

大学計算センタの運営における

ログ情報の利用

山 本 喜 一

慶應義塾大学 情報科学研究所

要旨

本論文では、大学計算センタのシステム評価の一手法について述べる。システム性能評価には、以下の三つの側面がある。

- (1) センタ管理者から見た評価。
- (2) 利用者側から見た評価。
- (3) 大学当局から見た評価。

今まで行なわれてきた評価は、特にハードウェア資源の利用率を測定し、それに基づく評価、予測を行なうことであつたが、ここでは上記三つの側面にそつた評価のための尺度をログ情報から収集し解析する方法について述べる。ログ情報はアカウントिंगのためにも使われており、一般には非常に大量のデータである。このようなデータを毎日集積し、長期間にわたる傾向を把握することはオペレーション上も不便なので、階層的なデータの縮小を備へたプログラムシステムを提案する。

1. 大学計算センタのモデル

大学における計算センタ、大学当局、利用者三者の関係は fig. 1 に示すとおりであり、計算センタの立場としては利用者からの要求に応じてサービスを提供し、大学当局に対してはシステムの稼働状況の報告、設備改善の要求を行なうことが必要である。⁽¹⁾ 利用者は計算センタのハードウェア、ソフトウェアを料金を支払って利用し、計算センタに対してサービス改善の要求を出す。一方大学当局は利用者からの要求に従つて予算を交付し利用者の研究成果を得る。また計算センタに対しては設備、人員も配備し、センタからの報告や要求に応じて運営を把握したり、予算を追加したしなければならない。以上述べたように、この三者の立場はそれぞれ異つており、それぞれの立場に応じた評価基準を持つてゐるはずである。したがつて計算センタとしてはそれぞれの立場に応じた評価のための情報を提供する必要がある。

現在使われている評価手法では、CPU、メモリなどのハードウェア資源の利用率を尺度としており、三様の立場をとるすべてに十分なものではない。利用者側からの評価基準を考えただけでも、利用可能な時間帯、利用料金、準備されているソフトウェアの種類と信頼性などの、システムの使いやすさはもうろん重要な評価尺度ではあるが、本論文では定量的に把握できる基準だけに着目して、三者それぞれに対する情報を得るための方法について述べる。このときに用ゐる方法にはハードウェアモニタ、ソフトウェ

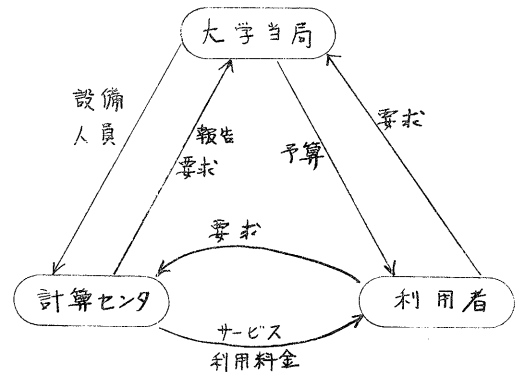


fig. 1 大学計算センタのモデル

アモニタなどの道具を使うものもあるが、最も安価で日常的に使うことができるものはログ情報である。一般に1つのランを投入してから処理を行ない出力が終了するまでの間には多量のログ情報が出力される。これらのログ情報はランの進行とともに作りだされ1つのファイルに集約されているので、1日の終わりにこのファイルから必要な情報を抽出しなければならない。既に述べたように、計算センタとしては大学当局、利用者に対しては情報を提供しなければならないが、そのためにはそれぞれが必要としている情報を知り、その情報をどのように収集するか問題となる。現在までの評価手法はほとんどのものが計算センタあるいはシステムそれ自身のための評価尺度を得るためのもので、稼働状況の把握やシステムの改善のためにこの尺度が使われている。ところが計算センタ、利用者、大学当局の三者の有機的な結合のなかで評価を行なうためには、得られた情報を従来の方法で解析するだけでなく、時間的季節的な変動を考慮した尺度を見逃すことはできない。

