CS615 - Aspects of System Administration

Filesystems, Disks, Storage

Department of Computer Science
Stevens Institute of Technology
Jan Schaumann
jschauma@stevens.edu
https://stevens.netmeister.org/615/
Current Events

https://dnsflagday.net/
Current Events

https://is.gd/mHmoLu/
https://is.gd/2r3DNp
Current Events

https://techcrunch.com/2019/01/29/facebook-project-atlas/
Staying up to date

- https://www.devopsweekly.com/
- https://sreweekly.com/
- https://www.nanog.org/
- https://puck.nether.net/mailman/listinfo/outages
- @nixCraft https://is.gd/cKEpWc
- https://www.cronweekly.com/
Let's review HW1

Running an instance:

```bash
$ aws ec2 run-instances --instance-type t1.micro \
    --key-name stevens --image-id ami-569ed93c
```
Let’s review HW1

Save yourself some typing:

$ alias instance='aws ec2 run-instances --key-name stevens --image-id'
$ instance ami-569ed93c --instance-type t1.micro
Let's review HW1

Make it permanent:

$ echo "alias instance='aws ec2 run-instances --key-name stevens --image-id'" \
   >> ~/.bashrc
$ . ~/.bashrc
$ alias
alias instance='aws ec2 run-instances --key-name stevens --image-id'
$
Let's review HW1

```
$ alias
alias console='aws ec2 get-console-output --instance-id'
alias hinstances='instances | grep ^INSTANCES | cut -f15'
alias instance='aws ec2 run-instances --key-name stevens --image-id'
alias instance-names='aws ec2 describe-instances --query Reservations[*].Instances[*].[PublicD
alias instances='aws ec2 describe-instances'
alias running='aws ec2 describe-instances --query Reservations[*].Instances[*].[InstanceId]
alias start-debian='instance ami-628ad918 --instance-type t2.micro'
alias start-fedora='instance ami-0d7a9cc499e108f74 --instance-type t1.micro'
alias start-freebsd='instance ami-d0b520b8 --instance-type t1.micro'
alias start-netbsd='instance ami-569ed93c --instance-type t1.micro'
alias start-netbsd-euw1='aws ec2 run-instances --key-name stevens-euw1 --image-id ami-a460c5d7
alias start-omnios='instance ami-0b2c7fd1c1f7e91d6 --instance-type t1.micro'
alias start-ubuntu='instance ami-78cef802 --instance-type t1.micro'
alias term-instances='aws ec2 terminate-instances --instance-ids'
$```

Lecture 02: Filesystems, Disks, Storage

February 5, 2019
Let's review HW1

ssh to an instance:

$ ssh -i ~/.ssh/ec2 root@<mumble>.compute-1.amazonaws.com
Let's review HW1

Let's save ourselves some typing:

$ cat >>~/.ssh/config <<EOF
> Host *.amazonaws.com
>     IdentityFile ~/.ssh/ec2
>     User root
> EOF
$ ssh <mumble>.compute-1.amazonaws.com
Let's review HW1

How do we know what host we’re connecting to?

$ ssh ec2-3-85-193-42.compute-1.amazonaws.com
The authenticity of host 'ec2-3-85-193-42.compute-1.amazonaws.com (3.85.193.42)' can't be established.
ECDSA key fingerprint is SHA256:lUB2XhoRshItvOVqrY05Lo1Mqzmz69DUP++GH3Yy1P0I.
Are you sure you want to continue connecting (yes/no)?
Let's review HW1

$ aws ec2 get-console-output --instance-id i-0aac317689367a7b8

[...]

e2: #################################################################################################################
e2: -----BEGIN SSH HOST KEY FINGERPRINTS-----
e2: 1024 SHA256:z5n5c0Pe0Kxhrw2Uxj6SY/kwTSk5IEZtveCMvUoBwHc root@ip-10-141-117-253.ec2.internal

e2: 521 SHA256:lUB2XhoRshItv0VqrY05Lo1MqmzG9DUP++GH3Yy1P0I root@ip-10-141-117-253.ec2.internal

e2: 256 SHA256:zENF+3vI3WWJC3iutGYypF3bKFCcJuTTmdDBRJJ9S14 root@ip-10-141-117-253.ec2.internal

e2: 2048 SHA256:PICqhhQyGFsjnw+TosFTY8cXcif2G9R+MxDxrhgaNdk root@ip-10-141-117-253.ec2.internal

e2: -----END SSH HOST KEY FINGERPRINTS-----
e2: #################################################################################################################
Let's review HW1

How do we know what host we're connecting to?

$ aws ec2 get-console-output --instance-id i-0aac317689367a7b8 | grep "^(ECDSA)" 
ec2: 521 SHA256:lUB2XhoRshItvOVqrY05Lo1Mqmz69DUP++GH3Yy1P0I root@ip-10-141-117-253.ec2.internal
$ ssh ec2-3-85-193-42.compute-1.amazonaws.com
The authenticity of host 'ec2-3-85-193-42.compute-1.amazonaws.com (3.85.193.42)' can't be established.
ECDSA key fingerprint is SHA256:lUB2XhoRshItvOVqrY05Lo1Mqmz69DUP++GH3Yy1P0I.
Are you sure you want to continue connecting (yes/no)? yes
Warning: Permanently added 'ec2-3-85-193-42.compute-1.amazonaws.com,3.85.193.42' (ECDSA)
NetBSD 7.0 (XEN3PAE_DOMU.201509250726Z)
Welcome to NetBSD - Amazon EC2 image!

[...]
We recommend that you create a non-root account and use su(1) for root access.
ip-10-141-117-253#
Let's review HW1

# uname -a
NetBSD ip-10-141-117-253.ec2.internal 7.0 NetBSD 7.0 (XEN3PAE_DOMU.201509250726Z) i38
Let's review HW1

