# 行业案例

Teradata是全球最早开始实施数据平台咨询与实施并获得成功的专业公司。三十余年来，已经在全球范围咨询和实施了一千多个大型数据平台系统，其中数据容量在PB级以上的就有数百家。

众多的咨询与实施成功案例，是确保项目顺利实施的基本条件。正因为这样，我们深信可以为上海期货交易所提供业界领先富有前瞻性的和切实落地可行的数据平台咨询方案。

在本项目应答书中，Teradata为上海期货交易所介绍了Teradata在国内众多的金融业数据平台咨询与实施案例，覆盖了以下各方面：国内金融机构使用Teradata构建数据平台的情况、国内金融机构数据治理咨询案例和Teradata统一大数据案例。

Teradata将为上海期货交易所在未来的发展道路上提供更多的数据平台规划和实施的经验，成为上海期货交易所与国内国际同业先进经验交流的桥梁和纽带。

## 证券业案例

### 中国证券登记结算公司案例

|  |  |
| --- | --- |
| **客户案例名称** | 中国证券登记结算公司 |
| **案例规模** | 116T（用户已使用）／260T（最大空间）  数百个分析应用 |
| **实施范围** | 详见案例说明部份 |
| **实施时间** | 2000年至今 |
| **数据仓库+大数据技术** | 数据仓库MPP |
| **数据治理咨询** | **无** |
| **系统架构** | EDW数据架构 |
| **主要技术特点** | 双系统互为灾备  MPP高性能并行处理  证券行业数据模型落地实施  Teradata数据平台咨询与实施方法论指导 |
| **现状** | 持续运行中 |
| **主要参与人员** | 钱某强、朱某翔、戴某华等 |
| **经验** | 证券行业咨询规划与实施经验，对数据源的理解、模型建设以及行业应用的快速构建 |

中国证券登记结算有限公司依据《中华人民共和国证券法》和《中华人民共和国公司法》组建。公司总资本为人民币12亿元，上海、深圳证券交易所是公司的两个股东，各持50％的股份。公司总部设在北京，下设上海、深圳两个分公司。中国证监会是公司的主管部门。

2001年3月30日，按照《证券法》关于证券登记结算集中统一运营的要求，经国务院同意，中国证监会批准，中国结算组建成立。同年9月，中国结算上海、深圳分公司正式成立。从2001年10月1日起，中国结算承接了原来隶属于上海和深圳证券交易所的全部登记结算业务，标志着全国集中统一的证券登记结算体制的组织架构已经基本形成。

公司的宗旨是，建立一个符合规范化、市场化和国际化要求，具有开放性、拓展性特点，有效防范市场风险和提高市场效率，能够更好地为中国证券市场未来发展服务的集中统一的证券登记结算体系。

按照《证券法》和《证券登记结算管理办法》的相关规定，中国结算履行下列职能：

* 证券账户、结算账户的设立和管理；
* 证券的存管和过户；
* 证券持有人名册登记及权益登记；
* 证券和资金的清算交收及相关管理；
* 受发行人的委托派发证券权益；
* 依法提供与证券登记结算业务有关的查询、信息、咨询和培训服务；
* 中国证监会批准的其他业务。

#### 项目背景

* **项目背景**

中国证券登记结算有限责任公司是我国证券交易市场唯一的证券登记结算机构，公司宗旨是建立一个符合规范化、市场化和国际化要求，具有开放性、拓展性特点，有效防范市场风险和提高市场效率，能够更好地为我国证券市场未来发展服务的集中统一的证券登记结算体系。围绕上述公司宗旨，公司把建设一个先进的、统一的数据处理中心和信息应用服务体系作为IT建设的重要任务。

随着证券市场的发展，证券监管机构、公司业务部门、证券市场对决策支持、风险防范和市场服务的需求越来越迫切，通常，这些服务需求都需要基于比较长期的历史数据，并且对数据的全面性和完整性的要求也很高，需要对沪、深两个证券市场的数据进行深入比较和分析。这些都需要建立一个跨沪、深两个证券市场的、全行业统一的、完整的、权威的、高效实用的数据仓库系统。

中国结算数据仓库系统将采集整理来自上海、深圳证券市场的登记结算业务基础数据，以及来自证券公司的客户保证金明细数据，并在其上开发综合应用，以满足监管机构、公司各业务部门及市场参与主体的各类数据需求。通过预定义报表、关键指标监控、即席查询和专题分析等多种形式，面向监管机构和公司管理部门，实现决策支持、风险防范等目标；有效利用数据资源，实现数据信息共享，通过网上自助服务和数据服务平台为证券公司、上市公司、投资人等提供信息服务。

* **发展历程**

中国结算数据仓库系统采用分步实施的手段进行建设，各个阶段具体实施的内容如下：

* 项目一期实施重点是搭建集中的、一致的、可扩展的数据仓库平台；建立以参与人为中心、基于详细数据、符合第三范式（3NF）、反映中国证券登记结算整体市场业务规则的逻辑数据模型；实现中国结算数据的集中存储；完成对业务术语及其定义的统一化、清算路径的统一化、市场参与主体编码的统一化和业务统计口径的统一化；实现基础信息查询和日常固定报表应用，向公司总部领导与各职能部门提供日常业务查询统计和监管报表下载。
* 项目二期实施重点从基础设施建设转移到应用建设上。在应用建设方面，在一期基础信息查询和日常固定报表应用的基础上，加强了信息查询功能，包括向中国证监会中央数据分析系统按照日、周、月、季、半年、年不同时点自动报送数据；向公司总部提供关键指标监控；为公司各职能部门提供即席查询和专题分析；完善面向上市公司的股东名册服务；建立面向结算参与人的登记结算信息查询服务；通过中国结算网站向投资人提供证券持有查询服务；实现对股东大会网上投票服务。在数据建设方面，将中国结算开放式基金TA系统的相关数据集成到数据仓库系统中。
* 项目三期实施重点主要是应用增强和市场服务，包括证券公司客户保证金数据报送系统建设；结算参与人风险分析系统建设；结算参与人数据服务系统建设；证监会机构监管综合信息系统配合建设；开放式基金系统管理人服务系统的建设；两市互查服务等。在数据建设方面，将证券公司客户保证金数据和上海历史数据集成到数据仓库系统中。
* 经过三期项目的持续发展，数据品种不断增加，应用功能逐渐丰富，系统价值持续提升。

到了2012年，中国证券登记结算有限责任公司数据仓库系统已初步成为一个数据全面、功能丰富、性能卓越的综合信息服务平台。数据仓库系统已发展到企业级数据仓库系统建设阶段，伴随着入仓数据以及应用的快速扩张，数据仓库系统的数据量以及用户访问量也飞速发展。EDW原有的空间性能、稳定性已经很难适应后续的发展，为保障EDW的稳定运行，高效地支撑应用系统的建设，并考虑到对原有应用系统建设的投资保护而实现新老系统之间应用的无缝移植，同时增强信息系统的抗风险能力，

2012年中期，中国证券登记结算有限责任公司决定对EDW生产系统进行升级改造，继续采用Teradata平台和“双系统架构”设计来建设同城备份系统。2013年，中国结算数据仓库系统实现了双系统架构，这是继“上海证券交易所”、“交通银行”、“北京地铁”之后，Teradata在国内的第四例数据仓库双系统成功案例。

#### 项目成果

* + **数据仓库系统的项目成果**
* ETL服务平台

ETL服务平台包括：ETL源数据获取及处理、ETL知识库、ETL Automation 、ETL错误处理服务、ETL监控等。

* 数据归并和整理服务

由于上海、深圳市场登记结算系统是彼此独立的，数据间没有关联关系。当需要统计投资人在整体市场上的信息时，需要从两个系统分别获取数据，再进行整合，无法直接获取。为了建立整个中国证券市场的单一视图，中国结算数据仓库系统进行了数据归并和整理，服务内容主要包括：统一沪深结算参与人编码，建立统一的结算参与人编码和沪深结算参与人编码之间的关系；整理上市公司、沪深证券信息，建立上市公司和沪深证券之间的关系；整理投资人统一编码及相关数据的对应关系等。

* 应用服务平台和信息服务平台

应用服务平台集成了企业应用资源，包括前端界面的整合和使用权限的整合；提供了统一的服务入口、个性化的页面展现，使用户对系统中集成的业务功能操作更方便、快捷。

信息服务平台是信息使用者和信息提供者的桥梁。用户可以通过信息服务平台访问或使用信息；信息服务平台提供与其他各种信息资源的接口，如数据仓库等。信息服务平台作为中国结算数据仓库系统对外提供信息服务的数据访问通道，隔离了内部系统和外部访问，使内部系统免受由于外部直接访问可能造成的危害，提高了安全性。

* 应用功能体系

中国结算数据仓库系统的应用功能分为：指标快报、登记结算信息统计分析系统、结算参与人风险分析系统、开放式基金管理人服务系统、专题分析系统、综合报表系统、对方市场查询服务、证监会机构监管综合信息系统、结算参与人数据服务系统、系统管理等。

* + **数据仓库“双系统架构”的建设成果**

在应用访问控制方面，采用EDW生产主机系统和灾备系统同时运行的“双系统”方案，在双系统方案，由各自独立的应用服务器提供访问接入，各应用的用户访问路由已经事先设定，当EDW生产主机发生故障时以手工切换IP地址方法将用户访问的路由指向灾备系统。

在数据同步方式方面，根据不同应用的容灾需求，确定不同的数据同步方式，多种数据同步方式有可能会被同时采用。

* 数据“双加载”：源数据都同时通过ETL服务器加载到EDW与灾备的两套生产系统中。
* 数据“单加载、再同步”：源数据首先加载到EDW中，然后通过数据同步将EDW中的数据复制到灾备系统中。

在设备选型方面，灾备系统平台与主数据仓库EDW系统架构保持一致，采用MPP海量并行处理体系，满足企业级数据仓库的应用需求，采用数据仓库一体机，具备线性扩展能力；

在双系统架构实现后，当任何一套系统不可用时，都可以将应用访问切换到可用的系统中。对于重要性高且时效性要求高的应用，应做到切换的时间最短，最大程度降低故障对应用访问和数据输出的影响。

* + **2015数据仓库生产系统扩容**

原数据仓库生产系统采用(4+2)6650H，经过3年多的运行，性能和容量都无法满足业务的需要。在性能方面，日间CPU使用率高达75-80%，峰值时段甚至达到85-95%，日间CPU使用率瞬间上冲到100%的情况时有发生，系统资源在支持日间频繁的查询和统计分析任务时已出现严重瓶颈。在容量方面，用户可用容量使用率已经接近100%，通过采取压缩DBC空间、对数据表进行多值压缩、调整PI倾斜表等手段获得部分空间。

解决EDW生产系统性能和容量不足的根本办法是启动扩容方案，即利用新系统代替现有旧系统，使性能获得一定的提升，容量满足近2年内需求。于是中登在2015年6月正式启动了扩容。

经过测算，中登决定采用16节点Teradata 2800作为扩容的新设备，16节点Teradata 2800作为数据仓库EDW生产系统，而替换下来的（4+2）6650作为数据集市生产系统。按照Teradata不同系列的定位，6系列面向以复杂灵活查询为主的应用，有高的磁盘I/O性能；2系列面向CPU密集型应用。而随着计算机技术发展，尤其是CPU性能、内存等技术的发展，使最新型号的2系列平台，在硬件的配置和性能上比几年前的6系列平台有大幅地增强。另外，新版本数据库带来的新特性，使最新型号的2系列平台在软件方面的行列混合存储、混合负载管理、系统管理等方面也大大优于几年前的6系列平台。

16节点Teradata 2800的非压缩数据库总空间达到了260TB，是原有系统的5倍。其它性能指标（如I/O吞吐、内存、CPU性能、网络连接速率）等也有较大的提升，综合评估，16节点Teradata 2800的综合性能是原有系统(4+2)6650H的2-3倍以上。

2015年10月，中登完成了数据仓库生产系统的扩容，16节点Teradata 2800正式成为EDW生产系统。扩容后系统总体性能有很大提高：加载约2倍左右，转换任务在4-5倍”，扩容过程达到了预期的效果。

### 上海证券交易所案例

|  |  |
| --- | --- |
| **客户案例名称** | 上海证券交易所案例 |
| **案例规模** | 涉及40+部门，累计发布1200+应用、数据量42T |
| **实施范围** | 详见案例说明部份 |
| **实施时间** | 2002年至今 |
| **数据仓库+大数据技术** | 数据仓库MPP、Aster |
| **数据治理咨询** | 目前正在规划数据治理咨询项目 |
| **系统架构** | EDW双活数据架构 |
| **主要技术特点** | 数据仓库整体规划与咨询 海量数据存储 并行计算 Teradata金融行业数据模型实施 |
| **现状** | 持续投产运行中 |
| **主要参与人员** | 肖某东、卫某毅、谢某静、陈某翀等 |
| **经验** | 证券行业咨询规划与实施经验， 对数据源的理解、模型建设以及行业应用的快速构建 |

#### 项目背景

作为上海证券交易所(以下简称“上证所”)波士顿咨询规划的新一代项目之一—“新一代信息系统”，也是上证所信息管理的核心项目数据仓库项目（以下简称“该项目”）在交易所领导的关心和支持下，以及全体项目组成员的努力下，经过所领导精心部署和有计划有步骤的安排，建成投产时是当时国内金融行业数据容量最大、数据内容最完整的数据仓库，同时也是亚太区最大的证券业数据仓库项目。

上证所的企业级数据仓库建成了一个涵盖从数据清洗、数据存储、数据整合、基本查询、统计应用、信息服务、数据分析与应用深化、挖掘和预测的企业智能信息平台，具体成功完成目标情况如下：

