

Nutzung SuGI-Cluster



Sebastian Breuers
eMail: s.breuers@uni-koeln.de

Inhalt

Einleitung.....	- 1 -
Besorgen eines D-Grid-Zertifikats	- 1 -
Anmeldung in VO lifescience	- 1 -
Nutzung des Clusters.....	- 1 -
Nutzung des Globus-Klienten.....	- 2 -
Nutzerzertifikat	- 2 -
Erstellen eines Proxy-Zertifikats.....	- 2 -
Submittieren eines Jobs an PBS/Torque Maui.....	- 3 -
xml-Beschreibung.....	- 3 -
Abschicken des Jobs	- 4 -
Terminationtime eines Jobs nachträglich ändern.....	- 5 -
Statuskontrolle und Steuerung	- 5 -
Nutzung der environment modules	- 5 -
Anzeigen der vorhandenen Module	- 6 -
Laden eines Moduls.....	- 6 -
Anzeigen der geladenen Module	- 6 -
Entladen eines Moduls.....	- 7 -
Entladen aller Module.....	- 7 -



Einleitung

Im Rahmen einer Sonderinvestition des Projektes SuGI (Sustainable Grid Infrastructure) wurden im Jahr 2007 drei Cluster an den Standorten Kaiserslautern, Kassel und Köln beschafft. Alle drei sind aktuell mit dem Klienten der Middleware Globus ToolKit 4 nutzbar. Nähere Informationen zur Hardware und zu den Zugangsdaten finden sich auf den Seiten des SuGI-Portals:

<http://portaldevel.sugi.uni-koeln.de/de/ueber-sugi/sugi-cluster-im-d-grid.html>

Um die Rechenleistung dieser Cluster nutzen zu können, müssen einige Vorbereitungen getroffen werden, die im Folgenden beschrieben sind.

Beantragen eines D-Grid-Zertifikats

Ein essentieller Schritt zur Nutzung von Ressourcen im Grid ist die Beschaffung eines D-Grid-Zertifikats. Informationen zur Anschaffung eines D-Grid-Zertifikats finden sich auf den Seiten des D-Grids:

<http://www.d-grid.de/index.php?id=243>

Über die Seite

<https://gridka-ca-sec.fzk.de/>

kann unter der Rubrik "**Erstantrag Persönliches Zertifikat**" ein D-Grid-Zertifikat beantragt werden. Bei der Organisation sollte die eigene Institution ausgewählt werden.

Nach Bearbeitung des Antrages werden Sie aufgefordert, sich als die angegebene Person persönlich auszuweisen. Nähere Informationen dazu werden Sie in der Aufforderungsemail Ihrer lokalen Registration Authority (RA) erhalten.

Anmeldung in VO lifescience

Grid-Nutzer sind häufig in virtuellen Zusammenschlüssen organisiert, den Virtuellen Organisationen (VO). Im Rahmen des SuGI-Projektes ist die VO lifescience gegründet worden. Mitglieder dieser VO sind berechtigt, die Cluster des SuGI-Projektes zu nutzen.

Um einer VO beizutreten, muss eine Anmeldung über die folgende Seite beim VO Management Registration Service (VOMRS) in Jülich erfolgen

<https://dispatch.fz-juelich.de:8814/D-Grid-VO-Member>

Für die Anmeldung muss bereits ein signiertes D-Grid-Zertifikat vorliegen und im Browser importiert sein.

In der Liste auf der angegebenen Seite muss bei der VO lifescience auf den Link "**member registration**" geklickt werden. Auf der folgenden Seite sollte den Anweisungen unter "**Registration (Phase I)**" gefolgt werden. Das Formular muss vollständig ausgefüllt und anschließend submittiert werden.

Nutzung des Clusters

Die Universität zu Köln stellt ihren Nutzern einen bereits vollständig eingerichteten Globus ToolKit 4-Klient-Rechner zur Verfügung. Globus ToolKit 4 ist eine Grid-Middleware, die zwischen dem Nutzer und den Ressourcen im Grid vermittelt. Der Rechner ist unter der Adresse

`grid.sugi.uni-koeln.de`

erreichbar. Um diesen Rechner für sich nutzen zu können, ist es nötig sich mit seinem UniNV-Account einzuloggen. Dieser ist für diesen Rechner freigeschaltet, wenn man sich mit

seinem Zertifikat bei der VO *lifescience* registriert hat. Nach dem Einloggen steht das UniNV-HPC-Homeverzeichnis zur Verfügung.

