Subject: RE: [RFC][v2][patch 0/12][CFQ-cgroup]Yet another I/O bandwidth controlling subsystem for CGroups bas
Posted by Satoshi UCHIDA on Tue, 27 May 2008 11:32:48 GMT
View Forum Message <> Reply to Message

Hi, Tsuruta-san.

> I'm looking forward to your report.
>

I report my tests.

My test shows following features.

 o The guaranteeing degrees are widely in write I/Os than in read I/Os for each environment.
 o Vasily's scheduler can guarantee I/O control.
    However, its guaranteeing degree is narrow.
      (in particular, at low priority)
 o Satoshi's scheduler can guarantee I/O control.
   However, guaranteeing degree is too small in write only and low priority case.
 o Write I/Os are faster than read I/Os.
   And, CFQ scheduler controls I/Os by time slice.
   So, guaranteeing degree is caused difference from estimating degree at requests level by the situation of read and write I/Os.

I'll continue testing variously.
I hope to improve I/O control scheduler through many tests.

Details of the tests are as follows:

Environment:
   Linux version 2.6.25-rc5-mm1 based.
     4 type:
        kernel with Vasily's scheduler
        kernel with Satoshi's scheduler
        Native kernel
        Native kernel and use ionice commands to each process.

   CPU0: Intel(R) Core(TM)2 CPU   6700 @ 2.66GHz stepping 6
   CPU1: Intel(R) Core(TM)2 CPU   6700 @ 2.66GHz stepping 6
   Memory: 4060180k/5242880k available (2653k kernel code, 132264k reserved, 1412k data, 356k init)
   scsi 3:0:0:0: Direct-Access   ATA   WDC WD2500JS-19N 10.0 PQ: 0 ANSI: 5
   sd 3:0:0:0: [sdb] 488282256 512-byte hardware sectors (250001 MB)
   sd 3:0:0:0: [sdb] Write Protect is off
sd 3:0:0:0: [sdb] Mode Sense: 00 3a 00 00
   sd 3:0:0:0: [sdb] Write cache: enabled, read cache: enabled, doesn't support DPO or FUA

sd 3:0:0:0: [sdb] 488282256 512-byte hardware sectors (250001 MB)
sd 3:0:0:0: [sdb] Write Protect is off
sd 3:0:0:0: [sdb] Mode Sense: 00 3a 00 00
sd 3:0:0:0: [sdb] Write cache: enabled, read cache: enabled, doesn't support DPO or FUA
sdb: sdb1 sdb2 sdb3 sdb4 < sdb5 >

Test 1:
Procedures:
- o Prepare 200 files which size is 250MB on 1 partition sdb3
- o Create 3 groups with priority 0, 4 and 7.
  - Estimate peformance in Vasily's scheduler : group-1 57.1% group-2 35.7% group-3 7.2%
  - Estimate peformance in Satoshi's scheduler : group-1 61.5% group-2 30.8% group-3 7.7%
  - Estimate peformance in Native Linux CFQ scheduler with ionice command : group-1 92.8% group-2 5.8% group-3 1.4%
- o Run many processes issuing random direct I/O with 4KB data on each files in 3 groups.
  - #1 Run 25 processes issuing read I/O only per groups.
  - #2 Run 25 processes issuing write I/O only per groups.
  - #3 Run 15 processes issuing read I/O and 15 processes issuing writeI/O only per groups.
- o Count up the number of I/Os which have done in 120 seconds.
- o Measure at 5 times in each situation.
- o Results is calculated by averatve of 5 times.

Results:

Vasily's scheduler
The number of I/Os (percentage to total I/Os)

group priority	group 1 7(highest)	group 2 4	group 3 0(lowest)	total I/Os
#1 Read	6077.2 (46.33%)	3680.8 (28.06%)	3360.2 (25.61%)	13118.2
#2 Write	10291.2 (53.13%)	5282.8 (27.27%)	3796.2 (19.60%)	19370.2
#3 Read&Write	7218.0 (49.86%)	4273.0 (29.52%)	2986.0 (20.63%)	14477.0
	(Read 45.52%)	(Read 51.96%)	(Read 52.63%)	(Read 48.89%)

Satoshi's scheduler
The number of I/Os (percentage to total I/O)

group priority	group 1 0(highest)	group 2 4	group 3 7(lowest)	total I/Os
#1 Read	9082.2 (60.90%)	4403.0 (29.53%)	1427.4 (9.57%)	14912.6
#2 Write	15449.0 (68.74%)	6144.2 (27.34%)	881.8 (3.92%)	22475.0
#3 Read&Write	11283.6 (65.35%)	4699.0 (27.21%)	1284.8 (7.44%)	17267.4
	(Read 41.08%)	(Read 47.84%)	(Read 57.07%)	(Read 44.11%)

