

#### 프로젝트 #4 스택의 구조 콜스택 복원하기

리눅스 시스템 및 커널 전문가 2기 김현성

© 1. 프로젝트 주제

2. 프로세스와 스택

3. 콜스택 복원하기

4. Cmm 스크립트

### 1 프로젝트 주제

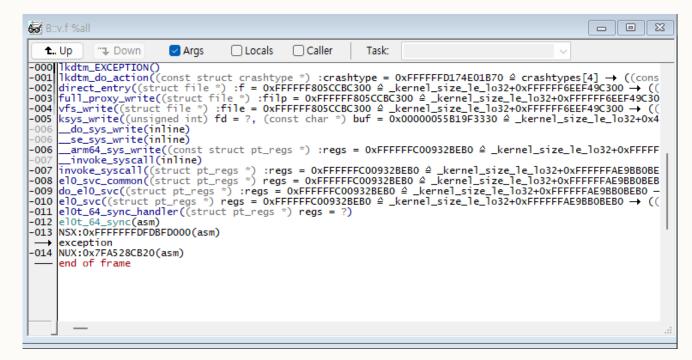


그림1. 콜스택

#### "스택 덤프를 이동하며 콜스택 복원시"

- ✓ 프로세스 스택 Corruption 시 손쉽게 디버깅 가능
- ✓ 스택 오버플로우 디버깅
- ✓ 프로세스의 콜스택을 보며 함수의 실행 흐름을 파악

### 2 프로세스와스택

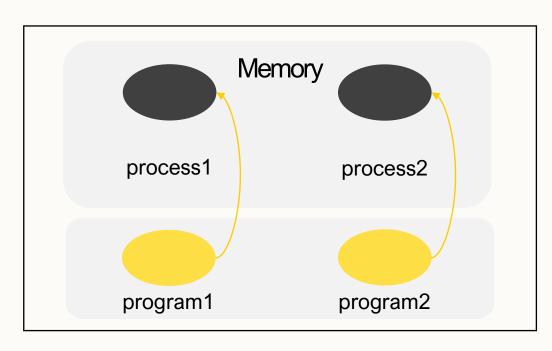


그림2. 프로세스

#### 프로그램 vs 프로세스

- ✓ 실행중인 프로그램
- ✓ 독립적인 실행
- ✓ 프로세스별 최소 1개의 스레드 생성

### 2 프로세스와스택

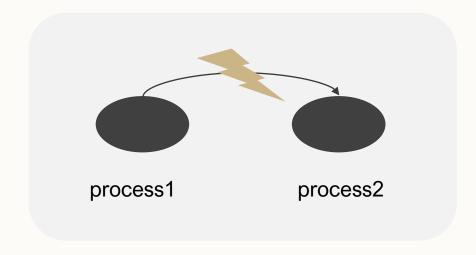




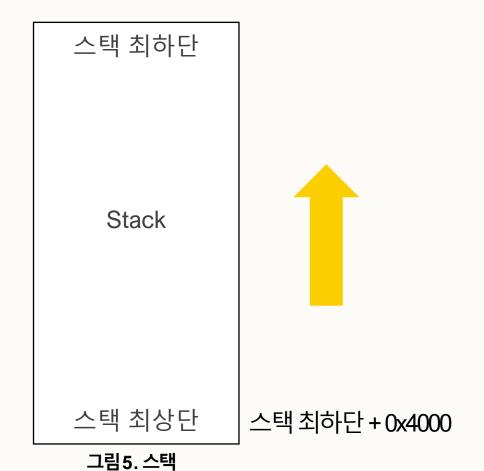
그림3. 프로세스간충돌

- OS는 개별 프로세스에게 독립적인 주소공간을 제공하여 프로세스간 충돌과 데이터 손실을 방지



그림4. 프로세스의 주소공간

### 2 프로세스와스택



#### 스택

"Key point"

- ✓ 스택은 높은 주소에서 낮은 주소로 자라난다.
- ✓ Armv8기반 리눅스 커널에서는 스택의 사이즈가 0x4000이다.

