Release Notes— Fluid Cache for SAN 2.1.0 for Linux Systems

Build Version: 2.1.0

Updated 5/10/16

Contents

Purpose of this Release

Fluid Cache for SAN for Linux 2.1.0 is a maintenance release with the following focus:

- Add support for RHEL 6.7
- Improve diagnostic capability when there are cluster issues
- Improve cluster shutdown robustness
- Fix stability issues and improve cluster operation

Note: Installation and Upgrade instructions for 2.1.0 are in the Dell Fluid Cache for SAN Version 2.1.0 Admin and Deployment Guide for Linux Systems.

Fixed Issues in this Release

Improve diagnostic capability
- Reduced log lines
- Reduce max log file size
- Consistent log format
- Stuck SCSI2 logging
- RoCE network verification
- Past anomaly detection

Cluster management usability
- CFM Shutdown Intelligence – do not kill all
- Marking blocks truly clean – quicker recovery time
- Improve handling of CFM crash during uncaching

Operational improvement
- Issues with write phase of compare and write
- Shutdown failure scenario
- MD Crash due to unresolved duplicates
- Performance degradation during unaligned I/Os
- Cache device could not be used

- I/O timeouts and errors after cache network failure
- Data loss in multiple failure scenario

Software and Hardware Compatibility

See Dell Fluid Cache for SAN Compatibility Matrix available at dell.com/CacheSolutions

Driver Required Version

Table 1   Driver Required Version

Element	Required Version
Mellanox ConnectX-3 OFED Drivers	RHEL 6.7: 3.2-1.0.1 [Download]

Linux Operating System Dependencies

To check dependencies that are currently installed on your system, run the following command: `rpm –qa`

Ensure that the following open source component packages are available in your system before you install Fluid Cache for SAN. See the operating system distribution media for the supported version of each package.

The following dependencies are required for RHEL, except as noted. The dependencies for SLES are similar. For Oracle Linux, contact Dell Customer Support.

- avahi-libs
- avahi-tools (RHEL)
- avahi-utils (SLES)
- bash
- glib2
- glibc-2.12.1.166.el6_7.1
- kernel
- libaio
- libibverbs*
- librdmacm*
- libuuid
- libxml2
- libxslt

- openssl
- pam
- perl
- perl-XML-LibXML
- perl-XML-NamespaceSupport
- perl-XML-SAX
- per-YAML-LibYAML
- python
- sg3_utils-libs
- sg3_utils
- shadow-utils
- xmlsec1-openssl
- xmlsec1
- zlib

* Installed when you install the Mellanox OFED package.

Installation Instructions

Refer to the Admin and *Deployment Guide* for complete installation instructions. However, before deploying Fluid Cache, you must verify the digital signature of the Fluid Cache package.

Verifying the Digital Signature of a Package

To verify a digital signature on a Linux system, you must have the Gnu Privacy Guard (GPG) package installed.

1. Get the Dell Linux public GnuPG key by searching http://pgp.mit.edu/ for this footprint: 0x1285491434d8786f

2. Import the public key by running the following command:
   `gpg --import <public_key>`

A message appears, informing you that the public key was processed but was unchanged.

3. Validate the Dell public key, if you haven't done so previously, by entering the following command: `gpg --list-keys`

4. Trust the Dell public key with the command `gpg –edit-key 23B66A9D` and then enter the command `trust`.

5. Decide how far you trust this user to correctly verify other users' keys (by looking at passports, checking fingerprints from different sources, and so on). Select the highest number/level of trust, and then enter the command `quit`.

6. Verify the file's digital signature by entering the following command:

```
gpg --verify <file_name>.sign <file_name>
```

A message appears, telling you that the verification resulted in a good signature.

NOTE: If you have not validated the key as shown in the previous step, additional messages appear to warn you that the key was not certified with a trusted signature.

Known Issues Arising from Third Party Software

The following issues are not caused by Fluid Cache itself, but arise from known issues in third party software.

Low Memory Handling Issues

ISSUE: When the system is low on memory, memory allocation errors, I/O timeouts, and EAGAIN errors (resource temporarily unavailable) may appear.

WORKAROUND: Add or free up RAM memory.

