AI in DevOps: What Happens When Ops Gets a Copilot

whoami

- pjay(Priyanshu Jain)
- 10+ years building platforms
- Principal Engineer @OkCredit(YCS18)
- Mountaineer (HMI alumni)
- pjay.in
- x.com/pjay_in

Agenda

- Managing Infra is getting harder

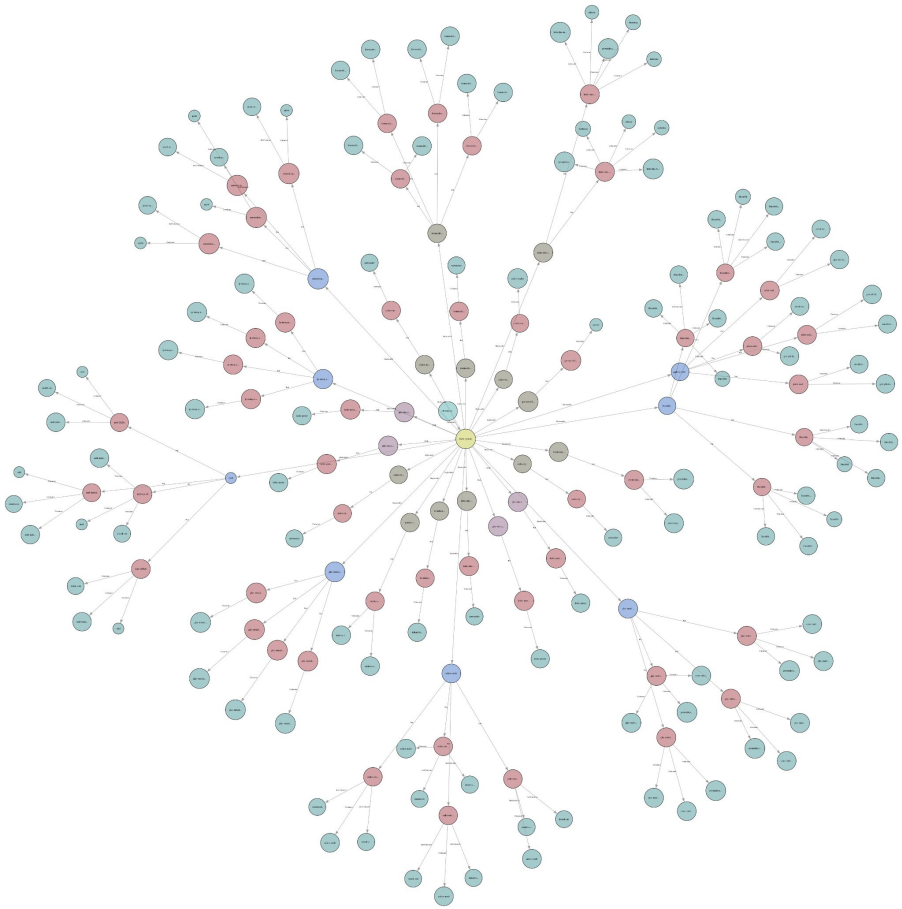
- AI-powered DevOps
  - Incident Response
  - Resource Management
  - Cost Optimization
  - Security & Compliance

- Going Forward

Current State of Infrastructure

It’s gone out of control

Infra is complex



This is what people expect from us

CI/CD, GitOps, and Infrastructure Automation:

- Lead the design, development, and optimization of CI/CD pipelines using Kubernetes-native tools (ArgoCD, GitHub Actions) to ensure rapid, reliable deployments.
- Drive Infrastructure-as-Code (IaC) initiatives using Terraform, CloudFormation, and Pulumi, ensuring consistent, automated, and reproducible infrastructure deployments.
- Advocate and implement GitOps best practices to manage Kubernetes configurations and application deployments.

Observability, Monitoring, and Incident Response:

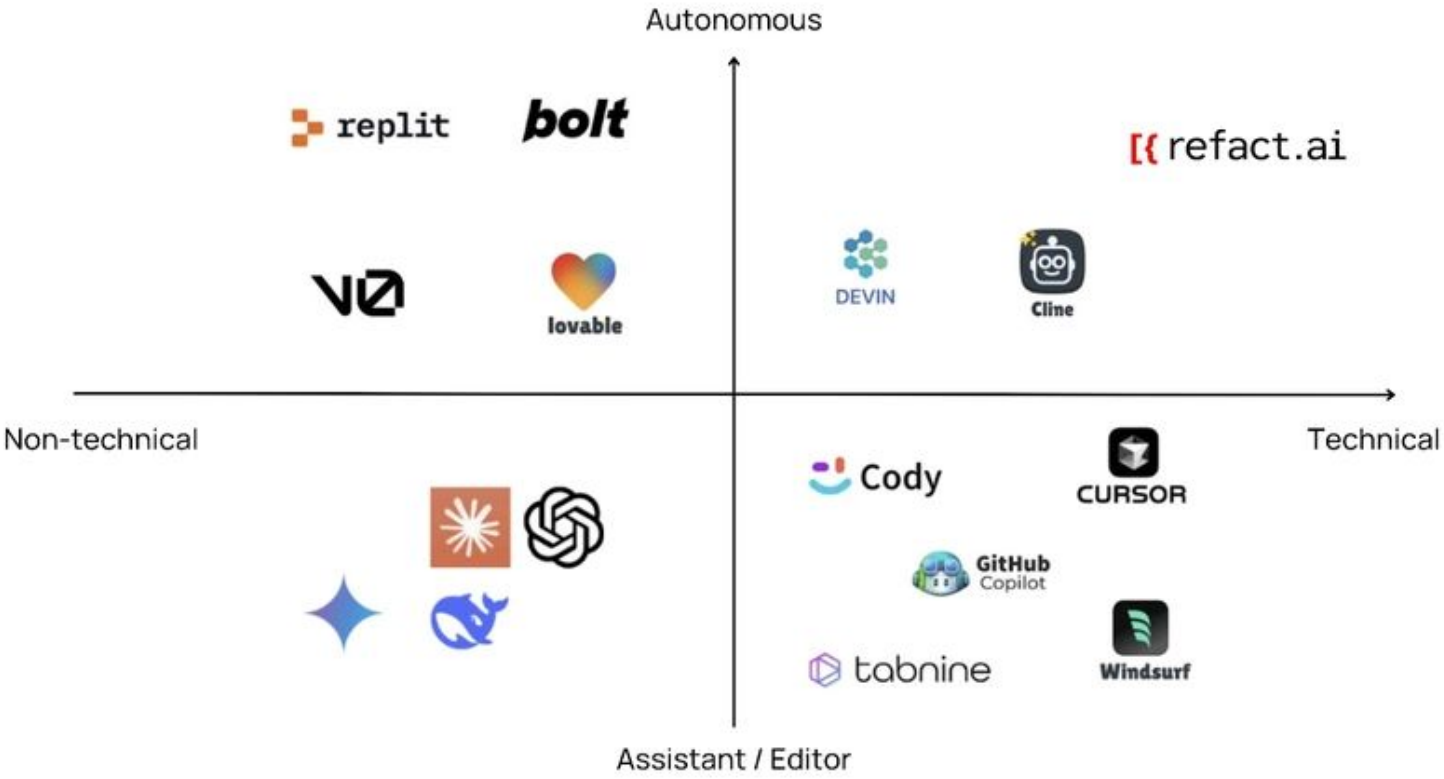
- Develop comprehensive monitoring, logging, and alerting systems (using Prometheus, Grafana, ELK, Datadog, etc.) that provide deep insights into system performance, including detailed latency metrics.
- Establish and refine SLOs/SLIs for ultra-low latency performance, and drive proactive incident management and post-mortem analyses.
- Continuously analyze system performance data to identify bottlenecks and implement improvements that enhance overall responsiveness.

