IBM ioMemory VSL 3.2.3

USER GUIDE FOR LINUX

MARCH 27, 2013

IBM

Legal Notices

Part Number: D0001565-004_1
Published: March 19, 2013

Introduction

Overview

Congratulations on your purchase of an IBM solid-state storage device. This guide explains how to install, troubleshoot, and maintain the software for your IBM High IOPS Adapters.

NOTE Throughout this manual, when you see a reference to an **IBM High IOPS Adapter**, you may substitute your particular device(s), such as an IBM High IOPS Adapter or each of the two IBM High IOPS Adapters of an IBM High IOPS Duo Adapter.

Attention **Products with Multiple Devices**: Some products, such as an IBM High IOPS Duo Adapter, are actually comprised of **multiple IBM High IOPS Adapters**. If your product consists of multiple IBM High IOPS Adapters, you will manage each IBM High IOPS Adapter as an independent device.

For example, if you have an IBM High IOPS Duo Adapter, you can independently attach, detach, and/or format each of the two IBM High IOPS Adapters. Each of the two devices will be presented as an individual device to your system.

About the IBM High IOPS Platform

The IBM High IOPS platform combines ioMemory VSL software with IBM High IOPS hardware to take enterprise applications and databases to the next level.

Performance

The IBM High IOPS platform provides consistent microsecond latency access for mixed workloads, multiple gigabytes per second access and hundreds of thousands of IOPS from a single product. The sophisticated IBM High IOPS architecture allows for nearly symmetrical read and write performance with best-in-class low queue depth performance, making the IBM High IOPS platform ideal across a wide variety of real world, high-performance enterprise environments.

The IBM High IOPS platform integrates with host system CPUs as flash memory to give multiple (and mostly idle) processor cores, direct and parallel access to the flash. The platform's cut-through architecture gives systems more work per unit of processing, and continues to deliver performance increases as CPU power increases.

Endurance

The IBM High IOPS platform offers best-in-class endurance in all capacities, which is crucial for caching and write-heavy databases and applications.

Reliability

The IBM High IOPS platform eliminates concerns about reliability like NAND failures and excessive wear. The all-new intelligent, self-healing feature called Adaptive Flashback provides complete, chip-level fault tolerance. Adaptive Flashback technology enables an IBM High IOPS product to repair itself after a single chip or a multi-chip failure without interrupting business continuity.

System Requirements

Please read the *IBM ioMemory VSL Release Notes* for more information on this release.

Hardware Requirements

- **Hardware Requirements**: These depend on your device (including device capacity, generation, and configuration). Please see the *IBM High IOPS Hardware Installation Guide* for requirements on the following:
  - PCIe Slot
  - Cooling
  - Power
- **Supported Devices**: Also see the *IBM High IOPS Hardware Installation Guide* for a list of supported IBM High IOPS Adapters.
- **RAM Requirements**: The *IBM ioMemory VSL Release Notes* contains memory (RAM) requirements for this version of the software.

For specific IBM High IOPS System x server configuration information and requirements, refer to the following URL: *http://www.ibm.com/support/entry/portal/docdisplay?lndocid=SERV-IOPS*

Supported Linux Distributions

- Red Hat Enterprise Linux (RHEL) 5.6, 5.7, 5.8, 6.0, 6.1, 6.2, 6.3
- SUSE Linux Enterprise Server (SLES) 10, 10 SP4, 11, 11 SP1, 11 SP2

Upgrading Legacy Adapters (IMPORTANT)

Please read these IBM High IOPS Adapter compatibility considerations.

Multiple High IOPS adapters are installed in a single system:

When multiple High IOPS Adapters are installed in the same server, all devices must operate with the same version of software. High IOPS adapters require matching firmware, drivers and utilities. This is a very important consideration when adding a new Second Generation High IOPS Adapter in a server where Legacy Adapters are deployed.

When Upgrading Legacy Adapters operating with a previous generation of software (1.2.x or v2.x), you must back up the data on the adapter before upgrading to prevent data loss. After upgrading the ioMemory VSL to version 3.x, the legacy adapters will not logically attach to the system until the firmware is also updated. Detailed instructions for upgrading software is provided in *Appendix F- Upgrading Devices from VSL 2.x to 3.x* of this user guide.

Upgrading from version 1.2.x or 2.x software to 3.x:

Upgrading Legacy adapters from 1.2.x software to version 3.1.1 offers a number of significant changes and improvements, however there are some important considerations

When performing an upgrade from 1.2.x to 3.x, you must perform a staged upgrade (upgrade to the 2.x software and firmware before upgrading to 3.x). The device driver name has also changed from fio-driver (version 1.2.x) to iomemory-vsl (2.x and above).

The upgrade process from 2.x to 3.x will require the adapter to be formatted. Formatting will remove all existing data from the card and the data must be restored after the update completes. Users must back up their data before proceeding with the upgrade process to version 3.x.

The firmware upgrade process updates and modifies important hardware settings that are not compatible with 1.2.x or 2.2.3 versions of software. Once updated, the card cannot be black-leveled to the previous versions of software. Please see the "change history" documentation for a complete list of new features, enhancements, and fixes.

Replacing a failed legacy High IOPS card and "mandatory" update requirements:

As the supply of legacy adapters diminishes from inventory, it becomes more likely that warranty replacement cards will transition to the newer versions of the High IOPS adapters. Replacement High IOPS cards and may require firmware updates to support the new or existing cards in the server.

Any situation when mixing the flash NAND technology occurs, the minimum version of software supported by the latest generation of hardware prevails. A mandatory upgrade of software is required to support the latest generation of hardware with backward compatibility to legacy cards in the server.

Change History's Update Recommendations:

Change histories files provide an ongoing list of changes to a series of software compatible with a family of hardware. Please review the change histories using the following guidelines as to how IBM recommends or suggests updates to code levels at the website below:

http://www.ibm.com/support/entry/portal/docdisplay?brand=5000008&lndocid=HELP-FIX

Software Installation

NOTE All commands require administrator privileges. Use sudo or log in as "root" to run the install.

Installation Overview

1. If you are installing this version of on a system with IBM High IOPS Adapters configured for ioMemory VSL software version 2.x installed, you must carefully follow the instructions in the *Appendix F- Upgrading Devices from VSL 2.x to 3.x* section.

   NOTE If you do not need to upgrade previous versions of IBM High IOPS Adapters to the 3.x.x firmware, but your system does have previous versions of the ioMemory VSL software installed, you will need to uninstall the ioMemory VSL software and the utilities. See the *Common Maintenance Tasks* section for instructions. Once you have uninstalled the packages, return to this page.

2. Install the latest version of the ioMemory VSL software. You can install the software as
   - A pre-compiled binary package
   - A source-to-build package
3. Install utilities and management software (included in driver installation instructions).
4. *Load the ioMemory VSL software* and *Set the Options*.
5. *Upgrade the Firmware* to the latest version, if needed (recommended). This applies to IBM High IOPS Adapters that may have a version of the firmware that is earlier than the latest version.

Installing RPM Packages

To install the Linux ioMemory VSL software and utilities:

1. You will need to install a version of the ioMemory VSL software that is built for your kernel. To determine what kernel version is running on your system, use the following command at a shell prompt:

```
$ uname -r
```

2. Compare your kernel version with the binary versions of the software available at *http://www.ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-65723* (follow that link and then select **IBM High IOPS software matrix**).
   - If there is a binary version of the software that corresponds to your kernel version, download that. **For example**:

     ```
     iomemory-vsl-<kernel-version>-<VSL-version>.x86_64.rpm
     ```

   - If there is no binary version of the software corresponding to your kernel, download the source package. **For example**:

     ```
     iomemory-vsl-<VSL-version>.src.rpm
     ```

     NOTE Exact package names may vary, depending on software and kernel version chosen.

     Use the source package that is made for your distribution. Source packages from other distributions are not guaranteed to work.

3. Download the support RPM packages you need from *http://www.ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-65723* (follow that link and then select **IBM High IOPS software matrix**). These packages provide utilities, firmware, and other files.
   - **Examples:**

Package	What is installed
`fio-util-<VSL-version>.x86_64.rpm`	ioMemory VSL utilities – **Recommended**
`fio-firmware-highiops-<version>.<date>-*.rpm`	Firmware archive package, installs the firmware archive file in a specific location (`/usr/share/fio/firmware`) – **Optional**
`highiops_<version>-<date>.fff`	Firmware archive file (not an installer package) used to upgrade the firmware, make note of where you store this file – **Recommended**
`libvsl-<version>.x86_64.rpm`	SDK libraries needed for management tools – **Recommended**, see *Monitoring and Managing Devices* for more information on available management tools.
`lib32vsl-<version>.i386.rpm`	SDK libraries needed for working with 32-bit management applications – **Optional** and only available for certain distributions. Attention You must have a full set of 32-bit system libraries installed (32-bit compatibility layer installed on your 64-bit system) before you can install this package.
`fio-common-<VSL-version>.x86_64.rpm`	Files required for the init script – **Recommended**
`fio-sysvinit-<VSL-version>.x86_64.rpm`	Init script – **Recommended**, see *Loading the ioMemory VSL Driver* for more information.

4. Change to the directory to where you downloaded the installation packages.
5. If needed, build the ioMemory VSL software from source:
   - **If you downloaded a binary version of the software**: skip the **Building the Software from Source** instructions below and continue to the next step.
   - **If you downloaded the software source package**: follow these **Building the Software from Source** instructions:

Building the Software from Source

You only need to follow these additional instructions if you downloaded the source package.

a. Install the prerequisite files for your kernel version.

NOTE Some of the prerequisite packages may already be in the default OS installation. If your system is not configured to get packages over the network, then you may need to mount your install CD/DVD.

- On RHEL 5/6, you need `kernel-devel`, `kernel-headers`, `rpm-build`, GCC4, and `rsync`.

```
$ yum install kernel-headers-`uname -r`
kernel-devel-`uname -r` gcc rsync rpm-build
make
```

This command will force `yum` to download the exact versions for your kernel. If the exact versions are no longer available in the repository, then you will have to manually search for them on the Internet.

- On SLES 10/11 you need `kernel-syms`, `make`, `rpm-build`, GCC4, and `rsync`.

```
$ zypper install kernel-syms make rpm gcc rsync
```

b. To build an RPM installation package for the current kernel, navigate to the directory with the downloaded source RPM file and run this command:

```
$ rpmbuild --rebuild
iomemory-vsl-<VSL-version>.src.rpm
```

Attention If your kernel is an Unbreakable Enterprise Kernel (UEK), you may also need to use the `--nodeps` option.

When using a .rpm source package for a non-running kernel, run this command:

```
$ rpmbuild --rebuild --define 'rpm_kernel_version
<kernel-version>'
iomemory-vsl-<VSL-version>.src.rpm
```

c. The new `RPM` package is located in a directory that is indicated in the output from the `rpmbuild` command. To find it, look for the "wrote" line. In the following example, the RPM packages are located in the `/usr/src/redhat/RPMS/x86_64/` directory.

```
...
Processing files:
iomemory-vsl-source-<version>-1.0.x86_64.rpm
Requires(rpmlib): rpmlib(PayloadFilesHavePrefix) <=
4.0-1 rpmlib(CompressedFileNames) <= 3.0.4-1
Obsoletes: iodrive-driver-source
Checking for unpackaged file(s):
/usr/lib/rpm/check-files
/var/tmp/iomemory-vsl-<version>-root
Wrote:
/usr/src/redhat/RPMS/x86_64/iomemory-vsl-2.6.18-128.el5
/usr/src/redhat/RPMS/x86_64/iomemory-vsl-source-<versi
```

In this example, `iomemory-vsl-2.6.18-128.el5-<version>-1.0.x86_64.rpm` is the package you will use.

d. Copy your custom-built software installation RPM package into the directory where you downloaded the installation packages and navigate to that directory.

e. Continue to the next step.

