HTCondor and Beyond: Research Computer at Syracuse University

Eric Sedore

ACIO – Information Technology and Services
HTCondor @ Syracuse (OrangeGrid)

- Scavenging compute time from ~3000 desktops on campus
- Virtual machines running in the background on idle desktops with HTCondor running inside the VM – tunneled to a private HTCondor network via VPN
- Condor Virtual Machine Coordinator (CVMC) – interfaces with HTCondor to launch and manage virtual machines
- Grown from 6000 to 10000+ cores over the last year
  - Added ~1000 dedicated cores
Updates to OrangeGrid

- Campus backbone upgrade to 40Gb and building links upgraded to 10Gb where node density is significant
- Methodology to support large VM’s (20GB+) on desktop clients
- Dynamically build virtual machine based on job requirements (CPU(s), memory)
- New controls to manage how many of a job type (with it’s associated VM) may run on each client (due to VM/jobs requiring heavy disk I/O or larger VM’s introducing memory pressure)
Focus Areas

1. Bridging the gap with the researcher
   - Better understand the researcher’s tool
   - Assist with scripting the “last mile” - finding the right level of investment
   - Does more assistance = more scholarly output?

2. Dynamically provision hardware resources
   - Researchers coming in all shapes and sizes
   - Compute resources are available to meet most needs, however given the diversity of research and the need for custom environments it is often in the wrong form
   - Focus on a virtualized research cloud
Chaos!
Research Cloud – backfill, priority, and oversubscription

- Attempting to blend single large machines, small clusters of largish machines (100’s), then typically smaller compute nodes to fill in the "cracks"
- Resource caps, pools, and weighting to try and place and manage load effectively
- Provide “hints” about what work should be prioritized and/or we can let the workloads “fight it out” controlled by the scheduler
- “flock of birds” when a large virtual machine (32 cores / 128 GB) goes active on a host as “load balancing” kicks in
- Current oversubscription is 3:1 – 4:1 or 5:1 appears possible though contention will become more likely (can we identify who cares if they run a little slower?)
An old, somewhat painful story - VM Management

- Three different hypervisors
  - Oracle VirtualBox - OrangeGrid
  - VMware ESX – research cloud
  - Microsoft Hyper-V – research cloud
- Four research VM’s distributed across the hypervisors (with multiple versions)
  - Default Linux node (Ubuntu)
  - LIGO (SL6)
  - 3D Render Node (Ubuntu)
  - BLAST Search (Ubuntu)
The Wonderful, Unpredictable World of Research Computing

- System Administrator
  - How many CPU’s, how much memory, how many IOPS, and what type of networking? Give me your requirements…
  - Do you need backups, how often, is there a time when load is low?
  - What kind of uptime do you need?
  - Just 60 terabytes?

- Researcher
  - I’d like to transform my research, how much can I have?
  - You do backups? Yes, of course I want backups. No, it is unpredictable and might be busy at any time
  - If it is up when I need it that is fine
  - Just 60 terabytes
Questions?