Scuba: Diving into data at Facebook

Lior Abraham, John Allen, Oleksandr Barykin, Vinayak Borkar, Bhuwan Chopra, Cipriam Gerea, Daniel Merl, Josh Metzler, David Reiss, Subbu Subramanium, Janet L. Wiener, Okay Zed

Presented by-
Robina Bhatia

OUTLINE

- Introduction
- Scuba Use cases
- Scuba Overview
- Experimental Evaluation
- Conclusion
- Discussion

INTRODUCTION

- Performance monitoring is a serious issue at Facebook.
- Scuba is a data management system Facebook uses for most of the real-time analysis.
- Scuba is fast, scalable, distributed, in-memory database.
- Since Scuba is memory bound, it expires the data at the same rate as it ingests it.
- Scuba provides an SQL Query interface and a GUI that produces time series graphs, pie charts etc.

INTRODUCTION

- In order to constrain the amount of data, Scuba allows rows to specify an optional sample_rate.
- Many teams at Facebook use Scuba:
  - Mobile development teams
  - Ads
  - Site Reliability
  - Bug Report

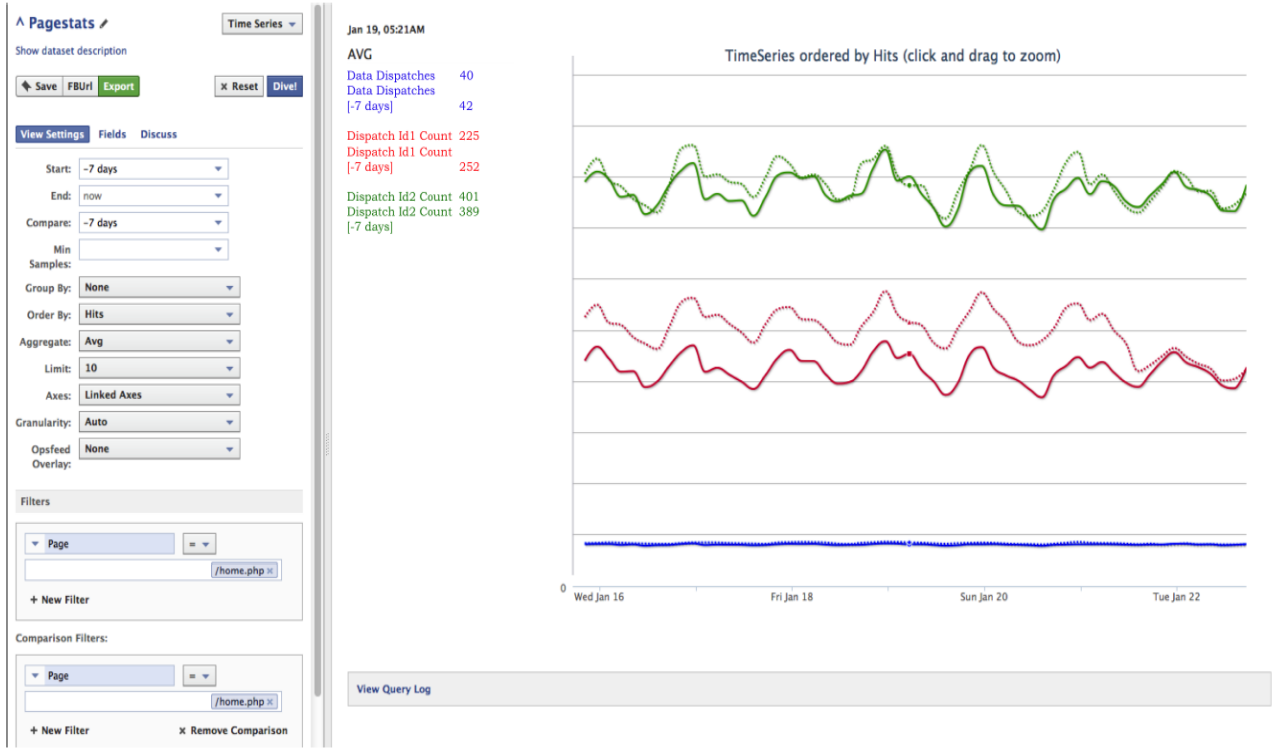

Figure 1: Scuba's web user interface. The query shown on the left side generates a time series graph with a week over week comparison of three columns related to Facebook page dispatches. The dotted lines represent the same days one week earlier. It is very easy to see daily and weekly cyclical behavior with these graphs.

