
AWS
re:Invent

ANT402

Application observability with Amazon Elasticsearch Service

Rajiv Taori
Principal PMT-Search Services

Thanigai Vellore
Software Engineer

Eugene Ma
Software Engineer

AWS re:Invent

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Agenda

- Introductions
- “Observability” challenge customers face
- How Amazon Elasticsearch Service can help
- Airbnb case study
- Introducing Trace Analytics: A new feature for distributed tracing
- Next Steps



Amazon Elasticsearch Service

Fully managed service that makes it easy to deploy, manage, and run Elasticsearch cost effectively with industry-leading reliability, scalability, and security.



Open Distro for Elasticsearch

100% Apache licensed, open-source distribution of Elasticsearch with plugins to deploy a secure, performant, Elasticsearch cluster.

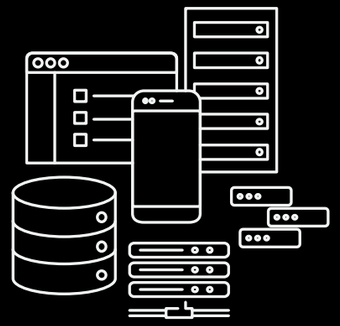
© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Machine-generated data is exploding



Applications and Services

- Services/Micro-services
- Web applications
- Business applications
- APIs



IT and DevOps

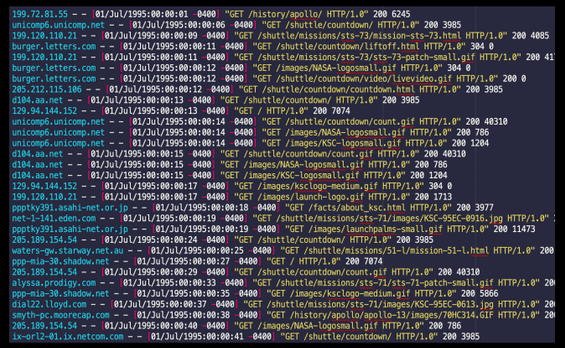
- Databases
- Load balancers
- Networking
- Servers



IoT and wireless

- Automotive
- Home devices
- Manufacturing
- Mobile applications

Finding the “needle in the haystack” is challenging













Amazon Elasticsearch Service delivers

- **Insights over analysis**
- **Open source not proprietary**
- **Machine Learning over rules-based**

AWS re:Invent

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Logging Infrastructure at Airbnb



Overview

Logs provide critical insights to better understand system behavior and user experience

