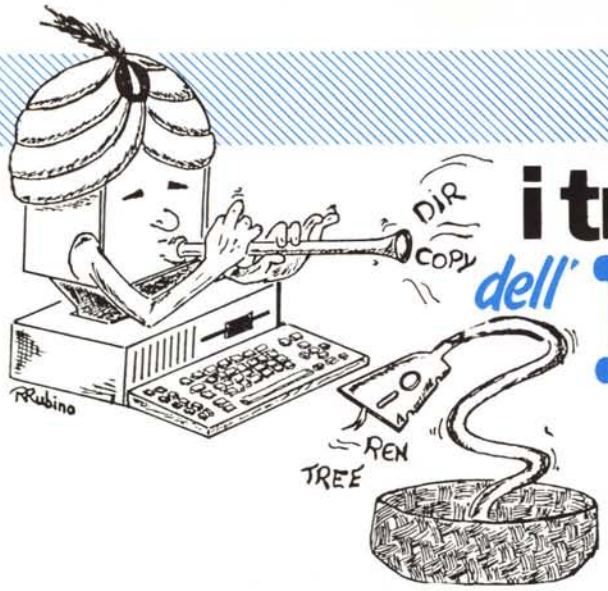


i trucchi dell' MS-DOS

a cura di Pierluigi Panunzi

Program Segment Prefix

Come abbiamo promesso nelle precedenti puntate, in questa parleremo di un'altra struttura creata e gestita dall'MS-DOS all'atto del caricamento di un programma e subito prima della sua esecuzione: «PSP» dalle iniziali delle parole «Program Segment Prefix», come apparirà subito da quanto diremo rappresenta un «prefisso» del segmento in cui è posto il programma appena caricato e da eseguire.

Riassunto delle puntate precedenti...

Nelle puntate precedenti abbiamo in particolare visto le differenze sostanziali tra un programma di tipo «.com» ed uno di tipo «.exe», tra le quali spicca la più importante che impone per i primi una limitazione a 64 kbyte nell'estensione; inoltre abbiamo visto che i «.exe» sono memorizzati su disco in maniera particolare, laddove il programma vero e proprio (l'immagine di come sarà caricata la memoria) è preceduto da una struttura tabellare denominata «Program Header», contenente tutte le informazioni necessarie al DOS per caricare correttamente il programma in memoria e per modificare in modo opportuno quelle locazioni di memoria («relocation items») che dipendono dalla posizione in cui il programma è caricato.

Viceversa, i programmi di tipo «.com» sono memorizzati su disco senza aggiunte o modificazioni e vengono caricati così come sono in memoria, per essere poi immediatamente eseguiti a partire da un entry point corrispondente all'indirizzo 0100H.

Caricato direttamente in memoria il «.com» oppure il «.exe», il DOS cede il controllo al programma, che verrà così eseguito.

In realtà, il DOS, prima di cedere definitivamente il controllo al programma, costruisce il cosiddetto PSP, che rappresenta una potente interfaccia «logica» tra il nostro programma ed il Sistema Operativo stesso.

Il caricamento e l'allocazione in memoria di un file

Supponiamo dunque di voler eseguire un nostro programma: se ci troviamo in «ambiente DOS» e cioè quando vediamo sul video il prompt «A:\>» o «C:\>» sappiamo che non dobbiamo far altro che digitare il nome del programma da eseguire.

Il «COMMAND.COM» (una nostra vecchia conoscenza), provvede a fare l'analisi sintattica del comando digitato, a riconoscere se si tratti di un comando di sistema (non è il nostro caso) oppure di un file residente su un disco.

Altra possibilità che abbiamo di eseguire un nostro programma, ma molto più complessa, è quella di stare già all'interno di un programma in corso di esecuzione, il quale deve in un certo istante eseguire il nostro programma.

Per chi non avesse idea ben chiara su quanto stiamo dicendo diamo subito un esempio: il ben noto programma «Lotus 1-2-3» può essere attivato in due modi, il primo semplicemente digitando «123», con il che il programma parte subito mentre il secondo modo si ha per mezzo del comando «acces», che viceversa ci fa comparire sul video un menu dal quale possiamo, tra l'altro, scegliere di far eseguire il nostro «123».

