

Next Session: 4:14pm

Scaling Answers by
Crashlytics with the
Lambda Architecture

Ed Solovey



flight

#twitterflight



Ed Solovey

Staff Software Engineer
Answers by Crashlytics

@edsolovey



Scaling Answers by Crashlytics with the Lambda Architecture

October 22, 2014

#twitterflight



Cannonball

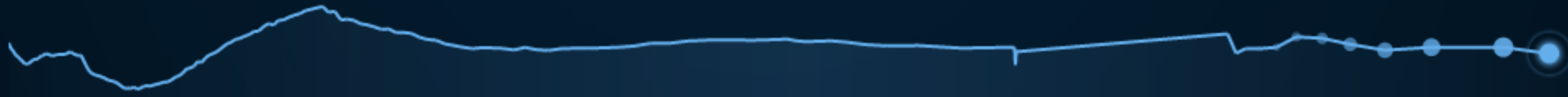
ios io.fabric.samples.Cannonball

Ed  SDK Canonical Sample App

-
-
-
-
-

ACTIVE USERS

78
RIGHT NOW



DAILY ACTIVE USERS

2.1 k

▲ 1.6%

DAILY NEW USERS

51

▼ -0.3%

MONTHLY ACTIVE USERS

6.6 k

▲ 1.4%

CRASH-FREE USERS

94.9 %

▼ -0.5%

SESSIONS

636.1 k

▼ -1.1%

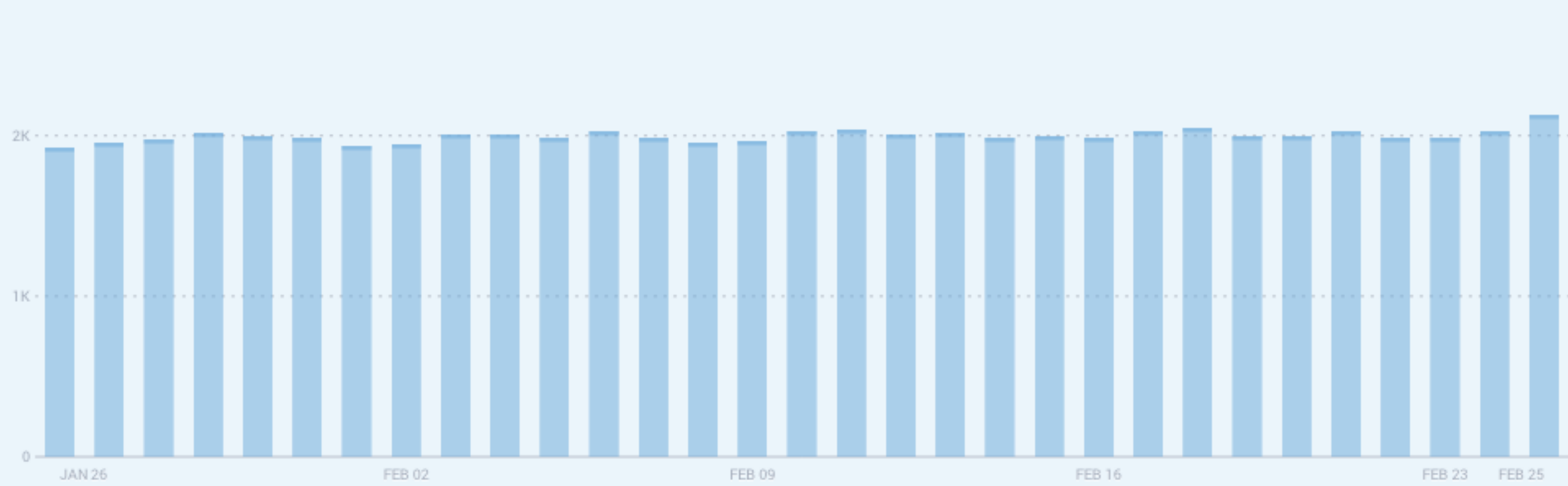
SESSION LENGTH

2:07 MIN

▲ 0.0%

DAILY ACTIVE USERS

2.1k
TODAY



-
-
-
-

Challenges

SCALABLE

FAULT-TOLERANT

ACTIONABLE

THOUSANDS
OF APPS

MILLIONS OF
MOBILE DEVICES

REAL-TIME

8000,0000

events every second

A Start: Stream Compute

800K
EVENTS /
SECOND

QUEUE



PERSIST
RESULTS

DB

samza

Spark

Stream Compute: Daily Active Users

```
{io.fabric.samples.cannonball,  
device-one,  
1412962201}
```



```
io.fabric.samples.cannonball  
-----  
device-one
```

Stream Compute: Daily Active Users

```
io.fabric.samples.cannonball  
-----  
device-one
```



```
{com.twitter.android,  
device-two,  
1412962263}
```

```
com.twitter.android  
-----  
device-two
```

Stream Compute: Daily Active Users



