CS 695

(*) containers.

- chroot / image files
- namespaces
- cgroups — mem — student 99% / faculty 1%
- (Linux) capabilities & secure computing

unshare
setns
chroot
mount

"escape the container"

minimize the attack surface.

(*) A3 is out.

(*) A4 is April.

(*) virtio followup via office hours.

(*) capabilities → root admin & user

- traditionally, bigger
- mechanisms to provide a larger varied set of privileges per process.
- embedded in linux source code.

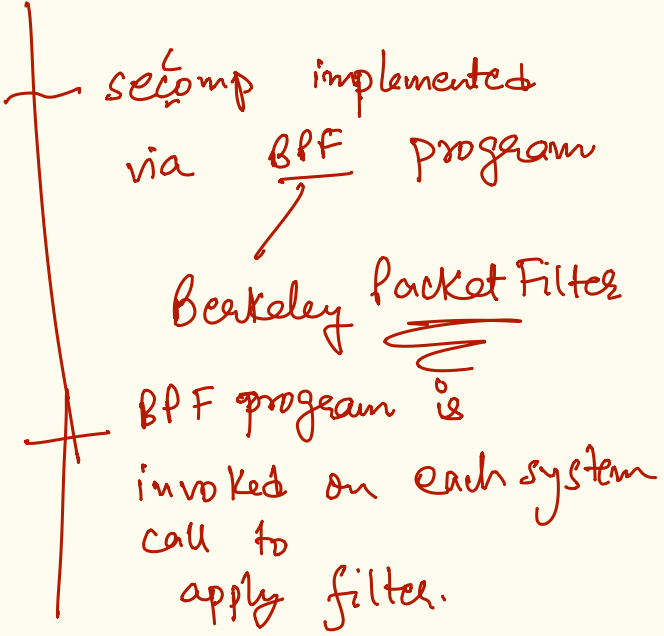
seccomp →

- secure computing
- ~ system call filter
- ~ strict, filter ← modes

strict: 4/5 system calls
- open
- read
- write
- exit
- ..

filter: customizable whitelist / backlist

— system call type & arguments.


seccomp implemented
via BPF program
Berkeley Packet Filter
BPF program is
invoked on each system
call to
apply filter.