6. Enter the following command to install the custom-built software package. Use the package name that you just copied/downloaded into that directory.

```
$ rpm -Uvh iomemory-vsl-<kernel-version>-<VSL-version>.x86_64.rpm
```

7. Enter the following commands to install the support files:

```
$ rpm -Uvh lib*.rpm
rpm -Uvh fio*.rpm
```

Attention If the installation of the `fio-util-<VSL-version>.x86_64.rpm` package fails due to missing dependencies, you will need to install the missing packages before you run the install command again, for example:

```
$ yum install lsof pciutils
```

The ioMemory VSL software and utilities are installed to the following locations:

Package Type	Installation Location
ioMemory VSL software	`/lib/modules/<kernel-version>/extra/fio/iomemory-vsl.ko`
Utilities	`/usr/bin`

NOTE You may also install the IBM High IOPS Management Application (optional GUI management software). IBM High IOPS Management Application and documentation are available as a separate download.

Once the packages are installed, continue to *Loading the ioMemory VSL Driver* later in the section.

Loading the ioMemory VSL Driver

To load the ioMemory VSL driver:

Run this command:

```
$ modprobe iomemory-vsl
```

NOTE The ioMemory VSL driver automatically loads at system boot. The IBM High IOPS Adapter is now available to the OS as `/dev/fiox`, where *x* is a letter (i.e., a, b, c, etc.).

To confirm the IBM High IOPS Adapter is attached, run the `fio-status` utility from the command line. The output lists each drive and its status (attached or not attached).

Attention If the IBM High IOPS Adapter is not automatically attaching, check the /etc/modprobe.d files to ensure that the `auto_attach` option is turned on (set to 1).

- For this command to work on SLES 10 systems, you must edit the `/etc/init.d/iomemory-vsl` file's init info and change `udev` to `boot.udev`. The file should look like this:

```
### BEGIN INIT INFO
# Provides:       iomemory-vsl
# Required-Start: boot.udev
```

- On SLES systems, you must also allow unsupported modules for this command to work.
  - **SLES 11 Update 2**: Modify the `/etc/modprobe.d/iomemory-vsl.conf` file and uncomment the appropriate line:

```
# To allow the ioMemory VSL driver to load on SLES11, uncomment
below
allow_unsupported_modules 1
```

  - **SLES 10 SP4**: Modify the `/etc/sysconfig/hardware/config` file so the `LOAD_UNSUPPORTED_MODULES_AUTOMATICALLY` sysconfig variable is set to `yes`, for example:

```
LOAD_UNSUPPORTED_MODULES_AUTOMATICALLY=yes
```

Controlling ioMemory VSL software Loading

You can control driver loading either through the init script or through udev.

In newer Linux distributions, users can rely on the udev device manager to automatically find and load drivers for their installed hardware at boot time, though udev can be disabled and the init script used in nearly all cases. We recommend using the init script to load the ioMemory VSL driver if you are managing a RAID array using LVM, mdadm, or Veritas Storage Foundation.

For older Linux distributions (such as SLES 10) without udev functionality, users must rely on a boot-time init script to load needed drivers.

Using the init Script

On systems where udev loading of the driver doesn't work, or is disabled, the init script may be enabled to load the driver at boot. On some distros it may be enabled by default.

NOTE The init Script is part of the `fio-sysvinit` package, which must be installed before you can enable it.

You can disable this loading of the driver with the following command:

```
$ chkconfig --del iomemory-vsl
```

To re-enable the driver loading in the init script, use the following command:

```
$ chkconfig --add iomemory-vsl
```

The ioMemory VSL software install process places an init script in `/etc/init.d/iomemory-vsl`. In turn, this script uses the setting options found in the options file in `/etc/sysconfig/iomemory-vsl`. The options file must have ENABLED set (non-zero) for the init script to be used:

```
ENABLED=1
```

The options file contains documentation for the various settings---two of which, `MOUNTS` and `KILL_PROCS_ON_UMOUNT`, are discussed further in the *Handling Driver Unloads* section below.

Mounting Filesystems when Using the init Script

Because the ioMemory VSL driver does not load by the standard means (in the `initrd`, or built into the kernel), using the standard method for mounting filesystems (`/etc/fstab`) for filesystems hosted on the IBM High IOPS Adapter does not work. To set up auto-mounting of a filesystem hosted on an IBM High IOPS Adapter:

1. Add the filesystem mounting command to `/etc/fstab` as normal.
2. You must add the '`noauto`' option and the '`0 0`' flag to `/etc/fstab` as in the two following sample entries.

```
/dev/fioa /mnt/fioa ext3 defaults,noauto 0 0
/dev/fiob1 /mnt/ioDrive ext3 defaults,noauto 0 0
```

(where the `a` in `fioa` can be `a`, `b`, `c`, etc., depending on how many IBM High IOPS Adapters you have installed in the system).

Attention Failure to add '`noauto 0 0`' to `fstab` may cause a boot failure.

To have the init script mount these drives after the driver is loaded and unmounted and before the driver is unloaded, add a list of mount points to the options file using the procedure documented there.

For the filesystem mounts shown in the earlier example, the line in the options file would look like this:

```
MOUNTS="/mnt/fioa /mnt/iodrive"
```

Using udev

On systems that rely on udev to load drivers, users need to modify an ioMemory VSL software options file if they want to prevent udev from auto-loading the ioMemory VSL at boot time. To do this, locate and edit the `/etc/modprobe.d/iomemory-vsl.conf` file that already has the following line:

```
# blacklist iomemory-vsl
```

To disable loading, remove the "#" from the line and save the file.

With the blacklist command in place, restart Linux. The ioMemory VSL driver will not be loaded by udev.

To restore the udev-loading of the driver, replace the "#" to comment out the line.

On either udev or init script systems

Users can disable the loading of the driver at boot time, and thus prevent the auto-attach process for diagnostic or troubleshooting purposes on either udev or init script systems. Follow the steps in the *Disabling Auto-Attach* section to disable or re-enable the auto-attach functionality.

Alternatively, you can prevent the ioMemory VSL driver from loading by appending the following parameter at the kernel command line of your boot loader:

```
iodrive=0
```

However, this method is not preferred as it prevents the driver from functioning at all, thus limiting the amount of troubleshooting you can perform.

IBM High IOPS Adapters and Multipath Storage

If you are using IBM High IOPS Adapters along with multipath storage, you must blacklist the IBM High IOPS Adapters to prevent `device-mapper` from trying to create a `dm-device` for each IBM High IOPS Adapter. **This must be done prior to activating `dm-multipath` and/or loading the driver**. If IBM High IOPS Adapters are not blacklisted, they will appear busy and you will not be able to attach, detach, or update the firmware on the devices.

To blacklist IBM High IOPS Adapters, edit the `/etc/multipath.conf` file and include the following:

```
blacklist {
devnode         "^fio[a-z]"
}
```

Handling Driver Unloads

Special consideration must be taken during VSL software driver unload time. By default, the init script searches for any processes holding open a mounted filesystem and kills them, thus allowing the filesystem to be unmounted. This behavior is controlled by the option `KILL_PROCS_ON_UMOUNT` in the options file. If these processes are not killed, the filesystem cannot be unmounted. This may keep the ioMemory VSL software from unloading cleanly, causing a significant delay on the subsequent boot.

Setting the ioMemory VSL Options

This section explains how to set ioMemory VSL software options. For more information about setting specific options, see *Appendix C- Using Module Parameters*.

One-Time Configuration

VSL software options can be set at install time, on the command line of either `insmod` or `modprobe`. For example, to set the `auto_attach` option to 0, run the command:

```
$ modprobe iomemory-vsl auto_attach=0
```

This option takes effect only for this load of the ioMemory VSL software; subsequent calls to `modprobe` or `insmod` will not have this option set.

Persistent Configuration

To maintain a persistent setting for an option, add the desired option to `/etc/modprobe.d/iomemory-vsl.conf` or a similar file. To prevent the IBM High IOPS Adapters from auto-attaching, add the following line to the `iomemory-vsl.conf` file:

```
options iomemory-vsl auto_attach=0
```

This option then takes effect for every subsequent ioMemory VSL software load, as well as on autoload of the VSL software during boot time.

Upgrading the Firmware

With the ioMemory VSL software loaded, you need to check to ensure that the IBM High IOPS Adapter's firmware is up-to-date. To do this, run the ***`fio-status`*** command-line utility.

If the output shows that the device is running in minimal mode, download the latest firmware from *http://www.ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-65723* (follow that link and then select **IBM High IOPS software matrix**), then use the IBM High IOPS Management Application software or the ***`fio-update-iodrive`*** utility to upgrade the firmware.

Attention **Upgrade Path**: There is a specific upgrade path that you must take when upgrading IBM High IOPS Adapter. Consult the *IBM ioMemory VSL Release Notes* for this ioMemory VSL software release before upgrading IBM High IOPS Adapters.

Your IBM High IOPS Adapter may have a minimum firmware label affixed (for example, "MIN FW: XXXXXX"). This label indicates the minimum version of the firmware that is compatible with your device.

Attention Do not attempt to downgrade the firmware on any IBM High IOPS Adapter, doing so may void your warranty.

When installing a new IBM High IOPS Adapter along with existing devices, you must upgrade all of the currently installed devices to the latest available versions of the firmware and ioMemory VSL software before installing the new devices.

Consult the *IBM ioMemory VSL Release Notes* for this ioMemory VSL software release for any upgrade considerations.

Upgrading VMware Guest OS

If you are using your IBM High IOPS Adapter with a VMware guest OS (using VMDirectPathIO), you must cycle the power on the host after you upgrade the device(s). Just restarting the virtual machine won't apply the change.

Virtual Controller Configuration

Depending on your use case and application, you may benefit from configuring supported devices to use Virtual Controller technology.

When configured, each physical IBM High IOPS Adapter is split into two (virtual) logical devices. Splitting the IBM High IOPS Adapter into two virtual devices has the following implications:

- **Latency**: There is no affect on latency.
- **Throughput**: The total peak I/O bandwidth of the device is approximately the same.
- **IOPS**: Depending on the use of the virtual devices (especially the average I/O size), the peak IOPS for each virtual device is about the same for a non-split device. In other words, the combined peak IOPS of the two virtual devices can be nearly double that of a non-split device. For details, see *Virtual Controller Conversion* in the *Maintenance* section.
- **Capacity:** Due to virtualization overhead, the combined capacity of the two virtual devices is slightly less than that of a single-controller device. See the *IBM ioMemory VSL Release Notes* for a list of compatible devices and their Virtual Controller capacities.

For more information on converting a device to use Virtual Controller technology, including device requirements, conversion steps, and additional considerations, see *Virtual Controller Conversion* in the *Maintenance* section.

Using the Device as Swap

To safely use the IBM High IOPS Adapter as swap space requires passing the preallocate_memory kernel module parameter. The recommended method for providing this parameter is to add the following line to the `/etc/modprobe.d/iomemory-vsl.conf` file:

```
options iomemory-vsl preallocate_memory=1072,4997,6710,10345
```

- Where 1072,4997,6710,10345 are serial numbers obtained from `fio-status`.

A 4K sector size format is required for swap—this reduces the ioMemory VSL software memory footprint. Use `fio-format` to format the IBM High IOPS Adapter with 4K sector sizes.

NOTE Be sure to provide the serial numbers for the IBM High IOPS Adapter, not an adapter, when applicable.

NOTE The `preallocate_memory` module parameter is necessary to have the device usable as swap space. See *Appendix C- Using Module Parameters* for more information on setting this parameter.

Attention You must have enough RAM available to enable the IBM High IOPS Adapter with pre-allocation enabled for use as swap. Attaching an IBM High IOPS Adapter, with pre-allocation enabled, without sufficient RAM may result in the loss of user processes and system instability.

