

Programmazione di Sistema e di Rete

Matteo Aprile
Professore: Franco Tommasi

INDICE

I	Libri di testo consigliati	1
II	Comandi utili	2
II-A	find: trovare tutti i file eseguibili	2
II-B	find: trovare file di intestazione del mac come stdio.h	2
II-C	lld: per capire che librerie usa il codice	2
II-D	gcc: per vedere tutta la gerarchia di file in una libreria	2
II-E	gcc: per vedere il codice con tutti i file importati	2
II-F	gcc -g: debugging debole	2
II-G	gcc -ggb: debugging forte	2
II-H	xattr: usato per i file che entrano in quarantena su MacOS	2
III	Variabili di sistema definite in .bashrc	2
III-A	INC	2
IV	Introduzione	2
IV-A	System call	2
IV-B	Programma Make	2
IV-C	Direttive di preprocessore	3
IV-D	Librerie	3
IV-E	Creazione librerie	4
IV-F	Aggiornamento librerie	4
V	System call	5
V-A	Funzioni e system call	5
V-B	Capire se una funzione è una system call	6
V-C	Numeri dei file descriptor	6
V-D	Meccanismi dei file	6
V-E	Unbuffered I/O	6
V-F	Fork & Exec	7
V-G	Thread	8
V-H	Gestione degli errori	8
V-I	Keyword in C ed accessi a variabili . .	8
V-J	Segnal e interrupt	9
V-K	Valori del tempo	9
VI	Gli standard	10
VI-A	Storia e basi	10
VI-B	Limiti	10
VI-C	Determinare l'allocazione	10
VII	File I/O	11
VII-A	Chiamata open	11

I. LIBRI DI TESTO CONSIGLIATI

- Advanced Programming in the Unix Environment, 3th ed, Stevens, Rago
- TCP/IP 1, Stevens (facoltativo)
- Unix Networking Programming the Socket Networking API, Stevens
- The Linux Programing Interface, Kerrisk
- maset
- Gapil Guida alla Programmazione in Linux, Simone Piccardi

II. COMANDI UTILI

A. *find*: trovare tutti i file eseguibili

```
1 $ find . -type f -perm -0100
2 ./standards/makeopt.awk
3 ./standards/makeconf.awk
4 ./proc/awkexample
5 ./systype.sh
6 ./advio/fixup.awk
```

B. *find*: trovare file di intestazione del mac come *stdio.h*

```
1 $ find /Applications/Xcode.app/ -name stdio.h 2>/dev
   /null
```

C. *ldd*: per capire che librerie usa il codice

```
1 $ ldd [nomevoci]e
```

D. *gcc*: per vedere tutta la gerarchia di file in una libreria

```
1 $ gcc -H lib.a
```

E. *gcc*: per vedere il codice con tutti i file importati

```
1 $ gcc -E file.c
```

F. *gcc -g*: debugging debole

```
1 $ gcc -g -ansi -I../include -Wall -DMACOS -
   D_DARWIN_C_SOURCE ls1.c -o ls1 -L../lib -lapue
```

G. *gcc -ggb*: debugging forte

```
1 $ gcc -ggb -ansi -I../include -Wall -DMACOS -
   D_DARWIN_C_SOURCE ls1.c -o ls1 -L../lib -lapue
```

H. *xattr*: usato per i file che entrano in quarantena su MacOS

```
1 $ xattr -d (delete) com.apple.quarantine [path sh]
```

III. VARIABILI DI SISTEMA DEFINITE IN .BASHRC

A. *INC*

```
1 INC="/Applications/Xcode.app/Contents/Developer/
   Platforms/MacOSX.platform/Developer/SDKs/MacOSX.
   sdk/usr/include/"
```

IV. INTRODUZIONE

A. *System call*

Sono uguali alle funzioni di libreria dal punto di vista sintattico, però cambia il modo di compilarle. Notare che non possono essere usati i nomi delle SC per delle function call.

Per poi poter "raccontare" tra umani le sequenze di bit che vengono mandate ai processori si usa **assembly**.

Sono effettivamente delle chiamate a funzioni ma poi dal codice assembly puoi capire che è una system call dato che ha dei meccanismi specifici.

Alcuni esempi di chiamate e registri:

- **eax** : registro dove metti il **numero della sc**
- **int 0x80**: avvisa il **kernel** che serve chiamare una sc
- **exit()**: chiudere un processo
- **write()**:

```
1 mov edx,4      ; lunghezza messaggio
2 mov ecx,msg    ; puntatore al messaggio
3 mov ebx,1      ; file descriptor
4 mov eax,4      ; numero della sc
5 int 0x80
```

dove nel **file descriptor** indichi a quale file devi mandare l'output. Questo viene usato dato che così non deve cercare il path ogni volta ma lo mantiene aperto riferendosi ad esso tramite un numero, cioè il più piccolo disponibile.

B. *Programma Make*

Quando viene avviato verifica la presenza di un file chiamato "Makefile", oppure si usa 'make -f'. In questo file ci sono le **regole di cosa fare per automatizzare delle azioni per un numero n di file**. Se, durante la compilazione di massa, una di queste da un errore il programma make si interrompe, per evitare ciò si usa '-i' (ignore).

Il Makefile andrà ad aggiornare una libreria andando a guardare se una delle 3 date di ultima modifica si sono aggiornate.

Andiamo a guardare **cosa contiene Makefile**:

```
1 DIRS = lib intro sockets advio daemons datafiles db
   environ \
2 fileio filedir ipc1 ipc2 proc pty relation
   signals standards \
3 stdio termios threadctl threads printer
   exercises
4
5 all:
6   for i in $(DIRS); do \
7     (cd $$i && echo "making $$i" && $(MAKE) ) ||
8     exit 1; \
9   done
10 clean:
11   for i in $(DIRS); do \
12     (cd $$i && echo "cleaning $$i" && $(MAKE)
13     clean) || exit 1; \
14   done
```

dove:

- **DIRS**: lo si associa alle **stringhe singole** che gli sono state associate
- **all**: nel ciclo for:
 - manda un comando in subshell

- \$\$i: riferimento alla variabile "i" del for + simbolo escape per il Makefile
- \$(MAKE): macro predefinita per i Makefile

La struttura è:

```
1 target: prerequisiti
2 rule
```

dove:

- **target**: è la cosa che si vuole fare, se essendo il primo target, sarà anche quello di default
- **prerequisiti**: file e/o target a loro volta
- **rule**: indica cosa può fare il target

Può capitare che prima di eseguire il Makefile ci sia uno script "configure".

