Rainbird: Real-time Analytics @Twitter

Kevin Weil -- @kevinweil
Product Lead for Revenue, Twitter
Agenda

‣ Why Real-time Analytics?
‣ Rainbird and Cassandra
‣ Production Uses at Twitter
‣ Open Source
My Background

- Mathematics and Physics at Harvard, Physics at Stanford
- **Tropos Networks** (city-wide wireless): mesh routing algorithms, GBs of data
- **Cooliris** (web media): Hadoop and Pig for analytics, TBs of data
- **Twitter**: Hadoop, Pig, HBase, Cassandra, data viz, social graph analysis, soon to be PBs of data
My Background

- Mathematics and Physics at Harvard, Physics at Stanford
- **Tropos Networks** (city-wide wireless): mesh routing algorithms, GBs of data
- **Cooliris** (web media): Hadoop and Pig for analytics, TBs of data
- **Twitter**: Hadoop, Pig, HBase, Cassandra, data viz, social graph analysis, soon to be PBs of data

Now revenue products!
Agenda

- Why Real-time Analytics?
- Rainbird and Cassandra
- Production Uses at Twitter
- Open Source
Why Real-time Analytics

- Twitter is real-time

http://twitpic.com/135xa - There's a plane in the Hudson. I'm on the ferry going to pick up the people. Crazy.
Why Real-time Analytics

- Twitter is real-time
- ... even in space
And My Personal Favorite

@THE_REAL_SHAQ
THE_REAL_SHAQ

Anybody in portland touches me rt now will get two tickets I'm at redstar cafe

26 Mar 09 via web  Favorite  Retweet  Reply
And My Personal Favorite

@THE_REAL_SHAQ

Anybody in portland touches me rt now will get two tickets I'm at redstar cafe

26 Mar 09 via web  Favorite  Retweet  Reply

@THE_REAL_SHAQ

100 people n the prtland area just came for tickets wow portland twitterers r niiiiice

26 Mar 09 via web  Favorite  Retweet  Reply
Real-time Reporting

- Discussion around ad-based revenue model
- Help shape the conversation in real-time with Promoted Tweets
Real-time Reporting

- Discussion around ad-based revenue model
- Help shape the conversation in real-time with Promoted Tweets
- Real-time reporting ties it all together
Agenda

- Why Real-time Analytics?
- **Rainbird and Cassandra**
- Production Uses at Twitter
- Open Source
Requirements

- Extremely high write volume
- Needs to scale to 100,000s of WPS
Requirements

‣ Extremely high write volume
  Needs to scale to 100,000s of WPS

‣ High read volume
  Needs to scale to 10,000s of RPS
Requirements

- Extremely high write volume
  - Needs to scale to 100,000s of WPS
- High read volume
  - Needs to scale to 10,000s of RPS
- Horizontally scalable (reads, storage, etc)
  - Needs to scale to 100+ TB
Requirements

- Extremely high write volume
  - Needs to scale to 100,000s of WPS
- High read volume
  - Needs to scale to 10,000s of RPS
- Horizontally scalable (reads, storage, etc)
  - Needs to scale to 100+ TB
- Low latency
  - Most reads <100 ms (esp. recent data)
Cassandra

- **Pro**: In-house expertise
- **Pro**: Open source Apache project
- **Pro**: Writes are extremely fast
- **Pro**: Horizontally scalable, low latency
- **Pro**: Other startup adoption (Digg, SimpleGeo)
Cassandra

- **Pro:** In-house expertise
- **Pro:** Open source Apache project
- **Pro:** Writes are extremely fast
- **Pro:** Horizontally scalable, low latency
- **Pro:** Other startup adoption (Digg, SimpleGeo)

- **Con:** It was really young (0.3a)
Cassandra

- **Pro**: Some dudes at Digg had already started working on distributed atomic counters in Cassandra
Cassandra

- **Pro**: Some dudes at Digg had already started working on distributed atomic counters in Cassandra

- Say hi to @kelvin
Cassandra

- **Pro:** Some dudes at Digg had already started working on distributed atomic counters in Cassandra

- Say hi to @kelvin

- And @lenn0x
Cassandra

- **Pro**: Some dudes at Digg had already started working on distributed atomic counters in Cassandra
- Say hi to @kelvin
- And @lenn0x
- A dude from Sweden began helping: @skr
Cassandra

- **Pro**: Some dudes at Digg had already started working on distributed atomic counters in Cassandra

- Say hi to @kelvin

- And @lenn0x

- A dude from Sweden began helping: @skr

- Now all at Twitter :)

Thursday, February 3, 2011
Rainbird

- It counts things. Really quickly.
- Layers on top of the distributed counters patch, CASSANDRA-1072
Rainbird

- It counts things. Really quickly.
- Layers on top of the distributed counters patch, CASSANDRA-1072
- Relies on Zookeeper, Cassandra, Scribe, Thrift
- Written in Scala
Rainbird Design

