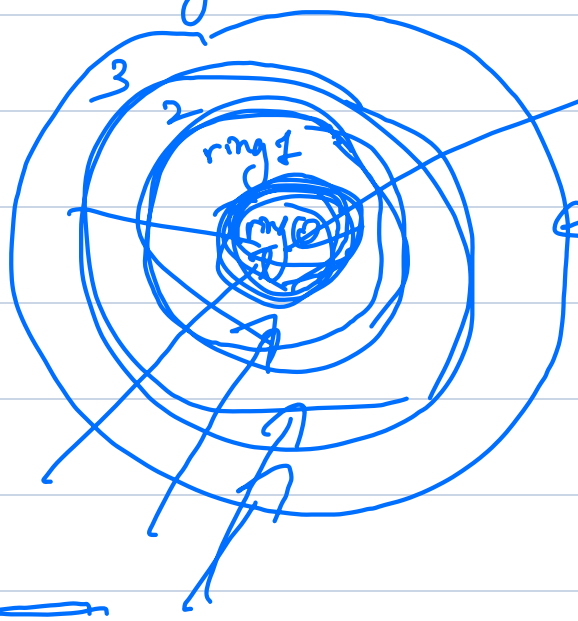
* Privilege Levels



most privilege (kernel)

least privilege (user space)

CS. (code segment) ⇒ last two bits for privilege level

SS. (stack segment) ⇒ stack access level
- kernel stack access
- user stack access
- Ring 1 or 2 stack access

* Modifying CS&SS is a one-way operation from ring 0 to ring 3.

* Going back from ring 3 to ring 0?

Software interrupt — INT 80 — pass argument. rcx. return value.

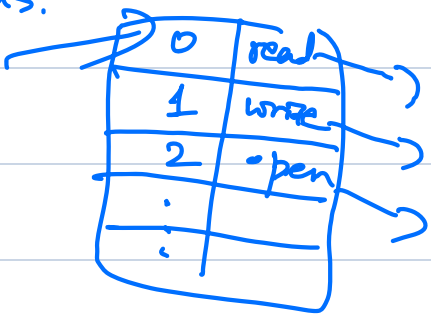
syscall / sysret — similar to interrupts.
- syscall rax. rdi ... pass arguments.
- ring 0 ⇒ syscall handler

MSR — address to the handler. register syscall handler.
↳ rcx : return value.

* Typical OS has many system calls.

read.
write
open.
close
⋮
gettimer
socket
connect
⋮
fork
clone
⋮

syscall 0 ←
1 ←
2 ←
⋮



sysenter / sysexit

* sysret ⇒ ring 3