# HareQL：快速HBase查詢工具的發展過程

Development of HBase Client and HareQL

**Mon-Fong Mike Jiang**
江孟峰
**Kuan-Yu Hubert Fan-Chiang**
范姜冠宇
**Tienyu Rebecca Lin**
林恬伃

# About Us



- **Providing IT solutions**
  - System development for big data solutions
  - Smart manufacturing related services
  - Financial data systems
  - Telecommunication data systems

  - We are the Cloudera certificated professional services team

- **Big data Product since 2011**
  - •Hare Data platform
  - •2013 HSP  Innovative Product Award
  - •2014 Golden Award of the TOP 100 Innovative Products
  - •Cloudera Certified Technology (Only one in Taiwan)

cloudera CERTIFIED

is-land
is-land Systems Inc.

# What is Hare

- It’s a NoSQL Database which is based on HBase
- Support SQL to HBase directly
- Provide DBMS-like Web UI
- Provide JDBC/ODBC and Restful Service

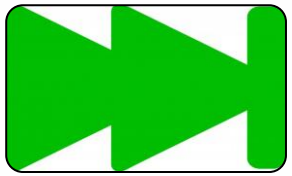
# Why Hare ?

## Easy

- Click and start to use it
- Friendly user interface
- To involve your big data rapidly

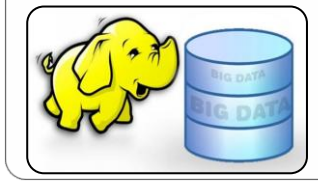
## Comfortable

- SQL language supported
- Data type management
- Multi-Cluster in one client



## Faster

- Quick access to the data in HBase
- Powerful query engine for better performance




## Compatible

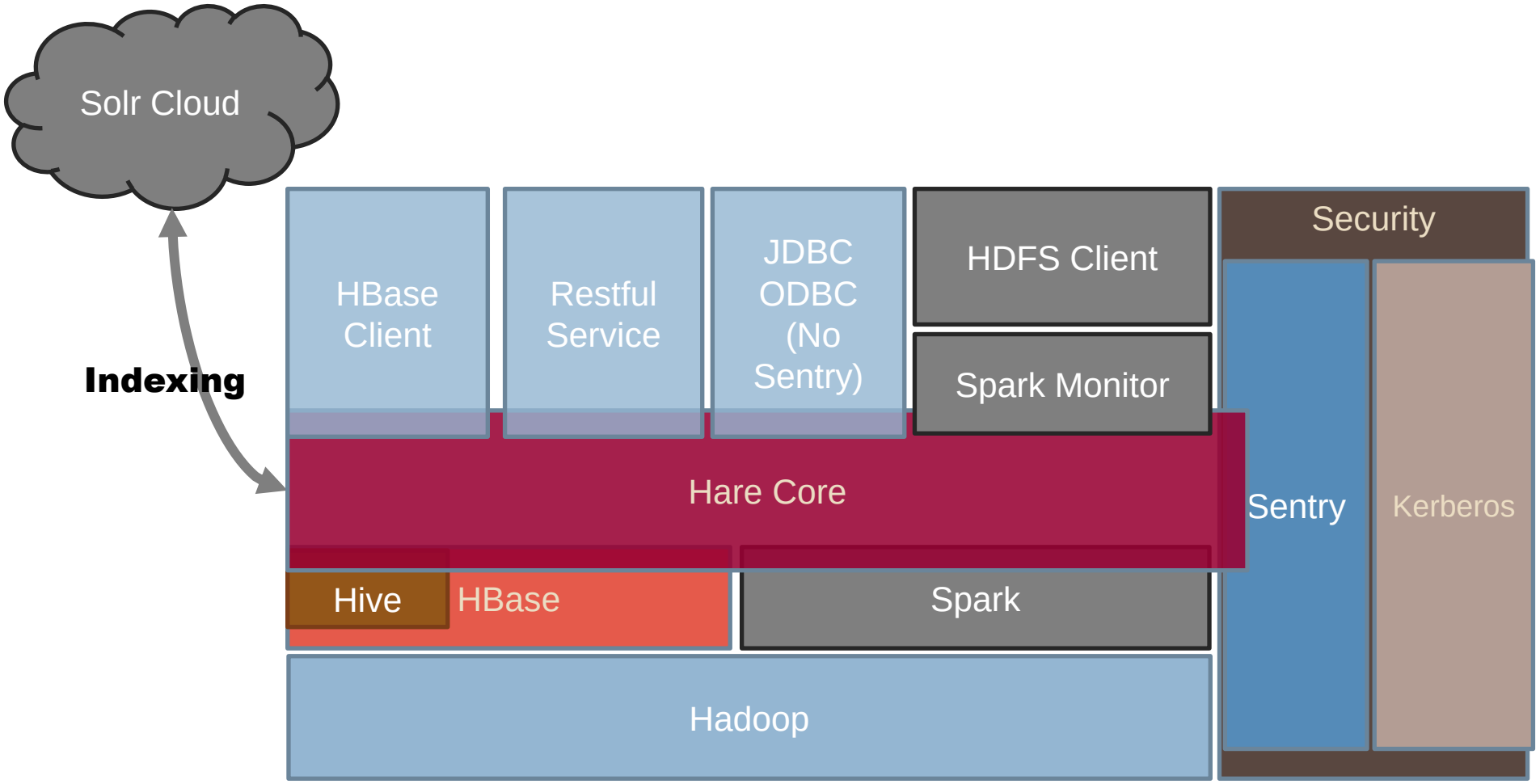
- Based on the Hadoop/HBase System
- Highly compatible in ecosystem

