ITSM CHRONICLES - S01 E01 - FIELD GUIDE

# When a Payroll System Goes Down at 3 AM

*A hypothetical outage. Real tools. Real lessons.*
*Six things your team can use on Monday.*

| WRITTEN BY | SERIES | TOOLS COVERED |
|---|---|---|
| **Rajiv Ramkrishnan** | **ITSM Chronicles** | **BMC Helix Suite** |

Twenty years in IT teaches you one thing fast. The technology is rarely the problem. The processes, the gaps, and the decisions people make under pressure - that is where things go wrong.

This document strips away any narrative. What you get here is the raw debrief of a hypothetical but entirely plausible scenario - what broke, why it broke, which tools actually helped, and what your team can take from it.

> **WHAT THIS IS NOT**
>
> This is not a product brochure. It is a practitioner's write-up grounded in real-world experience - and a quiet nod to the unsung ITSM heroes who have configured and run these platforms in production every day. These lessons, though fictional in narration, are inspired by true events. Not by a vendor slide deck.

SECTION 01

## The Problem - What Actually Happened

Imagine this: it is 03:00. A business-critical payroll and authentication platform goes down. Hundreds of thousands of users cannot log in. Payroll transactions are frozen. Automated monitoring catches the failure in 47 seconds - before a single user files a ticket.

The root cause is a change. An iPaaS connector update that shortened the OAuth token refresh interval from 3,600 seconds down to 300 seconds. Pushed in the early hours with partial change approval and no documented rollback plan.

> THE THUNDERING HERD PROBLEM
>
> When you shorten a token refresh interval across hundreds of thousands of sessions, they all hit the endpoint simultaneously at the next batch window. The Auth Gateway rate limiter - not updated in the CMDB - cannot handle the spike. It crashes. Payroll follows because it uses the same auth token.

This is a well-known failure pattern and entirely preventable. The change risk score was manually overridden. No test evidence. No rollback path.

**Why it matters:** One config value. One missing rollback plan. One CMDB entry not updated. That is all it takes.

SECTION 02

## The Tools - In Plain English

Four components from the BMC Helix platform are involved. Here is what each one actually does, in plain language.

**Helix ITSM** Your incident management nerve centre. It raises the P1, opens the bridge, links thousands of child tickets to the parent incident, and tracks every action. When the parent closes, all children follow automatically.

**Helix AIOps** The monitoring layer that catches the outage in 47 seconds - before a single user files a ticket. Runs synthetic probes continuously and fires automated P1 alerts when a service fails its health check.

**Helix iPaaS** The integration platform where the problem starts and where the fix is executed. The faulty connector lives here. The rollback workflow runs here. The remediation runbook is orchestrated from here.

**HelixGPT** The AI correlation engine. While engineers work the logs in isolation, HelixGPT ingests 17 signals across 4 service domains simultaneously and pinpoints the root cause in 22 minutes.

> THE HONEST TAKE ON AI
>
> HelixGPT does not replace anyone. It does what humans cannot do at scale - correlate events across multiple domains simultaneously. Your engineers still make the call. The AI cuts the time to insight.

SECTION 03

## How It Was Fixed - Step by Step

**T+00:00** Synthetic probe detects authentication endpoint failure. P1 auto-raised. Incident bridge opened. Stakeholders paged. Detection time: 47 seconds.

**T+02:00** Thousands of symptom tickets auto-linked as children to the parent P1. Engineers stop working individual tickets and focus on root cause.

**T+22:00** HelixGPT correlates 17 signals across 4 service domains. Root cause: thundering herd from shortened token refresh interval. Confidence: 91%.

**T+31:00** Emergency rollback change raised. Full approval from all required approvers. iPaaS rollback triggered. Token interval reverted to 3,600 seconds.

**T+43:00** Runbook executes. All Auth Gateway pods healthy. Jitter offset applied. CMDB updated with corrected rate limiter threshold.

**T+54:17** **Authentication returns HTTP 200. Payroll pipeline resumes. Incident resolved. Total MTTR: 54 minutes 17 seconds.**

> POST-INCIDENT REVIEW
>
> A PIR was filed the same day. Not optional. That is how you stop it happening twice.

SECTION 04

## Six Lessons Your Team Can Use on Monday

**1. Detect before your users do.**

Synthetic monitoring caught this in 47 seconds. Waiting for users to raise tickets means you are already behind.

**2. No rollback plan means no way out.**

Every change needs a documented rollback path. Not just high-risk ones. Every single one.

**3. CAB process exists for a reason.**

Partial approval and manual risk overrides are how you end up with a critical outage at 3 AM. Respect the process.

**4. AI correlation saves real time.**

HelixGPT found what humans missed in 22 minutes. Multi-domain correlation at scale is not a human-speed task.

**5. Your CMDB is your map. Keep it current.**

The rate limiter was not in the CMDB. Nobody knew to account for it. An inaccurate CMDB gives false confidence.

**6. File the PIR. Every time.**

Post-Incident Reviews are not admin overhead. They are how good teams stop repeating the same mistakes.

RR

**Rajiv Ramkrishnan**

ITSM professional, ITIL Expert, BMC Helix specialist. Twenty years fixing the gaps between people, processes, and IT systems. This is the Way: ITSM Chronicles is a fan-inspired series that uses storytelling to make ITSM lessons stick.

itsm.rajivramkrishnan.com