Collecting CPU MF (Counters) on z/OS

The purpose of this document is to describe the steps necessary to enable CPU MF Counters.

Background

With the System z10 and later processors there is a new hardware instrumentation facility, the "CPU Measurement Facility" (CPU MF) which can help with workload characterization and performance analysis. There is also a z/OS component "Hardware Instrumentation Services" (HIS) to support the hardware instrumentation facility (CPU MF).

There are 2 major modes for running CPU MF, 1) Counters and 2) Sampling. **Counters** mode is used to count events for each logical processor such as Cycles used, Instructions, and where data and instructions were sourced from the cache/memory hierarchy. (**Sampling** is used to take a snapshot of the CP information such as instruction address, primary address space number, and state information. This allows an analysis to map instruction addresses into modules or tasks and facilitates determination of hotspots.) Note the two modes are mutually exclusive, and are determined by the various options of the Modify command. These enablement instructions are for you to **run Counters mode only**.

Requirements

In order to use the CPU MF and HIS, the following is requirements must be installed on your Processor and z/OS operating systems

- **Requirements:** The machine must be a z10 or later and z/OS should be at the currently supported levels. While this document describes the steps necessary for z/OS, z/VM also supports CPU MF Counters and a detailed description on how to enable it can be found at **http://www.vm.ibm.com/perf/tips/cpumf.html**

Enablement Steps

This section documents the 3 Steps needed to enable CP MF and HIS, and collect the SMF 113s records:

1. Step 1 – Configuring the Processor to collect CPU Measurement Facility Data
2. Step 2 - Configuring z/OS to collect CPU Measurement Facility Data
3. Step 3 - Collecting CPU Measurement Facility Data (Counters only)

© 2019 IBM Corporation

Step 1 - Configuring the processor to collect CPU Measurement Facility Counters Data:

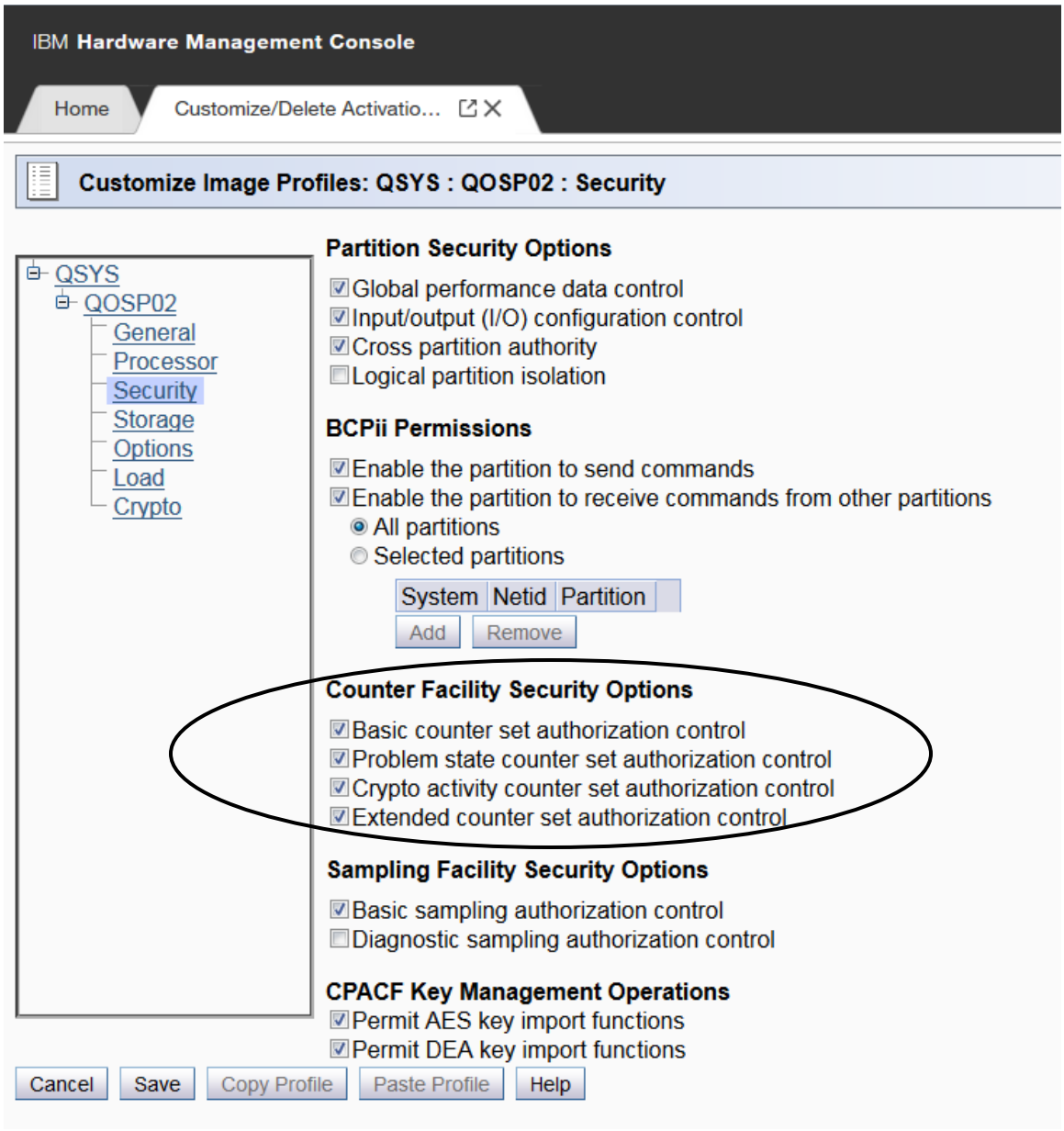
In the activation profile for the user's LPAR, access the security tab and verify that the following settings are checked. For Counters mode one can check Basic Counter, Problem State Counter, Crypto Activity Counter and Extended Counter within *Counter Facility Security Options.*