```
{io.fabric.samples.cannonball,  
device-three,  
1412962284}
```

```
io.fabric.samples.cannonball  
-----  
device-one  
device-three
```

```
com.twitter.android  
-----  
device-two
```

Stream Compute: Daily Active Users



```
{io.fabric.samples.cannonball,  
device-four,  
1412963002}
```

```
io.fabric.samples.cannonball  
-----  
device-one  
device-three  
device-four
```

```
com.twitter.android  
-----  
device-two
```

Two Major Problems

1. Fault tolerance

Recover From Failure: Batch Compute

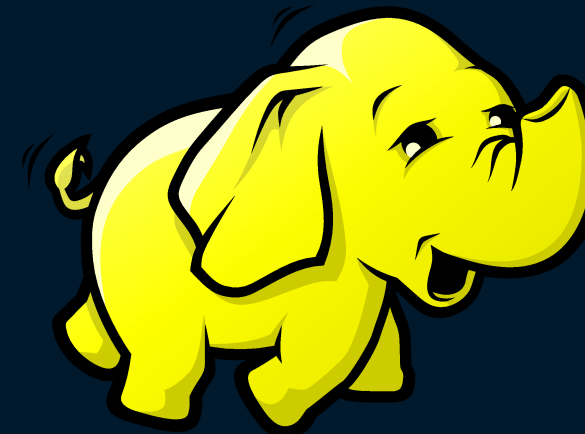
Archive Everything: Process in Batch

archive in parallel to speed processing
write to scalable store: S3, HDFS....

800K EVENTS /
SECOND

QUEUE

ARCHIVER



```
{ io.fabric.samples.cannonball,  
device-one, 1412962201 }
```

```
{ com.twitter.android,  
device-two, 1412962263 }
```

```
{ io.fabric.samples.cannonball,  
device-three, 1412962284 }
```

```
{ io.fabric.samples.cannonball,  
device-four, 1412962303 }
```

800K EVENTS /
SECOND

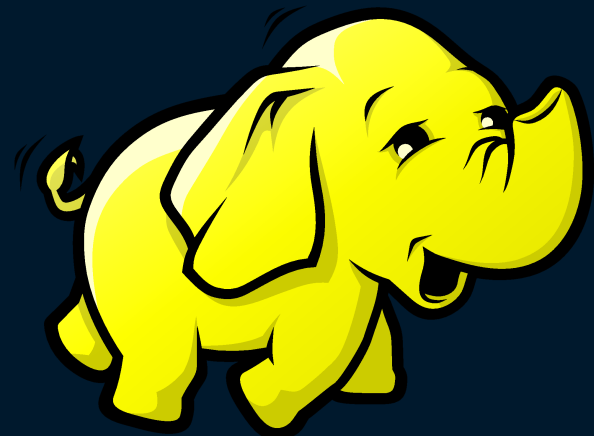
QUEUE



PERSIST
RESULTS

DB

Process in Batch



cascading

```
{io.fabric.samples.cannonball,  
device-one, 1412962201}
```

```
{com.twitter.android,  
device-two, 1412962263}
```

```
{io.fabric.samples.cannonball,  
device-three, 1412962284}
```

```
{io.fabric.samples.cannonball,  
device-four, 1412962303}
```