100s
of services

1000s
of engineers

~ 300
TB / day

Challenges



Integration

Log Formats

Compute Ecosystems

Logs Consumption



Operational

Dynamic

Traffic pattern

Maintenance Overhead



Organizational

Total Cost of Ownership

Tenancy

Governance

Principles



Modularity

- Separation of concerns
- Loose coupling



Efficiency

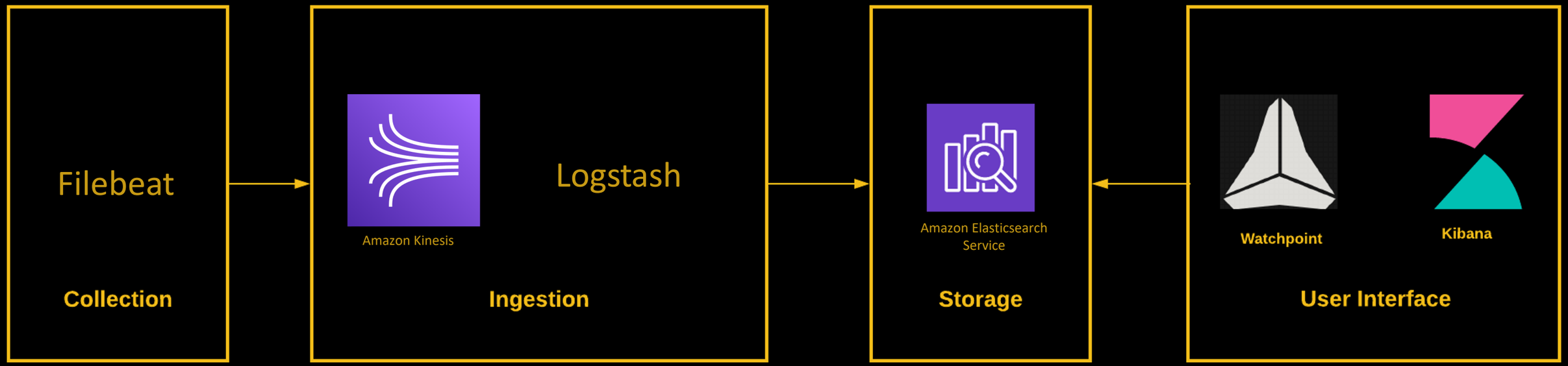
- Automation
- Self Healing



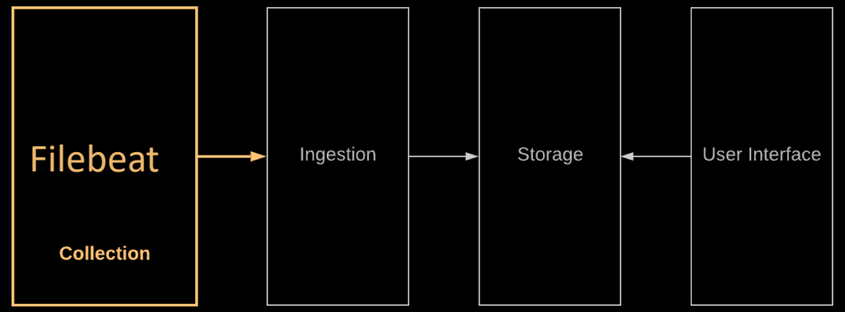
Multitenancy

- Consistency
- Governance

Architecture



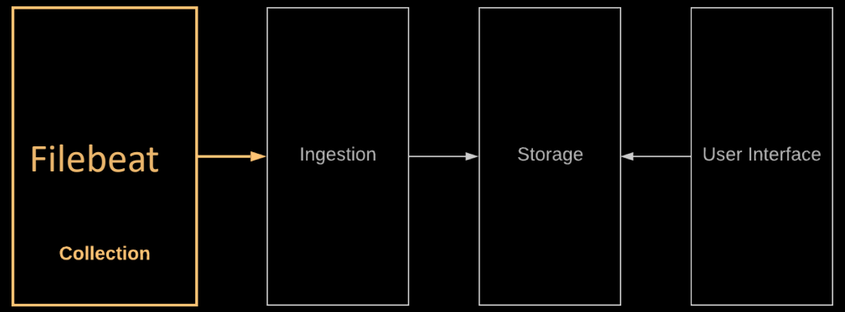
Collection



Filebeat

- Filebeat - standard log shipper
- Agent Deployment model
- Input: Local files (logs written to stdout)
- Output: Custom output plugin - Kinesis

Collection



Modularity

Structured Logging

Logs Enrichment

Efficiency

Batching Kinesis records

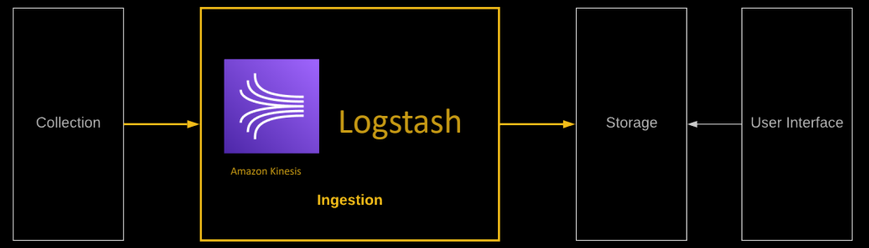
Closed loop Monitoring

Multi Tenancy

Schema Management

Rate Limiting

Ingestion



Amazon Kinesis

Buffers incoming logs

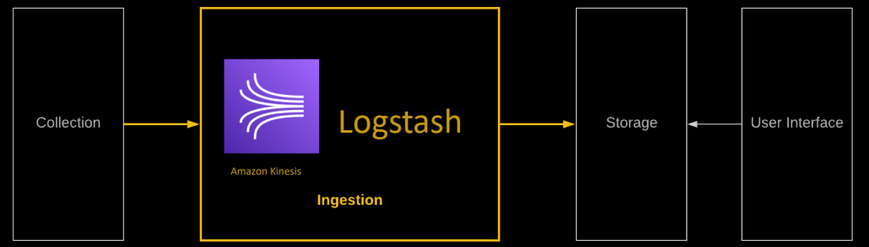
Fully managed

Logstash

Log enrichment and normalization

Bulk write to Elasticsearch

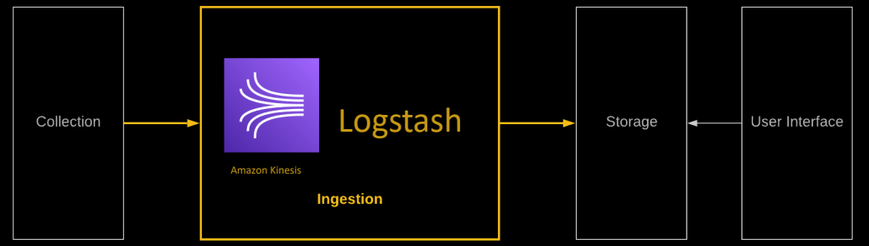
Ingestion



“Noisy neighbor” effect

How do we prevent a single tenant from monopolizing resources?

