



SOUNDCLOUD

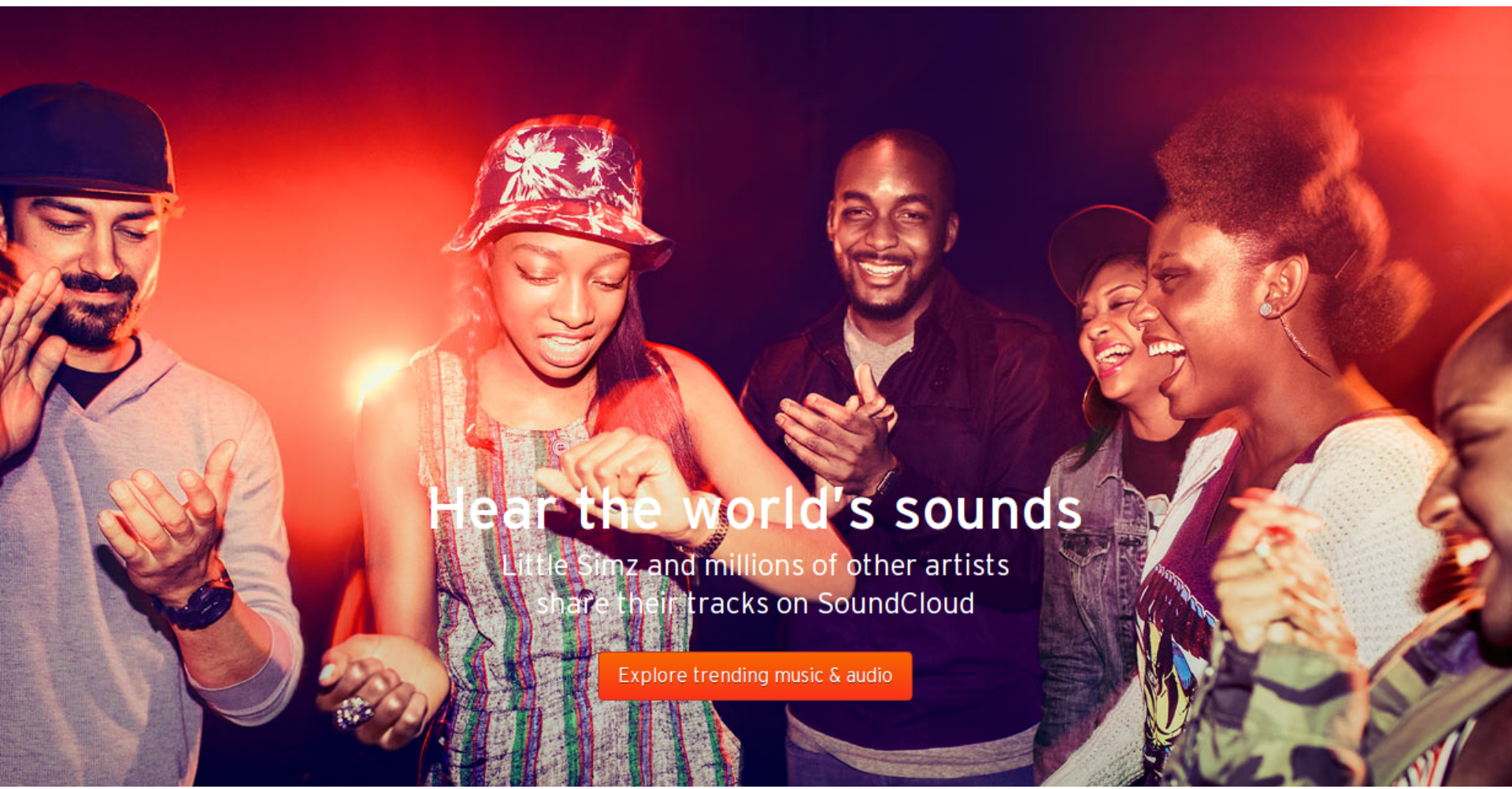


How SoundCloud Uses Cassandra

<http://bit.ly/qcon-2015-soundcloud-cassandra>

@greenEmilyMay

Emily Green



Hear the world's sounds

Little Simz and millions of other artists
share their tracks on SoundCloud

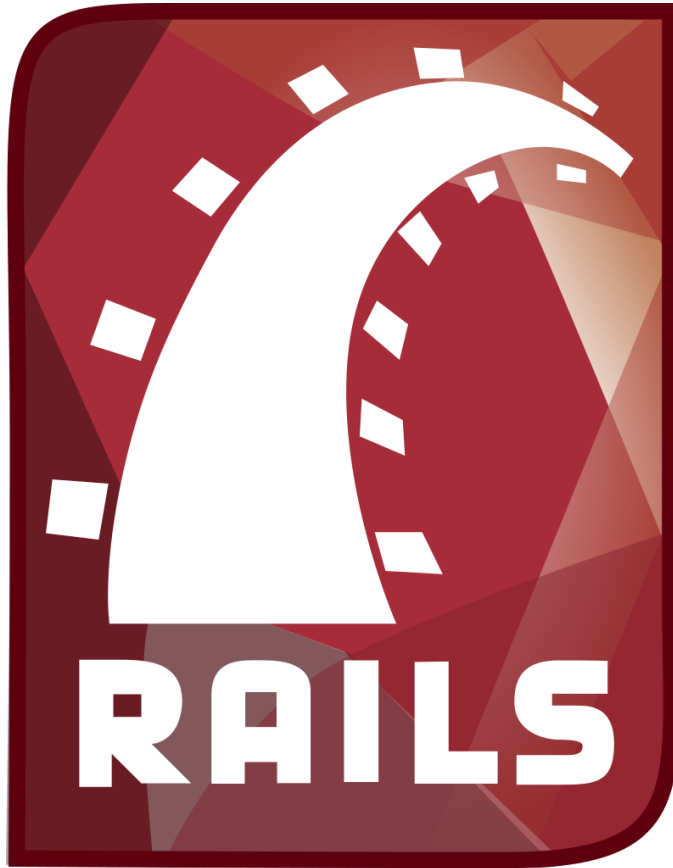
Explore trending music & audio

Coming up: history of data sharing

Coming up: history of data sharing

2 cassandra systems

EVOLUTION STEP 0: MONOLITH

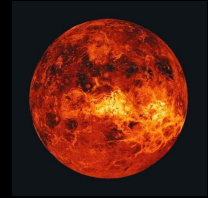
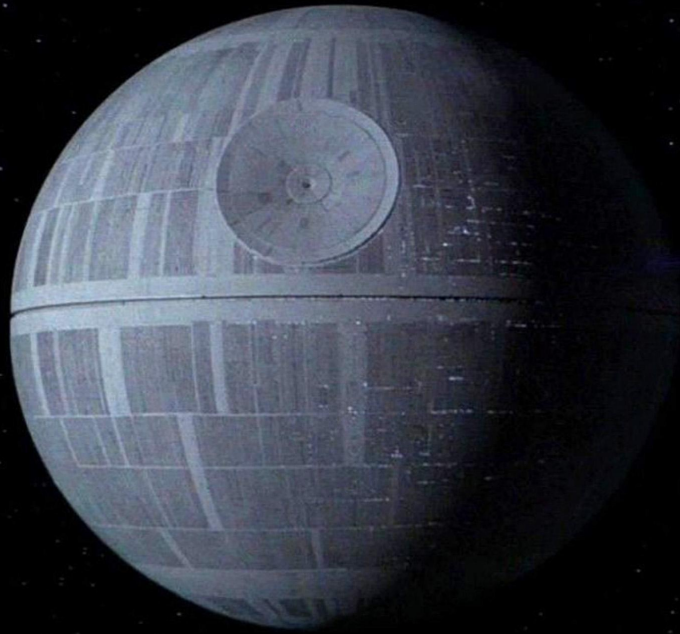




GET IN THE MOTHERSHIP

TheFilmTroll.com





EVOLUTION STEP 1: SHARE DATA BY READING FROM MYSQL SLAVES



Home

Collection

Search



Upload



emgreen



Search

Everything

Tracks

Playlists

People

Groups

Filter results

Added any time

Any length

To listen to

Filter by tag

Alternative # House # Pop

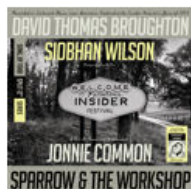
Folk # Weird pop # Dance

Singer-songwriter # Scotland

Podcast

72 tracks found

Siobhan Wilson

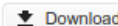
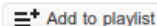


Song, by Toad

Siobhan Wilson - Dear God

8 months

#insider festival



5,762 | 104 | 20 | 15

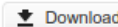


ANORAAK

Siobhan Wilson - Getting me down (Anoraak rework)

4 years

#Electropop



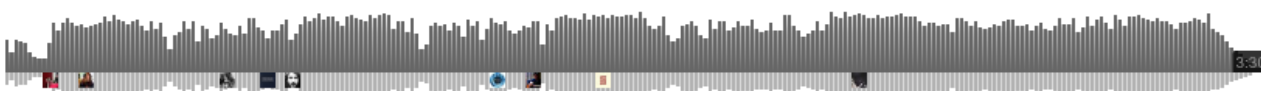
22,040 | 297 | 36 | 41



Siobhan Wilson Music

All Dressed Up - Siobhan Wilson

1 year