```bash
ip-10-141-117-253# ifconfig -a
xennet0: flags=8863<UP,BROADCAST,NOTRAILERS,Running,Simplex,Multicast> mtu 1500
    capabilities=2800<TCP4CSUM_Tx,UDP4CSUM_Tx>
    enabled=0
    address: 22:00:0a:8d:75:fd
    inet 10.141.117.253 netmask 0xffffff00 broadcast 10.141.117.255
    inet6 fe80::531e:e93c:d231:2b50%xennet0 prefixlen 64 scopeid 0x1
lo0: flags=8049<UP,Loopback,Running,Multicast> mtu 33184
    inet 127.0.0.1 netmask 0xff000000
    inet6 ::1 prefixlen 128
    inet6 fe80::1%lo0 prefixlen 64 scopeid 0x2
```
Let’s review HW1

Active Internet connections (including servers)

<table>
<thead>
<tr>
<th>Proto</th>
<th>Recv-Q</th>
<th>Send-Q</th>
<th>Local Address</th>
<th>Foreign Address</th>
<th>State</th>
</tr>
</thead>
<tbody>
<tr>
<td>tcp</td>
<td>0</td>
<td>0</td>
<td>10.141.117.253.22</td>
<td>207.172.174.21.38764</td>
<td>ESTABLISHED</td>
</tr>
<tr>
<td>tcp</td>
<td>0</td>
<td>0</td>
<td>*.22</td>
<td><em>.</em></td>
<td>LISTEN</td>
</tr>
<tr>
<td>udp</td>
<td>0</td>
<td>0</td>
<td>*.68</td>
<td><em>.</em></td>
<td></td>
</tr>
</tbody>
</table>

Active Internet6 connections (including servers)

<table>
<thead>
<tr>
<th>Proto</th>
<th>Recv-Q</th>
<th>Send-Q</th>
<th>Local Address</th>
<th>Foreign Address</th>
<th>(state)</th>
</tr>
</thead>
<tbody>
<tr>
<td>tcp6</td>
<td>0</td>
<td>0</td>
<td>*.22</td>
<td><em>.</em></td>
<td>LISTEN</td>
</tr>
</tbody>
</table>

Active UNIX domain sockets

<table>
<thead>
<tr>
<th>Address</th>
<th>Type</th>
<th>Recv-Q</th>
<th>Send-Q</th>
<th>Inode</th>
<th>Conn</th>
<th>Refs</th>
<th>Nextref</th>
<th>Addr</th>
</tr>
</thead>
<tbody>
<tr>
<td>c1dc3378</td>
<td>stream</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>c1dc33c8</td>
<td>0</td>
<td>0</td>
<td>[...]</td>
</tr>
</tbody>
</table>
Let's review HW1