#### task\_struct 구조체

- <u>/include/linux/sched.h</u>
- 프로세스의 속성 정보를 표현하는 구조체
- task\_struct구조체 필드의 stack에는 스택의 최하단 주소 값이다.

```
struct task struct {
#ifdef CONFIG_THREAD_INFO_IN_TASK
         * For reasons of header soup (see current_thread_info()), this
         * must be the first element of task_struct,
        struct thread info
                                       thread_info:
#endi f
       unsigned int
                                       state
       /* saved state for "spinlock sleepers" */
       unsigned int
                                      saved state:
         * This begins the randomizable portion of task_struct, Only
         * scheduling-critical items should be added above here.
        randomized_struct_fields_start
       void
                                       *stack
       refcount t
                                       usage:
       /* Per task flags (PF_*), defined further below: */
       unsigned int
                                       flags:
       unsigned int
                                       pt race:
```

그림 6. struct task\_struct

#### crash64> runq -m

```
CPU 0: [0 01:10:11.687] PID: 0 TASK: ffffffd17553fac0 COMMAND: "swapper/
0"
CPU 1: [0 01:10:11.687] PID: 0 TASK: ffffff80402b1e40 COMMAND: "swapper/
1"
CPU 2: [0 00:00:00.000] PID: 1591 TASK: ffffff805d66dac0 COMMAND: "bash"
CPU 3: [0 01:10:11.686] PID: 0 TASK: ffffff80402b5ac0 COMMAND: "swapper/
3"
```

그림 7. runq-m

crash utility 프로그램에서 bash 프로세스의 태스크 디스크립터 주소 확인

#### TRACE32> v.v %all (struct task\_struct\*)0xfffff805d66dac0

```
6 B::v.v %all (struct task_struct*)0xffffff805d66dac0
(struct task_struct *) (struct task_struct*)0xffffff805d66dac0 = 0xFFFFFF805D66DAC0 

Arrow kernel_size
⊕ (void *) stack = 0xFFFFFC009328000 = _kernel_size_le_lo32+0xFFFFFAE9BB08000 → ,
⊕ (struct __call_single_node) wake_entry = ((struct llist_node) llist = ((struct llist_node *) nex
(int) on_rq = 1 ≙ 0x1 ≙ 'NNNS'.
⊕ (struct sched_entity) se = ((struct load_weight) load = ((long unsigned int) weight = 1048576 ≧
(struct sched_rt_entity) rt = ((struct list_head) run_list = ((struct list_head *) next = 0xFFFF
(struct sched_dl_entity) dl = ((struct rb_node) rb_node = ((long unsigned int) __rb_parent_color
```

그림 8. task struct 구조체

TRACE32 프로그램에서 bash프로세스의 task struct구조체 압력 => 스택 주소 확인

#### TRACE32> v.f %all

```
- - X
B::v.f %all
                                                                                            Locals
                                                                                                                       Caller
       t...Up "3 Down
                                                                Args
                                                                                                                                                         Task:
 -000 Tkdtm_EXCEPTION()
 -001 lkdtm_do_action((const struct crashtype *) :crashtype = 0xFFFFFFD174E01B70 	crashtypes[4] 	→ ((cons
| direct_entry((struct file *) :f = 0xFFFFFF805CCBC300 \( \text{\rmsize} \) kernel_size_le_lo32+0xFFFFF6EEF49C300 \( \text{\rmsize} \) (\( \text{\rmsize} 
             __do_sys_write(inline)
             __se_sys_write(inline)
 -006 __arm64_sys_write((const struct pt_regs *) :regs = 0xFFFFFC00932BEB0 ≙ _kernel_size_le_lo32+0xFFFFF
             __invoke_syscall(inline)
 -008 el0_svc_common((struct pt_regs *) regs = 0xFFFFFFC00932BEB0 ≙ _kernel_size_le_lo32+0xFFFFFFAE9BB0BEB
 -009 do_el0_svc((struct pt_regs *) :regs = 0xFFFFFFC00932BEB0 ≙ _kernel_size_le_lo32+0xFFFFFFAE9BB0BEB0 -
-010 el0_svc((struct pt_regs *) regs = 0xFFFFFC00932BEB0 ≙ _kernel_size_le_lo32+0xFFFFFAE9BB0BEB0 → ((
-011 el0t_64_sync_handler((struct pt_regs *) regs = ?)
-012 el0t_64_sync(asm)
-013 NSX:0xFFFFFFFDFDBFD000(asm)
  → exception
-014 NUX:0x7FA528CB20(asm)
              end of frame
```

