The AIX Virtual Memory Manager (VMM)

Special Notices

Information in this document is correct to the best of our knowledge at the time of this writing. Please send feedback by fax to "AIXServ Information" at (512) 823-4009.

Please use this information with care. IBM will not be responsible for damages of any kind resulting from its use. The use of this information is the sole responsibility of the customer and depends on the customer's ability to evaluate and integrate this information into the customer's operational environment.

About This Document

This document addresses how RAM and paging space are used. The following information is applicable to AIX versions 3.2 and 4.1.

VMM Overview

The Virtual Memory Manager (VMM) services memory requests from the system and its applications. Virtual-memory segments are partitioned in units called pages; each page is either located in physical memory (RAM) or stored on disk until it is needed. AIX uses virtual memory in order to address more memory than is physically available in the system. The management of memory pages in RAM or on disk is handled by the VMM.

Real-Memory Management

In AIX, virtual-memory segments are partitioned into 4096-byte units called pages. Real memory is divided into 4096-byte page frames. The VMM has two major functions: 1) manage the allocation of page frames, and 2) resolve references to virtual-memory pages that are not currently in RAM (stored in paging space) or do not yet exist.

In order to accomplish its task, the VMM maintains a free list of available page frames. The VMM also uses a page-replacement algorithm to determine which virtual-memory pages currently in RAM will have their page frames reassigned to the free list. The page-replacement algorithm takes into account the existence of persistent vs. working segments, repaging, and VMM thresholds.

Free List

The VMM maintains a list of free page frames that it uses to satisfy page faults. The free list is made up of unallocated page frames. AIX tries to use all of RAM all of the time, except for a small amount which it maintains on the free list. To maintain this small amount of unallocated pages the VMM will use Page Outs and Page Steals to free up space and reassign those page frames to the free list. The virtual-memory pages whose page frames are to be reassigned are selected via the VMM's page-replacement algorithm.

Persistent vs. Working Memory Segments

AIX distinguishes between different types of memory segments, and to understand the Virtual Memory Manager, it is important to understand the difference between working and persistent segments. A persistent segment has a permanent storage location on disk. Files containing data or executable programs are mapped to persistent segments. When a Journaled File System (JFS) file is opened and accessed the file data is copied into RAM. VMM parameters control when physical memory frames allocated to persistent pages should be overwritten and used to store other data.

Working segments are transitory and exist only during their use by a process, and have no permanent disk storage location. Process stack and data regions are mapped to working segments and shared library text segments. Pages of working segments must also have disk storage locations to occupy when they cannot be kept in real memory. The disk paging space is used for this purpose. When a program exits, all of its working pages are placed back on the free list immediately.

Paging Space and Virtual Memory

Working pages in RAM that can be modified and paged out are assigned a corresponding slot in paging space. The allocated paging space will only be used if the page needs to be paged out. However, an allocated page in paging space cannot be used by another page. It remains reserved for a particular page for as long as it exists in virtual-memory. Since persistent pages are paged out to their location on disk from which they came, paging space does not need to be allocated for persistent pages residing in RAM.

The VMM has two modes for allocating paging space, late paging space allocation and early paging space allocation. The VMM's system default after the machine is booted is the late paging space allocation algorithm. Late allocation does not reserve paging space when a memory request for a working page is made. Instead, it approves the requests and only assigns paging space when the page is referenced and not until then. This method is more efficient because some programs allocate large amounts of virtual memory and then use only a small fraction of it.

VMM Memory Load Control Facility

When a process references a virtual-memory page that is on disk, because it either has been paged out or has never been read, the referenced page must be paged in, and this may cause one or more pages to be paged out if the number of available free page frames is low. The VMM attempts to steal page frames that have not been recently referenced, and thus unlikely to be referenced in the near future, via the page-replacement algorithm.

A successful page-replacement keeps the memory pages of all currently active processes in RAM, while the memory pages of inactive processes are paged out. However, when RAM is over- committed, it becomes difficult to choose pages for page out because they will be referenced in the near future by currently running processes. The result is that pages that will soon be referenced still get paged out and then paged in again later. When this happens, continuous paging in and paging out may occur if RAM is over-committed. This condition is called thrashing. The system spends most of its time paging in and paging out instead of executing useful instructions, and none of the active processes make any significant progress. The VMM has a memory load control algorithm that detects when the system is thrashing and then attempts to correct the condition.

