HPC Data Management: Or how (not) to handle your data in an HPC environment

Marty Kandes, Ph.D.

Computational & Data Science Research Specialist
High-Performance Computing User Services Group
San Diego Supercomputer Center
University of California, San Diego

CIML Summer Institute
Monday, June 27th, 2022
10:05 AM - 11:35 AM PT

tl;dr: Data has a lifecycle, data management is a lifestyle

Image Credit: Harvard Biomedical Data Management

HPC Data Management: Or how (not) to handle your data in an HPC environment

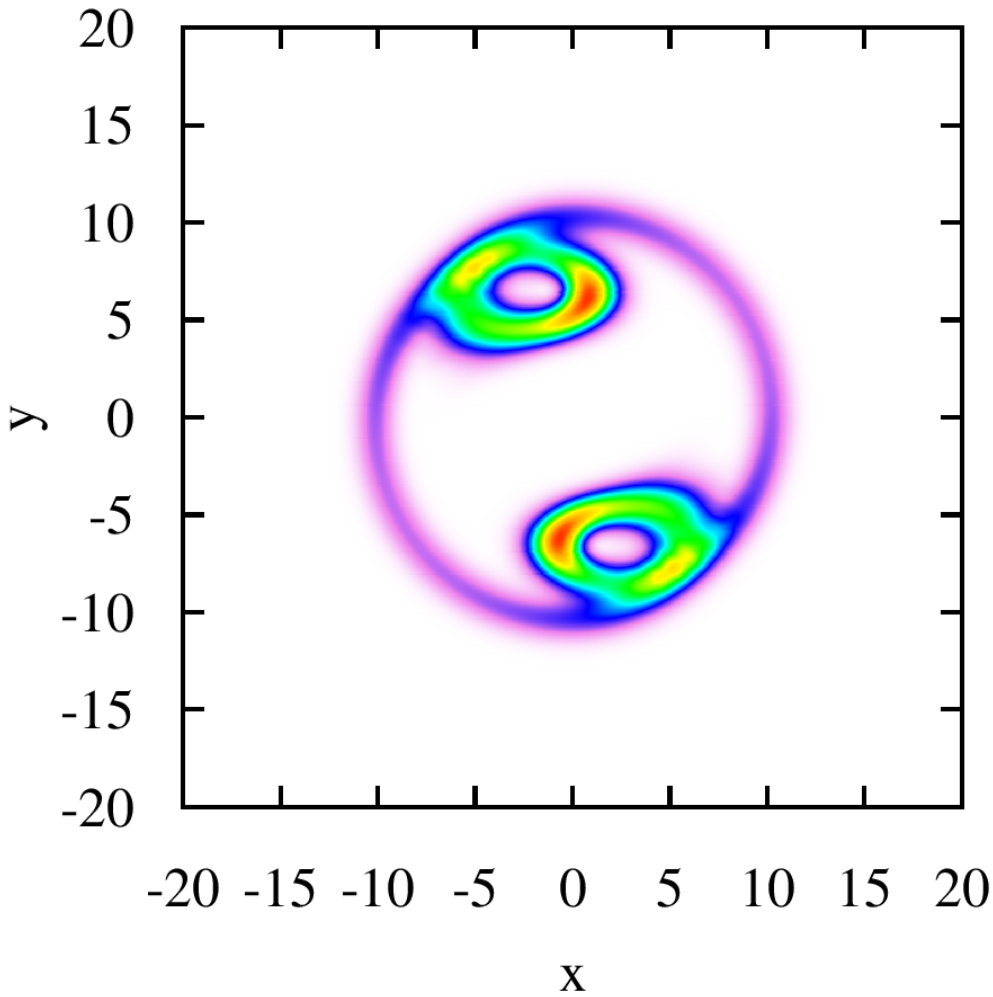
- Case Study: A simulation-visualization workflow
- Getting data to (and from) an HPC system
- Managing data on an HPC system
- Reading, writing, and preparing data for HPC
- Preventing data loss

Case Study: A simulation-visualization workflow

Exploring the Dynamics of a Quantum-Mechanical Compton Generator

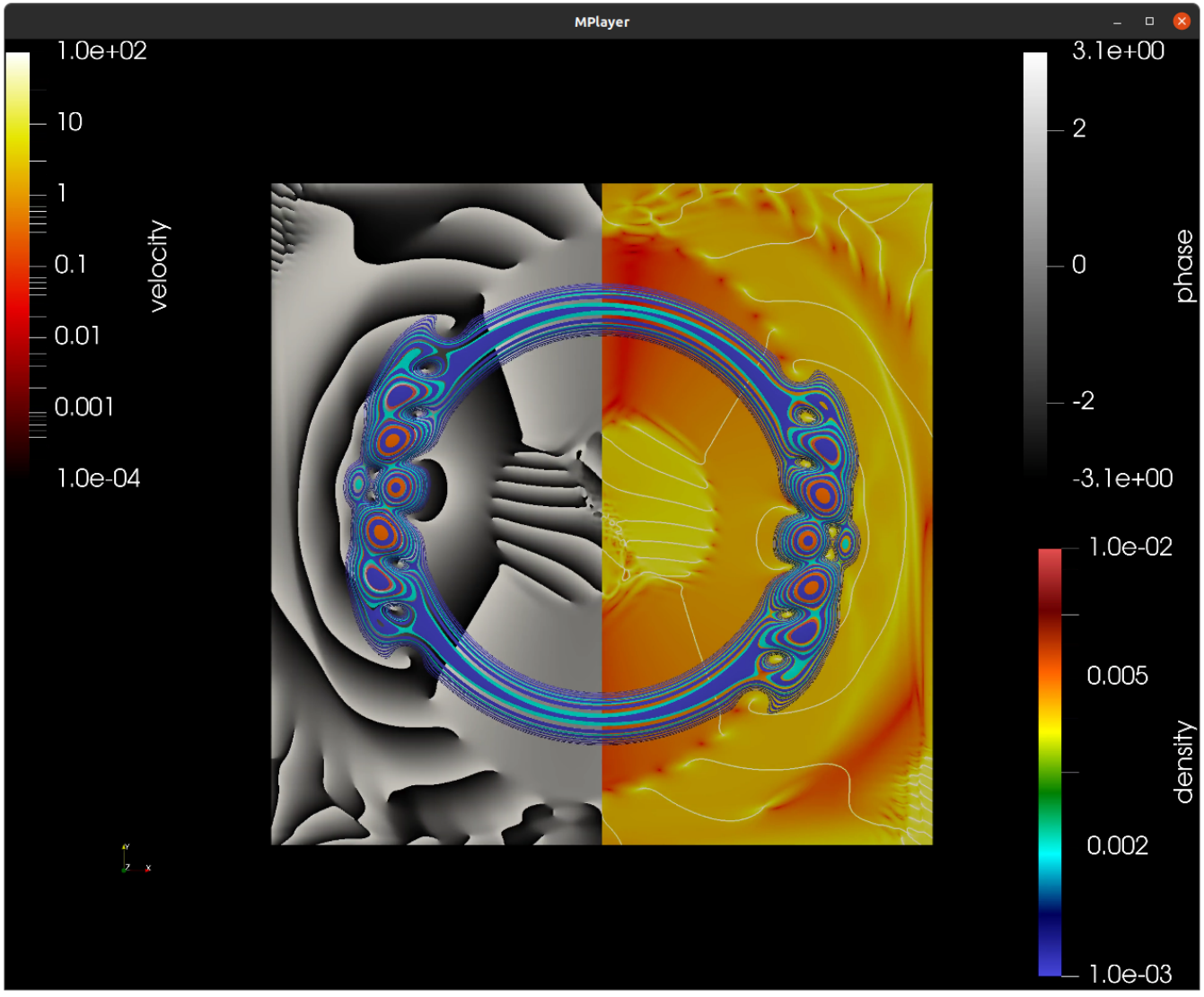
The objective of this XSEDE project is to explore the dynamics of a quantum-mechanical Compton generator through the use of numerical simulations. In this initial research allocation request, we aim to determine the fundamental mechanism by which the Compton-induced drift velocity manifests itself. We hypothesize that the drift velocity likely arises due to the nucleation of quantized vortices within the system. To test this hypothesis, we request **1.8M SUs on Comet for production simulation runs, 32 TB of project storage on Data Oasis to store simulation data, 375K SUs on Maverick for visualization and analysis of this simulated data**, and 100K SUs on both Comet and Stampede, respectively, for code development and testing.

Preliminary (Projected) 2D Visualization



movies/psi2xy.mp4

Full 3D Visualization



movies/gpse-v059-comet-compton-20170828-004-density-phase-velocity-xyz.mp4

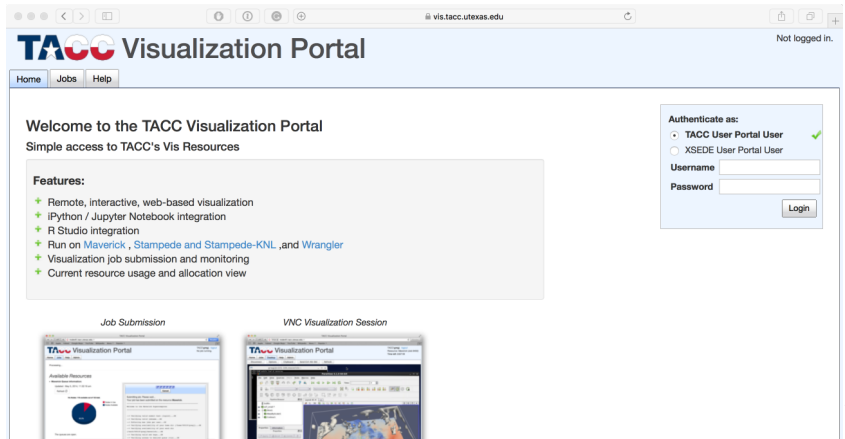
Simulation and Visualization Resources

Simulation on Comet @ SDSC

Visualization on Maverick @ TACC

mkandes@hardtack: ~/Dropbox/software/projects/gpse/source

```
995 ! --- BEGIN MAIN TIME PROPAGATION LOOP --------------------------------
996
997 !     Initialize simulation time
998       tN = t0
999
1000      IF (reanOn .EQV. .FALSE.) THEN ! start main simulation time
1001                                     ! propagation loop
1002
1003         DO n = 0, nTsteps
1004
1005            CALL MPI_BARRIER(MPI_COMM_WORLD, mpiError)
1006
1007 !          Compute partial relations to file from MPI_MASTER; write
1008 !          wave function and external potential to file from MPI_MASTER
1009            IF (MODULO(n, nTwrite) == 0) THEN
1010
1011 !             Compute partial base expectation values locally on each
1012 !             MPI process; reduce partial base expectation values from
1013 !             all MPI processes to MPI_MASTER to get full base
1014 !             expectation values
1015              CALL evua_compute_base(MPI_MASTER, mpiReal, mpiError, 1,&
1016                 & fdOrder, nXa, nXb, nXbc, nYa, nYb, nYbc, nZa, nZb, &
1017                 & nZbc, xO, yO, zO, dX, dY, dZ, gS, Xa, Ya, Za, &
1018                 & Vex3a, Psi3a)
1019
1020 !            Compute derived expectation values, uncertainties from base
```



Simulation data, I/O, and storage requirements

- What is the input data for a simulation?
- How much output data is generated during a simulation?
- What is the file format of the data?
- Does the simlation code use serial or parallel I/O methods?
- Where should the input and output data be stored?
- What are there storage limits on these filesystems?
- How should the data for a simulation be organized?
- Should the data be compressed?

Visualization data, I/O, and storage requirements

- How much simulation data needs to be visualized?
- Where should this simulation data be stored?
- What are there storage limits on these filesystems?
- How should the simulation data be transferred?
- Where should the visualization data be stored?
- Will the visualization data undergo further post-processing?

Simulation and data transfer workflow

- `code/run-gpse-comet.sh`
- `code/gpse.input`
- `code/compress-wave-functions.sh`
- `code/gridftp-transfer.sh`

Simulation batch job script

mkandes@hardtack: ~/Dropbox/work/ucsd/sdsc/presentations/2022/sdsc/ciml/data-...

```
#!/usr/bin/env bash

#SBATCH --job-name=gpse-v0.5.9-comet-compton-20170828-017
#SBATCH --account=sdu118
#SBATCH --partition=compute
#SBATCH --nodes=6
#SBATCH --ntasks-per-node=24
#SBATCH --cpus-per-task=1
#SBATCH --mem=120G
#SBATCH --time=48:00:00
#SBATCH --output=%x.o%j.%N

declare -xr LOCAL_SCRATCH_DIR="/scratch/${USER}/${SLURM_JOB_ID}"
declare -xr LUSTRE_SCRATCH_DIR="/oasis/scratch/comet/${USER}/temp_project/sdu118"
declare -xr LUSTRE_PROJECTS_DIR="/oasis/projects/nsf/sdu118/${USER}"

declare -xr COMPILER_MODULE='intel/2016.3.210'
declare -xr MPI_MODULE='intelmpi/2016.3.210'
declare -xr GNUTOOLS_MODULE='gnutools/2.69'

module purge
module load "${COMPILER_MODULE}"
module load "${MPI_MODULE}"
module load "${GNUTOOLS_MODULE}"
module list
export PATH="${HOME}/software/gpse/0.5.9/bin:${PATH}"
printenv

mkdir -p "${LUSTRE_SCRATCH_DIR}/${SLURM_JOB_NAME}"
cd "${LUSTRE_SCRATCH_DIR}/${SLURM_JOB_NAME}"
cp "${SLURM_SUBMIT_DIR}/gpse.input" ./

time -p mpirun -n "${SLURM_NTASKS}" gpse.x
                                                    1,1           All
```

Simulation input file

mkandes@hardtack: ~/Dropbox/work/ucsd/sdsc/presentations/2022/sdsc/ciml/data-management/code

```
&gpseIn
  itpOn     = .FALSE.     ! Perform imaginary time propagation? .TRUE. = Yes ; .FALSE. = No
  rk4Lambda = 2           ! 1 = Tan-Chen-1 ; 2 = Classical 4th-Order Runge-Kutta ; 3 = Tan-Chen Lambda-3 ; 4 = England ; 5 = Tan-Chen-5
  fdOrder   = 2           ! 2 = 2nd-Order Central Differences ( CD ); 4 = 4th-Order CD ; 6 = 6th-Order CD; 8 = 8th-Order CD
  nTsteps   = 155520      ! Total number of time steps in simulation
  nTwrite   = 648         ! Period of IO writes to disk; i.e., number of time steps between writes to disk
  nX        = 1152        ! Number of grid points along the x-axis
  nY        = 1152        ! Number of grid points along the y-axis
  nZ        = 576         ! Number of grid points along the z-axis
  dNx       = 1           ! Write out wave function only every dNx grid points along the x-axis
  dNy       = 1           ! Write out wave function only every dNy grid points along the y-axis
  dNz       = 1           ! Write out wave function only every dNz grid points along the z-axis
  t0        = 0.0         ! Time at the beginning of the simulation
  tF        = 30.0        ! Time at the end of the simulation; only use for compton generator
  xO        = 0.0         ! X-coordinate of the centre of the grid
  yO        = 0.0         ! Y-coordinate of the centre of the grid
  zO        = 0.0         ! Z-coordinate of the centre of the grid
  dT        = 1.929012345679e-4 ! Interval of a time step
  dX        = 0.02777777777 ! Distance between grid points along the x-axis
  dY        = 0.02777777777 ! Distance between grid points along the y-axis
  dZ        = 0.02777777777 ! Distance between grid points along the z-axis
  xOrrf     = 0.0         ! X-coordinate of the rotating reference frame's origin
  yOrrf     = 0.0         ! Y-coordinate of the rotating reference frame's origin
  zOrrf     = 0.0         ! Z-coordinate of the rotating reference frame's origin
  wX        = 0.0         ! X-component of the rotating reference frame's angular velocity vector
  wY        = 0.0         ! Y-component of the rotating reference frame's angular velocity vector
  wZ        = 0.5         ! Z-component of the rotating reference frame's angular velocity vector
  gS        = 0.0         ! Nonlinear atom-atom interaction coupling constant
  psiInput  = 1           ! 0 = No input wave function; 1 = Read wave function from .bin file; 2 = Read wave function from .vtk file ( not available yet )
  psiOutput = 3           ! 0 = No output wave function; 1 = Write wave function to .bin file; 2 = Write wave function to .vtk file; 5 = Write int den to .splot
  psiFileNo = 500         ! Set psiFileNo to input wave function file number
  psiInit   = 3           ! 0 = Isotropic 3D SHO ; 1 = Anisotropic 3D SHO ; 2 = Axisymmetric 3D SHO ; 3 = Approx 3D SHOR
  nXpsi     = 0           ! Degree of Hermite polynomial used to define anisotropic SHO wave function along x-axis
  nYpsi     = 0           ! Degree of Hermite polynomial used to define anisotropic SHO wave function along y-axis
  nZpsi     = 0           ! Degree of Hermite polynomial used to define both anisotropic and axially-symmetric SHO wave functions along z-axis
  nRpsi     = 0           ! Degree of (associated) Laguerre polynomials used to define radial components of isotropic and axially-symmetric SHO wave functions
  mLpsi     = 0           ! Projection of orbital angular momentum along z-axis for axially-symmetric SHO wave function
  xOpsi     = 0.0         ! X-coordinate of origin used to define initial wave function
  yOpsi     = 0.0         ! Y-coordinate of origin used to define initial wave function
  zOpsi     = 0.0         ! Z-coordinate of origin used to define initial wave function
  rOpsi     = 10.0        ! Radius of psiInit = 3 SHOR approx wavefunction
  wXpsi     = 0.0         ! Angular frequency of SHO potential along x-axis used to define anisotropic SHO wave function
  wYpsi     = 0.0         ! Angular frequency of SHO potential along y-axis used to define anisotropic SHO wave function
  wZpsi     = 1.0        ! Angular frequency of SHO potential along z-axis used to define both anisotropic and axially-symmetric SHO wave functions
  wRpsi     = 1.0         ! Radial angular frequency of axially-symmetric SHO or SHOR potential used to define axially-symmetric SHO or approx SHOR wave functions
  pXpsi     = 0.0         ! Initial linear momentum boost along x-axis
  pYpsi     = 0.0         ! Initial linear momentum boost along y-axis
"gpse.input" 64L, 5703C                                                                                    1,2           Top
```

Post-simulation data compression script

mkandes@hardtack: ~/Dropbox/work/ucsd/sdsc/presentations/2022/sdsc/ciml/data...

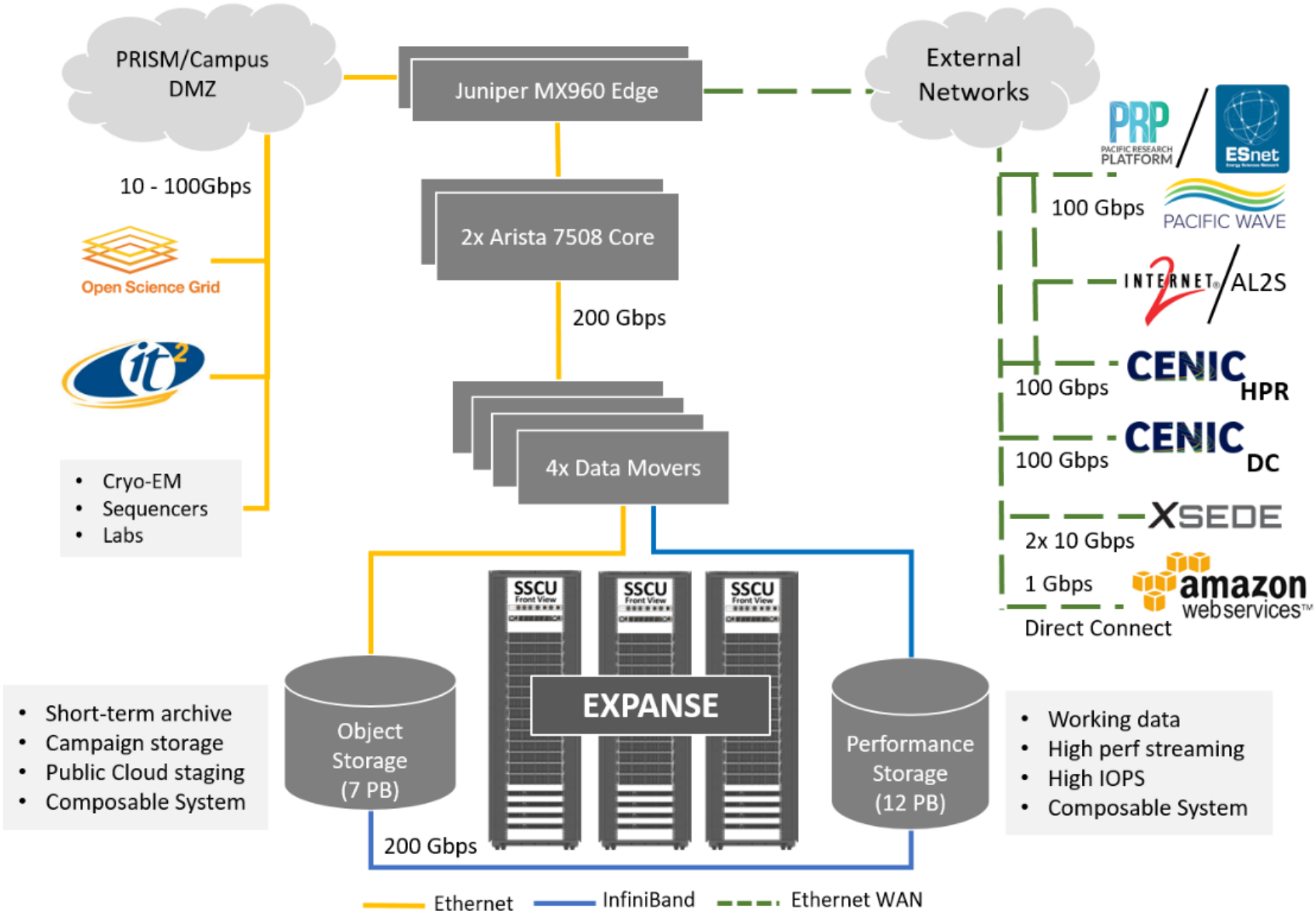
```
#!/usr/bin/env bash

declare -xr LUSTRE_SCRATCH_DIR="/oasis/scratch/comet/${USER}/temp_project"
declare -xr LUSTRE_PROJECTS_DIR="/oasis/projects/nsf/sdu118/${USER}"

declare -xr JOB_NAME='gpse-v0.5.9-comet-compton-20170828-013'

declare -xir START_TIMESTEP=1000
declare -xir END_TIMESTEP=1240

cd "${LUSTRE_SCRATCH_DIR}/sdu118/${JOB_NAME}"

for ((x="${START_TIMESTEP}"; x<="${END_TIMESTEP}"; x++)); do
  mkdir -p psi-${x}
  mv psi-${x}-*.vtk psi-${x}
  zip -r psi-${x}.zip psi-${x}
  rm -rf psi-${x}
done
                                                    1,1           All
```

Data transfer script

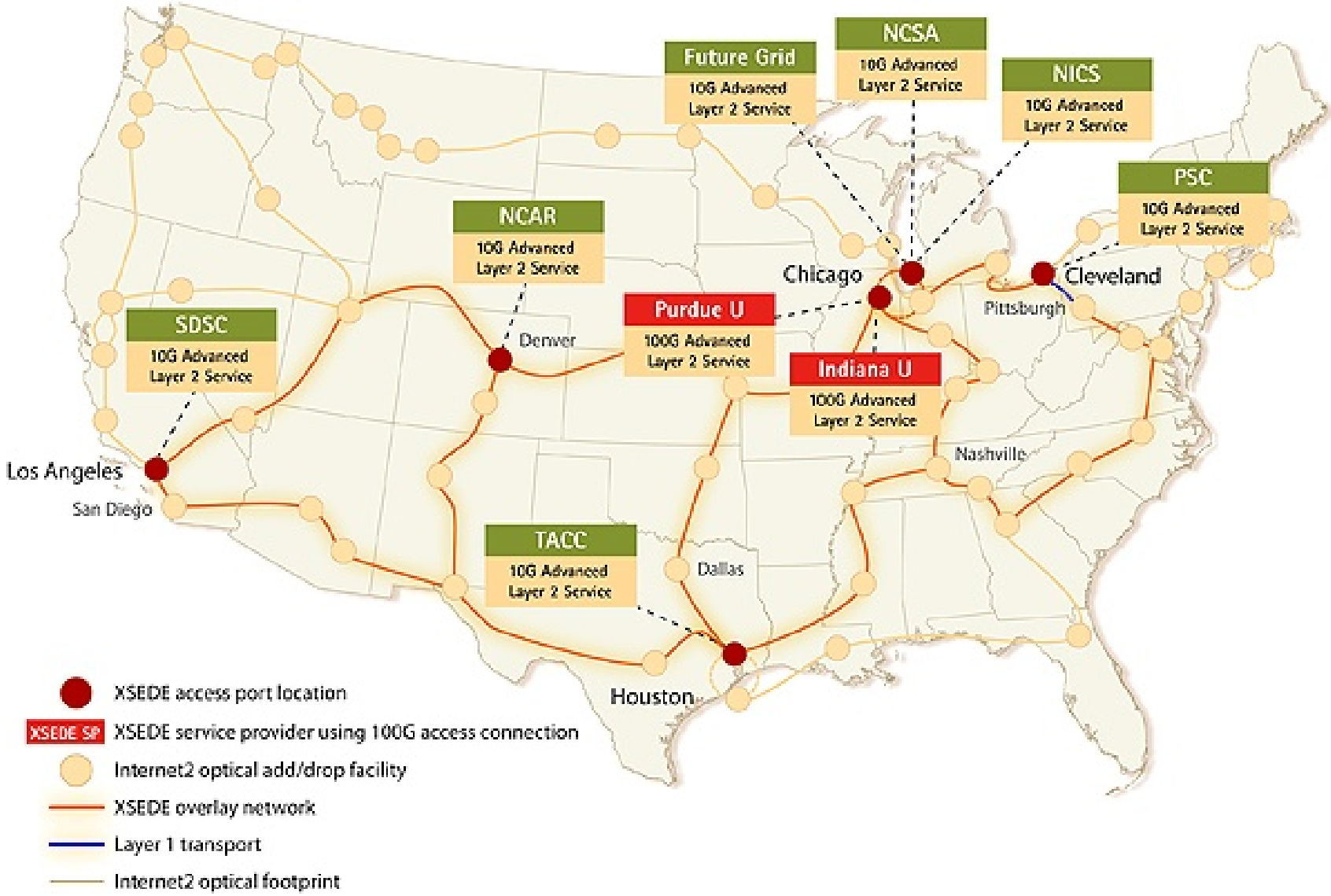
mkandes@hardtack: ~/Dropbox/work/ucsd/sdsc/presentations/2022/sdsc/ciml/data...

```
#!/usr/bin/env bash

declare -xr SOURCE_ENDPOINT='oasis-dm.sdsc.xsede.org:2811'
declare -xr SOURCE_DIR='/oasis/projects/nsf/sdu118/mkandes/gpse-v0.5.9-comet-com
pton-20170828-001'

declare -xr DESTINATION_ENDPOINT='gridftp.stampede2.tacc.xsede.org:2811'
declare -xr DESTINATION_DIR='/work/03216/mckandes/maverick/gpse-v0.5.9-comet-com
pton-20170828-001'

module load globus

for timestep in {1000..1024}; do
  globus-url-copy -vb -stripe -tcp-bs 8m -p 4 "gsiftp://${SOURCE_ENDPOINT}/${SOU
RCE_DIR}/psi-${timestep}.zip" "gsiftp://${DESTINATION_ENDPOINT}/${DESTINATION_DI
R}/psi-${timestep}.zip"
done
                                                              1,1           Top
```

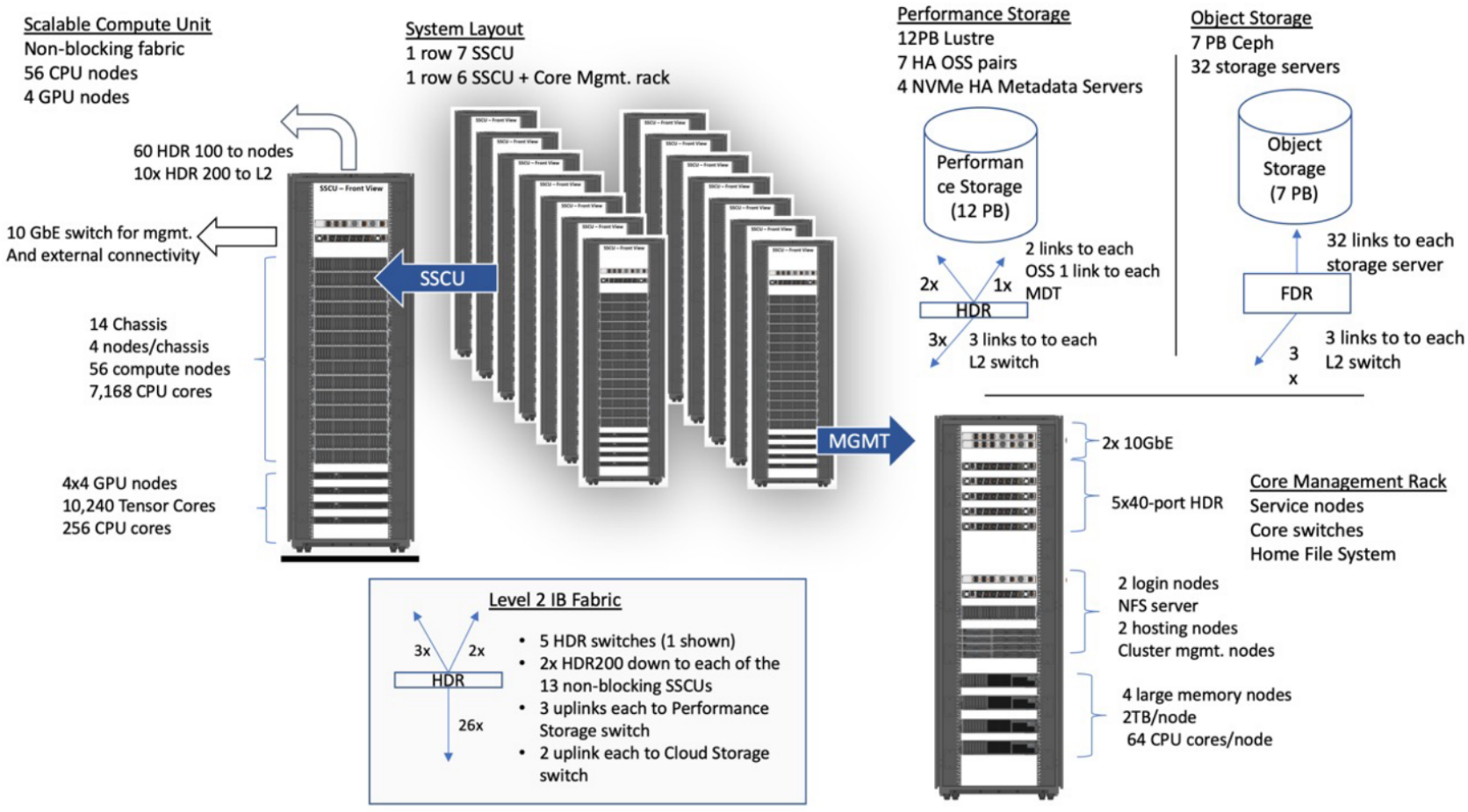
Expanse storage systems and external networks



