

Data Warehouse with ES in Datastory



吴文杰
@Lhfcws_宸风

目录

Contents

- 1 背景
- 2 Gaia
- 3 Cube
- 4 Q & A

Background

Extended SQL on ES

OLAP on ES

Final answer: 42



- 商业智能分析：
 - 数字营销
 - 数据分析洞察
 - 消费者连接
 -



新闻网站



论坛/垂直媒体



评论/跟帖



微博



博客



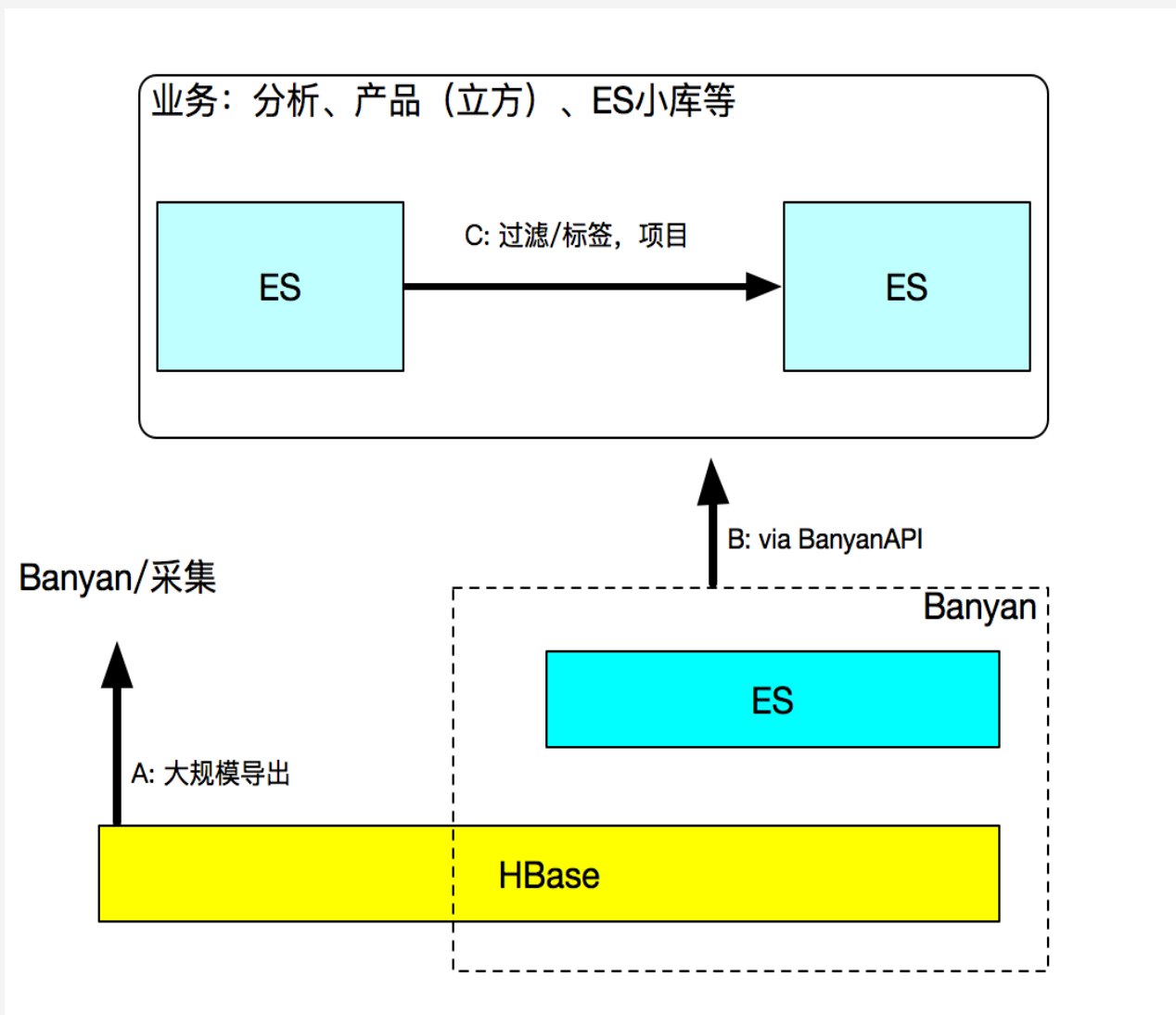
贴吧



A: Scan HBase

B: Query ES & Get HBase

C: Scroll ES & Write ES





- 离线取数与基础分析能力的构建 (Gaia)
- OLAP on ES 能力的构建 (Cube , 数说立方)





