Data Transfer Basics and Best Practices

Joon Kim, Ph.D.

PICSciE, Research Computing

Research Computing Bootcamp 2018

10/29/2018 – 11/01/2018

Why do we care?

Without good practice, you will waste time and effort

Time

1. Start data transfer using SCP at 10pm. Usually takes 10 hours.
2. At 2am, there was a brief 1-minute network outage. Transfer job aborted.
3. Arrive 8am in the morning. See the damage. Start again, which will take 10 hours.
4. **Lost a day of work.**

Effort

Why do we care?

**Without good practice,
you will waste time and effort**

1. Start data transfer using SCP at 10pm. Usually takes 10 hours.

Is that really the best?

Time

Effort

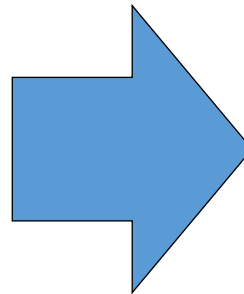**Focus on your research,
not on transferring data**

Time

Effort

Use case 1

- I have data at **Argonne National Lab** that I want to process & analyze at **Princeton HPC clusters**
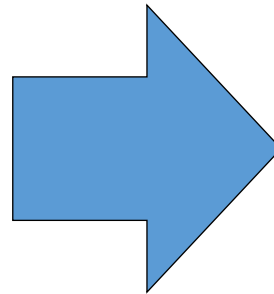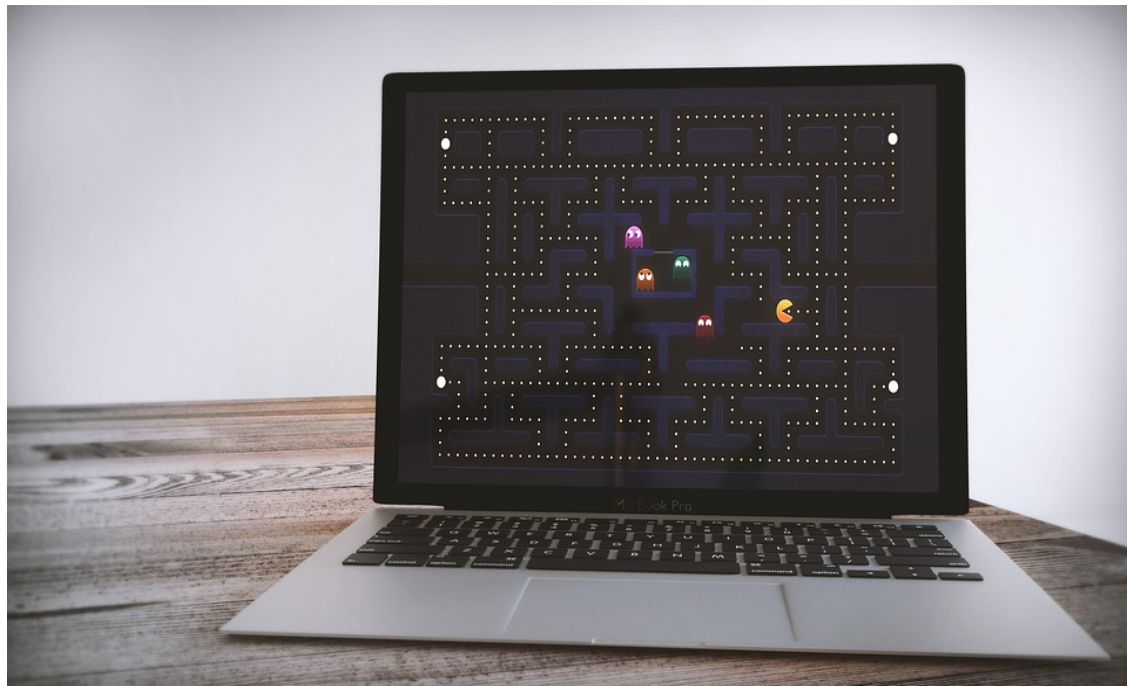
Argonne National Lab

Princeton RC HPC Clusters

Use case 2

- I have data on my **workstation/laptop** that I want to process & analyze at **Princeton HPC clusters**

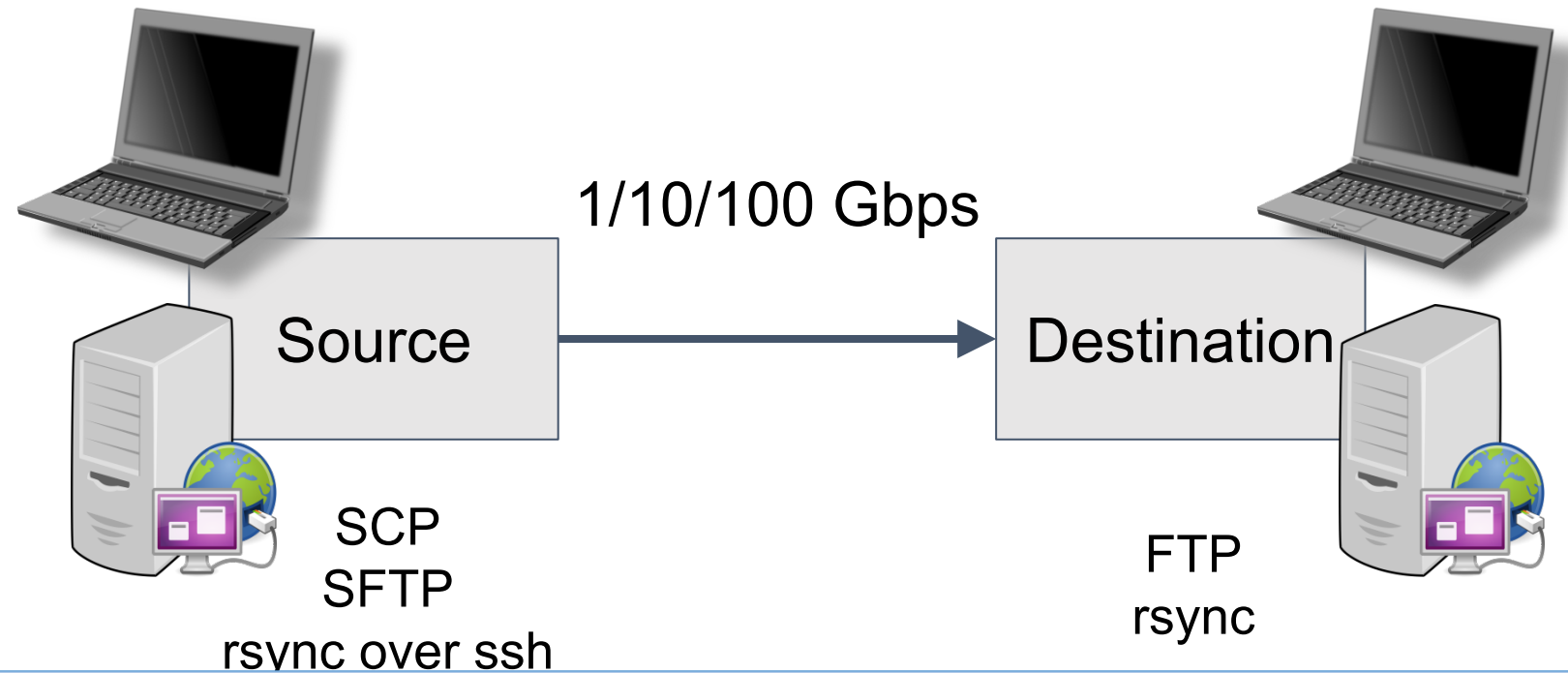workstation/laptop

Princeton RC HPC Clusters

Data Transfer Basics

Data transfer: Overview

Three key elements

- Endpoints

- Network

- Transfer tool

1/10/100 Gbps
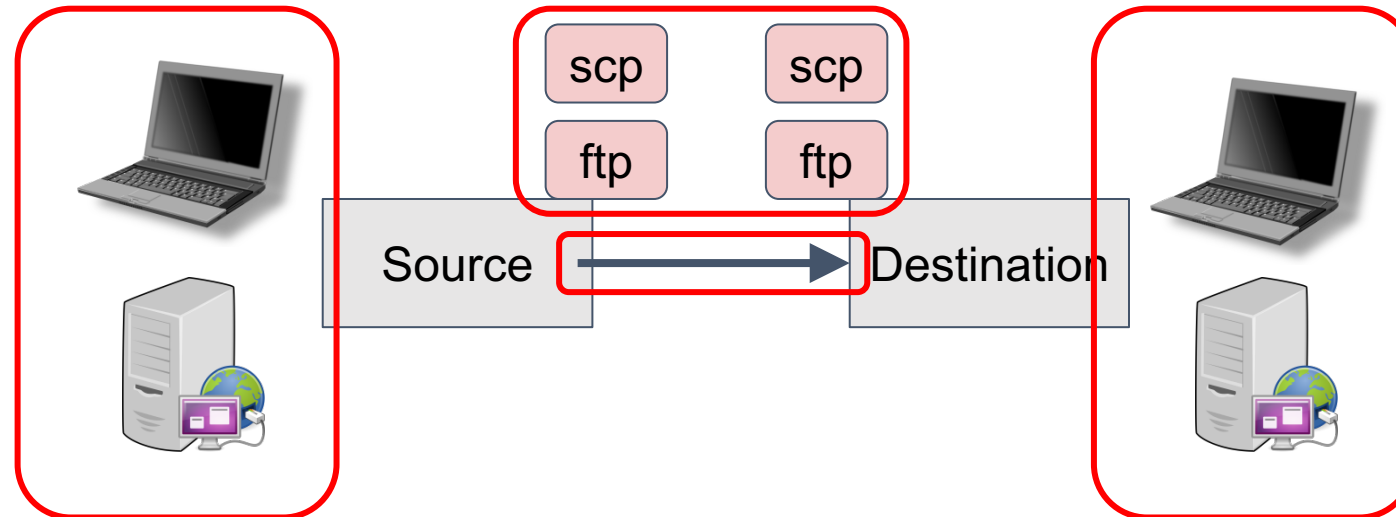
Source

Destination

SCP
SFTP
rsync over ssh

FTP
rsync

These will determine **how you transfer data** and **how fast it will be**

Why is my data transfer slow?