In questo caso dunque il programma «access.com» contiene al suo interno la chiamata al programma «123.exe», cosa che si ottiene, come vedremo nelle prossime puntate, per mezzo di un'apposita routine del DOS, chiamata in gergo «EXEC».

In entrambi i casi, dunque, sia da ambiente DOS che da programma per mezzo della EXEC, ecco che il sistema operativo sa che deve caricare in memoria un file al quale dovrà assegnare una certa zona di memoria: per tale scopo andrà dapprima a calcolare qual è la prima locazione di memoria libera.

Dal momento che si parla sempre di segmenti di programma (dato che il microprocessore è un 8088) ecco che il DOS assegnerà al programma il primo segmento libero (quello cioè ad indirizzo fisico il più basso possibile), che avrà così la denominazione di «Program Segment».

A questo punto il DOS creerà PSP, allocato a partire da un offset pari a 0000H, mentre successivamente, a partire dall'offset 0100H, verrà caricato il programma vero e proprio, secondo i meccanismi visti in precedenza.

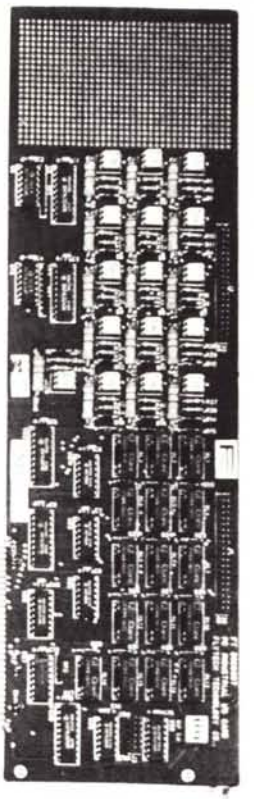
DISK-DRIVE
STAMPANTI
COMPUTER
MONITOR

LA CASA DEL COMPUTER

IMPORTAZIONE DIRETTA

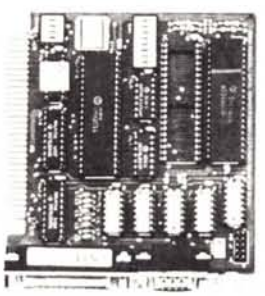
IL PIÙ VASTO ASSORTIMENTO DI ADD-ON CARDS PER PC/XT/AT

OLTRE **80** MODELLI... DIVERSI...



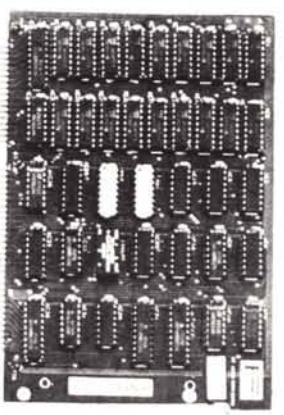
MODEM CARD
— Hayes compatibile
— CCITT V.21, V.22
— 300-1200 Bps
Cod. 11.9600

* **B.S.C. CARD**
* **RS-422**
* **4/8 SERIALI PER XENIX**



AT-PARALLEL/SERIAL
— 1 x Parallel Port
— 1 x Serial Port
Cod. 12.0300

* **AD-DA DM-P005** (uso industriale)
* **AD-DA 14** (uso industriale)
* **AD-DA FPC-010** (uso didattico)
* **AD CONVERTER** (alta velocità)
* **MULTI-DA** (alta velocità)



AT-128K RAM CARD
— Provvede ad espandere la memoria RAM da 512K a 640K
Cod. 12.0895

* **IC TEST**
* **SCHEDA PARLANTE**

INDUSTRIAL I/O
— 16 x Relay output
— 16 x Photo couple input
Cod. 11.8700

* **PAL WRITER** (standard Jedec)
* **B. PROM WRITER**
* **E. PROM WRITER 2716-27512** (da 1 a 10 textool)
* **8748-8749 WRITER**

