

Re: Stripe sizing and maxcontig

rate, not just the average, and measure what the database is doing rather than just your current filesystem parameters: does three 14kB reads really mean the database wanted 30kB, for instance? And what is your database block size?

Speak to your DBA. Are your measurements valid for each table? Can read-only tables be isolated? Are database log files random access? (Mind you, does Oracle still recommend SAME and hope for the best?)

My assumptions are : For random access, I'm assuming the best thing to do is to set the stripe size of a RAID volume to the average transfer size – so if I have, say, 4 disks in a volume they could all in theory be handling different requests. If I'm doing mostly sequential transfers, I'm better off setting the stripe size to (number of disks * transfer size), so a single transfer is spread out amongst as many disks as possible. Is this correct ?

Strictly, if you are using raid-5 then the discs cannot all be handling different requests because each request will involve more than one disc because of the parity.

Sequential transfers are generally favoured by wide stripes and big blocks whereas for random i/o the reverse tends to be true. Assuming random i/o of small size, it is probably best to set the stripe size to a bit larger than your i/o ops (say, 16kB) so they can be satisfied in one go.

There are some guidelines in the 3510 documentation on sun.com as well as limitations.

If I've got that right, my next question relates to maxcontig. If I set the stripe size of the RAID volume to 16Kb (closest to 14Kb value mentioned above), I know I then need to tune maxcontig accordingly. If I lower this from the default setting of 7 to 6, this means the system should bunch together transfers until they reach a size of 48Kb. If I have 4 disks in my RAID 5 volume (3 available volumes), and a stripe size of 16Kb so that the total stripe width is also 48Kb, does that mean that every transfer will be spread evenly across all disks ?

If you have three 4-disc raid-5 volumes then you have no online spares: is this wise?

Are you sure the default maxcontig is 7 and not 128? Or is this

Re: Stripe sizing and maxcontig

Re: Stripe sizing and maxcontig

a sparc vs x86 thing? Or is my memory failing? It depends whether you are doing sequential or random i/o. In any case, you may do better to mount the filesystem with forcedirectio and not worry about it.

This isn't really going anywhere so how about a change of tack?

Option 1: (the cynic's option) your db is not doing much so the 3510 will more than adequately keep up with it using all the default settings.

Option 2: take the opportunity to experiment with different parameters before you go live. Probably your db supplier publishes guidelines. This will have the additional benefit of making you familiar with the 3510 when a controller fails at 3 o'clock in the morning.

—

John.

.