- Memory Virtualization -

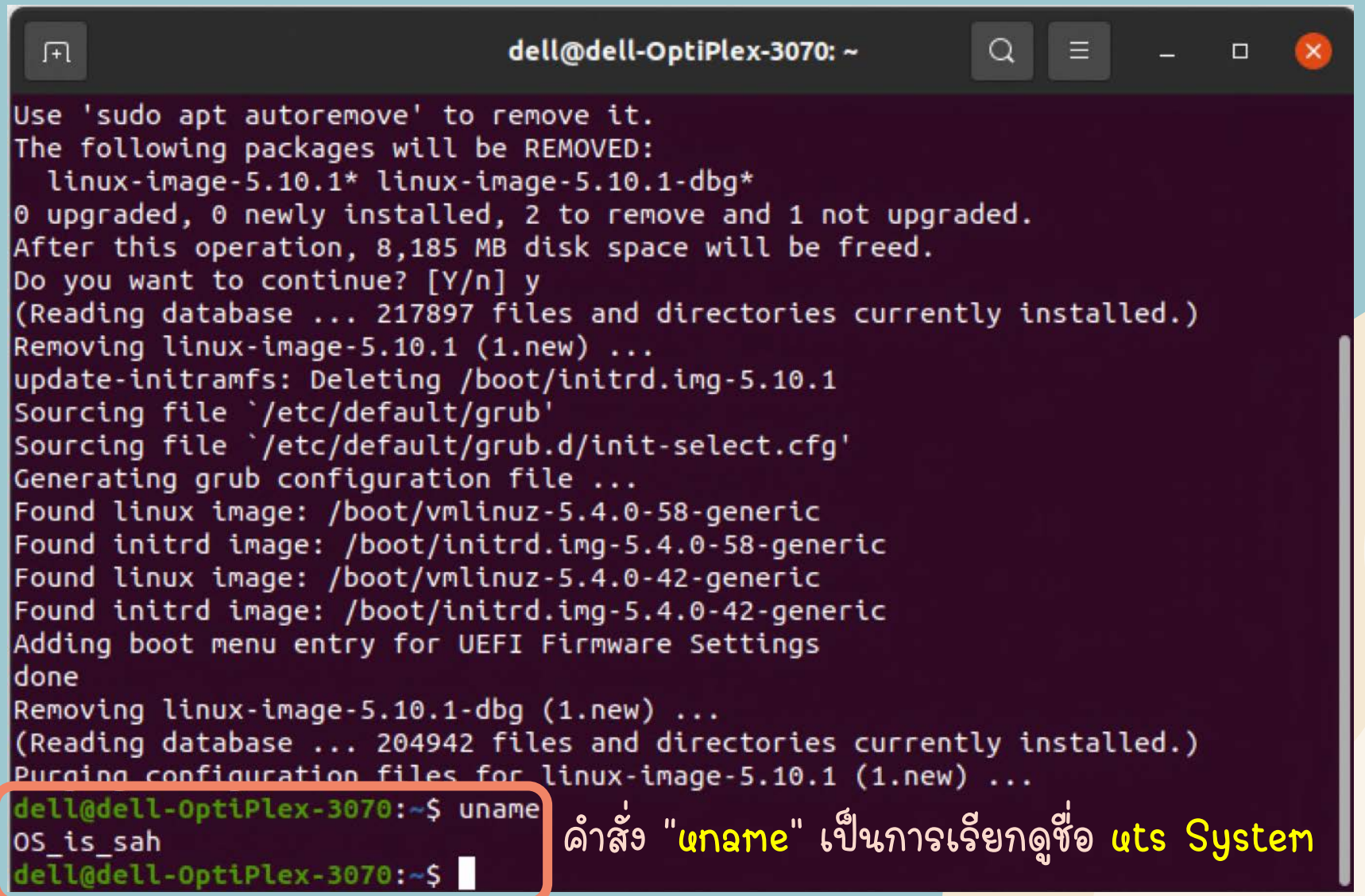
CPE326
Operating System

Group: OS is sah
By 7204_7205_7208_7218_7234

Mini Project

01

