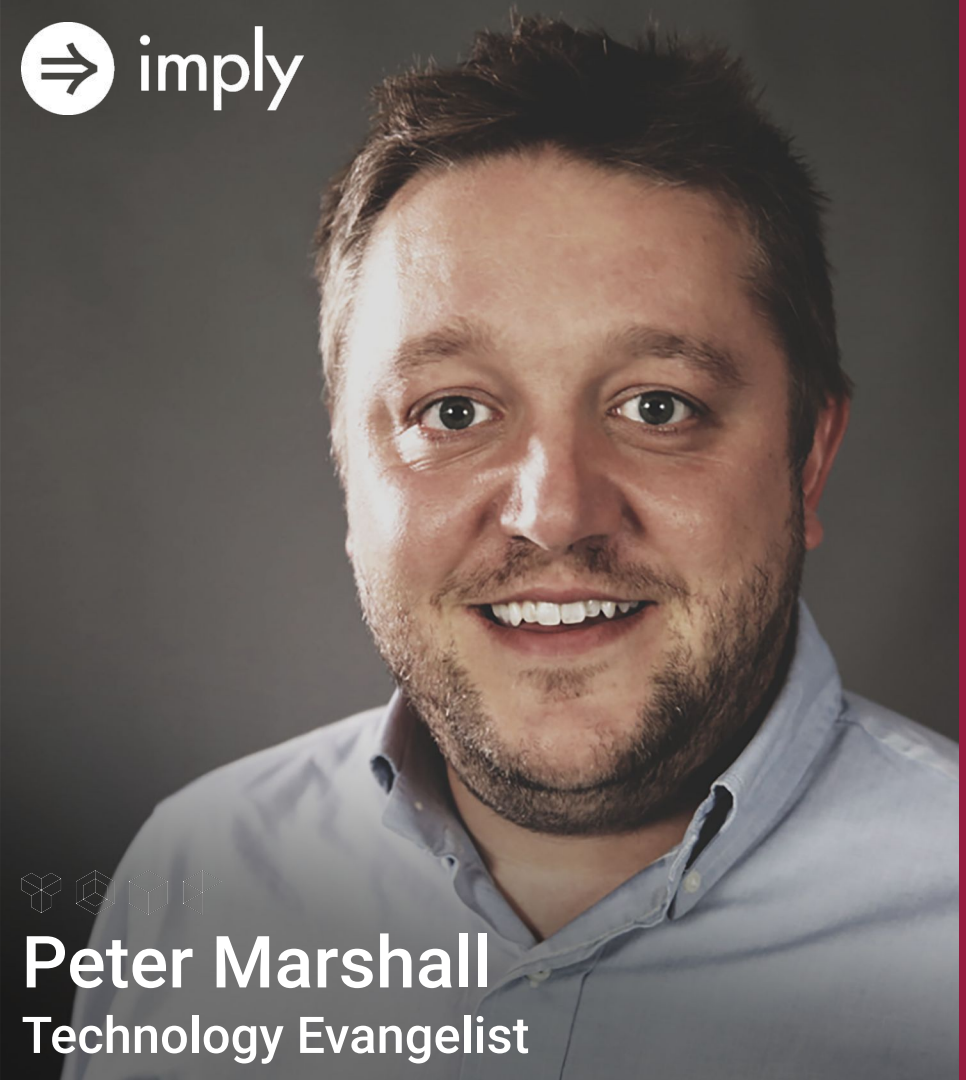
druid®

Under the covers...



Peter Marshall

Technology Evangelist

20 years in Enterprise Architecture

CRM, EDRM, ERP, EIP, Digital Services, Security, BI, RI, and MDM

BA Theology (!) and Computer Studies

TOGAF certified

Book collector & A/V buyer
Prime Timeline = *proper* timeline
#werk

petermarshall.io

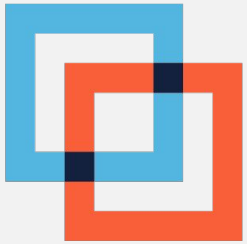
peter.marshall@imply.io

The **pain** that led to Druid being made

What the **components** of a cluster are

How Druid **optimises** data for on-demand statistics

Questions & Answers



METAMARKETS

unscalable	*poor query and ingestion scalability*
disjointed	*event data and batch data separately*
limiting	*expensive, product-constraining precalculation*
slow	*high latencies from production to presentation*

druid®

high performance

real-time

analytics

database

low-latency, distributed query execution and high throughput ingestion

event data (clickstream, network flows, user behavior, programmatic advertising, server metrics, IoT...)

counting, ranking, statistics...

highly-available, time-sharded, partitioned, columnar, indexed, compressed, versioned materialized view

2012 **Druid open sourced**

2013 **Druid 0.5** - Query prioritization, spatial indexing, group limits and ordering, indexing improvements

2014 **PyDruid** promoted

Druid 0.6 - Union, topN, alpha sorting, smarter brokers

2015 **Apache 2.0 License applied**

Druid 0.7.1 - Case sensitivity, pluggable metadata DB, LZ4 compression, improved balance, GROUP BY hashing

Druid 0.7.2

Community Project formed

2016 **Druid 0.8.3** - Broker performance improvements, Azure Blob storage support, TopN improvements, IPv6, more metrics

Druid 0.9.0 - Extensions overhaul, doubleMax / Min aggregators, Avro support, Graphite emitter, regex search, dimension ordering

Druid 0.9.1 - Query-time lookups, native Kafka indexing, statsD emitter, authorization, cost-based distribution, filter optimization

Druid 0.9.2 - New GROUP BY engine, roll-up made optional, Long filtering and encoding, datasketches, ORC support

2017 **Druid 010.0** - Apache Calcite SQL, Kerberos, LIKE, 2GB+ dims, GROUP BY improved, Ambari emitter, more DB support

First Powered By list includes Yahoo!, Walmart, TripleLift, PayPal, Netflix, Hulu, eBay, and AirBnB

Druid 0.10.1

Druid 0.11.0 - Double columns, TLS encryption, auth extensions, Redis cache, GROUP BY improvements, MORE SQL

2018 **Apache Druid® becomes an incubating Apache project**

Druid 0.12.0 - Incremental Kafka hand-off, compaction task, Quantiles sketch, basic auth, query queue improvements, MORE SQL

Druid 0.12.1 - Kerberos and Kafka improvements

Druid 0.12.2 - Ingestion improvements and better query caching

Druid 0.12.3 - Even more ingestion and query improvements

Druid 0.13 - Parallel batch, auto-compaction, SYSTEM, HyperLogLog, NULLs, blooms, new aggregators, OpenTSDB emitter

2019 **Druid 0.14** - New web console, Kinesis indexer, Parquet support, DogStatsD emitter, improved data balance

Druid 0.14.2 - Improved datasketch support

Druid 0.15 - Dataloader UI, moving averages, Moments datasketch, Orc core extensions, MORE SQL

Druid 0.16 - Server / SQL / Data UI, vectorization, minor compaction, GROUP BY arrays, Batch Shuffle, Docker image