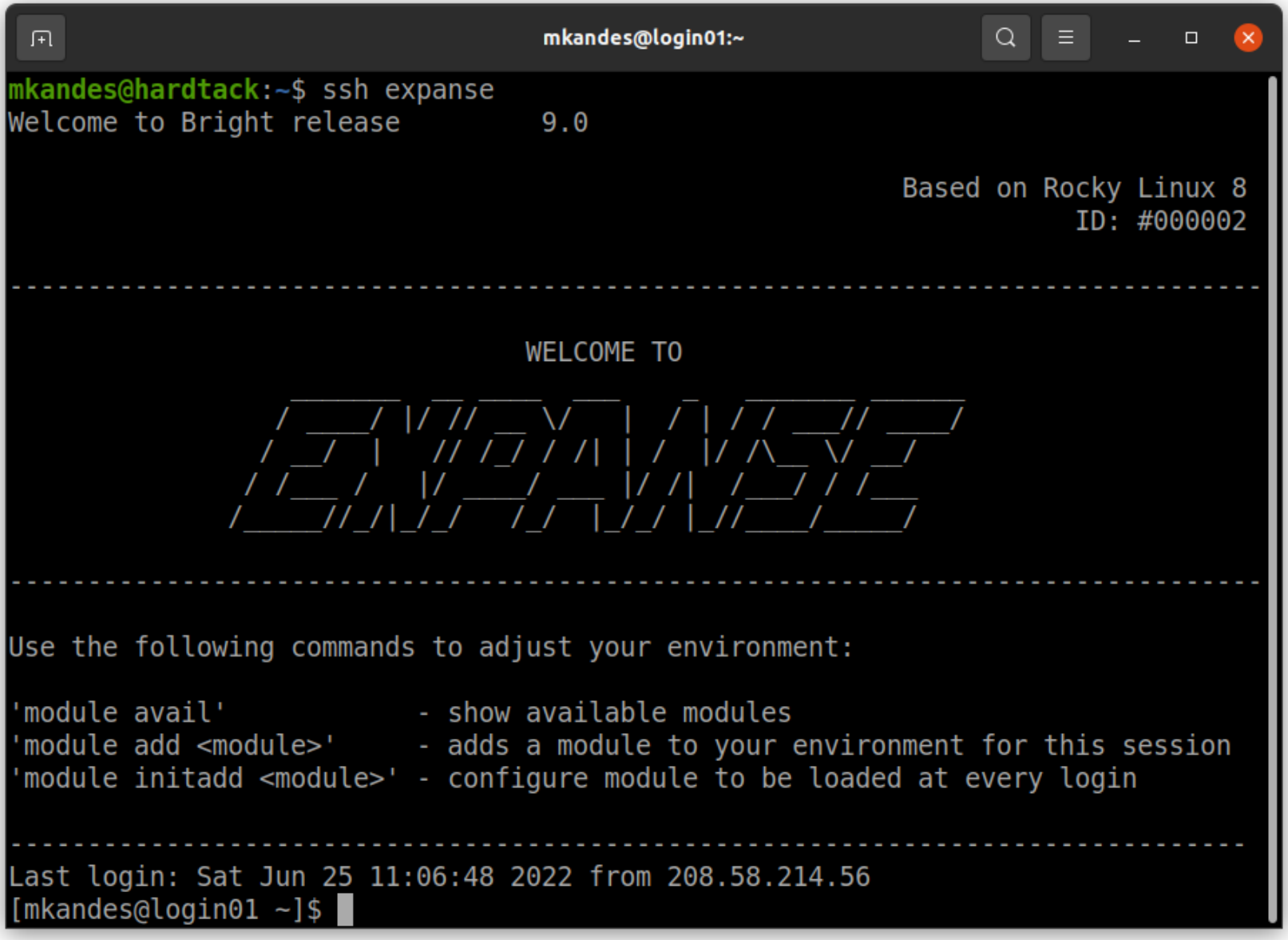
Research and Education Networks



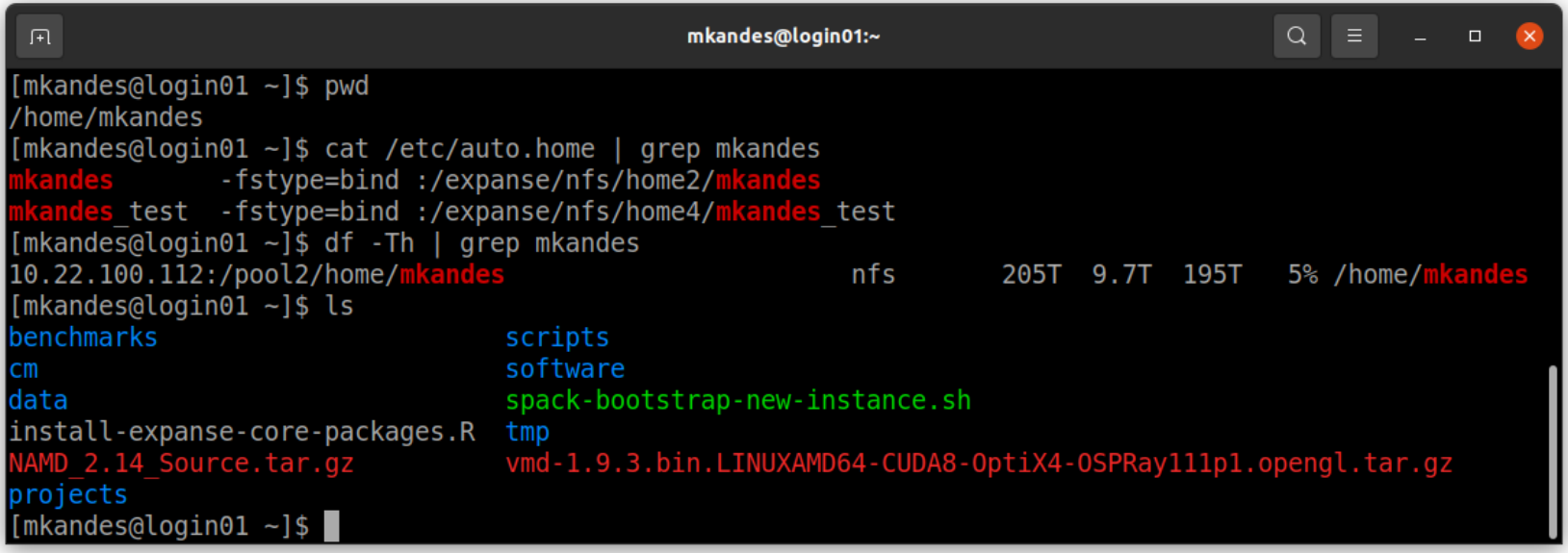
Expanse System Architecture



Expanse Login Nodes

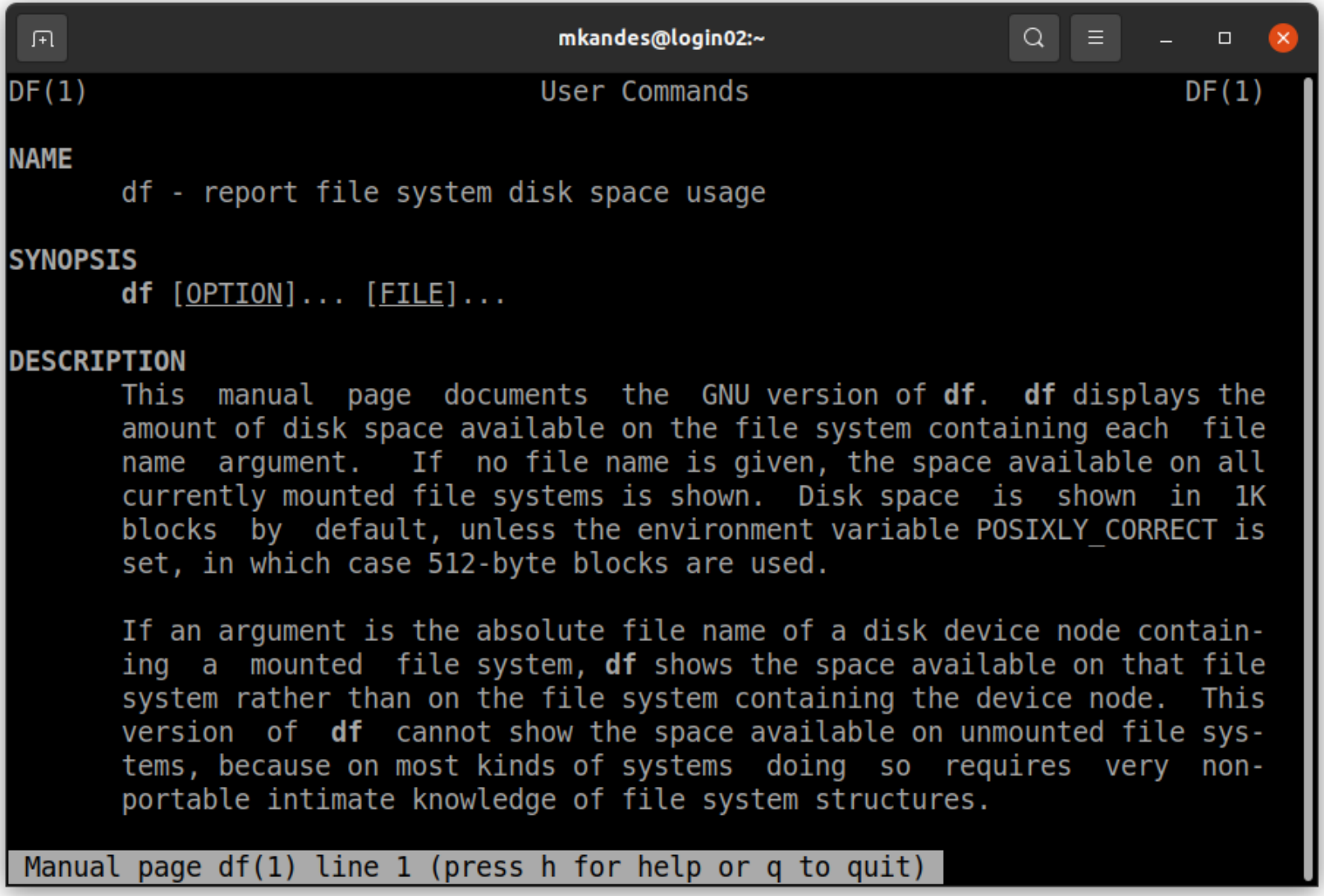


Expanse $HOME Network File System (NFS)



- Usage: Source code, compiled software binaries and libraries, small input/output files, batch job scripts, other important files such as containers
- Quota: approximately 100 GB
- Backup: ZFS snapshots performed once or twice per month
- Purge: 90 days after allocation expiration

df


mkandes@login02:~
DF(1) User Commands DF(1)
NAME
df - report file system disk space usage
SYNOPSIS
df [OPTION]... [FILE]...
DESCRIPTION
This manual page documents the GNU version of df. df displays the amount of disk space available on the file system containing each file name argument. If no file name is given, the space available on all currently mounted file systems is shown. Disk space is shown in 1K blocks by default, unless the environment variable POSIXLY_CORRECT is set, in which case 512-byte blocks are used.
If an argument is the absolute file name of a disk device node containing a mounted file system, df shows the space available on that file system rather than on the file system containing the device node. This version of df cannot show the space available on unmounted file systems, because on most kinds of systems doing so requires very nonportable intimate knowledge of file system structures.
Manual page df(1) line 1 (press h for help or q to quit)

Type(s) of I/O

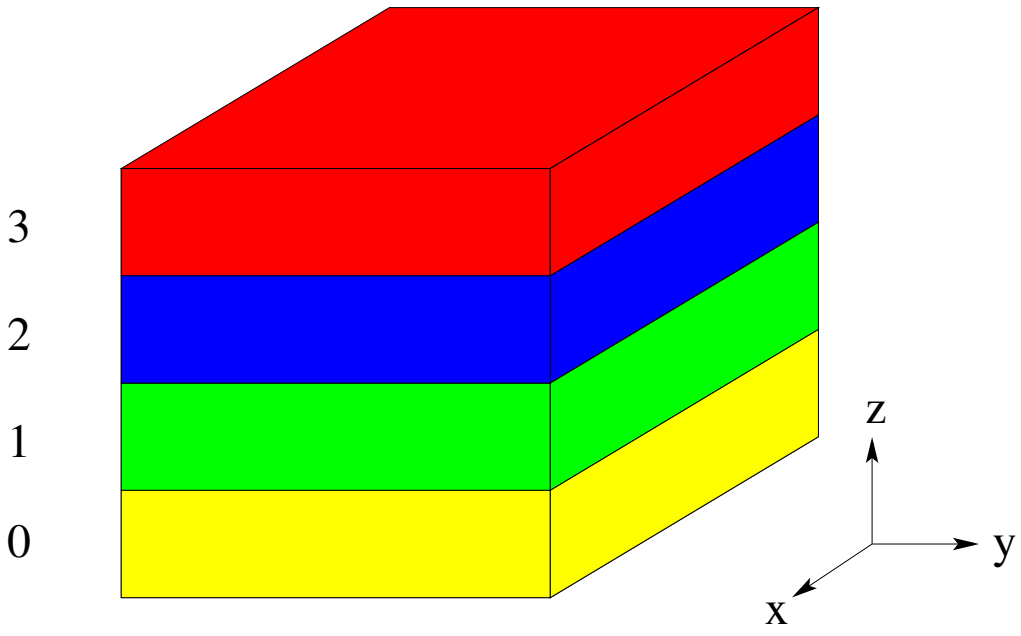
```
mkandes@login01:~/software/gpse/0.6.5/mpi-omp/gcc/10.2.0/openmpi/4.0.4/gpse/s...
[mkandes@login01 ~]$ cd ~/software/gpse/0.6.5/mpi-omp/gcc/10.2.0/openmpi/4.0.4/g
pse/
[mkandes@login01 gpse]$ ls
build  CHANGELOG  gpse.input  gpse.x  LICENSE  Makefile  README  source
[mkandes@login01 gpse]$ cd source/
[mkandes@login01 source]$ ls
evua.f90  grid.f90  io.f90    pmca.f90  rot.f90
gpse.f90  grk4.f90  math.f90  psi.f90   vex.f90
[mkandes@login01 source]$ grep -H psiOutput *.f90
gpse.f90:         & psiInput, psiOutput, psiFileNo, psiFileNoChkpt, psiInit, &
gpse.f90:               IF (psiOutput == 1) THEN
gpse.f90:               ELSE IF (psiOutput == 2) THEN
gpse.f90:               ELSE IF (psiOutput == 3) THEN
gpse.f90:               ELSE IF (psiOutput == 4) THEN
gpse.f90:               ELSE IF (psiOutput == 5) THEN
gpse.f90:          WRITE(UNIT=OUTPUT_UNIT, FMT=*) '#        psiOutput = ', psiOut
put
gpse.f90:       CALL MPI_BCAST(psiOutput, 1, mpiInt, mpiMaster, MPI_COMM_WORLD, m
piError)
psi.f90:       INTEGER, PUBLIC :: psiOutput = -1
psi.f90:!      psiOutput is a PUBLIC, INTEGER-valued variable that sets the file
[mkandes@login01 source]$
```

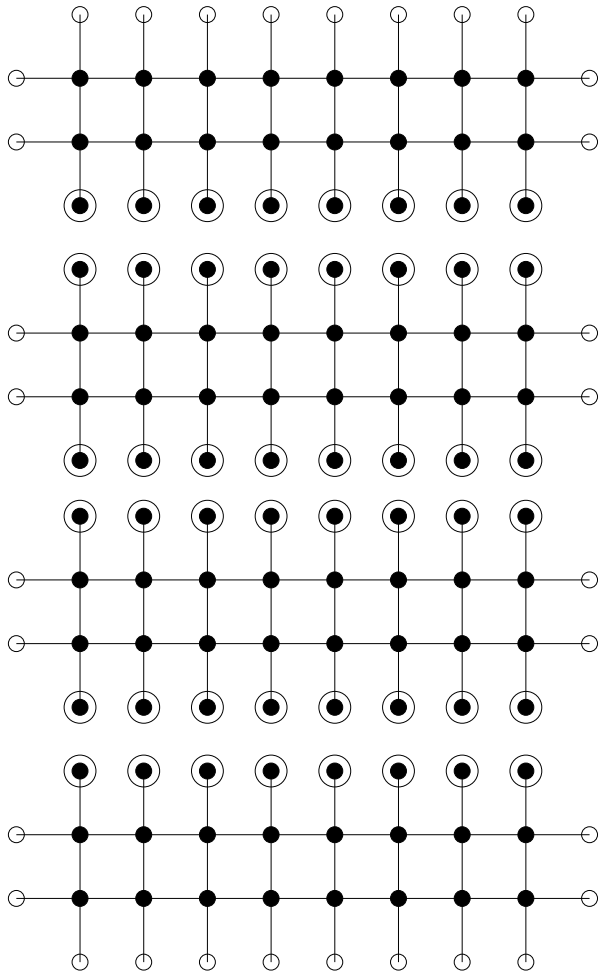
Simple parallel I/O - one file per process

mkandes@login01:~/software/gpse/0.6.5/mpi-omp/gcc/10.2.0/openmpi/4.0.4/gpse/s...

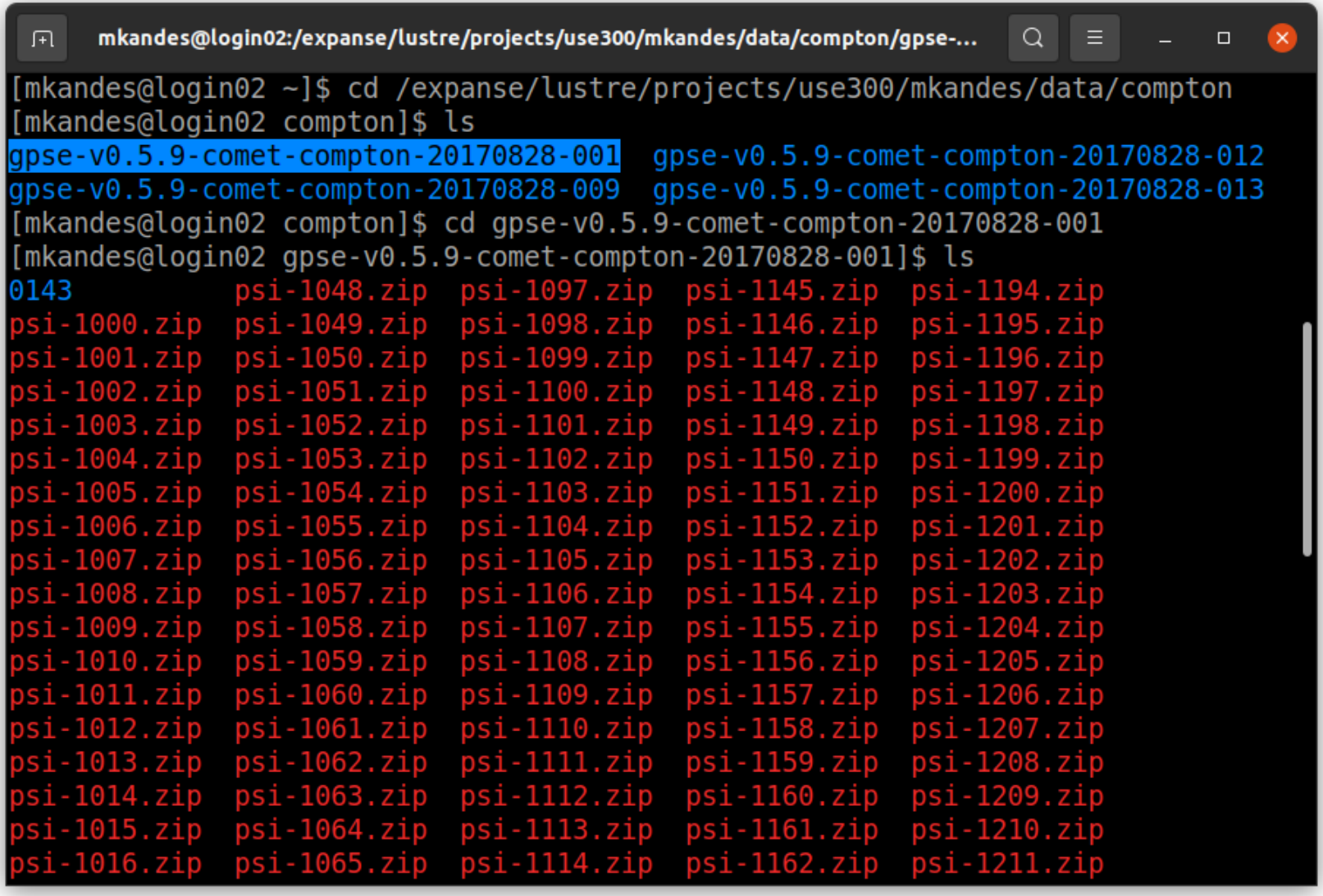
```
                END DO

                psiFileNo = psiFileNo + 1

!            Use simple parallel I/O to write out the complete wave
!            function simultaneously to many legacy VTK files that
!            only contain a single, one-dimensional slab of the wave
!            function overseen by an MPI process
             ELSE IF (psiOutput == 3) THEN

                CALL io_write_vtk('psi-', psiFileNo, mpiRank, nX, &
                   & nXa, nXb, nXbc, dNx, nY, nYa, nYb, nYbc, dNy, &
                   & nZ, nZa, nZb, nZbc, dNz, Xa, Ya, Za, Psi3a)
                psiFileNo = psiFileNo + 1

!            Use MPI-I/O to write the wave function out to a single
!            binary file in parallel
             ELSE IF (psiOutput == 4) THEN

                WRITE(UNIT=fileUnitChar, FMT='(I4.4)') psiFileNo
                CALL MPI_FILE_OPEN ( MPI_COMM_WORLD, &
                   & TRIM ( 'psi-'//fileUnitChar//'.bin' ), &
                   & MPI_MODE_CREATE + MPI_MODE_WRONLY, &
                                                    1155,25       53%
```

1D (Slab) Domain Decomposition

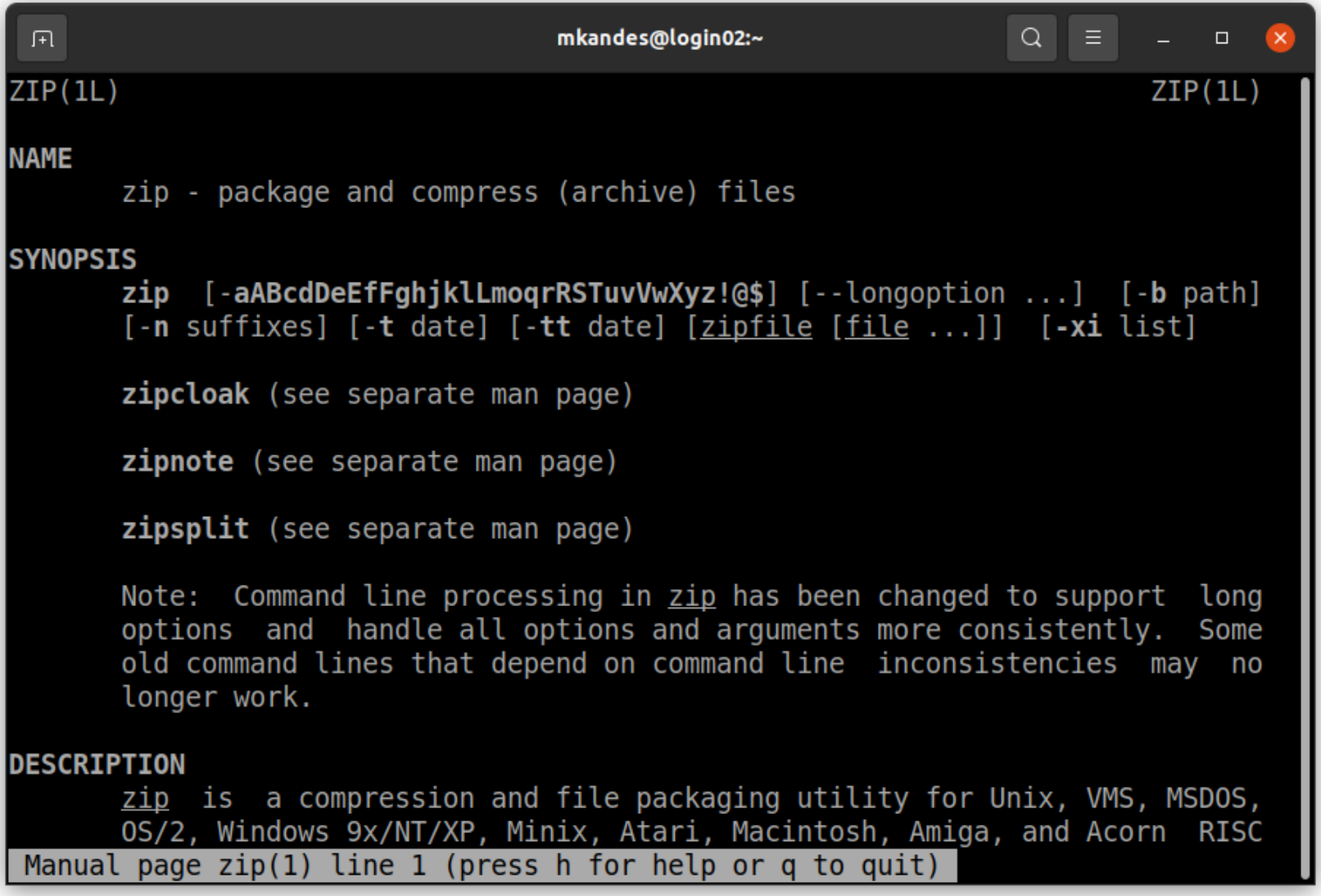




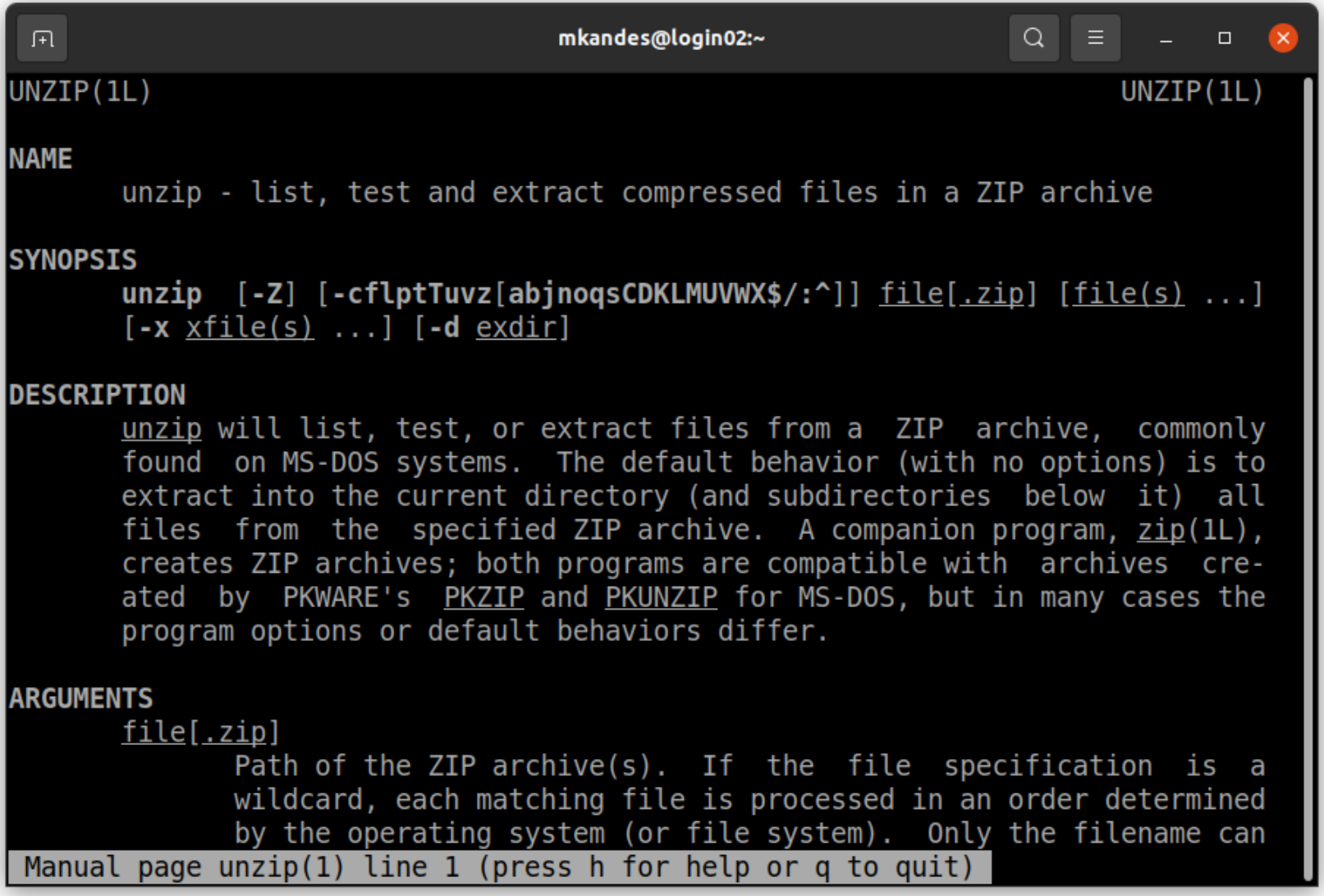
Simulation data - one zip archive per timestep



