



# Postgres中国技术大会 2017 ( P G 大象会 )

# PPT分享



## Postgre中国用户会

### ◆ QQ交流群

PostgreSQL专业2群: 100910388

PostgreSQL专业3群: 150657323

PostgreSQL专业4群: 461170054

◆ 文档翻译群: 309292849

◆ 欢迎投稿: [press@postgres.cn](mailto:press@postgres.cn)

### ◆ 微信公众号



官方微信公众号

### ◆ 新浪微博



官方微博

### ◆ 官方网站

[www.postgres.cn](http://www.postgres.cn)

### ◆ FaceBook

China PostgreSQL User Group

### ◆ Twitter

China PostgreSQL User Group



# 阿里云RDS、HDB PG多维存储 特性与案例



嘉宾：digoal  
公司：阿里云



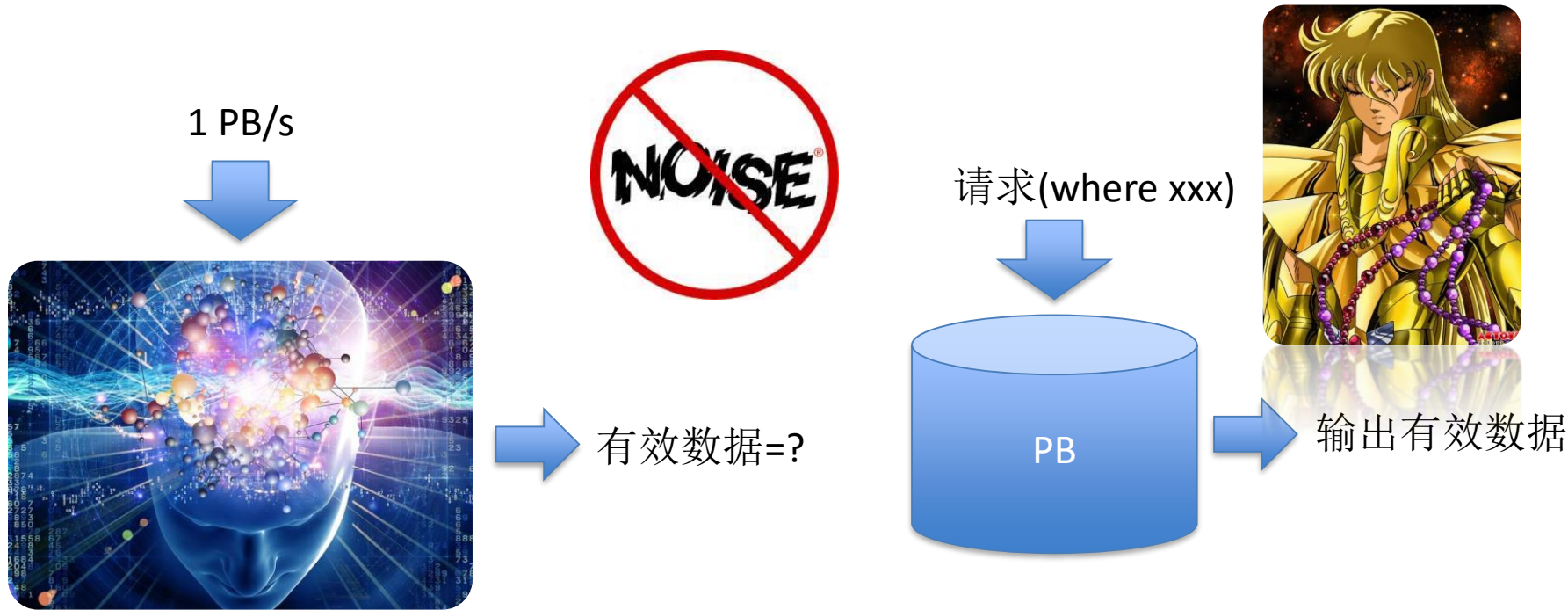


# 目录

- 什么是数据噪音
- 如何过滤噪音
- 什么是隐式噪音
- 如何过滤隐式噪音
- 聚集
- 切割
- 阿里云HybridDB for PG噪音过滤特性



# 什么是数据噪音





# 如何过滤噪音

- 索引
- 分区



# 什么是隐式噪音

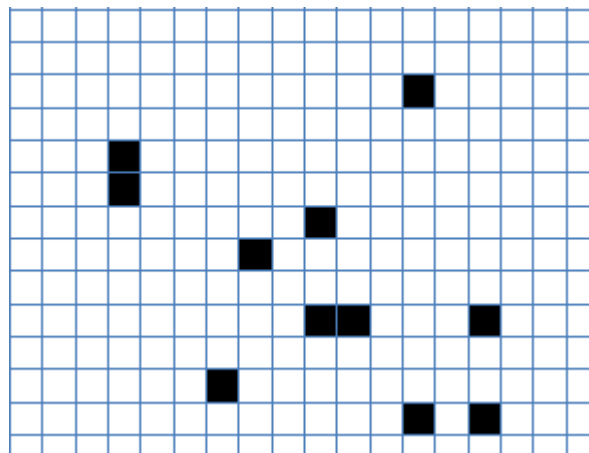
- 访问颗粒引入的噪音
  - 堆颗粒
  - 某些索引实现引入的颗粒





# 什么是隐式噪音

where xxx,  
索引访问



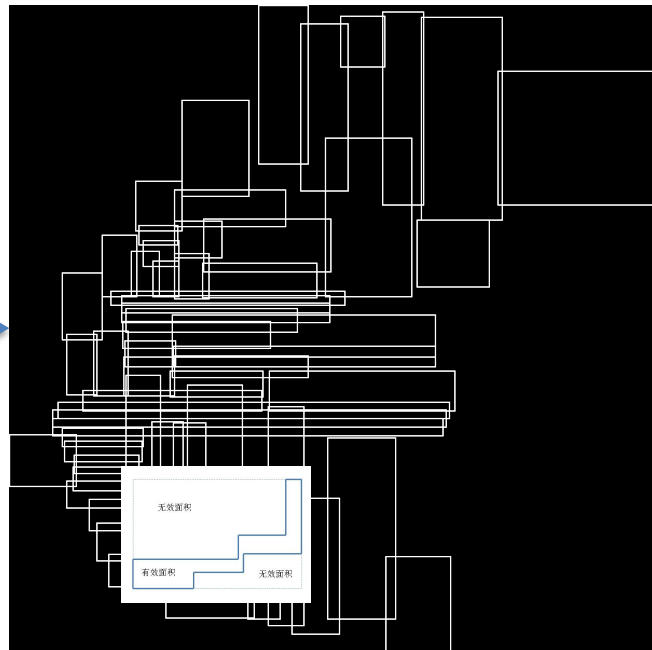
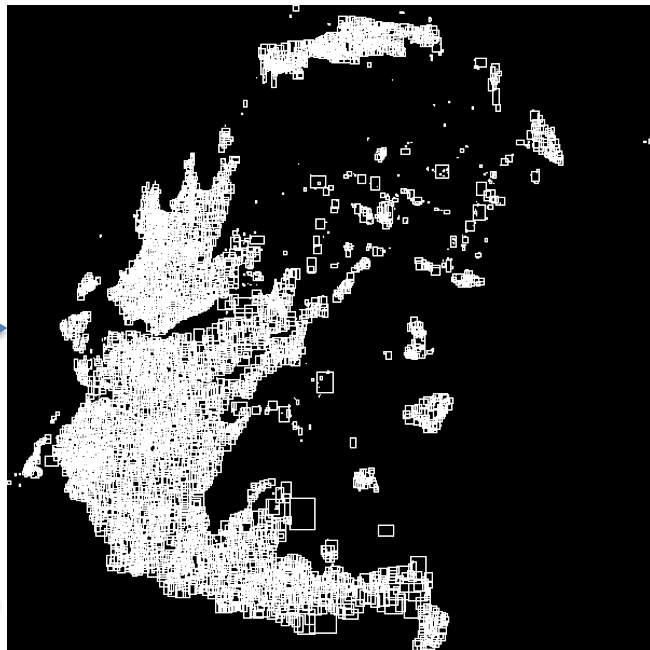
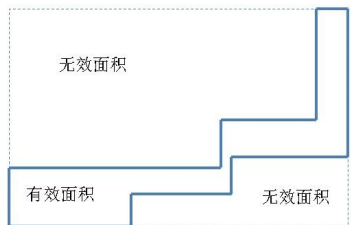
HEAP Storage



