



LinuxFest Northwest 2016

LinuxFest Northwest

April 24, 2016

Monitoring 101

Ilan Rabinovitch

Director, Technical Community

Datadog

```
$ finger ilan@datadog
```

```
[datadoghq.com]
```

```
Name: Ilan Rabinovitch
```

```
Role: Director, Technical Community
```

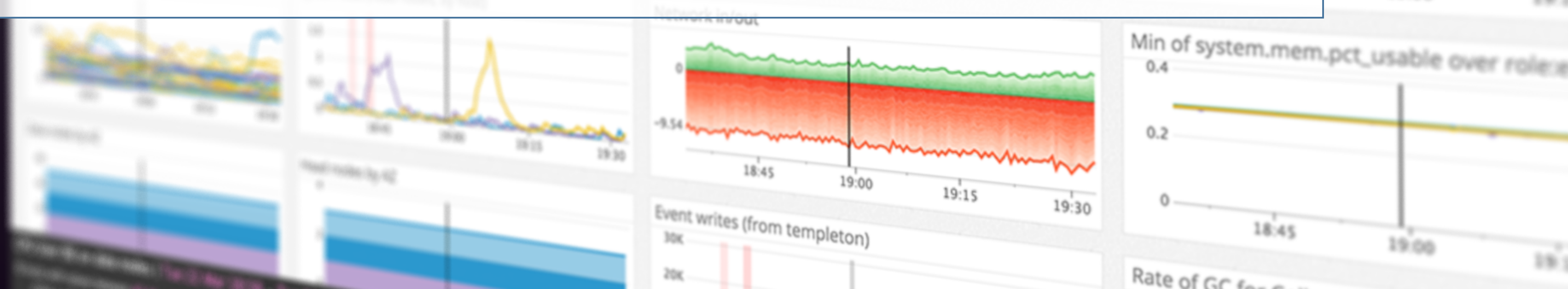
```
Interests:
```

- * Open Source
- * Large scale web operations
- * Monitoring and Metrics
- * Planning FL/OSS Community Events



Datadog Overview

- SaaS based infrastructure and app monitoring
- Open Source Agent
- Time series data (metrics and events)
- Processing nearly a trillion data points per day
- Intelligent Alerting
- We're hiring! (www.datadoghq.com/careers/)



Monitor Everything



AWS



Docker



CoreOS



Chef



Puppet



Github



Pagerduty



Nagios



Go



Postgres



Java



VMware



Redis



MySQL



Apache



Tomcat



MongoDB



New Relic

Operating Systems, Cloud Providers (AWS), Containers, Web Servers, Datastores, Caches, Queues and more...



Karthik Gaekwad

@iteration1

@datadog hi!! Can you help me monitor this resource allocation issue? I'm so tired...

@stackengine @datadoghq



Data @datadog · 4 Sep 2015

@iteration1 @stackengine @datadoghq The load balancers do not appear to be functioning properly on the brown server.



\$ cat ~/.plan

1. **Intro and Background:** What is DevOps?
2. **The Challenge:** Monitoring Dynamic Infrastructure
3. **Finding the Signal:** How do we know what to monitor?
4. **Implementation**

Our Focus Area

Culture
Automation
Metrics
Sharing



Damon Edwards and John Willis
DevOps Day LA

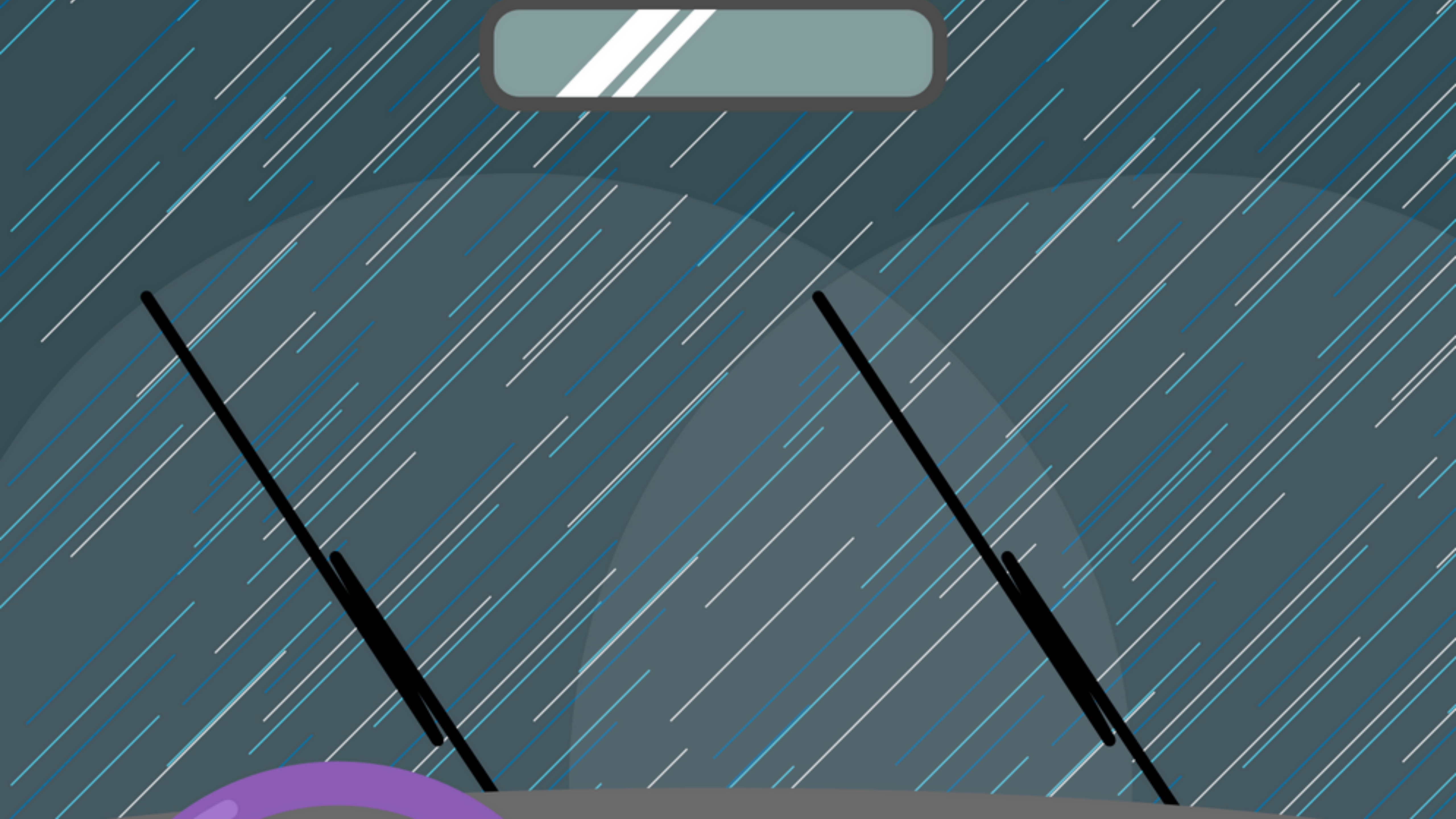
Culture

“organizations which design systems ...
are constrained to produce designs
which are copies of the communication
structures of these organizations”

- **Melvin E. Conway**









Honest Status Page @honest_update · 18 Sep 2015

We have no idea what's wrong so we're just gonna undo whatever we did last and whoever did it is the incident manager.

← 66 57 ...



Honest Status Page @honest_update · 20 Sep 2015

If there ain't an alert for it, it ain't broke!

← 33 28 ...

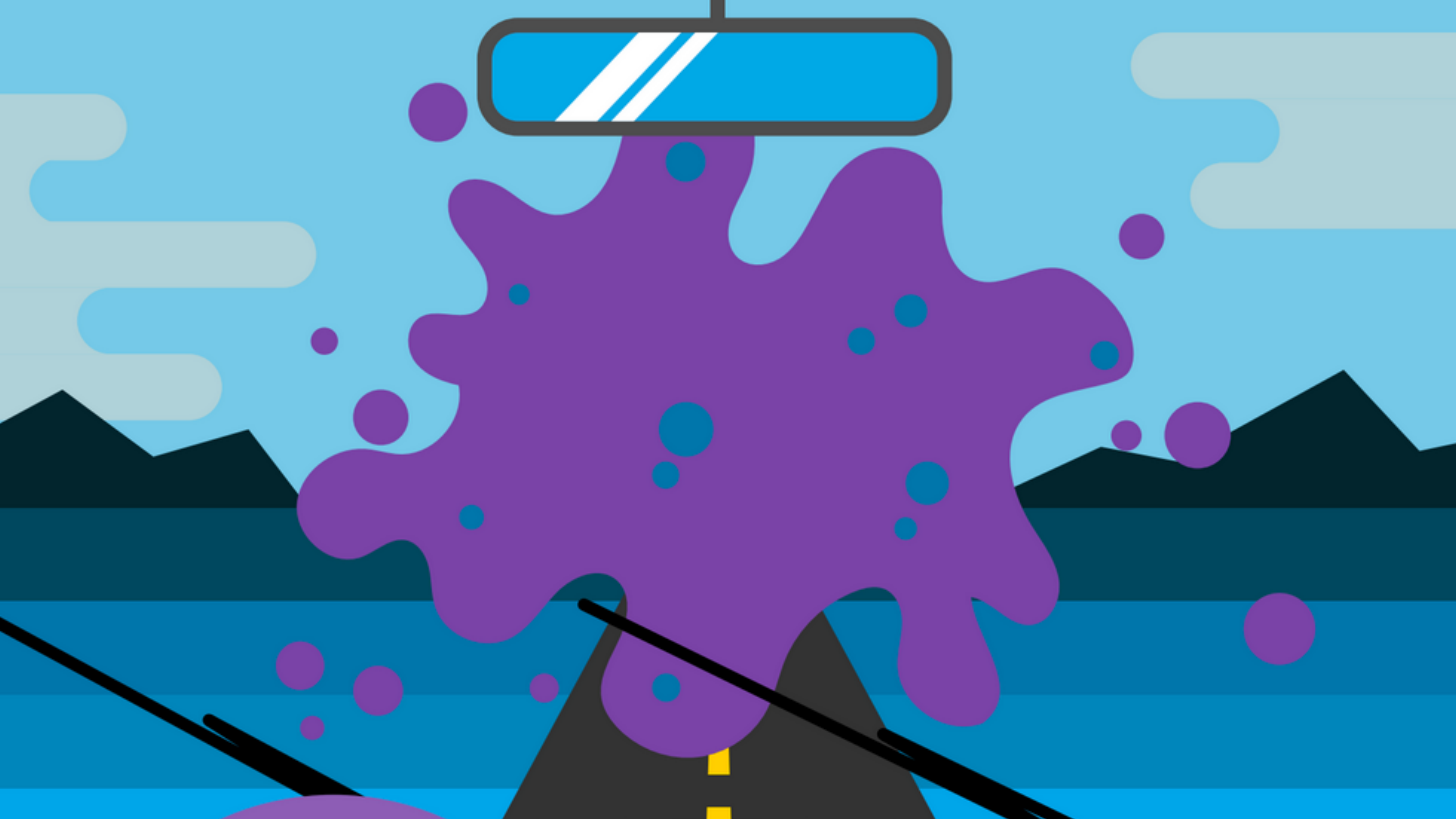


Honest Status Page @honest_update · 10 Mar 2015

Two words: unbounded queues.

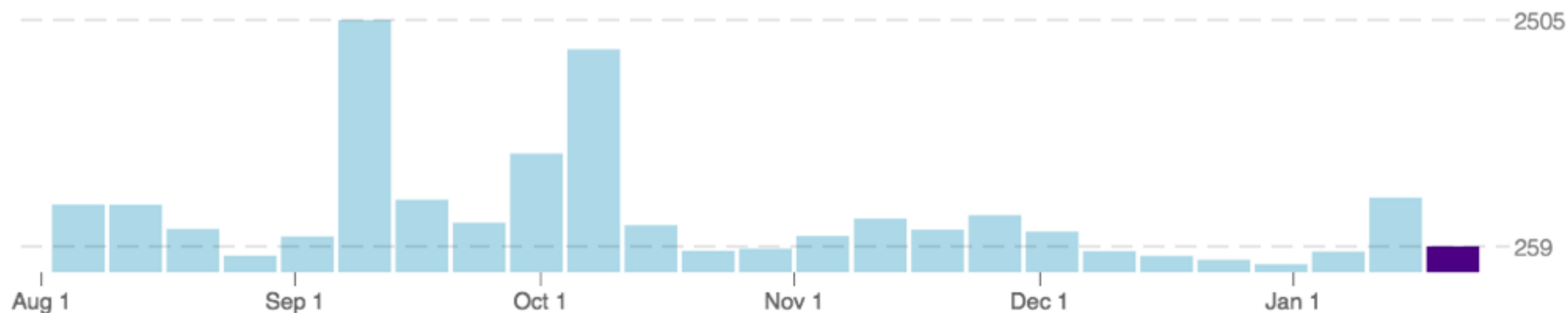
← 61 75 ...

