

云湖共生 · 数智未来

数据湖应用实践白皮书

阿里云基础产品事业部

阿里云计算平台事业部

阿里云数据库产品事业部

阿里云通用行业与电销事业部

联合出品

数据湖应用实践白皮书

1

相关结论

- 3 相关结论

2

数据湖的定义与架构

- 4 数据湖的定义
- 5 数据湖与数据仓库的对比
- 6 云湖共生，企业级数据湖

3

数据湖核心组件与方案介绍

- 7 核心组件
- 11 阿里云数据湖方案介绍

4

数据湖构建方案

- 14 数据湖的构建流程
- 15 数据湖典型构建方案

5

应用实践

- 19 海量数据交互式查询解决方案
- 20 构建分层模式混合数据湖
- 22 基于数据湖打造机器学习能力

6

客户声音

- 24 客户声音

7

研究方法

- 25 研究方法
- 26 附录 A：调查问卷

云湖共生 数智未来

数据湖应用实践白皮书

1 数据湖应用实践白皮书

相关结论



来自数据治理的挑战越发严峻

根据焦点小组讨论，85% 的用户认为构建或者计划构建数据湖的初衷是为了应对数据的快速增长以及治理的挑战。当下，随着数据量的快速扩张驱动企业组织采用更有效的数据管理方式。



数据湖的构建需要一套完善的流程

并非所有的企业组织都需要构建数据湖，数据湖的建设过程应该与业务紧密结合；但是数据湖的建设过程与传统的数据仓库，甚至是大热的数据中台应该是有所区别的。



基于云原生的数据湖将成为主流

随着企业越来越多的基于云计算来构建自己的业务系统，因此基于云原生的数据湖成为云中部署的理想业务负载手段。



数据湖的价值在于数据洞察

数据湖的最终价值在于能够高效的利用不同来源、数量巨大的数据，这使得组织构建数据湖或升级现有数据仓库，利用数据湖收集、存储各种数据并加以分析，帮助提取切实可行的洞察，在客户互动、创新选择、运营效率等方面带来益处。

2

数据湖应用实践白皮书

数据湖的定义与架构

1 数据湖的定义



【传统数据库】

面向大型制造业、银行、金融内部系统

大数据技术【探索期】

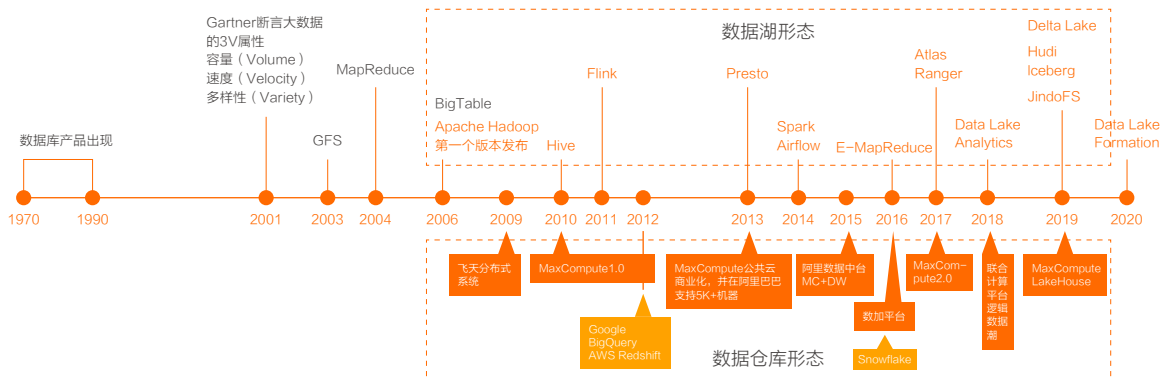
互联网时代到来，以分布式调度、存储为核心的基础设施建设时期，打破数据库以低成本大规模扩张问题。分布式计算模型为MapReduce

大数据技术【发展期】

开始关注开发效率，分布式计算模型针对场景细分，总体向SQL靠拢

大数据技术【普惠期】

开始关注投入企业生产必须的能力：工作流、安全、治理、规模、稳定性等，出现数据中台概念



从探索到普惠，大数据技术发展的 20 年

从概念上来说，“数据湖”已经被越来越多的人所接受。数据湖是以集中方式存储各种类型数据，提供弹性的容量和吞吐能力，能够覆盖广泛的数据源，支持多种计算与处理分析引擎直接对数据进行访问的统一存储平台。它能够实现数据分析、机器学习，数据访问和管理等细粒度的授权、审计等功能。

