

Apache Drill

Interactive Analysis of Large-Scale Datasets

Tomer Shiran

Latency Matters

- Ad-hoc analysis with interactive tools
- Real-time dashboards
- Event/trend detection
 - Network intrusions
 - Fraud
 - Failures



Big Data Processing

	Batch processing	Interactive analysis	Stream processing
Query runtime	Minutes to hours	Milliseconds to minutes	Never-ending
Data volume	TBs to PBs	GBs to PBs	Continuous stream
Programming model	MapReduce	Queries	DAG
Users	Developers	Analysts and developers	Developers
Google project	MapReduce	Dremel	
Open source project	Hadoop MapReduce		Storm and S4

Introducing Apache Drill...



GOOGLE DREMEL

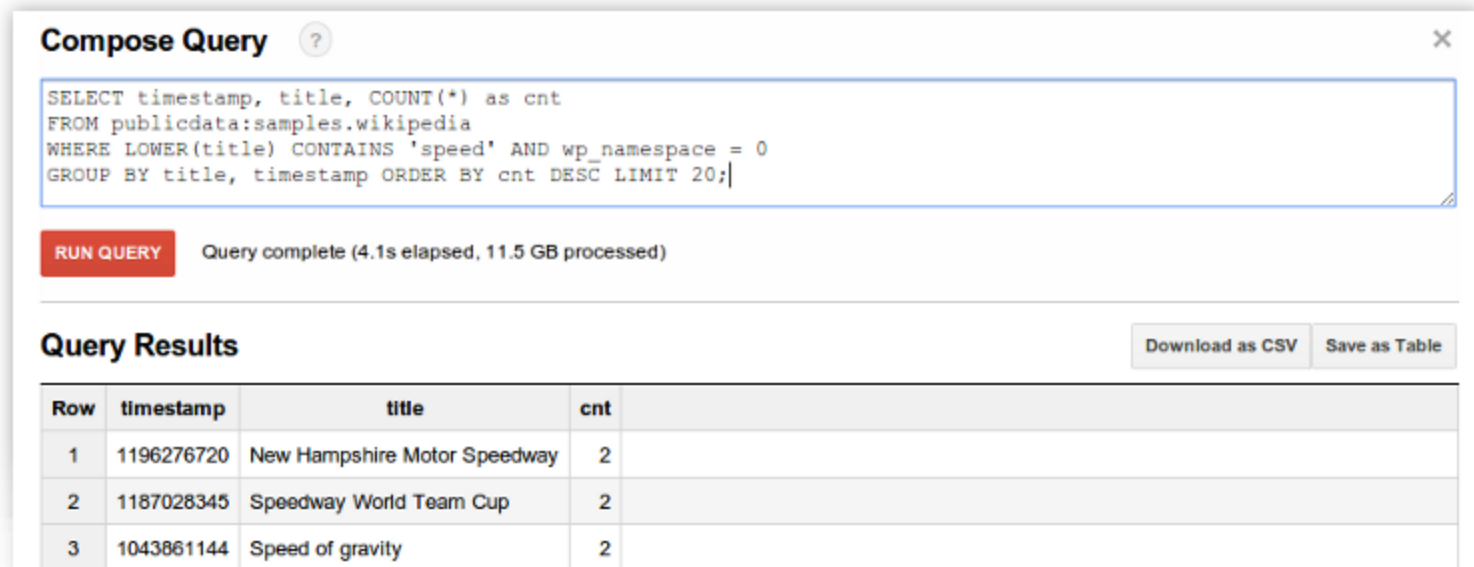
Google Dremel



- Interactive analysis of large-scale datasets
 - Trillion records at interactive speeds
 - Complementary to MapReduce
 - Used by thousands of Google employees
 - Paper published at VLDB 2010
 - Authors: Sergey Melnik, Andrey Gubarev, Jing Jing Long, Geoffrey Romer, Shiva Shivakumar, Matt Tolton, Theo Vassilakis
- Model
 - Nested data model with schema
 - Most data at Google is stored/transferred in Protocol Buffers
 - Normalization (to relational) is prohibitive
 - SQL-like query language with nested data support
- Implementation
 - Column-based storage and processing
 - In-situ data access (GFS and Bigtable)
 - Tree architecture as in Web search (and databases)

