Oozie jobs api can be used to get information about the workflows. We can use filter to get [Completed oozie jobs](#_Hlk521583681" \s "1,335,356,0,,Completed oozie jobs)

(filter=status=SUCCEEDED). Id attribute in the workflow object contains the unique id of the workflow.

A [Workflow’s details](#_Hlk521584066) can be obtained by querying with the unique id of the workflow.

Each workflow contains multiple actions (oozie Launcher, hive, terasort, teragen etc).

Each actions contains 2 attributes

externalId – job ID of Oozie Launcher job

ExternalChildId – List of Yarn job Id of action which oozie is performing eg (hive, teragen, terasort etc).

It is a list because there can be multiple jobs created in YARN for one action (hive)

For Map reduce action,

The externalId directly contains the mapreduce Job Id and does not contain LauncherJobId.

We can further use this externalChildId, externalId to further query yarn for job, task , container details etc

Oozie reading jobs approaches:

Completed oozie jobs:

<http://pndademocloud-hadoop-mgr-4:11000/oozie/v1/jobs> - Gives the first 50 jobs irrespective of status

<http://pndademocloud-hadoop-mgr-4:11000/oozie/v1/jobs?offset=1&len=1&filter=status%3DSUCCEEDED>

{

"total": 51,

"offset": 1,

"len": 1,

"workflows": [

{

"appName": "hivetera-wf-22g",

"externalId": null,

"conf": null,

"run": 0,

"acl": null,

"appPath": null,

"parentId": null,

"lastModTime": "Thu, 09 Aug 2018 18:13:03 GMT",

"consoleUrl": "http://pndademocloud-hadoop-mgr-4:11000/oozie?job=0000011-180809041031612-oozie-oozi-W",

"createdTime": "Thu, 09 Aug 2018 18:07:05 GMT",

"startTime": "Thu, 09 Aug 2018 18:07:05 GMT",

"toString": "Workflow id[0000011-180809041031612-oozie-oozi-W] status[SUCCEEDED]",

"id": "0000011-180809041031612-oozie-oozi-W",

"endTime": "Thu, 09 Aug 2018 18:13:03 GMT",

"user": "pnda",

"actions": [],

"status": "SUCCEEDED",

"group": null

}

]

}

Single Workflow’s details:

http://pndademocloud-hadoop-mgr-4:11000/oozie/v1/job/0000011-180809041031612-oozie-oozi-W

{

"appName": "hivetera-wf-22g",

"externalId": null,

"conf": "<configuration>\r\n <property>\r\n <name>examplesRoot</name>\r\n <value>ooziehivetestcode</value>\r\n </property>\r\n <property>\r\n <name>oozie.wf.application.path</name>\r\n <value>hdfs://10.25.14.83:8020/user/pnda/ooziehivetestcode/ooziehivetera20g</value>\r\n </property>\r\n <property>\r\n <name>oozie.use.system.libpath</name>\r\n <value>true</value>\r\n </property>\r\n <property>\r\n <name>queueName</name>\r\n <value>default</value>\r\n </property>\r\n <property>\r\n <name>user.name</name>\r\n <value>pnda</value>\r\n </property>\r\n <property>\r\n <name>jobTracker</name>\r\n <value>10.25.14.83:8088</value>\r\n </property>\r\n <property>\r\n <name>userName</name>\r\n <value>pnda</value>\r\n </property>\r\n <property>\r\n <name>mapreduce.job.user.name</name>\r\n <value>pnda</value>\r\n </property>\r\n <property>\r\n <name>nameNode</name>\r\n <value>hdfs://10.25.14.83:8020</value>\r\n </property>\r\n</configuration>",

"run": 0,

"acl": null,

"appPath": "hdfs://10.25.14.83:8020/user/pnda/ooziehivetestcode/ooziehivetera20g",

"parentId": null,

"lastModTime": "Thu, 09 Aug 2018 18:13:03 GMT",

"consoleUrl": "http://pndademocloud-hadoop-mgr-4:11000/oozie?job=0000011-180809041031612-oozie-oozi-W",

"createdTime": "Thu, 09 Aug 2018 18:07:05 GMT",

"startTime": "Thu, 09 Aug 2018 18:07:05 GMT",

"toString": "Workflow id[0000011-180809041031612-oozie-oozi-W] status[SUCCEEDED]",

