

# L'architettura Intel/AMD a 64 bit

G. Lettieri

2 Marzo 2023

## 1 Registri e istruzioni

I processori Intel/AMD a 64 bit sono una evoluzione dei precedenti processori a 32 bit, evoluzione a loro volta dei precedenti a 16 bit.

Lo *state* dell'elaboratore è dato dal contenuto dei registri del processore, dal contenuto della memoria e dal contenuto dei registri di I/O. Vediamoli di seguito.

All'interno del processore si trovano 16 registri di uso generale, i cui nomi sono mostrati in Figura 1 insieme al registro **rip**, che è l'Instruction Pointer, e al registro **rflags**, che è il registro dei flag. Tutti i registri sono grandi 64 bit. Ci sono anche altri registri di controllo, che vedremo successivamente.

È possibile anche riferire solo alcune parti di ogni registro, utilizzando un nome diverso, secondo la Tabella 1. In particolare, è possibile riferire i 32 bit meno significativi, i 16 bit meno significativi e gli 8 bit meno significativi. Per i soli registri **rax**, **rbx**, **rcx** ed **rdx** è possibile anche riferire gli 8 bit successivi a **al**, **bl**, **cl** e **dl**, utilizzando i nomi nell'ultima colonna della Tabella 1. Tuttavia ci sono delle complicate limitazioni sul loro utilizzo, e conviene ignorare l'esistenza di questi ulteriori sotto-registri.

Per quanto riguarda lo spazio di memoria, il processore può potenzialmente riferire  $2^{64}$  byte distinti utilizzando indirizzi di 64 bit. In pratica, però, quasi tutti i processori Intel/AMD ad oggi disponibili limitano i bit realmente utilizzabili a 48, ottenendo uno spazio di memoria di  $2^{48} \text{ B} = 256 \text{ TiB}^1$ , che è comunque molto grande. I più recenti offrono l'opzione di utilizzare 57 bit, corrispondenti ad uno spazio di memoria di  $2^{57} \text{ B} = 128 \text{ PiB}$ . I 48 (o 57) bit liberamente utilizzabili sono quelli meno significativi. Gli altri 16 (o 7) devono essere tutti uguali al bit numero 47 (o 56). Questo vuol dire che lo spazio di memoria si presenta come in Figura 2: si possono indirizzare soltanto due porzioni contigue, ciascuna grande  $2^{47}$  (caso con 48 bit, Fig. 2a) o  $2^{56}$  byte (caso con 57 bit, Fig. 2b), una all'inizio e l'altra alla fine dello spazio degli indirizzi. Gli indirizzi che rispettano questa regola sono detti in *forma canonica*. È il processore stesso a generare un errore se si tenta di utilizzare un indirizzo che non è in forma canonica.

Lo spazio di I/O, infine, è costituito da  $2^{16}$  locazioni di un byte (64 KiB).

---

<sup>1</sup>B sta per byte mentre Ki = 1024, Mi = 1024<sup>2</sup>, Gi = 1024<sup>3</sup>, Ti = 1024<sup>4</sup> e Pi = 1024<sup>5</sup>.

