

Re: Cached file read performance

Source: <http://unix.derkeiler.com/Mailing-Lists/FreeBSD/performance/2006-12/msg00011.html>

- *From:* Bruce Evans <bde@xxxxxxxxxxxx>
 - *Date:* Fri, 22 Dec 2006 22:18:04 +1100 (EST)
-

On Fri, 22 Dec 2006, Mark Kirkwood wrote:

I recently did some testing on the performance of cached reads using two (almost identical) systems, one running FreeBSD 6.2PRE and the other running Gentoo Linux – the latter acting as a control. I initially started a thread of the same name on –stable, but it was suggested I submit a mail here.

Linux has less bloat in the file system and (cached, at least) block i/o paths, so it won't be competed with fully any time soon. However, the differences shouldn't be more than a factor of 2.

conclusion I came to was that our (i.e FreeBSD) cached read performance (particularly for smaller block sizes) could perhaps be improved... now I'm

None was attached.

Machines
=====

- FreeBSD (6.2-PRERELEASE #7: Mon Nov 27 19:32:33 NZDT 2006):
- Supermicro P3TDER
 - 2xSL5QL 1.26 GHz PIII
 - 2xKingston PC133 RCC Registered 1GB DIMMS
 - 3Ware 7506 4x Maxtor Plus 9 ATA-133 7200 80G
 - Kernal GENERIC + SMP
 - /etc/malloc.conf -> >aj
 - ufs2 32k blocksize, 4K fragments

^^^^^^^^^^^^^^^^^^^^^^^^^^

Try using an unpessimized block size. Block sizes larger than BKVASIZE (default 16K) fragment the buffer cache virtual memory. However, I couldn't see much difference between block sizes of 16, 32 and 64K for a small (32MB) md-allocated file system with a simple test program.

Re: Cached file read performance

All versions got nearly 1/4 of bandwidth of main memory (800MB/S +-10% an an AthlonXP with ~PC3200 memory). On this system, half of the bandwidth of main memory is (apparently) unavailable for reads because it has to go through the CPU caches (only nontemporal writes go at full speed), and another 1/2 of the bandwidth is lost to system overheads, so 800MB/S is within a factor of 2 of the best possible.

– RAID0 256K stripe using txe driver

Gentoo (2.6.18-gentoo-r3):

- Supermicro P3TDER
- 2xSL5QL 1.26 GHz PIII
- 2xKingston PC133 RCC Registered 1GB DIMMS
- Promise TX4000 4x Maxtor plus 8 ATA-133 7200 40G
- default make CFLAGS (-O2 -march-i686)
- xfs stripe width 2
- RAID0 256K stripe using md driver (software RAID)

PIII's and PC133 are very slow these days. I could never get more than a couple of hundred MB/s main memory copy bandwidth out of PC100. PC133 and the read bandwidth are not much faster. The read bandwidth on freefall (800 MHz PIII) with a block size of 4MB is now 500MB/S for my best read methods.

Given the tests were about cached I/O, the differences in RAID controller and the disks themselves were seen as not significant (indeed booting the FreeBSD box with the Gentoo livecd and running the tests there confirmed this).

Yes, if the disk LED blinks then the test is invalid.

```
$ ./readtest /data0/dump/file 8192 0
random reads: 100000 of: 8192 bytes elapsed: 4.4477s io rate: 184186327 bytes/s
$ ./readtest /data0/dump/file 8192 1
sequential reads: 100000 of: 8192 bytes elapsed: 1.9797s io rate: 413804878 bytes/s
```

The speed seems to be limited mainly by main memory bandwidth for sequential reads and by system overheads for random reads.

```
$ ./readtest /data0/dump/file 32768 0
random reads: 25000 of: 32768 bytes elapsed: 2.0076s io rate: 408040469 bytes/s
$ ./readtest /data0/dump/file 32768 1
sequential reads: 25000 of: 32768 bytes elapsed: 1.7068s io rate: 479965034 bytes/s
```

Re: Cached file read performance

Now the difference is acceptably small. This also indicates that the system overhead for random accesses with non-large blocks is too large.

Gentoo:

```
$ ./readtest /data0/dump/file 8192 0
random reads: 100000 of: 8192 bytes elapsed: 1.2698s io rate: 645155193 bytes/s
$ ./readtest /data0/dump/file 8192 1
sequential reads: 100000 of: 8192 bytes elapsed: 1.1329s io rate: 723129371 bytes/s
```

:-(. I thought that PC133 couldn't go that fast even for a pure memory benchmark.

Bruce

freebsd-performance@xxxxxxxxxxxxx mailing list
<http://lists.freebsd.org/mailman/listinfo/freebsd-performance>

To unsubscribe, send any mail to "freebsd-performance-unsubscribe@xxxxxxxxxxxxx"