

大数据之路

阿里巴巴大数据实践

|| 阿里巴巴数据技术及产品部 著 ||

探索

架构

认知



中国工信出版集团



电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
http://www.phei.com.cn

大数据之路

阿里巴巴大数据实践

|| 阿里巴巴数据技术及产品部 著 ||

电子工业出版社

Publishing House of Electronics Industry

北京·BEIJING

内 容 简 介

在阿里巴巴集团内，数据人员面临的现实情况是：集团数据存储已经达到 EB 级别，部分单张表每天的数据记录数高达几千亿条；在 2016 年“双 11 购物狂欢节”的 24 小时中，支付金额达到了 1207 亿元人民币，支付峰值高达 12 万笔/秒，下单峰值达 17.5 万笔/秒，媒体直播大屏处理的总数据量高达百亿级别且所有数据都需要做到实时、准确地对外披露……巨大的信息量给数据采集、存储和计算都带来了极大的挑战。

《大数据之路——阿里巴巴大数据实践》就是在此背景下完成的。本书中讲到的阿里巴巴大数据系统架构，就是为了满足不断变化的业务需求，同时实现系统的高度扩展性、灵活性以及数据展现的高性能而设计的。

本书由阿里巴巴数据技术及产品部组织并完成写作，是阿里巴巴分享对大数据的认知，与生态伙伴共创数据智能的重要基石。相信本书中的实践和思考对同行会有很大的启发和借鉴意义。

本书著作权归淘宝（中国）软件有限公司所有，未经许可，不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有，侵权必究。

图书在版编目（CIP）数据

大数据之路：阿里巴巴大数据实践 / 阿里巴巴数据技术及产品部著. —北京：电子工业出版社，2017.7

ISBN 978-7-121-31438-4

I. ①大… II. ①阿… III. ①企业管理—数据管理 IV. ①F272.7

中国版本图书馆 CIP 数据核字（2017）第 094934 号

策划编辑：张彦红

责任编辑：葛 娜

印 刷：三河市双峰印刷装订有限公司

装 订：三河市双峰印刷装订有限公司

出版发行：电子工业出版社

北京市海淀区万寿路 173 信箱 邮编：100036

开 本：720×1000 1/16 印张：21 字数：338 千字

版 次：2017 年 7 月第 1 版

印 次：2017 年 9 月第 4 次印刷

印 数：8501~11000 册 定价：79.00 元

凡所购买电子工业出版社图书有缺损问题，请向购买书店调换。若书店售缺，请与本社发行部联系，联系及邮购电话：（010）88254888，88258888。

质量投诉请发邮件至 zlts@phei.com.cn，盗版侵权举报请发邮件至 dbqq@phei.com.cn。

本书咨询联系方式：（010）51260888-819，faq@phei.com.cn。

主编：阿里巴巴数据技术及产品部

阿里巴巴数据技术及产品部，定位于阿里集团数据中台，为阿里生态内外的业务、用户、中小企业提供全链路、全渠道的数据服务。作为阿里大数据战略的核心践行者，致力于让“大数据，赋能商业，创造价值”。经过多年的实践，数据技术及产品部已经构建了从底层的数据采集、数据处理，到挖掘算法、数据应用服务以及数据产品的全链路、标准化的大数据体系。通过这个体系，超过 EB 级别的海量数据能够高效融合，并以秒级的响应速度，服务并驱动阿里巴巴自身的业务和外部千万用户的发展。现在，阿里巴巴数据技术及产品部正通过技术和产品上的创新，探索全域数据的价值，将阿里在大数据上沉淀的能力对外分享，为各行各业的发展带来更多可能性。

“阿里巴巴数据技术及产品部”官方网站：

<https://dt.alibaba.com>

社交账号：Ali_Data



本书由阿里巴巴数据技术及产品部编著，特别感谢阿里巴巴集团各业务部门对本书编著工作给予的支持！

参与本书编写的有：罗金鹏、王静、王赛、王伟、王永伟、张磊等。

(注：以上排名不分先后)



序

大数据是什么？在过去的5年里，恐怕没有另外一个词比大数据更高频；也没有另外一个概念如大数据一样，被纷繁解读，著书立说。有趣的是，作为距离大数据最近的公司之一——尽管我们的初心或许和大数据没有直接关系——在关于大数据的理论和概念的争论中，阿里巴巴却鲜有高谈阔论。

