CMPE283: Virtualization
Assignment 1: Discovering VMX Features

Name: Utsav Rawat
(SJSU ID: 016664466)

Question:

Describe in detail the steps you used to complete the assignment. Consider your reader to be someone skilled in software development but otherwise unfamiliar with the assignment. Good answers to this question will be recipes that someone can follow to reproduce your development steps.

Answer:

Used Nested Virtualization in Google Cloud Platform (GCP), Nested virtualization lets us to run virtual machine (VM) instances inside of other VMs so that we can create our own virtualization environments.
Compute Engine VMs run on a physical host that has Google's security-hardened, KVM-based hypervisor. With nested virtualization, the physical host and its hypervisor are the level 0 (L0) environment. The L0 environment can host multiple level 1 (L1) VMs. On each L1 VM is another hypervisor, which is used to install the level 2 (L2) VMs

The following steps followed to develop and test the kernel module:

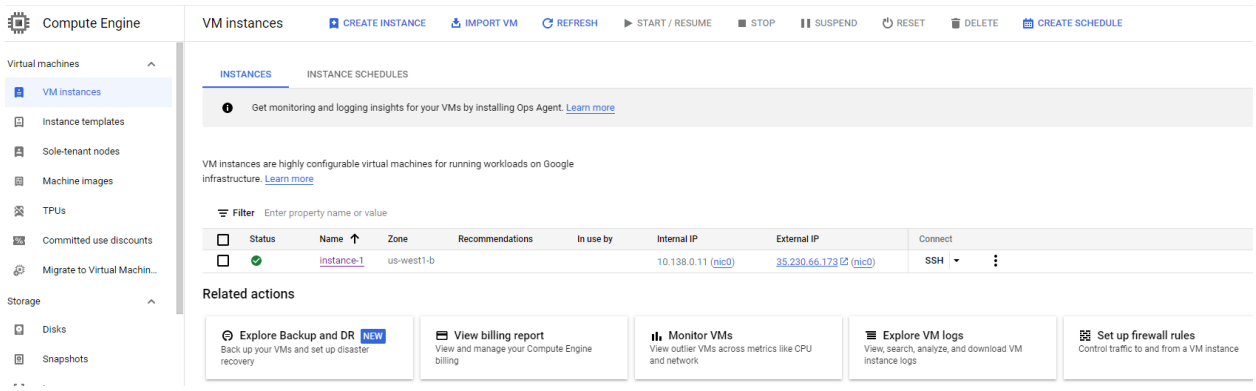
1. Create a GCP Free Tier Account which provides $300 USD Free credit to try various GCP Services.
2. You need a Credit/Debit card with international transactions enabled for verification purposes only. They will charge $1 USD while registration and then refund it.
3. Create a boot disk from a public image or from a custom image.

4. Open the Cloud Shell by clicking on the icon at the top right corner.

5. Create a custom image with the special license key that is required for nested virtualization.

   *gcloud compute images create IMAGE_NAME *
   *--source-disk DISK_NAME *
   *--source-disk-zone ZONE *
   --licenses
   https://www.googleapis.com/compute/v1/projects/vm-options/global/licenses/enable-vmx

6. Optional, delete the source disk after creating the image with the special license.
7. Create a VM that uses the new image with the special license.
   i. Name: For example “instance-1”.
   ii. Region: Select the Region where the custom image was created. example "us-west1-b"
   iii. Machine Family: General Purpose
   iv. Series: Select "N2" for lab
   v. Machine Type: Select “N2-standard-8 (8 vCPU, 32 GB Memory)” → This is recommended for lab setup; however you can select any other combination as per your requirement.
   vi. Then click the 'Change' button for the Boot Disk selection. You need to select your custom image "cmpe-283-image".
   vii. Then select the Firewall settings from the main screen and Select both HTTP and HTTPS traffic.
   viii. Click on Create button to create the Linux instance.



