



苏宁物流在实时大数据的最佳实践

苏宁易购-伍冠军



一、苏宁物流天眼系统介绍

二、苏宁物流实时技术架构演进

三、苏宁物流在实时大数据的最佳实践

四、性能调优分析和经验总结

苏宁物流天眼系统简介

物流天眼系统是苏宁物流集团面向物流及售后领域的**实时数据监控**、**多维度数据分析**和**大数据服务**的可视化平台。

实现对物流订单的**全链路异常实时跟踪**，物流作业的**全网实时监控**，仓储作业和车辆运输的**实时监控**。



天眼

数据由你掌握



绩效管理

目标管理导向，数据客观呈现，物流、售后管理全面提升



运营分析

多样分析方式，深入数据分析，管理决策有效支撑



数据报表

分类数据统计，多维自助分析，数据利用准确有效



数据服务

异常数据解析，开放对外服务，数据管理有理有据

全程实时监控22个系统及作业环节141个场景异常

易购下单



仓库出货



中转运输



配送销单



全流程异常环节

- 系统堵单
- 网络异常
- 接口不通
- 验证失败
- 系统宕机
-

系统

- 有单无货
- 作业延迟
- 商品破损
- 多货少货
- 料箱短缺
- 商品串码

作业

- 装箱不及时
- 错派漏派
- 装车不及时
- 包装破损
- 出站不及时
- 入站不及时

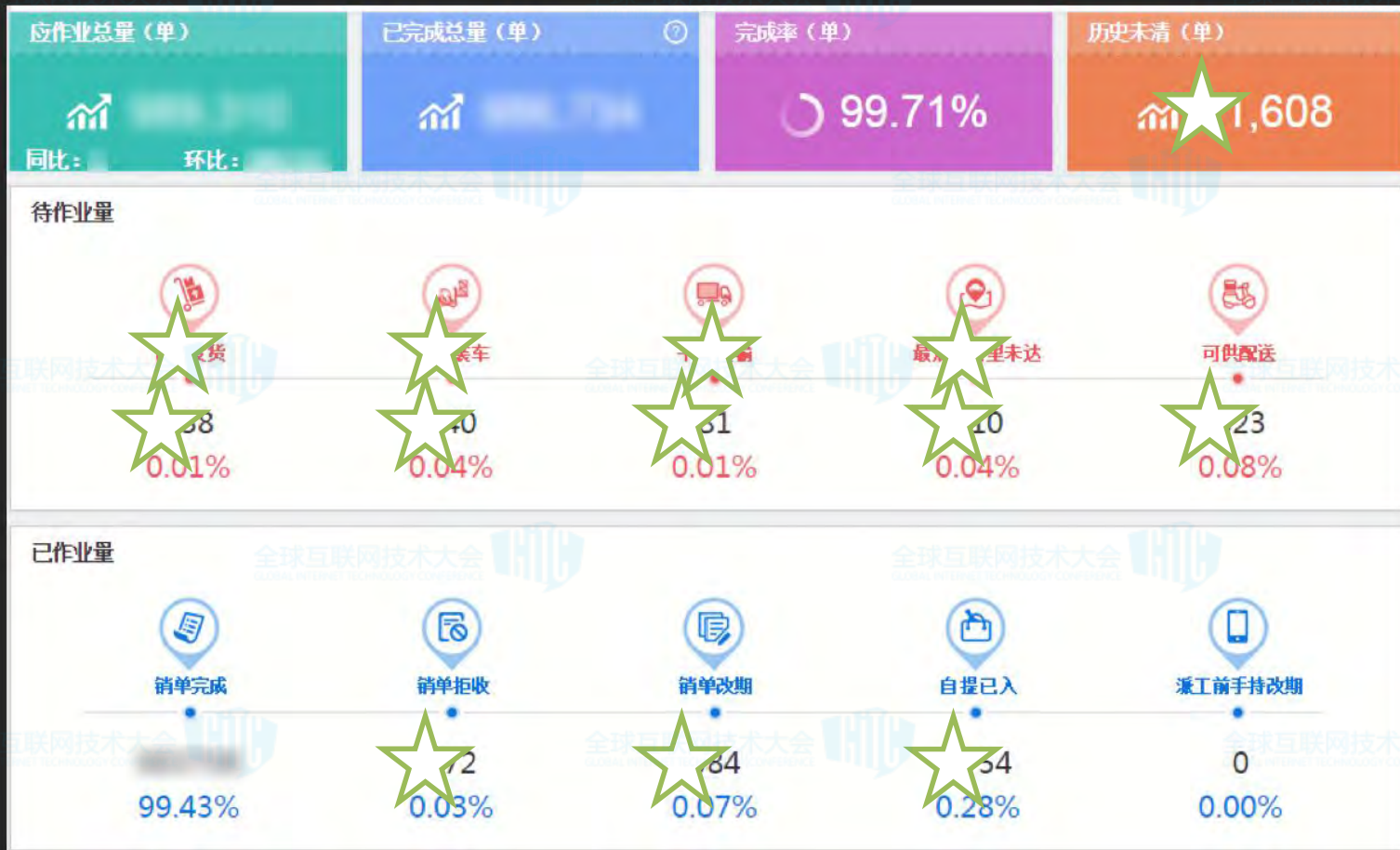
- 月台资源不足
-

- 卸货人力不足
- 开箱少货
- 开箱破损
