UNICORE Commandline Client: User Manual





UNICORE COMMANDLINE CLIENT: USER MANUAL

UNICORE Team

Document Version:1.0.0Component Version:8.0.0Date:18 03 2020

This work is co-funded by the EC EMI project under the FP7 Collaborative Projects Grant Agreement Nr. INFSO-RI-261611.

PDF BY DBLATEX

Contents

1	Ove	rview	1
2	Inst	allation and configuration	1
	2.1	Prerequisites	1
	2.2	Download	1
	2.3	Installation and configuration	2
	2.4	Preferences file	2
	2.5	Logging	3
	2.6	Installing UCC extensions	3
	2.7	Testing the installation	3
3	Gett	ting started with UCC	3
	3.1	Getting help	4
	3.2	Connecting	4
	3.3	List available sites	4
	3.4	Running your first job	4
	3.5	Listing your jobs	5
4	Con	nmon options to UCC	5
	4.1	User preferences	5
	4.2	Configuration file	6
	4.3	Credential and truststore options	7
	4.4	Trust store examples	12
	4.5	Support for oidc-agent	12
	4.6	Username and password	13
	4.7	Using Unity	13
	4.8	Client options	13
	4.9	Other options	16

5	Run	ning jobs	17
	5.1	Introduction	17
	5.2	Options overview	18
	5.3	Resource selection	18
	5.4	Processing jobs asynchronously	19
	5.5	Executing a command	20
6	Job	description format	20
	6.1	Overview	21
	6.2	Specifying the executable or application	23
	6.3	Arguments and Environment settings	23
	6.4	Application parameters	24
	6.5	Job data management	25
	6.6	Resources	29
	6.7	Miscellaneous options	30
7	Data	a management functions	31
7	Data 7.1	a management functions Specifying remote locations	31 31
7			
7	7.1	Specifying remote locations	31
7	7.1 7.2	Specifying remote locations	31 31
7	7.17.27.3	Specifying remote locations	313133
7 8	7.17.27.37.47.5	Specifying remote locations	 31 31 31 33 34
	7.17.27.37.47.5	Specifying remote locations	31 31 33 34 34
	 7.1 7.2 7.3 7.4 7.5 Met: 	Specifying remote locations	 31 31 33 34 34 35
	 7.1 7.2 7.3 7.4 7.5 Meta 8.1 8.2 	Specifying remote locations Data movement General commands Finding data Using the StorageFactory service	31 31 33 34 34 34 35 35
8	 7.1 7.2 7.3 7.4 7.5 Meta 8.1 8.2 	Specifying remote locations Data movement General commands Finding data Using the StorageFactory service adata management functions Basics Available commands	 31 31 33 34 34 35 36
8	 7.1 7.2 7.3 7.4 7.5 Meta 8.1 8.2 Wor 	Specifying remote locations Data movement General commands Finding data Using the StorageFactory service adata management functions Basics Available commands	 31 31 31 33 34 34 34 35 35 36 37
8	 7.1 7.2 7.3 7.4 7.5 Met: 8.1 8.2 Wor 9.1 	Specifying remote locations Data movement General commands Finding data Using the StorageFactory service adata management functions Basics Available commands	31 31 33 34 34 34 35 35 36 37 37

10	Batch processing	41
	10.1 Options	41
	10.2 Performance tuning options	42
	10.3 Resource selection in batch mode	42
11	The UCC shell	43
	11.1 Changing property settings	43
	11.2 Running an external command	44
	11.3 Exiting the shell	44
12	Sharing resources	44
	12.1 Editing ACLs	45
	12.2 Permission levels	45
13	UCC for site administrators	45
	13.1 Security considerations	46
	13.2 Admin commands	46
	13.3 Listing jobs, sites,	49
	13.4 Low-level operations	50
14	Scripting	50
	14.1 Script context	50
	14.2 Examples	51
15	Frequently asked questions	53
	15.1 Configuration	53
	15.2 Usage	53

1 Overview

The UNICORE Commandline client (UCC) is a full-featured client for the UNICORE middleware. UCC has client commands for all the UNICORE basic services and the UNICORE workflow system.

It offers the following functions

- · Job submission and management
- · Batch mode job submission and processing with many performance tuning options
- Data movement (upload, download, server-to-server copy, etc) using the UNICORE storage management functions and available data transfer protocols
- Storage functions (ls, mkdir, ...) including creation of storage instances via storage factories
- · Support for UNICORE workflow submission and management
- Support for the UNICORE metadata system
- Support for sharing UNICORE resources via ACLs
- · Information about the available services is provided via the "system-info" command
- Various utilities like a "shell" mode, low-level REST API operations and others
- Extensibility through custom commands and the possibility to run scripts written in the Groovy programming language
- Built-in help

Starting with Version 8 of the UCC, the UNICORE REST API is used exclusively for client-server communications.

For more information about UNICORE visit https://www.unicore.eu.

2 Installation and configuration

2.1 Prerequisites

To run UCC, you need a Java runtime version 8 or later (OpenJDK preferred).

2.2 Download

You can get the latest version from the SourceForge UNICORE download page.

2.3 Installation and configuration

To install, unpack the distribution in a directory of your choice. It's a good idea to add the bin/ directory to your PATH variable,

```
$> export PATH=$PATH:<UCC_HOME>/bin
```

where UCC_HOME is the directory you installed UCC in.

Note

Windows only Please do not install UCC into a directory containing spaces such as "Program files".

Also avoid long path names, this can lead to errors due to the Windows limit on command line length.

Setting environment variables can be done (as administrator) using the Control panel \rightarrow System \rightarrow Extras panel.

Though you can specify your keystore location and other parameters on the commandline, it is easiest to place this information in a file, so that you do not have to key in this information repeatedly.

2.4 Preferences file

UCC checks by default whether the file <userhome>/.ucc/preferences exists, and reads it.

A minimal example that specifies keystore, password and your preferred UNICORE registry URL would look as follows:

```
credential.path=<your keystore>
credential.password=<your password>
truststore.type=keystore
truststore.keystorePath=<your keystore>
truststore.keystorePassword=<your password>
client.serverHostnameChecking=NONE
registry=<your registry>
```

Please refer to Section 4 for a full description of available options.

Note

If you are worried about security, and do not want specify the password: UCC will ask for it if it is not given in the preferences or on the commandline.

Note

Windows only The preferences are usually searched in the "c:\Users\<user_name>\.ucc" folder.

To create the .ucc folder, you might have to use the command prompt "mkdir" command. When specifying paths in the preferences file, the backslash $\$ character needs to be written using an extra backslash $\$

For example, if you are using a local UNICORE installation for testing, you could use

```
registry=https://localhost:8080/DEMO-SITE/rest/core/registries/ ↔ default_registry
```

Note

If you wish to change the default property file location, you can set a Java VM property in the UCC start script, for example by editing the command that starts UCC

java -Ducc.preferences=<preferences location>

2.5 Logging

UCC writes some messages to the console, more if you choose the verbose mode (-v option). If you need real logging (e.g. when using the batch mode), you can edit the <UCC_HOME>/conf/logging.properties file, which configures the Log4J logging infrastructure used in UNICORE.

2.6 Installing UCC extensions

UCC can be extended with additional commands. It is enough to copy the libraries (.jar files) of the extension into a directory that is scanned by UCC: in general these are the UCC lib and the HOME/.ucc/lib directory.

2.7 Testing the installation

To test your UCC installation and to get information about the services available in the UNI-CORE system you're connecting to, do

```
$> ucc system-info -l -v
```

3 Getting started with UCC

Assuming you have successfully installed UCC, this section shows how to get going quickly.

3.1 Getting help

Calling UCC with the "-h" option will show the available options. To get a list of available commands, type

```
$> ucc -h
```

To get help on a specific command, type

```
$> ucc <command> -h
```

See also here for a list of common options.

3.2 Connecting

First, contact UNICORE and make sure you have access to some target systems.

```
$> ucc connect [options]
```

3.3 List available sites

Then, list the sites available to you using

```
$> ucc list-sites [options]
```

3.4 Running your first job

The UCC distribution contains samples that you can run. Let's run the "date" sample. The "-v" switch prints more info so you can see what's going on.

\$> ucc run [options] -v [UCC_HOME]/samples/date.u

Note

Look for UCC samples in the /usr/share/doc/unicore/ucc/samples directory,

This will run "date" on a randomly chosen site, and retrieve the output. To run on a particular site, use the "-s" option to specify a particular target system.

3.5 Listing your jobs

The command

```
$> ucc list-jobs
```

will print a list of jobs (actually their addresses) with their respective status (RUNNING, SUC-CESSFUL, etc)

4 Common options to UCC

The following table lists the options understood by most UCC commands. Most commands have additional options. You can always get a summary of all available options for a command by calling UCC with the "-h" or "--help" option, for example

\$> ucc run --help

Since it is not possible to give all the required options on the commandline, it is mandatory to create a preferences file containing e.g. your settings for keystore, registry etc.

