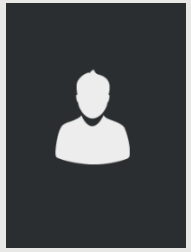
Ya100 技术详解

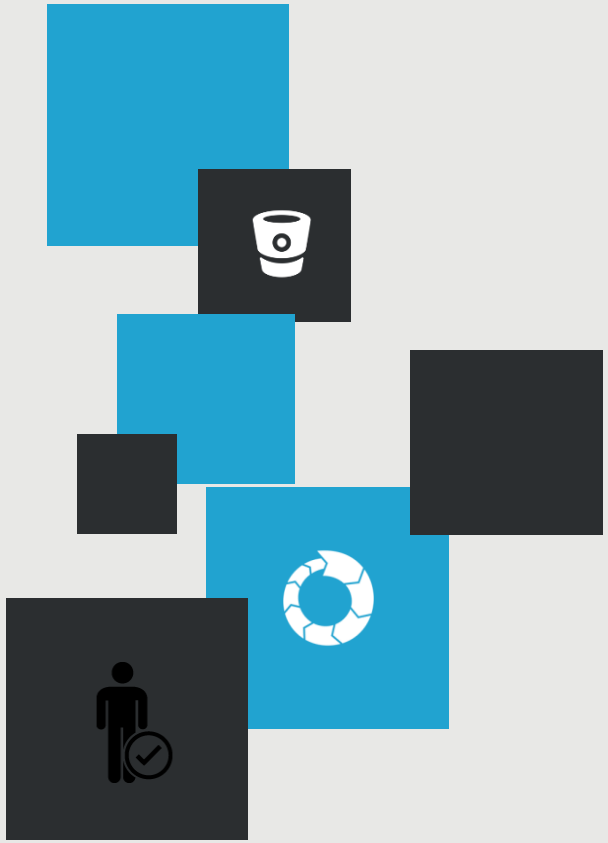
延云

一个比Spark-Parquet还快5~100倍的存储格式

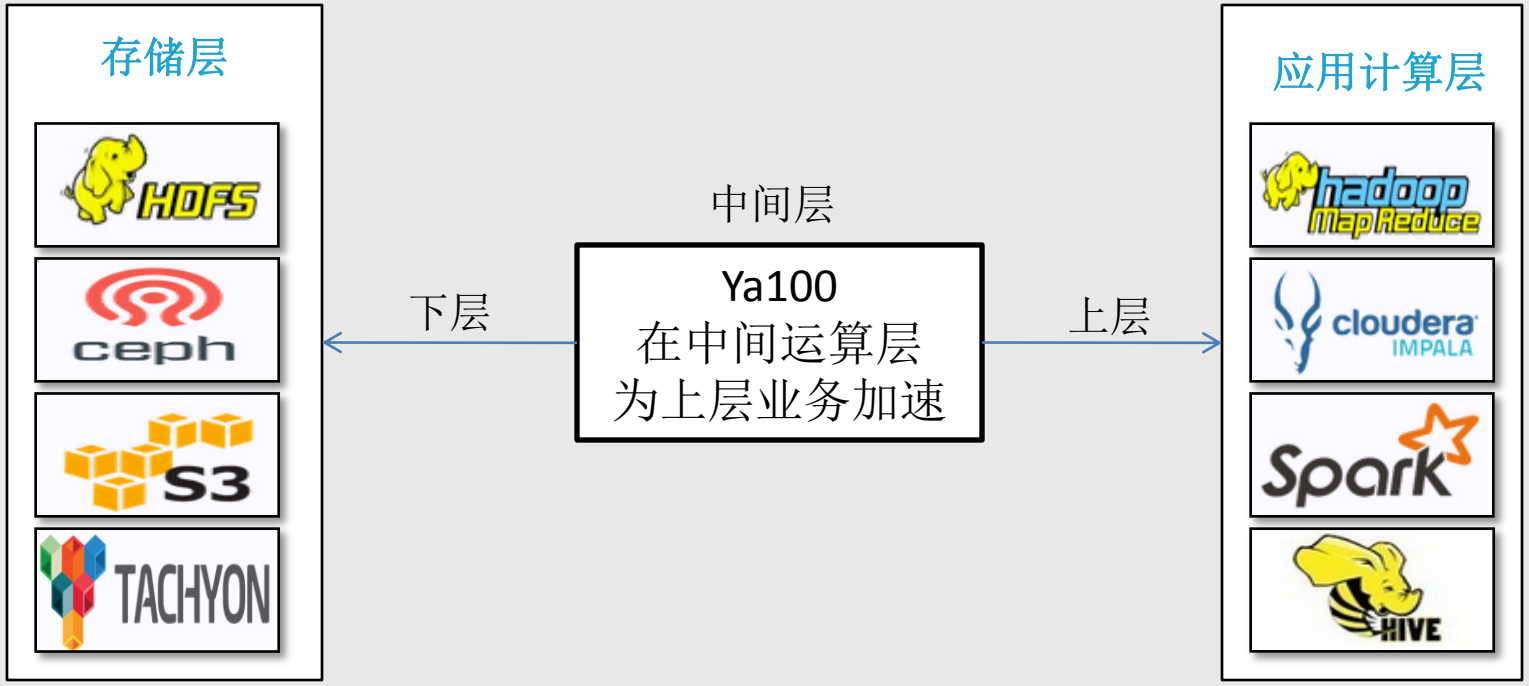


母延年
微信:ycloudnet
QQ群:171465049

YDB100大数据加速器介绍



- Spark SQL的一种新的存储格式。
- Ya100比Parquet格式快5~100倍.
- 任意维度组合,过滤，万亿数据秒级响应。



通过索引-不需要的行我们不读取

SQL场景：交易日志的查询

```
select
      （1）phonenum,usernick,ydb_grade,ydb_age,ydb_blood,ydb_zhiye,ydb_earn,ydb_day,amtlong
from spark_txt where
      （2）tradeid=' 20140122138702826 71'
limit 10;
```