In molti casi si ha un target "clean" che permette di pulire i file .o che sono inutili dopo la compilazione, o comunque qualsiasi tipo di file gli si voglia far eliminare. Questo tipo di target che non rappresentano un file, sono detti "phony" perchè fasulli, dato che non sono file ma sole parole

```
1 file: file.o lib.o
2
3 clean:
4 rm file.o
```

Abbiamo delle variabili automatiche per rendere il lavoro più facile:

- \$@: per riferirsi il target
- \$?: tutti i prerequisiti più recenti del target
- \$*: tutti i prerequisiti del target
- <https://www.gnu.org/software/make/manual/make.html#Automatic-Variables>

Un altro esempio di Makefile è:

```
1 ROOT=..
2 PLATFORM=$(shell $(ROOT)/systype.sh)
3 include $(ROOT)/Make.defines.$(PLATFORM)
4
5 PROGS = getcputc hello ls1 mycat shell1 shell2
6         testerror uidgid
7
8 all:      $(PROGS)
9
10 %:  %.c $(LIBAPUE)
11     $(CC) $(CFLAGS) $@.c -o $@ $(LDFLAGS) $(LDLIBS)
12
13 clean:
14     rm -f $(PROGS) $(TEMPFILES) *.o
15
16 include $(ROOT)/Make.libapue.inc
```

dove:

- ROOT: cwd
- PLATFORM: assumerà in valore del OS: macos/linux
- include: include un file
- PROGS: elenco dei programmi da usare
- %: target con nome variabile, indica un file
- %.c: target con nome variabile ma estensione .c
- \$(CC): indica il compilatore dove cc è un link simbolico a clang
- \$(CFLAGS) indica una macro predefinita dei default vuota che si può usare all'occorrenza

- all: target che prende in carico tutti i programmi che se saranno di tipo .c saranno presi in carico dal target successivo

Per la compilazione dei file, qualsiasi sia il linguaggio, make saprà come compilarlo grazie a tutte le definizioni di default presenti in:

```
1 $ make -p
```

notare che si può mettere un comando custom nelle rule del target

Nell'eventualità di voler aggiornare un solo file della libreria senza far aggiornare il resto ci basterà usare uno script che compila quel file passato a linea di comando.

```
1 $ gcc -ansi -I../include -Wall -DMACOS -
   D_DARWIN_C_SOURCE ${1}.c -o ${1} -L../lib -
   lapue
```

C. Direttive di preprocessore

Sono delle indicazioni date a gcc prima di iniziare la compilazione.

Iniziano tutte con '#':

- **#include**: serve ad includere delle librerie di sistema (<lib.h>) oppure di librerie fatte da noi e non in directory standard ("lib.h")
- **#define**:
 - permette di creare delle "macro", che vanno a sostituire una stringa con un'altra (es: #define BUFLLEN), può capitare che debbano essere definite delle macro prima che si compili il programma, in questi casi si usa scrivere es: '-DMACOS'
 - permette di creare delle "function like macro" (es: #define ABSOLUTE_VALUE(x) (((x<0)?-(x):(x)))
- **#ifdef**, **#ifndef**, **#endif**: usata per far accadere qualcosa nel caso un macro sia stata definita

```
1 #ifdef VAR
2 print("hello");
3 #endif
```

Per evitare che più file includano lo stesso si usano degli #ifndef in tutto il codice, in modo da evitare doppie definizioni.

D. Librerie

Durante la fase di compilazione creiamo dei file oggetto (.o) per ogni file in cui è scritta la descrizione delle funzioni di libreria (.c)

```
1 $ gcc -c bill.c
```

Si andrà poi a creare il prototipo della funzione (.h).

In fine tramite il linker si andranno ad unire tutti i file per crearne uno unico con tutte le definizioni delle funzioni incluse nelle librerie, di sistema e non, importate. Si vanno quindi a sciogliere tutti i riferimenti incrociati.

```
1 $ gcc -o program program.o bill.o
```

Per quanto riguarda le **funzioni di sistema** NON abbiamo il file sorgente ma abbiamo direttamente l'eseguibile. In compenso abbiamo un **file di libreria**, cioè un insieme di file oggetto linkati in un unico file, dove c'è il codice oggetto di tutte le funzioni.

Abbiamo **2 tipi di librerie**:

- **statiche**: è una **collezione di file oggetto** che hanno il codice compilato delle funzioni e che verranno **linkati al momento della compilazione**. Il programma che si crea sarà possibile essere eseguito solo sullo stesso OS.

Il **problema si ha nell'aggiornamento delle librerie al momento della scoperta di un bug**. Una volta corretto servirà ricevere la versione corretta per poter aggiornare il programma.

- **dynamic**: ricordano il concetto di plug-in, quindi **viene invocato a runtime e caricato in memoria** (es: aggiornamenti dei OS). **L'eseguibile non viene toccato la correzione avviene solo nella libreria**.

Il requisito maggiore è che chi si passa il codice debba avere lo stesso OS dell'altro utente. Notare che **non cambia il prototipo** dato che sennò bisognerà ricompilare l'intero programma.

In generale le **librerie statiche sono molto pericolose** infatti alcuni OS le aboliscono **per le questioni di sistema**. Su linux si ha come libreria statica 'lib.c' che è la libreria con le funzioni più usate in c. Per macos è stata abolita.

