

#####® ##### ## #####

```
##### ##### <rdivacky@FreeBSD.org>
#####: 43126
```

```
#####, #####, ##### #####, ##### ## ##### ## #—
##### ##### ## ##### ## ##### #####
##### ## ## ##### ##### ##/# ##### #####.
```

###, ##/2, #####, ##/2, #/390, ### #####  
 ## #####  
 #####, #####, ## #####.

##### ## # ##### ##### ## ### ##### #####.

##### ## # ##### ##### ## ##### #####.

##### ## # ##### ##### ##### ##### #####.

##### , ##### , ### ##### ### ## #####  
##### ## ##### , ##.

##### ## # ##### ##### ##### #####.

```
###, ### #####, ###, ### #####, ###, ###, ###,
###, #####, #####, #####, #####, #####
### ##### # ##### # # #,
###. # # # ##### # # # # #.
```

```
##### ## #### ##### ##### ## ##### ### ##### #  
##### ##### ##### ## ##### ## #####. #####  
##### ##### ## #####, ### ## #####  
##### ## ##### ## #####, ## #####  
##### ##### ## #####
```

2013#11#07 ## #####.

#####

[illegible]

##### ## #####

1. #####

2. # #### #####...

2



## 2.2.2. #####

##### (#, ) #####—  
#####. ##### ##  
#####. ##### ##  
#####.

## 2.2.3. #####

##### (#).  
#####(2).  
#####  
#####  
#####(3).  
#####(2).

## 2.2.4. #####

#####  
#####  
#####  
##### (1:1)  
##### (1:)

1:1

- #####
- ##### (#####)

- + #####

- + #####

1:1

- + #####
- + #####
- #####
- #####



#### ## #####?

### #### ## #### ## #### (#####) #####  
### ##### ## #### ##### #####. ##### ## ####  
## ##### ##### ## ##### ## #### ##### ## ####  
##### ## #### ## #### ##### ## #### ## ##### ##—  
##### ##### ## #### ##### ## #### #####. ## ####  
##### #### ## rdivacky\_linuxolator ##### ##### ## ####  
### ##### ##.

#### ##### ## ##### ##### ## ## ##### ## [2].

### 2.3.1. #####

##### ## ##### ## ####® ## #### ## ##### ## ## ##  
##### ## #### #: ##### ## #### #####. ##### ## ####  
##### ## #### #: # ##### ## # ####. ##### ## #### ##  
## #### ##### ## #### ##### ## ##### ##/#### ##  
### ##### ##### ## #### #386 ##### ## ####  
##### ## #### ##### ## ##### #####. #####  
### ##### [1] ## ####.

#### 2.3.1.1. #####

##### ## ## ##### ## #####, ##### ## #  
#### ## #####(2) #####. ##### # ##### sysentvec, #####  
##### ##. ## ##### ##### #####, #####—  
#####, ##### ## ##### (#####, #####—  
###, ##.). ##### ## ##### ##### ## #####—  
####, ## ## ## ##### ## ##### ##### ##—  
##### ## #####, ##### ## ##—  
##### ## ##.

#### 2.3.1.2. #####

##### ## ##### ## ##### 0x80 ##### %eax  
## ## # #####.

#### # ##### 0x80, ## int0x80 ##### ##—  
#### (##### ## sys/i386/i386/exception.s), ##### (##. #####  
## ## #####) ## # ## #####(2) (##### ## sys/i386/i386/trap.c),  
##### ## #####. ## #####  
### ##### (##### ## ## sysvec #####), ##### ## 32####  
## 64#### ## (##### ## ##), ##### ##—  
##### ## #####. ####, ## #####—  
##### ## ## ##### (##### ## ERESTART ## EJUSTRETURN #####).  
##### ## userret() ## ##### ## ##  
##### ## ## ## struct thread \*td,







#####® ##### ## #####

---

#### ## ## #####. ##### ## ##### 0x80 ## ##### ## ##### ##### ##  
##### ## ##### ## ## #####. ##### ##### ## ##### ## ## #####  
## ##### ## ##### #####. ##### ## ## #####® ##### #####  
##### ##### ## ## ##### ## ## ##### ## ##### ## #####:

1. ##### #> %ebx
2. ##### #> %ecx
3. ##### #> %edx
4. ##### #> %esi
5. ##### #> %edi
6. ##### #> %ebp

##### ## ## ##### ## ## ##, ##### ® ##### ##### #####—  
#### (#### ##### ## clone #####).

#### 2.4.1.2. #####

## ## ##### ## ##### ## arch/i386/kernel/traps.c ## ## ## #####  
#### ## arch/i386/kernel/entry.S, ##### ##### ## ## #####

#### 2.4.1.3. #####

##### ## ## ##### ## ##### ## ##### [###\(3\)](#), ##### ## ##  
##### ##### #####, ##### ##### ## ## #####—  
#####. ## ## ##### ##### ##### ## ##### ## ##### ##  
##### ## ## ##### ## ## ##### ##### ## ## #####.

#### 2.4.1.4. #####® #####

## ## 2.6 #####, ## #####® ##### ##### ##### ## ## ##—  
##### ® #####, ##### ##, ## ## #####. ## ## ##### ## ##  
##### ## #####, ## ## ## ##### (#####) [#####\(2\)](#) ##### ##  
#### #####. ##### ##### ## ##### ## ##### ## ##. ##### ##  
*NPTL* (## #####® #####) ##### ##### ## ## ##### (##  
##### 1:1 #####). ##### # ## ##### ## #####® 2.6 ##### ## clone  
##### (#### ##### ## ##### ##). #### #####  
# ## ## ##### ## ##### ## ## ##### #####  
#####. ## ##### ## # ## ##### ## ##### ## ## #####  
## ##### ## ##### # #####.

##### ##### ##:

- CLONE\_VM # ##### ##### #####
- CLONE\_FS # #####, ## ## #####



```
| CLONE_DETACHED
#endif

| 0);
```

```
#define CLONE_SIGNAL (CLONE_SIGHAND | CLONE_THREAD)
```

2.5. #####

##### ## ##### ## # #####, ##### ##### ## ##### ##### ## ##  
##### ## ## ##### ## ##, ## ##### ##### ##### ##### ## ## # #####  
##### ## # ##### ##### ##### ##### ## ## ##, ##### ##### #####  
## #####, ##### ##### ##### ## ## ##### ## ## ##### # #, #, ##### ##  
##### ##### ## ## ## ##### ##### ##### (###/#####) ##### ## ##  
## ## #####, ##### ##### ##### ##### ##### ## ##### ## ##  
(#####) #####, ##### ## ## ##### #####.

11





##### ## #####. ## #####—  
##### ## ##### ##### #####  
##### ## #####. ## #####, ## ##  
##### ##### ##### ##### (## #####  
#####) #####, ## #####—  
##### ## ##### #####. ## ## ##—  
## ## ##### ## #####  
##### (## ## #####, ##.). ## # ##### ##  
#####, ##### ## #####(9). ## ## ##, #####, #####  
#####) ##### ## ## ##  
#####).

[illegible]

#####  
#####  
#####  
#####  
#####  
#####

- #####  
#####

```
#####  ### #####  ## #####  #####  #####  #####  #####  #####  ## #####
#####  ### #####  ## #####  #####  #####  #####  #####  #####  #####  #####
```



- ##### # ##### # ##(9)
- ##### # ##### # #####(9)
- ##### # ##### # ####(9)

### 3.2.1.7. #####

- ##### (### #####)
- #####\_####
- #####\_###

3.2.1.8. #####

### 3.2.1.9. ##### ###/##### #####

16





#####  
#####(9) #####  
#####

### 3.2.3.2. ##\_#####

#####\_#####(9) #####  
#####  
#####  
#####  
#####  
#####

### 3.2.3.3. #####

- fgetvp # ##### # ##### # ##### # ##### # ##### # ##### # ##### # #####  
#####
- ##\_#####(9) # ##### # #####
- vn\_unlock # ##### # #####
- #####\_#####(9) # ##### # ##### # ##### # ##### # #####
- #####\_#####(9) # ##### # ##### # ##### # ##### # #####
- #####\_#####(9) # ##### # ##### # ##### # ##### # #####
- #####\_#####(9) # ##### # ##### # ##### # ##### # #####
- #####\_#####(9) # ##### # ##### # ##### # ##### # #####
- #####(9) # ##### # ##### # ##### # ##### # #####
- #####(9) # ##### # ##### # ##### # ##### # #####
- #####(9) # ##### # ##### # ##### # ##### # #####

