Fast-fail Policy for LSM Mirrored Volume I/O

Overview

This paper provides information to Tru64 UNIX customers who want to take advantage of the new fast-fail policy for LSM mirrored volume I/O, which was introduced in Tru64 UNIX Version 5.1B-3.

This paper will help you to answer the following questions:

- How will this policy help me?
- What changes were made to Tru64 UNIX to support this policy?
- When do I use this policy and what steps do I need to take?

The Glossary section provides a brief overview of the terminology used. It is assumed that readers have some knowledge of LSM and RAID levels; however basics will be covered as appropriate. The Resources section includes pointers to additional documentation.

Note that in this paper, the terms device, disk, and LUN are used interchangeably. In all cases, they either refer to one physical disk or a portion of a hardware RAID array, both of which are known as logical units.

Problem statement

The Logical Storage Manager (LSM) is a host-based storage management utility for the HP Tru64 UNIX operating system. This powerful software tool allows administrators to create virtual disks, called volumes, on top of UNIX physical disks in whichever supported RAID level is desired, be it RAID0 (striping), RAID1 (mirroring), RAID5, or a combination of mirroring and striping.

In the creation of volumes, LSM uses the LUN presented to Tru64 UNIX just as any other application or filesystem would. As such, it treats all LUNs the same, with details regarding the underlying storage configuration hidden from LSM.

Figure 1 illustrates an LSM volume called mirvol1. Volume mirvol1 has 2 mirrors, one of which uses LUN dsk1 for its storage while the other uses LUN dsk22 for its storage. Both dsk1 and dsk22 contain the same application data because mirrored volume writes must complete to both mirrors.

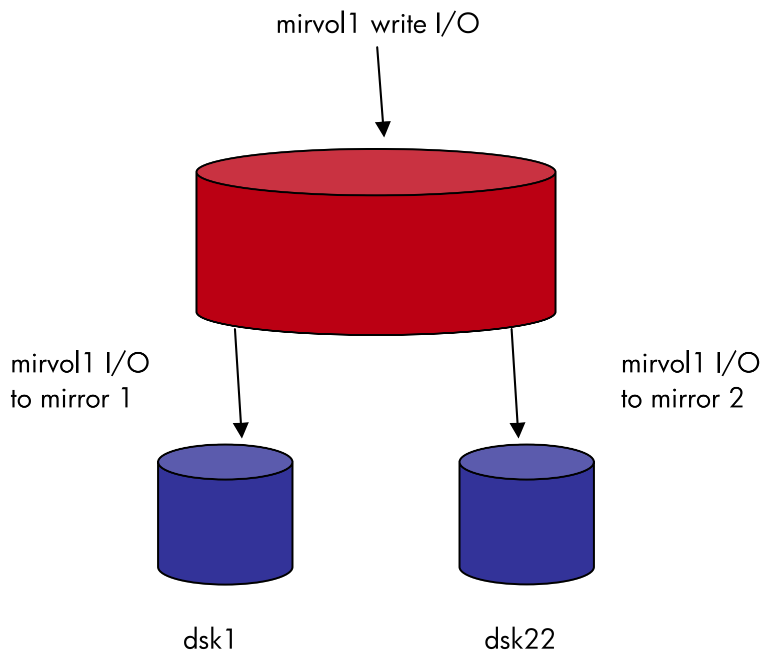

Figure 1

The I/O request has no knowledge of the application or the underlying storage devices. The Tru64 UNIX I/O subsystem is a layered architecture. The upper layers make requests to the underlying layers, which attempt to execute the request and return a response. The response can contain data and status. Each layer does not know about the implementation of the other layers.

Figure 2 illustrates the Tru64 UNIX mass storage I/O subsystem.

