

Ark Fabric 用户手册

1. Ark Fabric 概述

1.1 简介

Ark Fabric 数据编织系统可在不做数据迁移的情况下，屏蔽数据软件系统差异，连通数据存储孤岛，实现任何位置的多元化数据的即时连接与融合计算，使数据的采集、清洗、存储、处理、分析与可视化变得更加简便和高效，从而全面释放数据价值，提升业务决策质量。

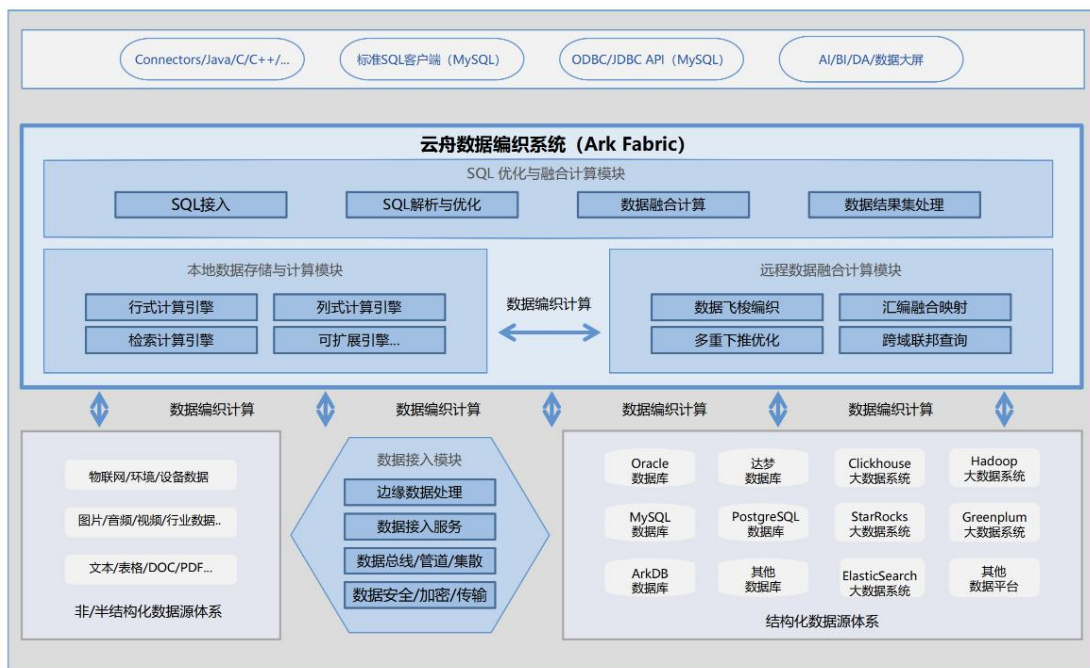
1.2 优势

Ark Fabric 作为一种全新的数据融合和管理理念，目标是为企业提供一个统一的交互访问方式，该访问方式独立于部署平台、数据流程、地理位置和架构方法。Ark Fabric 的重点在于对数据不需要移动，就可以确保企业内的所有数据都可在正确的时间、地点以实时的速度被正确地使用。它可以实现数据汇聚，数据收集、数据管理和共享加速，优化数据的使用，助力数据价值实现。Ark Fabric 旨在提高数据访问的便捷性和效率，帮助企业更好地利用数据进行智能化决策。

Ark Fabric 核心特性在于：其内置的引擎技术可动态改进数据使用，从而使得数据管理工作量减少 70%，加速企业价值实现时间。根据数据预测，Ark Fabric 的部署使用将可减少 50% 的人力数据管理成本，同时提高数据使用效率。

通过 Ark Fabric，用户可以获得更高效、准确、安全的数据处理和整合，并从中获得更好的业务竞争优势和经济利益。

1.3 架构



- 汇编级数据融合技术，在计算层实现对各种引擎的统一调度和执行优化。
- 独创的数据飞梭技术，实现跨引擎自适应分批数据抽取和过滤。
- 引擎可扩展架构设计，向上透明实现引擎扩展，可持续支持未知数据接入和计算。
- 本地结构化数据引擎与桥接现有数据软件相结合，实现数据灵活融合存储，统一计算。
- 通过可扩展协议技术及自由定制查询方式，覆盖更灵活的查询服务场景。

1.3.1 内计算引擎

1.3.1.1 ArkDB

ArkDB 作为 Ark Fabric 的内置计算引擎，是一款面向场景化需求的企业级云原生数据库。其特性主要体现在以下几个方面：简单易用、透明读写分离、计算存储分离、分布式存储、弹性伸缩、物理复制、数据分层治理、无锁备份和快速恢复、集群高可用、完整性事务支持。

1.3.1.2 ArkOLAP

ArkOLAP 作为 Ark Fabric 的内置计算引擎，通过 ArkDB 与 Clickhouse 配合实现 HTAP 的功能，对 ArkDB 的数据操作会被自动同步到 Clickhouse。

1.3.2 外计算引擎

Ark Fabric 具有强大的外计算引擎库，可以便捷的实现跨源异构数据的发现与访

问，实现对既有数据资源的能力整合。

已支持以下持数据引擎/平台：

- 1. MySQL
- 2. Oracle
- 3. 达梦数据库
- 4. Clickhouse

以下数据引擎/平台的支持处于 alpha 阶段：

- 1. ElasticSearch
- 2. StarRocks
- 3. PostgreSQL
- 4. GreenPlum

1.4 版本

Ark Fabric V1.1.0-beta.1

2. Ark Fabric 环境要求

2.1 硬件环境

2.1.1 最低配置（单节点）

硬件	参数
CPU	4 核/线程及以上，x86_64 或 Arm64
内存	8G 及以上
存储	128G 及以上的 SSD 本地存储
网络	1000Mb/s 及以上

2.1.2 推荐配置（单节点）

硬件	参数
CPU	32 核/线程及以上，x86_64 或 Arm64
内存	64G 及以上
存储	1T 及以上的 SSD 本地存储
网络	10Gb/s 及以上

2.1. 软件环境

操作系统要求：

操作系统参数	值
Linux 内核版本	3.10.0 或以上
发行版	CentOS, 银河麒麟 V10
CPU 架构	x86_64, Arm64

分布式存储系统要求：

Ceph	值
版本	15.2.17

3. Ark Fabric 功能使用

Ark Fabric 需通过 ArkDOS 系统，以界面交互方式实现相应的数据配置与管理行为。

3.1 ArkDOS 操作说明

3.1.1 数据配置

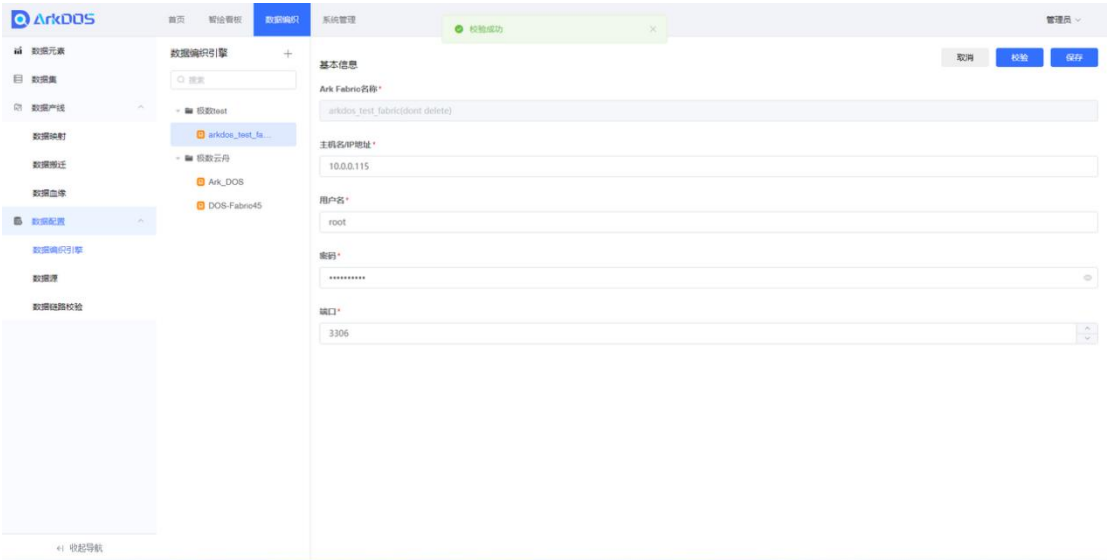
3.1.1.1 数据编织引擎

功能概述

管理 ArkDOS 至 Ark Fabric 的连接信息，系统支持多 Ark Fabric 管理。

功能说明

- a. 功能区域（左侧）：新建文件夹、添加数据编织引擎（即：Ark Fabric）、已添加的文件夹及 Ark Fabric 名称列表及其他操作功能（如：文件夹重命名、删除文件夹、删除、移动）。
- b. 内容区域（右侧）：展示已添加数据编织引擎的详细信息，点击校验可以查看当前连接信息状态。



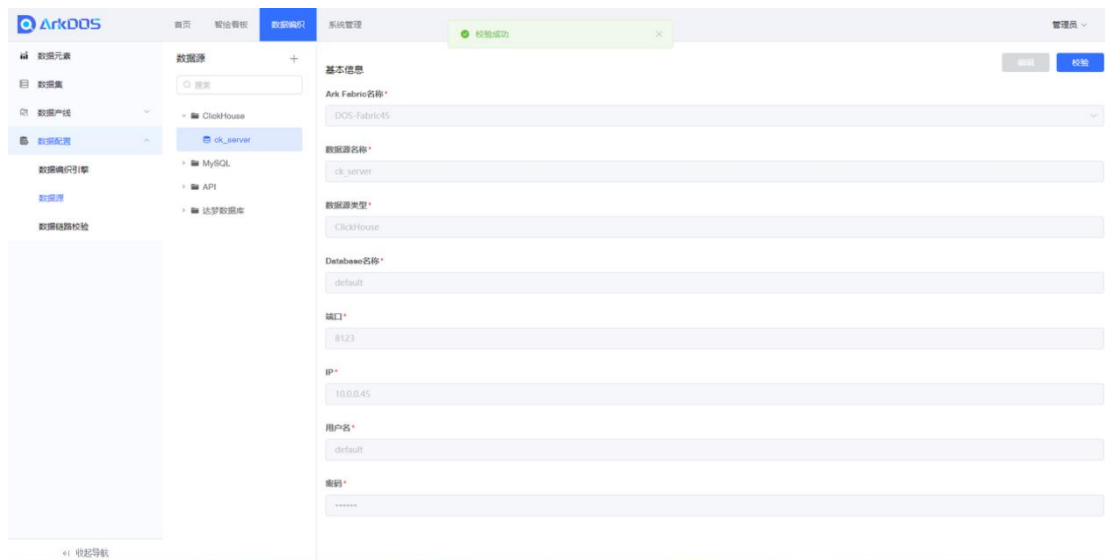
3.1.1.2 数据源

功能概述

管理 Ark Fabric 至各数据源类型连接信息，为后续数据分析操作提供数据来源。

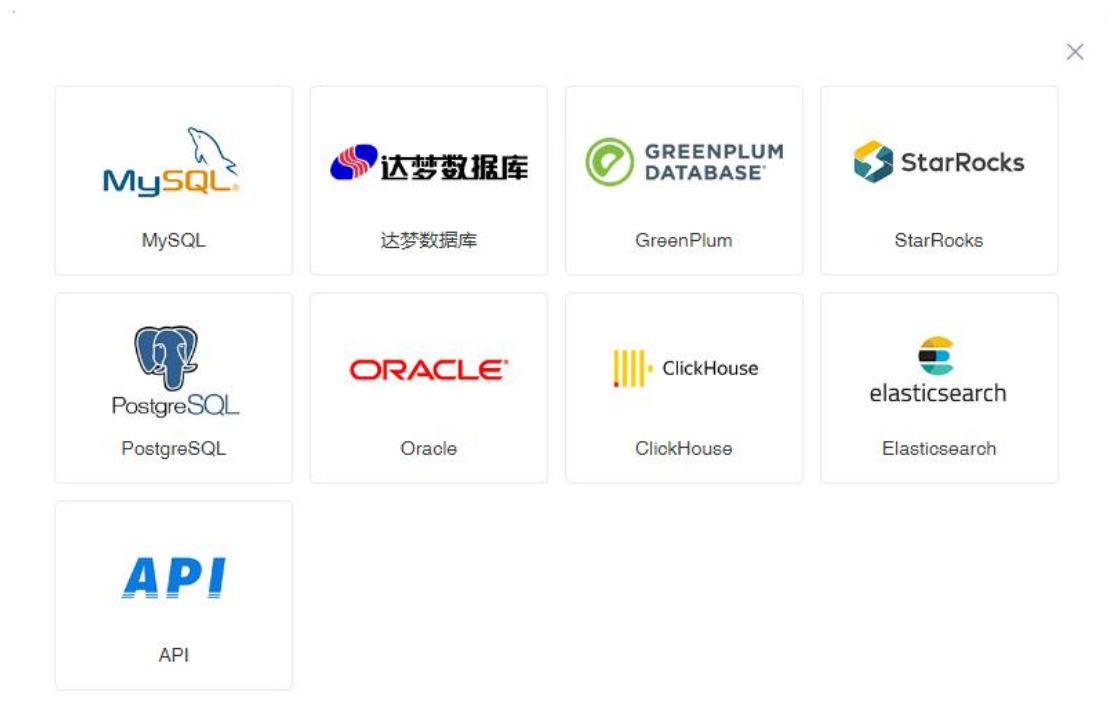
功能说明

- 功能区域（左侧）：新建数据源按钮、已新建完成的数据源配置会自动生成相应数据源类型的文件夹及数据源名称列表及其他操作功能（如：删除数据源）。
- 内容区域（右侧）：展示已添加数据源的详细信息，点击校验可以查看当前连接信息状态。