VMSTAT's "avm" Field

"avm" stands for "Active Virtual Memory" and not "Available Memory". "avm" is the number of 4K pages that are in use in paging space. The same idea is reflected in the PERCENT USED column of the 'lsps -s' command.

VMSTAT's "fre" Field

"fre" is the number of 4K pages that are currently on the free list. When an application terminates, all of its working pages are immediately returned to the free list. Its persistent pages, however, remain in RAM and are not added back to the free list until they are "stolen" by the VMM for other programs. Persistent pages are also freed if the corresponding file is deleted.

For this reason, the "fre" value may not indicate all the real memory that can be readily available for use by processes. If a page frame is needed, then persistent pages related to terminated applications are among the first to be handed over to another program.

The minimum number of pages that the Virtual Memory Manager keeps on the free list is determined by the "minfree" parameter of vmtune. If the number of pages on the free list drops below "minfree", the Virtual Memory Manager will steal pages until the free list has been restored to the maxfree value.

How The System is Using Memory

The 'svmon' command can be used to determine roughly how much memory the system is using.

Note: In AIX 4.1, PAIDE/6000 must be installed in order use svmon. Check to see if this is installed, by typing:

```
$ lslpp -l perfagent.tools.
```

To order, PAIDE/6000, call IBM DIRECT at 1-800-426-2255 or contact your local IBM representative.

As root, type 'svmon'. Under the pg space heading, the inuse field is the number of working pages that are in use in all of virtual memory.

Explanation of 'svmon' Output

memory:

size total size of memory in 4K pages

inuse number of pages in RAM that are in use by a process plus the number of persistent pages that belonged to a terminated process and are still resident in RAM. This value is the total size of memory minus the number of pages on the free list.

free number of pages on free list.

pin number of pages pinned in RAM (a pinned page is a page that is always resident in RAM and cannot be paged out)

in use:

work number of working pages in RAM

pers number of persistent pages in RAM

clnt number of client pages in RAM (client page is a remote file page)

pin:

work number of working pages pinned in RAM

pers number of persistent pages pinned in RAM

clnt number of client pages pinned in RAM

pgspace:

size total size of paging space in 4K pages

inuse total number of 'allocated' slots. (See explanation above on allocation of paging space).

To find out how much memory a process is using, type

```
$ svmon -P PID   (for one process)
or
$ svmon -Pau | more (for all processes)
```

To see the number of working pages unique to this process' private stack and data use in all of virtual memory look at the "work" type and description "private". The svmon output may also list several shared segments. For a complete picture, determine which segments are unique to an individual process and which are shared with other programs. Multiply the values by 4096 to get the number of bytes in memory the process is using. The number 4096 comes from the fact that each page is 4KB in size. You can also divide the number of pages by 256 in order to get megabytes.

03/27/96, 4FAX# 2449

Reader's Comments

Please fax this form to (512) 823-4009, attention "AIXServ Information". You may also e-mail comments to: elizabet@austin.ibm.com. These comments should include the same customer information requested below.

Use this form to tell us what you think about this document. If you have found errors in it, or if you want to express your opinion about it (such as organization, subject matter, appearance) or make suggestions for improvement, this is the form to use.

If you need technical assistance, contact your local branch office, point of sale, or 1-800-CALL-AIX (for information about support offerings). These services may be billable. Faxes on a variety of subjects may be ordered free of charge from 1-800-IBM-4FAX. Outside the U.S. call 415-855-4329 using a fax machine phone.

When you send comments to IBM, you grant IBM a nonexclusive right to use or distribute your comments in any way it believes appropriate without incurring any obligation to you.

Note: If you have a problem report or item number, supplying that number may help us determine why a procedure did or did not work in your specific situation.

Problem Report or Item #: Branch Office or Customer #:

Be sure to print your name and fax number below if you would like a reply:

Name: Fax Number:

END OF DOCUMENT (vmm.lxp, 4FAX# 2449)