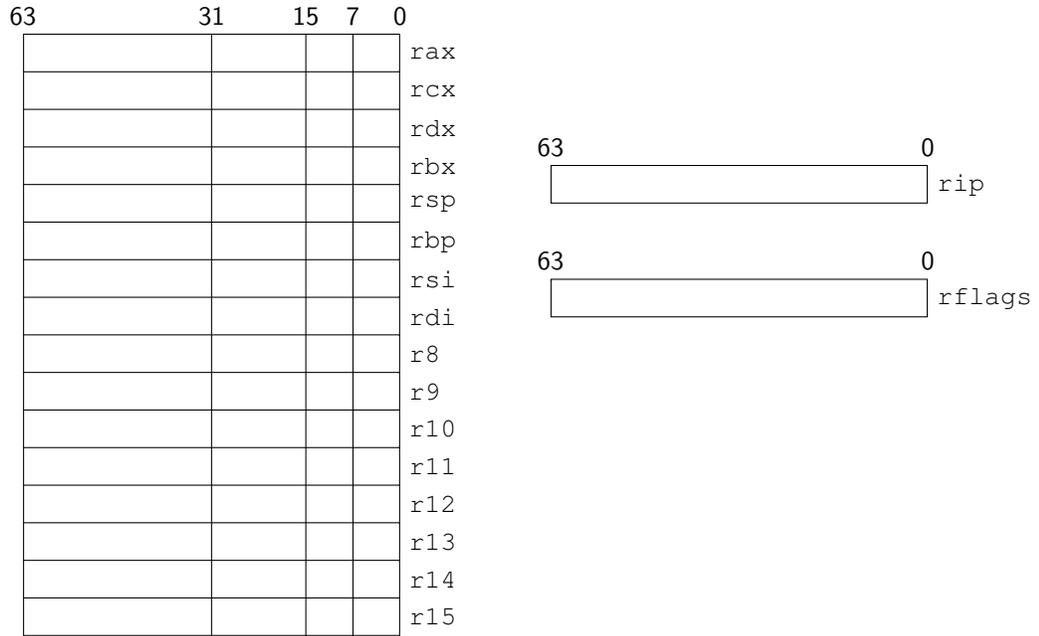


Figura 1: I registri del processore.

64b	32b	16b	8b	8b
rax	eax	ax	al	ah
rcx	ecx	cx	cl	ch
rdx	edx	dx	dl	dh
rbx	ebx	bx	bl	bh
rsp	esp	sp	spl	
rbp	ebp	bp	bpl	
rsi	esi	si	sil	
rdi	edi	di	dil	
r8	r8d	r8w	r8b	
r9	r9d	r9w	r9b	
r10	r10d	r10w	r10b	
r11	r11d	r11w	r11b	
r12	r12d	r12w	r12b	
r13	r13d	r13w	r13b	
r14	r14d	r14w	r14b	
r15	r15d	r15w	r15b	

Tabella 1: Nomi dei sotto-registri.

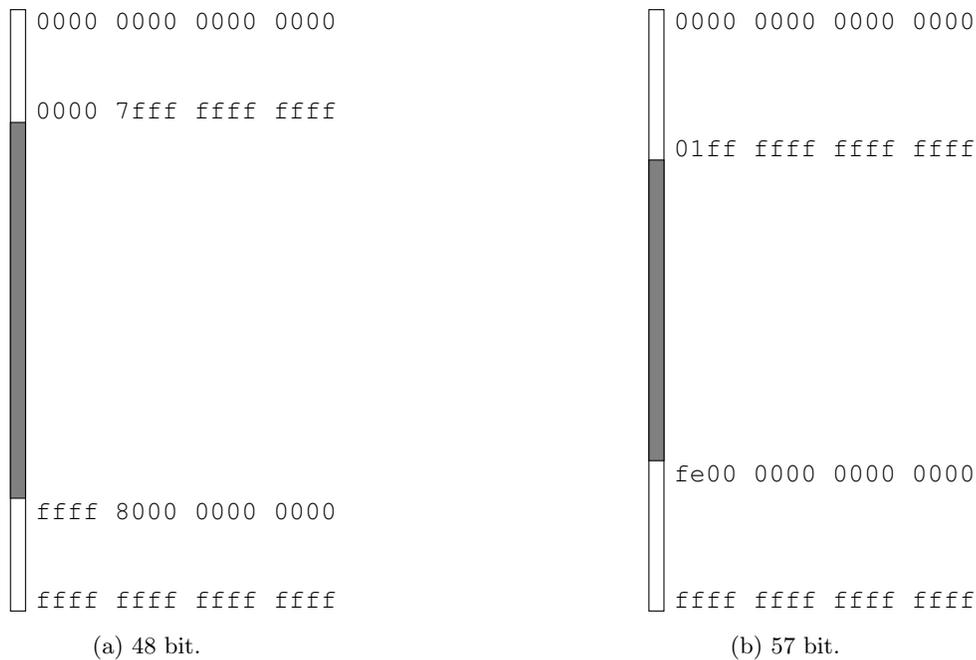


Figura 2: Spazio di indirizzamento di memoria (non in scala).

Le istruzioni riconosciute dal processore prevedono normalmente due operandi di ingresso e sovrascrivono il risultato sul secondo operando. La sintassi assembler è la seguente:

*codice-operativo primo-operando, secondo-operando*

Sono possibili tre modalità di indirizzamento degli operandi:

- **immediato:** l'operando è una costante contenuta nell'istruzione stessa (al massimo su 32 bit); in assembler l'operando deve essere preceduto dal carattere \$;
- **registro:** l'operando è contenuto in uno dei registri del processore; in assembler si usa il carattere % seguito dal nome del registro;
- **memoria:** l'operando è contenuto in memoria; l'istruzione deve dunque specificarne l'indirizzo.

