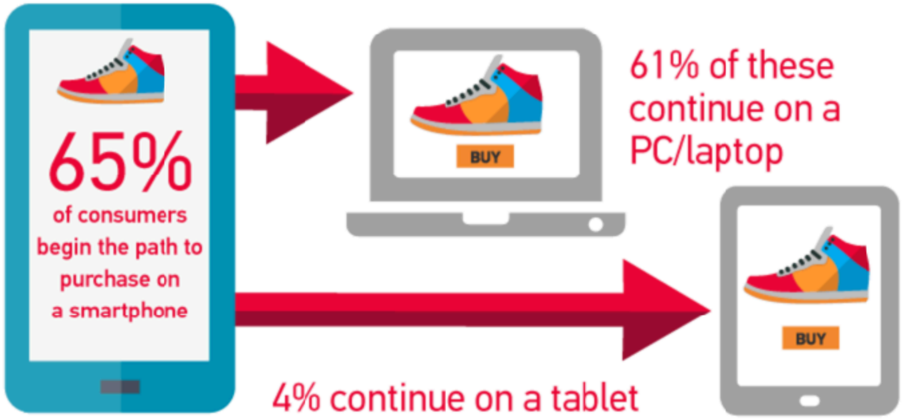
Cassandra @ Signal

How to handle billion requests a day with Cassandra without losing sleep

Aki Colović
Engineering Manager

7.2 devices/person

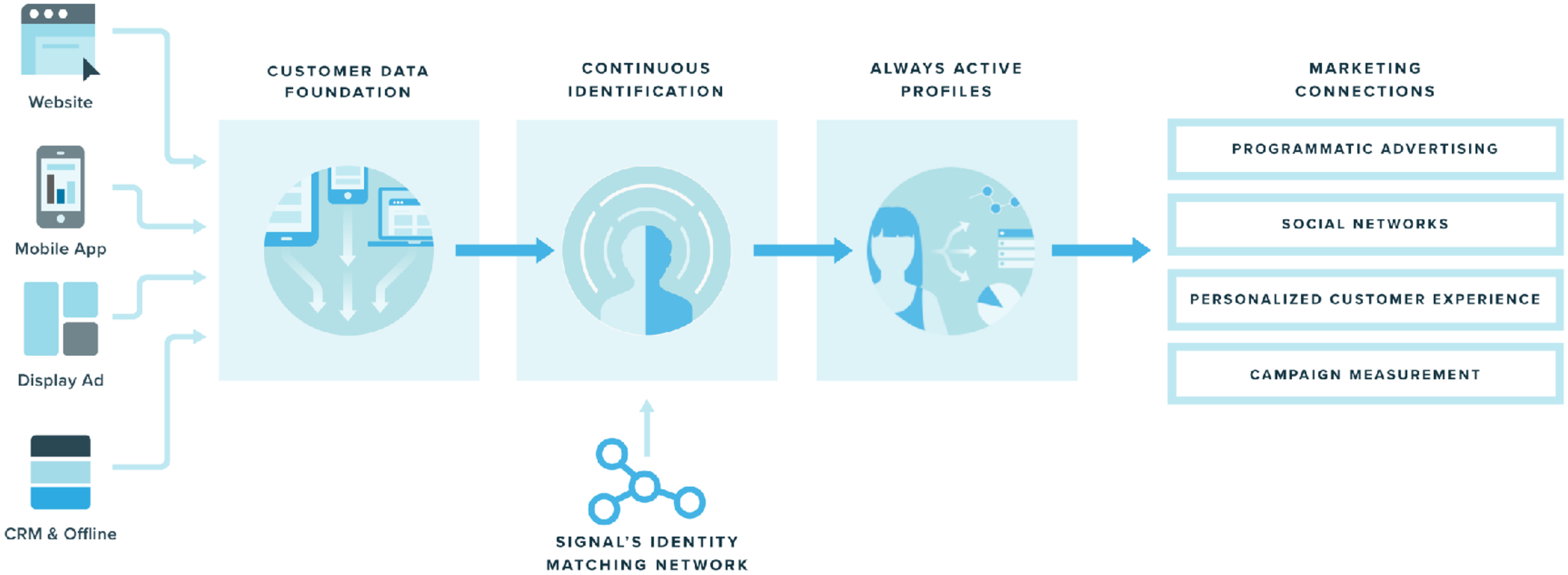
SIGNAL®



They do not have a single view of the customer:

Just 6% of global marketers say they have attained the single view of the customer that is necessary to provide relevant, seamless experiences.

