Arkansas High Performance Computing Center

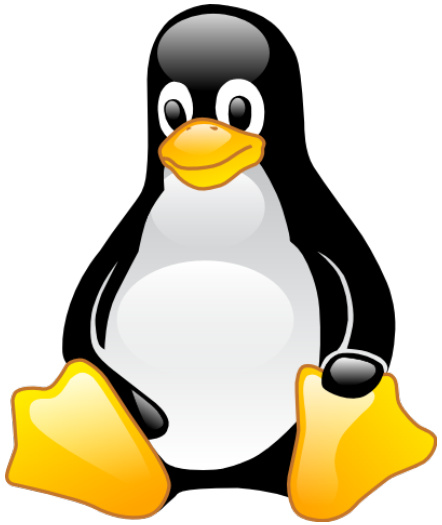
at the

University of Arkansas

UNIVERSITY OF ARKANSAS

AHPCC Workshop Series

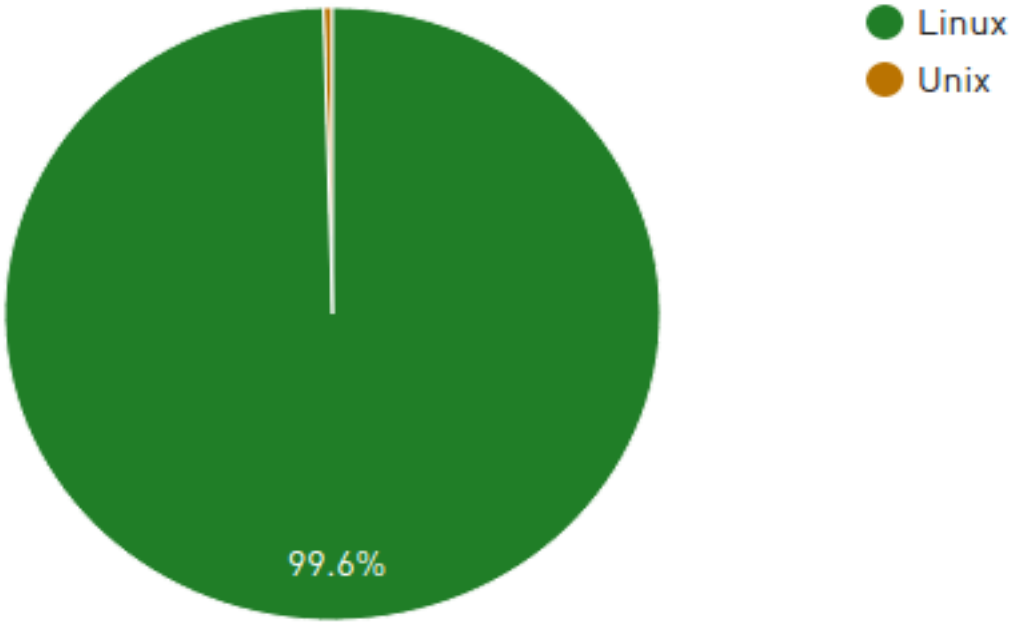
Introduction to Linux for HPC



Why Linux?

Compatible with many architectures
OS of choice for large scale computing
Free, high quality software tools and applications
Open standards/accessibility to source code
Development environment of choice for scientific research

Operating system Family Performance Share



Click a row in the table below to drill down.

OPERATING SYSTEM FAMILY	COUNT	SYSTEM SHARE (%)	RMAX (GFLOPS)	RPEAK (GFLOPS)	CORES
Linux	494	98.8	419,253,297	640,067,791	29,226,383
Unix	6	1.2	1,868,442	2,215,095	76,128

http://top500.org/statistics/list/

Linux for HPC

A few notes before beginning…

Linux is case sensitive! "A" and "a" are NOT the same thing.

It is best to avoid spaces in file naming. My Math Homework.txt should be My-Math-Homework.txt, or My_Math_Homework.txt.

Extensions in file names are ignored by Linux (but helpful to you…aka .txt, .sh, .dat, .pbs)

By default, Linux does not ask "Are You Sure" before performing operations. Be careful! (use –i if nervous)

No Trash Can or Waste Basket to retrieve accidentally deleted files from.

Control+C is your friend…interrupts a hung or persistent command.

