



——大数据与安全技术沙龙-长沙站——

大数据平台异构实时同步探索

唐田寿

- 上海新炬数据库专家
- 具有甲骨文高级认证、红帽认证的工程师
- 湖南师范大学本科&Oracle技术嘉宾
- 广东电力数据库主管、项目经理
- 长期服务移动、联通、电力行业，具有8年的数据库运维经验
- 擅长oracle 10G、11G、12c，有丰富的行业领域知识
- 某运营商大数据平台实时同步项目技术负责人

1 大数据概述

2 大数据时代下的 数据同步技术


3 某运营商实战案例



What is BigData

通常对于大数据的理解

数据量大 →  分布式存储

非结构化 →  Hadoop

产生速度快 →  流计算

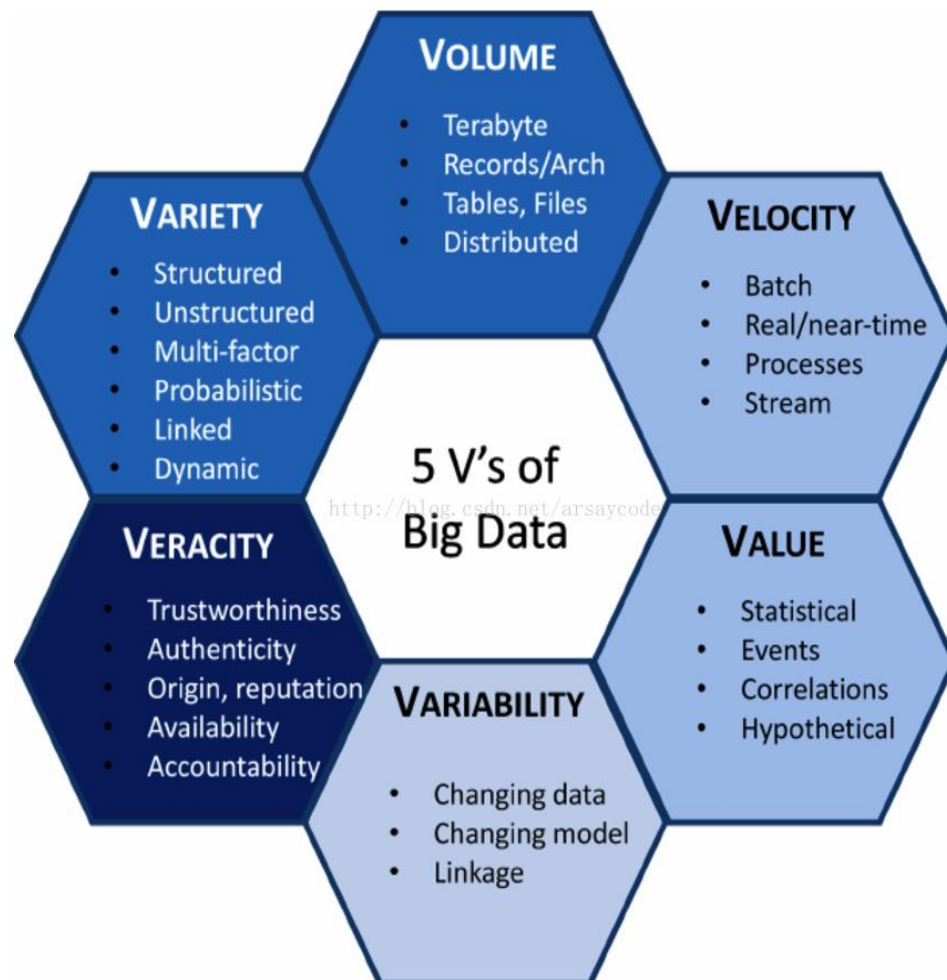
Pivotal 对于大数据的理解



单纯的数据与平台没有意义，
只有当相关的数据结合到客
户需求以及企业业务目标的
模型并通过极致友好的**应用**
渠道让用户真正受益，这才
是大数据技术的意义所在