การเปลี่ยนชื่อ Kernel บน Linux

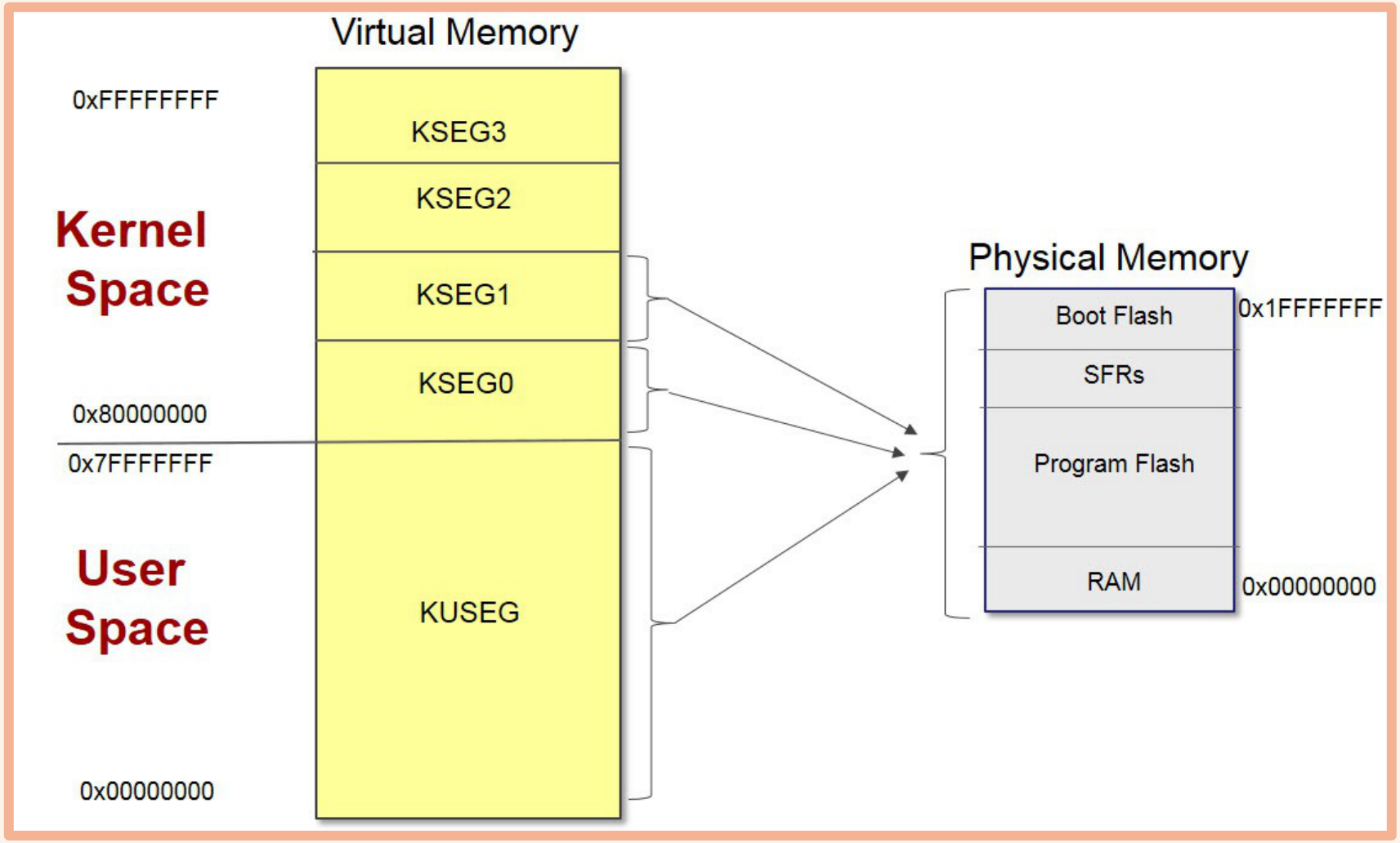


คำสั่ง "uname" เป็นการเรียกดูชื่อ uts System

Mini Project

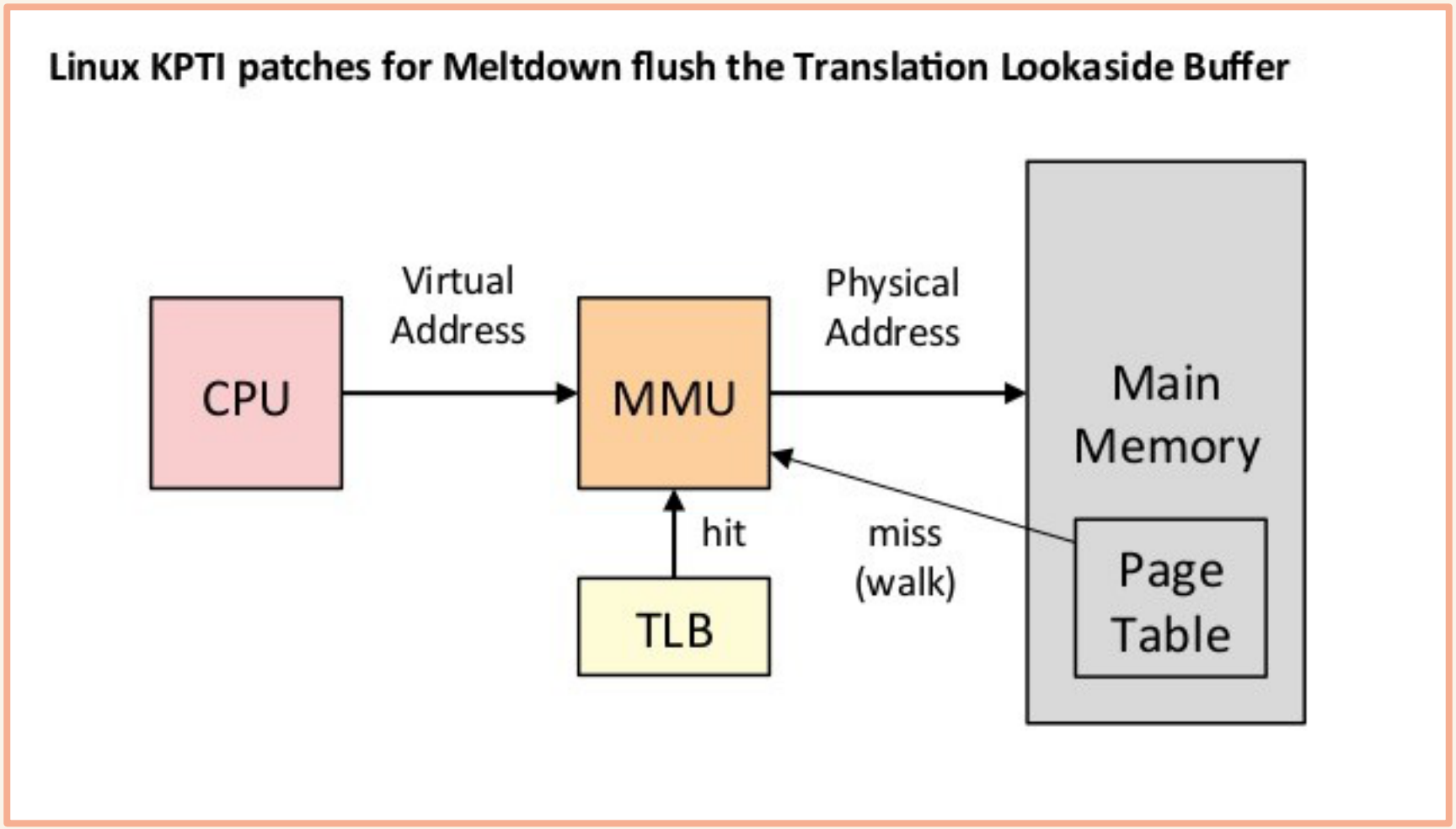