Figure from L. Abraham et al, *Scuba: Diving into Data at Facebook*, Proceedings of the VLDB Endowment, Vol 6, No. 11, August 2th-30th, 2013. Riva Del Garda, Trento, Italy.

SCUBA USE CASES

- Performance Monitoring
- Trend Analysis
- Pattern Matching

SCUBA USE CASES: *Performance Monitoring*

- The original and most common use of Scuba.
- Scuba dashboard displays:
  - Graphs
  - Number of cache requests
  - Network throughputs
- Performance bugs can often be spotted within minutes to hours of their introduction

SCUBA USE CASES: *Trend Analysis*

- User can look for trends in the data content
- User extracts sets of words from user posts and looks for spikes in word frequencies over time and across many dimensions.
- User can write new custom queries to try out new ideas for trend analysis.

SCUBA USE CASES: *Pattern Mining*

- It is used by a product specialist to analyze how different Facebook users respond to the changes in the website or mobile applications.
- Roll up queries run in milliseconds as compared to minutes they take in Hive.

SCUBA OVERVIEW

- Data model
- Data layout
- Data ingestion, distribution and lifetime
- Query model
- Query Execution

SCUBA OVERVIEW: *Data model*

- Scuba provides a standard table model to the users.
- Each table has rows contain the following types of data:
  - Integers
  - Strings
  - Sets of Strings
  - Vector of Strings

SCUBA OVERVIEW: *Data Layout*

Data Type	Compression type	Representation in row
Integer	Variable length	1-8 bytes
String	Dictionary	Index, uses number of bits for max dictionary index
String (alternate)	Uncompressed	4 bytes length + actual string
Sets of String	Dictionary	Fibonacci encoding of deltas between sorted indexes
Vectors of String	Dictionary	2 bytes count + 1 byte index size + each index

Figure 3: Data types and compression methods in Scuba.

Figure from L. Abraham et al, *Scuba: Diving into Data at Facebook*, Proceedings of the VLDB Endowment, Vol 6, No. 11, August 2th-30th, 2013. Riva Del Garda, Trento, Italy.

SCUBA OVERVIEW: *Data Layout*

- The data model of Scuba differs from Relational model in two ways:
  - There is no create table statement.
  - The columns within the table's rows may be sparsely populated; it is common for there to be two or three different row schemas within a table or for a column to change its type over time.