Nutzer, die keine Mitglieder der Universität zu Köln sind, haben die Möglichkeit, einen Globus-Toolkit-Klienten in Form einer virtuellen Maschine von den Seiten des SuGI-Portals herunterzuladen:

portal.sugi.uni-koeln.de

Nutzung des Globus-Klienten

Zur Nutzung des Globus-Klienten sind einige Konfigurationsschritte notwendig, wie z.B. die globus-konforme Lage der Zertifikatsinformationen.

Nutzerzertifikat

Sinnvoll genutzt werden kann der Klient erst, wenn den Komponenten des Globus-Klienten ein Nutzerzertifikat zur Verfügung steht. Dazu muss es unter einem bestimmten Namen und einer bestimmten Form im Heimverzeichnisunterordner `.globus` abgelegt werden. Falls er nicht existiert sollte er mit

```
mkdir .globus
```

angelegt werden. Achtung: Zu dem Verzeichnisnamen gehört der führende Punkt.

Das Zertifikat muss entweder als pkcs12-Keystore unter

```
~/ .globus/usercred.p12
```

abgelegt werden oder als privater und öffentlicher Schlüssel des Zertifikats unter

```
~/ .globus/usercert.pem
```

und

```
~/ .globus/userkey.pem
```

In allen Fällen dürfen die Zertifikatsdateien `usercred.p12`, `usercert.pem` und `userkey.pem` nur für den Nutzer lesbar sein, ansonsten gibt `grid-proxy-init` eine Fehlermeldung aus:

```
chmod 400 ~/ .globus/user*
```

Erstellen eines Proxy-Zertifikats

Durch Ausführen des Programms

```
grid-proxy-init
```

wird ein Proxy-Zertifikat erstellt. Dabei schaut das Programm im Verzeichnis `.globus` entweder nach dem pkcs12-Keystore `usercred.p12` oder nach dem privaten und öffentlichen Schlüsselpaar `usercert.pem` und `userkey.pem`. Findet es eines der beiden Zertifikate verlangt es entweder nach dem Exportpasswort des Keystores oder nach dem Passwort des privaten Schlüssels. Ohne weitere Angaben ist das Proxy-Zertifikat 12 Stunden gültig.

Die Gültigkeit kann mit `-valid hh:mm` bis auf die maximale Dauer, d.h. die Gültigkeit des D-Grid-Zertifikats, verlängert werden.

```
grid-proxy-init -valid 72:00
```

erzeugt also ein Proxy-Zertifikat, das 3 Tage gültig ist.

Mit dem Programm

```
grid-proxy-info
```

kann der Status des eigenen Proxy-Zertifikats abgefragt werden.

Submittieren eines Jobs an PBS/Torque Maui

Die beiden Ressourcen in Kassel und Köln verwenden PBS/Torque Maui als Batch-Scheduler-System. Zum Submittieren eines Jobs wird im Grid nicht das gewohnte `qsub` verwendet, sondern der Befehl

```
globusrun-ws -submit
```

Dieses Programm submittiert den Job in einer abstrakten Jobbeschreibung an den Globus-server, der ihn dann in einen konkreten PBS-Batchjob übersetzt.

Dabei wird das ursprüngliche Skript in ein anderes Skript verpackt. In diesem Wrapperskript stehen gesonderte PBS-Direktiven. Durch die Verpackung werden Ressourcenanforderungen, die über im eigenen Skript bereits vorhandene PBS-Direktiven an das Batchsystem vermittelt werden sollen, ignoriert. Um weiterhin die Prozessor- und Knotensteuerung nutzen zu können, muss das zu submittierende Skript in eine xml-Beschreibung eingefügt werden. Im Folgenden ist eine Demoversion für eine solche xml-Datei beschrieben.

