



了解openLookEng，从试用开始

廖登宏



目录

01

openLookKeng Tryme介绍

02

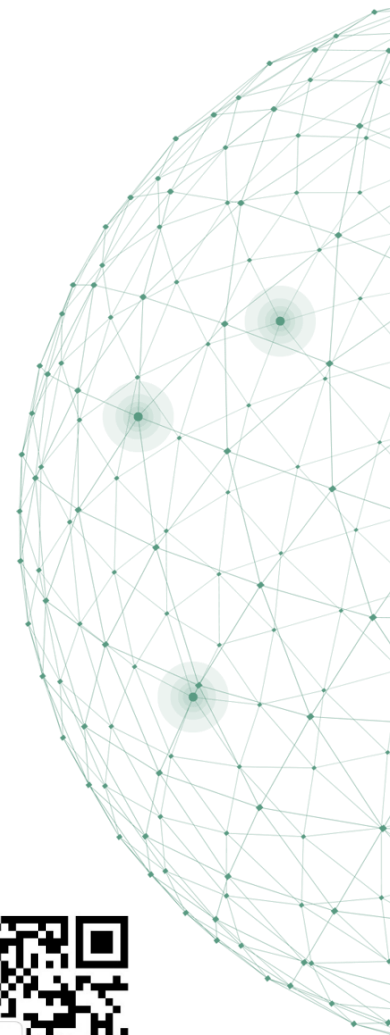
openLookKeng关键技术介绍与演示

03

openLookKeng在IDE中编译与运行



openLookKeng微信公众号 openLookKeng微信小助手



Tryme界面

<https://tryme.openlookeng.io/>

openLookEng lk

[Add Catalog](#) [Run](#) Catalog: system Schema: metadata

Search ...

- + hive
- + memory
- system
 - + information_schema
 - + jdbc
 - + metadata
 - runtime
 - nodes
 - queries
 - tasks
 - transactions
- + tpcds
- + tpch

1 `select * from analyze_properties;`

My saved queries All queries Results

`select * from analyze_properties`

| catalog_name | property_name | default_value | type | description | |
|--------------|---------------|---------------|-----------------------|---------------------------|--|
| hive | partitions | | array(array(varchar)) | Partitions to be analyzed | |

