
WHERE WILL YOUR DATA TAKE YOU?
.conf2012
splunk>
Copyright © 2012 Splunk Inc.

How to Integrate Splunk with any Data Solution

Julian Hyde (Optiq) @julianhyde

http://github.com/julianhyde/optiq
http://github.com/julianhyde/optiq-splunk

Splunk Worldwide Users Conference 2012

Why are we here?

I'm going to explain how to use Splunk to access all of the data in your enterprise.

And also to let people in your enterprise use data in Splunk.

This isn't easy. We'll be showing some raw technology – the new Optiq project and its Splunk adapter.

But it's open source, so you can all get your hands on it. :)

About me

- Database hacker
- Open source hacker
- Author of Mondrian (Pentaho Analysis)
- Startup fiend

http://www.flickr.com/photos/torkildr/3462606643


http://www.flickr.com/photos/sylvar/31436961/

“Big Data”

Right data, right time

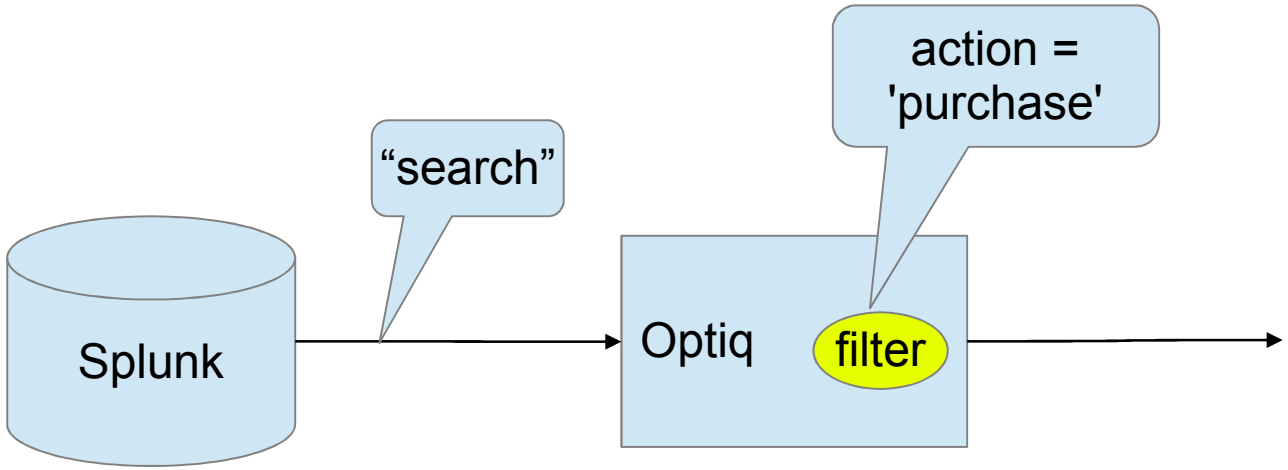
Diverse data sources / Performance / Suitable format

Example

Accessing Splunk data via SQL

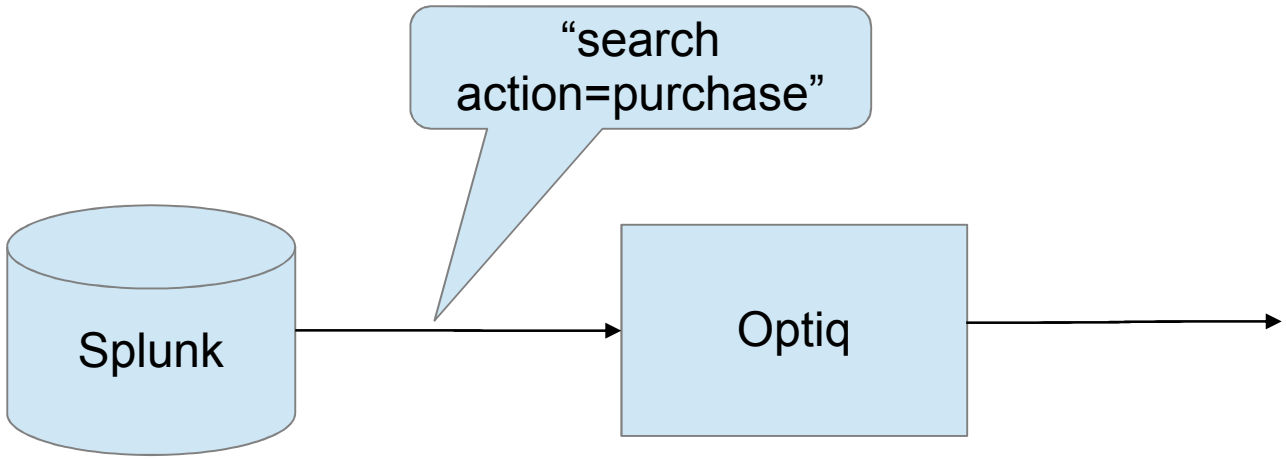
Sqlline (a standard JDBC client)