Consult the *IBM ioMemory VSL Release Notes* for RAM requirements with this version of the ioMemory VSL software.

NOTE The `preallocate_memory` parameter is recognized by the ioMemory VSL software at load time, but the requested memory is not actually allocated until the specified device is attached.

Using the Logical Volume Manager

The Logical Volume Manager (LVM) volume group management application handles mass storage devices like IBM High IOPS Adapters if you add the IBM High IOPS Adapter as a supported type:

1. Locate and edit the `/etc/lvm/lvm.conf` configuration file.
2. Add an entry similar to the following to that file:

```
types = [ "fio", 16 ]
```

The parameter "16" represents the maximum number of partitions supported by the device.

NOTE If using LVM or MD, do not use udev to load the ioMemory VSL driver. The init script will ensure that the LVM volumes and MD devices are detached before attempting to detach the IBM High IOPS Adapter.

Configuring RAID using the Logical Volume Manager

The simplest way to using your IBM High IOPS Adapters with LVM is to use the entire block device with it (for example: `/dev/fioa` rather than `/dev/fioa1`, `/dev/fioa2`, etc.). This way the block device does not need to partitioned ahead of time, though partitioning is also supported. The examples that follow assumes the entire block device is used.

Whether you plan to stripe (RAID 0) or mirror (RAID 1) the devices, the first two steps are the same:

1. First, create physical volumes, for example:

```
$ pvcreate /dev/fioa  /dev/fiob
```

2. Next add these physical volumes to a volume group, for example:

```
$ vgcreate iomemory_vg /dev/fioa /dev/fiob
```

Creating a Striped Volume (RAID 0) Using LVM

With the volume group created, you can create logical volumes within this volume group. In this instance, only one will be created and we name it `iomemory_lv`:

1. Create the striped volume using the `-i2` option for two stripes, for example:

```
$ lvcreate -l 100%VG -n iomemory_lv -i2 iomemory_vg
```

2. Create a file system on the newly created volume, for example:

```
$ mkfs.ext3 /dev/iomemory_vg/iomemory_lv
```

3. In order to make this persistently mount at boot time, edit the `/etc/sysconfig/iomemory-vsl` file by adding the volume group path under the example line:

```
# Example: LVM_VGS="/dev/vg0 /dev/vg1"
LVM_VGS="/dev/iomemory_vg"
```

   Be sure to just add the volume group path and not the logical volume as well. For more information on using the init script and the `/etc/sysconfig/iomemory-vsl` file, see *Loading the ioMemory VSL Driver* earlier in this guide.

Creating a Mirrored Volume (RAID 1) Using LVM

With the volume group created, you can create logical volumes within this volume group. In this instance, only one will be created and we name it `iomemory_lv`:

1. Create the mirrored volume using the `-m1` option for an original linear volume plus one copy, for example:

```
$ lvcreate -l 100%VG -n iomemory_lv -m1 --corelog iomemory_vg
```

   You can monitor the progress of the initial mirror synchronization using the lvs command. For example:

```
lvs -a -o +devices
```

   Attention With the `--corelog option`, LVM uses an in-memory region or log to track the state of the mirror legs. Since it is in-memory and lost at reboot, it must be regenerated at boot by scanning the mirror legs. For alternative configurations, consult the LVM documentation.

2. Create a file system on the newly created volume, for example:

```
$ mkfs.ext3 /dev/iomemory_vg/iomemory_lv
```

3. In order to make this persistently mount at boot time, edit the `/etc/sysconfig/iomemory-vsl` file by adding the volume group path under the example line:

```
# Example: LVM_VGS="/dev/vg0 /dev/vg1"
LVM_VGS="/dev/iomemory_vg"
```

   Be sure to just add the volume group path and not the logical volume as well. For more information on using the init script and the `/etc/sysconfig/iomemory-vsl` file, see *Loading the ioMemory VSL Driver* earlier in this guide.

Configuring RAID Using mdadm

You can configure two or more IBM High IOPS Adapters into a RAID array using software-based RAID solutions.

NOTE If you are using RAID1/Mirrored and one device fails, be sure to run `fio-format` on the replacement device (not the existing, good device) before rebuilding the RAID. Following are some examples of some common RAID configurations using the `mdadm` utility.

Attention The Linux kernel RAID 5 implementation performs poorly at high data rates. This is an issue in the Linux kernel. Alternatives include using RAID 10, or possibly a third-party RAID stack.

Mounting Arrays

Once you are done making your array (by following one of the configuration samples below), you must edit the `/etc/sysconfig/iomemory-vsl` file and add the array path under the example line. This will make the array mount at boot time.

In all of the examples below, we create arrays named `md0`. Here is is how you would edit the `/etc/sysconfig/iomemory-vsl` file to include the array using `md0` as an example (you add the array just below the example line in that file):

```
# Example: MD_ARRAYS="/dev/md0 /dev/md1"
MD_ARRAYS="/dev/md0"
```

For more information on using the init script and the `/etc/sysconfig/iomemory-vsl` file, see *Loading the ioMemory VSL Driver* earlier in this guide.

RAID 0

To create a striped set, where `fioa` and `fiob` are the two IBM High IOPS Adapters you want to stripe, run this command:

```
$ mdadm --create /dev/md0 --chunk=256 --level=0 --raid-devices=2 /dev/fioa
/dev/fiob
```

Making the Array Persistent (Existing after Restart)

NOTE On some versions of Linux, the configuration file is in `/etc/mdadm/mdadm.conf`, not `/etc/mdadm.conf`.

Inspect `/etc/mdadm.conf`. If there are one or more lines declaring the devices to inspect, make sure one of those lines specifies "partitions" as an option. If it does not, add a new DEVICE line to the file specifying "partitions" like this:

```
DEVICE partitions
```

Also add a device specifier for the fio IBM High IOPS Adapters:

```
DEVICE /dev/fio*
```

To see if any updates are needed to `/etc/mdadm.conf`, issue the following command:

```
$ mdadm --examine --scan
```

Compare the output of this command to what currently exists in `mdadm.conf` and add any needed sections to `/etc/mdadm.conf`.

NOTE For example, if the array consists of two devices, there will be three lines in the output of the command that are not present in the `mdadm.conf` file: one line for the array, and two device lines (one line for each device). Be sure to add those lines to the `mdadm.conf` so it matches the output of the command.

For further details please see the `mdadm` and `mdadm.conf` man pages for your distribution.

With these changes, on most systems the RAID 0 array will be created automatically upon restart. However, if you have problems accessing `/dev/md0` after restart, run the following command:

```
$ mdadm --assemble --scan
```

You may also want to disable udev loading of the ioMemory VSL driver, if needed, and use the init script provided for driver loading. Please see the *Using the Init Script* section of this guide for further details on how to use the init script.

NOTE In SLES 11, you may need to run the following commands to make sure these services are run on boot:

```
chkconfig boot.md on
chkconfig mdadmd on
```

RAID 1

To create a mirrored set using the two IBM High IOPS Adapters `fioa` and `fiob`, run this command:

```
$ mdadm --create /dev/md0 --level=1 --raid-devices=2 /dev/fioa /dev/fiob
```

RAID 10

To create a striped, mirrored array using four IBM High IOPS Adapters (fioa, fiob, fioc, and fiod), run this command:

```
$ mdadm --create /dev/md0 -v --chunk=256 --level=raid10 --raid-devices=4
/dev/fioa /dev/fiob /dev/fioc /dev/fiod
```

Building a RAID10 Across Multiple Devices

In a RAID10 configuration, sets of two disks are mirrored, and then those mirrors are striped. When setting up a RAID10 across multiple IBM High IOPS Adapters, it is best to make sure that no mirror resides solely on the two IBM High IOPS Adapters that comprise an a single product (such as an IBM High IOPS Duo Adapter).

In order to get the data to lay out properly,

- Use the `--layout=n2` option when creating the RAID10 (though it should be the default)
- Ensure that no two IBM High IOPS Adapters from the same device are listed side by side.

The following sample code shows some recommended configurations.

NOTE The following commands assume that all IBM High IOPS Adapters have been freshly formatted with the `fio-format` utility.

Attention The ordering of the fiox devices is critical.

```
# 2 Dual Device Products RAID10
$ mdadm --create --assume-clean --level=raid10 --layout=n2 -n 4 /dev/md0 \
 /dev/fioa /dev/fioc \
 /dev/fiob /dev/fiod
# Mirror groups are: fioa,fioc and fiob,fiod

# 3 Dual Device Products RAID10
$ mdadm --create --assume-clean --level=raid10 --layout=n2 -n 6 /dev/md0 \
 /dev/fioa /dev/fiod \
 /dev/fioc /dev/fiof \
 /dev/fioe /dev/fiob

# 4 Dual Device Products RAID10
$ mdadm --create --assume-clean --level=raid10 --layout=n2 -n 8 /dev/md0 \
 /dev/fioa /dev/fiod \
 /dev/fioc /dev/fiof \
 /dev/fioe /dev/fioh \
 /dev/fiog /dev/fiob

# 8 Dual Device Products RAID10
$ mdadm --create --assume-clean --level=raid10 --layout=n2 -n 16 /dev/md0 \
 /dev/fioa /dev/fiod \
 /dev/fioc /dev/fiof \
 /dev/fioe /dev/fioh \
 /dev/fiog /dev/fioj \
 /dev/fioi /dev/fiol \
 /dev/fiok /dev/fion \
 /dev/fiom /dev/fiop \
 /dev/fioo /dev/fiob
```

Understanding Discard (TRIM) Support

With this version of the ioMemory VSL software, Discard (also known as TRIM) is enabled by default.

Discard addresses an issue unique to solid-state storage. When a user deletes a file, the device does not recognize that it can reclaim the space. Instead the device assumes the data is valid.

Discard is a feature on newer filesystem releases. It informs the device of logical sectors that no longer contain valid user data. This allows the wear-leveling software to reclaim that space (as reserve) to handle future write operations.

Discard (TRIM) on Linux

Discard is enabled by default in the ioMemory VSL software release. However, for discard to be implemented, the Linux distribution must support this feature, and discard must be turned on.

In other words, if your Linux distribution supports discard, and discard is enabled on the system, then discard will be implemented on your IBM High IOPS Adapter.

Under Linux, discards are not limited to being created by the filesystem, discard requests can also be generated directly from userspace applications using the kernel's discard ioctl.

Attention There is a known issue that ext4 in Kernel.org 2.6.33 or earlier may silently corrupt data when discard is enabled. This has been fixed in many kernels provided by distribution vendors. Please check with your kernel provider to be sure your kernel properly supports discard. For more information, see the Errata in the *IBM ioMemory VSL Release Notes* for this version of the software

NOTE On Linux, MD and LVM do not currently pass discards to underlying devices. Thus any IBM High IOPS Adapter that is part of an MD or LVM array will not receive discards sent by the filesystem.

The LVM release included in Red Hat 6.1 supports passing discards for several targets, but not all (*RHEL 6.1 documentation*). Please see your distribution's documents for exact details.

Performance and Tuning

IBM High IOPS Adapters provide high bandwidth, high Input/Output per Second (IOPS), and are specifically designed to achieve low latency.

As IBM High IOPS Adapters improve IOPS and low latency, the device performance may be limited by operating system settings and BIOS configuration. These settings may need to be tuned to take advantage of the revolutionary performance of IBM High IOPS Adapters.

While IBM High IOPS Adapters generally perform well out of the box, this section describes some of the common areas where tuning may help achieve optimal performance.

Disable CPU Frequency Scaling

Dynamic Voltage and Frequency Scaling (DVFS) are power management techniques that adjust the CPU voltage and/or frequency to reduce power consumption by the CPU. These techniques help conserve power and reduce the heat generated by the CPU, but they adversely affect performance while the CPU transitions between low-power and high-performance states.

