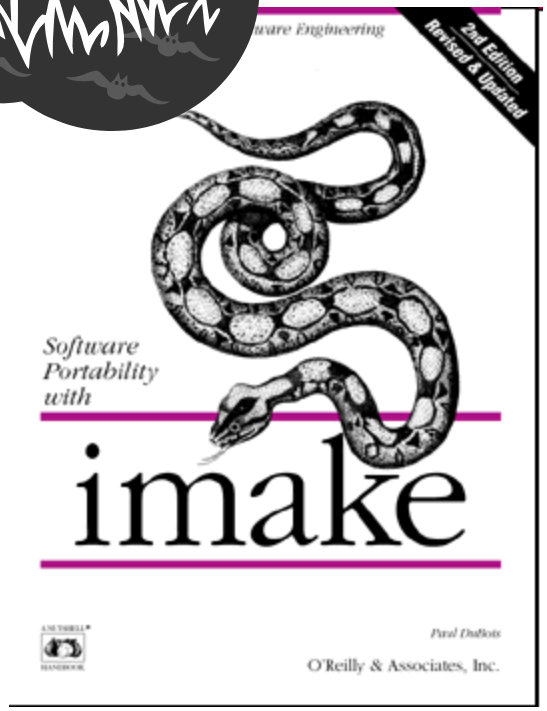
New goodies in v7.6

- Scalability enhancements (always…)
- File Transfer enhancements
- Grid/Cloud Universe enhancements
- Hierarchical Accounting Groups
- Keyboard detection on Vista/Win7
  - Just put "KBDD" in `Daemon_List`
- Sizeable amount of "Snow Leopard" work…

Extreme-Makeover – Build system edition!



Thanks Tim, the nightmares finally stopped

CMake
Cross-platform Make

Packaging via CPack, MSI via WiX

New ClassAds are now in Condor!

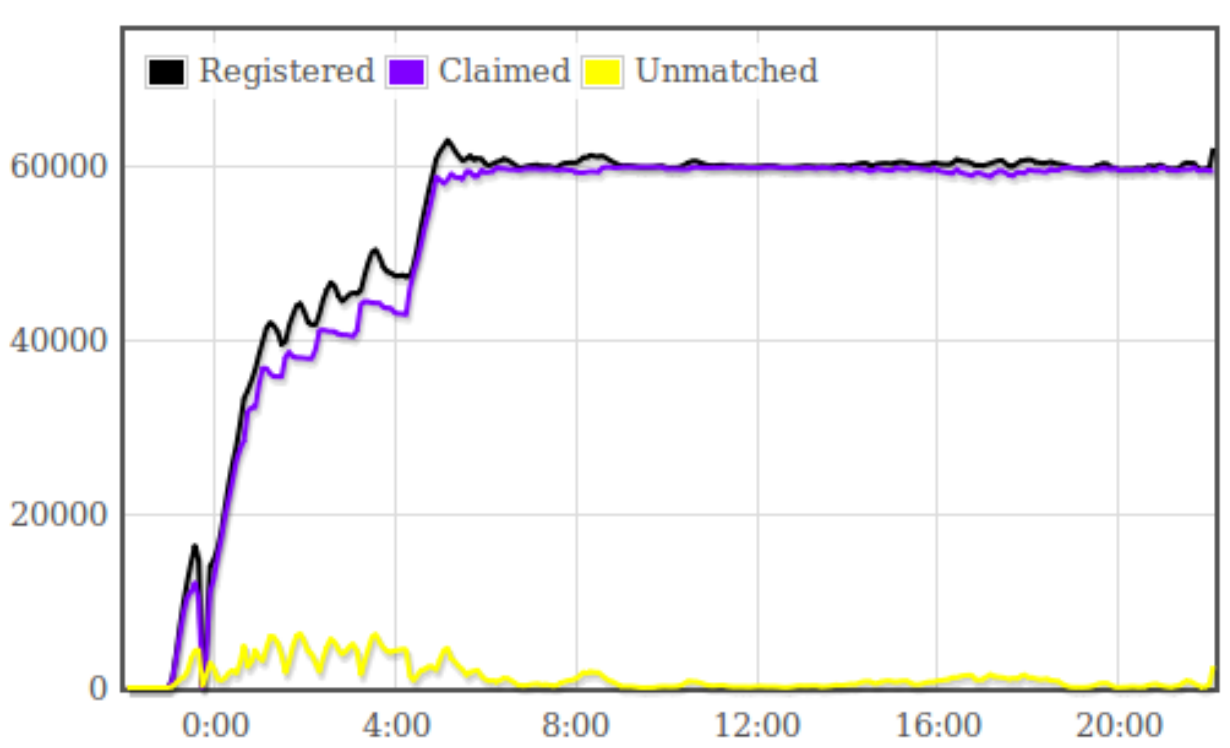
- Library in v7.5 / v7.6
  - Nothing user visible changes (we hope)… well, almost nothing
  - Logan's DAG Priority ?

