

The Flash Express Feature on IBM zEnterprise EC12 and z/OS exploitation of flash storage

The IBM zEnterprise EC12 (zEC12) server introduces the Flash Express feature (FC 0402) that provides a new PCIe I/O card that resides in the PCIe I/O drawer that implements internal Flash Solid State Disks (SSD). Each Flash Express card has a maximum capacity of 1.6 TB and is installed by pair to ensure availability and redundancy. A maximum of four pairs of cards can be installed on a zEC12 providing a maximum capacity of 6.4 TB of storage. In addition, data that resides on Flash SSD is encrypted through the use of smart cards and an integrated smart card reader on the primary and alternate Support Elements (SE). For complete and detailed instructions, see [IBM zEnterprise EC12 Technical Guide, SG24- 8049](#).

With the z/OS V1R13 RSM Enablement Offering web deliverable (FMID JBB778H) for z/OS V1R13, z/OS will exploit Flash through a new tier of memory called Storage Class Memory (SCM) for paging and SVC dump processing. This function is expected to provide faster paging and dump processing because flash storage is faster compared to hard disk storage. In addition to support for the existing large (1 MB) pages and frames, zEC12 supports *pageable* large pages when SCM is configured and allocated to z/OS.

Configuring Flash Storage on zEC12

Our zEC12 was delivered with two pairs of Flash Express cards. The cards were installed and configured online by our local hardware support personnel. This configuration task also formats the flash storage to prepare it for use by customer LPARs. There is no HCD/IOCP definition required for Flash Express installation. However, Flash Express cards are assigned PCHIDs according to their physical location in the I/O drawer even though they have no customer-usable ports.

You can allocate flash storage through a new dialog available on either the HMC or SE. Figure 1 shows the initial view of P91, our zEC12, after its Flash cards were configured

online

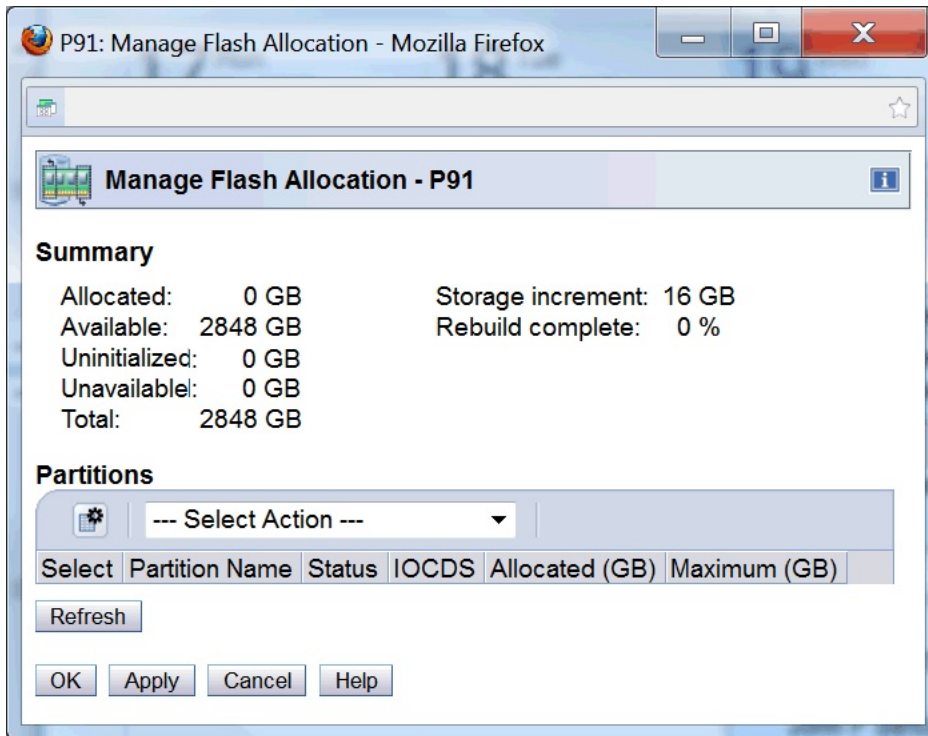


Figure 1: P91 Initial Manage Flash Allocation settings

Note in Figure 1 that with two pairs of cards only 2848GB of flash storage is available for customer allocation. The remaining storage is reserved for internal usage for flash firmware.

