Quattor@CERN

Véronique Lefébure
For CERN IT-FIO/FD

Quattor Workshop at CNAF Bologna
March 17-18 2008

Deployment status overview

- Quattor instances in production at CERN:
  - "Main instance": 7600 profiles
    - +1500 wrt October 2007
    - Of which ~1700 not-quattor-managed:
      - ~500 "standalone" machines (openlab, testbeds,…)
      - 470 OS not supported (solaris, windows)
      - 340 enclosures (twin systems)
      - 210 diskarrays
      - 60 diskshelves
    - In 140 clusters (+50 wrt March 2007)

Deployment status overview

- Quattor instances at CERN (cont'd):
  Linux for controls:
  - ATLAS, CMS, LHCb: 3 CDB's
    - Sitting in the experiments private network, used to configure the machines in the private network.
      - ATLAS: ~100 machines in their CDB, many of these are servers for diskless clients
      - CMS: 1070 nodes in their CDB
      - LHCb: ? nodes, diskless servers and diskless clients, making use of ncm-diskless_server.
  - + 1 shared, used by CMS and ATLAS
    - Based on Q1.1, using hierarchical package lists (comps2pan) and defaults templates.
    - Mainly for machines outside the private networks.
  - SWREP
      - Using the CC one for basic stuff, plus L4C for experiment specific software

Deployment status overview

- Desktops: *(Inputs from Jan Iven)*
  - CERN still uses a small subset of NCM components on desktops to configure side-wide defaults.

  →special requirements on the components - should be written in "cooperative" mode:
    - change only entries that are actually configured, ignore the rest. Do not remove comments. Do not remove options/entries not in the profile.
      - add a comment that these entries were changed by a tool (and that manual configuration may be pointless)
      - do not blindly overwrite config files, if nothing changes (as far as the component is concerned) = keep backup files and timestamps of the original modification intact.

Deployment status overview

- Desktops (cont'd):
  - CERN needs to target our default configuration from RPM better to machine actually on the CERN site (and not all SLC machines worldwide)
  - CERN is still using "lcm" (and not just ncm-ncd) since we need to track which components in the default profile have actually been selected (by the user) on a machine. CDB "active=true" is not sufficient for this.
  - still would love to see a hierarchical multi-source configuration:
    - look at local file (default, via RPM)
    - look at "default" in some "default" CDB
    - look at per-machine config in CDB
    - look at other local file (fine-grained local overrides - should replace /etc/nospam)
  - have written ncm-yum to preconfigure YUM (repositories, installed package groups, explicitly added RPMs) - not widely used.

Deployment status overview

- Main Instance Setup
  - CDB
    - Panc-v6 still used
      - Results on panc-v7 investigation: issue with the duplicate() functions. Fixed proposed by Cal but in v8. No time for testing that.
        - » Should CERN then skip v7 to go directly to v8 ? Timescale=?
    - Started to use Namespaces
      1. For Staging purpose (mid-October 2007):
         /prod, /preprod,/test
         - Successful, usefull, makes life much easier
      2. For organising templates (work in progress):
         1. /profiles, /repository
         2. /prod/pan, /prod/quattor
         3. /prod/components
         4. /prod/services
         5. …
         - Tedious and time consuming task

Véronique Lefébure
CERN IT-FIO/FD

Deployment status overview

- Main Instance Setup
  - CDB (cont'd)
    - Multi-core CPU setup: 4 cores
      - 33 minutes for recompiling 7600 profiles
      - Memory issues → 100 profiles per process
        - » Plan to go for either more powerfull hardware or multi-boxes architecture
        - » Maybe related to the big repository. To be checked.
  - SPMA and Swrep-soap
    - Implemented protected access to RHES repositories
    - Will most probably prepare new swrep for each VO
    - Plan to clean up the SWRep templates by adding an 'archive' modus as this is likely the cause for the slow panc performance and high memory consumption
    - Thinking of enforcing RPM signing (but are ALL rpms signed? Grid sw for ex.)

Deployment status overview

- Main Instance Setup
  - CDB2SQL and Oracle
    - CDB2SQL in the process of being improved:
      - Needs to be faster:
        - » Now: ~1 sec per profile → up to 2 hours for a full update !
        - » Plans for a x 3 improvement
    - CCM
      - ccm-2.0.2-1 deployed
      - SSL-based transfer of profiles partially deployed
    - “PrepareInstall”
      - Using YUM repository for base installation
      - Plan to use ccm to retreive profiles

Quattor development activities

- Conversion of CERN-CC templates to namespaces still to be done, planned for this year
- No more manpower at CERN available for developments (except CDB2SQL)

