

oozie investigate

OOZIE RERUN

BACKGROUND

Reprocessing is an important operational undertaking in any complex system

In a production environment, job reprocessing is a very common and critical activity.

There are three scenarios when a user needs to rerun the same job:

- The job failed due to a transient error.
- The job succeeded but the input data was bad.
- The application logic was flawed or there was a bug in the code.

WORKFLOW RERUN

重新运行作业，除了需要原有流程的id之外，还需要重新配置作业属性，它可以和原来的属性不一样，这里面有两个属性是必须要注意的，它们是重新提交作业必须具备的属性，如果不存在就会报错的，它们分别是 `oozie.wf.rerun.failnodes`和`oozie.wf.rerun.skip.nodes`。

这两个必须必须选择一个，第一个是自动运行失败的流程节点，第二个是需要跳过的节点，通过设置这个属性的话，如果已经运行成功的节点不在这个属性里面的话就可以让已经运行成功的节点再运行一遍了。

WORKFLOW RERUN – PRE-CONDITIONS 限制条件

- Workflow with id wfId should exist.
工作流id必须已经存在
- Workflow with id wfId should be in SUCCEEDED/KILLED/FAILED.
工作流的状态必须是终结状态，可以是（成功，被killed，失败）
- If specified , nodes in the config oozie.wf.rerun.skip.nodes must be completed successfully.
被skip的节点必须是运行上次运行成功的节点

WORKFLOW RERUN – BASIC PRINCIPLE 基本过程

- Reloads the configs.
重新加载配置文件job.properties etc.
- Creates a new Workflow Instance with the same wfId.
创建原来workflow id的一个新的实例
- Deletes the actions that are not skipped from the DB and copies data from old Workflow Instance to new one for skipped actions.
如果一个节点没有被skip，那么将删除这个节点在数据库中的信息，并且拷贝旧的workflow实例的数据到新的要skip的节点
- Action handler will skip the nodes given in the config with the same exit transition as before.
Action 处理器会跟进配置直接跳过相应的节点并且顺着原来转移方向继续执行下去。

WORKFLOW RERUN – 使用方法

oozie.wf.rerun.failnodes=true ,说明可以只跑失败的节点

rerun命令会执行fail node, 如果此workflow本来就是成功的, 那么rerun会立即返回成功

```
oozie job -rerun 0000004-160419071951949-oozie-oozi-W -config job.properties -Doozie.wf.rerun.failnodes=true
```

oozie.wf.rerun.failnodes=false ,说明可以重新跑成功的节点

```
oozie job -rerun 0000004-160419071951949-oozie-oozi-W -config job.properties -Doozie.wf.rerun.failnodes=false
```

skip nodes,被skip的节点在上次运行时必须是运行成功的节点, 也就是说只能skip成功的节点而不能是失败节点

```
oozie job -rerun 0000004-160419071951949-oozie-oozi-W -config job.properties -Doozie.wf.rerun.skip.nodes=node1,node3
```