ManeFrame II (M2) Introduction

Dr. Robert Kalescky
HPC Applications Scientist
Office of Information Technology and Center for Research Computing

Dr. John LaGrone
HPC Applications Scientist
Office of Information Technology and Center for Research Computing


OIT Research and Data Science Services
Center for Research Computing

Research and Data Science Services

- Provide research computing support
- High-Performance Computing - Dr. Robert Kalescky
- High-Performance Computing - Dr. John LaGrone
- Data Science - Dr. Eric Godat
- Data Science - Dr. Tue Vue
- Custom Devices (IOT, wearables, etc.) - Guillermo Vasquez

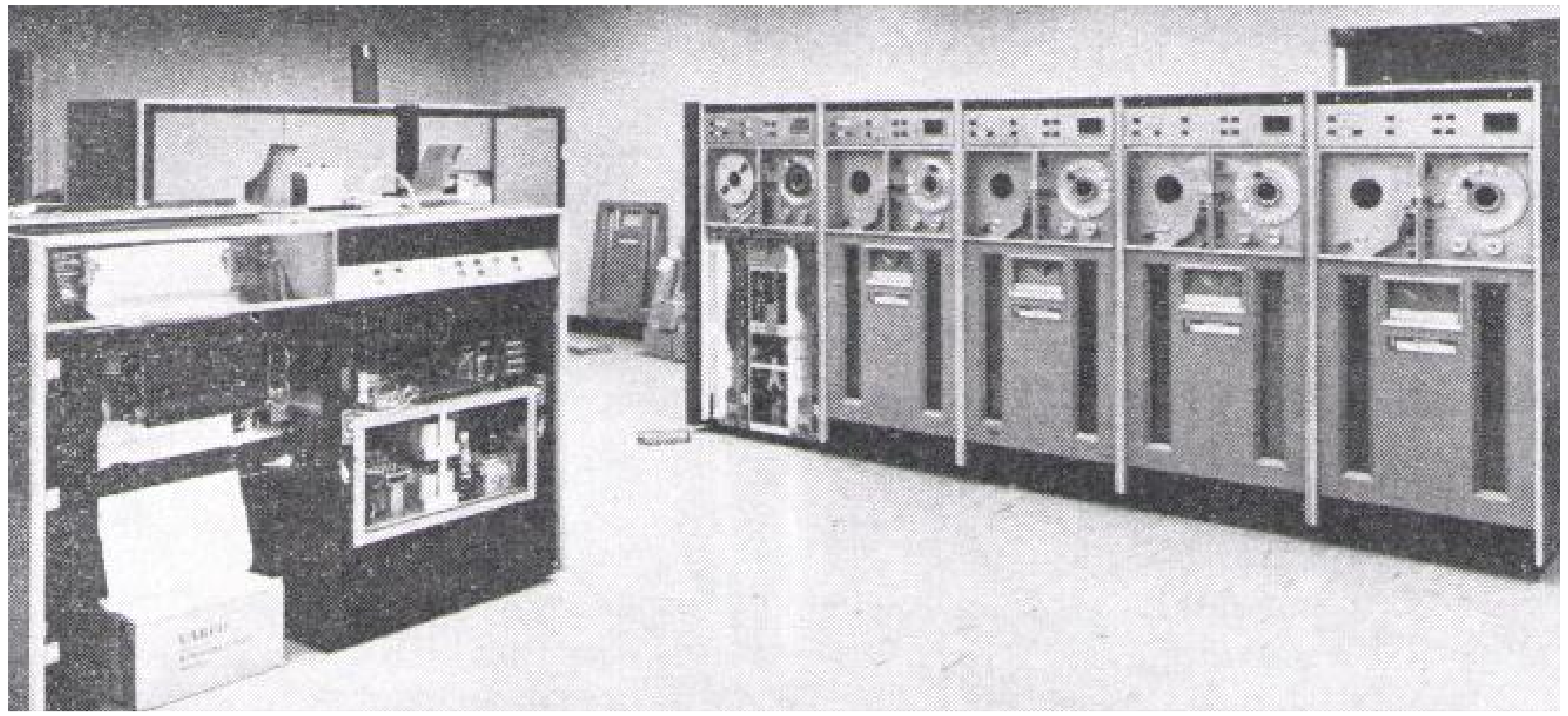
Center for Research Computing

- Maintains our primary shared resource for research computing, ManeFrame II (M2), in collaboration with OIT
- Provides research computing tools, support, and training to all faculty, staff, and students using research computing resources
- www.smu.edu/crc has documentation and news
- help@smu.edu or rkalescky@smu.edu for help
- Request an account at www.smu.edu/crc

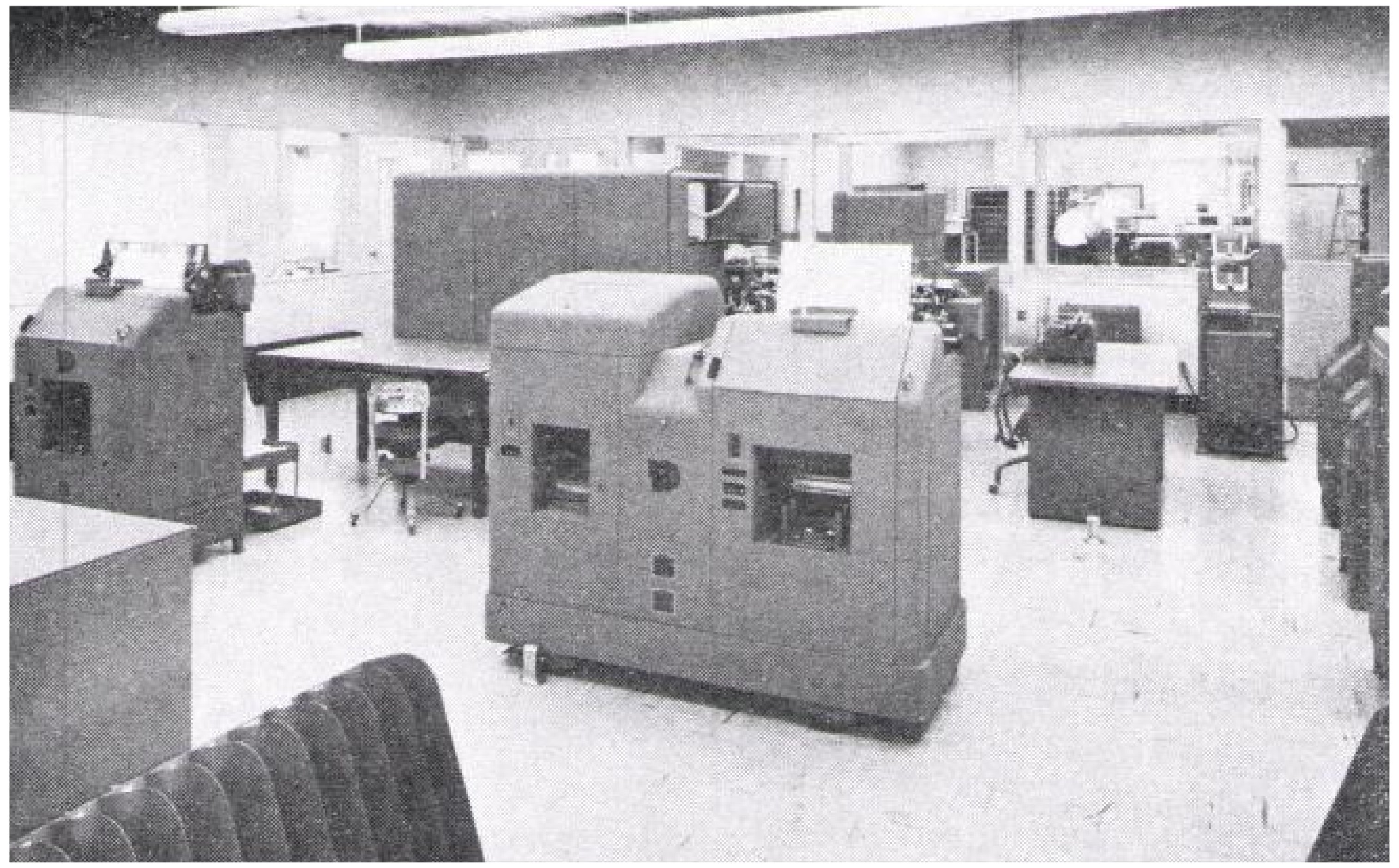
SMU's HPC History

UNIVAC Computing Center in 1957

Punch card machines



Printer (left) and magnetic tap storage (right)



UNIVAC 120 (back left)

High Performance Computing

Moore's Law

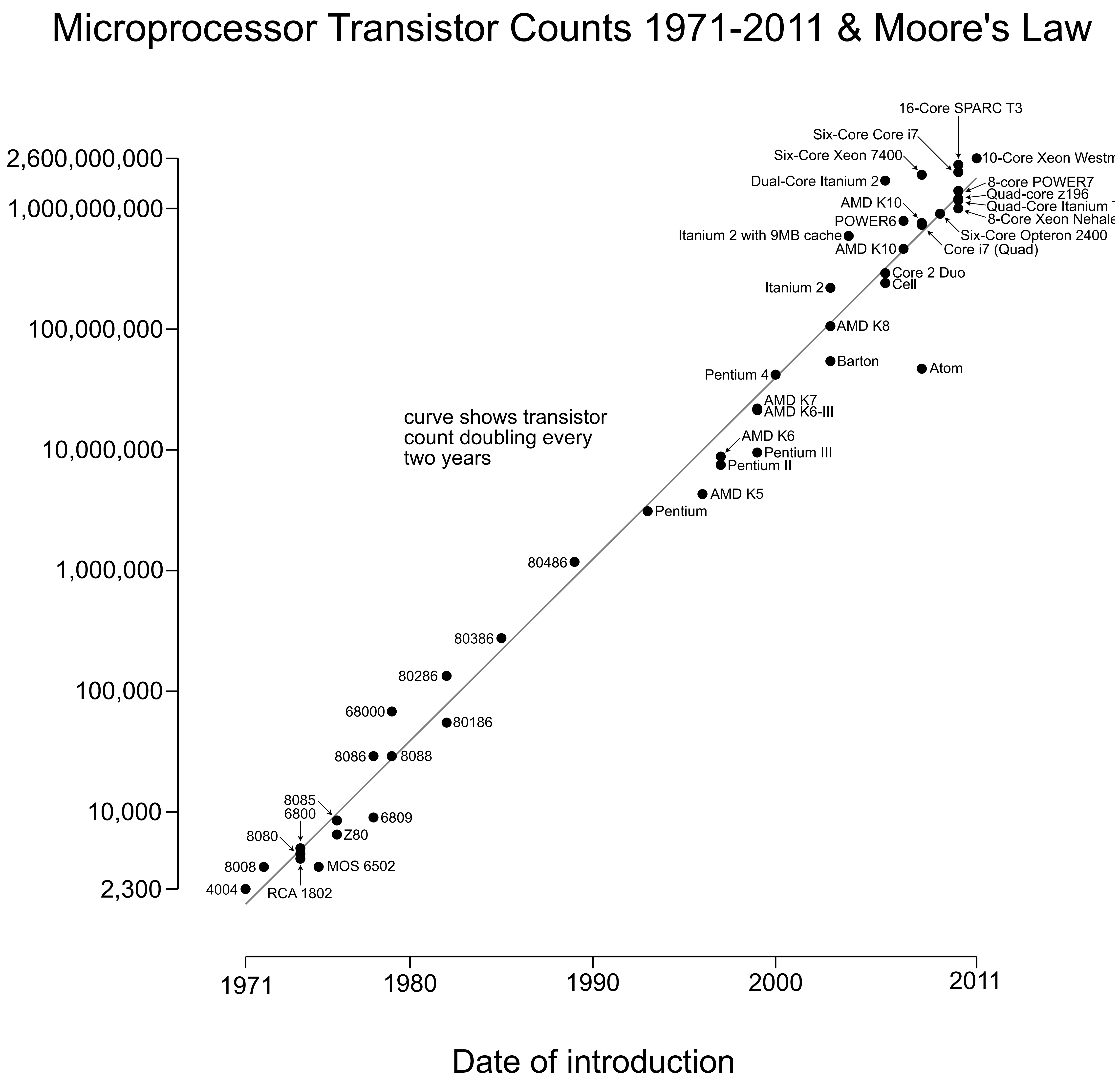
- Historically, we have depended on hardware advances to enable faster and larger simulations

