

Re: jumbograms (& em) & nfs a no go

Source: <http://unix.derkeiler.com/Mailing-Lists/FreeBSD/current/2004-01/2045.html>

From: Terry Lambert (tlambert2_at_mindspring.com)

Date: 11/01/03

Date: Sat, 01 Nov 2003 12:28:50 -0800

To: Michal Mertl <mime@traveller.cz>

Michal Mertl wrote:

> *On Fri, 31 Oct 2003, Terry Lambert wrote:*

> > *Michal Mertl wrote:*

> > > *I then left one computer at 4.9 and upgraded the other to 5.0. When I*

> > > *mount a partition from 5.0 machine I found out, that copying reliably*

> > > *works only from 5.0 to 4.9. The other way around I see messages 'em0:*

> > > *discard oversize frame (ether type 800 flags 3 len 67582 > max 6014)' on*

> > > *5.0 and the copying stalls. On 4.9 machine I later see 'nfs server*

> > > *10.0.0.2:/usr: not responding'. The interface is stuck for some time – can*

> > > *be revived by changing mtu back to 1500 and down/up sequence.*

> >

> > *Implies the sending host is not honoring the MTU restriction when*

> > *deciding whether or not to frag packets.*

>

> *Can you suggest what to do to find out what's really happening? I thought*

> *nfsd network part was mostly userland thus the same as ftpd (or better*

> *netperf) and should work.*

No. Traditionally (except in Linux), nfsd is a userland thing that calls a system call and never returns to user space. It exists in order to provide a process context for use by blocking calls in the kernel, specifically for use by `tsleep()`, `wakeup()`, and so on. In more modern UNIX systems, it's a kernel thread, and has no user space existence at all, or, on systems that will permit NFS to be turned off, and don't have the ability/desire to hang the kernel on/off state off the existence/nonexistence of active exports, it's a stub that tells the kernel to run the kernel thread(s).

The easiest way to find out what's happening is to grep the BSD sources where the message is coming from, and then work back to understand the code paths that permitted something that's 67582 bytes in size to get there in the first place.

Not looking at the absolutely newest sources, from memory, my original comment was based on "it had to come from the driver that way". This may or may not be a valid assumption, but it's

freebsd-current: Re: jumbograms (& em) & nfs a no go

at last a starting hypothe