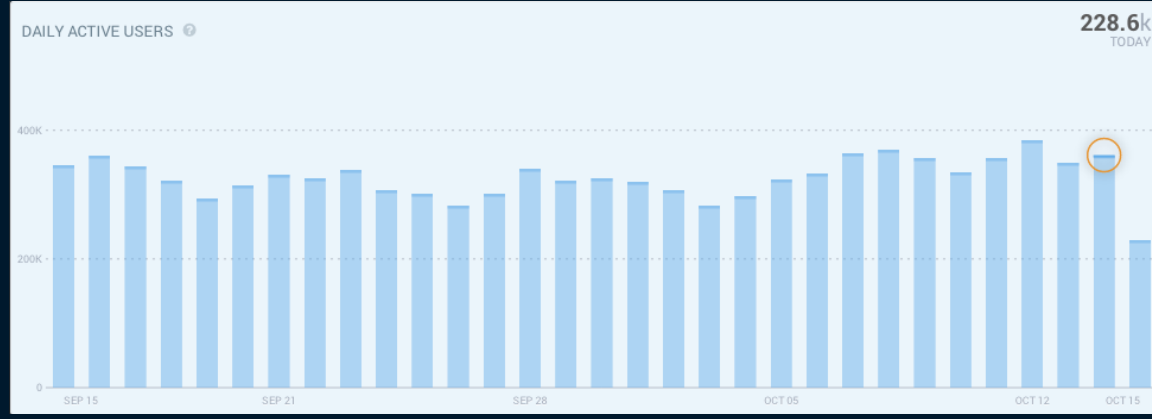
NIGHTLY

```
io.fabric.samples.cannonball,  
10/21/14 : 3  
com.twitter.android,  
10/21/14 : 1
```



Amazon Elastic
Map Reduce (EMR)

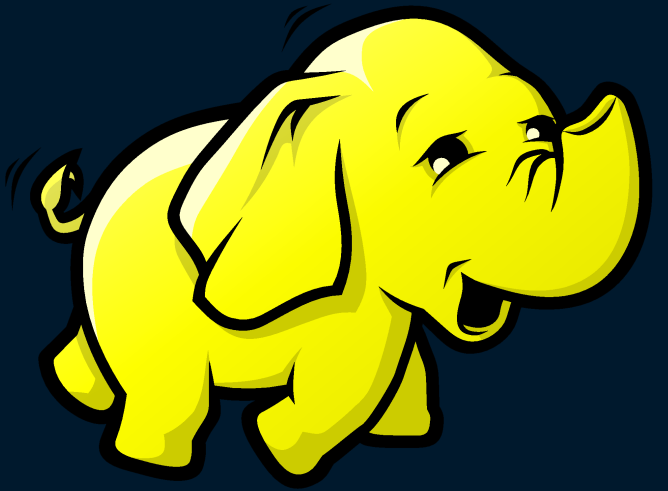
Fitting it Together: DAU



QUERY BATCH



```
cannonball, 10/21/14 : 3  
twitter, 10/21/14 : 1  
  
cannonball, 10/20/14 : 2  
twitter, 10/20/14 : 3  
cannonball, 10/19/14 : 1  
twitter, 10/19/14 : 7
```



```
timeseries :  
cannonball: 1,2,3,1  
twitter: 7,3,1,1
```

