

# HBase 用例分析

刘佳

liujia@pezy.cn

sina weibo:刘佳\_牛\_plus\_plus

# 提纲

- 数据魔方
  - 业务背景
  - 原有架构
  - 新架构设计实现及效果
- 访客直播间
  - 业务背景
  - 新架构设计实现及效果
- 广点通
  - 业务背景
  - 原有架构
  - 新架构设计实现及效果
- 其他工作



# 数据魔方-业务背景

- 面向B2C商户
- 数据可视化
- 把握行业热点



查询条件 所属分类 > 手机 > Nokia诺基亚 查看 取消

输入您要查询的产品

\*热门产品 展示该类目下50个热销产品，您可以在搜索框中查询全部产品  
热门产品: 5230XM N9 C5-00 5800W 5800Tube E63 C7-00 N97 E71 N95

\*属性选择目前可查看最近7天内的数据

适合送给谁: 男青年 中年男性 男青年 女青年 少女 中年女性 老年女性 老年男性 女童 少男

外观样式: 直板 滑盖 翻盖 旋转 其它

铃声: 普通 3290位 4090位 6490位 MP铃声 立体声铃声 19290位 390位 490位 1290位

摄像头: 10万 30万 130万 200万 300万 320万 100万 50万 500万 700万

屏幕颜色: 256色 4096色 5万5千色 26万 1600万 单色 其它颜色 266色

网络类型: GSM CDMA 小灵通 其它制式 其它制式 CDMA2000(3G) TD-SCDMA(3G) WCDMA(3G) 双模(GSM/CDMA) 双模(GSM/PHS)

主屏尺寸: 3.7英寸 4.5英寸 其它尺寸 其它尺寸 2.9英寸 1.9英寸 2.1英寸 3.2英寸 3.1英寸 1.8英寸

颜色分类: 透明 花色 橘色 白色 黄色 红色 酒红色 紫色 浅灰色 绿色

手机套餐: 套餐一 套餐二 套餐三 套餐四 套餐五 套餐六 套餐七 套餐八 官方标配

储存功能: SD卡 miniSD卡 TF/microSD卡 MMC卡 MSO记忆棒 不支持存储卡 RS-MMC卡 其它存储卡 内置存储

操作系统: Symbian Windows Mobile Palm/宋诺 Linux 无操作系统 其它操作系统 Android OS BlackBerry/黑莓 iPhone

高级功能: 无 大屏幕 QWERTY全键盘 GPS 蓝牙 手写输入 JAVA扩展 MP3播放 视频播放 电视播放

是否智能手机: 是 否

成色: 全新 9.9成新 9.5成新 9成新 8成新 7成及以下

售后服务: 全国联保 店铺三包 其它售后服务 自用(已过保)

