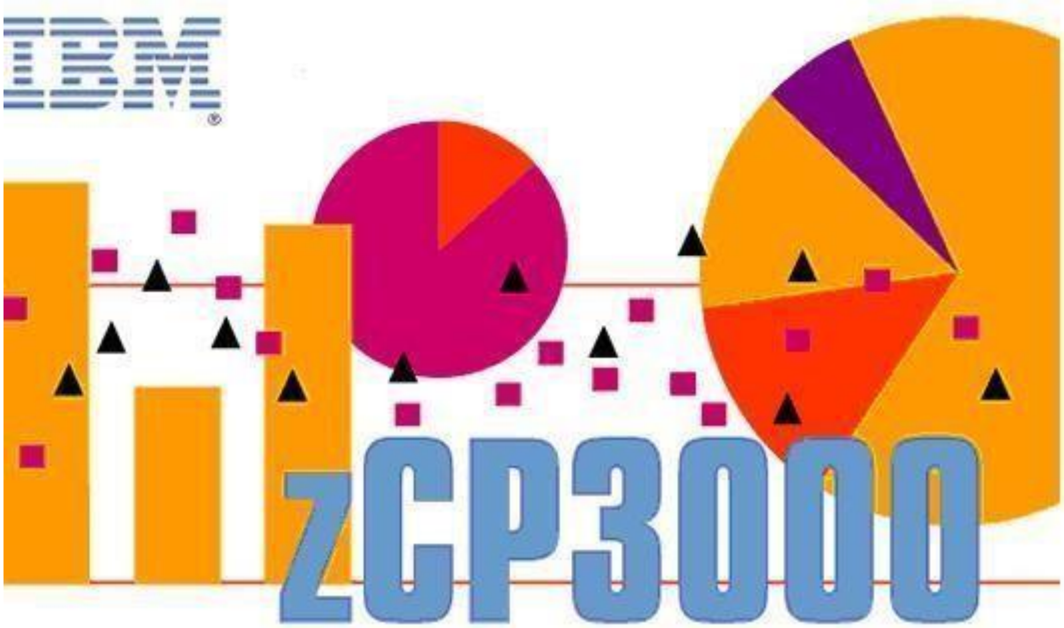
Data Extraction Program
For IBM *z* Batch Network Analyzer,
IBM *z*CP3000, and
IBM *z* Processor Capacity Reference

User's Guide

CP3KEXTR v4.36 07/19/2024



Shawn Lundvall

Joel Moss

Valerie Spencer

Contacts:

Comments, requirements, and questions may be directed to us at this address:

IBM CPS Tools Team	cpstools@us.ibm.com

You can download this document (**CP3KEXTR.pdf**) and the program package (**CP3KEXTR.Zip**) from one of the following sources:

https://ibm.biz/Bdmibu

The descriptions of output file formats and EDF vector names have been moved to a separate "Technical Reference" manual (that can be found at the same location as this User's Guide).

Table of Contents

Recent Program Versions

4.36	**07-19-24**	Update to the z16 RNI calculation Fix for a rare S0C1 condition encountered with 74(8) records
4.35	**05-14-24**	Exponent indicator changed to 'E' from 'e' Add SMFTIOE5 into the 14/15 records of the DAT file
4.33	**11-06-23**	Add NIO field. Add CPIT1P & CPIT2P Calculation updates for FCPI & SCPL1M & DMDPGV Remove SMF 23 processing
4.32	**10-10-23**	Add originating service class to reports if available Fixes to printing of large values. BCUMAP now reads more records before stopping.
4.31	**07-08-23**	Fixes for new zHyperLink™ fields Better processing for CUs with zero NED data WLCK for lock promotion Bug in rounding for time durations Better z/VM processing 42(6) fixes
4.30	**05-09-23**	42(6) fields switched to use new microsecond fields Added new zHyperLink™ fields from 74(8) record DASDIOV includes zHyperLink™ SYSSTCxx gets priority of 254 like SYSSTC New WRTD field for transactions finishing in .5 of goal Fixes for return codes etc.
4.29	**03-20-23**	Misc. fixes for Boost & zHyperLink™
4.28	**2-17-23**	New Fields for Synchronous I/O from 74.1 data New Fields for SMF 16 data Always put out input data statistics Message changes Various bug fixes
4.25	**11-07-22**	Addition of various DFSORT metrics to DATAnnn file. Fix to IOCP processing Various bug fixes
4.24	**8-29-22**	New VRPR section within EDF. This details topology of chip, DCM, Book, Drawer etc of processors.
4.23	**5-30-22**	z16 support. Various changes to calculations and formulae. Telum chip offload vectors

		CF structure async system duplexing flag Transactional CPU usage per service class
4.21	**06-07-21**	Fixes for issues with 113 synchronization and invalid data.
4.19	**04-23-21**	Fix for Issues with sample times Fix for incorrect PI values for Percentile Goals in period two and higher Accommodation for CPC serial number greater than 5 characters.
4.18	**03-08-21**	Fix for Issues with sample times Add vectors SLH and TPI for SMF70SLH & SMF70TPI. Message if DURATION greater than 70 or 113 interval. Stop. Message if SYSID moves between LPARs. Stop. Some changes to number of decimals
4.16	**02-19-21**	Add 2 new vectors, FRL9A= & FRL9M= with data from SMF71L9A & L9M Add new 'dat' type 1 field for SMF30HPT Add new vector LPIRDx= indicating setting of SMF70PFG bit 4 Fix for Cryptographic coprocessors
4.13	**02-12-21**	Fix case where 'dat' file time was incorrect Fix case where LPAR-related vectors would get an extra entry Fix rare 0C4 exception encountered with a 74(8) record
4.10	**12-22-20**	Implement processing of SMF 113 subtype 1 records Discontinue support for SMF 113 subtype 2 records Upgrade the algorithm that assigns SMF records to reporting intervals Add SMF 16 fields: iceMOSIZ and Rat = iceRcOUT / iceRcINP Change LpprPx calculation to ignore SMF70PDT value Remove INSTR vector Remove refs to 78.1 Remove SMF 23 mod & references

4.06	**07-30-20**	Remove the sections with the dat file and EDF formats from this "User's Guide" and place them in a separate "Technical Reference" document. Rework the BOOSTnn vector and add BOOSTFLA & BOOSTCLA Add two new input parms: SELDAT & LIMDAT Add a new SIISP (SIIS percent) vector and remove the EFA vector In the dat file: add SMF 16 data and add 2 new SMF 30 fields Add SMF type 16 to default TYPES= list

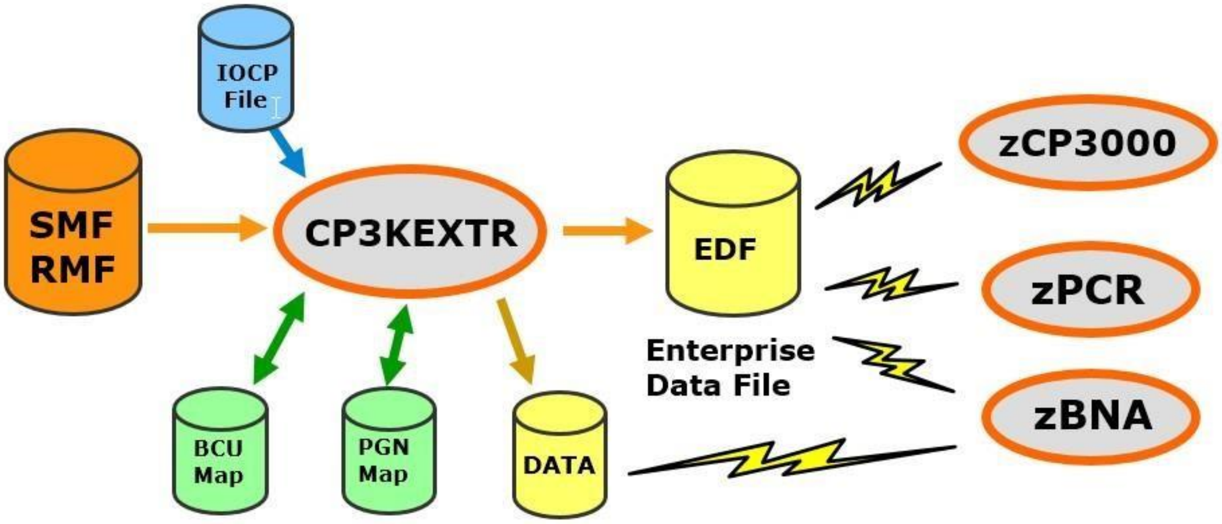
1. Introduction

The Extract program, CP3KEXTR, runs on a z/OS system where it reads and summarizes one or more SMF datasets for subsequent analysis by one of the CPS Tools Applications (IBM zCP3000, IBM zBNA, or IBM zPCR).

The program performs these functions:

- Parses the complex structure of the SMF data
- Extracts selected fields needed by the Tools applications
- Combines the data into a series of Extract "reporting periods"
- Produces a flat Enterprise Data File (EDF) that is read by the analysis applications
- Optionally writes a separate DAT file for zBNA with selected SMF fields
- Generates and/or uses several supplemental mapping files

Here is an overview of the data flow:



When the Extract program is run the default is that all input SMF records will be examined, but you can specify particular dates and times to select only certain data. You can also adjust the duration of the reporting period.

Program Inputs

The program uses makes use of several different inputs as follows:

SMF data	SMF (the System Management Facility) is a z/OS mechanism that records system and job-related information. The data is represented by a series of different record types each specializing in a different kind of information. The program selects specific record types based on the requirements of the run (more in **part 5**).
RMF Data	RMF (the Resource Measurement Facility) is a performance monitor for z/OS that collects operational data for later analysis. RMF data is placed into the SMF record stream using record types 70 to 79 (i.e., the SMF dataset includes the RMF data).
User Parameters	User inputs let you select a variety of processing options and identify what part of the input SMF data you want to examine. Most of the parameters are optional and many have default values that provide standard processing (see **part 3**).
BCU Map	Typically I/O control units with identical channel configurations are combined for reporting. This map identifies how (details in **part 6.1**).
PGN Map	The program combines service classes for reporting using the PGN map. Usually the program generates the map automatically and then uses it in processing, but if needed you can manually edit the file to adjust the combining (see **part 6.2**).
T30 Map	The T30 map provides a way for you to break out workloads in greater detail than the WLM service classes (details in **part 6.3**).
IOCP File	The IOCP (I/O control program) produces the IOCDS (I/O Control dataset). This contains details of directors, links, channel paths, units, and subsystems along with other system information. Here “IOCP File” refers to the systems IOCDS.