    +JobPrio = DAGManJobId*100 + NumRestarts
- Developers are ready & eager to take leverage new ads in next dev series - but are users ready? Are YOU ready?

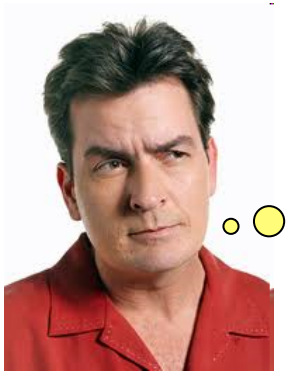
Scalability

- Everywhere
  - DaemonCore optimizations
  - Compiler optimizations enabled
- Schedd
  - Shadow recycling
  - Reduced fsync frequency (was a disk killer!)
  - Daemoncore optimizations
  - Reduced protocol overhead between shadows and the schedd
  - Asynchronous matchmaking(schedd would be unresponsive for O(10s))
- Collector
  - Moving to new classads - richer semantics and about 20% faster
  - Optimized classad insertion and removal
- Grid
  - Cream batching

of Cores managed by one Condor schedd: 60k



CHEP 2010 Paper
(slides at http://tinyurl.com/3nrawrw)



"Condor FASTBALL!"

An update on the scalability limits of the Condor batch system

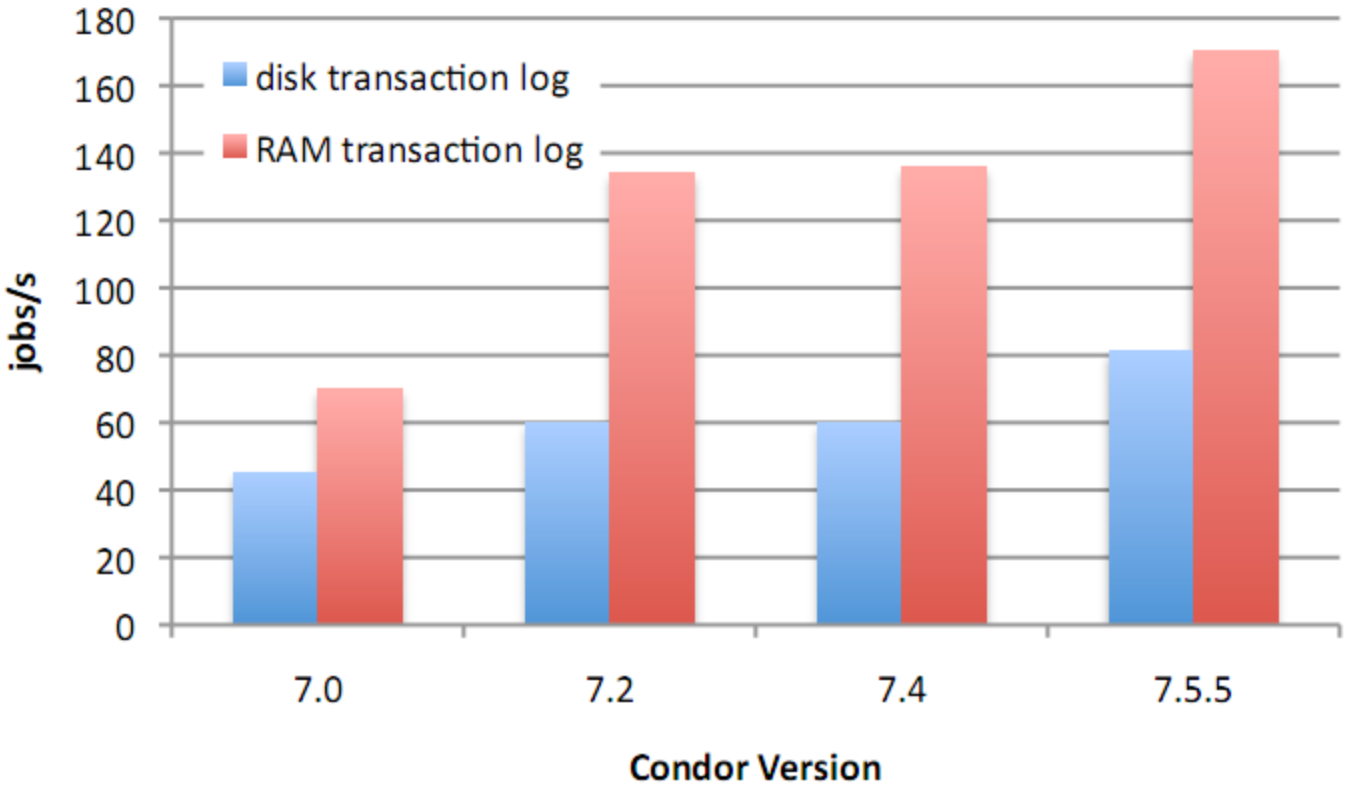
D Bradley[1], T St Clair[1], M Farrellee[1], Z Guo[1], M Livny[1], I Sfiligoi[2], T Tannenbaum[1]

[1]University of Wisconsin, Madison, WI, USA
[2]University of California, San Diego, La Jolla, CA, USA