- In 1965, Gordon Moore observed that the CPU and RAM transistor count about doubled each year

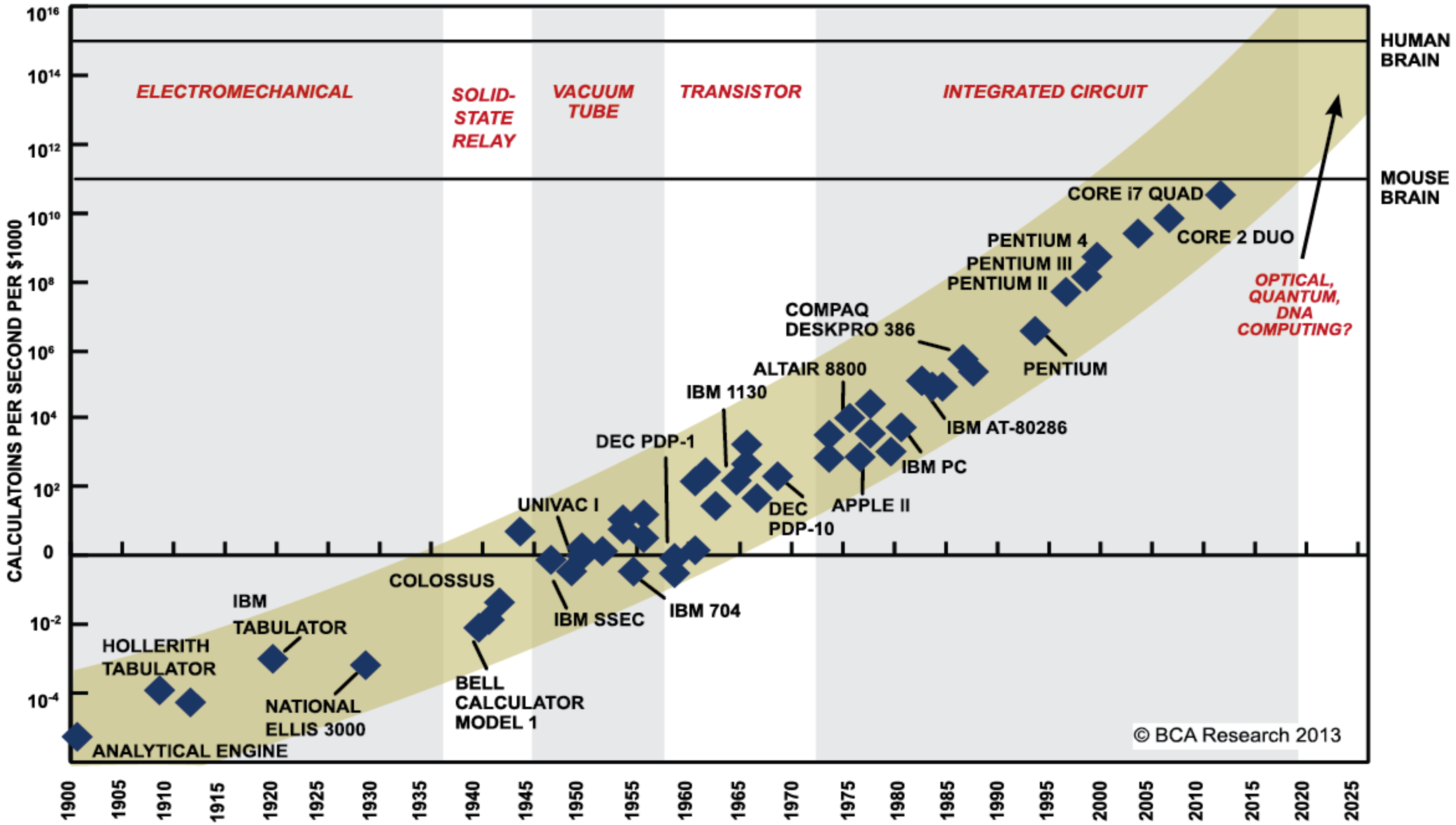
- "Moore's Law" has since been revised to a doubling once every 2 years, with startling accuracy

- Physical limits, e.g. power consumption, heat emission, and even the size of the atom, have currently stopped this expansion on individual processors, with speeds that have leveled off since around 2008

Growth in CPU Transistor Counts



Decrease in Cost Per FLOPS



SOURCE: RAY KURZWEIL, "THE SINGULARITY IS NEAR: WHEN HUMANS TRANSCEND BIOLOGY", P.67, *THE VIKING PRESS*, 2006. DATAPOINTS BETWEEN 2000 AND 2012 REPRESENT BCA ESTIMATES.

Additional Problems

- Growing performance discrepancy between processing units and data storage
  - 40% processor performance improvement per year
  - 10% RAM performance improvement per year
  - Discrepancy leads to memory-bound programs
    - CPU spends more and more time idle, waiting on data from
- Many simulations require incredible amounts of memory to achieve high-accuracy solutions (PDE, MD, QM, etc.)
  - These cannot fit on a single computer alone

The Parallel Solution

- Using multiple processing units (cores, etc.) simultaneously to perform a computation

- Use multiple computers to store data for large problems

- Essentially all modern CPUs have multiple cores

Unfortunately…

“I know how to make 4 horses pull a cart - I don't know how to make 1024 chickens do it.”

Enrico Clementi

What is High Performance Computing?

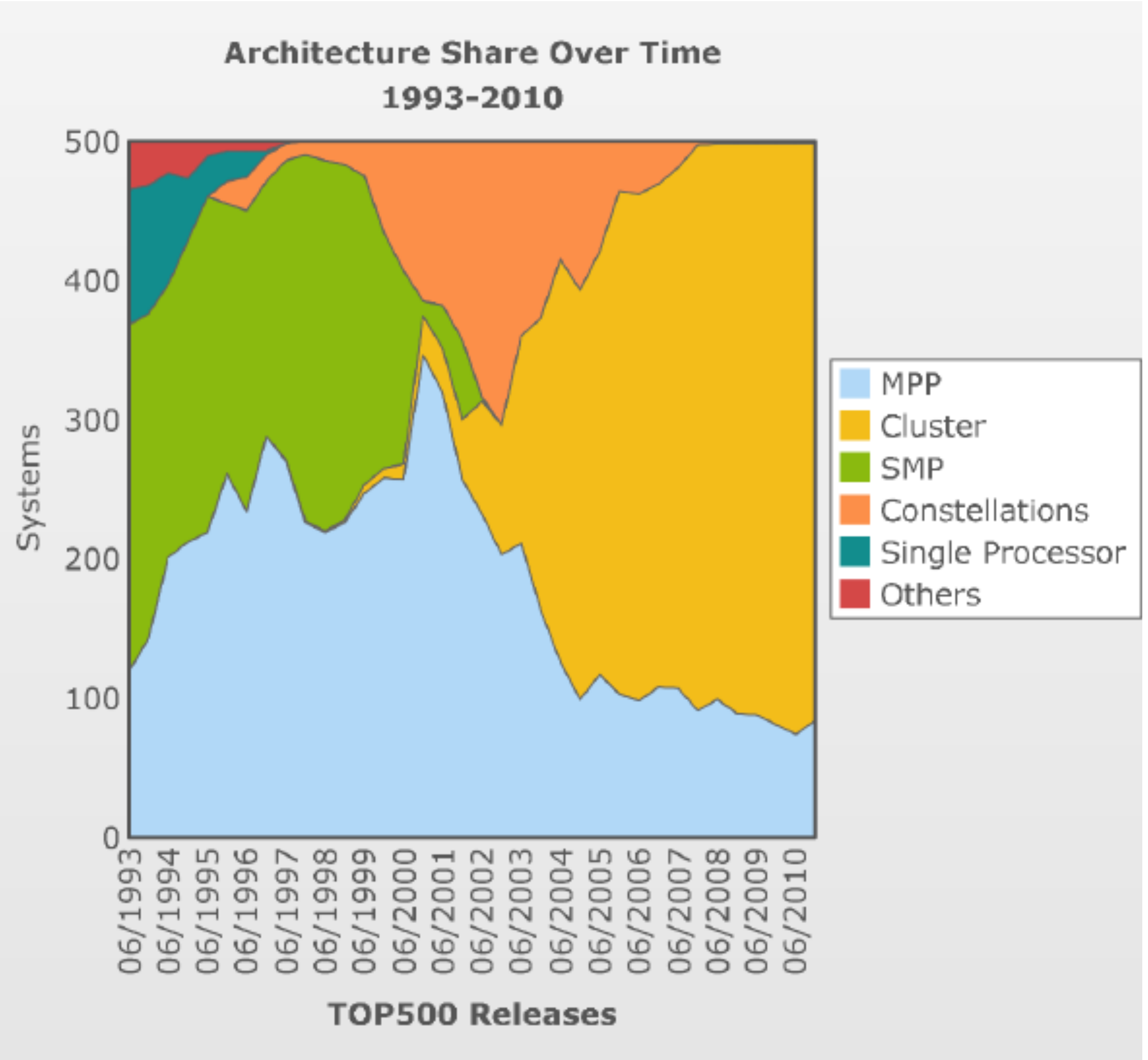
High performance computing (HPC) is the use of computing resources that are significantly more powerful than what is commonly available.*