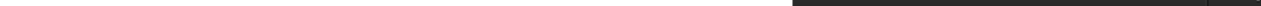
3,149 | 95 | 18 | 9



Siobhan Wilson Music

The Good Years - Siobhan Wilson (Cover of Karine Polwart and

2 years



Audio player controls: play, pause, skip, volume, and share icons.

**EVOLUTION STEP 2:
SHARE DATA BY
READING FROM MYSQL
SLAVES AND WITH
EVENTS**



Home

Collection

Search



Upload

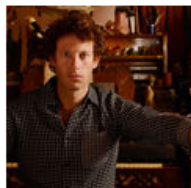


emgreen



Stream Explore

Hear the latest posts from the people you're following



samleesongs Folk Radio UK

Sam Lee And Friends KEXP NYC Live Session NPR

19 mins

#folk



Like Repost Add to playlist Share

326 | 20 | 3 | 4



Bitch Media

She-Hulk, Attorney at Law

10 hours

#podcast



Like Repost Add to playlist Share Download



Bitch Media

Lessons from Law and Order

10 hours

#podcast



Like Repost Add to playlist Share Download



Explore tracks by genre

Find new & established genres on SoundCloud Explore.

Explore tracks by genre

Who to follow

Refresh



Toni Romiti

15,732

Follow



Tamar Braxton

316,990

Follow



Curb Records

2,730,170 | 6

Follow

35 likes

View all



Lost Map

Rozi Plain - 'Actually'

6,257 | 206 | 32 | 14

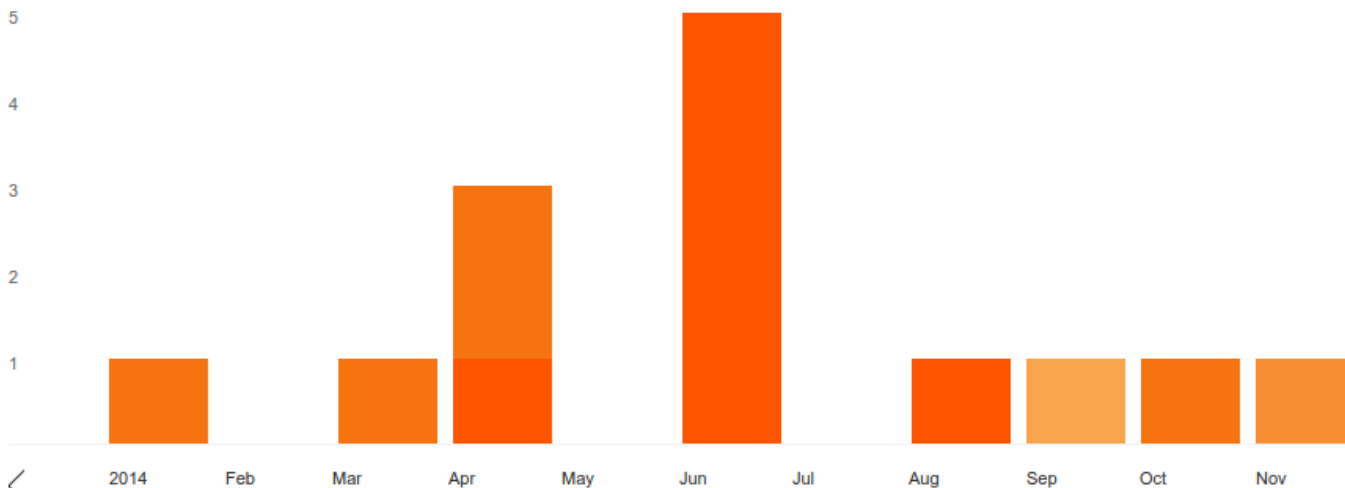


EVOLUTION STEP 3: SHARE DATA WITH EVENTS

SOME PEOPLE LISTEN


14 ▶ **Plays**
0 ♥ **Likes**
0 💬 **Comments**
0 ↻ **Reposts**
0 ⬇️ **Downloads**

Custom range
 Jan - Nov, 2014 ▼



Slightly clearer purring from AL.	7
Some More Alby Purring	5
Alby-purr-2	1
Alby-purr-1	1
Others	0