These power-savings techniques are known to have a negative impact on I/O latency and IOPS. When tuning for performance, you may benefit from reducing or disabling DVFS completely, even though this may increase power consumption.

DVFS, if available, is often configurable as part of your operating systems power management features as well as within your system's BIOS interface. Within the operating system and BIOS, DVFS features are often found under the Advanced Configuration and Power Interface (ACPI) sections; consult your computer documentation for details.

Limiting ACPI C-States

Newer processors have the ability to go into lower power modes when they are not fully utilized. These idle states are known as ACPI C-states. The `C0` state is the normal, full power, operating state. Higher C-states (`C1`, `C2`, `C3`, etc.) are lower power states.

While ACPI C-states save on power, they can have a negative impact on I/O latency and maximum IOPS. With each higher C-state, typically more processor functions are limited to save power, and it takes time to restore the processor to the `C0` state.

When tuning for maximum performance you may benefit from limiting the C-states or turning them off completely, even though this may increase power consumption.

Setting ACPI C-State Options

If your processor has ACPI C-states available, you can typically limit or disable them in the BIOS interface (sometimes referred to as a Setup Utility). APCI C-states may be part of of the Advanced Configuration and Power Interface (ACPI) menu. Consult your computer documentation for details.

C-States Under Linux

Newer Linux kernels have drivers that may attempt to enable APCI C-states even if they are disabled in the BIOS. You can limit the C-state in Linux (with or without the BIOS setting) by adding the following to the kernel boot options:

```
intel_idle.max_cstate=0 processor.max_cstate=0
```

In this example, the maximum C-state allowed will be `C0` (disabled).

Setting NUMA Affinity

Servers with a NUMA (Non-Uniform Memory Access) architecture may require special installation instructions in order to maximize IBM High IOPS Adapter performance. This includes most multi-socket servers.

On some servers with NUMA architecture, during system boot, the BIOS will not associate PCIe slots with the correct NUMA node. Incorrect mappings result in inefficient I/O handling that can significantly degrade performance. To prevent this, you must manually assign ioMemory devices optimally among the available NUMA nodes.

See *Appendix E- NUMA Configuration* for more information on setting this affinity.

Setting the Interrupt Handler Affinity

Device latency can be affected by placement of interrupts on NUMA systems. We recommend placing interrupts for a given device on the same NUMA node that the application is issuing I/O from. If the CPUs on this node are overwhelmed with user application tasks, in some cases it may benefit performance to move the the interrupts to a remote node to help load-balance the system.

Many operating systems will attempt to dynamically place interrupts across the nodes, and generally make good decisions.

Linux IRQ Balancing

In Linux this dynamic placement is called IRQ Balancing. You can check to see if the IRQ balancer is effective by checking `/proc/interrupts`. If the interrupts are unbalanced (too many device interrupts on one node) or on an overwhelmed node, you may need to stop the IRQ balancer and manually distribute the interrupts in order to balance the load and improve performance.

NOTE Restarting the IRQ Balancer after the ioMemory VSL software loads (and the IBM High IOPS Adapters are

attached) may resolve interrupt affinity issues. For example, run:

```
/etc/init.d/irq_balancer start
```

If that does not resolve the affinity issues, then we recommend manual pinning the device interrupts to specific nodes.

Hand-tuning interrupt placement in Linux is an advanced option that requires profiling of application performance on any given hardware. Please see your operating system documentation for information on how to pin specific device interrupts to specific nodes.

Maintenance

The ioMemory VSL software includes utilities for maintaining the device.

GUI Management

IBM High IOPS Management Application is a free browser-based solution for managing IBM High IOPS Adapters. IBM High IOPS Management Application and documentation are available as a separate download.

The IBM High IOPS Management Application can perform many management functions, including:

- Firmware upgrades
- Low-level formatting
- Attach and detach actions
- Device status and performance information
- Configure Swap and Paging
- Generate bug reports

Command-line Utilities

Several command-line utilities are included in the installation packages for managing your IBM High IOPS Adapter:

- `fio-attach`
- `fio-beacon`
- `fio-bugreport`
- `fio-detach`
- `fio-format`
- `fio-pci-check`
- `fio-status`
- `fio-sure-erase`
- `fio-update-iodrive`

For more information on command-line utilities, see *Appendix A- Utilities Reference*

Common Maintenance Tasks

The following are the most common tasks for maintaining your IBM High IOPS Adapter using command-line utilities.

NOTE All commands require administrator privileges. Log in as "root" or use sudo to run the commands.

NOTE If you came to this section from the *Software Installation* section, *return* to that section after you uninstall previous versions of the driver and utilities.

Unloading the ioMemory VSL driver

To unload the driver, run this command:

```
$ modprobe -r iomemory-vsl
```

Uninstalling the ioMemory VSL RPM Package

With versions 2.x and later (including 3.x releases) of the ioMemory VSL software, you must specify the kernel version of the package you are uninstalling. Run this command to find the installed driver packages:

```
$  rpm -qa | grep -i iomemory
```

Sample output:

```
iomemory-vsl-2.6.18-194.el5-2.2.2.82-1.0
```

Uninstall the ioMemory VSL software by running a command similar to this example (specify the kernel version of the driver you wish to uninstall):

```
$ rpm -e iomemory-vsl-2.6.18-194.el5-2.2.0.82-1.0
```

Uninstalling the ioMemory VSL Utilities and Other Support Packages

Uninstalling 2.x Support Packages

To uninstall the support RPM packages, run this command (adding or removing package names as needed):

```
$ rpm -e fio-util fio-snmp-agentx fio-common fio-firmware iomanager-gui
iomanager-jre libfio libfio-doc libfusionjni fio-sysvinit fio-smis
fio-snmp-mib libfio-dev
```

Uninstalling 3.x Support Packages

To uninstall the support RPM packages, run this command (adding or removing package names as needed):

```
$ rpm -e fio-util fio-snmp-agentx fio-common fio-firmware libvsl libvsl-doc
fio-sysvinit fio-smis fio-snmp-mib libvsl-dev
```

Virtual Controller Conversion

Converting your IBM High IOPS Adapter to a Virtual Controller configuration will split the IBM High IOPS Adapter into two logical devices.

For 512B I/Os, the combined IOPS performance of the two virtual devices is approximately double that of a single-controller device. For 4KB I/Os, there is more than an 80% improvement in IOPS performance with virtual devices. For 16KB and larger I/Os, there is no improvement of total IOPS performance over a non-Virtual Controller configuration.

Latency in the virtual devices is unaffected, and the combined bandwidth of the two virtual devices is the same as it would be without the split. Due to the overhead of an additional device, the combined capacity of the two virtual devices is slightly less than that of a single-controller device.

Splitting a single physical device into multiple virtualized devices, or merging multiple virtualized devices back to a single physical device requires a low-level format, which will erase all of the data on the device. Be sure to back up all of your data.

Supported Devices

Only relatively new devices (with few writes performed) may be split or merged. Devices with too much wear are unsuitable for converting to or from a Virtual Controller configuration. Merging virtual devices may also result in additional wear (depending on the wear differences of the two virtual devices).

To be suitable for splitting or merging, devices (including Virtual Controller devices) must have 90% or more of their remaining rated endurance of Petabytes Written (PBW). This rating as well as the current percentage remaining is visible in `fio-status` with the `-a` option. For example:

```
fio-status /dev/fct1 -a
...
 Rated PBW: 17.00 PB, 99.95% remaining
```

In the above example, the device is suitable for conversion because it has more than 90% of the rated PBW remaining.

If you attempt to merge or split a device that does not support Virtual Controller technology or a device that has too much wear, the update utility will not allow the conversion and the firmware upgrade will not take place. See the Release Notes for a list of devices that support Virtual Controller technology and their capacities after the conversion.

Multi-device Products

For products with more than one IBM High IOPS Adapter, such as an IBM High IOPS Duo Adapter, you must configure all of the IBM High IOPS Adapters to Virtual Controller technology at the same time. All of the devices must also be merged at the same time. For example, the two IBM High IOPS Adapters in an IBM High IOPS Duo Adapter will be converted into four virtual devices. The utility will not allow a conversion if you attempt to split or merge only one physical device in a multi-device product.

Splitting Controllers

Be sure to use firmware that supports Virtual Controller technology. Consult the Release Notes to determine if the firmware for that release supports Virtual Controller technology.

1. Back up all of your data. Because a low-level format is needed to complete the conversion, all of the user data on your device will be erased.
2. Use the `fio-update-iodrive` command-line utility to configure an IBM High IOPS Adapter to use Virtual Controller technology:
   - Use the `--split` option to split the controller.
   - Use the `-d` option to specify a device, otherwise all installed devices that can be split will be split.
   - Specify the firmware path, and check the *IBM ioMemory VSL Release Notes* to make sure the firmware supports Virtual Controller technology.

     Example:

     ```
     fio-update-iodrive --split -d /dev/fct0 <firmware-path>
     ```

     After rebooting, each physical device will be split into two virtual devices. Each IBM High IOPS Adapter will therefore split into two logical devices, each with a unique device path. For example, `/dev/fct0` may become `/dev/fct0` and `/dev/fct1`. You will manage each device as a unique device.
3. Reboot.
4. Load the ioMemory VSL driver.
5. Run `fio-status` to determine which devices need to be formatted.

6. Low-level format the device(s). For example:

```
fio-format /dev/fct0 /dev/fct1
```

Attention Formatting will erase all user data, be sure to back up your data. You can reverse the split by merging the controllers (without losing data) up until you format the virtual devices.

Merging Controllers

If your IBM High IOPS Adapter (including the two virtual devices) is suitable for merging, then you will be able to use the `fio-update-iodrive` utility to merge the virtual devices back into one physical device.

1. Back up all of your data. Because a low-level format is needed to complete the merge, all of the user data on your device will be erased.
2. Use the `fio-update-iodrive` command-line utility to configure the device for merging:
   - Use the `--merge` option to merge the virtual devices.
   - Use the `-d` option to specify a device.

     Attention The `fio-update-iodrive` utility only successfully works against one of the two virtual devices for each physical IBM High IOPS Adapter. Out of the two virtual devices, only the first virtual device (in terms of device numbering) is linked to the physical device (and the firmware). The second virtual device is not linked, and any firmware operation against that second virtual device will fail with this message:

     ```
     Error: Device '/dev/fctx' had an error while updating.
     This device does not support firmware update.
     ```

     This is expected, and the error will not affect the update/merge of the first (linked) virtual device. The update operation will complete on all devices that can merge and otherwise accept firmware changes.

   - Specify the firmware path, and check the *IBM ioMemory VSL Release Notes* to make sure the firmware supports Virtual Controller technology.

     Example:

     ```
     fio-update-iodrive --merge -d /dev/fct0 <firmware-path>
     ```

3. Reboot.
4. Load the ioMemory VSL driver.
5. Run `fio-status` to determine which devices need to be formatted.

6. Low-level format the device(s). For example:

```
fio-format /dev/fct0
```

Attention Formatting will erase all user data, be sure to back up your data. You can reverse the merge by splitting the controllers (without losing data) up until you format the merged device.

The IBM High IOPS Adapter is once again one logical device, and you will manage it as one device.

Disabling Auto-Attach

When the ioMemory VSL software is installed, it is configured to automatically attach any devices when the ioMemory VSL software is loaded. Sometimes you may want to disable the auto-attach feature. To do so:

1. Edit the following file:

```
/etc/modprobe.d/iomemory-vsl.conf
```

2. Add the following line to that file:

```
options iomemory-vsl auto_attach=0
```

3. Save the file. To re-enable auto-attach, simply edit the file and either remove that line or change it to the following:

```
options iomemory-vsl auto_attach=1
```

Unmanaged Shutdown Issues

Unmanaged shutdowns due to power loss or other circumstances can force the IBM High IOPS Adapter to perform a consistency check during the restart. This may take several minutes or more to complete.