2. 評価基準とその尺度

大学当局、利用者、計算センタのそれぞれの立場でシステムの評価基準は異なるが、それらの基準が達成されているかどうかを判定するための尺度として、どのようなものが必要かを考察する。

(1) 大学当局

大学の評価基準は投下費用に対応するだけの使用実績があかっているかどうかであり、これを判定するためには計算センタの利用状況を把握する必要がある。そのための情報としては年間のラン総数、1日当りの平均ラン数、季節的なラン数の変動、年間ラン総数の伸びなど、処理件数に関する統計値がともに使われ、さらには計算機利用料金に換算したシステムの利用状況、すなわち総収入、1人当りの使用金額、1ラン当りの使用金額などが必要となる。またシステムが実際にどのような目的で、どのような人々に利用されているかを知るためには学部別、所属別の利用者数、利用目的別のラン数などの統計値が必要となる。

大学の評価基準として、大学当局は現在のシステムが十分に利用され、しかもその能力が適切なものであることを判定しなければならない。このために、現在のシステムに対する負荷の程度を知る必要がある。もちろんラン数も負荷の1つの尺度ではあるが、さらに実際の稼働時間に関する各種統計(年間合計時間、平均、季節変動、ピーク時など)が必要となる。また各種の資源に関する統計によってシステムの構成が適切であるかどうかを知ることができる。

大学当局がこれらの尺度を利用するのは、主として年間予算の策定のためであり、その目的に合わせて考えれば、かなり長期間のデータを累積し加工して用意すること大切である。

(2) 利用者

利用者の評価基準の中心は、システムが使い易く、サービスが良いかどうかという点である。この基準は多分に主観的なものであるが非常に重要である。その尺度としては、利用料金、いつでも使えること、ソフトウェアが完備していること、ライブラリが整備されていること、フレームに対する処理が迅速であることなどが考えられる。計算センタとしては利用者に対してできるだけ多くの情報を提供し、利用者がその情報を利用して自分の仕事のスケジュールを

決めたり、ランの処理状況を知ったりできるようにすることが必要である。これらの情報は大きく二つに分けられる。まず、利用者のランなどの程度の時間で処理されるかを知らせるためには、現在のシステムなどの程度混雑しているかを知らせれば良い。また将来の仕事のスケジュールを立てるための情報としては、過去のシステムの稼働実績に基づいた負荷の変動状況を知らせる必要がある。この変動の状況は時間単位、日単位、月単位のどのレベルかに分け、ラン数、資源の利用率などにわたる統計値が必要となる。

フレームに対する応答のためには過去のランの実行経過を保存しておき、その中から特定のランに関する記録が抽出できるようにしておけばよい。

(3) 計算センタ

計算センタ自身としてのシステム評価基準は大別すると以下の三つになる。

(a) 与えられた設備を効率よく運用しているか。(当局の要求を満足しているか)。

(b) 利用者にとってサービスを提供しているか。(利用者の要求を満足しているか)。

(c) 将来に対する計画がうまく立案されているか。

これらの基準を評価するための尺度としては(1)、(2)で既に述べたものがあるが、より詳細な統計値を収集し、解析する必要がある。特に、負荷の程度を知るためには、大学当局、利用者のために提供される比較的粗いデータばかりでなく、ボトルネックを検出するために資源利用率や資源待ち行列の解析が必要である。またシステムの信頼度を知るためにはシステムダウン、周辺装置の故障に関する統計が必要となる。さらにシステムが効率よく運転されているかどうかを知るためには、利用率、スループットなどと共に、ラン数、アクティブタスク数、バックログ数などの経時変化を記録することも必要である。

以上述べたように、それぞれの立場で評価基準は異なり、そのための尺度も基準に応じて異っている。われわれはこの尺度としてログ情報から得られた統計値を利用することを提案し、次節以降ではこれらの尺度とログ情報との関連について述べる。

