

CIPS

Criteria di progetto

A cura di: R. Bettarini
G. Daniele
R. Ferrini
G. Fresta
P. Mogorovich

Copyright - Marzo 1979

by - CNUCE - Pisa

Istituto del Consiglio Nazionale delle Ricerche

CIPS: Criteri di Progetto

PREMESSA

E' stato osservato come in un numero abbastanza vasto di discipline scientifiche, il ricercatore abbia la necessita' di trattare masse di dati, di dimensioni a volte rilevanti, organizzati sotto forma di vettori generalmente a

1, 2 o 3 dimensioni.

La gestione, l'elaborazione e la presentazione dei dati sono i tre momenti piu' caratteristici di tale problema.

Si e' notato inoltre che il trattamento di questi dati, richiede un insieme di operazioni che sono per buona parte identiche anche in discipline diverse. Da qui l'idea di progettare un sistema operativo che permetta all'utente di accedere facilmente a tutte le routine di uso comune e di implementare, in tale sistema, le proprie con relativa facilità.

Tale sistema verra' nel seguito chiamato col nome di CIPS.

CRITERI GENERALI

Nella progettazione di CIPS sono stati seguiti alcuni criteri generali:

- 1) Efficienza. CIPS deve presentare caratteristiche di efficienza nei tre aspetti della gestione dei dati:
 - a) facilita' di reperimento dei dati di interesse dell'utente in un archivio, possibilmente conoscendo a priori le caratteristiche dei dati che si intende elaborare.
 - b) elaborazione dei dati stessi, la quale deve avere caratteristiche di velocita', di precisione e di non sovraccarico dell'ambiente di calcolo.
 - c) presentazione dei risultati, sia sotto l'aspetto del display grafico che della rappresentazione numerica.
- 2) Facilita' di implementazione. CIPS deve offrire all'utente una gamma quanto piu' vasta di possibilita', ma deve altresì permettere all'utente di inserire i

propri algoritmi in maniera non molto piu' complessa della scrittura di un normale programma applicativo.

- 3) Sicurezza dei dati trattati. Il trattamento di grandi quantita' di dati e' in genere piuttosto pesante dal punto di vista del calcolo. Si rende quindi necessario gestire i dati ed i risultati in modo da prevenire la loro distruzione per un errore di qualsiasi tipo.
- 4) Facilita' di accesso per l'utente. L'utente deve essere messo in condizioni di dedicare tutta la sua attenzione al problema che sta trattando, evitandogli nciosi problemi di formato di comandi e fornendogli la possibilita' di provare rapidamente algoritmi anche complessi. In quest'ottica e' preferibile l'adozione di un sistema di tipo conversazionale, con comandi semplici e ben documentati, raggruppabili se richiesto in "macro-comandi" capaci di eseguire piu' funzioni elementari.
- 5) Predizione d'errore. La durata dei calcoli ed il loro costo giustificano l'uso di criteri tali da prevedere, sia nel caso di transazioni semplici (comandi) sia nel caso di transazioni complesse (macro-comandi), la possibilita' di un errore. Gli errori, per quanto riguarda la loro predizione, possono essere di tipo numerico (per es. overflow) o di tipo logico (per es.

sequenze di comandi incompatibili). Inoltre, nel caso di errori non previsti e' opportuno provvedere CIPS di un'interfaccia col sistema operativo affinche' eventuali interruzioni con conseguente reset del sistema non provochino la distruzione dei dati.

- 6) Documentazione. L'utilizzo di CIPS e' assai facilitato se l'utente ha a disposizione una valida documentazione in linea per quanto riguarda l'utilizzo dello stesso. Si puo' inoltre pensare a rendere accessibile a CIPS un archivio di documenti di vario tipo, organizzati secondo un criterio di accesso di tipo standard (per es. STAIRS).

La struttura generale di CIPS che e' spiegata qui di seguito, e' stata realizzata trovando nell'ambiente VM/CMS una concreta applicazione dei principi di cui al paragrafo precedente.

GESTIONE DELLA MEMORIA

Nell'utilizzo di una macchina virtuale sotto VM, le aree di memoria disponibili all'utente sono A e B cfr. fig.1. L'area B puo' essere concettualmente divisa in due parti, C e D, in quanto l'utente puo' ottenere zone di memoria con due tecniche diverse, o da C verso indirizzi crescenti, o da D verso indirizzi decrescenti. Nell'utilizzo normale del CMS i programmi di utente sono caricati nella zona C.

