Commas Save Lives

Or At Least LinkedIn

TODD PALINO, LINKEDIN SITE ENGINEERING

Who Is This Guy?

- LINKEDIN SITE RELIABILITY ENGINEERING
- PRINCIPAL STAFF ENGINEER
- CO-AUTHOR, KAFKA: THE DEFINITIVE GUIDE
- MARATHON RUNNER







Incident Management at LinkedIn

TERMINOLOGY

GCN

Severity

TOOLS

Jira

Slack

Zoom

Google Docs

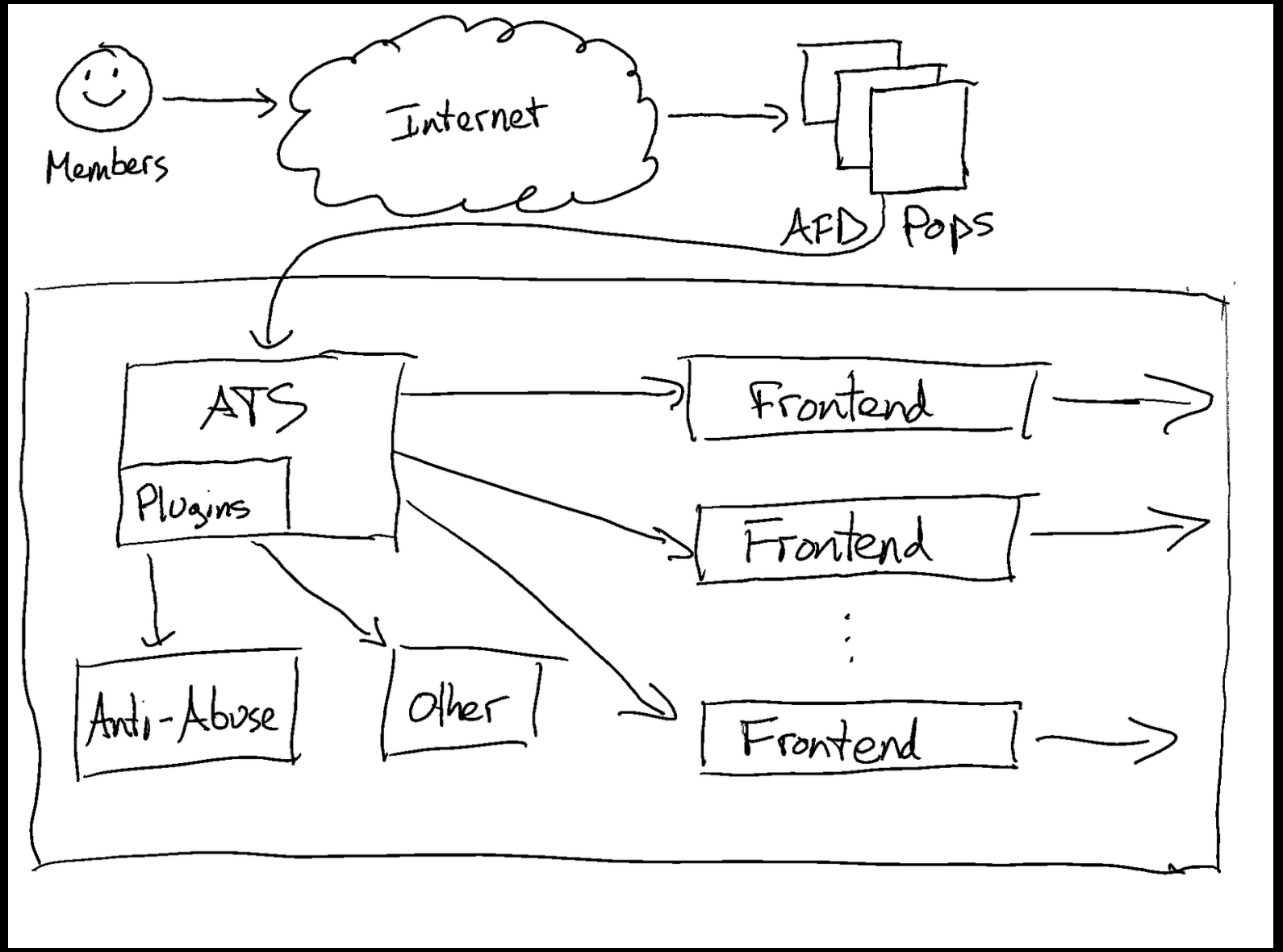
ROLES

SiteOps / NOC

Communications

On-call Engineers

Member Traffic Flow



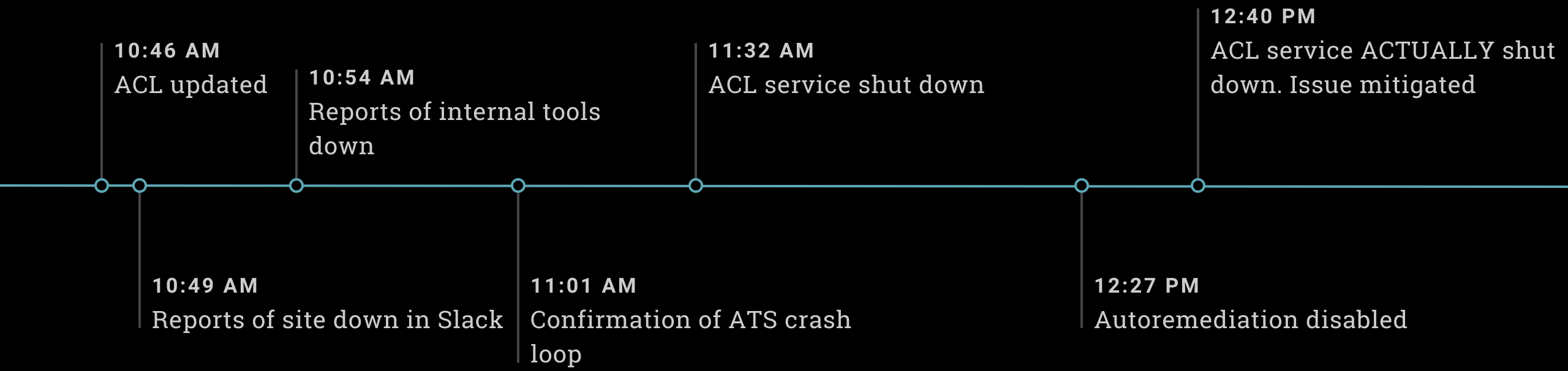
