Per chi si avvale di un sistema operativo GNU, gli strumenti per sviluppare codice per x86-16 sono costituiti dalla raccolta nota con il nome Dev86, la quale mette assieme il compilatore C Bcc,[1] l'assemblatore As86 e il «collegatore» Ld86, oltre a una libreria C adatta per produrre applicazioni per ELKS (*Embeddable Linux kernel subset*).

Considerato che strumenti del genere sono utili, presumibilmente, per realizzare un programma autonomo (*stand alone*) o il kernel di un sistema operativo, un programma di avvio facilita molto il lavoro e consente di concentrare l'attenzione su ciò che si vuole realizzare veramente. Per questo motivo, nel capitolo viene anche preso in considerazione Bootblocks per l'avvio di un sistema operativo da dischetti con file system Minix 1.

Preparazione

In una distribuzione GNU/Linux Debian sono disponibili i pacchetti '**bcc**', '**bin86**' e '**elks-libc**' che forniscono il necessario per la compilazione, ma in un altro sistema GNU può essere necessario procurarsi il pacchetto sorgente Dev86, dal quale si ottiene ciò che serve.

Se si è costretti a partire dai sorgenti di Dev86, una volta scaricato il pacchetto, questo può essere espanso in una directory qualunque nell'elaboratore GNU, come mostrato dall'esempio seguente, dove però, successivamente, si possano acquisire i privilegi dell'utente '**root**':

```
$ tar xzvf Dev86src-0.16.17.tar.gz [Invio]
```

Si ottiene la directory '`dev86-0.16.17/`' che si articola ulteriormente. Terminata l'installazione occorre compilare questi sorgenti e installarli. In questo caso si prevede di installare Dev86 a partire da '`/opt/dev86/`':

```
$ cd dev86-0.16.17 [Invio]

$ make PREFIX=/opt/dev86/ [Invio]
```

Viene richiesto di intervenire su alcuni indicatori (*flag*); in generale dovrebbe andare bene ciò che viene proposto in modo predefinito:

```
  1) (ON)    Library of bcc helper functions
  2) (ON)    Minimal syscalls for BIOS level
  3) (ON)    Unix error functions
  4) (ON)    Management for /etc/passwd /etc/group /etc/utmp
  5) (OFF)   Linux-i386 system call routines GCC
  6) (ON)    GNU termcap routines
  7) (ON)    Bcc 386 floating point
  8) (ON)    Linux-i386 system call routines
  9) (ON)    Example kernel include files and syscall.dat
 10) (ON)    Malloc routines
 11) (ON)    Various unix lib functions
 12) (ON)    Msdos system calls
 13) (ON)    Regular expression lib
 14) (ON)    Stdio package
 15) (ON)    String and memory manipulation
 16) (ON)    Linux-8086 system call routines
 17) (ON)    Termios functions
 18) (ON)    Unix time manipulation functions.

Select config option to flip [or quit] > quit [Invio]
```

Al termine della compilazione si passa all'installazione, cominciando dalla creazione della directory '`/opt/dev86/`'. Per fare questo

occorrono i privilegi dell'utente '**root**':

```
$ su [Invio]

...

# mkdir -p /opt/dev86 [Invio]

# make install [Invio]
```

Bcc

Bcc[2] è un compilatore C tradizionale, ovvero fatto per la vecchia sintassi, nota con la sigla K&R. Tuttavia, con l'ausilio di un programma esterno (di norma si tratta di Unproto di Wietse Venema, incluso nella distribuzione Dev86), può compilare sorgenti scritti nella forma di un C standard, pur non potendo disporre di tutte le funzionalità di un compilatore aggiornato.

```
bcc [opzioni] file_c...
```

Tabella u141.2. Alcune opzioni per l'uso di Bcc.

Opzione	Descrizione
`-ansi`	Si avvale di un programma esterno per poter accettare un sorgente scritto secondo le convenzioni attuali del linguaggio C, pur nei limiti di quanto Bcc può poi elaborare.
`-0`	Produce un codice adatto per CPU 8086/8088.
`-S`	Produce un file in linguaggio assemblatore, da usare poi con l'assemblatore As86. In mancanza dell'opzione '**-S**' o '**-c**', si ottiene direttamente un file eseguibile, con l'intervento automatico di As86 e di Ld86.
`-c`	Produce un file oggetto, da utilizzare poi con il collegatore Ld86. In mancanza dell'opzione '**-S**' o '**-c**', si ottiene direttamente un file eseguibile, con l'intervento automatico di As86 e di Ld86.
`-o` ***nome***	Produce un file con il nome specificato.
`-I`	Non utilizza i percorsi predefiniti per l'inclusione dei file di intestazione.
`-I`***percorso***	L'opzione '**-I**', a cui si attacca un percorso, aggiunge quel percorso a quelli usati per l'inclusione dei file di intestazione. Possono essere specificati più percorsi ripetendo l'uso dell'opzione.