is-land HARE HareDB.com

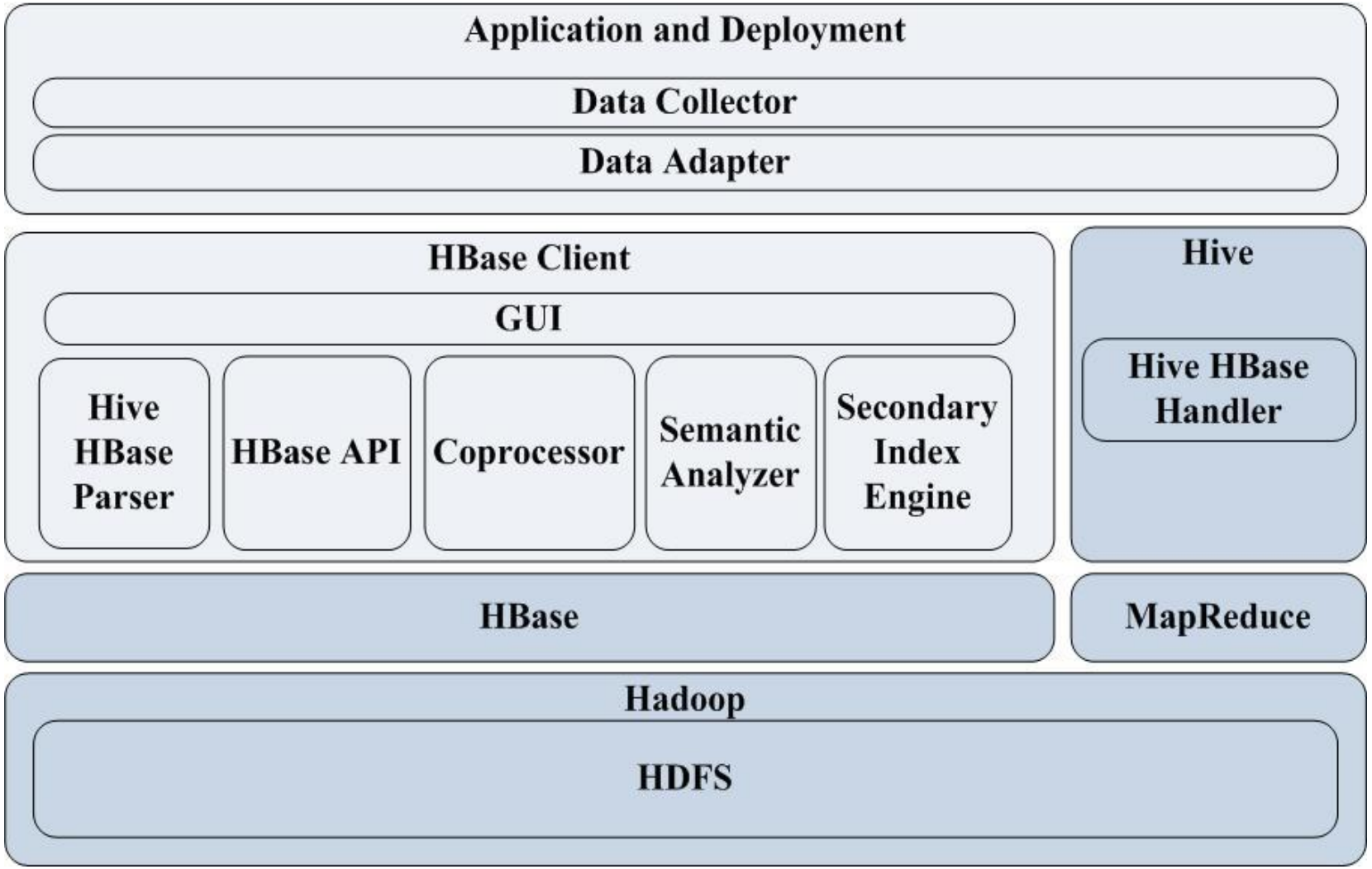
# Features

- **Easy use (Web UI)**
  - Easy install
  - Friendly UI
- **One Client ; Many Clusters (Connection Manager)**
- **Bulkload UI**
- **Meta Manager (Schema Manager)**
- **Relation between HBase Table and Hive Table**
- **HareQL (High Speed SQL Query in HBase)**
  - JDBC Driver
  - ODBC Driver (not support sentry)
  - Restful Services

