-

实时计算-产品白皮书



JDD深海大数据平台 出品

JDD深海大数据平台 出品

一个提供高性能、低延迟，处理实时流式大数据的计算服务平台

JDD深海大数据平台 出品一个提供高性能、低延迟，处理实时流式大数据的计算服务平台

目录

[1. 业务问题与挑战 2](#_Toc60324208)

[2. 产品定位与优势 2](#_Toc60324209)

[2.1. 产品定位 2](#_Toc60324210)

[2.2. 产品价值 2](#_Toc60324211)

[2.3. 产品优势 3](#_Toc60324212)

[2.4. 产品架构 3](#_Toc60324213)

[3. 应用场景 4](#_Toc60324214)

[4. 产品功能概述 5](#_Toc60324218)

[4.1. 项目管理 5](#_Toc60324219)

[4.2. 任务管理 5](#_Toc60324220)

[4.3. 告警策略 6](#_Toc60324221)

[4.4. 文件管理 7](#_Toc60324222)

[4.5. 数据权限 7](#_Toc60324223)

[4.6. 源管理 8](#_Toc60324224)

[4.7. 系统管理 8](#_Toc60324225)

[5. 行业客户案例 11](#_Toc60324226)

[5.1. 【金融行业】银行机构 11](#_Toc60324227)

[5.2. 【金融行业】某消费金融公司 12](#_Toc60324228)

[5.3. 【智能城市】智能AI防控项目 13](#_Toc60324229)

[6. 法律声明 14](#_Toc60324230)

1. 业务问题与挑战

随着互联网的快速发展，每天产生的数据正义指数级的速度增长，对这些大量数据的处理与分析有着巨大的价值。以Hadoop为代表的大数据技术的出现，可以很好地解决大量静态数据集的数据处理与分析，但很多数据都是实时产生的，当下的需求方更希望实时处理这些数据，这就需要流计算处理技术来实时处理这些数据，及时产出数据应用价值。

在当前大数据背景下，优化发展实时计算平台技术，可以有助于解决现有计算平台处理海量数据实时响应能力差的问题，提升计算平台处理大数据的速度，确保在大数据背景下，可以应用实时技术进行实时数据挖掘、实时指标计算等内容。

对于公司内部来说，建立统一的实时计算平台能够很好地应对各个业务条线的业务场景的不断变化创新和增多，而带来的业务对数据的时效性增高。从数据上看，当前实时数据与离线数据融合十分困难，无法满足批流结合的业务场景需要。

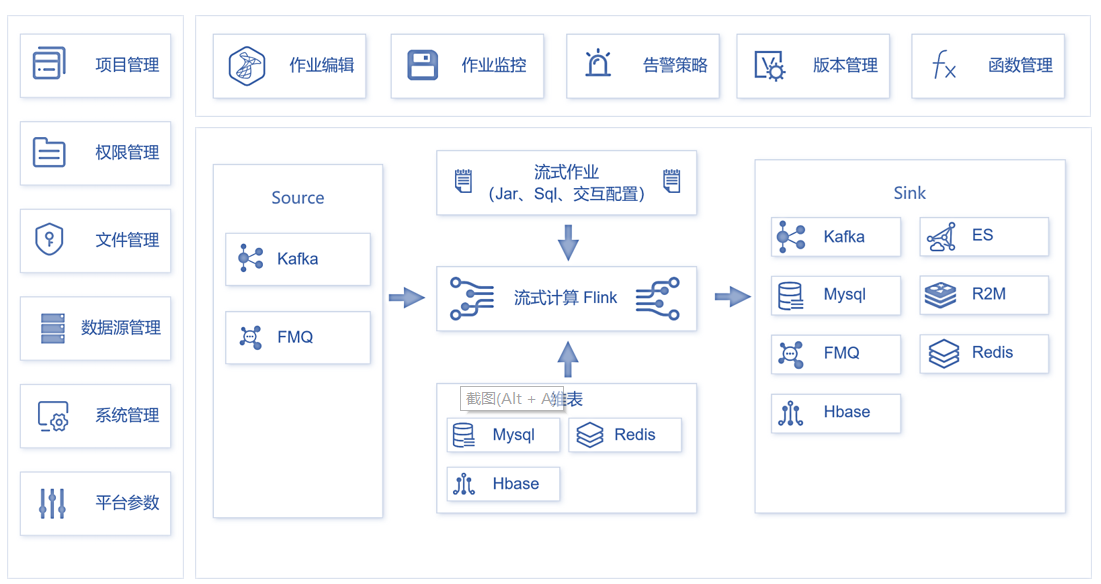
1. 产品定位与优势
   1. 产品定位

实时计算平台是一个提供高性能、低延迟，处理实时流式大数据的计算服务平台，是京东科技数据中台提供实时数据服务的核心平台。实时计算可以实现多种数据格式的接入或落地，通过对Flink的深度调优，在计算能力上轻松支持PB级毫秒级别的处理延迟，用户可以通过简单的SQL或自动化Sql配置，实现实时推荐、实时风控、实时指标等复杂的实时场景应用。

* 1. 产品价值
     1. 多数据源支持：支持各类常见数据源，部署简单功能轻便，可作为随时插拔的实时计算模块；
     2. 多开发方式：支持多种任务开发方式，jar、sql、自由交互配置的自动化sql，满足所有程度的开发者上手使用；
     3. 丰富的监控与告警：性能监控、数值监控，异常监控等丰富的监控场景，覆盖所有情况，并支持多种告警配置方式，简单便捷。
  2. 产品优势

平台的核心计算引擎是依托于Apache Flink技术来实现的，完美的继承了Flink的全部优点，并且对其不足之处也进行了改进。

* 低延迟：Flink采用纯流式计算(逐行计算)方式，并且使用异步snapshot来实现exactly once，snapshot不会影响到数据流的处理，最低可以达到毫秒级的延迟。
* 高吞吐：平台的实时计算引擎通过优化，单线程最大能达到每秒35万的数据处理能力，轻松应对任何大型活动的实时计算场景。
* 高容错：平台提供Checkpoint和Savepoint两种机制来保证任务的数据完成性；平台还提供了任务自动重启机制，保证任务的稳定性。
* 傻瓜化：提供了SQL编辑模式，只要懂得SQL编写的人，就可以用很简单的SQL逻辑完成非常复杂的实时计算代码逻辑，降低了开发门槛，也缩短了开发周期。
* 可视化：平台打破了实时任务监控暗盒情况，通过平台提供的任务监控，可以很清晰的知道当前的任务运行情况，数据流入流出情况，以及当前数据源流入和结果流出情况，让用户对自己的任务知根知底。
* 安全性：提供实时任务版本回滚对比，方便进行任务恢复、排查等；平台全部任务都为独立运行(隔离运行)，在任何情况下任务之间都不会出现相互影响的情况，保证任务的运行安全性;
  1. 产品架构