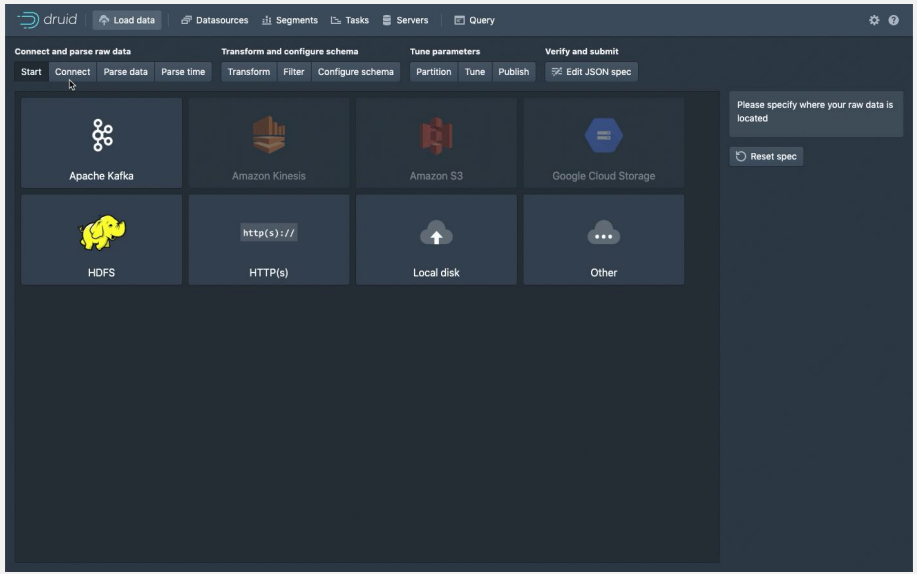
2020 **Druid 0.17 -** Superbatch, parallel auto-compaction, optimistic result merging, SQL NULLs, LDAP, Tasks sys table, lazy loading

Druid 0.18 - Joins

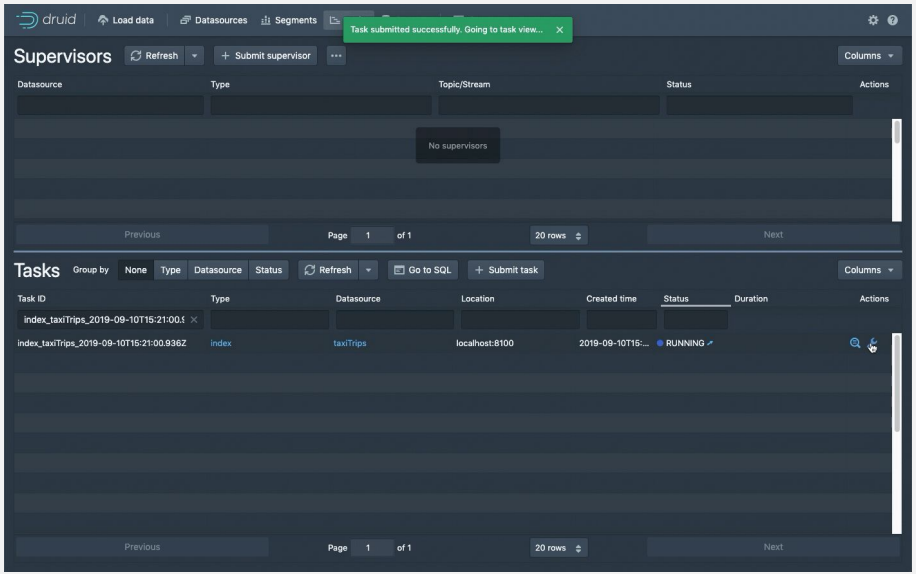
Druid 0.18.1 - Streaming ingest fix, HLL upgrade fix, improved ingestion, improved boolean filtering on ingestion

Druid 0.19 - GroupBy / Timeseries Vectorisation, Historical Index Table JOIN (beta), ranged batch ingestion append, Avro OCF, SQL inputSource, Apache Ranger authorization + Alibaba Cloud extension, RegEx LIKE

Build your ingestion spec in a GUI...



Monitor tasks, check segments, run SQL...



YAHOO!

imply

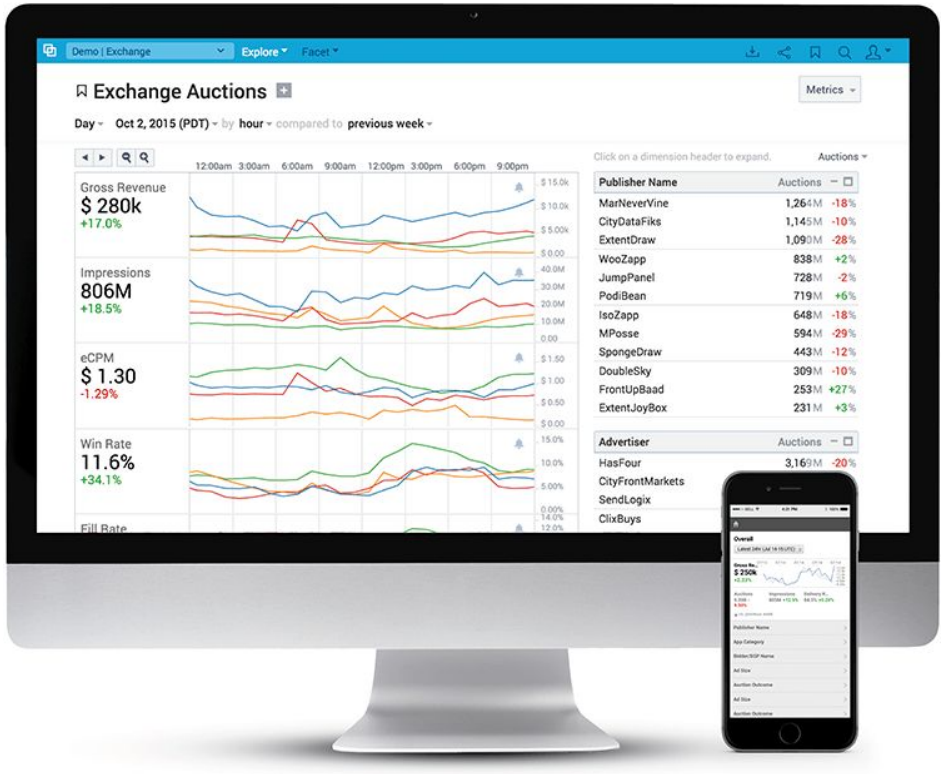
"It can do very large, OLAP-style processing on the fly in hundreds of milliseconds instead of precalculating everything...

"The performance is great ... some of the tables that we have internally in Druid have billions and billions of events in them, and we're scanning them in under a second."

TIM TULLY
VP Engineering

https://www.infoworld.com/article/2949168/hadoop/yahoo-struts-its-hadoop-stuff.html

METAMARKETS



Advertisers loaded impressions and clicks data and used web and mobile apps to optimize user and ad engagement

Billions of events per month

30+ dimensions

druid

druid

Production

Specialised collectors
Applications & APIs
Machine & Human Data
Environmental Sensors
Stream & Bulk Repositories

Delivery

PULSAR
amazon KINESIS
kafka

Store data for a limited time and deliver to a destination with guarantees.