### 3.2.3.4. #####

- fget # ##### # ##### # ##### # ##### # ##### # ##### # ##### # #####  
##### # ##### # #####
- fdrop # ##### # ##### # ##### # ##### # #####
- fhold # ##### # ##### # ##### # #####

## 4. #####® #####

#####  
#####

#####<sup>®</sup> #####

---

### ### ##### ##### ##### ##### #####. ## #####  
#####. #####, #####, #####, #####, #####. ##### #386 ##  
## ## ##### ##### ## ##### ##### ##### #####. ##  
#####. #####. #####. #####. #####. #####  
##### #386 ## ## #####. #.## ## ##### ## #####.

#### 4.1. #####

##### ## ##### ##### ## linux\_sysvec.c, ##### ##### ## ## ##—  
##### ## ## ## sysentvec #####. ##### # #####<sup>®</sup> ##### ##  
##### # #####, ## ##### ##### ##### ##### #####  
##### ## ## #####<sup>®</sup> ##.

##### 4.1.1. #####<sup>®</sup> #####

#####<sup>®</sup> ##### ##### ## ##### ## ##### (#### ## ## ## ##  
## 6 ##### ## #386) ##### ##### ## ## #####. ## #####<sup>®</sup> #####  
##### ## ## ##### ##### ##### ## ## #####. ## ##### ## ## ##—  
##### #: %ebx, %ecx, %edx, %esi, %edi, %ebp. ## ##### ## ## ## ## ##  
*most* ## ## #####. ##### (#### ##### clone) ##### # ##### ## ## ##  
##### ## ## ## ##### # ##### ##### ## ## linux\_clone #####.

##### 4.1.2. #####

##### ## ## ##### ## ## ## ## ## ##—  
##### ## syscalls.master. ## ## ## ## ## ##:

```
...
AUE_FORK STD { int linux_fork(void); -}
...
AUE_CLOSE NOPROTO { int close(int fd); -}
...
```

### ##### ##### #####. ### ##### ##### ## ## ##—  
#####. ### ##### ##### ##### #####. ## ## ##### STD,  
OBSOL, NOPROTO ### UNIMPL. STD ## # ##### ##### ##### ##—  
#####. OBSOL ## ##### ## ##### #####. NOPROTO #####  
##### ## ##### ##### ## ## ## ##### ## #####, ###. UNIM-  
PL ##### ## ## ##### ## ## ##### ## ## nosys ##### (# #####  
##### ## # ##### ##### ## ##### ##### ##### ## ##—  
### ENOSYS).

##### syscalls.master # ##### ##### #####: linux\_syscall.h, linux\_proto.h ##  
linux\_sysent.c. ### linux\_syscall.h ##### ##### ## ##### ## ##—  
##### ##., #.:

```
...
#define LINUX_SYS_linux_fork 2
```

#####

```
...
#define LINUX_SYS_close 6
...
```

```
### linux_proto.h #####
```

```
struct linux_fork_args {
    register_t dummy;
};
```

```
### #####, linux_sysent.c ##### ##### ##### #####  
##### # ##### # #####, #.:
```

```
{ 0, (sy_call_t *)linux_fork, AUE_FORK, NULL, 0, 0 -}, -/* 2 = linux_fork */
{ AS(close_args), (sy_call_t *)close, AUE_CLOSE, NULL, 0, 0 -}, -/* 6 = close */
```

```
## ### linux_fork ## ##### ## ##### ## ## ##  
## STD ##### ## #####, ##### ## ##### ## ## ##  
#####. ## ## ##### close ## ## ## ## ## ## ## ## ##(2) ## ## ##  
#####  
#####  
#####(2) ## ## ##.
```