E-mail: dan@hep.wisc.edu

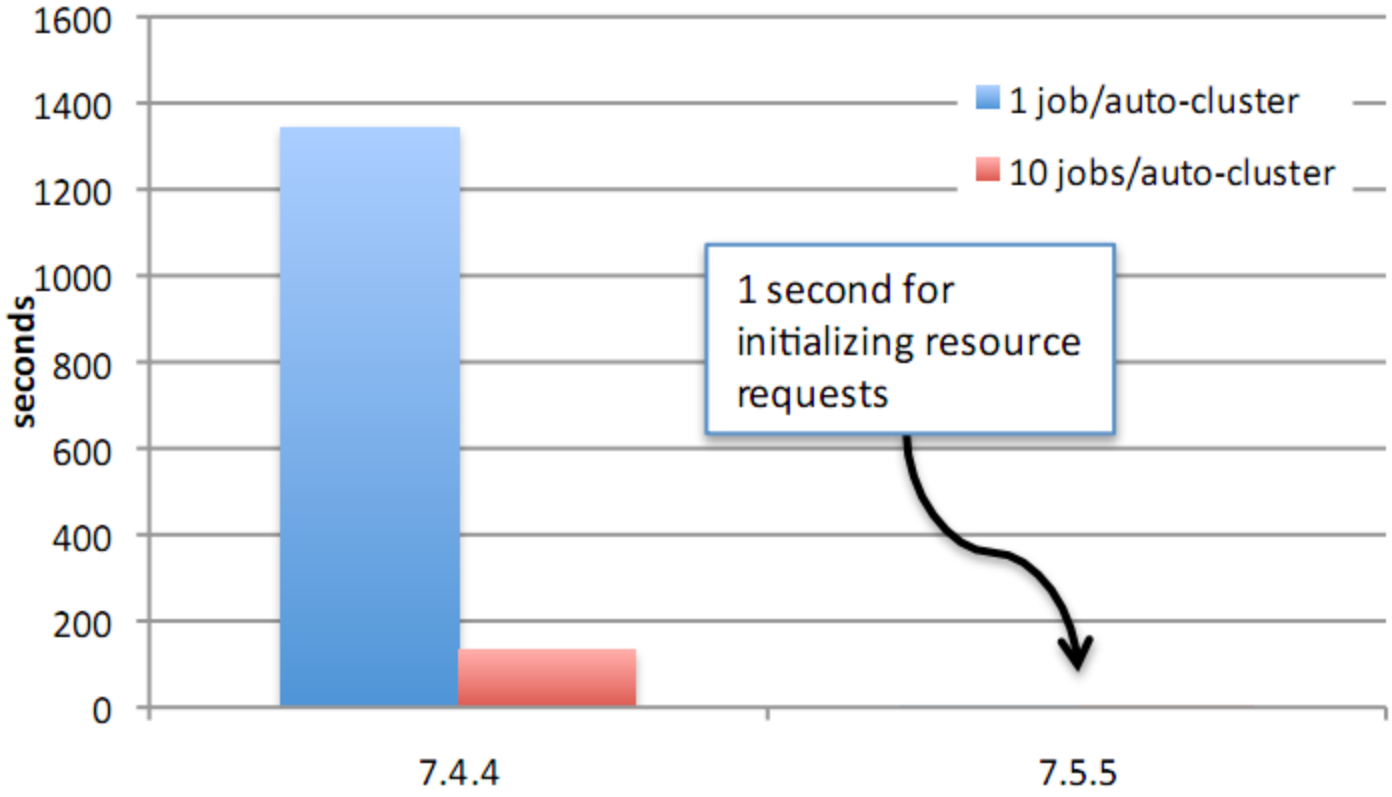
One Schedd: Maximum sustained job completion rate

(120k subsecond jobs, 'ideal' disk conditions)



Time spent in the schedd during negotiation

(same test… 1K jobs matched to 10K slot pool)
Schedd responsiveness improvement very noticeable, does other work



Condor File Transfer Hooks

- By default moves files between submit and execute hosts (shadow and starter).
- File Transfer Hooks - can have URLs grab files from anywhere
  - Globus Online, HDFS, all the usual like http:// …
- New – works for both stage-in and stage-out
- New - now works in VM Universe!
  - file:// URLs can be used to allow VM disk image files to be pre-staged on execute nodes

Contribution Modules

https://condor-wiki.cs.wisc.edu/index.cgi/wiki?p=ContribModules

- Quill, Stork, Aviary, DBQ, Remote Condor, QMF Management

Directory transfer

- Transfer of directories is now supported
- Just list them along with other files to be transferred