Most played tracks

	Slightly clearer purring from ...	7
2	Some More Alby Purring	5
3	Alby-purr-2	1
4	Alby-purr-1	1

Top countries Top cities

L	London, United Kingdom	3
2	Banbury, United Kingdom	2
3	Düsseldorf, Germany	1
4	Toronto, Canada	1
5	Hünxe, Germany	1
6	Runcorn, United Kingdom	1
7	Berlin, Germany	1

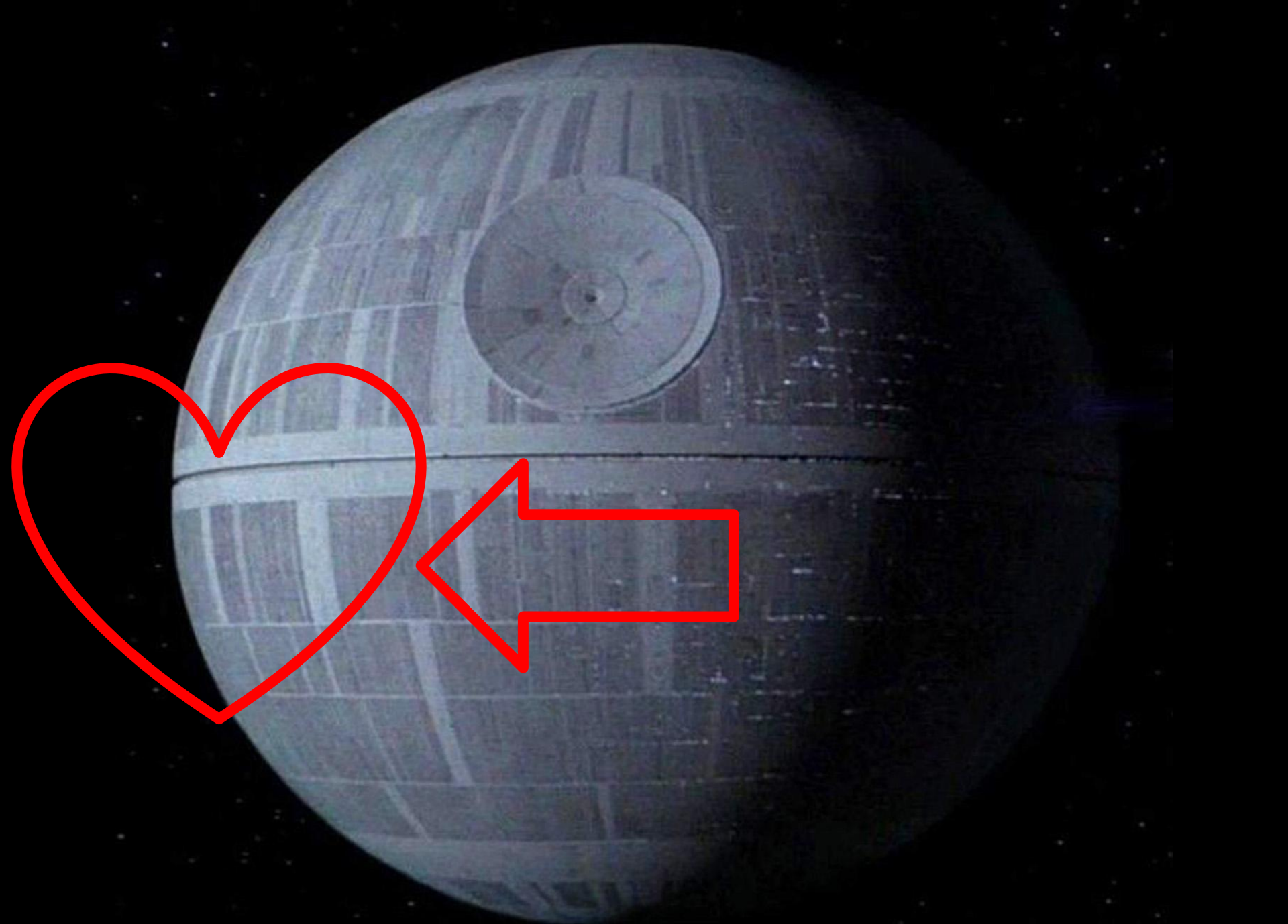
Who played the most

	Jarred Trainor	1
2	dan simpson	1
	Others	12

Websites Apps

S	SoundCloud.com	5
2	SoundCloud Visual Embed Player	5
3	Mobile Soundcloud	2
4	tancode	1
5	SoundCloud Mobile Web	1

