*I*n my previous job, I often wondered—How cool it would be if we can just automate JIRA ticket creation for every cloud incident? Security and Incident management are a constant challenge on cloud: alerts from all direction, engineers scrambling for context, and hours lost chasing down solutions already found in the past. I used to dream of a system smart enough to detect issues, diagnose them, and create detailed tickets with clear, actionable guidance—all without human intervention.

Back then, it felt more like wishful thinking. Fast forward to today at the American Red Cross I have engineered this solution.

I've built and deployed a JIRA Incident Creation System for AWS, powered by Amazon Bedrock (Claude & Nova FM), OpenAI embeddings for RAG, and robust AWS monitoring with CloudWatch & CloudTrail. And LangSmith for debugging. The system continuously tracks services of your choice eg. data pipelines, compute resources, and security boundaries. As something fails—be it a pipeline error, unexpected S3 access or file downloads or uploads of sensitive data, EC2 provisioning violation, or IAM security event—the system responds instantly.

It semantically searches a vector store through RAG of past incidents for similar issues and proven solutions. If it’s a completely new problem, Tavily’s API is used to find the latest best practices and remediation strategies on the web. The result? Detailed JIRA tickets are created automatically, with full context, RCA, and step-by-step remediation, and a brief Push Notification is sent to team—no manual work needed.

What once felt impossible now empowers teams to respond faster and eliminate alert fatigue.

How the Multi-Agent System Works:

⚙️ Supervisor Agent: Orchestrates the entire workflow, ensuring smooth handoffs and maintaining consistent state across agents.

🔍 Monitoring Agent: Continuously observes AWS CloudWatch and CloudTrail events, proactively detects anomalies, and generates detailed monitoring analysis.

🧠 Diagnosis Agent: Upon issue detection, agent uses RAG to 1st query a vector database for similar incidents and proven solutions. If a match is found, it instantly recommends the remedy, reducing response time. If no match, it automatically escalates to Tavily for new strategies, saving each result to memory so the system keeps learning.

📋 Resolution Agent: Generates comprehensive JIRA tickets with incident context, impact analysis, and detailed remediation procedures. And a Push Notification is sent to Developer.

The real breakthrough: the system learns from every incident, using RAG and vector memory to solve repeat issues faster each time.

If you're passionate about transforming operations with AI, automation, and cloud, I'd love to connect!

**#AWS** **#AmazonBedrock** **#Claude** **#AIEngineer** **#RAG** **#CloudWatch** **#LangSmith** **#MLOps**