```
ip-10-141-117-253# df -hi
Filesystem   Size  Used  Avail  %Cap  iUsed  iAvail  %iCap Mounted on
/dev/xbd1a   959M  485M  426M  53%    23240  507254 4%    /
/dev/xbd0a   246M  2.2M  231M  0%     15     65521 0%    /grub
kernfs       1.0K  1.0K     0B  100%  0      0     0%    /kern
ptyfs        1.0K  1.0K     0B  100%  0      0     0%    /dev/pts
procfs       4.0K  4.0K     0B  100%  15     517     2%    /proc
```

```
ip-10-141-117-253# mount
/dev/xbd1a on / type ffs (local)
/dev/xbd0a on /grub type ext2fs (local)
kernfs on /kern type kernfs (local)
ptyfs on /dev/pts type ptyfs (local)
procfs on /proc type procfs (local)
```
Let's review HW1

```bash
# fdisk /dev/xbd0
fdisk: primary partition table invalid, no magic in sector 0
fdisk: Cannot determine the number of heads
Disk: /dev/xbd0d
NetBSD disklabel disk geometry:
cylinders: 1024, heads: 1, sectors/track: 2048 (2048 sectors/cylinder)
total sectors: 2097152, bytes/sector: 512

BIOS disk geometry:
cylinders: 130, heads: 255, sectors/track: 63 (16065 sectors/cylinder)
total sectors: 2097152

Partitions aligned to 16065 sector boundaries, offset 63

Partition table:
0: <UNUSED>
1: <UNUSED>
2: <UNUSED>
3: <UNUSED>
Bootselector disabled.
No active partition.
Drive serial number: 0 (0x00000000)
```
Let's review HW1

# disklabel /dev/rxbd0
# /dev/rxbd0:
type: ESDI
disk: Xen Virtual ESDI
label: fictitious
flags:
bytes/sector: 512
sectors/track: 2048
tracks/cylinder: 1
sectors/cylinder: 2048
cylinders: 1024
total sectors: 2097152
rpm: 3600
interleave: 1
trackskew: 0
cylinderskew: 0
headswitch: 0         # microseconds
track-to-track seek: 0  # microseconds
drivedata: 0

4 partitions:
#     size   offset  ftype [filesize bsize cpg/sgs]
a: 2097152 0 4.2BSD 0 0 0  # (Cyl. 0 - 1023)
d: 2097152 0 unused 0 0 0  # (Cyl. 0 - 1023)
disklabel: boot block size 0
disklabel: super block size 0
Filesystems, Disks, Storage

$ ssh linux-lab.cs.stevens.edu
$ df -hT

$ dd if=/dev/zero of=/tmp/big bs=1G
[...]

Now try to ssh to that host...
File sizes are not always what they seem to be.

```sh
$ mkdir /tmp/${USER}
$ export LARGE=/tmp/${USER}/large
$ truncate -s $(df /tmp | awk '/^// { print $4; }')0000 ${LARGE}
$ ls -l ${LARGE}
$ du ${LARGE}
$ stat ${LARGE}
$ cp ${LARGE} ${LARGE}2
$ du ${LARGE}2
$ cat ${LARGE} > ${LARGE}2
$ ls -l ${LARGE}* 
$ du ${LARGE}*
```
Filesystems, Disks, Storage

How many files can be created on /tmp?