Although data written to the IBM High IOPS Adapter is not lost due to unmanaged shutdowns, important data structures may not have been properly committed to the device. This consistency check repairs these data structures.

Disabling the ioMemory VSL Software

The ioMemory VSL software automatically loads by default when the operating system starts. You can disable the ioMemory VSL software for diagnostic or troubleshooting purposes.

To disable auto-load, uninstall the VSL software to keep it from loading, or move it out of the `/lib/modules/<kernel_version>` directory.

Monitoring and Managing Devices

IBM provides many tools for managing your IBM High IOPS Adapters. These tools will allow you to monitor the devices for errors, warnings, and potential problems. They will also allow you to manage the devices including performing the following functions:

- Firmware upgrades
- Low-level formatting
- Attach and detach actions
- Device status and performance information
- Configuring Swap and Paging
- Generating bug reports

Management Tools

IBM has provided several tools for monitoring and managing IBM High IOPS Adapters. These include stand-alone tools that require no additional software and data-source tools that can be integrated with other applications.

Consider the descriptions of each tool to decide which tool (or combination of tools) best fits your needs.

Attention The ioMemory VSL software does print some error messages to the system logs, and while these messages are very useful for troubleshooting purposes, the VSL software log messages are not designed for continual monitoring purposes (as each is based on a variety of factors that could produce different log messages depending on environment and use case). For best results, use the tools described in this section to regularly monitor your devices.

Stand-alone Tools

These stand-alone tools do not require any additional software.

- **Command-line Utilities**: These utilities are run manually in the terminal. The `fio-status` utility provides status for all devices within a host. The other utilities allow you to perform other management functions. See *Appendix A- Utilities Reference* for full details.

- **IBM High IOPS Management Application**: The GUI browser-based IBM High IOPS Management Application allows you to monitor and manage every IBM High IOPS Adapter installed in multiple hosts across your network. It collects all of the alerts for all IBM High IOPS Adapters and displays them in the Alert Tab. You may also set up the IBM High IOPS Management Application to send email or SMS messages for specific types of alerts or all alerts. The software packages and documentation are available from *http://www.ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-65723* (follow that link and then select **IBM High IOPS software matrix**).

Data-source Tools

These data-source tools provide comprehensive data, just like the stand-alone tools, but they do require integration with additional software. At a minimum, some tools can interface with a browser. *However, the benefit of these tools is that they can be integrated into existing management software that is customized for your organization.*

These tools are available as separate downloads. See the *IBM ioMemory VSL Release Notes* for more information.

- **SNMP Subagent**: The IBM SNMP AgentX subagent allows you to monitor and manage your IBM High IOPS Adapters using the Simple Network Management Protocol. You can use a normal SNMP browser, or customize your existing application to interface with the subagent.
- **SMI-S CIM Provider**: The CIM provider allows you to monitor and manage your devices using the Common Information Model. You can use a normal CIM browser, or customize your existing application to interface with the CIM provider.
- **ioMemory VSL Management SDK**: This C programing API allows you to write customize applications for monitoring and managing IBM High IOPS Adapters.

Example Conditions to Monitor

This section gives examples of conditions you can monitor. It is intended as an introduction and not as a comprehensive reference. These conditions will have slightly different names, states, and values, depending on the tool you choose. For example, an SNMP MIB may have a different name than a SMI-S object or an API function.

In order to properly monitor these conditions, you should become familiar with the tool you choose to implement and read the documentation for that tool. You may also discover additional conditions that you wish to frequently monitor.

For quick reference, the possible states/values of these conditions are described as Normal (**GREEN**), Caution/Alert (**YELLOW**), or Error/Warning (**RED**). You may implement your own ranges of acceptable states/values, especially if you use a data-source tool.

Device Status

All of the monitoring tools return information on the status of the IBM High IOPS Adapters, including the following states:

GREEN	Attached

YELLOW	Detached, Busy (including: Detaching, Attaching, Scanning, Formatting, and Updating)
RED	Minimal Mode, Powerloss Protect Disabled

If the device is in Minimal Mode, the monitoring tool can display the reason for the Minimal Mode status.

Required Actions

If the device is in Minimal Mode, the action will depend on the reason. For example, if the reason is outdated firmware, then you will need to update the firmware.

Temperature

IBM High IOPS Adapters require adequate cooling. In order to prevent thermal damage, the ioMemory VSL software will start throttling write performance once the on-board controller reaches a specified temperature. If the controller temperature continues to rise, the software will shut down the device once the controller temperature reaches the maximum operating temperature.

These temperatures depend on the device. Newer IBM High IOPS Adapters have higher thermal tolerances. Consult the *IBM High IOPS Hardware Installation Guide* to determine the thermal tolerances of all devices you will monitor. **This table uses the thermal tolerances for newer devices** (93°C throttling, 100°C shutdown).

GREEN	<93°C
YELLOW	93-99°C
RED	100°C

You may wish to shift the conditions by a few degrees so the YELLOW condition exists before throttling occurs. For example:

GREEN	<90°C
YELLOW	90-96°C
RED	97°C

Required Actions

If the temperature is at or approaching the **YELLOW** condition, thermal mitigation steps may be necessary. Evaluate the server environment and system requirements necessary to operate the High IOPS adapters. Server operating conditions are documented in the user guides for the server and the requirement to operate High IOPS adapter is at the following website, which may include updates to uEFI and IMM code levels:
http://www.ibm.com/support/entry/portal/docdisplay?lndocid=SERV-IOPS

Health Reserves Percentage

IBM High IOPS Adapters are highly fault-tolerant storage subsystem with many levels of protection against component failure and the loss nature of solid-state storage. As in all storage subsystems, component failures may occur.

By pro-actively monitoring device age and health, you can ensure reliable performance over the intended product life. The following table describes the Health Reserve conditions.

GREEN	>10%
YELLOW	0-10%
RED	0%

At the 10% healthy threshold, a one-time warning is issued. At 0%, the device is considered unhealthy. It enters *write-reduced* mode. After the 0% threshold, the device will soon enter *read-only* mode.

For complete information on Health Reserve conditions and their impact on performance, see *Appendix B- Monitoring the Health of Devices*.

Required Actions

The device needs close monitoring as it approaches 0% reserves and goes into write-reduced mode, which will result in reduced write performance. Prepare to replace the device soon.

Write (Health Reserves) Status

In correlation with the Health Reserves Percentage, the management tools will return write states similar to these:

GREEN	Device is healthy
YELLOW	Device is getting close to entering reduced write mode.
RED	Device has entered reduced-write or read-only mode to preserve the flash from further wearout.

Required Actions

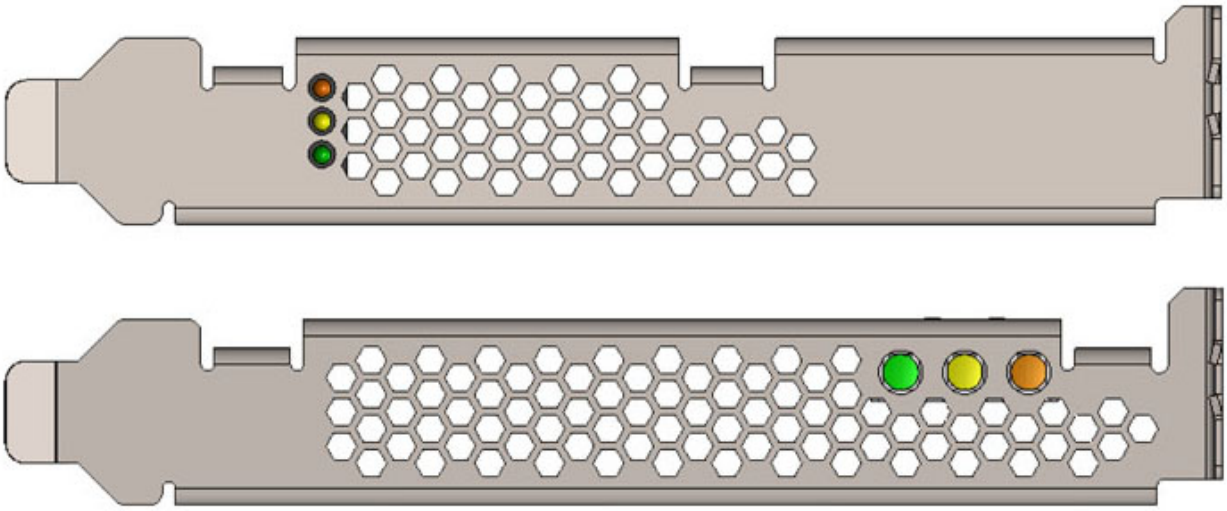
The device needs close monitoring as it approaches 0% reserves and goes into write-reduced mode, which will result in reduced write performance. Prepare to replace the device soon.

Device LED Indicators

If you have physical access to the devices, you can use the LED indicators to monitor their status.

Each IBM High IOPS Adapter includes three LEDs showing drive activity or error conditions. The LEDs on your

device should be similar to one of these configurations:



This table explains the information that these LEDs convey:

Green	Yellow	Amber	Indicates	Notes
O	O	O	Power off	
O	O	●	Power on. Problem with device, or driver not loaded (and device unattached)	Use fio-status to view problem, or load driver (and attach device)
●	O	O	Power on. Driver loaded (device may not be attached)	You may need to attach the device
●	(Flashing) ●	O	Writing (Rate indicates volume of writes)	Can appear in combination with the Read LED indication
(Flashing) ●	O	O	Reading (Rate indicates volume of reads)	Can appear in combination with the Write LED indication
●	●	●	Location Beacon	

Appendix A- Utilities Reference

The ioMemory VSL software installation packages include various command-line utilities, installed by default to `/usr/bin`. These provide a number of useful ways to access, test, and manipulate your device.

Attention There are some additional utilities installed in the `/usr/bin` directory that are not listed below. Those additional utilities are dependencies (used by the main ioMemory VSL utilities), and you should not use them directly unless Customer Support advises you to do so.

Utility	Purpose
`fio-attach`	Makes an IBM High IOPS Adapter available to the OS
`fio-beacon`	Lights the IBM High IOPS Adapter's external LEDs
`fio-bugreport`	Prepares a detailed report for use in troubleshooting problems
`fio-detach`	Temporarily removes an IBM High IOPS Adapter from OS access
`fio-format`	Used to perform a low-level format of an IBM High IOPS Adapter
`fio-pci-check`	Checks for errors on the PCI bus tree, specifically for IBM High IOPS Adapters
`fio-status`	Displays information about the device
`fio-sure-erase`	Clears or purges data from the device
`fio-update-iodrive`	Updates the IBM High IOPS Adapter's firmware

NOTE There are `-h` (Help) and `-v` (Version) options for all of the utilities. Also, `-h` and `-v` cause the utility to exit after displaying the information.

fio-attach

Description

Attaches the IBM High IOPS Adapter and makes it available to the operating system. This creates a block device in `/dev` named `fiox` (where `x` is `a`, `b`, `c`, etc.). You can then partition or format the IBM High IOPS Adapter, or set it up as part of a RAID array. The command displays a progress bar and percentage as it operates.

NOTE In most cases, the ioMemory VSL software automatically attaches the device on load and does a scan. You only need to run fio-attach if you ran fio-detach or if you set the ioMemory VSL software's auto_attach parameter to 0.

NOTE If the IBM High IOPS Adapter is in minimal mode, then auto-attach is disabled until the cause of the device being in minimal mode is fixed.

Syntax

```
fio-attach <device> [options]
```

where `<device>` is the name of the device node (`/dev/fctx`), where *x* indicates the device number: 0, 1, 2, etc. For example, `/dev/fct0` indicates the first IBM High IOPS Adapter installed on the system.