Machine configuration

Machine type

You must stop the VM instance to edit its machine type

n2-standard-8

vCPU	Memory
8	32 GB

CPU platform

Intel Cascade Lake

Display device

You must stop the VM instance to enable the display device.
Enable to use screen capturing and recording tools

☐ Enable display device

Storage

Boot disk

Name	instance-1
Image	cmpe-283-image
Size	200 GB
Interface type	SCSI
Type	SSD persistent disk
Encryption type	Google-managed
Mode	Boot, read/write
Snapshot schedule	None

Deletion rule

When deleting instance

○ Keep disk

◉ Delete disk

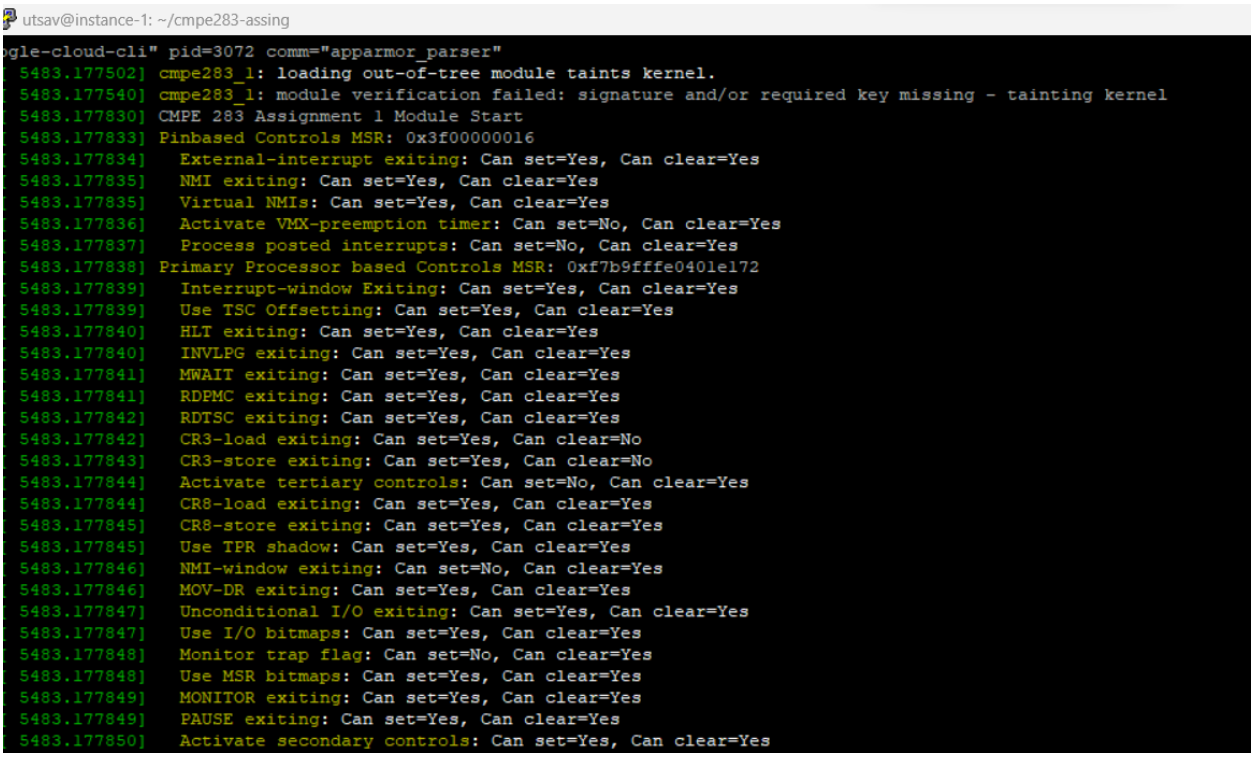
8. Now you can access the VM directly from the browser through SSH.
9. Create a new directory named "cmpe283-assing".
   >> mkdir cmpe283-assing
10. Copy the template "cmpe283-1.c" file and the template "Makefile" provided by the professor to the cmpe283-assing directory.
11. Modify “cmpe283-1.c" file, add all others 5 MSRs as explained in the assignment description:
    i. By referring SDM, created structures with name (description) and bit positions for Primary Processor based, Secondary Processor based, Tertiary Processor, Entry and Exit controls.
    ii. In order to detect and print VMX capabilities of CPU, the function report_capability ( ) is called with appropriate parameters passed in order to print Primary Processor based, Secondary Processor based, Tertiary Processor, Entry and Exit controls.
12. Run below command:
    >> apt install gcc make
    >> sudo apt-get linux-headers-$(uname -r)
13. Build the module using the following command inside the cmpe283-assing directory.
    >> make
14. Module is inserted into the kernel:
    >> sudo insmod ./cmpe283-1.ko
15. Verify that the VMX capabilities for all the MSRs are displayed in the logs command:
    >> sudo dmesg
16. when the module is removed with command:
    >> sudo rmmod cmpe283-1