How do it (wrong)




```
SELECT “source”, “product_id”
FROM “splunk”.”splunk”
WHERE “action” = 'purchase'
```

How do it (right)




```
SELECT “source”, “product_id”
FROM “splunk”.”splunk”
WHERE “action” = 'purchase'
```

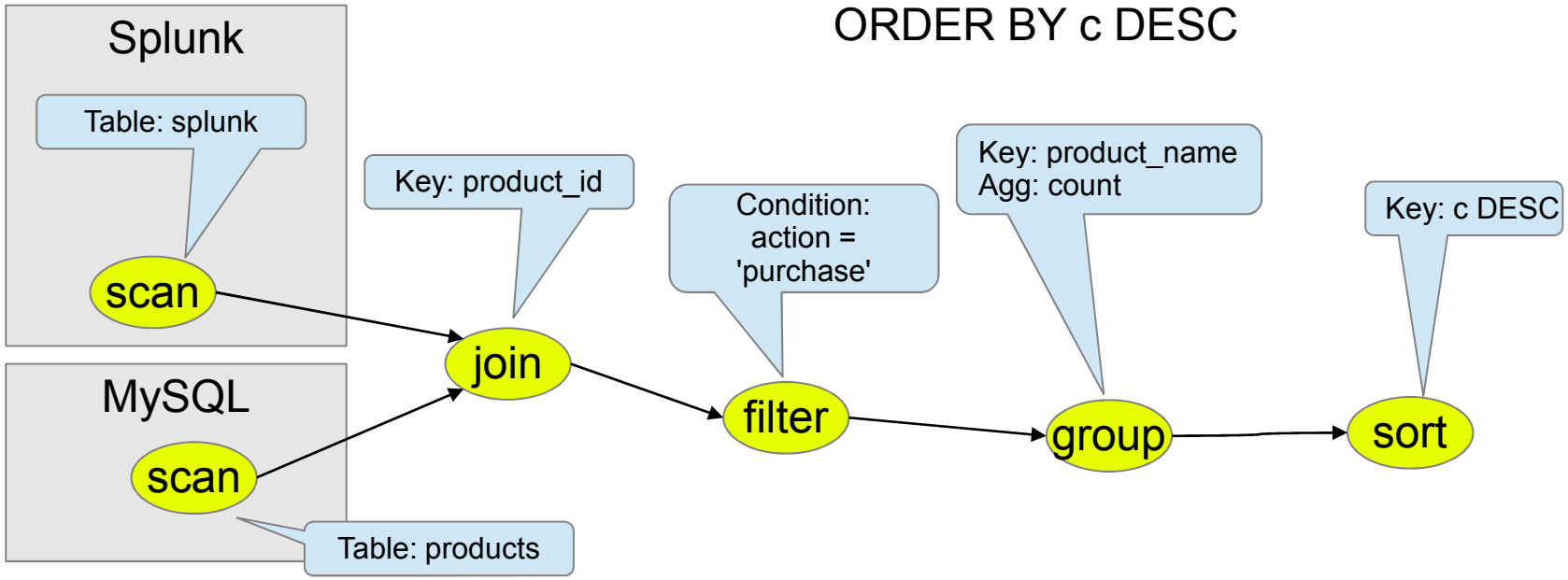
Example #2

Combining data from 2 sources (Splunk & MySQL)