Moving Desktops to Azure

ELIMINATE LINUX TOWERS

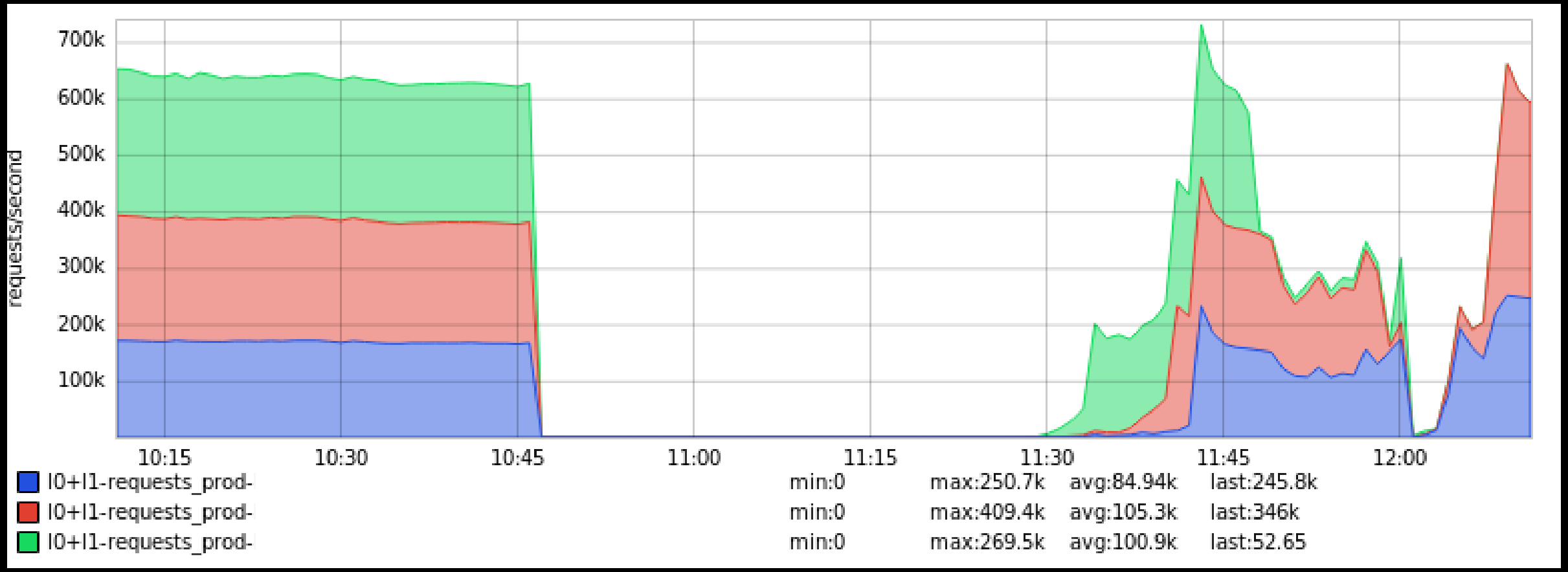
ACCELERATED BY PANDEMIC

ACCESS CONTROL UPDATES

Incident Timeline to Mitigation

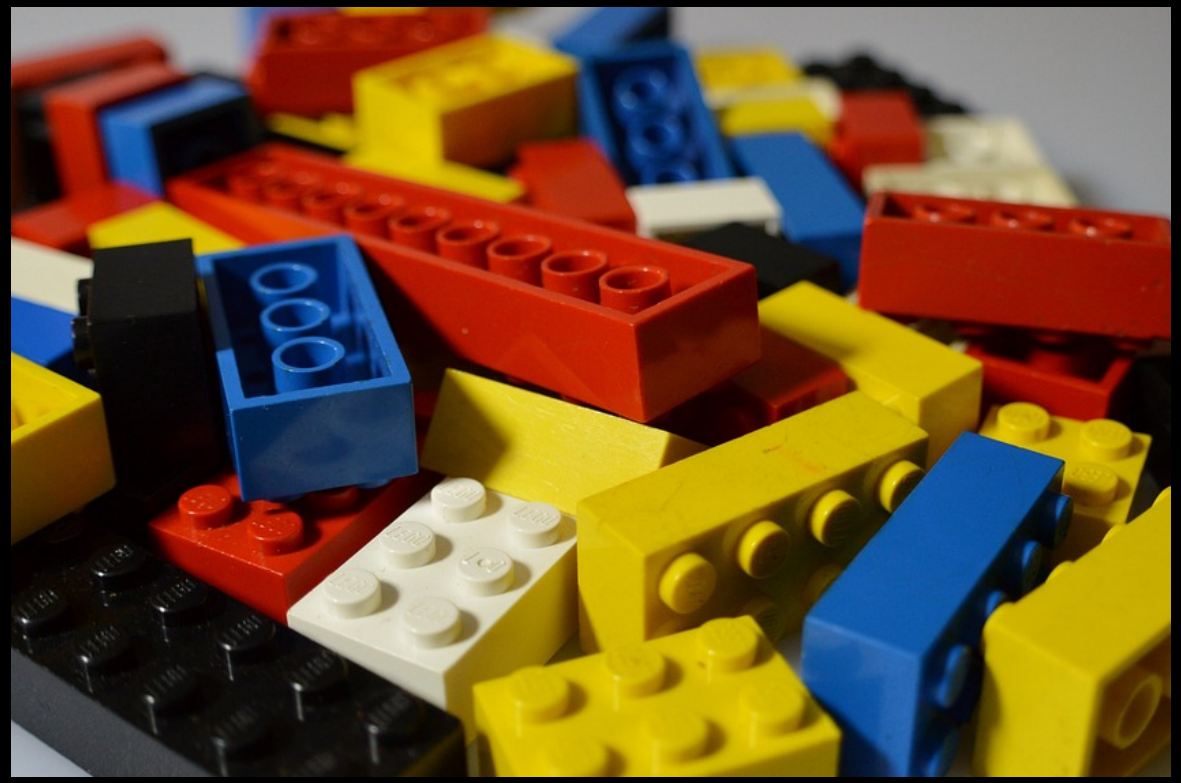


Oops


700k
600k
500k
400k
300k
200k
100k
requests/second
10:15
10:30
10:45
11:00
11:15
11:30
11:45
12:00
l0+l1-requests_prod-
min:0
max:250.7k
avg:84.94k
last:245.8k
l0+l1-requests_prod-
min:0
max:409.4k
avg:105.3k
last:346k
l0+l1-requests_prod-
min:0
