

# 鲲鹏在线数据服务-roc

## 产品简介

### 背景

你还在为每提供一份数据指标都需要开发一个接口甚至一个项目工程而发愁吗？

你还在为每改一个接口要走新建、变更、发布、上线的流程而心灰意冷吗？

你还在为提供的数据难以维护，牵一发而动全身的繁琐代码而发狂吗？

随着实时化技术的不断发展，在线实时数据的功能需求日益增多，为了避免每提供一次数据即需写一套接口的尴尬境地，统一数据服务平台-鲲鹏应允而生。

### 核心模块

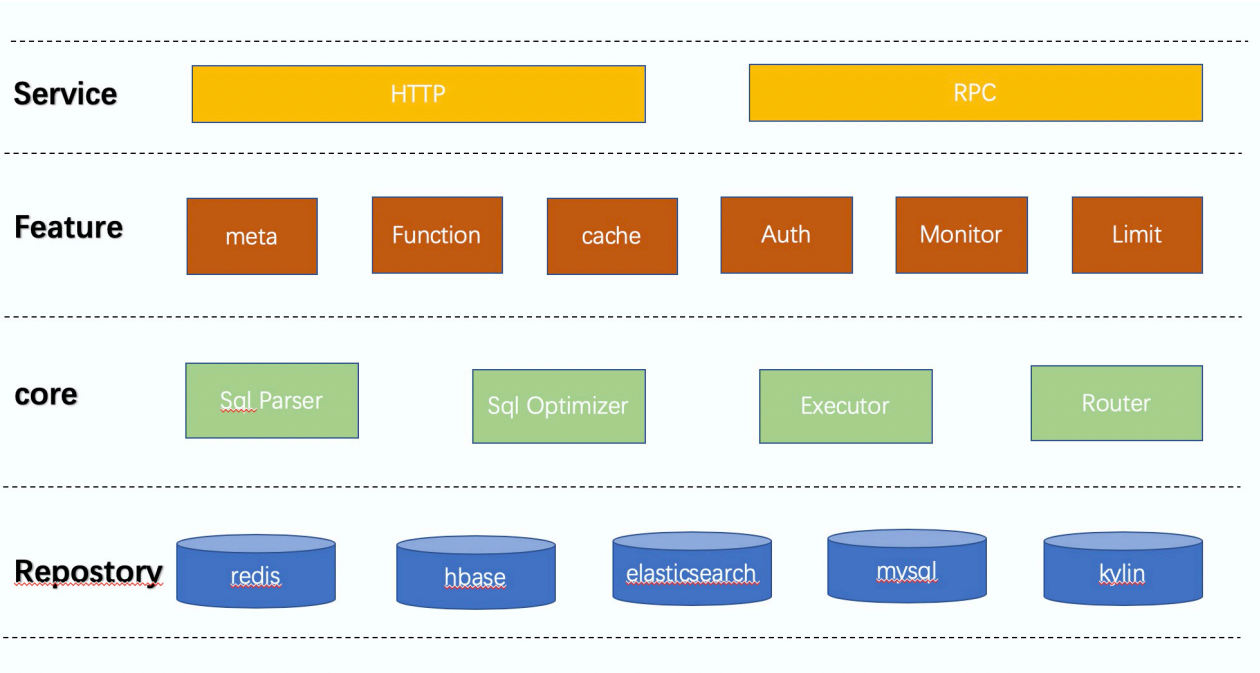
Roc-Engine：核心计算引擎，负责元数据定义，通过sql解析->元数据校验->语义分析->逻辑执行计划->物理执行计划->多种异构源查询获取数据。

Roc-Server：统一数据服务，负责接口化，元数据更新、权限校验、限流、监控、服务封装、初始化。

Roc-Manager：鲲鹏管理平台，负责用户管理，元数据管理，udf管理，监控大盘，权限、限流配置等。

## 架构大图

### 计算引擎



空间

空间单元

域

业务域

表

同步映射

扩展

版本

路由

数据预览

在线查询

函数

UDF

UDAF

UDTF

UDSF

上传

发布

版本

接口

token

参数配置

cache

Sql配置

用户管理

CRUD

权限配置

基础配置

限流

数据源

监控

服务诊断

集群状态

数据大盘

大图

调用趋势

## 概要设计

### 元数据

#### 空间

namespace 空间单元

属性	类型	含义
namespace_id	Long	空间id
name	String	空间名称
uid	Long	用户id
desc	String	描述
is_delete	Int	是否删除
create_time	Datetime	创建时间
update_time	Datetime	修改时间