```bash
$ ssh linux-lab.cs.stevens.edu
$ df -i /tmp
$ rm /tmp/${USER}/large2
$ df -i /tmp
```
Filesystems, Disks, Storage

$ ssh linux-lab.cs.stevens.edu
$ cd /tmp
$ df -i /tmp
$ touch newfile
$ cc -Wall ~jschauma/tmp/mkfiles.c
$ ./a.out /tmp/${USER}
$ ls -ld /tmp/${USER}
$ ls /tmp/${USER} | wc -l
$ touch newerfile
$ echo "hello hello hello" >> newfile
$ rm -fr /tmp/${USER}

See also: https://is.gd/nX07RR
Filesists, Disks, Storage

Important lessons:

File sizes are not always what they seem to be.

Error messages aren’t always what they seem to be!

All resources are finite.
Team Missions

Western Digital WD BLACK 3.5" PC HARD DRIVE
Western Digital WD BLUE 2.5" PC HARD DRIVE
Western Digital WD GREEN SATA PC SSD Solid State Drive
Western Digital WD RED NAS DRIVE
Filesystems, Disks, Storage

- basic disk concepts
- basic filesystem concepts
- file systems
Topics covered

- basic disk concepts
  - storage models
  - disk interfaces
  - physical disk structure
  - partitions
- basic filesystem concepts
- file systems
Topics covered

- basic disk concepts
  - storage models
  - disk interfaces
  - physical disk structure
  - partitions
- basic filesystem concepts
  - RAID
  - logical volume management
  - device formatting
- file systems
Topics covered

- basic disk concepts
  - storage models
  - disk interfaces
  - physical disk structure
  - partitions
- basic filesystem concepts
  - RAID
  - logical volume management
  - device formatting
- file systems
  - the UNIX filesystem or Berkeley Fast File System (FFS)
Basic Disk Concepts

Storage Models
Basic Disk Concepts: Storage Models

Direct Attached Storage (DAS)