# Software Stack

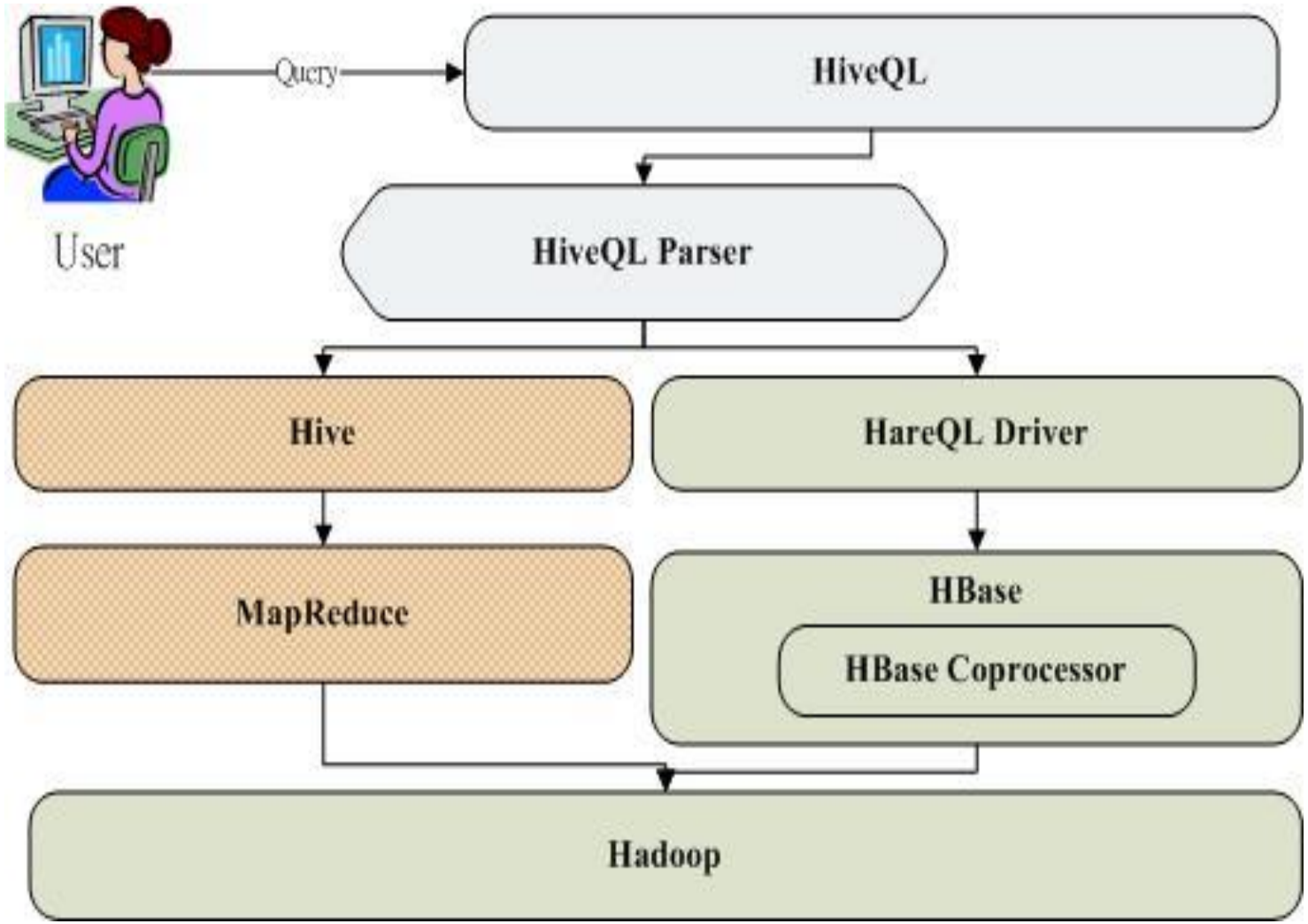


# System Architecture



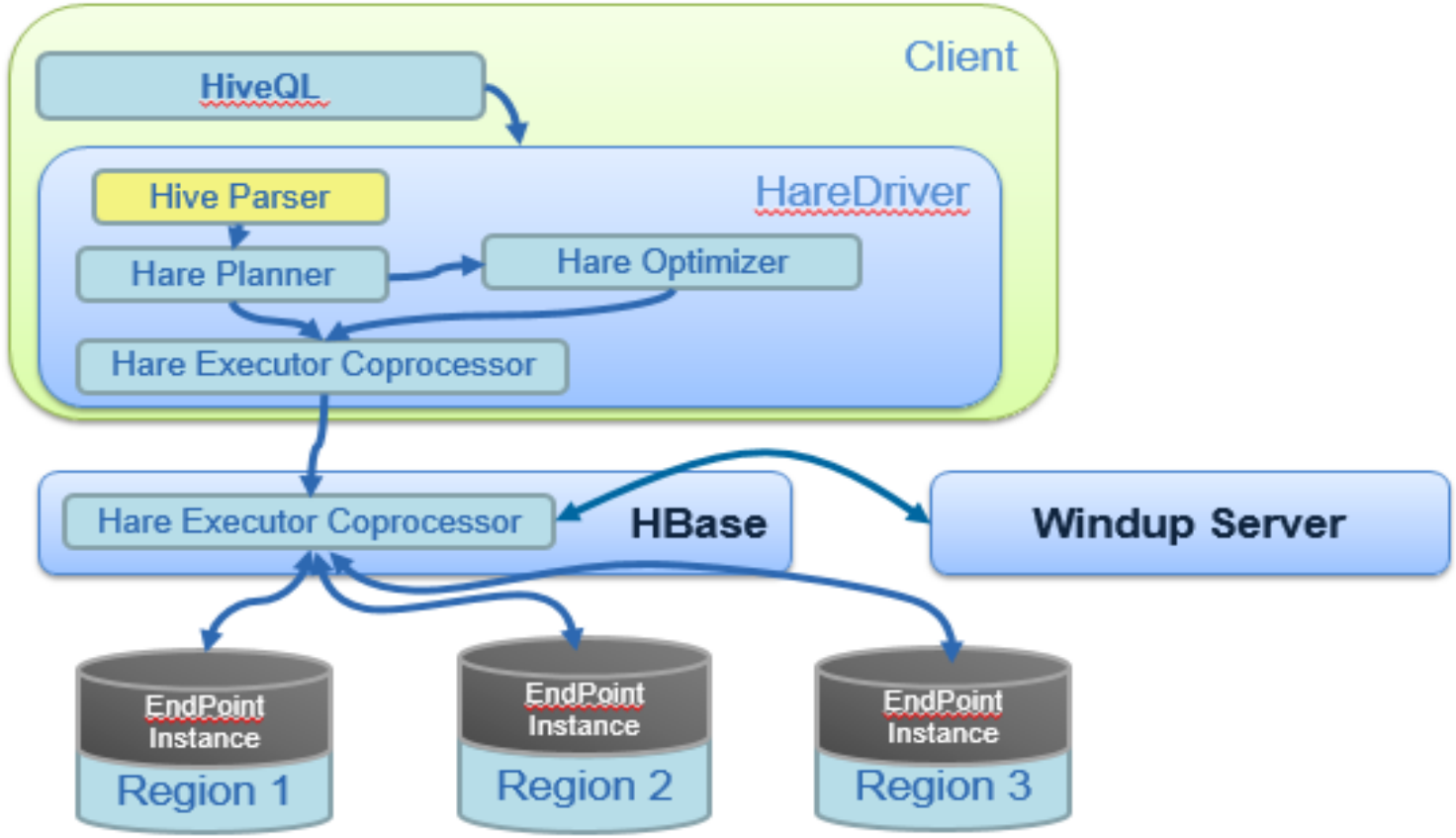
# HareQL

- **Hive: MapReduce**
- **We replace MapReduces in Hive to HBase coprocessors. We call the language “HareQL”.**
- **HareQL has some advantages as below.**
  - Low- latency
  - Query HBase table directly
  - High performance

