



基于HBase的 企业级大数据平台

目 录

- 1 公司及平台简介**
- 2 HBase为企业级平台提供的解决方案**
- 3 企业级业务应用支持**

我们是谁？

国内最大的保险公司

数亿客户

十几亿保单

上百万营销员

传统数据管理手段所面临的挑战

- 数据虽已集中，但未被集成和整合。
- 范式化模型带来的查询时需要大量join操作。
- 难以实现以客户为颗粒度的查询和分析
- 没有一个ROI很好的解决方案来实现我们的需要。

通过建立企业级大数据平台，我们期望能做到

- 企业级数据资产的整合和连接
- 基于该平台来支撑企业内部的各类数据应用
 - 面向营销员、客户的精确查询服务
 - 面向以客户为颗粒度的精准分析
 - 实现以客户为中心的风控体系
 -
- 以一个较低的成本来解决我们的各类需求

金字塔大数据平台——现况

规模	加工	数据	应用
5TB+高速内存 1PB+存储容量	每日加工上亿条 数据	数亿客户数据 十几亿保单数据	为公司管理层及 百万销售员等提 供数据查询和计 算服务

目录

- 1 公司及平台简介
- 2 HBase为企业级平台提供的解决方案
- 3 企业级业务应用支持



传统数据库存在
越来越多的问题

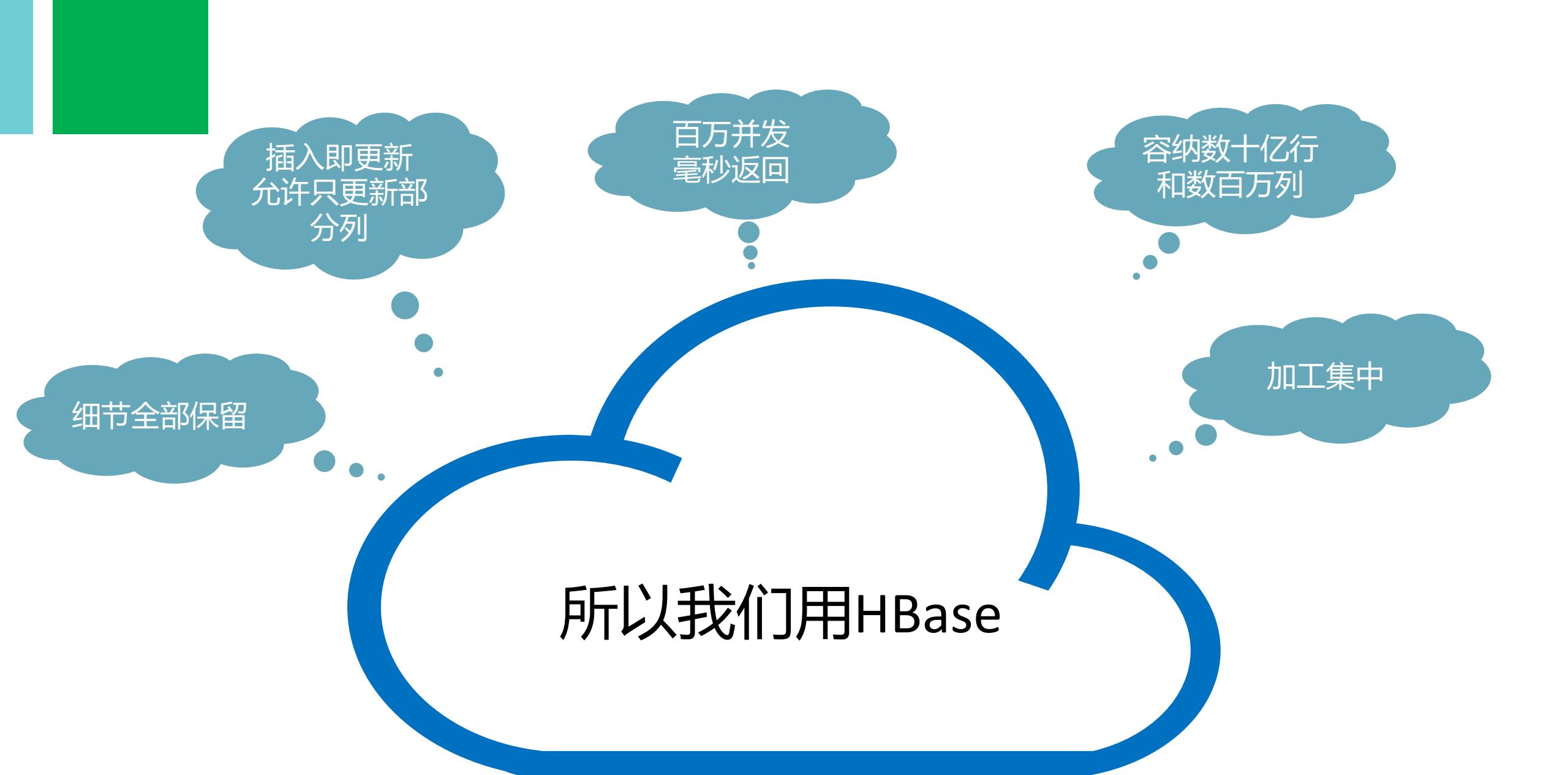
系统繁多

查询复杂

加工缓慢

丢失细节

更新困难



One Giant Table for All System

一张
宽表
存储
全国
系统
保单

容纳上亿行数据
与上百万列数据

HBASE

rowkey1(来自系统A)

rowkey2(来自系统A)

rowkey3(来自系统B)

rowkey4(来自系统C)



rowkeyn(来自系统D)