这个SQL由两部分组成

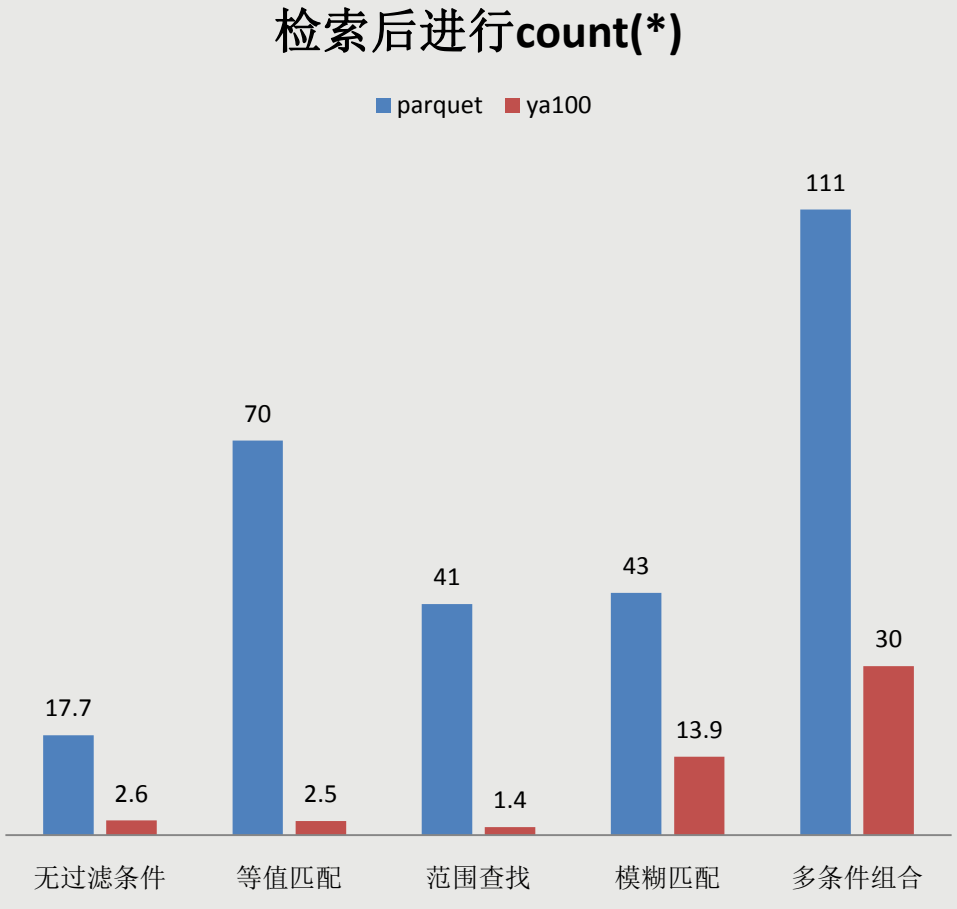
（1）红色部分是要返回的结果，有9个列。

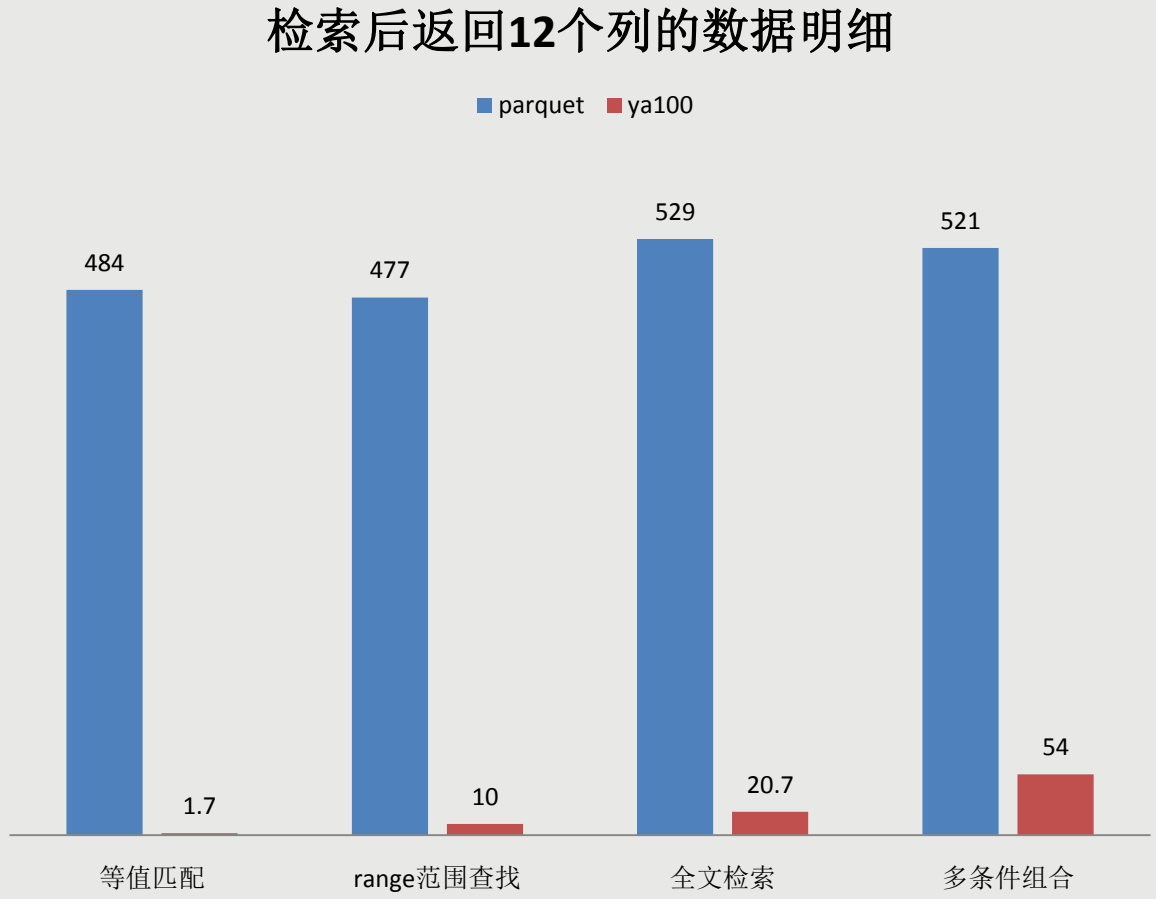
（2）蓝色部分是要进行筛选过滤的条件-交易号。

大家想一想，下面哪种计算方式快？

- 方案一：先将红色列的每行数据读过来，再按照蓝色部分的交易号筛选。
- 方案二：先按照蓝色部分的交易号筛选，在将红色列的满足条件的行读过来。

使用ya100的索引-与parquet的性能对比





非排序的列最后延迟读取

SQL场景：按照日志时间排序

```
select
      （1）phonenum,usernick,ydb_grade,ydb_age,ydb_blood,ydb_zhiye,ydb_earn,ydb_day,amtlong
from spark_txt
      （2）order by logtime desc
limit 10;
```

