System i™ integration with BladeCenter® and System x™

Aligning storage partitions for VMware® ESX Server on iSCSI attached integrated servers

Table of Contents

1 Introduction

The following content is an addendum to Redbook *VMware VI3 on BladeCenter and System x with System i*. This document will refer to specific sections of this Redbook.

Partition alignment is required for optimal disk I/O performance. When a guest operating system is hosted by VMware ESX Server 3 on iSCSI attached integrated servers, the following will affect the partition alignment:
- Storage space data offset
- VMFS partition offset (starting sector)
- Guest file system partition offset (starting sector). This is a file system partition within a virtual machine.

Various combinations of the above can be used to align storage.

2 Partition alignment practices

When aligning partitions on iSCSI attached integrated servers running VMware ESX Server 3, we recommend:
- 1 VMFS primary partition per storage space. The VMFS partition is the only partition on the storage space.
- 1 Guest file system primary partition per VMware virtual disk

By doing this, the partition alignment can be achieved by modifying the data offset of the storage space or the VMFS partition offset without modifying the guest file system partition offset within your virtual machine.

You may have multiple VMware virtual disks per storage space for guest operating systems that have the same storage space data offset and VMFS partition offset. The table below lists guest operating systems with the same data offset and partitions offsets within the same row. For example, partition alignment would be maintained when you have multiple VMware virtual disks on a storage space for Windows Server 2008 and Windows Vista. However, partition alignment would not be maintained with multiple VMware virtual disks on the same storage space if you had Windows Server 2008 and Linux.

The following table lists the recommended ways to achieve partition alignment on storage spaces.

i5/OS	Guest OS [1]	Storage Space Data Offset [2]	VMFS Partition Offset [3]	Guest File System Partition Offset [3]
V6R1	• Windows Server 2008 • Windows Vista	Align the first logical disk sector	Use default offset (128) when created with the VMware VI Client [5]	Use default offset (multiple of 8)
V6R1	• Legacy Windows [4] • Linux	Align the first logical partition sector	Use default offset (128) when created with the VMware VI Client [5]	Use default offset (63)
V5R4	• Windows Server 2008 • Windows Vista	Not applicable	Manually create with fdisk using offset of 63 [6]	Use default offset (multiple of 8)
V5R4	• Legacy Windows [4] • Linux	Not applicable	Use default offset (128) when created with VI Client [5]	Use default offset (63)

Notes:
1. For a list of all the operating systems supported by VMware ESX Server 3, see the *Guest Operating System Installation Guide* on the VMware documentation website http://www.vmware.com/support/pubs/
2. See the Storage Space Data Offset section below for a description of values.
3. The default partition offset values are in parenthesis.
4. Legacy Windows refers to Windows Server 2003, Windows XP, Windows 2000 and Windows NT 4.0
5. For instructions on creating a VMFS partition with the VMware Virtual Infrastructure Client, see section *7.1.6 Creating a VMFS datastore in a storage space with VirtualClient* of the Redbook.
6. For instructions on manually creating a VMFS partition, see *2.2 Manually create a VMFS partition specifying an offset* section below using the offset from the table above.

2.1 Storage Space Data Offset

In V6R1, data offset is a new configurable parameter that can be specified when the storage space is created. The storage space data offset can not be changed after the storage space is created so make sure you choose the correct offset as listed in the table above. To create a storage space, refer to section *7.1 Creating new storage spaces for Virtual Machines* in the Redbook.

The recommended values in the table and their meanings are below.

Align the first logical disk sector

This data offset will align sector 0 of the storage space. This type of data offset should be used when a partition offset is a multiple of 8. By default, partitions will be created with an offset that is a multiple of 8 on Windows Server 2008 and Windows Vista.

You can choose this type of data offset when creating a storage space with System i Navigator in the Advanced Data Offset dialog. If creating a storage space with the i5/OS command CRTNWSSTG, you would specify the value *ALIGNLGLDSK in the Data offset parameter.

Storage spaces created with Align the first logical disk sector can be identified with a Data offset value of 6 when looking at the storage space properties with System i Navigator or with i5/OS command DSPNWSSTG.

Align the first logical partition sector

This data offset will align sector 63 of the storage space. This type of data offset should be used when a partition offset is 63. By default, Legacy Windows and Linux partitions will be created with an offset of 63.

You can choose this type of data offset when creating a storage space with System i Navigator in the Advanced Data Offset dialog. If creating a storage space with the i5/OS command CRTNWSSTG, you would specify the value *ALIGNLGLPTN in the Data offset parameter.

Storage spaces created with Align the first logical partition sector can be identified with a Data offset value of 7 when looking at the storage space properties with System i Navigator or with i5/OS command WRKNWSSTG.