Option (short and long	Description
form)	
-c,configuration	Properties file containing your preferences. By default, a
<properties_file></properties_file>	file userhome/.ucc/preferences is checked.
-k,authenticatio:	n Manthematication method to use (default: X509)
<auth></auth>	
-o,output	Directory for any output produced (default is the current
<output_dir></output_dir>	directory)
-r,registry	The comma-separated list of URLs of UNICORE registries
<list_of_registry_< td=""><td>URLs></td></list_of_registry_<>	URLs>
-v,verbose	Verbose mode
-h,help	Print help message
-y,with-timing	Timing mode

Table 1: Common options for the UCC

4.1 User preferences

If you have multiple user IDs or are a member of multiple Unix Groups on the target system, you may wish to control the user attributes that are used when invoking UCC.

Here is a list of options related to user attributes.

Option (short and long form)	Description
-Z,preference	Select from your remote attributes (e.g. xlogin)

The preference option accepts multiple arguments of the form "<name>:<value>" where name

Name	Description
uid	Remote login
pgid	Primary group ID
supgids	Secondary group IDs (comma-separated)
role	UNICORE role (user, admin,)
VO	virtual organisation

Table 3: User attribute options

4.2 Configuration file

By default, UCC checks for the existence of a file <userhome/.ucc/preferences> and reads settings from there. As shown above, you can use a different file by specifying it on the commandline using the "-c" option.

The configuration file can contain default settings for many commandline options, which are given in the form <option name>=<value> where <option name> is the long form of the option. The property values may contain variables in the form $\{VAR_X\}$, which are automatically replaced with the environmental variable values with the same name. Additionally a special variable $\{UCC_CONFIG\}$ is recognized and is replaced with the absolute path of your configuration file.

For example, to set your keystore, truststore and registry, the file would contain the following settings

Note

To protect your passwords, you should make the file non-readable by others, for example on Unix using a command such as *chmod 600 preferences*

Note

If required passwords are not given in the properties file, they will be queried interactively.

4.3 Credential and truststore options

In general you'll want a truststore directory (or file) as well as a truststore file (or directory) containing trusted certificates. You might also need a keystore containing your identity in order to use UNICORE.

A full list of options related to credential and truststore management is available in the following table. You can also get them via the online help using

\$> ucc help-auth

Property name	Туре	Default value /	Description
		mandatory	
credential.path	filesystem path	mandatory to be set	Credential location. In case of <i>jks</i> , <i>pkcs12</i> and <i>pem</i> store it is the only location required. In case when credential is provided in two files, it is the certificate file path.
credential.format	[jks, pkcs12, der, pem]	-	Format of the credential. It is guessed when not given. Note that <i>pem</i> might be either a PEM keystore with certificates and keys (in PEM format) or a pair of PEM files (one with certificate and second with private key).
credential.passwo	rstring	-	Password required to load the credential.

Table 4:	Credential	properties
----------	------------	------------

Property name	Туре	Default value / mandatory	Description
credential.keyPa	athstring	-	Location of the private key if stored separately from the main credential (applicable for <i>pem</i> and <i>der</i> types only),
credential.keyPa	ass stoing	-	Private key password, which might be needed only for <i>jks</i> or <i>pkcs12</i> , if key is encrypted with different password then the main credential password.
credential.keyA	Liastring	-	Keystore alias of the key entry to be used. Can be ignored if the keystore contains only one key entry. Only applicable for <i>jks</i> and <i>pkcs12</i> .

Table 4: (continued)

Table 5: Truststore properties

Property name	Туре	Default	Description	
		value /		
		mandatory		
truststore.allowP	rþ æ lyLOW,	ALLOW	Controls whether proxy	
	DENY]		certificates are supported.	
truststore.type	[keystore,	mandatory	The truststore type.	
	openssl,	to be set		
	directory]			
truststore.update	I integer aumber	600	How often the truststore	
			should be reloaded, in	
			seconds. Set to negative	
			value to disable refreshing	
			at runtime. (runtime	
			updateable)	
Directory type settings				
truststore.direct	o integen numberor	Tibneout	Connection timeout for	
			fetching the remote CA	
			certificates in seconds.	

Property name	Туре	Default value / mandatory	Description
truststore.direct	o filgeisyistern path eF	ath	Directory where CA certificates should be cached, after downloading them from a remote source. Can be left undefined if no disk cache should be used. Note that directory should be secured, i.e. normal users should not be allowed
truststore.direct	o ₽₹EM¢Đ€R ∱g	PEM	to write to it. For directory truststore controls whether certificates are encoded in PEM or DER. Note that the PEM file can contain arbitrary number of concatenated, PEM-encoded certificates.
truststore.direct	olistInfcations. properties with a common prefix	*-	List of CA certificates locations. Can contain URLs, local files and wildcard expressions. (<i>runtime updateable</i>)
	Keystore ty	pe settings	-
truststore.keysto		-	The keystore type (jks, pkcs12) in case of truststore of keystore type.
truststore.keysto	_	-	The password of the keystore type truststore.
truststore.keysto	r stPiag h	-	The keystore path in case of truststore of keystore type.
	Openssl ty	pe settings	
truststore.openss			In case of openssl truststore, specifies whether the trust store is in openssl 1.0.0+ format (true) or older openssl 0.x format (false)

Table 5: (continued)

Property name	Туре	Default	Description
		value /	
		mandatory	
truststore.oper	nssl kGMOB&US_EUC	RHORMA DPMA	-
	EU-		truststore, controls which
	GRIDPMA_GLO	OBUS,	(and in which order)
	GLOBUS,		namespace checking rules
	EUGRIDPMA,		should be applied. The
		RIDPMA_REQ	URHQUIRE settings will
	EU-		cause that all configured
			Enamespace definitions files
	GLOBUS_REQ	JIRE,	must be present for each
	EU-		trusted CA certificate
	GRIDPMA_REC	QUIRE,	(otherwise checking will
	EU-		fail). The AND settings will
	GRIDPMA_AN	D_GLOBUS,	cause to check both existing
	EU-		namespace files. Otherwise
	GRIDPMA_AN	D_GLOBUS_R	EQNE IRES, found is checked
	IGNORE]		(in the order defined by the
			property).
truststore.oper	nsslfikesystem path	/etc/grid	-Detectority to be used for cates
			opeenssl truststore.
	Revocati	on settings	
truststore.crl0			Connection timeout for
			fetching the remote CRLs
			in seconds (not used for
			Openssl truststores).
truststore.crlI	Disk fikesystem path	-	Directory where CRLs
			should be cached, after
			downloading them from
			remote source. Can be left
			undefined if no disk cache
			should be used. Note that
			directory should be
			secured, i.e. normal users
			should not be allowed to
			write to it. Not used for
			Openssl truststores.
truststore.crlI	Locatistoofs.*	-	List of CRLs locations. Can
	properties with		contain URLs, local files
	a common		and wildcard expressions.
	prefix		Not used for Openssl
	Promit		truststores. (runtime
			updateable)
		1	puncuore)

Table 5: (continued)

Property name	Туре	Default	Description
		value /	
		mandatory	
truststore.crlMod		IF_VALID	General CRL handling
	IF_VALID,		mode. The IF_VALID
	IGNORE]		setting turns on CRL
			checking only in case the
			CRL is present.
truststore.crlUpd	a integet eurnber	600	How often CRLs should be
			updated, in seconds. Set to
			negative value to disable
			refreshing at runtime.
			(runtime updateable)
truststore.ocspCa	cintegeit number	3600	For how long the OCSP
			responses should be locally
			cached in seconds (this is a
			maximum value, responses
			won't be cached after
			expiration)
truststore.ocspDi	s filesystem path	-	If this property is defined
			then OCSP responses will
			be cached on disk in the
			defined folder.
truststore.ocspLc		<number></number>	Optional list of local OCSP
	properties with		responders
	a common		
	prefix		
truststore.ocspMc		IF_AVAILA	BGieneral OCSP ckecking
	IF_AVAILABLE,		mode. REQUIRE should
	IGNORE]		not be used unless it is
			guaranteed that for all
			certificates an OCSP
	1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1	10000	responder is defined.
truststore.ocspTi	menteger number	10000	Timeout for OCSP
	ICDE OCCE	0.000 05-	connections in miliseconds.
truststore.revoca		OCSP_CRL	Controls overal revocation
	OCSP_CRL]	C -] -	sources order
truststore.revoca	t fiture) Ba≇se]⊥	false	Controls whether all
			defined revocation sources
			should be always checked,
			even if the first one already
			confirmed that a checked
			certificate is not revoked.

Table 5: (continued)

4.4 Trust store examples

Here are some examples for commonly used trust store configurations.

Most commonly used is a directory (with a minimal set of options)

```
truststore.type=directory
truststore.directoryLocations.1=/trust/dir/*.pem
```

Java keystore used as a trust store:

```
truststore.type=keystore
truststore.keystorePath=/some/dir/truststore.jks
truststore.keystoreFormat=JKS
truststore.keystorePassword=xxxxxx
```

OpenSSL trust store

```
truststore.type=openssl
truststore.opensslPath=/truststores/openssl
truststore.opensslNsMode=EUGRIDPMA_GLOBUS_REQUIRE
truststore.updateInterval=1234
truststore.crlMode=IF_VALID
```

4.5 Support for oidc-agent

For infrastructures using OIDC authentication, UCC supports the *oidc-agent* tool that allows to interact with common OIDC servers to retrieve new access tokens.

Please visit https://github.com/indigo-dc/oidc-agent for more information.

