

点击上方蓝字关注“信息化时代”



博时基金信息技术部总经理 车宏原

车宏原介绍，基金公司传统的业务逻辑就是通过直销和代销渠道募集资金，对资金进行投资管理，为投资者创造价值。业务线有两条，一是募集资金，二是投资。募集资金主要靠渠道，传统渠道是以银行为主，现在互联网第三方销售、互联网直销渠道业务量也显著增加。

博时基金支撑业务的核心系统有三大类，一是募集资金端的注册登记系统，负责投资者的开户登记及清算;二是投资端的投资交易系统;三是估值系统，衔接募集资金端和投资端，给基金产品估值，让投资者可以按估值来买基金的份额。

车宏原介绍，就整个金融行业而言，基金公司的体量相对于银行、证券都比较小，早期的系统建设采用的是快速见效的模式。具体到数据平台的建设，根据数据量特点和业务需求，博时基金早期主要是采用分散的数据集市的模式来建数据中心，如投研数据中心、销售数据中心等。这种模式下，数据中心与业务应用是高度吻合的，但集市是分散的，数据不是集中的。其优点是起步容易、见效快、成本低;但缺点也很明显，就是数据到一定规模，系统就出现瓶颈，性能不够，可扩展性差。

近年来，在大资管、国际化和互联网金融背景下，基金公司业务已经发生很大的变

化，对IT的需求也发生了很大的变化。“行业初期只有传统的公募基金，提供5×8小时的服务，业务和系统相对都非常简单。但现在不一样了，”车宏原介绍，一方面，有公募、专户、年金、社保等多种形式的产品，有全球化投资，有股权、债权类投资子公司，以及财富管理、智能投顾等更丰富的服务形式。同时，互联网金融让投资者的日常生活与基金业务相关联，需要支持7×24小时的服务和巨大的业务量。最近的热点是金融科技，对云计算、大数据、人工智能等新技术的应用提出了很高的要求。这些因素都让博时基金的IT系统和计算需求越来越复杂，以数据集市模式来支撑如此的数据计算与分析需求，它越来越“力不从心”。

呼唤新一代数据平台

数据分析是为业务服务的。在募集资金端的数据分析是为了进一步了解市场，做精准营销和客户服务；在投资端，数据分析是为了提高投资业绩、提高投资回报水平；另外还有风险控制、产品规划设计、基金运营、公司整体管理决策等方方面面都需要数据分析。

基金行业过去是从建设效率角度考虑，多采用应用驱动的分散的数据集市，企业级统一的数据平台建设起步较晚。车宏原坦言，虽然已经普遍建立了市场营销类、研报表类、资讯类、管理类等数据分析型应用，但这些系统间相对独立，既有数据重叠，也有数据缺失，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。同时，各个部门分析需求各异，造成各种数据分析平台的野蛮生长，进而导致洞察片面、产生分析能力瓶颈以及管理上的高复杂性。

总而言之，数据价值未得到充分应用，博时基金需建设行业领先的企业级新一代数据平台。

另外一方面的背景是企业对数据价值的认识。越来越多的企业意识到数据是企业的核心资产，数据能驱动业务的发展、转型、和变革。这些理念和认识，也是新一代数据平台建设的驱动力，是新一代数据平台建设需求所考虑的因素。

新一代数据平台的建设目标是要综合利用各种技术工具，激活数据资产，提升数据的价值、提升数据的驱动力、提升科技的支撑力。为满足上述需求，新一代数据平台需要建立数据服务和数据驱动两大业务能力，建立数据管理、平台计算、平台服务三大平台支撑能力。在考虑系统建设目标的同时，需要考虑系统建设的约束条件，以实现最合适的投资回报。换句话说，车宏原介绍，它需要全面整合内外部数据，解决数据在企业内充分交流、共享和融合问题，解决重复开发和数据不一致等问题，需要实现在功能、性能、数据质量、数据服务、价值挖掘等方面能力的提升，同时还要实现对主数据的管理能力，实现对非结构化数据、半结构化数据、大数据的分析处理能力。

