

# CSGE602055 Operating Systems

## CSF2600505 Sistem Operasi

### Minggu 05: Virtual Memory

Rahmat M. Samik-Ibrahim

Universitas Indonesia

<http://rms46.vlsm.org/2/207.html>

REV097 13-NOV-2017

Minggu 00	29 Aug - 05 Sep 2017	Intro & Review
Minggu 01	07 Sep - 12 Sep 2017	IPR, SED, AWK, REGEX, & Scripting
Minggu 02	14 Sep - 19 Sep 2017	Protection, Security, Privacy, & C-language
Minggu 03	26 Sep - 30 Sep 2017	BIOS, Loader, Systemd, & I/O
Minggu 04	03 Okt - 07 Okt 2017	Addressing, Shared Lib, Pointer & I/O Programming
Minggu 05	10 Okt - 14 Okt 2017	Virtual Memory
Ming. UTS	15 Okt - 24 Okt 2017	
Minggu 06	26 Okt - 31 Okt 2017	Concurrency: Processes & Threads
Minggu 07	02 Nov - 07 Nov 2017	Synchronization
Minggu 08	09 Nov - 14 Nov 2017	Scheduling & Network Sockets Programming
Minggu 09	16 Nov - 21 Nov 2017	File System & Persistent Storage
Minggu 10	23 Nov - 28 Nov 2017	Special Topic: Retreat
Cadangan	30 Nov - 09 Des 2017	
Ming. UAS	10 Des - 23 Des 2017	

# Week 05: Memory

- 1 Start
- 2 Week 05
- 3 Memory
- 4 Paging
- 5 Translation
- 6 Memory
- 7 Hierarchical
- 8 VM
- 9 TOP
- 10 06-memory
- 11 The End

- Reference: (OSCE2e ch7/8) (UCB 11 12 13) (UDA P3L2) (OLD 06)
- Binding & Linking
  - Address Binding
  - Address Space: Logical & Physical
  - Dynamic & Static Linking
  - MMU: Memory Management Unit
  - Base and Limit Registers
  - Swapping
  - Mobile Systems Problem: no swap
- Memory Allocation
  - Contiguous Allocation
  - Multiple-variable-partition Allocation
  - First, Best, Worst Fit Allocation Strategy
- Fragmentation
  - External
  - Internal
  - Compaction

- Address Space
- Logical/Virtual Address
- Pages
- Page Number
- Page Offset
- Page Table
- PTE: Page Table Entry
- Page Flags: Valid/ Invalid
- TLBs: Translation Look-aside Buffers/ Associative Memory
- Physical Address
- Frames

# Address Translation Scheme

Address		Binary									
DEC	HEX	OFFSET	PG	OFF	PG	OFF	PAGE	OFF	PAGE	OFF	
00	00	00000	0	0000	00	000	000	00	0000	0	
01	01	00001	0	0001	00	001	000	01	0000	1	
02	02	00010	0	0010	00	010	000	10	0001	0	
03	03	00011	0	0011	00	011	000	11	0001	1	
04	04	00100	0	0100	00	100	001	00	0010	0	
05	05	00101	0	0101	00	101	001	01	0010	1	
06	06	00110	0	0110	00	110	001	10	0011	0	
07	07	00111	0	0111	00	111	001	11	0011	1	
08	08	01000	0	1000	01	000	010	00	0100	0	
09	09	01001	0	1001	01	001	010	01	0100	1	
10	0A	01010	0	1010	01	010	010	10	0101	0	
11	0B	01011	0	1011	01	011	010	11	0101	1	
12	0C	01100	0	1100	01	100	011	00	0110	0	
13	0D	01101	0	1101	01	101	011	01	0110	1	
14	0E	01110	0	1110	01	110	011	10	0111	0	
15	0F	01111	0	1111	01	111	011	11	0111	1	
16	10	10000	1	0000	10	000	100	00	1000	0	
17	11	10001	1	0001	10	001	100	01	1000	1	
18	12	10010	1	0010	10	010	100	10	1001	0	
19	13	10011	1	0011	10	011	100	11	1001	1	
20	14	10100	1	0100	10	100	101	00	1010	0	
21	15	10101	1	0101	10	101	101	01	1010	1	
22	16	10110	1	0110	10	110	101	10	1011	0	
23	17	10111	1	0111	10	111	101	11	1011	1	
24	18	11000	1	1000	11	000	110	00	1100	0	
25	19	11001	1	1001	11	001	110	01	1100	1	
26	1A	11010	1	1010	11	010	110	10	1101	0	
27	1B	11011	1	1011	11	011	110	11	1101	1	
28	1C	11100	1	1100	11	100	111	00	1110	0	
29	1D	11101	1	1101	11	101	111	01	1110	1	
30	1E	11110	1	1110	11	110	111	10	1111	0	
31	1F	11111	1	1111	11	111	111	11	1111	1	

