Subject: *SOLVED* disk io with vzdump / vzmigrate
Posted by goeldi on Thu, 24 May 2007 07:07:53 GMT
View Forum Message <> Reply to Message

When I make a vzdump or vzmigrate using the suspend or the online feature (which is very cool!) disk IO goes very high. It goes so high that the system is no more usable (average load more than 20!). This is a system with about 20GB of data to sync/dump.

For the suspend or online option, rsync is used. In rsync is an option --bwlimit (bandwidthlimit) which limits not only network bandwidth but disk io bandwidth too.

It would be cool to see an option like the --ssh option in vzmigrate to give it some rsync options. Or is there another way for this?

Subject: Re: disk io with vzdump / vzmigrate
Posted by rickb on Thu, 24 May 2007 08:16:11 GMT
View Forum Message <> Reply to Message

Linux itself has no way to do what you want (limit disk i/o bandwidth per process). However, the 2.6.18+ cfq scheduler supports i/o priorities, check out ionice and set the migrate pid to a low priority or the idle class.

Quote:
[root@arsenic ~]# /root/bin/ionice -h
Usage: ionice [OPTIONS] [COMMAND [ARG]...]
Sets or gets process io scheduling class and priority.

-n Class data (typically 0-7, lower being higher prio)
-c Scheduling class
1: realtime, 2: best-effort, 3: idle
-p Process pid
-h This help page

Jens Axboe <axboe@suse.de> (C) 2005

your disk i/o could be full, however you will not feel it as all other processes have significant priority. the vz tools do not set io priority yet directly, however you can re-ionice the process after it has been started.

good question!

Subject: Re: disk io with vzdump / vzmigrate
Posted by Vasily Tarasov on Fri, 25 May 2007 11:20:03 GMT

View Forum Message <> Reply to Message

Hello,

I want to ask you some questions:

1. Kernel you're using

2. You write "disk IO goes very high". Can you be more precise? Please, post tools output and other appropriate stuff.

3. You write "...the system is no more usable". What do you mean? The response of the system is very long? Can you give some certain examples?

4. You write "average load more than 20!". Sorry for my ignorance, where do you get this digit (20) and what it means?

Thank you,
Vasily.

Subject: Re: disk io with vzdump / vzmigrate
Posted by goeldi on Fri, 25 May 2007 12:03:29 GMT
View Forum Message <> Reply to Message

Vasily Tarasov wrote on Fri, 25 May 2007 07:20Hello,

I want to ask you some questions:

1. Kernel you're using

2.6.18-ovz028stab031.1-smp

Quote:2. You write "disk IO goes very high". Can you be more precise? Please, post tools output and other appropriate stuff.

This was only my interpretation.

Quote:3. You write "...the system is no more usable". What do you mean? The response of the system is very long? Can you give some certain examples?

Last night I made a vzdump of about 10GB in 6 VEs. Average system load (see it with command w or with top) went to over 50.0 for about 2 hours. I hid Ctrl-C and it took about half an hour to stop vzdump.

Quote:4. You write "average load more than 20!". Sorry for my ignorance, where do you get this digit (20) and what it means?

output of top or w
I see it in a gkrellm client too.

At the same time, cpu usage is about 0% - 1%

Subject: Re: disk io with vzdump / vzmigrate
Posted by Vasily Tarasov on Fri, 25 May 2007 12:16:29 GMT
View Forum Message <> Reply to Message

There was quite a big bug in disk I/O that causes high I/O load in our tests and it was fixed only in 032 kernel. Can you please try to apply this patch
http://git.openvz.org/?p=linux-2.6.18-openvz;a=commit;h=c7fa58b331d23c0796996dcab75c72053fafab5a
and then inform us, does it help?

Thank you,
Vasily.

Subject: Re: disk io with vzdump / vzmigrate
Posted by goeldi on Tue, 29 May 2007 06:30:46 GMT
View Forum Message <> Reply to Message

I cannot find ionice tool anywhere. Where do i get it?

Does you comment mean, that I may upgrade to kernel 2.6.20 to get this option?

Subject: Re: disk io with vzdump / vzmigrate
Posted by goeldi on Tue, 29 May 2007 06:37:45 GMT
View Forum Message <> Reply to Message

this is a productive system. I would prefer not to apply patches to the kernel[/coward]

BTW: Last night I tried to start rsync as a daemon on the target system with option --bwlimit=100, but now I have another problem (http://forum.openvz.org/index.php?t=tree&goto=13511&#msg_13511).

Would it be a possibility to start a bwlimited rsyncd on the target system? Does vzmigrate take notice of a rsync daemon?

And another question: I tried to first make a normal bwlimited rsync from source to target system. I made this because I thought that after that vzmigrate would sync only the changes. But obviously vzmigrate always transfers everything, because it ended with the error message, that the target directory already exists

---

Subject: Re: disk io with vzdump / vzmigrate
Posted by curx on Tue, 29 May 2007 07:10:23 GMT
View Forum Message <> Reply to Message

> I cannot find ionice tool anywhere. Where do i get it?

it's in package "schedutils"

Subject: Re: disk io with vzdump / vzmigrate
Posted by goeldi on Tue, 29 May 2007 07:30:56 GMT
View Forum Message <> Reply to Message

In schedutils I have 2 tools: chrt and taskset. So if I do this while vzmigrating:

chrt -p 99 <pid>

where <pid> is the pid of the rsync process, rsync on the target side will get a priority of 99?

Can you suggest me, where I get more USER information for priorities? Is it comparable with nice (highnumber=slow, lownumber=fast) and what is the highest/lowest priority I can give?

Sorry for these newbie questions, but nobody else seems to ask this :-/

Subject: Re: disk io with vzdump / vzmigrate
Posted by Vasily Tarasov on Tue, 29 May 2007 10:31:45 GMT
View Forum Message <> Reply to Message

If you have linux sources read about ionice tool in Documentation/block/ioprio.txt file.

HTH,
Vasily.

Subject: Re: disk io with vzdump / vzmigrate
Posted by Vasily Tarasov on Fri, 01 Jun 2007 07:10:46 GMT
View Forum Message <> Reply to Message

I close this thread as *SOLVED*, since initial problem is solved in 033 kernel (look at http://bugzilla.openvz.org/show_bug.cgi?id=602)

If you have other problems - please - post them as a separate thread.

Thank you,
Vasily.