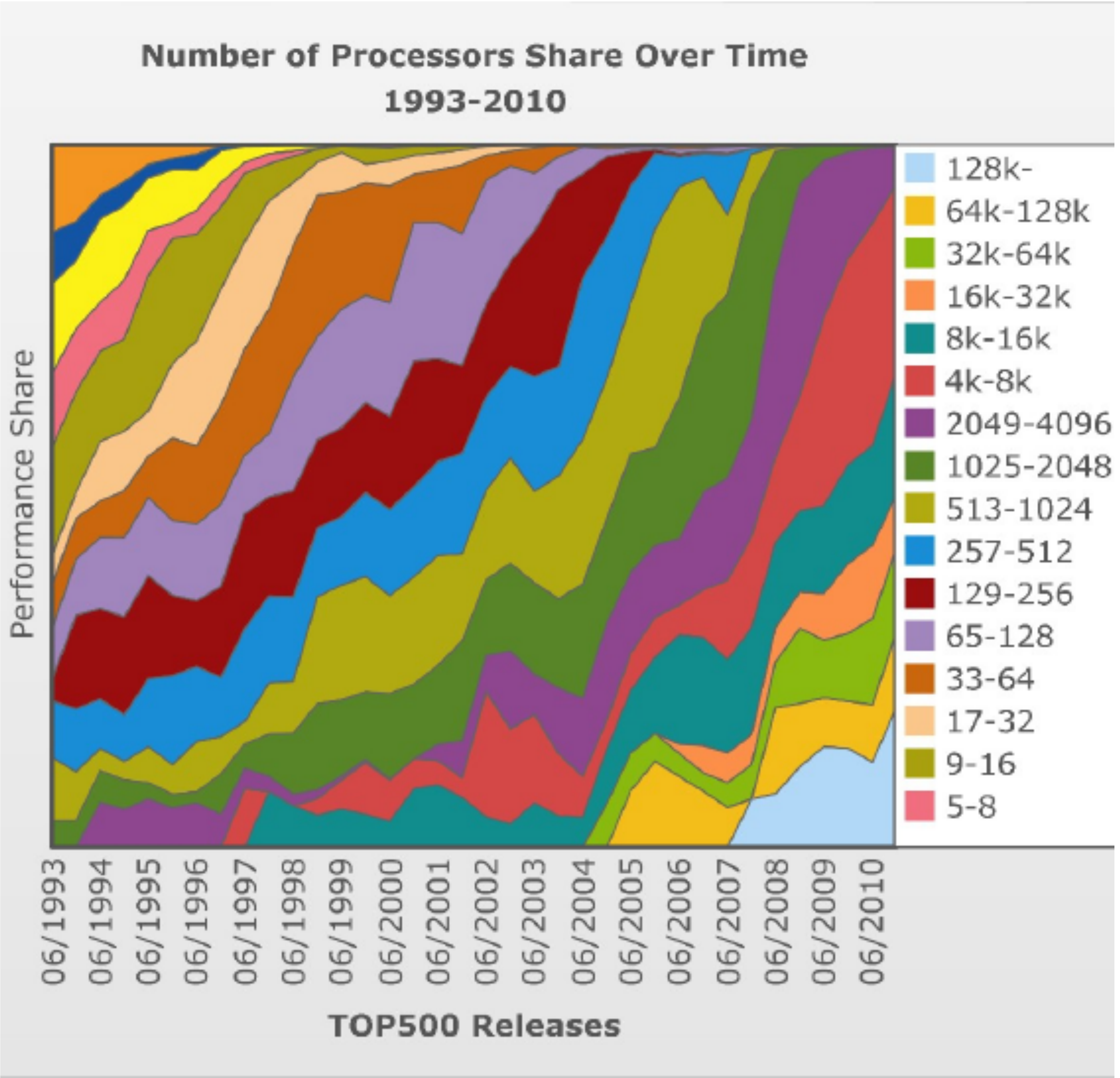
*As such, it's always a moving target.

Supercomputer Types

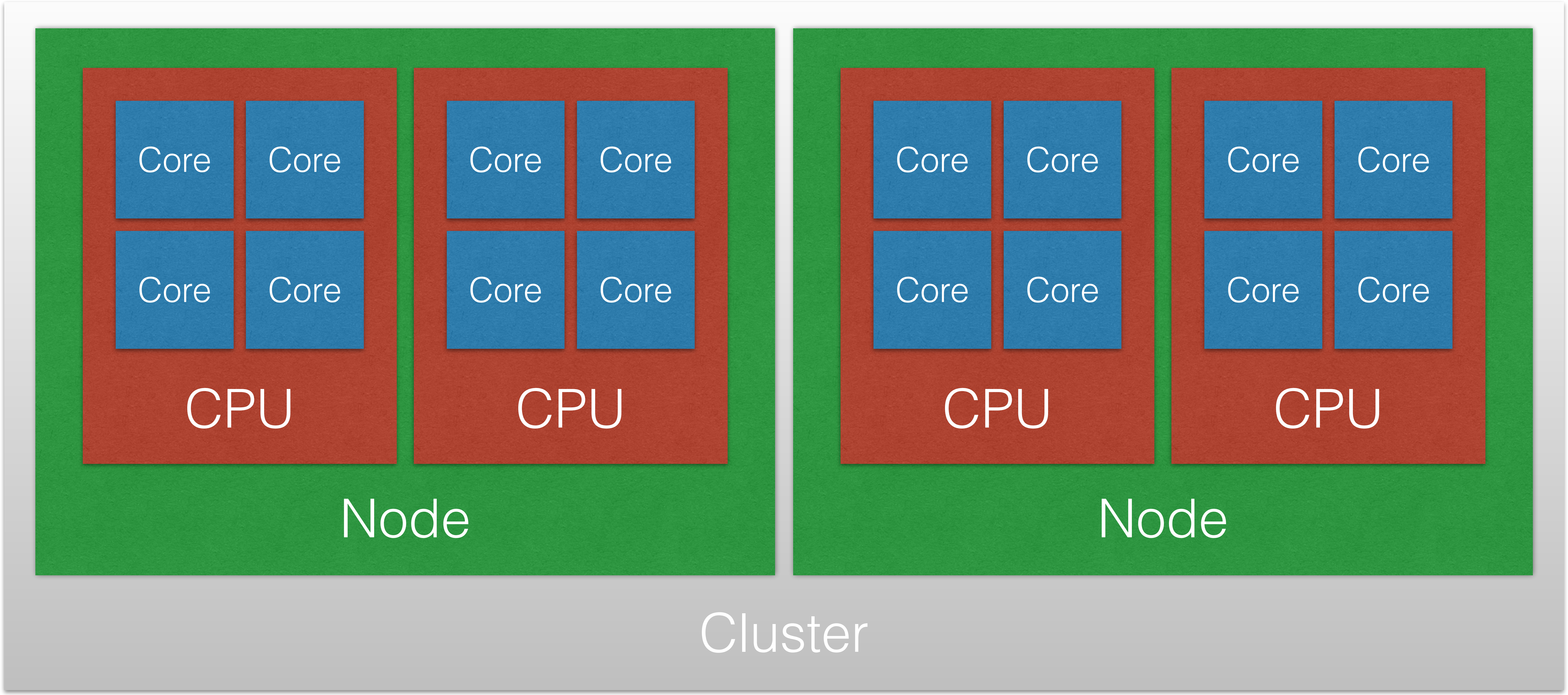
- Single processor
- Shared Memory Parallel (SMP)
- Massively Parallel Processors
- Constellations
- Clusters

History of Parallel Architectures



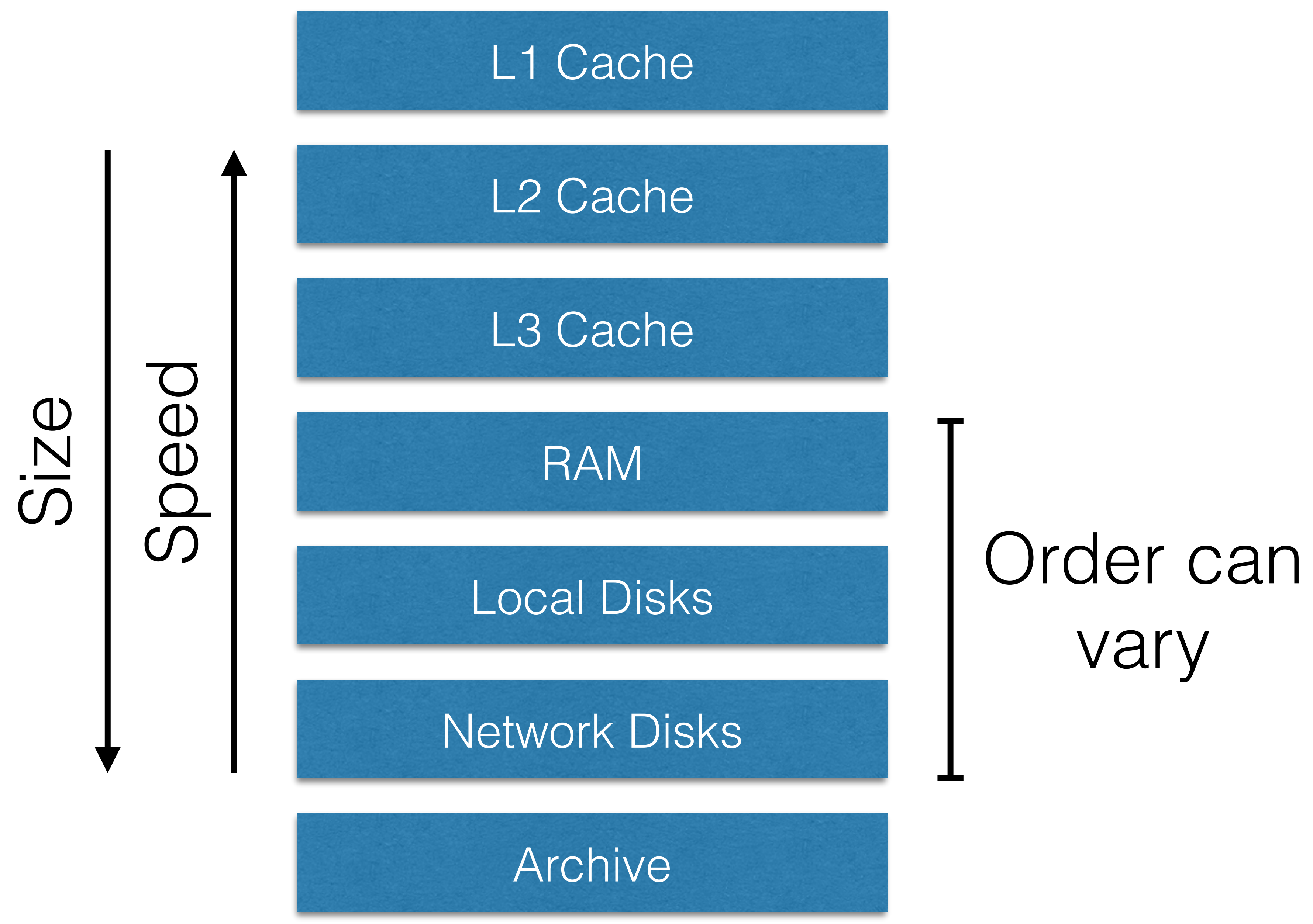


Cluster Compute Resources



Cluster Storage Resources

- Usually there is an inverse relationship between storage amount and speed
- Clusters and applications can distort the standard order
    - Aggregation, e.g. RAID
    - Network optimization to a specific topology
    - Applications can be configured and built for specific hardware
    - New hardware technologies



WELCOME TO

```
                       ___                                    _________
   /\/\   __ _ _ __   ___  / __\ __ __ _ _ __ ___   ___       /  _/  _/
  /    \ / _` | '_ \ / _ \/ _\| '__/ _` | '_ ` _ \ / _ \     / / / /
 / /\/\ \ (_| | | | |  __/ /  | | | (_| | | | | | |  __/   _/ / /_/ /
 \/    \/\__,_|_|_| |_|\___\/   |_|  \__,_|_| |_| |_|\___|  /___/___/
```

ManeFrame II

- High performance features:
  - CPUs
  - Accelerators
  - Networking technologies
  - Substantial memory per node
  - Interactive GPU-accelerated remote desktop experiences

ManeFrame II

- CentOS 7 operating system
- SLURM resource scheduler
- Lmod environment module system.
- Development tool chains including the GCC, LLVM, Intel, and PGI compiler suites
- Optimized high-level programming environments such as MATLAB, Python, and R

ManeFrame II Node Types

Type	Quantity	Cores	Memory [GB]	Additional Resources
Standard-Memory	176	36	256	
Medium-Memory-1	35	36	768	
Medium-Memory-2	4	24	768	3 TB SSD local scratch
High-Memory-1	5	36	1,536	
High-Memory-2	6	40	1,536	3 TB SSD local scratch
GPGPU-1	36	36	256	NVIDIA P100 GPU has 3,584 CUDA cores and 16 GB CoWoS
MIC-1	36	64	384	16 GB of high bandwidth (400 GB/s) stacked memory
VDI	5	36	256	NVIDIA Quadro M5000 GPU
v100x8	3	36	768	8 NVIDIA V100 GPUs with 5,120 CUDA cores and 32 GB CoWoS
Faculty Partner Nodes	3			Various research specific NVIDIA GPU configurations
ManeFrame II	**354**	**11,276**	**120 TB**	**2.8 PB storage and InfiniBand network**

