instaclustr

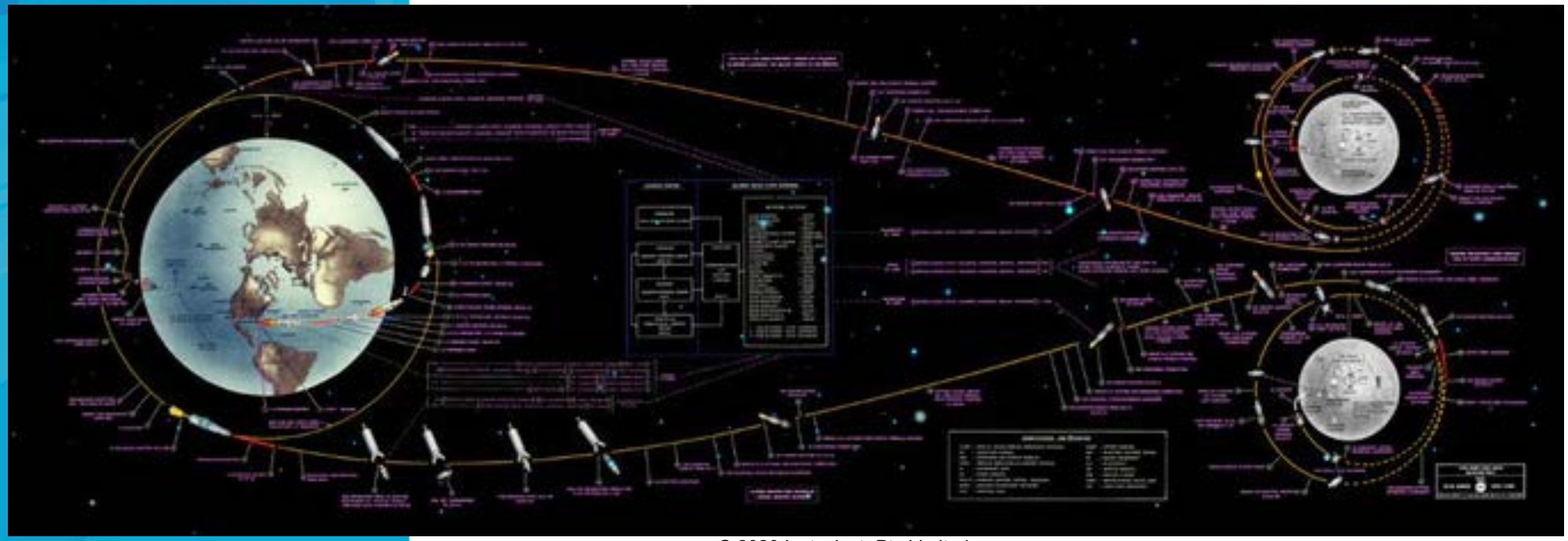
Kafka, Cassandra and Kubernetes at Scale –

Real-time Anomaly Detection on 19 Billion events a day

Paul Brebner
instaclustr.com Technology Evangelist

Overview

1. Wow! (headlines)
2. Why? (did we do it)
3. What? (does it do)
4. How? (does it work))
5. Well? (how well did it work)
6. So What?



1 Wow!?!

Headlines

instaclustr

Headlines 50 years ago

ELVIS The Memphis Record EXTRA

1969: YEAR IN REVIEW

ACADEMY AWARD WINNERS

GRAMMY WINNERS

COMEBACK

THE YEAR'S NO. 1 HITS

ELVIS PRESLEY COMMEMORATIVE ISSUE

- PRESLEY RETURNS HOME FOR FIRST MEMPHIS SESSIONS SINCE MID-1950's
- FIRST MAN WALKS ON MOON
- HALF MILLION IN ATTENDANCE AT WOODSTOCK FESTIVAL
- NEW YORK METS WIN WORLD SERIES
- NIXON SWORN IN AS NATION'S 37TH PRESIDENT
- JETS UPSET COLTS

instaclustr

Headlines 50 years ago

instaclustr

Headlines 51 years ago


ELVIS The Memphis Record EXTRA

1969: YEAR IN REVIEW

ACADEMY AWARD WINNERS

GRAMMY WINNERS

COMEBACK

PRESLEY RETURNS HOME FOR FIRST MEMPHIS SESSIONS SINCE MID-1950's

FIRST MAN WALKS ON MOON

HALF MILLION IN ATTENDANCE AT WOODSTOCK FESTIVAL

NEW YORK METS WIN WORLD SERIES

NIXON SWORN IN AS NATION'S 37TH PRESIDENT

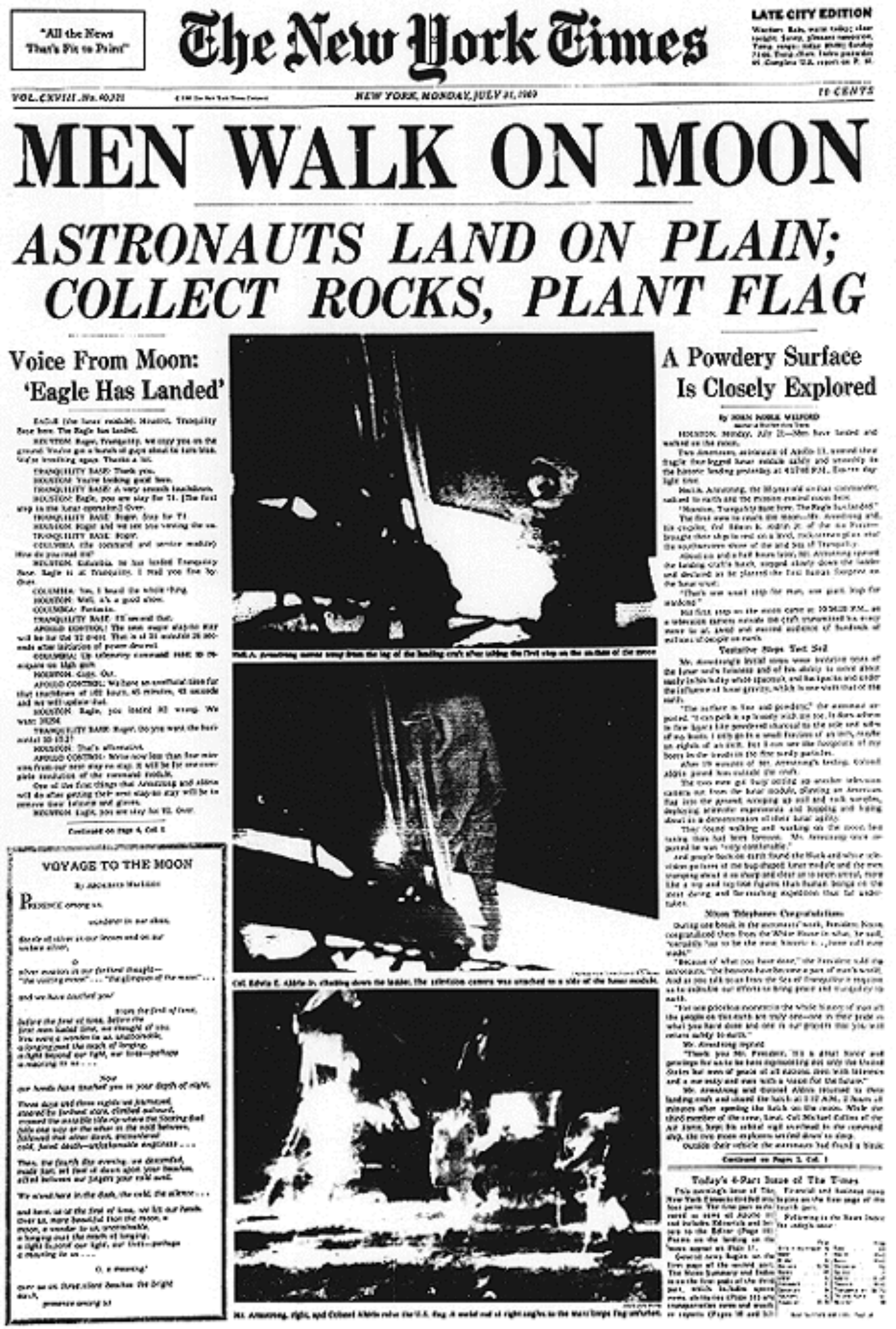
JETS UPSET COLTS

THE YEAR'S NO. 1 HITS

ELVIS PRESLEY COMMEMORATIVE ISSUE

instaclustr

Headlines 51 years ago


"All the News That's Fit to Print"

The New York Times

LATE CITY EDITION

10 CENTS

MEN WALK ON MOON

ASTRONAUTS LAND ON PLAIN; COLLECT ROCKS, PLANT FLAG

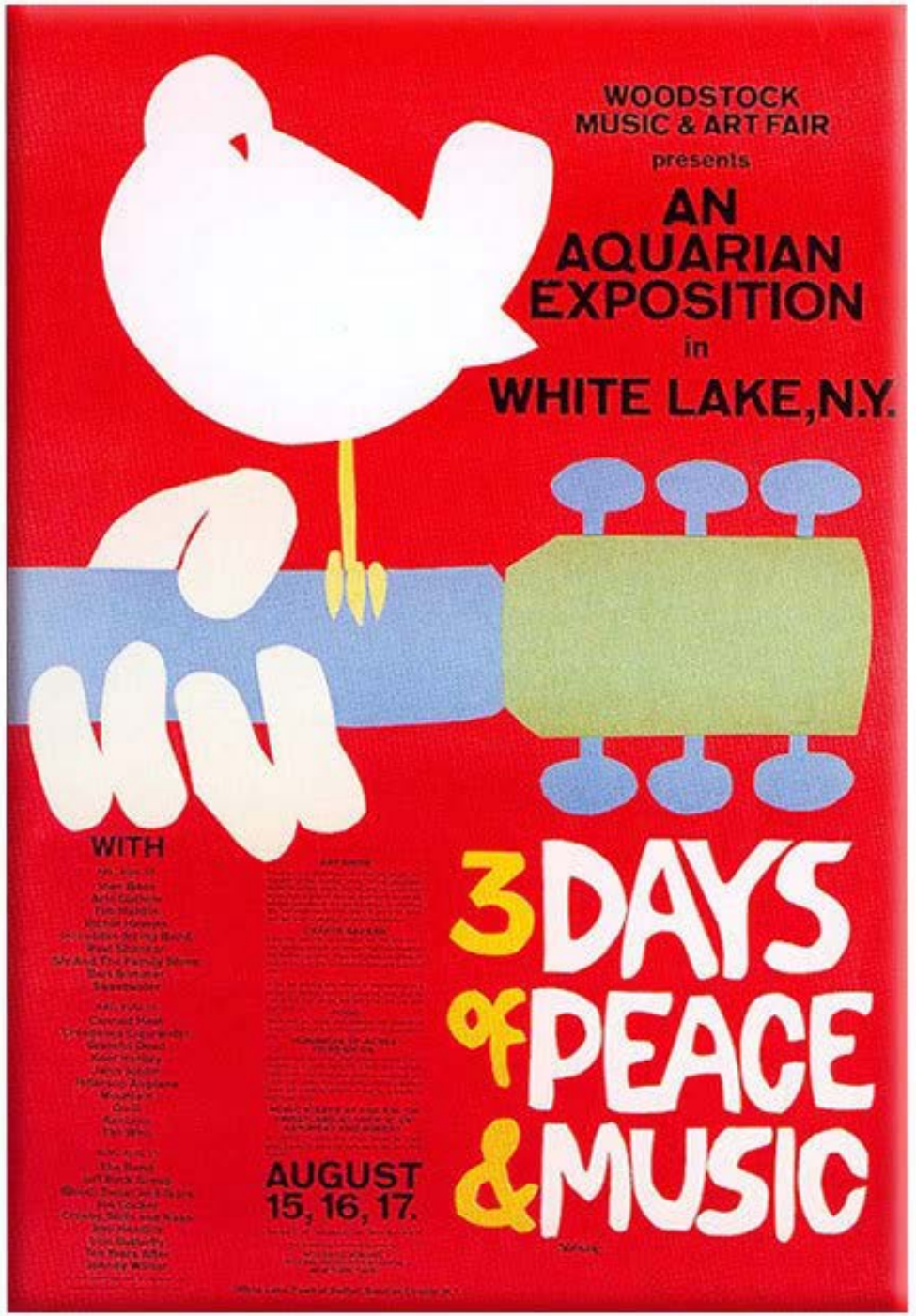
Voice From Moon: 'Eagle Has Landed'

A Powdery Surface Is Closely Explored

VOYAGE TO THE MOON

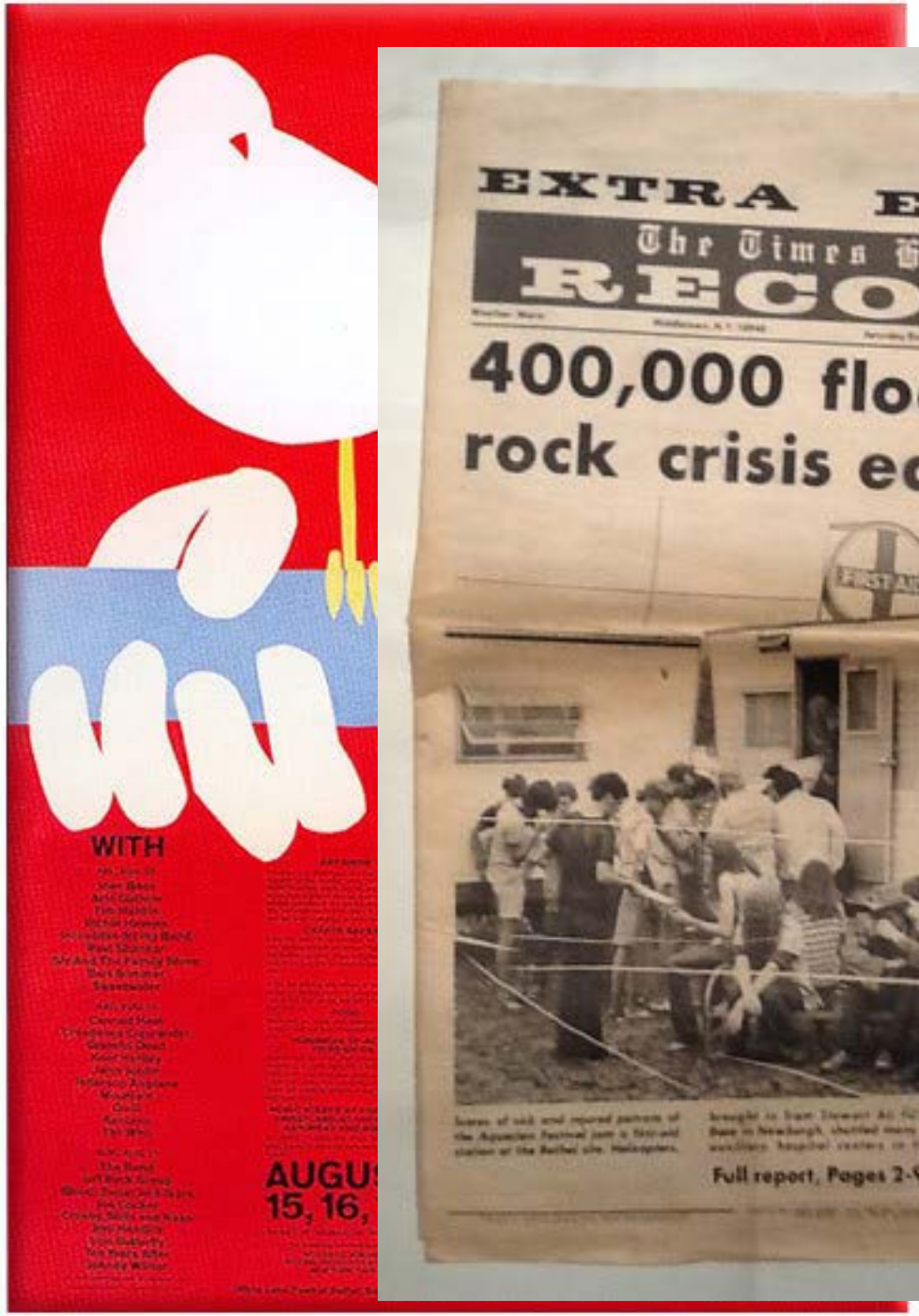
instaclustr

Headlines 51 years ago



instaclustr

Headlines 51 years ago



50,000 expected
1 Million descended on Woodstock
500,000 reached the venue

instaclustr

Forward 50 years

"All the News That's Fit to Print" — The Instaclustr Times — LATE CITY EDITION

VOL. CXVIII...No. 40,311 — NEW YORK, MONDAY, JULY 21, 1969 — 10 CENTS

19 Billion

Anomaly Checks A Day!

Instaclustr reveals

Massively Scalable!

Fast! Affordable!

Anomaly Detector Machine

Using Open Source

Apache Cassandra,

Apache Kafka,

Kubernetes & AWS

Press Release (2019)

instaclustr

Forward 50 years

"All the News That's Fit to Print" — The Instaclustr Times — LATE CITY EDITION

VOL. CXVIII ... No. 40,311 — NEW YORK, MONDAY, JULY 21, 1969 — 10 CENTS

19 Billion

Anomaly Checks A Day!

Instaclustr reveals

Massively Scalable!

Fast! Affordable!

Anomaly Detector Machine

Using Open Source

Apache Cassandra,

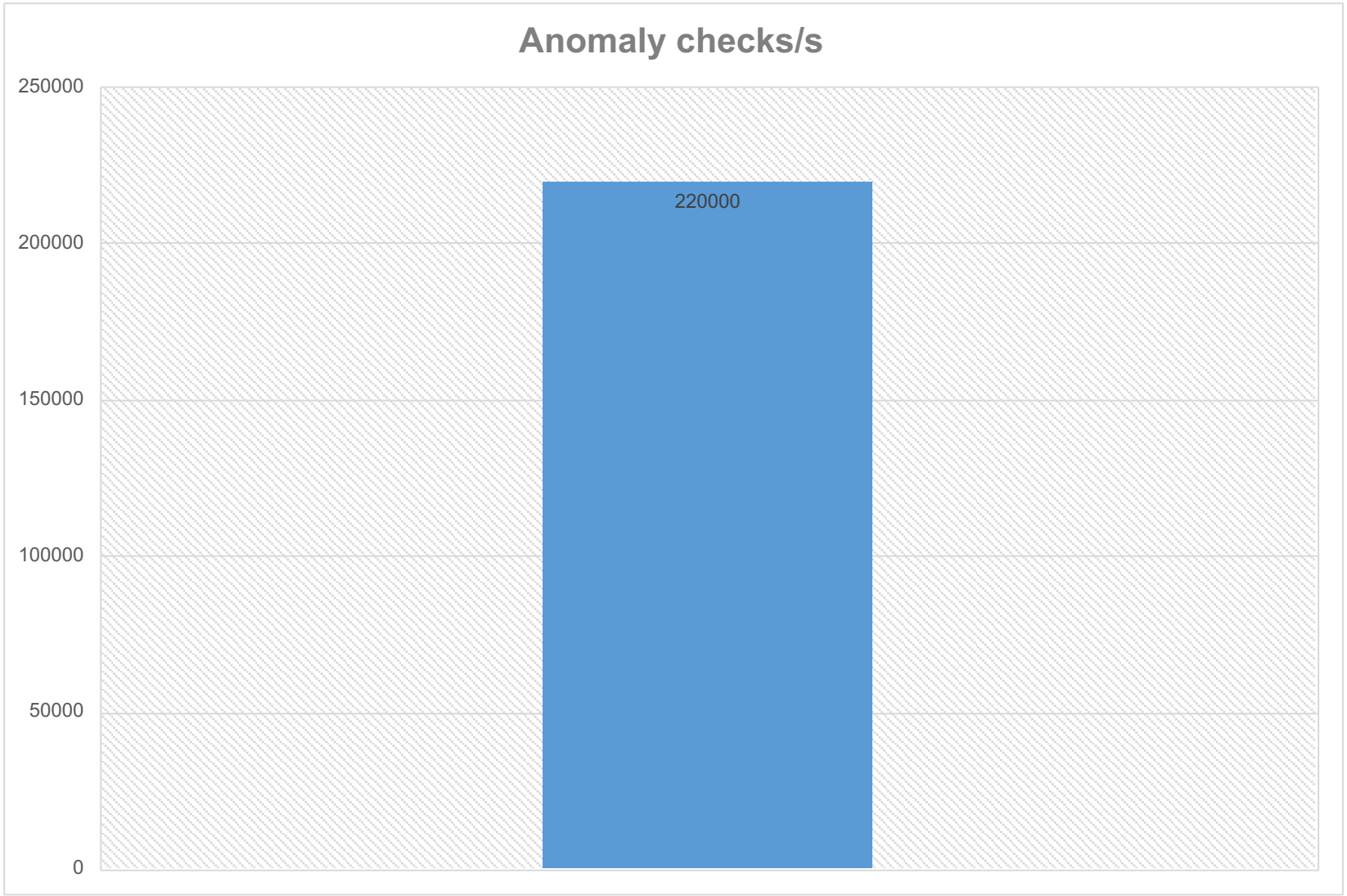
Apache Kafka,

Kubernetes & AWS

Is this a lot?

