

开放数据处理服务ODPS

SDK



SDK

概要

使用Maven的用户可以从[Maven库](#)中搜索"odps-sdk"获取不同版本的Java SDK：

包名	描述
odps-sdk-core	ODPS基础功能的主体接口，搜索关键词"odps-sdk-core"
odps-sdk-commons	一些Util封装，搜索关键词"odps-sdk-commons"
odps-sdk-udf	UDF功能的主体接口，搜索关键词"odps-sdk-udf"
odps-sdk-mapred	MapReduce Java SDK，搜索关键词"odps-sdk-mapred"
odps-sdk-graph	Graph Java SDK，搜索关键词"odps-sdk-graph"

AliyunAccount

阿里云认证账号。输入参数为accessId及accessKey，是阿里云用户的身份标识和认证密钥。此类用来初始化Odps。

Odps

ODPS SDK的入口，用户通过此类来获取项目空间下的所有对象集合，包括: Projects，Tables，Resources，Functions，Instances。用户可以通过传入 AliyunAccount 实例来构造ODPS对象。程序示例如下：

```
Account account = new AliyunAccount("my_access_id", "my_access_key");

Odps odps = new Odps(account);
String odpsUrl = "<your odps endpoint>";
odps.setEndpoint(odpsUrl);
odps.setDefaultProject("my_project");

for (Table t : odps.tables()) {
    ....
}
```

备注:

- ODPS提供了两个服务地址供用户选择。详细说明请参考 [Tunnel SDK简介](#)

Projects

ODPS中所有项目空间的集合。集合中的元素为 **Project** 。程序示例如下：

```
Account account = new AliyunAccount("my_access_id", "my_access_key");

Odps odps = new Odps(account);
String odpsUrl = "<your odps endpoint>";
odps.setEndpoint(odpsUrl);
Project p = odps.projects().get("my_exists");
p.reload();
Map<String, String> properties = prj.getProperties();
...
```

备注:

- ODPS提供了两个服务地址供用户选择。详细说明请参考 [Tunnel SDK简介](#)

Project

对项目空间信息的描述，可以通过 **Projects** 获取相应的项目空间。

SQLTask

用于运行、处理SQL任务的接口。可以通过run接口直接运行SQL。run接口返回 **Instance** 实例，通过 **Instance** 获取SQL的运行状态及运行结果。 程序示例如下，仅供参考：

```
Account account = new AliyunAccount("my_access_id", "my_access_key");
Odps odps = new Odps(account);
String odpsUrl = "<your odps endpoint>";
odps.setEndpoint(odpsUrl);

Instance instance = SQLTask.run(odps, "my_project", "select ...");
String id = instance.getId();

instance.waitForSuccess();

Set<String> taskNames = instance.getTaskNames();
for (String name : taskNames) {
    TaskSummary summary = instance.getTaskSummary(name);
    String s = summary.getSummaryText();
}

Map<String, String> results = instance.getTaskResults();
Map<String, TaskStatus> taskStatus = instance.getTaskStatus();

for (Entry<String, TaskStatus> status : taskStatus.entrySet()) {
    String result = results.get(status.getKey());
}
```

备注:

- 如果用户想创建表，需要通过SQLTask接口，而不是 Table 接口。用户需要将 创建表(CREATE TABLE) 的语句传入SQLTask。
- ODPS提供了两个服务地址供用户选择。详细说明请参考 Tunnel SDK简介

Instances

Instances

ODPS中所有实例(Instance)的集合。集合中的元素为 Instance 。程序示例如下：

```
Account account = new AliyunAccount("my_access_id", "my_access_key");

Odps odps = new Odps(account);
String odpsUrl = "<your odps endpoint>";
odps.setEndpoint(odpsUrl);
odps.setDefaultProject("my_project");

for (Instance i : odps.instances()) {
    ....
}
```

备注:

- ODPS提供了两个服务地址供用户选择。详细说明请参考 Tunnel SDK简介

Instance

对实例信息的描述，可以通过 Instances 获取相应的实例。程序示例如下：

```
Account account = new AliyunAccount("my_access_id", "my_access_key");

Odps odps = new Odps(account);
String odpsUrl = "<your odps endpoint>";
odps.setEndpoint(odpsUrl);
Instance ins = odps.instances().get("instance id");
Date startTime = instance.getStartTime();
Date endTime = instance.getEndTime();

...

Status instanceStatus = instance.getStatus();
String instanceStatusStr = null;
if (instanceStatus == Status.TERMINATED) {
    instanceStatusStr = TaskStatus.Status.SUCCESS.toString();
    Map<String, TaskStatus> taskStatus = instance.getTaskStatus();
    for (Entry<String, TaskStatus> status : taskStatus.entrySet()) {
```

```

        if (status.getValue().getStatus() != TaskStatus.Status.SUCCESS) {
            instanceStatusStr = status.getValue().getStatus().toString();
            break;
        }
    }
} else {
    instanceStatusStr = instanceStatus.toString();
}
...

TaskSummary summary = instance.getTaskSummary("instance name");
String s = summary.getSummaryText();