What is BigData



Volume: 巨大的数据量

Variety: 非结构化数据多样性

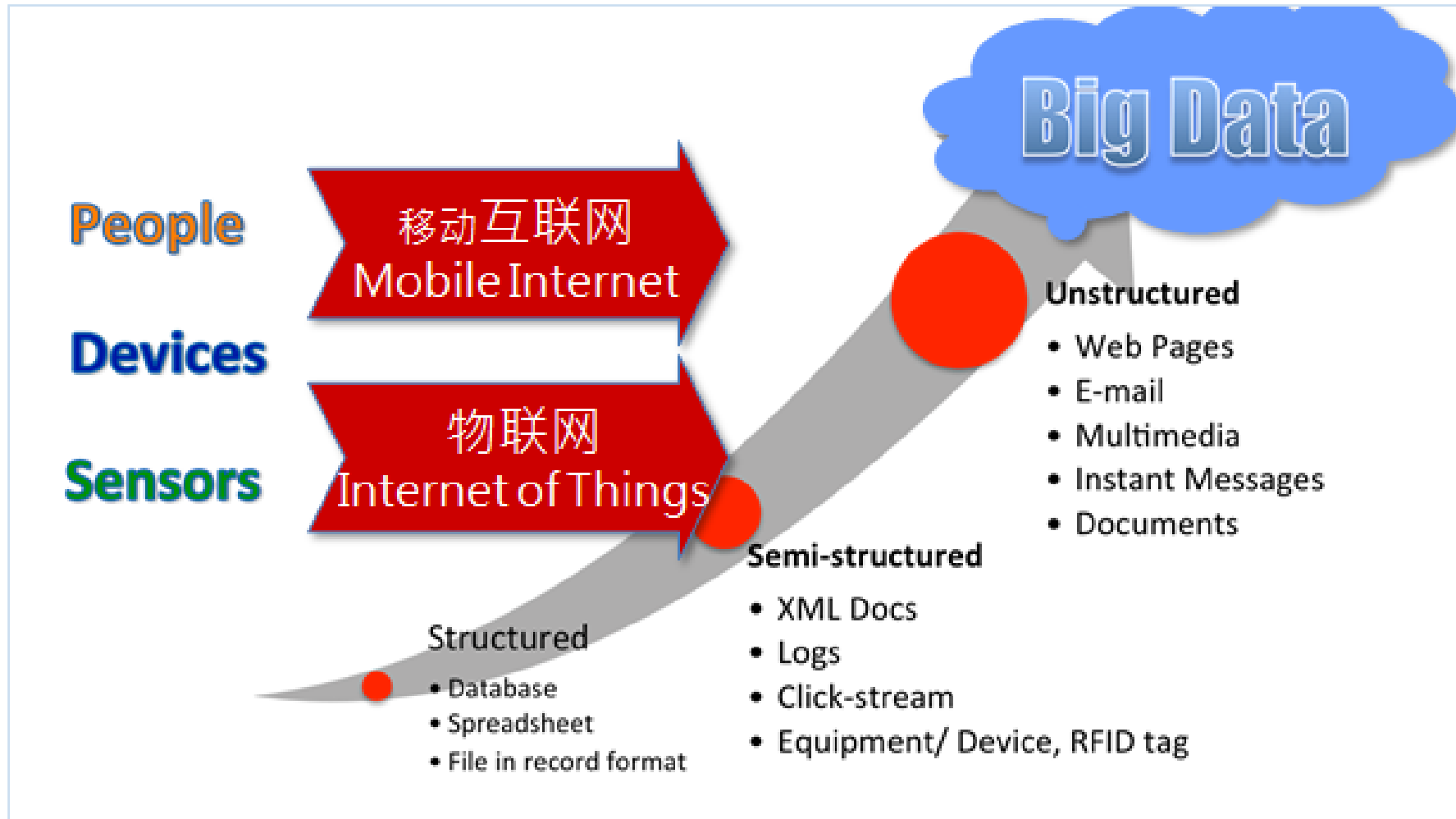
Velocity: 数据增长速度快

Value: 数据的价值性

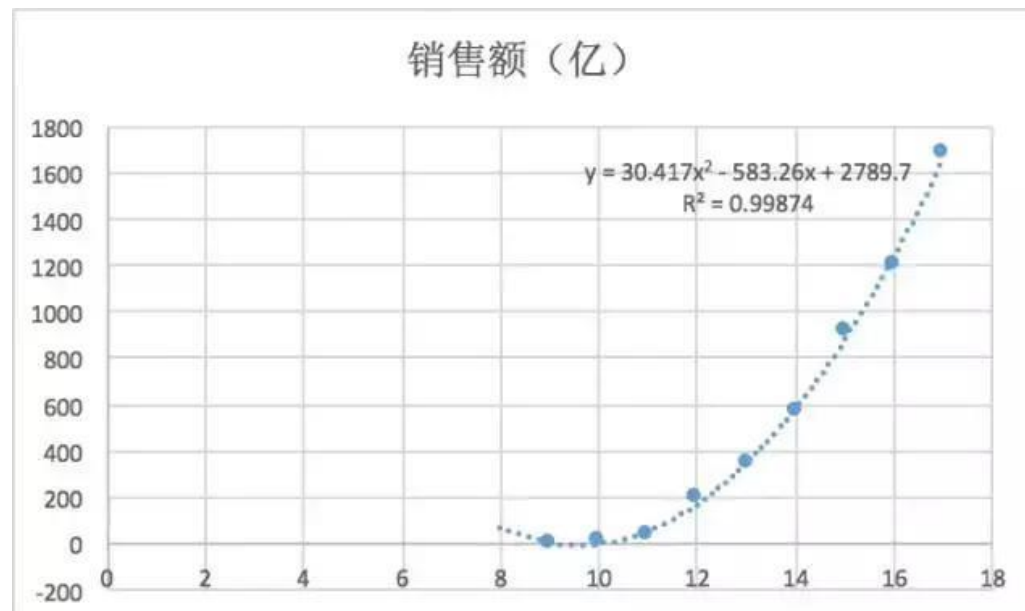
Vercity: 数据的真实性

Variability: 数据的可变性

How is BigData



why is BigData



2016年天猫“双11”全球购物狂欢节的成交总额：破**1000**亿元

2017年天猫“双11”全球购物狂欢节的成交总额：破**1500**亿元

2018年天猫“双11”全球购物狂欢节的成交总额：破**2146**亿元

一、主要技术：

1. Olap : Hadoop Hive,Hadoop Impala,Spark SQL
2. Oltp: HBase, MongoDB,cassderan
3. 搜索引擎: Solr, ElasticSearch
4. 机器学习: Spark Mllib,Hadoop Mahout ,Hama,GrahpX
5. 缓存: Redis, MemcachedDB
6. 可视化: Hue
7. ETL: Pig,Flume,Sqoop(数据采集)
8. 消息队列: Kafka
9. RPC: Thrift
10. 分布式协调管理: Zookeeper
11. 工作流: Ozzie

