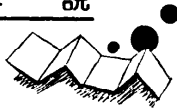


解説



パターン処理研究用共同利用システム  
EPICS の教訓†

川合 英俊††

はじめに

この研究システムのねらいは、複数の研究グループがそれぞれ専有する計算機を互いにつないで、半ば独立性を保ちながら研究成果の相乗的相互利用を行って研究を強力に進める場とその手段を提供することであった。計画から6年を経過した現在、おおむねそのねらいは達成したかに見え、特にパターンプロジェクト・プロトタイプの実現という使命を果たした。しかし、よいシステムのあり方としては今だに沢山の課題をばらんでいるので、それらの点を省みて今後の発展に向けて大方の批判を仰ぎたい。

1. システム構成

構成上の特徴は、中央にTSサービスを行う大形機において、周辺に個々の研究分野に応じた中小形機を配し、これらを星形にチャンネル結合した集中制御方式の異機種構内網としたことにある。

1.1 システム構成の拡充

3年の開発段階を終えた昭和50年の報告<sup>1)</sup>から以後の主な拡充を挙げると以下の通りである。

- 1) 中央を処理装置2台のSMS\*とした。従来の機器を片方(Bシステム)だけに接続し、もう片方(Tシステム)をTSサービス本位に運用することとした。各端末にはマニュアルスイッチがあってユーザはどちらをも選択できる。
- 2) 端末回線をミニコンやマイコンにも開放して、TS端末に音声、図形、漢字、バブルメモリなどの特殊な入出力機能を追加した。
- 3) 周辺の計算機のいくつかがリプレースされた。

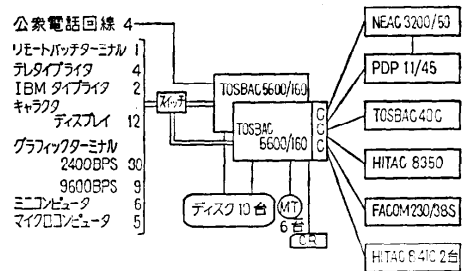


図-1 EPICS 2 構成図 (1979年1月)

- 4) 周辺計算機の構成を若干拡充した。
- 5) ソフトウェアの拡充も主として中央で行われた。TS機能の拡充のほか、LISPのあいつぐバージョンアップ、XPLの新設、標準画像形式パッケージ、漢字文編集ルーチンなどがその主なものである。

1.2 高速網 CCC システム

報告<sup>1),2)</sup>にあるとおり、中央と周辺とは250KB/Sの高速網CCCシステムで接続し、データファイルの転送とか遠隔プログラムとの会話などのハイレベルプロトコル処理管理プログラムをおいた。その後FORTRANサブルーチンやデータ転送ルーチンPCOM\*\*などに部分的に形を変えただけで、本質的には機能を拡張していない。

1.3 低速網端末回線の増設

構内回線は数次にわたって増設され、現在設置回線数は、公衆電話回線(300bps)4、タイプライタ用(300bps)10、グラフィック端末用(2,400bps)50、同(9,600bps)30に達し、年に2度程その10%に設置変更があるため、その管理は困難であった。端末制御プログラムに対して数多くの変更追加要求があいついだ。

2. システムデータの収集と利用形態

研究システムの負荷は定常業務でないので、業務に合わせて構成を最適化したり運用をはかったりすることが困難で、経年変化に合わせて調節するため常時システムデータを収集する必要があった。

† Retrospection on a Research System EPICS by Hidetoshi KAWAI (Information Sciences Division, Electrotechnical Laboratory).

†† 電子技術総合研究所パターン情報部

\* SMS: Shared Mass Storage. TOSBAC-5600の用語。主記憶でなく補助記憶を共用するマルチプロセッサシステムのこと。

\*\* 転送用データブロックのサイズを元の400Bから2KBに拡張した高速データ転送専用のユーティリティプログラム。

2.1 測定方法

測定項目は、処理量の経年変化のように長い年月にわたる息の長いものから、システム定数の調整やプログラム（システムおよびユーザの）バグを発見するための数秒ごとというこまかいものまで非常に多くの種類があるところに問題がある。こまかい測定のデータから大まかな動向を抽出する方法を用いた。

a. システムモニタ：VIDEO\*放送はテレビ有線放送で、システム負荷やプログラム進行状態を各端末の付近にあるモニタテレビに表示している。このうち毎分1回の抽出データは全コマンドの約25%を捕捉して表-1のような分布\*\*をよく表わす。

