墨天\$ 首页 资讯 活动 大会 学习 ✓ 文档 问答 服务 🚥 ✓ 💵 排行 ✓

deepseek C

○ 登录

注册

首页 / Debezium实战!一款不错的开源CDC工具



Debezium实战!一款不错的开源CDC工具



原创 🛭 范伟林 🕒 2025-03-18

• 448



一、前言



debezium属于kafka的一个connect插件,红帽的一个开源项目,有2.7和3.0两个稳定版本,可以根据kafka的版本自定义选择,kafka存在非常多connect插件,其中同类的商业版本为confluent。

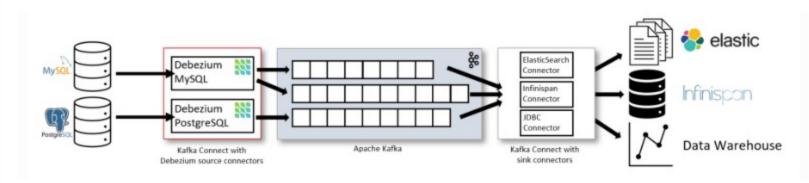
debezium捕获功能基本支持大部分主流的数据库MySQL、MariaDB、MongoDB、PostgreSQL、Oracle、SQL Server、Db2、Cassandra、Vitess、Spanner、Informix,入库方式本身有jdbc connector组件去消费kafka的topic,当然开发者也可以自行写一些程序去消费。

主要管理方式以kafka为主,connector使用REST的方式进行管理,debezium本身也有图形化管理工具debezium UI,也可以使用一些kafka的管理工

具,开源的也非常多。

基本原理为捕获的数据写入到kafka,之后 从kafka进行消费,本次实战以Oracle到Oracle同步测试来演示。

架构



二、kafka配置

1、配置JDK

mkdir /app/jdk

tar -zxvf jdk-11.0.25_linux-x64_bin.tar.gz

export JAVA_HOME=/app/jdk/jdk-11.0.25

export CLASSPATH=\$JAVA_HOME/lib/

export ZK_HOME=/app/zookeeper/apache-zookeeper-3.8.4-bin

 ${\tt export\ PATH=\$ZK_HOME/bin:\$JAVA_HOME/bin:\$PATH}$

2、配置zookeeper







最新文章

创建目录 mkdir -p /app/zookeeper 解压 tar -xvf apache-zookeeper-3.8.4-bin.tar.gz 参数配置 cd /app/zookeeper/apache-zookeeper-3.8.4-bin mkdir data cd /app/zookeeper/apache-zookeeper-3.8.4-bin/conf cp zoo_sample.cfg zoo.cfg vi zoo.cfg dataDir=/app/zookeeper/apache-zookeeper-3.8.4-bin/data 启动服务 cd /app/zookeeper/apache-zookeeper-3.8.4-bin/bin/ ./zkServer.sh start ./zkServer.sh status

查看

```
[fwl@kafka ~]$ zkServer.sh status
ZooKeeper JMX enabled by default
Using config: /app/zookeeper/apache-zookeeper-3.8.4-bin/bin/../conf/zoo.cfg
Client port found: 2181. Client address: localhost. Client SSL: false.
Mode: standalone
[fwl@kafka ~]$ jps
7458 Jps
7372 QuorumPeerMain
[fwl@kafka ~]$
```

Note:单机部署,使用默认端口为2181,自定义个数据目录,java进程名:QuorumPeerMain

3、配置kafka

```
创建目录
mkdir /app/kafka
解压
tar -zxvf kafka_2.13-2.8.2.tgz
配置参数
vi server.properties
log.dirs=/app/kafka/kafka_2.13-2.8.2/data/kafka-logs
num.partitions=1
default.replication.factor=1
服务启动
cd /app/kafka/kafka_2.13-2.8.2
bin/kafka-server-start.sh -daemon config/server.properties
```