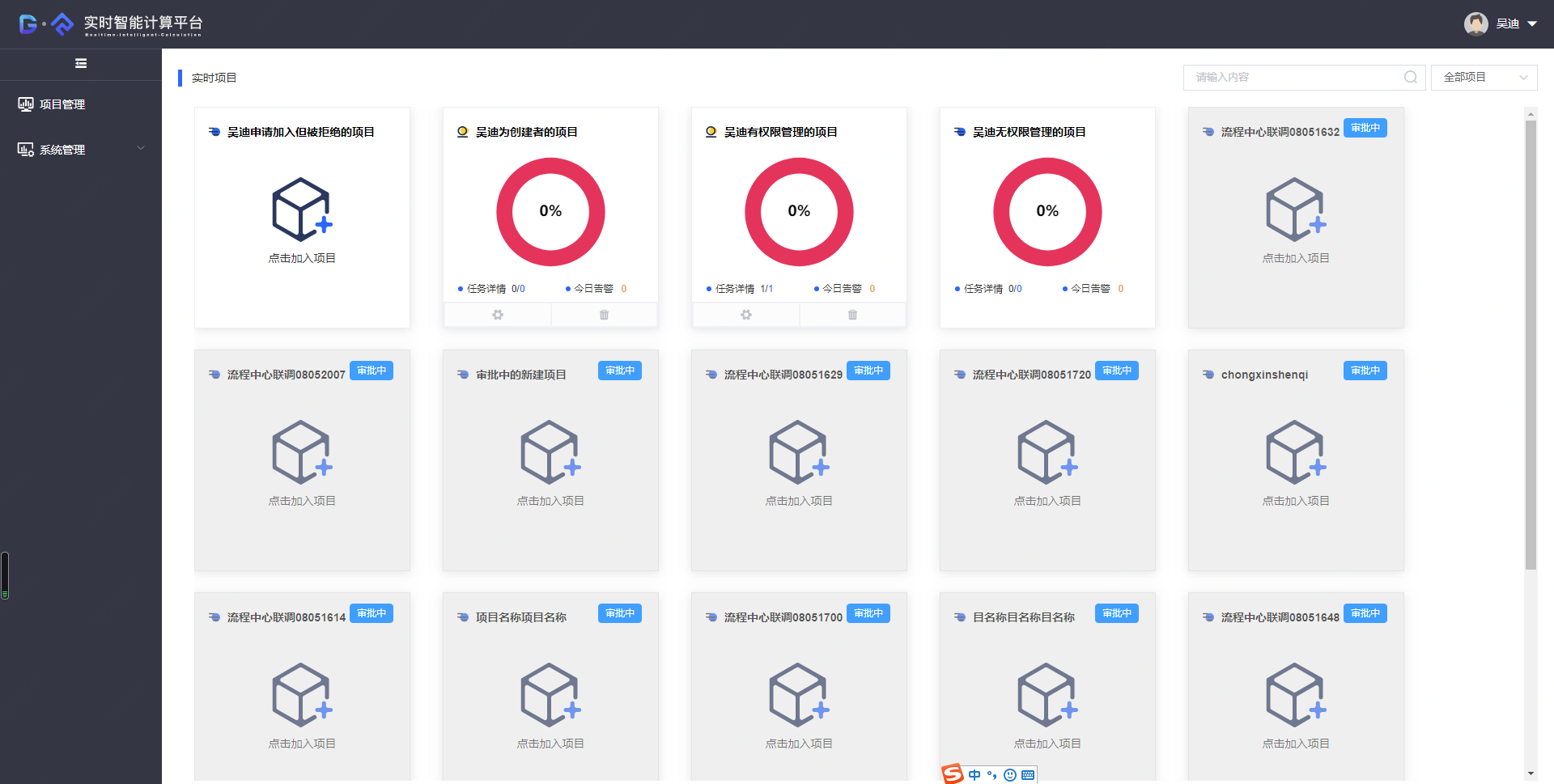


|  |  |
| --- | --- |
| 场景 | 场景描述 |
| 金融服务场景 | 支持实时指标、实时标签、实时报表的计算，并应用于实时风控策略、实时异常告警等金融高风险业务需求，支持接入各类外部数据实时消费 |
| 城市服务场景 | 支持实时流式计算，将各类实时监控、传感器等物联网数据整合计算并输出，全场景覆盖实时智能服务与监控 |
| 电商交易场景 | 支持实时推荐、实时分析等场景的计算，极大提升用户转化率，将最适合的物品实时推荐到用户面前 |
| 政企战略场景 | 实时决策、实时监测、实时大屏可将繁杂来源的复杂数据，实时分类、分析、处理计算后展示，帮助政企根据实时数据决策战略 |
| 媒体文娱行业 | 实时风控、实时挖掘、实时分析，帮助媒体文娱行业，在海量的资讯中，实时抓取热点方向与爆发问题，将行业时效性发挥极致 |