手机价格区间: 1001-2000元 拍卖 1000元以下 2001-3000元 3001-5000元 5001元以上

适合的送礼场合: 感恩回馈 学业有成 工作升迁 表达歉意 礼物拜寿 毕业送礼 生日送礼 婴儿诞生 乔迁新居 爱意表达

适合的送礼人物类型: 运动型 文艺型 事业型 时尚爱美丽 社交玩弄型 宅型 精明理财型 探索创新型 小资型

上市时间: 2011年 2009年 2008年 2010年 2004年 2005年 2006年 2007年 2003年以前

触摸屏: 电容式触摸屏 电阻式触摸屏

隐藏部分属性

1 手机 503,434,265 100.00% 1 手机 8.38%

1-1/1条 首页 上一页 1 下一页 末页

1-1/1条 首页 上一页 1 下一页 末页

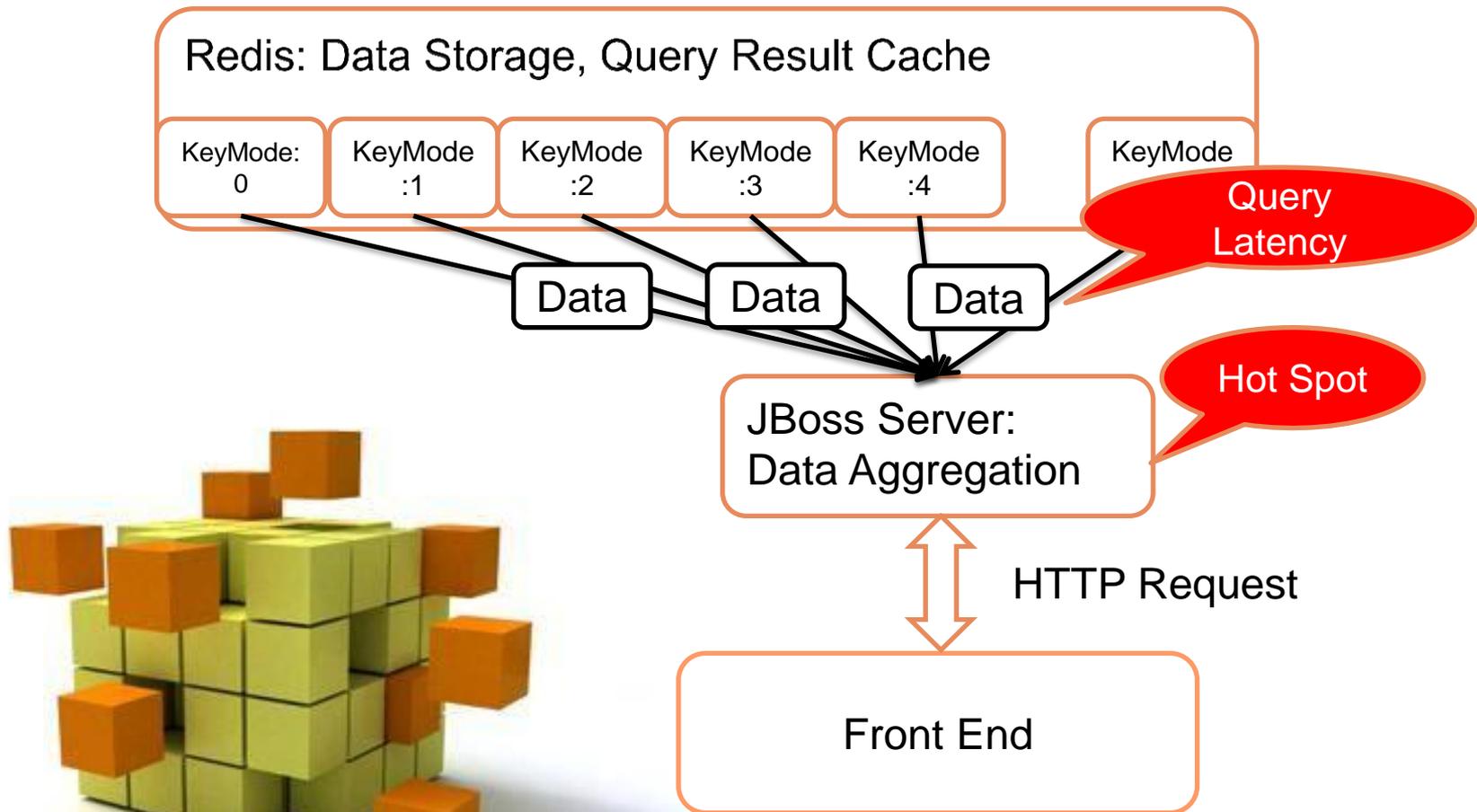
每页显示 10 条

序号	日期	成交金额(¥)	成交金额占比	成交人数	成交笔数	成交商品数	客单价(¥)
1	2011-04-03	74,493,736	14.78%	48,743	63,770	70,793	8.64

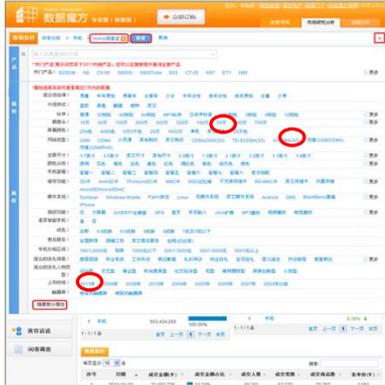
刘佳  
liujia@pezy.cn  
weibo: 刘佳\_牛\_plus\_plus



# 全属性查询原系统



# 全属性查询原系统



时间T : 20121004	时间T : 20121004	时间T : 20121004
类目C : 手机 003	类目C : 手机 003	类目C : 手机 003
属性名P ( 摄像头 : 045 )	属性名P ( 网络类型 : 032 )	属性名P ( 上市年份 : 012 )
属性值V ( 500万像素 : 172 )	属性值V ( WCDMA : 118 )	属性值V ( 2012 : 112 )



Key	Value
20121004,003,045,172	31 34 7
20121004,003,032,118	37 31 34 76 12 89 15
20121004,003,012,112	31 34 12 13