输出  
11条记录



# 什么是隐式噪音



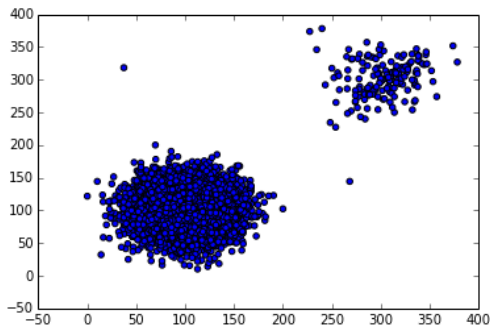
索引放大  
IO+CPU  
隐式噪音





# 如何过滤隐式噪音

- 聚集



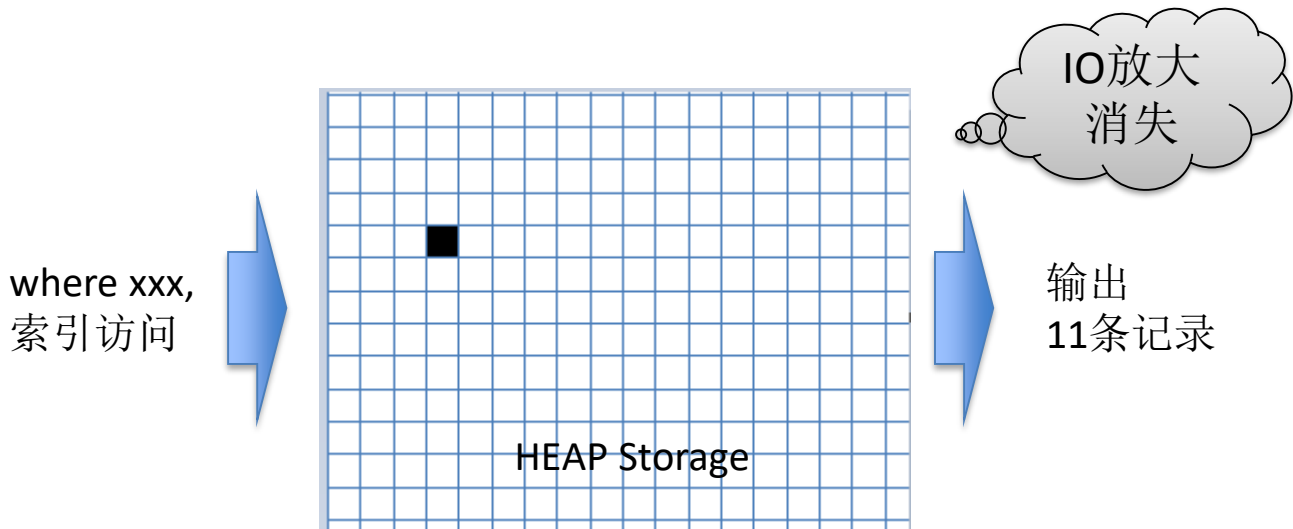
- 切割





# 一维聚集

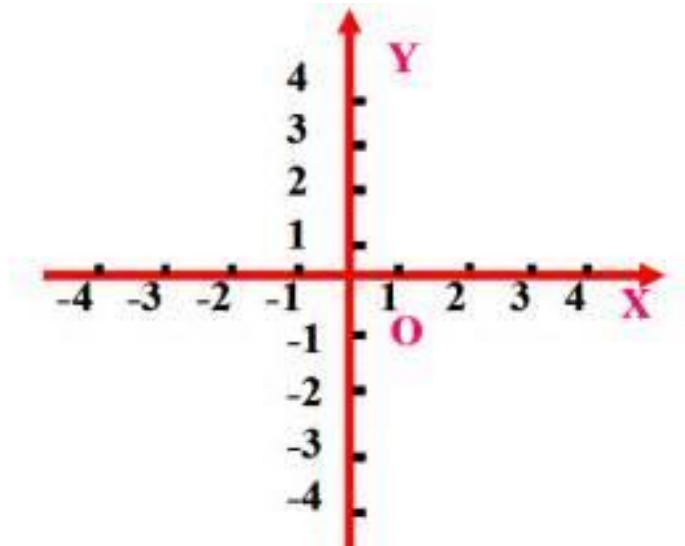
- where col1 =|... xxx; order by col1; .....
- CLUSTER [VERBOSE] table\_name [ USING index\_name ]





