MANAGING THIRD-PARTY SOFTWARE FOR THE LCG

Eric Poinsignon, Andreas Pfeiffer, Shaun Ashby, Pere Mato Vila, Ioannis Papadopoulos, CERN, Geneva, Switzerland, Saul Youssef, BU¸ Yannick Patois, CNRS

Abstract

The External Software Service of the LHC Computing Grid Project (LCG) Software Process & Infrastructure (SPI) project provides open source and public domain packages required by the LCG projects and experiments. Presently, more than 50 libraries and tools are provided for a set of platforms that is decided by the architect forum. All packages are installed following a standard procedure and are documented on the web. A set of scripts has been developed to ease new installations.

In addition to providing these packages, a software configuration management "toolbox" is provided and a distribution script. The toolbox contains a coherent set of combinations package-version for each release of a project. The distribution script manages the dependencies of the LCG projects so that users can easily download and install a release of a project including its depended packages. Emphasis has been put on ease of use for the end-user.

In the next section we describe the "Context" of the provided services. If you already know the context, jump directly to the "TWO SPI Services".

CONTEXT

The two services : "External Software" and "Distribution Service" are part of the SPI Project located at CERN [1] for the LCG [2].

LHC Computing Grid Project

The world's largest and most powerful particle accelerator, the Large Hadron Collider (LHC), is being constructed at CERN, the European Organization for Nuclear Research, near Geneva on the border between France and Switzerland.

The accelerator will start operation in 2007 and will be used to answer some of physics most fundamental questions by some 6,000 people from universities and laboratories all around the world. The computational requirements of the experiments that will use the LHC are enormous: 12-14 PetaBytes of data will be generated each year, the equivalent of more than 20 million CDs. Analyzing this will require the equivalent of 70,000 of today's fastest PC processors.

The goal of the LCG project is to meet these unprecedented computing needs by deploying a worldwide computational grid service, integrating the capacity of scientific computing centres spread across Europe, America and Asia into a virtual computing organization.

Applications Area

The Applications Area is one of four activity areas in the LHC Computing Grid Project. The Applications area develops and maintains part of the physics applications software and associated infrastructure, that is shared among the LHC experiments. The scope includes common applications software infrastructure, frameworks, libraries, and tools; common applications such as simulation and analysis toolkits; grid interfaces to the experiments; and assisting the integration and adaptation of physics applications software in the grid environment. [3]

The work of the applications area is conducted within projects. As of August 2003 there are five active projects: software process and infrastructure (SPI), persistency framework (POOL and conditions database), core libraries and services (SEAL), physicist interface (PI), and simulation.

SPI project

The goal of the Software Process & Infrastructure (SPI) project [4] is to provide to the development projects of the LCG:

- basic environment for physics SW development
- general scientific libraries and class libraries
- software development tools
- documentation tools and document templates
- compiler expertise
- support activity necessary to ensure that a common grid-enabled environment is available at all grid sites.

The reason for the Software Process & Infrastructure project is to enable consistency and homogeneity in the development of the different packages of the LCG Application Area.

TWO SPI SERVICES

This work in SPI is in the fields of "General scientific libraries and class libraries " and "Software Development Tools"*. The two services provided and presented here are: "External Software Service" and "Distribution

* Funded by Particle Physics and Astronomy Research Council, Swindon, UK

Service" (Figure 1). They are presented at CHEP04 as two posters, reference #277.