- Tryme的常用操作
- 常用的Catalog介绍

目录

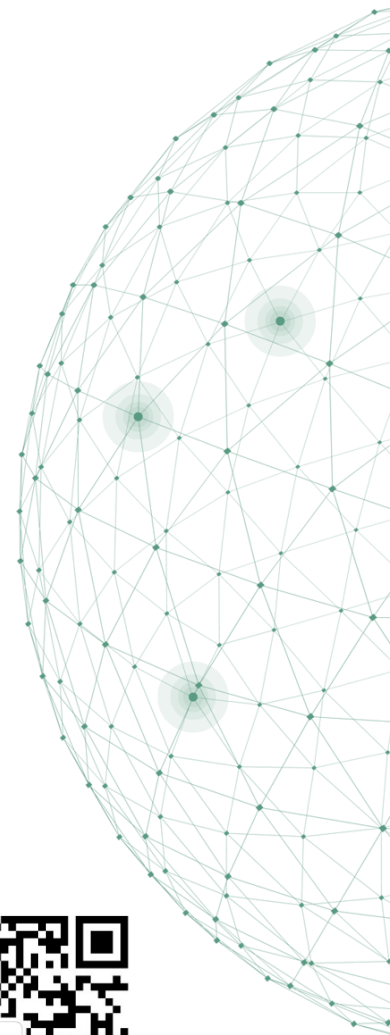
01 openLookKeng Tryme介绍

02 **openLookKeng关键技术介绍与演示**

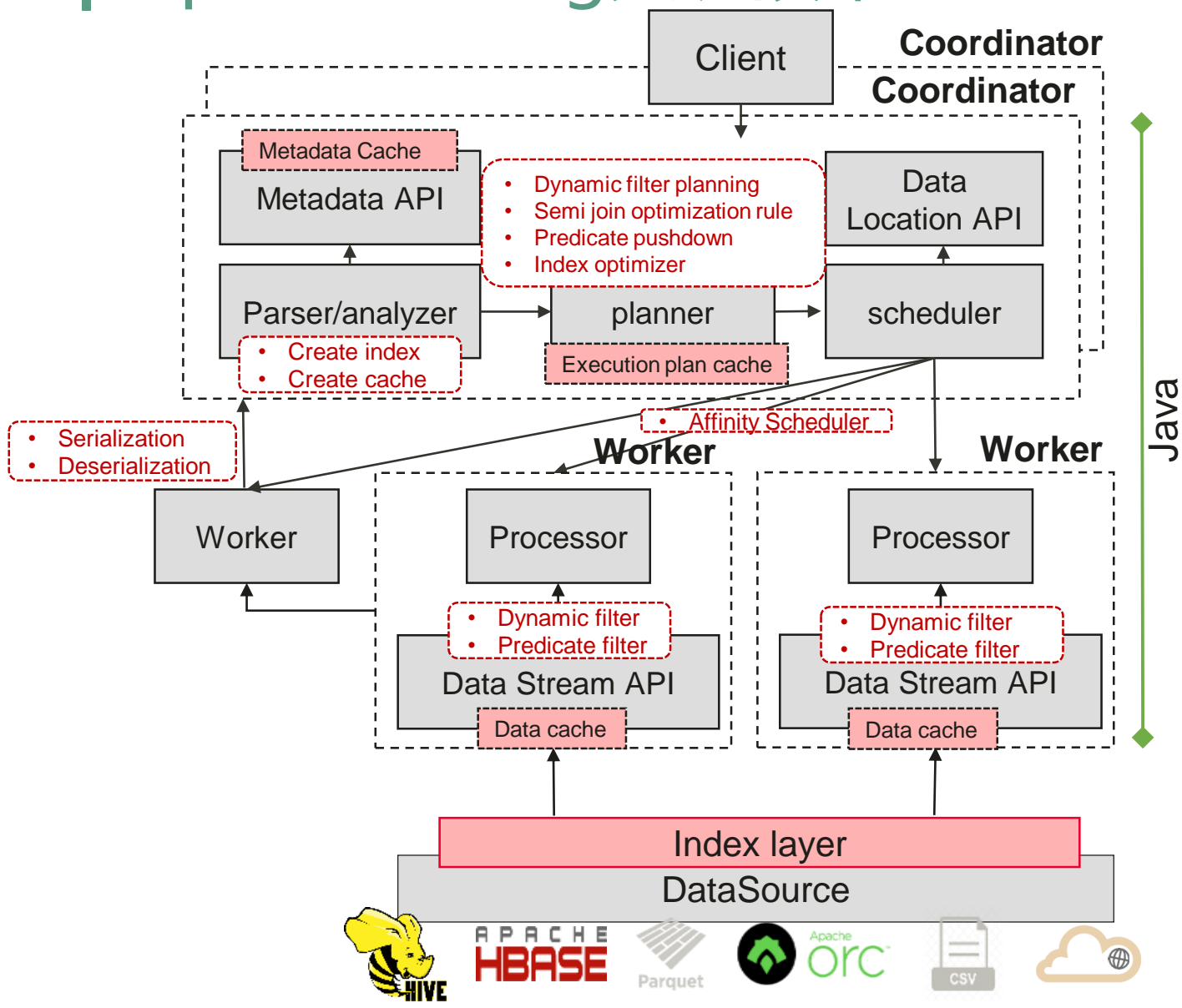
03 openLookKeng在IDE中编译与运行



openLookKeng微信公众号 openLookKeng微信小助手



openLookKeng关键技术



❑ 数据源侧，更适应openLookKeng

- 分桶/分区
- 小文件合并
- 查询字段排序

❑ 引擎层，增强交互式查询能力

- 缓存加速:

- 执行计划缓存
- 元数据缓存
- 增量列式缓存

- 优化器:

- 谓词下推
- **动态过滤**
- RBO&CBO

- 自适应调度器

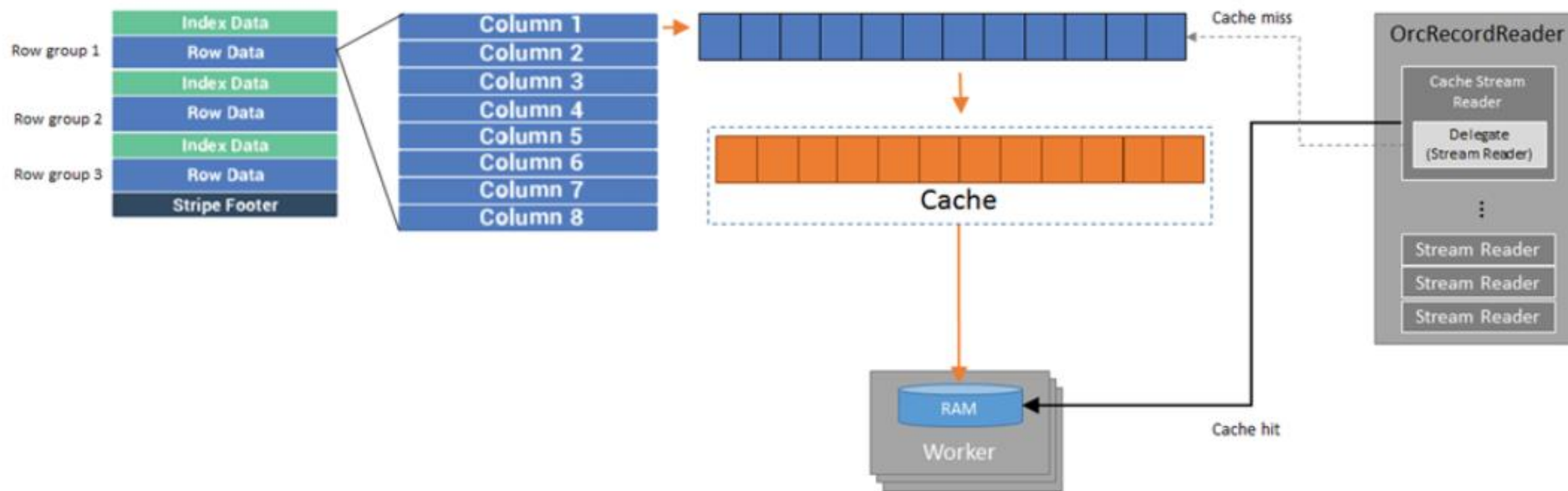
❑ 额外层，加速交互式查询

- **Heuristic index layer**

(bitmap/bloomfilter/min-max)

- Data cache layer