"id": "0000011-180809041031612-oozie-oozi-W",

"endTime": "Thu, 09 Aug 2018 18:13:03 GMT",

"user": "pnda",

"actions": [

{

"cred": null,

"userRetryMax": 0,

"trackerUri": "-",

"data": null,

"errorMessage": null,

"userRetryCount": 0,

"externalChildIDs": null,

"externalId": "-",

"errorCode": null,

"conf": "",

"type": ":START:",

"transition": "hive-node",

"retries": 0,

"consoleUrl": "-",

"stats": null,

"userRetryInterval": 10,

"name": ":start:",

"startTime": "Thu, 09 Aug 2018 18:07:05 GMT",

"toString": "Action name[:start:] status[OK]",

"id": "0000011-180809041031612-oozie-oozi-W@:start:",

"endTime": "Thu, 09 Aug 2018 18:07:05 GMT",

"externalStatus": "OK",

"status": "OK"

},

{

"cred": null,

"userRetryMax": 0,

"trackerUri": "10.25.14.83:8088",

"data": null,

"errorMessage": null,

"userRetryCount": 0,

"externalChildIDs": "job\_1533773070125\_0143,job\_1533773070125\_0147",

"externalId": "job\_1533773070125\_0142",

"errorCode": null,

"conf": "<hive xmlns=\"uri:oozie:hive-action:0.2\">\r\n <job-tracker>10.25.14.83:8088</job-tracker>\r\n <name-node>hdfs://10.25.14.83:8020</name-node>\r\n <prepare />\r\n <configuration>\r\n <property xmlns=\"\">\r\n <name>mapreduce.job.queuename</name>\r\n <value>default</value>\r\n <source>programatically</source>\r\n </property>\r\n <property xmlns=\"\">\r\n <name>mapred.job.queue.name</name>\r\n <value>default</value>\r\n <source>programatically</source>\r\n </property>\r\n </configuration>\r\n <script>tpch\_query1.sql</script>\r\n</hive>",

"type": "hive",

"transition": "HdfsDir",

"retries": 0,

"consoleUrl": "http://pndademocloud-hadoop-mgr-1:8088/proxy/application\_1533773070125\_0142/",

"stats": null,

"userRetryInterval": 10,

"name": "hive-node",

"startTime": "Thu, 09 Aug 2018 18:07:05 GMT",

"toString": "Action name[hive-node] status[OK]",

"id": "0000011-180809041031612-oozie-oozi-W@hive-node",

"endTime": "Thu, 09 Aug 2018 18:08:15 GMT",

"externalStatus": "SUCCEEDED",

"status": "OK"

},

{

"cred": null,

"userRetryMax": 0,

"trackerUri": "10.25.14.83:8088",

"data": null,

"errorMessage": null,

"userRetryCount": 0,

"externalChildIDs": null,

"externalId": "job\_1533773070125\_0150",

"errorCode": null,

"conf": "<shell xmlns=\"uri:oozie:shell-action:0.2\">\r\n <job-tracker>10.25.14.83:8088</job-tracker>\r\n <name-node>hdfs://10.25.14.83:8020</name-node>\r\n <configuration>\r\n <property xmlns=\"\">\r\n <name>mapreduce.job.queuename</name>\r\n <value>default</value>\r\n <source>programatically</source>\r\n </property>\r\n <property xmlns=\"\">\r\n <name>mapred.job.queue.name</name>\r\n <value>default</value>\r\n <source>programatically</source>\r\n </property>\r\n </configuration>\r\n <exec>hdfsdir.sh</exec>\r\n <env-var>HADOOP\_USER\_NAME=pnda</env-var>\r\n <file>lib/hdfsdir.sh#hdfsdir.sh</file>\r\n</shell>",

"type": "shell",

"transition": "TeraGen",

"retries": 0,

"consoleUrl": "http://pndademocloud-hadoop-mgr-1:8088/proxy/application\_1533773070125\_0150/",

"stats": null,

"userRetryInterval": 10,

"name": "HdfsDir",

"startTime": "Thu, 09 Aug 2018 18:08:15 GMT",

"toString": "Action name[HdfsDir] status[OK]",

"id": "0000011-180809041031612-oozie-oozi-W@HdfsDir",

"endTime": "Thu, 09 Aug 2018 18:08:36 GMT",

"externalStatus": "SUCCEEDED",

"status": "OK"