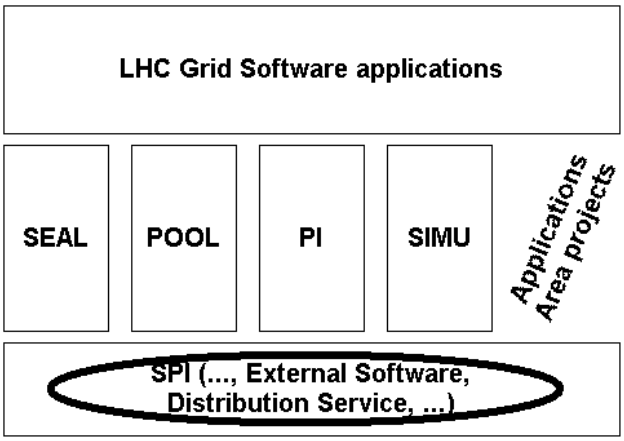


Figure 1: SPI location and presented services

External Software Service

The purpose of the External Software Service is to provide installations of open source and public domain packages, mainly software tools and libraries to LCG development teams. [5]

Its mandate was defined by the "Software Management Requirement Technical Assessment Groups (RTAG)" which recommended that External Software is used:

- All LCG projects must adopt the same set of tools, standards and procedures
- Adopt commonly used open-source or commercial software when easily available
- Avoid "do it yourself solutions"
- Avoid commercial software, if it gives licensing problems

In particular the service offers the following:

- Compilers
- HEP made packages
- Scientific libraries
- General tools
- Tests tools
- Database software
- Documentation generators
- XML parsers

More than 60 packages and tools are provided. Each of them is made available on a standard and common set of platforms. The list of platforms is decided by the Architect Forum. A unique name for each platform is created, composed of the system name and the associated compiler and its version. The 3 main platforms are Linux (RedHat and Scientific Linux for CERN), Windows and OSX. The current full list is:

- Rh73_gcc32
- Rh73_gcc323
- Rh73_icc80
- Rh73_ecc80
- Slc3_ia32_gcc323
- Slc3_ia64_gcc323
- Win32_vc71
- Osx10_gcc33

All the packages are installed in a directory tree at the AFS location: /afs/cern.ch/sw/lcg/external. They follow the same architecture of installation directories: <package name>/<version>/<platform>. As each package has a certain number of versions, each installed on the listed platforms, there is a total of over 450 installations.

All packages are compiled on CERN machines. The management of the installations is automated by a build process. The packages are built using pacman [7] and an internal python tool: spi_build_external. This last script follows XML configuration files describing the successive steps. This permits it to easily compile each package on any Unix platform (Figure 2 shows interaction between both programs).

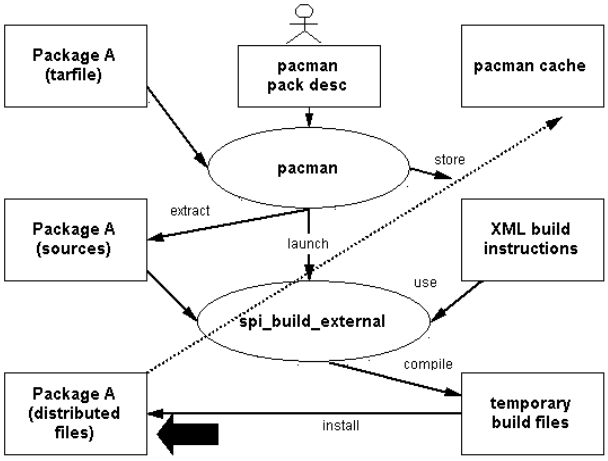


Figure 2: pacman and spi_build_external flowchart

All information concerning the External Software Service is documented and made available on the web [5]. Each package has a separate page with many information including Frequently Asked Questions (FAQ).

The External software Service is operational and its use is established in LHC experiments. In addition it is finding use in other projects and even outside of CERN. The European project EGEE is looking how to use it.

Distribution Service

The purpose of the Distribution Service is to provide distributions of the LCG projects deliverables (PI, POOL, SEAL, SIMULATION). [6]

Its deliverables are the combined set of software that includes the project software and the necessary dependant packages. These are tagged as dependant in the software configuration management files. Currently each package is distributed as a tarfile or an rpm file. The distribution script name is: lcg-installation-manager. It creates the

distribution files and can also download and install them locally (Figure 3).

