



CROSSGRID INSTALLATION GUIDE

WP3.4 DATA ACCESS OPTIMIZATION

WP 3.4

Document Filename:	CG3.4-v1.0-CYF-DataAccInstallGuide.doc
Work package:	WP 3.2-DataAccOpt
Partner(s):	CYF
Lead Partner:	CYF
Config ID:	CG3.4-v1.0-CYF-DataAccInstallGuide
Document classification:	PUBLIC

Abstract:

This document describe briefly how to install UDAL software developed by WP3.4 within the CrossGrid project. The document is written at high technical level and requires advance knowledge about Linux operating systems as well as distributed data processing.

Delivery Slip

	Name	Partner	Date	Signature
From	Łukasz Dutka	CYF	Oct 2004	
Verified by				
Approved by				

Document Log

Version	Date	Summary of changes	Author
1-0	20/10/2004	Inception of the document	Łukasz Dutka
	26/01/2005	Verified by the QE	Robert Pajak

CONTENTS

1. COPYRIGHT NOTICE	4
2. ABOUT THE SOFTWARE	5
2.1. SOFTWARE COMPONENTS.....	5
2.2. DEPENDENCIES.....	6
3. INSTALLATION IN THE CROSSGRID TESTBED	7
3.1. RPM LISTS FOR LCFG.....	7
3.2. PROFILE MODIFICATIONS FOR LCFG	7
3.3. MANUAL POST INSTALLATION STEPS.....	7
4. MANUAL INSTALLATION	8
4.1. DOWNLOAD.....	8
4.2. INSTALLATION FROM RPM.....	8
4.3. INSTALLATION FROM SOURCE	8
4.4. CONFIGURATION.....	8
4.4.1. <i>List of configuration files</i>	8
4.4.2. <i>Editing the configuration files</i>	9
4.4.3. <i>Startup scripts</i>	9
4.4.4. <i>Other requirements</i>	9
5. RUNNING AND TESTING	10
5.1. LOG FILES	10
6. THE EDG LICENSE AGREEMENT	11

1. COPYRIGHT NOTICE

Copyright (c) 2004 by ACC Cyfronet AGH, Poland. All rights reserved.

Use of this product is subject to the terms and licenses stated in the EDG license agreement. Please refer to Chapter 6 for details.

This research is partly funded by the European Commission IST-2001-32243 Project CrossGrid.

2. ABOUT THE SOFTWARE

Unified Data Access Layer (UDAL) provided by task 3.4 CrossGrid project provides a flexible architecture for storage nodes and storage centres. UDAL allows for a very flexible optimisation and control of the internal behaviour access to data stored in heterogeneous devices . It is fully adaptable for future purposes, but currently is used only for organizing the estimation of data access latency and bandwidth of internal storage nodes. However, it is important to notice that the provided solution could be also directly used for other purposes. The current version of UDAL is distributed together with a set of specialized components for data access cost estimation for data stored in secondary and tertiary storage.

2.1. SOFTWARE COMPONENTS

UDAL comprises of the following packages:

- `cg-wp3.4-unidal-1.3.4.RH7.3-1.0.i386.rpm`
contains main software which should be installed on every Storage Element node. This package, apart from generic components of Unified Data Access Layer, contains set of data access cost estimation components for data access cost estimation to data stored in secondary and tertiary storage.
- `cg-wp3.4-unidalclient-1.3.4.RH7.3-1.0.i386.rpm`
contains a set of client tools used for verification of installation correctness. This set is supposed for administrator or advanced developers rather than ordinary users. It can be installed on any client node.

In addition to generic UDAL packages we were partially involved in adaptation of EDG replica optimization service to CrossGrid needs as well as to UDAL interfaces and the specially adopted versions are distributed in the following packages:

- `edg-replica-optimization-2.1.4.cg-3.noarch.rpm`
- `edg-replica-optimization-client-2.1.4.cg-3.noarch.rpm`
- `edg-replica-optimization-config-2.1.4.cg-3.noarch.rpm`
- `edg-replica-optimization-interface-2.1.0-1.noarch.rpm`

The figure below depicts generic deployment of UDAL in a grid infrastructure.