To use oidc-agent, UCC has the following properties

Property name	Туре	Default value / mandatory	Description
oidc-agent.accoun	t string	mandatory to be set	Account short name.
oidc-agent.lifeti	mienteger >= 1	-	Minimum lifetime of the issued access token.
oidc-agent.scope	string	-	OpenID scope(s) to request.

Table	6.	Options	for	oidc-	agent
raute	υ.	Options	101	oruc	ugom

Note that you'll also need to configure a Unity server that will accept the OIDC token. Your config file would require at least:

```
authenticationMethod=oidc-agent
unity.address=https://<host>:<port>/unicore-soapidp/saml2unicoreidp ↔
-soap-oidc/AuthenticationService
oidc-agent.account=<oidc-agent account to be used>
```

4.6 Username and password

To authenticate with username and password, set the following

authenticationMethod=username
username=<your remote username>
password=<your remote password>

4.7 Using Unity

If your UNICORE installation is using the Unity identity management service (see http://www.unityidm.eu), you can setup UCC to use Unity. In this case you do not need a private key, only a truststore. UCC is configured for Unity using the following properties

```
authenticationMethod=unity
unity.address=https://<host>:<port>/unicore-soapidp/saml2unicoreidp ↔
    -soap/AuthenticationService
unity.username=<your Unity username>
unity.password=<your Unity password>
```

At minimum, you must specify the "authenticationMethod" and "unity.address" parameters. Password and username are optional: if you do not specify them, they will be queried interactively.

In some installations, you can use a so-called OIDC Bearer token instead of username and password. In this case specify only unity address and the token, like this:

```
authenticationMethod=unity
unity.address=https://<host>:<port>/unicore-soapidp/saml2unicoreidp ↔
    -soap-oidc/AuthenticationService
oauth2.bearerToken=<your oauth token>
```

4.8 Client options

The configuration file may also contain low-level options, for example if you need to specify connection timeouts, http proxies etc.

Property name	Туре	Default value / mandatory	Description
client.digitalSig	n [trug£fækse] ed	true	Controls whether signing of key web service requests should be performed.
client.httpAuthnE	n [tinieș false]	false	Whether HTTP basic authentication should be used.
client.httpPasswo	rstring	empty string	Password for use with HTTP basic authentication (if enabled).
client.httpUser	string	empty string	Username for use with HTTP basic authentication (if enabled).
client.inHandlers	string	empty string	Space separated list of additional handler class names for handling incoming WS messages
client.maxWsCallR		3	Controls how many times the client should try to call a failing web service. Note that only the transient failure reasons cause the retry. Note that value of 0 enables unlimited number of retries, while value of 1 means that only one call is tried.
client.messageLog		false	Controls whether messages should be logged (at INFO level).
client.outHandler		empty string	Space separated list of additional handler class names for handling outgoing WS messages
client.securitySe	s [triucen false]	true	Controls whether security sessions should be enabled.

Table 7: Client options

Property name	Туре	Default value / mandatory	Description
client.serverHost	WARN, FAIL]	WARN	Controls whether server's hostname should be checked for matching its certificate subject. This verification prevents man-in-the-middle attacks. If enabled WARN will only print warning in log, FAIL will close the connection.
client.sslAuthnEn		true	Controls whether SSL authentication of the client should be performed.
client.sslEnabled		true	Controls whether the SSL/TLS connection mode is enabled.
client.wsCallRetr	y integery number	10000	Amount of milliseconds to wait before retry of a failed web service call.
	HTTP clie	nt settings	
client.http.allow	- [thuep false]	true	If set to false, then the client will not use HTTP 1.1 data chunking.
client.http.conne		false	If set to true then the client will send connection close header, so the server will close the socket.
client.http.conne		20000	Timeout for the connection establishing (ms)
client.http.maxPe		6	How many connections per host can be made. Note: this is a limit for a single client object instance.
client.http.maxRe		3	Maximum number of allowed HTTP redirects.
client.http.maxTo		20	How many connections in total can be made. Note: this is a limit for a single client object instance.
client.http.socke	-	0	Socket timeout (ms)
	HTTP pro.	xy settings	

Table 7:	(continued)
----------	-------------

Property name	Туре	Default value / mandatory	Description
client.http.nonPr	o styiHg sts	-	Space (single) separated list of hosts, for which the HTTP proxy should not be used.
client.http.proxy	. staing word	-	Relevant only when using HTTP proxy: defines password for authentication to the proxy.
client.http.proxy	. string	-	Relevant only when using HTTP proxy: defines username for authentication to the proxy.
client.http.proxy	Hstring	-	If set then the HTTP proxy will be used, with this hostname.
client.http.proxy	₽ integer number	-	HTTP proxy port. If not defined then system property is consulted, and as a final fallback 80 is used.
client.http.proxy	Typing	HTTP	HTTP proxy type: HTTP or SOCKS.

Table 7: (continued)

4.9 Other options

The following table lists other options, that are more rarely used.

Table 8: Other options for the UCC

Property name	Description
blacklist	Comma separated <i>patterns</i> for sites / URLs to ignore
contact-registry	Do not attempt to contact the registry, even if one is
	configured

5 Running jobs

5.1 Introduction

The UCC can run jobs specified in the JSON job description format that is used by the UNI-CORE REST API, plus a few extensions related to handling of local files, submission options etc. See Section 6 for all the details.

In the following it is assumed that you have UCC installed Section 2 and tried some examples Section 3 .

For example, assume the file "myjob.u" looks as follows

```
{
  "ApplicationName": "Date",
  "ApplicationVersion": "1.0"
}
```

To run this through UCC, issue the following command

ucc run myjob.u

This will submit the job, wait for completion, download the stdout and stderr files, and place them in your default output directory. The run command has a number of options, to see all the possibilities use the built-in help:

ucc run -h

5.1.1 Controlling the output location and file names

Output files will be placed in the directory given by the "-o" option, if not given, the current directory is used. Also, file names will be put into a subdirectory named as the job id, to prevent accidental overwriting of existing files. This behaviour can be changed using the "-b" option. When "-b" is given on the command line, no subdirectory will be created.

5.1.2 Specifying the site

In the example above, a random site will be chosen to execute the job. To control it, you can use the "-s" option. This will accept the name of a target system. The target systems available to you can be listed by

ucc list-sites

5.1.3 Accessing a job's working directory

Using the UCC's data management functions, the job working directory can be accessed at any time after job submission. Please see section Section 7 for details.

5.2 Options overview

The following options are available when running jobs (see also the general options overview in Section 4.

Option (Short and long	Description
form)	
-a,asynchronous	Run asynchronously
-b,brief	Do not create a sub-directory for output files
-B,broker	Select the type of resource broker to use (see <i>run -h</i> for a
	list)
-d,dryRun	Only show candidate sites, but do not submit the job
-s,sitename	Site where the job shall be run
<site></site>	
-S,schedule	Schedule the submission of the job at the given time
<time></time>	
-o,output	Directory for any output produced (default is the current
<output_dir></output_dir>	directory)
-0,stdout	specify a name for the exported standard out (by default:
<stdout_name></stdout_name>	stdout)
-E,stderr	specify a name for the exported standard error (by default:
<stderr_name></stderr_name>	stderr)

Table 9: Jo	ob submission	options	for UCC
14010 / 10	00 000000000	opnomo	101 0 0 0

5.3 Resource selection

In general the user selects the execution site.

If no site is specified upon submission, UCC will select a matching site, where the requirements (resources, applications) are met.

In case there are other types of brokers available, they can be selected using the "-B" or "-broker" option.

• LOCAL (default): brokering is done by UCC itself

To see if other brokers exist, execute "ucc run -h", the available options will be listed in the help for the "-B" option.

5.4 Processing jobs asynchronously

In case of long-running jobs, you will want to run the job asynchronously, i.e. just submit the job, stage in any files and start it, in order to get the results later. UCC supports this, of course. The basic idea is that when submitting a job in asynchronous mode, a job descriptor file is written that contains the job's address, and any information about export files.

5.4.1 Asynchronous submission

Use the "-a" flag when submitting a job

ucc run -a <job file>

This will submit the job, stage-in any local files, start the job and exit. A job descriptor file (ending in ".job") will be written to your configured output directory.

5.4.2 Get the status of particular jobs

The command

ucc job-status <job_desc> <job_desc_2> ...

will retrieve the status of the given jobs. If not given on the command line, a job ID will be read from the console.

5.4.3 Download results

To get stdout, stderr and other files you have marked for export in your job description, do

ucc get-output -o <outdir> <job_desc>

Here, the option "-o" specifies the directory where to put the output, by default the current directory is used. As before, the job address can also be read from the console.

5.4.4 Referencing a job by its URL (endpoint address)

In case you want to check on a job not submitted through UCC, or in case you do not have the job descriptor file any more, you can also refer to a job given its URL. The "list-jobs" command will produce a list of all job URLs that you can access.

Note that in this case UCC will only retrieve stdout and stderr files. To download other result files, you'll have to use the data movement functions described in Section 7.

5.4.5 Scheduling job submission to the batch system

Sometimes a user wishes to control the time when a job is submitted to the batch queue, for example because she knows that a certain queue will be empty at that time.