ManeFrame File Systems

- $HOME
  - NFS file system, located in /users (e.g. /users/$USER)
  - When you log in to ManeFrame your home directory is where you will land in by default
  - This space should be used to write, edit, compile, and browse your programs and job submission scripts
  - $HOME space is restricted by quotas (200 GB)

ManeFrame II File Systems

$WORK

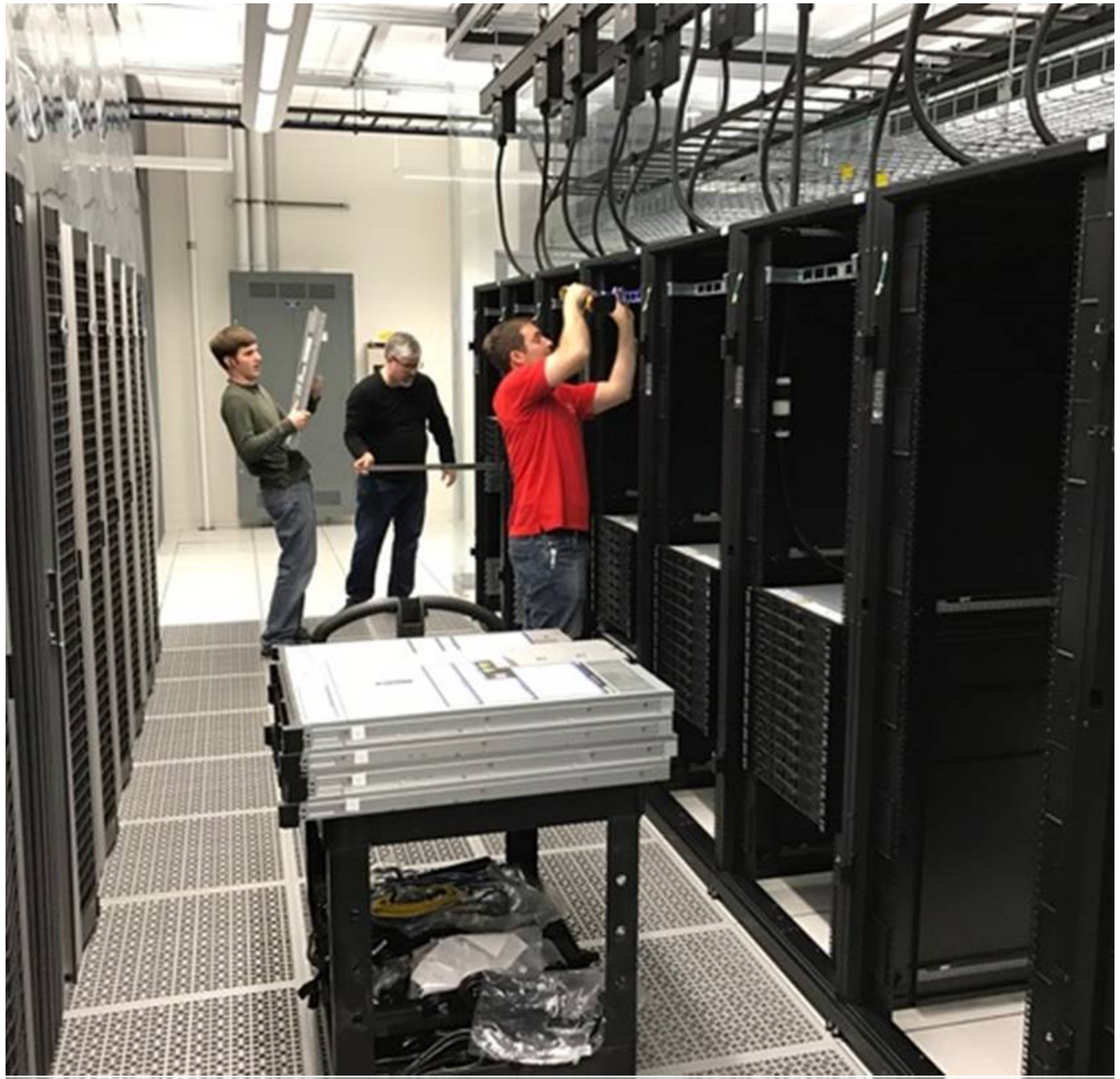
- NFS file system, located at /work/users/$USER
- $WORK space is restricted by quotas (8 TB)
- Intended for long term storage

ManeFrame File Systems

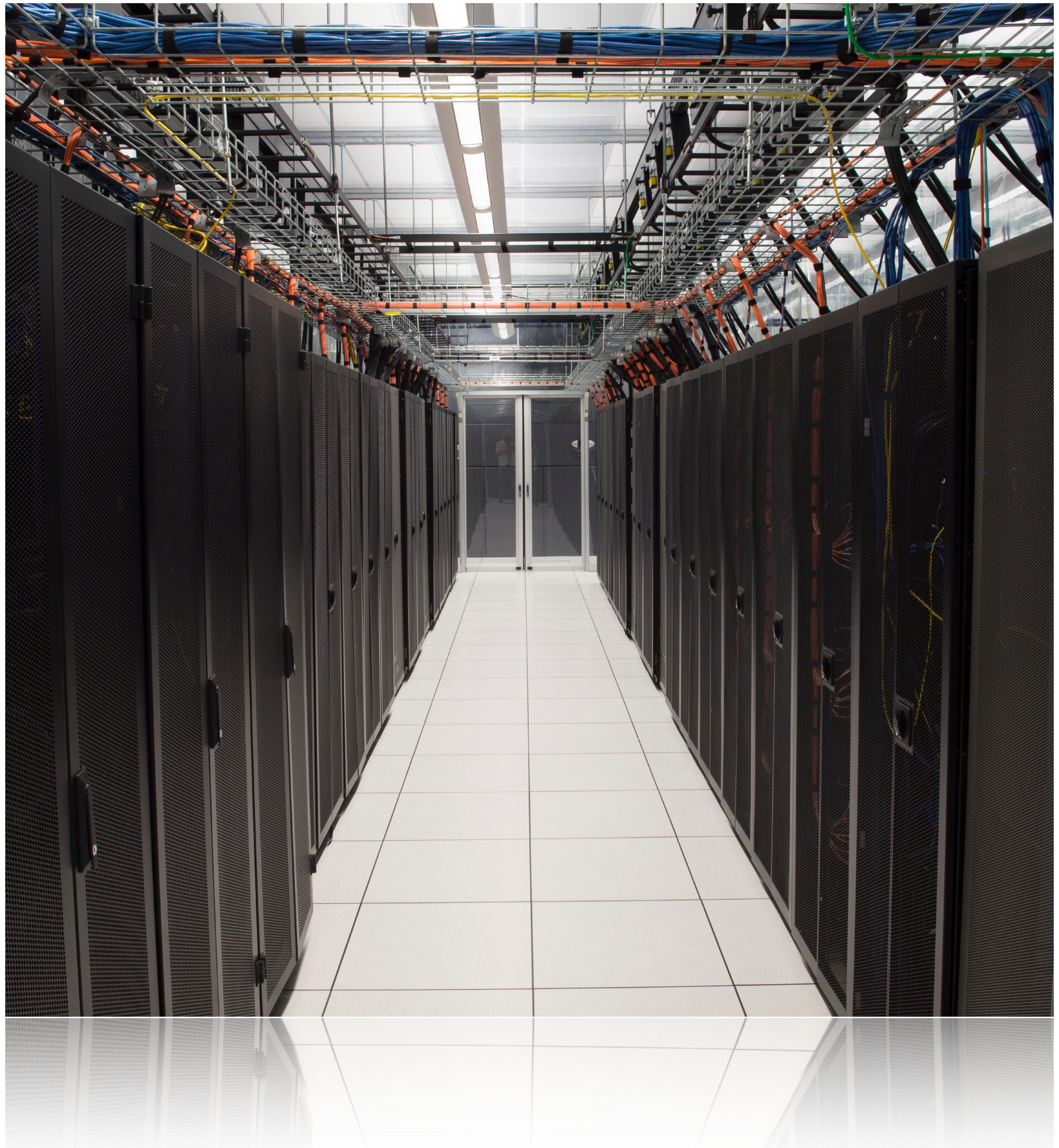
- $SCRATCH
  - 2.8 PB of scratch space is located in /scratch/users/ (e.g. /scratch/users/$USER)
  - High performance parallel Lustre file system
    - As a high-performance parallel filesystem, Lustre will not perform well if misused
  - $SCRATCH is a volatile file system
    - No guarantee that any of the files stored in $SCRATCH can be retrieved or restored in the event of an accidental delete, loss or failure of the filesystem

University Data Center (UDC)

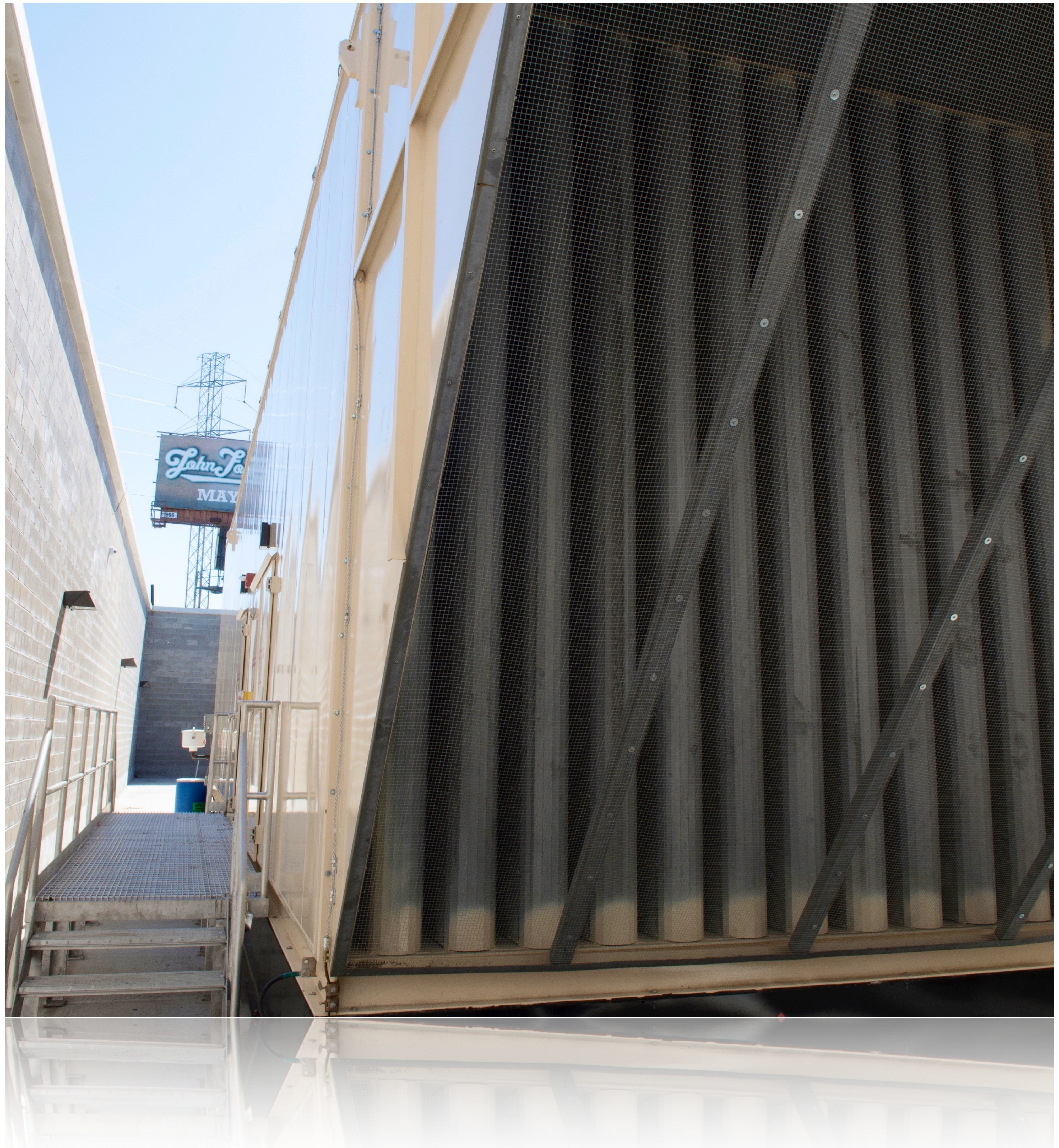
ManeFrame II



ManeFrame Hot Aisle



Cooling and Backup Generator



Accessing ManeFrame II

Accessing ManeFrame II

- SSH (secure shell)
    - Provides command line access to the login nodes
    - Graphical applications require X11-forwarding
    - M2 has five login nodes all accessible via m2.smu.edu