数据湖对存取的数据没有格式类型的限制，数据产生后，可以按照数据的原始内容和属性，直接存储到数据湖，无需在数据上传之前对数据进行任何的结构化处理。数据湖可以存储结构化数据（如关系型数据库中的表），半结构化数据（如 CSV、JSON、XML、日志等），非结构化数据（如电子邮件、文档、PDF 等）以及二进制数据（如图形、音频、视频等）。

2 数据湖与数据仓库的对比

数据湖 DATA LAKE	VS	数据仓库 DATA WAREHOUSE
存储原始数据，入湖简单、建设成本低	成本	数据提前建模，入仓要求高、建设成本高
数据体系松散、灵活度高	灵活性	数据体系严格、灵活度较低
数据治理较困难	治理难度	数据治理容易
数据种类丰富（结构化、半结构化、非结构化）	数据存储种类	数据种类较单一（结构化、半结构化）

数据湖涵盖的范围较广，在一些相关功能上与“数据仓库”概念类似，一些企业的管理、决策者也总是混淆两者的区别。其实，仅仅从产品应用场景上，数据湖、数据仓库就表现出了明显的不同：

首先，数据捕获时未定义架构。数据湖在功能上可以实现各种类型数据的存储，数据湖中的数据可以是非结构化的、未处理的形态，数据可以在确定需要使用时才会对数据进行对应处理与转换；而数据仓库则通常存放的是经过处理的、结构化的数据，数据存储的 Schema 在数据存储之前就需要被定义好。

其次，存储到数据湖中的数据通常会按照原始形态直接存储，随着业务和使用场景的发展，会使用不同的计算引擎对已经存储的数据进行分析与处理，数据湖中的数据在一个企业组织中通常会被多个不同应用、系统和部门使用和分析，覆盖的场景广泛并且范围也会动态延展，因此需要提供更多的灵活性以适应快速变化的应用场景；数据仓库中的数据通常使用场景是在数据收集期间就已经明确，数据仓库通常集中在 BI、业务、运营等商业决策相关场景中，数据仓库也可以把已经存在的数据转换到新场景，但在灵活性方面不如数据湖，需要有更多的数据转换时间和开发资源投入。

3 云湖共生，企业级数据湖



云原生时代的到来，也同样引领数据湖进入“云湖共生”新的阶段。在这样的背景下，阿里云推出基于云原生技术的企业级数据湖解决方案，该方案采用了存储计算分离架构，存储层基于阿里云对象存储 OSS 构建，并与阿里云数据湖分析（Data Lake Analytics 简称 DLA）、数据湖构建（Data Lake Formation 简称 DLF）、E-MapReduce（简称 EMR）、DataWorks（简称 DW）等计算引擎无缝对接，且兼容丰富的开源计算引擎生态，可满足大数据系统统一存储、海量规模的需求，与此同时，云原生数据湖可直接接入业务生产中心，业务系统中的原始数据、日志数据等可通过互联网直接入湖，无需经过中间环节处理，业务效率提升一倍以上，有力地驱动企业IT系统从成本中心转型为创新中心。

3

数据湖应用实践白皮书



数据湖核心组件与 方案介绍

1 核心组件



1.1 数据存储

随着数据来源的丰富和分布化，数据呈现出多样化的特点。与此同时，企业对数据分析功能的依赖也在激增。传统的集中数据存储和预定义方式已经无法满足业务场景快速变化的需求。数据湖可以有效集中存储各类未经处理、加工的数据，特别是从各种物联网智能设备捕获的数据。数据的统一存储是企业级数据湖构建的重要前提。数据湖为具有不同背景和工具偏好的不同客户提供统一存储服务。统一存储有助于更好地进行治理，维护和使用多种工具的能力，同时，具有以下特征：

- 可扩展性：企业数据湖充当整个组织或部门数据的集中数据存储，它必须可以解决跨企业间、不同部门间、不同业务间的数据互信和数据主权问题，同时能够按照容量的需求任意扩展。
- 数据可用性：数据的可用性和可靠性是企业制定决策的关键。跨多可用区的数据冗余机制有助于实现数据的高可用性，而多区域的数据复制可确保有效的灾难恢复，保证业务连续性。

- 安全性：在云或内部部署中，数据安全性意味着业务的安全性，因此，数据必须经过加密、不可篡改并且必须满足相关合规要求。安全性应该从一开始就进行设计，并且需要将其纳入非常基本的体系结构和设计中，进而在企业整体基础架构内部署和管理数据湖的安全性。比如存储层需要提供丰富的加密机制，具有灵活的密钥管理服务，并且具备防篡改能力。

