Executive summary of the 1st Quattor for Grid workshop on 2004-03-26

Twnty-two (22) people attended the technical and planning meeting for the use of Quattor configuration and installation components for use outside CERN.   The purpose of this meeting was to plan the porting effort of LCG-2 from LCFGng to Quattor in order to evaluate whether Quattor meets the needs of the involved sites.

During this meeting, these high-level decisions were taken:
-   An implementation of Quattor configuration objects for installing and configuring Grid services and service nodes will be undertaken
-   There is interest both in complete system installation solutions (most of the sites) and in a "light" version that installs grid services incrementally
-   The effort will be directed towards getting LCG-2 installed as soon as possible
-   It will include OS installation based on incremental lists adding service groups. It will not support the original fine-grained RedHat installgroups. The aim is to distribute a script or tools to generate the lists for the base OS installation; it is not intended to provide basic OS RPM lists and updates. The latter is a site issue.
-   The people involved will make an initial port to evaluate Quattor. However, this effort does not imply either that these people will either support the the ported components or provide configuration components for new grid services after the evaluation period.

A complete rundown of existing LCFGng objects resulted in a list of Quattor components to be made, and based on this list a priority ordering was established.
-   Getting worker nodes to work outside of CERN (i.e. use of PBS)
-   getting a working CE + CE information providers if needed
-   BDII information system
-   Storage Element
-   Resource Broker
-   other node types

Eight sites indicated a desire to participate in the first Quattor-Grid test bed:

LAL Orsay, FR
NIKHEF, Amsterdam, NL
RAL, Didcot, Ox, UK
UAM, Madrid, SP
CNAF, Bologna, IT
CERN-FIO, CERN,CH
CC-IN2P3, Lyon, FR (affirmation pending)
IFCA, Cantabria, SP (affirmation pending)

Moreover, effort has been committed by some of these sites to help in porting the initial set of grid configuration object to Quattor in order to reach a the goal of a completely Quattor-installed LCG-2 'site'. Four people will work on the initial port, with support from at least two others working on the Quattor core components and AII. Marginal but skilled effort is available to re-validate the use of Quattor light for incremental installation of grid services.
The EGEE JRA1 external testing sites NIKHEF and RAL expressed the desire to run a single infrastructure both for the production facility and this JRA1 test bed.

Additionally, technical discussion during the meeting clarified for the implementers of the initial port how to tackle common problems and shared configuration. These technical decisions are documented in the detailed minutes of the workshop.

Minutes of the 1st Quattor-for-Grid workshop of March 26th, 2004

German's presentation, first steps:
DavidG's is going to take minutes

Introduction round with institue/project affiliation , and what commitments can you do in porting the grid components

Cal: LAL/Orsay/ EGEE, commitments: LCFGng configuration for the latest EDG release, and would to want to do the same for Quattor. Both system and grid components.

German: CERN/FIO, coordinator of the Quattor project and will later show CERN commitments. Quattor should get used in other institutes, and have it portable and see it being used, since that will ease our job in poirting grid components.

David Groep: NIKHEF, looking for successor of LCFGng for both fabric management and grid installation. Having spent effort in the past on LCFG objects, will spent similar efforts in getting components ported to Quattor.

Alberto di Meglio: EGEE JRA1 integration activity. Cannot make yet a commitment, but of course interested in what is happening.

Maite: EGEE JRA1 middleware and testing. Not decided yet what tool will be used for the EGEE test bed, but wants to understand how quattor will evolve. Depending on today may consider providing resources. Too early to make firm commitments; more known in a few weeks.

Enrico Ferro: INFN grid and Grid.it, Wants to understand what is going to happen with Quattor after EDG, cannot yet make decisions but will decide after this meeting whether and how to contribute.

Louis Poncet: LCG deployment and certification. Commitments can be various. Knows very well LCFGng stuff, and expects to be able to convince his group to migrate to Quattor. This WS should give the arguments.

Rafael Angel Leiva: UAM, task leader for Quattor integration team (==1 person). UAM wants Quattor for grid configuration, and has already done something and wants to see Quattor everywhere.

Gonzalo Merino: PIC/Barcelona. Grid activities at PIC, has been using LCFGng for test bed and will soon decide on what to use for automatic configuration for the production facility. Cannot yet take commitments, but expects to be able to go back to lab and organize local support for such future commitments.

Daniel Cano: sees it as a good tool but currently their cluster is a mix because both in LCG, EGEE, X#, and local. Is shopping.

Marian Zurek, CERN and Krakow ATLAS trigger/DAQ group: looking around for a new solution to replace local scripts. Cannot yet make commitments, and will not do grid.