Processing

TEZ
samza
APACHE STORM
APACHE Spark

Enrich and clean the data either event-by-event or in batches, often without reduction in the size of the input.

Consolidation
Enrichment
Transformation & Stripping
Verification & Validation
Filtering & Sorting

Storage

Cassandra
APACHE HBASE
Lucene
Microsoft SQL Server
hadoop HDFS
ORACLE
MySQL
PostgreSQL

Long-term storage for processing later on, whether for day-to-day operations, tactical planning, or strategic outlook

Logical Reasoning
Feature & Structure Discovery
Segmentation & Classification
Recommendation
Forecasting & Prediction
Anomaly Detection
Automated Decision Making
Statistical Calculations

Query

APACHE DRILL
elasticsearch
Apache Solr
HIVE

Technologies built to speed up searching, aggregation, and filtering, shrinking the input data set.

Delivery

Real-time Analytics
BI Reporting & Dashboards
Buses, Queues & Databases
Search & Filtering UIs
Applications & APIs

druid

Query

LOCATE DATA • ISSUE QUERY • MERGE RESULTS

Data

INGEST DATA • STORE DATA • RESPOND TO QUERIES

Master

ISSUE TASKS • CATALOGUE DATA • MONITOR STATE

Metadata DB

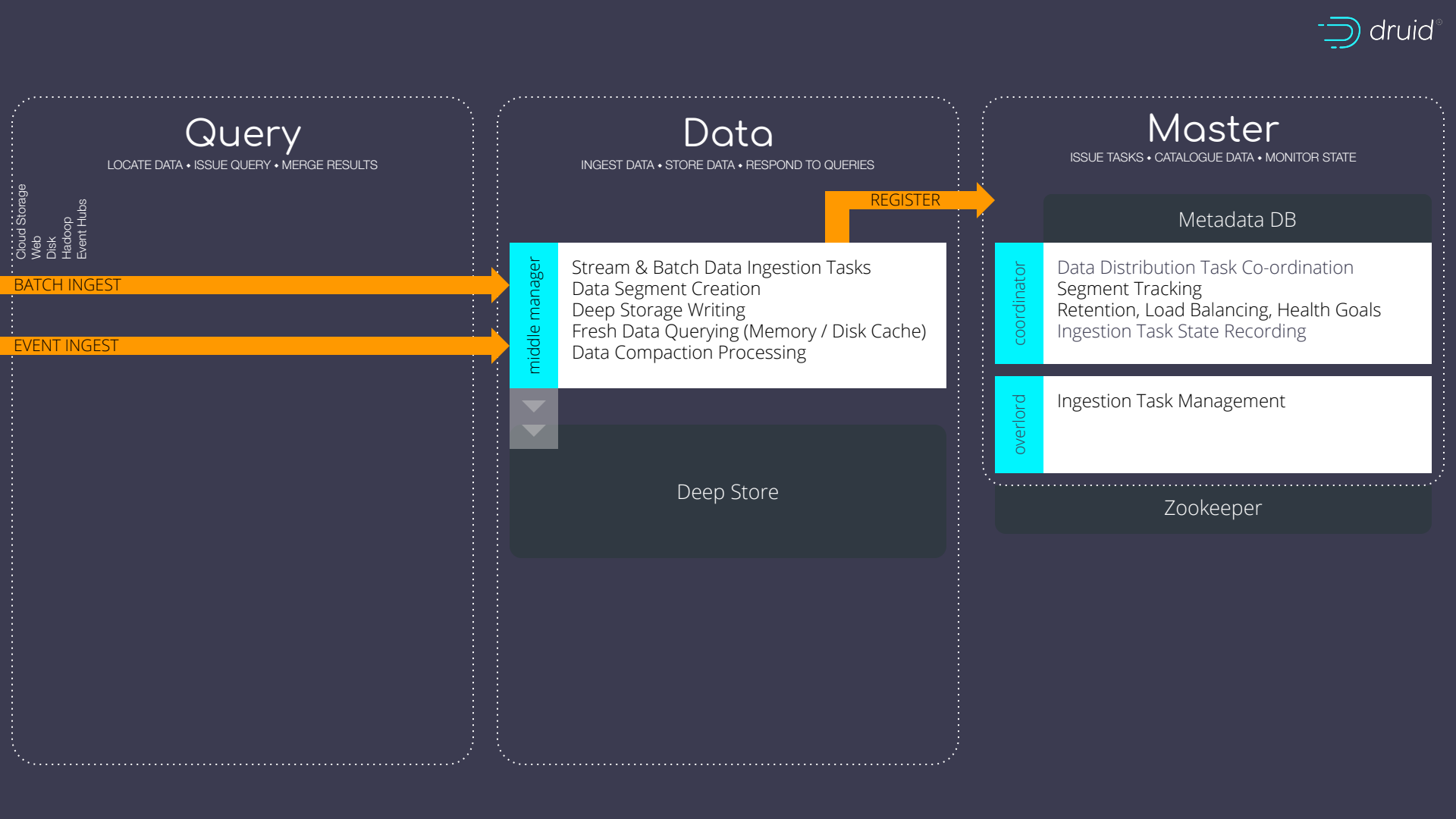
coordinator

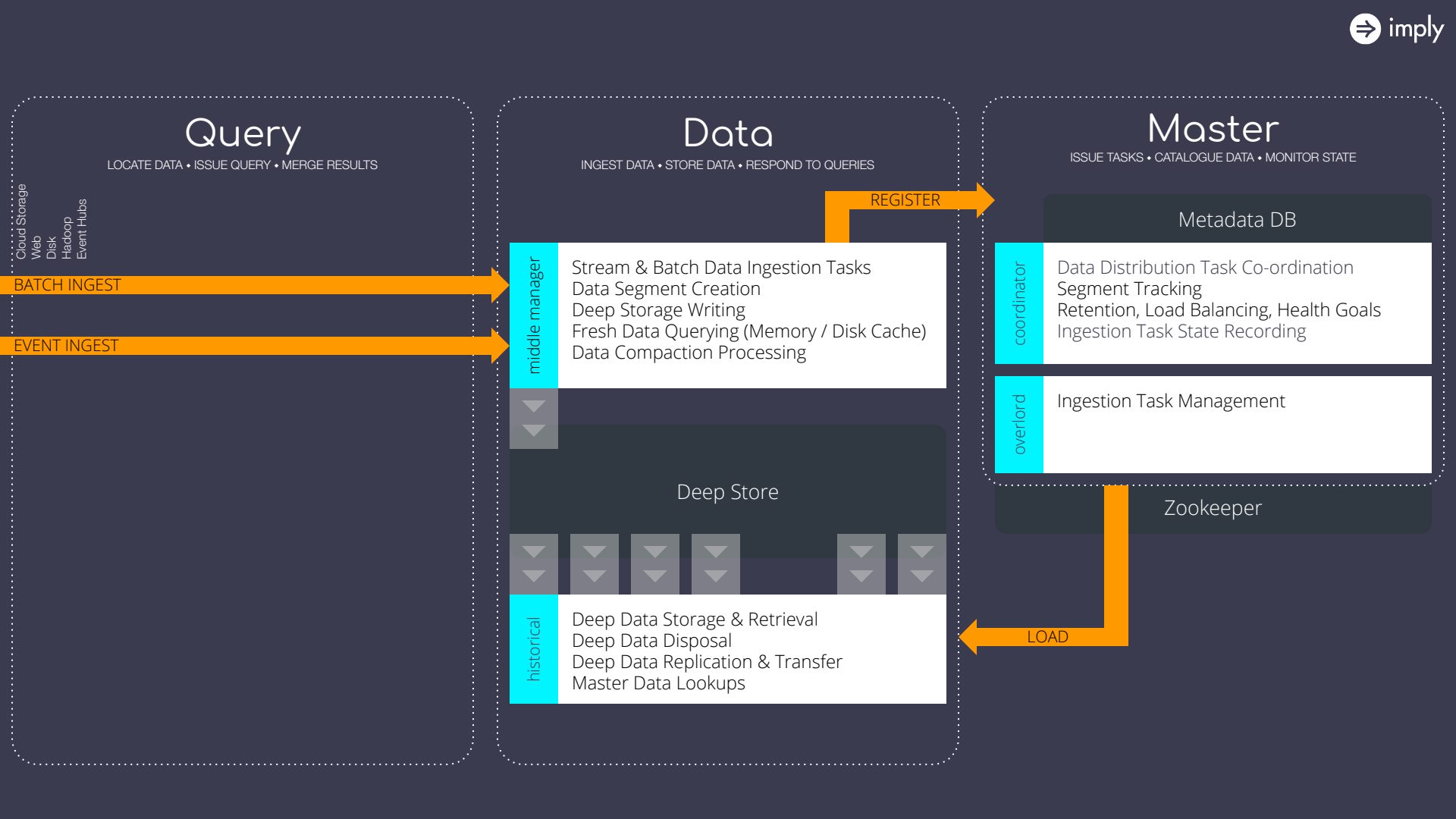
Data Distribution Task Co-ordination
Segment Tracking
Retention, Load Balancing, Health Goals
Ingestion Task State Recording