In V5R4, the data offset can not be specified for a storage space. The storage space assumes a data offset based on the first logical partition sector as described above.

2.2 Manually creating a VMFS partition specifying an offset.

Refer to the table above to determine if you must manually create the VMFS partition.

1. Determine the storage space sequence number of the storage space.
    a. In System i Navigator, expand **Integrated Server Administration, then Servers**
    b. Expand the plus sign next to one of the ESX servers linked to the new storage space. Click on **Linked Virtual Disks.**
    c. On the right pane, look for the storage space. Take note of the **Sequence Position** number next to the storage space name. In the example below, we will be using the storage space named Bigshared with the sequence number of 3. Note: Storage spaces with a sequence position of 1 or 2 can not be used for the VMFS datastore.
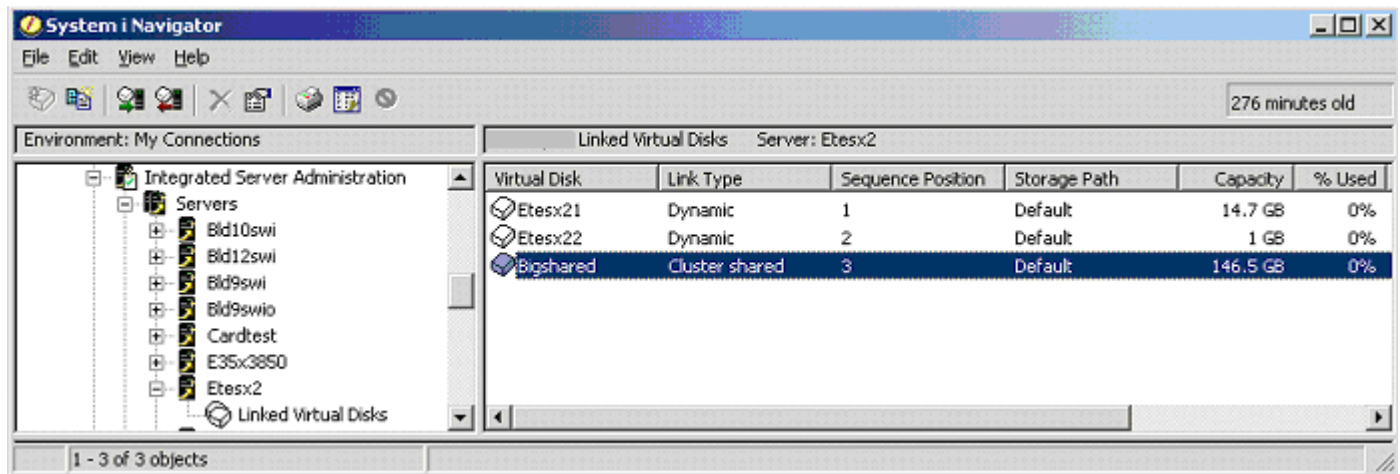
2. Ensure that the new disk is recognized by the ESX operating system. See section *7.1.4 Assigning the new storage space to all ESX servers in the Cluster* in the Redbook.
3. Identify the storage space on the ESX server.
     a. Sign on the ESX Server console as root.
     b. Run the command '**esxcfg-vmhbadevs**'. The output will show the storage devices where the first column is the VMkernel device name and the second column is the SCSI device name. The VMkernel device name has three numbers separated by colons where the third digit is the LUN. Subtracting one from the storage space sequence position is the LUN. For this example, storage space BIGSHARED has a sequence position of 3 and a LUN of 2. The output is below.

```
[root@etesx2 root]# esxcfg-vmhbadevs
vmhba1:0:0      /dev/sda
vmhba1:0:1      /dev/sdb
vmhba1:0:2      /dev/sdc
[root@etesx2 root]#
```

     c.   As an extra precaution, verify the storage space name associated with the disk on the ESX server. Run the command '**cat /proc/scsi/scsi**'. Look for the LUN number from the previous step associated with Vendor: IBM Model: VDASD xxxx where the 'xxxx' will be the storage space name. For this example, you can see that LUN 2 is associated with BIGSHARED. The output is below.

```
[root@etesx2 root]# cat /proc/scsi/scsi
Attached devices:
Host: scsi2 Channel: 00 Id: 00 Lun: 00
  Vendor: IBM      Model: VDASD ETESX21    Rev: 0001
  Type:   Direct-Access                    ANSI SCSI revision: 04
Host: scsi2 Channel: 00 Id: 00 Lun: 01
  Vendor: IBM      Model: VDASD ETESX22    Rev: 0001
  Type:   Direct-Access                    ANSI SCSI revision: 04
Host: scsi2 Channel: 00 Id: 00 Lun: 02
  Vendor: IBM      Model: VDASD BIGSHARED  Rev: 0001
  Type:   Direct-Access                    ANSI SCSI revision: 04
[root@etesx2 root]#
```

