



星环助力EMS迎战“双十一”

近年来，我国快递行业的规模随着电子商务的井喷出现了快速的扩张。据统计，全国现在已有上百万快递从业人员、十几万营业网点、几十万辆快递货车和一年近百亿件业务量。这些数字还在以每年超过40%的速率增长。天文数字的市场需求给快递公司带来了巨大的商机和前所未有的挑战。

首先，快递行业公司众多，彼此竞争极为激烈，在人工价格上升的趋势下还要维持服务的低价并且盈利，就必须严格控制成本。而每年的“双十一”会给公司的处理能力施加远高于平时的压力。想必各位买家都已经历了今年“双十一”后长达数天乃至一二周的漫长等待。因此，怎样缓解“双十一”的爆仓、避免快件变“慢件”是每个快递公司的难题。

“老问题新办法”，星环科技和华胜天成联手为国内EMS物流公司打造了面向未来的大数据Hadoop系统。依靠该系统强大的实时数据处理能力，物流公司的“双十一”效率明显比往年高很多。

在项目研究阶段，星环科技和华胜天成的开发人员发现快递的效率其实在于流程管理，而当流程全部计算机化后，快递公司面临的难题的答案其实就在自己手中——它就是快递的每一个生产环节都会产生的大量数据。监控这些数据进而对全国各处理中心的收寄和运载能力、出班投递计划做优化调整，公司就能降低成本。分析这些数据来对“双十一”的趋势做出预测，公司就能做好准备应对暴涨的需求。然而，快递生产环节中的数据具有数据量大、类型复杂、结构杂乱的特点，而上层应用对实时性要求很高，传统数据库在这样的情况下捉襟见肘，所以快递公司需要大数据解决方案。

问题与需求：

客户需要搭建一个大数据平台来对它在全国的揽投部、处理中心和集散中心的数据（包括已接收、留存件、已下段、未下段、已投递、未投递、揽收员、地址、已封发、已发运、未发运等等）进行处理。大数据平台需要将ESB（企业生产总



线)流来的数据实时动态加载进数据库,进行处理和统计并且实现实时数据查询。客户原先的技术在每次处理数据时都需要对所有数据进行一次全面ETL(萃取-转置-加载),查询延时在二十分钟以上,完全无法进行实时查询。

星环Hadoop发行版TDH方案实施效果



上图为TDH平台中的所有组件。在快递公司项目中,这些组件中起主要作用的是Transwarp Stream, Transwarp Hyperbase和Transwarp Inceptor。

Transwarp Stream 实时流处理引擎以Spark Streaming为基础提供强大的流处理能力,它拥有:

更强的表达能力: 支持DAG计算模型

丰富的输出方式: HBase, 告警页面, 实时展示页面

广泛的应用场景: 传感器网络处理, 服务监控, 反作弊

Transwarp Hyperbase 实时在线数据处理引擎以Apache HBase为基础,是企业建立高并发的在线业务系统的最佳选择。它拥有:

多种数据类型的支持: 支持结构化、半结构化和非结构化数据

高速处理能力: 延时在数毫秒到数百毫秒级,上亿并发

OLAP和批量统计: 支持高速的OLAP统计和SQL离线批处理

高效的图形计算: 提供图形构造API和专有的高效图算法

Transwarp Inceptor 内存分析引擎提供大数据的高速交互式SQL统计和R语言挖掘。它拥有:

更快的性能: 比Hadoop快10倍到100倍,比MPP快2倍到10倍

更强SQL支持: 兼容Oracle PL/SQL和HiveQL语法

更强分析能力: 支持R语言,提供更多的并行化算法

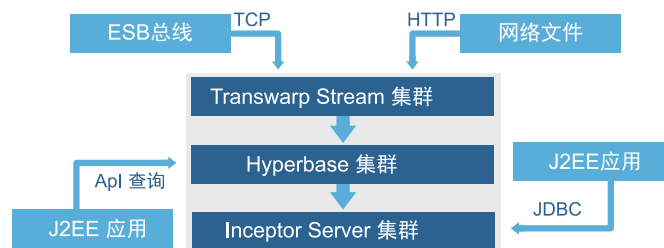
BI和报表工具: 支持Tableau,SAP BO, Oracle OBIEE

超高扩展能力: 线性扩展,支持GB到PB数据的快速处理

超强稳定性: 经过验证的稳定版本,7x24小时不间断运行

这三大组件在快递公司项目中的工作流程为(如下图所示):从ESB总线加载数据到Transwarp Stream 集群中进行处理;处理完毕后存入Transwarp Hyperbase;提供API给上层J2EE应用实时查询数据;也可以提供JDBC接口给上层J2EE应用从基于Hyperbase建立的Inceptor外表中实时查询数据。

Transwarp Data Hub (TDH) 典型应用案例



TDH通过此项目在国内乃至全球的快递行业中首次用大数据Hadoop发行版实现了生产环节数据的实时监控。效果如下：

数据实现实时导入：数据从ESB总线上流入、TDH处理完毕到查询结果显示的整个过程不超过1秒，导入平均速度为每秒1700条，仅需要用到TDH最高性能的1/500。

数据处理时间迅速：Transwarp Stream对单条数据的处理在毫秒级完成，平均一条记录从处理到进入Transwarp Hyperbase只需要1-2毫秒左右。

数据查询简单高效：合作方的J2EE应用要求的秒级查询能由TDH平台轻松胜任。而TDH提供的JavaAPI和JDBC接口可以非常简单地与现有系统进行集成从而实现数据查询，对于现有的企业级开发环境极其友好，应用迁移难度很小，快递公司的工作人员可以无学习成本地上手使用。

TDH可毫无延迟地反映快递公司遍布全国的海量业务的最真实情况，给了快递公司简单易用的工具来对业务的每个环节实时监控，使得快递公司在海量的快递业务中都可以快速精准地发现问题，如快件的积压、遗失、破损等，进而提高服务质量。TDH也可以根据最新的生产数据帮助快递公司随时调整优化投递计划，为公司降低成本。TDH更可以帮助快递公司分析趋势，为“双十一”和其他节假日带来的压力提前做好准备。

星环信息科技（上海）有限公司

🏠 地址：上海市徐汇区桂平路481号18幢3层301室（漕河泾新兴技术开发区）

✉ 邮编：200233 ☎ 电话：4008 079 976

🌐 网址：www.transwarp.io

TRANSWARP
DATA HUB