

# EDG 1.4 RPM package installation on DESY Linux 4 (S.u.S.E. 7.2) machines.

Andreas Gellrich<sup>1</sup>, Jacek Nowak<sup>2</sup>, Maxim Vorobiev<sup>1</sup>

<sup>1</sup> Deutsches Elektronen-Synchrotron (DESY), Notkestr. 85, 22607 Hamburg,  
Germany

*email:* [Andreas.Gellrich,Maxim.Vorobiev]@desy.de

<sup>2</sup> Institute of Computer Science, University of Mining and Metallurgy AGH, 30-059  
Kraków, Poland

*email:* jacek\_nowak@autograf.pl

## Abstract

The workstations at DESY (Deutsches Elektronen Synchrotron) run DESY Linux 4. It is a customized version of S.u.S.E Linux 7.2. It uses AFS user accounts and also most of DESY libraries and applications are shared through AFS. This customized Linux distribution has been used as the base for an EU DataGrid (EDG) 1.4 testbed installation. The installation was performed based on the manual installation from the "EDG installation guide". Binary RPM packages from the official EDG 1.4 distribution were used. As the only officially supported platform for EDG 1.4 is RedHat Linux 6.2, modifications to the installation procedure are necessary. The testbed at DESY proves that it is possible to install all main Grid nodes (WN, CE, SE, RB, BDII, RC, UI) on S.u.S.E. Linux 7.2 and not only on the official platform.

In this paper we would like to present the main issues which arose when trying to install EDG 1.4 on DESY Linux 4.

## 1 Introduction

DESY is a High Energy Physics laboratory located in Hamburg, Germany [5]. Four international experiments have been created to build and operate the HERA accelerator. Currently 700 scientists of 31 countries are involved. DESY has started a Grid testbed project in spring 2003 [7]. Its main purpose is to study the installation and operation of a Grid based on EDG middleware.

The testbed was installed using standard DESY Linux machines, which run a Linux operating system based on S.u.S.E 7.2. One possible application after finishing the testbed installation is Monte Carlo production for the HERA experiments. EDG (European Data Grid) version 1.4 [1], [2] has been chosen as the software base for this installation. To simplify installation and upgrade procedures we selected an EDG distribution based on RPM binary packages.

This selection implies the main problems we had to face during the installation. The only officially supported platform for EDG 1.4 is RedHat 6.2 and the RPM packages are designed for that system. However the installation was

completed without any changes to our standard DESY Linux installation and with only minor adjustments of the EDG packages.

In this paper we would like to present how to overcome the obstacles one has to face when installing EDG binary packages on a system based on S.u.S.E Linux 7.2. We believe that Grid testbed administrators who are trying to install EDG on platforms other than RedHat 6.2, can also benefit from our experience. This is an overview of the installation, detailed step-by-step instructions for each Grid node can be found in our installation guide [8].

The paper is organized as follows. In chapter 2 we will present the main features of DESY Linux 4 and compare it to RedHat 6.2. The installation process of the testbed will be described in chapter 3. Chapter 4 describes the future plans of the DESY Grid Testbed project.

## 2 System specification (RedHat 6.2 to DESY Linux 4 comparison)

DESY Linux 4 [9], [10] is currently the standard Linux distribution at DESY. It is used on desktops as well as on server machines. This distribution is based on S.u.S.E Linux 7.2 professional [6] and customized to the specific needs of computing at DESY. An automated installation and configuration system exists which allows for an easy setup of new DESY Linux 4 machines and a quick integration with the existing computing and security infrastructure. This includes AFS user accounts, and AFS sharing of applications and libraries common for all machines. Only about 250 of the almost 2000 S.u.S.E packages are installed locally on the PC's. The majority of GNU, HEP and commercial software is provided, maintained and updated synchronously in a central AFS repository. S.u.S.E Linux 7.2 which is used as the base for DESY Linux 4, features kernel version 2.4.18, glibc version 2.2.2 and RPM version 3.0.6.

RedHat 6.2 is a rather old but well tested and stable Linux distribution. EDG uses a version with kernel 2.2.24 and glibc 2.1.3. The EDG 1.4 binary packages are created with RPM version 3.0.5.

The differences in kernel and glibc versions, different locations or lack of many configuration scripts in DESY Linux 4 as compared to RedHat 6.2, were the cause of our doubts if it is possible to install EDG binary distribution on our system. On the other hand installing RedHat 6.2 on these machines would require a lot of extra work in order to integrate them with the existing system.

After some discussion we decided that the advantages DESY Linux has for us, make it worth to try to install EDG on an unmodified DESY Linux machine.