1. 应用场景
2. 产品功能概述
   1. 项目管理

项目管理模块中，用户可以创建一个实时项目，可以选择运行的集群和所需的项目总资源还可以对项目的成员进行管理。

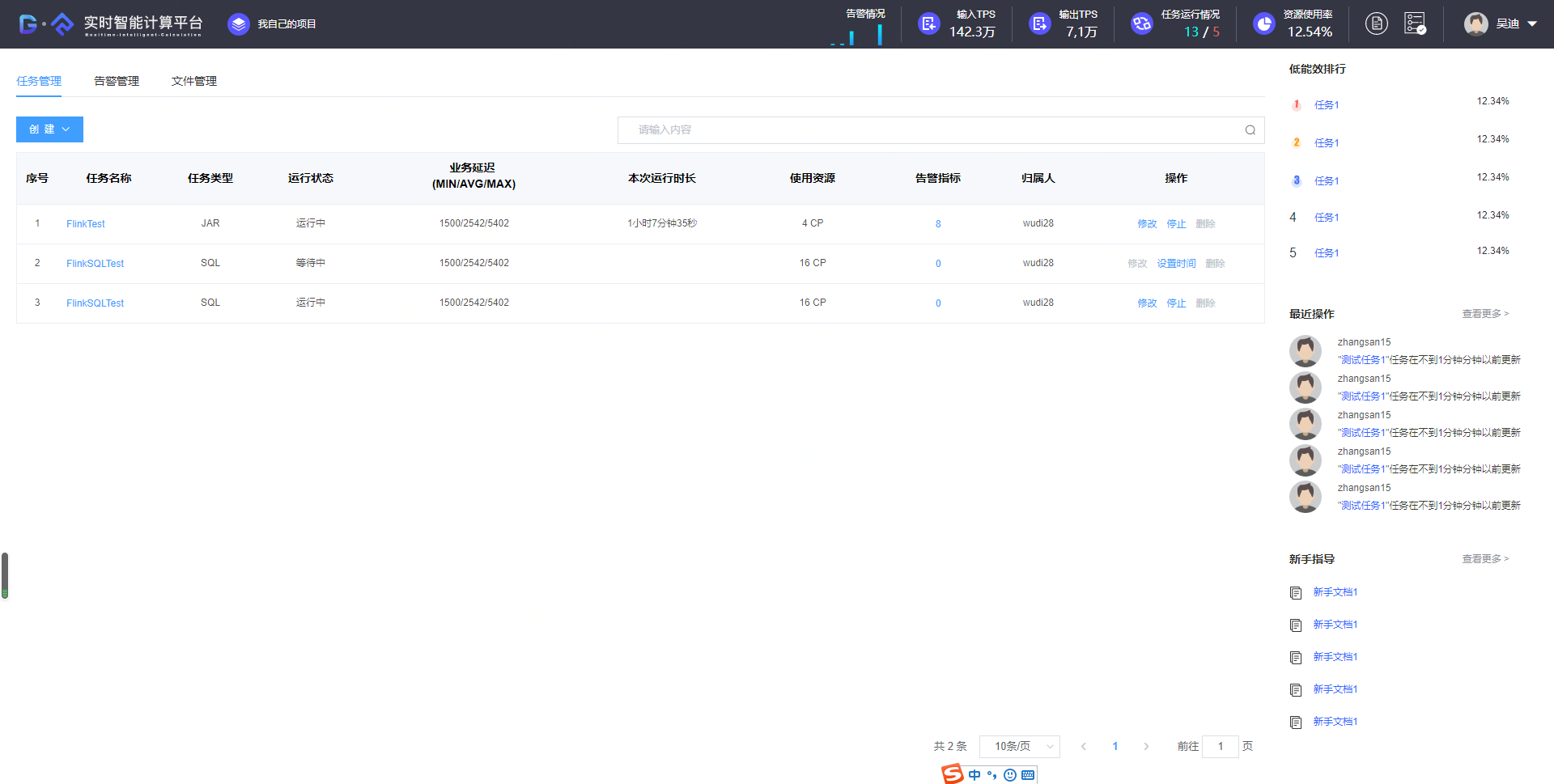
* + - * + 项目列表：



* 1. 任务管理

所有实时任务都必须存在于某一个项目之下。在任务管理模块中，用户可以创建，修改，删除，维护和查看实时任务，了解任务情况，目前支持jar任务、sql任务、语意任务。

* + - * + 任务列表：



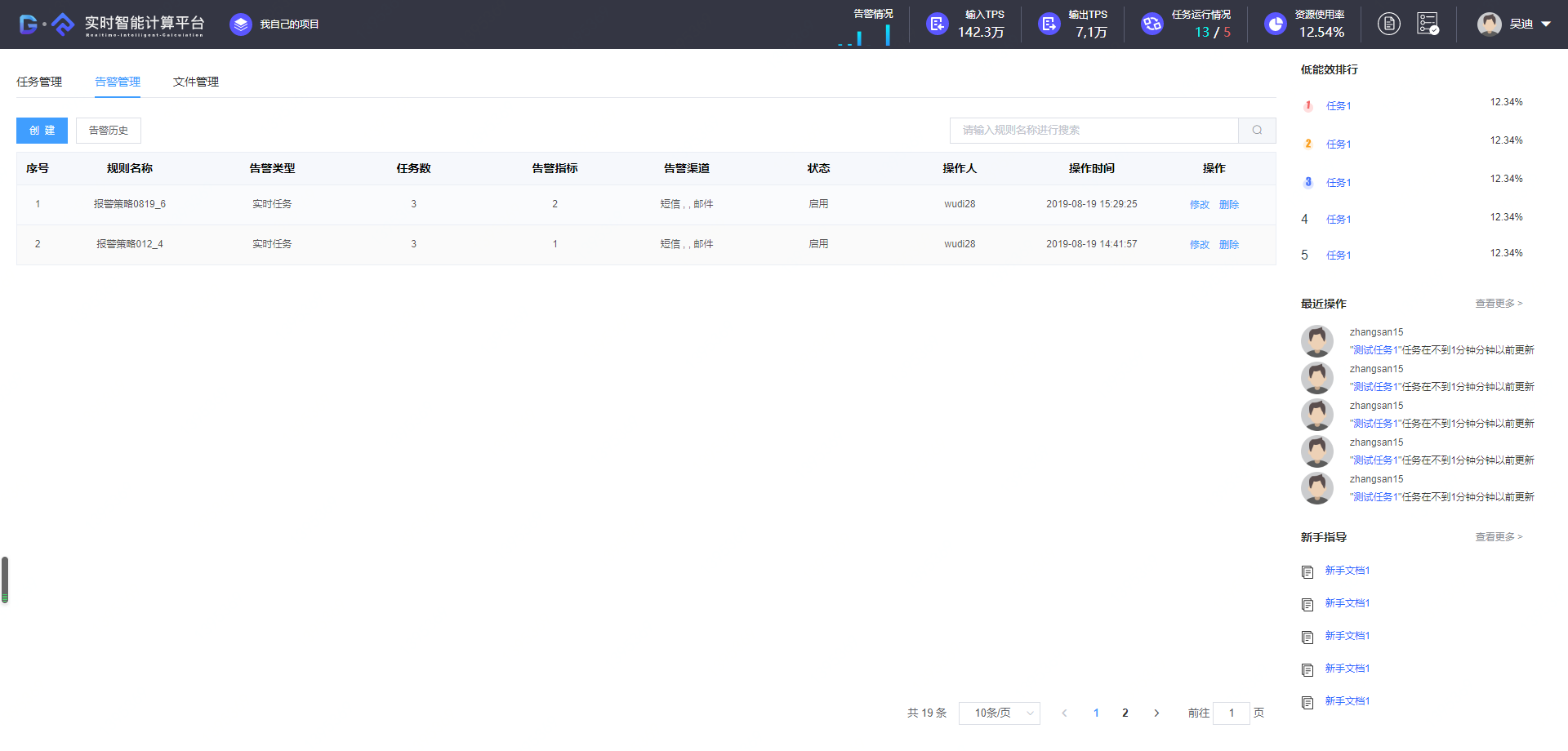
* + - * + 任务详情-运维监控：



* 1. 告警策略