OpenManage Installation Causes RDMA to Fail

ISSUE: On SLES 11 SP3, if you install OpenManage 7.4 after installing the Mellanox drivers, the installation of libmlx4-rdmav2 breaks the already installed RDMA configuration and Fluid Cache runs in TCP mode instead of RDMA.

WORKAROUND: Reinstall the Mellanox drivers after installing OpenManage and before installing Fluid Cache.

JDK Installer Generates Log Messages

ISSUE: Installing Fluid Cache after installing jdk-7u6-linux-x64.rpm causes the JDK installer to send harmless messages to log. This is a known Oracle and Java bug that occurs only on SLES.

WORKAROUND: Add a Required-Stop section to the jexec init script to stop the messages from being generated. In the file /etc/init.d/jexec, enter the following line:

```
# Required_Stop:
```

Volume Manager Disk Creation

ISSUE: Attempts to use Volume Manager to create a volume on a Fluid Cache disk fail with the following error: `Device /dev/fldc0 not found (or ignored by filtering)`

WORKAROUND: Volume Manager is not supported by Fluid Cache, and should not be used to create volumes out of Fluid Cache disks. Instead, create a volume on the backend store, and then cache that volume.

XFS Hangs During File System Creation

ISSUE: On RHEL 6.4, the XFS files system may become unresponsive. During creation of the file system, this may occur because of a deadlock between the xfssyncd and xfsbufd daemons. Consequently, several kernel threads become unresponsive and the XFS file system cannot be successfully created, leading to a kernel oops. A patch is available to prevent this situation by forcing the active XFS log onto a disk. For more details, see the bug report at https://bugzilla.redhat.com/show_bug.cgi?id=1014867.

WORKAROUND: A patch that addresses this issue is available at http://rhn.redhat.com/errata/RHSA-2013-1645.html.

XFS Hangs During Normal System Operation

ISSUE: On RHEL 6.4 and SLES 11 SP3, the XFS file system may become unresponsive when running a large number of file operations. For more details, see the bug report at http://oss.sgi.com/bugzilla/show_bug.cgi?id=922.

WORKAROUND: None

Btrfs Hangs During Normal System Operation

ISSUE: Under some circumstances, the Btrfs file system becomes unstable, goes into read-only mode, and becomes unresponsive.

WORKAROUND: None

Btrfs Hangs During Normal System Operation

ISSUE: When using the Btrfs file system on SLES 11 SP3, under some circumstances the kernel can become so busy dumping stack traces that the operating system starts sending non-maskable interrupts (NMI) and becomes unresponsive.

WORKAROUND: None

Ext4 Data Integrity Issues

ISSUE: When using the Ext4 file system, under some circumstances data at the end of a cached file may be lost, or data within a file may be lost.

WORKAROUND: None

Ext4 Buffer Cache Does Not Empty

ISSUE: When using the Ext4 file system, under certain circumstances the buffer cache may not completely empty.

WORKAROUND: None

Known Issues in Fluid Cache for SAN 2.0

The following are known issues in Fluid Cache for SAN 2.0.

Removing a Server from a Cluster

ISSUE: If you try to remove a server from a cluster before shutting down the server, Enterprise Manager displays a message saying that the action is not allowed.

WORKAROUND: Shut down the server. To shut down the server, in Enterprise Manager, locate the server, right-click, and select **Shut Down**. Then to remove a cluster from server, refer to the section "Removing a Server from a Cluster" in the Admin and *Deployment Guide.*

Removing Multiple Servers from a Cluster

ISSUE: Under some circumstances, removing more than one server at a time from a cluster may cause I/O to hang or data to be lost.

WORKAROUND: Avoid removing more than one server at a time from a cluster.

Replacing a Failed Node on a Three Node Cluster

ISSUE: If one of the nodes in a three node cluster fails, Enterprise Manager does not allow removal of the failed node, and displays a message that removing the node would result in fewer than the minimum of three nodes required. If the node is manually removed and returned to an operational state, Enterprise Manager does not allow it to rejoin the cluster, displaying a message that the node already belongs to the cluster.

WORKAROUND: Image a fourth node and add it to the cluster, then remove the failed node from the cluster.