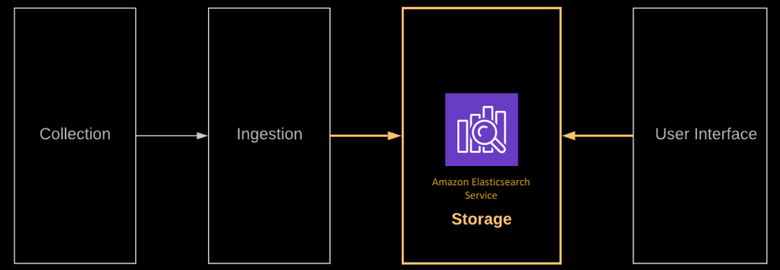
Ingestion



Multitenancy: Enforcing fairness

- Logging quotas with Kinesis
- Self-governance

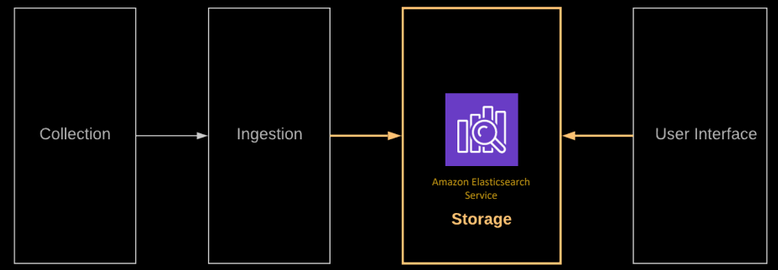
Storage



Amazon Elasticsearch Service

Decomposing monolithic clusters into many smaller clusters

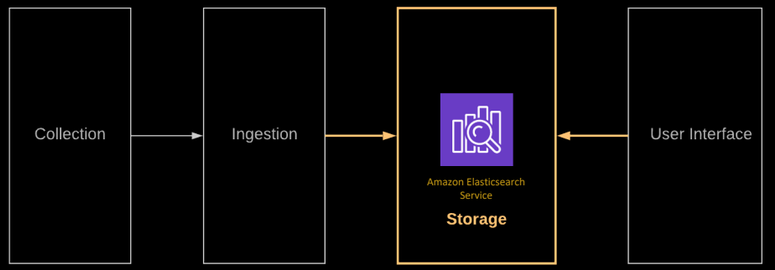
Storage



Efficiently managing many clusters

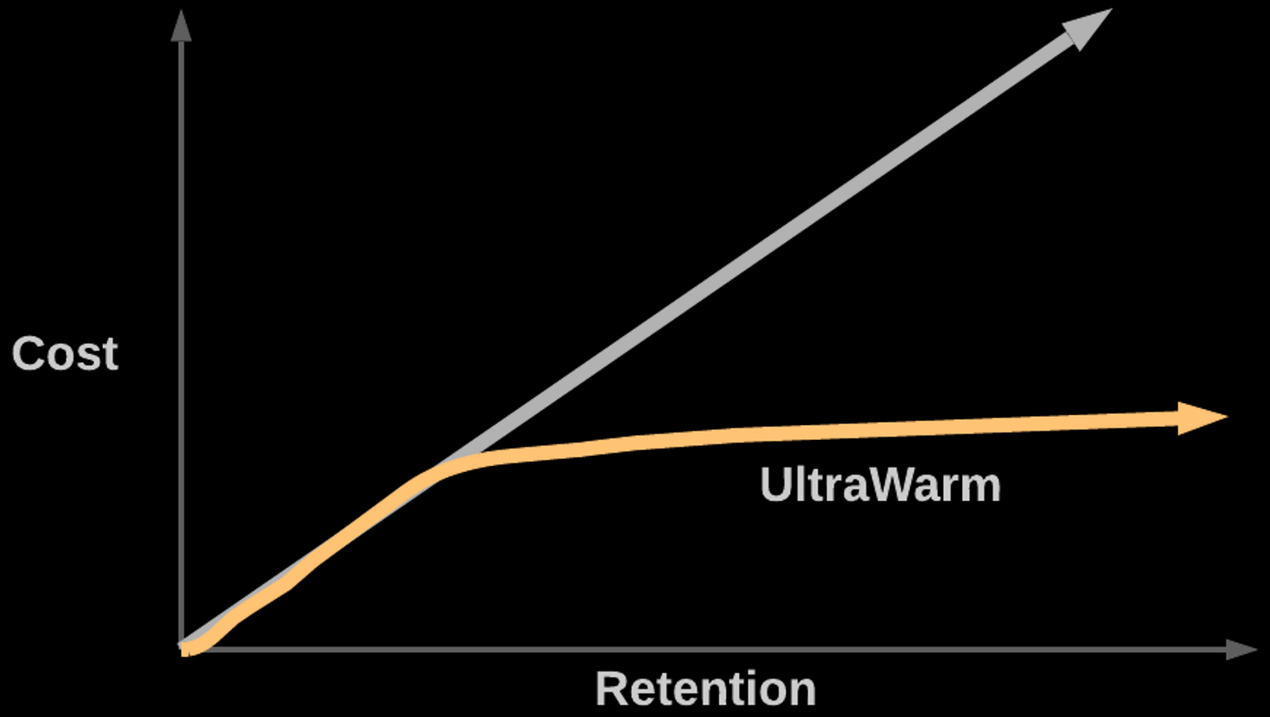
- Pets require special handling
- Cattle can share processes and tooling