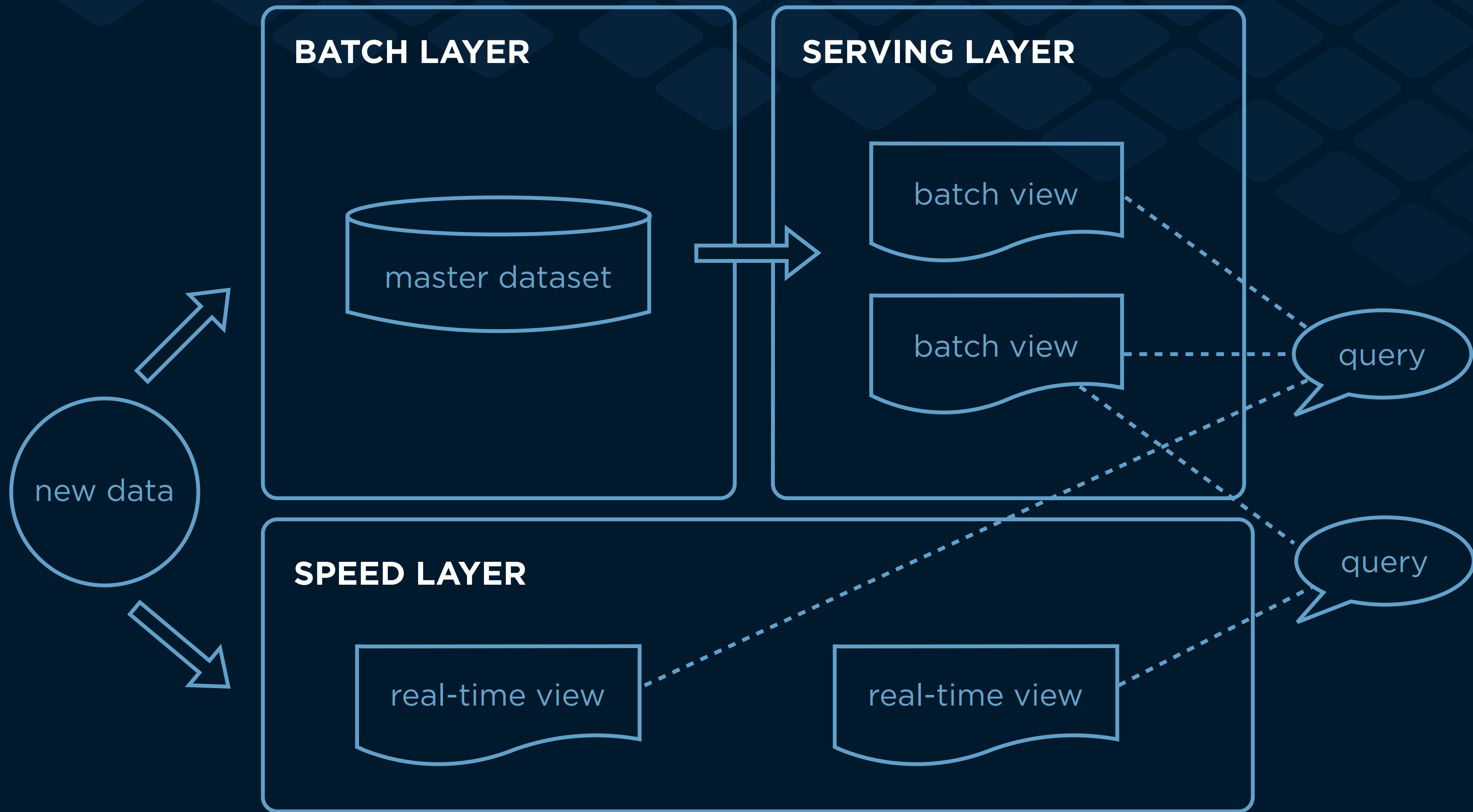
QUERY SPEED



COMBINE RESULTS

```
cannonball, 10/22/14 : 1  
twitter, 10/22/14 : 1
```

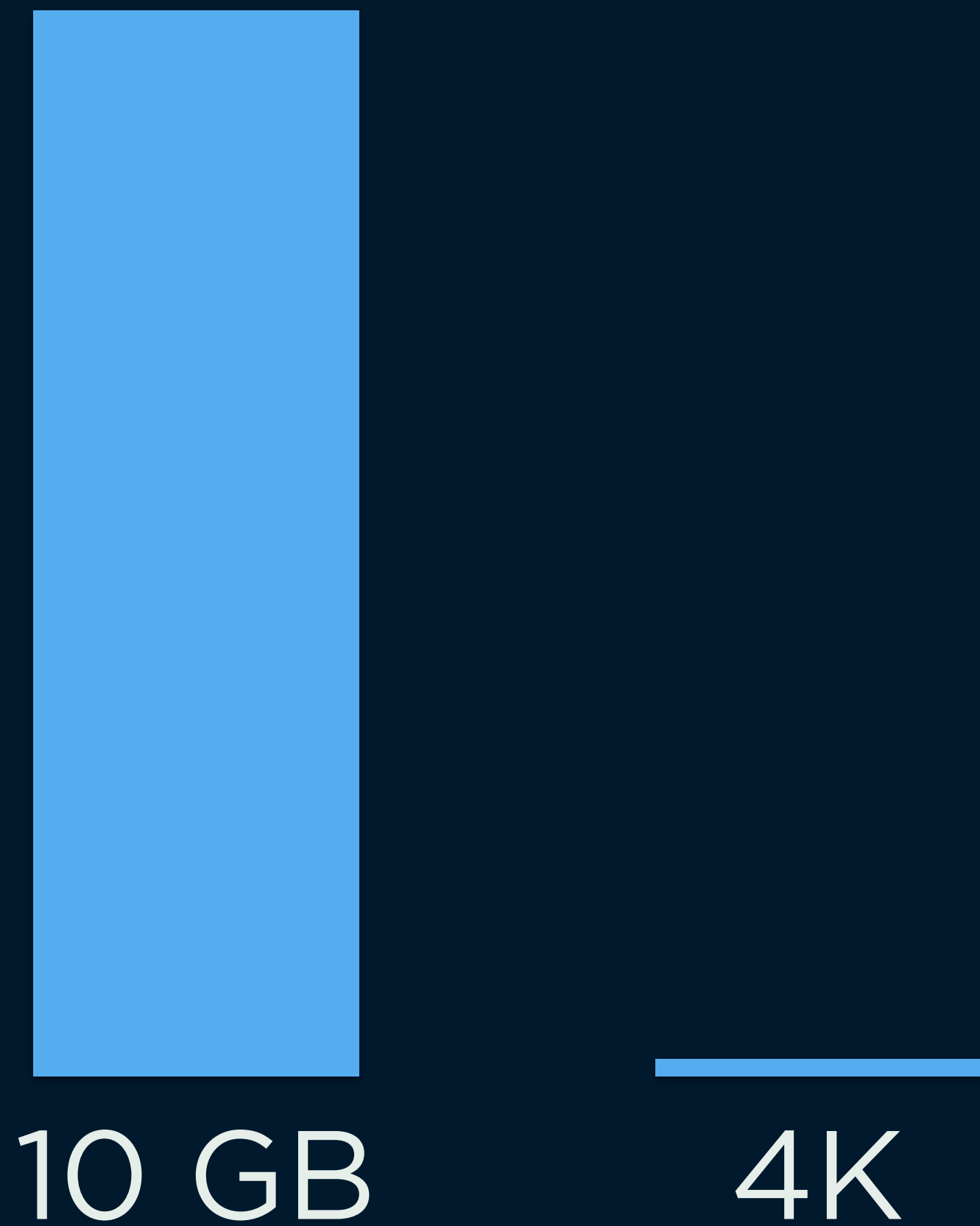




Two Major Problems

1. ~~Fault tolerance~~
2. Size of Raw Sets

HyperLogLog



Up to a billion of devices
Lose ~ 1% accuracy

HyperLogLog

Hash incoming device-id

Look for a pattern; e.g. longest sub-string of consecutive ones

Susceptible to skew

device-one

HASH



01001110101001100100100001101111000010101011

HyperLogLog

Can create multiple groups either by having many hash functions or bucketing incoming values

