

DataHub

快速入门

快速入门

快速指引

快速指引

本节主要描述使用DataHub Java SDK进行数据的读写。

准备工作

- 使用DataHub服务之前,需要注册阿里云云账号,利用阿里云账号的AccessId与AccessKey接入DataHub服务。

创建Project/Topic

- 登录DataHub Web Console页面，创建Topic。

发布数据/订阅数据

发布/订阅数据需要使用sdk完成，以下是Java SDK 2.12基础示例 DataHub Java SDK使用说明

```
String endpoint = "http://dh-cn-hangzhou.aliyuncs.com";
String accessId = "<YourAccessKeyId>";
String accessKey = "<YourAccessKeySecret>";

// 创建DataHubClient实例
DatahubClient datahubClient = DatahubClientBuilder.newBuilder()
    .setDatahubConfig(
        new DatahubConfig(endpoint,
            // 是否开启二进制传输，服务端2.12版本开始支持
            new AliyunAccount(accessId, accessKey), true))
    // 专有云使用出错尝试将参数设置为 false
    // HttpConfig可不设置，不设置时采用默认值
    .setHttpConfig(new HttpConfig().setConnTimeout(10000))
    .build();
```

```
String projectName = "<YourProjectName>";
String topicName = "<YourTopicName>";
// 可通过listShard接口获取shard列表，所有ACTIVE的shard均可使用，本例使用"0"
String shardId = "0";
// TUPLE类型的Topic需要设置schema，也可直接通过getTopic获取
RecordSchema schema = datahubClient.getTopic("test_project", "test_topic").getRecordSchema();
List<RecordEntry> recordEntries = new ArrayList<>();
RecordEntry entry = new RecordEntry(schema);

// 构造10条records
List<RecordEntry> recordEntries = new ArrayList<>();
for (int i = 0; i < 10; ++i) {
    RecordEntry recordEntry = new RecordEntry();
    // set attributes
    recordEntry.addAttribute("key1", "value1");
    // set tuple data
    TupleRecordData data = new TupleRecordData(schema);
    data.setField("field1", "HelloWorld");
    data.setField("field2", 1234567);
    recordEntry.setRecordData(data);

    recordEntries.add(recordEntry);
}

// 服务端从2.12版本开始支持，之前版本请使用putRecords接口
datahubClient.putRecordsByShard(projectName, topicName, shardId, recordEntries);

// 获取cursor, 这里根据时间获取游标
// 注: 正常情况下，getCursor只需在初始化时获取一次，然后使用getRecords的nextCursor进行下一次读取
String cursor = datahubClient.getCursor(projectName, topicName, shardId, CursorType.SYSTEM_TIME,
1455869335000).getCursor();

// 数据读取
// 每次限读100条，最大不可超过1000
int limit = 100;
while (true) {
    try {
        GetRecordsResult result = datahubClient.getRecords(projectName, topicName, shardId, schema, cursor, limit);
        // 如果有数据则处理，无数据需sleep后重新读取
        if (result.getRecordCount() > 0) {
            for (RecordEntry entry : result.getRecords()) {
                TupleRecordData data = (TupleRecordData) entry.getRecordData();
                System.out.println("field1:" + data.getField("field1"));
                System.out.println("field2:" + data.getField("field2"));
            }
        }
        // 拿到下一个游标
        cursor = result.getNextCursor();
    } catch (InvalidCursorException ex) {
        // 非法游标或游标已过期，建议重新定位后开始消费
        cursor = datahubClient.getCursor(projectName, topicName, shardId, CursorType.OLDEST).getCursor();
    }
}
```