



大数据时代运营商分析支撑域转型的实践与思考

刘南海, 雷蕾, 王睿

(中国移动通信集团广西有限公司, 广西 南宁 530022)

摘要: 大数据时代, 随着业务和管理模式向“数据驱动型”转变, 运营商分析支撑域的支撑模型和支撑模式也发生了转变。制定了分析支撑域规划, 如构建云化 ETL、MPP 数据库、能力服务中心、大数据运营支撑平台等 IT 基础设施, 实现转型。同时, 提出了分析支撑域在 IT、管理和核心竞争力方面的实施思路。

关键词: 分析支撑域; 大数据; Hadoop; MPP

中图分类号: TN915.07

文献标识码: A

doi: 10.11959/j.issn.1000-0801.2016226

Practice and thinking on the transition of telecom operator analysis support system in big data era

LIU Nanhai, LEI Lei, WANG Rui

China Mobile Group Guangxi Co., Ltd., Nanning 530022, China

Abstract: In big data era, with business and management mode changing to “data driven”, the support model and support pattern of telecom operator analysis support system has changed. With the ASS planning, cloud ETL, MPP DB, ability service center and big data operation platform were constructed. Also, implementation thinking in aspects of IT, management and core competitiveness of ASS were proposed.

Key words: analysis support system, big data, Hadoop, MPP

1 引言

伴随着电子信息技术的飞速发展, 计算机软硬件技术和网络不断更新换代, 关系到以互联网和数据信息处理为核心的行业信息化水平的日益提高。抽象表征人类个体及由个体组成的群体自身属性和外部行为的各类数据呈爆炸式增长, 推动人类社会进入大数据时代。以解决海量数据存储、计算、挖掘分析为核心的大数据技术的发展, 使得数据成为了一种全新的生产要素, 带动业务和管理模式的转变, 驱使工业经济向数据经济转型。

(1) 大数据技术的发展奠定了业务和管理模式转变的基础

“大数据”一般包含 4 个方面的意义: 数据容量 (volume) 大、数据类型 (variety) 多、数据处理速度 (velocity) 快、数据价值 (value) 密度低。“大数据技术”通常也针对这 4 个方面: 或解决数据存储效率低的问题; 或适应不同的数据类型 (结构化/半结构化/非结构化); 或采用不同的处理框架, 提升数据处理效率; 或使用不同的分析挖掘模型促进数据价值的转化。

大数据技术的不断发展使得数据将逐步成为与劳动力、土地等并列的生产要素。一方面, 数据作为对现实世界

的抽象和度量记载了各类物体的属性和之间的行为过程;另一方面,通过一定量的客观数据描述记载了事务发展的普遍规律,在一定条件下可以发掘成知识以供使用。这两个过程在实际业务和管理中的出现和发展,意味着业务和管理模式的转变。

(2)外部竞争和内部降本增效驱动业务和管理模式向“数据驱动型”转变

互联网企业 OTT(over the top)业务蚕食传统语音、短信收入,营收“剪刀差”“营改增”对利润产生的影响以及监管部门“大幅削减营销费用”的要求,倒逼掌握数据流动通道的运营商“降本增效”。数据作为新的生产要素融入现代化大生产的过程,势必促使业务和管理模式向“数据驱动型”转变。对于业务,可通过历史的销售数据挖掘产品和客户之间的潜在关系,用以指导产品销售,提升销售效率。对于管理,一方面精确衡量某个管理对象的静态、动态过程;另一方面基于历史预测未来,支持管理决策。

2 分析支撑域转型的实践

运营商务和管理模式向“数据驱动型”转变势必将对 IT 基础设施提出转型要求。作为企业内部进行数据采集、存储、分析、挖掘从而完成知识到价值转化的核心,分析支撑域的转型迫在眉睫。转型将涉及 IT 基础设施、管理、运营 3 个维度,其中,IT 基础设施的转型最为关键。通过分析支撑模型和支撑模式的转变,规划 IT 基础设施的演进并实施,支撑实际运营活动的开展,为大数据时代运营商分析支撑域持续转型奠定基础。具体如图 1 所示。

管理维度	· 组织是否采用了扁平化运作思路对支撑体系进行保障。 · 对于新技术的管理、论证是否有人去思考。 · 大数据模型逐渐成为驱动运营开展的核心,如何去管理? · ...
运营维度	· 运营越来越趋向于实时化,性能是否得到满足。是否能够闭环。 · 服务越来越趋向于自助化,效率上是否都得以满足。运营支撑能力是否能够灵活构建。 · 信息越来越多,数据质量如何得到保障? · ...
IT 维度	· 采用这么多新技术,系统是否足够稳健稳定。延展性如何? 响应是否及时。 · 软硬件资源是否能够满足技术扩展。 · 是否具备端到端元数据管理体系。 · ...

图 1 分析支撑域转型三维度

2.1 在公司 IT 架构中的位置

使用系统的 3 类角色(客户、客服代表、内部员工)分

别通过内外部门户和接触渠道与公司 IT 系统发生联系。公司内部 IT 系统按照承担的职责,将其划分为如下四大域。

(1)业务支撑域

包含客户关系管理(customer relationship management, CRM)子域和业务运营支撑系统(business and operation support system, BOSS)子域,面向产品、渠道、定价、促销等方面。

(2)运营支撑域

主要面向通信网络,既包含网络运营的功能,也包含对设备的维护和管理。

(3)管理支撑域

面向企业核心资产——人、财、物的配置、管理和效能管理。

(4)分析支撑域

面向上述三域提供专业的数据分析、决策支持。大数据技术引入前的主要系统是经营分析系统,利用接入业务支撑域的客户基础数据和计费账务数据提供定期数据分析报表和智能查询服务。

运营商 IT 分域构成如图 2 所示。

2.2 支撑模型和支撑模式的转变

支撑模型指业务需求、支撑能力要求和支撑方式的集合,是 IT 能力对业务需求进行匹配的抽象。通过分析确定支撑模型,对 IT 系统规划和实施具有指导意义,避免系统与业务目标的偏移;支撑模式指对一个具体需求内容的支撑过程的抽象,是 IT 人员使用 IT 能力实现业务目标的过程,更强调组织和管理。

业务和管理模式转向“数据驱动型”,对分析支撑域的支撑模型和支撑模式提出了新的要求。

对于支撑模型,在传统经营分析系统“报表展现”业务分析支撑模型的基础上,增加了“数据驱动运营”“数据价值输出”两类支撑模型。“数据驱动运营”面向企业内部,强调通过对历史数据的相关性进行挖掘形成知识,用知识实时分析当前数据,对后续未知进行预测,基于预测开展行动以体现知识和数据的价值。“数据价值输出”强调面向企业外部的数据开放,形成跨行业、跨企业的“数据驱动运营”,如图 3 所示。

对于支撑模式,由于“数据驱动型”的本质是从海量数据中挖掘知识并探索价值转化和实施途径,是一个持续“运行—评估—优化”过程。支撑模式也由传统的“需求—模型—设计—开发—测试—维护”模式向“问题—数据源—探索结果—识别模式—优化模型—假设—新问题”的