- HPC Portal
    - Provides web-based access to M2 resources
        - Web-based terminal access
        - JupyterLab (Notebooks)
        - RStudio
        - Remote Desktops

SSH & X11

- SSH
  - Linux, macOS, Windows, WSL is installed by default
  - Available via terminal applications

- X11
  - Linux
    - Installed by default
  - macOS
    - Apple provides X11 via the Xquartz open source project
    - https://www.xquartz.org
  - Windows
    - Xming
      - www.straightrunning.com/XmingNotes/

SSH and X11 for Windows

- As Windows is not a UNIX-like operating system standard utilities are not available by default
- Many SSH clients available with various features and styles
  - Putty
  - WinSCP
  - Tera Term
- X11 is available via Xming
- UNIX-like emulated environments
  - Ubuntu for Windows
    - Only on the most recent version of Windows 10
    - X11 not provided, but can use Xming
  - Cygwin
    - Can provide both SSH and X11

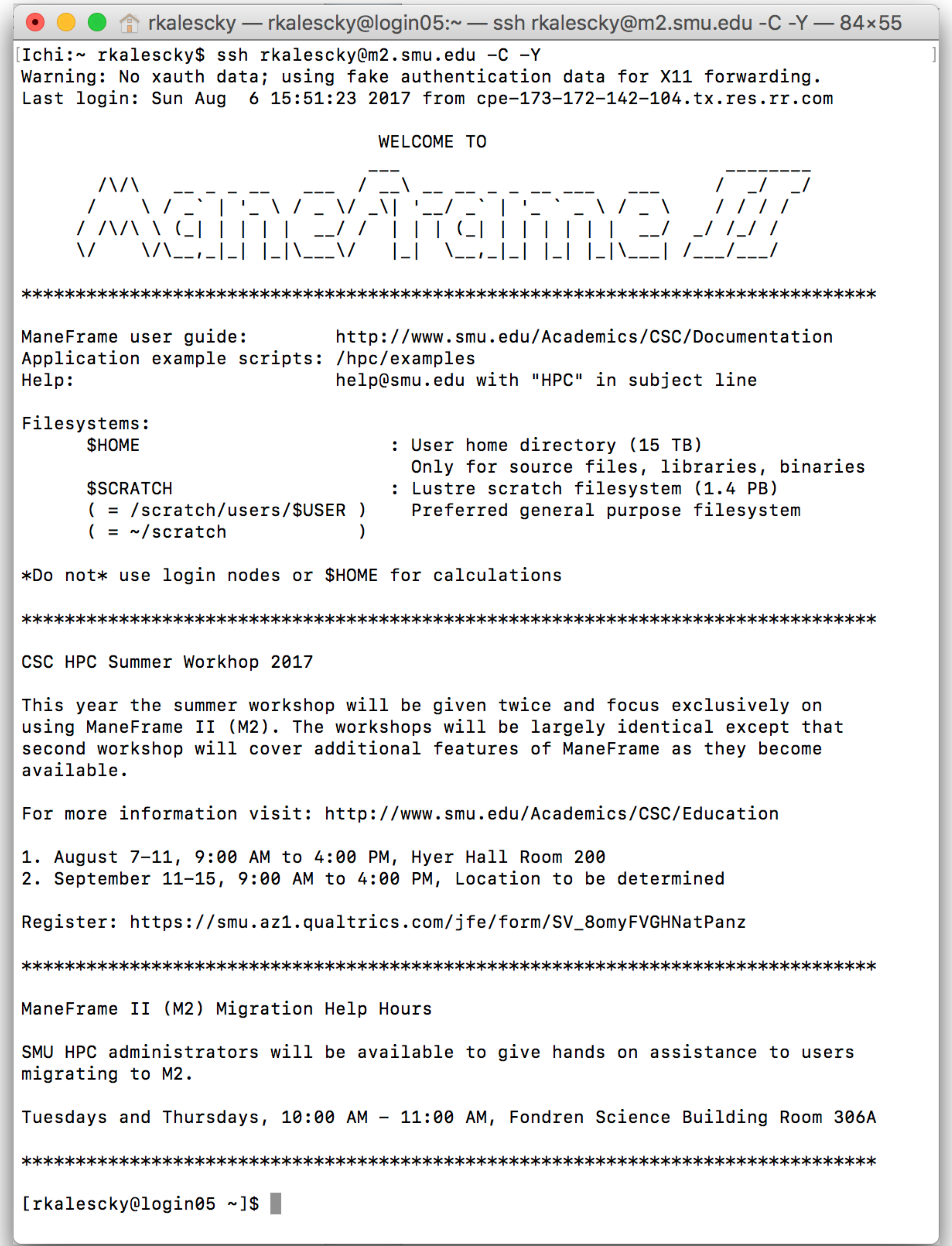
1. Open terminal application

2. At the command prompt, enter “ssh username@m2.smu.edu -C -X”

    - “username” is your ManeFrame username

    - “-C” requests compression, which can make the connection faster

    - “-X” requests X11 forwarding, which allows graphical programs to run from ManeFrame, but display on your computer



Xming and Putty Installation Notes

- Default installation options are fine for these installations

1. Install Xming
   - Upon initial launch Windows firewall may ask for permission, select allow access
2. Install Xming-fonts
3. Install Putty

2

1

3

Releases

Links and release status

Website Releases	Version	State/Notes	Released	MD5 signatures	Size MB
Xming Xming-x64	7.7.0.10	Website Release	14 Aug 2016	MD5 signatures	4.93 5.21
Xming-portablePuTTY Xming-portablePuTTY-x64	7.7.0.10	Website Release	9 Aug 2016	MD5 signatures	1.82 1.90
See Donations for how to obtain a Donor Password.					

Public Domain Releases	Version	State/Notes	Released	MD5 signature	Size MB
Xming-fonts	7.7.0.10	Public Domain	9 Aug 2016	ed1a0ab53688615bfec88ab399ae5470	31.1
Xming Xming-mesa	6.9.0.31	Public Domain	4 May 2007	4cd12b9bec0ae19b95584650bbaf534a e580debbf6110cfc4d8fcd20beb541c1	2.10 2.50

Website Snapshots	Version	State/Notes	Snapshot	MD5 signature	Size MB
Snapshot Xming Snapshot Xming-x64	7.7.0.11	Work in progress	7 Sep 2016 08:25	Not yet released	4.94 5.22
See Donations for how to obtain a Donor Password.					

Binaries

The latest release version (beta 0.67)

This will generally be a version we think is reasonably likely to work well. If you have a probl

For Windows on Intel x86

PuTTY:	putty.exe	(or by FTP)	(signature)
PuTTYtel:	puttytel.exe	(or by FTP)	(signature)
PSCP:	pscp.exe	(or by FTP)	(signature)
PSFTP:	psftp.exe	(or by FTP)	(signature)
Plink:	plink.exe	(or by FTP)	(signature)
Pageant:	pageant.exe	(or by FTP)	(signature)
PuTTYgen:	puttygen.exe	(or by FTP)	(signature)

A ZIP file containing all the binaries (except PuTTYtel), and also the help files

Zip file:	putty.zip	(or by FTP)	(signature)

A Windows MSI installer package for everything except PuTTYtel

Installer:	putty-0.67-installer.msi	(or by FTP)	(signature)

Legacy Inno Setup installer. Reportedly insecure! Use with caution, if the MSI fails.

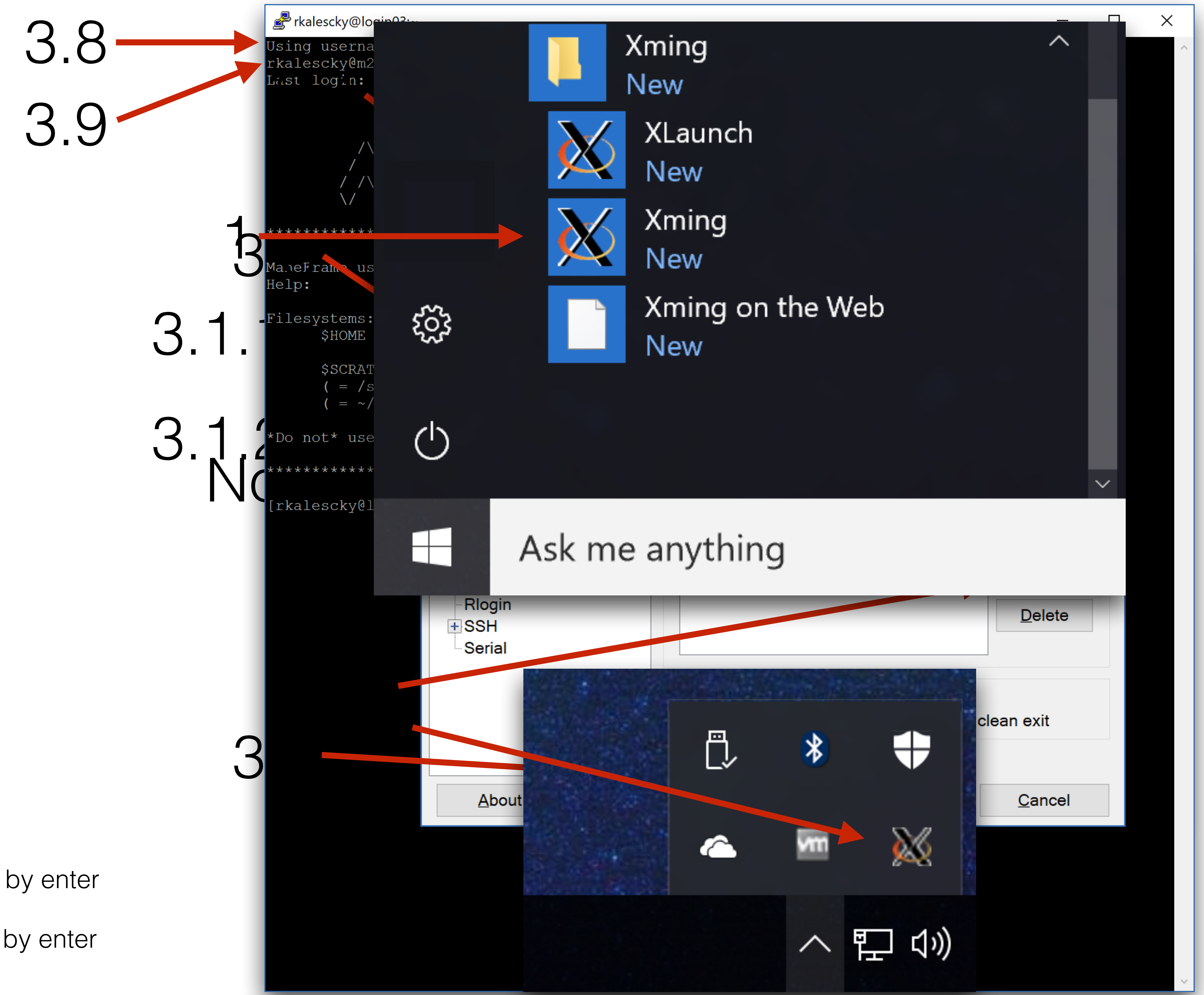
Legacy installer:	putty-0.67-installer.exe	(or by FTP)	(signature)

