KubeCon | CloudNativeCon

Europe 2021

Virtual

Forward Together »

KubeCon CloudNativeCon

Europe 2021

Virtual

Petabyte scale logging with Fluentd and Fluent Bit: A use case from Intuit

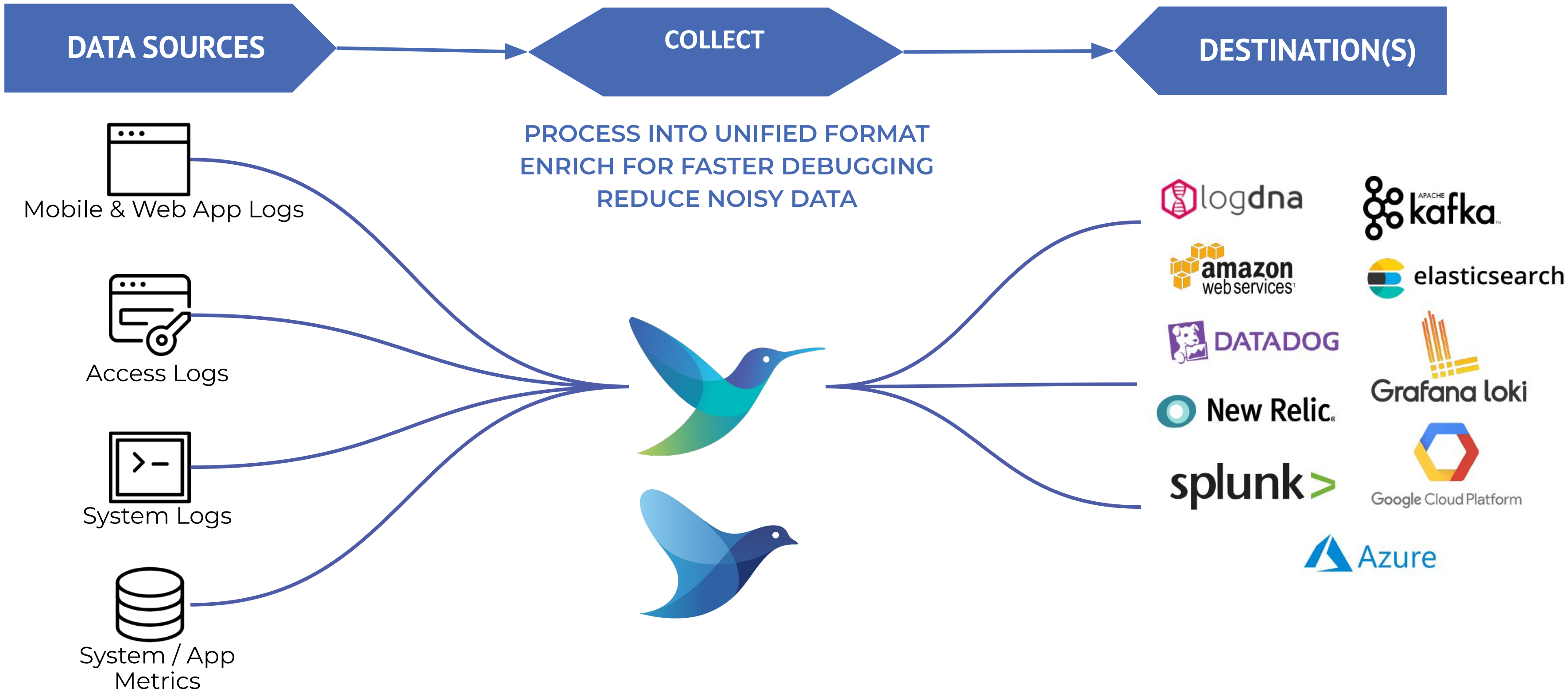
Hanzel Jesheen, Senior Software Engineer, Intuit
Anurag Gupta, Product, Calyptia

Who are we?

Anurag Gupta
Product
OSS Maintainer Fluent Bit
Calyptia

Hanzel Jesheen
Senior Software Engineer
Cloud Observability
Intuit

What are Fluentd / Fluent Bit?



