KWANGWOON UNIVERSITY

2022년 2학기 **운영체제실습** 6주차

Task Management

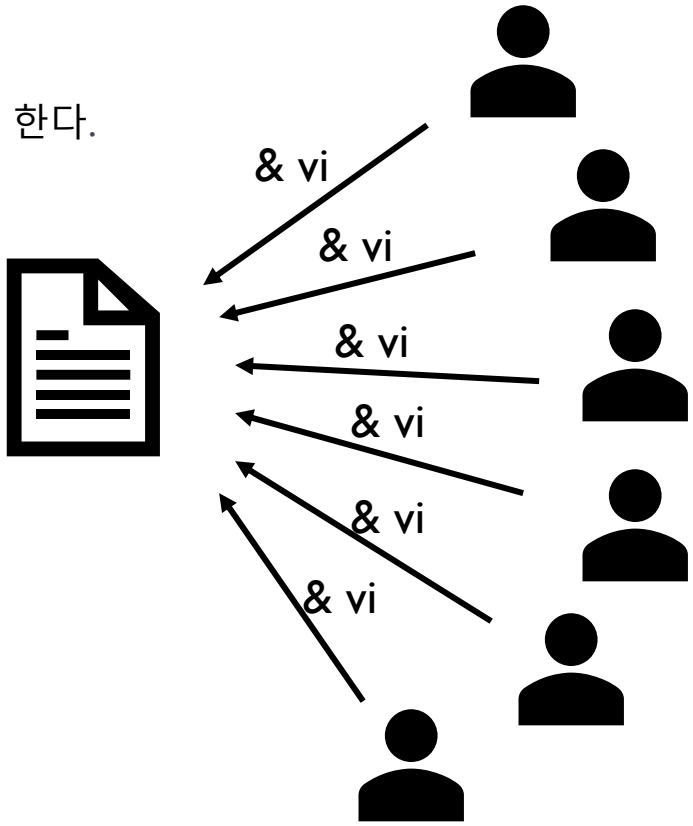
System Software Laboratory
School of Computer and Information Engineering
Kwangwoon Univ.

Contents

- **Process와 Process Descriptor의 이해**
- **task_struct 구조체**
- **실습 1**
  - 커널에 존재하는 모든 태스크들의 pid와 이름을 출력하여 확인하는 프로그램
- **프로세스 계보 (family)**

Process

- **6명의 사용자가 vi 프로그램을 동시에 사용하는 경우**
  - 각각 다른 프로세스가 6개 존재
  - 각각 다른 task_struct 구조체가 6개 존재
  - 리눅스 에서 프로세스는 Task 또는 Thread라 부르기도 한다.



Process와 Process Descriptor의 이해

- **Process Descriptor**
  - 변화하는 process의 모든 정보를 담고 있는 자료구조
    - General Term : **PCB** (Process Control Block)
    - Linux Specific : **task_struct** 라는 자료 구조를 사용
  - 즉, 리눅스에서 Process descriptor(프로세스 기술자)는 task_struct 자료 구조
  - 커널은 각 프로세스가 무엇을 하고 있는지 명확히 알아야 한다.
    - e.g.
      - 프로세스의 ID 및 우선 순위
      - 프로세스가 실행 상태
      - 프로세스가 할당 되어있는 주소 공간
      - 다룰 수 있는 파일
      - ……

task_struct 구조체 (1/9)

- **구조체를 통한 정보 관리**
  - 프로세스가 생성되면 task_struct 구조체를 통해 프로세스의모든 정보를 저장하고 관리
    - 모든 태스크들에게 하나씩 할당
    - include/linux/sched.h
    - 태스크 ID, 상태 정보, 가족 관계, 명령, 데이터, 시그널, 우선순위, CPU 사용량 및 파일 디스크립터 등 생성된 태스크의 모든 정보를 가짐
    - task_struct ***current**
      - 현재 실행되고 있는 태스크를 가리키는 매크로
      - in <include/asm-generic/current.h>

