

The background features a dark navy blue field with abstract, overlapping organic shapes in shades of magenta and purple. Two thin, light blue lines intersect diagonally across the upper right portion of the image.

AWS re:Invent

DECEMBER 2 – 6, 2024 | LAS VEGAS, NV

KUB321

Streamline Amazon EKS operations with generative AI

Lucas Duarte

Sr Containers Specialist SA
AWS

Sai Vennam

Principal Containers Specialist SA
AWS



Kubernetes operations can be time-consuming



Why customize foundation models (FMs) for Kubernetes?



**Adapt to
domain-specific
language**

e.g., rephrase complex error logs
into easy-to-understand
language



**Enhance
performance
for specific tasks**

e.g., auto-tune resource request
and limits when working with
Karpenter and HPA

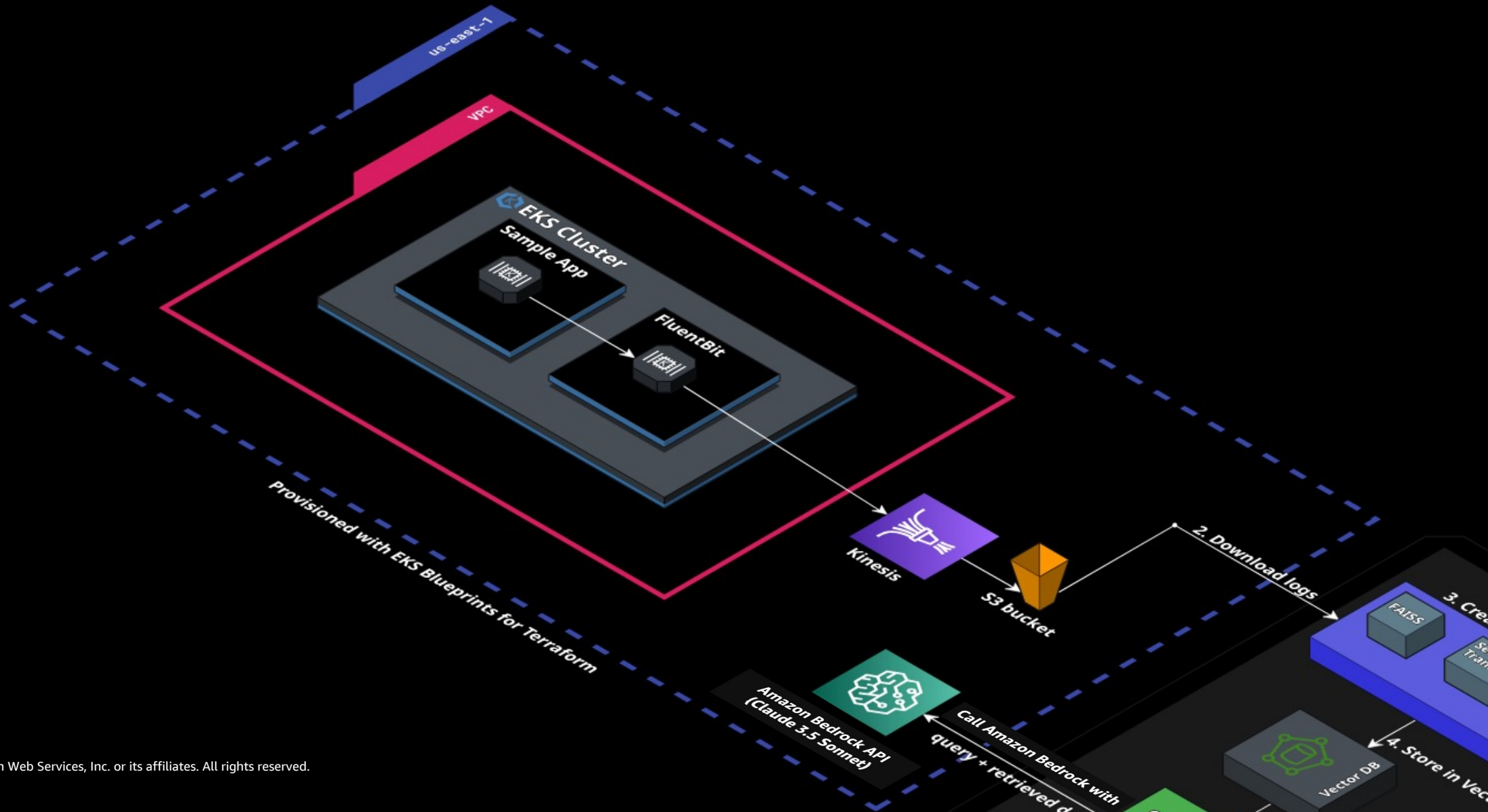


**Improve
context-awareness
in responses**

e.g., root cause analysis by
correlating logs and metrics
from multiple sources

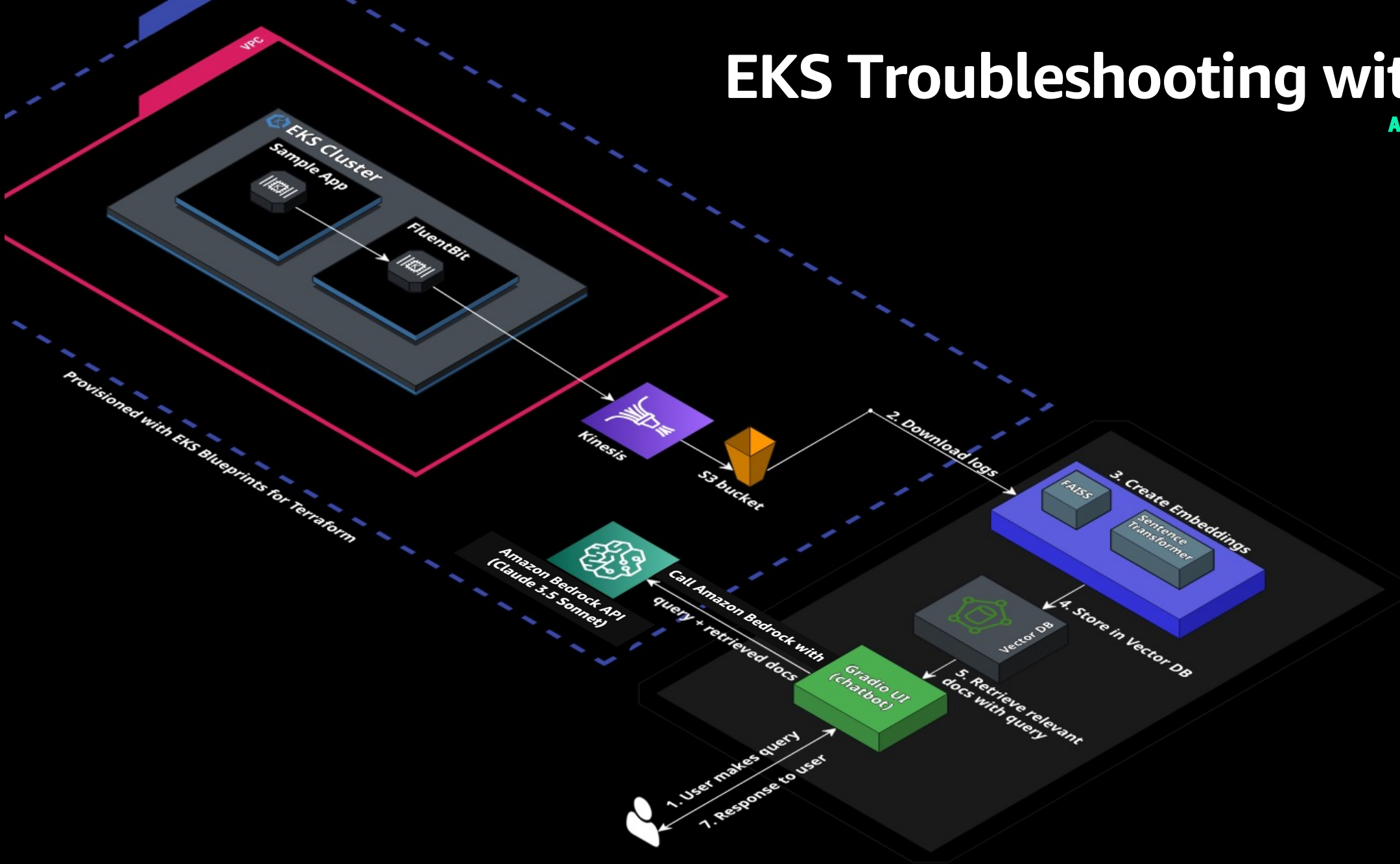
EKS Troubleshooting with LLMs

ARCHITECTURE

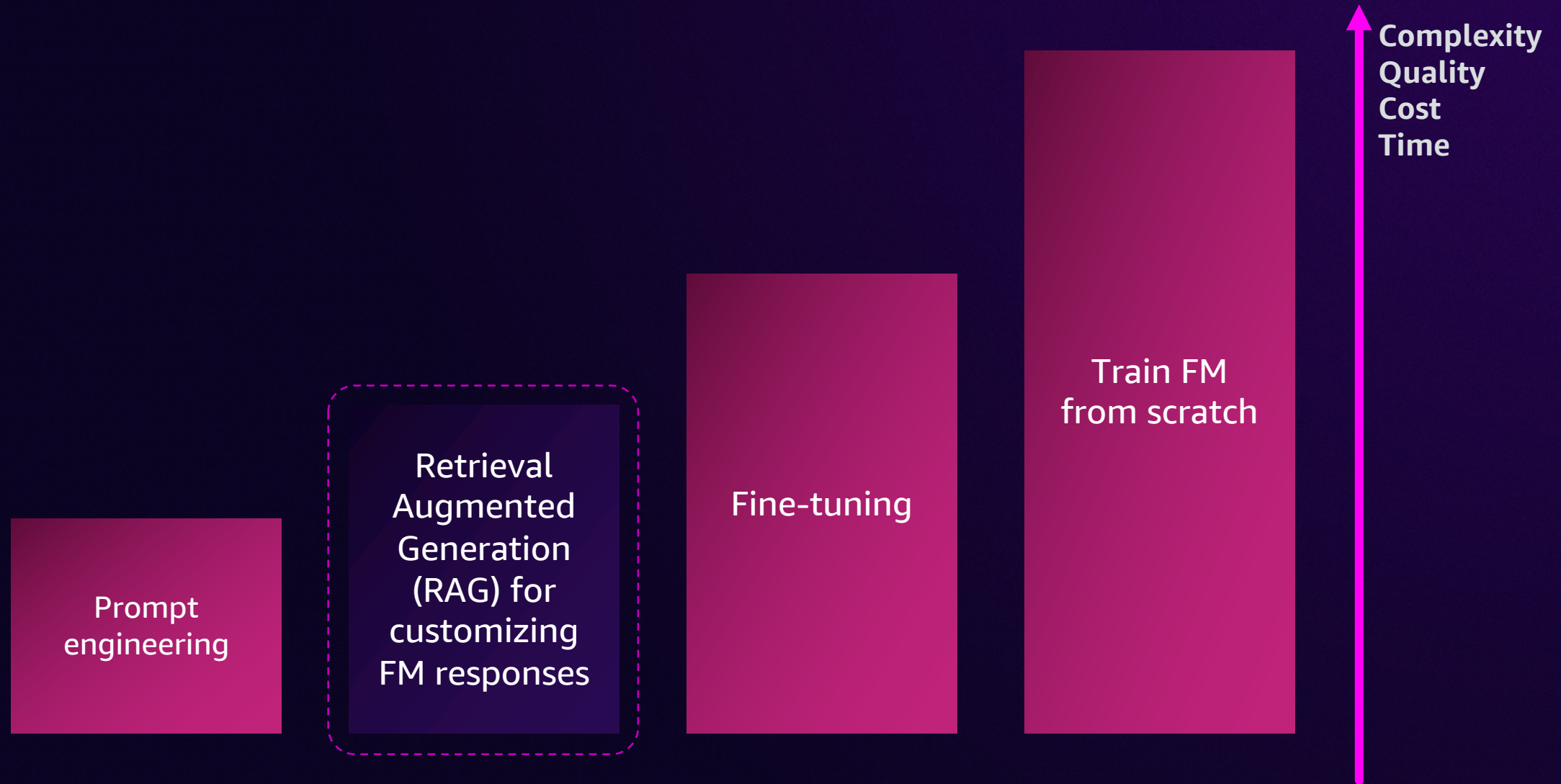


EKS Troubleshooting with LLMs

ARCHITECTURE



Common approaches for customizing FMs

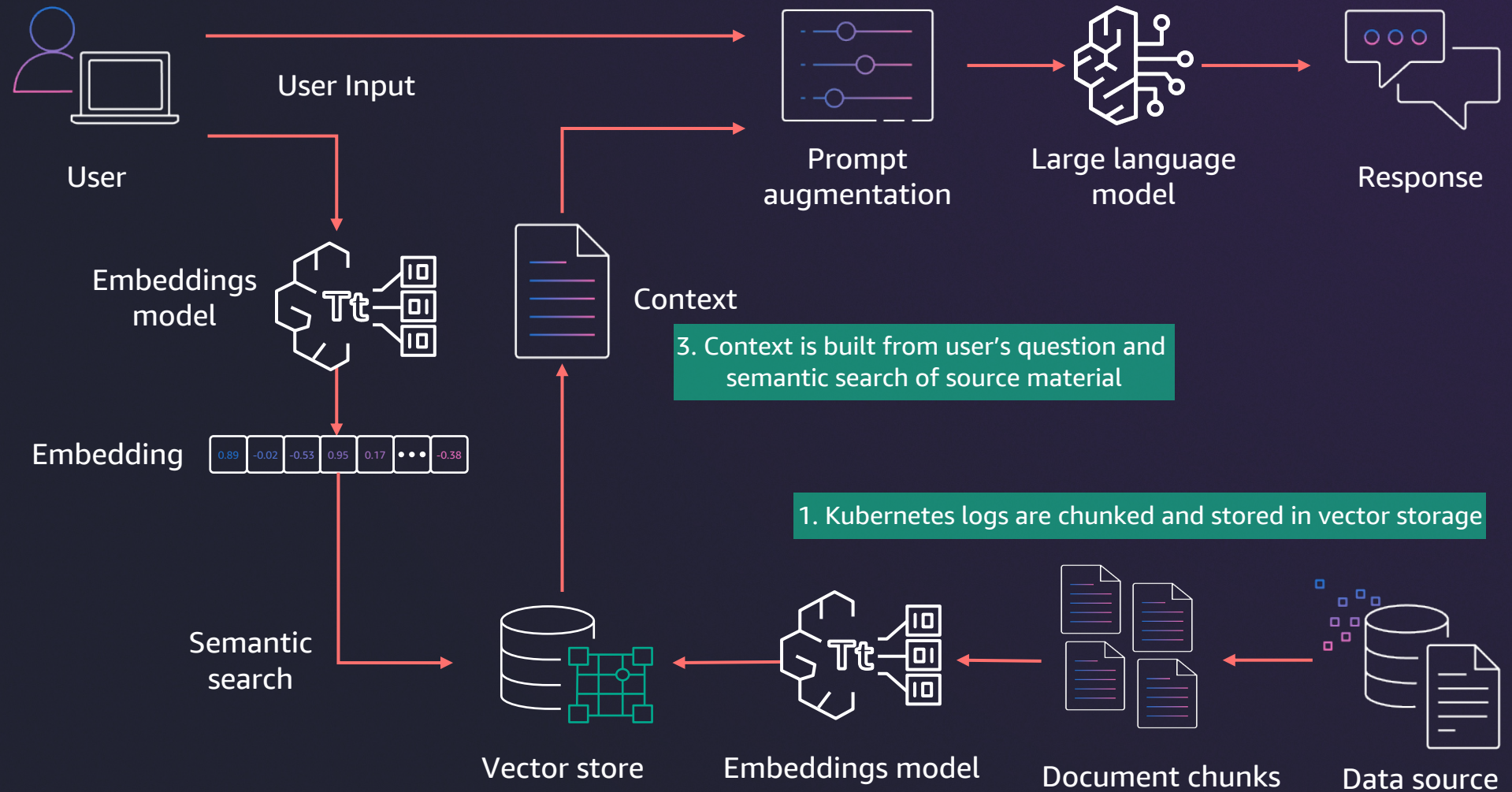


RAG in action

Text generation workflow

2. User asks a question: "Why is my app response time high?"