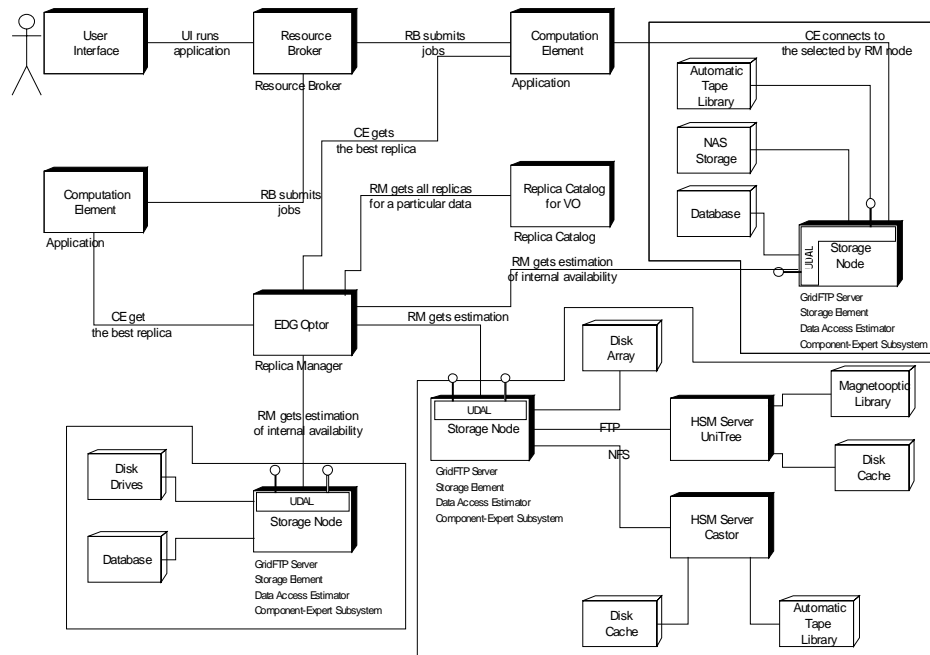


Fig. 1 UDAL Deployment Concept Diagram

2.2. DEPENDENCIES

UDAL is one of the lowest services in CrossGrid environment so it doesn't depend on other packages, except generic services provided by operating system and generic grid infrastructure. However, to exploit provided functionality it is required to install specially adopted version of EDG-Optor service.

3. INSTALLATION IN THE CROSSGRID TESTBED

The CrossGrid testbeds are managed by the LCFG deployment support tool. This tool allows an automatic installation of the software on all required nodes.

3.1. RPM LISTS FOR LCFG

Storage Element:

`cg-wp3.4-unidal-1.3.4.RH7.3-1.0.i386.rpm`

3.2. PROFILE MODIFICATIONS FOR LCFG

None

3.3. MANUAL POST INSTALLATION STEPS

The default configuration should be good for most cases but if a specific tertiary storage is installed then modification of this file `/opt/cg/etc/stelStor.conf` might be required. It affects only those SEs, which have tertiary storages installed in.

4. MANUAL INSTALLATION

4.1. DOWNLOAD

The RPMs of UDAL are available at:

- <https://savannah.fzk.de/distribution/crossgrid/autobuilt/i386-rh7.3-gcc2.95.2/wp3/RPMS/cg-wp3.4-unidal-1.3.4.RH7.3-1.0.i386.rpm>
- <https://savannah.fzk.de/distribution/crossgrid/autobuilt/i386-rh7.3-gcc2.95.2/wp3/RPMS/cg-wp3.4-unidalclient-1.3.4.RH7.3-1.0.i386.rpm>

4.2. INSTALLATION FROM RPM

The package `cg-wp3.4-unidal-1.3.4.RH7.3-1.0.i386.rpm` must be installed on every Storage Element (SE).