Per compilare con la versione dinamica non servono opzioni, per la statica si usa:

```
1 $ gcc -static
```

E. Creazione librerie

Per costruire una **libreria statica per MacOS**:

- 1) costruiamo il **file oggetto**:

```
1 $ gcc -c libprova.c
```

- 2) costruiamo la **libreria** (con ar=archive, c=create se lib.a non esiste):

```
1 $ ar rcs libprova.a libprova.o
```

- 3) costruire il **codice** che usa la libreria (con -Wall=verbose warning, -g=debugging, -c=create del file):

```
1 $ gcc -Wall -g -c useprova.c
```

- 4) **linker** per risolve le chiamate incrociate (con -L.=dove prendere la libreria, -l[nomelib]=usare la libreria):

```
1 $ gcc -g -o useprova useprova.o -L. -lprova
```

Per capire che libreria usa il codice si usa:

```
1 $ otool -L [nomecodice]
```

Per costruire una **libreria statica per Linux**:

- 1) costruiamo il **file oggetto**:

```
1 $ gcc -fPIC -Wall -g -c libprova.c
```

- 2) costruiamo la **libreria** (con 0.0=versione della libreria):

```
1 $ gcc -g -shared -Wl,-soname,libprova.so.0 -o  
libprova.so.0.0 libprova.o -lc
```

- 3) costruire il **link simbolico per aggiornare le librerie** senza aggiornare gli eseguibili e senza cambiare il nome del programma:

```
1 $ ln -sf libprova.so.0.0 libprova.so.0
```

- 4) **linker** per risolve le chiamate:

```
1 $ ln -sf libprova.so.0 libprova.so
```

Per capire che libreria usa il codice si usa:

```
1 $ ldd [nomecodice]
```

F. Aggiornamento librerie

Su **Linux** il sistema **andra' a prendere direttamente una libreria dinamica**, per evitare ciò e far trovare la nostra, basterà **impostare una variabile di ambiente**:

```
1 LD_LIBRARY_PATH='pwd' ldd useprova
```

Tipicamente la libreria viene distribuita nelle directory di sistema andandola ad "installare".

Su **MacOS** la libreria dinamica è un **dylib**:

```
1 $ gcc -dynamiclib libprova.c -o libprova.dylib
```

Quindi eseguendo il programma **trovera' la libreria controllando nella directory corrente** e quindi non serve creare la variabile di ambiente come su Linux.

i file di intestazione del mac come stdio.h per cercarla uso:

```
1 $ find /Applications/Xcode.app/ -name stdio.h 2>/dev  
/null
```

V. SYSTEM CALL

A. Funzioni e system call

Se prendiamo un funzionamento più semplice del comando "ls" potrebbe essere:

```
1 #include "apue.h"
2 #include <dirent.h>
3
4 int
5 main(int argc, char *argv[])
6 {
7     DIR                *dp;
8     struct dirent      *dirp;
9
10    if (argc != 2)
11        err_quit("usage: ls1 directory_name");
12
13    if ((dp = opendir(argv[1])) == NULL)
14        err_sys("can't open %s", argv[1]);
15    while ((dirp = readdir(dp)) != NULL)
16        printf("%s\n", dirp->d_name);
17
18    closedir(dp);
19    exit(0);
20 }
```

dove abbiamo che:

- **DIR**: struttura dati
- **struct dirent**: **tipo struttura** che contiene al suo interno diversi tipi di variabili.

Per capire se è una funzione di sistema lanciamo:

```
1 $ grep -rw "struct dirent" $INC
2
3 /Applications/Xcode.app/Contents/Developer/
  Platforms/MacOSX.platform/Developer/SDKs/
  MacOSX.sdk/usr/include//sys/dirent.h:struct
  dirent {
```

```
1 #ifndef _SYS_DIRENT_H
2 #define _SYS_DIRENT_H
3
4 #include <sys/_types.h>
5 #include <sys/cdefs.h>
6
7 #include <sys/_types/_ino_t.h>
8
9
10 #define __DARWIN_MAXNAMLEN      255
11
12 #pragma pack(4)
13
14 #if !__DARWIN_64_BIT_INO_T
15 struct dirent {
16     ino_t d_ino;                /* file
17     number of entry */
18     __uint16_t d_reclen;        /* length of
19     this record */
20     __uint8_t d_type;           /* file type
21     , see below */
22     __uint8_t d_namlen;        /* length of
23     string in d_name */
24     char d_name[__DARWIN_MAXNAMLEN + 1]; /*
25     name must be no longer than this */
26 };
27 #endif /* !__DARWIN_64_BIT_INO_T */
28
29 #pragma pack()
30
31 #define __DARWIN_MAXPATHLEN      1024
32
33 #define __DARWIN_STRUCT_DIRENTRY { \
```

```
29     __uint64_t d_ino;          /* file number of
30     entry */ \
31     __uint64_t d_seekoff;      /* seek offset (
32     optional, used by servers) */ \
33     __uint16_t d_reclen;       /* length of this
34     record */ \
35     __uint16_t d_namlen;       /* length of string
36     in d_name */ \
37     __uint8_t d_type;          /* file type, see
38     below */ \
39     char d_name[__DARWIN_MAXPATHLEN]; /*
40     entry name (up to MAXPATHLEN bytes) */ \
41 }
42
43 #if __DARWIN_64_BIT_INO_T
44 struct dirent __DARWIN_STRUCT_DIRENTRY;
45 #endif /* __DARWIN_64_BIT_INO_T */
46
47
48 #if !defined(_POSIX_C_SOURCE) || defined(
49     _DARWIN_C_SOURCE)
50 #define d_fileno      d_ino          /*
51     backward compatibility */
52 #define MAXNAMLEN      __DARWIN_MAXNAMLEN
53 /*
54 * File types
55 */
56 #define DT_UNKNOWN      0
57 #define DT_FIFO         1
58 #define DT_CHR          2
59 #define DT_DIR          4
60 #define DT_BLK          6
61 #define DT_REG          8
62 #define DT_LNK          10
63 #define DT_SOCK         12
64 #define DT_WHT          14
65
66 /*
67 * Convert between stat structure types and
68 * directory types.
69 */
70 #define IFTODT(mode)    (((mode) & 0170000) >>
71     12)
72 #define DTTOIF(dirtype) ((dirtype) << 12)
73 #endif
74
75 #endif /* _SYS_DIRENT_H */
```

dove vediamo che se la variabile "___DARWIN_64_BIT_INO_T" è stata definita avremo che la struttura di struct dirent è:

```
1 #define __DARWIN_STRUCT_DIRENTRY { \
2     __uint64_t d_ino;          /* file number of
3     entry */ \
4     __uint64_t d_seekoff;      /* seek offset (
5     optional, used by servers) */ \
6     __uint16_t d_reclen;       /* length of this
7     record */ \
8     __uint16_t d_namlen;       /* length of string
9     in d_name */ \
10    __uint8_t d_type;           /* file type, see
11    below */ \
12    char d_name[__DARWIN_MAXPATHLEN]; /*
13    entry name (up to MAXPATHLEN bytes) */ \
14 }
```