Program Outputs

These are the program output datasets:

EDF	The Enterprise Data File (EDF) is the primary output. It is a flat text file containing summarized data extracted from the SMF scan. The EDF is the input to other analysis applications like zBNA, zCP3000, and zPCR. An example is in **part 8.2**.
Print Output	The print listing shows the input parameters specified for the run along with any messages that were generated by the program.
BCU Map	The program will generate a BCU map if you specify the **BCU=AUTO** parameter, otherwise it will use an existing BCU file (details in **part 6.1**).
PGN Map	The program will generate a PGN map if you specify the **PGN=GOAL** parameter, otherwise it will an existing PGN file (details in **part 6.2**).

DATA File	The program can extract fields from the specific SMF record types and format the data into a flat text file for use by zBNA. The output file is specified by the DATA001 DD and can include fields from type 30, 42, 14, and 15 records. **part 7** gives details.

Reporting Interval

There are two different time intervals that must be considered when working with the Extract program – the RMF interval and the Extract reporting interval.

The **RMF Interval** is the time period (typically 15 minutes) during which z/OS accumulates and summarizes both software and hardware activities. At the end of this interval the system writes the collected information to the SMF dataset and, for many items, clears the accumulators to start again. The RMF interval is set by the system programmers and usually does not change.

The **Extract Reporting Interval** is the time period the Extract program uses to group the incoming SMF data. This is specified by the DURATION parameter and often is set to one hour (which gives more attractive graphs in zCP3000), but setting the reporting interval to be the same as the RMF interval is also common. Data from the SMF records that fall within an Extract reporting interval are combined and reduced to a single value (a total, an average, a bit indicator, etc.).

For example, consider one specific measurement item, the “start subchannel count”. Processing for this item will flow through the following steps (over-simplified for discussion):

- During the 15-minute RMF interval while the original applications are executing, this counter is incremented by the system for each start subchannel request.

- At the end of the RMF interval that count is written by the system to an RMF record.

- Later, when the Extract program is run it scans the RMF records.

- The “start subchannel count” from each RMF record (type 74) is totaled within the duration of an Extract reporting interval (e.g., records from 4 RMF intervals would be encountered for a 1 hour reporting interval).

- At the end of the Extract reporting interval the total is divided by the number of values giving an average.

- That average is appended to the string of values reported as the “start subchannel count” vector (DASDIOV) in the output EDF (resulting in one value in the vector for each Extract reporting period).

Summary -- for many data values the progression is to combine execution-time measurements into an RMF interval and then further combine these into an Extract reporting interval.

SMF Record Selection

It may be that the input SMF dataset contains information for only one partition for one particular time period. This is a simple case where you may want to accept and process everything in the input.

A more complicated situation is when the SMF input includes data for several partitions, perhaps generated over many days. In this case you can ask the program to select just certain records belonging to a particular partition or a specific time period. I.e., you can extract and summarize a subset of the original data. (see **part 3.2**).

2. Startup Guide

This section provides a checklist of the basic steps needed to get started running the Extract program. This is a simplified approach; more details are provided in **part 3** of this document.

2.1 Preparation

a. Have a **TSO** user ID and password on the z/OS system where the Extract will be run.

b. Identify the necessary parameters for the JES **JOB** card for that installation.

c. Get the SMF **SysId** of the system(s) for which we are Extracting data (see also **part 2.5**).

d. Decide when and for how long you wish to **sample** and set the Extract reporting interval accordingly (with the DURATION parameter). For performance analysis using the 15-minute intervals may be reasonable. However, for doing capacity planning one-hour intervals are a minimum. Depending upon the amount of data available different durations may be recommended (but nothing less than 5 minutes). If you don't know where to start use a week of one hour samples. More details are in **part 2.5** and **part 3.2**.

e. Know the dataset names of the SMF files for the samples you have chosen.

2.2 Research

If possible, sit down with the system programmer and discuss these points:

a. Review the SMFPRMxx member of 'SYS1.PARMLIB'. Verify that you are collecting the SMF record types you will need.

- For IBM zPCR types 70 through 75, 78 and 113(1) are required.
- For IBM zBNA also include 14, 15, 16, 30 (2,3,4,5), 42 (6)
- For IBM zCP3000 use the zPCR types and also optionally 42 (6).

b. Instructions for collecting SMF 113 records are at this link:
https://www.ibm.com/support/pages/node/6354583
Note: zCP3000 requires Basic and Extended counters.

c. Verify that you are doing INTERVAL processing.

d. Ask about the job that dumps off the SMF datasets. Sometimes that job will exclude certain types and they will not make it to the output dataset. Verify that the ones you will need are present. You can use either the standard SMF dump program, IFASMFDP, or if the installation uses log streams for the SMF data you can use IFASMFDL to select a subset of SMF record types.

e. Ask whether the SMF records are in sequence by date and time.

f. If possible, obtain an I/O configuration diagram for the installation.

g. If you will be using the FICON aggregation function of zCP3000, learn the name of the IOCP dataset and the name of the LPAR partition you will be studying.

2.3 Install the Program

a. Download the CP3KEXTR Extract Program to your PC

CP3KEXTR is an assembler language program that is distributed in LOADER format. The file **CP3KEXTR.zip** contains both the program and the JCL. Download this from one of the following sources:

IBM Employees:	https://ibm.biz/BdmKR4
Business Partners:	https://ibm.biz/Bdmibu
Clients:	https://ibm.biz/Bdmibu

b. On your PC, unpack **CP3KEXTR.zip** to obtain the CP3KEXTR.bin file. This is the file that will be uploaded to the TSO system.

c. **Sign on to TSO**

Log in to the TSO system (where you will run the Extract program)

d. **In TSO, Allocate the upload dataset**

This is the dataset into which you will upload the CP3KEXTR.bin file. In ISPF/PDF **3.2** choose **A**llocate. Then,

- Name the new dataset **CP3KEXTR.UPLOAD**.
- Set the RECFM to FB (or F) and the LRECL to 80.
- Specify a BLKSIZE of 0 for system determined block size. If system determined block size is not implemented zero will give you an error. In this case use 6400.
- Suggest a primary extent of three cylinders with a secondary extent of one and no directory blocks.

e. Upload the Extract program.

The CP3KEXTR.bin file in the PC contains binary characters and must be uploaded in binary mode. Use FTP to transfer the file. Open a command prompt window and switch to the directory that holds the CP3KEXTR.bin file. Then enter:

```
ftp wscmvs.washington.ibm.com        <-- (substitute your systems IP address)
Response:
User (wscmvs.washington.ibm.com:(none)): YourUid
Response:
331 Send password please.
Password: YourPw
Response:
230 YourUid is logged on. Working directory is “YourUid.”.
```

```
ftp> binary
220 Representation type is Image
ftp> put Cp3kExtr.bin CP3KEXTR.UPLOAD
Response:
200 Port request OK.
125 Storing data set YourUid.CP3KEXTR.UPLOAD
250 Transfer completed successfully.
ftp:344400 bytes sent in 0.49Seconds 708.64Kbytes/sec.
ftp> quit
```

f. In TSO, “Receive” the file

The CP3KEXTR.bin file from your PC is now in the TSO system. This step reformats it with the RECEIVE command into the CPSTOOLS.JCL data set.

In ISPF/PDF 6 type: **`RECEIVE INDSN(CP3KEXTR.UPLOAD)`**

The command will prompt you for the dataset name. Type:
`DA(CPSTOOLS.JCL) SPACE(2,2) CYLINDERS`

This will create the `CPSTOOLS.JCL` dataset (within your high-level qualifier).

2.4 Customize the JCL

In the CPSTOOLS.JCL dataset there is a "JCL" member. Edit this as described below.

a. Change the **JOB** card to meet the standards of your installation.

b. Adjust the 4 SET commands as needed to point to the data sets you will use.

2.5 Setup the Input Parameters

The Extract procedure is flexible and can support a number of variations in processing. These are selected using input parameters you provide via the **SYSIN001 DD**.

Most of the parameters are optional and default to functional values if not specified. However, the **SysId** of the system you want to study must be provided.

If you don’t know the **SysId** for the SMF data you can specify a dummy SYSID=XXXX (which won’t match anything). When there is no match the program skips its regular processing and instead prints a summary of the contents of the SMF dataset. This includes, among other things, a list of the SysId values found in the data. Put the following parameters (and only these) in the SYSIN001 input stream.

```
//SYSIN001 DD *
ENT='Customer Name'             (substitute a descriptive name)
SYSID=XXXX                       (specify a value of ‘XXXX’)
/*                              There is an output example in part 3.7
```

The following input parameters are the most commonly used. Additional parameters and variations are shown in **part 3**.

These 4 parameters are normally provided. The first 2 are always required:

ENT=’xxxxxx’	The Enterprise name, specified as a character string in single quotes (50 characters maximum). Imbedded blanks are permitted. Required.
SYSID=xxxx	Specifies the four-character JES SYSID of the system to be studied. Required Note – do not use quotes with this item.
BCU=AUTO	Specified to generate a BCU Mapping file. Normally you will include this. If you are interested in FICON Mapping you also want to specify an IOCP001 DD name pointing to the current IOCDS. This will generate a more complete BCUMAP suitable for FICON analysis. For the FICON analysis you should also use BCU=FICONAG and PART= to specify the LPAR partition name.
PGN=GOAL	Used to generate the Workload Mapping. Also normally included
SORT=YES	Specifies that the Extract should sort the incoming SMF data before processing. Since SMF records must be in sequence this parameter is typically included. Job run may be significantly shorter if SMF records are already properly sorted and this parameter is omitted.

The following parameters can be left out and the default for the run will be to process all SMF input data with an Extract reporting interval of 1 hour. Or you can specify the range of SMF data you want to process and adjust the reporting interval.

DURATION	Specifies the Extract reporting interval in hours or hours and minutes. This parameter defaults to 1 if not provided. The minimum is value is 5 minutes.
	E.g., DURATION=1 for one hour (the default)
	DURATION=2 for 2 hours
	DURATION=00:15 for 15 minutes

TIME	Specifies the times (the hours of the day) you want to study. If not specified, all times will be used. Be sure that the duration fits evenly into this time range. E.g., TIME=(00-24) to select all hours of the day (the default) TIME=(08-11) to select data from the 8th, 9th, and 10th hours. You may have multiple TIME= statements. For example, if you are using 2-hour samples you may have TIME=(08-12) and another TIME=(13-17).

DATE	Specifies the days of interest. If it is not specified all dates in the input SMF file will be used. There are several variations of the DATE statement – it can include a single date, a range of dates, for a list of separate dates.
	E.g., DATE=(09/25/15)
	DATE=(09/25/15-09/29/15)
	DATE=(09/25/15,09/26/15,09/27/15,09/28/15,09/29/15)

> The RMF interval must fit evenly within the Extract reporting DURATION and the Extract reporting DURATION must fit evenly within the specified TIME period.