3. ログ情報

2節で述べた評価尺度を得るための原始データを収集する方法には、ハードウェアモニタ、ソフトウェアモニタなどの特殊な道具が使われる。ところがこれらの道具は高価であったり、得られる原始データ量が莫大で、目的とする尺度を抽出するのに時間がかかったり、あるいは日常的に用いるにはシステムのオーバーヘッドが大きすぎる(数%から十数%増)という欠点がある。これに対して大型機のOSではシステムの状態を記録したり、それぞれのランの進行状況を知らせたり、アカウンティングのための情報を得るために、ログ情報あるいはアカウンティング情報とよぶ一連のデータをファイルに集積しているのが普通である。そこでわれわれは日常的な運用の際にも自動的に集積されているログ情報を加工、解析し評価のための尺度を得るために積極的に利用する方法について考察した。

われわれが使用しているUNIVAC 1106システム(OSはEXEC 8/E)では、ランが投入され出力が終了するまでの間に、ランの進行に伴ってfig. 2に示すよ

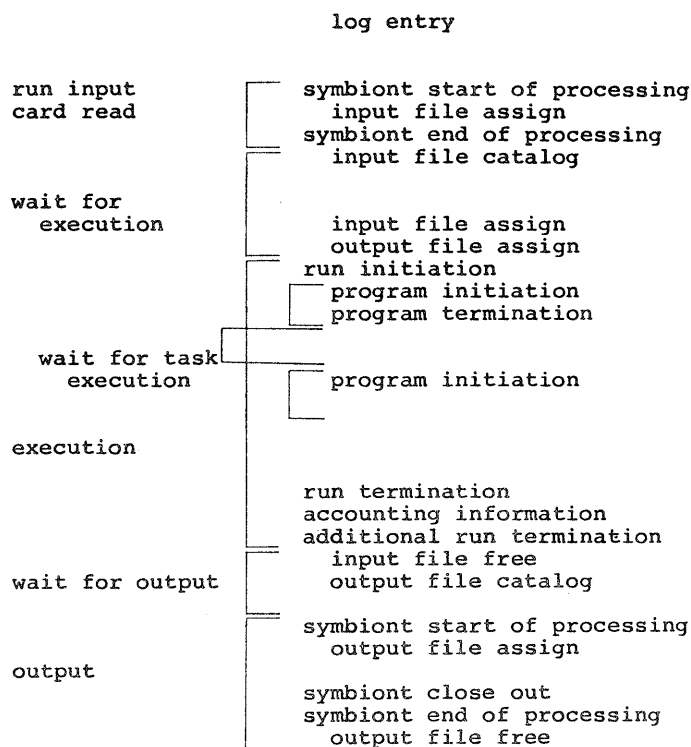


fig. 2 Run progress and corresponding log entries.

のログ情報がファイルに書き出されている。これらのログ情報は全部で24種類あり、それぞれのエントリは28語/エントリの固定長データである。ログ情報の一例を次頁のfig. 3に示す。

ログ情報のファイル(System Log File)への出力はランの進行に伴って行なわれるために、多重処理の状況ではいろいろなランの情報がファイルに入り乱れて入っている。したがって2節で述べたような尺度を得るためには、原始データとしてのSystem Log Fileを加工して必要な情報を抽出する必要がある。メーカーが提供しているシステムにはIBMのSMF(System Management Facility)およびSMF editor, ACOSのACE(Ac-