31|34|7

37|31|34|76|12|89|15

31|34|12|13



31|34

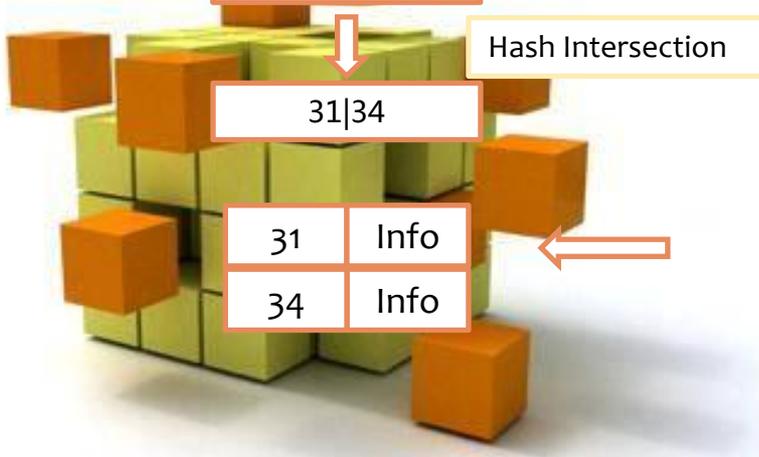
Hash Intersection

31 Info

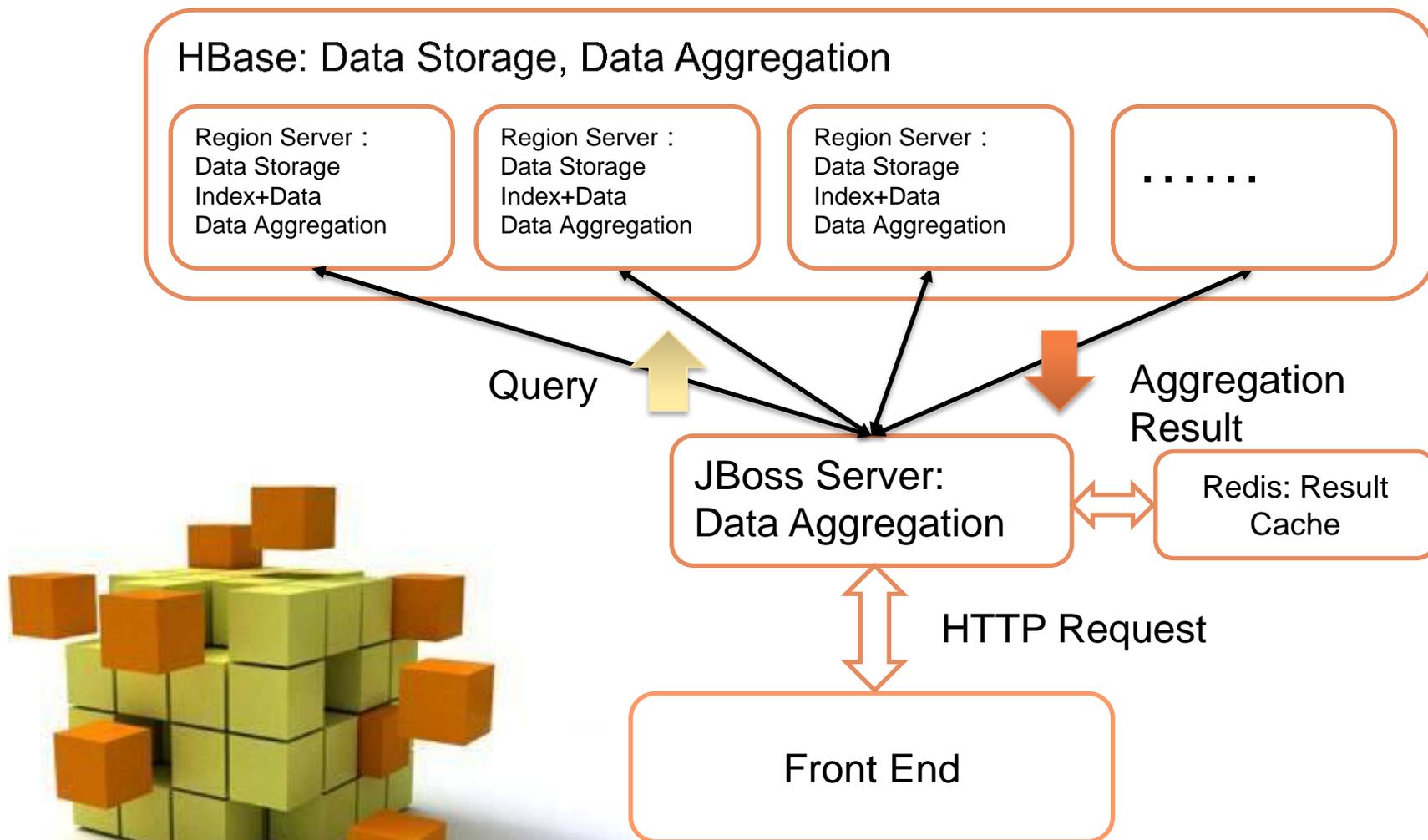
34 Info



Key	Value
12	Transaction, User, Goods
31	Transaction, User, Goods
34	Transaction, User, Goods



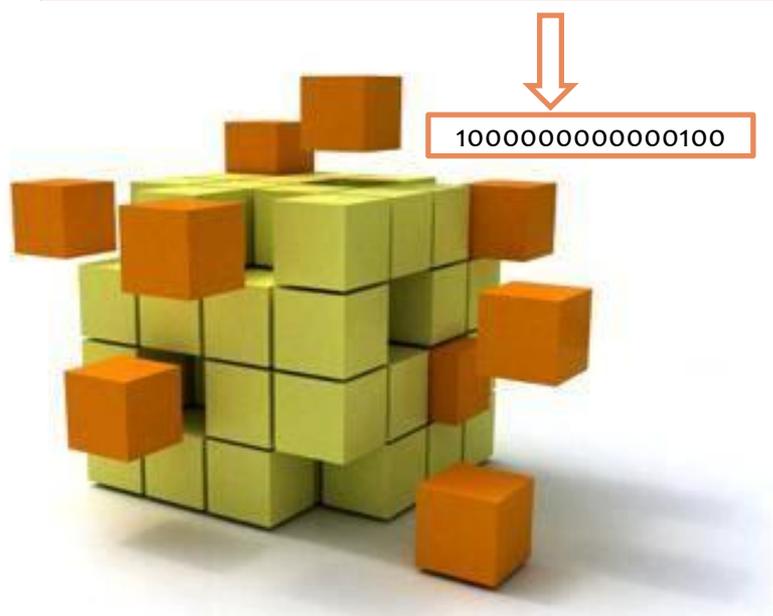
# 全属性查询新系统架构



# 全属性查询新系统逻辑流程

Index and Data

Row Key : TCPV	F1:c Bit Map	F2:c Trans ID	F3:c Trans Info
20121004,003,045,012	1000000010000100	31 34 7	Transaction, User, Goods
20121004,003,032,018	1001001000100100	37 31 34 76 12 89 15	Transaction, User, Goods
20121004,003,012,012	1011000000000100	31 34 12 13	Transaction, User, Goods



↓  
10000000000000100

↓  
31|34  
31|34|15  
31|34

↑  
31(0|1)=>34(0|74)

# 改造后效果

- 取原系统查询日志（2011.5.8~5.18），所有查询按原系统的响应时间分为三类统计平均响应时间。
- 测试前清空cache。
- 测试时HBase库中有25天数据，200G左右。
- 可查询时间从7天拓展到至少30天

查询类型	端计算未优化	端计算 FingerPrint优化	原系统响应时间	查询个数
<100ms	56.2842	36.2817	34.0306	1175
>100ms <1s	283.6209	238.5098	301.1503	460
>1s	1302.492	1120.949	2629.542	59