xml-Beschreibung

```
<job>
  <!-- Name des auszuführenden Skriptes -->
  <executable>skriptname</executable>

  <!-- Achtung hier wird der absolute Pfad auf dem Remote-
  rechner verlangt; eine flexiblere und portablere
  Lösung ist die Angabe in Abhängigkeit vom Heimver-
  zeichnis im Grid, das mit der Variablen
  ${GLOBUS_USER_HOME}
  angegeben werden kann -->
  <directory>/absolute/path/to/script/directory</directory>
  <argument>potential arguments to executable</argument>

  <!-- Achtung hier wird eine Datei mit absolutem Pfad auf
  dem Remoterechner verlangt; dir Verwendung von
  ${GLOBUS_USER_HOME}
  wird empfohlen. s.o. -->
  <stdout>/absolute/path/to/stdout/file</stdout>

  <!-- Achtung hier wird eine Datei mit absolutem Pfad auf
  dem Remoterechner verlangt Verwendung von
  ${GLOBUS_USER_HOME}
  empfohlen. s.o. -->
  <stderr>/absolute/path/to/stderr/file</stderr>

  <queue>default</queue>

  <maxWallTime>5</maxWallTime> <!-- integer in minutes -->

  <maxMemory>1000</maxMemory> <!-- integer in megabytes -->

  <!-- obligatorisch -->
  <jobType>single</jobType>
  <extensions>
    <!-- arbitrary jobname within batch system
```

```

        (optional)-->
        <name>jobname</name>
        <!-- PBS like alternative to <hostCount>
            and <cpusPerHost> -->
        <!--nodes> 2:ppn=2 </nodes-->
        <resourceAllocationGroup>
            <hostCount>2</hostCount>
            <cpusPerHost>2</cpusPerHost>
        </resourceAllocationGroup>
    </extensions>
</job>

```

Weitergehende Informationen zum **Job Description Scheme** finden sich auf folgender Seite http://www.globus.org/toolkit/docs/4.0/execution/wsgram/schemas/gram_job_description.html

Alternativ zu dem Element resourceAllocationGroup der Beschreibung kann auch das nodes-Element verwendet werden. Hierzu empfiehlt sich die Lektüre des Job Description Extension Supports des WS-GRAM

[http://www.globus.org/toolkit/docs/4.0/execution/wsgram/WS GRAM Job Desc Extensions.html](http://www.globus.org/toolkit/docs/4.0/execution/wsgram/WS_GRAM_Job_Desc_Extensions.html)

Ein Beispiel für diese Verwendung ist bereits mit den Kommentarzeichen eingefügt worden. Achtung: Die beschriebene Variante funktioniert nur, wenn das auszuführende Skript das Executable-Flag gesetzt hat.

Abschicken des Jobs

Mit dem Befehl

```

globusrun-ws -submit -factory $FACTORY \
-factory-type $FACTTYPE -f job.xml -o epr.file \
-term 'MM/DD/YY hh:mm' &

```

oder kurz

```

globusrun-ws -submit -F $FACTORY -Ft $FACTTYPE \
-f job.xml -o epr.file -term 'MM/DD/YY hh:mm' &

```

kann die Jobbeschreibung in der xml-Datei an den Globus ToolKit 4-Serverrechner geschickt werden.

- F steht dabei für den Grid-Rechner, auf den submittiert werden soll
- Ft gibt an, wie der Job weiterverarbeitet werden soll, hier in PBS-Anweisungen umschreiben.
- f gibt an, dass eine xml-basierte, abstrakte Jobbeschreibung verschickt wird
- o gibt den Namen der Endpoint Ressource (EPR)-Datei an. Über sie besteht direkte Verbindung zum Job im Grid
- term gibt den Endzeitpunkt an, bis zu dem der Job gültig ist. Diese Zeit sollte mit einer gewissen Toleranz berechnet werden und etwaige Wartezeiten im Batchsystem und die Laufzeit des Jobs berücksichtigen Falls dieser Wert nicht angegeben wird, vergibt Globus Toolkit einen Defaultwert von einem Tag. Bei einer zu kurz gewählten Zeit wird der Job abgebrochen.

Mit Hilfe des Befehls:

```

date "+%m/%d/%Y %k:%M" --date='4 years'

```

kann das Datum in der benötigten Form berechnet werden. Dabei wird das Datum in 4 Jahren ermittelt. Andere Angaben wie

```
date "+%m/%d/%Y %k:%M" --date='2 months'
```

für "in 2 Monaten" sind auch möglich.

