Logging and Metrics in Elastic Cloud: Drinking Our Own Champagne

Tim Banks, Site Reliability Engineer
Stephanie Jackson, Site Reliability Engineer

Elastic Cloud
1 March 2018

Tim - tw: @elchefe gh: timbanks
Stephanie - tw: @stejacks1 gh: stejacks

Topics Covered:

1 The Backstory: What we started with.

2 What logs? Who cares?

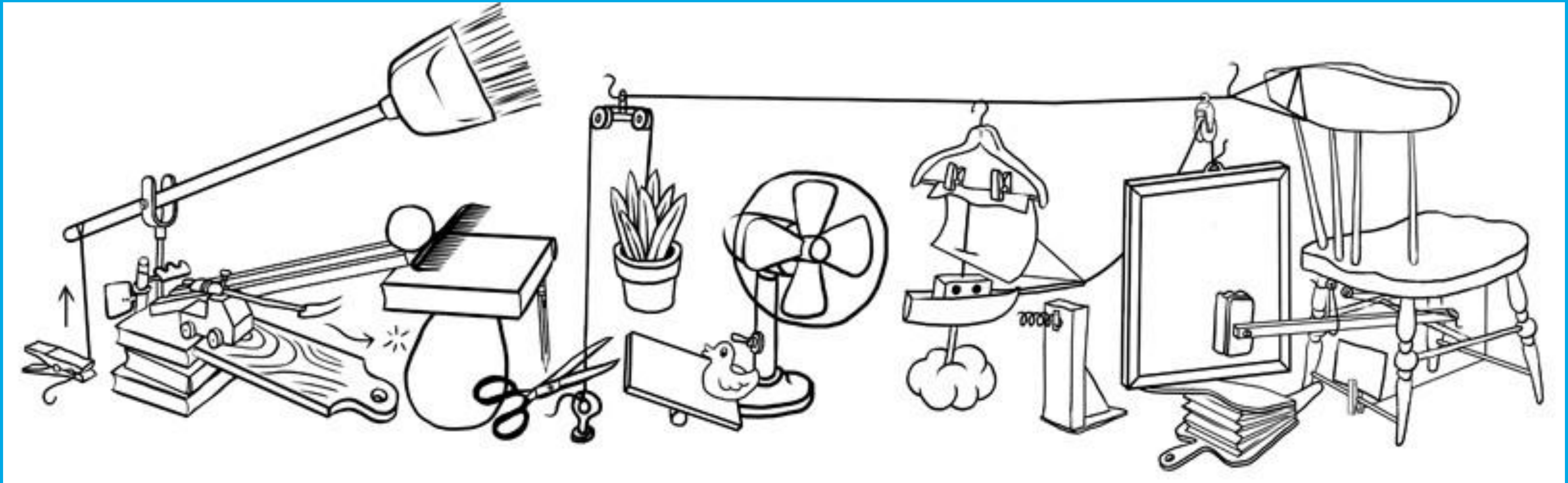
3 What are we going to do/what are our options?

4 Implementation

5 Plans for the future

The Backstory

What we started with



Legacy logging infrastructure architectural diagram

What We Started With

Legacy Logging Solution

- Datadog collectors
- Sysdig
- Elastic 1.7.x
- Kibana 4.x

What We Started With (continued)

Problems with the legacy logging solution

- Inconsistent data types (logs or metrics? Where did they go?)
- VERY old versions of the elastic stack
- Single point of failure for the cluster

What we started with (continued)

- No automated management
- Upgrading was nigh unto impossible
- Increasing technical debt

“This is suboptimal.”

Elastic Cloud SREs

What logs, and for who?

Logos? We don’t need no stinking logs.

Except, of course, we do.

- Cloud Infrastructure and Docker logs
- Elasticsearch and Kibana logs
- Metrics

So many different use cases

Cloud Team

- Cloud Infrastructure and Elasticsearch Logs
- Oncall and Break/Fix

Support

- Customer Logs and Errors
- Historical and Current Metrics

Customers

- Cluster logs
- Cluster metrics

What are our options?

Sources, Data stores, and queues, oh my.

It takes a village...

Requirements

- Using Elastic Cloud as Backend
- Elastic stack version released this decade
- Single Pane of Glass
- Ability to scale
- Cloud agnostic

cloud

Initial Discussions

- Should we use a central location or regional clusters?

- How will the data be sourced?

- Do we need an aggregation/pipeline tier?