# Memory (20 bits)

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
00000	A0	A1	A2	A3	A4	A5	A6	A7	A8	A9	AA	AB	AC	AD	AE	AF
00010	B0	B1	B2	B3	B4	B5	B6	B7	B8	B9	BA	BB	BC	BD	BE	BF
00020	C0	C1	C2	C3	C4	C5	C6	C7	C8	C9	CA	CB	CC	CD	CE	CF
00030	D0	D1	D2	D3	D4	D5	D6	D7	D8	D9	DA	DB	DC	DD	DE	DF
FFFF0																

# Hierarchical Page Table

- OPT: outer page table (P1)
- PT: page table (P2)
- Offset (D)
- Three-level Paging Scheme
- Hashed Page Tables
- Inverted Page Table
- Demand Paging
- Copy On Write (COW)



- Page Replacement Algorithm
  - Reference String
  - First-In-First-Out (FIFO)
  - Belady Anomaly
  - Optimal Algorithm
  - Least Recently Used (LRU)
  - LRU Implementation
  - Least Frequently Used (LFU)
  - Most Frequently Used (MFU)
- Frame Allocation
- Global vs. Local Allocation
- Non-Uniform Memory Access (NUMA)
- Working-Set Model
- Kernel
  - Buddy System Allocator
  - Slab Allocator

```
demo@badak: ~/git/demo/demos/week05-memory
root@... x rms46... x rms46... x @jem... x demo... x rms46... x rms46... x rms46... x rms46... x rms46... x rms46... x rms46... x rms46... x rms46... x
top - 12:20:10 up 7:49, 4 users, load average: 0.00, 0.01, 0.05
Tasks: 133 total, 1 running, 132 sleeping, 0 stopped, 0 zombie
%Cpu(s): 0.0 us, 0.0 sy, 0.0 ni,100.0 id, 0.0 wa, 0.0 hi, 0.0 si, 0.0 st
KiB Mem: 8197172 total, 402964 used, 7794208 free, 156120 buffers
KiB Swap: 683004 total, 0 used, 683004 free. 140104 cached Mem


```

PID	USER	PR	NI	VIRT	RES	SHR	S	%CPU	%MEM	TIME+	COMMAND
63	root	20	0	0	0	0	S	0.3	0.0	0:02.78	kworker/6:1
1	root	20	0	28828	4844	2932	S	0.0	0.1	0:01.12	systemd
2	root	20	0	0	0	0	S	0.0	0.0	0:00.00	kthreadd
3	root	20	0	0	0	0	S	0.0	0.0	0:07.34	ksoftirqd/0
5	root	0	-20	0	0	0	S	0.0	0.0	0:00.00	kworker/0:+
6	root	20	0	0	0	0	S	0.0	0.0	0:00.09	kworker/u1+
7	root	20	0	0	0	0	S	0.0	0.0	0:10.25	rcu_sched
8	root	20	0	0	0	0	S	0.0	0.0	0:00.00	rcu_bh
9	root	rt	0	0	0	0	S	0.0	0.0	0:00.00	migration/0
10	root	rt	0	0	0	0	S	0.0	0.0	0:00.15	watchdog/0
11	root	rt	0	0	0	0	S	0.0	0.0	0:00.14	watchdog/1
12	root	rt	0	0	0	0	S	0.0	0.0	0:00.00	migration/1
13	root	20	0	0	0	0	S	0.0	0.0	0:09.34	ksoftirqd/1
15	root	0	-20	0	0	0	S	0.0	0.0	0:00.00	kworker/1:+
16	root	rt	0	0	0	0	S	0.0	0.0	0:00.10	watchdog/2
17	root	rt	0	0	0	0	S	0.0	0.0	0:00.00	migration/2
18	root	20	0	0	0	0	S	0.0	0.0	0:09.84	ksoftirqd/2