그림 9. <del>콜</del>스택

v.f로 확인한 콜스택은 다음과 같다.

#### Stack 스택 최하단 0xFFFFFFC009328000 -000|lkdtm\_EXCEPTION() -001|lkdtm do action() -002|direct entry() -003|full proxy write() -004|vfs write() -005|ksys write() -006|\_do\_sys\_write(inline) -006 se\_sys\_write(inline) -006|\_arm64\_sys\_write() -007|\_invoke\_syscall(inline) -007|invoke\_syscall() -008|el0\_svc\_common() -009|do\_el0\_svc() -010|el0 svc() -011|el0t 64 sync handler() -012|el0t\_64\_sync(asm) -013|NSX:0xFFFFFFDFDBFD000(asm) -->|exception -014|NUX:0x7FA528CB20(asm) ---|end\_of\_frame 스택 최상단 0xFFFFFFC009328000+0x4000

그림 10. 콜스택 정리

#### TRACE32> v.f %all

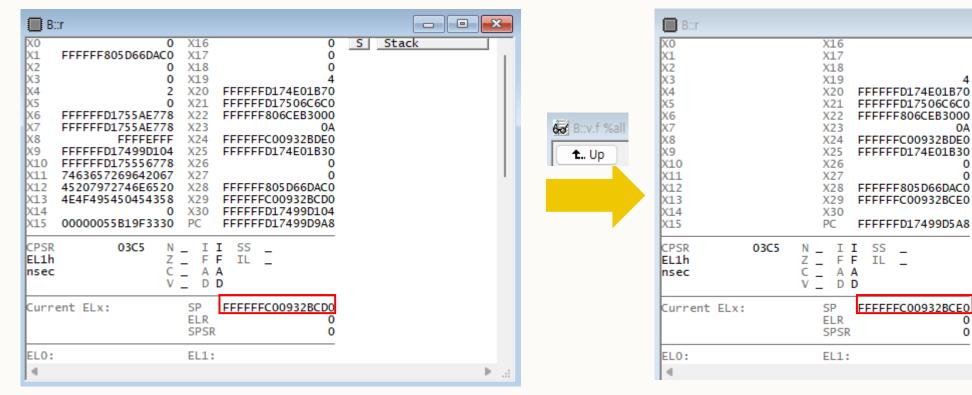


그림 11. sp레지스터

그림12. sp레지스터

\_ 0

S Stack

X

콜스택을 up 으로 이동시켰을 시 SP 레지스터 값이 변경된다.