```
ssh lab 'df -hiT /
```
Network Attached Storage (NAS)

ssh lab 'df -hiT /home/$(whoami)'

Lecture 02: Filesystems, Disks, Storage
February 5, 2019
Basic Disk Concepts: Storage Models

Storage Area Networks (SAN)
Basic Disk Concepts: Storage Models

Cloud Storage (Examples: EBS, S3)
Team Missions
Basic Disk Concepts: Storage Models: Cloud Storage

$ aws ec2 describe-instances
[...]
/dev/sda1 ebs None paravirtual
BLOCKDEVICEMAPPINGS /dev/sda
EBS 2014-01-25T20:18:19.000Z True attached vol-a0d000d6
[...]
create two instances
create a 1GB volume and attach it to one of the instances
create a new filesystem on the volume and mount it
create a file on the new filesystem
terminate the first instance
attach the volume to the second instance
retrieve the file from the volume via the second instance

If time permits, repeat using a Linux instance. Useful commands:

aws ec2 create-volume, aws ec2 attach-volume, format(1M),
newfs(1M), mount(8)
Basic Disk Concepts: Storage Models: Cloud Storage

$ aws ec2 create-volume --size 1 --availability-zone us-east-1d

 [...] 

$ aws ec2 attach-volume --volume-id vol-9d3aeaeb --instance-id \  i-dd74f0f3 --device /dev/sdh

 [...]

$ ssh <hostname>

# format

format> fdisk

format> verify

format> label

# newfs /dev/rdsk/c1t2160d0s0

[...]

# mount /dev/dsk/c1t2160d0s0 /mnt

# df -Th /mnt

[...]

# fstyp -v /dev/rdsk/c1t2160d0s0  | more

[...]
Basic Disk Concepts

Disk Devices
Basic Disk Concepts: Disk Devices
Basic Disk Concepts: Disk Devices

Security affects everything.

https://is.gd/bK0rwd
Basic Disk Concepts: Disk Devices
Basic Disk Concepts: Disk Devices
Basic Disk Concepts: Disk Devices
Basic Disk Concepts: Disk Devices
Basic Disk Concepts: Disk Devices
Basic Disk Concepts

Disk Interfaces
Basic Disk Concepts: Disk Interfaces: SCSI
Basic Disk Concepts: Disk Interfaces: ATA
Basic Disk Concepts: Disk Interfaces: ATA
Basic Disk Concepts: Disk Interfaces: Fibre Channel
Basic Disk Concepts: Disk Interfaces: Fibre Channel
Basic Disk Concepts: Disk Interfaces: Fibre Channel

- Fabric
- Loop
- Point-to-Point
Basic Disk Concepts: Disk Interfaces: Fibre Channel
Basic Disk Concepts: Disk Interfaces: Fibre Channel
Basic Disk Concepts: Disk Interfaces: SANs

- **ATA over Ethernet (AoE):**
  - create low-cost SAN
  - ATA encapsulated into Ethernet frames

- **Fibre Channel over Ethernet (FCoE):**
  - consolidate IP and FC/SAN networks
  - FC encapsulated into Ethernet frames

- ***oE:**
  - no TCP/IP overhead
  - restricted to a single Layer 2 network
  - no inherent security features

- **iSCSI**
  - SCSI encapsulated in TCP/IP packets
Team Missions
Basic Disk Concepts

Physical Disk Structure
Basic Disk Concepts: Disk Devices
Basic Disk Concepts: Disk Devices
Basic Disk Concepts: Disk Devices
Basic Disk Concepts: Physical Disk Structure
Basic Disk Concepts: Physical Disk Structure

Hard drive performance determined by:

- seek time
- rotational latency
- internal data rate
- a few other negligible factors (external data rate, command overhead, access time, etc.)
Basic Disk Concepts: Disk Devices
Basic Disk Concepts

Partitions
Basic Disk Concepts: Partitions

- Track/Cylinder
- Sector
- Heads: 8 Heads, 4 Platters
# Basic Disk Concepts: Partitions

NetBSD example (from `disklabel(8)`)

- Partition 'a': `/`
- Partition 'b': `swap`
- Partition 'e': `/home`

<table>
<thead>
<tr>
<th>#</th>
<th>size</th>
<th>offset</th>
<th>fstype</th>
<th>[fsize bsize cpg/sgs]</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>20972385</td>
<td>63</td>
<td>4.2BSD</td>
<td>4096 32768 1180 # (Cyl. 0*- 20805)</td>
</tr>
<tr>
<td>b</td>
<td>1048320</td>
<td>20972448</td>
<td>swap</td>
<td># (Cyl. 20806 - 21845)</td>
</tr>
<tr>
<td>c</td>
<td>78140097</td>
<td>63</td>
<td>unused</td>
<td>0 0 # (Cyl. 0*- 77519)</td>
</tr>
<tr>
<td>d</td>
<td>78140160</td>
<td>0</td>
<td>unused</td>
<td>0 0 # (Cyl. 0 - 77519)</td>
</tr>
<tr>
<td>e</td>
<td>56119392</td>
<td>22020768</td>
<td>4.2BSD</td>
<td>4096 32768 58528 # (Cyl. 21846 - 77519)</td>
</tr>
</tbody>
</table>
Basic Disk Concepts: Partitions

NetBSD example (from disklabel(8))