Follow @honest_update on Twitter



Pagerduty Incident Trends

Week by week



259 alerts
in the week of Nov 17

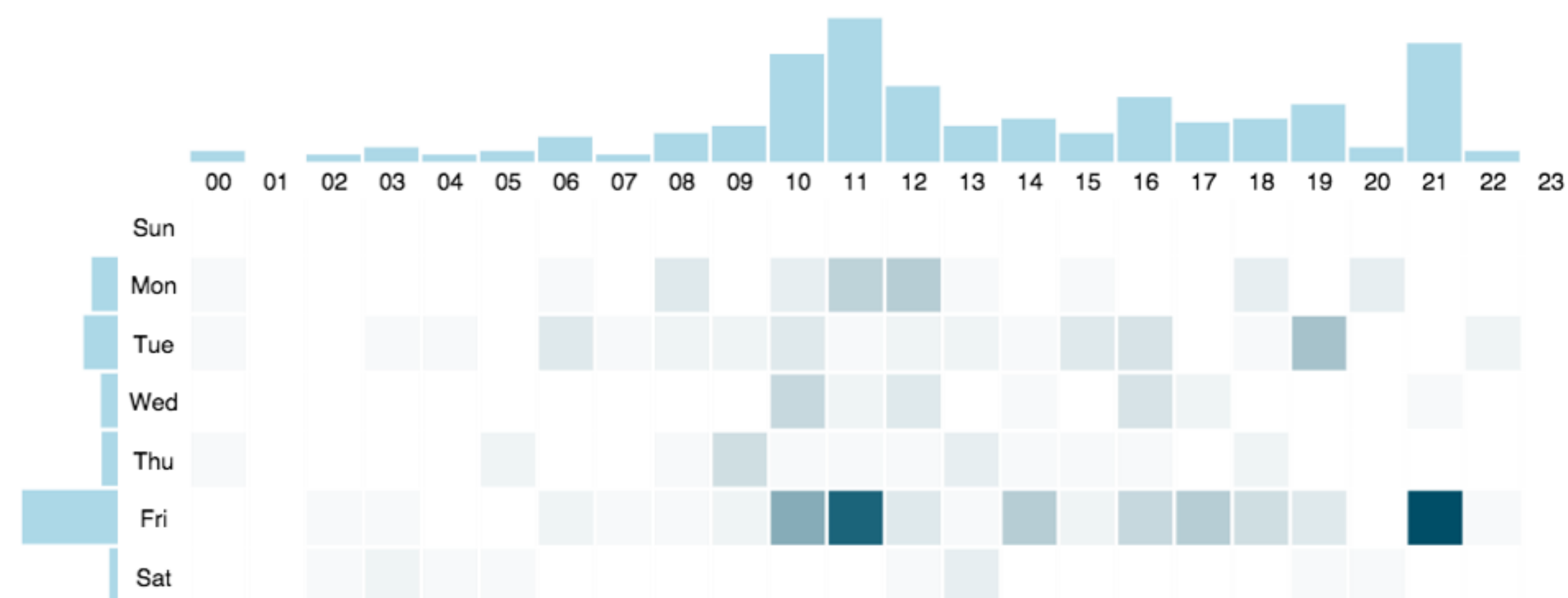
-65% from the previous week

Hadoop has alerted more than any other service for the last 3 weeks.

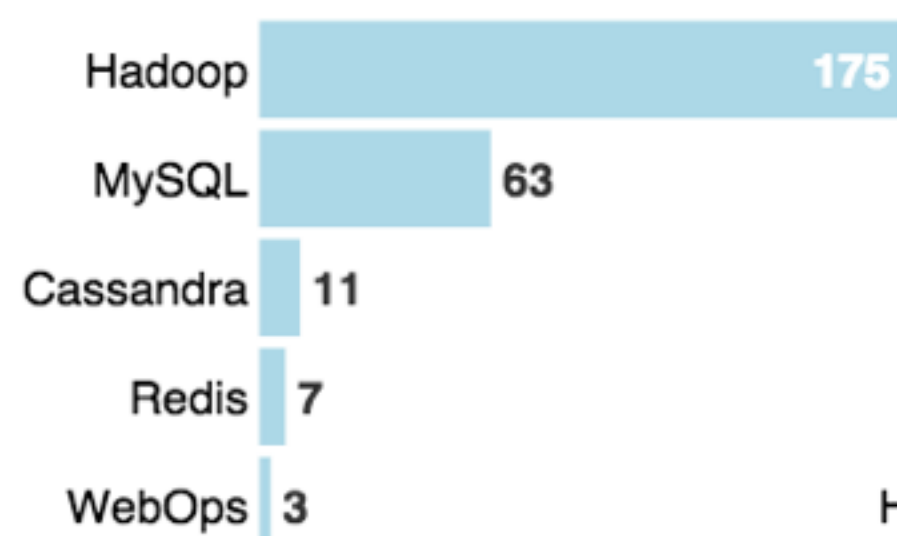
MySQL has alerted for the last 25 consecutive weeks.

MySQL alerted 63 times, **+1650%** from the previous week.