# RS Aggregation Frame : 进化

```
{ "field": [{"field": "alipay_trade_amt", "id": "123", "type": "sum"}, {"field": "buyer_provine_id", "id": "12333", "type": "distinctCount"}], "limited": [0, 20], "sort": [{"field": "shop_star", "reverse": "0"}], "group": ["shop_star"], "date": ["20101006", "20101008"], "key": ["50013865, 2726665, 18956796"] }
```

key:

```
["50012906, 21541, 38487", "50012906, 20000, 44618"]
```

single key query: ["50013865, 2726665, 18956796"] or mult-key query:

```
["50012906, 21541, 38487", "50012906, 20000, 44618"]
```

field:

```
alipay_trade_amt, "id": "31123", "type": "sum"}, {"field": "alipay_date", "id": "4322", "type": "original"}, {"field": "bu
```

```
{ "field": "alipay_trade_amt", "id": "31123", "type": "sum"}, {"field": "alipay_date", "id": "4322", "type": "original"}, {"field": "buyer_id", "id": "count", "type": "count"} ]
```

group:

```
["alipay_date"]
```

group: ["alipay\_date"] or no group : ["null"]

```
1 [{"count": "104", "4322": "20110724", "31123": "2233188"}, {"count": "90", "4322": "20110725", "31123": "1994442"}, {"count": "91", "4322": "20110726", "31123": "1972554"}, {"count": "76", "4322": "20110727", "31123": "1666422"}, {"count": "114", "4322": "20110728", "31123": "2568594"}, {"count": "121", "4322": "20110729", "31123": "2607544"}, {"count": "5240", "4322": "20110730", "31123": "81059242"}, {"count": "381", "4322": "20110731", "31123": "6616803"}, {"count": "180", "4322": "20110801", "31123": "4303538"}, {"count": "205", "4322": "20110802", "31123": "4249922"}, {"count": "205", "4322": "20110803", "31123": "4303436"}, {"count": "143", "4322": "20110804", "31123": "3095857"}, {"count": "127", "4322": "20110805", "31123": "2648857"}, {"count": "93", "4322": "20110806", "31123": "1941021"}, {"count": "122", "4322": "20110807", "31123": "2673488"}, {"count": "118", "4322": "20110808", "31123": "2506551"}, {"count": "138", "4322": "20110809", "31123": "2873995"}, {"count": "108", "4322": "20110810", "31123": "2402578"}, {"count": "101", "4322": "20110811", "31123": "2264940"}, {"count": "4191", "4322": "20110812", "31123": "72193169"}, {"count": "321", "4322": "20110813", "31123": "6107545"}, {"count": "154", "4322": "20110814", "31123": "3247651"}, {"count": "144", "4322": "20110815", "31123": "3095088"}, {"count": "131", "4322": "20110816", "31123": "2861767"}, {"count": "151", "4322": "20110817", "31123": "2867926"} ]
```

支持全表统计

服务端统计速度 15~50W R/S/Server or 400W R/S/Server.