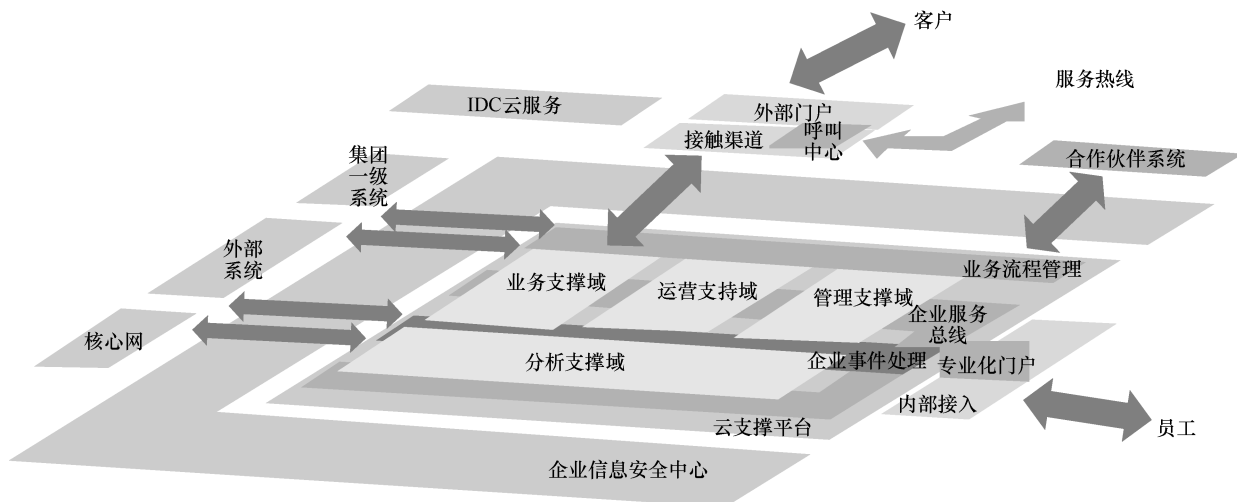


图2 运营商IT 分域构成

业务需求		支撑能力要求	支撑方式
业务分析 用于内部通报及决策	常规报表: 面向KPI指标和简单分析, 定期出报表, 监控业务发展趋势	传统数据仓库、报表能力	根据明确的需求, 在日常开发或工程中支撑
	手工分析: 由人工将提取后的数据再加工, 形成复杂多变的报表	基于传统仓库的分析工具, 强调操作性和算法封装	通过完善自助分析工具(SSA)支撑
	专题分析: 由专业人员的经验固化而成, 如多维财务分析	数据、模型、报表	根据明确的需求, 专题方式支撑
数据驱动运营 用于支撑公司内部运营	大数据关联分析: 在线、实时分析或挖掘, 支撑消费、位置类行为分析和预测	需要Hadoop、MPP等大数据处理技术	基于云化ETL、MPP、云挖掘能力, 以分析专题的形式支撑
	场景运营: 营销场景运营、四网协同运营等	需要对业务流程固化, 通过数据驱动流程各环节, 实现自动化运营	建设大数据运营支撑平台, 满足营销场景运营需求
数据价值输出 面向外部客户提供数据开放共享服务		在安全、可控的前提下, 向第三方提供能力服务, 打造新的商务模式	要求在营销场景运营能力的基础上改造

图3 分析支撑域的支撑模型

螺旋式模式转变,并且这种螺旋式支撑模式比起前者分界明显的“业务人员提需求,技术人员实现”更强调业务人员和技术人员的深度共同协作,具体如图4所示。

2.3 面向“数据驱动”的转型规划

分析支撑域的IT基础设施需适应支撑模型和支撑模式的转变。基于此,制定分析支撑域规划。支撑模型和支撑模式的分析经历了两个阶段:首先认识到“大数据关联分析”对数据存储、分析、处理方面的能力要求;然后才是“场景运营”对基础能力和应用整合形成的运营能力的要求。因此,分析支撑域规划也是两个阶段的过程。

2.3.1 规划演进一阶段

大数据技术引入之前,传统分析支撑域的范围极其有限。分析支撑系统和能力分散在BSS(business support

system)域、MSS(management support system)域、OSS(operation support system)域中。BSS域分析支撑相对集中,通过经营分析系统整合域内数据,形成统一客户画像,主要支撑域内业务分析;OSS域系统分析支撑能力相对分散,分专业建设,分析支撑能力相对较弱;MSS域同样存在分析支撑能力分散的问题,主要根据各类管理需求建设对应的系统,部分分析需求由经营分析系统支撑。

这种分析支撑能力的情形并不适应不断增长的数据存储、分析、处理方面的能力要求,主要体现在小型机架构下高扩容成本、离线分析架构无法支撑高并发量大数据的处理、分散数据访问、服务以及对四网协同、家庭宽带流量经营等跨域分析专题支撑不力等方面。为此,同时从跨域综合分析能力、大数据处理能力、需求和数据管理能力提

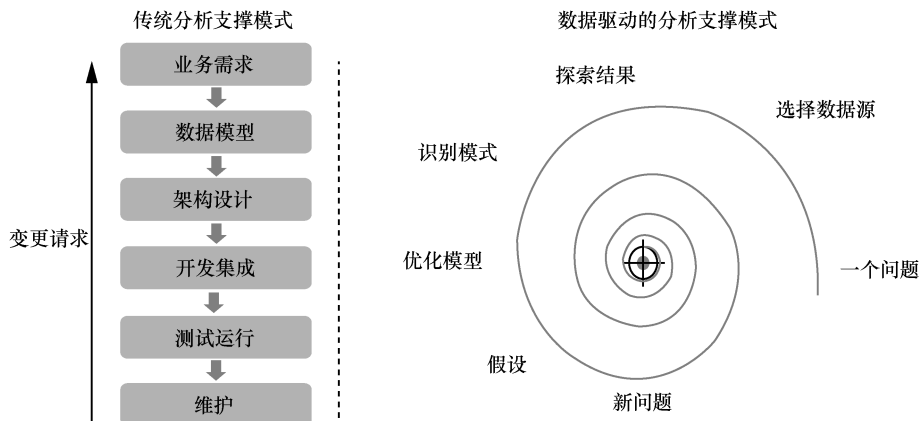


图4 分析支撑域的支撑模式转变

升方面着手,引入适用的新技术,按低成本、高效益的原则对决策分析的整个体系进行改造,提升能力、提高效益。

在逻辑架构上进行分层,统一数据中心、企业数据中心、能力服务中心、分析应用中心各司其职。在统一数据中心,引入基于 Hadoop 架构的云化 ETL,利用分布式文件存储降低成本,利用分布式批处理计算提升对数据源 ETL 过

程的执行效率,统一接入 BSS/OSS/MSS 三域数据;在企业数据中心,引入基于 MPP 的分布式数据库,利用分布式计算提升高度汇总数据关联计算的效率,利用无共享(share-nothing)架构提升扩容效率;在能力服务中心,面向上层应用抽象对底层数据操作和基础功能组件能力,支撑多个应用开发商开发不同的应用,实现应用的“百花齐放”。具体如图5所示。

