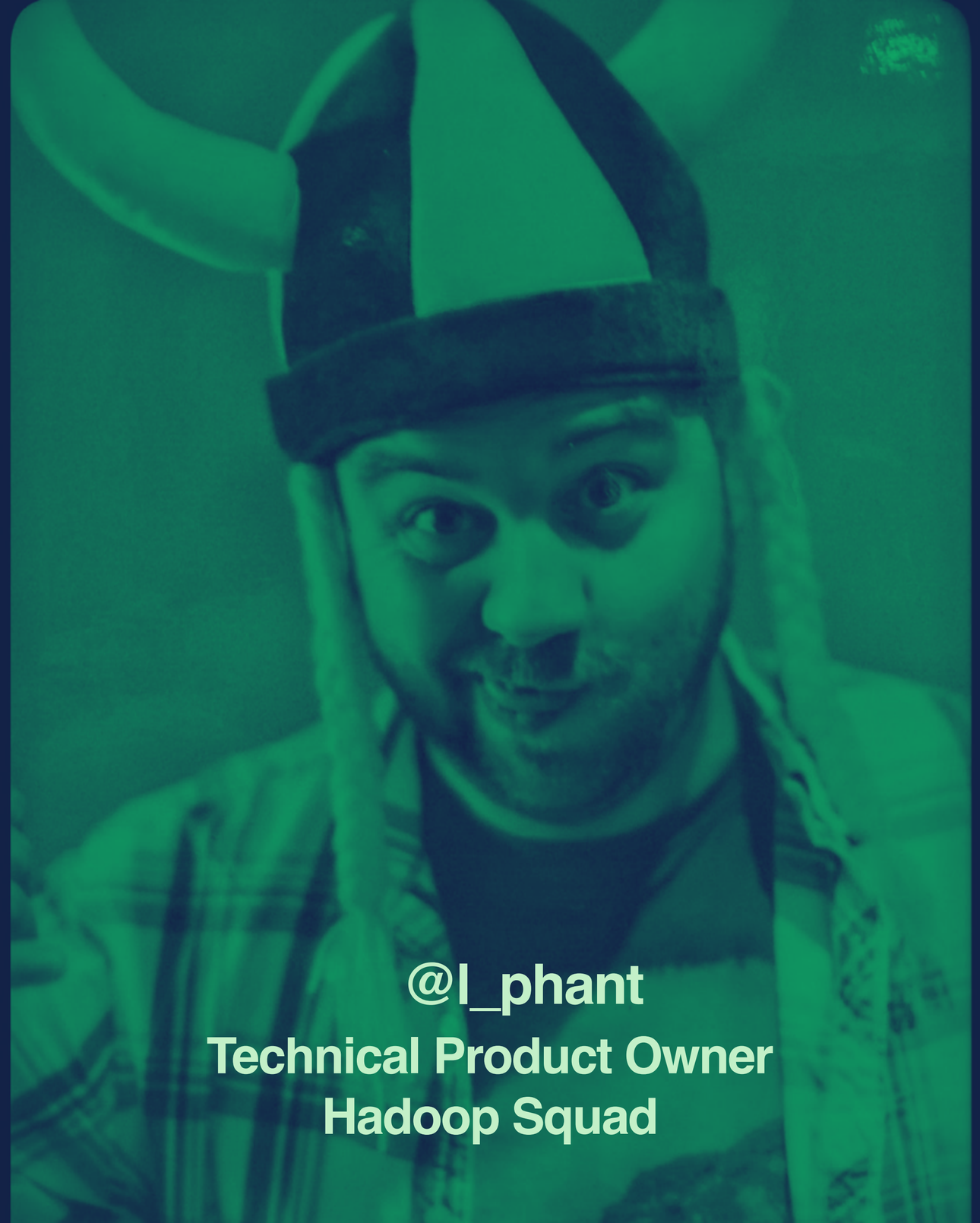
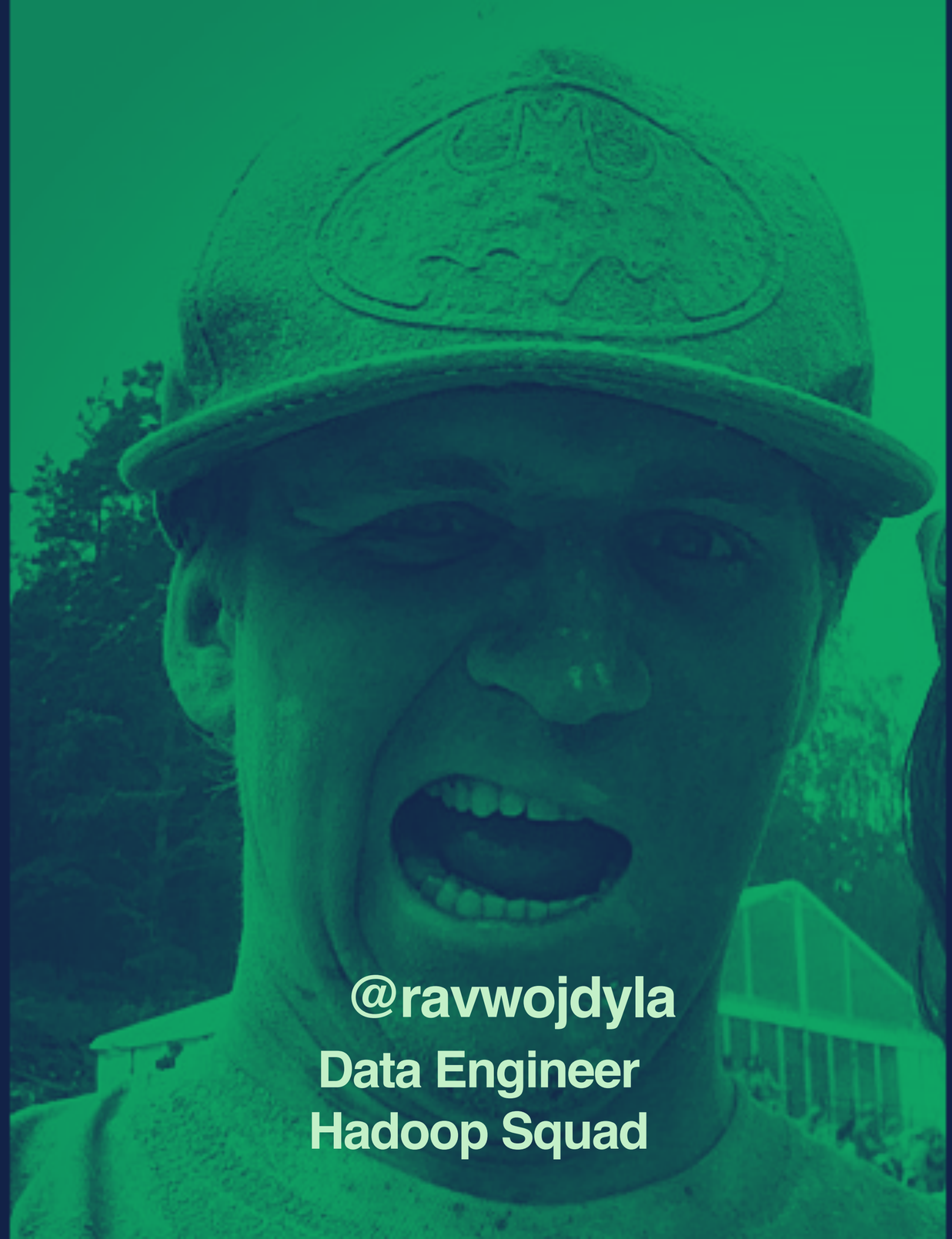
The Evolution of Hadoop at Spotify

Rafal Wojdyla (rav@spotify.com)
Josh Baer (jbx@spotify.com)



@l_phant

Technical Product Owner
Hadoop Squad



@ravwojdyla

Data Engineer
Hadoop Squad

Overview

- Growing Pains
- Gaining Focus
- The Future

Growing Pains

Spotify®

What is Spotify?