L'esempio seguente mostra la compilazione del file '`mio.c`', per produrre il file '`mio.s`', contenente il codice in linguaggio assemblatore. Per la compilazione, i file di intestazione vengono cercati esclusivamente in percorsi stabiliti: '`../include`' e '`../../include`'.

```
$ bcc -ansi -0 -S -o mio.s ↩
↪     -I -I../include -I../../include mio.c [Invio]
```

L'esempio successivo è simile, ma si produce il file oggetto '`mio.o`':

```
$ bcc -ansi -0 -c -o mio.s ↩
↪     -I -I../include -I../../include mio.c [Invio]
```

I nomi delle variabili e delle funzioni scritte in linguaggio C, si traducono nel linguaggio assemblatore in nomi preceduti dal trattino basso. Per esempio, la funzione ***main()***, diventa il simbolo '**_main**'. Per questa ragione, quando si scrivono porzioni di codice in linguaggio assemblatore da esportare, occorre ricordare di aggiungere un trattino basso all'inizio.

A meno di voler produrre programmi per il sistema operativo ELKS, il compilatore Bcc va utilizzato con l'opzione '**-c**', oppure '**-S**', per poter controllare i passaggi successivi, in particolare la fase di collegamento dei vari componenti.

As86

As86 è un assemblatore in linguaggio x86, adatto alla compilazione di quanto prodotto da Bcc. La sintassi usata da As86 è fondamentalmente quella Intel.

```
as86 [opzioni] file_s
```

Tabella u141.3. Alcune opzioni per l'uso di As86.

Opzione	Descrizione
`-0`	Compila nella modalità a 16 bit e avvisa quando incontra istruzioni che non sono adatte a CPU 8086/8088.
`-o` ***nome***	Produce un file oggetto con il nome specificato.
`-s` ***nome***	Produce un file di testo contenente l'elenco dei simboli individuati. Questa opzione può essere usata assieme a '**-o**'.
`-u`	Fa in modo che i simboli privi di una dichiarazione siano importati dall'esterno senza specificare il segmento.

L'esempio seguente rappresenta una situazione di utilizzo comune, in cui si produce il file oggetto '`mio.o`', a partire dal sorgente '`mio.s`':

```
$ as86 -u -0 -o mio.o mio.s [Invio]
```

Come precisato a proposito di Bcc, se si devono importare dei simboli dal codice C, occorre aggiungere un trattino basso all'inizio dei nomi.

Ld86

Ld86 è il «collegatore» (*linker*) associato a As86. La caratteristica di Ld86 è quella di poter produrre un eseguibile «impuro» (come viene definito nella sua pagina di manuale), per il quale il segmento usato dal codice è lo stesso usato per i dati (***CS***==***DS***==***SS***), oppure può tenere separati il codice e i dati in segmenti distinti (ma in tal caso vale ancora l'uguaglianza ***DS***==***SS***).

Nella pagina di manuale di Ld86 si parla di «I&D», ovvero di istruzioni e dati, che possono essere separati o meno.

```
ld86 [opzioni] file_o...
```

Tabella u141.4. Alcune opzioni per l'uso di Ld86.

Opzione	Descrizione
`-d`	Elimina l'intestazione dal file che va a essere generato. L'utilizzo di questa opzione implica l'uso di '**-s**'.
`-s`	Elimina i simboli.
`-o` ***nome***	Produce un file eseguibile con il nome specificato.
`-i`	Tiene separati il segmento usato dal codice rispetto a quello dei dati. In mancanza di questa opzione, il segmento è lo stesso.

L'esempio seguente mostra la creazione di un programma, privo di intestazione e di simboli, dove tutto viene così definito attraverso il codice in modo esplicito. In particolare, si presume che il file '`crt0.o`' sia realizzato in modo da definire esattamente la forma della prima parte del file eseguibile.

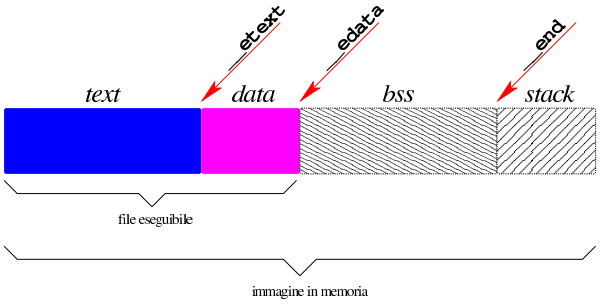
```
$ ld86 -d -s -o kimage crt0.o body.o main.o [Invio]
```

Ld86 definisce implicitamente dei simboli, raggiungibili dal codice che si scrive. Tra questi, sono molto importanti quelli seguenti, con cui è possibile determinare la collocazione in memoria del programma, distinguendo tra codice e dati:

Simbolo per As86	Nome per Bcc	Descrizione
`__etext`	`_etext`	Variabile a 16 bit contenente l'indirizzo conclusivo dell'area usata dal codice, nell'ambito del segmento in cui si colloca.
`__edata`	`_edata`	Variabile a 16 bit contenente l'indirizzo conclusivo dell'area usata dai dati inizializzati, nell'ambito del segmento in cui si colloca.
`__end`	`_end`	Variabile a 16 bit contenente l'indirizzo conclusivo dell'area usata dai dati non inizializzati (BSS), esclusa la pila, nell'ambito del segmento in cui si colloca.
`__segoff`	`_segoff`	Variabile a 16 bit contenente la distanza tra l'inizio del segmento codice e l'inizio di quello usato per i dati, espressa in multipli di 16 bit. In presenza di eseguibili in cui il segmento è lo stesso, questo valore è zero.