BAR CODE READER

* Legge tutti i codici a barre
* Emula la tastiera del PC/XT/AT
* Semplice da installare

AMPIA VARIETÀ DI

* DATA SWICHES
* SWITCH BOX
* CONVERTITORI DI PROTOCOLLO
* BUFFER 16/64/256 e 1MB
* PENNE OTTICHE
* CAVI STAMPANTI PARALL., SERIALI, ECC.
* ACCESSORISTICA PER CAVI SERIALI
* GRUPPI DI CONTINUITÀ

TELEFONATECI, NON POSSIAMO ELENCARVI TUTTO!

RICHIEDETECI IL CATALOGO - SCONTI AI SIG.RI RIVENDITORI

LA CASA DEL COMPUTER Via della Misericordia, 94 (sede) - PONTEDERA (Pisa)
Via T. Romagnola, 63 (magazzino) - FORNACETTE (Pisa) **Tel. 0587/422.022**

Vediamo dunque una prima differenza nei due casi «.com» e «.exe»: per i programmi del primo tipo, in generale più innocui, il fatto di iniziare ad un offset pari a 0100H è più che naturale, mentre sappiamo che gli «exe» possono iniziare ad un indirizzo qualsiasi, all'interno di un segmento e perciò un offset iniziale di 0100H riuscirebbe errato: come mai allora il tutto funziona bene?

È presto detto!

Per quanto riguarda i «.com» sappiamo che i quattro registri segmento (CS, DS, SS ed ES) a questo punto vengono inizializzati tutti allo stesso valore, proprio il valore del segmento assegnato dal DOS e che abbiamo indicato con «Program Segment».

Viceversa per i «.exe» sappiamo che i registri CS, IP nonché SS ed SP vengono alterati secondo i valori riportati all'interno del «Program Header», rispettivamente per definire il segmento di codice e di stack, a loro volta definiti dal «linker»: non abbiamo parlato degli altri due registri di segmento mancanti (DS ed ES), i quali per l'appunto, contengono proprio il valore del segmento dove è posto il PSP, il «Program Segment».

Ecco che dunque è spiegabile l'apparente incongruenza, grazie anche al modo di rappresentare gli indirizzi di memoria sotto forma «segment-offset».

Facciamo un esempio chiarificatore: supponiamo che il nostro «.exe» abbia come indirizzo iniziale proprio 0000H (che poteva essere qualsiasi, tanto il discorso non cambiava); supponiamo inoltre che il DOS assegni a tale programma il segmento «0A00H»: ciò vuol dire che avrà assegnato al nostro programma tutta la memoria che va dall'indirizzo 0A00H:0000H all'indirizzo 0A00H:FFFFH e cioè i 64 kbyte che vanno dall'indirizzo «fisico» 0A000H a 19FFFH (si ricordano i lettori come si calcolano gli indirizzi fisici?!).

A partire dunque da 0A00H:0000H il DOS porrà il PSP, seguito subito a ruota, a partire dall'indirizzo 0A00H:0100H, dal programma vero e proprio.

Ma noi sappiamo che il nostro programma deve partire da un offset nullo, ed allora invece di spostare tutto il programma, la cosa più ovvia e banale da fare è modificare il CS in modo tale che l'indirizzo completo di partenza del programma abbia offset nullo. In particolare, il DOS porrà il valore 0A00H nei registri DS ed ES, mentre nel CS metterà un valore pari a 0A10H, tale che (con un offset nullo) punti proprio al byte in cui è stato caricato il programma da eseguire.

Facendo i conti, infatti, l'indirizzo 0A00H:0100H a cui è stato caricato il programma (e perciò riferito al «Program Segment») è assolutamente identico (provare per credere!) all'indirizzo 0A10H:0000H fornito invece dalla coppia CS:IP.

I conti tornano, senza aver sprecato un solo byte di memoria...

In generale dunque si può notare che all'istante in cui viene dato il controllo al nostro programma «.exe» e cioè subito prima che venga eseguita la sua prima istruzione, il CS è pari al DS aumentato di 0010H, cosa che si può verificare ad esempio con il solito «debug».