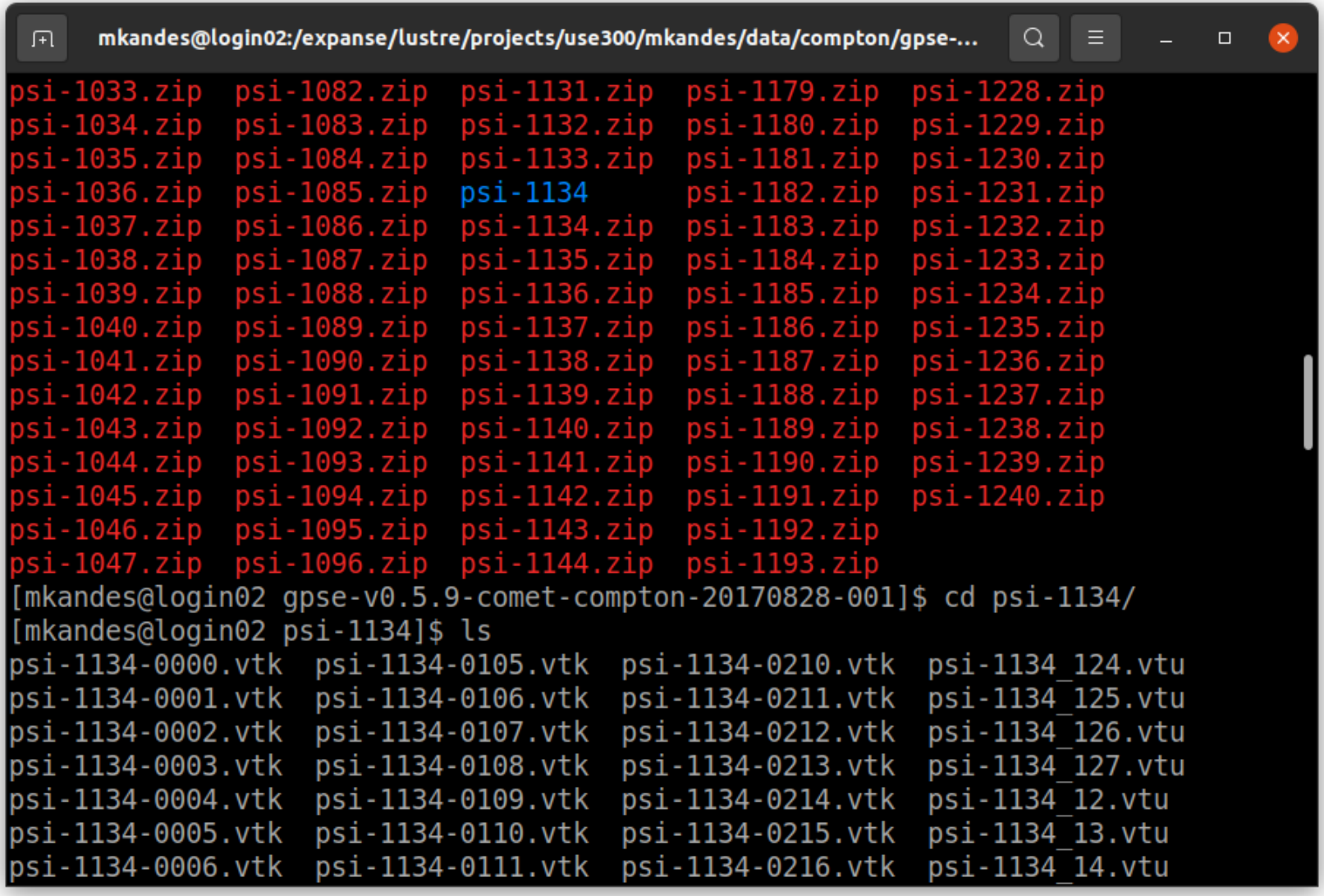
zip


mkandes@login02:~
ZIP(1L) ZIP(1L)
NAME
zip - package and compress (archive) files
SYNOPSIS
zip [-aABcdDeEfFghjklLmoqrRSTuvVwXyz!@$] [--longoption ...] [-b path]
[-n suffixes] [-t date] [-tt date] [zipfile [file ...]] [-xi list]
zipcloak (see separate man page)
zipnote (see separate man page)
zipsplit (see separate man page)
Note: Command line processing in zip has been changed to support long
options and handle all options and arguments more consistently. Some
old command lines that depend on command line inconsistencies may no
longer work.
DESCRIPTION
zip is a compression and file packaging utility for Unix, VMS, MSDOS,
OS/2, Windows 9x/NT/XP, Minix, Atari, Macintosh, Amiga, and Acorn RISC
Manual page zip(1) line 1 (press h for help or q to quit)

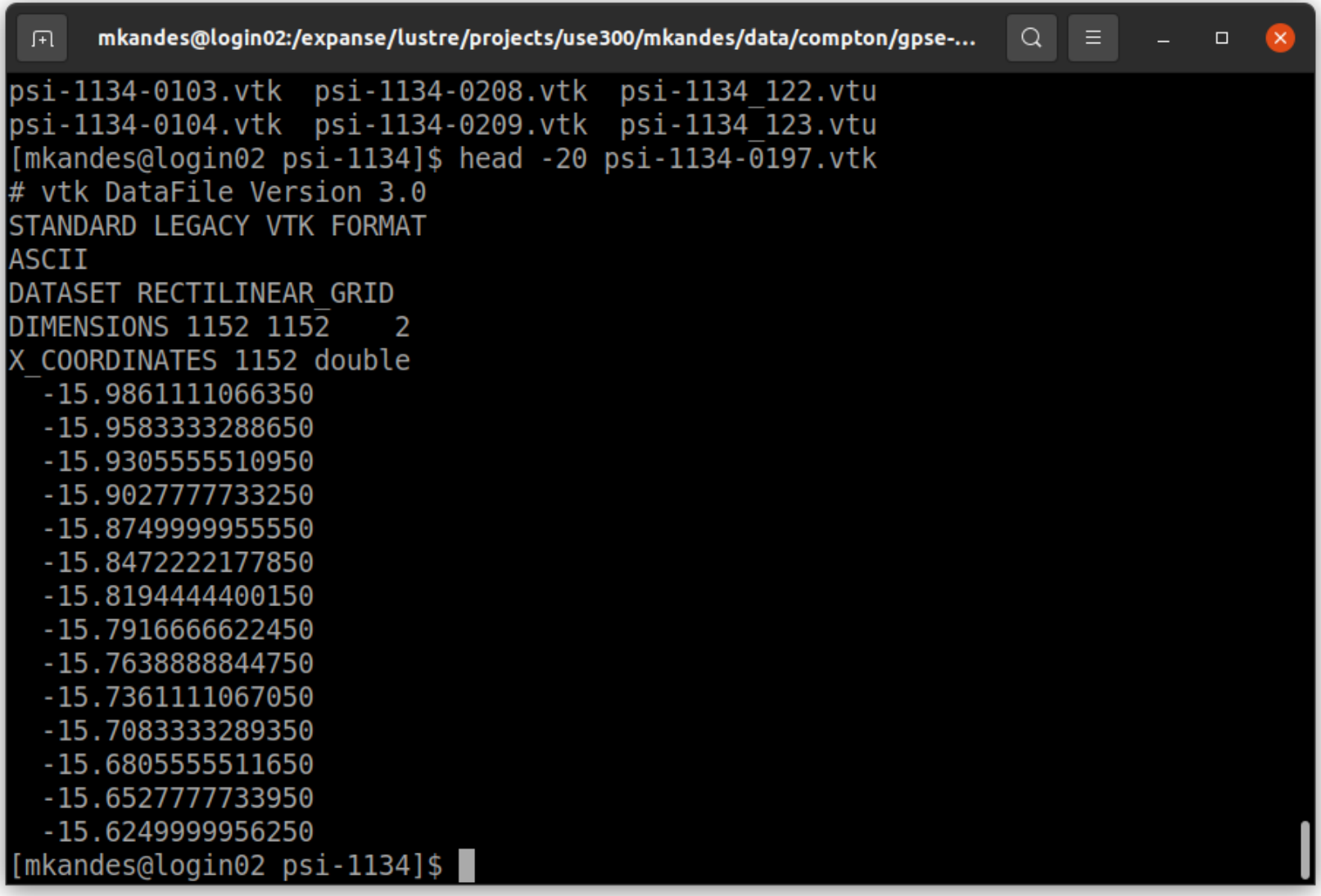
unzip


mkandes@login02:~
UNZIP(1L) UNZIP(1L)
NAME
unzip - list, test and extract compressed files in a ZIP archive
SYNOPSIS
unzip [-Z] [-cflptTuvz[abjnoqsCDKLMUVWX$/:^]] file[.zip] [file(s) ...]
[-x xfile(s) ...] [-d exdir]
DESCRIPTION
unzip will list, test, or extract files from a ZIP archive, commonly
found on MS-DOS systems. The default behavior (with no options) is to
extract into the current directory (and subdirectories below it) all
files from the specified ZIP archive. A companion program, zip(1L),
creates ZIP archives; both programs are compatible with archives cre-
ated by PKWARE's PKZIP and PKUNZIP for MS-DOS, but in many cases the
program options or default behaviors differ.
ARGUMENTS
file[.zip]
Path of the ZIP archive(s). If the file specification is a
wildcard, each matching file is processed in an order determined
by the operating system (or file system). Only the filename can
Manual page unzip(1) line 1 (press h for help or q to quit)

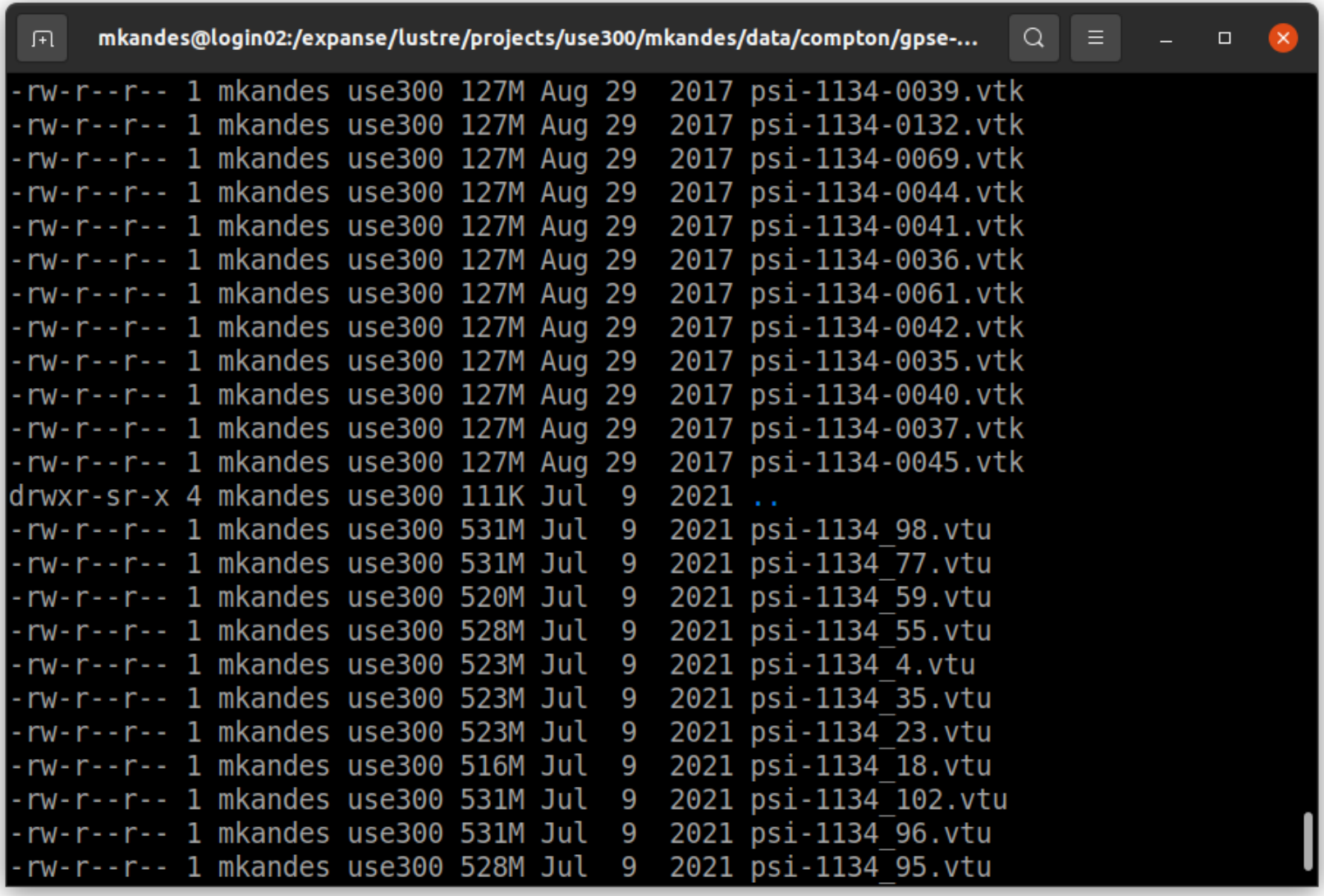
Simulation data - one vtk file per slab (process)



Simulation data - vtk file format



vtk vs. vtu

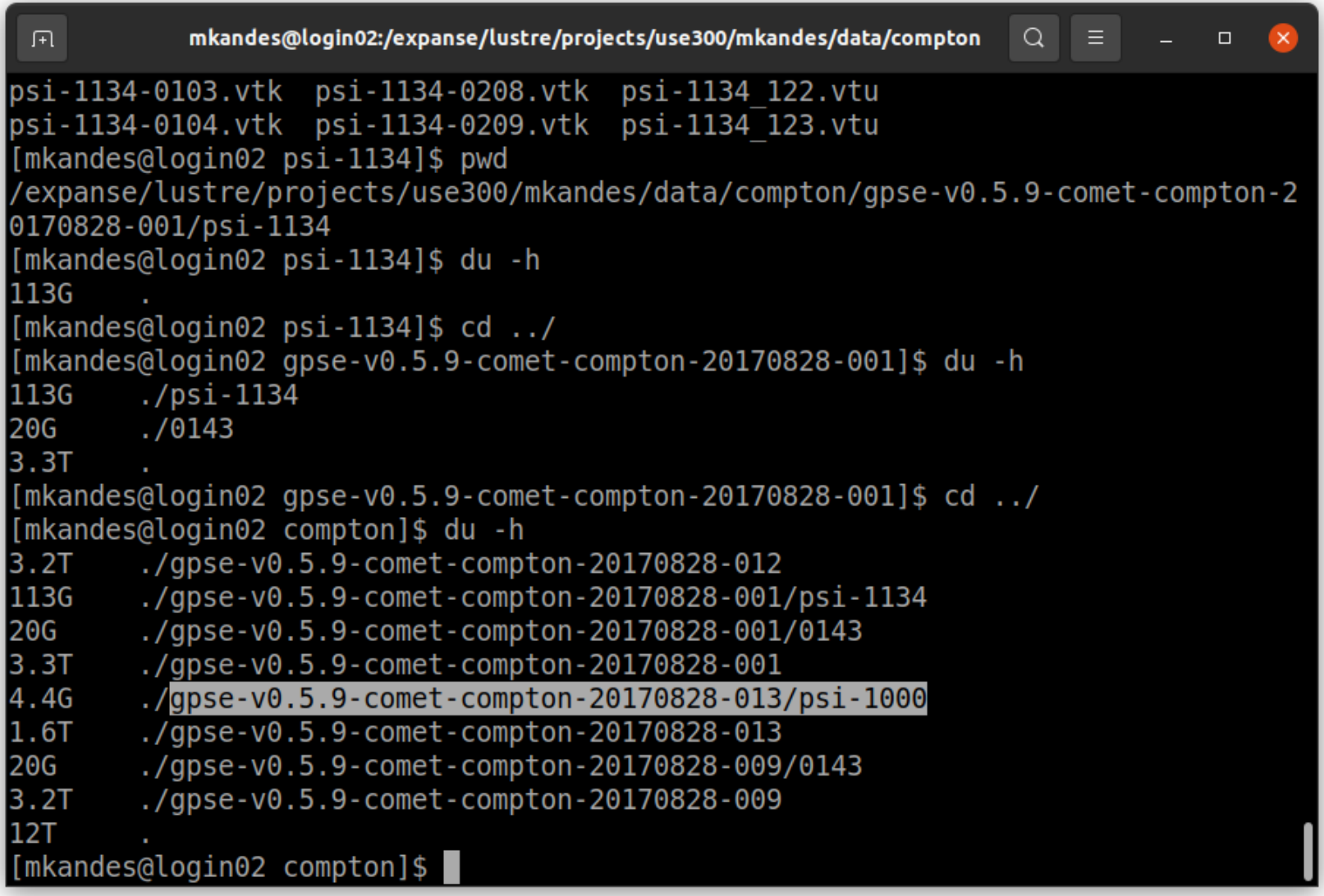

mkandes@login02:/expanse/lustre/projects/use300/mkandes/data/compton/gpse-...
-rw-r--r-- 1 mkandes use300 127M Aug 29 2017 psi-1134-0039.vtk
-rw-r--r-- 1 mkandes use300 127M Aug 29 2017 psi-1134-0132.vtk
-rw-r--r-- 1 mkandes use300 127M Aug 29 2017 psi-1134-0069.vtk
-rw-r--r-- 1 mkandes use300 127M Aug 29 2017 psi-1134-0044.vtk
-rw-r--r-- 1 mkandes use300 127M Aug 29 2017 psi-1134-0041.vtk
-rw-r--r-- 1 mkandes use300 127M Aug 29 2017 psi-1134-0036.vtk
-rw-r--r-- 1 mkandes use300 127M Aug 29 2017 psi-1134-0061.vtk
-rw-r--r-- 1 mkandes use300 127M Aug 29 2017 psi-1134-0042.vtk
-rw-r--r-- 1 mkandes use300 127M Aug 29 2017 psi-1134-0035.vtk
-rw-r--r-- 1 mkandes use300 127M Aug 29 2017 psi-1134-0040.vtk
-rw-r--r-- 1 mkandes use300 127M Aug 29 2017 psi-1134-0037.vtk
-rw-r--r-- 1 mkandes use300 127M Aug 29 2017 psi-1134-0045.vtk
drwxr-sr-x 4 mkandes use300 111K Jul 9 2021 ..
-rw-r--r-- 1 mkandes use300 531M Jul 9 2021 psi-1134_98.vtu
-rw-r--r-- 1 mkandes use300 531M Jul 9 2021 psi-1134_77.vtu
-rw-r--r-- 1 mkandes use300 520M Jul 9 2021 psi-1134_59.vtu
-rw-r--r-- 1 mkandes use300 528M Jul 9 2021 psi-1134_55.vtu
-rw-r--r-- 1 mkandes use300 523M Jul 9 2021 psi-1134_4.vtu
-rw-r--r-- 1 mkandes use300 523M Jul 9 2021 psi-1134_35.vtu
-rw-r--r-- 1 mkandes use300 523M Jul 9 2021 psi-1134_23.vtu
-rw-r--r-- 1 mkandes use300 516M Jul 9 2021 psi-1134_18.vtu
-rw-r--r-- 1 mkandes use300 531M Jul 9 2021 psi-1134_102.vtu
-rw-r--r-- 1 mkandes use300 531M Jul 9 2021 psi-1134_96.vtu
-rw-r--r-- 1 mkandes use300 528M Jul 9 2021 psi-1134_95.vtu

Visualization data - vtu file format

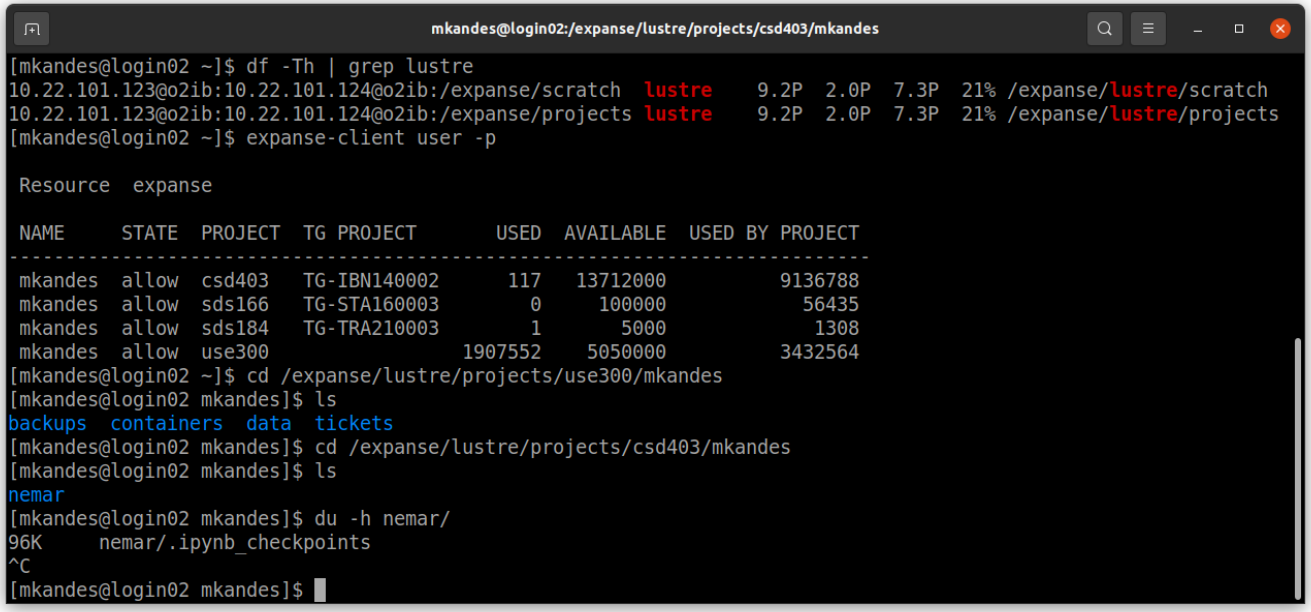
mkandes@login02:/expanse/lustre/projects/use300/mkandes/data/compton/gpse-...

```
[mkandes@login02 psi-1134]$ head -11 psi-1134_98.vtu
<VTKFile type="UnstructuredGrid" version="2.2" byte_order="LittleEndian" header_
type="UInt64">
  <UnstructuredGrid>
    <Piece NumberOfPoints="6223392"             NumberOfCells="3069504"
     >
      <PointData Scalars="RePsi" GlobalIds="___D3___GlobalNodeIds">
        <DataArray type="Float64" Name="ImPsi" format="appended" RangeMin="-0.00
12450472664"     RangeMax="0.00065853427111"     offset="0"                    />
        <DataArray type="Float64" Name="RePsi" format="appended" RangeMin="-0.00
093535214815"    RangeMax="0.001508906848"       offset="49787144"             />
        <DataArray type="Int64" IdType="1" Name="___D3___GlobalNodeIds" format="
appended" RangeMin="412086382"           RangeMax="686755584"           offset
="99574288"           />
        <DataArray type="UInt8" Name="vtkGhostType" format="appended" RangeMin="
0"                   RangeMax="1"                    offset="149361432"
   />
      </PointData>
      <CellData>
        <DataArray type="UInt8" Name="vtkGhostType" format="appended" RangeMin="
0"                   RangeMax="1"                    offset="155584832"
   />
[mkandes@login02 psi-1134]$
```

Simulation data storage requirements

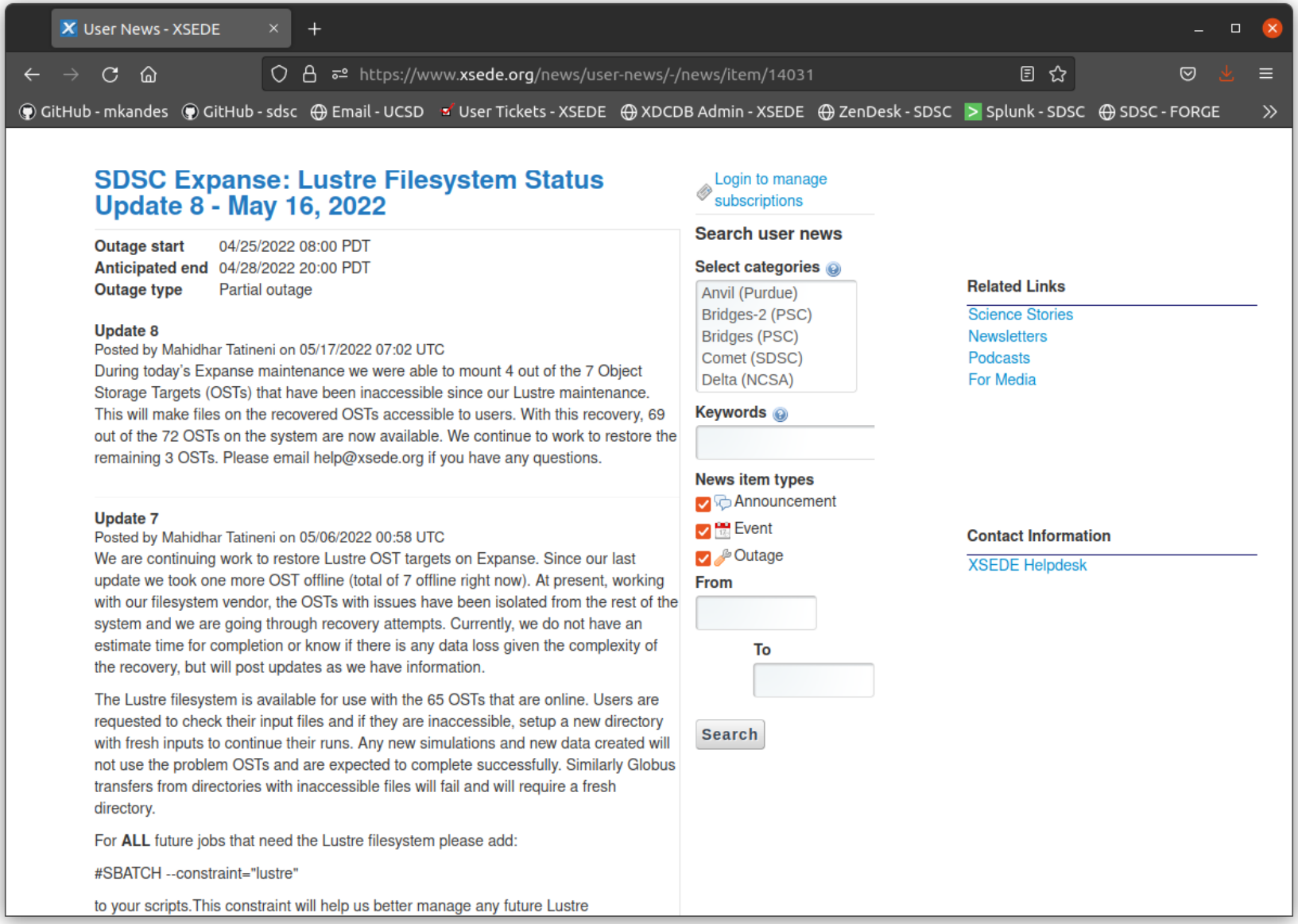


Expanse Lustre Filesystem - /expanse/lustre/projects

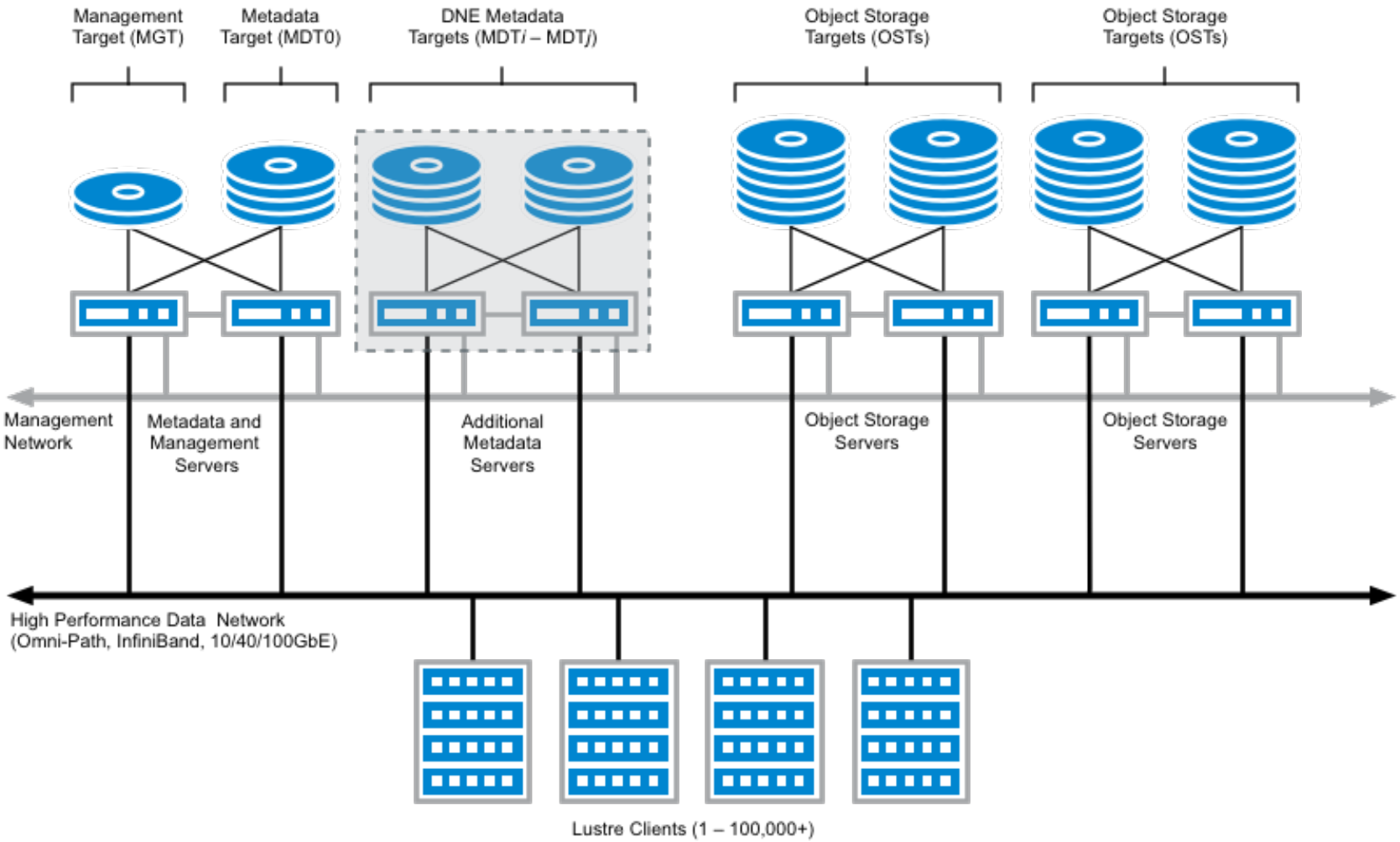


- Usage: Parallel (MPI) I/O, store large input/output files
- Quota: 500 GB per project/group (default)
- Backup: None
- Purge: 90 days after allocation expiration

