

ESGF Local Node Deployment Tutorial

This page shows how to deploy an ESGF Node that provides data, index and idp services and belongs to the esgf-test federation. The purpose of this node is to test the process of publication in the ESGF.

This page assumes that command are executed by the root user (not sudo).

Index

1. Prerequisites
2. Previous installation clean up
3. Installation from scratch
4. Configuration for publication
5. Publish the test dataset
6. Publish CORDEX datasets
7. Known issues
8. References

0. Prerequisites

1. You have to create a globus account - [?https://www.globusid.org/create](https://www.globusid.org/create)

1. Previous installation clean up

Execute `/usr/local/bin/esg-node stop` in order to stop the current ESGF services (in case they are running).

```
[root@spock ~]# /usr/local/bin/esg-node stop

EEEEEEEEEEEEEEEEEEEEEEEE SSSSSSSSSSSSSSS          GGGGGGGGGGGGGGFFFFFFFFFFFFFFFFFFFFFFFF
E::::::::::::::::::::E SS::::::::::::S          GGG::::::::::::GF::::::::::::F
E::::::::::::::::::::ES::::::::SSSSS::::S          GG::::::::::::GF::::::::::::F
EE::::::::EEEEEEEE::::ES::::S          SSSSSSS G::::GGGGGGG::::GFF:::::FFFFFFF::::F
 E::::E          EEEEEES::::S          G::::G          GGGGGG F::::F          FFFFFF
E::::E          S::::S          G::::G          F::::F
E::::EEEEEEEEEE S::::SSSS          G::::G          F:::::FFFFFFF
E::::::::::E          SS::::SSSS          G::::G          GGGGGGGGGG F:::::FFFFFF
E::::::::::E          SSS::::SS          G::::G          G::::G          F:::::FFFFFF
E::::EEEEEEEEEE          SSSSS::::S G::::G          GGGGG::::G F:::::FFFFFFF
E::::E          S::::SG::::G          G::::G          F::::F
E::::E          EEEEE          S::::S G::::G          G::::G          F::::F
EE::::EEEEEEEE::::ESSSSSS          S::::S G::::GGGGGGG::::GFF:::::FF
E::::::::::ES::::SSSSS::::S          GG::::::::::::GF:::::FF
E::::::::::ES:::::SS          GGG::::GGG::::GF:::::FF
EEEEEEEEEEEEEEEEEEEEEEEE SSSSSSSSSSSSS          GGGGGG GGGGFFFFFFFFFFFF.llnl.gov

Checking that you have root privs on spock.meteo.unican.es... [OK]
Checking requisites...

Using IP: 193.144.184.40
Stopping search services...
Using solr_workdir=/usr/local/src/esgf/workbench/esg/solr-5.5.3
Using solr_install_dir=/usr/local/solr-home/slave-8983
Using solr_data_dir=/esg/solr-index/slave-8983
Using solr_server_dir=/usr/local/solr
Using solr_logs_dir=/esg/solr-logs
Using esg_dist_url=http://esg-dn2.nsc.liu.se/esgf/dist
sudo: source: command not found
Sending stop command to Solr running on port 8983 ... waiting 5 seconds to allow Jetty process 16339 to stop gracefully.
```

```

Sending stop command to Solr running on port 8984 ... waiting 5 seconds to allow Jetty process 16554 to stop gracefully.
Stopping Globus Services for Data-Node... (GridFTP) stop_globus_services for datanode
globus-gridftp-server: unrecognized service
Stopping Globus Services for Index-Node... (MyProxy server) stop_globus_services for gateway
Stopping myproxy-server:                                [ OK ]
No MyProxy Process Currently Running...
Tomcat (jsvc) process is running...

stop tomcat: /usr/local/tomcat/bin/jsvc -pidfile /var/run/tomcat-jsvc.pid -stop org.apache.catalina.startup.Bootstrap
(please wait)
postmaster (pid 16024) is running...
Stopping postgresql service:                            [ OK ]
Stopping httpd:                                         [ OK ]
Running shutdown hooks...

-----
Running Node Services...
node type: [ data index idp compute ] (60)
-----
-----

```

Execute source /usr/local/bin/esg-purge.sh && esg-purge all

2. Installation from scratch

Change directory to /usr/local/bin/

```
[root@spock ~]# cd /usr/local/bin/
```

```

[root@spock bin]# wget -O esg-bootstrap http://distrib-coffee.ipsl.jussieu.fr/pub/esgf/dist/devel/esgf-installer/2.4/esg-b
[root@spock bin]# chmod 555 ./esg-bootstrap
[root@spock bin]# ./esg-bootstrap