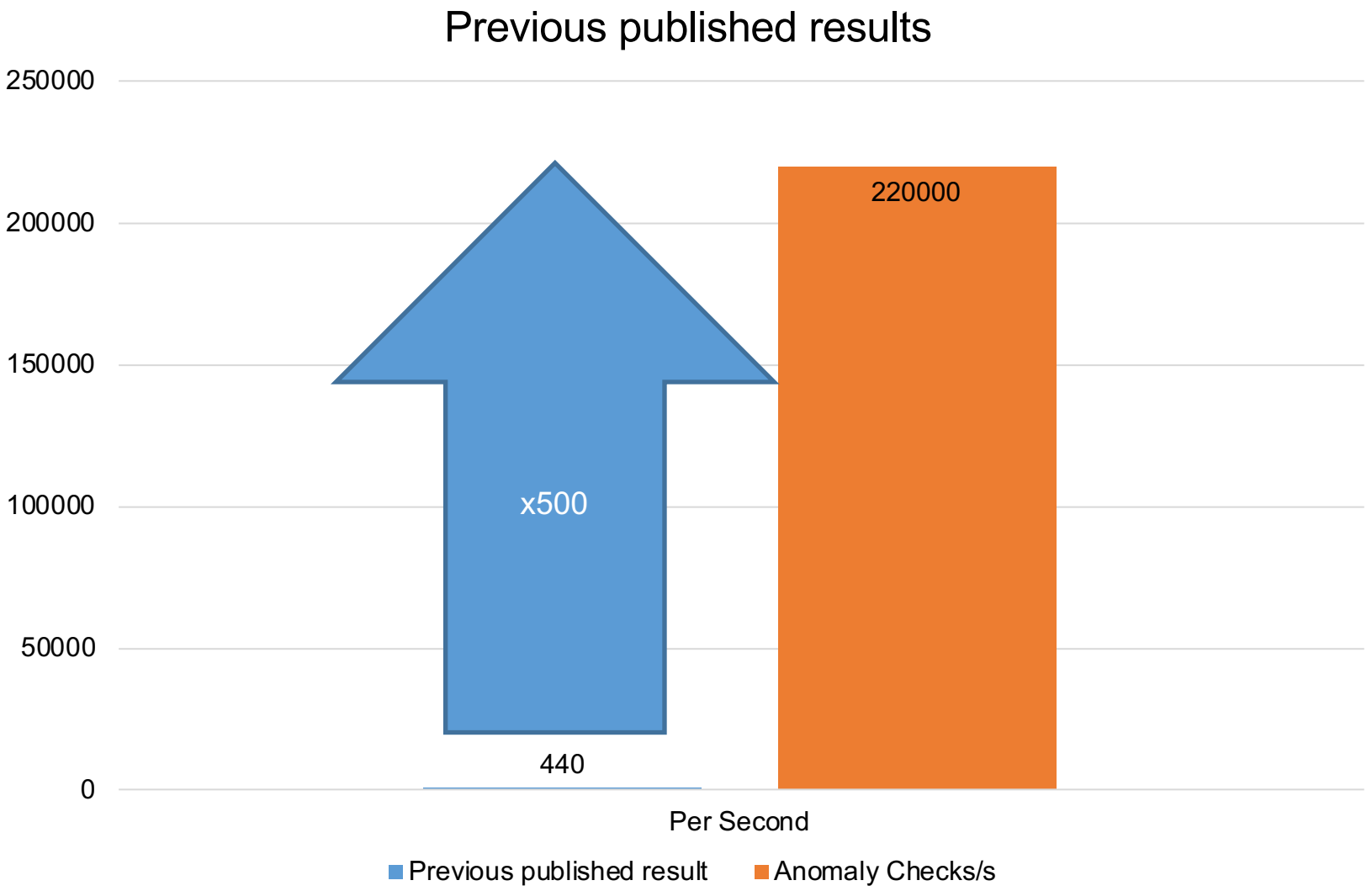
Headline Numbers Per Second

- 220,000 Anomaly checks Per Second



Headline Numbers Per Second

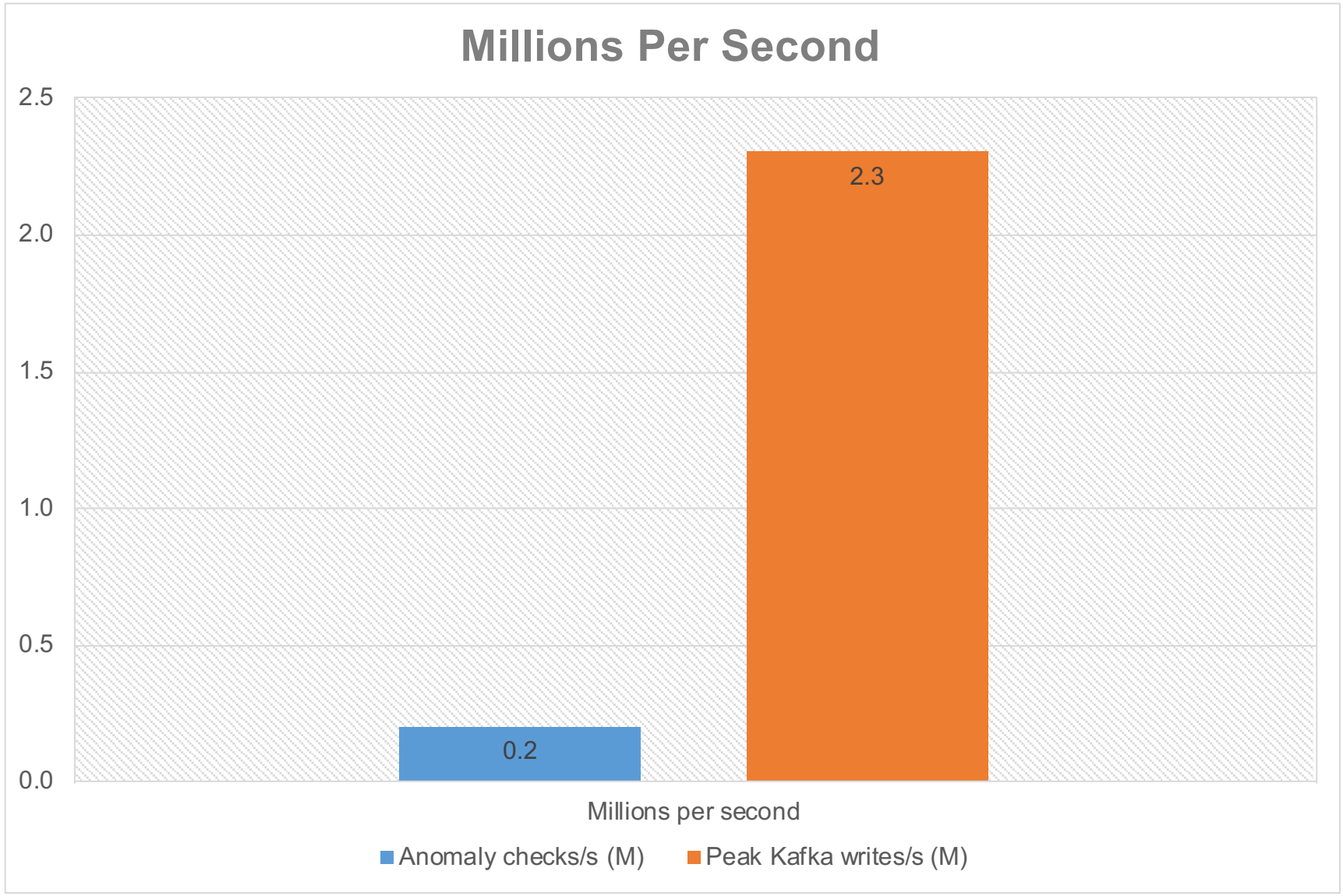
- 500x better than previously published results for similar system
- 2018, Kafka, Cassandra, Spark
- Bigger numbers?



instaclustr

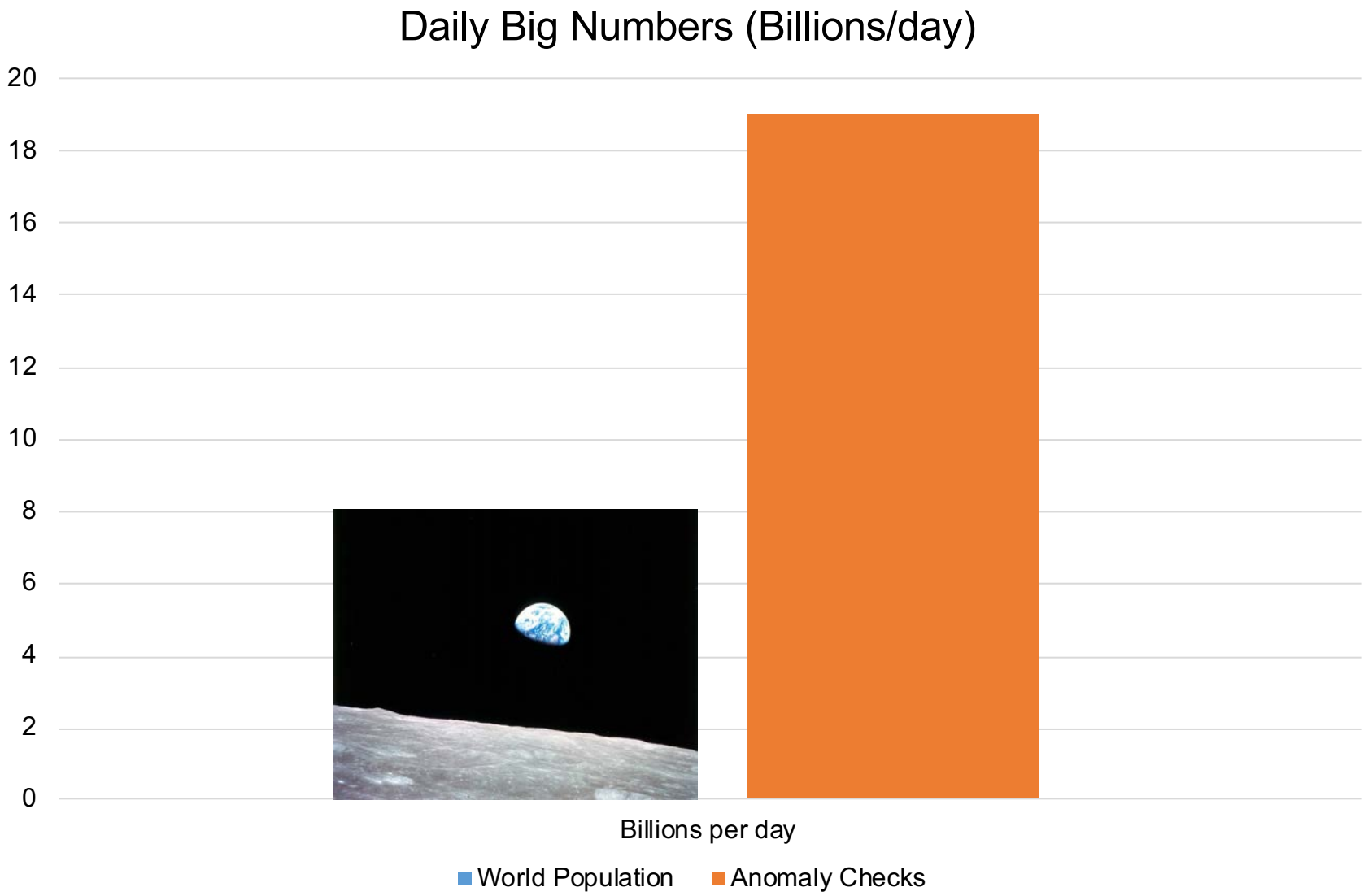
Headline Numbers Millions Per Second

- Peak 2.3 Million Kafka writes/s
- x10 rest of pipeline
- Kafka as a buffer, absorbs load spike



Headline Numbers Daily

- Planetary scale (population 7.7B)
- 19 Billion (1,000 Million) checks/day
- 2.5 events per person per day
- Had to stop somewhere, but no upper limit



2 Why?

Project Goals

Project Goals

Multiple (like Aussie Rules Football - AFL)

instaclustr

Project Goals

- Fast Data
- Real Time Streams processing
- < 1s RT

Project Goals

- Big Data
- Throughput and Size scale
- no upper limit
- big benchmark numbers

instaclustr

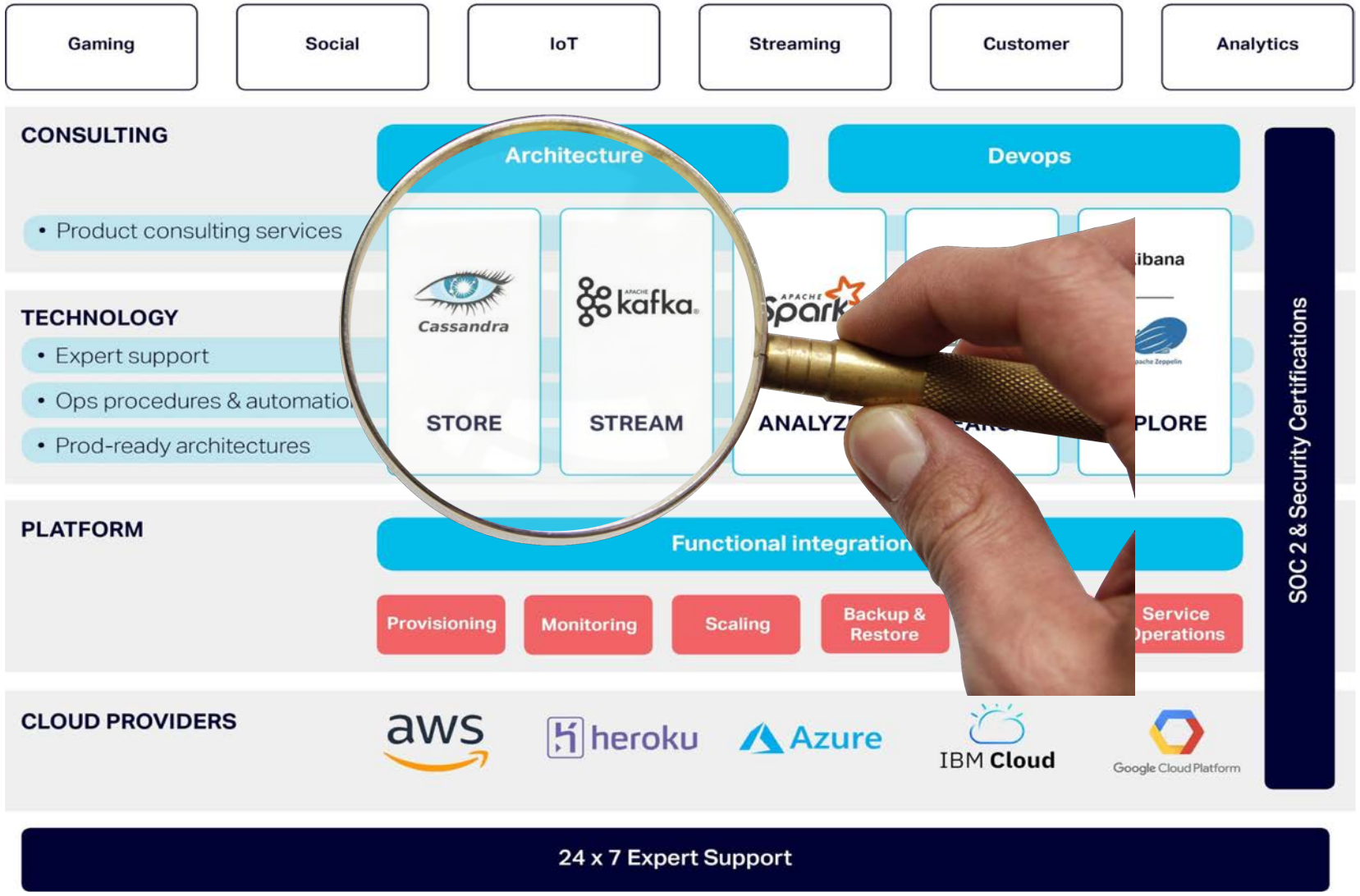
instaclustr

Cost Effective

- Incrementally scalable
- Only pay for what you use
- High benefit/cost ratio
- 1/2 car “Malcom” movie, 1986

Apache Kafka and Cassandra

- Technology - Kafka+Cassandra use case
- Platform - Instaclustr's Managed Platform
- Features - Provisioning, monitoring, scaling, and more

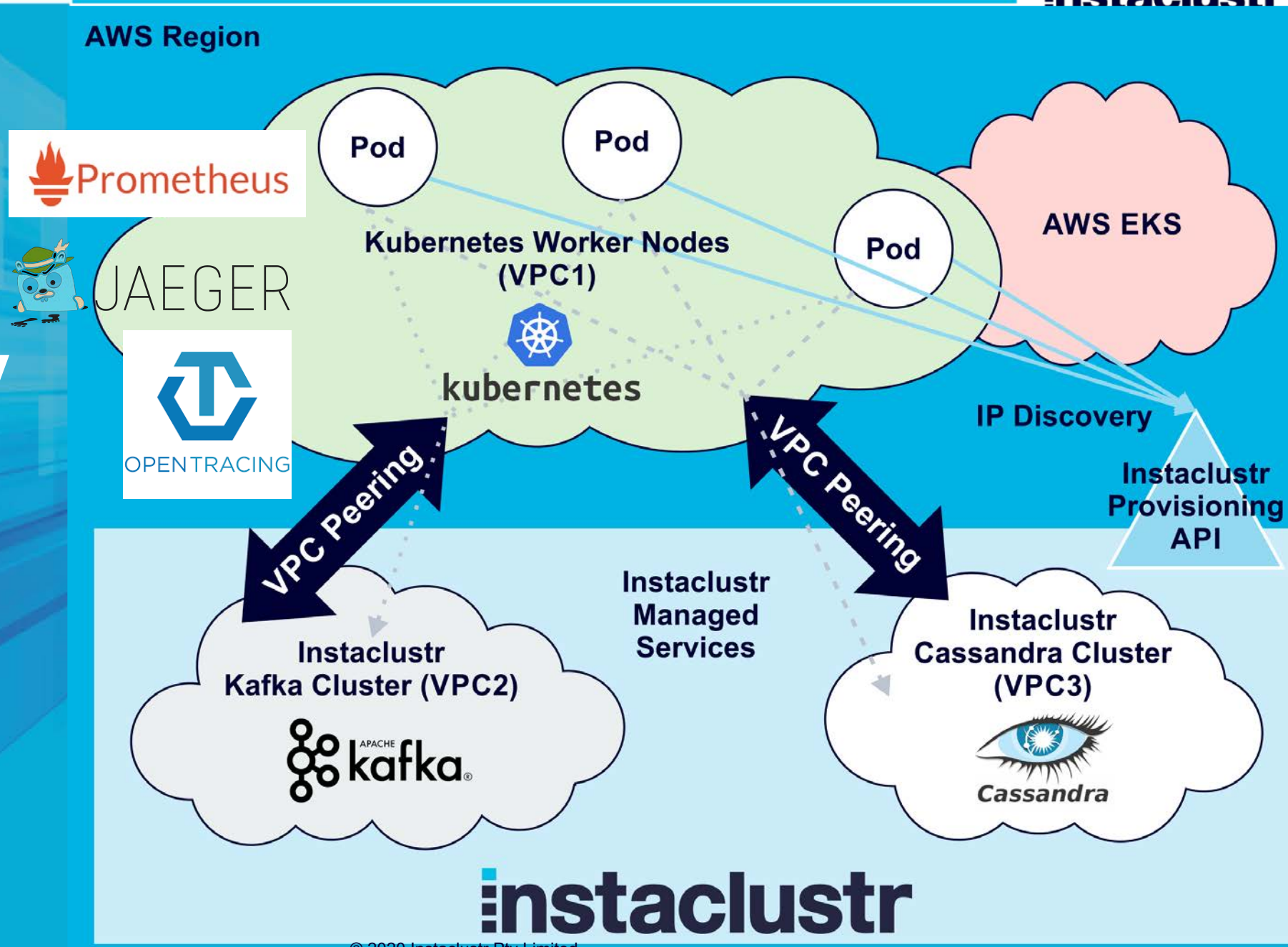


Kafka as a Buffer

- Cost effective for short load spikes
- E.g. Influx of unexpected festival goers
- Prevent overloading of rest of pipeline
- All events (eventually) processed

Application Automation and Observability

- Complementary technologies:
- Kubernetes (automation)
- Prometheus (monitoring)
- OpenTracing+
- Jaeger (tracing)



3 What?

Does it do?

instaclustr

What does it do? Anomaly Detection Use Case

Spot unusual events

instaclustr

“Man on Moon” headlines

- 400,000 people got them there
- JoAnn Morgan, Saturn 5 monitoring engineer
- Only woman in the control room for Apollo 11

instaclustr

Anomaly Detection Goals

Spot the difference
At speed and scale

instaclustr

Spot the difference at *speed*

- 1 second maximum
- Streams processing not batch

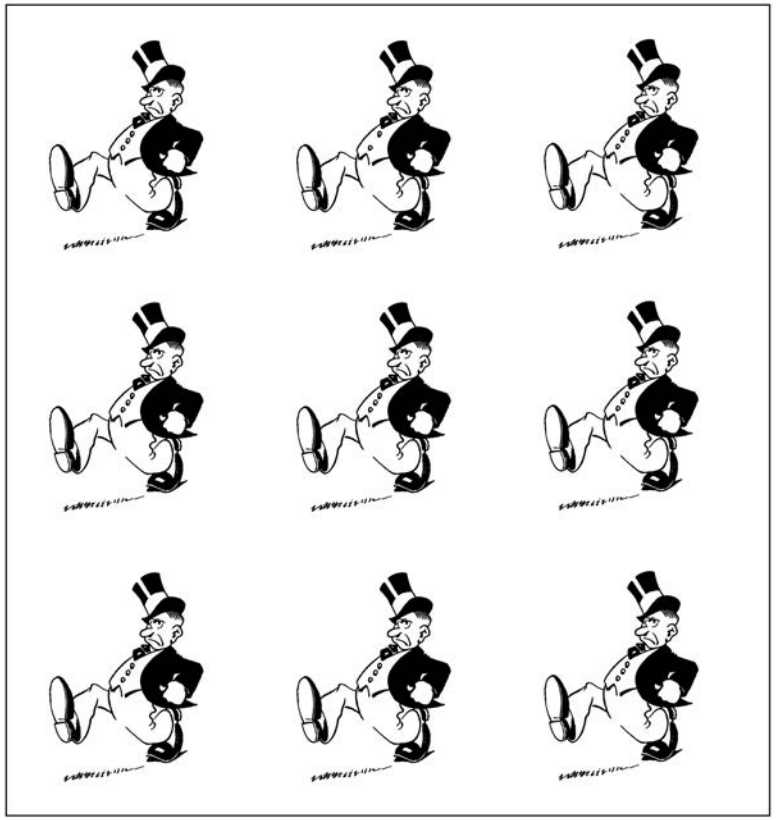
Find the odd one out

Spot the difference at *scale*

- Keys and Concurrency
- Multiple keys
- Need Big Data database
- For Storage and Processing capacity

SPOT THE ONE THAT'S DIFFERENT

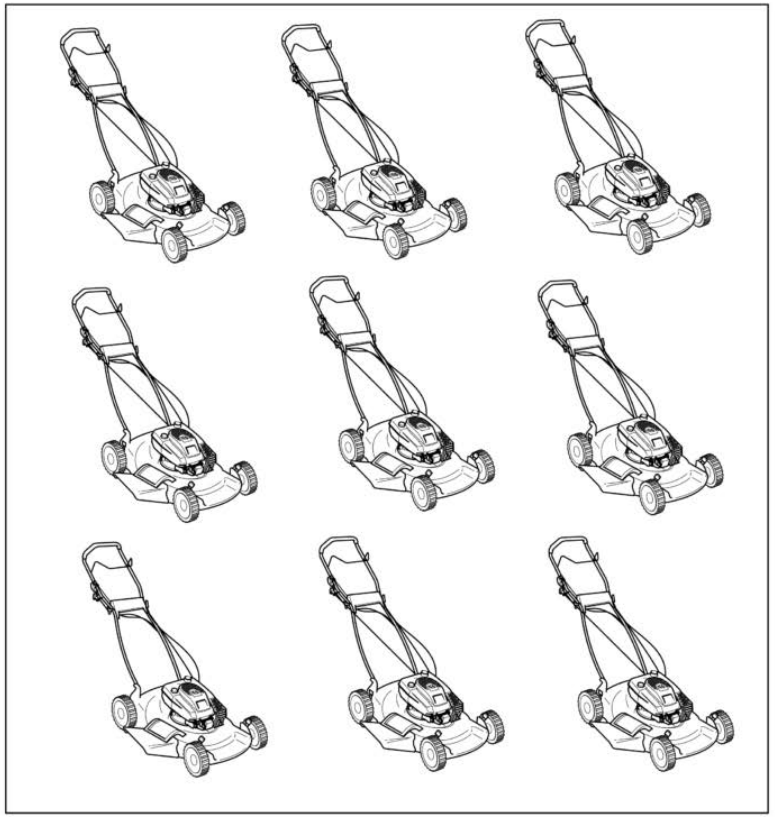
These men look like they are in a hurry. Can you spot the one that is different from the others?



www.easyfunpuzzles.com

SPOT THE ONE THAT'S DIFFERENT

Time to mow the lawn! Can you spot the one lawn mower that is different from the others?



