

SHARE PROGRAM LIBRARY AGENCY



PROGRAM NUMBER

052018

University of Miami

1365 MEMORIAL DRIVE - CORAL GABLES, FLORIDA
(305) - 284-6257

SPLA CONTROL NUMBER: 00243

(1) Program Number (to be filled by SPLA) 370D-05.2.018

(2) Title of Program MVS JOB SCHEDULER/DATA SET NAME ENQ CONFLICT MANAGER

(3) System Type(s) (Machine). IBM 370/303X

```
(4) Search Key(s) . . . . . (SMF EXITS) (MVS) (JES2)
                                (JOB SCHEDULING) (STEP-TERM STATISTICS)
                                (RESOURCE CONTROL)
                                (REQUEUER)
```

(5) Programming Systems/Languages ASSEMBLER (F)

(6) Primary Subject Code (SMF EXITS)

(7) Minimum System Requirements MVS/JES2

(8) New (N) or Revision (R) (if revision, show prior Program Number in Item 1) R

(9) Date of Submittal 7/7/81

(10) Documentation (number of original pages submitted) 26

(11) Author's Name and Address DAVID HALBIG
U.S. SENTATE COMPUTER CENTER
400 N. CAPITOL ST.

WASHINGTON, D.C. 20510

(12) Direct Technical Inquiries to Name & Address (if different than Author) (SAME AS ABOVE)

(13) Submitter's Installation Membership Code UCO

(14) **Abstract** (should contain sufficient information for a reader to determine the value of the program). Listed on the reverse side of this form are subjects which may serve as a guide for a descriptive abstract.

DISCLAIMER

Triangle Universities Computation Center (TUCC) serves solely as the distribution agent for contributed programs and does not test or maintain them. They are distributed essentially in the original form submitted by the author. Neither TUCC nor SHARE, INC., makes any warranty, expressed or implied, as to the documentation, function, or performance of the contributed programs.

SHARE PROGRAM LIBRARY SUBMITTAL FORM

Subject Guide:

- a. Purpose
- b. Programming Language used
- c. Version and modification level or release number
- d. Field of application
- e. Type of routine (main program, subroutine, etc.)
- f. Specific description of machine requirements

PURPOSE: To identify resource estimates of a batch job by scanning JCL and then place job into a JES job class based on an associated site-defined resource limit table. Additionally, requeues and holds batch jobs found to be in DSNAMES conflict. Once the conflict is relieved, the victim (requeued) job is automatically released. Also provides comprehensive statistics on CPU, storage, and I/O use (by DDNAME) at batch step termination.
PROGRAMMING LANGUAGE: Assembler(F); APPLICABILITY: MVS/JES2 3.7 NON-SE; TYPE OF ROUTINE: SMF EXITS (IEFUJV, IEFUJI, IEFU83, IEFUTL, IEFUSO, EIFACTRT) - The System Requires No Modifications to MVS or JES.
DISCLAIMER Triangle Universities Computation Center (TUCC) serves solely as the distribution agent for contributed programs and does not test or maintain them. They are distributed essentially in the original form submitted by the author. Neither TUCC nor SHARE, INC., makes any warranty, expressed or implied, as to the documentation, function, or performance of the contributed programs.
(Please attach additional pages if necessary) Total pages attached _____

An "Acknowledgement of Assistance" statement must be attached to this Submittal Form.

Permission to Publish

"I hereby give the SHARE Program Library Agency permission to reprint, reproduce, and distribute this program"

(15) Signature of Submitter and Date _____

(15) Signature of Installation Addressee _____

TAPESCAN 3.3 - GSFC TAPE ANALYSIS AND COPYING PROGRAM 12/09/80 17:19:26.8 INPUT VOL=SENATE (LAST MOD-04/05/78)

VOLUME TABLE OF CONTENTS FOR SENATE 9 TRACK SPLA

SEQ.	DATA SET NAME	RECFM	LRECL	BLKSIZE	DEN	TRTCH	MAX BLK	MIN BLK	AVG BLK	BLK COUNT	CREATED	JOB NAME	STEP	EXPIRES
0001	CP.JSC10LDD	FB	00080	06150	1600		02400	02400	02400	000001	11/14/80	CPUNLOAD	IEBGENER	00/00/98
0002	CP.JSC10JCL	VS	06176	06180	1600		06180	00060	02748	000019	11/14/80	CPUNLOAD	IEBCOPY	00/00/98
0003	CP.JSC10MAC	VS	06176	06180	1600		06180	00060	04133	000058	11/14/80	CPUNLOAD	IEBCOPY	00/00/98
0004	CP.JSC10SRC	VS	06176	06180	1600		06180	00060	05781	000128	11/14/80	CPUNLOAD	IEBCOPY	00/00/98

The distribution tape contains four partitioned data sets unloaded by IEBCOPY. These four pages show the output generated by the programs that created the tape.