二、主要产品：

Apache 开源hadoop Cloudera 企业hadoop
华为 FusionInsight 星环 星环大数据平台

大数据经典框架分析



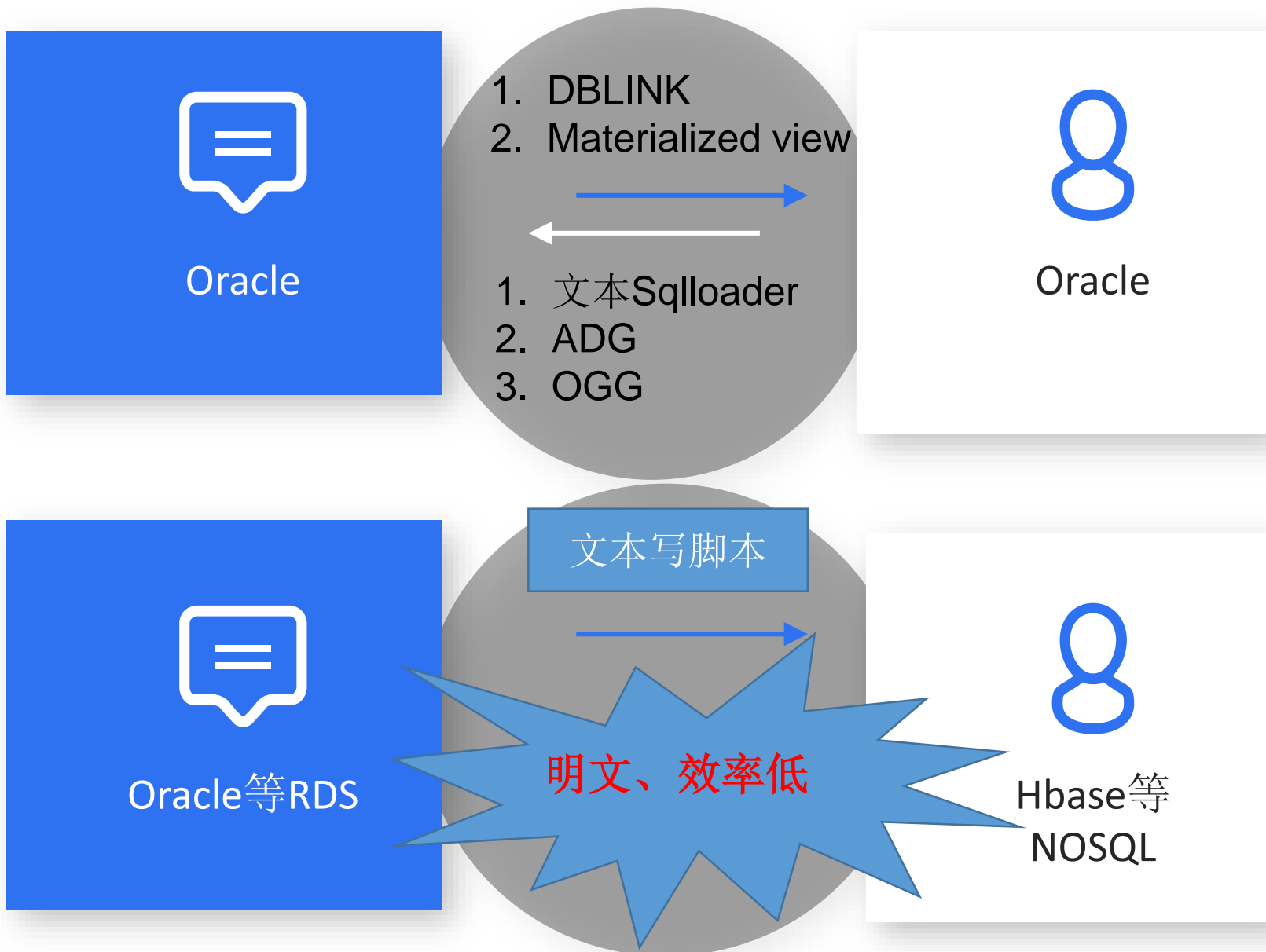