支持的数据源类型

- OLTP 型数据库：MySQL、达梦数据库（DM）、Oracle、PostgreSQL
- OLAP 型数据库：Elasticsearch、ClickHouse、StarRocks、GreenPlum
- API 数据源



配置 MySQL 数据源（示例）

1 前提条件

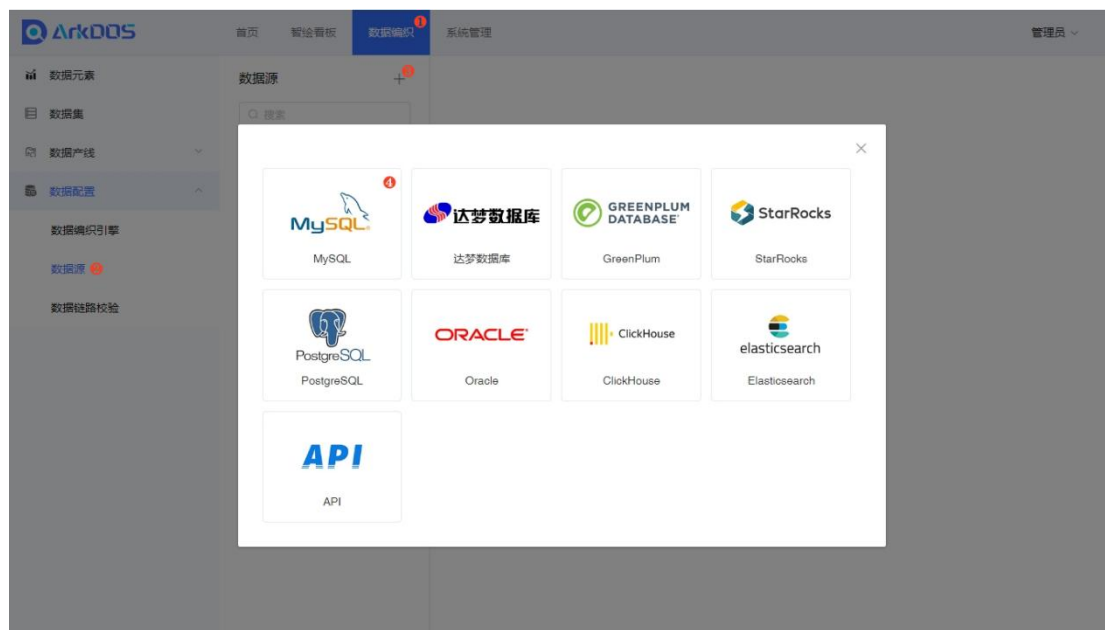
配置 MySQL 数据库之前，请收集以下信息：

- 数据库服务器的 IP 地址和端口号
- Database 名称
- 数据库用户名和密码

2 配置数据源连接步骤

步骤一：登入 ArkDOS 系统。

步骤二：按照以下步骤，选择 MySQL 图标。



步骤三：填入收集的 IP 、端口、数据库等相关的信息。

< 新建MySQL数据源 取消 校验 保存

基本信息

Ark Fabric名称 *

请选择Ark Fabric

数据源名称 *

请输入名称

数据源类型 *

MySQL

Database名称 *

请填写Database名称

主机名/IP地址 *

请输入主机名/IP地址

用户名 *

请输入用户名

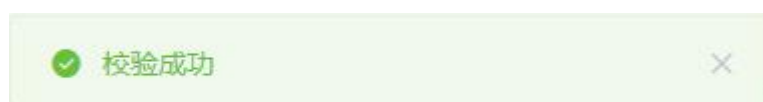
密码 *

请输入密码

端口 *

请输入端口

步骤四：数据源检验，校验成功后如下图所示，点击保存即可。



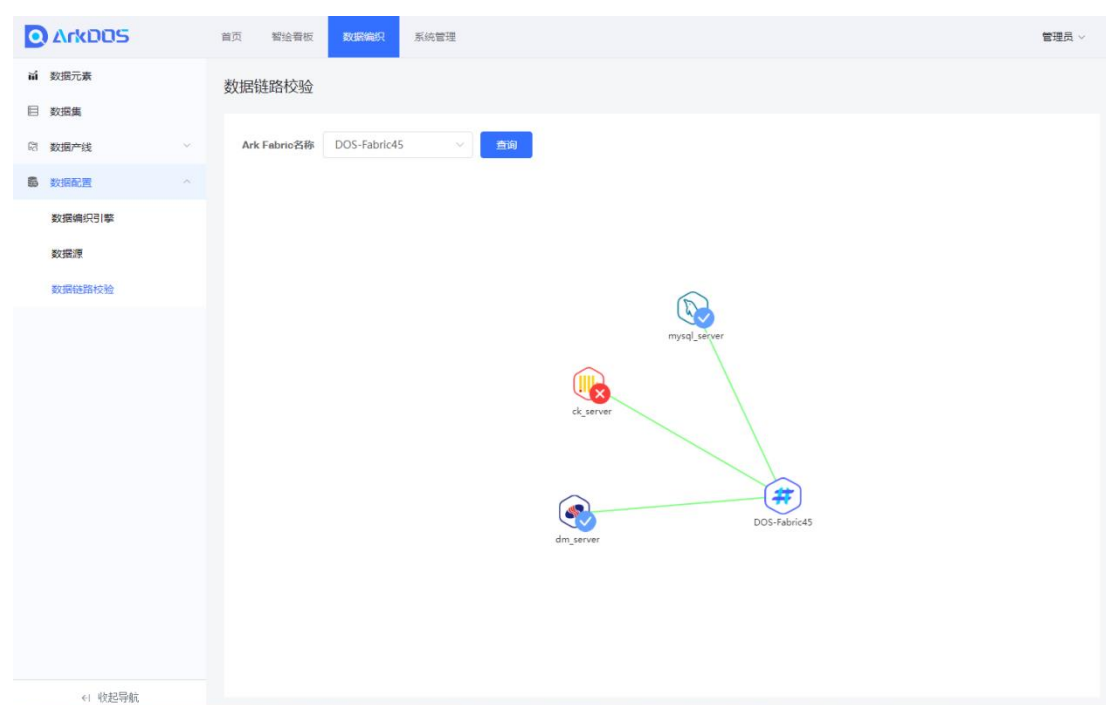
3.1.1.3 链路校验

功能概述

对 Ark Fabric 至数据源两端库表链路状态进行检验，协助用户掌握链路动态及排查问题。

功能说明

- a. 功能区域（上部）：提供选择 Ark Fabric 的筛选功能。
- b. 内容区域（下部）：展示链路拓扑图，展示范围：Ark Fabric 至数据源链路。
 - i. 展示从已选中的 Ark Fabric 至建立关系的数据源之间链路状态图；
 - ii. Ark Fabric 至数据源链路出现问题时，连接线路以红色虚线展示，数据源侧使用灰色图标，点击链路弹出错误信息；
 - iii. 数据源出现问题时，数据源使用红色图标，点击报错的数据源，从右侧滑动展示弹窗中显示错误信息。



3.1.2 数据产线

3.1.2.1 数据映射

功能概述

实现源数据库表至 Ark Fabric 库表的虚拟化，通过建立 Ark Fabric 表与源数据表的映射关系，为通过 Ark Fabric 实现跨数据库平台查询提供基础准备。

功能说明

- a. 功能区域（上部）：新建映射规则，提供源数据库名及日期范围筛选。

- b. 内容区域（下部）：展示已添加数据映射规则的执行结果信息。

The screenshot shows the ArkDQS Data Mapping page. The left sidebar contains navigation links: 数据元素, 数据集, 数据血缘, 数据映射 (selected), 数据迁移, 数据血缘, 数据配置, 数据索引引擎, 数据源, and 数据链路校验. The main content area is titled '数据映射' and features a '+ 新建映射规则' button and a '筛选' button. Below these is a table with 10 rows of mapping rules. The table columns are: Ark Fabric库名, 源数据库名, 映射情况, 规则描述, 提交时间, and 操作. The bottom of the page shows pagination: '共 10 条', '1' (selected), '10条/页', and '前往 1 页'.

Ark Fabric库名	源数据库名	映射情况	规则描述	提交时间	操作
mysql_test	zs0126	已映射完成1张数据库表	archives	2023-09-15 17:08:28	详情
fabric_manage	default	已映射完成1张数据库表	ck123	2023-09-04 15:54:06	详情
Ark_DOS_mysql	zs0623	已映射完成171张数据库表	数据源库表映射	2023-09-11 14:19:07	详情
fabric_manage	default	已映射成功2张数据库表...	CK	2023-08-30 10:30:03	详情
mysql_test	zs0126	已映射完成172张数据库表	房屋修缮	2023-09-05 14:19:39	详情
test	default	已映射完成1张数据库表	renwu1	2023-09-04 15:55:40	详情
dm	SYSDBA	已映射完成1张数据库表	dos_test1	2023-09-01 11:07:27	详情
dm	SYSDBA	已映射完成1张数据库表	dm_test	2023-09-01 10:55:45	详情
clickhouse	SYSDBA	已映射完成1张数据库表	test123	2023-08-23 15:43:51	详情
fabric_manage	SYSDBA	已映射完成17张数据库表	2022年中国GDP	2023-08-31 15:08:17	详情

新建映射规则

1 基本信息

- Ark Fabric 名称：选择数据映射规则需要存储在某个 Ark Fabric 中；
- 数据源：选择获取映射的数据来源；
- Ark Fabric 库名：选择存储映射关系的目的地。

2 映射规则

- 根据需要进行选择源数据库表，支持多选，支持自定义映射后的表名（表名不允许重复）；
- 根据需要进行配置已选源表的字段，可全选或部分选择。

新建映射规则

取消

确认

Ark Fabric名称

DOS-Fabric45

数据源

ck_server

Ark Fabric库名

fabric_manage

选择数据源表

	源表名 Q	自定义表名	配置字段
☒	lineitem	lineitem	已配置
☐	t1	t1	配置
☐	t3	t3	配置
☐	t4	t4	配置

配置表字段

	字段名 Q	数据类型
☐	l_orderkey	Int32
☒	l_partkey	Int32
☒	l_suppkey	Int32
☒	l_linenumbr	Int32
☒	l_quantity	Decimal(15, 2)
☒	l_extendedprice	Decimal(15, 2)
☐	l_discount	Decimal(15, 2)
☒	l_tax	Decimal(15, 2)
☒	l_returnflag	String

映射规则详情

映射规则详情

批量删除

Ark Fabric表名

请输入Ark Fabric表名

源数据表名

请输入源数据源表名

查询

	Ark Fabric名称	Ark Fabric库名	Ark Fabric表名	数据源类型	源数据库名	源数据表名	执行状态	操作
☐	DOS-Fabric45	mysql_test	yard_room_check_d...	MySQL	zs0126	yard_room_check_d...	执行完成	更新 删除
☐	DOS-Fabric45	mysql_test	key_observe_wall_list	MySQL	zs0126	key_observe_wall_list	执行完成	更新 删除
☐	DOS-Fabric45	mysql_test	yard_room_problem	MySQL	zs0126	yard_room_problem	执行完成	更新 删除
☐	DOS-Fabric45	mysql_test	yard_room_notificati...	MySQL	zs0126	yard_room_notificati...	执行完成	更新 删除
☐	DOS-Fabric45	mysql_test	yard_room_check_d...	MySQL	zs0126	yard_room_check_d...	执行完成	更新 删除
☐	DOS-Fabric45	mysql_test	yard_room_check_d...	MySQL	zs0126	yard_room_check_d...	执行完成	更新 删除
☐	DOS-Fabric45	mysql_test	yard_room_check_d...	MySQL	zs0126	yard_room_check_d...	执行完成	更新 删除
☐	DOS-Fabric45	mysql_test	yard_room_check_d...	MySQL	zs0126	yard_room_check_d...	执行完成	更新 删除
☐	DOS-Fabric45	mysql_test	yard_danger	MySQL	zs0126	yard_danger	执行完成	更新 删除
☐	DOS-Fabric45	mysql_test	yard_check_detail_...	MySQL	zs0126	yard_check_detail_...	执行完成	更新 删除

共 172 条

1

2

3

4

5

6

...