Where **are** the **bottlenecks**?

scp   scp

ftp   ftp

Source ———→ Destination
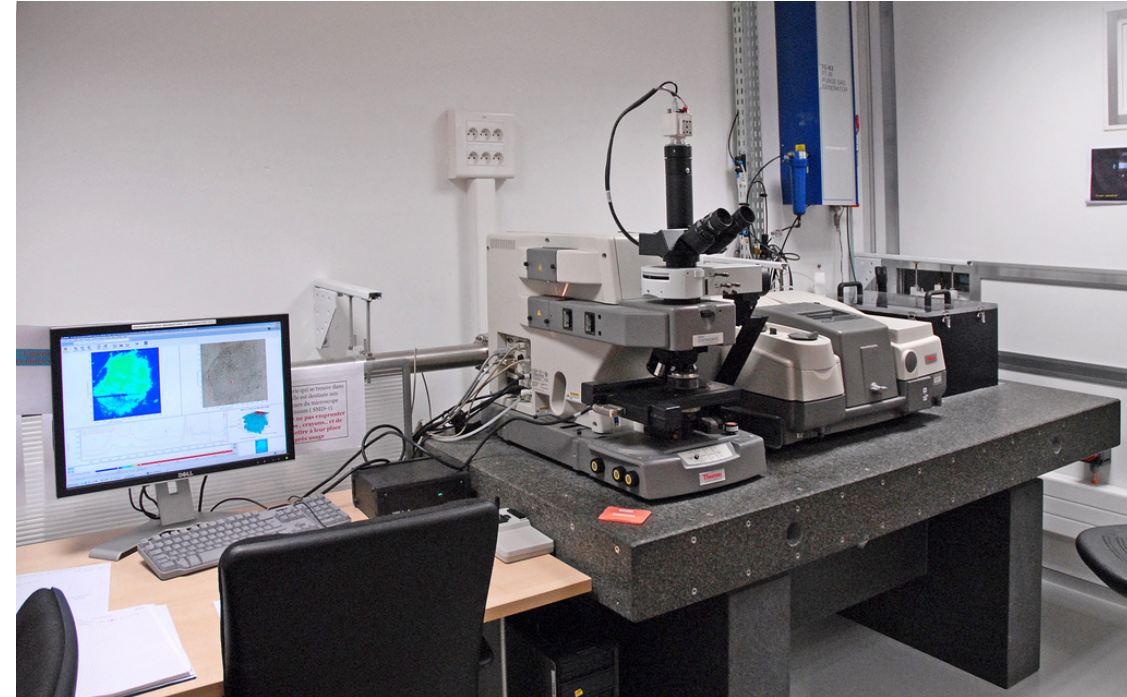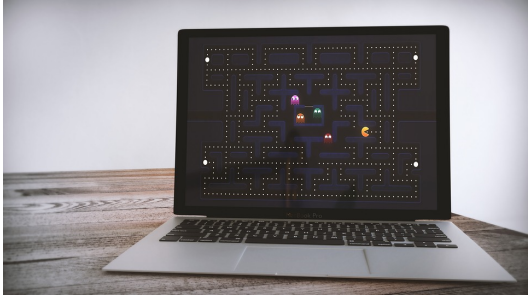
e.g., Server machine with 10 Gb/s connection
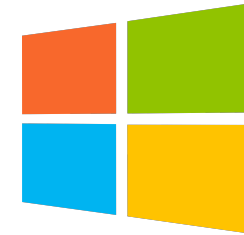**will not guarantee 10 Gb/s transfer speed**

Endpoint Examples

SMB file mount
NFS file mount

Z: drive
\\plabs.princeton.edu\data

1. Endpoints

- ## Operating system
  - ### Determines your interface and tool availability

- ## Resources
  - ### CPU: Higher clock speed > Number of cores
  - ### RAM: **> 32 GB** for a dedicated Data Transfer Node (DTN)
  - ### Disk I/O
    - #### Disk type (SATA, SSD), configuration (RAID), file system (ext4, GPFS)
    - #### Decent server with SATA, ext4, RAID, good transfer tool: ~ 4 Gb/s (500 MB/s)

You will likely get less than 1 Gb/s (125 MB/s) with your laptop, most desktops, and un-optimized servers

Example Data Transfer Node endpoint

Hardware description

- Motherboard: SuperMicro X9DRi-F

- CPU: 2 x Intel(R) Xeon Ivy Bridge E5-2643V2 3.5GHz 6 Cores (Total 12 Cores)

- Memory: 96G ((12) 8GB DDR3-1866MHz RAM ECC/REG)

- RAID: Adaptec ASR-81605ZQ (16 ports)

- 16 x Western Digital, model WD2500BHTZ-0

- Network Controller:
    - 10G: Intel X520, Myricom 10G-PCIE2-8C2-2S, Chelsio T5
    - 40G: Mellanox MCX312A-XCBT
    - 100G: Mellanox MCX455A-ECAT

System Configuration

We use the most recent CentOS-7 distribution of Linux, and have configured the data drives as RAID6

Performance Results for this configuration

Back-to-Back Testing using GridFTP

- memory to memory, 1 10GE NIC: 9.9 Gbps

- memory to memory, 4 10GE NICs: 39.5 Gbps

- disk to disk: 9.2 Gbps (1.2 GBytes/sec) using a single large file

- Endpoint system resource & configuration significantly affect transfer speed
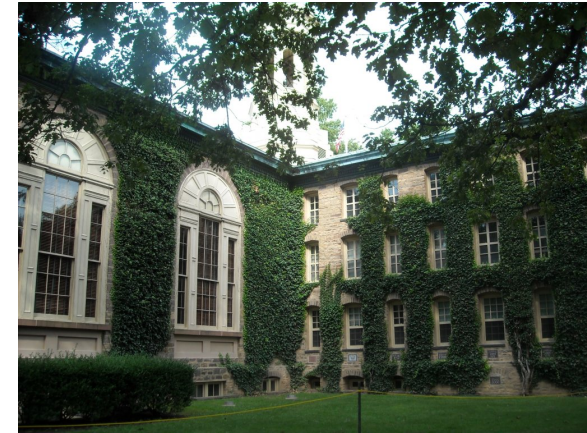
- Disk I/O can be a bottleneck

(ref: http://fasterdata.es.net/)

OFFICE of INFORMATION TECHNOLOGY

PICSciE

2. Network

- Network bandwidth
  - **Use wired connection when available**
  - **Get a good network card (NIC)**

10-100 Mbps

Wired: 1, 10, 40 Gbps
Wireless: < 100 Mbps

- Congestion (heavily dependent on time of day)
- Distance and latency
- "Things" along the way
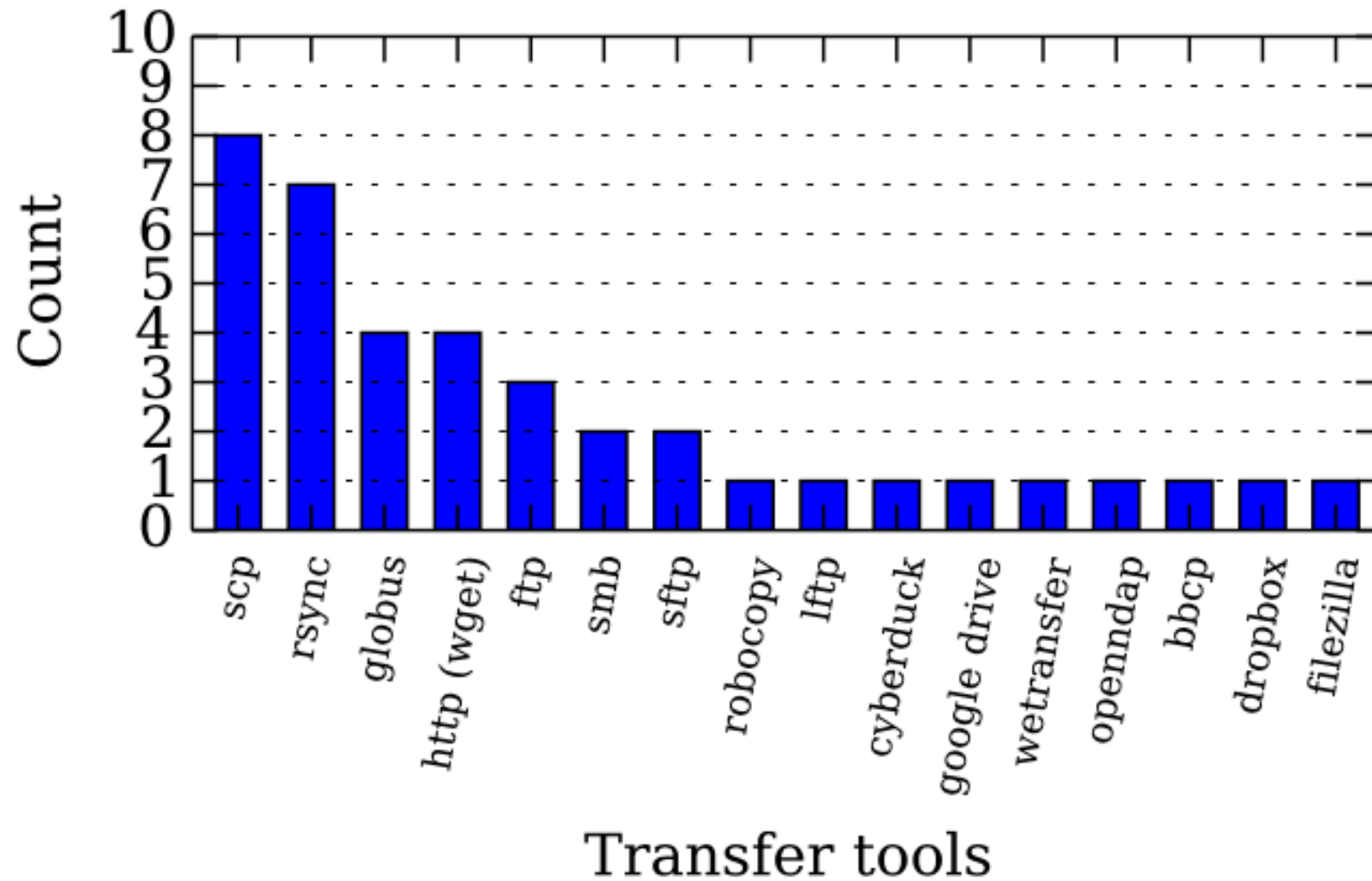 - Routers, switches, firewalls, NAT, security devices, …

