

Informatica 助力华海财险搭建 ODS 大数据平台



Informatica ETL 工具从口碑方面来说,周边的一些子公司、各同行企业使用后都认为比较好,华海财险在部署半年后也有同样的感受——性能和效率都有保障,而且能够扩展,可以兼顾未来三到五年的需求变化,这是一个可满足长远发展考虑的工具。

—— 华海财产保险股份有限公司数据管理部
谢德万处长

概述

华海财产保险股份有限公司(以下简称华海财险)采用Informatica ETL数据集成工具PowerCenter建立了统一的原始业务数据汇集区,可以将公司成立一年多以来到今后五年的业务数据在此统一采集、抽取和清洗。搭建数据仓库后,再建设统计分析应用,充分挖掘各业务系统的数据资源价值,从而促进业务的发展及创新。

公司及项目简介

华海财险是一家理念新、领域新的新公司,它以服务海洋经济为切入点,以互联网思维为先导,致力于成为一家具有海洋和互联网特色的全国性、综合性财产保险公司,华海财产保险股份有限公司的主要产品有车险、财产险、意外险、责任险和再保险等。

经过持续快速发展,华海财险目前已建成移动、网站、传统出单等涵盖多个领域的信息化系统。但随着业务的发展,整个公司对于各业务系统统一管理、运营监控以及各系统业务交叉合作的需求越来越迫切,因此,如何利用最新的技术在满足监管法规的前提下,为各系统业务的综合管理及紧密合作提供支持成为必须面对的问题。

数据驱动业务,华海财险的ODS大数据平台项目正是基于各系统业务数据的基础上创立的项目,旨在通过各系统数据信息的集中融合及合理共享为业务管理层提供集中统一的管理及决策支持信息,充分

挖掘各业务系统的数据资源价值,促进业务的创新和发展。在建设基础数据和数据共享平台时,华海财险采用Informatica ETL数据集成工具PowerCenter实现了不同系统中数据的统一采集、抽取,建立数据仓库,方便的实现了系统之间的数据对接。

面临挑战:建立数据仓库,提升数据抽取效率

华海财险的数据应用经过一年多的发展,并没有统一的原始业务数据汇集区,只有简单的基于少数核心业务系统的报表系统;在此基础上,建设了少量的统计分析应用。

各系统数据应用逻辑上相互隔离,各自独立,系统之间基本上不存在数据共享或整合。在物理上,各系统数据应用分布在不同物理位置及不同主机上,系统之间基本上不存在空间、线路、主机或存储的共用或共享。

原来负责数据抽取功能的ETL工具在性能等方面并不能满足发展所需,迫切需要性能、效率更高的ETL工具搭建数据仓库,来提升数据抽取、转换和加载性能,为下一步的大数据应用打下良好的基础,进而更好地为业务服务。

客户收益:实现了数据的不间断、自动化抽取

从2016年4月项目立项开始,华海财险经过需求调研、模型设计、ETL实施、报表



业务需求:

- 只有业务数据,没有统一的原始业务数据汇集区,无法集中采集和存储各部门数据。
- 加强数据管理能力,为进一步数据挖掘和数据分析做好数据准备。

挑战:

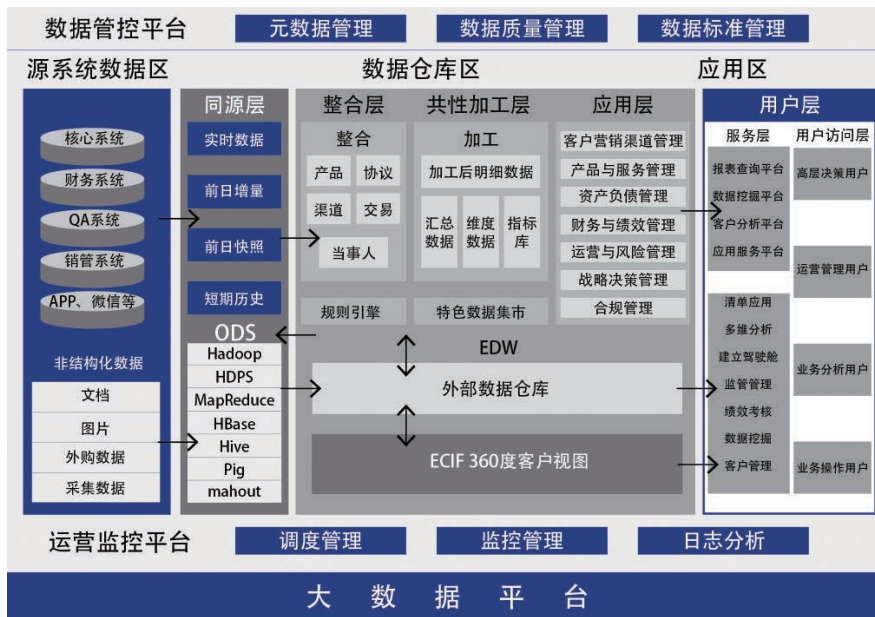
- 公司新成立,尚未有统一的数据架构设计,同时没有专业数据架构团队和数据管理标准,数据的创建和定义随意较大。
- 数据存放在不同的应用系统中,尚处于被动数据质量管理阶段,数据重复,较难以有效维持数据一致性,同一数据存在多种定义。
- 业务系统众多,需要确保ETL工具与各业务系统顺利对接。