- 跨区配送
- 车辆不足
- 货未到齐

-

物流作业的全网实时监控

苏宁物流



苏宁物流和天天快递当日订单量实时展示

苏宁物流



车辆线路的实时监控



全网妥投
率相比上
线前上升
20%

作业异常
比例下降
2.8%。

客诉率下
降0.5%

大数据实时监控和分析如何实现？

物流天眼系统整合了物流过程中的**订单、仓储、运输、配送、售后**等全流程的业务数据，随着苏宁易购线上订单的**急速增长**，如何高效的处理和分析大量复杂的物流业务数据并进行**实时监控和分析**？



一、苏宁物流天眼系统介绍

二、苏宁物流实时技术架构演进

三、苏宁物流在实时大数据的最佳实践

四、性能调优分析和经验总结

苏宁物流实时技术架构演进

Evolution of the technical architecture

1

2012-2014
传统架构

2

2015-2016
架构升级

3

2017
稳定快速

1

2014 传统架构

苏宁物流

前台展示: Cognos

业务应用

苏宁物流
实时报表

苏宁售后
实时报表

苏宁物流
离线报表

苏宁售后
离线报表

大数据的
服务

分析计算

DB2 sql

IBM datastage

Hql

数据存储

DB2

增量存储

Hive

hdfs

数据采集

ESB MQ

RSF实时同步

sqoop

数据来源

订单

仓储

运输

配送

售后

2

2015

架构升级 (去IOE)