- **if**: esegue un controllo sugli args. Notiamo che "err_quit" non è una funzione di sistema da:

```
1 $ grep -rw "err_quit" $INC
```

infatti non restituisce nulla. Deve allora essere una funzione di libreria creata da noi quindi non presente nella directory standard.

La funzione andrà a dare un messaggio di errore e poi esce dal programma.

- **opendir**: serve ad aprire una directory andandola a caricare nella RAM.
- **while**: leggiamo la directory e la inseriamo nella struttura che poi sarà richiamata tramite:

```
1 dirp->d_name
```

dove "d_name" è il nome dello slot in cui è contenuto il nome del file.

- **exit**: restituisce l'exit code del programma

B. Capire se una funzione è una system call

Andiamo a **vedere se e' una funzione o una system call tramite "man"**, lo si capisce tramite la dicitura in alto alla pagina del manuale:

- **Library Functions Manual**
- **System Calls Manual**

Abbiamo anche **esempi piu' particolari**, come fork, dove è indicata come system call ma in realtà le richiama ma in prima persona.

Potremo trovare i simboli di una libreria tramite:

```
1 $ nm lib.a
```

che ci fa vedere, per ogni file oggetto, i simboli associati per ogni funzione.

Le system call le troveremo in "\$INC/sys/syscall.h"

C. Numeri dei file descriptor

Prendiamo un esempio semplificato del comando "cat":

```
1 #include "apue.h"
2
3 #define BUFSIZE 4096
4
5 int
6 main(void)
7 {
8     int    n;
9     char   buf[BUFSIZE];
10
11     while ((n = read(STDIN_FILENO, buf, BUFSIZE)) >
12            0)
13         if (write(STDOUT_FILENO, buf, n) != n)
14             err_sys("write error");
15
16     if (n < 0)
17         err_sys("read error");
18
19     exit(0);
20 }
```

ogni processo ha 3 file descriptor usati 0, 1, 2.

- **BUFSIZE**: macro di preprocessore
- **read**: **system call** con parametri:
 - **STDIN_FILENO**: file descriptor per dire da quale "numero di byte leggere" si vuole leggere. Cioè per leggere dal file indicato nello standard input

– **buf**: indirizzo dell'inizio dell'array

– **BUFSIZE**: quando deve leggere

Restituisce il numero di char che ha letto, dato che potrebbe leggere meno byte di quelli richiesti nel caso in cui il file ne contenga di meno. Ad ogni sua iterazione **si ricorda la "posizione nel file"** che gli permette di non leggere sempre i primi n byte ma di ricominciare da dove ha lasciato.

- **write**: richiede gli stessi valori di read tranne per **STDOUT_FILENO** e **ritorna il numero byte effettivamente letti**

per capire **quanto vale STDIN_FILENO**:

```
1 $ grep -rw "STDIN_FILENO" $INC
2
3 /Applications/Xcode.app/Contents/Developer/Platforms
4 /MacOSX.platform/Developer/SDKs/MacOSX.sdk/usr/
5 include/unistd.h:#define STDIN_FILENO 0
6 /* standard input file descriptor */
7
8 /Applications/Xcode.app/Contents/Developer/Platforms
9 /MacOSX.platform/Developer/SDKs/MacOSX.sdk/usr/
10 include/asl.h: * asl_log_descriptor(c, m,
11 ASL_LEVEL_NOTICE, STDIN_FILENO,
12 ASL_LOG_DESCRIPTOR_READ);
```

Sappiamo che un processo per eseguire un programma, esegue prima una **fork** e poi con **exec** esegue il programma. Prima di eseguire la fork il **child chiude il file 1** e quando fa una **open**, la system call prenderà il file nel quale reindirizzare lo STDOUT e restituirà il numero 1.

Su questo sistema si basa UNIX infatti avviene anche con le pipe "—". Permette di creare programmi complessi unendo tanti piccoli programmi specializzati in un'unica funzione.

È molto importante capire che **i child ereditano i file descriptor dei parent** quindi non è necessario che il programma corrente faccia una open dei file descriptor.

D. Meccanismi dei file

Un file è in insieme di meccanismi:

- **apri**
- **leggi**
- **scrivi**
- **chiudi**

Questi meccanismi sono applicabili a file, cartelle, stampanti ecc..., solo che per ogni "tipo" **i 4 meccanismi si adeguano** a ciò che il caso particolare deve fare.

E. Unbuffered I/O

Le system call rappresentano una barriera tra kernel e programmi, dove avremo rispettivamente **due diverse modalità di utilizzo**:

- **kernel mode**: ha tutti i privilegi
- **user mode**: non può accedere a tutte le celle di memoria

Per **ottimizzare la scrittura sulla memoria da parte del kernel** si utilizza la libreria **STDIO** che incrementa le prestazioni dato che gestisce il passaggio di pacchetti con il kernel in modo da inviare dei pacchetti consistenti ogni tot e non piccoli

pacchetti soni secondo. Per fare ciò usa un **buffered i/o** che, una volta riempiti dei buffer, gli manda al kernel.

