

#####® ##### ## #####

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

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

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

#####.

#####.

```
##### ## # ##### ##### ## ### ##### #####.
```

#####,
##.

#####.

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

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

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

#####

```
#####  
##### (### ## ##### Linuxulator). ### #### #  
##### ## ##### ## #####® 2.6. ## #  
#####, ## #####® 2.6.16 ##### #  
#####-  
#####  
##### 2006 ## #  
#####.  
NPTL (#####® #####) #####  
#####, ##### TLS (#####), futexes (#####)
```

#####

1. #####

2.2.2.

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

2.2.3.

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

2.2.4.

#####® ##### ## ##### ## ## ## ##### #####
#####® ##### ## ##### ## ## ## ##### #####
#####. ##### ##
(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,

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

2.4. #### ## #####®

#####® ## # #####®##### ##### ##### ##### ## ##### #####, ## ##
##.
#####, #### ##### ##### ##### ##### ## ##
#####, #####® ## ##### ##### ##### ## ## ##### #####, ## ##
#####.

#####® ##### ##### ## 1991 ## # ##### ##### ## ##### ##
#####. ##### ## ## ##### ## ## ##### ## # #####
#####®##### #: #####, #####, #####, #####,
#####. ##### ## ##### ##### ##### ##
##.

2006 #####® ##### ## ## ## ##### ##### ## ## ##### ##### ##—
#####, #####,
#####, ##. #### ## ## ##### ##### ##### ## #####® ## ## ##
##—
#####.

#####® ##### ##### ## # ## ##### ##### #####. ## ##
#####® ####, ## ##
#####. ## ##### ##### ##—
#####® ##### ## ## ##### ##.#.#. ##### ##### ##—
2.6.#.#, ##### # ##### #####, ##### ## ##### ## ## ##
##.

[4].

2.4.1.

#####® ##### ## ##### #####® ##### ## ##### ## ## ## # #####
#####: ## ##### ## ## #####. ## ##### ## ## ##### ## ## ##:
#####. ## ##### ## ##### ## ## ##. ## #####
#####® 2.6 ## ## #386™ #####. #####
[3].

2.4.1.1.

#####® ## ##### (## #####) ##### syscallX ##### ## ##—
#####.
%eax ##### ## ##
0x80. ##### ##### ##### ## ##, #####
ermo ##### ## ## res ## -1

#####® ##### ## #####

#####. ##### ## ##### 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_FILES # ##### #####
- CLONE_SIGHAND # ##### ##### ##### ## #####
- CLONE_PARENT # #####
- CLONE_THREAD # ## ##### (##### #####)
- CLONE_NEWNS # ### #####
- CLONE_SYSVSEM # ##### #####
- CLONE_SETTLS # ##### ## #####
- CLONE_PARENT_SETTID # ### ## ##
- CLONE_CHILD_CLEAR_TID # ##### ## ##
- CLONE_CHILD_SETTID # ### ## ##

CLONE_PARENT ##### ## ##### ## #####. ##### ## #####
 ##### ## ##### # ##### ##### # ## ##### # ## ## #####—
 ## ## ##### ## ##### #####. CLONE_THREAD ##### #####
 ##### ## CLONE_PARENT, CLONE_VM ## CLONE_SIGHAND, ##### ## ## ##
 ## ## ## #####, ##### ##### ## ## ##### ## #####. CLONE_SETTLS ##### ## #####. ## CLONE_*_TID ## ##
 #####/##### ##### ## ## ## 0.

CLONE_THREAD ##### ## ##### ## ##### ## ## ##
 ##### #####. ## ##### (##### ## #####, #####
 ## ##### ## ##) ## # ##### ##### ## ## #####
 ##### ## ##### ##### ##### ## ## #####
 ##### #####. ## ## ##### ## ## ## ## ##
 ## ## ## ## ##### ## ## ## ##### # ##### ##

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

#####(3) ## ## ## ## ## ## ## ## ## ##
 ####:

```
int clone_flags = (CLONE_VM -| CLONE_FS -| CLONE_FILES -| CLONE_SIGNAL
-| CLONE_SETTLS -| CLONE_PARENT_SETTID
| CLONE_CHILD_CLEAR_TID -| CLONE_SYSVSEM
#if __ASSUME_NO_CLONE_DETACHED == 0
```

```
#####@ ##### ## #####
```

```
| CLONE_DETACHED
#endif

| 0);
```

```
### CLONE_SIGNAL ## #####
```

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

```
### 0 ##### ## ##### ## ##### ## ## ## #####
```

2.5.