overlord

Ingestion Task Management

Zookeeper


druid
Query
LOCATE DATA • ISSUE QUERY • MERGE RESULTS
Cloud Storage
Web
Disk
Hadoop
Event Hubs
BATCH INGEST
EVENT INGEST
Data
INGEST DATA • STORE DATA • RESPOND TO QUERIES
REGISTER
middle manager
Stream & Batch Data Ingestion Tasks
Data Segment Creation
Deep Storage Writing
Fresh Data Querying (Memory / Disk Cache)
Data Compaction Processing
Deep Store
Master
ISSUE TASKS • CATALOGUE DATA • MONITOR STATE
Metadata DB
coordinator
Data Distribution Task Co-ordination
Segment Tracking
Retention, Load Balancing, Health Goals
Ingestion Task State Recording
overlord
Ingestion Task Management
Zookeeper


imply
Query
LOCATE DATA • ISSUE QUERY • MERGE RESULTS
Cloud Storage
Web
Disk
Hadoop
Event Hubs
BATCH INGEST
EVENT INGEST
Data
INGEST DATA • STORE DATA • RESPOND TO QUERIES
REGISTER
middle manager
Stream & Batch Data Ingestion Tasks
Data Segment Creation
Deep Storage Writing
Fresh Data Querying (Memory / Disk Cache)
Data Compaction Processing
Deep Store
historical
Deep Data Storage & Retrieval
Deep Data Disposal
Deep Data Replication & Transfer
Master Data Lookups
Master
ISSUE TASKS • CATALOGUE DATA • MONITOR STATE
Metadata DB
coordinator
Data Distribution Task Co-ordination
Segment Tracking
Retention, Load Balancing, Health Goals
Ingestion Task State Recording
overlord
Ingestion Task Management
Zookeeper
LOAD

imply

Query

LOCATE DATA • ISSUE QUERY • MERGE RESULTS

Data

INGEST DATA • STORE DATA • RESPOND TO QUERIES

middle manager | middle manager | middle manager | middle manager | middle manager | middle manager

