Subject: Re: Re: [Announce] Kernel RHEL6 testing 042stab054.1
Posted by jjs - mainphrame on Sat, 07 Apr 2012 22:48:23 GMT
View Forum Message <> Reply to Message

Thanks for the pointer to the article, that's good info. I've checked my system, and I'm nowhere near the limit of space or inodes.

To further test, I create a ploop CT which contains the expected amount of disk space.
I then create a simfs CT with the same disk size settings, and it only has half the expected disk size.
I then create another ploop CT and it contains the expected amount of disk space.

If the 2nd CT which I created failed to get the requested disk space due to shortage on the system, then it's difficult to see how the 3rd CT could then get the full disk space requested. So there seems to be something funny going on with the disk size calculation of simfs CTs in stab54.2.

Joe

On Fri, Apr 6, 2012 at 1:49 PM, Kirill Kolyshkin <kolyshkin@gmail.com>wrote:

> This probably means your /vz partition has less space than the limit you
> set. There's an article on wiki explaining that in details, let me see...
> right, http://wiki.openvz.org/Disk_quota,_df_and_stat_weird_behavio ur

<jjs@mainphrame.com>

>
> Something definitely weird happening with simfs file sizes now:
>>
>> [root@mrmber ~]# vzctl set 777 --save --diskspace="20000000:24000000"
>> CT configuration saved to /etc/vz/conf/777.conf
>> [root@mrmber ~]# vzctl exec 777 df
>> Filesystem   1K-blocks   Used Available Use% Mounted on
>> /dev/simfs   5474372   710700   3205452 19% /
>> none   131072   4   131068   1% /dev
>> [root@mrmber ~]#
>>
>> ploop-based CTs seem fine.
>>
>> Joe
>>
>> On Thu, Apr 5, 2012 at 11:24 PM, jjs - mainphrame <jjs@mainphrame.com>wrote:
>>
>>> Look closer - there is breakage here. Normally there was a 10%
>>> difference between simfs and ploop, but this is different - this simfs CT

>>> has only 1/3 the advertised disk space...
>>>
>>> Joe
>>>
>>>
>>> On Thu, Apr 5, 2012 at 11:06 PM, Kirill Korotaev <dev@parallels.com>wrote:
>>>
>>>> Note, that ploop contains ext4 inode tables also (which are
>>>> preallocated by ext4), so ext4 reserves some space for its own needs.
>>>> Simfs however was limiting *pure* file space.
>>>>
>>>> Kirill
>>>>
>>>> On Apr 6, 2012, at 04:58 , jjs - mainphrame wrote:
>>>>
>>>> > However I am seeing an issue with the disk size inside the
>>>> simfs-based CT.
>>>> >
>>>> > In the vz conf files, all 3 CTs have the same diskspace setting:
>>>> >
>>>> > [root@mrmber ~]# grep -i diskspace /etc/vz/conf/77*conf
>>>> > /etc/vz/conf/771.conf:DISKSPACE="20000000:24000000"
>>>> > /etc/vz/conf/773.conf:DISKSPACE="20000000:24000000"
>>>> > /etc/vz/conf/775.conf:DISKSPACE="20000000:24000000"
>>>> >
>>>> > But in the actual CTs the one on simfs reports a significantly
>>>> smaller disk space than it did under previous kernels:
>>>> >
>>>> > [root@mrmber ~]# for i in `vzlist -1`; do echo $i; vzctl exec $i df;
>>>> done
>>>> > 771
>>>> > Filesystem 1K-blocks Used Available Use% Mounted on
>>>> > /dev/ploop0p1 23621500 939240 21482340 5% /
>>>> > none 262144 4 262140 1% /dev
>>>> > 773
>>>> > Filesystem 1K-blocks Used Available Use% Mounted on
>>>> > /dev/simfs 6216340 739656 3918464 16% /
>>>> > none 262144 4 262140 1% /dev
>>>> > 775
>>>> > Filesystem 1K-blocks Used Available Use% Mounted on
>>>> > /dev/ploop1p1 23628616 727664 21700952 4% /
>>>> > none 262144 4 262140 1% /dev
>>>> > [root@mrmber ~]#
>>>> >
>>>> > Looking in dmesg shows this:
>>>> >
>>>> > [2864.563423] CT: 773: started
>>>> > [2866.203628] device veth773.0 entered promiscuous mode

>>>> > [2866.203719] br0: port 3(veth773.0) entering learning state
>>>> > [2868.302300] ploop1:
>>>> > [2868.329086] GPT:Primary header thinks Alt. header is not at the
>>>> end of the disk.
>>>> > [2868.329099] GPT:47999999 != 48001023
>>>> > [2868.329104] GPT:Alternate GPT header not at the end of the disk.
>>>> > [2868.329111] GPT:47999999 != 48001023
>>>> > [2868.329115] GPT: Use GNU Parted to correct GPT errors.
>>>> > [2868.329128] p1
>>>> > [2868.333608] ploop1:
>>>> > [2868.337235] GPT:Primary header thinks Alt. header is not at the
>>>> end of the disk.
>>>> > [2868.337247] GPT:47999999 != 48001023
>>>> > [2868.337252] GPT:Alternate GPT header not at the end of the disk.
>>>> > [2868.337258] GPT:47999999 != 48001023
>>>> > [2868.337262] GPT: Use GNU Parted to correct GPT errors.
>>>> >
>>>> > I'm assuming that this disk damage occurred under the buggy stab54.1
>>>> kernel. I could destroy the container and create a replacement but I'd like
>>>> to make believe, for the time being, that it's valuable. Just out of
>>>> curiosity, what tools exist to fix this sort of thing? The log entries
>>>> recommend gparted, but I suspect I may not have much luck from inside the
>>>> CT with that. If this were PVC, there would obviously be more choices. You
>>>> thoughts?
>>>> >
>>>> > Joe
>>>> >
>>>> > On Thu, Apr 5, 2012 at 3:17 PM, jjs - mainphrame <jjs@mainphrame.com>
>>>> wrote:
>>>> > I'm happy to report that stab54.2 fixes the kernel panics I was
>>>> seeing in stab54.1 -
>>>> >
>>>> > Thanks for the serial console reminder, I'll work on setting that
>>>> up...
>>>> >
>>>> > Joe
>>>> >
>>>> > On Thu, Apr 5, 2012 at 3:47 AM, Kir Kolyshkin <kir@openvz.org> wrote:
>>>> > On 04/05/2012 08:48 AM, jjs - mainphrame wrote:
>>>> > Kernel stab53.5 was very stable for me under heavy load but with
>>>> stab54.1 I'm seeing hard lockups - the Alt-Sysrq keys don't work, only the
>>>> power or reset button will do the trick.
>>>> >
>>>> > I don't have a serial console set up so I'm not able to capture the
>>>> kernel panic message and backtrace. I think I'll need to get that set up in
>>>> order to go any further with this.
>>>> >
>>>> > 054.2 might fix the issue you are having. It is being uploaded at

>>>> the moment...
>>>> >
>>>> > Anyway, it's a good idea to have serial console set up. It greatly
>>>> improves chances to resolve kernel bugs.
>>>> http://wiki.openvz.org/Remote_console_setup just in case.
>>>> > <ATT00001.c>
>>>>
>>>>