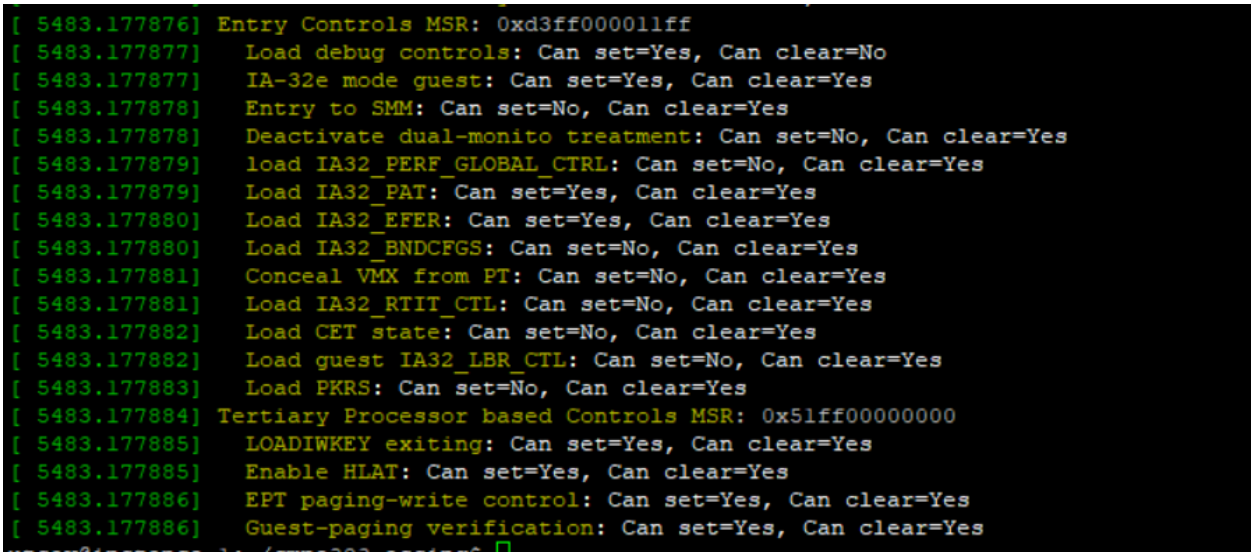
Output of dmesg Logs:

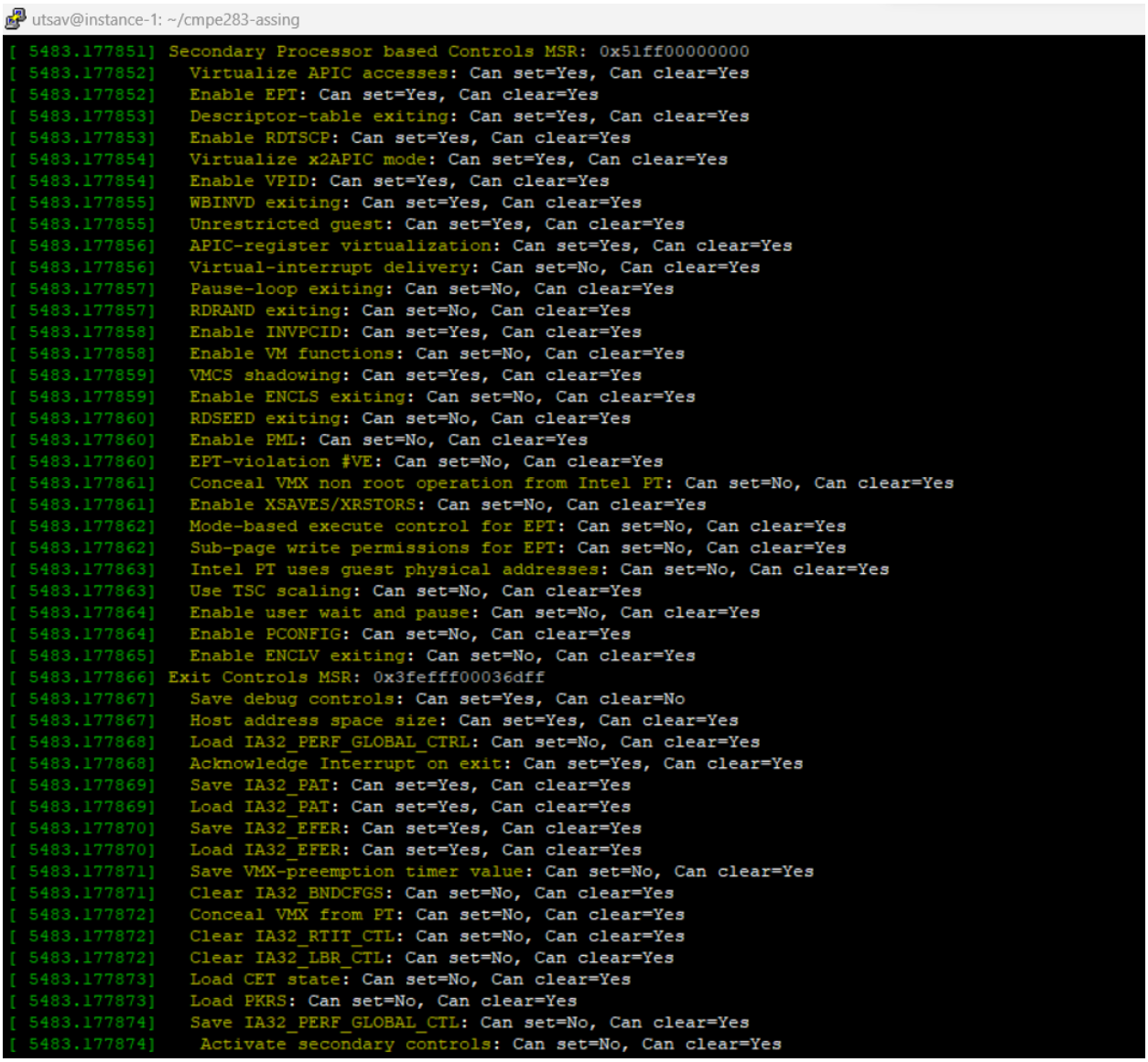
[4048.285552] CMPE 283 Assignment 1 Module Start
[4048.285555] Pinbased Controls MSR: 0x3f00000016
[4048.285556] External-interrupt exiting: Can set=Yes, Can clear=Yes
[4048.285556] NMI exiting: Can set=Yes, Can clear=Yes
[4048.285557] Virtual NMIs: Can set=Yes, Can clear=Yes
[4048.285557] Activate VMX-preemption timer: Can set=No, Can clear=Yes
[4048.285558] Process posted interrupts: Can set=No, Can clear=Yes
[4048.285559] Primary Processor based Controls MSR: 0xf7b9fffe0401e172
[4048.285560] Interrupt-window Exiting: Can set=Yes, Can clear=Yes
[4048.285560] Use TSC Offsetting: Can set=Yes, Can clear=Yes
[4048.285561] HLT exiting: Can set=Yes, Can clear=Yes
[4048.285562] INVLPG exiting: Can set=Yes, Can clear=Yes
[4048.285562] MWAIT exiting: Can set=Yes, Can clear=Yes
[4048.285563] RDPMC exiting: Can set=Yes, Can clear=Yes
[4048.285563] RDTSC exiting: Can set=Yes, Can clear=Yes
[4048.285564] CR3-load exiting: Can set=Yes, Can clear=No