- Design and evolve our cloud-native infrastructure (AWS/Kubernetes), ensuring availability, performance, and cost efficiency across regions and products.
- Build internal tools and platforms that help engineers deploy, monitor, and scale their services independently with minimal friction and maximum confidence.

Security Best Practices and Compliance:

- Implement and enforce robust security measures across the entire infrastructure, including container and network security best practices, encryption (in transit and at rest), and secure configuration management.
- Develop and maintain strict access control policies using RBAC, network segmentation, and automated compliance checks.
- Collaborate with security teams to conduct regular vulnerability assessments, penetration tests, and audits, ensuring adherence to industry standards and regulatory requirements.
- Integrate security into the CI/CD pipeline (DevSecOps) to identify and remediate risks early in the development lifecycle.

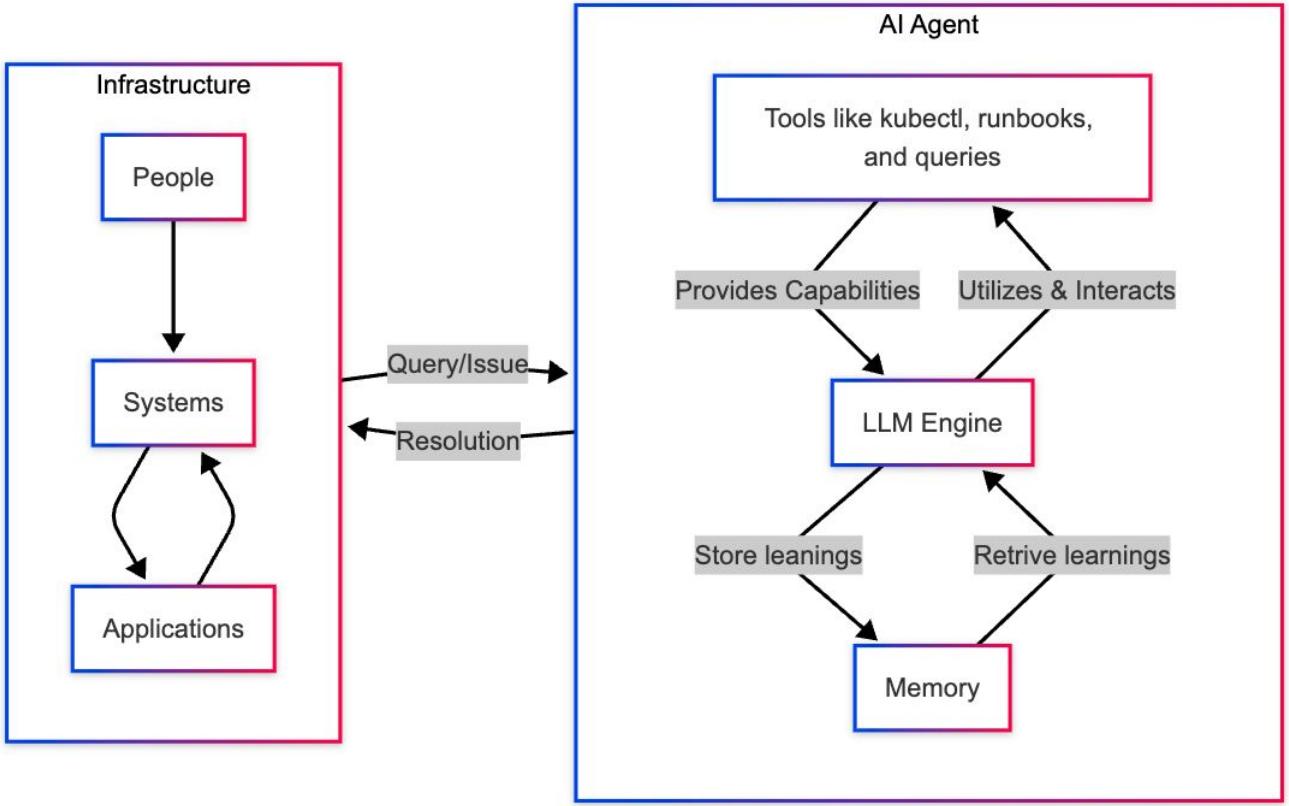
People are vibe coding



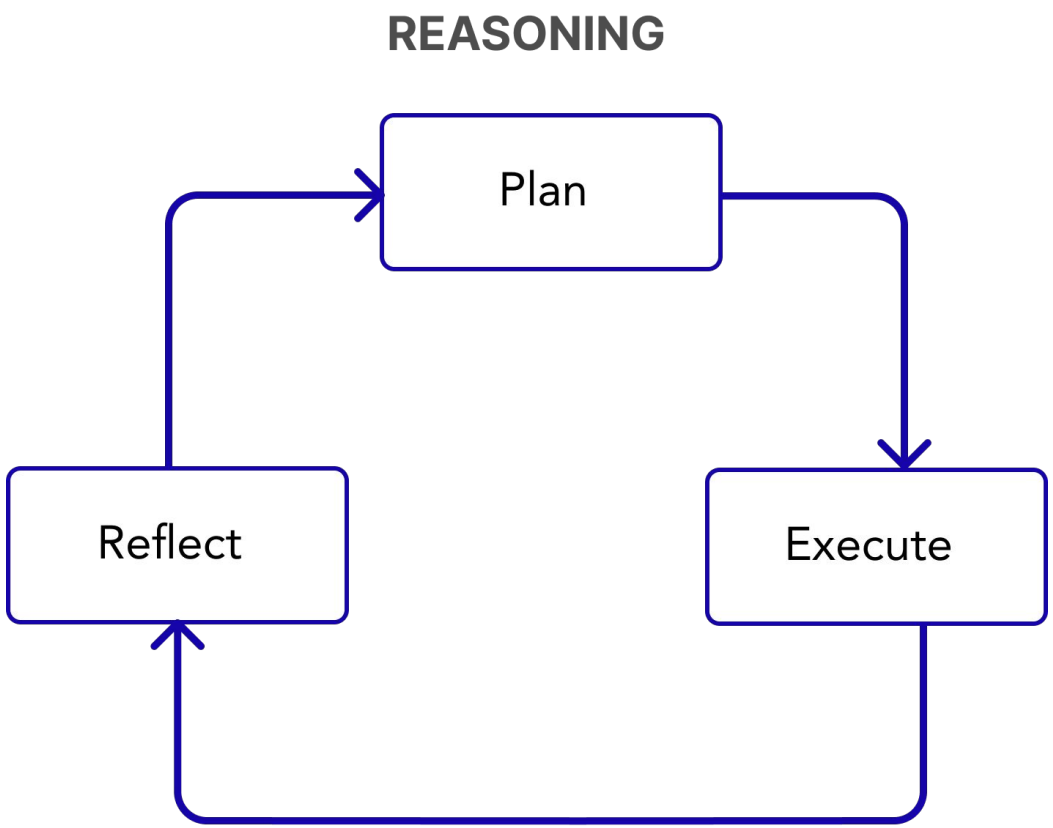
How can AI Agents help

AI agents can take the grunt work out of managing production environments

What are agents



Brain of an Agent



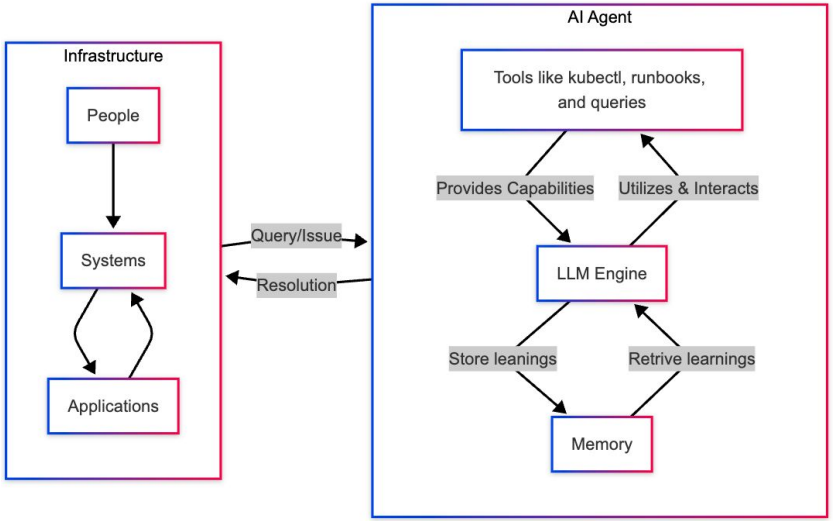
Why it works

Breaks problems into steps

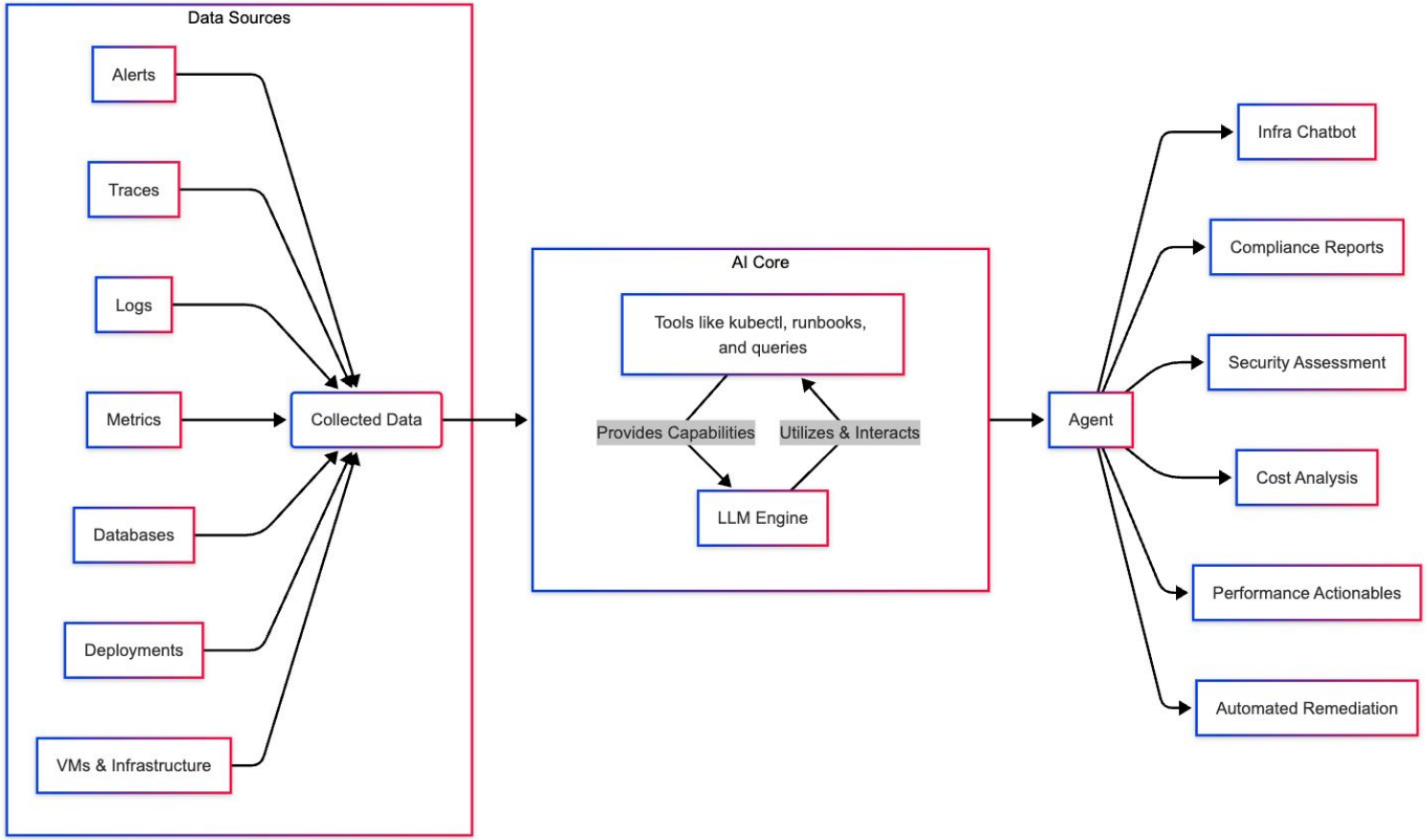
Uses all available tools

Learns from mistakes

Runs 24×7 at speed



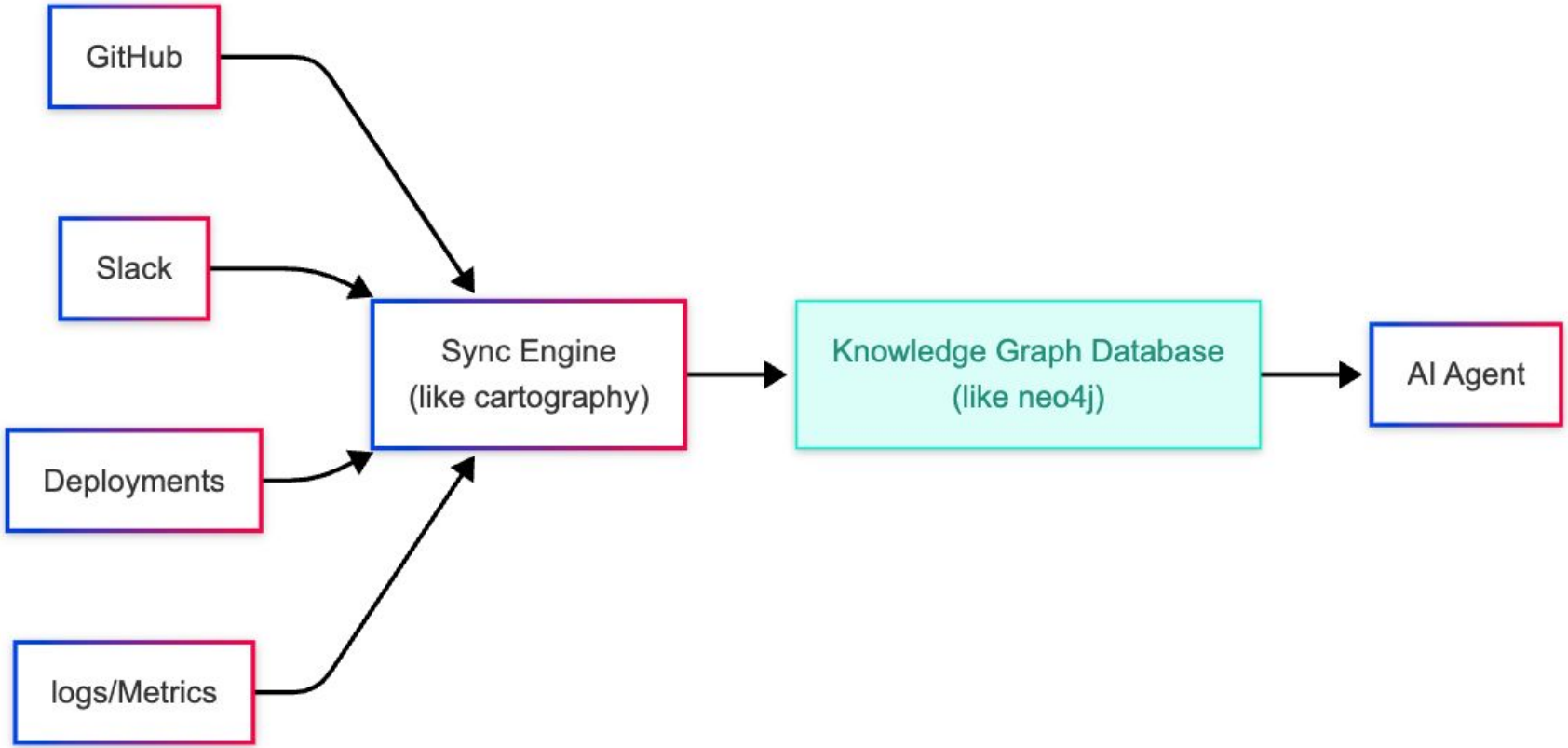
How this works for Infrastructure problems