产品与方案:

- Informatica PowerCenter

收益:

- 构建起一个有效的数据仓库,并在此基础上强化数据管理,促进各系统业务联动发展和创新
- 自动化数据处理流程提升数据抽取效率,负载均衡保障数据抽取不会中断
- 降低运维成本,运行稳定,几乎不需要人工维护



定制,再到测试和上线部署,花费了近半年的时间。目前,华海财险已成功部署了 ODS 大数据平台。

ODS 大数据平台项目是一个成功的智能商业平台的第一步,建立了统一的数据仓库,只有如此做,才能将业务系统和报表的数据源分开,使报表制作不影响业务系统的性能,且减少报表制作人员学习不同系统数据结构的难度。

PowerCenter 数据集成平台帮助华海财险实现了数据的统一集中管理,建立了规范的数据处理流程,从复杂的数据源结构中,实现了对各种结构化数据的自动化清理、抽取、转换,节约了运维管理成本,保障了数据质量,为系统集成和各个应用系统之间的数据共享提供基础。

- 采用先进的数据库技术,通过 Informatica PowerCenter 整合了公司不同数据来源的分散业务数据,构建起一个有效的数据交换枢纽和完整数据统一集中管理环境。
- 在数据流转过过程中,实现了各个环节

的自动化处理,确保业务系统的稳定安全,减少运维压力。即使在远端数据结构发生改变时,也无需人工运维。

- 可将数据路径追溯回最初来源或其最终目标,随时查阅相关的详细的数据和文档。

解决方案:两台服务器实现负载均衡,数据抽取完全自动化

在选择数据集成方案时,华海财险综合比较、测试了多家 ETL 方面的通用性平台。一方面,Informatica ETL 工具的口碑较好,技术成熟、市场占有率极高,很多保险业的同行都选择其解决方案,华海财险对它早有耳闻;另一方面,华海财险也对各个 ETL 工具进行了测试,Informatica PowerCenter 在性能、效率等方面脱颖而出。因此华海财险便毅然选择了 PowerCenter 打造其 ODS 大数据平台。

为了确保 ODS 大数据平台的稳定与热切换,该项目创新地实现了两台服务器之间的负载均衡,即任何一个节点出现问题或

宕机,另一台机器可以随时替代,相比传统一台服务器为主、另一台服务器作为备用的方式,该模式实现了真正的无缝热切换,连等待备份服务器开机、启动的时间都省去了。

在具体实施过程中,还曾出现原端数据结构改变时,数据抽取中断的问题,最后在 Informatica 及合作伙伴的合作下,通过相应的定制化开发,实现了前端的任何一个系统做系统架构的调整都不会影响数据抽取,整个过程完全自动化,无需运维人员参与,达到了节省人力成本的效果。

目前,华海财险对车险核心系统——车承保、车理赔、收付、财务等各个系统,进行全数据链的抽取,实现了完整的全公司数据统一集中管理环境。针对不同的数据建立起不同的数据管理模型,如一些不经常更新的数据,采用全量的数据抽取;对于保单、理赔等日常业务数据,采用增量的数据抽取方式;对于保费等实时信息,则实现了数据的实时抽取与展现。

PowerCenter 和业务系统的接口的耦合程度低,独立性强,允许现有系统无需改动或者能够将改动影响降到最低程度,即使面对最为复杂的数据环境也能及时、准确、经济有效地完成数据共享任务,完全保证了业务连续性。谢德万处长认为目前还处于打造框架的基础建设阶段,完成数据仓库的标准配置后,这一部署能够至少满足公司三到五年的发展,节省硬件成本,并降低与数据停机有关的成本和风险。

未来,随着用户规模和业务的进一步扩大,华海财险还计划集成更多的应用和数据源,逐步实现真正的大数据分析,在经营决策、指导业务、人力分析等方面发挥更大的价值,甚至通过数据深度挖掘、网销数据采集、关联分析实现精准营销,提高产能。谢德万处长表示这些中远期的目标将会逐步实现。



全球总部

地址: 美国加利福尼亚州红木城 Seaport 大街 2100 号
邮编: 94063 电话: 650.385.5000 传真: 650.385.5500

中国总部

地址: 北京市朝阳区建国门外大街乙十二号双子座大厦西塔 EF 层 06 室
邮编: 100022 电话: +86-10-5675-2000 传真: +86-10-5675-2030