#### 4.1.3. #####

```
#####® ##### ## ## ##, ## ## ## ## ##
##### ## ## ## ## ##. ## ## ## ## ##—
##### ## ## ##### ## ## DUMMY ####. ##### ##—
##### ## ## linux_dummy.c ## ## ## DUMMY(syscall), ##### ## ##
##### ##### ## ## ## #####
# ##### ##### ## ## #####. ## UNIMPL ##### ##
### ##### ## ## ## ## ##### ## ## ##
##### ## ## ##### ## #####.
```

## 4.2. #####

[illegible]

4.2.1. #####® #####

```
##### # SA_SIGINFO
## ##### linux_rt_sendsig #####. ## 
(## #####) #####, ## 
#####. ## 
#####. ## 
#####. ## 
##### (#####, ##### 
#####). #####, ## 
#####
```

#####<sup>®</sup> #####<sup>®</sup> ## #####

---

#### 4.2.2. #####\_##\_#####

#### ##### ## ##### ## linux\_send\_sig #### ## ##### ##### ##### ##### ## ##—  
#####. ## #### signinfo, ucontext, #### #####<sup>®</sup> #####. ## ##### ## ##### #####—  
#### ##### ##### ##### ##### ##### ## ## ##### ##### # ##### ## #####  
##### ##### ##### ##### ##### #####.

#### 4.2.3. #####\_#####

#### ##### ## ##### ## ##### ##### ## ##### #####. ## ##### #####  
##### ## ##### ## ##### ##### #####. ## ##### ##### ## ##### ##  
##### ##### #####.

#### 4.3. #####

#### #####<sup>®</sup> ##### ##### ##### ## #####(2) ##### ## ##### ## ##### ##—  
### ##### ## ##### #####. ##### ##### ##### ## ##### #####  
## ##### ##### ##### ##### ## ##### #####, ##### #####  
### ##### ## ## ##### ##### ##, ###. ### ##### ## ##### ##  
#### ## ##### ## ##### ## ##### ##### (##### ##  
#####(2) ##### ## ## ##### ##### (#####  
###). #####(2) ## # #####<sup>®</sup>##### ##### ## #####<sup>®</sup>## #####  
#####.

#####<sup>®</sup> ##### ## ##### ##### ## #####(2) ##### ## linux\_ptrace.c. ##  
##### ## ##### ##### ##### #####<sup>®</sup> ## ##### ## ##  
#####(2) ##### #####. ### ##### ## # ##### ##### ##—  
##### ## ##### ## ##### ## #####(2) #####. ### #####(2)  
##### ## ##### ##### #####<sup>®</sup> ## ##### ## ##### ##  
##### ## #####. ### #####, PT\_GETREGS ## #####<sup>®</sup> ##### ##  
#### ##### ## ## ##### ## ## ##### ## (#####)  
#####(2) #####, # ##### ## ##### ## #####<sup>®</sup> #####.

### #####(2) ##### ## ##### ## #####. #####  
#### ##### ## ## ##### strace (##### ## # #####(2) #####) ## ## ##—  
#####. ##### PT\_SYSCALL ## ## #####.

#### 4.4. #####

##### # #####<sup>®</sup> ##### ##### ## ## ##### ##### #####  
## ##### ##### ##### ## ## ##### #####. ##—  
##<sup>®</sup> ## ##### ##### ## ##### ## ##### # ## ##### ##  
#### ## ##### ##### #####:

```
static int
translate_traps(int signal, int trap_code)
{
```

```
if (signal != SIGBUS)
    return signal;

switch (trap_code) {

    case T_PROTFLT:
    case T_TSSFLT:
    case T_DOUBLEFLT:
        case T_PAGEFLT:
            return SIGSEGV;

    default:
        return signal;

    -}
}
```