## 3 Installation

The installation procedure was based on the official "*EDG 1.4 Installation Guide*" [4]. In this manual two ways of installation are described. The first one is an automated installation procedure which uses LCFG [11] the second one is manual installation. As LCFG installs a whole system from scratch, which is

RedHat 6.2 by default, it is not an option in our case. However there exists a modified version of this system called LCFG Lite, which has been designed to make installation of EDG on top of an existing system possible. Using LCFG Lite on DESY testbed is still being considered, but for the first installation manual instructions were chosen as they give greatest control of what is going on.

### 3.1 Installing the packages

The EDG RPM packages can be downloaded from the package repository [3]. They are grouped by node type (CE, WN, RB etc.) and by software package group (CA, Globus, EDG etc.). The software groups which were used in the installation are of DESY testbed: CA (Certification Authority), Globus, EDG, external and RedHat 6.2.

The CA group contains platform independent packages only and can be installed with a `'rpm -ivh *'` command. Also packages from the Globus group do not cause any problems and install smoothly. However, applying the same command to EDG packages, results in a long list of dependency errors. The exact list of package dependency errors for each Grid node can be found in the installation manual [8].

In the following sections all dependency errors are described and solutions to them are explained.

**Errors which can be safely ignored** A closer look at the dependency errors list has shown that some of them were reported wrongly. For example when installing packages for the RB node we got following errors:

```
libcrypto.so.0 is needed by openldap-ftree-2.0.18ft1.0.0pre3-1
libssl.so.0 is needed by openldap-ftree-2.0.18ft1.0.0pre3-1
```

although `libcrypto.so.0` and `libssl.so.0` are installed on the system. The reason for this kind of errors is that not all S.u.S.E RPM packages are organized in the same way as RedHat packages. Such errors can be safely ignored.

**Perl packages** A difference in format of version string was the reason of another problem. A perl 5.6.0 package comes with DESY Linux 4, still when trying to install some packages we get this error message: `perl >= 5.00503 is needed by edg-mkgridmap-1.0.9-2` This is because the string `'5.00503'` has only one dot and `'5.6.0'` has two. The greater or equal condition fails as 00503 equals 503 and is greater than 6. This dependency problem can be ignored if the package concerned is not a perl module. In case of a perl module, ignoring this problem would result in files being placed in wrong directories - `'perl5/5.00503'` directories instead of `'perl5/5.0.6'` directories - on our system. This can be solved by adding these directories to the perl module path or by rebuilding source RPM's with a modified version string. We have chosen to rebuild the packages.

**AFS related problems** Some other packages, mainly perl modules, could not be installed locally on the Grid nodes as DESY Linux 4 has an AFS based perl installation. New modules can be added only by the AFS package administrator. This can take some time, and we decided to rebuild the perl modules in order to be able to install them locally. A description on how to rebuild perl packages can be found on our projects web page.

**Dependency on initscripts** On nearly all Grid nodes a dependency on the initscripts package was reported. It contains RedHat specific scripts and configuration files. These files and scripts are incompatible with their S.u.S.E equivalents and the package can not be installed on DESY Linux. Some workarounds are necessary to compensate this fact. They will be described later.

**Dependency on Egcs** Egcs is required by some of EDG packages. It conflicts with the standard gcc installation on DESY Linux. Installing it would require modifications to the system which are unacceptable. The aim of making egcs part of the EDG distribution is to have a standardized compiler on all grid nodes. This feature is not needed to have a working testbed so we just ignored this dependency and deal with the problem some time later.

**Other dependency problems** On some nodes standard DESY Linux packages had to be removed in order to be able to install their EDG equivalents. For example on the Replica Catalog node OpenLDAP package had to be removed.

All other dependencies of EDG packages could be resolved by installing either standard S.u.S.E packages or packages from the external or RedHat 6.2 group from the EDG repository.

## 3.2 Starting services

The next step after installing the packages is to start Grid services on all Grid nodes. The lack of initscripts package, AFS based nature of DESY Linux and small differences between system utilities on S.u.S.E and RedHat require some tricks to make everything work. This will be described in the following section.

**Solving the *initscripts* issue** Some scripts starting daemons like *globus-gatekeeper* use the script */etc/init.d/functions* from initscripts. The attempt to use the original *functions* script failed. In particular the method *daemon()* did not work correctly. Fortunately it could be replaced by the command:  
`/sbin/startproc -l logfile-name`

Another incompatibility was caused by the */sbin/initlog* script which is used both by the *daemon()* function to start a program and log its output and also to log single messages. By replacing *daemon* with */sbin/startproc*, *initlog* is no longer used to start programs but it is still in some places used to log single

messages. A simple workaround for this was to create a fake initlog script which is only capable of logging single messages.