See **part 4** for an example of the JCL and the input statements.

2.6 Run the Extract

a. Submit the Job

Use the JCL member you just customized and submit the job. This will generate the output EDF file.

The following return codes will be issued:

0	No messages.
4	Messages of note have been issued.
8	Either important data (RMF types) were missing from the file or there is questionable data in some of the intervals (that you may want to exclude).
12	Something is critically wrong with the parameters, or no data fit within the SYSID and timeframe requested.
16	Something is drastically wrong and the program failed.
20	CP3KEXTR couldn't write to the PRINT001 DD.
225	No data was found that meets the input specification. This could happen if the specified SysId was not present in the SMF input dataset or if the date and time selections did not cover the time periods of the data.

Note: If no data was selected it might be because:

- The specified SysId wasn't present.
- No records were in the specified date/time range.
- The data is not sorted (add the sort work datasets to the JCL)
- If the data is not sorted you have to specify SORT=YES to get it to sort.

b. Check for Messages

Review the messages produced by the program -- they go into the PRINT001 DD statement and will be with your output JCL.

c. Download the EDF (and optionally the DATA file)

The EDF file will be in the member referenced by the EDF001 DD statement. For zBNA the DAT file will be in the dataset referenced by the DATA001 DD statement.

Download the EDF in ASCII mode and for zBNA processing also get the DAT file (also in ASCII mode). In cases where the DAT file is especially large, download the TRS (compressed) version instead (in binary mode). The JCL uses the following naming conventions:

File	*Mode*	*Usage*
XXXX.EDF	ASCII	The EDF output used by all CPS Tools.
XXXX.DAT	ASCII	The full uncompressed DAT file used by zBNA.

XXXX.TRS	BIN	The compressed DAT file. Download this instead when the uncompressed version is large. Note that zBNA can read the tersed file directly.

3. Program Input Parameters

This section describes the input parameters accepted by the Extract program (CP3KEXTR) through the SYSIN001 dataset. Note that if you are processing more than one LPAR in this Extract run there will be a SYSIN DD statement for each (e.g., SYSIN001, SYSIN002, etc.).

The **ENT** and **SYSID** parameters are required for all jobs while the others are used only if needed. The following conventions apply to coding parameters:

- One parameter per line
- Parameter text must begin in column 1
- The text may run up to column 72
- Comment lines are indicated by an asterisk '*' in column 1

You can review a sample of this type of input in **part 4.3**.

> Note that any of the Extract input streams may have comment lines. This includes:
>
> SYSIN, PGNMAP, BCUMAP or T30MAP. A comment is identified by an asterisk ***** in the first column (causing the remainder of the line to be ignored).

3.1 Required Parameters

There are two required parameters that must always be specified in each SYSIN dataset.

ENT='xxx xxx xxx'

The Enterprise name is specified as a character string in single quotes (up to 50 characters). Imbedded blanks are permitted. Multiple EDF files from different partitions and CECs for the same enterprise should have the same ENT parameter.

SYSID=xxxx

Specify the four-character JES SYSID as it appears in the SMF records. This is used to select only one particular SYSID from the SMF input.

Remember, if you don't know the **SysId** you can do a run with just the following two parameters and the program will print a summary of the contents of the SMF dataset showing the SysIds that are present.

```
//SYSIN001 DD *
ENT='Customer Name'           (substitute a descriptive name)
SYSID=XXXX                    (specify a value of XXXX)              /*
```

3.2 Reporting Interval Parameters

These parameters let you select or exclude certain dates and times. You can also specify the duration of the Extract reporting period.

DATE=(MM/DD/YY)

The specified date or date range will be used to select RMF records. Multiple DATE parameters can appear separately. If no DATE parameter is specified all dates in the input file will be used.

The parameter can be a single date, a list of dates, or a date range as follows:

- DATE=(MM/DD/YY)

- DATE=(MM/DD/YY-MM/DD/YY)
- DATE=(MM/DD/YY,MM/DD/YY,MM/DD/YY)

DURATION=hh

The DURATION value specifies the Extract reporting interval -- the time period in hours the Extract program uses to group the incoming SMF data. One or more RMF intervals will be summarized into each DURATION reporting interval.

If the DURATION is not specified the program uses a default of one hour, DURATION=1.

This parameter can be specified in whole hours or in hours and minutes:

- DURATION=hh
- DURATION=hh:mm

Remember, the RMF interval must fit evenly within the Extract reporting DURATION and the Extract reporting DURATION must fit evenly in the specified TIME period (next parameter).

For example, if your RMF has a 15-minute interval set your DURATION to 15 minutes or some multiple of 15 minutes. Specifying DURATION=01 (one hour) would evenly include 4 RMF intervals.

TIME=(hh-hh)

This parameter gives a range of hours. SMF data within this range will be accepted as input and data outside of this range will be ignored. The selection begins with the first value and runs up to the second value. Note that it is important that the DURATION of the Extract reporting interval fits evenly in the specified TIME period.

For example, if you specify TIME=(09-11) SMF data from the 9th and 10th hour will be included and would evenly contain a 1 hour reporting DURATION.

Multiple TIME parameters can be specified. The duration value must be specified first.

The time parameter can be given in whole hours or in hours and minutes:

- TIME=(hh-hh)
- TIME=(hh:mm-hh:mm)

The decision whether to include an RMF record is based on a time that represents the midpoint of the RMF interval. For example, if TIME=(09-11) is specified,

- An RMF record for an interval beginning at 10:55 with a duration of 15 minutes would not be included (because the midpoint time is after the 11:00 cutoff).
- A record beginning at 8:55 also with a duration of 15 minutes would be included (because its midpoint time is after 9:00).

With either form be sure that the DURATION fits evenly in this time window.

If you are using the SELECT statement you don’t need to also use the TIME parameter.

If neither TIME nor SELECT is specified the program defaults to TIME=(00-24), i.e., select all times.

SELECT DATE=MM/DD/YY,TIME=HH-HH
SELECT DATE=MM/DD/YY,TIME=HH:MM-HH:MM

EXCLUDE DATE=MM/DD/YY,TIME=HH-HH
EXCLUDE DATE=MM/DD/YY,TIME=HH:MM-HH:MM

These statements may be used to include or to exclude certain periods from analysis.

For example, if you are studying a customer's prime shift during a week but they are open late on Thursdays, you can add additional hours to the study for Thursday. In addition, if on Tuesday morning there was a hardware failure that made the data from 9-10 invalid you can EXCLUDE that data.

It is also possible to simply look at discrete times. For example, if the study of a week's data reveals that there are three peak hours it is possible to SELECT only those three hours for study.

3.3 Other Parameters

CECID/CPCID=cccccccc

Lets you specify up to eight characters for the CECID value that is included in the output EDF. If omitted the SYSID will be used for machines in basic mode. For machines in LPAR Mode CPCnnnnn will contain the 5-digit correct serial number.

NOTE: This only specifies a value for the EDF. It is not a selection item and has no effect on what SMF records are accepted in the run.

NOTE: It is best to let this default and rename the CEC once the data has been loaded into zCP3000.

CPUMOD=cccccccccc

A 10-character CPU model. If this is not specified a table lookup will be used with data in the RMF type 70 record.

PROCESS=YES/NO

This specifies whether you want the program to perform its full SMF processing.

NO	The program will accept the input parameters, generate the BCU and/or PGN maps, then stop without further processing. This is used in the case where you want to edit a BCU and/or PGN map before using it.
YES	The program will proceed through all the normal processing of the SMF data. This is the default. Note: If you have done manual editing of the BCU or PGN maps be sure to remove the BCU=AUTO and/or PGN=GOAL statements.

RMFINT=nn

Number of RMF intervals to be used when calculating the peak to average ratio. Any consecutive nn intervals will be considered. Default is one.

SORT=YES/NO

Coding YES will cause the program to do an internal sort on only the SMF data required for this study. In most cases the SMF data is already in time sequence and this sort is not required. The default is NO.

If you should see interval starts or durations which are strange or if CP3KEXTR should give you a message to sort the data then a sort is in order.

When using the internal sort the **REGION=** parameter should be adjusted as high as reasonable. Also, some thought should be given to the size of the sort work datasets.

SAVE=SORT=ALL/*SysId*

This is used to take an input (presumably tape) dataset and copy it to disk while sorting the data and eliminating duplicate records. The destination is described by the SMFSAVE DD statement.

ALL	All records (including every SysId) are written to disk
SysId	Only records for the specified SysId are written to disk

A report of the SysIds encountered, record types and subtypes encountered, and earliest and latest date and times will be written to the SYSPRINT DD statement.

Note that no other processing is permitted with this parameter. If this parameter is specified no other parameters are allowed except TYPES.

SUB30=n,n,n...

This is used with SMF records 30 processing to determine which type 30 subtypes are included in the DATA001 file. The default is SUB30=4,5. This parameter does not apply to the T30MAP processing which always uses type 2 and 3.

TYPES=14,15,16,30(2,3,4,5),42(6),70(1,2),71(1),72(3),
TYPES=73(1),74(1,4,5,8,9,10),75(1),78(3),89(1,2),110(1),113(1)

This parameter defines the records that are to be written to the disk dataset. It can only be specified with SAVE=SORT=. If this statement is omitted the default includes the types and subtypes shown above. Because the statement may exceed the 80 character input line length the record types can be specified as two separate statements (and the content of both will be combined by the program).

If you are not interested in using a particular record type you would include the TYPES statement as shown above but leave out that one type.

3.4 BCU - Parameters

BCU (basic control unit) is a term used in the Extract program to represent a physical I/O control unit -- the "box" that contains the LCUs (logical control units). The BCU map provides a mechanism to combine LCUs that are part of the same physical control unit.

BCU=AUTO/AUTOADD

The BCU=AUTO or AUTOADD option causes the program to create a new BCU Map and write it to the BCU001 dataset. This will combine BCUs that have the same channels. If not specified, an existing BCU map will be used. BCU=AUTO is the typical processing case, but since this is not the default it must be specified.

If IOCP data (IOCDS) is used it will combine BCUs with the same switches and links.

You must have type 74 subtype 1 and type 78 subtype 3 records. If you specify BCU=AUTO and these types aren't present the run will stop with an error message.

The mapping process is described in **part 6.1**. (Note that the map will be more complete if 74 subtype 5 records are available). If the IOCPnnn DD statement is found it will also be read to create a more complete BCUMAP.

Note that for shared DASD the BCUID may not be unique. This should be adjusted to be the same value on all systems.