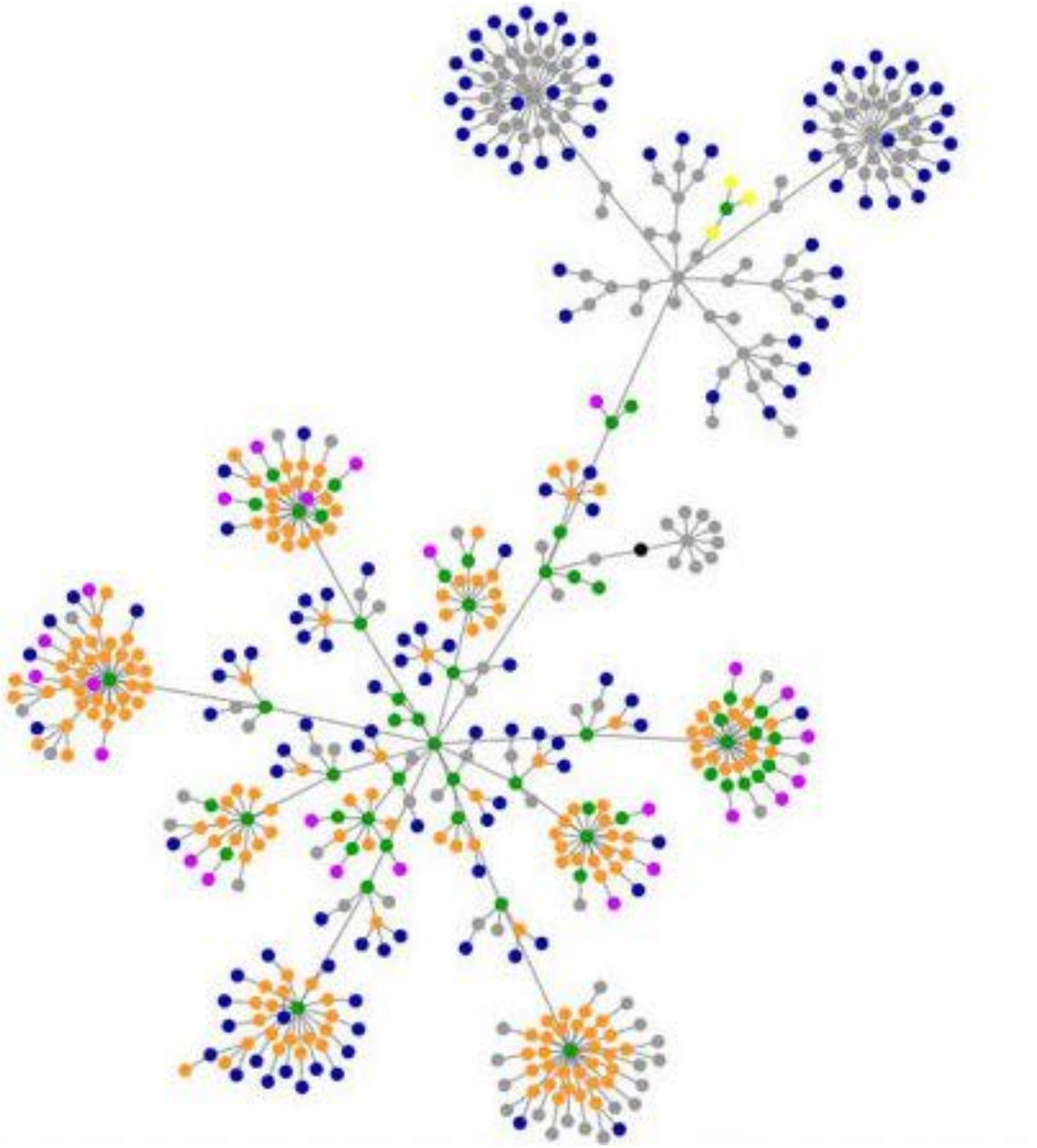
Platform for Digital Marketing



SIGNAL®

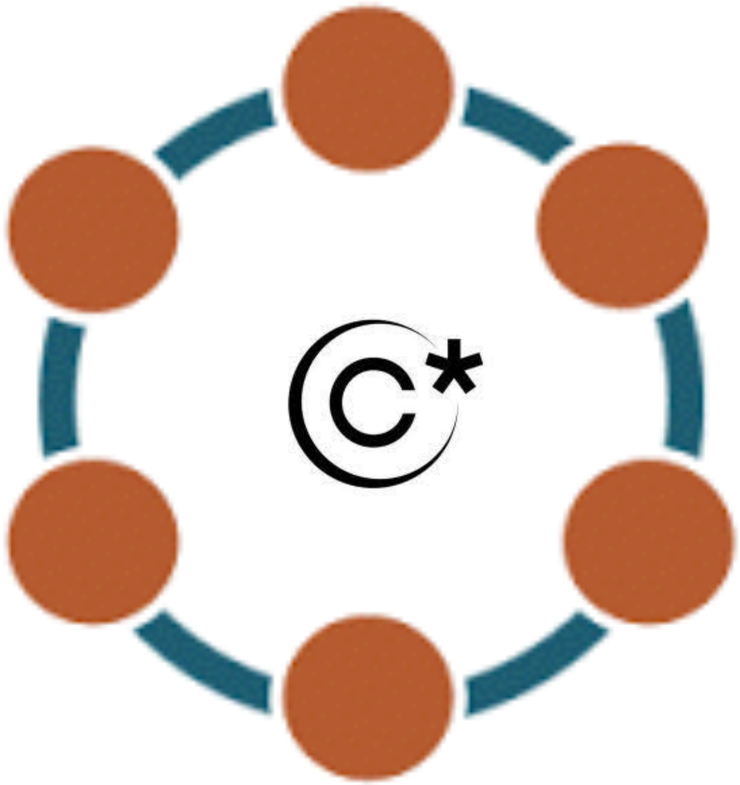
Our Partners

Signal’s Identity Graph



SIGNAL®

Apache Cassandra



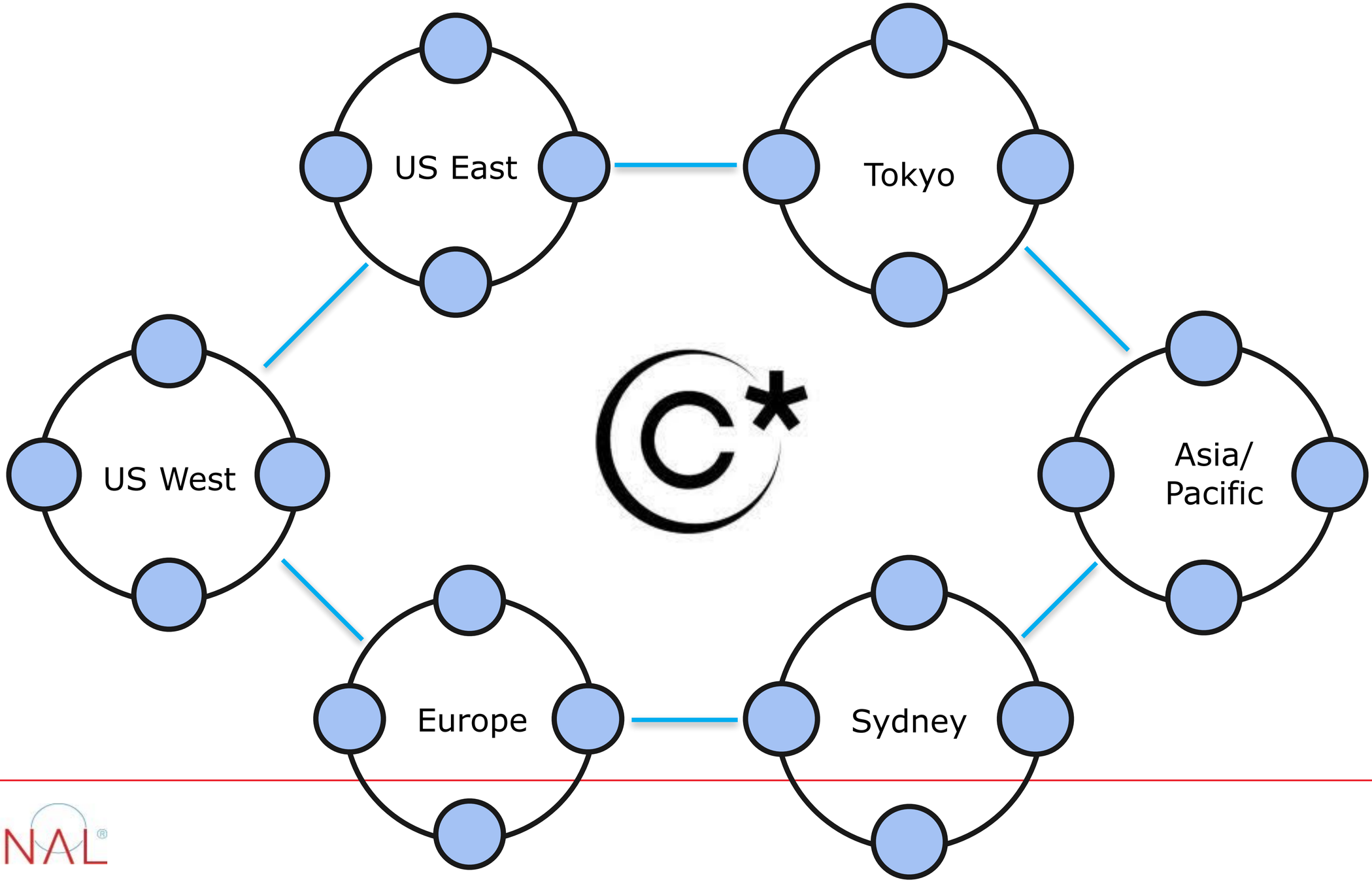
Cassandra @ Signal

500+ Nodes

~1.5 billion queries per day

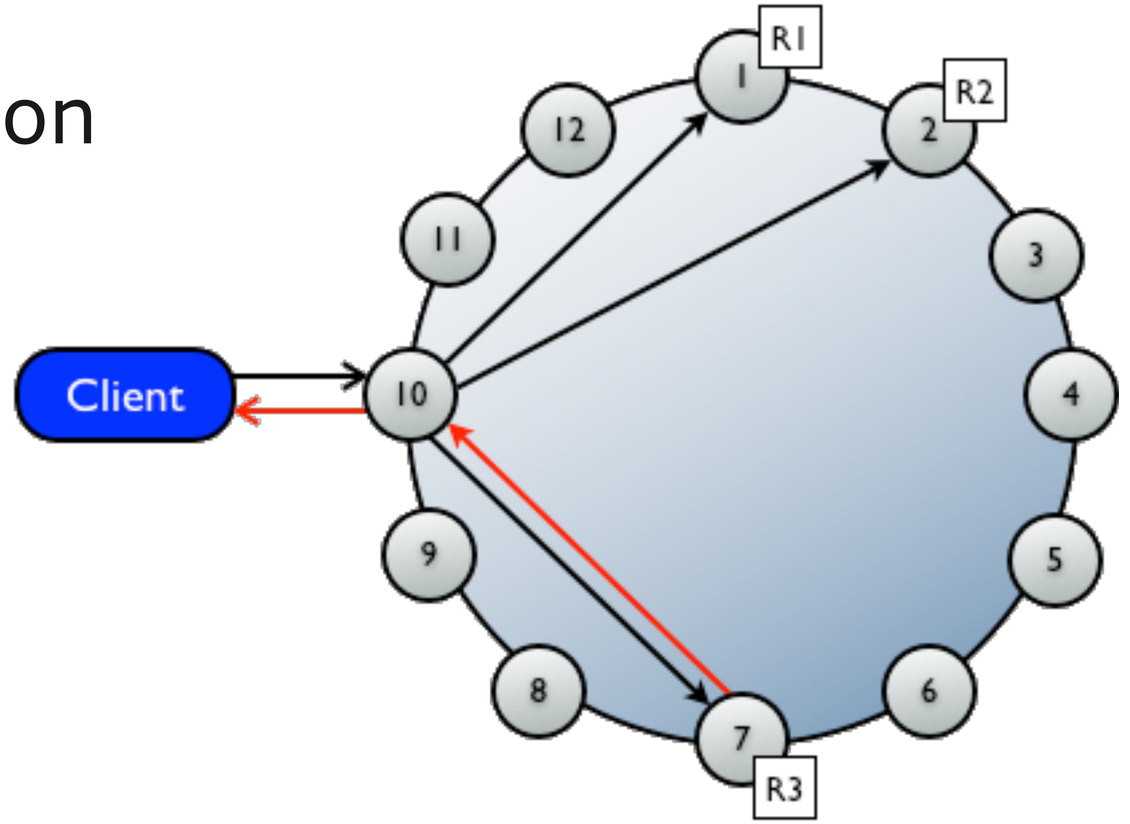