- 序列化&反序列化

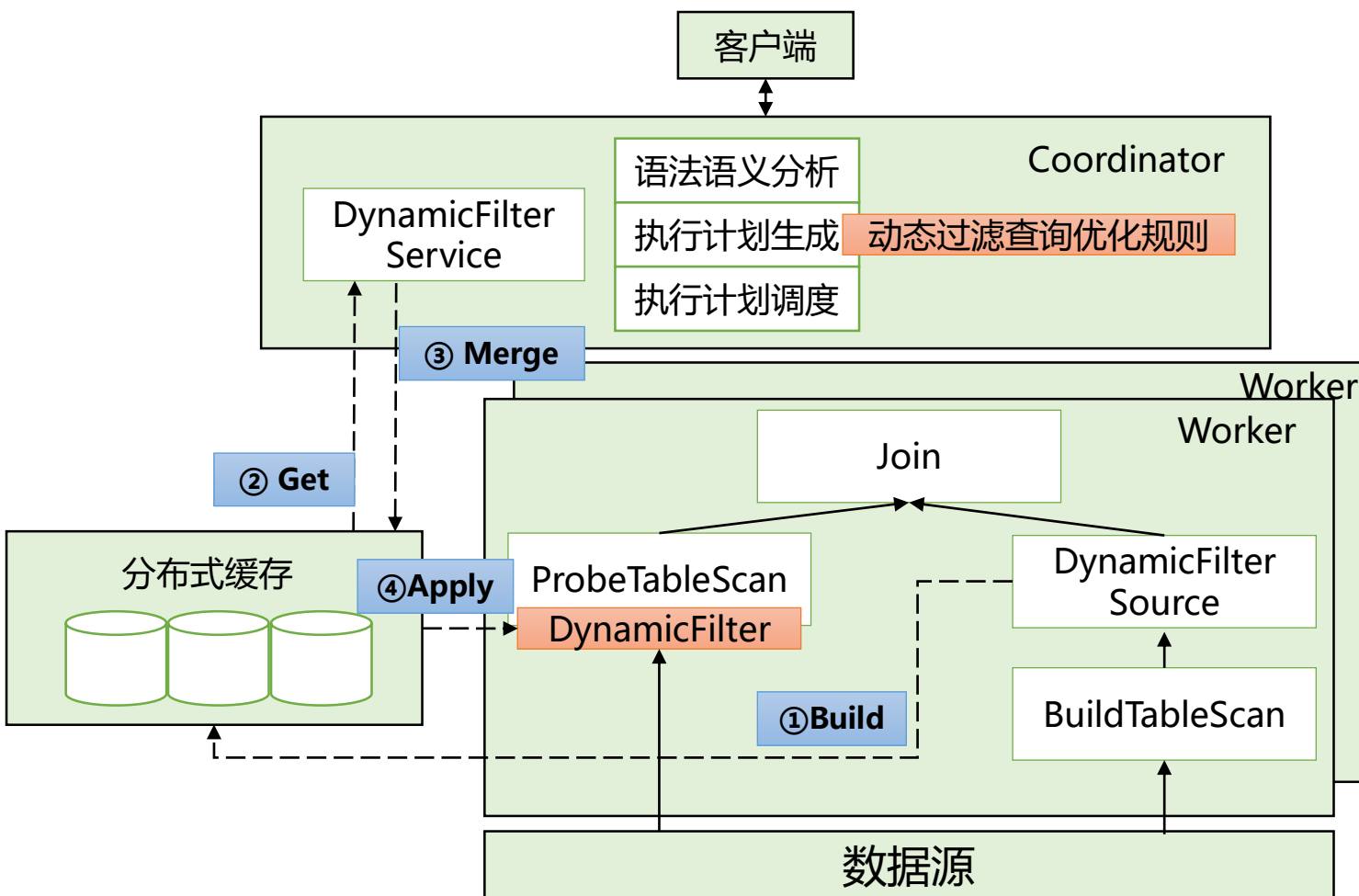
ORC Cache



```
SELECT count(*) FROM store_sales WHERE store_sales.ss_sold_date_sk BETWEEN 2451484 AND 2451513
```