The file `/etc/sysconfig/network` is also RedHat specific. It contains some basic information about network configuration. Below is an example.

```
NETWORKING=yes
HOSTNAME=grid007.desy.de
GATEWAY=131.169.223.16
GATEWAYDEV=eth0
NISDOMAIN=desy.de
```

As you can see it can be easily created for a given machine.

**condor\_master and the passwd file** A very disturbing thing happened when we tried to start the *condor\_master* daemon for the Resource Broker node - it crashed immediately causing segmentation fault. Further investigation has shown that it crashes while reading the `/etc/passwd` file. On DESY Linux machines this file contains AFS group entries. These entries start with '+@' characters. It seems that *condor\_master* daemon is not capable of reading such entries. The problem disappeared after we commented these entries out.

**AFS user accounts** AFS user accounts come with AFS home directories. Each Grid user has his `~/globus` file which is used by Grid services. When this file is placed on AFS however access to it is granted only to users which have a valid AFS token. Grid services when run by local users will have no access permission to this directory and will fail to do their job when invoked by an AFS user. The solution is to make `~/globus` a link to a local directory for each Grid user on all Grid nodes.

**Other problems** Some differences between S.u.S.E. 7.2 and RedHat 6.2 exist in the syntax of adding crontab entries and in the way services are added to *init*. So here the "EDG installation guide" can not be followed in a copy-and-paste manner.

## 4 Summary

### 4.1 Current status

As we have mentioned in the beginning the first reason for installing EDG 1.4 was to see if it is possible to install the system on DESY Linux machines. This goal has been accomplished All main grid nodes, including: CE, SE, RB, WN, RC, BDII, (for details see [7], have been installed. This was possible without any modifications to DESY Linux 4 and with only minor adjustments to the EDG binary distribution. The basic functionality of the Grid testbed has been tested successfully. No signs of binary incompatibility were found.

## 4.2 Future plans

With EDG 2.0 already available the natural next step will be to install this release. Much has changed since EDG 1.4 so new problems will certainly arise. Still we believe that our experience will be of great use and allow us to solve problems more quickly. We also hope to be finally able to test LCFG Lite as an automated way of testbed installation.

Currently it is being planned to port DESY HEP applications to the Grid. The first candidate is Monte Carlo production for DESY experiments. This will require to set up a new Grid enabled cluster of DESY Linux machines.

DESY will participate in the EDG successor EGEE. The Grid Testbed is the first attempt to establish a Grid infrastructure at DESY and will be the base for future Grid activities.

### Acknowledgements.

The DESY Grid Testbed project was initiated by the DESY IT Group in collaboration with the H1 experiment, represented by Ralf Gerhards.

We are very grateful to the H1 and HERA-B collaborations, especially to Alan Campbell and Brigit Lewendel, for supporting our testbed with Linux PC's. We would also like to thank Dave Kant from Queen Mary University of London for sharing with us his experience in EDG installation and administration.

## References

1. EDG Project Presentation, Mauro Draoli, Gianfranco Mascari, Roberto Puccinelli, 13/07/2001, [http://web.datagrid.cnr.it/introdocs/DataGrid-11-NOT-0103-1\\_1-Project\\_Presentation.pdf](http://web.datagrid.cnr.it/introdocs/DataGrid-11-NOT-0103-1_1-Project_Presentation.pdf)
2. The official EDG web page: <http://www.edg.org>
3. The EDG repository: <http://marianne.in2p3.fr/datagrid/repository>
4. EDG 1.4 Installation Guide, Integration Team (WP 6), January 17 2003, <http://marianne.in2p3.fr/datagrid/documentation/EDG-Installation-Guide.pdf>
5. The official DESY web page: <http://www.desy.de>
6. S.u.S.E Linux web page: <http://www.suse.de>
7. DESY Grid Testbed official web page: <http://www-it.desy.de/physics/projects/grid/>
8. DESY Grid Testbed installation instructions: <http://www-it.desy.de/physics/projects/grid/testbed/index.html>
9. After the First Five Years: Central Linux Support at DESY, Knut Woller, Thorsten Kleinwort, Peter Jung, CHEP 2001
10. Papers and Presentations on DESY Linux: <http://www-it.desy.de/systems/linux/papers/index.html>
11. EDG LCFG documentation web page: <http://datagrid.in2p3.fr/distribution/datagrid/wp4/edg-lcfg/documentation/>