- 存储任何内容：云存储对于文件格式、文件数量、存储容量没有限制，极大突破了 HDFS（Hadoop 分布式文件系统）因为 NameNode 设计机制无法支撑大量小文件的困境。应用无需对小文件进行合并处理，就可以直接存入云存储系统，系统的响应能力完全不会因为文件数量的增加有任何降低。

主要产品：

- 对象存储 OSS：阿里云对象存储 OSS 是一种海量、安全、低成本、高可靠的云存储服务，提供可达 99.999999999%（12 个 9）的数据持久性，99.995% 的全球领先的服务可用性。借助 RESTful API，对象存储 OSS 可以在互联网任何位置存储和访问，容量和处理能力可弹性扩展，支持从热到冷多种存储类型可全面优化存储成本。

1.2 数据加工

数据湖可以对接多种差异性的计算引擎，运行在不同负载之上，多种计算引擎都共享同一套存储系统。

对存储数据的类型提供了充足的灵活性，没有传统“入仓”的各种限制，数据产生后就能从对接的数据通道上传到数据湖，根据实际分析需求，再进行数据抽取（extract）、转换（transform）、加载（load），生成所需要的格式数据，生成的处理后数据可以再存储到数据湖中，供其他阶段或者分析中使用。这样的好处在于：一方面，数据的灵活性更高，可以让各类应用、智能物联网设备都可以轻松解决原数据的存储，而不需要设备消耗大量计算

资源进行转换，降低终端设备资源的能耗；另一方面，数据湖中数据可以与多种计算与分析平台结合使用，对于企业来说，计算存储分离的架构更灵活，在应对业务的快速变化时更加容易地构建应用平台和系统，提升效率，对于数据的分析，也可以更快速、更轻量，减少整体的成本投入。

主要产品：

- 数据湖构建：数据湖构建（Data Lake Formation，DLF）作为云原生数据湖架构核心组成部分，帮助用户简单快速地构建云原生数据湖解决方案。DLF 提供湖上元数据统一管理、企业级权限控制，并无缝对接多种计算引擎，打破数据孤岛，洞察业务价值。

- E-MapReduce：阿里云 E-MapReduce（EMR）是构建在阿里云云服务器 ECS 上的开源 Hadoop、Spark、HBase、Hive、Flink 生态大数据 PaaS 产品。提供用户在云上使用开源技术建设数据仓库、离线批处理、在线流式处理、即时查询、机器学习等场景下的大数据解决方案。

1.3 数据分析与治理

数据分析与治理是指对企业中数据的可用性、完整性和安全性的全面管理，主要取决于业务策略和技术实践。数据湖对接的各种计算引擎和计算生态，可以便捷的对集中存储的数据进行批量计算、机器学习、交互式查询。数据湖采用的 API 接口让数据湖实现多引擎的统一元数据管理和权限管理。移动应用、智能设备、Hadoop 计算生态、云原生服务都可以轻松实现对接。通过数据湖自建的事件机制能够更轻松的完成元数据采集，结合元数据管理数据服务，提升数据整体管理能力，让数据湖不会成为“数据沼泽”。

主要产品：

- MaxCompute：MaxCompute 是一项面向分析的大数据计算服务，它以 Serverless 架构提供快速、全托管的在线数据仓库服务，消除传统数据平台在资源扩展性和弹性方面的限制，最小化用户运维投入，使您经济并高效的分析处理海量数据。
- 云原生数据湖分析：云原生数据湖分析(Data Lake Analytics，简称 DLA) 采用全新的云原生+Serverless+数据库与大数据一体化架构，支持企业级权限管理、高效入湖、ETL、机器学习、流计算与交互式分析等。核心组件包括：统一Meta、Lakehouse、Serverless Spark、Serverless Presto，其中，Lakehouse融合数据仓库、数据湖的技术优势并有大量创新，解决DB、Kafka、SLS数据入湖难的问题；Serverless Spark与Presto引擎支持按需消费，总体分析成本可降低50%，并可以分析与集成对象存储(OSS)、HDFS、MySQL、Lindorm、TableStore、MongoDB、ES等云上15种数据源的数据。
- 机器学习 PAI：机器学习 PAI（Platform of Artificial Intelligence）是阿里云面向企业客户及开发者，提供轻量化、高性价比的云原生机器学习平台支持百亿特征、千亿样本规模加速训练，百余种落地场景，全面提升机器学习工程效率。
- 日志服务 SLS：阿里云日志服务（SLS）是一种日志大数据解决方案，一站式提供数据收集、清洗、分析、可视化和告警功能。全面提升海量日志处理能力，实时挖掘数据价值，智能助力研发/运维/运营/安全等场景。
- DataWorks：基于 MaxCompute/EMR/MC-Hologres 等大数据计算引擎，为客户提供专业高效、安全可靠的一站式大数据开发与治理平台，自带阿里巴巴数据中台与数据治理最佳实践，赋能各行业数字化转型。每天阿里巴巴集团内部有数万名数据/算法工程师正在使用 DataWorks，承担集团 99% 数据业务构建。