1 大数据概述

2 大数据时代下的 数据同步技术

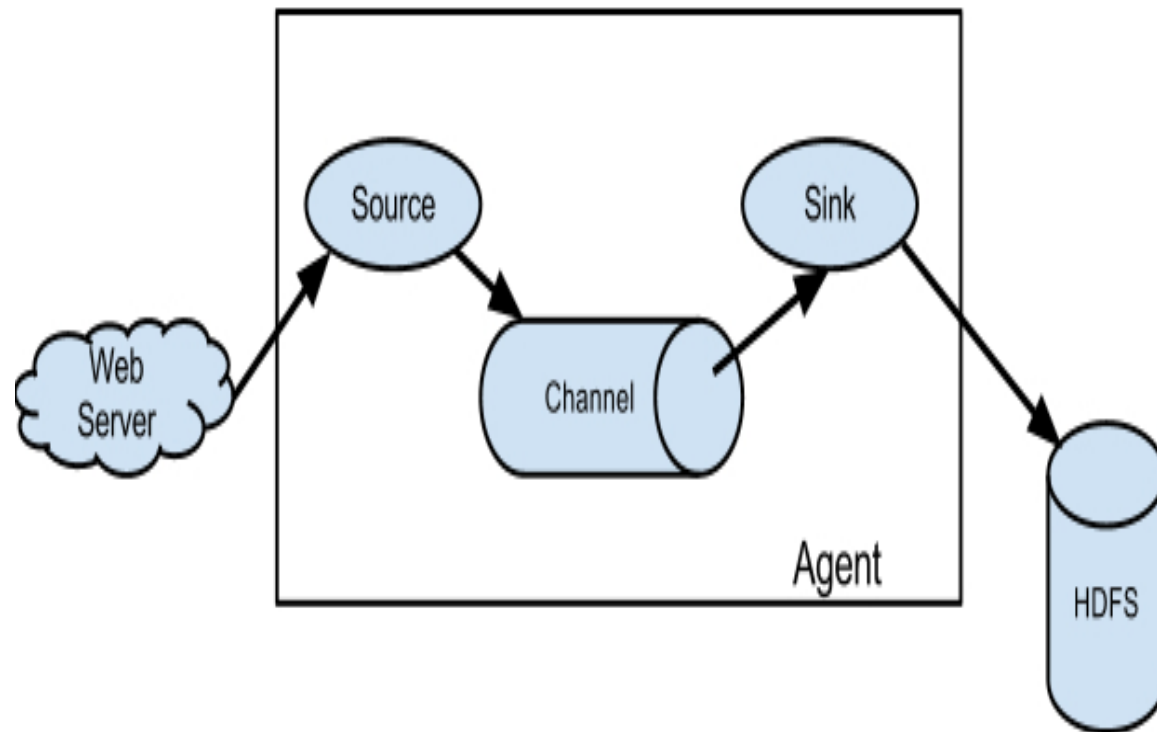
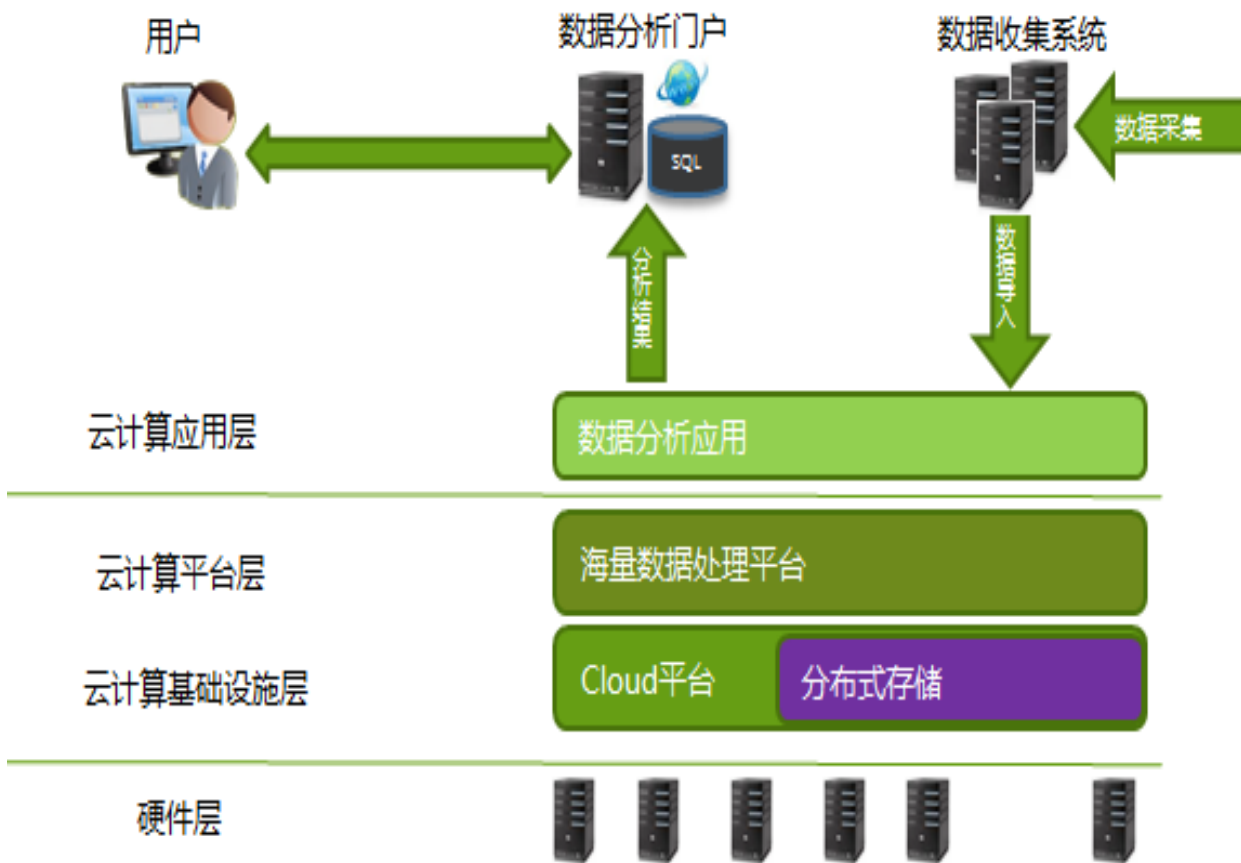
3 某运营商实战案例