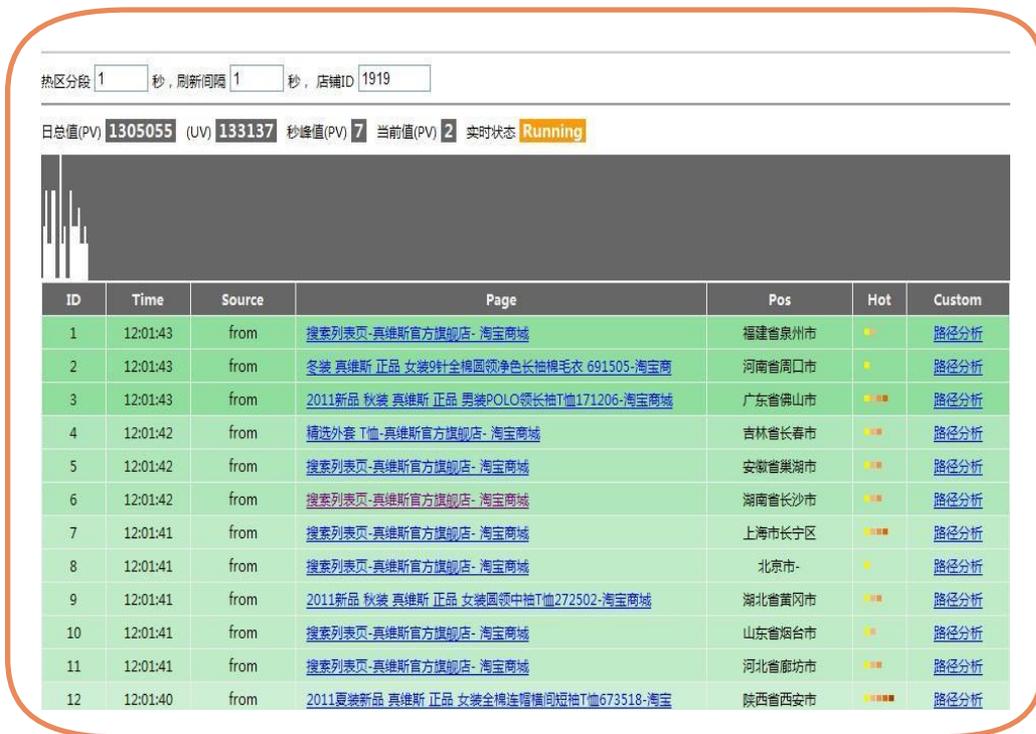
支持sum、avg、count、group by、sort

select 算子( F:C) where 条件 group by(F:C), sort by(F:C)查询。

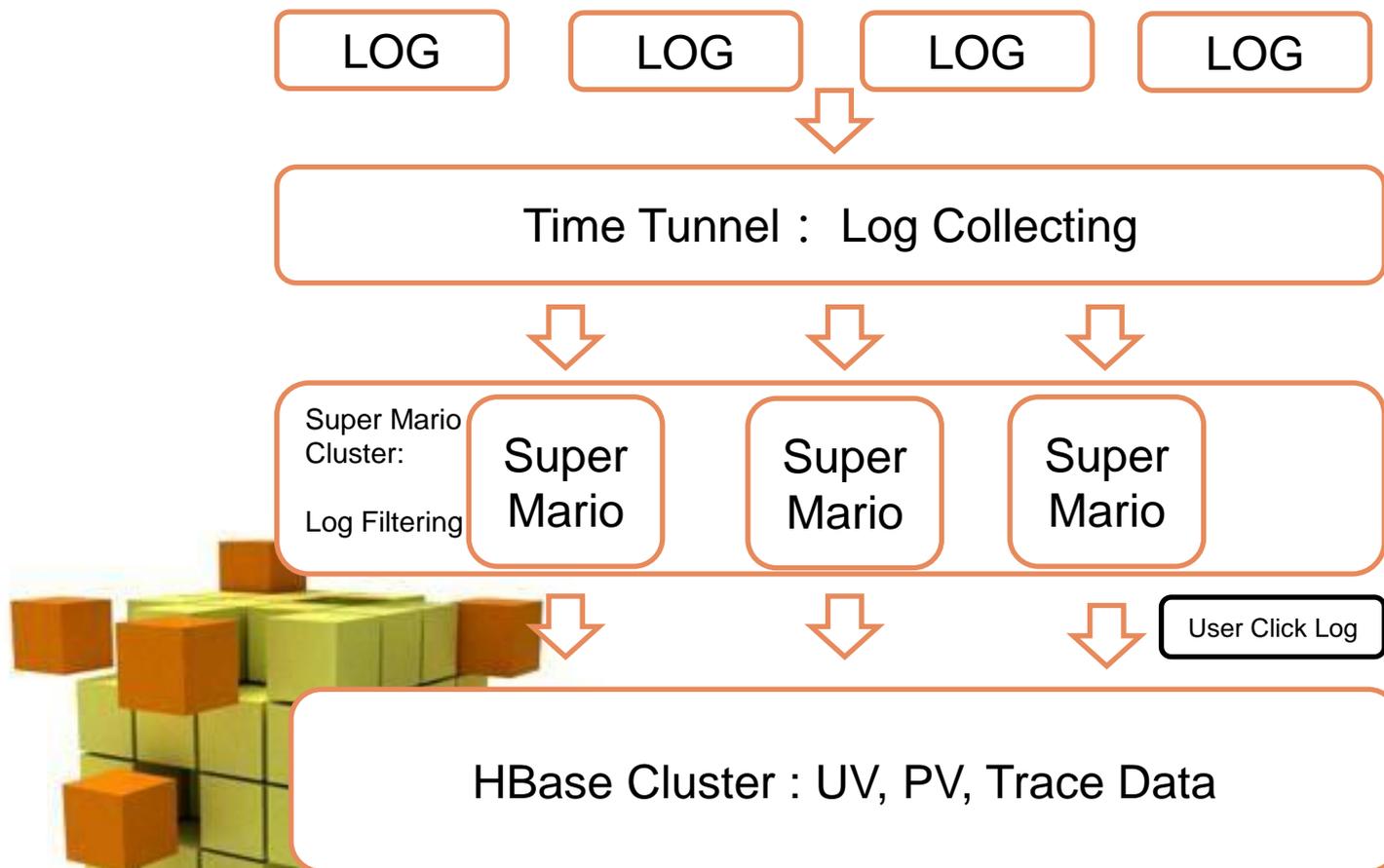
条件 [F:C>|<(rang)]and|or[F:C>|<(rang)]

# 访客直播间-业务背景

- 面向B2C商户
- 实时用户访问情况
- 掌控用户行为，精准营销



# 访客直播间-系统架构





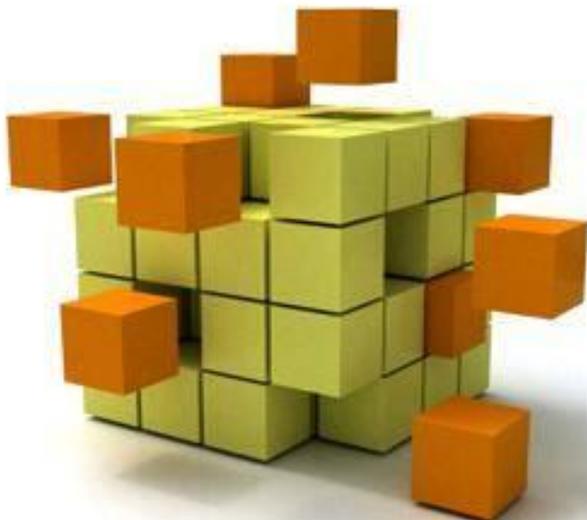
# 广点通-业务背景

- 面向广告主
- 根据用户组当前行为改变推荐策略
- 提高广告点击率



# 业务背景-挑战

- 一秒30000次请求 ( 30%流量, 峰值 )
- 一次请求200个候选广告数据
- 一个候选广告涉及7种CTR数据
- 一种CTR数据800字节 : 48小时的点击曝光计数
- 一次请求响应时间小于40ms
- 一次请求存储查询次数  $7 * 200 = 1400$
- 一次请求存储查询数据量  $7 * 200 * 800 = 1.06\text{MB}$
- 一秒存储层查询次数  $30000 * 200 * 7 = 42000000$  次/秒
- 一秒数据请求量  $30000 * 200 * 7 * 800 = 31\text{GB/秒}$  ( 约400张千兆卡 )

