

Data Archives

Current Usage (as of 18 Mar 2021)

	Total	Used	Avail
Isilon	22T	19T	3.7T
Dropbox	3,675 TB	85.96 TB (by all UC staff)	3,589.04 TB

Isilon for data that needs more frequent access.

Cybershake

Cybershake 20+ Inputs Table

We will keep the latest CS data on Isilon

	What is this data?	Source	VM (.cfgs) / vm_params. yaml	BB	IM	Note
22p12		Dropbox	Same as left	Same as left	Same as left	All good
22p11	100m Wellington region	Dropbox	Same as left	Same as left	Same as left /isilon/CS_IMs/v22p11	Source median missing
22p6	400m CS rerun of 20p4 /5 (Abandoned)					
22p4	100m last batch South Island	Dropbox	Same as left	Same as left	Same as left /isilon/CS_IMs/v22p4 /NeSI (faults not covered in KISTI runs)	All good
22p2	100m South Island batch	Dropbox	Same as left	Same as left	Same as left /isilon/CS_IMs/v22p2	All good except TeAnau IM (incomplete)
21p6	100m South Island Continuation (contains a few different batches in the one version)	Dropbox	Same as left	Same as left	Same as left /isilon/CS_IMs/v21p6	Many missing items. See Stocktake.csv .
21p1	First batch of 100m South Island	Dropbox	Same as left	Same as left	Same as left /isilon/CS_IMs/v21p1_im	All good
20p9		Dropbox	Same as left	Same as left	Same as left /isilon/CS_ims/	Missing all vm_params.yaml Stocktake.csv stating RELs csv /info are missing, but they are good. (TODO)
20p6 (=20p4_200)		Same as 20p4 /isilon/cybershake/v20p4/Sources	Dropbox Isilon: /isilon/cybershake/v20p4/200m_vmc ▪ both are missing AhuririR's nzvm.cfg file	Dropbox (8.0Tb) /isilon/cybershake/v20p6/	IM_calc – Dropbox /isilon/CS_ims/	
20p5		Dropbox (.csv/.info) RELXX.csv/.info are missing.	Dropbox	Dropbox	/isilon/CS_ims/	

20p4		Dropbox (.csv/.info) Isilon: /isilon/cybershake/v20p4/Sources (full)	Dropbox /isilon/cybershake/v20p4	Dropbox (117b)	/isilon/CS_ims/	
19p5		Full Srf/Stoch v19p5_sources – Dropbox (766Gb)	Dropbox (vm_params)	Dropbox (117b)	/isilon/CS_ims/	
18p6		Full Srf/Stoch v18p6_sources – Dropbox (846Gb)	Dropbox	Nearline /nesi/nearline /nesi00213 /RunFolder /to_nearline_18p6/ v18p6_waveform_backup.tar <i>(Can't retrieve from it) (43.67Tb)</i>	Dropbox (IMDB.h5) /isilon/CS_ims/	
18p5				Isilon: /isilon/cybershake/v18p5		
18p4				Isilon: /isilon/cybershake/v18p4		
17p9				Isilon: /isilon/cybershake/v17p9		

Note:

For a Cybershake version (eg. v20p9), we keep 3 items

- Sources: .csv and .info per fault and all realisations. All files packed into one tar file. v20p9_Sources.tar . (pre-v20p4, we kept all Srf and Stoch files)
- VMs: Only nzvm.cfg per fault. All files packed into one tar file. v20p9_VM_cfgs.tar
- BB: BB.bin files for all realisations packed into a tar file for each fault. (eg. AlpineF2K.tar)

An automated archive [script](#) can package all CS files and generate a Dropbox upload command that can be executed from NeSI login node (in "screen" session)

Dropbox has a maximum file size of 350Gb. The auto script will split the TAR files into multiple parts of 340Gb size. To join them (before extracting from TAR),

```
cat ${TARNAME}.parta* > ${TARNAME}.joined
tar xvf ${TARNAME}.joined
```

Empirical DB

Mostly on Isilon. Upload to Dropbox in progress

	location	notes
v20p4	/isilon/seistech/empiricals/v20p4 (Dropbox) Nearline : /nesi/nearline/nesi00213/seistech/empiricals/v20p4 (to be deleted)	ds, flt in db format Files below are > 340Gb and split parts are uploaded to Dropbox Br_10_ACTIVE_SHALLOW_ds.db CY_14_ACTIVE_SHALLOW_ds.db CB_14_ACTIVE_SHALLOW_ds.db BSSA_14_ACTIVE_SHALLOW_ds.db ASK_14_ACTIVE_SHALLOW_ds.db

v19p11	/isilon/seistech/empiricals/19p11_13stations (Dropbox) /isilon/seistech/empiricals/19p11 (Dropbox)	
v19p9	/isilon/seistech/empiricals/19p9_complete (Dropbox) /isilon/seistech/empiricals/19p9_1655 (Dropbox) /isilon/seistech/empiricals/19p9 (Dropbox)	
v18p6	/isilon/seistech/empiricals/18p6_new (Dropbox) /isilon/seistech/empiricals/18p6 (Dropbox) <u>Nearline</u> /nesi/nearline/nesi00213/empiricals/v18p6	Nearline versions are TAR per each IM type (eg. SA_0p5.tar, PGA.tar)
v18p5	<u>Nearline</u> /nesi/nearline/nesi00213/empiricals/v18p5 /nesi/nearline/nesi00213/empiricals/v18p5_no_subduction	Nearline versions are TAR per each IM type (eg. SA_0p5.tar, PGA.tar)

Old Staff Data

Nearline : /nesi/nearline/nesi00213/RunFolder/to_nearline/RunFolder.tar. Hobby, Asahn