18

 10条/页 前往 1 页

3.1.2.2 数据搬迁

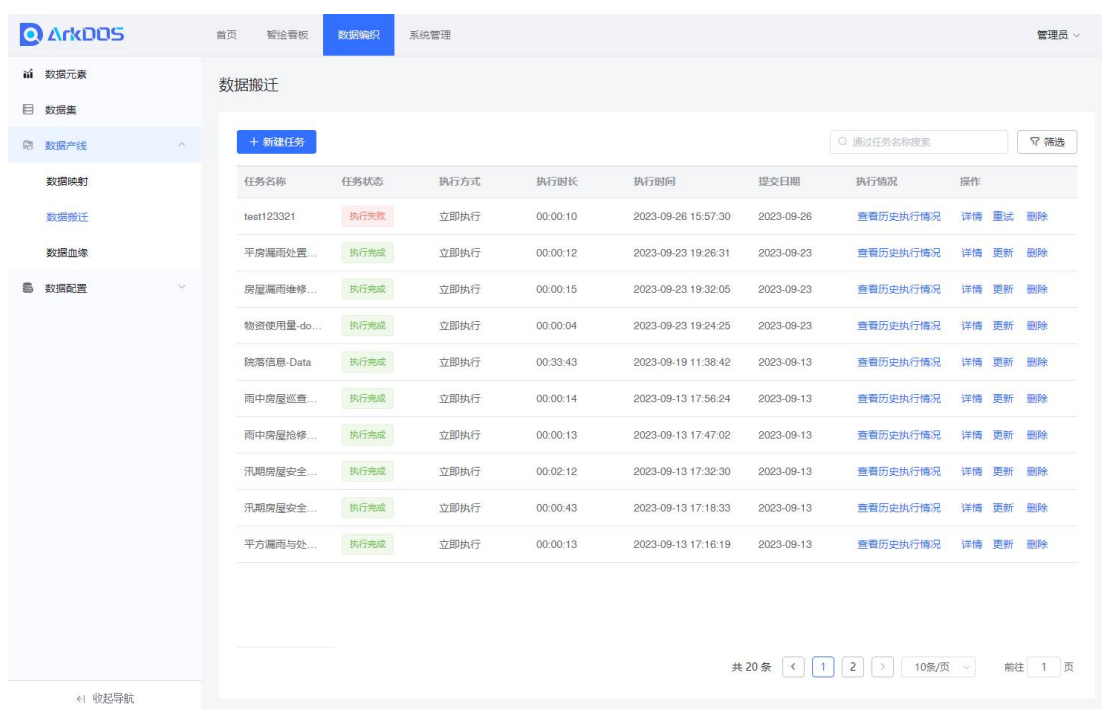
功能概述

实现源数据表至 Ark Fabric 表的数据本地化，为后续数据分析提供基于本地数据库的快速数据响应服务。

功能说明

- 功能区域（上部）：新建任务、通过任务名称搜索、执行状态筛选、日期范围筛选。

b. 内容区域（下部）：展示已添加数据产线任务的执行动态信息。



任务名称	任务状态	执行方式	执行时长	执行时间	提交日期	执行情况	操作
test123321	执行失败	立即执行	00:00:10	2023-09-26 15:57:30	2023-09-26	查看历史执行情况 详情 重试 删除	
平房漏雨处置...	执行完成	立即执行	00:00:12	2023-09-23 19:26:31	2023-09-23	查看历史执行情况 详情 更新 删除	
房屋漏雨维修...	执行完成	立即执行	00:00:15	2023-09-23 19:32:05	2023-09-23	查看历史执行情况 详情 更新 删除	
物资使用量-do...	执行完成	立即执行	00:00:04	2023-09-23 19:24:25	2023-09-23	查看历史执行情况 详情 更新 删除	
院落信息-Data	执行完成	立即执行	00:33:43	2023-09-19 11:38:42	2023-09-13	查看历史执行情况 详情 更新 删除	
雨中房屋巡查...	执行完成	立即执行	00:00:14	2023-09-13 17:56:24	2023-09-13	查看历史执行情况 详情 更新 删除	
雨中房屋抢修...	执行完成	立即执行	00:00:13	2023-09-13 17:47:02	2023-09-13	查看历史执行情况 详情 更新 删除	
汛期房屋安全...	执行完成	立即执行	00:02:12	2023-09-13 17:32:30	2023-09-13	查看历史执行情况 详情 更新 删除	
汛期房屋安全...	执行完成	立即执行	00:00:43	2023-09-13 17:18:33	2023-09-13	查看历史执行情况 详情 更新 删除	
平房漏雨与处...	执行完成	立即执行	00:00:13	2023-09-13 17:16:19	2023-09-13	查看历史执行情况 详情 更新 删除	

新建任务

1 任务信息

- 任务名称：填写数据搬迁的任务名称；
- Ark Fabric 名称：选择数据搬迁需要存储在某个 Ark Fabric 中；
- Ark Fabric 库名：选择数据存储的目地库；
- Ark Fabric 表名：选择数据存储的目的表；
- 存储引擎：

ArkDB（分布式存储）：适合存储海量数据的关系型数据库引擎，数据存储在分布式存储中，可动态扩容；

ArkDB（OLAP）：适合存储分析型的列式存储引擎，数据存储在 Clickhouse 数据库中；

ArkDB（本地存储）：适合存储少量数据的关系型数据库引擎，数据存储在本地磁盘，受限于磁盘大小及磁盘性能。

- 执行方式：立即执行或定时执行（可设定执行时间）。

2 数据准备

- 通过编写 SQL 语句查询出需要搬迁的数据；
- 可通过已选 Ark Fabric，实现跨源数据库与本地库的数据查询；
- 数据参考：帮助用户便捷查询已选 Ark Fabric 内库表及字段信息。

新建任务

取消提交

任务信息

任务名称

请输入任务名称

Ark Fabric名称

请选择Ark Fabric

Ark Fabric库名

请选择或输入新数据库名(回车确认)

Ark Fabric表名

请输入存储数据的Ark Fabric表名

存储引擎

请选择存储引擎

执行方式

立即执行

定时执行

执行时间

请选择执行时间

数据信息

1

运行结果 (显示 1000 条数据)

运行查询

点击运行查询
即可查看运行结果

数据参考

请选择Ark Fabric库

Ark Fabric表名

Ark Fabric字段名

暂无数据

暂无数据

任务详情

平房漏雨处置统计-dos45

任务信息

Ark Fabric名称

DOS-Fabric45

Ark Fabric库名

mysql_test

Ark Fabric表名

louyutongji_1

存储引擎

ArkDB(本地存储)

执行方式

立即执行

数据信息

1 SELECT

2 tp.company_id AS companyId,

3 c.short_name AS companyName,

4 COALESCE(SUM(rp.rain_leak_bungalow_number),0) AS rainLeakBungalowNum,

5 COALESCE(SUM(rp.reinforce_building_number),0) AS reinforceBuildingNum

6 FROM mysql_test.rain_patrol rp

7 LEFT JOIN mysql_test.company c ON rp.company_id = c.id

8 WHERE rp.is_deleted = 0

9 GROUP BY

10 rp.company_id

运行结果 (显示 1000 条数据)

运行查询

companyId	companyName	rainLeakBungalowNum	reinforceBuildingNum
1	京城集团	0.0	0.0
2	安宝门公司	334.0	333.0
3	交通口公司	285.0	281.0
4	东和公司	2.0	2.0
5	北新桥公司	99.0	99.0
6	修建公司	0.0	0.0
7	东四公司	252.0	243.0
8	蒙东公司	290.0	290.0
9	惠安公司	0.0	0.0
10	房产经营公司	0.0	0.0
11	福利达公司	0.0	0.0

数据参考

请选择Ark Fabric库

Ark Fabric表名

Ark Fabric字段名

暂无数据

暂无数据

3.1.2.3 数据血缘

功能概述

展示已完成数据虚拟化的映射规则，快速掌握 Ark Fabric 至数据源已建立的库表关系。

功能说明

13 / 48

- a. 功能区域（上部）：提供选择 Ark Fabric、Ark Fabric 库、Ark Fabric 表的筛选功能。
- b. 内容区域（下部）：展示数据映射规则及血缘图，展示范围：Ark Fabric 库表至源数据库表。
- i. 血缘表：展示 Ark Fabric 名称、Ark Fabric 库名、Ark Fabric 表名、数据源名称、数据源类型、源数据库名、源数据表名；
- ii. 血缘图：展示 Ark Fabric 至数据源的相关库表信息。

ArkDOS

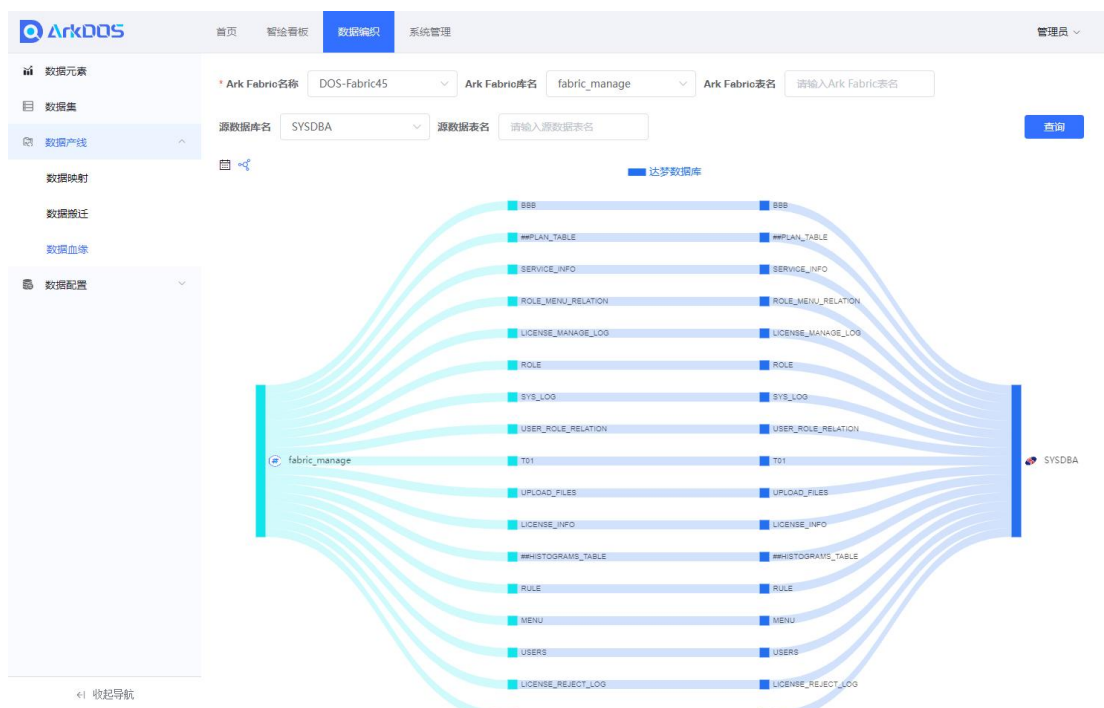
首页 数据血缘 数据映射 系统管理 管理员

Ark Fabric 名称 DOS-Fabric45 Ark Fabric 库名 请选择Ark Fabric库 源数据库名 请选择源数据库 查询

Ark Fabric名称	Ark Fabric库名	Ark Fabric表名	数据源名称	数据源类型	源数据库名	源数据表名
DOS-Fabric45	mysql_test	view_building_room_...	mysql_server	MySQL	zs0126	view_building_room_...
DOS-Fabric45	mysql_test	company	mysql_server	MySQL	zs0126	street_company_list
DOS-Fabric45	mysql_test	rain_patrol	mysql_server	MySQL	zs0126	rain_patrol_sign
DOS-Fabric45	mysql_test	company	mysql_server	MySQL	zs0126	street_company_list
DOS-Fabric45	mysql_test	house_leakage_deta...	mysql_server	MySQL	zs0126	house_leakage_deta...
DOS-Fabric45	mysql_test	company	mysql_server	MySQL	zs0126	street_company_list
DOS-Fabric45	mysql_test	prevention_material	mysql_server	MySQL	zs0126	prevention_material_...
DOS-Fabric45	mysql_test	prevention_material_...	mysql_server	MySQL	zs0126	prevention_material_...
DOS-Fabric45	mysql_test	prevention_material_...	mysql_server	MySQL	zs0126	prevention_material_...
DOS-Fabric45	mysql_test	archi	mysql_server	MySQL	zs0126	archives

共 221 条 1 2 3 4 5 6 ... 23 10条/页 前往 1 页

收起导航



3.1.2.4 数据集

功能概述

基于多类型源数据的数据查询结果，为下一步数据分析、报表制作或数据 API 服务进行相关的数据准备。

支持创建的数据集类型有：数据源数据集、SQL 数据集、Excel 数据集、API 数据集三种：

- a. 数据源数据集，源自数据元素编辑时，在用户选择数据源表后系统自动创建；
- b. SQL 数据集可通过编写 SQL 语句后将其查询结果作为数据集；
- c. Excel 数据集指通过导入本地 Excel 数据文件生成数据集；
- d. API 数据集是 API 数据源里的接口数据形成的数据集。