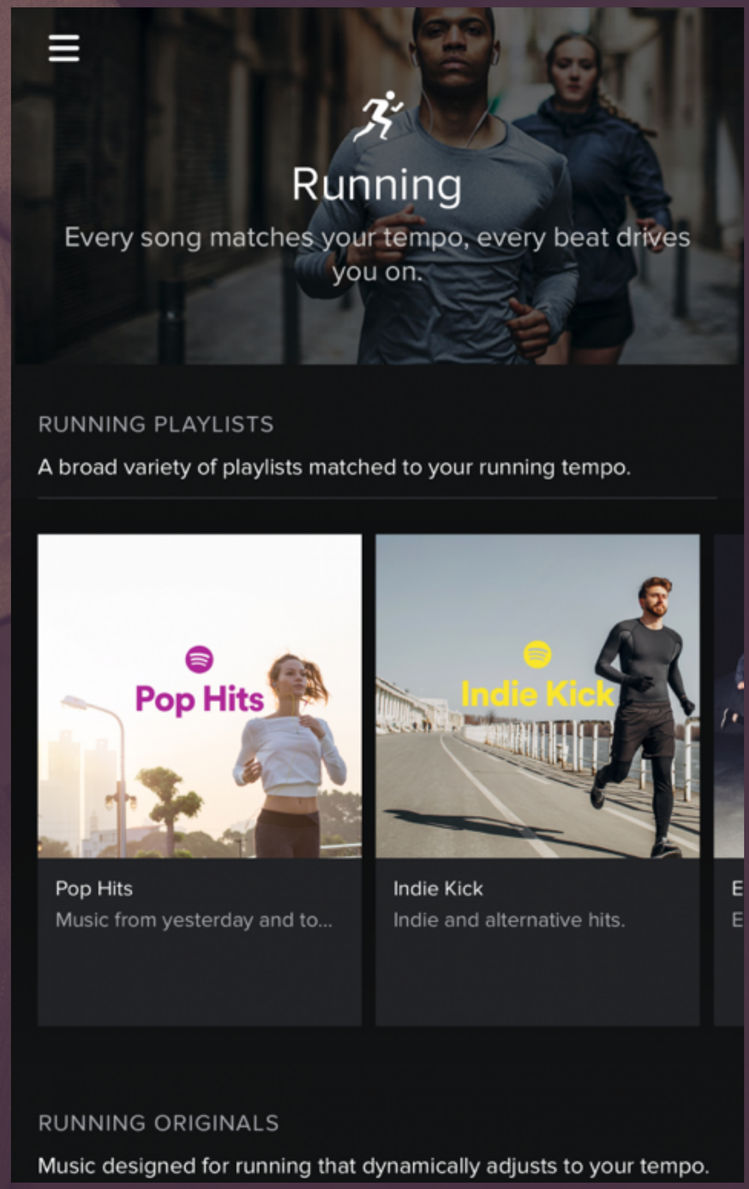
- Music Streaming Service
- Browse and Discover Millions of Songs, Artists and Albums
- Just announced
  - 75 Million Monthly Users
  - 20 Million Paid Subscribers

What is Spotify?

- **Data Infrastructure**
  - 1300 Hadoop Nodes
  - 47 PB Storage
  - 30 TB data ingested via Kafka/day
  - 400 TB generated by Hadoop/day

Powered by Data

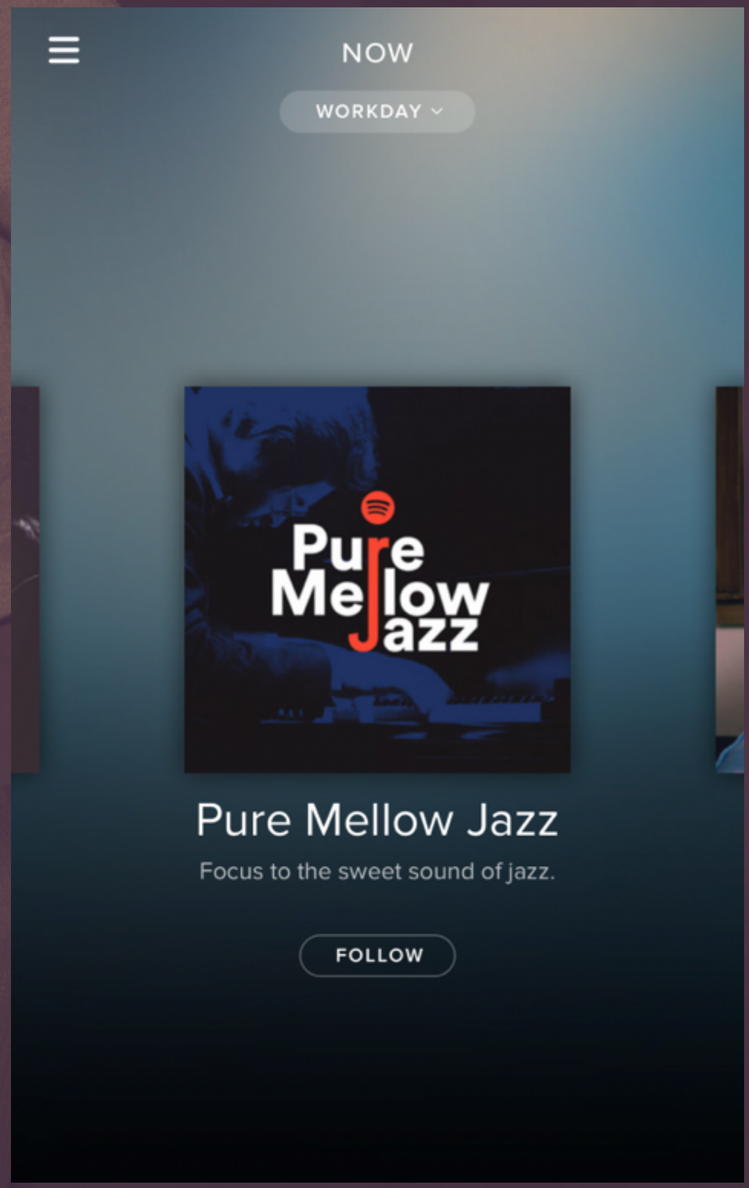
- **Running App**
  - **Matches music to running tempo**
- **Personalized running playlists in multiple tempos for millions of active users**



http://www.theverge.com/2015/6/1/8696659/spotify-running-is-great-for-discovery

Powered by Data

- **Now Page**
  - **Shows, podcasts and playlists based on day-parts**
  - **Personalized layout so you always have the right music for the right moment**



CHART

Netherlands Top 50

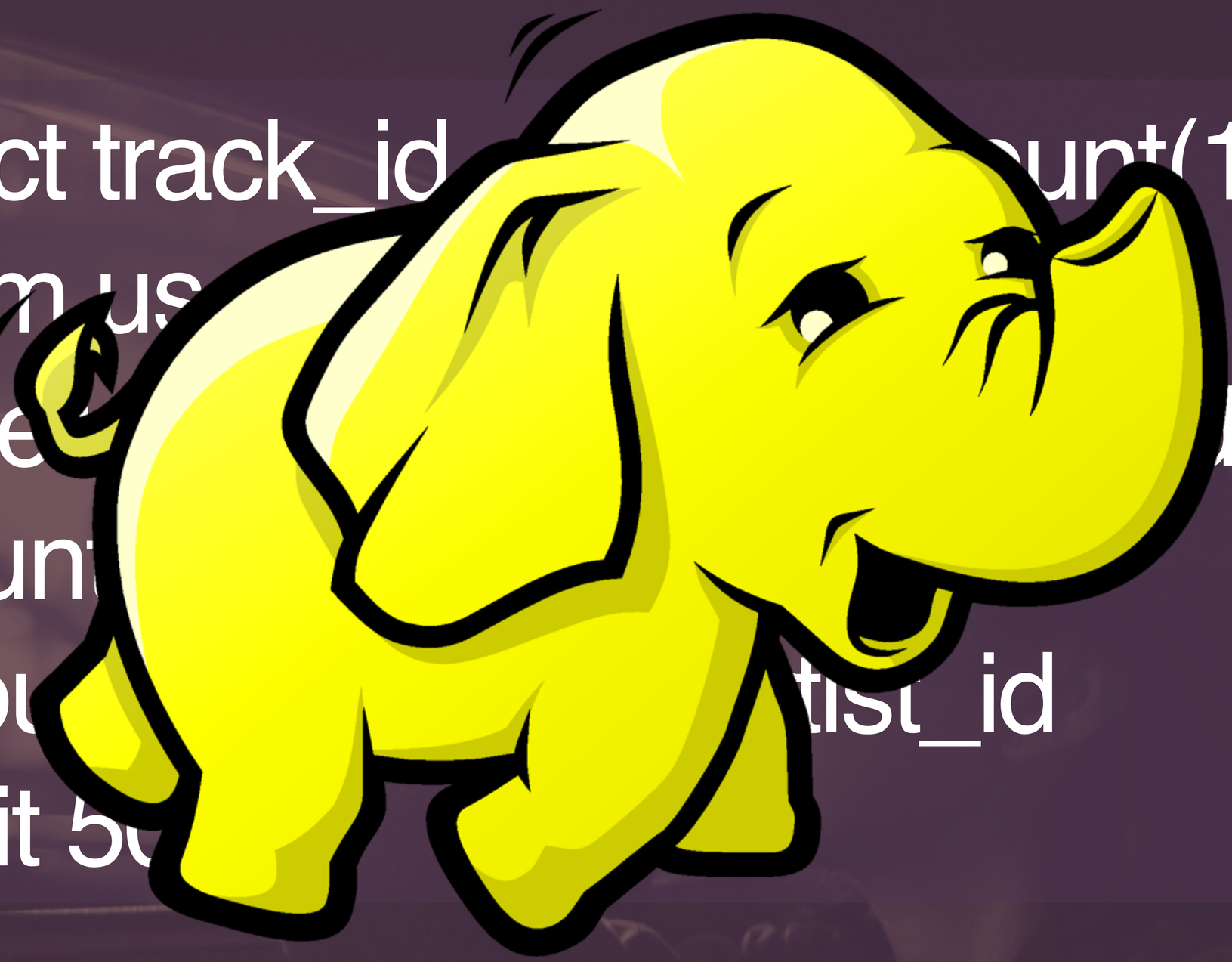
Your daily update of the most played tracks in Netherlands right now.