Davide Salomoni: NIKHEF.

Yannick Perret: IN2P3 Lyon. Sysadmin for grid, currently been using LCFGng, and it's time to change.Wants to apply it asap to grid, and looking around to replace all fabric management in IN2P3. In coming month trying to include BQS in LCG2, and wants to include BQS in the configuration system to make sure automatic installation of BQS is there in Quattor.

Veronique Lefebure: CERN/FIO. Quattor CDB management at CERN.

Antonio Retico and Alessandro Usai: manual installation guide for LCG software, cannot yet take commitments, but very interested since Quattor may be main customers.

Zhechka Toteva: Technical student in CMS, have been deploying and testing quattor for CMS desktop usage Wants to get new ideas out of this meeting and see where it is going.

SteveT: RAL, EGEE+LCG. Need it for fabric mgt at RAL and wants one solution for LCG and own stuff+EGEE JRA1. Will contribute to achieve that.

Andrea C: INFN CNAF, Tier1, expecting to install O(1000) machines. Need automatic installation and knows limitations of LCFGng and sees that Quattor is the solution. Can be a test site and fix bugs. Not yet able to commit because of other problems and lack of manpower. Local lobbying will start.

Ignacio Reguero and Manuel Guijarro: CERN Solaris port, about to certify Solaris 9 and will be using Quattor to install both CC systems and desktops. Not yet grid components.

Jose del Peso: UAM, same group. Development side of UAM computing. Plan to continue development of Quattor and use it to install farms. Expects an agreement between labs to continue the development and translate LCFGng objects to Quattor.

Judith: CERN/FIO thesis student on Quattor. Will join EGEE as a staff member in SA1.

Cal will be the chair, DavidG taking notes. It is a community-driven project and *not* CERN imposed. Cal will report to the GDB via Kors or JohnG.

Objectives: technical meeting, target is replacing LCFGng for Grid services AND to replace complete fabric management.   First step in an evaluation if Quattor can achieve this.

There must not be forks, thus documentation and agreement on the global schema is required. Ad hoc conventions are not sufficient to prevent divergence.

CERN is running fabric with Quattor, although some remnants of SUE are left. SUE will be gone in CEL3, replaced by writing system components. These components will be site independent in terms of configuration.

There were >100 components in EDG, but CERN is not going to provide all (like "auth").
Commitment to components needed at end of meeting.

CERN/FIO position: wants a "virtuous circle". No forks or branches please.
[eof of presentation]

Rafael UAM: is adding those core components that are not needed at CERN, like the AII.

Cal: changes to the code – adopt model of a development+a stable platform. Or is everyone on the bleeding edge? German: there should be a stable version for production. Cal: having a production system requires a commitment to support this stable version for some time.
Cal: to avoid forks need one place to put the code. German has some ideas and will discuss later.

Andrea: CERN relies still on SUE, but why? German: SUE takes care of some service configuration like AFS or LSF. Ignatio: is some late legacy, should be replaces by the NCM components. German: will go away in CEL3.

Ignatio: plan to manage desktops as well, currently making crude PoC to enable local admins to modify templates a bit via GUI

German: also CERN desktops will use NCM, but no representative here.

German+Rafael: Quattor is highly modular and you can mix NCM and legacy solutions in any combination.

Marian: what is security model. German: uses HTTP currently, maybe community can port to HTTPS.

General feeling is the LCG way of doing this exposes sensitive data to the world. CERN has a time window in which the root data is exposed after installing a node to limit exposure of that data even on the local network. German has details. System then gets two security levels. HTTPS will be very useful. Rafael will send around a link, and there is some info on the Quattor mailing list.

Cal: first, testing is needed outside CERN on all the development sites. There should be consensus on which distribution to start from (EDG2.1 or LCG-2).
Louis: components should be service oriented.
Cal: we need to decide what it is we want to reach, i.e., should be do BDII.
DavidG: EGEE will be based on LCG-2, so starting with LCG-2 makes more sense.

DavideS: easier for developers to start with something that can be tested easily (find transition problems). Thus: LCG-2
German: LCG-2

Decision: aim for LCG-2 system.

Cal: for initial testing agree on platform. Been using RH73, so stick with it?
Maite: where has it been tested. German: RH73, Solaris. Currently cert process for CEL3/RHEL3, has to finish end of this year.
Cal: we need it for the next few weeks. Need end-to-end solution soon on one platform.
German: as Quattor come with all the components, there may be some timing issues when CERN is getting rid of SUE. The SUE replacement will be used for CEL3 and not RH73. The changes may be made backportable to RH73 (end of May). If you want to prove it works for one service type that can be done. For a PoC you don't need all the components.
Cal: priorities are needed, to get at least one working.