- Banyan = ES索引 + HBase存储
- 构建 Hive on HBase/ES/ Banyan的能力
- 对不同存储的查询条件优化
- 提供ES特有的查询支持
- 提供拓展性的SQL表达能力



- StorageHandler 是Hive对接外部存储的核心类（对应SparkSQL的Relation类）
- 主要功能：
 - InputFormat / OutputFormat （如何读写）
 - MetaHook （如何读写Hive元数据）
 - Predicate Pushdown （下推优化、分解条件）



- InputSplit :
 - 按分片分隔，利用preference指定shard访问做到并行读取
- RecordReader:
 - 内部先scroll一批数据
 - 调next到当前批数据为空时，scroll新一批数据 （注意scroll timeout）
- BanyanEsStorageHandler -> BanyanEsInputFormat
 - scroll完ES后增加个读Hbase重新填充SearchHit



- **Sargable** = Search ARGument ABLE , 即SQL中可利用数据库自身索引优势对查询条件进行执行性能优化。
- **谓语句下推**：在实际数据读取和SQL实际执行之前预先执行条件语句进行预处理和过滤。
- **【案例】** SQL中的WHERE条件 , ORDER BY , GROUP BY, HAVING 等有时候可Sargable。
 - Sargable operators: =, >, <, >=, <=, BETWEEN, LIKE, IS [NOT] NULL
 - Sargable operators that rarely improve performance: <>, IN, OR, NOT IN, NOT LIKE
- <https://www.elastic.co/guide/en/elasticsearch/hadoop/master/spark.html#spark-pushdown>



- 对应 SparkSQL 的 SparkPlanner.pruneFilterProject

① 基于Hive的结点树ExprNodeDesc构建通用的结点树

【optional】

② 自顶向下查询可优化的操作符并进行优化

③ 若遇到不可优化的操作符

- 如果是逻辑连接符是AND 则跳过当前节点 并继续优化sibling节点
- 如果是OR则放弃优化

④ 将结点树转为存储可支持的格式

```
public static class DecomposedPredicate {  
    // 可下推部分的结点树  
    public ExprNodeGenericFuncDesc pushedPredicate;  
    // 下推后的序列化对象, 可在InputFormat被拿到  
    public Serializable pushedPredicateObject;  
    // 不可下推要交给Hive后面处理的结点树  
    public ExprNodeGenericFuncDesc residualPredicate;  
}
```



Gaia – es系UDF函数



elastic
中文社区

IT大咖说
知识共享平台

- es_match
- es_matchphrase
- es_term
- has_parent / has_child
- Agg支持
-
- UDF实现可能只是 return true , 只用来做下推查询。

```
1 SELECT * FROM t WHERE es_term(name, "ak"
2
3 {
4   "query": {
5     "bool": {
6       "filter": [
7         {
8           "terms": {
9             "name": [
10              "abel",
11              "simon"
12            ]
13          }
14        },
15        {
16          "range": {
17            "id": {
18              "gt": "1"
19            }
20          }
21        }
22      ]
23    }
24  }
25 }
```



- SemanticAnalyzerHook 拦截 ASTNode语法树
- 读取ES mapping , 重写 `CREATE` 语句

```
CREATE EXTERNAL TABLE es_tbl1 (id string)
TBLPROPERTIES (
    "es.nodes"="172.18.5.142:9200",
    "es.resource"="dt-rhino-weibo-mig-v6.2/weibo"
);
```

```
CREATE EXTERNAL TABLE es_tbl1 (id string) TBLPROPERTIES(
    "es.nodes"="172.18.5.142:9200",
    "es.resource"="dt-rhino-weibo-mig-v6.2/weibo"
);
```

```
SHOW CREATE TABLE es_tbl1;
```