2 New Entries • Last Updated: an hour ago

PLAY

#			SONG		ARTIST	DAILY PLAYS
1	–	+	Drank & Drugs	EXPLICIT	Lil' Kleine, Ronnie Flex	154,343
2	–	+	**Ain't Nobody (Loves Me Better)**		**Felix Jaehn, Jasmine Thompson**	128,221
3	–	+	**Lean On (feat. MØ & DJ Snake)**		**Major Lazer, MØ, DJ Snake**	122,443
4	–	+	**Stole the Show**		**Kygo, Parson James**	105,768
5	–	+	Parijs		Kenny B	101,985
6	–	+	**See You Again (feat. Charlie Puth)**		**Wiz Khalifa, Charlie Puth**	95,835
7	–	+	**Waiting For Love**		**Avicii**	92,805
8	–	+	**Policeman**		**Eva Simons, Konshens**	89,847
9	▲	+	**Hey Mama (feat. Nicki Minaj, Bebe Rexha & Afrojack)**		**David Guetta, Bebe Rexha, Nicki Minaj, Afro...**	84,661
10	▲	+	**El Perdón**		**Nicky Jam, Enrique Iglesias**	82,633
11	▼	+	**Want To Want Me**		**Jason Derulo**	80,176
12	–	+	**Where Are Ü Now (with Justin Bieber)**		**Jack Ü, Skrillex, Diplo, Justin Bieber**	75,597

```
select track_id                                    ount(1)
  from us
  whe                                              d
  coun
  grou                                          tist_id
  limit 5
```



“It’s *simple*, we just throw the data into Hadoop”

A naive data engineer

Moving Data to Hadoop

- Raw data is complicated
  - Often dirty
  - Evolving structure
  - Duplication all over
- Getting data to a central processing point is HARD

```
10.123.133.333 - - [Mon, 3 June 2015 11:31:33 GMT] "GET /api/admin/job/
aggregator/status HTTP/1.1" 200 1847 "https://my.analytics.app/admin"
"Mozilla/5.0 (Macintosh; Intel Mac OS X 10_10_4) AppleWebKit/537.36
(KHTML, like Gecko) Chrome/43.0.2357.81 Safari/537.36"

10.123.133.222 - - [Mon, 3 June 2015 11:31:43 GMT] "GET /api/admin/job/
aggregator/status HTTP/1.1" 200 1984 "https://my.analytics.app/admin"
"Mozilla/5.0 (Macintosh; Intel Mac OS X 10_10_4) AppleWebKit/537.36
(KHTML, like Gecko) Chrome/43.0.2357.81 Safari/537.36”

10.123.133.222 - - [Mon, 3 June 2015 11:33:02 GMT] "GET /dashboard/
courses/1291726 HTTP/1.1" 304 - "https://my.analytics.app/admin"
"Mozilla/5.0 (Macintosh; Intel Mac OS X 10_10_4) AppleWebKit/537.36
(KHTML, like Gecko) Chrome/43.0.2357.81 Safari/537.36"

10.321.145.111 - - [Mon, 3 June 2015 11:33:03 GMT] "GET /api/loggedInUser
HTTP/1.1" 304 - "https://my.analytics.app/dashboard/courses/1291726"
"Mozilla/5.0 (Macintosh; Intel Mac OS X 10_10_4) AppleWebKit/537.36
(KHTML, like Gecko) Chrome/43.0.2357.81 Safari/537.36"

10.112.322.111 - - [Mon, 3 June 2015 11:33:03 GMT] "POST /api/
instrumentation/events/new HTTP/1.1" 200 2 "https://my.analytics.app/
dashboard/courses/1291726" "Mozilla/5.0 (Macintosh; Intel Mac OS X
10_10_4) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/43.0.2357.81
Safari/537.36”

10.123.133.222 - - [Mon, 3 June 2015 11:33:02 GMT] "GET /dashboard/
courses/1291726 HTTP/1.1" 304 - "https://my.analytics.app/admin"
"Mozilla/5.0 (Macintosh; Intel Mac OS X 10_10_4) AppleWebKit/537.36
(KHTML, like Gecko) Chrome/43.0.2357.81 Safari/537.36"
```

LogArchiver

- Original method to transport logs from APs to HDFS
- Lasted from 2009 - 2013
- Relies on rsync/scp and cron to move files around


FAIL

ERR, LESSON?

kafka

Log -> HDFS latency reduced from hours to seconds!

Workflow Management Fail!

```
5 * * * *    spotify-core       hadoop jar merge_hourly_logs.jar
15 * * * *   spotify-core       hadoop jar aggregate_song_plays.jar
30 * * * *   spotify-analytics  hadoop jar merge_song_metadata.jar
0 1 * * *    spotify-core       hadoop jar daily_aggregate.jar
0 2 * * *    spotify-core       hadoop jar calculate_toplist.jar
```

https://github.com/spotify/luigi

```
[data-sci@sj-edge-a1 ~] $ hdfs dfs -ls /data
Found 3 items
drwxr-xr-x   - hdfs hadoop      0 2015-01-01 12:00 lake
drwxr-xr-x   - hdfs hadoop      0 2015-01-01 12:00 pond
drwxr-xr-x   - hdfs hadoop      0 2015-01-01 12:00 ocean
[data-sci@sj-edge-a1 ~] $ hdfs dfs -ls /data/lake
Found 1 items
drwxr-xr-x   - hdfs hadoop      1321451 2015-01-01 12:00 boats.txt
[data-sci@sj-edge-a1 ~] $ hdfs dfs -cat /data/lake/boats.txt
…
```

snakebite

https://github.com/spotify/snakebite

```
$ time for i in {1..100}; do hdfs dfs -ls / > /dev/null; done
real 3m32.014s
user 6m15.891s
sys  0m18.821s
```

```
$ time for i in {1..100}; do snakebite ls / > /dev/null; done
real 0m34.760s
user 0m29.962s
sys  0m4.512s
```