2.1 强大的数据存储引擎

数据湖一个重要目标是将所有企业数据能够集中存储，以供企业的各类应用在授权下进行访问。结合数据湖的这一设计目标以及元数据管理、自动化数据采集、自动化数据解析和处理等技术来解决各类与应用相关的日志埋点、采集与分析。

阿里云的数据湖底座对象存储 OSS 是基于阿里云自研的分布式存储引擎——盘古搭建，提供体系化的数据采集能力，支持结构化/半结构化/非结构化数据源。数据湖统一存储，提供了数据的管理能力。冷热分层的存储方式解决了数据分散在各个集群，需要在不同存储系统中反复拷贝等运维困扰。在大数据访问方面支持基于 Ranger 的数据湖权限管理，支持混合云方案，总体成本可降低近 50%。

此外，阿里云数据湖存储成本相对于高效云盘下降十倍以上，查询性能相对于传统对象存储提速三倍以上，并且查询引擎有着极高的弹性。通过云原生数据湖存储与计算分离特性，计算资源的弹性伸缩可以全部由实际计算任务的需求定义，无需按照业务峰值常驻计算资源。

2.2 与云原生平台的深入结合

阿里云云原生数据湖体系由数据湖对象存储 OSS、云原生数据湖分析 DLA、数据湖构建 DLF、E-MapReduce、DataWorks 等产品强强组合，在存储与计算分离架构下，提供“湖存储”、“湖加速”、“湖管理”、“湖计算”的企业级数据湖解决方案。

2.3 内部及外部的有效验证

经过二十年的发展，阿里巴巴集团已经成为一个数字经济体，阿里巴巴集团首先是阿里云数据湖产品的最佳实践者，后者支撑了阿里巴巴集团内部的电商、移动办公、文娱、物流、本地生活等各种复杂业务，建立了完善的自我实践机制，产品和方案得到有效的验证。同时，阿里云的数据湖方案也支撑了在线教育、互联网广告、新媒体、网络游戏等行业用户在快速发展过程中的实际业务需求，实现了技术的有效赋能。

目前，阿里云数据湖方案可支持 EB 级别的数据存储量，存储超过十万 Database、一亿 Table 以及十亿级别的 Partition，每天完成超过三十亿次的元数据服务请求，支持多个开源计算引擎以及多种云原生数仓引擎。



4 数据湖应用实践白皮书 ⑥

数据湖构建方案

1 数据湖的构建流程



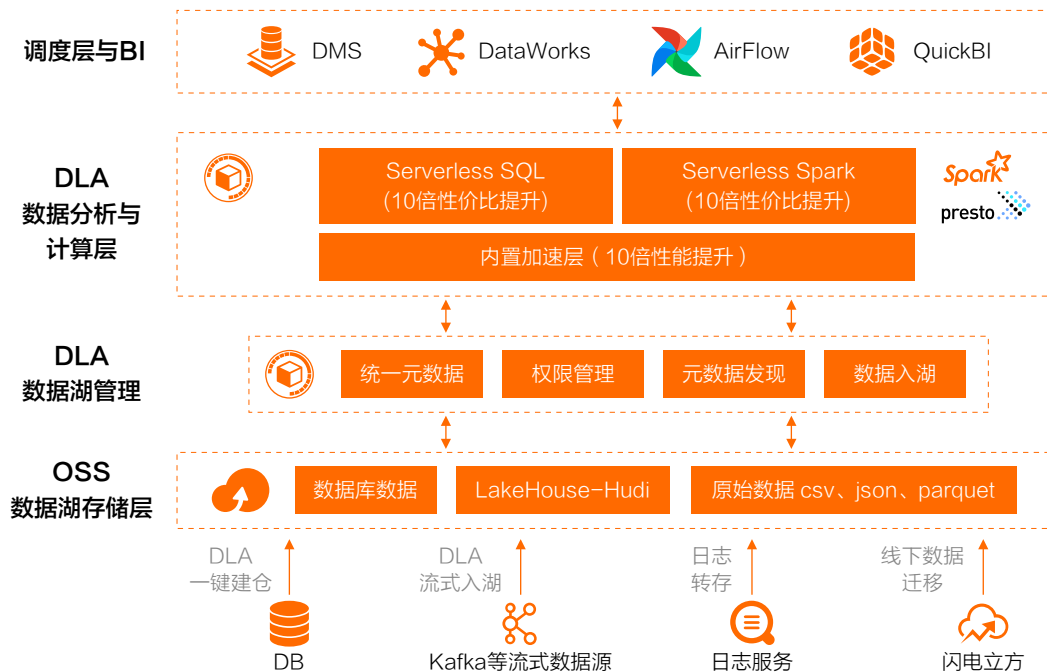
在数字化转型的当下，企业组织需要处理的数据越来越多、形式越来越杂，原有的数据存算架构的成本越来越高，但效率越来越低。在这种背景下，企业亟需一种成本更低且效率较高的方式来存算数据、访问数据，这是数据湖兴起的重要原因。同时，在数据湖的建设过程，技术应该与业务紧密结合，并且以一种更敏捷的方式去构建，根据相关业务及技术实践，我们建议采用如下步骤来构建数据湖：