You can specify multiple IBM High IOPS Adapters. For example, `/dev/fct1 /dev/fct2` indicates the second and third IBM High IOPS Adapters installed on the system.

Option	Description
`-r`	Force a metadata rescan. This may take an extended period of time, and is not normally required. Attention Only use this option when directed by Customer Support.
`-c`	Attach only if clean.
`-q`	Quiet: disables the display of the progress bar and percentage.

fio-bugreport

Description

Prepares a detailed report of the device for use in troubleshooting problems. The results are saved in the `/tmp` directory in the file that indicates the date and time the utility was run.

Example:

```
/tmp/fio-bugreport-20100121.173256-sdv9ko.tar.bz2
```

Syntax

```
fio-bugreport
```

Notes

This utility captures the current state of the device. When a performance or stability problem occurs with the device, run the `fio-bugreport` utility and contact Customer Support at *http://www.ibm.com/systems/support* for assistance in troubleshooting.

Sample Output

```
-bash-3.2# fio-bugreport
Collecting fio-status -a
Collecting fio-status
Collecting fio-pci-check
Collecting fio-pci-check -v
Collecting fio-read-lebmap /dev/fct0
Collecting fio-read-lebmap -x /dev/stdout/dev/fct0
Collecting fio-read-lebmap -t /dev/fct0
Collecting fio-get-erase-count/dev/fct0
Collecting fio-get-erase-count -b /dev/fct0
Collecting lspci
Collecting lspci -vvvvv
Collecting lspci -tv
Collecting messages file(s)
Collecting procfusion file(s)
Collecting lsmod
Collecting uname -a
Collecting hostname
Collecting sar -r
Collecting sar
Collecting sar -A
Collecting syslog file(s)
Collecting proc file(s)
Collecting procirq file(s)
Collecting dmidecode
Collecting rpm -qa iodrive*
Collecting find /lib/modules
Please attach the bugreport tar file
    /tmp/fio-bugreport-20090921.173256-sdv9ko.tar.bz2
  to your support case, including steps to reproduce the problem.
  If you do not have an open support case for this issue, please open a
support
  case with a problem description and then attach this file to your new
case.
```

For example, the filename for a bug report file named `/tmp/fiobugreport-20090921.173256-sdvk0.tar.bz2` indicates the following:

- Date (20090921)
- Time (173256, or 17:32:56)
- Misc. information (`sdv9ko.tar.bz2`)

fio-beacon

Description

Lights the IBM High IOPS Adapter's LEDs to locate the device. You should first detach the IBM High IOPS Adapter and then run `fio-beacon`.

Syntax

```
fio-beacon <device> [options]
```

where `<device>` is the name of the device node (`/dev/fctx`), where *x* indicates the card number: 0, 1, 2, etc. For example, `/dev/fct0` indicates the first IBM High IOPS Adapter installed on the system.

Options	Description
`-0`	Off: (Zero) Turns off the three LEDs
`-1`	On: Lights the three LEDs
`-p`	Prints the PCI bus ID of the device at `<device>` to standard output. Usage and error information may be written to standard output rather than to standard error.

fio-detach

Description

Detaches the IBM High IOPS Adapter and removes the corresponding `fctx` IBM High IOPS Adapter block device from the OS. The `fio-detach` utility waits until the device completes all read/write activity before executing the detach operation. By default, the command also displays a progress bar and percentage as it completes the detach.

Attention Before using this utility, ensure that the device you want to detach is **NOT** currently mounted and in use.

Syntax

```
fio-detach <device> [options]
```

where `<device>` is the name of the device node (`/dev/fctx`), where *x* indicates the card number: 0, 1, 2, etc. For example, `/dev/fct0` indicates the first IBM High IOPS Adapter installed on the system.

You can specify multiple IBM High IOPS Adapters. For example, `/dev/fct1 /dev/fct2` indicates the second and third IBM High IOPS Adapters installed on the system. You can also use a wildcard to indicate all IBM High IOPS Adapters on the system. For example, `/dev/fct*`

Options	Description

`-q`	Quiet: Disables the display of the progress bar and percentage.

Notes

With this version of ioMemory VSL software, attempting to detach an IBM High IOPS Adapter may fail with an error indicating that the device is busy. This typically may occur if the IBM High IOPS Adapter is part of a software RAID (0,1,5) volume, is mounted, or some process has the device open.

The tools fuser, mount, and lsof can be helpful in determining what is holding the device open.

fio-format

Description

NOTE IBM High IOPS Adapters ship pre-formatted, so `fio-format` is generally not required except to change the logical size or block size of a device, or to erase user data on a device. To ensure the user data is truly erased, use *`fio-sure-erase`*.

Performs a low-level format of the IBM High IOPS Adapter. By default, `fio-format` displays a progress-percentage indicator as it runs.

Attention Use this utility with care, as it deletes all user information on the device.

NOTE Using a larger block (sector) size, such as 4096 bytes, can significantly reduce worst-case ioMemory VSL host memory consumption. However, some applications are not compatible with non-512-byte sector sizes.

NOTE If you do not include the `-s` or `-o` options, the device size defaults to the advertised capacity. If used, the `-s` and `-o` options must include the size or percentage indicators.

Attention Do not interrupt the formatting! We recommend adding power backup to your system to prevent power failures during formatting. If formatting is interrupted, please contact Customer Support.

Syntax

```
fio-format [options] <device>
```

where `<device>` is the name of the device node (`/dev/fctx`), where *x* indicates the device number: 0, 1, 2, etc. For example, `/dev/fct0` indicates the first IBM High IOPS Adapter installed on the system.

Options	Description
`-b <size B|K>`	Set the block (sector) size, in bytes or KiBytes (base 2). The default is 512 bytes. For example: `-b 512B` or `-b 4K` (`B` in `512B` is optional).
`-f`	Force the format size, bypassing normal checks and warnings. This option may be needed in rare situations when `fio-format` does not proceed properly. (The "Are you sure?" prompt still appears unless you use the `-y` option.)
`-q`	Quiet mode: Disable the display of the progress-percentage indicator.

`-s <size M|G|T|%>`	Set the device capacity as a specific size (in TB, GB, or MB) or as a percentage of the advertised capacity: • `T` Number of terabytes (TB) to format • `G` Number of gigabytes (GB) to format • `M` Number of megabytes (MB) to format • `%` Percentage, such as 70% (the percent sign must be included).
`-o <size B|K|M|G|T|%>`	Over-format the device size (to greater than the advertised capacity), where the maximum size equals the maximum physical capacity. If a percentage is used, it corresponds to the maximum physical capacity of the device. (Size is required for the `-o` option; see the `-s` option above for size indicator descriptions.) Attention Before you use this option, please discuss your use case with Customer Support.
`-R`	Disable fast rescan on unclean shutdown to reclaim some reserve capacity.
`-y`	Auto-answer "yes" to all queries from the application (bypass prompts).

fio-pci-check

Description

Checks for errors on the PCI bus tree, specifically for IBM High IOPS Adapters. This utility displays the current status of each IBM High IOPS Adapter. It also prints the standard PCI Express error information and resets the state.

NOTE It is perfectly normal to see a few errors (perhaps as many as five) when `fio-pci-check` is initially run. Subsequent runs should reveal only one or two errors during several hours of operation.

Syntax

```
fio-pci-check [options]
```

Options	Description
`-d <value>`	1 = Disable the link; 0 = bring the link up (Not recommended)
`-e`	Enable PCI-e error reporting.
`-f`	Scan every device in the system.
`-i`	Print the device serial number. This option is invalid when the ioMemory VSL is loaded.
`-r`	Force the link to retrain.
`-v`	Verbose: Print extra data about the hardware.

fio-status

Description

Provides detailed information about the installed devices. This utility operates on either `fctx` or `fiox` devices. The utility depends on running as root and having the ioMemory VSL driver loaded. If no driver is loaded, a smaller set of status information is returned.

`fio-status` provides alerts for certain error modes, such as a minimal-mode, read-only mode, and write-reduced mode, describing what is causing the condition.

Syntax

```
fio-status [<device>] [<options>]
```

where `<device>` is the name of the device node (`/dev/fctx`), where *x* indicates the card number: 0, 1, 2, etc. For example, `/dev/fct0` indicates the first IBM High IOPS Adapter installed on the system.

If `<dev>` is not specified, `fio-status` displays information for all cards in the system. If the ioMemory VSL driver is not loaded, this parameter is ignored.

Options	Description
`-a`	Report all available information for each device.
`-e`	Show all errors and warnings for each device. This option is for diagnosing issues, and it hides other information such as format sizes.
`-c`	Count: Report only the number of IBM High IOPS Adapters installed.
`-d`	Show basic information set plus the total amount of data read and written (lifetime data volumes). This option is not necessary when the -a option is used.
`-fj`	Format JSON: creates the output in JSON format.
`-fx`	Format XML: creates the output in XML format.
`-u`	Show unavailable fields. Only valid with `-fj` or `-fx`.
`-U`	Show unavailable fields and details why. Only valid with `-fj` or `-fx`. NOTE Some `fio-status` fields are unavailable depending on the operating system or device. For example, some legacy fields are unavailable on newer IBM High IOPS Adapters.
`-F<field>`	Print the value for a single field (see the next option for field names). Requires that a device be specified. Multiple `-F` options may be specified.
`-l`	List the fields that can be individually accessed with `-F`.

Attention **Output Change**: The standard formatting of `fio-status` ouput has changed compared to the output from ioMemory VSL software version 2.x. This will affect any custom management tools that used the

output of this utility.

Basic Information: If no options are used, `fio-status` reports the following basic information:

- Number and type of devices installed in the system
- ioMemory VSL software version

Adapter information:

- Adapter type
- Product number
- External power status
- PCIe power limit threshold (if available)
- Connected IBM High IOPS Adapters

Block device information:

- Attach status
- Product name
- Product number
- Serial number
- PCIe address and slot
- Firmware version
- Size of the device, out of total capacity
- Internal temperature (average and maximum, since ioMemory VSL software load) in degrees Centigrade
- Health status: healthy, nearing wearout, write-reduced or read-only
- Reserve capacity (percentage)
- Warning capacity threshold (percentage)

Data Volume Information: If the -d option is used, the following data volume information is reported *in addition* to the basic information:

- Physical bytes written
- Physical bytes read

All Information: If the `-a` option is used, all information is printed, which includes the following information *in addition* to basic and data volume information:

Adapter information:

- Manufacturer number
- Part number
- Date of manufacture
- Power loss protection status
- PCIe bus voltage (avg, min, max)
- PCIe bus current (avg, max)
- PCIe bus power (avg, max)
- PCIe power limit threshold (watts)
- PCIe slot available power (watts)
- PCIe negotiated link information (lanes and throughput)

Block device information:

- Manufacturer's code
- Manufacturing date
- Vendor and sub-vendor information
- Format status and sector information (if device is attached)
- FPGA ID and Low-level format GUID
- PCIe slot available power
- PCIe negotiated link information
- Card temperature, in degrees Centigrade
- Internal voltage (avg and max)
- Auxiliary voltage (avg and max)
- Percentage of good blocks, data and metadata
- Lifetime data volume statistics
- RAM usage

Error Mode Information: If the ioMemory VSL software is in minimal mode, read-only mode, or write-reduced mode when `fio-status` is run, the following differences occur in the output:

- Attach status is "Status unknown: Driver is in MINIMAL MODE:"

- The reason for the minimal mode state is displayed (such as "Firmware is out of date. Update firmware.")
- "Geometry and capacity information not available." is displayed.
- No media health information is displayed.

fio-sure-erase

Attention As a best practice, do not use this utility if there are any IBM High IOPS Adapters installed in the system that you do not want to clear or purge. First remove any devices that you do not want to accidentally erase. Once the data is removed with this utility it is gone forever. **It is not recoverable.**

Attention Before you use this utility, be sure to back up any data that you wish to preserve.