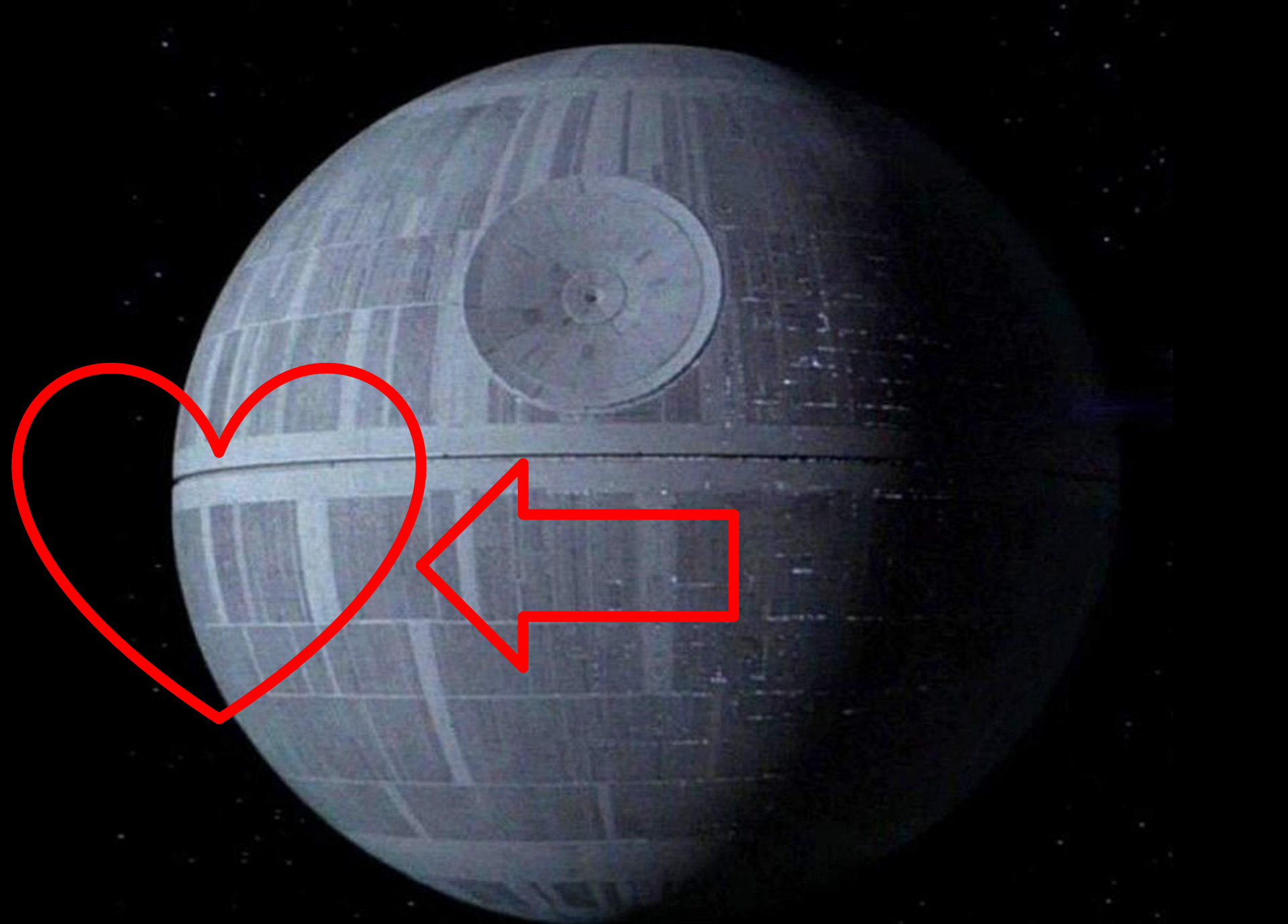
**EVOLUTION STEP 4:
SOMETHING OTHER
THAN MOTHERSHIP
OWNS SOME DATA**



GET OFF OUR



DATABASE



CONSISTENCY



AND STATE

Conclusion 1:
amazing as large
realised view on
events

Conclusion 2:
consistency takes
thought

Other uses:
fingerprinting

**Other uses:
fingerprinting
zipkin**

Coming up: history of data sharing

2 cassandra systems

THIS IS ALBY.

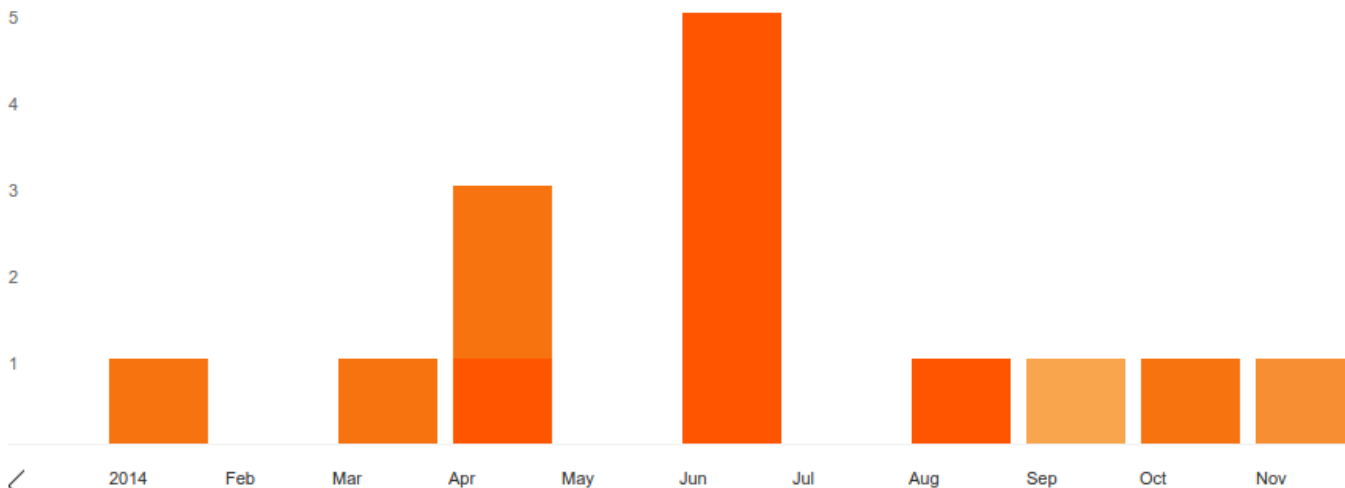


I UPLOAD HER SOUNDS

SOME PEOPLE LISTEN


14 Plays 0 Likes 0 Comments 0 Reposts 0 Downloads

Custom range
Jan - Nov, 2014 ▼



Slightly clearer purring from AL.	7
Some More Alby Purring	5
Alby-purr-2	1
Alby-purr-1	1
Others	0


Most played tracks

	Slightly clearer purring from ...	7
2	Some More Alby Purring	5
3	Alby-purr-2	1
4	Alby-purr-1	1

Top countries Top cities

L	London, United Kingdom	3
2	Banbury, United Kingdom	2
3	Düsseldorf, Germany	1
4	Toronto, Canada	1
5	Hünxe, Germany	1
6	Runcorn, United Kingdom	1
7	Berlin, Germany	1