Decision: start with RH73

German: walk-though of lxb1001.cern.ch.dump.txt

- Repository configuration schema is going to be redesigned
- There is no relocation (yet), rpmt transactions?
- "system/ccdbname" is CERN-local
- partitions and mounts are direct input to kickstart, "options" should have been "kickstartoptions"

NCM usage hints:
  - a template processor is not the solution for everything, especially if the file is not under the components full control (like e.g. ld.so.conf) – so it is not there, but see NCM::Check::lines
  - documentation must be in a separate file .pod.cin
  - must provide PAN template to define the data types "RESPONSIBLE" and "VERSION" must be in
  - each component is a perl module with 2 call-back methods (Configure and Unconfigure)
  - $NoAction indicates a fake mode (should be respected)

  Documentation now in pod file – see whether it can also be integrated with the component and the makefile adapted to extract and test whether non-zero

  - cdispd not used at CERN

  Example ncm-rm component
  - see use of LC::Process::execute()

  spma component
  - unescape function should have been in a library

- updateconffile is a good example on how to update a conffile selectively using NCM::Check::lines
- but the exception handling is hacky in the example

Code conventions exist and used by Rafael. Written by Lionel Cons, to be circulated by German. German: problem is that there are only 1.5 pages of NCM guidelines (found in the reading list).

Cal: grid components overview

Who is going to maintain the components? It is not clear. EGEE is not going to propose a tool, but there is a testing group that will need to install the software. The people who do the porting now will likely not commit to maintaining them indefinitely. From the one side, the software developers of services could also write the component, but then they might make a tight coupling between the component and the service like they did in EDG. On the other hand, if we split the two, effort should be found to maintain these components.

Cal: we need to get Quattor to a point where we can install a system and have it running. No clear decisions on this can be taken here, since it is a techies meeting.

Maite: a "light" mode of Quattor should be tested as well, since that would ease testing by other remote sites based on this light mode.

Component overview:
WP1 stuff: mainly file writers. Symlinking stuff is a problem thay me be separated out

WP2 stuff: 2 components completely unused, 6 remain, rm is implemented, 5 remain:
- Java secutiry : messy, since it changes tomcat configuration.
Problem: some components are messy because several components modified same service. Is porting more messy than writing a clean thing from scratch? Since services will be modilar as well, maybe just do it messy to get it running quickly.
- Lrc,rmc,ros: template processing things.
- Templating will make it very easy, so who is going to make a template component?
- Voms admin: calls external script and a config file.

WP3:
- information providers: source is unfindable. Maite: may have been changed by LCG
- r-gma: config file writer using templates. Does nasty tomcat config and restarts a lot of daemons in a specific order.

WP4:
- most of the comps at a low priority; only a fewneeded:
- ceinfo: config writer, Maite: part is in the infoscript itself
- lcas, lcmaps: config file writer -> **DavidG**

WP5:
- not used in LCG

WP6:
- lots of little ones, some some already edgcfg, etc.
- mkgridmap: config file writers
- myproxy:
- pbsknownhosts: conplex
- pbsexechost: trivial
- voenv
- poolaccounts

WP7:
- not used

gmond: ganglia monitoring: **SteveT**
paconfig: ??? (WP1)

LCG-2 config:
- edg-lcfg-lcg-disk
- "shift"
- ??
- ??

Priorities from Cal:
1. Getting worker nodes working outside CERN from scratch (pbsexechost)
2. CE to get it to work together
   correlation between CE batch config and the worker nodes: German: multiple
   options exist to make the information of the worker nodes visible by the master
   config. Rafael: quattor's 'server modules' can be used, e.g. notification by the
   CDB on the master node itself whenever a new node joins.
3. ceinfo provider
4. information system (BDII)
5. SE (gridFTP server via glbous.cfg)
6. RB

Then you can start testing. We should get to that point! That's agreed!

Other types: VOMS, RC, etc.

Development starting order:
- template functionality -> **SteveT**
- pbsexechost -> **DavidG**
- cron -> replaced by RPMs in /etc/cron.d e.g. for CRLs

- ceinfo
- mkgridmap
- globuscfg

- ssh config

- seinfo – rest of SE already done by globuscfg

- all the WP1 scripts

Discussion on whether we want filecopy or dirperm should be on the mailing list.

Firewall ports to be opened: convention per service component to specify what ports need to be opened.