Counter Facility Security Options

- Basic counter set authorization control
- Problem state counter set authorization control
- Crypto activity counter set authorization control
- Extended counter set authorization control

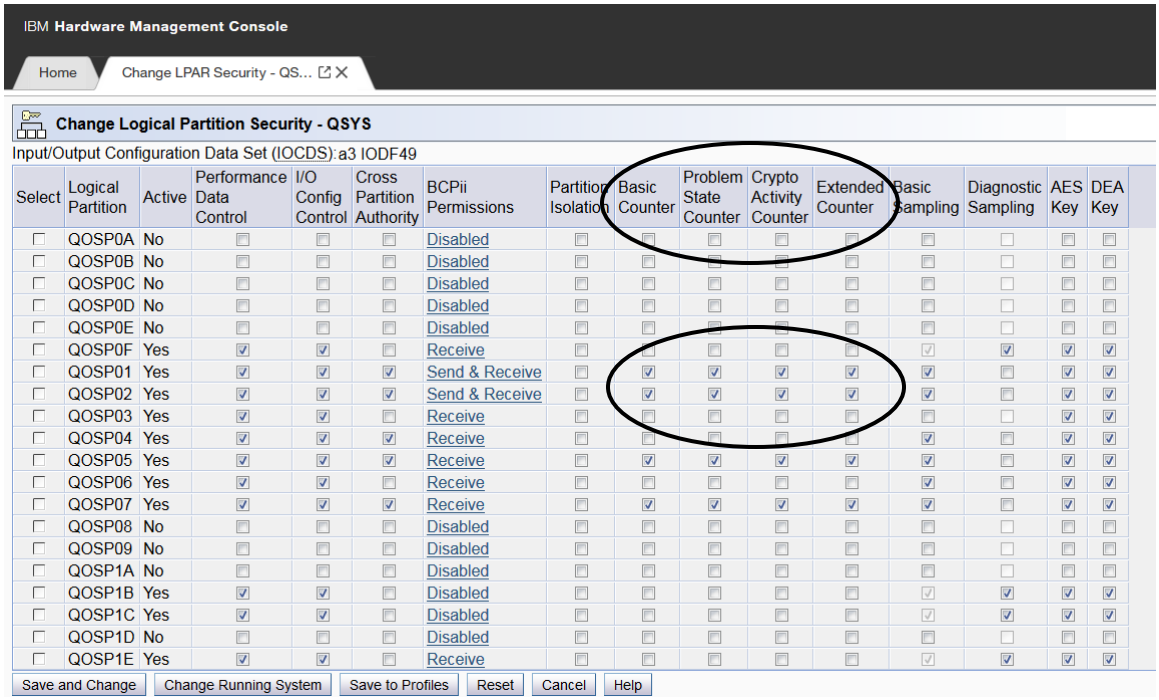
Here is a sample of the screen from the z14:

© 2019 IBM Corporation



If all the required settings are checked, proceed to activate the image. Otherwise, when the user attempts to start HIS Data Collection, the user will encounter an error message such as this one: **"HIS026I MODIFY HIS COMMAND CANNOT BE PROCESSED. BASIC COUNTER SET IS UNAUTHORIZED"**. You can also dynamically update the above security options from the HMC or on the SE via the Change LPAR Security icon if the LPAR(s) of interest is(are) already activated. Here is a sample of the screen from the z14:

© 2019 IBM Corporation



© 2019 IBM Corporation

Step 2 - Configuring z/OS to collect CPU Measurement Facility Data:

1. The z/OS operating system provides access to CPU Measurement Facility data through Hardware Instrumentation Services (HIS).

2. If you don't have a HIS member in SYS1.PROCLIB, copy the following procedure into the member, HIS, in your PROCLIB. This will be used to start Hardware Instrumentation Services (HIS). If you use ServerPac, the HIS procedure can be copied from ServerPac's PROCLIB.

```
//HIS  PROC
//HIS  EXEC PGM=HISINIT,REGION=0K,TIME=NOLIMIT
//SYSPRINT DD SYSOUT=*
//CMDFILE1 DD DUMMY
//CMDFILE2 DD DUMMY
```

3. Specify where to store the HIS output file: (optional step if OA43366 installed)

   - Starting in z/OS 2.1 and z/OS 2.2 with APAR OA43366 installed, the HIS output file is optional for HIS Counters, and the USS File System is not required. The SMF data will still be written. There is a new parameter "CNTFILE=NO" which will allow HIS to run without this USS File.
   - Define a userid for the HIS started task with an OMVS segment that specifies
     - any UID based on your naming rules,
     - a default home directory for output files.

     For example, you might define the HIS user ID as follows:
     ADDUSER HIS OMVS(UID(123) HOME('/HIS'))
     where **uid(123)** is the OMVS uid and **/HIS** is the default home directory. (Note that while OMVS access is required, no special authorization is needed. Also, any directory can be used for the HOME directory.)

   - Create the **/HIS** HOME directory (in a local file system) by issuing the following "**mkdir**" command under OMVS:
     "mkdir /HIS".

     Then 1) assign read/write/exec authority to the /user directory by issuing: **"chmod 777 /HIS**" or 2) assign write authority to the /HIS directory via "chmod 660 /HIS"

     In this case, **/HIS** will be the default directory where the HIS output file will be stored. (Even though Counters mode writes the data to SMF 113s records, this OMVS directory is still required. When HIS is finally

© 2019 IBM Corporation

stopped, it will write a very small amount of summary information into this directory, as a *.CNT file.

4. Counters data is captured in SMF 113s records. You need to ensure that you are collecting these SMF records as specified in your SMFPRMxx member.

   - Start collecting SMF Type 113 records by adding 113 to the SYS statement and if necessary to the SUBSYS(STC(TYPE statements in the SMFPRMxx member(s) for each LPAR for which data is being collected. SMF type 113, subtype 2 records are provided by the Hardware Instrumentation Services (HIS) support. z/OS 2.1 also captures subtype 1 records (deltas only) and we recommend that you capture these also.

   - Dynamically pick up the changes to the SMFPRMxx member(s) by issuing SET SMF=xx command on each LPAR of interest.

   - Verify that the SMF type 113 records are being collected by issuing D SMF,O command on each LPAR of interest. Look for output similar to the ones below:

     SYS(TYPE(0,2,3,6:10,14,15,22:24,26,30,32,33,41,42,47:48,59,60:69,70:79,80:83,85,88,89,90:91,94,98,100:103,108,110,113,115:117,130,134,148:151,161,200,244,245)) – PARMLIB

     SUBSYS(STC,TYPE(23,42,70:79,99:103,110,113)) -- PARMLIB

Step - 3 Collecting CPU Measurement Facility Data:

1. Issue the following z/OS Command to start the Hardware Instrumentation Services (HIS) Address Space :
   S HIS

   The following message indicates a successful start:

   HIS002I HIS INITIALIZATION COMPLETE

   If the correct Driver is not installed on the system, error message “HIS031I REQUESTED FUNCTION IS UNSUPPORTED BY HARDWARE” will be reported upon unsuccessful completion of the z/OS start HIS Address Space Command.

