If you have your own Columnar format, stop now and use Parquet

😛

Julien Le Dem
Principal Architect, Dremio
VP Apache Parquet

About Dremio

- Enabling self-service data discovery, exploration and analysis on modern data
- Founded in June 2015
- Building on open source technologies including Drill, Parquet, Spark
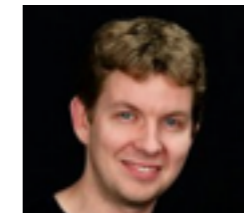
Top Silicon Valley VCs

LIGHTSPEED VENTURE PARTNERS

Redpoint

Tomer Shiran
Founder & CEO

- Apache Drill Founder
- MapR (VP Product); Microsoft; IBM Research
- Carnegie Mellon, Technion

Jacques Nadeau
Founder & CTO

- Apache Drill PMC Chair
- Recognized SQL & NoSQL expert
- Quigo (AOL); Offermatica (ADBE); aQuantive (MSFT)

Julien Le Dem
Architect

- Apache Parquet Founder
- Apache Pig PMC Member
- Twitter (Lead, Analytics Data Pipeline); Yahoo! (Architect)

dremio

Background of Parquet

- **Twitter's data**
  - Lots of data: Instrumentation, User graph, Derived data, ...
  - Complex: deeply nested structures

- **Analytics infrastructure**:
  - Several 1000s nodes Hadoop clusters
  - Log collection to HDFS in Thrift

- **Parquet**
  - Columnar: space and query efficient
  - Inspired from the Google Dremel Paper
 - supports complex data
  - interoperable

Caillebotte: The Parquet Planers

Parquet timeline

- **Fall 2012:** Twitter & Cloudera's Impala team merge efforts to develop columnar formats.

- **March 2013:** OSS announcement; Criteo signs on for Hive integration.

- **July 2013:** 1.0 release. 18 contributors from more than 5 organizations.

- **August 2013:** Drill chose Parquet as its primary storage format.

- **May 2014:** Apache Incubator. 40+ contributors, 18 with 1000+ LOC. 26 incremental releases.

- **Apr 2015:** Parquet graduates from the Apache Incubator.