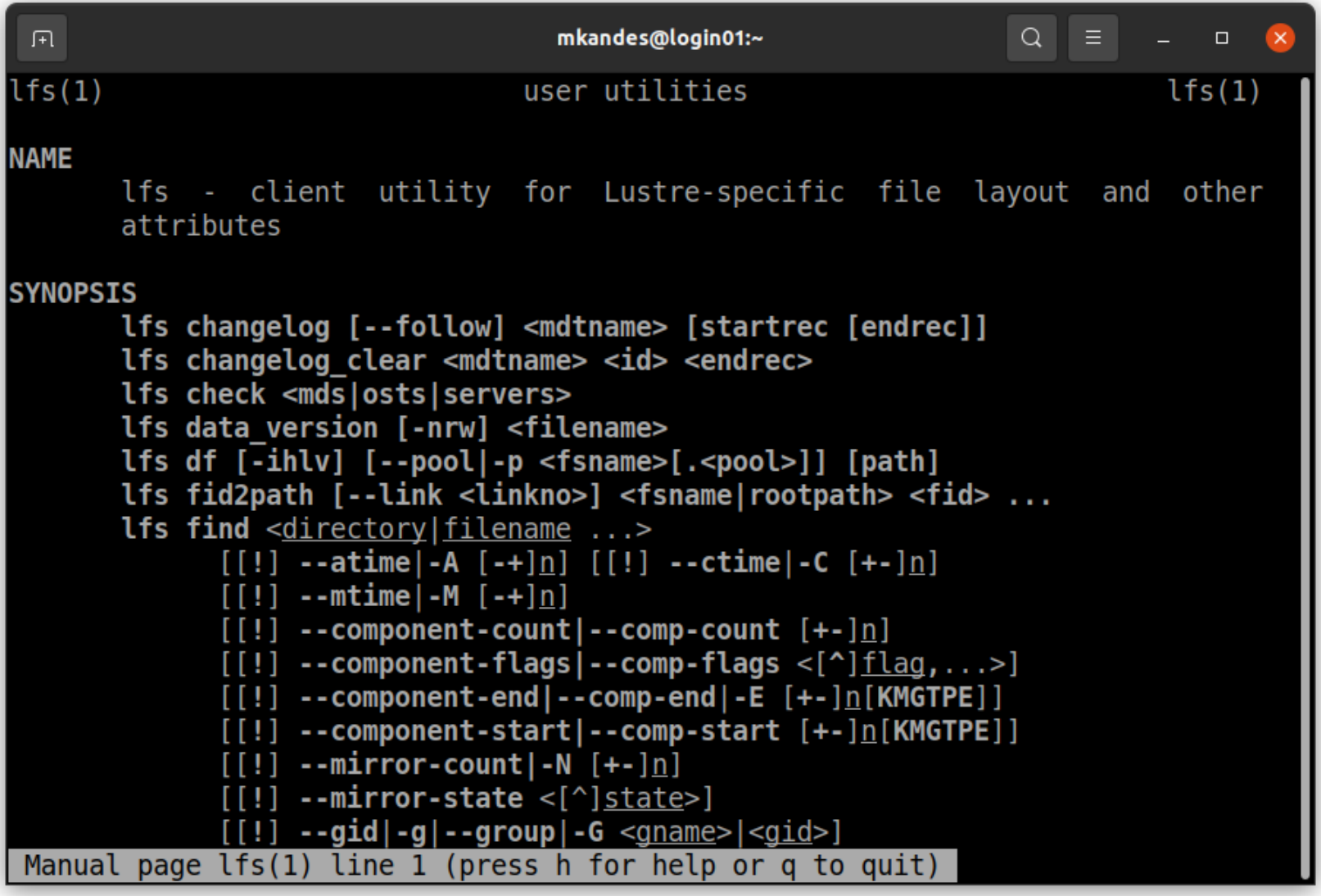
Murphy's Law - Lustre Edition



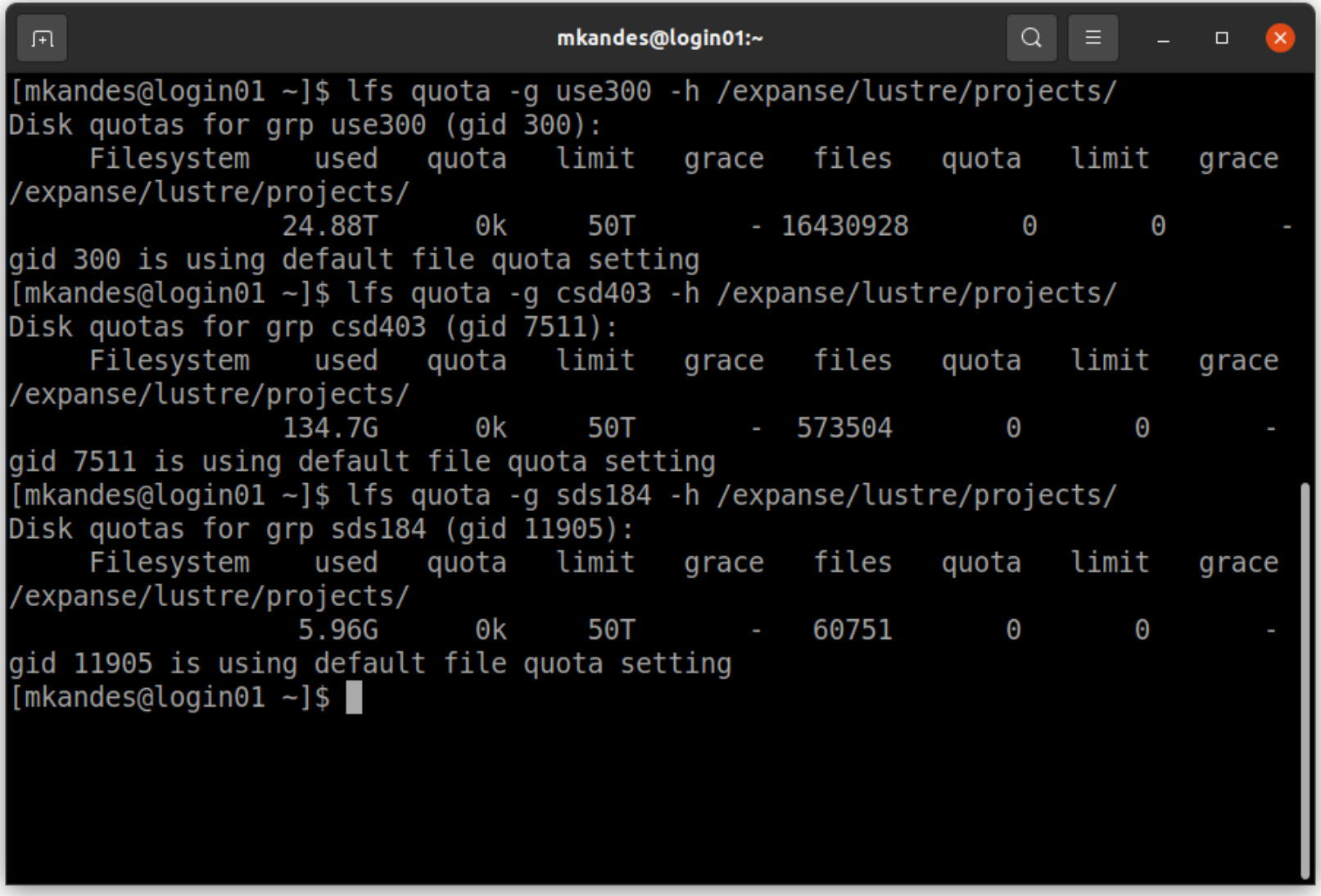
Lustre Filesystem Architecture



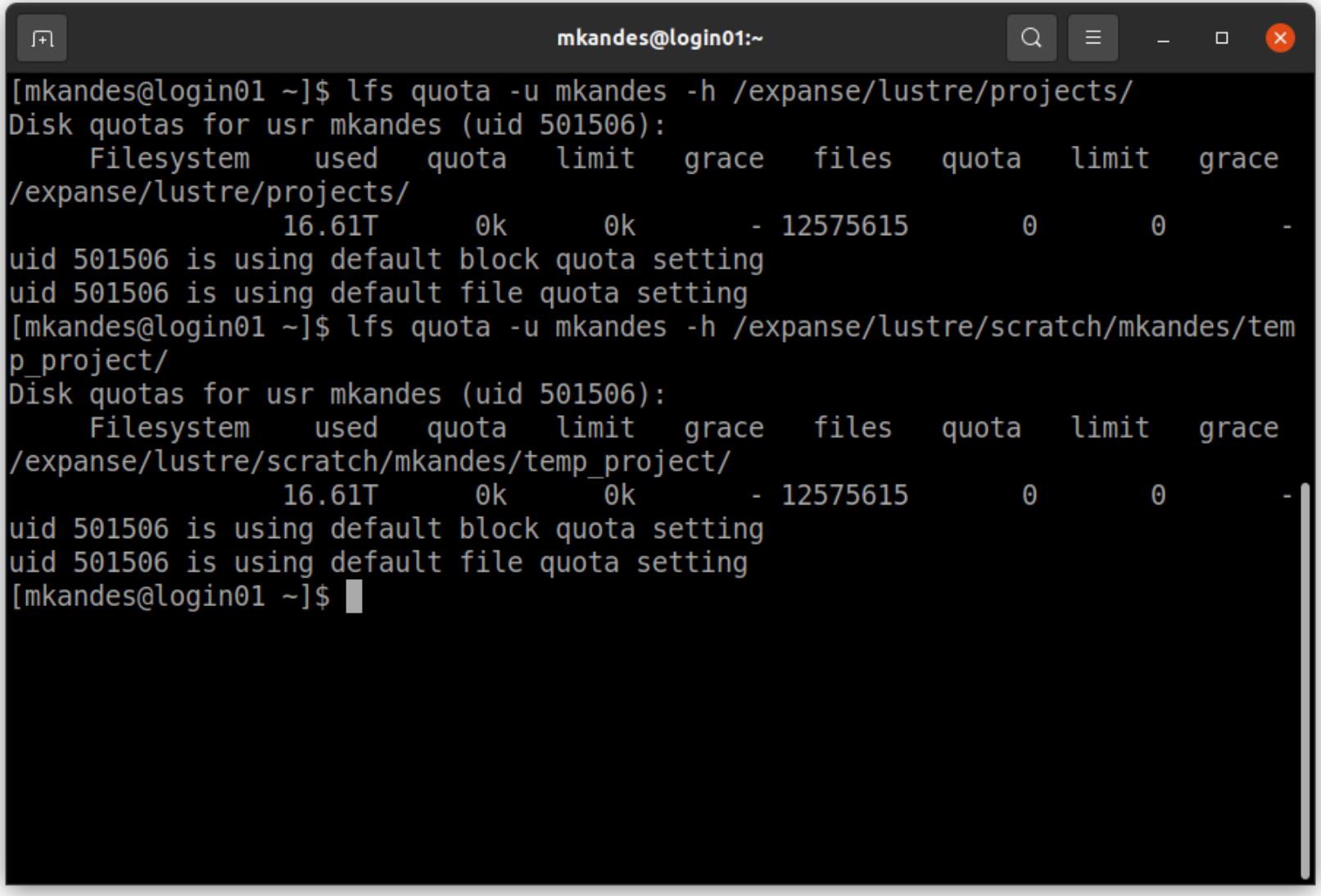
lfs


mkandes@login01:~
lfs(1) user utilities lfs(1)
NAME
lfs - client utility for Lustre-specific file layout and other attributes
SYNOPSIS
lfs changelog [--follow] <mdtname> [startrec [endrec]]
lfs changelog_clear <mdtname> <id> <endrec>
lfs check <mds|osts|servers>
lfs data_version [-nrw] <filename>
lfs df [-ihlv] [--pool|-p <fsname>[.<pool>]] [path]
lfs fid2path [--link <linkno>] <fsname|rootpath> <fid> ...
lfs find <directory|filename ...>
[[!] --atime|-A [-+]n] [[!] --ctime|-C [+-]n]
[[!] --mtime|-M [-+]n]
[[!] --component-count|--comp-count [+-]n]
[[!] --component-flags|--comp-flags <[^]flag,...>]
[[!] --component-end|--comp-end|-E [+-]n[KMGTPE]]
[[!] --component-start|--comp-start [+-]n[KMGTPE]]
[[!] --mirror-count|-N [+-]n]
[[!] --mirror-state <[^]state>]
[[!] --gid|-g|--group|-G <gname>|<gid>]
Manual page lfs(1) line 1 (press h for help or q to quit)

lfs quota (by group)



lfs quota (by user)



Expanse Lustre Filesystem - `/expanse/lustre/scratch`

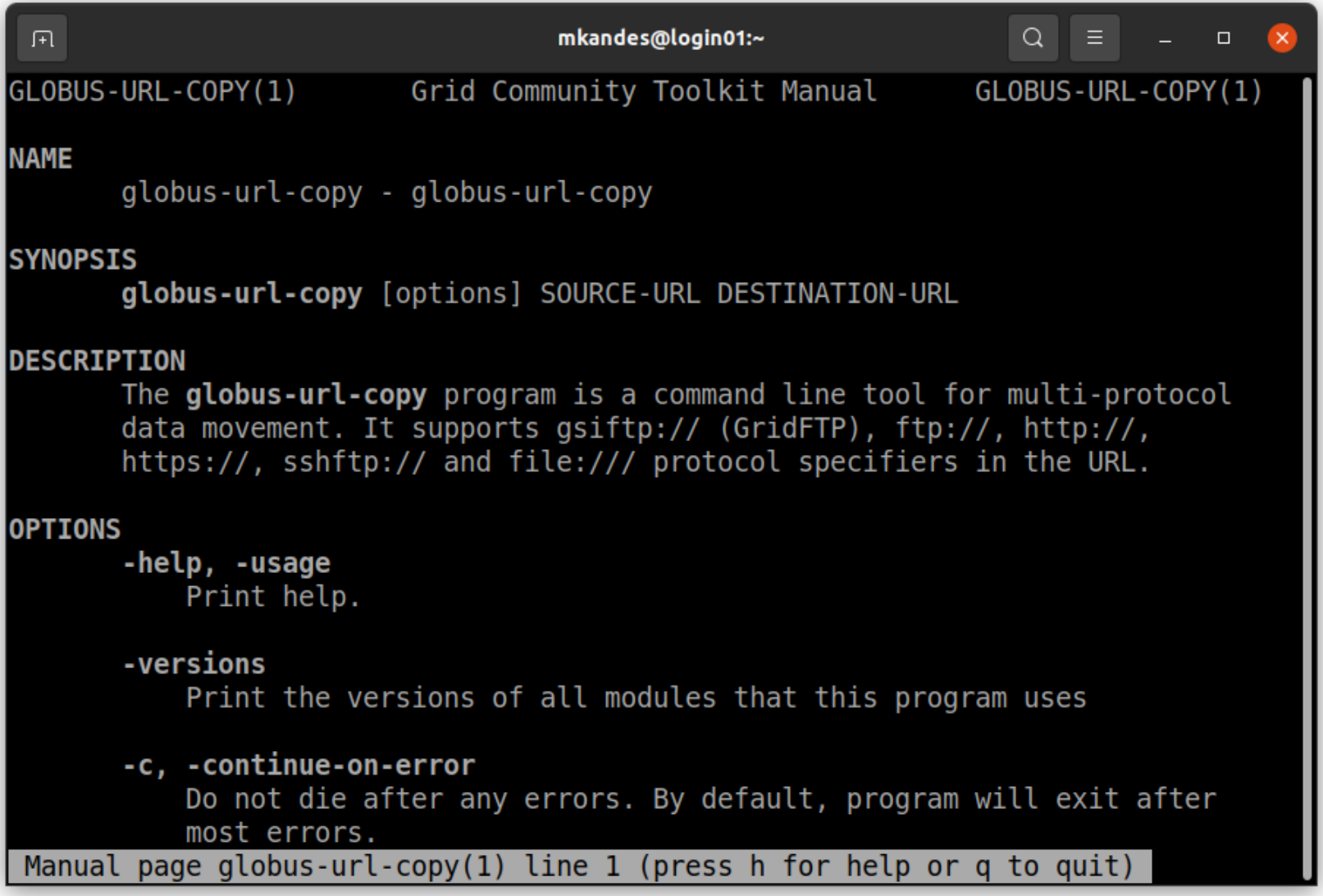
- Usage: Parallel (MPI) I/O, storing large (temporary) files
- Quota: Up to 10 TB per user (default)
- Backup: None
- Purge: 90 days after file creation

Data transfer script - GridFTP

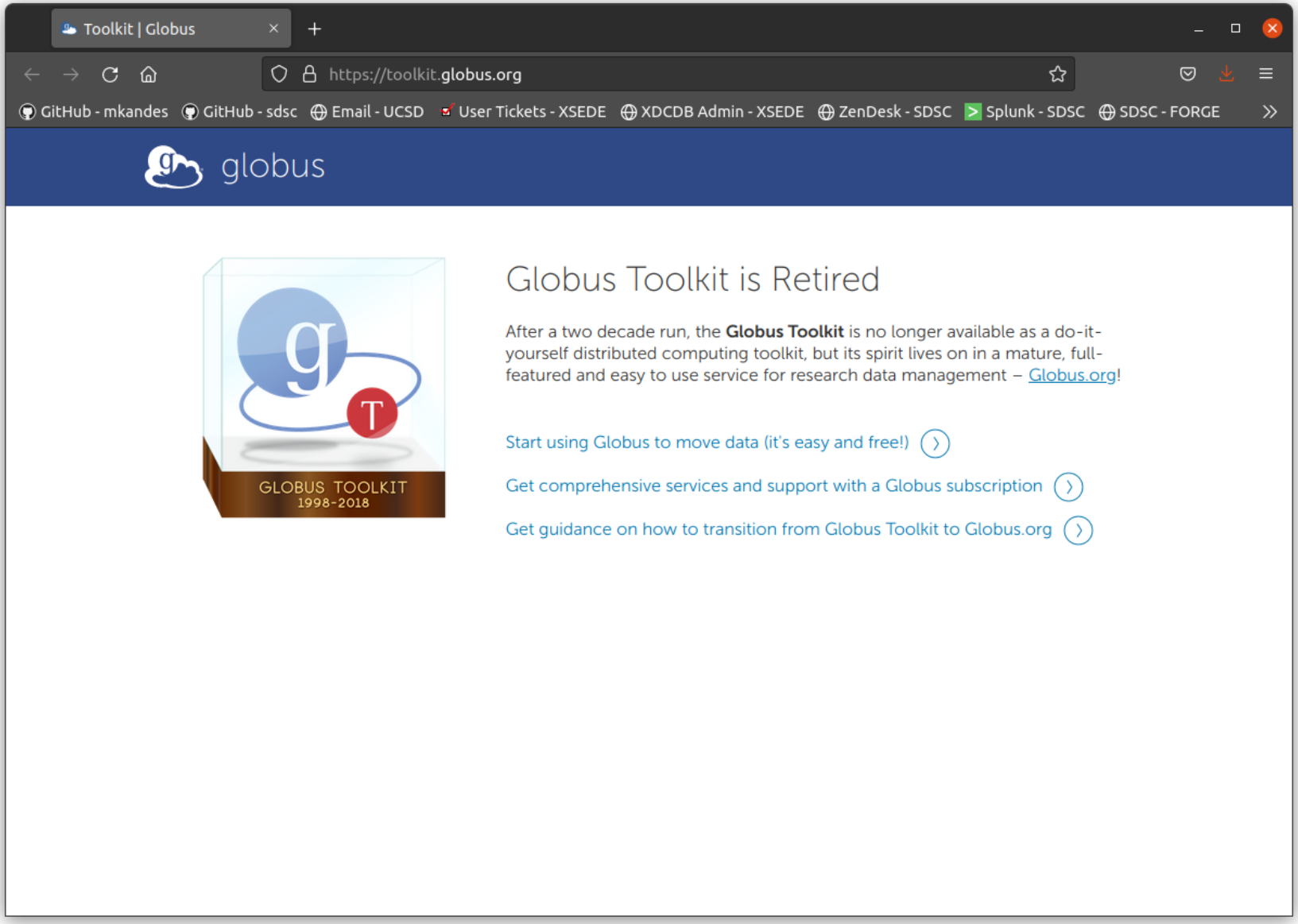
mkandes@hardtack: ~/Dropbox/work/ucsd/sdsc/presentations/2022/sdsc/ciml/data...

```
#!/usr/bin/env bash

declare -xr SOURCE_ENDPOINT='oasis-dm.sdsc.xsede.org:2811'
declare -xr SOURCE_DIR='/oasis/projects/nsf/sdu118/mkandes/gpse-v0.5.9-comet-com
pton-20170828-001'

declare -xr DESTINATION_ENDPOINT='gridftp.stampede2.tacc.xsede.org:2811'
declare -xr DESTINATION_DIR='/work/03216/mckandes/maverick/gpse-v0.5.9-comet-com
pton-20170828-001'

module load globus

for timestep in {1000..1024}; do
  globus-url-copy -vb -stripe -tcp-bs 8m -p 4 "gsiftp://${SOURCE_ENDPOINT}/${SOU
RCE_DIR}/psi-${timestep}.zip" "gsiftp://${DESTINATION_ENDPOINT}/${DESTINATION_DI
R}/psi-${timestep}.zip"
done
                                                              1,1           Top
```