02

Memory virtualization

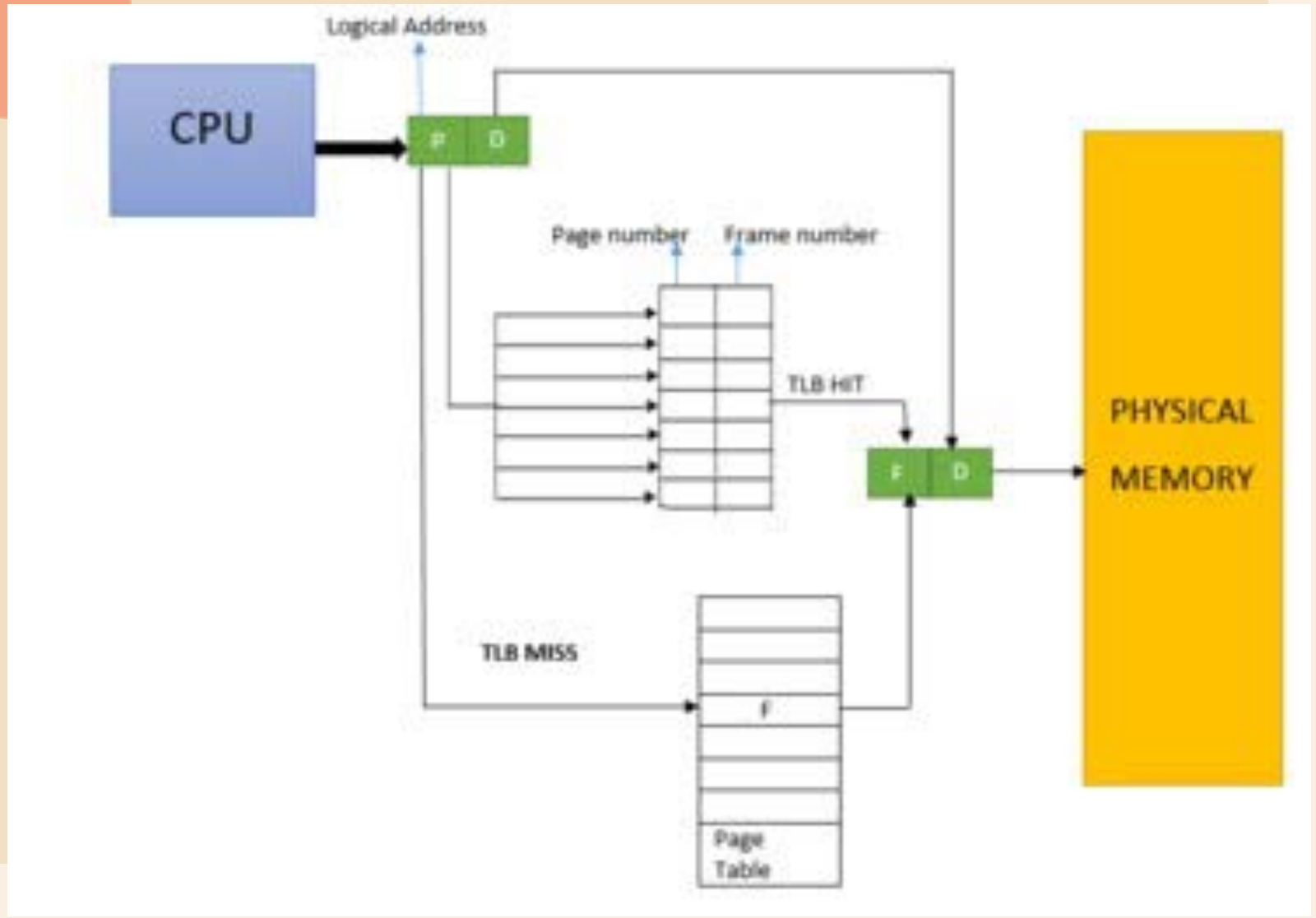


- User Space
- Kernel Space
- Physical Memory

Memory Management Unit (MMU)