"." means "right here", ".." means "back up one"

"~" is shortcut for "home directory"

TAB command completion

Arrow Up last command search

Linux for HPC

"ssh" - creates secure connection to remote system
"pwd" - lists your current location on system
"ls" - shows content of current directory
"man" - get help and usage for a particular command
"df" - shows disk space used
"du" - shows disk usage in given directory
"cd" - changes to a different directory
"mkdir" - make new directory or folder
"nano" - text editor
"more" - allows viewing of text files without editor ("less" – includes scrolling)
"mv" - moves or renames a file
"cp" - copies a file to another name or location
"wc" - counts characters, words, or lines in a file
"head" - grab lines at top of text file
"tail" - grab lines at bottom of text file
"grep" - searches for specific words or strings
"diff" - compares two files to show differences
"rm" - deletes files and directories
"wget" - command to transfer files via web protocol
"scp" - transfer protocol for file transfer, related to ssh
"history" - list of commands executed previously

Linux for HPC

ssh - OpenSSH SSH client (remote login program)

```
uaf120305:~ eCycle$ ssh jpummil@trestles.uark.edu
jpummil@trestles.uark.edu's password:
Last login: Mon Dec  7 09:34:21 2015 from zenlb4
[jpummil@tres-l2 ~]$
```

Linux for HPC

pwd – print name of current/working directory (folder)

```
[jpummil@tres-l2 ~]$ pwd
/home/jpummil
```

Linux for HPC

ls – list directory (folder) contents

```
[jpummil@tres-l2 ~]$ ls
Apps  intel  perl5  test

[jpummil@tres-l2 ~]$ ls -lh
total 16K
drwxrwxr-x 21 jpummil jpummil 4.0K Nov  7 16:28 Apps
drwxr-xr-x  3 jpummil jpummil 4.0K Nov  7 16:31 intel
drwxrwxr-x  5 jpummil jpummil 4.0K Nov  6 11:21 perl5
drwxrwxr-x  2 jpummil jpummil 4.0K Dec  8 10:28 test
```

Linux for HPC

man – format and display the online manual pages

[jpummil@tres-l2 ~]$ **man** ls

NAME
ls - list directory contents

SYNOPSIS
ls [OPTION]... [FILE]...

DESCRIPTION
List information about the FILEs (the current directory by default). Sort entries alphabetically if none of **-cftuvSUX** nor **--sort**.

Mandatory arguments to long options are mandatory for short options too.

-a, **--all**
do not ignore entries starting with .

Linux for HPC

df – report file system disk space usage

```
[jpummil@tres-l2 ~]$ df -h
Filesystem            Size  Used Avail Use% Mounted on
/dev/sda3             901G  3.2G  852G   1% /
tmpfs                  16G  8.0K   16G   1% /dev/shm
/dev/sda1             248M   43M  193M  19% /boot
172.17.6.15@o2ib:/storaget
                       70T   12M   70T   1% /storage2
172.17.6.15@o2ib:/scratch2
                       16T  1.7T   14T  11% /scratch
172.17.6.19@o2ib:/storagep
                       70T  113G   70T   1% /storage
172.17.6.19@o2ib:/storaged
                       70T   12M   70T   1% /storaged
```

Linux for HPC

du – estimate file space usage

```
[jpummil@tres-l2 ~]$ du -sh
2.4G        .
[jpummil@tres-l2 ~]$ du -sh work
2.5M        perl5
```

Linux for HPC

rm – remove files or directories

```
[jpummil@tres-l2 ~]$ ls
Apps  files  file.txt  intel  perl5  phone.txt  work

[jpummil@tres-l2 ~]$ touch file.txt

[jpummil@tres-l2 ~]$ ls
Apps  files  file.txt  intel  perl5  phone.txt  work

[jpummil@tres-l2 ~]$ rm file.txt

[jpummil@tres-l2 ~]$ ls
Apps  files  intel  perl5  phone.txt  work
```

(need rm –R to delete directories with files in them)

Linux for HPC

mkdir/rmdir (rm –R) – make/remove directories (folders)

```
[jpummil@tres-l2 ~]$ ls
Apps  intel  perl5

[jpummil@tres-l2 ~]$ mkdir test

[jpummil@tres-l2 ~]$ ls
Apps  intel  perl5  test

[jpummil@tres-l2 ~]$ rmdir test

[jpummil@tres-l2 ~]$ ls
Apps  intel  perl5

[jpummil@tres-l2 ~]$ mkdir work
```