In generale, la struttura di un file eseguibile prodotto da Ld86 è composta inizialmente dal codice, quindi continua con i dati inizializzati. Lo spazio dei dati non inizializzati non fa parte del file e deve essere previsto quando si carica il file in memoria, per metterlo in esecuzione; inoltre, lo stesso va fatto per la pila (*stack*) dei dati, la quale deve collocarsi dopo tale area.

Figura u141.6. Contenuto di un programma privo di intestazione e di simboli.



Bootblocks

Il pacchetto Bootblocks[3] consente di avviare un sistema per elaboratori x86-16, contenuto in un dischetto o in una partizione del disco fisso, con il kernel inserito nello stesso file system. Il pacchetto viene distribuito assieme agli strumenti di sviluppo Dev86, ma non viene compilato automaticamente assieme a quelli. Si trova precisamente nella sottodirectory '`bootblocks/`' dei sorgenti di Dev86. Si compila in modo molto semplice con il comando '**`make`**':

```
# cd sorgenti_dev86/bootblocks [Invio]

# make [Invio]
```

Dalla compilazione si ottengono diversi file e sono utili in particolare:

File	Descrizione
'**`makeboot`**'	programma per l'installazione del settore di avvio, da usare attraverso un sistema GNU/Linux comune;
'**`makeboot.com`**'	programma analogo a '**`makeboot`**', da usare con un sistema Dos.

Per avviare un programma autonomo o un kernel vero e proprio, in un dischetto con file system Minix 1, è sufficiente copiare tale file in modo che si trovi nella directory principale con il nome '`boot`', oppure si crea la directory '`/boot/`' e vi si colloca il file con il nome che si preferisce.

Supponendo di utilizzare un sistema GNU/Linux, supponendo di

avere preparato il dischetto Minix (con i nomi al massimo di 14 byte) contenente tutto quello che serve, soprattutto con la directory '`/boot/`' o con il file '`boot`', se questo dischetto risulta inserito nell'unità corrispondente al file di dispositivo '`/dev/fd0`', **senza essere stato innestato**, si può eseguire il comando seguente, tenendo conto che il programma '**`makeboot`**' si presume collocato in una directory prevista tra i vari percorsi della variabile di ambiente '**`PATH`**':

```
# makeboot minix /dev/fd0 [Invio]

  Wrote sector 0
  Wrote sector 1
```

Se il programma si accorge che il settore di avvio del dischetto contiene già qualcosa, si rifiuta di procedere, a meno di usare l'opzione '**`-f`**':

```
# makeboot -f minix /dev/fd0 [Invio]

  Boot block isn't empty, zap it first
  Wrote sector 0
  Wrote sector 1
```

È importante sapere che questo programma di avvio colloca in memoria il programma o il kernel da avviare a partire dall'indirizzo efficace 10000_{16}. Pertanto, dopo l'avvio effettivo, rimane inutilizzato lo spazio di memoria da 00500_{16} $0FFFF_{16}$.

Riferimenti

- Andrew S. Tanenbaum, *Operating Systems: Design and Implementation*, **prima edizione**, 1987, Prentice-Hall, ISBN 0-13-637406-9
  Appendice B: *introduction to the IBM PC*
- MAD, *Assembly tutorial*
  http://www.xs4all.nl/~smit/asm01001.htm
- Wikipedia, *x86 instruction listings*
  http://en.wikipedia.org/wiki/X86_instruction_listings
- *The x86 Interrupt List, aka "Ralf Brown's Interrupt List", "RBIL"*
  http://www.cs.cmu.edu/~ralf/files.html
- *Computer interrupt*
  http://wayback.archive.org/web/20040101000000/http://calab.kaist.ac.kr/~hyoon/courses/cs310_2001fa01ll/micro17.ppt*
  http://www.ece.msstate.edu/~reese/EE3724/lectures/interrupt/interrupt.pdf
- BiosCentral, *BIOS data area*
  http://www.bioscentral.com/misc/bda.htm
- Robert de Bath, *Linux 8086 development environment*
  http://homepage.ntlworld.com/robert.debath/
  http://homepage.ntlworld.com/robert.debath/dev86/

[1] **Bcc, As86, Ld86** GNU GPL

[2] Va specificato che si tratta del compilatore Bcc di Bruce Evans, perché con questo nome o con questa sigla si trovano più facilmente riferimenti a compilatori C diversi, per esempio quello di Borland.

[3] **Bootblocks** GNU GPL