告警策略主要用配置任务的相关告警策略，可以配置全局告警，也可以配置个别任务的告警阈值。

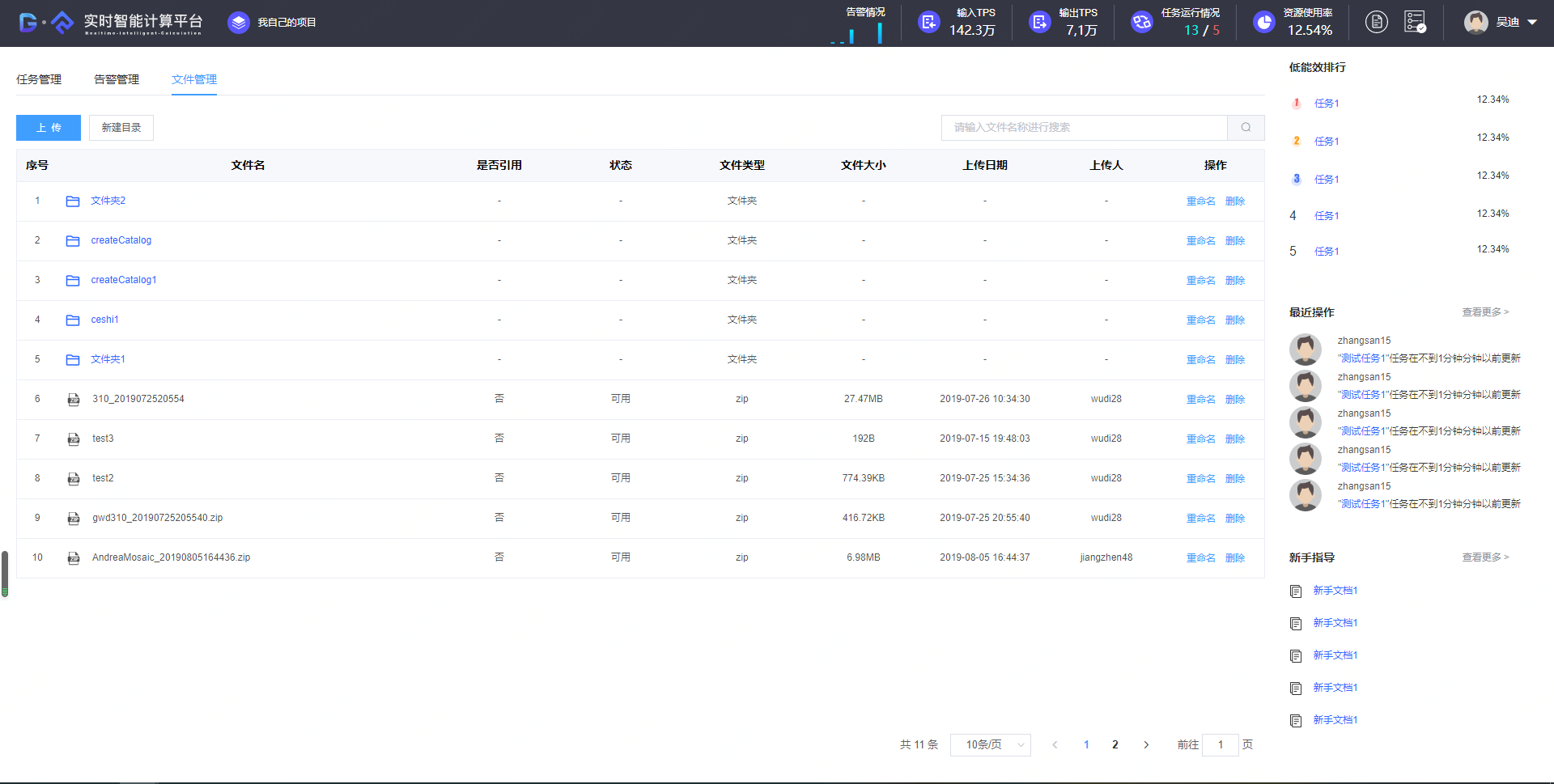
* + - * + 策略列表：



* 1. 文件管理

文件管理用于用户创建Jar包类型实时任务时，对Jar包进行预上传处理，用户可在此模块提前上传相应的Jar包用于实时任务创建或修改，还可以对历史Jar包进行管理。

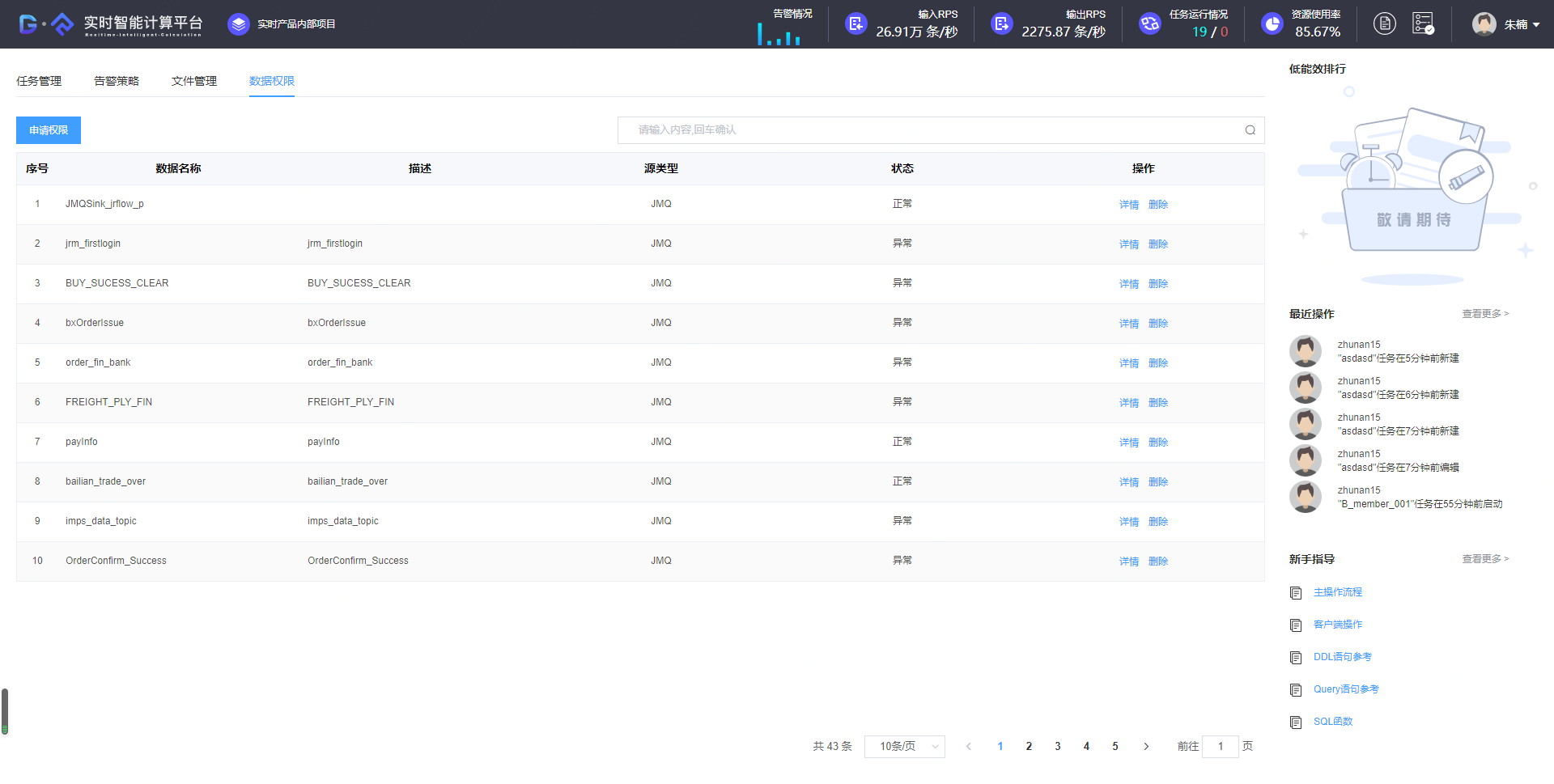
* + - * + 文件列表：



* 1. 数据权限

主要用于项目中使用SQL类型实时计算任务时，指定topic的访问权限申请。

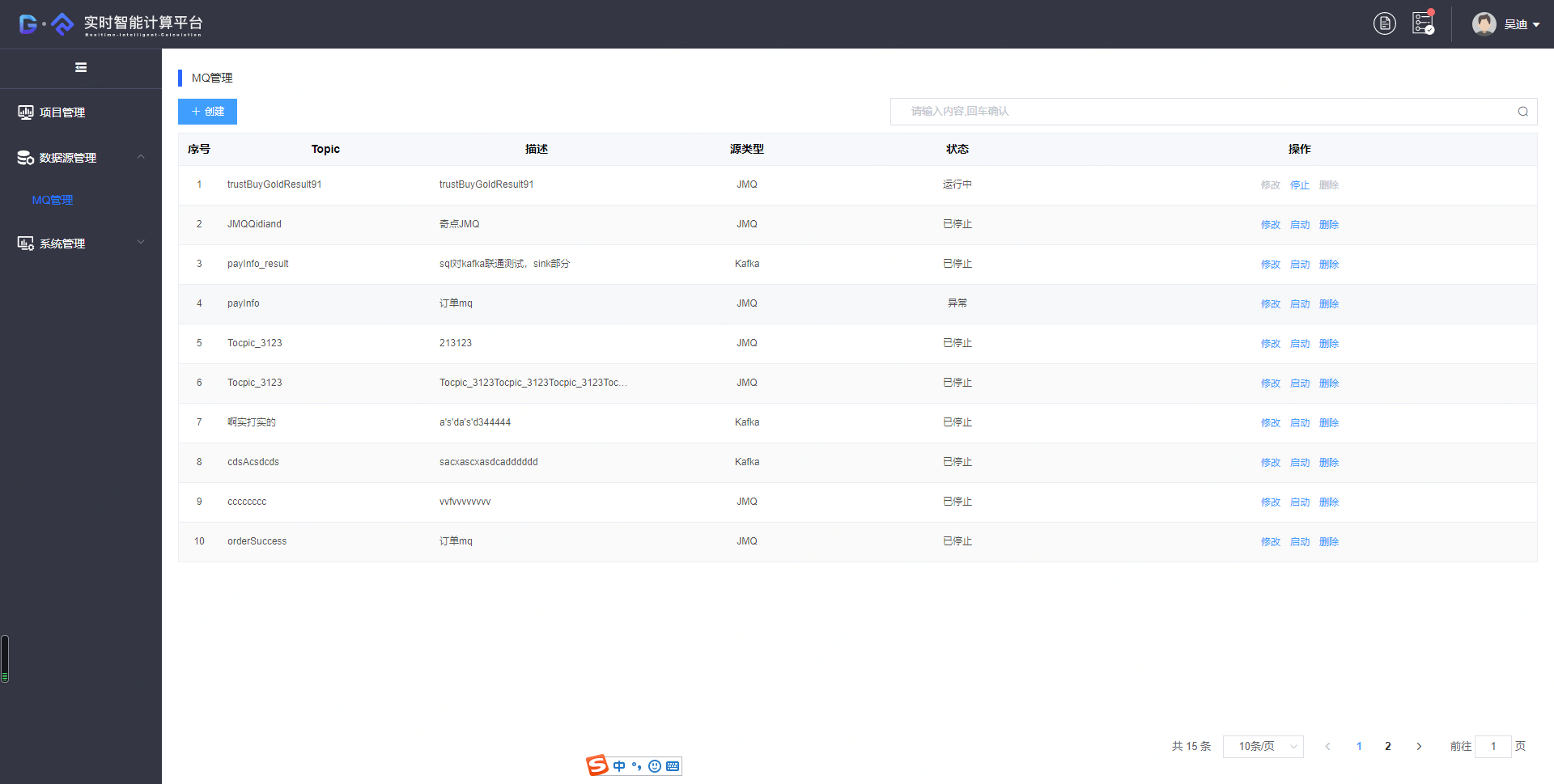