A user with a system programmer role can manage the allocations using the select action dropdown menu in the Partitions table in the Manage Flash Allocation panel. For example, selecting Add Allocation allows the user to choose an existing LPAR to allocate initial and maximum amounts in 16 GB increments. Figure 2 shows an example of the New Flash Allocation panel.

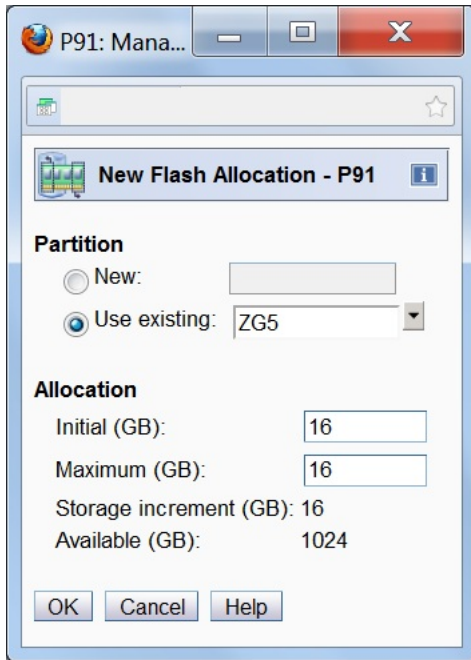


Figure 2: New Flash Allocation panel

Note that when allocating flash storage to LPARs using the New Flash Allocation panel, you must be careful to select the correct LPAR. The order of Existing Partitions in the New Flash Allocation might list non-flash-supporting images first (such as CFs), which can lead to the accidental allocation of flash storage to those LPARs.

Our complete LPAR allocation of Flash is shown in Figure 3:

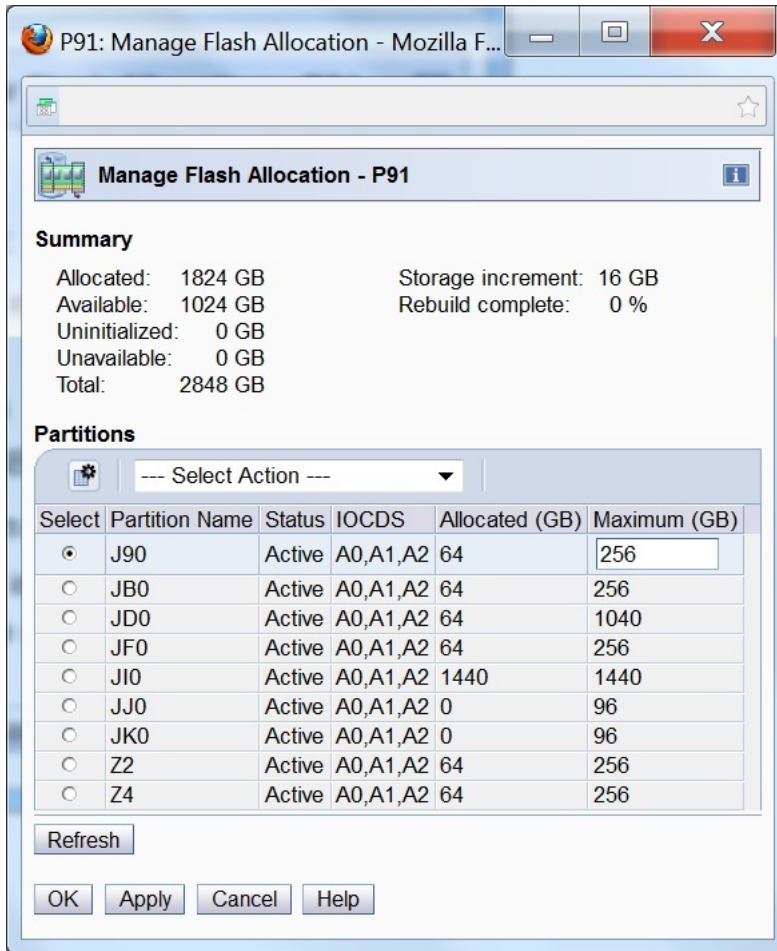


Figure 3: P91 Flash Allocations

Configuring storage class memory on z/OS on zEC12

Flash can be allocated before the LPARs are activated and detected by z/OS during IPL, or configured on dynamically after IPL. You can specify a new PAGESCM keyword in IEASYSxx with the z/OS V1R13 RSM Enablement installed to control the use of SCM. PAGESCM defaults to ALL to allow allocation of all increments specified in the Manage Flash Allocation panel while specifying PAGESCM=NONE prevents the use of any SCM. The system issues new message IAR031I during IPL to indicate the amount of SCM that was allocated to the LPAR based on the PAGESCM value. For example, on a system with nothing specified for PAGESCM and 64G of SCM allocated to the LPAR, message IAR031I reports

```
IAR031I USE OF STORAGE-CLASS MEMORY FOR PAGING IS ENABLED - PAGESCM=ALL
, ONLINE=0006536M
```