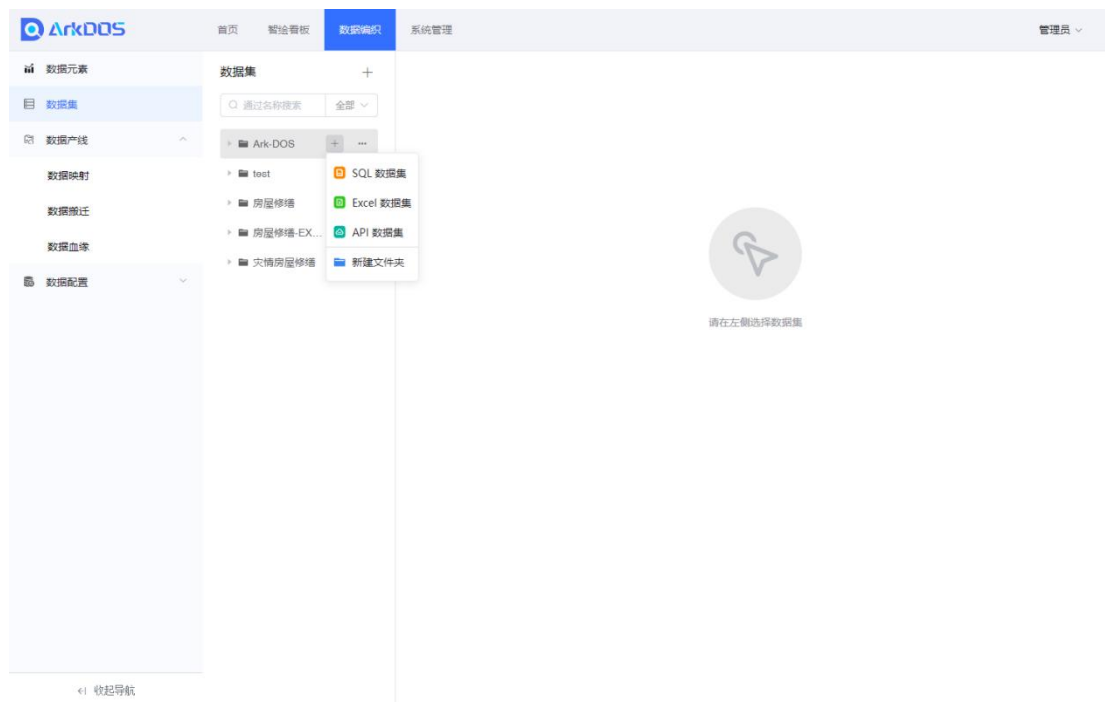
功能说明

- a. 功能区域（左侧）：新建文件夹、新建 SQL 数据集、新建 Excel 数据集、新建关联数据集、新建 API 数据集，已新建完成的数据集文件夹及数据集名称列表及其他操作功能（如：文件夹重命名、删除文件夹、移动文件夹、删除数据集、移动数据集）。
- b. 内容区域（右侧）：展示已添加数据集的数据预览、字段管理、导出为 API、数据集导出、编辑数据集等功能。

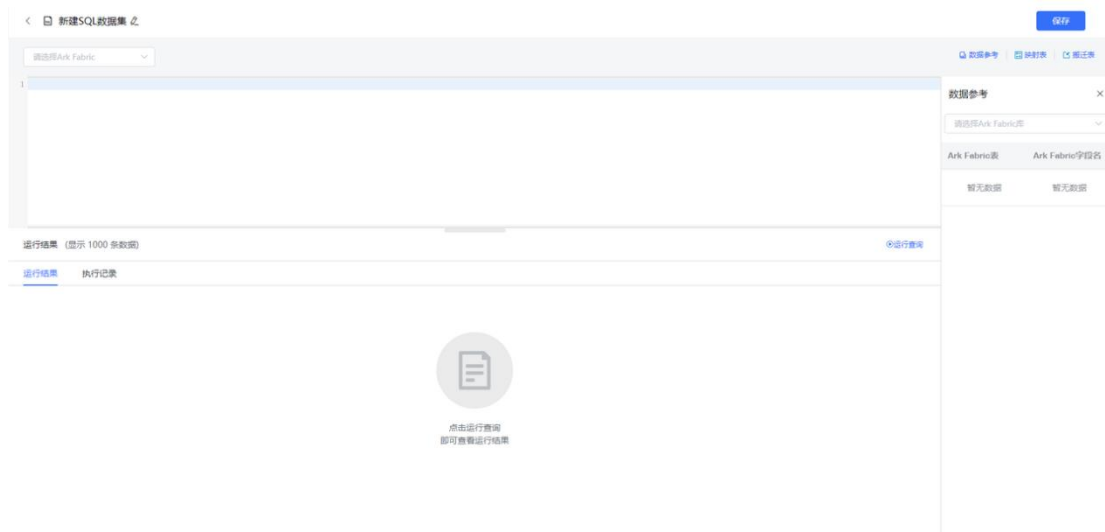
SQL 数据集

1 新建 SQL 数据集

如下图所示，点击【添加数据集】并选择【SQL 数据集】。



2 编辑数据集



3 输入正确的 SQL 语句

- 【序号 1】位置选择 Ark Fabric;
- 【序号 2】位置为 SQL 语句输入区，在此区域输入正确的 SQL 语句;
- 【序号 3】点击可展示预览数据。

ArkDOS 4.5

数据参考 数据表 备注表

```
1 SELECT par.company_id as companyId,c.short_name as companyName,
2 ROUND(SUM(CASE par.material_id WHEN 1 THEN par.num ELSE 0 END),0) as bianZhiLai,
3 ROUND(SUM(CASE par.material_id WHEN 2 THEN par.num ELSE 0 END),0) as shanGaiCaiLiao,
4 SUM(CASE par.material_id WHEN 3 THEN par.num ELSE 0 END) as qianSi,
5 ROUND(SUM(CASE par.material_id WHEN 4 THEN par.num ELSE 0 END),2) as muCai,
6 SUM(CASE par.material_id WHEN 5 THEN par.num ELSE 0 END) as shashiCaiLiao,
7 ROUND(SUM(CASE par.material_id WHEN 6 THEN par.num ELSE 0 END),2) as jiZhuao,
8 SUM(CASE par.material_id WHEN 7 THEN par.num ELSE 0 END) as yuShuChe,
9 SUM(CASE par.material_id WHEN 8 THEN par.num ELSE 0 END) as daJiaoZuoYong,
10 SUM(CASE par.material_id WHEN 9 THEN par.num ELSE 0 END) as chuiBeng,
11 ROUND(SUM(CASE par.material_id WHEN 10 THEN par.num ELSE 0 END), 0) as faBanJi
12 FROM fwez_test_prevention_material_reserve pa par LEFT JOIN fwez_test_prevention_material_reserve par ON par.reserve_id=par.id
13 LEFT JOIN fwez_test_prevention_material pa on par.material_id=pa.id LEFT JOIN fwez_test_company c on par.company_id=c.id
14 WHERE par.reserve_id IS NOT NULL
```

运行结果 (显示 1000 条数据)

运行结果 执行记录

companyId	companyName	bianZhiLai	shanGaiCaiLiao	qianSi	muCai	shashiCaiLiao	jiZhuao
2	安恒门公司	250.0	20.0	150.0	0.0	2.0	5.0
3	交里口公司	300.0	40.0	100.0	3.0	6.0	3.0
4	东和公司	200.0	10.0	300.0	2.0	2.0	0.0
5	北新桥公司	300.0	20.0	100.0	5.0	1.0	0.0
6	修建公司	500.0	15.0	400.0	10.0	10.0	10.0
7	东园公司	500.0	20.0	10.0	3.0	5.0	5.0
8	康东公司	90.0	40.0	20.0	2.0	1.5	0.0
10	康产益善公司	100.0	2.0	50.0	0.3	1.0	0.0
13	朝阳门公司	200.0	25.0	200.0	5.0	3.0	2.6
15	康西公司	450.0	90.0	150.0	4.0	4.0	5.2

SQL 数据集支持“数据参考”功能，如下图所示；

选择数据源后，点击【数据参考】，在右侧弹框中可查看当前 Ark Fabric 下的数据表及字段信息，选择具体数据表，支持一键复制表名、字段名，方便编写 SQL 语句。

ArkDOS 4.5

数据参考 数据表 备注表

```
1 SELECT par.company_id as companyId,c.short_name as companyName,
2 ROUND(SUM(CASE par.material_id WHEN 1 THEN par.num ELSE 0 END),0) as bianZhiLai,
3 ROUND(SUM(CASE par.material_id WHEN 2 THEN par.num ELSE 0 END),0) as shanGaiCaiLiao,
4 SUM(CASE par.material_id WHEN 3 THEN par.num ELSE 0 END) as qianSi,
5 ROUND(SUM(CASE par.material_id WHEN 4 THEN par.num ELSE 0 END),2) as muCai,
6 SUM(CASE par.material_id WHEN 5 THEN par.num ELSE 0 END) as shashiCaiLiao,
7 ROUND(SUM(CASE par.material_id WHEN 6 THEN par.num ELSE 0 END),2) as jiZhuao,
8 SUM(CASE par.material_id WHEN 7 THEN par.num ELSE 0 END) as yuShuChe,
9 SUM(CASE par.material_id WHEN 8 THEN par.num ELSE 0 END) as daJiaoZuoYong,
10 SUM(CASE par.material_id WHEN 9 THEN par.num ELSE 0 END) as chuiBeng,
11 ROUND(SUM(CASE par.material_id WHEN 10 THEN par.num ELSE 0 END), 0) as faBanJi
12 FROM fwez_test_prevention_material_reserve pa par LEFT JOIN fwez_test_prevention_material_reserve par ON par.reserve_id=par.id
13 LEFT JOIN fwez_test_prevention_material pa on par.material_id=pa.id LEFT JOIN fwez_test_company c on par.company_id=c.id
14 WHERE par.reserve_id IS NOT NULL
```

运行结果 (显示 1000 条数据)

运行结果 执行记录

companyId	companyName	bianZhiLai	shanGaiCaiLiao	qianSi	muCai	shashiCaiLiao	jiZhuao
2	安恒门公司	250.0	20.0	150.0	0.0	2.0	5.0
3	交里口公司	300.0	40.0	100.0	3.0	6.0	3.0
4	东和公司	200.0	10.0	300.0	2.0	2.0	0.0
5	北新桥公司	300.0	20.0	100.0	5.0	1.0	0.0
6	修建公司	500.0	15.0	400.0	10.0	10.0	10.0
7	东园公司	500.0	20.0	10.0	3.0	5.0	5.0
8	康东公司	90.0	40.0	20.0	2.0	1.5	0.0
10	康产益善公司	100.0	2.0	50.0	0.3	1.0	0.0
13	朝阳门公司	200.0	25.0	200.0	5.0	3.0	2.6
15	康西公司	450.0	90.0	150.0	4.0	4.0	5.2

4 保存 SQL 数据集

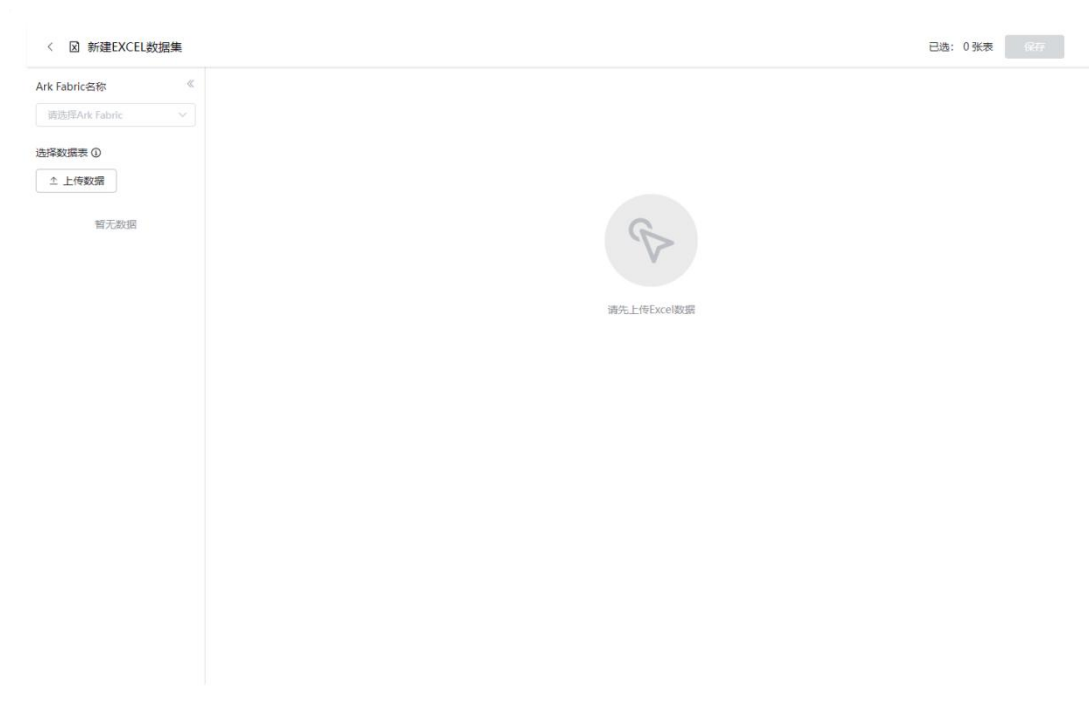
点击【保存】，SQL 数据集添加成功。

Excel 数据集

1 上传 Excel 与 CSV 文件

支持上传 Excel 文件；

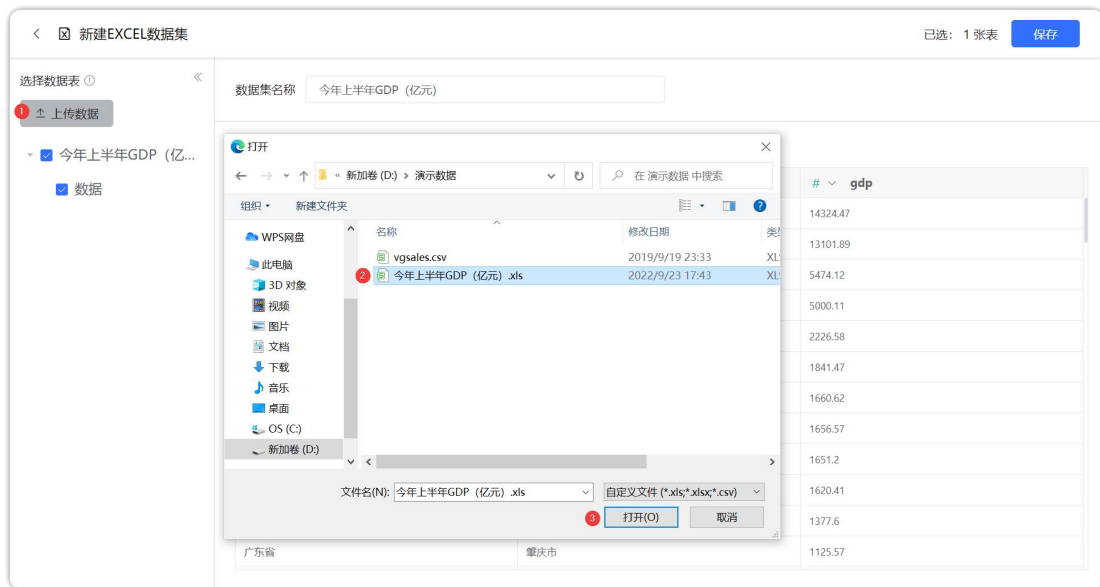
如下图所示，点击【添加数据集】并选择【Excel 数据集】，选择【Ark Fabric】，点击【上传文件】，将 Excel 文件上传到 ArkDOS。