SYSTEM SUPPORT UTILITIES JENIN111 PAGE

TAPE INIT--SER=SENATE,OWNER=SPLA*,DISP=REHND 00015000
VOLISENATEO SPLA

DATA SET UTILITY--GENERATE PAGE

IEB3521 WARNING : OUTPUT RECFM/LRECL/BLKSIZE COPIED FROM INPUT
PROCESSING ENDED AT EOD

IEBCOPY MESSAGES AND CONTROL STATEMENTS PAGE 0001

00040000

C 0=TAPEJCL,I=DISKJCL

IEB1671 FOLLOWING MEMBER(S) UNLOADED FROM INPUT DATA SET REFERENCED BY DISKJCL

IEB1541 ASSEMBLE HAS BEEN SUCCESSFULLY UNLOADED
IEB1541 DOCUMENT HAS BEEN SUCCESSFULLY UNLOADED
IEB1541 FORMAT HAS BEEN SUCCESSFULLY UNLOADED
IEB1541 IEALPAJ1 HAS BEEN SUCCESSFULLY UNLOADED
IEB1541 IEALPAJ2 HAS BEEN SUCCESSFULLY UNLOADED
IEB1541 IEALPAJ3 HAS BEEN SUCCESSFULLY UNLOADED
IEB1541 INIT HAS BEEN SUCCESSFULLY UNLOADED
IEB1541 JES2 HAS BEEN SUCCESSFULLY UNLOADED
IEB1541 LINKEDIT HAS BEEN SUCCESSFULLY UNLOADED
IEB1541 LINKSE2 HAS BEEN SUCCESSFULLY UNLOADED
IEB1541 LNKLS101 HAS BEEN SUCCESSFULLY UNLOADED
IEB1541 LOAD HAS BEEN SUCCESSFULLY UNLOADED
IEB1541 REQUEUER HAS BEEN SUCCESSFULLY UNLOADED
IEB1541 SMP HAS BEEN SUCCESSFULLY UNLOADED
IEB1541 UNLOAD HAS BEEN SUCCESSFULLY UNLOADED

C U=TEMPMAC,I=COPYLIB 00041000
 S M=ACTRIMAP 00042000
 S M=COMMEXIT 00043000
 S M=CPENG 00044000
 S M=CKOSSMEM 00045000
 S M=ECFOLLOW 00046000
 S M=EXITCORE 00047000
 S M=HASHDSN 00048000
 S M=JESCKFNC 00049000
 S M=JESRLEFNC 00050000
 S M=KEGECQU 00051000
 S M=KNMANCHR 00052000
 S M=SMF10TAB 00053000
 S M=SMF239 00054000
 S M=TABLE 00055000
 S M=TRINSRCH 00056000
 S M=DELETE 00057000
 S M=TRTSRCH 00058000
 S M=TSQSRCH 00059000
 S M=UJIMAP 00060000
 S M=UJVMAP 00061000
 S M=USIPARM 00062000
 S M=USOMAP 00063000
 S M=USOCORE 00064010
 S M=UTLCORE 00065020
 S M=UB3CORE 00066030

FOLLOWING MEMBER(S) COPIED FROM INPUT DATA SET REFERENCED BY COPYLIB

IEB1671
 IEB1541 ACTRIMAP HAS BEEN SUCCESSFULLY COPIED
 IEB1541 COMMEXIT HAS BEEN SUCCESSFULLY COPIED
 IEB1541 CPENG HAS BEEN SUCCESSFULLY COPIED
 IEB1541 CKOSSMEM HAS BEEN SUCCESSFULLY COPIED
 IEB1541 ECFOLLOW HAS BEEN SUCCESSFULLY COPIED
 IEB1541 EXITCORE HAS BEEN SUCCESSFULLY COPIED
 IEB1541 HASHDSN HAS BEEN SUCCESSFULLY COPIED
 IEB1541 JESCKFNC HAS BEEN SUCCESSFULLY COPIED
 IEB1541 JESRLEFNC HAS BEEN SUCCESSFULLY COPIED
 IEB1541 KEGECQU HAS BEEN SUCCESSFULLY COPIED
 IEB1541 KNMANCHR HAS BEEN SUCCESSFULLY COPIED
 IEB1541 SMF10TAB HAS BEEN SUCCESSFULLY COPIED
 IEB1541 SMF239 HAS BEEN SUCCESSFULLY COPIED
 IEB1541 TABLE HAS BEEN SUCCESSFULLY COPIED
 IEB1541 TRINSRCH HAS BEEN SUCCESSFULLY COPIED
 IEB1541 DELETE HAS BEEN SUCCESSFULLY COPIED
 IEB1541 TRTSRCH HAS BEEN SUCCESSFULLY COPIED
 IEB1541 TSQSRCH HAS BEEN SUCCESSFULLY COPIED
 IEB1541 UJIMAP HAS BEEN SUCCESSFULLY COPIED
 IEB1541 UJVMAP HAS BEEN SUCCESSFULLY COPIED
 IEB1541 USIPARM HAS BEEN SUCCESSFULLY COPIED
 IEB1541 USOCORE HAS BEEN SUCCESSFULLY COPIED
 IEB1541 USOMAP HAS BEEN SUCCESSFULLY COPIED
 IEB1541 UTLCORE HAS BEEN SUCCESSFULLY COPIED

IEB1541 U83CORE HAS BEEN SUCCESSFULLY COPIED
IEB1441 THERE ARE 000005 UNUSED TRACKS IN OUTPUT DATA SET REFERENCED BY TEMPMAC
IEB1491 THERE ARE 000008 UNUSED DIRECTORY BLOCKS IN OUTPUT DIRECTORY

00064000
00065000
00066000

C=0=TEMPMAC,I=SCIPVMAC
S M=IHADECB
S M=IEFCVRWA

IEB1671 FOLLOWING MEMBER(S) COPIED FROM INPUT DATA SET REFERENCED BY SCIPVMAC
IEB1541 IEFCVRWA HAS BEEN SUCCESSFULLY COPIED
IEB1541 IHADECB HAS BEEN SUCCESSFULLY COPIED
IEB1441 THERE ARE 000003 UNUSED TRACKS IN OUTPUT DATA SET REFERENCED BY TEMPMAC
IEB1491 THERE ARE 000006 UNUSED DIRECTORY BLOCKS IN OUTPUT DIRECTORY