Checksums for all the above files

MD5:	md5sums	(or by FTP)	(signature)
SHA-1:	sha1sums	(or by FTP)	(signature)
SHA-256:	sha256sums	(or by FTP)	(signature)
SHA-512:	sha512sums	(or by FTP)	(signature)

Accessing ManeFrame II: Putty

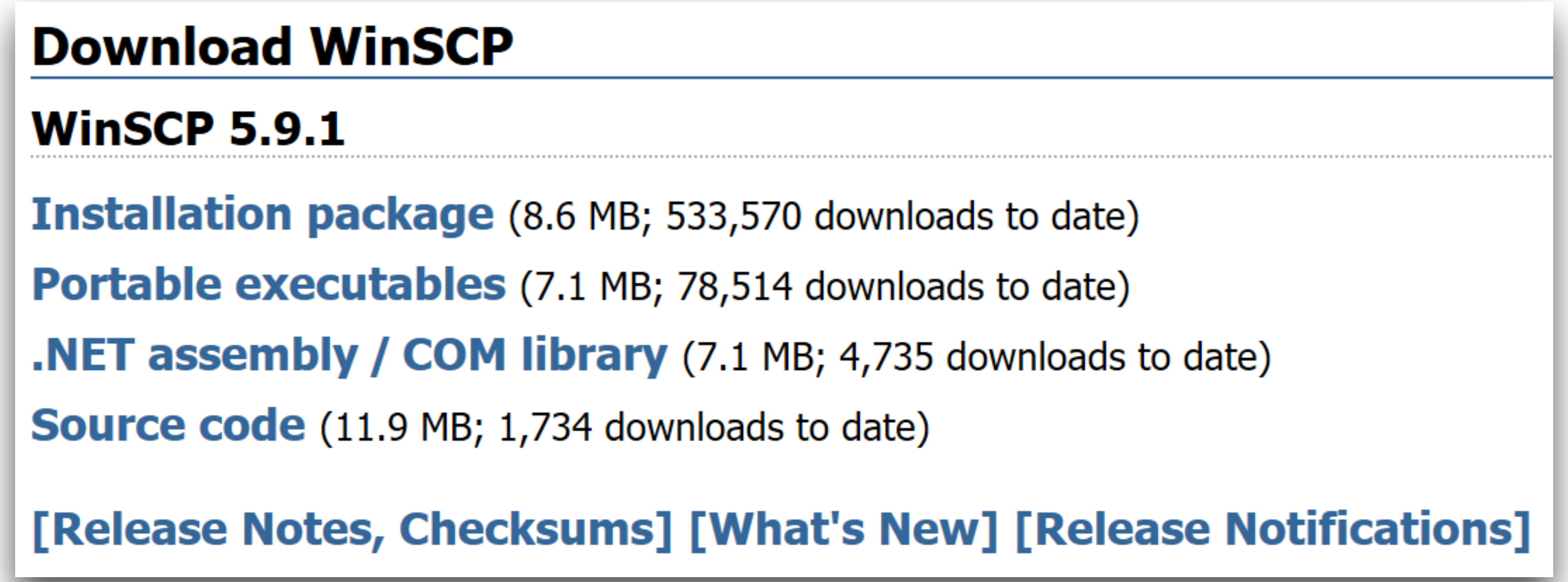
1. Start the Xming program in Start Menu
    - The Xming program runs from the System Tray
2. Start the Putty program in the Start Menu
3. In the Putty window:
    1. In the “Category:” box:
        1. Scroll to the bottom and select “+” next to “SSH”
        2. Select “X11”
    2. In the “X11 forwarding” section, select “Enable X11 forwarding”
    3. In the “Category:” box, scroll to the top and select “Session”
    4. In the “Host Name” field, type “m2.smu.edu”
    5. In the “Saved Sessions” field, type “ManeFrame II (M2)”
    6. Press “Save”
    7. Press “Open”
        - Select “Yes” if you presented with a “Putty Security Alert” window
    8. At the command prompt, type your ManeFrame username followed by enter
    9. At the command prompt, type your ManeFrame password followed by enter

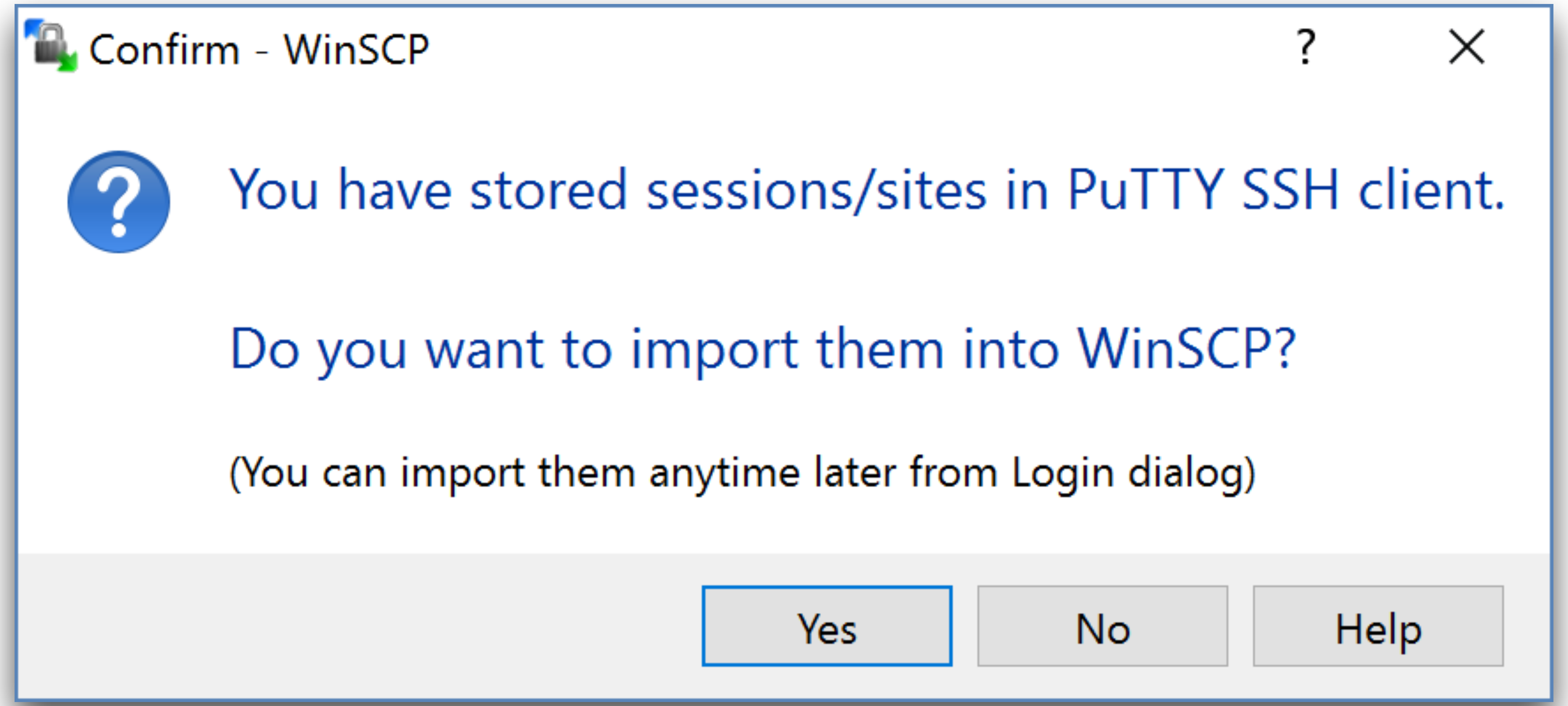


Transferring Files

- It is easy to transfer files to and from ManeFrame II using an SFTP client
- Web
  - HPC Portal
- macOS
  - Cyberduck
  - Transmit
- Windows
  - WinSCP
  - Cyberduck
- UNIX-(like) systems also have command line options available