La struttura interna del PSP

Abbiamo riportato nella tabella sottostante una sintesi della struttura del PSP.

Iniziamo l'analisi, indirizzo per indirizzo, escludendo i campi «riservati», dei quali nulla è dato sapere, in quanto utilizzati esclusivamente dall'MS-DOS.

Offset 0000H: è una word contenente il codice operativo dell'istruzione Assembler «INT 20H», quella generalmente usata da un programma scritto in Assembler per ritornare al DOS, proprio alla fine del programma stesso. In particolare questa istruzione posta all'offset 0000H del segmento di codice fa sì che si possa rientrare al DOS in maniera alternativa effettuando un salto (JMP) a tale indirizzo, a patto che (e questo è di fondamentale importanza) il nostro programma Assembler al termine non sia saltato ad un segmento di codice differente da quello di partenza.

Comunque per terminare correttamente un programma Assembler è consigliabile utilizzare l'istruzione già vista (INT 20H), oppure l'analoga funzione dell'MS-DOS richiamabile da INT 21H con il registro AH azzerato; meglio ancora è terminare il programma, specie se di tipo «.exe» con una chiamata alla funzione 4CH (valore da porre in AH) dell'MS-DOS (ancora attivata con INT 21H).

Offset 0002H: si tratta di una word che indica il segmento relativo alla prima locazione di memoria non utilizzata dal DOS e perciò ci consente di sapere quanta memoria ha a disposizione il nostro programma, semplicemente sottraendo il valore espresso (moltiplicato per 16 in quanto l'unità di misura adottata è il «paragrafo») dall'ammontare della memoria del nostro personal (in generale 640 kbyte).

Bisogna stare attenti perché questo calcolo è valido solo se nella memoria non sono presenti dei dischi virtuali (RAMDISK, VDISK, ecc.) che di solito vengono posti e resi residenti in indirizzi «alti».

Struttura del PSP (Program Segment Prefix).

offset	significato
0000 0001	Chiamata al DOS (INT 20H)
0002 0003	dimensione della memoria (in paragrafi)
0004	riservato
0005 ... 0009	chiamata al dispatcher del DOS
000A ... 000D	indirizzo "Terminate"
000E ... 0011	indirizzo "Ctrl-Break"
0012 ... 0015	indirizzo "Critical Error Handler"
0016 ... 0028	riservati
002C 002D	segmento dell'"environment"
002E ... 004F	riservati
0050 0051	chiamata al DOS (INT 21H)
0052 ... 005B	riservati
005C ... 006B	primo FCB (File Control Block)
006C ... 007F	secondo FCB
0080 0081 ... 00FF	lunghezza stringa di comando stringa di comando

AVETE MAI PENSATO CHE...

DISK-DRIVE · STAMPANTI · COMPUTER · MONITOR

LA CASA DEL COMPUTER

IMPORTAZIONE DIRETTA

LA C.D.C. importa direttamente dai costruttori di INTERFACCE, MAIN BOARD, TASTIERE, CASES, ecc. **solo le parti staccate** per garantire il meglio della produzione orientale ed inoltre ASSEMBLA in proprio effetuando un TEST PRELIMINARE DI FUNZIONAMENTO.

LA C.D.C. inserisce sui propri PC/XT/AT* da SEMPRE solo ed esclusivamente i DRIVE CHINON che sono sinonimo di qualità, silenziosità, ed affidabilità.

LA C.D.C. è organizzata in modo da avere SEMPRE pronto a magazzino quanto Vi occorre e può effettuare spedizioni ANCHE IN GIORNATA (SERVIZIO RAPIDO PER LE ISOLE 24 ORE IN PREPAGATO).

LA C.D.C. GARANTISCE i propri prodotti con la sostituzione immediata o riparazione ANCHE DOPO IL PERIODO DI GARANZIA (servizio HALF COST).