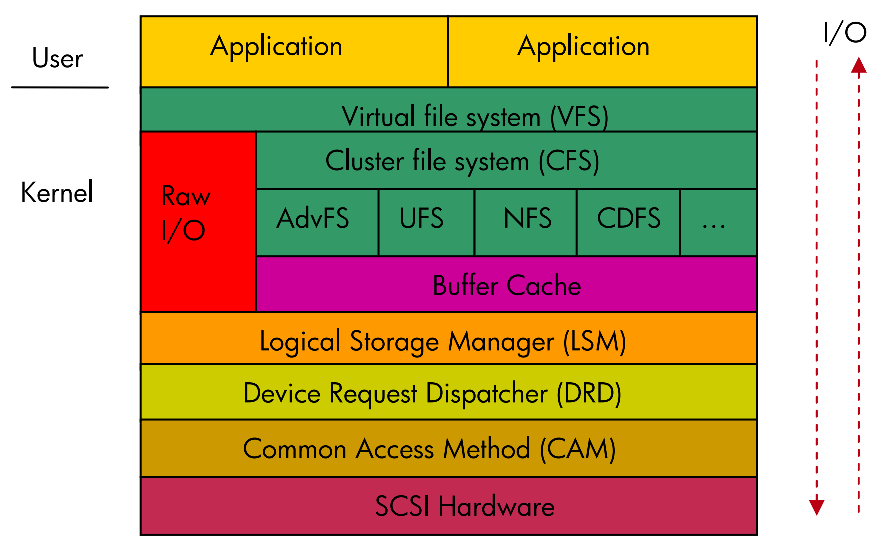

Figure 2

As illustrated, applications in user space issue I/O to the underlying layers. The Cluster File System (CFS) and Device Request Dispatcher (DRD) layers are available in TruCluster environments only. The other layers exist in both single-system and TruCluster environments. LSM is an optional layer in either environment.

With respect to disk errors, the Tru64 UNIX I/O subsystem is designed to avoid returning an error on I/O requests. As such, failed I/O requests may be retried numerous times.

Retries are incorporated because in some cases, if given enough time, the hardware can correct a disk error using its own error correction mechanisms, such as trying to restore data on a different disk location. Alternatively, the I/O could have failed due to an intermittent error, such as from a temporary bus voltage problem. Furthermore, the device could be a badly behaving device that is returning misleading errors. As the I/O subsystem could get the same error from hardware in all of these cases, retries are used to account for the range of possibilities. As these problems bear investigation, errors and retries are logged for administrator use.

From an application perspective, retries can result in long delays in which the application appears hung as I/O errors are logged and retries are logged and serviced.

The underlying CAM layer may retry numerous times to avoid failing the I/O request. In a TruCluster environment, the DRD layer may retry as well, searching for another cluster member to access the device without error. The hardware subsystem may also have retried a number of times before it passed the error to CAM.

For improved performance, if a faulty device is used in an LSM mirrored volume, administrators expect that LSM will fail that mirror and route the I/O to another mirror immediately. This allows the hardware problem to be resolved without impacting mirrored volume I/O. However, LSM depends on the lower layers to return an I/O status. The lower layers have no idea LSM is being used to provide redundancy. If they did, for some errors, retries could be limited as LSM has another device to service the request. This is the crux of the new fast-fail policy for LSM mirrored volume I/O.

The fast-fail policy provides a mechanism for device errors to be reported faster which allows LSM to route I/O to another mirror faster.

Changes made to support fast-fail

To support the fast-fail policy for LSM mirrored volume I/O, LSM introduced a new kernel tunable parameter, *lsm_fastfail_enable*, which can be set for all mirrored volume I/O. Setting this parameter to 1 enables the fast-fail policy for LSM mirrored volumes; setting it to 0 (the default value) disables the policy.

Administrators can enable or disable it at any time using the *sysconfig* command. See the Tru64 UNIX *sysconfig(8)* reference page for information about this command.

When fast-fail is enabled, LSM embeds a fast-fail signal in mirrored volume I/O requests for use by the underlying layers. If an I/O failure is encountered, code in the DRD (for TruClusters) and CAM layers of the Tru64 UNIX I/O subsystem limit retries accordingly.

Anytime an I/O request is failed as a result of the fast-fail policy, CAM includes the following error message in the binary error log, */var/adm/binary.errlog*.

    Fast Fail Requested, not retrying

This information is viewable using EVM, WEBES (Web-Based Enterprise Services) or SEA (System Event Analyzer, formally Compaq Analyzer) as preferred. See Resources for more information.

LSM ensures that fast-fail is disabled on I/O to mirrored volumes that contain only one enabled plex.

Furthermore, if fast-fail is enabled and a mirrored volume I/O request fails, LSM will retry that request with fast-fail disabled. This would occur if every mirror used in that volume reported an error for that request. LSM will retry once, reporting retries via a console message such as the following.

> lsm:volio: read error on volume mirvol1, retry with 'fastfail' off

This message is recorded in the */var/adm/messages* file, which is also viewable via the Tru64 UNIX Event Manager (EVM).

The CAM and DRD layers then try the I/O request again without fast-fail enabled. This avoids failing an application's I/O request unless absolutely necessary.