transfer_input_files = input.txt,scripts,libs

transfer_output_files = output.txt,images

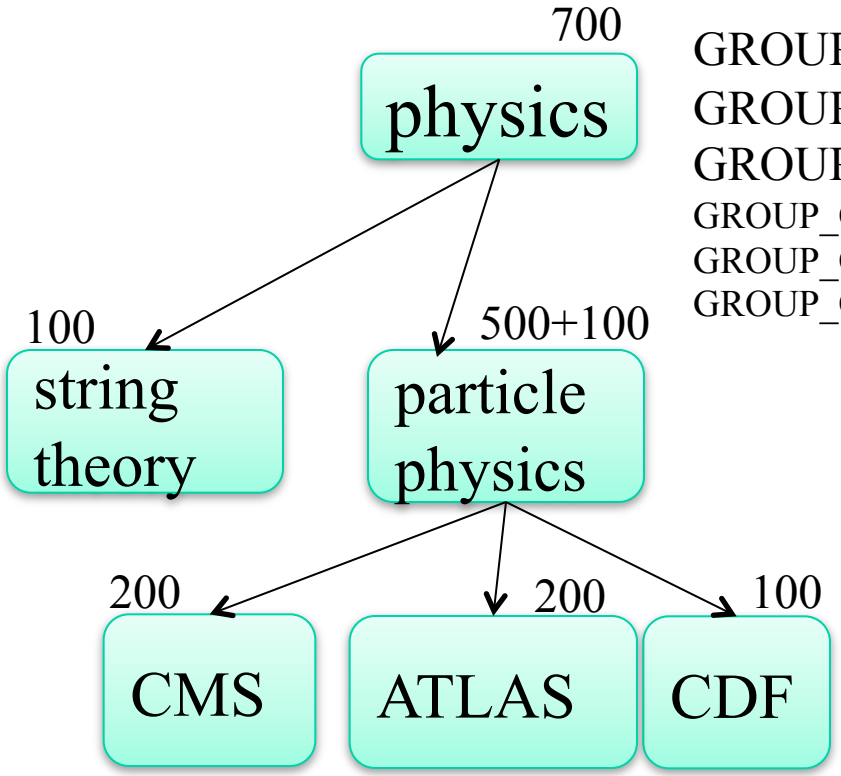
Directory transfer

- When output is auto-detected (i.e. transfer_output_files not specified), directories are *ignored*.
- This behavior may change in future versions.

Directory transfer

- Works in Condor-C and non-grid universes.
- Does not currently work for Globus jobs.

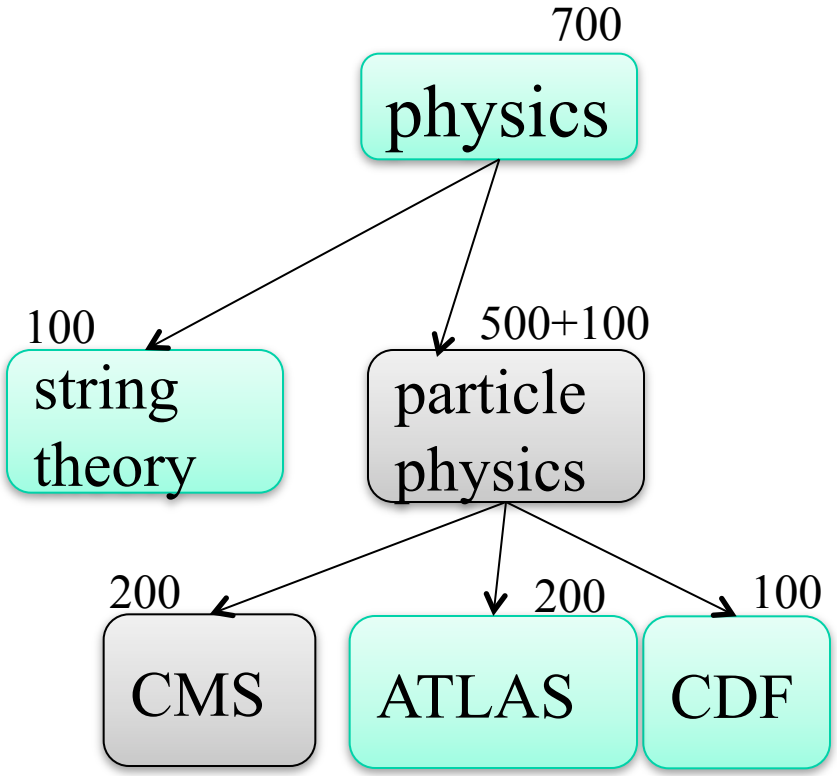
Hierarchical Group Quotas




```
GROUP_QUOTA_physics = 700
GROUP_QUOTA_physics.string_theory = 100
GROUP_QUOTA_physics.particle_physics = 600
GROUP_QUOTA_physics.particle_physics.CMS = 200
GROUP_QUOTA_physics.particle_physics.ATLAS = 200
GROUP_QUOTA_physics.particle_physics.CDF = 100
```

group.sub-group.sub-sub-group…

Hierarchical Group Quotas



Groups configured to accept surplus will share it in proportion to their quota.