公司原始
数据库

系统A

系统B

系统C

系统D

SQL Server

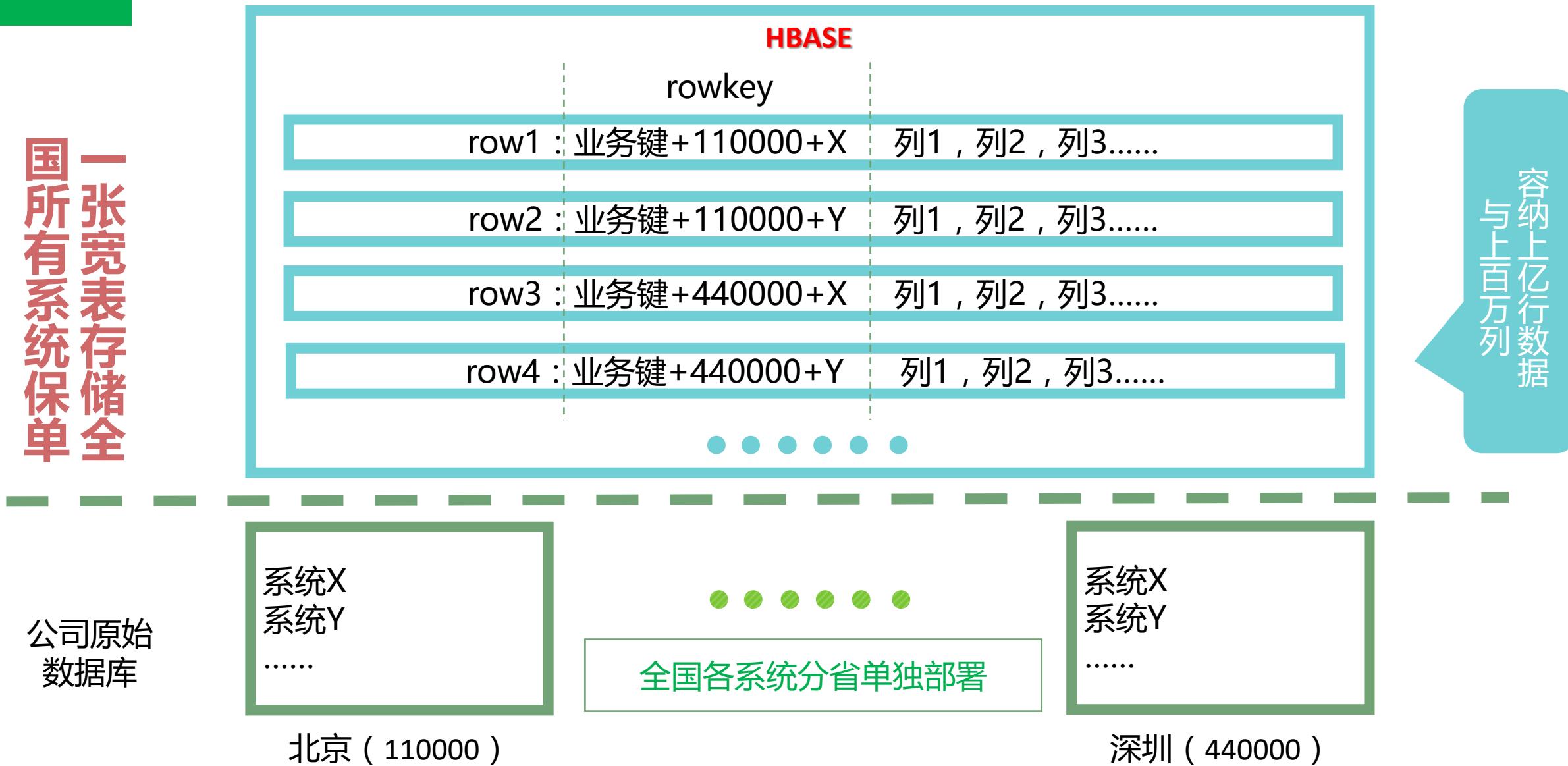
Oracal

MongoDB

infomix

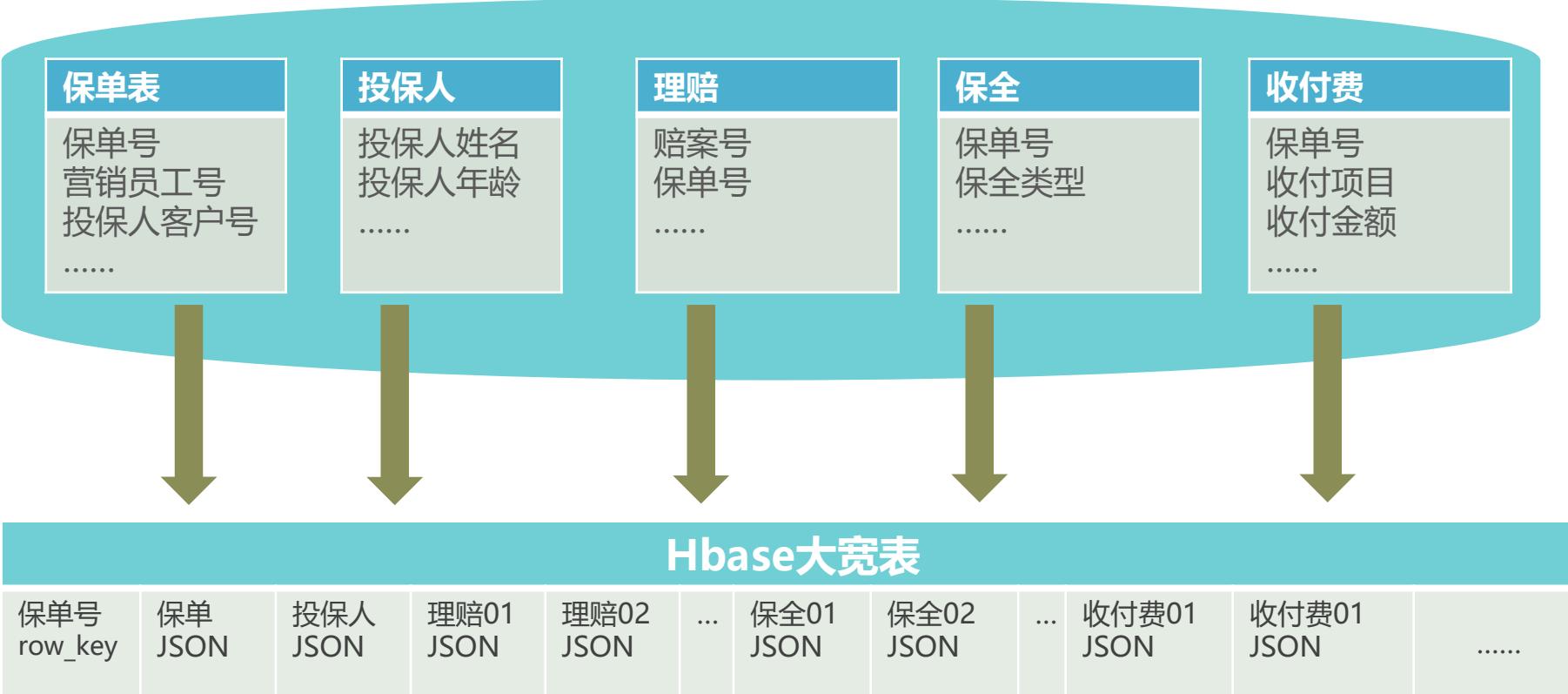
国所有系统保全
一张宽表存储全

One Giant Table for All System



One Row for One Contract

数据按行聚集



增加效率
提前 JOIN
ON

新增数据仅
需添加新列

One Column for One Original Row

一个程序即可计算
全部保单指标

原始表

收付费表

保单号	收付流水号	收付项目	收付金额
001	11	A	100
001	12	B	200
001	13	C	-100

彻底解决
更新难、
更新慢的
问题。

“11”是收付费表主键值

行内容以JSON的方式进行存储

保单号 <i>row_key</i>	收付费11	收付费12	收付费13
001	{"流水号": "11", "收付项目": "A", "收付金额": "100"}	{"流水号": "12", "收付项目": "B", "收付金额": "200"}	{"流水号": "13", "收付项目": "C", "收付金额": "-100"}