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

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

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

#####. ##### ## ##### ## ##—

! ## ## ## ## ## ##, ##
#####. ## ## ##—
#####. ## ##
#####. ##

#####

(# #)

#

(###) #####

(##)

[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. #####®

#####

#####[®] #####

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

4.1.

linux_sysvec.c, ##### ##### ## ## ##—
sysentvec #####. ##### # #####[®] ##### ##
#####, ## ##### ##### ##### ##### #####
#####[®] ##.

4.1.1. #####[®]

#####[®] ##### ##### ## ##### ##### (#### ## ## ## ##
6 ##### ## #386) ##### ##### #####. #####[®] #####
#####. ##### ## ## ##—
#: %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.

#####®
linux_sendsig

4.2.1. #####®

SA_SIGINFO
linux_rt_sendsig #####
(## ##)

sigreturn #####
#####

#####[®] #####[®] ## #####

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

linux_send_sig #### ## ##### ##### ##### ##### ## ##—
#####. ## #### signinfo, ucontext, #### #####[®] #####. ## ##### ## ##### #####—

#####.

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

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

4.3.

#####[®] ##### ##### ##### ## #####(2) ##### ## ##### ## ##### ##—
#####. #### ##### ##### ## ##### #####
#####, #### ##### #####,
##, ###. ### ##### ## ##### ##
(##### ## #####).
#####(2) ##### ## ## ## ##### ##### ##### ## ## ##### ##### (#####
###.). #####(2) ## # #####[®]##### ##### ##### ## ##### ##[®]## ##### ##
#####.

#####[®] ##### ## ##### ##### ##### ## #####(2) ##### ## linux_ptrace.c. ##
#####[®] ## ##### ## ## #####
#####(2) ##### ##### #####. ### ##### ## # ##### ##### ##—
#####(2) #####. ### #####(2)
#####[®] ## ##### ## ##### ##
#####. ### #####, PT_GETREGS ## #####[®] ##### ## #####
(#####)
#####(2) #####, # ##### ## ## ##### ## #####[®] #####.

#####(2) ##### ##### ## ##### ## #####. #####
strace (##### ## # #####(2) #####) ## ## ##—
#####. #### PT_SYSCALL ## ## #####.

4.4.

#####[®] ##### ##### ## ## ##### ##### ##### #####
#####. ##—
##[®] ## ##### ##### ## ##### ## ##### # ##### ## ##### #####
#####:

```
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;

    -}
}
```

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

[illegible]

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

#####® 2.6 ### #####
###. ##### 2.6. ### #####® ##### ##### ## ## #####


```
##### ## ## ##### ## # #### ## ### 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. ##### ## #####

#####

- 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_QUEUE, FUTEX_CMP_QUEUE ### FUTEX_WAKE_OP.
```

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

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

5.4.4.1. ##### ### / ##### ###

5.4.4.2.

31

#####

#####(9) ## #####. ## #####. #####(9) ##—
(#####) ## working_proc #####
#####. ## futex_sleep #####. ##
futex_wake ## wp_new_futex ##. ##
#####.

5.4.4.3. #####_####

futex_wake #####.
#####®, #####

futex_sleep.

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 ## openat #####. ##### ##### ##### ## #####
#####.

#####® ##### ## ## ##### #####: 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 ## #####

```

#####[®] #####

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

#####

##/##### , ##/##-
, #####/##### ## ## ## ## ## ## ## ## ##-
2.6 ##### ## ## ## ##-

##[®] #####TM ##### ## ## ## ## ##
epoll ##### ## ## ##### ## ## ##[®]
2.6.

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

6.2.

epoll ## inotify #####.

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

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

, ## #####[®] ##### ## ## ## ## ## ## ## ## ## ## ##
#####. ##### ## ## ## ## ## ## ## ##-

##.

6.3.

(## #####):

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

#####

- ##### <jkim@FreeBSD.org>
- ##### <netchild@FreeBSD.org>
- ##### <ssouhlal@FreeBSD.org>
- ## ####
- ##### <davidxu@FreeBSD.org>

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

7.

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