#### TRACE32> d.v %y.ll 0xFFFFFC009328000+0x4000

```
Q [B::d.v %y.ll 0xFFFFFFC009328000+0x4000]
                                                                                                                      address | data
                                                                     symbol
                                                  value
   NSD: FFFFFFC00932BDF0
                                                  0xFFFFFFC00932BE00
                                                  0xFFFFFD174227D10 \vmlinux\kernel__syscall\invoke_syscall+0x48
   NSD:FFFFFFC00932BDF8
                         10 7D 22 74 D1 FF FF FF
   NSD:FFFFFFC00932BE00
                         30 BE 32 09 CO FF FF FF
                                                  0xFFFFFFC00932BE30
                         48 7E 22 74 D1 FF FF FF
                                                                     \\vmlinux\kernel__syscall\el0_svc_common+0x60
   NSD:FFFFFFC00932BE08
                                                  0xFFFFFFD174227E48
   NSD:FFFFFFC00932BE10
                                                  0xFFFFFFC00932BEB0
   NSD:FFFFFFC00932BE18
                         40 00 00 00 00 00 00 00
   NSD:FFFFFFC00932BE20
                         FO 16 D7 74 D1 FF FF FF
                                                  0xFFFFFD174D716F0
                                                                     \\vmlinux\arch__arm64__kernel__sys\sys_call_table
                                                                     \\vmlinux\entry common\el0t 64 sync handler+0xA4
                                                  0xFFFFFFD174D4DA24
                         24 DA D4 74 D1 FF FF FF
                         70 BE 32 09 CO FF FF FF
                                                  0xFFFFFFC00932BE70
                         34 7F 22 74 D1 FF FF FF
                                                  0xFFFFFD174227F34 \\vmlinux\kernel__syscall\do_el0_svc+0x28
   NSD:FFFFFFC00932BE38
                                                  0xFFFFFC00932BEB0
   NSD:FFFFFFC00932BE40
                         BO BE 32 09 CO FF FF FF
                         00 20 55 86 AF FF FF FF
                                                  0xFFFFFFAF86552000
                         FF FF FF FF 00 00 00 00
   NSD:FFFFFFC00932BE50
                                                  0xFFFFFFFF
                         20 CB 28 A5 7F 00 00 00
   NSD:FFFFFFC00932BE58
                                                  0x7FA528CB20
   NSD:FFFFFFC00932BE60
                         00 00 00 20 00 00 00 00
                                                  0x20000000
                         00 20 55 86 AF FF FF FF
                                                  0xFFFFFFAF86552000
   NSD:FFFFFFC00932BE70
                         80 BE 32 09 CO FF FF FF
                                                  0xFFFFFFC00932BE80
   NSD:FFFFFFC00932BE78
                         2C D3 D4 74 D1 FF FF FF
                                                  0xFFFFFD174D4D32C \\vmlinux\entry_common\el0_svc+0x2C
   NSD:FFFFFFC00932BE80
                         A0 BE 32 09 C0 FF FF FF
                                                  0xFFFFFFC00932BEA0
                         24 DA D4 74 D1 FF FF FF
                                                  0xFFFFFD174D4DA24 \\vmlinux\entry_common\el0t_64_sync_handler+0xA4
   NSD:FFFFFFC00932BE90
                         00 00 00 00 00 00 00
   NSD:FFFFFFC00932BE98
                         CO DA D4 74 D1 FF FF FF
                                                  0xFFFFFD174D4DAC0 \\vmlinux\entry_common\el0t_64_irg_handler+0x10
   NSD:FFFFFFC00932BEA0
                                                  0xFFFFFC00932BFE0
                         E0 BF 32 09 C0 FF FF FF
   NSD:FFFFFFC00932BEA8
                         10 16 21 74 D1 FF FF FF
                                                  0xFFFFFD174211610 \\vmlinux\Global\el0t_64_sync+0x1A0
   NSD:FFFFFFC00932BEB0
                         01 00 00 00 00 00 00 00 0x1
```