- Aggregators buffer for 1m
- Intelligent flush to Cassandra
- Query servers read once written
- 1m is configurable
struct Event
{
  1: i32 timestamp,
  2: string category,
  3: list<string> key,
  4: i64 value,
  5: optional set<Property> properties,
  6: optional map<Property, i64> propertiesWithCounts
}
Rainbird Data Structures

struct Event
{
  1: i32 timestamp,
  2: string category,
  3: list<string> key,
  4: i64 value,
  5: optional set<Property> properties,
  6: optional map<Property, i64> propertiesWithCounts
}

Unix timestamp of event
struct Event
{
  1: i32 timestamp,
  2: string category,
  3: list<string> key,
  4: i64 value,
  5: optional set<Property> properties,
  6: optional map<Property, i64> propertiesWithCounts
}
Rainbird Data Structures

```cpp
struct Event {
    1: i32 timestamp,
    2: string category,
    3: list<string> key,
    4: i64 value,
    5: optional set<Property> properties,
    6: optional map<Property, i64> propertiesWithCounts
}
```
struct Event {
    1: i32 timestamp,
    2: string category,
    3: list<string> key,
    4: i64 value,
    5: optional set<Property> properties,
    6: optional map<Property, i64> propertiesWithCounts
}
struct Event
{
    1: i32 timestamp,
    2: string category,
    3: list<string> key,
    4: i64 value,
    5: optional set<Property> properties,
    6: optional map<Property, i64> propertiesWithCounts
}
Hierarchical Aggregation

- Say we’re counting Promoted Tweet impressions
  - category = pti
  - keys = [advertiser_id, campaign_id, tweet_id]
  - count = 1

- Rainbird automatically increments the count for
  - [advertiser_id, campaign_id, tweet_id]
  - [advertiser_id, campaign_id]
  - [advertiser_id]

- Means fast queries over each level of hierarchy
- Configurable in rainbird.conf, or dynamically via ZK
Hierarchical Aggregation

- Another example: tracking URL shortener tweets/clicks
  - full URL = http://music.amazon.com/some_really_long_path
  - keys = [com, amazon, music, full URL]
  - count = 1

Rainbird automatically increments the count for
  - [com, amazon, music, full URL]
  - [com, amazon, music]
  - [com, amazon]
  - [com]

- Means we can count clicks on full URLs
- And automatically aggregate over domains and subdomains!
Hierarchical Aggregation

- Another example: tracking URL shortener tweets/clicks
  - full URL = http://music.amazon.com/some_really_long_path
  - keys = [com, amazon, music, full URL]
  - count = 1

Rainbird automatically increments the count for
  - [com, amazon, music, full URL]
  - [com, amazon, music]
  - [com, amazon]
  - [com]

Means we can count clicks on full URLs
And automatically aggregate over domains and subdomains!

Thursday, February 3, 2011
Hierarchical Aggregation

Another example: tracking URL shortener tweets/clicks

full URL = http://music.amazon.com/some_really_long_path

keys = [com, amazon, music, full URL]

count = 1

Rainbird automatically increments the count for

- [com, amazon, music, full URL]
- [com, amazon, music]
- [com, amazon]
- [com]

Means we can count clicks on full URLs

And automatically aggregate over domains and subdomains!

How many people tweeted any music.amazon.com URL?
Hierarchical Aggregation

- Another example: tracking URL shortener tweets/clicks
- full URL = http://music.amazon.com/some_really_long_path
- keys = [com, amazon, music, full URL]
- count = 1

Rainbird automatically increments the count for
- [com, amazon, music, full URL]
- [com, amazon, music]
- [com, amazon]
- [com]

Means we can count clicks on full URLs
And automatically aggregate over domains and subdomains!

How many people tweeted any amazon.com URL?
Hierarchical Aggregation

- Another example: tracking URL shortener tweets/clicks
  - full URL = http://music.amazon.com/some_really_long_path
  - keys = [com, amazon, music, full URL]
  - count = 1

- Rainbird automatically increments the count for
  - [com, amazon, music, full URL]
  - [com, amazon, music]
  - [com, amazon]
  - [com]

  How many people tweeted any .com URL?