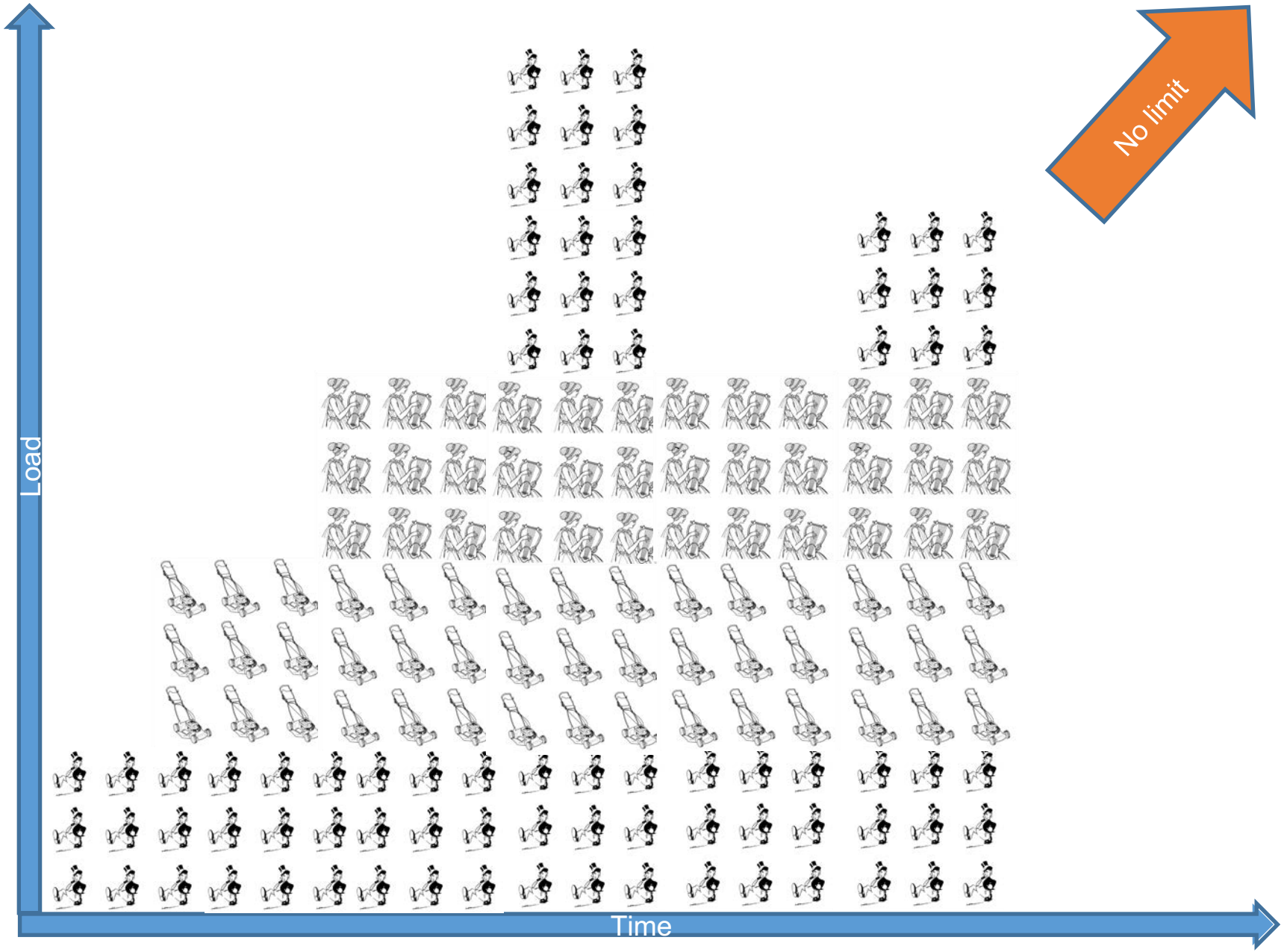
www.easyfunpuzzles.com

Original lawn mower image from OCAL7Clikr.com accessed at http://www.clker.com/cliparts/2/9/e/7/11954368581269955423.lic.lac_Push_Mower.svg

instaclustr

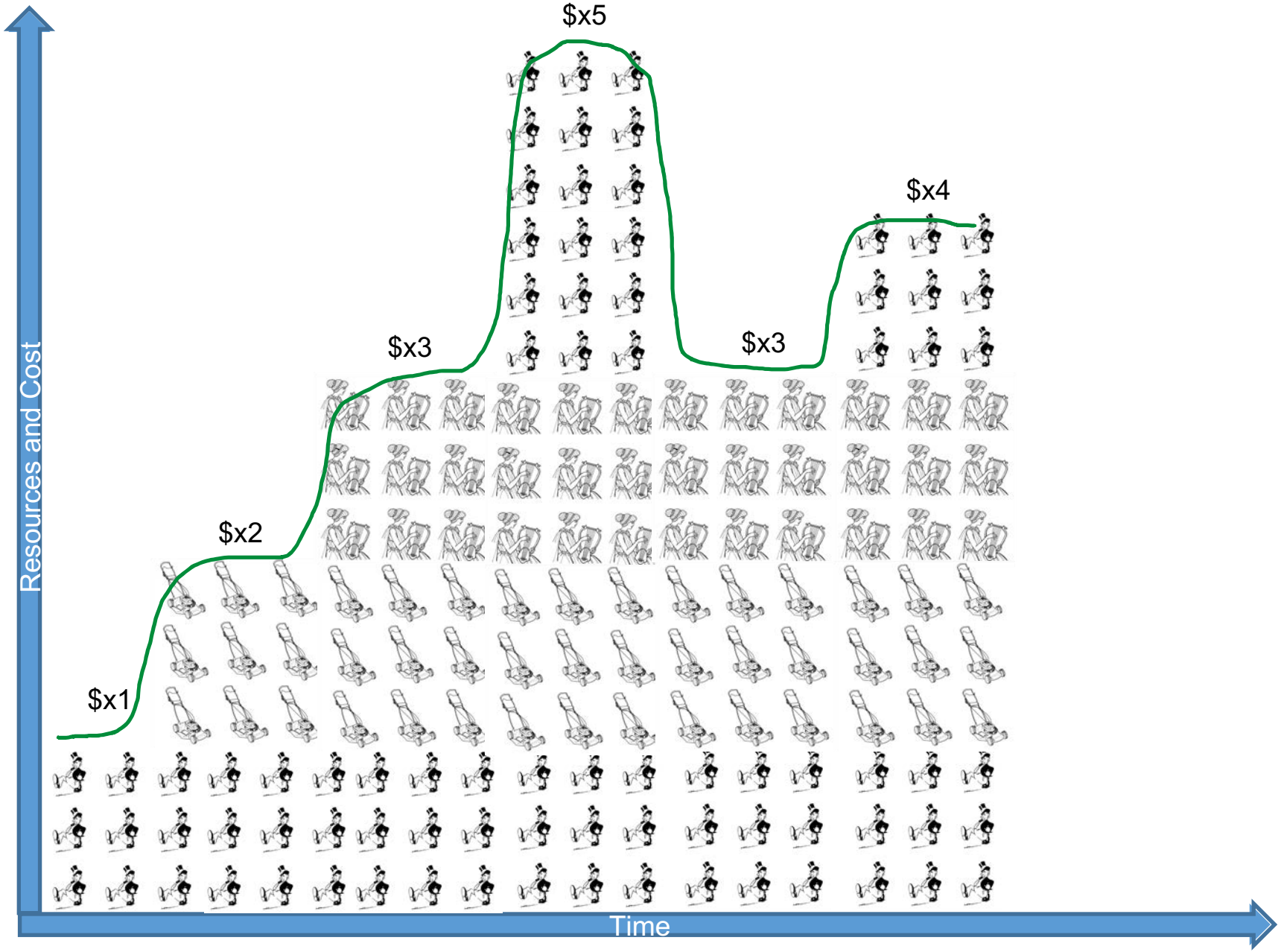
Scalability

- Massive load (data velocity)
- Increasing load
- No upper bound
- Load spikes



Affordability

- Linear resource scalability
- Elastic, on-demand
- Incremental resources and cost with changing load



instaclustr

Anomaly Detection Use Cases

Many and varied

Infrastructure monitoring

instaclustr

Anomaly Detection Use Cases

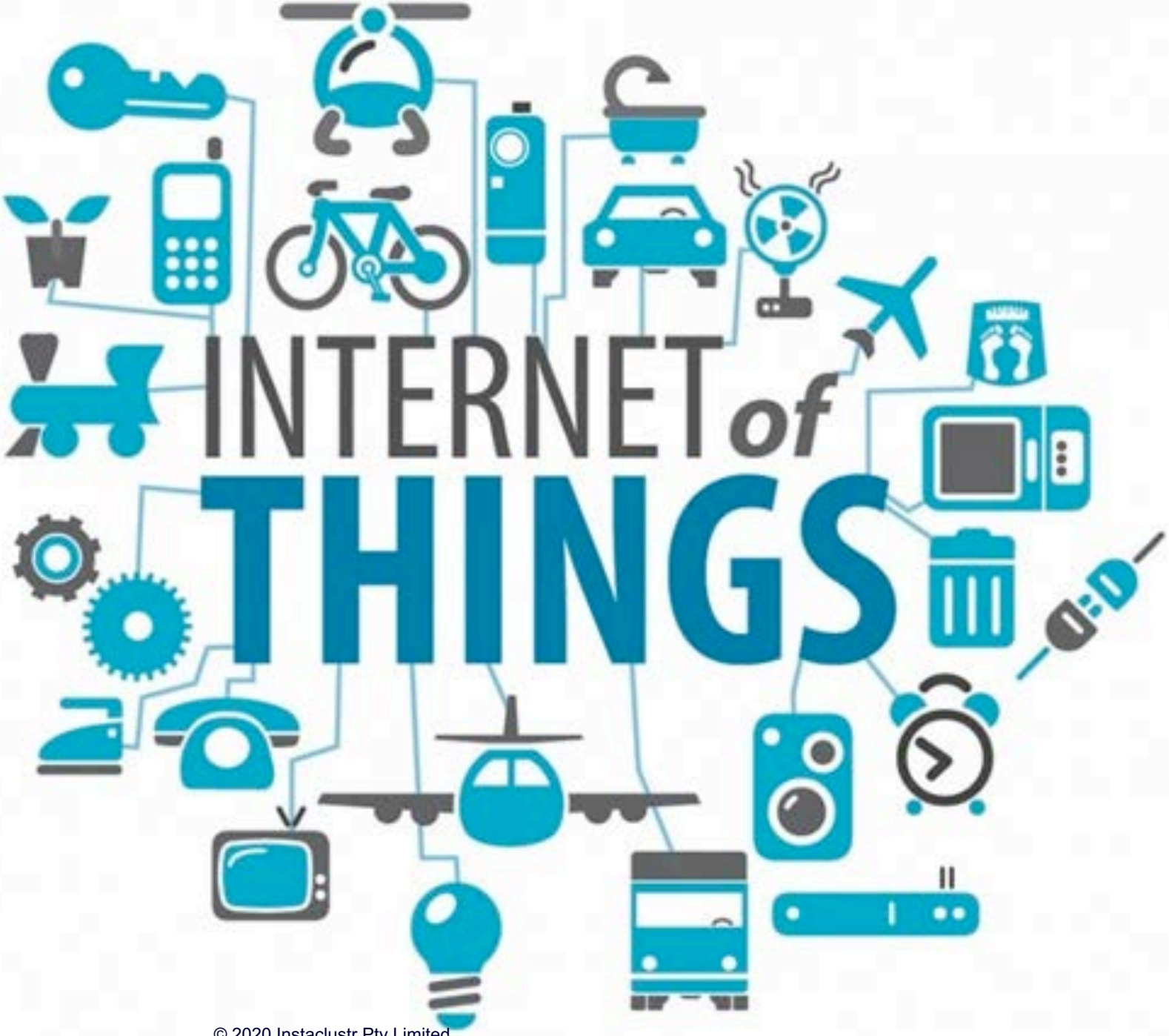
Many and varied

Application Monitoring

Anomaly Detection Use Cases

Many and varied

IoT



instaclustr

Anomaly Detection Use Cases

Many and varied

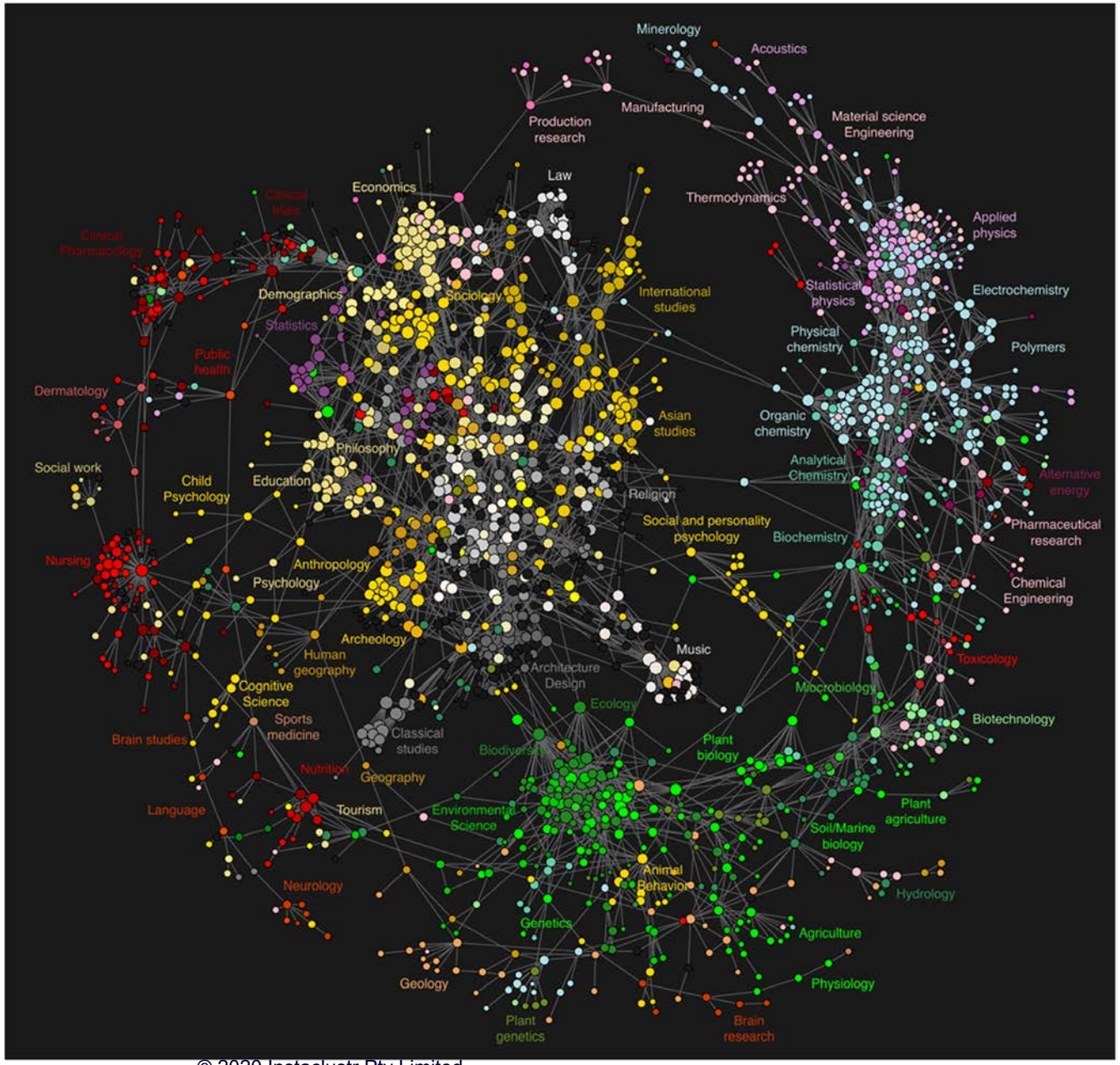
Finance fraud detection

instaclustr

Anomaly Detection Use Cases

Many and varied

Clickstream analytics



instaclustr

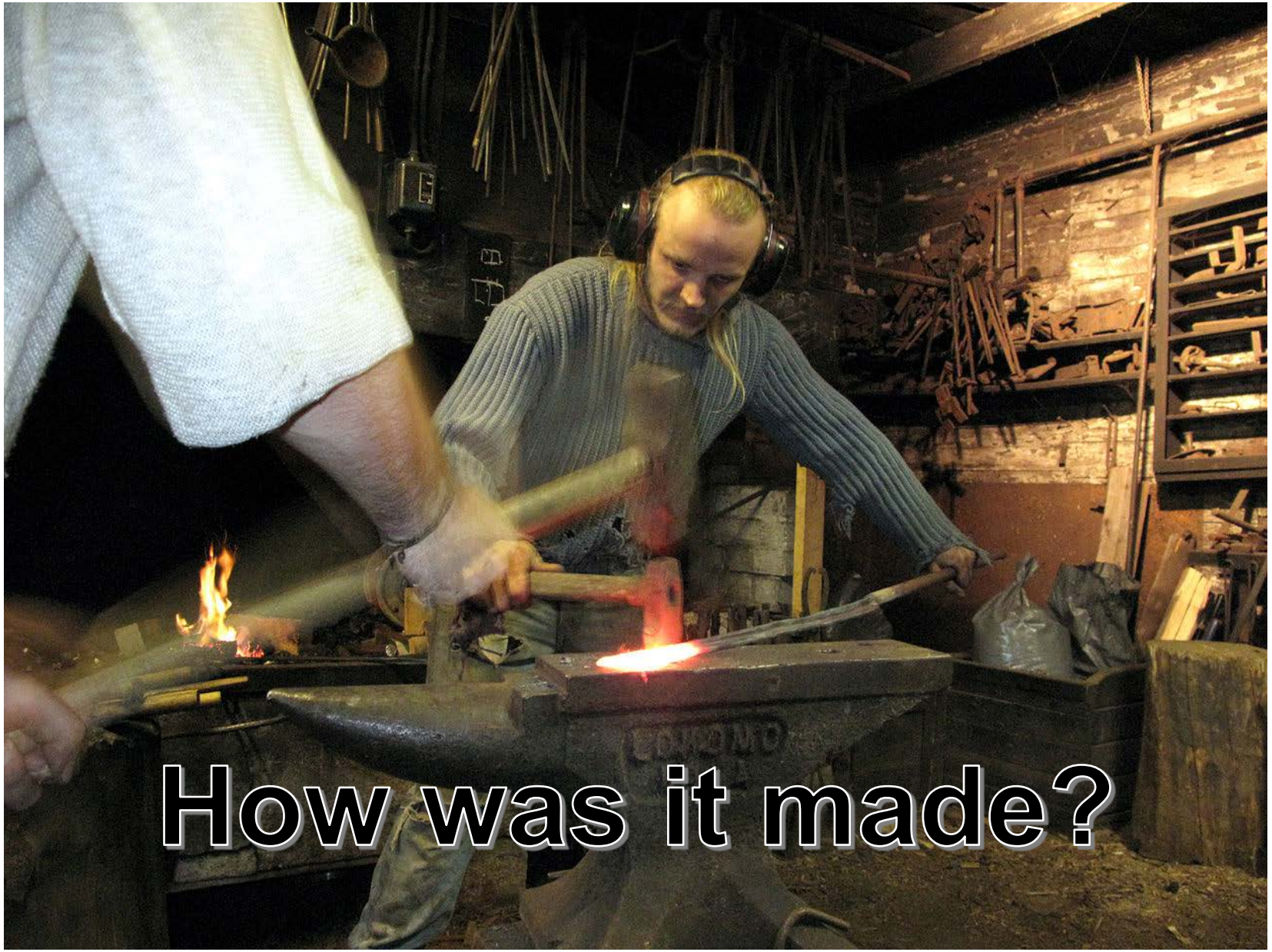
Anomaly Detection Use Cases

Many and varied

Drone deliveries

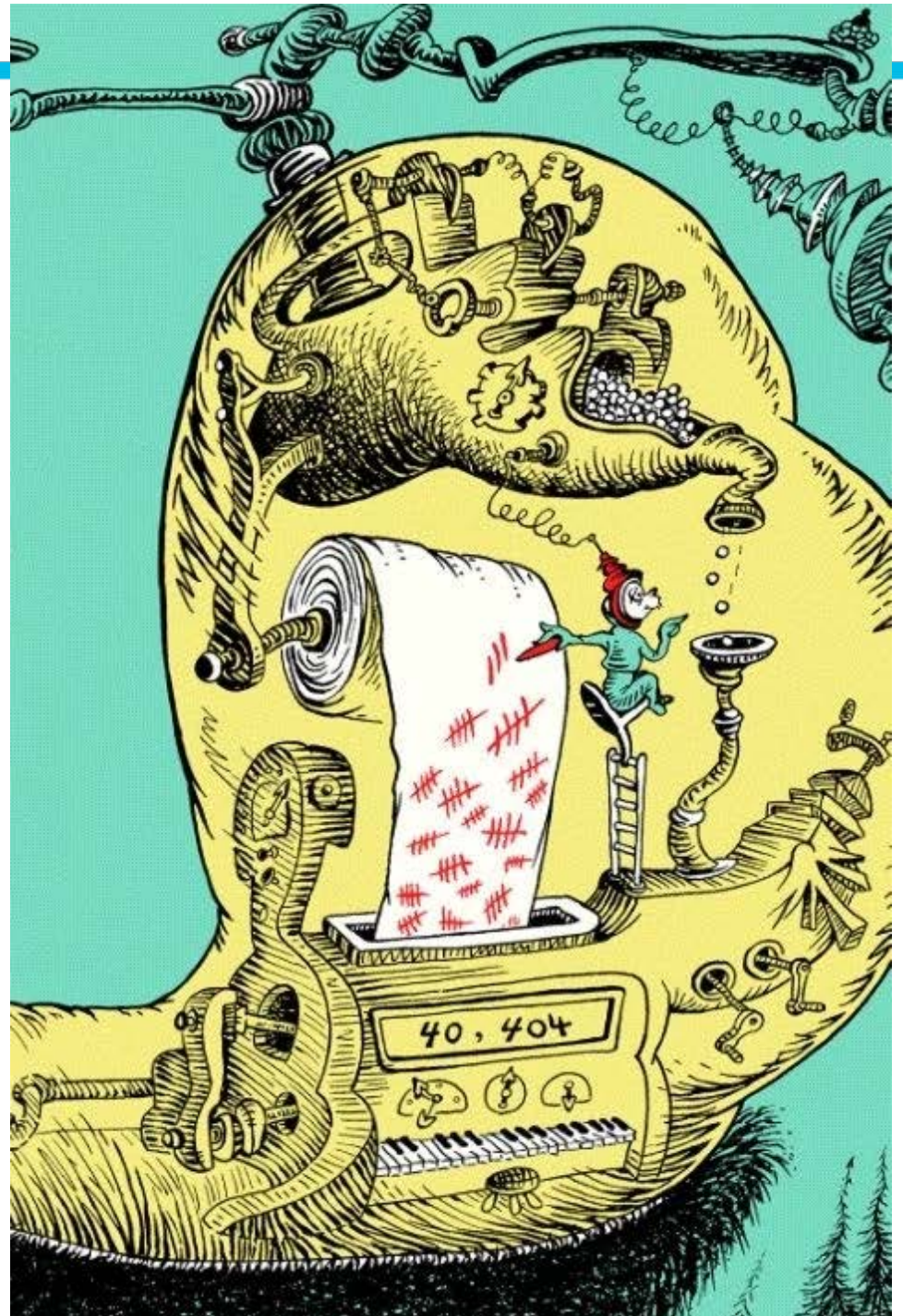
4 How does it work?