2. Once HIS has been successfully started, issue the Modify command to start Counters and begin collecting SMF 113s.

© 2019 IBM Corporation

Issue the following z/OS Command to start CPU Measurement Facility Data Collection. We recommend that CPU MF should be permanently enabled since it requires so little space for the SMF 113 records and has very low overhead. Using Counters, CPU MF writes a SMF 113 every 15 minutes (as the default). It is recommended to use the SI=SYNC parameter to set the interval the SMF Global Sync value. While not required anymore with APAR OA43366, an OMVS file will be written once CPU MF Counters have been ended. It is a very small amount of information. The recommendation is to run with Basic and Extended Counters, and to use the command below.

To start CPU Measurement Facility data collection for **Counters ONLY**:

Starting in August 2017 the recommendation is to also enable Crypto Counters, in addition to the Basic and Extended counter sets.

F HIS,B,TT='Text',CTRONLY,CTR=(B,E,C),SI=SYNC,CNTFILE=NO"

- The "**CTRONLY"** option indicates to HIS that ONLY Counters are to be collected.
- In this example, for **CTR=(B,E,C)**, will include 3 Counters sets, the "**B**" token indicates that the **B**asic counter set is to be collected, and the "**E**" token indicates that the **E**xtended counter set is to be collected, and the "**C**" token indicates the **C**rypto counter set is to be collected.
- SI=SYNC sets the SMF Interval to "Synch" with the SMF Global recording interval of your system, so that SMF 113s are written at the same time and interval as other SMF records.
- The CNTFILE=NO indicates that the USS file for HIS output will not be used (and thus the PATH='/HIS' parameter is not needed).

For z14s and z15s, the recommendation is to collect all counter sets.

F HIS,B,TT='Text',CTRONLY,CTR=ALL,SI=SYNC,CNTFILE=NO

- In this example the CTR=ALL will collect all counter sets (Basic, Problem, Crypto, Extended and MT-Diagnostic)

Successful start of CPU Measurement Facility data collection is indicated by the following message:

HIS011I HIS DATA COLLECTION STARTED

You an issue a Display HIS command to view the status and confirm the options

D HIS

© 2019 IBM Corporation

Here is an example of the output from the Display HIS command where HIS had been modified with this specific command

```
F HIS,B,TT='BEC',CTRONLY,CTR=(B,E,C),SI=SYNC,CNTFILE=NO

HIS011I HIS DATA COLLECTION STARTED

D HIS

HIS015I 13.15.43 DISPLAY HIS 559
HIS     002A ACTIVE
COMMAND: MODIFY HIS,B,TT='BEC',CTRONLY,CTR=(B,E,C),SI=SYNC,CNTFILE=NO
START TIME:  2017/11/02 13:15:15
END TIME:    ----/--/-- --:--:--
COMPLETION STATUS: --------
COUNTER VERSION NUMBER 1: 3   COUNTER VERSION NUMBER 2: 5
COMMAND PARAMETER VALUES USED:
 TITLE=  BEC
 PATH=   .
 DURATION= NOLIMIT
 DATALOSS= IGNORE
 STATECHANGE= SAVE
 SMP= NO
 CNT= YES
 COUNTER SET= BASIC,CRYPTO-ACTIVITY,EXTENDED
 CNTINTVAL= SYNC
 CNTFILE= NO
 MAP= NO
```

Note: If any set up errors are encountered such as an option error, then, data collection can be stopped by issuing an "F HIS,E" command and restarting data collection with the correct options.

Note: For more information about collecting HIS data, please see the manual, z/OS MVS System Commands, for details about using HIS commands. http://www.ibm.com/systems/z/os/zos/library/bkserv/v2r1pdf/index.html#IEA

You can run the IFASMFDP program to ensure that you are collecting SMF 113s in the SMF data sets.

3. If you need to stop collecting CPU MF Counters, issue the following z/OS command to stop CPU Measurement Facility data collection:

   F HIS,E

   When the data collection measurement is ended for Counters, HIS writes a summary file to the UNIX System Services data set in the '/HIS' path.