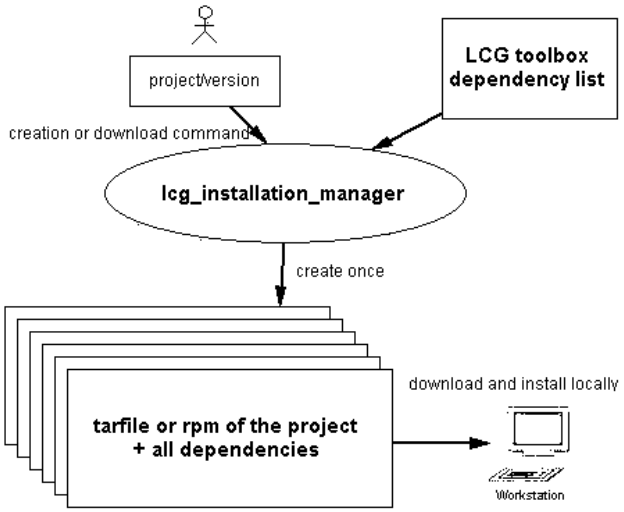


Figure 3: distribution script

As more and more configuration files appeared around, it was decided to centralize the different configuration information in a set of XML files. Then, any specific configuration file, like the scram toolbox, could be generated from those XML files. The purpose is to only maintain a centralized set of XML files (Figure 4).

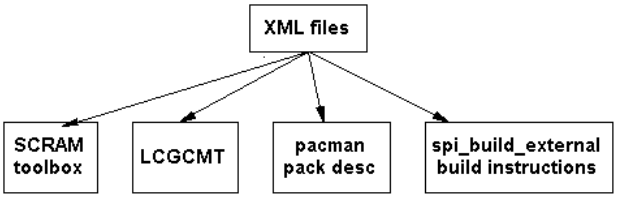


Figure 4: Dependence of configuration files.

Once all the distribution files are prepared, a download index page [6] show all the available projects (Figure 5).

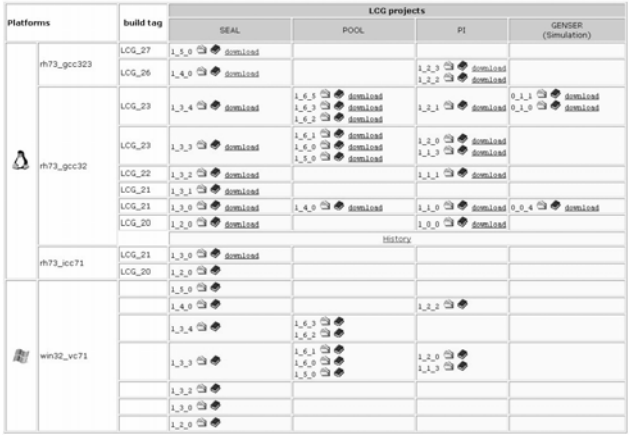
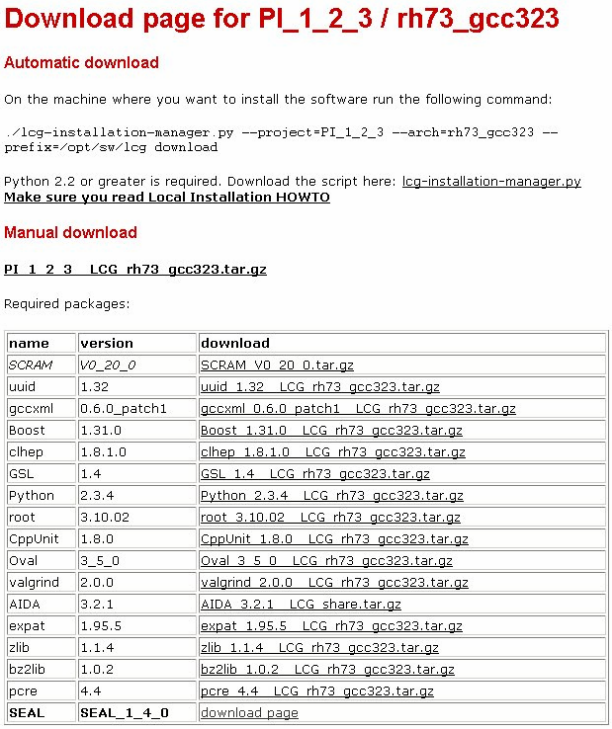