- Anomaly Detection
- Architecture
- Technologies



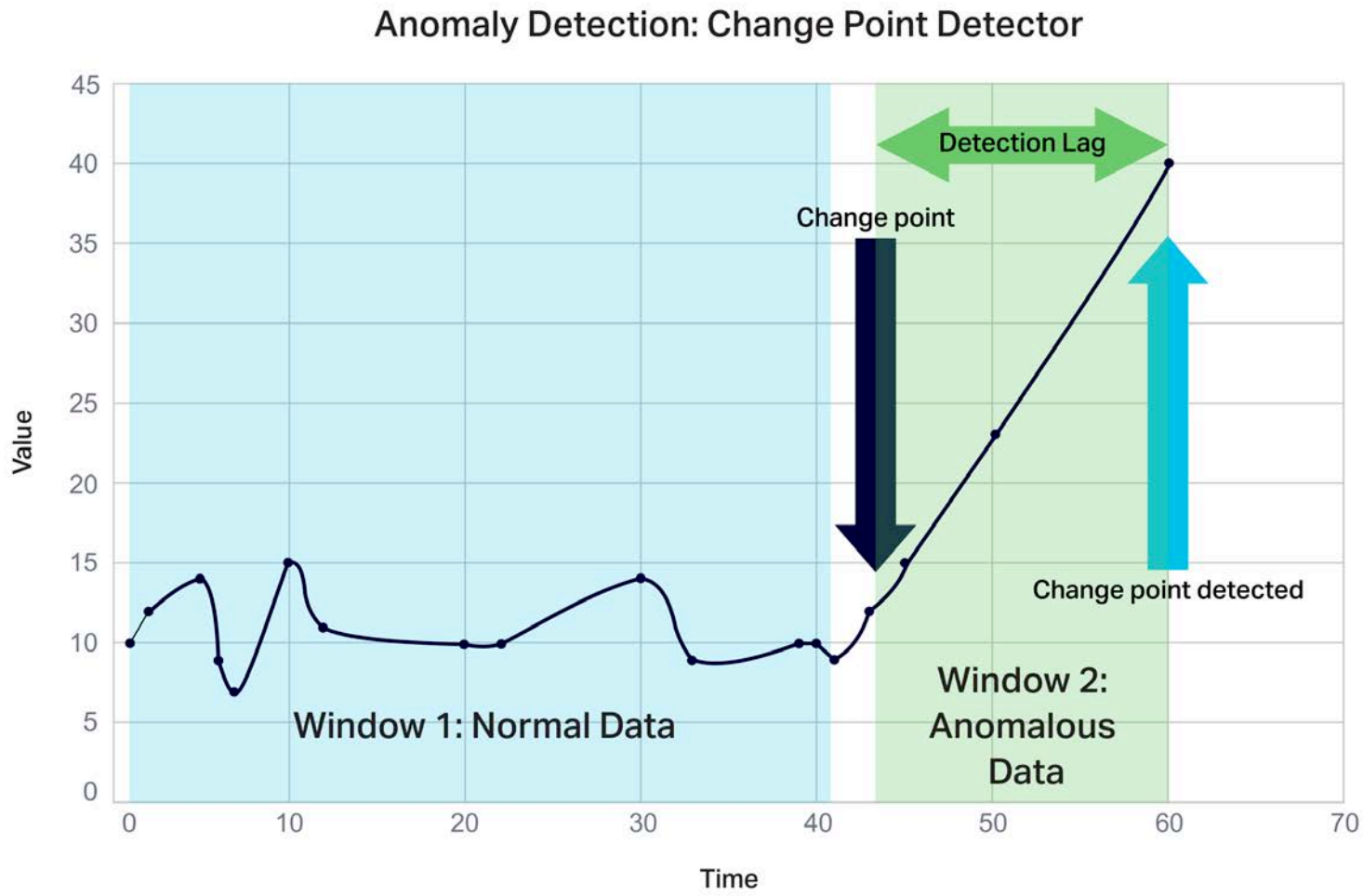
Is this our machine?

- The Audio-Telly-o-Tally-o Count
- Streams processing machine for counting sleepers
- We've advanced from this 1960's technology



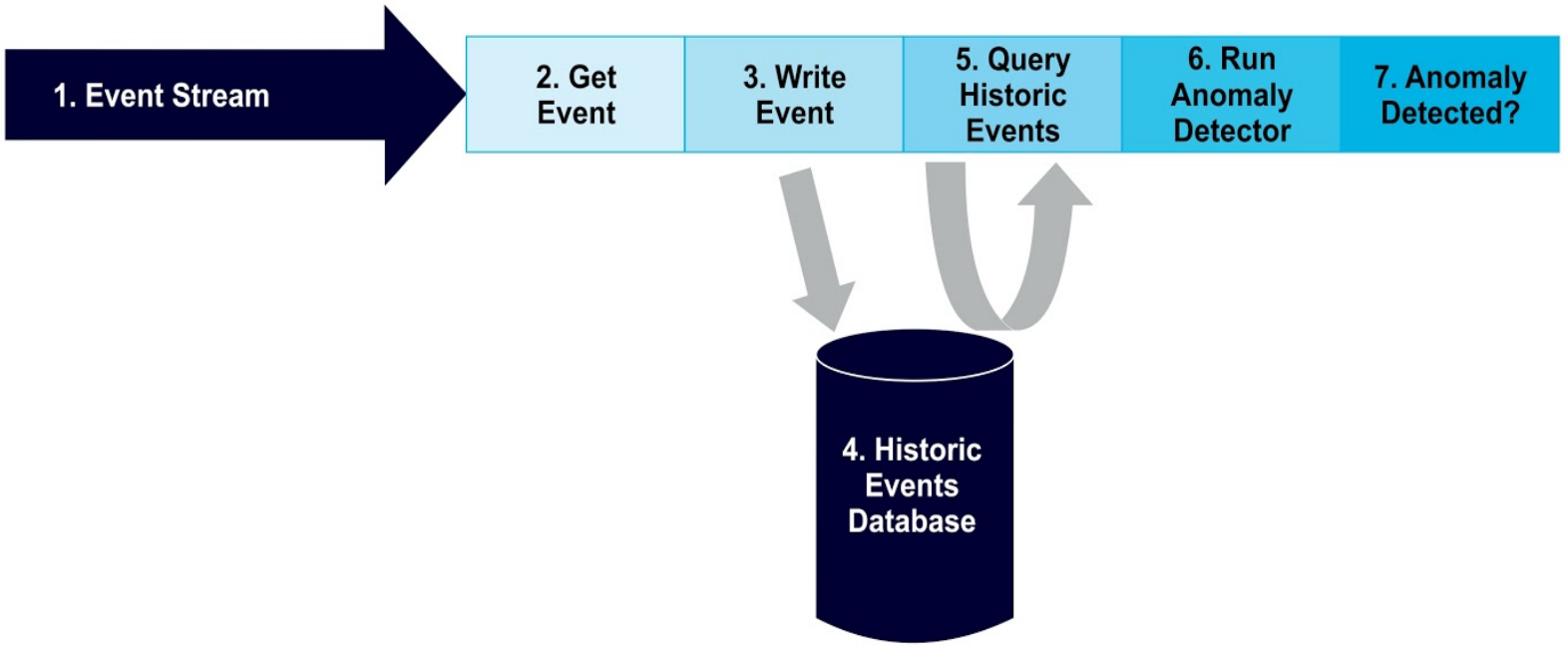
How does it work?

- CUSUM (Cumulative Sum Control Chart)
- Statistical analysis of historical data



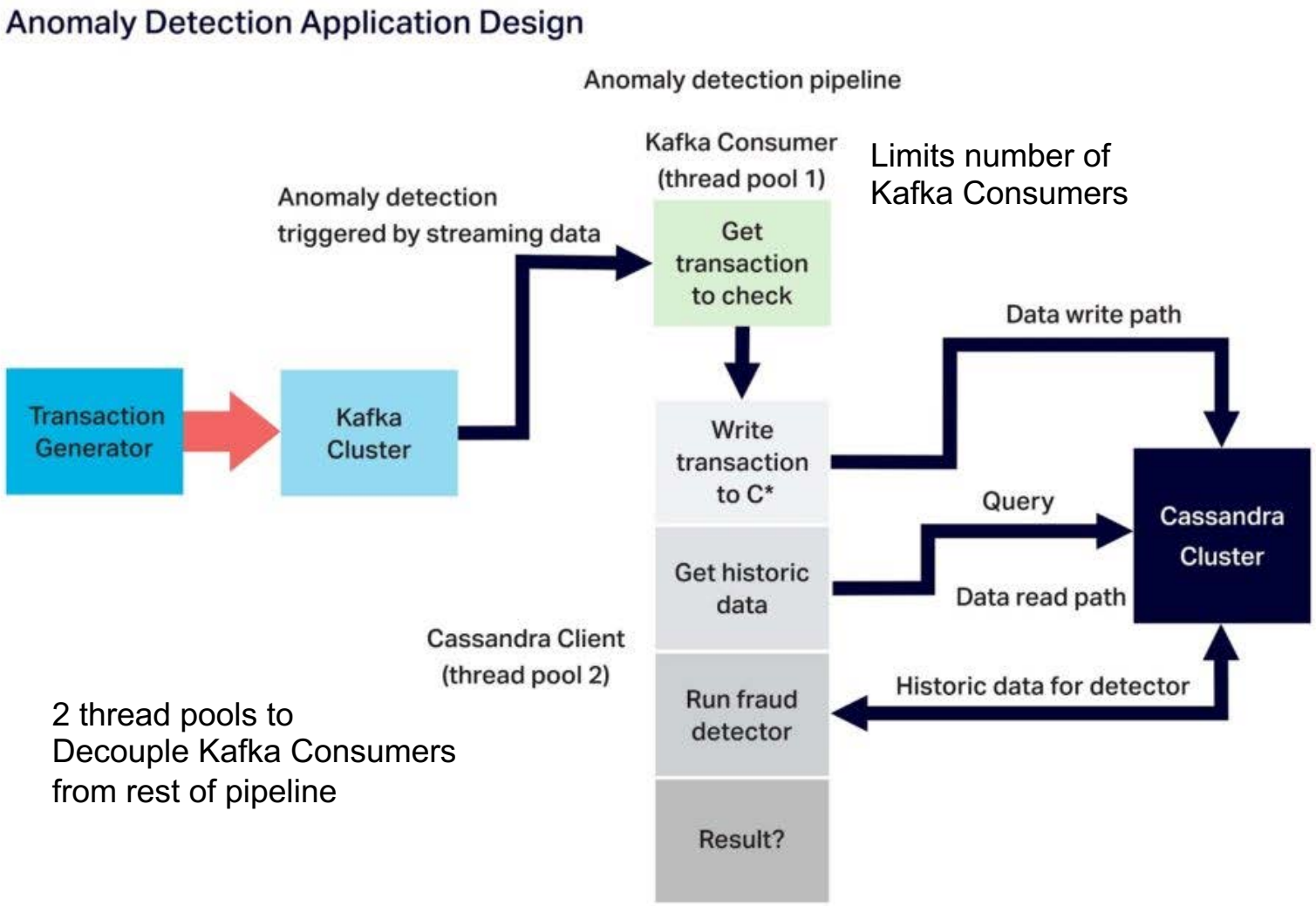
Logical steps

(1) Events arrive in a stream
(2) Get the next event from the stream
(3) Write the event to the database (4)
(5) Query the historic data from the database (4)
(6) If there are sufficient observations, run the anomaly detector
(7) Was a potential anomaly detected? Take appropriate action.



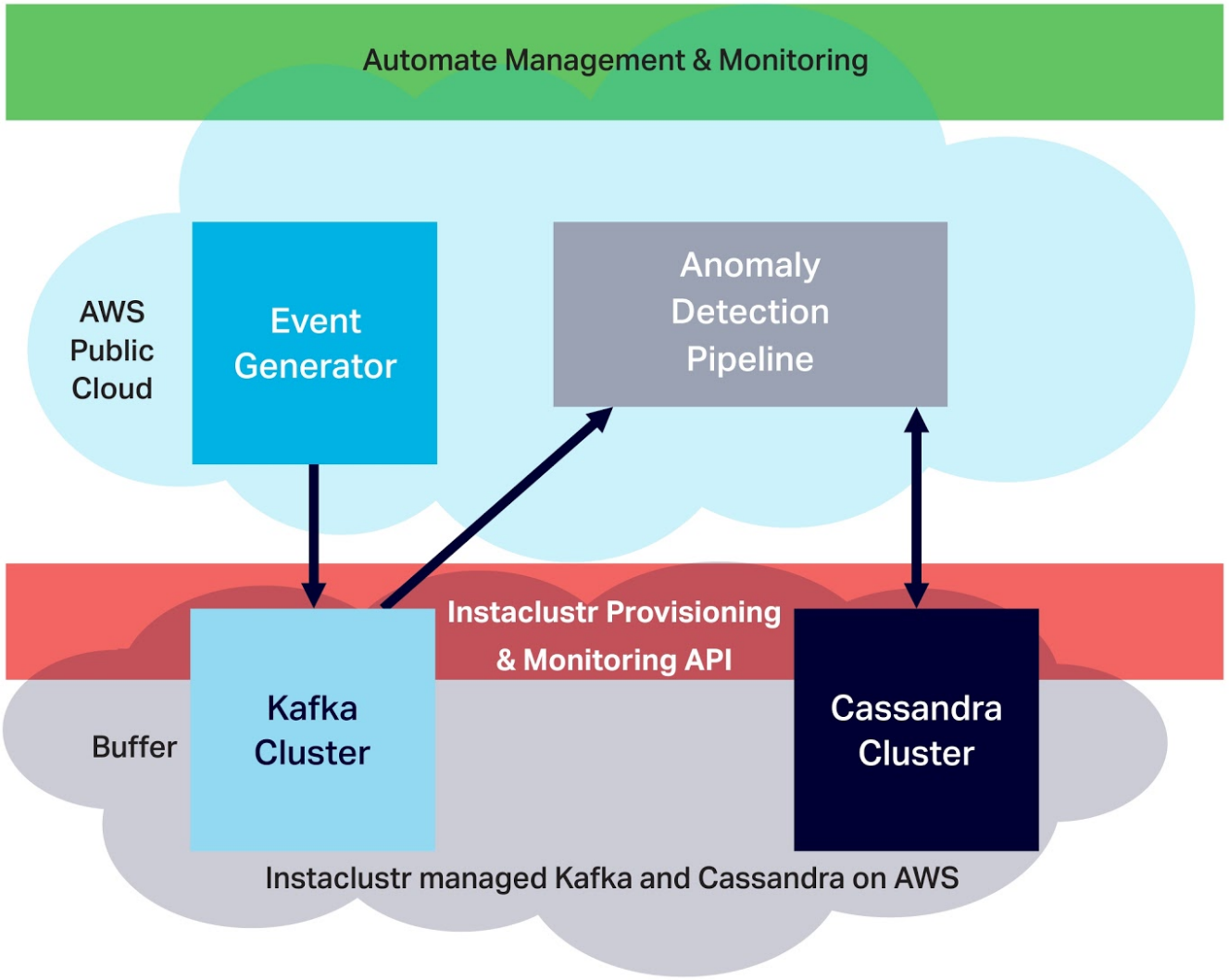
Pipeline Design

- Design, showing interaction with Kafka and Cassandra Clusters
- Load generator, detector pipeline
- 2 thread pools
- To constrain the number Kafka consumers (→ Kafka partitions)



Cloud Deployment Context

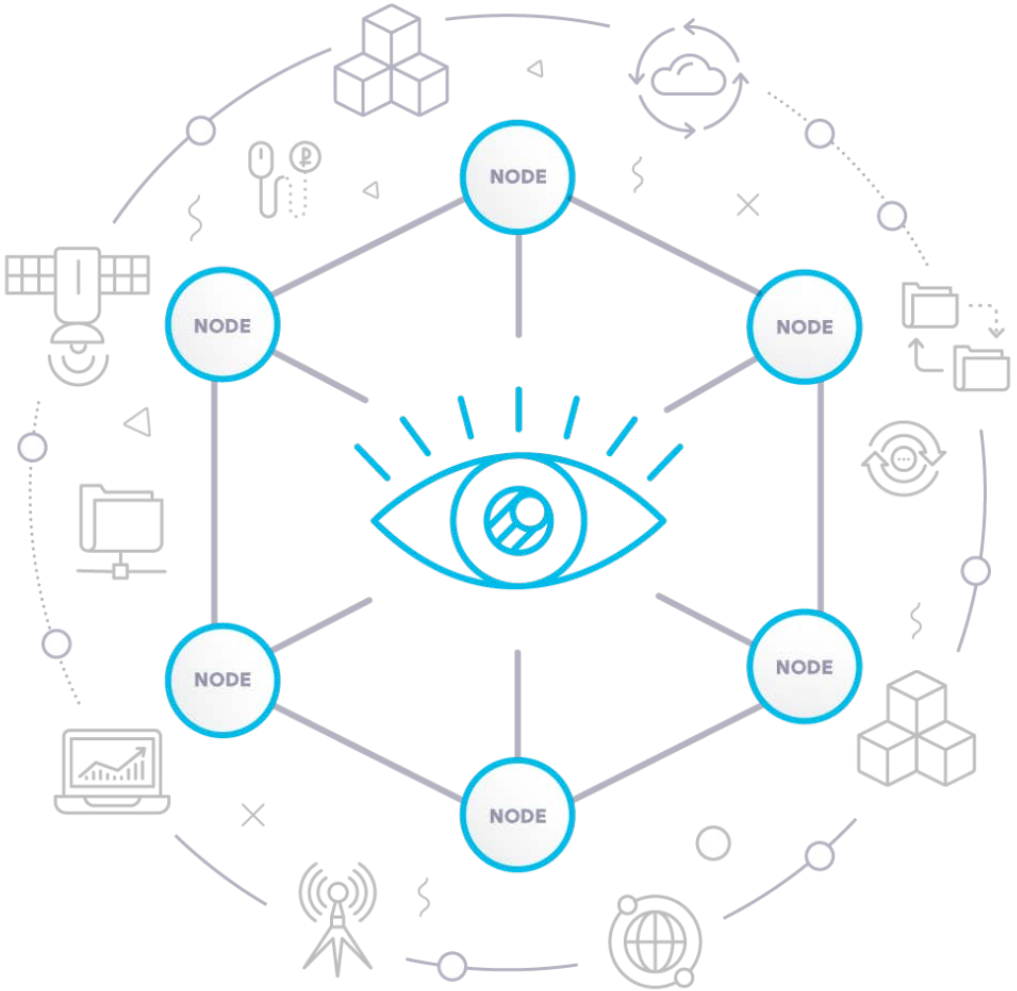
- Kafka and Cassandra clusters managed by Instaclustr
- Application in AWS



instaclustr

Cassandra

- Open Source
- NoSQL Database
- Masterless ring architecture & partitioned data for
- Linear scalability
- High availability
- Fast writes
- Powerful queries with indexes



Instaclustr Managed Apache Cassandra

Benefits

- Optimised for low latency/high throughput
- Automated Provisioning, Monitoring, Management
- SOC2 certified
- Multiple cloud providers
- 24/7 Technical support
- Automated Health Checks
- Dynamic scaling
- Zero downtime migrations
- New! Certified Apache Cassandra
  - Key highlights of the Certification Report include:
    - Performance testing (latency and throughput) comparing the current version to previous versions
    - 24-hour soak testing (including repairs and replaces)
    - Testing against popular drivers

What is Kafka?

Distributed streams processing

instaclustr

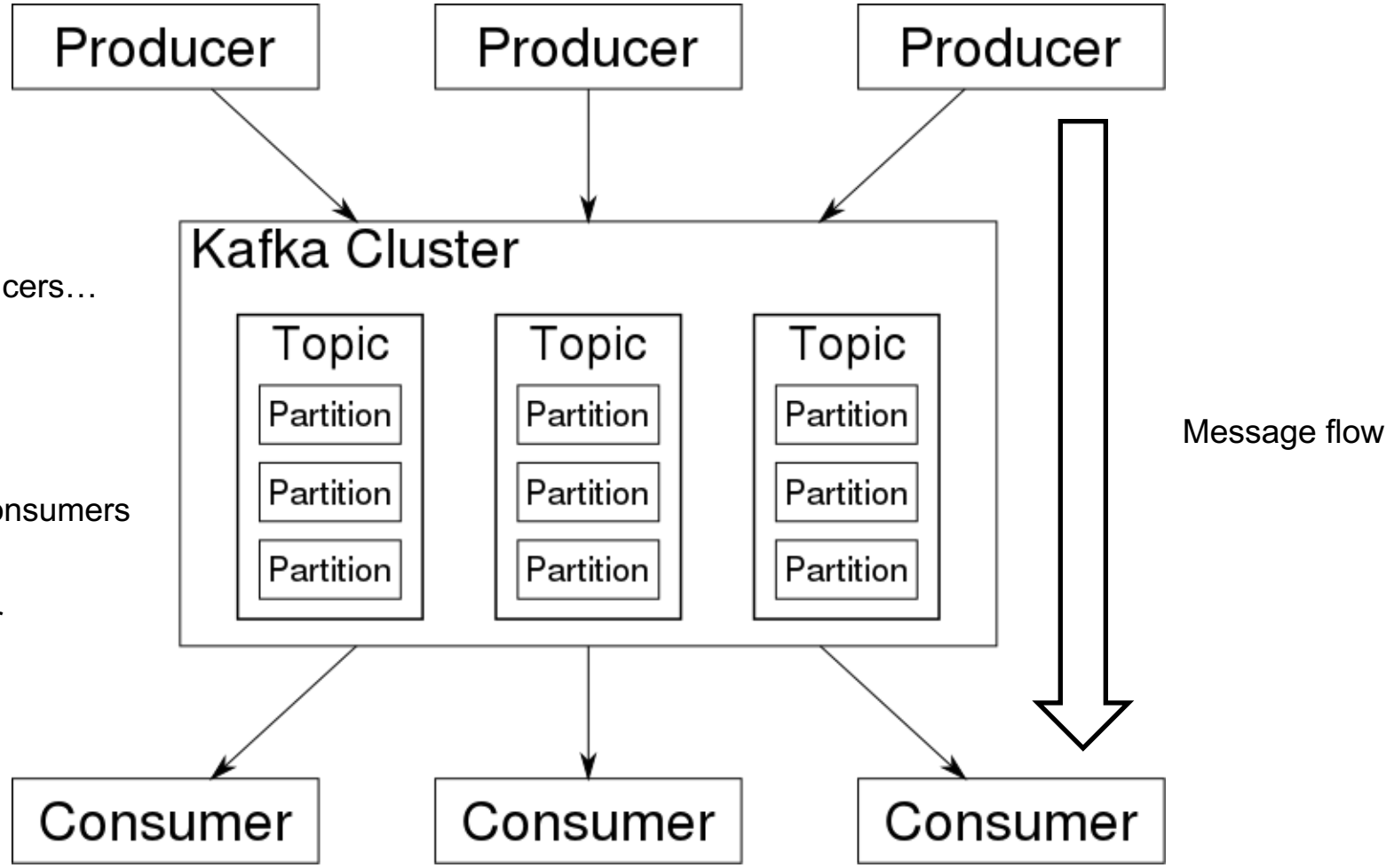
APACHE kafka®

1 Distributed Producers…

2 Send Messages

3 To Distributed Consumers

4 Via Kafka Cluster



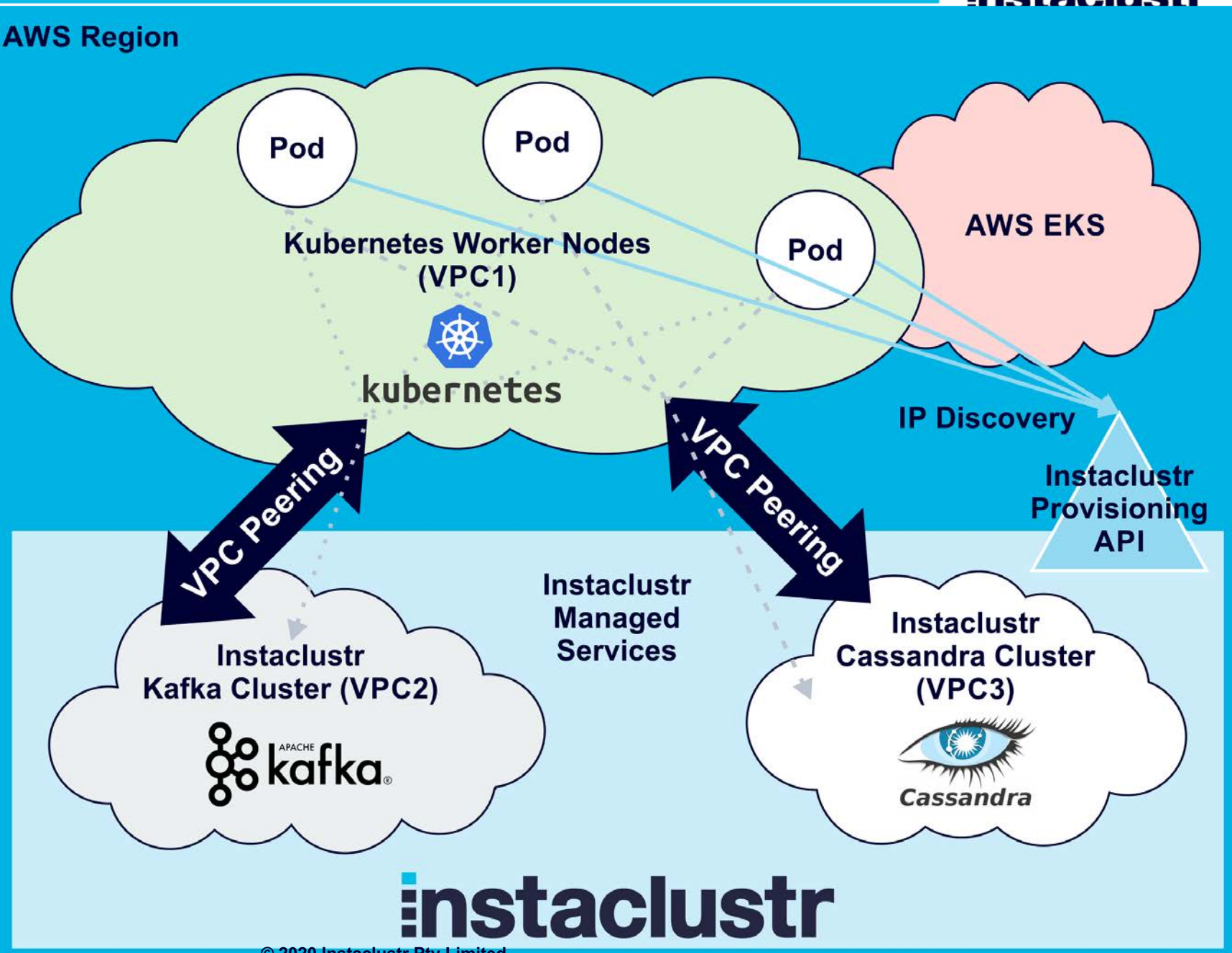
Kafka

Key Benefits

- Fast – high throughput and low latency
- Scalable – horizontally scalable, just add nodes and partitions
- Reliable – distributed and fault tolerant
- Zero data loss
- Open Source
- Heterogeneous data sources and sinks
- Available as an Instaclustr Managed service