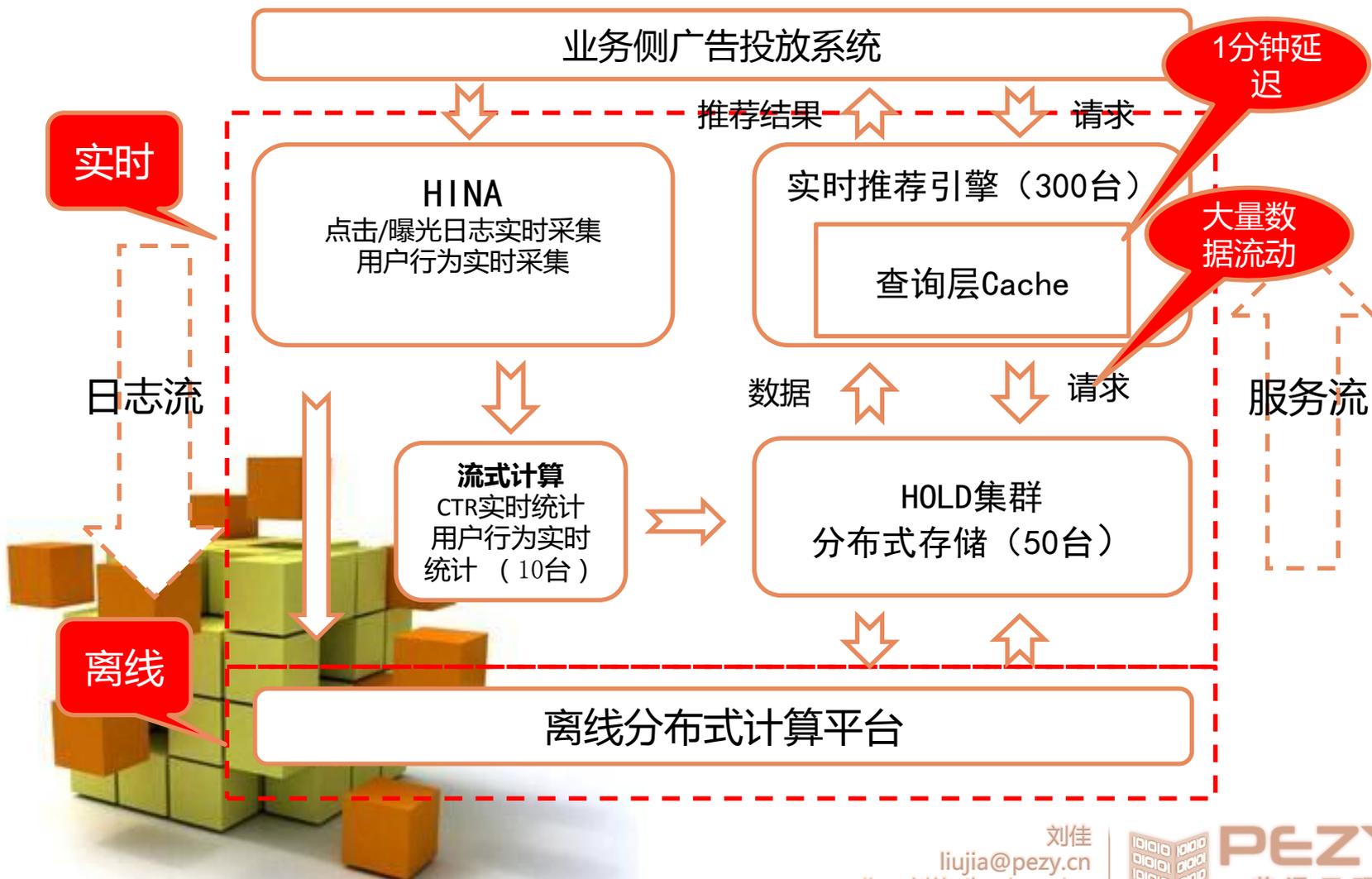


备注:

1、现网实时查询系统, 2012.8.16日总流量12.1亿, 峰值2.4w/s, 均值1.4w/s, 为1.7倍

2、测试设备: 350C1 + 50 B6

# 原系统架构



# 新实验架构

业务侧广告投放系统

HBase Client

Simple RPC



HINA

点击/曝光日志实时采集  
用户行为实时采集

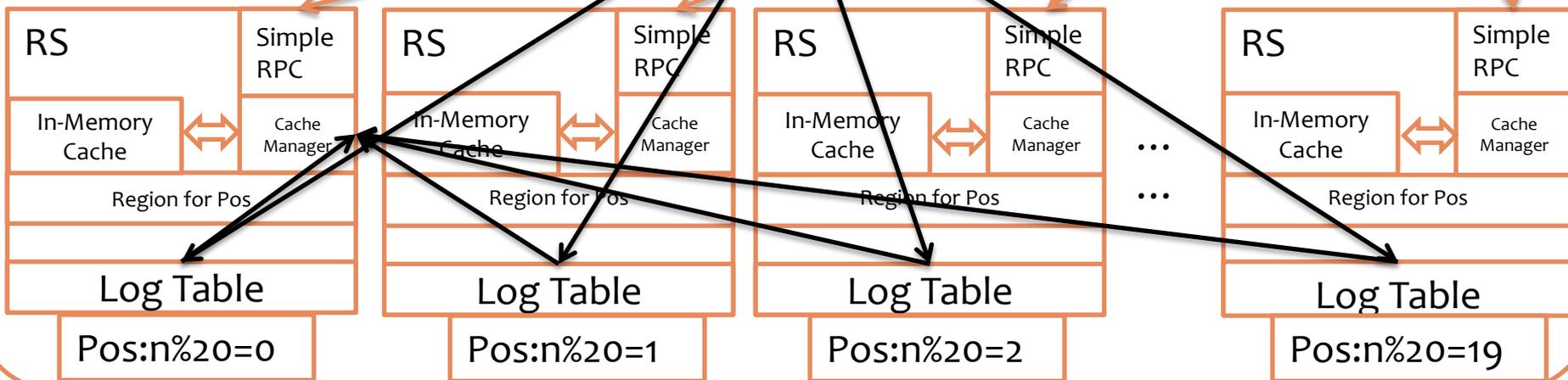


流式计算  
清洗日志  
(10台)

