Organizing Logical Volumes to Increase I/O Performance

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.

Note: The information in this document has **not** been verified for AIX 4.1.

About This Document

This document was created from HONE ITEM Q494228 and is applicable to AIX 3.2.

SOLUTION:

The INTRA-physical volume allocation policy describes the location of the logical volume on the disk. This may be set when the LV is created (with mklv) or changed with the "chlv" command. The values may be CENTER, MIDDLE or EDGE. The disk is laid out as follows: (EDGE 3 MIDDLE 3 CENTER 3 MIDDLE 3 EDGE)

To change the INTRA policy, use the "chlv" command as follows:

```
chlv -a {c m e} LVname
```

Logical volumes with the most active I/O should be placed in the CENTER, while inactive ones should reside on the edges. Note that the logical volume might be split across the different edges when allocated. Since it is assumed that the disk heads generally linger in the CENTER of the disk, it would take equal time for the heads to move to the EDGEs to capture data from the least active LVs. This would only impact performance if the data on the EDGES is read frequently or in large blocks.

The INTER-physical volume allocation policy should also be considered when organizing the LVs on the disk. This policy will determine whether an LV can

be extended across multiple PVs (disks). To increase performance, the value should be MAXIMUM, indicating that the LV should be extended across the maximum number of PVs. This, however, will decrease the availability of the LV should ANY one of the PVs become unavailable. To restrict an LV to a single PV (or the minimum number of PVs which will hold the LV, use MINIMUM allocation).

To change the INTER policy, use the "chlv" command as follows:

```
chlv -e {x m} LVname
```

where "x" = MAXIMUM and "m" = MINIMUM allocation policy.

After the policies have been set for all logical volumes involved, the LVs may be reorganized by using the "reorgvg" command. This command will reorganize all physical partitions to match the allocation policies assigned to each LV. Since the system cannot always fill all of the requirements, it will put the LVs as close as possible to the positions specified. Use the command:

```
reorgvg VGname LVname1 LVname2 ...
```

This is where VGname is the name of the volume group in which the LVs are contained, and LVname1, LVname2, etc are the names of the logical volumes to be reorganized. Note that the order of the LVnames indicates the precedence of the organization of the LVs. The first one in the list will get the highest consideration in its allocation position. If you do not list any LVnames, then ALL LVs will be reorganized. Note that, in this case, the precedence will simply be determined by the order in which the LVs were originally created.

Determining LVM Configuration Based on Performance and

Availability

(From HONE ITEM Q494227)

SOLUTION:

In general, whenever LVM is configured to provide better performance, the availability of the system is impacted.

To configure the system for the highest PERFORMANCE, follow these guidelines when creating Logical Volumes:

1. Create Logical volumes with only 1 copy (do not mirror partitions). If you should select to mirror partitions, then set the following:

   a. Scheduling Policy: PARALLEL

   b. Allocate each logical partition copy on a separate disk: YES

   The parallel-write scheduling policy starts the write operation for all the physical partitions in a logical partition at the same time. When the write

operation to the physical partition that takes longest to complete finishes, the write operation returns.

The YES in the allocation policy sets the STRICT variable in the mklv command run by SMIT. This setting ensures that no two copies of a logical volume partition will reside on the same physical volume if it can be avoided.

2. Write-verify: NO

   Do not perform a follow-up read to each write for verification.

3. INTRA-policy should be set to one of the following:

   a. CENTER: If heavy I/O to the LV

   b. MIDDLE: For moderate I/O

   c. EDGE: If there is little I/O to this LV

4. INTER-policy: MAXIMUM

   This will spread each logical volume across as many physical volumes as possible, thereby allowing the work of reading or writing to a logical volume to be shared among several physical volumes.

   To configure the system for the highest AVAILABILITY, follow these guidelines when creating Logical Volumes:

   a. Use 3 copies (mirror partitions twice).

   b. Write-verify: YES Verify each write by performing a follow-up read.

   c. INTER-policy: MINIMUM This specifies that only as many physical volumes as there are copies should be used to contain the logical volume.

   d. Scheduling Policy: SEQUENTIAL This requires the write of each copy to be performed sequentially, increasing the likeliness that at least one copy completes prior to a system failure.

   e. Allocate each logical partition copy on a separate disk: YES

This will not allow any two copies of the same data to reside on the same disk.

Also consider these factors which may increase system availability:

1. Include at least 3 physical volumes in a volume group. This will allow a quorum (majority) of PVs in the Volume Group to be maintained in the event of one of the PVs becoming unavailable.

2. If mirroring partitions, mirror the copies onto physical volumes that are attached to separate busses, adapters and power supplies. Upon failure of one of these components, the copies may still be maintained, since they are not attached to the failing device.

EQUAL Item: 31XGC

Abstract: What is the best way to eliminate fragmentation on the RS6000 file systems?

R: To determine the physical partition fragmentation on a disk, such as hdisk0, you would issue the command:

```
"lspv -l hdisk0"
```

This will list the logical volumes that are on that disk, the number of physical partitions from that logical volume that are on that disk, and how those physical partitions are distributed on the disk. Under "DISTRIBUTION" you can see how many physical partitions are in the outer edge, outer middle, center, inner middle, and inner edge sections of the disk.

To view the actual order of physical partitions on the disk, you would issue the "lspv -M hdisk0" command. You will be able to see which logical volumes are on the fragmented portions of the disk. The numbers shown in the output are the physical partition number on the disk (physical volume) and the physical partition number on the logical volume.

If you have a logical volume on fragmented portions of a disk, you may reorganize your disk. You would want the logical volumes with the highest I/O to be located nearest the center of the disk.

Although you can check the fragmentation of physical partitions, you cannot check the fragmentation of your actual data within a file system. This fragmentation is caused by continual creation and removal of files. If you notice that performance has degraded over a period of time, then it is quite possible that there is fragmentation within your file system.

The only way to alleviate this type of fragmentation is to backup the file system by filename, delete the file system, then recreate the file system with the "mkfs" command, and then restore the file system data.

08/29/96,

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:

08/29/96,

END OF DOCUMENT (disk.performance.lxp,)