* + - * + 权限列表



* 1. 源管理

管理输入的数据源，主要用于SQL与实时指标任务的统一数据接入功能，目前可以支持FMQ和Kafka

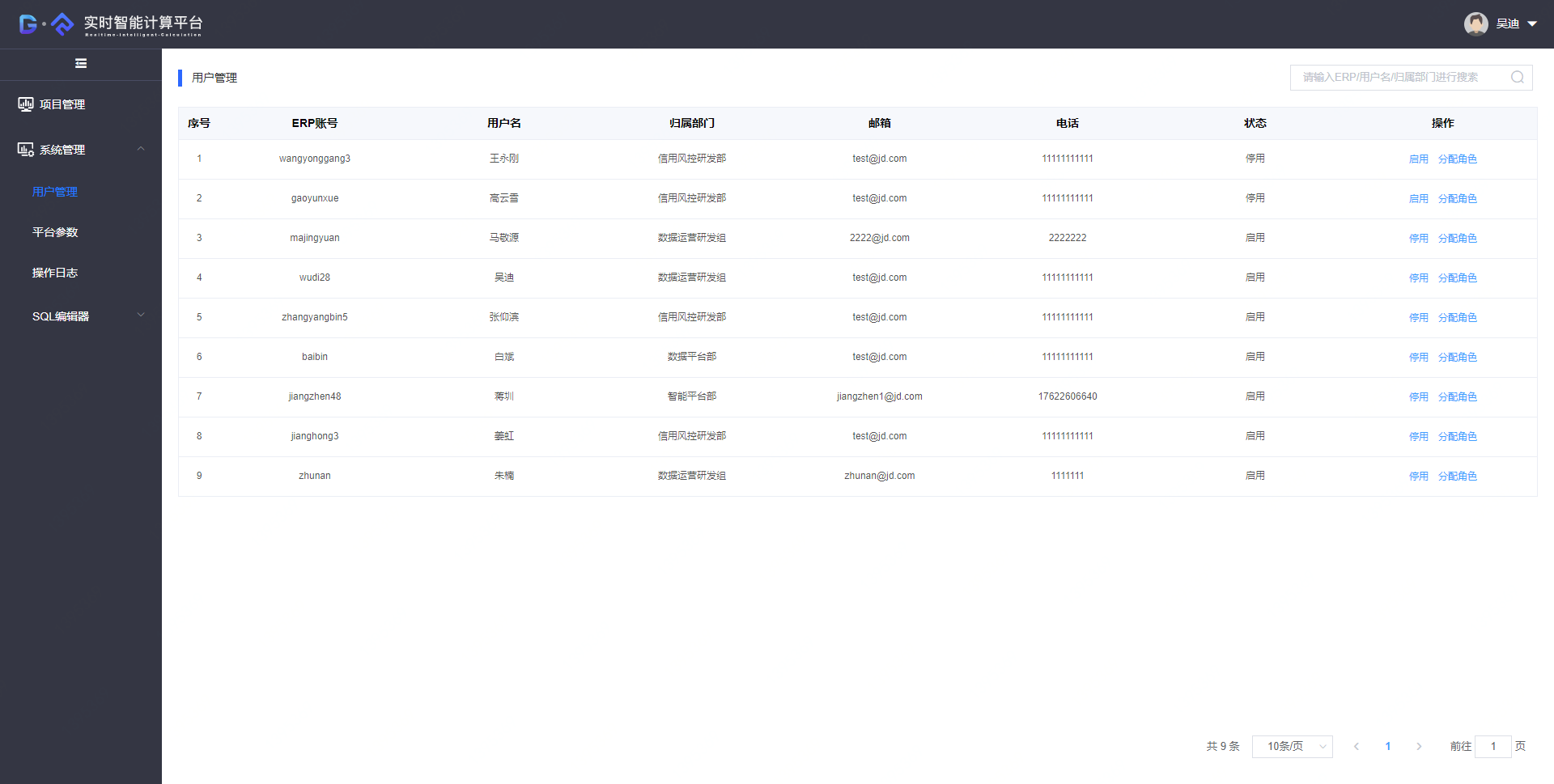
* + - * + MQ列表：



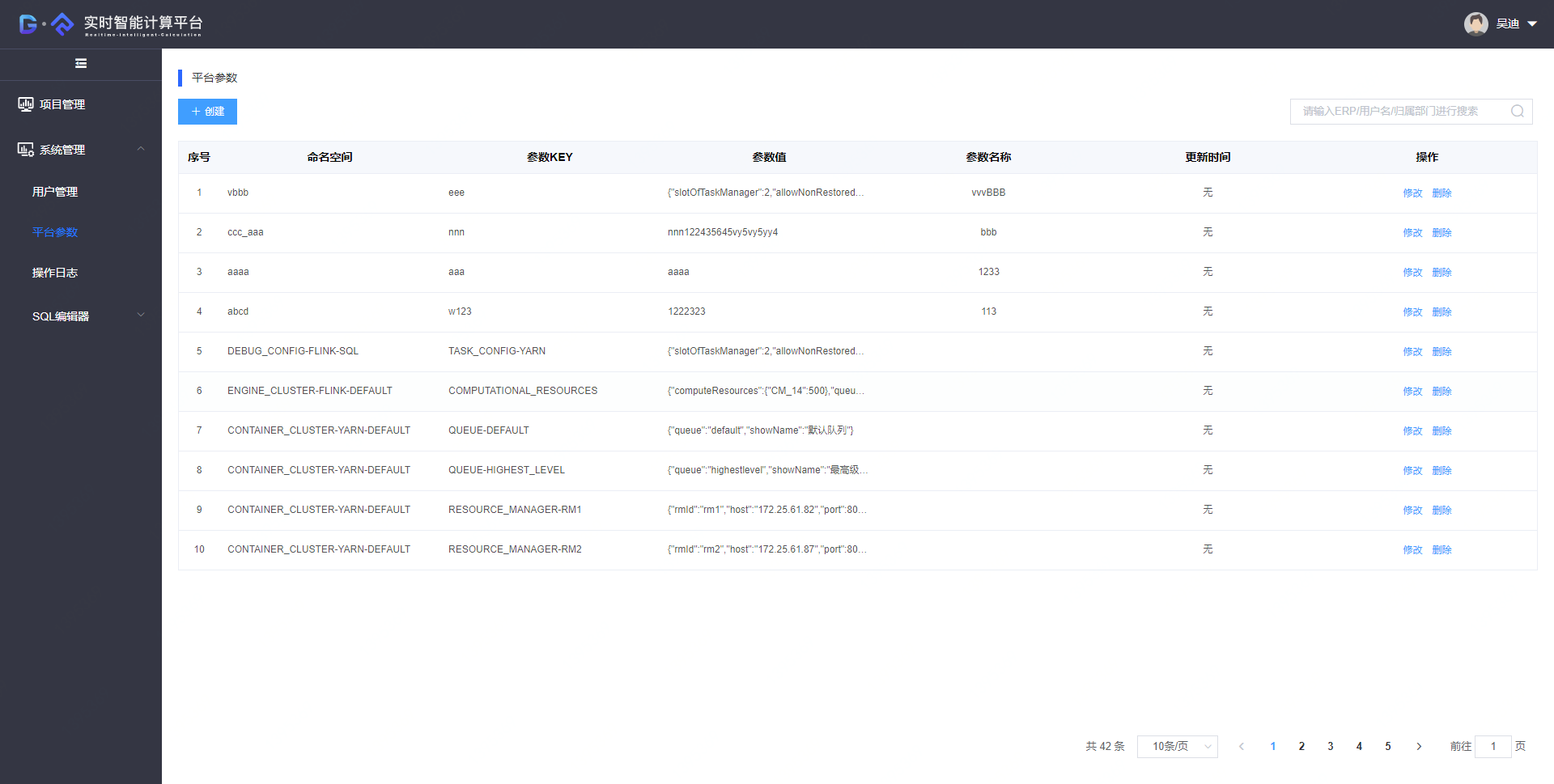
* 1. 系统管理

主要用于平台的基本运维、SQL开发等系统设置，并可查看平台所有人的操作记录

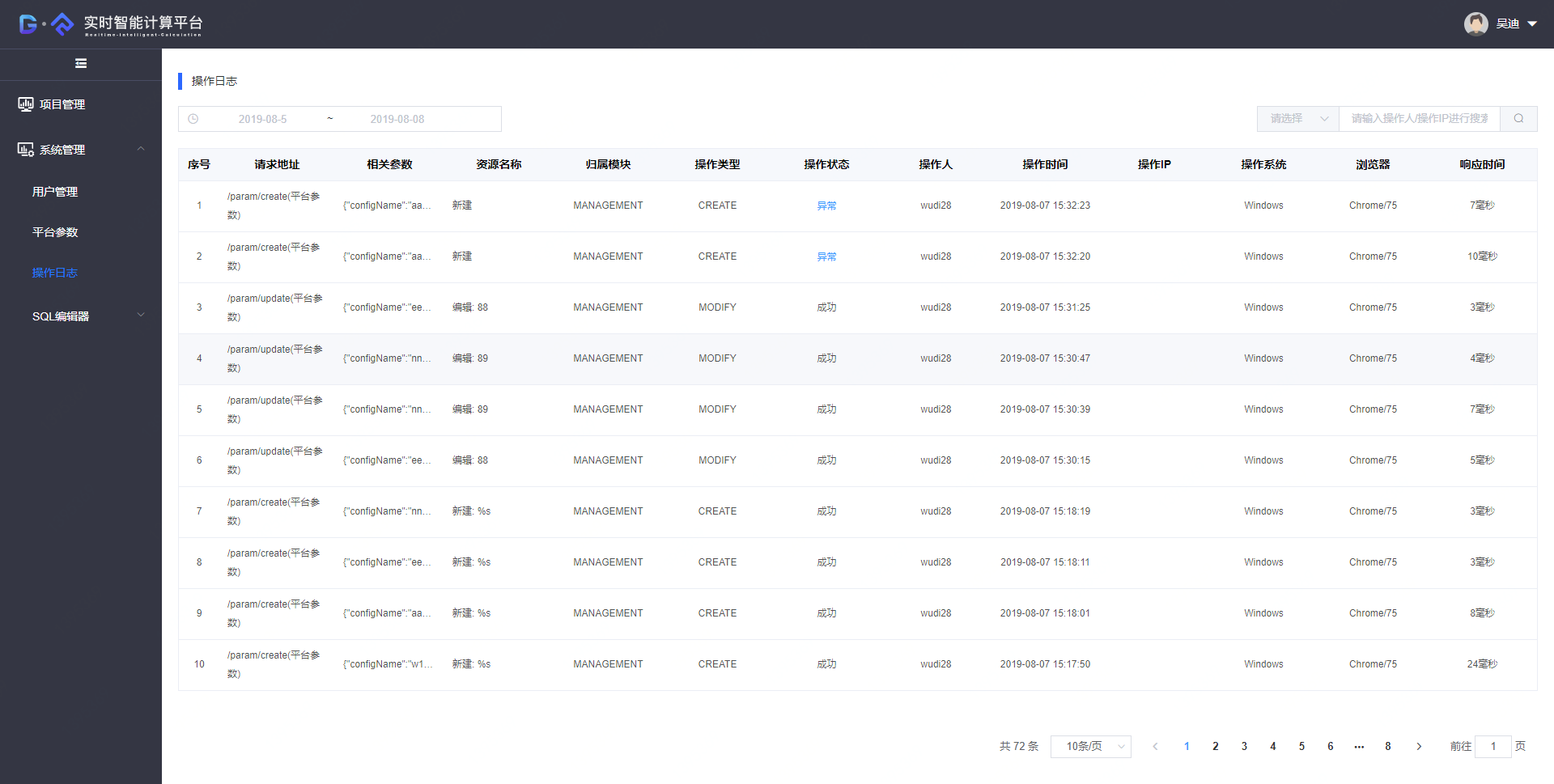
* + 1. 用户管理



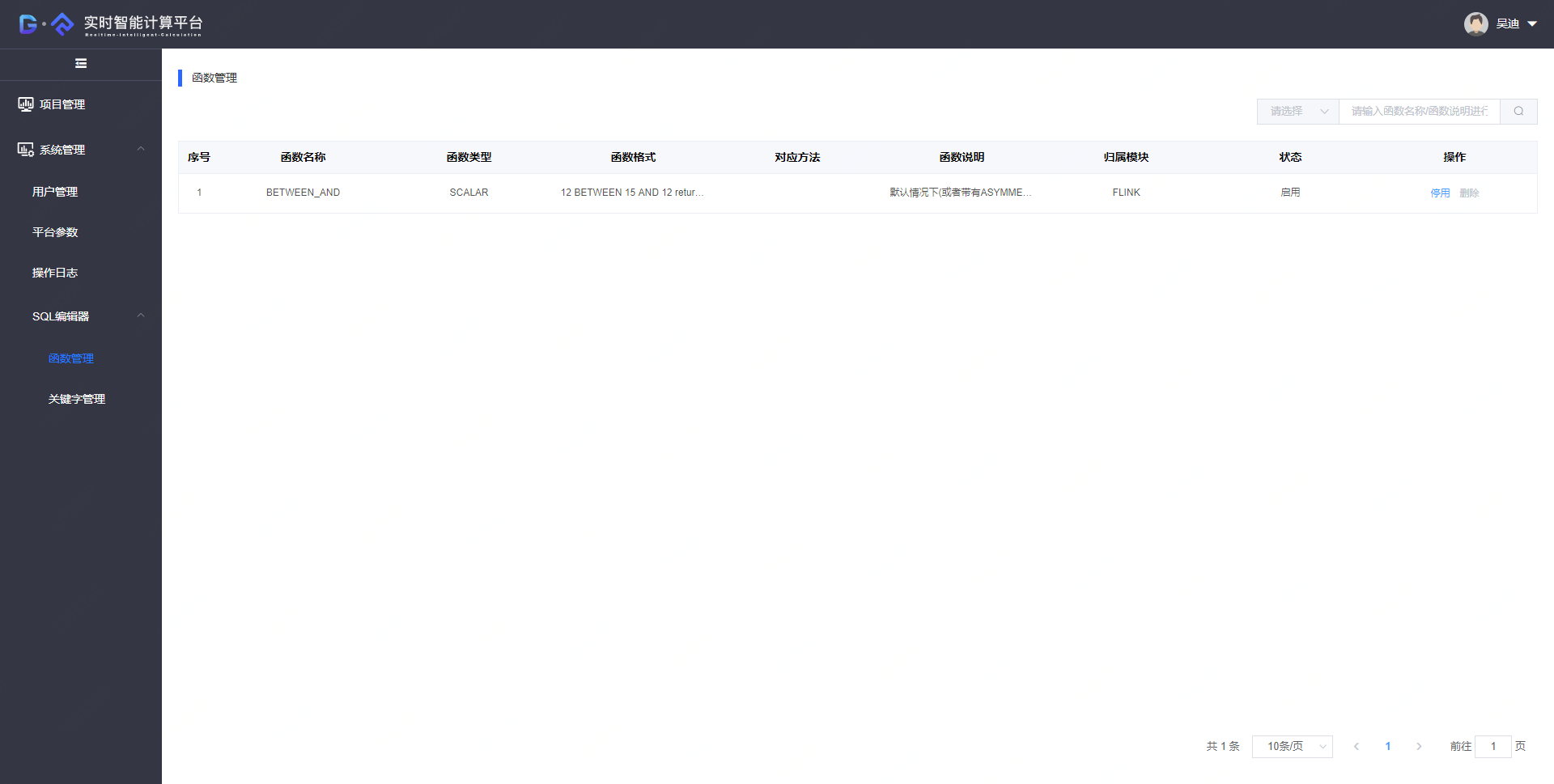