Il caso di indirizzamento di memoria è il più complicato, in quanto è possibile chiedere al processore di *calcolare* l'indirizzo richiesto (ricordarsi che gli indirizzi sono numeri). Ci sono due modalità principali di calcolo dell'indirizzo, con vari sotto casi:

- La modalità con spiazamento, base, indice e scala, che usa la seguente sintassi:

*spiazamento (base, indice, scala)*

dove:

- *spiazamento* è una costante (8 o 32 bit, con segno)
- *base* è il nome di un registro a 64b;
- *indice* è il nome di un registro a 64b;
- *scala* può valere 1, 2, 4 oppure 8.

Il processore calcherà l'indirizzo sommando lo *spiazamento* (esteso con segno a 64b), il contenuto del registro *base* e il contenuto del registro *indice* moltiplicato per la *scala*.

Ci sono vari casi particolari, vediamo i più comuni:

- se lo spiazamento è zero, si può omettere;
- se la scala è 1, si può omettere insieme alla seconda virgola;
- si può omettere tutta la parte tra parentesi; in questo caso l'indirizzo coincide con lo spiazamento e l'indirizzamento è detto *diretto*;
- si possono omettere l'indice e la scala, con le relative virgole; in questo caso l'indirizzo è ottenuto sommando lo spiazamento (esteso con segno a 64b) e il contenuto del registro *base*; se si omette anche lo spiazamento, l'indirizzamento è detto *indiretto*.

- Relativa a **rip**, con la seguente sintassi:

*spiazamento (%rip)*

dove *spiazamento* è una costante (8 o 32 bit, con segno). Il processore calcherà l'indirizzo sommando *spiazamento* (esteso con segno a 64b) al contenuto di **rip**.

Si noti che il secondo operando non può essere di tipo immediato e che al più un operando può essere di tipo memoria.

Solo l'istruzione `movabs` ammette operandi immediati di 64b e spiazamenti di 64b, e solo nei seguenti formati:

- `movabs $costante, %registro`: copia la *costante* nel *registro*;
- `movabs spiazamento, %rax`: leggi dall'indirizzo *spiazamento* in memoria e scrivi in **rax** (o nei suoi sotto-registri).
- `movabs %rax, spiazamento`: scrivi il contenuto di **rax** (o dei suoi sotto-registri) all'indirizzo *spiazamento* in memoria.

Ogni istruzione può lavorare con operandi di 8, 16, 32 o 64 bit. Se uno dei due operandi è un registro l'assemblatore può dedurre automaticamente la dimensione, altrimenti è necessario specificarla aggiungendo uno dei seguenti suffissi al codice operativo dell'istruzione:

- `b` per operandi di tipo byte;
- `w` per operandi di tipo “parola” (*word*, due byte).
- `l` per operandi di tipo “parola lunga” (*long word*, quattro byte).
- `q` per operandi di tipo “parola quadrupla” (*quad word*, otto byte).

Si consiglia di aggiungere sempre il suffisso, anche quando non serve.

## 2 Esempio di programma

Prepariamo un semplice esempio di programma da fare eseguire al nostro elaboratore. Dato che ci interessa principalmente sapere cosa succede durante l'esecuzione all'interno dell'elaboratore, usiamo il linguaggio assembler, che è il più vicino al linguaggio macchina: ogni istruzione di assembler si traduce in una istruzione di linguaggio macchina. Sempre per lo stesso motivo, ci conviene pensare che tutta la procedura di preparazione del programma in linguaggio macchina (scrittura del file sorgente, assemblamento, collegamento) sia svolta all'esterno del calcolatore, su un altro calcolatore o in qualunque altro modo, anche se in pratica usiamo lo stesso calcolatore per fare tutto. L'esecuzione del nostro programma comincia dal momento in cui esso è stato caricato in memoria; a quel punto, esiste solo il linguaggio macchina e tutta la procedura precedente non conta più.