Deep Store

historical | historical | historical | historical | historical | historical

Master

ISSUE TASKS • CATALOGUE DATA • MONITOR STATE

Metadata DB

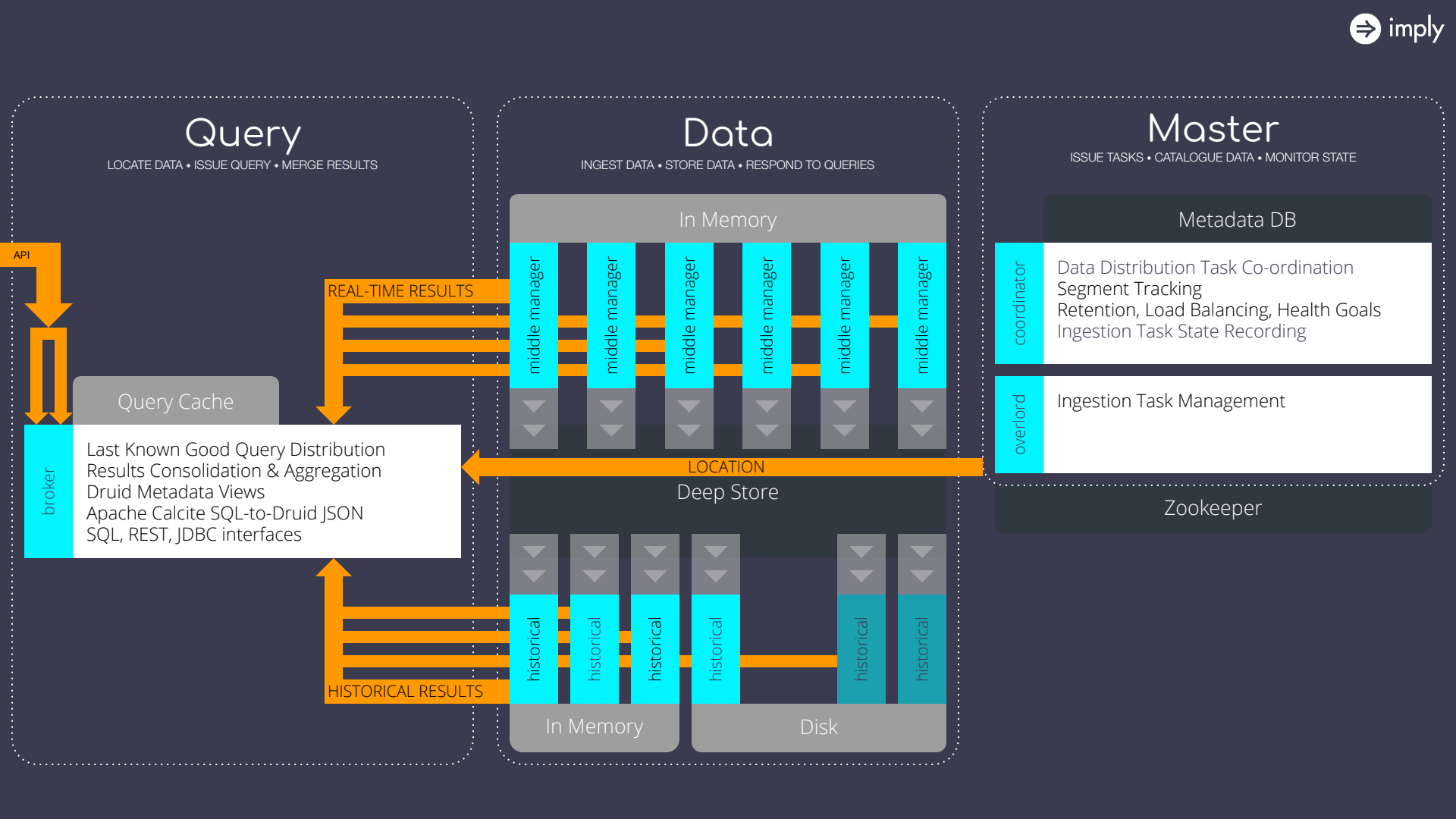
coordinator

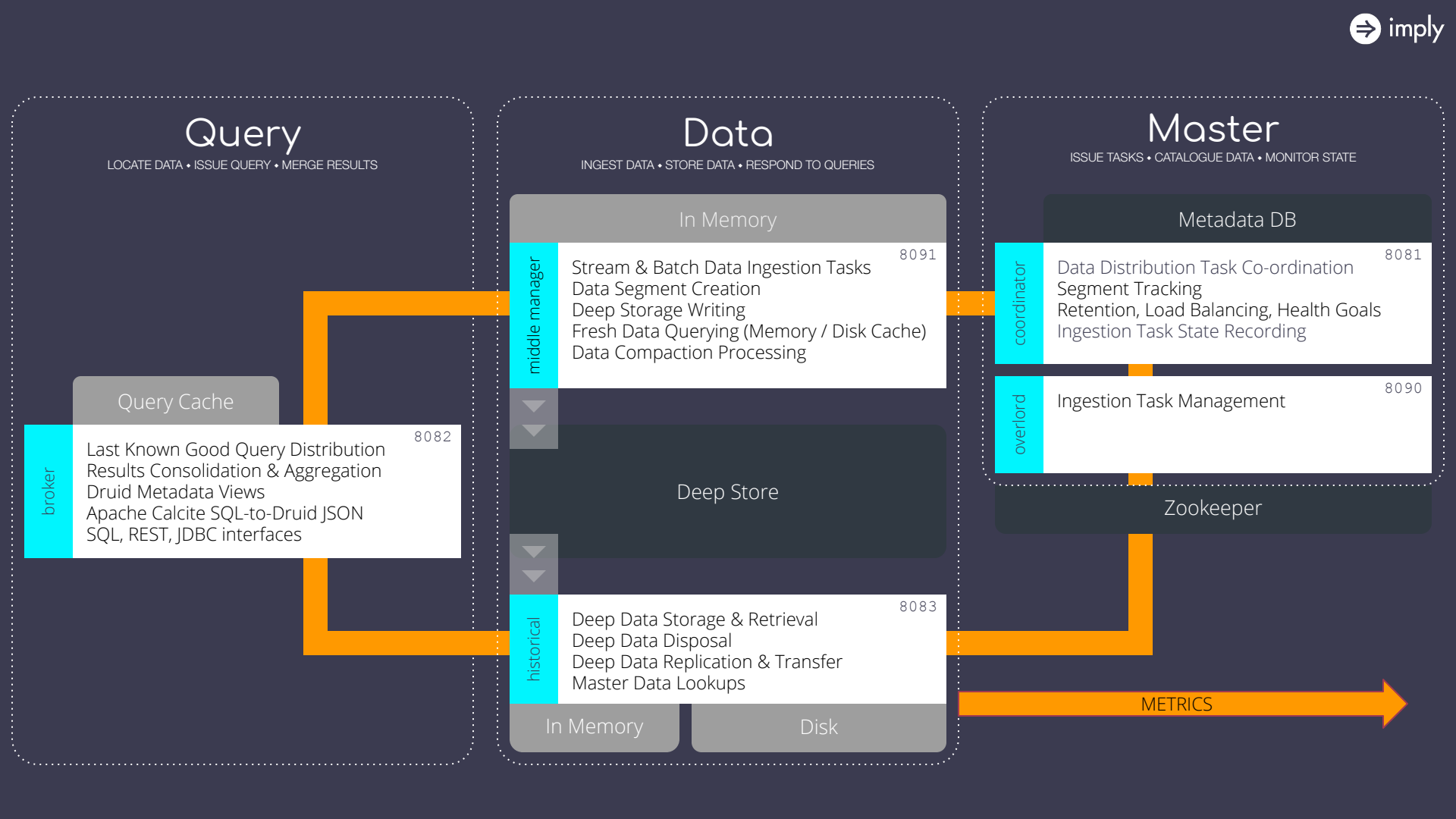
Data Distribution Task Co-ordination
Segment Tracking
Retention, Load Balancing, Health Goals
Ingestion Task State Recording

overlord

Ingestion Task Management

Zookeeper


imply
Query
LOCATE DATA • ISSUE QUERY • MERGE RESULTS
Data
INGEST DATA • STORE DATA • RESPOND TO QUERIES
Master
ISSUE TASKS • CATALOGUE DATA • MONITOR STATE
API
REAL-TIME RESULTS
In Memory
middle manager
middle manager
middle manager
middle manager
middle manager
middle manager
Metadata DB
coordinator
Data Distribution Task Co-ordination
Segment Tracking
Retention, Load Balancing, Health Goals
Ingestion Task State Recording
overlord
Ingestion Task Management
Query Cache
broker
Last Known Good Query Distribution
Results Consolidation & Aggregation
Druid Metadata Views
Apache Calcite SQL-to-Druid JSON
SQL, REST, JDBC interfaces
LOCATION
Deep Store
Zookeeper
historical
historical
historical
historical
historical
historical
HISTORICAL RESULTS
In Memory
Disk


Query
LOCATE DATA • ISSUE QUERY • MERGE RESULTS
Query Cache
broker
8082
Last Known Good Query Distribution
Results Consolidation & Aggregation
Druid Metadata Views
Apache Calcite SQL-to-Druid JSON
SQL, REST, JDBC interfaces
Data
INGEST DATA • STORE DATA • RESPOND TO QUERIES
In Memory
middle manager
8091
Stream & Batch Data Ingestion Tasks
Data Segment Creation
Deep Storage Writing
Fresh Data Querying (Memory / Disk Cache)
Data Compaction Processing
Deep Store
historical
8083
Deep Data Storage & Retrieval
Deep Data Disposal
Deep Data Replication & Transfer
Master Data Lookups
In Memory
Disk
Master
ISSUE TASKS • CATALOGUE DATA • MONITOR STATE
Metadata DB
coordinator
8081
Data Distribution Task Co-ordination
Segment Tracking
Retention, Load Balancing, Health Goals
Ingestion Task State Recording
overlord
8090
Ingestion Task Management
Zookeeper
METRICS