Things you don't have much control over

What transfer tool do you use?

Secure Copy (SCP)

- **Secure Copy (SCP)**
  - Uses SSH for authentication and data transfer (TCP port 22)
  - If you can SSH into a place, SCP mostly just works.
  - Unix-based systems (including Mac OS X): Should have it by default
  - Windows: WinSCP (https://winscp.net/eng/download.php)

Secure Copy (SCP)

- **Secure Copy**
  - Uses SSH for
  - Unix-based
  - Windows: W

rsync

- **rsync (rsync over SSH)**
  - Sync files and directories between two endpoints. (e.g., backups, only transfer new files)
  - **Careful with "--delete" option (this *mirrors* directories)**
  - Unix-based systems (including Mac OS X): Should have it by default
  - Windows: via Cygwin or DeltaCopy (never tried myself)

```
hyojoonk@hyojoonkmacwired:~/globus$ rsync -a -h --verbose ./test tigressdata2:
Password:
Duo two-factor login for hyojoonk

Enter a passcode or select one of the following options:

 1. Duo Push to XXX-XXX-
 2. Phone call to XXX-XXX-
 3. SMS passcodes to XXX-XXX-

Passcode or option (1-3): 1
building file list ... done
test/

sent 84 bytes   received 26 bytes   4.68 bytes/sec
total size is 0   speedup is 0.00
hyojoonk@hyojoonkmacwired:~/globus$
```

File Transfer Protocol (FTP)

- **'Secure' File Transfer Protocol ('S'FTP)**
  - Need FTP server running on the receiving end.
  - SFTP is more secure. Use if you want encrypted transfer.
  - Unix-based systems (including Mac OS X): Should have it by default
  - Windows: FileZilla (https://filezilla-project.org)

```
hyojoonk@hyojoonkmacwired:~/globus$ sftp tigressdata2:
Password:
Duo two-factor login for hyojoonk

Enter a passcode or select one of the following options:

 1. Duo Push to XXX-XXX-
 2. Phone call to XXX-XXX-
 3. SMS passcodes to XXX-XXX-

Passcode or option (1-3): 1
Connected to tigressdata2.
Changing to: /home/hyojoonk/
sftp> ls
10G.dat        rrddump.xml    test
sftp> get rrd
File "/home/hyojoonk/rrd" not found.
sftp> get rrddump.xml
Fetching /home/hyojoonk/rrddump.xml to rrddump.xml
/home/hyojoonk/rrddump.xml                    100%  249KB   2.7MB/s   00:00
sftp>
```

File Transfer Protocol (FTP)

- **'Secure' Fi**
  - Widely us
  - SFTP is r
  - Unix-base
  - Windows

File Transfer Protocol (FTP)

- Two-factor authentication with FileZilla
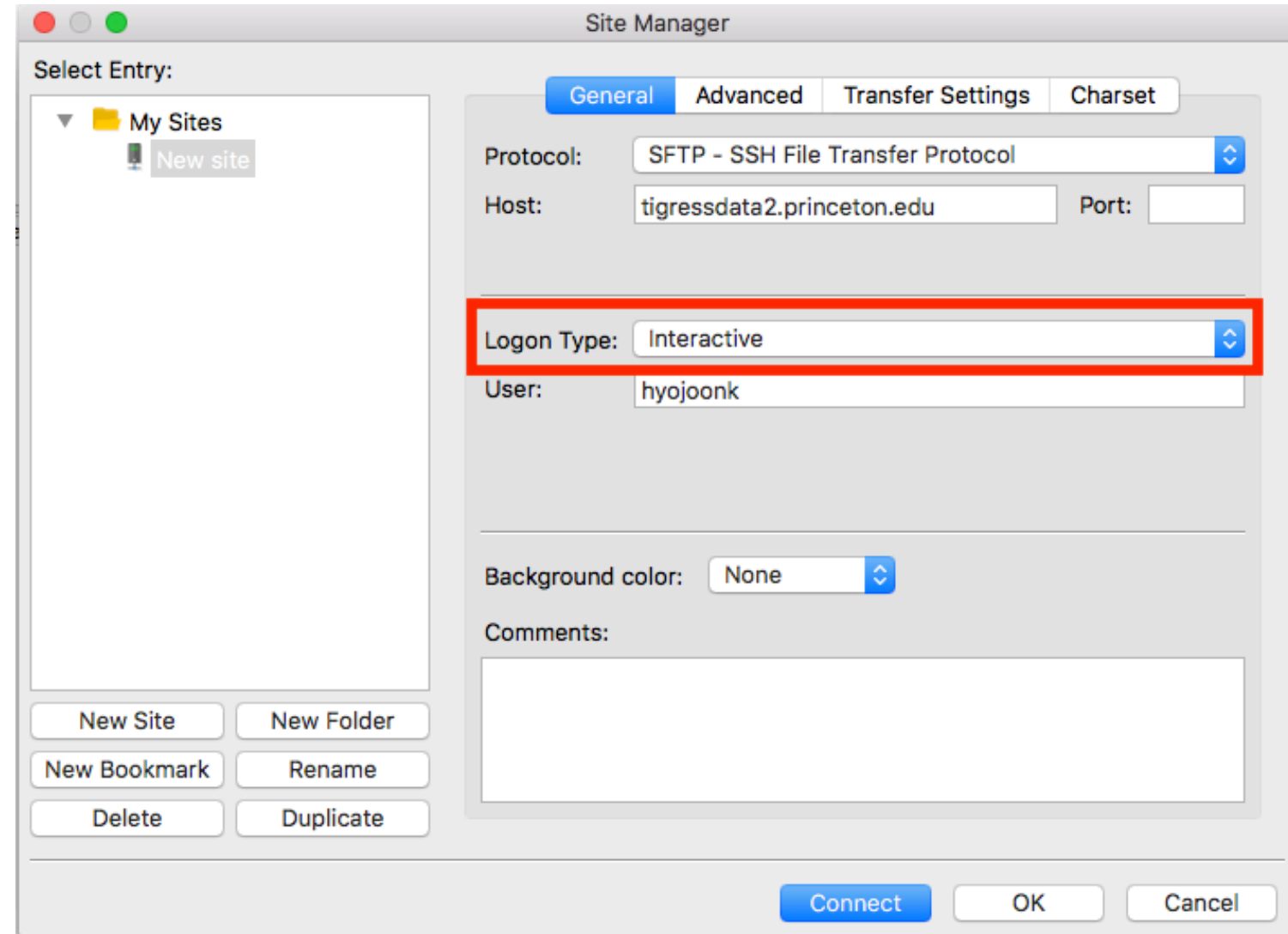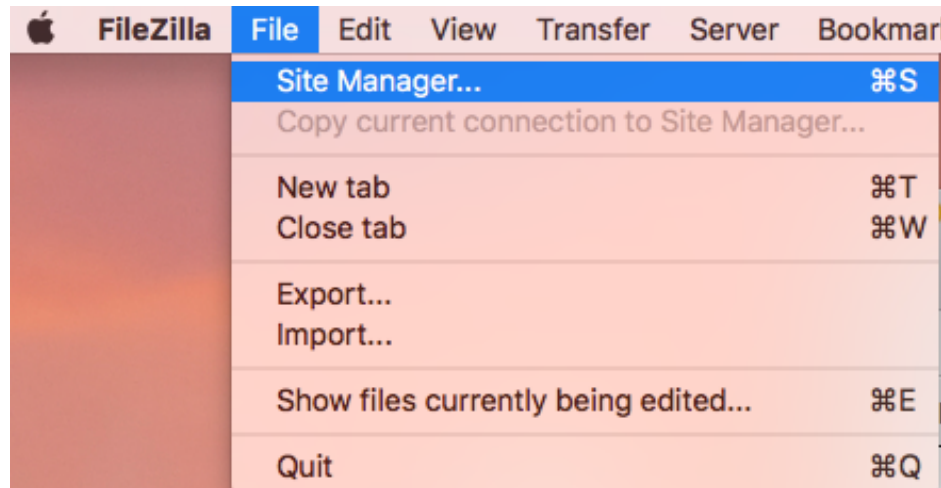
Browser click, wget, curl (HTTP/HTTPS)

- Web browser click

- wget (in terminal)
  - $ wget http://downloads.sourceforge.net/gnuwin32/wget-1.11.4-1-src.zip

- curl (in terminal)
  - $ curl -o data_src.zip http://downloads.sourceforge.net/gnuwin32/wget-1.11.4-1-src.zip

These tools are okay, but not always

- Great compatibility. Widely available.
- Small datasets. Quick transfers. (< 10 mins)

- Large bulk data transfers.
- Transfers on unreliable connections and hosts.