* + 1. 平台参数



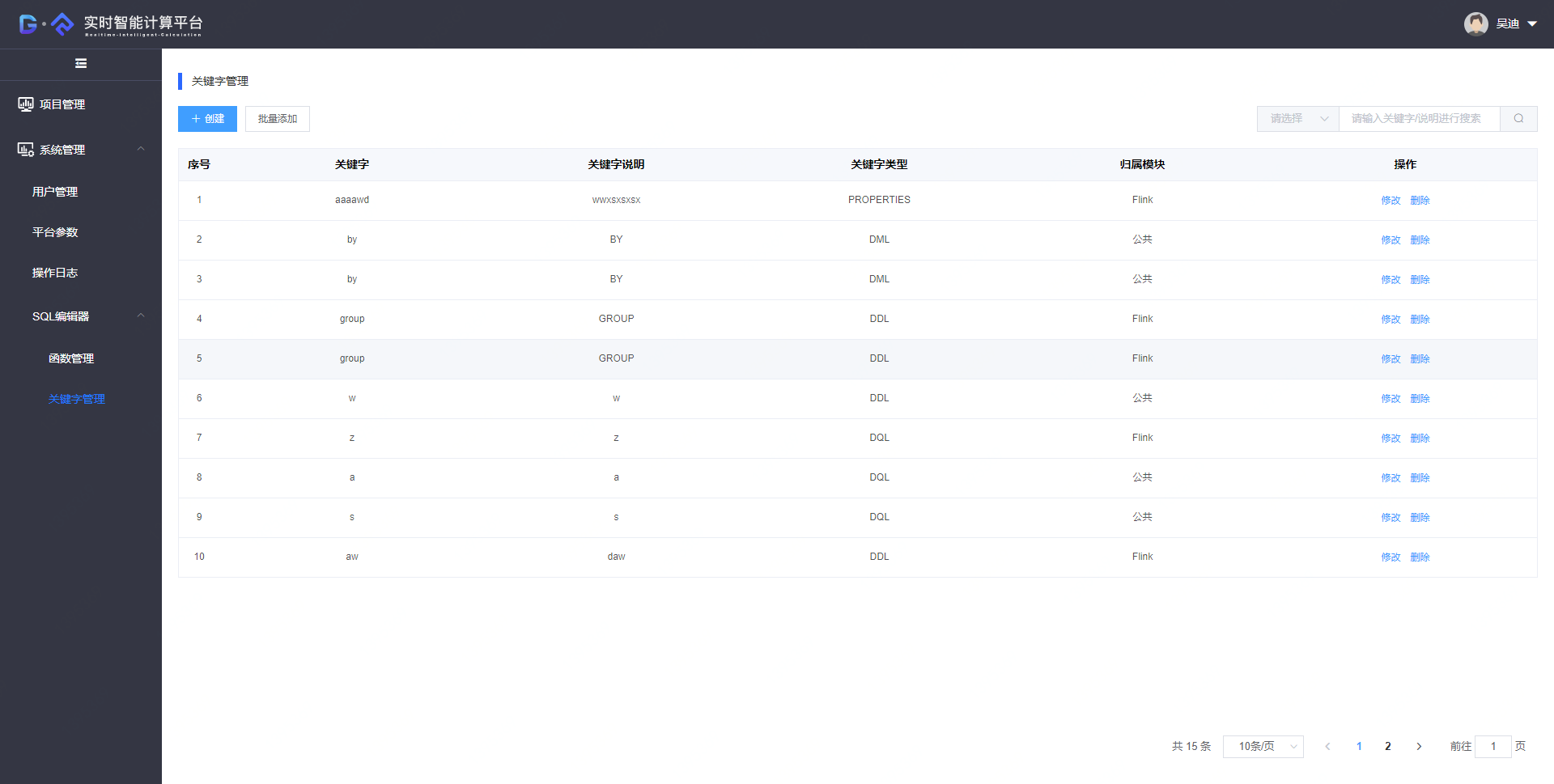
* + 1. 操作日志



* + 1. SQL-IDE配置
       - * 函数管理：

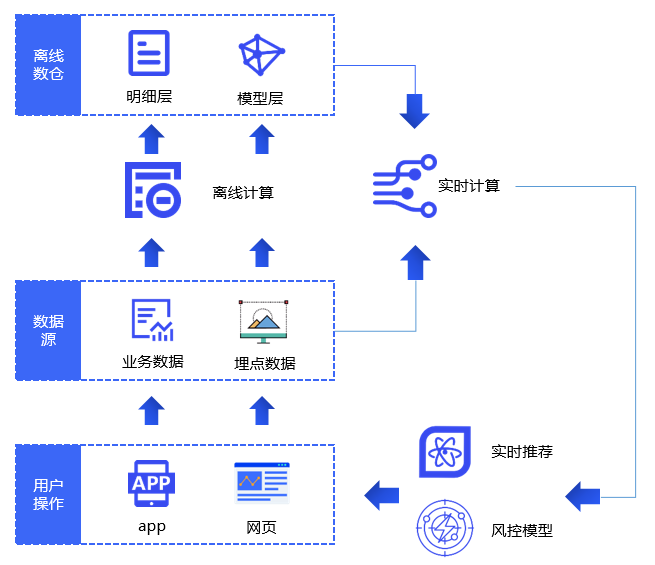


* + - * + 关键字管理：



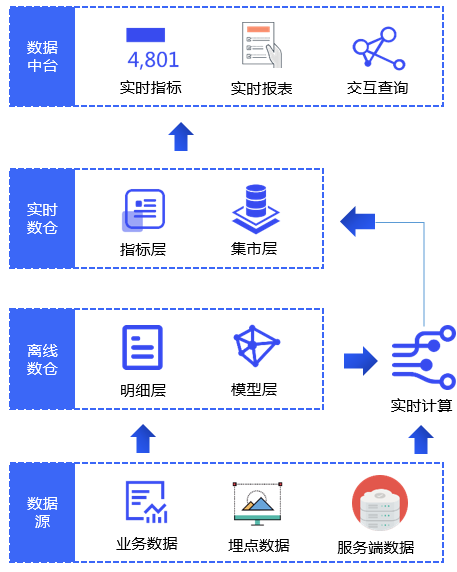
1. 行业客户案例
   1. 【金融行业】银行机构

某银行机构实时计算部署后，实时获取机构业务库数据、埋点数据、服务端数据，将机构T+1风控模型升级为实时模型，实时更新风控策略，提高用户反馈与贷前风险识别，同时支持实时推荐，为客户推荐最佳金融产品，提升产品交易转化40%。该银行在实时推荐与风控场景下业务流程如下图：