这个SQL由两部分组成

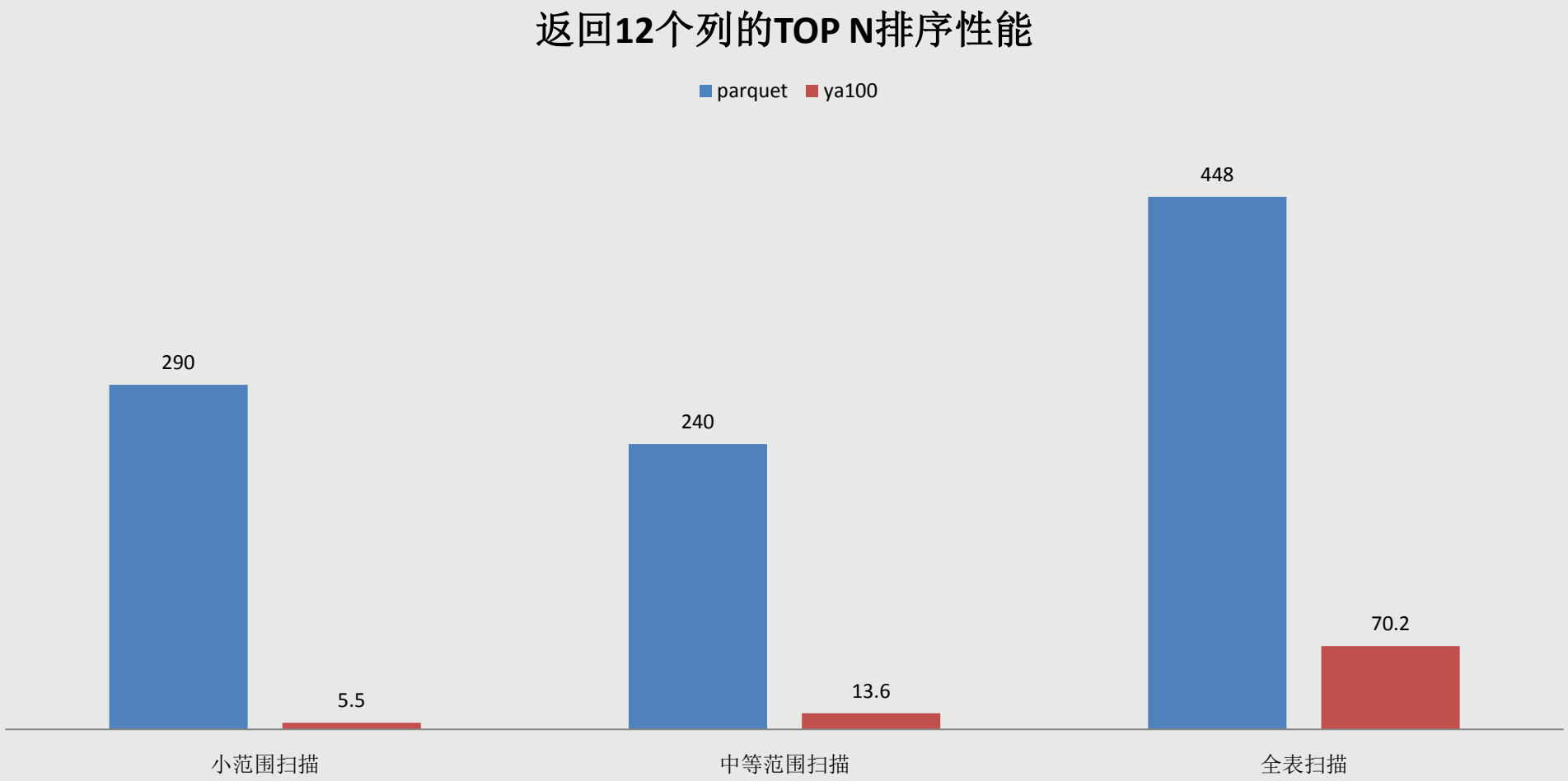
（1）红色部分是要返回的结果，有9个列。

（2）蓝色部分是要进行排序取TOP的列。

大家想一想，下面哪种计算方式快？

- 方案一：先将**红色列的每行数据**读过来后，再按照**蓝色部分的时间**排序。
- 方案二：先按照**蓝色部分的时间**的排序，在排序后剩余的满足条件的10行**红色列的数据**读过来。

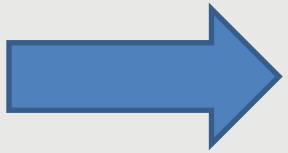
延迟读取方式-与parquet在TOP N排序上的性能对比



采用标签代替原始值-进行分组与排序

原始值
冲锋衣裤
速干衣裤
羽绒服/棉服
休闲衣裤
抓绒衣裤
软壳衣裤
户外风衣
功能内衣
军迷服饰
越野跑鞋
沙滩/凉拖
滑雪服

给原始值排重后
由小到大加标签



标签值	字典值
0	冲锋衣裤
1	速干衣裤
2	羽绒服/棉服
3	休闲衣裤
4	抓绒衣裤
5	软壳衣裤
6	户外风衣
7	功能内衣
8	军迷服饰
9	越野跑鞋
10	沙滩/凉拖
11	滑雪服