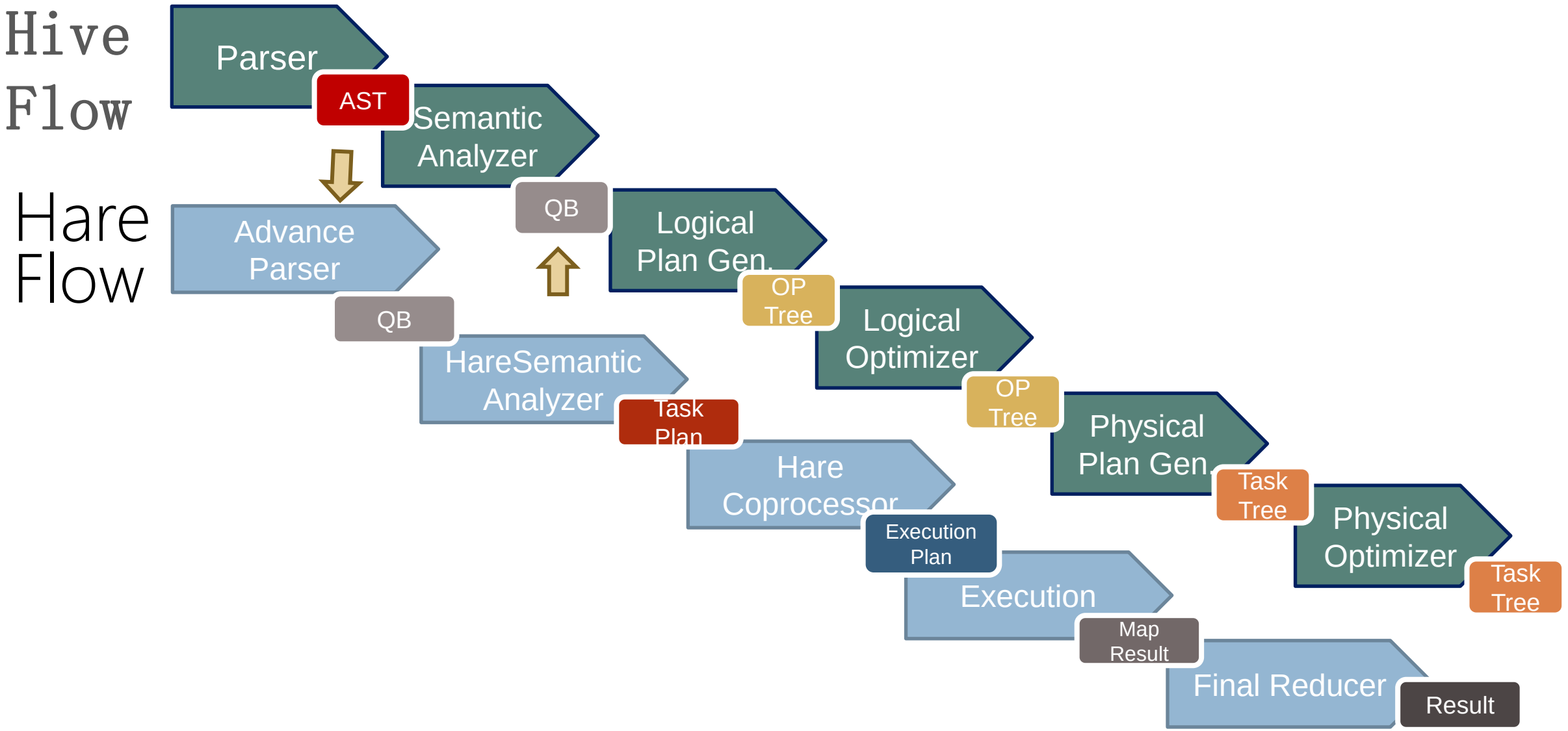


# HareQL Architecture

- **Hive Parser make us support HiveQL**

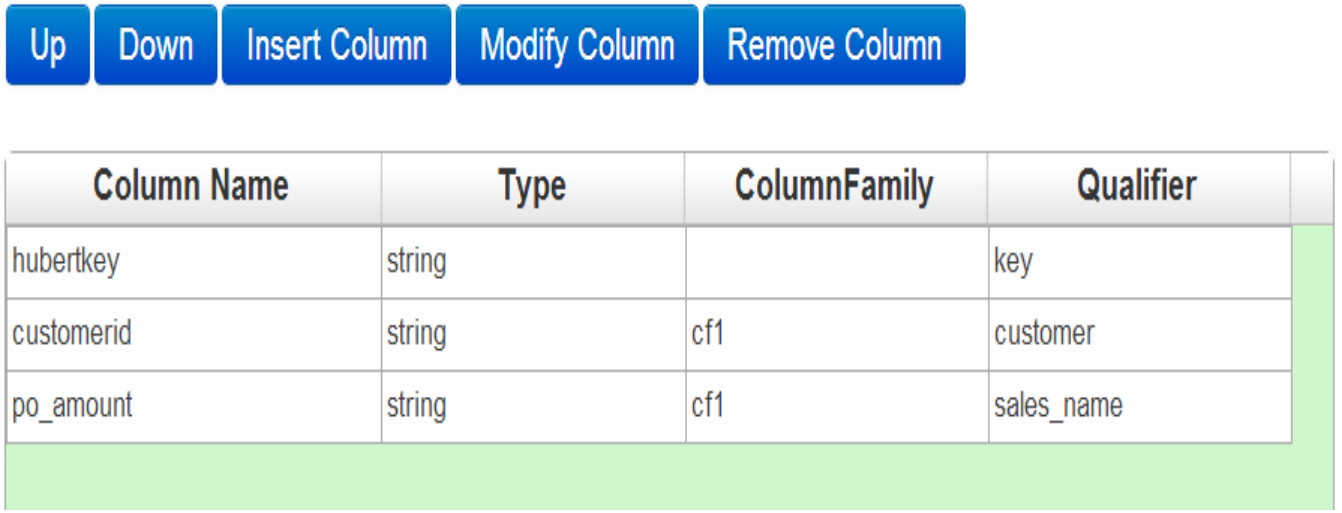


# From Hive to Hare

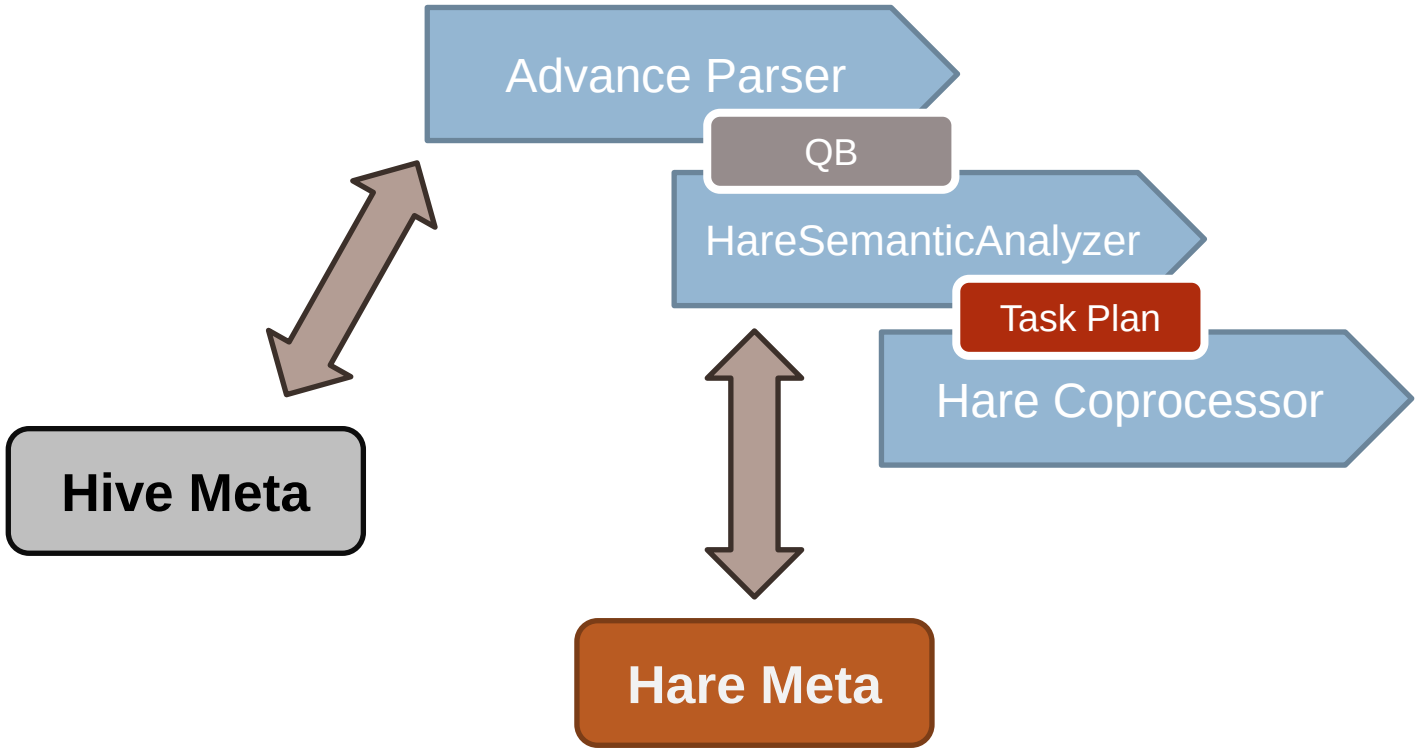


# Metadata

- As we know, anything that can be converted to an array of bytes can be stored in HBase. However, we have to convert the data back correctly, or we can't recognize the data.
- We integrated meta-store of Hive to HBase Client. We call the data type of HBase column "Meta data".



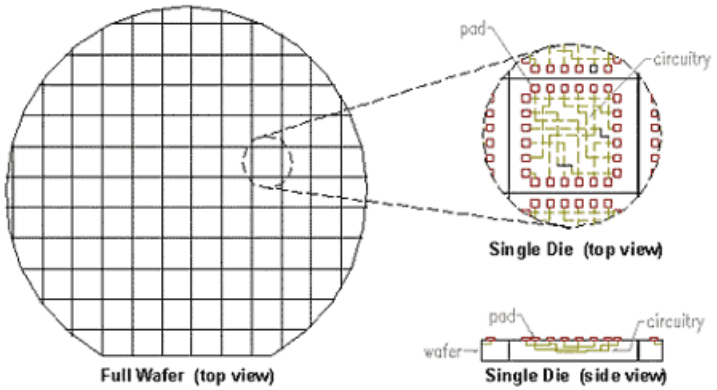
# When to get Metadata ?