SCUBA OVERVIEW: *Data ingestion and distribution*

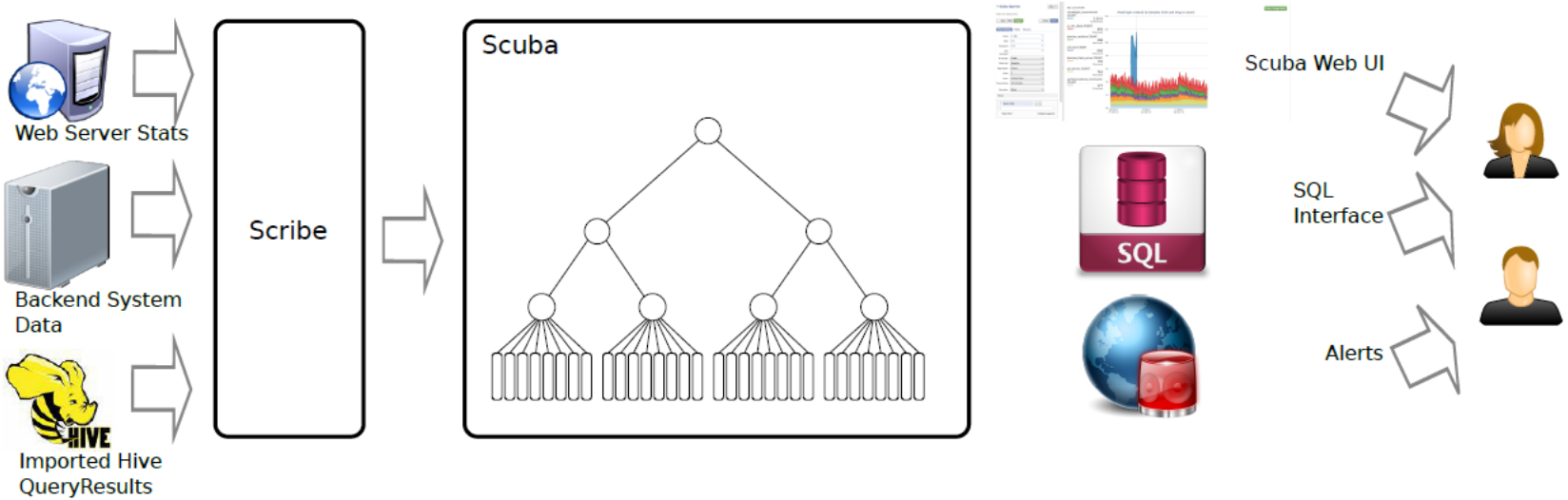

Figure 2: Scuba system architecture: from data ingestion on the left to user queries on the right.

Figure from L. Abraham et al, *Scuba: Diving into Data at Facebook*, Proceedings of the VLDB Endowment, Vol 6, No. 11, August 2th-30th, 2013. Riva Del Garda, Trento, Italy.

SCUBA OVERVIEW: *Data lifetime*

- Old data is deleted at the same rate at which new data is received.
- Data can be pruned for one of the two reasons:
  - Age
  - Space

SCUBA OVERVIEW: *Query Model*

- Scuba provides three query interfaces:
  - Scuba Web UI
  - SQL Interface
  - The Thrift-based API
- Scuba queries have the expressive power of the following SQL query

```
SELECT column, column, ...,
  aggregate(column), aggregate(column), ...
FROM table
WHERE time >= min-timestamp
  AND time <= max-timestamp
 [AND condition ...]
GROUP BY column, column, ...
ORDER BY aggregate(column)
LIMIT number
```

SCUBA OVERVIEW: *Query model*

- Any comparison to a string may include a regular expression.

- Joins are not supported in Scuba.

SCUBA OVERVIEW: *Query Execution*

The following interact with each other via Thrift:

- The Scuba Client
- The Root Aggregator
- The Intermediate Aggregator
- The Leaf Aggregator

