OneDrive on HPC

IMPORTANT POINTS:

- Officially, JCU's OneDrive for Business platform hasn't been provisioned for housing research data, however JCU staff/students may choose to place their research data on OneDrive.
- OneDrive software installed on HPC is NOT provided or supported by Microsoft. Technology Solutions staff do not use routinely use OneDrive on HPC - two of us have successfully run the commands below once.
- HPC storage does NOT exist to serve as a sync client to your OneDrive files.   The OneDrive software on HPC should only be used to upload files on HPC to OneDrive - e.g., for data protection or archiving (long-term).
- A long-term storage option for JCU research data is being considered. AARNet CloudStor is the most likely platform at this time. HPC cluster storage may be "scratch space" only, possibly with routine deletion of "old" files configured.

Background:

Extensive documentation on the linux onedrive connector capabilities can be found at https://github.com/abraunegg/onedrive/blob/master/docs/USAGE.md - the information provided below is a very brief introduction for getting started. At the time of writing this document, HPC staff had only completed very limited testing of the Linux interface to OneDrive. The onedrive command doesn't have the capabilities provided by Windows/MacOS OneDrive clients, but it allows you to copy files housed on HPC to OneDrive.

Brief Documentation

A command line based interface to OneDrive has been installed on HPC cluster servers. There is some configuration required before you start using the command(s) provided.

1. Create a "onedrive" directory somewhere within your scratch space directory.

```
mkdir -p -m 700 /scratch/`whoami`/onedrive
```

2. Modify your environment by loading the onedrive module.

```
module load onedrive
```

3. Authenticate with OneDrive:

   a. Issue the command

   ```
   onedrive
   ```

   b. You should get a response something like:

   ```
   [jc012345@ln01 ~]$ onedrive

   Authorize this app visiting:

   https://.....

   Enter the response uri:
   ```

   c. Copy and paste the URL into a browser (on your desktop/laptop computer) and hit the Enter or Return key to load content. You will then need to authenticate to OneDrive - for most JCU staff/students SSO (single sign on) should take care of it for you. Once authentication has been successful, a blank page should be returned (may change in future). You need to copy the URI/URL that was returned (where you pasted the URL earlier) and paste the contents as a response to the `Enter the response uri:` prompt. You should receive a message that indicates successful authentication.
4. Now it's time to configure your use of `onedrive` on HPC. The following is very similar to a configuration that has been tested by HPC staff for use of OneDrive as an archive for files on HPC.

```
mkdir -p ~/.config/onedrive
cat <<EOF >~/.config/onedrive/config
sync_dir = "/scratch/`whoami`/onedrive"
skip_file = "~*|.~*|*.tmp"
skip_dotfiles = "true"
skip_symlinks = "true"
upload_only = "true"
no_remote_delete = "true"
remove_source_files = "false"
EOF
```

5. To confirm your configuration (may be larger than the contents of your configuration file), execute

```
onedrive --display-config
```

6. You are now ready to upload files to or download files from OneDrive, remembering that HPC storage is for research files only - all other files should be moved to another location (removed from HPC). For uploading files to OneDrive, copy/move files to the "`sync_dir`" you configured - e.g., to move all *.dat files in your current working directory, use `mv *.dat /scratch/`whoami`/onedrive`.
7. Two aliases have been created for upload/download convenience. To upload files from your "`sync_dir`", simply use the command:

```
onedrive-upload
```

8.   The progress of uploads will show up on your terminal. The command should tell you whether transfers were successful or unsuccessful. When finished, you may choose to verify success accessing your OneDrive storage from your desktop/laptop computer.

Downloading from OneDrive

At this stage, it looks like download capabilities are tied to sync'ing everything from OneDrive. Files you have uploaded to OneDrive from another personal /work computing device *may not be permitted* on HPC (research files only). Therefore, you should **not** be sync'ing (on HPC servers) data with OneDrive. Point in time downloads from OneDrive to HPC can be done, providing you delete all content that *isn't* related to your research.

Access to Teams/Sharepoint Libraries

The linux onedrive client installed on HPC systems is capable of working with Teams/Sharepoint libraries. However, due to JCU's O365 storage limit, files on HPC should **not** be sync'ed with Teams/Sharepoint document libraries. Such libraries can be synced to other computing devices (e.g., your desktop computer).

HPC filesystems

The future architecture of the JCU HPC cluster is under review - based on a government & university drive to utilize 'cloud' infrastructure where practical & sustainable. Technology solutions (including HPC) staff are investigation options for a future HPC cluster where >=95% of storage is likely to be scratch space, with routine deletion of old files. The

```
no_remote_delete = "true"
```

configuration entry (in `~/.config/onedrive/config`) will be essential if this future state becomes reality.