KBA1129 How to Generate a non-maskable interrupt NMI in VMware ESX

Symptoms

The symptoms of a server hang are that the system is not reachable via NFS, the GUI, SSH, console logins, etc. and that the hypervisor still indicates the virtual machine (VM) is running. The server may respond to ping depending on the nature of the hang. Should such a condition arise where the system is otherwise unreachable, a non-maskable interrupt (NMI) may be sent from the hypervisor to cause the Delphix operating system (DxOS) to kernel panic and generate a crash dump. The resulting crash dump can be collected by Delphix Support for further analysis.

If the system does not respond, retry the procedure. The final recourse is to reset or power on/off the system which will not generate a core and reduce potential for root cause analysis.

It is important to note that this procedure will not be successful in all cases. VM hangs may occur for a variety of reasons related to the guest operating system, or ESX hypervisor issues. The following procedure is a best-effort to collect system state information at the time of a VM hang.

Procedure

VMware ESX 5.x, 6.x

1. Login to ESX via SSH

2. At the prompt use the command `esxcli vm process list` to get the list of VMs and record the world id

3. Once the world ID is obtained, execute the following command to initiate the NMI: "vmdumper <world id> nmi". In the example below, the Delphix Engine name is known to be "example5023", which can be used with "grep" to reduce output for parsing.

Example of esxcli and vmdumper commands

```
~ # esxcli vm process list | grep -A 10 example
example5023
   World ID: 5678754
   Process ID: 0
   VMX Cartel ID: 5678753
   UUID: 56 4d 67 e6 38 d9 70 27-b9 06 56 4c 77 a9 5b 9d
   Display Name: delphix5023
   Config File: /vmfs/volumes/6c25682a-d47ef09e/dlpx-5.0.2.3-55/dlpx-5.0.2.3-
```
No output beyond "Sending NMI to guest..." is expected. The command prompt should typically return within a few seconds.

4. Other methods are detailed in VMware KB article How to send NMI to Guest OS on ESXi 6.x (2149185).

---

VMware ESX 4

NMIs can only be sent on ESX from the SSH command-line. First attempt to connect to the ESX system via SSH; if that fails, enable SSH using the following sequence of steps:

1. On the ESX system's console: Press ALT-F1 and a console log should be displayed. Type 'unsupported' to access the VMware "Tech Support Mode". The text entered will not be visible.

2. A password prompt should appear and the root password may be entered to gain CLI access.

3. Edit (vi) /etc/inetd.conf, search for line beginning with #ssh, and uncomment the line

```
~ # vi /etc/inetd.conf
```

4. Identify the inetd process id (pid) using the command ps | grep inetd
   
   ```
   ~ # ps | grep inetd
   1541 1541 busyboxinetd
   ```

5. Restart the inetd process using the pid obtained in previous step using kill -HUP <pid>
   
   ```
   ~ # kill -HUP 1541
   ```

To send the NMI, do the following:

1. Login to ESX via SSH

2. At the ESX command prompt use the command vm-support -x to get the list of VMs and note the vmid belonging to the Delphix Engine
   
   "vm-support -x" example

```
[root@esxserver ~]# vm-support -x
VMware ESX Support Script 1.30
Available worlds to debug:
vmid=4305RHEL Oracle Source
```
Once the vmid is collected from the previous step, use vmdumper to generate NMI:

```
Generate NMI with "vmdumper"
[root@esxserver ~]# vmdumper 4308 nmi
```

Advice from VMware

VMware has a number of knowledge base articles of their own relating to diagnosing unresponsive VMs and generating NMIs for those systems. The following articles are especially relevant to the processes discussed above.

- Troubleshooting a virtual machine that has stopped responding
- Using Tech Support Mode in ESXi 4.1, 5.x, 6.x
- Using Tech Support Mode in ESXi 4.0, 3.5

Monitoring Delphix VM Console

During the NMI process, it is helpful to observe (and record if possible) the Delphix VM console behavior to confirm the NMI is received, and to observe the DxOS panic and reboot. Details of the output may vary but the DxOS panic indicated with "NMI received" is expected in every instance if the operation is successful. If the console indicates the dump has reached 100%, the desired diagnostic information should be retrievable.

Example Delphix VM console output

```
panic[cpu0]/thread=ffffff000b805c40: NMI received
ffffff000b805aa0 fffffff000b805b7f ()
ffffff000b805ad0 unix:av_dispatch_nmivect+34 ()
ffffff000b805ae0 unix:nmiint+152 ()
ffffff000b805be0 unix:mach_cpu_idle+6 ()
ffffff000b805c00 unix:cpu_idle+11a ()
ffffff000b805c20 unix:thread_start+8 ()
syncing file systems... done
dumping to /dev/zvol/dsk/rpool/dump, offset 65536, content: kernel
0:03 100% done
100% done: 119831 pages dumped, dump succeeded
rebooting...
```

If any issue is encountered during the DxOS panic dump, the process may take up to 2 hours to time out. If the "NMI received" activity is NOT observed in the console, the Engine may still be rebooted but this indicates diagnostic information may not be generated as the NMI was not registered by the Engine. If you are receiving messages on the console and it stops responding for a long period of time (>30m), particularly at any point during syncing file systems or incrementing the amount of pages dumped, it is likely that it will not complete. You can observe the VM from the hypervisor to try to make a determination of whether there's enough activity to warrant giving it additional time or proceed with resetting the VM.