Deploying Apache Kafka
Exabyte Scale Data Coordination

James Mountifield
Enterprise Sales Engineer - Sumo Logic
Agenda
- Who am I
- Who are Sumo Logic
- Collection Architecture
- Lessons Learned
- Some Cool Stats
Who am I?
Who are Sumo Logic?

Build, run and secure your AWS, Azure, Google Cloud Platform or Hybrid applications with Sumo Logic, a cloud-native, machine data analytics service for log management and time series metrics.

One platform for real-time, Continuous Intelligence.

100+ PETABYTES DATA ANALYZED DAILY
20+ MILLION SEARCHES PERFORMED DAILY
1500+ CUSTOMERS
Scale: Exponential Growth

(100 petabytes) / (300 bytes) = ~ 3.33... \times 10^{14}
Log events in the “live” dataset

Daily Stats:
- 100+ Petabytes Analyzed
- 30+ Million Queries Run
- 500+ Trillion Records Examined
Architecture Context Prerequisites

- Cloud Native / SaaS Architecture built on AWS

- The service is NOT a public interface, because ...

- Constantly subject to new services, modifications of existing services, new technology
  - EC2 -> Lambda
  - ELB -> ALB

- The “service” including it’s APIs are independent of its architecture
- Sumo Logic is deployed across several regions, all Availability Zones
  - North America (North Virginia, Oregon, GovCloud)
  - Europe (Frankfurt, Dublin)
  - APAC (Sydney)

- True multi tenant with relevant security built in from the beginning
Our Collection Architecture
Our Collection Architecture
Our Collection Architecture
Demo Time
Thank You