Linux for HPC

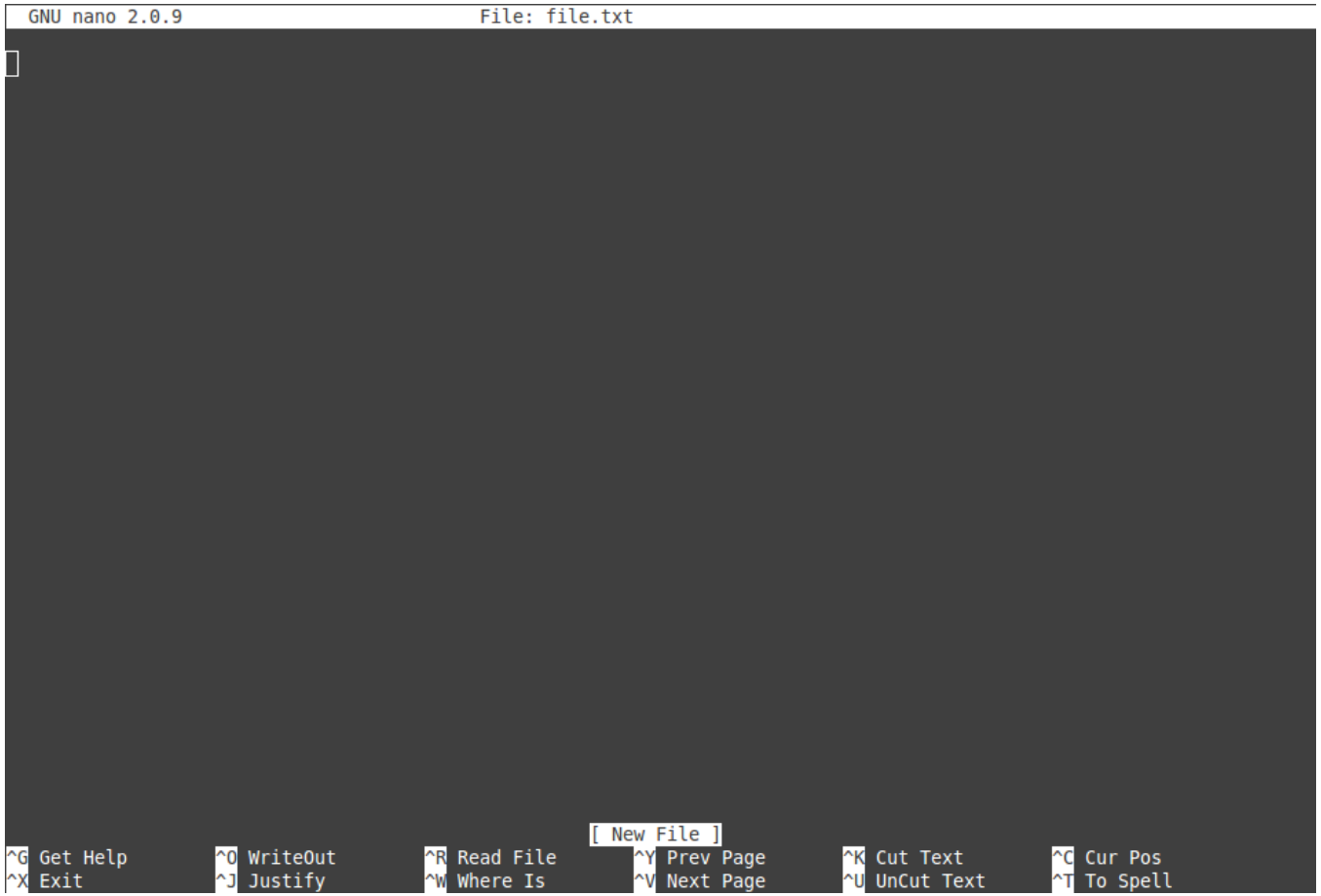
cd – change to a different directory (folder)

```
[jpummil@tres-l2 ~]$ pwd
/home/jpummil

[jpummil@tres-l2 ~]$ ls
Apps  intel  perl5 work

[jpummil@tres-l2 ~]$ cd work

[jpummil@tres-l2 perl5]$ pwd
/home/jpummil/work

[jpummil@tres-l2 perl5]$ ls

[jpummil@tres-l2 perl5]$ cd ..

[jpummil@tres-l2 ~]$ pwd
/home/jpummil
```

