Subject: Process Ids
Posted by tscofield on Fri, 30 Apr 2010 03:25:18 GMT
View Forum Message <> Reply to Message

On my hardware node I have set the pid_max > 32768 in oder to support a large number of virtual machines. I am running into an issue with some older applications that are failing when the pid they have is > 32768. Is there a way to force the VE to only use PID's < 32768 while still having an elevated pid_max on the hardware node?

Subject: Re: Process Ids
Posted by maratrus on Fri, 30 Apr 2010 12:42:55 GMT
View Forum Message <> Reply to Message

Hello,

could you please specify what kind of applications are falling with pids > 32768?

Subject: Re: Process Ids
Posted by tscofield on Fri, 30 Apr 2010 13:40:42 GMT
View Forum Message <> Reply to Message

We initially set the pid_max on the hw node to 128k. We were seeing two issues with that setting. One of the errors was either with sudo or ps.
$sudo ps -ef |grep 105753
stat: Value too large for defined data type
$

The other error is with a commercial application, BEA's Tuxedo while trying to run the tmshutdown command
tmshutdown: internal error: CMDTUX_CAT:766: ERROR: must run on master node

We have changed the pid_max to 64k and we haven't seen the issue with sudo or ps, but the error with tmshutdown persists. If we drop the pid_max back down to 32k the tmshutdown command works as expected.

Our hardware node is running RedHat Enterprise 5.3 and the 2.6.18-128.1.1.el5.028stab062.3 kernel
The VE is running RedHat Enterprise 3.9, below is the config

NUMPROC="8192:8192"
AVNUMPROC="41212:41212"
NUMTCPSOCK="82424:82424"
NUMOTHERSOCK="82424:82424"
VMGUARPAGES="3046400:9223372036854775807"

```
# Secondary parameters
KMEMSIZE="3376091955:3713701150"
TCPSNDBUF="787755281:1125363985"
TCPRCVBUF="787755281:1125363985"
OTHERSOCKBUF="393877640:731486344"
DGRAMRCVBUF="393877640:393877640"
OOMGUARPAGES="3046400:9223372036854775807"
PRIVVMPAGES="3046400:3145728"

# Auxiliary parameters
LOCKEDPAGES="164848:164848"
SHMPAGES="989089:989089"
PHYSPAGES="0:9223372036854775807"
NUMFILE="1318784:1318784"
NUMFLOCK="1000:1100"
NUMPTY="512:512"
NUMSIGINFO="1024:1024"
DCACHESIZE="737494741:759619584"
NUMIPTENT="200:200"
DISKSPACE="104857600:104857600"
DISKINODES="29301100:32231211"
CPUUNITS="900"
FEATURES="nfs:on "
IP_ADDRESS="10.152.106.39"
VE_ROOT="/vz/root/$VEID"
VE_PRIVATE="/vz/private/$VEID"
CPUS="2"
```

Subject: Re: Process Ids
Posted by maratrus on Fri, 30 Apr 2010 17:47:21 GMT
View Forum Message <> Reply to Message

Hello,

I attached a patch against ovzkernel-2.6.18-164.15.1.el5.028stab068.9 kernel.

Could you please test it with your application?

File Attachments
1) diff-ve-virt-pidmax-20100430, downloaded 456 times

Subject: Re: Process Ids

Posted by tscofield on Tue, 04 May 2010 21:14:04 GMT
View Forum Message <> Reply to Message

We are still getting the error and the process ID's within the VE is still going above 32k. Are there any parameters that need to be set on the VE config file to limit the pid_max?

Subject: Re: Process Ids
Posted by maratrus on Wed, 05 May 2010 14:02:35 GMT
View Forum Message <> Reply to Message

Hello,

I suppose you to change pid_max via sysctl pid_max (/proc/sys/kernel/pid_max). By default, pid_max is the same
as on the HN but you can change it writing to the file mentioned above.
Does it suit you?

Subject: Re: Process Ids
Posted by tscofield on Wed, 05 May 2010 16:21:14 GMT
View Forum Message <> Reply to Message

Running the "sysctl -w kernel.pid_max=32768" command in the VE seems to have solved the problem. Is this a parameter that could be put in the config file or would it need to be part of the VE start script.

Subject: Re: Process Ids
Posted by maratrus on Thu, 06 May 2010 08:50:57 GMT
View Forum Message <> Reply to Message

Hello,

this parameter is not a part of a VE's configuration file.
You'd better put it in VE's start script.

Subject: Re: Process Ids
Posted by tscofield on Thu, 06 May 2010 15:12:50 GMT
View Forum Message <> Reply to Message

Thanks. When would you expect these patches to be incorporated into the stable kernel?

Subject: Re: Process Ids
Posted by maratrus on Fri, 07 May 2010 12:48:37 GMT
View Forum Message <> Reply to Message

Could you please answer one more question?

Quote:
$sudo ps -ef |grep 105753
stat: Value too large for defined data type

Do you run 32-bit VEs on the 64bit HN?

Quote:
Thanks. When would you expect these patches to be incorporated into the stable kernel?

Look, the feature you want to be in the kernel is not a widely demanded one. Frankly speaking, I cannot remember the case when it was needed. BUT, that doesn't mean that it will not be included in OpenVZ kernel. First of all. I suggest you to file a new bug report and describe the problem you have. You can also attach the patch I proposed you to use. It helps maintainers to include the feature much faster. Another point I want to note is that if you are running 32-bit VEs on the 64-bit HN and ps(!) complains about huge PID then such a bug MUST BE fixed. And dynamically adjusted pid_max values is a possible solution.

Subject: Re: Process Ids
Posted by tscofield on Fri, 07 May 2010 12:52:48 GMT
View Forum Message <> Reply to Message

Yes we are running 32 bit VE on a 64 bit HN. I will go ahead and open a bug as you suggested.
Thanks.