Otherwise, if enabled, fast-fail operates quietly. It adds minimal overhead.

Design details

Changes were made to the LSM, DRD and CAM layers of the Tru64 UNIX I/O subsystem to support fast-fail. The following sections describe how the fast-fail policy performs at these layers.

LSM

Plex types

The fast-fail policy is available for LSM mirrored (RAID1) volumes only. The mirror can be simple, concatenated, or striped (RAID0).

Simple, concatenated, or striped refers to how the data is spread across the mirrors of a mirrored volume. The following explains these layouts in more detail.

In LSM vernacular, a mirror is called a plex. Each plex is made up of one or more subdisks. A subdisk is defined as an offset and length onto a physical disk. Either an entire disk or a disk partition can be added to LSM. In either case, subdisks are created from this physical storage.

Simple plexes have one subdisk. Concatenated plexes store data across two or more subdisks with data beginning on one subdisk, continuing on the next in a linear fashion. Striped plexes store data across two or more subdisks using a specified stripe width to interleave I/O requests across multiple disks for improved performance. Both concatenated and striped plexes allow for virtual disks with a capacity greater than a single disk.

Figures 3, 4, and 5 illustrate these mirrored volume layouts.

Figure 3 illustrates a mirrored volume with a simple plex layout. The volume size shown is measured in blocks (512 bytes).

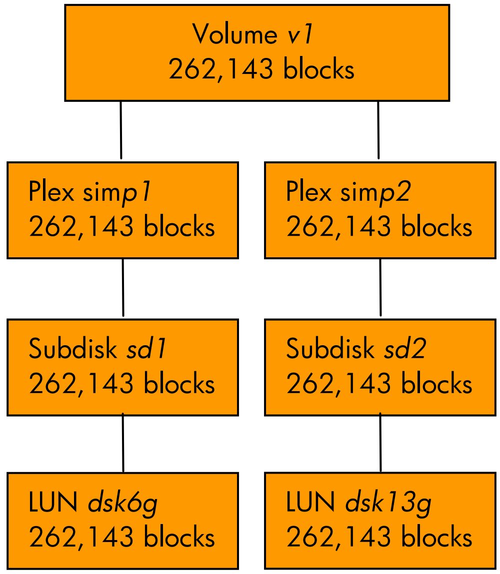

Figure 3

Figure 4 illustrates a mirrored volume with a concatenated plex layout across two LUNs. The volume size shown is measured in blocks (512 bytes). The volume and LUN size is small for illustrative purposes.

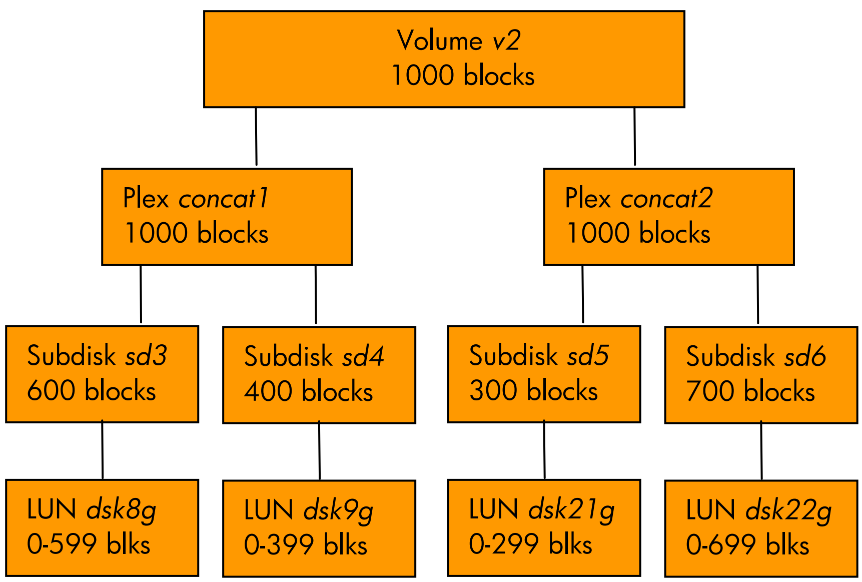

Figure 4

Figure 5 illustrates a mirrored volume with a striped layout across two LUNs. The stripe width used is the current default stripe width for striped plexes, 64K. The volume size shown is measured in blocks (512 bytes). The volume and LUN size is small for illustrative purposes.