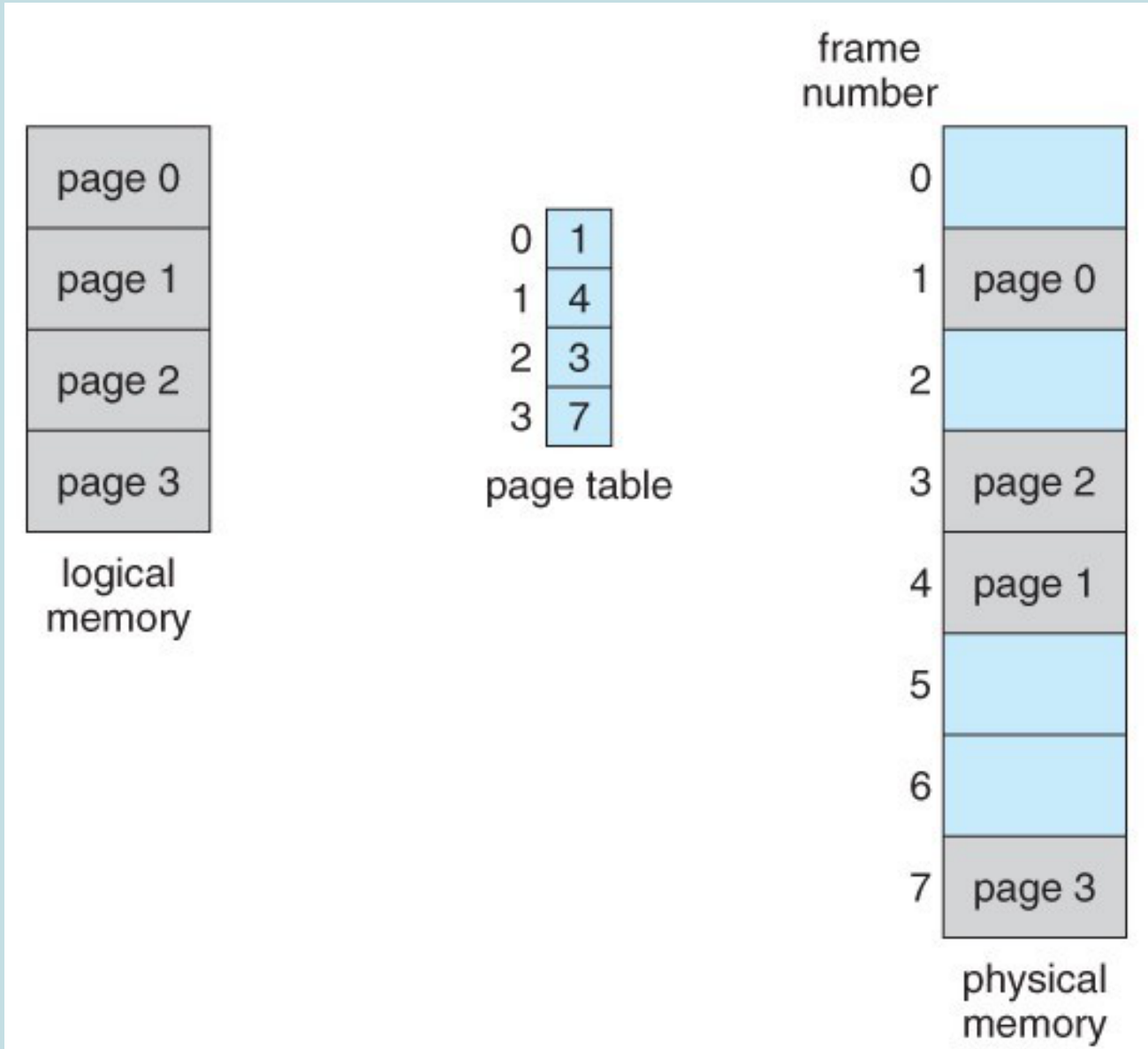


TLB

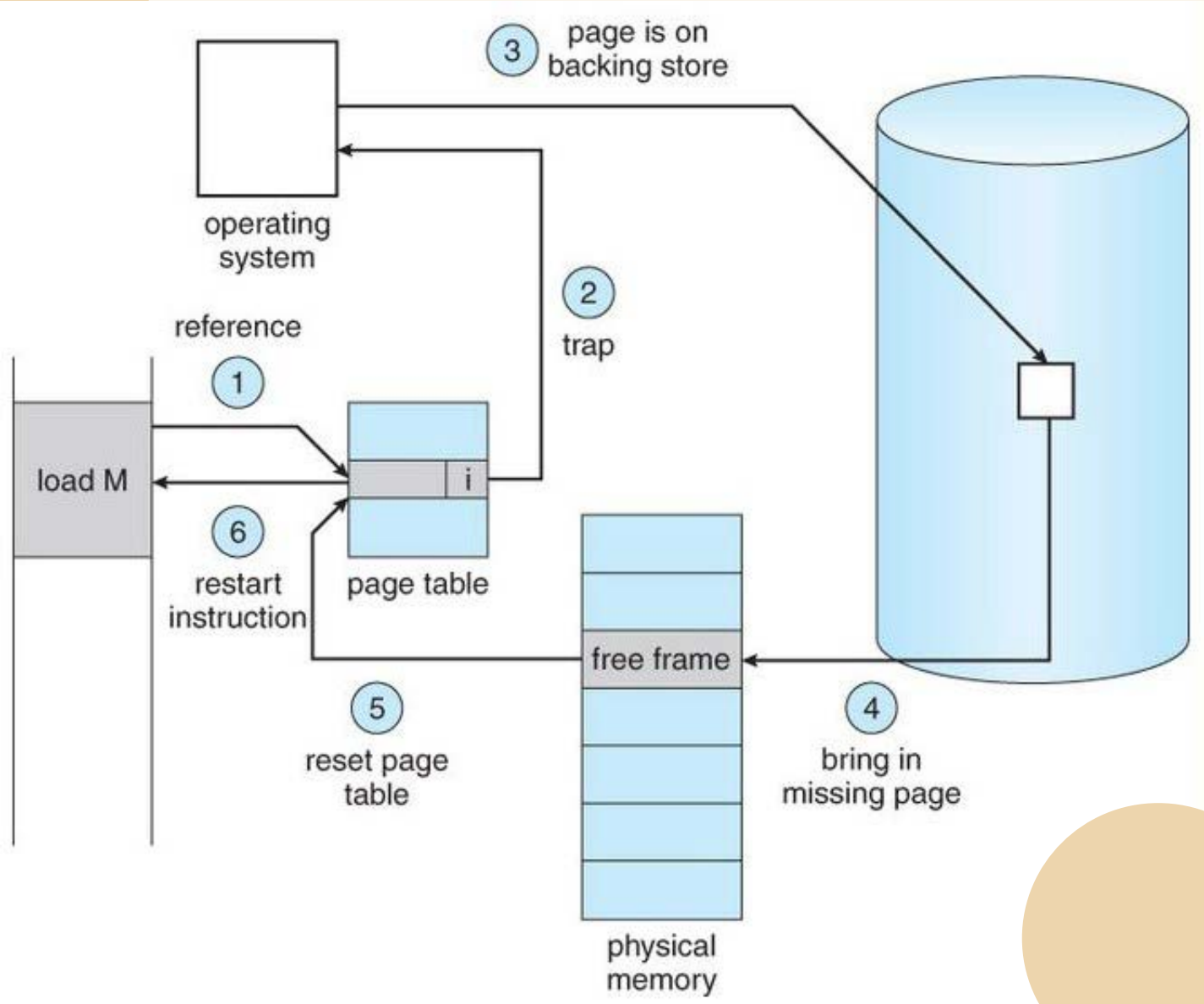

Logical Address
CPU
P
D
Page number
Frame number
TLB HIT
F
D
PHYSICAL
MEMORY
TLB MISS
F
Page
Table

Page

- Default Size ของ Page = 4096 Byte
- Linux page size ($getconf PAGESIZE)



Page
HIT&FAULT



Q&A

01

Where ae your topic source code located in the Ubuntu kernel files?

