Submitting large number of Jobs on Xanadu in array format

Below is the example of passing multiple .fastq or .fq files through FASTQC by submitting array jobs. The requirement for the array job to work in the script below is that all the files are present in the same directory.

(STEP1): export the file names

```
export FILES=($(ls -1 *.fastq))  # Exporting list of all the
fastq files from the directory
```

(Step2): Compose the script and save as `fastqc.sh`

```
#!/bin/bash
#SBATCH --job-name=array
#SBATCH --output=array_%A_%a.out
#SBATCH --error=array_%A_%a.err
#SBATCH --partition=general
#SBATCH --qos=general
#SBATCH --mem=4G
#SBATCH -c 4

Module load fastqc

# grab out filename from the array exported from our 'parent'
shell in step 1
FILENAME=${FILES[$SLURM_ARRAY_TASK_ID]}
# make & move into new directory, and run!
# Option 1 : Each output file has a separate output directory
mkdir ${FILENAME}_out
fastqc -t 4 -o ./${FILENAME}_out path/to/$FILENAME

# Option 2 : Fastqc outputs of all the files saved in the same
directory
mkdir fastqc_out
fastqc -t 4 -o ./fastqc_out path/to/$FILENAME

Use one of the options
```

(Step3): Submitting Array job

```
# get size of array
NUMFASTQ=${#FILES[@]}
# now subtract 1 as we have to use zero-based indexing (first
cell is 0)
ZBNUMFASTQ=$(($NUMFASTQ - 1))
# now submit to SLURM
if [ $ZBNUMFASTQ -ge 0 ]; then
sbatch --array=0-$ZBNUMFASTQ%10 fastqc.sh
fi
```

%10 : This command will run your script NUMFASTQ times with at most 10 copies running at any one time. (The "%10" part is optional and the value has to be less than NUMFASTQ

Set this to **50** if you are planning to run **more than 50** jobs.

#NUMFASTQ : Represent total number of fastq files to be processed.