The History of SMF

A. It uses volume and data set information from the SMF data set to
create and update records in the inventory data sets. There is one
inventory for direct access resources and one for tape.

B. From volume information in the inventories it produces a volume
inventory report and SCRATCH volume reports.

C.   From data set information in the inventory it produces a variety of
data set reports. The data set inventory records are selected and
arranged   according to such parameters as data set name, volume
indentification, account number, expiration date and last usage date.

D. It creates control cards for IEHPROGM to SCRATCH data sets which
have expired.

E.   It creates control cards for the Resource Management Utility to
REMOVE data set inventory records from the inventories for data sets
which have been SCRATCHed or DELETEd."

"F.   It accepts control cards to remove volume and data set information
from the inventories.

G. It can compress the inventories.

Each inventory is a partitioned data set (PDS). Each directory entry in
the PDS reflects a particular volume.   The member itself contains
information about data sets on that volume.

This was followed by twelve pages identifying each field in each
inventory record and its source SMF record (5,14,15,17,18,19, or 20)."

and then the JCL:

"2.1.5.6 Job Control Language and Control Statements Used by the Utility

The Resource Management Utility can be invoked by the following job
control language


```
//jobname  JOB   positional parameters and
                   keywords
//stepname EXEC  PGM=utility name
//SYSPRINT DD    parameters describing
                  SYSOUT device
//SYSDATIN DD    parameters describing direct
                  access inventory PDS
//SYSTAINV DD    parameters describing tape
                  inventory PDS
//SYSMAN   DD    parameters describing the
                  SMF data set
//SYSREPRT DD    parameters describing report
                  device
//SYSPUNCH DD    parameters describing punch
```

```
                   data set
//SYSIN    DD    parameters describing control
                 statement data set "
```

"The control statement data set (described on the SYSIN DD statement)
contains one or more of the following operations:

```
  A. UPDATE
  B. REPORT
  C. REMOVE
  D. COMPRESS"
```

Then followed twelve pages detail the syntax and use of the many
operands for each operation.

The sixteen error messages produced by the Resource Managment Utility,
IEH901E through IEH916I are then documented, and give the first clue as
to the planned name for the Resource Management Utility program name of

                         IEHMAN!

"2.3.2.2 Resource Management Utility Requirements

The utility program is designed to be in a planned overlay structure
with a minimum of 15K required for executable code at one time. In
addition, core storage is required for buffers.   The number of bytes
needed for buffers is device dependent. Maximum buffer sizes (in bytes)
are shown in the table below:

Device	Type	Maximum buffer size
2301	Drum	21,000
2302	Drum	5,400
2303	Drum	5,300
2311	Disk	4,000
2314	Disk	7,900
2321	Data Cell	2,400"

And even a performance evaluation of the utility:

"The direct access inventory has entries for ten volumes, each with nine
extensions, giving a total of 100 members in the directory. Each member
is 3500 bytes long and contains 15 data set inventory records, giving a
total of 1500 data set entries.

An UPDATE is estimated to take 9 seconds.

A REPORT VOLID operation, which produces a list of direct access volume
inventory information is estimated to take 1 second.

A REPORT DS operation, which produces an unordered list of all data sets
on the printer, is estimated to take 82 seconds. Further time would be
required to produce a sorted list.

A prototype of the utility was coded and tested in several different versions. MVTTRACE was used to obtain timings of the final version described below.

SYS1.MAN data set resided on a 9-track 2400 tape drive, model 3, and was recorded at 800 bpi. The direct access inventory resided on a 2311 device with a member blocksize of 3200 bytes. The printer was a 1403 device with 120 print positions. The CPU was a model 50, the system MVT, Release 14."

"The first UPDATE run required 12.2 minutes. This initialized the inventory data sets.

A. The input stream from SYS1.MAN contained the following records arranged as if coming from an MVT environment:

```
      60 job commencement           (type 20)
      60 job termination            (type 5)
    1500 old non-temporary          (type 14)
    1500 new non-temporary          (type 15)
      60 temporary                  (type 16)
      60 SCRATCH                    (type 17)
      20 RENAME                     (type 18)
      80 direct access volume dismount (type 19)
```

The resulting direct access inventory contained information about 1862 data sets, of which 39 had been SCRATCHed. There were 197 directory names in the directory, of which 148 were base members and 49 were extension members. The average number of data sets per volume was 12.6. The average length of data set inventory records was 208 bytes.