Transfer tools: Single vs multi stream

Single stream
- scp
- ftp
- rsync

src → dst

Multi stream
- GridFTP
- BBCP

src ⇒ dst

- Less packet loss (w/ dups)
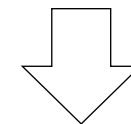- Better utilization of link

Faster transfer speed

Transfer tools: scp vs. GridFTP

Sample Results: disk-to-disk testing from Berkeley, CA to Argonne, IL (near Chicago). RTT = 53 ms, network capacity = 10Gbps, RAID = 4 disks, RAID Level-0. **Note that to get more than 1 Gbps (125 MB/s) disk to disk requires RAID.**

Tool	Throughput	Downloading 500 GB data
scp	140 Mbps (17.5 MB/s)	8 hours
HPN patched scp, 1 disk	760 Mbps (95 MB/s)	
HPN patched scp, RAID disk	1.2 Gbps (150 MB/s)	
GridFTP, 1 stream, 1 disk	760 Mbps (95 MB/s)	1.5 hours
GridFTP, 1 stream, RAID disk	1.4 Gbps (175 MB/s)	
GridFTP, 4 streams, RAID disk	5.4 Gbps (675 MB/s)	
GridFTP, 8 streams, RAID disk	6.6 Gbps (825 MB/s)	10 minutes

(ref: http://fasterdata.es.net/data-transfer-tools/)

OFFICE of INFORMATION TECHNOLOGY

PICSciE

24

Tools that might perform better than SCP

- **Globus/GridFTP (later)**

- **BBCP (**http://www.slac.stanford.edu/~abh/bbcp/**)**
  - Mac OS X, Linux-based systems. SSH-based access control
  - Both endpoints need the tool installed
  - "$ bbcp -V -s 16 /local/path/largefile.tar remotesystem:/remote/path/largefile.tar"
  - More info
    - http://www.nersc.gov/users/storage-and-file-systems/transferring-data/bbcp/

- **Fast Data Transfer (FDT)**
  - Java-based tool from Caltech & CERN (http://monalisa.cern.ch/FDT/)
  - Can theoretically run in any Operating System, including Windows
  - Need server-side running in server mode
  - "$ java -jar ./fdt.jar -ss 1M -P 10 -c remotehost.domain.uci.edu ~/file.633M -d /userdata/hjm"

Tools that might perform better than SCP

- **Aria2c** (https://aria2.github.io)
  - Faster http/https, ftp, sftp, BitTorrent, and Metalink download tool (x4 faster)
  - Windows, Mac, Linux, Android App
  - http: "$ ./aria2c -x4 -k1M http://foo.com/foo.zip"

- **LFTP (**http://lftp.tech**)**
  - Faster download (get) speed (2-5x) for ftp, http, sftp, fish, torrent. Upload (put) speed is same.
  - Compatible with normal FTP, HTTP servers.
  - Mac OS X, Linux-based systems. (apt-get install lftp; yum install lftp; brew install lftp)
  - ftp: "$ lftp ftp://speedtest.tele2.net"
  - http: "$ lftp -e 'pget -n 5 foo.zip' http://foo.com/"

- **HPN-patched SCP/SSH** (https://www.psc.edu/hpn-ssh)
  - Need both ends to be patched.

OFFICE of
INFORMATION TECHNOLOGY

PICSciE

Aspera

- Proprietary transfer tool owned by IBM (https://asperasoft.com)
  - Server is not free, but client program is free

- Uses Fast and Secure Protocol (FASP). Max ~10 Gb/s.

- Aspera servers at:
  - Amazon
  - European Nucleotide Archive (ENA)
  - US National Institutes of Health National Center for Biotechnology Information (NIH NCBI)

How to select a transfer tool

• **Transfer takes less than 10 mins** • **Not that frequent**	**Other than that, your transfer job is a noticeable chunk in your workflow**
SCP (WinSCP) **FTP (FileZilla)** **rsync**	**Globus (GridFTP)** **BBCP** **LFTP** **Aria2c** **FDT** **…**

OFFICE of INFORMATION TECHNOLOGY

PICSciE

(extra) Transfer settings: Encryption

Tool	Encrypted Control	Encrypted Data
FTP HTTP (even password-based access)		
BBCP BBFTP Globus/GridFTP	✔	
SCP SFTP rsync over SSH Globus/GridFTP with encryption-on HTTPS	✔	✔

Data encryption provides best security,
but negatively impacts transfer speed (10-50% slower)

# Summary and Best Practices

- **Data transfer speed is affected by: Endpoints, network, and transfer tool**

- **Know the limitation of your endpoints**

- **Used wired instead of wireless for large transfers**

- **Seek better transfer tools if transfer takes > 10 minutes and happens frequently**
  - e.g., Globus, BBCP

- **Ask for help**
  - Your department IT staff
  - About using RC resources: cses@princeton.edu

OFFICE of INFORMATION TECHNOLOGY

PICSciE

Transfer to/from Princeton RC HPC clusters

- I have data on my **workstation/laptop** that I want to process & analyze at **Princeton HPC clusters**
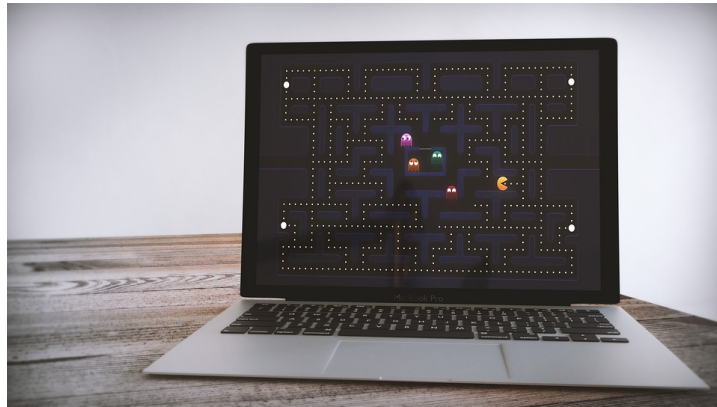
workstation/laptop
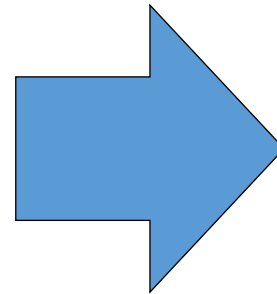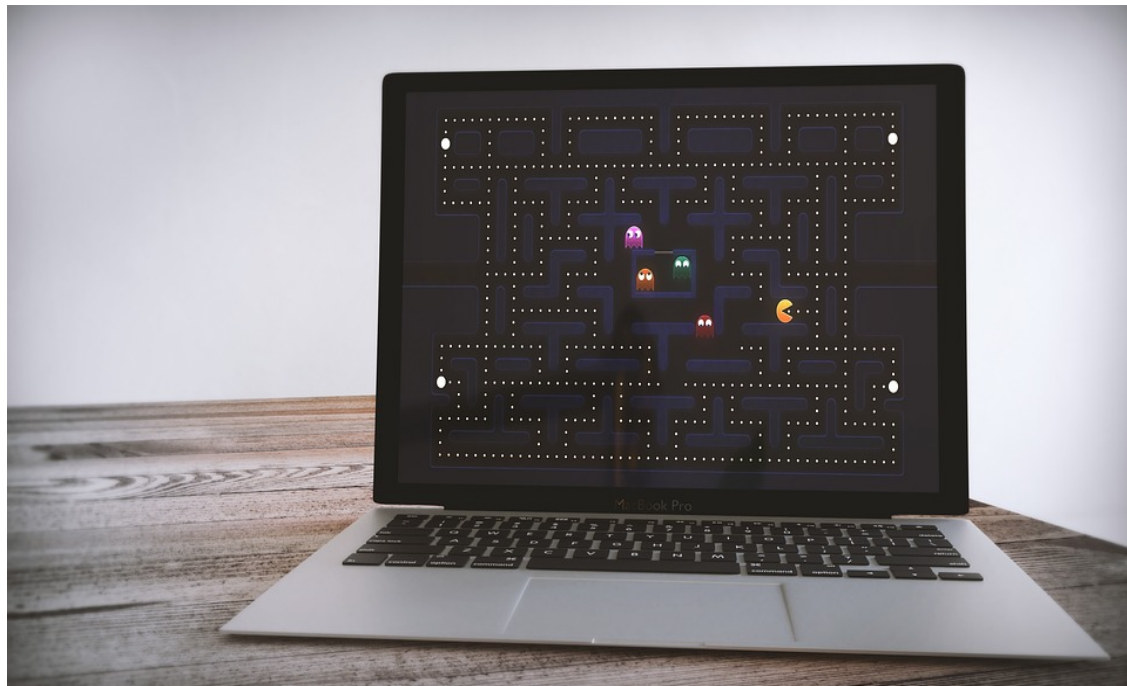
Princeton RC HPC Clusters