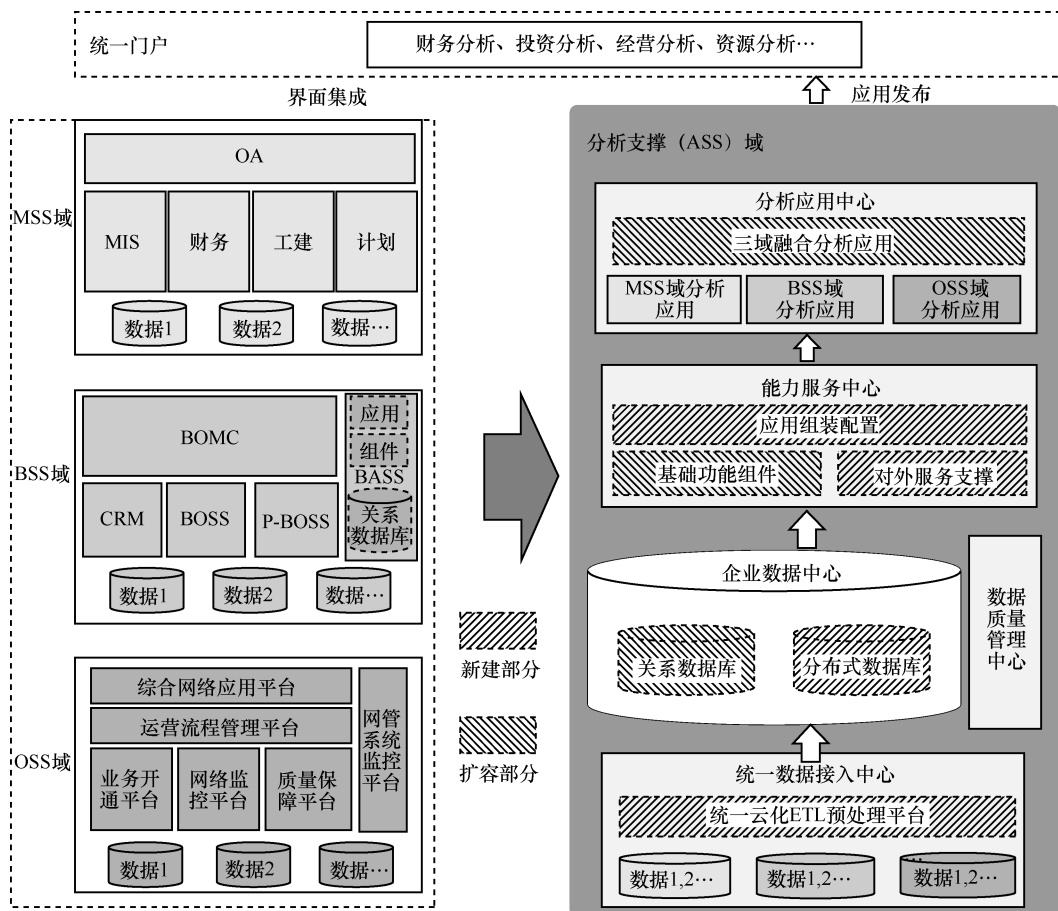


图5 分析支撑域规划演进一



2.3.2 规划演进二阶段

云化 ETL、MPP 数据库一定程度上提升了数据存储、分析、处理的基础能力,使得数据转化为知识成为可能,能力服务中心也为引入外部厂商开发应用提供了开放环境,促进了知识(各类应用的业务意义正是知识的体现)的“百花齐放”。

知识到价值的转化离不开运营。数据蕴含的知识通常只是表征一个事物的属性或者其活动过程的规律,需要通过运营才能转化为价值。比如,“挖掘出满足某些条件的客户有极大可能购买某产品”是知识,可以以客户标签的应用形式存在,需要提取客户清单,选择合适时机对其开展营销,产生了产品交易,收了客户的钱才形成价值。这个过程就是一种运营。

于是,为了提升知识到价值的转化效率,在一阶段的应用中心上规划了运营中心。运营中心一方面整合应用中心的知识;另一方面连接外部使能系统(如 BSS 域的销售渠道、OSS 域的控制用户网络服务策略的 PCRF 网元),打

通知识到价值的转换渠道,提供一站式的运营支撑。具体如图 6 所示。

2.4 IT 基础设施构建

遵循两阶段的规划演进,构建云化 ETL、MPP 数据库、能力服务中心、大数据运营支撑平台四大 IT 基础设施,支持“数据驱动”转型。

2.4.1 云化 ETL

ETL(extract-transform-load)指数据抽取、转换、装载的过程。作为分析支撑域的基础,能够按照具体规则将分散于各业务系统的数据进行轻度汇总后集成入数据仓库,为上层分析应用提供数据支撑。

大数据技术引入前,传统的 ETL 过程基于“小型机+盘阵”的架构,由于与数据仓库中的高度汇总以及面向应用的数据计算共享计算和存储能力,在数据量激增的大数据时代出现性能瓶颈。并且 CPU 和存储的扩容与性能的提升已经出现强烈的非线性关系,如图 7 所示。因此迫切需要引入 MPP 架构的 ETL 能力。

