

Big Data Infrastructure at Spotify

Wouter de Bie

Team Lead Data Infrastructure



June 12, 2013

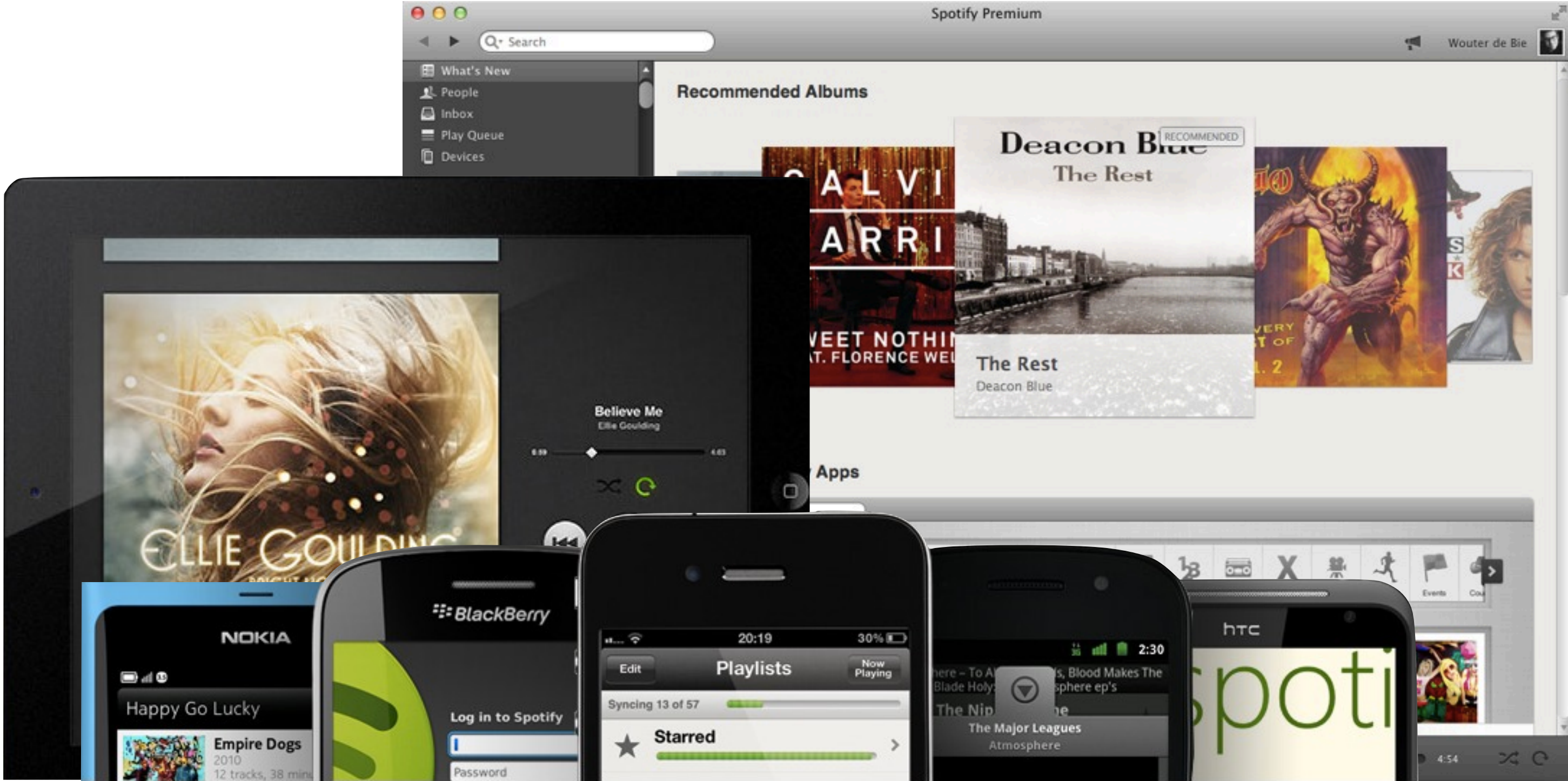
Agenda

Let's talk about Data Infrastructure, how we did it, what we learned and how we've failed

- Some **Context**
- Why **Data?**
- **Use** Cases
- Our **Infrastructure**
- Lesson **learned**



Spotify? Spotify!