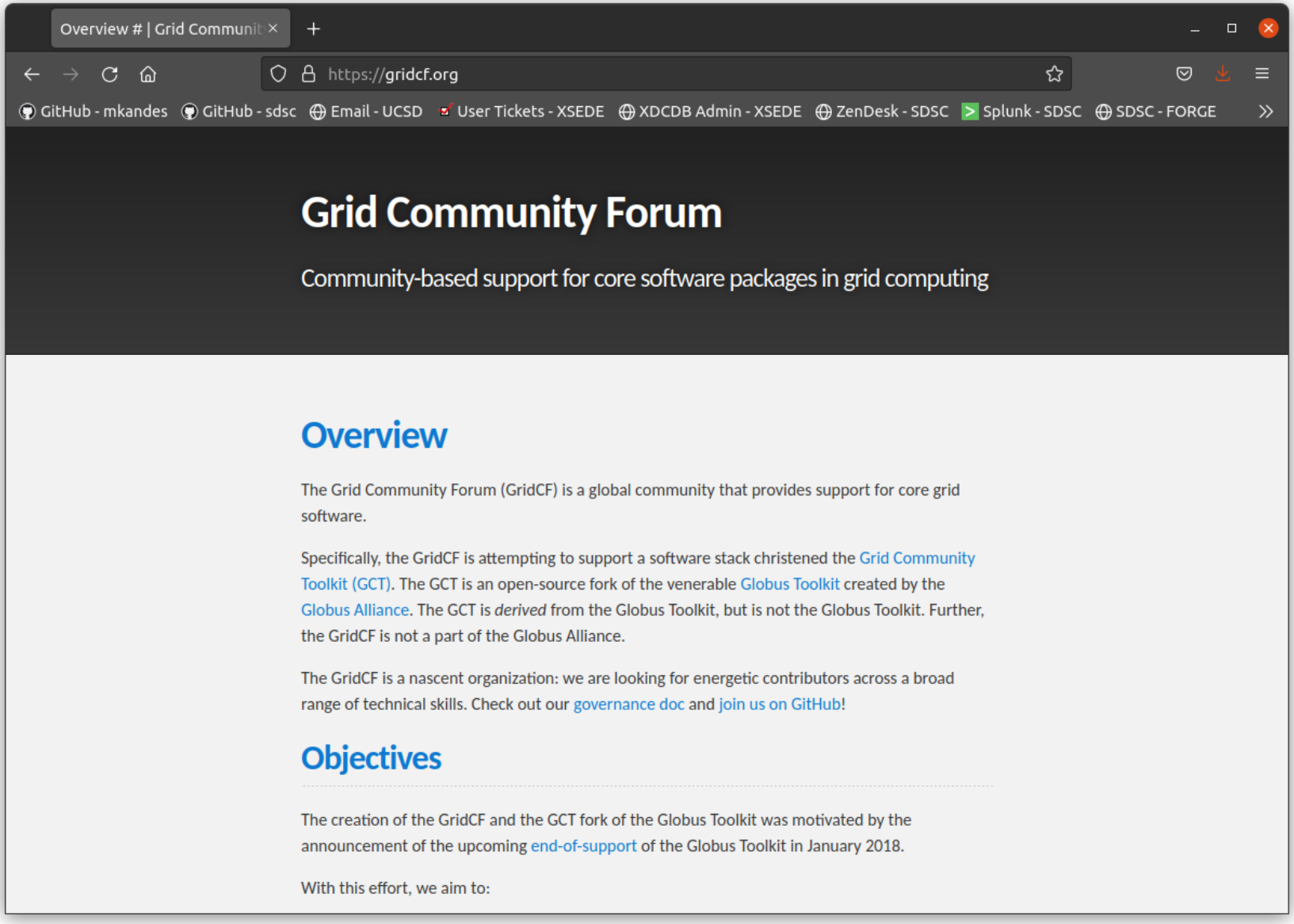
globus-url-copy



Globus Toolkit



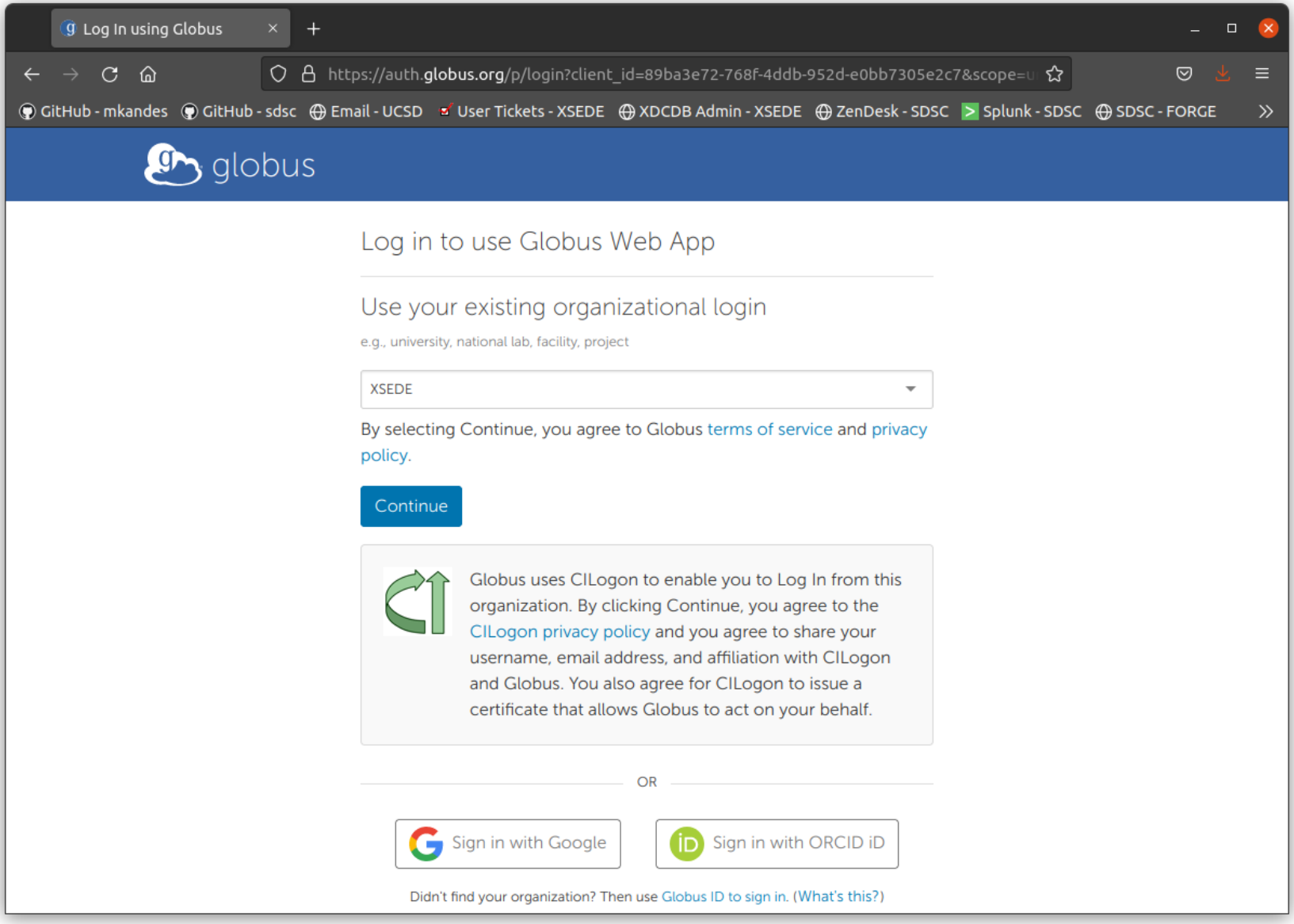
Grid Community Toolkit



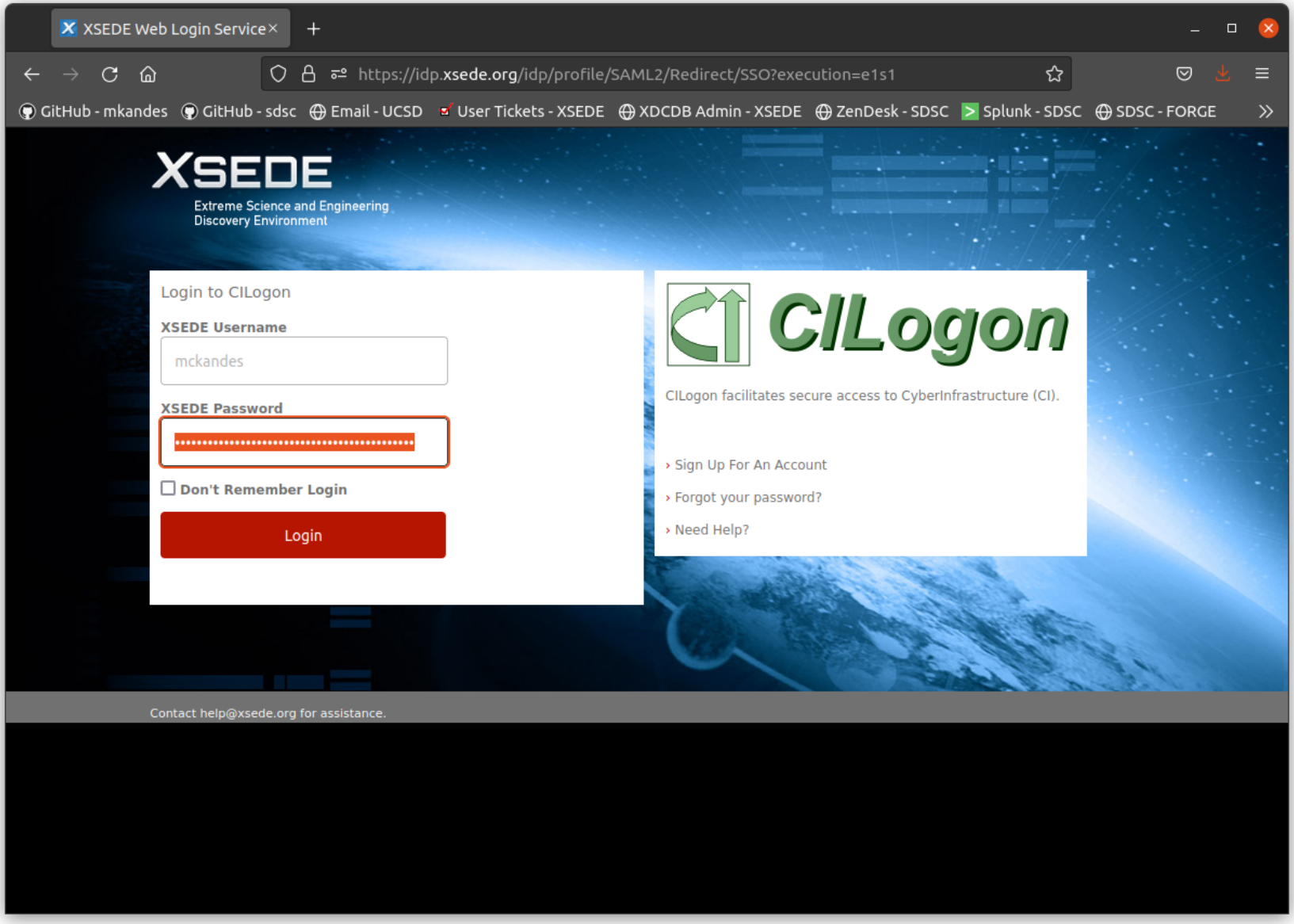
Globus


Research data managem
https://www.globus.org
GitHub - mkandes
GitHub - sdsc
Email - UCSD
User Tickets - XSEDE
XDCDB Admin - XSEDE
ZenDesk - SDSC
Splunk - SDSC
SDSC - FORGE
globus
a uchicago non-profit service
I Want To...
Pricing
Resources
Support
About
Log In
Introducing Flows
Our new service that accelerates discovery
Auth
Transfer
Web Form
Describe
Identifier
Share
Search
GET CREDENTIALS
TRANSFER DATA
USER INPUT
GET METADATA
MINT DOI
SET POLICY
CREATE INDEX
LEARN MORE ABOUT FLOWS
Research data management simplified.
TRANSFER
SHARE
BUILD

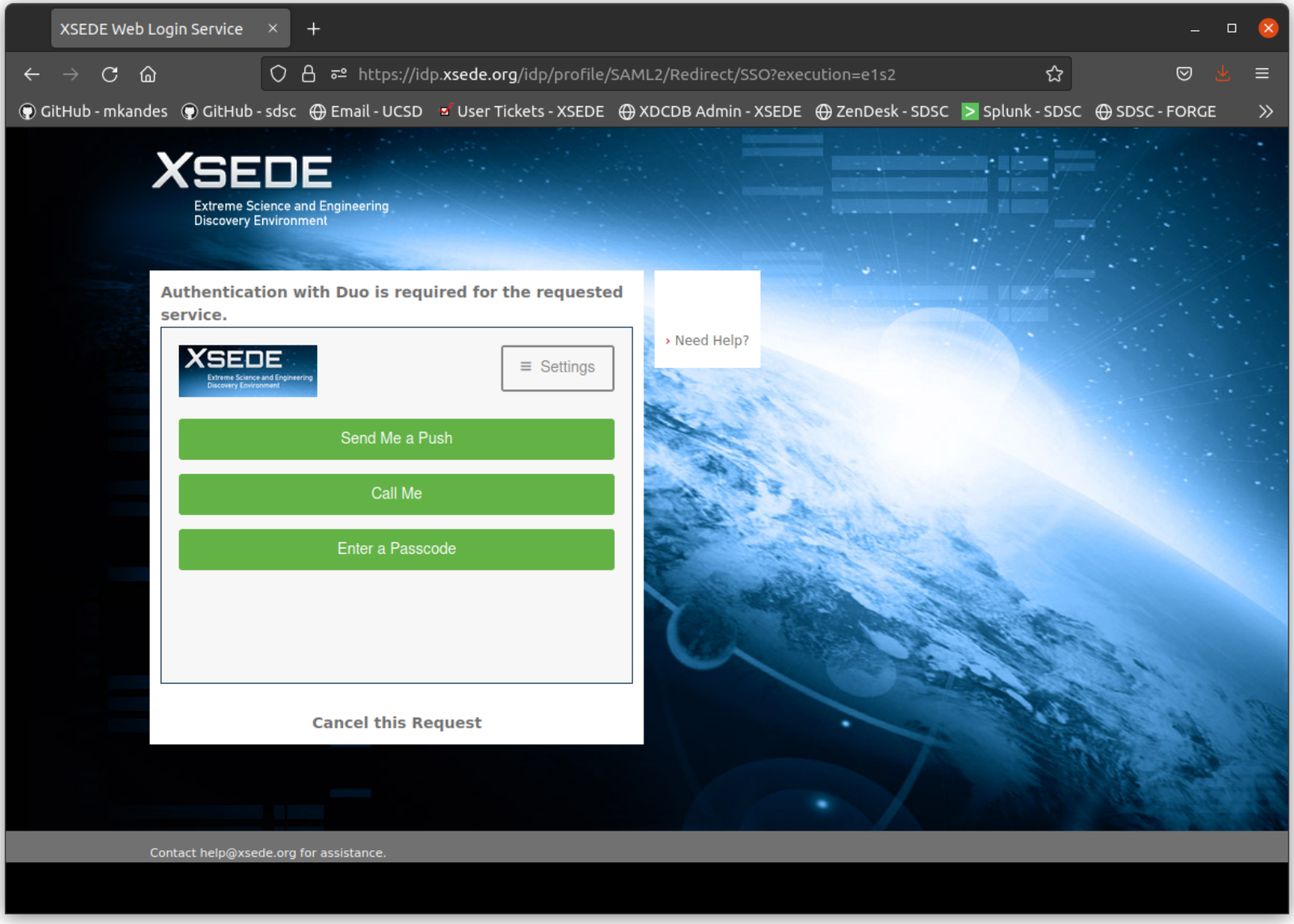
Globus Log In



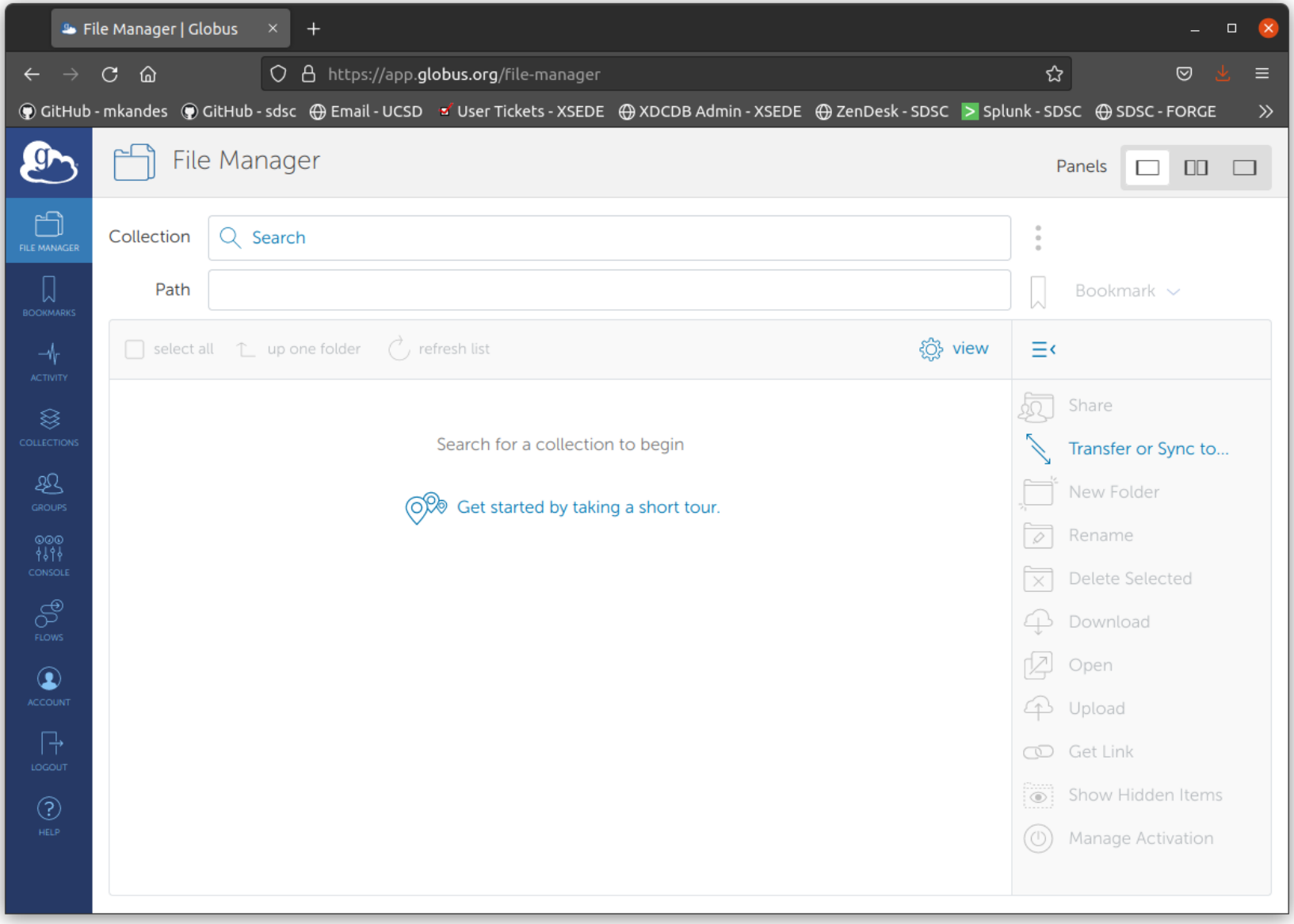
XSEDE CILogon



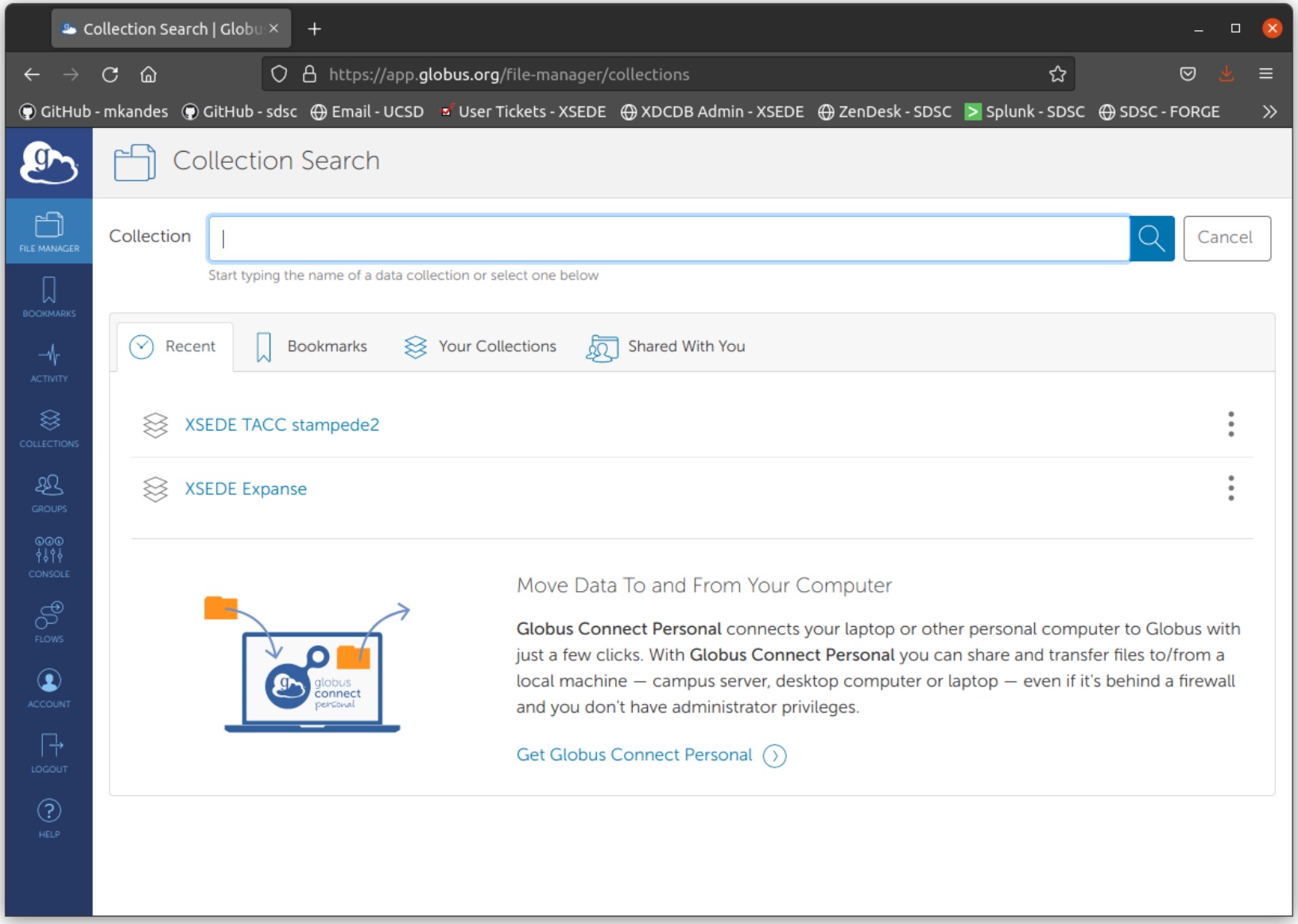
XSEDE Duo Two-Factor Authentication



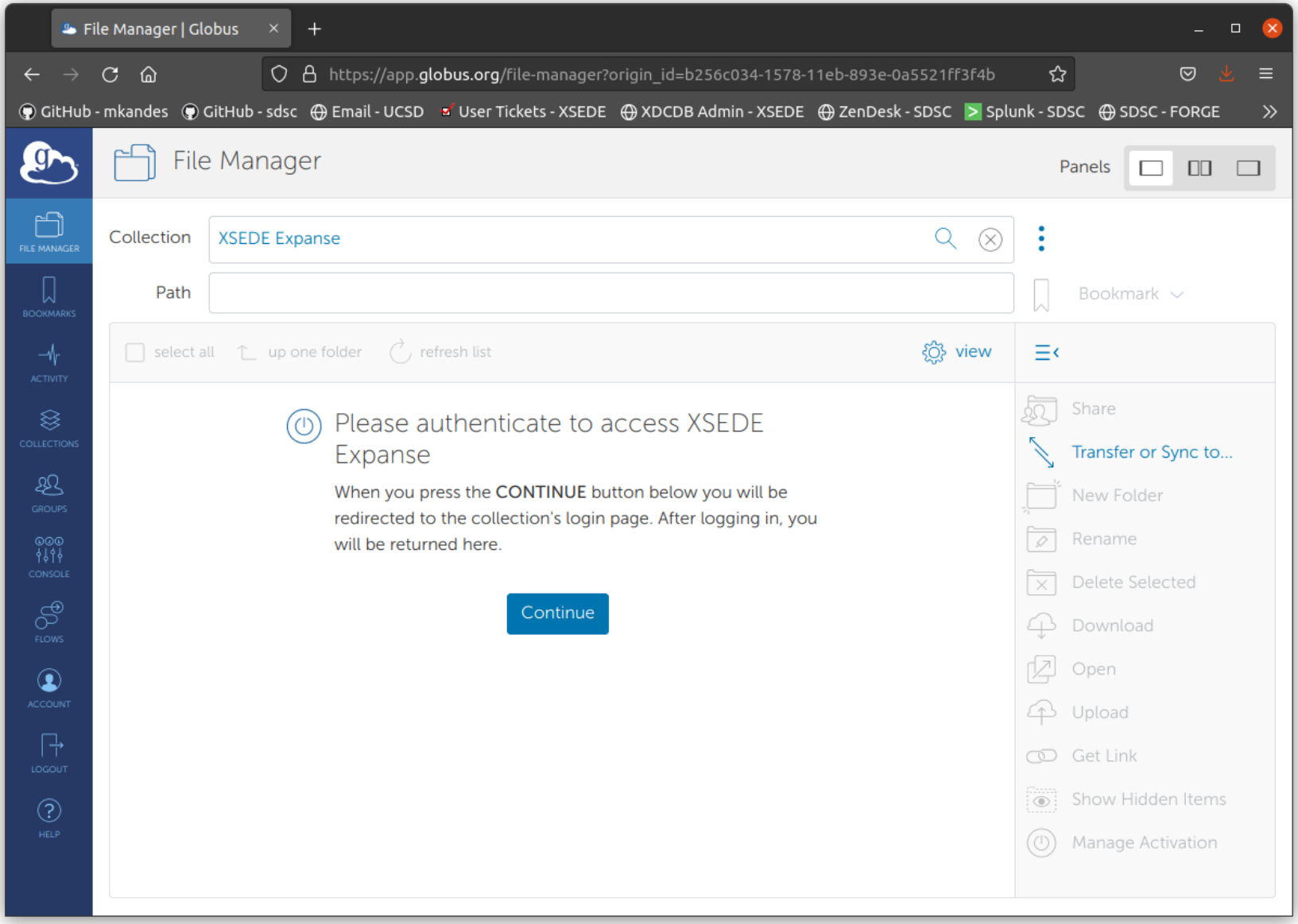
Globus File Manager



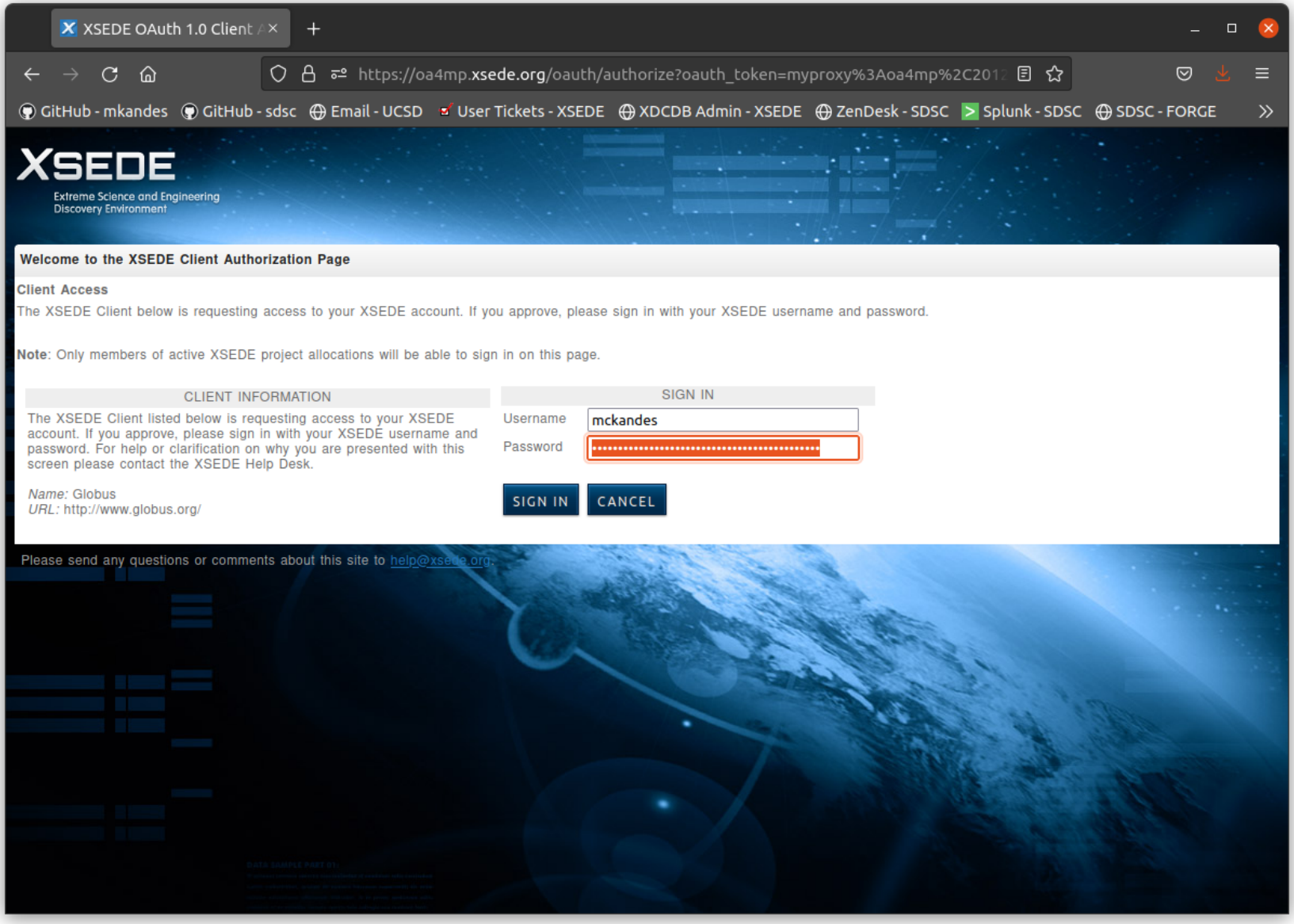
Globus Collection Search



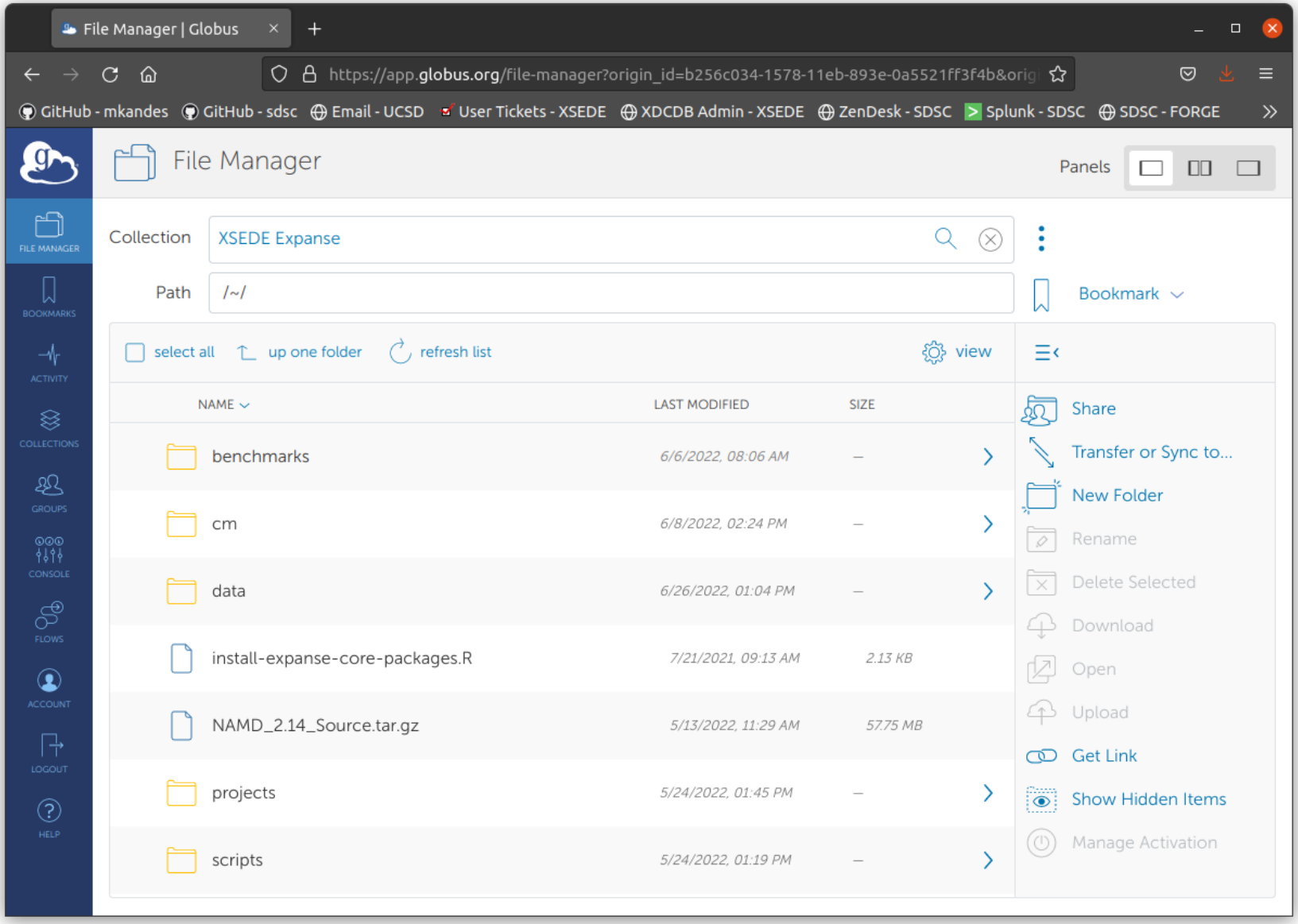
XSEDE Expanse Globus Collection / Endpoint