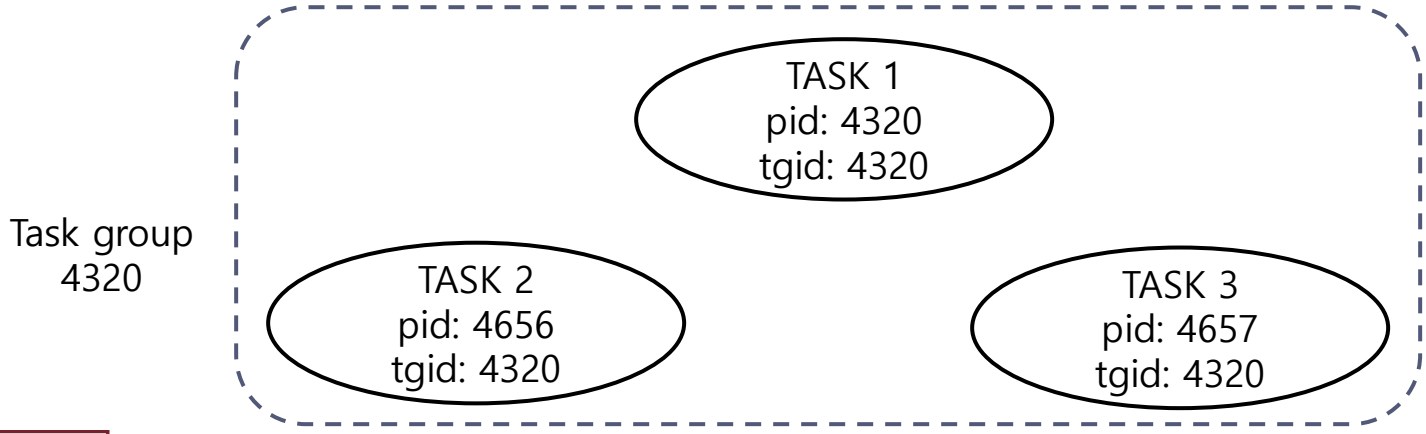
 # task_struct 구조체

task_struct 구조체 (2/9)

- **태스크 식별 정보에 관련된 변수, 함수들**

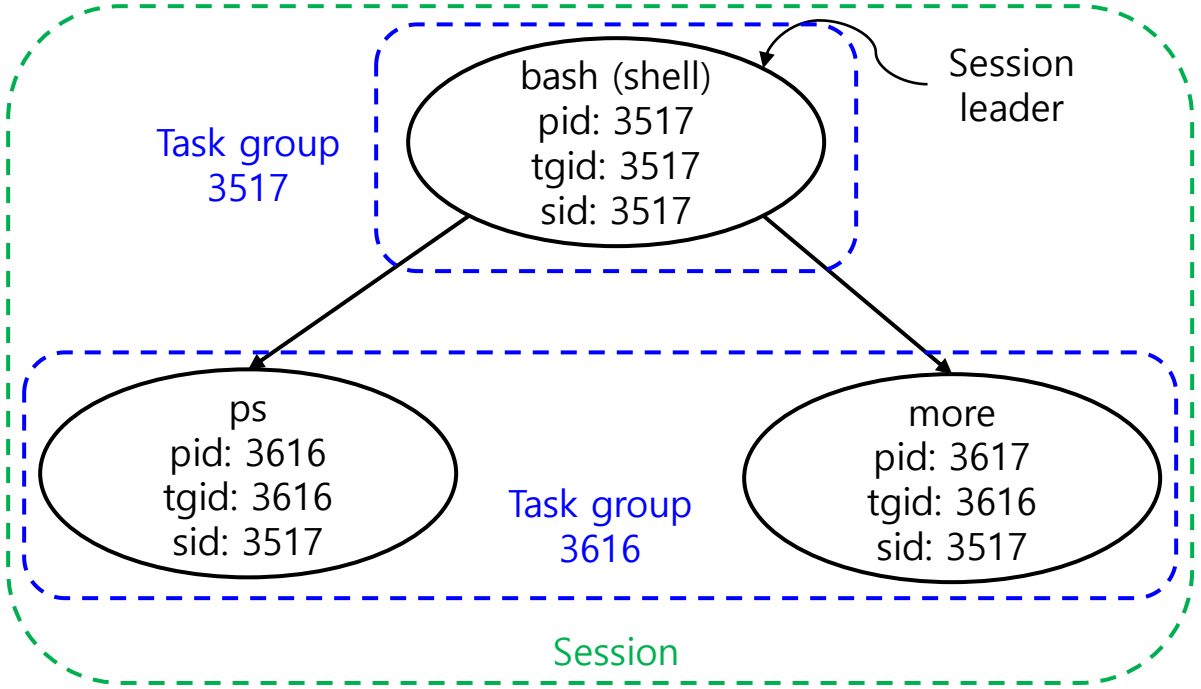
```
pid_t pid;                                  // Thread(Lightweight Process) ID
pid_t tgid;                                 // Thread Group(Process) ID
struct hlist_node pid_links[PIDTYPE_MAX];
```