[4048.285564] CR3-store exiting: Can set=Yes, Can clear=No
[4048.285565] Activate tertiary controls: Can set=No, Can clear=Yes
[4048.285565] CR8-load exiting: Can set=Yes, Can clear=Yes
[4048.285566] CR8-store exiting: Can set=Yes, Can clear=Yes
[4048.285566] Use TPR shadow: Can set=Yes, Can clear=Yes
[4048.285567] NMI-window exiting: Can set=No, Can clear=Yes
[4048.285567] MOV-DR exiting: Can set=Yes, Can clear=Yes
[4048.285568] Unconditional I/O exiting: Can set=Yes, Can clear=Yes
[4048.285568] Use I/O bitmaps: Can set=Yes, Can clear=Yes
[4048.285569] Monitor trap flag: Can set=No, Can clear=Yes
[4048.285569] Use MSR bitmaps: Can set=Yes, Can clear=Yes
[4048.285569] MONITOR exiting: Can set=Yes, Can clear=Yes
[4048.285570] PAUSE exiting: Can set=Yes, Can clear=Yes
[4048.285570] Activate secondary controls: Can set=Yes, Can clear=Yes
[4048.285572] Secondary Processor based Controls MSR: 0x51ff00000000
[4048.285573] Virtualize APIC accesses: Can set=Yes, Can clear=Yes
[4048.285573] Enable EPT: Can set=Yes, Can clear=Yes
[4048.285574] Descriptor-table exiting: Can set=Yes, Can clear=Yes
[4048.285574] Enable RDTSCP: Can set=Yes, Can clear=Yes
[4048.285575] Virtualize x2APIC mode: Can set=Yes, Can clear=Yes
[4048.285575] Enable VPID: Can set=Yes, Can clear=Yes
[4048.285576] WBINVD exiting: Can set=Yes, Can clear=Yes
[4048.285576] Unrestricted guest: Can set=Yes, Can clear=Yes
[4048.285577] APIC-register virtualization: Can set=Yes, Can clear=Yes
[4048.285577] Virtual-interrupt delivery: Can set=No, Can clear=Yes
[4048.285578] Pause-loop exiting: Can set=No, Can clear=Yes
[4048.285578] RDRAND exiting: Can set=No, Can clear=Yes
[4048.285578] Enable INVPCID: Can set=Yes, Can clear=Yes
[4048.285579] Enable VM functions: Can set=No, Can clear=Yes
[4048.285579] VMCS shadowing: Can set=Yes, Can clear=Yes
[4048.285580] Enable ENCLS exiting: Can set=No, Can clear=Yes
[4048.285580] RDSEED exiting: Can set=No, Can clear=Yes
[4048.285581] Enable PML: Can set=No, Can clear=Yes
[4048.285581] EPT-violation #VE: Can set=No, Can clear=Yes
[4048.285582] Conceal VMX non root operation from Intel PT: Can set=No, Can clear=Yes
[4048.285582] Enable XSAVES/XRSTORS: Can set=No, Can clear=Yes
[4048.285583] Mode-based execute control for EPT: Can set=No, Can clear=Yes
[4048.285583] Sub-page write permissions for EPT: Can set=No, Can clear=Yes
[4048.285584] Intel PT uses guest physical addresses: Can set=No, Can clear=Yes
[4048.285584] Use TSC scaling: Can set=No, Can clear=Yes
[4048.285585] Enable user wait and pause: Can set=No, Can clear=Yes
[4048.285585] Enable PCONFIG: Can set=No, Can clear=Yes
[4048.285586] Enable ENCLV exiting: Can set=No, Can clear=Yes
[4048.285587] Exit Controls MSR: 0x3fefff00036dff

