操作系统实验

杨明

201605130117

Content

- 1. Debug Framework
- 2. Linux 0.11 plus
- 3. Trouble-shooting

Debug Framework

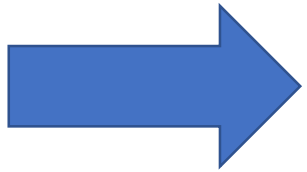
- Adding a new instruction
  - EMIT：0F 3C

```
mov eax, id
.byte 0x0F
.byte 0x3C
```

Linux 0.11 plus　swap

```
int free_page_tables(unsigned long from,unsigned long size)
```

```
for (nr=0 ; nr<1024 ; nr++) {
    if (1 & *pg_table)
        free_page(0xfffff000 & *pg_table);
    *pg_table = 0;
    pg_table++;
}
```

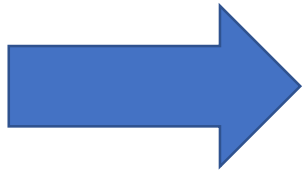


```
for (nr=0 ; nr<1024 ; nr++) {
    if (*pg_table) {
        if (1 & *pg_table)
            free_page(0xfffff000 & *pg_table);
        else
            swap_free(*pg_table >> 1);
        *pg_table = 0;
    }
    pg_table++;
}
```

Linux 0.11 plus　swap

```
int copy_page_tables(unsigned long from,unsigned long to,long size)
```

```
for ( ; nr-- > 0 ; from_page_table++,to_page_table++) {
  this_page = *from_page_table;
  if (!(1 & this_page))
    continue;
  this_page &= ~2;
  *to_page_table = this_page;
  if (this_page > LOW_MEM) {
    *from_page_table = this_page;
    this_page -= LOW_MEM;
    this_page >>= 12;
    mem_map[this_page]++;
  }
}
```



```
for ( ; nr-- > 0 ; from_page_table++,to_page_table++) {
  this_page = *from_page_table;
  if (!this_page)
    continue;
  if (!(1 & this_page)) {
    if (!(new_page = get_free_page()))
      return -1;
    read_swap_page(this_page>>1, (char *) new_page);
    *to_page_table = this_page;
    *from_page_table = new_page | (PAGE_DIRTY | 7);
    continue;
  }
  this_page &= ~2;
  *to_page_table = this_page;
  if (this_page > LOW_MEM) {
    *from_page_table = this_page;
    this_page -= LOW_MEM;
    this_page >>= 12;
    mem_map[this_page]++;
  }
}
```

Linux 0.11 plus　swap

```
void do_no_page(unsigned long error_code,unsigned long address)
```

```
page = *(unsigned long *) ((address >> 20) & 0xffc);
if (page & 1) {
  page &= 0xfffff000;
  page += (address >> 10) & 0xffc;
  tmp = *(unsigned long *) page;
  if (tmp && !(1 & tmp)) {
    swap_in((unsigned long *) page);
    return;
  }
}
```

Linux 0.11 plus mmap

- Add new sys call

```
fn_ptr sys_call_table[] = { sys_setup, sys_exit, sys_fork, sys_read,
sys_write, sys_open, sys_close, sys_waitpid, sys_creat, sys_link,
sys_unlink, sys_execve, sys_chdir, sys_time, sys_mknod, sys_chmod,
sys_chown, sys_break, sys_stat, sys_lseek, sys_getpid, sys_mount,
sys_umount, sys_setuid, sys_getuid, sys_stime, sys_ptrace, sys_alarm,
sys_fstat, sys_pause, sys_utime, sys_stty, sys_gtty, sys_access,
sys_nice, sys_ftime, sys_sync, sys_kill, sys_rename, sys_mkdir,
sys_rmdir, sys_dup, sys_pipe, sys_times, sys_prof, sys_brk, sys_setgid,
sys_getgid, sys_signal, sys_geteuid, sys_getegid, sys_acct, sys_phys,
sys_lock, sys_ioctl, sys_fcntl, sys_mpx, sys_setpgid, sys_ulimit,
sys_uname, sys_umask, sys_chroot, sys_ustat, sys_dup2, sys_getppid,
sys_getpgrp, sys_setsid, sys_sigaction, sys_sgetmask, sys_ssetmask,
sys_setreuid,sys_setregid, sys_sigsuspend, sys_sigpending,
sys_sethostname, sys_setrlimit, sys_getrlimit, sys_getrusage,
sys_gettimeofday,  sys_settimeofday, sys_getgroups, sys_setgroups,
sys_select, sys_symlink, sys_lstat, sys_readlink, sys_uselib,
sys_swapon, sys_reboot, sys_readdir, sys_mmap, sys_munmap,
sys_truncate, sys_ftruncate, sys_fchmod, sys_fchown, sys_getpriority,
sys_setpriority, sys_profil, sys_statfs, sys_fstatfs, sys_ioperm,
sys_socketcall, sys_syslog };

/* So we don't have to do any more manual updating.... */
int NR_syscalls = sizeof(sys_call_table)/sizeof(fn_ptr);
```

Linux 0.11 plus mmap

- Add new sys call

```
static caddr_t
mmap_chr(unsigned long addr, size_t len, int prot, int flags,
    struct inode *inode, unsigned long off)
{…
}

caddr_t
sys_mmap(unsigned long *buffer)
{…
}

int sys_munmap(unsigned long addr, size_t len)
{…
}
```