表-1 RUN コマンドの分布 (1978年11月Tシステム)

CPU 時間 (2*10 <sup>-4</sup> 時間)	コマンド数	コアサイズ ×4 KB	コマンド数
n			
16~	8	70~	310
~16	40	~70	274
~14	202	~60	203
~12	436	~50	361
~10	559	~40	392
~8	837	~30	564
~6	525	~20	782
~4	226	~10	405
~2	186		
~0	292		

表-2 TS コマンドの分布 (1978年11月下旬 Tシステムの日平均)

	数	CPU 時間
全コマンド	3,478	3.8(H)
内訳 (%)		
システムコマンド	9.9	0.1
ファイル操作	45.9	2.1
その他	21.5	0.5
RUN コマンド	22.5	97.1
内訳 (%)		
リモートバッチ	2.1	0.1
BASIC	0.4	0.1
FORTRAN ゴー	24.0	62.7
同上コンパイルゴー	31.5	7.6
WESTRAN	0.5	0.0
XPL	0.2	0.0
PASCAL	0.2	0.0
LISP	41.1	29.3

\* 5秒ごとに割込んで、各ユーザのプログラムサイズ、CPU 時間などを表しテレビ放送するプログラム。  
 \*\* 表-1 の n=8 は約9秒つまり約一千万命令にあたり '意味のある小実験' という心理的直観に対応している。  
 \*\*\* ミニコン PDP 11/40 を端末回線につないでいる。  
 \*\*\*\* 表-3 でいう資源消費量とは、各種のリソース使用量を加えたものであるが、100KB のプログラムを CPU 時間にして約9秒（一千万命令）走らせるとほぼ1単位を消費する。これはちょうど心理的直観の一仕事にあたっている。

b. アニメーション：測定用端末\*\*\*によるシステム監視システムは、刻々と変わるシステム負荷の様子を即時的かつインタラティブにグラフ表示する。

c. TS 集計：毎日サービス終了時にコマンド種別集計を記録している。例は表-2 の通り。日変化の例として TS プログラム全体のコアサイズを図-2 に示した。

d. 月集計：バッチ処理についてはアカウントルーチンが完備して、TS 端末からの起動されたものを含めてよく集計できる。中央と周辺との間の計算機間通信についてはハイレベルプロトコルのシーケンス別頻度を集計している。年間変動の例を図-3 に示した。

2.2 利用状態

a. オンラインファイルの占有状態

昭和53年12月現在での登録ユーザ数は163で、内訳は表-3\*\*\*\*のおりである。大きなファイルは数が少なく、また激しく計算機を使う人の人数も少ない。少数の者に資源が集中していることがよくわかる。

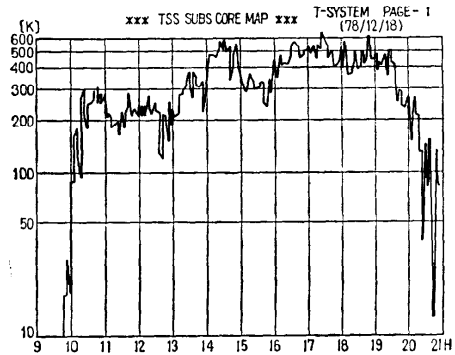


図-2 TS プログラム全量の日変化 (×4 B)

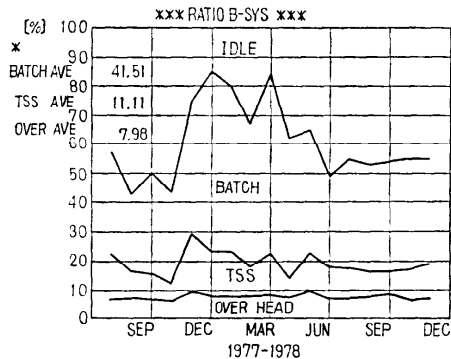


図-3 月平均 CPU 時間の年間変動 (実時間比)

表-3 ファイル占有の状態 (昭和53年12月)

ユーザの占有量 MB	ユーザ数 計 160	過去一年間の 資源消費量 $\times 10^4$	ユーザ数 計 164
51.2~		16~	0
~51.2	1	~16	1
~25.6	12	~8	11
~12.8	12	~4	6
~6.4	16	~2	10
~3.2	121	~1	136

