTEL  
OBSERVABILITY AI ASSISTANT



**EMPOWER**

**UPSKILL**

People

**OBSERVE**

**DECIDE & REACT**

# RESOURCES

- <https://sre.google/sre-book/postmortem-culture/>
- <https://github.com/elastic/elasticsearch/>
- <https://www.elastic.co/guide/en/elasticsearch/reference/current/esql.html>
- <https://www.elastic.co/elasticsearch/ai-assistant>
- <https://www.elastic.co/search-labs>
- <https://docs.elastic.co/serverless>
- <https://www.elastic.co/blog/elastic-donation-proposal-to-contribute-profiling-agent-to-opentelemetry>
- <https://www.elastic.co/observability/opentelemetry>

QUESTIONS?  
THANK YOU

YOUR FEEDBACK HERE!