[4048.285588] Save debug controls: Can set=Yes, Can clear=No
[4048.285588] Host address space size: Can set=Yes, Can clear=Yes
[4048.285589] Load IA32_PERF_GLOBAL_CTRL: Can set=No, Can clear=Yes
[4048.285589] Acknowledge Interrupt on exit: Can set=Yes, Can clear=Yes
[4048.285590] Save IA32_PAT: Can set=Yes, Can clear=Yes
[4048.285590] Load IA32_PAT: Can set=Yes, Can clear=Yes
[4048.285590] Save IA32_EFER: Can set=Yes, Can clear=Yes
[4048.285591] Load IA32_EFER: Can set=Yes, Can clear=Yes
[4048.285591] Save VMX-preemption timer value: Can set=No, Can clear=Yes
[4048.285592] Clear IA32_BNDCFGS: Can set=No, Can clear=Yes
[4048.285592] Conceal VMX from PT: Can set=No, Can clear=Yes
[4048.285593] Clear IA32_RTIT_CTL: Can set=No, Can clear=Yes
[4048.285593] Clear IA32_LBR_CTL: Can set=No, Can clear=Yes
[4048.285594] Load CET state: Can set=No, Can clear=Yes
[4048.285594] Load PKRS: Can set=No, Can clear=Yes
[4048.285595] Save IA32_PERF_GLOBAL_CTL: Can set=No, Can clear=Yes
[4048.285595]   Activate secondary controls: Can set=No, Can clear=Yes
[4048.285597] Entry Controls MSR: 0xd3ff000011ff
[4048.285597] Load debug controls: Can set=Yes, Can clear=No
[4048.285598] IA-32e mode guest: Can set=Yes, Can clear=Yes
[4048.285598] Entry to SMM: Can set=No, Can clear=Yes
[4048.285599] Deactivate dual-monito treatment: Can set=No, Can clear=Yes
[4048.285600] load IA32_PERF_GLOBAL_CTRL: Can set=No, Can clear=Yes
[4048.285600] Load IA32_PAT: Can set=Yes, Can clear=Yes
[4048.285600] Load IA32_EFER: Can set=Yes, Can clear=Yes
[4048.285601] Load IA32_BNDCFGS: Can set=No, Can clear=Yes
[4048.285601] Conceal VMX from PT: Can set=No, Can clear=Yes
[4048.285602] Load IA32_RTIT_CTL: Can set=No, Can clear=Yes
[4048.285602] Load CET state: Can set=No, Can clear=Yes
[4048.285603] Load guest IA32_LBR_CTL: Can set=No, Can clear=Yes
[4048.285603] Load PKRS: Can set=No, Can clear=Yes
[4048.285605] Tertiary Processor based Controls MSR: 0x51ff00000000
[4048.285605] LOADIWKEY exiting: Can set=Yes, Can clear=Yes
[4048.285606] Enable HLAT: Can set=Yes, Can clear=Yes
[4048.285606] EPT paging-write control: Can set=Yes, Can clear=Yes
[4048.285607] Guest-paging verification: Can set=Yes, Can clear=Yes