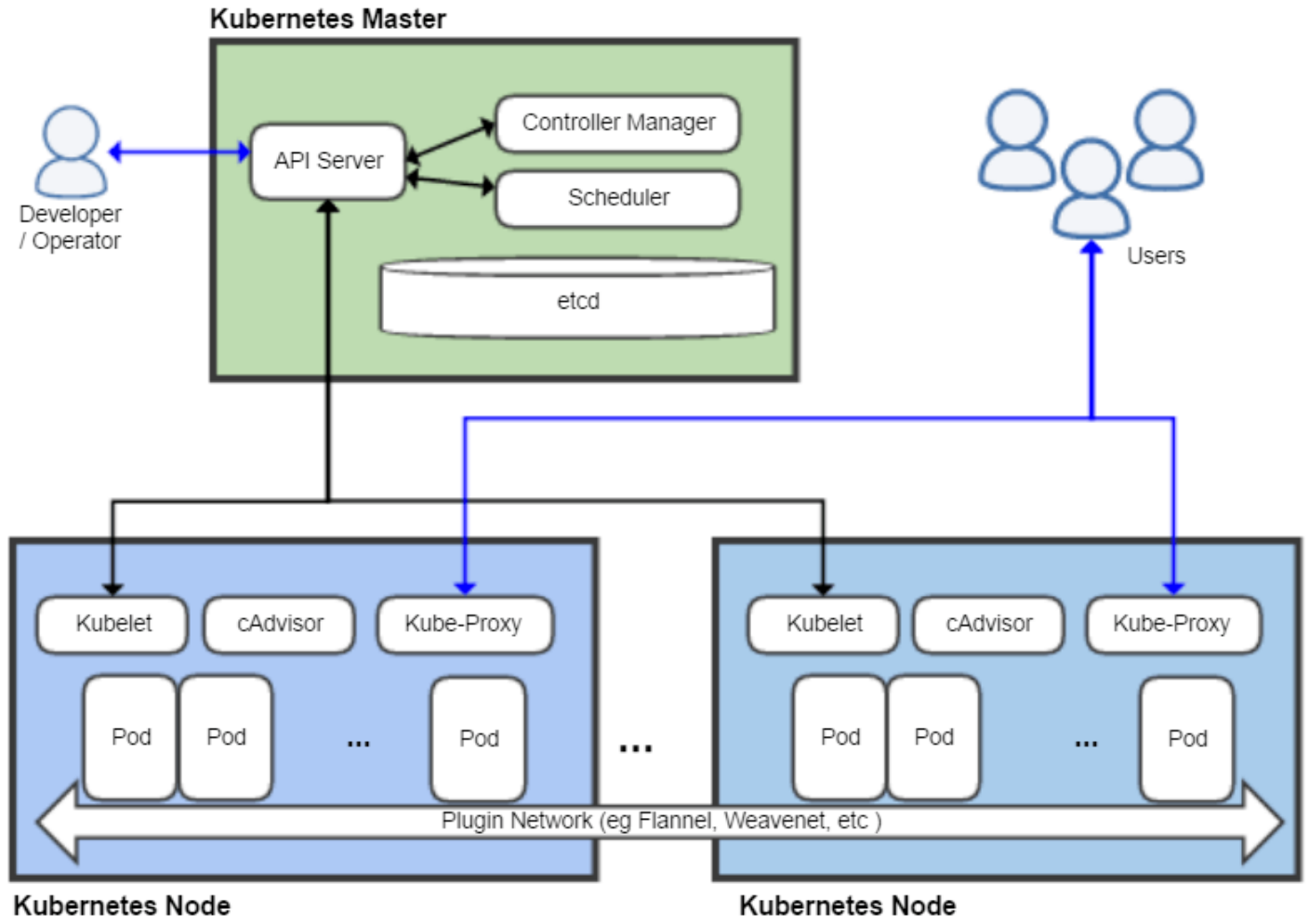
Application Automation with Kubernetes

- AWS EKS
- Kafka load generator and Anomaly Detection Pipeline deployed on worker nodes



Kubernetes

- An automation system for the management, scaling and deployment of containerized applications
- Master/worker Nodes architecture
- Pods are units of concurrency



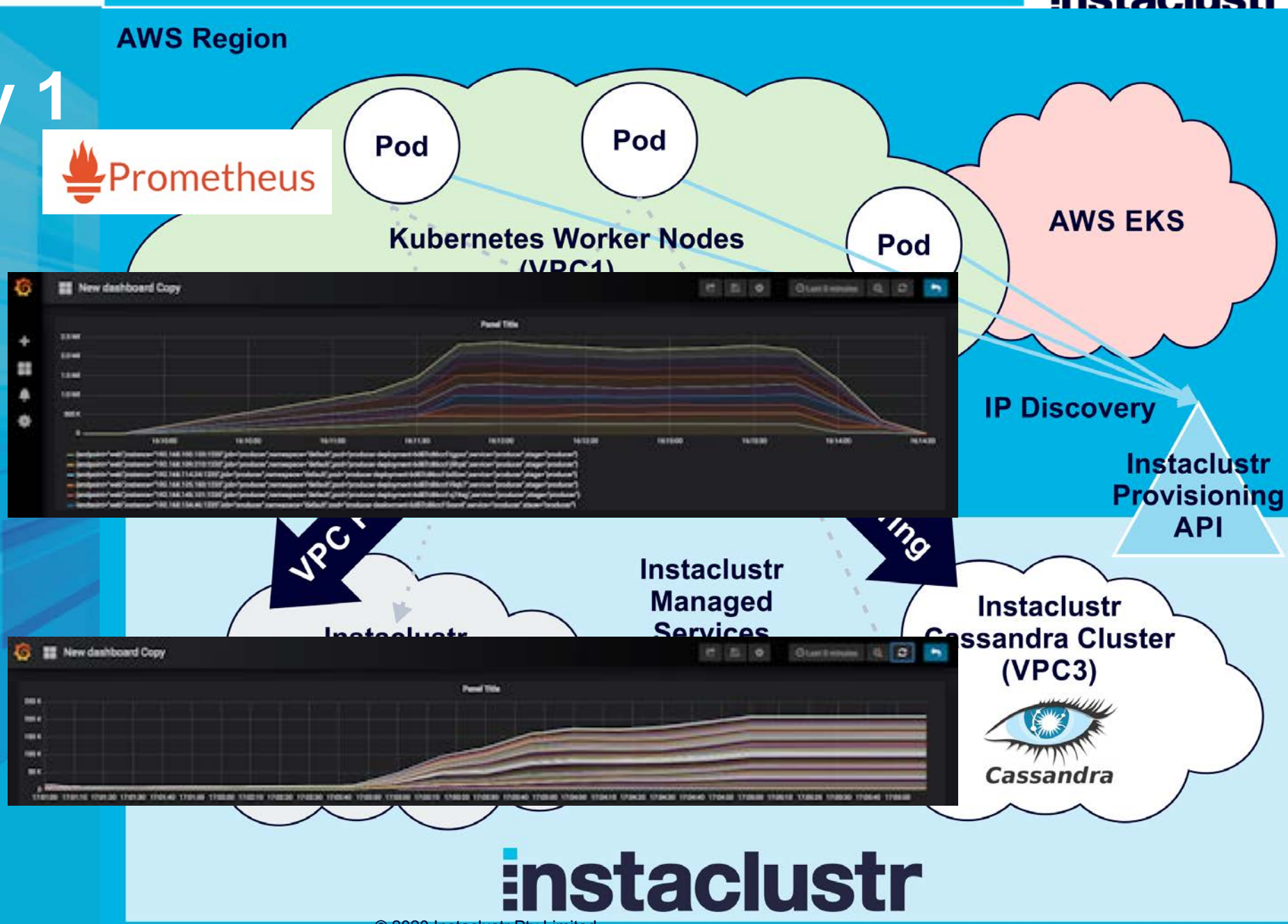
Kubernetes

Benefits

- Open Source
- Cloud provider and programming language agnostic
- Develop and test code locally, then deploy at scale
- Helps with resource management – deploy application to Kubernetes and it manages scaling up/down and keeping application alive
- More powerful frameworks built on Kubernetes APIs are becoming available

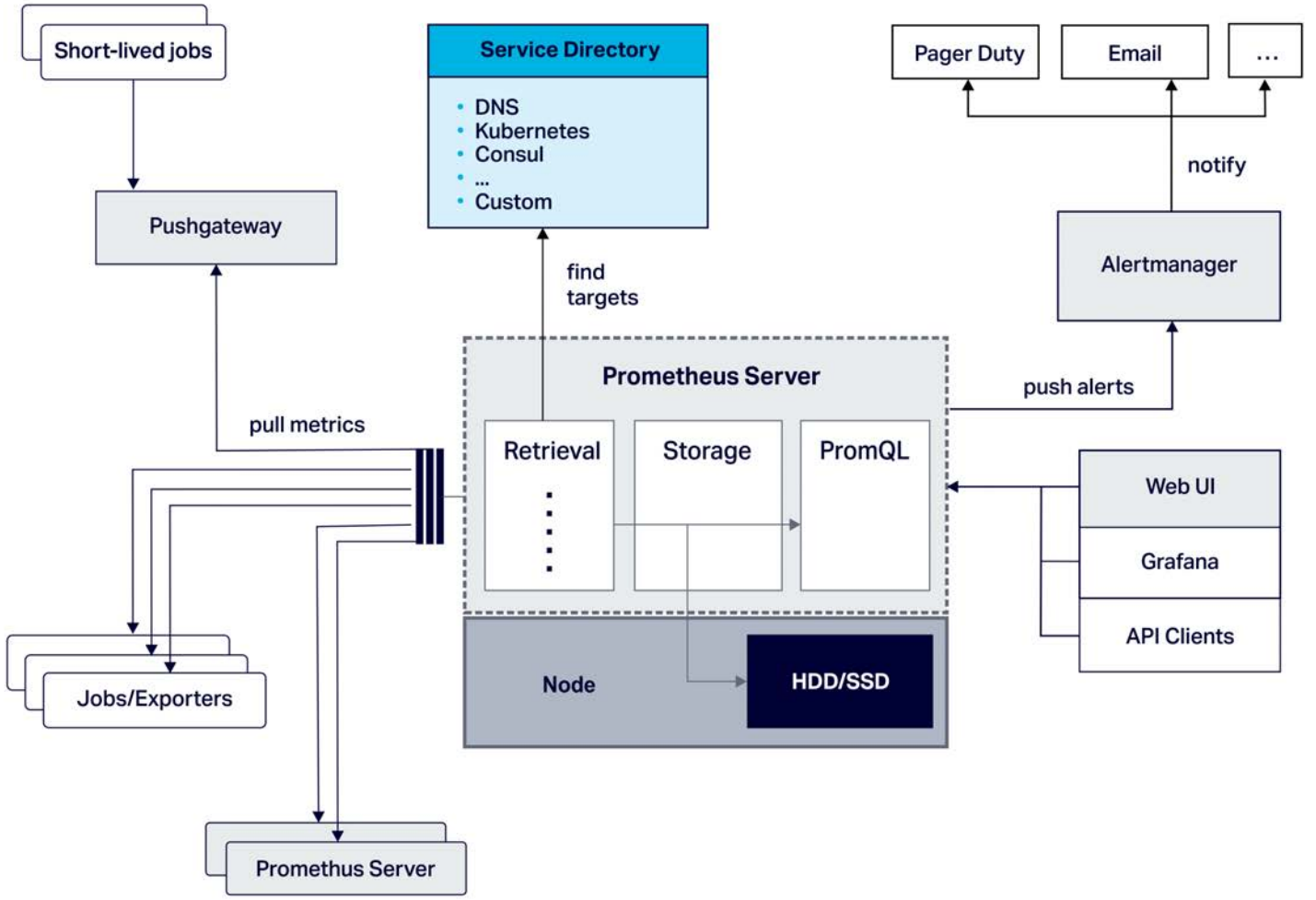
Observability 1 Prometheus Monitoring

- Ran using Kubernetes Prometheus Operator
- Grafana for graphing
- Used to debug, tune, and observe business metrics (TPS, RT) from 100 Pods



Prometheus Architecture

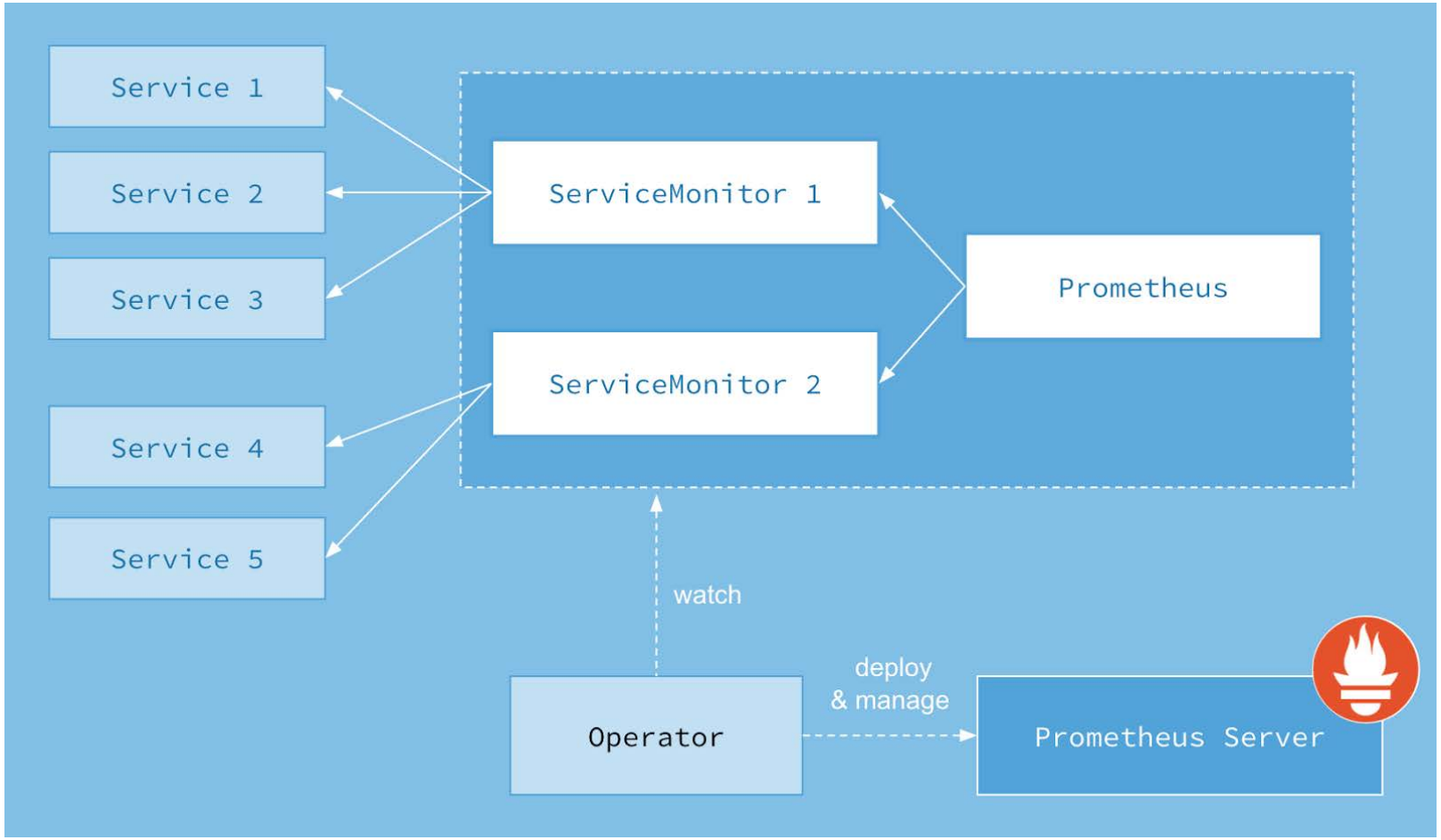
- Monitoring of applications and servers
- Instrumentation
- Pull-based
- Architecture & Components…



Prometheus Operator

In production on Kubernetes

Use Prometheus Operator to manage application complexity and dynamics



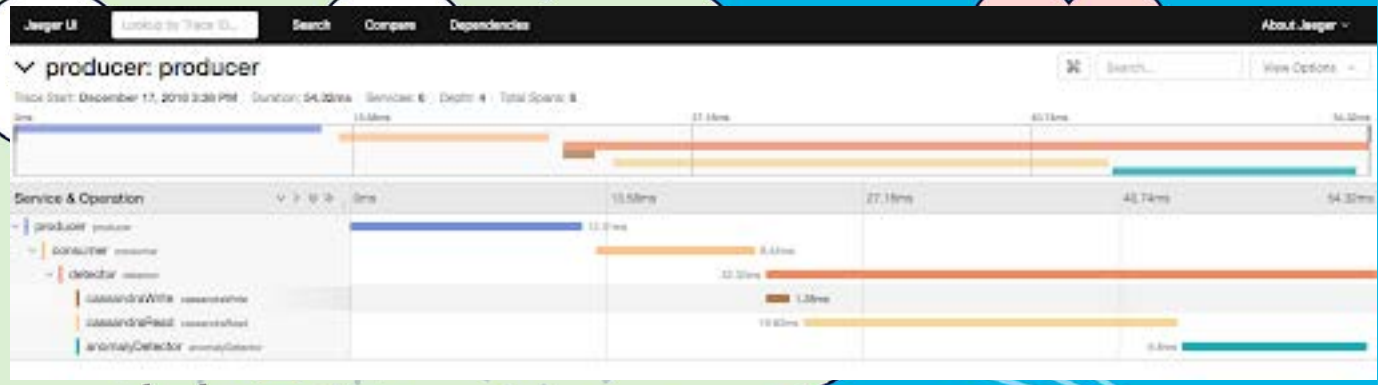
Observability 2 Tracing with OpenTracing and Jaeger

- Single traces
- Topology of system
- Even though this example has simple topology, valuable for debugging