因为自知而敬畏，因为敬畏而谦逊。甚至在大数据这个概念出现很久之前，阿里巴巴就不得不直面、认知、探索，并架构和大数据有关的一切。数据作为一个生态级的平台企业最直接的沉淀，亦是最基本的再生产资料。如果没有基于大数据的人工智能的应用，淘宝根本不可能面对每天亿级的用户访问数量。因此，仅仅因为本能，阿里巴巴一开始就自然生长在这样一个数据的黑洞中，并且被越来越多、越来越密集的数据风暴裹挟。阿里巴巴在大数据方面所做的各种艰苦努力，其实就是力图对抗这种无序和复杂的熵增，从中梳理结构，提炼价值。

这是一个历经磨炼、也卓有成效的长期过程。如书中所提到的，阿里巴巴不仅数据量超宇宙级，而且更是因为业务场景的复杂和多元化，其面对着甚至超过 Google 和 Facebook 的更复杂的难题。大部分时候，阿里巴巴都是在无人区艰难跋涉。每一组功能和逻辑，每一套架构与系统，都与业务和场景息息相关。这个黑洞膨胀之快，以至于大部分时候都是在出现痛点从而刺激了架构升级。换言之，大数据系统——如果我们非要用一个系统去描述的话——其复杂度之高，是几乎不可能在一开始就完整和完美地进行自上而下定义和设计的。从需求→设计→迭代→

升华为理论，在无数次的迭代进化中，我们对大数据的理解才逐渐成形，慢慢能够在将数据黑洞为我所用的抗争中扳回一局。

这个系统生长和进化的过程实际上已经暗暗揭示了阿里巴巴对大数据真髓的理解。大、快、多样性只是表象，大数据的真正价值在于生命性和生态性。阿里巴巴称之为“活数据”。活数据是全本记录、实时驱动决策和迭代，其价值是随着使用场景和方式动态变化的。简单地把数据定义为正/负资产都太简单。数据也不是会枯竭的能源。数据可以被重复使用，并在使用中升值；数据与数据链接可能会像核反应一样产生价值的聚变。数据使用和数据聚变又产生新的数据。活数据的基础设施就需要来承载、管理和促进这个生态体的最大价值实现（以及相应的成本最小化）。丰富的数据形式、多样化的参与角色和动机，以及迥异的计算场景都使得这个系统的复杂度无限升级。阿里巴巴的大数据之路就是在深刻理解这种复杂性的基础上，摸索到了一些重要的秩序和原理，并通过技术架构来验证和夯实。

如果说互联网实现了人人互联和通信，并没有深度地协同计算，那么这样的一个大数据平台和架构就是一张升级的、智能的互联网。这是人类自己设计出来的复杂的信息处理系统，同时也将是真正意义上人类智力大联合的基础设施。这是一个伟大的蓝图，我们敬畏其复杂度和潜能。《大数据之路——阿里巴巴大数据实践》便是阿里巴巴分享对大数据的认知、与世界共创数据智能的重要基石。数据技术及产品部作为阿里巴巴集团的数据中台，一直致力为阿里巴巴集团内、外提供大数据方面的系统服务，承载了阿里巴巴集团大数据梦想至关重要的数据平台建设。相信他们的实践和思考对同行会有很大的启发和借鉴意义。

曾鸣教授

阿里巴巴集团学术委员会主席

湖畔大学教务长

2017年5月

轻松注册成为博文视点社区用户 (www.broadview.com.cn), 扫码直达本书页面。

- **提交勘误**: 您对书中内容的修改意见可在 [提交勘误](#) 处提交, 若被采纳, 将获赠博文视点社区积分 (在您购买电子书时, 积分可用来抵扣相应金额)。
- **交流互动**: 在页面下方 [读者评论](#) 处留下您的疑问或观点, 与我们和其他读者一同学习交流。