Linux 0.11 plus mmap

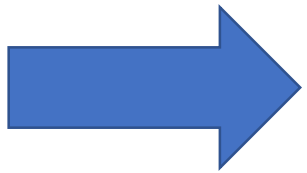
- Modify some sys structs

```
struct file {
  unsigned short f_mode;
  unsigned short f_flags;
  unsigned short f_count;
  unsigned short f_reada;
  struct inode * f_inode;
  struct file_operations * f_op;
  off_t f_pos;
};
```

Linux 0.11 plus mmap

- Modify some sys structs

```
struct m_inode {
  unsigned short i_mode;
  unsigned short i_uid;
  unsigned long i_size;
  unsigned long i_mtime;
  unsigned char i_gid;
  unsigned char i_nlinks;
  unsigned short i_zone[9];
/* these are in memory also */
  struct task_struct * i_wait;
  struct task_struct * i_wait2; /* for pipes */
  unsigned long i_atime;
  unsigned long i_ctime;
  unsigned short i_dev;
  unsigned short i_num;
  unsigned short i_count;
  unsigned char i_lock;
  unsigned char i_dirt;
  unsigned char i_pipe;
  unsigned char i_mount;
  unsigned char i_seek;
  unsigned char i_update;
};
```



```
struct inode {
  dev_t i_dev;
  ino_t i_ino;
  umode_t i_mode;
  nlink_t i_nlink;
  uid_t i_uid;
  gid_t i_gid;
  dev_t i_rdev;
  off_t i_size;
  time_t  i_atime;
  time_t  i_mtime;
  time_t  i_ctime;
  unsigned long i_data[16];
  struct inode_operations * i_op;
  struct super_block * i_sb;
  struct task_struct * i_wait;
  struct task_struct * i_wait2; /* for pipes */
  unsigned short i_count;
  unsigned char i_lock;
  unsigned char i_dirt;
  unsigned char i_pipe;
  unsigned char i_mount;
  unsigned char i_seek;
  unsigned char i_update;
};
```

Trouble-shooting

- 1. page fault

```
/* gdbstub.cc:487 */
if (last_stop_reason == GDBSTUB_EXECUTION_BREAKPOINT ||
            last_stop_reason == GDBSTUB_TRACE)
{
  write_signal(&buf[1], SIGTRAP);
}
else if (last_stop_reason == GDBSTUB_STOP_NO_REASON)
{
  write_signal(&buf[1], SIGSEGV);
}
else
{
  write_signal(&buf[1], 0);
}
```

Trouble-shooting

- 2. not owner

配合 `coreutils` 源码中的 `remove.c` 逆向 `rm`,发现Linux 0.11 缺少系统调用 `lstat`

```
.text:000002C0 ; ---------------------------------------------------------------------------
.text:000002C0
.text:000002C0 loc_2C0:                                ; CODE XREF: sub_280+17↑j
.text:000002C0                                         ; sub_280+23↑j ...
.text:000002C0                 lea     eax, [ebp+var_20]
.text:000002C3                 push    eax             ; statbuf
.text:000002C4                 push    path            ; filename
.text:000002CA                 call    lstat
.text:000002CF                 add     esp, 8
.text:000002D2                 test    eax, eax
.text:000002D4                 jz      short loc_30C
.text:000002D6                 cmp     dword_6FD4, 2
.text:000002DD                 jnz     short loc_2EC
.text:000002DF                 cmp     dword_6F4C, 0
.text:000002E6                 jz      short loc_2EC
.text:000002E8                 xor     eax, eax
.text:000002EA                 jmp     short locret_339
.text:000002EC ; ---------------------------------------------------------------------------
```

Troble-shooting

- 2. not owner
  - 添加新的系统调用

```
/* include/linux/sys.h */
extern int sys_lstat();
extern int sys_nop();


fn_ptr sys_call_table[] = { sys_setup, sys_exit, sys_fork, sys_read,
sys_write, sys_open, sys_close, sys_waitpid, sys_creat, sys_link,
sys_unlink, sys_execve, sys_chdir, sys_time, sys_mknod, sys_chmod,
sys_chown, sys_break, sys_stat, sys_lseek, sys_getpid, sys_mount,
sys_umount, sys_setuid, sys_getuid, sys_stime, sys_ptrace, sys_alarm,
sys_fstat, sys_pause, sys_utime, sys_stty, sys_gtty, sys_access,
sys_nice, sys_ftime, sys_sync, sys_kill, sys_rename, sys_mkdir,
sys_rmdir, sys_dup, sys_pipe, sys_times, sys_prof, sys_brk, sys_setgid,
sys_getgid, sys_signal, sys_geteuid, sys_getegid, sys_acct, sys_phys,
sys_lock, sys_ioctl, sys_fcntl, sys_mpx, sys_setpgid, sys_ulimit,
sys_uname, sys_umask, sys_chroot, sys_ustat, sys_dup2, sys_getppid,
sys_getpgrp, sys_setsid, sys_sigaction, sys_sgetmask, sys_ssetmask,
sys_setreuid,sys_setregid, sys_nop, sys_nop, sys_nop,
sys_nop, sys_nop, sys_nop, sys_nop,
sys_nop, sys_nop, sys_nop, sys_nop, sys_nop,
sys_lstat, sys_nop, sys_nop };

/* So we don't have to do any more manual updating.... */
int NR_syscalls = sizeof(sys_call_table)/sizeof(fn_ptr);
```

Trouble-shooting

- 2. not owner
  - 添加新的系统调用

```
/* kernel/system_call.s */
cmpl NR_syscalls,%eax
jae bad_sys_call
```

```
/* fs/stat.c */
int sys_lstat(char * filename, struct stat * statbuf)
{
        return sys_stat(filename, statbuf);
}
```

Thanks