Also possible: 3 or more sources; 3-way joins; unions

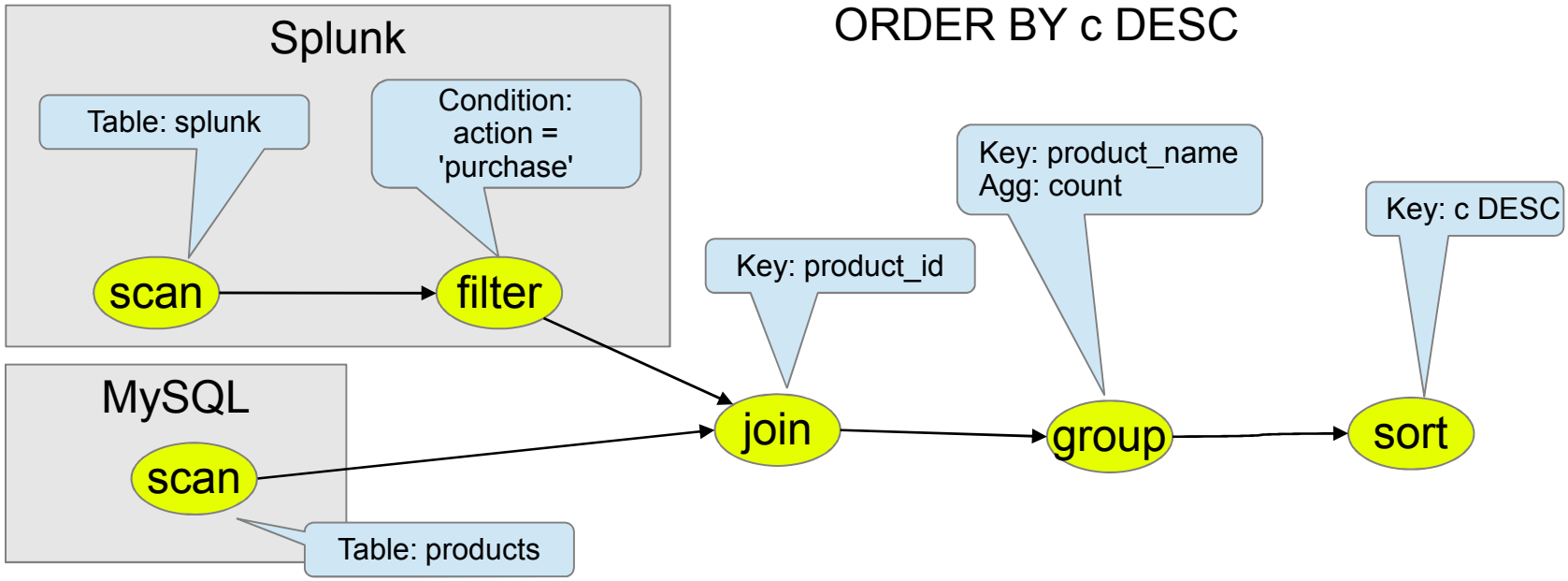
Expression tree

```
SELECT p.“product_name”, COUNT(*) AS c
FROM “splunk”.”splunk” AS s
    JOIN “mysql”.”products” AS p
    ON s.”product_id” = p.”product_id”
WHERE s.“action” = 'purchase'
GROUP BY p.”product_name”
ORDER BY c DESC
```