00067000

C=TAPEMAC,I=TEMPMAC

IEB1671 FOLLOWING MEMBER(S) UNLOADED FROM INPUT DATA SET REFERENCED BY TEMPMAC
IEB1541 ACTRTMAP HAS BEEN SUCCESSFULLY UNLOADED
IEB1541 COMEXIT HAS BEEN SUCCESSFULLY UNLOADED
IEB1541 CPENQ HAS BEEN SUCCESSFULLY UNLOADED
IEB1541 CROSSEM HAS BEEN SUCCESSFULLY UNLOADED
IEB1541 ECFOLLOW HAS BEEN SUCCESSFULLY UNLOADED
IEB1541 EXITCORE HAS BEEN SUCCESSFULLY UNLOADED
IEB1541 HASHDSN HAS BEEN SUCCESSFULLY UNLOADED
IEB1541 IEFCVRWA HAS BEEN SUCCESSFULLY UNLOADED
IEB1541 IHADECB HAS BEEN SUCCESSFULLY UNLOADED
IEB1541 JESCKFNC HAS BEEN SUCCESSFULLY UNLOADED
IEB1541 JESRLEFNC HAS BEEN SUCCESSFULLY UNLOADED
IEB1541 JESRECORD HAS BEEN SUCCESSFULLY UNLOADED
IEB1541 REGLEQ HAS BEEN SUCCESSFULLY UNLOADED
IEB1541 RMNANGHR HAS BEEN SUCCESSFULLY UNLOADED
IEB1541 SMFIOTAB HAS BEEN SUCCESSFULLY UNLOADED
IEB1541 SMF239 HAS BEEN SUCCESSFULLY UNLOADED
IEB1541 TABLE HAS BEEN SUCCESSFULLY UNLOADED
IEB1541 TRINSRCH HAS BEEN SUCCESSFULLY UNLOADED
IEB1541 TDELETE HAS BEEN SUCCESSFULLY UNLOADED
IEB1541 TPRTSRCH HAS BEEN SUCCESSFULLY UNLOADED
IEB1541 TSECSRCH HAS BEEN SUCCESSFULLY UNLOADED
IEB1541 UJIMAP HAS BEEN SUCCESSFULLY UNLOADED
IEB1541 UJVMAP HAS BEEN SUCCESSFULLY UNLOADED
IEB1541 USIPARM HAS BEEN SUCCESSFULLY UNLOADED
IEB1541 USOCORE HAS BEEN SUCCESSFULLY UNLOADED
IEB1541 USUMAP HAS BEEN SUCCESSFULLY UNLOADED
IEB1541 UTLCORE HAS BEEN SUCCESSFULLY UNLOADED
IEB1541 U83CORE HAS BEEN SUCCESSFULLY UNLOADED

IEBCOPY MESSAGES AND CONTROL STATEMENTS

C O=TAPE SRC,1=JSC30SRC
 S M=DOCUMENT
 S M=IEFUJV
 S M=IEFUJ1
 S M=IEFUS1
 S M=IEFU83
 S M=IEFUTL
 S M=IEFUS0
 S M=IEFACTRT
 S M=IEQUEUER
 S M=IEQUEINIT
 S M=IEQUETEST

DATA SET REFERENCED BY JSC30SRC

IEB1671 FOLLOWING MEMBER(S) UNLOADED FROM INPUT
 IEB1541 DOCUMENT HAS BEEN SUCCESSFULLY UNLOADED
 IEB1541 IEFACTRT HAS BEEN SUCCESSFULLY UNLOADED
 IEB1541 IEFUJ1 HAS BEEN SUCCESSFULLY UNLOADED
 IEB1541 IEFUJV HAS BEEN SUCCESSFULLY UNLOADED
 IEB1541 IEFUS1 HAS BEEN SUCCESSFULLY UNLOADED
 IEB1541 IEFUS0 HAS BEEN SUCCESSFULLY UNLOADED
 IEB1541 IEFUTL HAS BEEN SUCCESSFULLY UNLOADED
 IEB1541 IEFU83 HAS BEEN SUCCESSFULLY UNLOADED
 IEB1541 IEQUEUER HAS BEEN SUCCESSFULLY UNLOADED
 IEB1541 IEQUEINIT HAS BEEN SUCCESSFULLY UNLOADED
 IEB1541 IEQUETEST HAS BEEN SUCCESSFULLY UNLOADED
 IEB1471 END OF JOB -00 WAS HIGHEST SEVERITY CODE

IEB1471

MVS JOB SCHEDULER

Both throughput optimization and turnaround guarantees require accurate and enforceable resource estimates. To accomplish this, a system has been developed which relies exclusively on user JCL to determine the resource estimate. JCL inherently defines such things as:

- o JOB CPU time limit
- o max # of tape drives used per step
- o max amount of temporary DASD space used per step
- o max amount of virtual storage

For throughput optimization, these estimates are used to schedule a system-wide workload which uses each resource (memory, CPU, I/O path and devices) to its fullest, without over-committing that resource.

Likewise, in making a turnaround time guarantee, the system elicits a promise, in terms of an upper limit on resource use, from the job.

For example, if a batch job exceeds its estimate of DASD space, it will be ABENDED with a D37 or B37. Likewise, if a job exceeds its CPU time estimate, it will be ABENDED with a 322:

The discussion which follows describes what characteristics of a job are interrogated by the scheduler and how these characteristics are to be understood.