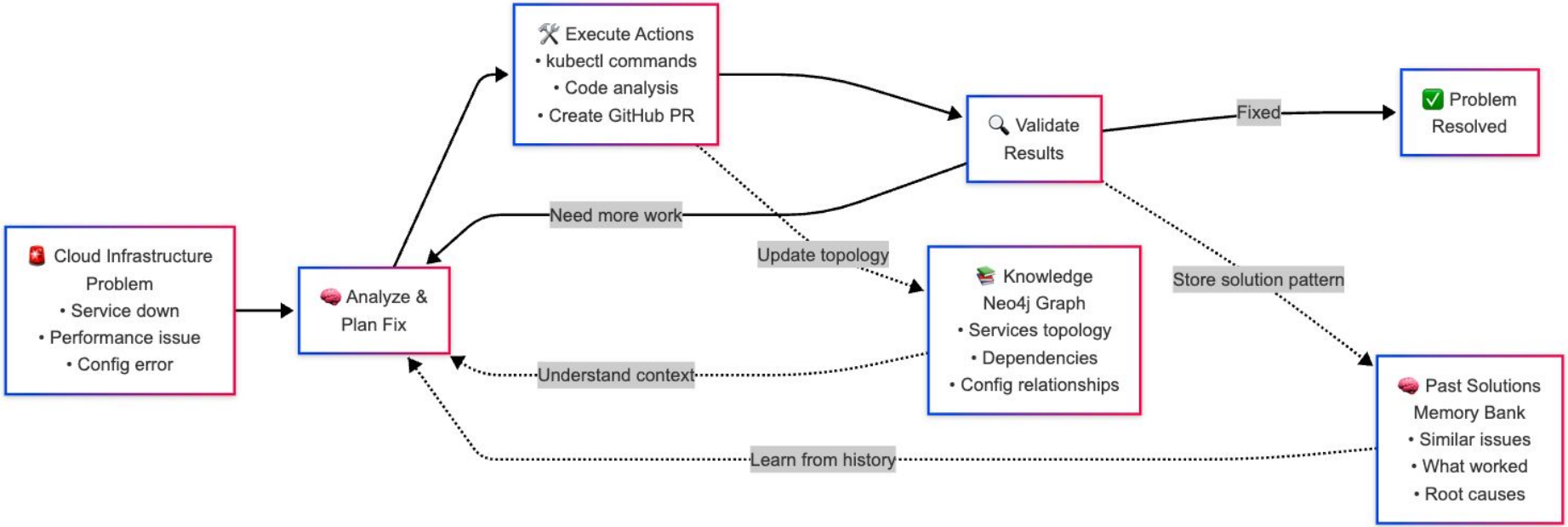
Incident Response

Machines fixing machines with human in the loop

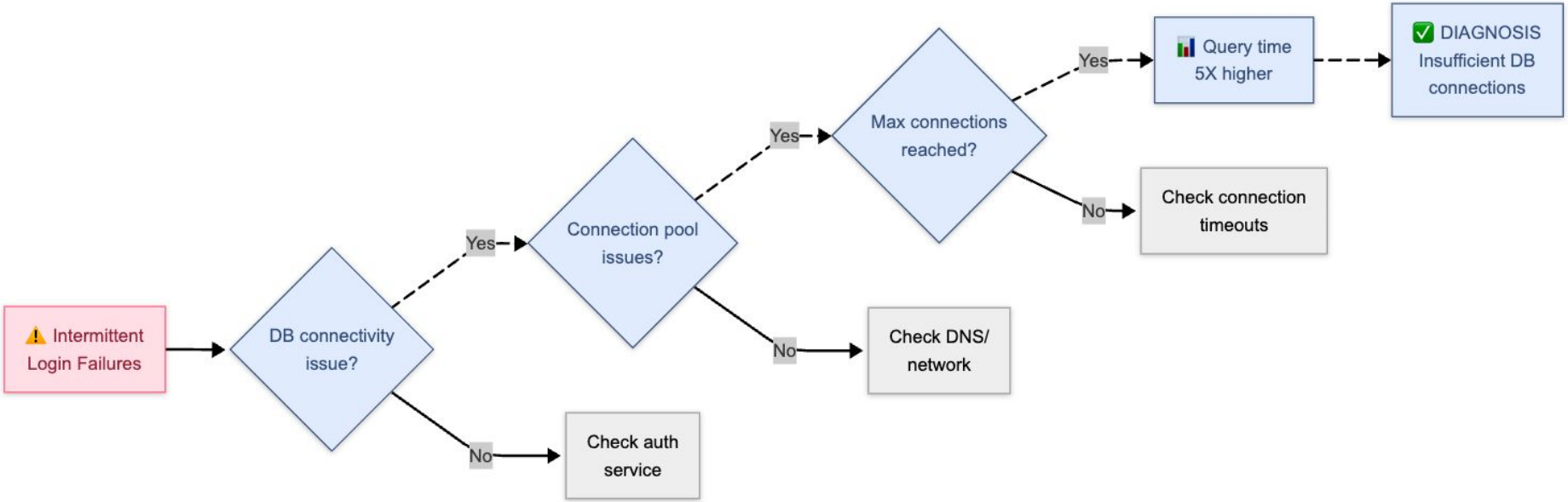
Knowledge Graph



Diagnosis pipeline



Example


⚠️ Intermittent Login Failures
DB connectivity issue?
Yes
No
Check auth service
Connection pool issues?
Yes
No
Check DNS/ network
Max connections reached?
Yes
No
Check connection timeouts
📊 Query time 5X higher
✅ DIAGNOSIS Insufficient DB connections

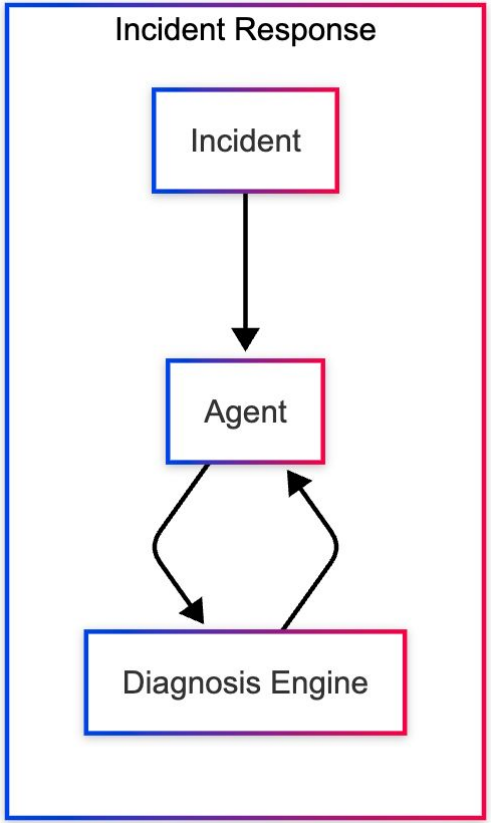
Overall

Automates specific solutions via tools

Remembers successful fixes

Gets smarter over time

As models get better, it can even use more generic tools to solve unknown problems