90.96+ TB of data

SIGNAL®


US East
Tokyo
US West
C*
Asia/
Pacific
Europe
Sydney

What is Cassandra DB?

- distributed NoSQL database
- tunable consistency
- asynchronous replication
- decentralized
- high availability
- easy to scale

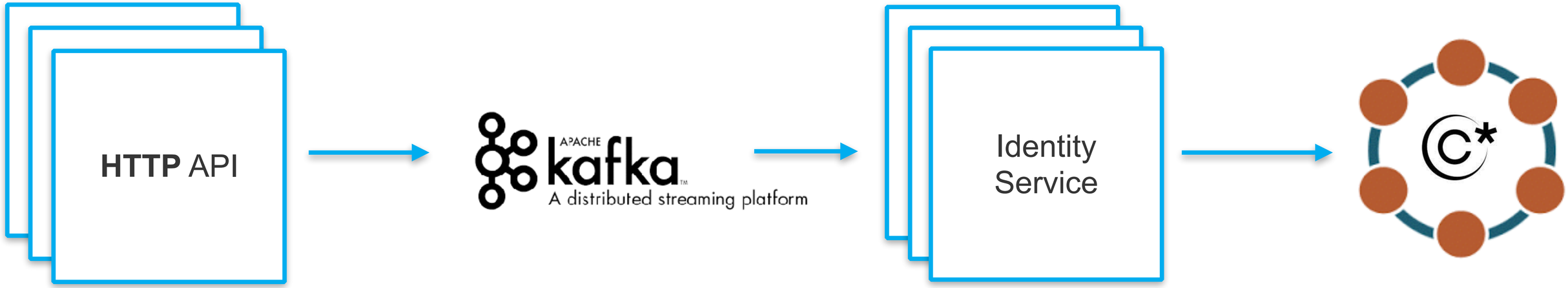


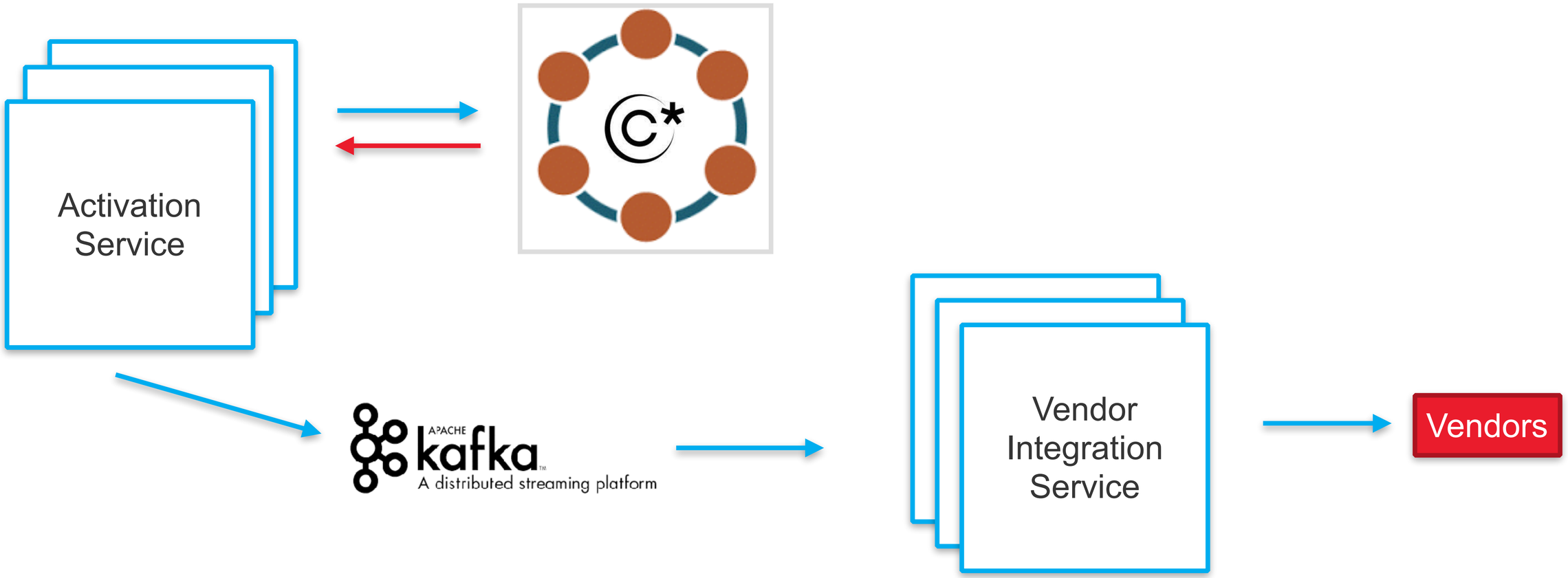
Dealing With Failure

We've survived …

- Amazon swallowing nodes
- Amazon rebooting the internet
- Disk Failures
- Corrupt SSTables
- Intra and inter region network issues

Data Ingestion




Activation Service
C*
APACHE kafka
A distributed streaming platform
Vendor Integration Service
Vendors

What About Downsides?