CPU time - the CPU time is obtained from one of two places: the TIME= keyword on the JOB card or the TIME= keyword (or its default) on each EXEC statement. If the CPU time is specified via the TIME= keyword on the JOB card, it takes precedence over the SUM of CPU time from all the EXEC statements. In cases where no TIME= keyword is on the JOB card, an estimate of CPU time is made by adding all STEP-LEVEL CPU times. The default STEP-LEVEL CPU time is 8 hours, at our site so specify the TIME= keyword on your job card.

Virtual Storage - the estimated virtual storage use is found in one of two places: the REGION= keyword on the JOB card or the highest REGION= value found on any step. If no value is specified, the default is 640K at our site.

If a REGION= value is specified on the JOB card, it establishes an upper limit on the virtual storage any STEP may use, even if a step specifies a higher value.

Print Lines - the estimated number of output print lines is obtained from the /*JOBPARM L= keyword. The value specified, such as L=10, represents the number of 1,000's of lines, or 10,000 lines in this case. If no value is specified, the default is L=70 or 70,000 lines.

Tape drive usage - the maximum number of drives used in any step is obtained from JCL DD statements. The maximum number of drives is broken into two categories:

6250/1600 (UNIT = TAPE)
1600/800 (UNIT = DUALDEN)

The routine which obtains this number will correctly identify the following tape-use cases:

- o tapes passed from one step to another (tape will be correctly counted in the receiving step)
- o cataloged tape data set (OS catalog will be searched to identify the device type of the file)
- o multiple devices per DD statement, such as UNIT=(TAPE,2)
- o UNIT=AFF unit affinity use of drives is not counted against the total drive use for that step
- o VOL=REF= backward references permit identification of the device type of the referred - to file
- o explicit use of UNIT=TAPE, UNIT=DUALDEN, and all compatible esoteric and generic names

DASD Space - the maximum allocation of DASD space in a step is obtained from the JCL DD statements. Each SPACE= request in a job's JCL stream is converted from the unit of allocation (i.e. CYL, TRK, ABSTR, aveblk) into an equivalent number of 1,000-byte blocks. This normalizing operation permits easy comparison of space requests regardless of device type (3330-11, 3350) and type of allocation (CYL, TRK, ABSTR, aveblk). The routines which obtain the space estimate correctly identify the following space - use cases:

- o DASD space passed from one step to another - the space passed into the step is counted against the space use for that step.

- o Secondary allocations - the value given as the secondary allocation quantity is doubled and added to the primary - this mechanism accounts for the potential use of extents by the DASD file.
- o Explicit use of UNIT=SYSWK, UNIT=SYSDA, and all compatible esoteric and generic names

VIO space - the maximum allocation of VIO space in a step is obtained from the JCL DD statements. The handling of SPACE= requests is identical to that for real DASD, given above.

Teleprocessing devices - the maximum number of teleprocessing devices used in any one step is obtained by scanning all DD statements in a step. The routines look for the esoteric names UNIT=TELE2 and UNIT=BSC3 and compatible esoteric and generic names

Direct unit assignment - the maximum number of direct unit assignments (e.g. UNIT = 381) made in any one step is obtained by scanning all DD statements in a step.

Unit Record Device assignment - the maximum number of unit record devices assigned in any one step is obtained by scanning all DD statements, e.g. UNIT=SYSRDR

Job Priority - the job selection priority of a job is obtained from the PRTY= keyword on the job card. If no value is specified, the priority is 0.

Performance Group - In MVS, all work is assigned a performance group, dictating how that job is to receive resources once the job is initiated. The performance group specification is given as a number 1 thru 9. There is no correlation between a high number and high performance. The numbering system is site-defined. The performance group specification is obtained from the PERFORM= keyword on the JOB card, or from the highest PERFORM= value found on any step.

After the Scheduler identifies all the resources to be used by a job, it then places the job into an appropriate job class (see the job class limit table example in exhibit # 3). This capability for placing jobs into job classes means the scheduler can be implemented and the class structure changed at any time without changes to user JCL. As important as the class limits is the initiator setup: by careful selection of classes in the initiator strings, tape drive use and turnaround guarantees can be very successfully managed. Exhibit # 4 shows our current initiator setup. Note the initiator strings change twice during the day (8:00 A.M. and 6:00 P.M.) these changes occur without operator intervention. This is made possible by the use of JES automatic commands. The IPO documentation from IBM describes how to automatically re-initiate such automatic commands at midnight.

After some experience was gained with the Scheduler, it was found the SRM was superfluous as a batch job control mechanism. We have therefore designed our IPS to make swapping very unlikely in our batch performance groups. (high ISV values; minimum MPL set equal to # of initiators).

Other observations:

Before the Scheduler was turned on, seminars were held for applications programmers on how to 'beat' the scheduler. Techniques such as use of UNIT=AFF on tape drives, how to compute DASD SPACE requirements, and how to reduce CPU consumption by reblocking files were described.

The Scheduler identification process was then described in detail, so individuals understood how they could qualify for the more desirable classes (short turnaround). The results, after a months use of the Scheduler, were remarkable:

- o Significant numbers of programmers modified their JCL and programs on their own in order to qualify for better turnaround and concurrently reduce the total system workload.
- o Tape drives allocation recovery, which accounted for 10% of all job's elapsed time, virtually disappeared.
- o Turnaround, especially on low - resource non - setup jobs, improved dramatically.
- o Turnaround in general became much more consistent.

