Managing a growing campus pool

Eric Sedore
essedore@syr.edu
Quick context

- Scavenging CPU time from 2000 desktops on campus
- Growth from 3000 to 6000+ cores over the last year
- Varied types of research (in all dimensions, run time, data needed per job, data access type (HTCondor transfer/NFS))
- Challenge feeding data to the compute nodes
  - VPN tunnel bandwidth
  - SAN Storage / Virtual server environment
  - Campus network
- Exploring steering jobs via average availability of each node
Steering jobs

- More Traditional – CPU/Memory/Application requirements
- Jobs that are not easily checkpointed and have longer run times
- Dynamic average giving the job a view into the amount of scavenging time (in hours) of a node – publish via classad
  - sessionlength = "44"
  - sessionlength = "9"
  - sessionlength = "10"
  - sessionlength = "41"
- Add to job requirements or ranking: TARGET.sessionlength >= "4"
- Data gathered through parsing logs from the last seven days (from Condor VM Coordinator)
Dealing with growth

- Data delivery infrastructure - horizontal and vertical scaling
- Network infrastructure upgrades – upgrading 1 Gb links to 10 Gb links
- Scaling out VPN end points – working toward a model, 1000-1500 per endpoint
- Throttling vs. bottleneck (protecting network links where there are high compute node populations)
Pushing NFS

- ~6300 nodes supported on a single NFS server (65000+ open files)
- Can run out of CPU at peak times (current server has 16 cores) – context switching / I/O wait
- Virtual Machine running in ESX 5.x – SAN storage (15K spindles on an IBM 5300)

- 3.0.0-21-virtual #35-Ubuntu SMP Fri May 25 18:35:12 UTC 2012 x86_64 x86_64 x86_64 GNU/Linux
Pool up and running – supporting the load