z/OS 1.9: Large Page Support

Kathy Walsh
IBM Advanced Technical Support

Trademarks

The following are trademarks of the International Business Machines Corporation in the United States and/or other countries.

DB2*
DB2 Connect
DB2 Universal Database
e-business logo
GDPS*
Geographically Dispersed Parallel Sysplex
HyperSwap
IBM*
IBM eServer
IBM logo*
Parallel Sysplex*
System z
Tivoli*
VM/ESA*
WebSphere*
z/OS*
z/VM*
zSeries*

* Registered trademarks of IBM Corporation

The following are trademarks or registered trademarks of other companies.

Intel is a registered trademark of the Intel Corporation in the United States, other countries or both.

Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.

Java and all Java-related trademarks and logos are trademarks of Sun Microsystems, Inc., in the United States and other countries.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Microsoft, Windows and Windows NT are registered trademarks of Microsoft Corporation.

SET and Secure Electronic Transaction are trademarks owned by SET Secure Electronic Transaction LLC.

* All other products may be trademarks or registered trademarks of their respective companies.

Notes:

Performance is in Internal Throughput Rate (ITR) ratio based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput improvements equivalent to the performance ratios stated here.

IBM hardware products are manufactured from new parts, or new and serviceable used parts. Regardless, our warranty terms apply.

All customer examples cited or described in this presentation are presented as illustrations of the manner in which some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.

This publication was produced in the United States. IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change without notice. Consult your local IBM business contact for information on the product or services available in your area.

All statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

Information about non-IBM products is obtained from the manufacturers of those products or their published announcements. IBM has not tested those products and cannot confirm the performance, compatibility, or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

Prices subject to change without notice. Contact your IBM representative or Business Partner for the most current pricing in your geography.

Agenda

- **Large Page Performance Considerations**
- **Overview of z/OS Large Page Implementation**
  - ***IEASYSxx Changes***
  - ***IARV64 Changes***
  - ***RMF Changes***

This document was created with material provided by Elpida Tzortzatos, IBM Corp, z/OS Core Technical Design

Large Page Performance Considerations

- **Large Page support is a special purpose performance improvement feature**
  - *Not recommended for general use*
  - *Large page usage provides performance value to a select set of applications, primarily long running, memory access intensive applications*
- **Some applications can be severely degraded by the use of large pages**
  - *Short lived processes with small working sets are usually not good candidates for large pages*
- **The system will not "rush" to build a 1MB page**
  - *When the large pages are requested it may take awhile for the request to be satisfied*

Enable Large Pages via IEASYSxx

- **New keyword in IEASYSxx called LFArea specifies the amount of real storage to be used for Large Pages**
  - ***LFArea=(xx%| xxxxxxM | xxxxxxG)***
    - **xx% - Percentage of online storage at IPL to be used for Large Pages**
    - **xxxxxxM or xxxxxxG – Amount of online storage at IPL to be used for Large Pages**

z/OS Large Page Support Overview

- **The z/OS Large Page size is 1 Megabyte**
- ***Both 4K and 1MB page sizes are supported***
- **If the system is constrained for 4K pages, free 1MB pages will be used to back 4K page requests**
- **Pages can later be coalesced into 1MB pages**
- **1MB Large Pages will NOT be reconfigurable**

z/OS Large Page Support Overview

- **Large Pages are backed by 256 contiguous 4K real storage frames**
- **Large Page support is available only for 64-Bit Virtual Private storage**
- **The request to use 1MB pages is expressed as an attribute of the memory object**
  - ***Memory Object Pages are backed at allocation time***
    - **The entire "usable" virtual address range of a memory object backed by large pages will be backed when the virtual memory object is obtained**
      - Guard Areas will not be backed
  - ***Memory Object is implicitly fixed at allocation time***

z/OS Large Page Paging Support

- **Large Pages will be NOT be pageable**
  - ***Large Pages will be treated as fixed pages and will not be paged out to AUX***
  - ***Primary reason is because certain control structures used for paging to Aux are not created for large pages***
  - ***Other reasons include:***
    - **The great delay in paging in a 1MB page**
    - **The expense of constructing 1MB of contiguous real storage from 256 4K frames when required**

z/OS Large Page Support Overview

- **The following IARV64 requests will operate on Large Pages**

GETSTOR	**Creates a 64-Bit private memory object that can be backed by Large Pages**
DETACH	**Frees a memory object that can be backed by Large Pages**
CHANGEGUARD	**Changes the amount of guard area within the specified memory object**
LIST	***Provides information about memory objects***

- **IARV64 has a new keyword called PAGEFRAMESIZE=**
  - ***4K (default)***
  - ***1M (Requests 1MB pages)***
  - ***MAX (Requests the largest page size supported)***
- **All APIs enhanced for Large Page support will require the requestor to be authorized**

Large Page Support Overview…

- **The following example shows how a memory object that is to be backed by Large Pages is created:**

```
IARV64 REQUEST=GETSTOR                                  +
                SEGMENTS=THREE_SEG,                     +
                ORIGIN=VIRTT64_ADDR,                    +
                PAGEFRAMMESIZE=1MEG,                    +
                GUARDLOC=LOW,                           +
                GUARDSIZE=ONE_SEG,                      +
                SVCDUMPRGN=NO,                          +
                COND=YES
```

IARV64 GETSTOR Request Operation

2097 Processor	z/OS Release	Issue IARV64 GETSTOR with Pageframesize=
Yes	z/OS 1.9 with OA20902 applied	Request will be accepted
Yes	Pre z/OS 1.9 or z/OS 1.9 without OA20902	Request will be rejected
No	z/OS 1.9 with OA20902 applied	Request will be rejected when pageframesize=1M Request will be backed with 4K pages when pageframesize=MAX
No	Pre z/OS 1.9 or z/OS 1.9 without OA20902	Request will be rejected

RMF Enhancements for z10 Processor – OA12774

- **Postprocessor Paging Activity Report**
  - ***Frame and Slot count section extended with a new section on memory object usage***
    - **Minimum, maximum and average numbers of 1 MB frames backed in real storage**
    - **Number of large memory objects allocated in the system**
- **RMF Monitor III Central Storage Summary**
  - ***STORR (Storage Resource Delays) and STORS (Storage Delay Summary) reports displays the total number of large memory objects allocated in the system and the number of 1 MB frames backed in real storage***
  - ***STORF (Storage Frames) report displays the number of 1 MB frames per address space backed in real storage***