CQL != SQL

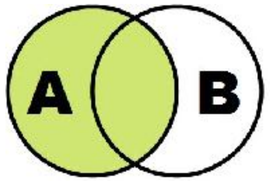
Cassandra Query Language

```
CREATE TABLE profiledata (
  btid uuid,
  identifier text,
  data map<text, text>,
  internal map<text, text>,
  uids map<text, text>,
  PRIMARY KEY (btid, identifier)
) WITH
  bloom_filter_fp_chance=0.010000 AND
  caching='KEYS_ONLY' AND
```

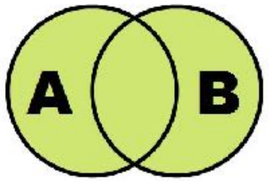
```
SELECT FROM profiledata where btid=1a4b0130-7e0a-4b97-a368-91a5b439f6ae;

UPDATE profiledata_index SET btid=d4a44d4-53f8-495e-8db2-464889375ae3 WHERE entity_prefix ='s'
```

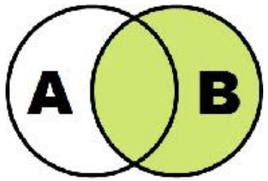
SIGNAL®



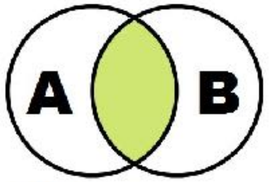

```
SELECT *
FROM A
LEFT JOIN B
ON A.id = B.id
```



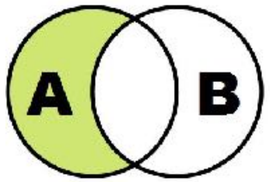

```
SELECT *
FROM A
FULL OUTER JOIN B
ON A.id = B.id
```



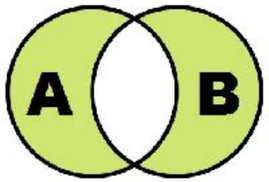

```
SELECT *
FROM A
RIGHT JOIN B
ON A.id = B.id
```



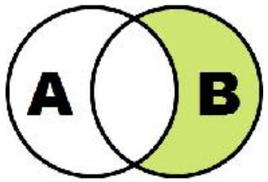

```
SELECT *
FROM A
INNER JOIN B
ON A.id = B.id
```


```
SELECT *
FROM A
LEFT JOIN B
ON A.id = B.id
WHERE B.id IS NULL
```




```
SELECT *
FROM A
FULL OUTER JOIN B
ON A.id = B.id
WHERE A.id IS NULL
OR B.id IS NULL
```

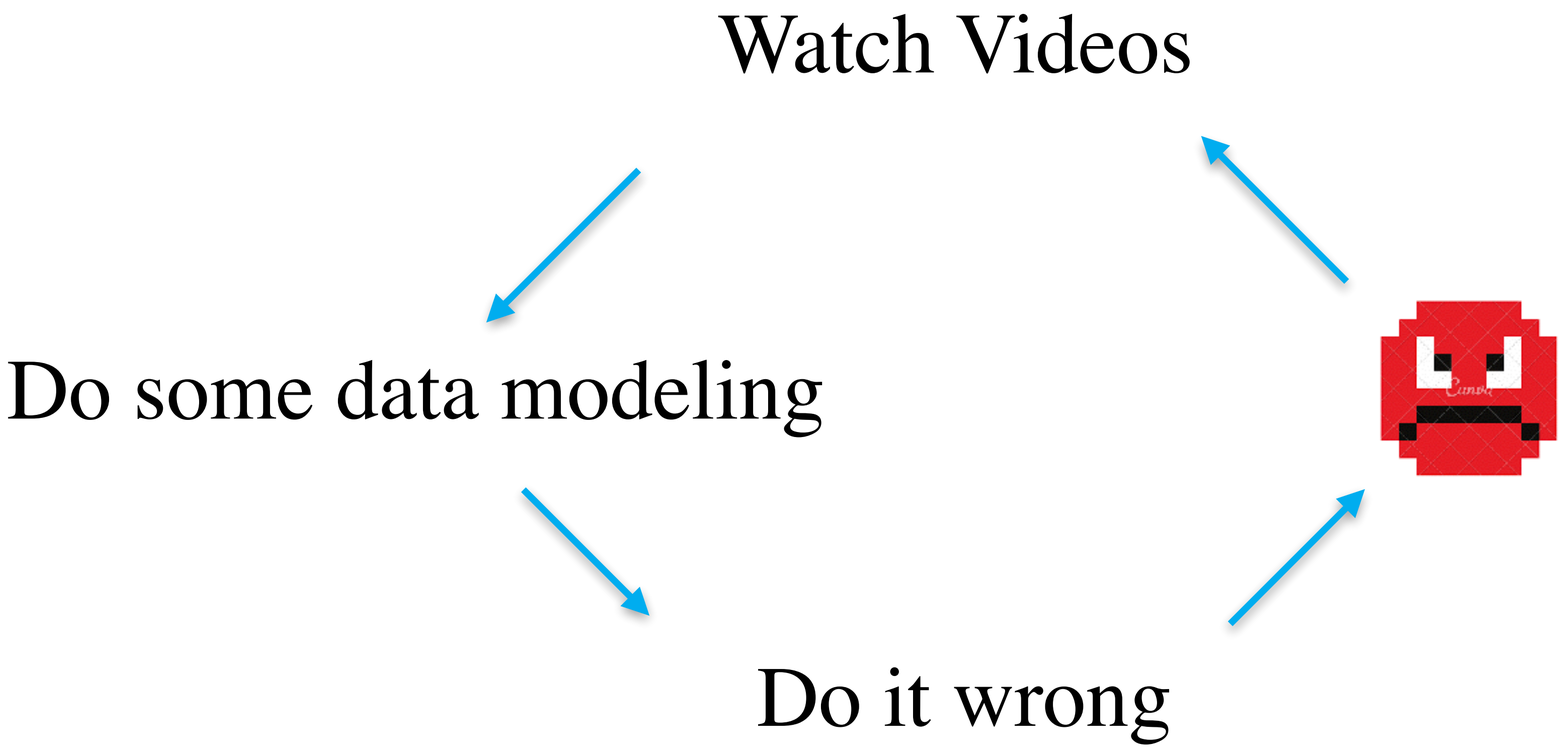



```
SELECT *
FROM A
RIGHT JOIN B
ON A.id = B.id
WHERE A.id IS NULL
```