Dynamic Filtering

依靠join条件以及build侧表读出的数据，运行时生成动态过滤条件（dynamic filters），应用到probe侧表的table scan阶段，从而减少参与join操作的数据量，有效地减少IO读取与网络传输

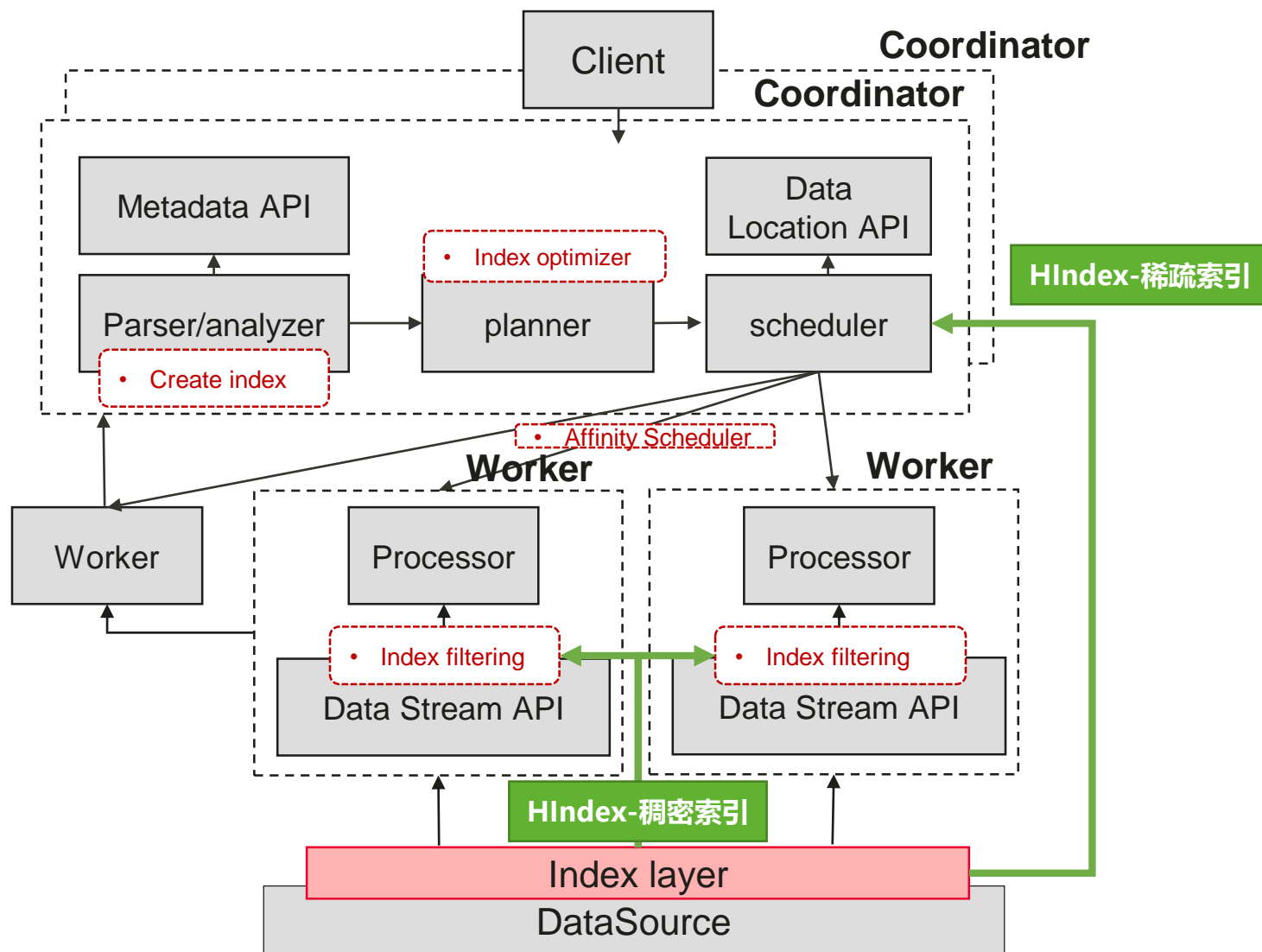


Dynamic Filtering

- 添加DynamicFilterSource算子，搜集build侧数据
- 依赖分布式缓存进行DF的处理
- 适用于inner join & right join
- **适用于join选择率较高的场景**

```
set session enable_dynamic_filtering = true;
set session join_distribution_type = 'BROADCAST';
SELECT count(*) FROM hive.tpcds.catalog_sales t2
join hive.tpcds.date_dim as t1 on t2.
cs_sold_date_sk= t1. d_date_sk AND t1.d_year =
2001;
```