Transfer to/from Princeton RC HPC clusters

```
$ scp yum.log hyojoonk@della.princeton.edu:/tigress/hyojoonk/yum.log
Password:
```

Within campus or VPN

SCP

SFTP

rsync

Globus

BBCP (tigressdata)

Aspera (client install)

Some other system

Princeton RC HPC cluster nodes
(tiger, della, perseus, nobel, adroit, tigressdata)

- Using BBCP to transfer dataset that you plan to use on Della
  - Della itself does not have BBCP program installed

```
$ bbcp yum.log hyojoonk@della.princeton.edu:~/yum.log
Password:
```

```
bash: bbcp: command not found
bbcp: bbcp unexpectedly terminated on della.princeton.edu
```
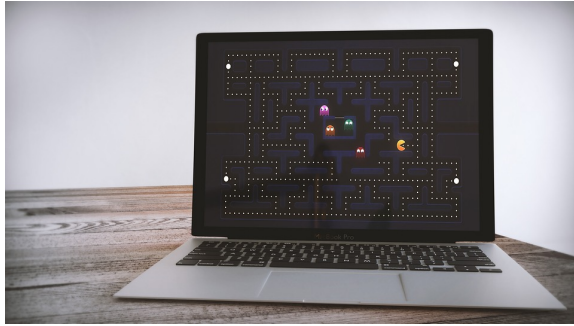
  - 1. Use BBCP to "Tigressdata" and use /tigress/<NetID>/ as destination

```
$ bbcp yum.log hyojoonk@tigressdata.princeton.edu:/tigress/hyojoonk/yum.log
```

  - 2. SSH into Della and access file from /tigress/<NetID>/

```
[hyojoonk@della5 ~]$ ls -la /tigress/hyojoonk/yum.log
-rw-r--r--. 1 hyojoonk cses 39132 Sep 13 12:52 /tigress/hyojoonk/yum.log
```

Quiz: should I find another tool?

10 GB data

Wired Macbook at Princeton
(1 Gb/s)

Tigressdata at Princeton
(10 Gb/s)

Will not be faster

```
hyojoonk@hyojoonkmacwired:~/globus$ scp 10G.dat tigressdata2:
hyojoonk@tigressdata2's password:
10G.dat                                     100% 9537MB  104.5MB/s   01:31
hyojoonk@hyojoonkmacwired:~/globus$
```

836 Mb/s

Globus

JoonLaptop to Princeton TIGRESS
transfer completed an hour ago

Overview		Event Log

Task ID 8612141a-9f1e-11e8-96e1-0a6d4e044368

Owner Hyojoon Kim (hyojoonk@princeton.edu)

Source JoonLaptop ⓘ
        owner: hyojoonk@princeton.edu

Files 1
Directories 0
Bytes Transferred 10 GB
Effective Speed 92.12 MB/s

736.96 Mb/s

OFFICE of
INFORMATION TECHNO

PICSciE

34

Quiz 2: should I find another tool?

100 GB data, 5 times per day

Some DTN at Princeton
(10 Gb/s)

Tigressdata at Princeton
(10 Gb/s)

Yes!

```
[hyojoonk@test-dtn test]$ scp 100G__.dat tigressdata2:/tigress/hyojoonk/100G.dat
hyojoonk@tigressdata2's password:
100G_b.dat                                      100%   93GB  144.5MB/s    11:00
[hyojoonk@test-dtn test]$
```

1.15 Gb/s (11 mins)

BBCP

```
[hyojoonk@test-dtn test]$ bbcp -V -s 8 100G__.dat tigressdata2:/tigress/hyojoonk/100G2.dat
Password:
Source test-dtn.princeton.edu using a final send window of 1548360
1 file copied at effectively 612.1 MB/s
[hyojoonk@test-dtn test]$
```

4.89 Gb/s (< 3 mins)

About Globus

What is Globus?

- Fast, reliable data transfer and management service

- Uses **GridFTP** underneath

- Main advantages
  - Fast transfer speed (multi-stream)
  - Convenient to use: "Fire-and-Forget"

OFFICE of INFORMATION TECHNOLOGY

Transfer tools: scp vs. GridFTP

Sample Results: disk-to-disk testing from Berkeley, CA to Argonne, IL (near Chicago). RTT = 53 ms, network capacity = 10Gbps, RAID = 4 disks, RAID Level-0. **Note that to get more than 1 Gbps (125 MB/s) disk to disk requires RAID.**

		Downloading 500 GB data
Tool	**Throughput**	
scp	140 Mbps (17.5 MB/s)	**8 hours**
HPN patched scp, 1 disk	760 Mbps (95 MB/s)	
HPN patched scp, RAID disk	1.2 Gbps (150 MB/s)	
GridFTP, 1 stream, 1 disk	760 Mbps (95 MB/s)	**1.5 hours**
GridFTP, 1 stream, RAID disk	1.4 Gbps (175 MB/s)	
GridFTP, 4 streams, RAID disk	5.4 Gbps (675 MB/s)	
GridFTP, 8 streams, RAID disk	6.6 Gbps (825 MB/s)	**10 minutes**

(ref: http://fasterdata.es.net/data-transfer-tools/)

OFFICE of INFORMATION TECHNOLOGY

PICSciE

38

Easy to use

- Use a web browser to request a transfer job

- "Fire-and-Forget"

- Get email notification when transfer is complete (or was unsuccessful)

1. Select "from" and "to"

2. Select file or directory

3. Click

4. Wait

optional step

OFFICE of INFORMATION TECHNOLOGY

How it works

1. Go to https://www.globus.org.

2. Pick two endpoints. Submit transfer request.

3. Dataset is transferred between two endpoints.
   - Your machine's web browser is a "remote control"
   - But, your machine can be an endpoint too

4. Get notification when transfer is done.

3. Data

A   **B**

globus
Globus.org

1. Select endpoints
2. Request transfer

4. Get notification
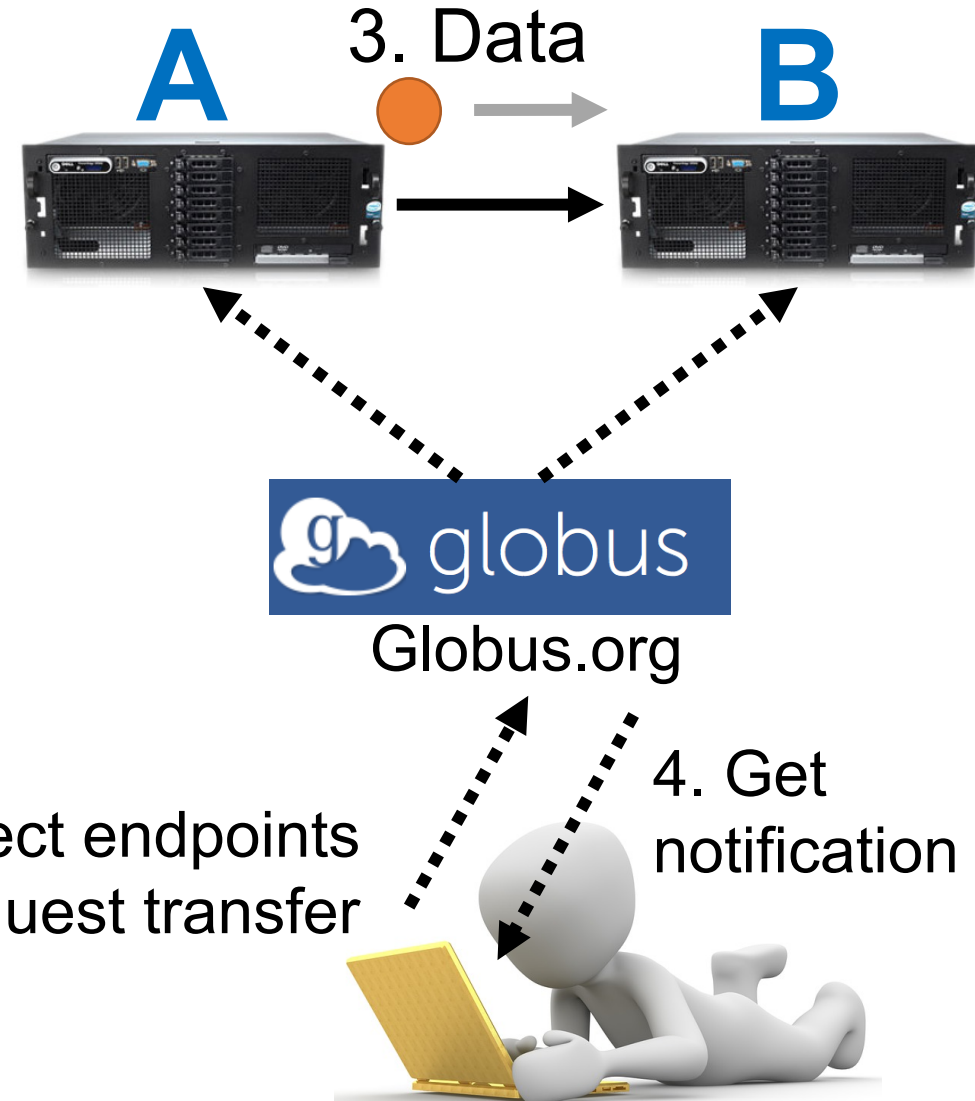
- **To use the Globus service**
  - Account and a web browser. Use your Princeton Net ID.
  - Access to source and destination endpoints