4. Create the partition with fdisk. In our example from step 3, we are going to create the partition on device /dev/sdc. Issue the command 'fdisk /dev/sdc' with the following choices.
     a. **Command** (m for help): type **p** to print the partition table. If there are existing partitions, type **d** to delete the partitions. If this is a newly created storage space, there will not be any partitions.
     b. **Command** (m for help): type **n** for new partition
     c. **Command action**: type **p** for primary partition
     d. **Partition number**: type **1**
     e. **First cylinder**: press **Enter** to accept the default
     f. **Last cylinder**: press **Enter** to accept the default
     g. **Command** (m for help): type **t** to set the partition's system ID
     h. **Hex code**: type **fb** to set the partition system ID to fb (VMware VMFS volume)
     i. **Command** (m for help): type **x** to go int expert mode
     j. **Expert command**: type **b** to adjust the starting sector number

k. **Partition number**: type **1** to choose partition 1
l. **New beginning of data**: type in the desired sector offset number. In our example we'll use 63.
m. **Expert command**: type **w** to write label and partition information to disk

Below is the command example where we are creating the VMFS partition using an offset of 63. The storage space is newly created so there were no existing partitions.

```
[root@etesx2 root]# fdisk /dev/sdc
Command (m for help): p

Disk /dev/sdc: 157.2 GB, 157292029440 bytes
255 heads, 63 sectors/track, 19123 cylinders
Units = cylinders of 16065 * 512 = 8225280 bytes

   Device Boot    Start        End     Blocks   Id  System
Command (m for help): n
Command action
   e   extended
   p   primary partition (1-4)
p
Partition number (1-4): 1
First cylinder (1-19123, default 1):
Using default value 1
Last cylinder or +size or +sizeM or +sizeK (1- 19123, default 19123):
Using default value 19123

Command (m for help): t
Selected partition 1
Hex code (type L to list codes): fb

Command (m for help):Command (m for help): x

Expert command (m for help): b
Partition number (1-4): 1
New beginning of data (62- 307210994, default 62): 63

Expert command (m for help): w
[root@etesx2 root]#
```

5. Format the partition with VMFS. Enter the following command.
```
vmkfstools –C vmfs3 –b 1m –S <volumeLabel> <partition>
```
For <volumeLabel>, use the storage space name. For <partition>, use the VMkernel device name with ':1' appended for the partition number. For this example, the partition name will be vmhba1:0:2:1 and the storage space name is BIGSHARED. Note that when executing the command, a ':' must be preceded by a '\' so that the partition name is interpreted correctly.
```
[root@etesx2 root]# vmkfstools –C vmfs3 –b 1m –S BIGSHARED vmhba1\:0\:2\:1
Creating file system on "vmhba0:0:2:1" with blockSize 307210994 and volume label
"BIGSHARED".
Successfully created new volume: 48092853-06fc748c-c934-00145e7a4c11
[root@etesx2 root]#
```

3 Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:

IBM Director of Licensing
IBM Corporation
North Castle Drive
Armonk, NY 10504-1785
U.S.A.

For license inquiries regarding double-byte (DBCS) information, contact the IBM Intellectual Property Department in your country or send inquiries, in writing, to:

IBM World Trade Asia Corporation
Licensing
2-31 Roppongi 3-chome, Minato-ku
Tokyo 106-0032, Japan

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Licensees of this program who wish to have information about it for the purpose of enabling: (i) the exchange of information between independently created programs and other programs (including this one) and (ii) the mutual use of the information which has been exchanged, should contact:

IBM Corporation
Software Interoperability Coordinator, Department YBWA
3605 Highway 52 N
Rochester, MN 55901
U.S.A.

Such information may be available, subject to appropriate terms and conditions, including in some cases, payment of a fee.

The licensed program described in this document and all licensed material available for it are provided by IBM under terms of the IBM Customer Agreement, IBM International Program License Agreement or any equivalent agreement between us.

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements

may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of 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.

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

This information is for planning purposes only. The information herein is subject to change before the products described become available.

If you are viewing this information softcopy, the photographs and color illustrations may not appear.

Trademarks

BladeCenter®, IBM®, the IBM logo, IBM System i™, IBM System x™ and i5/OS™ are trademarks or registered trademarks of International Business Machines in the United States, other countries, or both.

VMware® is a trademark or registered trademark of VMware, Inc. in the United States, other countries, or both.

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

Microsoft®, Windows® and Windows Server™ are trademarks or registered trademarks of Microsoft Corporation in the United States, other countries, or both.

Other company, product or service names may be trademarks or service marks of others.

End of Document