gfs2-uevents.txt

uevents and GFS2

During the lifetime of a GFS2 mount, a number of uevents are generated.
This document explains what the events are and what they are used
for (by gfs_controld in gfs2-utils).

A list of GFS2 uevents

1. ADD

The ADD event occurs at mount time. It will always be the first
uevent generated by the newly created filesystem. If the mount
is successful, an ONLINE uevent will follow. If it is not successful
then a REMOVE uevent will follow.

The ADD uevent has two environment variables: SPECTATOR=[0|1]
and RDONLY=[0|1] that specify the spectator status (a read-only mount
with no journal assigned), and read-only (with journal assigned) status
of the filesystem respectively.

2. ONLINE

The ONLINE uevent is generated after a successful mount or remount. It
has the same environment variables as the ADD uevent. The ONLINE
uevent, along with the two environment variables for spectator and
RDONLY are a relatively recent addition (2.6.32-rc+) and will not
be generated by older kernels.

3. CHANGE

The CHANGE uevent is used in two places. One is when reporting the
successful mount of the filesystem by the first node (FIRSTMOUNT=Done).
This is used as a signal by gfs_controld that it is then ok for other
nodes in the cluster to mount the filesystem.

The other CHANGE uevent is used to inform of the completion
of journal recovery for one of the filesystems journals. It has
two environment variables, JID= which specifies the journal id which
has just been recovered, and RECOVERY=[Done|Failed] to indicate the
success (or otherwise) of the operation. These uevents are generated
for every journal recovered, whether it is during the initial mount
process or as the result of gfs_controld requesting a specific journal
recovery via the /sys/fs/gfs2/<fsname>/lock_module/recovery file.

Because the CHANGE uevent was used (in early versions of gfs_controld)
without checking the environment variables to discover the state, we
cannot add any more functions to it without running the risk of
someone using an older version of the user tools and breaking their
cluster. For this reason the ONLINE uevent was used when adding a new
uevent for a successful mount or remount.

4. OFFLINE

The OFFLINE uevent is only generated due to filesystem errors and is used

as part of the "withdraw" mechanism. Currently this doesn't give any information about what the error is, which is something that needs to be fixed.

5. REMOVE

The REMOVE uevent is generated at the end of an unsuccessful mount or at the end of a umount of the filesystem. All REMOVE uevents will have been preceeded by at least an ADD uevent for the same fileystem, and unlike the other uevents is generated automatically by the kernel's kobject subsystem.

Information common to all GFS2 uevents (uevent environment variables)
---

1. LOCKTABLE=

The LOCKTABLE is a string, as supplied on the mount command line (locktable=) or via fstab. It is used as a filesystem label as well as providing the information for a lock_dlm mount to be able to join the cluster.

2. LOCKPROTO=

The LOCKPROTO is a string, and its value depends on what is set on the mount command line, or via fstab. It will be either lock_nolock or lock_dlm. In the future other lock managers may be supported.

3. JOURNALID=

If a journal is in use by the filesystem (journals are not assigned for spectator mounts) then this will give the numeric journal id in all GFS2 uevents.

4. UUID=

With recent versions of gfs2-utils, mkfs.gfs2 writes a UUID into the filesystem superblock. If it exists, this will be included in every uevent relating to the filesystem.