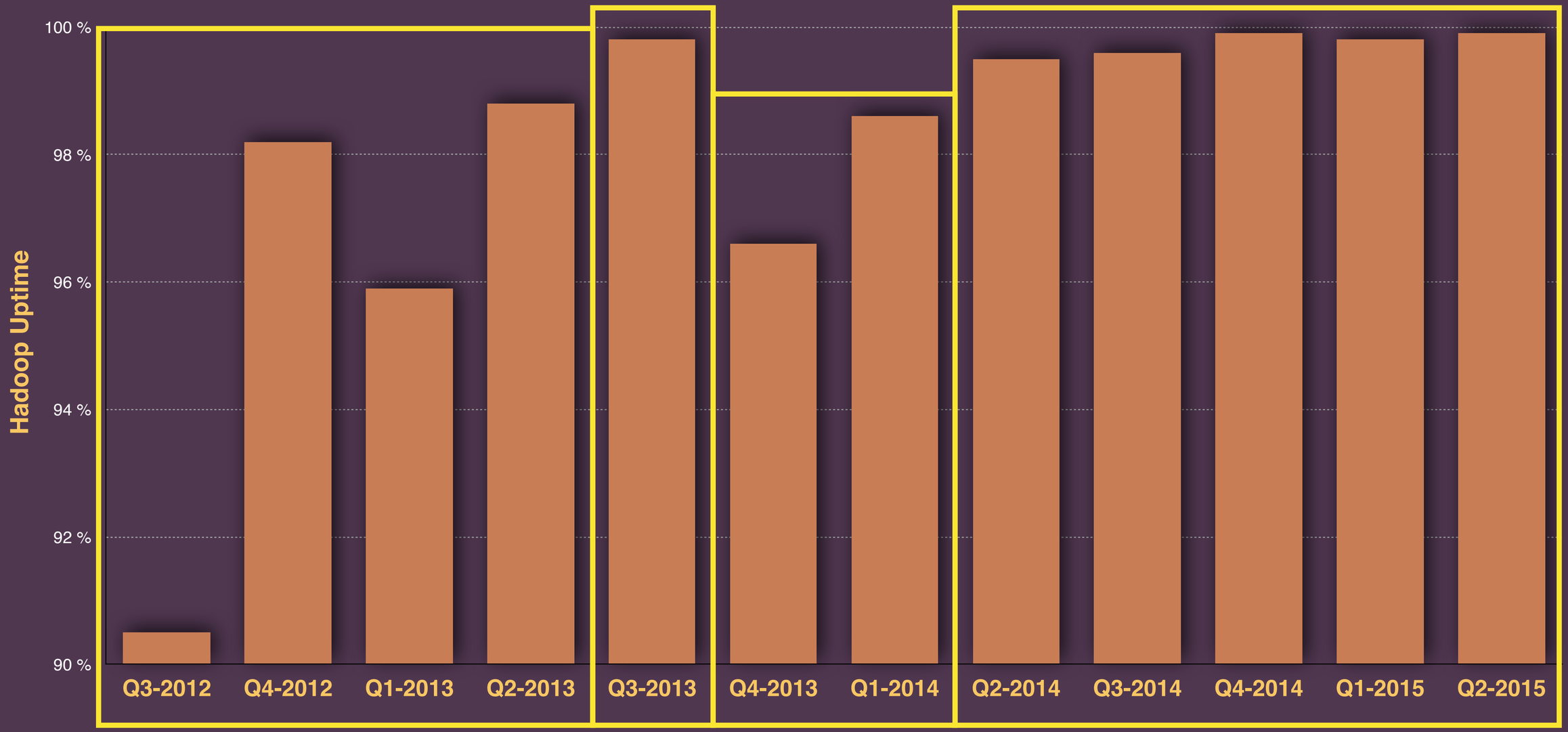
Gaining Focus

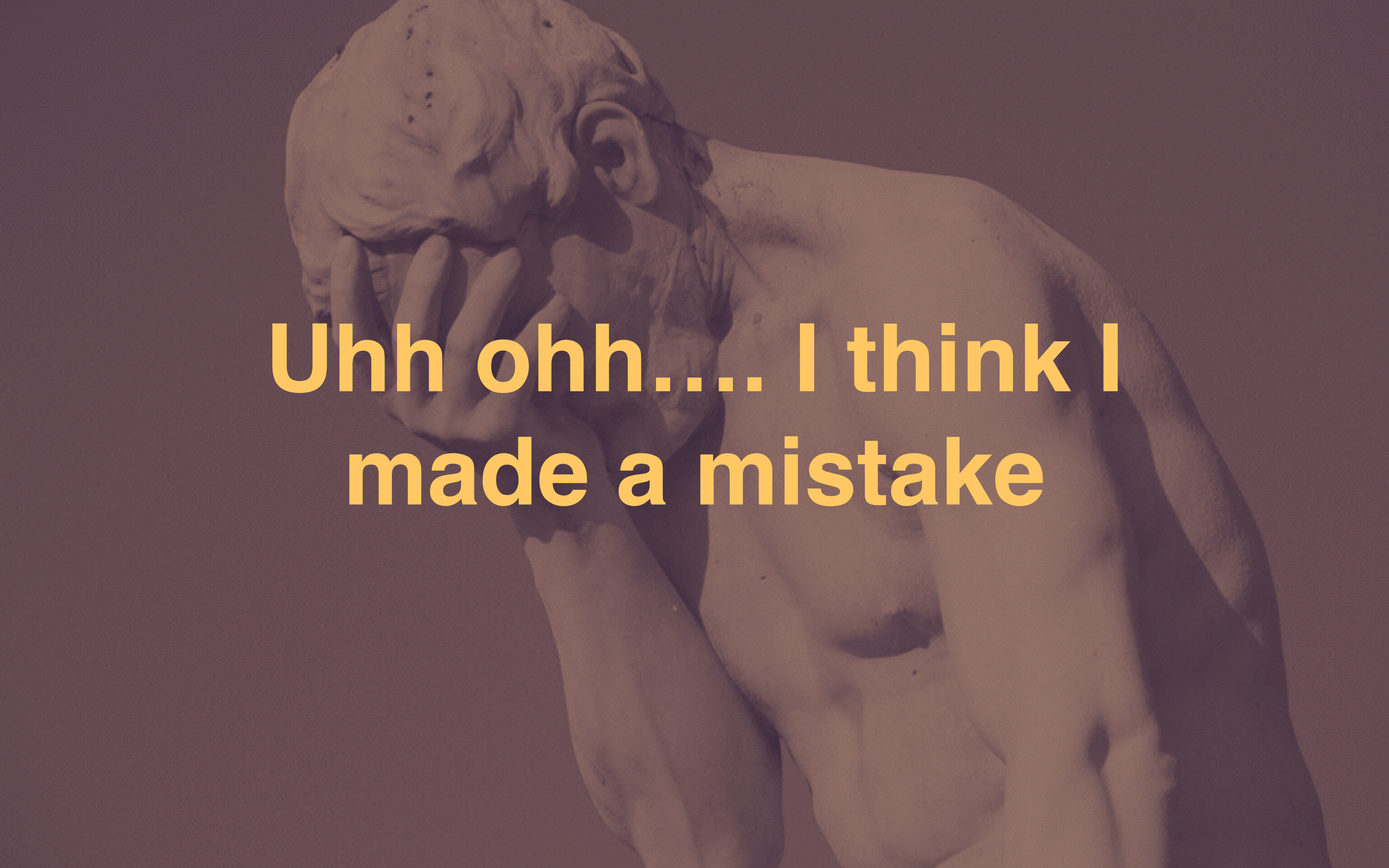
Spotify®

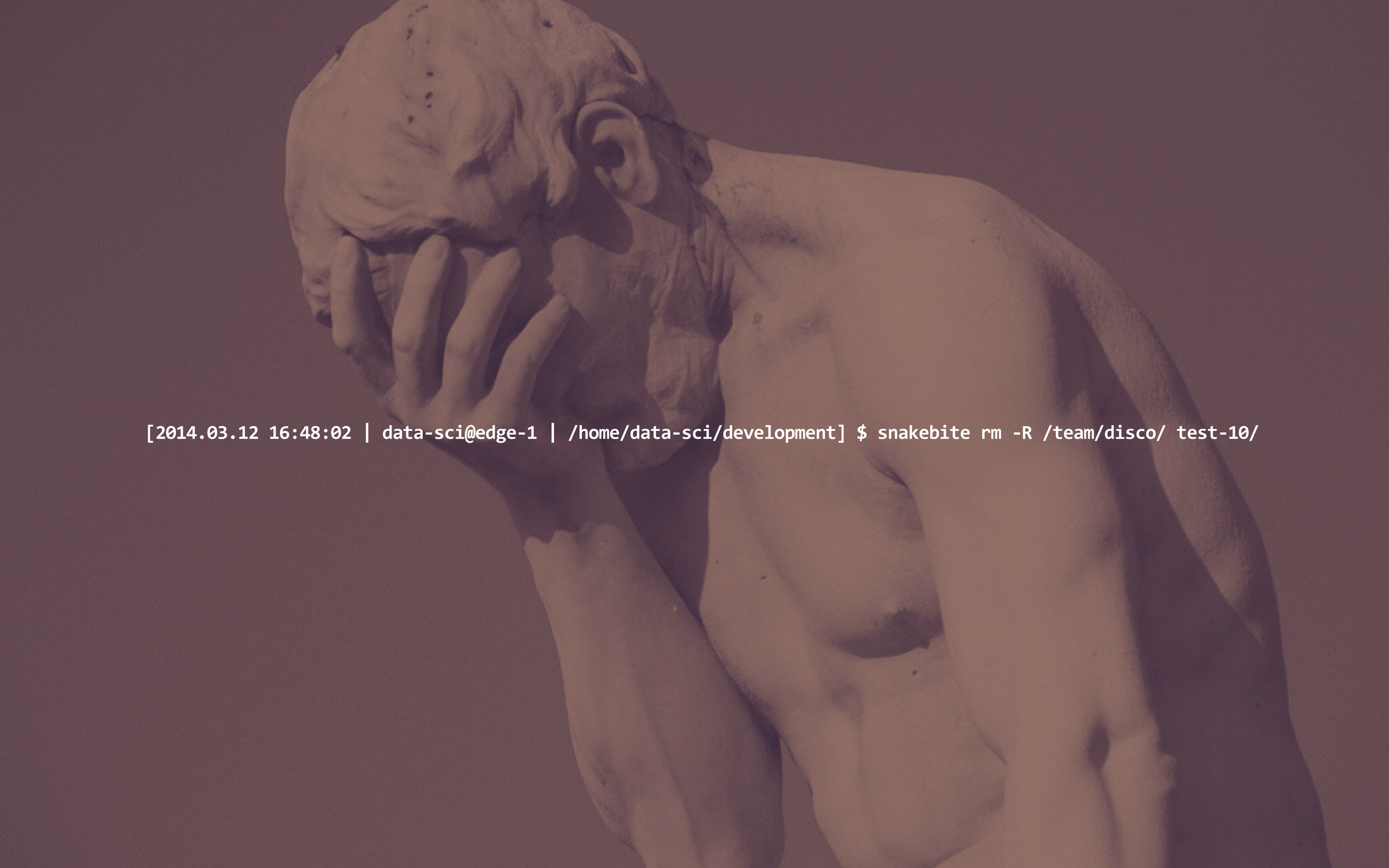
Hadoop Availability

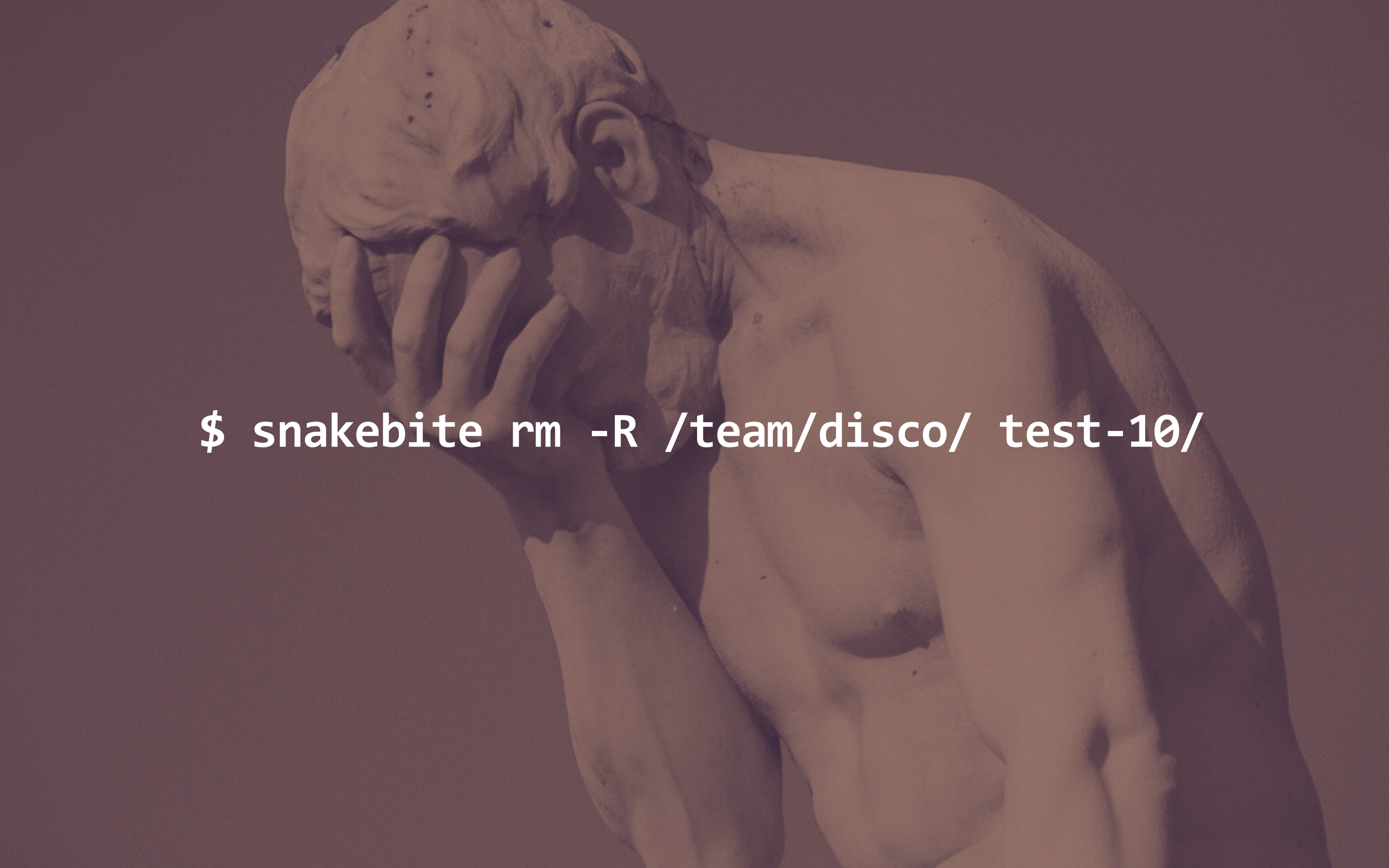
- In 2013:
  - Hadoop expanded to 200 nodes
  - Critical but not very reliable