Specification	**BCUID in EDF**
BCU=**AUTO**	The BCU manufacturer (**IBM**) plus a hyphen (**-**) plus the 5-digit device serial number. If no VOLSER is found the name will default to "**L**" plus the LCU number.

BCU=**AUTOADD**	"**A**" concatenated with the 4-digit hex address of the lowest numbered device on this BCU.

BCU=NO/DASD/TAPE/OTHER/ALL

These choices are used to request processing of an existing BCU map that will be read in from the BCU001 DD (i.e., you did not specify BCU=AUTO). This maps the volume data to BCUs by SYSID and address. Default is ALL.

You can specify the type of the BCU entries you want to process as follows:

NO	Skip BCU processing
DASD	Only process the BCU macros in the BCU map
TAPE	Only process the BCU**T** macros in the BCU map
OTHER	Only process the BCU**O** macros in the BCU map
ALL	Process BCUD, BCUT, and BCUO
DASD,TAPE	Process both DASD and TAPE. (Or any combination of DASD, TAPE, or OTHER separated by commas)

BCU=LCU/COM

You can specify BCU=**COM** as an additional option for BCU=AUTO to choose an alternate combining strategy. Put this specification on a separate line either before or after the BCU=AUTO statement. The processing distinction is:

Specification	**Method for Combining**
BCU=**LCU** (default)	Combine different LCUs that the program can determine are part of the same physical control unit.
BCU=**COM**	Combine different control units that share the same set of paths (i.e., the same CHPIDS and links). ESCON control units that share the same directors are combined. Control units on copper channels where the control units are daisy chained are combined.

If you will be using zCP3000 to analyze shared DASD, it will expect that a given BCU will have the same BCUID from each system. This will more likely happen automatically with BCU=AUTO than with BCU=AUTOADD.

BCU Summary

The BCU map can be generated automatically and used in the same run (case 1 below) or it can be generated, manually edited, and then used in subsequent processing (2 & 3 below). These are the parameters needed to handle these cases:

	Case	***Specify these parameters***	
1	**Fully Automatic** Generate the map and then use it	**BCU=AUTO**	to generate the BCU map
		PROCESS=YES	generate the map and then do the normal SMF processing
		BCU=ALL	use everything in the new map
2	**Generate a map** to manually edit	**BCU=AUTO** **PROCESS=NO**	to generate the BCU map to stop after making the map
3	**Use existing map** after editing	**omit** BCU=AUTO	don't make a new map (it would write over the edited one)
		PROCESS=YES	do the normal SMF processing and use the edited BCU map
		BCU=ALL	use everything in the map

BCU=FICONAG

There may be cases in FICON studies where it is desired that parts of physical boxes not be combined. These cases are when different groups of LCUs connect to different groups of CHPIDs. If you specify BCU=FICONAG these LCU groups will be kept separate in the BCUMAP.

PART=nnnnnnnn

This parameter specifies the partition name which is used to process the IOCP file in an LPAR environment. It is only used with BCU=AUTO when an IOCPnnn DD statement is provided. The default is the SYSID specification.

SHOWACT=ALL/NONE

Specifying NONE will suppress all actuator data. zCP3000 does not yet support files produced by using a value here but SHOWACT=.001 will cause only those actuators which have I/O rates of .001 or greater to be output. NOT YET TO BE USED WITH zCP3000. <>

3.5 PGN - Workload Parameters

zCP3000 users frequently want to combine service classes to a smaller number for executive reporting. The Extract program provides a facility to do this. It can also, probably more simply, be done once in zCP3000.

The PGN Mapping mechanism provides a way to combine service classes for reporting. The program does this automatically and you can make manual adjustments if needed. However, the zCP3000 application handles this so manual adjustments are usually not necessary.

The PGN001 DD statement must be specified with the JCL to process workload data. If it is not present, there will be no PGN processing regardless of the input parameters. See also **part 6.2**

PGN=GOAL

This option asks the program to automatically build a PGN map into the PGN001 DD file. It will scan the SMF data, read the SMF Type 72 subtype 3 records, do the combining, and write the map. If record type 72 subtype 3 data is not present the program will issue an error message and terminate the run.

Normally you include this parameter. But when you leave it out the program uses an existing PGN map, giving you the opportunity to make manual adjustments. See also **part 6.2c**.

Case 1 below is where the PGN map is generated automatically and then used in the same run. Case 2 & 3 below is where the map is generated, manually edited, and then used in later processing. These are the parameters needed to handle these cases:

	Case	***Specify these parameters***	
1	**Fully Automatic** Generate the map and then use it	**PGN=GOAL** **PROCESS=YES**	generate the PGN map do the normal SMF processing and use the new PGN map
2	**Generate a map** to manually edit	**PGN=GOAL** **PROCESS=NO**	generate the PGN map stop after making the map
3	**Use existing map** after editing	**omit** PGN=GOAL	don’t make a new map (it would write over the edited one)

		PROCESS=YES	do the normal SMF processing and use the new PGN map

TYPE42=YES/NO

This option will build records for each of the workloads describing the dataset usage by business unit. This data will be built from type 42 subtype 6 records. The default is YES.

There are some assumptions for the type 42 records processing.

- A PGNMAP was supplied.
- BCU=DASD or ALL was specified.
- SHOWACT=NONE was not specified

3.6 Convert Tape to Disk

The Extract Program has a facility to copy an input dataset (usually from tape), sort the records removing duplicates, and write them to an output dataset (usually on disk). The destination is described by the SMFSAVE DD statement.

As part of this processing a report will be written to the PRINT001 DD statement which lists:

- All SYSIDs encountered
- For each SYSID, the record types and subtypes encountered
- For each record type and subtype the earliest and latest time and date

No other processing can take place in a tape to disk run.

The input dataset must be allocated to the SMFIN DD statement. The output (disk) dataset will be written to the SMFSAVE DD statement. Give the job as much region and SORTWKnn’s as you can.

SAVE=SORT=ALL SAVE=SORT=*SysId*

No other statements except for TYPES described below are permitted with the SAVE=SORT option.

=ALL	All records (including every SysId) are written to disk. Use this if you have data with multiple SYSIDs and you would like to keep the data for distinct SYSIDs separate.
=*SysId*	Only records for the specified SYSID are written to disk

For example, if you received a tape with data for SYSA, SYSB and SYSC, you would need to make three separate passes:

- one with SAVE=SORT=SYS**A** producing the first SMFSAVE dataset, then
- another pass with SAVE=SORT=SYS**B** producing the 2nd different SMFSAVE dataset •
    and a final pass with SAVE=SORT=SYS**C**

By default the program will collect and sort the following record types:

TYPES=14,15,16,30(2,3,4,5),42(6),70(1,2),71(1),72(3),
TYPES=73(1),74(1,4,5,8,9,10),75(1),78(3),89(1,2),110(1),113(1)

You may change this list adding or subtracting record types. This is the list of all records that the Extract Program can process so if you are not interested in using one particular record type you would include the TYPES statement shown above but leave out that one type.

Because the statement may exceed the 80 character input line length the record types can be specified as two separate statements (and the content of both will be combined by the program).

For example, if you are pressed for space and know that you will not be processing the type 30 records you could place this statement in the input stream and remove the “30(2,3,4,5)” part. Then the output dataset would not contain type 30 records. This is the only other statement permitted with the SAVE=SORT= parameter.

3.7 Summary Output

If you use a SysId that doesn’t match anything in the incoming SMF data the program skips its regular processing and instead prints a summary of the contents of the SMF dataset. (Note that you can force this by deliberately specifying SYSID=xxxx). Here is an example of that output:

```
******* zCP3000 EXTRACT PROGRAM   10/04/18 V3.86   ******
      RUNDATE=10/04/18  RUNTIME=15:31:28.98
      SMFDSN=XXXXXX.SMF234.TEST
      ------ PROCESSING SYSIN001 DATASET ---->
ENT='Customer Name'
SYSID=XXXX
<----------- SYSIN001 DATASET PROCESSED
CP2KEXT: NO BCUMAP PROCESSING REQUESTED
CP2KPR10: NO PGNMAP PROCESSING REQUESTED
CP2KT30M: NO T30M001  PROCESSING REQUESTED
CP3KEXTR: NO DATA FOUND TO PROCESS
CP2KUTIL: CEC 2097-742     S/N (LAST 4 DIGITS): 9F30
              LPAR AQFT     SYSID AQFT    GMT OFFSET  -4:00
              LPAR VICTEST
              LPAR VMTOOL1
              LPAR AQCF1
              LPAR AQHO
              LPAR AQLINX
              LPAR HOCF4
              LPAR GDLVM7
              LPAR LNXVM14
              LPAR POKVMXA1
              LPAR PHYSICAL
CP2KUTIL SYSID: AQFT MVS: ZV010900 RMF: V7R1M9
CP2KUTIL RMF DURATION: 30 MINUTES
CP2KUTIL: GOAL MODE DATA. POLICY PRIMSHFT DESC PLPSC Production policy
    TYPE 002         1 RECORDS FROM: 06/23/18 09:10 - 06/23/18 09:10
    TYPE 003         1 RECORDS FROM: 06/23/18 09:26 - 06/23/18 09:26
    TYPE 070 01     10 RECORDS FROM: 06/20/18 09:00 - 06/20/18 13:30
    TYPE 070 02     10 RECORDS FROM: 06/20/18 09:00 - 06/20/18 13:30
    TYPE 071 01     10 RECORDS FROM: 06/20/18 09:00 - 06/20/18 13:30
    TYPE 072 03    360 RECORDS FROM: 06/20/18 09:00 - 06/20/18 13:30
    TYPE 072 04     10 RECORDS FROM: 06/20/18 09:00 - 06/20/18 13:30
    TYPE 073 01     10 RECORDS FROM: 06/20/18 09:00 - 06/20/18 13:30
    TYPE 074 01    370 RECORDS FROM: 06/20/18 09:00 - 06/20/18 13:30
    TYPE 074 02     10 RECORDS FROM: 06/20/18 09:00 - 06/20/18 13:30
    TYPE 074 03     10 RECORDS FROM: 06/20/18 09:00 - 06/20/18 13:30
    TYPE 074 04     20 RECORDS FROM: 06/20/18 09:00 - 06/20/18 13:30
    TYPE 074 05   1390 RECORDS FROM: 06/20/18 09:00 - 06/20/18 13:30
    TYPE 074 06     10 RECORDS FROM: 06/20/18 09:00 - 06/20/18 13:30
    TYPE 075 01     70 RECORDS FROM: 06/20/18 09:00 - 06/20/18 13:30
    TYPE 076 01    120 RECORDS FROM: 06/20/18 09:00 - 06/20/18 13:30
    TYPE 077 01     10 RECORDS FROM: 06/20/18 09:00 - 06/20/18 13:30
    TYPE 078 02     10 RECORDS FROM: 06/20/18 09:00 - 06/20/18 13:30
    TYPE 078 03     20 RECORDS FROM: 06/20/18 09:00 - 06/20/18 13:30
```

4. JCL

Modify the sample JCL in this document to meet the requirements of you installation.