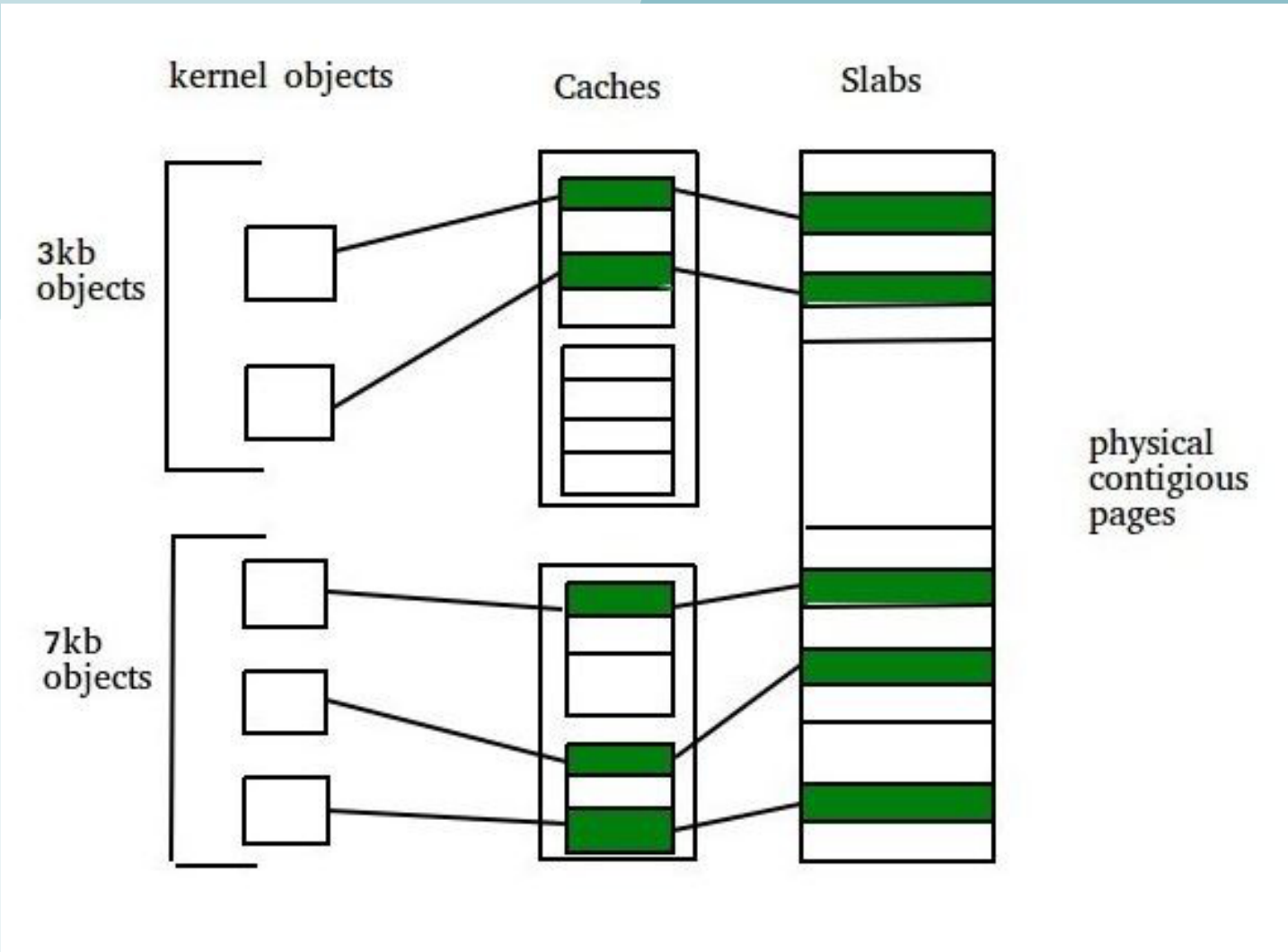
```
dell@dell-OptiPlex-3070:~/kernel/linux-5.10.1$ ls mm
backing-dev.c              memblock.o                 pgalloc-track.h
backing-dev.o              memcontrol.c               pgtable-generic.c
balloon_compaction.c       memcontrol.o               pgtable-generic.o
balloon_compaction.o       memfd.c                    process_vm_access.c
built-in.a                 memfd.o                    process_vm_access.o
cleancache.c               memory.c                   ptdump.c
cleancache.o               memory-failure.c           ptdump.o
cma.c                      memory-failure.o           readahead.c
cma_debug.c                memory_hotplug.c           readahead.o
cma.h                      memory_hotplug.o           rmap.c
cma.o                      memory.o                   rmap.o
compaction.c               mempolicy.c                rodata_test.c
compaction.o               mempolicy.o                shmem.c
debug.c                    mempool.c                  shmem.o
debug.o                    mempool.o                  shuffle.c
debug_page_ref.c           memremap.c                 shuffle.h
debug_vm_pgtable.c         memremap.o                 shuffle.o
dmapool.c                  memtest.c                  slab.c
dmapool.o                  memtest.o                  slab_common.c
early_ioremap.c            migrate.c                  slab_common.o
early_ioremap.o            migrate.o                  slab.h
fadvise.c                  mincore.c                  slob.c
fadvise.o                  mincore.o                  slub.c
failslab.c                 mlock.c                    slub.o
filemap.c                  mlock.o                    sparse.c
filemap.o                  mmap.c                     sparse.o
frame_vector.c             mmap.o                     sparse-vmemmap.c
frame_vector.o             mm_init.c                  sparse-vmemmap.o
frontswap.c                mm_init.o                  swap.c
frontswap.o                mmu_gather.c               swap_cgroup.c
gup_benchmark.c            mmu_gather.o               swap_cgroup.o
gup.c                      mmu_notifier.c             swapfile.c
gup.o                      mmu_notifier.o             swapfile.o
highmem.c                  mmzone.c                   swap.o
highmem.o                  mmzone.o                   swap_slots.c
hmm.c                      modules.order              swap_slots.o
hmm.o                      mprotect.c                 swap_state.c
huge_memory.c              mprotect.o                 swap_state.o
huge_memory.o              mremap.c                   truncate.c
hugetlb.c                  mremap.o                   truncate.o
hugetlb_cgroup.c           msync.c                    usercopy.c
hugetlb_cgroup.o           msync.o                    usercopy.o
hugetlb.o                  nommu.c                    userfaultfd.c
hwpoison-inject.c          oom_kill.c  "oom_killer"   userfaultfd.o
hwpoison-inject.ko         oom_kill.o                 util.c
hwpoison-inject.mod        page_alloc.c               util.o
hwpoison-inject.mod.c      page_alloc.o               vmacache.c
hwpoison-inject.mod.o      page_counter.c             vmacache.o
hwpoison-inject.o          page_counter.o             vmalloc.c
init-mm.c                  page_ext.c                 vmalloc.o
init-mm.o                  page_idle.c                vmpressure.c
internal.h                 page_idle.o                vmpressure.o
interval_tree.c            page_io.c                  vmscan.c
interval_tree.o            page_io.o                  vmscan.o
ioremap.c                  page_isolation.c           vmstat.c
ioremap.o                  page_isolation.o           vmstat.o
kasan                      page_owner.c               workingset.c
Kconfig                    page_poison.c              workingset.o
Kconfig.debug              page_poison.o              z3fold.c
khugepaged.c               page_reporting.c           z3fold.ko
khugepaged.o               page_reporting.h           z3fold.mod
kmemleak.c                 page_reporting.o           z3fold.mod.c
ksm.c                      page_vma_mapped.c          z3fold.mod.o
ksm.o                      page_vma_mapped.o          z3fold.o
list_lru.c                 pagewalk.c                 zbud.c
list_lru.o                 pagewalk.o                 zbud.o
maccess.c                  page-writeback.c           zpool.c
maccess.o                  page-writeback.o           zpool.o
madvise.c                  percpu.c                   zsmalloc.c
madvise.o                  percpu-internal.h          zsmalloc.o
Makefile                   percpu-km.c                zswap.c
mapping_dirty_helpers.c    percpu.o                   zswap.o
mapping_dirty_helpers.o    percpu-stats.c
memblock.c                 percpu-vm.c
```

Q&A

02

Which files are related to the topics?