The job class 'Z' is a catch-all class to trap jobs which fit in no other category. One initiator is always set aside for such jobs, since predictably crises will occur which require circumvention of the Scheduler. (Jobs are most easily forced into 'Z' by use of PRTY = 3 on the job card). Make sure, however, that you keep close track of jobs running in 'Z': Since class 'Z' jobs are usually high - resource users outside normal scheduling control, they will wreak havoc with jobs in the other initiators.

Don't be afraid to experiment with various class limit values and initiator setups; since such changes require no user JCL changes, their only impact is on turnaround time.

JOB RESOURCE ESTIMATE SUMMARY

The resource identification made by the scheduler is summarized on the JES JOB LOG of every job. A description of the resource summary is given below: (SEE EXHIBIT #1).

1. Programmer name Programmer name, as given on the job card.
2. Jobname Jobname, as given on the job card.
3. Job class Job class as assigned by the scheduler. Based on the resource estimate, the scheduler assigns a job class.
4. Performance group number- performance group number, as given on the PERFORM= keyword.
5. CPU time CPU time estimate, given in hours, minutes, and seconds. The above example shows 10 minutes, 0 seconds.
6. K - CORE The limit on the virtual storage to be used by the job given in K.
7. KLNES The estimate of print lines given in 1,000's of lines.
8. 6250 The maximum number of 6250 - bpi tape drives to be used by any one step.
9. 1600 The maximum number of 1600/800 bpi tape drives to be used by any one step.
10. 1K-DASD The maximum amount of DASD space allocated in any one step, given in 1K - block equivalents.
11. 1K-VIO The maximum amount of VIO space allocated in any one step, given 1K-block equivalents.
12. TP The maximum number of teleprocessing devices assigned in any one step.
13. DU The maximum number of direct unit assignments made in any one step.

14. UR

The maximum number of unit record assignments made in any one step.

15. PRTY

The job selection priority for the job, as given on the PRTY= keyword.

STEP TERMINATION STATISTICS

In the earlier seminars on the Scheduler, it was very apparent the application staff needed a readily accessible and comprehensive summary of I/O, CPU, and Storage Activity within each batch job step. Such a summary was necessary to target or direct tuning efforts. With such a need in mind, an SMF exit (IEFACTRT) was greatly expanded and such statistics were produced as part of the deallocation messages in each batch job's JCL. See exhibit # 1 for an example of these statistics.

<u>FIELD</u>	<u>MEANING</u>
USED	<p>The maximum amount of virtual storage actually used by the step; if running V=R, the field will be labelled "REQUESTED", and will reflect the amount of <u>real</u> storage requested (may not actually be used).</p>
LSQA+SWA	<p>LSQA = Local System Queue Area SWA = Scheduler Work Area. The LSQA contains tables and queues associated with your address space.</p> <p>It includes;</p> <ul style="list-style-type: none">o page tableso region control task tableso timer queue elements <p>The LSQA is normally about 20K in length. It is always 'fixed', which is significant if a system is real-storage constrained. The area is fixed for the duration of the job step/task.</p> <p>The SWA contains control blocks that exist from task initiation to task termination. It includes control blocks and tables created during JCL interpretation and used by the initiator during job step scheduling.</p>
AVE W/S SIZE	<p>Average Working Set Size is the average size of the group of pages in <u>real</u> storage.</p>
SWAP CNT	<p>The number of times the task was swapped out to external storage. Swapping is the mechanism used by the SRM (System Resource Manager) to break bottlenecks. Each swap</p>

can be expensive: 60,000 instructions or about 51 msec on a 158-3.

PAGE-INS

The number of pages moved into page frames in real storage. This count reflects physical I/O to move the page from auxiliary storage to real storage.

PAGE RECLAIMS

The count of page frames which are disconnected (stolen) from an address space, but are retrieved for re-use before being re-allocated to another address space. Although no physical I/O is reflected in this count, the 'cost' in CPU resources is about the same as a page-in operation.

PRV, CSA, VIO

PRV - user private area - the region in which your program executes.

CSA - Common Service Area - CSA contains pageable data areas that are addressable by all active virtual storage address spaces. It contains LOG, SMF, and reply buffers, to name a few.

VIO - Virtual I/O is a mechanism used to simulate real I/O that might otherwise have been required for a temporary file (e.g. compiler work file). The data 'written' by the program (e.g. compiler) is merely transferred within virtual storage, and when the data is 'read' back in, is again transferred within virtual storage back to the program's buffer areas. The auxiliary storage manager (ASM) reduces the impact of VIO by paging out long-term unreferenced VIO pages when under real storage stress.

SWAP

The number of pages involved in all swapping operations during step executions. In a non-MVS/SE environment, reclaim of pages of a swapped-out address space cannot occur.

TCB TIME

Task Control Block Time - includes all CPU time spent in

- o executing instructions in the problem program

- o executing instructions in the 'front end' of SVC's (supervisor call).

(SVC's, which perform pre - defined system tasks, such as GETMAIN or ENQ, are of two types: single stage and two - stage. In single stage SVC's, such as GETMAIN, 'front-end' refers to the entire SVC; in two-stage SVC's, such as EXCP, 'front-end' refers the first stage, which is initiation of an action, such as issuance of a SIO instruction. The second stage involves processing the interrupt when the action is complete, such as when an I/O interrupt occurs).

SRB TIME

System Request Block Time. SRB time includes time spent in:

- o Page Resolution/Swap Control
- o I/O interrupt processing (see 'two-stage SVC's above);
- o Timer queue element processing;
- o Interregion communication;

As a rule of thumb, most SRB time is accumulated in I/O interrupt processing.

