

Re: Response issues on GS1280, VMS 7.3-2

Source: <http://unix.derkeiler.com/Newsgroups/comp.os.vms/2005-07/0825.html>

From: Keith Parris (keithparris_NOSPAM_at_yahoo.com)

Date: 07/13/05

Date: Wed, 13 Jul 2005 15:14:32 GMT

Lee wrote:

> *Five-node Gigabit Ethernet VMScLuster across three sites.*

Do you have LAVC\$FAILURE_ANALYSIS in place? This would help you determine if transient network problems as a contributing factor.

For a baseline measurement in multi-site clusters, I usually like to run LOCKTIME.COM from [KP_CLUSTERTOOLS] on the V6 Freeware CD to get the inter-site link latencies. Do you have results from that?

> *Third party SAN disk environment.*

Do you see a lot of disk mount verifications?

> *Early this year, the cluster was at VMS 7.3-2 running on four individual ES45's. Since migrating from the ES45 nodes to four nodes on two GS1280's in May-2005, interactive users have been experiencing intermittent several-second periods of slow response. The situation is occurring on all four production nodes. Symptoms are more pronounced and wide-spread during peak periods (mid-morning, noon). I myself notice occasional lags of several seconds after entering a command in DCL.*

Is this just any arbitrary DCL command, or something dealing with specific file(s) or disk(s)? If it's any arbitrary DCL command, then maybe a CPU shortage is involved, including things like saturation of the primary CPU in interrupt state (\$MONITOR MODES/ALL could help check for that).

The key in this sort of situation is to find out what a process is waiting on while it's hung. Looking at process states is a good start: CPU-bound processes tend to show up in CUR or COM state; various resource-wait states can also be informative. I also find it useful to look for lock queues, as locks tend to be held across I/O operations so any slow I/O devices tend to build up lock queues. (Lock queues can be detected using Availability Manager / DECamds and their Lock Contention data gathering facility, or LCKQUE from [KP_LOCKTOOLS] on the V6 Freeware CD.) I saw a post from someone here not long ago who had an SDA

extension which could determine what a process was waiting on — that would be handy.

Do all 4 production nodes run the same application mix? Have you looked into the possibility of remastering of large lock trees as a contributing factor for the hangs? \$MONITOR RLOCK lets you view lock tree remastering rates.

> *OBSV #2 No resource hogs have been found on any of the nodes.*

OK.

> *OBSV #3 Sequential snapshots of the nodes show many processes in/out of MUTEX. The processes in MUTEX range widely, from OPCOM to production users. These processes slide in and out of MUTEX so quickly that there is inadequate time to determine the reason for the MUTEX state.*

The SDA extension MTX can be used for mutex tracing.

> *23D7E642 _TNA4059: SUSP 0 3636 0 00:03:51.03 4264*

Any idea what suspended this interactive process?

> *OpenVMS V7.3-2 on node D 12-JUL-2005 14:29:25.56 Uptime 38 21:10:50*
> *Pid Process Name State Pri I/O CPU Page flts*
> *Pages*
> *23F8CA29 _TNA866: RWSCS 4 8664 0 00:00:07.98 4622*

RWSCS often indicates a process waiting for a lock request.

> *OBSV #4 HP has identified one main problem as being in logical name translation.*
> *Here's the status from the four nodes (from MONITOR IO).*
> *CUR AVE MIN MAX*
> *Log Name Translation Rate 198.66 906.97 0.00 9845.33*
> *Log Name Translation Rate 3902.00 3896.64 0.00 15286.00*
> *Log Name Translation Rate 2077.00 1341.27 0.00 13067.33*
> *Log Name Translation Rate 1690.66 621.39 0.00 3901.33*

The SDA extension LNM can be used for logical name translation tracing.

> *On the ES45's, I could execute a procedure containing 1000 logical name translations in a split second. On the GS1280 nodes, the same procedure requires from several to 10 seconds.*

ES45s can have at most 4 CPUs and their path to memory is quite short, so the memory subsystem is quite fast. GS1280s can scale to many more CPUs and while the EV7 on-chip memory interface makes its memory subsystem amazingly fast for that size/scale of system, some types of memory operations are going to be faster on the ES45.

> *OBSV #5 I ran Autogen with feedback and a couple of items stood out.*

...

> *MSCP_BUFFER parameter information:*

> *Feedback information.*

> *Old value was 1300, New value is 1300*

> *MSCP server I/O rate: 367 I/Os per 10 sec.*

> *I/Os that waited for buffer space: 1564*

> *I/Os that fragmented into multiple transfers: 3276*

You have plenty of memory, and it appears you may be using VMS MSCP
Serving to access disks at each of the two main sites from the opposite
site, so it certainly wouldn't hurt to add a MIN_MSCP_BUFFER line in
MODPARAMS.DAT to raise MSCP_BUFFER and eliminate the need for requests
to wait for buffer space or be fragmented into multiple transfers.