53年12月末のファイル数は17,605個で、カタログの数は1,045であった。平均の呼出回数は74.3回で、80%の呼出しは多い方から20%のファイルに集中して、80-20則に合致している\*。ファイルの占有量は、昭和47年10月からスタートして53年12月に832MBとはぼ一直線に増加している。

カタログの深さなどのファイルシステム構造や、ファイルの寿命については調査中であるが、少数の型に分類できるものと多くのユーザが推測している。

#### b. BシステムとTシステム

それぞれバッチ本位、TSS本位に運用されている。その差はディスクを除きCCCを含む入出力機器はすべてBに接続している点と、主記憶のTS領域の上限をBで小さく(400KB)Tで大きく(880KB)とっている点、CPUのディスパッチ優先権がBではバッチジョブそれも入出力バウンドジョブに、TではTSプログラムにある点である。TS領域の伸縮速度は同じで、3秒に80KBずつ伸び10秒に40KBずつ縮むようにシステム定数を調整してある。計算機間通信を行うプログラムはBのバッチ領域でしか、実行することができない。

以下の数字は大雑把なので精度は10%程度と考えてほしい。日常のログオンユーザ数はBで10、Tで20程度で、毎月Bに2,700、Tに3,700のログオンがある。投入されるコマンドは月にBで50,000、Tで80,000だから、ひとつのログオンあたり約20個前後のコマンドを投入していることになる点では古い状態<sup>3)</sup>とかわりない。ログオン時刻分布はB、T共に午後2~3時に他の時刻より1.5~2倍のゆるやかな頂をもった山をみせる。ログオン時間Xとログオン数Yとの間には

$$X \propto \log Y$$

の関係があり、月集計でみるかぎり、B、T共に上式からのずれは小さい。ここでXは30分単位、底は1/2である。つまり全ログオンの半数は数分以下、残

表-4 ハイレベルプロトコルシーケンスの分布

	50年	53年
月平均バッチジョブ数	6,400	3,300
そのうちのユーザプログラム数	5,000	1,500
平均コアサイズ (KB)	104	80
平均CPU時間 (秒)	30.5	92.8
月平均EPICSジョブ数	2,000	31
EPICSジョブの内訳 (%)		
接続の確認	4	5
リモートバッチ	2	3
逆リモートバッチ	51	15
データファイル転送	41	75
計算機の連動	2	2

りの1/2が30分以上なのである。

表-2にも示したように全コマンドの22.5%(48年は9%)がプログラムランであり、その6%(Tでは2.1%)がバッチジョブの起動であり、さらにその2%(月30個に相当)がいわゆるEPICSジョブの実行である。EPICSジョブとは計算機間通信のうちハイレベルプロトコルを用いるもので、PCOMなどのユーティリティは含んでいない。

#### c. 計算機間通信処理

昭和53年後半の6ヵ月に188個のEPICSジョブが観測された。50年には毎月2,000個が観測されていたので、両方の内訳を表-4に示した。比率はデータファイルの転送が激増しており、PCOMもよく日常使われていることも考え合わせると、複数の計算機を連動して実験する場面は逆に激減していることがわかる。そのような総合的連動実験はデモンストレーションに限られ、研究的な実験では実験者の介入がもっとはなはだしいのである。

### 3. EPICS の教訓

#### 3.1 転送速度

##### a. 両計算機の主記憶相互間

まず実測値は静かな(他のプログラムが無い)状態で、TOSBAC 5600からHITAC 8350へ200KB/S、逆向きが182KB/Sであった。CCCが4Bを転送するのに要する時間( $\mu s$ )は、CCC通過(13)、ケーブル遅延(1)、TOSBACコアアクセス(1)、HITACコアアクセス(1.2 $\times$ 4)の和となり、かっこの数字を仮定すると19.8でこのときの転送速度202KB/Sは実測値をよく説明する。

##### b. ユーザプログラム相互間

ハイレベルプロトコルは双方の管理プログラムによって処理されるが、384Bずつブロッキングしている

\* 電総研の宮川正弘氏の発見と検証による。

ので、Lバイト転送するとき

$$D+2E+J+N+\left(\frac{L}{384}-1\right)\left(E+R+\frac{422}{T}\right)$$

だけの時間 (ms) がかかる。ここで

D: 割込禁止の解除通知処理 (22)

