Subject: Re: [PATCH 2/3] i/o bandwidth controller infrastructure
Posted by Andrea Righi on Mon, 30 Jun 2008 16:10:54 GMT
View Forum Message <> Reply to Message

Andrea Righi wrote:
> Andrew Morton wrote:
>> On Fri, 27 Jun 2008 00:36:46 +0200
>> Andrea Righi <righi.andrea@gmail.com> wrote:
>>
>>>> Does all this code treat /dev/sda1 as a separate device from /dev/sda2?
>>>> If so, that would be broken.
>>> Yes, all the partitions are treated as separate devices with
>>> (potentially) different limiting rules, but I don't understand why it
>>> would be broken... dev_t has both minor and major numbers, so it would
>>> be possible to select single partitions as well.
>> Well it's functionally broken, isn't it? A physical disk has a fixed
>> IO bandwidth and when the administrator wants to partition that
>> bandwidth amongst control groups he will need to consider the entire
>> device when doing so?
>>
>> I mean, the whole point of this feature and of control groups as a
>> whole is isolation. But /dev/sda1 and /dev/sda2 are very much _not_
>> isolated. Whereas /dev/sda and /dev/sdb are (to a large degree)
>> isolated.
>
> well... yes, sounds reasonable. In this case we could just ignore the
> minor number and consider only major number as the key to identify a
> specific block device (both for userspace<->kernel interface and when
> accounting/throttling i/o requests).

oops.. no, this is obviously wrong. So, I dunno if it would be better to
add complexity in cgroup_io_throttle() to identify the disk a partition
belongs or to just use the struct block_device as key, instead of dev_t,
as you intially suggested. I'll investigate.

-Andrea

_____
Containers mailing list
Containers@lists.linux-foundation.org
https://lists.linux-foundation.org/mailman/listinfo/containers