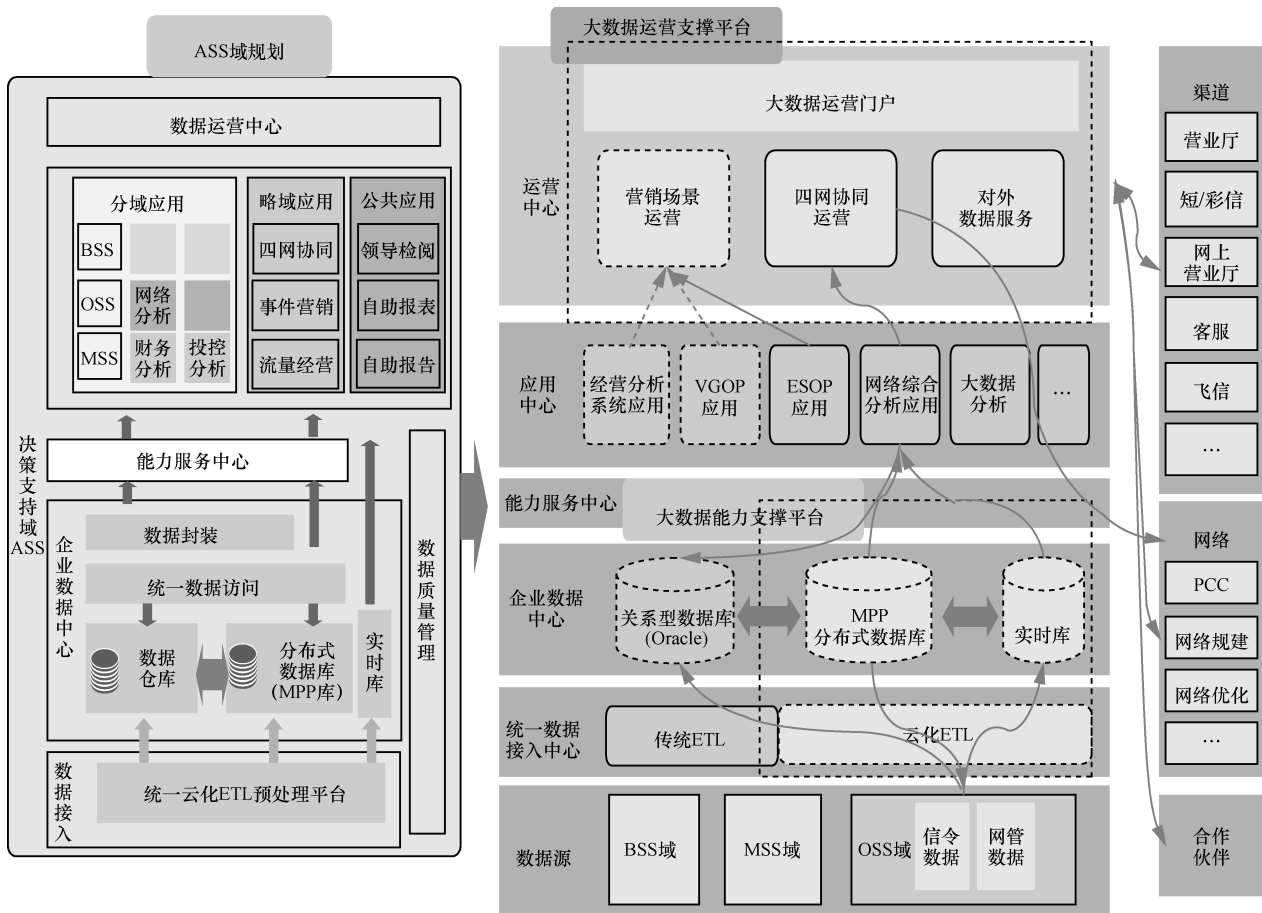


图 6 分析支撑域规划演进二

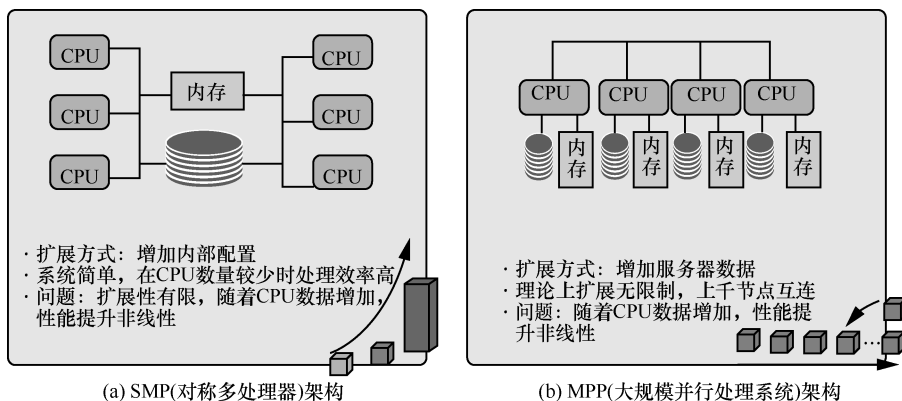


图7 SMP和MPP架构

云化 ETL 包含两个层面,“云化”指的是采用 MPP 方式的硬件架构,并且在软件框架上采用了“云计算”类相关的技术,适用于 ETL 的过程。云化 ETL 的核心是 Hadoop,如图 8 所示。



图8 云化 ETL 相关组件结构

(1)Hadoop

是一个分布式存储和计算的框架,广泛用于海量数据的存储和处理。包含 HDFS(hadoop distributed file system)、YARN(yet another resource negotiator)、MapReduce、HBase、Hive 等组件。

(2)HDFS

是一个适合运行在通用硬件之上的、具备高度容错特性、支持高吞吐量数据访问的分布式文件系统,适合大规模数据集应用。

(3)YARN

是一个分布式的资源管理系统,用以提高分布式集群环境下的资源利用率,这些资源包括内存、I/O、网络、磁盘等。在它上面可以部署 MapReduce 等各种分布式计算框架。

(4)MapReduce

是分布式大型计算框架,支持 MapReduce 编程模型,高度适应数据处理的 ETL 过程。

(5)HBase

是面向列(column-oriented)、适合存储海量非结构化数据或半结构化数据、高可靠、高性能、可灵活扩展伸缩、

支持实时数据读写的分布式存储系统。

(6)Hive

是建立在 Hadoop 之上的数据仓库解决方案,支持将结构化的数据文件映射为一张表,提供 HQL(hive SQL)实现方便高效的数据查询,底层数据存储 in HDFS 上。Hive 的本质是将 HQL 转换为 MapReduce 程序去执行,使不熟悉 MapReduce 的用户很方便地利用 HQL 进行数据 ETL 操作。

各组件按照图 9 的方式以具体的进程实例分布在各种物理节点上,形成具有高度扩展型的 IT 基础设施。BDI 管理集群向外提供图形化的集成管理界面,可以对 ETL 的数据处理流程进行管理和配置。管理阶段中的 3 个主节点作为整个分布式集群管理的核心,维持资源调用的一致性和高可靠性,确保工作节点的工作。同时,HDFS、YARN、HBase、Hive 的管理组件进程也部署在管理节点上,与部署在工作节点上的工作组件等一系列静态进程构成统一的 ETL 处理工作环境。依靠这些进程,部署和分配 ETL 数据处理任务后,在工作节点产生一系列的动态进程(由 HDFS 等组件的管理进程生成)完成具体的数据处理任务。

云化 ETL 的硬件部署情况如图 10 所示。本期部署了 72 个节点的集群,包含 4 个 BDI 节点、4 个管理节点和 64 个工作节点。节点内部使用万兆网络相连,外部管理控制使用吉比特网络。特别地,云化 ETL 处理后的轻度汇总数据将通过万兆交换机与 MPP 数据库集群相连,由 MPP 数据库集群完成面向分析应用的高度汇总过程。

2.4.2 MPP 数据库

如果说云化 ETL 解决的是传统数据仓库 SMP 共享存储架构下,ETL 和轻度汇总数据存储和处理过程性能不足的问题。对应地,传统数据仓库的高度汇总和关联分析需要依靠 MPP 架构的数据库来解决。因为云化 ETL 使用的