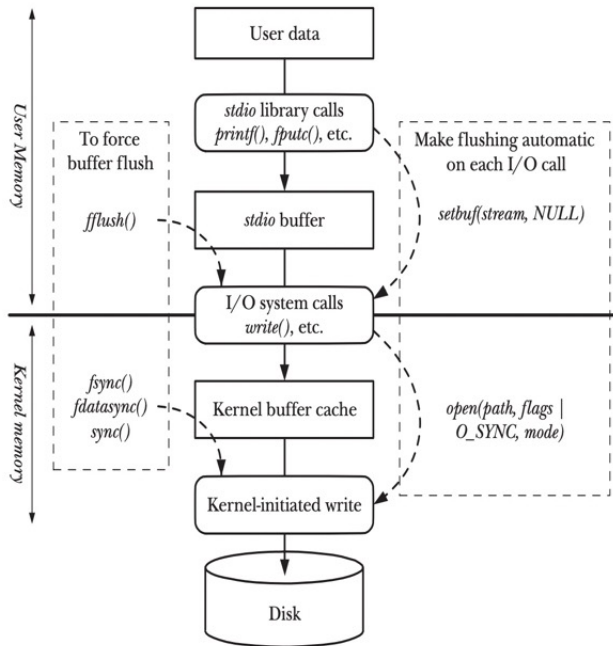


Figura 1. Schema unbuffered I/O

F. Fork & Exec

Prendiamo il codice di shell1.c che crea uno schell dal quale poter eseguire programmi:

```
1 #include "apue.h"
2 #include <sys/wait.h>
3
4 int
5 main(void)
6 {
7     char    buf[MAXLINE];    /* from apue.h */
8     pid_t   pid;
9     int     status;
10
11     printf("%% "); /* print prompt (printf requires
12                    %% to print %) */
13     while (fgets(buf, MAXLINE, stdin) != NULL) {
14         if (buf[strlen(buf) - 1] == '\n')
15             buf[strlen(buf) - 1] = 0; /* replace
16             newline with null */
17
18         if ((pid = fork()) < 0) {
19             err_sys("fork error");
20         } else if (pid == 0) { /* child */
21             execlp(buf, buf, (char *)0);
22             err_ret("couldn't execute: %s", buf);
23             exit(127);
24         }
25
26         /* parent */
27         if ((pid = waitpid(pid, &status, 0)) < 0)
28             err_sys("waitpid error");
29         printf("%% ");
30     }
31     exit(0);
32 }
```

avremo allora:

- **fgets**: funzione dello **stdout** che legge la stringa che dai prima di dare invio, quando una system call viene interrotta la fgets restituisce null facendo fermare il loop. Ha come argomenti:

- **buf**: buffer nel quale mettere la stringa
- **MAXLINE**: proviene da una nostra libreria

```
1 $ grep -rw "MAXLINE" include/
2 Binary file include//apue.h.gch matches
3 include//apue.h:#define MAXLINE 4096
   /* max line length */
```

- **stdin**: presente in stdio ed è una **struttura file** che **definisce uno standard input** tramite un puntatore ad un "file"

- **if 1**: permette di avere un null dove prima avevamo \n
- **if 2**: abbiamo una fork() che dopo che **viene invocata ritorna 2 volte**, questo perché andrà a creare 2 bash identici con memorie uguali nei contenuti ma indipendenti, l'unico cambiamento è il pid. fork() andrà quindi a restituire 0 nel child e il pid del child al parent tramite getpid().

Se pid < 0 vorrà dire che la fork è fallita. Se pid = 0 vorrà dire che siamo nel child.

Il che è molto importante dato che **il codice verrà eseguito sia dal child che dal parent**, e sarà contenuto nella memoria virtuale che hanno i programmi grande 2^{32} o 2^{64} in base all'OS.

Appena viene eseguita l'**exec**, lo spazio di memoria **viene azzerato** ma a discrezione del programma, vengono salvate alcune variabili di ambiente.

execlp: serve a far **eseguire un codice** (buf) del quale abbiamo il sorgente e l'eseguibile

if 3: serve a far andare avanti il parent.

waitpid: aspetta il child nel caso impieghi troppo tempo ad eseguire la sua azione, **tenendo appeso il prompt**. Con argomenti:

- **pid**: pid del child
- **&status**: reindirizza l'exit code del child, quando finisce, nella variabile status

Processi

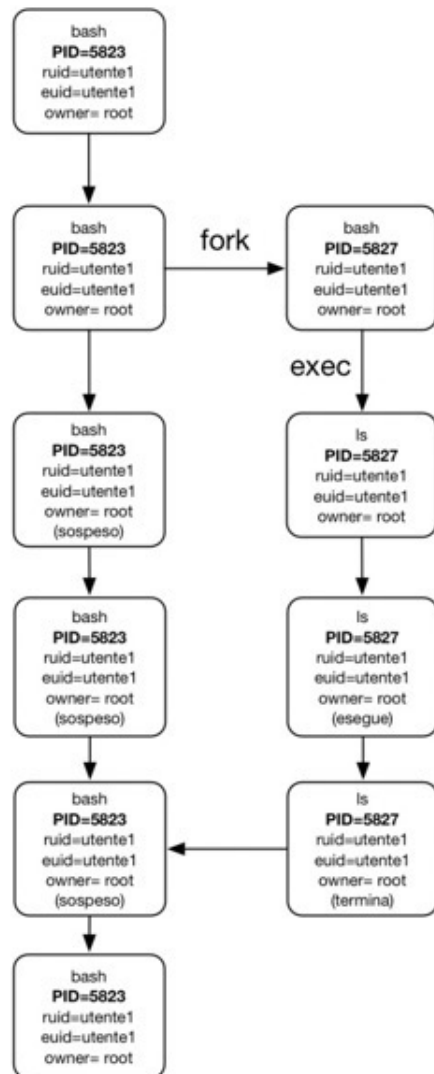


Figura 2. Esecuzione di fork ed exec

G. Thread

Sono dei **processi con lo stesso spazio di memoria del parent**. Per evitare che ognuno scriva dove vuole, **avviene una sincronizzazione tra le thread**. Questo metodo viene usato nelle macchine unicore per poter svolgere più operazione "contemporaneamente".

Tutto questo è **orchestrato dal kernel** che gestisce il **time shearing**.