Partition 'a': / 10 GB
Partition 'b': swap
Partition 'e': /home 26 GB

<table>
<thead>
<tr>
<th>#</th>
<th>size</th>
<th>offset</th>
<th>fstype</th>
<th>[fsize bsize cpg/sgs]</th>
</tr>
</thead>
<tbody>
<tr>
<td>a:</td>
<td>20972385</td>
<td>63</td>
<td>4.2BSD</td>
<td>4096 32768 1180      # (Cyl. 0*- 20805)</td>
</tr>
<tr>
<td>b:</td>
<td>1048320</td>
<td>20972448</td>
<td>swap</td>
<td># (Cyl. 20806 - 21845)</td>
</tr>
<tr>
<td>c:</td>
<td>78140097</td>
<td>63</td>
<td>unused</td>
<td>0 0                   # (Cyl. 0* - 77519)</td>
</tr>
<tr>
<td>d:</td>
<td>78140160</td>
<td>0</td>
<td>unused</td>
<td>0 0                   # (Cyl. 0 - 77519)</td>
</tr>
<tr>
<td>e:</td>
<td>56119392</td>
<td>22020768</td>
<td>4.2BSD</td>
<td>4096 32768 58528     # (Cyl. 21846 - 77519)</td>
</tr>
</tbody>
</table>
## Basic Disk Concepts: Partitions

**Solaris example (from `format(1m)`):**

Current partition table (original):
Total disk cylinders available: 38758 + 2 (reserved cylinders)

<table>
<thead>
<tr>
<th>Part</th>
<th>Tag</th>
<th>Flag</th>
<th>Cylinders</th>
<th>Size</th>
<th>Blocks</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>root</td>
<td>wm</td>
<td>3 - 3764</td>
<td>3.62GB</td>
<td>(3762/0/0) 7584192</td>
</tr>
<tr>
<td>1</td>
<td>swap</td>
<td>wu</td>
<td>3765 - 4364</td>
<td>590.62MB</td>
<td>(600/0/0) 1209600</td>
</tr>
<tr>
<td>2</td>
<td>backup</td>
<td>wm</td>
<td>0 - 38757</td>
<td>37.26GB</td>
<td>(38758/0/0) 78136128</td>
</tr>
<tr>
<td>3</td>
<td>unassigned</td>
<td>wm</td>
<td>0</td>
<td>0</td>
<td>(0/0/0) 0</td>
</tr>
<tr>
<td>4</td>
<td>unassigned</td>
<td>wm</td>
<td>0</td>
<td>0</td>
<td>(0/0/0) 0</td>
</tr>
<tr>
<td>5</td>
<td>unassigned</td>
<td>wm</td>
<td>0</td>
<td>0</td>
<td>(0/0/0) 0</td>
</tr>
<tr>
<td>6</td>
<td>unassigned</td>
<td>wm</td>
<td>0</td>
<td>0</td>
<td>(0/0/0) 0</td>
</tr>
<tr>
<td>7</td>
<td>home</td>
<td>wm</td>
<td>4365 - 38757</td>
<td>33.06GB</td>
<td>(34393/0/0) 69336288</td>
</tr>
<tr>
<td>8</td>
<td>boot</td>
<td>wu</td>
<td>0</td>
<td>0.98MB</td>
<td>(1/0/0) 2016</td>
</tr>
<tr>
<td>9</td>
<td>alternates</td>
<td>wu</td>
<td>1 - 2</td>
<td>1.97MB</td>
<td>(2/0/0) 4032</td>
</tr>
</tbody>
</table>
### Basic Disk Concepts: Partitions

Linux example (from `fdisk(8)`):

Disk /dev/sda: 80.0 GB, 80000000000 bytes
255 heads, 63 sectors/track, 9726 cylinders
Units = cylinders of 16065 * 512 = 8225280 bytes

<table>
<thead>
<tr>
<th>Device</th>
<th>Boot</th>
<th>Start</th>
<th>End</th>
<th>Blocks</th>
<th>Id</th>
<th>System</th>
</tr>
</thead>
<tbody>
<tr>
<td>/dev/sda1</td>
<td>*</td>
<td>1</td>
<td>33</td>
<td>265041</td>
<td>83</td>
<td>Linux</td>
</tr>
<tr>
<td>/dev/sda2</td>
<td></td>
<td>34</td>
<td>9726</td>
<td>77859022+</td>
<td>83</td>
<td>Linux</td>
</tr>
</tbody>
</table>
Basic Disk and Filesystem Concepts: RAID and Logical Volumes

- allow file systems to be larger than the physical size of a disk
- increase I/O performance when *striped*
- fault tolerant when *mirrored* or *plexed*
Team Missions
Hooray!

5 Minute Break
Homework

Repeat the examples from class. Make sure you understand the commands and how they relate to the concepts we discussed. Repeat for a different OS, for example:

- ami-0d7a9cc499e108f74 – Fedora 29
- ami-00e61b69944d5d2a3 – FreeBSD 12.0
- ami-569ed93c – NetBSD 7.0
- ami-0b2c7fd1c1f7e91d6 – OmniOS 5.11

Remember to *shut down* your EC2 instances and to *delete* any unused ESB volumes!
Reading

- https://is.gd/5mndwA
- https://is.gd/ig4QP5
- https://is.gd/9YeIKh
Reading

Disk Interfaces:

- **SCSI:**
  - `scsi(4), scsictl(8)`

- **ATA:**
  - [https://en.wikipedia.org/wiki/Advanced_Technology_Attachment](https://en.wikipedia.org/wiki/Advanced_Technology_Attachment)
Reading

Disk Interfaces:

- **Serial attached SCSI:**

- **Fibre Channel:**
  - https://en.wikipedia.org/wiki/Fibrechannel

- **AoE, FCoE, iSCSI:**
  - https://en.wikipedia.org/wiki/ATA_over_Ethernet
Reading

Basic Disk Concepts:
- disklabel(8), fdisk(8)
- format(1m)

RAID:

Basic Filesystem Concepts:
- https://is.gd/8KHnQj
- https://is.gd/wGgJ0e
- newfs(8)

NFS: https://is.gd/70yqMZ