

Re: Apparent strange disk behaviour in 6.0

Source: <http://unix.derkeiler.com/Mailing-Lists/FreeBSD/current/2005-07/1379.html>

From: Julian Elischer (julian_at_elischer.org)

Date: 07/30/05

Date: Sat, 30 Jul 2005 12:47:36 -0700
To: Poul-Henning Kamp <phk@haven.freebsd.dk>

Poul-Henning Kamp wrote:

> In message <20050730171536.GA740@uk.tiscali.com>, Brian Candler writes:

>

>> On Sat, Jul 30, 2005 at 03:29:27AM -0700, Julian Elischer wrote:

>>

>>> The snapshot below is typical when doing tar from one drive to another..

>>> (tar c -C /disk1 f- .|tar x -C /disk2 -f -)

>>>

>>> dT: 1.052 flag_1 1000000us sizeof 240 i -1

>>> L(q) ops/s r/s kBps ms/r w/s kBps ms/w d/s kBps ms/d %busy Name

>>> 0 405 405 1057 0.2 0 0 0.0 0 0 0.0 9.8| ad0

>>> 0 405 405 1057 0.3 0 0 0.0 0 0 0.0 11.0| ad0s2

>>> 0 866 3 46 0.4 863 8459 0.7 0 0 0.0 63.8| da0

>>> 25 866 3 46 0.5 863 8459 0.8 0 0 0.0 66.1| da0s1

>>> 0 405 405 1057 0.3 0 0 0.0 0 0 0.0 12.1| ad0s2f

>>> 195 866 3 46 0.5 863 8459 0.8 0 0 0.0 68.1| da0s1d

>>>

>>> even though the process should be disk limited neither of the disks is

>>> anywhere

>>> near 100%.

>>

>> One IDE disk doing 405 reads per second (2.5ms per seek) is pretty good.

>

>

> Sorry, but your reasoning is terminally wrong.

>

> The 405 reads/sec takes only 0.2 msec per read, including transfer,

> seek and other overhead.

>

> You can read this number directly in the "ms/r" column.

>

> It follows effortlessly that the reads do not result in random seeks.

>

>

>> But if really is only 12.1% busy (which the 0.3 ms/r implies),

>

>

freebsd-current: Re: Apparent strange disk behaviour in 6.0

- > *"busy %" numbers is *NOT* a valid measure of disk throughput, please do*
- > *not pay attention to such numbers!*
- >
- > *A disk can be 100% busy and still be able to accept 128 times more*
- > *traffic:*
- >
- > *read sector 0*
- > *read sector N*
- > *read sector 1*
- > *read sector N-1*
- > ...
- >
- > *will keep the disk 100% busy without getting much done.*
- >
- > *read sector 0*
- > *read sector 1*
- > *read sector N-1*
- > *read sector N*
- > *read sector 2*
- > *read sector 3*
- > *read sector N-3*
- > *read sector N-2*
- > ...
- >
- > *Will get twice as much done and still keep the disk 100% busy.*
- >
- >
- > *If you want to know how busy your disk is, simply look in the ms/r*
- > *and ms/r columns and decide if you can live with that average*
- > *transaction time. If it is too high for your liking, then your*
- > *disk is too busy.*
- >
- > *If you want to do quantitative predictions, you need to do the*
- > *queue-theory thing on those numbers.*
- >
- > *If you know your queue-theory, you also know why busy% is*
- > *a pointless measurement: It represents the amount of time*
- > *where the queue is non-empty. It doesn't say anything about*
- > *how quickly the queue drains or fills.*

exactly.. I'm trying to work out why teh read and write queues are empty for so much time in a transaction that SHOULD be disk bound....

there should ALWAYS be something waiting to be read on teh IDE drive (since it is slower or on the SCSI drive, since it has to do more work (sync meta writes)? (no soft updates)

Interestingly, if I do
mount -u -oasync on the raid, teh number of transactions per second goes down but the throughput remains about the same. suggesting that the ATA drive is teh bottleneck, but that continues to report lots of time with no queued

freebsd-current: Re: Apparent strange disk behaviour in 6.0

requests.

>

freebsd-current@freebsd.org mailing list

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

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