H. Gestione degli errori

Per convenzione una funzione ritorna 0 se è andato tutto bene. Ci sono delle eccezioni, come la read, che ritorna il numero di byte letti.

Ogni valore possibile ritornato è specificato nel manuale:

```

1 $ man 2 intro
2 ...
3 1 EPERM Operation not permitted. An attempt was made
  to perform an operation limited to processes
  with appropriate privileges or to the owner of a
  file or other resources.
4
5 2 ENOENT No such file or directory. A component of a
  specified pathname did not exist, or the
  pathname was an empty string.
6 ...
  
```

Per le **system call**, quando avviene un errore, si **avvalora la variabile "errno"** che può essere consultata in un programma con

```
1 extern int errno;
```

Con "errno" bisogna tenere in conto che:

- 1) **non viene svuotata quando passiamo l'errore**. Quindi per sapere quando è stato dato un errore bisogna andare a consultarla quando la system call viene invocata
- 2) **vale 0** se non usata

Per la gestione degli errori useremo:

- **strerror**: restituisce la **stringa del valore** di errno
- **perror**: legge errno e stampa un messaggio a piacere

```

1 #include "apue.h"
2 #include <errno.h>
3
4 int
5 main(int argc, char *argv[])
6 {
7     fprintf(stderr, "EACCES: %s\n", strerror(EACCES)
8 );
9     errno = ENOENT;
10    perror(argv[0]);
11    exit(0);
  
```

dove:

- **fprintf**: stampa un errore allo standard specificato

I. Keyword in C ed accessi a variabili

In C le variabili con la keyword **"const"** serve a dare l'accesso ad una variabile ma in **sola lettura**. Ciò che la frena è il compilatore.

Se una **funzione vuole lavorare anche al suo esterno** usiamo:

- 1) attraverso il **passaggio di variabili dai parametri**
- 2) vengono passati gli **indirizzi delle variabili** abilitandone la scrittura
- 3) tramite **variabili globali**

```
1 const int *x
```

Un'altra keyword è **"restrict"** che serve a **non far creare copie di una variabile**, questo per fare in modo che il compilatore non si possa confondere con le copie di quella variabile.

```
1 int *restrict x
```


J. Segnal e interrupt

Guardiamo il file shell2.c :

```
1 #include "apue.h"
2 #include <sys/wait.h>
3
4 static void sig_int(int);      /* our signal-
   catching function */
5
6 int
7 main(void)
8 {
9     char    buf[MAXLINE];      /* from apue.h */
10    pid_t    pid;
11    int      status;
12
13    if (signal(SIGINT, sig_int) == SIG_ERR)
14        err_sys("signal error");
15
16    printf("%% "); /* print prompt (printf requires
   %% to print %) */
17    while (fgets(buf, MAXLINE, stdin) != NULL) {
18        if (buf[strlen(buf) - 1] == '\n')
19            buf[strlen(buf) - 1] = 0; /* replace
   newline with null */
20
21        if ((pid = fork()) < 0) {
22            err_sys("fork error");
23        } else if (pid == 0) { /* child */
24            execlp(buf, buf, (char *)0);
25            err_ret("couldn't execute: %s", buf);
26            exit(127);
27        }
28
29        /* parent */
30        if ((pid = waitpid(pid, &status, 0)) < 0)
31            err_sys("waitpid error");
32        printf("%% ");
33    }
34    exit(0);
35 }
36
37 void
38 sig_int(int signo)
39 {
40     printf("interrupt\n%% ");
41 }
```

dove rispetto a shell1 abbiamo come differenze:

- dichiarazione di funzione per **gestire un segnale**
- if 1: gestisce un segnale di errore
 - **signal**: gli diciamo che se arriva il segnale SIGINT allora eseguire la nostra funzione sig_int. In caso contrario, viene restituito SIG_ERR e la variabile globale errno viene impostata per indicare l'errore.
 - **sig_int**: avrà in ingresso un numero intero, che rappresenta il segnale, dato dal kernel.
- definizione della funzione dove viene gestito tramite una stampa

I **processi si interfacciano con il modo esterno tramite** delle funzionalità dette **Inter Process Communication (IPC)** alcune di queste sono i **segnali**.

I segnali **fanno parte dei software interrupt** dove il kernel:

- 1) **interrompe l'esecuzione** di un processo
- 2) **esegue il codice** definito per quell'interrupt
- 3) se non ci sono danni al processo questo **riprende da dove era stato interrotto**

Per quanto riguarda le **hardware interrupt** intendiamo i segnali di interrupt, dati dalle periferiche, che arrivano al processore. **L'interrupt e' un numero identificativo** che fa capire la natura di quell'interrupt tramite una **tabella degli interrupt** dove ad ogni numero corrisponde l'**indirizzo ad una routine**.

Il kernel invia tutti i segnali che saranno **causati da condizioni particolari o accessi a memoria non autorizzati** ma alle quali ha accesso (es: malloc che richiede l'accesso a memoria).

Come programmatori bisognerà **occuparsi di gestire i segnali** tramite:

- 1) azione di **default**
- 2) **ignorare** il segnale
- 3) gestione del segnale **scrivendo il signal endler**

Tramite il manuale di "signal" possiamo vedere tutti i tipi di segnale che esistono. Come si può notare alcuni hanno le diciture:

- **terminate process**: termina il processo
- **create core image**: effettua una fotocopia del core prima di terminare il processo in modo da poter effettuare un debug

Potremo **inviare i segnali tramite il comando "kill"**. Per esempio:

- SIGINT: ctrl C
- SIGQUIT: ctrl \
- SIGSTOP: Ctrl Z

K. Valori del tempo

Il sistema gestisce il tempo in secondi a partire dalla **Epoch 01/01/1970** e per ogni processo dà:

- **clock time**: **tempo effettivo di esecuzione** del processo da quando è nato a quando è terminato
- **user clock time**: tempo che **non richiede l'intervento del kernel** (speso dalla CPU)
- **system clock time**: tempo che **richiede l'intervento del kernel**

dove **il clock time non è la somma di user e system** dato che non si contano i processi che intervengono nel mezzo. La loro somma può essere maggiore del clock time se abbiamo un processore multicore dato che somma il tempo dai diversi core (il tempo potrebbe essere diverso in base al core).

```
1 $ time [nome programma]
2
3 ...
4 real    0m0.006s
5 user    0m0.001s
6 sys     0m0.005s
```

VI. GLI STANDARD

A. Storia e basi

La **standardizzazione di UNIX** e' iniziata nel 1988 facendo affidamento ad alcuni **standard di C** dato che fa usi di interfacce e prototipi.