To schedule a job, you can either use the "-S" option to the ucc "run" command:

```
ucc run -S "12:24" ...
```

Alternatively, you can specify the start time in your job file using the "Not before" key word

```
{
   "Not before": "12:30",
}
```

In both cases, the specified start time can be given in the brief "HH:mm" (hours and minutes) format shown above, or in the full ISO 8601 format including year, date, time and time zone:

```
{
    "Not before": "2011-12-24T12:30:00+0200",
}
```

5.5 Executing a command

If you just want to execute a simple command remotely (i.e. without data staging, resource specifications etc), you can use the "exec" command.

This will run the given command remotely (similarly to "ssh"), and print the output to the console. You can specify the site with the "-s" option. If you do not specify the site, a random site will be chosen.

UNICORE will run the command on the login node, it will not be submitted to the batch system.

For example, try

ucc exec /bin/date

Watch out to properly escape any arguments, in order not to interfere with the arguments to UCC.

6 Job description format

UNICORE uses a JSON format that allows you to specify the application or executable you want to run, arguments and environment settings, any files to stage in from remote servers or the local machine and any result files to stage out.

Several complete job samples can be found in the "samples" directory of the distribution. On Linux, check also the /usr/share/unicore/ucc/samples directory.

Comments are (inofficially!) possible using the "#" hash character, as in Unix shell scrips.

Note

Note: quotes "" are needed around the keys and values in case special characters (like : or /") appear, if in doubt use quotes!

To view an example job showing most of the available options, run

ucc run -H

(most of the options shown are not mandatory, of course)

Note

UCC also accepts jobs in the JSDL format that is used internally in UNICORE. To do this, use the "-j" option when submitting the job.

Usually, a UNICORE job file describe a single batch job on the target system. However there is a feature called "parameter sweep" which leads to the creation of multiple batch jobs from a single "template" job. UCC can also create these "sweep jobs", as described in the relevant parts of the job description. Note that a "sweep job" still is treated as a single job by UNICORE. Sweep jobs are very useful if you need to run jobs that are highly similar, and only differ by a parameter setting or even by a different input file.

6.1 Overview

UNICORE's job description consists of multiple parts

- an *Imports* section listing data to be staged in to the job's working directory from remote storage locations or the client's file system
- pre-processing
- a section describing the main executable
- · post-processing
- an *Exports* section listing result files to be staged out to remote storage locations
- a Resources section stating any resource requirements like job runtime or number of nodes
- a number of additional elements for setting the job name, or defining tags for the job

Here is a table listing the supported elements, these will be described in more detail below.

Tag	Туре	Description
ApplicationName	String	Application name
ApplicationVersion	String	Application version
Executable	String	Command line
Arguments	List of strings	Command line arguments
Environment	Map of strings	Environment values
Parameters	Мар	Application parameters
Optional	IgnoreNonZeroExitCode	"true"/"false"
Don't fail the job if app		
exits with non-zero exit		
code (default: false)		
User precommand	String	Pre-processing
RunUserPrecommandOn	Login'Norde"/"false"	Pre-processing is done on
		login node (default: true)
User postcommand	String	Post-processing
RunUserPostcommandO		Post-processing is done on
		login node (default: true)
RunUserPostcommandO	nLogih tNæd e"false"	Post-processing is done on
		login node (default: true)
Resources	Мар	The job's resource requests
Project	String	Accounting project
Project	String	Accounting project
Imports	List of imports	Stage-in / data import
Exports	List of exports	Stage-out / data export
Exports	List of exports	Stage-out / data export
Job type	normal, interactive	Whether to run the job via
V 1		the batch system (<i>normal</i> ,
		default) or on a login node
		(interactive)
Login node	String	For <i>interactive</i> jobs, select
0	6	a login node (by name, as
		configured server side.
		Wildcards * and ? can be
		used)
Tags	List of strings	Job tags
Notification	String	URL to send job status
	<i>o</i>	change notifications to (via
		HTTP POST)
User email	String	User email to send
	<i>o</i>	notifications to (if the batch
		system supports it)
Site	String	UNICORE site name to run
		the job at (only applicatible
		if using a broker!)

Table 10: UNICORE JSON job description

Table 10: (continued)

Tag	Туре	Description
Name	String	Job name

6.2 Specifying the executable or application

You can specify a UNICORE application by name and (optional) version, or using a (machine dependent) path to an executable file.

```
{
    "ApplicationName": "Date",
    "ApplicationVersion": "1.0",
}
```

Note the comma-separation and the curly braces. To directly call an executable,

```
{
    "Executable": "/bin/date",
}
```

6.3 Arguments and Environment settings

Arguments and environment settings are specified using a list of String values. Here is an example.

```
{
   "Executable": "/bin/ls",
   "Arguments": ["-l", "-t"],
   "Environment": [ "PATH=/bin:$PATH", "FOO=bar" ],
}
```

6.3.1 Argument sweeps

To create a sweep over an Argument setting by replacing the value by a sweep specification. This can be either a simple list:

```
"Arguments": [
{ "Values": ["-o 1", "-o 2", "-o 3"] },
],
```

or a range:

```
"Arguments": {
    "-o", { "From": "1", "To": "3", "Step" : "1" },
},
```

where the From, To and Step parameters are floating point or integer numbers.

6.4 Application parameters

In UNICORE, parameters for applications are often transferred in the form of environment variables. For example, the POVRay application has a large set of parameters to specify image width, height and many more. In UCC, you can specify these parameters in a very simple way using the "Parameters" keyword:

```
{
   "ApplicationName": "POVRay",
   "Parameters": {
    "WIDTH": "640",
    "HEIGHT": "480",
    "DEBUG": "",
    },
}
```

Note that an "empty" parameter (which does not have a value) needs to be written with an explicit empty string due to the limitations of the JSON syntax.

6.4.1 Parameter sweeps

You can sweep over application parameters by replacing the parameter value by a sweep specification. The replacement can be either a simple list:

```
"Parameters": {
    "WIDTH": { "Values": ["240", "480", "960"] },
},
```

or a range:

```
"Parameters": {
   "WIDTH": { "From": "240", "To": "960", "Step": "240" },
},
```

where the From, To and Step parameters are floating point or integer numbers.

6.5 Job data management

In general your job will require data files, either from your client machine, or from some remote location. An important concept in UNICORE is the job's workspace, which is the default location into which files are placed. The same applies to result files: by default, files will be downloaded from the job's workspace.

However, other remote storage locations are supported, too.

The remote location can be given as a full UNICORE URI, or using the more user friendly (but slower) "unicore://" notation. Read more on remote locations in Section 7.

Local files can be given as an absolute or relative path; in the latter case the configured output directory will be used as base directory.

6.5.1 Importing files into the job workspace

To import files from your local computer or from remote sites to the job's working directory on the remote UNICORE server, there's the "Imports" keyword. Here is an example Imports section which demonstrates some of the possibilities.

```
{
"Imports": [
# import a local file from the client machine
# into the job workspace
 { "From": "/work/data/fileName", "To": "fileName" },
#
# import a set of local files from the client machine
# into the job workspace
 { "From": "/work/data/pdf/*.pdf", "To": "/" },
# import a remote file from a UNICORE storage. The real address
# will be resolved by UCC
 { "From": "unicore://DEMO-SITE/Home/testfile", "To": "testfile" },
# import a remote file from a UNICORE storage using the UFTP \, \leftrightarrow \,
   protocol
 { "From": "UFTP:https://gw:8080/DEMO-SITE/services/ ↔
    StorageManagement?res=Home#testfile",
   "To": "testfile" },
```

```
# create a symlink from a file on the compute machine to the job ↔
workspace
{ "From": "link:/work/data/testfile", "To": "linked-file" },
# copy a file on the compute machine to the job workspace
{ "From": "link:/work/data/testfile", "To": "copied-file" },
],
}
```

If for some reason an import fails, but you want the job to run anyway, there is a flag "FailOn-Error" that can be set to "false" :

```
"Imports": [
#
# do not fail on errors for this import:
#
{ "From": "/work/data/fileName",
    "To": "fileName",
    "FailOnError": "false",
},
```

],

Note

UCC supports simple wild cards ("*" and "?") for importing and exporting files

Supported protocols for imports

- file: // : copy file(s) from the remote machine into the job dir
- link://: symlink file/dir from the remote machine into the job dir
- unicore://: resolve location pointing to some UNICORE server

Using "inline" data to import a file into the job workspace

For short import files, it can be convenient to place the data directly into the job descrition, which can speed up and simplify the job submission process.

Here is an example:

```
"Imports": [
  { "From": "inline://dummy",
    "To": "uspaceFileName",
    "Data": "this is some test data", },
]
```

The "From" URL has to start with "inline://"

Sweeping over a stage-in file

{

You can also sweep over files, i.e. create multiple batch jobs that differ by one imported file. To achieve this, replace the "From" parameter by list of values, for example:

Note that only a single stage-in can be sweeped over in this way, and that this will not work with files imported from your local client machine.