Here, unused particle physics surplus is shared by ATLAS and CDF.

```
GROUP_ACCEPT_SURPLUS_physics.particle_physics.ATLAS = true
GROUP_ACCEPT_SURPLUS_physics.particle_physics.CDF = true
```

Changed behavior: AUTOREGROUP

- GROUP_AUTOREGROUP is not identical to behavior prior to 7.5.6
- Now it is equivalent to GROUP_ACCEPT_SURPLUS
- Sharing between users in different groups determined by group quotas, *not* user priorities.

Network Port Usage

- Condor v7.4 needs a lot of open network ports for incoming connections
  - Schedd: 5 + 5*NumRunningJobs
  - Startd: 5 + 5*NumSlots
- Not a pleasant firewall situation.
- CCB can make the schedd or the startd (but not both) turn these into outgoing ports instead of incoming

Have Condor listen on just *one* port per machine

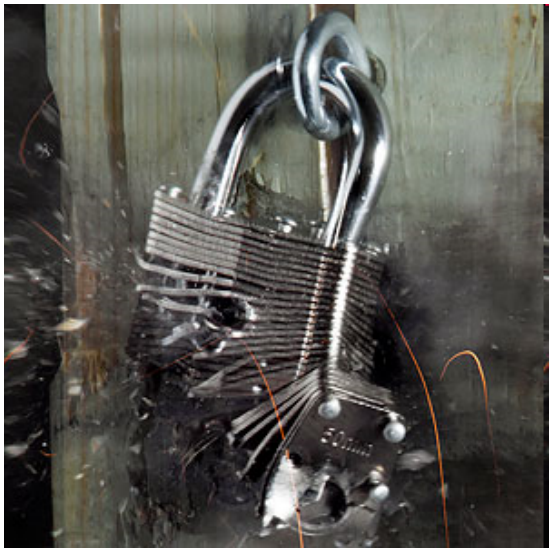
condor_shared_port

- All daemons on a machine can share one incoming port
  - Simplifies firewall or port forwarding config
  - Improves scalability
  - Running now on both Unix and Windows

USE_SHARED_PORT = True

DAEMON_LIST = … SHARED_PORT

Death to Locks!



- LOCK_DEBUG_LOG_TO_APPEND
  - Defaults to False on Unix
  - Relies on Posix O_APPEND semantics
  - Big gain w/ many running jobs
  - Only will lock on rotation
- CREATE_LOCKS_ON_LOCAL_DISK
  - Defaults to True
  - Lock files for event logs to /tmp
  - Relief for event logs in user home directories (aka on NFS)
  - Next step: only lock on rotation

Deltacloud Grid Type

- Project sponsored by Red Hat and Apache
- Has its own simple protocol
- Translates requests into protocol of target cloud service
  - Many protocols supported
  - Wider reach than EC2 protocol (e.g. GoGrid, Rackable, RHEV, …)

Igor's Talk Last Year…

Condor-G basically insecure!

- It takes a lot of trust to use Cond
  - At le
- If any
  your c
  - And nt that!
- The g r that vanilla Condor is mu ore secure
  - Condor team tells me remote admins can only access/modify files in the submit directory

Fixed!

You have to log in, right?

Madison, Apr 2010 Igor Sfiligoi 16 UCSD

DAGMan

- Condor_hold/condor_remove of DAGMan job now works correctly (7.5.6).
- $MAX_RETRIES for script argument (7.5.6).
- A bunch of new config macros to match existing condor_submit_dag command-line arguments (7.5.6).
- Condor_hold/condor_remove of DAGMan job now works correctly (7.5.6).
- Jobstate.log file (7.5.5).
- Node status file (7.5.4).
- The new configuration variable DAGMAN_MAX_JOB_HOLDS specifies the maximum number of times a DAG node job is allowed to go on hold (7.5.4).
- Category throttles in splices can be overridden by higher levels in the DAG splicing structure (7.5.3).
- Lazy creation of submit files for nested DAGs (7.5.2).