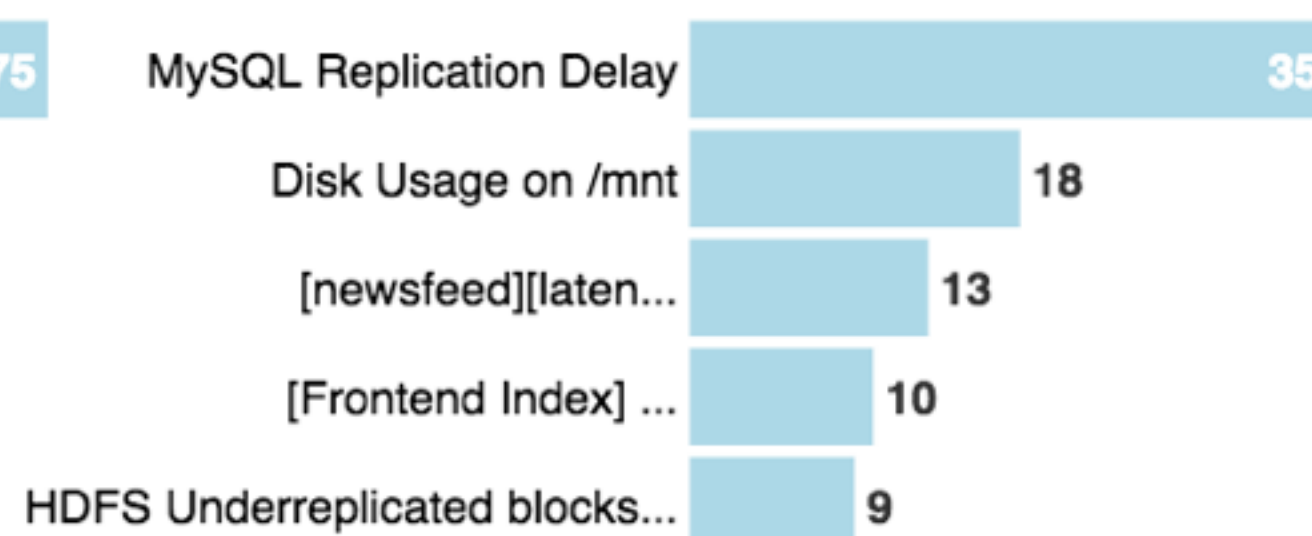
By day and hour



Top alerting services



Top alerting incidents



Sharing

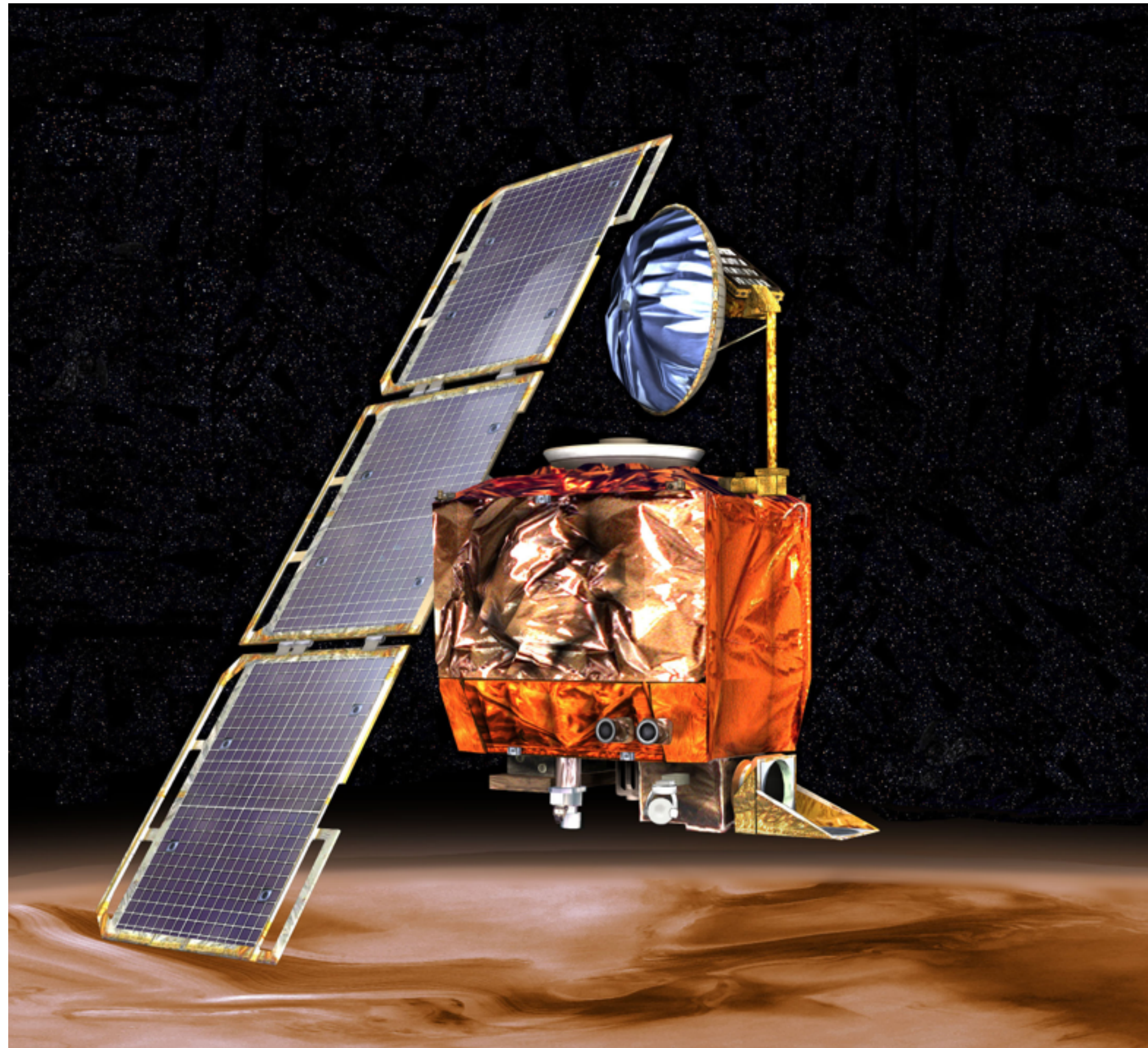


Looping Back on Culture

Describe the problem as your
“enemy” not each other

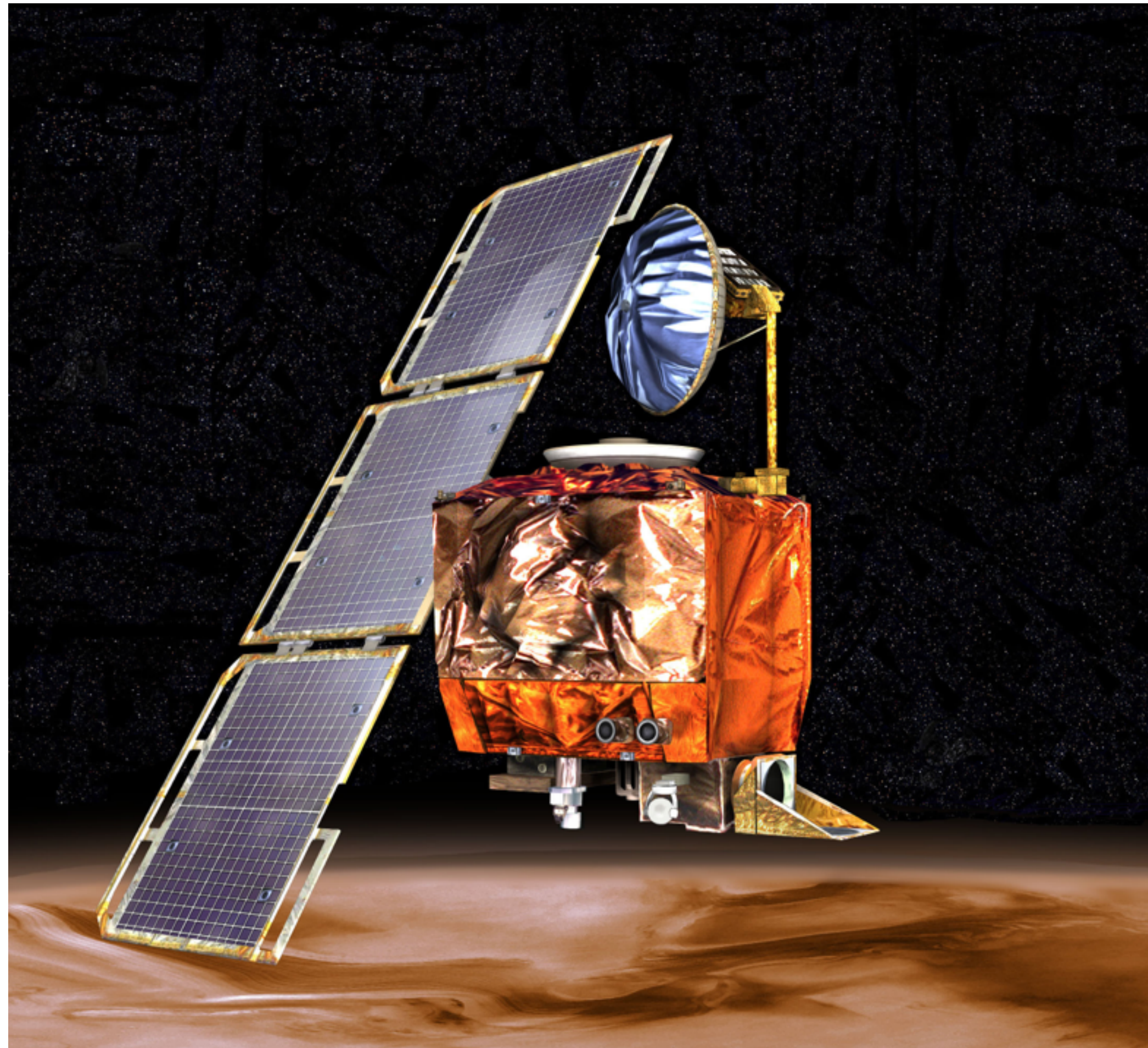
Learn Together

Sharing



Using and Sharing the same metrics and measurements across teams is key to avoiding misunderstandings.

Sharing



Using and Sharing the same metrics and measurements across teams is key to avoiding misunderstandings.

Our Focus Area

Culture
Automation

Metrics
Sharing



Damon Edwards and John Willis
DevOps Day LA

**Collecting data is cheap;
not having it when you
need it can be expensive**

Instrument all the things!



You're in the cloud and it's everything you dreamed of!



Managed
Databases



Autoscaling



Container
Orchestration



Infinite Storage



Private Clouds

The background features a gradient from yellow to orange to red. It is populated with several white storks carrying large, 3D Docker container icons. The containers are white with blue accents and are suspended by their top handles. The text is centered in the middle of the image.

8 SURPRISING FACTS ABOUT REAL
DOCKER
ADOPTION

More info at: www.datadoghq.com/docker-adoption/



Honest Status Page

@honest_update

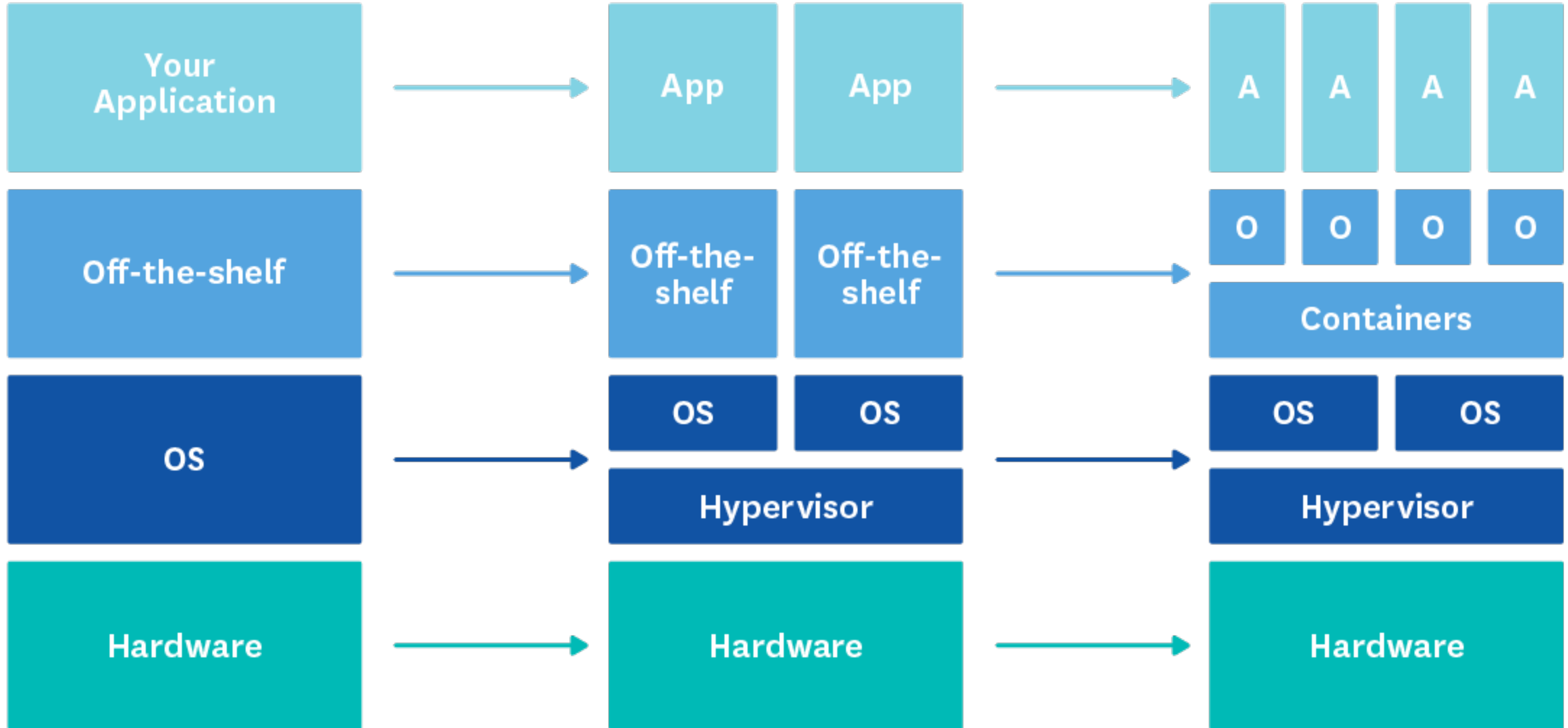


Following

Between containers, cluster managers and virtual machines we've lost track of where our code even is. [#inception](#)

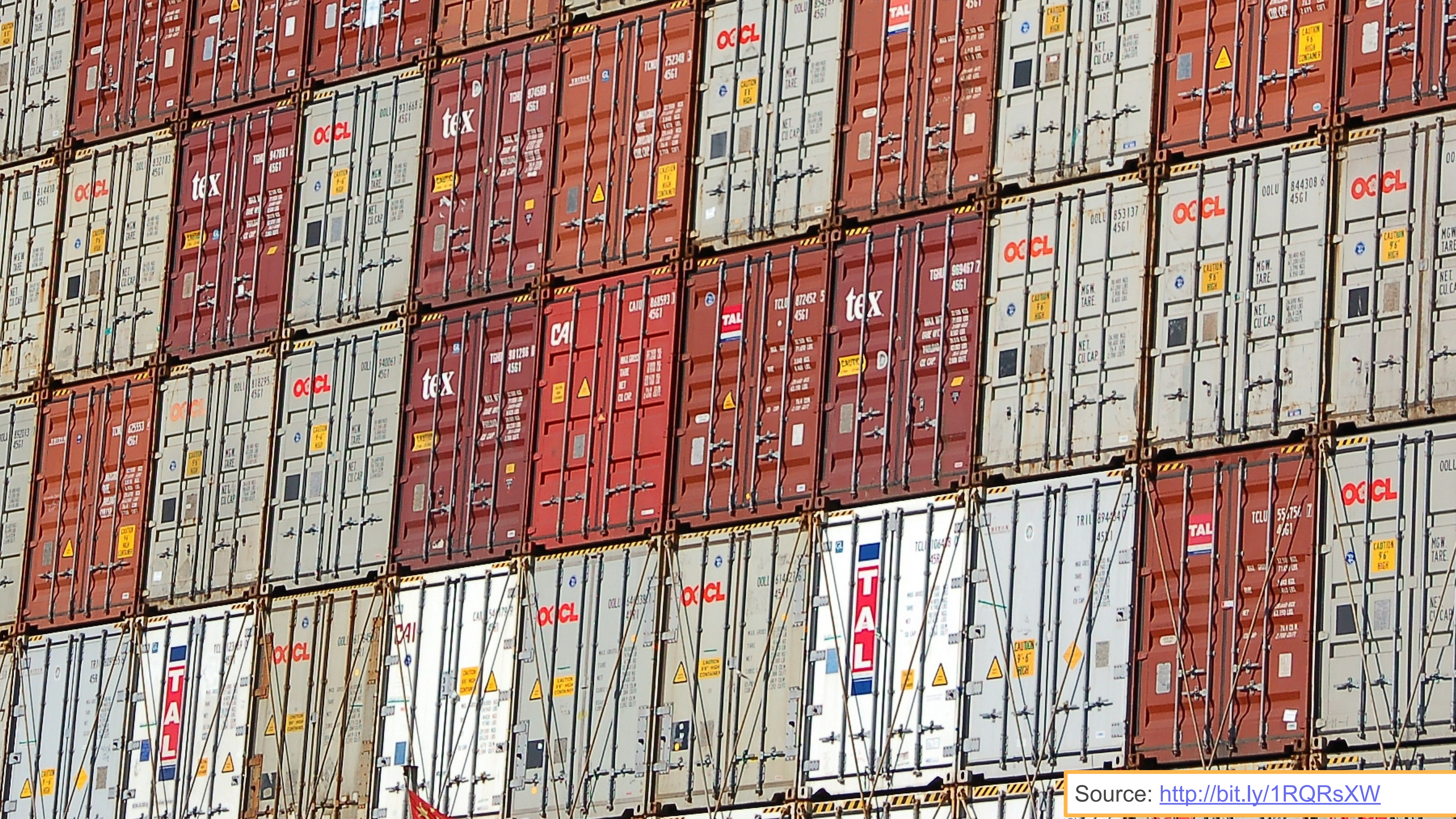
A = Application

o = Off-the-shelf Component





Source: <http://bit.ly/1SvbuP>



Source: <http://bit.ly/1RQRsXW>

Operational Complexity Increases with..

- Number of things to measure
- Velocity of change

How much we measure?