Figure: top

## TOP (2)

```
demo@badak: ~/git/demos/week05-memory  
root@... x rms46... x rms46... x @jem... x demo... x rms46... x rms46... x rms46... x rms46... x rms46... x rms46... x rms46...  
Fields Management for window 1:Def, whose current sort field is %CPU  
Navigate with Up/Dn, Right selects for move then <Enter> or Left commits,  
'd' or <Space> toggles display, 's' sets sort. Use 'q' or <Esc> to end!  
  
* PID = Process Id PGRP = Process Group vMj = Major Faults  
* USER = Effective Use TTY = Controlling T vMn = Minor Faults  
* PR = Priority TPGID = Tty Process G USED = Res+Swap Size  
* NI = Nice Value SID = Session Id nsIPC = IPC namespace  
* VIRT = Virtual Image nTH = Number of Thr nsMNT = MNT namespace  
* RES = Resident Size P = Last Used Cpu nsNET = NET namespace  
* SHR = Shared Memory TIME = CPU Time nsPID = PID namespace  
* S = Process Statu SWAP = Swapped Size nsUSER = USER namespac  
* %CPU = CPU Usage CODE = Code Size (Ki nsUTS = UTS namespace  
* %MEM = Memory Usage DATA = Data+Stack (K  
* TIME+ = CPU Time, hun nMaj = Major Page Fa  
* COMMAND = Command Name/ nMin = Minor Page Fa  
PPID = Parent Proces nDRT = Dirty Pages C  
UID = Effective Use WCHAN = Sleeping in F  
RUID = Real User Id Flags = Task Flags <s  
RUSER = Real User Nam CGROUPS = Control Group  
SUID = Saved User Id SUPGIDS = Supp Groups I  
SUSER = Saved User Na SUPGRPS = Supp Groups N  
GID = Group Id TGID = Thread Group  
GROUP = Group Name ENVIRON = Environment v
```

Figure: "h" = help

## TOP (3)

```
demo@badak: ~/git/demos/week05-memory
```

```
root@... x rms46... x rms46... x @jem... x demo... x rms46... x rms46... x rms46... x rms46... x rms46... x rms46... x rms46... x rms46... x rms46...
```

### Fields Management for window 1:Defj, whose current sort field is %CPU

Navigate with Up/Dn, Right selects for move then <Enter> or Left commits,  
'd' or <Space> toggles display, 's' sets sort. Use 'q' or <Esc> to end!

* PID = Process Id	PGRP = Process Group	vMj = Major Faults
USER = Effective User	TTY = Controlling T	vMn = Minor Faults
PR = Priority	TPGID = Tty Process G	* USED = Res+Swap Size
NI = Nice Value	SID = Session Id	nsIPC = IPC namespace
* VIRT = Virtual Image	nTH = Number of Thr	nsMNT = MNT namespace
* RES = Resident Size	P = Last Used Cpu	nsNET = NET namespace
* SHR = Shared Memory	TIME = CPU Time	nsPID = PID namespace
S = Process Statu	* SWAP = Swapped Size	nsUSER = USER namespac
%CPU = CPU Usage	* CODE = Code Size (Ki	nsUTS = UTS namespace
* %MEM = Memory Usage	* DATA = Data+Stack (K	
TIME+ = CPU Time, hun	nMaj = Major Page Fa	
COMMAND = Command Name/	nMin = Minor Page Fa	
* PPID = Parent Proces	nDRT = Dirty Pages C	
UID = Effective Use	WCHAN = Sleeping in F	
RUID = Real User Id	Flags = Task Flags <s	
RUSER = Real User Nam	CGROUPS = Control Group	
SUID = Saved User Id	SUPGIDS = Supp Groups I	
SUSER = Saved User Na	SUPGRPS = Supp Groups N	
GID = Group Id	TGID = Thread Group	
GROUP = Group Name	ENVIRON = Environment v	