Native Linux CFQ scheduler
The number of I/Os (percentage to total I/O)

group	group 1	group 2	group 3	total
#1 Read	4362.2 (34.94%)	3864.4 (30.95%)	4259.8 (34.12%)	12486.4
#2 Write	6918.4 (37.23%)	5894.0 (31.71%)	5772.0 (31.06%)	18584.4
#3 Read&Write	4701.2 (33.62%)	4788.0 (34.24%)	4496.0 (32.15%)	13985.2
	(Read 45.85%)	(Read 48.99%)	(Read 51.28%)	(Read 48.67%)

Native Linux CFQ scheduler with ionice command
The number of I/Os (percentage to total I/O)

group	group 1	group 2	group 3	total
priority	0(highest)	4	7(lowest)	I/Os
#1 Read	12844.2 (85.34%)	1544.8 (10.26%)	661.8 (4.40%)	15050.8
#2 Write	24858.4 (92.44%)	1568.4 (5.83%)	463.4 (1.72%)	26890.2
#3 Read&Write	16205.4 (85.53%)	2016.8 (10.64%)	725.6 (3.83%)	18947.8
	(Read 37.49%)	(Read 57.97%)	(Read 56.62%)	(Read 40.40%)

Test 2:
Procedures:
- o Prepare 200 files which size is 250MB on 1 partition sdb3
- o Create 3 groups with priority 0, 4 and 7.
- o Run many processes issuing random direct I/O with 4KB data on each files in 3 groups.
  #1 Run 25 processes issuing read I/O only in group 1 and group 2 and run 25 processes issuing write I/O only in group 3.
  (This pattern is represent by "R-R-W".)
  #2 Run 25 processes issuing read I/O only in group 1 and group 3 and run 25 processes issuing write I/O only in group 2.
  (This pattern is represent by "R-W-R".)
  #3 Run 25 processes issuing read I/O only in group 2 and group 3 and run 25 processes issuing write I/O only in group 1.
  (This pattern is represent by "R-R-W".)
- o Count up the number of I/Os which have done in 120 seconds.
- o Measure at 5 times in each situation.
- o Results is calculated by averatve of 5 times.

Results:
Vasily's scheduler

The number of I/Os (percentage to total I/Os)

group priority	group 1 7(highest)	group 2 4	group 3 0(lowest)	total I/Os
#1 R-R-W	8828.2 (52.46%)	4372.2 (25.98%)	3628.8 (21.56%)	16829.2
#2 R-W-R	5510.2 (35.01%)	6584.6 (41.83%)	3646.0 (23.16%)	15740.8
#3 W-R-R	6400.4 (41.91%)	3856.4 (25.25%)	5016.4 (32.84%)	15273.2

Results shows peculiar case in test #2.
I/O counts in group 2 are 5911, 4895, 8498, 9300 and 4319.
In third and fourth time, I/O counts are huge.
An average of first, second and fifth results is following.

group priority	group 1 7(highest)	group 2 4	group 3 0(lowest)	total I/Os
#2 R-W-R	6285.7 (41.82%)	5041.7 (33.54%)	3702.7 (24.64%)	15030.0

$B!! (B

Satoshi's scheduler
The number of I/Os (percentage to total I/O)

group priority	group 1 0(highest)	group 2 4	group 3 7(lowest)	total I/Os
#1 R-R-W	9178.6 (61.95%)	4510.8 (30.44%)	1127.4 (7.61%)	14816.8
#2 R-W-R	9398.0 (56.29%)	6152.2 (36.85%)	1146.4 (6.87%)	16696.6
#3 W-R-R	15527.0 (72.38%)	4544.8 (21.19%)	1380.0 (6.43%)	21451.8

Native Linux CFQ scheduler
The number of I/Os (percentage to total I/O)

group priority	group 1	group 2	group 3	total I/Os
#1 R-R-W	4622.4 (33.15%)	4739.2 (33.98%)	4583.4 (32.87%)	13945.0
#2 R-W-R	4610.6 (31.72%)	5502.2 (37.85%)	4422.4 (30.43%)	14535.2
#3 W-R-R	5518.0 (37.57%)	4734.2 (32.24%)	4433.2 (30.19%)	14685.4

Native Linux CFQ scheduler with ionice
The number of I/Os (percentage to total I/O)

group	group 1	group 2	group 3	total
priority	0(highest)	4	7(lowest)	I/Os
#1 Read	12619.4 (84.49%)	1537.4 (10.29%)	779.2 (5.22%)	14936.0
#2 Write	12724.4 (81.44%)	2276.6 (14.57%)	623.2 (3.99%)	15624.2
#3 Read&Write	24442.8 (91.75%)	1592.6 (5.98%)	604.6 (2.27%)	26640.0

Thanks,
Satoshi UCHIDA.