Storage

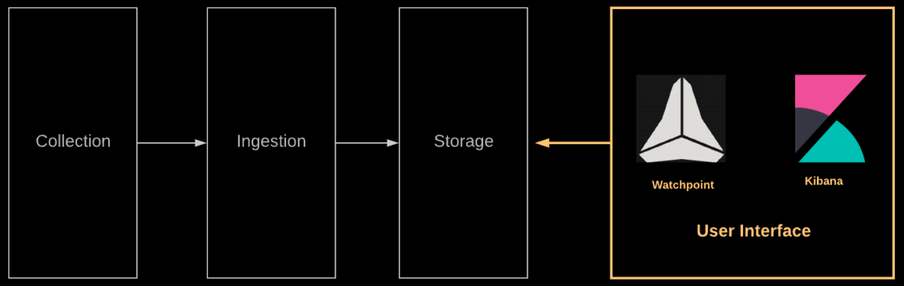


Maximizing cost efficiency

- Compression
- UltraWarm
- Automated rightsizing
- Configurable retention



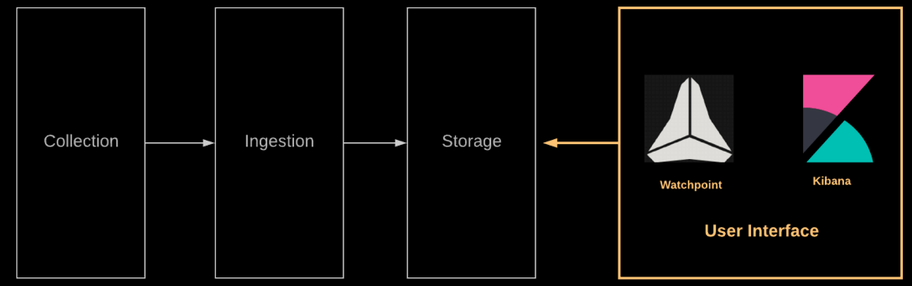
User Interface



Watchpoint

- Federation across clusters
- Simple, user-friendly UI

User Interface



Kibana

- Powerful tool for advanced users
- Deep links from Watchpoint

Impact



Integration

Integrates with diverse ecosystems

Integrates with traces and metrics



Operational

Cost efficient and elastic

0.5 FTE to keep the lights on

3 9’s of reliability, subsecond ingest latency



Organizational

Sustainable log volume growth

Self-governance

Appreciations



Rong Hu
Observability engineer

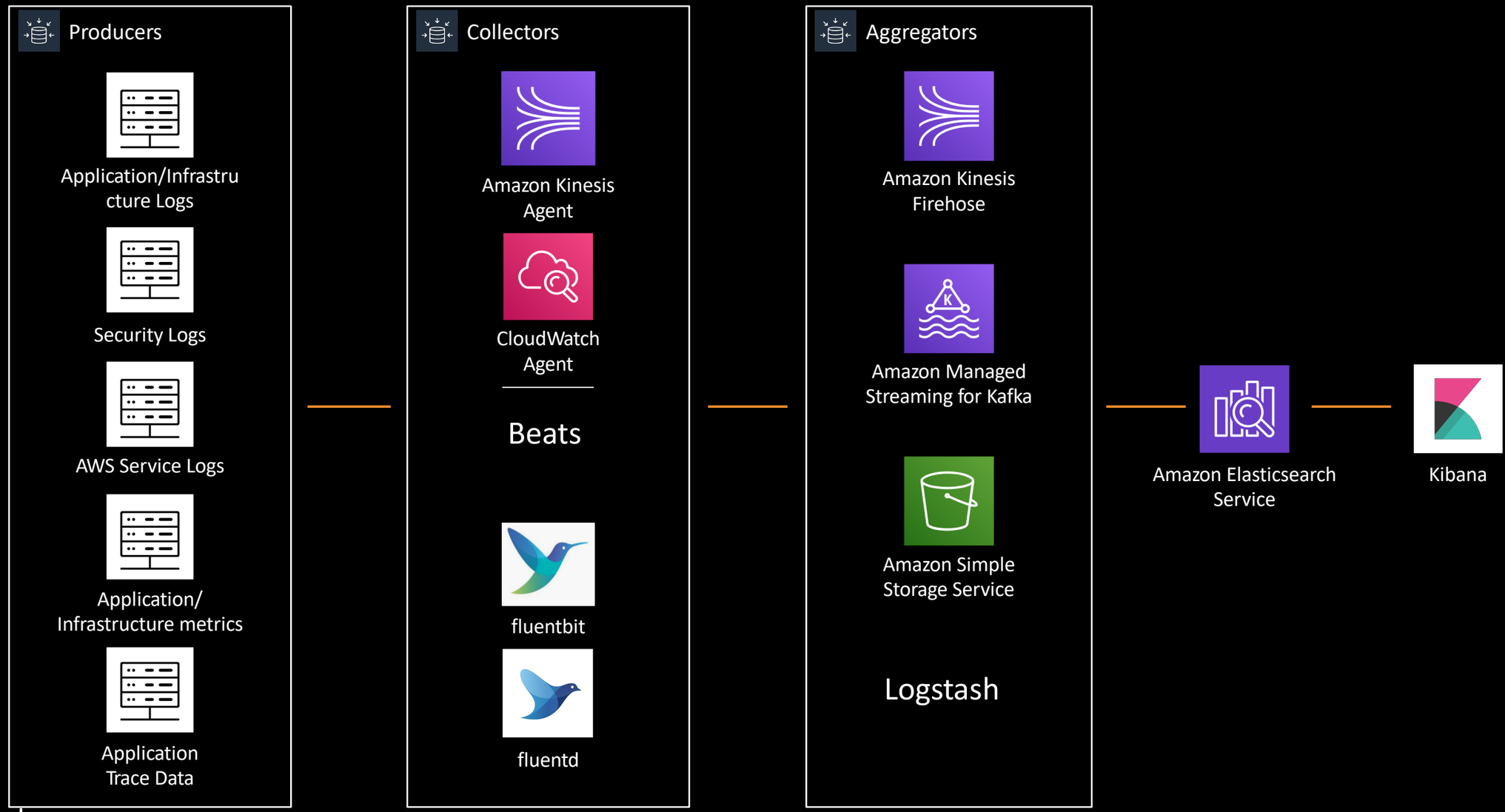


Xuan Lu
Observability engineer



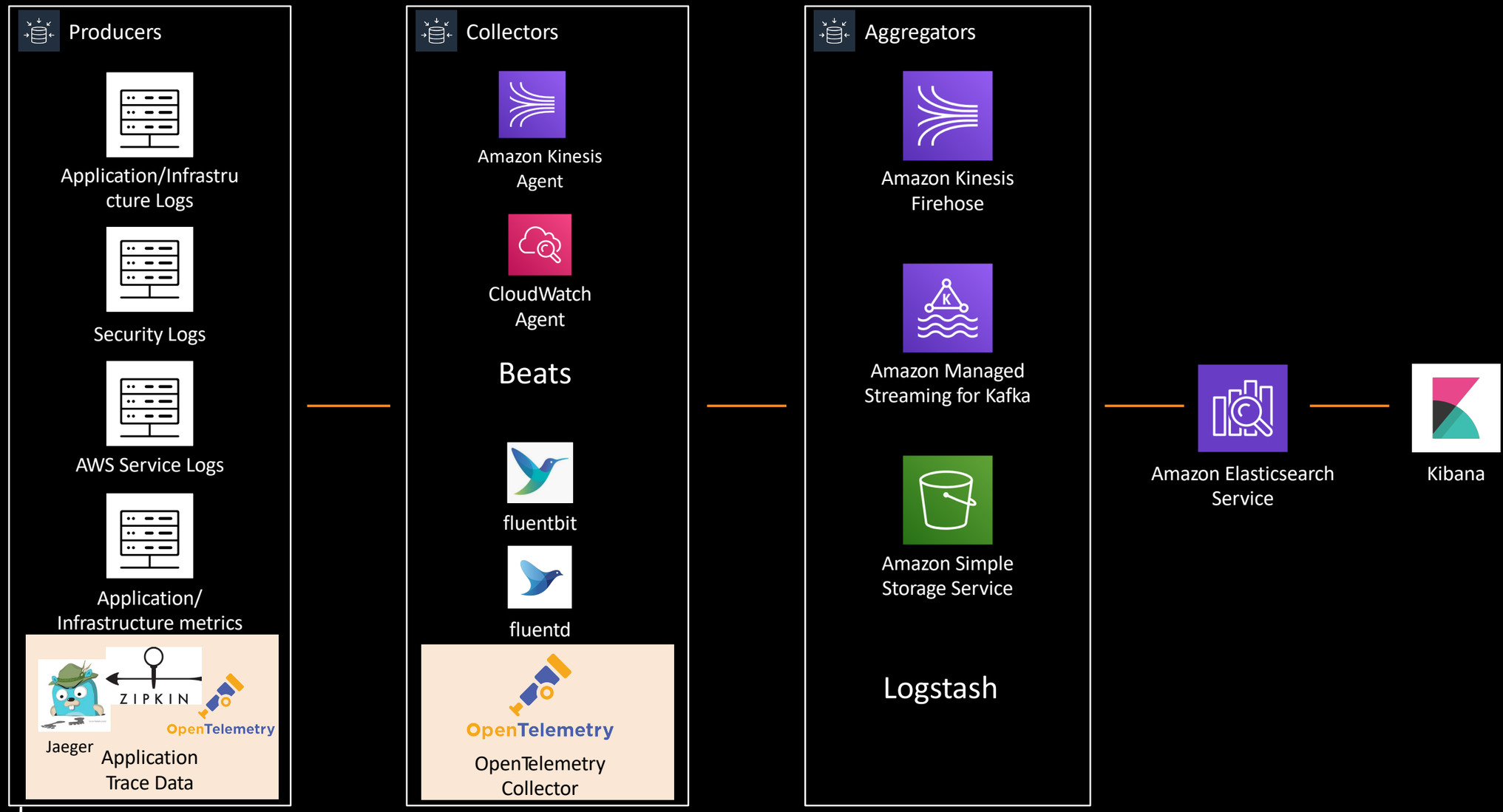
Amulya Sharma
AWS Solution Architect