counting Cafeteria Management Program)/MDC(Multi-display Control Program), FACOMのSMF, UNIVACのQUOTAなどのアカウンティング情報およびログ情報の収集編集システムや、ソフトウェアモニタとしてIBMのRMF(Resource Management Facility), UNIVACのSIP(System Improving Program)などがあるが、ソフトウェアモニタ以外のプログラムでは、主としてシステムの稼働状況を記録することかその目的であり、ログ情報に基づいて総ラン数や平均CPU時間と求めるというような静的な解析を行なうことが主体である。ところが2節に述べたような多様な評価目的のための尺度を得るためには、このような従来の解析ばかりでなく、例えばバックログ数の時間変動を知るために、ランの入力時刻と実行開始時刻を時系列としてとらえ、その個数の差からその時刻でのバックログ数を求めるというように、ログ情報を負荷のトレースデータとみなして動的な解析を行なうことが大切である。特に大宇計算センタの場合には、ランの入力を制御することかほとんど不可能なため、負荷の状態の変化が激しく短期的に見た場合には1日単位の平均的な尺度ではあまり意味を持たないことが多く、既製のログ解析プログラムでは能力が不足している。したがって時間的な変動を抽出できるような解析プログラムが必要となる。また月、学期、年というような長期的な視野に立った場合には、ラン単位に編集したログ情報を扱うにしても莫大な量となり、直接これ进行处理することは事実上不可能である。したがって日、週、月というような単位で原始データを縮小し、長期的な解析に必要なデータを抽出するようなシステムも必要と

な っ て く る 。

0	4	24	
1			
2			program name
3			
4			version name
5			
6			program termination time/date*
7			cpu time (200μs)
8			0
9			ER/CC charge (200μs)
10			
11			I/O transfer word count
12			for I/O device group
13			1 to 10
14			
15			
16			
17			
18			
19			SUPs (200μs)
20			coreblock SUPs
21			voluntary delay time (200μs)
22			realtime mode time
23			reentry address
24			condition word
25			log entry time/date*
26			
27			runid

* indicates this word has the following format.

mm	dd	yy	second
----	----	----	--------

SUPs (Standard Unit of Processing)

$SUPs = RTC + CER + CXCC + \Sigma (IOG + LT) * FG$

RTC : CPU time

CER : charge for ER request

CXCC : charge for control card

IOG : I/O transfer word count

LT : f(device access time)

FG : conversion factor

(depend on transfer rate)

fig. 3 An example of a log entry
(program termination log)

4. データ解析のためのプログラムシステムと出力項目

われわれの提案するシステムは計算センタ，利用者，大学当局のそれぞれの立場に立って適切な評価が下せるような尺度とできるだけの努力で提供することを目指す。そのために尺度を得る単位を日，週，月（暦月ではなく4週を1月とする），年の4つに分け，それぞれの単位ごとにデータの縮小を行ないオペレーションが早くかつ容易に行なえるようになっていく。データの縮小の程度は1日平均600ラン，1ラン当りの平均タスク数を10個として，System Log Fileは約22,200セクタ（1セクタ=28語）の大きさである。これを集約プログラム（collection, sequencing and compaction）を適当に用いることにより約3,500セクタに縮小する。（1日分のログファイルが1200枚のテープ1本だとすると，同じテープに1週間分の情報を蓄えられる。）この縮小ファイル（daily log）を用いて1日単位の要約情報を約1K語にまとめこれを要約ファイル（summary file）に保存する。1週間分の縮小ファイルと要約ファイルを用いて週間集計（weekly summary）を行ないやはり約1K語の週間要約情報（weekly summary file）にまとめる。4週分の週間要約情報

報から月間要約情報 (monthly summary file) を作り、これを12ヵ月分まとめたものを用いて年間要約情報 (annual summary file) を作成する。

このシステムの概要は fig. 4 に示すとおりであり、この中で最も重要な部分は集約プログラムによって作成する縮小ファイルと日報集計 (daily) によって作られる要約情報である。縮小ファイルの内部構成は fig. 5 (a) から fig. 5 (e) に示すとおりで、ランゴとにまとめてありしかも到着時刻順に並べてある。このファイルには1日当り3009ランまで収容可能で実用上十分な大きさを持っている。

日報集計処理のために必要なCPU時間は、

ラン総数が約600個のとき、System Log file の収集と並びかえに約5分、縮小ファイルへの集約に約30秒、要約情報の出力および日報の作成に約2秒を要し、毎日の処理のターンアラウンドは約30分である。