查看

实用性技巧,19c快速克隆DBHOME目录 986浏览

2023-11-23

Oracle-ACFS模块安装

2023-09-23 748浏览

小技巧,审计Oracle登录失败的会话

2023-08-30

1417浏览

SAP HANA DB 2.0 sp3单机安装+SUSE 1

2.3

2023-08-24 1883浏览

目录

- 一、前言
- 二、kafka配置
- 三、debezium配置
- 四、抽取-source connector配置
- 五、入库-sink connector配置

```
[fwl@kafka kafka_2.13-2.8.2]$ jps
8065 Kafka
8084 Jps
7372 QuorumPeerMain
[fwl@kafka kafka_2.13-2.8.2]$ [
```

Note:单机版部署,使用默认端口9092

三、debezium配置

```
新建插件目录
cd /app/kafka/kafka_2.13-2.8.2
mkdir plugins
解压到插件目录
cd /app/kafka/kafka_2.13-2.8.2/plugins
tar -zxvf debezium-connector-oracle-2.7.3.Final-plugin.tar.gz
tar -zxvf debezium-connector-jdbc-2.7.3.Final-plugin.tar.gz
查看
[fwl@kafka plugins]$ ll
total 65952
drwxrwxr-x 2 fwl fwl
                          4096 Oct 16 23:29 debezium-connector-jdbc
-rw-r--r-- 1 root root 32787029 Oct 16 05:07 debezium-connector-jdbc-2.7.3.Final-plugin.tar
                          4096 Oct 16 09:03 debezium-connector-oracle
drwxrwxr-x 2 fwl fwl
-rw-r--r-- 1 root root 34731726 Oct 16 05:07 debezium-connector-oracle-2.7.3.Final-plugin.t
[fwl@kafka plugins]$
修改connect配置文件
cd /app/kafka/kafka_2.13-2.8.2/config
vi connect-distributed.properties
plugin.path=/app/kafka/kafka_2.13-2.8.2/plugins
启动
cd /app/kafka/kafka_2.13-2.8.2
bin/connect-distributed.sh -daemon config/connect-distributed.properties
```

Note: connect有2种方式standalone和distributed,默认端口8083,可以配置启动多个connect

查看进程

[fwl@kafka kafka_2.13-2.8.2]\$ jps 8065 Kafka 8786 ConnectDistributed 8905 Jps 7372 QuorumPeerMain [fwl@kafka kafka_2.13-2.8.2]\$

查看插件

```
[fwl@kafka kafka_2.13-2.8.2]$ curl -s localhost:8083/connector-plugins|jq
    "class": "io.debezium.connector.jdbc.JdbcSinkConnector",
    "type": "sink",
    "version": "2.7.3.Final"
   "class": "io.debezium.connector.oracle.OracleConnector",
   "type": "source",
    "version": "2.7.3.Final"
    "class": "org.apache.kafka.connect.file.FileStreamSinkConnector",
    "type": "sink",
    "version": "2.8.2"
 },
    "class": "org.apache.kafka.connect.file.FileStreamSourceConnector",
    "type": "source",
    "version": "2.8.2"
 },
 {
   "class": "org.apache.kafka.connect.mirror.MirrorCheckpointConnector",
    "type": "source",
    "version": "1"
 },
    "class": "org.apache.kafka.connect.mirror.MirrorHeartbeatConnector",
    "type": "source",
    "version": "1"
 },
    "class": "org.apache.kafka.connect.mirror.MirrorSourceConnector",
    "type": "source",
    "version": "1"
 }
[fwl@kafka kafka_2.13-2.8.2]$
```