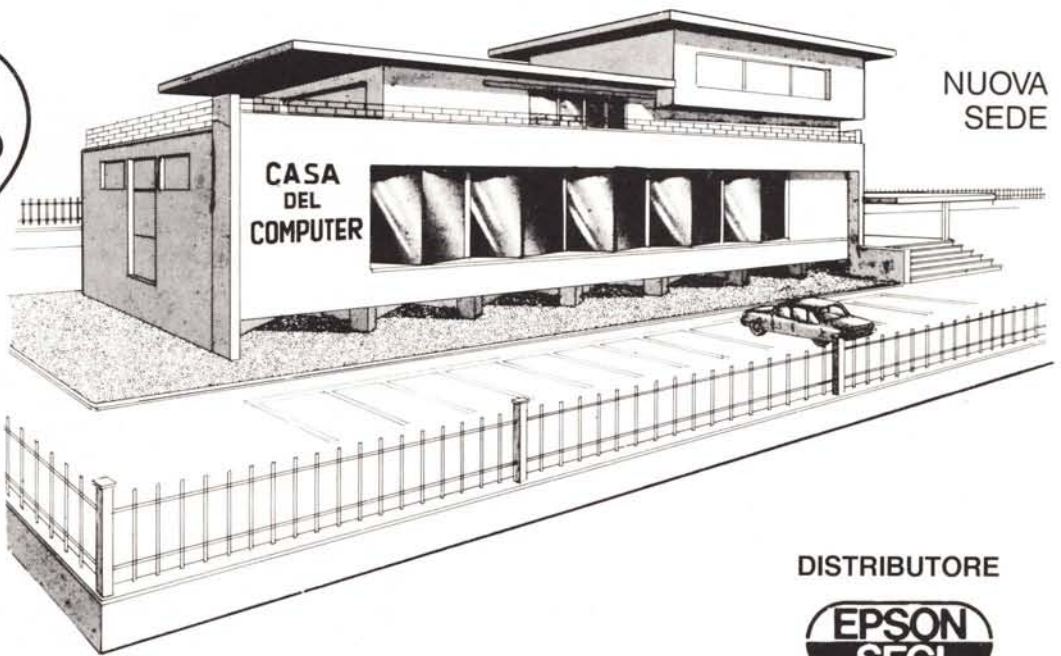
LA C.D.C. ha tutti i pezzi di ricambio a magazzino degli articoli di propria importazione che vengono conservati per minimo 5 ANNI.

VELOCI SPEDIZIONI IN TUTTA ITALIA

SPESSO È MEGLIO SPENDERE QUALCOSA IN PIÙ PER SPENDERE MENO...

... PENSATECI...!!!

1° CASH & CARRY ALL'INGROSSO



NUOVA SEDE

DEPOSITI:

BOLOGNA

TELETEX s.r.l. - Via Emilia, 51
Anzola Emilia (Bo) - Tel. 051/734485

AGENZIE:

ROMA

H2S s.r.l.
Via Assisi, 80
Tel. 06/7883697

MILANO

C.S.M. SISTEM s.r.l.
Via Valsolda, 21
Tel. 02/8435685

TORINO

R.M. PROFESSIONAL
Via Accademia Albertina, 35/C
Tel. 011/510173

DISTRIBUTORE

EPSON SEGI

- PC / XT / AT COMPATIBILI
- PC / XT PORTATILI
- INTERFACCE x APPLE/IBM

- MODEM
- STAMPANTI
- FLOPPY DISK DRIVE
- HARD DISK
- STREAMER
- MONITOR
- DISKETTE

LA CASA DEL COMPUTER Via della Misericordia, 94 (sede) - PONTEDERA (Pisa)
Via T. Romagnola, 63 (magazzino) - FORNACETTE (Pisa) **Tel. 0587/422.022**

RICHIEDETECI IL CATALOGO E PREVENTIVI OGGI STESSO!!

Offset 0005H: si tratta di cinque byte rappresentanti in prima analisi una «long JMP» alla routine del DOS che effettua l'analisi di un comando «dispatcher»): questo almeno secondo quanto riporta il «Technical Reference Manual», in quanto andando a vedere con il debugger il contenuto di questo campo abbiamo trovato i seguenti byte

```
9A 2E CD 39 F3
```

corrispondenti invece ad una CALL «lunga» ad un indirizzo «strano».