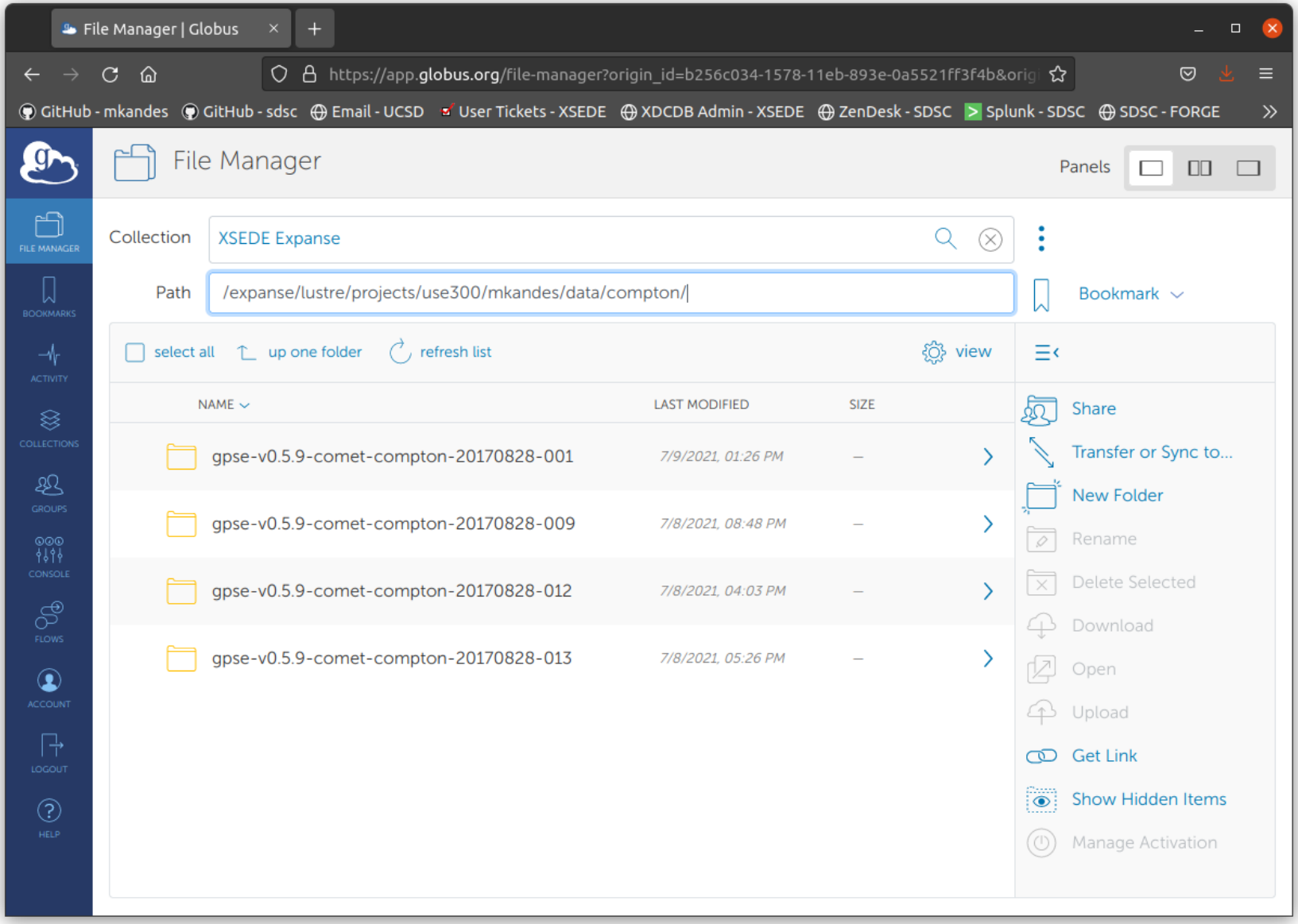
XSEDE Client Authorization



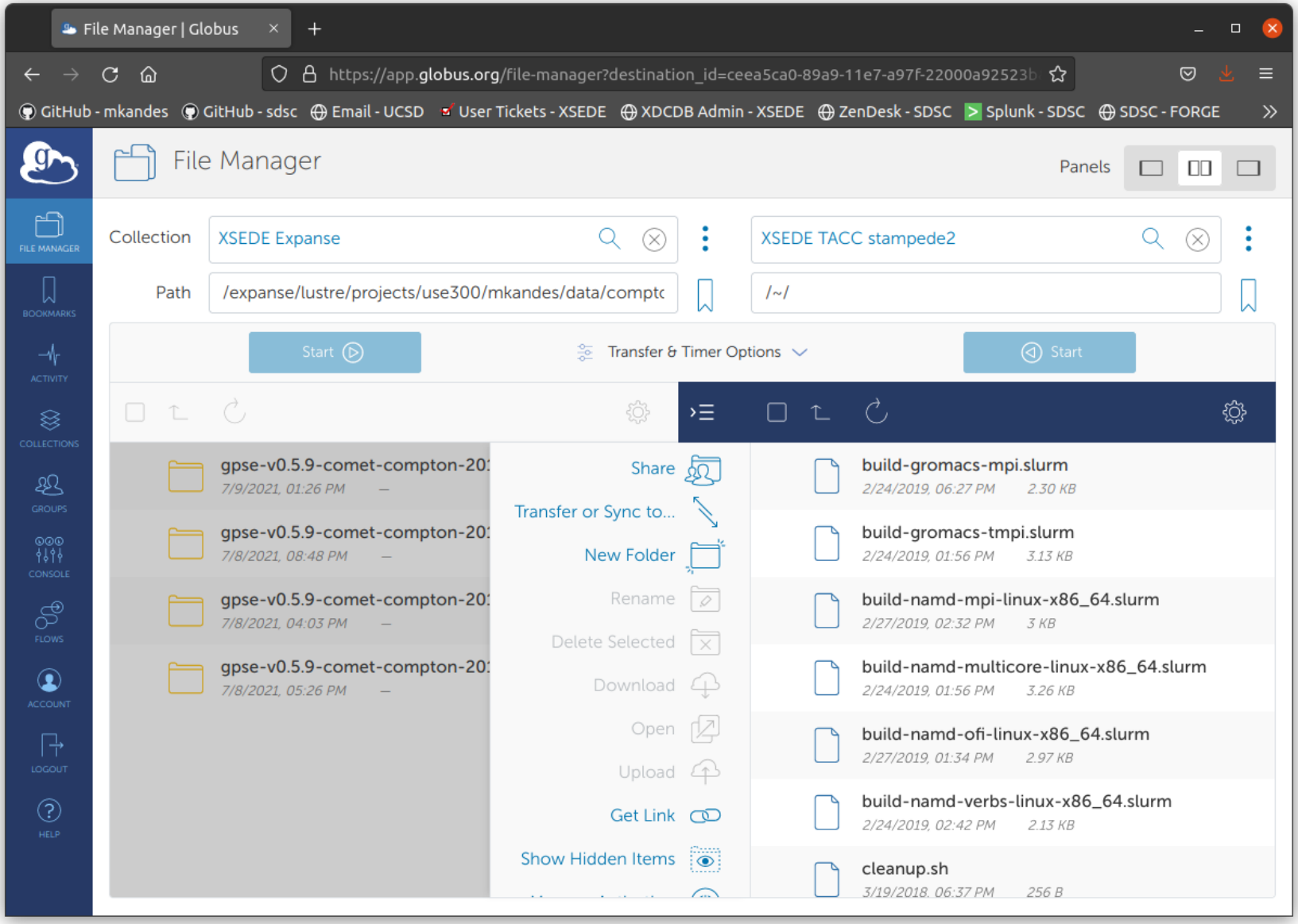
XSEDE Expanse - $HOME



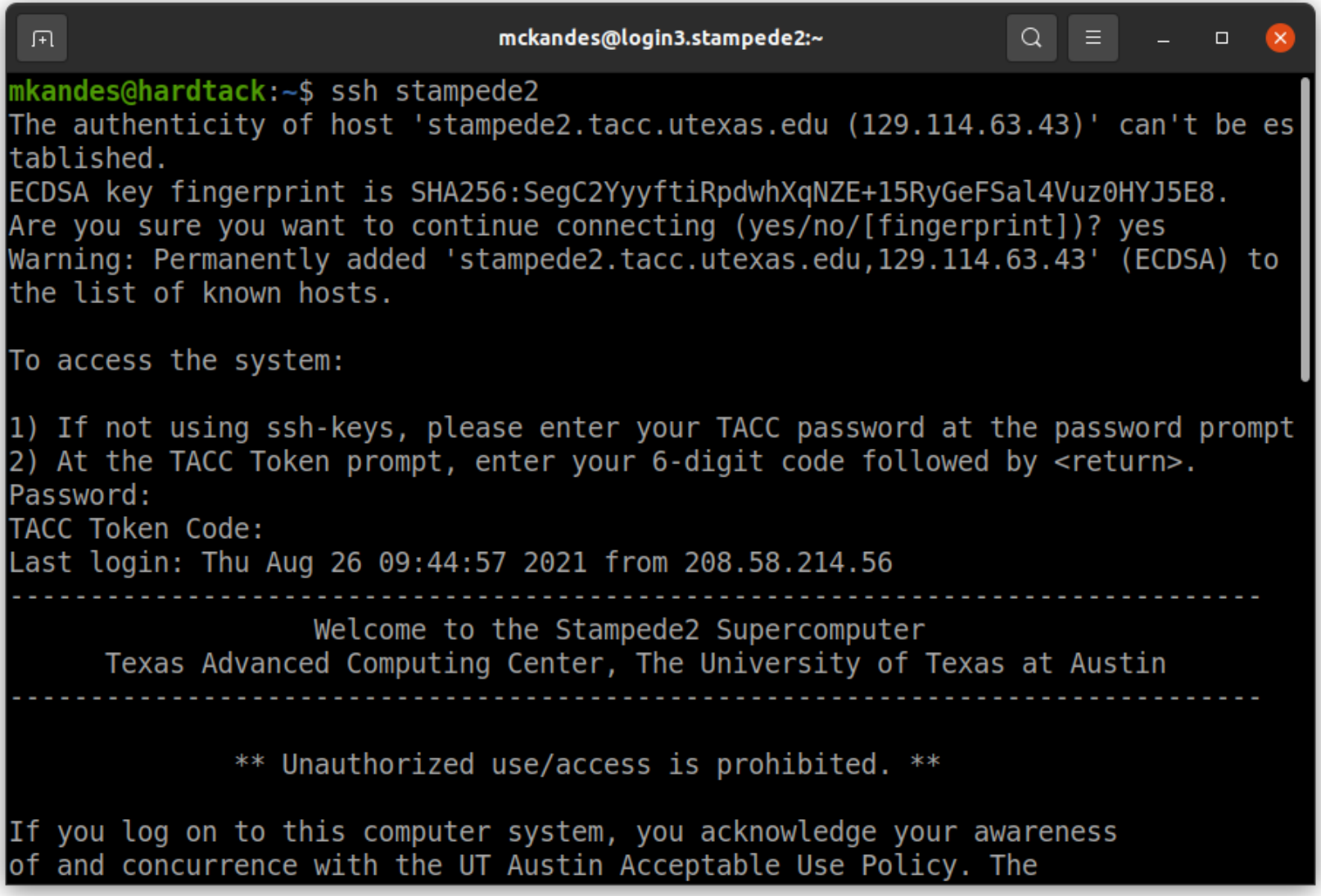
XSEDE Expanse - /expanse/lustre/projects