日報として出力する項目は使用形態の相違から、バッチ (リモートバッチを含む)、優先度Dのバッチ (出力は翌日渡し)、デマンドの3種に大別しておりその内容は table 1 に示すとおりである。日報のうちランの個数に関する要約情報の出力例を fig. 6 に、バックログ数の時間変動の出力例を fig. 7 に示す。週報、月報に関しても基本的には table 1 の内容と同じものについての要約情報を出力するか日報ほどの細かさは必要ないので fig. 8 に示すような形式をとる。また変動に関しては詳細なものよりは1週間を単位として見たときに有効な情報が得られるように考慮している。たとえばバックログ数の変動の場合には1時間当りの平均バックログ数をおめ fig. 9 に示すように印刷している。

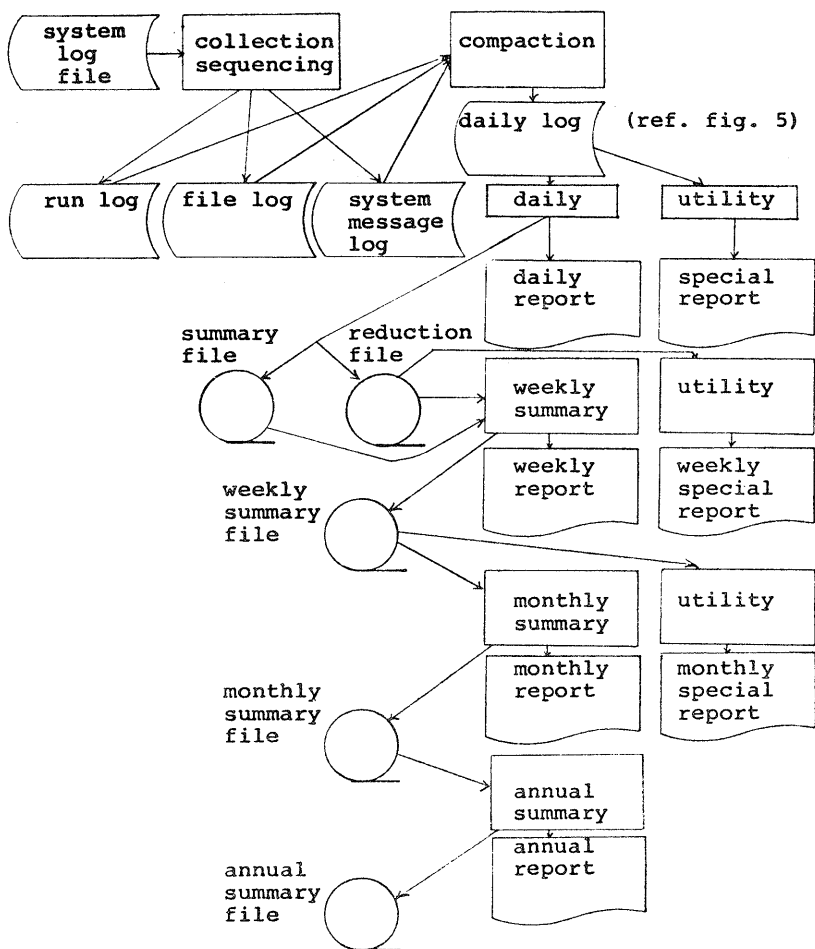


fig. 4 Program and file system for log analysis

- table 1
Items on the daily report
1. Runs
 - 1.1 Number
 - number of runs processed
 - printed pages
 - read cards
 - number of tasks
 - used tape units
 - I/O transfer word count
 - 1.2 Time
 - CPU time / run (/task)
 - turnaround time
 - waiting time
 - for start of execution
 - for output printing
 - for task initiation
 2. Resource utilization
 - CPU, core, printer, card reader,
 - tape unit, disc/drum,
 - demand terminal
 3. Variation
 - number of runs
 - input, active, backlogged
 - turnaround time
 - resource utilization
 - CPU, core, disc/drum