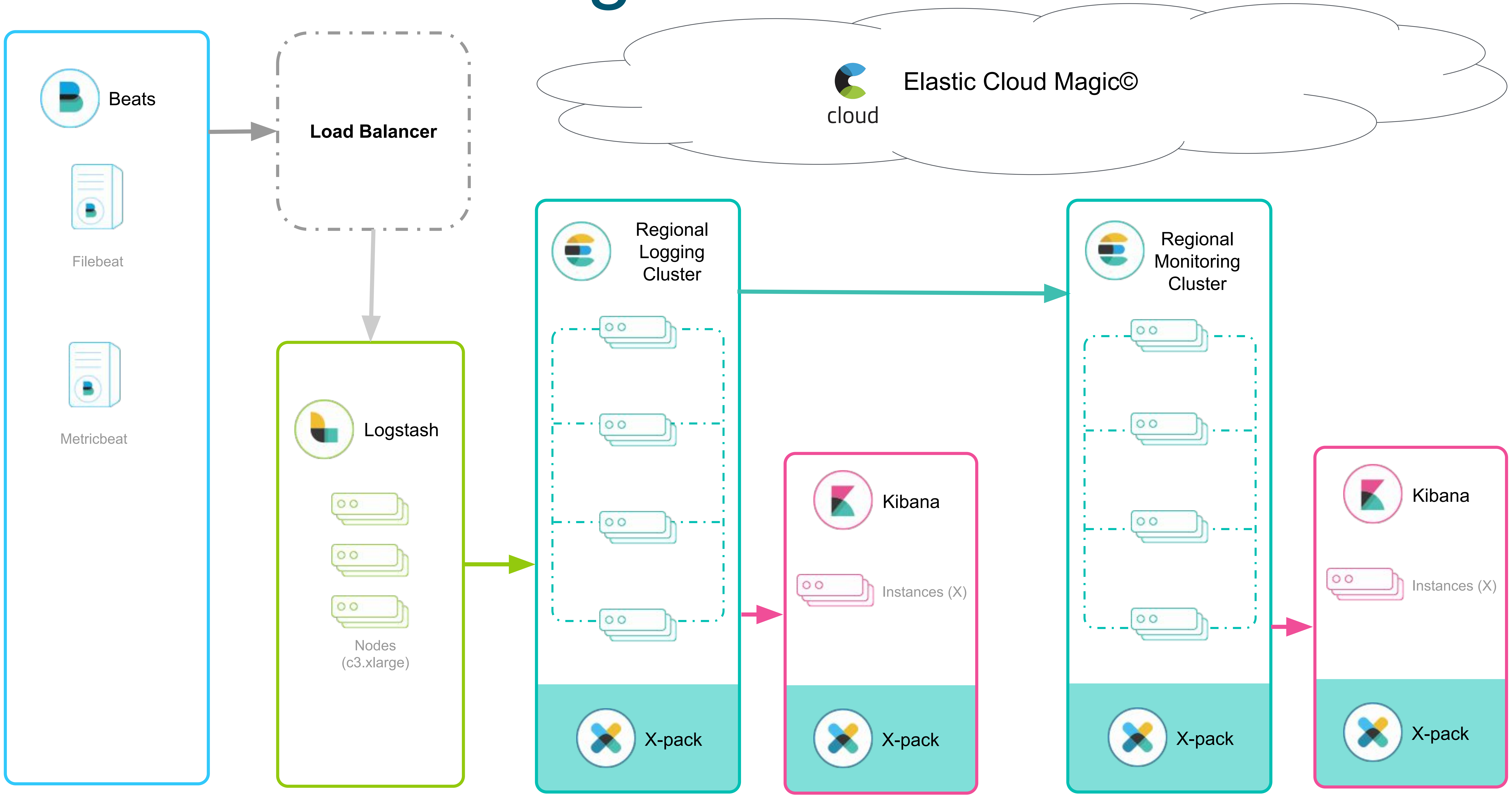
- What will we do for long-term data storage?

- What tools will be used for visualization and management?