4. Context + original user input is fed to FM on Amazon Bedrock





Claude 3.5 Sonnet

MOST INTELLIGENT MODEL PRICED FOR HIGH-VOLUME, USER-FACING USE CASES

Core features

- Raises the industry bar for intelligence, outperforming other models on a wide range of evaluations
- Excels in coding, reasoning, and writing
- Industry-leading vision capabilities, especially with charts and data visualizations

Top skills

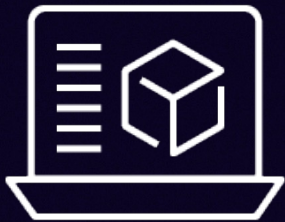
- Advanced reasoning
- Code generation
- Code translation
- Agentic tool use
- Visual reasoning
- Creative writing

Top use cases

- Software development, including migrations, code fixes, and code translations
- Data science, especially navigating unstructured data
- Customer support, with strong reasoning around ticket classification, triage, and resolution

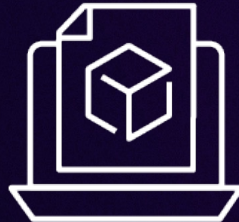
Continue your Amazon EKS learning

Learn at your
own pace



Take the **Amazon EKS workshop** to expand your EKS skills

Increase your
knowledge



Use the **Best Practices Guide** to build your Kubernetes knowledge

Earn Amazon
EKS badge



Demonstrate your knowledge by achieving **digital badges**



<https://github.com/aws-samples/reinvent24>

Session resources



github.com/aws-samples/reinvent24
sessions/KUB321

Thank you!



Please complete the session survey in the mobile app