# Hare Restful Service

- Table manipulation
- Row manipulation
- Bulkload data
- Sending SQL
- Scanning
- Metadata manipulation

# Semiconductor application

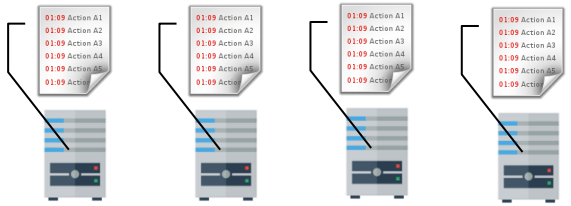


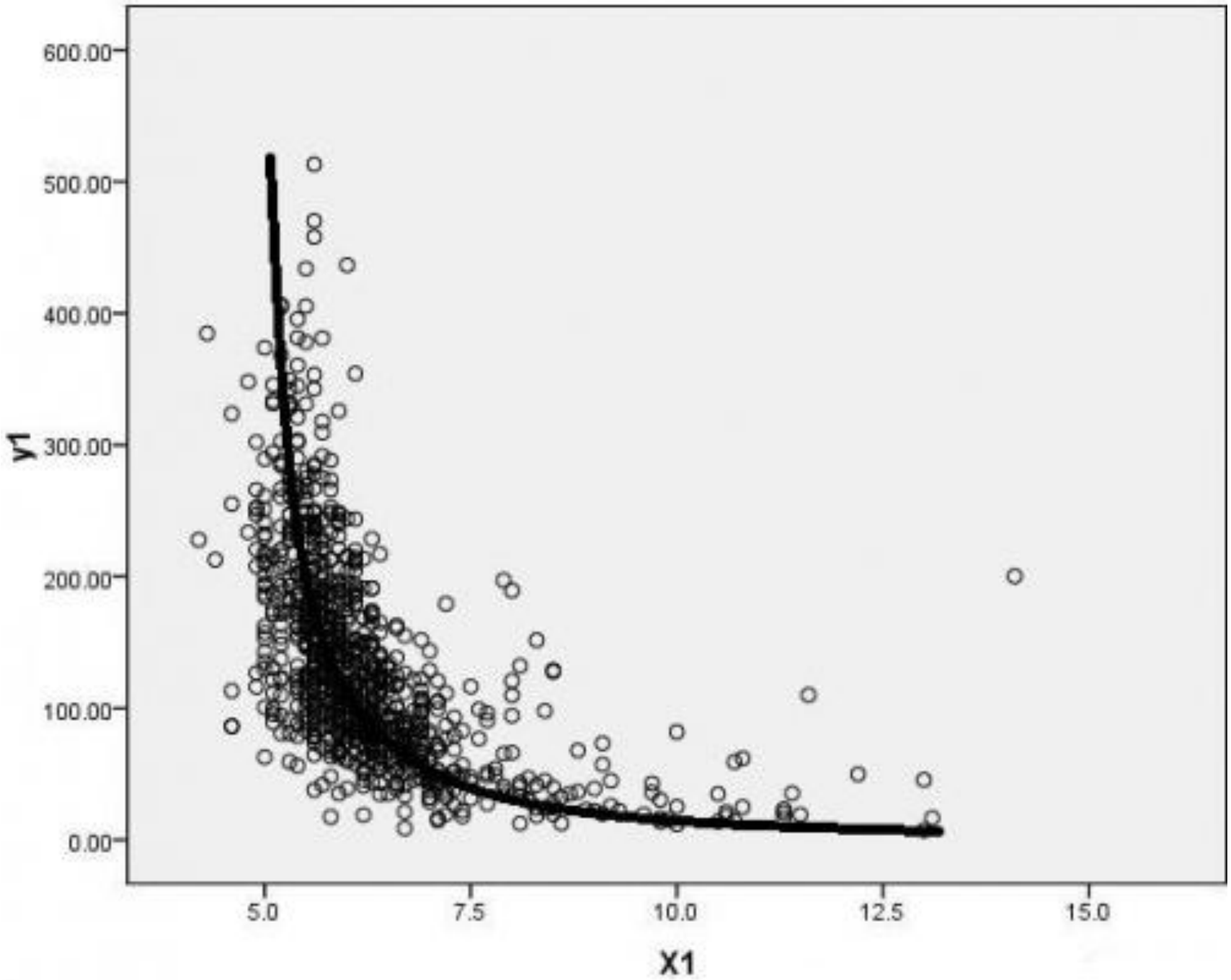

## Structured Data

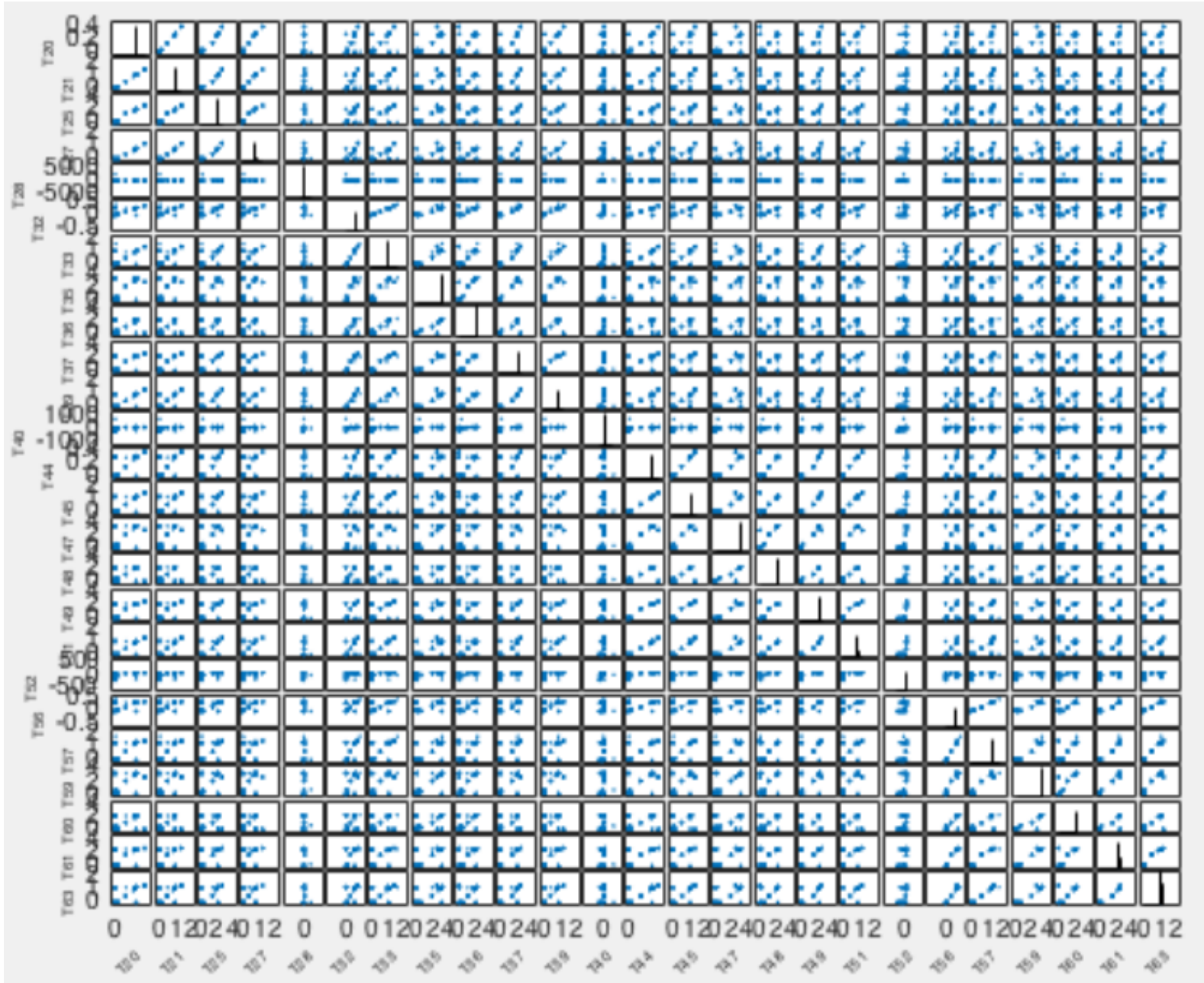
SQM System | SPC System | EDA System | PLM System