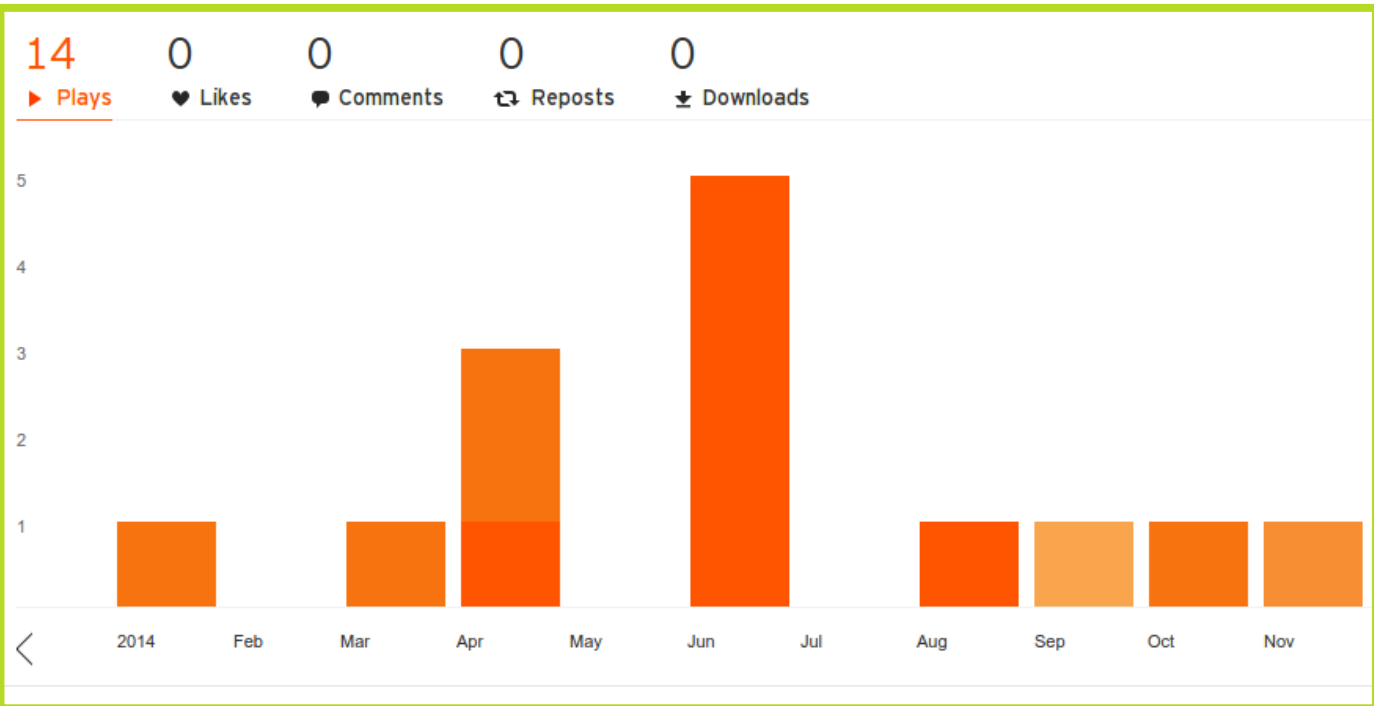
Who played the most

	Jarred Trainor	1
2	dan simpson	1
	Others	12

Websites Apps

S	SoundCloud.com	5
2	SoundCloud Visual Embed Player	5
3	Mobile Soundcloud	2
4	tancode	1
5	SoundCloud Mobile Web	1

SOME PEOPLE LISTEN




Custom range

Jan - Nov, 2014 ▼

Slightly clearer purring from AL.	7
Some More Alby Purring	5
Alby-purr-2	1
Alby-purr-1	1
Others	0


Most played tracks

	Slightly clearer purring from ...	7
2	Some More Alby Purring	5
3	Alby-purr-2	1
4	Alby-purr-1	1

Top countries Top cities

L	London, United Kingdom	3
2	Banbury, United Kingdom	2
3	Düsseldorf, Germany	1
4	Toronto, Canada	1
5	Hünxe, Germany	1
6	Runcorn, United Kingdom	1
7	Berlin, Germany	1

Who played the most

	Jarred Trainor	1
2	dan simpson	1
	Others	12

Websites Apps

S	SoundCloud.com	5
2	SoundCloud Visual Embed Player	5
3	Mobile Soundcloud	2
4	tancode	1
5	SoundCloud Mobile Web	1

SOME PEOPLE LISTEN

14

▶ Plays

0

♥ Likes

0

💬 Comments

0

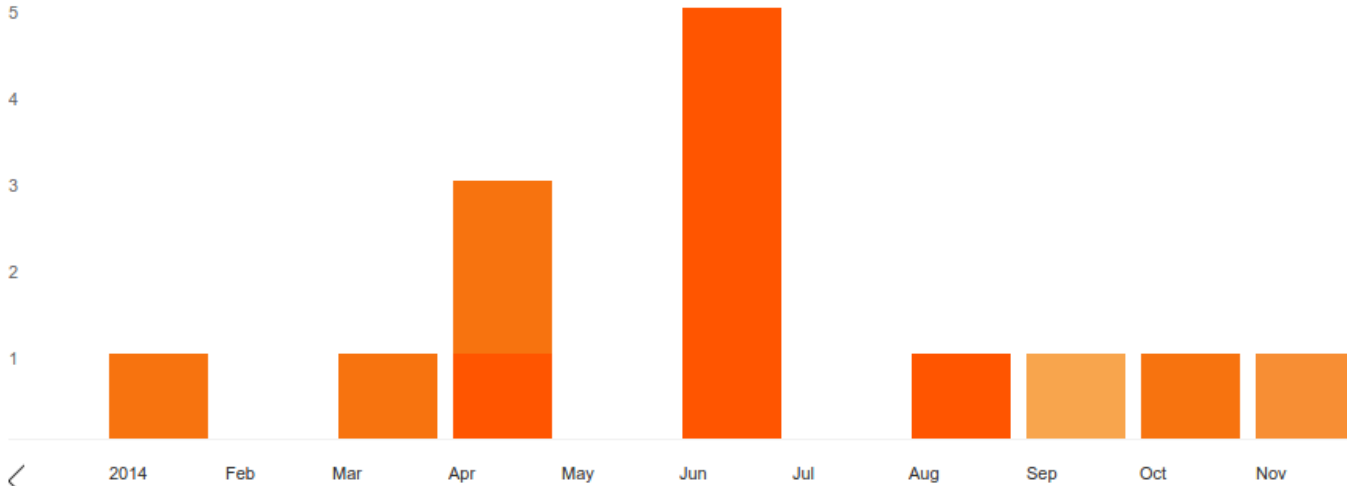
↻ Reposts

0

⬇ Downloads


Custom range

Jan - Nov, 2014 ▼



Slightly clearer purring from AL.	7
Some More Alby Purring	5
Alby-purr-2	1
Alby-purr-1	1
Others	0


Most played tracks

	Slightly clearer purring from ...	7
2	Some More Alby Purring	5
3	Alby-purr-2	1
4	Alby-purr-1	1

Top countries Top cities

L	London, United Kingdom	3
2	Banbury, United Kingdom	2
3	Düsseldorf, Germany	1
4	Toronto, Canada	1
5	Hünxe, Germany	1
6	Runcorn, United Kingdom	1
7	Berlin, Germany	1

Who played the most

	Jarred Trainor	1
2	dan simpson	1
	Others	12

Websites Apps

S	SoundCloud.com	5
2	SoundCloud Visual Embed Player	5
3	Mobile Soundcloud	2
4	tancode	1
5	SoundCloud Mobile Web	1