  - 리눅스 커널에서의 태스크 식별
    - 프로세스와 스레드 모두 task_struct로 관리 (공유, 접근제어의 차이)
    - 시스템에 존재하는 태스크를 구분하기 위해, 각 태스크에 pid를 할당
    - 한 프로세스 내에서 생성된 스레드들은 동일한 tgid를 가지는 그룹을 형성하고, 각각의 스레드들은 유일한 pid를 가짐



task_struct 구조체 (3/9)

- **태스크 식별 정보에 관련된 변수, 함수들**
  - 태스크 그룹과 세션 리더의 예
    - 사용자가 콘솔로 로그인하여 shell을 띄웠다고 가정
      - shell도 하나의 태스크이므로 자신의 PID를 가지며, 하나의 태스크 그룹을 형성
      - #ps | more 명령을 실행하면 각 명령에 대해 각각 PID 값을 부여 받고, 이 두 명령은 하나의 태스크 그룹을 형성
      - 두 개의 태스크 그룹은 하나의 세션을 이루며 이때 shell은 세션 리더가 됨



task_struct 구조체 (4/9)

- **태스크들에 대한 사용자의 접근 제어에 관련된 변수, 함수, 매크로**
  - 태스크가 생성되면 사용자의 ID와 사용자가 속한 그룹의 ID가 등록됨

```
struct cred *cred;
gid_t GROUP_AT(struct group_info *gi, int i);
```