JAEGER

OPENTRACING

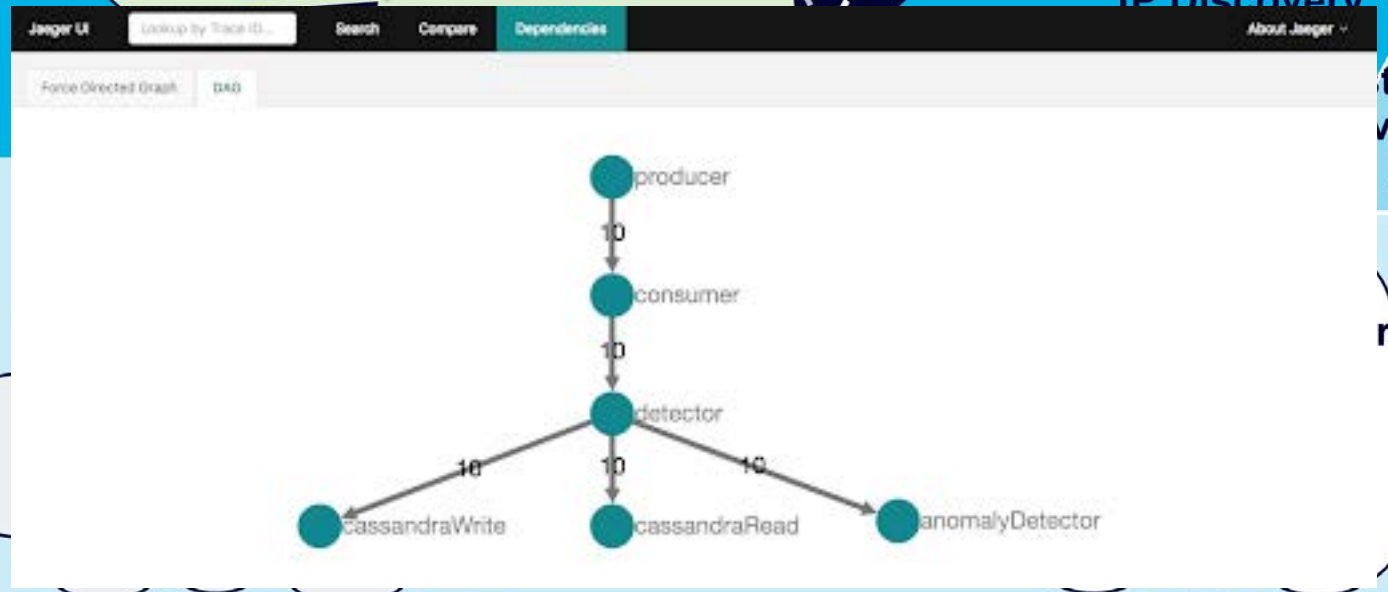
AWS Region



kubernetes

IP Discovery

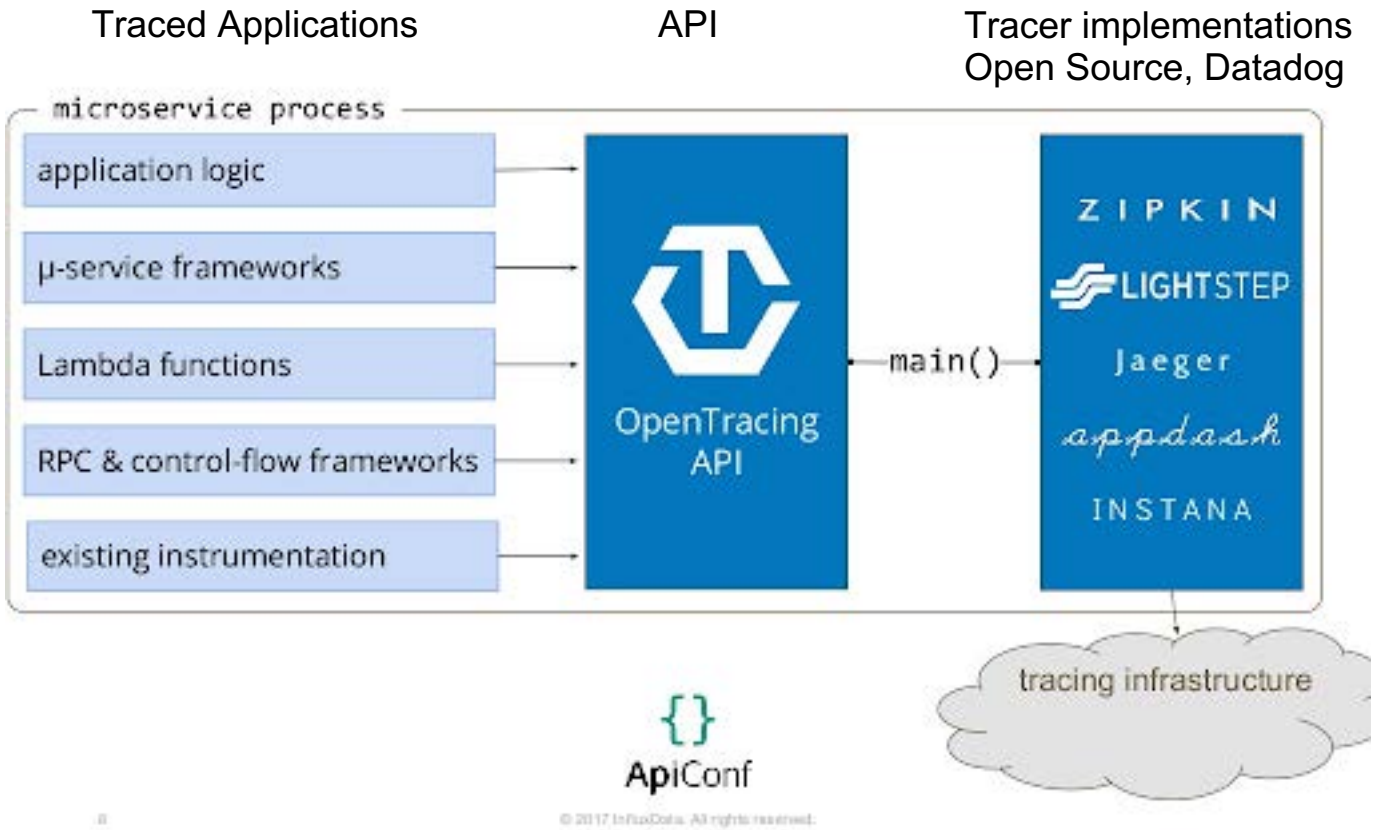
taclustr
visioning
API



OpenTracing

Standard API for distributed tracing

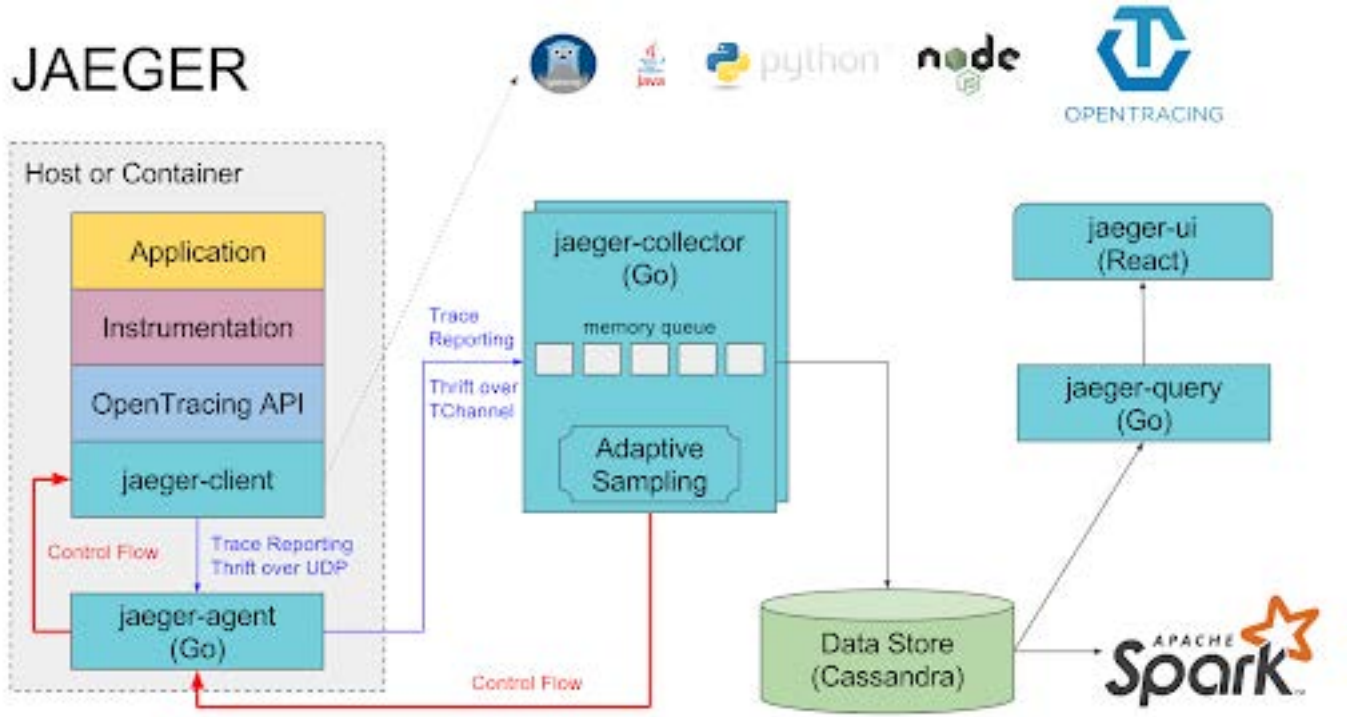
- Specification, not implementation
- Need
  - Application instrumentation
  - OpenTracing tracer



Jaeger Tracer

Open Source Tracer
Uber/CNCF

Scalable architecture



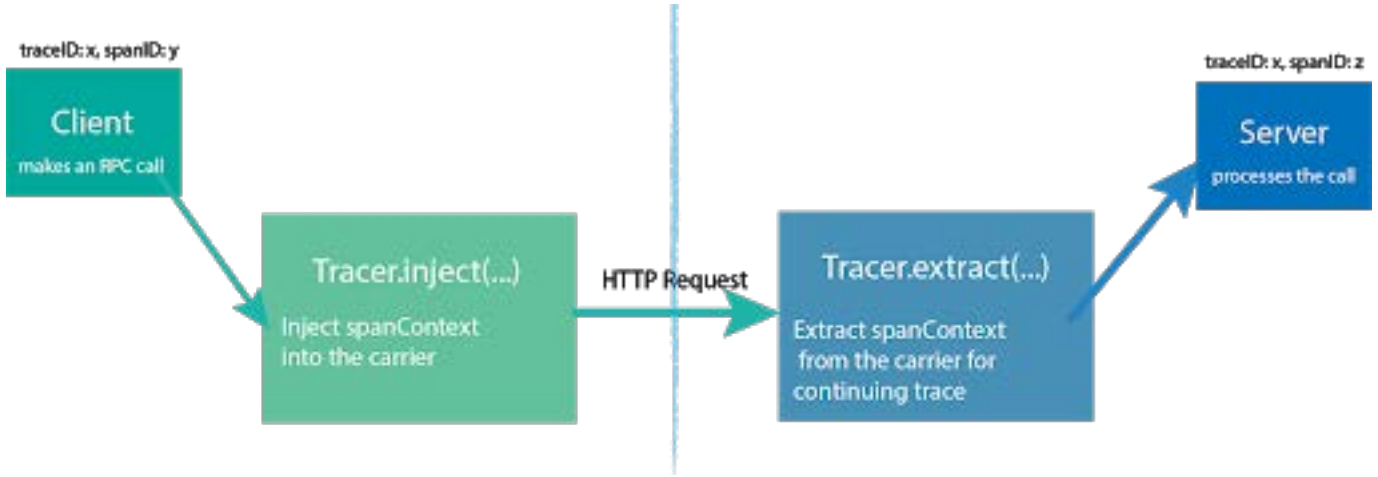
JAEGER

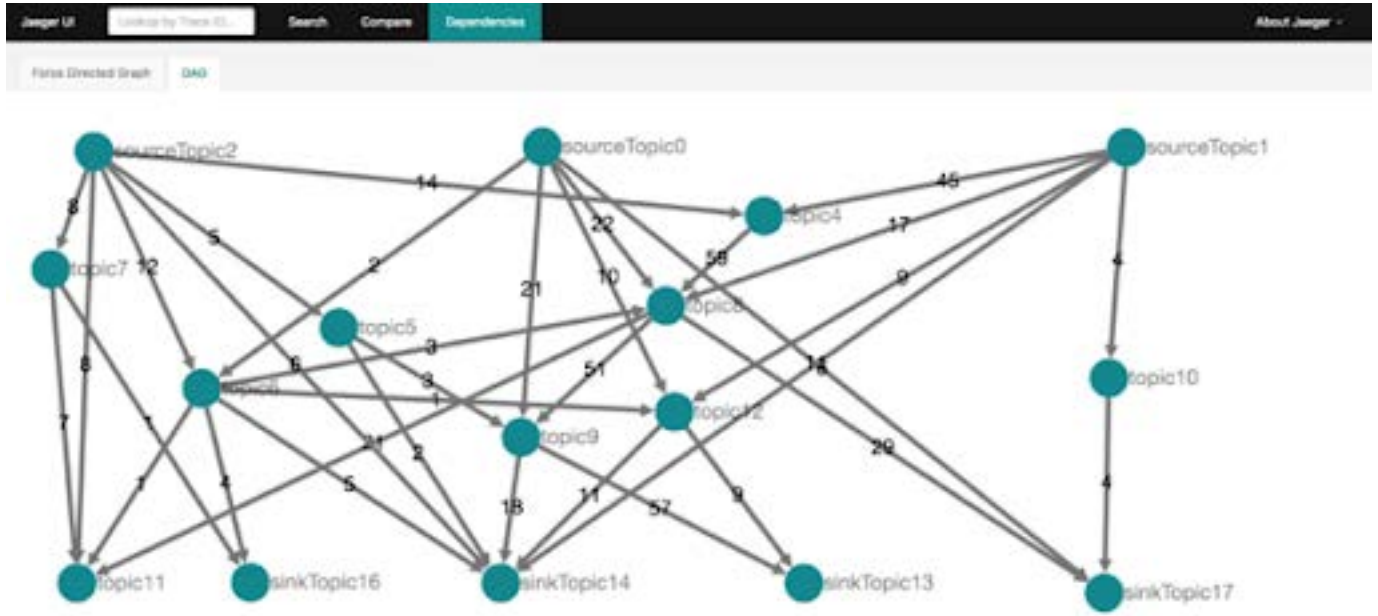
Tracing across Kafka topics

More complex example:

discovering event flows across multiple topics

E.g. Kafka ESB





5 How *well* did it work? Scaling Out

From 3 to ???
Cassandra nodes

How well did it work? Scaling Out

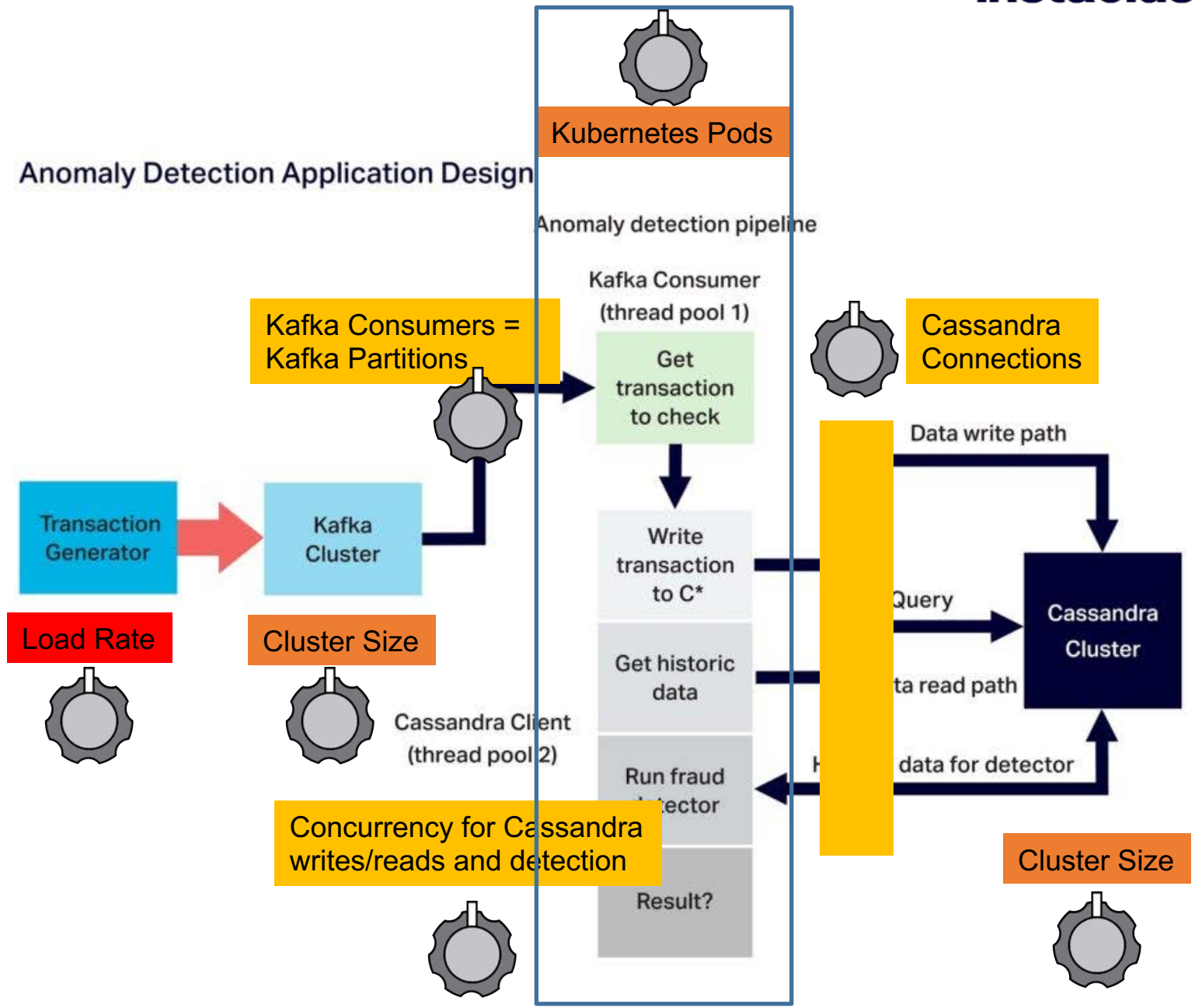
From 3 to ???
Cassandra nodes

Due to 1:1 read/write ratio, decreased compression chunk size to 1KB

“La Jamais Contente”, first car to reach 100 km/h in 1899 (electric, 68hp)

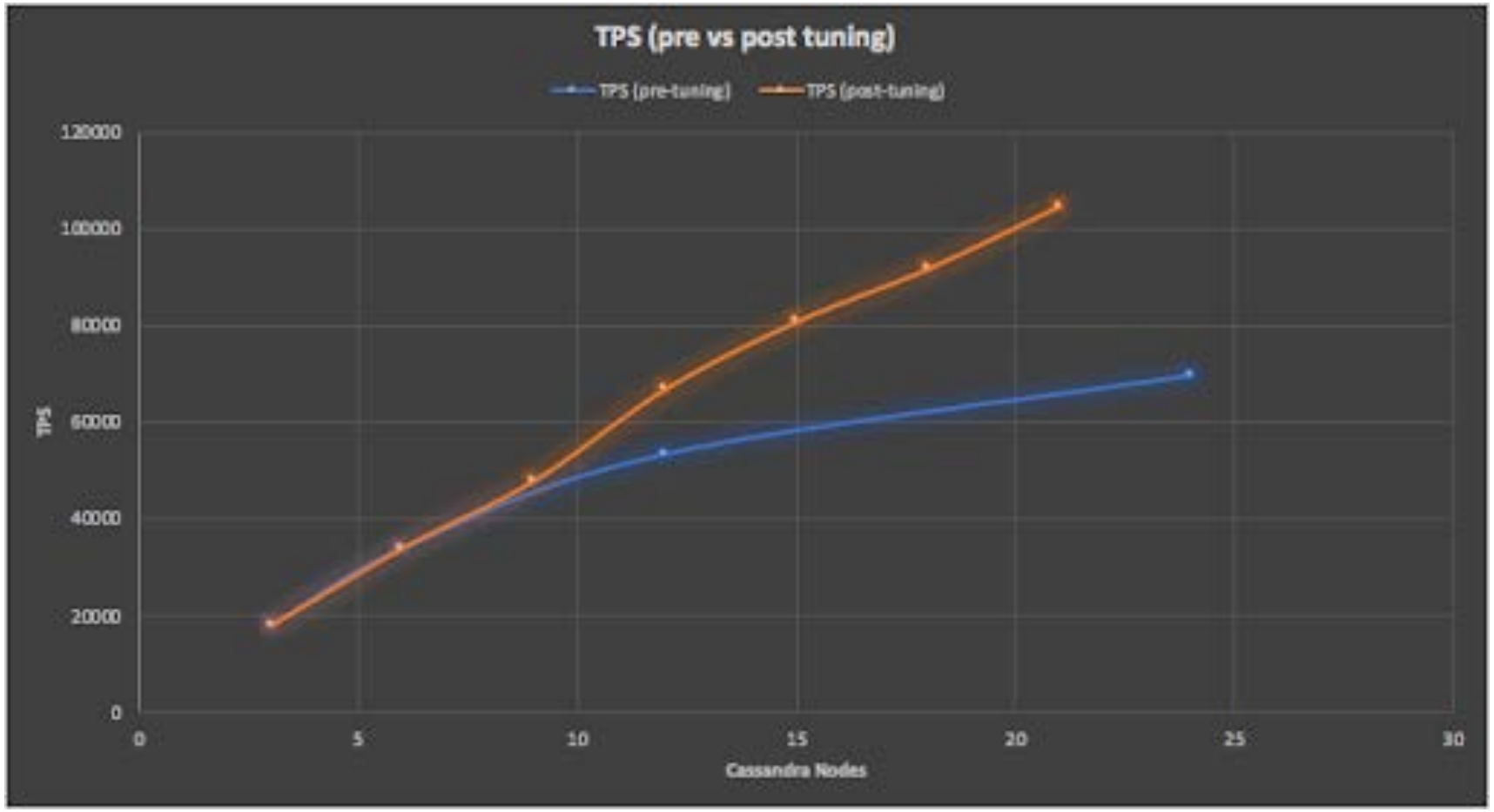
Scaling Knobs

- Load generator (red)
- Cluster sizes and worker pods (orange)
- Thread pools, partitions and connections (yellow)



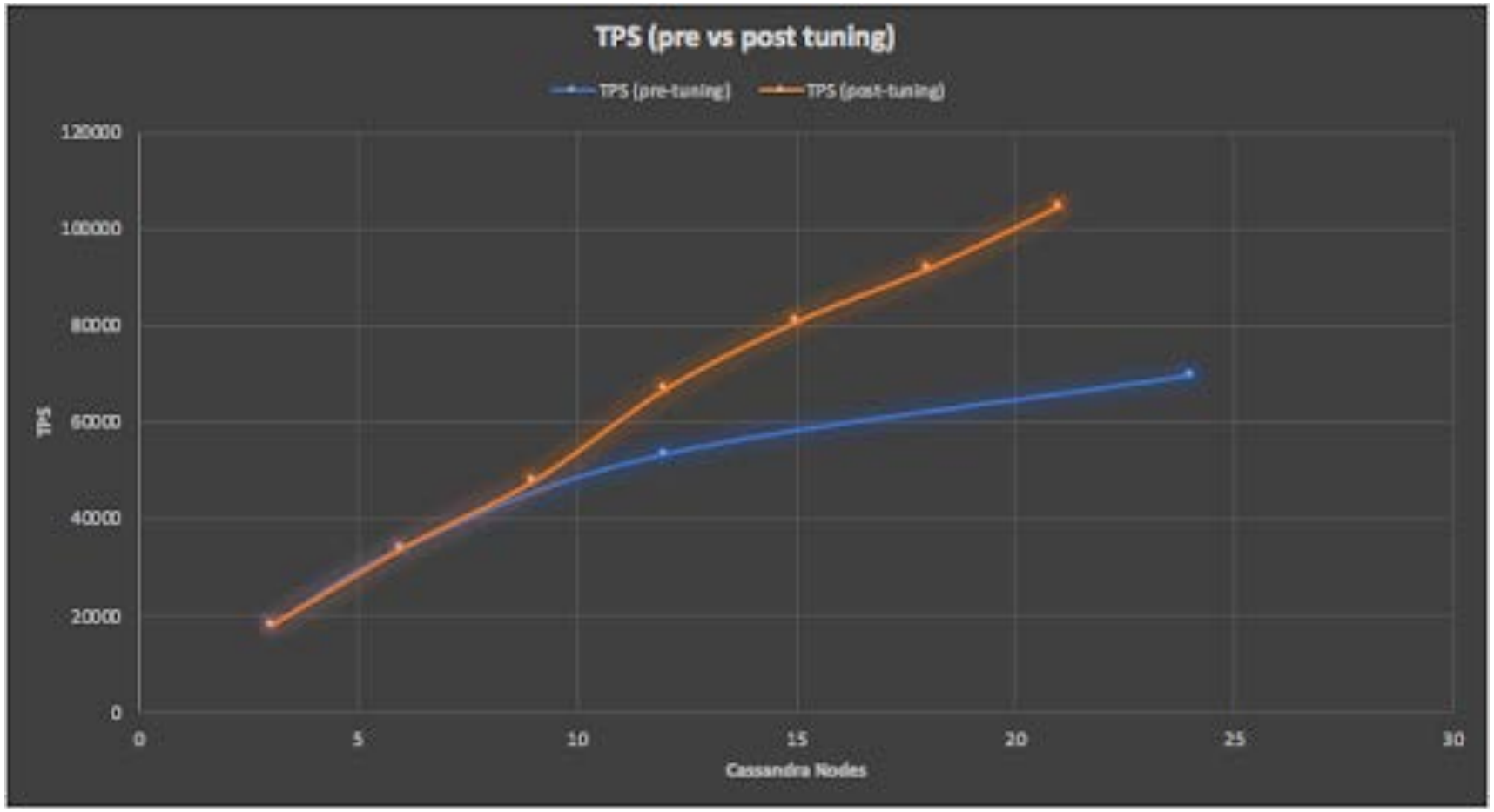
Cassandra scalability

- Kubernetes → easy to scale application, just increase Pods
- First attempt, tuned for 3 node Cassandra cluster then scaled out to 24 nodes
- Whoops (blue line)



Cassandra scalability - better

- Then tuned knobs (*thread pools, Pods* and Cassandra *connections)* to maximize throughput for *each* configuration (orange line)
- Also tuned Kafka…



Minimize Cassandra Connections but *maximize* detector thread pool (pool 2) concurrency