RES TIME

Transaction Residency Time - Under MVS a transaction represents the level at which the SRM monitors and controls cumulative service consumption by an address space. For batch jobs, a transaction corresponds to a cluster of contiguous job steps under one performance group. The transaction residency time is the amount of transaction elapsed time spent in a non-swapped state, minus allocation and deallocation elapsed time.

ALLOC TIME

Allocation Time - elapsed time between step initiation time and when the problem program receives control. The time is spent by DADSM in locating and assigning DASD space and by job management in assigning exclusive-use UCB's, such as tape drives.

UNRC'D TIME

Un-recorded time - the time reported here is a computed value and is an estimate of non-TCB time. The elements which make it up include time spent in:

- o I/O interrupt processing;
- o swapping;
- o paging.

TCB/SRB

This ratio is helpful in determining if a significant amount of step CPU time is spent in I/O processing.

RES/TCB

This ratio is helpful in determining if the step, while swapped in, is being given adequate access to the CPU.

ATV/RES

This ratio is helpful in determining if a step is suffering long elapsed times primarily because of swap - outs.

PFM GRP

Performance Group. Each job or job step requests a specific (or is assigned a default) performance group number. At the SCC, default performance group numbers are assigned by job class. The performance group characteristics are assigned by the IPS (Installation Performance Standard). The IPS dictates the amount of access each performance group is to be given to the CPU, under varying workload conditions, and under varying amounts of total service requested by the job or job step.

PRTY

Dispatching Priority - when a task is permitted to be dispatched by the SRM, this value states the relative priority of this task over all others currently dispatched. The value in this field may have one of four values:

MTW(nn) - MTW is the mean-time-to-wait group, which corresponds to the old (VS/1) definition of DDG.

ROTATE - a group where all member tasks are given service on a round - robin basis.

FIXED(nn) - a group with a fixed priority relative to all other groups.

DISP(nn) - if a task is outside control of the SRM, its dispatching priority is shown; all other field values shown above refer to dispatching priorities controlled by the SRM.

STEP NO

Self - defining

STEP NAME	Self defining
-----------	---------------

COND CODE Self - defining - if a step abends or is
not executed, the field may have the format:

```
o    ABEND          Uxxx - user abend
o    ABEND          Sxxx - system abend
o    NOT EXEC      (BYP)
```

BYP means the step was not executed because a condition code test or step restart.

SERVICE UNITS Self - defining

I/O COUNTS

The I/O counts, and unit and device type assignments are in the same order as the DD statements for the step. Read them one column at a time. Since the system is generated with alternate channels, '25i' and '15i' in the UCB field refer to the same physical device.

Below the I/O statistics area are two (2) additional sections, one for VSAM statistics by DDNAME, the other for NON - VSAM statistics by DDNAME. The header line in each case contains three (3) strings of numbers, given as (XXX/YYY/ZZZ) where

XXX - total number of slots available for storing this DDNAME data.

YYY - total number of slots used by this step.

ZZZ - total number of additional slots required (if any) if available slots are not adequate.

For NON - VSAM file statistics the following items are provided:

DDNAME - DDNAME as it appears in the JCL. The DDNAME entries are ordered alphabetically; no entry is present if the associated file was never opened.

DEVICE - the device type (interpreted) on which the associated file resides.

VOLSER - VOLSER of the associated file.

OPNTYP - type of OPEN (e.g. INPUT, OUTPUT, OUTIN,
etc).

OPNCNT - number of times the associated file was opened via the DDNAME.

BUFNO - DSORG,OPTCD,RECFM,LRECL,BLKSIZE, - DCB characteristics of the associated file.
Because of shortcomings of the SMF data on

which these numbers are based, the BUFNO value is often invalid.

BLOCK-CNT - total number of blocks transferred via the DD statement regardless of the number of OPEN's.

EXTS - number of extents, if a DASD - based file.

TRKALC - number of tracks allocated, if a DASD - based file.

TRKUSD - number of tracks used if a DASD - based file. Because of shortcomings of the SMF data on which this number is based, this value is sometimes not available or is invalid.

TRKRLS - number of tracks released from file, if a DASD - based file.

For VSAM file statistics, the following items are provided:

DDNAME - DDNAME as it appears in the JCL. The DDNAME entries are ordered alphabetically; no entry is provided if the associated file is never OPENed.

CPNT - Component type - DATA or INDEX component. If the VSAM KSDS is OPENed at the CLUSTER level, there will be one entry each for the DATA and INDEX components.

STAT - Component status - the data on which the VSAM statistics display is based provides two (2) items of information for each OPENed component:

CURR - cumulative information to -
date, but not including the
current OPEN.

CHNG - information on the current
OPEN only.

LEVELS - number of levels in the index component.

EXTENTS - number of extents in the associated component.

RECORDS - total number of records in the associated component.

DELETES - total number of records deleted from this component.

INSERTS - total number of records added to this component.

UPDATES -	total number of records modified in this component.
GETS -	total number of records retrieved from the component.
UNUS-CI -	total number of unused Control Intervals in this component.
SPLT - CI -	total number of CI - splits which have occurred in this component.
SPLT - CA -	total number of CA - splits which have occurred in this component.
BLKCNT -	total number of EXCP's issued against this component.

DSNAME CONTENTION REQUEUE FACILITY

Normal MVS handling of DSN conflicts blocks JES initiators until the conflict is relieved. To prevent this blockage, the REQUEUER was developed. The REQUEUER detects DSN conflicts before MVS does and requeues and holds jobs which otherwise would wait for data sets.