Some **Context**

- Spotify started in **2006**
- Now **850+** employees, **250+** engineers
- **26 million** monthly active users
- **20+ million** tracks available
- **4** data centers across the globe
- **12** data engineers building a **platform** for easy access to data



Why data?

We play music, right?



Why data?

We were the **first** company to do free music **streaming**. But now **everybody** can do it.



Reporting

Business Analytics

Operational Analytics

Product features

Use Cases

We're a data-driven company, so data is used almost
everywhere

Reporting

- Reporting to **labels, licensors, partners** and **advertisers**
- We support **our partners**



Business Analytics

- Analyzing **growth, user behavior, sign-up funnels**, etc
- Company **KPIs**
- **A/B** testing
- **NPS** analysis
- **Segmentation** analysis



Operational metrics

- **Root** cause analysis
- **Latency** analysis
- Better **capacity planning** (servers, people, bandwidth)



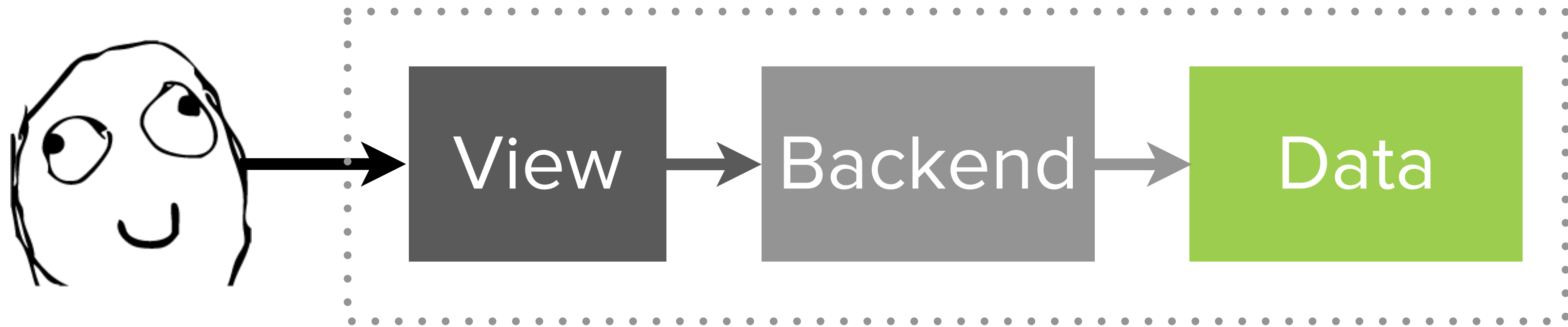
Product features

- **Radio**
- **Top** lists
- **Recommendations** (better than external parties, because of the amount of data)



Everybody should be
able to use data!

So why Data **Infrastructure**?