max:269.5k
avg:100.9k
last:52.65

What Was the Root Cause?

~~What Was the Root Cause?~~ THIS IS A MYTH!



NEVER ONE THING

HUMAN ERROR IS THE SMALLEST COMPONENT

Contributing Factors

CHANGE PROCESS

- Manually updated JSON
- No review process
- Update was missing a comma (TRIGGER)

LACK OF VALIDATION

- UI did not validate input
- App that loaded the list to the DB did not validate
- ATS was supposed to validate, but there was a bug

ATS ISOLATION

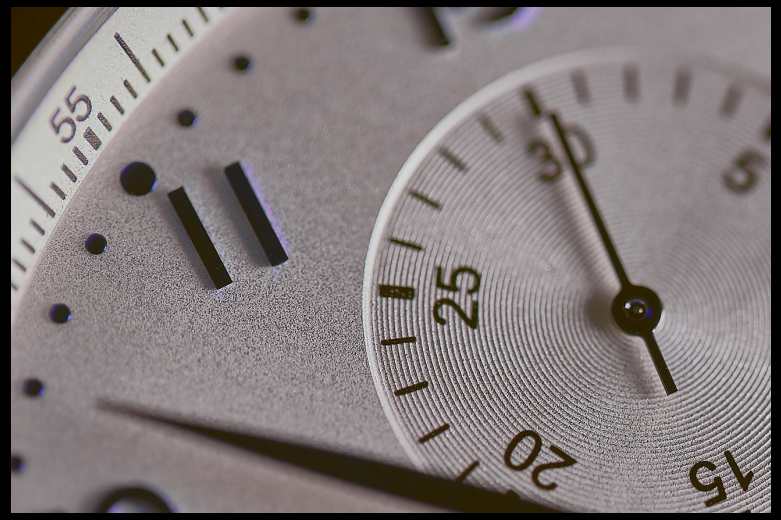
- External and Internal tiers are separate clusters
- The configurations are supposed to be isolated
- Both tiers loaded the same anti-abuse list (although internal does not need it)

NO FALLBACK

- ATS went into a core dump loop
- The app and its infrastructure could not detect this and revert to a last known good state

Working The Incident

What Went Right?



TIME TO DETECTION

PROBLEM IDENTIFICATION

INCIDENT TOOLING

Too Many Cooks

WHAT ELSE IS THERE TO DO?