NOTE After using `fio-sure-erase`, format the device using *`fio-format`* before using the device again.

Attention If the device is in Read-only mode, perform a format using `fio-format` before running `fio-sure-erase`. If the device is in Minimal mode, then `fio-sure-erase` cannot erase the device. Updating the firmware may take the device out of Minimal Mode. If the device remains in Minimal mode, contact Customer Support at *http://www.ibm.com/systems/support* for further assistance.

In order to run `fio-sure-erase`, the block device **must be detached**. See the *`fio-detach`* section for more information.

Description

The `fio-sure-erase` is a command-line utility that securely removes data from IBM High IOPS Adapters. It complies with the "Clear" and "Purge" level of destruction from the following standards:

1. DOD 5220.22-M - Comply with instructions for Flash EPROM
2. NIST SP800-88- Comply with instructions for Flash EPROM

See below for more information on Clear and Purge support.

Syntax

```
fio-sure-erase [options] <device>
```

Where `<device>` is the name of the device node (`/dev/fctx`), where *x* indicates the card number: 0, 1, 2, etc. For example, `/dev/fct0` indicates the first IBM High IOPS Adapter installed on the system. Use *`fio-status`* to view this device node.

NOTE **Products with Multiple Devices**: `fio-sure-erase` works on individual IBM High IOPS Adapters. For example, if you are planning to purge an IBM High IOPS Duo Adapter, you will need to perform this operation on each of the product's two IBM High IOPS Adapters.

Options	Description

`-p`	Purge instead of Clear: performs a write followed by an erase. For more information on Purge, see below. Attention Purging the device may take hours to accomplish, depending on the size of the device that needs to be purged.
`-y`	No confirmation: do not require a yes/no response to execute the utility.
`-t`	Do not preserve current format parameters, including device and sector size (reset to default).
`-q`	Quiet: do not display the status bar.

NOTE If you run `fio-sure-erase` with no options, a Clear is performed. For more information, see below.

Each block of memory consists of uniform 1 bits or 0 bits.

Clear Support

A "Clear" is the default state of running `fio-sure-erase` (with no options), and refers to the act of performing a full low-level erase (every cell pushed to "1") of the entire NAND media, including retired erase blocks.

Metadata that is required for operation will not be destroyed (media event log, erase counts, physical bytes read/written, performance and thermal history), but any user-specific metadata will be destroyed.

The following describes the steps taken in the Clear operation:

1. Creates a unity map of every addressable block (this allows `fio-sure-erase` to address every block, including previously unmapped bad blocks).
2. For each block, performs an erase cycle (every cell is pushed to "1").
3. Restores the bad block map.
4. Formats the device (the purpose of this is to make the device usable again, the utility erases all of the headers during the clear).

Purge Support

A "Purge" is implemented by using the `-p` option with `fio-sure-erase`. Purge refers to the act of first overwriting the entire NAND media (including retired erase blocks) with a single character (every cell written to logical "0"), and then performing a full chip erase (every cell pushed to "1") across all media (including retired erase blocks).

Metadata that is required for operation will **not** be destroyed (media event log, erase counts, physical bytes read/written, performance and thermal history), but any user-specific metadata will be destroyed.

The following describes the steps taken in the Purge operation:

1. Creates a unity map of every addressable block (this allows `fio-sure-erase` to address every block, including previously unmapped bad blocks).
2. For each block, performs a write cycle (every cell written to "0").

3. For each block, performs an erase cycle (every cell pushed to "1").
4. Restores the bad block map.
5. Formats the drive (the purpose of this is to make the drive usable again, the utility erases all of the headers during the clear).

fio-update-iodrive

Attention You should back up the data on the IBM High IOPS Adapter prior to any upgrade as a precaution.

Description

Updates the IBM High IOPS Adapter's firmware. This utility scans the PCIe bus for all IBM High IOPS Adapters and updates them. A progress bar and percentage are shown for each device as the update completes.

Attention It is extremely important that the power not be turned off during a firmware upgrade, as this could cause device failure. If a UPS is not already in place, consider adding one to the system prior to performing a firmware upgrade.

Attention Note that when running multiple firmware upgrades in sequence, it is critical to load the ioMemory VSL driver after each firmware upgrade step. Otherwise the on-device format will not be changed, and there will be data loss.

Attention Do not use this utility to downgrade the IBM High IOPS Adapter to an earlier version of the firmware. Doing so may result in data loss and void your warranty.

Attention The default action (without using the `-d` or `-s` option) is to upgrade all IBM High IOPS Adapters with the firmware contained in the `fusion_<version>-<date>.fff` firmware archive file. Confirm that all devices need the upgrade prior to running the update. If in doubt, use the `-p` (Pretend) option to view the possible results of the update.

Attention You must detach all IBM High IOPS Adapters before updating the firmware.

Attention **Upgrade Path**: There is a specific upgrade path that you must take when upgrading IBM High IOPS Adapter. Consult the *IBM ioMemory VSL Release Notes* for this ioMemory VSL software release before upgrading IBM High IOPS Adapters.

NOTE If you receive an error message when updating the firmware that instructs you to update the midprom information, contact Customer Support.

To update one or more specific devices:

- If the ioMemory VSL driver is loaded, use the `-d` option with the device number.

Syntax

```
fio-update-iodrive [options] <firmware-path>
```

where `<firmaware-path>` is the full path to the firmware archive file (`highiops_<version>-<date>.fff`) available at *http://www.ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-65723* (follow that link and then select **IBM**

High IOPS software matrix). If you downloaded the `.fff` firmware archive file, then the firmware is wherever you stored it. If you installed the firmware from the firmware package, the default path is `/usr/share/fio/firmware/`. This parameter is required.

Options	Description
`-d`	Updates the specified devices (by `fctx`, where *x* is the number of the device shown in `fio-status`). If this option is not specified, all devices are updated. Attention Use the `-d` or `-s` options with care, as updating the wrong IBM High IOPS Adapter could damage your device.
`-f`	Force upgrade (used primarily to downgrade to an earlier firmware version). If the ioMemory VSL driver is not loaded, this option also requires the `-s` option. Attention Use the -f option with care, as it could damage your card.
`-l`	List the firmware available in the archive.
`-p`	Pretend: Shows what updates would be done. However, the actual firmware is not modified.
`-c`	Clears locks placed on a device.
`-q`	Runs the update process without displaying the progress bar or percentage.
`-y`	Confirm all warning messages.
`-s`	Updates the devices in the specified slots using '*' as a wildcard for devices. The slots are identified in the following PCIe format (as shown in lspci): `[[[[<domain>]:]<bus>]:][<slot>][.[<func>]]`
`--split`	Split the IBM High IOPS Adapter into virtual devices.
`--merge`	Merge the virtual devices of an IBM High IOPS Adapter.

Appendix B- Monitoring the Health of Devices

This section describes how the health of IBM High IOPS Adapters can be measured and monitored in order to safeguard data and prolong device lifetime.

NAND Flash and Component Failure

An IBM High IOPS Adapter is a highly fault-tolerant storage subsystem that provides many levels of protection against component failure and the loss nature of solid-state storage. As in all storage subsystems, component failures may occur.

By pro-actively monitoring device age and health, you can ensure reliable performance over the intended product life.

Health Metrics

The ioMemory VSL software manages block retirement using pre-determined retirement thresholds. The IBM High IOPS Management Application and the `fio-status` utilities show a health indicator that starts at 100 and counts down to 0. As certain thresholds are crossed, various actions are taken.

At the 10% healthy threshold, a one-time warning is issued. See the *Health Monitoring Techniques* section below for methods for capturing this alarm event.

At 0%, the device is considered unhealthy. It enters *write-reduced* mode, which somewhat prolongs its lifespan so data can be safely migrated off. In this state the IBM High IOPS Adapter behaves normally, except for the reduced write performance.

After the 0% threshold, the device will soon enter *read-only* mode – any attempt to write to the IBM High IOPS Adapter causes an error. Some filesystems may require special mount options in order to mount a read-only block device in addition to specifying that the mount should be read-only.

For example, under Linux, `ext3` requires that `"-o ro,noload"` is used. The `"noload"` option tells the filesystem to not try and replay the journal.

Read-only mode should be considered a final opportunity to migrate data off the device, as device failure is more likely with continued use.

The IBM High IOPS Adapter may enter failure mode. In this case, the device is offline and inaccessible. This can be caused by an internal catastrophic failure, improper firmware upgrade procedures, or device wearout.

NOTE For service or warranty-related questions, contact the company form which you purchased the device.

NOTE For products with multiple IBM High IOPS Adapters, these modes are maintained independently for each device.

Health Monitoring Techniques

`fio-status`: Output from the `fio-status` utility shows the health percentage and device state. These items are referenced as `"Media status"` in the sample output below.

```
Found 3 ioMemory devices in this system
Fusion-io driver version: 3.0.6 build 364

Adapter: Single Adapter
        Fusion-io ioDrive 1.30TB, Product Number:F00-001-1T30-CS-0001,
SN:1133D0248, FIO SN:1134D9565
...
Media status: Healthy; Reserves: 100.00%, warn at 10.00%; Data: 99.12%
Lifetime data volumes:
   Physical bytes written: 6,423,563,326,064
   Physical bytes read   : 5,509,006,756,312
```

IBM High IOPS Management Application: In the Device Report tab, look for the Reserve Space percentage in the right column. The higher the percentage, the healthier the drive is likely to be.

The following Health Status messages are produced by the `fio-status` utility:

- Healthy
- Read-only
- Reduced-write
- Unknown

Software RAID and Health Monitoring

Software RAID stacks are typically designed to detect and mitigate the failure modes of traditional storage media. The IBM High IOPS Adapter attempts to fail as gracefully as possible, and these new failure mechanisms are compatible with existing software RAID stacks. An IBM High IOPS Adapter in a RAID group will fail to receive data at a sufficient rate if a) the device is in a write-reduced state, and b) it is participating in a write-heavy workload. In this case, the device will be evicted from the RAID group. A device in read-only mode will be evicted when write I/Os are returned from the device as failed. Catastrophic failures are detected and handled just as though they are on traditional storage devices.

Appendix C- Using Module Parameters

The following table describes the module parameters you can set by editing the `/etc/modprobe.d/iomemory-vsl.conf` file and changing their values.

Each module parameter in the configuration file must be preceded by `options iomemory-vsl`. The `/etc/modprobe.d/iomemory-vsl.conf` file has some example parameters that are commented out. You may use these examples as templates and/or uncomment them in order to use them.

NOTE These changes must be completed before the ioMemory VSL software is loaded in order to take effect.

Module Parameter	Default (min/max)	Description
`auto_attach`	1	`1` = Always attach the device on driver load. `0` = Don't attach the device on driver load.
`fio_dev_wait_timeout_secs`	30	Number of seconds to wait for `/dev/fio*` files to show up during driver load. For systems not using `udev`, this should be set to `0` to disable the timeout and avoid an unneeded pause during driver load.
`force_minimal_mode`	0	`1` = Force minimal mode on the device. `0` = Do not force minimal mode on the device.
`numa_node_override`	Nothing Selected	A list of `<affinity specification>` couplets that specify the affinity settings of all devices in the system. Each item in the couplet is separated by a colon, and each couplet set is separated by a comma. Where each `<affinity specification>` couplet has the following syntax: `<device-id>=<node-number>` See *Appendix E- NUMA Configuration* for more information on using this parameter.
`parallel_attach`	1	`1` = Enable parallel attach of multiple devices. `0` = Disable parallel attach of multiple devices.
`preallocate_memory`	No devices selected	For the selected devices, pre-allocate all memory necessary to have the drive usable as swap space. Where the `<value>` for this parameter is a comma-separated list of device serial numbers.