},

{

"cred": null,

"userRetryMax": 0,

"trackerUri": "10.25.14.83:8088",

"data": null,

"errorMessage": null,

"userRetryCount": 0,

"externalChildIDs": null,

"externalId": "job\_1533773070125\_0153",

"errorCode": null,

"conf": "<shell xmlns=\"uri:oozie:shell-action:0.2\">\r\n <job-tracker>10.25.14.83:8088</job-tracker>\r\n <name-node>hdfs://10.25.14.83:8020</name-node>\r\n <configuration>\r\n <property xmlns=\"\">\r\n <name>mapreduce.job.queuename</name>\r\n <value>default</value>\r\n <source>programatically</source>\r\n </property>\r\n <property xmlns=\"\">\r\n <name>mapred.job.queue.name</name>\r\n <value>default</value>\r\n <source>programatically</source>\r\n </property>\r\n </configuration>\r\n <exec>teragen.sh</exec>\r\n <env-var>HADOOP\_USER\_NAME=pnda</env-var>\r\n <file>lib/teragen.sh#teragen.sh</file>\r\n</shell>",

"type": "shell",

"transition": "TeraSort",

"retries": 0,

"consoleUrl": "http://pndademocloud-hadoop-mgr-1:8088/proxy/application\_1533773070125\_0153/",

"stats": null,

"userRetryInterval": 10,

"name": "TeraGen",

"startTime": "Thu, 09 Aug 2018 18:08:36 GMT",

"toString": "Action name[TeraGen] status[OK]",

"id": "0000011-180809041031612-oozie-oozi-W@TeraGen",

"endTime": "Thu, 09 Aug 2018 18:12:22 GMT",

"externalStatus": "SUCCEEDED",

"status": "OK"

},

{

"cred": null,

"userRetryMax": 0,

"trackerUri": "10.25.14.83:8088",

"data": null,

"errorMessage": null,

"userRetryCount": 0,

"externalChildIDs": null,

"externalId": "job\_1533773070125\_0163",

"errorCode": null,

"conf": "<shell xmlns=\"uri:oozie:shell-action:0.2\">\r\n <job-tracker>10.25.14.83:8088</job-tracker>\r\n <name-node>hdfs://10.25.14.83:8020</name-node>\r\n <configuration>\r\n <property xmlns=\"\">\r\n <name>mapreduce.job.queuename</name>\r\n <value>default</value>\r\n <source>programatically</source>\r\n </property>\r\n <property xmlns=\"\">\r\n <name>mapred.job.queue.name</name>\r\n <value>default</value>\r\n <source>programatically</source>\r\n </property>\r\n </configuration>\r\n <exec>terasort.sh</exec>\r\n <env-var>HADOOP\_USER\_NAME=pnda</env-var>\r\n <file>lib/terasort.sh#terasort.sh</file>\r\n</shell>",

"type": "shell",

"transition": "end",

"retries": 0,

"consoleUrl": "http://pndademocloud-hadoop-mgr-1:8088/proxy/application\_1533773070125\_0163/",

"stats": null,

"userRetryInterval": 10,

"name": "TeraSort",

"startTime": "Thu, 09 Aug 2018 18:12:22 GMT",

"toString": "Action name[TeraSort] status[OK]",

"id": "0000011-180809041031612-oozie-oozi-W@TeraSort",

"endTime": "Thu, 09 Aug 2018 18:13:03 GMT",

"externalStatus": "SUCCEEDED",

"status": "OK"

},

{

"cred": null,

"userRetryMax": 0,

"trackerUri": "-",

"data": null,

"errorMessage": null,

"userRetryCount": 0,

"externalChildIDs": null,

"externalId": "-",

"errorCode": null,

"conf": "",

"type": ":END:",

"transition": null,

"retries": 0,

"consoleUrl": "-",

"stats": null,

"userRetryInterval": 10,

"name": "end",

"startTime": "Thu, 09 Aug 2018 18:13:03 GMT",

"toString": "Action name[end] status[OK]",

"id": "0000011-180809041031612-oozie-oozi-W@end",

"endTime": "Thu, 09 Aug 2018 18:13:03 GMT",

"externalStatus": "OK",

"status": "OK"

}

],

"status": "SUCCEEDED",

"group": null

}

DocumentTypes:

**oozieWorkflows**

appName

id

lastModTime

createdTime

startTime

endTime

status

"plugin

\_documentType

\_tag\_appName

**oozieWorkflowActions:**

id

startTime

endTime

externalId – Launcher oozie yarn job id. This can be empty if oozie launches only one mapreduce job in YARN. Oozie reports externalId as empty and only contains externalChildID for such tasks.

externalChildID – Contains Yarn job id of mapreduce tasks(if there are multiple map reduce we will send multiple documents to Elasticsearch with each containing one externalChildID. All other fields between these documents will remain the same)

status – workflow oozie task status

externalStatus – Yarn Job status

errorCode

name

type – hive or mapreduce etc

retries

\_plugin

\_documentType

\_tag\_appName