PROGRAMMING

PROGRAMMING

PROGRAMMING



Im in ur daturcenturz



**Our data
infrastructure**

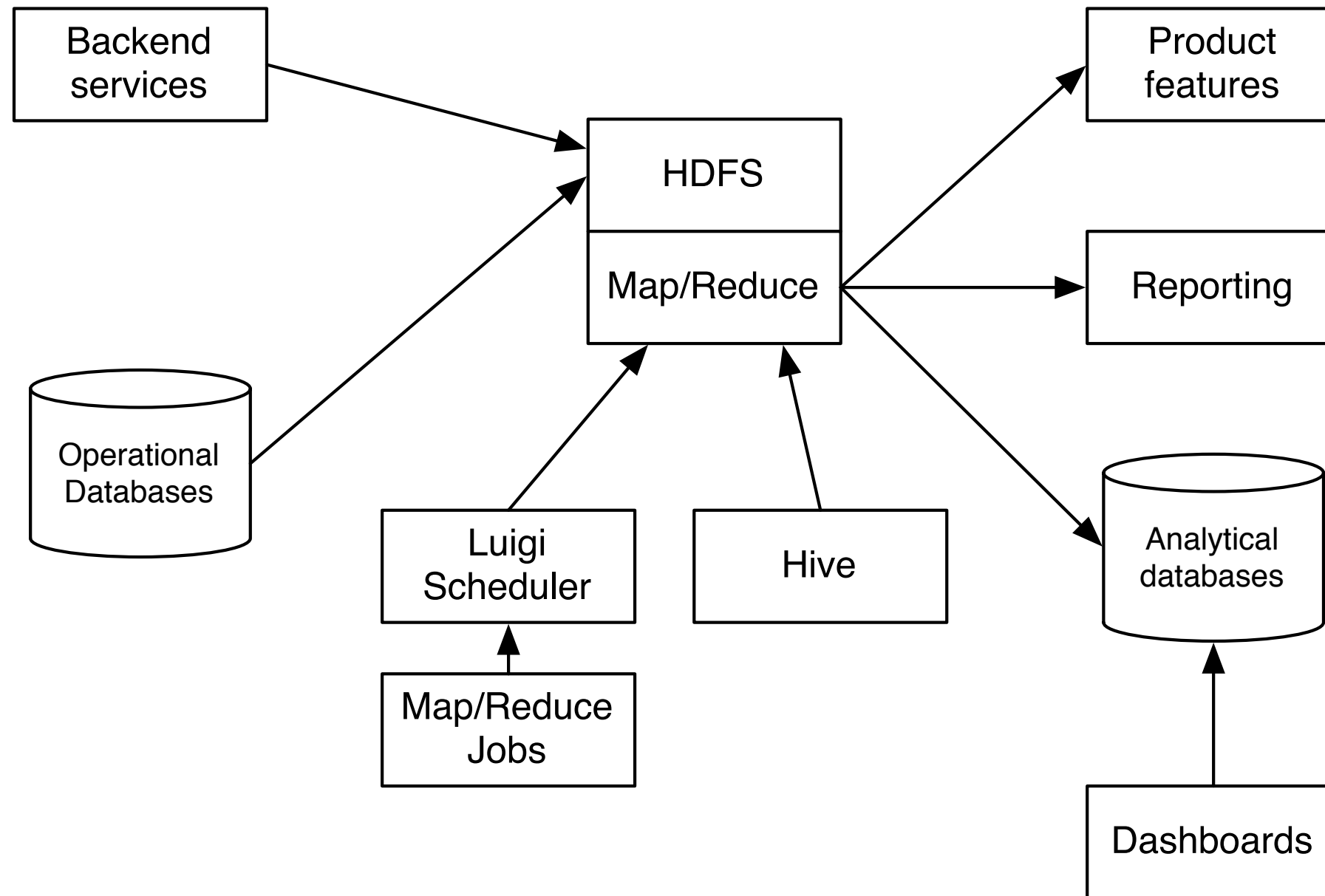
upgredin ur hadoopz



Some **geeky** numbers

- **600 GB** of compressed data from users per day
- **150 GB** of data from services per day
- **4 TB** of data generated in Hadoop each day
- **190 node** Hadoop cluster
- Soon **690 nodes**
- **4 PB** of storage capacity (soon **28 PB**)

Spotify's data infrastructure



The **thee pillars** of our Data Infrastructure



Kafka

Collection

Hadoop

Processing

Databases

Analytics/Visualization

Data **collection**



June 12, 2013

Data collection

Kafka: High volume pub-sub system

- Started with a **store-and-forward** system
- Evaluated Apache **Flume**
- Currently from **Backend-to-HDFS**, but in the future **Backend-to-Backend**
- It was almost a good fit, but...