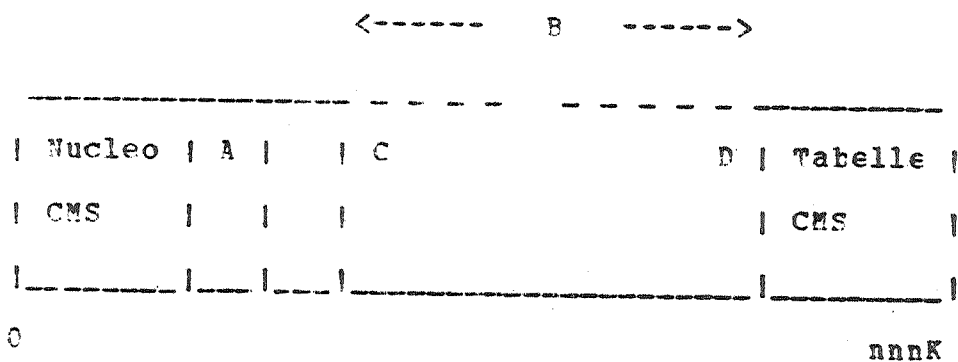


fig. 1

Al momento di iniziare una tipica seduta, l'utente deve avere una configurazione di macchina virtuale del tipo:

- 1) Disco utente (accesso "A")
- 2) Disco CIPS (read-only) (accesso "E")
- 3) Disco di lavoro (accesso D) (facoltativo)
- 4) Unita' nastro dedicate (secondo necessita')
- 5) Qualsiasi altro disco con accesso diverso da quelli gia' specificati.

Nel momento in cui l'utente inizia una seduta CIPS, tramite un comando di "START" vengono svolte alcune funzioni tali da garantire una perfetta compatibilita' col CMS e vengono acquisite aree di memoria nelle zone contrassegnate con "A" e con "D".

Vengono inoltre caricate routines di servizio che resteranno permanentemente residenti. In queste aree di memoria vengono inserite informazioni relative al tipo di seduta, al terminale usato ecc. (area GF), e altre relative alle matrici che vengono usate (area ICB).

Questi ultimi blocchi di controllo contengono informazioni quali l'indirizzo iniziale della matrice, le sue dimensioni, stringhe di spiegazione, rappresentazione interna della matrice ecc. Le matrici sono identificate da un nome simbolico definito dall'utente. Poiche' i vari ICB non sono necessariamente di lunghezza costante, sono

organizzati secondo puntatori a catena.

La definizione di un'immagine da parte dell'utente consiste nel costruire un ICB (a richiesta secondo formati standard o meno) e nel riservare nell'area "B" di memoria lo spazio opportuno a contenere i dati.

Accanto alla catena degli ICB e' caricata un'ulteriore catena di dati (RCB) che verranno descritti in seguito.

La memoria dell'area B e' organizzata nel modo seguente. L'area C (dal margine sinistro verso destra) e' riservata al caricamento dei programmi; l'area D (dal margine destro verso indirizzi decrescenti) contiene nell'ordine:

- a) routine considerate di frequente utilizzo e quindi permanentemente residenti;
- b) blocchi di controllo;
- c) i dati veri e propri.

Questa gestione di memoria e' organizzata in modo da proteggere i dati in essa contenuti. La protezione e' del tipo hardware. Per quanto riguarda i dati, le matrici che vengono richieste dei vari programmi in lettura restano in modo "read-only", mentre quelle su cui si scrive sono sbloccate per il solo periodo in cui la routine interessata resta attiva. Un'organizzazione di questo tipo richiede che i dati siano allineati al piu' piccolo segmento di memoria la cui chiave di protezione sia indipendente dagli altri;

tale segmento e' di 2048 byte. Si puo' quindi immaginare di avere uno spreco medio di 1k byte per matrice.

Per quanto riguarda i blocchi di controllo (ICB, PCB e GP) la protezione e' ancora piu' rigorosa. Le routine applicative non sono infatti in grado di modificarli, ma devono richiedere tale servizio al supervisore CIPS.

La gestione di masse di dati di dimensioni considerevoli direttamente nella memoria principale e' un problema delicato.