前台展示: Cognos → Echarts

业务应用

苏宁物流
实时报表苏宁售后
实时报表实时包裹
推荐售后自动化
派工天天快递
实时报表

分析计算

DB2 sql

datastage

spark
streaming

Hql

Spark sql

数据存储

DB2

postgresql

redis

Hive

hbase

数据采集

ESB MQ

kafka

RSF同步

flume

数据来源

订单

仓库

运输

配送

售后

3

2017 稳定快速

苏宁物流





一、苏宁物流天眼系统介绍

二、苏宁物流实时技术架构演进

三、苏宁物流在实时大数据的最佳实践

四、性能调优分析和经验总结

苏宁物流在实时大数据的最佳实践

1

全网实时监控报表的整体迁移（架构升级）

2

物流订单双11大促实时监控

3

天天快递数据实时查询下载

4

实时箱包推荐

全网实时监控报表的整体迁移（架构升级）

苏宁物流

实现目标

实时频率
从2小时提
升到1分钟

大数据的
处理能力

上百张报
表的同时
迁移

稳定性和
扩展性高

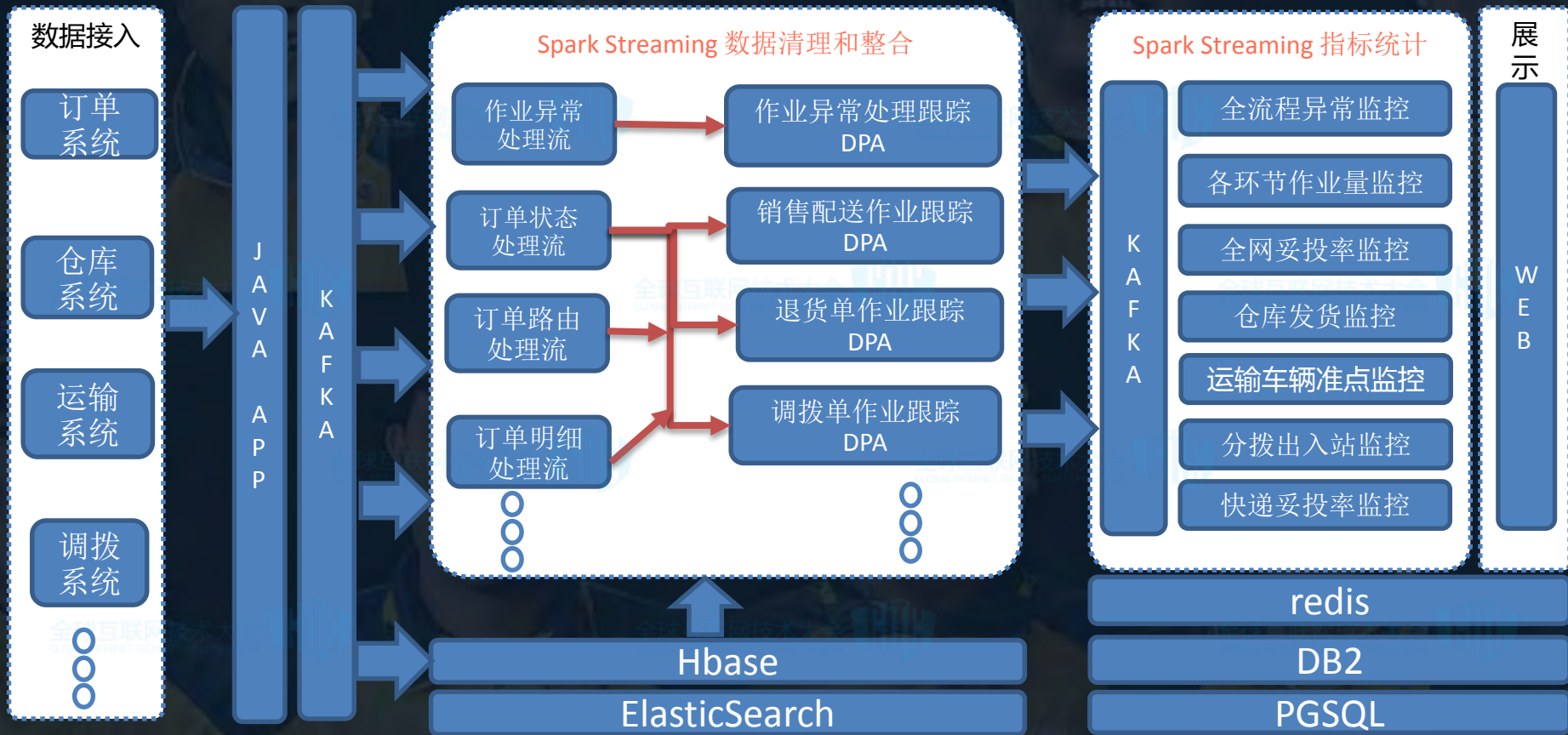
实现前提

迁移
成本
要低

实现
周期
要短

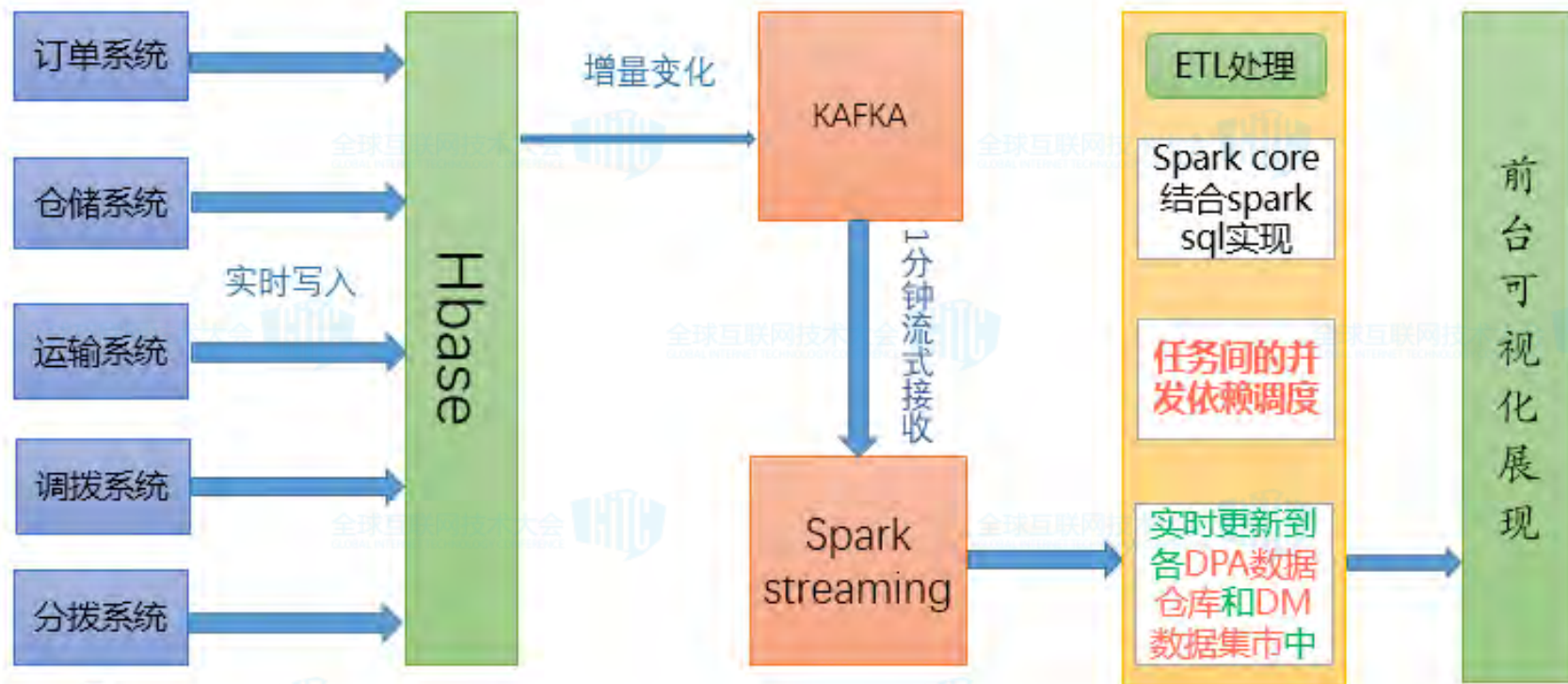
数据整合类

& 指标统计类



全网实时监控报表的整体迁移-技术架构设计

苏宁物流



全网实时监控报表的整体迁移-效果展现

苏宁物流

一个spark streaming任务搞定所有迁移的核心报表

每天实时更新上百张hbase的表，
上百亿的数据，
上百T的空间

Hbase的qps峰值

360
万条/秒

实时性能从2个小时提高到了1分钟以内。

1分钟处理的数据量达到上千万，
压测性能满足双11大促要求

Spark的处理峰值

100
万条/秒

物流订单双11大促实时监控-订单总量

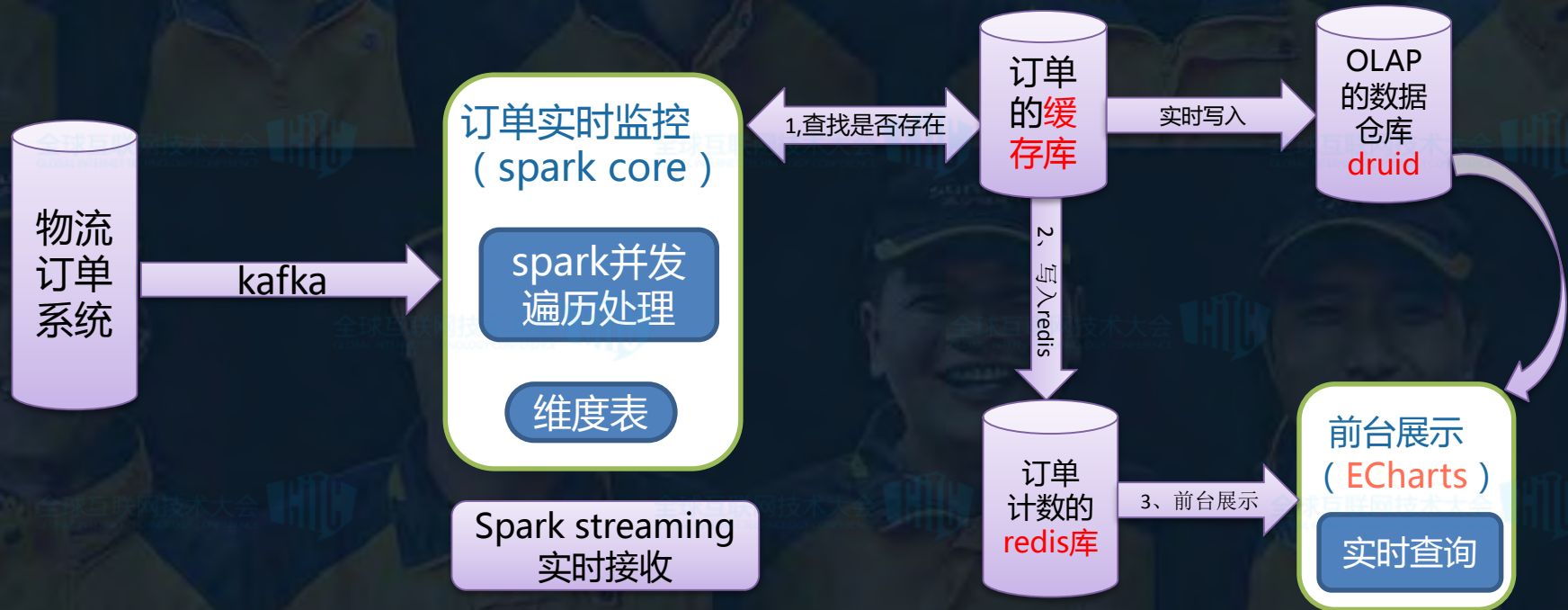