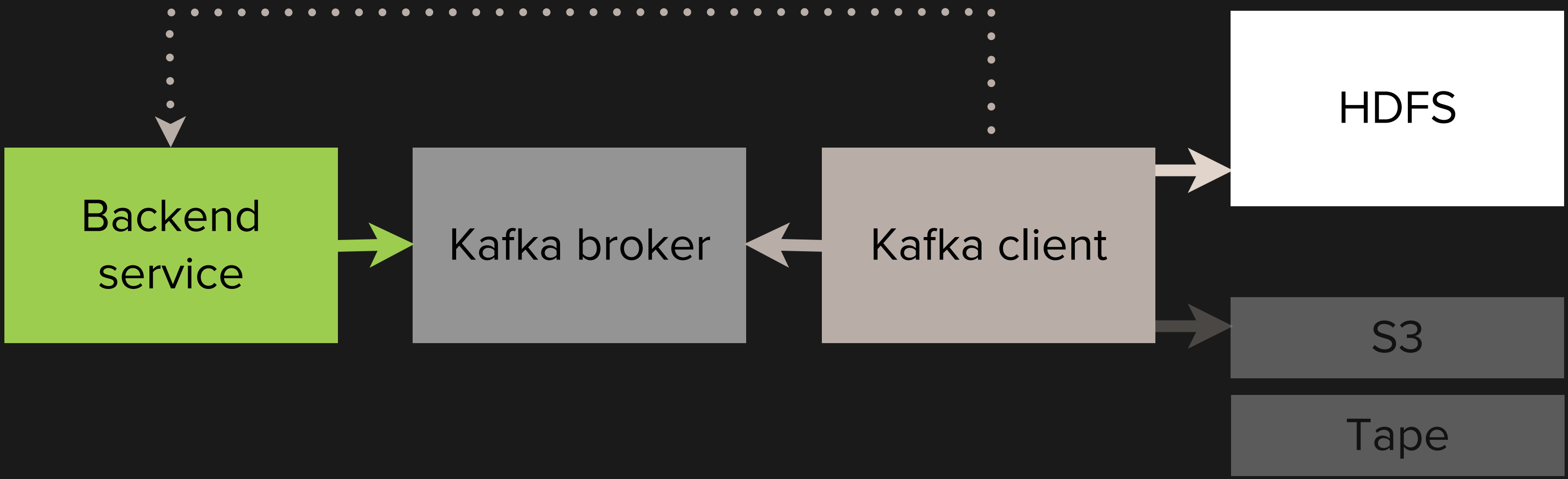
Guaranteed delivery

Kafka doesn't provide **message acknowledgements**..

- ... at least, not in **0.7 (stable)**
- **0.8** has support, but **no end-to-end** acknowledgements
- **A track streamed == monetary transaction**

Hacking Kafka

acknowledgements



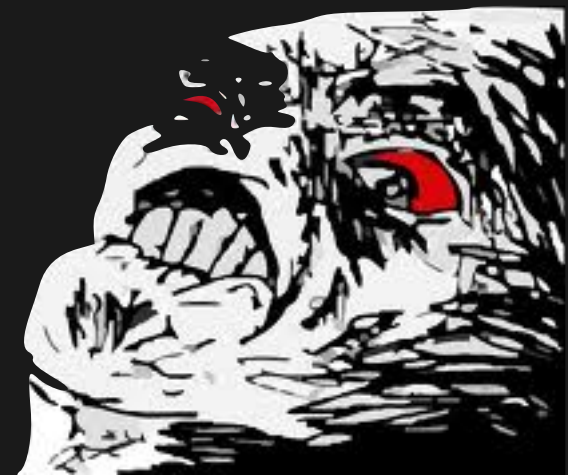
Dumping databases

Not only do we have log files, but also production databases

- Using **Sqoop** for dumping PostgreSQL
- Map/Reduce job for dumping **Cassandra**
- For large DB's we parse application logs and only dump **deltas**

Failures (or the hard stuff)

- We started with a **store-and-forward** system that didn't scale
- Kafka has multiple **components** (client, broker, ZooKeeper)
- Internet **weather**
- As with many large Java systems: **Garbage collection**





Hadoop: our trusted elephant

Scheduling

We wrote and **open-sourced** our own scheduler: Luigi

- **Nothing** suitable out there.. (unless you really, really like the XML hell of **Oozie**)
- <https://github.com/spotify/luigi>
- Written in **Python**
- **Generic** scheduler and dependency system that supports Python M/R, Pig and Hive

Map/Reduce languages

Python with Hadoop Streaming

- Pros: **fast development**, many Spotify libraries available
- Cons: **slower than Java**, no access to Hadoop API

Java

- Pros: **fast**, access to Hadoop API
- Cons: **verbose** language, not many Spotify libraries available

PIG

- Pros: very **small scripts**, **faster** than streaming
- Cons: **yet another language** to learn, not many Spotify libs available

Hive

- Pros: **SQL like** syntax (easy for non-programmers) and relational data model
- Cons: more **moving parts** (not well suited for a whole pipe line)

Scaling Hadoop at Spotify

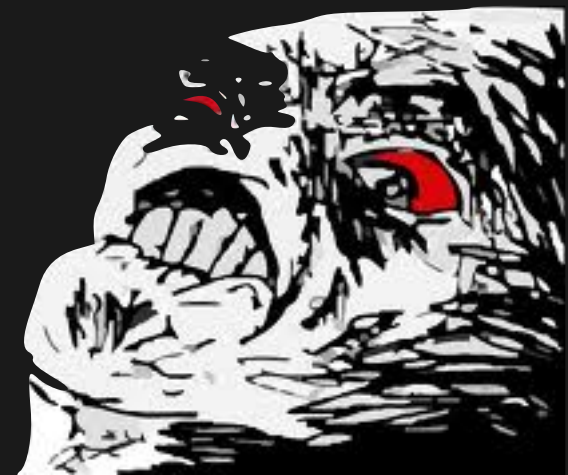
Our Journey

- Started with a small (scrap metal) cluster of **37 servers**
- Moved to Amazon **Elastic Map/Reduce** (EMR) and S3 to quickly scale
- Built an in-house cluster of **60 nodes** because of EMR costs
- Capacity planning every 6 months, grown to **190 nodes** today
- Just ordered **500 more** nodes
- Put in place **data-retention** policy and data archive