页面入口: <http://www.broadview.com.cn/31438>



目 录

| | |
|------------------------|----|
| 第 1 章 总述 | 1 |
| 第 1 篇 数据技术篇 | |
| 第 2 章 日志采集 | 8 |
| 2.1 浏览器的页面日志采集 | 8 |
| 2.1.1 页面浏览日志采集流程 | 9 |
| 2.1.2 页面交互日志采集 | 14 |
| 2.1.3 页面日志的服务器端清洗和预处理 | 15 |
| 2.2 无线客户端的日志采集 | 16 |
| 2.2.1 页面事件 | 17 |
| 2.2.2 控件点击及其他事件 | 18 |
| 2.2.3 特殊场景 | 19 |
| 2.2.4 H5 & Native 日志统一 | 20 |
| 2.2.5 设备标识 | 22 |
| 2.2.6 日志传输 | 23 |
| 2.3 日志采集的挑战 | 24 |
| 2.3.1 典型场景 | 24 |
| 2.3.2 大促保障 | 26 |
| 第 3 章 数据同步 | 29 |
| 3.1 数据同步基础 | 29 |
| 3.1.1 直连同步 | 30 |

| | | |
|--------------|----------------|-----------|
| 3.1.2 | 数据文件同步 | 30 |
| 3.1.3 | 数据库日志解析同步 | 31 |
| 3.2 | 阿里数据仓库的同步方式 | 35 |
| 3.2.1 | 批量数据同步 | 35 |
| 3.2.2 | 实时数据同步 | 37 |
| 3.3 | 数据同步遇到的问题与解决方案 | 39 |
| 3.3.1 | 分库分表的处理 | 39 |
| 3.3.2 | 高效同步和批量同步 | 41 |
| 3.3.3 | 增量与全量同步的合并 | 42 |
| 3.3.4 | 同步性能的处理 | 43 |
| 3.3.5 | 数据漂移的处理 | 45 |
| 第 4 章 | 离线数据开发 | 48 |
| 4.1 | 数据开发平台 | 48 |
| 4.1.1 | 统一计算平台 | 49 |
| 4.1.2 | 统一开发平台 | 53 |
| 4.2 | 任务调度系统 | 58 |
| 4.2.1 | 背景 | 58 |
| 4.2.2 | 介绍 | 59 |
| 4.2.3 | 特点及应用 | 64 |
| 第 5 章 | 实时技术 | 68 |
| 5.1 | 简介 | 69 |
| 5.2 | 流式技术架构 | 71 |
| 5.2.1 | 数据采集 | 72 |
| 5.2.2 | 数据处理 | 74 |
| 5.2.3 | 数据存储 | 78 |
| 5.2.4 | 数据服务 | 80 |
| 5.3 | 流式数据模型 | 80 |
| 5.3.1 | 数据分层 | 80 |
| 5.3.2 | 多流关联 | 83 |
| 5.3.3 | 维表使用 | 84 |
| 5.4 | 大促挑战&保障 | 86 |

| | | |
|--------------------|------------------|------------|
| 5.4.1 | 大促特征 | 86 |
| 5.4.2 | 大促保障 | 87 |
| 第 6 章 | 数据服务 | 91 |
| 6.1 | 服务架构演进 | 91 |
| 6.1.1 | DWSOA | 92 |
| 6.1.2 | OpenAPI | 93 |
| 6.1.3 | SmartDQ | 94 |
| 6.1.4 | 统一的数据服务层 | 96 |
| 6.2 | 技术架构 | 97 |
| 6.2.1 | SmartDQ | 97 |
| 6.2.2 | iPush | 100 |
| 6.2.3 | Lego | 101 |
| 6.2.4 | uTiming | 102 |
| 6.3 | 最佳实践 | 103 |
| 6.3.1 | 性能 | 103 |
| 6.3.2 | 稳定性 | 111 |
| 第 7 章 | 数据挖掘 | 116 |
| 7.1 | 数据挖掘概述 | 116 |
| 7.2 | 数据挖掘算法平台 | 117 |
| 7.3 | 数据挖掘中台体系 | 119 |
| 7.3.1 | 挖掘数据中台 | 120 |
| 7.3.2 | 挖掘算法中台 | 122 |
| 7.4 | 数据挖掘案例 | 123 |
| 7.4.1 | 用户画像 | 123 |
| 7.4.2 | 互联网反作弊 | 125 |
| 第 2 篇 数据模型篇 | | |
| 第 8 章 | 大数据领域建模综述 | 130 |
| 8.1 | 为什么需要数据建模 | 130 |
| 8.2 | 关系数据库系统和数据仓库 | 131 |

