

JobSub - Bug #6314

problem with dag generation

05/21/2014 10:27 AM - Dennis Box

Status:	Closed	Start date:	05/21/2014
Priority:	Normal	Due date:	
Assignee:	Dennis Box	% Done:	100%
Category:	JobSub Tools	Estimated time:	0.00 hour
Target version:		Spent time:	3.00 hours
First Occurred:		Stakeholders:	greenlee@fnal.gov wolbers@fnal.gov kirby@fnal.gov
Occurs In:			

Description

Hello Dennis,

Does condor_dagman kill worker jobs? I think I am seeing that when I use dagNabbit.py. At any rate, jobs are dying more and faster with dagNabbit.py than otherwise. Here is the path of an example dagman.out file:

```
/uboone/data/users/condor-tmp/greenlee/submit.20140519_145138.dag.dagman.out
```

I don't have the original submit.dag that goes with the above log file, but it is very similar to this one:

```
/uboone/app/users/greenlee/work/v1_01_00/reco3D/muoniso_cc1_reco3D/submit.dag
```

Thanks,

Herb

History

#1 - 06/04/2014 10:01 AM - Dennis Box

- Stakeholders updated

Hello Dennis,

I'd like to call your attention to another dagman log file that I ran yesterday and today. Actually, this dagman.out is still being updated. The condor_dagman job is still running. One or more jobs are still alive (held) -- don't know why.

The name of the log file is here:

```
/uboone/data/users/condor-tmp/greenlee/submit.20140603_220346.dag.dagman.out
```

The key section is the following, I believe:

```
06/04/14 01:09:44 Currently monitoring 1 Condor log file(s)
```

```
06/04/14 01:09:44 Event: ULOG_IMAGE_SIZE for Condor Node Jb_2_0 (16547952.49.0)
```

```
06/04/14 01:09:44 Event: ULOG_IMAGE_SIZE for Condor Node Jb_2_0 (16547952.57.0)
```

```
06/04/14 01:09:44 Event: ULOG_JOB_ABORTED for Condor Node Jb_2_0 (16547952.46.0)
```

```
06/04/14 01:09:44 Executing: condor_rm -const DAGManJobId' '==' '16547927' '&&' 'ClusterId' '==' '16547952
```

```
06/04/14 01:09:44 Running: condor_rm -const DAGManJobId' '==' '16547927' '&&' 'ClusterId' '==' '16547952
```

```
06/04/14 01:09:45 Number of idle job procs: 0
```

```
06/04/14 01:09:45 Event: ULOG_IMAGE_SIZE for Condor Node Jb_2_0 (16547952.45.0)
```

```
06/04/14 01:09:55 Currently monitoring 1 Condor log file(s)
```

```
06/04/14 01:09:55 Event: ULOG_SHADOW_EXCEPTION for Condor Node Jb_2_0 (16547952.7.0)
```

```
06/04/14 01:09:55 Number of idle job procs: 1
```

```
06/04/14 01:09:55 Event: ULOG_JOB_HELD for Condor Node Jb_2_0 (16547952.7.0)
```

```
06/04/14 01:09:55 Number of idle job procs: 1
```

06/04/14 01:09:55 Event: ULOG_JOB_EVICTED for Condor Node Jb_2_0 (16547952.2.0)
06/04/14 01:09:55 Number of idle job procs: 1
06/04/14 01:09:55 Event: ULOG_JOB_EVICTED for Condor Node Jb_2_0 (16547952.0.0)
06/04/14 01:09:55 Number of idle job procs: 1
06/04/14 01:09:55 Event: ULOG_JOB_ABORTED for Condor Node Jb_2_0 (16547952.2.0)
06/04/14 01:09:55 Number of idle job procs: 0
06/04/14 01:09:55 Event: ULOG_JOB_ABORTED for Condor Node Jb_2_0 (16547952.0.0)
06/04/14 01:09:55 Number of idle job procs: 0
.
.
.

If I look in the cluster log file here:

/uboone/data/users/condor-tmp/greenlee/reco3D-cosmic_gaus_reco3D.sh_20140603_220346_30402_0_1.log

at the same time there is this:

```
009 (16547952.046.000) 06/04 01:09:44 Job was aborted by the user.  
The system macro SYSTEM_PERIODIC_REMOVE expression '((NumJobStarts > 3) || (ImageSize >= 40000  
00) || (JobStatus 2 && JobUniverse 5 && ((CurrentTime - EnteredCurrentStatus) > 86400*3)))'  
evaluated to TRUE
```

So, one worker job used too much memory (ImageSize) and got killed. But this event triggered dagman, or condor, or something to kill the entire job cluster. This is obviously very inconvenient, and I would say unacceptable. Unless there is a way to change this, we can't use dag.

Here's the .dag file that went with this job, which I will leave for a while.