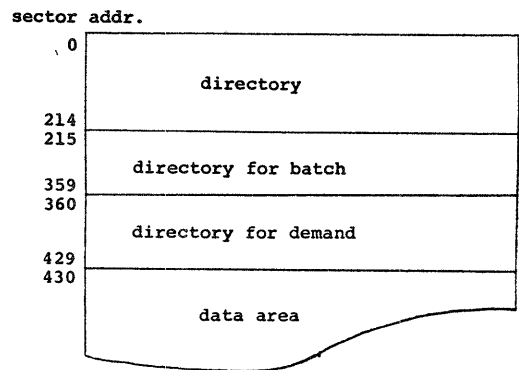


fig. 5 (a) layout of the daily log file

sector addr.	
0	header
	month day year
	number of runs number of sectors
	start sector addr. number of sectors
	number of words
	.
	.
214	run-1
215	

fig. 5 (b) contents of the directory

word count

1	runid	
2	type	priority
3	number of tasks	number of assignments
4	start of execution	end of execution
5	arrival time	departure time
6	number of input cards	coreblock SUPs
7	estimated CPU SUPs	actual CPU SUPs
8	output time	actual CPU time
9	estimated print pages	actual printed pages
10	input site	
11	block-1 I/O transfer word count[drum]	
12	block-2 I/O transfer word count	
13	block-3 I/O transfer word count[disc]	
14	block-4 I/O transfer word count[tape]	
15	user number	
16		project number
17	charge	
18	used disc track	
19	used drum track	
20	not used now	

fig. 5 (c) run table

5. 出力項目と評価尺度との関連

大学当局に対しては基本的には年報作成プログラムによって評価尺度を出力する。処理件数に関する統計及び資源の利用率に関する年間累計、季節変動などについては本プログラムシステムの流に乗っており問題なく処理できる。一方所属別、利用目的別の利用状況および利用料金に関する処理は現在のところシステムに含まれていないのぞ別途考慮する必要がある。ただし必要な情報は縮小ファイル (daily) のランテーブルに含まれているので単に累積をとる操作を追加すればよい。

利用者の評価尺度の一つである現在のシステムの混雑の程度を知らせることはできないが、前日の日報あるいは前

の週の週報によって現状を予測してもらうことはできる。一方将来の仕事のスケジュールの発注では各種の変動に関する出力によって全体的な傾向をかなり適確に把握してもらうことができている。またクレーンに対する処理のためには、縮小ファイルがランゴとにまとめてあり、しかも到着時刻順に並べてあるので対象とするランの全履歴を簡単に出力できる。

計算センタ自身としては稼働状況を把握するためにほぼ十分な尺度が得られている。特に変動に関する統計は負荷の変動に敏感に反応しており単なる平均値や合計よりもシステムの状況を明確に反映している。

6. おわりに

大学計算センタでシステム評価のためにログ情報を利用して得られる評価尺度と、計算センタ、利用者、大学当局の三つの立場に応じた評価基準とその尺度との関係を整理し、その尺度を得るためのシステムについて述べた。現在までに完成しているのは週間集計プログラムまでの部分であるが、今後の作業にさらばファイルがはるかに小さくなり、操作が単純になりしかも高速になった。

[謝辞]

本研究の最初から適切な助言をいただいた工学部管理工学科の浦昭二教授に感謝いたします。またシステムの製作は当研究所の白石幸男氏および工学部管理工学科の八友田和幸、松林毅両君の大きな努力のたまものでありここに感謝いたします。

[参考文献]

(1) Donald N. Streeter, 'The Scientific Process and the Computer', John & Wiley, 1974.