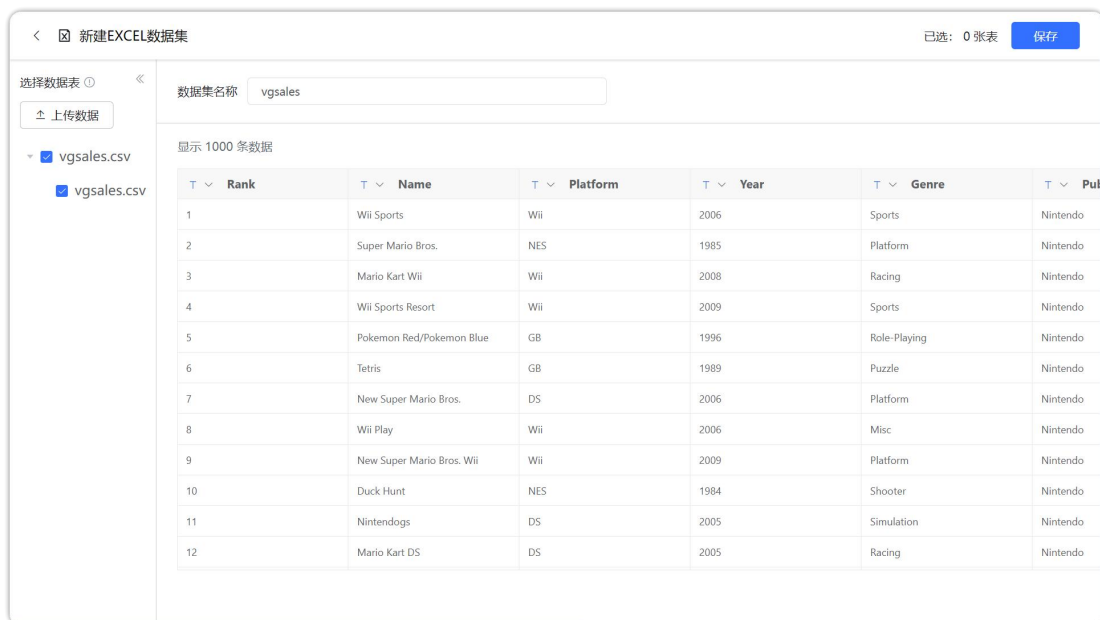
注意：Excel 必须满足以下要求的格式，否则无法成功导入：

- 1.文件中不能存在合并单元格；
- 2.文件的第一行为标题行，不能为空，不能为日期型；
- 3.Excel 文件大小请确保在 500M 以内。





支持上传 CSV 格式文件。

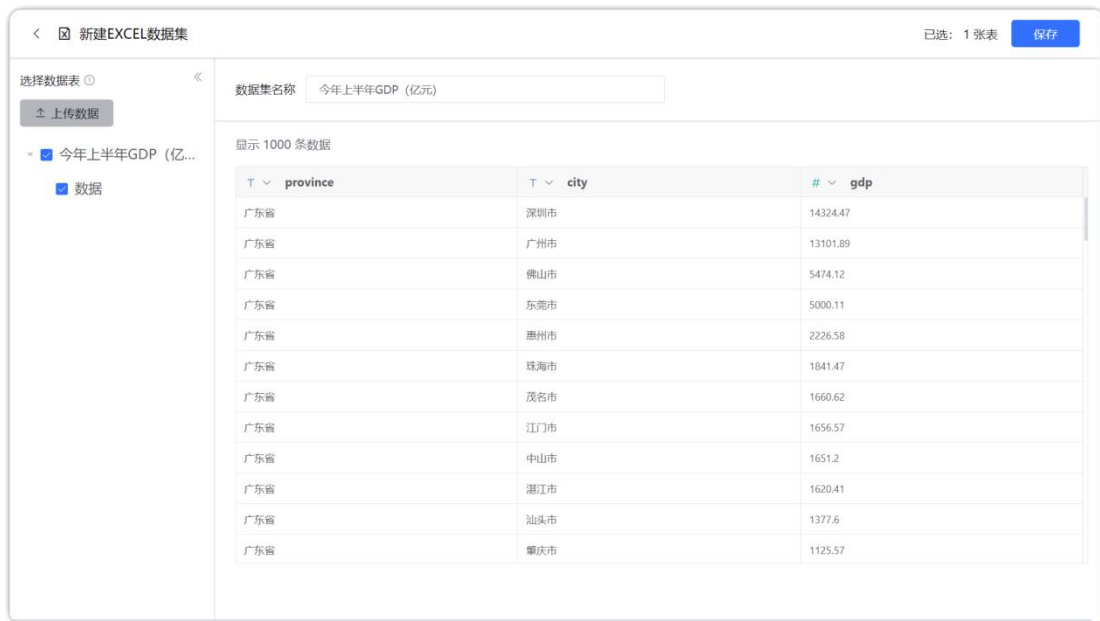


2 保存 Excel 数据集

以上传 Excel 文件为示例；

选择 Excel 文件，同一个 Excel 中可以勾选多个 sheet 页，填入数据集名称，可调整对应的字段类型；

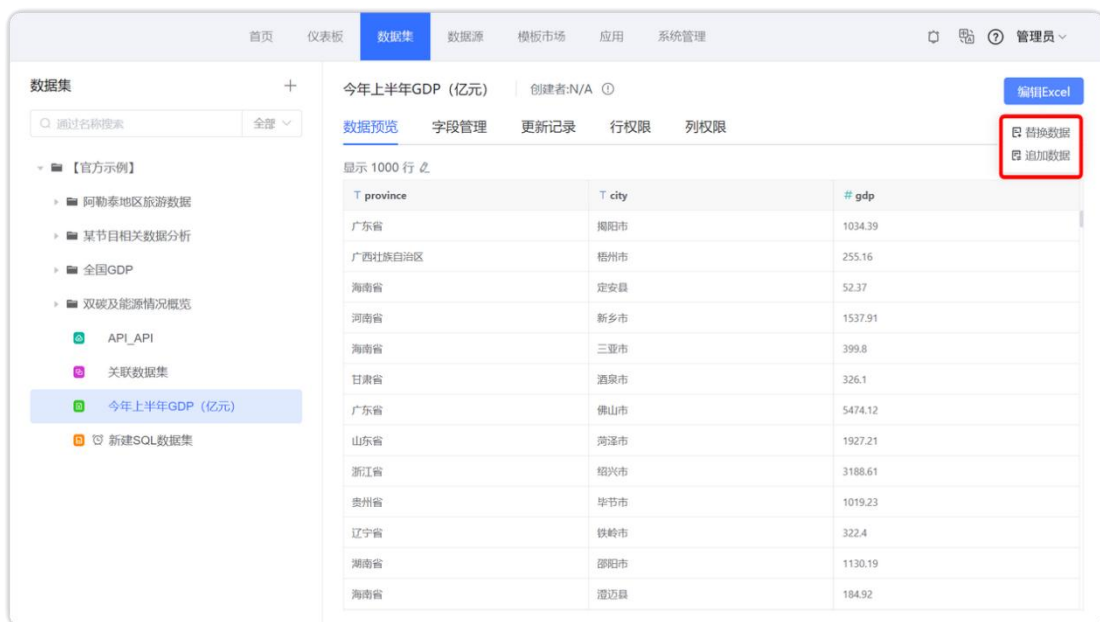
如下图所示，Excel 数据展示在【数据预览】区，若预览符合预期，点击【保存】，生成 Excel 数据集。



3 编辑 Excel 数据集

以上传 Excel 文件为示例；

Excel 数据集可以进行追加或替换，追加即不改变原有数据，增加新的数据，替换即重新替换掉原有的数据。



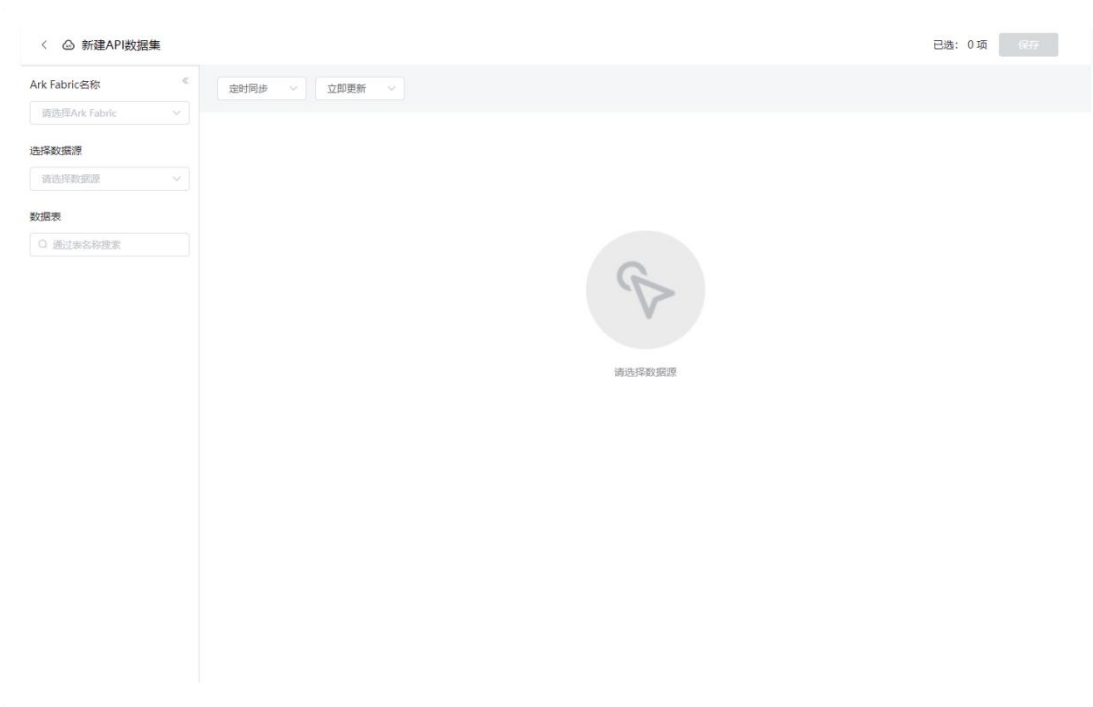
API 数据集

1 添加 API 数据集

点击【添加数据集】并选择【API 数据集】，选择【数据源】，选择【数据表】，数据存在 Ark Fabric 中。

2 保存 API 数据集

如下图所示，选定 API 数据表后，点击【保存】，生成 API 数据集。



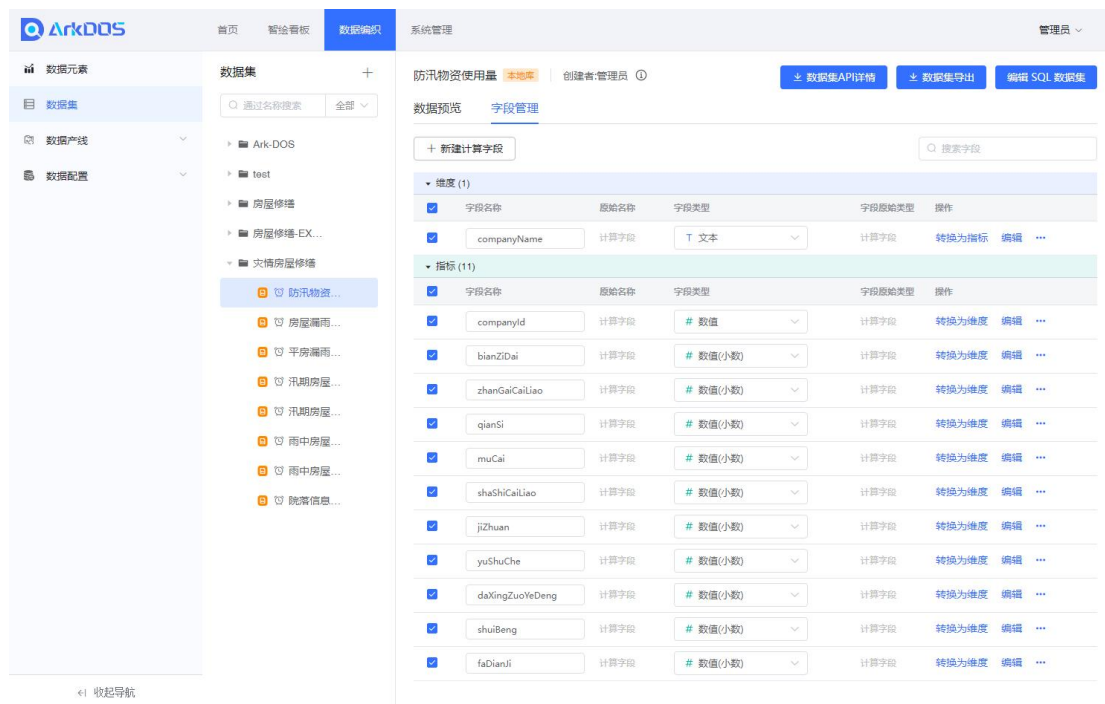
数据集功能设计

1 数据集预览

点击左侧数据表，右侧默认展示预览数据。

从字段管理 Tab 页，切换至数据预览 Tab 页时，若字段管理中没有任何变更，则不向后端发起数据请求；若【字段管理】中的发生变更操作，包括：新建、编辑、删除计算字段，对字段进行任意修改（比如是否选中、字段名、类型、维度指标转换），则向后端重新请求数据；