REAL-TIME COUNT TIMESERIES

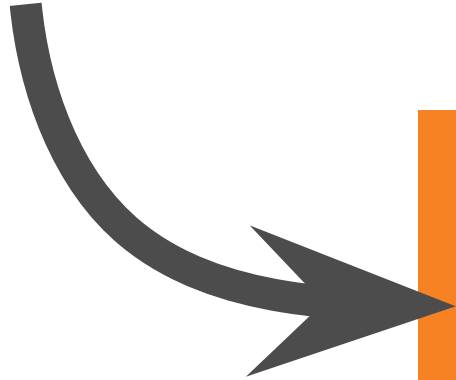


BATCH TOP-K

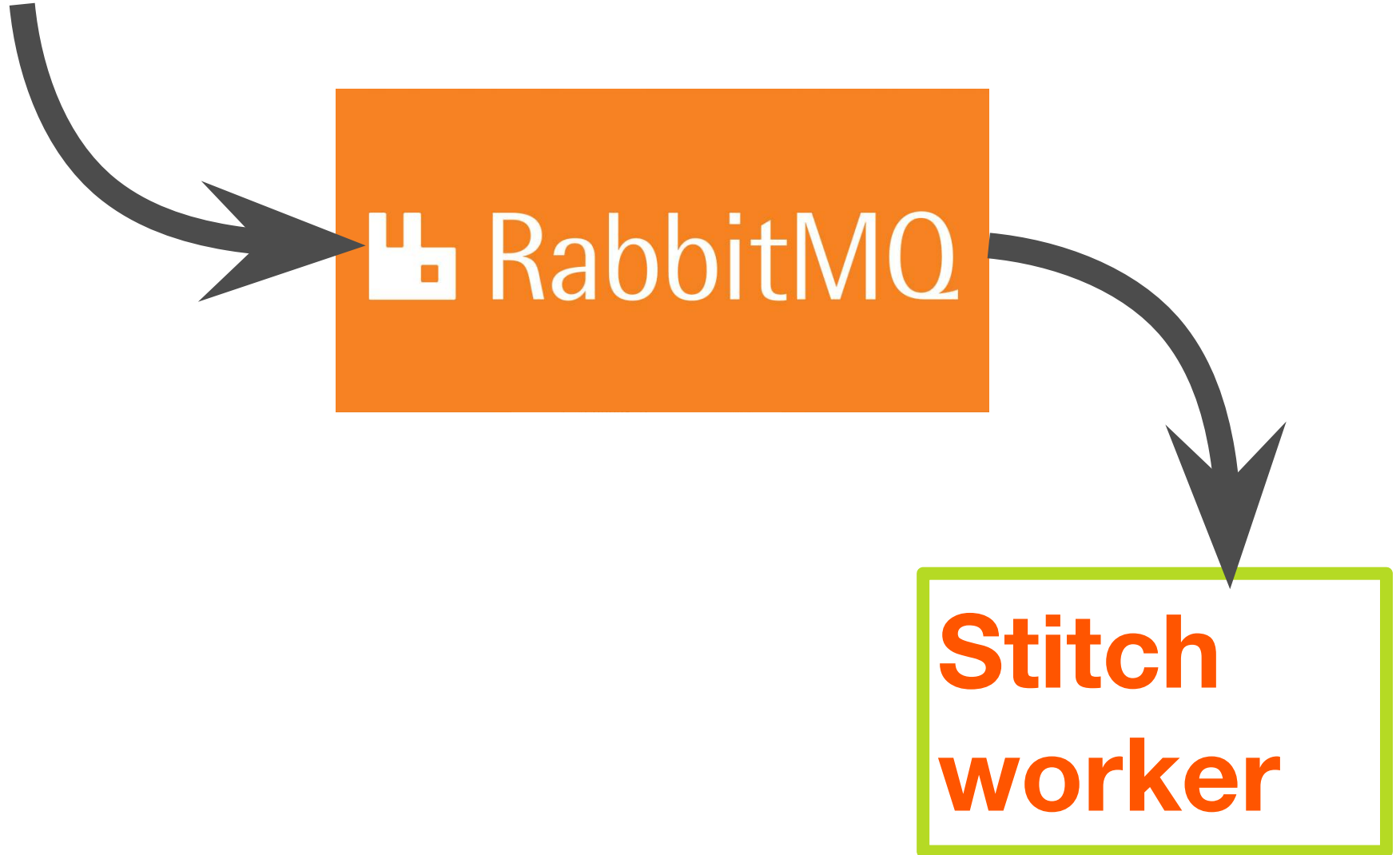
```
{ play: {  
  at: 2014-11-12 15:40,  
  trackId: 1234, ownerId: 5678 }}
```