Observability Data Flow



© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Observability Data Flow – What's new



© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

What is OpenTelemetry?



- OpenTelemetry is a Linux Foundation Cloud Native Computing Foundation (CNCF) open source project
- Formed in May 2019 from a merger of the OpenTracing and OpenCensus projects
- Mission to provide
  - Libraries, agents, and other components to comprehensively unify monitoring, managing and debugging of applications and services
  - Supports 11 languages and growing
- Captures metrics, distributed traces, resource metadata, and logs from applications and services

AWS Distro for OpenTelemetry

A secure, production-ready open source distribution supported by AWS



- Upstream-first distro of the popular CNCF project
- Certified by AWS for security and predictability
- Backed by AWS support
- One-click deploy and configure from AWS Container and AWS Lambda consoles
- Exporters for AWS monitoring solutions including - CloudWatch, X-Ray, Amazon Elasticsearch Service, and Partner Solutions

Next Steps

- ODFE home page: https://opendistro.github.io/for-elasticsearch/
- Developer Portal: https://github.com/opendistro-for-elasticsearch
- Roadmap: https://github.com/orgs/opendistro-for-elasticsearch/projects/3
- Feature requests, bugs, Discussion Forum: https://discuss.opendistrocommunity.dev/c/trace-analytics/49

Thank you!

Rajiv Taori

Principle PMT – Search Services

Eugene Ma

Software Engineer

Thanigai Vellore

Software Engineer

AWS re:Invent

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.



Please complete the session survey

AWS re:Invent

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.