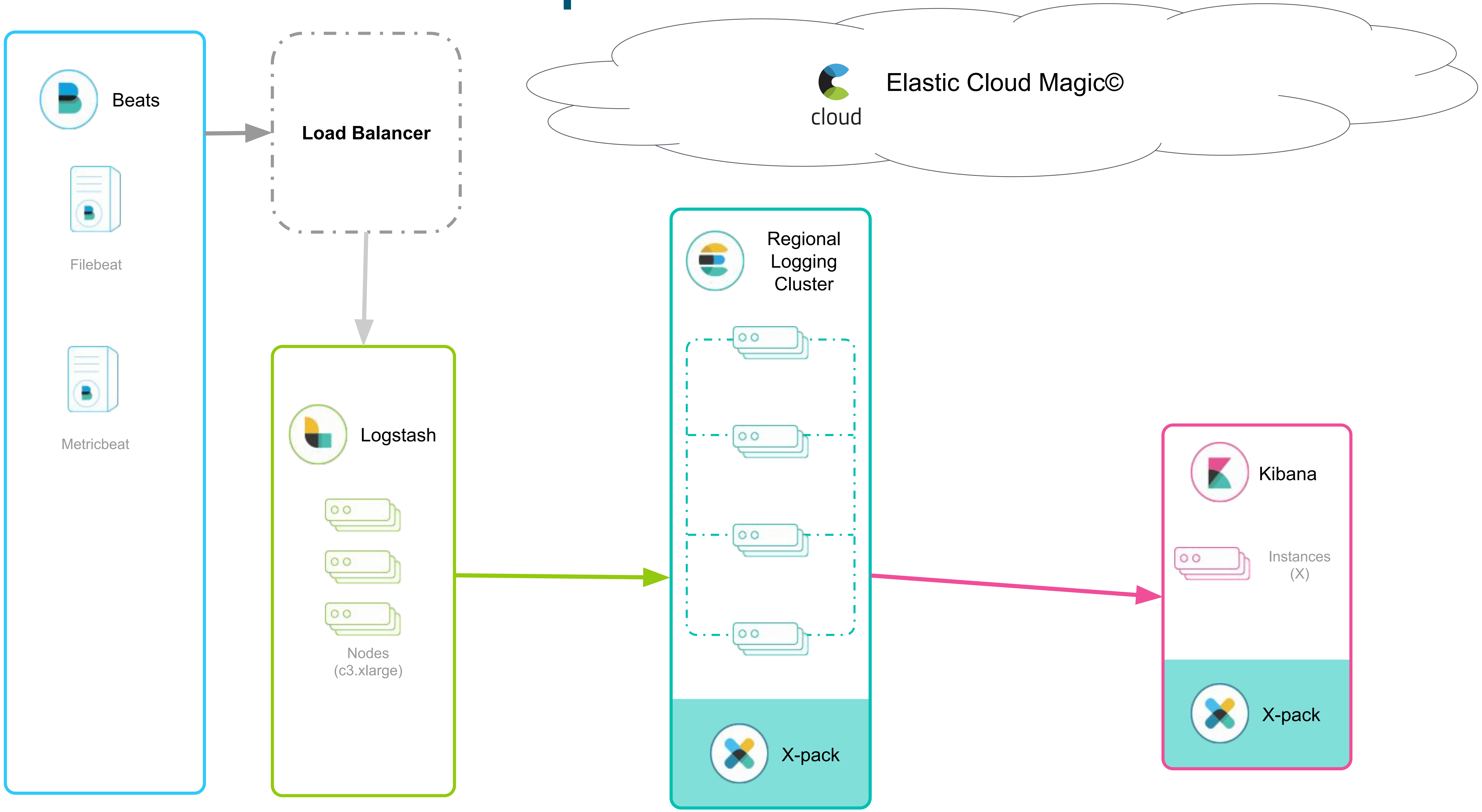
The New, Better, Shinier Logging

Our architecture and lessons learned

Architecture Diagram



Proof of Concept



Initial Rollout

- Small to Big
- Brand new shiny clusters
- Dashboards and watches and alerting, oh my!
- Limited retention

Lessons learned

And a few more grey hairs earned

- Really Big Regions Cause Really Big Problems.
- Monitoring is vital
- We Found Bugs(™)

Current state

HUUUUUUGE

- We’re #1
- Split backends for large regions
- Multi-cloud
- No SPOG

How big is this monster really?

As of February 20, 2018

- 8 AWS regions
- 4.2 TB of filebeat data per day
- 1.2 TB of metricbeat data per day
- 250,000 requests per second
- 40 billion + documents
- Hundreds of logstash processes

Future plans

New toys, better reliability and more

Future Plans

We will use all their toys!

- ML! APM!
- Sliders!
- SPOG
- Capacity Planning

More Questions?

Visit us at the AMA

elastic

www.elastic.co