Subject: Re: Re: [RFC] Virtualization steps
Posted by Herbert Poetzl on Thu, 13 Apr 2006 13:42:39 GMT
View Forum Message <> Reply to Message

On Thu, Apr 13, 2006 at 10:52:19AM +0400, Kirill Korotaev wrote:
> Herbert,
>
> Thanks a lot for the details, I will give it a try once again. Looks
> like fairness in this scenario simply requires sched_hard settings.

hmm, not precisely, it's a cpu limit you described
and that is what this configuration does, for fair
scheduling you need to activate the indle skip and
configure it in a similar way ...

> Herbert... I don't know why you've decided that my goal is to prove
> that your scheduler is bad or not precise. My goal is simply to
> investigate different approaches and make some measurements.

fair enough ...

> I suppose you can benefit from such a volunteer, don't you think so?

well, if the 'results' and 'methods' will be made
public, I can, until now all I got was something
along the lines:

 "Linux-VServer is not stable! WE (swsoft?) have
  a secret but essential test suite running two
  weeks to confirm that OUR kernels ARE stable,
  and Linux-VServer will never pass those tests,
  but of course, we can't tell you what kind of
  tests or what results we got"

which doesn't help me anything and which, to be
honest, does not sound very friendly either ...

> Anyway, thanks again and don't be cycled on the idea that OpenVZ are
> so cruel bad guys :)

but what about the Virtuozzo(tm) guys? :)
I'm really trying not to generalize here ...

best,
Herbert

> Thanks,
> Kirill

>
> >well, your mistake seems to be that you probably haven't
> >tested this yet, because with the following (simple)
> >setups I seem to get what you consider impossible
> >(of course, not as precise as your scheduler does it)
> >
> >
> >vcontext --create --xid 100 ./cpuhog -n 1 100 &
> >vcontext --create --xid 200 ./cpuhog -n 1 200 &
> >vcontext --create --xid 300 ./cpuhog -n 1 300 &
> >
> >vsched --xid 100 --fill-rate 1 --interval 6
> >vsched --xid 200 --fill-rate 2 --interval 6
> >vsched --xid 300 --fill-rate 3 --interval 6
> >
> >vattribute --xid 100 --flag sched_hard
> >vattribute --xid 200 --flag sched_hard
> >vattribute --xid 300 --flag sched_hard
> >
> >
> > PID USER PR NI VIRT RES SHR S %CPU %MEM TIME+ COMMAND
> > 39 root 25 0 1304 248 200 R 74 0.1 0:46.16 ./cpuhog -n 1
> > 300 38 root 25 0 1308 252 200 H 53 0.1 0:34.06 ./cpuhog
> > -n 1 200 37 root 25 0 1308 252 200 H 28 0.1 0:19.53
> > ./cpuhog -n 1 100 46 root 0 0 1804 912 736 R 1 0.4
> > 0:02.14 top -cid 20
> >and here the other way round:
> >
> >vsched --xid 100 --fill-rate 3 --interval 6
> >vsched --xid 200 --fill-rate 2 --interval 6
> >vsched --xid 300 --fill-rate 1 --interval 6
> >
> > PID USER PR NI VIRT RES SHR S %CPU %MEM TIME+ COMMAND
> > 36 root 25 0 1304 248 200 R 75 0.1 0:58.41 ./cpuhog -n 1
> > 100 37 root 25 0 1308 252 200 H 54 0.1 0:42.77 ./cpuhog
> > -n 1 200 38 root 25 0 1308 252 200 R 29 0.1 0:25.30
> > ./cpuhog -n 1 300 45 root 0 0 1804 912 736 R 1 0.4
> > 0:02.26 top -cid 20
> >
> >note that this was done on a virtual dual cpu
> >machine (QEMU 8.0) with 2.6.16-vs2.1.1-rc16 and
> >that there were roughly 25% idle time, which I'm
> >unable to explain atm ...
> >
> >feel free to jump on that fact, but I consider
> >it unimportant for now ...
> >
> >best,

> >Herbert
> >
> >
> >>Thanks,
> >>Kirill
> >
> >
>