Subject: How does OpenVZ manage memory?
Posted by rsauvat on Mon, 23 Aug 2010 14:50:31 GMT
View Forum Message <> Reply to Message

Hi ,

I have a question regarding how the memory is processed in openvz or maybe it is just on Linux in general.

I am running about 10 containers on a host. All containers are web servers running the same distribution (Gentoo) and even the exact sames binaries, as I have mount (with bind option) the binary and library directory from one VM to the others. So I have the same containers except for user data (mostly mysql databases). I have remarked that if I add all the memory used by each vm and compare it to the used memory in the physical host it is almost twice a big. If I add all used memory in each vm I get 8098 Mo. The free command without buffers and cache report 3687Mo.

My guess is that Linux is sharing the memory for all processes so when a process require an already loaded library or binary It doesn't really take more space in RAM. When inside a container, it doesn't know about the shared memory space and display the memory usage as if the processes in the container where the only ones.

If anyone can infirm or confirm my guess it would be a great help was I am trying to understand how memory is used.

I have an other question as well. When a container is set to use for example 5Go of RAM and 1Go of Swap. If this container uses 5,5Go of memory and the host as enough free ram for the 0,5Go, does the 0,5Go get allocated in ram or in swap space?

If any as the knowledge to answer my questions it would be great.

Regards.

Subject: Re: How does OpenVZ manage memory?
Posted by mike on Wed, 11 May 2011 09:01:30 GMT
View Forum Message <> Reply to Message

You might want to read up on mmap syscall and to google for "copy on write" approach; lwn.net articles well might be preferred in search results.

BTW one can use -o bind,ro with 2.6.32 so that getting root and compromising a binary in a container won't get all the rest compromised -- you might want to either bindmount off the "maintenance" VE which won't be accessible from network (except from your host or even by vzctl enter only), or prepare a chroot which isn't running as a VE altogether (it might be less convenient with Gentoo).