2 数据湖典型构建方案



(1) 基于DLA构建的数据湖方案



应用场景

在游戏、社交、电商、资讯等互联网应用场景中有大量计算分析、个性化推荐、离线交互式查询等需求来提升产品的创新能力，在此类场景中，数据规模快速增长，数据来源渠道多样化，包括但不限于应用自身产生的数据、采集的各类日志数据、数据库中抽取的各类数据，计算根据业务的需求随时进行、动态扩展，基于阿里云数据湖方案，可以使用DLA云原生的服务与引擎，解决数据管理、分析、计算问题，能够进一步发挥存算分离的优势，让数据的管理与分析更加简单。

优势与价值

阿里云对象存储 OSS 是阿里云对外提供的海量、安全和高可靠的云存储服务，能支撑 EB 规模的数据湖，客户无需考虑存储量扩容。

云原生数据湖分析 DLA 无需购买任何资源，即可运行标准的 SQL 分析数据，DLA 可使用云原生的 Spark 服务，动态创建计算资源参与计算。DLA 可以提供一站式数据库入湖、元数据管理、元数据自动发现、支持数据库数据同步到对象存储 OSS 中，并且提供统一开放的 Meta 服务对对象存储 OSS 数据进行管理，支持库表权限，此外，利用元数据爬取功能，可以一键创建对象存储 OSS 上的元数据信息。

(2) 基于 EMR 构建的数据湖方案



- 应用场景

Hadoop 发行的各个版本中提供了众多开源计算引擎，随着计算引擎与数据存储解耦的趋势不断增强，Hadoop 存储与计算耦合成为可能，同时云上 AI 技术的发展进一步降低数据分析的成本，提升数据分析效率。这些都为开源数据湖的发展提供良好契机。

- 优势与价值

阿里云对象存储 OSS 是阿里云对外提供的海量、安全和高可靠的云存储服务，能支撑 EB 规模的数据湖，客户无需考虑存储量扩容，同时对象存储 OSS 提供的冷热分层能力，可以根据数据热度进行分层进而降低数据存储成本。

对象存储 OSS 无缝对接 EMR Hive、Spark、Presto、Impala 等大数据处理引擎，提供弹性吞吐能力和空间，通过 OSS 加速器功能或和 EMR JindoFS 的结合使用可以输出高达上百 Gbps 读取吞吐能力。同时，对象存储 OSS 支持多种数据通道，全面覆盖日志、消息、数据库、HDFS 各种数据源，支持通过 Ranger 提供与开源体系一致的权限管理。

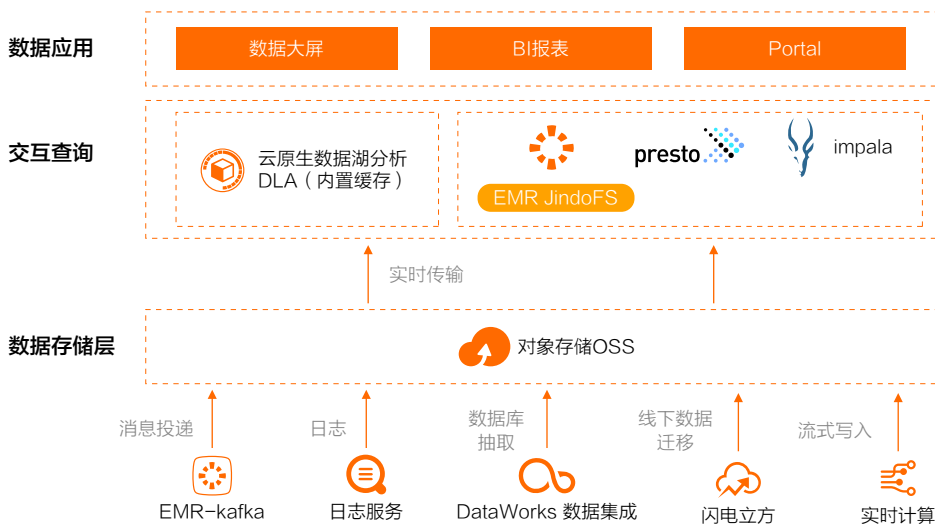
OSS 与 EMR 的结合还解决了传统 Hadoop 在扩展性、运维模式、成本优化方面的难题。按需建立集群，避免集群长时间处于低负载状态，资源利用率得到大幅提升。