La procedura inizia preparando un file di testo che contiene il programma in linguaggio assembler. Si consideri il file di Figura 3. I numeri di riga servono per riferimento e non fanno parte del contenuto del file. L'assemblatore produce una o più sequenze di byte da caricare in memoria, in base alle direttive o alle istruzioni contenute nel file sorgente. Le parole chiave che cominciano per “.” sono *direttive* e servono a chiedere all'assemblatore di svolgere vari compiti. Alla riga 1 troviamo la direttiva **.data** che chiede all'assemblatore di aggiungere alla *sezione data* ciò che segue nel file fino alla prossima direttiva che specifica una nuova sezione (in questo caso, la direttiva **.text** alla riga 8). Le sezioni sono sequenze di byte che possono essere caricate indipendentemente. La sezione “data” verrà caricata in una zona di memoria accessibile sia in lettura che in scrittura, mentre la sezione “text” in una zona di sola lettura. Normalmente la sezione data contiene variabili, mentre la sezione text contiene codice, ma niente di tutto ciò è imposto o controllato dall'assemblatore.

Alla riga 2 troviamo la definizione di una *etichetta* (`num1` in questo caso). Le etichette servono a dare un nome all'indirizzo del primo byte che le segue. Abbiamo bisogno delle etichette perché, avendo delegato all'assemblatore e al collegatore il compito di decidere dove caricare le sezioni, non sappiamo gli

```

1  .data
2  num1:
3      .quad    0x1122334455667788
4  num2:
5      .quad    0x9900aabbccddeeff
6  risu:
7      .quad    -1
8  .text
9  .global main
10 main:
11     movabsq $num1, %rax
12     movq (%rax), %rcx
13     movabsq $num2, %rax
14     movq (%rax), %rbx
15     addq %rbx, %rcx
16     movabsq $risu, %rax
17     movq %rcx, (%rax)
18     ret

```

Figura 3: Un esempio di programma scritto in Assembler GNU per x86\_64.

indirizzi delle entità che definiamo. Tramite le etichette possiamo riferirci a tali indirizzi simbolicamente, e lasciare che siano poi l'assemblatore e il collegatore a sostituirle con i veri indirizzi.

Alla riga 3 troviamo la direttiva **.quad** seguita da un numero. La direttiva chiede all'assemblatore di riservare 8 byte (a partire dal punto della sezione in cui è arrivato) e di inizializzarli con il numero specificato. Possiamo specificare il numero in varie basi (esadecimale, in questo caso): l'assemblatore provvederà a convertire in ogni caso il numero in binario. Più in generale è possibile riservare e inizializzare una sequenza di quad, semplicemente facendo seguire **.quad** da una lista di numeri separati da virgole. Altre direttive di questo tipo sono **.byte**, per riservare e inizializzare una sequenza di byte, **.long** (per doppie parole) o **.word** (parole).

L'effetto delle righe 2 e 3 è di allocare e inizializzare una variabile che occupa una parola quadrupla e darle il nome `num1`. Le righe 4-5 e 5-6 fanno la stessa cosa con `num2` e `risu`.

Alla riga 8 diciamo all'assemblatore che quanto segue deve essere aggiunto alla sezione "text".

Alla riga 9 troviamo la direttiva **.global** seguita dall'etichetta `main`. Con questa direttiva stiamo chiedendo di rendere questa etichetta visibile al collegatore. Il collegatore cercherà questa etichetta per identificare la prima istruzione del nostro programma<sup>2</sup> Il registro **rip** verrà inizializzato con questo indirizzo

<sup>2</sup>Vedremo in seguito che il realtà il collegatore cerca l'etichetta `start`, che normalmente si trova in un file oggetto che viene automaticamente collegato al nostro programma; il codice

quando il programma dovrà essere eseguito.

Alla riga 10 definiamo l'etichetta `main`. Si noti che non c'è differenza sintattica tra la definizione di `num1`, `num2` e `risu` da una parte e `main` dall'altra, anche se nelle nostre intenzioni le prime sono variabili mentre `main` è un'etichetta del programma. Non è casuale, in quanto per l'assemblatore non c'è alcuna differenza tra queste etichette. In ogni caso l'etichetta serve a dare un nome all'indirizzo del byte che la segue. L'assemblatore non segnalerà alcun errore se proviamo a saltare a `num1` o se proviamo a leggere o scrivere all'indirizzo `main`.