[LUNCH]

RPM lists for machine types also needed: Cal can do that.
Cal's Question: how to deal with underlying OS. Past: based on rpm comps file with the groups preserved. People could turn on/off groups. We could try to do the same with Quattor, but one problem arises: since Quattor only includes templates once, the list of groups should be defined before the rpm list is called (in LCFGng, the list can be iterated repeatedly with the new groups defined).
German: done at CERN without something that resembles that RPM groups file. CERN has a hierarchical solution: base, then interactive client list, a disk-server add-on list, etc. Thus a list of "standard" system is defined at the machine/cluster level (not based on services). As showed in demo.

No sites present actually used fine-grained RedHat groups.

Decision: the redhat installgroups need not be supported, but only tell which groups have been used to make a installset

It would be nice to have a script to generate the lists and some initial ports exist.
Sites may have a list of grid services only to feed into SPMA-light in user mode. Those sites should then provide some effort. No firm commitments on that yet.

Decision: Cal will make a base list and some additions (like NFS) to put on top. Grid lists exists for clients and services exist already and the scructure will be preserved.

Example: mkgridmap. How to concert to NCM.

[TEXT FROM CAL]

organization of the schema for edg-mkgridmap

```
/system/grid/vo/<name>/ldapuri          STRING
/system/grid/vo/<name>/user             STRING
/system/grid/vo/<name>/enable           BOOL
/system/grid/auth/<name>/ldapuri        STRING
/system/grid/auth/<name>/enable         BOOL
```

"allow" and "deny" pattern are local software configuration and are not in the global schema:

        /software/components/edg-mkgridmap/allow   STRING
        /software/components/edg-mkgridmap/deny   STRING
        /software/components/edg-mkgridmap/overwrite   BOOL

but check whether hyphens are allowed.
The grid-mapfile-local configuration will start as component-local configuration, and may later be moved out to the global part of the schema. Of course then there will be a need for backwards-compatibility and the component should look in both locations.

        /software/components/edg-mkgridmap/gmflocal   STRING
        /software/components/edg-mkgridmap/locals/   LIST
        /software/components/edg-mkgridmap/locals/<n>/subject   STRING
        /software/components/edg-mkgridmap/locals/<n>/user   STRING

        /software/components/edg-mkgridmap/refreshPeriodSec   LONG (unit: seconds)

Units must be added to the name of the resource using standard unit names
Capitalisation examples:
        refreshPeriodSec
        sizeMB
units are to be capitalised as are normal words.

The mkgridmap LCFGng component is a template processor. The syntax for the templates should be slightly different, especially the "for" loops. Needed constructs are "if" and "for".

Testing facilities and committed sites:

LAL   charles.loomis@cern.ch
NIKHEF   grid.sysadmin@nikhef.nl
RAL   S.m.Traylen@rl.ac.uk
UAM   angel.leiva@uam.es
CNAF Andrea Chierici <andrea.chierici@cnaf.infn.it>
CERN FIO (very likely) german.cancio@cern.ch
CC-IN2P3 (likely) yperret@in2p3.fr
IFCA (likely) cano@ifca.unican.es

Template generator	SteveT
RPM lists	Cal
PBS exechost	DavidG
Mkgridmap	SteveT

Ceinfo	
Ce locallogger	
BDII infosystem	
Seinfo	
Se	
SSH configuration	SteveT
RB	
Lcas	DavidG
lcmaps	DavidG
Dirperm	DavidG

This means we will first get a WN and CE and then see further. At least it must work outside CERN (like using PBS/Torque).
For the components that already exist, German can do some work to make them generic.

For Quattor "light" Maite is interested in it. May help test it if time/resources permit.

Cal: core Quattor software is useable, installguide needs to be improved a bit but it is not blocking.

Repository for new code development.
Options:
1. use Marianne since IN2P3 is willing to keep it open for Quattor
2. use CERN CVS, but you need a CERN account
   since everyone except for Yannick Perret has one already. But he can get it.
This should be resolved soon.

Then, access to the tree at CERN is managed using AFS ACLs, and then we don't need a split. **German has to check with Manuel**.

Bugs: Marianne is closed, so we use Savannah. Savanah repository has already been setup, **German will set it up again and enable bugzilla**. Louis will help there.
Submitting anonymous bugs to anonymous owners need to be checked in the new version of bugzilla.

Mailing lists. If it's really Quattor core, use project-quattor@cern.ch. If its related to the grid components ot testing and support, use **hep-project-quattor-grid@cern.ch**
The latter is also for the announcements.

If questions are asked on the Quattor-grid list, do not by default expect German to answer it! It is self-help.

If ever a tutorial is needed, it will cost German a lot of work. Not unless a new need is identified.