

DevOps Agent & WGU

Ethan Hall | Manager, Cloud Engineering
Jarad Dingman | Staff Site Reliability Engineer

Western Governors University

A nonprofit, fully online university serving roughly 200,000 students.

~200,000

Students served nationwide

Nonprofit

Fully online university

Technology

Is the university experience itself

Always on

Students depend on systems being available when they need them

WHY IT MATTERS

Reliability is directly connected to student success.

Before the DevOps Agent

Every investigation waits on people to assemble before it can begin.

- 1 Dynatrace detects a problem
- 2 Incident is opened
- 3 Engineers and service owners are contacted
- 4 Teams assemble in a bridge or chat
- 5 Logs, metrics, deployments, and dependencies are reviewed
- 6 Hypotheses are tested
- 7 Root cause and mitigation are identified

How the solution works

Dynatrace provides the signal. The AWS DevOps Agent expands the investigation.

- 1 Dynatrace detects the problem
- 2 Investigation is triggered automatically
- 3 Topology and telemetry are analyzed
- 4 AWS resources and relationships are examined
- 5 Findings, root cause, and mitigation returned

Human-controlled, agent-assisted



Investigation 2026-09-21T15:35:35.161Z

Add feedback
Ask for human support

Triggered by User at 2026-09-21 09:36:09 MDT

Summary

Investigation timeline
Root cause
Mitigation plan

Platform-wide outage: RabbitMQ broker unreachable via NLB ([REDACTED] API + 8 services) complete Severity: critical

Incident

One platform-wide event — 59 of 65 active Dynatrace problems, synchronized ~15:01Z onset across 8+ independently-owned backend services plus the [REDACTED] API. ~17 global HTTP-monitor outages indicate real user impact. The client-side MyWGU JavaScript errors (which Davis mislabeled 'Ping') are a downstream manifestation of the backend failure, not a separate CDN issue.

Cause

CONFIRMED ROOT CAUSE — the shared self-managed RabbitMQ cluster [REDACTED] (3x EC2 in [REDACTED]/us-west-2 across us-west-2a/2b/2c, fronted by internal NLB [REDACTED] on port 5672) gracefully shut down at 15:01:37Z (connection_forced 'shutdown' across vhosts /wgu, /ne, /fdp, /test, /), then the rabbit application stopped (mnesia exited ~15:07Z) and got stuck in a cluster restart requiring rabbitmqctl force_boot (~15:07–15:09Z, quorum/last-disc-node loss). The OS hosts and beam.smp process stayed up (availability read 1.0, CPU still flowing) but the 5672 listener stopped accepting connections — so all 3 NLB targets went unhealthy on 5672 and every AMQP client got java.net.ConnectException: Connection refused from ~15:01Z. All 8 backend victims (503 via RabbitHealthIndicator, or 500 from the event publisher) and the Degree Plan API share this one signature. Ruled out: security-group change (SG sg-0a27f14276940172d intact), active ARC zonal shift (none; the zonal_shift_delegated_to_dns override is stale), Amazon MQ (no managed brokers anywhere), PingFederate backend, and any GitHub deploy in the window.

Mitigation

Restore the RabbitMQ cluster — complete the force_boot / restart the rabbit application on the 3 nodes so the 5672 listener accepts connections and the NLB targets return healthy. All 3 nodes are SSM Online, so recovery can be driven via SSM Run Command; dependent services recover automatically once the broker is reachable.

Findings & evidence

root cause

RabbitMQ cluster shut down and got stuck in a force_boot restart, leaving all NLB targets unhealthy on 5672

hypothesis

Client-side JavaScript failure on the Ping/MyWGU SSO login front door — ruled out

hypothesis

GitHub deployment ruled out as trigger — ruled out

symptom

Platform-wide correlated failure storm (59 of 65 active problems) — 2026-09-21T15:01:00Z → ongoing

symptom

[REDACTED] API increased failure rate (P-26097240) — 2026-09-21T15:19:00Z → ongoing

gap

PingFederate AWS account [REDACTED] inaccessible

After the DevOps Agent

Investigation starts before the bridge fills up.

77%

Year-over-year reduction in mean time to resolution

~28 min

Average resolution time for recent priority incidents

99.83%

FY26 Q4 system availability

14 hours

Annualized downtime, against a 30-hour goal

THE MOST IMPORTANT CHANGE

The investigation can produce an answer before all the engineers are in the room.

Questions