- Created a ‘squad’ with two missions:
  - Migrate to a new distribution with Yarn
  - Make Hadoop reliable

How did we do?

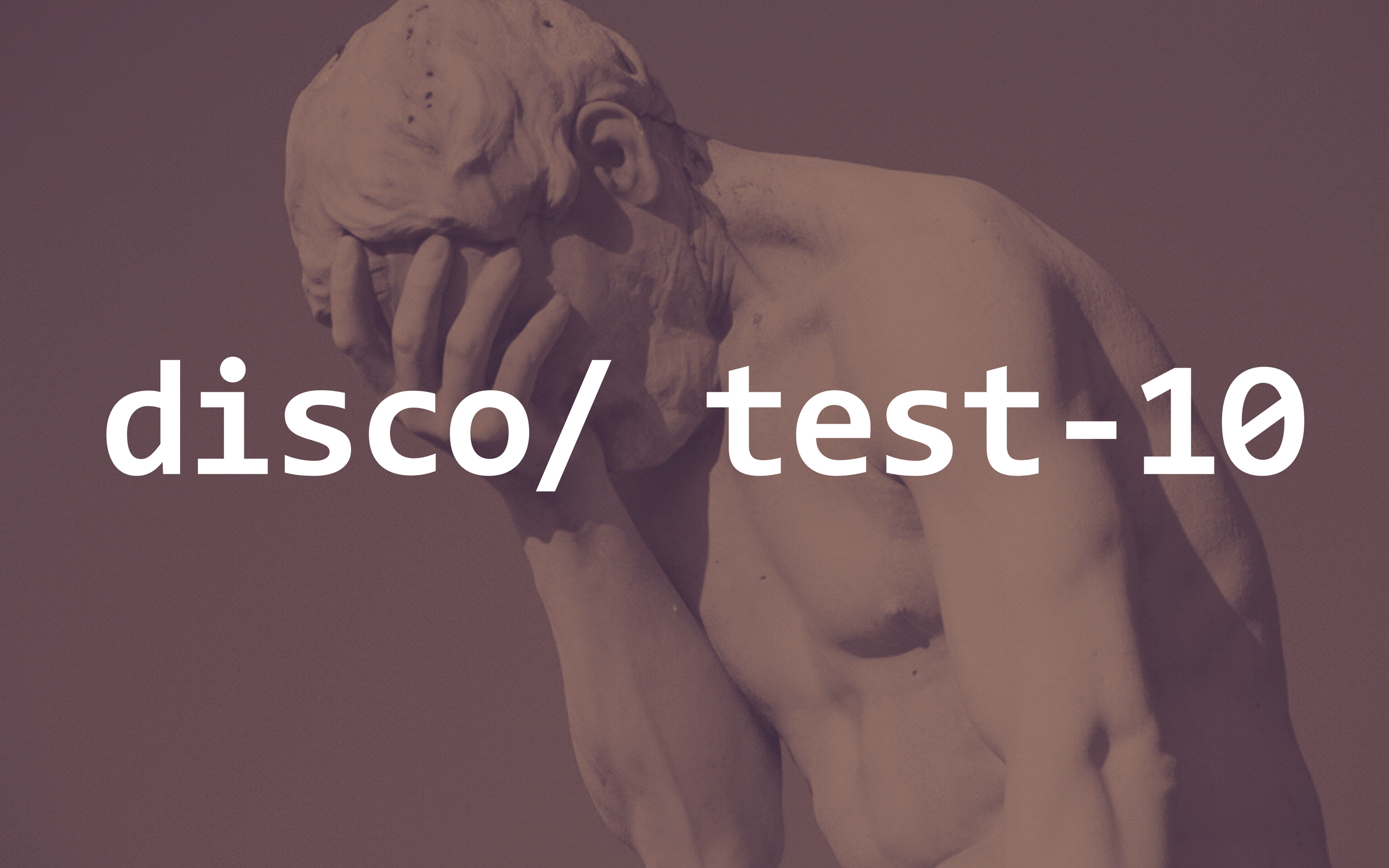

Hadoop Uptime
100 %
98 %
96 %
94 %
92 %
90 %
Q3-2012
Q4-2012
Q1-2013
Q2-2013
Q3-2013
Q4-2013
Q1-2014
Q2-2014
Q3-2014
Q4-2014
Q1-2015
Q2-2015