File Systems Table (fstab) Settings Prevent Server from Starting

ISSUE: The server can fail to boot because of errors related to setting non-zero fs_passno field in the /etc/fstab file. Setting the force fsck option causes the system boot to fail. This is because the devices do not exist at boot time when fstab is processed. When the error occurs, the following message is displayed:

```
/dev/fldc0 is mounted. e2fsck: Cannot continue, aborting.

/dev/fldc1 is mounted. e2fsck: Cannot continue, aborting.

*** An error occurred during the file system check.

*** Dropping you to a shell; the system will reboot

*** when you leave the shell.
```

The sixth field (fsck option) is used by the fsck (8) program to determine the order in which file system checks are done at reboot time. If the sixth field is set to zero, these system check errors do not occur during reboot. Therefore it is recommended that for any fldc-related user entry in /etc/fstab, the fifth field (dump option) and the sixth field (fsck option) are both set to zero. In case the sixth field is not zero, it is very likely the above Linux error message occurs during reboot.

WORKAROUND: Remount the root file system as read/write (mount -n -o remount, rw /) and then comment out the Fluid Cache disk device entries from fstab, as shown:

```
#/dev/fldc0 /root/fldc/mnt ext2 defaults 1 3
```

Then change:

```
/dev/fldc1 /root/fldc/mnt2 ext2 defaults 1 4
```

either to:

```
/dev/fldc1 /root/fldc/mnt2 ext2 defaults 1 0
```

or to:

```
#/dev/fldc1 /root/fldc/mnt2 ext2 defaults 1 4
```

Root File System Fills Up

ISSUE: Fluid Cache may fail and require manual intervention if the root file system on a node becomes full. The cluster may still be operational, but error messages are generated and Fluid Cache is not fully functional.

WORKAROUND: Please contact Dell Customer Support for the steps required to return the node to full functionality.

Configuration Errors after Mapping Volume Fails

ISSUE: When mapping a volume to a Fluid Cache node, if the operation times out or fails to complete normally, under some circumstances a partial configuration is created on the node, and Enterprise Manager mistakenly shows an apparently normal volume mapping. When the mapping is later completed normally, this partial configuration remains, and may interfere with administrative actions. For instance, deleting the cluster may fail because Storage Center has a record of a volume still in use by the cluster.

WORKAROUND: Contact Dell Customer Support for assistance in resolving this issue.

Timeouts on PowerEdge R820 Servers

ISSUE: On SLES 11 SP3, under some circumstances, I/O loads could cause CPU usage spikes that decrease performance on PowerEdge R820 Servers.

WORKAROUND: None

GPT Partition Table Corruption

ISSUE: Partprobe reports GUID Partition Table (GPT) corruption and device mapper errors on all other nodes after making partition table changes to cached LUNs on one node.

WORKAROUND: Do not modify partition tables on cached LUNs. Modifying partition tables on shared volumes is not supported.

Use of Kpartx to Partition Volumes

ISSUE: When using the kpartx command to partition volumes, the partition table on shared block devices is in an inconsistent state across the cluster. For instance, partitions created on a LUN on one node may not be seen by other nodes.

WORKAROUND: Kpartx is not supported for partitioning a cached volume. Use the partprobe command instead.

Syslog Error Messsages

ISSUE: The avahi-daemon writes numerous messages to syslog about receiving a packet from an invalid interface when there is another mDNS stack (such as the one that Oracle RAC installs) installed and running on the same node as Fluid Cache.

WORKAROUND: Move the other mDNS stack of the application to another node, if possible. Otherwise, it is possible to patch and recompile the avahi-daemon to stop sending this message. Contact Dell Customer Support for instructions on patching the avahi-daemon or to obtain a replacement avahi-daemon.

Stopping the Fluid Cache Service Creates Segmentation Fault

ISSUE: On SLES 11 SP3, stopping the Fluid Cache service results in a segmentation fault because of a conflict in the Avahi binary rpm.

WORKAROUND: Contact Dell Customer Support to obtain a patch with which to recompile the Avahi binary rpm, or to obtain a pre-compiled version of the Avahi binary rpm. You may also contact SLES to see whether a version of Avahi is available that is based on source version 0.6.24 or greater.