DATA SOURCES

COLLECT

DESTINATION(S)

PROCESS INTO UNIFIED FORMAT
ENRICH FOR FASTER DEBUGGING
REDUCE NOISY DATA

Mobile & Web App Logs

Access Logs

System Logs

System / App Metrics

Challenges for logs @ scale

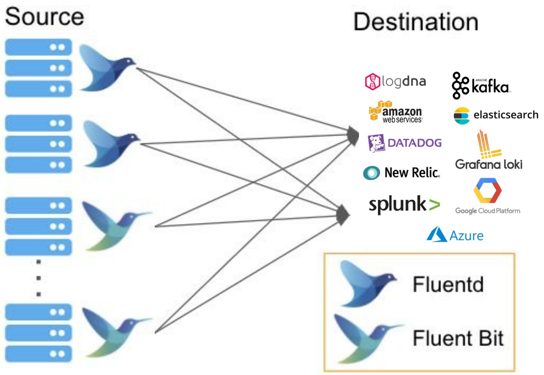
- **High Scale can equal high costs!!**
- **Reliability and buffering**
- **Networking**
- **Event Throughput**
- **Security**
  - Securing sensitive information.
  - Securing the data transition.
- **Operationality**
  - Minimizing log collector operations in data source.

Challenges for logs @ scale

- **High Scale can equal high costs!!** - Filtering, Parsing, compression
- **Reliability and buffering** - Filesystem and Memory buffers
- **Networking** - Configurable retry mechanisms, Backpressure handling
- **Event Throughput** - Multi-worker configuration
- **Security** - TLS in transit
  - Securing sensitive information.
  - Securing the data transition.
- **Operationality** - Forwarder / Aggregator architecture
  - Minimizing log collector operations in data source.

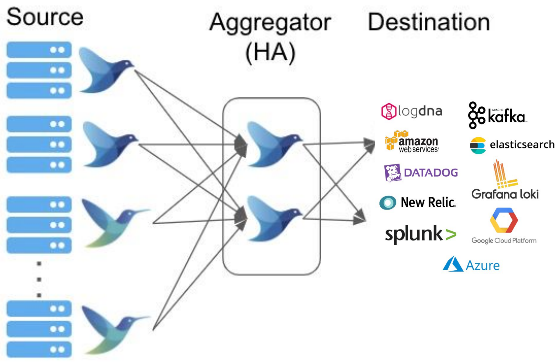
Common Architecture

Forwarder only



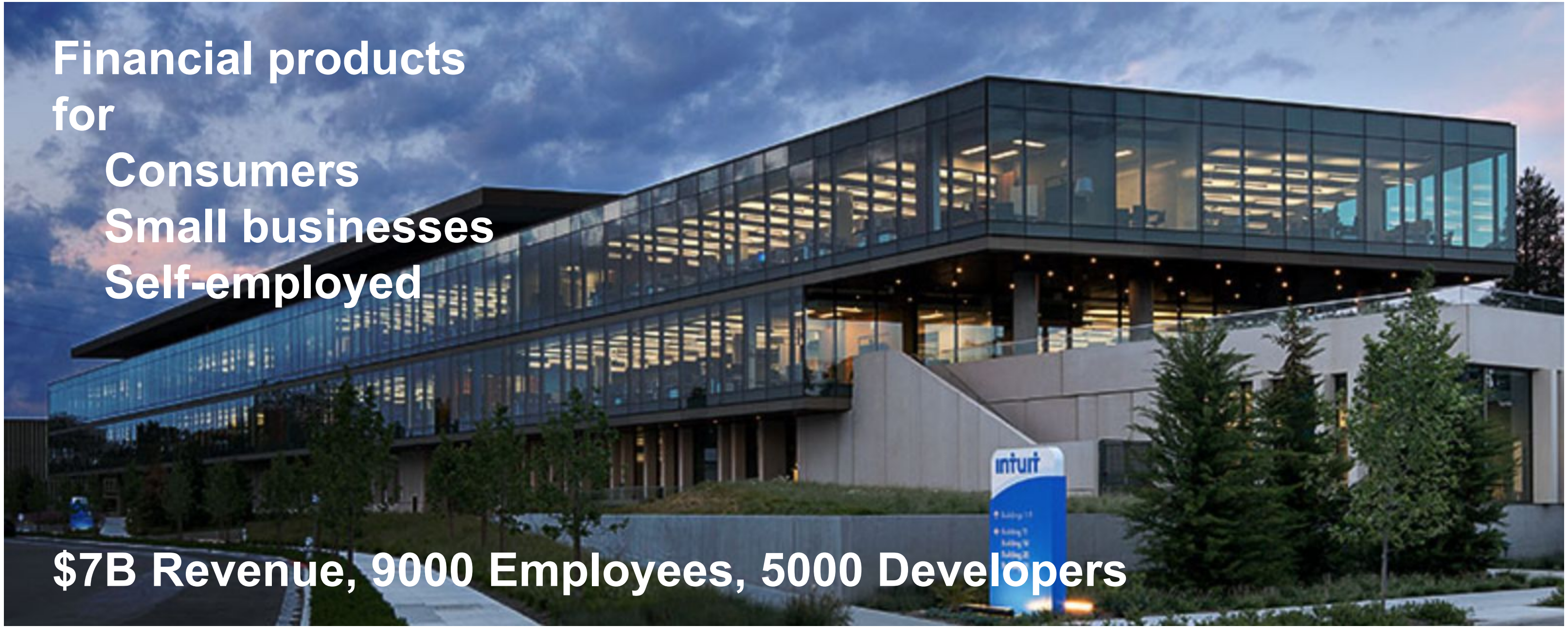
- **Advantages**
  - No aggregator is needed; each forwarder handles backpressure.
- **Disadvantages**
  - Hard to change configuration across a fleet of forwarder (E.g., adding another backend or processing)
  - Hard to add more end destinations if needed