   Successful stop is indicated by a message similar to the following:

   HIS022I HIS DATA COLLECTION IS ENDING. 464

© 2019 IBM Corporation

OUTPUT FILE PREFIX: SYSHIS20080413.223225

The file which contains the Counters will have a name whose format is similar to this : SYSHIS20080413.223225.CNT **(For Counters ONLY)**

If you stop HIS when there was no USS file, CNTFILE=NO, then you'll see the following:

HIS022I HIS DATA COLLECTION IS ENDING. 920
OUTPUT FILE PREFIX:
TITLE= 'CPUMF Counter'
HIS012I HIS DATA COLLECTION ENDED

We recommend that once you'd enabled HIS and modified it to start data collection that you do not end it, and that you keep it collecting SMF 113s continuously.

If you have any questions, send a note to John Burg, jpburg@us.ibm.com

© 2019 IBM Corporation

Appendix

Questions

Q: Can the z/OS HIS maintenance be applied to a non z10 GA2 machine, such as a z9 and HIS started on that processor?

A: Yes, it can be applied and all the z/OS HIS steps completed in advance, including starting HIS. What will happen is that HIS will issue the following message and then end.

```
$HASP373 HIS      STARTED
IEF403I HIS - STARTED - TIME=21.09.36
HIS031I REQUESTED FUNCTION IS UNSUPPORTED BY HARDWARE
Jobname  Procstep Stepname  CPU Time     EXCPs     RC
HIS      STARTING HIS       00:00:00         5     00
IEF404I HIS - ENDED - TIME=21.09.36
```

Q: Is the HIS task something that we run on an on-demand basis, or all the time on every system?
A: Experience has shown that there is very little overhead and a very small amount of space is required for the SMF 113s records, so our recommendations is to enable CPU MF Counters on all of your partitions and continuously run it and collect SMF 113s.

Q: Is the OMVS directory required for Counters mode?
A: With APAR OA43366 on z/OS 2.1 and /OS 2.2, the OMVS directory file is not required (e.g. CNTFILE=NO). Without APAR OA43366, even though CPU MF Counters writes SMF 113s records, the OMVS directly is still required, as summary information is written to it once Counters collection is ended.

z/OS Operational Considerations

As discuss above we recommend that HIS and CPU MF be continuous running and collecting SMF 113s. Many customers use “auto operations” to help ensure HIS is started

© 2019 IBM Corporation

and enabled correctly. Consider including the he S HIS command in COMMNDxx, and in addition use the following z/OS system messages to trigger commands or actions in auto operations processes.

Condition:	Once HIS is started, enable Counters and begin collecting SMF 113s
System Message:	**HIS002I HIS INITIALIZATION COMPLETE**
Issue Command:	**"F HIS,B,TT='runid',PATH='/HIS/',CTRONLY,CTR=(B,E),SI=SYNC"**

Condition:	HIS stopped recording data but HIS stayed active
System Message:	**HIS025I** DATA COLLECTION IS PREMATURELY ENDING.
Action	**Recycle HIS**

Additional User Experience – Dec 2014

Below is feedback from an internal IBM installation that implemented CPU MF Counters. The purpose is to share their experience and information that may be useful to others.

For additional reference see the z/OS MVS System Commands book section on "Starting, configuring and stopping hardware event data collection" – the URL below is for the z/OS 2.1 bookshelf:
http://www.ibm.com/systems/z/os/zos/library/bkserv/v2r1pdf/index.html#IEA

There are 2 steps needed to activate HIS:

- Section 1: HMC/SE changes for the LPAR
- Section 2: Setup the HIS started task and file system