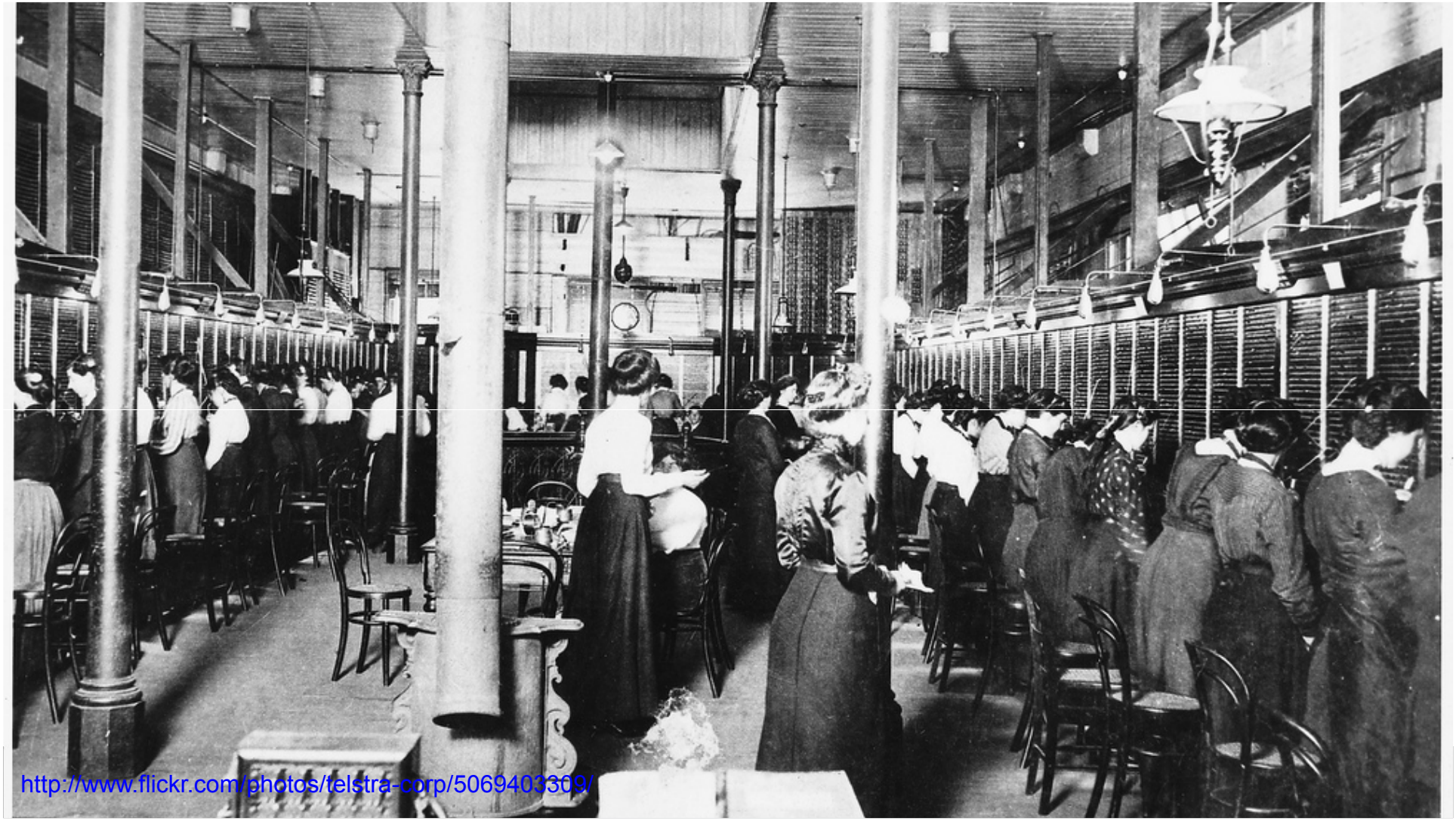
Expression tree (optimized)

```
SELECT p.“product_name”, COUNT(*) AS c
FROM “splunk”.”splunk” AS s
    JOIN “mysql”.”products” AS p
    ON s.”product_id” = p.”product_id”
WHERE s.“action” = 'purchase'
GROUP BY p.”product_name”
ORDER BY c DESC
```