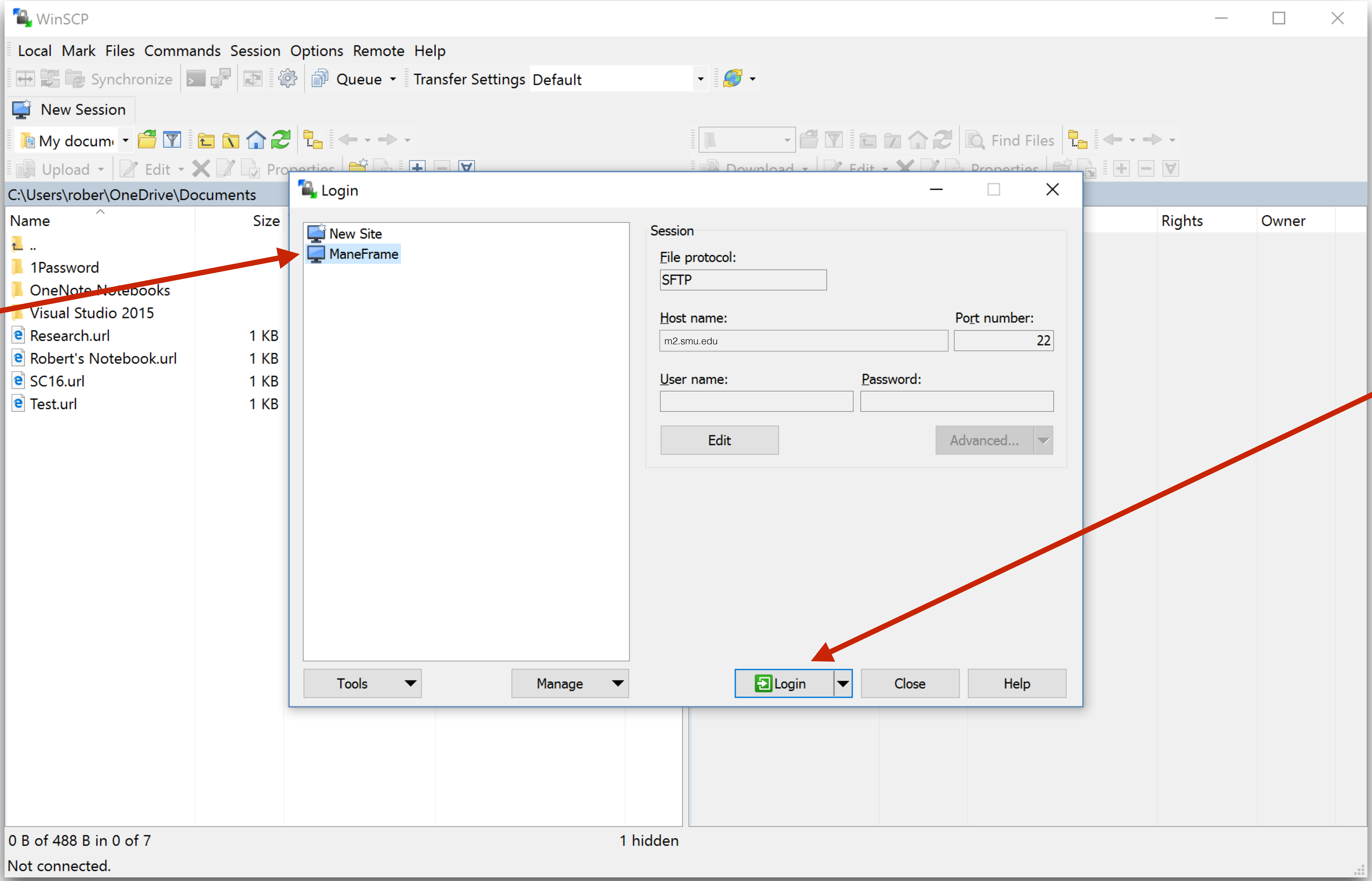
WinSCP Installation





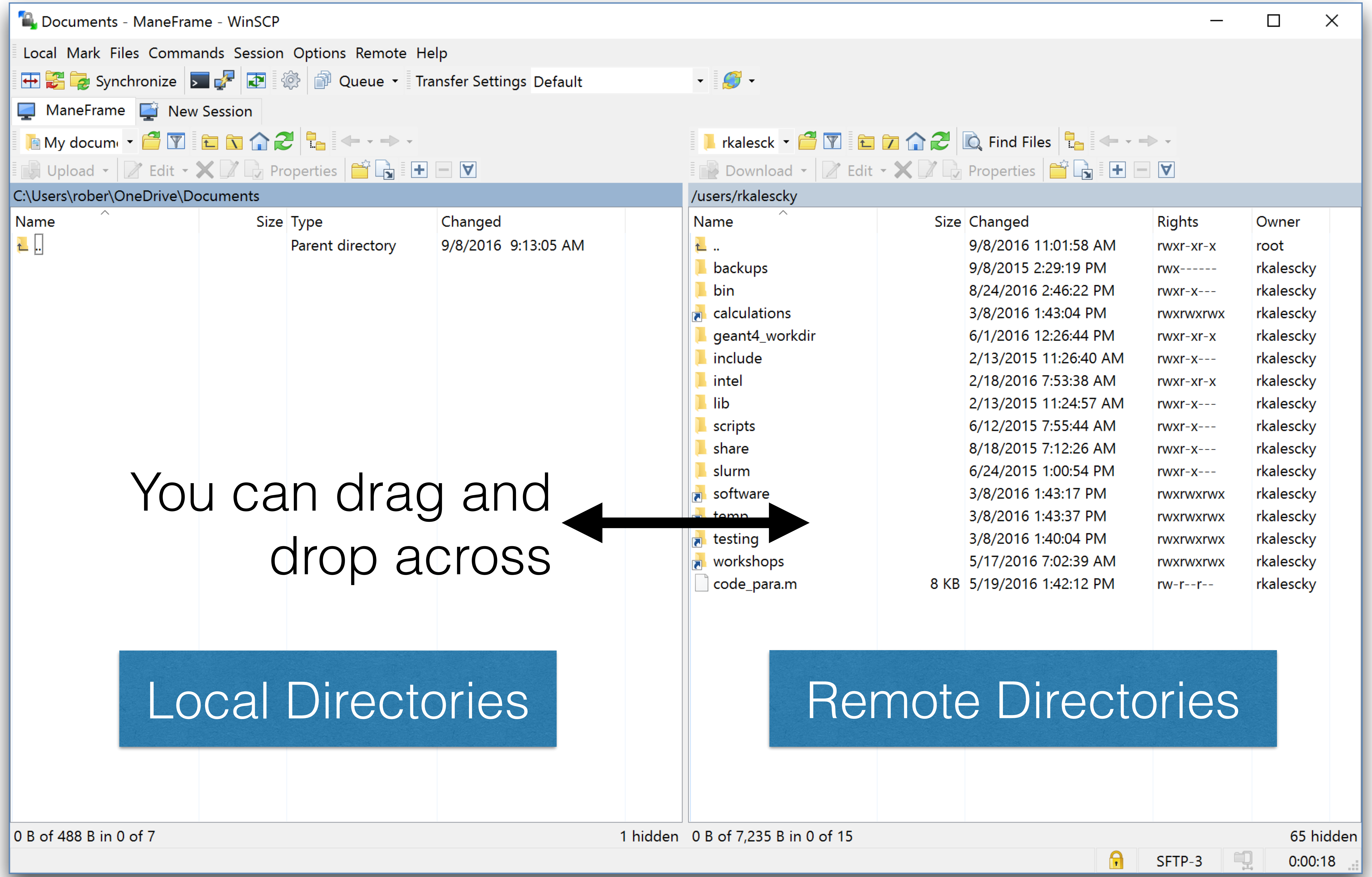
Using WinSCP

Can be imported from Putty



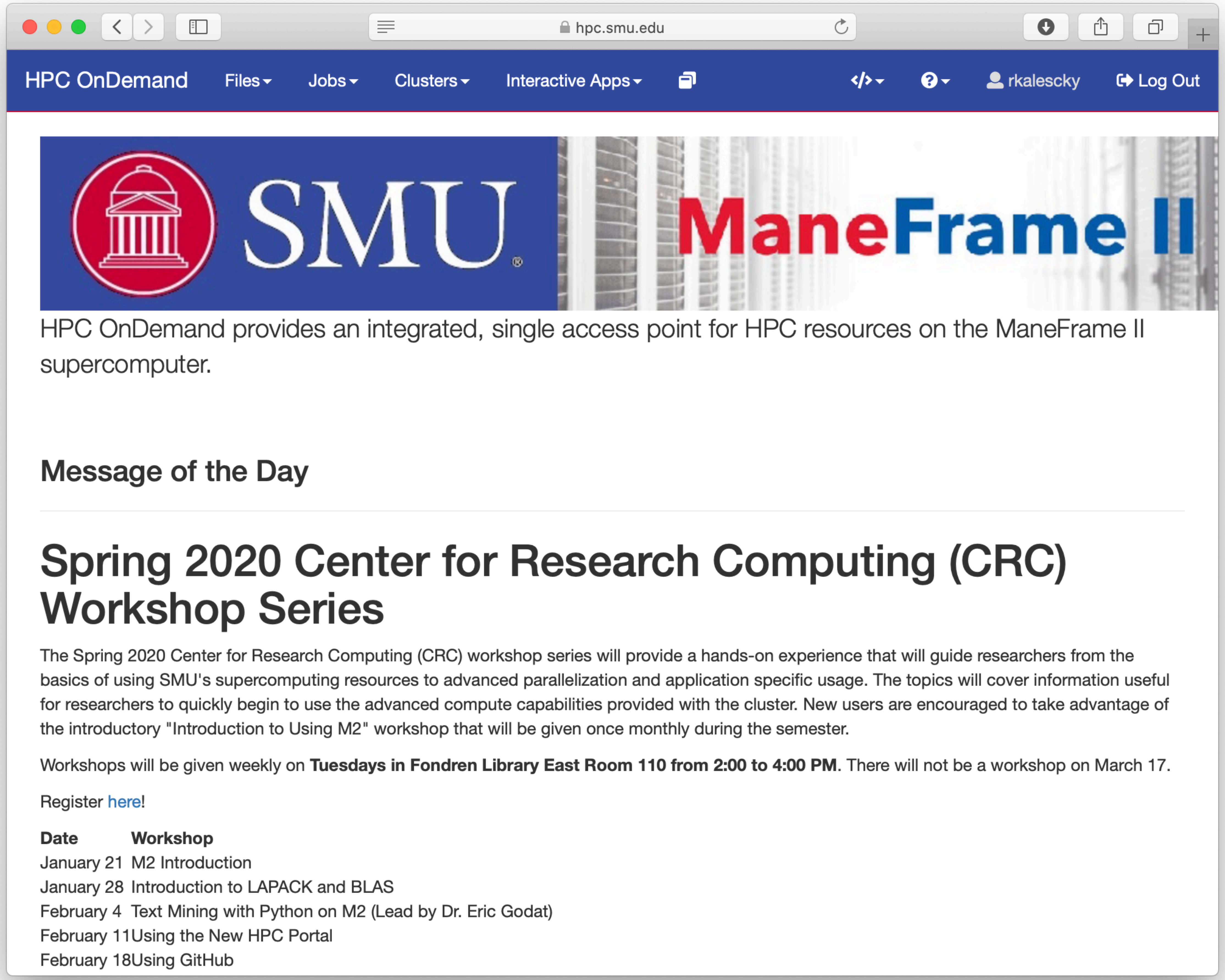
You will be prompted for your password

Using WinSCP



HPC Portal: hpc.smu.edu

- Provides an integrated, single access point for HPC resources on the ManeFrame II (M2) supercomputer.
- Accessing the Portal
  - Access to the HPC portal requires an existing M2 account.
  - Go to hpc.smu.edu.
  - Sign in using your SMU ID and SMU password
- Documentation
  - http://s2.smu.edu/hpc/documentation/portal.html



Using ManeFrame II

UNIX Tutorial

- UNIX tutorial
  - http://s2.smu.edu/hpc/documentation/unix.html

Accessing Applications and Tools

Software Available on M2

```
----------------------------------------------------------------------------- /hpc/modules/applications -----------------------------------------------------------------------------
   abcluster/2.0                  compliance/3.0.2            geant4/10.06.p01           mesa/r12115             psi4/1.3.2-1           (D)    singularity/3.0.3    (D)
   aimall/17.01.25                comsol/5.2a                 gmt/5.4.3                  molcas/8.2              python/2                      singularity/3.5.3
   amber/16.15             (D)    comsol/5.4                  gnuplot/5.2.2              molpro/2015.1.24        python/3               (D)    snap/7.0.2
   amber/16/cpu                   comsol/5.5           (D)    gromacs/2019.4             molpro/2018.1.1  (D)    q-chem/5.0                    snid/5.0
   amber/16/gpu                   crystal/17/1.0.2            gurobi/8.0.1               mplus/8                 q-chem/5.1             (D)    spack
   ansys/18.2/electronics         ctioga2/0.14.1              hdfview/3.0.0              mplus/8.2        (D)    q-chem/5.2.1                  star/2.5.3a
   ansys/19.1/electronics         cuda/10.0                   idl/8.0                    namd/2.12/cpu    (D)    quantum-espresso/6.4.1        stata/mp-14
   ansys/19.1/workbench    (D)    cudnn/7.6.4                 jags/4.3.0                 namd/2.12/gpu           r/3.4.0                       stata/mp-15          (D)
   arrayfire/3.5.0                demon/6.0.2                 julia/0.5.2                namd/2.13_cpu           r/3.4.1                       tcad/2019
   autodock_vina/1.1.2            demon_nano/2019-06-24       julia/0.6.0                namd/2.13_gpu           r/3.4.1-1              (D)    tensorflow/1.2_gpu
   awscli/2                       ds9/7.5                     julia/1.0.1         (D)    nbo/6.0                 r/3.4.4                       tensorflow/1.4_gpu
   babel/2.4.1                    firefox/76                  knitro/11.1.0              nccl/2.4.8              r/3.5.2                       tensorflow/1.10_gpu
   bazel/0.26.1                   gamess/2017r1               lammps/aug17               omnia/1                 r/3.6.2                       tensorflow/2.0_gpu   (D)
   bazel/0.29.1            (D)    gams/25.1.2                 lammps/2016.11.17   (D)    openmm/7.2              rstudio/1.1.383               tensorrt/5.1.5
   bedtools2/2.26.0               gauss/17.0.3                macaulay2/1.11             orca/4.0.0              rstudio/1.1.423               texlive/2017
   borg/1.8                       gaussian/g03e               madgraph/2.6.3.2           orca/4.0.1       (D)    rstudio/1.1.442        (D)    valgrind/3.12.0
   cfour/2.00beta                 gaussian/g09d               mallet/2.0.8               orca/4.1.1              samtools/1.6                  vasp/5.4.4/cpu
   cfour/2.1               (D)    gaussian/g16a               mathematica/11.1.1         paraview/5.4.1          sapt/2016.1                   vasp/5.4.4/gpu       (D)
   charmm/c41b1                   gaussian/g16b               matlab/r2016b              period04/1.2.0          sas/9.4.update6               vmd/1.9.3/egl
   clocksim/2014                  gaussian/g16c        (D)    matlab/r2017a              periodic_nbo/1.1        sas/9.4                (D)    vmd/1.9.3/ogl        (D)
   cmake/3.11.1                   gaussview/6.0.16            matlab/r2018b       (D)    psi4/1.3.2              sextractor/2.19.5

------------------------------------------------------------------------------- /hpc/modules/compilers -------------------------------------------------------------------------------
   gcc-4.8.5    gcc-7.3    intel-17.0.4    intel-2020.0    pgi-12.10    pgi-16.5    pgi-17.7     pgi-18.4
   gcc-6.3      gcc-9.2    intel-18.0.1    llvm-4.0.0      pgi-13.2     pgi-17.4    pgi-18.10    pgi-19.5
```