SCUBA OVERVIEW: *Query Execution*

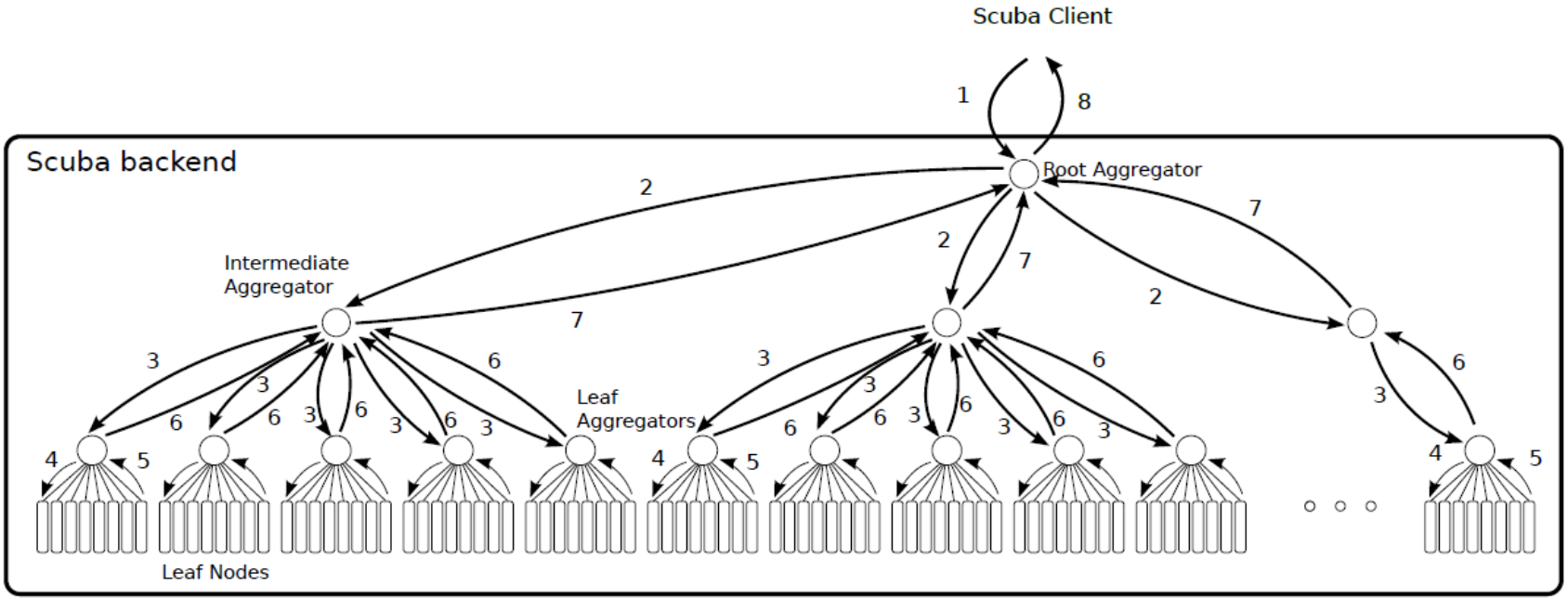

Figure 7: Step-wise breakdown of executing a Scuba query.

Figure from L. Abraham et al, *Scuba: Diving into Data at Facebook*, Proceedings of the VLDB Endowment, Vol 6, No. 11, August 2th-30th, 2013. Riva Del Garda, Trento, Italy.

EXPERIMENTAL EVALUATION

- For the experimental evaluation, Scuba's speed up and scale up was measured on a test cluster of 160 machines.
- Two queries were run:

```
SELECT count(*), SUM(column1) as sum1,
   SUM(column2) as sum2
FROM mytable
WHERE time >= now()-3*3600
```

```
SELECT count(*), sum(column1) as sum1,
   service,
   (time - now())/60*60 + now() as minute,
FROM mytable
WHERE time >= now()-3*3600
   and time <= now()
GROUP BY service, minute
ORDER BY sum1 DESC
LIMIT 1800
```

EXPERIMENTAL EVALUATION: *Single Client Experiments*

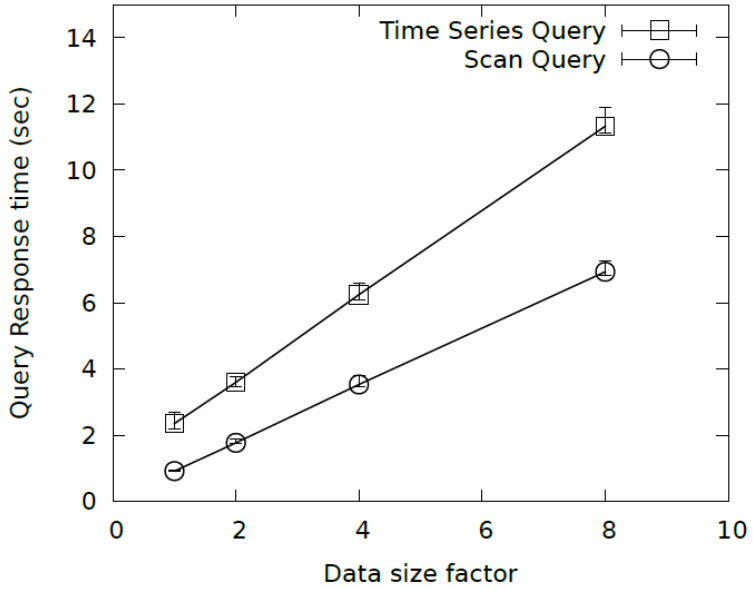

Figure 10: Measuring speed up as the amount of data increases.

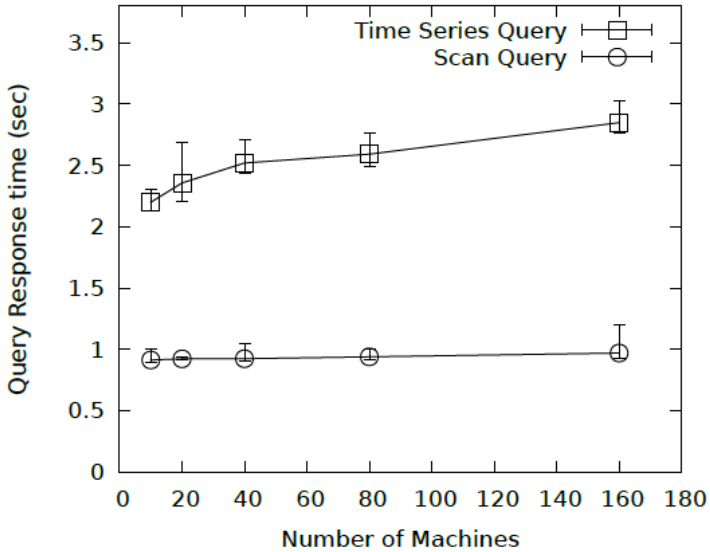

Figure 11: Measuring scale up as the number of machines increases.