Query executed successfully. Affected rows : -1



createtab_stmt

```
CREATE EXTERNAL TABLE `es_tbl1` (
  `uid` string COMMENT 'from deserializer',
  `attitudes_count` bigint COMMENT 'from deserializer',
  `self_content_length` bigint COMMENT 'from deserializer',
  `msg_type` smallint COMMENT 'from deserializer',
  `repost_comment_count` bigint COMMENT 'from deserializer',
  `emoji` array<string> COMMENT 'from deserializer',
  `id` string COMMENT 'from deserializer',
  `fingerprint` string COMMENT 'from deserializer'.
```

Took 1 sec. Last updated by admin at November 25 2017, 10:41:03 AM.



```
SELECT id
FROM ds_banyan_newsforum_content
WHERE codetable ("kw_include(content, #{关键词}) AND id > '1' ",
"/tmp/codetable_test.csv")
LIMIT 10;
```

```
SELECT id
FROM ds_banyan_newsforum_content
WHERE codetable("kw_include(content, #{关键词}) AND id > '1' ",
"/tmp/codetable_test.csv")
LIMIT 10;
```

Query executed successfully. Affected rows : -1



Line Chart	
id	
2146c95cd8158f5e2084a4ea68d1c24d	
1e4fdb0b2ad484686db276ffeb366665	
68b6e4db505944f34988e65dc31ca8aa	
e1cac5cb29a1636b3e7f80256ddc2e23	
5d958e417837035d9ab1d997cbaa7bc8	
bd316281b75a29f6d7aa4112531fab7f	
267ce1b5ddb277a078d383fdb27cde23	
bde006418e8f2e08ac0abac3411a5b25	
a9bb0b96aae115141c4186f21fd889a	



```
SET      sql1 = kw_include(title, #{关键词});

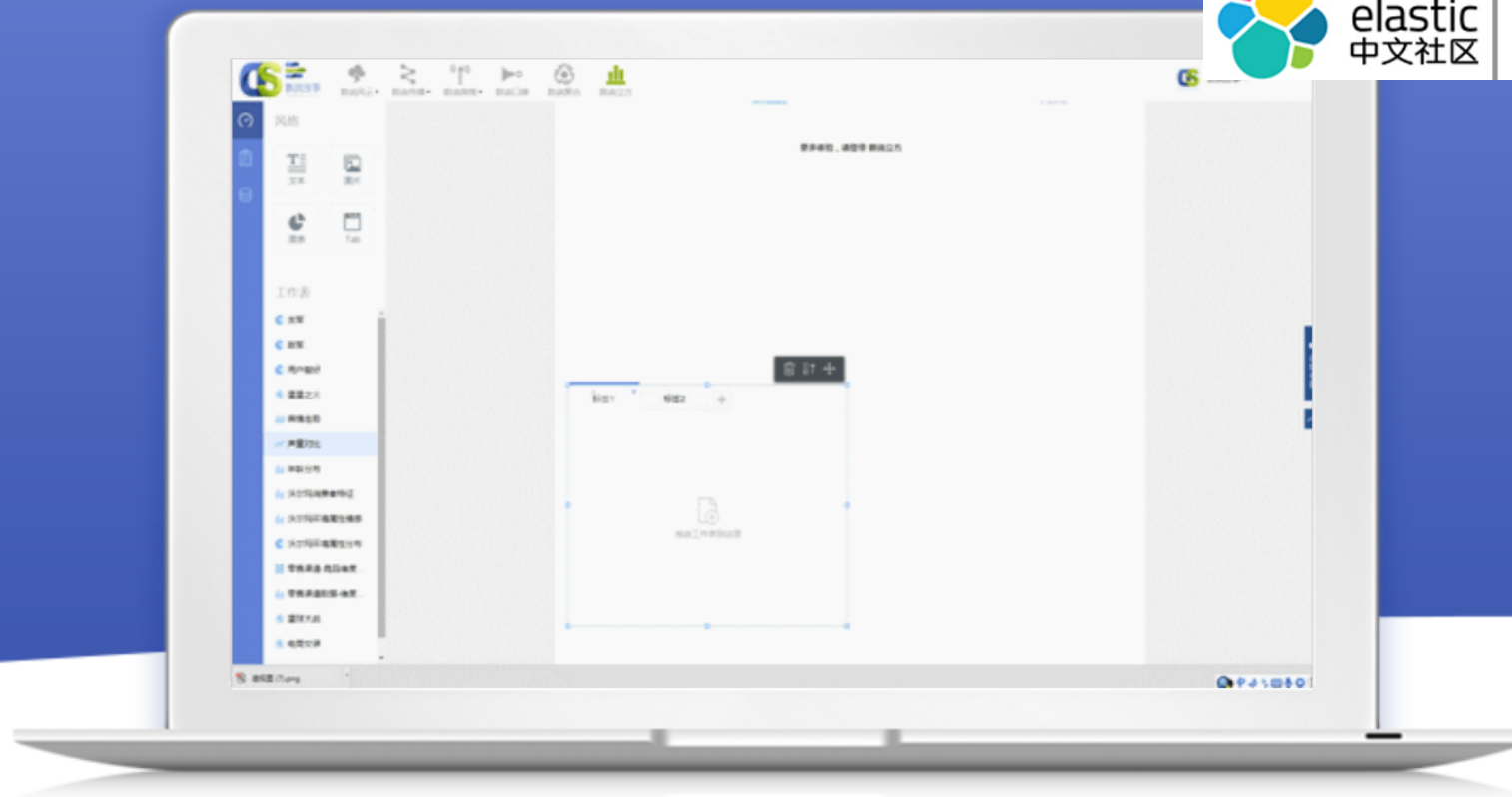
SELECT ct_script("if ( translate_groovy('${hiveconf:sql1}') ) _return = #{字段长度}",
                "/tmp/codetable_test.csv", struct("title", title)) as field_len
FROM    ds_banyan_newsforum_content
WHERE   codetable("${hiveconf:sql1}", "/tmp/codetable_test.csv", "OR")
LIMIT   5;
```