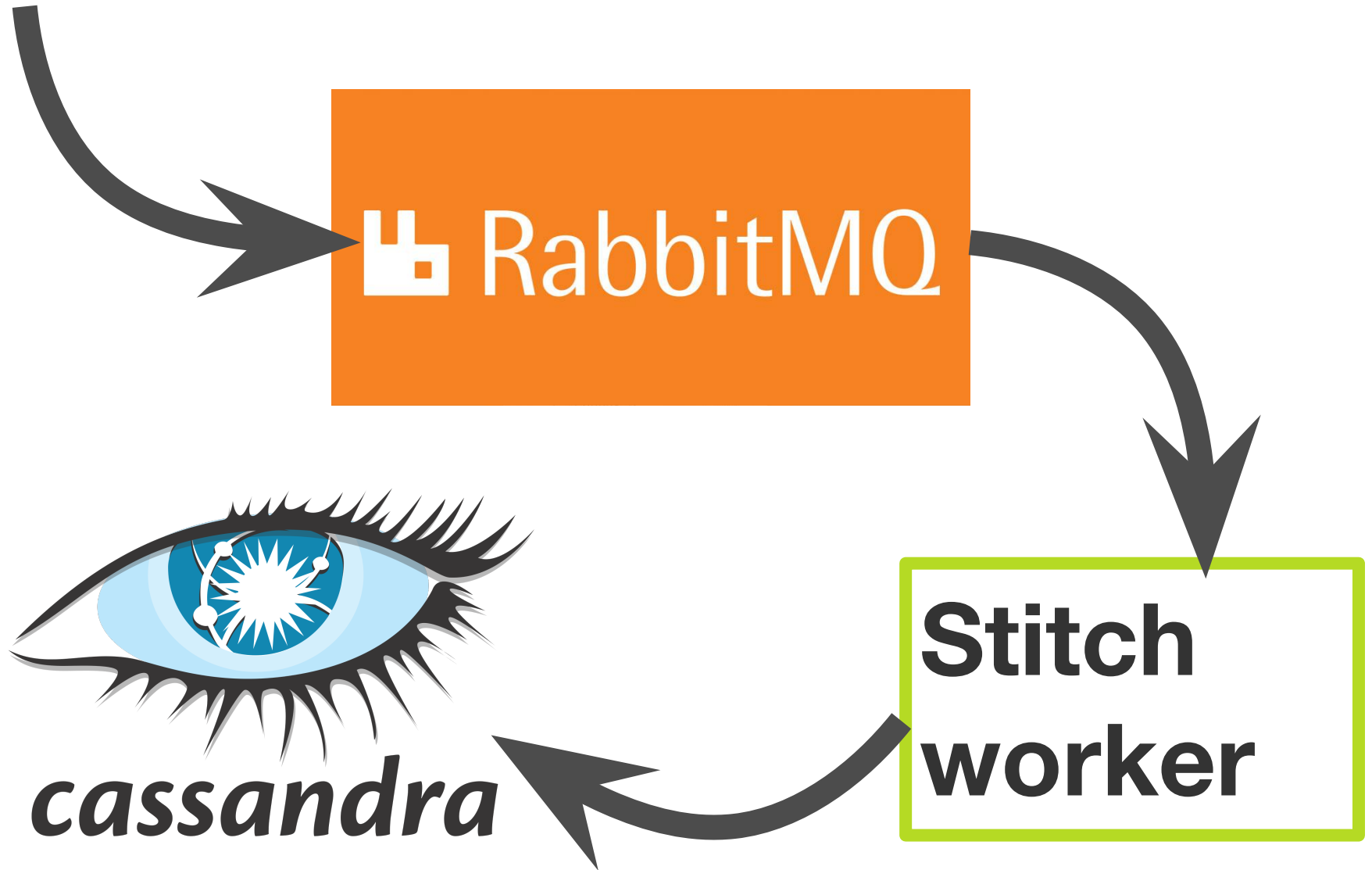
```
{ play: {  
  at: 2014-11-12 15:40,  
  trackId: 1234, ownerId: 5678 }}
```



```
{ play: {  
  at: 2014-11-12 15:40,  
  trackId: 1234, ownerId: 5678 }}
```



```
{ play: {  
  at: 2014-11-12 15:40,  
  trackId: 1234, ownerId: 5678 }}
```





**Stitch
web**

cassandra

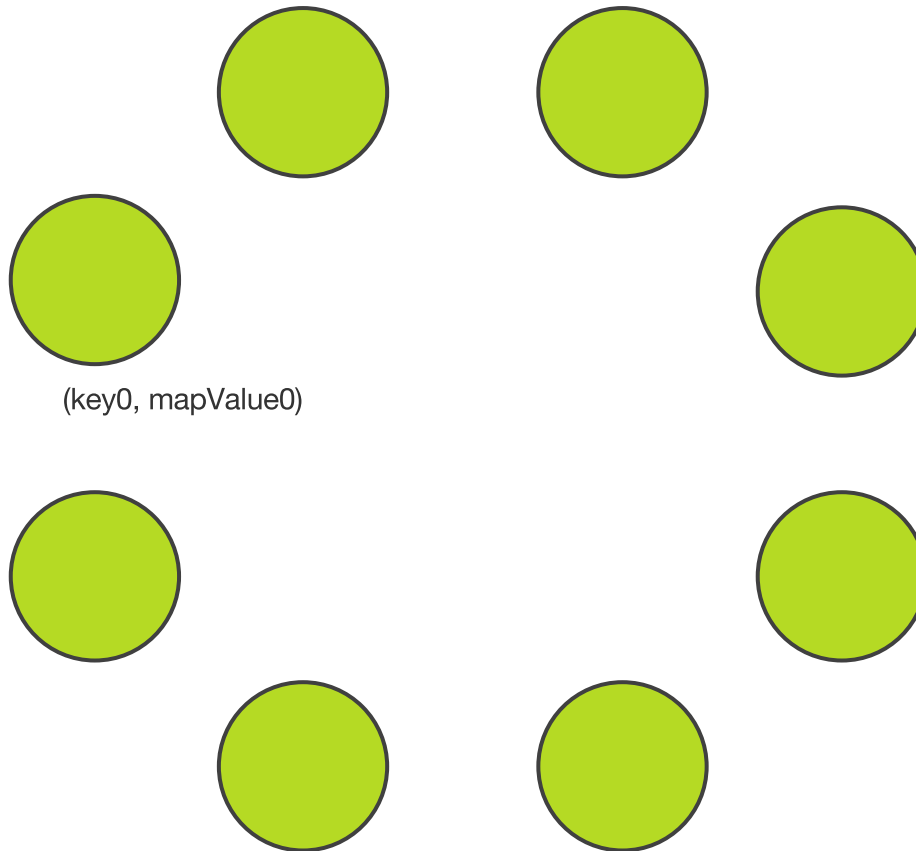
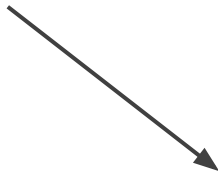
Map[A, Map[B, C]]

Map[A, Map[B, C]]

(key0, mapValue0)



hash(key0)



Map[Id,Map[TimeBucket,Count]]

Map[Id,Map[TimeBucket,Count]]

183 -> (
1380585600000 -> 59904,
1383264000000 -> 62195
)

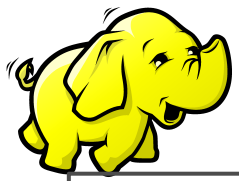
Map[Id,Map[TimeBucket,Count]]

183 -> (
 1380585600000 -> 59904,
 1383264000000 -> 62195
)

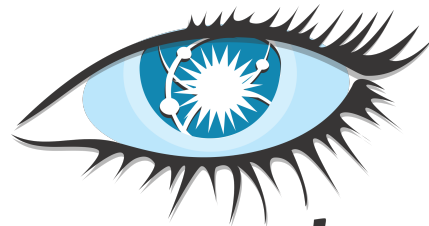
Map[Id,Map[TimeBucket,Count]]

183 -> (
 1380585600000 -> 59904,
 1383264000000 -> 62195
)