Figure 5: Download index page

For each “download” link, one can get a download page containing all the dependencies and how to get them in one row (Figure 6).



Download page for PI_1_2_3 / rh73_gcc323

Automatic download

On the machine where you want to install the software run the following command:

```
./lcg-installation-manager.py --project=PI_1_2_3 --arch=rh73_gcc323 --
prefix=/opt/sw/lcg download
```

Python 2.2 or greater is required. Download the script here: lcg-installation-manager.py
Make sure you read Local Installation HOWTO

Manual download

PI_1_2_3__LCG_rh73_gcc323.tar.gz

Required packages:

name	version	download
SCRAM	*V0_20_0*	SCRAM_V0_20_0.tar.gz
uuid	1.32	uuid_1.32__LCG_rh73_gcc323.tar.gz
gccxml	0.6.0_patch1	gccxml_0.6.0_patch1__LCG_rh73_gcc323.tar.gz
Boost	1.31.0	Boost_1.31.0__LCG_rh73_gcc323.tar.gz
clhep	1.8.1.0	clhep_1.8.1.0__LCG_rh73_gcc323.tar.gz
GSL	1.4	GSL_1.4__LCG_rh73_gcc323.tar.gz
Python	2.3.4	Python_2.3.4__LCG_rh73_gcc323.tar.gz
root	3.10.02	root_3.10.02__LCG_rh73_gcc323.tar.gz
CppUnit	1.8.0	CppUnit_1.8.0__LCG_rh73_gcc323.tar.gz
Oval	3_5_0	Oval_3_5_0__LCG_rh73_gcc323.tar.gz
valgrind	2.0.0	valgrind_2.0.0__LCG_rh73_gcc323.tar.gz
AIDA	3.2.1	AIDA_3.2.1__LCG_share.tar.gz
expat	1.95.5	expat_1.95.5__LCG_rh73_gcc323.tar.gz
zlib	1.1.4	zlib_1.1.4__LCG_rh73_gcc323.tar.gz
bz2lib	1.0.2	bz2lib_1.0.2__LCG_rh73_gcc323.tar.gz
pcre	4.4	pcre_4.4__LCG_rh73_gcc323.tar.gz
SEAL	**SEAL_1_4_0**	download page

Figure 6: Example of a download page

The download and local installation part of the script expands all the needed packages in the same architecture as defined under AFS: <package>/<version>/<platform>. The end user just has to define the root directory of his or her choice. A full installation is very easy.

Finally, the last step is to provide the distribution files for production as well as the distribution files for programmers.

WEB

The number of supported packages is large enough (more than 50) to require us to automate the web site. Shortly, the web will be automatically generated. The information will be stored in a regularly updated database and the web pages will be generated via PHP. [5]

ACKNOWLEDGEMENTS

Thanks to the many people in LCG, IT and PH experiments who participated to installations, configuration, script writing, web correction, advices, help, support, poster and articles.

REFERENCES

[1] CERN: http://www.cern.ch/
[2] LCG: http://cern.ch/LCG/
[3] Application Area: http://lcgapp.cern.ch/project/
[4] SPI: http://spi.cern.ch/
[5] External Software: http://spi.cern.ch/extsoft/
[6] Distribution Service: http://spi.cern.ch/lcgsoft/
[7] Pacman: http://physics.bu.edu/~youssef/pacman/