Google BigQuery

- Hosted Dremel (Dremel as a Service)
- CLI (bq) and Web UI
- Import data from Google Cloud Storage or local files
 - Files must be in CSV format
 - Nested data not supported [yet] except built-in datasets
 - Schema definition required



The screenshot shows the 'Compose Query' interface in Google BigQuery. At the top, there is a text input field containing a SQL query: `SELECT timestamp, title, COUNT(*) as cnt FROM publicdata:samples.wikipedia WHERE LOWER(title) CONTAINS 'speed' AND wp_namespace = 0 GROUP BY title, timestamp ORDER BY cnt DESC LIMIT 20;`. Below the query field is a red 'RUN QUERY' button and a status message: 'Query complete (4.1s elapsed, 11.5 GB processed)'. Underneath, the 'Query Results' section is visible, featuring a table with three columns: 'Row', 'timestamp', 'title', and 'cnt'. The table contains three rows of data. To the right of the table, there are two buttons: 'Download as CSV' and 'Save as Table'.

```
SELECT timestamp, title, COUNT(*) as cnt
FROM publicdata:samples.wikipedia
WHERE LOWER(title) CONTAINS 'speed' AND wp_namespace = 0
GROUP BY title, timestamp ORDER BY cnt DESC LIMIT 20;
```

Query Results Download as CSV Save as Table

Row	timestamp	title	cnt
1	1196276720	New Hampshire Motor Speedway	2
2	1187028345	Speedway World Team Cup	2
3	1043861144	Speed of gravity	2

APACHE DRILL

Nested Data Model

- The data model in Dremel is Protocol Buffers
 - Nested
 - Schema
- Apache Drill is designed to support multiple data models
 - Schema: Apache Avro, Protocol Buffers, ...
 - Schema-less: JSON, BSON, ...
- Flat records are supported as a special case of nested data
 - CSV, TSV, ...

Avro IDL

```
enum Gender {
  MALE, FEMALE
}

record User {
  string name;
  Gender gender;
  long followers;
}
```

JSON

```
{
  "name": "Tomer",
  "gender": "Male",
  "followers": 100
}
{
  "name": "Maya",
  "gender": "Female",
  "followers": 200,
  "zip": "94305"
}
```


Nested Query Languages

- DrQL
 - SQL-like query language for nested data
 - Compatible with Google BigQuery/Dremel
 - BigQuery applications should work with Drill
 - Designed to support efficient column-based processing
 - No record assembly during query processing
- Mongo Query Language
 - `{ $query: {x: 3, y: "abc"}, $orderby: {x: 1}}`
- Other languages/programming models can plug in