苏宁物流



物流订单双11大促实时监控-技术架构实现

苏宁物流

Kafka+Spark streaming+hbase+redis+druid



天天快递数据实时查询下载-实现目标

苏宁物流

支持上万人同时
进行实时查
询下载

全国分拨中心
和上十万个快
递点全覆盖

实时下载的峰值

80
万条/秒

可同时支持单
个快递点上百
万未清数据的
下载

下载时长控制
在30秒内，数
据少时达到毫
秒级

天天快递数据实时查询下载-技术架构实现

苏宁易购

Kafka+Spark streaming+hbase+elasticsearch(组合查询下载)



实时箱包
推荐

售后实时
智能派工

库存智能
ABC

物流实时
分拨





一、苏宁物流天眼系统介绍

二、苏宁物流实时技术架构演进

三、苏宁物流在实时大数据的最佳实践

四、性能调优分析和经验总结

1

Spark任务性能调优

2

经验总结

1: 开发调优

2: 资源调优

3: 数据倾斜调优

原则1：对多次使用的RDD或DataFrame进行持久化，共享同一个RDD

原则2：1) 经过filter算子过后使用coalesce优化分区数量。
2) 分区少并且数据量大是通过repartition重分区增大并发。

原则3：读写hbase或插入数据库时使用foreachPartition代替foreach
并且使用批量读取和插入，可大幅提升读写性能。

原则4：尽量避免shuffle算子，大表关联小表可使用Broadcast+Hashmap的方式

Spark Executor端的内存主要分为三块：

第一块是：分配给shuffle数据的内存，默认占20%；

第二块是：分配给RDD缓存的内存，默认占60%

第三块是：程序执行预留的内存，默认也是20%

Status	Locality Level	Executor ID / Host	Launch Time	Duration	GC Time	Write Time	Shuffle Write Size / Records