四、抽取-source connector配置

捕获Oracle支持3种方式,LogMiner、XStream API和OpenLogReplicator,以下使用logminer的方式。

```
配置数据库
alter database archivelog;
alter database force logging;
alter database add supplemental log data;
创建的CDC专用用户
create tablespace dbzumtbs datafile '/oracle/app/oracle/oradata/fwldb/dbzumtbs01.dbf' size
create user dbzum identified by "dbzum" default tablespace dbzumtbs quota unlimited on dbzu
grant dba to dbzum;
开启附加日志
ALTER TABLE scott.nokey05 ADD SUPPLEMENTAL LOG DATA (PRIMARY KEY, UNIQUE) COLUMNS;
抽取connector配置
curl -X POST http://localhost:8083/connectors -H "Content-Type: application/json" -d '{
"name": "scott_source01",
"config": {
"connector.class" : "io.debezium.connector.oracle.OracleConnector",
"tasks.max" : "1",
"topic.prefix": "fwldb",
"database.hostname" : "192.168.137.12",
"database.port": "1521",
"database.user" : "dbzum",
"database.password" : "dbzum",
"database.dbname" : "fwldb",
"snapshot.mode": "no_data",
"table.include.list" : "scott.fwl_t1",
"log.mining.strategy":"online_catalog",
"schema.history.internal.kafka.bootstrap.servers" : "192.168.137.11:9092",
"schema.history.internal.kafka.topic": "fwldb"
}
}'
测试
insert into FWL_T1 values(1, 'aa');
update FWL_T1 set name='bb' where id = 1;
```

日志查看

```
[2025-03-18 09:50:50,245] INFO [Producer clientId=fwldb-schemahistory] Resetting the last seen epoch of partition fwtob-8 to 8 since the associated topicio changes in the control of the last seen epoch of partition fwtob-8 to 8 since the associated topicio changes in the last seen epoch of partition fwtob-8 to 8 since the associated topicio changes in the last seen epoch of partition fwtob-8 to 8 since the associated topicio changes in the last seen epoch of partition fwtob-8 to 8 since the associated topicio changes in the last seen epoch of partition fwtob-8 to 8 since the associated topicio changes in the last seen epoch of partition fwtob-8 to 8 since the associated topicio changes in the last seen epoch of partition fwtob-8 to 8 since the associated topicio changes in the last seen epoch of partition fwtob-8 to 8 since the associated topicio changes in the last seen epoch of partition fwtob-8 to 8 since the associated topicio changes in the last seen epoch of partition fwtob-8 to 8 since the associated topicio changes in the last seen epoch of partition fwtob-8 to 8 since the associated topicio changes in the last seen epoch of partition fwtob-8 to 8 since the associated topicio changes in the last seen epoch of partition fwtob-8 to 8 since the associated topicio changes in the last seen epoch of partition fwtob-8 to 8 since the associated topicio changes in the last seen epoch of partition fwtob-8 since topicio changes in the last seen epoch of partition fwtob-9 to 9 since the associated topicio changed from null to 8dJH5ahmTiSAAu_ZpcERAA (or 2025-03-18 99:50:50,459) INFO (changes clientid-connector-producer-scott_source)-8 seetting the last seen epoch of partition fwldb-0 to 9 since the associated topicio changed from null to 8dJH5ahmTiSAAu_ZpcERAA (or 2025-03-18 99:50:50,459) INFO (producer clientid-connector-producer-scott_source)-9 seetting the last seen epoch of partition fwldb-0 to 9 since the associated topicio changes from null to 8dJH5ahmTiSAAu_ZpcERAA (or 2025-03-18 99:50:50,459) INFO (producer 
  alert日志查看
  LOGMINER: Begin mining logfile for session -2147483391 thread 1 sequence 484, /oracle/app/oracle/oradata/fwldb/redo01.log
  LOGMINER: summary for session# = 2147483905
LOGMINER: StartScn: 3728767 (0x0000.0038e57f)
  LOGMINER: EndScn: 3728774 (0x0000.0038e586)
  LOGMINER: HighConsumedScn: 0
  LOGMINER: session_flag: 0x0
  LOGMINER: Read buffers: 8
  LOGMINER: Memory LWM: limit 10M, LWM 8M, 80%
  LOGMINER: Memory Release Limit: 0M
  LOGMINER: Begin mining logfile for session -2147483391 thread 1 sequence 484, /oracle/app/oracle/oradata/fwldb/redo01.log
  LOGMINER: summary for session# = 2147483905
  LOGMINER: StartScn: 3728773 (0x0000.0038e585)
  LOGMINER: EndScn: 3728780 (0x0000.0038e58c)
  LOGMINER: HighConsumedScn: 0
  LOGMINER: session_flag: 0x0
  LOGMINER: Read buffers: 8
  LOGMINER: Memory LWM: limit 10M, LWM 8M, 80%
  LOGMINER: Memory Release Limit: 0M
LOGMINER: Begin mining logfile for session -2147483391 thread 1 sequence 484, /oracle/app/oracle/oradata/fwldb/redo01.log
  LOGMINER: summary for session# = 2147483905
  LOGMINER: StartScn: 3728779 (0x0000.0038e58b)
  LOGMINER: EndScn: 3728786 (0x0000.0038e592)
  LOGMINER: HighConsumedScn: 0
  LOGMINER: session_flag: \theta x \theta
  LOGMINER: Read buffers: 8
```