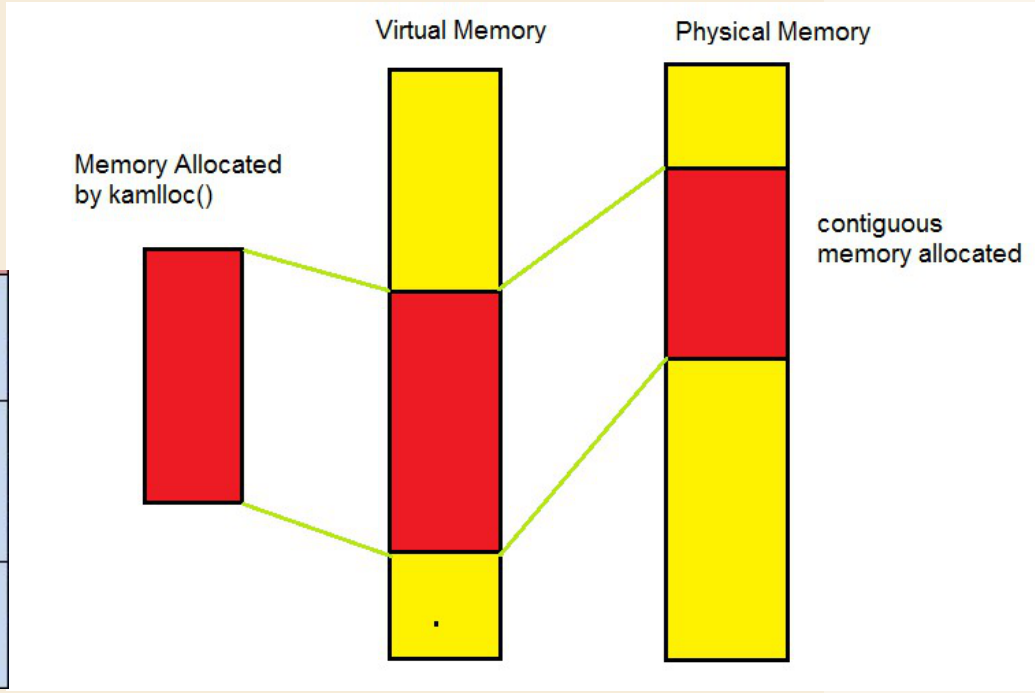
Slab

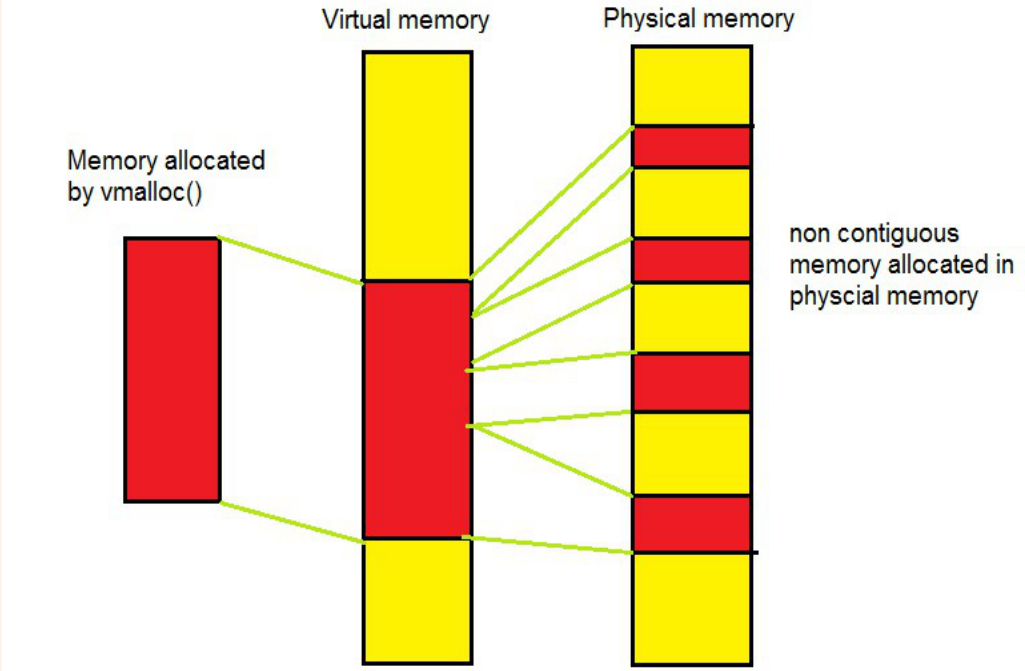

kernel objects
Caches
Slabs
3kb objects
7kb objects
physical contigious pages

Kmalloc & Vmalloc

• Kmalloc

Kmalloc() allocates contiguous memory in physical memory always.
Kmalloc() performance is better than vmalloc().
Kmalloc() with GFP_ATOMIC flag can be used in allocating memory in interrupt handler.