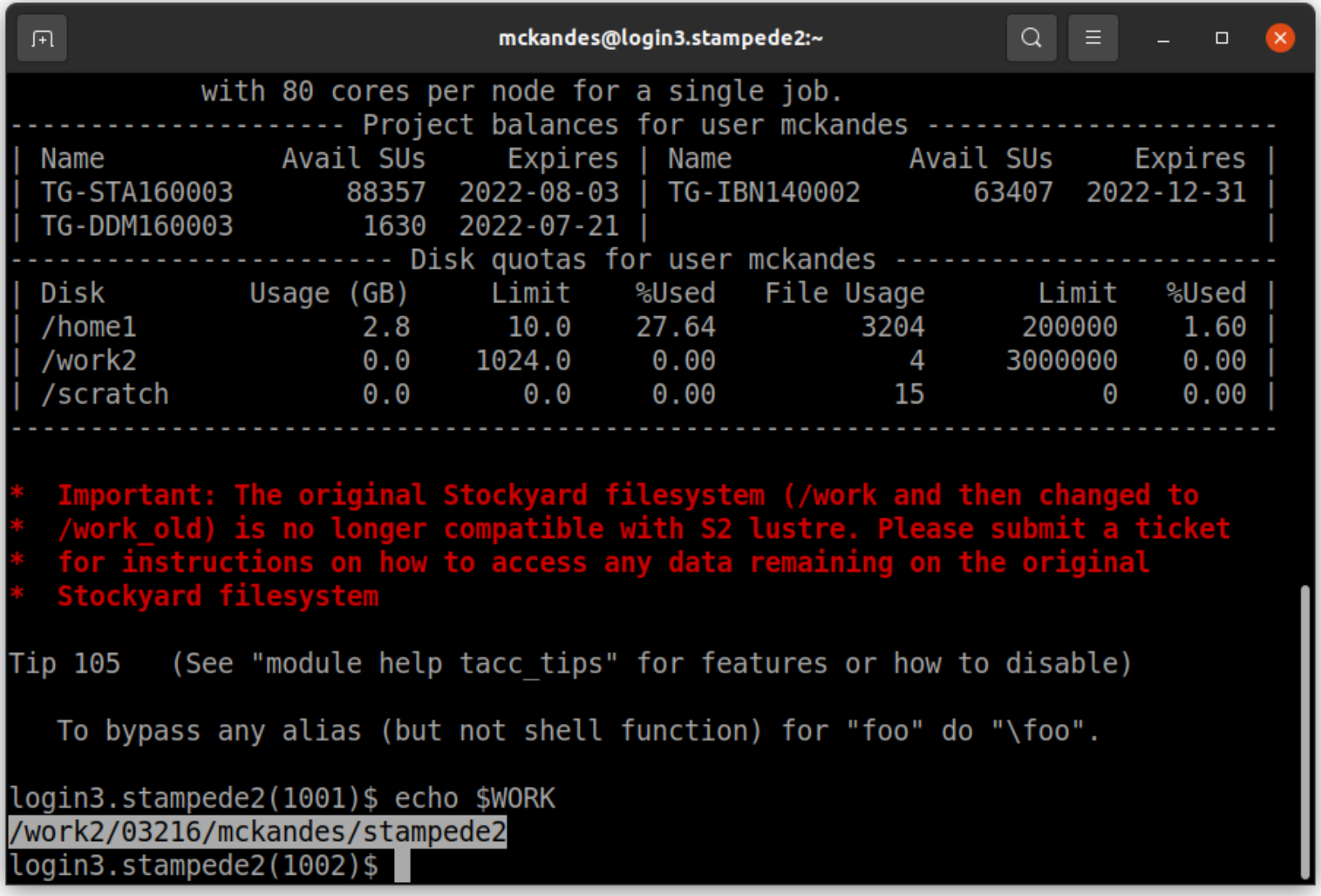
XSEDE Expanse — XSEDE Stampede2



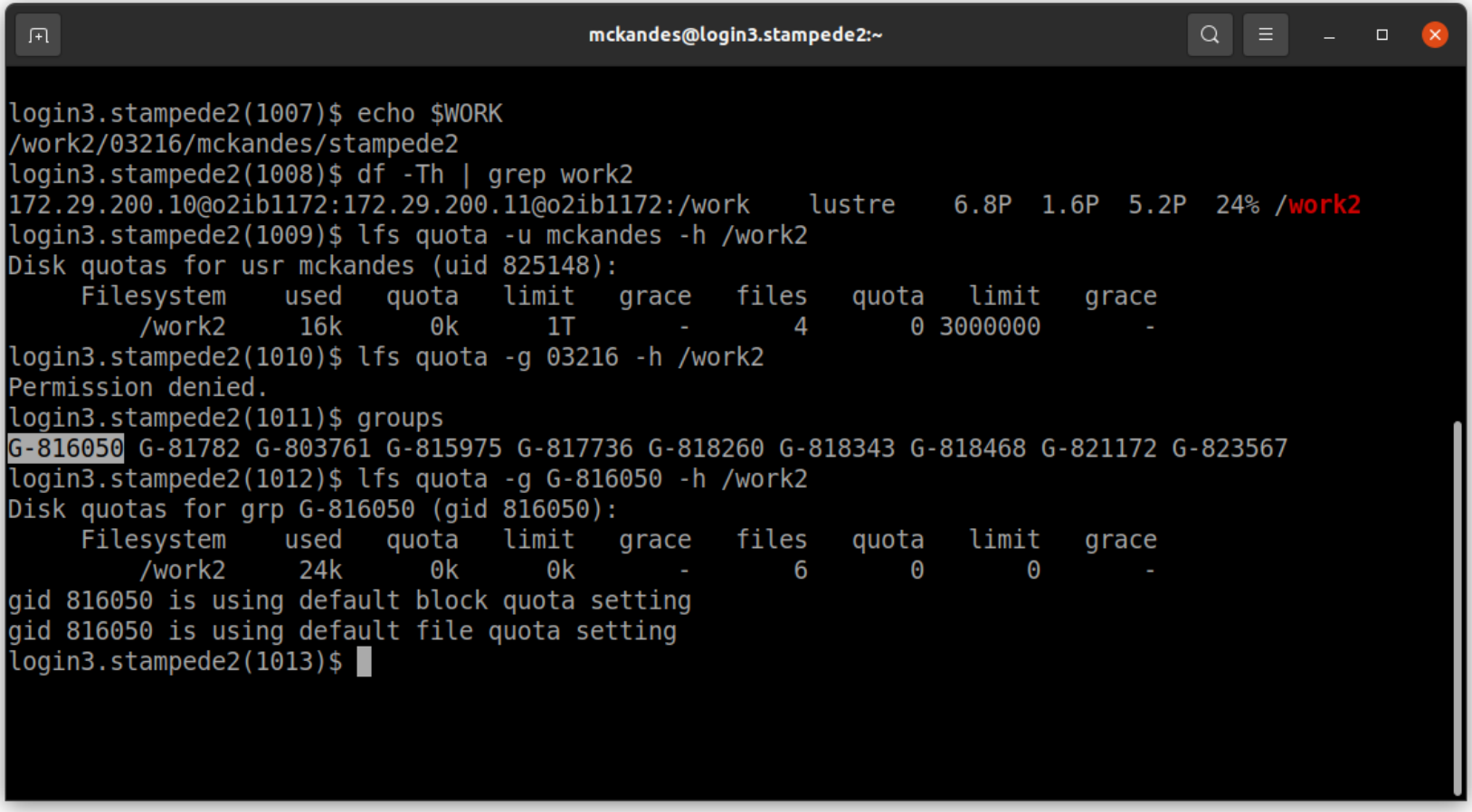
Stampede2 Login Nodes - 2FA



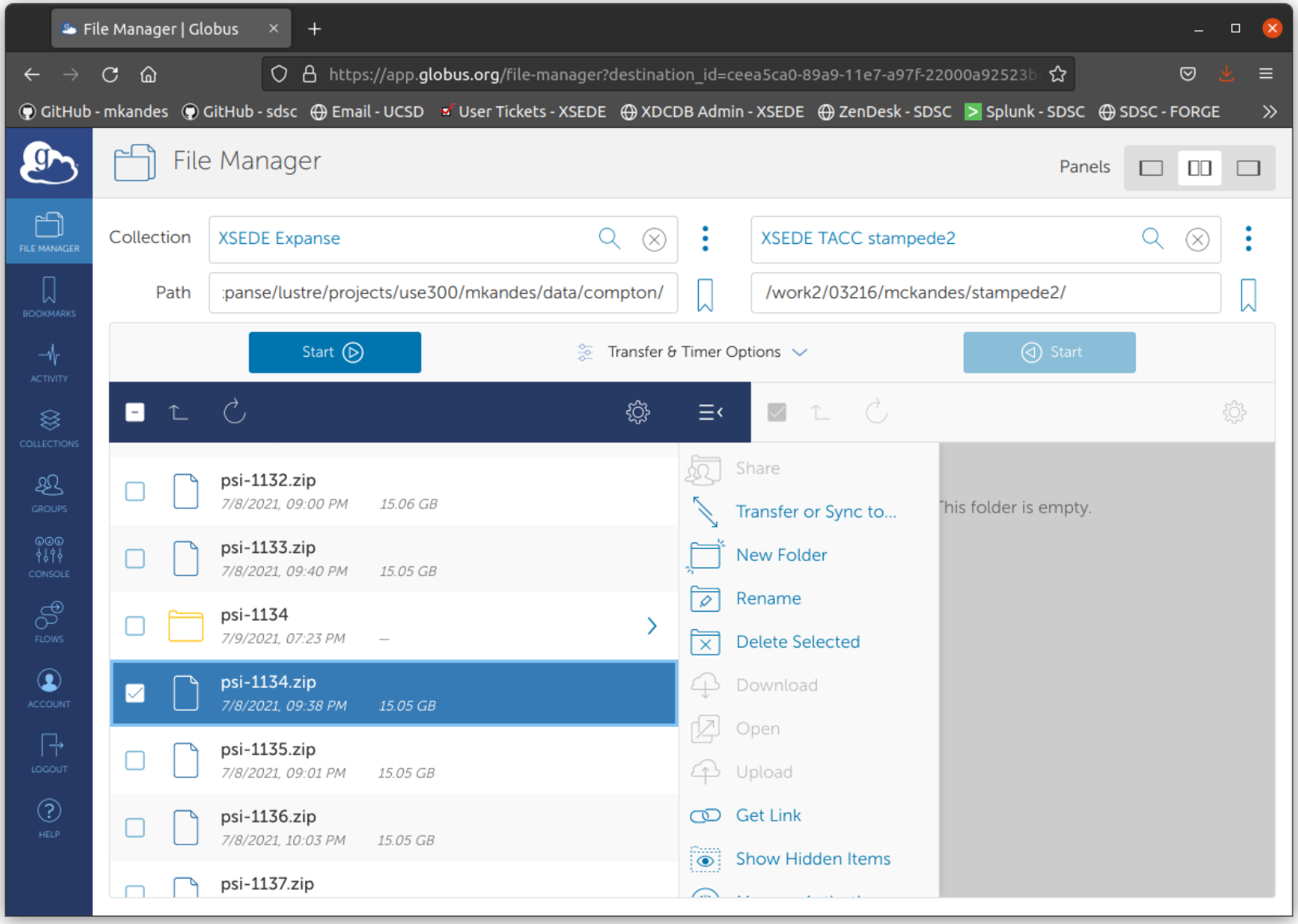
Stampede2 Login Nodes - Allocations



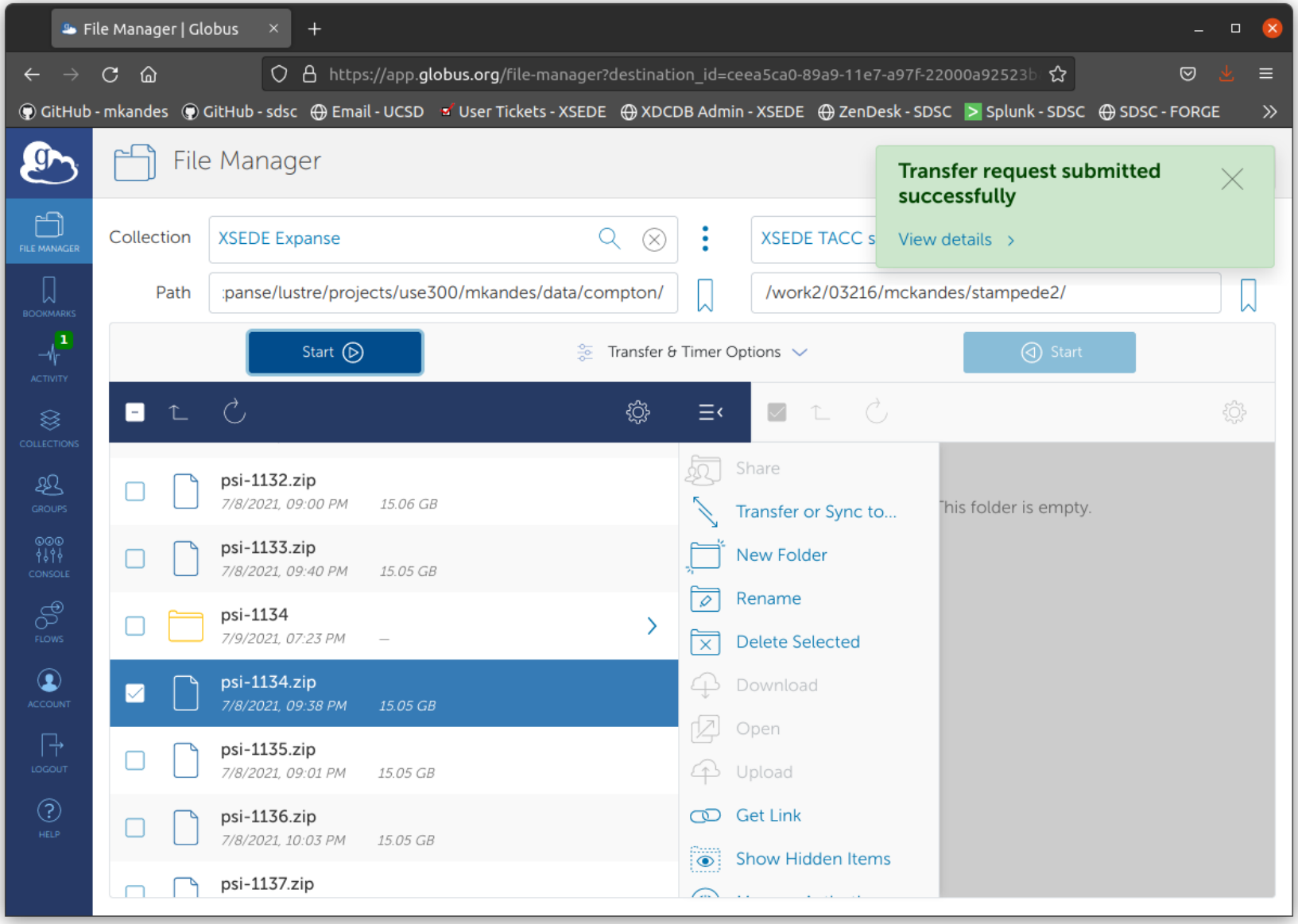
Stampede2 $WORK Lustre Filesystem



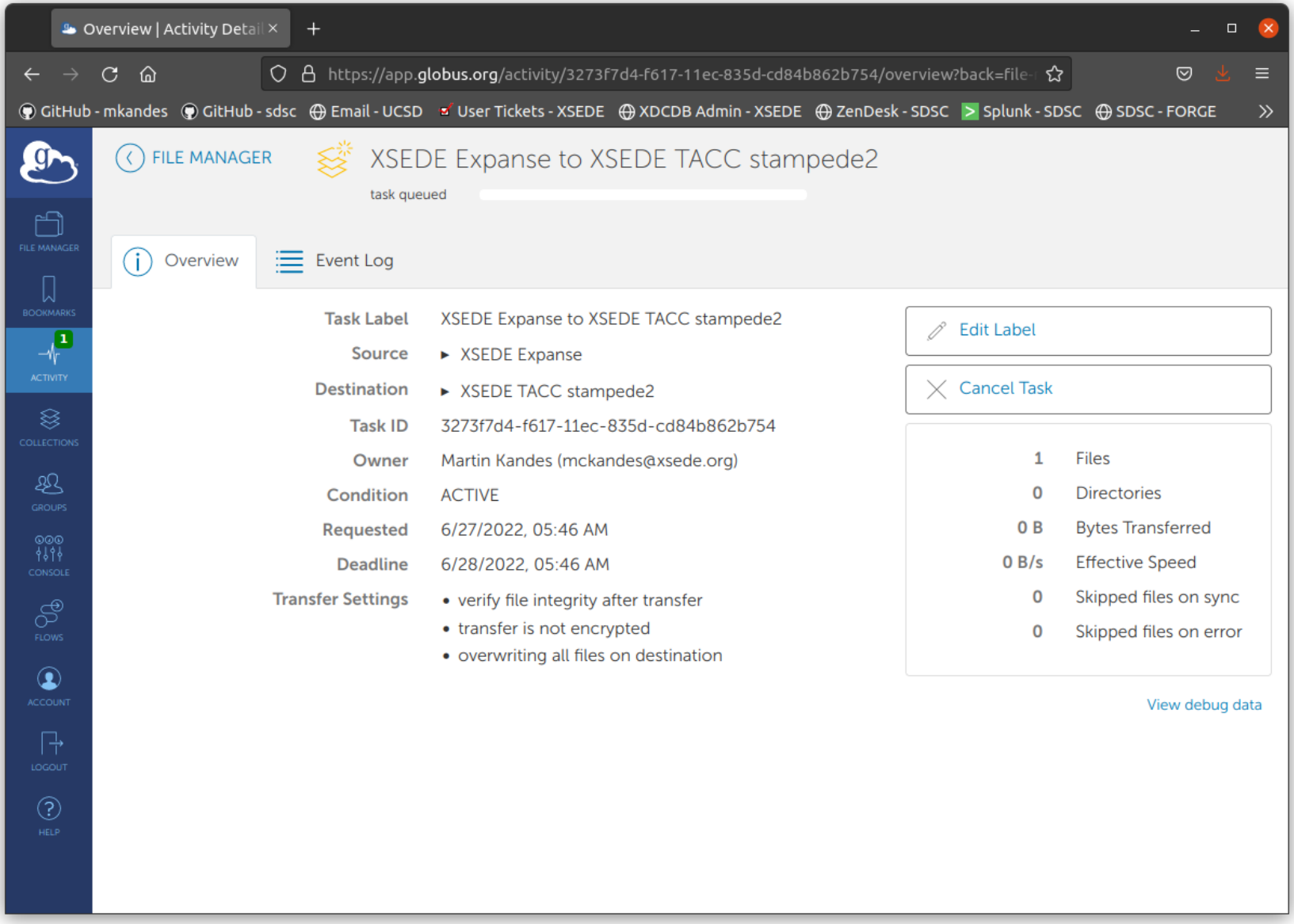
Globus File Transfer - Setup



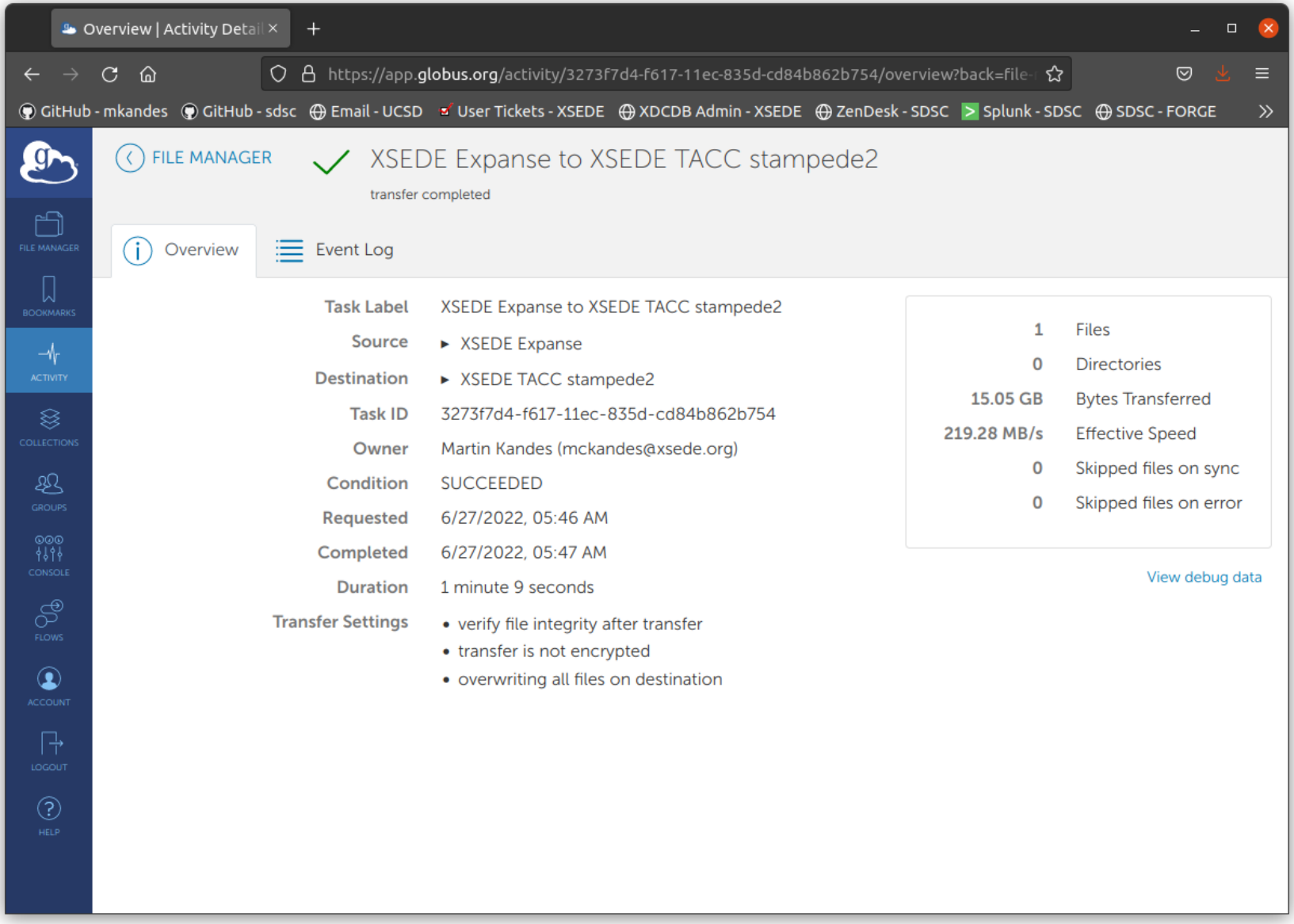
Globus File Transfer - Initiated



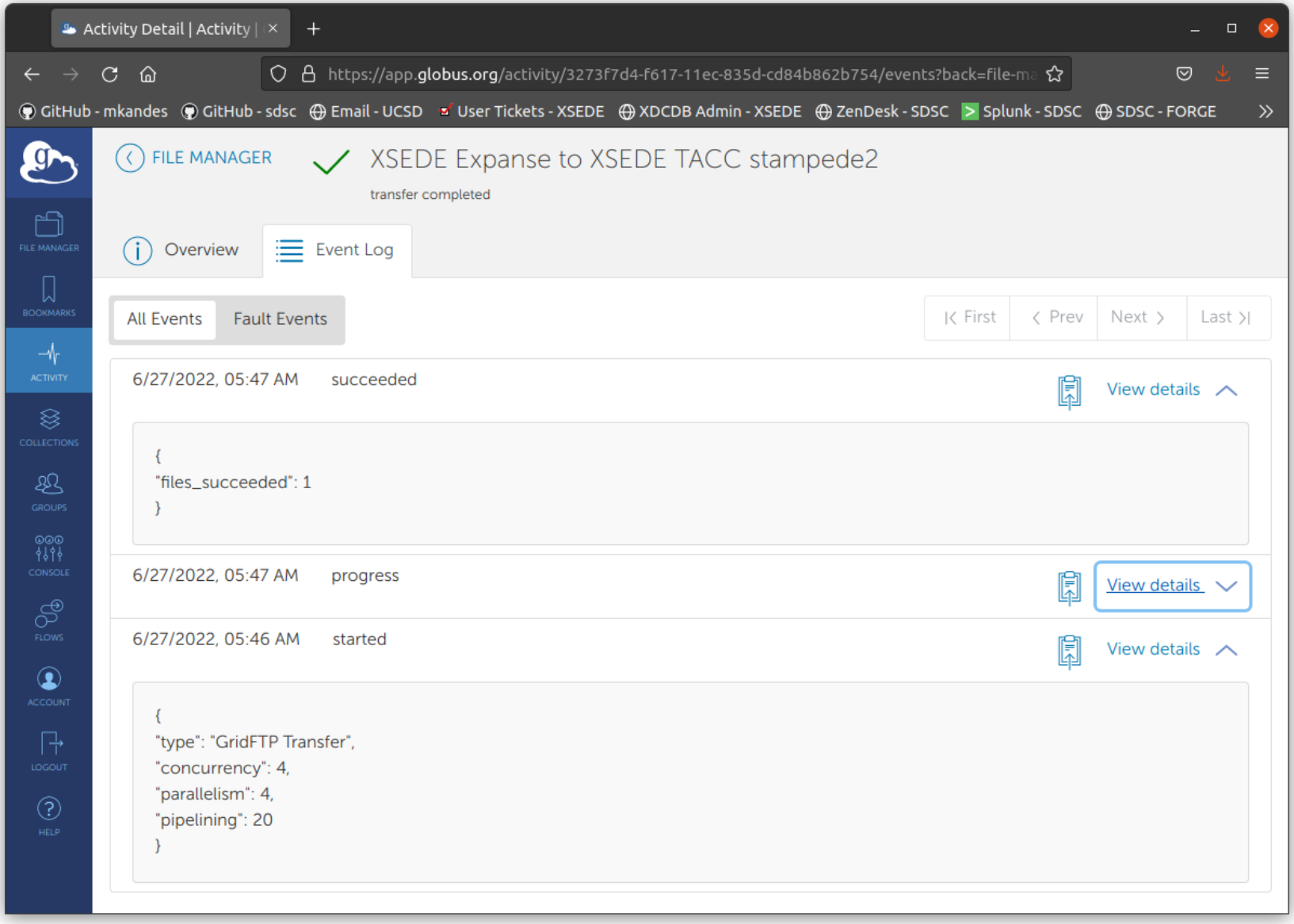
Globus File Transfer - Details



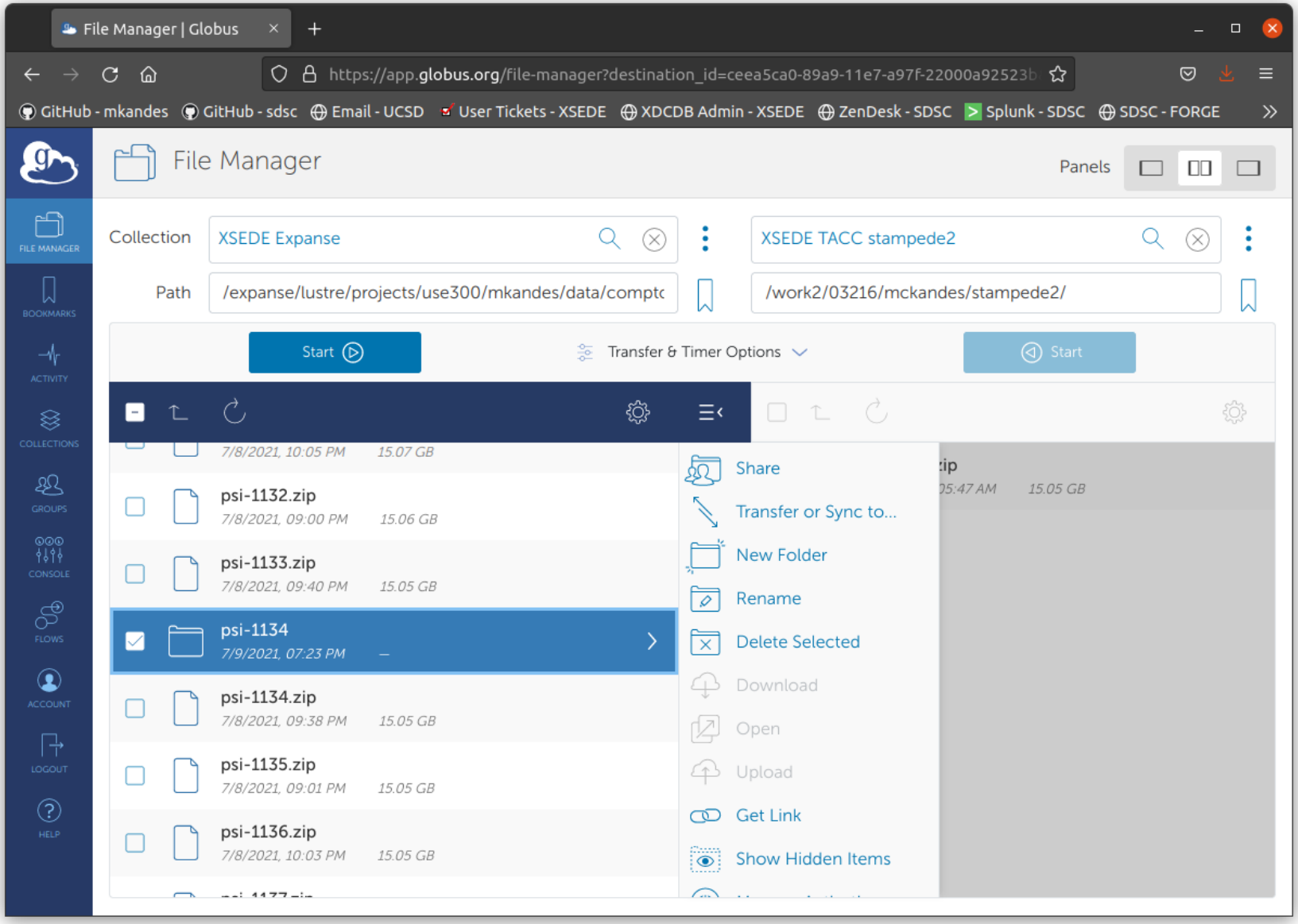
Globus File Transfer - Completed



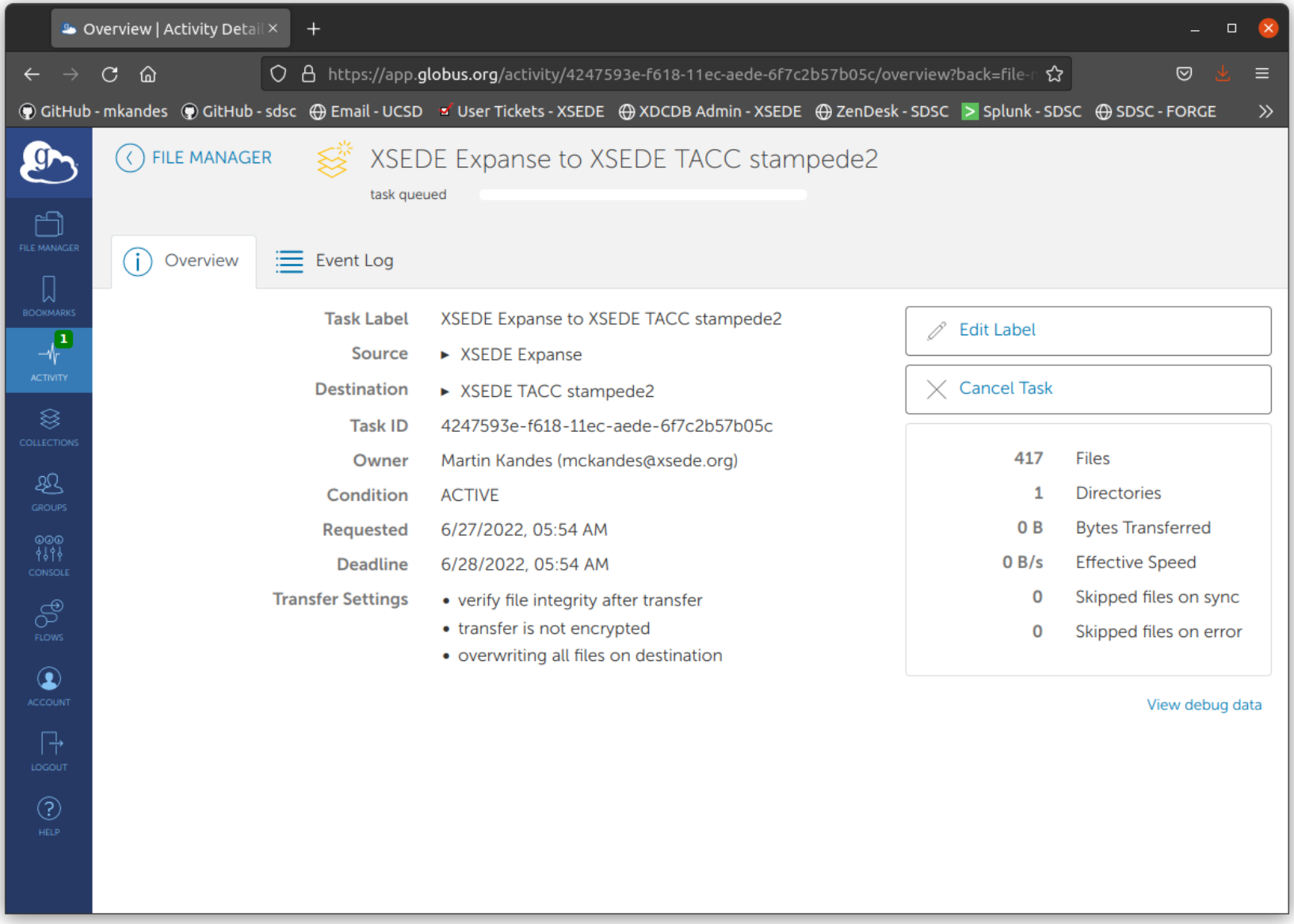
Globus File Transfer - Event Log



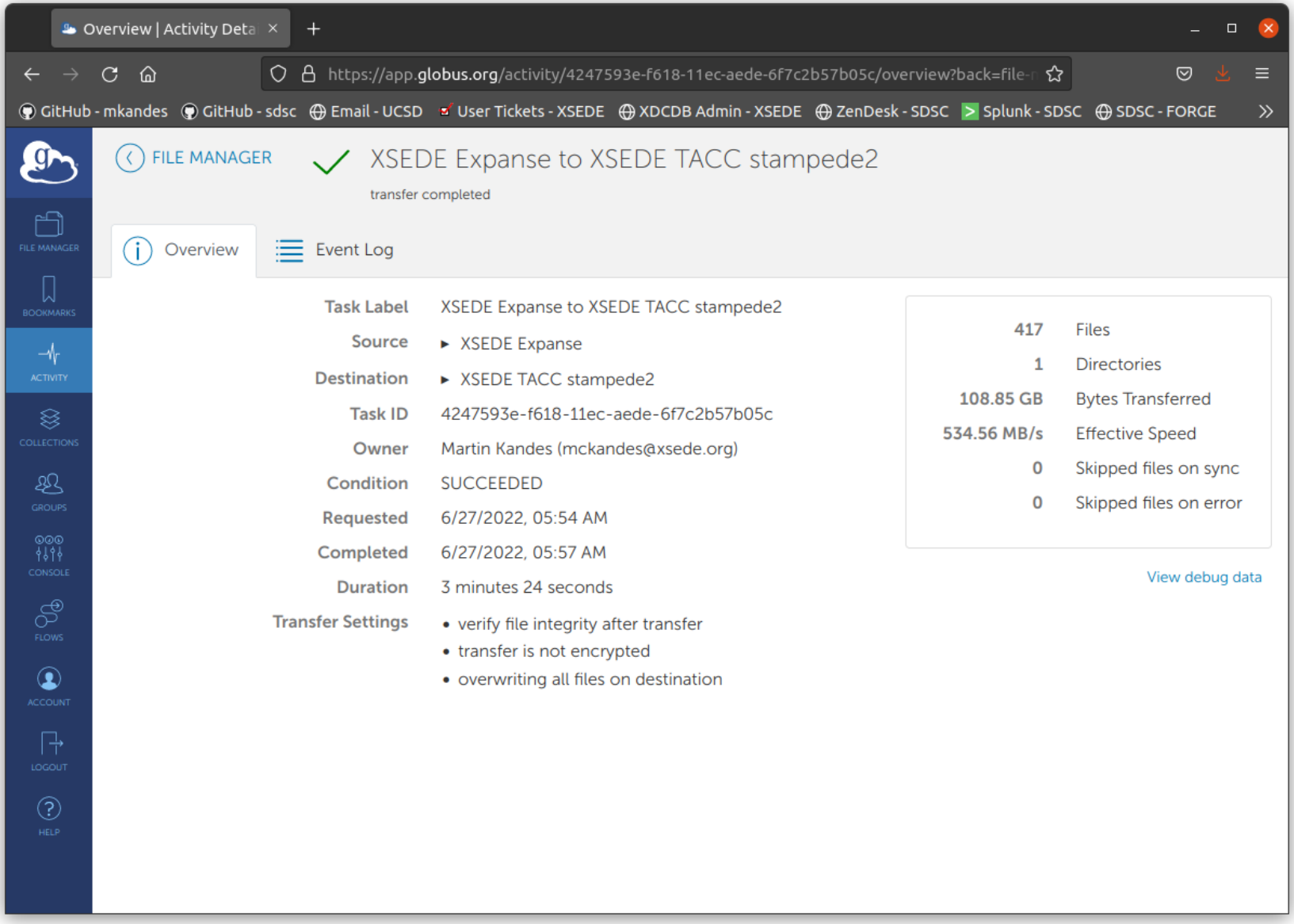
Globus Directory Transfer - Setup



Globus Directory Transfer - Queued



Globus Directory Transfer - Completed



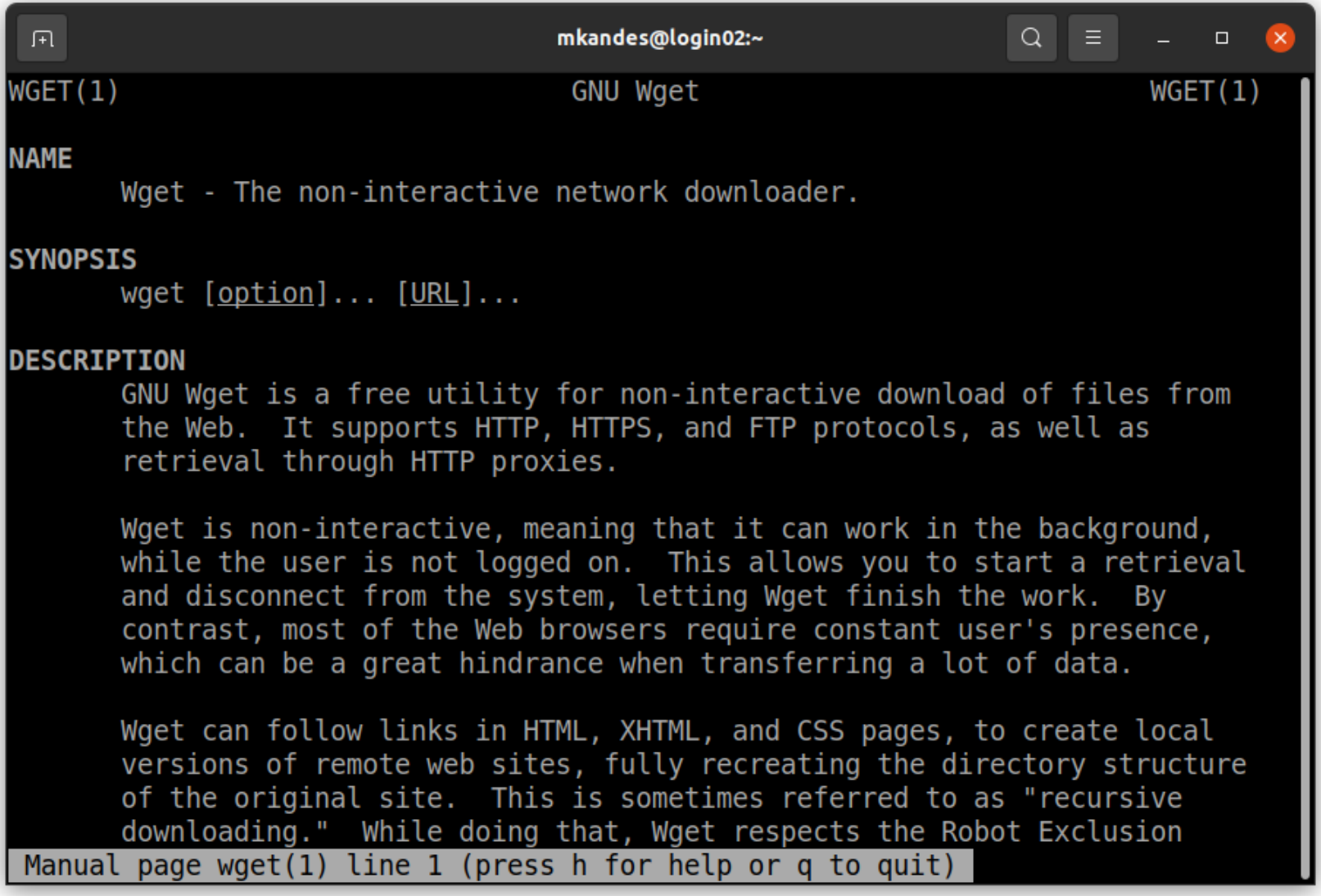
Getting data to (and from) an HPC system

- Transfer: `wget`, curl, `scp`, `sftp`, `rclone`
- Archiving and Compression: `gzip`, `gunzip`, `tar`
- Integrity: `md5sum`, `sha256sum`

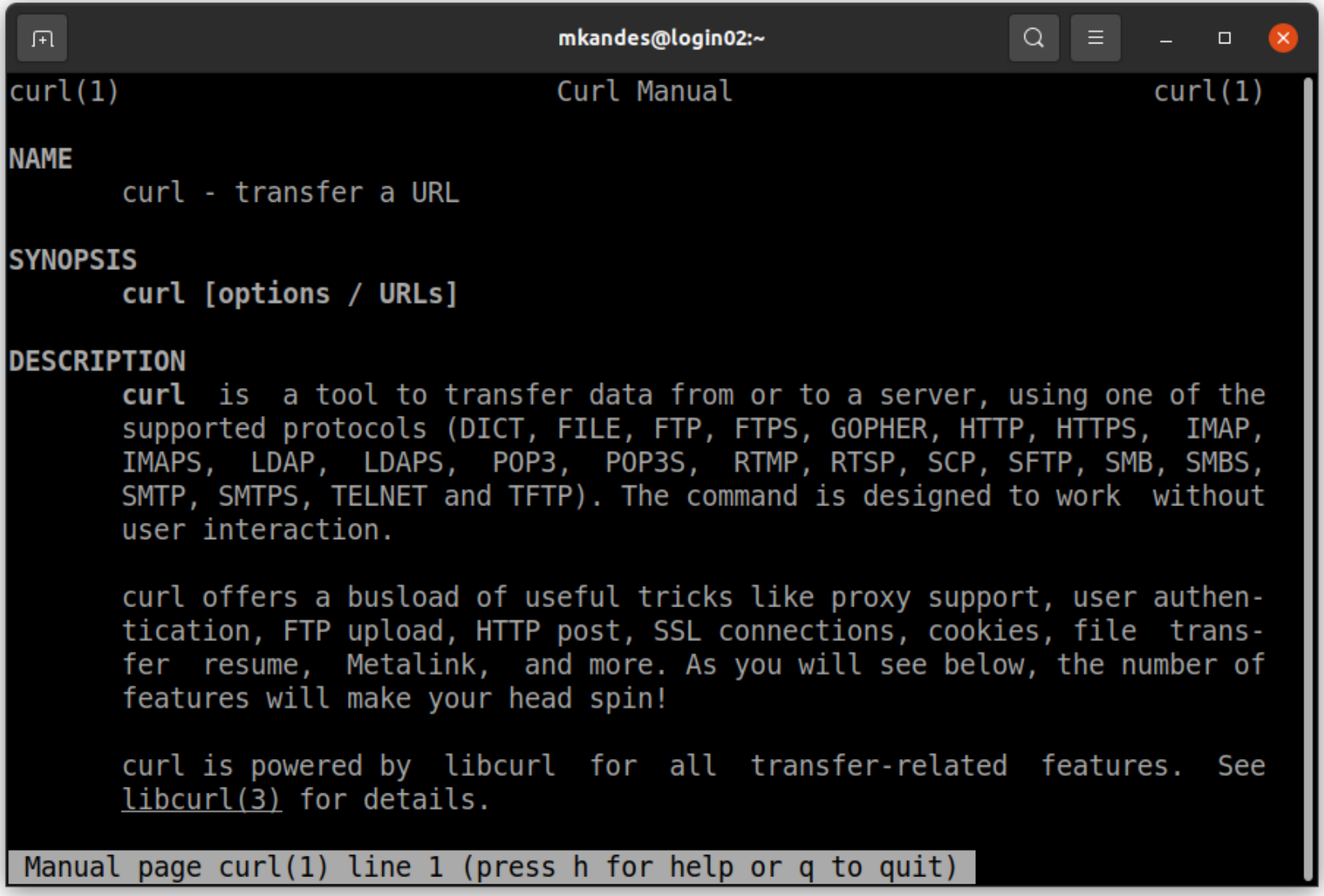
Example Dataset

https://ftp.ncbi.nlm.nih.gov/gene/DATA/gene_info.gz

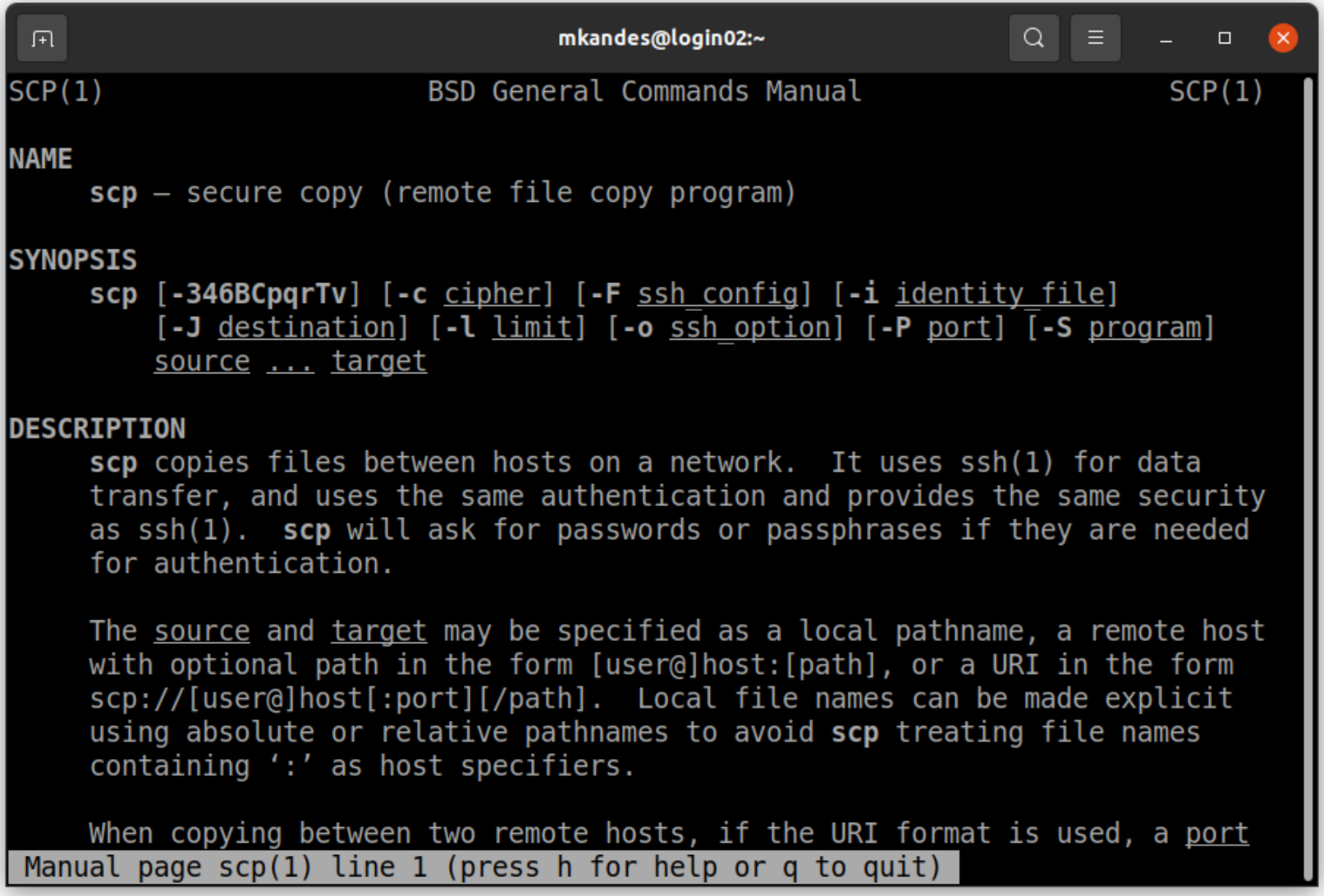
wget


mkandes@login02:~
WGET(1) GNU Wget WGET(1)
NAME
Wget - The non-interactive network downloader.
SYNOPSIS
wget [option]... [URL]...
DESCRIPTION
GNU Wget is a free utility for non-interactive download of files from the Web. It supports HTTP, HTTPS, and FTP protocols, as well as retrieval through HTTP proxies.
Wget is non-interactive, meaning that it can work in the background, while the user is not logged on. This allows you to start a retrieval and disconnect from the system, letting Wget finish the work. By contrast, most of the Web browsers require constant user's presence, which can be a great hindrance when transferring a lot of data.
Wget can follow links in HTML, XHTML, and CSS pages, to create local versions of remote web sites, fully recreating the directory structure of the original site. This is sometimes referred to as "recursive downloading." While doing that, Wget respects the Robot Exclusion
Manual page wget(1) line 1 (press h for help or q to quit)