For testing purpose `cg-wp3.4-unidal-1.3.4.RH7.3-1.0.i386.rpm` can be installed on other machine but in most cases it is not necessary.

Both rpms install binaries and configuration files under `/opt/cg` directory.

4.3. INSTALLATION FROM SOURCE

1. Make sure you are working with RedHat Linux 7.3. UDAL was developed and tested for that operating system but it should work properly with other ones.
2. Make sure you have `gcc 2.95`.
3. Get the source code of UDAL available at the following URL:
 - <https://savannah.fzk.de/distribution/crossgrid/autobuilt/i386-rh7.3-gcc2.95.2/wp3/SRPMS/cg-wp3.4-unidal-1.3.4.RH7.3-1.0.src.rpm>
 - or tarball
<https://savannah.fzk.de/distribution/crossgrid/autobuilt/i386-rh7.3-gcc2.95.2/wp3/SOURCES/cg-wp3.4-unidal-1.3.4.RH7.3.tar.gz>
4. Unpack and go to the directory of the sources
5. Run `configure` command.
6. Run `make install` command.
7. To build your own rpm please run `make rpm` command.

4.4. CONFIGURATION

4.4.1. List of configuration files

UDAL installed on SE has a few configuration files which in most cases do not need to be modified.

- `daes.conf` – this is a configuration file for Data Access Estimator (DAES) service. This demon provides SOAP interface for Optor and other clients. The configuration file allows adjusting path, debug level and TCP ports, for the DAES server as well as for CEXS interface, which DAES is client to.
- `cexs.conf` – this is a configuration for Component Expert Subsystem (CEXS) service. CEXS is a demon selecting best CEComponent in the specific context. This file allows configuring paths, debug level and TCP ports.

- `compsCont.conf` – this is configuration of component registry used by CEXS. And the file shouldn't be modified.
- `stel.conf` – this is a configuration file for STEL demon provided information about current storage configuration. This configuration file adjusts TCP ports and important paths.
- `stelStor.conf` - this configuration file contains information about current storage system configuration, type of storages and mounting points of devices. It should be modified if the internal configuration of storage is more complicated then ordinary secondary storage device.

4.4.2. Editing the configuration files

Normally there is no need to modify anything after UDAL installation. Users can modify logging level, path or specific ports of demons. However, it is important to remember that if the TCP port is changed of a specific demon it must be done in all configuration files: `daes.conf`, `stel.conf`, `cexs.conf`.

4.4.3. Startup scripts

UDAL can be started, stopped or restarted by `/opt/cg/etc/init.d/cg-unidal` script. The parameters `start|stop|restart` are allowed.

4.4.4. Other requirements

4.4.4.1. Environment

None

4.4.4.2. Users

None

4.4.4.3. Ports

TCP port 18001 must be accessible to DAES client outside SE. But in most cases EDG-Optor is the only client, so it can be accessible for that machine.

7.1.1.1. Certificates

N/A

4.4.4.4. Folders

UDAL in standard configuration is installed under `/cg/opt`. Since it logs file are stored in `/cg/opt/var/log` it is important that `/cg/opt/var/log` directory exists. Otherwise the service will not started at all.

5. RUNNING AND TESTING

1. Make sure the demons are running or run `cg-unidal start` command in case they aren't.
2. Ask infrastructure for data access cost estimation for non-existing file located in your SE, where UDAL is installed in. The session should look similar to the one below.

```
[bin]$ cg-daes-client -p 18001 estDAB file:///etc/non-existingfile
DataAccessBandwidth for file:///etc/non-existingfile1 is -1
```

3. Ask for estimation cost to file existing in your Storage Element. The session should look similar to the one below:

```
[bin]$ cg-daes-client -p 18001 estDAB file:///etc/file1
DataAccessBandwidth for file:///etc/file1 is 3
```

5.1. LOG FILES

Unified Data Access Layer provides a wide logging facility. The logging precision can be adjusted by modified log level in `cexs.conf`, `stel.conf` and `daes.conf` configuration files. Logs are stored into `/opt/cg/var/log` directory. The path can be modified as well.