word count

1	task name	
2		
3	actual CPU time	elapsed time
4	actual CPU SUPs	coreblock SUPs
5	initiation time	termination time
6	wait time	CPU utilization
7	block-1 I/O transfer word count[drum]	
8	block-2 I/O transfer word count	
9	block-3 I/O transfer word count[disc]	
10	block-4 I/O transfer word count[tape]	

fig. 5 (d) task table

word count

1	file name	
2		
3	type	granule count

fig. 5 (e) assignment table

***** DAILY REPORT *****
 DATE : MON. 4/16/79
 * NUMBER OF BACK-LOG RUNS *
 * * * * *

 DATE 4/19/79
 NUMBER OF RUNS 527
 NUMBER OF BATCH-RUN 451
 * 1-1

RUNS	TOTAL	RLT.RUNS	AVERAGE	AVERAGE*	STD.DEV.	STD.DEV.*	MAX
BATCH STD.	302						
BATCH D-OP.	246						
DEMAND	7						
PRINTED PAGES	3447		11.4		48.9		726
BATCH STD.	1614		6.6		10.6		106
BATCH D-OP.	512		73.1		139.3		414
DEMAND	1321		72.0		137.2		1117
READ CARDS	28582		94.7		172.6		1117
BATCH STD.	20553		83.5		157.9		972
BATCH D-OP.	2016		288.0		172.1		431
DEMAND	6011		222.7		206.7		1117
TASKS	2023		6.7		13.8		194
BATCH STD.	1112		4.5		3.7		41
BATCH D-OP.	47		17.6		31.1		17
DEMAND	864		7.3		7.7		194
USED TAPES	100		1.4		1.4		6
BATCH STD.	92		2.4		1.8		6
BATCH D-OP.	2		2.0		1.0		3
DEMAND	4		1.1		1.0		1
CPU-TIME (SEC.)	3281.3 (3319.5)		10.9 (1.6)		26.4 (8.0)		241.9 (177.7)
BATCH STD.	2168.9 (2168.2)		8.8 (1.9)		22.2 (8.6)		241.9 (177.7)
BATCH D-OP.	652.1 (652.1)		93.2 (13.9)		58.0 (28.0)		204.0 (99.0)
DEMAND	460.3 (499.2)		9.4 (1.6)		19.1 (2.0)		15418
TURN-AROUND (SEC.)	149602		507.1		1474.7		2745
BATCH STD.	34207		1220.4		372.1		15418
DEMAND	6548395		216.8		3146.3		22163
OPEN-WAIT (SEC.)	3836		880.6		1575.2		2276
BATCH STD.	61645		880.6		5611.1		22163
BATCH D-OP.	25972		88.0		186.8		1587
OUTPUT-WAIT (SEC.)	21730		88.3		165.3		1340
BATCH STD.	4242		86.6		266.7		1587
DEMAND							
TASK-WAIT (SEC.)	12800		6.3		19.6		265
BATCH STD.	711		2.9		2.9		63
BATCH D-OP.	22		1.5		5.5		1
DEMAND	12067		14.0		28.0		265
TASK-ELAPSE (SEC.)	106106		52.4		264.3		5411
BATCH STD.	27637		24.9		78.8		910
BATCH D-OP.	1881		40.0		68.2		344
DEMAND	76588		88.6		391.2		5411
DRUM-I/O (KW.)	28712.5		98.4		150.2		1443.8
BATCH STD.	1818.6		246		85.2		544.6
BATCH D-OP.	3073.6		175.8		153.8		598.6
DEMAND	3999.8		285		183.6		1443.8
DISC-I/O (KW.)	357946.8		1185.3		4182.6		435143.8
BATCH STD.	292157.8		229		289.2		435143.8
BATCH D-OP.	6670.4		11187.6		4552.4		435143.8
DEMAND	59118.4		952.9		4766.4		10975.9
TAPES-I/O (KW.)	109735.2		1206.5		1987.1		17520.1
BATCH STD.	103251.9		363.4		1832.8		17520.1
BATCH D-OP.	1604.1		419.7		2006.1		17520.1
DEMAND	4875.2		229.2		392.5		17520.1
			99.6		1219.8		4406.5

fig. 6 An example of the daily report

fig. 7 Variation of backlogs
 (daily)

DATES:
MON. 4/16/79
TUE. 4/17/79
WED. 4/18/79
THU. 4/19/79
FRI. 4/20/79
SAT. 4/21/79

fig. 8 An example of the weekly report

[illegible]

- 10 -