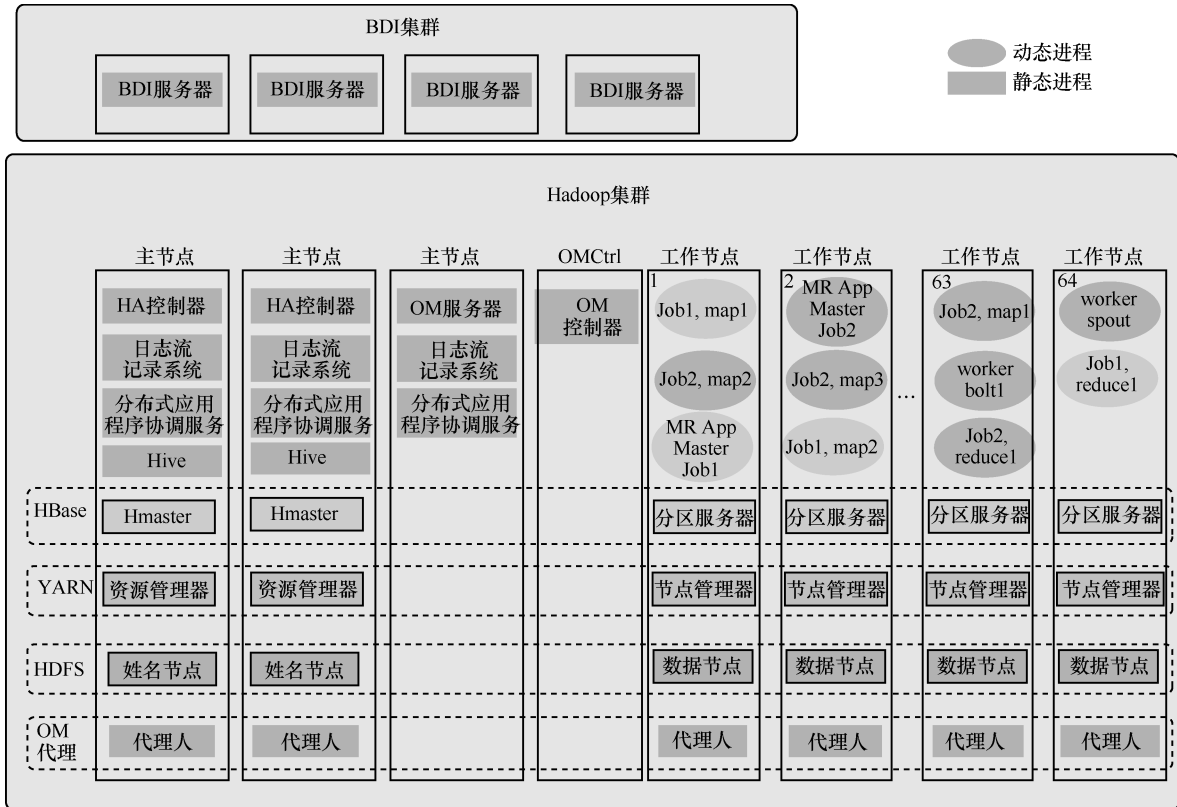


图9 云化ETL组件进程部署

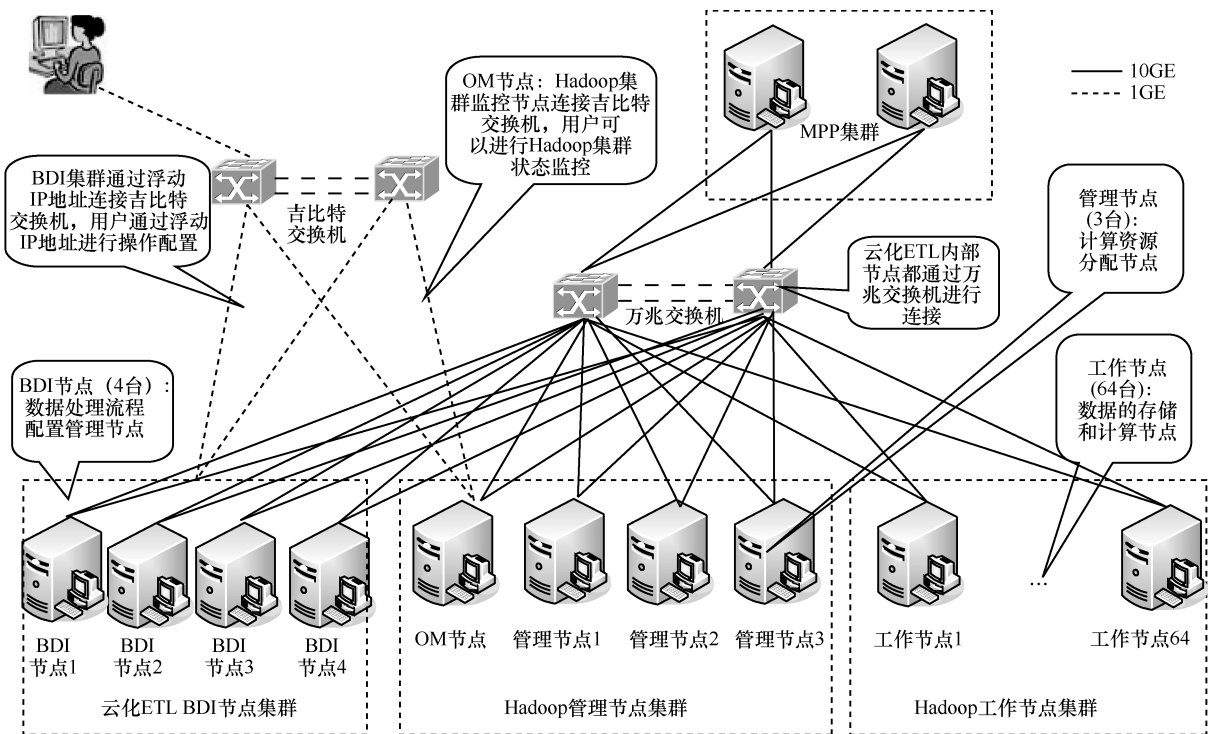


图10 云化ETL硬件部署

Hadoop 的 MapReduce 过程更多偏向于离线数据处理,并不适用于多表关联分析。

实际中采用的是 GBase 8a 产品,如图 11 所示,这是一个无管理节点的节点对等的架构,每个节点内的 CPU 不能直接访问另一个节点的内存,节点之间的信息交互通过节点互联网实现。这种无共享(share-nothing)架构,使得资源的水平扩展比较容易实现。

在各节点内部,通过 GCluster 组件管理元数据,对客户端的 SQL 请求映射分布式地执行计划并调度实施,各节点计算完成后,将各自部分的结果汇总在一起得到最终的结果,返回客户端。GCluster 仅能访问节点内部的数据,跨节点数据关联由 GCluster 之间通过高速网络进行。具体如

图 12 所示。

MPP 数据库在各节点实现多表关联计算的过程中,性能的核心在于选择合适的数据分布键使得需要进行关联的表数据能均匀分布在所有节点相同的数据分区中,这将减少扩节点跨分区的数据连接,极大发挥多节点并行处理的作用。

如图 13 所示,对于 cust 表和 sales 表的数据,多需要根据 cust_id 进行关联。那么将 cust_id 作为分布键对 cust 表和 sales 表的数据进行散列分布后,cust_id 相同的数据将被分配至相同的节点或者是节点相同的数据分区。如此进行二表关联时,仅需要在分区和节点内部进行关联计算,如此极大降低了跨节点/分区的网络数据消耗。