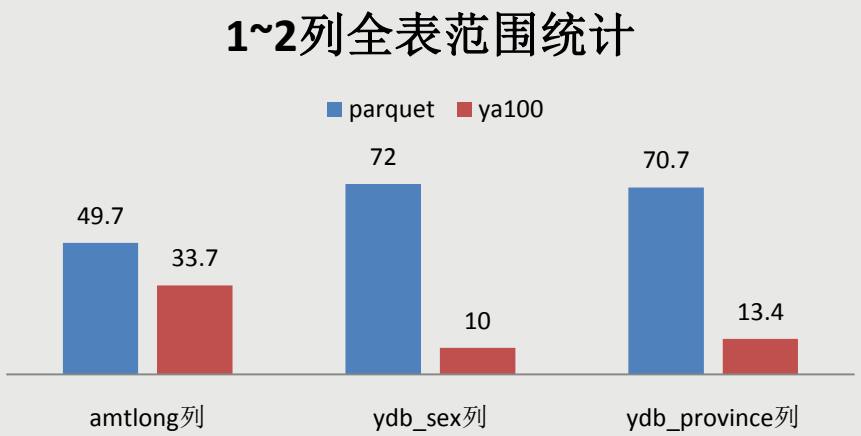
优点

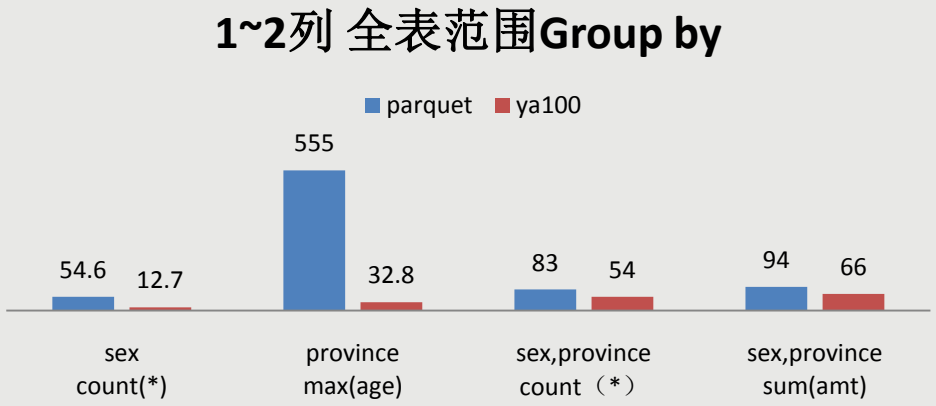
1. 真实数据每个列的值多少有重复值，原始值仅存储一份，可以减少存储空间占用。
2. 标签值采用定长存储，可以随机读取。
3. Group by分组计算的时候，使用标签代替原始值，数值型计算速度比字符串的计算速度快很多。
4. 标签值的大小原始值的大小是对应的，故排序的时候也仅读取标签进行排序。
5. 标签比原始值占的内存少。

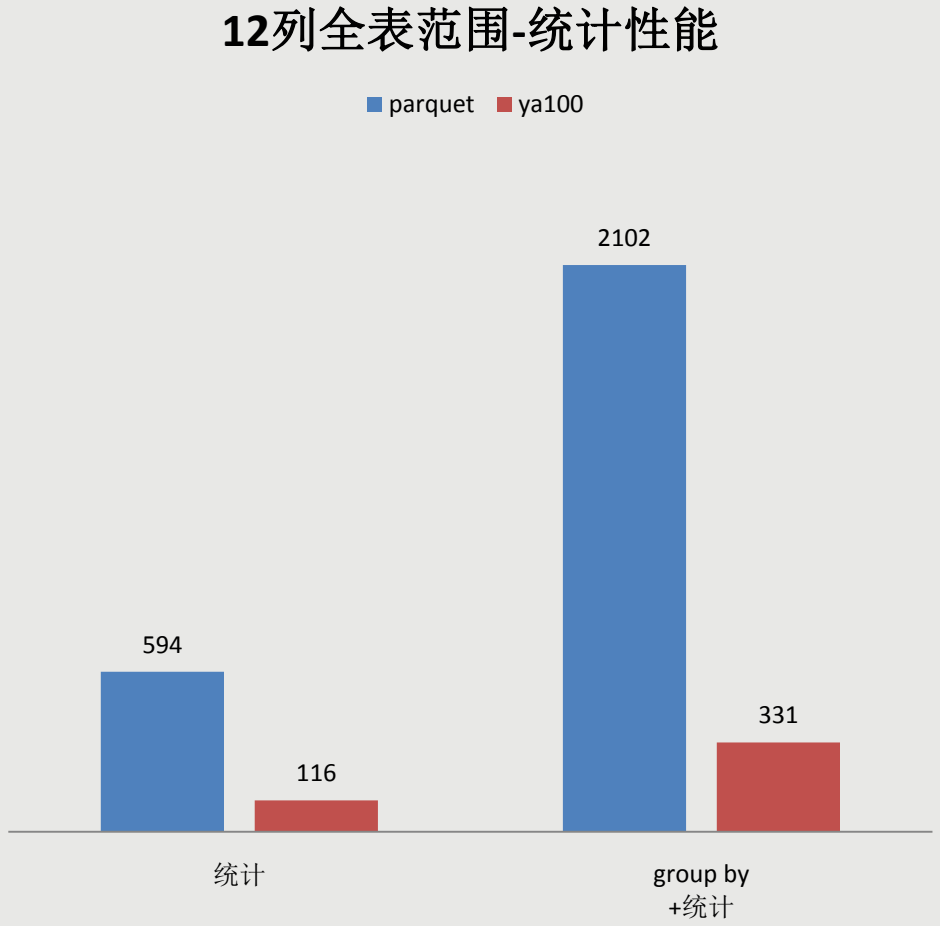
缺点

1. 如果数据重复值很低，存储空间相反比原始数据大。
2. 如果重复值很低，且查询逻辑需要大量的根据标签值获取原始值的操作的时候，性能比原始值慢。

使用标签标签技术的ya100 与parquet的在group by性能上的对比







多种优化合体后更猛-性能提升数百倍

万亿数据，秒级响应

Ya100的价格-最低每年4200元

周期	价格（人民币）	公测结束前付费6折	折后价格
一年	10000	2016年6月15日前为公测期，公测期完全免费。在公测期结束前付费可在上述价格的基础上 6 折。	6000
两年	18000 （平均每年9000）		10800 （平均每年5400）
三年	24000 （平均每年8000）		14400（平均每年4800）
五年及以上	N*7000（平均每年7000）		N*4200（平均每年4200）