Kafka Scaling

Kubernetes Pods x Kafka Consumer threads

→

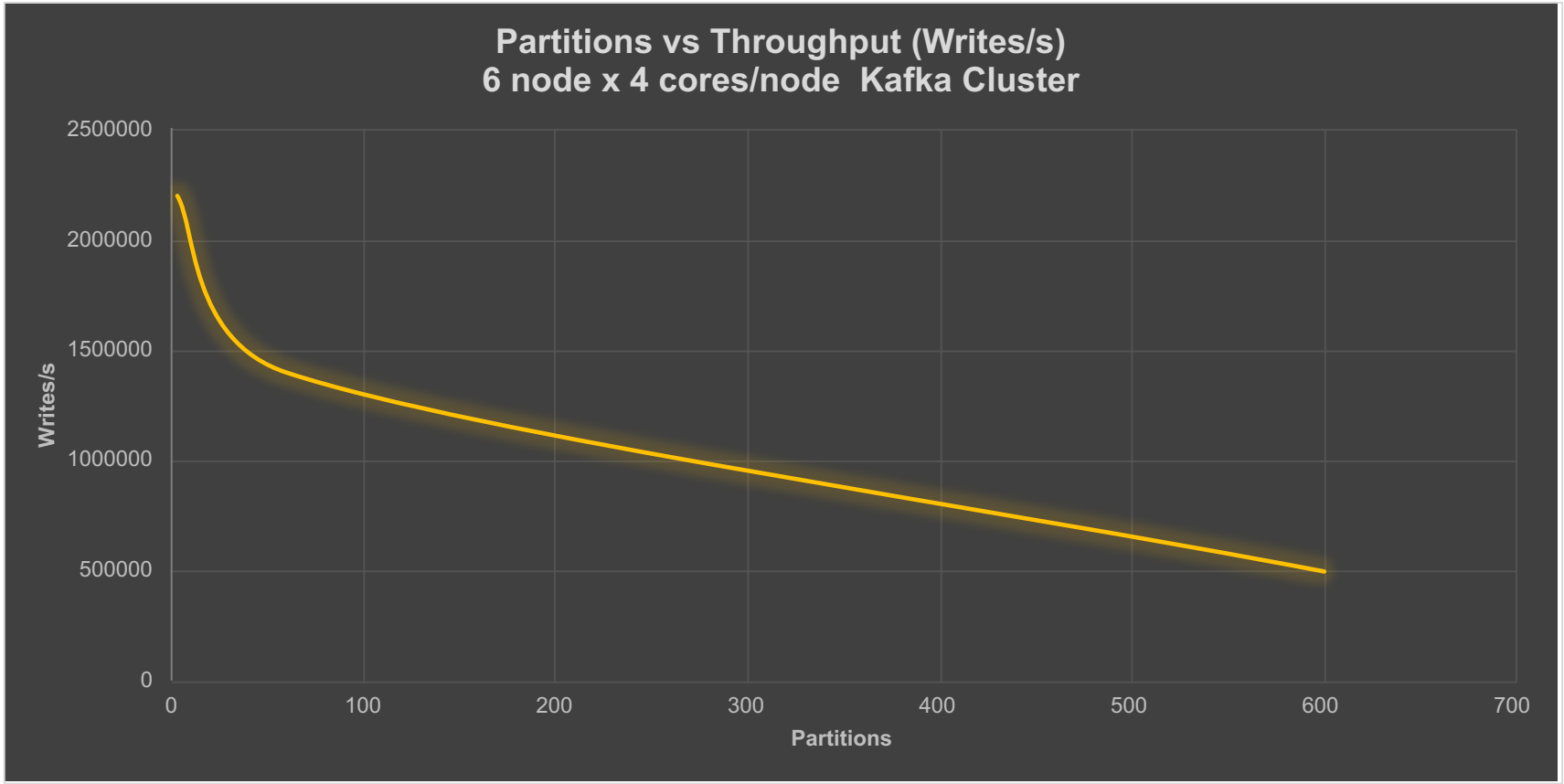
More Kafka Consumers

→

More Kafka Partitions

→

Lower Throughput!

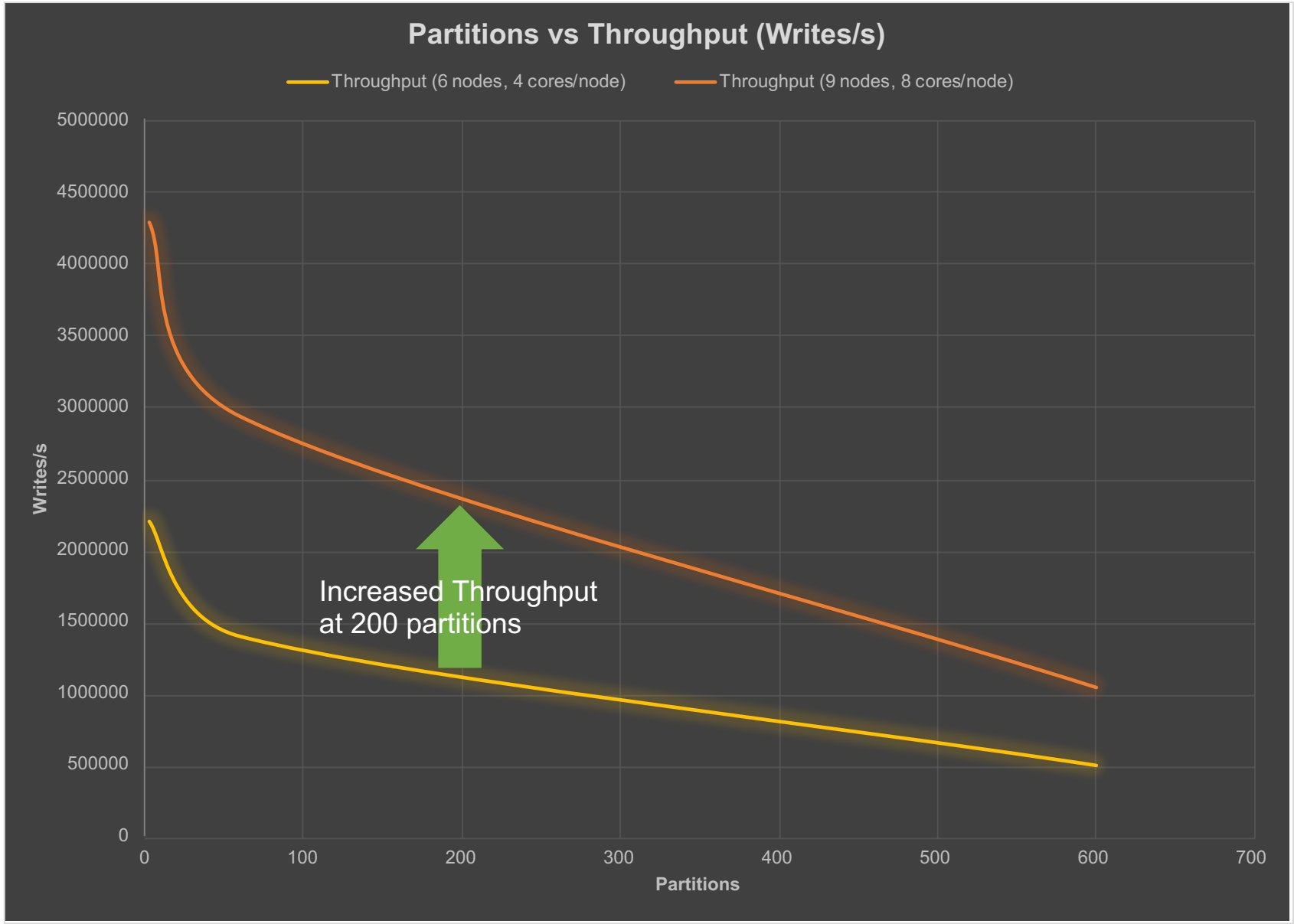


Kafka Scaling - better

Solutions?

Bigger Kafka cluster

Kafka tuning?
num.replica.fetchers = 1 by default, may help to increase



Final system resources

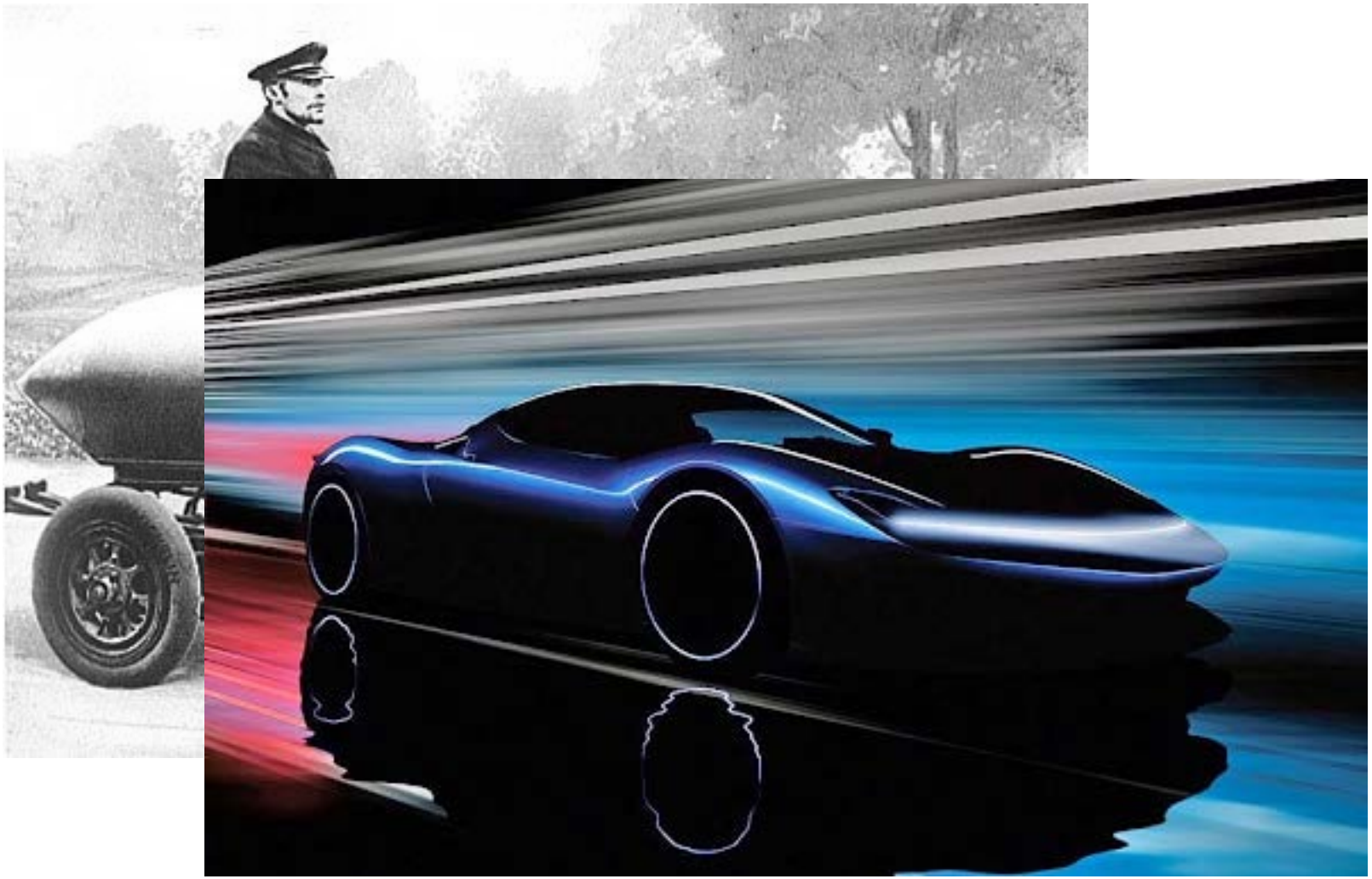
Cluster Details (all running in AWS, US East North Virginia)

- Instaclustr managed Kafka – EBS: high throughput 1500, 9 x r4.2xlarge-1500 (1,500 GB Disk, 61 GB RAM, 8 cores), Apache Kafka 2.1.0, Replication Factor=3
- Instaclustr managed Cassandra – Extra Large, 48 x i3.2xlarge (1769 GB SSD, 61 GB RAM, 8 cores), Apache Cassandra 3.11.3, Replication Factor=3
- AWS EKS Kubernetes Worker Nodes – 2 x c5.18xlarge (72 cores, 144 GB RAM, 25 Gbps network), Kubernetes Version 1.10, Platform Version eks.3

instaclustr

Scaling Out

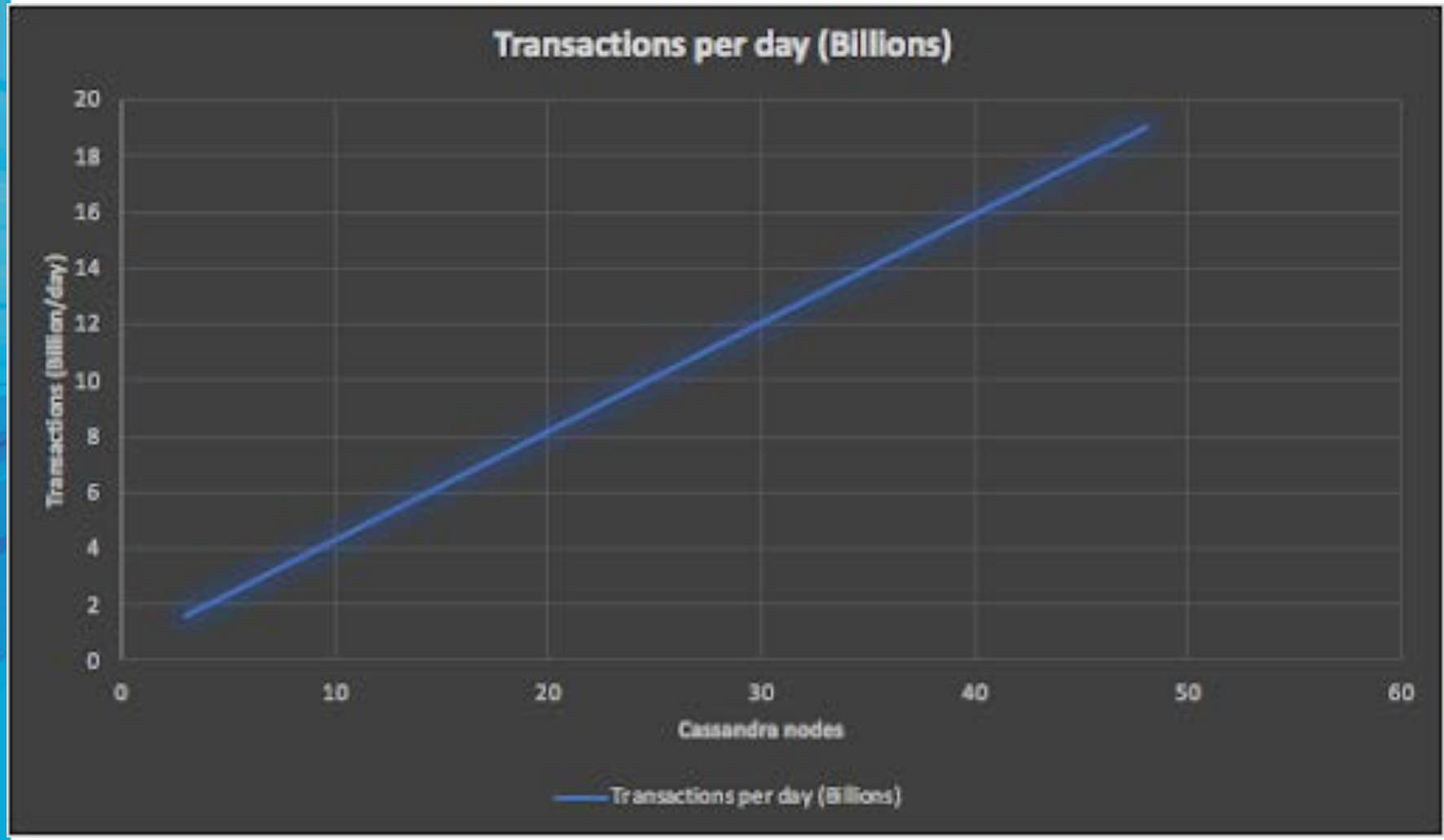
From 3 to ??
Cassandra nodes



“Pininfarina Battista” the fastest car in the world (2019)
0-100 kph in 2 seconds, top speed 350 kph (electric, 1,900hp).

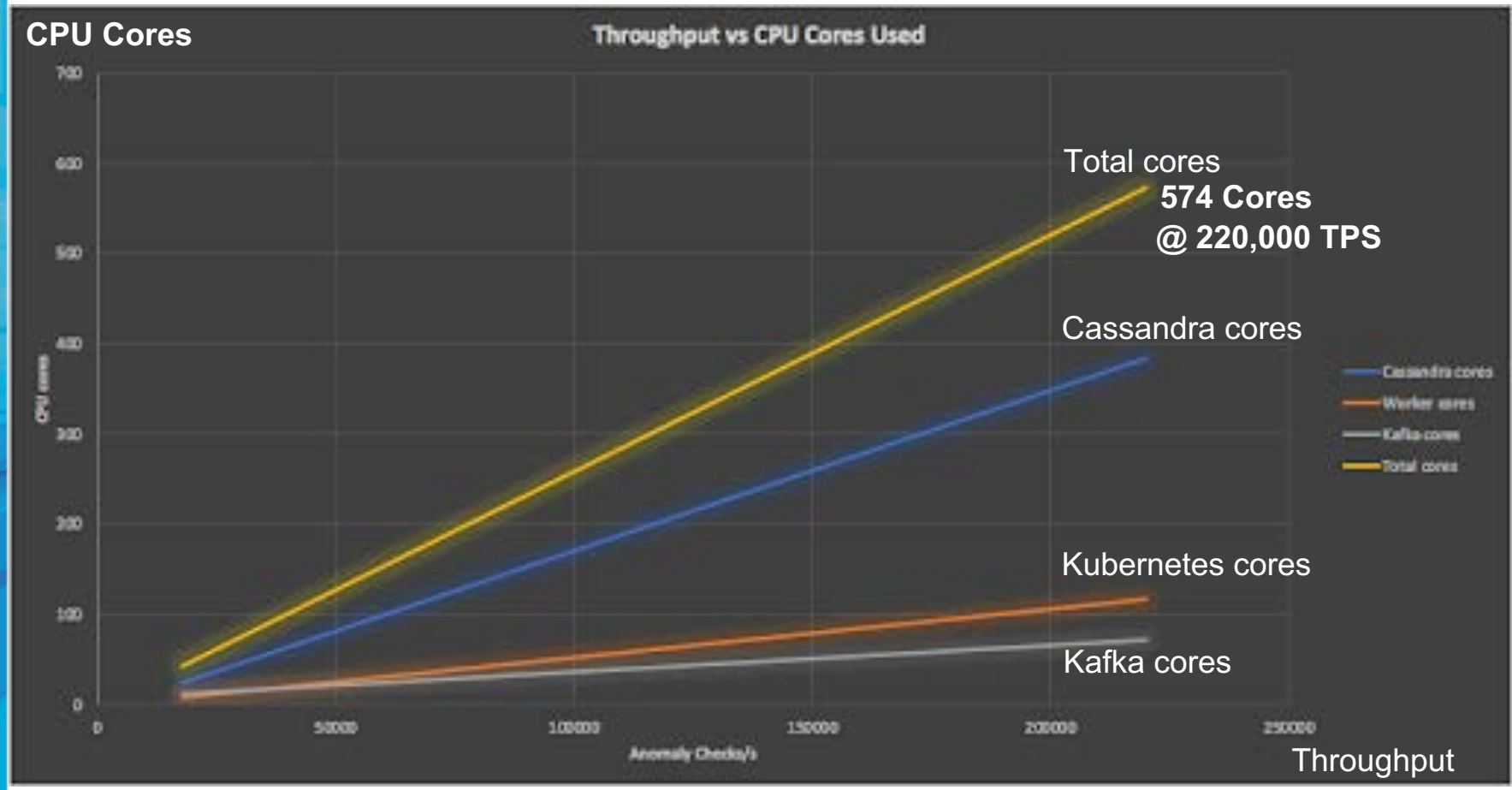
Scaling Out

- From 3 to 48 Cassandra Nodes
- 1.9 to19 Billion checks/day
- No upper limit



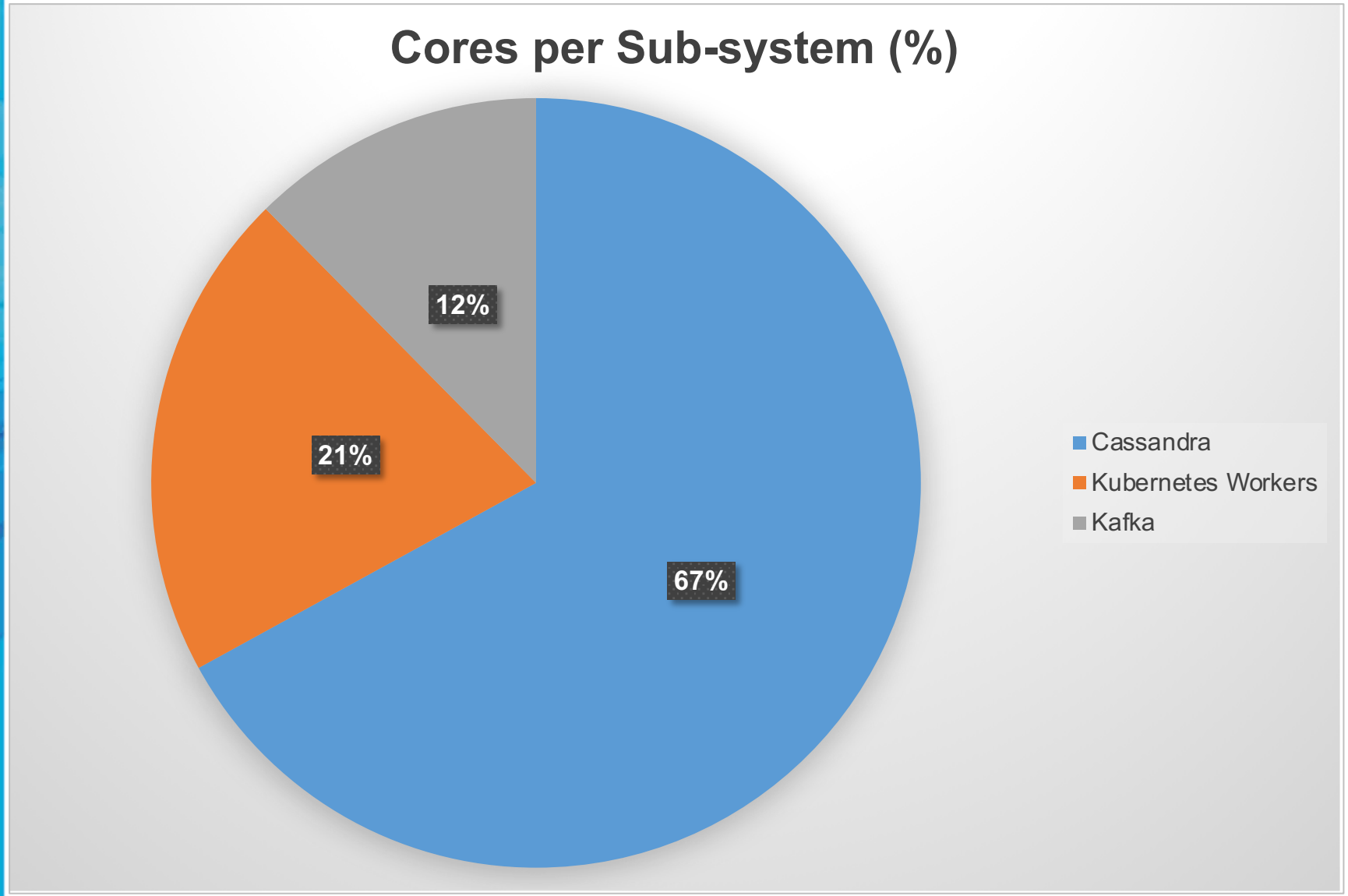
Resources

- Throughout (checks per second) vs cores for each subsystem:
- Cassandra > Workers > Kafka
- Maximum 574