utsav@instance-1: ~/cmpe283-assing
gle-cloud-cli" pid=3072 comm="apparmor_parser"
5483.177502] cmpe283_1: loading out-of-tree module taints kernel.
5483.177540] cmpe283_1: module verification failed: signature and/or required key missing - tainting kernel
5483.177830] CMPE 283 Assignment 1 Module Start
5483.177833] Pinbased Controls MSR: 0x3f00000016
5483.177834] External-interrupt exiting: Can set=Yes, Can clear=Yes
5483.177835] NMI exiting: Can set=Yes, Can clear=Yes
5483.177835] Virtual NMIs: Can set=Yes, Can clear=Yes
5483.177836] Activate VMX-preemption timer: Can set=No, Can clear=Yes
5483.177837] Process posted interrupts: Can set=No, Can clear=Yes
5483.177838] Primary Processor based Controls MSR: 0xf7b9fffe0401e172
5483.177839] Interrupt-window Exiting: Can set=Yes, Can clear=Yes
5483.177839] Use TSC Offsetting: Can set=Yes, Can clear=Yes
5483.177840] HLT exiting: Can set=Yes, Can clear=Yes
5483.177840] INVLPG exiting: Can set=Yes, Can clear=Yes
5483.177841] MWAIT exiting: Can set=Yes, Can clear=Yes
5483.177841] RDPMC exiting: Can set=Yes, Can clear=Yes
5483.177842] RDTSC exiting: Can set=Yes, Can clear=Yes
5483.177842] CR3-load exiting: Can set=Yes, Can clear=No
5483.177843] CR3-store exiting: Can set=Yes, Can clear=No
5483.177844] Activate tertiary controls: Can set=No, Can clear=Yes
5483.177844] CR8-load exiting: Can set=Yes, Can clear=Yes
5483.177845] CR8-store exiting: Can set=Yes, Can clear=Yes
5483.177845] Use TPR shadow: Can set=Yes, Can clear=Yes
5483.177846] NMI-window exiting: Can set=No, Can clear=Yes
5483.177846] MOV-DR exiting: Can set=Yes, Can clear=Yes
5483.177847] Unconditional I/O exiting: Can set=Yes, Can clear=Yes
5483.177847] Use I/O bitmaps: Can set=Yes, Can clear=Yes
5483.177848] Monitor trap flag: Can set=No, Can clear=Yes
5483.177848] Use MSR bitmaps: Can set=Yes, Can clear=Yes
5483.177849] MONITOR exiting: Can set=Yes, Can clear=Yes
5483.177849] PAUSE exiting: Can set=Yes, Can clear=Yes
5483.177850] Activate secondary controls: Can set=Yes, Can clear=Yes


[5483.177876] Entry Controls MSR: 0xd3ff000011ff
[5483.177877] Load debug controls: Can set=Yes, Can clear=No
[5483.177877] IA-32e mode guest: Can set=Yes, Can clear=Yes
[5483.177878] Entry to SMM: Can set=No, Can clear=Yes
[5483.177878] Deactivate dual-monito treatment: Can set=No, Can clear=Yes
[5483.177879] load IA32_PERF_GLOBAL_CTRL: Can set=No, Can clear=Yes
[5483.177879] Load IA32_PAT: Can set=Yes, Can clear=Yes
[5483.177880] Load IA32_EFER: Can set=Yes, Can clear=Yes
[5483.177880] Load IA32_BNDCFGS: Can set=No, Can clear=Yes
[5483.177881] Conceal VMX from PT: Can set=No, Can clear=Yes
[5483.177881] Load IA32_RTIT_CTL: Can set=No, Can clear=Yes
[5483.177882] Load CET state: Can set=No, Can clear=Yes
[5483.177882] Load guest IA32_LBR_CTL: Can set=No, Can clear=Yes
[5483.177883] Load PKRS: Can set=No, Can clear=Yes
[5483.177884] Tertiary Processor based Controls MSR: 0x51ff00000000
[5483.177885] LOADIWKEY exiting: Can set=Yes, Can clear=Yes
[5483.177885] Enable HLAT: Can set=Yes, Can clear=Yes
[5483.177886] EPT paging-write control: Can set=Yes, Can clear=Yes
[5483.177886] Guest-paging verification: Can set=Yes, Can clear=Yes