© 2019 IBM Corporation

Section 1 – Change LPAR Security options for the LPAR on the HMC/SE

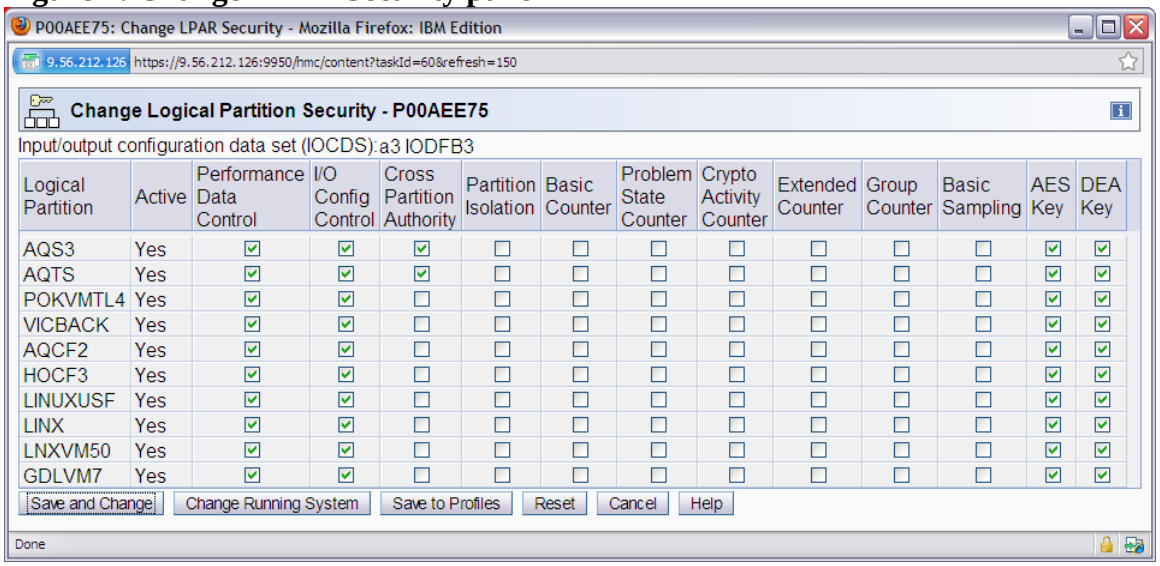
To enable hardware instrumentation on the LPAR, we need to enable the counters. To do this we need to logon to the HMC and get into Single Object Operations (SOO) and change the LPAR Security Controls.

Logon to the HMC:

- Double click on **Defined CPCs**
- Highlight the CEC
- Right Click on the CEC, select **Recovery** -> **Single Object Operations**
- Double click on CPC
- Right click on the CPC (usually P00serial number), select **CPC Operational Customization** -> **Change LPAR Security**

The screen should look like this:

Figure 1: Change LPAR Security panel



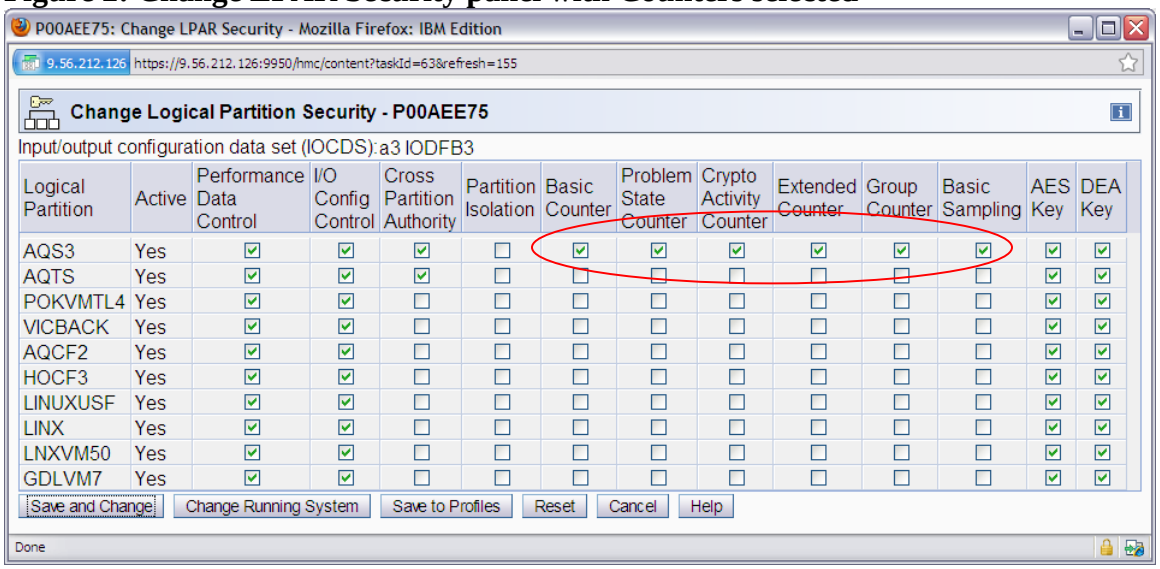
Depending on what counters you want, select the appropriate ones. We will setup HIS to only collect the Basic and Extended counters, but in case we need the other counters later, we will select all of them:

- Basic Counter
- Problem State Counter
- Crypto Activity Counter
- Extended Counter
- Group Counter
- Basic Sampling

Then click on button <Change Running System>:

© 2019 IBM Corporation

Figure 2: Change LPAR Security panel with Counters selected



This is will activate the changes to the LPAR immediately and will NOT save the values to the activation profiles, so the updates will be lost the next time the LPAR is activated via the activation image profile. If you now want to save to the activation profiles, select <Save to Profiles>. If you want the changes to made immediately and save to the activation profile, select <Save and Change>.

Section 2 – Setup HIS to collect SMF 113 records all the time

If you want to setup HIS to collect SMF 113 records all the time you can follow the procedure below.

A few configuration points we want to mention:

- We decided to use the CMDFILE1 DD statement and point to a dataset that will contain the HIS options. We felt it was easier than having automations code the start options (and this way if we wanted to change it we wouldn't have to involve the automations team). We chose to use SYS1.PARMLIB(HISPRMxx) to store the HIS options. There is no formal HISPRMxx support for HIS, we just chose SYS1.PARMLIB since we store the other MVS configuration parameters there as well. You can use any dataset you wish, just ensure that the HIS started task id has READ access to the dataset.
- You will need to have automations add the HIS start and stop commands to the system IPL startup and shutdown.
- We also added an entry to cron to delete files in the HIS file system, so it doesn't fill up.

© 2019 IBM Corporation

1. Copy the HIS proc from SYS1.PROCLIB to your PROCLIB

2. Modify the HIS proc by changing the CMDFILE1 DD statement to point to the dataset name that will have the HIS options:

   ```
   //CMDFILE1 DD DSN=SYS1.PARMLIB(HISPRMXX),DISP=SHR,FREE=CLOSE
   ```

   Note: We added FREE=CLOSE to the dataset name in the CMDFILE1 DD, so HIS does not retain allocation of the dataset while the task is up (i.e. in our case, HIS had SYS1.PARMLIB allocated with DISP=SHR while the task was up). By coding FREE=CLOSE, HIS will release the allocation of the dataset once hardware event data collection begins, in other words once F HIS,B… is issued. The one drawback is if you need to restart HIS collection, you will need to issue P HIS and then S HIS (vs. F HIS,E, then restarting hardware event data collection via F HIS,B …). The reason is because once we start HIS counters via F HIS,B … the system will free up the dataset from HIS allocation, so HIS will no longer 'point' to the dataset in the CMDFILE1 DD statement.

3. Update SYS1.PARMLIB(HISPRMxx):

   TT='*sysname*',CTRONLY,CTRSET=(B,E),SI=SYNC

   TT or TITLE is an optional keyword. We chose to set TT to the system name, i.e. TT='AQHO'

   CTRSET=(B,E) only gathers basic and extended counters (see section 3 for more information on counter types).

   SI=SYNC sets the SMF interval to "Synch" with the SMF global recording interval of the system, so the SMF 113 records are written at the same time and interval as the other SMF records.

   See the z/OS System Commands book for additional options.

4. Define a started task id for HIS with an OMVS segment: (HOME=/var/his), non-0 UID and a file system at least (100,50) cylinders.

   We had 2 considerations for the zFS:

   (a) We wanted the HIS file system to be unique on each system in the sysplex. When HIS writes a file to the USS file system, the filenames didn't seem to have the LPAR name embedded in the filename, so it wasn't easy to see which files were associated with which LPAR. Granted for counters, we really wouldn't look at the files, but in case we needed to turn on sampling in the future we would the data isolated for each LPAR.

© 2019 IBM Corporation

We chose to use /var/his. Every system in the sysplex has its own /var file system, so /var/his would be unique.