Basic Module Commands

Command	Description	Example Usage
module avail	List available modules	`module avail`
module load	Load modules	`module load python`
module list	List loaded modules	`module list`
module unload	Unload modules	`module unload gcc`
module purge	Unload all modules	`module purge`
module spider	Search for modules matching text	`module spider python`
module save	Save loaded modules	`module save`

Module Tutorial

http://s2.smu.edu/hpc/documentation/modules.html

Submitting Jobs Using Slurm

Common Classes of Applications

- Serial
  - Users run short-lived single-threaded calculations in batches
  - Not uncommon to see batches of greater than 10K
  - Batches can be submitted individually, via SLURM’s array feature, or bundled to use a whole node
- Shared-Memory
  - Users run long-lived computationally intensive jobs with various memory requirements
  - Not uncommon to see run times of several months
- Distributed-Memory
  - Users run computationally intensive jobs that span nodes
  - Not uncommon to see hundreds of nodes used for single job

Common Classes of Applications

- Serial
    - Simple calculations frequently done *en masse*
    - Usually done to get statistics over many trials
    - R, MATLAB, Python, ROOT, Autodock
- Shared-Memory
    - Complex calculations with demanding compute requirements
    - Quantum chemistry and molecular dynamics packages (e.g. CFOUR, Gaussian, NAMD, CHARMM)
- Distributed-Memory
    - Distributed programs usually via MPI
    - Many parallel computing research codes
    - Some packages such as CFOUR, CHARMM, LAMMPS

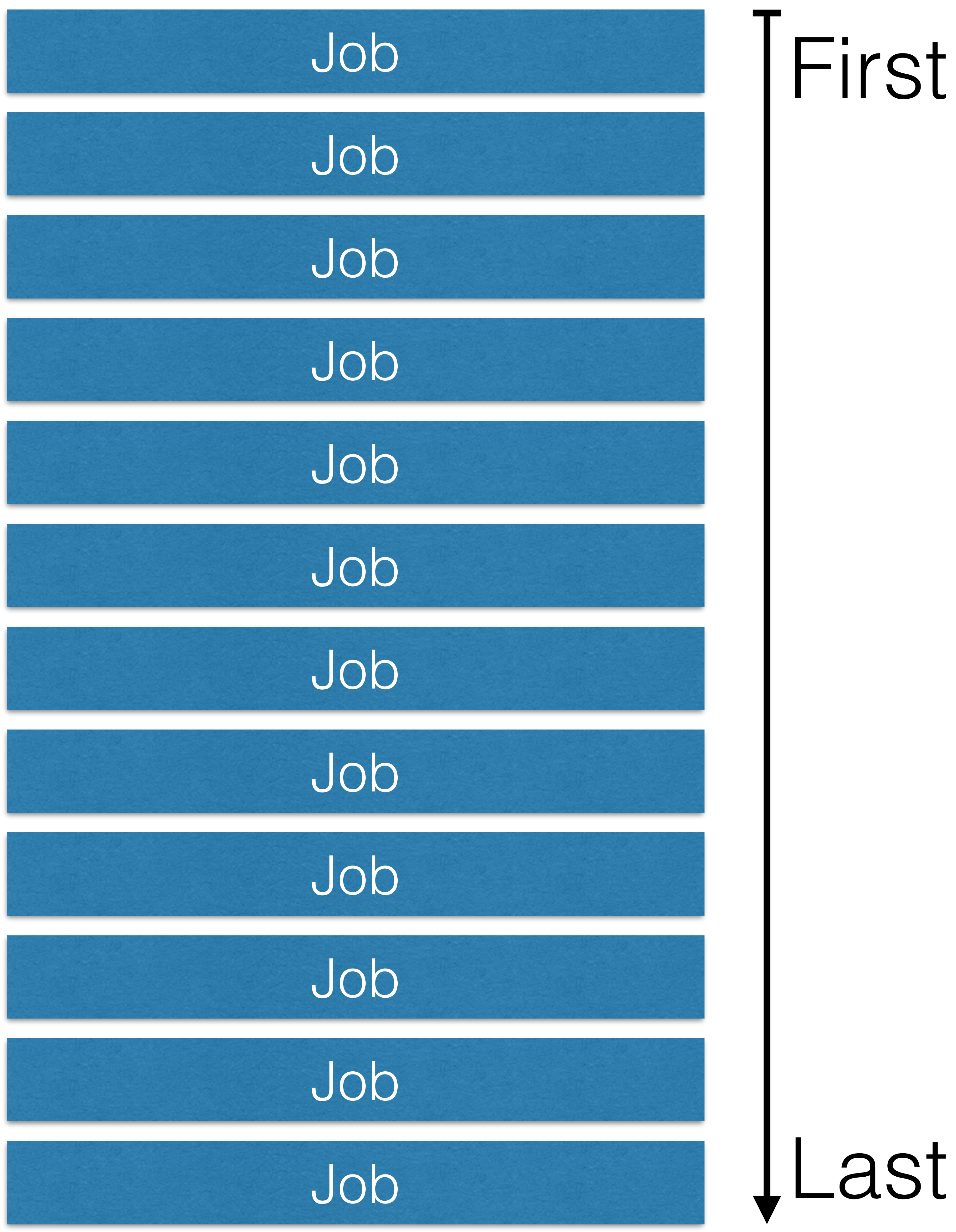
Resource Management

- Compute resources are shared by all users

- Queue systems manage access to specific resources

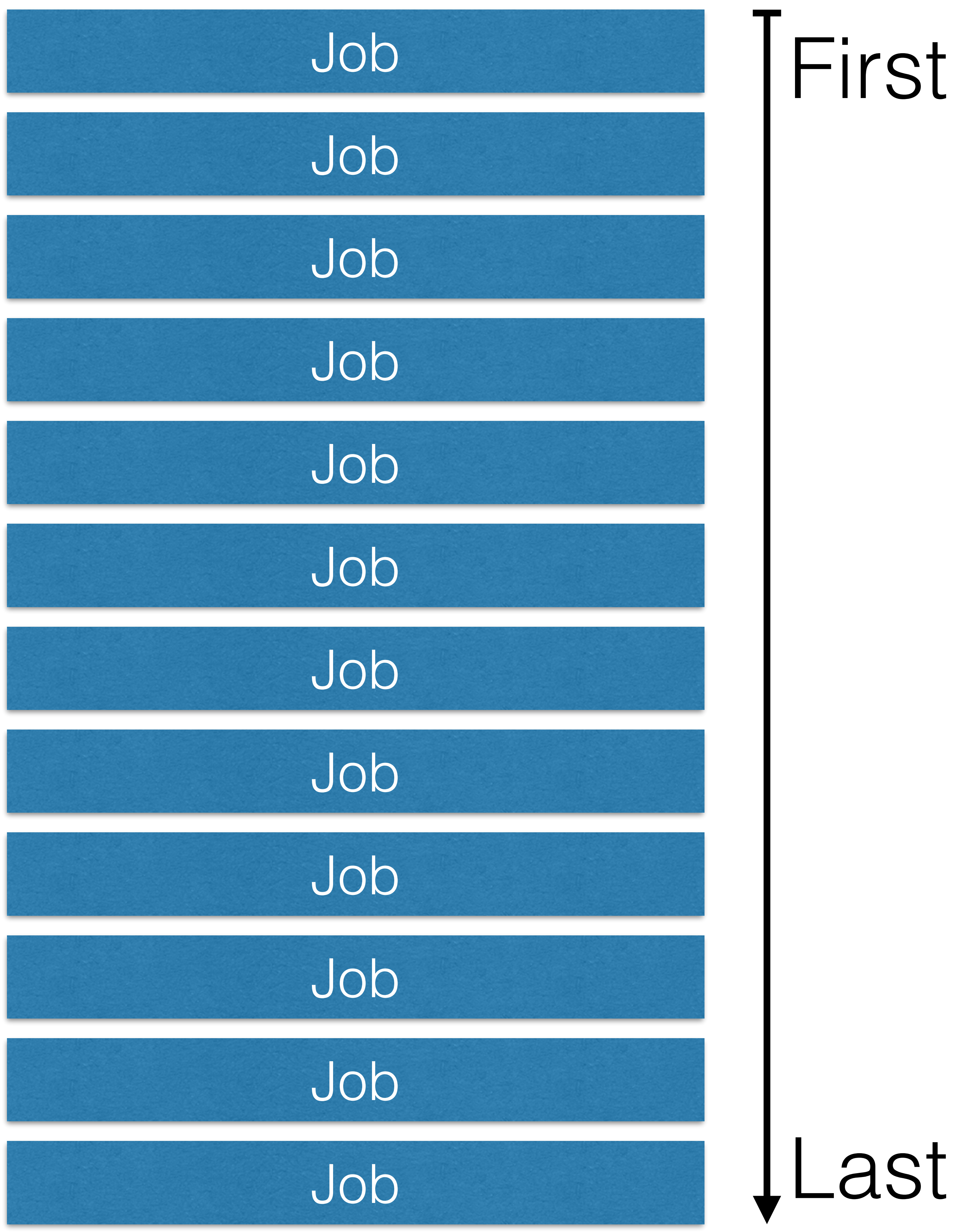
- There are many queue systems around: Moab, Torque, LoadLeveler, Condor, Oracle Grid Engine, Argent Job Scheduler, Platform LSF, etc.

- ManeFrame uses Simple Linux Utility for Resource Management (SLURM)



Resource Management

- ManeFrame resource allocation is governed by a fair-share factor
  - Preference is given to those who have submitted the fewest jobs
  - Prevents a single user from “blocking” subsequent job submissions of other users
- Allocations can be requested interactively or as a batch job
- Queues (or partitions) are used to distinguish resources by type and the maximum run time



ManeFrame II Partitions (Queues)

Queue	Quantity [Nodes]	Exclusivity	Duration
development	2	Partial	2 hours
htc	52	Shared	1 day
standard-mem-s	80	Exclusive	1 day
standard-mem-m	24	Exclusive	1 week
standard-mem-l	18	Exclusive	1 month
medium-mem-1-s	20	Exclusive	1 day
medium-mem-1-m	10	Exclusive	1 week
medium-mem-1-l	5	Exclusive	1 month
medium-mem-2	4	Exclusive	2 weeks
high-mem-1	5	Exclusive	2 weeks
high-mem-2	6	Exclusive	2 weeks
mic	36	Exclusive	1 week
gpgpu-1	36	Exclusive	1 week
v100x8	3	Shared	1 week
dcv	5	Exclusive	1 day

Basic SLURM Commands

Command	Description	Example Usage
sinfo	Displays partition and node state information	`sinfo`
squeue	Displays queue state information	`squeue -u $USER`
sbatch	Submits batch script	`sbatch job.submit`
srun	Requests resources for interactive use	`srun -p development --exclusive --pty $SHELL`
scancel	Cancel jobs	`scancel 12345678`

Need help or have questions?

rkalescky@smu.edu