Two existing operator commands have been updated to show information related to SCM for a given system.

1. The DISPLAY M command has been updated to show the status of a system's SCM, including displaying details for each online increment:

```
-RO J90,D M=SCM
IEE174I 15.50.16 DISPLAY M
STORAGE-CLASS MEMORY STATUS
256G DEFINED
ONLINE
  0G-64G
192G OFFLINE-AVAILABLE
12% IN USE
SCM INCREMENT SIZE IS 16G
```

```
-RO J90,D M=SCM(DETAIL)
IEE174I 15.52.34 DISPLAY M
STORAGE-CLASS MEMORY STATUS - INCREMENT DETAIL
256G DEFINED
ADDRESS  IN USE  STATUS
      0G      0%  ONLINE
     16G      0%  ONLINE
     32G     27%  ONLINE
     48G     20%  ONLINE
ONLINE: 64G  OFFLINE-AVAILABLE: 192G  PENDING OFFLINE: 0G
12% IN USE
SCM INCREMENT SIZE IS 16G
```

2. The Auxiliary Storage DISPLAY ASM command has been updated to identify the SCM blocks that the system is currently using, in addition those in use on page data sets.

```
-RO J90,D ASM
IEE200I 15.55.54 DISPLAY ASM
TYPE      FULL  STAT  DEV  DATASET NAME
PLPA      18%   OK   8571  SYS1.J90.PLPA
COMMON    0%   OK   8570  SYS1.J90.COMMON
LOCAL     5%   OK   8570  SYS1.J90.LOCAL
LOCAL     5%   OK   8570  SYS1.J90.LOCAL2
LOCAL     5%   OK   8571  SYS1.J90.LOCAL3
LOCAL     5%   OK   8571  SYS1.J90.LOCAL4
LOCAL     5%   OK   8572  SYS1.J90.LOCAL5
SCM       8%   OK   N/A   N/A
PAGEDEL  COMMAND IS NOT ACTIVE
```

```
-RO J90,D ASM,SCM
IEE207I 15.57.56 DISPLAY ASM
STATUS    FULL      SIZE          USED          IN-ERROR
IN-USE    8%          16,777,216    1,458,024      0
```

Allocating new flash storage increments to an LPAR that is already active with the z/OS V1R13 RSM Enablement Offering installed will result in the system detecting and reporting the new SCM values through new message IAR034I:

```
IAR034I ADDITIONAL STORAGE-CLASS MEMORY DETECTED
```

```
-RO J90,D M=SCM
IEE174I 15.58.55 DISPLAY M
STORAGE-CLASS MEMORY STATUS
272G DEFINED
ONLINE
  0G-64G
208G OFFLINE-AVAILABLE
12% IN USE
SCM INCREMENT SIZE IS 16G
```

You might notice that the IN USE % reported by D M=SCM is higher than the % FULL value reported by D ASM. This is because ASM creates a pool of available SCM blockids for quick allocation that is considered in-use by the D M command processor.

You can configure offline SCM increments online to an active system using the CONFIG command with the new SCM option. The reconfiguration value that you specify must be a multiple of the SCM increment size.

```
RO J90,CF SCM(16G),ONLINE
IEE195I SCM LOCATIONS 64G TO 80G ONLINE
IEE712I CONFIG PROCESSING COMPLETE
```

```
-RO J90,D M=SCM(DETAIL)
IEE174I 15.59.45 DISPLAY M
STORAGE-CLASS MEMORY STATUS - INCREMENT DETAIL
272G DEFINED
ADDRESS  IN USE  STATUS
   0G      0%  ONLINE
  16G      0%  ONLINE
  32G     27%  ONLINE
  48G     20%  ONLINE
  64G      0%  ONLINE
ONLINE: 80G  OFFLINE-AVAILABLE: 192G  PENDING OFFLINE: 0G
9% IN USE
SCM INCREMENT SIZE IS 16G
```