Contributors
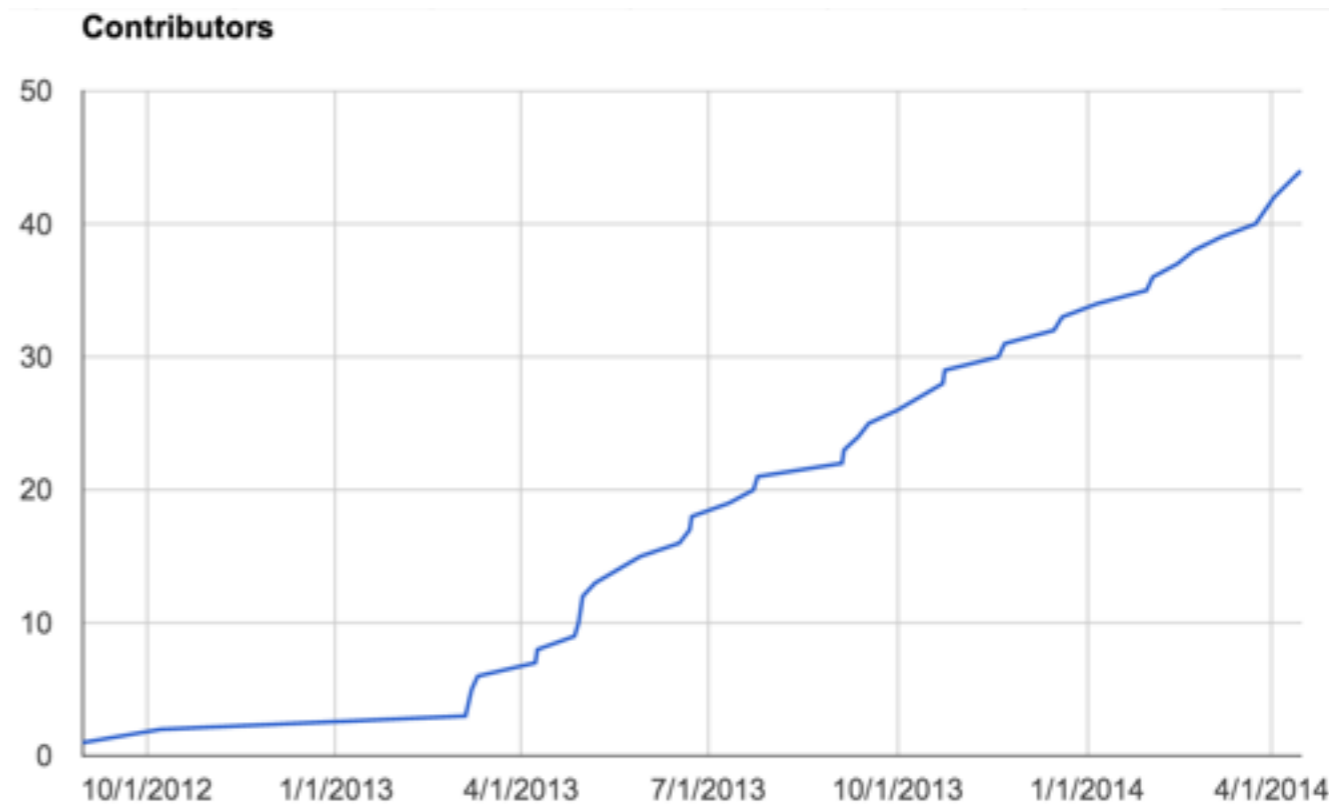
dremio

What does Parquet do?

Interoperability

Space efficiency

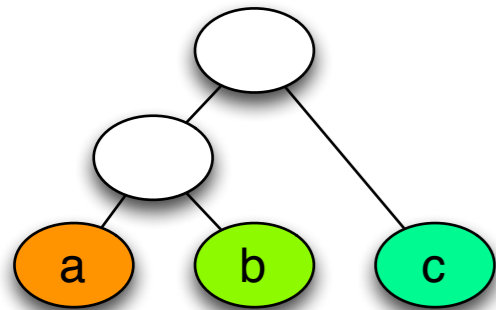**Query efficiency**

dremio

Columnar storage

Nested schema

Logical table representation

a	b	c
a1	b1	c1
a2	b2	c2
a3	b3	c3
a4	b4	c4
a5	b5	c5

On Disk:

Row layout

a1	b1	c1	a2	b2	c2	a3	b3	c3	a4	b4	c4	a5	b5	c5

Column layout

a1	a2	a3	a4	a5	b1	b2	b3	b4	b5	c1	c2	c3	c4	c5

encoded chunk	encoded chunk	encoded chunk
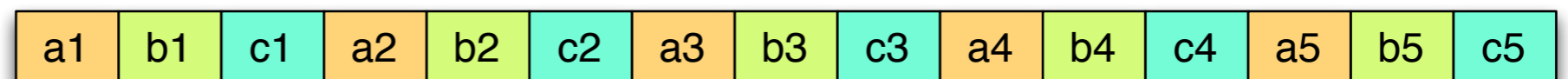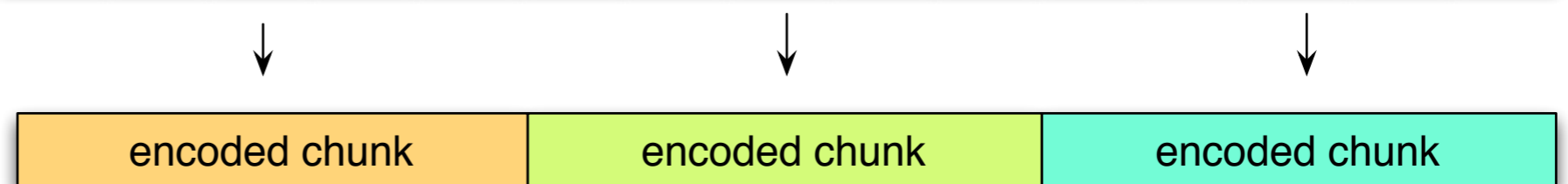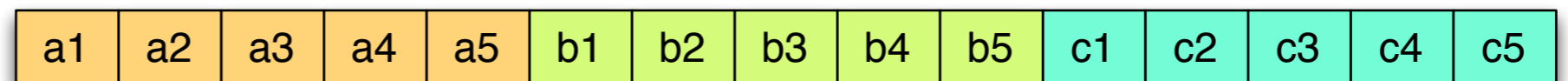
Encodings: Dictionary, RLE, Delta, Prefix

dremio

Nested representation

Borrowed from the Google Dremel paper

Schema:

