Executive Summary: Beyond Detection - Evaluating LLM-Based Multi-Agent Systems for Real-Time Incident Comprehension and Recommendation

Philip Drammeh, M.Eng.
Independent Researcher
October 29, 2025

https://myantfarm.ai

ABSTRACT

Organizations operating digital infrastructure face a growing gap between detection and comprehension of incidents. While monitoring tools generate alerts within seconds, understanding why an outage occurred and what to do next often takes minutes or hours—causing revenue loss, SLA penalties, and customer dissatisfaction. Our study aims to assess this latency by evaluating large language model (LLM)–based multi-agent reasoning for real-time incident comprehension and decision support.

I. BUSINESS PROBLEM

Enterprises in telecom, cloud, and IoT depend on distributed observability systems such as Datadog, Jira, and Slack. The volume and heterogeneity of data make root-cause isolation slow and cognitively expensive, often leading to significant revenue impact from outages and degraded customer experience. Traditional dashboards visualize data but rarely deliver actionable narratives. This study investigates how orchestrated LLM agents can reduce **incident comprehension latency**—the time between signal and usable understanding.

II. APPROACH

A reproducible Docker-based simulation was built with four core services: the Evaluator, Multi-Agent Coordinator, LLM backend (quantized Llama 3.2 via Ollama), and Analyzer. Three experimental conditions were tested—C1 (baseline/manual), C2 (single-agent), and C3 (multi-agent orchestration). Performance was evaluated using two metrics: **Time-to-Usable Understanding ($\mathbf{T_{2U}}$)** and **Decision Quality (DQ)**.

III. KEY FINDINGS

Across 348 simulated trials, the multi-agent condition (C3) reduced comprehension latency by **58%** and improved decision quality by **48%** compared with baseline processes. Results converged after 30–40 trials, indicating robustness of improvement. The architecture establishes a foundation for measurable gains in reliability, interpretability, and time-to-decision.

IV. STRATEGIC RELEVANCE

Incident comprehension delay directly translates to operational cost. Reducing T_{2U} by 50–60% yields material savings in downtime and human triage effort. Integrating governed access (**Model Context Protocol, MCP**) and evidence grounding (**Retrieval-Augmented Generation, RAG**) ensures enterprise readiness without compromising security or traceability. The framework is domain-agnostic and extendable to Industrial IoT and critical infrastructure monitoring.

V. CONCLUSION AND NEXT STEPS

Our study demonstrates that orchestrated LLM agents can transform incident response from reactive detection to proactive comprehension. While MCP and RAG were conceptually integrated but not active during testing, they represent the next step toward deployable, auditable AIOps environments where noise becomes actionable narrative in seconds. Next-phase testing will extend the framework to live telemetry using MCP for secure data mediation and RAG for context retrieval. These integrations will enable quantitative ROI analysis through reduced Mean Time to Mitigate (MTTM), SLA compliance improvement, and automation efficiency. A 10–15% improvement in comprehension efficiency at enterprise scale could translate into multimillion-dollar savings annually, bridging academic rigor with measurable operational impact.

Contact: philip.drammeh@gmail.com
Repository: https://github.com/Phildram1/myantfarm-assets