Hadoop failures

- “We just need good **developers**” - No, we need Hadoop **experts**
- We underestimated the **complexity** of Hadoop
- You can throw **money** at the problem of scaling, but at our scale, it pays off to **optimize**
- Give people **easy tools** early on



Lessons learned

4+ years of Hadoop taught us

- Hadoop has brought us **very far**. We would never be able to handle the current volume with a “cheap” RDBMS
- “**Commodity hardware**” doesn’t mean **cheap hardware**
- Hadoop isn’t a **silver bullet**
- Hadoop is a **complex** system that needs love and care
- You will have to **extend** Hadoop (and eco-system components) to tailor it to your needs

Databases and visualization

Databases: used for **aggregates**

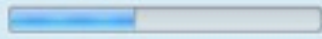

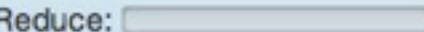
- Aggregates from Hadoop are put into **PostgreSQL** or **Cassandra**
- PostgreSQL powering **dashboards** and empowering analysts
- Cassandra for **columnar** sets (Spotify Analytics for labels)
- Databases are used for **low-latency** access by systems or analysts

Spotify Data Warehouse

Spotify Data Warehouse Jobs Data Streams Backends Wouter de Bie Help

keyword in Id Title Description Query Username Regex Case sensitive | Show: New Queued Running Success Failed Aborted

Limit: 20 Updated At ↑ Go

| Id | Title | Backend | Created by | Created at | Update at | State | Actions |
|--|---|-----------------|-----------------|----------------------|----------------------|---|----------------------------------|
| 10967 | User Search | [Hive] Hive | thomasm | 05 Jun 2013 07:17:11 | 05 Jun 2013 07:17:59 | Success 0 (3.5 sec) | Run Clone Result Download Delete |
| 10966 | Desktop Notis for WP (SE) | [Hive] Hive | tombai | 05 Jun 2013 07:15:11 | 05 Jun 2013 07:15:25 |  | Abort Clone |
| Stage 1 of 1 Job ID: job_201305010824_292108 42 mappers, 1 reducers Map:  Reduce:  | | | | | | | |
| 10810 | RememberMe statistics | [Hive] Hive | martinlljeqvist | 31 May 2013 04:14:33 | 05 Jun 2013 06:57:24 | Success 13 (118.6 sec) | Run Clone Result Download Delete |
| 10965 | select date, country, window size, | [DB] statistics | preifors | 05 Jun 2013 05:42:13 | 05 Jun 2013 06:04:03 | Success 24,233 (1309.8 sec) | Run Clone Result Download Delete |
| 10835 | Trending artist list take 2 end of 2012 | [Hive] Hive | samantha | 31 May 2013 11:24:22 | 05 Jun 2013 05:31:47 | Success 10,000 (2811.6 sec) | Run Clone Result Download Delete |
| 10956 | KPN MOB_HI DAU WAU DAU | [Hive] Hive | tynan | 04 Jun 2013 16:54:55 | 05 Jun 2013 00:39:08 | Success 452 (6944.0 sec) | Run Clone Result Download Delete |

https://datawarehouse.spotify.net

Spotify Data Warehouse

Spotify Data Warehouse Jobs ▾ Data Streams ▾ Backends ▾ Wouter de Bie ▾ Help

State: Success [View results](#) | Email not sent. Please contact system admin.