B. The second UPDATE required 19.7 minutes."

In a separate section of the Final Specification the error expectations of SMF were predicted:

"6.4.2.1 APAR records for IEBPRTPCH, a utility program estimated to be of a similar size (13,000 bytes) were used to project expected errors and severity.

6.4.2.3 All Severity 1 errors should be resolved before integration. Only two OS utility programs have ever received Severity 1 APAR's after release.

and the accompanying table showed

	After Customer Ship	
Expected:	6 months	18 months
Severity 2, 3, 4 errors	7	15
Man hours to correct	40	40
Machine hours to correct	5	5"

And then along came the Final "Final":

```
       ----------------------------------------------
                    S/360-OS-1010-01-Pok
            Final Programming Functional Specifications
                    IBM Operating System/360
              Data Set Management and D.A.  Space
                       Accounting (Subset 2)
                  Internally Dated June 25, 1969.
       ----------------------------------------------
```

This document was much thinner than its predecessor, and states on its
cover letter:

"This revision of the referenced specification contains modifications in
the following major areas.

1. The Resource Management Utility program description has been deleted.

        THUS DIED SMF Reporting.

This final specification also eliminated the type 16 (temporary data
set) by redesigning the type 14 and 15 records to what we now have.

In addition, design specifications that have departed from the initial
specifications were identified:

"1.3.2 From Initial Specifications: These specifications depart from
the Objectives and Initial Specifications in the following respects:

 a. No I/O device timing will be performed.

 c. The TCT of Subset 1 will be expanded to include information
    previously intended for a new control block (the IOCT). This will
    eliminate the need to modify the DEB."

The final error table was now modified

"   After Customer Ship
 Expected: 6 months   18 months

Severity 2, 3, 4 errors   5 (was 7)   20 (was 15)
Man hours required   4 (was 40)   10 (was 40)
Machine hours required   4 (was 5)   10 (was 5)"

            "Systems Management Utility – IEHMAN"

The IEHMAN Planning Guide describes the same features that were
specified in the SMF design for Subset 2, called the "Resource
Management Utility"!

Not only did the IEHMAN Planning Guide exist within IBM, but through an

error, a small number of copies were actually shipped from Mechanicsburg, Pa., to some IBM customer sites!

Once the error was discovered, IBM immediately recalled all copies!

Ken Plambeck was the author of the IEHMAN package.

The project had 10-12 people 12 hrs/day, but the product was killed when the manager could not get sponsors.

Even though IEHMAN was never announced nor ever released, it showed that IBM designers, even in 1966-1967, knew that analysis tools would be needed to exploit the SMF data.


```
                   -------------------------------------------------
                     In May, 1969, C28-6712-0, Planning for System
                      Mangement Facilities (62 pages) described
                     the version of SMF planned for Release 18.
                   -------------------------------------------------
                   but the real announcement of availability was:
                   -------------------------------------------------
                          IBM System 360/Operating System
                                 Release 18 Guide
                                   GC28-6718-1
                                    July, 1969
                   -------------------------------------------------
```

"System Management Facilities

    SMF is a new feature of the operating system, selectable at system
    generation in conjunction with the MVT configuration. SMF may not
    be specified in conjunction with Model 65 Multiprocessing.

SMF provides two basic functions:

  Collecting and recording system-, job-, and step-related data for each
  job processed by the system.

  Monitoring job processing via exits from the control program to
  installation-provided routines at   specific   points   during   the
  processing of each job."

Data Collection: SMF writes 13 types of records to a data set resident
on either tape or direct access. The records contain such information
as: system configuration, job and job step identification, CPU wait
time, CPU usage by each job and job step, and I/O device usage and main
storage usage by each job step.   The writing of SMF records may be
supressed at IPL. An installation-provided routine can periodically
analyze the SMF dataset to evaluate system efficiency, performance, and
usage.

Job Monitoring: SMF provides five exits within the control program to
installation routines:

  From the reader/interpreter of the job just before each job control

statement is interpreted.

From the initiator/terminator when a job is selected for initiation.

From the initiator/terminator when a step is selected for initiation.

From the intiator/terminator when a step and/or job is terminated.