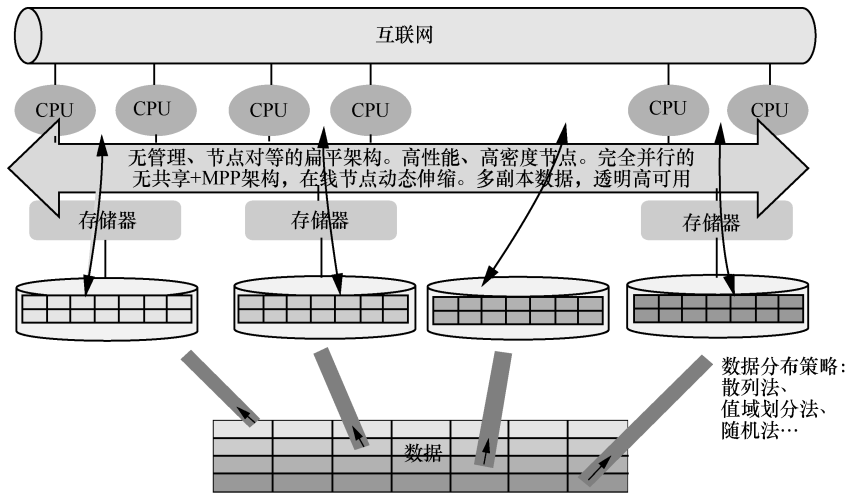


图 11 MPP 数据库架构

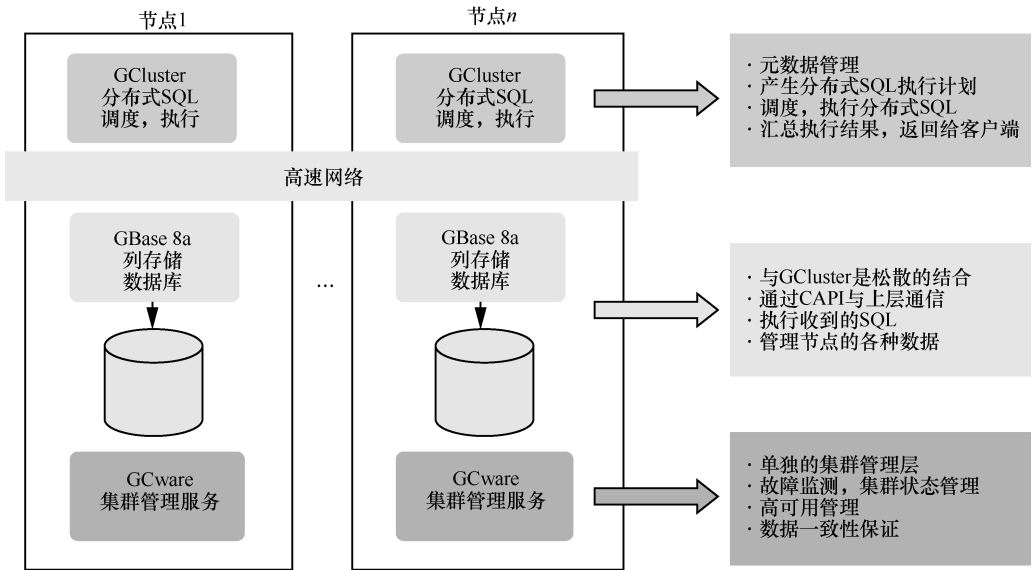


图 12 MPP 数据库对等节点内部组件

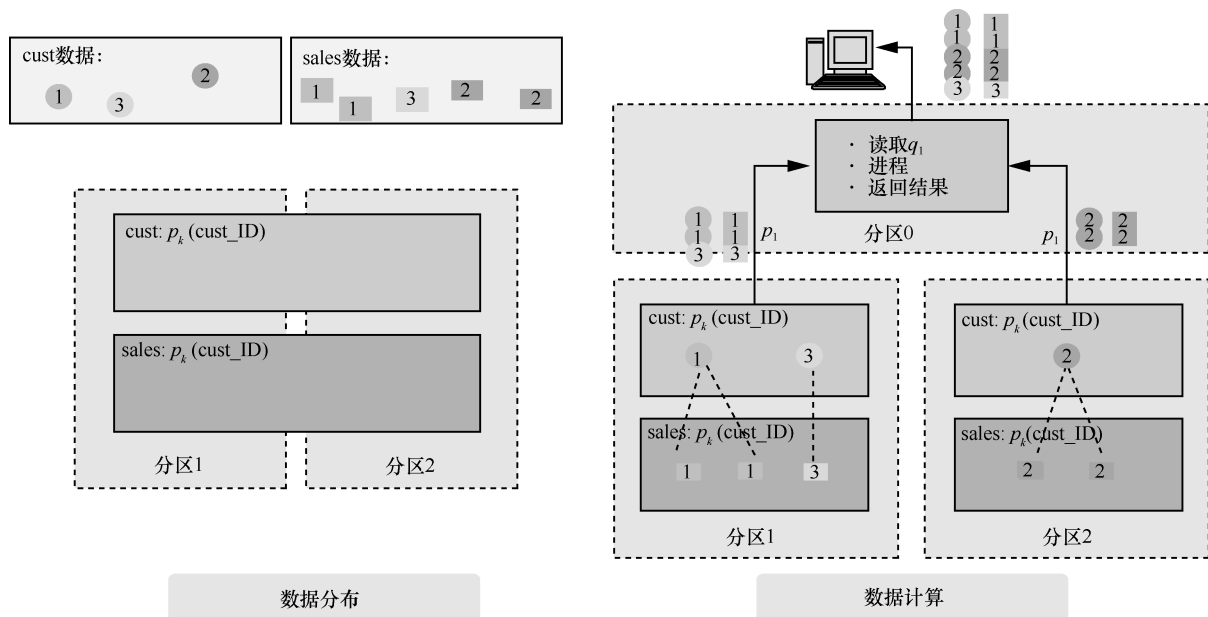


图 13 MPP 数据库数据分布和关联计算

实际部署中,按照 46 个计算节点、2 个数据分发节点的方式部署 MPP 数据库集群。48 个节点内部使用万兆网络相连,其中,2 个数据分发节点与云化 ETL 相连,外部应用和监控管理平台通过吉比特网络接入,如图 14 所示。

2.4.3 能力服务中心

云化 ETL 和 MPP 数据库针对的是数据层面的问题,数

据按照一定的规则面向业务领域进行了构建。为了实现数据到知识的转化,需要开发各类分析应用,这个应用开发的过程通常是极其个性化和专业化的,也有不同的开发商专注于某个具体的分析应用领域。出于降本增效的考虑,引入开发商之间的竞争,实现应用的“百花齐放”,因此构建能力服务中心,向应用开发商提供统一的数据服务,如图 15 所示。