Uhh ohh.... I think I
made a mistake

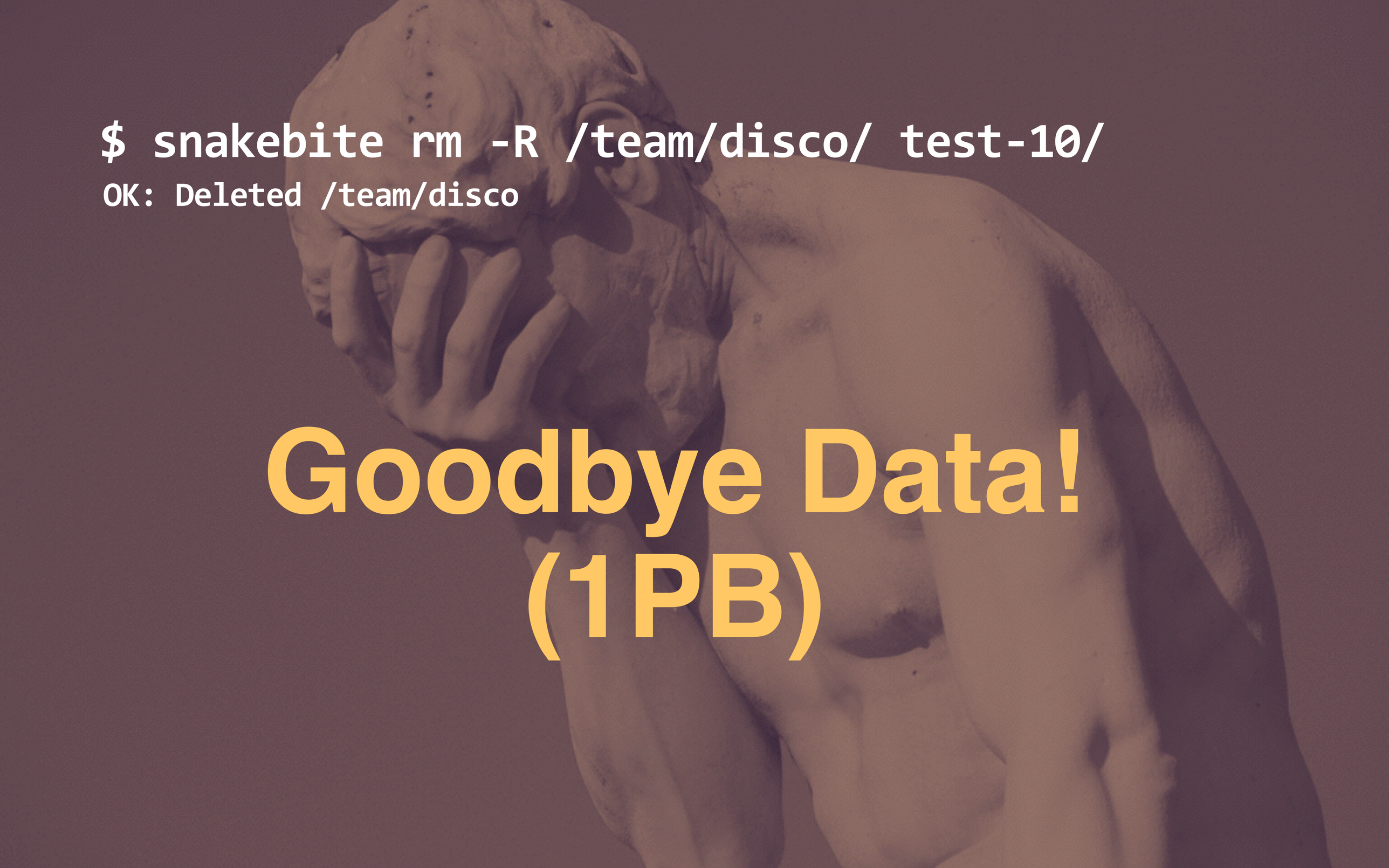

[2014.03.12 16:48:02 | data-sci@edge-1 | /home/data-sci/development] $ snakebite rm -R /team/disco/ test-10/



$ snakebite rm -R /team/disco/ test-10/


disco/ test-10

MOTHER OF GOD


$ snakebite rm -R /team/disco/ test-10/
OK: Deleted /team/disco
Goodbye Data!
(1PB)

Lessons Learned

- “Sit on your hands before you type” - Wouter de Bie
- Users will always want to retain data!
- Remove superusers from ‘edgenodes’
- Moving to trash = client-side implementation

The Wild Wild West

Pre-Production

Going from Python to Crunch

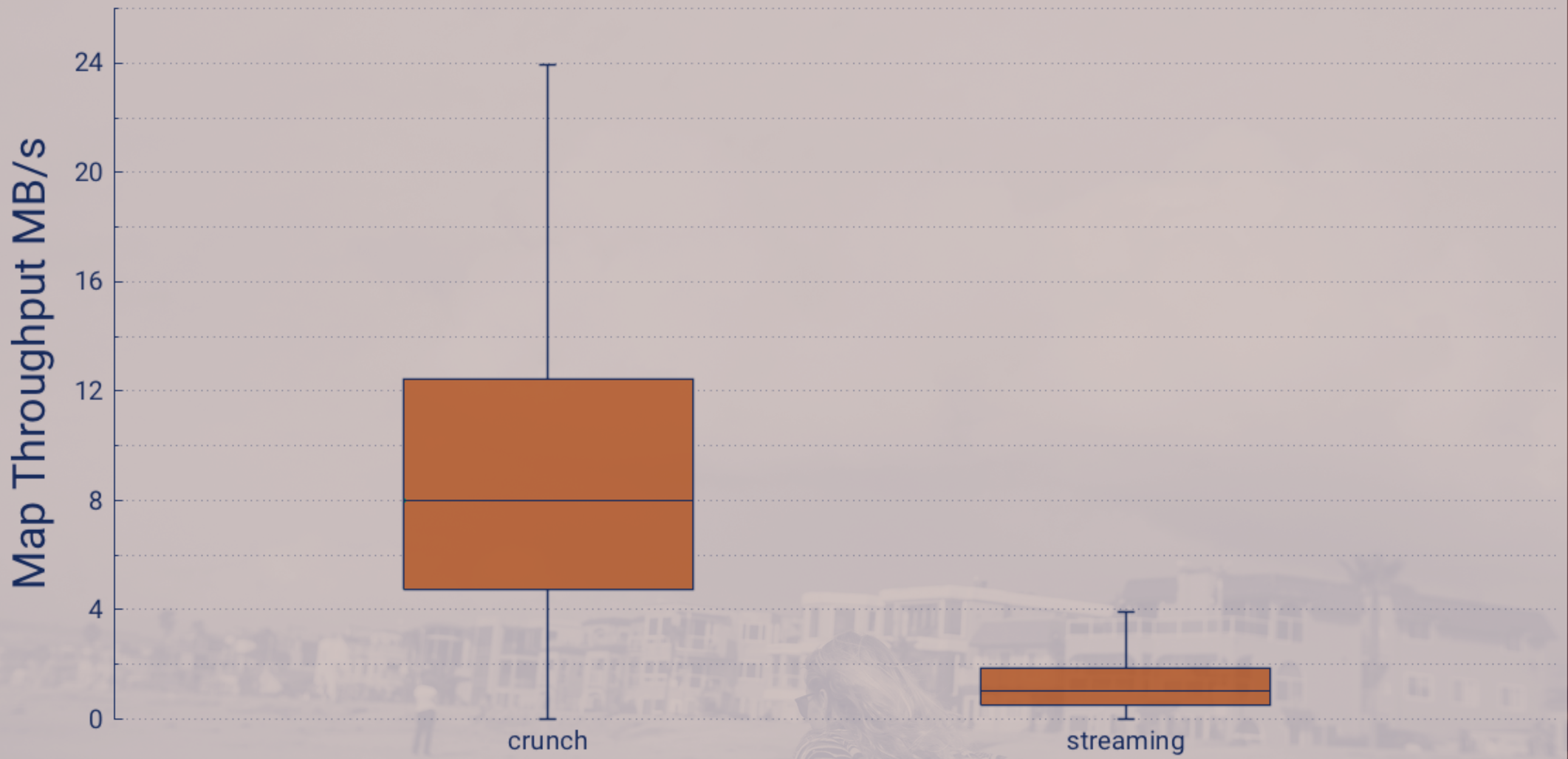
- Most of our jobs were Hadoop (python) streaming
- Lots of failures, slow performance
- Had to find a better way