宽表

新增记录
仅需要在
HBase中
追加新列
即可。

One Json for All Content

原系统表

保单号	字段1	字段2	字段3	字段4	字段5
11	A1	A2	A3	A4	A5
12	B1	B2	B3	B4	B5
13	C1	C2	C3	C4	C5



金字塔大数据平台

11 column=ci:保单信息, value={保单号:11,字段1:A1,字段2:A2,字段3:A3,字段4:A4}

12 column=ci:保单信息, value={保单号:12,字段1:B1,字段2:B2,字段3:B3,字段4:B4}

13 column=ci:保单信息, value={保单号:13,字段1:C1,字段2:C2,字段3:C3,字段4:C4}

谨防丢失
保留细节

目 录

- 1 公司及平台简介
- 2 **HBase**为企业级平台提供的解决方案
- 3 企业级业务应用支持

客户资源管理

客户画像：支持数亿客户上千标签的高效刷新，精准刻画老客户

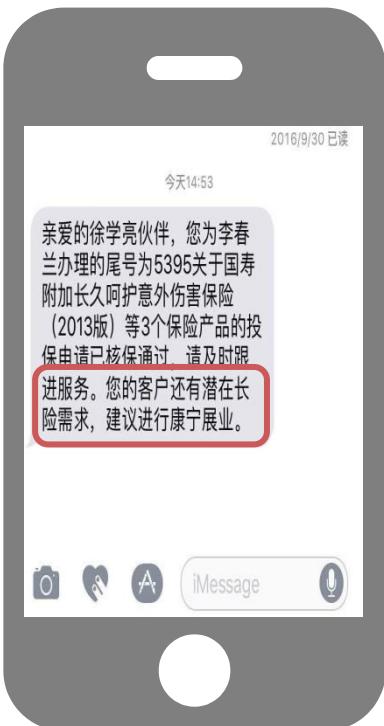


精准营销

购买能力分析

分析已有客户的再次购买能力，
并推送给前台

计算客户的购买能力并将结果推送给营销员



提高展业成功率



客户开发平台

实现对已有客户的分类筛选、定向推送和追踪

