

freebsd-arch: Re: simplify disksort, please review.

Re: simplify disksort, please review.

Source: <http://unix.derkeiler.com/Mailing-Lists/FreeBSD/arch/2005-06/0044.html>

From: Pawel Jakub Dawidek (pjd_at_FreeBSD.org)

Date: 06/09/05

Date: Thu, 9 Jun 2005 21:30:08 +0200

To: Jeff Roberson <jroberson@chesapeake.net>

On Wed, Jun 08, 2005 at 04:43:03PM -0400, Jeff Roberson wrote:

+> <http://www.chesapeake.net/~jroberson/disksort.diff>

+>

+> Our disksort algorithm used to be complicated by the BIO_ORDERED flag,
+> which could cause us to make some exceptions in the sorting. When the
+> ordered support was removed we never simplified the algorithm. The patch
+> above gets rid of the switch point and associated logic. It's now a
+> simple hinted insertion sort with a one way scan. Since it's a fairly
+> central algorithm, I'd appreciate a review.

It is not related to you patch, but I think our disk sort is just too trivial.

The one example of how the order can be broken (`write(offset, size)`):

```
write(1024, 512)
write(0, 2048)
```

It'll be sorted to:

```
write(0, 2048)
write(1024, 512)
```

because we're sorting only based on offset.

AFAIR we now require from file system not to send next request to the same area until previous request is not finished, so we should be safe (for now).

--

Pawel Jakub Dawidek
pjd@FreeBSD.org
FreeBSD committer

<http://www.wheel.pl>
<http://www.FreeBSD.org>
Am I Evil? Yes, I Am!

- application/pgp-signature attachment: [stored](#)

Re: simplify disksort, please review.