连接器查看

LOGMINER: Memory LWM: limit 10M, LWM 8M, 80%

LOGMINER: summary for session# = 2147483905 LOGMINER: StartScn: 3728785 (0x0000.0038e591) LOGMINER: EndScn: 3728792 (0x0000.0038e598)

LOGMINER: Memory LWM: limit 10M, LWM 8M, 80%

LOGMINER: summary for session# = 2147483905 LOGMINER: StartScn: 3728791 (0x0000.0038e597) LOGMINER: EndScn: 3728798 (0x0000.0038e59e)

LOGMINER: Memory LWM: limit 10M, LWM 8M, 80%

LOGMINER: Memory Release Limit: 0M

LOGMINER: Memory Release Limit: 0M

LOGMINER: Memory Release Limit: 0M

Tue Mar 18 17:52:44 2025

LOGMINER: HighConsumedScn: 0 LOGMINER: session_flag: 0x0 LOGMINER: Read buffers: 8

LOGMINER: HighConsumedScn: 0 LOGMINER: session_flag: 0x0 LOGMINER: Read buffers: 8

```
[fwl@kafka kafka_2.13-2.8.2]$ curl -s localhost:8083/connectors/scott_source01/status|jq
{
    "name": "scott_source01",
    "connector": {
        "state": "RUNNING",
        "worker_id": "192.168.137.11:8083"
},
    "tasks": [
        {
            "id": 0,
            "state": "RUNNING",
            "worker_id": "192.168.137.11:8083"
        }
        }
        ,
        "type": "source"
}
```

LOGMINER: Begin mining logfile for session -2147483391 thread 1 sequence 484, /oracle/app/oracle/oradata/fwldb/redo01.log

LOGMINER: Begin mining logfile for session -2147483391 thread 1 sequence 484, /oracle/app/oracle/oradata/fwldb/redo01.log

LOGMINER: Begin mining logfile for session -2147483391 thread 1 sequence 484, /oracle/app/oracle/oradata/fwldb/redo01.log