1 instance

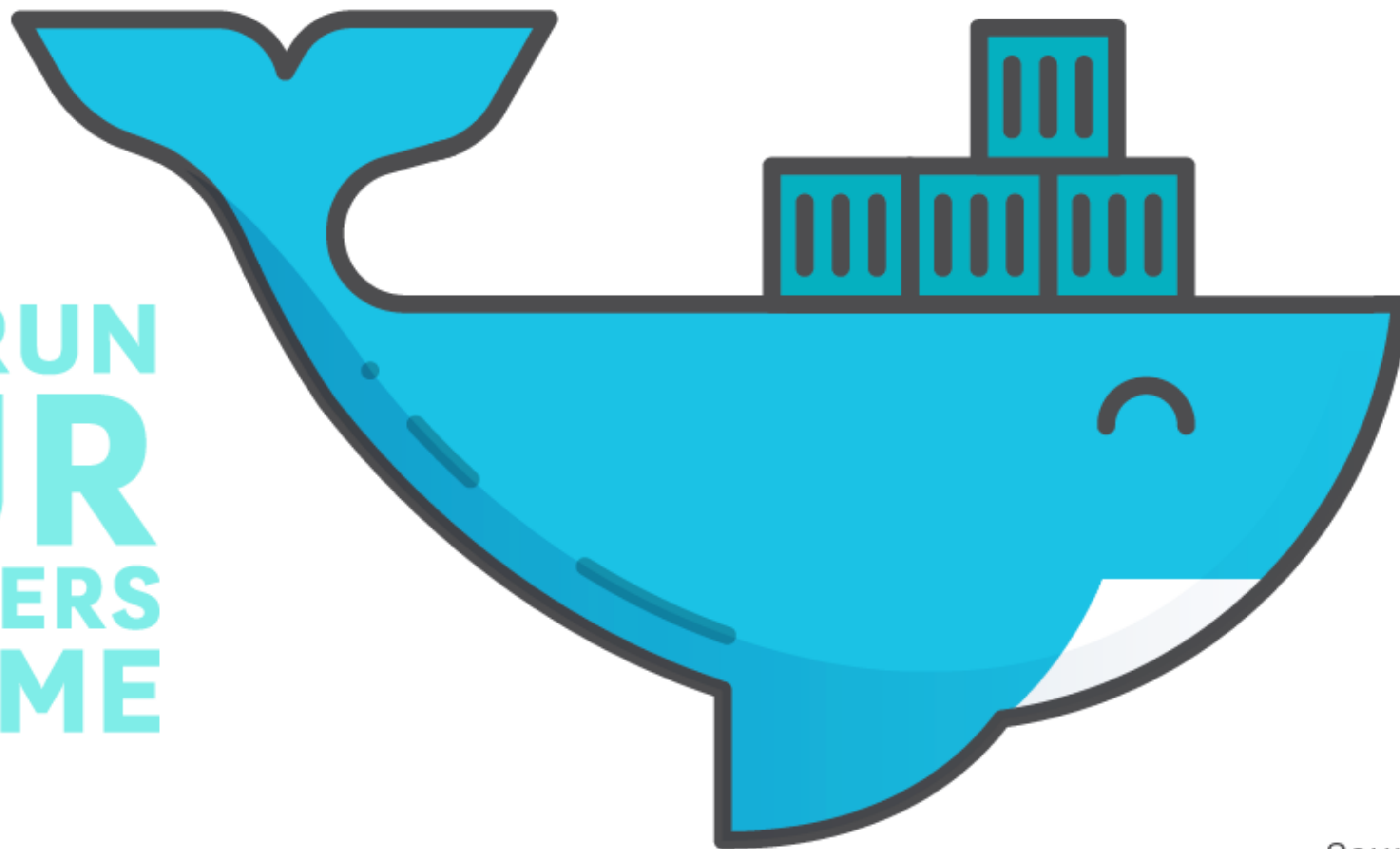
- 10 metrics from CloudWatch

1 operating system (e.g., Linux)

- 100 metrics

50~ metrics per application

HOSTS RUN
FOUR
CONTAINERS
AT A TIME



Source: Datadog

How much we measure?

1 instance

- 10 metrics from CloudWatch

1 operating system (e.g., Linux)

- 100 metrics

50~ metrics per application

N containers

- $150 * N$ metrics

Operational Complexity

100

instances



400

containers

Operational Complexity: Scale

160

metrics per host



640

metrics per host

Assuming 4 containers per host

Operational Complexity: Scale

100

instances



64,000

metrics

Assuming 4 containers per host

How much we measure?

1 instance

- 10 metrics from application

1 operating system (e.g. Linux)

- 100 metrics

50~100 metrics per application

N containers

- $150 * N$ metrics

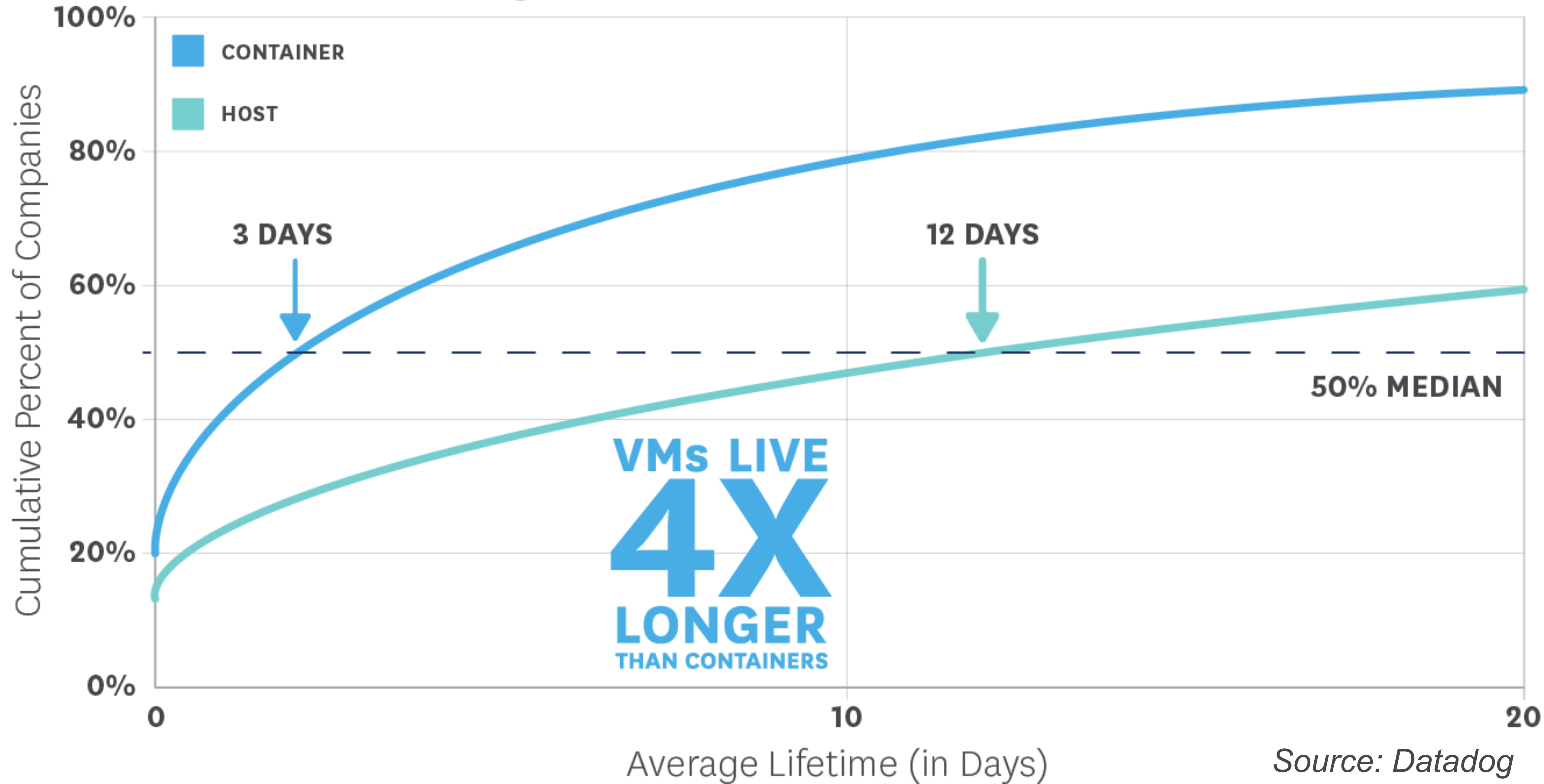
Metrics Overload!

Operational Complexity Increases with..