Forwarders with aggregators



- **Advantages**
  - Less resource utilization on the edge devices
  - Allow processing to scale independently on the aggregator tier.
  - Easy to add more backends (configuration change in aggregator vs. all forwarders)
- **Disadvantages**
  - Dedicated resources required for an aggregation instance

intuit

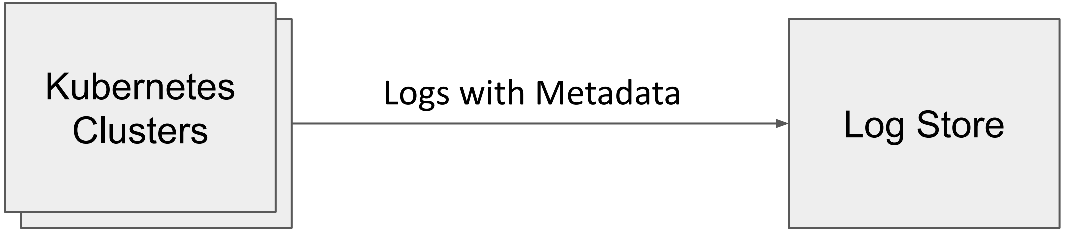


Financial products for
- Consumers
- Small businesses
- Self-employed

$7B Revenue, 9000 Employees, 5000 Developers

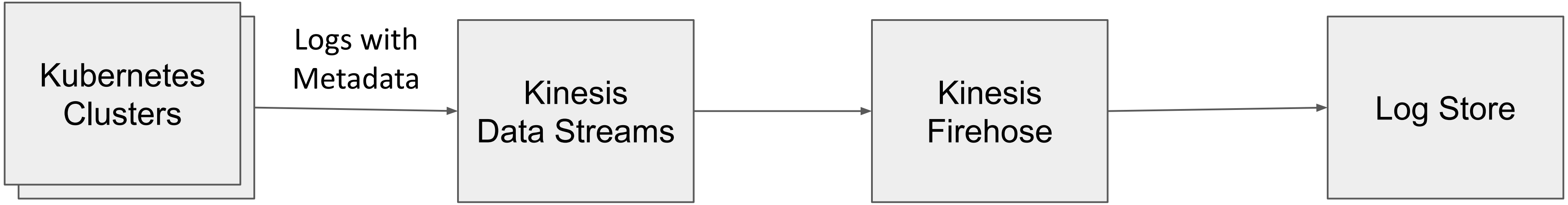
turbotax. quickbooks. mint.

Logging for Kubernetes



- At Intuit, log analytics is a core capability that is offered by a centrally hosted log store.
- Among others, containers running on Kubernetes are a major log sources. There are 100+ Kubernetes clusters hosting 2000+ services.
- Fluentd processes running as daemonsets are used to collect and forward the logs to the store.
- Log events are enriched with metadata to generate insightful correlations and improve search experience.
- High throughput & Low latency pipeline is desirable.

Streaming Pipeline



- Highly distributed log sources that are spread across 100+ VPCs. Transfer data between Kubernetes VPCs and log store VPC.
- Durable, fault tolerant, and scalable log data pipeline is required.
- Ability to fan out the data to multiple stores to solve for additional requirements like security, compliance etc.