6.5.2 Exporting result files from the job workspace

To export files from the job's working directory to your local machine or to some remote storage, use the "Exports" keyword. Here is an example Exports section that specifies two exports:

```
"Exports": [
    #this exports all png files to a local directory
    { "From": "*.png", "To": "/home/me/images/" },

    #this exports a single file to a to local directory
    #failure of this data transfer will be ignored
    { "From": "error.log", "To": "/home/me/logs/error.log", " ↔
        FailOnError": "false", },

    #this exports to a UNICORE storage
    { "From": "stdout", "To": "unicore://DEMO-SITE/Home/results/ ↔
        myjob/stdout" },
}
```

As a special case, UCC also supports downloading files from other UNICORE storages using the Exports keyword:

```
{
    "Exports": [
    #this exports a file from a UNICORE storage
    { "From": "unicore://DEMO-SITE/Work/somefile", "To": "/home/me/ ↔
        somefile" },
  ]
}
```

The protocol to be used for imports and exports can be chosen using the "Preferred Protocols" entry, containing a space-separated list of protocols:

```
"Preferred protocols": "UFTP BFT",
}
```

If not specified or not available at the remote site, BFT will be used.

6.5.3 Specifying credentials for data staging

Some data staging protocols supported by UNICORE require credentials such as username and password. To pass username and password to the server, the syntax is as follows

```
"Imports": [
    { "From": "ftp://someserver:25/some/file", "To": "input_data",
        "Credentials": { "Username": "myname", "Password": " ↔
        mypassword" },
    },
]
```

and similarly for exports.

{

Servers 7.9.0 and later also support OAuth Bearer token for HTTP data transfers.

```
{
   "Imports": [
    { "From": "https://someserver/some/file", "To": "input_data",
        "Credentials": { "BearerToken": "some_token" },
    },
    ]
}
```

You can leave the token value empty, "BearerToken": "", if the server already has your token by some other means.

6.5.4 Redirecting standard input

If you want to have your application or executable read its standard input from a file, you can use the following

"Stdin": "filename",

then the standard input will come from the file named "filename" in the job working directory.

6.6 Resources

A job definition can have a Resources section specifying the resources to request on the remote system. For example

```
"Resources": {
    "Runtime": "12h",
    "Nodes": "8",
    "Queue" : "fast",
}
```

UNICORE has the following built-in resource names.

Resource name	Description	
Runtime	Job runtime (wall time) (in seconds, append "min", "h", or	
	"d" for other units)	
Queue	Batch system queue / partition to use	
Nodes	Number of nodes	
CPUs	Total number of CPUs	
CPUsPerNode	Number of CPUs per node	
Memory	Memory per node	
Reservation	Batch system reservation ID	
NodeConstraints	Batch system node constraints	

In addition, sites may define custom resources, which you can use, too.

6.7 Miscellaneous options

6.7.1 Site name

To specify on which site (if available) the job should be run (e.g. for UCC's batch mode)

```
"Site": "DEMO-SITE",
```

If you do not specify anything UCC will select a site that will match your requirements (at least those that UCC checks for). You can also set the site during job submission as an option to "ucc run ...".

6.7.2 Specifying a project

If the system you're submitting to requires a project name for accounting purposes, you can specify the account (or project) you want to charge the job to using the "Project" tag:

```
"Project" : "my_project",
```

6.7.3 Job tags

To set job tags that help you find / filter jobs later, use the "Tags" keyword

```
"Tags": [ "production", "train1", "my_tag" ],
```

6.7.4 Specifying a URL for receiving notifications

The UNICORE/X server can send out notifications when the job enters the RUNNING and/or DONE state.

To enable this, add the URL of the receiving service to your job:

```
"Notification" : "https://your-service-url",
```

UNICORE will send an authenticated HTTPS POST message to this URL, with JSON content.

```
"href" : "https://unicore-url/rest/core/jobs/job-uuid",
"status" : "RUNNING",
"statusMessage" : ""
```

The "status" field will be RUNNING when the user application starts executing, and "SUC-CESSFUL" / "FAILED" when the job has finished.

```
"href" : "https://unicore-url/rest/core/jobs/job-uuid",
"status" : "SUCCESSFUL",
"statusMessage" : "",
"exitCode" : 0
```

Do not expect "realtime" behaviour here, as UNICORE has a certain delay (typically 30 to 60 seconds, depending on the server configuration) until "noticing" job status changes on the batch system.

6.7.5 Specifying the user email for batch system notifications

Some batch systems support sending email upon completion of jobs. To specify your email, use

```
"User email" : "foo@bar.org" ,
```

6.7.6 Specifying the job name

The job name can be set simply by

"Name": "Test job",

7 Data management functions

UCC offers access to all the data management functions in UNICORE. You can upload or download data from a remote server, initiate a server-to-server transfer, create directories and so on.

7.1 Specifying remote locations

Remote locations are via URIs that includes protocol, storage server (host/port), site name, and filename, for example

```
BFT:https://mygateway:8080/SITE/rest/core/storages/HOME/files/ ↔
  my_file
```

which specifies a file named "/my_file" on the storage instance "https://mygateway:8080/SITE/rest/core/storages/HOME", using the BFT protocol.

Paths are always relative to the storage root, not the root of the actual file system.

The protocol is optional, and will default to "BFT" if not given.

7.2 Data movement

7.2.1 cp

The cp command is a generic command for copying source file(s) to a target destination, where source and target can be remote locations or files on the local machine. Wild card characters * and ? are supported.

Examples for client-server transfers:

```
ucc cp data/*.pdf https://server/rest/core/storages/SHARE/files/ ↔
pdfs
ucc cp https://server/rest/core/storages/SHARE/files/pdfs .
```

The "-R" option allows to choose whether subdirectories are to be copied, too.

The "-X" option allows to resume a previous transfer. Missing data will be appended to an existing target file (if the chosen protocol supports it).

Examples for server-server transfer:

```
ucc cp https://server/rest/core/storages/SHARE/files/*.pdf \
    https://otherserver/rest/core/storages/WORK/data/
```

For server-to-server transfers, the cp command supports several additional options.

The "-S" option allows to schedule a transfer for a certain time. For example

ucc cp -S "23:00" ...

The format is simply "HH:mm" (hours and minutes). Alternatively you can give the time in the full ISO 8601 format including year, date, time and time zone:

ucc cp -S "2011-12-24T12:30:00+0200" ...

Another useful option is "-a" which will execute the server-server transfer asynchronously, i.e. the client will not wait for the transfer to finish.

7.2.2 copy-file-status

This will print the status of the given data transfer. As argument, it expects a file name containing the transfer reference, or directly the reference.

Example (for Unix) which captures the reference into a shell variable:

```
export ID=$(ucc cp -a ...
ucc copy-file-status $ID
```

7.2.3 Specifying the file transfer protocol

To use a different protocol from the default BFT, you can use the "-P" option to specify your preferred protocol. UCC will try to match them with the capabilities of the storage and use the first match. Your preferred protocol can also be listed in your preferences file using the "protocols" key:

protocols=UFTP

Note

If necessary, you can specify additional filetransfer options in your preferences file as well. For example, to use the UFTP protocol you may need to specify the client host address and the number of parallel streams explicitely:

```
uftp.client.host=your_client_ip_address
uftp.streams=2
# encrypt data (at the cost of performance)
uftp.encryption=true
# compress data
uftp.compression=true
```

Use the special value "all" to enable all available client IP addresses for UFTP.

uftp.client.host=all

You can also override the UFTP server host, which can be useful in case the UFTP server is accessible via multiple network interfaces:

uftp.server.host=myhost.com

UCC will try to use reasonable defaults for any missing parameters.

7.3 General commands

7.3.1 mkdir

This will create a directory (including required parent directories) remotely.

Example

```
ucc mkdir https://mygateway:8080/SITE/rest/core/storages/HOME/files ↔ /pdfs
```

7.3.2 rm

This will remove a file or directory remotely. By default, UCC will ask for a confirmation. Use the "--quiet" or "-q" option to disable this confirmation (e.g. when using this command in scripts).

Example

```
ucc rm https://mygateway:8080/SITE/rest/core/storages/HOME/files/ ↔ pdfs
```

7.3.3 rename

This will rename/move a remote file/directory on the same storage.

Examples

```
ucc rename https://mygateway:8080/SITE/rest/core/storages/HOME/ ↔ files/data/fool.pdf /files/data/foo2.pdf
```

will rename the file "foo1.pdf" to "foo2.pdf"

7.3.4 stat

This command shows full information on a certain file or directory. Add the "-m" flag to also print user-defined metadata.

Example

```
ucc stat -m https://mygateway:8080/SITE/rest/core/storages/HOME/ ↔ files/foo.txt
```

7.4 Finding data

7.4.1 ls

This will list a remote directory. Useful options are: "-l" (detailed output), "-H" (human-friendly) and "-R" (recurse). Example:

```
ucc ls -l -H https://mygateway:8080/SITE/rest/core/storages/HOME/
```

If the storage supports metadata, you can get the metadata of a single file using "ls -l -m":

```
ucc ls -l -m https://mygateway:8080/SITE/rest/core/storages/HOME/. \leftrightarrow bashrc
```

7.5 Using the StorageFactory service

UNICORE sites may allow users to dynamically create storage resources, which even can be linked to special back-end systems like Apache HDFS, iRODS, or cloud storage like Amazon S3.

You can find out if there are sites supporting this "StorageFactory" service either by running the *system-info -l* command, or better using

```
$> ucc create-storage -i
```

This will list the available StorageFactory services and also show which types of storage are supported and how much space is left on each of them.

UCC supports creating storages via the create-storage command. The simple

\$> ucc create-storage

will create a new storage resource using the default storage type at some site.

Usually you want to control at least where the storage is created. Additionally, the type of storage and some parameters can be passed to UCC.