From the timer second level interruption handler (SLIH) if a specified
CPU or wait time limit for a job or step is exceeded."

If no wait time limit is specified, the default value provided by IBM
is 15 minutes; in previous releases the default value was 30 minutes.

A new macro instruction (SMFWTM) may be used in exit routines to write
records to the SMF data set. A test procedure (TESTEXIT) is provided in
SYS1.SAMPLIB to facilitate routine testing.

You must also provide analysis and report routines to process the SMF
data set."


```
              --------------------------------------------
                    IBM System/360 Operating System
                         Consolidated Document
                             Release 18
                             GY28-6681-2
                      Third Edition (October, 1969)
              --------------------------------------------
```

"Updated Version of Release 18, designated as Release 18.6, may now be
ordered from PID.

Approximately 40 PTFs have been applied to the distribution libraries,
correcting more than 80 APARs.

(Release 18.0 text:)

A second release of SMF will support MFT, M65 Multiprocessing and Remote
Job Entry.

Graphics Job Processing (originally planned for the second release of
SMF) is supported in this initial release of SMF.

Primary Control Program (PCP) and Conversational Remote Job Entry (CRJE)
are not supported.

Performance

If SMF is chosen at System Generation, performance will be reduced.

The performance reduction will be dependent upon the installation's use
of the following:

  . SMF buffer size
  . Device used for the SMF data set

. SMF data set allocation size
. Number of jobs run per day
. Execution time of the installation exits

The performance reduction to the system when all System Management
Facilities are functioning can be less than 5%.

Storage Requirements

A fixed amount of main storage is required when SMF is chosen at System
Generation. In MVT a maximum of 1500 bytes is added to the main
resident storage. Supervisor Queue Space is used for data collection
tables, new job queue entries, and the user defined SMF buffer. The
variable storage depends on the number of active jobs in the system and
the SMF options chosen."


```
             --------------------------------------------------
                    IBM System/360 Operating System:
                            Release 19 Guide
                              GC28-6733-1
                               June 1970
             --------------------------------------------------
```

Announced Subset 2 of SMF and support for MFT and M65 Multiprogramming
and RJE with all twenty-one SMF records (0-15,17-20) with one new
record,

  Type 13 - MFT Partitions

and one new Exit

  (IEFUSO) for SYSOUT data set limit exceeded


```
             -------------------------------------------------
                    System Programming Guide Release 19
             -------------------------------------------------
```

"Example of SYS1.MANX Space Requirements

		ID	bytes
1	IPL per day	0	
20	Devices online	8,19	
2	Vary Onlines per Hour	9	
2	Vary Offlines per Hour	11	
1	Device Allocation per Hour	10	
1	Scratch Dataset per 4 Hours	17	
1	Rename Dataset per 4 Hours	18	
24	Accumulated Wait Time	1	
	Total for these records		2,025
Job Processing			
1	Start of Job	20	
1	End of Job	5	

```
1  Dismount 2 Volumes per job     19
3  Steps per job                  4
Step Processing
3  DD statements per step         4
2  Input datasets per step        14
2  Output datasets per step       15
2  Output writers per step        6
 Total for one job                        6,303
 Total for 48 jobs in 4 hours             302,688
SMF headers
  Record Descriptor Words                 5,044
  Block Descriptor Words                  1,684

TOTAL SMF DATA FOR 4 HOURS:               311,411"
```

```
            --------------------------------------------
                 IBM System/360 Operating System:
                        Release 20 Guide
                          GC28-6730-0
                          January 1971
            --------------------------------------------
```

Announced 20.0 with support for S/360 and new S/370 155 processor (S/370 165 in April) and indicates to expect 20.6 in 6-8 months.

```
            --------------------------------------------
                 IBM System/360 Operating System
                   System Management Facilities
                          GC28-6712-4
                   (Fourth Edition, February 1971)
            --------------------------------------------
```

Applied to Release 20.1 and identifies the eleven new SMF records created with Release 20 (now totaling 31 record types):

```
 Type 21 - ESV Record
 Type 30 - Start TSO Record
 Type 31 - TIOC Initialization Record
 Type 32 - Driver Record
 Type 33 - Driver Modify Record
 Type 34 - TSO Step Termination
 Type 35 - TSO Logoff Record
 Type 38 - Initial TSO Configuration Record
 Type 40 - Dynamic Allocation Record (not documented!)
 Type 41 - Modify TSO Record
 Type 42 - Stop TSO Record
```

```
            --------------------------------------------
                 IBM System/360 Operating System:
                        Release 21 Guide
                          GC28-6730-2
                   March, 1972 (Release 21.0)
            --------------------------------------------
```