E: TOSBAC の管理プログラム走行時間 (4.6)

J: 同期を待ち合わせる再行処理 (5.0)

N: HITAC の割込み処理 (38)

R: バッファ割つけオーバーヘッド(総体的に55.3)

T: コアツコアのスループット(aでいう 202)

422: ヘッダ付ブロック長

を代入すると1ブロックあたり約 60ms となって、このときの転送速度 6.4KB/S は実測値 5.2~5.3KB/S をよく説明する。

### c. 日常の測定値

周辺と中央との間を 64KB のデータが往復するのに 3~6分かかり、前項の説明の 20 倍にあたる。中央ではユーザプログラムが主記憶割り付け待ちをするからである。この時間を最悪のとき 20 分以内とするようジョブスケジューラの定数を調節している。ブロック長を大きく (2KB) して、かつこの点を改善したのが PCOM で、760KB をよく 20 秒で転送する。

### 3.2 ハイレベルプロトコルシーケンスの分布

表-3 の RUN コマンド率が昭和 48 年の 9% から徐々に増えたように、計算機の使用量は増加の一途を辿っているのに、表-4 のように EPICS ジョブは激減している。計算機間通信処理機能のユーザ数が数分の一に減ったこともあるが、実験のあり方がデモストレーションを除くと、中央での処理、周辺での入出力、中央と周辺との間のデータ転送の3つの場面に分化してきたからである。

研究分野はハイレベルプロトコルシーケンスの好みには関係がなく、データ長に関係がある。シーケンスの好みは設備の設置環境に関係している<sup>4)</sup>。PCOM が日常よく使われることと、表-4 のデータ転送比率が増加していることは、周辺が特殊入出力に単能化してきていることを表わしている。このような推移は、周辺がシングルユーザベースの OS をもっていることと、その OS のバージョンアップに伴って、計算機間通信処理管理プログラムを修正することが重荷であることによって制約を受けているのである。

周辺の装置は、その管理主体も中央から分散しており、いつでも万遍なく多くのユーザに使用可能とはなっていない。したがってユーザは利用可能になるのを

待つことなく、すべてリソースの共用は中央においてファイルを複写することを基本とすることとなる。

### 3.3 資源と負荷の偏在

どのユーザも EPICS 内の資源を万遍なく使うわけでないことは負荷がシステム内に偏在することにもなる。表-4 のほかに次のような例もある。30 分に一回づつサンプルして端末の使用頻度分布を調べてみると、約 60 台のうち 10 台に使用頻度の 1/2 が集中している。表-5 は、主だった CPU 時間消費が少数のプログラム実行に偏っていることを示している。

特に 2,000 秒 (2 ギガ命令) 以上の処理を行うのは、シミュレーションやダイナミックプログラミング手法を用いて大量のデータについてアルゴリズムを検証しているほんの数名の者に限られる。面白いことに、そのような者が憑かれたように計算機にしがみつくとのは高々一週間から 10 日位しか持続しない。表-5 でジョブ当りの強制スワップ回数がプログラム走行時間に比例しているのは納得しやすいことである。

### 3.4 開発の手順

計算機通信プログラムの開発の苦しさは、その報告<sup>2)</sup>の息の長い文章がせつせつと語るとおりである。

ほかに数十万円から数千万円のソフトウェア開発が毎年行われた。その成否の鍵は開発者と試用者の連繋にあり、設計のねらいは大胆で試用はねばり強くなければならない。代表的な成功例には記号処理用の LISP<sup>5)</sup>、漢字処理システム、マイクロコン用のラインローダ PULCOM などがある。共用されるプログラムはそのコメント文に開発者の思いやりと愛情がこめられていることが不可欠の要件である。そのためにはシステムの運用が予算的にも安定していなければならない。ユーザは長期的に安定なシステムでなければ愛着をもつことができないからである。

表-5 長いプログラムの数 (53 年 11 月 T システム)

CPU 時間 (2* 秒) n	実行度数	強制スワップ回数
12~	2	368
11~	5	1,676
10~	24	2,517
9~	25	1,510
8~	85	2,087
7~	186	2,269
6~	263	1,616
5~	344	1,172
4~	526	1,013
3~	606	592
2~	996	523
1~	1,483	255
0~	1,430	170