Linux for HPC

nano – Nano's ANOther editor

[jpummil@tres-l2 ~]$ nano file.txt



Linux for HPC

more – file perusal filter for crt viewing

```
[jpummil@tres-l2 ~]$ more file.txt
apple
banana
orange
lemon
lime
```

Linux for HPC

mv – move (rename) files

```
[jpummil@tres-l2 ~]$ ls
Apps  file.txt  intel  perl5

[jpummil@tres-l2 ~]$ mv file.txt fruit.txt

[jpummil@tres-l2 ~]$ ls
Apps  fruit.txt  intel  perl5

[jpummil@tres-l2 ~]$ more fruit.txt
apple
banana
orange
lemon
lime
```

Linux for HPC

cp – copy files and directories

```
[jpummil@tres-l2 ~]$ ls
Apps  fruit.txt  intel  perl5

[jpummil@tres-l2 ~]$ cp fruit.txt old-fruit.txt

[jpummil@tres-l2 ~]$ ls -l
total 13
drwxrwxr-x 21 jpummil jpummil 4096 Nov  7 16:28 Apps
-rw-rw-r--  1 jpummil jpummil   31 Dec  8 12:38 fruit.txt
drwxr-xr-x  3 jpummil jpummil 4096 Nov  7 16:31 intel
-rw-rw-r--  1 jpummil jpummil   31 Dec  8 12:52 old-fruit.txt
drwxrwxr-x  5 jpummil jpummil 4096 Nov  6 11:21 perl5

[jpummil@tres-l2 ~]$ more old-fruit.txt
apple
banana
orange
lemon
lime
```

Linux for HPC

wc – print newline, word, and byte counts for each file

```
[jpummil@tres-l2 ~]$ wc fruit.txt
 5  5 31 fruit.txt

[jpummil@tres-l2 ~]$ wc old-fruit.txt
 5  5 31 old-fruit.txt
```

Linux for HPC

head – output the first part of files

```
[jpummil@tres-l2 ~]$ more fruit.txt
apple
banana
orange
lemon
Lime

[jpummil@tres-l2 ~]$ head -n 2 fruit.txt
apple
banana
```

Linux for HPC

tail – output the last part of files

```
[jpummil@tres-l2 ~]$ more fruit.txt
apple
banana
orange
lemon
Lime

[jpummil@tres-l2 ~]$ tail -n 2 fruit.txt
lemon
Lime
```

Linux for HPC

grep – print lines matching a pattern

```
[jpummil@tres-l2 ~]$ ls
Apps  fruit.txt  intel  old-fruit.txt  perl5  phone.txt

[jpummil@tres-l2 ~]$ more phone.txt
123-456-7890
987-654-3210
654-098-4321

[jpummil@tres-l2 ~]$ grep 123 phone.txt
123-456-7890
```

(grep –i, ignore case grep –v print all but matching)

Linux for HPC

sort – sort lines of files

```
[jpummil@tres-l2 ~]$ more phone.txt
123-456-7890
987-654-3210
654-098-4321

[jpummil@tres-l2 ~]$ sort phone.txt
123-456-7890
654-098-4321
987-654-3210

[jpummil@tres-l2 ~]$ sort -r phone.txt
987-654-3210
654-098-4321
123-456-7890
```

Linux for HPC

diff – compare files line by line

```
[jpummil@tres-l2 ~]$ more fruit.txt
apple
banana
orange
Lemon
lime

[jpummil@tres-l2 ~]$ more fruit2.txt
apple
banana
orange
cherry
lime

[jpummil@tres-l2 ~]$ diff fruit.txt fruit2.txt
4c4
< lemon
---
> cherry
```

