

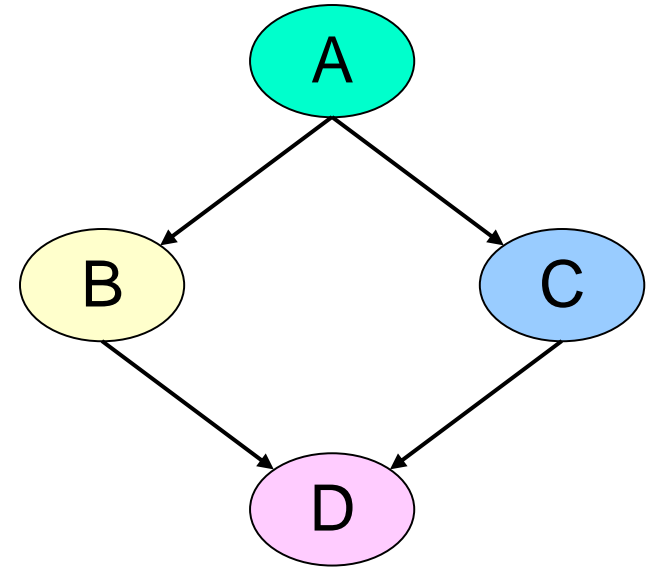
What's New in DAGMan

HTCondor Week 2013

Kent Wenger

DAGMan Introduction

- › Directed Acyclic Graph Manager
- › HTCondor's workflow management tool
- › User specifies dependencies between jobs, DAGMan manages them automatically
- › Features to handle many variations/special cases
- › Scales to very large workflows
- › Handles many error conditions
- › See Tuesday tutorial slides



Workflow log file

- › Greatly reduces DAGMan's file descriptor usage
- › User log names in submit files can include macros
- › Control with `DAGMAN_ALWAYS_USE_NODE_LOG, -dont_use_default_node_log`
- › New in 7.9.0
- › Must be disabled with pre-7.9.0 schedd
- › Bug caused HTCondor-C jobs to fail (7.9.0-7.9.5)
- › If `dagman_log` is specified in submit file, DAGMan's file "wins"

Top-level VARS setting applies to splices

› foo.dag:

```
Splice bar bar.dag
```

```
Vars bar+baz state="Wisconsin"
```

› bar.dag

```
Job baz baz.sub
```

› baz.sub:

```
arguments = $(state)
```

› Not for sub-DAGs

› New in 7.9.0

Set job attributes w/ VARS

- › Begin macroname with a “+” character to define a ClassAd attribute
- › For example, the following VARS specification:

```
Vars NodeE +A="\bob\""
```

would allow the HTCondor submit description file for NodeE to use the following line:

```
arguments = "$$([A])"
```

- › Like **+A="bob"** in submit file
- › Doesn't work for scheduler universe
- › New in 7.9.4

Suppressing emails from node jobs

- › Config:
`DAGMAN_SUPPRESS_NOTIFICATION`
- › Command line:
 - `-suppress_notification`
 - `-dont_suppress_notification`
- › Default is suppressing notification
- › New in 7.9.1
- › Default for all jobs became NEVER in 7.9.2

Status in DAGMan's ClassAd

```
> condor_q -1 59 | grep DAG_  
DAG_Status = 0  
DAG_InRecovery = 0  
DAG_NodesUnready = 1  
DAG_NodesReady = 4  
DAG_NodesPrerun = 2  
DAG_NodesQueued = 1  
DAG_NodesPostrun = 1  
DAG_NodesDone = 3  
DAG_NodesFailed = 0  
DAG_NodesTotal = 12
```

- › Sub-DAGs count as one node
- › New in 7.9.5

More info in node status file

...

Nodes total: 12

Nodes done: 8

Nodes pre: 0

Nodes queued: 3

Nodes post: 0

Nodes ready: 0

Nodes un-ready: 1

Nodes failed: 0

...

> New in 7.9.3

DAGMAN_USE_STRICT defaults to 1

- › Questionable settings that might cause subtle problems become immediate fatal errors (instead of just warnings)
- › **DAGMAN_USE_STRICT** range is 0-3
- › Default setting of 1 is new in 7.9.4 (was 0)

Log files in /tmp are errors

- › Default node log or individual job logs
- › Can cause DAG to fail (because /tmp may get cleared out)
- › Setting `DAGMAN_USE_STRICT` to 0 allows DAG to run (dangerously)
- › New in 7.9.4

Minor changes

- › `DAGMAN_LOG_ON_NFS_IS_ERROR` is ignored when both `CREATE_LOCKS_ON_LOCAL_DISK` and `ENABLE_USERLOG_LOCKING` are True
- › DAGMan will now try twice to write a POST script terminated event, rather than trying once and exiting

Relevant Links

- › DAGMan:

<http://research.cs.wisc.edu/htcondor/dagman/dagman.html>

- › For more questions:

htcondor-admin@cs.wisc.edu