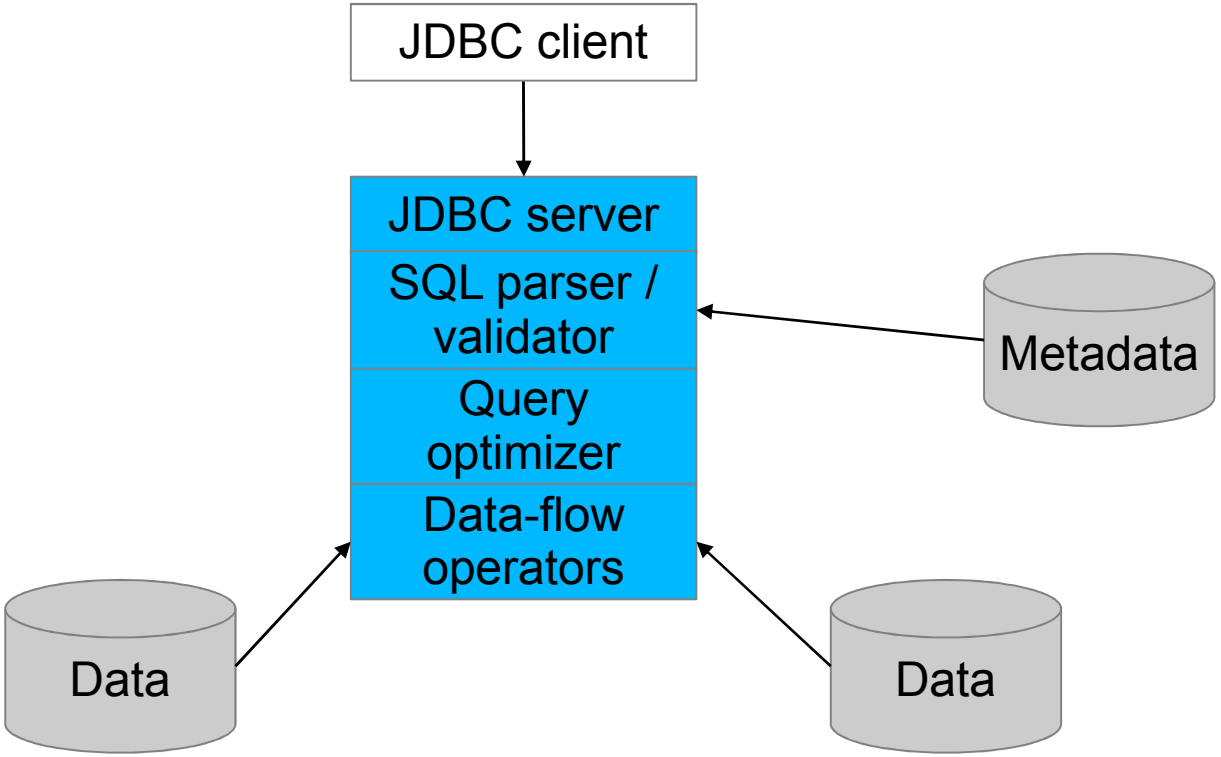


Optiq is not a database.