Linux for HPC

history – list of previously executed commands

```
[jpummil@tres-l2 ~]$ history | tail -n 18
 1110  wget http://home.iae.nl/users/mhx/flops.c
 1111  ls
 1112  which gcc
 1113  gcc -o flops flops.c
 1114  more flops.c
 1115  gcc -DUNIX -o flops flops.c
 1116  ls
 1117  ./flops
 1118  ls
 1119  rm flops
 1120  clear
 1121  ls
 1122  scp flops.c jpummil@comp.uark.edu:
 1123  scp jpummil@comp.uark.edu:flops.c .
 1124  scp jpummil@comp.uark.edu:flops.c jpummil@sublime.uark.edu:
 1125  man history
 1126  clear
 1127  history | tail -n 20
```

Linux for HPC

Merging files

> and >> operators for re-directing output to create new files

> Will create new file, but subsequent use will overwrite existing data

```
[jpummil@tres-l2 ~]$ more phone.txt > merged.txt
[jpummil@tres-l2 ~]$ more merged.txt
123-456-7890
987-654-3210
654-098-4321
[jpummil@tres-l2 ~]$ more fruit.txt > merged.txt
[jpummil@tres-l2 ~]$ more merged.txt
apple
banana
orange
lemon
lime
```

Linux for HPC

Merging files

> and >> operators for re-directing output to create new files

>> Will also create new file, but will instead append new data to existing data

```
[jpummil@tres-l2 ~]$ more phone.txt >> merged.txt
[jpummil@tres-l2 ~]$ more merged.txt
123-456-7890
987-654-3210
654-098-4321
[jpummil@tres-l2 ~]$ more fruit.txt >> merged.txt
[jpummil@tres-l2 ~]$ more merged.txt
123-456-7890
987-654-3210
654-098-4321
apple
banana
orange
lemon
lime
```

Linux for HPC

“Piping” commands

| is called the “pipe” in Linux (above the \ on the far right side of keyboard)

| allows you to join commands together in a string to do multiple operations

```
[jpummil@tres-l2 ~]$ grep -i 123 phone.txt > temp.txt

[jpummil@tres-l2 ~]$ wc -c temp.txt
13 temp.txt
```

***Could be written as one command using the pipe (|)….

```
[jpummil@tres-l2 ~]$ grep -i 123 phone.txt | wc -c
13
```

Linux for HPC

Using “wildcards”

* Is considered a matching character or “wildcard” in linux

```
[jpummil@tres-l2 ~]$ ls
Apps  extra.txt  flops.c  fruit2.txt  fruit.txt  intel  perl5  phone.txt  temp2.txt  temp.txt

[jpummil@tres-l2 ~]$ ls *.txt
extra.txt  fruit2.txt  fruit.txt  phone.txt  temp2.txt  temp.txt

[jpummil@tres-l2 ~]$ ls fr*.txt
fruit2.txt  fruit.txt

[jpummil@tres-l2 ~]$ rm te*.txt

[jpummil@tres-l2 ~]$ ls *.txt
extra.txt  fruit2.txt  fruit.txt  phone.txt
```

(Reference also both “?” and “[]” wildcards for even more flexibility)

Linux for HPC

wget – The non-interactive network downloader

```
[jpummil@tres-l2 ~]$ ls
Apps  fruit2.txt  fruit.txt  intel  perl5  phone.txt

[jpummil@tres-l2 ~]$ wget http://home.iae.nl/users/mhx/flops.c
--2015-12-08 13:28:28--  http://home.iae.nl/users/mhx/flops.c
Resolving home.iae.nl... 212.61.15.15
Connecting to home.iae.nl|212.61.15.15|:80... connected.
HTTP request sent, awaiting response... 200 OK
Length: 34942 (34K) [text/x-c]
Saving to: “flops.c”

100%[======================================================================>] 34,942      127K/s   in 0.3s

2015-12-08 13:28:29 (127 KB/s) - “flops.c” saved [34942/34942]

[jpummil@tres-l2 ~]$ ls
Apps  flops.c  fruit2.txt  fruit.txt  intel  perl5  phone.txt
```

Linux for HPC

scp – secure copy (remote file copy program)

From trestles.uark.edu (local) to comp.uark.edu (remote):

```
[jpummil@tres-l2 ~]$ scp flops.c jpummil@comp.uark.edu:
Password:
flops.c                                       100%   34KB  34.1KB/s   00:00
```

From comp.uark.edu (remote) to trestles.uark.edu (local):

```
[jpummil@tres-l2 ~]$ scp jpummil@comp.uark.edu:flops.c .
Password:
flops.c                                       100%   34KB  34.1KB/s   00:00
```

(Remember….SOURCE, then DESTINATION)

Linux for HPC

Compressing and archiving files

Lots of files become tedious to maintain over time and they also take up lots of space on the system. For these reasons and others, it is nice to be able to organize and simplify things on occasion.