IBM is expected to support dynamic reconfiguration of SCM increments at a later date. Until then, to remove SCM increments from an LPAR will require deactivating the image and using the Remove Allocation option from the Manage Flash Allocation panel on the SE or HMC. The firmware processing to deallocate increments can take several minutes depending on how much flash storage is being deallocated, and you might see "Uninitialized" or "Unavailable" amounts during that period; as result the system might be temporarily unable to allocate those amounts to any other LPAR. However, this

condition is transient. Figure 4 highlights the unavailable fields while this transition was in progress.

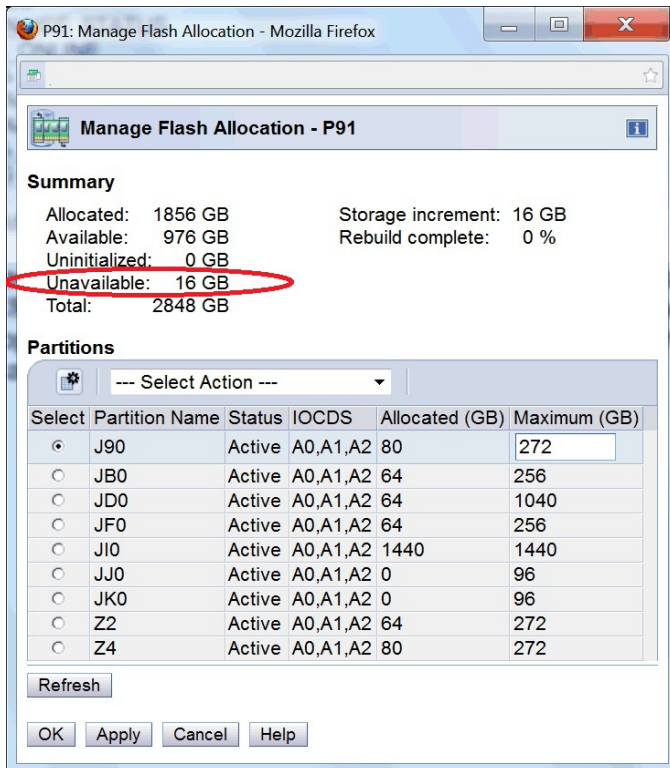


Figure 4: P91 Flash Allocations

RMF support for Storage Class Memory

APAR OA38660 introduces RMF support to report on SCM statistics in the following reports:

- Monitor I Paging Activity
- Monitor I Page Data Set Activity
- Monitor II Page Data Set Activity (PGSP)

SMF records 71 and 75 are changed with APAR OA38660 to provide this new support. In addition, new overview conditions are added based on SMF record 71.

The following images show snippets of these updated reports to highlight the new SCM statistics. Let's start with the real-time Monitor II Page Data Set Activity report.

RMF - PGSP Page Data Set Activity

Line 1 of 8

CPU= 67/ 12 UIC= 65K PR= 0

System= J90 Total

S	VOLUME	DEV	DEV	%SLOTS	PAGE	I/O REQ	AVG PAGES	17:33:22	
T	SERIAL	NUM	TYPE	IN USE	TRAN TIME	RATE	PER I/O	V	DATA SET NAME
P	PAG901	8571	33909	18.16	0.000	0.000	0.000		SYS1.J90.PLPA
C	PAG900	8570	33909	0.41	0.000	0.000	0.000		SYS1.J90.COMMON
L	PAG900	8570	33909	5.70	0.000	0.000	0.000	Y	SYS1.J90.LOCAL
L	PAG900	8570	33909	5.47	0.000	0.000	0.000	Y	SYS1.J90.LOCAL2
L	PAG901	8571	33909	5.08	0.000	0.000	0.000	Y	SYS1.J90.LOCAL3
L	PAG901	8571	33909	5.11	0.000	0.000	0.000	Y	SYS1.J90.LOCAL4
L	PAG902	8572	33909	5.31	0.000	0.000	0.000	Y	SYS1.J90.LOCAL5
S	N/A	N/A	N/A	6.96	0.000	408.000	1.000		N/A

Figure 5: Monitor II Page Data Set Activity

Note the new line of type S in Figure 05 that represents SCM. All DASD-related fields are N/A for this line. Note also that because of its superior performance, the majority of paging activity is going to SCM.