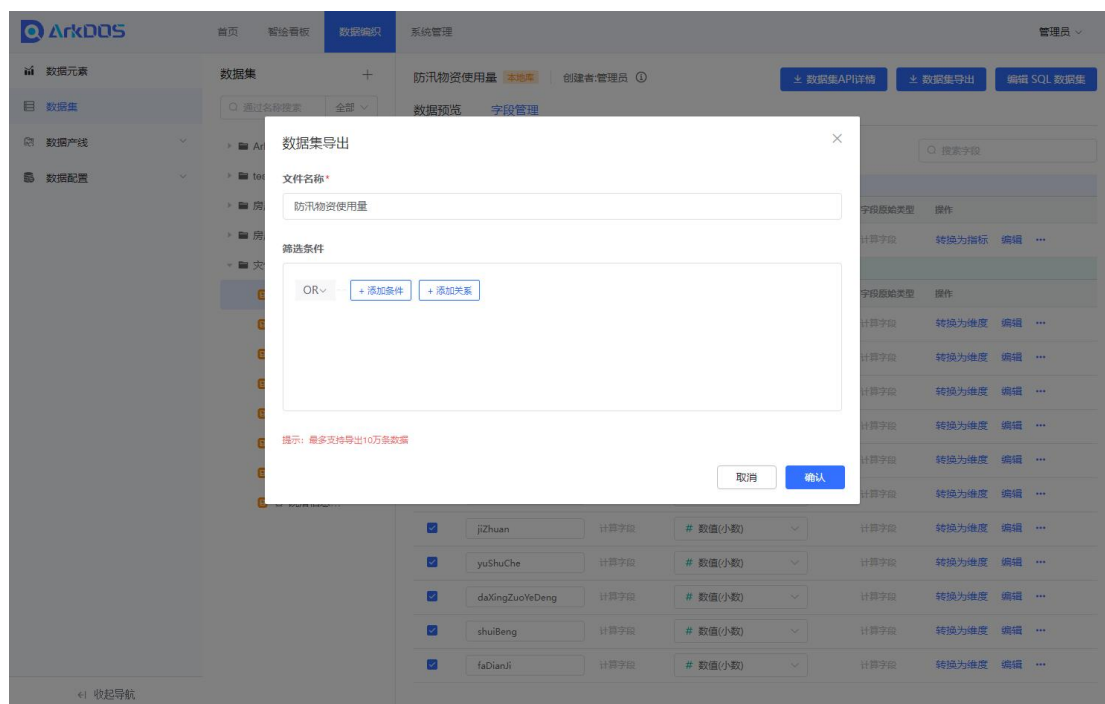
是否变更的判定标准：在字段管理 Tab 页中，向后端发起任何修改请求，即判定为发生变更。



2 数据集导出

支持按条件将数据导出 Excel。

注意：最多支持导出 10 万条数据。



3 字段管理

字段重命名

切换到字段管理页，修改字段名。

字段选中

字段显示支持全选/全不选，也可单独勾选。

更换字段类型

如下图所示，支持更换字段类型。

日期字段解析

示例：数据库中以 2000/01/01 文本存储的数据，由于格式差异性，日期类过滤组件无法过滤出来这些数据；

可将该文本字段改为时间类型，并选择 Y/m/d 格式（具体格式看数据源），则日期类过滤组件可以过滤出来这些数据。

更换指标维度

如下图所示，支持将字段转换为维度或指标。

新建计算字段

点击【新建计算字段】按钮，弹出新建计算字段页面，在此页面输入字段名，字段表达式，通过引用函数，生成新的数据集字段。

- 【序号 1】在此输入新建字段的名称；
- 【序号 2】字段表达式由引用的字段和函数组成，也可手动添加四则运算符号；
- 【序号 3】选择新建字段的数据类型，包括维度或指标；
- 【序号 4】根据引用的字段及函数计算结果，调整新建字段的字段类型；
- 【序号 5】引用字段栏，点击数据集原有的字段，字段表达式栏引用到对应的字段；
- 【序号 6】函数栏，使用的函数支持对应数据集的数据库函数。

The screenshot shows the '新建计算字段' (New Calculation Field) dialog box. It contains the following elements:

- Field Name (1):** A text input field containing '返回绝对值'.
- Field Expression (2):** A text input field containing 'ABS([id])'.
- Data Type (3):** Radio buttons for '维度' (Dimension) and '指标' (Indicator), with '指标' selected.
- Field Type (4):** Radio buttons for '文本' (Text), '时间' (Time), '数值' (Number), '数值(小数)' (Number (Decimal)), and '地理位置' (Geographic Location), with '数值' selected.
- Click to Reference Field (5):** A list of fields including '# id', '# 2021GDP', '# 2020GDP', and '# increase'.
- Click to Reference Function (6):** A list of functions including 'ABS(x)', 'ACOS(x)', 'ASIN(x)', 'ATAN(x)', 'CASE expr WHEN v1 THEN r1 [WHEN v2 THEN v2] [ELSE m] END', 'CEIL(x)', 'CHAR_LENGTH(str)', 'CONCAT(s1,s2,...)', 'COS(x)', and 'COT(x)'.
- Buttons:** '取消' (Cancel) and '确认' (Confirm) buttons at the bottom right.

请注意： 字段表达式支持所有的数据库函数以及简单的加减乘除，不仅是右侧列出的示例函数。比如：直连模式，数据源为 MySQL 数据库，那么字段表达式中可输入所有 MySQL 数据库的函数，包括 IF、CASE WHEN 等用法。已增加的计算字段不支持二次计算。

复制字段

支持对字段进行【复制】操作，复制后的字段相当于新建一个计算字段，复制后的字段，字段名：原字段名_Copy。

注意： 仅数据集字段管理的字段支持复制

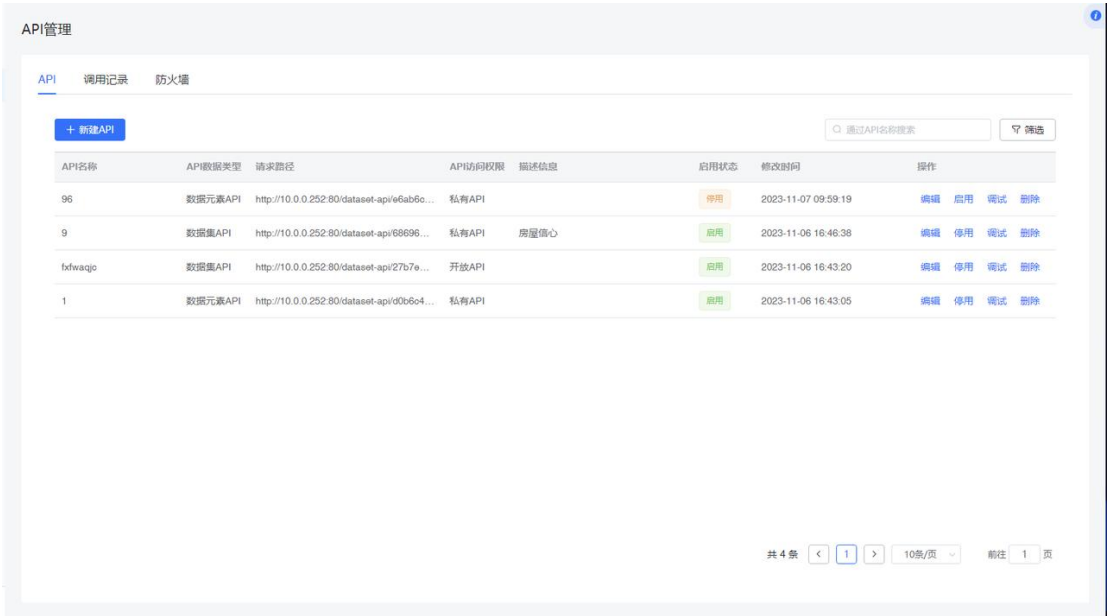
3.1.2.5 API 管理

功能概述

统一管理系统对外提供的数据服务。提供动态创建与发布数据接口的功能，可设置接口权限并监控外部系统对接口的调用。

功能说明

- a. 功能区域（上部）：API/调用记录/防火墙页面切换 Tab，新建 API 按钮、API 名称搜索框、筛选框（API 数据类型、访问权限、启用状态）。
- b. 内容区域（下部）：展示已经保存的 API 数据信息列表，展示 API 名称、API 数据类型、请求路径、访问权限、描述信息、启用状态、修改时间、操作（编辑、停用、调试、删除）。



新建 API

点击【新建 API】，选择【API 数据类型】，继续选择{已选数据类型下的某个数据集或数据元素}，点击下一步进入【API 接口信息】编辑，填写“接口名称”、系统自

动生成请求路径、描述信息、设置访问权限（私有 API 需要配置 token）、选择启用状态，保存即可完成相应配置。

1

2

API数据来源API接口信息

API数据类型*

☒ 数据集API

☐ 数据元素API

API数据来源*

汛期房屋安全检查-平房-Data

▼

取消

下一步

1

2

API数据来源API接口信息

接口名称 *

fxfwaqjc

请求路径 *

http://10.0.0.252:80/dataset-api/http://10.0.0.252:80/dataset-api/27b7e117-5bb9-46e

描述信息

请输入内容

访问权限 *

☐ 私有API ☒ 开放API

访问权限说明

请求私有接口时,需要把token值放入header的Authorization字段中携带,才可以访问成功。(如果是开放接口,不需要设置header)