- Number of things to measure
- Velocity of change



Average Lifetimes of Hosts and Containers



Source: Datadog

Host half-life

hours,
days,
months



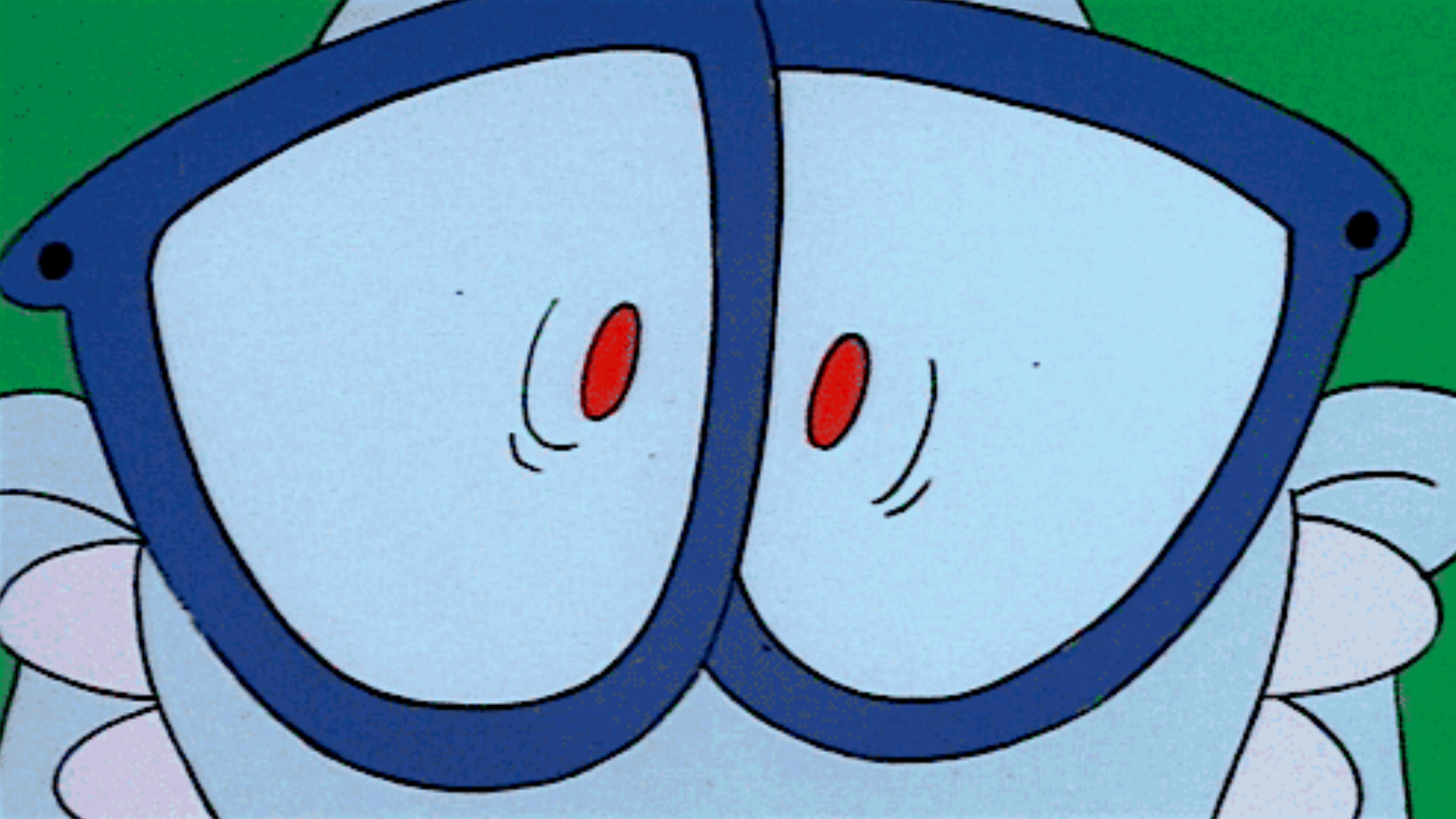
Container half-life

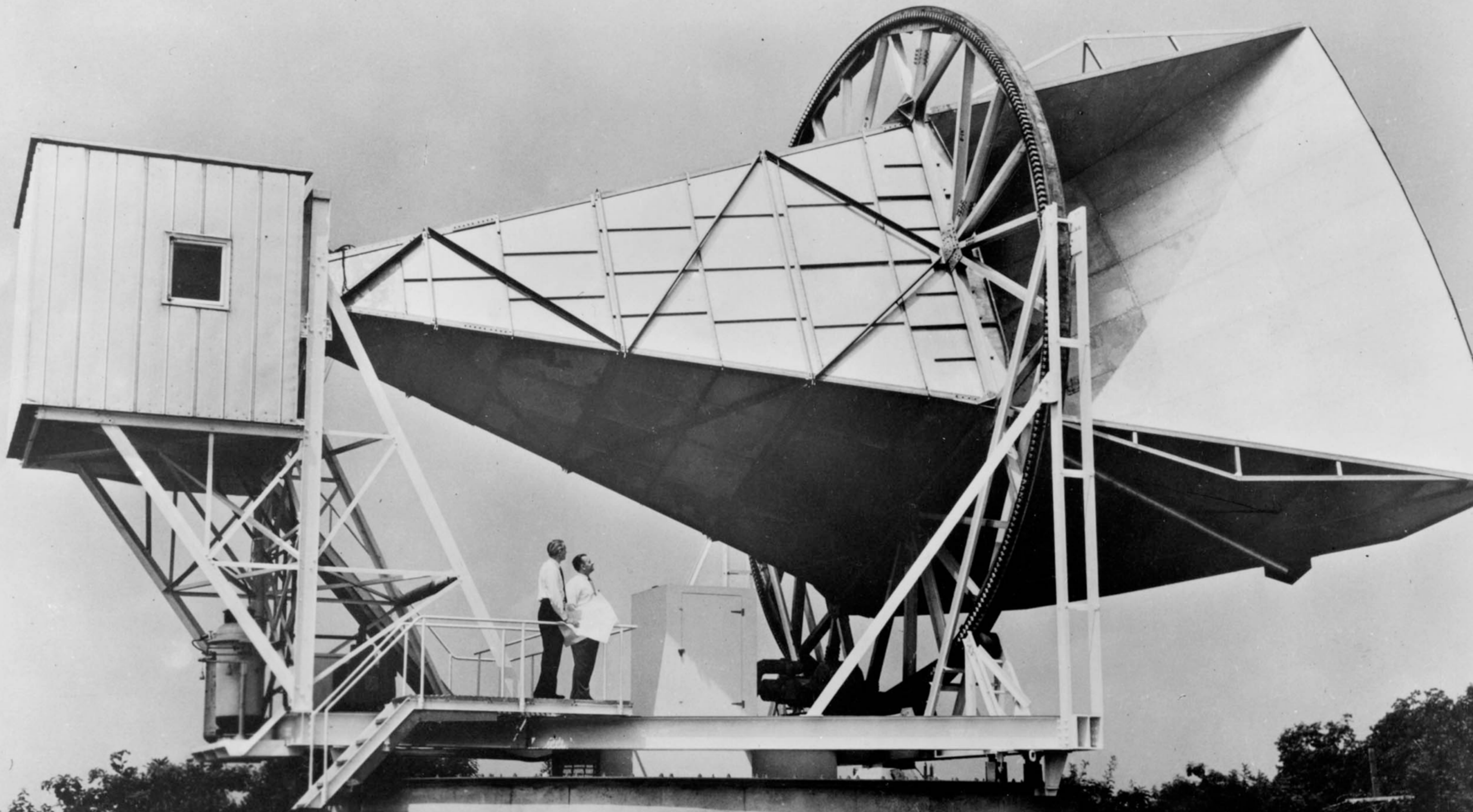
minutes,
hours,
days

Operational Complexity Increases with..

- Number of things to measure
- Velocity of change







A stylized night landscape illustration. In the foreground, a hot air balloon with a green and yellow checkered pattern floats on the left. The background features a dark blue sky with a crescent moon and stars, a range of pink mountains, and a dark blue lake with green lily pads and reeds. A white rectangular box with a light blue gradient is centered over the scene, containing the main title and a breadcrumb trail.

Monitoring 101: Alerting on what matters

series / theory / alerting / monitoring / monitoring-101

More Details at: <http://www.datadoghq.com/blog/monitoring-101-alerting/>

Finding Signal - Categorizing Your Metrics

WORK METRICS

RESOURCE METRICS

EVENTS



THE COMMONWEALTH TRUST (INDIA) LTD

COMMONWEALTH TRUST

WORK METRICS

THROUGHPUT

SUCCESS

ERROR

PERFORMANCE



RESOURCE METRICS

UTILIZATION

SATURATION

ERROR

AVAILABILITY



EVENTS

CODE CHANGES

ALERTS

SCALING EVENTS

ETC



Examples: NGINX - Metrics

Work Metrics:

- **Requests Per Second**
- **Dropped Connections**
- **Request Time**
- **Error Rates (4xx or 5xx)**
- **Success (2xx)**

Resource Metrics:

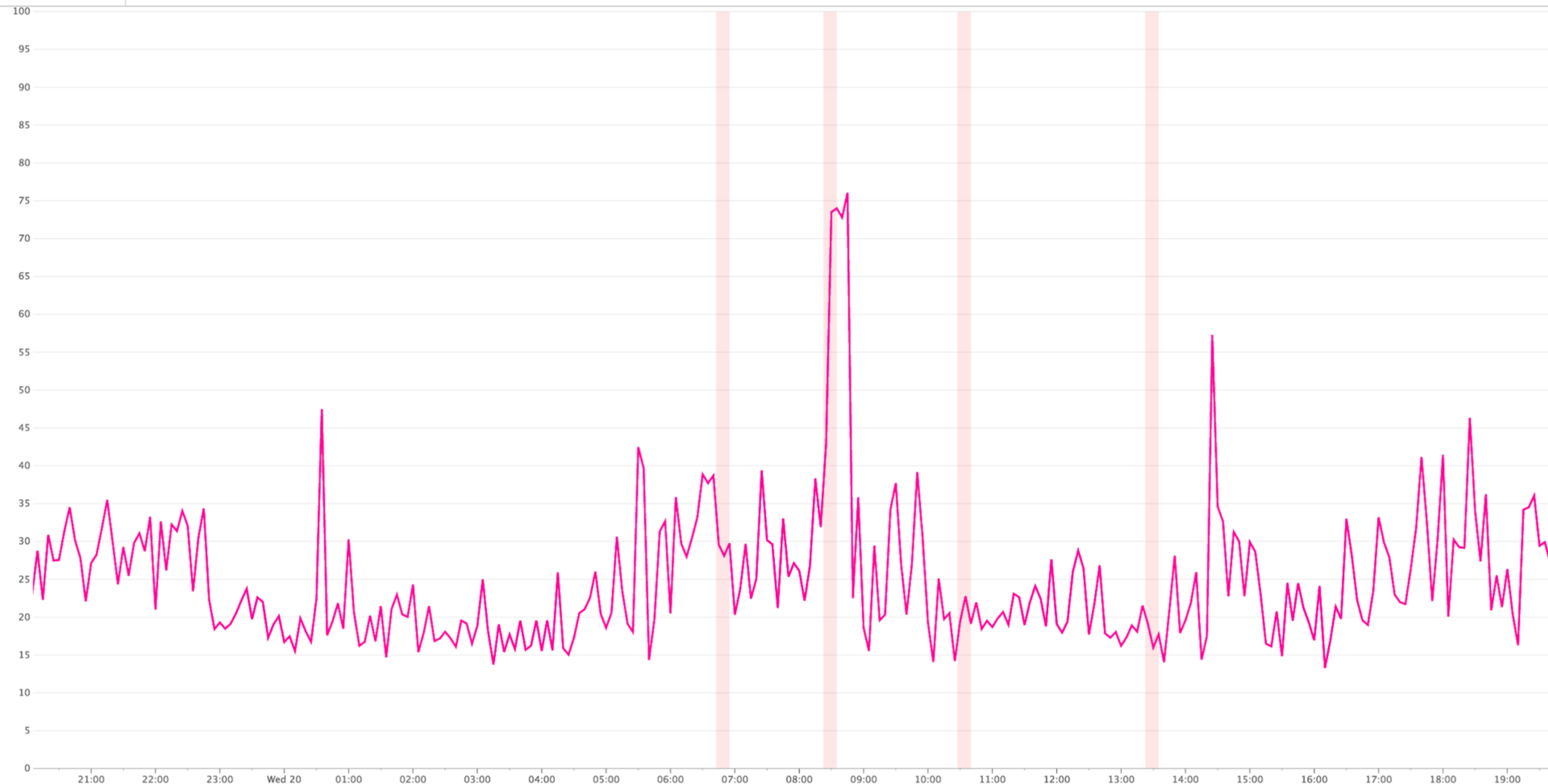
- **Disk I/O**
- **Memory**
- **CPU**
- **Queue Length**

Examples: NGINX - Events

- **Configuration Change**
- **Code Deployment**
- **Service Started / Stopped**
- **etc**

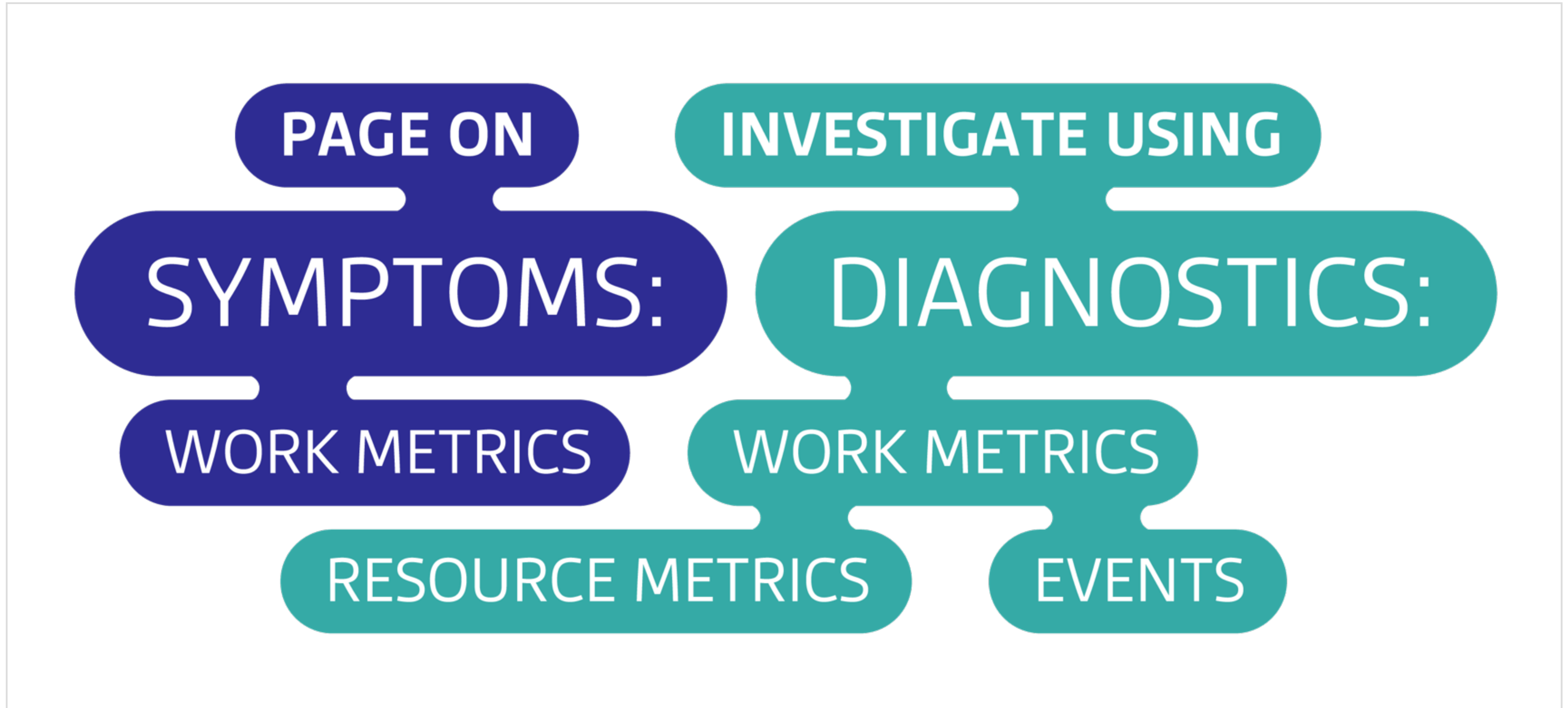
I/O wait (%)

Show 1d The Past Day



**When to let a sleeping
engineer lie?**

When to alert?