| | | |
|---------------|------------------------------|------------|
| 8.3 | 从 OLTP 和 OLAP 系统的区别看模型方法论的选择 | 132 |
| 8.4 | 典型的数据仓库建模方法论 | 132 |
| 8.4.1 | ER 模型 | 132 |
| 8.4.2 | 维度模型 | 133 |
| 8.4.3 | Data Vault 模型 | 134 |
| 8.4.4 | Anchor 模型 | 135 |
| 8.5 | 阿里巴巴数据模型实践综述 | 136 |
| 第 9 章 | 阿里巴巴数据整合及管理体系 | 138 |
| 9.1 | 概述 | 138 |
| 9.1.1 | 定位及价值 | 139 |
| 9.1.2 | 体系架构 | 139 |
| 9.2 | 规范定义 | 140 |
| 9.2.1 | 名词术语 | 141 |
| 9.2.2 | 指标体系 | 141 |
| 9.3 | 模型设计 | 148 |
| 9.3.1 | 指导理论 | 148 |
| 9.3.2 | 模型层次 | 148 |
| 9.3.3 | 基本原则 | 150 |
| 9.4 | 模型实施 | 152 |
| 9.4.1 | 业界常用的模型实施过程 | 152 |
| 9.4.2 | OneData 实施过程 | 154 |
| 第 10 章 | 维度设计 | 159 |
| 10.1 | 维度设计基础 | 159 |
| 10.1.1 | 维度的基本概念 | 159 |
| 10.1.2 | 维度的基本设计方法 | 160 |
| 10.1.3 | 维度的层次结构 | 162 |
| 10.1.4 | 规范化和反规范化 | 163 |
| 10.1.5 | 一致性维度和交叉探查 | 165 |
| 10.2 | 维度设计高级主题 | 166 |
| 10.2.1 | 维度整合 | 166 |
| 10.2.2 | 水平拆分 | 169 |

| | | |
|---------------|--------------|------------|
| 10.2.3 | 垂直拆分 | 170 |
| 10.2.4 | 历史归档 | 171 |
| 10.3 | 维度变化 | 172 |
| 10.3.1 | 缓慢变化维 | 172 |
| 10.3.2 | 快照维表 | 174 |
| 10.3.3 | 极限存储 | 175 |
| 10.3.4 | 微型维度 | 178 |
| 10.4 | 特殊维度 | 180 |
| 10.4.1 | 递归层次 | 180 |
| 10.4.2 | 行为维度 | 184 |
| 10.4.3 | 多值维度 | 185 |
| 10.4.4 | 多值属性 | 187 |
| 10.4.5 | 杂项维度 | 188 |
| 第 11 章 | 事实表设计 | 190 |
| 11.1 | 事实表基础 | 190 |
| 11.1.1 | 事实表特性 | 190 |
| 11.1.2 | 事实表设计原则 | 191 |
| 11.1.3 | 事实表设计方法 | 193 |
| 11.2 | 事务事实表 | 196 |
| 11.2.1 | 设计过程 | 196 |
| 11.2.2 | 单事务事实表 | 200 |
| 11.2.3 | 多事务事实表 | 202 |
| 11.2.4 | 两种事实表对比 | 206 |
| 11.2.5 | 父子事实的处理方式 | 208 |
| 11.2.6 | 事实的设计准则 | 209 |
| 11.3 | 周期快照事实表 | 210 |
| 11.3.1 | 特性 | 210 |
| 11.3.2 | 实例 | 212 |
| 11.3.3 | 注意事项 | 217 |
| 11.4 | 累积快照事实表 | 218 |
| 11.4.1 | 设计过程 | 218 |
| 11.4.2 | 特点 | 221 |

| | | |
|--------|----------|-----|
| 11.4.3 | 特殊处理 | 223 |
| 11.4.4 | 物理实现 | 225 |
| 11.5 | 三种事实表的比较 | 227 |
| 11.6 | 无事实的事实表 | 228 |
| 11.7 | 聚集型事实表 | 228 |
| 11.7.1 | 聚集的基本原则 | 229 |
| 11.7.2 | 聚集的基本步骤 | 229 |
| 11.7.3 | 阿里公共汇总层 | 230 |
| 11.7.4 | 聚集补充说明 | 234 |