As an example, creating a storage of type "S3" would look like this

\$> ucc create-storage -t S3 accessKey=... secretKey=...

You can also read parameters from a file. Say you have your S3 keys in a file *s3.properties*, then you can use the following syntax

```
$> ucc create-storage -t S3 @s3.properties
```

You can also mix this with the normal *key=value* syntax, or mix it like this:

```
$> ucc create-storage -t S3 accessKey=@s3.accessKey secretKey=@s3. ↔
secretKey
```

The last version *key*=@*file* causes just the value to be read from the named file.

8 Metadata management functions

UCC offers a simple interface to access the metadata management service in UNICORE.

8.1 Basics

The metadata functions are all accessed via a single UCC command metadata. The actual operation to be performed is given with the "-C" (i.e. "command") option.

The storage to be operated upon is given using the "-s" option, alternatively the "-m" option can be used to directly give the metadata service URL.

In addition to the URL, the name of the target file on the storage is required.

Metadata is represented in JSON format. The metadata operations usually read metadata from a file (or write results to file), which is specified using the "-f" option.

In the following examples, <STORAGE> denotes the URL of a storage capable of handling metadata.

8.2 Available commands

8.2.1 creating metadata

To create metadata, a file in JSON format is required containing key-value pairs. For example, edit the file "meta.json" to contain:

{
 foo: bar
}

Say we have a file "test" on our storage, then you can create metadata as follows

ucc metadata -C create -f meta.json -s <STORAGE> /test

If you now look at the file with "ls -l -m",

ucc ls -l -m <STORAGE>/test

you should get something like this:

```
-rw- 3344 2011-06-27 22:32 /test {
    "foo": "bar",
    "resourceName": "/test"
}
```

8.2.2 reading metadata

Apart from the "ls -l -m" used above, there is also an explicit "read" command, which can write the metadata to a file as well.

ucc metadata -C read -s <STORAGE> /test -f out.json

The "-f" option is optional.

8.2.3 updating metadata

Using update, the given metadata is merged with any existing metadata. Say we have a file x.json containing:

{ x: y }

we can append this to the existing metadata

ucc metadata -C update -s <STORAGE> /test -f x.json

Check that the metadata has indeed been appended.

8.2.4 deleting metadata

Explicitely deleting is also possible:

ucc metadata -C delete -s <STORAGE> /test

Check that the metadata has indeed been deleted.

8.2.5 searching

Searching requires a search string (according to the rules of Apache Lucene), and is triggered by the "search" command:

ucc metadata -C search -q "foo" -s <STORAGE> /

8.2.6 triggering metadata extraction

To trigger the extraction of metadata on the server, use the "start-extract" command:

```
ucc metadata -C start-extract -s <STORAGE> /
```

In this case the "/" denotes the base path from which to start the extraction process.

9 Workflows

9.1 Introduction

Note

In version 8.0.0 the workflow system support is not available it will be re-added when the version 8.0 of the workflow system has been released

UCC supports the UNICORE workflow system and allows to submit workflows to the workflow engine and manage them.

The workflows are executed server-side, and UCC is used only for submitting, managing data and getting results.

9.2 Command overview

The following commands are provided. More details and examples follow below.

- workflow-submit : submit a workflow file
- workflow-control: abort or resume a running workflow
- list-workflows : list information about workflows

9.3 Basic use

To check the availability of workflow services, issue the following command

```
ucc system-info -1
```

This should show at least an accessible workflow engine and service orchestrator.

The distribution contains some example workflow files in the <[?]> directory that you can edit and submit.

ucc workflow-submit yourworkflow.swf

which will submit the workflow and print the address of the workflow to standard output. To get the workflow status,

ucc list-workflows <workflow_address>

To list all your workflows, you can use the <[?]> command without an explicit workflow address

ucc list-workflows -1

9.4 Managing workflow data

During workflow execution, data files will be produced that the workflow system will move to a storage location that is accessible for the individual jobs. Usually this location is created automatically by UCC before the workflow is submitted, using a special UNICORE service called a storage factory. If you want to influence this decision, UCC allows to select the storage factory to be used via the "-f" option to the workflow-submit command:

ucc workflow-submit -f <factory-url> <workflow-file>

You can check the available factories with the system-info command. If not specified, UCC will use the fist storage factory it finds in the registry.

In general, storages that are dynamically created will be deleted when the workflow is deleted. To persistently store data, you need to make sure to export important result files to a persistent location (e.g. your Home on UNICORE site, or a persistent storage).

Alternatively you can directly specify a storage URL, either using the convenient "u6://..." notation, or as a real network URL:

```
ucc workflow-submit -S u6://MY-SITE/Home <workflow-file>
ucc workflow-submit -S https://my-gateway/SITE/services/ ↔
StorageManagement?res=myuser-Home <workflow-file>
```

9.4.1 Importing local data for use by a workflow

If you have local files that need to be imported before starting the workflow, you have to specify this using a normal UCC job file that contains only an "Imports" section:

```
#stage-in specification
Imports: [
   {From: local-file.sh, To: "c9m:${WORKFLOW_ID}/input.sh"}
],
}
```

When submitting the workflow, add the "-u <filename>" option to specify the imports file.

This will cause UCC to copy the local file "local-file.sh" to the workflow storage space. You can refer to this file in your workflow using the "global" name "c9m:...", say in a script activity:

```
....
<jsdl:DataStaging>
<jsdl:CreationFlag>overwrite</jsdl:CreationFlag>
<jsdl:FileName>input.sh</jsdl:FileName>
<jsdl:Source>
<jsdl:URI>c9m:${WORKFLOW_ID}/input.sh</jsdl:URI>
</jsdl:Source>
</jsdl:DataStaging>
....
```

The workflow system will resolve the name at runtime and your file will be used. This allows you to group your files by workflow ID.

9.4.2 Downloading output files

You can use the usual get-file command to download files using the "global IDs" used by the workflow engine. Hint: the workflow-info command will list the files that are produced by the workflow.

9.4.3 Workflow templates

If the workflow's Documentation section contains Argument definitions, the corresponding replacement will be done by reading parameter values from the .u file. These so-called workflow templates can be a very simple and safe way to make adjustments in complex workflows # before submission. As an example, consider the following simple example workflow

```
<s:Workflow xmlns:s="http://www.chemomentum.org/workflow/simple"
xmlns:jsdl="http://schemas.ggf.org/jsdl/2005/11/jsdl"
```

```
xmlns:jsdl-p="http://schemas.ggf.org/jsdl/2005/11/jsdl- ↔
             posix"
          xmlns:jsdl-u="http://www.unicore.eu/unicore/jsdl- ↔
             extensions">
 <s:Documentation>
 <!-- template parameter definitions -->
 <jsdl-u:Argument>
   <jsdl-u:Name>UC_PRECISION</jsdl-u:Name>
   <jsdl-u:ArgumentMetadata>
      <jsdl-u:Type>choice</jsdl-u:Type>
      <jsdl-u:ValidValue>low</jsdl-u:ValidValue>
      <jsdl-u:ValidValue>medium</jsdl-u:ValidValue>
       <jsdl-u:ValidValue>high</jsdl-u:ValidValue>
   </jsdl-u:ArgumentMetadata>
 </jsdl-u:Argument>
 </s:Documentation>
 <s:Activity Id="date" Name="JSDL">
  <s:JSDL>
     <jsdl:JobDescription>
        <jsdl:Application>
          <jsdl:ApplicationName>ComputeIt</jsdl:ApplicationName>
          <jsdl-p:POSIXApplication>
            <jsdl-p:Argument>--precision ${UC_PRECISION}</jsdl-p: ↔
               Argument>
          </jsdl-p:POSIXApplication>
       </jsdl:Application>
     </jsdl:JobDescription>
    </s:JSDL>
 </s:Activity>
</s:Workflow>
```

This defines a parameter UC_PRECISION which is used as application argument in a job. The ArgumentMetadata follow the same specification as application metadata in the UNICORE IDB. The UCC will now try to take the value for UC_PRECISION from the supplied .u file. As an example, the .u file could contain

```
"UC_PRECISION" : "high",
}
```

Before submission, UCC will then replace this value, resulting in the workflow

```
xmlns:jsdl-u="http://www.unicore.eu/unicore/jsdl- ↔
             extensions">
  <s:Documentation>
  <!-- ... -->
  </s:Documentation>
  <s:Activity Id="date" Name="JSDL">
   <s:JSDL>
      <jsdl:JobDescription>
        <jsdl:Application>
          <jsdl:ApplicationName>ComputeIt</jsdl:ApplicationName>
          <jsdl-p:POSIXApplication>
            <jsdl-p:Argument>--precision high</jsdl-p:Argument>
          </jsdl-p:POSIXApplication>
        </jsdl:Application>
      </jsdl:JobDescription>
    </s:JSDL>
  </s:Activity>
</s:Workflow>
```

As this is just string replacement, variables can be used anywhere in the workflow, including application definitions, file staging etc.

10 Batch processing

The *batch* command allows you to run many jobs without having to start UCC each time. You can control how many jobs should go to which site. This allows efficient job processing, while putting some load on the client machine. If you need to take the client offline, you should consider using the workflow system instead, which also allows efficient high-throughput processing.