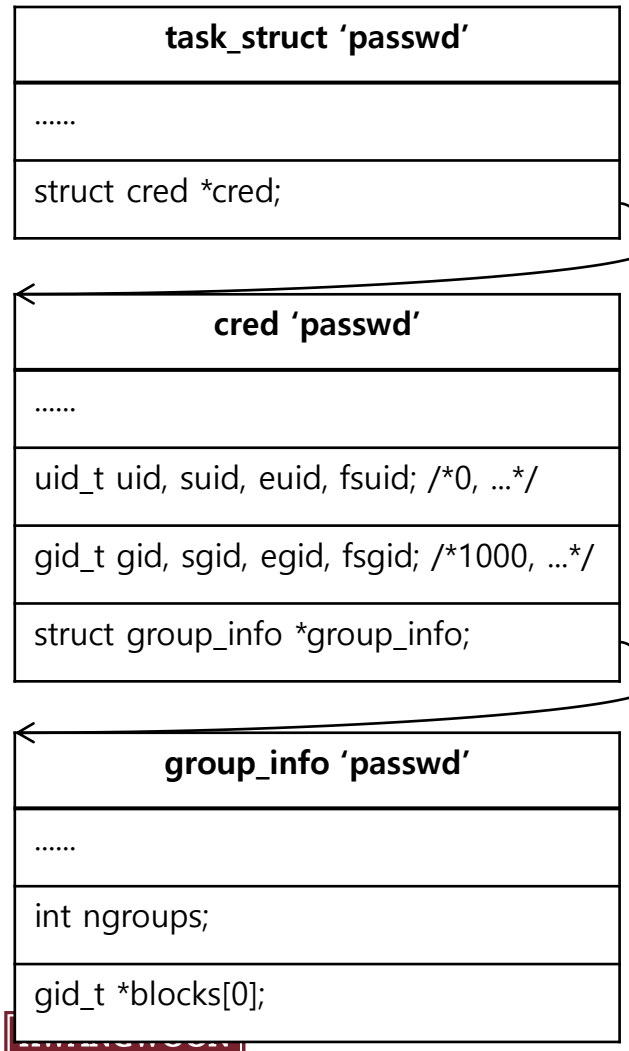
  - struct cred
    - 권한 관련 변수들이 저장된 구조체.
    - (uid, suid, euid, fsuid, gid sgid, egid, fguid)

  - struct group_info
    - 그룹들의 정보를 가진 구조체.

```
for(i=0; i < current->cred->group_info->ngroup; ++i )
{
        gid = GROUP_AT(current->cred->group_info, i);
}
```

task_struct 구조체 (5/9)

- **task_struct, cred, group_info의 관계**



$ sudo ps -A -o pid,ppid,uid,suid,euid,fsuid,gid,sgid,fsgid,command

```
sslab@ubuntu:~$ passwd &
[3] 4008
sslab@ubuntu:~$ Changing password for sslab.
(current) UNIX password: 123
123: command not found

[3]+  Stopped                 passwd

[3]+  Stopped                 passwd
sslab@ubuntu:~$ sudo ps -A –o pid,ppid,uid,suid,euid,fsuid,gid,sgid,fsgid,command
PID   PPID   UID  SUID  EUID FSUID   GID  SGID FSGID COMMAND
3282   2297     0     0     0     0  1000  1000  1000 passwd
3294   2297     0     0     0     0  1000  1000  1000 passwd
4012   2297     0     0     0     0  1000  1000  1000 sudo ps -A –o pid,ppid,uid,suid,euid,
4013   4012     0     0     0     0     0     0     0 ps -A -o pid,ppid,uid,suid,euid,fsuid,gid,s
```

suid (saved uid) : 권환 전환을 지원하는데 사용, uid 값을 저장

euid (effective uid) : 프로세스가 파일에 대해 가지는 권한
:파일에 접근 시 euidd를 통해 파일 접근 허용 여부 결정

fsuid (filesystem uid) : 파일시스템 접근 제어 용도로 사용됨

task_struct 구조체 (6/9)