No Joins, wait what?

SIGNAL®

Data Modelling Pains



Final thoughts on modeling…

"The best way to approach data modeling for Cassandra is to start with your queries and work backwards from there. Think about the actions your application needs to perform, how you want to access the data, and then design column families to support those access patterns."

Secondary Index

- works well for low cardinality data
- does not scale
- material views(c* 3.0+)

Deleting is painful at scale

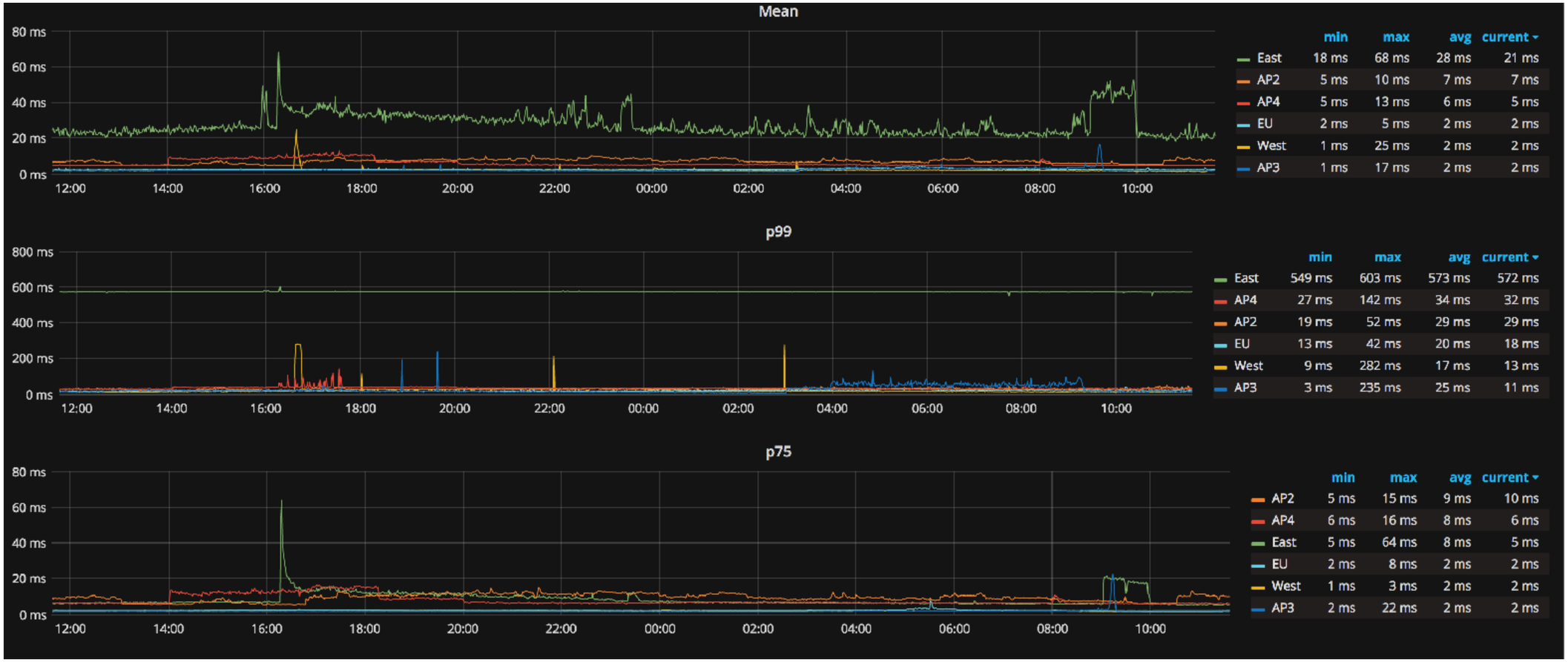
- tombstones
- disk not reclaimed immediately
- increased I/O pressure
- sometimes big expensive scans needed