#####® ##### ## #####

#####  
#####  
#####  
#####  
#####  
#####

5.2. #####® 2.6 #####

5.2.1. ##### ## 2.6 #####

5.2.2. #####® ##### ### ##### #####

```
##### ## # ##### ## # clone ##### # ##### #  
##### (#### #) #### # ##### 1:1 #####.
```

```
struct linux_emuldata {
    pid_t pid;

    int *child_set_tid; /* in clone(): Child.s TID to set on clone */
    int *child_clear_tid; /* in clone(): Child.s TID to clear on exit */

    struct linux_emuldata_shared *shared;

    int pdeath_signal; /* parent death signal */

    LIST_ENTRY(linux_emuldata) threads; /* list of linux threads */
};
```

```
struct linux_emuldata_shared {
    int refs;

    pid_t group_pid;

    LIST_HEAD(, linux_emuldata) threads; /* head of list of linux threads */
};
```

24



#####® ##### ## #####

### linux\_emuldata ##### ## ## ##### ##### em\_find. ### ##-  
##### ## ## ##### ##:

```
struct linux_emuldata *em_find(struct proc *, int locked);
```

####, proc ## ## ##### ## ##### ##### ##### ##### ## ## ##  
##### ##### ##### ## ##### ## ##### ## ##. ### ##### ##### ##  
EMUL\_DOLOCK ### EMUL\_DOUNLOCK. ##### ##### #####.

### 5.2.3. #####

##### ## ## ##### ##### ##### ##### ## ##### ## ## ##### ##  
## ##### ##### #####® ## ##### ## ##### ##### #####. ## ## ##  
### #####. ##### ##### ## ##### ## ## (=####) ### ## (=###) ## #####  
##### ## #####. ### ##### ## ##### ## ##### ## ##### (## #####) ## =  
### ## ## = ##### > ##### ## ## ## ##### ## ##### PID = shared -> group\_pid  
### TID = proc -> p\_pid. ### ## ##### ## linux\_emuldata structure ## # ##### ##.

### ##### ##### #####, #####, ##### #####. ##### ## ## ##/####  
#####. ## ##### ## ##### ## child\_clear\_tid ### child\_set\_tid ## ##### ##  
###.

### 5.2.4. #####

### clone ##### ## ## ## ##### ## ##### ## #####®. ### #####  
##### ## ##:

```
int linux_clone(1_int flags, void *stack, void *parent_tidptr, int dummy,  
void *child_tidptr);
```

### flags ##### ##### ##### ##### ##### ##### ## #####.  
## ##### #####, #####® ## ##### ##### ##### ##### ##-  
#####, ## ##### ## ##### ##### ##### ##### ##### ## ## ##.  
#### ## ## ## flags ##### ## ## ##### ##### ## ## ##### #####.  
### stack ##### ## ##NULL #####, ##### ## ##### ##### ## ## ## NULL  
## ## ##### ## ##### ##### ## ##### ##### ##### (#.#. ## #####  
####(2) #####). ### parent\_tidptr ##### ## ## ## ## #####  
### ##### ## (#.#. ##### ##) ##### ## ##### ## #####  
### ## ## ##### ##. ### dummy ##### ## ## ##### ## ## ##  
##### ##### ## ## ##### ## #386. ## ## ## #####  
#### ## ## ## ##### ## ## ## ##### ## ## ## ## ## ##. ###  
child\_tidptr ##### ## ## ## ## ##### ## ##### ## ## ## ##  
##### ##### ## ## ## ##### #####.

### ##### ##### ##### ## ##### ##### ##### ## ##  
##### ##. ### #####, CLONE\_VM ##### ## ##### (##### ## ##), ##. ###  
##### ## ## CLONE\_FS ### CLONE\_FILES ##### ##### ## ## #####

```
### clone ##### ## #### ## #### ## #### ## ##### ##### #####(2) ###
#####(2) #####. ##### #### # #### ## 2.6 ##### #### clone ## #####
#####(2) ### #####(2) #####.
```

```
##### ## ##### ## ## ##### ##### ## ## ##### # ## ##
##### ## #####. ##### ## ## #####: emul_lock ##### ## ##### ##### ##
linux_emuldata ### emul_shared_lock ##### ## ##### linux_emuldata_shared. ### emul_lock ## #
##### ##### ##### ##### emul_shared_lock ## # ##### ##### sx_lock.
##### ## ## ##### ##### ## ## ##### ##### ##### ## ## ##
## ## ##### ## ##### #####.
```