- Means we can count clicks on full URLs
- And automatically aggregate over domains and subdomains!
Temporal Aggregation

- Rainbird also does (configurable) temporal aggregation
- Each count is kept minutely, but also denormalized hourly, daily, and all time
- Gives us quick counts at varying granularities with no large scans at read time
- Trading storage for latency
Multiple Formulas

- So far we have talked about sums
- Could also store counts (1 for each event)
- ... which gives us a mean
- And sums of squares (count * count for each event)
- ... which gives us a standard deviation
- And min/max as well

- Configure this per-category in rainbird.conf
Rainbird

- Write 100,000s of events per second, each with hierarchical structure
- Query with minutely granularity over any level of the hierarchy, get back a time series
- Or query all time values
- Or query all time means, standard deviations
- Latency < 100ms
Agenda

‣ Why Real-time Analytics?
‣ Rainbird and Cassandra
‣ Production Uses at Twitter
‣ Open Source
Production Uses

- It turns out we need to count things all the time
- As soon as we had this service, we started finding all sorts of use cases for it
  - Promoted Products
  - Tweeted URLs, by domain/subdomain
  - Per-user Tweet interactions (fav, RT, follow)
  - Arbitrary terms in Tweets
  - Clicks on t.co URLs
Use Cases

- Promoted Tweet Analytics
Production Uses

- Promoted Tweet Analytics

Each different metric is part of the key hierarchy
Production Uses

- Promoted Tweet Analytics

Uses the temporal aggregation to quickly show different levels of granularity.

<table>
<thead>
<tr>
<th>Campaign</th>
<th>Start Date</th>
<th>Status</th>
<th>Impressions</th>
<th>Clicks</th>
<th>Retweets</th>
<th>Replies</th>
<th>Eng’ment Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>World Cup USA v Ghana</td>
<td>April 3, 2010</td>
<td>EXPIRED</td>
<td>155,647</td>
<td>2,794</td>
<td>75</td>
<td>36</td>
<td>1.87%</td>
</tr>
<tr>
<td>Cristiano Ronaldo</td>
<td>April 9, 2010</td>
<td>EXPIRED</td>
<td>120,944</td>
<td>3,574</td>
<td>116</td>
<td>31</td>
<td>3.08%</td>
</tr>
<tr>
<td>#worldcup</td>
<td>March 19, 2010</td>
<td>EXPIRED</td>
<td>116,637</td>
<td>5,969</td>
<td>217</td>
<td>31</td>
<td>5.33%</td>
</tr>
<tr>
<td>Donovan</td>
<td>March 18, 2010</td>
<td>EXPIRED</td>
<td>41,931</td>
<td>2,529</td>
<td>76</td>
<td>12</td>
<td>6.24%</td>
</tr>
<tr>
<td>World Cup Germany v Argentina</td>
<td>March 20, 2010</td>
<td>EXPIRED</td>
<td>34,163</td>
<td>387</td>
<td>19</td>
<td>8</td>
<td>1.21%</td>
</tr>
<tr>
<td>#wc2010</td>
<td>March 22, 2010</td>
<td>EXPIRED</td>
<td>11,039</td>
<td>793</td>
<td>18</td>
<td>6</td>
<td>7.40%</td>
</tr>
</tbody>
</table>
Production Uses

‣ Promoted Tweet Analytics

Data can be historical, or from 60 seconds ago

Thursday, February 3, 2011
Production Uses

› Internal Monitoring and Alerting

- We require operational reporting on all internal services
- Needs to be real-time, but also want longer-term aggregates
- Hierarchical, too: [stat, datacenter, service, machine]
Production Uses

- Tweet Button Counts

Tweet Button counts are requested many many times each day from across the web

- Uses the all time field
Agenda

‣ Why Real-time Analytics?
‣ Rainbird and Cassandra
‣ Production Uses at Twitter
‣ Open Source
Open Source?

- Yes!
Open Source?

- Yes! ... but not yet
Open Source?

- Yes! ... but not yet
- Relies on unreleased version of Cassandra
Open Source?

- Yes! ... but not yet
- Relies on unreleased version of Cassandra
- ... but the counters patch is committed in trunk (0.8)
Open Source?

- Yes! ... but not yet
- Relies on unreleased version of Cassandra
- ... but the counters patch is committed in trunk (0.8)
- ... also relies on some internal frameworks we need to open source
Open Source?

- Yes! ... but not yet
- Relies on unreleased version of Cassandra
- ... but the counters patch is committed in trunk (0.8)
- ... also relies on some internal frameworks we need to open source
- It will happen
Open Source?

- Yes! ... but not yet
- Relies on unreleased version of Cassandra
- ... but the counters patch is committed in trunk (0.8)
- ... also relies on some internal frameworks we need to open source
- It will happen
- See http://github.com/twitter for proof of how much Twitter ❤ open source
Team

- John Corwin (@johnxorz)
- Adam Samet (@damnitsamet)
- Johan Oskarsson (@skr)
- Kelvin Kakugawa (@kelvin)
- Chris Goffinet (@lenn0x)
- Steve Jiang (@sjiang)
- Kevin Weil (@kevinweil)
If You Only Remember One Slide...

- **Rainbird** is a distributed, high-volume counting service built on top of Cassandra.
- Write 100,000s events per second, query it with hierarchy and multiple time granularities, returns results in <100 ms.
- Used by Twitter for multiple products internally, including our Promoted Products, operational monitoring and Tweet Button.
- Will be open sourced so the community can use and improve it!
Questions?

Follow me: @kevinweil