Testing at Scale

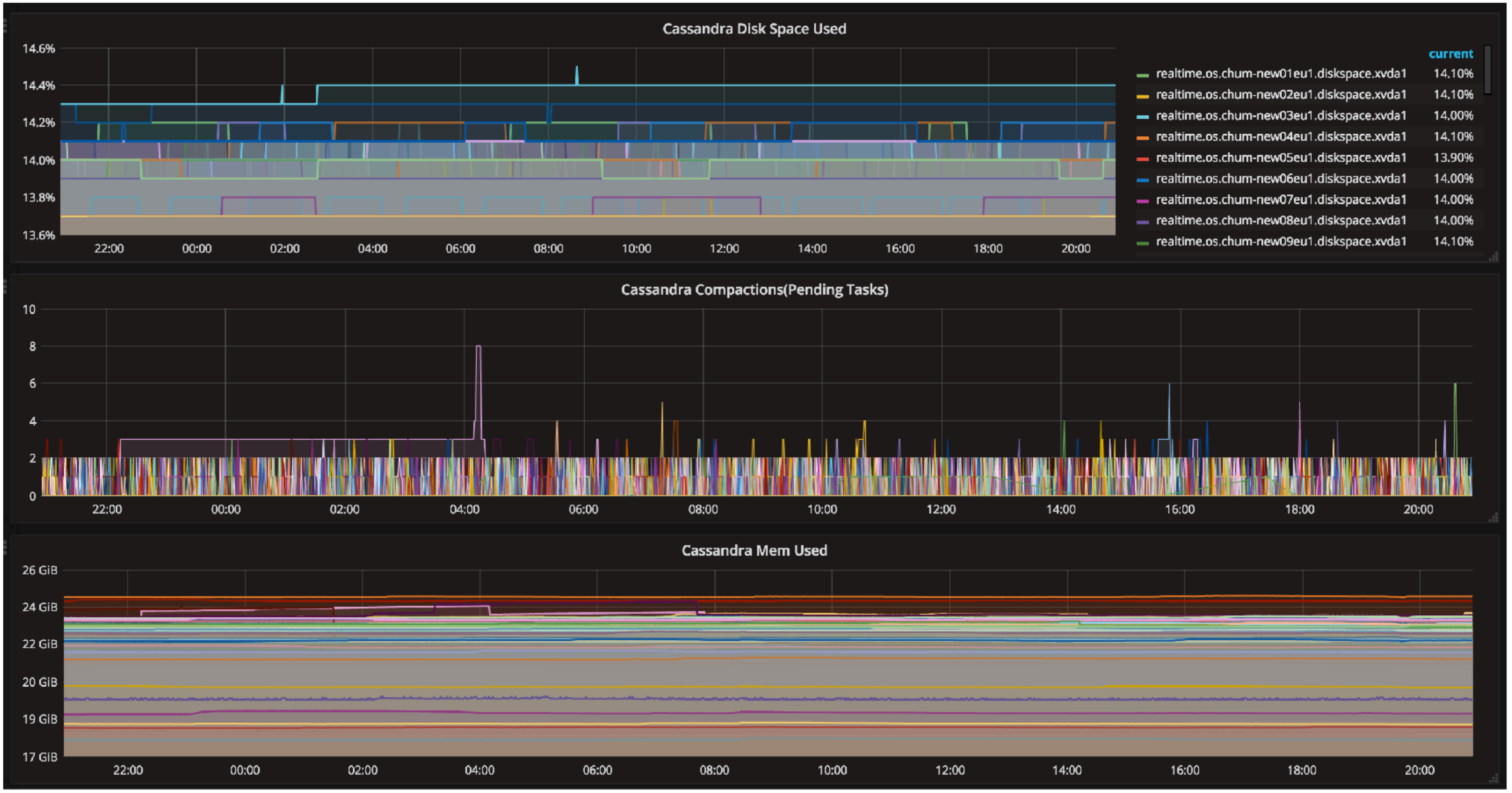
- test, test, test
- https://github.com/jsevellec/cassandra-unit
- https://github.com/Netflix/chaosmonkey