HBase  
RPC

Ad  
Recommendation  
Request

HBase Cluster

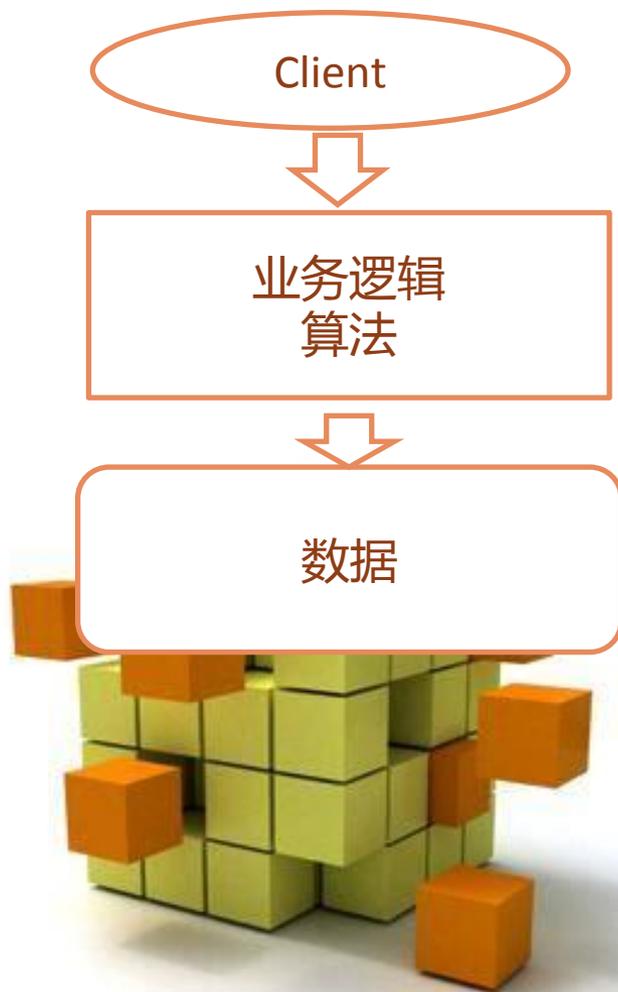


数据冗余，压缩层次



# 方案效果对比

160B6，支撑10亿流量



<<48倍

A VS B

10B6，支撑30亿流量



反思数据魔方



# 其它工作

- 运维管理工具
  - 多租户：访问控制，服务器分组，应用隔离，分等级服务 (CoS)
- 功能扩展
  - 索引，计算，SQL，JDBC，ETL
- 性能提升
  - 写入并行化，客户端RPC连接数，随机读优化

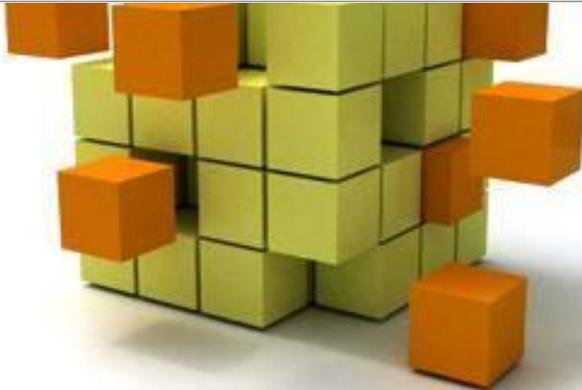
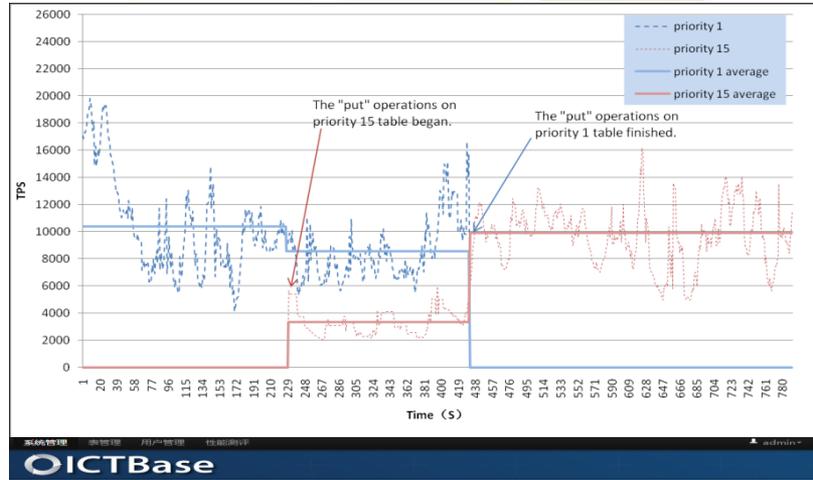
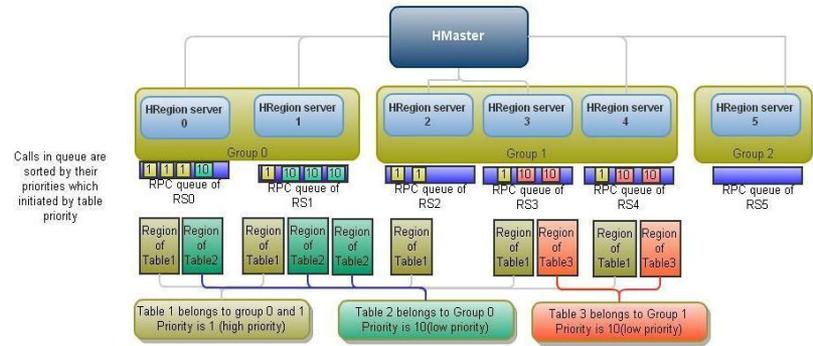
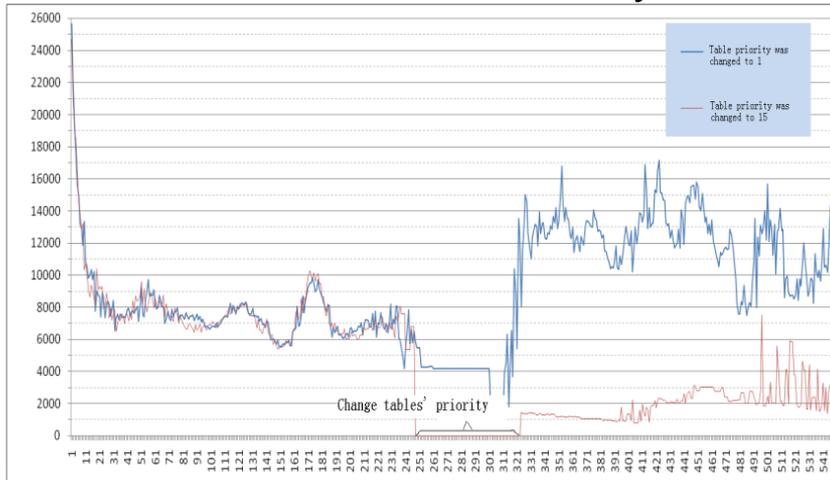


# 其它工作：资源管理

Isolation and allocation: HBase-4120....