(b) The size of the zFS. The size of the zFS varies with the type of counters you will collect and the frequency of how often you stop and start collection and what System z machine you are running on. Every time HIS is stopped, it writes a small file to the file system. In our case, for just basic and extended counters, the file was 12078 bytes for a z/OS 2.1 LPAR on a zEC12 and 8802 bytes for z196. The file size depends on how many CPs the LPAR has and what System z processor the LPAR is running on. The System zEC12 has more counters than the System z196. Both of these LPARs have 5 CPs and 2 zIIP CPs.

Table 1 - Size of counter file for basic and extended counters on zEC12 vs. z196

	Size of counter file on zEC12	Size of counter file on z196
z/OS 2.1 LPAR with 5 GP CPs + 2 zIIP CPs	12078	8802

Mount file system (OMVS.ZFS.&SYSNAME.VAR.HIS) with AUTOMOVE=UNMOUNT to /var/his. Ensure the permission bits for /var/his are 755 with OWNER=HIS.

If necessary, define a STARTED profile for HIS:
RDEF STARTED HIS.* STDATA(USER(HIS) GROUP(TASKS))
SETR REFRESH RACLIST(STARTED)

5. Update BPXPRMxx

6. Verify/update SMFPRMxx is collecting SMF record type 113

   Issue D SMF,O to verify 113 records are being collected. Depending on how your SMFPRMxx member is coded, you may not see type 113 records listed, but it doesn't mean that you are not collecting them. For example, you may collect all records, but can code NOTYPE to exclude certain records.

7. Start HIS:
   S HIS

   You should see message:
   HIS002I HIS INITIALIZATION COMPLETE

8. Issue command to start collecting counter data:

   F HIS,BEGIN,DDNAME=CMDFILE1

© 2019 IBM Corporation

You should see messages:
HIS033I HIS SERVICE PARAMETER(S) ACCEPTED
HIS011I HIS DATA COLLECTION STARTED

If you receive the following error message, then you need to verify the settings are correct on the HMC/SE as noted in Section 1:

F HIS,BEGIN,DDNAME=CMDFILE1
HIS026I MODIFY HIS COMMAND CANNOT BE PROCESSED.
BASIC COUNTER SET IS UNAUTHORIZED

9. Display HIS:
D HIS

Verify the items in blue, i.e. title and counter set = Basic,Extended

D HIS
HIS015I 10.05.15 DISPLAY HIS 309
HIS 00C6 ACTIVE
COMMAND: MODIFY HIS,BEGIN,DDNAME=CMDFILE1
START TIME: 2014/08/18 10:05:11
END TIME: ----/--/-- --:--:--
COMPLETION STATUS: --------
FILE PREFIX: SYSHIS20140818.100511.
COUNTER VERSION NUMBER 1: 1 COUNTER VERSION NUMBER 2: 2
COMMAND PARAMETER VALUES USED:
TITLE=AQHO
PATH= .
DDNAME= CMDFILE1
COUNTER SET= BASIC,EXTENDED
DURATION= NOLIMIT
CTRONLY
DATALOSS= IGNORE
STATECHANGE= SAVE
SMFINTVAL= SYNC
HISSERV STATUS: ACTIVE
EVENT
AUTHORIZED= BASIC,PROBLEM-STATE,CRYPTO-ACTIVITY,EXTENDED,ZOS
ENABLED= BASIC,EXTENDED
SAMPLE
AUTHORIZED= BASIC
ENABLED= NONE
BUFCNT= 4 (PAGES/PROCESSOR)
SAMPFREQ= 800000 (SAMPLES/MINUTE)

© 2019 IBM Corporation

PROFILER
NAME START QUERY SAMPLE S F
HISPROF 2014/08/18 10:05:11 00:00:00.000006 ------

10. Add an entry to cron to delete files in /var/his after a certain number of days – we chose 14 days. Our cron entry runs Monday – Friday at 00:40:
40 00 * * 1-5 /usr/bin/skulker -rl /var/logs/skulker.log /var/his 14

11. Add HIS to automations
    HIS has a dependency on OMVS, so it should be a child of OMVS in automations.

    Start HIS: S HIS
    After msg HIS002I HIS INITIALIZATION COMPLETE is issued, start HIS collection:
    F HIS,BEGIN,DDNAME=CMDFILE1

    Stop HIS:
    P HIS

12. After you have setup HIS and HIS is up and running, verify you are collecting SMF 113 records by running PGM=IFASMFDP (SMF dump).

© 2019 IBM Corporation