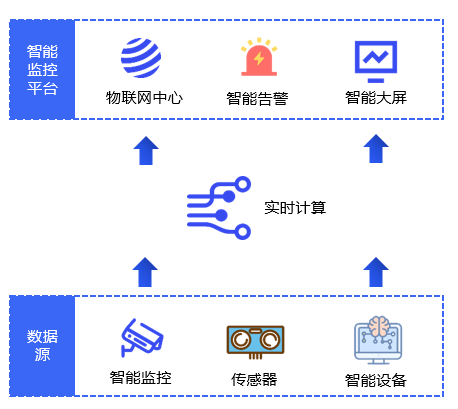
* 1. 【金融行业】某消费金融公司

某消费金融公司实时计算部署后，通过将离线数仓中的汇总数据与实时流计算中的数据融合计算，为指标层、集市层提供实时数据，并展示与应用在数据中台的指标与报表中，让运营活动、核心指标等，均可实时计算展示，与交互查询分析，让业务决策快人一步。该消费金融公司在实时数仓与实时指标场景下的业务流程如下图：



* 1. 【智能城市】智能AI防控项目

某智能AI防控项目实时计算部署后，通过实时解析、清洗Iot相关数据，并加以识别与计算后，将数据实时输出到智能监控平台，因物联网对数据延迟要求极高，采用流式计算，可达到毫秒级的计算延迟，让整个项目达到预想的RPS吞吐要求，让园区内所有监控设备的数据实时分析与展示，覆盖智能生活，智能服务等场景。该智能AI防控项目在智能城市与物联网场景下的业务流程图如下：



1. 法律声明

暂无



电话：18733448927

邮箱：zhaoziyun@jd.com

电话：18733448927

邮箱：zhaoziyun@jd.com

一站式大数据解决方案

深海大数据平台一站式大数据解决方案

深海大数据平台

咚咚扫码加入深海大数据平台

咚咚扫码加入

大数据平台交流群

微信扫码加入咚咚扫码加入

大数据平台交流群

微信扫码加入

深海大数据平台

微信扫码加入

深海大数据平台