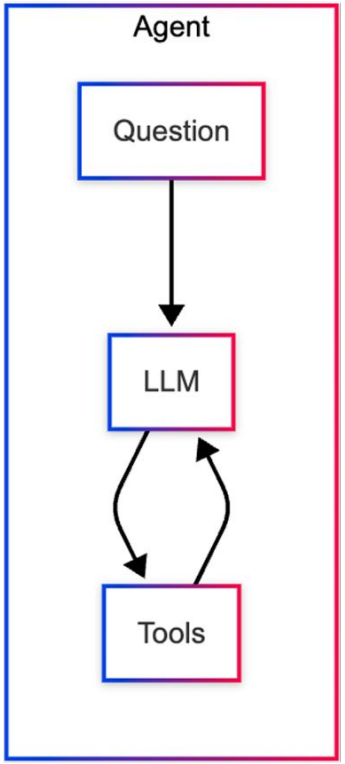
Resource Management

Answer questions, generate code, and assist with everyday DevOps tasks

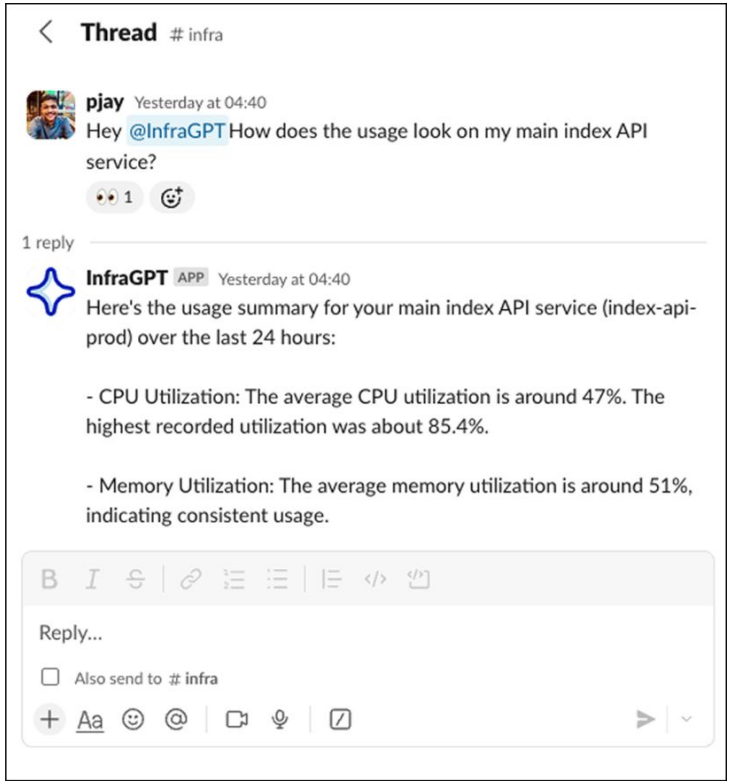
Answer questions

Knowledge Graph Database

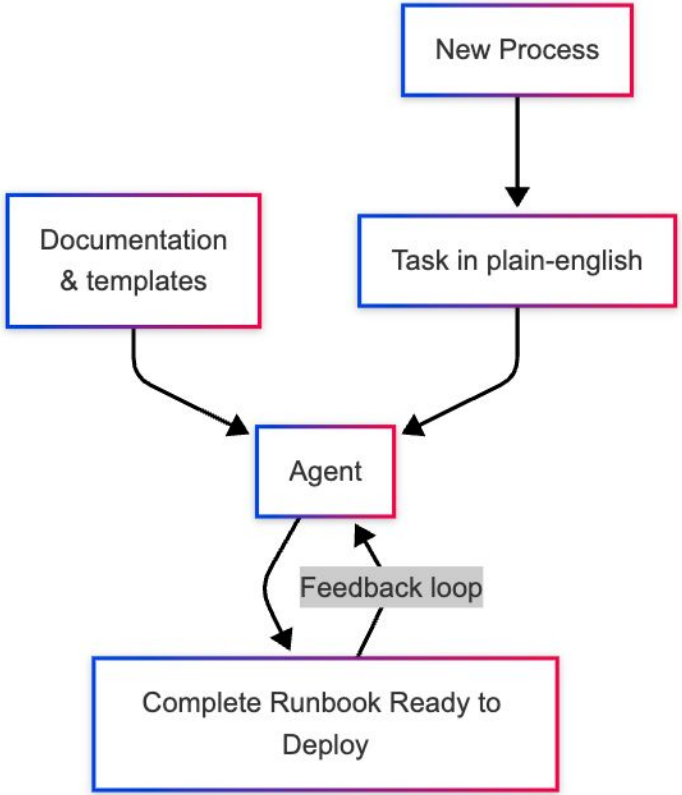
+



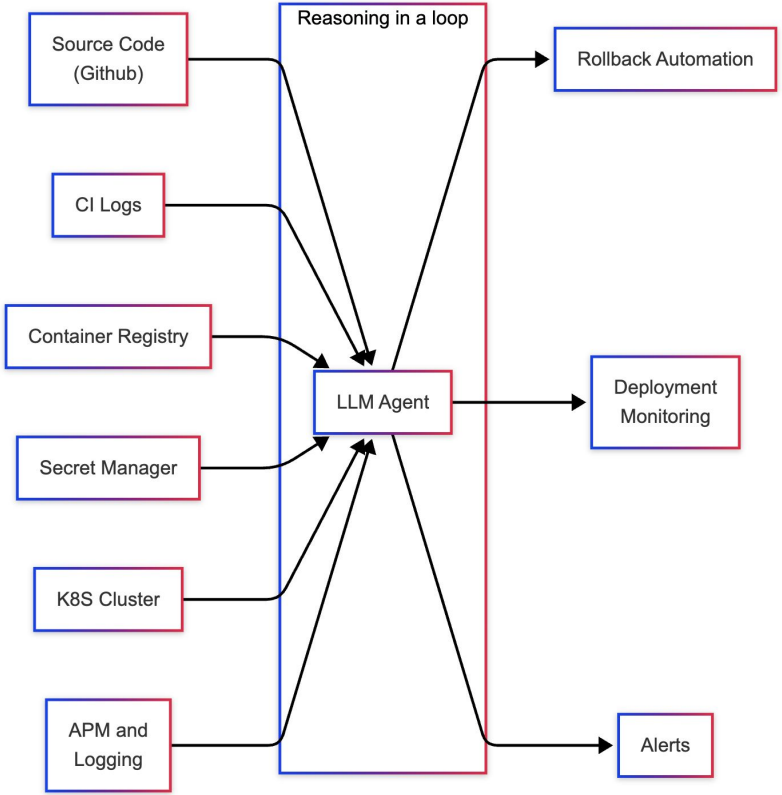
=



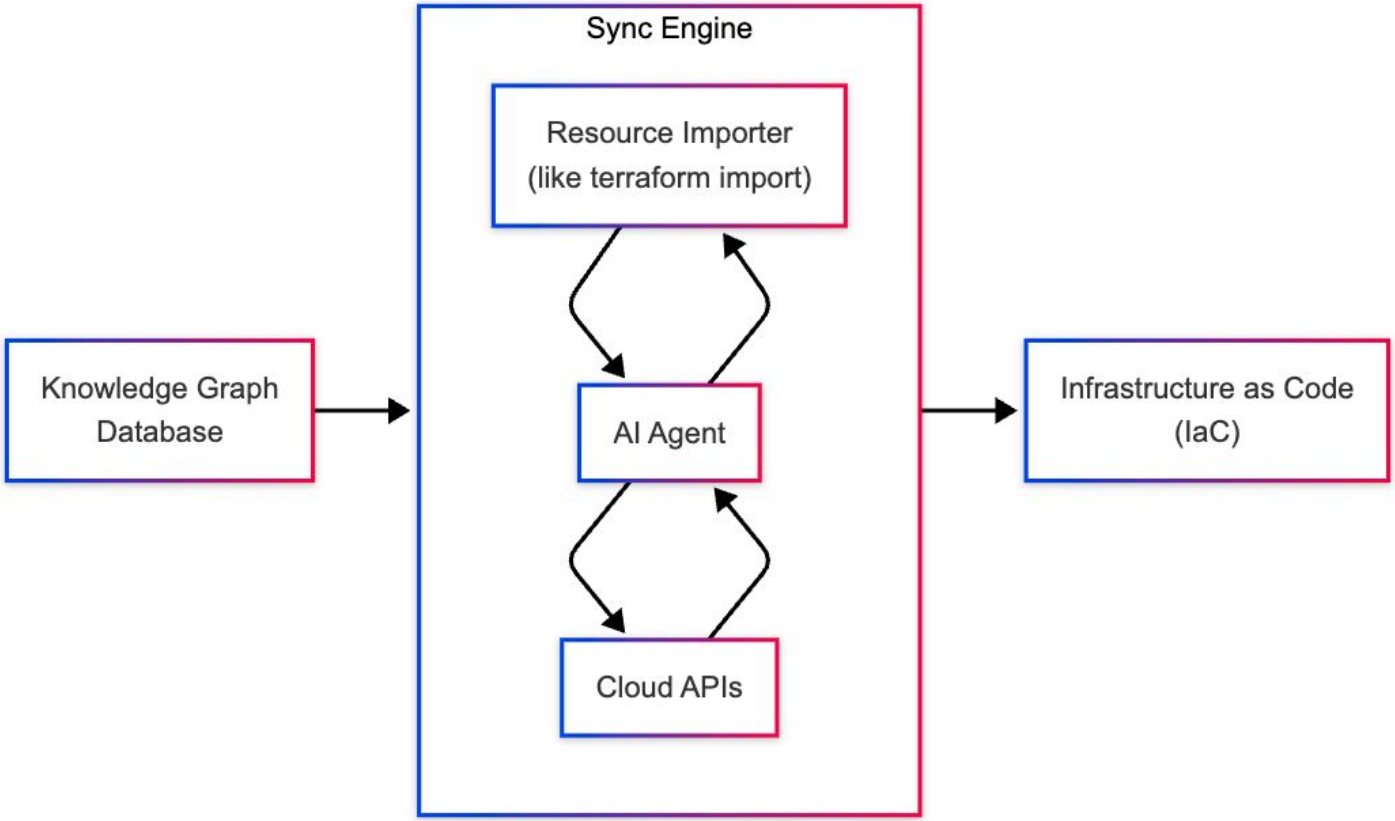
Automated Runbooks - (codegen)



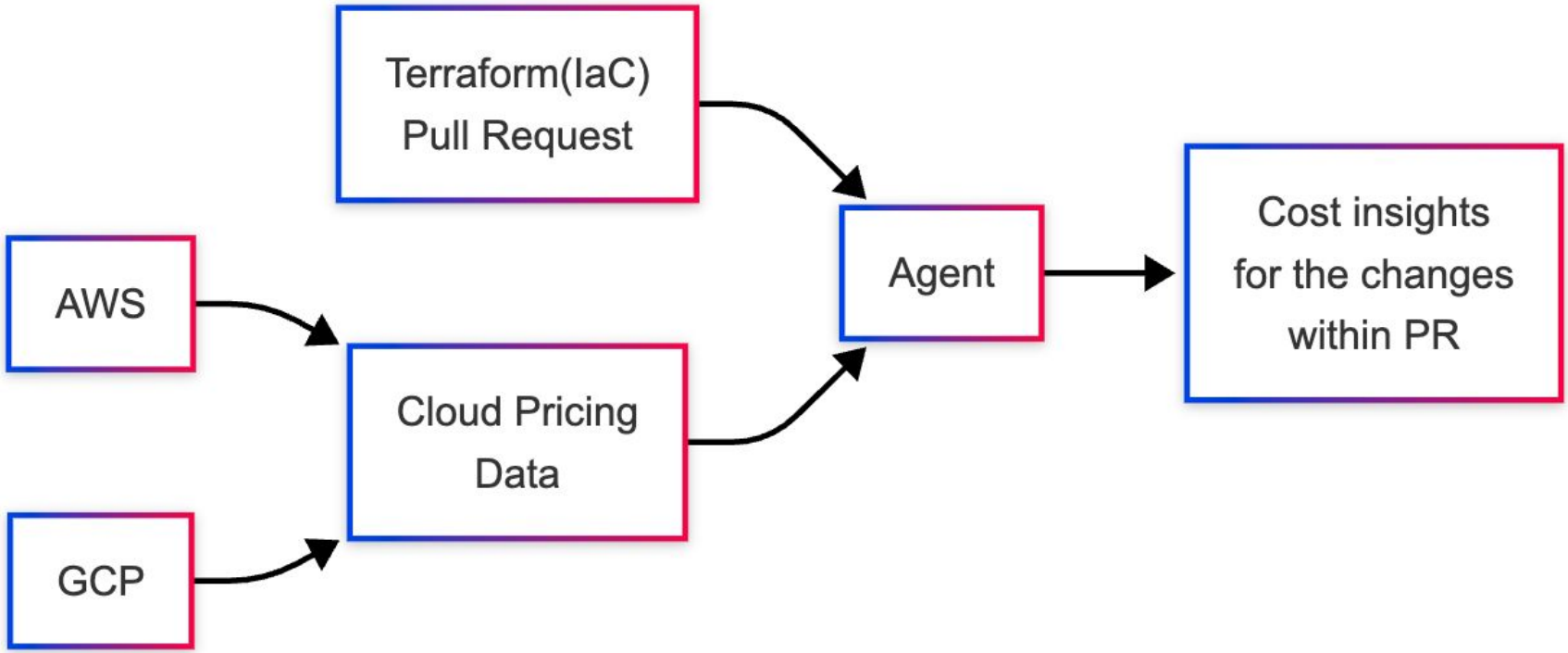
Supercharged delivery pipeline



Iac Drift Updates



Proactive Cost Monitoring



Cost Optimization

Doing this every month will have a compounding effect

Cost Optimization :Methods

1. Deleting unused resources
2. Resizing resources
3. Autoscaling
4. Optimizing non-production environment
5. Optimize code cost
6. Optimizing architecture cost
7. Optimizing network cost
8. Switching to open-source software
9. Pooled subscription billing into cloud bill

OkCredit(> 2M MAUs)