### 3.5 処理の個人化

システムの構成要素が分散していることは、各要素が他の要素を共用し合うことでもある。連動実験の経験はプロトタイプに成果利用されたが、EPICS でひとつの周辺のユーザが他人の管理している別の周辺要素を使う機能があるのに、日常これを用いた実験が数多く見うけられないのは、当初の見込みが外れた一例である。これは周辺において、個人的な資源と公共的な資源を区別しないことから起こったことと思われる。そのような OS の保守が困難だったからであろう。つまり周辺の処理は個人化されているのである。

中央のファイルや印刷機は複数の周辺システムや端末によって共用されており、多重使用環境に適応する OS も備わっている。中央ファイル経由の間接的連動によってでも共同利用が促進されると研究は質的にも高度化しよく推進される。障害を避けて直接的連動を進めるには、資源は個人的用途のものとは公共向けのものとに分けて管理されていなければならない。したがって、他人の計算機の資源を使える途があるだけでは不十分で、ユーザは公共的な資源を手元にかき集め、処理中はこれらを個人化 (personalize) できなければならない。EPICS ではこの点での配慮に欠けていたため、ニュース 100 通、メモ 50 通というドキュメントの精力的配布にも拘わらず、資源に公共性をもたせるための管理プログラムが完備していないことが障害になり、大規模な連動実験へ進展しなかった。大規模な連動実験はパターンプロジェクト・プロトタイプによるデモンストレーションとして実が結ばれようとしている<sup>6)</sup>。

逆に、個人的な処理はその計算機を独占し勝ちのため、これを端末規模に限ればシステムに障害を起さない。こうして、端末の知能化と相まって個人処理用の計算機が端末に位置するようになる。EPICS でも、バッチベースの高速計算機間通信は減少して、インタラクションベースの低速端末通信を備えたマイコンやミニコンがとみに数を増して来た。代表例にはパターンプロジェクトで開発されたマイコン PULMIC によるものがあり、すでに和文処理用、文字読取用、LISP マシン用、バブルメモリ用データベースマシンなどが具体化されつつある。一方、個人用計算機は機器構成に制約が多いことから、端末通信の高速化が望まれ、新しく構内高速網が計画されるに至った。

### おわりに

EPICS は一面ではその使命をすでに果してリングバスによるパターンプロジェクト・プロトタイプに成果利用された<sup>6),7)</sup>。研究システムでは業務が定常でないにも拘わらず、ボトルネックを解析することはよく行われるのに、経年変化を多角的にとらえることがおろそかになり勝ちである。ここに散逸しかけていた不十分なデータをもとに分析を試みた。要は機能の実現よりもむしろ公共的資源の確立とその運用管理主体の明確化にあるといえよう。

経年変化のデータを保管することは容易でないもので、このような反省の機会を与えられたことは幸運である。分散システムが充分機能するためには、ユーザから独立した公共資源の運用責任体制ができていて、システム提供者と末端ユーザとの両側からその責任を問われるという環境が大事である。システム要素間に通信経路を構築することよりも、システム要素の自立的な成長に適応してこの経路を有効ならしめるよう保守し、かつ新しい知識をユーザに吸収させるところに分散システムの困難さがある。最後に、ここにふれたデータの収集には多くの方々の協力があつたことを特に記して謝意を表わす。

### 参考文献

- 1) 川合：パターン情報処理システム EPICS, 情報処理, 16-2, 昭和50年2月。
- 2) 川合ほか：パターン情報処理研究システム——計算機複合体 EPICS——, PIPS-R-No. 7, EPICS-15-G, 電総研, 昭和49年11月。
- 3) 川合, 岸田：EPICS-TSS の実態, 第15回情報処学会大会 304, 昭和49年12月。
- 4) Kawai, H. and Fujimasa, A.: Use of High Level Protocol in EPICS, Trans. of IECE. E 59-6, June 1976.
- 5) LISP User's Manual, EPICS-5-ON-4, 電総研, Mar. 31, 1978.
- 6) Nishino, H.: PIPS PROJECT—BACKGROUND AND OUTLINE—Proc. of 4th IJCP, Nov. 1978.
- 7) Okuda, N. Kunikyo, T. and Kaji, T.: Ring Century Bus—an Experimental High Speed Channel for Computer Communications, Proc. of 4th ICC, Sept. 1978.

(昭和54年1月9日受付)