Moving from Python to Crunch

- Investigated several frameworks*
- Selected Crunch:
  - Real types - compile time error detection, better testability
  - Higher level API - let the framework optimize for you
  - Better performance #JVM_FTW

*thewit.ch/scalding_crunchy_pig

Crunch vs Hadoop Streaming Benchmark

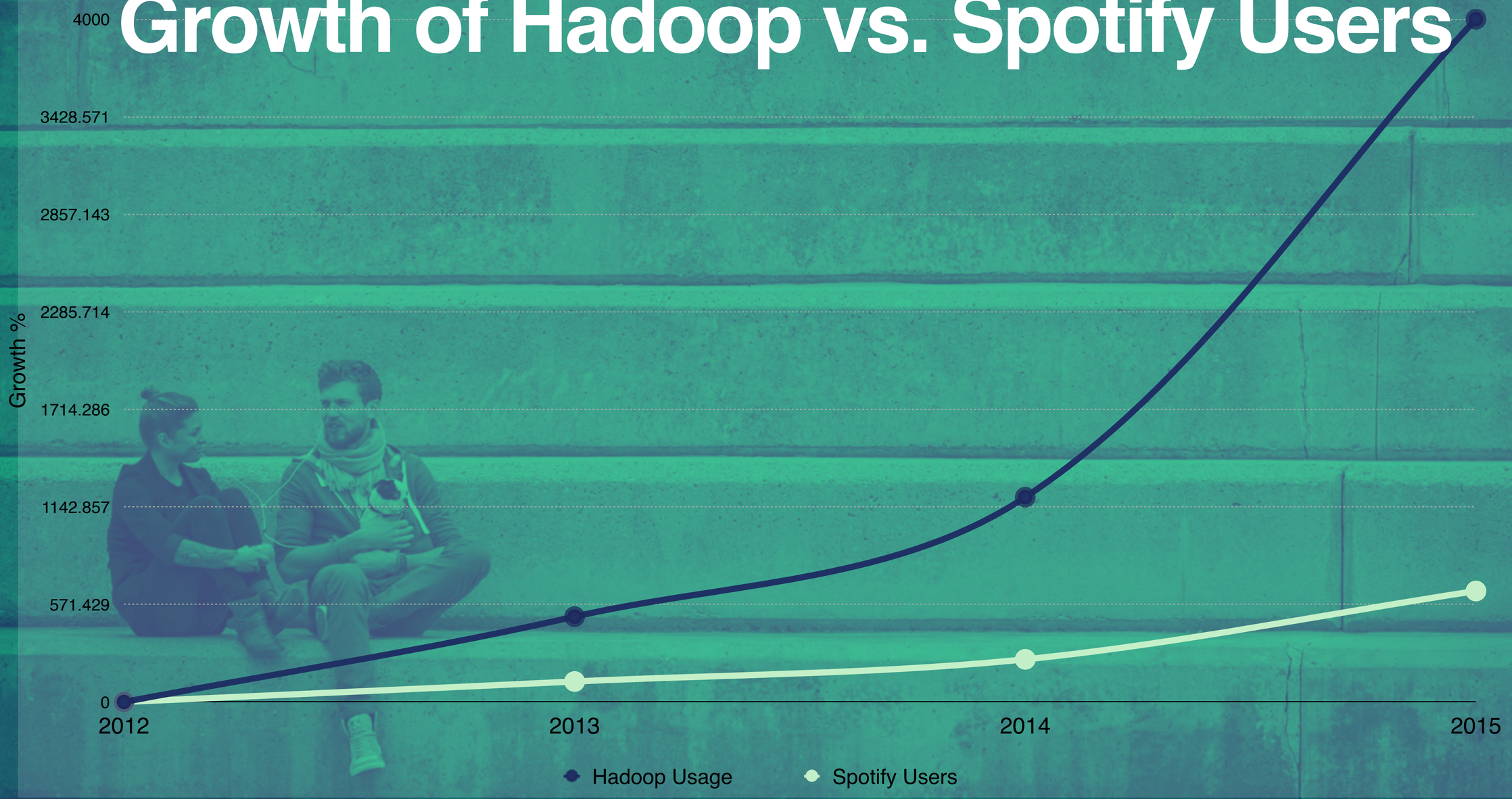


Let’s Review

- Getting data into Hadoop
- Deploying data pipelines
- Increasing availability and reliability of infrastructure
- Killing it with performance

The Future

Growth of Hadoop vs. Spotify Users



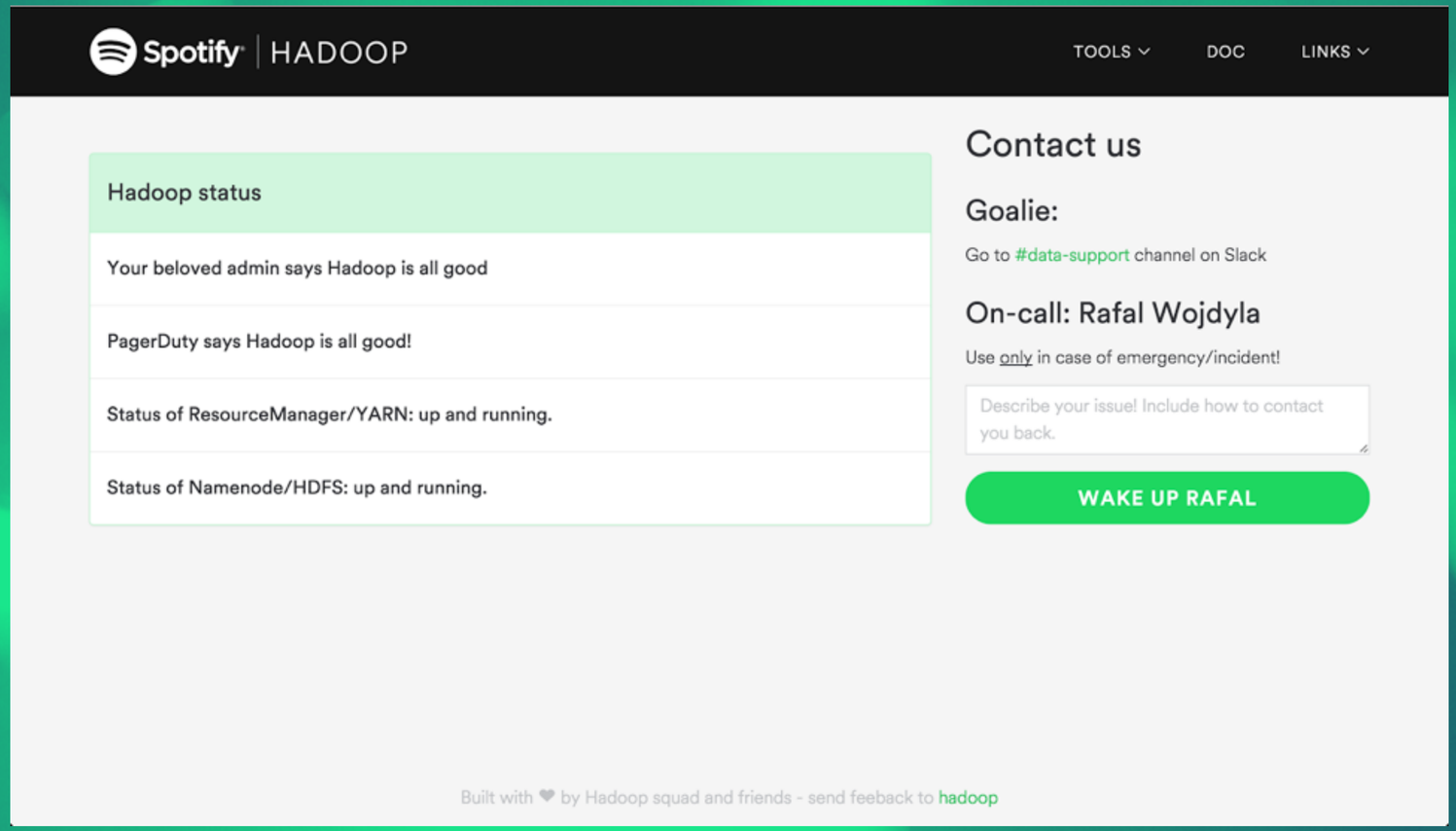
Explosive Growth

- **Increased Spotify Users**
  - More users -> more data -> longer running jobs
- **Increased Use Cases**
  - Beyond simple analytics
- **Increased Engineers**
  - Adding data scientists and data engineers