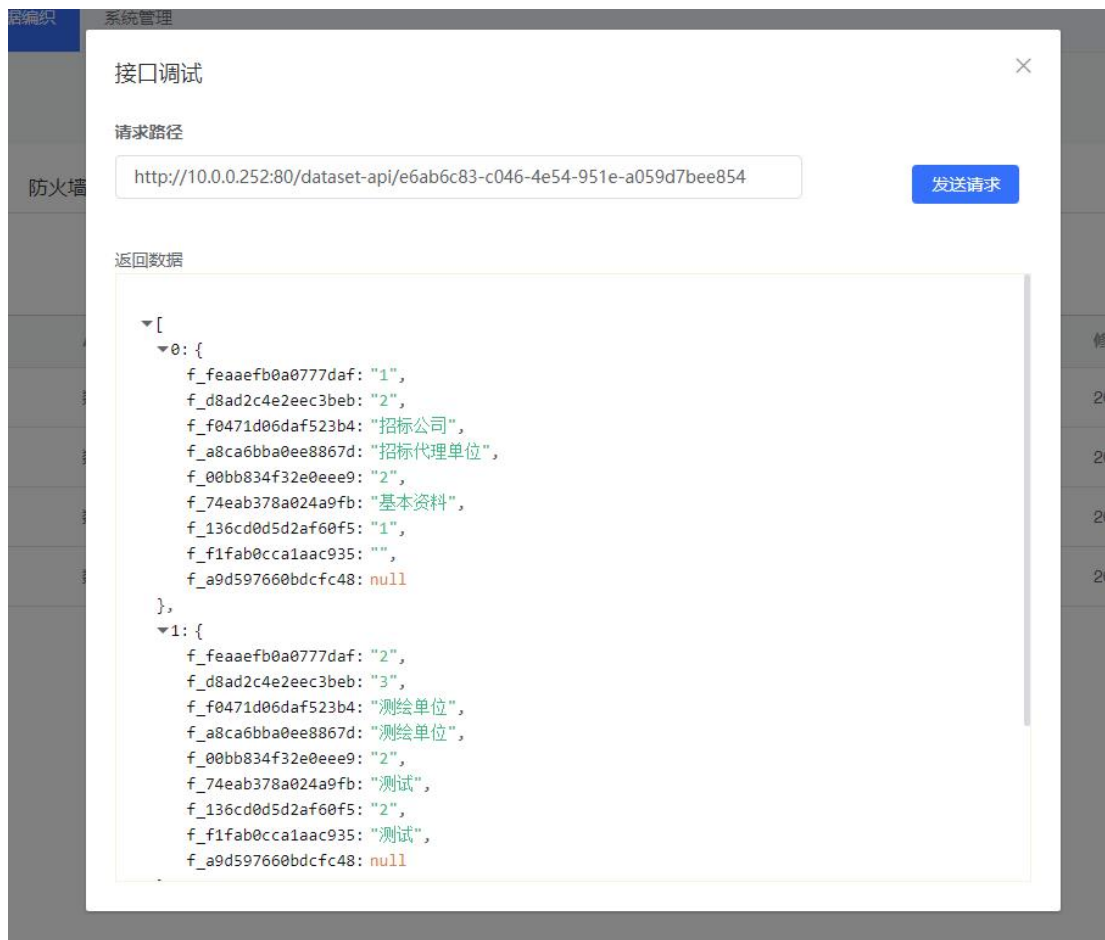
启用状态 *

☒ 立即启用 ☐ 暂时停用

返回保存

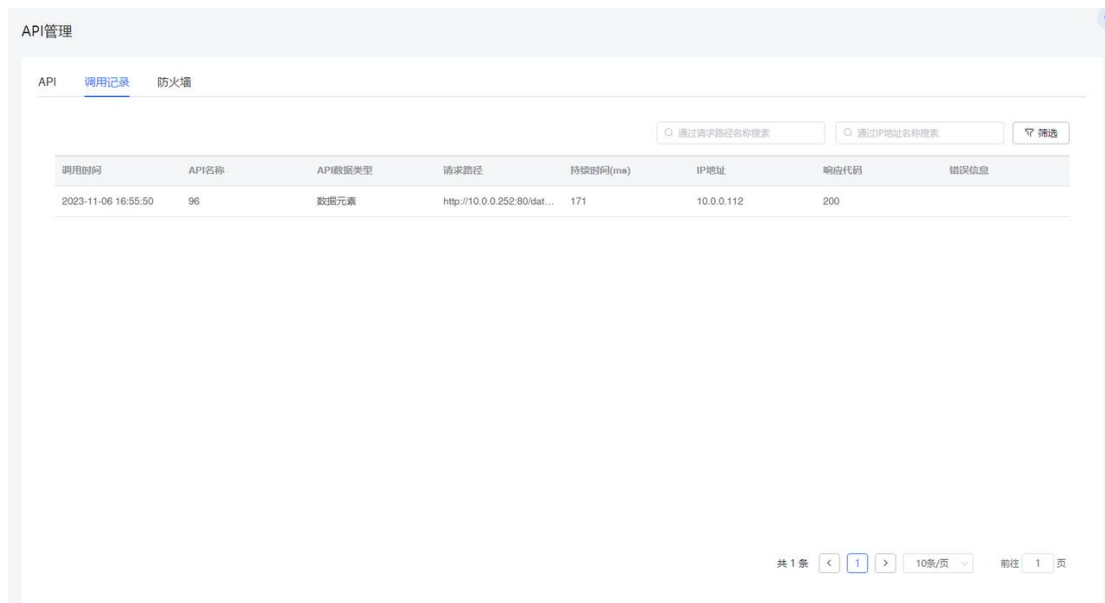
调试

点击某条 API 操作区的“调试”按钮，在弹窗中点击“发送请求”，即可查看到相应的返回数据。



调用记录

展示过往所有 API 的外部调用记录，提供相应的“请求路径”、IP 地址、日期范围等筛选条件。



防火墙

通过以 IP 列表的方式设定“黑名单/白名单”，实现可控的外部 IP 访问系统 API 服务。

防火墙

防火墙状态 ☒

访问权限 ☒ 黑名单 ☐ 白名单

说明信息

除了黑名单列表中的IP禁止访问API,其他IP一律允许访问

IP地址

请输入内容

保存

4. 附录

4.1 内计算引擎（ArkDB）常见问题

通过 `log_error` 指定错误日志文件的路径，错误发生时可以从该文件中查看相关信息。

与 arkdb 相关的常见错误信息，如下表所示：

错误信息	原因及处理办法
arkdb_cluster_namespace is needed, please set it	实例的 my.cnf 没有设置 arkdb_cluster_namespace，请设置之后再启动。
slave can not connect to master, slave starting abort:	读节点无法连接写节点，导致 slave abort，请检查写节点是否已经启动，并且读节点的 my.cnf 指定的写节点主机名称、用户名、密码、端口是否正确。
checked write heartbeat, Do you already have another arkdb server running on this pool	被系统通过 heartbeat 检查出当前打开了两个主节点，建议检查 my.cnf 配置参数。

write page error:....File too large	检查 ceph 存储中对单个对象是否有大小限制，如果有，将限制取消。
-------------------------------------	------------------------------------

4.2 内计算引擎（ArkDB）参数说明

4.2.1 启动参数

参数	值	含义
arkdb_confirm_single_write	0 或 1	默认值为 0，1 表示启动时不锁 ceph 对象，正常情况无需此参数，此参数一般在启动失败后使用。 如在一个 namespace 上已经存在写节点，造成无法启动。添加此参数即可启动。 在写节点挂掉时，且挂之前未能清除 namespace 锁，再次启动时需用此参数抢锁。

4.2.2 配置参数

读节点连接写节点需要的参数。

参数	值	含义
arkdb_master_hostname	字符串	写节点的 host 名称。
arkdb_master_username	字符串	登录写节点的用户名。
arkdb_master_password	字符串	登录写节点的密码。
arkdb_master_port	数值	写节点 mysql 端口。

arkdb_server_mode	slave	读节点模式。
-------------------	-------	--------

写节点连接读节点需要的参数。

参数	值	含义
arkdb_slave_username	字符串	登录读节点的用户名。
arkdb_slave_password	字符串	登录读节点的密码。
arkdb_server_mode	master	写节点模式。

以下是只读参数，需要在 my.cnf 中配置，主从要一致。

arkdb 系统参数变量	值	含义
arkdb_backend_config_file	字符串	实例连接 ceph 使用的配置文件。
arkdb_cluster_namespace	字符串	实例在 ceph 中使用的 pool 名称。
arkdb_enabled	布尔值	默认为 off，on 表示已经开启 arkdb 分布式模式。
arkdb_flow_control_max_delayed_lsn	整型	如主从复制延迟值大于当前值，则开始限流。 在限流后，如延迟值小于当前值，则自动关闭限流机制。
arkdb_flow_control_check_period	整型	默认值为 5。设置限流检查间隔时间，单位为秒。即每 5 秒检查一次限流和延迟状态，再采取措施。
arkdb_master_parallel_recovery	布尔值	默认值为 1。控制是否开启数据库并行恢复功能。如开启，则主库在数据库

		启动时，使用并行恢复来做 REDO 恢复，相比原来的单线程模式，会有所增速。
arkdb_master_retry_before_giveup	整型	默认值为 5。控制主库将一个从库踢出集群时，重试次数。当主库连不上从库时，会使用此参数判断是否将从库放弃（踢出）。
arkdb_pc_sleep_factor	布尔值	控制 page cleaner 线程的 sleep 因子，设置值越小刷盘越快。
arkdb_redo_storage_object_size	整型	默认值为 10M。设置在存储层，redo 日志文件被切割的对象大小。
arkdb_undo_storage_object_size	整型	默认值为 1G。设置在存储层，undo 文件被切割的对象大小。
arkdb_index_storage_object_size	整型	默认值为 16K。设置在存储层，ibd(ibdata)文件被切割的对象大小。
arkdb_slave_max_read_distance	整型	默认值为 10000000（10M），最大 1G，最小值为 0。在从库中设置主从延迟复制的延迟字节数。
arkdb_wait_lsn_advanced_lwm	整型	配合上一参数延迟复制功能，如复制时延迟字节数值未达到，即判断此时日志增长速度超过此参数值，则继续等待其达到所

		设置的延迟字节数值。如速度未达到此参数值，则不等待直接复制。
arkdb_log_writer_pause_max_time	整型	默认值为 50 微秒，最小值为 0，最大值为 1024。设置每次日志刷盘前，需要等待的时间。
arkdb_apply_thread_count	整型	最小值为 2，最大值为 1000。设置从库复制系统中应用线程的个数，即并发线程数。
arkdb_apply_thread_queue_size	整型	默认值为 10000，最小值为 1000，最大值为 100000。设置从库复制时，应用线程对应的工作队列长度。
arkdb_apply_logic_queue_size	整型	默认值为 10000，最小值为 1000，最大值为 100000。设置逻辑日志应用队列长度。
arkdb_access_not_running_slave	BOOL	默认值为 OFF。在未启动复制功能的从节点中执行语句，常见情况下是不安全的，可能会引发 ArkDB 数据错误。 如确认不会产生问题，或可接受问题发生造成的影响，可设置为 ON。
arkdb_local_state_uuid	UUID	标识一个节点。在集群不同节点之间相互交互时使用，集群中每个节点的值都不尽相同。

arkdb_log_file_on_cloud	BOOL	设置是否将 redo log 放到本地存储。此为实验性参数，为保证数据的正确性和 HA，请保持为默认值。
arkdb_master_wait_before_giveup	整型	控制读节点在无响应时，被踢出集群的时间长度，单位为秒。
arkdb_new_namespace	整型	用以辅助 ArkDB 测试事宜，普通用户无需关注。
arkdb_pre_read_log_size	整型	控制 REDO 读取时，所使用的异步 IO 队列大小。
arkdb_replication_adaptive_undo_weight	浮点型	控制多线程物理复制模块中 UNDO 与 Btree 线程初始比例。
arkdb_reset_slave	命令	清除从节点的 Buffer pool 页面，并且将 arkdb_slave_applied_lsn 置为 0。
arkdb_restatistic_table_now	命令	重新统计表属性。
arkdb_slave_report_to_master	命令	向写节点汇报自己的信息。
arkdb_start_cluster_with_snap	命令	用一个快照来启动。
arkdb_statistic_table_enabled	bool	ArkDB 表统计功能。
arkdb_adaptive_origin_now	命令	此命令一般不会用到。在需要更好的兼容原生 MySQL Server 时，将

		ArkDB 的数据通过 arkbak 备份到本地磁盘后，需要修改数据中的相关标识。
--	--	---

注：arkdb_enabled 设置为 on 之后初始化，数据将进入到 ceph 存储中，之后不能再改为 off。

4.2.3 推荐参数

参数名	推荐值	说明
innodb_log_recent_close_d_size	33554432	提升性能。
innodb_log_recent_writte_n_size	33554432	提升性能。
innodb_log_write_ahead_size	16384	提升性能。
innodb_write_io_threads	20	提升写入能力。
innodb_read_io_threads	20	提升读取能力。
innodb_doublewrite	OFF	在共享存储上，doublewrite 可以不用。

4.2.4 系统变量

所有配置参数均可查询，如：

```
mysql> select @@arkdb_enabled;
+-----+
| @@arkdb_enabled |
+-----+
|          1 |
+-----+
1 row in set (0.00 sec)
```

只读变量，见下表所示：

参数	值	含义
arkdb_cluster_uuid	uuid 值	集群唯一标识号。用以识别从库在启动时拿到的 UUID 是不是和 master 的 UUID 一致，防止出现数据错误等问题，只读参数，无需设置。
arkdb_master_make_clone_checkpoint	布尔值	内部使用。
arkdb_master_newest_checkpoint_lsn	数值	写节点最近检查点 LSN。
arkdb_master_newest_flushed_log_lsn	数值	写节点最近已刷盘的日志 LSN。
arkdb_master_newest_log_lsn	数值	最新的 LSN。
arkdb_slave_applied_lsn	数值	读节点应用 LSN。
arkdb_slave_replication_running	数值 0 或 1	读节点复制状态。
arkdb_server_status	字符串 running 或 startup	读节点运行状态。
arkdb_slave_delayed_bytes	数值	读节点应用 LSN 与写节点最新 LSN 的差值。

可修改的接口，见下表所示：

参数	值	含义
arkdb_server_mode	字符串 master 或 slave	arkdb 模式。
arkdb_slave_check_logs_period	数值	读节点读取日志记录的间隔时间。

arkdb_slave_md1_lock_timeout	数值	读节点查询 md1 锁超时时间。
arkdb_start_slave	数值 0 或 1	设置为 1 启动从节点复制。
arkdb_stop_slave	数值 0 或 1	设置为 1 停止从节点复制。
arkdb_master_newest_oldest_lsn_diff	数值	设置为 0 关闭热点页处理，设置其他值触发热点页处理。

4.3 内计算引擎（ArkOLAP）参数配置

参数名称	类型	默认值	说明
arkolap_assemble_thread_count	整型	10	控制插入到后端 Clickhouse 时合并数据的线程个数。
arkolap_assemble_thread_queue_size	整型	10000	控制每一个合并线程，在合并数据时所使用的缓存区大小。
arkolap_connection_addresses	字符串		设置后端 Clickhouse 服务所在 IP 和端口。如包括多个分片，则可通过 ip:port,ip:port 格式指定。
arkolap_connection_hostname	字符串		如只存在单个 Clickhouse 服务，其地址可用此参数