Better Logging Pipeline

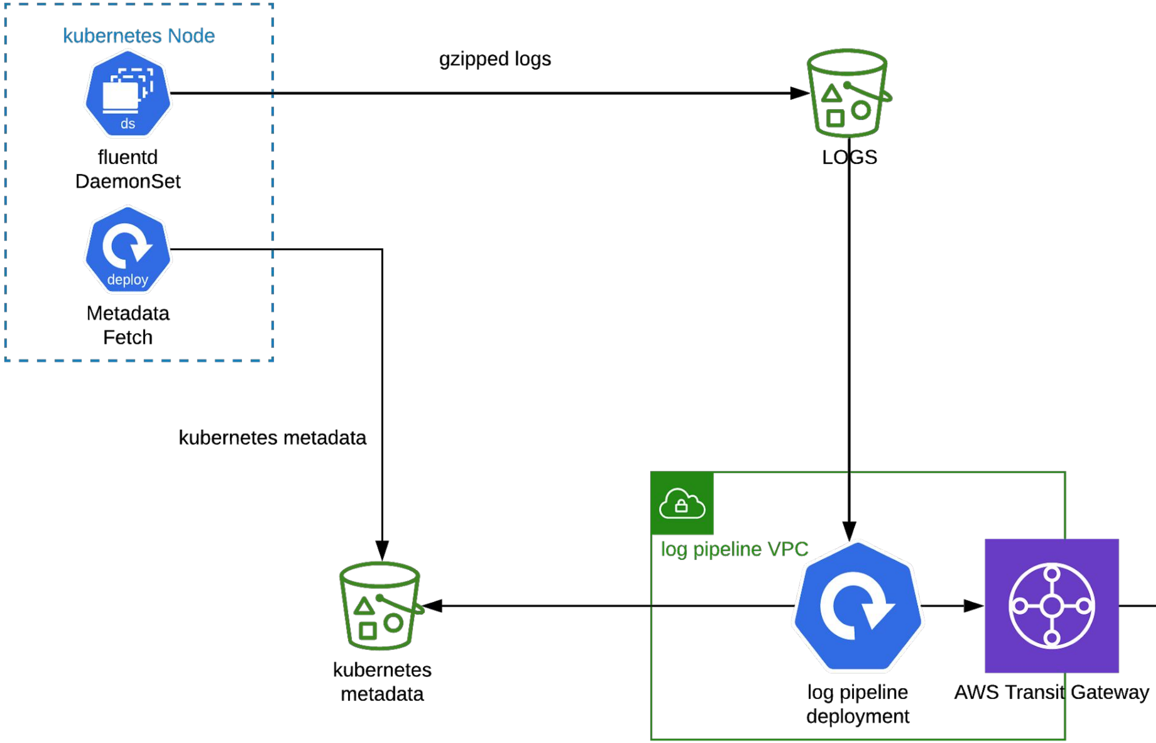
Challenges faced with streaming transport

- Multiline events need to be identified and packaged as a single record at the source. This added to the work that fluentd process has to do at the collection time. This severely limited the collection throughput.
- Metadata enrichment at the source added to the fluentd workload.
- Dequeuing from the stream required additional hop adding to latency as well as cost.

Problems to be solved

- Increase collection throughput with minimal overhead.
- Low end-to-end latency to transport between source and target.
- Reduce the cost to maintain the pipeline.

S3 Pipeline: Architecture



S3 Pipeline: Improvements

Target: Increase Throughput

- Minimize the work done by fluentd.
- Avoid multi-line detection
  - Eliminate the need for CPU intensive timestamp parsing.
  - Maintain the chronology of events and offload multiline detection task to the log store.
- Avoid Metadata Enrichment
  - Export kubernetes metadata from each cluster and enrich log events in transit.