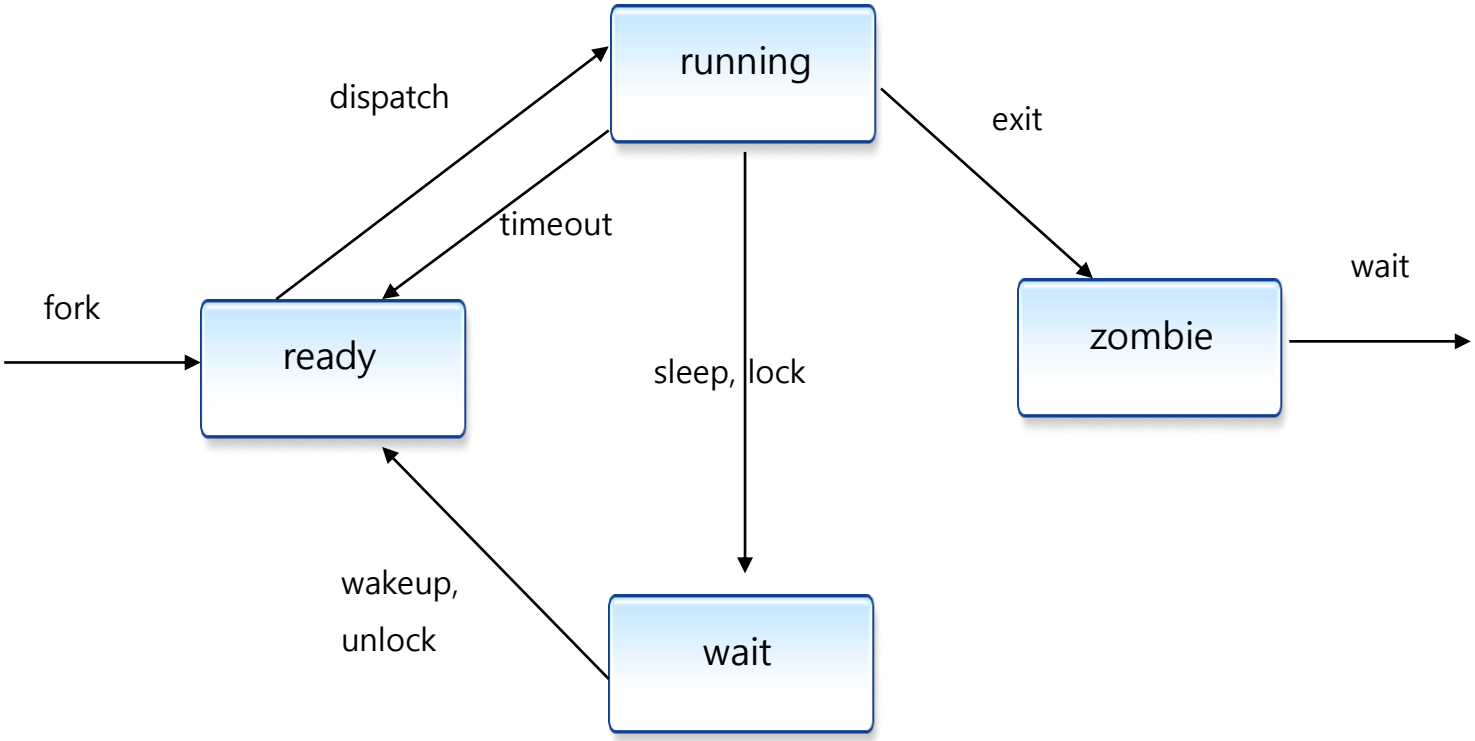
- 테스크 상태 정보에 관련된 변수
  - state 변수

```
#define TASK_RUNNING 0
#define TASK_INTERRUPTIBLE 1
#define TASK_UNINTERRUPTIBLE 2
#define TASK_STOPPED 4
#define EXIT_ZOMBIE 16
```

    - TASK_RUNNING
      - 실행 중이거나 준비 상태
    - TASK_INTERRUPT
      - 하드웨어나 시스템 자원을 사용할 수 있을 때까지 기다리고 있는 대기 상태
      - 예) wait for a semaphore
    - TASK_UNINTERRUPT
      - 하드웨어적인 조건을 기다리는 상태, 시그널을 받아도 무시
      - 예) process가 장치 파일을 열 때 해당 장치 드라이버가 자신이 다룰 하드웨어 장치가 있는지 조사할 때, memory swapping
    - TASK_STOPPED
      - 수행 중단 상태 (시그널을 받거나 트레이싱 등)
    - EXIT_ZOMBIE
      - process 실행은 종료했지만 아직 process의 자원을 반환하지 않은 상태

task_struct 구조체 (7/9)

- 테스크 상태와 전이



task_struct 구조체 (8/9)