그림 13. 스택 덤프

d.v명령어를 입력하여 스택 주소의 덤프값을 확인

#### TRACE32> r.init

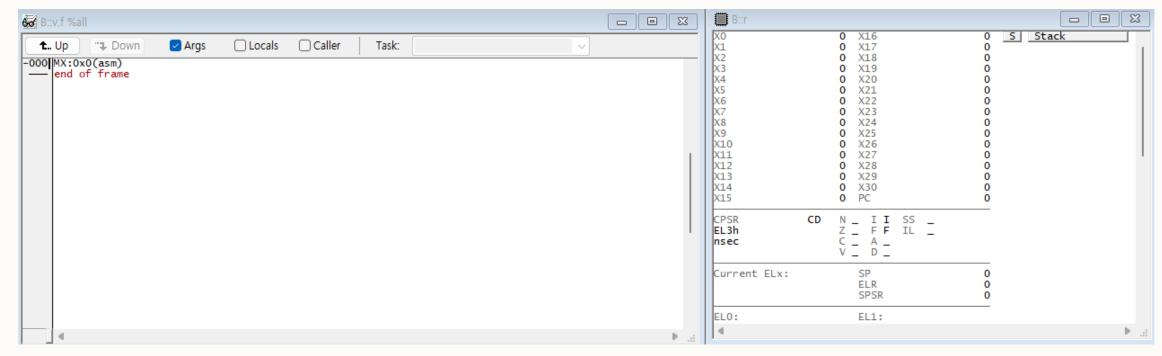
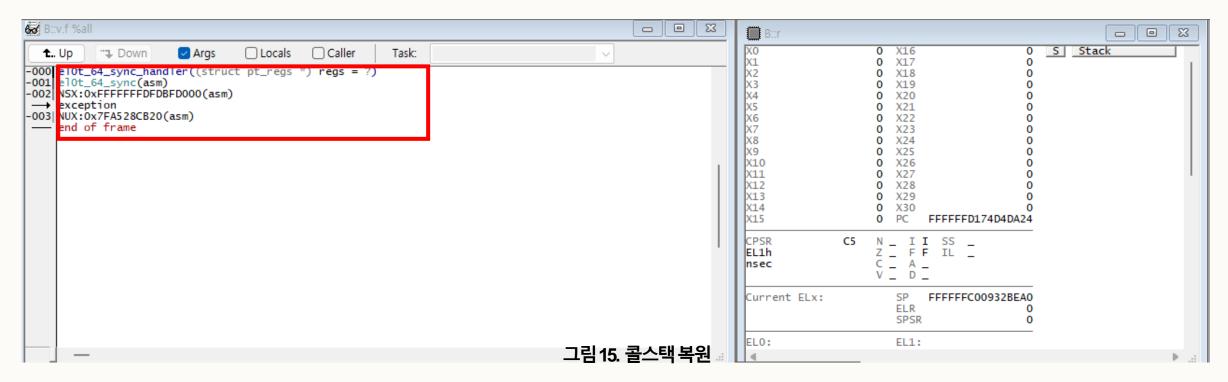


그림 14. 레지스터 초기화

콜스택 복원 실습을 위해 레지스터 값을 초기화 시켜주어 콜스택을 삭제

TRACE32> r.s pc 0xFFFFFD174D4DA24 r.s sp 0xFFFFFC00932BEA0 r.s x29 0xFFFFFC00932BEA0



Arm 아키텍처 calling convention 규약에 맞게 pc, sp, x29 레지스터에 값 입력

#### 전역변수를 선언하는 소스코드

```
9 Global &g_init_task_addr
10 Global &g_init_task_list_addr
11 Global &g_task_struct_start_addr
12 Global &g_task_struct_name
13 Global &g_task_struct_state
14 Global &g_task_struct_thread // added
15
16 Global &g_task_struct_list_next
17 Global &g_task_struct_list_next_temp
18 Global &g_task_struct_list_offset
19 Global &g_task_struct_list_offset
20 Global &g_task_struct_state_offset
21 Global &g_task_struct_thread_offset // added
```