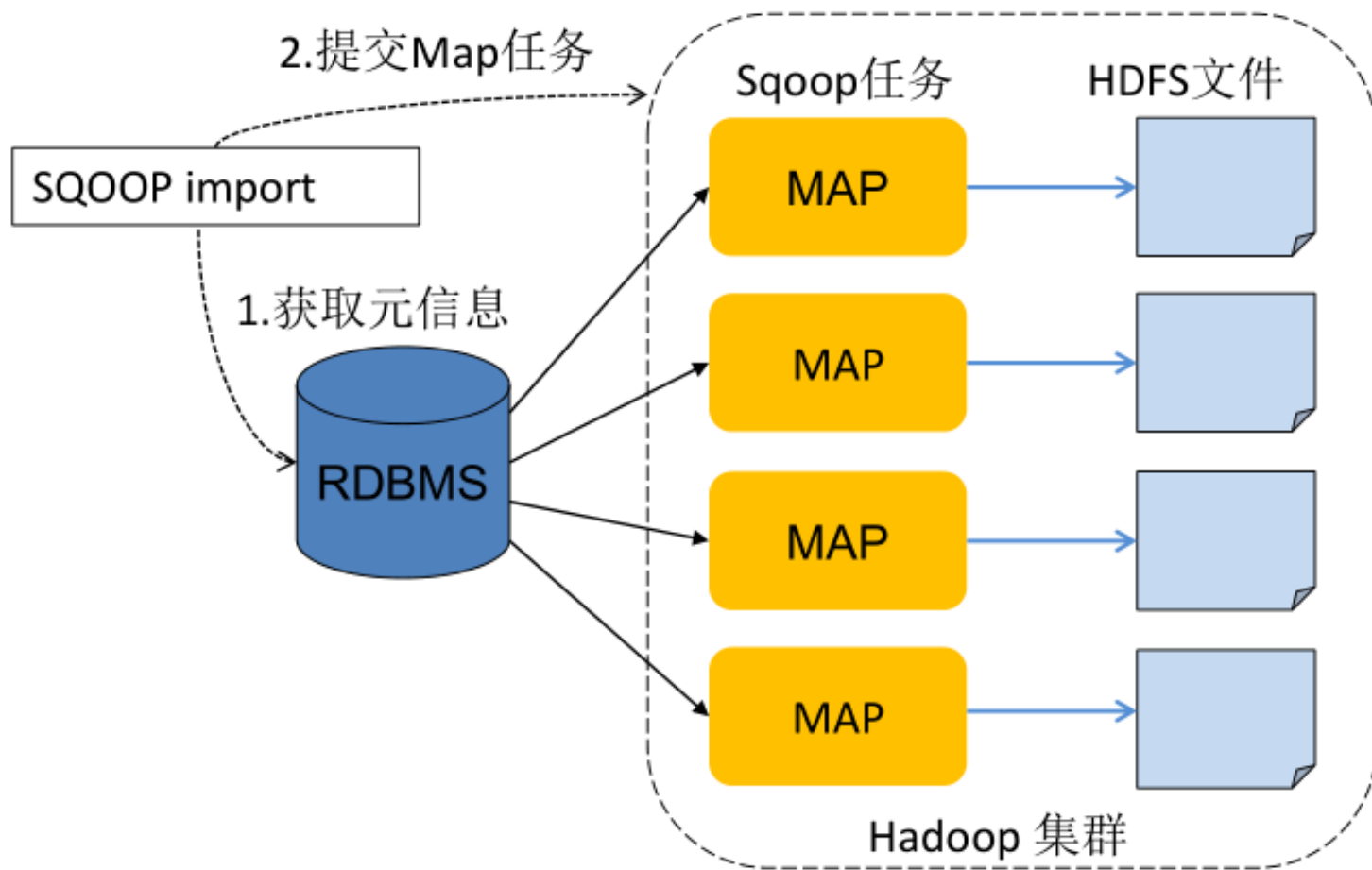
传统技术：脚本方式



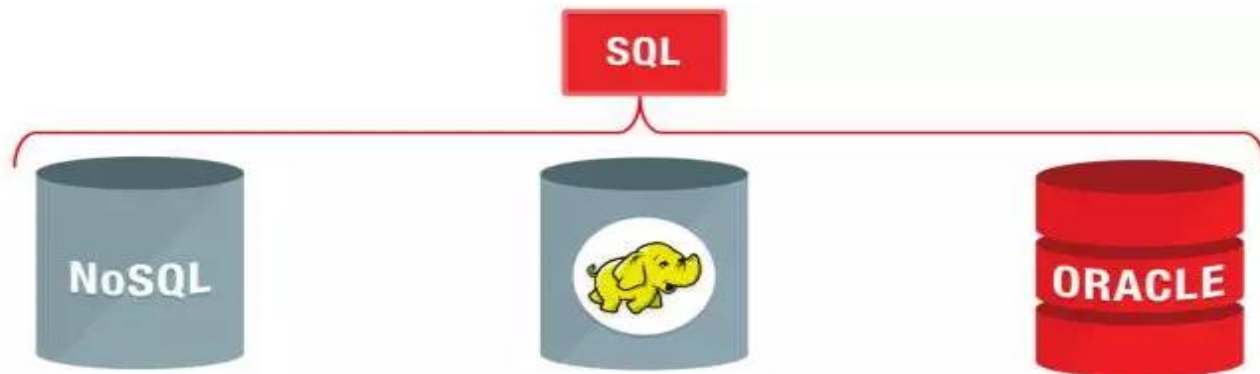
开源技术一：FLUME日志采集技术



开源技术二：Sqoop异构实时同步技术

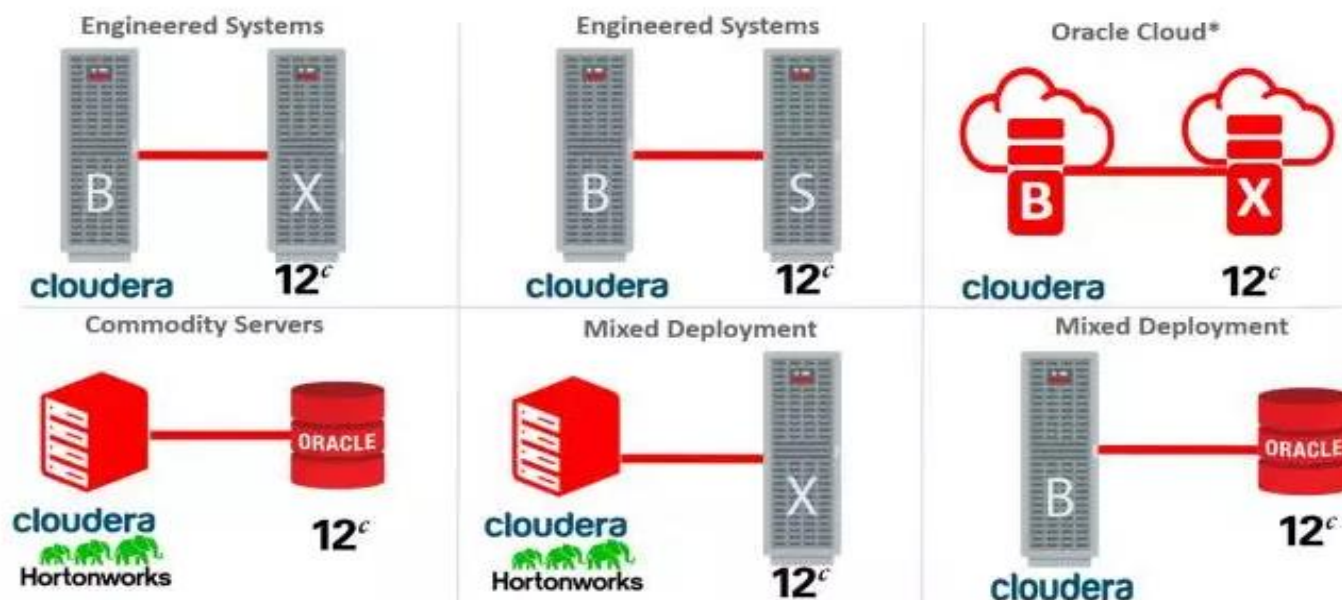


新型同步技术一：oracle 12c BDS

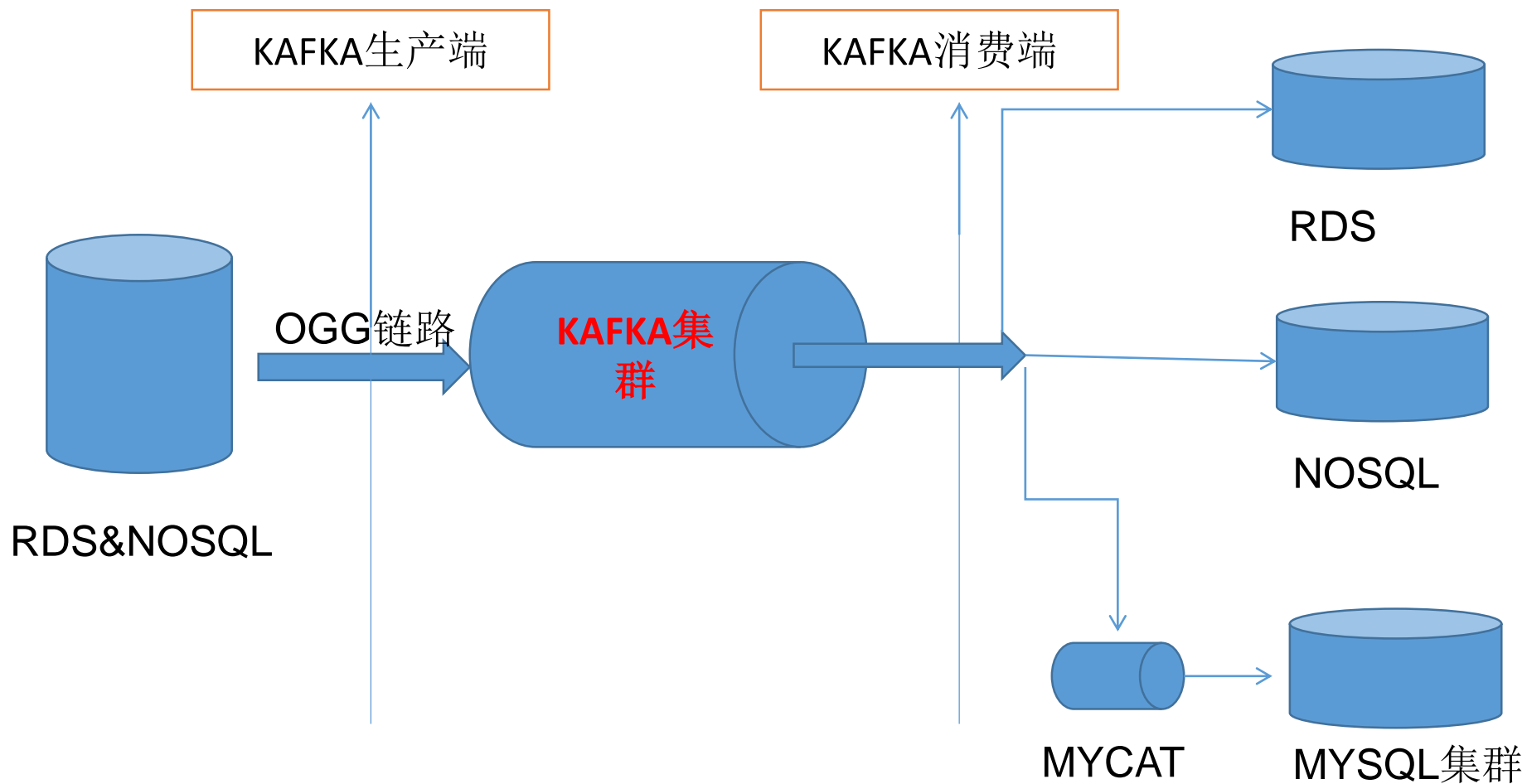