#####

[illegible]



#####

5.3.4. ##### ## #####® ###

#### 5.3.4.1. #386

```
##### set_thread_area #####  
### clone ## ### clone. #####  
#####. ####  
#####, #####  
##### set_thread_area #####.  
#####® #####.  
#####  
#####  
#####® ## ##. ## 6, 3 ## #1. ## 6 ## #-  
#####®, 3 ## ## #1 #####.  
##### 3 ## ## ## ##  
#####  
##### (#####) ##  
##### 1. #####  
#####® ##### (#.#. #####)  
#####. #####  
##. ##### (#####) ##  
%gs ##### load_gs. #####  
#####. ### CLONE_SETTLS #####  
##### load_gs ## ##. ##### (#####  
3) ## #####® #####  
#####® #####.
```

#### 5.3.4.2. ###64

```
### ##64 ##### ## ##### ## ## 386 ### ## ##### ## ##### ##
32### ##### ##### ## ## ##### (##### ## ##### 32### ##
##### ##) ## ## ## ## ## ## # ##### ## ##### ## ## ##### ##
##### ## (#### # #### ##### ## ## 32### ## ##). ##### ## ## ##
## ##### ## ##### ## ## ## ## ## ##### #####). ##### ## ##
##### ## ## ## ##### ##### #####.
```

#### 5.4. #####

5.4.1. ##### ## #####

```
##### # ## # ##### ® ##### # # #:#-
#### # #####, ##### # ##### 
### # ##### # ##### # # .
## # #####® ##### # # -
#####. ##### 
## ### ##### # #. # # 1:(#####) #### #-
##### # # ##### # # # # # (#)
```



- FUTEX\_WAIT
- FUTEX\_WAKE
- FUTEX\_FD
- FUTEX\_REQUEUE
- FUTEX\_CMP\_REQUEUE
- FUTEX\_WAKE\_OP

#### 5.4.3.1. #####\_####

#### ##### uaddr ### ##### val ## #####. ## ###, EWOLDBLOCK ## #####, ##### ## ##### ## ## ##### ## #####. ## ## ##### timeout ## ##### ## ##### ## ## ##### ## ## #####, ##### ## ##### ## #####.

#### 5.4.3.2. #####\_####

#### ##### # ##### ## uaddr ### ##### ## val ##### ##### ## #####. #####.

#### 5.4.3.3. #####\_##

#### ##### # ##### ##### # ##### #####.

#### 5.4.3.4. #####\_#####

#### ##### val ##### ## ##### ## uaddr, ##### ##, ## ##### val2 ##### ## ##### ## ##### ## uaddr2.

#### 5.4.3.5. #####\_###\_#####

#### ##### ## Futex\_REQUEUE ## ## ##### val3 ##### ## val #####.

#### 5.4.3.6. #####\_####\_##

#### ##### ## val3 (##### #####) ## uaddr. ##### ## val ##### ## uaddr ## ## ##### ## ##### # ##### ## val2 ##### ## uaddr2.

### ##### Futex\_WAKE\_OP:

- FUTEX\_OP\_SET
- FUTEX\_OP\_ADD
- FUTEX\_OP\_OR

#####@ ##### ## #####

---

- FUTEX\_OP\_AND
- FUTEX\_OP\_XOR



####

```
##### ## ## val2 ##### ## ## ##### #####. ### val2 ##
##### ##### ## struct timespec *timeout ##### ## #####
FUTEX_REQUEUE, FUTEX_CMP_REQUEUE ### FUTEX_WAKE_OP.
```

#### 5.4.4. ##### ## #####

```
### ##### ## ##### ## ##### ## ##### ## ##### ## ##.
## ## ##### ## linux_futex.c ### linux_futex.h #####. ### futex ##### ## ##:
```

```
struct futex {
    void *f_uaddr;
    int f_refcount;

    LIST_ENTRY(futex) f_list;

    TAILQ_HEAD(If_waiting_proc, waiting_proc) f_waiting_proc;
};
```

```
### ## ##### waiting_proc ##:
```

```
struct waiting_proc {

    struct thread *wp_t;

    struct futex *wp_new_futex;

    TAILQ_ENTRY(waiting_proc) wp_list;
};
```