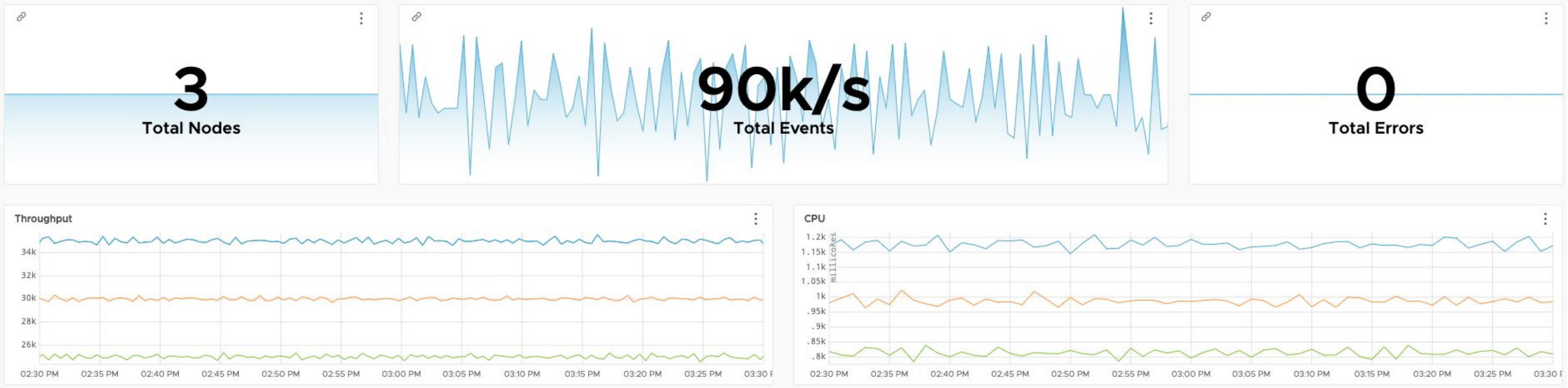
S3 Pipeline: Improvements

Target: Minimize Latency & Reduce Cost

- Network transfer cost is the highest component. So, reducing the data transferred will reduce both cost and latency.
- Fluentd writes compressed data to s3 (~10X compression) and it’s written to Log Store as is. So, the data is always compressed in transit and decompression happens at the log store.
- Metadata is applied in batch and need not be added to each log event.
- AWS Transit Gateway to transfer data between Log Pipeline VPC and Log Store VPC.

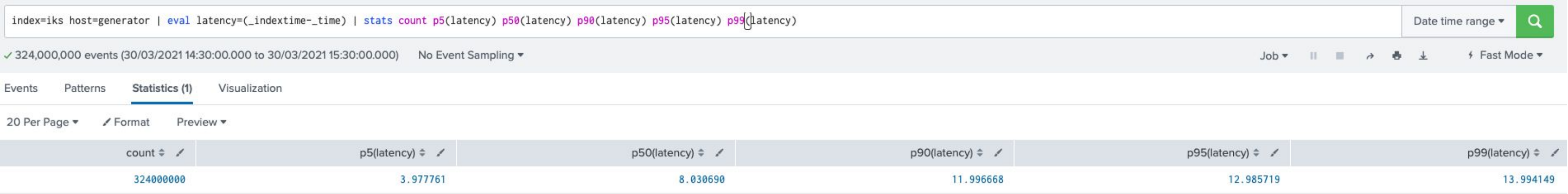
S3 Pipeline: Demo

Throughput



S3 Pipeline: Demo

End-to-end Latency



```
index=iks host=generator | eval latency=(_indextime-_time) | stats count p5(latency) p50(latency) p90(latency) p95(latency) p99(latency)
```

✓ 324,000,000 events (30/03/2021 14:30:00.000 to 30/03/2021 15:30:00.000) No Event Sampling ▾

count	p5(latency)	p50(latency)	p90(latency)	p95(latency)	p99(latency)
324000000	3.977761	8.030690	11.996668	12.985719	13.994149

- Total Events: 324 Million (90,000 events/s over 1 hour)
- Latency
  - 5th Percentile: < 4 seconds
  - Median / 50th Percentile: ~ 8 seconds
  - 90th Percentile: < 12 seconds
  - 95th Percentile: < 13 seconds
  - 99th Percentile: < 14 seconds

S3 Pipeline: Results

Throughput

6x

- Supports **6-times** throughput, compared to the streaming pipeline, at **30,000 events per seconds per node** while consuming just 1.2 core CPU.
- Supporting more than **1 GB/s** of log data across the pipeline for a single cluster.

Latency

⅓

- Median End-to-end latency cut down to less than **one-third** from 30 seconds to just 8 seconds.
- For **99th Percentile**, the latency is cut down by more than **75%**.

Cost

92%

- More than **92%** cost saved when compared to streaming pipeline.
- More than **$50,000** saved for every PB transported.