* 体系架构：目前数据仓库建立了双中心灾备系统，双系统可以互为主/备切换对外提供服务，实现了同城异地双中心运营管理；数据仓库系统包含了数据加载及输出、数据仓库存储、数据仓库门户(Portal)、元数据管理(MDS)、数据仓库管理(DWM)、非结构化数据管理等子系统。
* 数据存储；整个数据仓库的数据涵盖了1990年上证所建所以来的全部历史数据及最新增量数据，范围包括交易核心数据和交易相关数据，类型则包含结构化数据和非结构化数据，密级分类刚包含机密数据、秘密数据、开公数据。
* 信息服务：数据仓库系统目前为上证所仅次于核心交易系统的运营管理系统，面向所内、监管部门、证券市场提供各种数据服务，如：上证所月报、证监会数据报送、TopView数据加工、公开信息披露等；系统包含了查询、报表、OLAP等应用达1000多个，部署了九大应用专区，用户数超过500人，遍及40个所内部门。
* 架构扩展及分析应用深化：最早构建的数据仓库系统作为所内SAS分析的基础数据库提供数据挖掘分析的信息平台，同时支持市场监管、新产品研发的核心数据统计分析，为规范市场和产品创新提供了有力的决策支持；发展到今天，Teradata也在不断的引入新的大数据平台和分析技术，丰富和完善交易所数据生态环境，应用了Teradata Aster、开源R等技术，与证券监管业务结合，展开一系列的数据分析与探索应用开发，充分利用数据资源，提升业务数据分析价值。
* 系统管理：建成了数据仓库管理系统、元数据管理系统、非结构化数据管理系统等辅助数据仓库管理，同时制定了一系列配套的项目和系统开发维护管理制度。

#### 项目成果

**建设实施历程和成果**

* 概述

上证所数据仓库的建设到目前为止在大阶段上经过了数据仓库一期、数据仓库二期、信息增值服务建设、数据仓库灾备系统建设、架构扩展及数据分析应用深化五个过程，整个建设过程遵循合理的规划、科学的部署原则，完成了数据集中存储、应用平台搭建、信息集中服务、安全运营保障、数据分析应用深化的建设历程。

* 数据仓库一期建设
* 建设目标

数据仓库一期的建设目标是完成上证所基本信息平台的搭建。通过集中、统一的数据中心对所有结构化信息和非结构化信息的整合机制，最大程度地保证了上证所自建所以来所积累的全部历史信息，以及以后新增的所有有价值数据的完整性和可用性。

* 建设过程

上证所于2002年启动数据仓库一期(交易历史数据存储项目)的建设，经过将近一年的持续努力，正式建成了国内证券行业数据容量最大、数据内容最完整的TB级数据仓库一期工程。数据仓库一期的建成标志着上证所构建起了基本信息平台，通过集中、统一的数据中心对所有结构化信息和非结构化信息的整合机制，最大程度地保证了上证所自建所以来所积累的全部历史信息，以及以后新增的所有有价值数据的完整性和可用性。

交易历史数据存储项目于2002年10月正式进入项目实施阶段，上证所抽调精兵强将，组成项目组，制定了项目管理、文档规范、安全保密、变更控制、汇报制度等；专门设立项目管理委员会，所有重大事项均通过管理委员会进行决策。根据项目内容，划分了介质、数据分析、加载、应用等多条主线。全部实施过程经过了信息调查分析、业务需求分析、逻辑数据模型设计、体系结构设计、物理数据库设计、ETL设计开发、历史数据加载、备份、业务应用设计开发、系统测试、验收移交、用户培训及知识转移等各个环节。

由于前期做了充分的准备，采用科学的工作方法。在2003年初，上证所就建立了自动数据加载机制，并开始对2003年新产生的每日增量交易数据进行抽取、转换、加载；一方面不断清洗历史存量数据，另一方面根据用户的反馈，不断完善各种功能。2003年11月，开发工作全部完成，其后的测试、验收、全面试运行均取得了圆满成功。

* 建设成果

在科学的逻辑数据模型基础上，上证所通过恢复历史备份的数千张光盘和磁带，完成了1990年以来所有交易历史数据的整合与存储；通过对多个交易相关业务子系统的处理，完成了交易相关上市公司、会员、基金、债券、指数、板块、统计数据等信息的整合；通过设计对非结构化数据的整合方案，完成了交易所8大业务部门长期积累的数百万页非结构化文档信息的存储；并且通过建立了数据自动加载机制，完成了对于以后日常业务增量信息的自动加载处理。

* 数据仓库二期建设
* 建设目标

在完成数据仓库一期建设的基础上，于2005年启动数据仓库二期建设，数据仓库二期建设目标是在数据仓库一期的基础上开发信息服务平台及应用服务体系，使信息产生价值为所内外提供数据加工、统计分析、市场监管、决策支持服务，为进一步发挥数据仓库的价值服务打下良好的基础。

* 建设过程

数据仓库二期项目于2005年1月从业务部门的集中需求访谈开始，向所内9大业务部门共进行了32次的需求访谈，访问了业务部门领导和业务人员共40余人次，共收集了511个原始需求。项目组对原始需求进行了归纳整理，建立了全所级的10大需求分类，并在部门原始需求的基础上归纳整理形成311个全所级作为应用开发的业务需求。

在2005年6月底完成了第一阶段应用的开发工作，与此同时面向业务部门进行了大规模的应用推广工作，向业务部门使用人员全面地介绍数据仓库系统的操作和分析的使用方法，经过推广工作体现了系统各项开发技术都已成熟，业务部门对数据仓库系统有了全面的认识。

项目于2005年10完成了第二阶段大规模的开发工作，于此同时元数据管理子项目、非结构化数据管理子项目、数据仓库管理子项目也已经完成，并在生产环境下完成了全面的部署。2005年10月至11月期间进行了大规模的集成测试工作，对系统非功能性和功能性的各项指标全面进行了测试，为系统的投产打下了良好的基础。

项目于2005年12月之后进入试运行，于2006年2月转入了正式运行。系统运行的情况完全达到了项目的各项指标的要求，同时也作为上证所建成15周年的所庆献礼。

* 建设成果

数据仓库二期项目的建设完成了前端应用开发共完成了600多个包含报表、查询、多维分析、KPI等多种形式的应用，为上证所各业务部门提供统计报表、数据查询、关联分析等信息服务；数据仓库二期的Portal集成了系统的各个子项目的功能，包括：前端应用、元数据管理、非结构化数据管理、数据仓库管理等。通过统一的门户入口，用户根据权限分配情况可使用数据仓库二期提供的各项服务，实现单点登入的功能；对用户使用数据仓库系统的用户进行统一的管理；元数据管理子项目包含了元数据管理和数据管理，提供了完整的元数据管理和应用功能，包括：信息浏览、流程管理、在线帮助、影响分析、安全管理等内容。

* 信息增值服务建设
* 建设目标

通过数据仓库二期的建设，塔建的信息服务的基础平台已初步显现出数据仓库关联分析、数据统计的威力；上证所于2006年基于数据仓库平台逐步加大信息增值服务的开发，完成了Top-View信息产品加工、统计报表移植、债券统计系统移植、异动股票统计、交易公开信息披露统计、计费系统移植、权证信息披露等。

* 建设过程

在完成基础数据存储平台、应用服务平台搭建的基础上，上证所不断扩大数据仓库系统的应用支持服务；面向所外提供各种信息发布和增值服务支持，如：异动股票信息发布、权证信息披露、交易公开信息披露、Top-View信息产品数据加工、上证所统计月报/季报/年鉴等，使上证所数据仓库系统与证券市场息息相连，特别是Top-View信息产品数据加工通过向市场投资者提供信息增值服务为上证所带来了可观的经济创收；面向市场监管单位提供证监会数据报送、证监会统计报告、证监局外派机构监管信息统计等；面向所内除了提供前端应用查询、分析外，向所内提供的数据加工服务有：成交概貌统计、指数概况统计、债券信息统计、会员计费统计、决策数据支持等，同时向SAS信息挖掘系统提供数据分析支持。

* 建设成果

通过信息增值服务的建设，上证所一方面基于数据仓库统一视图的数据管理提供最权威的数据统计信息，确保数据统计口径的一致性和真实性，另一方面减少了重复投资代价。在确保数据仓库系统安全稳定运行的基础上，所内相关的统计报表系统逐步迁移到数据仓库的应用平台上，通过单一的数据中心向辐射提供各方面的信息支持服务，通过单一的门户网站集成所内相关的应用系统向所内相关人员提供信息查询服务。

* 数据仓库灾备系统建设

由于上证所数据仓库系统越来越成为企业运营不可缺失的支持平台，一旦数据仓库系统停止服务将有可能影响整个证券市场的正常运转，因此建立双中心的运营保障数据仓库体系对确保证券市场稳定发展是非常必要的。

于2007年上证所开始着手数据仓库灾备系统的建设，通过建立两套数据仓库系统实现双中心可主/备切换，达到同城异地灾备的目的，从而确保数据仓库系统提供持续稳定的对外服务，目前该项目在建设过程中。

通过双加载实现主从系统数据的同步，为确保双系统数据的一致性在系统之间增加数据一致性检查，在数据加载的事前、事中、事后进行数据的一致性检查，确保双系统的数据相一致，建制两套对等的应用服务系统使其可任意切换数据库服务，以松偶合方式增强系统的容错能力，任何一套数据仓库系统都可独立运行，通过简单的设置实现主/从系统的切换，达到互为备份的目的。

* 架构扩展及数据分析应用深化

上证所数据仓库系统在上证所领导的规划下以科学的发展观逐步实现新一代信息系统的建设，根据上证所信息系统长期发展的战略规划，结合Teradata数据平台技术不断的发展，Teradata也向交易所提出了建设逻辑数据平台的建议，并逐步落地，提供了大数据技术Teradata Aster平台软件、开源R软件等深化交易所的各项数据探索分析与监管应用等等，从广度和深度上不断拓展和加深数据生态系统的建设。

在现有数据服务对象和服务内容的基础上，为更大地发挥上证所数据平台的价值，Teradata 向交易所提供了数据治理咨询规划等方案，完善数据平台生态环境的构建，扩大数据仓库服务的对象和内容；在服务对象方面，在考虑信息安全的前提下，进行面向全市场提供数据服务的探研，使市场参与者能够直接或间接受益于数据平台的服务；在服务内容方面将不断丰富上市公司、会员公司、证券市场的服务内容，为更好地服务于市场提供全面的服务。

### 中国证监会监测中心案例

|  |  |
| --- | --- |
| **客户案例名称** | 中国证监会监测中心案例 |
| **案例规模** | 100T |
| **实施范围** | 详见案例说明部份 |
| **实施时间** | 2011年至今 |
| **数据仓库+大数据技术** | 数据仓库MPP |
| **数据治理咨询** | 无 |
| **系统架构** | EDW数据架构 |
| **主要技术特点** | 数据仓库整体规划与咨询 海量数据存储 并行计算 Teradata金融行业数据模型实施 |
| **现状** | 持续投产运行中 |
| **主要参与人员** | 乔某尉、朱某翔、戴某华等 |
| **经验** | 证券行业咨询规划与实施经验， 对数据源的理解、模型建设以及行业应用的快速构建 |

#### 项目背景

中国资本市场建立二十年来，从无到有、从小到大、逐步发展壮大。尤其是近年来，随着市场规模的急剧扩大，机构数量的显著增多，投资品种日益丰富，创业板、融资融券、股指期货等相继推出，资本市场为我国经济发展和综合国力提升作出了巨大贡献。

资本市场是国民经济的“晴雨表”，其变化牵动着广大投资者、各种社会舆论、国内外金融机构乃至国家领导的目光。因此，对市场整体进行高效的分析与监测，及时发现市场运行中潜藏的系统性风险，为管理层决策提供有效支持十分重要。

中国证监会“资本市场运行监测系统”，是一个企业级的数据仓库项目，它将要求采用先进的信息技术，建立一个覆盖资本市场的运行监测系统，以期及时有效地监控各种市场信息，根据现代金融模型对海量数据进行跟踪预测、分析判断和风险预警。“资本市场运行监测系统”将整合上海交易所、深圳交易所、各期货交易所、中国登记结算公司、证监会内部相关业务系统、外部数据供应商等提供的金融市场信息数据，搭建跨系统集成、跨部门共享、跨应用关联的统一数据平台，基于资本市场逻辑数据模型，实现历史3+1年数据的集中存储。

#### 项目成果

资本市场运行监测数据仓库项目的总体目标是：建立覆盖资本市场的全方位动态监测系统，通过获取行业自律机构、存管银行、证券期货经营机构、国内金融资讯机构以及国际重要市场等各资本市场参与主体的基础数据，搭建一个涵盖国内与国际宏观经济、股票、基金、债权、期货市场的多功能分析平台。利用多市场、多层次、多时点的海量信息数据，全方位地监测资本市场运行动态，多维度地分析各方数据，提高宏观经济和资本市场运行分析能力，更加有效地监测资本市场总体运行情况，及时防范和化解系统性风险，为国家制定金融监管和市场发展政策提供决策支持依据。

2011年，证监会经过对国内多家数据仓库厂商平台的综合评价，最终选择了Teradata作为资本市场数据仓库平台。

资本市场运行监测数据仓库项目一期的功能性建设包括了如下内容：

* + **数据仓库平台建设**

1、数据仓库平台建设主要包括：

* 集中资本市场的数据资源，基于资本市场逻辑数据模型，实现历史3+1年数据的集中存储，以实现资源共享和一致的业务视图；
* 建立一个集中的、一致的、可扩展的中央数据仓库基础设施；
* 建设从多个数据源进行数据采集的网络和平台；
* 建立业务数据批量交换规范，在此基础上开发一套ETL(数据抽取、转换和加载)程序，形成标准的批量信息交换接口和统一任务调度机制，为系统之间交换数据以及从数据仓库获得所需要的数据提供统一的ETL平台；
* 建立数据质量检查和跟踪机制，对数据质量和冲突报告进行集中管理，分项跟踪，由数据质量专门管理人员将数据质量问题拆分给责任人落实；
* 建立一套系统安全机制，确保只有授权的用户才能访问与之权限相对应的数据；

2、数据仓库平台的功能主要包括：

* 采集源系统每日数据并导入数据仓库系统。
* 数据仓库系统根据设定的数据清洗、转换、加载和数据质量检查规则对导入的数据进行数据清洗、转换、加载和数据质量检查，并基于企业数据模型对数据进行集中存储与管理。
* 数据仓库系统进一步根据业务应用系统的要求，对数据进行加工处理，以便为业务报表、查询应用和专题分析等提供必要的汇总数据。
  + **统计分析应用建设**

统计分析业务应用建设主要包括：

* 建立统计分析应用门户（Portal）

建立统计分析型应用的统一门户，方便业务人员使用和访问。

* 业务统计与分析应用建设

提供各业务指标的定义，汇总层级和展现方式。推荐具体的报表工具，实施报表平台。提供由数据仓库进行后端数据支持，前端工具展现的整体实现方式。

业务统计与分析应用包括各类固定报表、灵活查询和统计分析以及专题分析等。具体包括如下内容：

* 满足现有的证监会日常统计工作报表应用开发；
* 灵活查询和统计分析，包括市场综合分析、投资品种分析、投资者分析等；
* 专题分析，如盈亏分析等。
* 管理者驾驶舱(KPI Dashboard)