Scaling Machines: Easy
Scaling People: Hard

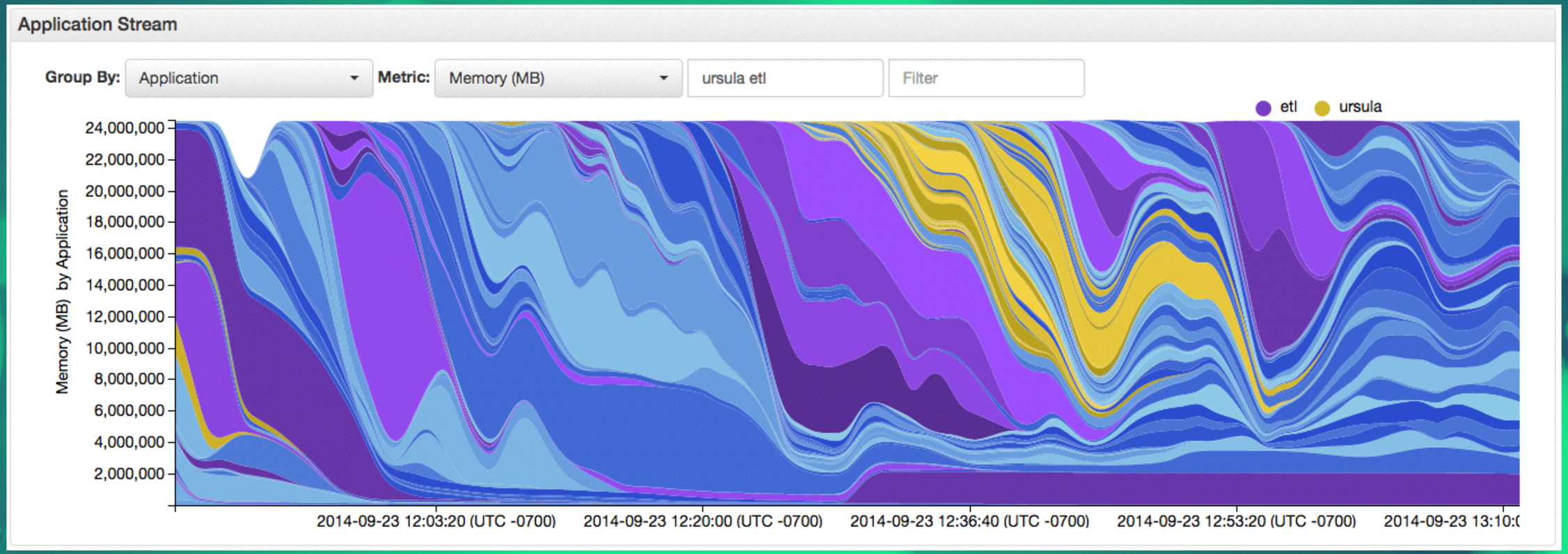
User Feedback: Automate it!

hadoop.spotify.net



Single entry point to information

Inviso



Developed by Netflix: https://github.com/Netflix/inviso

Hadoop Report Card

- Contains Statistics
- Guidelines and Best Practices
- Sent Quarterly

Real Time Use Cases

- Expanding our use of Storm for:
  - Targeting Ads based on genres
  - Quicker recommendations
- More information:
  - https://labs.spotify.com/2015/01/05/how-spotify-scales-apache-storm/

Takeaways

- There’s no golden path
  - No perfect solutions, only ones that work now!
- Big Data is constantly evolving
  - Don’t be afraid to rebuild and replace!


Join The Band!
Engineers needed in NYC, Stockholm
http://spotify.com/jobs

Bonus Slides

Hardware Profiles

- 190 nodes:

  Intel Xeon X5675 @ 3.07GHz (12 physical + HT)

  32GB RAM, 12x2TB disks

- 690 nodes:

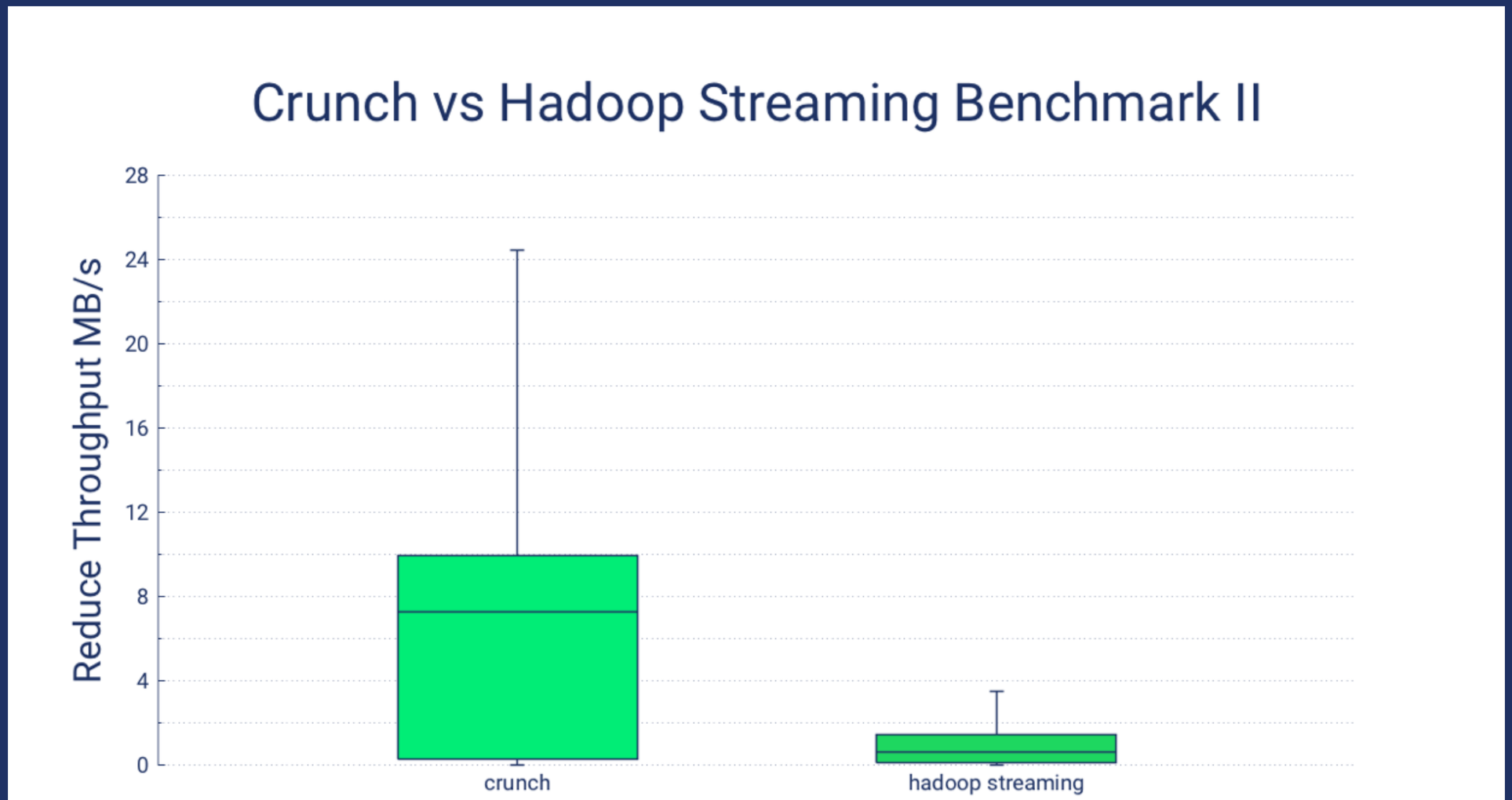
  Intel Xeon E5-2630L 0 @ 2.00GHz (12 physical + HT)

  64GB RAM, 12x4TB disks

- 400 nodes:

  Intel Xeon E5-2630L v2 @ 2.40GHz (12 physical + HT)

  96GB RAM, 12x4TB disks


Crunch vs Hadoop Streaming Benchmark II
Reduce Throughput MB/s
28
24
20
16
12
8
4
0
crunch
hadoop streaming