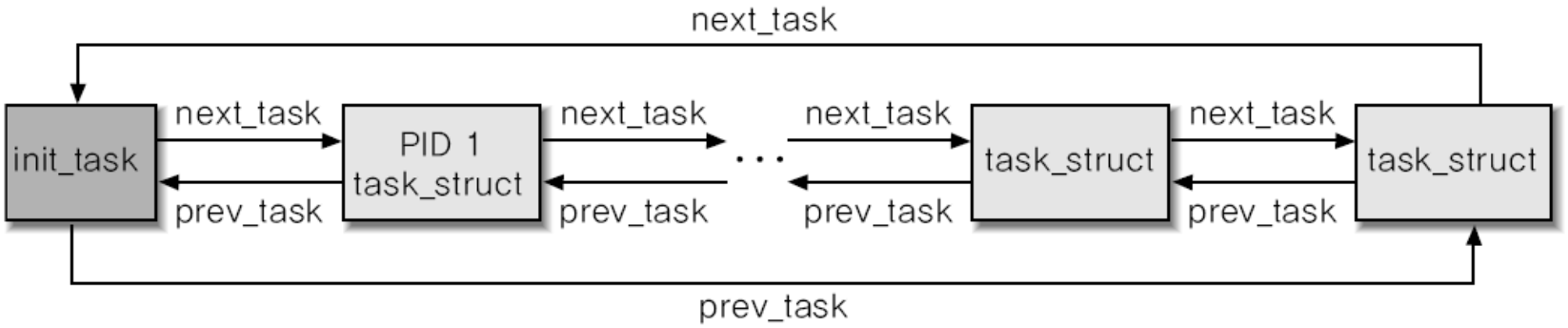
- **테스크 관계와 관련된 변수들**

```
struct task_struct *real_parent, *parent;
```

  - real_parent : 원래 부모 태스크(실제로 fork한 task)
  - parent : 현재 부모 태스크(SIGCHLD signal을 받는 task)

```
struct list_head tasks, children, sibling;
```

  - 커널에 존재하는 모든 태스크들은 원형 이중 연결 리스트로 연결



task_struct 구조체 (9/9)

- **그 밖의 변수들**
  - 스케줄링 정보
  - 시그널 정보
  - 메모리 정보
  - 파일 정보
  - 문맥 교환 정보
  - 시간 정보
  - 자원 사용 정보 등

실습 1 (1/4)

- **커널에 존재하는 모든 태스크들의 pid와 이름을 출력하여 확인하는 프로그램**
  - 사용할 변수

```
struct task_struct {
        ...
    pid_t pid;                  // task 식별자
        ...
     char comm[TASK_COMM_LEN]; /* executable name excluding path
                                  - access with [gs]et_task_comm (which lock
                                    it with task_lock())
        ...                       - initialized normally by setup_new_exec */
 };
```

  - 사용할 매크로
    - include/linux/sched/signal.h

```
#define next_task(p) \
        list_entry_rcu((p)->tasks.next, struct task_struct, tasks)
```

실습 1 (2/4)

- 커널에 존재하는 모든 태스크들의 pid와 이름을 출력하여 확인하는 프로그램

```
sslab@ubuntu:~$ mkdir taskTset
sslab@ubuntu:~$ cd taskTset/
sslab@ubuntu:~/taskTset$ vi display.c
```

```c
#include <linux/module.h>
#include <linux/sched.h>
#include <linux/init_task.h>

int displaytask_init(void)
{
        struct task_struct *findtask = &init_task;

        do
        {
                printk("%s[%d] ->", findtask->comm, findtask->pid);
                findtask = next_task(findtask);
        }
        while((findtask->pid != init_task.pid));
        printk("%s[%d]\n", findtask->comm, findtask->pid);
        return 0;

}

void displaytask_exit(void)
{
}

module_init(displaytask_init);
module_exit(displaytask_exit);
MODULE_LICENSE("GPL");
```