Look at man pages to view the commands “tar” and “gzip”

Files “bundled together” into one package will have a .tar suffix.

If they have also been compressed with gzip, they will be .tar.gz or .tgz

```
[jpummil@tres-l1 test]$ ls
flops2.c  flops3.c  flops4.c  flops5.c  flops.c
```

Linux for HPC

Compressing and archiving files

```
[jpummil@tres-l1 test]$ tar cf flops.tar flops*.c

[jpummil@tres-l1 test]$ ls -lh
total 3.0K
-rw-rw-r-- 1 jpummil jpummil  35K Dec  9 13:07 flops2.c
-rw-rw-r-- 1 jpummil jpummil  35K Dec  9 13:08 flops3.c
-rw-rw-r-- 1 jpummil jpummil  35K Dec  9 13:08 flops4.c
-rw-rw-r-- 1 jpummil jpummil  35K Dec  9 13:08 flops5.c
-rw-rw-r-- 1 jpummil jpummil  35K Dec  9 13:07 flops.c
-rw-rw-r-- 1 jpummil jpummil 180K Dec  9 13:10 flops.tar

[jpummil@tres-l1 test]$ gzip flops.tar

[jpummil@tres-l1 test]$ ls -lh
total 3.0K
-rw-rw-r-- 1 jpummil jpummil 35K Dec  9 13:07 flops2.c
-rw-rw-r-- 1 jpummil jpummil 35K Dec  9 13:08 flops3.c
-rw-rw-r-- 1 jpummil jpummil 35K Dec  9 13:08 flops4.c
-rw-rw-r-- 1 jpummil jpummil 35K Dec  9 13:08 flops5.c
-rw-rw-r-- 1 jpummil jpummil 35K Dec  9 13:07 flops.c
-rw-rw-r-- 1 jpummil jpummil 37K Dec  9 13:10 flops.tar.gz
```

Linux for HPC

Compressing and archiving files

So now…we have BOTH the originals as well as the compressed archive. What now?

```
[jpummil@tres-l1 test]$ tar tvf flops.tar.gz
-rw-rw-r-- jpummil/jpummil 34942 2015-12-09 13:07 flops2.c
-rw-rw-r-- jpummil/jpummil 34942 2015-12-09 13:08 flops3.c
-rw-rw-r-- jpummil/jpummil 34942 2015-12-09 13:08 flops4.c
-rw-rw-r-- jpummil/jpummil 34942 2015-12-09 13:08 flops5.c
-rw-rw-r-- jpummil/jpummil 34942 2015-12-09 13:07 flops.c

[jpummil@tres-l1 test]$ rm flops*.c

[jpummil@tres-l1 test]$ ls -lh
total 512
-rw-rw-r-- 1 jpummil jpummil 37K Dec  9 13:21 flops.tar.gz
```

One neat, tidy bundle…and about 20% of the original disk usage!

To restore when needed, simply ”tar xvzf flops.tar.gz”….

Linux for HPC

Final Exercises

- Use Nano to generate a few unique text files and save them
- Practice copying and renaming the files
- Practice merging the files, grepping for unique information, sorting them in different ways.
- Practice archiving and compressing the files.
- Practice remote copying the files to and from comp.uark.edu

PRACTICE! PRACTICE! PRACTICE!

Linux for HPC

Additional Helpful Linux Links

Command line tips and tricks: http://commandlinefu.com/commands/browse

Learning more command line: http://linuxcommand.org

Linux Shell Scripting Tutorial: http://bash.cyberciti.biz/guide/Main_Page

Advanced Shell Scripting: http://tldp.org/LDP/abs/html

Daily Linux News Articles: http://lxer.com

UofA Linux User Group: http://ualug.uark.edu/about (just sign up and use the email to query other Linux users on campus)

Linux for HPC

User Support Contact for AHPCC

HPC-SUPPORT@listserv.uark.edu

AHPCC Web Site

http://hpc.uark.edu

AHPCC on Twitter

@ua_ahpcc