Adrian Cole

@adrianfcoble

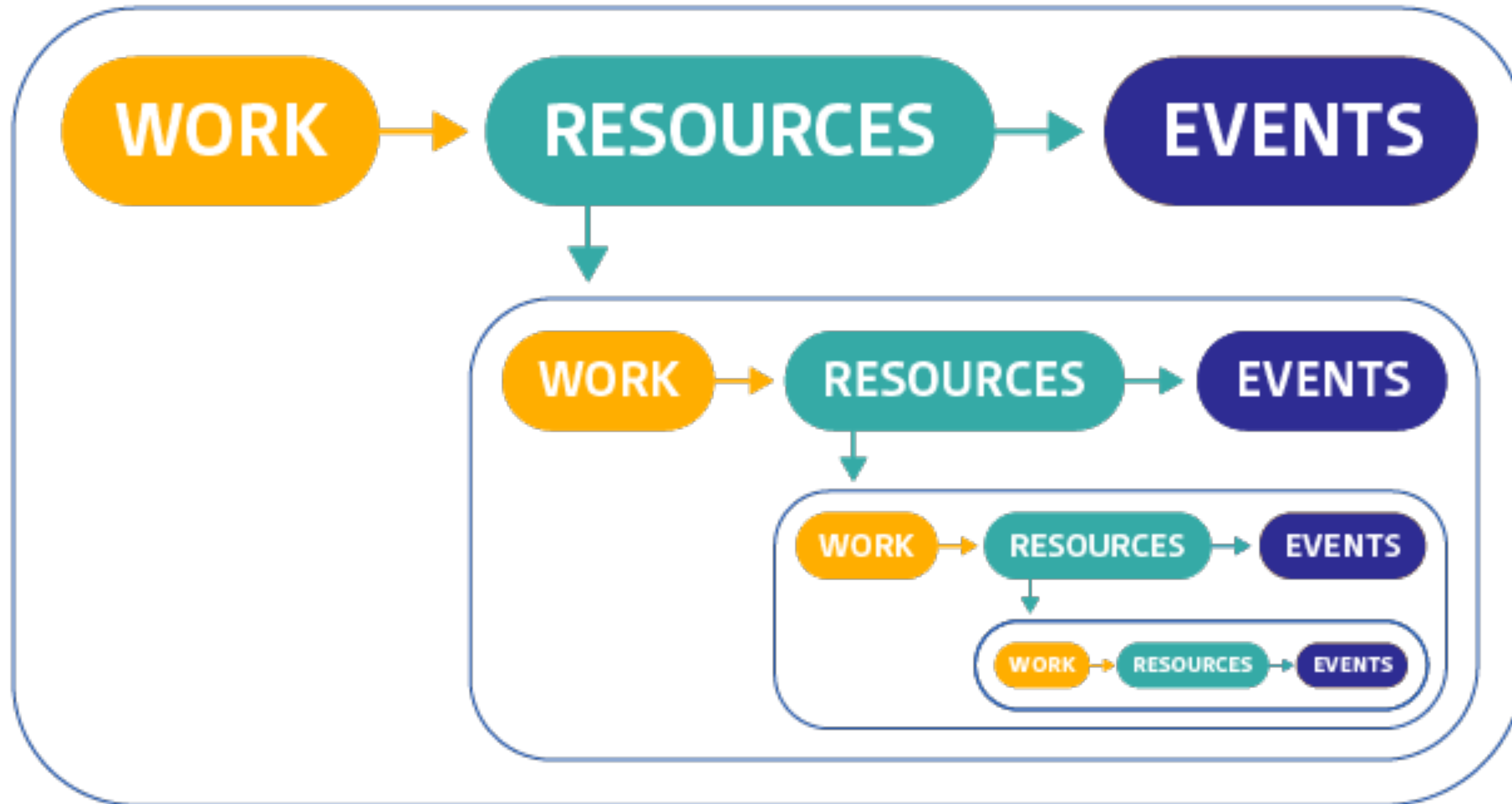


Following

Q: Are we losing money?

A: Can't answer that, but I can tell you what average CPU usage was 5ish mins ago..