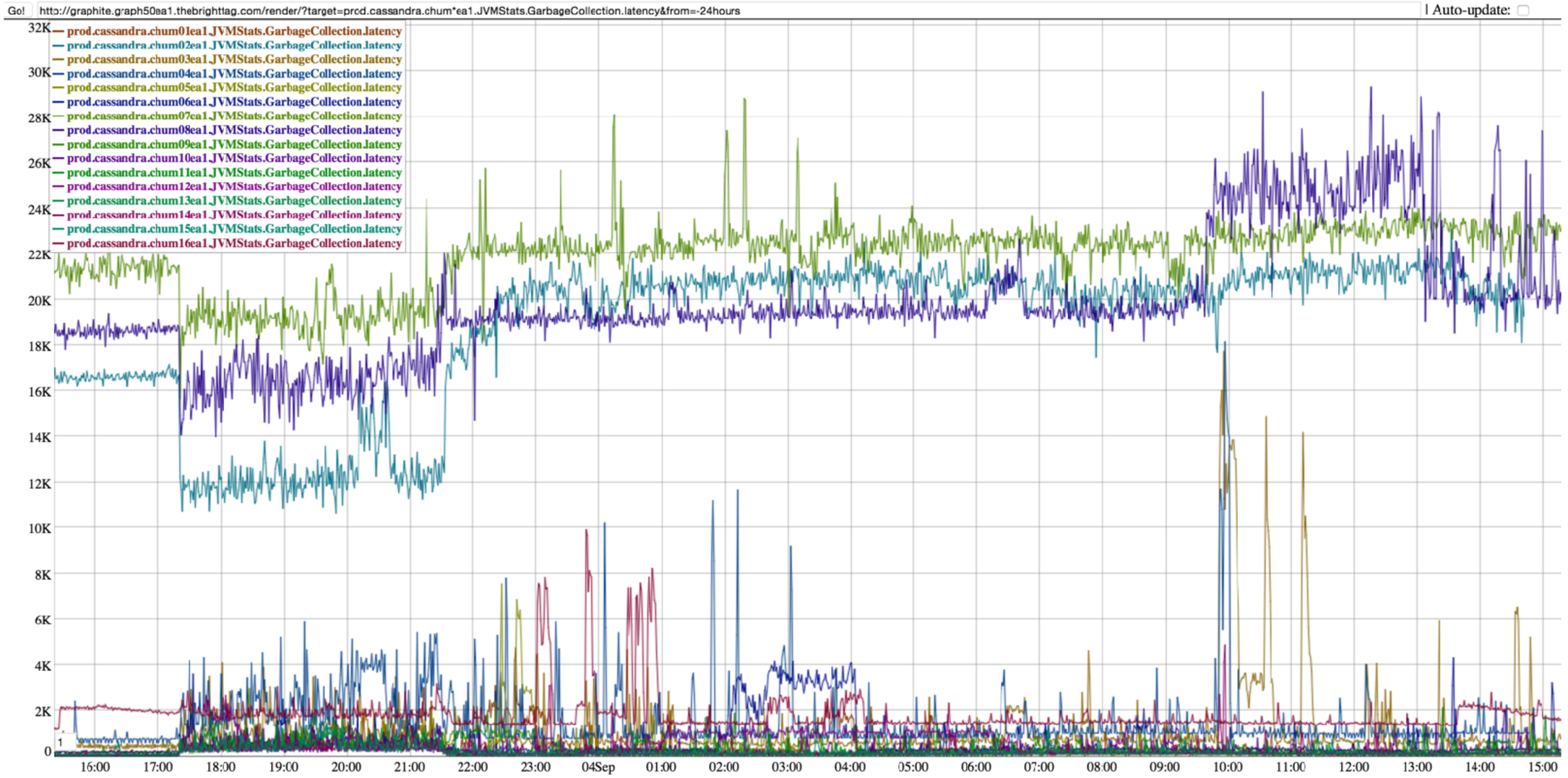
Monitoring will set you free!

Monitor Read Latency



SIGNAL®


Cassandra Disk Space Used
14.6%
14.4%
14.2%
14.0%
13.8%
13.6%
22:00
00:00
02:00
04:00
06:00
08:00
10:00
12:00
14:00
16:00
18:00
20:00
current
realtime.os.chum-new01eu1.diskspace.xvda1 14.10%
realtime.os.chum-new02eu1.diskspace.xvda1 14.10%
realtime.os.chum-new03eu1.diskspace.xvda1 14.00%
realtime.os.chum-new04eu1.diskspace.xvda1 14.10%
realtime.os.chum-new05eu1.diskspace.xvda1 13.90%
realtime.os.chum-new06eu1.diskspace.xvda1 14.00%
realtime.os.chum-new07eu1.diskspace.xvda1 14.00%
realtime.os.chum-new08eu1.diskspace.xvda1 14.00%
realtime.os.chum-new09eu1.diskspace.xvda1 14.10%
Cassandra Compactions(Pending Tasks)
10
8
6
4
2
0
22:00
00:00
02:00
04:00
06:00
08:00
10:00
12:00
14:00
16:00
18:00
20:00
Cassandra Mem Used
26 GiB
24 GiB
22 GiB
20 GiB
19 GiB
17 GiB
22:00
00:00
02:00
04:00
06:00
08:00
10:00
12:00
14:00
16:00
18:00
20:00


Go!
http://graphite.graph50ea1.thebrighttag.com/render/?target=prod.cassandra.chum*ea1.JVMStats.GarbageCollection.latency&from=-24hours
Auto-update:
prod.cassandra.chum01ea1.JVMStats.GarbageCollection.latency
prod.cassandra.chum02ea1.JVMStats.GarbageCollection.latency
prod.cassandra.chum03ea1.JVMStats.GarbageCollection.latency
prod.cassandra.chum04ea1.JVMStats.GarbageCollection.latency
prod.cassandra.chum05ea1.JVMStats.GarbageCollection.latency
prod.cassandra.chum06ea1.JVMStats.GarbageCollection.latency
prod.cassandra.chum07ea1.JVMStats.GarbageCollection.latency
prod.cassandra.chum08ea1.JVMStats.GarbageCollection.latency
prod.cassandra.chum09ea1.JVMStats.GarbageCollection.latency
prod.cassandra.chum10ea1.JVMStats.GarbageCollection.latency
prod.cassandra.chum11ea1.JVMStats.GarbageCollection.latency
prod.cassandra.chum12ea1.JVMStats.GarbageCollection.latency
prod.cassandra.chum13ea1.JVMStats.GarbageCollection.latency
prod.cassandra.chum14ea1.JVMStats.GarbageCollection.latency
prod.cassandra.chum15ea1.JVMStats.GarbageCollection.latency
prod.cassandra.chum16ea1.JVMStats.GarbageCollection.latency
32K
30K
28K
26K
24K
22K
20K
18K
16K
14K
12K
10K
8K
6K
4K
2K
0
16:00
17:00
18:00
19:00
20:00
21:00
22:00
23:00
04Sep
01:00
02:00
03:00
04:00
05:00
06:00
07:00
08:00
09:00
10:00
11:00
12:00
13:00
14:00
15:00


THE
APACHE®
SOFTWARE FOUNDATION


KEEP
CALM
'CAUSE YOU'RE
NOT
ALONE


DATASTAX

signal.co/careers

Hiring in ***Chicago*** and ***NYC*** office

SIGNAL®

Thank You!

acolovic@signal.co

@akicolovic

CASSANDRA SUMMIT 2015

SIGNAL®