Query [Meta](#) Backend Hive Name Trending artist list take 2 end of 2012

Find A/a RegEx Replace All

```

1 select
2 p.artist,
3 p.track,
4 p.country,
5 sum(p.play_30s) as streams
6 from
7 (select
8   a.gid,
9   a.country,
10  m.track_gid,
11  get_json_object(m.json_data, '$.album.artistname') as artist,
12  get_json_object(m.json_data, '$.track.name') as track,
13  sum(a.play30s) as play_30s
14 from
15  aggregated_plays a
16 join
17  metadata m
18 on a.dt >= 20120701 and a.dt <= 20121231 and a.country = 'US' and m.dt = 20130529 and a.gid=m.track_gid
19 group by
20  a.gid,
21  a.country,
22  m.track_gid,
23  get_json_object(m.json_data, '$.album.artistname'),
24  get_json_object(m.json_data, '$.track.name')) p
25 group by
26  p.artist

```

TABLES

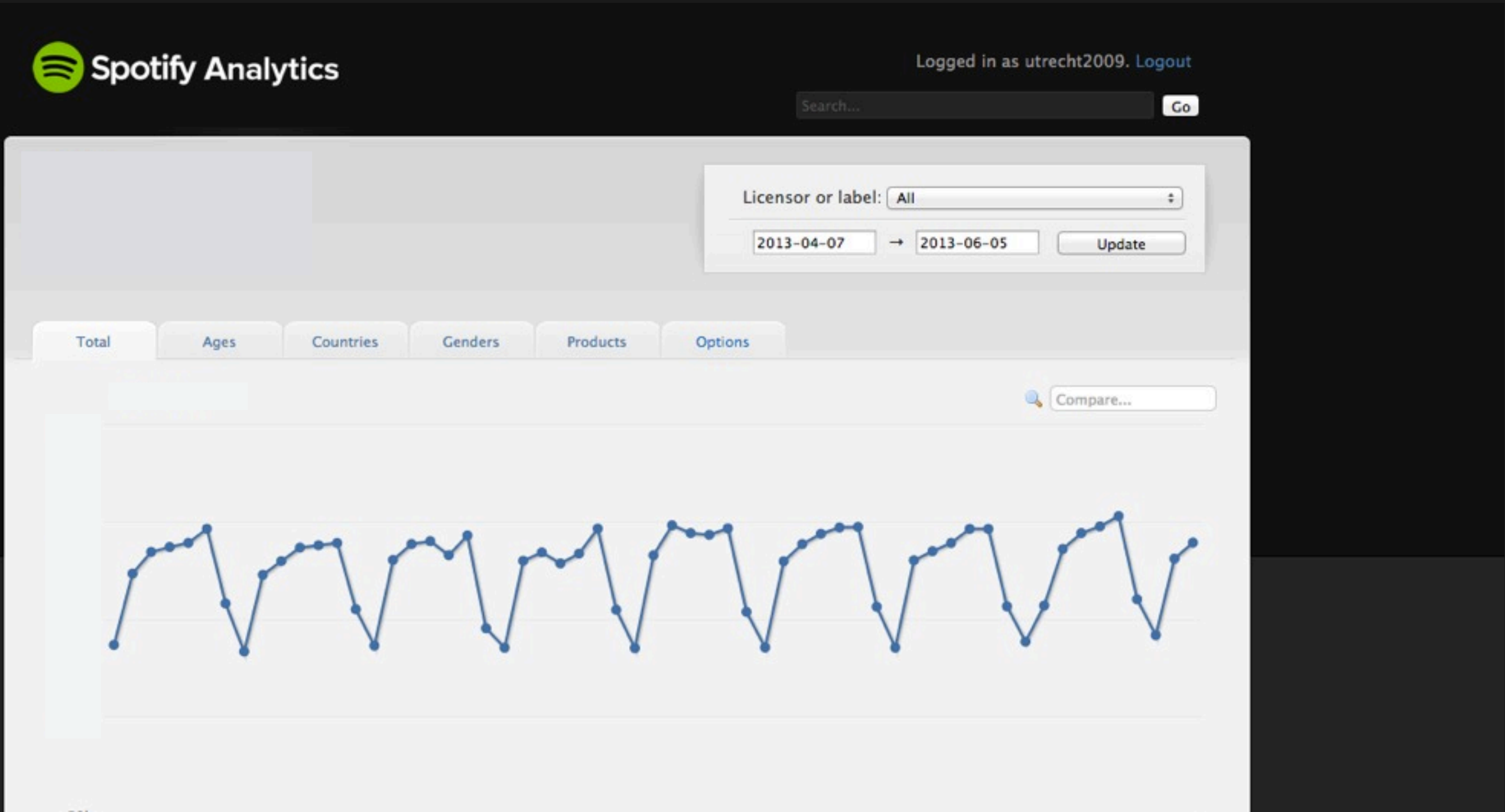
- activated_social
- active_users
- activeusers_demographics
- ad_item_impressions
- add_credit_event
- added_playlist
- aggregated_city_plays
- aggregated_download
- aggregated_plays
- album_info
- album_name
- ap_log_abnormal_exit
- ap_log_abnormal_exit_tsv
- ap_log_aggregated_request_latencies
- ap_log_application_key