CERN-CC-specific activities

- Secure profiles transfer over SSL
  - Planned to be fully deployed this year
- Xen-based virtualisation
  - Used more and more, but needs a more flexible template structure
  - Still not fully automated
- More and more CDB users (eg. Vobox Application Mgrs)

  → Acl management issue
    - Namespaces /cluster will help. Currently:
      - 37 acl groups, /var/lib/cdb/auth/cdb.acls: ~9000 lines
    - Slow, no support for wildcards

  → Long apparent commit times because users queue
    - Request for more info before running the commit
    - Can we handle several sessions in parallel when it is knows that they don't interfere? Access to dependency info ?
- SMS (State Mgmt System) needs access to CDB acl's

CERN-CC-specific activities

- Update to Quattor 1.3 core templates:
  - Work in progress
  - performance issue with push and npush (massively used at CERN by the “package_default”)
  - structure_interface: **switchmedium** and **mtu**: CERN specific → include structure_annotation ?
- Plan for a CDB template structure convention
  - To allow broad usage of automating tools
  - Proposal: see next slide
- Reviewing how to better handle non-quattor-managed objects
  - Get rid of fake configuration data such as “kernelversion = ‘---’ “
  - Proposal: see slides
- Migration of service data from CDB to SDB

CERN-CC-specific activities

- Need for an agreed profile template structure
- Current proposal from the Fabric Service team:

```
object template profile_vocms01;
  define ELFMS_OS = "slc4";
  define ELFMS_ARCHITECTURE = "x86_64";
  define ELFMS_SVCCLASS = "vobox";
  define ELFMS_RESOURCE = "cms";
  define ELFMS_CUSTOMIZATION = "dbs";
# Only FS can edit
  include pro_hardware_xxxx;
  include netinfo_vocms01;
  include "pro_type_" + ELFMS_SVCCLASS;
# Application manager can also edit
  include "pro_custom_" + ELFMS_RESOURCE + "_" + ELFMS_CUSTOMIZATION;
  include "pro_custom_" + value("/system/hostname");
# Only FS can edit
  include "pro_resource_" + ELFMS_RESOURCE;
# The importance can only be set through the web-form
       "/system/importance" = 60;
```


Véronique Lefébure
CERN IT-FIO/FD

```
template pro_type_vobox;
  [..]
  "/system/cluster/name" =
       ELFMS_SVCCLASS + "_" + ELFMS_RESOURCE;
```

```
template pro_resource_cms;

  "/system/accounting/name" = ELFMS_RESOURCE;
```


Véronique Lefébure
CERN IT-FIO/FD

```
template pro_custom_cms_dbs;

# CMS ACLs,etc …
   "/system/cluster/usercontact" = "Lee.Lueking@cern.ch";
   "/system/rootmail" =
   value(split(",","/system/cluster/usercontact"));
   [..]
#specific system configuration
    include pro_filepartition_cms_dbs;
#specific packages
    include pro_service_php;
    include pro_service_grid_ui;
```

CERN-CC-specific activities

- Better handling of non-quattor-managed objects (like twin enclosures)
  - Get rid of things like:
    - `"/system/sitelease" = "---";`
    - `"/system/kernel/version" = "---";`
    - `"/system/filesystems" = nlist();`
    - `"/software/components/spma/active" = false ;`
    - `"/software/components/grub/active" = false ;`
    - `"/software/packages" = nlist();`
    - `"/software/repositories" = list();`
    - `"/system/network/domainname" = "cern.ch";`
    - `"/system/network/hostname"="";`
    - `"/system/network/interfaces" = nlist();`
    - `"/hardware/cards/nic/0/hwid" = "FF-FF-FF-FF-FF-FF";`

- Proposal:
  - Move from :

```
template pro_declaration_profile_base;

  include pro_declaration_all_types;
  include pro_declaration_functions;
  include pro_declaration_system_functions;

  type "/" = {
    "hardware" : hardware_quattor_type
    "software" : SOFTWARE
    "system"   : system_quattor_managed_type
  };
```

  - to :

```
template pro_declaration_profile_base;

  include pro_declaration_all_types;
  include pro_declaration_functions;
  include pro_declaration_system_functions;
```


Véronique Lefébure
CERN IT-FIO/FD

– and let the “/” be defined at the cluster level :

```
template configuration_level_full;

include pro_declaration_profile_type;

  type "/"  =  {
    "hardware" : hardware_quattor_type
    "software" : SOFTWARE
    "system"   : system_quattor_managed_type
  };
  "/system/quattor_managed" = "yes";
```

```
template configuration_level_nonetwork;

  type "/" = {
    "hardware" : hardware_quattor_type
    "system"   : system_no_network_type
  };
  "/system/quattor_managed" = "no";
```

That’s it

Véronique Lefébure
CERN IT-FIO/FD