### 郑州商品期货交易所

|  |  |
| --- | --- |
| **客户案例名称** | 郑州商品期货交易所 |
| **案例规模** | 源系统：4个 支持应用：9个业务部门，报表60多张，多维分析主题2个， 数据量：9T |
| **实施范围** | 详见案例说明部份 |
| **实施时间** | 2015年1月-2016年5月 |
| **数据仓库+大数据技术** | 数据仓库MPP |
| **数据治理咨询** | 无 |
| **系统架构** | EDW数据架构 |
| **主要技术特点** | 数据仓库整体规划与咨询 海量数据存储 并行计算 Teradata金融行业数据模型 |
| **现状** | 已经投产运行 |
| **主要参与人员** | 项目经理：赵某军 咨询顾问：刑某莉 模型师：孟某脉、黄某辉 ETL：吉某鹏、刘某文 |
| **经验** | 1、期货行业咨询规划与实施经验，对数据源的理解、 2、金融模型在期货行业的客户化经验 |

#### 客户简介

郑州商品交易所（以下简称郑商所）是经国务院批准成立的我国首家期货市场试点单位。郑商所隶属中国证券监督管理委员会管理。

郑商所按照《期货交易管理条例》和《期货交易所管理办法》履行职能。依据《郑州商品交易所章程》、《郑州商品交易所交易规则》及其实施细则和办法实行自律性管理。遵循公开、公平、公正和诚实信用的原则，为期货合约集中竞价交易提供场所、设施及相关服务，对期货交易进行市场一线监管，防范市场风险，安全组织交易。

郑商所实行会员制。会员大会是郑商所权力机构，由全体会员组成。理事会是会员大会常设机构，下设战略发展、品种、监察、交易、交割、会员资格审查、调解、财务、技术、结算等10个专门委员会。截至2016年底，郑商所共有会员196家，分布在全国26个省（市）、自治区。其中期货公司会员155家，占会员总数的79%；非期货公司会员41家，占会员总数的21%。

#### 项目建设成果

2014年，郑州商品交易数据仓库系统采用了Teradata公司的数据仓库平台产品。该项目是要在建立金融业务逻辑数据模型的基础上，整合企业各业务系统中的数据，构建企业级数据仓库，需要整合的源业务系统包括：交易结算系统、会员服务系统、品种信息数据库、市场监察系统、官方网站，并提供数据管理平台，包括元数据系统等、数据质量系统等。

通过建设数据仓库系统，到达如下总体目标：

一、基本平台建设方面：

1. 建设企业级数据仓库模型，该模型以金融行业为蓝本，兼顾期货行业特点，具有稳定性、可扩展性、灵活性以及前瞻性。模型具有清晰的主题结构，将覆盖整个交易所现有以及目前可预见的全部业务。

2. 建设企业级数据仓库，将逻辑模型固化，该数据仓库存储结构合理、层次明确，对各类数据进行统一采集、存储、清理、转换、汇总，形成企业统一数据字典以及统一数据口径。

3. 建设数据仓库统一管控平台，对数据仓库日常管理具有监控、管理功能，形成统一的数据质量、元数据管理机制。

二、应用建设方面：

建设数据集市、集市模型、指标库、统一报表平台、即席查询、行为分析、领导驾驶舱、统一门户等，满足各业务部门在能对业务准确并及时了解其运行状态，对日常业务数据进行分析和统计的需要，

该项目由天睿信科提供了主要的软件开发实施服务。一期完成的工作包括：

* 数据仓库体系架构设计
* 信息调研
* 逻辑数据模型设计
* 物理数据模型设计
* ETL任务设计与开发
* ETL调度平台建议
* 元数据管理系统设计与开发
* 数据质量管理系统设计与开发
* 历史数据补采

### 申万宏源证券案例

|  |  |
| --- | --- |
| **客户案例名称** | 申万宏源证券案例 |
| **案例规模** | 先后接入38个源系统（原宏源31个，原申万7个） 数据量：25T 应用：13个业务部门，542个应用接口 |
| **实施范围** | 详见案例说明部份 |
| **实施时间** | 原宏源数据仓库：2011年-2012 申万宏源数据整合：2016年1月-至今 |
| **数据仓库+大数据技术** | 数据仓库MPP |
| **数据治理咨询** | 无 |
| **系统架构** | EDW双活数据架构 |
| **主要技术特点** | 数据仓库整体规划与咨询 海量数据存储 并行计算 Teradata金融行业数据模型 异地双活系统架构的实施 |
| **现状** | 刚刚整合完原申万和原宏源数据入数据仓库，已经上线 |
| **主要参与人员** | 原宏源数据仓库：梅某雁、刑某莉等 申万宏源数据整合： 项目经理：赵某军 咨询顾问：刑某莉 模型师：孟某脉、易某李 ETL：吉某鹏、李某沛、熊某文 应用接口开发：张某学、左某校 |
| **经验** | 1、证券行业咨询规划与实施经验，对数据源的理解 2、证券公司合并重组后的数据整合入仓经验 3、证券行业模型建设以及行业应用的快速构建 4、异地双活系统的建设经验 |

#### 项目背景

宏源证券经过多年的IT建设，已建立了经纪业务类、营销服务类、自营类、资管类、投行类、资讯类、其它管理类等多个信息系统。系统之间相互独立，既有数据重叠，更有数据缺失，不能形成共享整合的企业级数据平台，无法为公司分析决策提供完整的信息。

数据中心是支持证券公司实现基于数据的精细化营销、服务和管理从而提升核心竞争力的基础。只有建立一个企业级的数据中心才能高质量的满足持续不断的数据应用需求。为此，需要尽快建设以客户为中心、共享整合的企业数据中心,将客户数据和业务数据集中整合到统一的数据基础架构之中，并尝试运用多维分析、数据挖掘等商业智能工具进行客户分析和业务分析，发掘价值客户、预警客户流失、定向产品营销等，帮助证券公司提升营销服务和风险管控的能力。

#### 项目成果

宏源证券企业级数据中心的建设是一个长期持续的过程，宏源数据中心项目采用“统筹规划、分步实施”的策略，分阶段有序推进：

一期项目：建立技术平台，实现数据整合。其主要目标，一是为企业数据中心选择高效、可持续发展的技术平台，二是建立良好的可扩展的客户模型和业务数据模型，三是实现客户和业务数据整合，并在此数据中心的基础上开发业务急需的示范应用。客户和业务数据整合应优先建立统一的客户视图并实现经纪业务数据整合，在此基础上再逐步将其他业务和信息系统的数据整合到企业数据中心。

二期项目：持续完善企业数据中心。将各类依赖于统一企业数据的信息系统的数据来源迁移到企业数据中心。

三期项目：基于数据中心的业务应用与创新。整合的企业数据中心可为客户分析和业务分析提供全面支持，基于这些分析可以开发出各种新的业务和服务产品，推动公司创新发展。

2011年，宏源证券经过对国内多家数据仓库厂商平台的综合评价，最终选择了Teradata作为数据中心仓库平台。

## 银行业案例

### 工商银行案例

|  |  |
| --- | --- |
| **客户案例名称** | 工商银行 |
| **案例规模** | 纳入113个源业务系统3430+张数据表，覆盖存款、贷款、信用卡、票据、基金、债券、理财、结售汇、贵金属、电子银行、金融衍生产品等领域，截止2015年底，积累了744T用户数据，建立了80多个应用集市，为绩效考核、营销、综合统计等65个应用提供3487+数据接口服务 |
| **实施范围** | 详见案例说明部份 |
| **实施时间** | 2005年至今 |
| **数据仓库+大数据技术** | 数据仓库、Aster、Hadoop |
| **数据治理咨询** | 2008年开始数据治理咨询 |
| **系统架构** | EDW双活 UDA架构：EDW+Aster+Hadoop |
| **主要技术特点** | 数据仓库整体规划与咨询； 数据仓库成熟度评估 海量数据存储 并行计算 Teradata金融行业数据模型 |
| **现状** | 持续运维运行中； |
| **主要参与人员** | 陈某峰、谢某群等 |
| **经验** | 全国性大型银行业数据平台咨询规划与实施经验， 对数据源的理解、模型建设以及行业应用的构建 |

#### 项目背景

中国工商银行于2000年底起着手启动全行数据仓库建设。由于中国工商银行数据仓库建设初期业务需求迫切，为了加速建设步伐，确保周期短、投资小、见效快，中国工商银行经过分析论证，决定采用先建数据集市、后建数据仓库的方案进行数据仓库建设。目前已经构建了个人客户关系管理系统、法人客户关系管理系统、业绩价值管理系统、审计信息系统、风险监控系统、利率管理系统等六个数据集市。这些集市为中国工商银行带来了巨大效益。

但与此同时，因这几个数据集市的应用分析结果只能建立在来自不同业务系统的、不一致的数据源基础上，在后台分析型应用方面的现状是缺乏共享的基础数据平台，缺乏一致的业务信息视图，无法有效地支撑与解决日益增长的分析型业务应用的需求，使得各分析型子系统之间存在一定程度的信息冗余和对相同数据重复操作，从而带来一些操作处理、存储管理的冗余，加大了开发和维护成本。

2004年10月份，中国工商银行启动了数据仓库逻辑数据模型客户化建设项目，整个项目建设历时5个月，基本形成了中国工商银行自有知识产权的数据仓库逻辑数据模型。

2005年3月，在顺利完成数据仓库逻辑数据模型建设项目后，中国工商银行启动了数据仓库物理平台选型测试的任务。该任务历时8个月，于05年11月顺利完成。物理平台选型测试工作作为中国工商银行数据仓库物理平台产品选型的参考、考察数据仓库产品厂商的技术水平、服务能力和实施经验的任务，为中国工商银行数据仓库物理平台的选型奠定了坚实的基础。

在进行数据仓库商务的同时，为保证后续数据仓库项目的顺利实施，2006年12月，中国工商银行启动了数据仓库体系建设规划研究项目，从架构、功能、规模等方面对数据仓库和数据集市进行规划，以指导后续数据仓库建设工作。

2007年4月，中国工商银行数据仓库实施项目启动，到2008年1月正式上线试运行，开发测试工作历时10个月，已经顺利完成投产工作和系统上线。

截至到目前，中国工商银行数据仓库建设已经进入第三期，纳入113个源业务系统3430+张数据表，源系统覆盖范围达到90%以上。基础数据平台模型实体2072张，支持下游17个应用，160个数据接口。



#### 项目成果

* + **一期成果**
* 设计并建立企业级基础数据平台基础设施；
* 设计数据管控体系与数据基础数据平台管理模式；
* 建立基础数据平台开发、测试、运行架构；
* 对FS-LDM7.0进行客户化；
* 分析上游共107个应用，其中重点分析应用68个，最终决定本期入仓52个，其中主机23个，开放平台29个；分析源表9438张，其中确定入仓的表1186张，其中存量方式下载687个，增量方式下载499个；候选表1367张；
* 开发3个应用，非零售信用风险内部评级、资本金管理、指标库。
* 建立元数据信息管理平台；
* 建立数据质量检查平台；
* **二期成果**
* 继续增加整合10个源系统，追加源表近400张，使得入仓源系统总数达到62个，源表总数达到1566张；
* 对数据模型进行优化；
* 对ETL调度进行优化；
* 开发支持银税报表应用
* 开发支持金融监管报表（1104)
* 开发支持ATM流程优化应用
* 开发支持非居民存款业务统计应用
* 开发支持个人目标客户监控分析应用
* 开发支持专业化经营、系统化管理 改革及绩效评估应用
* 开发支持数据质量管理平台应用
* 开发支持PCRM数据源切换
* 开发支持法人不良贷款欠款扣收应用
* 开发支持个人客户内部评级应用
* 开发支持理财自动化报表应用
* 开发支持业务运营风险管理系统
* 开发支持存贷通理财协议项目
* 开发支持个人客户财富管理
* 开发支持境内机构布局优化模型评估
* **三期成果**
* 开发支持CIIS-PCRS联动查询
* 开发支持个人征信数据源切换
* 开发支持私人银行客户关系管理及投资组合管理
* 开发支持新一代全球现金管理

#### 实施效益

经过近两年的持续开发和建设，中国工商银行数据仓库系统在不断的发展壮大，随着整合源系统数目的增加，仓库已经积累了大量的数据，成为全行整合度最高、信息量最大的系统之一。支持的应用的数目也从一期的3个扩充到现在的17个，并初步建立起比较完善的元数据管理、数据质量体系。当然中国工商银行数据仓库的数据规模之大、复杂程度之高也是国内绝无仅有的，所以尽管经历了历时两年的开发，系统仍然处在一个逐步开放阶段，应用开发方式仍以加工对外供数支持为主，即席查询也刚刚开始对总行各部门开放，正准备逐步向分行推广。但必竟已是初见成效，特别是对17个应用的数据支持。具体实现效益如下：