Oracle_Hive

Oracle_HDFS



新型同步技术二：Goldengate+Kafka



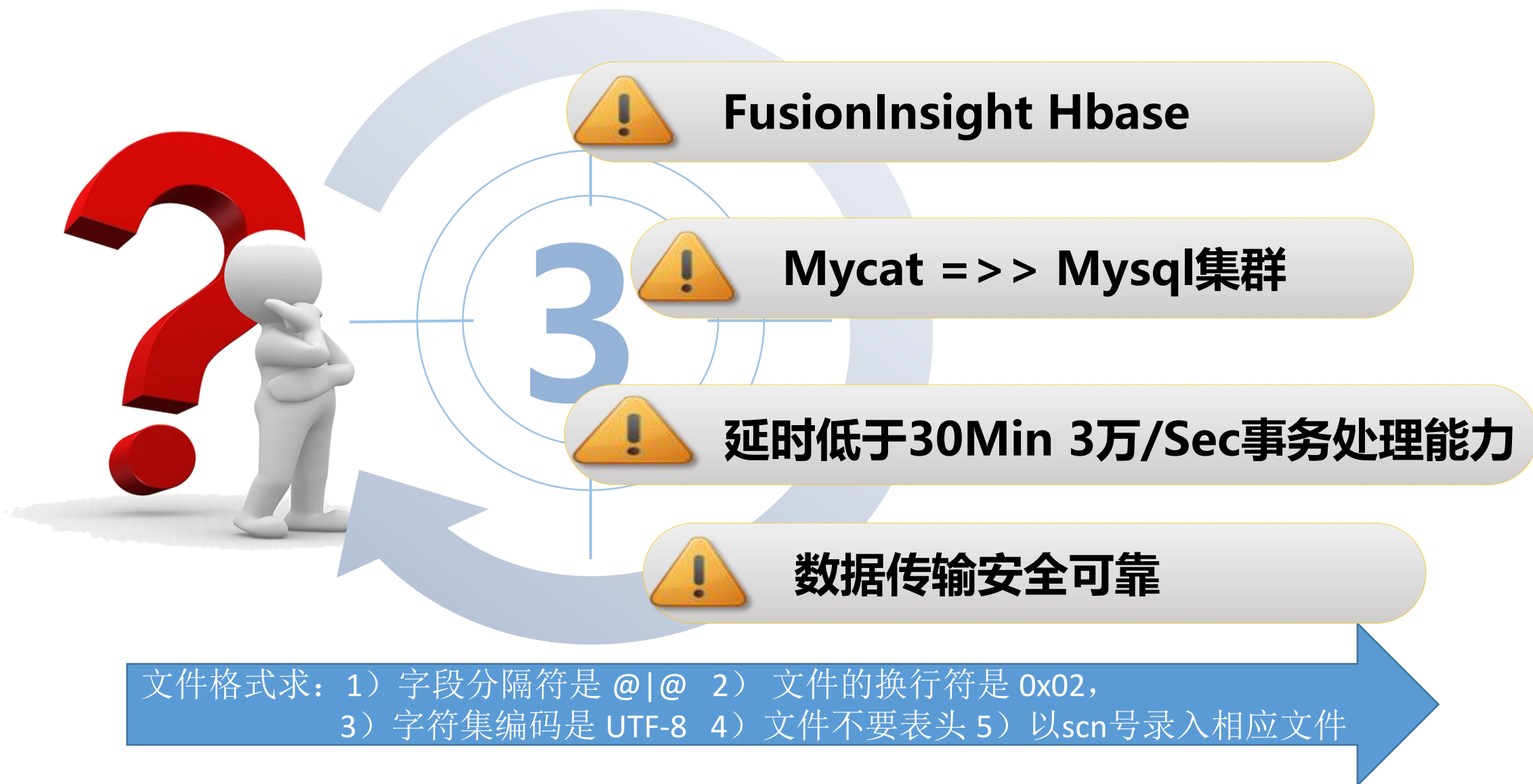
1 大数据概述

2 大数据时代下的 数据同步技术

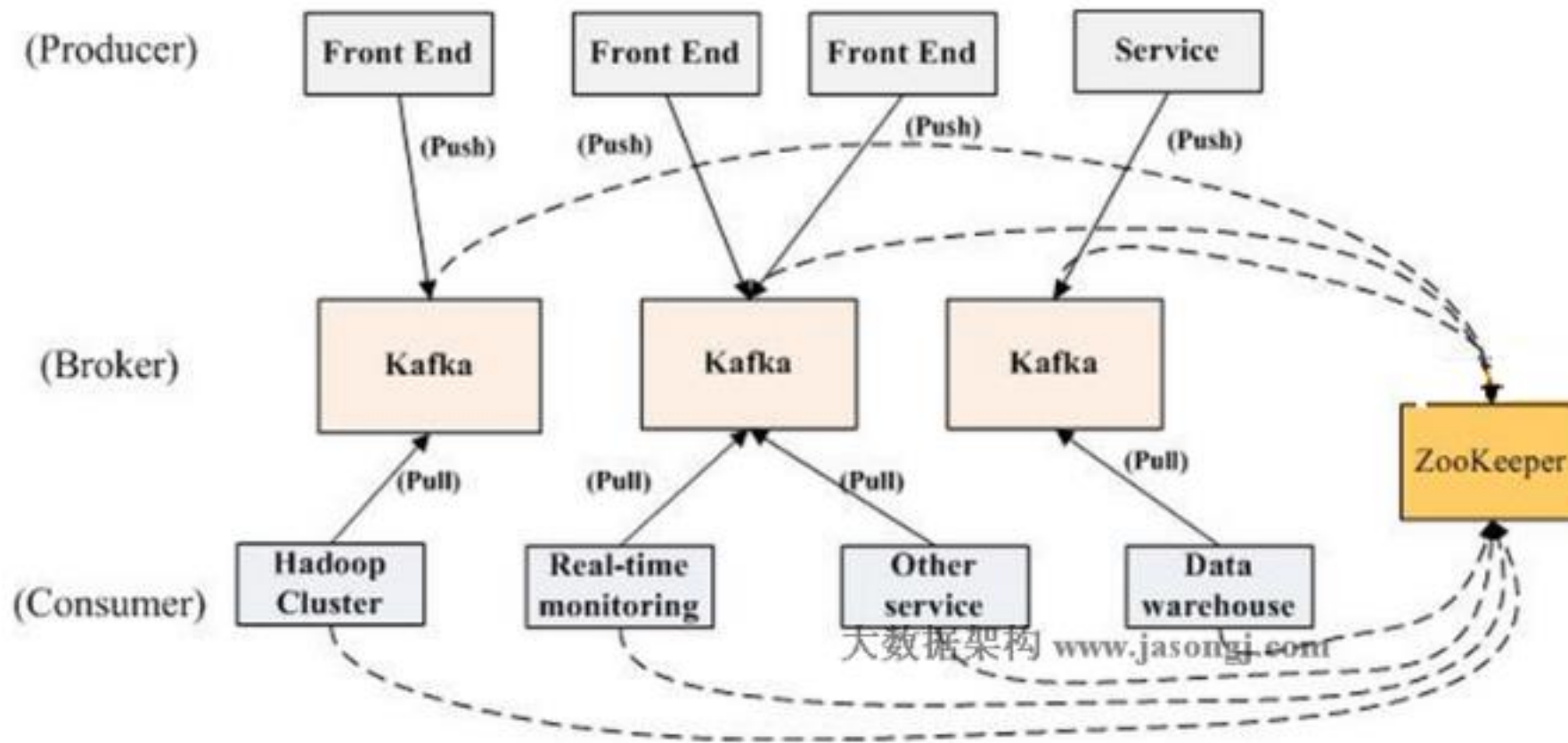
3 某运营商实战案例



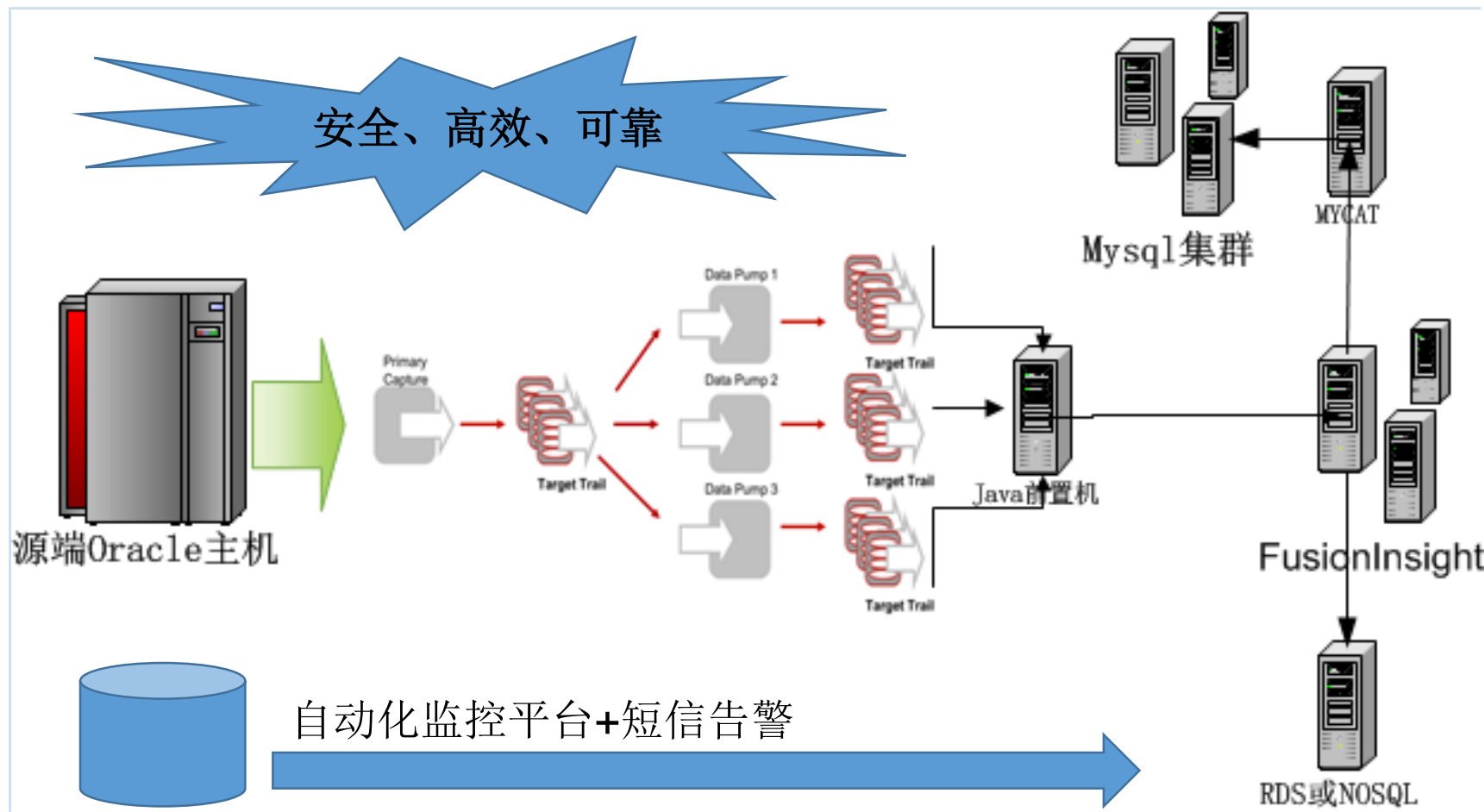
客户需求-----How you do?



KAFKA基本架构



某运营商集成同步架构



Initload(初始化数据) + CDC (增量数据)

项目难点及解决方法

1 方案架构

→ 根据客户需求，对可能满足要求的方法进行论证测试



确立**OGG+Kafka**架构

2 INITLOAD+CDC

→ 确立能生成指定格式并支持ORACLE的工具,对spool,写脚本等进行测试



确立 sqlldr + scn+接口开发

3 数据安全高效

→ 与厂商多次测试不同接口及SSL协议支持



OGG本身压缩加密+华为客户端数字证书KEY认证模式



The logo for DBAplus, featuring the letters 'DBA' in red, blue, and orange respectively, followed by 'plus' in green. A thin white horizontal line is positioned below the letters.

DBAplus

www.dbaplus.cn

THANK YOU!