A Virtual Comet

HTCondor Week 2017

May 3 2017

Edgar Fajardo
On behalf of OSG Software and Technology
Working in Comet

What my friends think I do
What Instagram thinks I do
What I think I do
What my boss thinks I do
What is Comet?

“HPC for the long tail of Science”
Where is Comet?

• Comet is located at the San Diego Super Computing Center in La Jolla, California. On the UCSD Campus.
## Comet by the numbers

<table>
<thead>
<tr>
<th>System Component</th>
<th>Specs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Racks</td>
<td>27</td>
</tr>
<tr>
<td>Computes nodes per rack</td>
<td>72</td>
</tr>
<tr>
<td>Cores per Node</td>
<td>24 x Dual socket Haswell 12 core @ 2.5GHz</td>
</tr>
<tr>
<td>Ram per Node</td>
<td>128GB</td>
</tr>
<tr>
<td>Total number of Cores</td>
<td><strong>47000</strong></td>
</tr>
</tbody>
</table>
Three ways of accessing a Comet

1. Usual “old” way submitting jobs to SLURM batch system
2. Virtual Cluster Interface
3. Science Gateways (web portals for science domains)
Virtual Cluster Interface

• A “cloud” like API to get resources and monitor them.
• For example how to request a VM:

```plaintext
cm comet start osg --count=1 --walltime=2d --allocation=csd428
```

Cluster Name

Comet allocation

XSEDE Allocation cannot be used, yet.
How Comet/Condor integration works

Hosted at UCSD T2

HTCondor -CE

- job1: +project_Name="allocation1" +CometOnly=True
- job2: +project_Name="allocation1" +CometOnly=True
- job3: +project_Name="allocation1" +CometOnly=True

Cloudmesh

start/stop VM

condor_q

Black Box

Central Manager

Virtual Cluster

vm-1/2/3

Job1

Job2

Job3
Where does OSG kick in?

Glideins can get into Comet using the already existing UCSD T2 grid infrastructure.
From the Comet Virtual Cluster Head Node

- Puppet Master / Foreman
- DHCP
- HTCondor Central Manager
- Squid Proxy

~Same puppet config as the T2
It looks just as another UCSD entry in GlideinWMS:

```xml
<entry name="OSG_T2_US_UCSD_gw6_Comet_mcore" auth_method="grid_proxy+project_id" enabled="True" gatekeeper="osg-gw-6.t2.ucsd.edu osg-gw-6.t2.ucsd.edu:9619" gridtype="condor" trust_domain="grid" verbosity="std" work_dir="Condor">
  <config>
    <max_jobs>
      <default_per_frontend glideins="5000" held="50" idle="100"/>
      <per_entry glideins="10000" held="1000" idle="4000"/>
      <per_frontends/>
    </max_jobs>
    <release max_per_cycle="20" sleep="0.2"/>
    <remove max_per_cycle="5" sleep="0.2"/>
    <restrictions require_glidein_glexec_use="False" require_voms_proxy="False"/>
    <submit cluster_size="10" max_per_cycle="25" sleep="2" slots_layout="fixed">
      <submit_attr name="+CometOnly" value="True"/>
      <submit_attr name="+maxMemory" value="98304"/>
      <submit_attr name="+xcount" value="24"/>
    </submit_attr>
  </submit>
</config>
```

Allocation base

Do not run at UCSD

24 cores per Pilot
Achievements

• Successfully ran LIGO, Xenon1T, IceCube, CMS Production and CMS UCSD user jobs in the Virtual
In last 30 days, it has delivered 845k CPU Hours with the OSG VC.
Achievements

OSG: 3rd largest user of Comet
Scavenged Used Cycles

OSG Comet Virtual Cluster would like to make use of unused cycles… free science

Comet available nodes shown in dark blue… 7 days in December 2016
Scavenged Used Cycles

OSG Comet Virtual Cluster would like to make use of *unused cycles*…

*Comet available nodes shown in dark blue… 7 days in February 2017… where did they all go?*
Acknowledgements

• Trevor Cooper, Dmitry Mishin (SDSC) and the whole Comet team.
• Fugang Wang and Gregor von Laszewski (Indiana University) for the troubleshooting in the Comet Cloudmesh.
• Terrence Martin (UCSD) for the full integration setup and help debugging the network infrastructure at Comet Virtual Cluster.
• HTCondor team. THANKS !!!!!
Edgar’s Wishlist

These are things that would make my life easier:

• Support for Condor Metrics to InfluxDB
• Support for Python bindings in with python2.7 and pip install.
• Conditional flocking. (Even with just submitter ads)
• A job Ad that tells me the CPU efficiency of a job.
• Declare a resource uncontested (so its use does not count towards priorities).
Questions?

Contact us at:

1-900-OSG-Comet-Masters
Just Kidding

Contact us:

emfajard@ucsd.edu

Thank You
Comet Network Architecture
InfiniBand compute, Ethernet Storage

- Node-Local Storage: 320 GB, 72 HSWL, 18 racks
- Management
- Gateway
- Login Data Mover
- Home File Systems VM Image Repository
- FDR 36p
- 18 switches
- FDR 72
- 2*36 FDR
- 36 Ethernet Bridges (4 x 18-port each)
- Arista 40Gbps (2x)
- Arista 40Gbps (2x)
- Internet 2
- Research and Education Network Access Data Movers
- Juniper 100 Gbps
- Data Movers

Performance Storage
- 7.7 PB, 200 GB/s
- 32 storage servers

Durable Storage
- 6 PB, 100 GB/s
- 64 storage servers

Additional Support Components (not shown for clarity)
- Ethernet Mgmt Network (10 GbE)