Esso e' senz'altro il modo piu' veloce di gestire tali dati, ma solo se si hanno a disposizione le risorse necessarie. Poiche' si puo' pensare che un limite ragionevole alla memoria virtuale utilizzata possa essere circa 2-Mbyte (salvo casi in cui il sistema e' particolarmente scarico), si vede che, se si ha a che fare con matrici di grosse dimensioni (per es. 512 x 512 elementi) la risorsa "memoria" si esaurisce rapidamente. Per questo i dati in memoria sono organizzati secondo vari formati; tipicamente Fullword (4 byte), Halfword (2 byte) e Byte (1 byte).

Qualora il ricorso a rappresentazioni del tipo Halfword o Byte fosse non sufficiente, o non possibile per il tipo di operazioni da effettuare, occorre provvedere con tecniche di Dump temporaneo. Il Dump temporaneo puo' essere di due tipi: manuale o automatico; nel primo caso e' l'utente che si preoccupa di liberare un'area di memoria causando la scrittura su disco di alcuni dati; associata alla tecnica di

dump e' la tecnica di compattazione dei dati per evitare frammentazione della memoria.

La tecnica di dump automatico funziona nel modo seguente. Ogni richiesta di definizione di matrice e' soggetta ad un controllo che ne verifica il possibile soddisfacimento. Nel caso positivo la richiesta e' soddisfatta. In caso negativo si provvede a scaricare automaticamente su disco una o piu' matrici, secondo criteri di minore utilizzo, per far posto alla nuova matrice. Lo stesso criterio e' seguito quando la richiesta di una matrice non e' a livello di definizione ma di utilizzo e la matrice stessa si trova su disco, per un dump precedente. I dati relativi alla posizione di una matrice sono nell'ICB.

GESTIONE DEI PROGRAMMI

Il fatto che CIPS sia attivo in un determinato momento non compromette affatto il contemporaneo uso del CMS.

Quest'ultimo puo' perfettamente svolgere le sue funzioni, salvo il fatto di avere una disponibilita' di memoria piu' limitata. Questo puo' portare in pratica ad una situazione in cui l'utente scrive un programma e lo compila (in ambiente CMS), lo prova su dati disponibili in memoria (in ambiente CIPS) e se necessario lo modifica ancora in ambiente CMS senza uscire dall'ambiente CIPS. I comandi CIPS si distinguono da quelli CMS per il fatto di essere preceduti dalla stringa "CIPS". Quindi "COPY ..." sara' un normale comando CMS, mentre "CIPS COPY ..." sara' una funzione CIPS.

Riprendiamo in esame quest'ultimo comando, "CIPS COPY parametri". Questa stringa sara' analizzata da un interprete che riceverà da essa, da informazioni contenute nei parametri e da altre ricavate nell'area "GF", il nome di un file di tipo text adatto a svolgere la funzione richiesta (per es. COPIA TEXT). Sorge ora il problema di interfacciare questo file con i dati indicati nei parametri. Per questo il file "COPIA TEXT" deve essere

corredato da due files COPIA RCB1 e COPIA RCB2 di cui il primo obbligatorio.

Il supervisore CIPS, integrando i dati ottenuti dal file "COPIA RCB1" e dai parametri associati al comando, seguendo un certo criterio gerarchico e' in grado di fornire al programma "COPIA" la "P-list" necessaria alla sua esecuzione. Oltre a questo svolge anche altre funzioni, quali rendere accessibili i dati (modificando la chiave di protezione degli stessi) e riblccarli al termine dell'esecuzione, ecc. Tra i dati contenuti nel file di tipo "RCB1" vi sono alcuni controlli numerici sui parametri (per es. range di validita' ecc.).

Il file di tipo "RCB2" relativo ad una routine, file che e' facoltativo, contiene informazioni di controllo del seguente tipo:

- a) relazioni di validita' tra i parametri della routine;
- b) relazioni di validita' tra la routine e altre routine chiamate precedentemente o successivamente;
- c) relazioni di validita' tra i parametri della routine e le altre routine (e/o i loro parametri) chiamate precedentemente o successivamente.

Questi dati servono a fare una verifica un po' piu' complessa sulla possibile presenza di errori nella chiamata di routine; le informazioni di cui ai punti b) e c) vengono usate nel caso di macro-comandi.

Questi ultimi sono sequenze di comandi, che svolgono piu' funzioni elementari, e sono raggruppati, per comodita', in un unico file. Le informazioni di cui ai precedenti punti b) e c) permettono quindi una verifica di coerenza di sequenze di comandi e di parametri associati.