* 非零售信用风险内部评级：是非零售信用风险内部评级业务应用体系的重要构成部分，其主要目的是为非零售信用风险内部评级模型研发、模型有效性后验及优化、评级体系运行监控、以及信用风险项下的压力测试等业务管理应用提供数据平台支持，并逐步构建以相关数据支持为基础的业务管理应用；
* 资本金：接口包括法人贷款的信用风险经济资本计量，个人贷款的信用风险经济资本计量，资金业务、无息资产和表外资产信用风险经济资本计量、债券投资的信用风险经济资本计量，票据贴现的信用风险经济资本计量，银行卡透支的信用风险经济资本计量，操作风险经济资本计量，其他风险经济资本计量，法人贷款关联客户信息文件，个人贷款合同和相关合同文件；
* 指标库：建立起中国工商银行的常用报表指标库，利用数据仓库强大的计算引挚进行计算，业务部门可以方便的取用计算好的指标数据，也可以方便的定制修改指标定义，形成全行标准的指标库；
* 银税报表：定期为税务部门提供企业的资产负债表、现金流量表，损益表等数据；
* 金融监管报表（1104)：将1104监管报表移植到数据仓库上来，使得1104报表数据更加准确、完整；
* ATM流程优化：统计各地区、各机构的自助设备配备情况，以及自助设备在柜台客户分流方面的具体作用，从而分析自助设备配备的合理性，为自助设备配置优化提供参考依据；
* PCRM数据源切换：将原来在SAS上开发和PCRM移植到数据仓库，利用数据仓库强大的计算引挚进行处理，原来的PCRM系统只能处理部分高端客户，准确性和完整性能稍差些，而且存在性能问题，移植到EDW以后，这些问题就不存在了；
* 非居民存款业务统计：根据外汇管理局的要求，将人民币非居民个人存款纳入热钱监测范围。实现3个接口：非居民个人存款汇总、港澳非居民存款汇总、台湾非居民存款汇总；
* 个人目标客户监控分析：二级支行、支行、二级分行、分行、总行潜力客户统计；
* 专业化经营、系统化管理、改革及绩效评估：网点中间业务销售额统计、客户经理中间业务销售额统计；
* 法人不良贷款欠款扣收：不良客户欠款扣收二期07年10月投产，基本实现不良个人客户与其关联人、帐销案存客户在全国的资产和存款信息自动匹配、自动下发功能。包括欠款明细、存款明细、资产信息、关联企业存款信息、关联企业资产信息、月初客户信息、法人代表存款明细；
* 个人客户内部评级：一期做了非零售客户的内部评级，二期继续开发零售客户的内部评级，也是按照BASEL II的要求，支持个人客户信用风险内部评级模型研发、模型有效性后验及优化、评级体系运行监控、以及信用风险项下的压力测试等业务管理应用提供数据平台支持，并逐步构建以相关数据支持为基础的业务管理应用；所用的信息项主要包括借据信息、客户信息、借据还款信息、逾期借据信息等；
* 理财自动化报表：实现理财产品经营信息自动从EDW和其他业务系统的自动抽取，并且根据业务规则生成各种理财产品的基本信息，各理财产品分机构营销情况，分渠道营销情况，分客户营销情况，贡献度分析等。目前实现的理财产品包括：自主发行，代理销售，代客理财；
* 业务运营风险管理系统：通过统计个人活期账户交易情况来计算个人活期账户资金偏离度；
* 存贷通理财协议项目：主要统计存贷通理财协议的各类信息；
* 个人客户财富管理：为资产在100万以上的财富客户提供专属服务，EDW提供财富客户金融资产统计、签约、流失、金融资产销售、开销户统计；
* 境内机构布局优化模型评估：提供对私、对公总体情况，对私、对公账户情况，网点绩效考核统计；
* 基于仓库的即席查询也在逐步推广，目前，已对总行包括管信部、信用卡部等7个业务部门开放了即席查询功能，目前正准备逐步向分行做试点推广工作。

### 招商银行案例

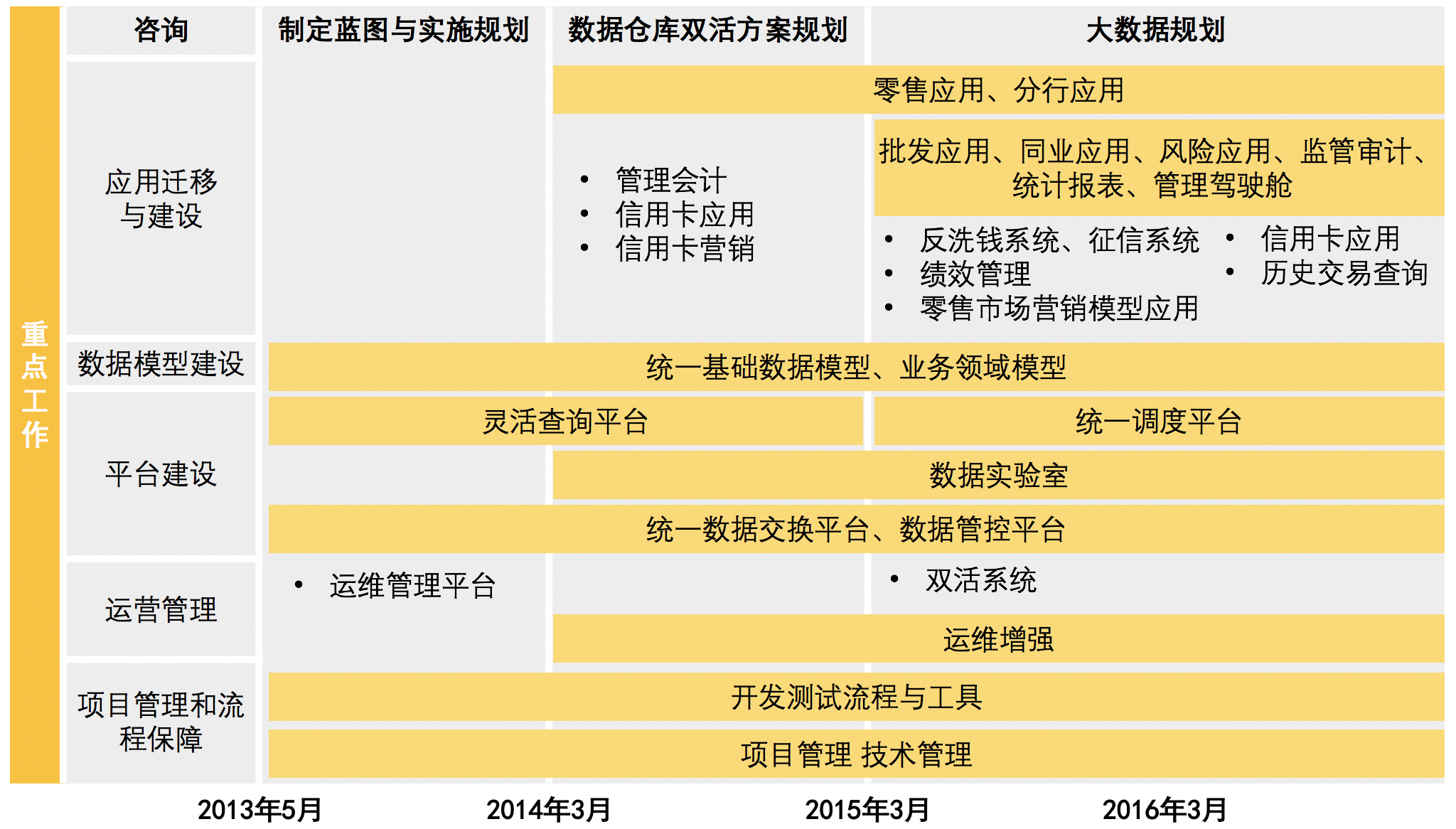
|  |  |
| --- | --- |
| **客户案例名称** | 招商银行 |
| **案例规模** | 累计发布60+应用，数据量615T |
| **实施范围** | 详见案例说明部份 |
| **实施时间** | 2013.06至今 |
| **数据仓库+大数据技术** | 数据仓库MPP |
| **数据治理咨询** | 目前正在规划数据治理咨询项目 |
| **系统架构** | EDW双活数据架构 |
| **主要技术特点** | 数据平台整体咨询规划与项目实施 海量数据存储与并行计算 Teradata金融行业数据模型实施 |
| **现状** | 持续投产运行中 |
| **主要参与人员** | 陈某宇、曾某标、寇某毅、佘某斌、覃某刚、李某斌、李某琛、丁某淦、李某杰等 |
| **经验** | 全国性大型银行业咨询规划与实施经验， 对数据源的理解、模型建设以及行业应用的快速构建、IBM DB2平台应用迁移 |

#### 项目背景

在数据经济时代，数据已经成为企业的重要资产。如何发挥数据的价值，是数据在管理决策、业务创新、日常运营等方面产生更大的效益，是每个企业所面临的机遇和挑战。招商银行在很早之前已经意识到数据对银行发展的重要性，并于1998年建立了当时国内第一套可用的数据仓库系统。

经过多年的使用，随着业务应用和数据量的不断发展，招商银行在数据仓库系统的架构、可用性、负载管理能力、线性扩展能力等方面都遇到了挑战。为此，招商银行启动了新一代数据仓库平台的选型工作，并于2013年1月决定选择Teradata 6690企业级动态数据平台作为招商银行新一代数据仓库平台。有了强大的平台能力支持，招商银行着手新一代数据仓库的规划和实施工作。

招商银行数据仓库项目发展历程如下：



招商银行新一代数据仓库项目自2013年5月启动后，历经数据平台咨询、制定蓝图实施规划、数据仓库双活方案规划、大数据规划四个阶段，围绕应用迁移与建设、数据模型建设、平台建设、运营管理四大方面进行，通过招商银行与Teradata联合制定的数据仓库项目管理和运营管理规范，保障数据平台咨询和项目落地实施顺利开展与高质量成果交付。

#### 项目成果

招商银行第一期项目的成果包含实施规划成果及2013年项目实施建设成果两大部分。

* + **实施规划成果**

完成招商银行数据仓库新平台实施总体规划，包括基于Teradata平台的技术、信息、运维等架构规划，制定了包含项目实施（含应用迁移）策略、线路、方法、总体计划等内容的整体实施方案；

* + **实施建设成果**

在2013年内依据数据仓库新平台实施总体规划完成第一期实施，实现数据仓库新平台投产，其中包含以下内容：

* 遵循招商银行数据标准，完成基于FS-LDM V12.0的本期数据建模及其ETL开发，本期数据建模需要实现数据装载范围包括：当前招商银行数据仓库中操作型数据存储（ODS）和统一基础数据模型（UDM）已经采集装载且常用的业务系统源数据，涉及数据源表的数量不超过1500个；
* 建立与数据仓库新平台配套的ETL调度和数据抽取方法和工具；
* 建立与数据仓库新平台配套的元数据管理和数据质量管理的规范、流程和工具；
* 建立与数据仓库新平台配套的开发、测试、投产的规范、流程和工具；
* 建立与数据仓库新平台配套的运行维护的规范、流程和工具；
* 了解招商银行数据仓库应用现状，分析应用改进和业务提升机会，制定应用迁移计划，并实现初始应用上线。完成建设需求。

### 建设银行

|  |  |
| --- | --- |
| **客户案例名称** | 建设银行 |
| **案例规模** | 70+集市，数据量700T |
| **实施范围** | 详见案例说明部份 |
| **实施时间** | 2005年至今 |
| **数据仓库+大数据技术** | 数据仓库MPP |
| **数据治理咨询** | 2008年开始数据治理咨询 |
| **系统架构** | EDW双系统架构（2750+6650） |
| **主要技术特点** | 海量数据存储 并行计算 Teradata金融行业数据模型 |
| **现状** | 持续投产运行中 |
| **主要参与人员** | 黄某林、赖某聪、徐某兰等 |
| **经验** | 建行新一代仓库体系架构下的实施经验 |

#### 项目背景

中国建设银行于2000年底起着手启动全行数据仓库建设。由于中国建设银行行数据仓库建设初期业务需求迫切，为了加速建设步伐，确保周期短、投资小、见效快，中国建设银行经过分析论证，决定采用先建数据集市、后建数据仓库的方案进行数据仓库建设。目前已经构建了个人客户关系管理系统、法人客户关系管理系统、业绩价值管理系统、审计信息系统、风险监控系统、利率管理系统等六个数据集市。这些集市为中国建设银行带来了巨大效益。

但与此同时，因这几个数据集市的应用分析结果只能建立在来自不同业务系统的、不一致的数据源基础上，在后台分析型应用方面的现状是缺乏共享的基础数据平台，缺乏一致的业务信息视图，无法有效地支撑与解决日益增长的分析型业务应用的需求，使得各分析型子系统之间存在一定程度的信息冗余和对相同数据重复操作，从而带来一些操作处理、存储管理的冗余，加大了开发和维护成本。

2004年10月份，中国建设银行启动了数据仓库逻辑数据模型客户化建设项目，整个项目建设历时5个月，基本形成了中国建设银行自有知识产权的数据仓库逻辑数据模型。

2005年3月，在顺利完成数据仓库逻辑数据模型建设项目后，中国建设银行启动了数据仓库物理平台选型测试的任务。该任务历时8个月，于05年11月顺利完成。物理平台选型测试工作作为中国建设银行数据仓库物理平台产品选型的参考、考察数据仓库产品厂商的技术水平、服务能力和实施经验的任务，为中国建设银行数据仓库物理平台的选型奠定了坚实的基础。

在进行数据仓库商务的同时，为保证后续数据仓库项目的顺利实施，2006年12月，中国建设银行启动了数据仓库体系建设规划研究项目，从架构、功能、规模等方面对数据仓库和数据集市进行规划，以指导后续数据仓库建设工作。

中国建设银行启动数据仓库源于几个方面的要求：

* + **来自外部的要求**

必须符合相关上市规则、条例和会计准则的要求：

* 国内：财政部、证监会、银监会、人民银行的各种法规、准则和特定要求，如1104监管报表、反洗钱报表
* 香港：香港财务报告准则或国际财务报告准则、公司条例、香港联合交易所的证券上市规则
* 公众股东、债券持有人及证券交易所：要求上市公司作大量的信息披露，包括年度报告、中期报告及持续披露（比如重大收购、出售资产事项、关联交易，其他对股价可能产生重大影响的信息）的内容，如关联交易披露。
  + **来自内部的要求**
* 要求信息管理部门：
* 提升对产品创新和客户服务的支持能力
* 提升经营管理决策的支持能力
* 提升风险与内控管理的支持能力
* 提升制定业务政策、提高管理效率所需数据信息的支持能力
* 提升企业绩效管理所需要的数据支持能力
* 提升对最高决策所需要的信息支持能力

2005年3月，中国建设银行启动了数据仓库物理平台选型测试的任务。2006年12月，中国建设银行启动了数据仓库体系建设规划研究项目，从架构、功能、规模等方面对数据仓库和数据集市进行规划，以指导后续数据仓库建设工作。

2007年4月，中国建设银行数据仓库实施项目启动，到2008年1月正式上线试运行，开发测试工作历时10个月，已经顺利完成投产工作和系统上线。

#### 项目成果

* **一期项目实施成果**
* 设计并建立企业级基础数据平台基础设施；
* 对FS-LDM7.0进行客户化；
* 设计数据管控体系与数据基础数据平台管理模式；
* 建立基础数据平台开发、测试、运行架构；
* 整合8个源系统，建立了企业级的协议、团体、产品、资产、地域数据信息；
* 开发4个应用，1104监管报表、管理决策仪表盘、资产负债信息管理系统数据支持、业务指标库（SMIS）。
* **项目建设历程**