			单独指定。
arkolap_connection_password	字符串		如只存在单个 Clickhouse 服务，其登录密码可用此参数单独指定。
arkolap_connection_port	字符串		如只存在单个 Clickhouse 服务，其登录端口可以用此参数单独指定。
arkolap_connection_username	字符串		如只存在单个 Clickhouse 服务，其登录用户名可用此参数单独指定。
arkolap_distributed_write	BOOL	ON	开启分布式写入的模式。
arkolap_enabled	BOOL	ON	OLAP 存储引擎的总开关。如关闭，将无法创建对应存储引擎表，也不能写入，仅可查询。
arkolap_execute_thread_count	整型	10	控制合并之后的数据写入操作在执行阶段的线程个数。不同表的写入操作才可并行，相同表将会串行。
arkolap_execute_thread_queue_size	整型	10000	控制合并之后写入线程所使用的缓存大小。
arkolap_import_table	命令		将后端 Clickhouse 中的库表结构转换

			成 Arkolap 引擎并导入，指定的 Value 格式为“db_name=test;table_name=t1;”。
arkolap_max_merge_batch_size	整型	10000	控制在单条写入状态下，arkolap 存储引擎在将其合并时的最大行数。如达到此参数设定的值，将会执行后端写入操作。
arkolap_max_sync_cache_size	整型	50000	控制在 arkolap 引擎中缓存的数据量大小。如果超过此参数值数据量，则会触发引擎的流量控制，进而阻塞应用程序的写入操作，同时在日志中打印警告“arkolap cache overflow”。
arkolap_max_write_delay	整型	10	控制写入超时时间。如在写入合并时，行数尚未达到 arkolap_max_merge_batch_size 值，但已达到 arkolap_max_write_delay 所设定的超时时间，将会执行写入操作。
arkolap_optimize_level	ENUM	OLAPONLY	默认值为 OLAPONLY，表示只要存在

			arkolap 引擎的表，则直接使用 olap 查询。 控制在 htap 查询时，如何修改执行计划来得到更好的查询效率。 其它参数包括 AUTO、ALWAYS，这两种参数尚未完全支持。
arkolap_query_batch_size	整型	10000	控制在查询后端 Clickhouse 时，每一批查询的数据行数，即对应的 limit 大小。如此参数值过大，会导致分批数据过大，进而影响性能。
arkolap_reload_olap_funcs	命令		
arkolap_sync_error_mode	整型	1	控制在 arkolap 写入并同步到后端 Clickhouse 时，如出现错误，对应的容错模式。如果是 1，则停止同步；如果是 0，则忽略错误继续同步。
arkolap_sync_skip_counter	整型	0	此参数仅作用于 session 级别，用以控制在同步数据出错后，需跳过几

			个同步数据任务。 可用于处理同步错误，亦可用于跳过ddl，如 drop table 时不想删除后端 clickhouse 表，可设置 set session arkolap_sync_skip_counter=1;在 drop table 时将只删除本地表。
arkolap_sync_skip_from_binlog	BOOL	ON	控制是否忽略由 Binlog 产生的数据更新同步。如设置为 1，则不考虑用 Binlog 来补数据；如设置为 0，则在中断后，使用 Binlog 中 GTID 来控制数据一致性。尽可能保证同步写入的本地 InnoDB 或 ArkDB 引擎数据和后端数据的一致性。
arkolap_sync_start_at	整型	0	控制开启同步时的位置，此位置用 Binlog 中 GTID 后面的数字来指定。
arkolap_sync_state	ENUM	stopped	表示当前同步系统的状态，包括三个值，分别是 stopped（同步停步），prepare（处于准备阶

			段)， running （正常同步状态）。
arkolap_sync_drop_table	BOOL	OFF	默认值为 OFF 。为避免误操作， drop table 时不会 sync 到后端 clickhouse ，支持 session 级别设置。
arkolap_sync_create_table	BOOL	ON	默认值为 ON 。设置为 OFF 时， create table 不会 sync 到后端 clickhouse ，支持 session 级别设置。

4.4 Ark Fabric 参数配置

参数名称	类型	默认值	说明
arkfabric_dm_fetch_batch_size	整型	10000	在访问指定数据库时，每批数据的行数限定。如不满足参数设置值，则全部一次性返回；如超过参数设置值，则会分批查询。
arkfabric_clickhouse_fetch_batch_size	整型	10000	在访问指定数据库时，每批数据的行数限定。如不满足参数设置值，则全部一次性返回；如超过参数设置值，则会分批查询。

arkfabric_oracle_fetch_batch_size	整型	10000	在访问指定数据库时，每批数据的行数限定。如不满足参数设置值，则全部一次性返回；如超过参数设置值，则会分批查询。
arkfabric_mysql_fetch_batch_size	整型	10000	在访问指定数据库时，每批数据的行数限定。如不满足参数设置值，则全部一次性返回；如超过参数设置值，则会分批查询。
arkfabric_pg_fetch_batch_size	整型	10000	在访问指定数据库时，每批数据的行数限定。如不满足参数设置值，则全部一次性返回；如超过参数设置值，则会分批查询。
arkfabric_starrocks_fetch_batch_size	整型	10000	在访问指定数据库时，每批数据的行数限定。如不满足参数设置值，则全部一次性返回；如超过参数设置值，则会分批查询。
arkfabric_es_fetch_batch_size	整型	10000	在访问指定数据库时，每批数据的行数限定。如不满足参数设置值，则全部一次性返回；如超过参数设置值，则会分批查询。
arkfabric_engine_pushdown	BOOL	ON	在执行查询语句时，是否将语句中的条件下推到存储引擎。

arkfabric_connection_pool_max_size	整型	50	内部连接池最大值。 odbc 连接串和用户名的组合可表示为一个连接池。设置后，在此组合下最多有 arkfabric_connection_pool_max_size 个连接。
arkfabric_optimize_level	ENUM (auto, never, single_user_pushdown)	ON	在执行查询语句时，是否实施对整个语句的优化，用以提升整体性能。如值为 single_user_pushdown，表示在认定两个表是否同一个数据源时，除使用 host 和 port 外，增加了对 user 名字的判断。如是不同的用户名，则不会被认定为相同数据源。 此参数只在达梦数据库和 Oracle 数据库中生效，为防止在这两种数据库中，不同用户下存在同名表时，实际只会查一个表的情况。
arkfabric_connection_pool_max_size	整形	50	控制 arkfabric 与后端数据库建立连接池的最大连接个数，值的范围是（10，1000）。设置此参数后，当前已经建立的连接池不会生效。如想让其生效，需要通过 flush tables 命令将已经缓存的表清空。再使用表时，会自动建立

			一个新设置值的连接池。
arkfabric_license_expire_date	字符串	空	arkfabric 产品 License 截止时间。
arkfabric_license_file_name	字符串	空	arkfabric 的 license 文件全路径。需要指定一个正确的 key 文件后，arkfabric 才可以被正常启动，如需在线更新 key 文件，可指定一个新的 key 文件路径，更新此参数即可。
arkfabric_read_only	布尔值	ON	控制是否可在 arkfabric 表中使用 INSERT、UPDATE、DELETE、TRUNCATE，四种修改数据的操作。默认值为开启，在执行这些操作时，将会报错。
arkfabric_clickhouse_driver_file_name	字符串	/usr/lib64/libclickhouseodbcw.so	指定 clickhouse 数据库的 odbc 驱动程序文件所在路径。参数可以动态修改，但修改后只能在下次重连时才生效。如想强制重连，则需在执行 flush tables 或 select arkfabric_connection_pool_clear()函数后，下同。
arkfabric_dameng_driver_file_name	字符串	/usr/lib64/mysql/drivers/dameng/libdodbc.so	指定达梦数据库的 odbc 驱动程序文件所在路径。

arkfabric_greenplum_driver_file_name	字符串	/usr/pgsql-15/lib/psqlodbc*.so	指定 greenplum 数据库的 odbc 驱动程序文件所在路径。
arkfabric_mysql_driver_file_name	字符串	/usr/lib64/libmyodbc8w.so	指定 MySQL 数据库的 odbc 驱动程序文件所在路径。
arkfabric_oracle_driver_file_name	字符串	/usr/lib/oracle/19.19/client64/lib/libsqora.so.19.1	指定 Oracle 数据库的 odbc 驱动程序文件所在路径。
arkfabric_starrocks_driver_file_name	字符串	/usr/lib64/libmyodbc8w.so	指定 Starrocks 数据库的 odbc 驱动程序文件所在路径。
arkfabric_postgres_driver_file_name	字符串	/usr/pgsql-15/lib/psqlodbc*.so	指定 Postgres 数据库的 odbc 驱动程序文件所在路径。
arkfabric_odbc_driver_connect_mode	布尔	off	指定 arkfabric 连接后端数据库的模式。如是 on 则无需配置 odbc.ini 文件，如是 off 则需要配置此文件。
arkfabric_query_timeout	整型	900	设置 arkfabric 在后端数据库中查询数据时，语句执行的超时时间。执行时间如果超过这个参数所设置的值，将自动返回错误信息并提示超时。目前仅对 arkfabric_clickhouse 引擎有效。

4.5 Ark Fabric 常见问题

驱动找不到 (参照问题如 `Can't open lib`

`‘/usr/lib/oracle/19.19/client64/lib/libsqora.so.19.1’: file not found`)

建议从如下几方面排查:

- a. 检查文件是否真实存在, 如 `ls -l /usr/lib/oracle/19.19/client64/lib/libsqora.so.19.1` 是否能列出驱动;
- b. 检查 `selinux` 的安全策略, 使用 `getenforce` 命令检查当前模式: 不能是 `Enforcing` 模式, 可以是 `Disabled`, `Permissive` 模式;
- c. `LD_LIBRARY_PATH` 环境变量是否配置正确, 对于 `systemd` 启动的 `Ark Fabric`, 需要在 `/etc/sysconfig/mysql` 文件中正确配置使其包含 `LD_LIBRARY_PATH=/usr/lib/oracle/19.19/client64/lib`。

4.6 Ceph 集群的设置和检查

4.6.1 授权

检查 `Ceph` 集群中是否已存在专用的用户, 如不存在则可使用下面的方法创建 `client.arkfabric` :

Bash

```
ceph auth get-or-create client.arkfabric \
    mon 'allow *' \
    mds 'allow *' \
    mgr 'allow *' \
    osd 'allow *' \
    -o /etc/ceph/ceph.client.arkfabric.keyring
```

创建完成后, 即可使用 `arkfabric` 用户访问 `ceph`。

可使用如下 `python` 脚本校验是否有权访问 `ceph`(按需修改):

Python

```
#!/usr/bin/env python3
# -*- coding: UTF-8 -*-
# 需要安装 python3-rados (15.2.17)软件包, 以 centos7 下为例:
# yum install python3-rados

import rados, sys

cluster = rados.Rados(conffile = '/etc/ceph/ceph.conf',
```

```

        conf = dict (keyring =
'/etc/ceph/ceph.client.arkfabric.keyring'),
        name = 'client.arkfabric')
print ("librados version: " , str(cluster.version()))
cluster.connect()
print ("Cluster ID: " , cluster.get_fsid())
print ("Cluster Statistics:")
print ("=====")
cluster_stats = cluster.get_cluster_stats()
for key, value in cluster_stats.items():
    print (key, value)

```

4.6.2 检查写入对象限制

可使用如下 python 脚本检查 Ceph 集群是否允许写入 4G 大小的 rados 对象:

```

Python
#!/usr/bin/env python3
# -*- coding: UTF-8 -*-

# 需要安装 python3-rados (15.2.17)软件包，以 centos7 下为例：
# yum install python3-rados

import rados

try:
    cluster = rados.Rados(conffile = '/etc/ceph/ceph.conf',
        conf = dict (keyring =
'/etc/ceph/ceph.client.arkfabric.keyring'),
        name = 'client.arkfabric')
except TypeError as e:
    print('Argument validation error: ', e)
    raise e
print("Created cluster handle.")

try:
    cluster.connect()
except Exception as e:
    print("connection error: ", e)
    raise e
finally:
    print("Connected to the cluster.")

poolname="test_obj_size"

```

```

print("Create Pool:", poolname)
print("-----")
cluster.create_pool(poolname)
ioctx = cluster.open_ioctx(poolname)
ioctx.write("hw", b"!"*1024, 0)

for i in range(1024*4 - 1):
    ioctx.append("hw", b"!"*1024*1024)
    print(ioctx.stat("hw"))
ioctx.remove_object("hw")
cluster.delete_pool(poolname)

```

或者使用 rados 命令测试:

Bash

需要安装 ceph-common (15.2.17)软件包(包含 rados 命令), 以 centos7 下为例:

```
# yum install ceph-common
```

```
dd if=/dev/zero of=testfile1 bs=1024 count=$((4*1024*1024))
```

```

rados --keyring=/etc/ceph/ceph.client.arkfabric.keyring \
      --name=client.arkfabric \
      --conf=/etc/ceph/ceph.conf \
      --pool=test_obj_size \
      put testfile1 testfile1

```

如果不允许, 需要修改 Ceph 集群配置:

Plain Text

```
→ # vim /etc/ceph/ceph.conf
```

```
[global]
```

```
    osd max object size = 4294967295
```