Announced the REC parameter and minor change in format of SMF records,
and indicated that 21.6 would be along in 4-6 months.

```
         --------------------------------------------
                IBM System/360 Operating System:
                       Release 21 Guide
                        GC28-6730-04
                  August, 1972 (Release 21.6)
         --------------------------------------------
```

was announced, with no SMF enhancements.

In October, 1972, I first used the Statistical Analysis System, SAS, to
read SMF records from OS/360 Release 20.6 at State Farm Insurance. At
that time, SAS cost $100! By early 1973, I reported our successful
results to user groups (SAM, TESDATA) and ACM (SIGSIM, Chicago). John
Chapman demanded that I present at SHARE.

In March, 1974 (SHARE XLII, Houston), in the CME project, I presented
SMF/SAS; CME scheduled an open session for the Chicago SHARE.

That summer, IBM announced SGP, their Statistics Gathering Package,
written by Bill Tetzloff. The open session at the August, 1974, SHARE
meeting in Chicago began with Bill's SGP product description followed by
State Farm Auto's presentation of their use of SAS and SMF data. Over
750 people (half of SHARE attendance) were present! While SGP was truly
a valiant IBM effort to provide reporting from an extract of a fixed set
of SMF fields, it was not easily tailored, was cast in concrete, and
appeared inflexible.   This audience then saw the SAS language for the
first time, and saw SAS actually used for productive SMF analysis.   At
the end, one attendee pointedly asked IBM, "Now that you have seen SAS,
is there any reason why we should still consider SGP?"

This discovery that the SAS language was highly suited for analysis of
SMF and RMF data led to many SHARE sessions that demonstrated that SMF
data analysis really did save money and answered data center
management's questions (response, capacity, service objectives, etc.).
SAS was the tool that made SMF analysis practical, in 1974.

The early releases of MVS became available:

```
 VS/2   Announced August 1972

 ?/73?  MVS 2.2 MF-1 Type 70-77,79
 ?/75?  MVS 3.0
 5/76   MVS 3.7
 8/76   SU7 Supervisor 2
 8/76   SU20 RMF
 9/76   SU11 TSO CMD. Package
 9/76   SU13 TSO/VTAM
 3/78   SU58 TSO/VTAM Level 2
 3/78   SU50 MVS SE1 (for 3.7)
 7/78   SU78 TSO Session Manager Rel. 2
```

```
3/79   SU65 MVS SE1.1 (for 3.7)
3/79   MVS 3.8 (SCP2)
8/79   SU74 MVS SE2 (for 3.8)
        Type 23, 30, 32, 90
1/85   NPDA V3 R2 (Type 37)
2/85   DFSORT R7 (Type 16)
3/85   NLDM R3 370 (Type 39)
7/86   DB2 (Type 100,101,102)
```

By the mid 1970s, large TSO shops (several hundred concurrent logged on users) began noting degradation due to the BSAM SMF writer:

 - ENQ/DEQ was used by all SMF records, a very slow, serialized server for every logical record written.

 - TCLOSE to update VTOC after every block was written moved the disk drive arm - the drive shook like a washing machine!

 - WRITE VERIFY doubled the I/O time

In addition, the 1975 SHARE-IBM meeting in New York was called because users were not migrating from MVT to MVS, and were blaming SMF accounting issues, technical issues about actual numbers in the records (eg., more absolute CPU time, the loss of "storage measurement" a la "K-Core Hours" from MVT, the inclusion of PCI counts into SMF EXCP buckets) caused, Ron Hankison, then IBM Representative to SHARE MVS Group, to become the local SMF expert within IBM.

From Ron Hankison 1989 interview:

Accounting requirements from SHARE and GUIDE had built up, and nothing had changed since the original OS implementation (VS 2.2 and VS 1.6 had only added a few fields regarding paging and Virtual storage).

TCLOSE was out of control, the constraint was apparent and there was a backlog of user requirements.