RUNNING	PROCESS_LOCAL	19 / slave483-prd3.cnsuning.com	2017/11/12 02:01:35	11 min	0.7 s	5 s	1707.4 MB / 3591837
SUCCESS	PROCESS_LOCAL	26 / slave973-prd3.cnsuning.com	2017/11/12 02:01:35	32 s	0.1 s	0.6 s	54.0 MB / 131958
SUCCESS	PROCESS_LOCAL	1 / slave131-prd3.cnsuning.com	2017/11/12 02:01:35	10 s	0.2 s	0.6 s	4.5 MB / 14077
SUCCESS	PROCESS_LOCAL	6 / slave1075-prd3.cnsuning.com	2017/11/12 02:01:35	11 s	46 ms	0.3 s	20.6 MB / 44114

如何解决上图数据量突增的情况下单个executor内存溢出。

- 1：适当提高`executor-memory`的值，设置6G
- 2：适当提高`spark.shuffle.memoryFraction`的内存比例，设置0.4
- 3：适当提高`num-executors`，提高executor的个数，提高并行度。

如果有driver端的OOM

- 1：适当提高`driver-memory`的值。

1：数据倾斜的现象：绝大多数task执行得都非常快，但个别task执行极慢，并且可能出现内存溢出的情况。

2：如何定位数据倾斜的代码：数据倾斜只会发生在shuffle过程中。可以通过Spark Web UI清晰的定位到哪个task执行时出现了数据倾斜

查看导致数据倾斜Key的分布情况：可以先使用sample抽样的方式，减少运算量，然后使用countByKey算子统计出每个key出现的次数，降序即可查看到哪些Key出现了数据倾斜。