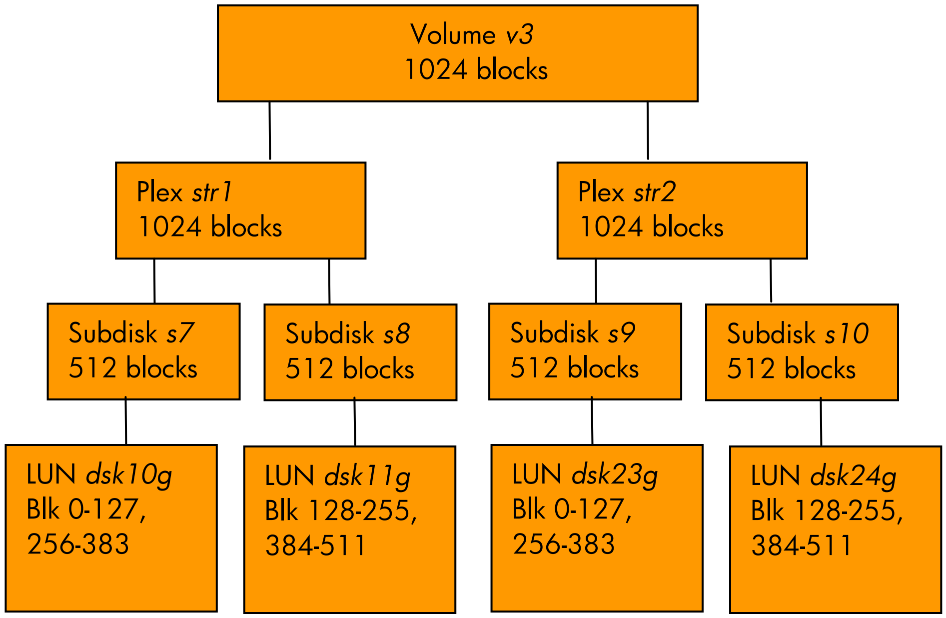

Figure 5

If there is only one plex, the plex layout still applies although there is no redundancy.

The fast-fail policy applies to all three plex layouts for mirrored volumes.

This policy is currently not available for LSM RAID5 volumes because they are not supported in a TruCluster environment. The Tru64 UNIX *Logical Storage Manager* manual (see Resources) provides details on this alternative to RAID1 mirroring. Because its performance in a TruCluster environment has been poor, it is currently not available in TruCluster solutions.

Mirrored volume I/O

No matter the plex type, on a write request to an LSM mirrored volume, LSM will return the request to the caller when the write request has completed to all plexes of the volume. If an error is reported, LSM will detach the faulty plex. This will ensure that the data across plexes are kept consistent.

This function does not change for the fast-fail policy. Instead, on some write request errors, fast-fail allows the underlying I/O subsystem to return the error faster, allowing LSM to process plex detaches faster and continue I/O on surviving plexes.

On a read request to an LSM mirrored volume, LSM will choose one plex to handle the request. Administrators can set one of three read policies on mirrored volumes.

If the *prefer* policy is set, LSM will use the specified preferred plex for all read requests. If the *round* policy is set, LSM will issue read requests in a round-robin fashion across plexes. If the default *select* policy is set, LSM will favor striped plexes on read requests, otherwise defaulting to *round*.

Using the read policy, LSM will issue a read request to the appropriate plex of the volume. If that request returns an error, LSM will try another plex, continuing until a plex can service the request or until all of the volume's plexes are exhausted. Plexes that reported an error will be detached.

This process does not change for the fast-fail policy. As with write requests, on some read request errors, fast-fail enables the underlying I/O subsystem to return the error faster, allowing LSM to process plex detaches faster, continuing I/O on surviving plexes.

LSM offers a configurable writeback policy for read requests. This policy is on by default, but can be manually turned off at any time. If enabled, LSM will retry read request failures in an effort to avoid associated plex detaches.

The method is to find another plex in the volume that can service the read request and turn that read request into a write request for those plexes that returned a read error. If the writeback succeeds, the failed plexes are re-read and if that process succeeds, the plexes will not be detached. If the writeback or re-read fails, the plex detach process will continue. If enabled, fast-fail will apply to writeback-on-read-failure I/O as well. It is rarely, if ever, turned off.

If fast-fail is enabled and every plex of a mirrored volume reported an error for an I/O request, LSM will retry that request with fast-fail turned off. Fast-fail is designed to respond to mirrored volume I/O requests faster by finding a plex to service the request faster. It is not designed to fail a mirrored volume I/O request faster.