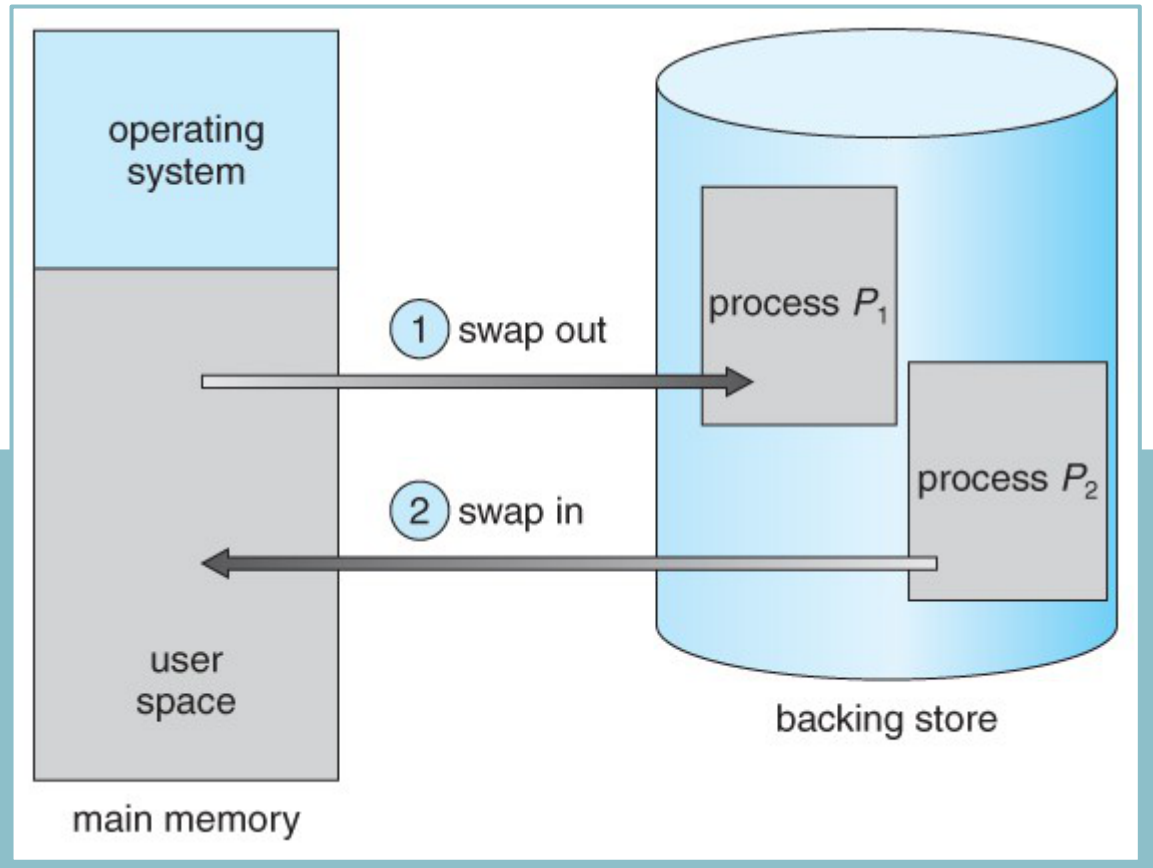


• vmalloc

Vmalloc() doesn't give any assurance of contiguous allocation of memory in physical memory.
Vmalloc() performance is low as compared to kmalloc() as there is overhead of mapping of non-contiguous physical memory in contiguous virtual memory.
Vmalloc() can not be used to allocate memory in interrupt handler.

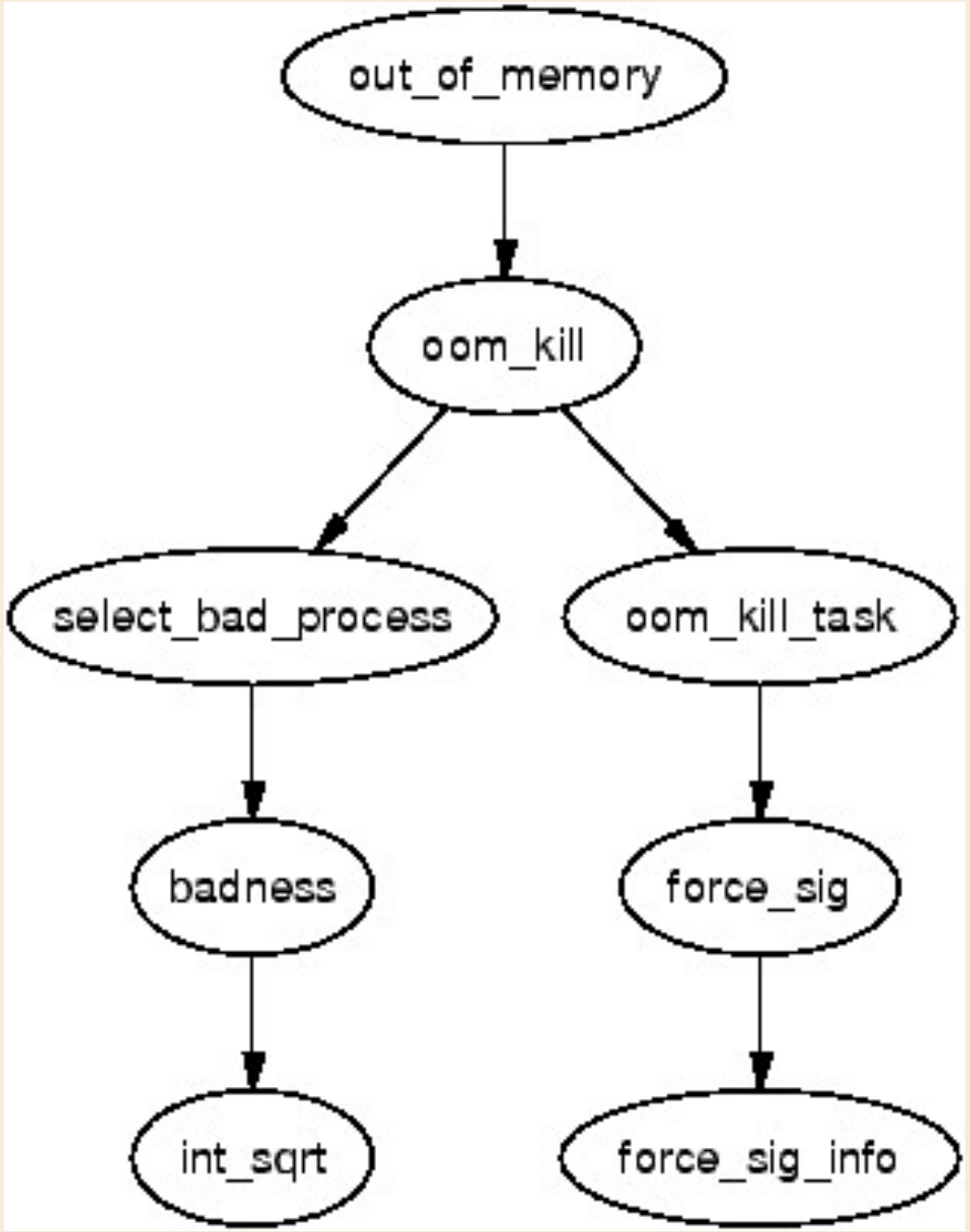
Swapping

- **Swapping out : Remove TLB**



- **Swapping in : Create TLB**

Oom_killer


out_of_memory
oom_kill
select_bad_process
oom_kill_task
badness
force_sig
int_sqrt
force_sig_info