Alle righe 11–18 c'è il programma vero e proprio. Ogni riga contiene una istruzione per il processore. Quello che l'assemblatore farà sarà di tradurre ogni riga nella corrispondente sequenza di byte di linguaggio macchina, andando così a formare la sezione `text`. Si noti che questo non è molto diverso da quanto l'assemblatore fa con la direttiva `.quad` (o `.byte`, etc.): si tratta in ogni caso di costruire a poco a poco la sequenza dei byte che compongono la sezione che poi verrà caricata in memoria. Di fatto, niente vieta di usare le direttive `.quad` etc. nella sezione `text` o di scrivere una istruzione nella sezione `data`. In ogni caso si ottengono dei byte, e questi non sono di per sé né dati, né istruzioni: è solo nel momento in cui la CPU li utilizzerà che verranno interpretati in un modo o in un altro.

Il programma vuole calcolare la somma dei due numeri memorizzati agli indirizzi `num1` e `num2` e scrivere il risultato all'indirizzo `risu`. Per farlo, carica il primo numero nel registro `rcx` (righe 11–12), il secondo numero nel registro `rbx` (righe 13–14), li somma scrivendo il risultato in `rcx` (riga 15), infine copia il risultato in `risu` (righe 16–17). Infine, termina alla riga 18<sup>3</sup>

Osserviamo la riga 11. L'obiettivo è di caricare l'indirizzo della prima variabile in `rax`, in modo da poter poi caricare il primo numero tramite indirizzamento indiretto (riga 12). Il primo operando è di tipo *immediato* (lo si riconosce dal carattere `$`). Non ci si lasci confondere dall'uso dell'etichetta: l'etichetta sparirà e verrà sostituita dal suo valore numerico (si ricordi che gli indirizzi sono numeri). L'istruzione sta caricando una costante in `rax`. Questa costante (avendo usato l'etichetta `num1`) non è altro che l'indirizzo del primo numero da sommare. Stiamo usando `movabs` perché, non sapendo dove il programma verrà caricato, non possiamo sapere se tale indirizzo può essere contenuto in soli 32 bit, e `movabs` è l'unica istruzione che accetta operandi immediati a 64 bit.

Si noti che al posto delle istruzioni 12 e 13 avremmo potuto scrivere

```
movabsq num1, %rax
movq %rax, %rcx
```

In questo caso il primo operando della `movabs` è di tipo memoria (lo si riconosce dal fatto che non inizia né con `$`, né con `%`). L'istruzione sta ordinando al processore di eseguire una operazione di lettura in memoria di 8 byte (dal momento che la destinazione è `rax`) a partire dall'indirizzo `num1` (di nuovo: nel

in `start` esegue alcune inizializzazioni e poi chiama `main`.

<sup>3</sup>In realtà ritorna al codice `start` che aveva chiamato `main`; `start` conterrà le istruzioni che chiedono al sistema operativo di terminare l'esecuzione.

linguaggio macchina finale `num1` sparirà e al suo posto ci sarà il vero indirizzamento). Questa istruzione può caricare solo in **`rax`**, quindi abbiamo bisogno della successiva per copiare il valore letto in **`rcx`**, dove lo volevamo.

Stesse considerazioni valgono per le righe 13–14 e, nella direzione opposta, per le righe 16–17.

## 2.1 Assemblamento ed esecuzione in modo “nativo”

Se l’elaboratore che stiamo usando usa l’architettura Intel/AMD a 64 bit, possiamo assemblare ed eseguire il programma direttamente.

Supponiamo che il file in cui abbiamo scritto il programma si chiami `sum.s`. Per assemblarlo lanciamo il comando

```
as sum.s
```

Se non ci sono errori di sintassi, l’assemblatore non stampa niente e produce il file `a.out` (assembler output). È possibile cambiare il nome del file prodotto usando l’opzione `-o` seguita dal nome desiderato. Conviene farlo, visto che anche il collegatore usa il nome `a.out` come default:

```
as sum.s -o sum.o -g
```

(Abbiamo aggiunto anche l’opzione `-g`, che include nel file le informazioni utilizzate dal debugger.)

Tale file deve essere poi passato al collegatore per produrre l’eseguibile (invochiamo `g++`, che a sua volta invocherà il collegatore con i parametri opportuni):

```
g++ sum.o -o sum -g
```

(Anche qui abbiamo usato l’opzione `-o` per specificare il nome dell’eseguibile e l’opzione `-g` per il debugger.)