Recurse until you find root cause



How does your current monitoring

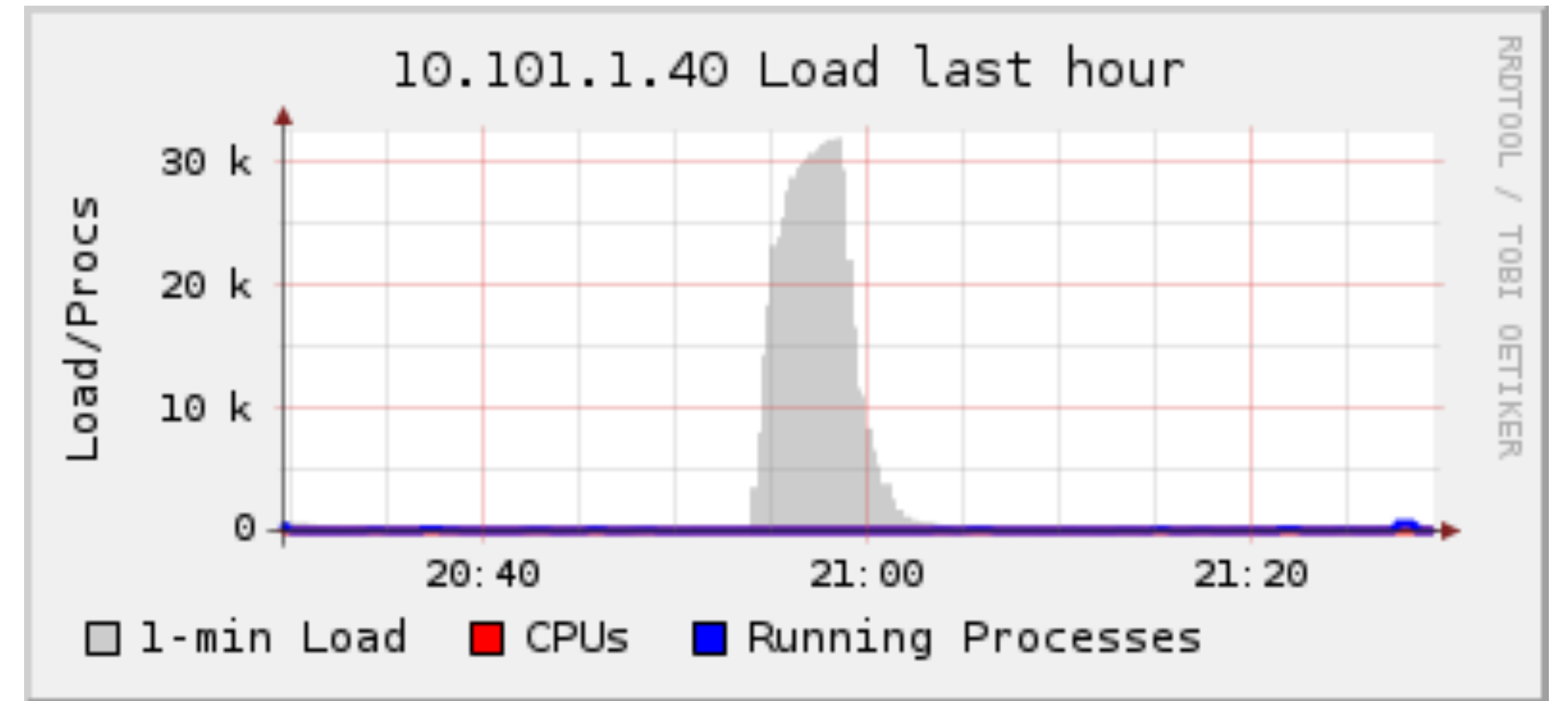


Too Many Tools

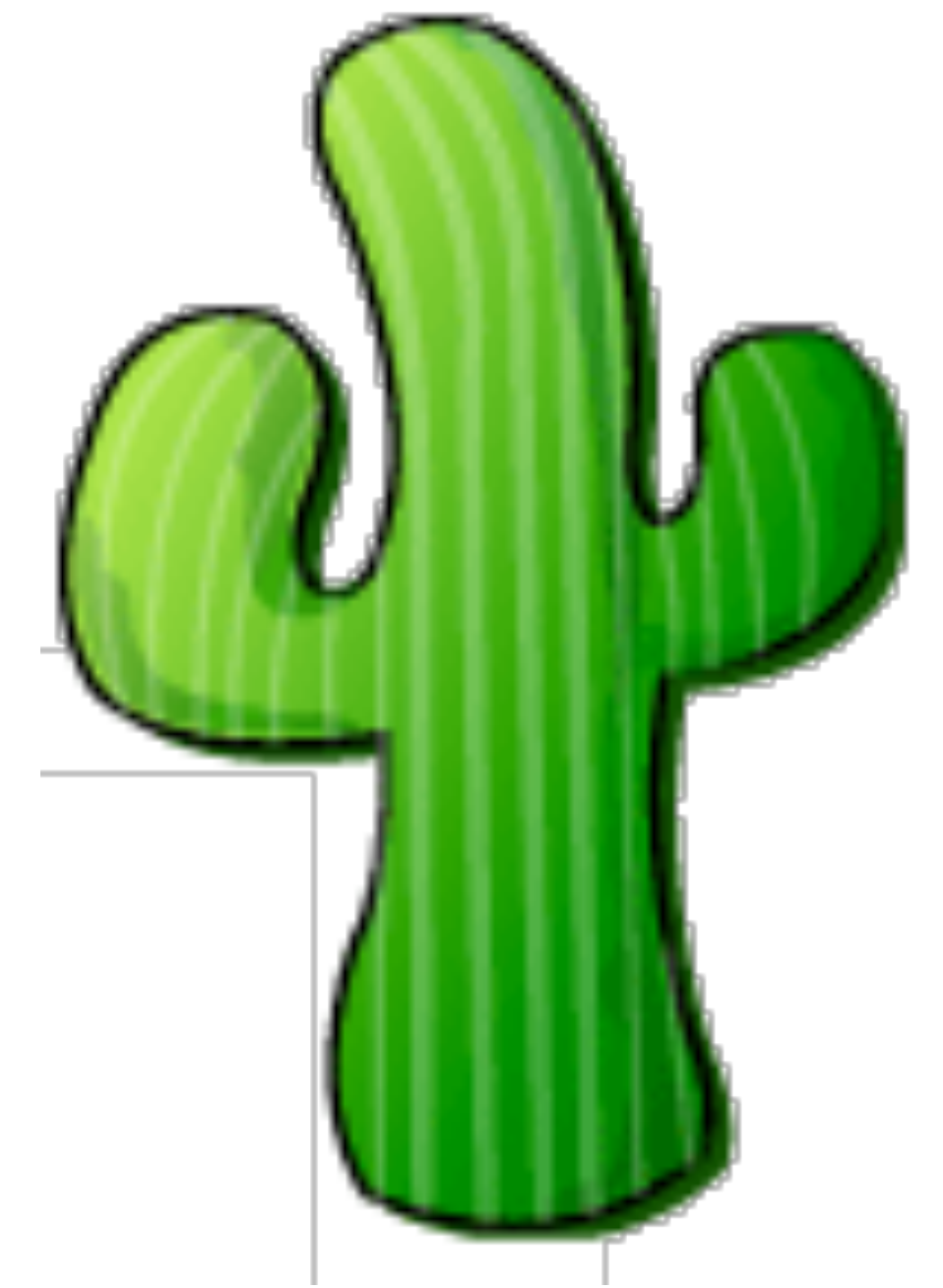
Trending vs Alerting

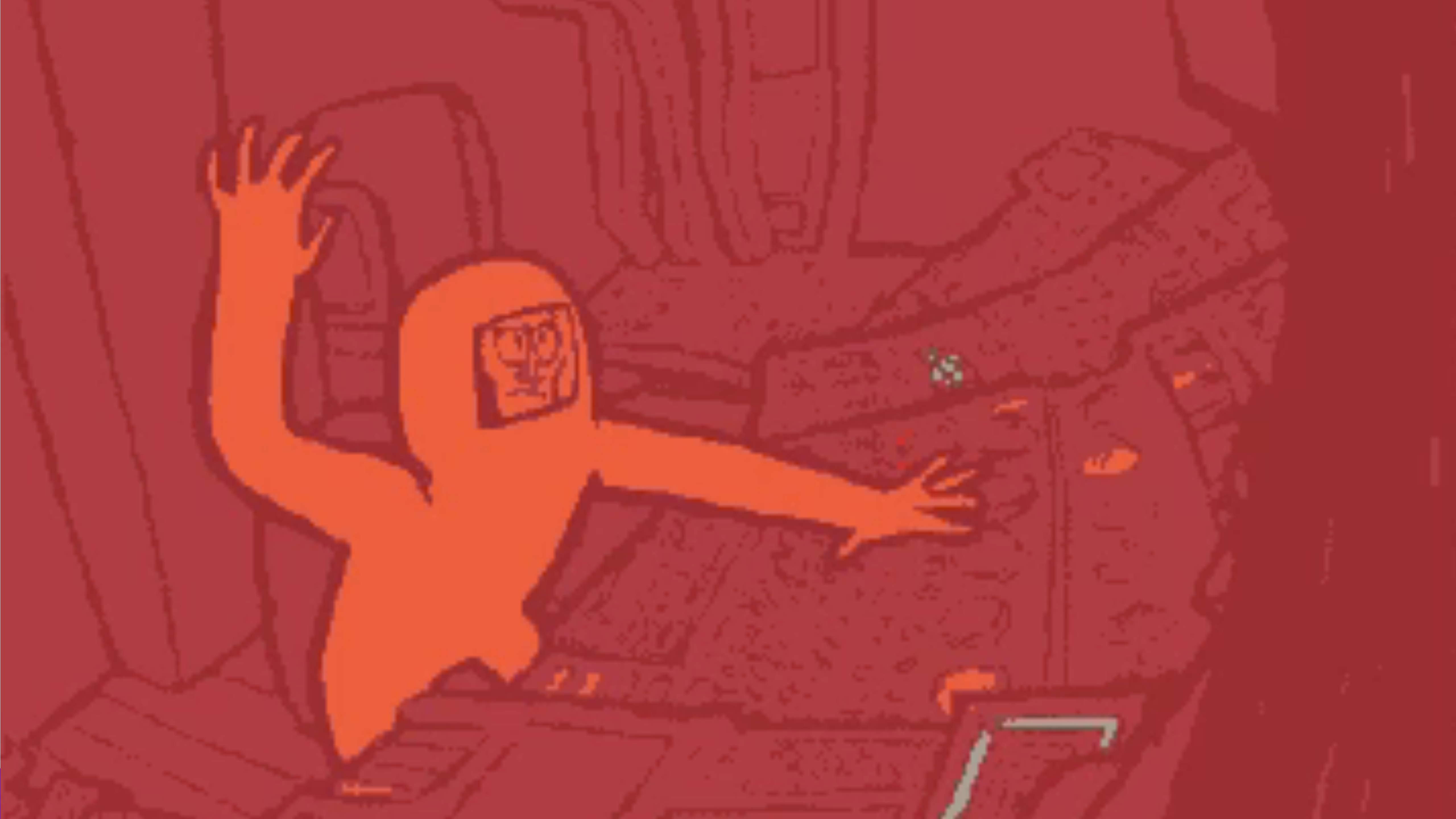
Many Point Solutions

How do they all fit together?



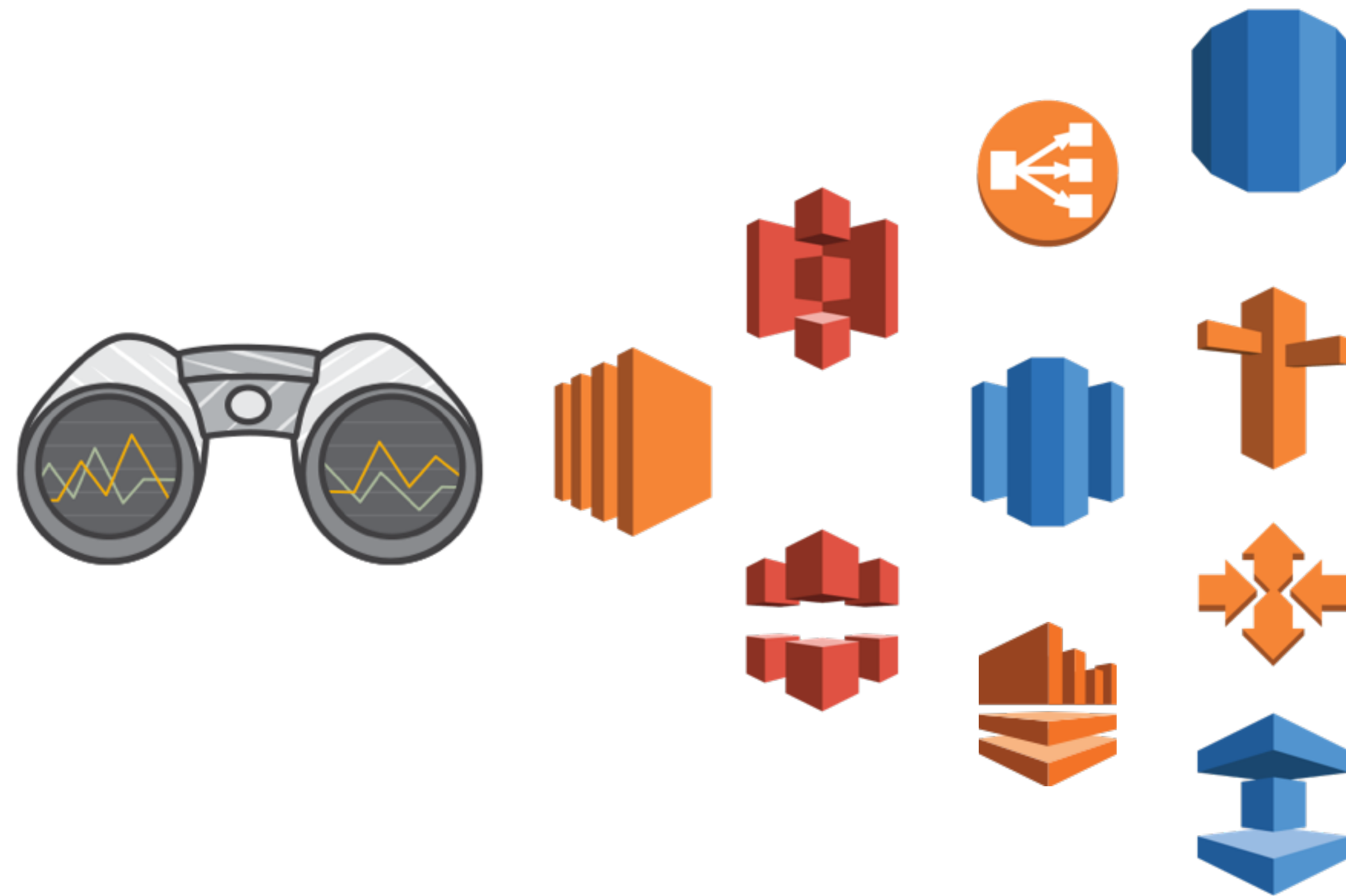
Nagios®





Too Many Tools

Pick tools that let you aggregate many data sources



Cryptic Alerts

Date: January 1, 2011 11:06:41 AM GMT+09:00
From: Nagios Monitoring user
Subject: **** PROBLEM Host Alert: dbserver1 is DOWN ****
To: support@frank4dd.com

***** Nagios *****

Notification Type: PROBLEM
Host: dbserver1
State: DOWN

Info: CRITICAL - Host Unreachable

Date/Time: Sat Jan 1 11:06:41 JST 2011

Cryptic Alerts

Date: January 1, 2011 11:06:41 AM GMT+09:00
From: Nagios Monitoring user
Subject: **** PROBLEM Host Alert: dbserver1 is DOWN ****
To: support@frank4dd.com

***** Nagios *****

Notification Type: PROBLEM
Host: dbserver1
State: DOWN

Info: CRITICAL - Host Unreachable

Date/Time: Sat Jan 1 11:06:41 2011

WHAT?

3 Say what's happening

 Preview

 Edit

Home Page Response Time is Above 3s

Our home page is taking more than 5s to load. Our customers generally get board and abandon at 5s. Lets investigate before users become frustrated.

- Do we need more capacity? check web tier resource utilization on this dashboard (<http://.....>)
- Spin up more resources with this [jenkins job](http://jenkins.com/job/deploy-it-all/)(<http://jenkins.com/job/deploy-it-all/>)
- Check this wiki page for more ideas.
- Still stuck? Call Jenny! 867-5309

Informative and Actionable Alerts

Why is this important?

What do I do about it?

Who do I call next if I get stuck?

EVERY ALERT MUST BE ACTIONABLE

Averages Are Lies

You can't provision for your average traffic.

Keep the real data.

There are **3** kinds of lies:



