



HTCondor Private Cloud Integration

Andrew Lahiff

STFC Rutherford Appleton Laboratory

European HTCondor Site Admins Meeting 2014

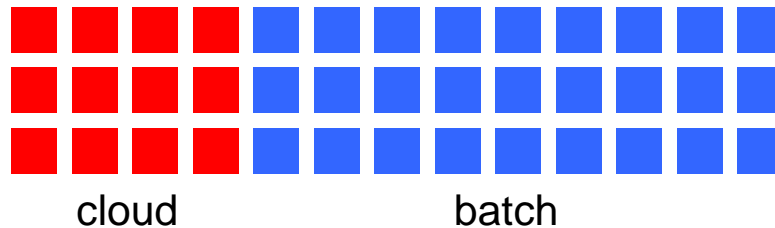
Clouds & batch systems

- Some sites might be (or be thinking about)
 - Moving some/all services to an internal private cloud
 - Setting up a private cloud in parallel to their batch system
 - Have access to a private cloud
- Problem: partitioned resources
 - Worker nodes (batch system)
 - Hypervisors (cloud)
- Likely to be a common situation at sites providing both grid & cloud resources

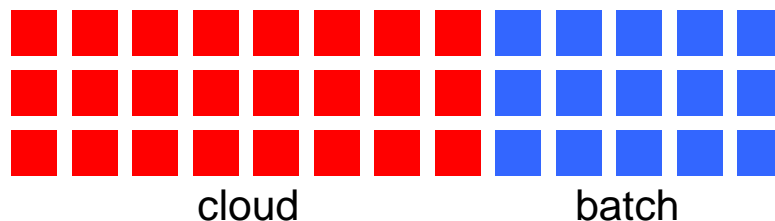


Clouds & batch systems

- Ideal situation: completely dynamic
 - If batch system busy but cloud not busy
 - Expand batch system into the cloud



- If cloud busy but batch system not busy
 - Expand size of cloud, reduce amount of batch system resources



Clouds & batch systems

- Could just manually create virtual WNs
 - But it would be better to dynamically create them as they are needed
- Need some method of creating virtual WNs
 - Could use existing products
 - glideinWMS
 - Cloud Scheduler
 - Write your own
 - A few examples on github
 - Cloud autoscaling service
 - Try to use existing HTCondor functionality
 - ...



Using HTCondor

- Advertise appropriate offline ClassAd(s) to the collector
 - Hostname used is a random string
 - In our use case these represents types of VMs, rather than specific machines
 - E.g. for VO-specific VMs, have an offline ClassAd for each type of VM
- `condor_rooster`
 - Enable this daemon
 - Configure to run appropriate command to instantiate a VM
 - HTCondor pool password inserted into the VM
 - Volatile disks created on hypervisor's local disk for job scratch area

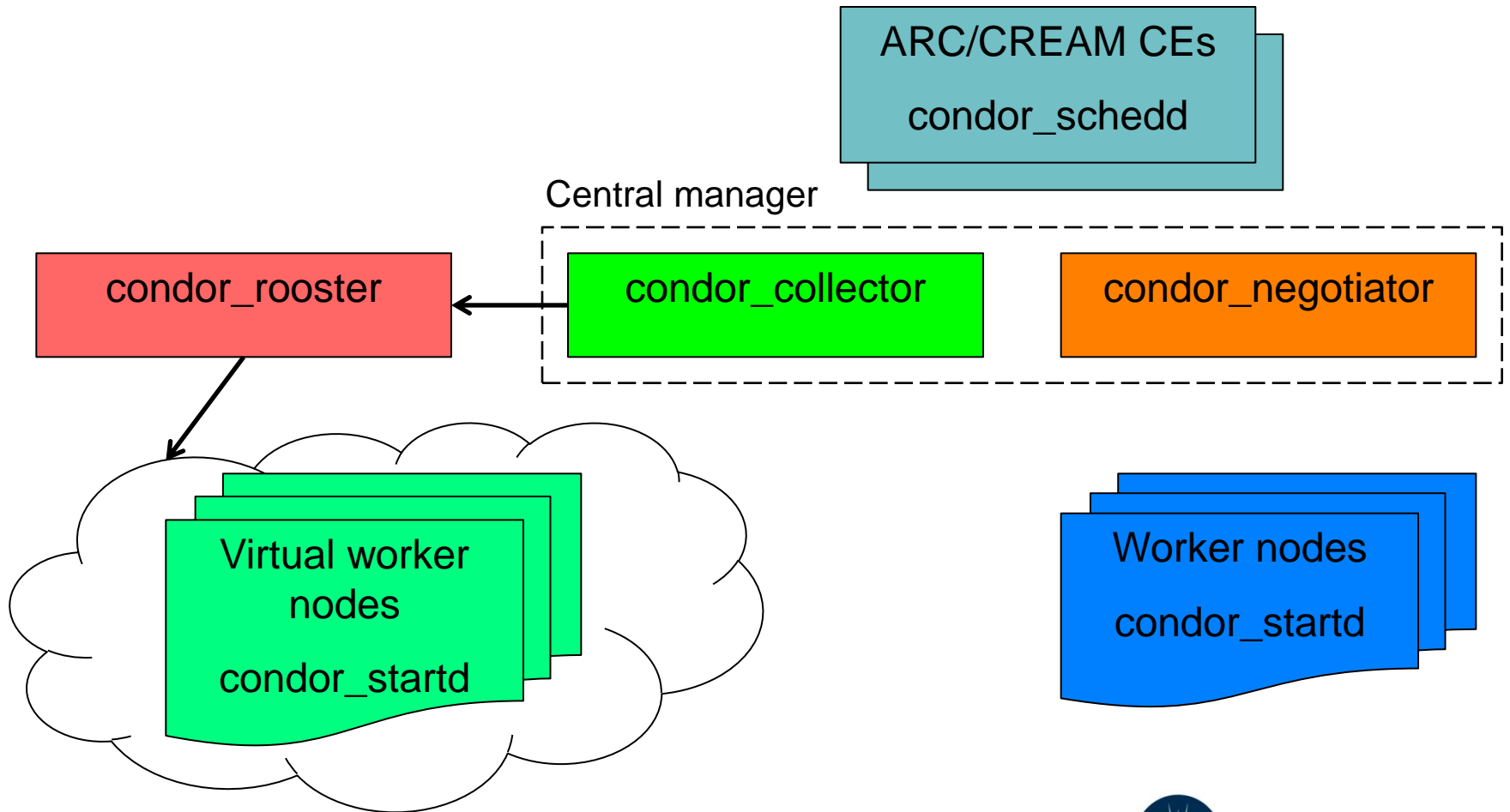


Using HTCondor

- When there are idle jobs
 - Negotiator can match jobs to the offline ClassAd
 - Configured so that online machines are preferred to offline
 - condor_rooster daemon notices this match
 - Instantiates a VM
 - Image used is a setup as a worker node, with HTCondor installed
 - VM starts up & joins the HTCondor pool



Provisioning worker nodes



VM lifetime

- Using short-lived VMs
 - Only accept jobs for a limited time period before shutting down
- HTCondor on the worker node controls everything
 - START expression
 - New jobs allowed to start only for a limited time period since the VM was instantiated
 - New jobs allowed to start only if the VM is healthy
 - Startd cron (healthcheck)
 - HIBERNATE expression
 - VM is shutdown after machine has been idle for too long