- **Globus endpoint types**
  - Globus Connect Server
    - Administrator providing Globus service
    - Linux (CentOS, Fedora, Red Hat, Scientific Linux, SuSE, Debian/Ubuntu)
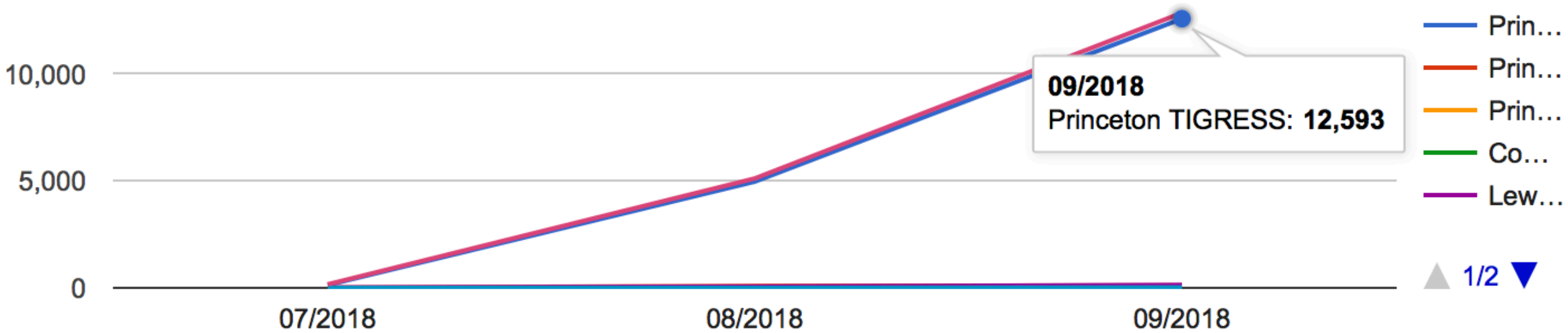 - Globus Connect Personal
    - Researcher transferring data to/from own workstation/laptop
    - Mac, Linux, Windows

OFFICE of
INFORMATION TECHNOLOGY

PICSciE

Globus has good coverage

- Universities

- DOE national labs
  - ESnet at CERN, ANL, LBNL, LLNL, LANL, ORNL, and PNNL

- National computing facilities
  - NERSC, NCSA, SDSC …

- Federal agencies
  - NIH, USDA, NASA/JPL, USGS …

- Over 50,000 registered endpoints at over 500 institutions worldwide

OFFICE of INFORMATION TECHNOLOGY

PICSciE

Number of tranfers per month

09/2018
Princeton TIGRESS: **12,593**

Activity stats per target DTN

DTN	Total # of Transfers (to+from)	Total # of files	Total size(GB)
Princeton Astrophysics DTN	12	12	189
Princeton TIGRESS FRNN	1	2,013,045	874.74
Lewis-Sigler Institute Core DTN	15	1,014	320.17
NERSC DTN	4	4	91.44
Princeton Neuroscience Institute DTN	137	1,102,625	29,436.05
NERSC Cori	17	1,479	304.07
Private endpoint	9,705	8,459,453	181,425.33
XSEDE TACC stampede2	20	30,400	6,418.53
XSEDE PSC bridges	8	74	107.86
pppl#pppl-globus2	1	2,889	10,306.84
XSEDE Comet	7	58	78.26
ncsa#BlueWaters	8	76	35.69
DeepWang2	8	136	4,472.71
cineca#MARCONI	3	11	199.66
Simons Foundation Data Server	11	273,572	3,782.33
ncsa#Nearline	1	32	978.63
Princeton TIGRESS Akey_eGTeX_exome_seq	4	560	5,287.04
NCAR RDA Dataset Archive	2	1,596	282.46
XSEDE Ranch	1	18	521.88
NCAR GLADE	24	16,061	1,604.34

Physics DTN   **Tigress DTN**   **Lewis-Sigler DTN**   **PNI DTN**   **CS DTN**

Globus Connect Server (GridFTP)

10 Gbps (Mar. 2016 -)
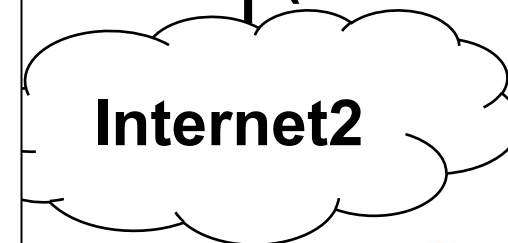
10 Gbps (Nov. 2015 -)
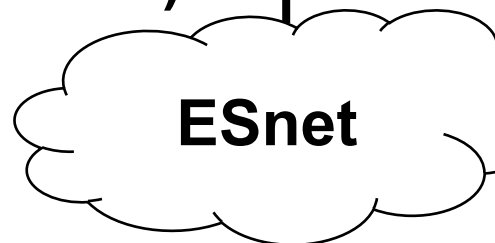
ESnet

Internet2

Data Transfer Nodes (DTNs)
- High-end servers
- 10 Gbps connections
- Tuned and optimized. RAID configured
- Globus installed
- Supported

OFFICE of INFORMATION TECHNOLOGY

PICSciE

45

Data Transfer Nodes w/ Globus @ Princeton

- **Tigress DTN (Princeton TIGRESS)**
  - Contact: cses@princeton.edu

- **LSI DTN (Lewis-Sigler Institute Core DTN)**
  - Contact: csgenome@princeton.edu

- **PNI DTN (Princeton Neuroscience Institute DTN)**
  - Contact: pnihelp@princeton.edu

- **CS DTN (Computer Science Department DTN)**
  - Contact: csstaff@cs.princeton.edu

- **Physics DTN (Princeton Physics DTN)**
  - Contact: Vinod Gupta (vinod@Princeton.EDU), Sumit Saluja (ssaluja@Princeton.edu)

- **Astrophysics DTN (Princeton Astrophysics DTN)**
  - Contact: help@astro.princeton.edu

OFFICE of INFORMATION TECHNOLOGY

PICSciE

Contact your departmental staff

- About existing DTNs
  - Check if you can use these DTNs

- About having a departmental DTN

Internet

10 Gb/s

Transfer via Globus

If you have an account on RC systems, you can use Globus and Tigress DTN

Tigress DTN

scratch

scratch

scratch

/tigress/<ID>   /tiger/scratch/gpfs/<ID>   /della/scratch/gpfs/<ID>   /perseus/scratch/gpfs/<ID>
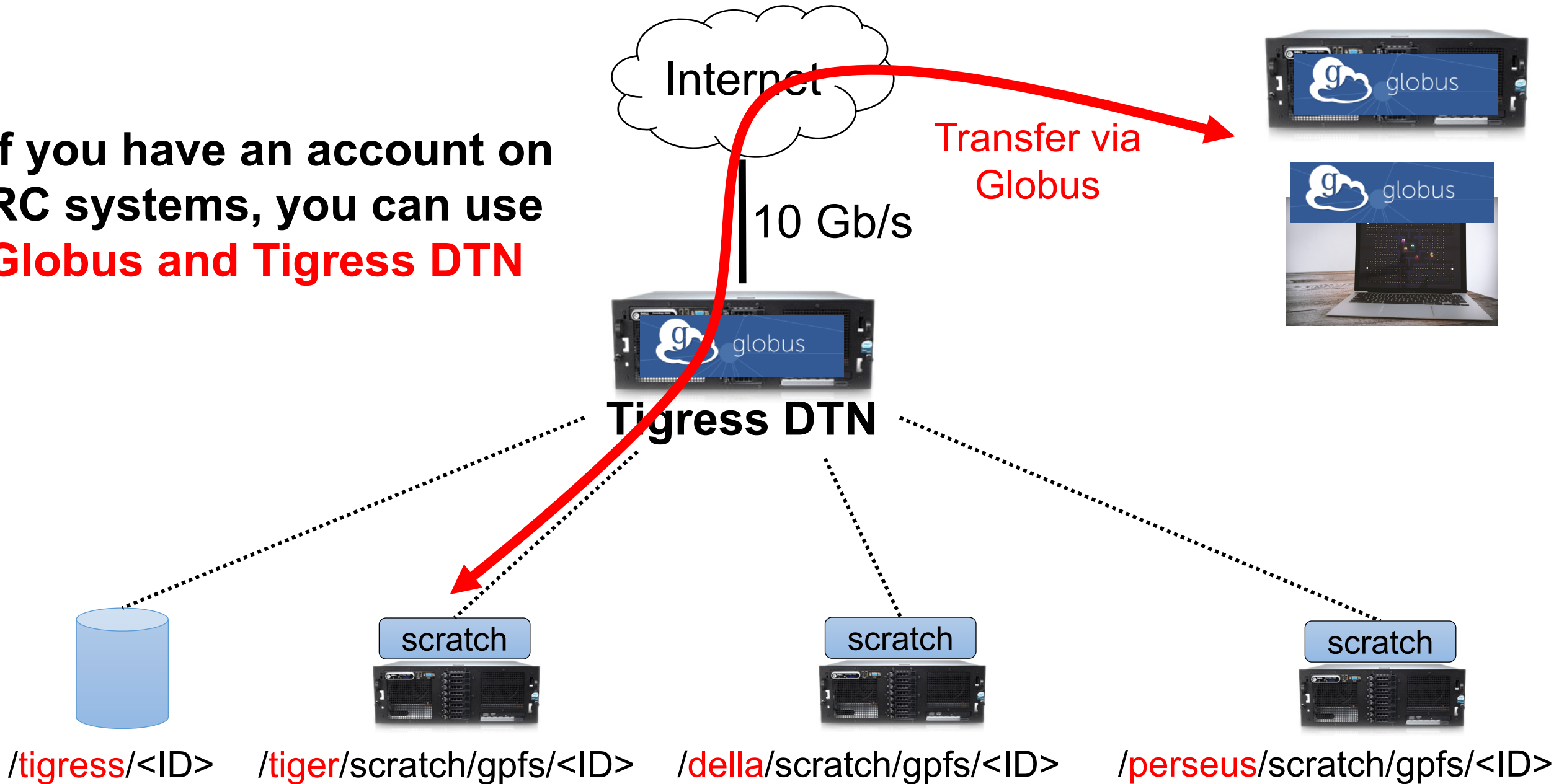
Demo

