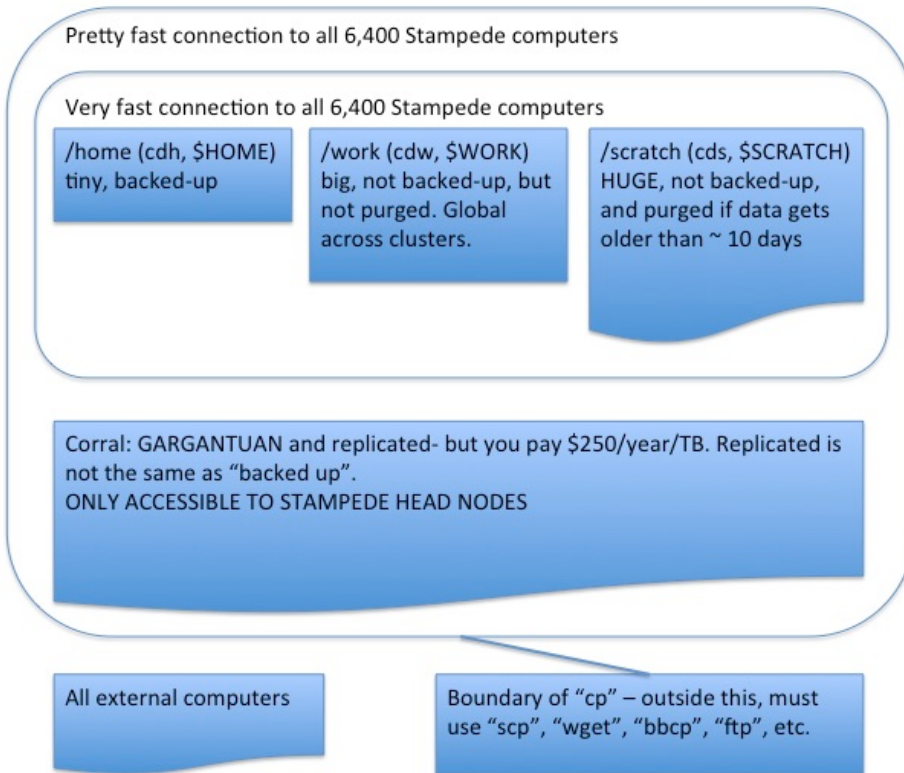# Stampede2 Directories

Diagram of stampede2 directories and what connects to what, and how fast

Stampede2 is a collection of 4,200 KNL nodes (computers), 1736 SKX nodes (computers) connected to three file servers, each with unique characteristics.

You need to understand the file servers to know how to use them.

Pretty fast connection to all 6,400 Stampede computers

Very fast connection to all 6,400 Stampede computers

/home (cdh, $HOME)
tiny, backed-up

/work (cdw, $WORK)
big, not backed-up, but
not purged. Global
across clusters.

/scratch (cds, $SCRATCH)
HUGE, not backed-up,
and purged if data gets
older than ~ 10 days

Corral: GARGANTUAN and replicated- but you pay $250/year/TB. Replicated is
not the same as "backed up".
ONLY ACCESSIBLE TO STAMPEDE HEAD NODES

All external computers

Boundary of "cp" – outside this, must
use "scp", "wget", "bbcp", "ftp", etc.

	$HOME	$WORK2	$SCRATCH
Purged?	No	No	Files can be purged if not accessed for 10 days.
Backed Up?	Yes	No	No
Capacity	10GB	1TB	Basically infinite. 8.5 PB
Command to Access	cdh	cdw2	cds
Purpose	Store Executables	Store Files	Run Jobs

Executables that aren't available on TACC through the "module" command should be stored in $HOME.

If you plan to be using a set of files frequently or would like to save the results of a job, they should be stored in $WORK.

If you're going to run a job, it's a good idea to keep your input files in a directory in $WORK2 and copy them to a directory in $SCRATCH where you plan to run your job.

This example command might help a bit:

```
cp $WORK2/my_fastq_data/*fastq $SCRATCH/my_project/
```

Stampede2's /home and /scratch file systems are mounted and visible only on Stampede2, but the work file system mounted on Stampede2 is part of the global file system hosted on Stockyard. This means /work and /work2 are visible on common among TACC clusters : lonestar5, stampede2, and frontera.

General Guidelines to reduce File I/O load on TACC:

TACC staff now recommends that you run your jobs out of the $SCRATCH **file system instead of the global** $WORK **file system.** Copy input files to $SCRATCH, run analyses, output should be written to $SCRATCH. Copy results to $WORK when done.

General Guidelines to when transferring files to/from TACC:

1. Don't do too many (>3) simultaneous file transfers.
2. If you need to transfer recursive files (directories within directories, lots of small files), create a TAR archive before transferring.

Now let's go on to look at how jobs are run on stampede2.