### Figure: Moving Fields

## TOP (4)

```
demo@badak: ~/glt/demo/demos/week05-memory  
root@... x rms46... x rms46... x @jem... x demo... x rms46... x rms46... x rms46... x rms46... x rms46... x rms46... x rms46... x rms46...  
Fields Management for window 1:Def, whose current sort field is %CPU  
Navigate with Up/Dn, Right selects for move then <Enter> or Left commits,  
'd' or <Space> toggles display, 's' sets sort. Use 'q' or <Esc> to end!  
  
* PID = Process Id SUID = Saved User Id ENVIRON = Environment v  
* PPID = Parent Proces SUSER = Saved User Na vMj = Major Faults  
* %MEM = Memory Usage GID = Group Id vMn = Minor Faults  
* VIRT = Virtual Image GROUP = Group Name nsIPC = IPC namespace  
* RES = Resident Size PGRP = Process Group nsMNT = MNT namespace  
* SHR = Shared Memory TTY = Controlling T nsNET = NET namespace  
* SWAP = Swapped Size TPGID = Tty Process G nsPID = PID namespace  
* CODE = Code Size (Ki SID = Session Id nsUSER = USER namespac  
* DATA = Data+Stack (K nTH = Number of Thr nsUTS = UTS namespace  
* USED = Res+Swap Size P = Last Used Cpu  
USER = Effective Use TIME = CPU Time  
PR = Priority nMaj = Major Page Fa  
NI = Nice Value nMin = Minor Page Fa  
S = Process Statu nDRT = Dirty Pages C  
%CPU = CPU Usage WCHAN = Sleeping in F  
TIME+ = CPU Time, hun Flags = Task Flags <s  
COMMAND = Command Name/ CGROUPS = Control Group  
UID = Effective Use SUPGIDS = Supp Groups I  
RUID = Real User Id SUPGRPS = Supp Groups N  
RUSER = Real User Nam TGID = Thread Group
```

### Figure: Moving Fields

# TOP (5)



demo@badak: ~/glt/demo/demos/week05-memory

root@... x rms46... x rms46... x @jem... x demo... x rms46... x rms46... x rms46... x rms46... x rms46... x rms46... x rms46... x rms46... x

KiB Mem: 8197172 total, 417256 used, 7779916 free, 156744 buffers  
KiB Swap: 683004 total, 0 used, 683004 free. 140200 cached Mem

PID	PPID	%MEM	VIRT	RES	SHR	SWAP	CODE	DATA	USED
3	2	0.0	0	0	0	0	0	0	0
4110	4108	0.1	115584	4776	3352	0	752	1128	4776
4129	3534	0.0	13020	3068	2384	0	184	808	3068
1	0	0.1	28828	4844	2932	0	1160	2152	4844
2	0	0.0	0	0	0	0	0	0	0
5	2	0.0	0	0	0	0	0	0	0
6	2	0.0	0	0	0	0	0	0	0
7	2	0.0	0	0	0	0	0	0	0
8	2	0.0	0	0	0	0	0	0	0
9	2	0.0	0	0	0	0	0	0	0
10	2	0.0	0	0	0	0	0	0	0
11	2	0.0	0	0	0	0	0	0	0
12	2	0.0	0	0	0	0	0	0	0
13	2	0.0	0	0	0	0	0	0	0
15	2	0.0	0	0	0	0	0	0	0
16	2	0.0	0	0	0	0	0	0	0
17	2	0.0	0	0	0	0	0	0	0
18	2	0.0	0	0	0	0	0	0	0
20	2	0.0	0	0	0	0	0	0	0
21	2	0.0	0	0	0	0	0	0	0