Figure from L. Abraham et al, *Scuba: Diving into Data at Facebook*, Proceedings of the VLDB Endowment, Vol 6, No. 11, August 2th-30th, 2013. Riva Del Garda, Trento, Italy.

EXPERIMENTAL EVALUATION: *Multi client experiments*

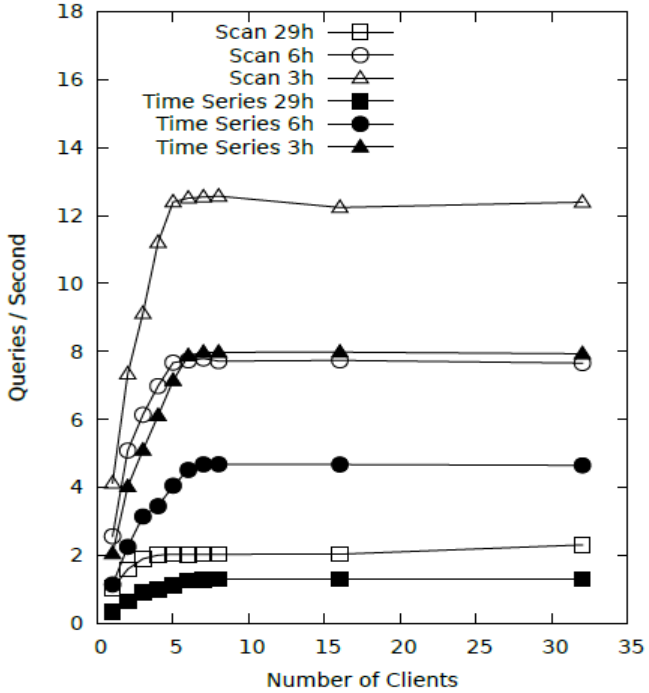

Figure 12: Measuring throughput as the number of clients increases.

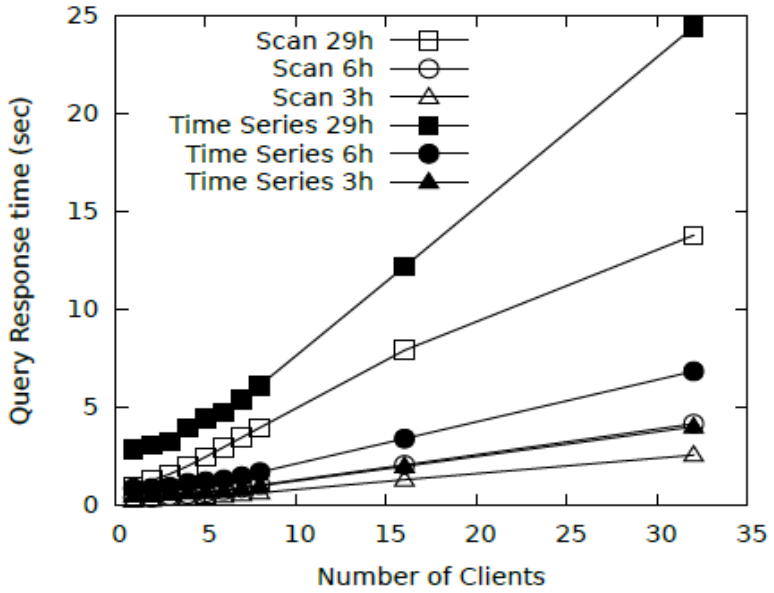

Figure 13: Measuring latency as the number of clients increases.

Figure from L. Abraham et al, *Scuba: Diving into Data at Facebook*, Proceedings of the VLDB Endowment, Vol 6, No. 11, August 2th-30th, 2013. Riva Del Garda, Trento, Italy.

CONCLUSION

- Scuba prunes data as fast as it ingests data.
- Scuba expects that tables will contain sampled data.
- No schema declaration is required.
- A table can contain rows with different schemas
- Scuba has about a dozen different ways to visualize data.
- Queries are run with best effort availability.
- Scuba is not intended to be a complete SQL database.

DISCUSSION

- Joins are not supported in Scuba.
- Why no support for float data types?
- How important is accuracy for Stream processing systems? (response time vs accuracy)