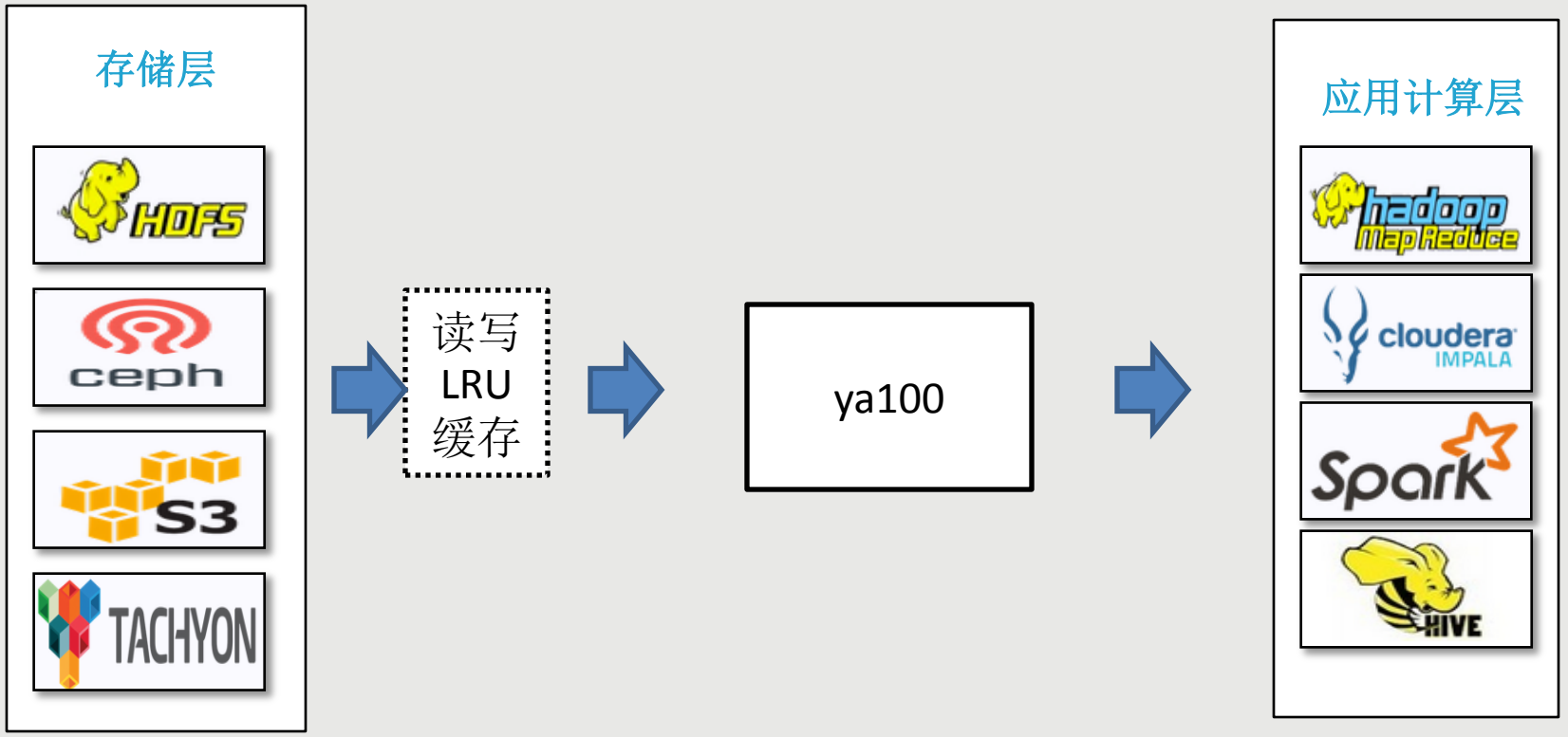
注：以上价格 均不会对数据量大小、集群规模 、查询并发等有任何限制，大家随便玩。