SCM System | MES System | FDC System | ERP System

## Un-Structured Data




600.00
500.00
400.00
300.00
200.00
100.00
0.00
y1
5.0
7.5
10.0
12.5
15.0
X1

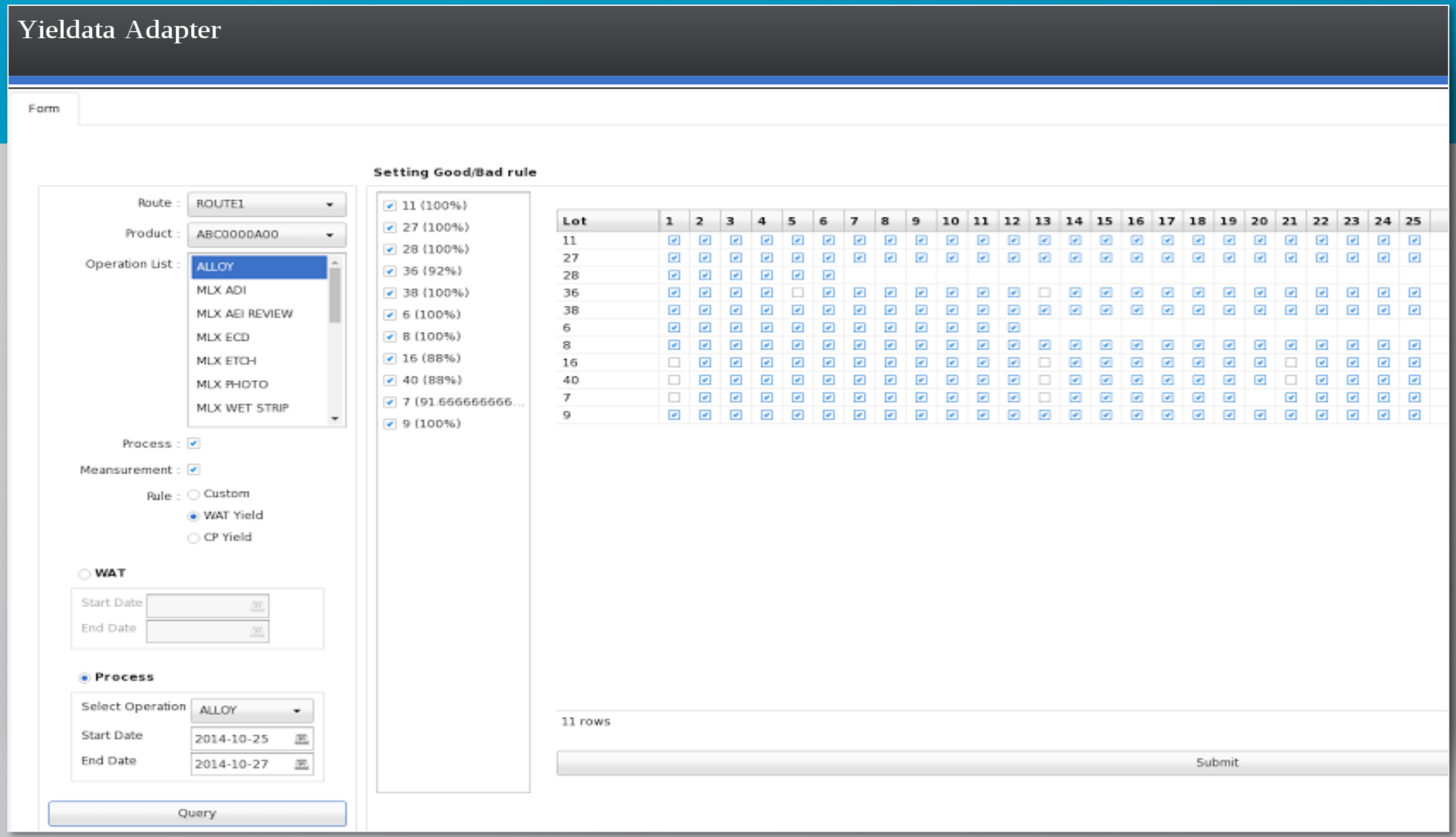


# Schema Design in HBase

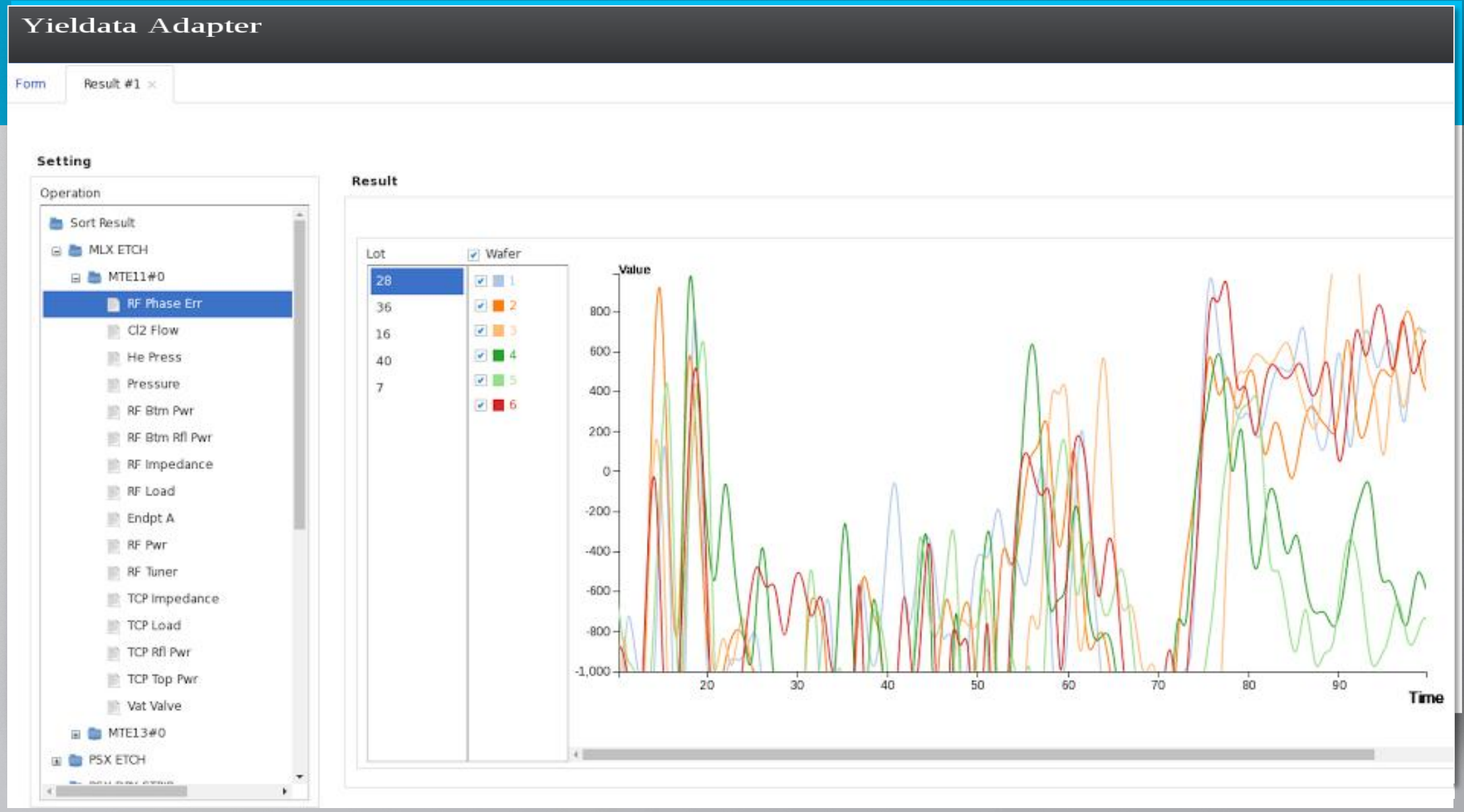
- Designed the row key according to the access request

| Table | fdcTrace | | |
|---|---|---|---|
| Rowkey | [salt][Lot perfix][OperNo][EQID][ChamberID][Stb slot][DateTime][ID] | | |
| | name | type & size | desc |
| | salt | char | value: 64+week_no(1..52) |
| | Lot prefix | String(8) | ex: AC418075 |
| | OperNo | String(64) | ex: 010.00010 |
| | EQID | String(64) | ex: OXED204 |
| | ChamberID | String(64) | ex: A |
| | Stb slot | String(2) | ex: 02 |
| | DateTime | String(15) | ex: 140622000036630 => 2014-06-22 00:00:36.630 |
| | ID | String(10) | ex: 155560 |

# Application – Yieldata



# Yieldata – Root Cause Ranking



# Thank you


- is-land Systems Inc.
- Company：www.is-land.com.tw
- Big Data：www.HareDB.com
- Email：service@haredb.com
- Addr : 新竹科學園區展業二路4號3樓