

ログインノード/バッチジョブからグリッド上のファイルにアクセスする方法

ログインノード/バッチジョブからグリッド上のファイルにアクセスする方法

グリッド上のファイルはグリッドジョブ(pathena, prun等で投入)からアクセスしたり、`rucio download`でローカルにダウンロードすることが基本ですが、ログインノード(login.icepp.jp、lxplus.cern.ch)やバッチジョブ(地域解析センターのHTCondor)からアクセスすることもできます。

ATLASLOCALGROUPDISK上のファイルの一覧の取得

まずはグリッド上のファイルがどのグリッドサイトのどのディレクトリに保存されているかを確認する必要があります。 `rucio list-file-replicas DATASET_NAME` でファイルの保存場所を確認できます。

```
$ export  
ATLAS_LOCAL_ROOT_BASE=/cvmfs/atlas.cern.ch/repo/AtlasLocalRootBase  
$ source ${ATLAS_LOCAL_ROOT_BASE}/user/atlasLocalSetup.sh  
$ lsetup rucio  
$ rucio list-file-rePLICas  
user.masaito.data17_13TeV.00340072.physics_Main.merge.DAOD_SUSY6.r10426_  
+-----+  
- - - + - - - - - + - - - - - + -  
- - - - -  
- - - - -  
- - - - -  
- - - - - +  
| SCOPE          | NAME  
| FILESIZE      | ADLER32    | RSE: REPLICA  
|  
| -----+-----  
- - - - - + - - - - - + - - - - - + -  
- - - - -  
- - - - -  
- - - - -  
- - - - - |  
| user.masaito |  
user.masaito.17346017.EXT0._000136.DAOD_SUSY6.test.pool.root |  
322.901 MB | 68d4e998 | TOKYO-LCG2_LOCALGROUPDISK: davs://lcg-  
se01.icpp.jp:18443/dpm/icpp.jp/home/atlas/atlaslocalgroupdisk/rucio/us  
|
```

上のコマンド出力例だと、lcg-

```
se01.icepp.jp:18443/dpm/icepp.jp/home/atlas/atlaslocalgroupdisk/rucio/user/masaito/8f/9c/user.masaito.17346017.EXT0._000136.DAOD_SUSY6.test.pool.root がファイルの保存場所です。
```

プロキシ証明書の準備と設置

グリッド上のファイルにアクセスするためには適切なグリッド(プロキシ)証明書が必要になります。ログインノードからグリッド上のファイルにアクセスしたいときは以下のようにしてプロキシ証明書を作成できます。

```
$ export
ATLAS_LOCAL_ROOT_BASE=/cvmfs/atlas.cern.ch/repo/ATLASLocalRootBase
$ source ${ATLAS_LOCAL_ROOT_BASE}/user/atlasLocalSetup.sh
$ lsetup emi
$ voms-proxy-init -voms atlas
```

defaultではプロキシ証明書は /tmp/x509up_u{UID} に作成されます。

バッチジョブからグリッド上のファイルにアクセスしたいときは注意が必要です。通常バッチジョブからログインノード上の'/tmp/'にアクセスできないためです。グリッド証明書にアクセスできるように、バッチノードが見える場所(例えば \$HOME 以下)に証明書を設置する必要があります。ログインノード上で以下のようなコマンドを実行することでプロキシ証明書が \$HOME/.globus/gridproxy.cert に設置されます。

```
$ export
ATLAS_LOCAL_ROOT_BASE=/cvmfs/atlas.cern.ch/repo/ATLASLocalRootBase
$ source ${ATLAS_LOCAL_ROOT_BASE}/user/atlasLocalSetup.sh
$ lsetup emi
$ voms-proxy-init -voms atlas -out $HOME/.globus/gridproxy.cert
```

この証明書をバッチジョブの中で環境変数として指定することで、バッチジョブ中でGRID上のファイルにアクセスできるようになります。

```
$ export X509_USER_PROXY=${HOME}/.globus/gridproxy.cert
$ export
X509_CERT_DIR=/cvmfs/atlas.cern.ch/repo/ATLASLocalRootBase/etc/grid-security-emi/certificates
```

プロキシ証明書は他のユーザーに悪用されないようにパーミッション設定に注意する必要があります。上記の方法で作成したプロキシ証明書はdefaultで'rw——(600)'となっていて、他のユーザーから中身が読み取れないようになっていますが、証明書の権限は常に注意するようにしてください。

ファイルの読み込み

ROOTの `TFile::Open` を通してファイルにアクセスすることができます。

- WebDAVでのアクセス

```
TFile::Open("https://lcg-  
se01.icepp.jp:18443/dpm/icepp.jp/home/atlas/atlaslocalgroupdisk/rucio/
```

- xrootdでのアクセス

```
TFile::Open("root://lcg-  
se01.icepp.jp:1094//dpm/icepp.jp/home/atlas/atlaslocalgroupdisk/rucio/
```

注: 使用できるプロトコル(WebDAV, xrootd等)はサイトによって異なります。

`rucio list-file-replicas` で表示されたプロトコルは少なくとも使えるはずで
す。

バッチジョブで実行する場合は、環境変数(`X509_USER_PROXY`、 `X509_CERT_DIR`)が
引き継がれていることを確認するか、もしくは手動で明示的に設定(`export ~~~`)す
るようにしてください。

HTCondorでの例

以下のようなROOTマクロでグリッド上のファイルの読み込みをしてみます。

read.cc

```
void read(){  
    auto f = TFile::Open("https://lcg-  
se01.icepp.jp:18443/dpm/icepp.jp/home/atlas/atlaslocalgroupdisk/rucio/us  
  
    f->ls();  
  
    auto t = (TTree*)f->Get("CollectionTree");  
    std::cout << "# of entries in CollectionTree = " << t->  
>GetEntries() << std::endl;  
  
    return;  
}
```

submitファイルでは、必要な環境変数が定義されるように注意が必要です。以下で
は、 `getenv` を使う例、 `environment` を使う例を記載しています。

getenv を使う例

`getenv = True` を指定することで、ログインノードでの環境変数
(`X509_USER_PROXY`、 `X509_CERT_DIR` を含む)が引き継がれます。

execute.sdf

```
executable = /usr/bin/root
getenv      = True
arguments   = "read.cc"
output      = out.out
error       = out.err
log         = log.log
job_queue   = q4m
queue
```

environment を使う例

environment で必要な環境変数を個別に指定することも可能です。また、HTCondor はx509ユーザー証明書用のオプション(x509userproxy, use_x509userproxy)もあり、こちらを使用することで、ログインノード上の証明書がバッチノードにコピーされます。(つまり、証明書をバッチノードから見える場所に置く必要がなくなります。)

execute.sdf

```
executable = /usr/bin/root
arguments   = "read.cc"
use_x509userproxy = true
x509userproxy = $ENV(HOME)/.globus/gridproxy.cert
environment =
"X509_CERT_DIR=/cvmfs/atlas.cern.ch/repo/ATLASLocalRootBase/etc/grid-
security-emi/certificates"
output      = out.out
error       = out.err
log         = log.log
job_queue   = q4m
queue
```

ジョブ投入例

```
condor_submit execute.sdf
```

© 2024 ICEPP, the University of Tokyo.

Reproduction of the article, figures and tables on this page is allowed as specified in the CC-BY-4.0 license.