第3篇 数据管理篇

| | | |
|--------|--------------|-----|
| 第12章 | 元数据 | 236 |
| 12.1 | 元数据概述 | 236 |
| 12.1.1 | 元数据定义 | 236 |
| 12.1.2 | 元数据价值 | 237 |
| 12.1.3 | 统一元数据体系建设 | 238 |
| 12.2 | 元数据应用 | 239 |
| 12.2.1 | Data Profile | 239 |
| 12.2.2 | 元数据门户 | 241 |
| 12.2.3 | 应用链路分析 | 241 |
| 12.2.4 | 数据建模 | 242 |
| 12.2.5 | 驱动 ETL 开发 | 243 |
| 第13章 | 计算管理 | 245 |
| 13.1 | 系统优化 | 245 |
| 13.1.1 | HBO | 246 |
| 13.1.2 | CBO | 249 |
| 13.2 | 任务优化 | 256 |
| 13.2.1 | Map 倾斜 | 257 |
| 13.2.2 | Join 倾斜 | 261 |
| 13.2.3 | Reduce 倾斜 | 269 |

| | |
|--------------------|-----|
| 第 14 章 存储和成本管理 | 275 |
| 14.1 数据压缩 | 275 |
| 14.2 数据重分布 | 276 |
| 14.3 存储治理项优化 | 277 |
| 14.4 生命周期管理 | 278 |
| 14.4.1 生命周期管理策略 | 278 |
| 14.4.2 通用的生命周期管理矩阵 | 280 |
| 14.5 数据成本计量 | 283 |
| 14.6 数据使用计费 | 284 |
| 第 15 章 数据质量 | 285 |
| 15.1 数据质量保障原则 | 285 |
| 15.2 数据质量方法概述 | 287 |
| 15.2.1 消费场景知晓 | 289 |
| 15.2.2 数据加工过程卡点校验 | 292 |
| 15.2.3 风险点监控 | 295 |
| 15.2.4 质量衡量 | 299 |
| 第 4 篇 数据应用篇 | |
| 第 16 章 数据应用 | 304 |
| 16.1 生意参谋 | 305 |
| 16.1.1 背景概述 | 305 |
| 16.1.2 功能架构与技术能力 | 307 |
| 16.1.3 商家应用实践 | 310 |
| 16.2 对内数据产品平台 | 311 |
| 16.2.1 定位 | 311 |
| 16.2.2 产品建设历程 | 312 |
| 16.2.3 整体架构介绍 | 316 |
| 附录 A 本书插图索引 | 319 |

第1章

总述

2014年，马云提出，“人类正从IT时代走向DT时代”。如果说在IT时代是以自我控制、自我管理为主，那么到了DT（Data Technology）时代，则是以服务大众、激发生产力为主。以互联网（或者物联网）、云计算、大数据和人工智能为代表的新技术革命正在渗透至各行各业，悄悄地改变着我们的生活。

在DT时代，人们比以往任何时候更能收集到更丰富的数据。IDC的报告显示：预计到2020年，全球数据总量将超过40ZB（相当于40万亿GB），这一数据量是2011年的22倍！正在呈“爆炸式”增长的数据，其潜在的巨大价值有待发掘。数据作为一种新的能源，正在发生聚变，变革着我们的生产和生活，催生了当下大数据行业发展热火朝天的盛景。

但是如果不能对这些数据进行有序、有结构地分类组织和存储，如果不能有效利用并发掘它，继而产生价值，那么它同时也成为一场“灾难”。无序、无结构的数据犹如堆积如山的垃圾，给企业带来的是令人咋舌的高额成本。

在阿里巴巴集团内，我们面临的现实情况是：集团数据存储达到

EB 级别，部分单张表每天的数据记录数高达几千亿条；在 2016 年“双 11 购物狂欢节”的 24 小时中，支付金额达到了 1207 亿元人民币，支付峰值高达 12 万笔/秒，下单峰值达 17.5 万笔/秒，媒体直播大屏处理的总数据量高达百亿且所有数据都需要做到实时、准确地对外披露……这些给数据采集、存储和计算都带来了极大的挑战。

在阿里内部，数据工程师每天要面对百万级规模的离线数据处理工作。阿里大数据井喷式的爆发，加大了数据模型、数据研发、数据质量和运维保障工作的难度。

同时，日益丰富的业态，也带来了各种各样、纷繁复杂的数据需求。如何有效地满足来自员工、商家、合作伙伴等多样化的需求，提高他们对数据使用的满意度，是数据服务和数据产品需要面对的挑战。

如何建设高效的数据模型和体系，使数据易用，避免重复建设和数据不一致性，保证数据的规范性；如何提供高效易用的数据开发工具；如何做好数据质量保障；如何有效管理和控制日益增长的存储和计算消耗；如何保证数据服务的稳定，保证其性能；如何设计有效的数据产品高效赋能于外部客户和内部员工……这些都给大数据系统的建设提出了更多复杂的要求。