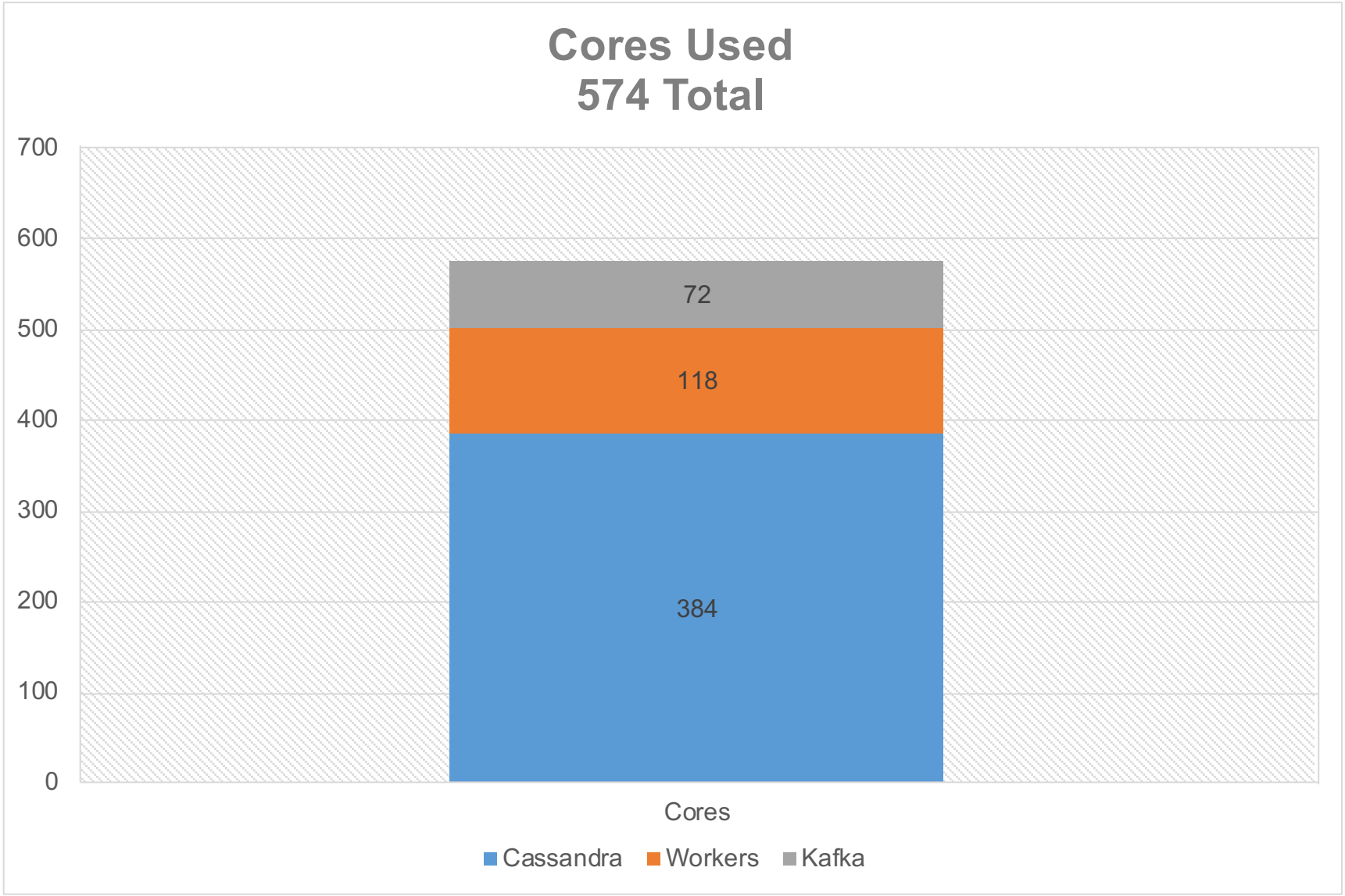
Cores used - balance

Cassandra (67%) > Kubernetes (21%) > Kafka (12%)



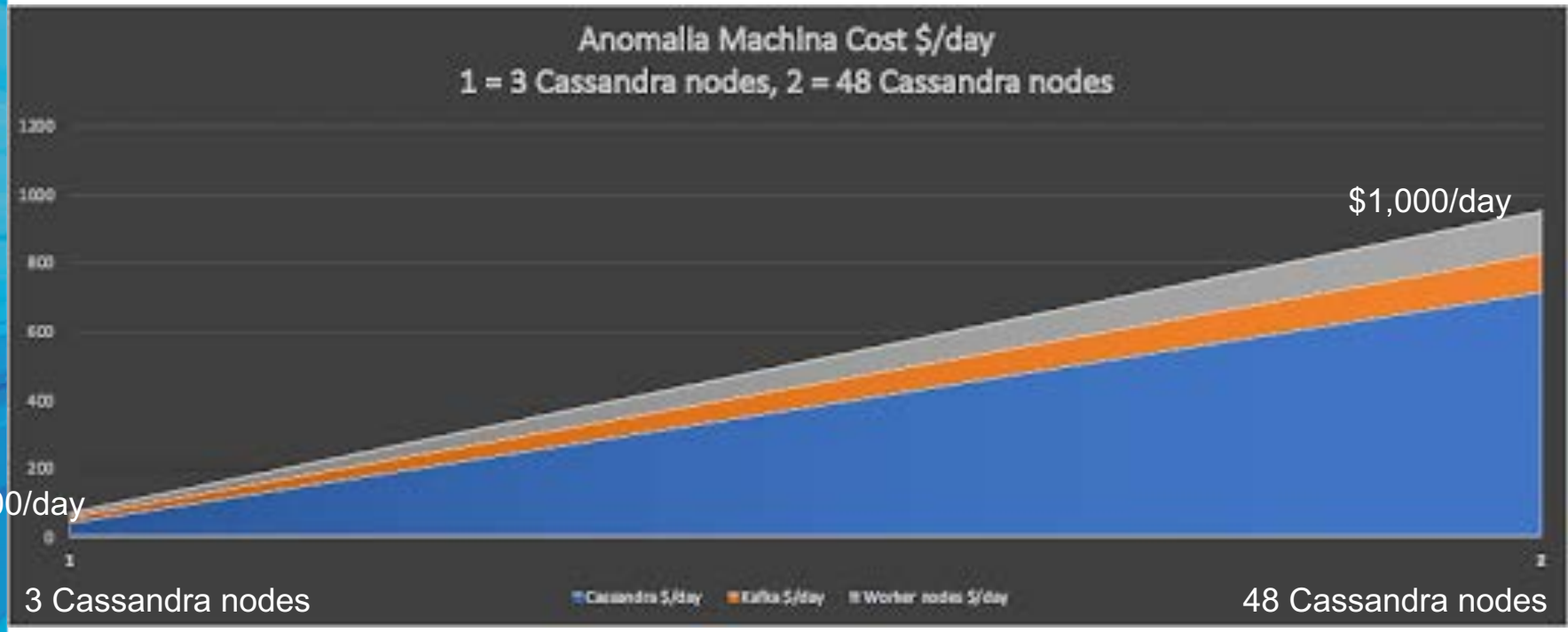
Maximum cores used

Cassandra 384 +
Workers 118 +
Kafka 72 =
574 Cores Total



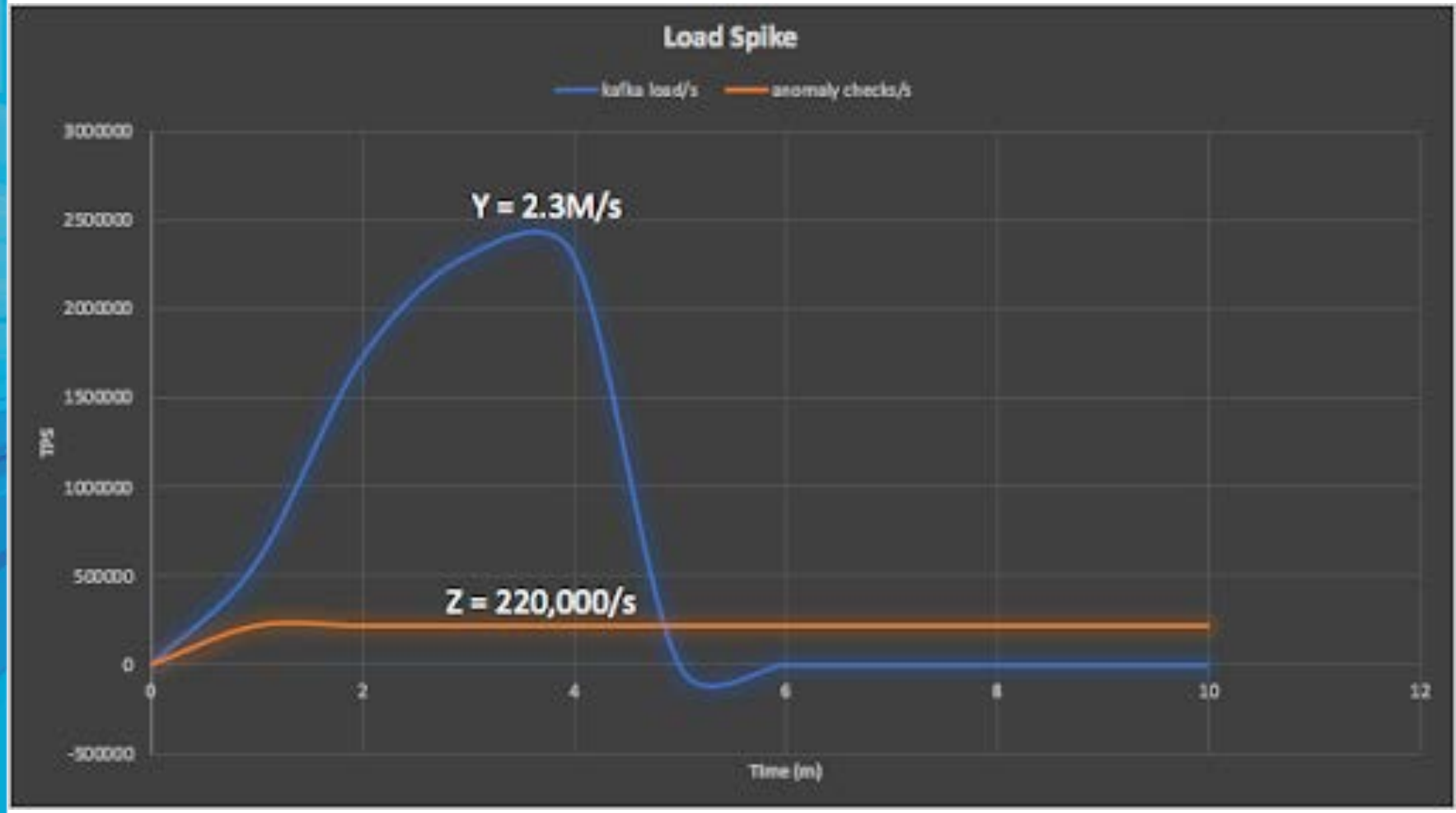
Cost – Affordability at scale

- Operational $ (AWS instances) only
- Total $1,000/day
- Can be scaled with incremental cost change



Kafka as a Buffer

- Kafka acts as a buffer, can process 10x the Cassandra capacity
- 2.3M/s vs 220,000/s
- Cheaper than increasing Cassandra capacity x10



6 So What?

Some Takeaways

Takeaways

Technical

- Kubernetes (+AWS EKS) enabled automation (deployment, scaling, monitoring) of the application
  - Some effort to understand and setup
  - But once working it makes application deployment fast, scalable, repeatable and low cost
- Prometheus and OpenTracing+Jaeger critical for debugging, tuning and reporting application performance and scalability
  - Tricky to monitor applications in Kubernetes, but using the Kubernetes Operators automates the monitoring configuration
- To achieve near linear scalability and ***maximize throughput*** need to optimize pipeline, by tuning thread pools and number of Kubernetes Pods to:
  - Minimize: Cassandra *Connections*
  - Minimize: Kafka Consumers → Kafka *Partitions*
  - Maximize: Detector thread pool concurrency

Takeaways

Business

- Kafka+Cassandra enable Fast Streaming+Storage at Scale
- Instaclustr Managed Kafka+Cassandra service
  - Makes it easy to automate cluster provisioning (creation/deletion/scaling), and monitoring
  - Highly available SLAs
  - Proactive cluster monitoring, alerting and maintenance
- Affordability at Scale
  - Low cost Open Source and Commodity Cloud infrastructure
  - only pay for what you use, application and Kafka+Cassandra clusters scale linearly with load so cost only increases incrementally
- Application can be easily resized (scaled up and down) for any workload, no upper limit
- Lots more use cases using Kafka+Cassandra

Newsflash!

Geospatial Anomaly Detection

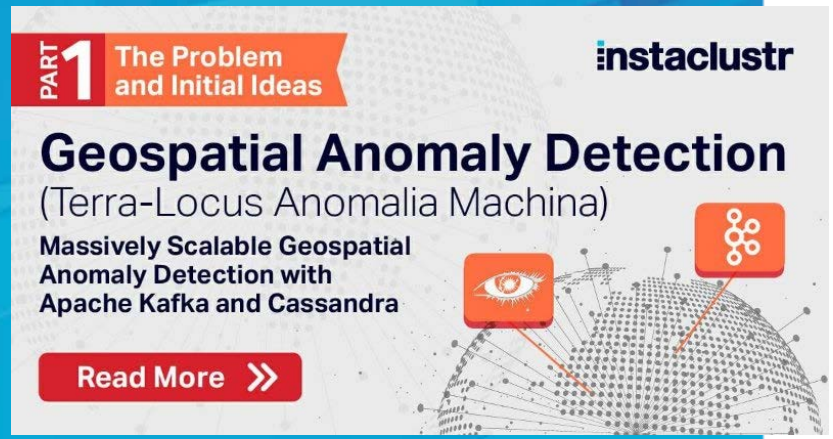
Newsflash!

Geospatial Anomaly Detection

Compared performance of multiple Spatial representations and Cassandra implementations

- Extensions to detect anomalies over time and *space*
  - E.g. is an event unusual relative to nearest 50 neighbours?
- How to find neighbours using
  - Distance between Latitude/longitude points
  - Bounding Box
  - Geohashes
  - 3D (including 3D Geohashes)
- Using different Cassandra implementations
  - Clustering columns
  - Secondary indexes
  - Denormalized multiple tables
  - Cassandra Lucene Index Plugin

Further information



- The complete *Anomalia Machina* Blog Series (10 Parts):
  - Massive scale Kafka and Cassandra deployment for real-time anomaly detection: 19 Billion events per day https://www.instaclustr.com/massive-scale-kafka-cassandra-real-time-anomaly-detection/
- Latest 4-part *Geospatial* Anomaly Detection blogs:
  - https://www.instaclustr.com/geospatial-anomaly-detection-with-kafka-cassandra/
- The Open Source Anomalia Machina Code
  - https://github.com/instaclustr/AnomaliaMachina
- All of Paul's Blogs
  - https://www.instaclustr.com/paul-brebner/

instaclustr

Some Anomalies are easy to detect


Evening News Noon

Moon men fight for their lives 200,000 miles from earth

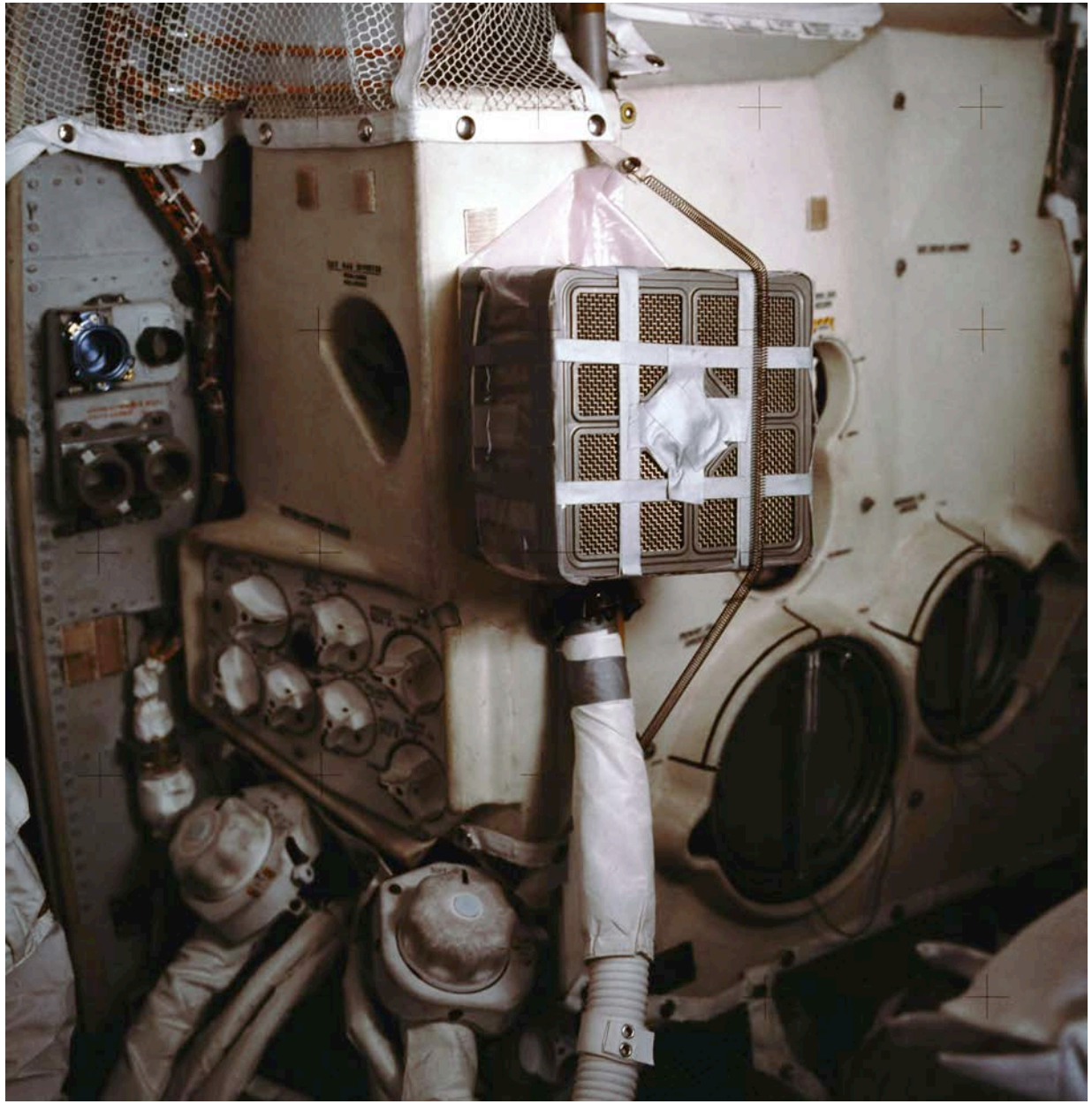
EXPLOSION IN APOLLO 13!

"Houston, we've had a problem … we had a pretty large bang"
April 14, 1970

instaclustr

But hard to fix

The ultimate space hack – the Supplemental $C0_2$ Removal System



instaclustr

Other anomalies are harder to detect

Earth now has 2 Moons!

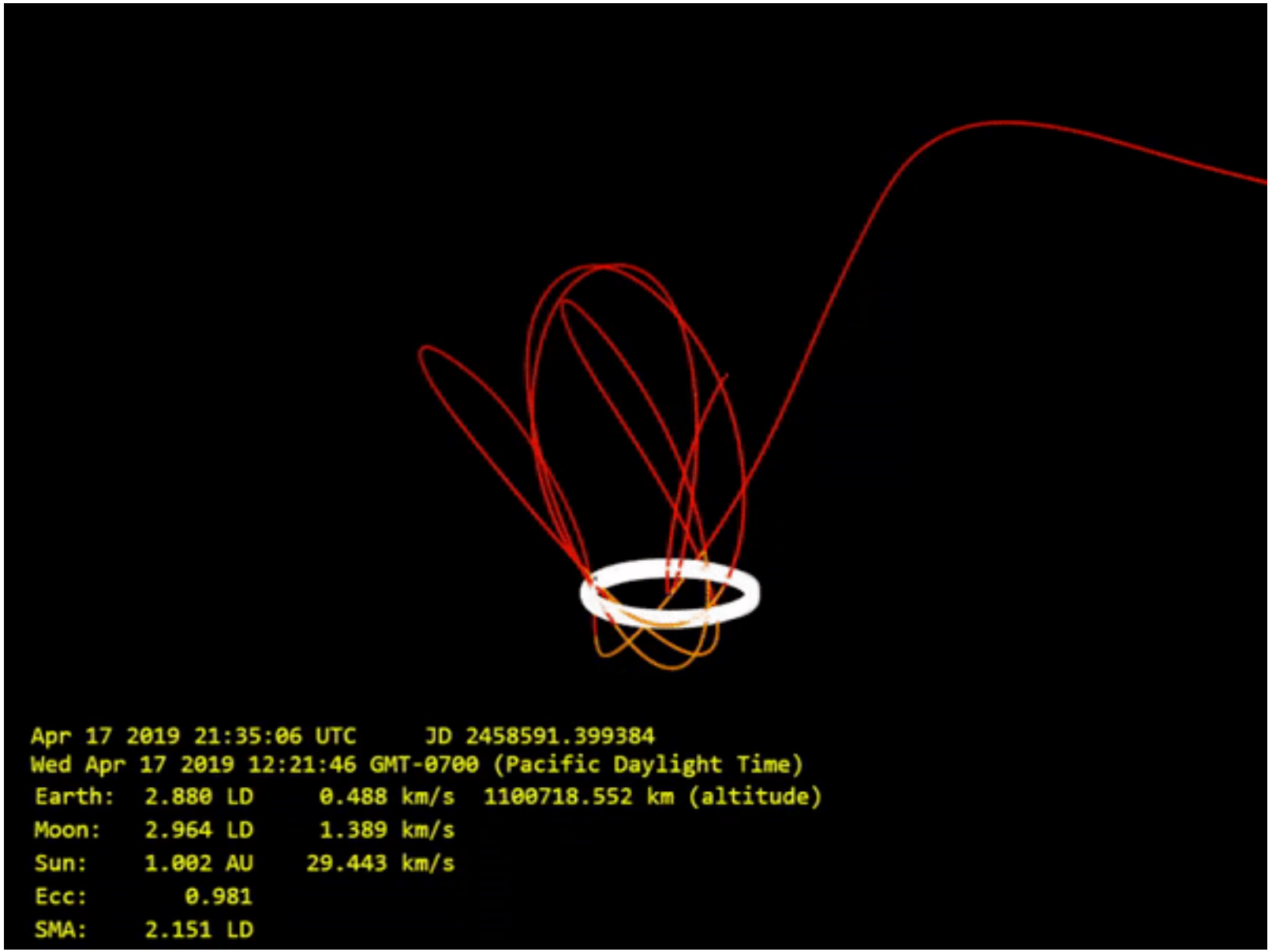
Mini-moon, car size, temporary!

instaclustr

Try out the Instaclustr Managed Platform for Open Source

www.instaclustr.com/platform/

It's easy to detect complex spatio-temporal anomalies reliably at scale with Kafka, Cassandra & Kubernetes



THE END