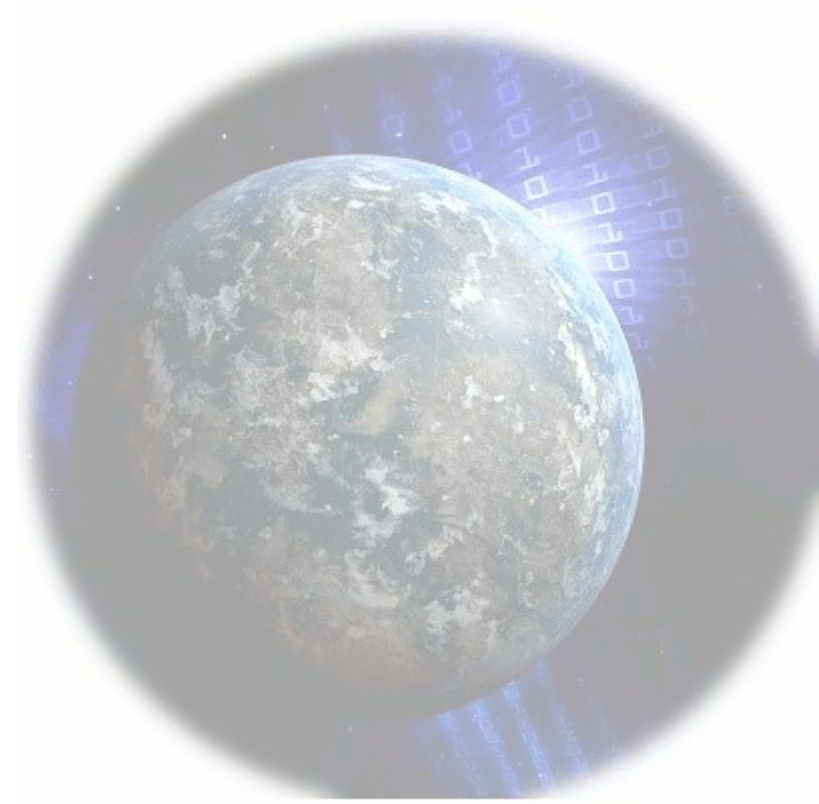# PhEDEx Installation Tutorial

T-3 Workshop August 2010
Carl Lundstedt

Goals

- Understand what we're trying to do

- Installing the client software

- Phedex Configuration

- Starting/Stopping Agents

What does it mean to "Install Phedex"

• PhEDEx at your site is a set of agents that query a central database for tasks.

• These agents need to 'understand' your site's layout and be able to present their tasks as an authorized user (you).

• This understanding is done via configuration files. The construction of these files is the hard part

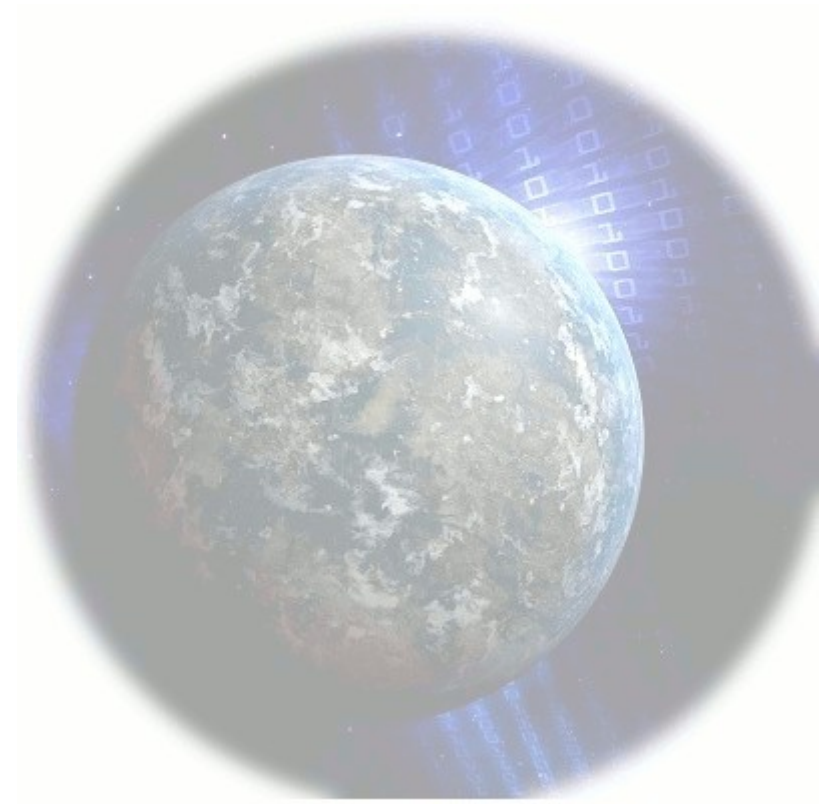# What Phedex needs to be able to do @ a Tier3:

- Get files, provide files, remove files, verify downloads.

- It accomplishes these task with agents. You need, as a minimum 4 agents configured and running:

Filedownload
FileExport
FileRemove
BlockDownloadVerify

Installing the client software

- Create a fresh cmsphedex account on the machine (or virtual machine)

- AS ROOT: Install Nebraska's RPM repo to get the srm wrapper script
rpm -i http://t2.unl.edu/store/rpms/SL4/x86_64/Nebraska-repo-0.1-1.noarch.rpm

- Install the OSG client package to help with Certificate management
Follow:
https://twiki.grid.iu.edu/bin/view/ReleaseDocumentation/ClientInstallationGuide
(watch out for selinux, I installed ours to /opt/osg/osg-1.2)

- It's helpful to define a large amount of stuff in your user's .bashrc
(More on that later)

- Log into the server that will serve as your Phedex machine as the user you wish to manage the transfers e.g. cmsphedex

- Change directory to cmsphedex's home.
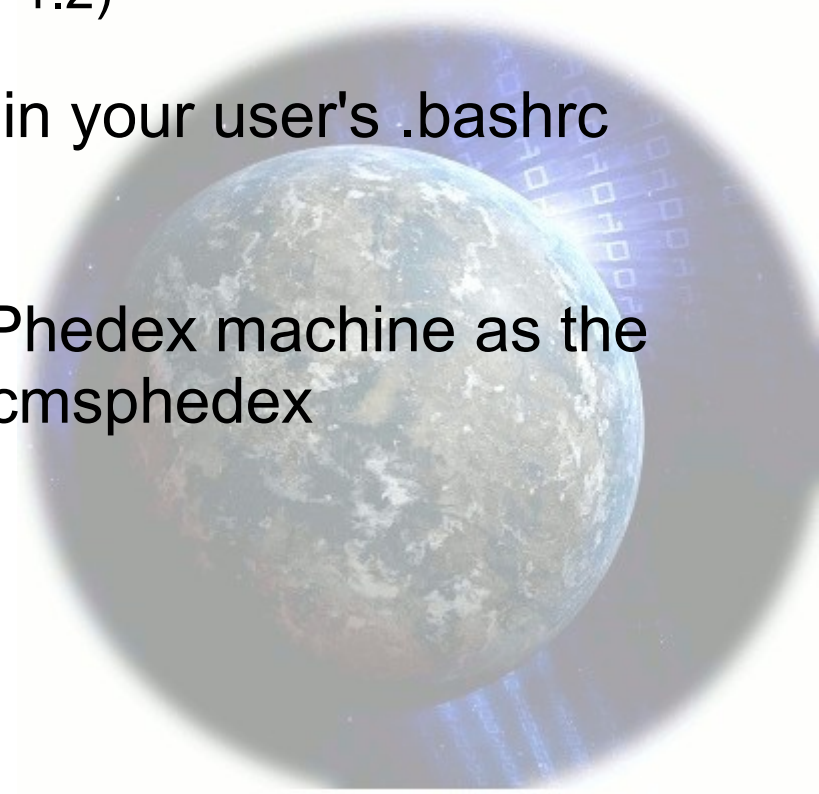
THE MEAT of Installation

