Shooting for the sky: Testing the limits of condor

HTCondor Week 2015

21 May 2015

Edgar Fajardo
On behalf of OSG Software and Technology
Acknowledgement

Although I am the one presenting. This work is a product of a collaborative effort from:

• HTCondor development team
• GlideInWMS development team
• HTCondor UW CHTC (A Moate), UCSD T2 (Terrence M) and CMS FNAL T1 (Burt H), who provided the hardware for the testbed.
• CMS Computing and WLCG sites provided the worker node resources.
• OSG Software team.
Vanilla Condor on a Slide

Legend:
- Schedd
- Central Manager
- Worker node

Proudly providing HTC for more than 20 years

HTCondor Week 2015
GlideInWMS on a slide

Happily working for more than 6 years !!!

Legend:
- Schedd
- Central Manger
- Worker node

HTCondor Week 2015
Vanilla HTCondor vs GlideInWMS in the numbers

<table>
<thead>
<tr>
<th></th>
<th>Vanilla Condor</th>
<th>GlideinWMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of nodes</td>
<td>O(1k)</td>
<td>O(10k)</td>
</tr>
<tr>
<td># Different types of machines</td>
<td>O(10)</td>
<td>O(100)</td>
</tr>
<tr>
<td>Location of Schedds and WN</td>
<td>Private</td>
<td>WAN</td>
</tr>
</tbody>
</table>
The Challenge:

Hi OSG Software folks

For Run II, we would like to have a single global pool of 200,000 running jobs, u think Condor can handle it?

Wow, We would expect so. Nobody has tried that big of a pool

Nobody has found the Higgs before we and ATLAS did

Touche

Btw, u can use some of my WLCG resources to test. About 20k slots

Just 20k? Didn’t u say u want a pool of 200k? Nvm we will figure it out.
The Challenge in the numbers

<table>
<thead>
<tr>
<th>CMS Requirements for LHC Run II</th>
</tr>
</thead>
<tbody>
<tr>
<td># Startd</td>
</tr>
<tr>
<td>Autoclusters (types of machines)</td>
</tr>
<tr>
<td>Number of Schedds</td>
</tr>
<tr>
<td>High Availability</td>
</tr>
</tbody>
</table>

Challenge Accepted !!!
How to get 200k slots?

Gathering from a commercial cloud for example*:

\[
\frac{0.013}{\text{hour} \times \text{core}} \times \frac{24 \text{ hours}}{\text{day}} \times \frac{30 \text{ days}}{\text{month}} \times 200\text{k cores} = \frac{1,872,000}{\text{month}}
\]

Without using your credit card

Our Solution: The UberGlideIn

*At spot pricing and without using the Uber GlideIn
Uber GlideIn

Normal GlideIn

1 Core

GlideIn

One Master

One Startd

This works because jobs are sleep jobs.

pilot_cpu_benchmark.py

UberGlideIn

1 Core

GlideIn

Master 1

Master 2

Master 3

Master n

Startd_1

Stard_2

Startd_3

Startd_n

HTCondor Week 2015
Moreover we wanted them distributed all over.

Why?

In production, the network latency of having Startd all over the WAN is known to cause problems. See [Bradley et all](#).
Does it work?

Yes it does !!!
Well most of the time it works:

Hi OSG Software folks

About those tests you are running at our site. Just FYI u brought down our firewall, while using > 10k ports

Ohh dude, sorry about that. We will fix it

Mental note: Talk to the HTCondor dev team to reduce the long lived TCP connections from the WN to the outside (Collector, Negotiator …)
Now we have the resources, let's test:

- For ~ 3 months
- Report/Brainstorm
- Test
- Fix

Ganglia features were key. THANKS!!!

HTCondor Week 2015

GANGLIAD_REQUIREMENTS = MyType != "Machine"
Did it work? YES !!!
“It was not all a bed of roses”, maybe it was...
HTCondor Improvements

For more details see Todd’s talk on What’s new in HTCondor 2015?

• Non blocking GSI authentication at the Collector
• Shared port at the worker node. In general reduce # of long lived TCP connections.
• Removed file locking at the Schedd
• Reduced incoming TCP connections at the Schedd
• Batched resources request from the Collector
Scale Improvements throughout “History”

Max parallel Running Jobs single HTCondor Pool in latest series

HTCondor Series

- 7.0.1
- 7.1.3
- 7.3.1
- 7.5.5
- 8.3.0
- 8.3.1
- 8.3.2

Max Parallel Running Jobs

- 200
- 1000
- 2000
- 9000
- 95000
- 150000
- 200000

CHEP 08

CHEP 09

CHEP 2015
Ahh, One more thing ...

Hi Edgar,

Since u and Jeff are working on the scaling tests, what about we scale test our new rockstar: The HTCondor CE.

About the size of UNL sleeper pool
~16k parallel running jobs?

Sounds good, Which levels are u looking for?

Are u kidding me? That is twice as much of what any OSG site needs?
Step back: What is a CE?
HTCondor-CE in a slide
What is HTCondor-CE?

Condor + Configuration

For more details see OSG AHM 2015 or CHEP 2015 talks.
Did we make it?
YES !!!

After some development cycles:

16 k Jobs achieved

UNL sleeper batch started to swap

Better than any Gram5 known to date
HTCondor CE In the Numbers:

<table>
<thead>
<tr>
<th></th>
<th>HTCondor-CE</th>
<th>GRAM 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Best max running jobs</td>
<td>16k*</td>
<td>10k</td>
</tr>
<tr>
<td>Network Port usage (per running job)</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>StartUp Rate</td>
<td>70 jobs/min</td>
<td>55 jobs/min*</td>
</tr>
</tbody>
</table>

*Disclaimer: This tests were done on different hardware with 5 years in between them.
Conclusions

• “The OSG Software team, in conjunction with HTCondor and GlideinWMS development teams have collaborated to push the scalability limits of a single HTCondor pool “
• The HTCondor-CE is ready to rock and roll
Questions?

Contact us at:

1-900-scale-masters
Just Kidding

Contact us:

osg-software@opensciencegrid.org