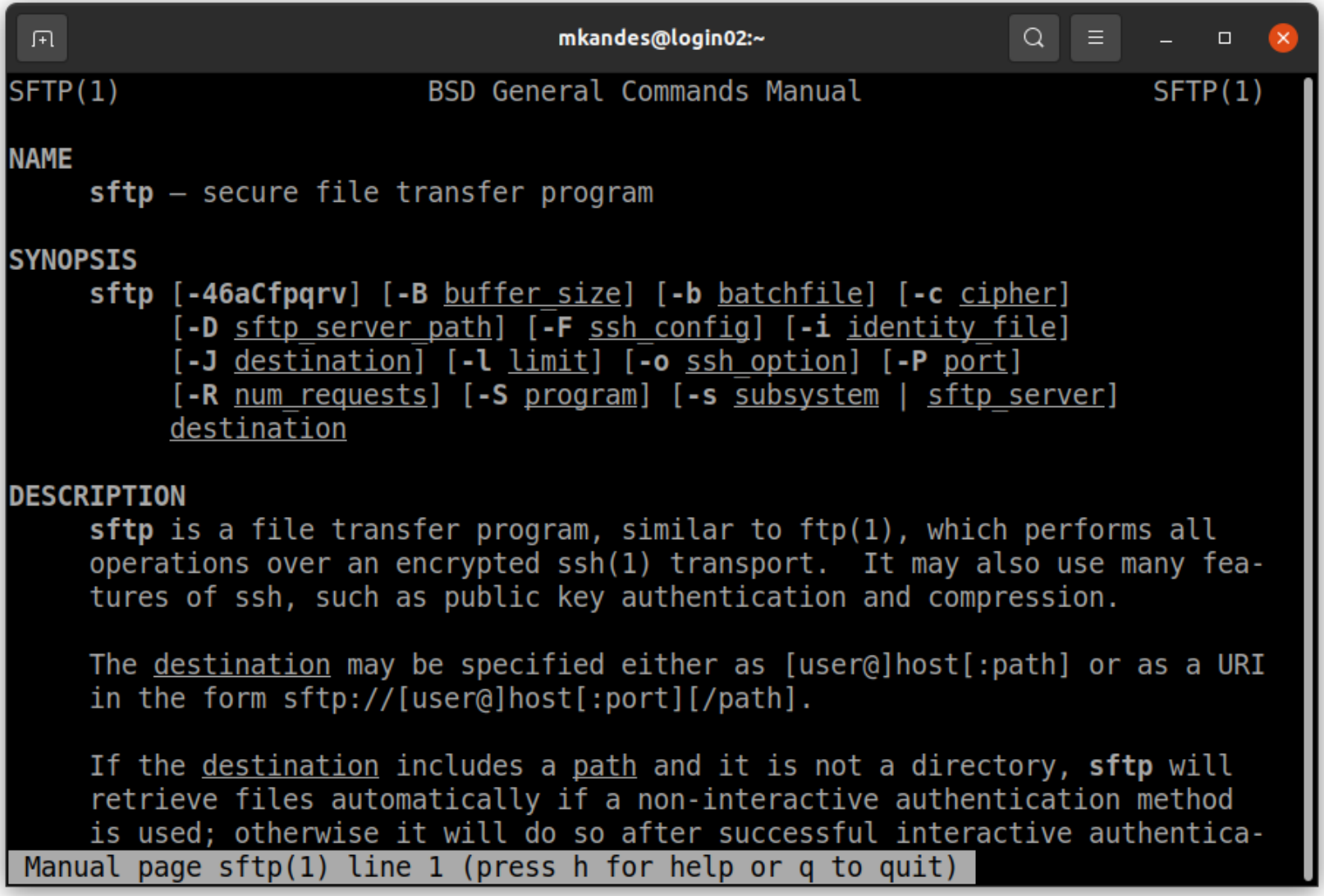
curl


mkandes@login02:~
curl(1) Curl Manual curl(1)
NAME
curl - transfer a URL
SYNOPSIS
curl [options / URLs]
DESCRIPTION
curl is a tool to transfer data from or to a server, using one of the supported protocols (DICT, FILE, FTP, FTPS, GOPHER, HTTP, HTTPS, IMAP, IMAPS, LDAP, LDAPS, POP3, POP3S, RTMP, RTSP, SCP, SFTP, SMB, SMBS, SMTP, SMTPS, TELNET and TFTP). The command is designed to work without user interaction.
curl offers a busload of useful tricks like proxy support, user authentication, FTP upload, HTTP post, SSL connections, cookies, file transfer resume, Metalink, and more. As you will see below, the number of features will make your head spin!
curl is powered by libcurl for all transfer-related features. See libcurl(3) for details.
Manual page curl(1) line 1 (press h for help or q to quit)

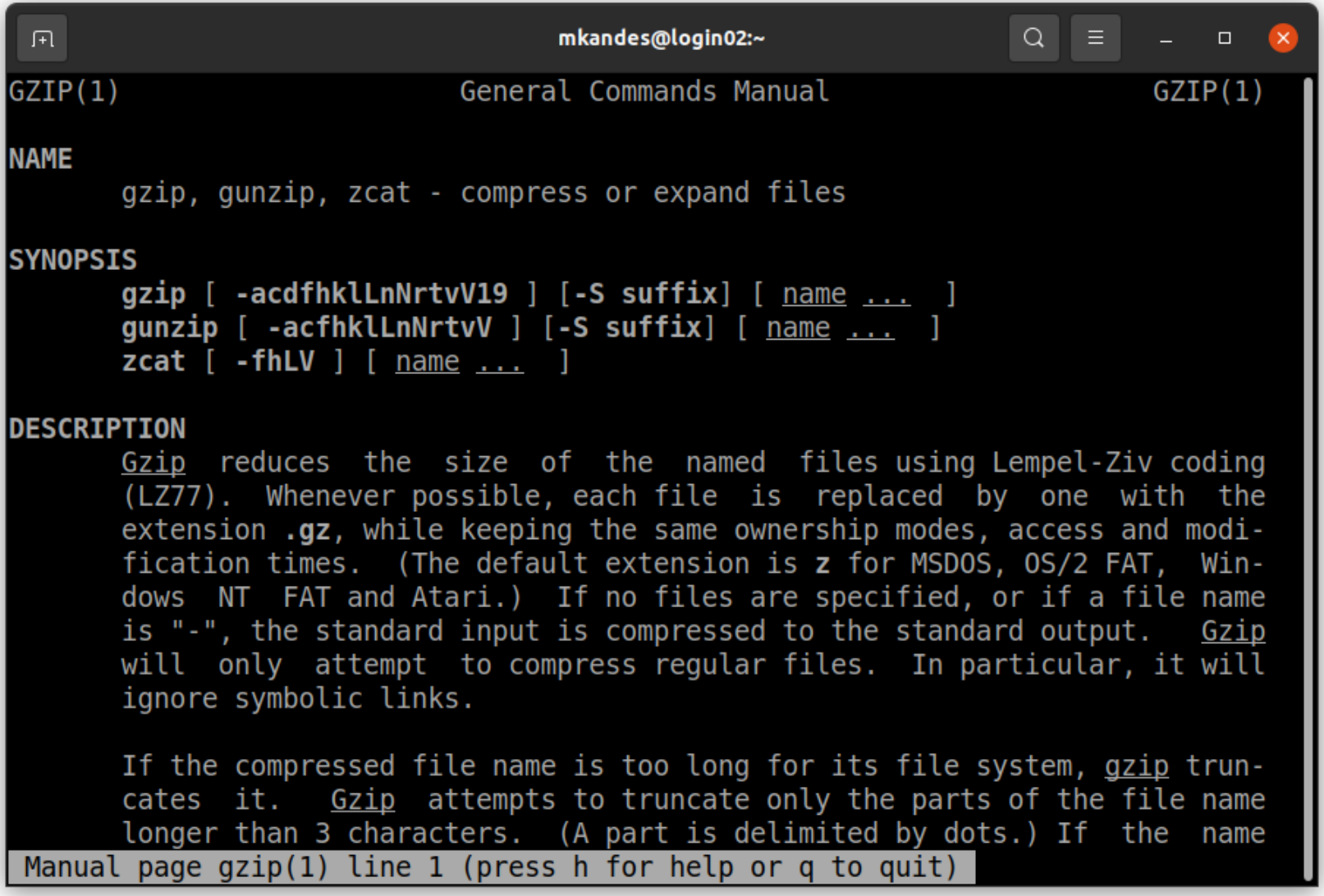
scp


mkandes@login02:~
SCP(1) BSD General Commands Manual SCP(1)
NAME
scp – secure copy (remote file copy program)
SYNOPSIS
scp [-346BCpqrTv] [-c cipher] [-F ssh_config] [-i identity_file]
[-J destination] [-l limit] [-o ssh_option] [-P port] [-S program]
source ... target
DESCRIPTION
scp copies files between hosts on a network. It uses ssh(1) for data
transfer, and uses the same authentication and provides the same security
as ssh(1). scp will ask for passwords or passphrases if they are needed
for authentication.
The source and target may be specified as a local pathname, a remote host
with optional path in the form [user@]host:[path], or a URI in the form
scp://[user@]host[:port][/path]. Local file names can be made explicit
using absolute or relative pathnames to avoid scp treating file names
containing ':' as host specifiers.
When copying between two remote hosts, if the URI format is used, a port
Manual page scp(1) line 1 (press h for help or q to quit)

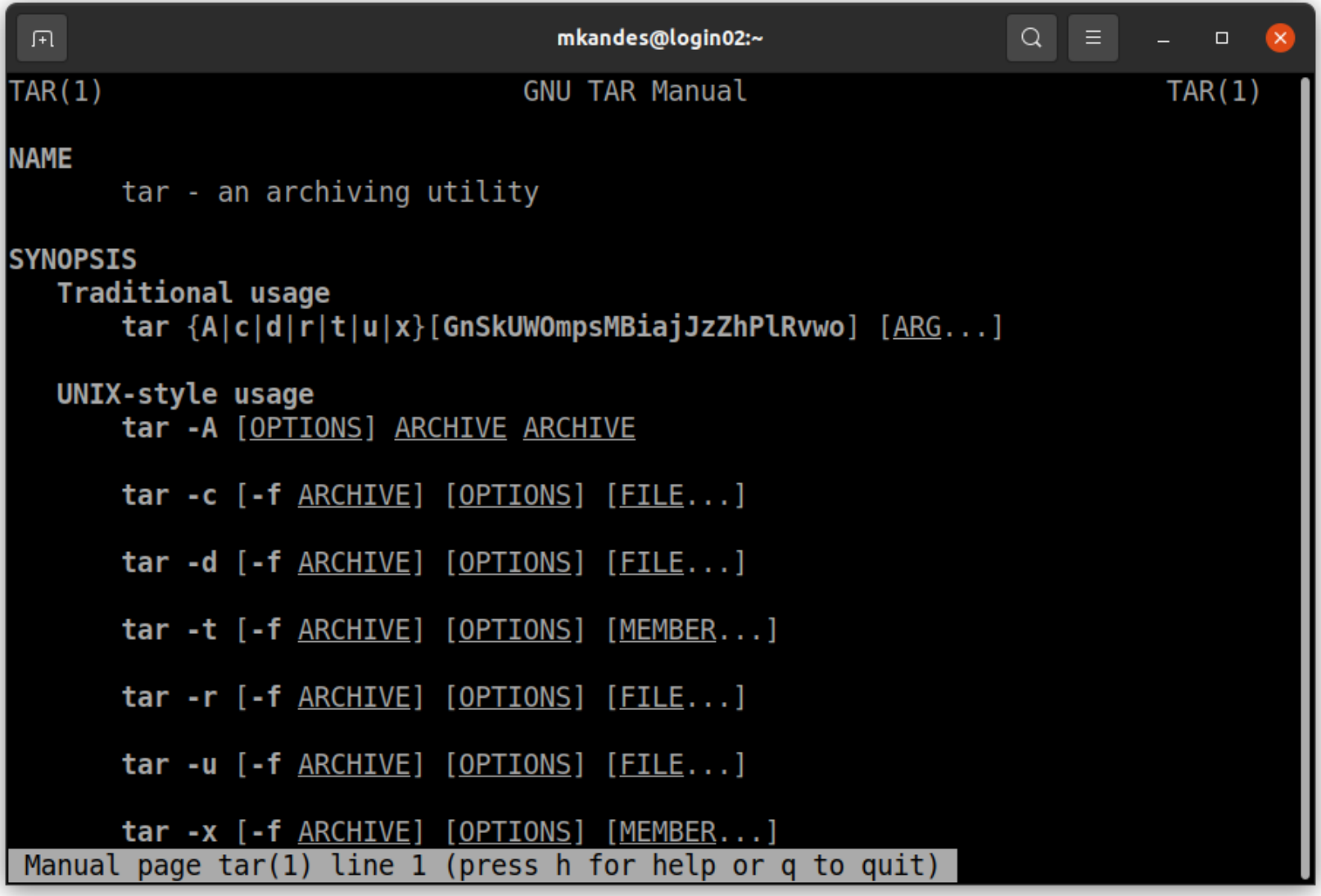
sftp


mkandes@login02:~
SFTP(1) BSD General Commands Manual SFTP(1)
NAME
sftp — secure file transfer program
SYNOPSIS
sftp [-46aCfpqrv] [-B buffer_size] [-b batchfile] [-c cipher]
[-D sftp_server_path] [-F ssh_config] [-i identity_file]
[-J destination] [-l limit] [-o ssh_option] [-P port]
[-R num_requests] [-S program] [-s subsystem | sftp_server]
destination
DESCRIPTION
sftp is a file transfer program, similar to ftp(1), which performs all
operations over an encrypted ssh(1) transport. It may also use many fea-
tures of ssh, such as public key authentication and compression.
The destination may be specified either as [user@]host[:path] or as a URI
in the form sftp://[user@]host[:port][/path].
If the destination includes a path and it is not a directory, sftp will
retrieve files automatically if a non-interactive authentication method
is used; otherwise it will do so after successful interactive authentica-
Manual page sftp(1) line 1 (press h for help or q to quit)

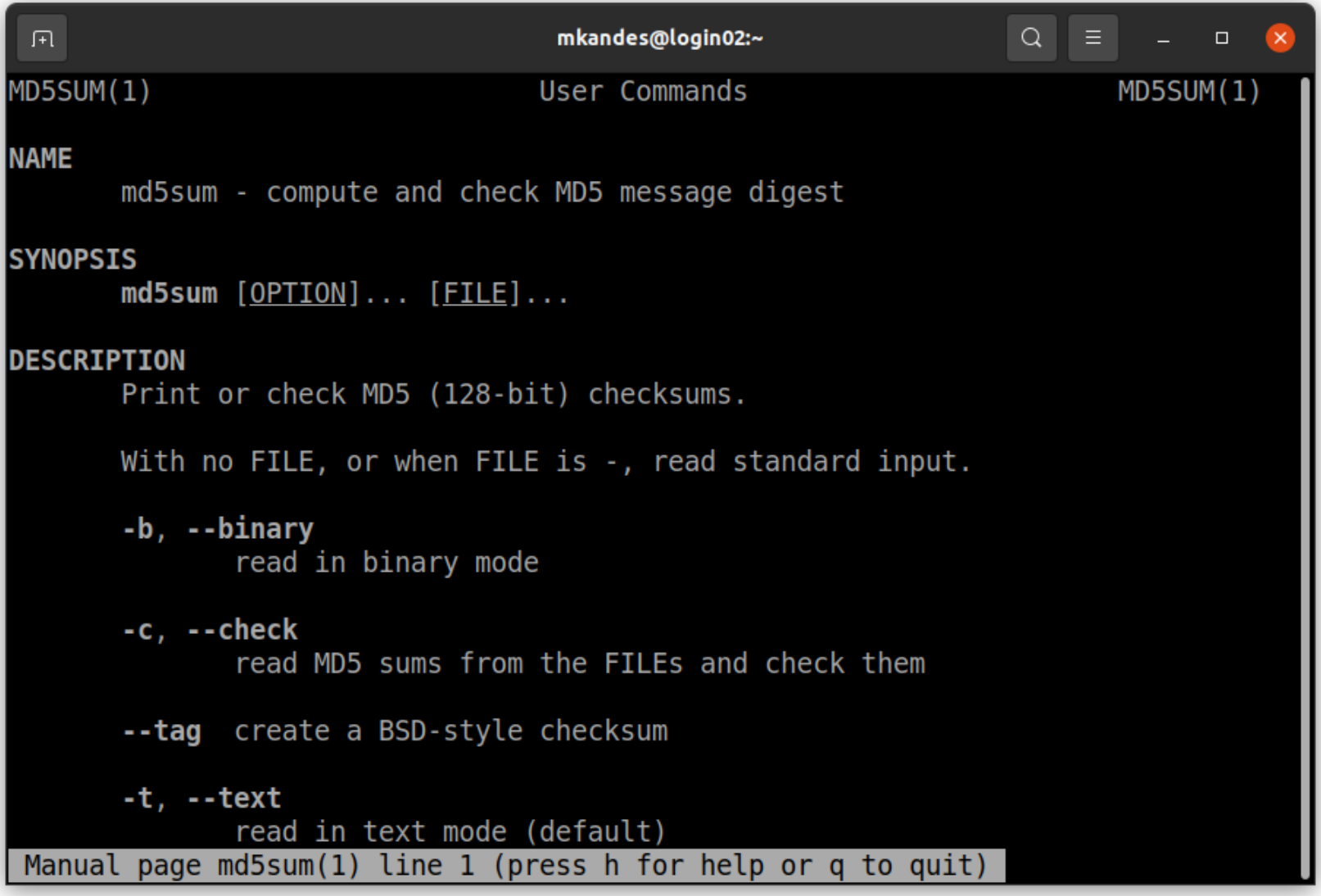
gzip


mkandes@login02:~
GZIP(1) General Commands Manual GZIP(1)
NAME
gzip, gunzip, zcat - compress or expand files
SYNOPSIS
gzip [-acdfhklLnNrtvV19] [-S suffix] [name ...]
gunzip [-acfhklLnNrtvV] [-S suffix] [name ...]
zcat [-fhLV] [name ...]
DESCRIPTION
Gzip reduces the size of the named files using Lempel-Ziv coding (LZ77). Whenever possible, each file is replaced by one with the extension .gz, while keeping the same ownership modes, access and modification times. (The default extension is z for MSDOS, OS/2 FAT, Windows NT FAT and Atari.) If no files are specified, or if a file name is "-", the standard input is compressed to the standard output. Gzip will only attempt to compress regular files. In particular, it will ignore symbolic links.
If the compressed file name is too long for its file system, gzip truncates it. Gzip attempts to truncate only the parts of the file name longer than 3 characters. (A part is delimited by dots.) If the name
Manual page gzip(1) line 1 (press h for help or q to quit)

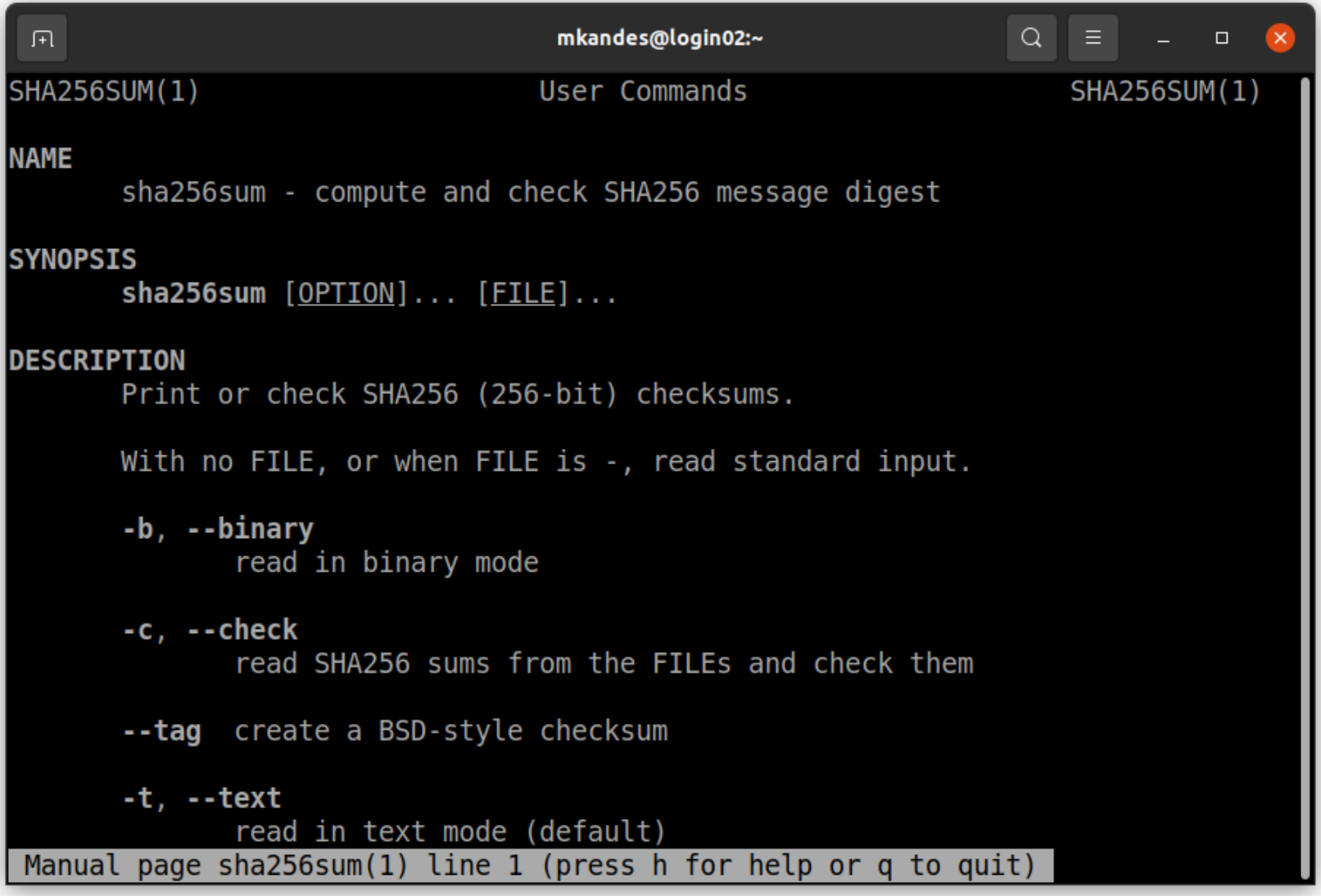
tar


mkandes@login02:~
TAR(1) GNU TAR Manual TAR(1)
NAME
tar - an archiving utility
SYNOPSIS
Traditional usage
tar {A|c|d|r|t|u|x}[GnSkUWOmpsMBiajJzZhPlRvwo] [ARG...]
UNIX-style usage
tar -A [OPTIONS] ARCHIVE ARCHIVE
tar -c [-f ARCHIVE] [OPTIONS] [FILE...]
tar -d [-f ARCHIVE] [OPTIONS] [FILE...]
tar -t [-f ARCHIVE] [OPTIONS] [MEMBER...]
tar -r [-f ARCHIVE] [OPTIONS] [FILE...]
tar -u [-f ARCHIVE] [OPTIONS] [FILE...]
tar -x [-f ARCHIVE] [OPTIONS] [MEMBER...]
Manual page tar(1) line 1 (press h for help or q to quit)

md5sum


mkandes@login02:~
MD5SUM(1) User Commands MD5SUM(1)
NAME
md5sum - compute and check MD5 message digest
SYNOPSIS
md5sum [OPTION]... [FILE]...
DESCRIPTION
Print or check MD5 (128-bit) checksums.
With no FILE, or when FILE is -, read standard input.
-b, --binary
read in binary mode
-c, --check
read MD5 sums from the FILEs and check them
--tag create a BSD-style checksum
-t, --text
read in text mode (default)
Manual page md5sum(1) line 1 (press h for help or q to quit)

sha256sum


mkandes@login02:~
SHA256SUM(1) User Commands SHA256SUM(1)
NAME
sha256sum - compute and check SHA256 message digest
SYNOPSIS
sha256sum [OPTION]... [FILE]...
DESCRIPTION
Print or check SHA256 (256-bit) checksums.
With no FILE, or when FILE is -, read standard input.
-b, --binary
read in binary mode
-c, --check
read SHA256 sums from the FILEs and check them
--tag create a BSD-style checksum
-t, --text
read in text mode (default)
Manual page sha256sum(1) line 1 (press h for help or q to quit)

Managing data on an HPC system

Use local `/scratch` for data intensive I/O

Reading, writing, and preparing data for HPC

Pre-process raw datasets into ML framework native formats

Preventing data loss

Back up your data regularly! Try `rsync`

Questions?

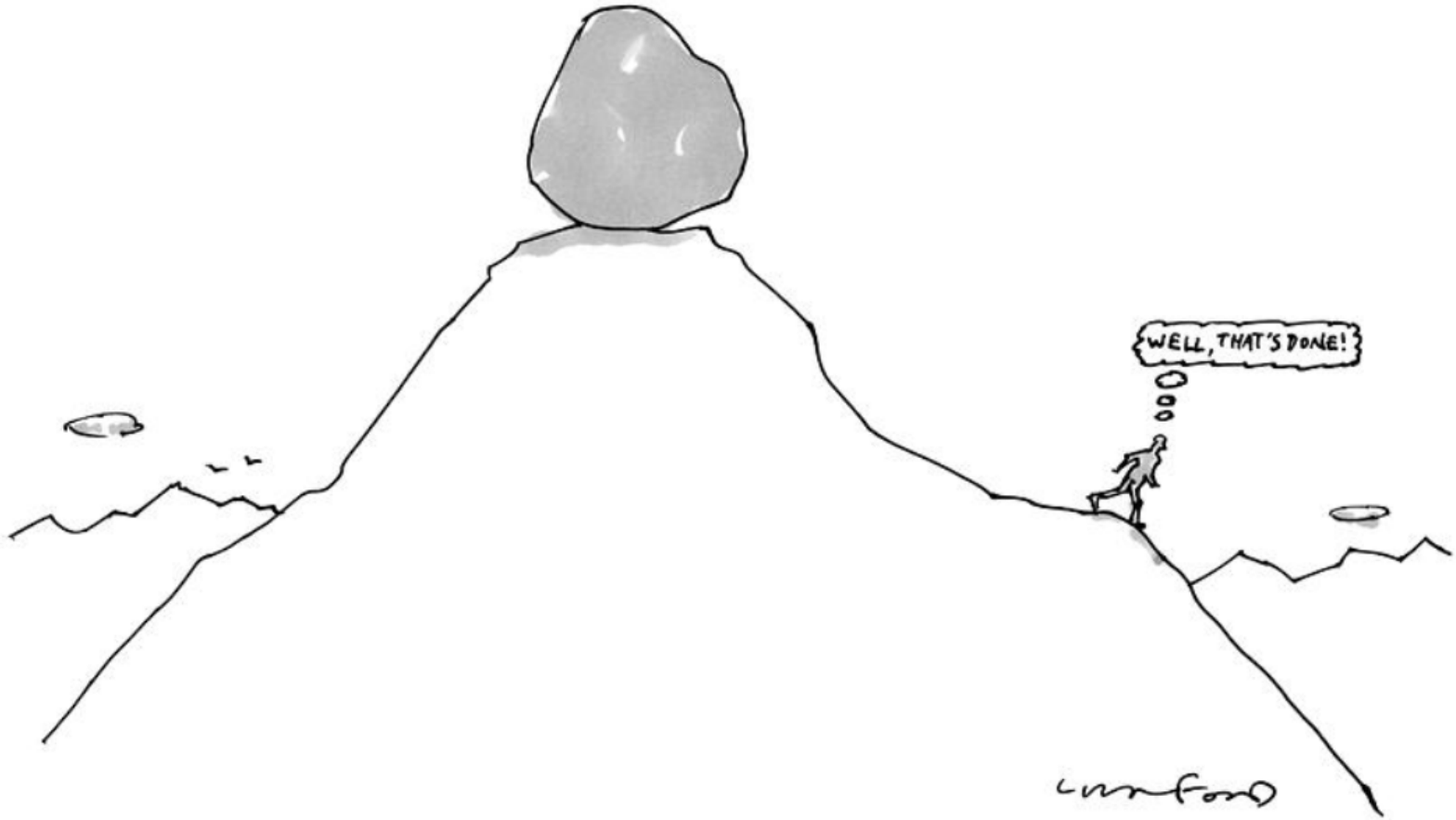


Image Credit: New Yorker - M. Crawford