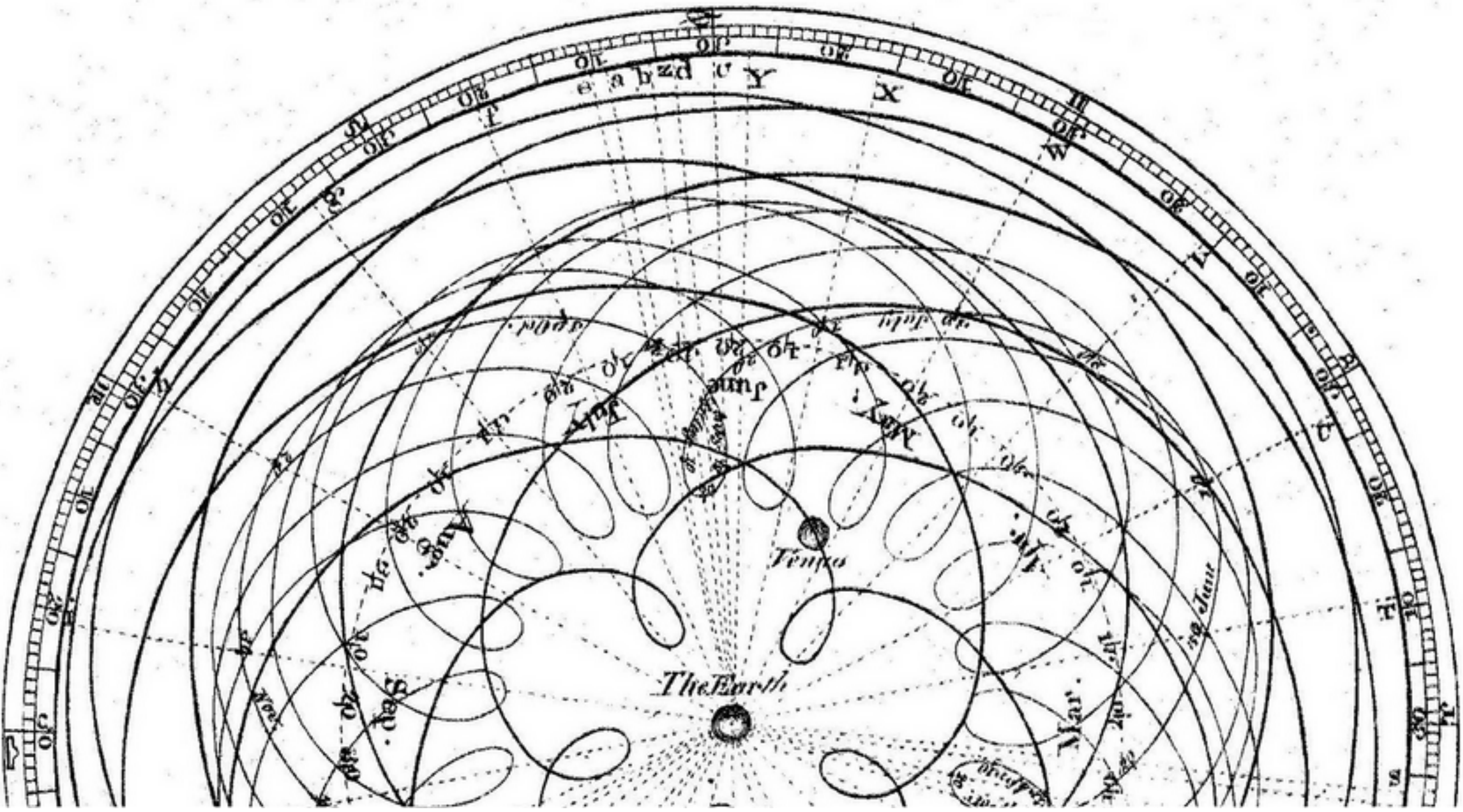
Static vs Dynamic

Static configurations tracking dynamic infrastructure

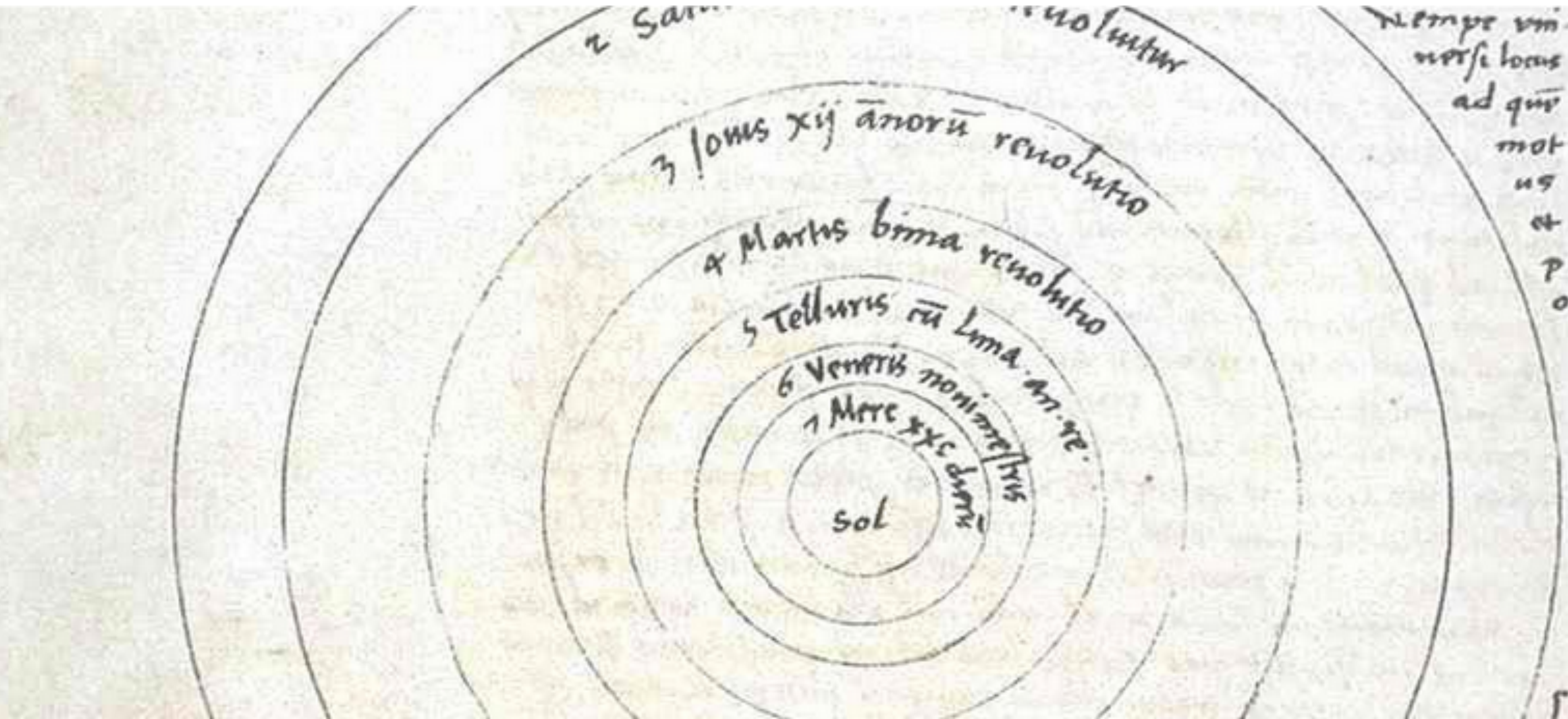
```
1 host_name snowflake1
2 address 192.168.1.13
3 check_command check-host-alive
4 max_check_attempts 10
5 notification_interval 120
6 notification_period 24x7
7 notification_options d,u,r
8 }
```



Host Centric



Service Centric



Tags All the Way Down

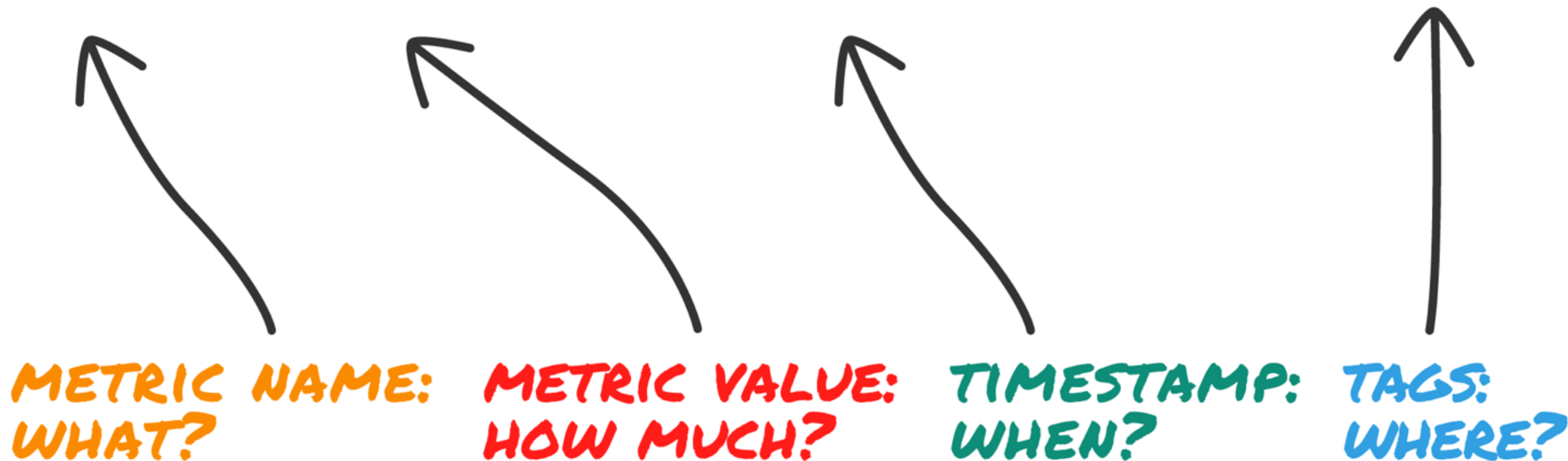
Launch Instance **Connect** **Actions** ▾

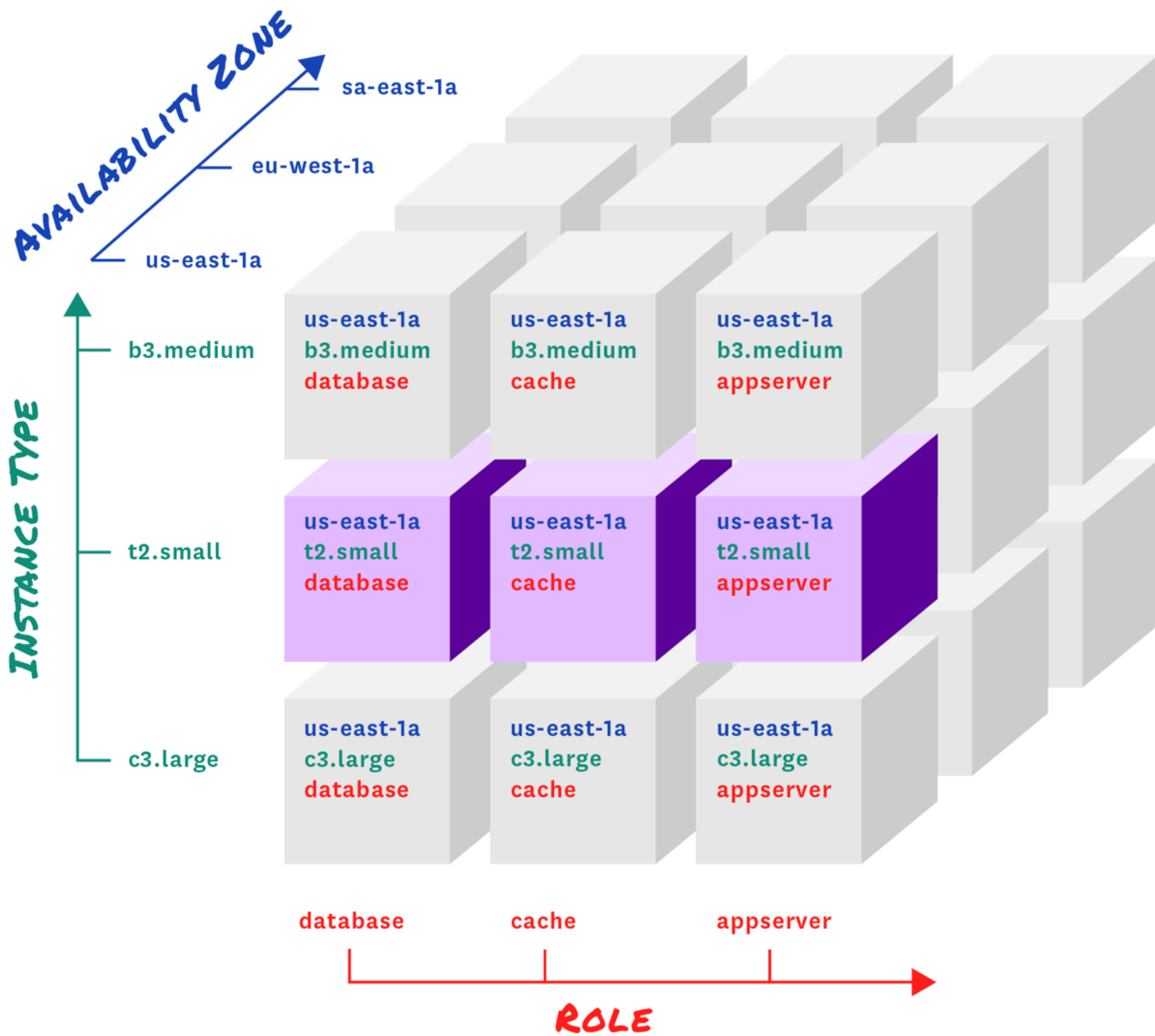
🔍 Filter by tags and attributes or search by keyword

<input type="checkbox"/>	Name ▾	Instance ID ▾	Instance Type ▾	Availability Zone
<input type="checkbox"/>	m3.medium	us-east-1a
<input type="checkbox"/>	r3.large	us-east-1c
<input type="checkbox"/>	m3.xlarge	us-east-1a
<input type="checkbox"/>	t2.micro	us-east-1a

DATAPPOINT

SYSTEM.NET.BYTES_RCVD 4 2016-03-02 15:00:00 [FILE-SERVER]





Asking Better Questions

“Monitor all containers running image web in region us-west-2 across all availability zones that use more than 1.5x the average memory on c3.xlarge”

Asking Better Questions

“90% of all web requests are taking more than 0.5s to process and respond.”

Custom Metrics

- Instrument custom applications
- You know your key transactions best.
- Use async protocols like Elys' STATSD



STATSD



DATADOG



Source: <http://bit.ly/1NoW6aj>

Resources

[Monitoring 101: Alerting](https://www.datadoghq.com/blog/monitoring-101-alerting/)

<https://www.datadoghq.com/blog/monitoring-101-alerting/>

[Monitoring 101: Collecting the Right Data](https://www.datadoghq.com/blog/monitoring-101-collecting-data/)

<https://www.datadoghq.com/blog/monitoring-101-collecting-data/>

[Monitoring 101: Investigating performance issues](https://www.datadoghq.com/blog/monitoring-101-investigation/)

<https://www.datadoghq.com/blog/monitoring-101-investigation/>

[The Power of Tagged Metrics](https://www.datadoghq.com/blog/the-power-of-tagged-metrics/)

<https://www.datadoghq.com/blog/the-power-of-tagged-metrics/>

[Monitoring Sucks Project](https://github.com/monitoringsucks/)

<https://github.com/monitoringsucks/>