Mini Project

03

Q&A

03

How DO they work?

```
dell@dell-OptiPlex-3070:~/kernel/linux-5.10.1/cpe326$ ./processMaps.out
PID: 10431
command: cat /proc/10431/maps
5619eca6c000-5619eca6d000 r--p 00000000 08:02 52301277                   /home/dell/kernel/linux-5.10.1/cpe326/processMaps.out
5619eca6d000-5619eca6e000 r-xp 00001000 08:02 52301277                   /home/dell/kernel/linux-5.10.1/cpe326/processMaps.out
5619eca6e000-5619eca6f000 r--p 00002000 08:02 52301277                   /home/dell/kernel/linux-5.10.1/cpe326/processMaps.out
5619eca6f000-5619eca70000 r--p 00002000 08:02 52301277                   /home/dell/kernel/linux-5.10.1/cpe326/processMaps.out
5619eca70000-5619eca71000 rw-p 00003000 08:02 52301277                   /home/dell/kernel/linux-5.10.1/cpe326/processMaps.out
5619ed9e8000-5619eda09000 rw-p 00000000 00:00 0                          [heap]
7f8321f38000-7f8321f5d000 r--p 00000000 08:02 14682249                   /usr/lib/x86_64-linux-gnu/libc-2.31.so
7f8321f5d000-7f83220d5000 r-xp 00025000 08:02 14682249                   /usr/lib/x86_64-linux-gnu/libc-2.31.so
7f83220d5000-7f832211f000 r--p 0019d000 08:02 14682249                   /usr/lib/x86_64-linux-gnu/libc-2.31.so
7f832211f000-7f8322120000 ---p 001e7000 08:02 14682249                   /usr/lib/x86_64-linux-gnu/libc-2.31.so
7f8322120000-7f8322123000 r--p 001e7000 08:02 14682249                   /usr/lib/x86_64-linux-gnu/libc-2.31.so
7f8322123000-7f8322126000 rw-p 001ea000 08:02 14682249                   /usr/lib/x86_64-linux-gnu/libc-2.31.so
7f8322126000-7f832212c000 rw-p 00000000 00:00 0
7f832213f000-7f8322140000 r--p 00000000 08:02 14682245                   /usr/lib/x86_64-linux-gnu/ld-2.31.so
7f8322140000-7f8322163000 r-xp 00001000 08:02 14682245                   /usr/lib/x86_64-linux-gnu/ld-2.31.so
7f8322163000-7f832216b000 r--p 00024000 08:02 14682245                   /usr/lib/x86_64-linux-gnu/ld-2.31.so
7f832216c000-7f832216d000 r--p 0002c000 08:02 14682245                   /usr/lib/x86_64-linux-gnu/ld-2.31.so
7f832216d000-7f832216e000 rw-p 0002d000 08:02 14682245                   /usr/lib/x86_64-linux-gnu/ld-2.31.so
7f832216e000-7f832216f000 rw-p 00000000 00:00 0
7ffccb3f1000-7ffccb412000 rw-p 00000000 00:00 0                          [stack]
7ffccb586000-7ffccb589000 r--p 00000000 00:00 0                          [vvar]
7ffccb589000-7ffccb58a000 r-xp 00000000 00:00 0                          [vdso]
ffffffffff600000-ffffffffff601000 --xp 00000000 00:00 0                  [vsyscall]
```

```
dell@dell-OptiPlex-3070:~/kernel/linux-5.10.1/cpe326$ ./demorun.out &
[2] 10467
dell@dell-OptiPlex-3070:~/kernel/linux-5.10.1/cpe326$ pmap 10467 -x
10467:   ./demorun.out
Address           Kbytes     RSS   Dirty Mode  Mapping
000055b3364c5000       4       4       0 r---- demorun.out
000055b3364c6000       4       4       0 r-x-- demorun.out
000055b3364c7000       4       0       0 r---- demorun.out
000055b3364c8000       4       4       4 r---- demorun.out
000055b3364c9000       4       4       4 rw--- demorun.out
00007f69e3955000     148     144       0 r---- libc-2.31.so
00007f69e397a000    1504     524       0 r-x-- libc-2.31.so
00007f69e3af2000     296      64       0 r---- libc-2.31.so
00007f69e3b3c000       4       0       0 ----- libc-2.31.so
00007f69e3b3d000      12      12      12 r---- libc-2.31.so
00007f69e3b40000      12      12      12 rw--- libc-2.31.so
00007f69e3b43000      24      20      20 rw---   [ anon ]
00007f69e3b5c000       4       4       0 r---- ld-2.31.so
00007f69e3b5d000     140     140       0 r-x-- ld-2.31.so
00007f69e3b80000      32      32       0 r---- ld-2.31.so
00007f69e3b89000       4       4       4 r---- ld-2.31.so
00007f69e3b8a000       4       4       4 rw--- ld-2.31.so
00007f69e3b8b000       4       4       4 rw---   [ anon ]
00007fff72ee2000     132      16      16 rw---   [ stack ]
00007fff72f39000      12       0       0 r----   [ anon ]
00007fff72f3c000       4       4       0 r-x--   [ anon ]
ffffffffff600000       4       0       0 --x--   [ anon ]
---------------- ------- ------- -------
total kB            2360    1000      80
```

```
dell@dell-OptiPlex-3070:~/kernel/linux-5.10.1/cpe326$ ./zoneSpecify.out
Node 0, zone      DMA
  pages free     3975
        min      34
        low      42
        high     50
        spanned  4095
        present  3998
--
Node 0, zone    DMA32
  pages free     715570
        min      6012
        low      7515
        high     9018
        spanned  1044480
        present  739699
--
Node 0, zone   Normal
  pages free     145755
        min      10849
        low      13561
        high     16273
        spanned  1304576
        present  1304576
--
Node 0, zone  Movable
  pages free     0
        min      0
        low      0
        high     0
        spanned  0
        present  0
--
Node 0, zone   Device
  pages free     0
        min      0
        low      0
        high     0
        spanned  0
        present  0
```

Thank you.