Note:支持Oracle大部分版本,CDB和non-CDB,数据初始化方式可以使用数据泵,debezium本身也支持数据初始化,支持DDL同步

五、入库-sink connector配置

"before": null,

```
连接器配置
 curl -X POST http://localhost:8083/connectors -H "Content-Type: application/json" -d '{
 "name": "scott_sink01",
 "config": {
 "connector.class": "io.debezium.connector.jdbc.JdbcSinkConnector",
 "tasks.max": "1",
 "connection.url":"jdbc:oracle:thin:@192.168.137.12:1521/testdb",
 "connection.username": "fwl",
 "connection.password": "fwl",
 "table.name.format": "fwl.fwl_t1",
 "insert.mode": "upsert",
 "delete.enabled":"true",
 "primary.key.mode": "record_key",
 "schema.evolution": "basic",
 "topics": "fwldb.SCOTT.FWL_T1"
 }
 }'
连接器状态查看
[fwl@kafka kafka_2.13-2.8.2]$ curl -s localhost:8083/connectors/scott_sink01/status|jq
  "name": "scott_sink01",
  "connector": {
    "state": "RUNNING",
    "worker_id": "192.168.137.11:8083"
  },
  "tasks": [
     "id": 0,
     "state": "RUNNING",
     "worker_id": "192.168.137.11:8083"
  "type": "sink"
[fwl@kafka kafka_2.13-2.8.2]$ [
Note:可以利用正则匹配多个topic入库
六、数据检查
检查插入和更新
SQL> select * from scott.fwl_t1;
           ID NAME
            1 bb
SQL>
 SQL> select * from fwl.fwl_t1;
 ID
 1
 bb
 SQL>
topic信息查看
     "version": 2
   "payload": {
```

```
"after": {
      "ID": "1",
      "NAME": "aa"
    },
    "source": {
      "version": "2.7.3.Final",
      "connector": "oracle",
      "name": "fwldb",
      "ts ms": 1742334975000,
      "snapshot": "false",
      "db": "FWLDB",
      "sequence": null,
      "ts us": 1742334975000000,
      "ts ns": 17423349750000000000,
      "schema": "SCOTT",
      "table": "FWL_T1",
      "txId": "060002005a0e0000",
      "scn": "3729260",
      "commit_scn": "3729267",
      "lcr position": null,
      "rs id": "0x0001e4.000012c6.01b0",
      "ssn": 0,
      "redo_thread": 1,
      "user_name": "SCOTT",
      "redo sql": null,
      "row id": "AAAaxEAAEAAAAFrAAA"
    },
    "transaction": null,
    "op": "c",
    "ts ms": 1742306181827,
    "ts us": 1742306181827592,
    "ts_ns": 1742306181827592000
  }
}
  },
  "payload": {
    "before": {
      "ID": "1",
      "NAME": "aa"
    },
    "after": {
      "ID": "1",
      "NAME": "bb"
    },
    "source": {
       "version": "2.7.3.Final",
       "connector": "oracle",
       "name": "fwldb",
      "ts ms": 1742335057000,
      "snapshot": "false",
      "db": "FWLDB",
       "sequence": null,
      "ts_us": 1742335057000000,
       "ts ns": 17423350570000000000,
       "schema": "SCOTT",
      "table": "FWL_T1",
      "txId": "080000008f0d0000",
       "scn": "3729430",
      "commit_scn": "3729437",
      "lcr_position": null,
      "rs id": "0x0001e4.00001301.0010",
      "ssn": 0,
      "redo_thread": 1,
      "user_name": "SCOTT",
      "redo_sql": null,
      "row id": "AAAaxEAAEAAAAFrAAA"
    },
    "transaction": null,
     "op": "u",
     "ts ms": 1742306263662,
     "ts us": 1742306263662434,
    "ts_ns": 1742306263662434000
```

```
Processed a total of 1 messages
[ful@bafba bafba 2 12 2 2 21¢ |
最后检查删除
SQL> delete scott.fwl_t1 where id = 1;
1 row deleted.
SQL> commit;
Commit complete.
SQL> select * from scott.fwl_t1;
no rows selected
SQL>
 SQL> select * from fwl.fwl_t1;
 no rows selected
 SQL>
topic信息查看
 "payload": {
   "before": {
     "ID": "1",
     "NAME": "bb"
   },
   "after": null,
   "source": {
     "version": "2.7.3.Final",
     "connector": "oracle",
     "name": "fwldb",
     "ts_ms": 1742335644000,
     "snapshot": "false",
     "db": "FWLDB",
     "sequence": null,
     "ts us": 1742335644000000,
     "ts ns": 17423356440000000000,
     "schema": "SCOTT",
     "table": "FWL_T1",
     "txId": "06001f00380e0000",
     "scn": "3731008",
     "commit_scn": "3731009",
     "lcr_position": null,
     "rs id": "0x0001e4.00002a0b.0010",
     "ssn": 0,
     "redo_thread": 1,
     "user_name": "SCOTT",
     "redo_sql": null,
     "row_id": "AAAaxEAAEAAAAFrAAA"
   },
   "transaction": null,
   "op": "d",
   "ts_ms": 1742306848644,
   "ts us": 1742306848644706,
   "ts_ns": 1742306848644706000
 }
'rocessed a total of 1 messages
fwl@kafka kafka_2.13-2.8.2]$
Note:可以看到数据同步正常。

≪ 墨力计划 debezium
```