Ya100定价每年4200元是一个什么概念？

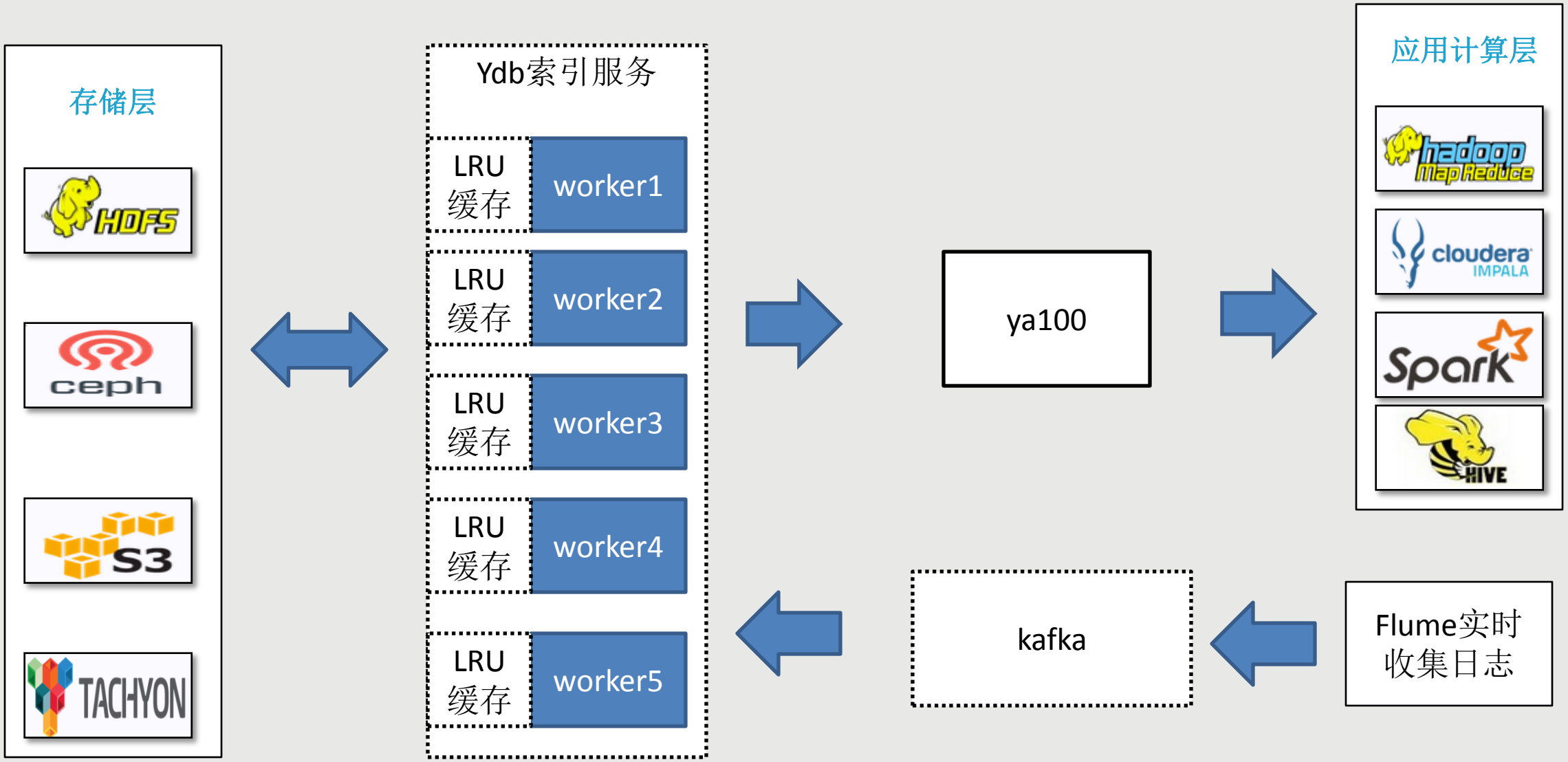
需要bat年薪100万技能水平的的大数据研发工程师4~5个，研发3年，研发过程中需采购数十台服务器进行测试。

研发成功率我们先不说，仅工资与服务器成本与4200元每年的定价相比ya100就是白菜价。

YA100普通运行模式（离线模式）



YA100与Ydb运行模式-数据实时导入-1分钟内可见



YA100的安装方法

一、依赖的jar包的配置

export SPARK_CLASSPATH=$SPARK_CLASSPATH:/data/ycloud/spark_ydb/ya100.jar

二、spark启动示例

./start-thriftserver.sh --master yarn-client --executor-memory 990m --executor-cores 2 --num-executors 16

建议的spark版本为 spark1.6

三、配置内置函数

```
create  function Yfilter as 'cn.net.ycloud.ydb.handle.fun.Yfilter';
create  function Ytop10000 as 'cn.net.ycloud.ydb.handle.fun.Ytop10000';
create  function Ycombine as 'cn.net.ycloud.ydb.handle.fun.Ycombine';
create  function Ycount as 'cn.net.ycloud.ydb.handle.fun.Ycount';
create  function Ymax as 'cn.net.ycloud.ydb.handle.fun.Ymax';
create  function Ymin as 'cn.net.ycloud.ydb.handle.fun.Ymin';
create  function Yavg as 'cn.net.ycloud.ydb.handle.fun.Yavg';
create  function Ysum as 'cn.net.ycloud.ydb.handle.fun.Ysum';
create  function Ymaxstring as 'cn.net.ycloud.ydb.handle.fun.Ymaxstring';
create  function Yminstring as 'cn.net.ycloud.ydb.handle.fun.Yminstring‘;
```

YA100表的数据类型

类型	解释
string	字符串类型，该类型不分词，通常用来存储比较短的字符串，如类目
int	整形32位
long	整形64位
double	Double类型
float	Float类型
text	字符串类型，单字分词，通常用来存储长文本
textcjk	字符串类型，CJK分词，通常用来存储长文本
tint	整形32位-适合大范围的range过滤查询
tlong	整形64位-适合大范围的range过滤查询
tdouble	Double类型-适合大范围的range过滤查询
tfloat	Float类型-适合大范围的range过滤查询

YA100表的创建

```
CREATE external  table spark_ya100(
 phonenum bigint, usernick string, ydb_sex string,  ydb_province string, ydb_grade string, ydb_age
string, ydb_blood string, ydb_zhiye string, ydb_earn string, ydb_prefer string, ydb_consume string, ydb_day
string, amtdouble double,amtlong int, content string, ydbpartion string, ya100_pipe string
) partitioned by (dt string)
STORED BY 'cn.net.ycloud.ydb.handle.Ya100StorageHandler'
LOCATION  '/data/ydb/shu_ya100'
TBLPROPERTIES(
"ya100.handler.table.name"="ydb_example_shu",
"ya100.handler.schema"="phonenum long,usernick string,ydb_sex string,ydb_province string,ydb_grade
string,ydb_age string,ydb_blood string,ydb_zhiye string,ydb_earn string,ydb_prefer string,ydb_consume
string,ydb_day string, amtdouble double,amtlong int,content textcjk,ydbpartion string,ya100_pipe string"
)
```

*//*注，ydbpartion为ydb的分区字段与ya100_pipe 为ya100的管道字段，必须要都要创建，否则有些功能会被限制使用。

ya100.handler.table.name为预留的key,不同的表之间要区分开来

ya100.handler.schema 为内置的索引的数据类型，名字要与spark的字段名称一致。

YA100表的数据导入

一、设置导入参数

```
set hive.mapred.supports.subdirectories=true;
set mapred.min.split.size=2147483648;
```

