Subject: HDD Perfromance Problems (ext4 journalling?)
Posted by chr107D3o on Mon, 24 Aug 2015 12:47:17 GMT
View Forum Message <> Reply to Message

Hi,

I am running an OpenVZ DB Server (Debian 7, latest RHEL6 Kernel) and am having some issues with performance.

The server has around 8 DB instances, 5 of them postgresql the remaining mysql. The postgresql is one large DB handled by a dedicated DB container, the mysql are all local. The mysql are doing lots of consolidation for a datawarehouse and are working on rather large DBs x00GB+ (all on SSD).

HW: 24cores E5, 64GB Ram, various 1TB SSDs.

The moment one of the VZ mysql instance starts doing its consolidation work, Load Average goes up to 8.xx.
That is to be somewhat expected, when I dig into it, it start looking like a configuration issue:

htop
Load average: 7.99 7.71 7.62
CPU Avg[|||||||||||||||| 20.5%]
so the bottleneck is clearly not CPU bound

iotop
Total DISK READ: 12.03 M/s | Total DISK WRITE: 16.00 M/s
TID PRIO USER DISK READ DISK WRITE SWAPIN IO> COMMAND

2644 be/3 root 0.00 B/s 0.00 B/s 0.00 % 28.75 % [jbd2/dm-4-8]
2648 be/3 root 0.00 B/s 0.00 B/s 0.00 % 14.54 % [jbd2/dm-1-8]
587681 be/4 statd 11.69 M/s 0.00 B/s 0.00 % 9.25 % mysqld --basedir=/usr
--datadir=/var/lib/mysql.mnt --plugin-dir=/usr/lib/mysql/plugin --log-error=/var/log/mysql/error.log
--pid-file=/var/run/mysqld/mysqld.pid --socket=/var/run/mysqld/mysqld.sock --port=3306
580639 be/3 root 0.00 B/s 0.00 B/s 0.00 % 7.76 % [jbd2/ploop43140]
583993 be/3 root 0.00 B/s 33.81 K/s 0.00 % 5.15 % [jbd2/ploop27542]
39032 be/4 sshd 0.00 B/s 1337.41 K/s 0.00 % 4.58 % postgres: wal writer process
that looks more like it, especially when I take a look at the first two lines:

the jbd2/dm-1-8 & jbd2/dm-4-8 seem to be ext4 journalling and takes on average 40% of the IO load - thats a lot!
(cant post link at archlinux)

There are some posts out there (link above plus some stackexchange posts) where this is attributed to bugs in earlier kernel versions.

I could try to disable ext4 journalling, but first would like to see if someone had similar issues and how he solved it

thanks