#### 각 필드의 오프셋을 저장하는 코드

```
23     ; find the offset

24     &g_task_struct_list_offset=address.offset(v.address(((struct task_struct)0x0).tasks))

25     &g_task_struct_name_offset=address.offset(v.address(((struct task_struct)0x0).comm))

26     &g_task_struct_state_offset=address.offset(v.address(((struct task_struct)0x0).__state))

28     &g_task_struct_thread_offset=address.offset(v.address(((struct task_struct)0x0).__state))

29     &g_task_struct_thread_offset=address.offset(v.address(((struct task_struct)0x0).thread))

20     &g_task_struct_thread_offset=address.offset(v.address(((struct task_struct)0x0).thread))

21     &g_task_struct_thread_offset=address.offset(v.address(((struct task_struct)0x0).thread))

22     &g_task_struct_thread_offset=address.offset(v.address(((struct task_struct)0x0).thread))

23     &g_task_struct_thread_offset=address.offset(v.address(((struct task_struct)0x0).thread))

24     &g_task_struct_thread_offset=address.offset(v.address(((struct task_struct)0x0).thread))

25     &g_task_struct_thread_offset=address.offset(v.address(((struct task_struct)0x0).thread))

26     &g_task_struct_thread_offset=address.offset(v.address(((struct task_struct)0x0).thread))

26     &g_task_struct_thread_offset=address.offset(v.address(((struct task_struct)0x0).thread))

27     &g_task_struct_thread_offset=address.offset(v.address(((struct task_struct)0x0).thread))

28     &g_task_struct_thread_offset=address.offset(v.address(((struct task_struct)0x0).thread))

29     &g_task_struct_thread_offset=address.offset(v.address(((struct task_struct)0x0).thread))

29     &g_task_struct_thread_offset=address.offset(v.address(((struct task_struct)0x0).thread))

29     &g_task_struct_thread_offset=address.offset(v.address(((struct task_struct)0x0).thread))

20     &g_task_struct_thread_offset=address.offset(v.address(((struct task_struct)0x0)).thread(((struct task_struct)0x0).thread(((struct task_struct)0x0))

20     &g_task_struct_thread_offset=address(((struct task_struct)0x0).thread(((struct task_struct)0x0))

21     &g_tas
```

그림 17. 오프셋 값저장

각 필드들의 오프셋을 해당 전역변수에 저장한다.

#### Linkedlist를 순회하며 show thread 함수로 이동

```
while &g init task list addr!=&g task struct list next
37
38
         &g task struct start addr=&g task struct list next-&g task struct list offset
39
40
         &g_task_struct_name=&g_task_struct_start_addr+&g_task_struct_name_offset
         &g_task_struct_state=&g_task_struct_start_addr+&g_task_struct_state_offset
42
43
         print "process name: " data.string(d:&g task struct name)
44
         gosub show_thread &g_task_struct_start_addr
45
46
47
         &g_task_struct_list_next_temp=data.quad(d:&g_task_struct_list_next)
48
         &g task struct list next=&g task struct list next temp // for debugging purpose
```

그림 18. Linked list 순회

While문을 돌며 task라는 필드의 포인팅 하고 있는 Linked list를 순회하며 프로세스 이름을 출력하며, gosub 구문을 통해 show\_thread 함수로 이동한다.

```
54
55
                 show thread //added
56
57
58
59
     show thread:
60
       entry &task descriptor addr
61
       LOCAL &x19 &fp &sp &pc
62
63
       &x19=v.value( ((struct task struct *)&task descriptor addr)->thread.cpu context.x19)
       &fp=v.value( ((struct task_struct *)&task_descriptor_addr)->thread.cpu_context.fp)
64
       &sp=v.value( ((struct task struct *) &task descriptor addr)->thread.cpu context.sp)
65
       &pc=v.value( ((struct task struct *)&task descriptor addr)->thread.cpu context.pc)
66
67
     // print "x19: " data.string(d:&g task struct name)
68
     // print "fp: " data.string(d:&g task struct name)
69
     // print "sp: " data.string(d:&g task struct name)
70
     // print "pc: " data.string(d:&g task struct name)
72
73
       print "x19: " format.hex(1., &x19)
74
       r.s x19 &x19
75
       print "fp: " format.hex(1., &fp)
76
       r.s x29 &fp
77
       print "sp: " format.hex(1., &sp)
78
       r.s sp &sp
79
       print "pc: " format.hex(1., &pc)
80
       r.s pc &pc
81
82
     RETURN
83
84
     enddo
```