Heuristic index

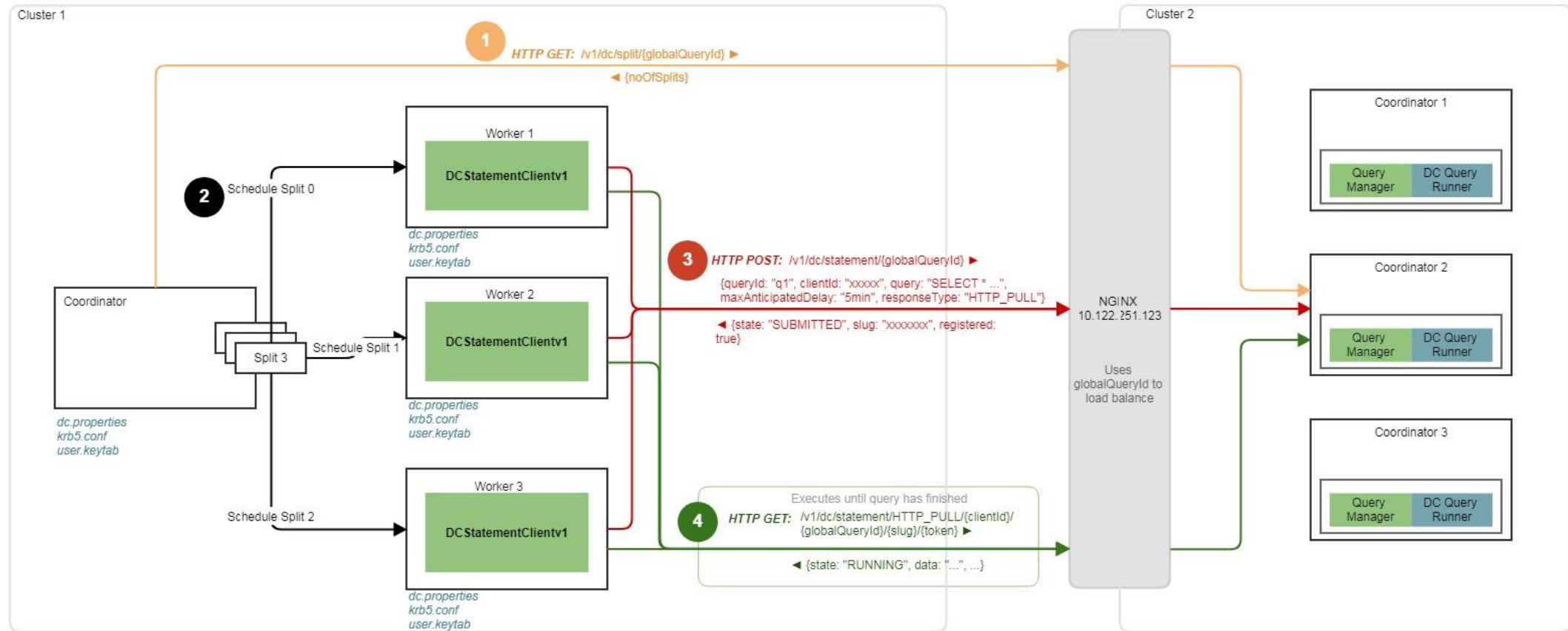


Heuristic index

- 提供统一的索引框架
- 支持多种索引结构
 - 稀疏索引: Bloomfilter、Min-Max
 - 稠密索引: Bitmap
- 任务调度阶段:
 - 裁剪Split,减少调度到Worker的任务数
 - 支持基于索引的亲和性调度
- 数据读取阶段:
 - 减少加载到计算侧内存的数据量

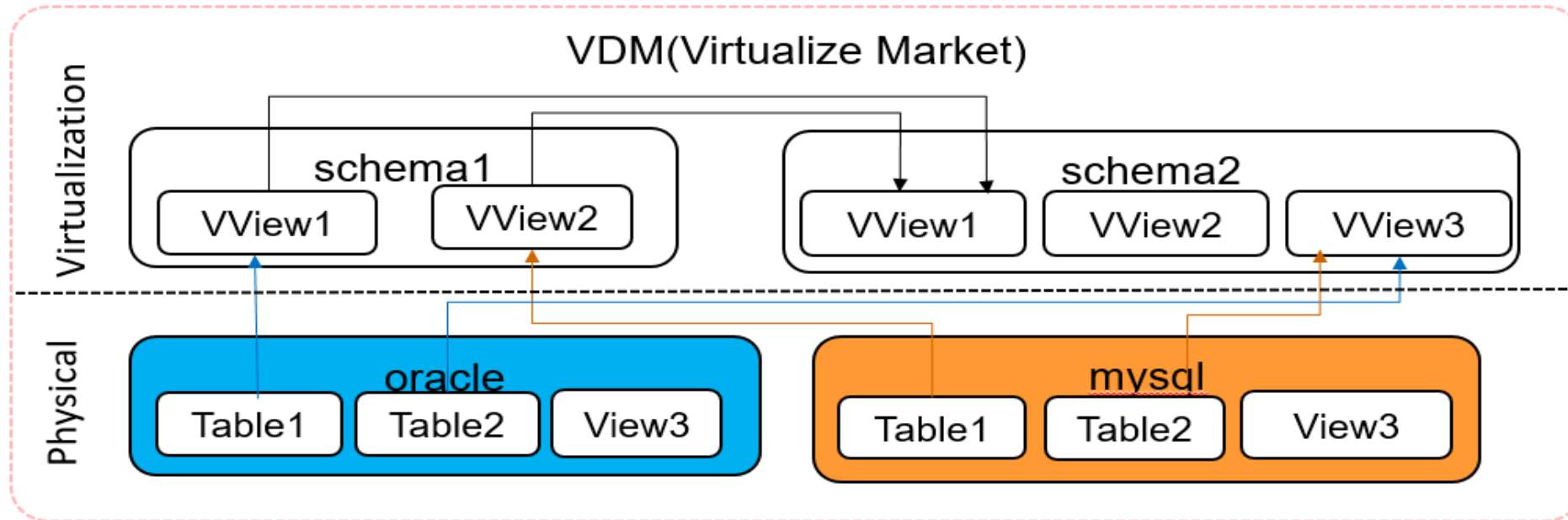
```
select count(*) from store_sales  
where ss_item_sk=123;
```


DC Connector



```
SELECT count(*) FROM dc1.tpcds.sf1.call_center as t1, hive.tpcds.catalog_sales t2 WHERE t1. cc_call_center_sk = t2. cs_call_center_sk ;
```