In definitiva abbiamo gli standard:

- Posix.1-2001 / SUSv3: (<http://pubs.opengroup.org/onlinepubs/009604599/>)
- Posix.1-2008 / SUSv4: più usato in ambiti di automazioni aziendali, infatti sono specializzate sullo scambio di informazione in segnali realtime. Per questo la sua certificazione non è stata presa da nessuno se non fa un IBM. (<http://pubs.opengroup.org/onlinepubs/9699919799/>)

(PS: le versioni sono back compatibili quindi se settiamo `-D_XOPEN_SOURCE=700` non precludiamo la SUSv3)

Nonostante gli standard **ogni OS fa delle sue modifiche su alcune cose esterne alle SUS**.

Per verificare il tipo di standard su un applicativo (`_XOPEN_SOURCE`) o un sistema (`_XOPEN_VERSION`), si fa affidamento alle **"feature test macros"** consultabili dai **codici di intestazione .h**. Per esempio `_XOPEN_SOURCE` impostata a 600 o 700 indica SUSv3 o SUSv4.

```
1 -D_XOPEN_SOURCE=600
```

in questo modo potremo allora andare a compilare tutti i programmi conformi su qualsiasi OS.

B. Limiti

Abbiamo dei **limiti di compilazione** che possono essere visti nei file di intestazione.

Possiamo visualizzare i runtime limit che tramite le funzioni: `sysconf` (es: lunghezza massima del nome dei file che dipende dal filesystem può capirlo tramite `pathconf` su un file qualunque di quel filesystem)

- **sysconf**: usato per **determinare il valore corrente di un limite**
- **pathconf**: da **informazioni sul file system** e per fare ciò gli serve poter arrivare ad un qualunque file del filesystem
- **fpathconf**: come `pathconf` ma **prende anche il file descriptor**

Tutti e 3 prendono come **parametro**:

```
1 int name
```

che **restituisce una chiave** in base a cosa si vuole indagare. In pratica fanno riferimento ad un nome simbolico che si riferisce ad un valore. Tutte queste chiamate fanno sì di avere più **portabilità**. Se queste chiamate sono fatte da file include **vincono sempre quelli di sysconf**. Saranno precedute da `"_SC_"` per i `sysconf` e da `"_PC_"` per i `pathconf`.

C. Determinare l'allocazione

Supponendo di avere **bisogno di uno spazio** dove mettere un nome di file (path) per poterlo gestire. Tramite la funzione `"path_alloc"` ci faremo **restituire un puntatore ad una memoria capace di contenere il massimo dei caratteri** gestibili dal sistema (es: $2^{32} \rightarrow 32$ bit, $2^{64} \rightarrow 64$ bit).

Per vedere qual è la lunghezza usiamo la variabile limite:

```
1 pathconf(_PC_NAME_MAX)
```

in genere avremo `NAME_MAX = 255`.

Lo stesso lavoro di riferimenti simbolici uno dopo l'altro è `"pid_t"`, per poter **lasciare piu' liberta' al programmatore**:

```
1 $ grep -rw "pid_t" $INC | grep typedef
2
3 /sys/_types/_pid_t.h:typedef __darwin_pid_t pid_t;
4 /sys/_types.h:typedef __uint32_t __darwin_id_t;
5
6 $ grep -rw "__uint32_t" $INC | grep typedef
7
8 /i386/_types.h:typedef unsigned int __uint32_t;
```

tutti questi rimandi sono dati dalla **portabilità**.

VII. FILE I/O

A. Chiamata open

I file I/O sono le **funzioni che gestisce buffered I/O** ed in contrapposizione da quelle della libreria "stdlib". La chiamata open() fa parte di queste funzioni, i suoi argomenti sono:

- **arg**: path assoluto o relativo
- **flag**: bit che indica l'**attivazione di alcune modalita'**. Si avrà allora a **settare il bit della flag a 1**:
- **mode**: serve a dare i **privilegi con cui i file deve essere creato (da usare solo nella creazione del file)**

```
1 open(file, O_RDONLY | O_APPEND | O_CREAT | O_TRUNC  
    , file_mode)
```

avremo allora: 11000001010

con O_RDONLY: 2, O_APPEND: 8, O_CREAT: 512,
O_TRUNC: 1024

Appena **aperto il file questo avrà la "current position" a 0**, cioè ad inizio file. Iniziando a scrivere avremo che alla prima lettura la nostra current position si sarà spostata fino a dopo abbiamo finito di scrivere.

Per quanto riguarda **read e write i bit delle flag** abbiamo un modo "scomodo":

```
1 #define O_RDONLY      0x0000      /* open for  
    reading only */  
2 #define O_WRONLY      0x0001      /* open for  
    writing only */  
3 #define O_RDWR        0x0002      /* open for  
    reading and writing */  
4 #define O_ACCMODE      0x0003      /* mask for  
    above modes */
```

sono dette quindi **maschere** per leggere o scrivere.

Una variante è **openat()** che prende il file descriptor (passato dalla open() sulla directory) di una directory per passare il path relativo a quella directory. Abbiamo quindi la **possibilità di accedere a delle directory anche senza avere i privilegi**. ???

il permesso di esecuzione delle directory serve a cercare un nome in una directory. se il permesso è disattivato non si riesce ad arrivare alla foglia, se invece apri una directory il sistema sa già a quale inod andare dato che si salva quello al posto di tutto il path. quindi fai diventare relativi i privilegi per quella directory.

questa chiamata è interessante per quando le thread vogliono lavorare in un loro ambiente.

alla openat è associata una flag alla directory aperta dalla open, allora usiamo la flag O_DIRECTORY. nelle flag vediamo la differenza tra UNIX e gli standard dato che ce ne sono un certo numero di base e poi ogni UNIX ne aggiunge altri. Gli standard gli troviamo nel manuale del SUSv3 e SUSv4.