druid®

The Druid Story

A Druid Cluster

Data in Druid

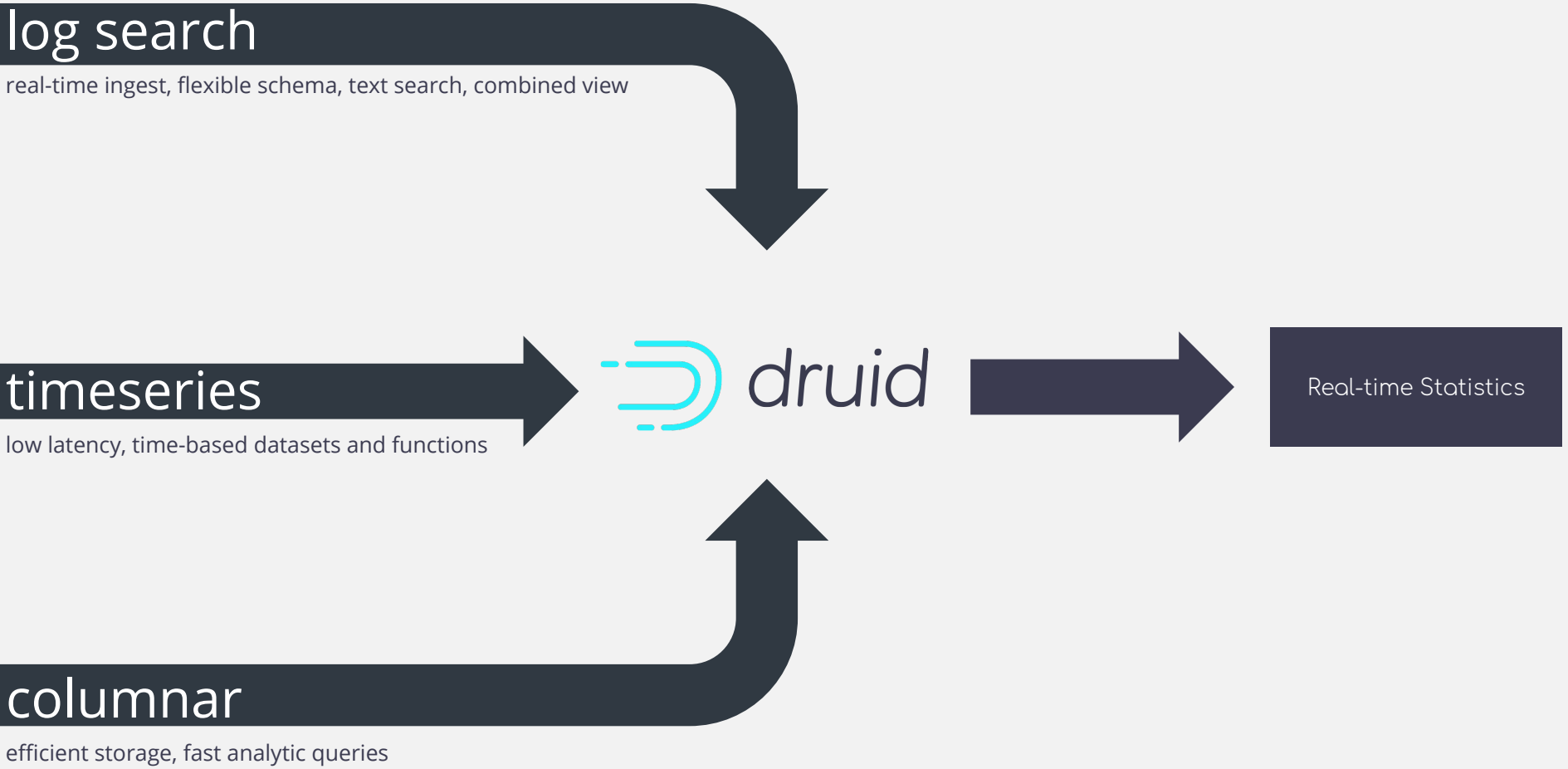
druid

- The Druid Story
- A Druid Cluster
- Data in Druid

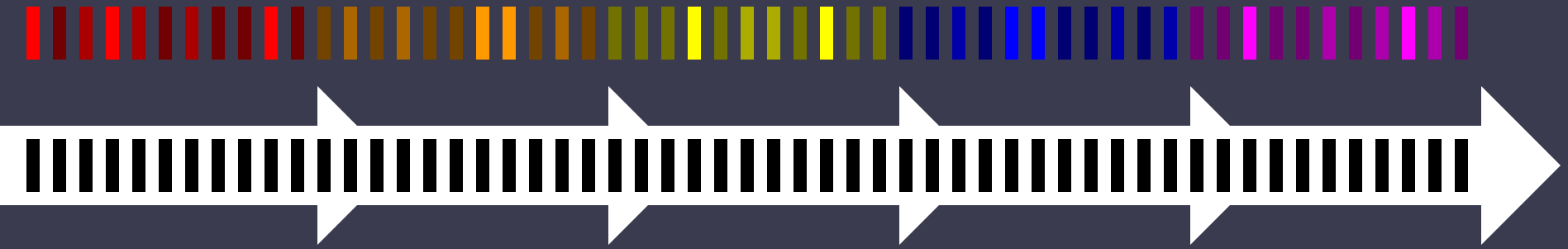
What are Druid's data optimisations?

Why are those optimisations so cool?

What types of data work best?


log search
real-time ingest, flexible schema, text search, combined view
timeseries
low latency, time-based datasets and functions
columnar
efficient storage, fast analytic queries
druid
Real-time Statistics

Source Format
Segment Period
Dimensions
Row Filters
Transforms



Source Format
Segment Period
Dimensions
Row Filters
Transforms

Columnar
Dictionary Encoded
Bitmap Indexed
Compressed

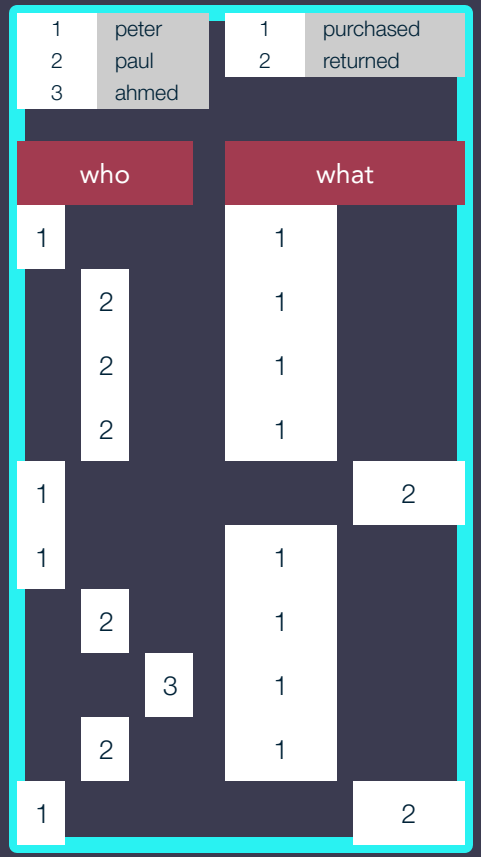
imply

when	who	what	stockChange	satisfaction
12:30	peter	purchased	5	5
12:32	paul	purchased	6	3
12:33	paul	purchased	9	4
12:35	paul	purchased	2	2
12:40	peter	returned	-2	1
12:42	peter	purchased	3	2
12:45	paul	purchased	8	3
12:50	ahmed	purchased	2	4
12:52	paul	purchased	4	2
12:55	peter	returned	-5	1