/uboone/app/users/greenlee/work/v02_00_01/reco3D/cosmic_gaus_reco3D/submit.dag

Is there a way we can change our .dag configuration, or invoke dagNabbit.py differently? Or maybe dagNabbit.py can be updated to do something smarter?

Thanks,

Herb

On Mon, 2 Jun 2014, Herbert Greenlee wrote:

Hello Dennis,

Here's an existing dagman.out file that had many killed jobs:

/uboone/data/users/condor-tmp/greenlee/submit.20140529_133808.dag.dagman.out

In this one, you can see many jobs being killed within a few seconds of each other, around 05/29/14 14:00.

Herb

On Mon, 2 Jun 2014, Dennis Box wrote:

Hi Herb,

First, I didn't get a chance to look at the dag in this email and now it is gone, sorry.

I have never encountered condor_dagman killing off worker jobs unless there are errors in the dag that prevent them from running.

I do notice that job 16533460 you submitted this morning is failing, with most of the jobs running into a SYSTEM_PERIODIC_REMOVE condition, either too many restarts or too much memory most likely. I can't read the log files because of permissions, can you chmod them so I can take a look?

Thanks,
Dennis

On 5/21/14 10:20 AM, Herbert Greenlee wrote:

Hello Dennis,

Does condor_dagman kill worker jobs? I think I am seeing that when I use dagNabbit.py. At any rate, jobs are dying more and faster with dagNabbit.py than otherwise. Here is the path of an example dagman.out

file:

/uboone/data/users/condor-tmp/greenlee/submit.20140519_145138.dag.dagman.out

I don't have the original submit.dag that goes with the above log file,
but it is very similar to this one:

/uboone/app/users/greenlee/work/v1_01_00/reco3D/muoniso_cc1_reco3D/submit.dag

Thanks,

Herb

#2 - 06/04/2014 10:06 AM - Dennis Box

Added Herb, Steve, and Mike as watchers to this ticket.

#3 - 06/04/2014 10:16 AM - Dennis Box

Hi Herb,

Can you change the permissions of /uboone/data/users/condor-tmp/greenlee/stop-reco3D-cosmic_gaus_reco3D.sh_20140603_220346_30408_0_1*
so I can read them?

Thanks,
Dennis

#4 - 06/04/2014 11:49 AM - Dennis Box

Hi Herb,

I think I understand what is going on.
You structured your dagNabbit job like so:

```
<serial>  
jobsub -n reco3D-cosmic_gaus_reco3D.sh #1 job one dag node  
jobsub -n -N 59 reco3D-cosmic_gaus_reco3D.sh #59 jobs but one dag node  
jobsub -n stop-reco3D-cosmic_gaus_reco3D.sh #1 job one dag node  
</serial>
```

and if one of the jobs in the middle node dies the entire node gets killed. The condor people would most likely argue that this is correct behavior.

If you restructure your job like this:

```
<serial>  
jobsub -n reco3D-cosmic_gaus_reco3D.sh #1 job one dag node  
</serial>  
<parallel>  
jobsub -n reco3D-cosmic_gaus_reco3D.sh #1 job node 1 of 59  
jobsub -n reco3D-cosmic_gaus_reco3D.sh #1 job node 2 of 59  
.....  
.....  
.....  
jobsub -n reco3D-cosmic_gaus_reco3D.sh #1 job node 58 of 59  
jobsub -n reco3D-cosmic_gaus_reco3D.sh #1 job node 59 of 59  
</parallel>  
<serial>  
jobsub -n stop-reco3D-cosmic_gaus_reco3D.sh #1 job one dag node  
</serial>
```

Then your generated dag would probably still not run to completion, but most of the reco3D-cosmic_gaus jobs would complete, and you would get a rescue dag that you could run once you found the problem with the ones that failed.

It would be possible to add a feature to automate this using an attribute for the <parallel> element, I imagine the input file could look something something like

```
<serial>  
jobsub -n reco3D-cosmic_gaus_reco3D.sh  
</serial>  
<parallel copies="59">  
jobsub -n reco3D-cosmic_gaus_reco3D.sh  
</parallel copies="59">  
<serial>  
jobsub -n stop-reco3D-cosmic_gaus_reco3D.sh
```

</serial>

Where dagNabbit would run the jobsub command N times when it saw the copies="N" flag.

I'm not sure how long it would take to get this feature into dagNabbit, its probably easy but I would have to look.

In the meantime, could you try restructuring your job to verify that it gives you the behavior you want?

Cheers
Dennis

#5 - 06/06/2014 05:24 PM - Dennis Box

- *Status changed from New to Resolved*

- *% Done changed from 0 to 100*

closing ticket

#6 - 07/29/2014 10:45 AM - Dennis Box

- *Status changed from Resolved to Closed*