# 二维聚集

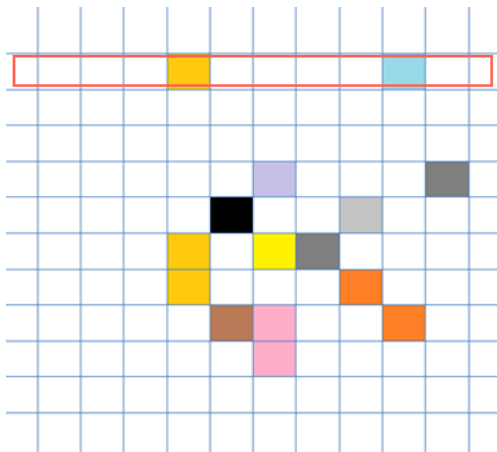
- Where  $a=|...xxx$  AND/OR  $b=|...xxx$ ;
- `row_number()` over (order by col1) as rn1, `row_number()` over (order by col2) as rn2.



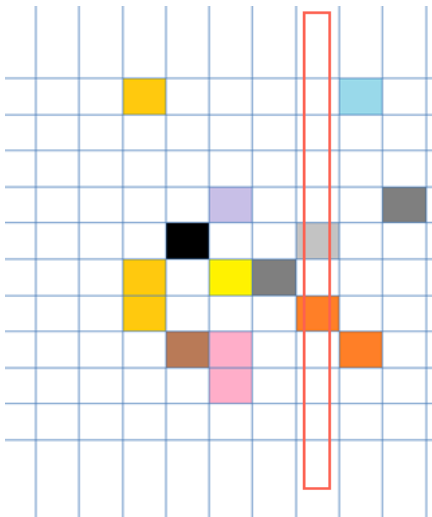


# 二维聚集

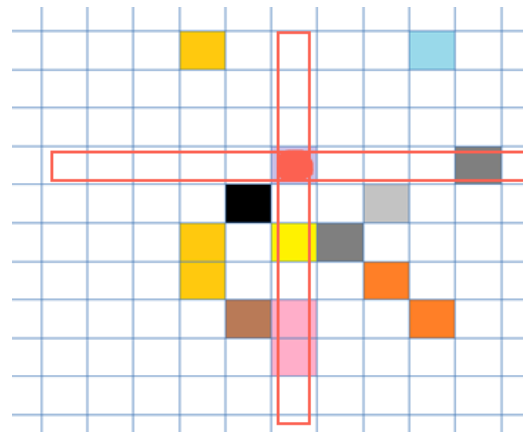
col1 = |... ?;



col2 = |... ?;



col1 = |... ? AND col2 = |...?;

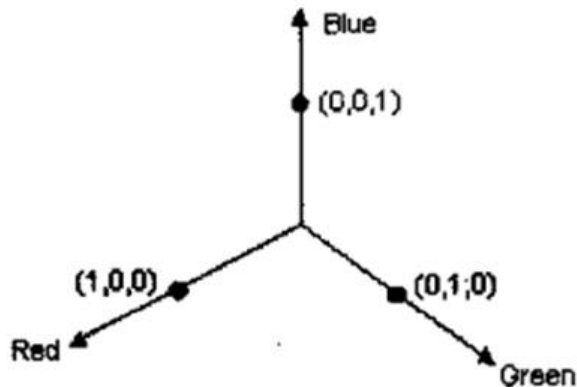


HOW about col1 = |... ? OR col2 = |...?;