- I have data at **CERN (Geneva)** that I want to process & analyze at **Princeton HPC clusters**

CERN Large Hardon Collider
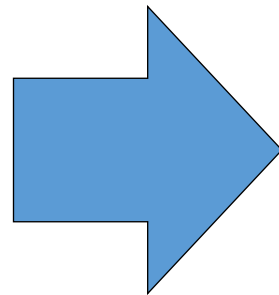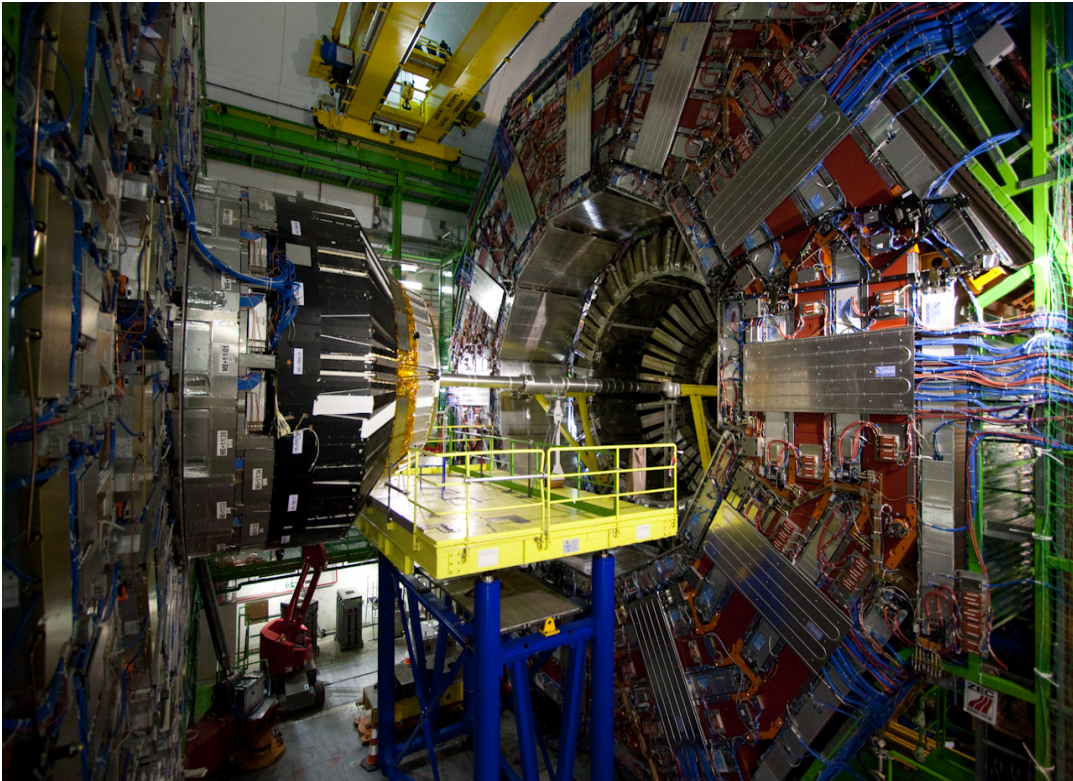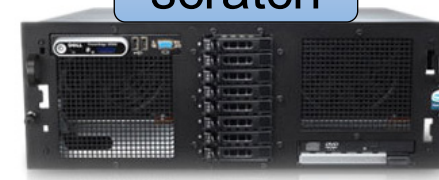
Princeton HPC Clusters

Use case 1

Globus.org

Internet

CERN DTN
Globus endpoint
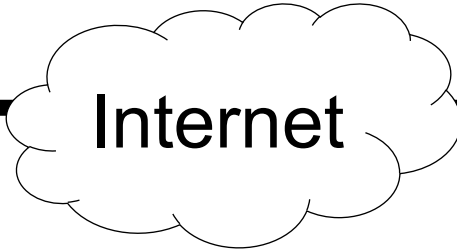
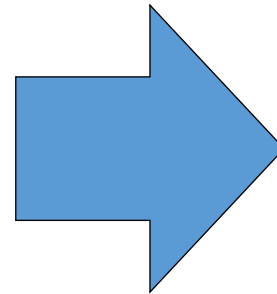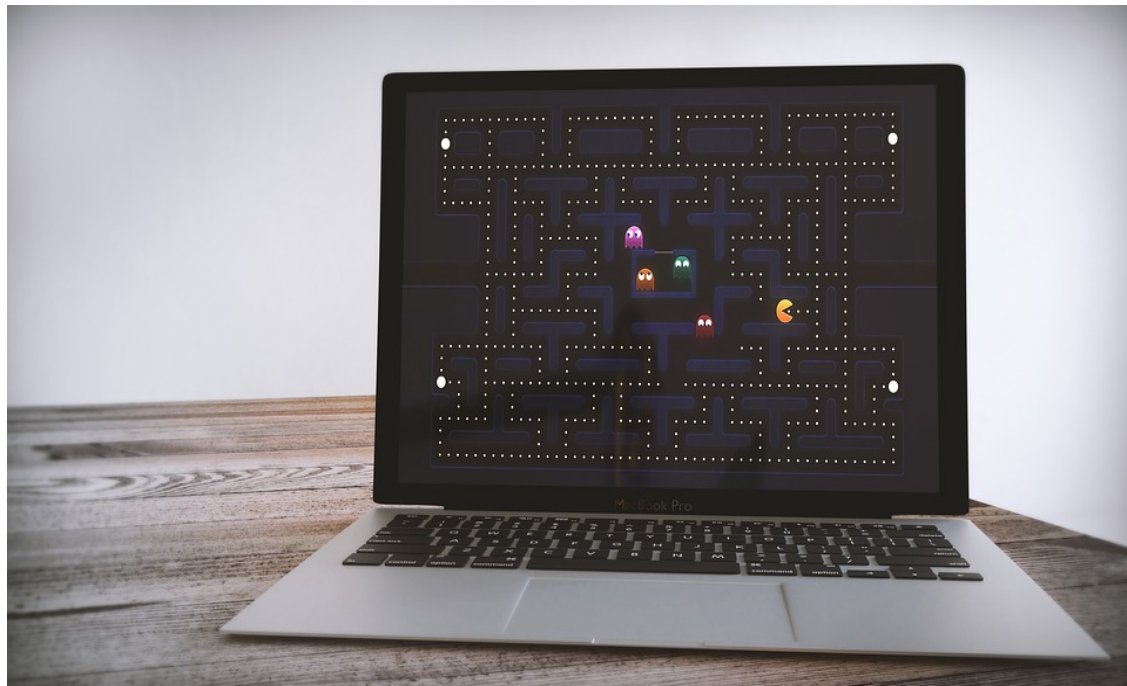Princeton Tigress DTN
Globus endpoint

