SL(C) 5 Migration at CERN

Ulrich Schwickerath, Ricardo Silva
CERN IT, 1211 Geneve 23, Switzerland
E-mail: Ulrich.Schwickerath@cern.ch, Ricardo.Silva@cern.ch

Abstract. Most LCG sites are currently running on SL(C)4. However, this operating system is already rather old, and it is becoming difficult to get the required hardware drivers, to get the best out of recent hardware. A possible way out is the migration to SL(C)5 based systems where possible, in combination with virtualization methods. The former is typically possible for nodes where the software to run the services is available and tested, while the latter offers a possibility to make use of the new hardware platforms whilst maintaining operating system compatibility.

Since autumn 2008, CERN has offered public interactive and batch worker nodes for evaluation to the experiments. For the Grid environment, access is granted by a dedicated CEs.

The status of the evaluation, feedback received from the experiments and the status of the migration will be reviewed, and the status of virtualization of services at CERN will be reported. Beyond this, the migration to a new operating system also offers an excellent opportunity to upgrade the fabric infrastructure used to manage the servers.

1. Introduction
When taking the decision on which operating system should be the production one during the first phase of LHC operations, several things have to be taken into account. While the most important argument for the experiments is the stability of the computing infrastructure, resource providers also have to make sure that proper support can be provided, and that the available resources are optimally used and that access to those resources is provided in a secure way.

2. Why change ?
2.1. Context and life-cycle
Scientific Linux CERN [1] is a customized version of the Scientific Linux [2] which is based RedHat Enterprise server [3]. As of writing this document, Scientific Linux 4 (SL4) is the main operating system on all Grid sides. It is based on RedHat Enterprise Server 4. The Production Phase 1 for this operating system has ended in February 2009. Support for SL4 will end in October 2010. Due to the delay of LHC data taking, there is the risk that the first hot data needs to be processed on an operating system which is no longer supported.

The alternative is a switch to SL5, which is based on RedHat Enterprise Server 5. The end of production phase for this operating system has been announced for March 2011. The delay of LHC data taking offers an opportunity to switch to this newer operating system, allowing for a stable running during the first phase of LHC operations.
2.2. Hardware support

Recent hardware requires Linux drivers which happen to be present only in recent kernel versions. If these drivers are not provided by RedHat, they need to be back ported either by the Scientific Linux developers, or by the site Linux experts which is a time consuming activity. Moreover, with each update, it has to be repeated before the new patches can go into production. This causes additional delays, because the ports must be tested before they can be deployed into production, to avoid instabilities of the computing infrastructure.

An example for missing hardware support in SL(C)4 is the new memory architecture of the recently introduced Intel Nehalem processors. This new generation of processor can be used effectively only on recent Linux kernel version, like the one included in SL(C)5.

2.3. Software support

Software support is another field of concern. New features and some bug fixes are not available for SL(C)4. Recent compilers, specifically gcc4, offer better performance of experiment code. The system compiler for SL(C)5 is gcc4.1, a version of gcc4.3 is shipped with SL(C)5.3 and above. Within the architects forum, gcc 4.3 has been selected by the experiments as the compiler they want to use [4].

Virtualization is often mentioned as a solution to computing issues. CERN has started several projects related this technology, and is actively looking into possibilities to use virtualization both for batch worker nodes and critical services. While different virtualization technologies are being evaluated, a common feature is that they are all based on recent operating system versions for the hypervisors machines. As an example, KVM support is only available in recent RHE5 versions and derived operating systems, such as SLC5. For the guests it is planned to support at least all currently supported operating system versions. The exact outline of these project is beyond the scope of this paper.

3. Opportunities

The change of the base operating system is a major change in the computing infrastructure. CERN IT has to provide computing services in production quality which must not be interrupted. This makes it difficult to do major internal changes. Instead of doing a big-bang migration of all services, the new infrastructure builds up besides the existing production system. At a later stage, previously existing resources are migrated step by step to the new system.