Then take mean of groups

device-one

HASH

01001110101001100100100001101111000010101011

four 1's

device-two

HASH

00101010011100100100010010010010101010001000

three 1's

device-three

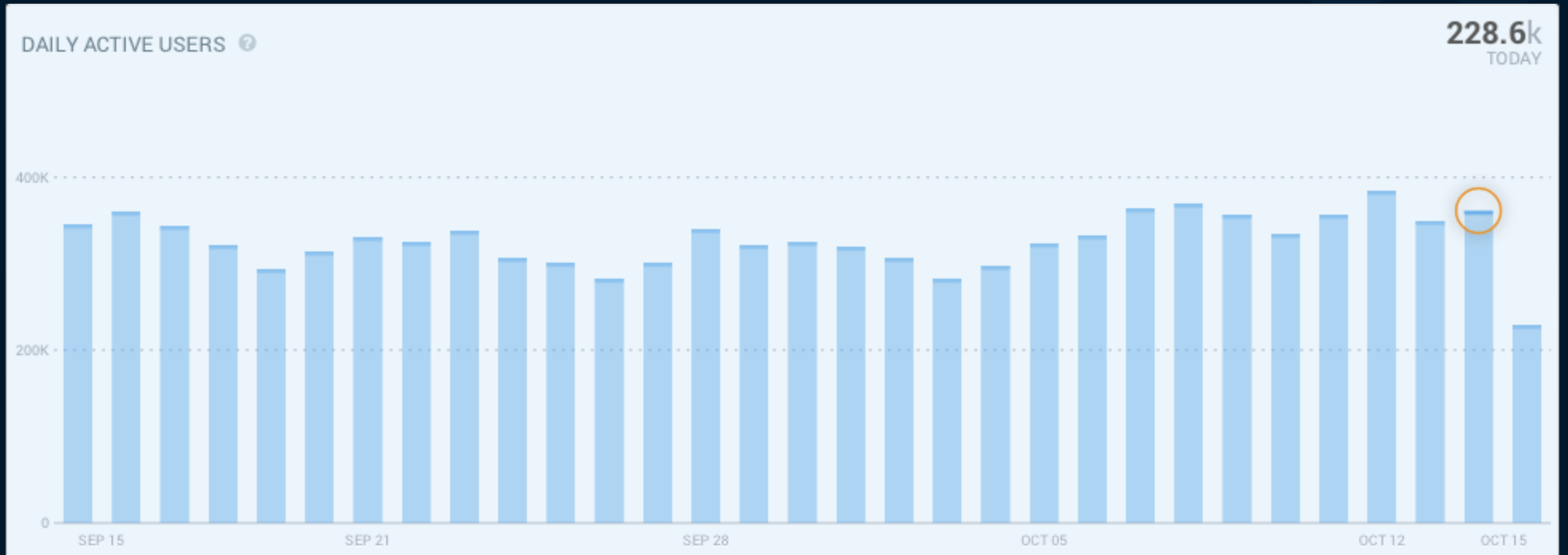
HASH

00101010010100100100010011010010101010001000

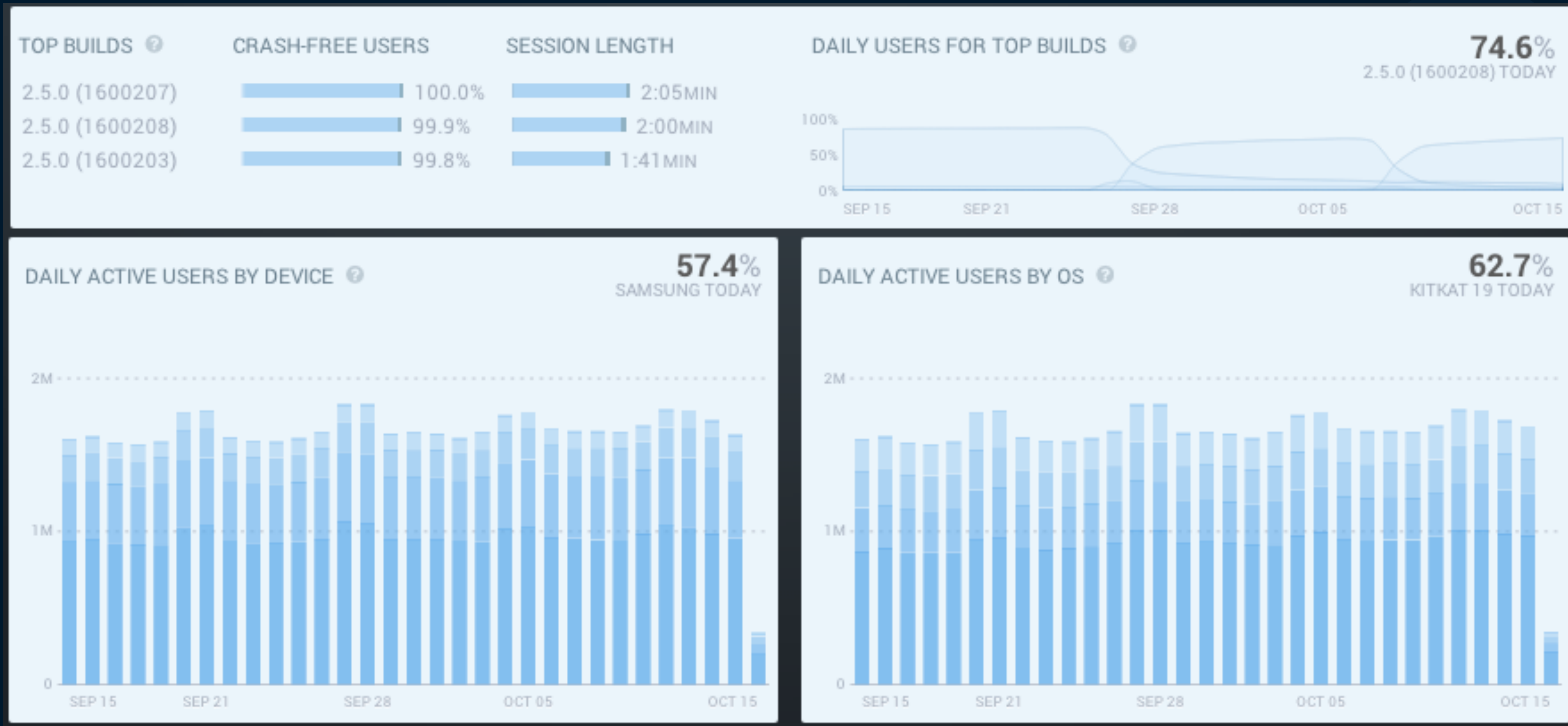
two 1's

What Can We Build With These Tools?