TOO MUCH DIRECTION

NO STRONG INCIDENT COMMANDER

Break glass hammer


SOLAR CHARGE LEVEL
WALL·E

Post-Incident

Fixing the Technology

TRAFFIC SERVER

Audit and fix isolation

Fix validation bug and add unit tests

Add monitoring for config validation errors

CHANGE VALIDATION

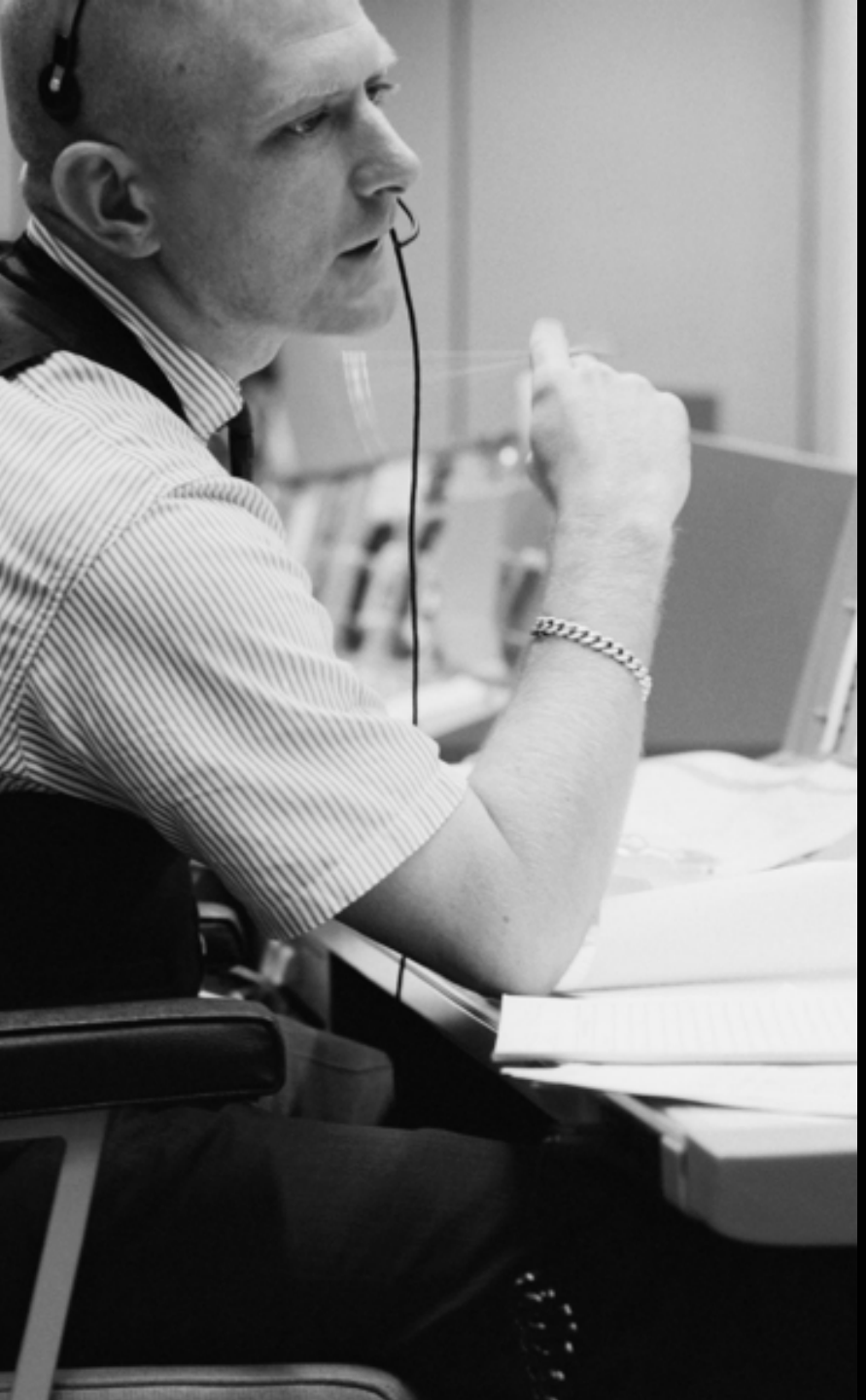
Update process for allowlist changes

Add format checking in multiple apps

AUTO-REMEDIATION

Define “break glass” procedures

Follow deployment pauses



Incident Commander Role

- INTERACTING WITH SENIOR LEADERSHIP REQUIRES A CERTAIN SET OF SKILLS
- CREATE AN IC ROTATION OF SENIOR LEADERS TO RUN THE LARGEST INCIDENTS
- USE THAT ROTATION AND TRAINING TO BRIDGE THE GAP TO A PERMANENT ROLE

Incident Response Overhaul

OLD, ORGANIC PROCESS

UNWRITTEN KNOWLEDGE

POOR FOLLOW-UP AND ANALYTICS

Future of Incident Response

DEDICATED TEAM

jeli

SINGLE INCIDENT FLOW

DEEPER ANALYSIS


Q