中国建设银行数据平台项目从2005年正式启动建设，经历了多年建设，已经具有相当的规模。建设历程如下：

* 2006年9月份完成了一期EDW&MIS系统的建设工作，系统初始配置为8个5250节点，可用数据容量3T
* 2007年3月完成一期硬件扩容项目，扩容后系统配置升级为6个5450H节点，可用数据容量由3T增加至11T
* 2008年7月完成二期扩容项目，扩容后系统配置升级为（18+1）个5500H节点，可用数据容量增加至110T。
* 2011年进行三期扩容项目，可用数据容量将增加至250T以上，裸硬盘容量超过900TB。
* 2013年进行四期扩容项目，可用数据容量将增加至610T以上，裸硬盘容量超过2.2PB。
* **案例项目实施效益**

经过持续的开发和建设，中国建设银行建立了企业级的数据平台系统，并不断扩展数据服务和应用支持能力，建立、完善了企业级的协议、客户、产品、资产、地址、事件、渠道数据信息，提高了数据的共享度，可为各类分析型应用提供数据支持。总行信息中心、资产负债管理部、合规部、个人金融部都受益于基础数据平台的数据支持。

鉴于基础数据平台的建设是一个持续不断的过程，同时基础数据平台还要不间断的支持分析型应用。技术部和信息中心即达成一致并请示行领导同意，未来基础数据平台系统建设通过每年立项，以滚动开发的方式进行。通过滚动开发机制的建立，逐步建立一个相对稳定的基础数据平台管理和应用支持服务团队，不断完善和丰富基础数据平台的内容，推动数据质量的持续改善，实现对应用需求的快速响应。通过项目的实施，初步形成了中国建设银行基础数据平台支持服务团队的雏形。目前已经建立了综合管理、需求管理、数据分析、技术实施、测试、运行维护和数据质量小组，并通过需求管理工单、应用任务分配单、组间协同工单和数据质量跟踪改进工单等方式，保证了基础数据平台的不断完善，满足管理决策需求。

根据2009年所做的估算，若2009年第三期结束时的35个应用采用开放平台替代实现，相应服务器、存储设备和维护费用约9000万元，与仓库的现有设备投资基本相当。

将大数据量应用加工移到数据仓库，间接节约了设备投入：

* 营运稽核：节省了分行每年约200万的运维成本和约1300万的设备投入。
* 反洗钱：节省了分行每年约200万的运维成本和约1065万的设备资源。

截止到目前，数据仓库系统中的用户数据量为700T，每日从78个源系统抽取数据，日增量数据约600G，数据仓库模型约1292张表，支持下游66个应用。

### 中行信用卡中心

|  |  |
| --- | --- |
| **客户案例名称** | 中行信用卡中心 |
| **案例规模** | 40T |
| **实施范围** | 详见案例说明部份 |
| **实施时间** | 2004年至今 |
| **数据仓库+大数据技术** | 数据仓库MPP |
| **数据治理咨询** | 2012年开始数据治理咨询与实施工作 |
| **系统架构** | EDW架构 |
| **主要技术特点** | 数据平台整体咨询规划与项目实施 海量数据存储与并行计算 Teradata金融行业数据模型实施 |
| **现状** | 持续投产运行中 |
| **主要参与人员** | 张某健、宋某华等 |
| **经验** | 全国性大型银行业咨询规划与实施经验，对数据源的理解、模型建设以及行业应用的快速构建 |

#### 项目背景

中国银行卡中心数据仓库建设项目自2003年启动，成功完成了数据仓库基础信息分析平台的建设，还建设了以Teradata TCRM软件平台为支撑的市场营销子系统及包括客户行为分析子系统在内的七个数据挖掘模型应用。此外，银行卡数据仓库还实施了机构信息管理与维护（管理全辖的机构编码信息及其他一些公用信息）、风险准备金提拨（计提风险准备金）、发卡/收单分润、人行征信数据报送、MIS Portal等应用。数据仓库系统已经成为银行卡中心业务关键系统之一，对于支撑中国银行卡业务的发展承担越来越重要的职能。

#### 项目成果

* **系统建设历程**
* 中行信用卡中心2004年初完成数据仓库一期的建设，系统初始配置为2个5350节点，可用数据容量600G
* 2006年，系统进行第一次扩容，新增2个5400节点与原有节点共存，可用数据容量增加至3T
* 2009年，系统配置扩展成2个5555H，可用数据容量增加到9T
* 2013年，系统配置扩展成2个6650H，可用数据容量增加到52T
* **案例项目实施效益**

经过多年的建设，中国银行银行卡中心已建立了完整的银行卡数据整合、业务分析、市场营销与客户服务体系。通过业界最先进的技术和专业的实施，中国银行银行卡中心不仅具备了及时高效、全方位的精细化管理能力，同时具备了完整的、闭环的客户关系管理能力，实现了数据驱动业务、信息带动发展的经营管理水平，构成了一个业务与信息相互促进、营销与服务协同工作、共同支持银行卡业务运营的完整的信息技术体系。

* 客户单一视图的建立，支持客户的产品交叉持有分析，对满足信用条件的目标客户进行交叉销售，分析客户的行为习惯，在满足信用条件的情况下进行产品升级，多种产品组合可以巩固客户关系，降低流失率；
* 各种客户信息和交易信息数据导入DW，洞悉客户及客户行为，支持客户细分、客户生命周期管理、为已有客户提供个性化与差异化服务；
* DW整合发卡、交换、财务等各个源系统的数据，为业务管理部门提供更为及时、准确的MIS报表、KPI等经营管理数据；
* 基于帐户级的收入、风险、成本计算，支持到分行、销售团队、客户经理的分润与绩效考核，实现基于客户价值的绩效考核和分行分润；
* 负面客户信息的整理应用，作为新卡审批和额度核准的参考，降低银行发卡风险；
* 统一规范地满足监管单位（人民银行、外管局等）的监管要求，快速且低成本地实现各级监管机构的监管需求。
* 数据仓库作为中国银行信用卡信息平台，广泛使用于中国银行总分行的信息人员，截止2008年5月，总行数据仓库BI用户数达到50个左右，分布在市场、风险、财务、经管等部门，一般BI人员可以看到PUBLIC组的数据，如需看到敏感数据（如身份证号、地址），需要得到超级用户授权。另外，Portal发布报表总数为298个，总行、32分行的Portal用户数达到170，权限管理上，总行用户可看到全行数据，分行用户只能看到分行的数据。

中国银行银行卡中心数据仓库中，不仅包括了银行卡的数据，还包括了个人金融的全部数据。

截至2012年底，中国银行银行卡中心数据仓库实现从近20个源系统的取数；数据仓库应用数量达到了43个，包括如下12类应用：管理信息系统、经营指标快报、客户关系管理、分润和绩效考核、监管报送、数据接口、数据挖掘、灵活查询、公司卡数据分析、收单业务分析应用、运行监控。其中仅管理信息系统的报表数量达到300多张。

### 上海浦东发展银行案例

|  |  |
| --- | --- |
| **客户案例名称** | 上海浦东发展银行 |
| **案例规模** | 数据仓库已经集成了78个总行源系统、18个总行参数、9个分行参数上传的业务数据； 基于数据仓库已经建设了65个总行应用和27个分行应用，满足了总分行各个部门对业务分析和数据服务的需求。包括监管类应用、分析管理应用、历史数据查询应用、运营系统的数据服务、分行数据分发服务等； 数据容量规模147TB； |
| **实施范围** | 详见案例说明部份 |
| **实施时间** | 2007年至今 |
| **数据仓库+大数据技术** | 数据仓库、Aster、DataLab |
| **数据治理咨询** | 2013年进行数据架构及治理咨询 |
| **系统架构** | 逻辑数据仓库混搭架构：EDW+Aster+DataLab |
| **主要技术特点** | 数据仓库整体规划与咨询； 数据仓库成熟度评估 海量数据存储 并行计算 Teradata金融行业数据模型 |
| **现状** | 持续运维运行中； 配合进行新一代数据仓库建设；数据仓库优化，以及进行CoD扩容规划 |
| **主要参与人员** | PM：汪某俊， 模型师：梅某燕 架构师：孙某兵等 |
| **经验** | 银行业数据平台咨询规划与实施经验， 对数据源的理解、模型建设以及行业应用的快速构建 |

#### 客户简介

上海浦东发展银行是 1992 年 8 月 28 日经中国人民银行批准设立、于 1993 年 1 月 9 日正式开业的股份制商业银行，总行设在上海。

经中国人民银行、中国证监会正式批准，上海浦东发展银行于 1999 年获准公开发行Ａ股股票，并在上海证券交易所正式挂牌上市（股票交易代码： 600000 ），这是中国银行业改革的一项重大举措。注册资本金达 39.15亿元，良好的业绩和诚信经营的声誉使浦发银行业已成为中国证券市场中一家备受关注和尊敬的上市公司。

主营业务主要包括：吸收公众存款、发放短期、中期和长期贷款、办理结算、办理票据贴现、发行金融债券、代理发行、代理兑付、承销政府债券、买卖政府债券、同业拆借、提供信用证服务及担保;代理收付款项及代理保险业务；提供保管箱服务。外汇存款；外汇贷款；外汇汇款；外汇兑换；国际结算；同业外汇拆借；外汇票据的承兑和贴现；外汇借款；外汇担保；结汇、售汇；买卖和代理买卖股票以外的外币有价证券；代客外汇买卖；资信调查、咨询、见证业务；人民币和外币信用卡业务；离岸金融业务；经中国人民银行批准的其他业务。

根据2007年核心资本排名，公司2008年7月获评英国《银行家》杂志全球1000强银行排名第176位，比前年提前了15位排名；2008年6月公司被英国《金融时报》评为全球市值500强企业第422位。公司还被国际评级研究机构 RepuTex（崇德）评为“中国十佳可持续发展企业”，被中国扶贫基金会评选为“2008中国民生行动先锋”。在中国信息协会、中国服务贸易协会联合主办的2007-2008中国最佳客户服务评选中，公司95528客户服务中心荣获“2007-2008中国最佳客户服务中心奖”。公司在《经济观察报》主办的“2007年度中国最佳银行评选”中荣膺“2007最佳企业伙伴银行”奖项。由中国《金融时报》主办的“2008中国最佳金融机构排行榜”中荣膺“最佳风险控制银行”和“最具成长性银行”；由《首席财务官》杂志社主办“2008年度中国CFO最信赖的银行评选”中荣获“最佳公司金融品牌奖”、“最佳投行业务奖”和“最佳贸易金融奖”。2008年12月，被《21世纪经济报道》的“2008年亚洲银行竞争力排名”中名列综合竞争力15强。

公司自开业以来，经营规模不断扩大，网点连年增加，经营效益持续提高。 截至2008年末，公司已在全国49个城市开设了32家分行，机构网点总数增至491家。2008年底，公司总资产为13094.25亿元，本外币一般贷款余额为6975.65亿元，本外币一般存款余额为9472.94亿元，2008年公司实现税后利润125.16亿元。

目前浦发银行拥有个人和公司客户数约2000万，帐户数约4500万。

#### 项目背景与发展历程

* **POC (2005.10 – 2006.1)**

Teradata EDW解决方案与项目实施方法论验证，项目管理能力与初步应用验证

* **数据仓库首期 与EDA项目（2006.3 – 2007.3）**

企业级数据仓库平台搭建，完成CIM、1104等应用项目

* **数据仓库运维及应用推进（2007.4 – 2008.3）**

EDW运行维护，实施包括东方卡、总账指标、支付结算报表等专题应用

* **全行应用推广（2008. 4 - ）**

EDW运行维护，分行应用推广

#### 项目实施成果

* **POC (2005.10 – 2006.1)**
* 卡业务分析：
* 东方卡发卡情况统计表
* 东方卡存款结构统计表
* 公司业务 ：
* 公司业务产品分类统计查询
* 公司业务客户分析模型
* 财务分析：
* 流动性比例监测表
* 风险分析：
* 不良贷款金额及不良贷款率查询
* 反洗钱：
* 人民币个人客户分散转入集中转出可疑交易
* 人民币个人客户集中转入分散转出可疑交易
* 人民币公司客户分散转入集中转出可疑交易
* 人民币公司客户集中转入分散转出可疑交易
* 外币个人客户分散汇入集中汇出可疑交易
* 外币个人客户集中汇入分散汇出可疑交易
* **数据仓库首期与EDA项目（2006.3 – 2007.3）**
* 系统层面
* 完成项目范围内的数据加载，建立统一的、共享的数据仓库平台；
* 完成DW Automation对第三方工具（DataStage）的统一任务调度与监控管理；
* 完成系统监控到HP OpenView的集成；
* 技术层面
* 建立具有先进性、灵活性和可扩展性的企业级数据仓库的总体架构；
* 建立一套完善的ETL体系，完成数据的抽取、加载与转换。
* 应用层面
* 完成CIM应用系统的实施；
* POC当前已验证的应用成果转产；
* 完成1104银监会监管报表报送。
* 管理层面
* 通过Teradata数据仓库系统管理手段完善系统管理；
* 借助于数据仓库项目建设和维护方法论，完善项目不同阶段的管理。
* 规划层面（EDA）

在当前EDW平台基础之上建立有效的数据管理框架，建立有效的管理组织架构，制定应用发展规划，制定IT长远规划，重点解决以下几个方面的问题：

* 数据管理框架
* 管理组织架构
* IT长远规划
* 应用发展规划
* **数据仓库运维及应用推进（2007.4 – 2008.3）**
* 数据仓库运维
* 数据仓库生产环境和开发测试环境的日常维护管理（DBA）
* ETL系统的日常维护支持
* 现有数据仓库应用系统的日常维护支持
* 源业务系统变化引起的数据仓库变更
* 基于数据仓库的数据请求支持
* 基于数据仓库的应用开发支持
* 其它浦发银行数据仓库维护团队负责人根据实际需求的安排所开展的与数据仓库运行维护紧密相关的工作
* 应用推进

面向报表开发的应用数据复用层的规划设计及实现；

* 报表应用数据集市的设计与开发；
* 前端报表展现的设计与开发，从而构建一套基于浦发银行数据仓库平台，面向全行的全新报表应用体系；
* 形成基于数据仓库平台的符合EDA规范的应用开发规范和流程（包括需求、开发、变更、数据质量流程和文档体系）以及配套的InfoUser、PowerUser、AppUser的用户体系；
* 实现东方卡/代发、支付结算、总帐指标体系等主题分析应用。
* **全行应用推广（2008. 4 - ）**
* 数据仓库应用全行范围推广，开放数据中间层给分行用户，同时协助开发基于数据仓库的各类分行应用，如分行绩效考核、ATM使用统计分析等。
* 全面建设对公客户关系管理系统和零售客户关系管理系统；
* 全面支持内评、风险加权资产等风险管理相关的项目；