실습 1 (3/4)

- **커널에 존재하는 모든 태스크들의 pid와 이름을 출력하여 확인하는 프로그램 (cont'd)**
  - Makefile

```
obj-m := display.o

KDIR := /lib/modules/$(shell uname -r)/build
PWD := $(shell pwd)

all:
	$(MAKE) -C $(KDIR) SUBDIRS=$(PWD) modules
clean:
	$(MAKE) -C $(KDIR) SUBDIRS=$(PWD) clean
```

```
sslab@ubuntu:~/taskTset$ sudo make
make -C /lib/modules/4.19.67-SSLAB/build SUBDIRS=/home/sslab/taskTset  modules
make[1]: Entering directory '/home/sslab/Downloads/linux-4.19.67'
  Building modules, stage 2.
  MODPOST 1 modules
make[1]: Leaving directory '/home/sslab/Downloads/linux-4.19.67'
sslab@ubuntu:~/taskTset$ sudo insmod display.ko
sslab@ubuntu:~/taskTset$ sudo rmmod display.ko
```

실습 1 (4/4)

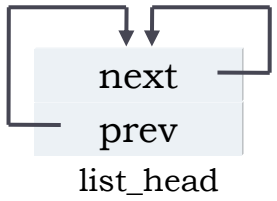
- **커널에 존재하는 모든 태스크들의 pid와 이름을 출력하여 확인하는 프로그램 (cont'd)**
  - 결과 화면

```
sslab@ubuntu:~/taskTset$ dmesg
......
[ 4571.620051] evolution-addre[1932] ->
[ 4571.620051] gnome-terminal-[1982] ->
[ 4571.620052] bash[1988] ->
[ 4571.620052] gvfsd-network[2003] ->
[ 4571.620052] sh[2042] ->
[ 4571.620053] zeitgeist-daemo[2046] ->
[ 4571.620053] zeitgeist-fts[2053] ->
[ 4571.620054] zeitgeist-datah[2055] ->
[ 4571.620054] gvfsd-dnssd[2080] ->
[ 4571.620055] update-notifier[2123] ->
[ 4571.620055] deja-dup-monito[2153] ->
[ 4571.620055] gvfsd-metadata[2200] ->
[ 4571.620056] kworker/4:0[2220] ->
[ 4571.620056] kworker/2:1[2340] ->
[ 4571.620057] kworker/0:0[2358] ->
[ 4571.620057] kworker/3:0[2373] ->
[ 4571.620057] kworker/5:0[2381] ->
[ 4571.620058] kworker/0:1[3795] ->
[ 4571.620058] kworker/2:2[3846] ->
[ 4571.620059] kworker/4:2[3847] ->
[ 4571.620059] kworker/u256:0[3888] ->
[ 4571.620059] kworker/1:2[3901] ->
[ 4571.620060] kworker/u256:2[3912] ->
[ 4571.620060] kworker/u256:1[3966] ->
[ 4571.620060] kworker/5:1[3989] ->
[ 4571.620061] sudo[4438] ->
[ 4571.620061] insmod[4439] ->
[ 4571.620061] swapper/0[0]
sslab@ubuntu:~/taskTset$
```

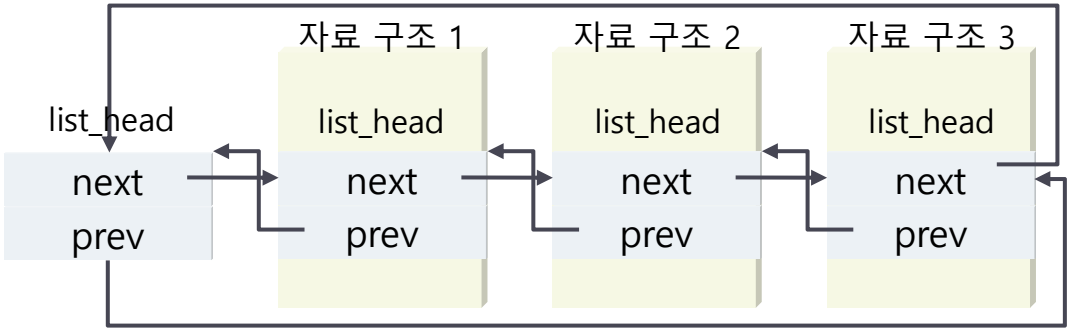
프로세스 계보 (family) (1/5)