Execute Check Query Save Delet

Spotify Analytics Dashboards



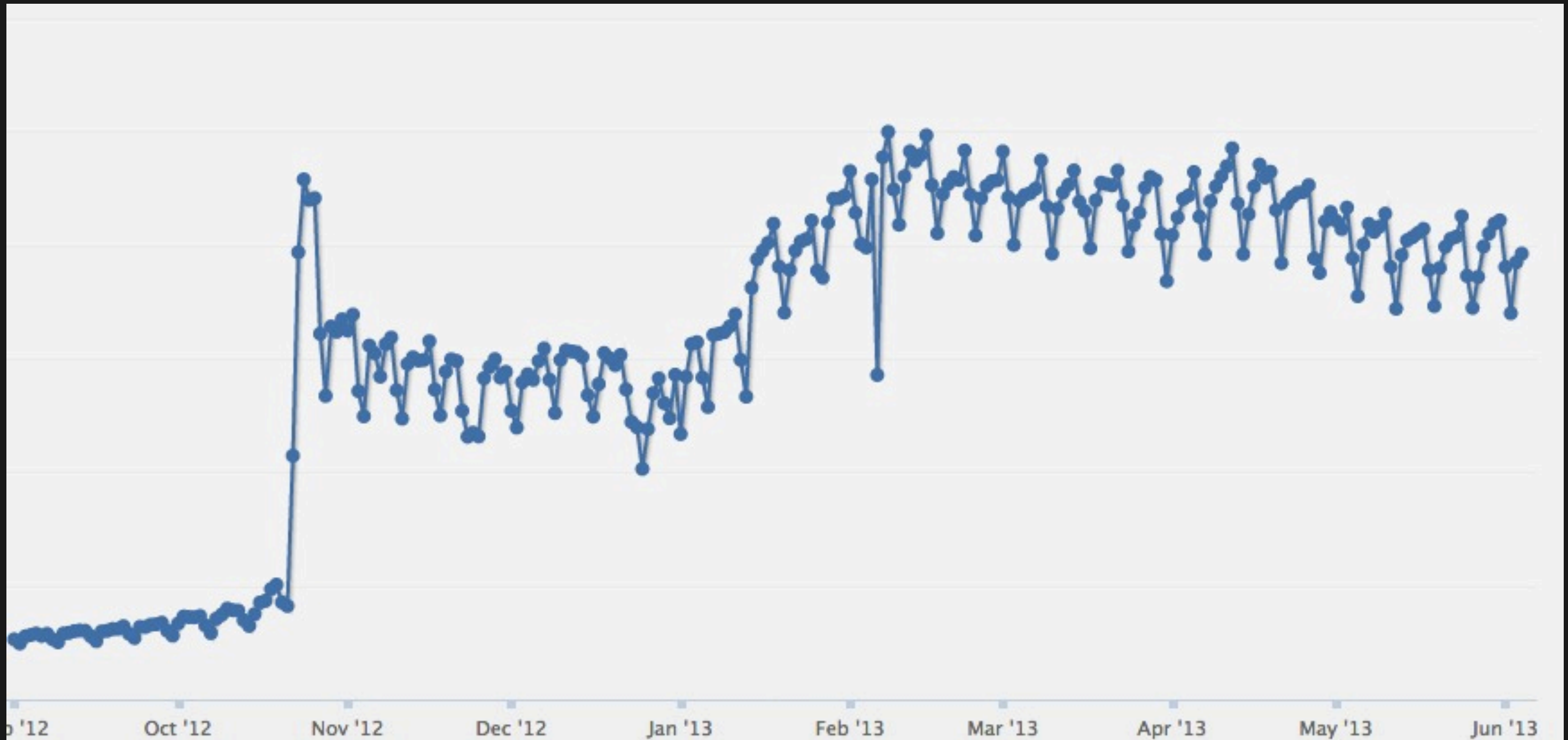
Spotify Analytics



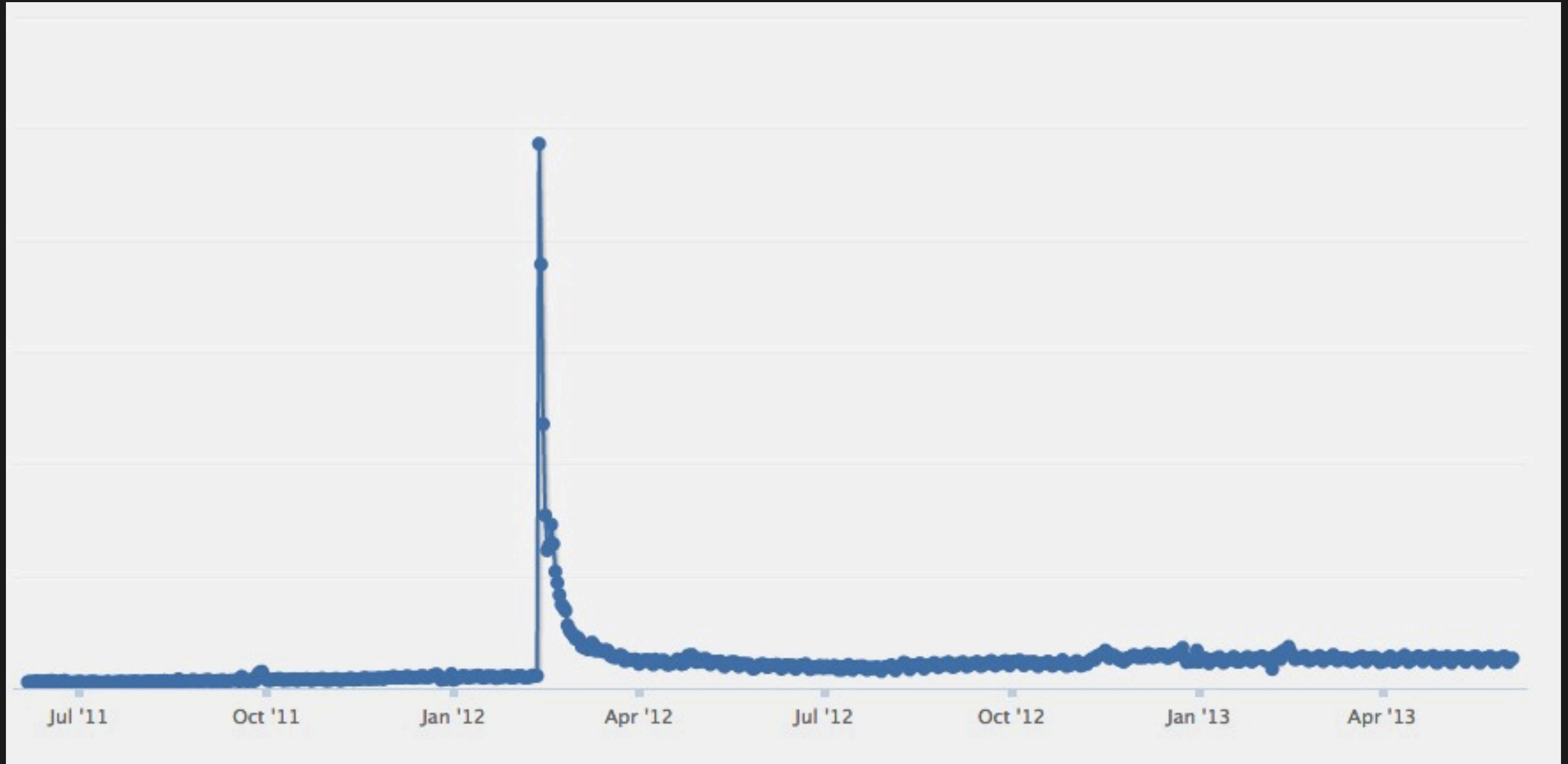
Spotify Analytics: Daft Punk



Spotify **Analytics**: Kendrick Lamar - Gábor was right :)



Spotify Analytics: Whitney Houston



Lessons learned

- Data is **crucial** for our business
- **Hadoop** and other **cutting-edge** technology work pretty well for us
- Hiring **technical analysts** was a good idea!
- There is no “**one-size-fits-all**” data product
- Spotify has the “**build it ourselves**” mindset. Sometimes it’s better to buy then build

Want to join the band?

Check out <http://www.spotify.com/jobs> or [@Spotifyjobs](https://twitter.com/Spotifyjobs) for more information.

Or mail: wouter@spotify.com

Or twitter: [@xinit](https://twitter.com/xinit)



June 12, 2013