This means for some errors, with fast-fail enabled, it might take longer to report a failed mirrored volume I/O request. This time should not be significant compared to a volume I/O failure reported without fast-fail because retries were limited initially. Furthermore, this retry is designed to give the hardware an opportunity to return a successful I/O request, falling in line with the Tru64 UNIX I/O subsystem goal of not returning a device error to the application unless absolutely necessary.

As always, system logs, such as the binary error log and the */var/adm/messages* file, will be useful in determining the cause of a problem.

If the fast-fail policy is not enabled, I/O will be treated normally with no additional impact. I/Os completing successfully will ignore the fast-fail policy, which applies to I/O failures only. Fast-fail does not apply to mirrored volumes with only one enabled plex.

In summation, on some underlying device errors of an LSM mirrored volume, the fast-fail policy will allow LSM to return the I/O request to the caller faster, because DRD (for TruClusters) and CAM will limit the number of underlying retries. As always, LSM will detach failed plexes. After the underlying I/O access problem has been resolved, lost redundancy can be restored at any time. See LSM mirrored volume recovery for additional information.

DRD

By design, DRD routes I/O requests to a member in a TruCluster environment which has direct connectivity to the device. When encountering errors, DRD decides whether the I/O request should be routed to a remote member or returned with the error condition.

If fast-fail is enabled and an I/O request completes with an error, DRD will route that request to another cluster member only if the error status indicates the device is not directly accessible by the current cluster member. If that is the case, DRD will route that request to another cluster member which will retry the I/O. LSM is unaware of this retry. Provided a cluster member is found that can complete the request successfully, LSM will not report an error.

For other errors, if fast-fail is enabled, DRD will return the failed request to LSM immediately. As the error was not the result of a connectivity error, there is no indication that retrying the request on another cluster member will improve the chances of the I/O completing successfully. Furthermore, the retry on other cluster members may increase the time for the I/O to complete, causing the application to not meet its timing requirements.

CAM

If fast-fail is enabled and an I/O request completes with an error, CAM may limit the number of retries it performs. CAM will not limit all retries.

<u>CAM algorithms not changed for fast-fail</u>

By the nature of SCSI there are certain error conditions that are normal and expected. I/O requests reporting these conditions will be retried normally even if fast-fail is enabled.

For example, getting a device ready for I/O is part of a normal retry process to which fast-fail would not apply.

On multi-path devices, path failures will be retried normally, meaning that attempts to access the data will be made using another available path to the device. Only when all paths are exhausted will the I/O be returned with a failure.

Error conditions that are typically transient and not indicative of an actual failure, such as a lost packet on a fabric or a busy device, will be retried normally.

CAM has limits on its retries, as determined by its design and by extensive testing. Most errors are retried a finite number of times, the number of which depends on the error reported. The time required for each retry can vary according to the error code returned, the hardware subsystem, and the amount of I/O. The errors for which fast-fail does not apply are considered transient conditions that do not indicate a hardware problem.

<u>CAM algorithms changed for fast-fail</u>

If a device indicates that bad block replacement (BBR) is necessary, BBR will not be performed if fast-fail is enabled. CAM BBR recovery on device failure can be costly. Furthermore, modern drives do their own BBR and should not be asking CAM to do so.

In addition, all other error cases that may indicate a device is faulty will not be retried if fast-fail is enabled. For some error cases, retries will be limited, generally cut in half. For others, CAM will not retry at all.

LSM mirrored volume recovery

A mirrored volume can have up to 32 plexes, although most customers configure two or to a lesser extent three mirrors per volume. The fast-fail policy may fail plexes faster to improve application response time, but it will do so at the cost of compromising redundancy.

Customers using fast-fail may want to adjust the number of plexes, using a higher count if feasible to lessen the impact of reduced redundancy. It is always advisable to spread mirrors across controllers (also known as HBA's, host bus adaptors) for improved availability.

The fast-fail policy for mirrored volume I/O affects normal reads and writes to mirrored volumes. This does not include I/O to mirrored volumes in LSM recovery mode. There are two forms of LSM mirrored volume recovery: recovery after a system crash and recovery after device failure(s), both of which are described below.

Mirrored volume recovery after a system crash

If a computer crashes and a process on that computer was issuing I/O to a mirrored volume, LSM puts that volume in a special recovery mode.