Daily cloud costs :-

Nov 1st, 2021 Cost

₹368,496.67

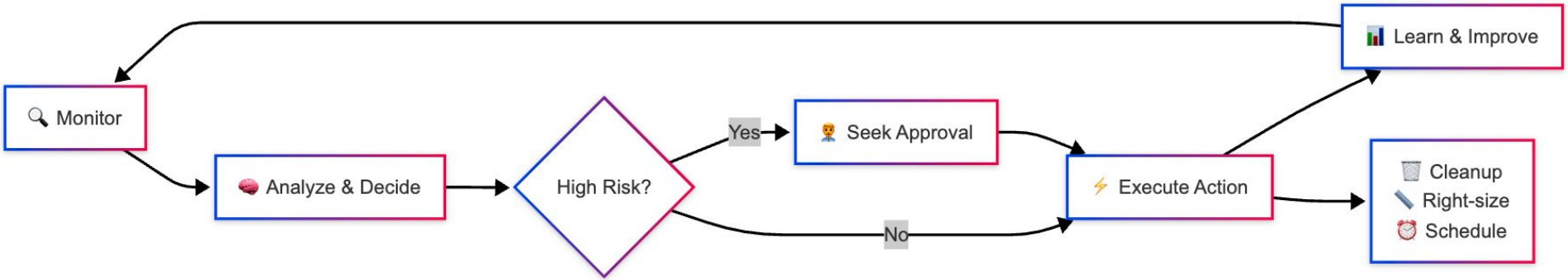
Sept 1st, 2024 Cost

₹47,832.70

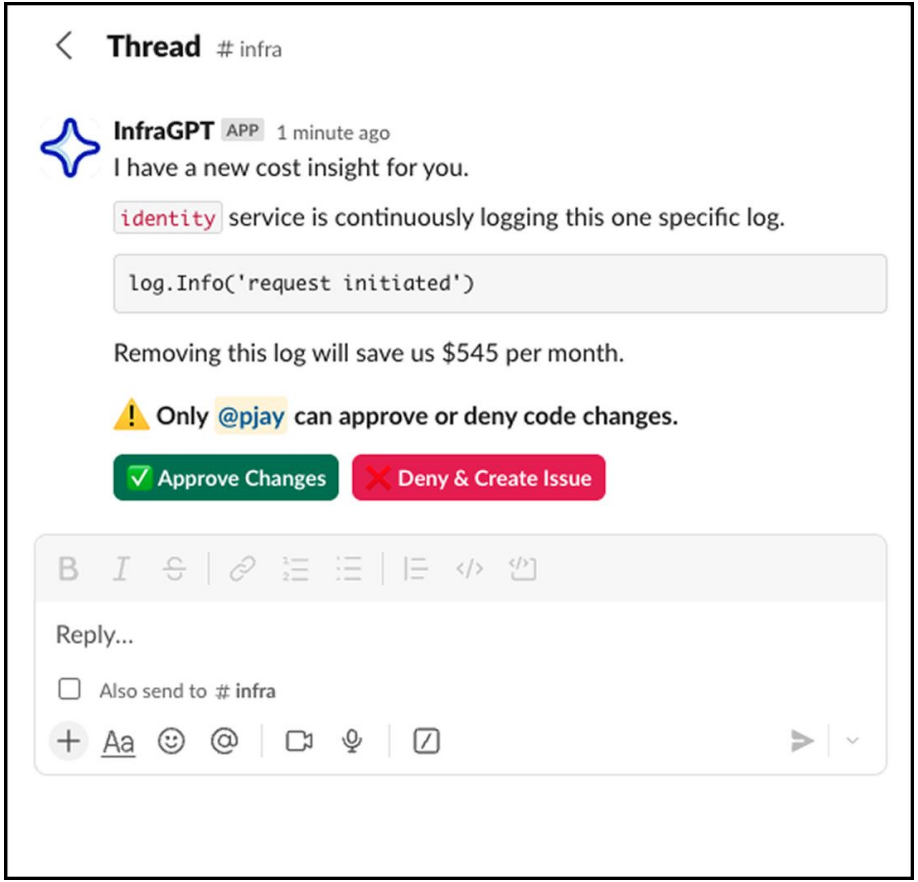
(≈87% drop)

Cost Optimization: Fundamentals

Cost Optimization: Agentic Pipeline



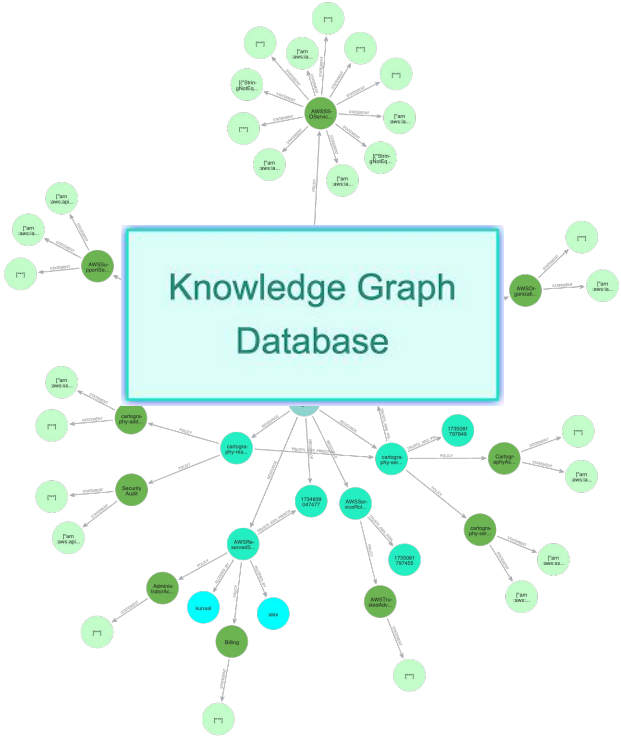
Cost Optimization: Insights example



Security & Compliance

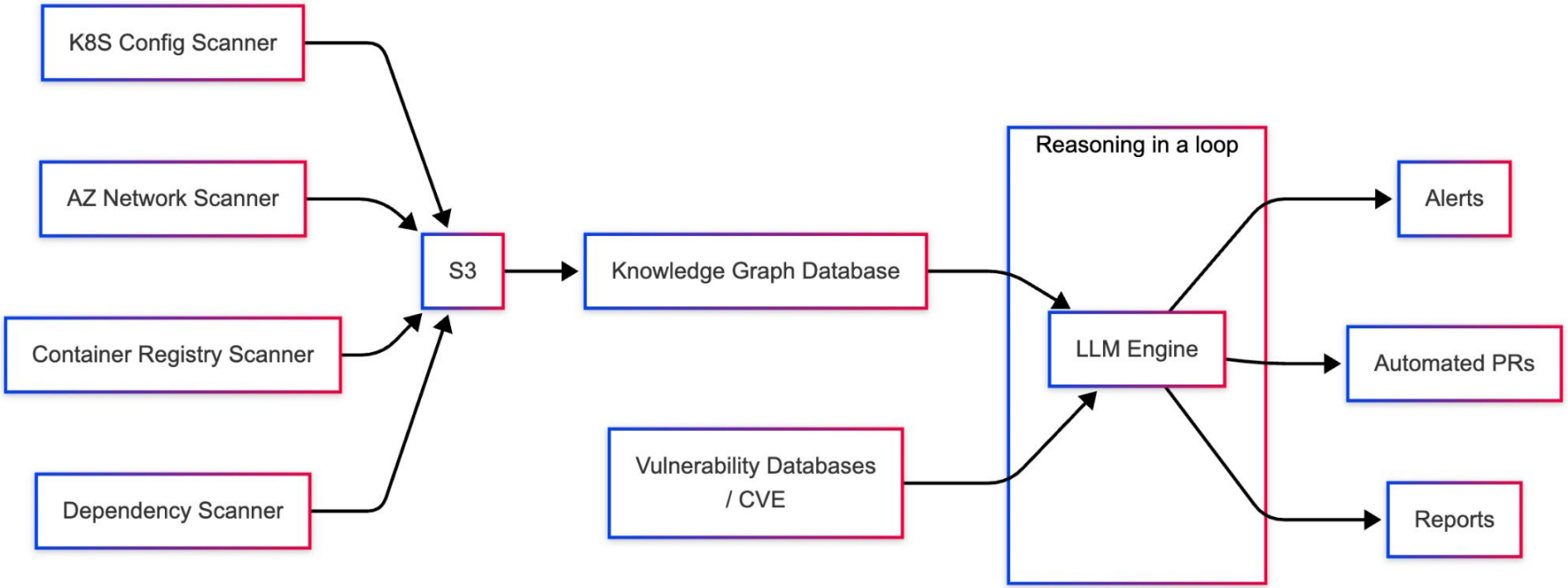
Unfair advantage for infrastructure teams managing security

Answer Contextual Queries



- Which s3 buckets allow public read or write access?
- Which credentials haven't been rotated in last 90 days?
- Which VMs are running out of date OS Version?