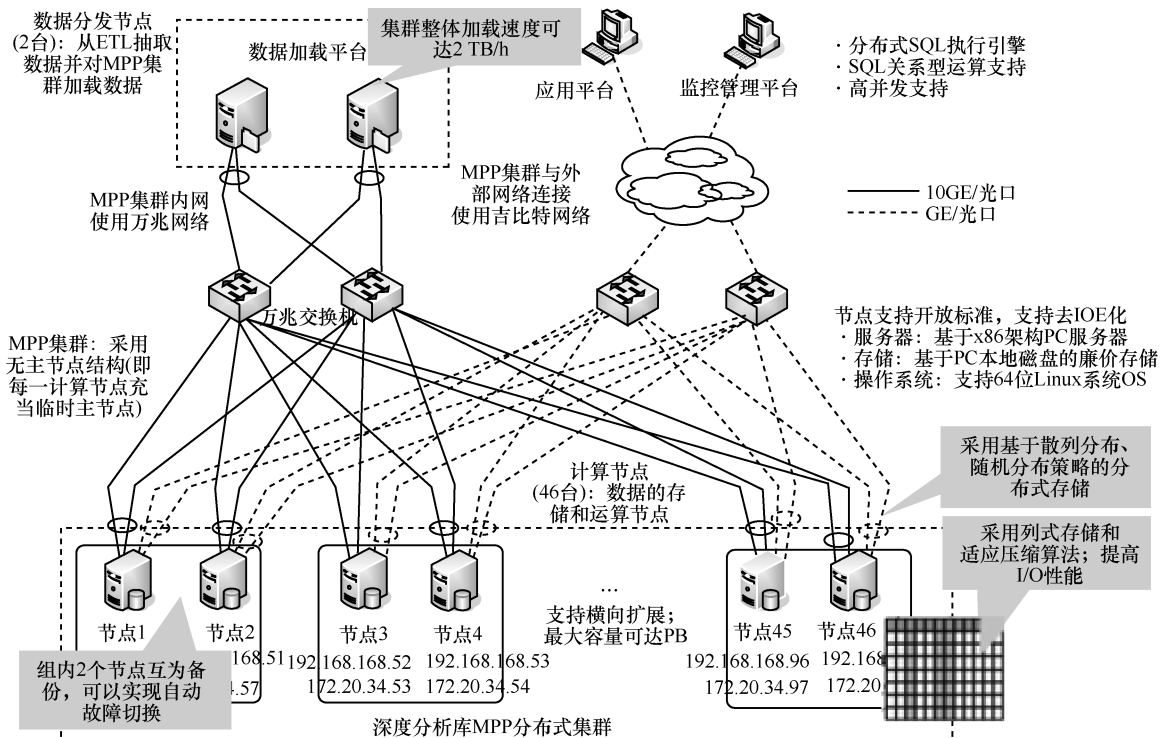


图 14 MPP 数据库部署

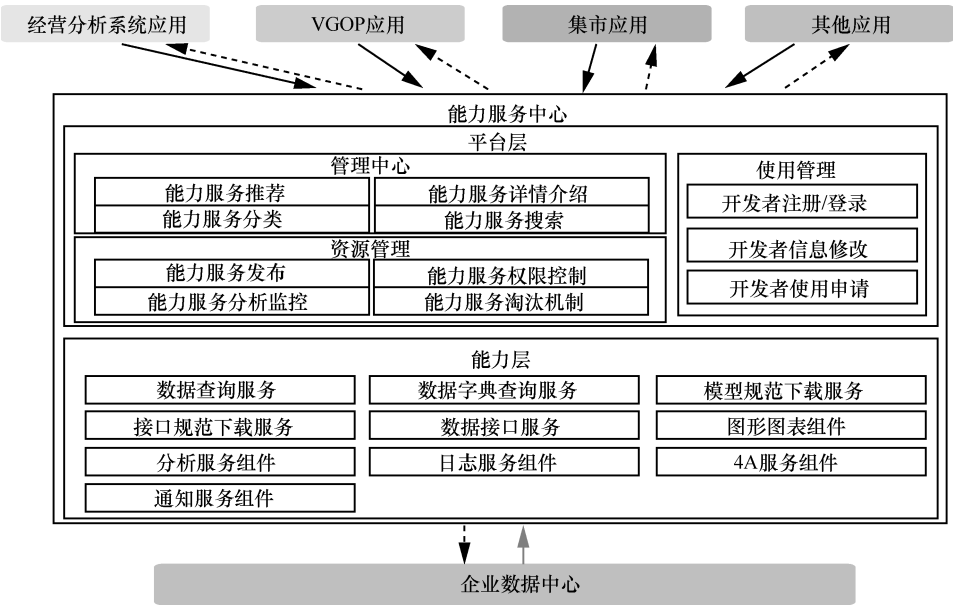


图 15 能力服务中心的定位

能力服务中心具体的功能架构如图 16 所示，分为展示层、业务层、服务层和数据层。

展示层、业务层主要负责展现能力服务资源管理、能力服务使用管理、能力服务管理中心的界面，控制页面跳转。

服务层提供统一的开发规范和数据服务，支持多种形式的的能力服务组件；提供用户权限、应用权限管理及鉴权机制，确保数据安全性及服务可靠性；引入负载均衡及基于内存的数据缓存机制，提高查询效率，保障服务的及时响应。

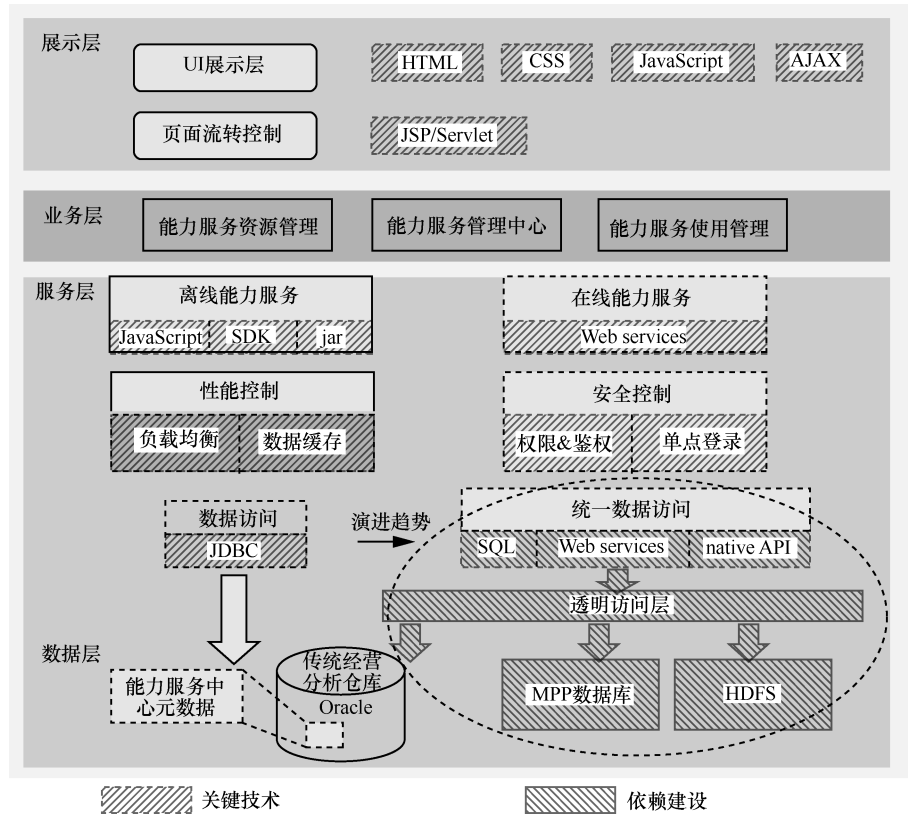


图 16 能力服务中心的功能架构



数据层封装底层数据仓库,包括传统 Oracle 数据库、Hadoop 集群和 MPP 数据库,通过透明数据层访问异构数据库。

2.4.4 大数据运营支撑平台

云化 ETL 和 MPP 数据库解决基础数据能力问题,各类应用揭示了数据所蕴含的知识。知识到价值的转化需要通过运营来实现。事实上,甚至在未引入大数据技术前,就已经有基于传统 Oracle 关系型数据库的精准营销运营场景,业务人员通过客户画像指定业务口径,从经营分析系统中提取客户号码清单,给客户群发推荐短信,客户接到短信后去营业厅办理业务完成销售,实现了数据到价值的转化。在这个过程中,客户数据客观存在,也通过客户画像(应用)形成知识,如果业务人员不指定业务口径(系统知识与人的知识结合),不提取客户号码去群发短信,客户没有进行业务订购,不形成价值。

大数据运营支撑平台面向具体的运营场景,比如面向市场的大数据销售模板运营、面向网络的四网协同运营、面向外部的对外数据服务等。通过连接各应用的功能模块将涉及的运营环节进行整合,提供一站式运营,如图 17 所示。

由于目前仅大数据销售模板运营较为成熟,所以大数据运营支撑平台当前主要面向数据驱动营销进行构建,后续扩展支撑其他方面的运营。

“以产品为抓手,以销售任务为导向”的大数据销售模板运营的业务模板包含客户(customer)、产品(product)、渠道(channel)、时机(time)4个要素。强调业务(产品)、客户、场景(时间、空间)及营销话术等要素协同一体,推动业务(产品)产生增量效益。

基于此业务模型,按照图 18 所示的功能框架构建大

数据运营支撑平台。对产品/内容库、客户标签库、营销平台、事件库、营销渠道进行功能优化改造,并用大数据运营门户将运营流程串联起来,提供面向运营的一站式支撑,提升“数据—知识—价值”的转化效率。

2.5 运营效果

在大数据运营支撑平台整合支撑销售模板运营之前,在没有大数据技术支撑的传统 Oracle 数据库和大量定制化配置开发、手工数据传递等的支撑下,销售模板运营取得了不错的效果。见表 1,从 2013 年 7 月到 2014 年 7 月底,累计营销 1.25 亿次,成功营销客户 242.7 万人,累计销售收入达 3 348 万元。

有理由相信,随着分析支撑域云化 ETL、MPP 数据库、能力服务中心、大数据运营支撑平台等 IT 基础设施形成生产力,数据的存储、分析、处理能力增强,知识的提炼、知识到价值的转化过程将得到固化,更多的运营场景被发掘,结合组织和管理模式的配套,将有力地驱动运营商业务和管理模式向“数据驱动型”转变。

3 分析支撑域转型的思考

大数据技术的发展、企业外部竞争和内部管理的要求驱动着运营商业务和管理模式向“数据驱动型”转变。相对应,分析支撑域的支撑模型和支撑模式也在发生转变。基于这种转变,制定分析支撑域转型规划,构建云化 ETL、MPP 数据库、能力服务中心、大数据运营支撑平台等 IT 基础设施。这些平台和功能已陆续上线,需要运营和使用才能发挥能力,形成生产力。系统和平台的上线仅仅是整个分析支撑域转型乃至业务和管理模式转型的起点。