4.1 DD Statement Summary

DD Name	*Purpose*	*Required*
SYSIN001	Input - parameters to the Extract program.	Always
SMFIN	Input - SMF dataset	Always
EDF001	Output - EDF file. Logical record length is 80	Always
PRINT001	Output - CP3KEXTR messages	Always
BCU001	The BCU mapping. LRECL is 80.	
PGN001	The service class mapping file. LRECL is 80.	
DATA001	Dataset for extracts of type 30, 42, 14, & 15 records	for zBNA
T30M001	The workload decomposition group mapping to be used with type 30 record processing. LRECL is 80.	
IOCP001	The IOCP dataset, also known as the IOCDS (I/O Control Dataset). LRECL is 80.	
SYSLOUT	LOADER messages	
SORTMSGS	Sort Messages	if SORT=YES
SORTWK0x	Sort Work Space	if SORT=YES
SMFSAVE	SMF Records to be saved for analysis in future runs	if SAVE= is specified

The next page has example JCL for CP3KEXTR that will produce output for all CPS Tools. It is contained in the "JOB" member of CPSTOOLS.JCL. Make these adjustments to customize it:

- Change the JOB card to meet the requirements of your installation
- Adjust the SMF SET command to specify the input SMF data set
- Adjust EDF, DAT, and TRS SET commands.
- In the SYSIN parameters, specify the SYSID and ENTerprise name to use for the run.

This example JCL is in the "JOB" member of CPSTOOLS.JCL and will produce output for all CPS Tools including zBNA. If the Extract run is for zCP3000 or zPCR the DATA001 DD statement and TERS step can be omitted.

4.2 JCL to Run the Extract

This example JCL is in the "JOB" member of CPSTOOLS.JCL and will produce output for all CPS Tools. If the Extract run is for zCP3000 or zPCR the DATA001 DD statement and TERS step can be omitted.

```
//#USERID  JOB
(????,????),MSGLEVEL=1,MSGCLASS=O,NOTIFY=????????
//*
//* THESE SET PARAMETERS MUST BE GIVEN VALUES
//*
//  SET SMF=XXXX.XXXX.XXXX             SMF INPUT
FILE
//  SET EDF=XXXX.XXXX.EDF              EDF FILE
OUTPUT FOR ANY CPS TOOL
//  SET DAT=XXXX.XXXX.DAT              DAT FILE
OUTPUT FOR ZBNA OR ZMCAT
//  SET TRS=XXXX.XXXX.TRS              TERSED
VERSION OF OUTPUT DAT FILE //*
//*------------------------- REMOVE FILES THAT WILL
BE REALLOCATED  //     EXEC PGM=IEFBR14
//DD1   DD
DISP=(MOD,DELETE),UNIT=SYSDA,SPACE=(TRK,1),DSN=&EDF
//DD2   DD
DISP=(MOD,DELETE),UNIT=SYSDA,SPACE=(TRK,1),DSN=&DAT
//DD3   DD
DISP=(MOD,DELETE),UNIT=SYSDA,SPACE=(TRK,1),DSN=&TRS  //*
//*------------------------- RUN THE EXTRACT PROGRAM
//EXTR    EXEC PGM=LOADER
//SMFIN    DD  DISP=SHR,DSN=&SMF
//*        DD  DISP=SHR,DSN=XXXX.XXXX     ADDITIONAL DATA SETS AS NEEDED
//*
//EDF001   DD
SPACE=(CYL,(10,100),RLSE),DISP=(,CATLG),
//
UNIT=(SYSDA,2),DCB=(RECFM=FB,LRECL=80),DSN=&EDF
//*
//* PGN MAP REQUIRES 72(3) RECORDS, IF NONE PRESENT COMMENT OUT PGN001
//* BCU MAP REQUIRES 74(1) RECORDS, IF NONE PRESENT COMMENT OUT BCU001
```

Run the extract and create all necessary data sets

these 4 SET

Add additional DD statements as needed for multiple SMF data sets

Temporary MAP files for Performance Group and BCU Map processing

```
//PGN001   DD  DISP=SHR,DSN=HLQ.CPSTOOLS.JCL(PGNMAP)
//BCU001   DD  DISP=SHR,DSN=HLQ.CPSTOOLS.JCL(BCUMAP)  //*
//* DATA FILE IS FOR ZBNA ONLY, MAY COMMENT OUT IF NOT A ZBNA RUN
//DATA001  DD  SPACE=(CYL,(200,500),RLSE),DISP=(,CATLG),
//             UNIT=SYSDA,DSNTYPE=LARGE,DSN=&DAT
//*
//PRINT001 DD  SYSOUT=*
//SYSOUT   DD  SYSOUT=*
//SYSLOUT  DD  SYSOUT=*                      MESSAGES FROM
MVS LOADER
//SYSLIN   DD  DISP=SHR,DSN=HLQ.CPSTOOLS.JCL(ZOBJEXTR)
//*
//*-------------------------- EXTRACT PROGRAM INPUT
PARAMETERS
//SYSIN001 DD  *
ENT='YOUR ENTERPRISE'      YOUR ENTERPRISE NAME, SPACES OK,
USE QUOTES
SYSID=XXXX                 SMF SYSID, 1-4 CHARS, NO QUOTES
PGN=GOAL                   GENERATE PGNMAP
BCU=AUTO                   GENERATE BCUMAP
SORT=YES
/*
//*---------------- TERSE THE DATA FILE TO MINIMIZE NETWORK BANDWIDTH,
//*                 MAY REMOVE THIS STEP IF NOT
ZBNA RUN
//*                 RC=10 IF NO RECORDS PRESENT
//TERS    EXEC PGM=TRSMAIN,PARM=SPACK
//SYSPRINT DD  SYSOUT=*
//INFILE   DD  DISP=SHR,DSN=&DAT
//OUTFILE  DD  DISP=(,CATLG,DELETE),SPACE=(CYL,(50,500),RLSE),
//             UNIT=SYSDA,DSN=&TRS
```

Output **DAT** file for zBNA

SMF records for this SYSID will be processed

Step to compress the output DAT file

4.3 Return Codes

These are the Extract Program return codes. Non-zero return codes can be found in the SYSOUT files along with the message generating the non-zero return code.

0	No notifications
4	Messages of note have been issued. Check the message, but usually the EDF output is OK.
8	A warning that messages have been issued that you should review. Either important SMF/RMF record types were missing from the input file or there is questionable data in some of the intervals (that you may want to exclude). EDF output will be usable but adjustments may produce better results.
12	Something is critically wrong with the parameters or no data fit within the SYSID and timeframe requested.
16	Something is drastically wrong and this run has failed.
20	CP3KEXTR couldn't write to the PRINT001 DD.
225	No data was found that meets the input specification.

4.4 JCL for SORT

This JCL is required to run RMF Post Processor. The MODS= statement is required only for the post processor.

```
//WICKSS JOB (????,????),MSGLEVEL=1,MSGCLASS=O,NOTIFY=WICKS,
//           REGION=4000K
//SORT    EXEC PGM=ICEMAN                        RUN DFSORT
//SYSOUT    DD  SYSOUT=*
//SYSPRINT  DD  SYSOUT=*
//SORTMSGS  DD  SYSOUT=*
//*
//SYSLMOD   DD   UNIT=SYSDA,SPACE=(3600,(20,20,1))
//SYSLIN    DD   UNIT=SYSDA,SPACE=(80,(10,10))
//SORTLIB   DD   DSNAME=SYS1.SORTLIB,DISP=SHR
//SYSUT1    DD   UNIT=(SYSDA,SEP=(SORTLIB,SYSLMOD,SYSLIN)),            X
//             SPACE=(1000,(60,20))   //*
//SORTIN  DD DISP=SHR,DSN=SYS1.MANDATA
//SORTOUT DD DISP=OLD,DSN=WICKS.RMF.CTS51
//*ORTOUT DD DSN=&&SORTED,DISP=(NEW,PASS),
//*       SPACE=(CYL,(100,50)),UNIT=SYSDA
//SORTWK01 DD DSN=&&SORT1,SPACE=(CYL,(125,125)),UNIT=SYSDA
//SORTWK02 DD DSN=&&SORT1,SPACE=(CYL,(125,125)),UNIT=SYSDA
//* SORT  BY    DATE, TIME, RECTYPE
//*
//SYSIN DD *
 SORT FIELDS=(11,4,CH,A,
7,4,BI,A,
               6,1,BI,A),
      EQUALS,
      FILSZ=E400000
  MODS E15=(ERBPPE15,500),E35=(ERBPPE35,500)    /*
```

4.5 JCL to Create Multiple Output Files in One Run

The program has the ability to create multiple output EDF files from a single pass of the SMF input data. For example, you could create an EDF for prime shift, second shift, and all shifts with a single pass of the data, creating three EDF files with different time parameters in a single run.

To do this you will use two different groups of DD statements:

- DD statements used for all “subruns”
- DD statements specific to a particular subrun.

An example of the JCL to accomplish this task is on the next page.

4.5a DDs used for all “subruns”

These are specified once and apply to all subruns.

SORTMSGS	Sort Messages. Required for SORT=YES
SORTWK0x	Sort Work Space. May be required for SORT=YES
SMFIN	This is the input SMF dataset. All subruns will process this same data.
SMFSAVE	SMF Records to be saved for analysis in future runs. Required if SAVE=SORT= is specified
SYSLOUT	LOADER messages

4.5b DDs Specific to a Particular Subrun

These DD statements are associated with each separate subrun. Each DD name ends in a three-digit number the program uses to correlate them. I.e., all the DD statements related to subrun 1 end in “001”, the ones for subrun 2 end in “002”, etc.

SYSIN001	Input parameters to the Extract program for a subrun
EDF001	Is the EDF output file. Logical record length is 80
PRINT001	CP3KEXTR messages
BCU001	The BCU mapping. LRECL is 80.
PGN001	The service class mapping file. LRECL is 80.
DATA001	Dataset for extracts of type 30, 42, 14, & 15 records
T30M001	The workload decomposition group mapping to be used with type 30 record processing. LRECL is 80.
IOCP001	The IOCP dataset. LRECL is 80.

The program will load the parameters for all of your subruns. If errors are found in any of the inputs the run will stop. If all parameters are successfully processed the PGNMAPs and BCUMAPs will be generated and processed. Assuming no errors are found in this stage the program will continue. The input parameters will apply only to the “subrun” in which they are specified with one exception -- “SORT=YES” will cause the input to be sorted for all “subruns”.

