

Re: [RFC][PATCH] inotify 0.13.1

Source: <http://linux.derkeiler.com/Mailing-Lists/Kernel/2004-10/2055.html>

From: Robert Love (rml_at_novell.com)

Date: 10/08/04

To: John McCutchan <ttb@tentacle.dhs.org>

Date: Thu, 07 Oct 2004 19:37:36 -0400

OK.

I am working on redoing the inotify locking wrt inodes.

Most of the time that we grab `inode->i_lock`, we don't need it. We do need it when allocating or freeing `inode->inotify_data`. We also need it when first obtaining the `inotify_data` reference count (see below).

So we need an `inotify_data` reference count to track its lifetime. We already have `->watcher_count`, which functions pretty well as this ref count, but we need to formalize it, make it `atomic_t`, and always grab it under `->i_lock`. Freeing the structure is then a consequence of the ref count dropping to zero via `atomic_dec_and_test()`.

Then we can add `inotify_data->lock` to protect the contents of `inotify_data` from concurrent access and not use `i_lock`.

Anyone dereferencing `inotify_data` (or setting/freeing it for the first time) needs to hold a ref count. This is an existing bug.

We can hopefully end up only taking `inode->i_lock` when grabbing this ref count, and no other time. We won't have to nest it with `dev->lock` any more.

That is the plan, anyhow. Patch later tonight or tomorrow, with luck.

Robert Love

-

To unsubscribe from this list: send the line "unsubscribe linux-kernel" in the body of a message to majordomo@vger.kernel.org

More majordomo info at <http://vger.kernel.org/majordomo-info.html>

Please read the FAQ at <http://www.tux.org/lkml/>