#### 域

所属数据域，例如订单域、用户域、商品域等。

属性	类型	含义
region_id	Long	数据域Id
namespace_id	Long	所属空间id
name	String	数据域名
uid	Long	用户id
desc	String	描述
is_delete	Int	是否删除
create_time	Datetime	创建时间
update_time	Datetime	修改时间

### 表（逻辑表）

逻辑表，可映射多个物理表。

属性	类型	含义
table_id	Long	逻辑表Id
region_id	Long	数据域Id
namespace_id	Long	所属空间id
name	String	表名
uid	Long	管理员id
desc	String	描述
is_delete	Int	是否删除
create_time	Datetime	创建时间
update_time	Datetime	修改时间

### 表（物理表）

物理实际存储表

属性	类型	含义
table_id	Long	物理表
logic_table_id	Long	逻辑表id
source_id	Long	数据源Id
name	String	表名
uid	Long	管理员id
is_delete	Int	是否删除
create_time	Datetime	创建时间
update_time	Datetime	修改时间

## 路由

逻辑表到物理表路由策略

属性	类型	含义
router_id	Long	数据域Id
logic_table_id	Long	逻辑表id
physics_table_id	Long	物理表Id
type	String	支持time，默default 认逻辑表和物理表为一对一
name	String	路由名称
start	Long	距当前时间开始时间长度
start	Long	距当前时间结束时间长度
unit	String	时间单位支持 秒 分钟 小时 天
uid	Long	管理员id
is_delete	Int	是否删除
create_time	Datetime	创建时间
update_time	Datetime	修改时间

## 属性

字段属性

属性	类型	含义
field_id	Long	属性Id
logic_table_id	Long	逻辑表Id
region_id	Long	数据域Id
namespace_id	Long	空间Id
name	String	属性名称
type	String	类型：Int、Long、String、Date、DateTime、Double等等
is_pk	Int	是否是主键 具有顺序
is_route	Int	是否为路由属性，暂时支持time route
is_index	Int	是否是索引
desc	String	字段描述
uid	Long	管理员id
is_delete	Int	是否删除
create_time	Datetime	创建时间
update_time	Datetime	修改时间

## 数据源

数据源，目前支持redis、hbase、ealsticsearch、kylin、mysql

属性	类型	含义
source_id	Long	数据源id
name	String	数据源名称唯一不可重复
type	String	类型Hbase、Redis、Mysql、Kylin、Elasticsearch
is_private	Int	是否私有
props	String	json 链接配置信息
desc	String	描述
uid	Long	负责人ID
desc	String	描述
is_delete	Int	是否删除
create_time	Datetime	创建时间
update_time	Datetime	修改时间

## 工程模块

### roc-engine

鲲鹏计算引擎，此工程为SDK，封装为jar。

### roc-engine-common

公共工具

### roc-engine-meta

元数据，只负责元数据定义，namespace、region、table、field等

### roc-engine-parser

计算引擎解析器，处理抽象语法树，语义分析，绑定元数据信息，校验等。

### roc-engine-optimizer

计算引擎逻辑计划定义，优化器，谓词下推，路由计划等。

### roc-engine-executor

计算引擎执行器，通过逻辑计划进行sql化到source源中。

### roc-engine-source

计算引擎source源，例如hbase、es、mysql、redis等

## **roc-engine-function**

计算引擎内部函数实现

## **roc-engine-sei**

服务扩展接口，不负责实现，例如元数据更新接口、自定义source插件，udf抽象等。

## **roc-server**

### **roc-server-rest**

http入口

### **roc-server-service**

接口实现

### **roc-server-cache**

缓存层实现，定制缓存策略，lru内存、redis等

### **roc-server-sdk**

封装友好的sdk

### **roc-server-api**

接口定义、rocContext上下文定义

### **roc-server-trace**

监控链路追踪 rt、慢查询、报错查询等进行上报

### **roc-server-common**

公共组件

### **roc-server-fetcher**

元数据更新，function更新

## **roc-manager**

### **roc-manager-rest**

http接口

### **roc-manager-service**

业务实现

## **roc-manager-repostory**

持久化

## **roc-manager-domain**

实体定义

## **roc-manager-common**

公共组件