- 이중 연결 리스트 (doubly linked list)

```
struct list_head {
     struct list_head *next, *prev;
 };
```



(a) 비어있는 이중 연결 리스트



(b) 몇 개의 자료 구조를 가진 이중 연결 리스트

프로세스 계보 (family) (2/5)

- **리스트를 처리하는 매크로 및 함수**
  - list_add(n, p)
    - p 바로 다음에 n을 삽입.
  - list_add_tail(n, p)
    - p 바로 앞에 n을 삽입.
  - list_del(p)
    - p를 삭제.
  - list_empty(p)
    - 헤드가 p인 리스트가 비어있는지를 검사.
  - list_entry(p, t, m)
    - 이름이 m이고 주소가 p인 list_head 필드를 포함한 t타입 자료구조의 주소를 반환.
  - list_for_each(p, h)
    - 헤드의 주소가 h로 지정된 모든 원소를 순환.
    - 각 원소의 list_head 자료 구조의 주소가 p에 반환

프로세스 계보 (family) (3/5)

- **부모 프로세스의 디스크립터를 얻으려면**
  - 부모 프로세스의 task_struct에 대한 포인터 ➔ **parent**

```
struct task_struct *my_parent = current->parent;
```

- **자식 프로세스들의 디스크립터를 얻으려면**
  - 자식 프로세스의 task_struct에 대한 포인터 ➔ **children**

```
struct task_struct *task;
struct list_head *list;

list_for_each(list, &current->children) {
        task = list_entry(list, struct task_struct, sibling);
        /* task now points to one of current's children */
}
```

```
list_for_each_entry(task, &current->children, list) {
        /* task now points to one of current's children */
}
```

프로세스 계보 (family) (4/5)

- **조상 프로세스를 끝까지 따라가려면**
  - init 프로세스의 task_struct에 대한 포인터

```
struct task_struct *task;
for (task = current; task != &init_task; task = task->parent)
        ;
/* task now points to init */
```

- **전체 리스트에서 다음 프로세스를 얻으려면**
  - next_task(task);

```
#define next_task(p) \ list_entry_rcu((p)->tasks.next, struct task_struct, tasks)
```

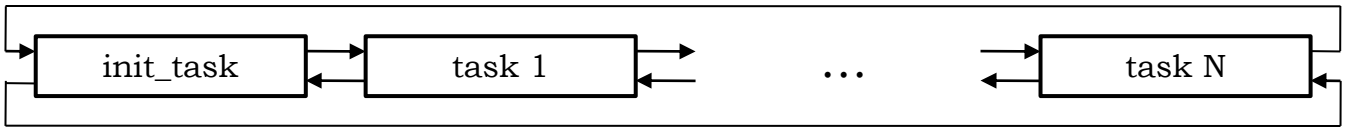
    - 이전 프로세스를 얻으려면 tasks의 prev 변수를 이용

프로세스 계보 (family) (5/5)

- **시스템 내 모든 프로세스를 출력하려면**
  - for_each_process(task)
    - 리스트에 있는 모든 프로세스를 탐색

```
struct task_struct *task;

for_each_process(task) {
        /* 각 태스크의 이름과 PID 출력 */
        printk("%s[%d]₩n", task->comm, task->pid);
}
```




```
#define for_each_process(p) ₩
    for (p = &init_task ; (p = next_task(p)) != &init_task ; )
```