数据湖构建 DLF 服务是为了解决构建数据湖过程中用户对数据资产的管理需求。DLF 可以为 OSS 存储的数据提供统一的元数据视图和统一的权限管理，并提供实时数据入湖和清洗模板，为上层的数据分析引擎提供生产级别的元数据服务。

结合 DataWorks 数据综合治理为数据湖的用户提供统一的数据视图，掌握数据资产的现状，提高获取数据的效率，保障数据安全的合规，提升数据查询的分析效率，有效支撑离线大数据仓库的构建，数据联邦的查询和分析处理以及海量数据的低频交互式查询和智能报表的构建。

5 数据湖应用实践白皮书 应用实践

1 海量数据交互式查询解决方案



1.1 应用场景

在一些业务中，数据源可能来自多个渠道，且实时产生的数据中有大量的数据是半结构化数据，需要对存储的数据做 OLAP（联机分析处理）及查询。对于数据的分析可能是来自于临时或者在某些时间段内的需求，例如周报、月报等历史数据分析，需要能够对接流式计算产生的数据，并且分析结果需要在对应的应用中可视化展示。

1.2 价值与优势

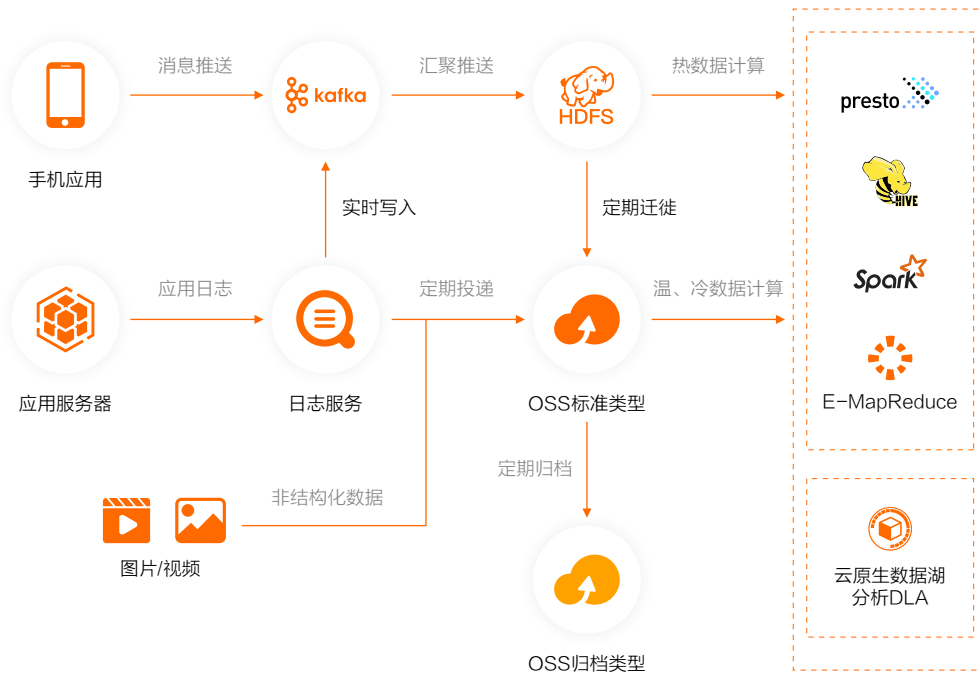
首先，该解决方案可以使用包括阿里云数据湖分析产品 DLA 或 EMR 服务提供的分布式查询引擎对已有数据进行查询，能够更加灵活、更快的获得查询结果；