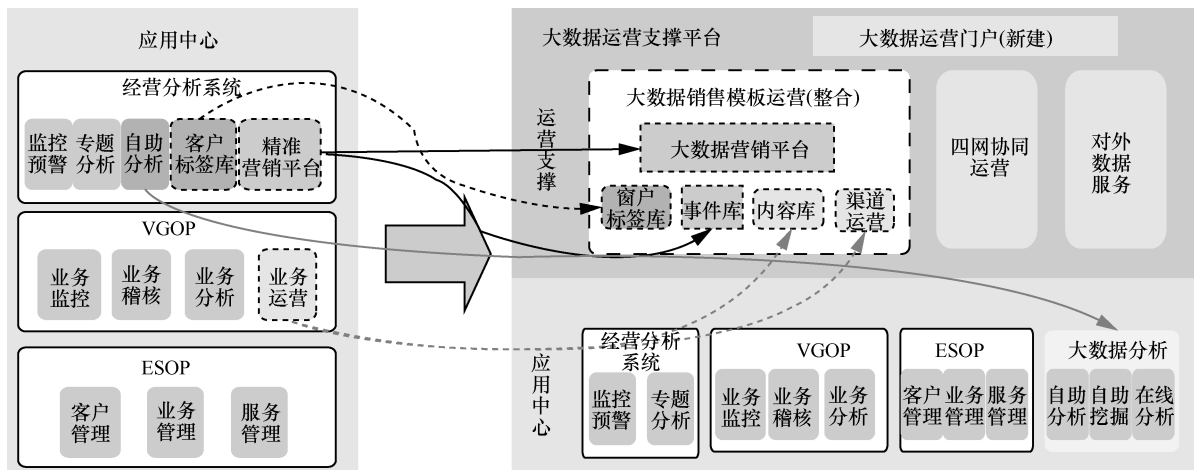


图 17 大数据运营支撑平台的定位

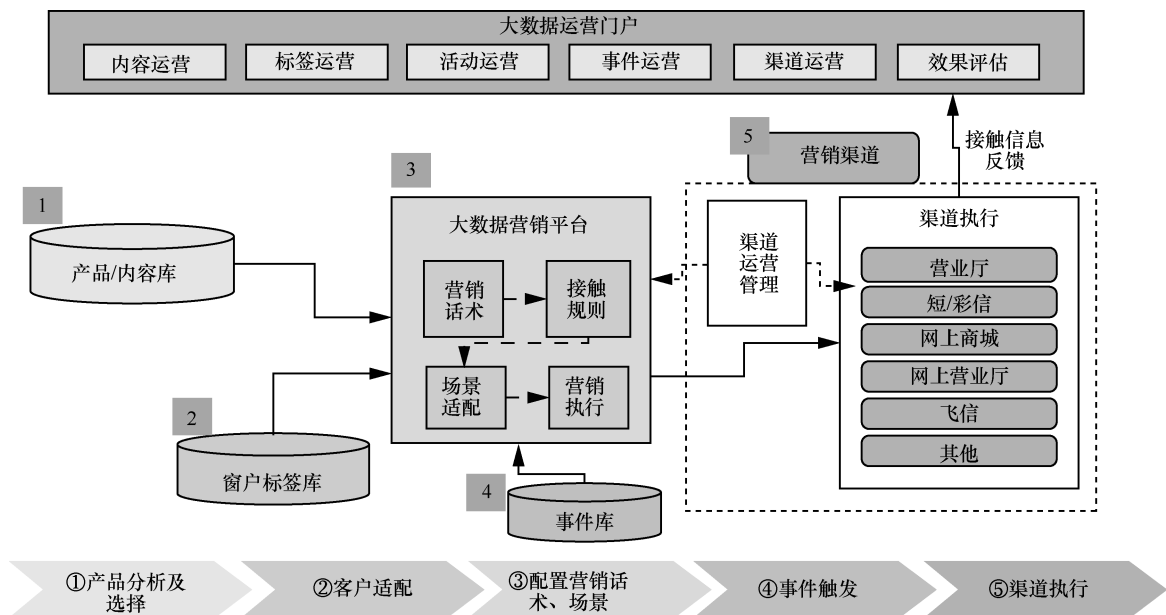


图 18 大数据运营支撑平台功能框架

表 1 大数据技术使用前销售模板的运营效果

销售模板类型	销售模板数量/个	短信下发信量/万条	营销响应客户量/万人	营销成功客户量/万人	营销响应率	营销成功率	当月收入/万元	累计收入/万元
流量套餐增值业务类销售模板	47	2 705.60	219.82	209.06	8.12%	7.73%	841.72	2 854.60
非流量套餐增值业务类销售模板	17	945.64	1.00	2.39	0.11%	0.25%	1.47	4.34
语音产品类销售模板	8	6 365.55	24.35	30.34	0.38%	0.48%	126.891 2	489.686 9
应用客户端类销售模板	5	1 724.79	46.64	0.64	2.70%	0.04%	0	0
终端类销售模板	4	176.46	0.16	0	0.09%	—	0	0
服务提醒类销售模板	10	573.70	5.61	0.27	0.98%	0.05%	0	0
合计	91	12 491.74	297.58	242.70	2.38%	1.94%	970.09	3 348.63

3.1 实时处理

传统分析支撑域的数据源多以日为周期(每天可以从数据源系统获得前一天的数据),基于历史数据进行统计分析,对实时性要求不高,无法支撑如“实时/准实时事件驱动营销”之类的场景。引入云化 ETL 和 MPP 数据库后,一定程度上提升了数据提取、汇总和查询的效率,仍无法支撑实时性要求高的场景。需要规划引入流处理、内存数据库等 IT 基础设施,并且积极推动前端数据源系统向实时消息机制的转型。

3.2 数据资产化

数据驱动的运营使得数据成为一种资产。需要考虑以资产的方式进行管理,协调考虑数据生命周期、数据价值

评估、数据口径、数据存储、知识挖掘、数据使用、数据安全等各方面。

3.3 模型挖掘

数据需要通过挖掘才能成为知识,知识通常以一定模型算法的形式存在,这正是分析支撑域的核心竞争力。大数据技术使得对海量数据的挖掘处理成为可能,迫切需在组织机构、管理流程、人才培养方面统筹规划,建立模型挖掘团队,开展挖掘研究,提升核心掌控力。

4 结束语

大数据时代,业务和管理模式逐步向“数据驱动”转型,对电信运营商的 IT 支撑提出了转型要求。广西

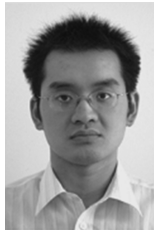


移动在技术架构上采用“Hadoop”和“MPP”的混搭,实现数据生命周期各种能力的服务化改造以及“数据驱动营销”作为应用层的优先切入,确保 IT 支撑转型与核心业务支撑的顺利衔接,转型实践路径具有推广借鉴价值。

参考文献:

- [1] 郑毅. 证析——大数据与基于证据的决策 [M]. 北京: 华夏出版社, 2012.
ZHENG Y. Analytics: on big data and evidence-based decision[M]. Beijing: Huaxia Publishing House, 2012.
- [2] BILL F. 驾驭大数据[M]. 北京: 人民邮电出版社, 2013.
BILL F. Taming the big data tidal wave [M]. Beijing: Posts & Telecom Press, 2013.
- [3] 徐子沛. 大数据: 正在到来的数据革命 [M]. 桂林: 广西师范大学出版社, 2013.
XU Z P. The big data revolution [M]. Guilin: Guangxi Normal University Press, 2013.
- [4] 刘军. Hadoop 大数据处理[M]. 北京: 人民邮电出版社, 2013.
LIU J. Hadoop big data processing [M]. Beijing: Posts & Telecom Press, 2013.

[作者简介]



刘南海(1982-),男,中国移动通信集团广西有限公司信息系统部 IT 专家、工程师,主要承担 BI/大数据系统的 IT 规划实施建设及运营工作,主要研究方向为大数据 DaaS、PaaS、SaaS 的服务化和企业级数据治理。



雷蕾(1978-),女,中国移动通信集团广西有限公司信息系统部规划建设室经理、工程师,主要研究方向为云计算、大数据。



王睿(1980-),男,中国移动通信集团广西有限公司信息系统部大数据开发支撑室经理、工程师,主要研究方向为大数据、IT 架构、云计算、网络安全管控。