#### 实施效益

浦发银行EDW自2006.12.1日投产运行以来，已经能够对以下应用提供数据支持。

* 浦发银行1104项目
* 反洗钱
* 人行集中金融统计信息系统
* 个人理财与对公OCRM
* 综合对账单
* 对公客户关系管理系统
* 零售客户关系管理系统
* 内部评级和风险加强资产项目
* 其他（运营监测、审计、分行绩效等）

数据仓库还陆续支持了一些紧急灵活查询分析需求，包括地震赈灾分析、重庆分行贷款逾期分析、个人房贷断供分析、伪造卡交易信息分析、轻松理财卡和借记卡年费定价分析等工作，在为日常的业务运营和决策提供了高质量的数据和高效率的支撑。

在数据仓库之上的分析应用开展地如火如荼之时，浦发银行的业务也在井然有序的发展，在这一年中成功推出了资产证券化业务、第三方存管业务、利多多业务等等。为跟上业务发展的步伐、及时地支持新业务应用分析、有效地积淀历史数据，数据仓库这一年中成功地进行了多次有效地扩展，将新业务也纳入到数据仓库之中，顺利解决了应用分析及时性的问题。数据仓库为客观准确地掌握业务发展状况提供有力的后盾。

2008年以来，浦发银行为充分发挥数据仓库的巨大优势和蕴含的业务价值，决定在全行范围内进行推广应用，包括在一定程度上开放数据仓库对全行的模型和数据支持。基于数据仓库的新的应用开发项目如内评、分行绩效考核、客户关系管理等也在如火如荼的展开。

为保证浦发银行不断发展的经营分析和决策需求，公司和浦发的业务也在不断发展，围绕数据仓库基础平台和数据仓库管理信息系统的服务团队也不断成长，目前，团队（含合作伙伴）规模已经达到35人，服务内容包括数据仓库基础平台项目和3个应用建设项目，项目都采用Teradata总包的模式。

### 交通银行案例

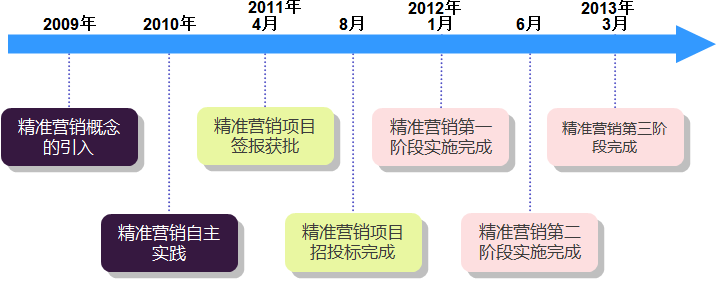
|  |  |
| --- | --- |
| **客户案例名称** | 交通银行 |
| **案例规模** | 客户化10个主题、859个实体；为77个业务应用提供数据支持； 数据规模约200TB； |
| **实施范围** | 详见案例说明部份 |
| **实施时间** | 2006年至今 |
| **数据仓库+大数据技术** | 数据仓库、Aster、Hadoop |
| **数据治理咨询** | 2012年数据标准和精准营销咨询 |
| **系统架构** | 数据仓库双系统 |
| **主要技术特点** | 数据仓库整体规划与咨询； 数据仓库成熟度评估 海量数据存储 并行计算 Teradata金融行业数据模型 |
| **现状** | 持续运维运行中； 配合进行新一代数据仓库建设； |
| **主要参与人员** | PM：佘某斌 模型师：高某华 架构师：谢某群等 |
| **经验** | 银行业数据平台咨询规划与实施经验， 对数据源的理解、模型建设以及行业应用的快速构建 |

#### 客户简介

交通银行精准营销项目全称为“零售客户信息管理与精准营销项目”。该项目由交通银行总行个人金融业务部为主办部门，零售板块、科技板块等相关部门共同参与的全行性的营销管理项目。通过该项目的实施，建立交通银行零售客户的全方位视图，全面实现营销决策科学化、营销管理流程化、客户营销精确化。

交通银行精准营销项目目标是建立银行零售客户的全方位视图，全面实现营销决策科学化、营销管理流程化、客户营销精确化 ，推荐适合产品和服务以满足客户需求，最终提升营销效率和效益，实现“以客户为中心”的战略转型，打造最佳“财富管理银行”。

交行精准营销项目的发展历程如下：



交行精准营销项目主要围绕以下四大能力进行建设：

* + 客户信息管理能力：包括个人客户数据集市建设，客户信息整合与质量提升咨询。
  + 客户深入洞察能力：包括客户细分咨询，业务分析数量模型咨询。
  + 营销活动管理能力：包括事件式营销分析和流程导入试点，营销活动和过程管理咨询。
  + 渠道整合管理能力：包括CRM系统建设咨询。

#### 项目成果

通过分析客户交易数据，筛选 “存款大幅波动”和“新增代发客户”中高营销成功率的客户名单，同时整合卡中心、核心、零贷、电银、基金各系统联系信息，通过OCRM推送实时自动发送给前线客户经理跟进营销，实现了客户经理由熟客营销向生客营销的转变，极大提高营销效率。

* + 在营销试点的应用当中，通过整合多个系统的联系方式，客户经理的营销接触率平均达到80%以上，外呼中心的试点名单接通率提升了1倍。
  + 项目在开展客户信息整合等基础工作的同时，也注重业务的同步实践，确保早出成果。在上海分行率先进行的营销试点取得了16.8%的成功率，比对照组提高了56%。
  + 在营销试点的过程中，除了营销理念和方法有很大转变之外，项目组也很重视对营销流程的自动化支持，以提高工作效率。从项目启动以来，对OCRM系统进行了3次40个功能项的新增/改造。

### 中信银行案例

|  |  |
| --- | --- |
| **客户案例名称** | 中信银行 |
| **案例规模** | 截止2016.12，共纳入107个源系统数据，支持12个仓内集市，52个仓外应用系统，数据规模99T |
| **实施范围** | 详见案例说明部份 |
| **实施时间** | 2014.04至今 |
| **数据仓库+大数据技术** | 数据仓库MPP、Aster |
| **数据治理咨询** | 2011年开始数据治理咨询与实施工作 |
| **系统架构** | EDW数据架构 |
| **主要技术特点** | 数据仓库整体规划与咨询 海量数据存储 并行计算 Teradata金融行业数据模型 |
| **现状** | 持续投产运行中 |
| **主要参与人员** | 项目经理：丁某博 业务专家：殷某辉 模型师：高某莉 架构师：曹某科 |
| **经验** | 金融行业咨询规划与实施经验，对金融行业源系统的理解、模型的建设以及行业应用的快速构建，以及项目实施过程中遇到的各种问题是如何处理的，促进提升项目的实施质量 |

#### 项目背景

中信银行在2014年1月，综合内外部因素，结合新核心“1＋2＋9”系统工程的良好契机，启动了基础性、战略性的企业级数据中心建设项目，旨在推动数据共享和分析应用，促进数据标准化和数据质量治理，在数据架构、数据管控和相关应用领域达到同业平均水平。

2014年4月，中信银行宣布采用Teradata数据仓库平台作为该行企业级数据中心平台的主系统。

#### 项目成果

项目主要目标为：

1、搭建全行统一、共享的数据集成平台，采用统一模型策略，实现对基础明细数据的整合，对通用共性指标的加工，同时实现数据标准化和数据质量检查。

2、数据集成平台规划为统计报表、监管报送、客户分析、管理会计、绩效考核、风险管理等管理分析型应用提供数据支持服务，并逐步有序对管理分析型系统进行整合。

3、基于数据集成平台搭建即席查询环境，为总分行管理分析类用户提供即席查询服务，满足灵活多变、临时性、一次性的分析需求。

### 兴业银行案例

|  |  |
| --- | --- |
| **客户案例名称** | 兴业银行 |
| **案例规模** | 接入13个源系统，客户化10个主题，3651个实体；数据规模约70TB |
| **实施范围** | 详见案例说明部份 |
| **实施时间** | 2005年至今 |
| **数据仓库+大数据技术** | 数据仓库MPP |
| **数据治理咨询** | 2011年数据标准咨询 |
| **系统架构** | 数据仓库系统 |
| **主要技术特点** | 数据仓库整体规划与咨询； 精准营销咨询； 海量数据存储 并行计算 Teradata金融行业数据模型 |
| **现状** | 持续运维运行中； |
| **主要参与人员** | PM：李某杰 模型师：刘某茜 架构师：谢某群等 |
| **经验** | 银行业数据平台咨询规划与实施经验， 对数据源的理解、模型建设以及行业应用的快速构建 |

#### 客户简介

兴业银行成立于 1988 年 8 月，是经国务院、中国人民银行批准成立的首批股份制商业银行之一，总行设在福建省福州市。

本行经营范围包括：存款、贷款；国内外结算；票据承兑与贴现；发行金融债券；代理发行、代理兑付、承销政府债券；买卖政府债券、金融债券；同业拆借；买卖、代理买卖外汇；银行卡业务；信用证服务及担保；代理收付款项及代理保险业务；保管箱服务；结售汇业务；短期融资券承销、基金托管、金融衍生产品交易等及其他经国务院银行业监督管理机构批准的业务。

截至2009年9月末，兴业银行资产总额为12663.12亿元，股东权益为558.48亿元，不良贷款比率为0.61%。今年1至9月累计实现税后利润95.72亿元。根据英国《银行家》杂志2009年7月发布的全球银行1000强排名，兴业银行按总资产排名列第108位，按一级资本排名117位。 根据美国《福布斯》发布的2009全球上市公司2000强排名，兴业银行综合排名第389位，在307家上榜的全球银行中排名第62位。

已在全国主要城市设立了 41家分行、 480 多家分支机构，在上海、北京设立了资金营运中心、零售银行管理总部、信用卡中心、资产托管部、大型客户业务部、投资银行部、期货金融部、可持续金融中心等总行经营性机构，建立了网上银行“兴业在线”、电话银行“95561”和手机银行“无线兴业”，与全球1000多家银行建立了代理关系。

#### 项目背景与发展历程

* 兴业银行企业级数据仓库（首期）项目的目标是：
* 客户化完成兴业银行企业级数据仓库逻辑数据模型，形成全行客户、产品、机构、渠道等主题的基础数据标准规范
* 完成数据仓库硬软件平台建设，并物理化客户相关主题数据模型
* 完成客户相关主题模型的行内业务数据加载
* 项目进展情况

企业级数据仓库项目建设项目一期工作已经完成。集成以下企业业务系统信息。

* 本外币核心业务系统
* 中间业务平台
* 信用卡系统
* 个人外汇实盘买卖系统
* 银证平台
* 开放式基金代销系统
* 商业汇票系统
* 信贷管理系统
* 财务管理系统
* 个人信贷系统
* 营销服务系统

目前EDW系统处在运行维护阶段，除了新增源系统，配合源业务系统作模型的调整外，系统也开始着手在仓库平台上搭建一些应用系统，比较重要的是反洗钱系统和营销客户二期系统。其中营销二期系统涉及几乎所有的对私客户系统和数据，原有的营销一期系统的资产负债数据的计算、客户评级都完全转移到仓库中计算，同时还负责二期新的模拟利润的数据的计算，使仓库成为了对私客户数据的真正的权威数据。

* **2009年事件式营销项目(EBM)**

2009年，Teradata为兴业银行提供了事件式营销咨询和试点工作，先后在4家分行开展了20多个事件的设计、开发和营销推广工作，在营销执行中，获得了远高于传统营销的营销效果。未来将在全行全面推广事件式营销。

* **2013年建设IDA**

项目的总体目标是构建兴业银行企业级自助取数平台，满足业务部门对数据要求的自主性、时效性、灵活性，方便业务部门随时随心地获取到所需要的各类业务数据，支持精细化管理中对数据探索的需求。

* **2015年**

#### 项目实施成果

通过企业级数据仓库一期系统的建设，达到了以下成果：

* + **业务层面**

通过金融逻辑数据模型客户化建立业务数据单一视图，建立起一套《兴业银行基础业务数据定义规范》

* + **系统层面**

完成规定的行内业务系统的数据加载，建立统一的、共享的基础数据平台

* + **技术层面**

建立具有先进性、灵活性和可扩展性的企业级数据仓库的总体结构

建立一套完善的ETL体系完成数据抽取、清洗、加载与转换

提供源业务系统数据的暂存和访问接口

提供面向应用和数据集市的数据访问接口

* + **管理层面**

通过系统管理和元数据管理等手段完善系统管理。

借助于数据仓库项目建设和维护方法论，建立起EDW长期发展的管理架构和运作流程。

### 光大银行案例

|  |  |
| --- | --- |
| **客户案例名称** | 光大银行 |
| **案例规模** | 70个源系统，34个应用，EDW数据量120T |
| **实施范围** | 详见案例说明部份 |
| **实施时间** | 2006年至今 |
| **数据仓库+大数据技术** | Teradata MPP、Greenplum、Hadoop |
| **数据治理咨询** | 2008年开始数据治理咨询与实施工作 |
| **系统架构** | EDW架构、UDA架构 |
| **主要技术特点** | 数据仓库整体规划与咨询； Teradata金融行业数据模型 |
| **现状** | 持续投产运行中 |
| **主要参与人员** | 王某华、邓某庚、陈某燕等 |
| **经验** | 基于TADD的银行业数据仓库开发经验，对银行数据的理解以及模型建设。 |

#### 客户简介

中国光大银行成立于1992年8月，1997年1月完成股份制改造，成为国内第一家国有控股并有国际金融组织参股的全国性股份制商业银行。多年来,伴随着中国经济和金融业的发展进程，光大银行不断改革创新、锐意进取，始终把自身发展与国民经济的增长紧密结合，在为社会提供优质金融服务的同时，取得了良好的经营业绩，逐步形成了与现代商业银行相适应的多元化股权结构、日益完善的公司治理与经营机制、比较先进的科技支持系统、素质较高的员工队伍、布局合理的机构网络、范围广泛的同业合作等优势。