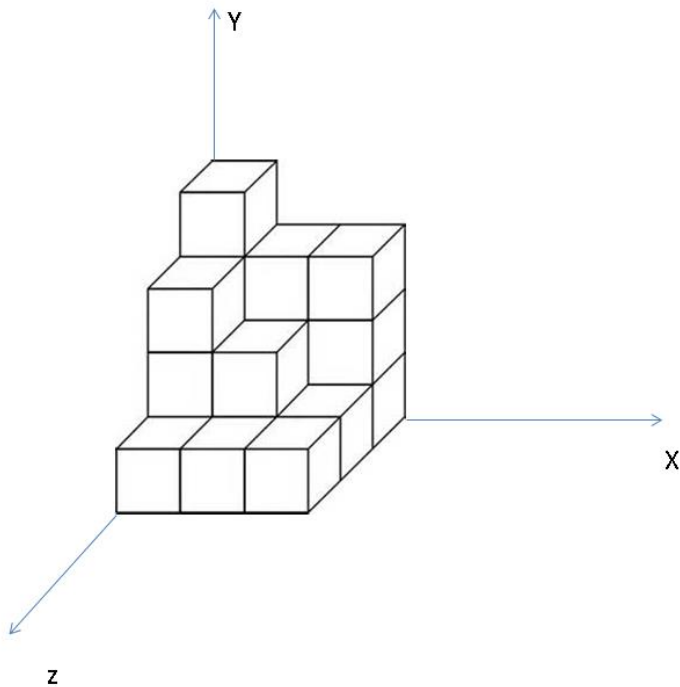
# 三维聚集

- Where  $a=|\dots xxx \text{ AND} | \text{OR } b=|\dots xxx \text{ AND} | \text{OR } c=|\dots xxx;$
- $\text{row\_number() over (order by col1)}$  as rn1,  $\text{row\_number() over (order by col2)}$  as rn2,  $\text{row\_number() over (order by col3)}$  as rn3.





# 三维聚集





# 多维聚集

- Where a=|...xxx AND/OR b=|...xxx AND/OR c=|...xxx AND/OR .....
- row\_number() over (order by col1) as rn1, row\_number() over (order by col2) as rn2, row\_number() over (order by col3) as rn3, row\_number() over (.....).....



# 多维聚集







# 聚集效果

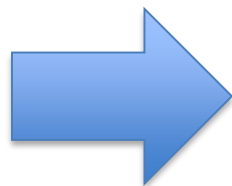
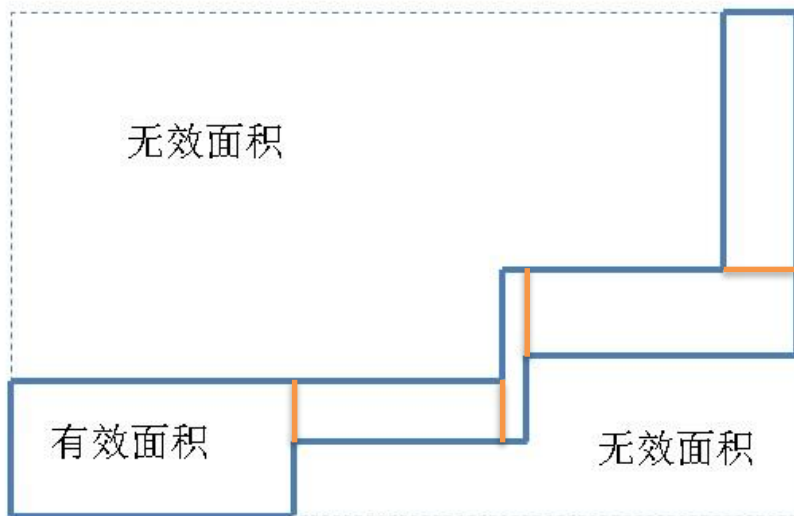
- multi-column BRIN index



过滤性



# 空间切割



切割效果：  
100%或接近100%  
有效面积。



# 阿里云HybridDB for PG噪音过滤特性

- sortkey
- metascan
  - 优于 BRIN

|    |   |   |   |    |   |   |    |   |   |   |   |
|----|---|---|---|----|---|---|----|---|---|---|---|
| 10 | 4 | 8 | 5 | 12 | 2 | 6 | 11 | 3 | 9 | 7 | 1 |
|----|---|---|---|----|---|---|----|---|---|---|---|



- 明天下午
  - 《阿里云HybridDB for PostgreSQL列存优化》



# 例子

- 聚集

- 空间数据

- 标准数据

- 切割

- 空间圈选+透视

轨迹数据、时空透视

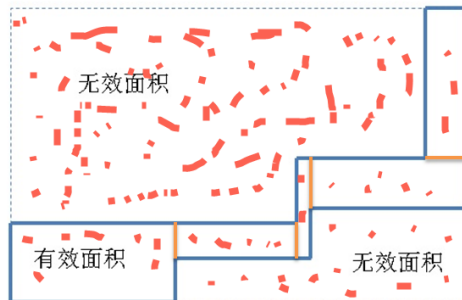
1、时间分区

2、空间聚集(cluster using geohash order)

IoT, 股票数据、实时搜索, 无冷数据

1、schema less

2、UDF





Thanks!