In a single-system environment, when the computer reboots and LSM is restarted, any mirrored volume that was actively being used will be recovered. In a TruCluster environment, another cluster member will start this recovery if the volume's underlying disks are still available. If the disks are on a bus private to the member that crashed, that member must be accessible to access that storage. Therefore, LSM must wait until that member is rebooted to initiate mirrored volume recovery. LSM uses its metadata to determine if any mirrored volumes need this recovery.

The special recovery mode used is called read-writeback mode. If a mirrored volume is in read-writeback mode, an LSM process reads a portion of a plex of that volume and writes what it read to the same area on other plexes of that volume. Then that process reads the next portion from another plex of that volume and writes what it read across the other plexes. It continues this work back and forth across all plexes until all volume data has been synchronized. Working across plexes in this fashion is called round-robin access.

Mirrored volume recovery ensures that plexes have the same data. For system crashes, all LSM can do is ensure that its plexes are consistent. The I/O subsystem does not know if I/O in flight during the crash completed or not. When the plexes have been resynchronized, the volume is taken out of read-writeback mode.

Resynchronizing a mirrored volume can be a lengthy process depending on the size of the volume. To help, LSM offers a logging policy for mirrored volumes called dirty region logging (DRL). The *Logical Storage Manager* manual details this function, providing examples on enabling it.

If DRL is enabled, regions of the mirrored volume being written to are logged in the DRL. If the system crashes, only those regions marked dirty in the log are resynchronized.

The fast-fail policy does not apply to volumes in recovery mode after a system crash or to DRL I/O; however, the DRL restriction may be relaxed in the future (see Future considerations).

Mirrored volume recovery after a disk failure

As with recovery after a system crash, the fast-fail policy does not apply to mirrored volumes in recovery after a disk failure.

If an I/O request to a plex of a mirrored volume fails, LSM must detach the faulty plex to avoid inconsistencies among mirrors. The plex remains associated with the mirrored volume, but it is in LSM's "detach" state, so I/O will not be issued to it.

On a full LSM restart (manual or system reboot), if the failed LUN is available and does not return an error, LSM will restore this redundancy automatically when the volume is started. If the device remains inaccessible, LSM will terminate recovery. If the device is producing intermittent errors, LSM may succeed in initiating recovery, although I/O performance may be impacted before recovery is terminated.

A way to avoid automatic recovery on LSM restart is via administrative action to disassociate the faulty plex from its volume. If the hardware problem has been resolved before an LSM restart, this should not be necessary.

If LSMs hot spare support is enabled, LSM will automatically recover an underlying disk failure of a mirrored volume using spare LSM storage. The spare storage ideally comes from devices previously added to LSM as a spare device; however, if no spare devices are available, free, unreserved space in the corresponding diskgroup is used.

Because fast-fail does not apply to volumes in recovery mode, spare devices should be chosen wisely or hot-spare should not be enabled at all (the default). If hot-spare support is enabled along with fast-fail, spare devices should optimally be on a different controller than a mirrored volume. If the original failure was due to a faulty controller, this avoids I/O delays during recovery.

If LSM does not initiate mirrored volume recovery automatically after a disk failure, the administrator will need to do so. This recovery can be initiated at any time. The *Logical Storage Manager* manual details how to recover a volume after a disk failure or otherwise reconfigure a mirrored volume on-line.

If the hardware problem has been resolved, mirrored volume recovery should run smoothly.

Whether automatic or user-initiated, this recovery is a plex attach process, which copies data from a good plex to the new plex. The time it takes to attach the plex can be significant depending on the size of the volume, however LSM mirrored volume recovery occurs on-line, so application I/O will continue.

While the fast-fail policy does not apply to LSM mirrored volumes in recovery after a disk failure, this restriction may be relaxed in the future.

<u>I/O during mirrored volume recovery</u>

During either type of mirrored volume recovery, the volume is accessible and will accept new I/O requests. As recovery I/O will be in progress, there will be degradation in application I/O performance during this recovery time. The significance of this degradation depends on the application requirements. It can be measured using the Tru64 UNIX *iostat* or *volstat* commands. See the *iostat(1)* and *volstat(8)* reference pages for more information.

Furthermore, the *Logical Storage Manager* manual details options to LSM mirrored volume recovery commands to help tweak I/O performance during recovery. These commands (*volrecover, volume, volplex*) also have reference pages that explain these options. Administrators can adjust the amount of data recovered at one time and the delays between recovery operations as appropriate.