Task 디스크립터의 시작 주소를 인 자로 받아 thread.cpu\_context의 레 지스터의 값을 직접적으로 접근

#### TRACE32> pedit C:\troubleshooting\_vmcore\EXCEPTION\scripts/\*

```
(B::wr.we.PLIST)
                                                                                                                                    źM∮ Step
                             : Up
                                         2 Continue
                                                          źII Stop
                                                                      ź≣ł Enddo
                                                                                     ≨∰ Skip
                                                                                                 2 Breakpoints
            35 |&g_init_task_list_addr=&g_init_task_addr+&g_task_struct_list_offset
            37 &g_task_struct_list_next=data.quad(d:&g_init_task_list_addr)
                while &g_init_task_list_addr!=&g_task_struct_list_next
            40
            41
                          &q_task_struct_start_addr=&q_task_struct_list_next-&q_task_struct_list_offset
            43
                          &g_task_struct_name=&g_task_struct_start_addr+&g_task_struct_name_offset
            44
                          &g_task_struct_state=&g_task_struct_start_addr+&g_task_struct_state_offset
            46
                          print "process name: " data.string(d:&g_task_struct_name)
            48
                          gosub show_thread &g_task_struct_start_addr // added
            50
                          &g_task_struct_list_next_temp=data.quad(d:&g_task_struct_list_next)
                               show_thread //added
                  entry &task_descriptor_addr
            61
                 LOCAL &x19 &fp &sp &pc
                 &x19=v.value( ((struct task_struct *)&task_descriptor_addr)->thread.cpu_context.x19)
&fp=v.value( ((struct task_struct *)&task_descriptor_addr)->thread.cpu_context.fp)
&sp=v.value( ((struct task_struct *)&task_descriptor_addr)->thread.cpu_context.sp)
                  &pc=v.value( ((struct task_struct *)&task_descriptor_addr)->thread.cpu_context.pc)
                   print "x19: " data.string(d:&g_task_struct_name)
print "fp: " data.string(d:&g_task_struct_name)
print "sp: " data.string(d:&g_task_struct_name)
print "pc: " data.string(d:&g_task_struct_name)
                 print "x19: " format.hex(1., &x19)
                 r.s x19 &x19
            75
                                  " format.hex(1., &fp)
                  print "fp:
           76
                 r.s x29 &fp
            77
78
                                 " format.hex(1., &sp)
                  print "sp:
                 r.s sp &sp
                                " format.hex(1., &pc)
                  print "pc:
```

#### Cmm 스크립트를 수행하여 콜스택이 복원된 모습

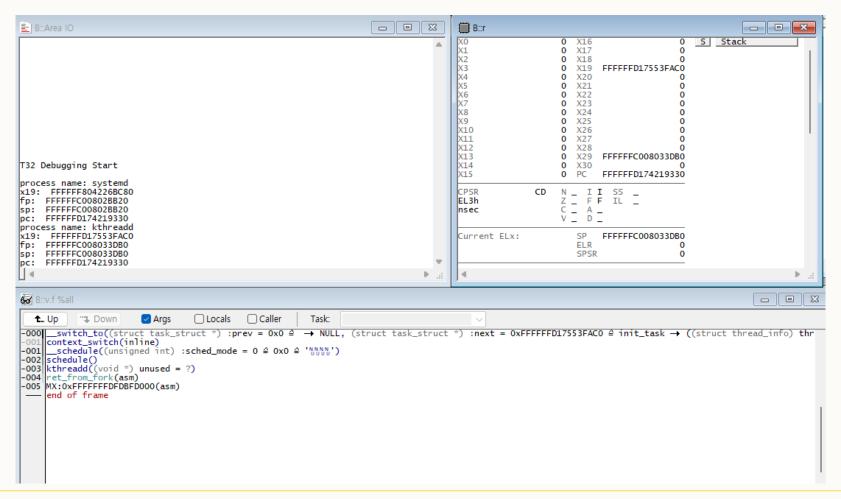


그림 21. 콜스택복원

# Thank you