Anomaly Detection



Dimensional Analysis



Attributes

SCALABLE

FAULT-TOLERANT

ACTIONABLE

THOUSANDS
OF APPS

MILLIONS OF
MOBILE DEVICES

REAL-TIME

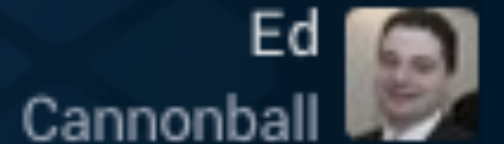
Real-time Active Devices



Cannonball

iOS io.fabric.samples.Cannonball

Ed



Cannonball



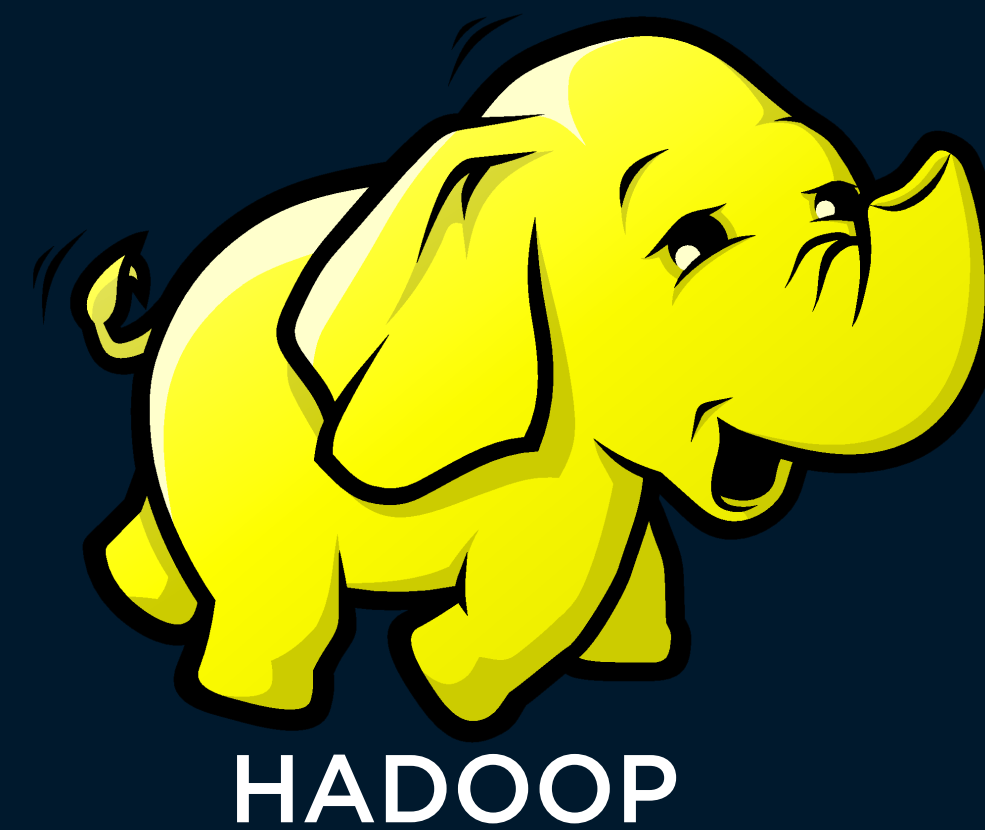
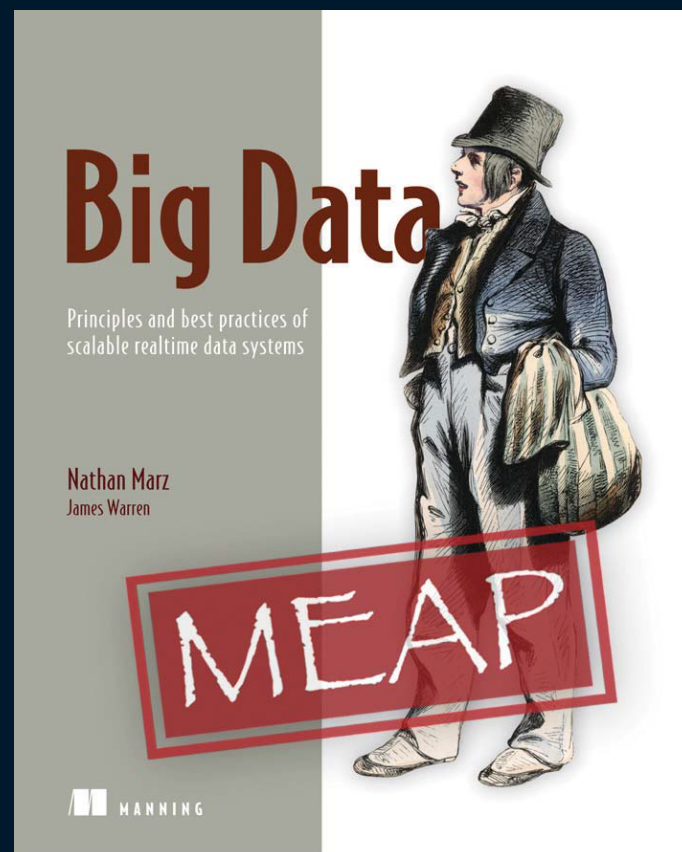
ACTIVE USERS ⓘ

87

RIGHT NOW



Learn More!





Thank You

@edsolovey