The Monitor I Page Data Set Activity report provides similar new statistics for SCM.

PAGE DATA SET ACTIVITY															
z/OS V1R13				SYSTEM ID J90				DATE 10/30/2012				INTERVAL 15.00.009			
				RPT VERSION V1R13 RMF				TIME 17.30.00				CYCLE 1.000 SECONDS			
NUMBER OF SAMPLES = 900				PAGE DATA SET AND SCM USAGE											
PAGE	SPACE	VOLUME	DEV	DEVICE	SLOTS	----	SLOTS USED	---	BAD	%	PAGE	NUMBER	PAGES	V	
TYPE	SERIAL	NUM	TYPE	ALLOC	MIN	MAX	AVG	SLOTS	USE	TRANS	IO REQ	XFER	D	I	
														DATA SET NAME	
PLPA	PAG901	8571	33909	71999	13078	13078	13078	0	0.00	0.000	0	0	0	SYS1.J90.PLPA	
COMMON	PAG900	8570	33909	71999	297	297	297	0	0.00	0.000	0	0	0	SYS1.J90.COMMON	
LOCAL	PAG900	8570	33909	575999	32815	32815	32815	0	0.00	0.000	0	0	0	Y	SYS1.J90.LOCAL
LOCAL	PAG900	8570	33909	575999	31509	31509	31509	0	0.00	0.000	0	0	0	Y	SYS1.J90.LOCAL2
LOCAL	PAG901	8571	33909	575999	29248	29248	29248	0	0.00	0.000	0	0	0	Y	SYS1.J90.LOCAL3
LOCAL	PAG901	8571	33909	575999	29442	29442	29442	0	0.00	0.000	0	0	0	Y	SYS1.J90.LOCAL4
LOCAL	PAG902	8572	33909	575999	30610	30610	30610	0	0.00	0.000	1	1	1	Y	SYS1.J90.LOCAL5
SCM	N/A	N/A	N/A	20972K	1459K	1459K	1459K	0	0.89	0.000	51,770	51,789		N/A	

Figure 6: Monitor I Page Data Set Activity

Finally the Paging Activity report now reports statistics on SCM paging blocks, as well as the SHARED, HIGH SHARE, and COMMON frames backed by SCM slots. Figure 7 and Figure 8 show these new fields.

SHARED FRAMES / SLOTS	TOTAL	CENTRAL STORAGE		FIXED TOT	FIXED BEL	AUX DASD	AUX SCM
MIN	89,586		66,721	349	0	0	12,484
MAX	89,638		66,773	350	0	0	12,486
AVG	89,620		66,754	350	0	0	12,484
LOCAL PAGE DATA SET SLOTS	TOTAL	AVAILABLE	BAD	NON-VIO	VIO		
MIN	2,879,995	2,879,995	0	0	0		
MAX	2,879,995	2,879,995	0	0	0		
AVG	2,879,995	2,879,995	0	0	0		
SCM PAGING BLOCKS	TOTAL	AVAILABLE	BAD	IN-USE			
MIN	16777216	16600133	0	173,338			
MAX	16777216	16603878	0	177,083			
AVG	16777216	16603713	0	173,503			

Figure 7 Monitor I Paging Activity 1

HIGH SHARED FRAMES	TOTAL	CENTRAL STORAGE		AUX DASD	AUX SCM
MIN	136902.1M		156,895	0	208
MAX	136902.1M		161,632	0	208
AVG	136902.1M		159,106	0	208
HIGH COMMON FRAMES	TOTAL	CENTRAL STORAGE	FIXED 4K	AUX DASD	AUX SCM
MIN	17301504	78,576	4,612	0	3
MAX	17301504	78,579	5,266	0	6
AVG	17301504	78,577	4,741	0	5

Figure 8: Monitor I Paging Activity 2

Additional information related to Flash Storage and Storage Class Memory can be found in latest versions of the following Redbooks:

- [IBM zEnterprise System Connectivity Handbook, SG24-5444](#)
- [IBM zEnterprise EC12 Technical Introduction, SG24-8050](#)
- [IBM zEnterprise EC12 Technical Guide, SG24 8049](#)

You should also consult the PSP bucket for information on the z/OS V1R13 RSM Enablement Offering web deliverable (FMID JBB778H) for z/OS V1R13.