Subject: Re: Container Test Campaign
Posted by kir on Mon, 03 Jul 2006 08:14:55 GMT
View Forum Message <> Reply to Message

Clement,

Thanks for sharing the results! A few comments...

(1) General

1.1 It would be nice to run vmstat (say, vmstat 10) for the duration of the tests, and put the vmstat output logs to the site.

1.2 Can you tell how you run the tests. I am particularly interested in
- how many iterations do you do?
- what result do you choose from those iterations?
- how reproducible are the results?
- are you rebooting the box between the iterations?
- are you reformatting the partition used for filesystem testing?
- what settings are you using (such as kernel vm params)?
- did you stop cron daemons before running the test?
- are you using the same test binaries across all the participants?
- etc. etc...

Basically, the detailed description of a process would be nice to have, in order to catch possible problems. There are a lot of tiny things which are influencing the results. For example, in linux kernels 2.4 binding the NIC IRQ to a single CPU on an SMP system boosts network performance by about 15%! Sure this is not relevant here, it's just an example.

1.3 Would be nice to have diffs between different kernel configs.

(2) OpenVZ specifics

2.1 Concerning the tests running inside an OpenVZ VE, the problem is there is a (default) set of resource limits applied to each VE. Basically one should tailor those limits to suit the applications running, OR, for the purpose of testing, just set those limits to some very high values so they will never be reached.

For example, the tbench test is probably failed to finish because it hits the limits for privvmpages, tcpsndbuf and tcprcvbuf. I have increased the limits for those parameters and the test was finished successfully. Also, dbench test could hit the disk quota limit for a VE.

Some more info is available at http://wiki.openvz.org/Resource_management

2.2 For OpenVZ specifically, it would be nice to collect
/proc/user_beancounters output before and after the test.

Clément Calmels wrote:
> Hi,
>
> A first round about virtualisation benchmarks can be found here:
> http://lxc.sourceforge.net/bench/
> These benchmarks run with vanilla kernels and the patched versions of
> well know virtualisation solutions: VServer and OpenVZ. Some benchs also
> run inside the virtual 'guest' but we ran into trouble trying to run
> some of them... probably virtual 'guest' configuration issues... we will
> trying to fix them...
> The metacluster migration solution (formely a Meiosys company produt)
> was added as it seems that the checkpoint/restart topic is close to the
> virtualisation's one (OpenVZ now provides a checkpoint/restart
> capability).
> For the moment, benchmarks only ran on xeon platform but we expect more
> architecture soon. Besides the 'classic' benchs used, more network
> oriented benchs will be added. Netpipe between two virtual 'guests' for
> example. We hope we will be able to provide results concerning the
> virtual 'guest' scalability, running several 'guest' at the same time.
>
> Best regards,
>
>
> Le mercredi 07 juin 2006 à 16:20 +0200, Clement Calmels a écrit :
>
>> Hello !
>>
>> I'm part of a team of IBMers working on lightweight containers and we
>> are going to start a new test campaign. Candidates are vserver,
>> vserver context, namespaces (being pushed upstream), openvz, mcr (our
>> simple container dedicated to migration) and eventually xen.
>>
>> We will focus on the performance overhead but we are also interested in
>> checkpoint/restart and live migration. A last topic would be how well
>> the
>> resource managment criteria are met, but that's extra for the moment.
>>
>> We plan on measuring performance overhead by comparing the results on
>> a vanilla kernel with a partial and with a complete virtual
>> environment. By partial, we mean the patched kernel and a 'namespace'
>> virtualisation.
>>
>> Test tools
>> ----------
>> o For network performance :

>>
>> * netpipe (http://www.scl.ameslab.gov/netpipe/)
>> * netperf (http://www.netperf.org/netperf/NetperfPage.html)
>> * tbench (http://samba.org/ftp/tridge/dbench/README)
>>
>> o Filesystem :
>>
>>   * dbench (http://samba.org/ftp/tridge/dbench/README)
>>   * iozone (http://www.iozone.org/)
>>
>> o General
>>
>>   * kernbench (http://ck.kolivas.org/kernbench/) stress cpu and
>>   filesystem through kernel compilation
>>   * More 'real world' application could be used, feel free to submit
>>   candidates...
>>
>> We have experience on C/R and migration so we'll start with our own
>> scenario, migrating oracle under load. The load is generated by DOTS
>> (http://ltp.sourceforge.net/dotshowWe ran into trouble trying to run sto.php).
>>
>> If you could provided us some material on what has already been done :
>> URL, bench tools, scenarios. We'll try to compile them in. configuration
>> hints and tuning are most welcome if they are reasonable.
>>
>> Results, tools, scenarios will be published on lxc.sf.net . We will
>> set up the testing environment so as to be able to accept new
>> versions, patches, test tools and rerun the all on demand. Results,
>> tools, scenarios will be published on lxc.sf.net.
>>
>> thanks !
>>
>> Clement,
>>