

## Re: newfs and mount vs. half-baked disks

**Source:** <http://unix.derkeiler.com/Mailing-Lists/FreeBSD/arch/2003-11/0030.html>

---

**From:** Kirk McKusick ([mckusick\\_at\\_beastie.mckusick.com](mailto:mckusick_at_beastie.mckusick.com))

**Date:** 11/10/03

To: Bruce Evans <[bde@zeta.org.au](mailto:bde@zeta.org.au)>  
Date: Sun, 09 Nov 2003 15:51:55 -0800

> *Date: Sat, 8 Nov 2003 19:35:16 +1100 (EST)*  
> *From: Bruce Evans <[bde@zeta.org.au](mailto:bde@zeta.org.au)>*  
> *X-X-Sender: [bde@gamplex.bde.org](mailto:bde@gamplex.bde.org)*  
> *To: Kirk McKusick <[mckusick@mckusick.com](mailto:mckusick@mckusick.com)>*  
> *Cc: [arch@freebsd.org](mailto:arch@freebsd.org)*  
> *Subject: Re: newfs and mount vs. half-baked disks*  
> *X-ASK-Info: Whitelist match*  
>  
> *On Fri, 7 Nov 2003, Kirk McKusick wrote:*  
>  
>> *From: Bruce Evans <[bde@zeta.org.au](mailto:bde@zeta.org.au)>*  
>> *> ...*  
>> *> Some cases can be discerned anyway, depending on how much erasing of*  
>> *> metadata newfs does when it starts. E.g., if there was an ffs file*  
>> *> system on the disk, then this will be recorded in the disk label unless*  
>> *> that feature has been axed). newfs doesn't rewrite the label until*  
>> *> just before it finishes, so the old label can be looked at to determine*  
>> *> what was there. Writing the label last may be a bug though. Perhaps*  
>> *> newfs should erase all the primary metadata for the old filesystem*  
>> *> when it starts so as to minimise the window where there may be*  
>> *> conflicting metadata.*  
>>  
>> *You cannot depend on the disk label as the disklabel is going away*  
>> *or at least being wholly overhauled with GEOM. In particular, the*  
>> *existing disk label only has a 2^32 block count which is insufficient*  
>> *for filesystems larger than 2Tb.*  
>  
> *I don't use GEOM, so the label won't be going away for me. Anyway, there*  
> *is no dependency (the label is just one of the things that one might*  
> *examine to recover a crashed disk), and any overaul by GEOM would have*  
> *to duplicate the functionality of storing metadata about the superblocks*  
> *somewhere outside the superblocks. (I actually store metadata about file*  
> *systems in (backups of) disk files in /var/backups. Normal backups*  
> *provide inadequate backups of metadata.)*

I agree that there needs to be some place that the additional information

freebsd-arch: Re: newfs and mount vs. half-baked disks

be stored to let fsck find the alternate superblocks. However, it seems unlikely that it will be the disk label as FreeBSD is looking to go to one of the existing disk label standards which do not have space reserved for that purpose. So, my guess is that some of the "boot" area will have to be reserved to that purpose.

- > *The block count is in units of sector size, so disks much larger than*
- > *2TB can be supported by disklabel using (fake if necessary) sector sizes*
- > *larger than 512. File systems need to use similarly large block (fragment*
- > *for ffs) sizes, and some patches are needed for reading superblocks if*
- > *the sector size is larger than 8K. Since ffs uses a block size of 16K*
- > *by default, a sector size of 16K are not unreasonable and this is sufficient*
- > *for disks smaller than 64TB.*
- >
- > *Bruce*

Actually, FFS requires its fragment size be no smaller than the sector size (since it presumes that it cannot do read/write in smaller than sector sizes). So, on a 16K filesystem, you get 2K fragments. So your hack only gets you to 8TB which is not going to last long at current disk growth rates.

Kirk McKusick

---

freebsd-arch@freebsd.org mailing list

<http://lists.freebsd.org/mailman/listinfo/freebsd-arch>

To unsubscribe, send any mail to "freebsd-arch-unsubscribe@freebsd.org"