VDM



```
create schema vdm1.schema1;
CREATE VIEW vdm1.schema1.view1 AS (SELECT * FROM tpcds.sf1.call_center);
CREATE VIEW vdm1.schema1.view2 AS (SELECT * FROM dc1.tpcds.sf1.call_center as t1,
hive.tpcds.catalog_sales t2 WHERE t1. cc_call_center_sk = t2. cs_call_center_sk);
SELECT count(*) FROM vdm1.schema1.view1 as v1, vdm1.schema1.view2 as v2 WHERE
v1.cc_call_center_id = v2.cc_call_center_id;
```

| openLookEng Tryme 小结

- 免安装部署，直接上手试用
- 易用的Sql Editor，智能补全
- 多用户集群资源弹性伸缩
- 关联Gitee、Github账户登录
- 可直接体验动态过滤、启发式索引、Cache、跨DC等高价特性

目录

01

openLookKeng Tryme介绍

02

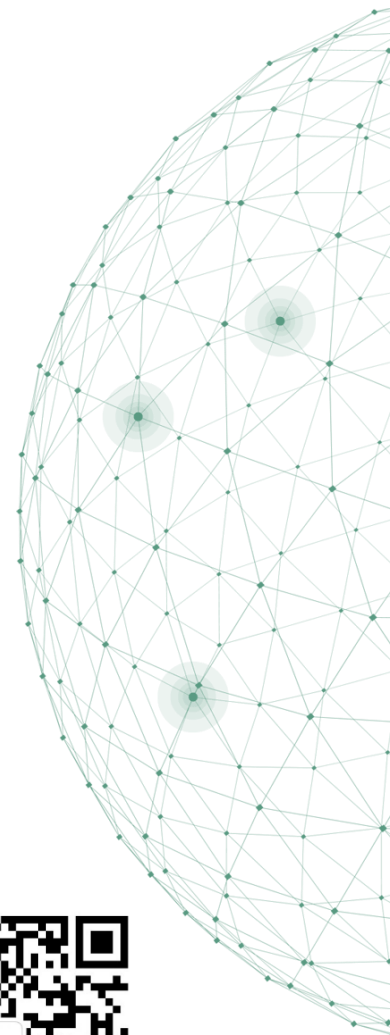
openLookKeng关键技术介绍与演示

03

openLookKeng在IDEA中编译与运行



openLookKeng微信公众号 openLookKeng微信小助手



在Linux和Windows下的IDEA上编译

- Maven仓库

```
<mirror>
  <id>huaweicloud</id>
  <mirrorOf>*</mirrorOf>
  <url>http://mirrors.huaweicloud.com/repository/maven</url>
</mirror>
```

- Linux: mvn clean package -DskipTests

- Windows:

- 修改hetu-maven-plugin, 将"/"修改为File.separatorChar

```
String className = classPath.substring(0, classPath.length() - 6).replace(File.separatorChar, '.');
```

- 设置IDEA的类unix命令行
 - 注释掉presto-root的pom文件中的presto-docs文件
- ```
<!--module>presto-docs</module-->
```

## 在Linux下的IDEA上运行

- Presto comes with sample configuration that should work out-of-the-box for development. Use the following options to create a run configuration:
- Main Class: `io.prestosql.server.PrestoServer`
- VM Options: `-ea -XX:+UseG1GC -XX:G1HeapRegionSize=32M -XX:+UseGCOverheadLimit -XX:+ExplicitGCInvokesConcurrent -Xmx2G -Dconfig=etc/config.properties -Dlog.levels-file=etc/log.properties`
- Working directory: `$MODULE_DIR$`
- Use classpath of module: `presto-main`

# | 在Windows下的IDEA上运行

- 修改config.properties的plugin绑定路径
- 修改PrestoSystemRequirements.java类
- 修改catalog，删除hive.properties
- 配置启动程序入口

## | 小作业（二选一）

1. 在openLookEng Tryme环境上，体验Cache、动态过滤、启发式索引、DC Connector以及VDM特性，观察特性生效后的效果并截图
2. 在Windows/Linux/Mac环境的Intellij Idea上编译并运行openLookEng，使用CLI连接到服务端，运行任一查询并截图

请将截图发给小助手，领取小礼品哈！



openLookEng微信公众号 openLookEng微信小助手



# openLookKeng, Big Data Simplified



openLookKeng微信公众号



openLookKeng微信小助手