4.5c JCL for Multiple Outputs

This is an example of a job with 3 subruns.

```
//VS1PASS JOB (????,????),MSGLEVEL=1,MSGCLASS=O,NOTIFY=VSPENCE,          //
REGION=64M                                                     //*
//* THIS IS JCL TO RUN THE EXTRACT PROGRAM
//* TO PRODUCE MULTIPLE EDF OUTPUT FILES
//* IN A SINGLE RUN
//*--------------------
//EXTR     EXEC  PGM=LOADER
 //*                  PRINT DATA SETS FOR MESSAGES FROM YOUR RUN
//PRINT001  DD   SYSOUT=*
//PRINT002  DD   SYSOUT=*
//PRINT003  DD   SYSOUT=*
 //*                  INPUT SMF DATASET
//SMFIN     DD   DISP=SHR,DSN=A11PTG.ACXIOM.SMF                    //*
//*---------------- OUTPUT DATASETS FOR EDFI FILES
//EDF001    DD   DISP=SHR,DSN=VSPENCE.CPSTOOLS.JCL(EDF1S)          One EDF output
for
//EDF002    DD   DISP=SHR,DSN=VSPENCE.CPSTOOLS.JCL(EDF2S)             each subrun
//EDF003    DD   DISP=SHR,DSN=VSPENCE.CPSTOOLS.JCL(EDF3S)  //*
//*---------------- OUTPUT DATASETS FOR WORKLOAD GROUPING DATASET
//PGN001    DD   DISP=SHR,DSN=VSPENCE.CPSTOOLS.JCL(PGNEDF1S)
//PGN002    DD   DISP=SHR,DSN=VSPENCE.CPSTOOLS.JCL(PGNEDF2S)          //PGN003
DD   DISP=SHR,DSN=VSPENCE.CPSTOOLS.JCL(PGNEDF3S)            //*
//*---------------- OUTPUT DATASETS FOR AUTO BCU MAP
//BCU001    DD   DISP=SHR,DSN=VSPENCE.CPSTOOLS.JCL(BCUEDF1S)
//BCU002    DD   DISP=SHR,DSN=VSPENCE.CPSTOOLS.JCL(BCUEDF2S)
//BCU003    DD   DISP=SHR,DSN=VSPENCE.CPSTOOLS.JCL(BCUEDF3S)          //*
//*---------------- COMPAT MODE PROCESSING OF IEAIPS TO GENERATE PGNMAP
//*IPS001   DD   DSN=SYS1.PARMLIB(IEAIPS00),DISP=SHR
//*                    INPUT DATASET FOR IOCP PROCESSING (FICON)
//*IOCP001  DD   DISP=SHR,DSN=PATTRSN.CPSTOOLS.JCL(IOCP)              //*
//*SORTMSGS DD  SYSOUT=*
//SYSOUT    DD  SYSOUT=*                                              //*
//SORTWK01  DD  UNIT=SYSDA,SPACE=(CYL,(200))
//SORTWK02  DD  UNIT=SYSDA,SPACE=(CYL,(200))
//SORTWK03  DD  UNIT=SYSDA,SPACE=(CYL,(200))
//SORTWK04  DD  UNIT=SYSDA,SPACE=(CYL,(200))
//*-------------------- # OUTPUT DATASET FOR TYPE30 RECORDS
//*   NOTE:  DCB INFORMATION FOR THESE DATASETS IS SUBJECT TO CHANGE.
//*          SO IT IS BEST TO LET THE PROGRAM ALLOCATE THEM.
//*TYP30001 DD   SPACE=(CYL,(20,10),RLSE),DISP=(,CATLG),UNIT=SYSDA,
//*            DSN=PATTRSN.TYPE30.DATA
//* # OUTPUT DATASET FOR SAVE= PARAMETER (SAVE CP2K'S SMF RECORDS)
//*SMFSAVE  DD   SPACE=(CYL,(20,10),RLSE),DISP=(,CATLG),UNIT=SYSDA,
//*            BLKSIZE=8192,DSN=PATTRSN.SMF.SAVE                          //*
# OPTIONAL USER DUMP DATASET
//*SYSUDUMP DD   SYSOUT=*
 //*
```

Input parameters for the **first** subrun

```
 //*-------------------
//SYSIN001  DD   *
ENT=SHIFT 1 DATA IN A SINGLE PASS     ###  YOUR COMPANY NAME
SYSID=SYSA                            ###  JES SYSID
DURATION=00:15                  RMF/SMF DURATION FOR GATHER
*                               DURATION >= RMF INTERVAL
TIME=(08:00-15:00)              CONTINGUOUS PERIOD OF INTEREST
 DATE=(03/28/20-04/03/20)              ###  CONTINGUOUS DAYS
PGN=GOAL
BCU=AUTO                                                        *SHOWACT=NONE
SHOWACT=NONE REDUCES OUTPUT SIZE.
SORT=YES                   IF RECORDS ARE IN SORTED ORDER               *
OMITTING SORT=YES WILL SPEED THINGS UP.
PROCESS=YES
```

Input parameters for

```
 /*
```

the

```
//*--------------------
```

second subrun

```
//SYSIN002  DD   *
ENT=SHIFT 2 DATA IN A SINGLE PASS      ###  YOUR COMPANY NAME
SYSID=SYSA                             ###  JES SYSID
DURATION=00:15                  RMF/SMF DURATION FOR GATHER
*                               DURATION >= RMF INTERVAL
TIME=(15:00-23:00)              CONTINGUOUS PERIOD OF INTEREST
 DATE=(03/28/20-04/03/20)               ###  CONTINGUOUS DAYS
PGN=GOAL
BCU=AUTO                                                        *SHOWACT=NONE
SHOWACT=NONE REDUCES OUTPUT SIZE.
SORT=YES                   IF RECORDS ARE IN SORTED ORDER               *
OMITTING SORT=YES WILL SPEED THINGS UP.
PROCESS=YES
 /*
```

Input parameters for the

```
//*--------------------
```

third subrun

```
//SYSIN003  DD   *
ENT=SHIFT 3 DATA IN A SINGLE PASS     ###  YOUR COMPANY NAME
SYSID=SYSA                            ###  JES SYSID
DURATION=00:15                  RMF/SMF DURATION FOR GATHER
*                               DURATION >= RMF INTERVAL
TIME=(00:00-08:00)              CONTINGUOUS PERIOD OF INTEREST
 DATE=(03/28/20-04/03/20)              ###  CONTINGUOUS DAYS
PGN=GOAL
BCU=AUTO                                                        *SHOWACT=NONE
SHOWACT=NONE REDUCES OUTPUT SIZE.
SORT=YES                   IF RECORDS ARE IN SORTED ORDER               *
OMITTING SORT=YES WILL SPEED THINGS UP.
PROCESS=YES
```

```
/*
//*-------------------- MESSAGES FROM THE MVS LOADER
//SYSLOUT  DD    SYSOUT=*
//* THE OBJECT MODULE OF THE PROGRAM
//SYSLIN    DD   DISP=SHR,DSN=PATTRSN.CPSTOOLS.JCL(ZOBJEXTR)
```

5. SMF Input Data Requirements

The following SMF record types and subtypes are used by the Extract program:

Record	Type	CPS Tool	Subtypes Used
14	Input DS Activity	For **zBNA**	
15	Output DS Activity	For **zBNA**	
16	DFSORT Statistics		
30	Common Address Space Work (accounting information)	For **zBNA**, and for T30MAP processing or creating the DATA001 dataset	**2** - Activity since previous interval end **3** - Activity for interval before step end **4** - Step totals **5** - Job termination
42	DFSMS Statistics (Data Facility Storage Management Subsystem)	For **zBNA** If TYPE42=YES also for **zCP3000**	**6** - DASD data set level I/O statistics
70	CPU activity record	Required, except for type 30 processing	**1** - CPU, LPAR, coupling facility data **2** - Data for cryptographic coprocessors
71	Paging Activity	Only for system paging statistics.	**1** (has only one subtype)
72	Workload Activity	for PGN=GOAL	**3** – Workload activity data
73	Channel Path Activity	for BCU=AUTO	**1** (has only one subtype)

74	Device Activity	for BCU=AUTO, 74.9 for **zBNA**	**1** - Device activity **4** - Coupling facility **5** - Cache subsystem **8** - Enterprise Disk System **9** - PCI Express Activity **10** - Extended EADM statistics
75	Page Data Set Activity	For auxiliary storage information	**1** (has only one subtype)
78	I/O Queuing Activity	for BCU=AUTO	**3** – I/O queuing & HiperPAV
89	Product usage data		**1** - Usage Data **2** - State Data
110	CICS Data		**1**
113	Hardware capacity	all tools	Contains hardware capacity, reporting, and statistics for IBM System z10 or later machines. **1** – Hardware data event counters

Note: the Extract Program will eliminate duplicate records in the input stream. (Occasionally customer procedures are such that records are duplicated in the SMF input stream). If records are discarded a message will indicate how many have been discarded.

Note: If you have CMF (Boole and Babbage's equivalent of monitor I) data it is successfully processed by Extract Program as well.

6. Map Files

There are three supplemental map files that can help provide additional control in processing.

Map	Use
BCU	Combine control units with identical channel configurations
PGN	Combine service classes for reporting
T30	Break out workloads in greater detail using type 30 records

6.1 BCU Map

BCU (basic control unit) is a term used in the Extract program to represent a physical control unit -- the “box” that may contain multiple LCUs (logical control units). The BCU map provides a mechanism to combine LCUs that are part of the same physical control unit. I.e., a number of LCUs of a single current technology control unit will be reported as only one control unit.

6.1a How the Program Builds the BCU Map

When you request that a BCU Map be built by specifying **BCU=AUTO** the processing is as follows (type 74 subtype 1 records are required to do BCU processing. If not present you should omit the BCU DD statement and the BCU=AUTO parameter).

1. If you provide a DD for the IOCP001 DD the associated dataset with **IOCP** data is read in. Normally the IOCDS is only needed for the FICON aggregation function.

2. Then in all cases the SMF data is read in.

   - Type 74 subtype 1 - for device information (required)
   - Type 74 subtype 5 - for control unit caches and names
   - Type 78 subtype 1 - for LCU to path information

3. Next combining of Disk LCUs into BCUs takes place. The first combining takes place using NED (node element descriptor) data. NED data is control unit/device data that a device sends to the host in response to a particular command. Control unit NED data is in 74.5 records and device NED data is in 74.1 records.

4. Using the NED data the program can determine what a physical box is (i.e., the BCU) and will combine all LCUs and channels for a given NED Id into that one BCU.

