

# Table of Contents

<b>Central Harvester Instances.....</b>	<b>1</b>
Technical documentation.....	1
Monitoring.....	1
Harvester machines.....	1
Important paths and files.....	1
.Harvester.Service.....	2
Condor Schedd machines.....	3
Restarting condor.....	3
Important paths and files.....	3
Troubleshooting.....	3

# Central Harvester Instances

## Technical documentation

For a technical description of Harvester components please visit the Harvester github wiki [wiki](#)

## Monitoring

- Harvester dashboard in BigPanDA [wiki](#)
- Harvester node monitoring [wiki](#) to see CPU, memory and disk usage
- Harvester service monitoring [wiki](#) is integrated with the other ADC components. Currently the harvester and schedd monitoring checks the processes are running and diskspace is not reaching a critical usage (80% and 90% thresholds).

## Harvester machines

You can connect with the usual atlpan user

Name	HarvesterID	Description
aipanda170	cern_cloud	Pre-production node.
aipanda171, aipanda172	CERN_central_A	Grid PQs: CERN T0, CA, FR, NL, RU, US. Submits via remote schedds
aipanda173, aipanda174	CERN_central_B	Grid PQs: CERN, DE, ES, IT, TW, UK. Submits via remote schedds
aipanda175	CERN_central_0	Submits to P1. It contains a local MySQL database and local schedd.
aipanda177, aipanda178	CERN_central_1	Submits to special resources like CloudSchedulers. Submit to local schedd.

## Important paths and files

Now Python 3 becomes standard for central Harvester.

Some Harvester servers are migrated to run with Python 3 in virtual environment (venv), while some nodes are still running with Python 2, installed in system python (no venv), which will be migrated.

Please note the path difference of files below between different python version.

Files under /cephfs are shared across harvester nodes and schedd nodes, which are used when harvester submits via remote schedds.

Filename	Description
/var/log/harvester	All the harvester logs of the various agents
/opt/harvester	Main directory of Harvester python virtual environment (python 3 only)
/opt/harvester/etc/panda/panda_harvester.cfg (python 2: /usr/etc/panda/panda_harvester.cfg)	General configuration of subcomponents, DB connection, etc.

/opt/harvester/etc/panda/panda_queueconfig.json (python 2: /usr/etc/panda/panda_queueconfig.json)	Queue configuration
/data/atlpan/harvester_common/ , /cephfs/atlpan/harvester/harvester_common/	Condor sdf templates and other files needed by Harvester
/data/atlpan/harvester_wdirs/\${harvesterID}/XX/YY/\${workerID} , /cephfs/atlpan/harvester/harvester_wdirs/\${harvesterID}/XX/YY/\${workerID}/	Worker directories: sdf file submitted to Condor for each job and other files. Where XXYY are the last 4 digits of workerID.
/data/atlpan/harvester_worker_dir/ , /cephfs/atlpan/harvester/harvester_worker_dir/	Deprecated and removed. (worker directory)
/data/atlpan/condor_logs/	Local condor and pilot logs for each job
/data1/atlpan/condor_logs/	On schedd nodes: condor and pilot logs for each job; on harvester nodes: dummy folders/files but must exist

## Harvester Service

To start, stop, or reload service

Python 3 (with venv):

```
[root@<machine>]# /opt/harvester/etc/rc.d/init.d/panda_harvester-uwsgi start
[root@<machine>]# /opt/harvester/etc/rc.d/init.d/panda_harvester-uwsgi stop
[root@<machine>]# /opt/harvester/etc/rc.d/init.d/panda_harvester-uwsgi reload
```

Python 2 (without venv):

```
[root@<machine>]# /usr/etc/rc.d/init.d/panda_harvester-uwsgi start
[root@<machine>]# /usr/etc/rc.d/init.d/panda_harvester-uwsgi stop
[root@<machine>]# /usr/etc/rc.d/init.d/panda_harvester-uwsgi reload
```

The most important tables in the DB structure can be found here [↗](#), although it's under constant evolution. The DB configuration can always be found in the panda-harvester.cfg file, but currently you can connect to them like this:

Read-only account atlas-ro for debugging:

- On aipanda171,172: # mysql -h dbod-harv.cern.ch -P 5501 -u atlas-ro -p HARVESTER
- On aipanda173,174: # mysql -h dbod-harv2.cern.ch -P 5500 -u atlas-ro -p HARVESTER
- On aipanda175: # mysql -u atlas-ro -p harvester
- On aipanda177,178: # mysql -h dbod-harv-cl.cern.ch -P 5500 -u atlas-ro -p harvester

Password is in /cephfs/atlpan/harvester/mysql-passwd

Important paths and files

Note the database names, DB hostname, and ports are different on nodes. Check configuration in [db] schema in harvester.cfg

You will need to know the password to connect. Run only queries you understand and where you know what you are doing.

If write permission really needed, use the account of harvester service, Check harvester.cfg for account user and password.

## Condor Schedd machines

These are external condor schedd nodes harvester nodes submit through. You can connect with the usual atlpn user

Name
aipanda023
aipanda024
aipanda156
aipanda183
aipanda184

They are interchangeable.

## Restarting condor

```
[root@]# systemctl restart condor
```

## Important paths and files

Filename	Description
/etc/condor/	Condor configuration, our configuration usually goes under config.d
/var/log/condor/	Condor logs, one per each agent

## Troubleshooting

Condor schedd related [↗](#)

more...

FahuiLin - 2019-03-14

---

This topic: PanDA > CentralHarvesterInstances  
Topic revision: r13 - 2019-09-26 - IvanGlushkov



Copyright &© 2008-2020 by the contributing authors. All material on this collaboration platform is the property of the contributing authors.

Ideas, requests, problems regarding TWiki? Send feedback