The Evolution of Hadoop at Spotify
Through Failures and Pain

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

Note: Our views are our own and don't necessarily represent those of Spotify.
Overview

• Growing Pains (2009-2012)
• Gaining Focus (2013 - 2014)
• The Future (2015+)
Our First Major Hadoop Bug
Cluster 1.0
What is Spotify?

- Music Streaming Service
- Browse and Discover Millions of Songs, Artists and Albums
- Launched in October 2008
- December 2014:
  - 60 Million Monthly Users
  - 15 Million Paid Subscribers
What is Spotify?

- Data Infrastructure:
  - 1300 Hadoop Nodes
  - 42 PB Storage
  - 20 TB data ingested via Kafka/day
  - 200 TB generated by Hadoop/day
select artist_id, count(1) 
from user_activities 
where play_seconds > 30 
group by artist_id;
select artist_id, count(1)
from user_activities
where play_seconds > 30
group by artist_id;
0 * * * * spotify-core hadoop jar hourly_import.jar
15 * * * * spotify-core hadoop jar hourly_listeners.jar
30 * * * * spotify-analytics hadoop jar user_funnel_hourly.jar
* 1 * * * spotify-core hadoop jar daily_aggregate.jar
* 2 * * * spotify-core hadoop jar calculate_royalties.jar
*/2 22 * * * spotify-radio hadoop jar generate_radio.jar
0  *  *  *  *  spotify-core  hadoop jar hourly_import.jar
15 *  *  *  *  spotify-core  hadoop jar hourly_listeners.jar
30 *  *  *  *  spotify-analytics hadoop jar user_funnel_hourly.jar
  * 1  *  *  *  spotify-core  hadoop jar daily_aggregate.jar
  * 2  *  *  *  spotify-core  hadoop jar calculate_royalties.jar
  */2  22  *  *  *  spotify-radio  hadoop jar generate_radio.jar
Handles the ‘plumbing’ for Hadoop jobs

https://github.com/spotify/luigi
Handles the ‘plumbing’ for Hadoop jobs

https://github.com/spotify/luigi
To the Cloud!
# sudo addgroup hadoop
# sudo adduser –ingroup hadoop hdfs
# sudo adduser –ingroup hadoop yarn
# cp /tmp/configs/*.xml /etc/hadoop/conf/
# apt-get update
...
[hdfs@sj-hadoop-b20 ~] $ apt-get install hadoop-hdfs-datanode
...
[yarn@sj-hadoop-b20 ~] $ apt-get install hadoop-yarn-nodemanager
# sudo addgroup hadoop
# sudo adduser -ingroup hadoop hdfs
# sudo adduser -ingroup hadoop yarn
# cp /tmp/configs/*.xml /etc/hadoop/conf/
# apt-get update
...
[hdfs@sj-hadoop-b20 ~] $ apt-get install hadoop-hdfs-datanode
...
[yarn@sj-hadoop-b20 ~] $ apt-get install hadoop-yarn-nodemanager
Automated Config Management
(via Puppet)
[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

[Several hdfs commands and outputs related to listing directory contents]

[Several hdfs commands and outputs related to listing directory contents]

[Several hdfs commands and outputs related to listing directory contents]

[Several hdfs commands and outputs related to listing directory contents]

[Several hdfs commands and outputs related to listing directory contents]
[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
...

$ time for i in {1..100}; do hadoop fs -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

(2013-2014)
Forming a team

- In 2013, expanded to 200 nodes
- Hadoop critical
- Needed a team totally focused on it
- Created a ‘squad’ with two missions:
  - Migrate to a new distribution with Yarn
  - Make Hadoop reliable
Hadoop Availability by Quarter

Availability %

Q3-2012  Q4-2012  Q1-2013  Q2-2013  Q3-2013  Q4-2013  Q1-2014  Q2-2014  Q3-2014  Q4-2014  Q1-2015

Hadoop ownerless  Upgrades  Getting there
Alerting
Alerting

- Alert on service level problems (i.e. no jobs running)
- Keep your alarm channel clean. Beware of alert fatigue.
Uhh ohh.....
I think I made a mistake
[data-sci@sj-edge-a1 ~] $ snakebite rm -R /team/disco/ CF/test-10/
[data-sci@sj-edge-a1 ~] $ snakebite rm -R /team/disco/ CF/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 Cluster

- Same hardware profile as production cluster
- Similar configuration
- Staging environment
- Reliable
Automated Testing
Moving Data
Apache Falcon

- Features:
  - Data discovery
  - Lineage
  - Lifecycle management
  - More
  - We use it for data movement
  - Uses Oozie behind the scenes
Improving Performance

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

• Investigated several frameworks
• Selected Crunch:
  • Real types!
  • Higher level API
  • Easier to test
  • Better performance  #JVM_FTW

*Dave Whiting’s analysis of systems: http://thewit.ch/scalding_crunchy_pig
Let’s Review
The Future

(2015+)
Note: Spotify users is based on publicly released numbers only.
Explosive Growth

- Increased Spotify Users
- Increased use cases
- Increased Engineers
Scaling machines: easy
Scaling people: hard
User Feedback: Automate IT!
Data Management
Raynor

- Data-discovery tool
- Luigi Integration
- Find and browse datasets
- View schemas
- Trace lineage

- Open-source plans? :-(

Two Takeaways

• **Automate Everything**
  • More time to play FIFA
  • build cool tools

• **Listen to your users**
  • Fail fast, don’t be afraid to scrap work
Join the Band!

Engineers wanted in NYC & Stockholm

http://spotify.com/jobs