5. If BCU=FICONAG is specified only LCUs with identical channel configurations are combined based on NED data. This normally results in 2 BCUs for a shark control unit.

The result of this is that we should have essentially one BCU per box (unless we have used BCU=FICONAG). If BCU=COM has been specified additional combining will take place. Since we are combining control units with identical channel configurations it means that control units on copper channels where the control units are daisy chained are combined.

Also if BCU=COM is specified ESCON control units that share the same directors are combined. (This does not happen with the IOCDS option because channel information also includes the ESCON switch and port. These cannot be identical and are not combined.)

If there are 74 subtype 5 records the NED ID determines the type/model, manufacturer, and serial number. Here accurate combining is assured. Without 74 subtype 5 data there is device level NED data in the 74 subtype 1 records but it is at the device level. If they are all reporting the same for all devices on a control unit it is possible to correctly combine the devices.

The BCU map provides an address map for z/OS data which identifies the BCU characteristics. Three types of statements can be in the BCU map:

- TIMESTMP has environmental information and will be first
- BCUx where x is either blank, T or O defines the BCU characteristics.
- ADDRS defines address ranges.

It is important that the BCUx statement labels match across the different systems so that zCP3000 can correlate them. An example is in **part 4.3**. Here are the OpCodes and operands used in the BCU map file:

OpCode	BCU Operands	
TIMESTMP	Defines Environmental information,	Generated for:
	SYSID=	BCU=IOCP and BCU=AUTO. Required
	PART=	BCU=IOCP. The partition name in the IOCP definition
	IOCPDATE=	BCU=IOCP. The date from the IOCP definition if present
	IOCPTIME=	BCU=IOCP. The time from the IOCP definition if present
	IODFDATE=	BCU=IOCP. Last date of the IODF file from type 78 records
	IODFTIME=	BCU=IOCP. The time of the IODF file from type 78 records
BCU **BCUT** **BCUO**	The BCU identifier ends with: blank for DASD, **T** or Tape, and **O** for other. These opcodes **will each have a label** -- a character string starting in position 1 to uniquely identify this BCU across the enterprise. For shared DASD this string should be the same for each SYSID mapping.	
	CTYPE=	Required for RMF before 4.2.0. Device type, e.g. 3990-3
	CACHE=	Cache size if controller is cached. Defaults to 0
	NVS=	Non-volatile storage size if controller is cached. Defaults to 0

	CHPID=	The chpids used by this BCU
	SWCHS=	The switches corresponding to the CHPIDs
	LINKS=	The switch links corresponding to the CHPIDs
	CHPTYP=	The CHPID types corresponding to the CHPIDs
ADDRS	The address identifier includes these sub-fields:	
	AD=	The range of addresses defined by this line
	LCU=	The LCU number for these addresses
	RMF=	Y or N whether RMF data was found for this range of devices
	DEVT=	The device type (e.g. 3380K)

6.1b BCU Map – Sample Output with BCU=AUTO

```
         TIMESTMP SYSID=PRD0                                         00010001
VSPSSD0  BCUD  CTYPE=3990-2,CACH=0,NVS=0,                           X00020002
   MFR=HTC,SER=03111,                                               X00020003
   CHPID=(4C,4D,7C)                                                  00020004
         ADDRS AD=0200-0202,LCU=009C,RMF=Y,DEVT=33903                00020005
VDEVTB7  BCUD  CTYPE=9393-2,CACH=2048,NVS=8,                        X00030006
   MFR=IBM,SER=24325,PLNT=91,                                       X00030007
   CHPID=(50,51,52,53,6E,6F,80,81)                                   00030008
         ADDRS AD=D600-D63F,LCU=01AC,RMF=Y,DEVT=33903,SSID=0061      00030009
         ADDRS AD=D640-D67F,LCU=01B0,RMF=Y,DEVT=33903,SSID=0062      00030010
         ADDRS AD=D680-D6BF,LCU=01B4,RMF=Y,DEVT=33903,SSID=0063      00030011
         ADDRS AD=D6C0-D6FF,LCU=01B8,RMF=Y,DEVT=33903,SSID=0064      00030012
VDEVT29  BCUD  CTYPE=9393-2,CACH=1024,NVS=8,                        X00040013
   MFR=IBM,SER=23774,PLNT=91,                                       X00040014
   CHPID=(50,51,52,53,6E,6F,80,81)                                   00040015
         ADDRS AD=D700-D73F,LCU=01BC,RMF=Y,DEVT=33903,SSID=0071      00040016
         ADDRS AD=D740-D77F,LCU=01C0,RMF=Y,DEVT=33903,SSID=0072      00040017
          ADDRS AD=D780-D7BF,LCU=01C4,RMF=Y,DEVT=33903,SSID=0073      00040018
 ADDRS AD=D7C0-D7FF,LCU=01C8,RMF=Y,DEVT=33903,SSID=0074      00040019
```

6.2 PGN Map

The PGNMAP (Performance Group Name Mapping) is a work file for the WLM (workload management) description of the workload and service class (business units). (The earlier performance group name convention has given way to goal mode operation but the old name is still being used here).

The PGN Mapping mechanism includes a way for you to manually combine service classes for reporting. However, the newer zCP3000 application will handle this automatically so manual adjustments are usually not necessary.

6.2a PGN Map – Reporting Classes

Reporting classes are included in the EDF and can be viewed in zCP3000. Analysis of most aspects of workloads can be done with these.

Reporting classes can be used in place of a service class but be sure to understand how the reporting classes are set up. A reporting class must correspond to a given workload class. Since this does not always happen, problems can occur.

If you are going to make manual adjustments to the service classes the following steps outline a procedure to verify that you have good data. (An example is in **part 6.2b**).

1. Do one standard Extract run with **PGN=GOAL**. This will build a PGNMAP file.
2. Open the PGNMAP file and find the entries for the reporting performance groups – they will contain "**RPT=1**".
3. Change the reporting classes you want to "**RPT=0**". Then comment out the corresponding service classes.
4. Also comment out the workload called "**THE REST**" – it will be at the very end of the list of workloads.
5. Remove PGN=GOAL and run the Extract again.
6. Load the two files **separately** into zCP3000. If everything lined up correctly the capture ratio will be identical in each of the files. If this is not the case the reporting performance groups do not correspond directly with the service class.

Note: The reporting classes that have been updated in the PGNMAP file will not be available in the reporting class view in zCP3000. In GOAL Mode it is defined by the WORKLOAD name (R723MWNM) and the Service Class name (R723MCMN). The Workload and Service Class data is mapped to a unique description.

(PGN=GOAL) builds a description of WORKLOAD.SCLASS as a default. Each description defines a unique business unit. To combine data from different workload and service class, make their descriptions be the same.

In the example in **part 6.2b** both "SSYSTEM.SYSTEM" and "SYSTEM.SYSSTC" are combined into one unit called "STUFF".

A mapping for each SYSID can be provided or where a performance group maps to a business unit for all SYSIDs to be processed an "*" is permitted indicating any SYSID. One workload name can now appear across multiple SYSIDs in zCP3000.

The PRTY (priority) field is the relative priority of the workload. This is not so easy to establish in GOAL mode. CP3KEXTR will artificially construct a relative priority value from the R723CIMP (the importance) field. It will be 11-R723CIMP. It is an optional field.

The PGN mapping file contains the following parameters. Two lines are allowed for each. An asterisk for a workload name means put anything not yet processed into this bucket. (See "The Rest" in the example).

SYSID=	A character string to identify the SMF ID of the system image corresponding to this data. If this is valid for all images an '*' may be used.
PGN=	This is ignored. PGN=0 is provided by PGN=GOAL as a default.
PER=	An optional specification of performance group period. A specification of PER=* will group all periods. This parameter is optional.
PRTY=	Relative priority of the performance group. Parameter is optional and defaults to 0.
DESC=	A unique character string describing the business unit. Field length for CP3KEXTR is 48, but zCP3000 will only display the first 26 characters.
WORKLOAD=	A unique character string describing the workload name as it would have appeared in an RMF workload report. The field length is eight characters. If this is WORKLOAD=* then all data not already processed will be collected here (this should be the last entry in the PGNMAP file). See the example for “The Rest”.
SCLASS=	A unique character string describing the service class name as it would have appeared in an RMF workload report. Field length is 8 characters.

Example of a single PGN map entry:

```
SYSID=X001 PGN=0 PRTY=9    DESC='CICSTEST.FBTEST  '
 PER=*  WORKLOAD=CICSTEST SCLASS=FBTEST
```

6.2b PGN Map – Sample Using Goal Mode

```
*******************************************************************
*  SYSID= MUST BEGIN IN COL 1 TO INDICATE START OF RECORD
*  DESC= THE DESCRIPTION IS THE WORKLOAD NAME.SERVICE CLASS
*  PGN= IS UNUSED IN GOAL MODE (GOAL MODE IS WHEN WORKLOAD IS NOT BLANK)
*  PRTY= IN GOAL MODE IS 10-S723CIMP RELATIVE IMPORTANCE OF GOAL
*  SCLASS= IS SERVICE CLASS NAME S723MCNM
*  REPORT CLASSES ARE COMMENTED OUT
*******************************************************************
*
```

```
SYSID=X001 PGN=0 PRTY=9    DESC='CICSTEST.FBTEST
'
 PER=*  WORKLOAD=CICSTEST SCLASS=FBTEST
*
SYSID=X001 PGN=0 PRTY=8    DESC='TSOTEST.TSOT1
'
 PER=1  WORKLOAD=TSOTEST  SCLASS=TSOT
*
SYSID=X001 PGN=0 PRTY=8    DESC='TSOTEST.TSOT2
'
 PER=2  WORKLOAD=TSOTEST  SCLASS=TSOT
*
SYSID=X001 PGN=0 PRTY=7    DESC='TSOTEST.TSOT3
'
 PER=3  WORKLOAD=TSOTEST  SCLASS=TSOT
*
SYSID=X001 PGN=0 PRTY=6    DESC='TSOTEST.TSOT4
'
 PER=4  WORKLOAD=TSOTEST  SCLASS=TSOT
*
SYSID=X001 PGN=0 PRTY=0    DESC='TSOTEST.TSOT*
*
'
 PER=*  WORKLOAD=TSOTEST  SCLASS=TSOT
```

```
SYSID=X001 PGN=0 PRTY=0    DESC='STUF
PER=*  WORKLOAD=SYSTEM   SCLASS=SYST
*
```

```
SYSID=X001 PGN=0 PRTY=0
DESC='STUF
 PER=*  WORKLOAD=SYSTEM
SCLASS=SYSS
*
SYSID=X001 PGN=0
PRTY=0    DESC='The
Rest'
 PER=*
WORKLO
AD=*
SCLASS
=any *  .....
```