注：

ya100因为生成的最终不是一个文件，而是一个目录，这个目录下会有索引，所以要通过subdirectories参数设置spark支持子目录，否则向表里导入数据会报错失败。

set mapred.min.split.size=2147483648 是为了控制map的个数，防止生成的索引文件数量太多。

二、开始导入数据

```
insert into table spark_ya100 partition (dt='1200million')
 select
phonenum,usernick,ydb_sex,ydb_province,ydb_grade,ydb_age,ydb_blood,ydb_zhiye,ydb_earn,ydb_prefer,ydb_consume,ydb_
day,amtdouble,amtlong,content,2,1
from spark_txt where dt='1200million';
```

YA100表的预留的ya100_pipe管道列的作用

通过ya100_pipe列，可以把命令送下去，用原生ya100来计算，最后结果返回来。

我们在创建spark表的时候会预留一个列ya100_pipe，该列为系统预留的管道列，通过该管道，用户可以通过SQL，直接穿透Spark、hive的计算层，直接操作原生的ya100，减少输出给Spark或hive的数据规模大小，从实现更高的性能。

1. 在ya100里过滤掉后的数据，不会在抛出给spark，也节省了数据传输时间。
2. 针对有排序的场景，可以设置按某种方式排序，只返回TOP N条记录，可以提升性能。
3. 针对有数据聚合的场景，直接在ya100里聚合，提升性能。

利用索引，不需要的行直接在ya100里过滤掉

一、基本写法

```
set ya100.spark.filter.ydb_example_shu=ydb_sex='女' or ydb_province='辽宁' or ydb_day>='20151217';
```

有些时候，会遇到转义的问题，这里的value值，可以进行urlencode处理。

如：

```
set ya100.spark.filter.ydb_example_shu=ydbpartion%3D%2720151110%27+and+%28ydb_sex%3D%27%E5%A5%B3%27+or+ydb_province%3D%27%E8%BE%BD%E5%AE%81%27+or+ydb_day%3E%3D%27201512 17%27%29;
```

二、完整SQL使用示例

```
set ya100.spark.filter.ydb_example_shu=ydb_sex='女' or ydb_province='辽宁' or ydb_day>='20151217';
select * from spark_ya100 where dt= ‘1200million’ limit 10;
```

支持的过滤条件以及写法-普通过滤

一、等值匹配：

如 qq=‘165162897’

二、支持 in操作，

如：indexnum in (1,2,3)

三、>,<,>=,<=,区间查询的写法

clickcount >=10 and clickcount <=11

四、对于带有范围的过滤筛选，使用下面的方式能提升查询效率

indexnum like '({0 TO 11}) '　　不包含边界值

indexnum like '([10 TO 11]) '　　包含边界值

五、不等于的写法

label<>'l_14' and label<>'l_15'

六、过滤条件可以进行and与or的组合

indexnum='1' or indexnum='2' or (clickcount >=5 and clickcount <=11)

支持的过滤条件以及写法-全文检索

一、文件检索

text或testcjk类型可以使用全文检索功能
text或testcjk类型因分词的原因，不能再进行group by与order by、与统计了。

以“中华人民共和国”为例
如果是text类型，采用单字分词，分词结果为
中-华-人-民-共-和-国
content='中' content=' 中华' content='中人' 都可以匹配到这条记录
如果是textcjk类型这个类型采用二元分词 分词结果就是
中华-华人-人民-民共-共和-和国
contentcjk=' 中' 匹配不到结果
contentcjk =' 中华' 可以匹配到结果
contentcjk =' 中人' 匹配不到结果

二、模糊匹配

对于string类型的列，可以使用like功能进行模糊匹配，但是性能特别低，要谨慎使用。
label like 'l*' 注意如果该列排重后的值数量特别多，禁止将通配符写在最前面，否则查询性能很低
label like 'l_??' 如要要进行类似QQ号码的定长匹配，可以用?来代替*，*是匹配0~多个字符，而?则表示只匹配一个字符

直接通过ya100进行TOP N排序-非排序的列最后延迟读取

```
set ya100.spark.top10000.ydb_example_shu=ydb_age desc limit 10;
set ya100.spark.top10000.ydb_example_shu=ydb_sex desc,ydb_province limit 100；
set ya100.spark.top10000.ydb_example_shu=* limit 10;
```

注意，当前版本limit的最大值为10000 （每个索引最大返回1000）
列的名字 写*表示 不需要排序，但只返回前N条就可以了

完整SQL示例

```
set ya100.spark.top10000.ydb_example_shu=phonenum desc limit 10;
select * from spark_ya100 where dt='1200million' order by phonenum desc limit 10;
```

管道操作三：使用Ycombine在ya100层进行高性能的聚合

示例一

```
set ya100.spark.combine.ydb_example_shu=*,ydb_age;
select
ydb_province,Ycount('*',ya100_pipe), Ycount('ydb_age',ya100_pipe), Ymaxstring('ydb_age'
,ya100_pipe), Yminstring('ydb_age',ya100_pipe)
 from spark_ya100  group by ydb_province limit 10
```

示例二

```
set ya100.spark.combine.ydb_example_shu=amtlong;
select ydb_sex, ydb_province,Ysum('amtlong',ya100_pipe) as cnt from spark_ya100  group
by ydb_sex, ydb_province  limit 10
```

更多ya100的SQL使用示例一

- **等值匹配+count(*)**

```
set ya100.spark.filter.ydb_example_shu=phonenum='13870282671'  and usernick='江峻熙';
set ya100.spark.combine.ydb_example_shu=*;
set ya100.spark.top10000.ydb_example_shu=;
select Ycount('*',ya100_pipe) from spark_ya100 where dt='100million' limit 10;
```

- **按照手机号TOP N排序**

```
set ya100.spark.filter.ydb_example_shu= amtlong like '([1090 TO 1100] )' and amtdouble like
'([1090 TO 1100] )';
set ya100.spark.combine.ydb_example_shu=;
set ya100.spark.top10000.ydb_example_shu=phonenum desc limit 10;
select * from spark_ya100 where dt='100million' order by phonenum desc limit 10;
```