Infatti disassemblando l'istruzione si ha

```
CALL OF339H:0CD2EH
```

che non corrisponde ad alcun punto utile della ROM contenente il BIOS. Qualche lettore ci può venire in aiuto?

Comunque lungi dall'essere utile per effettuare la chiamata al dispatcher stesso (infatti ciò è vivamente sconsigliato dal già citato manuale) il valore indicante l'offset dell'indirizzo a cui saltare (posto all'offset 0006H del PSP, subito dopo l'op-code posto all'offset 0005H del PSP) può servire a farci sapere effettivamente quanti byte il DOS può assegnare al nostro programma: il meccanismo relativo è alquanto complicato e tralasciamo senz'altro di parlarne, per non appesantire ulteriormente questa puntata già di per sé parecchio «indigesta».

Offset 000AH, 000EH, 0012H: si tratta di tre «vettori» (formati cioè da tre coppie «offset: segment», e perciò sono tre double-word), indicanti l'indirizzo di tre routine (altrimenti attivabili rispettivamente con INT 22H, INT 23H e INT 24H) relative ad altrettante routine, delle quali parleremo nelle prossime puntate, che gestiscono rispettivamente:

- il ritorno al DOS (o al programma chiamante nel caso di lancio di un programma tramite la funzione EXEC);
- l'esecuzione di una particolare routine all'atto della pressione dei tasti «Ctrl-Break» durante l'esecuzione di un programma;
- una routine di gestione di errori che impediscono il corretto funzionamento del DOS (errori dell'unità a dischi in primo luogo).

Offset 002CH: si tratta di una word rappresentante il valore del segmento in cui è posto dal DOS il cosiddetto «environment» attuale, che il programma da eseguire eredita dal programma chiamante (il DOS o il programma che ha utilizzato la funzione EXEC). Ne parleremo diffusamente nelle prossime puntate, mentre ora diciamo solo che si tratta di quel complesso di stringhe utilizzate dal DOS per passare informazioni da un programma all'altro.

Per intenderci si tratta di stringhe tipo

```
PATH = C:\PROVA
```

impostate a livello DOS con il comando «path», come pure tutte le associazioni effettuate per mezzo del comando «set», associazioni che costituiscono in un certo senso il «corredo di informazioni» che un programma «trasmette» ai programmi da esso lanciati.

Torneremo come detto su questo strano, ma utile modo: ricordiamo che già a livello BASICA è possibile gestirlo per mezzo dell'istruzione «ENVIRONMENT».

Offset 0050H: si tratta di una coppia di byte tutto sommato inutile in quanto rappresenta il codice operativo dell'istruzione Assembler INT 21H, la chiave per attivare una delle tantissime funzioni dell'MS-DOS (ponendone in AH il numero identificativo). Tale campo può servire a quei programmatori che invece di usare la semplice INT 21H vogliono effettuare una CALL («lunga» e cioè dotata di offset e segmento) all'offset 0005H del segmento in cui si trova il PSP: una spiegazione sulla presenza di tale campo (come pure di altri campi apparentemente inutili) crediamo che possa essere data solamente dai progettisti dell'MS-DOS...

Offset 005CH, 006CH: si tratta di due campi relativi ai cosiddetti FCB (File Control Block) di altrettanti file che il nostro programma può richiedere nella linea di comando, all'atto dell'esecuzione del programma stesso: anche in questo caso le informazioni riportate in questi due campi sono in un certo senso «reliquie del passato» mantenute per un certo grado di compatibilità verso l'ormai vecchio sistema operativo CP/M, che guarda caso allocava proprio all'indirizzo 005CH e 006CH le indicazioni relative ai nomi dei file sui quali il programma in esecuzione poteva lavorare.