```

备注：

- ODPS提供了两个服务地址供用户选择。详细说明请参考 Tunnel SDK简介。

Tables

ODPS中所有表的集合。集合中的元素为 **Table** 。程序示例如下：

```

Account account = new AliyunAccount("my_access_id", "my_access_key");

Odps odps = new Odps(account);
String odpsUrl = "<your odps endpoint>";
odps.setEndpoint(odpsUrl);
odps.setDefaultProject("my_project");

for (Table t : odps.tables()) {
    ....
}

```

备注:

- ODPS提供了两个服务地址供用户选择。详细说明请参考 Tunnel SDK简介

Table

对表信息的描述，可以通过 **Tables** 获取相应的表。程序示例如下：

```

Account account = new AliyunAccount("my_access_id", "my_access_key");

Odps odps = new Odps(account);
String odpsUrl = "<your odps endpoint>";
odps.setEndpoint(odpsUrl);
Table t = odps.tables().get("table name");
t.reload();
Partition part = t.getPartition(new PartitionSpec(tableSpec[1]));
part.reload();

```

...

备注:

- ODPS提供了两个服务地址供用户选择。详细说明请参考 Tunnel SDK简介

Resources

ODPS中所有资源的集合。集合中的元素为 `Resource` 。程序示例如下：

```
Account account = new AliyunAccount("my_access_id", "my_access_key");

Odps odps = new Odps(account);
String odpsUrl = "<your odps endpoint>";
odps.setEndpoint(odpsUrl);
odps.setDefaultProject("my_project");

for (Resource r : odps.resources()) {
    ....
}
```

备注:

- ODPS提供了两个服务地址供用户选择。详细说明请参考 Tunnel SDK简介

Resource

对资源信息的描述，可以通过 `Resources` 获取相应的资源。程序示例如下：

```
Account account = new AliyunAccount("my_access_id", "my_access_key");

Odps odps = new Odps(account);
String odpsUrl = "<your odps endpoint>";
odps.setEndpoint(odpsUrl);
Resource r = odps.resources().get("resource name");
r.reload();

if (r.getType() == Resource.Type.TABLE) {
    TableResource tr = new TableResource(r);
    String tableSource = tr.getSourceTable().getProject() + "."
        + tr.getSourceTable().getName();
    if (tr.getSourceTablePartition() != null) {
        tableSource += " partition(" + tr.getSourceTablePartition().toString()
            + ")";
    }
    ....
}
```

备注:

- ODPS提供了两个服务地址供用户选择。详细说明请参考 Tunnel SDK简介

一个创建文件资源的示例：

```
String projectName = "my_porject";

String source = "my_local_file.txt";
File file = new File(source);
InputStream is = new FileInputStream(file);

FileResource resource = new FileResource();
String name = file.getName();
resource.setName(name);

odps.resources().create(projectName, resource, is);
```

一个创建表资源的示例：

```
TableResource resource = new TableResource(tableName, tablePrj, partitionSpec);
//resource.setName(INVALID_USER_TABLE);
resource.setName("table_resource_name");
odps.resources().update(projectName, resource);
```

Functions

ODPS中所有函数的集合。集合中的元素为 Function 。程序示例如下：

```
Account account = new AliyunAccount("my_access_id", "my_access_key");

Odps odps = new Odps(account);
String odpsUrl = "<your odps endpoint>";
odps.setEndpoint(odpsUrl);
odps.setDefaultProject("my_project");

for (Function f : odps.functions()) {
    ....
}
```

备注:

- ODPS提供了两个服务地址供用户选择。详细说明请参考 Tunnel SDK简介

Function

对函数信息的描述，可以通过 Functions 获取相应的函数。程序示例如下：

```
Account account = new AliyunAccount("my_access_id", "my_access_key");
```

```
Odps odps = new Odps(account);
String odpsUrl = "<your odps endpoint>";
odps.setEndpoint(odpsUrl);

Function f = odps.functions().get("function name");
List<Resource> resources = f.getResources();
```

备注:

- ODPS提供了两个服务地址供用户选择。详细说明请参考 [Tunnel SDK简介](#)

一个创建函数的示例：

```
String resources = "xxx:xxx";
String classType = "com.aliyun.odps.mapred.open.example.WordCount";

ArrayList<String> resourceList = new ArrayList<String>();
for (String r : resources.split(":")) {
    resourceList.add(r);
}

Function func = new Function();
func.setName(name);
func.setClassType(classType);
func.setResources(resourceList);

odps.functions().create(projectName, func);
```