client

parser

analyer

planner

scheduler

SELECT \* FROM tb

WHERE tb.id=1009

file1

file2

filen

.....

split

split

split

.....

spilt

split

.....

spilt

2000+

***INDEX***

col `id`

split

split

split

.....

split

<100

result

**DATA**

***Bloom Index***

SELECT \* FROM tb

WHERE tb.name='Marry'

stripe

column1

column2

column3

column4

batch1

batch2

......

batchn

Orc RecordRead

Stream Reader

Stream Reader

Stream Reader

Stream Reader

Stream Reader

Stream Reader

orc pageSource

get next batch

Page

get next page

Marry:

Jacker:

0

0

0

0

0

0

1

0

0

...

1

0

1

0

1

0

1

0

1

1

0

...

1

0

bitmap:

name

......

.......

***BitMap Index***

SELECT \* FROM tb

WHERE tb.sale<3000

......

stripe1:

sale:min 4000 max 39887

stripe2:

sale:min 2000 max 90000

stripe3:

sale:min 3001 max 28988

stripe2:

sale:min 2 max 2980

......

Total num：>100000

record in:

<3000

***Mix-Max Index***

DATA

woker

operator

woker

operator

woker

operator

INDEX MOUDLE