RS Group

Table and Method Priority



group01

10 Region Servers 0 Regions

host27:60020	host34:60020	host25:60020	host37:60020
host31:60020	host22:60020	host32:60020	host40:60020
host35:60020	host24:60020		

group0

10 Region Servers 11 Regions (Default Group)

host28:60020	host38:60020	host26:60020	host36:60020
host23:60020	host21:60020	host30:60020	host33:60020
host29:60020	host39:60020		

+ Add Group

刘佳  
liujia@pezy.cn  
weibo: 刘佳\_牛\_plus\_plus



# 其它工作：访问控制



# 其它工作：功能扩展

- 功能扩展
  - 索引, 计算, SQL, JDBC, ETL

系统管理 表管理 用户管理 性能测评 admin

OICTBase

+ 创建表

100w

列簇信息

列簇	最大版本	压缩	在内存	快缓存	存活时间	布隆过滤器
CF1	3	NONE	false	false	2147483647	NONE
列		类型				
col1		INT				
col2		LONG				
col3		DOUBLE				
col4		STRING				
CF2	3	NONE	false	false	2147483647	NONE
列		类型				
col5		STRING				
col6		STRING				
col7		STRING				
col8		INT				

索引信息

列簇	列	索引类型	附加列
CF2	col5	SECONDARYINDEX	--
CF2	col7	IMPSECONDARYINDEX	(family:CF1, columns:col4)

系统管理 表管理 用户管理 admin

OICTBase

+ 创建表

test\_jdbc

```
1 select count(CF1:col2), avg(CF1:col3), count(CF1:col2)/avg(CF1:col3) from 1e where key='abd' and key='abd' CF1:col1:500 and CF2:col16='Maria' group by CF2:col6
```

查询

groupName	count(CF1:col2)	avg(CF1:col3)	count(CF1:col2)/avg(CF1:col3)
Bush	28	510.8208834542025	0.05481373394654945
Cook	26	470.4902195668796	0.05526151005632994
Fox	20	579.3842785090766	0.03451940403261509
Green	23	548.8520182414186	0.04190564894649471
Hall	23	486.0440115827724	0.04732081756362333
Hill	21	536.0309719460781	0.03917684070336236
Mill	17	446.27189739743034	0.038093368861316715
Miller	29	561.4097160528783	0.05165567885766433
Sharp	16	438.4178295528804	0.036494866133335795

系统管理 表管理 用户管理 性能测评 admin

OICTBase

+ 创建表

100w

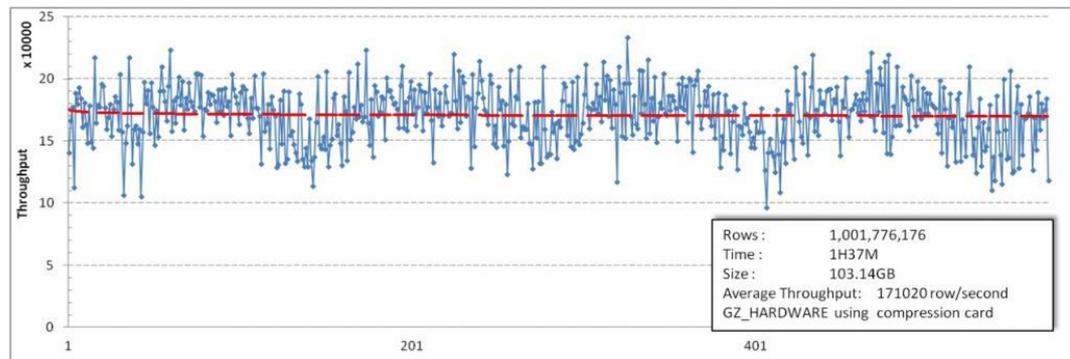
```
1 select * from 100w where CF1:col1>500 and CF2:col5='Maria'
```

查询

	CF1:col1	CF1:col2	CF1:col3	CF1:col4	CF2:col5
51ef017659e4504	780	427666017515536889	970.7522104826297	2012-11-22 00:23:07	Maria
9fb53e7e2b2d68a	675	4614625406549243510	777.8509510139554	2012-11-23 15:28:40	Maria
30b8e57877688d3	945	4254944304286112016	303.17415601547236	2012-11-22 18:27:17	Maria
169408d0bbd0285	614	6248712494567630463	261.7165146279625	2012-11-24 05:09:43	Maria
9830e2061e6a44	749	6559542934257137374	924.9923408389884	2012-11-25 13:57:04	Maria
42d379f124d6f2	680	5936576895411589923	828.603623233309	2012-11-26 06:22:53	Maria
7132882812b498b	634	6271260504142121746	797.070686733911	2012-11-05 19:49:20	Maria



# 其它工作：性能优化



- 写入：并行化优化17万/秒 8超16 CPU，16超32 CPU 30万/秒（No HLog，写HLog 25%下降），硬件压缩。
- 随机读：Pread。HDFS优化：Shortcut，Special Dir，(HDFS长连接)。Region and HFile Block management
- 性能是后期做**大规模**Join（跨表操作）、Sort、Group（排序、中间结果暂存）的基础。



参：

<http://blog.csdn.net/kalaamong/article/details/7324894>

刘佳  
liujia@pezy.cn  
weibo: 刘佳\_牛\_plus\_plus



# 致谢

- 预祝今年大会圆满成功。
- 感谢大家，请提问。