Il programma può essere infine caricato ed eseguito scrivendo

```
./sum
```

In questo caso il programma non produce alcun output, perché ci siamo limitati a scrivere il risultato in memoria, senza inviarlo ad alcuna periferica di I/O. Possiamo però vederlo in funzione utilizzando il debugger. Scriviamo

```
gdb sum
```

`gdb` è un debugger a riga di comando per il quale esistono varie interfacce grafiche (per esempio, `ddd`). Possiede anche una semplice interfaccia semi-grafica incorporata; la si può far partire premendo prima i tasti “Ctrl+x” e poi il tasto “a”. Per mettere in funzione il nostro programma settiamo prima un breakpoint all’etichetta `main` scrivendo `b main` (invio), poi facciamo partire l’esecuzione scrivendo `r` (e poi invio). Possiamo far avanzare il programma di una istruzione alla volta scrivendo `ni` (invio). Per esaminare il contenuto dei registri possiamo cambiare il layout dell’interfaccia con il comando `layout reg` (invio). In qualunque momento possiamo terminare il debugger scrivendo `q` (invio).

gdb ha molti comandi, si consiglia di consultare qualche guida in rete o l'help incorporato (comando `help`).

## 2.2 Assemblamento ed esecuzione in modalità “cross”

Se abbiamo installato libce<sup>4</sup> e QEMU<sup>5</sup> possiamo anche assemblare il nostro programma in modalità “cross” ed eseguirlo sulla macchina virtuale QEMU. Con modalità “cross” si intende appunto che usiamo un assembler/collegatore che gira sul nostro elaboratore, ma produce codice che deve essere caricato su un'altra macchina, che può a questo punto avere anche una architettura diversa. *Nota bene:* se il nostro elaboratore è un Mac Silicon non abbiamo altra scelta se non operare in modalità cross, in quanto la CPU del nostro Mac comprende un linguaggio macchina completamente diverso da quello Intel/AMD che abbiamo usato nel nostro programma. Attenzione: ciò è vero *anche* se stiamo usando Linux dentro una macchina virtuale VMware sul Mac Silicon. Infatti, il tipo di virtualizzazione usata da VMware (*hardware-assisted*) permette alla VM di girare velocemente, ma non di cambiare l'architettura di base della CPU. QEMU, invece, usa una tecnica (*binary translation*) che da una parte è piú lenta, ma dall'altra permette di emulare qualsiasi architettura.

La libreria libce contiene alcuni script che semplificano tutte le operazioni. Per usarli è necessario scrivere i file del programma in una directory in cui non ci sia altro, per es.:

```
mkdir sum
cd sum
```

Il file `sum.s` deve trovarsi in questa directory e tutti i comandi vanno lanciati da questa directory. Lo script `compile` (senza argomenti) provvede a compilare/assemblare e collegare tutti i file contenuti nella directory corrente, usando le opzioni corrette per la modalità cross. Se non ci sono errori il risultato è un file `a.out`, sempre nella directory corrente. Questo file non è un eseguibile per il nostro elaboratore, ma deve invece essere caricato nella macchina QEMU. Per farlo basta lanciare lo script `boot`. Si noti che in questo caso la macchina parte e si spegne velocemente, in quanto è configurata per spegnersi non appena termina il programma. Per questo esempio (visto che comunque il programma non stampa niente) conviene far partire la macchina in modalità “debug” nel seguente modo:

```
boot -g
```

Ora la macchina parte e si mette in attesa del collegamento di un debugger esterno. Per collegarlo dobbiamo aprire un secondo terminale e portarci nella stessa directory da cui abbiamo fatto partire la macchina (`sum` nel nostro caso), quindi lanciare lo script `debug` (senza argomenti). Partirà una istanza di `gdb` che si collegherà alla macchina QEMU, permettendoci di esaminare lo stato e

<sup>4</sup>[https://calcolatori.iet.unipi.it/istruzioni\\_libce.php](https://calcolatori.iet.unipi.it/istruzioni_libce.php)

<sup>5</sup>[https://calcolatori.iet.unipi.it/istruzioni\\_qemu.php](https://calcolatori.iet.unipi.it/istruzioni_qemu.php)

di controllare l'esecuzione. In questo caso `gdb` è stato configurato in modo da mostrare più informazioni. In particolare, mostra automaticamente il contenuto dei registri generali, il codice sorgente e la parte superiore della pila. Per il resto può essere usato come il normale `gdb`.