查询客户

开发追踪

筛选客户

推荐客户

保障分析

保障分析： 分析客户已有保障，试算客户应有保障

产品介绍

对公司200多总险种进行大数据建模分析，从身故、伤残、重疾、医疗、养老5个角度分析客户保障情况。

应用于手机端APP



发现客户保障缺失
提供销售契机



大数据智能核保

风险保额预警：实现以客户为中心的累计风险保额计算

产品介绍

利用大数据平台整合客户、保单数据，进行以客户为中心的风险保额计算，为公司核保系统提供强有力的数据支撑。

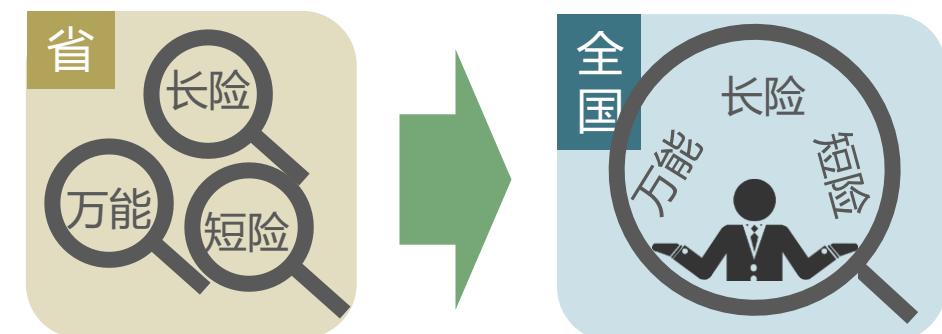
两百多个保单粒度风险保额

三百多个客户粒度风险保额

- 长险寿险净风险保额
- 长险重疾险净风险保额
- 长险护理险净风险保额
- 长险客运交通工具意外伤残净风险保额
- 长险航空意外伤残净风险保额
- 长险重大自然灾害意外伤残净风险保额

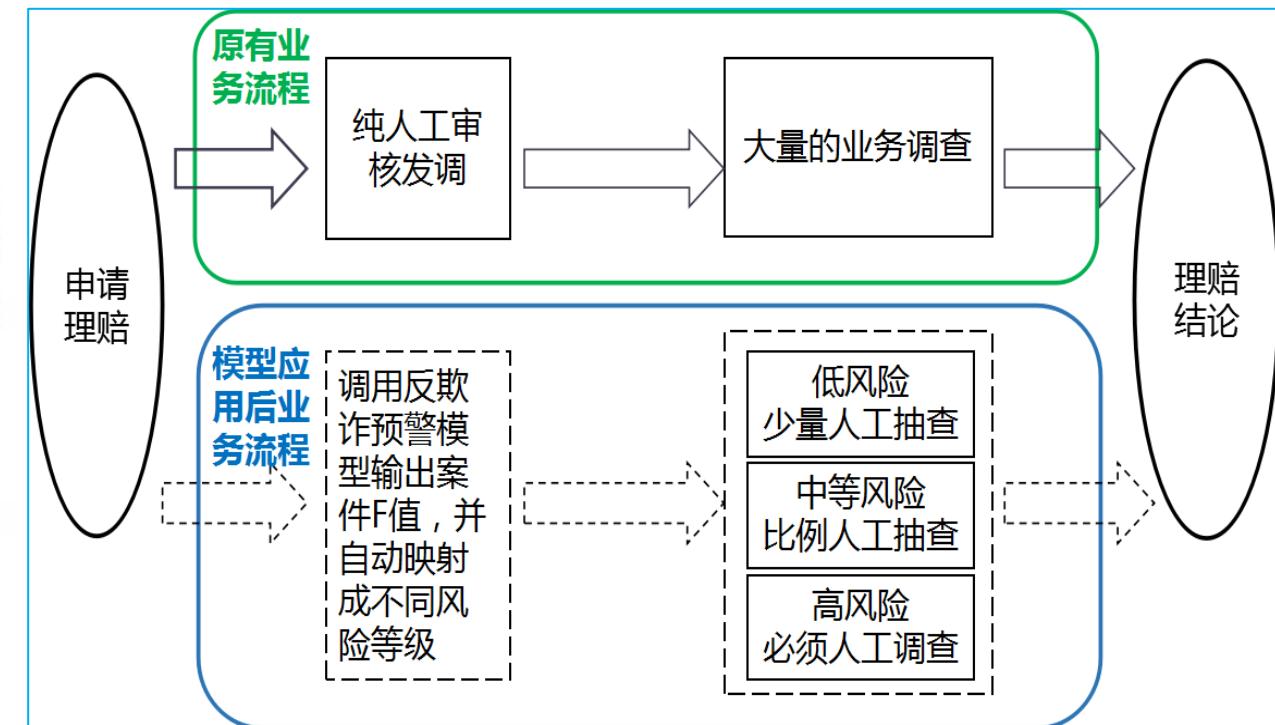
.....

整合多系统全国数据 **提升**核保智能化作业



智能理赔反欺诈

智能理赔反欺诈： 实现毫秒级反馈欺诈风险评估结果



下一步工作



不断丰富和深入基于HBase的各类应用



通过主备集群来提高业务连续性



通过完善大数据应用来实现实时统计分析

敬请指正