- **amtlong的统计**

```
set ya100.spark.filter.ydb_example_shu= amtlong like '([1090 TO 1100] )' and amtdouble like
'([1090 TO 1100] )';
set ya100.spark.combine.ydb_example_shu=amtlong;
set ya100.spark.top10000.ydb_example_shu=;
select Ycount( ‘amtlong’ ,ya100_pipe) ,Ysum( ‘amtlong’ ,ya100_pipe)
,Yavg( ‘amtlong’ ,ya100_pipe), Ymax( ‘amtlong’ ,ya100_pipe), Ymin( ‘amtlong’ ,ya100_pipe)  from
spark_ya100 where dt= ‘100million’  limit 10;
```

更多ya100的SQL使用示例二

•多列TOP N排序

```
set ya100.spark.filter.ydb_example_shu= amtlong like '([1090 TO 1100] )' and amtdouble like '([1090 TO 1100] )';
set ya100.spark.combine.ydb_example_shu=;
set ya100.spark.top10000.ydb_example_shu= ydb_age desc, ydb_sex;
select ydb_sex, ydb_age from spark_ya100 where dt='100million'  order  by  ydb_age desc, ydb_sex  limit 10;
```

•多列group by统计

```
set ya100.spark.filter.ydb_example_shu=;
set
ya100.spark.combine.ydb_example_shu=ydb_age,ydb_blood,ydb_zhiye,ydb_earn,ydb_prefer,ydb_consume,ydb_day,amtlong;
set ya100.spark.top10000.ydb_example_shu=;
select
ydb_sex,ydb_province,ydb_grade, Ymaxstring( ‘ydb_age’ ,ya100_pipe),Ymaxstring( ‘ydb_blood’ ,ya100_pipe),Ymaxstring( ‘ydb_zhiye’ ,ya100_pipe),Ymaxstring( ‘ydb_earn’ ,ya100_pipe),Ymaxstring( ‘ydb_prefer’ ,ya100_pipe),Ymaxstring( ‘ydb_consume’ ,ya100_pipe),Ymaxstring( ‘ydb_day’ ,ya100_pipe),Ysum( ‘amtlong’ ,ya100_pipe) as cnt from spark_ya100 where dt= ‘100million’   group by ydb_sex,ydb_province,ydb_grade order by cnt desc limit 10
```

通过YA100-创建与ydb连接映射表

```
CREATE external  table spark_ydb(
 phonenum bigint, usernick string, ydb_sex string,  ydb_province string, ydb_grade string, ydb_age string, ydb_blood
string, ydb_zhiye string, ydb_earn string, ydb_prefer string, ydb_consume string, ydb_day string, amtdouble double,amtlong
int, content string, ydbpartion string, ya100_pipe string
)
STORED BY 'cn.net.ycloud.ydb.handle.Ya100StorageHandler'
LOCATION  '/data/ydb/shu_ydb'
TBLPROPERTIES(
"ya100.handler.table.name"="ydb_example_shu",
"ya100.handler.master"="101.200.130.48:8080",
"ya100.handler.columns.mapping"="phonenum,usernick,ydb_sex,ydb_province,ydb_grade,ydb_age,ydb_blood,ydb_zhiye,ydb_earn,ydb_pref
er,ydb_consume,ydb_day,amtdouble,amtlong,content,ydbpartion,ya100_pipe")
```

注

ydbpartion为ydb的分区字段，查询的时候必须执行ydbpartion,而spark表则没有分区字段

ya100_pipe 为ya100的管道字段，必须要都要创建，否则有些功能会被限制使用。

ya100.handler.table.name为ydb系统里的表名字

ya100.handler.columns.mapping 为ydb系统内的表与spark的表之间的映射。

通过Ya100与ydb连接后的查询示例一

该方式除了ydb的分区参数ydbpartion必须传递外与普通ya100没有任何区别
Ydb本身可以实时导入，也意味着ya100可以借助ydb来进行实时数据的导入。

示例一 基本统计使用

```
set ya100.spark.filter.ydb_example_shu=ydbpartion='20150811';
set ya100.spark.combine.ydb_example_shu=*,amtdouble,ydb_province;
set ya100.spark.top10000.ydb_example_shu=;
select Ycount('*',ya100_pipe) ,Ysum('amtdouble', ya100_pipe) ,Ymaxstring('ydb_province', ya100_pipe)
from spark_ydb limit 10;
```

示例二 group by count（*）

```
set ya100.spark.filter.ydb_example_shu= ydbpartion= '20150811'and amtlong like '([1090 TO 1100] )' ;
set ya100.spark.combine.ydb_example_shu=*;
set ya100.spark.top10000.ydb_example_shu=;
select ydb_sex, ydb_province,Ycount('*',ya100_pipe) as cnt from spark_ydb  group by ydb_sex, ydb_province  order by cnt
desc limit 10;
```

通过Ya100与ydb连接后的查询示例二

示例三 group by sum

```
set ya100.spark.filter.ydb_example_shu= ydbpartion='20150811'and amtlong like '([1090 TO 1100] )' ;
set ya100.spark.combine.ydb_example_shu=amtdouble;
set ya100.spark.top10000.ydb_example_shu=;
select ydb_sex, ydb_province,Ycount('amtdouble',ya100_pipe) as cnt,Ysum('amtdouble',ya100_pipe)
from spark_ydb  group by ydb_sex, ydb_province  order by cnt desc limit 10;
```

示例四 top N查询

```
set ya100.spark.filter.ydb_example_shu= ydbpartion= '20150811'and amtlong like '([1090 TO 1100] )' ;
set ya100.spark.combine.ydb_example_shu=;
set ya100.spark.top10000.ydb_example_shu=* limit 10;
select ydb_sex, phonenum,amtlong,amtdouble from spark_ydb  limit 10;
```

示例五 按照amtdouble, amtlong排序 top N查询

```
set ya100.spark.filter.ydb_example_shu= ydbpartion= '20150811'and amtlong like '([1090 TO 1100] )' ;
set ya100.spark.combine.ydb_example_shu=;
set ya100.spark.top10000.ydb_example_shu=amtdouble desc ,amtlong limit 10;
select ydb_sex, phonenum,amtlong,amtdouble from spark_ydb order by amtdouble desc ,amtlong  limit 10;
```

相关资源

- 测试报告以及软件获取
  http://ycloud.net.cn/yyydb
- 延云官网
  http://ycloud.net.cn
- QQ群:171465049

谢谢大家！