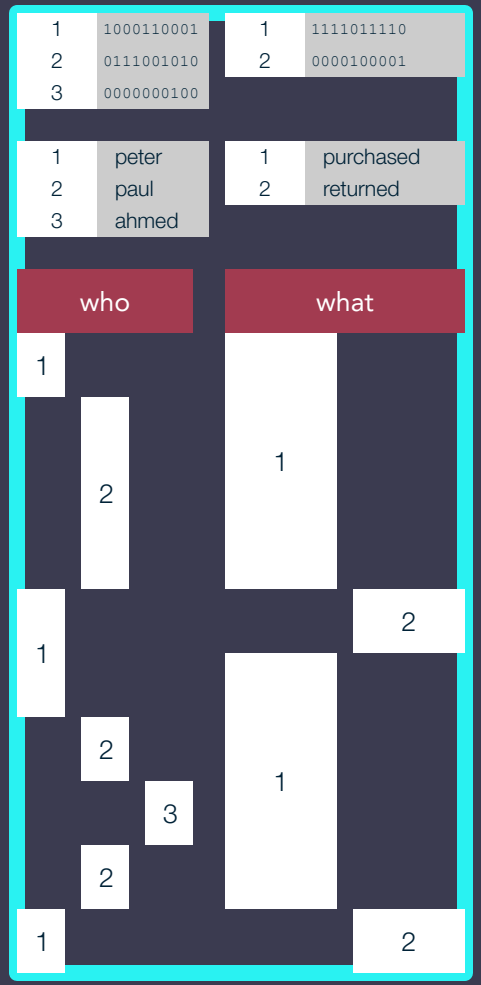
when	who	what	stockChange	satisfaction
12:30	peter	purchased	5	5
12:32	paul	purchased	6	3
12:33	paul	purchased	9	4
12:35	paul	purchased	2	2
12:40	peter	returned	-2	1
12:42	peter	purchased	3	2
12:45	paul	purchased	8	3
12:50	ahmed	purchased	2	4
12:52	paul	purchased	4	2
12:55	peter	returned	-5	1

imply

when	who	what	stockChange	satisfaction
12:30	peter	purchased	5	5
12:32	paul	purchased	6	3
12:33	paul	purchased	9	4
12:35	paul	purchased	2	2
12:40	peter	returned	-2	1
12:42	peter	purchased	3	2
12:45	paul	purchased	8	3
12:50	ahmed	purchased	2	4
12:52	paul	purchased	4	2
12:55	peter	returned	-5	1



when	who	what	stockChange	satisfaction
12:30	1	1	5	5
12:32	2	1	6	3
12:33	2	1	9	4
12:35	2	1	2	2
12:40	1	2	-2	1
12:42	1	1	3	2
12:45	2	1	8	3
12:50	3	1	2	4
12:52	2	1	4	2
12:55	1	2	-5	1



when	who	what	stockChange	satisfaction
12:30	1	1	5	5
12:32	2		6	3
12:33			9	4
12:35			2	2
12:40	1	2	-2	1
12:42		1	3	2
12:45	2		8	3
12:50	3		2	4
12:52	2		4	2
12:55	1	2	-5	1

imply

Dimensions

Filters ~ Sorting ~ Functions ~ Aggregate ~ Top N ~ Group By ~ Bucketing

Windowed Aggregates

Metrics ~ Sets ~ Approx Count / Quantiles

when	who	what	stockChange	satisfaction
12:30	1	1	5	5
12:32	2		6	3
12:33			9	4
12:35			2	2
12:40	1	2	-2	1
12:42		1	3	2
12:45	2		8	3
12:50	3		2	4
12:52	2		4	2
12:55	1	2	-5	1