If fast-fail is enabled, it is possible that additional retries underneath LSM could have corrected the problem and avoided a plex detach and the need for mirrored volume recovery. It is arguable, however, that a hardware problem has been identified and should be investigated.

The fast-fail policy is targeted to those environments with strict application response time requirements in which the possibility of long I/O delays are less tolerable and hardware problems must be investigated and resolved immediately.

That said, it is not expected that hardware will fail often. Statistics on the mean-time-between-failures (MTBF) are available in most device hardware documentation. However if a failure occurs, field service can determine if a device is actually bad by issuing I/O tests and examining system logs. Device errors are not necessarily because a disk is faulty: the cause could be bus or adapter problems. There is a large array of possible hardware problems. System logs are available to help isolate the issue.

The fast-fail policy gives the administrator the option of moving storage off suspect hardware without impacting mirrored volume I/O in progress.

Fast-fail performance

Fast-fail testing for V5.1B-3 was exercised in the following environments.

- Stand-alone system
- 2-member TruCluster environment
- 4-member TruCluster environment

Individual members included DS20E, ES45D, GS160, and GS1280 AlphaServers.

For fast-fail testing, HP's test lab covered both parallel SCSI and fibre-channel storage using a variety of host adapters and a variety of storage facilities, including direct-attach disks, MSA1000, HSG80, and HSV110.

The fast-fail policy was enabled and disabled across multiple runs of many TruCluster, file system, LSM and raw I/O tests using LSM volumes. Failures were instrumented.

Tests injected device errors directly or by disabling ports on a SAN switch. HP's test lab also tested on a system with known erratic hardware. A bus was failing, injecting intermittent costly errors.

For performance numbers, after creating a mirrored volume, a test command was used to simulate a hardware device error. The Tru64 UNIX *dd* command was used to read a block of data through the LSM volume.

With fast-fail disabled, the read request took roughly 55 seconds and with it enabled, the request completed in roughly three seconds. This was consistent across multiple tests in both single-system and TruCluster environments, proving a considerable performance enhancement.

Furthermore, with fast-fail enabled, LSM immediately detached plexes on write errors, avoiding possible excessive I/O delays. This was on a cluster using an erratic bus in its hardware subsystem. Eventually this bus failed so badly that fast-fail was of no benefit. However, a failure noted early allows administrators to make appropriate decisions on the positioning of a mirrored volume's data before I/O delays become too costly.

Summary

The new fast-fail policy for LSM mirrored volume I/O offers an effective solution for some environments, but it has trade-offs that should be evaluated before enabling this policy.

On the one hand, when faced with underlying disk errors, it will allow LSM to fail a plex faster and keep I/O moving, avoiding potentially costly I/O delays on device error.

Conversely, it may fail plexes too quickly, not giving the hardware enough time to correct the problem. This may expose errors not normally seen this early by the administrator.

If a computing environment does not have active administration, it is arguable that Tru64 UNIX should make every effort to try to correct I/O problems, logging data for an administrator to review when available.

Such delays, however, may not be tolerable for some environments and potentially faulty hardware must be evaluated at the first sign of problems. If the configuration has active administration and strict application response times, fast-fail will be especially useful.

The errors targeted for fast-fail are those that indicate a hardware problem rather than something more transient, but there is no guarantee. As with any suspect hardware problem, administrative action will be required to determine what is wrong and how to address it, be it changing the hardware subsystem or adjusting timing requirements. With fast-fail enabled, this investigation can proceed offline without impacting mirrored volume I/O.

While aging hardware can be the most problematic, new hardware can offer challenges as well. It can fail in new ways which could result in excessive retries and corresponding delays in the I/O subsystem. Such problems provide many challenges, often putting the burden on the administrator to make appropriate configuration changes based on information reported by the I/O subsystem. The fast-fail policy for LSM mirrored volume I/O is designed to help administrators troubleshoot without a heavy degradation in mirrored volume I/O performance.

These trade-offs should be evaluated carefully when considering using this new enhancement to LSM mirrored volume I/O.

Related work

The HSG80 V8.8 firmware includes a related enhancement.

Beginning in V8.8, if an LSM mirrored volume uses LUNs from an HSG80 for its plexes, administrators can set the new HOST_REDUNDANT switch for those LUNs. Some errors will be handled differently, reporting an error status that will not be retried by CAM. This will allow for faster response time on device error.

See Resources for more information.

Future considerations

A few enhancements to the fast-fail policy for LSM mirrored volume I/O are under consideration.