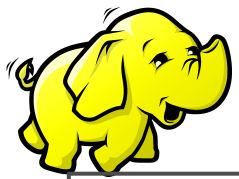




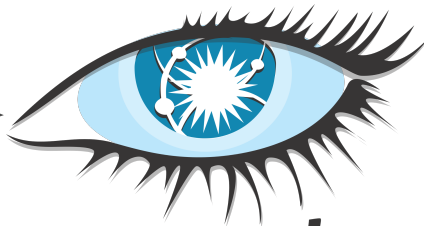
play.log



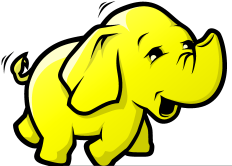
cassandra



Stitch restorator



cassandra

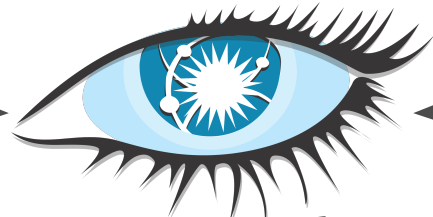


play.log

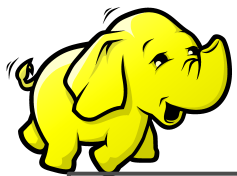


Stitch restorator

Stitch worker

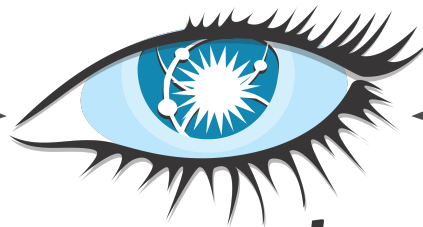


cassandra



Stitch restorator

Stitch worker



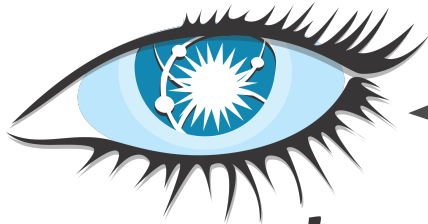
cassandra



Stitch web

 **RabbitMQ**

Stitch worker



cassandra

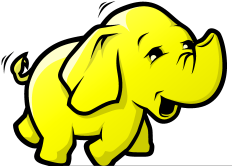


Stitch web

 **RabbitMQ**

Stitch worker





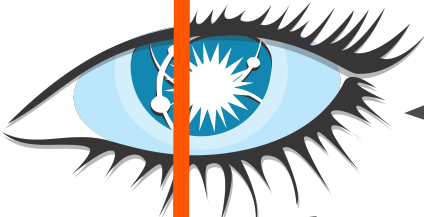
play.log

Stitch web

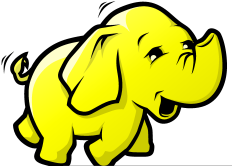
 RabbitMQ

Stitch batch

Stitch worker



cassandra



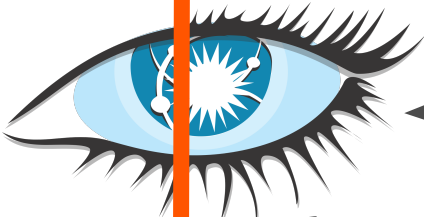
play.log

Stitch web

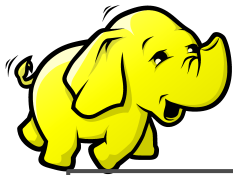
 RabbitMQ

Stitch batch

Stitch worker

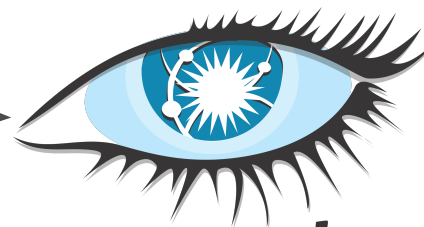


cassandra

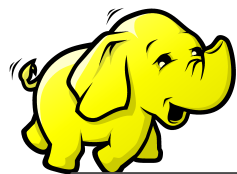


play.log

Inches job



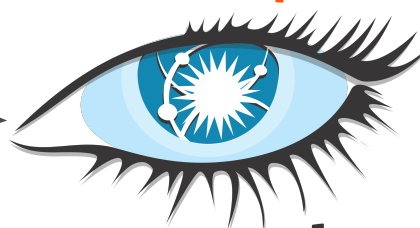
cassandra



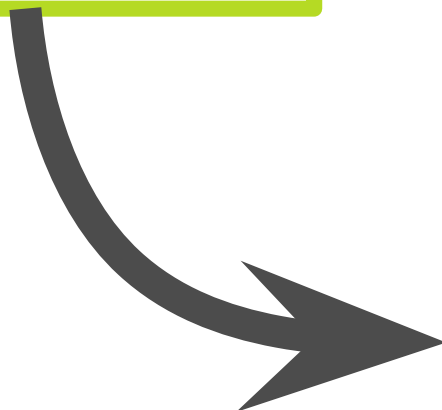
play.log

Inches job

Inches web



cassandra



Map[Id,Map[TimeBucket+Dim,Count]]

183 -> (
1380585600000|London -> 531,
1380585600000|Paris -> 651,
1383264000000|Berlin -> 123,
1383264000000|London -> 823
)

Map[Id,Map[TimeBucket+Dim,Count]]

183 -> (
1380585600000|London -> 531,
1380585600000|Paris -> 651,
1383264000000|Berlin -> 123,
1383264000000|London -> 823
)

Map[Id,Map[TimeBucket+Dim,Count]]

183 -> (
1380585600000|London -> 531,
1380585600000|Paris -> 651,
1383264000000|Berlin -> 123,
1383264000000|London -> 823
)

Map[Id,Map[TimeBucket+Pos,Dim+Count]]

**183 -> (
1380585600000|2 -> London|531,
1380585600000|1 -> Paris|651,
1383264000000|2 -> Berlin|123,
1383264000000|1 -> London|823
)**

Map[Id,Map[TimeBucket+Pos,Dim+Count]]

183 -> (

1380585600000|2 -> London|531,

1380585600000|1 -> Paris|651,

1383264000000|2 -> Berlin|123,

1383264000000|1 -> London|823

)

+ 183 -> (

1380585600000|1 -> Paris|651,

1383264000000|1 -> Berlin|123

)

Map[Id,Map[TimeBucket+Pos,Dim+Count]]

183 -> (
1380585600000|2 -> London|531,
1380585600000|1 -> Paris|651,
1383264000000|2 -> Berlin|123,
1383264000000|1 -> Berlin|123
)

Map[Id,Map[TimeBucket+Pos,Dim+Count]]

183 -> (
 1380585600000 -> binary_a,
 1383264000000 -> binary_b
)

Conclusion:
know your query

Thank you

<http://bit.ly/qcon-2015-soundcloud-cassandra>

@greenemilymay

em@soundcloud.com