O_CREAT: flag per dire che si vuole creare un file. una trappola che c'è è la builtin "umask" usata per dare dei permessi di base. questa umask è presente in ogni processo e la eredita dal padre (ma può comunque modificarla). ci restituisce un numero ottale (perché inizia con 0) e dice di togliere la scrittura ai gruppi

```
1 $ umask  
2 0022  
3  
4 $ ll file  
5 -rw-r--r-- 1 docente staff 0 14 Ott 09:02 file
```

quindi la umask taglia i permessi dei file. quindi se diamo un 666 nella creazione di un file con open ci ritroveremo 644.

altra trappola: se creo un file in sola scrittura e lettura ecc, potremo usare una chiamata per togliere dei flag in corsa.

un'altra trappola: la cosa strana è che se apri un file in rw e ottieni i privilegi, se i privilegi vengono cambiati dalla linea di comando togliendo w se vado a scrivere in quel file perché aperto con la open in precedenza, quel file comunque puoi continuare a scriverlo dato che il permesso non viene revocato

O_TRUNC: nella creazione del file se il file esiste ed è aperto per rw viene azzerato il file (se vuoi aprire un file nuovo ed esiste già il vecchio viene azzerato)

O_EXCL: nella creazione del file se in file già esiste fa fallire la chiamata. quando si fa una read potrebbe essere appesa per giorni, ma se il file viene aperto con questo file, se non c'è nulla da leggere o scrivere ritorna e non lo lascia appeso

chiamata lseek: in un dato momento se hai un file descriptor sta in un certo punto del file (se non lo hai toccato stai all'inizio). lseek sui file già aperti permette di spostare la "current position"

1 arg: file descriptor 2. offset per dire dove ci si vuole spostare 3. whence: ("da dove") può essere: SEEK_SET valore preciso di dove andare. SEEK_CUR può essere negativo ed indica il gap dopo il quale andare cioè di quanto spostarsi, SEEK_END valore rispetto alla fine del file sarà negativo ma se metto valore positivo crei un nuovo dato che salta alcuni blocchi. nell'inod viene indicato il file buchi compresi, l'inod ricodiamo che ha 256 byte dove un po' sono usati per i metadati e nel resto abbiamo dei puntatori a dei blocchi che sono a loro volta blocchi di puntatori ecc quindi non c'è il concetto di buco coi blocchi perché semplicemente ci sono un po' di blocchi da una parte e un po' dall'altra

Un esempio di buco lo creiamo:

```
1 #include "apue.h"  
2 #include <fcntl.h>  
3  
4 char    buf1[] = "abcdefghij";  
5 char    buf2[] = "ABCDEFGHIJ";  
6  
7 int  
8 main(void)  
9 {  
10     int    fd;  
11  
12     if ((fd = creat("file.hole", FILE_MODE)) < 0)  
13         err_sys("creat error");  
14  
15     if (write(fd, buf1, 10) != 10)  
16         err_sys("buf1 write error");  
17     /* offset now = 10 */  
18  
19     if (lseek(fd, 16384, SEEK_SET) == -1)  
20         err_sys("lseek error");  
21     /* offset now = 16384 */  
22  
23     if (write(fd, buf2, 10) != 10)  
24         err_sys("buf2 write error");  
25     /* offset now = 16394 */  
26  
27     exit(0);  
28 }
```

avremo:

```
1 $ xxd file.hole
2 00000000: 6162 6364 6566 6768 696a 0000 0000 0000
   abcdefghij.....
3 00000010: 0000 0000 0000 0000 0000 0000 0000 0000
   .....
4 00000020: 0000 0000 0000 0000 0000 0000 0000 0000
   .....
5 00000030: 0000 0000 0000 0000 0000 0000 0000 0000
   .....
6 00000040: 0000 0000 0000 0000 0000 0000 0000 0000
   .....
7 00000050: 0000 0000 0000 0000 0000 0000 0000 0000
   .....
8 00000060: 0000 0000 0000 0000 0000 0000 0000 0000
   .....
9 00000070: 4142 4344 4546 4748 494a
   ABCDEFGHIJ
```

abbiamo che la memoria effettiva sul disco è:

```
1 $ du file.hole
2 3208    file.hole
```

invece la size del file viene dichiarata:

```
1 $ stat -x file.hole
2  File: "file.hole"
3  Size: 1638410      FileType: Regular File
4  Mode: (0644/-rw-r--r--)  Uid: ( 501/
   matt) Gid: ( 20/  staff)
5 Device: 1,16      Inode: 27657406  Links: 1
6 Access: Fri Oct 14 09:59:17 2022
7 Modify: Fri Oct 14 09:57:59 2022
8 Change: Fri Oct 14 09:57:59 2022
9 Birth: Fri Oct 14 09:47:24 2022
```

esistono dei file descriptor seekble e non che cerca di capire se è un file regolare o meno:

```
1 #include "apue.h"
2
3 int
4 main(void)
5 {
6     if (lseek(STDIN_FILENO, 0, SEEK_CUR) == -1)
7         printf("cannot seek\n");
8     else
9         printf("seek OK\n");
10    exit(0);
11 }
```

ci sono più casi in cui il numero di byte letto è minore di quello chiesto: 1. possiamo avere un valore di ritorno di 50 se chiedo di leggere 100 perchè ci sono solo 50 byte 2. quando legge da un terminale: la read ritorna quando dai invio 3. quando legge da una rete: puoi leggere 100 byte ma se ne leggi 10 puoi gestire tu cosa fare, in genere se non arriva nulla la read rimane in attesa e non gli arrivano byte rimane appesa se non arrivano rimane appesa e poi interrompe quando non sente più nulla 4. quando legge da una pipe: se non arriva nulla rimane appesa, se arriva qualcosa interrompe e restituisce quello che ha letto 5. quando interrotta da un segnale e alcuni dati sono stati letti gli restituisce