其次，对象存储 OSS 可以支持多种交互式查询服务和引擎，对应的数据分析产品可以直接查询对象存储 OSS 中的数据；

此外，该方案提供包括 OSS 加速器、EMR JindoFS 和 DLA SQL 内置加速在内的多种加速方案，进一步提升交互式查询的访问速度。

2 构建分层模式混合数据湖





2.1 应用场景

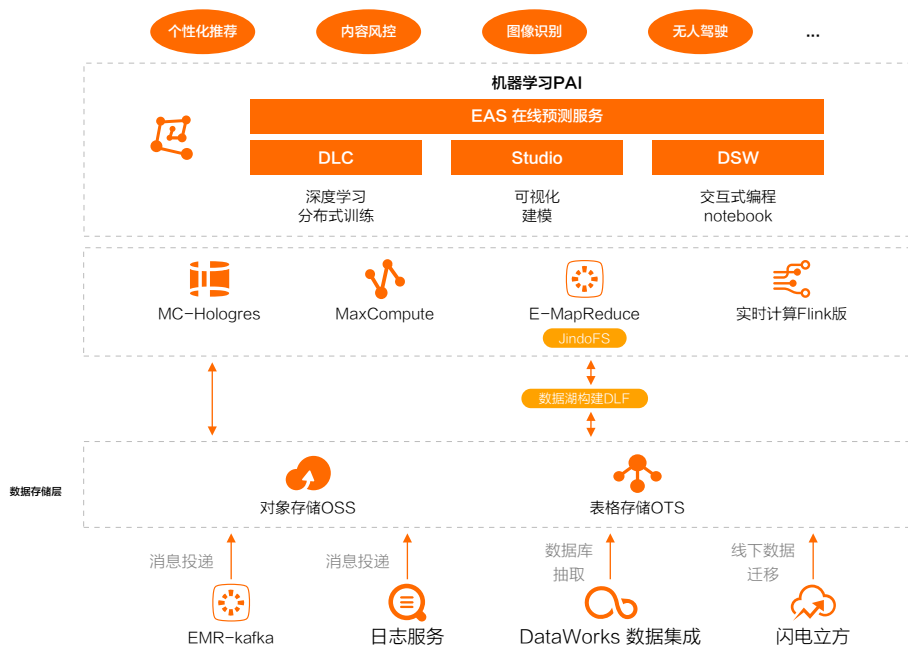
在一些数据基数大、数据随着时间推移访问热度降低的场景中，存在存储成本持续优化的需求，同时计算和存储增速往往无法匹配，扩容存储资源可能会导致空闲的计算资源产生。

2.2 价值与优势

数据湖可以实现冷热数据分层存储，显著优化成本，避免出现存储扩容所产生的计算资源闲置现象，极大降低资源规划管理难度。

存储到 OSS 的冷数据，可以直接使用 EMR、DLA 等多种云原生计算引擎进行数据处理，对于原有计算任务几乎无需调整，就可以直接使用。

3 基于数据湖打造机器学习能力



3.1 应用场景

在互联网电商、游戏、社交等行业中存在推荐、风控、预测等在内的机器学习场景，基于 CPU 资源的在线预测服务和基于 GPU 资源的在线推理场景以及包括图像、人脸、语音、NLP、无人驾驶行业的模型训练等深度学习场景。

在这些场景中，需要面对数据量大、模型训练性能差、时间长，训练出的模型效果差等问题，并且在线资源的弹性扩容预测困难，基于 GPU 资源的推理服务由于 QPS 性能不高，也会导致资源消耗过大。

3.2 价值与优势

阿里云企业级数据湖充分发挥机器学习平台 PAI 的技术优势。首先，PAI-DLC 分布式训练让训练速度提升近 400%，让训练时长缩短近 75%，相同资源规模下，更省时间，按量付费的方式也更节省资源成本。同时，个性化推荐和风控等解决方案，为企业提供 AI 专家服务和保障业务。

其次，在推理场景中 PAI Blade 和 PAI EAS 模型在线部署组合使用，实现自动化弹性伸缩，按量付费无须担心资源浪费，同时支持多模型的管理和 A/B 测试，可以更好的选择效果优质的模型。在 GPU 资源不变的条件下，推理并发量 QPS 提升 1 倍，在客户业务量不变的情况下，GPU 资源最高节省 50%。

6 数据湖应用实践白皮书 ⑥

客户声音

“ 在选择阿里云之前，我们曾经使用自建的数据集群解决方案，但却面临着自建数据集群运维成本高，弹性伸缩难的问题，同时，由于数据类型复杂以及分布地区众多，数据的综合治理成本很高，阿里云的服务帮助我们很好地解决了以上的问题。

——易点天下 CTO 王一舟 ”