```

Your directory should look like this:

```
[root@spock bin]# ls
esg-bootstrap  esg-functions  esg-init  esg-node  esg-purge.sh  jar_security_scan  setup-autoinstall
```

Check your node's version:

```

[root@spock bin]# ./esg-node --version
Version: v2.4.24-master-release
Release: Bifrost
Earth Systems Grid Federation (http://esgf.llnl.gov)
ESGF Node Installation Script

```

Set node's type:

```

[root@spock bin]# ./esg-node --set-type data idp index
node type set to: [ index data idp ] (28)

```

Install the node:

```
[root@spock bin]# ./esg-node --install
```

Please select the ESGF distribution mirror for this installation (fastest to slowest):

```

-----
[1] http://dist.ceda.ac.uk/esgf

```

```

[2] http://esg-dn2.nsc.liu.se/esgf
[3] http://aims1.llnl.gov/esgf
[4] http://distrib-coffee.ipsl.jussieu.fr/pub/esgf
-----
select [1] > 1

```

```

Are you ready to begin the installation? [Y/n]
Configured host IP address does not match available IPs...
Detected multiple IP addresses bound to this host...
Please select the IP address to use for this installation
-----
[0] : 193.xxx.xxx.xxx
[1] : 192.xxx.xxx.xxx
-----
select [] > (select the one that fits your case)

```

```

Welcome to the ESGF Node installation program! :-)

What is the fully qualified domain name of this node? [spock.meteo.unican.es]:
What is the admin password to use for this installation? (alpha-numeric only) []:
Please re-enter password:
What is the name of your organization? [unican]:
Please give this node a "short" name: []: unican
Please give this node a more descriptive "long" name []: unican
What is the namespace to use for this node? (set to your reverse fqdn - Ex: "gov.llnl") [es.unican.meteo]:
What peer group(s) will this node participate in? (esgf-test|esgf-prod) [esgf-test]:
What is the default peer to this node? [spock.meteo.unican.es]:
What is the hostname of the node do you plan to publish to? [spock.meteo.unican.es]:
What email address should notifications be sent as? []:
Is the database external to this node? [y/N]:
Please enter the database connection string...
(form: postgresql://[username]@[host]:[port]/esgct)
What is the database connection string? [postgresql://dbsuper@localhost:5432/esgct]: postgresql://
entered: postgresql://dbsuper@localhost:5432/esgct
What is the (low priv) db account for publisher? [esgct]:
What is the db password for publisher user (esgct)? []:

```

```

Enter password for postgres user dbsuper:
Re-enter password for postgres user dbsuper:
Please Enter PostgreSQL port number [5432]:>

```

```

Would you like a "system" or "user" publisher configuration:
-----
*[1] : System
[2] : User
-----
[C] : (Custom)
-----
select [1] >

You have selected: 1
Publisher configuration file -> [/esg/config/esgct/esg.ini]

Is this correct? [Y/n]
Your publisher configuration file will be: /esg/config/esgct/esg.ini
What is your organization's id? [unican]:

```

```

Would you like to use the DN: (OU=ESGF.ORG, O=ESGF) ? [Y/n]:
...

```

Please enter the password for this keystore :

Enter a single ip address which would be cleared to access admin restricted pages.
You will be prompted if you want to enter more ip-addresses

Do you wish to allow further ips? y/n
n

Create user credentials
Please enter username for tomcat [dnode_user]:
Please enter password for user, "dnode_user" [*****]:
Would you like to add another user? [y/N]:

Please Enter the public (i.e. routable) IP address of this host [193.xxx.xxx.xxx]:>
Do you wish to use an external IDP peer?(N/y):

Do you want to continue with the Globus installation and setup? [Y/n] :
Do you want to register the MyProxy server with Globus? [Y/n]:
Please provide a Globus username []: YOUR-GLOBUS-USER
Globus password []:

When finished, you should see something like this:

```
Running Node Services...
node type: [ data index idp ] (29)
-----
myproxy-s 23071    root    5u  IPv4 1526752      0t0  TCP *:7512 (LISTEN)
java      26088    solr   28u  IPv6 1591850      0t0  TCP 127.0.0.1:7983 (LISTEN)
java      26088    solr   92u  IPv6 1591986      0t0  TCP *:8983 (LISTEN)
java      26257    solr   28u  IPv6 1592730      0t0  TCP 127.0.0.1:7984 (LISTEN)
java      26257    solr   92u  IPv6 1593098      0t0  TCP *:8984 (LISTEN)
postmaste 29509    postgres 3u  IPv6 1449862      0t0  TCP [::1]:5432 (LISTEN)
postmaste 29509    postgres 4u  IPv4 1449863      0t0  TCP 127.0.0.1:5432 (LISTEN)
httpd     12706    root    4u  IPv6 1512235      0t0  TCP *:80 (LISTEN)
-----

Finished!...
In order to see if this node has been installed properly you may direct your browser to:
http://spock.meteo.unican.es/thredds
http://spock.meteo.unican.es/esg-orp
http://spock.meteo.unican.es/

Your peer group membership -- : [esgf-test]
Your specified "default" peer : [spock.meteo.unican.es]
Your specified "index" peer - : [spock.meteo.unican.es] (url = http://spock.meteo.unican.es/)
Your specified "idp" peer --- : [spock.meteo.unican.es] (name = SPOCK.METEO.UNICAN.ES)
Your temporary certificates have been placed in /etc/tempcerts
You can install them by executing this : esg-node --install-keypair /etc/tempcerts/hostcert.pem /etc/tempcerts/hostkey.pem
When prompted for the chainfile, specify: /etc/tempcerts/cacert.pem

[Note: Use UNIX group permissions on /esg/content/thredds/esgcert to enable users to be able to publish thredds catalogs fr
%> chgrp -R <appropriate unix group for publishing users> /esg/content/thredds

-----
Administrators of this node should subscribe to the
esgf-node-admins@lists.llnl.gov by sending email to: majordomo@lists.llnl.gov
with the body: subscribe esgf-node-admins
```

```
-----
v2.4.24-master-release
```

```
Writing additional settings to db. If these settings already exist, psql will report an error, but ok to disregard.
ERROR: insert or update on table "permission" violates foreign key constraint "permission_user_id_fkey"
DETAIL: Key (user_id)=(1) is not present in table "user".
Node installation is complete.
```

Execute the following:

```
[root@spock bin]# ./esg-node --install-keypair /etc/tempcerts/hostcert.pem /etc/tempcerts/hostkey.pem
...
Please set the password for this keystore :
Please re-enter the password for this keystore:
...
certfile> /etc/tempcerts/cacert.pem
certfile>
...
Is the above information correct? [Y/n]
Is the above information correct? [Y/n]
```

Restart the node:

```
[root@spock bin]# ./esg-node restart
```

Check that everything works ([?https://github.com/ESGF/esgf-installer/wiki/ESGF-Post-Installation-Tests](https://github.com/ESGF/esgf-installer/wiki/ESGF-Post-Installation-Tests)).

If the CoG portal does not work follow the instructions on [?https://www.earthsystemcog.org/projects/cog/install_or_upgrade](https://www.earthsystemcog.org/projects/cog/install_or_upgrade).

Now you should be able to log in the CoG portal using the openid "[?https://spock.meteo.unican.es/esgf-idp/openid/rootAdmin](https://spock.meteo.unican.es/esgf-idp/openid/rootAdmin)" and the password that you chose in the installation process.

Configuration for publishing

The installation process should have created a user in the postgres database, named rootAdmin. You can check it by running `psql -U dbsuper -d esgset` (to access the postgres cli) and visualizing the table `esgf_security.user`.

```
esgset=# select * from esgf_security.user;
id | firstname | middlename | lastname | email | username | password | dn |
ion_token | notification_code
-----+-----+-----+-----+-----+-----+-----+-----+-----
1 | Admin | | User | emailOfTheAdmin | rootAdmin | hashOfThePassword | | https://domain/esgf-idp-
-b50e-d43692adc5e5 |
```

In order to test the publication, create a new user using the CoG web interface ([?https://\[index_node_fgdn\]](https://[index_node_fgdn])). You should click on 'Create Account' and fill the form. Once the user is created using the CoG interface, it should be visible in the `esgf_security.user` table of the postgres database.

```
esgset=# select * from esgf_security.user;
id | firstname | middlename | lastname | email | username | password | dn |
ion_token | notification_code
-----+-----+-----+-----+-----+-----+-----+-----+-----
1 | Admin | | User | emailOfTheAdmin | rootAdmin | hashOfThePassword | |
-b50e-d43692adc5e5 | 0
2 | zequi | | cimadevilla | emailOfZequi | zequi | hashOfThePassword | |
```

Once the user is created, create permissions and roles as follows:

(reference documentation - [?https://acme-climate.atlassian.net/wiki/display/ESGF/Guide+to+ESGF+Publishing+and+Best+Practices](https://acme-climate.atlassian.net/wiki/display/ESGF/Guide+to+ESGF+Publishing+and+Best+Practices))

```

esgctl=# select * from esgf_security.role;
id | name | description
-----+-----+-----
1 | super | Super User
2 | none | None
3 | default | Standard
4 | publisher | Data Publisher
5 | admin | Group Administrator
6 | user | user role
(6 rows)

esgctl=# select * from esgf_security.group;
id | name | description | visible | automatic_approval
-----+-----+-----+-----+-----
1 | wheel | Administrator Group | t | t
2 | test_group | test group | t | t
3 | cordex_group | cordex group | t | t
(3 rows)

esgctl=# select * from esgf_security.permission;
user_id | group_id | role_id | approved
-----+-----+-----+-----
2 | 2 | 4 | t
2 | 2 | 6 | t
2 | 3 | 6 | t
2 | 3 | 4 | t
(4 rows)

```

Add the following elements to /esg/config/esgf_policies_local.xml

```

<policy resource=".*test.*" attribute_type="test_group" attribute_value="user" action="Read"/>
<policy resource=".*test.*" attribute_type="test_group" attribute_value="publisher" action="Write"/>
<policy resource=".*cordex.*" attribute_type="cordex_group" attribute_value="user" action="Read"/>
<policy resource=".*cordex.*" attribute_type="cordex_group" attribute_value="publisher" action="Write"/>

```

Add the following elements to /esg/config/esgf_ats_static.xml

```

<attribute
  type="test_group"
  attributeService="https://spock.meteo.unican.es/esgf-idp/saml/soap/secure/attributeService.htm"
  description="Test group for test data"
  registrationService="https://spock.meteo.unican.es/esgf-idp/secure/registrationService.htm"/>

<attribute
  type="cordex_group"
  attributeService="https://spock.meteo.unican.es/esgf-idp/saml/soap/secure/attributeService.htm"
  description="Test group for cordex data"
  registrationService="https://spock.meteo.unican.es/esgf-idp/secure/registrationService.htm"/>

```

Generate your credentials for publication - globus certificate

```
myproxy-logon [ -b ] -s <openid_server> -l <your_esgf_username> -p 7512 -t 72 -o $HOME/.globus/certificate-file
```

The certificate is valid for 72 hours when specified by -t. If you are publishing for the first time, you will need to mkdir \$HOME/.globus and use -b to bootstrap its trustroots with the server. The esgf_username is the simply the username portion of your openid rather than the entire openid string, e.g. sashakames, not [?https://pcmdi.llnl.gov/esgf-idp/openid/sashakames](https://pcmdi.llnl.gov/esgf-idp/openid/sashakames)

Publish the test dataset

For esgprep and esgpublish to be available, execute `source /etc/esg.env`.

```
[root@spock ~]# esgprep mapfile --project test /esg/data/test/
Collecting files      : 1 files
Mapfile(s) generation: 100% |????????????????????????????????????????????????????????????????????| 1/1 files
Mapfile(s) generated : 1 (see /root/mapfiles)
```

```
[root@spock ~]# esgpublish --service fileservice --map mapfiles/test.test.map --project test --thredds --publish --offline
INFO      2017-06-02 14:59:48,405 Replacing files in dataset: test.test, version 1
INFO      2017-06-02 14:59:48,413 File /esg/data/test/sft1f.nc exists, skipping
INFO      2017-06-02 14:59:48,416 New dataset version = 2
INFO      2017-06-02 14:59:48,430 Adding file info to database
INFO      2017-06-02 14:59:48,469 Writing THREDDS catalog /esg/content/thredds/esgcat/1/test.test.v2.xml
INFO      2017-06-02 14:59:48,522 Writing THREDDS ESG master catalog /esg/content/thredds/esgcat/catalog.xml
INFO      2017-06-02 14:59:48,533 Reinitializing THREDDS server
INFO      2017-06-02 14:59:48,830 Publishing: test.test
INFO      2017-06-02 14:59:49,871   Result: SUCCESSFUL
```

Notes:

1. --map must point to the file generated by esgprep mapfile
2. --thredds publish data to the data node
3. --publish publish data to the index node
4. --offline is required for publish the test dataset (Why?)
5. This publication works out of the box because esgf installs by default the required /esg/config/esgcat/esg.test.ini file.