Columns:
docid
links.backward
links.forward
name.language.code
name.language.country
name.url

https://blog.twitter.com/2013/dremel-made-simple-with-parquet
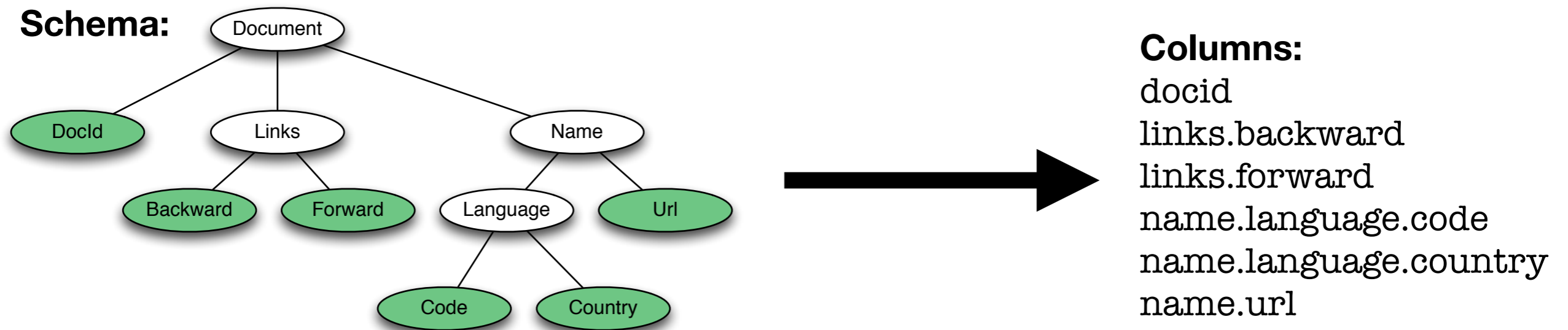
Statistics

Vertical partitioning
(projection push down)

+

Horizontal partitioning
(predicate push down)

=

Read only the data
you need!

a	b	c
a1	b1	c1
a2	b2	c2
a3	b3	c3
a4	b4	c4
a5	b5	c5

+

a	b	c
a1	b1	c1
a2	b2	c2
a3	b3	c3
a4	b4	c4
a5	b5	c5

=

a	b	c
a1	b1	c1
a2	b2	c2
a3	b3	c3
a4	b4	c4
a5	b5	c5

dremio

Interoperability
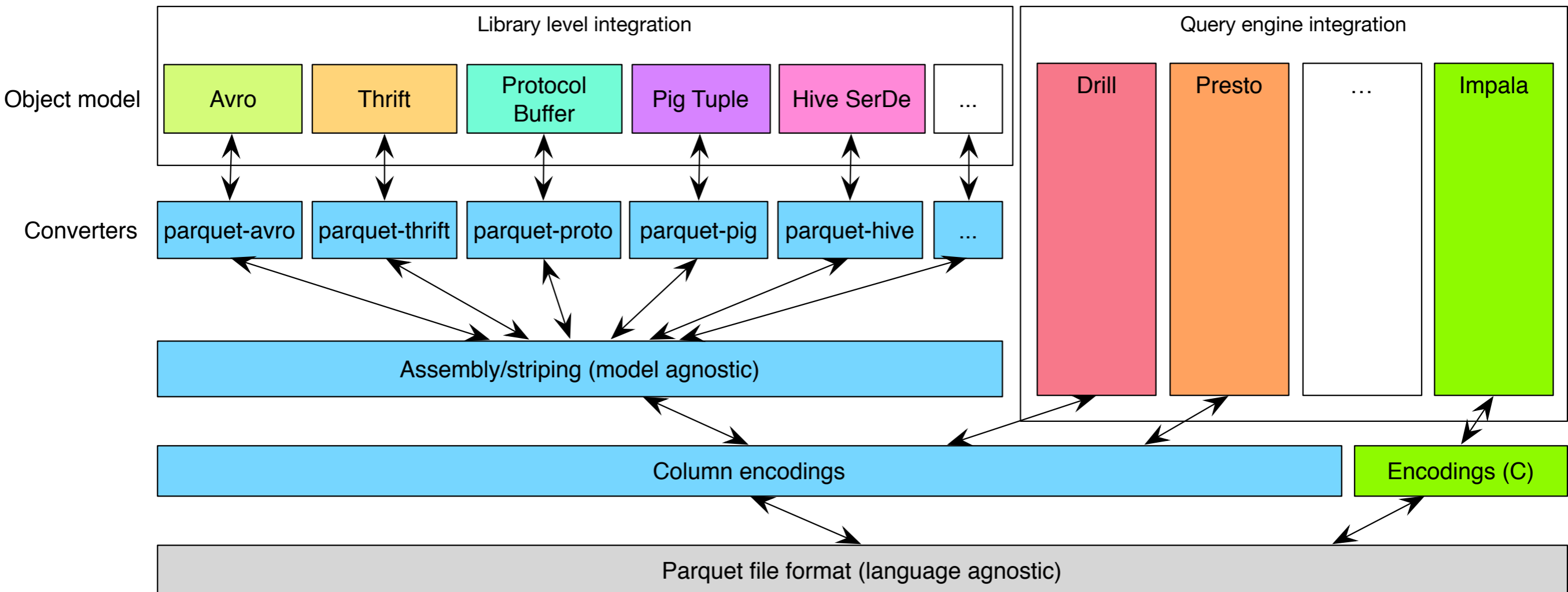
Query engines, frameworks and libraries integrated with Parquet (non exhaustive)

