HEPiX Virtualisation working group

Andrea Chierici
INFN-CNAF
Workshop CCR 2010

Outline

- Introduction
- Working assumptions
- Sub-tasks description
- References and links

Introduction

- The objective is to enable virtual machine images created at one site to be used at other HEPiX (and WLGC) sites.
- Agreement required on how images
  - are generated securely and with traceability
  - are transmitted securely and efficiently
  - can be expired/revoked if necessary
  - can be customised to meet individual site requirements (contextualisation)
  - can be used with different hypervisors

Working assumptions

- Images are generated by some authorized or trusted process
  - Some sites may accept “random” user generated images, but most won’t
  - No root access by end user during image generation
- Images are “contextualized” to connect to local site workload management system
  - But at least one site (other than CERN…) is interested in seeing images connect directly to experiment workload management system.
  - Recipient site controls how “payload” ends up in the image

Generation (1)

- The security-related policy requirements for the generation and endorsement of trusted virtual machine (VM) images for use on the Grid has already been defined
- The aim is to enable Grid Sites to trust and instantiate endorsed VM images that have been generated elsewhere.

Generation (2)

- **Endorser**:
  - Confirms that a particular VM complete image has been produced according to the requirements of the policy
  - States that the image can be trusted.
- An Endorser should be one of a limited number of authorised and trusted individuals appointed either by a VO or a Site.
  - The appointing VO or Site must assume responsibility for the actions of the Endorser and must ensure that he/she is aware of the requirements of the policy.

Virtual Machine Image Catalogue

- VMIC records details of virtual machine images distributed to the sites to be run on the site's local hardware.
  - Gives sites a central point of control over the images run at their site
  - Provides a mechanism to control (trust) who may subscribe images to the site, and to block images that are deprecated for some reason.

Endorsed vs approved

- Endorsed (endorser decision):
  - Role defined in the policy document
  - Scope: VMI production & maintenance
- Approved (site decision)
  - Marks the VMI “valid for use” by the site
  - Scope: operating the VMI
- For a VMI to run, it must be both:
  - Endorsed by an endorser (i.e. part of the VMIC endorsed)
  - Approved by the local site

Transmission

- Recommendation for basic transport protocol(s) to be supported
  - Prescriptive for sites wishing to generate images
- Current model is “tagged images” distributed in manner akin to mechanism used for VO software today
- Proposal for optional protocols to improve transmission efficiency
  - E.g. transmission of only differences w.r.t. a reference image
  - Interested in protocols such as bitTorrent
- Will not comment on intra-site image transmission

Expiry & Revocation

- Status a little unclear
- “Image Revocation List” à la CRL?
  - Technical proposal required
- Image endorses required to revoke images in case of security issues and the like

Contextualization

- Proposal for mechanism allowing site to configure image
  - File system mounted at image instantiation and automated invocation of scripts on the file system during the initialization.
  - Final job/payload will not execute as root
- Restrictions on aspects sites are allowed to configure
  - No changes to C compiler, Perl, Python, … to be allowed
- Contentious issue is kernel patching (not allowed)

Support for multiple hypervisors

- Recommendations/recipe(s) to enable sites to generate images that can be used with a range of hypervisors
  - Only KVM and XEN (both para and full)
  - Limited to sl5 for both host and guest
- Already tested extensively
  - KVM integration in sl5.4 helpful

Summary

- A year ago, sites were rejecting any possibility of running remotely generated virtual machine images.
- Today, we have the skeleton of a scheme that will enable sites to treat trusted VM images exactly as normal worker nodes.
  - This enables
    - VOs to be 100% sure of the worker node environment
    - (potentially) inclusion in the VM image of the pilot job framework enabling “cloud like” submission of work to sites.
- **Active involvement of VOs is now highly desirable** as we move towards delivering a proof-of-concept system.
- Nothing in what is being done
  - prevents sites that wish to do so from implementing Amazon EC2-style instantiation of user generated images, or
  - precludes use of CERNvm.

References and links

- Hepix-virtualisation@cern.ch
- VM generation policy draft
  - http://www.jspg.org/wiki/Policy_Trusted_Virtual_Machines
- Virtualisation workshop upcoming