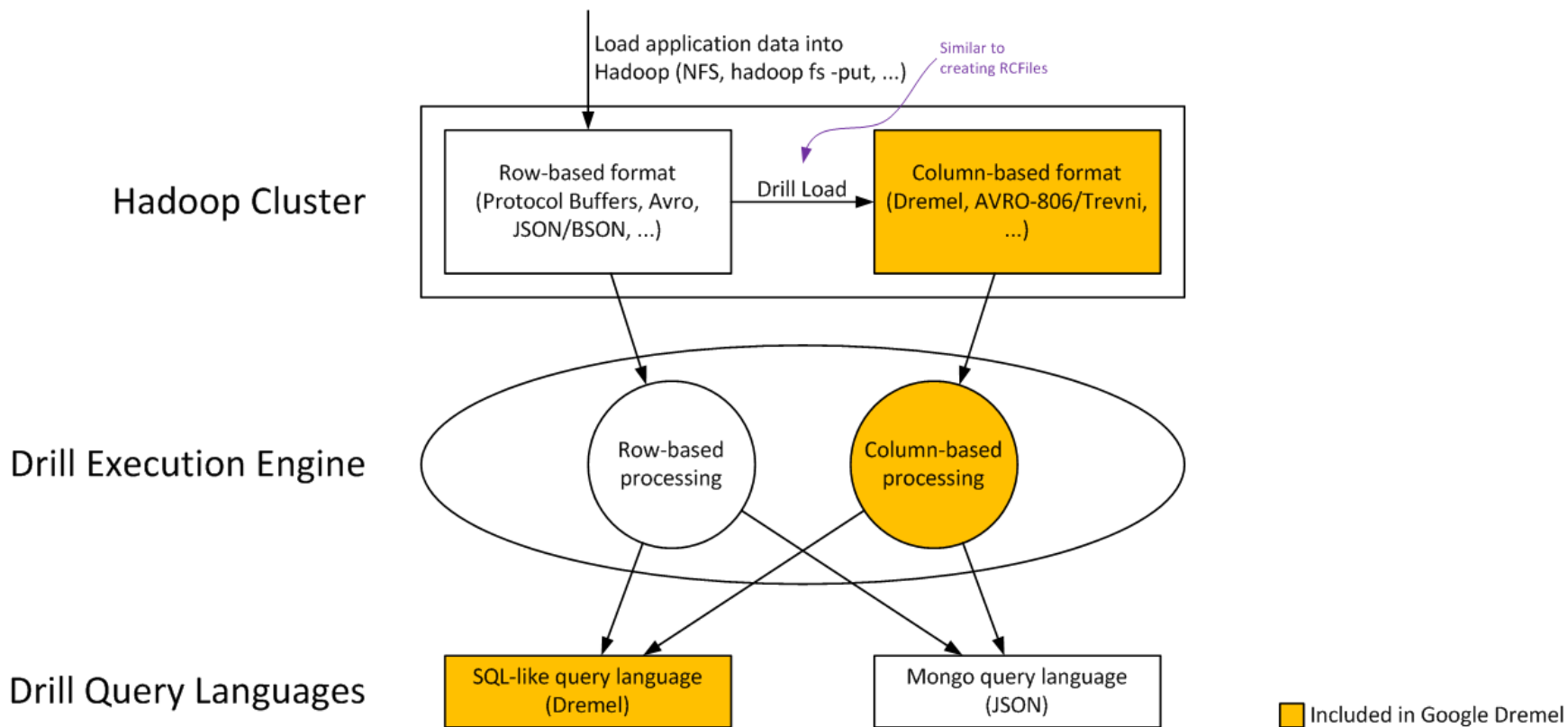
DrQL Example

```
DocId: 10
Links
  Forward: 20
  Forward: 40
  Forward: 60
Name
  Language
    Code: 'en-us'
    Country: 'us'
  Language
    Code: 'en'
  Url: 'http://A'
Name
  Url: 'http://B'
Name
  Language
    Code: 'en-gb'
    Country: 'gb'
```

```
SELECT DocId AS Id,
       COUNT(Name.Language.Code) WITHIN Name AS Cnt,
       Name.Url + ',' + Name.Language.Code AS Str
FROM t
WHERE REGEXP(Name.Url, '^http') AND DocId < 20;
```

```
Id: 10
Name
  Cnt: 2
  Language
  Str:
'http://A,en-us'
  Str:
'http://A,en'
Name
  Cnt: 0
```

Data Flow



Architecture

- Nested query languages
 - Pluggable model
 - DrQL
 - Mongo Query Language
- Distributed execution engine
 - Extensible model (eg, Dryad)
 - Low-latency
 - Fault tolerant
 - Column-based and row-based processing
- Nested data formats
 - Pluggable model
 - Column-based (Dremel, AVRO-806/Trevni, RCFile) and row-based (Protocol Buffers, Avro, JSON, BSON, CSV)
 - Schema (Protocol Buffers/Dremel, Avro/AVRO-806/Trevni, CSV) and schema-less (JSON, BSON)
- Scalable data sources
 - Pluggable model
 - Hadoop
 - NoSQL



Design Principles

Flexible

- Pluggable query languages
- Extensible execution engine
- Pluggable data formats
 - Column-based and row-based
 - Schema and schema-less
- Pluggable data sources

Easy

- Unzip and run
- Zero configuration
- Reverse DNS not needed
- IP addresses can change
- Clear and concise log messages

Dependable

- No SPOF
- Instant recovery from crashes

Fast

- C/C++ core with Java support
- Min latency and max throughput (limited only by hardware)
- Full column-based data support including operators

Hadoop Integration

- Hadoop data sources
 - Hadoop FileSystem API (HDFS/MapR-FS)
 - HBase
- Hadoop data formats
 - Apache Avro
 - RCFile
- MapReduce-based tools to create column-based formats
- Hive-based query language and optimizer
- Table registry in Hcatalog
- Run long-running services in YARN