3：数据倾斜的解决方案：

- 1) 过滤少量导致倾斜的key (key为空值的情况，不影响结果的情况下)
- 2) 提高shuffle操作的并行度，增大内存（只能缓解）
- 3) 大表和小表关联可使用map join或Broadcast的方式。
- 4) 采样倾斜key并分拆join操作。

一个优秀的系统离不开以下四大特性

稳定性

扩展性

健壮性

高性能

问题现象

旧交货单：1342281927 （最后一位是0到9的随机数）
主键设计：为了避免写入热点,将交货单反转:7291822431

源端业务系统进行升级改造，交货单发生变化
新交货单：★S0006940886510101（后四位是固定的）

问题原因

此时问题来了，反转后都是以0101开头，写入热点明显。

解决方案

临时方案：OMS开头的交货单,后四位移到最前面后再反转。
永久方案：

问题现象

大促期间,某个时间点**数据量剧增**, kafka堆积告警, spark streaming任务运行超时, 系统压力增大, 超出系统负载时甚至会出现**内存溢出**, **任务失败**的情况。

问题原因

由于spark streaming**单个批次的处理能力是有限的**, 比如压测的性能**5秒处理50万**, 当5秒内接收数据增大到500万的时候, 自然就会出现任务运行超时甚至失败。

解决方案

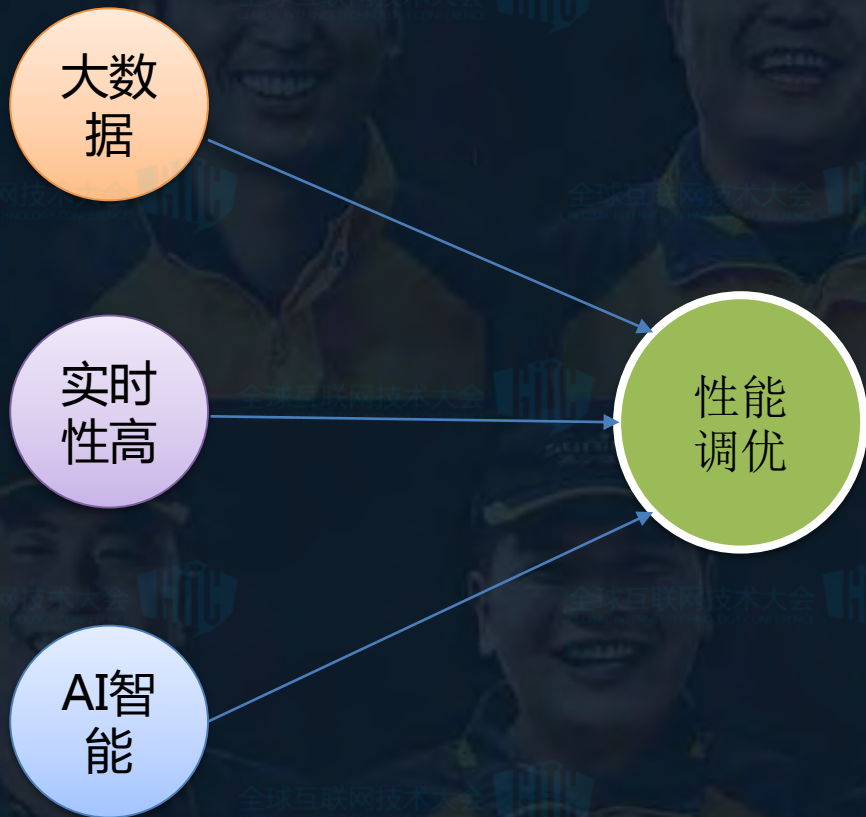
任务限流-参数配置

```
conf.set("spark.streaming.kafka.maxRatePerPartition","1250")  
maxRatePerPartition = (压测QPS目标) / (分区数*单个报文批量数)
```

平台的容错：选择kafka、hbase、spark、pgsql等分布式组件

任务的容错：保证数据不丢失和幂等性，丢失的情况下快速恢复。

架构的容错：架构流程的某个环节出错，有备选替代的方案。



苏宁物流-未来展望

六 数据驱动苏宁物流裂变



招聘大数据实时开发工程师

Spark/Flink
流式计算方向

AI智能
算法方向



联系邮箱: 399611722@qq.com

T 谢谢
HANK YOU!