##### 5.4.4.1. #####\_### / #####\_###

```
# ##### ## ##### ##### ## futex_get #####, ##### ##### # ##### ##
##### ## ##### ## ##### ## ## ##### # ## #####. ##### ##### # #####
##### ## ## ## ##### futex_put #####, ##### ##### # ##### ##### ##
##### ## ## ## ##### ##### ## ## #####.
```

##### 5.4.4.2. #####\_#####

```
#### # ##### ##### # ##### ## ##### # working_proc ##### ##
##### ##### ## ## ##### ## ##### ##### ## ## #####
```

#####

```
# #####(9) ## #####. ### ##### ## #####. ##### #####(9) ##-
##### (### ##### ## ## ##) ### working_proc ##### ##
##### ## ## ## #####. ### ## ## ## futex_sleep #####. ## ##
##### ## futex_wake ## ## wp_new_futex ## ## ##. ##### ##
##### ## ## ## #####.
```

5.4.4.3. #####

[illegible]

5.4.4.4. ##### ### ##

```
### FUTEX_WAKE_OP ##### ## #####. ##### ## ##### ## #####
## ##### uaddr ### uaddr2 ##### ## ##### ## ##### ##### val3 ##
uaddr2. ##### val ##### ## ## ##### ## ##### ## ## ## #####
##### ##### ## ##### ## val2 (#.#, timeout) ##### ## ## #####.
```

5.4.4.5. #####

```
##### encoded_op ### uaddr. ### #####-
##### ###, #####, ###
#####. ### ##### ## ## ##:
```

```
oldval = *uaddr2
*uaddr2 = oldval OP oparg
```

```
##### # #####. ##### # ##### # # # # # uaddr # #####  
### ## ##### # ####. ### ##### ##### ##### # # # # #—  
#### oldval # ##### # cmparg ##### # # #####.
```

5.4.4.6. #####

```
##### sx_lock ### ##### (#-
##### sx_lock). ##### ## ##### ##
## #### #.
```

5.5. #####

##### # ## ##### # ##### ##### ##### ##### ##### # #  
##### ##### ##### ##### ##### # # # ##### # ##### ##### # #  
##### ##### ##### # # #



#####® ##### #

### 5.5.1. #####

##### 2.6.16 #####, #####. #####  
##### (openat #####) #####  
##### dirfd #####. #####  
#####, #. ##### filename #####  
##### dirfd #####, #####  
#####. ##### dirfd #####  
#####. ##### dirfd ##### AT\_FDCWD.  
##### openat #####:

file descriptor 123 = -/tmp/foo/, current working directory = -/tmp/

```
openat(123, -/tmp/bah\, flags, mode) /* opens -/tmp/bah */
openat(123, bah\, flags, mode) /* opens -/tmp/foo/bah */
openat(AT_FDCWD, bah\, flags, mode) /* opens -/tmp/bah */
openat(stdio, bah\, flags, mode) /* returns error because stdio is not a directory */
```

#####  
#####  
#####. #####  
##### #. ##### open(/tmp/foo/bah., flags, mode) #####  
#####  
##### /tmp/foo/. #####. ##### /tmp/foo #####  
##### dirfd #####. #####  
#####.

#####® ##### #: linux\_openat, linux\_mkdirat, linux\_mknodat,  
linux\_fchownat, linux\_futimesat, linux\_fstatat64, linux\_unlinkat, linux\_renameat, linux\_linkat, linux\_symlinkat,  
linux\_readlinkat, linux\_fchmodat ### linux\_faccessat. #####  
#####(9) #####.

