Virtual Pools for Interactive Analysis and Software Development through an Integrated Cloud Environment

Claudio Grandi (INFN Bologna)
A. Italiano and D. Salomoni (INFN-CNAF)
Introduction

• Most sites that offer remote access also have local users that need:
  – interactive access
  – local access to the batch system
  – user interfaces to access the distributed infrastructure

• A classic model would be to statically allocate machines for the interactive access
  – These machines would have local access to the batch system and user interface software
  – In the worst case if special configurations are needed for the batch nodes used locally also the local batch farm nodes would need to be allocated statically

• We describe here a solution based on virtualization to dynamically allocate resources to different tasks regardless of the needed configuration
Classical model

- Bastion
- Desktop
- Interactive login
- Interactive pool
- Local submission
- Local batch farm
- Grid worker nodes
- Grid submission
- CE
- /opt
- /home
- /data
- SE
Idle slots are not allocated to any specific task
Description of the project

- Build on the experience of the WNoDeS (Worker Nodes on Demand Service) project
  - Instantiate virtual machines (VM) with the desired profile on demand
    - Including VO specific customization, e.g. file system mounts, etc...
  - The batch system steers resources allocation
  - In production at the INFN-CNAF Tier-1 with 2000 VMs

- A specific client is installed on the local desktops and bastion hosts
  - Instantiate the VM with the desired number of cores and memory
    - Throttle bandwidth to avoid abuse by individual users
  - Open an interactive shell on that VM

- When exiting the machine is destroyed

- Tests at INFN-CNAF on the INFN-Bologna Tier-3
• Joint project of INFN and Università di Bologna
  – Supports local research groups including LHC experiments
• Integrated with the INFN-CNAF Tier-1 infrastructure
  – ~ 50 dual quad-core boxes accessible through CNAF-LSF
  – 150 TB on CNAF-GPFS
  – The boundaries of the Tier-1 and Tier-3 are virtual
    • Each site can expand into the other according to policies in LSF
    • The profile of the virtual machine defines to which site it belongs
  – Computing and Storage Elements are independent
• Differently from the Tier-1 also offers interactive services to local users
• Currently in a prototype phase
WNoDeS

Batch job execution on real or virtual machines instantiated on demand

Grid jobs executed on virtual machines instantiated using the runtime environment required in the job JDL (CREAM CE)

Virtual Interactive Pools (VIP)

Interactive access to virtual machines instantiated on demand

Batch Job Execution

On demand computing via OCCI API and AuthN + AuthZ layer

Cloud Computing

Grid Integration
Virtual Interactive Pools

The VIP command line interface is build on top of the WNoDeS cloud computing interface

- The user requests the instantiation of a VM via the VIP CLI
  - The CLI allows to specify the requirements (RAM, CPU, bandwidth, mount points)
  - It is possible to wait until the login is possible or to get back a request ID that may be used to check the status of the request and eventually login on the virtualized resource
- The request goes through the authN and authZ layer of WNoDeS
- WNoDeS submits a fake batch job with the specified requirements
- The batch job is dispatched on the WNoDeS *bait host*
  - The bait host runs a process that “attracts” jobs
- On the bait host the batch job execution is immediately suspended and the process to instantiate the VM starts
- When the VM is ready and reachable via ssh, WNoDeS customizes the machine, e.g. it creates the mount points
- The VM is now ready to be used and the associated batch job is resumed
- At logout WNoDeS automatically kills the batch job and destroys the VM
- Wall clock time spent by the bait job is used for accounting purposes
Starting a session...

X application running on the VM

image name

number of cores

memory

mount points
...all files I need are there...

- my home directory
- my files
- Experiment software
- AFS
- cms data
- /home fs
...it looks like what I asked for!

Two cores

1500 GB RAM
Outlook

• Address the use case of a user needing to save the VM status
  – Currently each time the user starts an interactive session gets a fresh virtual machine
    • The only persistent data is on the mounted file systems
  – The use case may be useful e.g. to test services
  – In general it requires root access
  – A snapshot of the VM is saved on the storage area before destroying it. The saved image may be requested via the VIP CLI

• Package the client for easy installation also on personal workstations

• Optimize the resource allocation time
  – Access to an already instantiated machine with default configuration
    • Fast – only 1 or 2 machines available for each VO
  – Instantiation of a machine with custom configuration on reserved slots
    • Less than 1’. Defined number of slots reserved for that scope for all VOs
  – Instantiation of a machine out of the standard batch pool
Web site: http://web.infn.it/wnodes

Mailing list: wnodes@lists.infn.it

At CHEP 2010:

– A. Italiano *WNoDeS, a tool for integrated Grid/Cloud access and computing farm virtualization*
  Tue 19th, 16:45; Computing Fabrics and Networking Technologies

– D. Salomoni *An Authentication Gateway for integrated Grid and Cloud access*
  Wed 20th, 12:30; Grid and Cloud Middleware