etl过程中断监控通知解决方案

2020年2月24日

22:18

**背景:**

在进行数据处理过程中,存在不可控风险(如源进行字段调整后与目标数据字段不匹配,字段长度不匹配),导致etl进程无法进行抽数操作,该问题发生时,无相关预警,不能在问题发生时进行及时处理.

**需求:**

在进行etl处理的过程中,出现etl过程问题后系统监控到问题并立刻将问题通过邮件发送,建立高效的数据问题反馈机制.

**风险:**

该邮件功能需办公网内有自建的邮件服务器,且可通过代理方式进行相关的邮件发送.

**具体方案:**

**一. 开发需要监控的etl抽数脚本,这里建立一个简单的实例**

1. **创建数据库抽数源表及目标表**

目标表中对字段长度进行修改,用于抛出异常.

计算机生成了可选文字:
localhost
老informationschema
老performanceschema
对象《s。凵r匚巳。稱tm@．
添加字的：适入字的
字的索弓分键
descrlptlon
目标表
巨动涕
囗无符号
囗盾充零
》怡rg巳t。稱tmat.
乒
删涂字的
注释SQL顸览
0
sourceofetlemail@
个上移下移
255
1024
老testetl
arg
老world
。稱tm

1. **建立需要监控的etl抽数脚本,如下:**

将source表的数据抽到target表中

计算机生成了可选文字:
`testetl
100
入

1. **建立抽数工作流脚本,如下:**

在监控的抽数脚本下,若该脚本发生错误,则进行邮件发送步骤.

计算机生成了可选文字:
testetl
>．囗10
START
除表中的茫
需要监控的抽数脚本
发送鄶
常抽数流程
错误时执行发送邮件
操作

具体邮件步骤设置如下:

地址:

计算机生成了可选文字:
#作业名称：
她址服务器§0地仁
收仁人
收仁人她址：
发仁人
cloudman@foxmail.com
回短名称鬈删e」oberror
发，老她址：goxyun@163.com
回短她址
确定囤）
OHelp

服务器设置:

计算机生成了可选文字:
#作业名称：
她址服务器§0地仁
#服务器
SMTP*%*•smtp.163.com
二号：25
用户验证？
用户名：goxyun@163.com
便甲安全验证？囗
确定囤）
OHelp

邮件消息设置:

计算机生成了可选文字:
#作业名称：
她址服务器§0地仁
信0里帝凸期？
．只发送0丨主释？
便甲HTML囗
[@河二二二二二二二二二二二二二二二二二二二二二二二二二三]
管优先囗
优[@二二二二二二二二二二二二二二二二二二二二二二二二二〔二二三]
[@二二二二二二二二二二二二二二二二二二二二二二二二二〔二二三]
主题kettleerror
注释·bi銃中e出壬““±现措誤
0凸志适到““位行苷
E志为““
确定囤）
OHelp

1. **使用kitchen.bat 工具进行调用,模拟windows任务调用情况:**

计算机生成了可选文字:
C:\WINDOWS\system32\cmd.exe
：\et《t00《\data—integration>kitchen.bat/fi《eD：\et《t00《\et《test\testet上kjb

结果如下:

计算机生成了可选文字:
C:\WINDOWS\system32\cmd.exe
2020/02/2423．45：17一表输出．
0一
2020/02/2423．45：17一表输出
．0一
2020/02/2423．45：17一表输出
．0一Caused
》ongforCO》umn'keyiatrow
2020/02/2423：45：17一表输出
．0一
tionsMappijava:104)
2020/02/2423：45：17一表《翁出．
0一
ent.java:970)
2020/02/2423：45：17一表《翁出．
0一
Statement.java:1109〕
2020/02/2423．45：17一表《翁出．0
Statement.java849〕
．17一表输出．0
atorg.pentaho.di．core.database.Database.emptyAndCo《(Database.java:14C3〕
3more
com.mysql.cj.jdbc.exceptions.Mysq《DataTruncation：Datatruncation：Datat00
by．
at
at
at
at
com.mysq《。cj.jdbc.exceptions.SQLExceptionsMapping.trans《ateException(SQLExcep
com.mysq《。cj.jdbc.引ientPreparedStatement.execute》nterna《(C《ientPreparedState
com.mysq《。cj.jdbc.引ientPreparedStatement.executeUpdate《nterna《(C《ientPrepare
com.mysq《。cj.jdbc.引ientPreparedStatement.executeBatchSeria《《y(C《ientPrepared
6more
2020/02/2423．45
2020/02/2423．45
2020/02/2423．45．18
2020/02/2423．45．18
2020/02/2423．45．18
2020/02/2423．45．18
2020/02/2423．45．18
D·et《t00《
2020/02/2423．45．17一testet《
2020/02/2423．45．17一testet》
2020/02/2423．45．17一testet》
2020/02/2423．45．17一testet《
2020/02/2423．45．17一testet《
2020/02/2423．45．18一testet》
2020/02/2423．45．18一Kitchen
．17一表输出．0一完成处理〔0，0，
一转换被检测
一ERROR〔version7．1．0．0一12，bui尾1from2017一05一1617．1&02bybui
R=3，W。0，U=0，E—l)
错误被检测到！
错误被检测到！
data—inte
一转换正在杀死其他步！
ERROR〔version7．1．0．0一12，bui尾1from2017一05一1617．1&02bybui
一开始项[发送邮件]
一完成作业项[发送邮件]（结果=[true])
rue])
testet》
丿C
一完成作业项[删除目标表中的数据]（结果=[true])
testet》
一任务执行完毕
testet》
Finished！
Kitchen
Kitchen一Start=2020/02/2423：45：01．561．St=2020/02/2423：45：1&309
Processingendedafter16seconds.
ration)

1. **邮箱中出现通知:**

计算机生成了可选文字:
Msil?aiFcfoå
Hia<cloudman@foxmail.com/
0箱苜页1设_扶
转发
kettleerror
发讎人：kettle_joberror<goxyun@163.com/
时河：2D20年224（星期一）说上11：45
#人：
系统中et]任务粼粼**土现误
体日志到***到萱进行音
0
际记力．
移的到“
反馈建议]帮助中心]退±
篁本》囗
0通
收侔箱（349）
星际#
鄄件（3）
垃圾箱（1刀
QQ鄶件订阅（2〕
早他箱
历》记事本
地仁收
仁中转站
鹫卡》启片
亞河
[空]
作业名
作业项
：testetl
到发送那件
上次的貊果：
作业项号
读取行
写入行
输土行
2
正在等cn.gravatar.cam的应“