Installation of the software has been 'scripted' by Doug Johnson from CU. Get the script from his server:

wget   http://www-hep.colorado.edu/computing/PhEDEx-lcg-cp/InstallPhEDEx
chmod +x InstallPhEDEx

Download java binary install kit:

wget http://unlhepcl.unl.edu/cl/jre-6u21-linux-x64.bin

Edit InstallPhEDEx to customize for your site:

phedexHome="/home/cmsphedex"   (?)
javaFile="$phedexHome/jre-6u21-linux-x64.bin"

Execute the script. Wait.

./InstallPhEDEx

Install the T3 srmcp utilities from Nebraska (remember setting up the yum repo? You'll have to do this as root.)

yum install srmcpT3

THE MEAT of Configuraton

mkdir -p ~/SITECONF/T3_US_mysite/PhEDEx

cp ~/PHEDEX/Custom/Template/* ~/SITECONF/T3_US_mysite/PhEDEx

NOW THE TRULY MYSTERIOUS PART!!

Edit the Config files to match your site.
cd ~/SITECONF/T3_US_mysite/PhEDEx/
cp Config Config.debug
Edit Config.debug, below is Nebraska's common header.

```
PHEDEX_BASE=/grid_home/cmsphedex;
PHEDEX_INSTANCE=Debug;
PHEDEX_LABEL=NEBRASKA;
PHEDEX_SITE=T2_US_Nebraska;
PHEDEX_NODE=${PHEDEX_SITE};
PHEDEX_VERSION=3_3_1;
PHEDEX_OS_VERSION=slc5_amd64_gcc434;
PHEDEX_X509_USER_PROXY=$PHEDEX_BASE/gridcert/proxy.cert;
```

THE MEAT of Configuraton

Configure the agents for the Debug instance.

ConfigPart.SRMDownload

```
### AGENT LABEL=download-srm PROGRAM=Toolkit/Transfer/FileDownload
 -db            ${PHEDEX_DBPARAM}
 -nodes          ${PHEDEX_NODE}
 -ignore         '%CERN_MSS%'
 -accept         '%'
 -delete         ${PHEDEX_CONFIG}/FileDownloadDelete
 -validate       ${PHEDEX_CONFIG}/FileDownloadVerify,-d
 -backend        SRM
 -protocols      'srmv2'
 -batch-files    1
 -command /usr/bin/srmcp_wrapper,-2,-x509_user_proxy=${X509_USER_PROXY},
          -debug=true,-retry_num=2,-x509_user_trusted_certificates=
          /opt/osg/osg-1.2/globus/TRUSTED_CA
```

May not want all!

-d = delete on arrival
Don't use in PROD instance!

Look at other site's configs to configure your other agents.

http://cmssw.cvs.cern.ch/cgi-bin/cmssw.cgi/COMP/SITECONF/

A quick note on **clients**. For this tutorial I used Nebraska's T3 srmcp wrapper. This will probably 'work' for US transfers, but you may wish to change your client
lcg-cp or srm-cp may be more appropriate. We can help

lcg-cp may be needed to contact CERN, you can setup a CERN only agent if needed.

srm-cp has trouble with Bestman

srm-copy may be what you're after in the end.