In the case of CERN, the major differences between the SLC4 and the SLC5 installations are

- security enhancements
  - reduced number of software packages and daemons. The new policy is to install only what is really needed.
  - improved firewall rules. This is specifically important for interactive nodes which are exposed to a large number of users. Privileged access to batch worker nodes is restricted to connections from trusted machines
  - reduction of people with administrative privileges
  - use of SELinux by default on all machines
- improved box management
  - CERN uses the Quattor tool kit to manage boxes. Internal management has been moved to a different structure. The new schema is heavily based on name spaces, which allows to make better use of ACLs on templates.
  - Enforcing a more service oriented approach
  - 64bit operating system only

1 CERN is running the vast majority of its SLC4 resources already under an 64bit OS since long. This is not necessarily the case at other centers though, which is relevant for GRID applications.
• user management
  – retirement of kerberos 4
  – use of LDAP for user management

Some of these changes are significant and not necessarily transparent to the users. This is particularly true for the new security measures listed above.

4. Issues
The main issues reported by the users so far in the migration process have their origin in the enhanced security measures which are applied. Specifically, SELinux has caused several issues, mainly due to problems in the software stack which are being fixed by the developers. As a work around, the problematic SELinux rules have been identified and relaxed temporarily, to allow for backward compatibility during the transition phase. Also missing software packages were reported. When setting up the initial SLC5 configuration, CERN started from a minimal software configuration, to reduce the risk of system compromises because of buggy packages which are not needed anyway. Feedback from the experiments is essential here, and the software configuration is being extended on request of the users. Unexpected dependencies on rather old packages like CERNLIB were discovered as well. As CERNLIB is still in use by non-LHC experiments as well, an updated version has been provided.

Experiments require the ability to run 32bit SLC4 compiled code on 64bit SLC5. For this it is necessary to identify and install the corresponding 32bit compatibility libraries. Not all of them are provided upstream though. The way to ensure consistent package lists has been discussed at various places and occasions. Packages which are not provided upstream could be distributed with the grid middleware. The presences of all other packages could be ensured by providing an empty meta rpm [6], to be distributed with the worker node software, which contains all the required dependencies.

A major step in the migration is the switch of the alias to CERNs interactive login facilities from SLC4 to SLC5. This interactive cluster serves also as a user interface (UI) to the Grid [5]. The availability of a corresponding SLC5 version of the GRID UI is a prerequisite to this step, and it is delayed. It has been agreed with the experiments that the switch of the alias will only be done once the majority of resources in the Grid have been migrated to the new operating system. This prerequisite is considered necessary to protect users who compile their code on the interactive cluster, and then submit the resulting binaries for execution as a GRID job.

5. Status at CERN
CERN has decided to deploy all new CPU resources on SLC5. This way, the available resources under SLC5 have exceeded the available resources on SLC4 for the first time in early May 2009. The evolution of public computing resources at CERN, expressed in kSi2k, is shown in 1. In early May 2009, 2 LCG CEs as well as 2 CREAM CEs were available as gateways to the new resources. The migration of older machines from SLC4 to SLC5 has started in autumn 2009.

The interactive login cluster has been prepared for the alias switch as well. In early May, 18 dual Quadcore boxes are available behind the alias lx64slc5.cern.ch, and accessible for the users.

6. Summary
Switching the operating system from SLC4 to SLC5 is a necessary step to ensure stable running of computing resources during the first phase of LHC data taking. CERN has decided to perform a rolling migration, which allows users to migrate slowly from the old operating system to the new one, while verifying their results. At CERN, the migration is also used to modernize and harden the infrastructure which turned out to be a painful but necessary step.
Significant resources are available at CERN for use by all registered users, including the GRID users.

7. References

[1] http://linux.web.cern.ch/linux/scientific.shtml
[2] https://www.scientificlinux.org
[3] http://www.redhat.com/security/updates/errata/
[4] http://lcgapp.cern.ch/project/mgmt/af.html
[5] http://glite.web.cern.ch/glite/packages/R3.2/x86_64/updates.asp
[6] https://twiki.cern.ch/twiki/bin/view/LCG/SL5DependencyRPM