Figure: Memory Information

# 06-memory

```
#define MSIZE1 0x10000
#define MSIZE2 0x20000
#define MSIZE3 0x50000
#define MSIZE4 0x100000
#define MSIZE5 0x1000000
#define MSIZE6 0x10000000
#define LINE 75
#define MAXSTR 80
#include <stdio.h>
#include <stdlib.h>
#include <unistd.h>
#include <sys/types.h>

void printLine(int line) {
    while(line-- > 0) putchar('x');
    putchar('\n');
    fflush(NULL);
}

void main (void) {
    char strSYS2[MAXSTR], strSYS1[MAXSTR];
    char* chrStr = &strSYS1[0];
    int ii, myPID = getpid();
    sprintf(strSYS2, "top -b -n 1 -p%d | tail -5", myPID);
    sprintf(strSYS1, "top -b -n 1 -p%d | tail -1", myPID);
    printf("MSIZE1 (10k) MSIZE2 (20k) MSIZE3 (50k) MSIZE4 (100k) MSIZE5 (1M) MSIZE6 (10M) MSIZE1\n");
    printLine(LINE);
}
```

## 06-memory (2)

```
// sprintf(strSYS2, "top -b -n 1 -p%d | tail -5", myPID);
// sprintf(strSYS1, "top -b -n 1 -p%d | tail -1", myPID);
system(strSYS2);          /* (1) */
chrStr = malloc(MSIZE1);
system(strSYS1);          /* (2) */
free(chrStr);
chrStr = malloc(MSIZE2);
system(strSYS1);          /* (3) */
free(chrStr);
chrStr = malloc(MSIZE3);
system(strSYS1);          /* (4) */
free(chrStr);
chrStr = malloc(MSIZE4);
system(strSYS1);          /* (5) */
free(chrStr);
chrStr = malloc(MSIZE5);
for (ii = 0; ii < MSIZE5; ii++) {
    chrStr[ii]='a';
}
system(strSYS1);          /* (6) */
free(chrStr);
chrStr = malloc(MSIZE6);
system(strSYS1);          /* (7) */
free(chrStr);
chrStr = malloc(MSIZE1);
for (ii = 0; ii < MSIZE1; ii++) {
    chrStr[ii]='a';
}
system(strSYS1);          /* (8) */
free(chrStr);
printLine(LINE);
}
```



## 06-memory (2)

```
>>>>> $ 06-memory
```

(1) START            (2) MSIZE1=10k            (3) MSIZE2=20k    (4) MSIZE3=50k  
(5) MSIZE4=100k    (6) MSIZE5=1M (F)    (7) MSIZE6=10M    (8) MSIZE1=10k (F)

[illegible]

```
KiB Mem:  8197160 total,  341564 used,  7855596 free,  50776 buffers
```

```
KiB Swap: 683004 total, 0 used, 683004 free. 195692 cached
```

PID	PPID	%MEM	VIRT	RES	SHR	SWAP	CODE	DATA	USED
1567	1185	0.0	4172	688	612	0	4	320	688 (1)
1567	1185	0.0	4172	688	612	0	4	320	688 (2)
1567	1185	0.0	4172	688	612	0	4	320	688 (3)
1567	1185	0.0	4496	688	612	0	4	644	688 (4)
1567	1185	0.0	5200	1212	1116	0	4	1348	1212 (5)
1567	1185	0.2	20560	17576	1116	0	4	16708	17576 (6)
1567	1185	0.0	266320	1212	1116	0	4	262468	1212 (7)
1567	1185	0.0	4172	1212	1116	0	4	320	1212 (8)

[illegible]

# The End

- This is the end of the presentation.