Subject: RE: [RFC][v2][patch 0/12][CFQ-cgroup]Yet another I/O bandwidth controlling subsystem for CGroups bas
Posted by Satoshi UCHIDA on Fri, 23 May 2008 02:53:50 GMT
View Forum Message <> Reply to Message

Hi, Tsuruta-san,

Thanks for your test.

>
> Uchida-san said,
>
> > In the test #2 and #3, did you use direct write?
> > I guess you have used the non-direct write I/O (using cache).
>
> I answered "Yes," but actually I did not use direct write I/O, because
> I ran these tests on Xen-HVM. Xen-HVM backend driver doesn't use direct
> I/O for actual disk operations even though guest OS uses direct I/O.
>

Where did you build expanded CFQ schedulers?
I guess that schedulers can be control I/Os if it is built on guest OS,
But not if on Dom0.
I guess you built on Dom0 so that you could not control I/O. (maybe, you say)

> So, I retested with the new testing environment and got good results.
> The number of I/Os is proportioned according to the priority levels.
>

Ok.
I'm testing both systems and get similar results.
I will report my test in next week.

> Details of the tests are as follows:
>
> Envirionment:
> Linux version 2.6.25-rc2-mm1 based.
> CPU0: Intel(R) Core(TM)2 CPU 6600 @ 2.40GHz stepping 06
> CPU1: Intel(R) Core(TM)2 CPU 6600 @ 2.40GHz stepping 06
> Memory: 2063568k/2088576k available (2085k kernel code, 23684k
> reserved, 911k data, 240k init, 1171072k highmem)
> scsi 1:0:0:0: Direct-Access ATA WDC WD2500JS-55N 10.0 PQ: 0
> ANSI: 5
> sd 1:0:0:0: [sdb] 488397168 512-byte hardware sectors (250059 MB)
> sd 1:0:0:0: [sdb] Write Protect is off

> sd 1:0:0:0: [sdb] Mode Sense: 00 3a 00 00
> sd 1:0:0:0: [sdb] Write cache: disabled, read cache: enabled,
> doesn't support DPO or FUA
> sdb: sdb1 sdb2 sdb3 sdb4 < sdb5 sdb6 sdb7 sdb8 sdb9 sdb10 sdb11
> sdb12 sdb13 sdb14 sdb15 >
>
> Procedures:
> o Prepare 3 partitions sdb5, sdb6 and sdb7.
> o Run 100 processes issuing random direct I/O with 4KB data on each
>   partitions.
> o Run 3 tests:
>   #1 issuing read I/O only.
>   #2 issuing write I/O only.
>   #3 sdb5 and sdb6 are read, sdb7 is write.
> o Count up the number of I/Os which have done in 60 seconds.
>
> Results:
>   Vasily's scheduler
>   The number of I/Os (percentage to total I/Os)
>
> ---
> | partition | sdb5 | sdb6 | sdb7 | total
> |
> | priority | 7(highest) | 4 | 0(lowest) | I/Os
> |
>
> |--------------+-------------+-------------+-------------|--------
> |
> | #1 read | 3383(35%) | 3164(33%) | 3142(32%) | 9689
> |
> | #2 write | 3017(42%) | 2372(33%) | 1851(26%) | 7240
> |
> | #3 read&write | 4300(36%) | 3127(27%) | 1521(17%) | 8948
> |
>
> ---
>
>   Satoshi's scheduler
>   The number of I/Os (percentage to total I/O)
>
> ---
> | partition | sdb5 | sdb6 | sdb7 | total
> |
> | priority | 0(highest) | 4 | 7(lowest) | I/Os
> |
>
> |--------------+-------------+-------------+-------------|--------
> |

> | #1 read | 3907(47%) | 3126(38%) | 1260(15%) | 8293
> |
> | #2 write | 3389(41%) | 3024(36%) | 1901(23%) | 8314
> |
> | #3 read&write | 5028(53%) | 3961(42%) | 441(5%) | 9430
> |
>
> ----------
>
> Thanks,
> Ryo Tsuruta

Thanks,
Satoshi UCHIDA.