Assume you have a bunch of jobs in UCC's job description format (Section 6) stored in a directory *jobs*. The output should go to a directory *out*. You can run them all through UCC using a single invocation as follows:

ucc batch -i jobs -o out

As job files, UCC will accept files ending in ".u"

10.1 Options

You can run in "follow" mode, where UCC will watch the input directory, and will process new files as they arrive.

ucc batch -f -i jobs -o out

UCC can also process JSDL files, to batch-process these, use the "-j" option:

ucc batch -j -i jobs -o out

10.2 Performance tuning options

Getting the most performance out of UCC and the UNICORE installation can be a challenging task. Sending too many jobs to a site might decrease throughput, sometimes the client machine can be the limiting factor, etc.

You should experiment a bit to get the best performance for your specific setup. UCC has many options available for tuning. Here is an overview.

Option (short and long	Description			
form)				
-K,keep	Do not delete finished jobs on the server. By default,			
	finished jobs are destroyed.			
-m,max	Limit on jobs submitted by UCC at one time (default: 100)			
<maxrunningjobs></maxrunningjobs>				
-t,threads	Number of threads to be used for processing (default: 4)			
<numthreads></numthreads>				
-u,update	Minimum time in milliseconds between status requests on a			
<updateinterval></updateinterval>	single job (Default: 1000)			
-R,noResourceCheclDo not check if the necessary application is available on the				
	target system (will increase performance a bit)			
-X,noFetchOutcome Do not fetch standard output and error				
-S,submitOnly	Only submit the jobs, do not wait for them to finish			
-M,maxNewJobs	Limit the number of job submissions (default: 100)			
-s,sitename	Specify which site to use			
-W,siteWeights	Specify a file containing site weights			

Table 12: Tuning options for the UCC batch mode

10.3 Resource selection in batch mode

By default, the UCC batch mode will select a random site for running a job. You can modify the selection in different ways.

- using the "-s" option or a "Site: <sitename>," entry in the job file, you can specify the site directly
- use the "-W" option to specify a file containing site weights.

Say you have two sites where one site is a big cluster and the other a small cluster. To send more jobs to the big cluster, you can use the site weights file,

```
#example site weights file for use with "ucc batch -W ..."
BIG-CLUSTER = 100
SMALL-CLUSTER = 10
#send no jobs to this site
BAD-CLUSTER = 0
# set default weight (for any sites not specified here)
UCC_DEFAULT_SITE_WEIGHT = 10
```

This would tell UCC to send 10 times more jobs to the "BIG-CLUSTER" site, and send no jobs to the "BAD-CLUSTER". All other sites would get weight "10", i.e. the same as "SMALL-CLUSTER".

11 The UCC shell

If you want to run a larger number of UCC commands, the overhead of starting the Java VM or checking the registry may become annoying. For this scenario, UCC offers a "shell" that allows the user to enter UCC commands interactively.

It is started by

ucc shell <options>

If you want to process a list of commands from a file instead of typing them, you can start the shell like this

ucc shell -f commandsfile

or on Unix you can use the redirection features

ucc shell < commandsfile

11.1 Changing property settings

To change a property setting in shell mode, you can use the *set* command. Without additional arguments, current properties are listed:

```
ucc>set
registry=https://...
output=/tmp
...
```

To set one or more properties, add space separated key=value strings:

```
ucc>set output=/work registry=https://....
```

You can also clear a property (set it to null) by using unset

```
ucc>unset registry
```

commands. You can use to make commands shorter and more readable. It's also useful to pre-set certain things in your preferences file.

For example

```
ucc> set S1=https://myserver/my_site/rest/core/storages/HOME
ucc> ls -l ${S1}
```

11.2 Running an external command

You can run an external command via the "system" shell command. For example

ucc> system vi job.u

11.3 Exiting the shell

To exit, type exit or press CTRL-D

12 Sharing resources

Accessing UNICORE resources (jobs, storages, ...) is usually only possible when you "own" the resource or when there are special server-side policies in place that allow you access.

Starting with server version 7.3, UNICORE supports ACLs on a per-service instance basis. This means, that you can give other users access to your target systems, jobs, storages,

For example, you might have access to an S3 cloud storage via UNICORE, and you want to securely share data on this resource. Or, you want to allow others to check job status, or even allow them to abort jobs.

Note that to access actual **files** the permissions on file system level still need to match. Usually this is achieved by using Unix groups.

12.1 Editing ACLs

The ACLs are managed via the "share" command. Use the basic

```
ucc share <URL>
```

to showe the current ACL for the given resource, where "URL" is the full WSRF service URL of the resource, e.g.

ucc share https://localhost:8080/DEMO-SITE/rest/core/storages/HOME

To add an ACL entry, use

ucc share ACE1 ACE2 ... <URL>

where "ACE" is an access control entry expressed in a simple format:

```
[read|modify]:[DN|VO|GROUP|UID]:[value]
```

For example to give "modify" permission to a user whose UNIX user id on the target system is "test", you would use

ucc share modify:UID:test <URL>

To delete entries, use the "-d" option

ucc share -d modify:UID:test <URL>

To delete all entries, use the "-b" option

ucc share -b <URL>

12.2 Permission levels

The permissions controlled by ACLs are as follows

- read : access resource properties
- modify : perform actions e.g. job submission or creating a file export

Only the owner of a resource can edit the ACL or destroy the resource.

13 UCC for site administrators

UCC can be used for administrative and user support tasks, like checking server status, or getting the full details of a user job.

13.1 Security considerations

Usually, each UNICORE user has only access to his or her own resources (such as jobs). For administrative use, you will need to aquire administrator privileges. There are two ways to achieve this.

- create dedicated user credentials (e.g. a certificate) and map them to the role "admin" (in the XUUDB, or whatever attribute source you are using). This method is recommended if you want to remotely administrate UNICORE/X.
- use the server keystore (of the UNICORE/X server you want to administrate) as UCC keystore. This will also give you administrator privileges. For this you will need to be logged on to the UNICORE/X server.

13.2 Admin commands

UCC has dedicated commands for accessing the "AdminService" of a UNICORE/X container. To get started, try

```
ucc admin-info -l
```

UCC will try to access the admin service on each availabe UNICORE/X server. For each server, a list of statistical and performance data will be listed.

It will also list the available admin commands for each server, with a short description of their parameters. For example, here is a sample output:

```
default_admin
  Services:
   TargetSystemFactoryService[1]
    . . .
 Monitors:
   use.externalConnectionStatus.REST_UnitySAMLAuthenticator: OK
    use.security.overview: ServerIdentity: CN=Demo UNICORE/X,O= ↔
       UNICORE, C=EU; Expires: Thu Sep 09 12:01:19 CEST 2032; ←
       IssuedBy: CN=Demo CA,O=UNICORE,C=EU
    . . . .
 Metrics:
   use.externalConnectionStatus.REST_UnitySAMLAuthenticator: OK
   use.rest.callFrequency: 0.016677196376660174
    . . .
  Available commands:
    ShowJobDetails : parameters: jobID, [xnjsReference]
    ShowServerUsageOverview : parameters: [clientDN]
    ToggleResourceAvailability : 'resources' - comma separated list \leftrightarrow
        of IDs
   ToggleJobSubmission : parameters: [message]
    ToggleBESJobSubmission :
```

To invoke a command, the "admin-runcommand" is used. It can take optional parameters.

13.2.1 Disabling/enabling job submission

For example, it is possible to disable/enable job submission to the server, using the *ToggleJob-Submission* command, which can take an optional message:

ucc admin-runcommand ToggleJobSubmission message="Maintenance"

The service will reply:

SUCCESS, service reply: OK - job submission is disabled

If a user now tries to submit, she will receive an error message on submission. Running the command again will re-enable the service.

ucc admin-runcommand ToggleJobSubmission message="Maintenance" SUCCESS, service reply: OK - job submission is now enabled