Project goals were to:

   Fix the performance constraint
   Add functionality
   Restructure after 5-10 years experience

Estimated project by doubling the then-known size of SMF (Source Lines of Code) and used that (and IBM internal estimates) for the actual estimated man-hours.

The final implementation took four times as much as the estimate, requiring 12 people 1.5 to 2 years for SE2 SMF Writer redesign.

Because no one else in VS 2 cared much about SMF, the development was independent, which allowed many things to be considered and many went in and out of the final SMF enhancements.

Primary developers were:

    Barb Butler
    Bill McTiernan
    Steve Rosengarn
    Joe Winterton

                            Programming Objectives and
                  Initial Programming Functional Specifications
                            MVS Accounting Facility
                                August 1, 1977

"1.4 Summary of Specifications

The rewrite of SMF will resolve the numerous, known problems with SMF.
The performance of SMF will be much improved by the elimination of
bottlenecks during the collection and writing of the data.   Additional
data will be available for recording that fills in several known
exposures for resource utilization and system activities.   Flexibility
will be built in that allows the installation improved control over the
recording and make tradeoffs between the integrity of the data collected
and   the   performance impact.   Complexity will be reduced by the
capability of establishing a common file to contain all the pertinent
information needed to manage the installation. Additional useability
items will make this package very appealing to DP installations."

1.5 Summary of RAS Characteristics

Due to the critical nature of the data being handled, minimizing the
opssibility of loss of data will be stressed in the design of this
package. The improvements described in this document will be covered by
either an FRR or ESTAE routine to ensure that loss of data is held to a
minimum in the event of a failure in the SMF component.   Optional
facilities will be provided to minimize loss of data due to system
failures. Programming techniques known to have demonstrated improved
quality will be used during implementation and test."

2. User Requirements Addressed
        Started Task Reporting
        Interval Reporting
        Performance/Overhead of Data Collection
        Recording Selectivity
        Proliferation of Data and Records
        Installation Tracking Completeness
        Accounting Direction
        Proliferation of Tools

7. Performance

The overall reduction of SMF overhead and its impact on system thruput
will be significant in those environments that have a high volume of
data that is recorded to SMF. The performance improvement of the
collection routine and its branch entry capability will provide a much
improved interface for those components that should be recording to SMF

but were concerned about the SMF bottlenecks.

The path length for a Write to the SMF data set is approximately 60,000 instructions. The frequency of that event is installation dependent but probably averages about once every 13 logical records. The size and frequency of record receipts will be increasing rapidly as processor speeds increase. The revised path length by using the VSAM ICIP in SRB mode is approximately 2500 instructions, resulting in a reduction of 57,500 instructions per Write."

        Summary of the author's opinions:

1. SHARE was instrumental in the creation of SMF.

2. Users had a fair idea of what data was needed.

3. IBM designers in 1966 knew more than users of the range of data that we would ultimately need.

4. The iterative process between IBM designers and IBM users provided realistic validation of the design before implementation.

5. Designer's specifications and wants exceeded the funding and time for implementation.

6. The 1966-1969 IBM design and implementation process was better structured, managed, and documented than your company's most recent managed software project in 1989.

7. Based on this example of IBM design practices of twenty years ago, imagine the detail we would find in today's IBM design documents!

8. SMF made IBM pre-eminent in the business of real data processing by giving DP management actual measures of the resource usage and the service (response) for each user. DP could then "manage by objectives" and prove to the president the value and costs of the services DP provided the company. No other vendor of hardware and software has provided the accurate measurements that IBM has given its customers.

9. As good as it is, it still isn't perfect:

   MVS/ESA added three important CPU measures to
   the type 30 (task level) record,
       RCT - Region Control Task CPU
       SLIH - Second Level Interrupt Handler CPU
       HPT - Hiperspace CPU
   but we do not have these same CPU measures
   in the type 72 (performance group) record.

   SHARE REQUIREMENT, ANY ONE?

Contributors:

With sincere thanks for the dedicated developers designers and users of

SMF data, I especially salute these contributors to this research:

Barry Merrill

Merrilly yours,

 Herbert W. Barry Merrill, PhD
 President-Programmer
 Merrill Consultants
 MXG Software
 10717 Cromwell Drive   technical questions: support@mxg.com
 Dallas, TX 75229
 http://www.mxg.com   admin questions: admin@mxg.com