field_len	
4	2
4	2
4	
4	
4	



- 【业务方】
 - 减少写代码的开销和出错率
 - 更友好的筛选取数工具
 - 为后续的数据分析提供基础（Zeppelin等）
- 【平台方】
 - 统一管控计算资源（ES task ✖）

分析从此开始 观点从未停止



垂直搜索

分布式搜索技术，
让你轻松拥有社交数据



BI报表

海量数据的实时在线分析、
丰富的可视化效果



仪表盘

专业数据可视化呈现，
非专业人员也可轻松搭建

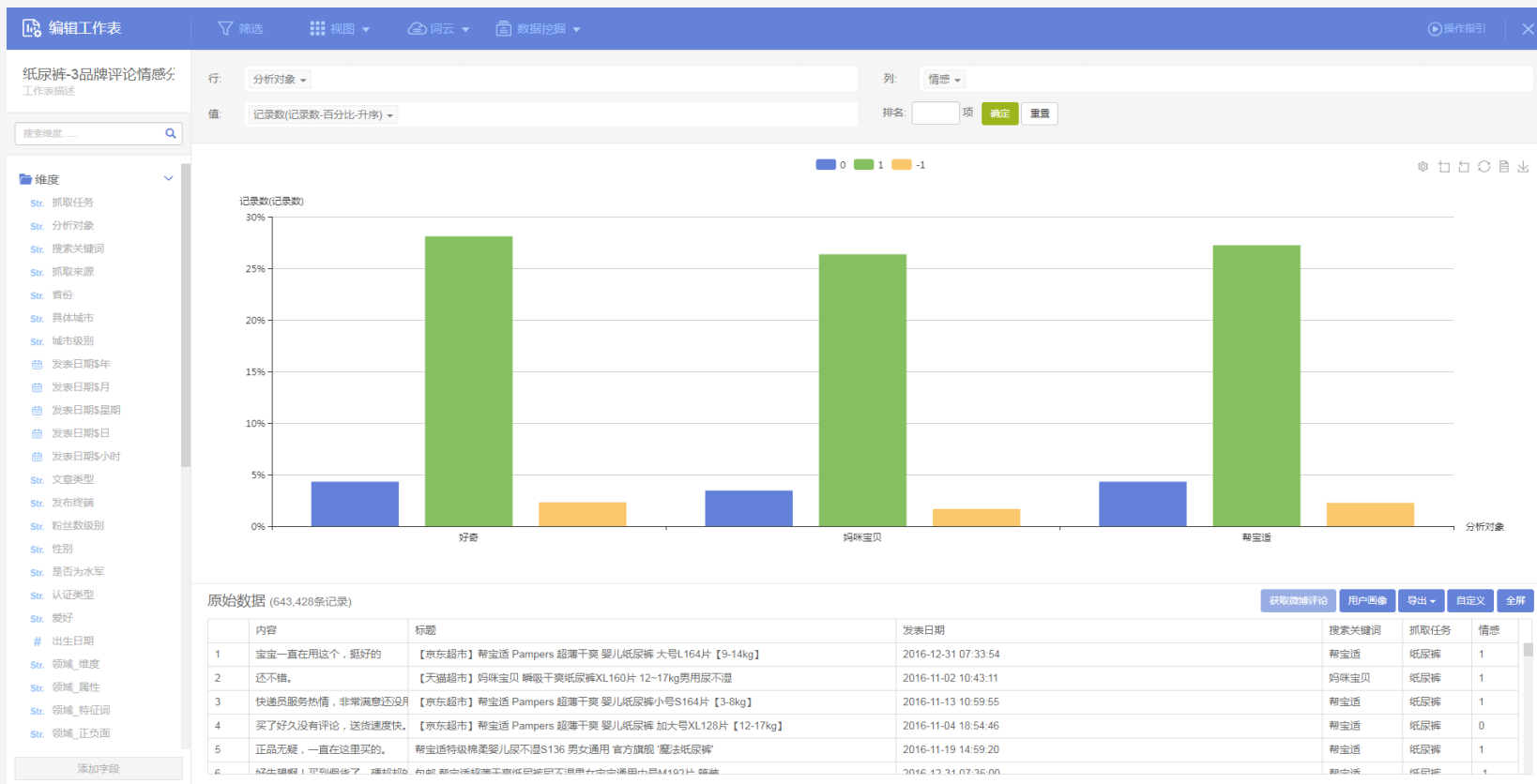


机器学习

集数据处理、特征工程、建模、
文本挖掘为一体的机器学习平台



- 数说立方(Cube) 是数说故事自研的基于ES的OLAP产品，可提供非技术人员自由的导入数据、维度透视、统计分析等功能。





- 即席查询
 - 实时
 - 灵活度高
- 占用空间小
 - ES索引代替维度表
- 全文检索支持
 - Lucene支持



- es-hadoop
 - 导入、刷库
- Schema 自动识别
- ES Agg 进行统计聚合
 - 维度选择
- Cardinality



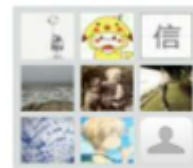
- 建议使用latest stable的 es-hadoop （旧版还是有些隐性的bug，新版代码更清晰）
- 注意一些特殊字段(suggest, array<string>, nested等) 可能会有坑或不兼容等
- es-hadoop支持跨版本ES的读写
- es-rest 提供客户端统一接口读写不同版本ES的能力



- ES 的数组字段不易识别 (var $\rightarrow$ array<var>)
 - 抽样数据进行schema调整

Cube – ES相关技术点 – Agg & Cardinality

- 什么样的数据字段可以成为维度？（日期？数值？字符串？）
 - Cardinality查询，设定阈值过滤
- HyperLogLog++: precision_threshold = 10000



Elasticsearch交流群



该二维码7天内(11月30日前)有效, 重新进入
将更新