Publish CORDEX datasets

See [CORDEXPublication](#)

Known issues during installation

#error "Pycogp requires PostgreSQL client library (libpq) >= 9.1

This error occurs sometimes during installation but removing the node and installing it from scratch seems to solve it...

```
Traceback (most recent call last):
  File "setup.py", line 110, in <module>
    """
  File "/usr/local/uvcdat/2.2.0/lib/python2.7/distutils/core.py", line 111, in setup
    _setup_distribution = dist = klass(attrs)
  File "/usr/local/uvcdat/2.2.0/lib/python2.7/site-packages/setuptools-1.4-py2.7.egg/setuptools/dist.py", line 239, in __init__
  File "/usr/local/uvcdat/2.2.0/lib/python2.7/site-packages/setuptools-1.4-py2.7.egg/setuptools/dist.py", line 268, in fetch_egg
  File "/usr/local/uvcdat/2.2.0/lib/python2.7/site-packages/setuptools-1.4-py2.7.egg/pkg_resources.py", line 568, in resolve
  File "/usr/local/uvcdat/2.2.0/lib/python2.7/site-packages/setuptools-1.4-py2.7.egg/pkg_resources.py", line 806, in best_match
  File "/usr/local/uvcdat/2.2.0/lib/python2.7/site-packages/setuptools-1.4-py2.7.egg/pkg_resources.py", line 818, in obtain
  File "/usr/local/uvcdat/2.2.0/lib/python2.7/site-packages/setuptools-1.4-py2.7.egg/setuptools/dist.py", line 318, in fetch_egg
  File "/usr/local/uvcdat/2.2.0/lib/python2.7/site-packages/setuptools-1.4-py2.7.egg/setuptools/command/easy_install.py", line 1
  File "/usr/local/uvcdat/2.2.0/lib/python2.7/site-packages/setuptools-1.4-py2.7.egg/setuptools/command/easy_install.py", line 1
  File "/usr/local/uvcdat/2.2.0/lib/python2.7/site-packages/setuptools-1.4-py2.7.egg/setuptools/command/easy_install.py", line 1
  File "/usr/local/uvcdat/2.2.0/lib/python2.7/site-packages/setuptools-1.4-py2.7.egg/setuptools/command/easy_install.py", line 1
  File "/usr/local/uvcdat/2.2.0/lib/python2.7/site-packages/setuptools-1.4-py2.7.egg/setuptools/command/easy_install.py", line 1
distutils.errors.DistutilsError: Setup script exited with error: command 'gcc' failed with exit status 1

Sorry...
This action did not complete successfully
Please re-run this task until successful before continuing further

Also please review the installation FAQ it may assist you
```

```
https://github.com/ESGF/esgf.github.io/wiki/ESGFNode%7CFAQ
```

Failed building wheel for Pillow

This error seems unavoidable but it also seems that it doesn't affect the esgf functionality.

Installing a custom certificate in the ESGF Node

You should own your certificate file (hostcert.crt) and your private key (hostkey.key). Your /etc/httpd/conf/esgf-httpd.conf must reference your certificate and key:

```
228     SSLVerifyClient optional
229     SSLVerifyDepth 10
230     SSLCertificateFile /etc/certs/hostcert.crt
231     #SSLCACertificateFile /etc/certs/esgf-ca-bundle.crt
232     SSLCertificateKeyFile /etc/certs/hostkey.key
233     #SSLCertificateChainFile /etc/certs/cachain.pem
234     SSLOptions +StdEnvVars +ExportCertData
```

Then you have to import your certificate and your key into your tomcat keystore (located in /esg/config/tomcat/ and named esg-truststore.ts and keystore-tomcat). They are configurated in /usr/local/tomcat/conf/server.xml.

1. If the self-signed certificate is installed in keystore-tomcat, remove it with `keytool -delete -alias ALIAS -keystore keystore-tomcat`, where alias can be obtained with `keytool -v -list -keystore keystore-tomcat`.
1. Execute `# openssl pkcs12 -export -in /etc/certs/hostcert.crt -inkey /etc/certs/hostkey.key -out server.p12 -name my-esgf-node -CAfile /etc/certs/hostcert.crt -caname root and keytool -importkeystore -deststorepass PASSWORD -destkeypass PASSWORD -destkeystore keystore-tomcat -srckeystore server.p12 -srcstoretype PKCS12 -srcstorepass PASSWORD -alias my-esgf-node`
1. Ensure it has been correctly installed with `keytool -v -list -keystore keystore-tomcat`.
1. Restart the node: `esg-node restart`
1. More info in [?Stackoverflow](#)

References

- [?ESGF Installation From Scratch](#)