Automated Vulnerability assessment



Automated Vulnerability assessment: Research Example

Sean Heelan
@seanhn

Bug 3 github.com/torvalds/linux...

Fun fact: I found this one using o3 as a backend when evaluating if it was able to find a previous bug I had found. In ~100 runs it showed up twice. Afterwards I checked if Sonnet 3.7 could find it, and it can, but with 2.5x more false positives.

> **Sean Heelan** @seanhn · Apr 15
> Week 2 bug 2 was another remote UAF in ksmbd. This one is post-auth, although guest accounts can hit it. It's a straightforward "You forgot to null the reference" in the kerberos auth path. See commit message for details. github.com/smfrench/smb3-...

6:53 PM · Apr 30, 2025 · **13.5K** Views

AI / BUG HUNTING / LINUX KERNEL

How I used o3 to find CVE-2025-37899, a remote zeroday vulnerability in the Linux kernel's SMB implementation

MAY 22, 2025 SEANHN 8 COMMENTS

Follow @seanhn

In this post I'll show you how I found a zeroday vulnerability in the Linux kernel using OpenAI's o3 model. I found the vulnerability with nothing more complicated than the o3 API – no scaffolding, no agentic frameworks, no tool use.

Sean Heelan
@seanhn

I wrote-up how I used o3 to find CVE-2025-37899, a remote zeroday vulnerability in the Linux kernel's SMB implementation. Link to the blog post below 👇

4:12 PM · May 22, 2025 · **91.2K** Views

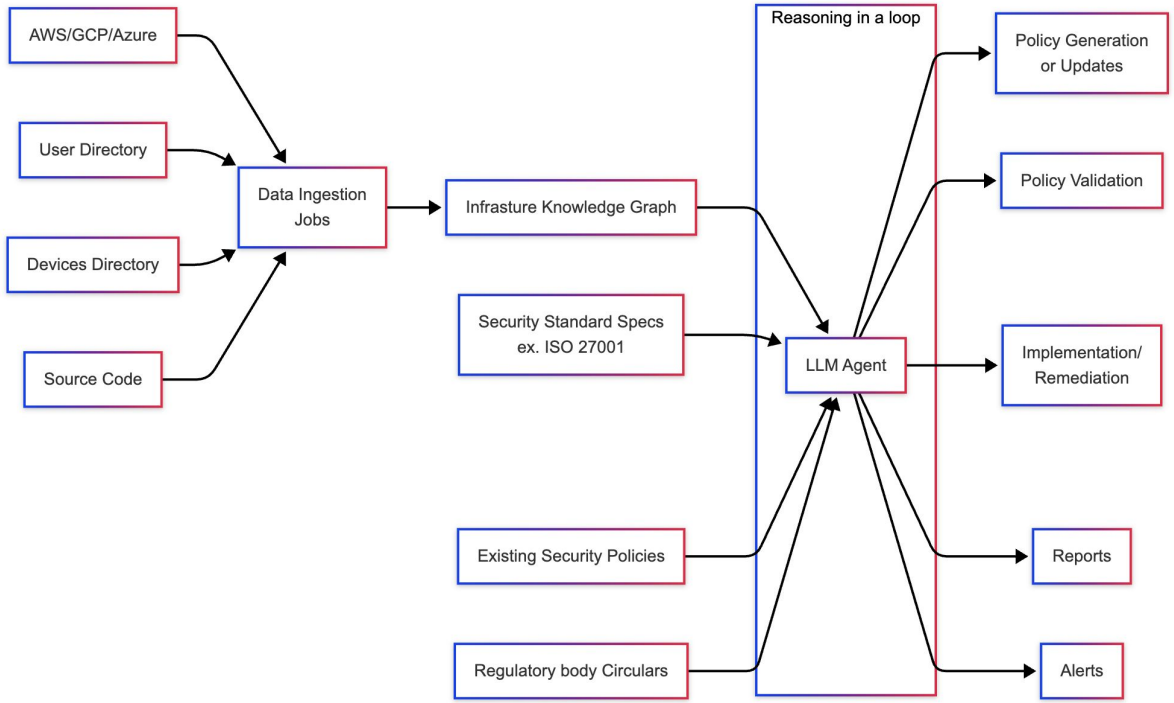
Sean used my LLM tool to help find the bug! He ran it against the prompts he shared in this GitHub repo using the following command:

```
llm --sf system_prompt_uafs.prompt              \
    -f session_setup_code.prompt                \
    -f ksmbd_explainer.prompt                   \
    -f session_setup_context_explainer.prompt   \
    -f audit_request.prompt
```

Sean ran the same prompt 100 times, so I'm glad he was using the new, more efficient fragments mechanism.

o3 found his first, known vulnerability 8/100 times - but found the brand new one in just 1 out of the 100 runs it performed with a larger context.

Compliance & Governance



Measuring Success

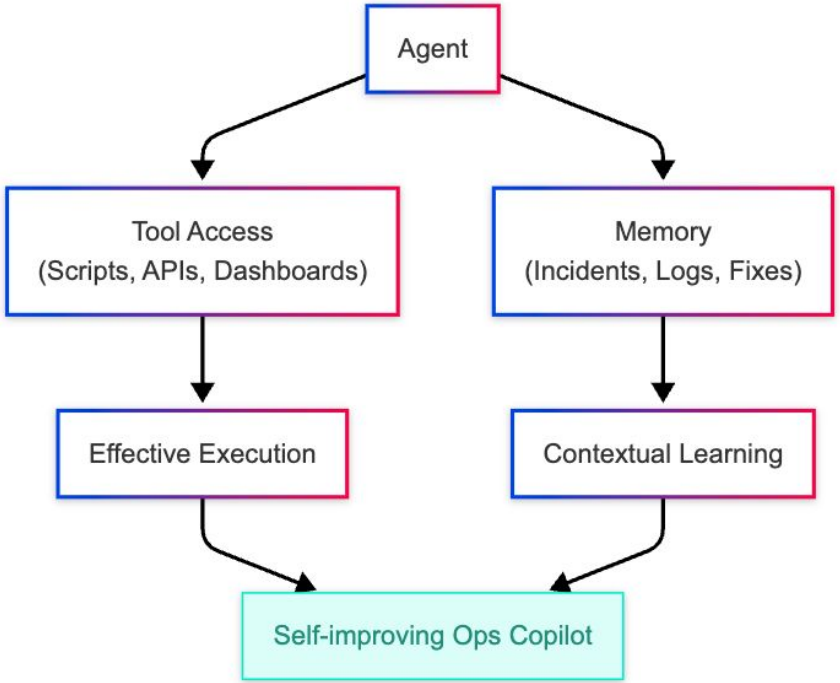
1. 📈 MTTR Reduction
2. 💰 Cost Savings
3. ⚡ Deployment Frequency
4. 😊 Team Satisfaction
5. 🛡️ Security Posture

Summary

Effective Agent =

Tool access × Strong memory

"It's not how smart the agent is it's how well it uses tools and remembers"



Way Forward

Towards self healing, self managing infrastructure

POLL: Which DevOps tasks do you hate and wish AI could handle for you?

url.pjay.in/k8sug

Questions?

Reach me at pjay.in