本书介绍的阿里巴巴大数据系统架构，就是为了满足不断变化的业务需求，同时实现系统的高度扩展性、灵活性以及数据展现的高性能而设计的。

如图 1.1 所示是阿里巴巴大数据系统体系架构图，从图中可以清晰地看到数据体系主要分为数据采集、数据计算、数据服务和数据应用四大层次。

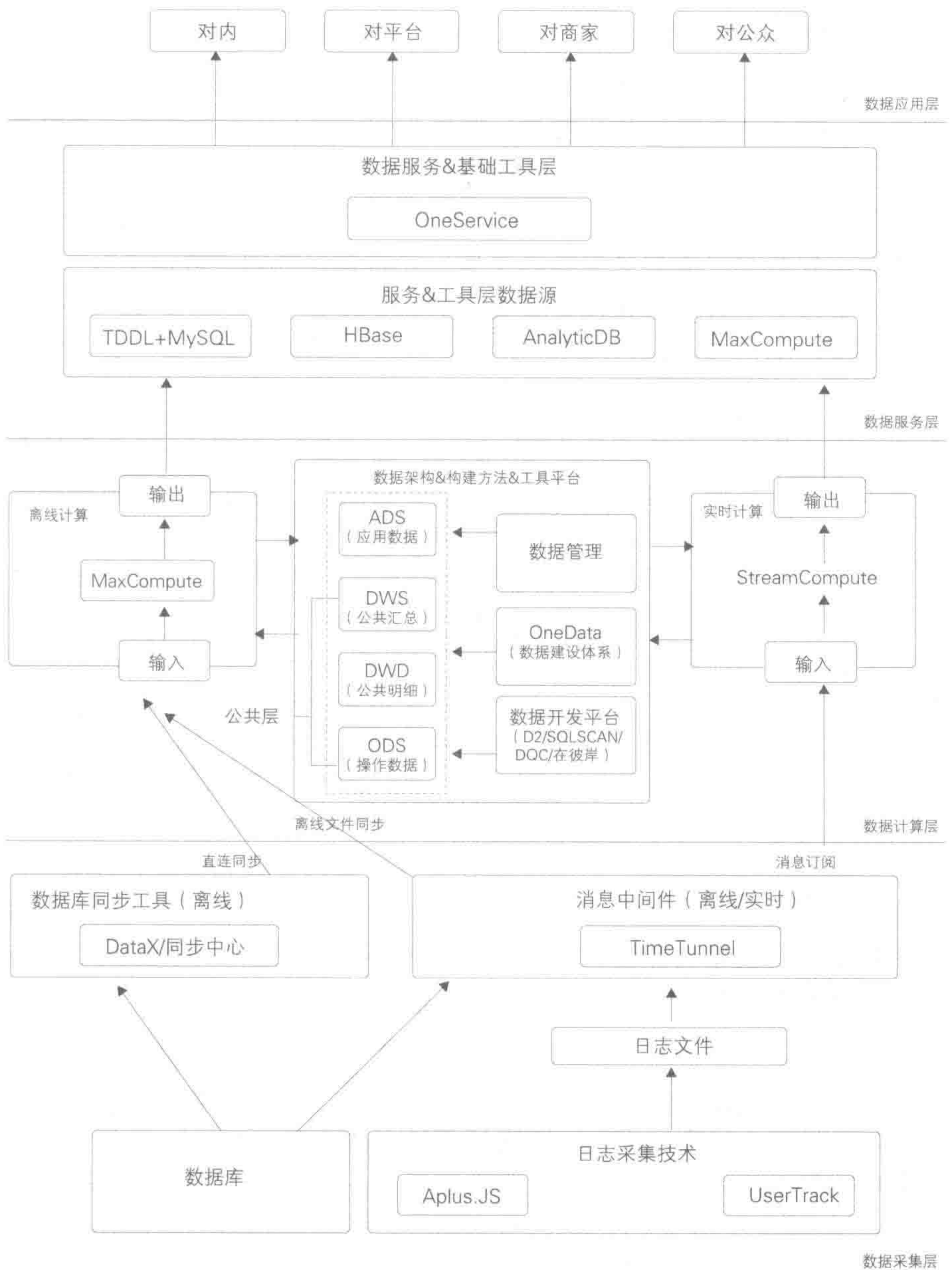


图 1.1 阿里巴巴大数据系统体系架构图