It may be useful for administrators to have control over the volumes for which fast-fail is enabled. Therefore, the fast-fail policy for mirrored volume I/O may be settable on a per-volume basis in a future patch kit.

Another consideration is to allow fast-fail to be enabled for I/O to DRL subdisks if they are mirrored and for RAID5 volumes in a single-system environment. Furthermore, on mirrored volume recovery after a disk failure, fast-fail may apply to volumes with more than two mirrors or for any number of mirrors for some errors. This is under investigation.

In addition, if fast-fail is enabled, retries may be further limited in additional CAM device error paths. Work will be on-going as the I/O subsystem is continually tweaked for improved performance.

Finally, allowing the administrator to set timers for LSM mirrored volume I/O is under investigation. Historically, this has been problematic and may cause confusion as guidelines are not always clear.

Comments are welcomed. The Resources section includes the appropriate URL.

Resources

Tru64 UNIX homepage:
http://h30097.www3.hp.com/

Tru64 UNIX documentation:
http://h30097.www3.hp.com/docs/

Tru64 UNIX TruCluster homepage:
http://h30097.www3.hp.com/unix/reliabilityvsavailability.htm

Tru64 UNIX LSM user's guide:
http://h30097.www3.hp.com/docs/base_doc/DOCUMENTATION/V51B_HTML/ARH9BDTE/TITLE.HTM

Configuring LSM for Maximum Performance (Whitepaper):
http://h30097.www3.hp.com/tiplsm.html

Tru64 UNIX tools and other software products:
http://h30097.www3.hp.com/products.html

Web Based Enterprise Services (WEBES):
http://h18023.www1.hp.com/support/svctools/webes/

HP Hardware Library:
http://cybrary.inet.cpqcorp.net/

HSG80 V8.8 Reference Guide:
http://h200001.www2.hp.com/bc/docs/support/SupportManual/c00309747/c00309747.pdf

Feedback
http://h30097.www3.hp.com/comments.html

Glossary

CAM
ANSI standard defining an operating system interface to SCSI hardware; CAM (Common Access Method) manages pools of buffers used to perform physical disk I/O

DRD
A layer in the Tru64 UNIX operating system I/O stack for TruCluster environments; DRD (Device Request Dispatcher) allows cluster members to have cluster-wide access to all disk, tape, floppy, cdrom, and dvdrom

Fabric
Switched fibre-channel topology using anywhere from a single switch or router function to multiple interconnected switches or routers to deliver a frame of I/O data from a source node to a destination node; a node in this context can be a computer system, RAID array controller, disk device, tape device, or other fabric connected device

Fibre-channel
Serial bus connecting high speed storage devices to computers; Tru64 UNIX uses the Fibre-Channel Protocol (FCP) for SCSI for Fibre Channel use as the physical interface

LSM
The Logical Storage Manager (LSM) is a host-based storage management utility for HPs Tru64 UNIX operating system, managing disk storage on-line in both single-system and TruCluster environments

LSM DRL
Dirty Region Logging (DRL) is used to speed up mirrored volume recovery time after a system crash by recovering mirrored volume regions marked dirty in a log rather than recovering the entire volume

LSM plex
A mirror in an LSM volume comprised of one or more subdisks

LSM plex attach
Adding a mirror to a volume, synchronizing it with the volume’s data

LSM subdisk
Basic building block of a volume; an offset and length on a physical disk

LSM volume
Virtual disk treated by applications as a physical disk however multiple physical disks can be used in the creation of a volume for fault tolerance (redundancy) and/or improved I/O performance; software RAID

LSM volume recovery
Resynchronizing plexes after a system crash or restoring redundancy by re-attaching a previously failed plex

LUN
Logical Unit Number (LUN) used to identify a physical disk or collection of physical disks presented to the operating system as one device (a logical unit); the LUN is used to identify the unit itself in addition to the numeric identifier

RAID
Redundant Array of Independent (or Inexpensive) Disks; a category of disk drives that use two or more drives in combination for fault tolerance (redundancy) and improved performance

RAID0
RAID level where data is striped across a set of disks for improved performance

RAID1
RAID level where data is mirrored across a set of disks for improved availability

RAID5
RAID level where data is striped across a set of disks for improved performance and parity is distributed across the disks as well for redundancy

TruCluster Server Product
HPs single-system image clustering technology allowing multiple AlphaServer systems running the Tru64 UNIX operating system to share resources and act as a single virtual machine; TruCluster technology provides high availability and improved performance in particular for distributed parallel processing applications