Query engines:
Hive, Impala, HAWQ,
IBM Big SQL, Drill, Tajo,
Pig, Presto, SparkSQL

Frameworks:
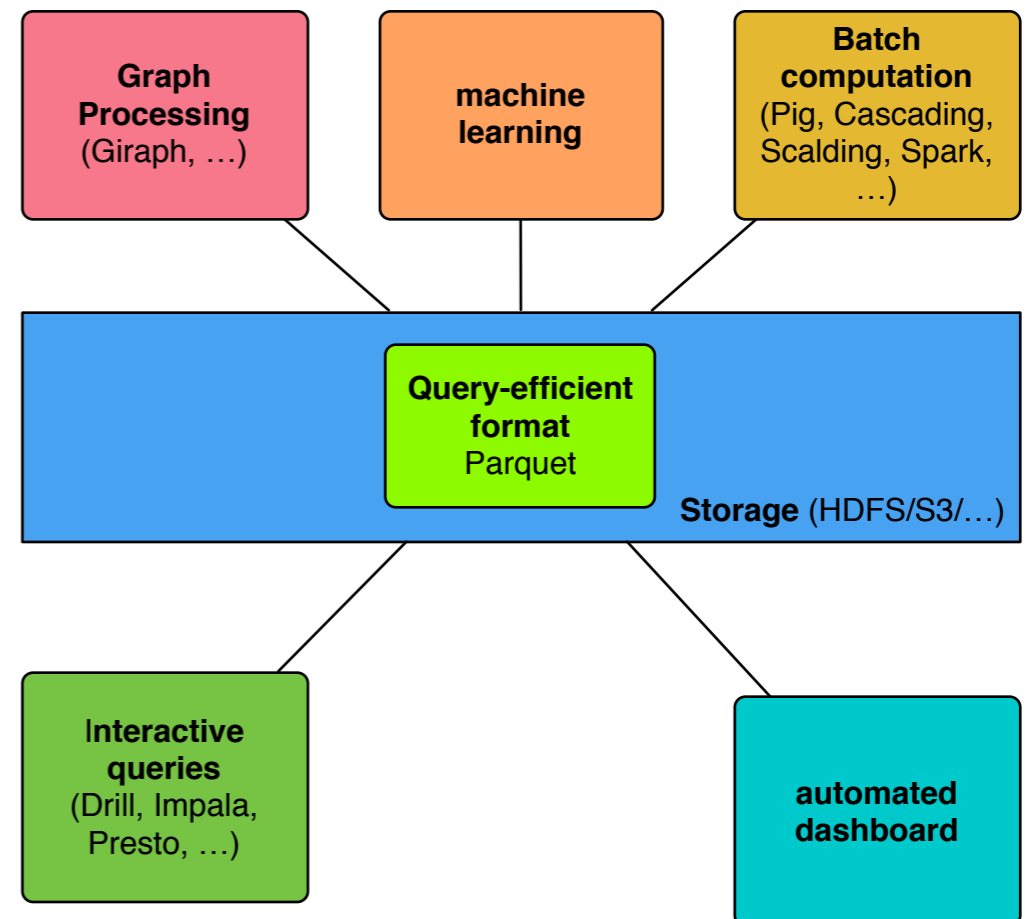Spark, MapReduce, Cascading,
Crunch, Scalding, Kite

Data Models:
Avro, Thrift, ProtocolBuffers,
POJOs

Loose coupling

- Users don't want to load their data into every tool.

- Many tools are available and show up every day.

- The cost of trying a new tool should be minimal

Get involved

Twitter:
- @ApacheParquet

Mailing list:
- dev@parquet.apache.org

Github repo:
- https://github.com/apache/parquet-mr

Parquet sync ups:
- Regular meetings on google hangout

dremio