13.2.2 Getting job details

To get the full job details (for example in user support), try

ucc admin-runcommand ShowJobDetails jobID=<unique_jobid>

for example

```
ucc admin-runcommand ShowJobDetails jobID=cdfdafc5-0274-464d-ac4a ↔
    -463f46c942fa
SUCCESS, service reply: Job information for cdfdafc5-0274-464d-ac4a \leftrightarrow
    -463f46c942fa
{Info=Action ID
                      : cdfdafc5-0274-464d-ac4a-463f46c942fa
              : JSDL
Action type
Status
               : DONE (trans.: none)
Result
               : SUCCESSFUL [Success.]
Owner
                : CN=Demo User, O=UNICORE, C=EU
Exec. Definition: <JobDefinition xmlns="http://schemas.ggf.org/jsdl \leftrightarrow
   /2005/11/jsdl">
  <JobDescription>
    <JobIdentification>
      <JobName>Date</JobName>
    </JobIdentification>
    <Application>
      <jsdl:POSIXApplication xmlns:jsdl="http://schemas.ggf.org/ ↔
          jsdl/2005/11/jsdl-posix">
        <jsdl:Executable>/bin/date</jsdl:Executable>
      </jsdl:POSIXApplication>
```

```
</Application>
    <Resources/>
  </JobDescription>
</JobDefinition>
Orig. Definition: <JobDefinition xmlns="http://schemas.ggf.org/jsdl ↔
    /2005/11/jsdl">
  <JobDescription>
    <JobIdentification>
      <JobName>Date</JobName>
    </JobIdentification>
    <Application>
      <ApplicationName>Date</ApplicationName>
      <POSIXApplication xmlns="http://schemas.ggf.org/jsdl/2005/11/ <->
          jsdl-posix"/>
    </Application>
    <Resources/>
  </JobDescription>
</JobDefinition>
Processing context: de.fzj.unicore.xnjs.ems. ←
   ProcessingContext@4308cff0
Application Info: AppInfo for Date 1.0
Job log:
Thu Sep 08 09:25:03 CEST 2016: Created with ID cdfdafc5-0274-464d- ↔
   ac4a-463f46c942fa
Thu Sep 08 09:25:03 CEST 2016: Created with type 'JSDL'
Thu Sep 08 09:25:03 CEST 2016: Client: Name: CN=Demo User,O=UNICORE ↔
   ,C=EU
Xlogin: uid: [schuller], gids: [addingOSgroups: true]
Role: user: role from attribute source
Security tokens: User name: CN=Demo User,O=UNICORE,C=EU
Consignor DN: CN=Demo User,O=UNICORE,C=EU
Delegation to consignor status: true, core delegation status: true
Message signature status: UNCHECKED
Client's original IP: 127.0.0.1
Thu Sep 08 09:25:04 CEST 2016: Using default execution environment.
Thu Sep 08 09:25:04 CEST 2016: No staging in needed.
Thu Sep 08 09:25:04 CEST 2016: Status set to READY.
Thu Sep 08 09:25:04 CEST 2016: Status set to PENDING.
Thu Sep 08 09:25:04 CEST 2016: Incarnated resources: [CPUsPerNode ↔
   =1.0, MemoryPerNode=2.68435456E8, ArraySize=1, Nodes=1.0, <
   ArrayLimit=64, CPUTime=3600.0]
Thu Sep 08 09:25:04 CEST 2016: Command is:
#!/bin/sh
# ....
chmod u+x /bin/date 2> /dev/null
rm -f /opt/unicore-servers/FILESPACE/cdfdafc5-0274-464d-ac4a-463 ↔
   f46c942fa//UNICORE_SCRIPT_EXIT_CODE
 /bin/date
```

echo \$? > /opt/unicore-servers/FILESPACE/cdfdafc5-0274-464d-ac4a ↔

```
-463f46c942fa//UNICORE_SCRIPT_EXIT_CODE

Thu Sep 08 09:25:04 CEST 2016: TSI reply: submission OK.

Thu Sep 08 09:25:04 CEST 2016: Submitted to classic TSI as [ ↔

schuller NONE] with BSSID=3269519504309

Thu Sep 08 09:25:13 CEST 2016: Exit code 0

Thu Sep 08 09:25:13 CEST 2016: Job completed on BSS.

Thu Sep 08 09:25:14 CEST 2016: Status set to DONE.

Thu Sep 08 09:25:14 CEST 2016: Result: Success.

Thu Sep 08 09:25:14 CEST 2016: Total: 10.01 sec., Stage-in: 0.05 ↔

sec., Stage-out: 0.00 sec., Datamovement: 0 %}
```

Thus you can get a full view of what the user submitted and what was executed.

13.3 Listing jobs, sites, ...

You can also use all normal UCC commands to access the server. Note however that due to the authentication and authorisation system in UNICORE, this may not always work as expected: the "admin" user might not have the required Unix permissions to access files, list directories etc.

The UCC commands that list server-side things (list-jobs etc) accept a filtering option, that can be used to limit the results of the operation. Filtering works on the XML resource properties of the resource in question.

Filtering is enabled by the "-f" or "--filter" option of the form

-f XMLNAME OPERATOR VALUE

where XMLNAME is the name of an XML Element from the WSRF resource properties document.

For example, to list all running jobs:

ucc list-jobs -f Status equals RUNNING

To list all jobs submitted on Nov 13, 2007:

ucc list-jobs -f SubmissionTime contains 2007-11-13

etc.

Table 13: Filtering options

Operator (long and	Description
short form)	
equals, eq	String equality (ignoring case)

Operator (long and	Description
short form)	
notequals, neq	String inequality (ignoring case)
contains, c	Substring match
notcontains, nc	substring non-match
greaterthan, gt	Lexical comparison
lessthan, lt	Lexical comparison

Table 13: (continued)

13.4 Low-level operations

UCC supports several low-level operations using the "wsrf" command.

To destroy a resource,

ucc wsrf destroy <Address>

To get a complete property listing (i.e. print the XML resource property document)

ucc wsrf getproperties <Address>

To extend the lifetime of a resource

ucc wsrf extend <Address> <Days>

These commands can be abbreviated, e.g. + ucc wsrf d <Address>

14 Scripting

UCC can execute Groovy scripts. Groovy (http://groovy.codehaus.org) is a dynamic scripting language similar to Python or Ruby, but very closely integrated with Java. The scripting facility can be used for automation tasks or implementation of custom commands, but it needs a bit of insight into how UNICORE and UCC work.

14.1 Script context

Your Groovy scripts can access some predefined variables that are summarized in the following table

variable	description	Java type
registry	A preconfigured client for	de.fzj.unicore.uas.client.IRegistryQuery
	accessing the registry	
configurationProvider	Security configuration	de.fzj.unicore.ucc.authn.UCCConfigurationProvider
	provider (keystore, etc)	
registryURL	the URL of the registry	java.lang.String
messageWriter	for writing messages to the	de.fzj.unicore.ucc.MessageWriter
	user	
commandLine	the command line	org.apache.commons.cli.CommandLine
properties	defaults from the user's	java.util.Properties
	properties file	

Table 14: Variables accessible for scripts	Table 14	1: Variables	accessible	for	scripts
--	----------	--------------	------------	-----	---------

14.2 Examples

Some example Groovy scripts can be found in the samples/ directory of the UCC distribution.

Here is a script that will delete all your finished jobs (use at your own risk):

```
Groovy example: delete all your jobs
```

```
/*
* remove all jobs
*/
//import UNICORE/X client classes
import de.fzj.unicore.uas.client.*;
import eu.unicore.security.wsutil.client.authn. \hookleftarrow
    DelegationSpecification
//helper: kills a job with the given status (SUCCESSFUL, RUNNING, \, \leftrightarrow \,
    ...)
def kill(job, statuscode) {
   if (job.status.toString() == statuscode)job.destroy()
}
// {\tt iterate} over TSSs and remove all jobs
def lister = new de.fzj.unicore.uas.lookup.SiteLister(registry, \leftarrow
    configurationProvider)
lister.each {
    it.jobs.each{
        messageWriter.message "Job at "+it.address.stringValue
```

```
securityProperties = configurationProvider. ↔
   getClientConfiguration(it,DelegationSpecification. ↔
   STANDARD)
  kill(new JobClient(it, securityProperties),"SUCCESSFUL")
}
```

Groovy example: list available storages

}

```
/*
* list available storages
*/
import de.fzj.unicore.uas.client.*
import eu.unicore.security.wsutil.client.authn. ↔
   DelegationSpecification
import javax.xml.namespace.QName
//porttype of storage service
def SMSPORT=new QName("http://unigrids.org/2006/04/services/sms"," ↔
   StorageManagement")
//method to extract storage name from a storage client
def findName(epr) {
  securityProperties = configurationProvider.getClientConfiguration ↔
      (epr, DelegationSpecification.STANDARD)
  sms=new StorageClient(epr, securityProperties)
  return sms.resourcePropertiesDocument.storageProperties. ↔
     fileSystem.name
}
//list storages from registry
registry.listAccessibleServices(SMSPORT).each {
    name=findName(it)
    messageWriter.message "Storage <"+name+"> at "+it.address. ↔
       stringValue
}
//list storages attached to target systems
def lister = new de.fzj.unicore.uas.lookup.SiteLister(registry, \leftrightarrow
   configurationProvider)
lister.each {
    it.storages.each {
        name=findName(it)
           messageWriter.message "Storage <"+name+"> at "+it. ↔
                address.stringValue
    }
}
```

15 Frequently asked questions

15.1 Configuration

15.1.1 Do I really have to store my password in the preferences file? Isn't this insecure?

Putting the password in a file or giving it as a commandline parameter can be considered insecure. The file could be read by others, and the commandline parameters may be visible in for example in the output of the *ps* command. Thus, UCC will simply ask for the password in case you did not specify it.

15.1.2 How can I enable more detailed logging?

UCC uses log4j, by default the configuration is done in <UCC_HOME>/conf/logging.properties You can edit this file and increase the logging levels, choose to log to a file or to the console, etc.

15.2 Usage

15.2.1 Can I use multiple registries with UCC?

Yes. Simply use a comma-separated list of URLs for the "-c" option. However, you may only use a single key/truststore, so all registries (and sites listed in them) must accept the same security credentials.

15.2.2 Can I upload and execute my own executable?

Yes. Check Section 5.

15.2.3 Can I use UCC to list the contents of the registry?

Using the *rest* command (and the UNIX *jq* utility for formatting the output), this is very easy, for example

```
ucc rest get https://localhost:8080/DEMO-SITE/rest/core/registries/ ↔ default_registry | jq
```

will list the content of the registry.

15.2.4 I get strange errors related to security

Please read the general UNICORE FAQ on www.unicore.eu[the UNICORE website] which contains descriptions of many common errors.