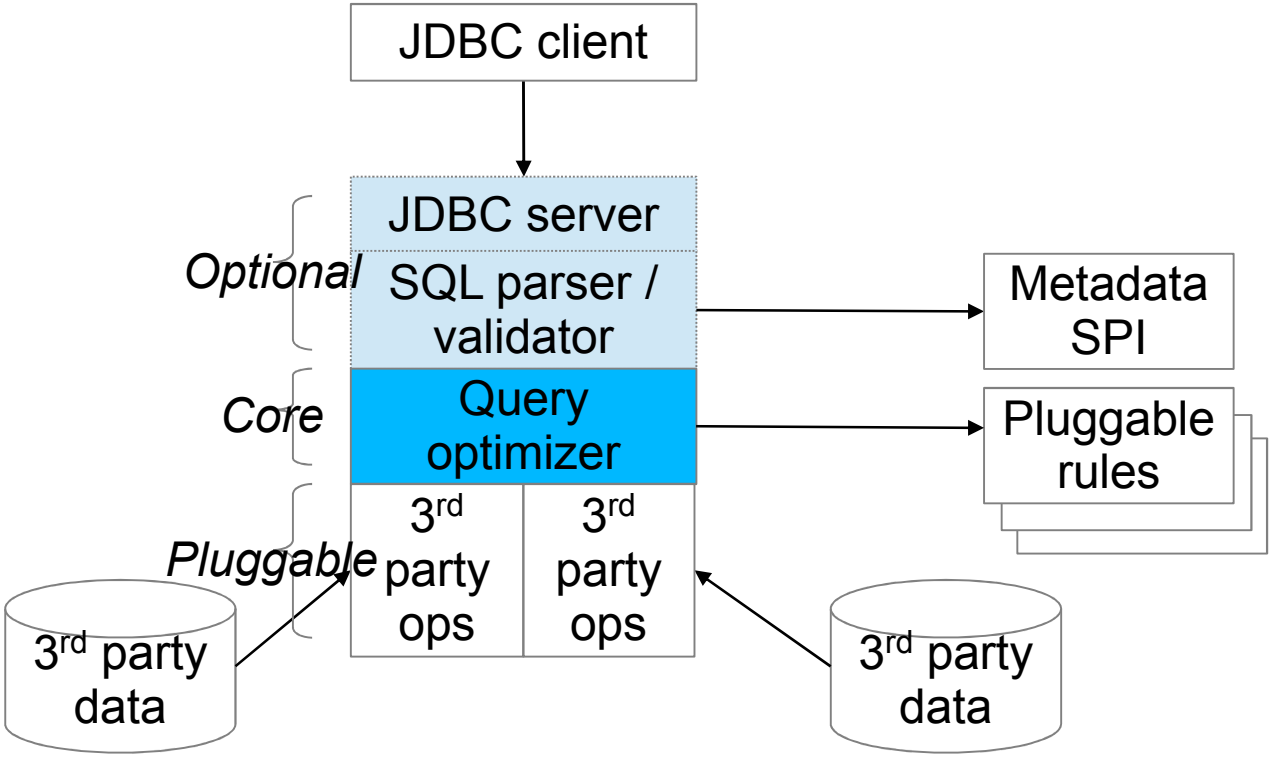
http://www.flickr.com/photos/torkildr/3462606643


http://www.flickr.com/photos/telstra-corp/5069403309/

Conventional database architecture



Optiq architecture



What is Optiq?

A really, really smart JDBC driver

Framework

Potential core of a data management system

Writing an adapter

Driver – if you want a vanity URL like “jdbc:splunk:”

Schema – describes what tables exist (Splunk has just one)

Table – what are the columns, and how to get the data. (Splunk's table has any column you like... just ask for it.)

Operators (optional) – non-relational operations

Rules (optional, but recommended) – improve efficiency by changing the question

Parser (optional) – to query via a language other than SQL

Splunk Adapter

Rules for pushing down filters, projections

The tricky bit: changed the validator to allow tables to have any column

To be written: rules for pushing down aggregations, joins

(What you've seen today is in github.)

Would be really nice if... Splunk pushed down filters, projections, aggregations from its search pipeline to the MySQL connector. (Currently you have to hand-write a SQL statement.)


http://www.flickr.com/photos/walkercarpenter/4697637143/

Optiq roadmap ideas

Mondrian use Optiq to read from data sources such as Splunk

Kettle integration (read/write SQL to ETL)

Adapters: Cascading, MongoDB, Hbase, Apache Drill, …?

Front-ends: linq4j, Scala SLICK, Java8 streams

Contributions

Conclusions

Liberate your data!

Optiq is a framework

Build & share Optiq adapters

Questions?

@julianhyde

http://julianhyde.blogspot.com

http://github.com/julianhyde/optiq

http://github.com/julianhyde/optiq-splunk

Additional material: The following queries were used in the demo

```
select s."source", s."sourcetype"
    from "splunk"."splunk" as s;
```

```
select s."source", s."sourcetype",
    s."action" from
    "splunk"."splunk" as s
where s."action" = 'purchase';
```

```
select s."source", s."sourcetype",
```

```
select * from "mysql"."products";
```

```
select p."product_name",
    s."action"
from "splunk"."splunk" as s
 join "mysql"."products" as p
 on s."product_id" =
    p."product_id";
```