Running UDAL produces the following set of files:

- `cexs.pid` – Process ID of CEXS demon.
- `cexs.std` – captured standard output of CEXS demon.
- `cexs_expert.log` – log produced by built-in expert system.
- `cexs_kernel.log` – log produced by CEXS demon.
- `daes.log` – log produced by DAES demon.
- `daes.pid` – process ID of DAES demon.
- `daes.std` – captured standard output of DAES demon.
- `stel.log` – log produced by STEL.
- `stel.pid` – process ID of STEL demon.
- `stel.std` – captured standard output of STEL.

6. THE EDG LICENSE AGREEMENT

Copyright (c) 2005 CrossGrid. All rights reserved.

This software includes voluntary contributions made to the CrossGrid Project. For more information on CrossGrid, please see <http://www.eu-crossgrid.org>.

Installation, use, reproduction, display, modification and redistribution of this software, with or without modification, in source and binary forms, are permitted. Any exercise of rights under this license by you or your sub-licensees is subject to the following conditions:

1. Redistributions of this software, with or without modification, must reproduce the above copyright notice and the above license statement as well as this list of conditions, in the software, the user documentation and any other materials provided with the software.

2. The user documentation, if any, included with a redistribution, must include the following notice: “This product includes software developed by the CrossGrid Project (<http://www.eu-crossgrid.org>).”

Alternatively, if that is where third-party acknowledgments normally appear, this acknowledgment must be reproduced in the software itself.

3. The names “CrossGrid” and “CG” may not be used to endorse or promote software, or products derived therefrom, except with prior written permission by cgooffice@cyfronet.krakow.pl.

4. You are under no obligation to provide anyone with any bug fixes, patches, upgrades or other modifications, enhancements or derivatives of the features, functionality or performance of this software that you may develop. However, if you publish or distribute your modifications, enhancements or derivative works without contemporaneously requiring users to enter into a separate written license agreement, then you are deemed to have granted participants in the CrossGrid Project a worldwide, non-exclusive, royalty-free, perpetual license to install, use, reproduce, display, modify, redistribute and sub-license your modifications, enhancements or derivative works, whether in binary or source code form, under the license conditions stated in this list of conditions.

5. DISCLAIMER

THIS SOFTWARE IS PROVIDED BY THE CROSSGRID PROJECT AND CONTRIBUTORS “AS IS” AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY, OF SATISFACTORY QUALITY, AND FITNESS FOR A PARTICULAR PURPOSE OR USE ARE DISCLAIMED. THE CROSSGRID PROJECT AND CONTRIBUTORS MAKE NO REPRESENTATION THAT THE SOFTWARE, MODIFICATIONS, ENHANCEMENTS OR DERIVATIVE WORKS THEREOF, WILL NOT INFRINGE ANY PATENT, COPYRIGHT, TRADE SECRET OR OTHER PROPRIETARY RIGHT.

6. LIMITATION OF LIABILITY

CROSSGRID INSTALLATION GUIDE
WP3.4 Data Access Optimization

THE CROSSGRID PROJECT AND CONTRIBUTORS SHALL HAVE NO LIABILITY TO LICENSEE OR OTHER PERSONS FOR DIRECT, INDIRECT, SPECIAL, INCIDENTAL, CONSEQUENTIAL, EXEMPLARY, OR PUNITIVE DAMAGES OF ANY CHARACTER INCLUDING, WITHOUT LIMITATION, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES, LOSS OF USE, DATA OR PROFITS, OR BUSINESS INTERRUPTION, HOWEVER CAUSED AND ON ANY THEORY OF CONTRACT, WARRANTY, TORT (INCLUDING NEGLIGENCE), PRODUCT LIABILITY OR OTHERWISE, ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.