`tintr_hw_wait`	0 (0, 255)	Interval (microseconds) to wait between hardware interrupts. Also known as interrupt coalescing. `0` is off.
`use_workqueue`	0 (0 or 3)	Linux only: `3` = use standard OS I/O elevators; `0` = bypass.

NOTE Other than `preallocate_memory`, module parameters are global — they apply to all IBM devices in the computer.

Appendix F- Upgrading Devices from VSL 2.x to 3.x

This version of the ioMemory VSL software supports new features, including the latest generation of High IOPS architecture and improved Flashback protection. These features require the latest version of the IBM firmware. Every IBM High IOPS Adapter in a system running 3.1.x or later must be upgraded to the latest version of the firmware.

For example, if you have a system running 2.x ioMemory VSL software with IBM High IOPS Adapters previously installed, and you want to install new IBM High IOPS Adapters (that require the latest version of the firmware), then you will need to upgrade all of the existing devices to the latest firmware version.

Attention You cannot revert a device's firmware to an earlier version once you have upgraded the device (without voiding your warranty). If you experience problems with your upgrade, please contact Customer Support at *http://www.ibm.com/systems/support*.

Attention Upgrading devices (previously configured for VSL 2.x.x) to work with VSL 3.x.x will require a low-level media format of the device. No user data will be maintained during the process. Be sure to backup all data as instructed.

Attention **Upgrade Path**: Depending on the current firmware version of your devices, you may need to upgrade your device's firmware multiple times in order to preserve internal structures. Consult the ioMemory VSL software for the upgrade path. Visit *http://www.ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-65723* (follow that link and then select **IBM High IOPS software matrix**) for all of the required software and firmware versions.

For more information on upgrading from one version to the next, see the *IBM ioMemory VSL Release Notes* (available at *http://www.ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-65723* (follow that link and then select **IBM High IOPS software matrix**)) for the version you will upgrade the device to. Then follow the upgrade instructions in that version's user guide for your operating system (including the firmware update instructions).

Upgrade Procedure

Be sure to follow the upgrade path in the *IBM ioMemory VSL Release Notes*. Make sure that all previously installed IBM High IOPS Adapters are updated with the appropriate firmware.

⚠ If you plan to use IBM High IOPS Adapters and IBM High IOPS Adapters in the same host, perform this upgrade on all existing IBM High IOPS Adapters **before** installing the new IBM High IOPS Adapters.

1. Prepare each existing IBM High IOPS Adapter for upgrade.
    a. Backup user data on each device.

> The upgrade process will require a low-level media format of the device. No user data will be maintained during the process; be sure to make a complete backup.

Use a backup method of your choice. For best results, use software and backup devices that have proven effective in the past. Do not backup the data onto another IBM High IOPS Adapter on the same system. The back up must be to a local disk or to an externally attached volume.

    b. Run the *fio-bugreport* utility and save the output. This will capture the device information for each device in the system. This device information will be useful in troubleshooting any upgrade issues. Sample command:

```
fio-bugreport
```

    c. Detach IBM High IOPS Adapters, for example:

```
fio-detach /dev/fct*
```

For more information, see *fio-detach*

2. Unload the current ioMemory VSL driver, for example:

```
$ modprobe -r iomemory-vsl
```

3. Uninstall the 2.x ioMemory VSL software.
    a. To uninstall the software, you must specify the kernel version of the package you are uninstalling. Run the appropriate command to find the installed packages:
        - RPM command:

```
$  rpm -qa | grep -i iomemory
```

        Sample output:

```
iomemory-vsl-2.6.18-194.el5-2.2.2.82-1.0
```

    b. Uninstall the ioMemory VSL software by running a command similar to this example (specify the kernel version of the package you wish to uninstall):
        - Sample RPM command:

```
$ rpm -e iomemory-vsl-2.6.18-194.el5-2.2.0.82-1.0
```

    c. Uninstall the utilties:
        - Sample RPM command:

```
$ rpm -e fio-util fio-snmp-agentx fio-common fio-firmware
iomanager-gui iomanager-jre libfio libfio-doc libfusionjni
fio-sysvinit fio-smis fio-snmp-mib libfio-dev
```

4. Install the new ioMemory VSL software and related packages.
    a. Download the ioMemory VSL software binary package for your kernel and all supporting packages at *http://www.ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-65723* (follow that link and then select **IBM High IOPS software matrix**)

> If you don't see a binary for your kernel, follow the instructions in the **Building the ioMemory VSL from Source** section of *Installing RPM Packages*. To see your current kernel version, run:
>
> ```
> uname -r
> ```

    b. Install the ioMemory VSL software and utilities using the appropriate commands:
        - RPM commands:

```
rpm -Uvh
iomemory-vsl-<kernel-version>-<VSL-version>.x86_64.rpm
rpm -Uvh lib*.rpm
rpm -Uvh fio*.rpm
```

See *Installing RPM Packages* for full instructions on installing those packages.

    c. Reboot the system.

5. Update the firmware on each device to the latest version using `fio-update-iodrive`.

> **Prevent Power Loss**
> Take measures to prevent power loss during the update, such as a UPS. Power loss during an update may result in device failure. For all warnings, alerts, and options pertaining to this utility, see the *`fio-update-iodrive`* utility reference in the appendix.

Sample syntax:

```
fio-update-iodrive <firmware-path>
```

Where `<firmare-path>` is the full path to the firmware archive file (`highiops_<version>-<date>.fff`) available at *http://www.ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-65723* (follow that link and then select **IBM High IOPS software matrix**). This command will update all of the devices to the selected firmware. If you wish to update specific devices, consult the *utility reference* for more options.

6. Reboot the system

7. Load the ioMemory VSL software, for example:

```
$ modprobe iomemory-vsl
```

For more information, see *Loading the ioMemory VSL Driver*

If run, `fio-status` will warn that the upgraded devices are missing a lebmap. This is expected, and will be fixed in the next step.

Destructive Step
Running `fio-format` in the next step will erase the entire device, including user data. Once this format is started, the device cannot be downgraded to the 2.x driver without voiding your warranty. If you experience problems with your upgrade, please contact Customer Support at *http://www.ibm.com/systems/support*.

8. Format each device using `fio-format`, for example:

```
fio-format <device>
```

You will be prompted to confirm you wish to erase all data on the device.

The format may take an extended period of time, depending on the wear on the device.

9. Attach all IBM High IOPS Adapters, for example:

```
fio-attach /dev/fct*
```

10. Check the status of all devices using fio-status, for example:

```
fio-status -a
```

Your IBM High IOPS Adapters have now been successfully upgraded for this version of the ioMemory VSL software. You may now install any IBM High IOPS Adapters.

Appendix E- NUMA Configuration

About NUMA Architecture

Servers with a NUMA (Non-Uniform Memory Access) architecture may require special installation instructions in order to maximize IBM High IOPS Adapter performance. This includes most multi-socket servers.

On some servers with NUMA architecture, during system boot, the BIOS will not associate PCIe slots with the correct NUMA node. Incorrect mappings result in inefficient I/O handling that can significantly degrade performance. To prevent this, you must manually assign IBM High IOPS Adapters optimally among the available NUMA nodes.

Attention The example below shows the final implementation of custom affinity settings. This implementation required an analysis of the specific system, including the system architecture, type and number of IBM High IOPS Adapters installed, and the particular PCIe slots that were used. Your particular circumstances will require a custom analysis of your set-up. This analysis requires understanding of your system's NUMA architecture compared to your particular installation.

Your actual settings may be different than the example below, depending on your server configuration. In order to create the correct settings for your specific system, use `fio-status` to list all of the devices (fct numbers). Next, use `fio-beacon` to identify each of the devices in their respective PCIe slots. Then use the example below of setting the `numa_node_override` parameter as a template and modify it for your particular system.

Configuring your IBM High IOPS Adapters for servers with NUMA architecture requires the use of the `numa_node_override` parameter by modifying the `iomemory-vsl.conf` file.

numa_node_override Parameter

The `numa_node_override` parameter is a list of `<affinity specification>` couplets that specify the affinity settings of all devices in the system. Each item in the couplet is separated by a colon, and each couplet set is separated by a comma.

Syntax:

```
numa_node_override=<affinity specification>[,<affinity specification>...]
```

Where each `<affinity specification>` has the following syntax:

```
<fct-number>:<node-number>
```

Simple Example:

```
numa_node_override=fct4:1,fct5:0,fct7:2,fct9:3
```

Has the effect of creating :

Device	Node/Group	Processor Affinity
fct4	node 1	all processors in node 1
fct5	node 0	all processors in node 0
fct7	node 2	all processors in node 2
fct9	node 3	all processors in node 3

Advanced Configuration Example

This sample server has 4 NUMA nodes with 8 hyper-threaded cores per node (16 logical processors per node, a total of 64 logical processors in the system). This system also uses the expansion configuration and has 11 PCIe expansion slots. During system boot, the system's BIOS will assign PCIe slots 1-6 to NUMA node 2 and PCIe slots 7-11 to NUMA node 0. NUMA nodes 1 and 3 will have no assigned PCIe slots. This creates a load balancing problem in the system when IBM High IOPS Adapters are under heavy traffic. Specifically, during these periods of high use, half of the CPUs in the system will sit idle while the other half of the CPUs are 100% utilized, thus limiting the throughput of the IBM High IOPS Adapters.

To avoid this problem, you must manually configure the affinity of the IBM High IOPS Adapters using the `numa_node_override` parameter to distribute the work load across all NUMA nodes. This parameter will override the default behavior of the ioMemory VSL software. For more information about the `numa_node_override` parameter, refer to the syntax explanation above.

What follows is an example of how to manually configure 10 IBM High IOPS Duo Adapters (each with two IBM High IOPS Adapters). Slot 1 is a Generation 1 PCI-e slot, so it is not compatible with an IBM High IOPS Duo Adapter. Thus we can fill slots 2-11 with IBM High IOPS Duo Adapters.

NOTE Because each IBM High IOPS Duo Adapter has two IBM High IOPS Adapters, there are two device numbers for each IBM High IOPS Duo Adapter (one for each IBM High IOPS Adapter). There will therefore be two device numbers for each slot.

When the system boots, the default BIOS NUMA node assignments are:

BIOS Assigned NUMA Node	PCI-e Slots	FCT device numbers	Processor Affinity
0	7-11	8,9,13,14,18,19,23,24,28,29	all processors in the node
1	none	none	none
2	2-6	135,136,140,141,145,146,150,151,155,156	all processors in the node

3	none	none	none

Here, the BIOS creates a load imbalance by assigning the cards to only two NUMA nodes in the system. In order to balance the work load, we want to make the following manual settings:

Assigned NUMA Node	PCI-e Slots	FCT device numbers	Processor Affinity
0	7-9	8,9,13,14,18,19	all processors in the node
1	10-11	23,24,28,29	all processors in the node
2	2-3	135,136,140,141	all processors in the node
3	4-6	145,146,150,151,155,156	all processors in the node

In order to configure the ioMemory VSL software with these override settings, set the `numa_node_override` parameter with the following string:

```
numa_node_override=fct8:0,fct9:0,fct13:0,fct14:0,fct18:0,fct19:0,
fct23:1,fct24:1,fct28:1,fct29:1,fct135:2,fct136:2,fct140:2,fct141:2,
fct145:3,fct146:3,fct150:3,fct151:3,fct155:3,fct156:3
```

Attention The above example contains line breaks for formatting purposes. There would be no line breaks in a real implementation of the `numa_node_override` parameter.

IBM Support

IBM High IOPS Adapter software and documentation are available on the web at the following address:

http://www.ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-65723 (follow that link and then select **IBM High IOPS software matrix**).

IBM part number 00D2423