These will be combined for summarization because the descriptions match

6.2c PGN Map – Sample Using Reporting Classes

```
******************************************************************
*  SYSID= MUST BEGIN IN COL 1 TO INDICATE  START OF RECORD
*  DESC, THE DESCRIPTION IS THE WORKLOAD NAME || SERVICE CLASS
*  PRTY IN GOAL MODE IS 10-S723CIMP RELATIVE IMPORTANCE OF GOAL
*  SCLASS IS  SERVICE CLASS NAME S723MCNM
*  RPT = 1 IS REPORTING PERFORMANCE GROUPS
******************************************************************  *
SYSID=PROD       RPT=0 PRTY=33   DESC='BATCHPRD.BATCHHOT    '
 PER=*  WORKLOAD=BATCHPRD SCLASS=BATCHHOT
*
SYSID=PROD       RPT=0 PRTY=21   DESC='BATCHPRD.BATCHPRD    '
 PER=*  WORKLOAD=BATCHPRD SCLASS=BATCHPRD  *
SYSID=PROD       RPT=0 PRTY=50   DESC='DDFDB2.DDFDB2.1      '  PER=1
WORKLOAD=DDFDB2   SCLASS=DDFDB2
SYSID=PROD       RPT=0 PRTY=34   DESC='DDFDB2.DDFDB2.2      '  PER=2
WORKLOAD=DDFDB2   SCLASS=DDFDB2
SYSID=PROD       RPT=0 PRTY=22   DESC='DDFDB2.DDFDB2.3      '
 PER=3  WORKLOAD=DDFDB2   SCLASS=DDFDB2
SYSID=PROD       RPT=0 PRTY=0    DESC='DISCRET.DISCRET      '
 PER=*  WORKLOAD=DISCRET  SCLASS=DISCRET  *

*YSID=PROD       RPT=0 PRTY=46   DESC='STC.STCDB2HI         '
*PER=*  WORKLOAD=STC      SCLASS=STCDB2HI                       *
                                                                SYSID=PROD       RPT=0
PRTY=54   DESC='STC.STCDB2LO         '
 PER=*  WORKLOAD=STC      SCLASS=STCDB2LO  *

*YSID=PROD       RPT=0 PRTY=54   DESC='STC.STCDB2MD         '
*PER=*  WORKLOAD=STC      SCLASS=STCDB2MD
```

```
*
SYSID=PROD       RPT=0 PRTY=50   DESC='TSO.TSO.1            '
 PER=1  WORKLOAD=TSO      SCLASS=TSO
*
SYSID=PROD       RPT=0 PRTY=255  DESC='SYSTEM.SYSTEM        '
 PER=*  WORKLOAD=SYSTEM   SCLASS=SYSTEM  *
*YSID=PROD       RPT=0 PRTY=254  DESC='SYSTEM.SYSSTC        '
*PER=*  WORKLOAD=SYSTEM   SCLASS=SYSSTC
*
SYSID=PROD       RPT=1 PRTY=21   DESC='.BATCHPRD            '
 PER=*  WORKLOAD=         SCLASS=BATCHPRD  *
SYSID=PROD       RPT=0 PRTY=46   DESC='.STCRTC              '
 PER=*  WORKLOAD=         SCLASS=STCRTC
*
SYSID=PROD       RPT=1 PRTY=50   DESC='.TSO.1               '
 PER=1  WORKLOAD=         SCLASS=TSO  *
*YSID=*    PRT=0 PRTY=0    DESC='THE REST                   '
*PER=*  WORKLOAD=*        SCLASS=*        RPT=0
```

6.3 T30 Map

There are cases where the breakdowns by workloads is not as granular we would like. In cases like these we can create a mapping using parameters in Type 30 (work unit accounting) records to break out workloads in greater detail than is possible with WLM Service Classes. To do this the customer must be collecting type 30 interval records (Subtype 2 and 3).

There are some differences between the workload data generated by using SMF Type 72 via a PGNMAP and using SMF Type 30 via a T30MAP.

- The T30 data does not include a lot of the SYSTEM information since SMF 30 is only recording address spaces started after SMF. That’s not too much of a problem for the business unit capture ratio but can be for the overall system CR. If you use T72 data, the SYSTEM work is often deleted anyway.

- The T30 data offers you the opportunity to break down the workload in zCP3000 to a much finer granularity. Often all of CICS is put into one service class which means that it will appear as one workload in zCP3000. You can split the CICS workload by using T30 data and jobname as the classification of work.

- You should remember to make the DUR parameter is larger than the type 30 interval.

This option will generate an additional set of workloads like those resulting from the PGNMAP with a different header in the EDF file. Once in zCP3000 you can only process one of these at a time. By default it will be the PGN001 data, but you can set the profile option “TYPE 30” in zCP3000 so it will read the T30M001 instead.

This file is very similar to the PGN map except there is no way to automatically generate one. Therefore, you will need to code this mapping file yourself. Refer to **part 6.3a** . The program looks through this map and the first match it finds will be the entry where the data is accumulated.

Each entry begins with SYSID= in column one and can be at most two lines long. It can have at most one field for job name or RACF userid etc. (**Note**: WLM= and SCN= must both be specified together.)

For example JOBN=PROD* or JOBN=TEST%% will collect either all jobs which begin with PROD or all jobs that begin with TEST and two characters. To combine these entries in the output the DESC fields for both of these should be the same.

A mapping for each SYSID can be provided or where a performance group maps to a business unit for all SYSIDs to be processed, an "*" is permitted indicating any SYSID.

One workload name can now appear across multiple SYSIDs in zCP3000.

The PRTY (priority) field is a relative priority of the workload. You will need to somehow arrive at values for the different business units. It is an optional field.

If a type 30 record is found which does not satisfy the criteria of one of the entries, the program will create a new entry for this data. This entry will have a description of "workload name.service class name".

The PGN mapping file contains the following parameters. Two lines are allowed. SYSID in column 1 begins the entry.

SYSID=	A character string to identify the SMF ID of the system image corresponding to this data. If this is valid for all images, an '*' may be used.
PRTY=	Relative priority of the performance group. Parameter is optional and defaults to 0.
DESC=	A unique character string describing the business unit. Field length for CP3KEXTR is 48, but zCP3000 will only display the first 26 characters. Workload descriptions must be unique for each SYSID.
JOBN=	The job name indicated in the type 30 record.
CLASS=	The job class indicated in the type 30 record.
USER=	The RACF userid in the type 30 record.
RACF=	The RACF group in the type 30 record.
PROG=	The program name in the type 30 record.
WLM=/SCN=	The workload manager name and the service class name. Both must be specified. An example is shown in **part 6.3a**.

6.3a T30 MAP – Sample Data Set

```
SYSID=ASYS    DESC='AP'   JOBN=PAP*
SYSID=ASYS    DESC='BE'   JOBN=PBE*
SYSID=ASYS    DESC='GL'   JOBN=PGL*
SYSID=ASYS    DESC='PC'   JOBN=PPC*
SYSID=ASYS    DESC='PO'   JOBN=PPO*
SYSID=ASYS    DESC='TR'   JOBN=PTR*
SYSID=ASYS    DESC='REST' JOBN=*
```

7. “DAT” Output File

The Extract Program can create the DATA001 file containing selected fields from SMF record types **30**, **42**, **14**, **15, 16,** and **120**. This is a flat text file used by the **zBNA** program. Note that the description of output file formats has been moved to a separate "Technical Reference" manual (that can be found at the same location as this User's Guide).

When these records are encountered in the SMF input the program checks for a DATA001 DD statement. If present the SMF fields are extracted and written to this file. Some of the same data may also be included in the EDF file.

The first character of each record is a number that indicates the data type (30, 42, 14, or 15) and all records and self defining records will start with this same code.

code	SMF type	contains
0	**info**	Extract run identification data: SMF dataset name, RunDate, RunTime, Extract Version
1	**30**	Accounting information
2	**42**	DFSMS (Data Facility Storage Management Subsystem) statistics and configuration
3	**14,15**	Dataset activity – Input (14) & Output (15)
4	**120 (12)**	Java batch details
5	**16**	LSORT information

Each line for a self-defining field entry includes the following fields, separated by commas:

code	The record type code (1,2, or 3). Always position 1.
*****	Constant – marker for self-defining entry. Always position 2.
NAME=	The name of the field
N or C	Whether the output is numeric or character. (Note the output in the file will always be in character format.)

COLS=	The column in the output in which this field will start
COLE=	The column in the output in which this field will end.

Here is an example of the self defining record data:

```
1*NAME=SMF30DTE,N,COLS=1,COLE=6
1*NAME=SMF30TME,N,COLS=7,COLE=12
     1*NAME=SMF30STP,C,COLS=13,COLE=13
```

By default only type 30 subtype 4 records (step totals) and subtype 5 records (job totals) are processed. However by using an additional parameter, “SUB30=2,3,4,5” you may specify which other subtypes are to be included.

For example, SUB30=2,3 indicates that you want just subtypes 2 and 3 and will exclude subtypes 4 and 5 (i.e., specifying the SUB30 parameter overrides the default subtype 4,5). The subtype of the record is included in the output so the post processing program can know the source.

8. EDF Output

The EDF (Enterprise Data File) is the primary output of the Extract program. It is a flat text file that contains the summarized information drawn from the SMF scan. It is an input to other analysis applications like zBNA, zCP3000, and zPCR.

Note: The list describing the EDF vector names has been moved to a separate "Technical Reference" manual (that can be found at the same location as this User's Guide).

8.1 EDF Format

The EDF is a text file with 80-character fixed-size records divided into named sections. A section name begins in the first position of a line while subsequent section content always starts in position 7.   Each record is organized as follows:

Column

1-4	Section name.
5	"S" when a section is specified. (This value will change when zCP3000 uses this format for output.)
6	Always blank
7-72	Information in the form KEYWORD=data appropriate to the section.

The section content consists of a series of expressions, one for each target SMF field. An expression has a key name followed by an equal sign followed by one or more values. Some have just a single value but more often there will be a series of multiple values (a vector) that is made up of the key name, the equal sign, and the following sequence of items:

- Average of the individual values
- Minimum value
- Maximum value
- Standard deviation of the individual values
- Count of values: the number of individual sample values (see below)
- List of individual sample values – one for each reporting interval (see below)

The first 5 items, the summary values, are always present in a multi-item vector. If the individual sample values are not all the same then the list of individual sample values will be present and will follow the count.

But if all the sample values are the same the list of individual sample values will be suppressed. In this case the minimum and the maximum will be equal and the count will indicate the number of sample values but the sample values will not be included. Just the 5 summary numbers will be present.