imply

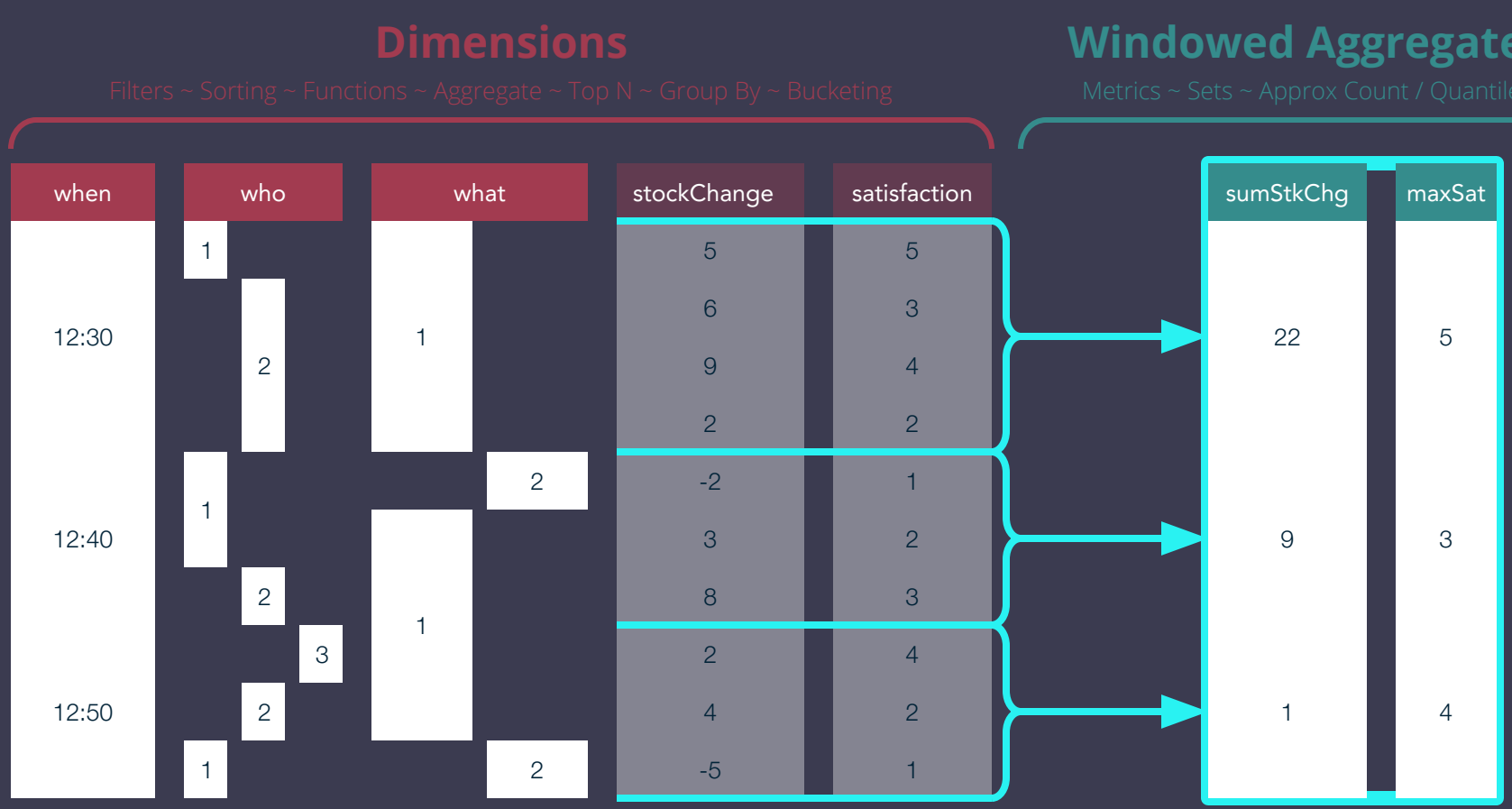
Dimensions

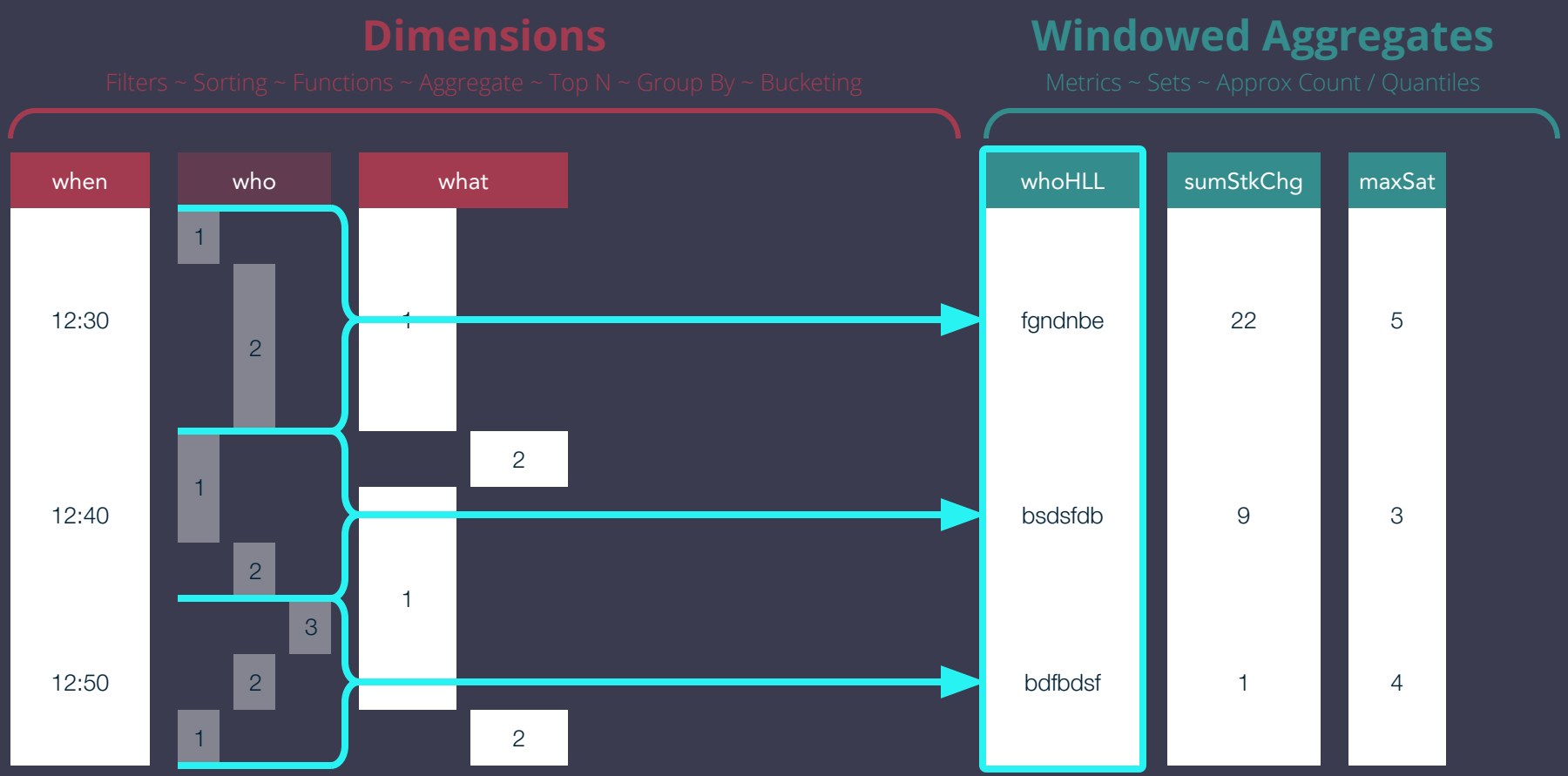
Filters ~ Sorting ~ Functions ~ Aggregate ~ Top N ~ Group By ~ Bucketing

Windowed Aggregates

Metrics ~ Sets ~ Approx Count / Quantiles

when	who	what	stockChange	satisfaction
12:30	1	1	5	5
	2		6	3
			9	4
			2	2
12:40	1	2	-2	1
	2	1	3	2
	3		8	3
			2	4
12:50	2		4	2
	1	2	-5	1


Dimensions
Filters ~ Sorting ~ Functions ~ Aggregate ~ Top N ~ Group By ~ Bucketing
Windowed Aggregates
Metrics ~ Sets ~ Approx Count / Quantiles
when
who
what
stockChange
satisfaction
sumStkChg
maxSat
12:30
12:40
12:50
1
2
1
2
1
2
3
1
2
1
2
5
6
9
2
-2
3
8
2
4
-5
5
3
4
2
1
2
3
4
22
9
1
5
3
4


Dimensions
Filters ~ Sorting ~ Functions ~ Aggregate ~ Top N ~ Group By ~ Bucketing
Windowed Aggregates
Metrics ~ Sets ~ Approx Count / Quantiles
when
who
what
whoHLL
sumStkChg
maxSat
12:30
12:40
12:50
1
2
1
2
1
2
3
2
1
2
1
2
fgndnbe
22
5
bsdsfdb
9
3
bdfbdsf
1
4

wn	wt
12:30	1
12:40	2
	1
12:50	2

wH	ssc	ms
fgndnbe	22	5
bsdsfdb	9	3
bdfbdsf	1	4

druid

when	who	what	stockChange	satisfaction
12:30	peter	purchased	5	5
12:32	paul	purchased	6	3
12:33	paul	purchased	9	4
12:35	paul	purchased	2	2
12:40	peter	returned	-2	1
12:42	peter	purchased	3	2
12:45	paul	purchased	8	3
12:50	ahmed	purchased	2	4
12:52	paul	purchased	4	2
12:55	peter	returned	-5	1

wn	wt
12:30	1
12:40	2
12:50	1
	2

wH	ssc	ms
fgndnbe	22	5
bsdsfdb	9	3
bdfbdsf	1	4


druid®
+
community

The **pain** that led to Druid being made

What the **components** of a cluster are

How Druid **optimises** data for on-demand statistics

Questions & Answers

druid®

http://druid.apache.org

Apache Distribution
https://github.com/apache/druid

Imply Distribution
https://imply.io/get-started

Druid Community
https://druid.apache.org/community/

Google Groups Druid User Forum
https://groups.google.com/

Meetup Groups
https://www.meetup.com/pro/apache-druid/

ASF Slack
#druid

Twitter
@druidio

Add Apache Druid
as a skill