The facility works two (2) different ways, depending on whether the DSNAME holder (conflictor) is a batch job or a TSO session.

- 1) Resource holder (conflictor) is a TSO session:

The resource requestor (victim) is requeued and put on HOLD. Messages are sent to the TSO user:

```
JOB xxxxxxxx HAS BEEN REQUEUED AND HELD
PLEASE FREE yyy_____yy
```

In addition, messages are sent to the console operator:

```
IEFUJI-08-I xxxxxxxx REQUEUED BECAUSE
zzzzzzzz HOLDS DSN yyy_____yy
```

```
IEFUJI-09-I xxxxxxxx NOT ELIGIBLE
FOR AUTO-RELEASE
```

As noted in the console message, a batch job requeued because of a DSNAME conflict with a TSO user is not automatically released when the conflict is removed. The console operator must release the batch job.

- 2) Resource holder (conflictor) is a batch job:

The resource requestor (victim) is requeued and put on HOLD. A message is sent to the console operator:

```
IEFUJI-08-I xxxxxxxx REQUEUED
BECAUSE zzzzzzzz HOLDS DSN yy_____yy
```

When the conflict is removed (conflictor job ends) the requestor (victim) job is AUTOMATICALLY released.

To implement this requeue facility it was necessary to provide a continuously - running system task, called REQUEUER. Normally REQUEUER will require no attention by the operators or system programmers, but there are features both groups should be aware of:

When the REQUEUER task is started at IPL, it verifies the CURRENT status of all jobs previously involved in resource conflicts before the system went down. The exceptions to this are all jobs requeued because of conflicts with TSO sessions. During this verification process, jobs which qualify may be automatically released.

If it is desired to repeat this verification process at any time, the following command should be issued:

F REQUEUER,VERIFY

If the REQUEUER task is suspected of having difficulty, the following command should be entered:

F REQUEUER,DUMP

This command formats and dumps relevant control blocks to the printer. It does NOT terminate the REQUEUER task.

To stop the REQUEUER for any reason, enter:

P REQUEUER

If the REQUEUER is stopped, MVS reverts back to its former way of handling DSNAMES conflicts.

REQUEUER SYSTEM MESSAGES

Message: IEFUJI-01-W-FATAL ERROR-CALL SYSTEMS PROGRAMMING
Cause: The Requeuer system has detected an error which
has caused an individual job to be failed.
Action: Notify Systems Programming immediately.
Record the name and JES Job # of the failing job.

Message: IEFUJI-02-I-OPEN FAILED FOR SNAPDUMP DD STATEMENT
Cause: The SNAPDUMP DD statement is missing
from the initiator procedure.
Action: Notify Systems Programming immediately.

Message: IEFUJI-03-I-SNAP OF DIAGNOSTIC AREAS FAILED
Cause: Logic error is SNAP routine.
Action: Notify Systems Programming immediately.

Message: IEFUJI-04-I REQUEUER SYSTEM SUSPENDED IN LONG-TERM WAIT
Cause: Logic error or difficulty with disk pack containing
REQUEUER Resource Control file.
Action: Contact Systems Programming immediately.
Note Jobname and JES Job # of failing job.

Message: IEFUJI-05-I-ERROR IN REQUEUER TASK
Cause: Probable logic error in REQUEUER system task.
Action: Notify Systems Programming immediately.
Note Jobname and JES Job # of failing job.

Message: IEFUJI-06-I-ERROR WHILE ATTEMPTING TO REQUEUE JOB
Cause: Error detected in the JES IEFSSREQ interface.
Action: Contact Systems Programming immediately.

Message: IEFUJI-08-I-xxxxxxx REQUESTED
BECAUSE zzzzzzzz HOLDS DSN yyy____yyy
Cause: Requeuer system has detected a DSNAME
conflict and has requested the DSNAME requestor.
Action: None

Message: IEFUJI-09-I-xxxxxxx NOT ELIGIBLE FOR AUTO-RELEASE
Cause: The resource holder (conflictor) is a TSO session.
Action: After the TSO user has released the DSNAME
causing the conflict, release this job.

Message: REQUEUER-04-I-W-FATAL ERROR - CALL SYSTEMS PROGRAMMING
Cause: Logic error in REQUEUER task.
Action: Call Systems Programming immediately.

Message: REQUEUER-01-I-OPEN FAILED FOR SNAPDUMP DD STATEMENT
Cause: The SNAPDUMP DD statement is missing.
from the REQUEUER procedure.
Action: Call Systems Programming immediately.

Message: REQUEUER-02-I-SNAP FAILED
Cause: Logic error in SNAP dump routine.
Action: Call Systems Programming immediately.

Message: REQUEUER-03-I-OPEN FAILED FOR RESFILE DD STATEMENT
Cause: RESFILE DD statement is missing from
the REQUEUER procedure.
Action: Call Systems Programming immediately.

INSTALLATION INSTRUCTIONS

- o Punch/list first data set on tape:

```
//SI EXEC PGM=IEBPTPCH
//SYSPRINT DD SYSOUT=A
//SYSUT1 DD UNIT=TAPE,DISP=SHR,LABEL=(1,SL,EXPDT=98000
// DSN=CP.JSC10LOD,VOL=SER=SENATE
//SYSUT2 DD SYSOUT=B,DCB=FUNC=I
//SYSIN DD *
      PUNCH TYPORG=PS,MAXFLDS=10
      RECORD FIELD=(80)
```

- o After customizing JCL for local use, load files 2 thru 4 from tape to disk.
- o Using member ASSEMBLE from library CP.JSC10JCL, perform assemblies of all source.
- o Customize source as necessary:

IEFUJV -

- o Update resource limit table if desired - lines 851000 thru 899000
- o Update device name lookup table lines 914000 thru 951000. For esoteric names (e.g. SLOWTAPE, DSOBTAPE, SYSWK, etc) you must first get a hex dump of your local device name table.