In realtà con le versioni 2.00 e successive dell'MS-DOS l'uso dell'FCB per gestire un file all'interno di un programma Assembler viene calorosamente sconsigliato grazie all'introduzione di nuove funzioni del DOS, per le quali il nome del file da gestire viene associato all'inizio ed una word detta «file handler», che viceversa viene utilizzata da tutte le funzioni che operano sul file: in parole povere, dovendo «aprire» un file, scriverci qualcosa all'interno e successivamente «chiuderlo», «anticamente» si usava l'FCB in tutte e tre le operazioni, il che costringeva il DOS stesso a controllare ogni volta che si scriveva all'interno di un file che il file stesso fosse stato già aperto.

Ora invece con la funzione di OPEN si associa il «file handler» al file di cui si fornisce il nome, mentre nelle funzioni successive si fa riferimento all'«handler» che ci garantisce che il file sia stato correttamente aperto.

Inutile dire che su questi argomenti ritorneremo largamente in dettaglio, data la loro fondamentale importanza.

Offset 0080H, 0081H: si tratta di un campo lungo 128 byte, condiviso da due funzioni.

La prima funzione è relativa alla memorizzazione della stringa di comando impostata all'atto dell'esecuzione del programma stesso (anche questa è una «reliquia» del vecchio CP/M), fatto che consente al programma in corso di esecuzione di conoscere i parametri impostati all'atto dell'attivazione del programma stesso: ad esempio se abbiamo impostato il comando

```
convert prog1.com prog2.com /x
```

all'interno del campo in esame troveremo, a partire dall'offset 0081H la seguente stringa

```
PROG1.COM PROG2.COM /X
```

e cioè tutto quanto digitato subito dopo il nome del programma attivato. Nel byte posto all'offset 0080H invece c'è la lunghezza in byte di tale stringa (nel nostro caso il valore 22H). C'è da aggiungere che eventuali caratteri di «input/output redirection» e di «piping» (i caratteri «<», «>» e «|») non vengono riportati in tale campo in quanto tale funzione di «redirection» viene gestita in maniera trasparente dall'MS-DOS e perciò deve risultare invisibile.

Altra caratteristica è che non appare nemmeno il nome del programma attivato, il quale dunque non potrà assolutamente sapere «come si chiama», sempre se ciò abbia un certo senso...

La seconda funzione associata a tale campo, che costringe ad un eventuale salvataggio delle informazioni in esso contenute che altrimenti andrebbero perse, è relativa all'area di default di trasferimento dati dall'unità a dischi, area chiamata DTA («Disk Transfer Area»), utilizzata da alcune funzioni del DOS.

Terminiamo dunque questa onerosa puntata consigliando ai lettori interessati di tenere bene a mente le informazioni sin qui riportate, in quanto si ritroveranno parecchie volte nelle puntate successive.

MC

OLTRE 3.000 CLIENTI SODDISFATTI HANNO ACQUISTATO

IL PIÙ VELOCE

PC/AT 286 ESISTENTE SUL MERCATO

DISK-DRIVE
STAMPANTI
COMPUTER
MONITOR

LA CASA DEL COMPUTER

IMPORTAZIONE DIRETTA

- SPEED UTILITY 13.1 MHz
- ZERO WAIT STATE 6/10 MHz
- DRAM 41256-100

DISPONIBILE ANCHE IN VERSIONE «BABY»: COMPACT 286

NON DIMENTICATE

CHE ABBIAMO SEMPRE PRONTA CONSEGNA A MAGAZZINO CON PREZZI IMBATTIBILI

* TURBO XT 4,77/8 MHz (versione economica)
* TURBO XT 4,77/10 MHz con NEC V-20
* PC PORTATILI BONDWELL 8
* PC TRASPORTABILI MITAC-VISO

SUPER SCONTI PER ORDINI SUPERIORI A 30 UNITÀ

SONO STATI SENSIBILMENTE RIDOTTI I PREZZI DI VENDITA

LA CASA DEL COMPUTER Via della Misericordia, 94 (sede) - PONTEDERA (Pisa)
Via T. Romagnola, 63 (magazzino) - FORNACETTE (Pisa) Tel. 0587/422.022