“ 阿里云提供的数据湖解决方案满足了加和 ReachMax 的流量业务日均 150 亿条、近 200 列的数据分析，以及大量的临时数据统计需求，支撑了每年高达 50 亿人民币流量决策和分析服务，利用 OSS 的多进多出、分层计费、全类型备份的能力，帮助加和大幅简化了数据存储和应用的复杂度，相比之前的自建解决方案，需求的平均成本降低了 30%，并大幅度减少了数据处理时间，提升了需求承接的弹性，让我们能够有更多的时间和精力专注在业务的创新和流量价值的提升上。

——加和科技 CTO 王可攀 ”

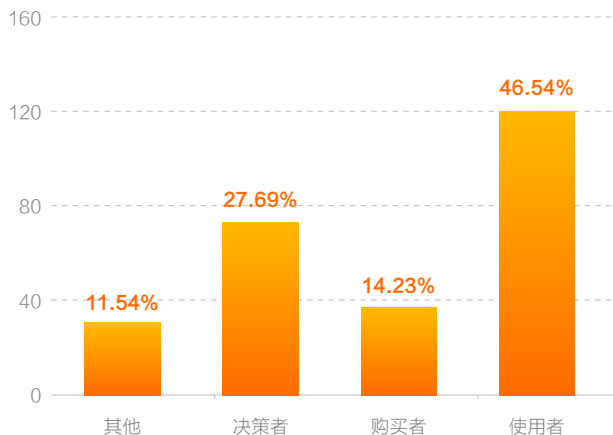
“ 英语流利说使用了阿里云 OSS 对象存储构建数据湖，支撑千万级规模用户的使用，应对流利说的“中国人英语语音数据库”容量高速增长和复杂度不断提升的分析性能的持续需求，借助于计算与存储的弹性伸缩能力，让运维管理难度大幅度下降的同时，也给整体成本带来了有效的优化。

——英语流利说 CTO 胡哲人 ”

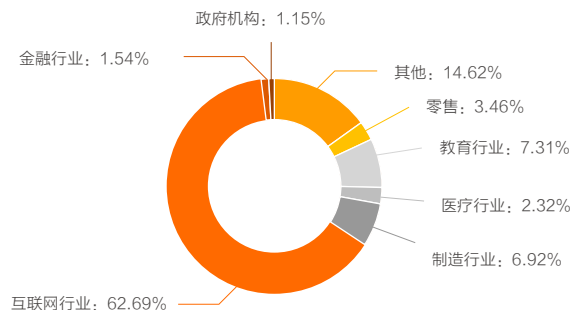
7 数据湖应用实践白皮书

研究方法

为了更深入地了解数据湖在客户业务中的实际应用，我们对多家公司的相关负责人进行了面对面的访谈，并采访了多位阿里云解决方案架构师，收集与阿里云数据湖方案相关的反馈，同时，采用调查问卷的方式，收集了三百多位阿里云用户的实际反馈。



受访者角色分布



受访企业行业分布

附录 A：调查问卷

1、促使贵公司/组织采用数据湖解决方案的最主要因素是什么？

- 现有方案投入成本过高
- 现有方案资源调配及管理耗时太长
- 希望在数据治理上获得更多价值
- 基于云上新业务的出现
- 其他

2、在数据湖解决方案的选择上会更关注什么？

- 部署、使用的便捷性
- 数据维护/迁移/管理成本
- 弹性扩容，业务快速部署
- 一套完善的构建体系
- 方案的云原生能力
- 其他

3、在使用阿里云数据湖方案之后，贵公司/组织获得了哪些预期之外的好处？

- 业务更流畅，客户满意度提升
- 内部员工满意度（如IT系统使用人员）提升
- 低于预期的成本投入
- 快速的服务响应
- 丰富的产品生态，让业务开展更为顺利
- 其他

4、未来一年贵公司/组织在数据湖应用的预算将会？

- 增加100%以上
- 增加100%以下
- 减少50%以下
- 基本持平
- 减少50%以上

5、您认为阿里云数据湖解决方案在哪些方面需要提升与创新？

•

6、贵公司/组织规模：

- 少于50人
- 50到150人
- 150到500人
- 500人到1000人
- 1000人以上

7、贵公司/组织所处的行业：

- 金融行业
- 互联网行业
- 制造行业
- 其他
- 医疗行业
- 教育行业
- 政府机构

8、您的职责是：

- 使用及维护者
- 选择及购买者
- CIO或总负责人
- 其他



 阿里云

欢迎打开钉钉扫码加入
查看更多数字化转型解决方案

本白皮书及其内容的版权，属于阿里云计算有限公司所有或已获得合法授权；未经阿里云计算有限公司书面授权许可，任何人不得复制、修改、转载、摘编或以其它任何方式使用本白皮书的全部或部分内容。