Note that the 4-byte UCB type field value is unlike the values for generic names. Remove/replace all of the original esoteric names from the device name lookup table. Add entries for each esoteric name defined in your local DEVNAMET.

- o Update the esoteric name masks, lines 988000 thru 1005030. The mask entries are ordered according to the index number found in each esoteric name entry in the DEVNAMET. The mask value is 'exclusive OR' ed' against the value from the device name lookup table; the result is a valid UCBTYP field value. An example will clear up this a little bit:

'SYSWK' is an esoteric name defined at our site. The table entry for their esoteric name, line 934000, gives a UCBTYP value of X'00032000'. This UCBTYP value is copied directly out of our DEVNAMET. The third byte (hex value X'20') gives the device class of the esoteric name - DASD in this case. The second byte (hex value X'03') is an index value. It means this esoteric name is the third esoteric name defined in the I/O GEN.

This index value is used by the job scheduler to locate a valid UCBTYPE field. In the case of SYSWK, its associated UCBTYPE mask is in the esoteric name mask table at line 990000. Note that this mask value must be 'exclusive OR'ed against the UCBTYPE field value gotten from the device name lookup table

X'00032000' line 934000
(exclusive OR)
X'3053000D' line 990000
X'3050200D'

VIO entries are a special case: the result of the 'exclusive OR' operation must result in a valid UCBTYPE value for bytes 3 and 4, but bytes 1 and 2 must be X'FFFF'. (The author recognizes this procedure is cumbersome and will clean it up considerably in later releases).

- o Update the DASD attribute table lines 1020000 thru 1025000 as necessary.
- o Build your own acct # validation system, if desired after line # 1722000. Note one of the bits in the JMR (mapped by macro COMMEXIT), CEPIDF01, is used by subsequent exits (IEFUTL and IEFUSO) to distinguish between a production and test job. If the bit is off, the associated job is considered production and SYSOUT and TIME = limits will not be enforced (extension will be granted and warning message issued).
- o Modify routing codes used for picking ticket - line 3389000. Jobs which are placed on "HOLD" by the exit (see job resource table lines 851000 thru 899000) or by the programmers (TYPRUN=HOLD) will cause the scheduler to produce a 'picking ticket' using route code = 15. The picking ticket size is the same as continuous form tractor feed card stock available from your IBM representative. This picking ticket facility will require a dedicated output - only console. In addition to listing a resource summary for the job, it lists the:
 - o # of scratch tapes the job will need.
 - o the volsters of specific tape volumes the job will need.
- o Disable the 'HOLD' and class - assignment feature, if desired - lines 3582000 and 3578000. Initially you will want to disable these features to permit programmers to get accustomed to the new resource summaries provided on the JES2 job log. As soon as you wish to use the scheduler to assign job classes, re-enable these instructions.

IEFACTRT

- o Modify the UCBTYP name lookup table, as necessary - lines 475000 thru 501000.
- o Modify LOGO - lines 1571000 and 1582000.
- o Reassemble any exits, as required.
- o Using member FORMAT of PDS CP.JSC10JCL, initialize the REQUEUER resource control file. If you have a shared DASD system with MSI from Allen Services installed, you will need only 1 resource control file. If you do not have MSI and are running a multi - CPU environment, you will need 1 resource control file per CPU. Read the notes in the JCL carefully.
- o Update your JES2 and INIT procs using members JES2 and INIT of PDS CP.JSC10JCL as guides.
- o Install member REQUEUER from PDS CP.JSC10JCL in SYS1.PROCLIB
- o Run member SMP in PDS CP.JSC10JCL to install exits into your operating system.
- o Install member IEALPAJ2 of PDS CP.JSC10JCL into SYS1.PARMLIB.
- o Place library CP.JSC10LNK in the authorized library list (IEAAPF00).
- o Re - ipl with MLPA = (J2)

If you wish to not contaminate your system LPA, perform all of the above EXCEPT the SMP run, then

- o Run member LINKEDIT of PDS CP.JSC10JCL.
- o Install member IEALPAJ1 of PDS CP.JSC10JCL into SYS1.PARMLIB
- o Put library CP.JSC10LNK in your link list.
- o Re - ipl with MLPA = (J1).

DGH:ek

[illegible]

[illegible]

Initiator setup 8:00 A.M. to 6:00 P.M.

I1	= Q	(drained)
I2	= R	(drained)
I3	= S	(drained)
I4	= N	(started)
I5	= O	(started)
I6	= P	(started)
I7	= ABC	(started)
I8	= ABCEFG	(started)
I9	= AEIJGKBC	(started)
I10	= AEIJGKBC	(started)
I11	= AEIJGKBC	(started)
I12	= ZA	(started)

Initiator Setup 6:00 P.M. to 8:00 A.M.

I1	= Q	(drained)
I2	= R	(drained)
I3	= S	(drained)
I4	= N	(started)
I5	= O	(started)
I6	= P	(started)
I7	= ABCDT	(started)
I8	= ABCDTEFGHU	(started)
I9	= AEIJGKBCTUV	(started)
I10	= AEIJGKBCTUV	(started)
I11	= AEIJGKBCTUV	(started)
I12	= ZA	(started)

DGH:ek