#### 5.5.1.1. #####

#####(9) ##### (#####)  
##### dirfd ##### nameidata #####  
#####  
##### dirfd #####  
##### dirfd ##### AT\_FDCWD ##### dvp ##### nameidata  
##### NULL ##### dirfd #####  
#####  
#####. #####  
#####(9) ##### dvp ##### dp #####  
#####(9) #####. #####(9) #####  
#####  
##### openat #####:

openat() ---> kern\_openat() ---> vn\_open() --> namei()

#####

```
### ##### kern_open ### vn_open #### ## ##### ## ##### #####  
dirfd #####. ## ##### ##### ## ##### ##### ##### #####  
##### ## ##### ## ## ##### ## ## ##### #####. #####  
##### ##### ## ##### ##### ## ## #####. #####  
##### ##.
```

### 5.5.2. #####

[illegible]

### 5.5.3. #####

#####  
#####. #####  
#####. #####  
##### (#####). #####  
#####. #####  
#####.

6. #####

6.1. #####

```

##### 2007 #####® ##### ## ##### ## ##### ## ##-
##### 2.6.16 ##### #####. ##### ##### #####, #####
##### ## #####, ##### #####, ##### epoll ## inotify ##
##### ## ## ## ## ##### ##. ##### ## ## ## ##### ##
##### ##### ## ## ##® ##### ##### ## ##### #####
##### ##### 4 ## 2.6.16 ## ##### ##### ##### ## #####

```

#####<sup>®</sup> #####

---

#### ##### 6 ## 2.6.16. ### ##### 6 #####\_#### ### #####  
##### ##### ##### ## ## ##### ##### ## #####  
##### ## ##### ## ## ##### ## #####

## ## ## ## ## ## ##### ##### ##/##### , ##/##-  
##### , #####/##### ## ## ## ## ## ## ## ## ##-  
##### ##### ## ##### 2.6 ##### ## ## ## ##-  
##### ## ##### ## ## ## ## ## ## ## ## ##  
##### ## ## ## ## ##<sup>®</sup> #####<sup>TM</sup> ##### ## ## ## ## ##  
## ## ##### ## epoll ##### ## ## ##### ## ## ##<sup>®</sup>  
##### 2.6.

## ## ## ##### 2.6.16 ##### ## ##### ##### 7.0 ## #-  
##### ## ##### ## 2.6 ##### ##### ## #####  
#### ## ##### ## ##### 6 #####\_#### , ##### ## #####

## 6.2. #####

##### ##### ##### ## ##### ## ##### #####  
## ## ## ## ## ## ## ## ## ## ## ## ##  
##### ## epoll ## inotify #####.

## ## ## ## ## ## ## ## ## ## ## ## ## ## ## ##  
## ## ## ## ## 2.6 ##### ## ## ## ## ## 6 ##  
##### #####\_#### ##### ## ## ## ## 4 ## ## ##  
## ##.

## ##### ##### ## ## ## ## ## ## ## ## ## ## ## ## ##  
## ## ## ## 2.6 ##### ## ## ## ## ## ## ## ##-  
##. ##### ## ## ## ## ## ## ## 2.6 #####.

##### , ## ##<sup>®</sup> ##### ## ## ## ## ## ## ## ## ## ##  
##### ##### #####. ##### ## ## ## ## ## ## ##-  
##### ##### ## ## ## ## ## ## ## ## ## ## ## ## ## ## ##  
##### ## ## ## ## ## ## ## ## ## ## ## ## ## ## ##.

## 6.3. ####

# ##### ## ##### ## ## (## #####):

- ##### <jhb@FreeBSD.org>
- ##### <kib@FreeBSD.org>
- #####
- #####

#####

---

- ##### <[jkim@FreeBSD.org](mailto:jkim@FreeBSD.org)>
- ##### <[netchild@FreeBSD.org](mailto:netchild@FreeBSD.org)>
- ##### <[ssouhlal@FreeBSD.org](mailto:ssouhlal@FreeBSD.org)>
- ## #####
- ##### <[davidxu@FreeBSD.org](mailto:davidxu@FreeBSD.org)>

# ##### ## ##### ## ##### ## ##### ## #####, ##### ## #####  
#####.

## 7. #####

1. ##### # ##### #. #####. ##### ## ##### ##  
## #####. #####, 2005.
2. <http://www.FreeBSD.org>
3. <http://tldp.org>
4. <http://www.linux.org>