Zulässige Werte für \$FACTORY und \$FACTTYPE im SuGI-Cluster-Verbund sind

Standort	\$FACTORY	\$FACTTYPE
Kaiserslautern	gt4.sugi.uni-kl.de	SGE
Kassel	sugi-ks.hrz.uni-kassel.de	PBS
Köln	submit.sugi.uni-koeln.de	PBS

Terminationtime eines Jobs nachträglich ändern

Wurde die Life- bzw. Terminationtime nicht korrekt gesetzt, kann sie mit dem folgenden Befehl:

```
wsrf-set-termination-time -e job.epr 604800
```

korrigiert werden. Dabei ist `job.epr` die mit `-o` beim submit erzeugte Datei und `604800` die Anzahl der Sekunden, die der Job vom aktuellen Zeitpunkt ab gültig sein soll. Alternativ kann auch das Schlüsselwort 'infinity' verwendet werden

```
wsrf-set-termination-time -e job.epr 'infinity'
```

Statuskontrolle und Steuerung

Unter der Verwendung der End Point Resource-Datei, die beim Abschicken des Jobs mit

```
globusrun-ws -submit -o epr.file
```

erzeugt wurde, kann der Status der Jobs im Grid abgefragt bzw. kontrolliert werden. Mit

```
globusrun-ws -status -j epr.file
```

oder

```
globusrun-ws -monitor -j epr.file
```

kann der Status des Jobs abgefragt werden und mit dem Job nachträglich interagiert werden.

```
globusrun-ws -kill -j epr.file
```

beendet den Job.

Nutzung der *environment modules*

Über die `environment modules` werden auf den SuGI-Clustern die Umgebungsvariablen für bestimmte Tools gesetzt und die Programm- (`$PATH`) und Bibliothekspfade (`$LD_LIBRARY_PATH`) für das jeweilige Tool angepasst.

Auf die Funktionen der `environment modules`, die von Ihrem Ressourcenprovider für Sie eingerichtet und konfiguriert wurden, kann mit dem Programm `module` zugegriffen werden.

Um das Vorhandensein der `module` Konfigurationspfade zu gewährleisten, gibt es zwei Optionen:

1. Als Shebang Programm kann in dem auszuführenden Skript

```
#!/bin/bash -l
```

genutzt werden, um eine Login-Shell zu starten. Dabei wird dann auch die Umgebung für die Modules richtig gesetzt.

2. Im auszuführenden Skript wird die Initialisierung durchgeführt, indem `/etc/profile.d/modules.sh` gesourcet wird:

```
. /etc/profile.d/modules.sh
```

Im Folgenden werden einige häufig benutzte Befehle für die `environment modules` erläutert.

Anzeigen der vorhandenen Module

Mit dem Befehl

```
module avail
```

werden die zur Verfügung stehenden Module angezeigt

Beispielausgabe

```
----- /opt/rrzk/modulefiles -----  
gnu-compilers/4.2.0                pathscale-compilers/3.1  
gnu-compilers/system(default)      pgi-compilers/5.1-6  
intel-compilers/10.0               pgi-compilers/5.2-4
```

Die Ausgabe zeigt die vorhandenen Modulnamen und die zur Verfügung stehenden Versionen.

Laden eines Moduls

Mit dem Befehl

```
module load '<modulename>'
```

kann ein Modul geladen werden. Wird der Aufruf ohne eine Versionsangabe gemacht, wird automatisch die alphabetisch höchste genommen, das ist bei der verwendeten Versionsnamenkonvention die aktuellste.

Beispielaufruf

```
module load gnu-compilers
```

Laden eines Moduls in einer bestimmten Version

Mit dem Befehl

```
module load '<modulename>/<version>'
```

kann ein Module in einer bestimmten Version geladen werden. Diese Version muss vorhanden sein und kann der Ausgabe von `module avail` entnommen werden.

Beispielaufruf

```
module load gnu-compilers/4.2.0
```

Anzeigen der geladenen Module

Mit dem Befehl

```
module list
```

kann die Liste der geladenen Module angezeigt werden.

Beispielausgabe

```
Currently Loaded Modulefiles:  
1) gnu-compilers/4.2.0
```

Entladen eines Moduls

Um die Änderungen, die durch das Laden eines Moduls entstanden sind, wieder zu entfernen wird der Befehl

```
module unload 'modulename'
```

verwendet. Die Angabe einer Versionsnummer ist hier nicht zwingend erforderlich.

Beispielaufruf

```
module unload gnu-compilers
```

Entladen aller Module

Mit dem Befehl

```
module clear
```

und der anschließenden Bestätigung der Eingabeaufforderung mit 'y' können alle Module wieder entladen werden.