THE MEAT of Configuraton

Install your database credentials. There are instructions included in the email you'll get from CERN. In the end you should have a DBParam file in ~/gridcert that will have Sections for each instance. e.g.

```
Section            Debug/YOURAWSOMESITENAME
Interface          Oracle
Database           cms_transfermgmt_sc
AuthDBUsername     cms_transfermgmt_sc_writer
AuthDBPassword     <KEEP THIS SECRET!!!>
AuthRole           phedex_nebraska_debug
AuthRolePassword   <KEEP THIS SECRET TOO!!!!>
ConnectionLife     86400
LogConnection      on
LogSQL             off
```

The TFC and You

• Phedex and CMSSW need to be able to translate between "LFNs" and "PFNs" (CMSSW speaks LFN, your storage site speaks PFN)

• The ruleset is outlined in your Trivial File Catalog (TFC) this MUST be checked into CVS The TFC is a flat text file named storage.xml located in SITECONF/<SITENAME>/PhEDEx

Basic Format:

```
<storage-mapping>
  <lfn-to-pfn protocol="file"
      path-match="/+store/(.*)" result="file:/some/disk/path/store/$1"/>
  <pfn-to-lfn protocol="file"
      path-match="file:/+some/disk/path/store/(.*)" result="/store/$1"/>
</storage-mapping>
```

• The best way to build your TFC is to 'cheat' and look in CVS for comparable sites and build your TFC by example. The TFC is heavily site specific, you will need to build these.

Now for your first datasets

• Subscribe to a loadtest in the debug instance. We'll have to help you here. There's an analog hole here. You have to have sites create an injection for you.

• In order to do the CRAB tutorial from this workshop you'll need:
 /JetMETTau/Run2010A-Jun14thReReco_v2/RECO (750GB)

• Goto the Phedex pages and create a subscription request
https://cmsweb.cern.ch/phedex/prod/Info::Main

You need your cert loaded into your browser.

For Prod and Debug Create a 'Transfer Request'

Starting your agents

- Generate a grid-proxy
  grid-*proxy-init -valid 190:00 -voms cms*

- Copy your new proxy to the spot you indicated in your agent config, I put Nebraska's at ~/gridcert/proxy.cert

      cp /tmp/x509up_u<user id> ~/gridcert/proxy.cert

- Setup the environment (if you haven't yet, e.g. in your log in)

source $sw/$myarch/cms/PHEDEX/PHEDEX_$version/etc/profile.d/env.sh

- Start the agent

Master -config ~/SITECONF/T3_US_<SiteName>/PhEDEx/Config.Debug start

Check logs to see if your agents have started.
DEBUGGING STARTS. COME BACK TOMORROW.

Now for your first Prod dataset

Making Your Environment Happy

• Setup your login shell to automatically source the appropriate setup scripts.

• Here's a snippet from my login shell for Nebraska's cmsphedex account

```
export CVSROOT=:pserver:anonymous@cmscvs.cern.ch:/cvs_server/repositories/CMSSW

export sw=$HOME/sw
export version=3_3_1
export myarch=slc5_amd64_gcc434
source $sw/$myarch/cms/PHEDEX/PHEDEX_$version/etc/profile.d/env.sh
PATH="/grid_home/cmsphedex/sw/slc5_amd64_gcc434/cms/PHEDEX/PHEDEX_3_3_1/
    Utilities:$PATH"
export PATH
source /opt/osg/osg-1.2/setup.sh
```

Get these things done prior to the workshop. Some things need a little lead time. (Credit Paul Rossman)

1. Become CERN user and CMS member

- http://ph-dep-usersoffice.web.cern.ch/ph-dep-UsersOffice/

2. Acquire a CERN computer account

- http://it-div.web.cern.ch/it-div/documents/ComputerUsage/CompAccountRegistrationForm-English.pdf

3. Gain access rights in CMS CVS

- https://savannah.cern.ch/support/?func=additem&group=cmscompinfrasup&assigned_to=3534&category_id=107

4. Acquire a certificate from DOEGrids cert

- http://www.doegrids.org/

5. Register with LCG

- http://lcg.web.cern.ch/LCG/registration.htm

6. Register your DOEGrid cert with LCG User Registration

- https://ca.cern.ch/ca/

7. Get a HyperNews account

- https://twiki.cern.ch/twiki/bin/view/CMS/PhedexAdminDocHyperNewsForums