scratch

/della/scratch/gpfs/hyojoonk/

CERN file server

51

Use case 2

- I have data on **my laptop** that I want to process & analyze at **Princeton HPC clusters**
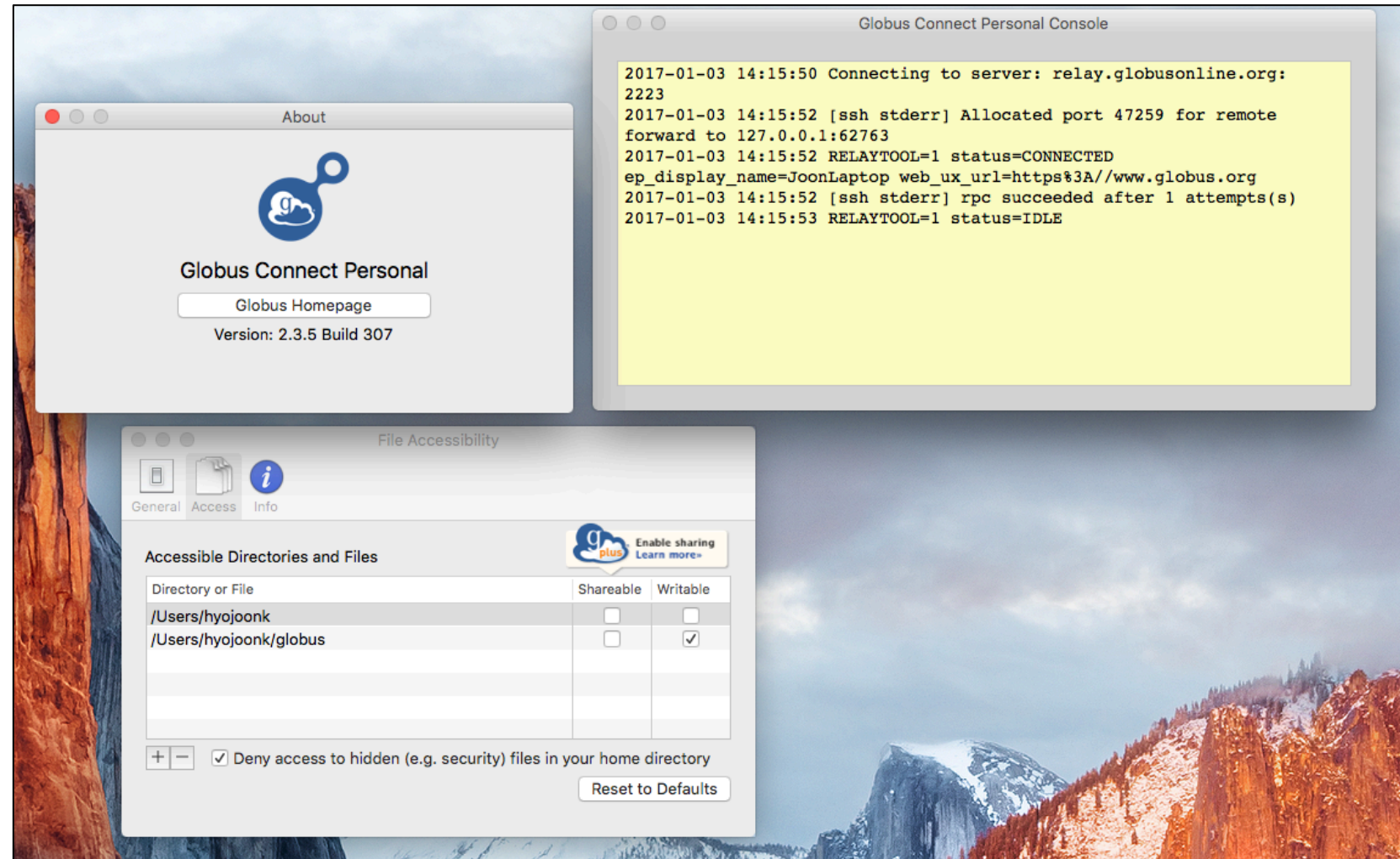
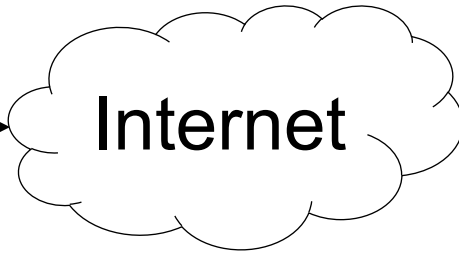My laptop

Princeton HPC Clusters
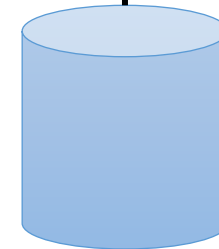
Globus Connect Personal

- Make your own machine a Globus endpoint
  - Mac, Windows, Linux

- You are the administrator for your own Globus endpoint

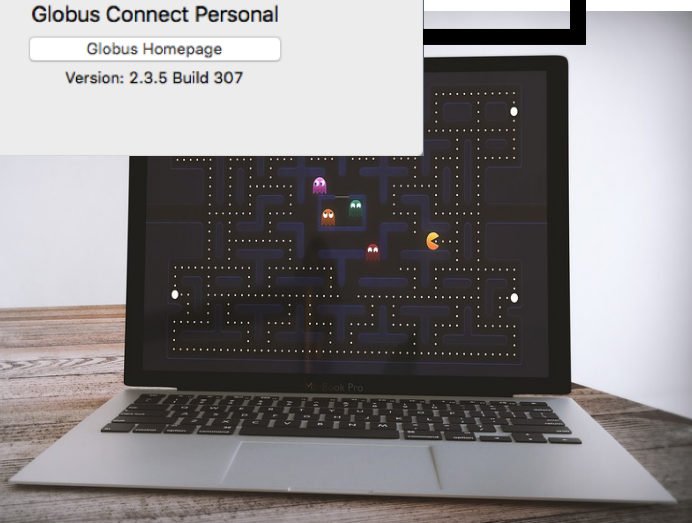- Limited performance (# of streams), but convenient!

Globus.org

Internet

Princeton Tigress DTN
Globus endpoint

/tigress/hyojoonk/

My Laptop

Additional information about Globus

- **Sharing:** Share file or directory with other Globus users
  - https://docs.globus.org/how-to/share-files/
  - For Globus sharing on Tigress DTN, contact cses@princeton.edu

- **Scripting:** Command Line Interface
  - https://docs.globus.org/cli/

- **Even more scripting/programming:** Python SDK
  - https://globus-sdk-python.readthedocs.io/en/stable/

- More information about Globus:
  - https://docs.globus.org/

"Data Transfer Basics and Globus Transfer Tool Tutorial"
(Spring, 2019)

https://researchcomputing.princeton.edu/education/workshops

OFFICE of
INFORMATION TECHNOLOGY

PICSciE

Summary and Takeaways

- **Small and quick transfers**: basic transfer tools (e.g.,SCP, SFTP (FileZilla), rsync, etc)

- **Large (and small) transfers**: Data Transfer Nodes (DTNs) and Globus when possible

- **Large transfers, but Globus is unavailable**: BBCP, FDT, Aspera, LFTP, Aria2c, etc

- **Know your environment and limitations** Endpoints, network, transfer tool

- **Don't hesitate to reach out to us**

OFFICE of INFORMATION TECHNOLOGY
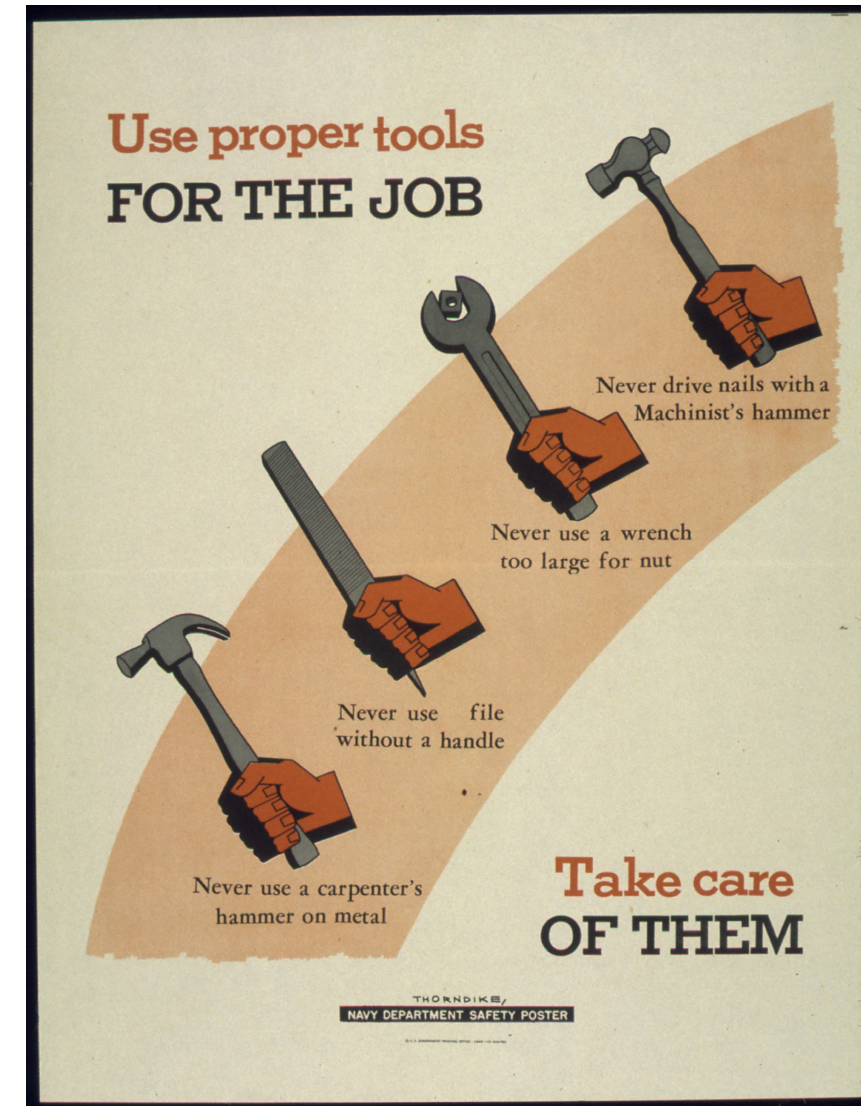
Q&A

joonk@princeton.edu
Computational Science and Engineering Support: cses@princeton.edu
Research Computing Website: https://researchcomputing.princeton.edu

Help sessions:
Lewis Science Library 347
Tuesday 10-11am
Thursday 2-3pm (visualization emphasis)