#### 项目背景与发展历程

目前，光大银行IT建设正处于从数据大集中阶段向应用集成和金融创新阶段的过渡。当前，光大银行信息化的重点基本上围绕着核心业务系统的建设和改造、渠道整合技术的提升以及各种管理系统的建设而展开。2008年，如何集成各应用系统，加快渠道系统与银行核心业务系统、后台管理系统的连接，如何遵从“以账户为中心”向“以客户为中心”转化的趋势，打造适应银行未来发展的信息系统，如何加快信息系统的开发节奏，统一系统的数据规范——这些都已成为新一轮银行信息化建设所必须面对的重点课题。

EDW项目的本质在于数据管理与数据应用的有效整合。为了在高速增长期内实现高利润率和最大化股东价值，光大银行从2006年起加快了信息智能化建设的步伐，构建全行统一的基础数据平台，即企业级的数据仓库系统，作为全行信息化建设的基础。企业级数据仓库系统的建设是一个持续不断的过程。在项目一、二期，重点建设光大银行企业级逻辑数据模型，并在此基础上整合光大银行核心业务系统和十多个外围交易处理系统的数据，建成全行的基础数据平台，为光大银行的决策支持和管理信息系统以及监管报表系统提供数据支持。在本期中不断地扩充数据源并扩展逻辑数据模型，增强和完善基础数据平台的数据支持能力。

从数据管理的角度来看，数据仓库系统实现了全行账户、客户、交易数据及大部分历史数据的集中统一存放和管理，从宏观上避免了数据的冗余存放、重复抽取和处理，无论在系统运行效率还是总体投资方面都可以为光大银行带来巨大的经济效益。

从数据应用的角度来看，基础数据平台在前期支持了众多分析应用与管理应用的基础上，本期将继续为资产负债管理与市场风险管理以及新会计准则系统等重大里程碑式全行级应用系统提供持续的数据支持。另外作为统一的全行级别数据提供平台，屏蔽源业务系统的重大升级改造，将分析性应用系统受到的影响减至最少。同时，本期基础数据平台项目将继续支持各个业务部门提出的随机业务查询， 为各项业务的持续发展提供信息支持。

#### 项目实施成果

* 纳入数据仓库的源系统范围：
* 票据系统
* CALL CENTER
* 网银
* 第三方存管
* 信用卡
* 管理会计
* 核心
* 信贷风险管理系统
* 个贷系统
* 国结系统
* 保理系统
* 总行大前置系统
* 银证通系统
* 个人外汇买卖
* 理财产品销售
* 理财产品登记
* 电子支付
* 手机银行
* 信用卡授信监控
* 信用卡审批
* 数据仓库支持的应用系统：

信用卡中心-业务决策支持系统

* 对公操作型CRM系统
* 内部评级法
* 零售操作型/分析型CRM系统
* 电子渠道BI应用项目
* 国结统计分析系统
* 银监会-1104非现场稽核系统
* 反洗钱监测上报系统
* 人民银行-金融统计大集中系统
* 人民银行-支付结算报表系统
* 人民银行-国际收支申报系统
* 管理会计系统（SAP）数据接口
* 资产负债管理系统（ALM）数据接口
* 关键指标KPI系统
* 汇总层模型设计与实施 （共37张表，本期继续进行中）
* 基础数据平台元数据实施
* 数据仓库管理办法编写
* 数据仓库ROI分析报告
* 信息平台规划编写

### 民生银行案例

|  |  |
| --- | --- |
| **客户案例名称** | 民生银行 |
| **案例规模** | 数据量160T |
| **实施范围** | 详见案例说明部份 |
| **实施时间** | 2003年至今 |
| **数据仓库+大数据技术** | 数据仓库MPP |
| **数据治理咨询** | 无 |
| **系统架构** | EDW架构 |
| **主要技术特点** | 数据仓库整体规划与咨询； Teradata金融行业数据模型 |
| **现状** | 持续投产运行中 |
| **主要参与人员** | 王某华、邓某庚、陈某燕等 |
| **经验** | 基于TADD的银行业数据仓库开发经验，对银行数据的理解以及模型建设。 |

#### 客户简介

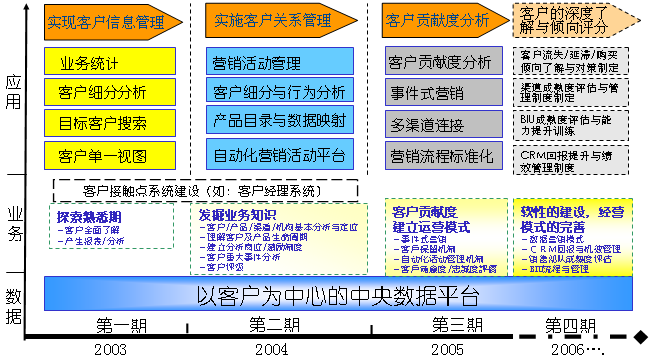
中国民生银行于1996年1月12日在北京正式成立，是我国首家主要由非公有制企业入股的全国性股份制商业银行，2000年12月19日，中国民生银行A股股票（600016）在上海证券兴业银行挂牌上市。

截止2007年末，中国民生银行总资产规模达到9197亿元，存款总额6712亿元，贷款总额5549亿元，不良贷款率1.22%。2004年，实现主营业务收入253亿元，利润总额92亿元，净利润63亿元。

截止2007年末，中国民生银行已在华北地区包括总行及北京、太原、石家庄和天津分行；华东地区包括上海、杭州、宁波、南京、济南、苏州、青岛、温州分行；华南地区包括福州、广州、深圳、泉州、厦门分行和汕头直属支行；其他地区包括西安、大连、重庆、成都、武汉、昆明分行等地设立了23家分行、1家直属支行，在香港设立1家代表处，机构总数量为327个，与境外80多个国家和地区的830多家银行建立了代理行关系。

#### 项目背景与发展历程

* 根据民生银行发展规划，在2001年民生银行提出“八大系统”的建设思路。
* 在2002年，“八大系统”的规划工作正式展开，根据民生银行五年业务发展战略要求，提出了需要增强的业务和IT能力，包括：业务定价、绩效考评、信贷风险管理、客户信息管理、人力资源管理和培训、以及IT管理。
* 在2003年，中国民生银行重塑其管理体系，决定同时启动“八大系统”。即客户信息管理系统、客户经理管理系统、客户服务中心系统、业务流程系统、风险评价系统、业务定价系统、信用卡系统和行员培训系统。
* 在2003年4月，正式启动客户信息和客户关系管理系统（CIM/CRM），并由Teradata实施，同时负责其企业级数据仓库的建设。使CIM/CRM系统建设成为8大系统建设的重中之重。

****

数据仓库建设发展规划如下：

2003年（项目一期）：建立以客户为中心的中央数据仓库平台，实现客户信息的统一收集和管理，将客户信息提升为民生银行的无价资产，并在此基础上建立客户信息管理平台，实现客户单一视图、目标客户搜索、客户细分管理和初步的业务统计分析报表，为实施客户关系管理平台奠定基础。

2004年（项目二期）：重点开始实施客户关系管理，在一期的建设成果上，实施分析型 CRM 解决方案，并与民生银行操作型 CRM 无缝连接，完成呼叫中心的渠道整合，搭建自动化营销的基础平台。

2005年（项目三期）：为了能更好地加强客户关系管理，重点开始建立贡献度分析系统，提升深度分析客户贡献度的能力，帮助实现客户保留、管理产品及服务定价。

2006年（项目回顾期）：巩固前三期成果，酝酿下一个三年规划。充分利用日渐成熟的企业级数据仓库，在此基础上开发更高层次的分析型应用，使信息面得到进一步提升，更好地支持民生银行总行以及各分行的运营和决策。

2008年1月，民生银行以新核心再造工程为契机，适时与Teradata再次合作启动了ECIF系统建设项目，希望能利用Teradata专业服务团队在金融数据模型方面的强大整合能力及Teradata在动态企业智能（AEI）方面的先进技术，帮助民生实现再次超越。

#### 项目实施成果

* **一期CIM**

通过建立CIM系统，对民生银行的数据源进行了详细分析，将不同系统的数据加以整合并组织到统一的数据仓库平台。

结合CIM系统，Teradata和民生银行一起设计了企业级数据仓库逻辑数据模型（CMBC-LDM），完成了民生银行核心业务系统和其他13个外围产品系统的数据抽取、清洗、加载和转换（ECTL）工作，截止2003年底14个源系统的数据，都能以T＋1的方式进入数据仓库。同时对民生银行众多系统的业务代码进行了梳理和规范，建成了民生银行企业级数据仓库基础环境，初步形成了客户单一视图，并能够对全行客户进行客户细分和多维分析，并完成了统一客户信息管理、目标客户搜索、客户细分、业务统计分析等应用。

主要成果包括：

1) 结合FS-LDM建成民生银行企业级数据仓库，整合14个源系统数据。

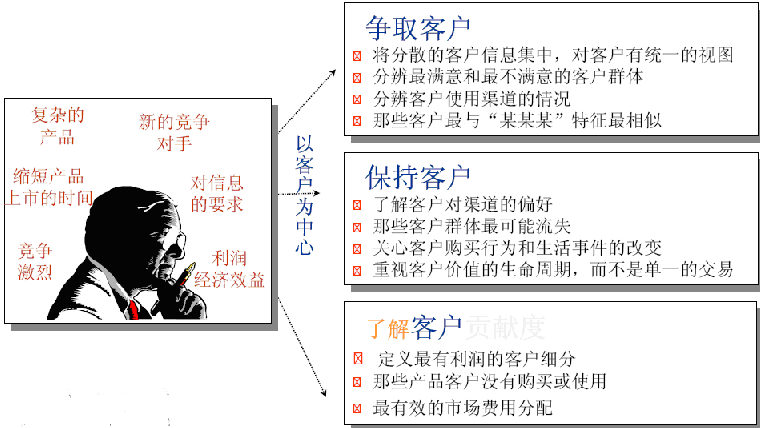
2) 建成全行客户统一视图，并提供对客户360度视角分析的功能，对客户的基本资料、财务信息、关联信息、资信状况、与银行业务往来信息等进行模块化设计，将这些相关信息进行统一的存储和管理，并提供集成统一的访问界面。客户包括民生银行的公司客户、个人客户和同业客户。

3) 终端用户可以自己通过客户细分/客户分析/目标客户搜索定位客户，得到相关客户名单，进一步分析客户，提升企业决策效率与质量，决策人能够自由地获取需要的数据，而不需依赖IT人员。

4) 减轻了OLTP系统支援决策分析报表的负荷。

* **二期CRM**

通过实施Teradata客户关系管理解决方案，民生银行建立了数据仓库-客户细分-客户分析-自动化营销-效果评估-数据仓库这样一个闭环的营销过程，搭建民生银行理想的市场营销环境，实现了更好的客户保留和客户获取、增强交叉销售的效果，并能防范营销过程中过度打扰现象的出现，优化了市场营销活动效率，大大提高了目标销售的成功率。

1. 确立客户关系管理的策略

2) 建立适应客户关系管理的组织架构

3) 建设能支持市场营销活动的分析型的客户关系管理系统

4) 建立与客户服务渠道的连接，实现完善的渠道整合

5) 业务价值-从产品导向走向客户导向

* **三期VA & PMM**

1) 利润贡献度分析VA

* 帮助民生银行建立一套适合自身特色的全行一致的资金转移计价、成本分配、风险准备等方法论
* 利用净利息收入 (NIR)、其他收入 (OR)、直接费用(DE)、间接费用 (IE)、风险准备 (RP)等主要因子进行利润贡献度计算。
* 在(帐户)详细交易数据的基础上，实现先进的帐户、客户、产品、渠道和机构的利润贡献度分析与应用

2) 绩效指标管理PMM

结合数据仓库的优势，Teradata为民生银行提供了一套灵活的绩效指标展现及分析功能，支持对企业经营指标和考核指标的直观展示及管理。

考核指标主要面向民生的高层领导，其特点是以预警方式综合显示各指标完成计划的情况，以便行领导可以在一个界面中能够一目了然地监控某时期各分行对各考核指标的完成情况。而经营指标则面向具体的分析人员，其特点则是使用多种分析图形从细节角度展示指标的具体值，以便分析人员能够从较细的角度查看指标的发展情况。

* **四期回顾与应用增强**
* 贡献度分析结果测试及数据源问题跟踪
* 利润贡献度分析系统（VA）数据及应用增强
* 配合实施大额反洗钱监控系统
* EDW数据增强（系统升级：金卡V2.0）
* 基础模型和应用模型Review
* 数据质量监控系统优化及上线
* 管理会计SAP系统整合压力测试
* EDW数据增强（系统改造：银证通）
* 产品目录维护及新增产品加载
* 对公授信额度贷款相关的模型和ETL调整
* 特殊审批业务跟踪评估系统(客户单一视图贡献度明晰跟踪功能扩展)
* CIM系统测试和优化
* 零售部考核项目
* EDW数据增强（贷记卡系统整合）
* MIS系统历史数据移植
* 个贷五级分类项目
* 积分系统项目
* 电子银行渠道整合项目数据调研
* EDW系统扩容规划

#### 实施效益

从目前的EDW运行情况来看，数据仓库在民生银行实实在在的给业务部门的业务发展和管理提供了帮助，数据平台也在为行内多个系统提供数据。数据仓库正在发挥全行基础数据平台的重要作用。

总体来看，数据仓库平台系统建设在以下方面取得良好的效益：

* 节约重复投资

从经济效益角度来看，数据仓库建设对民生银行的数据源进行了详细分析，完成了民生银行核心业务系统、外围产品系约30余个系统的数据抽取、清洗、加载和转换（ETL）工作，基本实现全行帐户、客户和交易数据集中统一存放和管理，建成了民生银行企业级数据仓库基础环境，从宏观来看基本避免了对数据的冗余存放，和重复处理，在设备方面可为行里节约大量投资。

* 促进管理水平

在实施客户关系管理系统的过程中，不仅仅进行技术的开发，同时在如何利用客户关系管理系统进行市场营销方面也做了大量的推动和培训工作。这些工作促进了各业务部门的营销观念和管理方式的进一步完善和更新，此外在建设数据平台和客户信息整合的过程中，针对一些数据质量和数据缺口的问题，项目组与各科技和业务部门紧密配合，对业务流程的规范，数据管理的完善方面也发挥了作用，可以说，在一定程度上，通过建设本系统，促进了银行管理效益的提高。

* 提升社会效益

通过实施全行统一的数据平台和具有比较先进的营销理念的客户关系管理系统，增加了银行的科技含量，为实现把民生银行打造为领先银行的目标提供了必要的科技基础，进一步提升了民生银行在公众中的形相，一定程度上为银行带了潜在的社会效益。