utsav@instance-1: ~/cmpe283-assing
[5483.177851] Secondary Processor based Controls MSR: 0x51ff00000000
[5483.177852]   Virtualize APIC accesses: Can set=Yes, Can clear=Yes
[5483.177852]   Enable EPT: Can set=Yes, Can clear=Yes
[5483.177853]   Descriptor-table exiting: Can set=Yes, Can clear=Yes
[5483.177853]   Enable RDTSCP: Can set=Yes, Can clear=Yes
[5483.177854]   Virtualize x2APIC mode: Can set=Yes, Can clear=Yes
[5483.177854]   Enable VPID: Can set=Yes, Can clear=Yes
[5483.177855]   WBINVD exiting: Can set=Yes, Can clear=Yes
[5483.177855]   Unrestricted guest: Can set=Yes, Can clear=Yes
[5483.177856]   APIC-register virtualization: Can set=Yes, Can clear=Yes
[5483.177856]   Virtual-interrupt delivery: Can set=No, Can clear=Yes
[5483.177857]   Pause-loop exiting: Can set=No, Can clear=Yes
[5483.177857]   RDRAND exiting: Can set=No, Can clear=Yes
[5483.177858]   Enable INVPCID: Can set=Yes, Can clear=Yes
[5483.177858]   Enable VM functions: Can set=No, Can clear=Yes
[5483.177859]   VMCS shadowing: Can set=Yes, Can clear=Yes
[5483.177859]   Enable ENCLS exiting: Can set=No, Can clear=Yes
[5483.177860]   RDSEED exiting: Can set=No, Can clear=Yes
[5483.177860]   Enable PML: Can set=No, Can clear=Yes
[5483.177860]   EPT-violation #VE: Can set=No, Can clear=Yes
[5483.177861]   Conceal VMX non root operation from Intel PT: Can set=No, Can clear=Yes
[5483.177861]   Enable XSAVES/XRSTORS: Can set=No, Can clear=Yes
[5483.177862]   Mode-based execute control for EPT: Can set=No, Can clear=Yes
[5483.177862]   Sub-page write permissions for EPT: Can set=No, Can clear=Yes
[5483.177863]   Intel PT uses guest physical addresses: Can set=No, Can clear=Yes
[5483.177863]   Use TSC scaling: Can set=No, Can clear=Yes
[5483.177864]   Enable user wait and pause: Can set=No, Can clear=Yes
[5483.177864]   Enable PCONFIG: Can set=No, Can clear=Yes
[5483.177865]   Enable ENCLV exiting: Can set=No, Can clear=Yes
[5483.177866] Exit Controls MSR: 0x3fefff00036dff
[5483.177867]   Save debug controls: Can set=Yes, Can clear=No
[5483.177867]   Host address space size: Can set=Yes, Can clear=Yes
[5483.177868]   Load IA32_PERF_GLOBAL_CTRL: Can set=No, Can clear=Yes
[5483.177868]   Acknowledge Interrupt on exit: Can set=Yes, Can clear=Yes
[5483.177869]   Save IA32_PAT: Can set=Yes, Can clear=Yes
[5483.177869]   Load IA32_PAT: Can set=Yes, Can clear=Yes
[5483.177870]   Save IA32_EFER: Can set=Yes, Can clear=Yes
[5483.177870]   Load IA32_EFER: Can set=Yes, Can clear=Yes
[5483.177871]   Save VMX-preemption timer value: Can set=No, Can clear=Yes
[5483.177871]   Clear IA32_BNDCFGS: Can set=No, Can clear=Yes
[5483.177872]   Conceal VMX from PT: Can set=No, Can clear=Yes
[5483.177872]   Clear IA32_RTIT_CTL: Can set=No, Can clear=Yes
[5483.177872]   Clear IA32_LBR_CTL: Can set=No, Can clear=Yes
[5483.177873]   Load CET state: Can set=No, Can clear=Yes
[5483.177873]   Load PKRS: Can set=No, Can clear=Yes
[5483.177874]   Save IA32_PERF_GLOBAL_CTL: Can set=No, Can clear=Yes
[5483.177874]   Activate secondary controls: Can set=No, Can clear=Yes