L'utente che desidera usare una sua routine in ambiente CIPS dovra' semplicemente organizzare la sua routine in "SUBROUTINE" e scrivere almeno il corrispondente file RCB1. Questo e' scritto in un apposito linguaggio ad alto livello e tradotto tramite un apposito compilatore.

I files RCB1, di dimensioni estremamente ridotte, vengono conservati in memoria, dopo che sono stati letti una volta, e di tale area si tiene il catalogo. Una richiesta di RCB1 viene soddisfatta cercando nell'area degli RCB1 la presenza del blocco interessato; se tale ricerca fallisce, l'RCB1 viene letto da disco.

CIPS e' naturalmente fornito di una sua libreria, facilmente implementabile. Nella scrittura di programmi appartenenti a tale libreria vengono seguite tutte le tecniche necesarie ad ottenere la maggiore velocita' possibile durante l'esecuzione.

Il problema della velocita' di esecuzione e del non sovraccarico del sistema centrale, uniti a considerazioni sull'hardware disponibile, fanno intravedere, come ottimale, una configurazione di stazione formata da un terminale connesso al VM370, un terminale video per la presentazione dei dati, e di un minicalcolatore che svolge sia le funzioni

di interfaccia tra il video e l'elaboratore centrale, sia funzioni di elaborazioni autonome quando possibile.

L'elaborazione di grosse quantità di dati pone problemi anche dal punto di vista economico. Non ci si può permettere di perdere risultati che sono costati vari minuti di tempo di calcolo, per un errore di programma. Oltre quindi alla protezione dei dati in memoria e alla verifica dei comandi e del loro formato, esiste anche la possibilità per CIPS di prendere il controllo nel momento in cui si verifica un'interruzione. A questo punto, sfruttando anche una serie di informazioni immagazzinate nell'area "GP", si è in grado tramite una semplice procedura di recuperare la situazione al momento precedente l'esecuzione della routine che ha causato l'interruzione.

GESTIONE DEI RAPPORTI COL RESTO DELL'AMBIENTE

Esiste inoltre una serie di possibilita' che permettono ad un utente di usare CIPS con particolari facilitazioni. Vediamone alcune. La gestione di un archivio di immagini e' un problema difficilmente definibile a priori. Nel caso di immagini ottenute con tecniche di telerilevamento (si fa riferimento in particolare ad immagini Landsat), i dati vanno necessariamente tenuti su un nastro magnetico; e' tuttavia utile avere a disposizione un file ad accesso veloce (su disco magnetico) in cui, di ogni immagine, sono definite, per esempio le caratteristiche geografiche e quelle atmosferiche al momento della ripresa.

Questo archivio di descrizioni puo' essere organizzato in modo tale che la richiesta di studio di un zona porti all'elenco delle immagini disponibili e al successivo montaggio del nastro senza richiesta esplicita da parte dell'utente.

La presentazione di tali immagini puo' essere inoltre effettuata usando reticoli che permettano rapidamente di raggiungere la zona desiderata all'interno del "frame".

Per quanto riguarda la gestione di serie di documenti (per es. descrizione di utilizzo del sistema e delle sue

routine applicative, documentazione su tecniche di analisi dei dati, ecc.) e' ragionevole pensare che tali documenti siano accessibili secondo criteri standard (per es. STAIRS).

Poiche' tali funzioni sono svolte in ambiente VM/370 da macchine virtuali e cio' dedicate, si puo' pensare ad un'interfaccia tra tali macchine virtuali e quella su cui opera CIPS. Un'interfaccia di questo genere potrebbe ad esempio sfruttare alcune tecniche di scambio di aree di memoria tra macchine virtuali (VMCF).

CONCLUSIONI

E' stata studiata, a grandi linee, la struttura di un sistema per elaborazione di dati (generalmente matrici a varie dimensioni). Durante tale studio e' stata tenuta presente l'esigenza di avere uno strumento sicuro, comodo, flessibile, e facilmente implementabile.

La realizzazione di un tale sistema e' possibile anche per gradi, risolvendo prima alcuni problemi che rendono CIPS operativo, e raffinando in una seconda fase le varie tecniche.

L'ambiente operativo in cui tale sistema presenta maggiori vantaggi ci sembra essere il sistema VM/CMS sotto il quale e' in fase di implementazione.

STAMPATO PRESSO IL
SERVIZIO TECNOGRAFICO
DEL CNUCE