具体而言，对数据平台的建设需求，车宏原从几个方面来梳理。在营销和客服方面，需要支持渠道分析、销售业绩归属分析、客户画像、客户行为分析、精准营销、客户收益分析、智能客服、智能投顾等功能;在产品管理方面支持产品设计、产品比较分析、市场评价、FOF、智能投顾等功能;在合规和风险方面，需要实现投资组合风险分析和预警、投资组合绩效归因分析、统一的监管报表等;在基金运营方面需要做到运营类查询和报表、操作性风险警示;在经营管理方面要实现各项经营指标分析、管理驾驶舱、多维分析报表;在投研和量化方面要支撑宏观策略研究、投资组合分析、投资辅助决策、智能投顾、量化模型训练与验证、大数据指数研究、舆情与市场热度分析等。

MPP+ Hadoop搭平台一期建设见成果

要做什么定了，那怎么做呢?经过综合考量，博时基金决定用以MPP为主，MPP和Hadoop有效互补的技术架构来搭建统一的数据平台，从而平衡成本、性能、可扩展性等多方面因素。

就像所有的系统选型一样，博时基金数据平台选型也是经过了对多家技术产品的论证对比。MPP有好多家供应商，Hadoop技术方案发展也很快，最终选择了以Teradata的MPP为核心，做传统的核心业务的支持，以Hadoop为辅助，做半结构化和非结构化数据的处理以及历史数据的储存和处理。

为什么会选择Teradata?“我们看重的是平台、模型和方法论等方面的综合能力，Teradata在这方面全球领先。”车宏原说，“建数据平台是个非常大的工程，难度会很大、周期会很长，数据平台建设偏离目标、周期拖得很长、成本失控，甚至项目失败的案例并不少见。先进的模型、科学的方法论和成熟的经验，可以为我们的成功提供保障。”

博时基金企业数据平台的建设，采用了数据仓库之父Bill Inmon提出的范式建模方法论，这种方法需要一次性整合企业数据、初期建设难度大，但是有利于对业务的长期支持、特别是对业务变化的支持。

范式建模方法对数据模型的要求非常高，在建设过程中，博时基金结合自身的经验积累对Teradata的金融行业数据模型进行改进，形成适合基金行业的数据模型，填补了行业空白。

事实证明，博时基金的选择是正确的。“我们仅用不到九个月就完成了第一期建设，完全实现项目计划的要求，完成了对18个内部系统和6个外部数据源的整合，外部数据源包括国内的资讯系统和全球的资讯系统。数据库表有1100多个，数据的字段有将近3万个。这么大的工作量，如果没有一定的套路来做，很难在这么短时间

完成，还可能做着做着就乱了。是正确的决策让我们顺利地完成了数据整合。”

对于博时基金大数据平台一期建设的成果，车宏原从以下几个方面来总结：

第一，数据整合效果明显。全部内部系统和外部数据源的数据入仓，为企业决策提供了完整的视图。领导驾驶舱可以通过简洁直观的可视化展现，及时掌握运营状况，销售、投研等多数据源的整合，形象化凸显的重点指标，同时可以实现交叉、分组、多维度数据钻取。

第二，建立指标中心，解决了以前系统规划的不足。过去，复杂指标数据都在软件代码中定义，而新的指标中心作为数据平台一个清晰的数据层次，涵盖了公司的主要业务范围和相关数据，对内部和外部数据进行指标化处理，支持维度和指标的自定义，解决了过去架构下重复开发、数据不一致、统计口径不一致、计算性能差等问题。

第三，形成了统一的报表中心。系统使用统一的查询报表工具，建立了完善统一的数据权限管理；将分析类的数据查询、数据报表进行统一管理，而日志类、明细类、操作型或业务系统单独使用的查询报表仍保留在原系统，降低业务系统间的耦合。平台进行整体的分层设计、维度设计和主题设计，提供常用的自助查询和自助报表，以及少量自助多维分析。

第四，支持多种数据服务方式。除了报表类固定形式的数据查询外，可以提供Excel插件、BI工具、软件开发工具访问等形式支持自助分析服务，满足开发和业务部门自助取数、自助分析、深度开发的需要。

接下来，博时基金大数据平台将继续做那些事情？车宏原透露，最高优先级的工作是对传统核心业务的数据支持。

除此以外，企业数据平台还可以支持基于人工智能等技术的创新探索。可以开展很多这样的课题，例如在更多的客户属性字段、更大数据量的处理能力、更高质量的客户数据基础上，采用人工智能技术进行客户细分和客户行为预测，从而进一步支持精准营销、客户服务和智能投顾等业务。

CCiD 赛迪

中国信息化周报

影 响 力 就 是 生 命 力

www.cio360.net

中国信息化周报（信息化时代）

微信号：CIO360

官网：www.cio360.net