* 营销观念更新

针对零售部和客服中心等业务部门，除了系统推广的统一培训，项目组还分别组织零售部进行客户细分、沟通活动管理及客户行为分析的相关培训及实践训练，项目组对总行零售客服中心人员进行了沟通活动模块的培训工作，可以说在推动和帮助民生银行的业务单位利用客户关系管理系统进行市场营销工作方面确实取得了一些成效，也得到零售、客服等业务部门的认可。

* 数据质量改进和完善

通过EDW系统实施也进一步推进了民生银行源系统数据质量的改进和完善，一些源系统，如MA管理会计系统等，针对数据仓库实施过程中发现的数据质量问题单独立项进行了系统改进和增强改造。使得全行总体数据质量向着良性循环的方向发展。

## 保险业案例

### 中国人保财险公司案例

|  |  |
| --- | --- |
| **客户案例名称** | 中国人保财险 |
| **案例规模** | 25T |
| **实施范围** | 详见案例说明部份 |
| **实施时间** | 2011年至今 |
| **数据仓库+大数据技术** | 数据仓库MPP |
| **数据治理咨询** | 无 |
| **系统架构** | EDW架构 |
| **主要技术特点** | 数据平台整体咨询规划与项目实施 海量数据存储与并行计算 Teradata金融行业数据模型实施 |
| **现状** | 持续投产运行中 |
| **主要参与人员** | 高某力、郝某军等 |
| **经验** | 大型保险公司数据平台咨询规划与实施经验， 对数据源的理解、模型建设以及行业应用的构建 |

#### 项目背景

中国人民财产保险股份有限公司（简称中国人保财险公司PICC P&C），作为中国国内最大的非寿险保险机构，原先采用Oracle数据库作为数据仓库系统，集中了全辖30多个省市分公司的数据，建设了十多个报表和分析型业务应用（例如财险、车险、理赔、承保、责任险、财务、销售管理、农险、清算）。原数据仓库系统的用户数据规模约为25TB，日增量数据量为几个GB。

受Oracle平台性能和容量限制，原系统只能容纳生产系统60%的数据，仅能存储3年的历史数据，在性能和容量上均无法满足日益增长的业务发展需要。例如在数据加载效率、报表生成效率方面均大大超出正常的业务允许时间范围，给监管报表报送、各项分析统计都产生了极大影响，浪费了大量时间、人力、物力。

2011年5月，该保险公司对Teradata、Oracle Exadata、IBM Netezza、EMC/Greenplum进行POC测试，在各个场景的测试结果中，Teradata均表现优异。2011年6月，该保险公司正式招标，随后Teradata动态企业级数据仓库平台在与Oracle Exadata、IBM Netezza、EMC/Greenplum的竞争中胜出，成为该保险公司选择的数据仓库新平台。

#### 项目成果

Teradata数据仓库新平台，为(4+2)6650数据仓库平台，提供的用户可用空间为60TB。6650数据仓库平台，是Teradata公司最新推出的动态数据仓库系统，是具备同时支持传统硬盘HDD和固态硬盘SSD能力的平台，它具有最强大的处理能力、高可用性、可扩展性，同时具有良好的多代共存投资保护能力。

6650平台采用了Teradata 最新的13.10版本数据库。Teradata 13.10数据库是全球首例时态数据库，每次修改、加载或更新数据后，再也无需手工管理和更新有效日期，而是由系统自动记录。此外，将时序智能置入 Teradata 优化器 (Teradata Optimizer) 意味着用户无需使用复杂的 SQL 或结构化查询语言，便可获得其基于时间的业务问题的解答。自动化数据管理及简化的查询可减少数据错误、优化应用程序开发，并能面向企业用户开放时序数据分析。

Teradata 13.10数据库内嵌并集成了地理信息系统GIS，使得平台在处理地理信息方面更一致、更快速、更容易，能为企业提供将精确地理编码功能与地理空间数据（geospatial data）、以及商业智能与分析架构相结合的有效方法。

项目一期，Teradata为该保险公司建立了设计合理的、可靠稳健的6650数据仓库平台，并将原Oracle数据仓库中的数据迁移到6650数据仓库平台。

Teradata 6650数据仓库平台，为该保险公司提供了业界最强大的处理能力、高可用性、可扩展性，以及良好的多代共存投资保护能力，为该保险公司未来数据仓库分析业务的发展打下坚实的基础。目前该保险公司已把绝大部分的管理分析类业务应用加入数据仓库，并把生产系统中所有的历史数据都装入数据仓库，用户数据量将成倍地增长，该保险公司成为国内数据仓库用户数据量最大的保险公司之一。随着数据仓库应用的不断丰富和完善，大大提升了该保险公司在客户管理、运营管理、风险管理、财务管理、信息管理等各方面的综合能力。

**一、迁移过程：大大提高了数据仓库的准确性和稳定性**

在项目启动之后，Teradata派出了熟悉Oracle数据库和Teradata数据库的技术专家负责迁移项目的整体实施，历时三个月，成功将数据仓库从Oracle平台迁移到了Teradata平台上。在迁移的过程中，在保持业务逻辑一致不变的前提下，项目团队发现并修正了原Oracle平台下的一些错误和系统缺陷，优化之处多达100余处，大大提高数据仓库的准确性和稳定性。

1、借此机会修正源系统中的业务数据

* 大小写不一致，例如：在原系统中的保险单据号有的是大写有的是小写，同存于原Oracle系统中，实际业务上是同一笔；
* 原Oracle系统中的数据包含“回车”、“换行”、“TAB键”以及一些不可见字符；
* 原Oracle系统中对数据存在前空格与后空格的情况，如参数表中地区代码“01”与 “ 01”（带前空格）/“01 ”（带后空格）；

2、优化原系统中调度

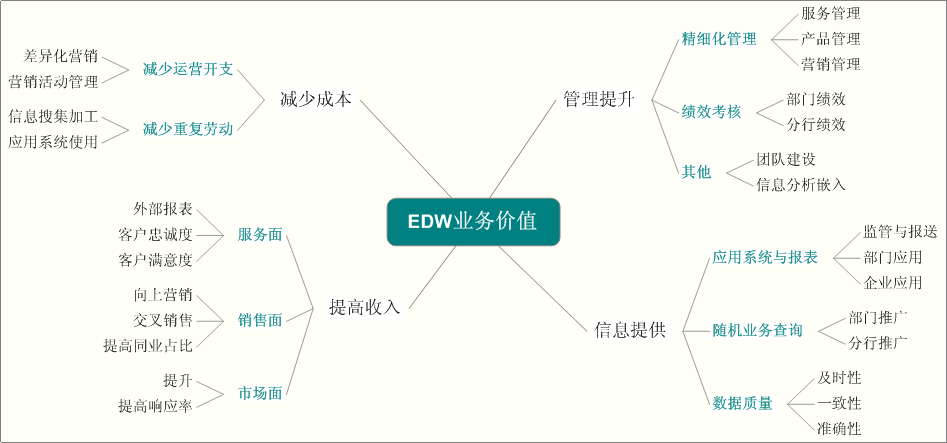
* 对原系统的任务调度做了全面分析整理；
* 将一些串行任务改成并行；
* 优化任务间的依赖关系；

3、修正原系统中的一些错误逻辑

* 例如原Oracle中使用Rownum/Rowid；
* 一些公认的错误语法——错误的关联更新；

**二、迁移之后：继续增强数据处理、分析和管理能力**

该保险公司借助于Teradata公司在数据仓库诸多实施案例中积累的管理经验和业务经验，不断地提高管理水平、强化信息提供能力，减少经营成本，提高运营收入水平，并继续奠定在同业数据仓库领域的领先地位。



2012年12月，该保险公司再次采购“Teradata数据管控平台”，将进一步加强数据处理、分析和管理的能力，增加企业范围内数据治理方面的技术能力和管理规范，从而节约公司资源、提升数据访问效率和完善数据信息安全性。一期将实现元数据管理工具、数据归档工具、ETL工具升级等功能和服务。

### 新华人寿保险公司案例

|  |  |
| --- | --- |
| **客户案例名称** | 新华人寿保险 |
| **案例规模** | 上游数据源NCI新、老核心、13下游应用系统，300多应用 |
| **实施范围** | 参见案例说明 |
| **实施时间** | 2013年至今 |
| **数据仓库+大数据技术** | 数据仓库MPP |
| **数据治理咨询** | 无 |
| **系统架构** | 准实时+批量ODS架构 |
| **主要技术特点** | 准实时数据处理 保险行业逻辑数据模型 |
| **现状** | 持续投产运行中 |
| **主要参与人员** | 王某明、刘某东等 |
| **经验** | 寿险行业数据模型设计 保险行业准实时ODS系统构建 |

#### 项目背景

新华人寿保险于2013年8月采用Teradata数据仓库平台作为全行级“数据集成平台”的主系统。平台建设目的是为了通过汇总、整合企业数据，搭建统一、高质量、企业级的数据与信息服务平台，改变现有系统间分散孤立的IT架构状况；提升统计报表、数据分析和决策支持的速度和能力，满足企业对内加强数据管控、降低运营成本，对外提升品牌竞争能力、增加营销收入等要求。具体来看，平台系统建设的目标可以包括：

* + 符合新华保险公司IT规划要求，符合新核心项目群的总体规划和设计要求；
  + 改变当前系统与系统间点对点的数据交互模式，建立统一的企业级数据服务平台；
  + 对数据集成平台中的数据实现标准化、规范化管理；
  + 数据集成平台中的基础数据应满足相关系统的数据要求；
  + 数据集成平台要满足系统间批量数据交换需求及准实时系统查询需求；
  + 配合新系统上线，解决相关系统与核心系统的数据接口问题；

#### 项目成果

新平台上线后实现如下项目成果：

**一、消除对核心系统的依赖，实现系统解耦**

通过新平台的建设，实现数据接口和交换的一致化、稳定化，实现系统解耦。

* + 提供企业数据模型，屏蔽数据供需双方的变化及其相互影响
  + 对数据的交换和接口进行一致化处理，提供一个公共的平台，实现数据供方和需方的系统解耦

通过新平台实现数据接口和交换的一致化、稳定化，使得新、老核心业务系统切换的过程中间，能够有一个集中的点来屏蔽对下游系统的影响。

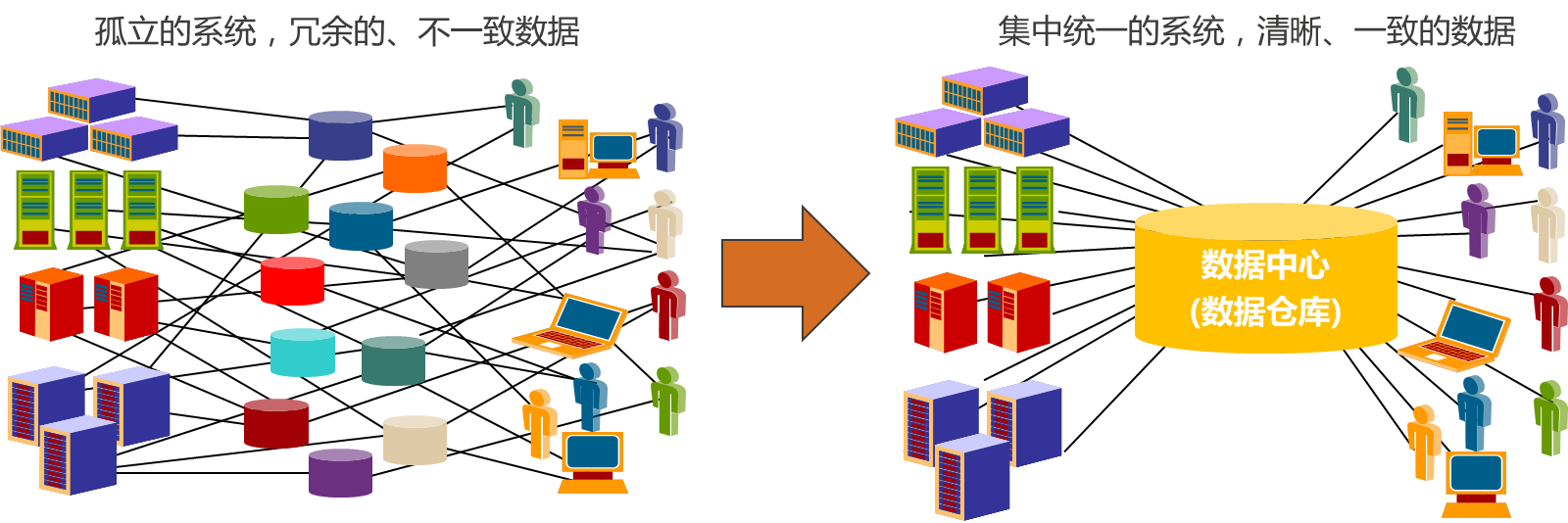
新平台提供了：

* + 提供整合的企业级数据模型，屏蔽数据供需双方的变化及其相互影响
  + 对数据的交换和接口进行一致化处理，提供一个公共的平台，实现数据供方和需方的系统解耦。

**二、对企业数据进行集成，消除网状数据交换**

通过新平台的建设，对企业数据进行集成，消除网状数据交换，将能够更好的为新华保险提供：

* + 统一的基础数据和更好更快的决策基础
  + 保证了数据的一致性、完整性
  + 更低的成本
  + 以及更好更及时的决策



**三、建立数据交换与应用的企业标准，实现企业数据的治理与管控**

原来新华保险的数据采集、传输、交换、共享、应用缺乏一个一致的标准，数据在企业内部流转过程中，产生了大量的不一致和冗余，数据质量难以保障。严重影响对数据的使用和对业务的支撑。

通过新平台的建设，建立数据交换与应用的企业标准，实现企业数据的治理与管控。以新平台为依托，从数据模型、元数据和数据质量入手，建立企业数据的交换与应用标准。

**四、提高数据交换和使用的及时性，为一线生产经营提供动态智能**

通过新平台的建设，提高数据（信息）交换和使用的及时性，Teradata为新华保险提供动态智能（AI）实践。动态智能（AI）指的是为不仅仅为企业管理决策者，同时为企业一线员工、分公司业务人员等提供基于数据事实的（准）实时智能服务。即通过数据集成平台，为分公司、一线运营人员提供接近实时的精细化管理、精准营销管理支撑功能。